

THE EVOLVING CHROMATIN AND TRANSCRIPTIONAL LANDSCAPES – EMERGING METHODS, TOOLS AND TECHNIQUES

EDITED BY: Jean-Philippe Lambert, Steve Bilodeau and Christoph Franz Kurat
PUBLISHED IN: *Frontiers in Cell and Developmental Biology* and
Frontiers in Genetics



frontiers

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence.

The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714

ISBN 978-2-88971-921-1

DOI 10.3389/978-2-88971-921-1

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

THE EVOLVING CHROMATIN AND TRANSCRIPTIONAL LANDSCAPES – EMERGING METHODS, TOOLS AND TECHNIQUES

Topic Editors:

Jean-Philippe Lambert, Laval University, Canada

Steve Bilodeau, Laval University, Canada

Christoph Franz Kurat, Ludwig Maximilian University of Munich, Germany

Citation: Lambert, J.-P., Bilodeau, S., Kurat, C. F., eds. (2021). The Evolving Chromatin and Transcriptional Landscapes – Emerging Methods, Tools and Techniques. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88971-921-1

Table of Contents

- 04 Editorial: The Evolving Chromatin and Transcriptional Landscapes—Emerging Methods, Tools and Techniques**
Steve Bilodeau, Christoph Franz Kurat and Jean-Philippe Lambert
- 06 Histone H3 Lysine 56 Acetylation Is Required for Formation of Normal Levels of Meiotic DNA Breaks in *S. cerevisiae***
Zsolt Karányi, Lilla Hornyák and Lóránt Székvölgyi
- 15 A Degenerate Peptide Library Approach to Reveal Sequence Determinants of Methyllysine-Driven Protein Interactions**
Ariana Kupai, Robert M. Vaughan, Bradley M. Dickson and Scott B. Rothbart
- 28 H3K4me1 Distribution Predicts Transcription State and Poising at Promoters**
Sunhee Bae and Bluma J. Lesch
- 39 The Use of Mononucleosome Immunoprecipitation for Analysis of Combinatorial Histone Post-translational Modifications and Purification of Nucleosome-Interacting Proteins**
Kashif Aziz Khan, Marlee K. Ng and Peter Cheung
- 55 Proximity Labeling Techniques to Study Chromatin**
Henning Ummethum and Stephan Hamperl
- 68 Exploring the Histone Acetylation Cycle in the Protozoan Model *Tetrahymena thermophila***
Suzanne Wahab, Alejandro Saettone, Syed Nabeel-Shah, Nora Dannah and Jeffrey Fillingham
- 82 A New View of Genome Organization Through RNA Directed Interactions**
Gabriel Khelifi and Samer M. I. Hussein
- 88 Interactions With Histone H3 & Tools to Study Them**
William A. Scott and Eric I. Campos
- 109 Structural Paradigms in the Recognition of the Nucleosome Core Particle by Histone Lysine Methyltransferases**
Ashley Janna, Hossein Davarinejad, Monika Joshi and Jean-Francois Couture
- 121 Methods to Study Intracellular Movement and Localization of the Nucleotide Excision Repair Proteins at the DNA Lesions in Mammalian Cells**
Mihaela Robu, Rashmi G. Shah and Girish M. Shah
- 140 Approaches to Study Native Chromatin-Modifying Complex Activities and Functions**
Maxime Galloy, Catherine Lachance, Xue Cheng, Félix Distéfano-Gagné, Jacques Côté and Amelie Fradet-Turcotte



Editorial: The Evolving Chromatin and Transcriptional Landscapes—Emerging Methods, Tools and Techniques

Steve Bilodeau^{1,2,3,4*†}, Christoph Franz Kurat^{5*†} and Jean-Philippe Lambert^{2,3,6,7*†}

¹ Centre de Recherche du CHU de Québec – Université Laval, Axe Oncologie, Québec, QC, Canada, ² Centre de Recherche sur le Cancer de L'Université Laval, Québec, QC, Canada, ³ Centre de Recherche en Données Massives de L'Université Laval, Québec, QC, Canada, ⁴ Département de Biologie Moléculaire, Biochimie Médicale et Pathologie, Faculté de Médecine, Université Laval, Québec, QC, Canada, ⁵ Molecular Biology Division, Biomedical Center Munich, Ludwig-Maximilians-Universität, Munich, Germany, ⁶ Centre de Recherche du CHU de Québec – Université Laval, Axe Endocrinologie et Néphrologie, Québec, QC, Canada, ⁷ Département de Médecine Moléculaire, Faculté de Médecine, Université Laval, Québec, QC, Canada

OPEN ACCESS

Edited and reviewed by:

Michael E. Symonds,
University of Nottingham,
United Kingdom

*Correspondence:

Steve Bilodeau
steve.bilodeau@
crchudequebec.ulaval.ca
Christoph Franz Kurat
christoph.kurat@bmc.med.lmu.de
Jean-Philippe Lambert
jean-philippe.lambert@
crchudequebec.ulaval.ca

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 24 September 2021

Accepted: 07 October 2021

Published: 01 November 2021

Citation:

Bilodeau S, Kurat CF and Lambert J-P
(2021) Editorial: The Evolving
Chromatin and Transcriptional
Landscapes—Emerging Methods,
Tools and Techniques.
Front. Cell Dev. Biol. 9:782776.
doi: 10.3389/fcell.2021.782776

Keywords: chromatin, chromatin immunoprecipitation (ChIP), transcription, epigenetics, histones

Editorial on the Research Topic

The Evolving Chromatin and Transcriptional Landscapes—Emerging Methods, Tools and Techniques

Mechanisms controlling the packaging of the genetic material into chromatin are central for normal and disease development. At the core of the chromatin structure, the DNA is wrapped around histone proteins to create nucleosomes which are constantly modified and acted upon to allow for effective regulation of transcription, DNA repair, replication and maintenance of the cellular state. Accordingly, in recent years, multiple chromatin modifiers and remodelers have emerged as causal factors and promising drug targets for numerous pathologies (Hogg et al., 2020; Bhat et al., 2021). As such, an in-depth understanding of the mechanisms required for effective regulation of chromatin states during normal and disease development is essential.

The advent of effective sequencing technologies has enabled rapid progress in our understanding of chromatin biology. For example, the original article by Bae and Lesch made use of the chromatin immunoprecipitation coupled to sequencing (ChIP-seq) technique to highlight bimodal patterns of H3K4me1 at active promoters flanked by H3K4me3. Interestingly, a unimodal pattern was found to coincide with H3K4me3 and H3K27me3 at poised promoters. Furthermore, emerging sequencing techniques were the basis of the thought-provoking opinion article of Khelifi and Hussein on the roles of RNA directed interactions on genome organization. The authors postulate that two distinct functional groups of long non-coding RNA (lncRNA) respectively operate locally on the structure of chromatin itself and promote long-range chromatin interactions and bridging events.

This Frontiers Research Topic reports significant progresses toward the systematic deployment of complementary approaches to sequencing techniques. One example is the development of degenerated methylated lysine-oriented peptide libraries (Kme-OPL), which enables the specificity of Kme reader modules to be defined. In a research article, Kupai et al. describe the development of Kme-OPL and its use for the characterization of Kme reader modules to reveal the specificity or promiscuity of Kme reader modules. Similarly, Janna et al. details the biochemical and structural studies of the crosstalk between PTMs which enable a molecular understanding of the positive impact of histone H2B ubiquitylation on the methylation of H3K79 and H3K4. This is furthered

by Scott and Campos who discuss the numerous tools to characterize histone H3 and its partners. One such tool, proximity dependent biotinylation is highlighted by Ummethum and Hamperl.

In this Frontiers Research Topic, three detailed protocols by Aziz Khan et al.; Galloy et al.; and Robu et al. promote the effective characterization of chromatin and its effectors. In their step-by-step protocol, Aziz Khan et al. describe how to isolate large amounts of nucleosomes from mammalian cells for downstream characterization. Galloy et al. focuses on chromatin remodelers and provided two distinct protocols to permit large-scale purification of chromatin remodeling complexes and the use of an anchor-away system in human cells. Lastly, Robu et al. reports step-by-step protocols to study proteins involved in nucleotide excision repair (NER) localization at DNA lesions. These methods all contribute to the characterization of the dynamic nature of the interplays shaping the chromatin environment.

Further, the need for effective model systems to study chromatin was also highlighted in this Frontier Research Topic. In a brief research report, Karányi et al. revisited the roles of H3K56ac during meiotic recombination. Working in the atypical SK1 *Saccharomyces cerevisiae* strain, a strain well-adapted to synchronous sporulation (Borner and Cha, 2015), the authors employed classical yeast genetics in combination with ChIP-seq to reveal the requirement for H3K56ac to produce normal levels of double strand breaks in recombination hotspot regions. In a review article, Wahab et al. highlight the *Tetrahymena thermophila* model, its unique biology and its use to study Kac-dependent processes. Historically *Tetrahymena* has enabled the identification of the first lysine acetyltransferase (Brownell et al., 1996). The authors propose that it is perfectly suited to uncover novel mechanisms impacting chromatin structures and functions when coupled to modern techniques.

REFERENCES

- Bhat, K. P., Umit Kaniskan, H., Jin, J., and Gozani, O. (2021). Epigenetics and beyond: targeting writers of protein lysine methylation to treat disease. *Nat. Rev. Drug Discov.* 20, 265–286. doi: 10.1038/s41573-020-00108-x
- Borner, G. V., and Cha, R. S. (2015). Induction and analysis of synchronous meiotic yeast cultures. *Cold Spring Harb. Protoc.* 2015, 908–913. doi: 10.1101/pdb.prot085035
- Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y., et al. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84, 843–851. doi: 10.1016/s0092-8674(00)81063-6
- Hogg, S. J., Beavis, P. A., Dawson, M. A., and Johnstone, R. W. (2020). Targeting the epigenetic regulation of antitumour immunity. *Nat. Rev. Drug Discov.* 19, 776–800. doi: 10.1038/s41573-020-0077-5

While sequencing-based methods remain the dominant approach to study chromatin biology, exciting new tools and techniques are emerging to complement them. Together, these approaches will allow for a more detailed understanding of chromatin biology and transcriptional regulation. We believe that this Frontiers Research Topic will support this endeavor.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

Research in the Bilodeau Laboratory was funded by funds from the Canada Research Chair in Transcriptional Genomics [Grant #950-228321]; a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2019-06490), a Project Grant from the Canadian Institutes of Health Research (PJT-451568) and an operating grant from the Cancer Research Society and Genome Quebec (840369). Research in the Kurat Laboratory was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation; Project-ID 213249687–SFB 1064). Research in the Lambert Laboratory was funded by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2017-06124), a Project Grant from the Canadian Institutes of Health Research (PJT-168969) and an operating Grant from the Cancer Research Society and Genome Quebec (25123).

ACKNOWLEDGMENTS

We thank all contributing authors and reviewers for their support to the Research Topic.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Bilodeau, Kurat and Lambert. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Histone H3 Lysine 56 Acetylation Is Required for Formation of Normal Levels of Meiotic DNA Breaks in *S. cerevisiae*

Zsolt Karányi^{1,2}, Lilla Hornyák¹ and Lóránt Székvölgyi^{1*}†

¹ MTA-DE Momentum Genome Architecture and Recombination Research Group, Department of Biochemistry and Molecular Biology, Faculty of Medicine, University of Debrecen, Debrecen, Hungary, ² Department of Internal Medicine, Faculty of Medicine, University of Debrecen, Debrecen, Hungary

OPEN ACCESS

Edited by:

Jean-Philippe Lambert,
Laval University, Canada

Reviewed by:

Jeffrey Fillingham,
Ryerson University, Canada
Craig Peterson,
University of Massachusetts Medical
School, United States

*Correspondence:

Lóránt Székvölgyi
lorantsz@med.unideb.hu

†ORCID:

Lóránt Székvölgyi
orcid.org/0000-0002-7529-0319

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 25 October 2019

Accepted: 12 December 2019

Published: 10 January 2020

Citation:

Karányi Z, Hornyák L and
Székvölgyi L (2020) Histone H3 Lysine
56 Acetylation Is Required
for Formation of Normal Levels
of Meiotic DNA Breaks
in *S. cerevisiae*.
Front. Cell Dev. Biol. 7:364.
doi: 10.3389/fcell.2019.00364

Meiotic recombination is initiated by Spo11-catalyzed DNA double-strand breaks (DSBs) that are promoted by histone modifications and histone modifying enzymes. Herein we investigated the role of histone H3 lysine 56 acetylation (H3K56ac) located near the entry/exit points of the DNA in the globular H3 domain. We generated a series of mutant cells (*asf1* Δ , *rtt109* Δ , *hst3/4* Δ , and H3K56A) in which the endogenous level of H3K56ac was manipulated and tracked during meiotic growth. We show that complete loss or increased abundance of H3K56ac in these mutants allows timely entry into meiosis and sporulation and does not impair S phase progression, first and second meiotic cell divisions, and spore viability. In the *asf1* Δ , *rtt109* Δ , *hst3/4* Δ mutants, DSBs and crossovers form normal levels with a short (60-min) delay at the *HIS4-LEU2* artificial recombination hotspot, however, DSB formation shows a \sim threefold decrease in the H3K56A mutant at the natural *BUD23-ARE1* hotspot. The latter DSB phenotype, showing significant DSB reduction in the H3K56A mutant, was also observed at DSB sites using genome-wide mapping of Rfa1-coated single-stranded DNA flanking DSBs (RPA ChIP). Parallel mapping of H3K56-acetylated histones in wild type cells revealed strong depletion of the H3K56ac ChIP signal over Spo11-oligo DSBs, albeit most H3K56-acetylated histones were enriched adjacent to the identified RPA ChIP binding sites. Taken together, these associations demonstrate a prominent role of H3 lysine 56 acetylation in the formation of DNA breaks within recombination hotspot regions.

Keywords: recombination, DNA break, meiosis, histone modification, H3K56 acetylation

INTRODUCTION

Meiosis is a cellular differentiation process which is accompanied by high levels of recombination between the homologous chromosomes, initiated by DNA double-strand breaks (DSBs) catalyzed by Spo11 and accessory factors (Keeney et al., 1997; Székvölgyi and Nicolas, 2010; Székvölgyi et al., 2015). In *S. cerevisiae*, meiotic DSBs are controlled by the elements of chromatin structure. This involves a complex interplay between DNA sequence composition, local chromatin status, nucleosome occupancy, and transcription factor binding (Pan et al., 2011). Among these factors, histone modifications represent an important layer which has only recently been explored in detail. The most well-characterized histone modification is H3K4me3 that appears to be essential for recombination sites in most organisms. In the current model, H3K4me3 is deposited by Set1C

and becomes recognized by Spp1 (the PHD finger subunit of Set1C), which leads to tethering of DSB sites to the chromosome axis that undergo Spo11-mediated cleavage (Acquaviva et al., 2013; Sommermeyer et al., 2013; Karányi et al., 2018). A similar mechanism has been proposed in mammals involving the meiosis-specific H3K4 methylase Prdm9 (Baudat et al., 2010; Parpanov et al., 2010), however, CXXC1 (the yeast ortholog of Spp1) is apparently not essential for the association of H3K4 tri-methylated recombination sites with the DSB machinery (Tian et al., 2018).

The widely localized H3K4me3 mark has the virtue of initiating recombination at numerous places in the genome, however, other histone modifications are also needed to keep recombination flexible for the diversity of recombinant haplotypes (Szekvolgyi and Nicolas, 2010). These “alternative” pathways remain to be clarified to better understand the plasticity of crossover patterning. Most chemical modifications are concentrated at the N-termini of histones and are not expected to alter the structure of nucleosomes (White et al., 2001; Biswas et al., 2011). However, modifications of histone core domains can directly change nucleosome structure, which is well-established biochemically (Luger et al., 1997; Biswas et al., 2011) but its functional relevance is less understood.

In the current study, we focused on the role of histone H3 lysine 56 acetylation located near the entry/exit points of the DNA in the globular H3 domain, predicted to destabilize the histone/DNA contact (Buning and Van Noort, 2010; Simon et al., 2011). H3K56ac is a transient chromatin signal showing rapid turnover and is closely linked to DNA replication and histone eviction during transcription (Rufange et al., 2007; Watanabe et al., 2013). Functional studies in mitotically proliferating yeast cells revealed that H3K56ac enables the assembly and disassembly of nucleosomes during DNA synthesis and upon transcriptional activation. Furthermore, the histone chaperone Asf1 [carrying the H3K56-specific acetyltransferase Rtt109 (Abshiru et al., 2013)] and the histone residue H3K56 were found to be necessary for meiotic S phase progression and for the development of reproductive capacity in yeast and mouse models (Recht et al., 2006; Govin et al., 2010; Messiaen et al., 2016). However, the mechanism of action of H3K56 acetylation has not been fully elucidated and remains to be clarified. Therefore, we applied a functional approach in meiotic *S. cerevisiae* cells to modify the natural levels of H3K56ac to study the biochemical phases of meiosis. Our results demonstrate that H3K56ac is necessary for formation of normal levels of DSBs within recombination hotspot regions.

METHODS

All methods are available as **Supplementary Material**.

RESULTS

We generated mutant cells in which the endogenous level of H3K56ac was modified and tracked during meiotic growth

(**Figure 1A**). In wild-type cells, the level of H3K56ac reached a peak during DSB formation (at ~4–5 h in SPM) and then rapidly dropped, similar to earlier western blot results (Recht et al., 2006). The H3K56ac signal disappeared almost completely from *asf1Δ* and *rtt109Δ* deletion mutants during the meiotic time course (~fourfold reduction compared to wild type), in agreement with the crucial role of Asf1 and Rtt109 in the deposition of this epigenetic tag (Tsubota et al., 2007). The K56ac signal increased and remained high in the *hst3/4Δ* double-mutant that prevents deacetylation of H3K56 (Celic et al., 2006). In mitotic growth conditions, we could not detect an H3K56 acetylation signal, except for the *hst3/4Δ* mutant (**Figure 1A**), in which H3K56ac persists throughout the cell cycle (Celic et al., 2006). This is in accordance with the transient nature of H3K56 acetylation in asynchronously proliferating mitotic cells, mainly associated with newly replicated chromatin regions (Simoneau et al., 2015).

The correlation between H3K56ac dynamics and the phase of meiotic DSBs prompted us to analyze the progress of S phase, DSB formation, sporulation efficiency, and spore viability in cells with various H3K56ac levels. FACS analysis revealed a 30–60 min delay in G1/S phase progression in mutant cells, however, all mutants reached the G2/M phase within 6 h in SPM (**Figures 1B,C**). We then performed a meiotic time course up to 20 h. The results reported in **Figures 1D,E** show that sporulation efficiencies and spore viabilities do not differ between mutant and wild-type cells (less than 1.5-fold change was observed in the number of tetrads and viable spores). We conclude that absence or elevated levels of H3K56ac is dispensable for meiotic S phase progression, MI/MII cell divisions, and spore viability.

We next analyzed whether DSBs were affected in mutants (with *Sae2⁺/Rad50⁺* background). DSB levels were monitored at the *HIS4-LEU2* hotspot (Xu and Kleckner, 1995) using Southern blot hybridization (**Figure 2A**). All mutants produced comparable DSB levels relative to wild-type (between 6 and 8% based on the global maxima of the curves), with a ~60 min delay in *asf1Δ* and *hst3/4Δ* mutants (**Figure 2C**). The delay of DNA breaks in *asf1Δ* and *hst3/4Δ* cells suggests that DSBs appear and disappear with a slightly modified kinetics in these mutants. Maximal DSB reduction was observed in the *rtt109Δ* strain (1.5-fold reduction). Importantly, meiotic DSBs detected in the mutants were highly recombinogenic (R1/R2 recombinants are highlighted in **Figure 2B**) forming wild-type levels of crossovers (~10–11%, **Figure 2C**). These results suggest that DSBs detected at *HIS4-LEU2* were properly processed to form mature recombination products.

To obtain mechanistic insights whether the modifiable H3K56 residue influences DSB formation, we constructed a plasmid shuffle system that allows the expression of wild-type H3K56 (H3 ctrl) and unmodifiable H3K56A as the only source of histone H3 (**Figure 3A**). Like wild-type cells, H3K56A mutants progressed synchronously through the meiotic S phase (**Figure 3B**) and showed normal transcriptome dynamics with no difference in gene expression during meiosis (**Figure 3C** and **Supplementary Table S1**). “Core”

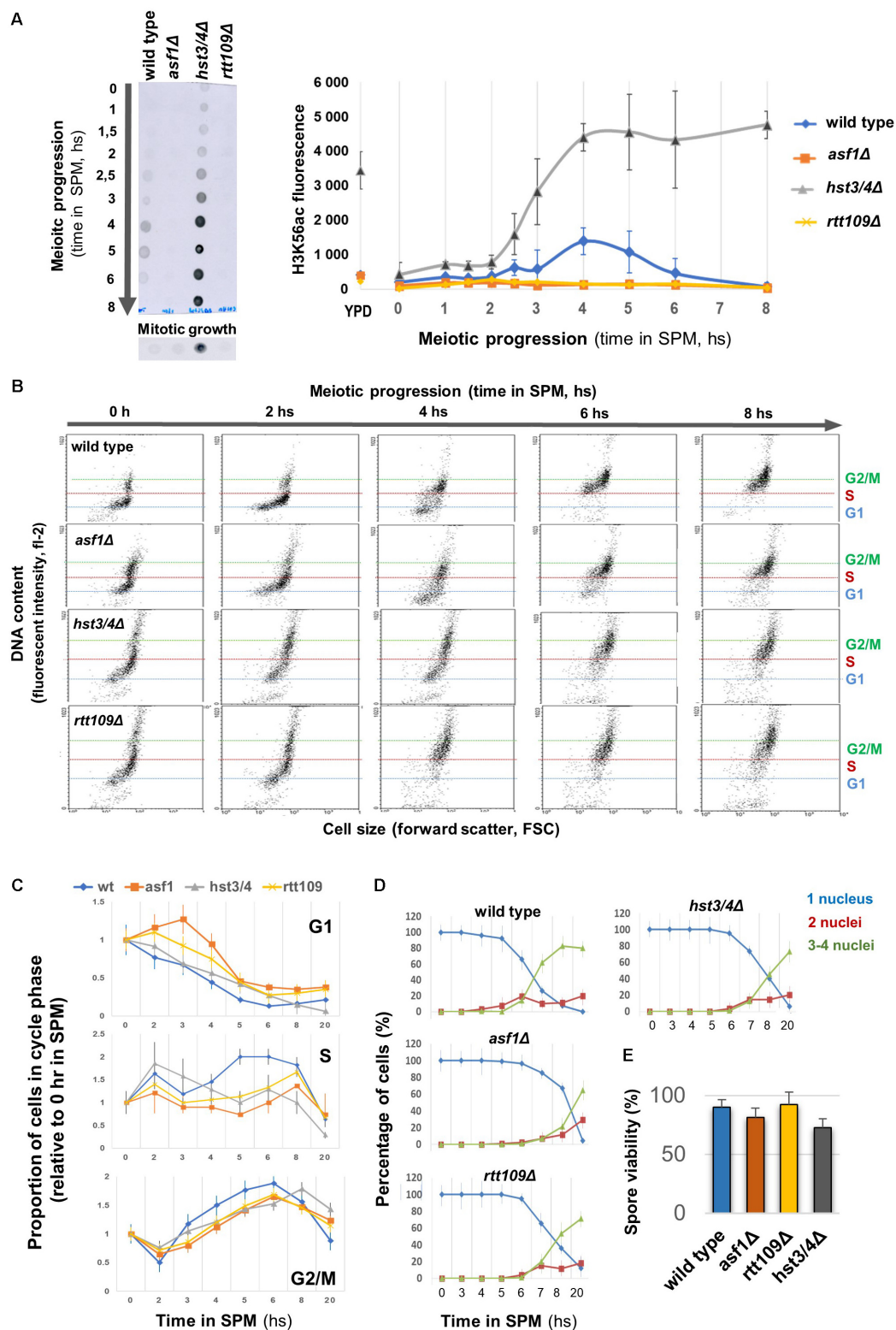


FIGURE 1 | Different H3K56ac levels allow normal meiotic progression and sporulation. **(A)** Left: Representative dot blot showing H3K56ac levels in wild type, *asf1Δ*, *rtt109Δ*, and *hst3/4Δ* cells. Right: Quantification of H3K56ac levels. YPD: mitotic growth. **(B)** Representative FACS profile of sporulating strains, showing DNA content in terms of cell size. G1, S, G2/M phases are indicated. **(C)** Proportion of cells in G1, S, and G2/M based on meiotic FACS profiles. Data were normalized to 0 h in SPM. **(D)** Sporulation efficiency. Nuclei were stained with DAPI and cells were scored for nucleus count. **(E)** Spore viability (50 tetrads per strain were counted). Panels **(A, C–E)** show the mean of three independent replicates. Error bars: SEM.

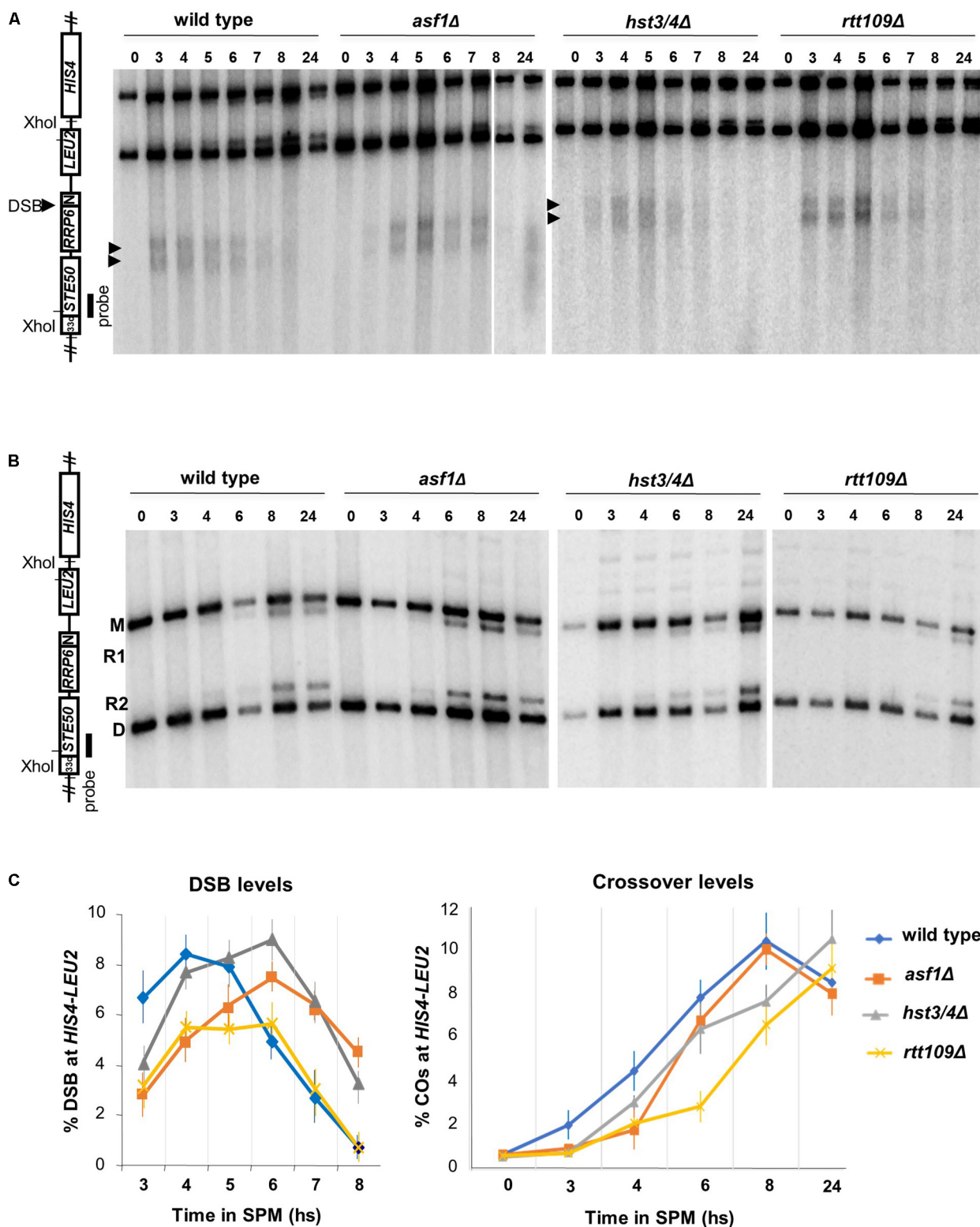


FIGURE 2 | Southern blot analysis of DSB and crossover formation at the *HIS4-LEU2* hotspot. **(A)** Representative blot showing *XhoI*-digested DNA from wild-type, *asf1Δ*, *hst3/4Δ*, and *rtt109Δ* cells. **(B)** Representative blot showing recombinant products (R1, R2) and parental bands (M, D) at *HIS4-LEU2*. **(C)** Quantification of DSB levels [% DSB signal/total lane signal] and recombination frequencies [% (R1 + R2) signal/total lane signal]. Mean values are from three independent replicates. Error bars: SEM.

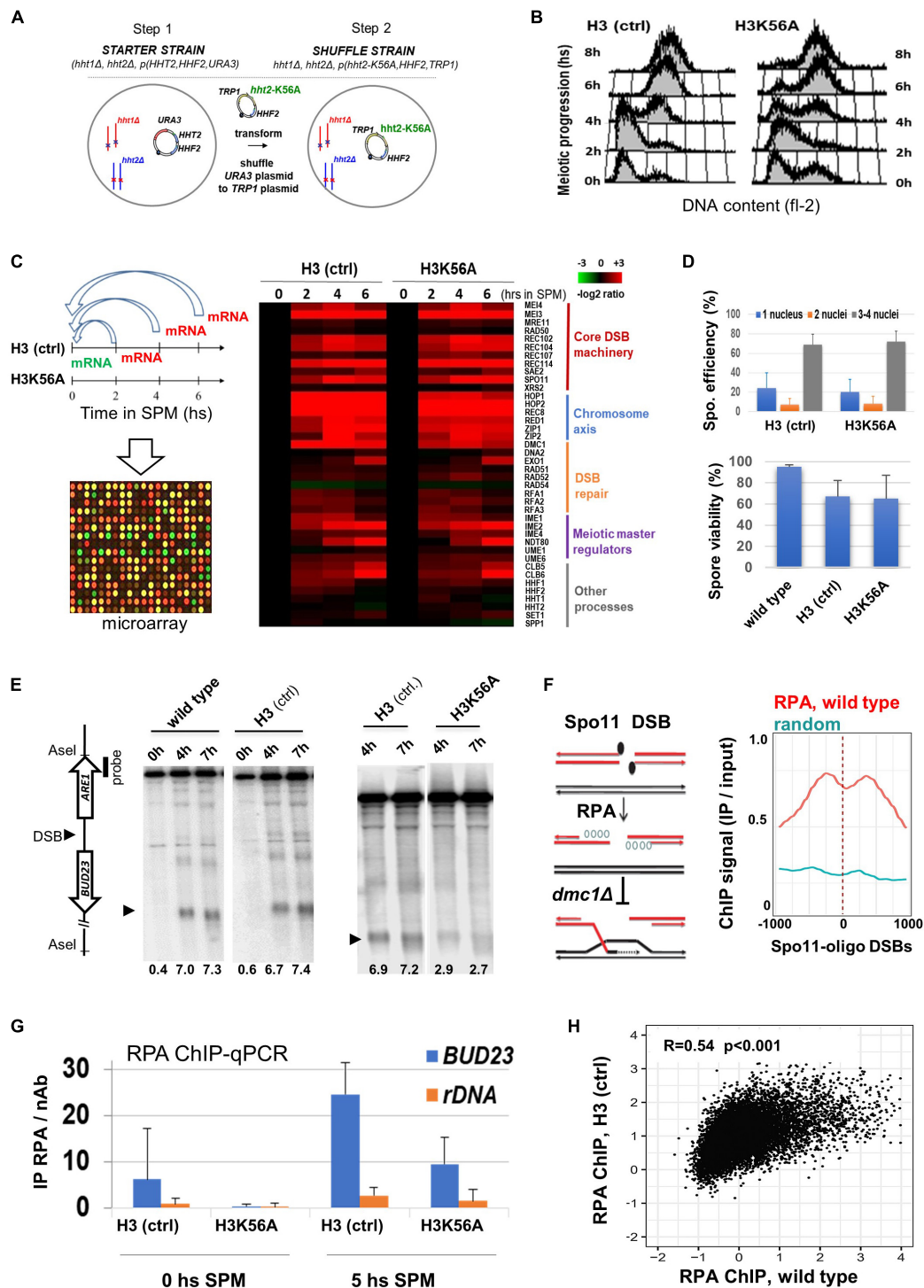


FIGURE 3 | Sporulation and meiotic DSB formation in a H3K56A mutant. (A) H3K56A cells were constructed by plasmid shuffling. **(B)** S phase progression in H3K56A and H3 (ctrl) cells measured by FACS. **(C)** Meiotic transcriptome dynamics in H3 (ctrl) and H3K56A strains. Left: Relative mRNA levels as measured by transcriptome microarrays at 2, 4, 6 h in SPM compared to 0 h in SPM. Right: Heatmap highlighting a selection of "core" genes that govern the process meiosis. The warmer the color, the greater the degree of induction. Full details are provided in **Supplementary Table S1**. **(D)** Sporulation efficiency and spore viability of H3K56A, H3 (ctrl), and wild-type strains. Error bars: SEM. **(E)** DSBs within *BUD23-ARE1*. Representative blot shows *Asel*-digested gDNA from wild type, H3 (ctrl) and H3K56A cells. DSB levels (below the blots): % DSB signal/total lane signal. **(F)** Left: Outline of the RPA ChIP approach, showing resected ssDNAs covered by RPA, which accumulates in *dmc1Δ* cells. Right: RPA enrichment over Spo11-oligo DSB sites and random sites in wild-type SK1 cells. Wild-type RPA data are from Borde et al. (2009). **(G)** ChIP-qPCR analysis of RPA enrichment at 0 and 5 h in SPM in the *BUD23* hotspot and the *rDNA* negative control site (known to lack DSBs). Error bars: SEM. **(H)** Correlation of RPA ChIP profiles in H3 (ctrl) and wild-type cells. R: Pearson correlation coefficient.

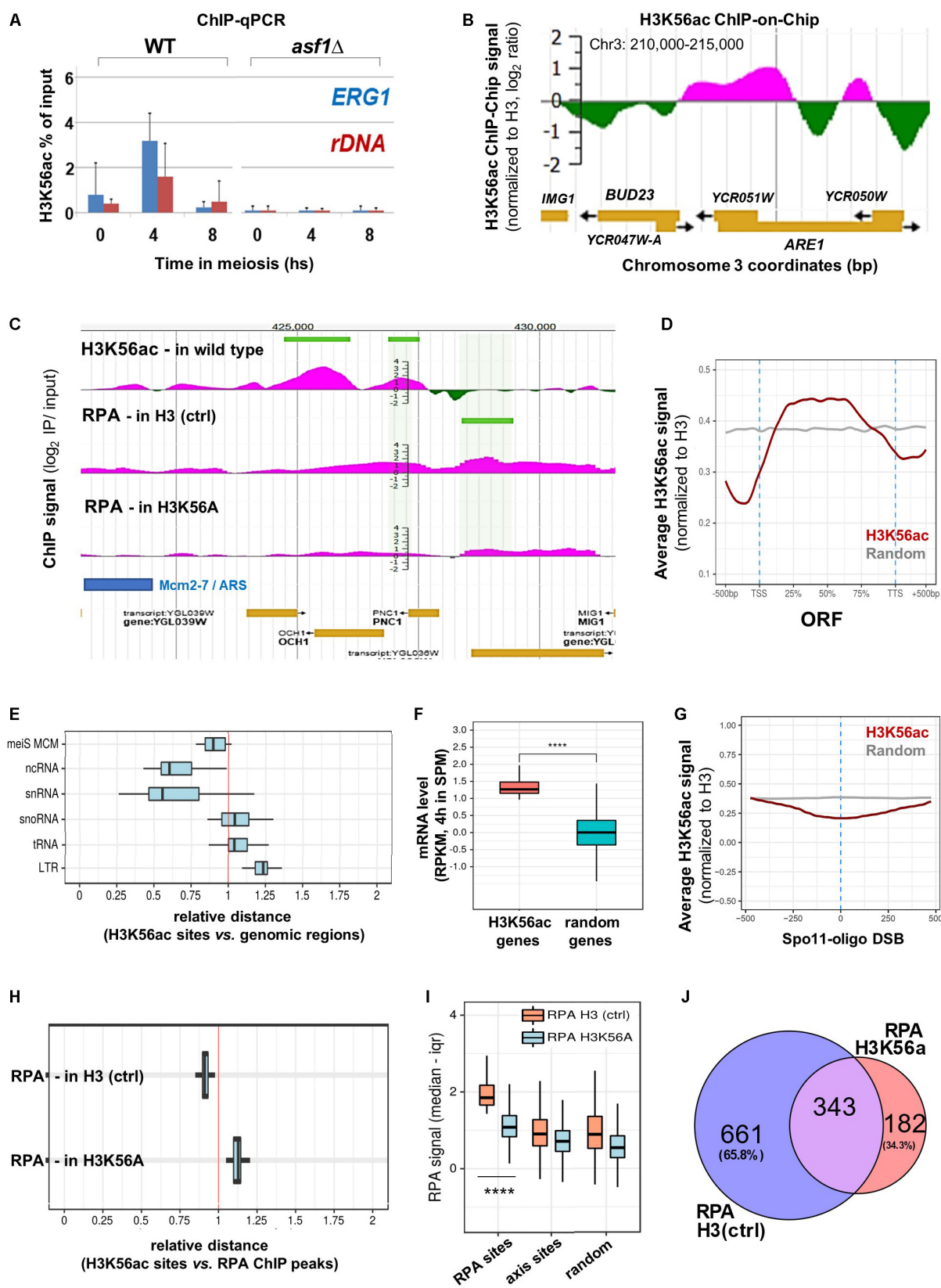


FIGURE 4 | Continued

FIGURE 4 | Genome-wide mapping of H3K56ac and RPA. **(A)** H3K56ac capture using an anti-H3K56ac antibody. H3 (C-ter) ChIP was performed to normalize to H3 occupancy. ChIP-qPCR enrichments were measured at *ERG1* and *rDNA*. Error bars: SEM. **(B)** H3K56ac signal detected by microarray (at 5 h in SPM). IP/input ratios were normalized to H3 occupancy (Borde et al., 2009). **(C)** Genome browser snapshot showing the distribution of H3K56ac in wild type cells and RPA enrichment in H3 (ctrl) and H3K56A cells. Peaks are highlighted by horizontal green lines. **(D)** Metagene profile of H3K56ac over protein-coding genes. Red curve shows the median H3K56ac signal. Gray curve: random signal. **(E)** Relative distance of H3K56ac peaks from various genomic regions. Vertical red line: random distances. Shift toward the left or right sectors: clustering or repulsion between H3K56ac peaks and the studied genomic elements. Mei MCM: meiotic replication origins (Blitzblau et al., 2012). **(F)** Genes associated with high H3K56ac levels show increased mRNA expression compared to random genes. Statistical significance: $p < 0.0001$ (Mann-Whitney *U* test). RNA-seq data are from Brar et al. (2012). **(G)** Depletion of H3K56ac over Spo11-oligo DSBs (Mohibullah and Keeney, 2017). Gray line: random signal. **(H)** Relative distance of H3K56ac peaks from RPA peaks in H3 (ctrl) and H3K56A cells. Vertical red line: random distance. **(I)** Distribution of RPA ChIP signal at RPA binding sites, axis sites (Mer2; Panizza et al., 2011), and random sites in H3 (ctrl) and H3K56A cells. Statistical significance: $p < 0.0001$ (Mann-Whitney *U* test). **(J)** Overlap of RPA peaks in H3 (ctrl) and H3K56A cells.

DSB, chromosome axis, and repair genes were properly transcribed in H3K56A cells, including *SPO11* and *RFA1* (Supplementary Figure S1), which were both essential for subsequent meiotic DSB mapping using RPA ChIP (see later). Sporulation efficiency and spore viability of the H3K56A mutant was not different from H3 (ctrl) expressing wild-type H3 (Figure 3D).

Next, we analyzed DSB formation in H3 *dmc1Δ* and H3K56A *dmc1Δ* cells that accumulate unrepaired DSB ends. DSB frequencies were determined by Southern blot within the natural hotspot region *BUD23-ARE1* (Figure 3E). Quantification of DSBs in wild-type and H3 (ctrl) strains confirmed the correct location, timing and frequency of DSBs in plasmid shuffle cells, demonstrating that our system accurately reports DSB formation. Importantly, a threefold reduction of DSB levels was observed in the H3K56A mutant, which was subsequently confirmed by an independent RPA ChIP method capturing Rfa1-covered ssDNA flanking Spo11-oligo DSBs (Figures 3F–H). In plasmid shuffle cells, RPA levels increased in the *BUD23* hotspot region by 5 h in SPM when DSBs are formed, and H3K56A mutants showed a ~threefold decrease in RPA levels relative to H3 (ctrl) (Figure 3G), consistent with our Southern blot results (Figure 3E). In addition, the RPA profiles of H3 (ctrl) and wild-type cells (Borde et al., 2009) were positively correlated (Figure 3H), whereas the RPA binding sites did not correlate with the binding of Mcm2-7 replicative helicase that marks meiotic DNA replication (Supplementary Figure S2). These results collectively demonstrate that RPA enrichment is an adequate indicator of meiotic DSB locations and frequencies.

Notably, the *BUD23* and *ERG1* hotspot regions were flanked by H3K56ac at the time of DSB formation as measured by H3K56ac ChIP-qPCR and ChIP-chip (Figures 4A,B). We note that the H3K56ac ChIP signal was undetectable in *asf1Δ* cells that are deficient in H3K56 acetylation (Figure 1A), confirming the specificity of our ChIP assay. Furthermore, the H3K56ac signal was also depleted in wild-type cells within the recombinationally cold *rDNA* region (Figure 4A). The same associations were observed at the genomic scale (Figure 4C and *JBrowse* link), suggesting that the presence of H3K56 residue and/or deposition of H3K56-acetylated nucleosomes near DSB hotspots is required for complete DSB formation. Our H3K56ac ChIP results show that H3K56 acetylation is preferentially associated with protein-coding ORFs (Figure 4D) and is depleted from promoters and transcription termination

sites [in line with published ChIP-qPCR data in mitotic cells (Schneider et al., 2006)]. The relative distance of H3K56ac from several genomic elements showed a clear non-random distribution (Figure 4E) as H3K56ac peaks were preferentially associated with Mcm2-7 helicase binding sites (Blitzblau et al., 2012), ncRNA and snRNA genes, but were further away from snoRNAs, tRNAs, and LTR retrotransposons compared to a computer-randomized chromosomal distribution (vertical red line in the figure). The proximity of H3K56ac peaks to Mcm2-7 binding sites agrees with the suggested role of H3K56 acetylation in marking nascent (replicating) chromatin upon replicative and repair DNA synthesis (Yu et al., 2012). Regarding meiotic gene expression (Figure 4F), protein-coding genes enriched in H3K56ac showed significantly higher mRNA expression levels in meiosis (at 4 h in SPM) than random genes [$p < 0.0001$; RNA-seq data are from Brar et al. (2012)]. Increased expression of H3K56-acetylated genes supports previous results in mitotic cells (Schneider et al., 2006; Williams et al., 2008), indicating that H3K56ac is an active chromatin mark associated with transcription.

Regarding meiotic DSB sites, chromosomal distribution of the H3K56ac signal showed a depletion over Spo11-oligo DSBs (Figure 4G), which precisely mark recombination hotspots. However, genomic positions of H3K56-acetylated histones were preferentially enriched adjacent to the identified RPA binding sites relative to random distance (Figure 4H, RPA H3 ctrl). Importantly, the RPA ChIP signal detected in the H3K56A mutant was strongly reduced at the RPA sites identified in the H3 (ctrl) strain ($p < 0.0001$, Mann-Whitney *U* test), as opposed to chromosome axis sites (Mer2-tagged; Panizza et al., 2011) and randomly selected sites that did not differ between control and mutant strains (Figure 4I). The preferential decrease of RPA ChIP signal at RPA binding sites in the H3K56A mutant demonstrates the specificity of H3K56A substitution mutation for recombination hotspot regions, accumulating hyper-resected ssDNA in the vicinity of DSBs. The specific decrease in the number of RPA peaks in the H3K56A mutant is highlighted in Figure 4J. This Venn diagram analysis shows that 65.8% of RPA peaks were eliminated by the H3K56A mutation (661 sites out of 1004), whereas 343 RPA sites were not affected or the signal was even increased (182 peaks). The latter RPA binding sites may represent unscheduled DSBs that are not related to H3K56 acetylation, which is clearly apparent from the increased relative distance between H3K56ac histones and RPA ChIP peaks detected in the H3K56A mutant (Figure 4H, RPA-H3K56A).

DISCUSSION

The above functional results highlight the association of H3K56 acetylation and meiotic DSB formation, suggesting that H3K56-acetylated histones are required to produce normal levels of DSBs within recombination hotspot regions. Nevertheless, the exact molecular mechanism underlying the DSB-promoting effect of H3K56 acetylation has yet to be clarified. The cause of the observed differences between the two mutant systems (histone modifying enzyme deletion vs. histone mutation) is currently not known. We obtained complementary results that are fully consistent with recent data identifying differences in the RNAPII binding profile of *rtt109Δ* and H3K56R mutants (Topal et al., 2019). Based on the genome-scale analysis of H3K56A cells, we propose that lack of H3K56 acetylation affects the stability or turnover rate of the well-positioned first nucleosomes flanking DSB sites (consistent with Kaplan et al. (2008)), and this may reduce the efficiency of Spo11-catalyzed DNA cleavage. Moreover, the absence of H3K56ac mark could reduce the rate of DSB end dissociation from nucleosomes flanking DSB sites, impeding the timely resection and processing of DSB ends. Alternatively, the H3K56A mutation may exert its effect indirectly on DSB formation, however, this is probably independent of changes in transcription since no differential gene expression was detected in the H3K56A mutant. A further possibility could be that H3K56 acetylation promotes the interaction of H3K4me3, Spp1, and Mer2 during the loop tethering process. The recently identified epistatic relationship of H3K56ac and H3K4me3 supports this hypothesis and seems particularly important in this regard (Voicheck et al., 2018). H3K56ac was found to act upstream of Set1C and H3K4 methylation, generating complementary H3K56ac/H3K4me3 histone modification patterns along newly replicated chromatin. These functional relationships could readily allow close cooperation between H3K56ac and H3K4me3 during meiotic recombination, especially because newly replicated chromatin is rich in H3K56ac and DNA replication is mechanically coupled to meiotic DSB formation (Murakami and Keeney, 2014). We assume that these complex spatial interactions occur in the

context of 3D chromatin structure. This could be detected by chromosome conformation capture methods (Dekker et al., 2017). Future use of these C-based approaches, together with mutant analyses, is expected to provide a deeper understanding of meiotic DSB formation with regards to the role of H3K4 methylation, H3K56 acetylation, and other potentially relevant histone modifications.

DATA AVAILABILITY STATEMENT

Datasets generated for this study can be accessed in **Supplementary Table S1** and via JBrowse (login: h3k56ac, password: Mozaic4, <http://geneart.med.unideb.hu/pub/h3k56ac>). Raw data are available at GEO (GSE37487).

AUTHOR CONTRIBUTIONS

ZK and LH analyzed the genomic data. LS performed the experiments and wrote the manuscript.

FUNDING

LS received funding from HAS-Lendület-LP2015-9/2015, NKFIH-NNF-130913, and GINOP-2.3.2-15-2016-00024.

ACKNOWLEDGMENTS

The authors thank Alain Nicolas (Institut Curie) for the conceptual support and critical comments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2019.00364/full#supplementary-material>

REFERENCES

- Abshiru, N., Ippersiel, K., Tang, Y., Yuan, H., Marmorstein, R., Verreault, A., et al. (2013). Chaperone-mediated acetylation of histones by Rtt109 identified by quantitative proteomics. *J. Proteomics* 81, 80–90. doi: 10.1016/j.jprot.2012.09.026
- Acquaviva, L., Székely, L., Dichtl, B., Dichtl, S. B., de La Roche Saint André, C., Nicolas, A., et al. (2013). The COMPASS subunit Spp1 links histone methylation to initiation of meiotic recombination. *Science* 339, 215–218. doi: 10.1126/science.1225739
- Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., et al. (2010). PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* 327, 836–840. doi: 10.1126/science.1183439
- Biswas, M., Voltz, K., Smith, J. C., and Langowski, J. (2011). Role of histone tails in structural stability of the nucleosome. *PLoS Comput. Biol.* 7: e1002279. doi: 10.1371/journal.pcbi.1002279
- Blitzblau, H. G., Chan, C. S., Hochwagen, A., and Bell, S. P. (2012). Separation of DNA replication from the assembly of break-competent meiotic chromosomes. *PLoS Genet.* 8: e1002643. doi: 10.1371/journal.pgen.1002643
- Borde, V., Robine, N., Lin, W., Bonfils, S., Géli, V., and Nicolas, A. (2009). Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites. *EMBO J.* 28, 99–111. doi: 10.1038/emboj.2008.257
- Brar, G. A., Yassour, M., Friedman, N., Regev, A., Ingolia, N. T., and Weissman, J. S. (2012). High-resolution view of the yeast meiotic program revealed by ribosome profiling. *Science* 335, 552–557. doi: 10.1126/science.1215110
- Buning, R., and Van Noort, J. (2010). Single-pair FRET experiments on nucleosome conformational dynamics. *Biochimie* 92, 1729–1740. doi: 10.1016/j.biochi.2010.08.010
- Celic, I., Masumoto, H., Griffith, W. P., Meluh, P., Cotter, R. J., Boeke, J. D., et al. (2006). The Siruins Hst3 and Hst4 preserve genome integrity by controlling Histone H3 Lysine 56 Deacetylation. *Curr. Biol.* 16, 1280–1289. doi: 10.1016/j.cub.2006.06.023
- Dekker, J., Belmont, A. S., Guttman, M., Leshyk, V. O., Lis, J. T., Lomvardas, S., et al. (2017). The 4D nucleome project. *Nature* 549, 219–226. doi: 10.1038/nature23884
- Govin, J., Dorsey, J., Gaucher, J., Rousseaux, S., Khochbin, S., and Berger, S. L. (2010). Systematic screen reveals new functional dynamics of histones H3 and H4 during gametogenesis. *Genes Dev.* 24, 1772–1786. doi: 10.1101/gad.1954910

- Kaplan, T., Liu, C. L., Erkmann, J. A., Holik, J., Grunstein, M., Kaufman, P. D., et al. (2008). Cell cycle- and chaperone-mediated regulation of H3K56ac incorporation in yeast. *PLoS Genet.* 4:e1000270. doi: 10.1371/journal.pgen.1000270
- Karányi, Z., Halász, L., Acquaviva, L., Jónás, D., Hetey, S., Boros-Oláh, B., et al. (2018). Nuclear dynamics of the Set1C subunit Spp1 prepares meiotic recombination sites for break formation. *J. Cell Biol.* 217, 3398–3415. doi: 10.1083/jcb.201712122
- Keeney, S., Giroux, C. N., and Kleckner, N. (1997). Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* 88, 375–384. doi: 10.1016/s0092-8674(00)81876-0
- Luger, K., Rechsteiner, T., Flaus, A., Waye, M., and Richmond, T. J. (1997). Characterization of nucleosome core particles containing histone proteins made in bacteria. *J. Mol. Biol.* 272, 301–311. doi: 10.1006/jmbi.1997.1235
- Messiaen, S., Guiard, J., Aigueperse, C., Fliniaux, I., Tourpin, S., Barroca, V., et al. (2016). Loss of the histone chaperone ASF1B reduces female reproductive capacity in mice. *Reproduction* 151, 477–489. doi: 10.1530/REP-15-0327
- Mohibullah, N., and Keeney, S. (2017). Numerical and spatial patterning of yeast meiotic DNA breaks by Tel1. *Genome Res.* 27, 278–288. doi: 10.1101/gr.213587.116
- Murakami, H., and Keeney, S. (2014). Temporospatial coordination of meiotic dna replication and recombination via DDK recruitment to replisomes. *Cell* 158, 861–873. doi: 10.1016/j.cell.2014.06.028
- Pan, J., Sasaki, M., Kniewel, R., Murakami, H., Blitzblau, H. G., Tischfield, S. E., et al. (2011). A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. *Cell* 144, 719–731. doi: 10.1016/j.cell.2011.02.009
- Panizza, S., Mendoza, M. A., Berlinger, M., Huang, L., Nicolas, A., Shirahige, K., et al. (2011). Spo11-accessory proteins link double-strand break sites to the chromosome axis in early meiotic recombination. *Cell* 146, 372–383. doi: 10.1016/j.cell.2011.07.003
- Parpanov, E. D., Petkov, P. M., and Paigen, K. (2010). Prdm9 controls activation of mammalian recombination hotspots. *Science* 327:835. doi: 10.1126/science.1181495
- Recht, J., Tsubota, T., Tanny, J. C., Diaz, R. L., Berger, J. M., Zhang, X., et al. (2006). Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6988–6993. doi: 10.1073/pnas.0601676103
- Rufange, A., Jacques, P. E., Bhat, W., Robert, F., and Nourani, A. (2007). Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. *Mol. Cell* 27, 393–405. doi: 10.1016/j.molcel.2007.07.011
- Schneider, J., Bajwa, P., Johnson, F. C., Bhaumik, S. R., and Shilatifard, A. (2006). Rtt109 is required for proper H3K56 acetylation: a chromatin mark associated with the elongating RNA polymerase II. *J. Biol. Chem.* 281, 37270–37274. doi: 10.1074/jbc.C600265200
- Simon, M., North, J. A., Shimko, J. C., Forties, R. A., Ferdinand, M. B., Manohar, M., et al. (2011). Histone fold modifications control nucleosome unwrapping and disassembly. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12711–12716. doi: 10.1073/pnas.1106264108
- Simoneau, A., Delgoushaie, N., Celic, I., Dai, J., Abshiru, N., Costantino, S., et al. (2015). Interplay between histone H3 lysine 56 deacetylation and chromatin modifiers in response to DNA damage. *Genetics* 200, 185–205. doi: 10.1534/genetics.115.175919
- Sommermeier, V., Beneut, C., Chaplais, E., Serrentino, M. E., and Borde, V. (2013). Spp1, a member of the Set1 Complex, promotes meiotic DSB formation in promoters by tethering histone H3K4 methylation sites to chromosome axes. *Mol. Cell* 49, 43–54. doi: 10.1016/j.molcel.2012.11.008
- Szekvolgyi, L., and Nicolas, A. (2010). From meiosis to postmeiotic events: homologous recombination is obligatory but flexible. *FEBS J.* 277, 571–589. doi: 10.1111/j.1742-4658.2009.07502.x
- Szekvolgyi, L., and Nicolas, A. (2010). From meiosis to postmeiotic events: homologous recombination is obligatory but flexible. *FEBS J.* 277, 571–589. doi: 10.1111/j.1742-4658.2009.07502.x
- Szekvolgyi, L., Ohta, K., and Nicolas, A. (2015). Initiation of meiotic homologous recombination: flexibility, Impact of Histone modifications, and chromatin remodelling. *Cold Spring Harb. Perspect. Biol.* 7, 447–462. doi: 10.1101/cshperspect.a016527
- Tian, H., Billings, T., and Petkov, P. M. (2018). CXXC1 is redundant for normal DNA double-strand break formation and meiotic recombination in mouse. *PLoS Genet.* 14:e1007657. doi: 10.1371/journal.pgen.1007657
- Topal, S., Vasseur, P., Radman-Livaja, M., and Peterson, C. L. (2019). Distinct transcriptional roles for Histone H3-K56 acetylation during the cell cycle in Yeast. *Nat. Commun.* 10, 1–13. doi: 10.1038/s41467-019-12400-5
- Tsubota, T., Berndsen, C. E., Erkmann, J. A., Smith, C. L., Yang, L., Freitas, M. A., et al. (2007). Histone H3-K56 acetylation is catalyzed by histone chaperone-dependent complexes. *Mol. Cell* 25, 703–712. doi: 10.1016/j.molcel.2007.02.006
- Voichek, Y., Mittelman, K., Gordon, Y., Bar-Ziv, R., Lifshitz Smit, D., Shenhav, R., et al. (2018). Epigenetic control of expression homeostasis during replication is stabilized by the replication checkpoint. *Mol. Cell* 70, 1121.e9–1133.e9. doi: 10.1016/j.molcel.2018.05.015
- Watanabe, S., Radman-Livaja, M., Rando, O. J., and Peterson, C. L. (2013). A histone acetylation switch regulates H2A.Z deposition by the SWR-C remodeling enzyme. *Science* 340, 195–199. doi: 10.1126/science.1229758
- White, C. L., Suto, R. K., and Luger, K. (2001). Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *EMBO J.* 20, 5207–5218. doi: 10.1093/emboj/20.18.5207
- Williams, S. K., Truong, D., and Tyler, J. K. (2008). Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9000–9005. doi: 10.1073/pnas.0800057105
- Xu, L., and Kleckner, N. (1995). Sequence non-specific double-strand breaks and interhomolog interactions prior to double-strand break formation at a meiotic recombination hot spot in yeast. *EMBO J.* 14, 5115–5128. doi: 10.1002/j.1460-2075.1995.tb00194.x
- Yu, Y., Song, C., Zhang, Q., DiMaggio, P. A., Garcia, B. A., York, A., et al. (2012). Histone H3 lysine 56 methylation regulates DNA replication through its interaction with PCNA. *Mol. Cell* 46, 7–17. doi: 10.1016/j.molcel.2012.01.019

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Karányi, Hornyák and Székvolgyi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



A Degenerate Peptide Library Approach to Reveal Sequence Determinants of Methyllysine-Driven Protein Interactions

Ariana Kupai, Robert M. Vaughan, Bradley M. Dickson and Scott B. Rothbart*

Center for Epigenetics, Van Andel Institute, Grand Rapids, MI, United States

OPEN ACCESS

Edited by:

Jean-Philippe Lambert,
Laval University, Canada

Reviewed by:

Mark T. Bedford,
The University of Texas MD Anderson
Cancer Center, United States

Joshua J. Hamey,
University of New South Wales,
Australia

*Correspondence:

Scott B. Rothbart
scott.rothbart@vai.org

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 20 February 2020

Accepted: 23 March 2020

Published: 09 April 2020

Citation:

Kupai A, Vaughan RM,
Dickson BM and Rothbart SB (2020)
A Degenerate Peptide Library
Approach to Reveal Sequence
Determinants of Methyllysine-Driven
Protein Interactions.
Front. Cell Dev. Biol. 8:241.
doi: 10.3389/fcell.2020.00241

Lysine methylation facilitates protein-protein interactions through the activity of methyllysine (Kme) “reader” proteins. Functions of Kme readers have historically been studied in the context of histone interactions, where readers aid in chromatin-templated processes such as transcription, DNA replication and repair. However, there is growing evidence that Kme readers also function through interactions with non-histone proteins. To facilitate expanded study of Kme reader activities, we developed a high-throughput binding assay to reveal the sequence determinants of Kme-driven protein interactions. The assay queries a degenerate methylated lysine-oriented peptide library (Kme-OPL) to identify the key residues that modulate reader binding. The assay recapitulated methyl order and amino acid sequence preferences associated with histone Kme readers. The assay also revealed methylated sequences that bound Kme readers with higher affinity than histones. Proteome-wide scoring was applied to assay results to help prioritize future study of Kme reader interactions. The platform was also used to design sequences that directed specificity among closely related reader domains, an application which may have utility in the development of peptidomimetic inhibitors. Furthermore, we used the platform to identify binding determinants of site-specific histone Kme antibodies and surprisingly revealed that only a few amino acids drove epitope recognition. Collectively, these studies introduce and validate a rapid, unbiased, and high-throughput binding assay for Kme readers, and we envision its use as a resource for expanding the study of Kme-driven protein interactions.

Keywords: lysine methylation, reader domains, functional proteomics, non-histone proteins, lysine-orientated peptide libraries

Abbreviations: BPTF-BRD-PHD, Bromodomain Plant Homeodomain Finger Transcription Factor Bromodomain-Plant Homeodomain; CBX, polycomb chromobox; CDYL1b chromo, Chromodomain Y-like protein 1b chromodomain; CDYL2 chromo, Chromodomain Y-like protein 2 chromodomain; DDO1 PHD, Death-Inducer Obliterator 1 Plant Homeodomain; GST, Glutathione S Transferase; JMJD2a TTD, Jumonji domain-containing protein 2a Tandem Tudor domain; Kme, methyllysine; Kme-OPL, methyllysine-oriented peptide library; L3MBTL1 3xMBT, Lethal 3 Malignant Brain Tumor-like protein 1 3 Malignant Brain Tumor domains; L3MBTL3 3xMBT, Lethal 3 malignant brain tumor-like protein 3 3 malignant brain tumor domains; MPP8 chromo, M Phase Phosphoprotein 8 chromodomain; PCL1 Tudor, Polycomb-like Protein 1; PHF20 Tudor, Plant Homeodomain Finger Protein 20; RFU, relative fluorescence units; UHRF1 TTD, Ubiquitin-like containing PHD and RING finger domains 1 Tandem Tudor domain; 53BP1 TTD, p53 Binding Protein 1 Tandem Tudor domain.

INTRODUCTION

Lysines can be mono-, di-, or tri-methylated on the sidechain ϵ -amino group (Ambler and Rees, 1959; Alix et al., 1979) and this post-translational modification can be “read” by proteins that contain methyllysine (Kme) binding domains (e.g., chromo, Tudor, MBT, PHD, etc.) (Liu et al., 2012). The first discovered Kme reader was heterochromatin protein 1 (HP1), whose chromodomain binds tri-methylated lysine 9 on histone H3 (H3K9me3) and facilitates HP1-mediated gene silencing (Bannister et al., 2001). Since this turn of the century discovery, more than 200 Kme reader proteins have been identified (Liu et al., 2012). Nearly all of these proteins have been studied as histone Kme readers and have been linked to various chromatin-associated functions like transcriptional regulation (Wozniak and Strahl, 2014), DNA repair (Botuyan et al., 2006) and DNA replication (Kuo et al., 2012).

The study of Kme reader-protein interactions is expanding beyond histones (Cornett et al., 2019). For example, M-phase phosphoprotein 8 (MPP8) is a Kme reader that, like HP1, was linked to gene silencing through recognition of H3K9me3 through its chromodomain (Kokura et al., 2010). MPP8 also functions in gene regulation through interactions with non-histone proteins like DNA methyltransferase 3a (DNMT3a) (Chang et al., 2011) and activating transcription factor 7-interacting protein 1 (ATF7IP) (Tsusaka et al., 2018). Other Kme readers, including HP1 (Liu et al., 2013), also have reported non-histone interactions (Cui et al., 2012; Ferry et al., 2017). Lysine methylation has been detected on over 3,000 unique human proteins (Hornbeck et al., 2015) but functions associated with Kmes are limited. This gap in knowledge has persisted in part because few technologies can directly associate proteins with Kme readers (Ong and Mann, 2006; Guo et al., 2014).

Here, we describe the development of a high-throughput assay for rapid, *in vitro* determination of where a Kme reader may bind in the proteome. The method identifies Kme-driven interactions by screening a Kme reader against a methyllysine-oriented peptide library (Kme-OPL) (Figure 1A). The OPL synthetic strategy is modified from the development of positional scanning peptide libraries (Houghten et al., 1991), and variations have been successfully applied to the study of other signaling processes, including phosphorylation and arginine methylation (Creixell et al., 2015; Gayatri et al., 2016). The degeneracy of the peptide library allows for the survey of all amino acid sequence combinations (excluding cysteine) minus to plus three (P-3/+3) from a central Kme. The assay informs on methyl order (Kme0, Kme1, Kme2, Kme3) preference and amino acid context, two key determinants of Kme reader interactions. Amino acid preferences are used to rank all lysine-centered motifs in the human proteome for each Kme reader, and these data are made available as a communal resource to help facilitate the identification of new Kme driven-protein interactions (Figure 1B). Additionally, Kme-specific antibodies can be used in place of Kme readers in this assay. Here, we report the use of the Kme-OPL assay for detecting the preferred methyl order of binding for multiple Kme readers, determining the optimal amino acid context for Kme reader

binding, and revealing the binding determinants of histone Kme-specific antibodies.

MATERIALS AND METHODS

Recombinant Protein Production

Plasmids encoding N-terminal GST fusions of each reader domain (Supplementary Table S1) were transformed into BL21 *E. coli* and protein expression was induced with 0.5 mM IPTG at 16°C for 6–16 h. Induced bacterial pellets were suspended in 30 mL cold 1× PBS supplemented with 1 mM DTT and 1 mM PMSF. Next, bacteria were incubated with lysozyme (Thermo #89833 LOT#ta262343) and 1 μ L of Pierce universal nuclease (Pierce #88702 LOT#00775219) on ice for 30 min followed by 3 rounds of sonication (30 s sonication with 10 s rest, all on ice) using a Qsonica ultrasonic processor (500 W 20 kHz with 1/8” microtip) at 40% amplitude. Lysed bacteria were centrifuged at 38465 rcf for 45 min at 4°C. Cleared supernatant was incubated with 5 mL of Glutathione resin (Thermo #16101 LOT#UD285112) with rotation at 4°C for 16 h. Bound protein was washed 3× with 10 mL cold 1× PBS and eluted twice with 10 mL of 25 mM HEPES pH 7.5, 10 mM L-glutathione (Sigma), and 100 mM NaCl. Protein was concentrated by centrifugation at 1,500 rpm on a Sorvall Legend X1 centrifuge in Amicon Ultra-15 centrifugal filter units (UFC#903024). Protein was resuspended in 25 mM HEPES pH 7.5 and 100 mM NaCl and concentrated by centrifugation as above three times. Protein was quantified by absorbance measurement at 280 nm divided by the computed extinction coefficient (ExpASy) (Gasteiger et al., 2003) of the GST-tagged protein domain.

Kme-OPL Reader Assay

Kme-OPL pools and sets were synthesized by PepScan as C-terminal PEG-biotin conjugates. Binding reactions were performed in 384 deep-well plates (Axygen #P-384-240SQ-C-S). The general procedure per reaction well was as follows. First, 2 μ L streptavidin magnetic bead slurry (Pierce, #88817) was washed with Buffer 1 (100 mM NaCl, 25 mM HEPES, pH 7.5, 0.5% BSA (w/v), 0.1% NP-40). Then, 2 μ g of peptide pool or set in water was added to washed beads and incubated for 30 min. Reactions were then collected by centrifugation at 2,000 rpm on a Sorvall Legend X1 centrifuge for 2 min prior to being placed on a plate magnet (Alpaqua A001222 LOT#1442). Solution was aspirated and beads were resuspended in 100 μ L Buffer 1. These four preceding steps comprised one wash. A second wash was performed, and beads were resuspended in 100 μ L GST tagged protein at 125 pmol per well in Buffer 2 (200 mM NaCl, 25 mM HEPES, pH 7.5, 0.5% BSA (w/v), 0.1% NP-40). Following a 30 min incubation, the well was washed 2× and beads were resuspended in 100 μ L of a 1:4,000 dilution of primary anti-GST antibody (Sigma #7781) in Buffer 1. Following another 30 min incubation, the well was washed 2× and beads were resuspended in 100 μ L of a 1:5,000 dilution of secondary anti-rabbit Alexafluor 488 (Invitrogen #A11034) for 30 min in Buffer 1. The well was again washed 2×, and beads were resuspended in 60 μ L of Buffer 1. 40 μ L was then transferred to a black 384-well plate (Corning #3575), and fluorescence intensity

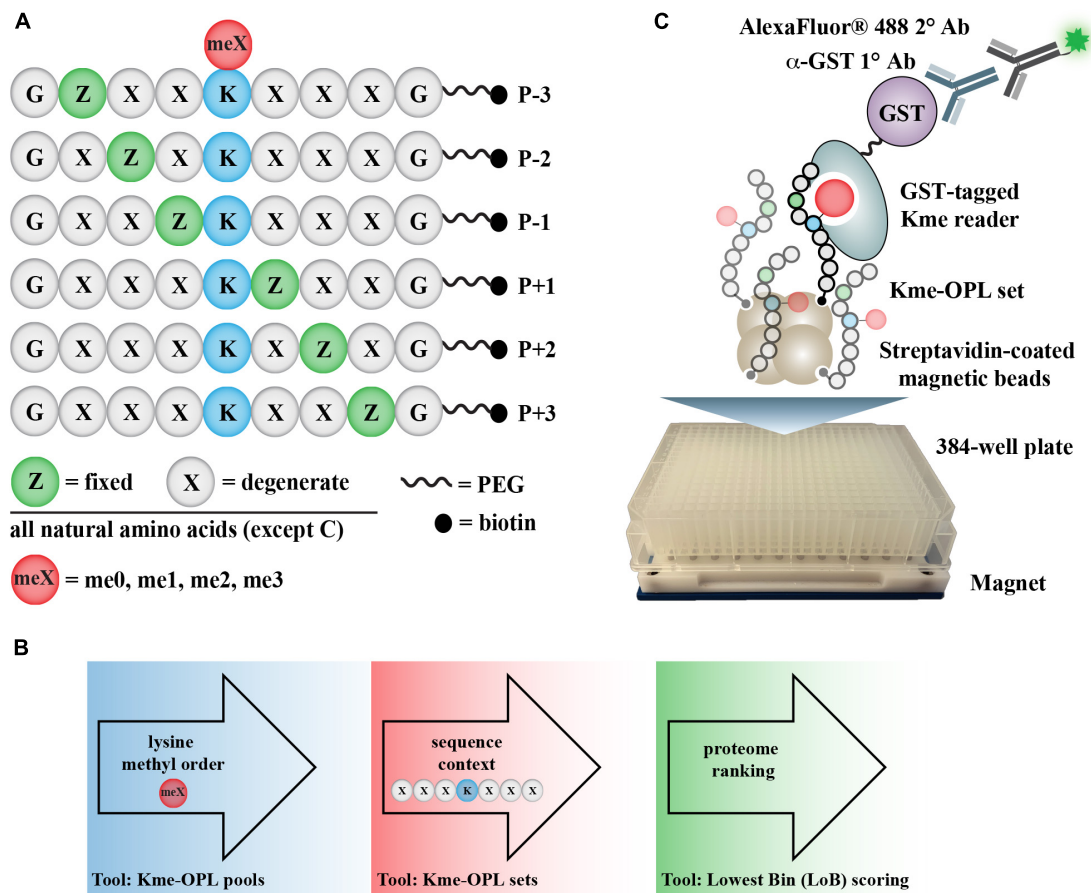


FIGURE 1 | Kme-OPL assay overview. **(A)** Design of Kme-OPL. **(B)** Schematic of assay workflow. **(C)** Cartoon of the magnetic bead pulldown assay developed for screening the activities of Kme readers. PEG, polyethylene glycol.

(485 ± 10 nm excitation filter and 528 ± 10 nm emission filter) was measured with a Synergy HT plate reader (Biotek). All steps were performed at room temperature, as cold incubations did not increase signal (data not shown). All incubation steps after bead resuspensions were performed with the plate on a shaker. For full library screens, peptide-bound beads in Buffer 1 were kept at 4°C for no more than 1 week. The INTEGRA assist plus pipetting robot was used for dispensing all buffers, protein, and antibodies as well as for washing steps. The primary anti-GST antibody alone gave appreciable, position-specific signal on Kme-OPL (Supplementary Figure S2). Therefore, all Kme-OPL reader profiles were performed in parallel with GST control reactions that were subtracted from signals obtained with GST-tagged readers.

Fluorescence Polarization Assay

Peptides functionalized with N-terminal 5-carboxyfluorescein (FAM) were synthesized by Genscript. All 7-mer motifs were synthesized with flanking glycines to mimic the Kme-OPL design. Binding assays were done in black 384 well plates (Corning #3575). Protein was serially diluted with 10 nM FAM peptide in FP assay buffer (25 mM HEPES pH 7.5, 100 mM NaCl,

0.05% NP-40). Polarization was measured on a Synergy Neo fluorescence plate reader (Biotek) with a 485 ± 10 nm excitation filter and a 528 ± 10 nm emission filter. Measurements were scaled to the last dilution of protein with a requested polarization of 20 milli-polarization units (mP). Anisotropy units (A) were calculated using the equation $A = (2P)/(3-P)$. Dissociation constants were determined by non-linear regression analysis of anisotropy curves by specific binding with Hill slope in GraphPad version 8.3.0.

Histone Peptide Microarrays

Peptide microarrays were fabricated using an Aushon 2470 microarrayer and used as described (Cornett et al., 2017) with the following modifications. Protein and antibody hybridization steps were performed in buffer containing 25 mM HEPES pH 7.5, 100 mM NaCl, 0.5% BSA (w/v), and 0.1% NP-40. Slides were washed 3×5 min in PBS supplemented with 0.1% Tween-20 between each hybridization step. Antibodies used were primary anti-GST (Sigma #7781, 1:2,000 dilution) and an AlexaFluor 647-labeled secondary antibody (Life Technologies A-21245, 1:5,000 dilution). Arrays were scanned using an Innopsys InnoScan 110AL microarray scanner and analyzed using ArrayNinja

(Dickson et al., 2016). Full lists of peptides queried by array analysis are in **Supplementary Table S2**.

Biotinylated Peptide Pulldowns

Biotinylated peptides were synthesized by the High Throughput Peptide Synthesis and Array Core Facility at UNC Chapel Hill. HEK293 cells were lysed in CSK Buffer (10 mM Pipes pH 7.0, 300 mM sucrose, 100 mM NaCl, 3 mM MgCl₂) supplemented with 0.1% Triton X-100, Roche Complete EDTA-free protease inhibitor tablet (#11 873 580 001), and Sigma phosphatase inhibitor cocktail 3 (#P0044) for 30 min on ice. Lysates were pre-cleared with 200 μ L streptavidin magnetic bead slurry (Pierce #88817) with rotation at room temperature for 30 min. 25 μ L bead slurry was washed in buffer containing 25 mM HEPES pH 7.5, 100 mM NaCl, 0.5% BSA (w/v), and 0.1% NP-40 and were then complexed with 50 μ g biotinylated peptide for 1 h at room temperature. 50 μ g of pre-cleared lysate was added to beads conjugated with peptide and volume was brought up to 500 μ L with pulldown buffer. Following a 4-h incubation at 4°C with rotation, beads were washed 3 \times 5 min with 500 μ L of wash buffer containing 25 mM HEPES, pH 7.5, 200 mM NaCl, 0.5% BSA (w/v), and 0.1% NP-40. Peptides and protein were eluted in 25 μ L of 1 \times SDS loading dye by heating at 95°C for 5 min prior to loading on a 7% polyacrylamide gel for SDS-PAGE. Transfer to PVDF membrane was performed at 45 mA for 90 min using a Hoefer TE77X semi-dry transfer unit. Membranes were blocked with 5% BSA in PBS-T for 15 min prior to incubation with a 1:1,000 dilution of anti-CDYL2 antibody (ab183854 LOT:GR240986-6) in blocking buffer overnight at 4°C with rotation. Membranes were washed 3 \times 5 min with 1 \times PBS-T and incubated in a 1:10,000 dilution of HRP-conjugated secondary antibody (GE #NA934V) in blocking buffer. Membranes were exposed to ECL substrate (Pierce 32209) following 3 \times 5 min washes with 1 \times PBS-T and imaged with Kodak \times omat. Images were quantified with ImageJ Version 1.52.

RESULTS

A Kme Reader Assay Querying Kme-OPL

The Kme-OPL reader platform is a plate-based magnetic bead pulldown assay read out by fluorescence intensity (**Figure 1C**). The library is oriented around a central lysine, which can have one of four possible methyl orders (Kme0, Kme1, Kme2, or Kme3) (**Figure 1A**). Within each methyl order, the library is organized into 114 Kme-OPL sets, where each set has one amino acid fixed in one position. All other positions contain 19 amino acids in an equimolar, degenerate mix. Cysteine is excluded due to incompatibility with the synthetic approach. The peptides are biotinylated, which allows for binding to streptavidin magnetic beads. The beads and Kme-OPL sets are first complexed, and then a recombinant GST-tagged Kme reader is added (**Figure 1C**). Next, a primary GST antibody followed by a secondary antibody conjugated to a fluorophore are added. Binding is read out by fluorescence intensity measurements. The optimization of several assay components is present in **Supplementary Figure S1** and further detailed in section “Materials and Methods.”

Methyl Order Preferences for Histone Kme Readers Are Recapitulated With Kme-OPL

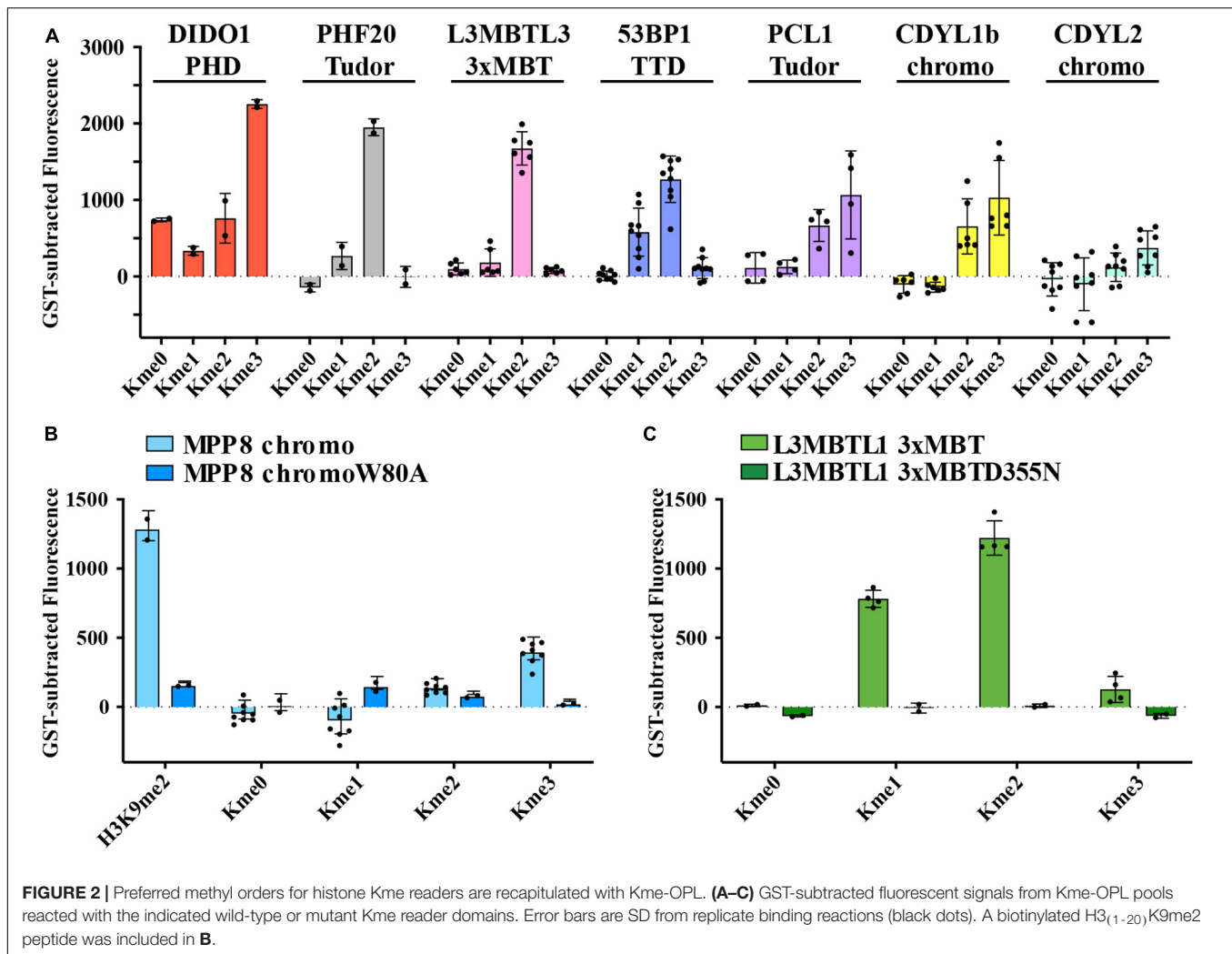
Kme readers have been reported to prefer the same lysine methyl order on histone and non-histone proteins (Cui et al., 2012; Liu et al., 2013; Ferry et al., 2017). We first tested whether the Kme-OPL platform could detect Kme reader methyl order preference. To measure preferred methyl order, we synthesized Kme-OPL pools, where all peptides sets with the same methyl order are combined into a single pulldown reaction. We queried nine reader domains known to bind histone Kmes (**Table 1**, **Supplementary Figures S3A, S4A** and **Figure 5C**). Each measurement is reported as a GST subtracted value (**Figure 2**). Most values had a simultaneous GST measurement subtracted. A small subset of experiments had high GST signals for unknown reason (**Supplementary Figure S2**). For experiments without a simultaneous GST measurement, we inferred whether low or high background signal should be subtracted based on the signal from the Kme0-OPL pool. In each case, binding to Kme-OPL pools was consistent with reported histone methyl order preferences (**Figure 2A**). To further test if signal was dependent on Kme binding, we assayed mutant forms of MPP8 chromo and L3MBTL1 3xMBT which had single amino acid substitutions known to disrupt their interactions with Kmes (Li et al., 2007; Chang et al., 2011). In both mutants, signal intensities were reduced to GST background levels (**Figures 2B,C**). We note variability in max signals with Kme-OPL pools, which we interpret as either weak overall affinity or high affinity to a limited set of peptides. Later, we resolve this mixed interpretation.

Kme-OPL Reports on Sequence Determinants of Kme Reader Specificity

We next used the Kme-OPL platform to determine how the amino acid sequence surrounding the Kme modulated reader binding (Kme-OPL profile). We used this sequence data in conjunction with Lowest Bin (LoB) scoring

TABLE 1 | Reported histone interactions for Kme readers queried in **Figure 2**.

Protein domain	Associated histone mark
MPP8 chromo	H3K9me2/3 (Kokura et al., 2010)
L3MBTL1 3xMBT	H1bK26me1/2 (Trojer et al., 2007) H4K20me1/2 (Min et al., 2007)
DIDO1 PHD	H3K4me3 (Gatchalian et al., 2016)
PHF20 Tudor	H3K4me2 (Klein et al., 2016) H4K20me2 (Klein et al., 2016)
L3MBTL3 3xMBT	Many Kme2 (Nady et al., 2012)
53BP1 TTD	H4K20me1/2 (Botuyan et al., 2006; Hartlerode et al., 2012) H3K18me2 (Shanle et al., 2017) H3K36me2 (Tong et al., 2015)
PCL1 Tudor	H3K36me3 (Cai et al., 2013)
CDYL1b chromo	H3K9me2/3 (Franz et al., 2009; Escamilla-Del-Arenal et al., 2013) H3K27me3 (Vermeulen et al., 2010)
CDYL2 chromo	H3K9me3 (Fischle et al., 2008) H3K27me3 (Fischle et al., 2008)



(Cornett et al., 2018) to predict where these readers may bind in the proteome (Figure 1B). For these studies, MPP8 and CDYL2 chromodomains were chosen because aspects of their amino acid binding preferences are reported elsewhere (Li et al., 2011; Barnash et al., 2016), and these data were consistent with Kme-OPL profiles generated with these readers.

The MPP8 chromodomain structure is a three stranded antiparallel β sheet with a C terminal α helix (Li et al., 2011). In the H3K9 tail sequence, Q5, T6, and A7 interact with the residues V58, F59, E60 and V61 in MPP8 and induce creation of another β strand (β 1), forming a β hairpin (Li et al., 2011; Figure 3A). H3S10 forms a non-backbone hydrogen bond with MPP8 residue E91. For residues succeeding S10, no interactions are observed. Because conformational induction of the β hairpin is essential for Kme binding, a specific sequence context that will recapitulate these contacts is necessary. These data suggest the MPP8 Kme3-OPL pool signal is low (Figure 2A) because a limited number of pool peptides can induce this conformational change.

The MPP8 chromo Kme-OPL profile had clear position preferences that aligned with the crystal structure (Figure 3B and Supplementary Figures S3B,C). As with Kme-OPL pools

measurements, Kme-OPL profiles were GST background subtracted (Supplementary Figure S2B). In the MPP8 Kme3-OPL profile, P-3 slightly preferred basic or aromatic amino acids. P-2 strongly favored A/G. The preference of P-2 toward smaller amino acids is likely due to this position's location inside the β hairpin. P-1, a position that performs Van der Waals interactions (Li et al., 2011), preferred K/R/I. The same P-2 and P-1 preferences have been reported for the chromodomains of CBX proteins (Kaustov et al., 2011). CBX2/3/5/6/7/8 all had P-2 in a small hydrophobic pocket that could only fit alanine or smaller residues, and CBX7 preferred P-1 R/I/L/F/Y/V. The conserved preferences across chromodomains further support our results for MPP8. Continuing with the MPP8 profile, P+1 strongly favored S/T, likely because of the ability of these amino acids to form a non-backbone hydrogen bond. MPP8 chromo did not have a P+2 amino acid preference, which is in accordance with no contacts being made in this position in the H3 co-structure (Figure 3A). Although there are also no contacts being made in P+3 in the co-structure, our assay revealed a P+3 preference for lysine (Figure 3B). Figure 3C shows a histogram of MPP8 Kme3-OPL set signals and provides

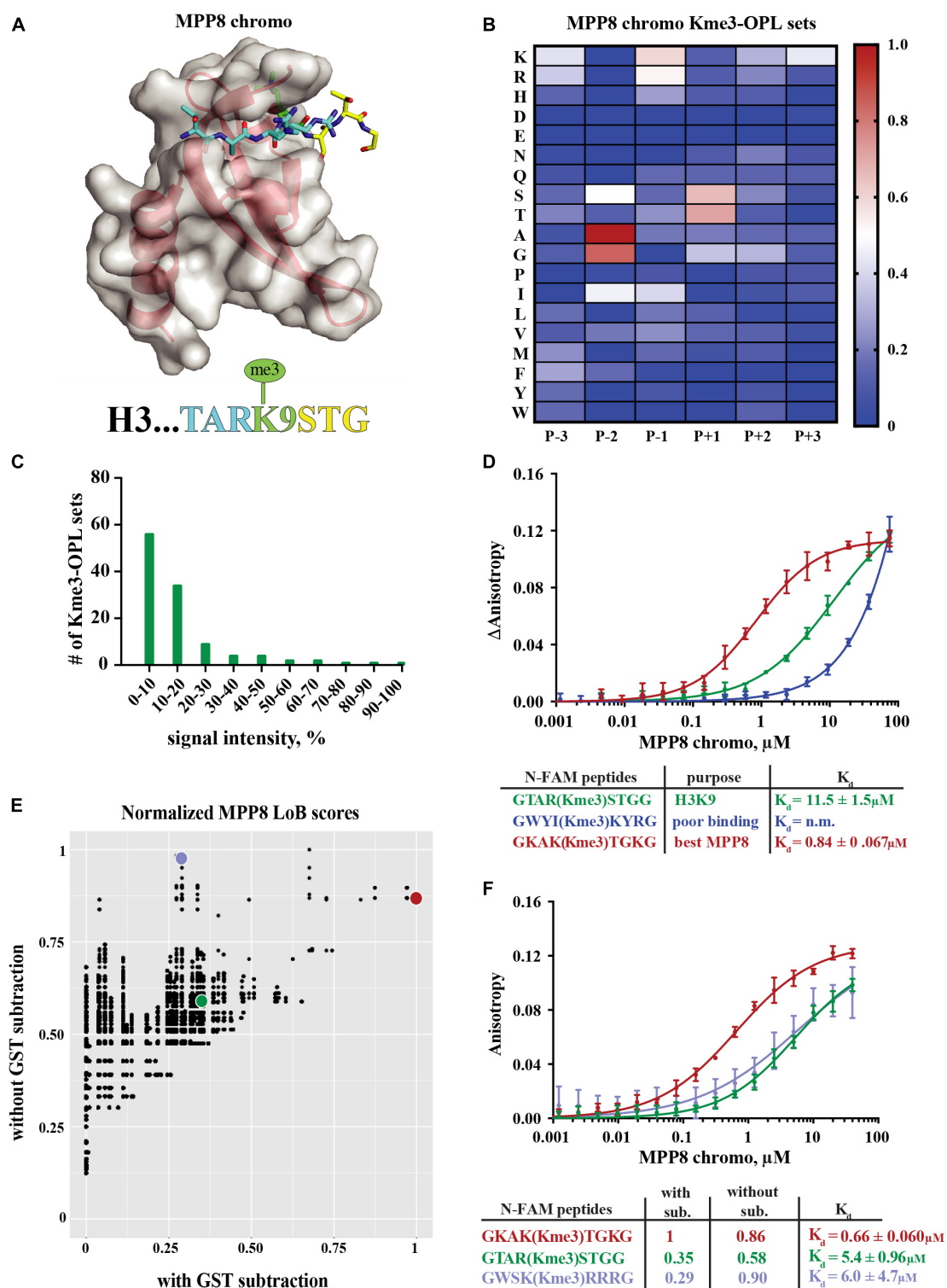


FIGURE 3 | Binding determinants of MPP8 chromo and validation of Lowest Bin (LoB) scoring. **(A)** MPP8 chromo co-structure with an H3K9me3 peptide (PDB:3QO2). **(B)** MPP8 chromo Kme3-OPL profile. Each Kme3-OPL set is shown in the heatmap as an average of 4 replicate GST subtracted fluorescence measurements, and data is scaled from 0 (no signal, blue) to 1 (highest average signal, red). **(C)** MPP8 chromo specificity profile. Specificity is graphed as the number of Kme3-OPL sets with a given signal intensity range. **(D)** MPP8 chromo fluorescence polarization. Data points are plotted as an average of 4 measurements from 2 independent experiments. Error is SD **(E)** Scatterplot of normalized MPP8 LoB scores from data with GST subtraction vs without GST subtraction. Red, green, and blue points correspond to same colored peptides as in **F**. **(F)** Fluorescence polarization of MPP8 chromo. Data points are plotted as an average of 5 replicate measurements from 2 independent experiments. Error is SD.

an easy way of determining if a protein is sequence specific. We consider MPP8 chromo to be sequence specific because only a few Kme3-OPL sets had high signals while the majority were shifted toward lower values. Collectively, the MPP8 chromo Kme3-OPL profile showed preference for several amino acids that would be predicted from the structure of MPP8 chromo bound to H3K9me3.

We next asked whether we could use the MPP8 Kme3-OPL profile to predict an optimal binding sequence. We predicted the best binding sequence, KAK(Kme3)TGK, by choosing the Kme3-OPL set with the highest signal in each position. We compared this sequence to the sequence surrounding H3K9me3, TAR(Kme3)STG, and also to a predicted poor binding sequence, WYI(Kme3)KYR, chosen by picking Kme3-OPL sets with low signals in each position. We measured the K_d of the MPP8 chromo interaction with each peptide using fluorescence polarization (FP). The predicted poor binding peptide had a K_d that was too weak to be determined (Figure 3D). The best predicted peptide had a K_d of $0.84 \pm 0.067 \mu\text{M}$, binding ten-fold tighter than the H3K9me3 peptide. Since this tight-binding peptide was present in the Kme3-OPL pools, the low overall signal of MPP8 toward the Kme3-OPL pool was unlikely due to low affinity to the entire pool (Figure 2). Rather, MPP8 bound strongly to only a few sequences that were diluted in the pools, resulting in a lower Kme-OPL pool signal.

In order to relate Kme-OPL profiles to the human proteome, we used our previously developed LoB scoring function (Cornett et al., 2018). LoB scoring ranks all lysine centered seven-mers in the proteome from most to least likely to bind to a given reader. This was originally developed to identify lysine methyltransferase substrates but can also be applied to identify Kme reader interactions. LoB scoring minimizes false positives by having the lowest Kme-OPL set dictate the score. All LoB scoring is deposited at https://github.com/ariana-kupai/LoB_scores as GST background subtracted values. The reasons for GST subtraction are described below.

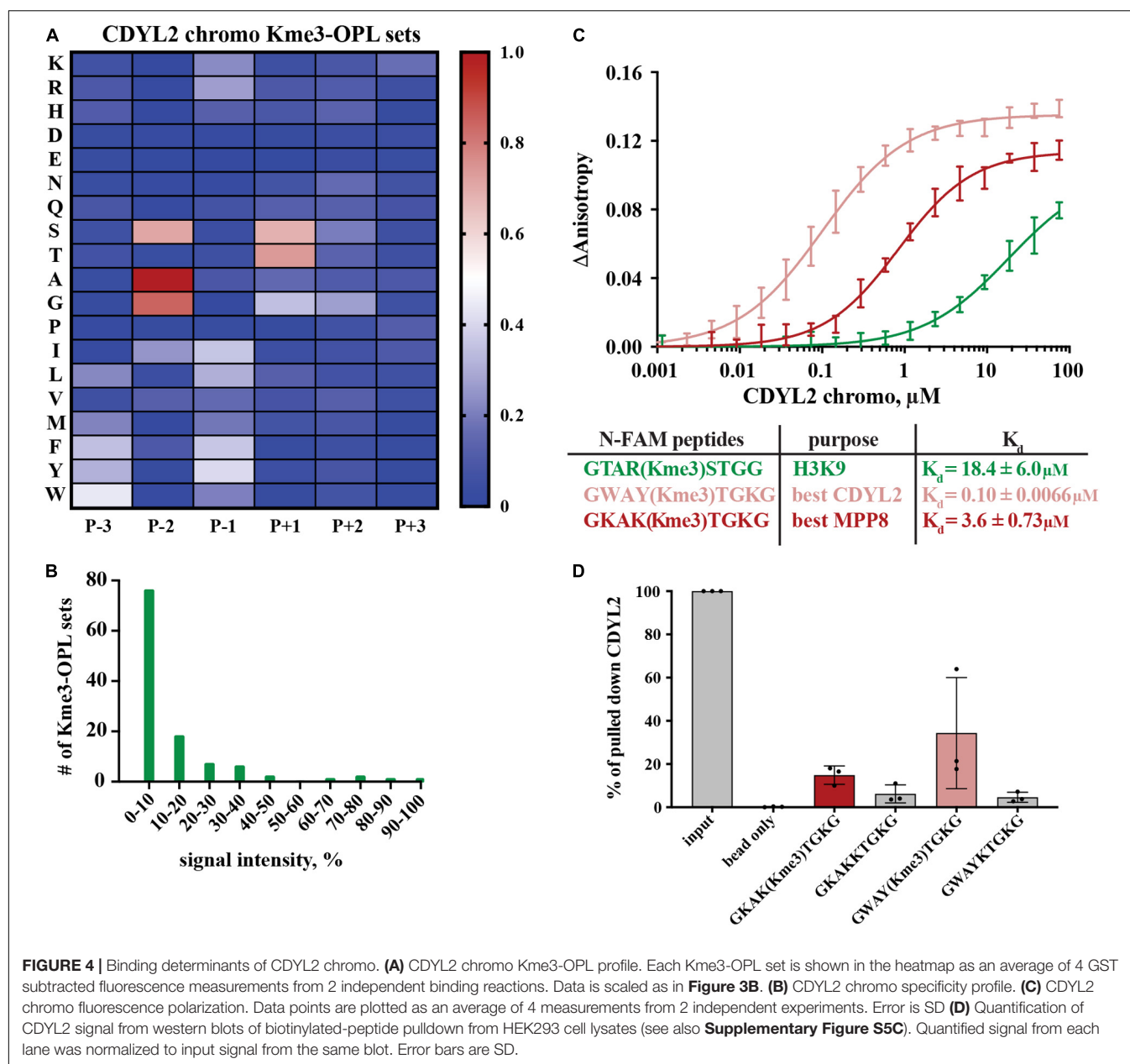
LoB scores generated from MPP8 chromo screening were plotted with and without GST subtraction and normalized to their respective highest score to facilitate comparisons (Figure 3E). Each dot on the scatterplot represents a lysine-centered seven-mer in the proteome. Three peptides were chosen with ranging LoB scores (TAR(Kme3)STG, KAK(Kme3)TGK, and WSK(Kme3)RRR) for comparison in FP binding assays (Figure 3F). With MPP8 chromo, TAR(Kme3)STG had a K_d of $5.4 \pm 0.96 \mu\text{M}$, KAK(Kme3)TGK had a K_d of $0.66 \pm 0.060 \mu\text{M}$, and the K_d for WSK(Kme3)RRR was $6.0 \pm 4.7 \mu\text{M}$. The FP binding results led us to conclude that background subtracted LoB scores were more reflective of *in vitro* binding constants. Consequently, we have reported LoB scoring only on GST subtracted Kme-OPL profiles, which should help further reduce selection of false positives for downstream studies. Of note, the preferred sequence for MPP8 chromo, KAKKTGK, mapped to Calcium permeable stress-gated cation channel 1 (CSC1) in the human proteome. However, because CSC1 localizes to the plasma membrane and MPP8 is found in the nucleus, this interaction is not likely to be physiologically relevant.

MPP8 and CDYL2 chromodomains have similar structures and both recognize H3K9me2/me3 (Fischle et al., 2008; **Supplementary Figures S3A, S4A**). We next sought to compare Kme-OPL profiles of these closely related proteins. Certain positional binding preferences were conserved between MPP8 and CDYL2 chromodomains, as anticipated from their recognition of the same histone Kme. In both Kme-OPL profiles, P-2 was the most selective position, favoring A/G (**Figures 3B, 4A** and **Supplementary Figure S4B**). Both proteins also favored P+1 S/T. CDYL2 had a more specific profile than MPP8 (**Figures 3C, 4B**). MPP8 signals tapered off while CDYL2 signals had a bimodal distribution, signifying more amino acids promoted or inhibited binding.

Peptide scaffolds are the basis for some Kme reader antagonists (James et al., 2013; Simhadri et al., 2014; Stuckey et al., 2016). Notably, the Kme-OPL profile of CDYL2 highlighted similar characteristics of amino acids in the CDYL2 Kme peptidomimetic inhibitor, UNC4991 (Barnash et al., 2016). UNC4991 has P-3 F, consistent with the preference the Kme3-OPL profile showed for P-3 non-polar aromatic residues (Figure 4A). UNC4991 has P-2 A and our assay showed preference for P-2 A/G. UNC4991 has P-1 F and our assay showed preference for P-1 non-polar residues. UNC4991 has P+1 T and our assay had P+1 preference for S/T. UNC4991 lacks P+2 and P+3 residues, consistent with the lack of amino acid preference in these positions on Kme-OPL. Using the Kme-OPL assay, we converged upon the same characteristics of amino acids that promoted peptide binding to CDYL2. This further validates our CDYL2 chromo Kme3-OPL profile and demonstrates the potential utility of the Kme-OPL platform for designing peptide-based inhibitors.

We next used the Kme3-OPL profile for CDYL2 to predict an optimal binding sequence. The predicted best binding sequence was WAY(Kme3)TGK, which had a K_d of $0.10 \pm 0.0066 \mu\text{M}$ as measured by FP. This peptide, which does not map to any human protein, bound 100-fold tighter than the H3K9me3 peptide (Figure 4C). We also measured the K_d of CDYL2 chromo with KAK(Kme3)TGK, the best MPP8 peptide. This interaction had a K_d of $3.6 \pm 0.73 \mu\text{M}$. We functionalized these sequences (methylated and unmethylated) with biotin and performed peptide pulldowns for CDYL2 from HEK293 cell lysates. A lysate titration and western blot images from three independent experiments are in **Supplementary Figure S5**. In three replicate experiments, the tri-methylated sequences pulled down more CDYL2 than the unmethylated sequences. In one of the replicates, WAY(Kme3)TGK pulled down more CDYL2 than KAK(Kme3)TGK, consistent with the *in vitro* observation (Figure 4D). These pulldowns were performed with 3 independent preparations of cell lysate. Therefore, variables like cell cycle distribution of the bulk population, protein posttranslational modifications, and abundance of competitively binding proteins cannot be ruled out as variables impacting the reproducibility of these and other pulldowns from cell extracts.

We also measured the K_d of MPP8 chromo with WAY(Kme3)TGK, the best CDYL2 peptide (**Supplementary Figure S4C**). The K_d s of MPP8 chromo with the best MPP8



peptide and the best CDYL2 peptide were very similar (**Figure 3D**). Comparatively, MPP8 is a less specific reader than CDYL2. P-2 A and P(+1 T may have greater impact than other positions for driving interactions with MPP8, making the MPP8 and CDYL2 best peptides equally strong binding sequences. Collectively, these results show the Kme-OPL platform can be used to identify preferred amino acid sequences for very specific reader domains.

Kme-OPL Recognized Promiscuity in Kme Reader Binding

Kme-OPL reported amino acid binding preferences for sequence specific Kme readers, so we next sought to determine what

Kme-OPL would report for non-specific readers. L3MBTL3 is a promiscuous Kme2 reader whose 3xMBT domain binds to many Kme2 histone contexts *in vitro* (Nady et al., 2012). L3MBTL3 was also classified as a promiscuous Kme2 reader by our assay. The Kme2-OPL profile for L3MBTL3 3xMBT tolerated all residues (**Figure 5A**), and most Kme-OPL sets had high signals (**Supplementary Figures S6A,B** and **Figure 5B**). These results are consistent with previous studies that show surrounding amino acids do not impact L3MBTL3's mechanism of Kme recognition (Li et al., 2007) or potency of the L3MBTL3 peptidomimetic inhibitor UNC1215 (James et al., 2013); both of which lack protein-peptide contacts outside of the Kme (**Supplementary Figure S6C**). Also consistent with a previous report (Nady et al., 2012), the lowest signals in the L3MBTL3

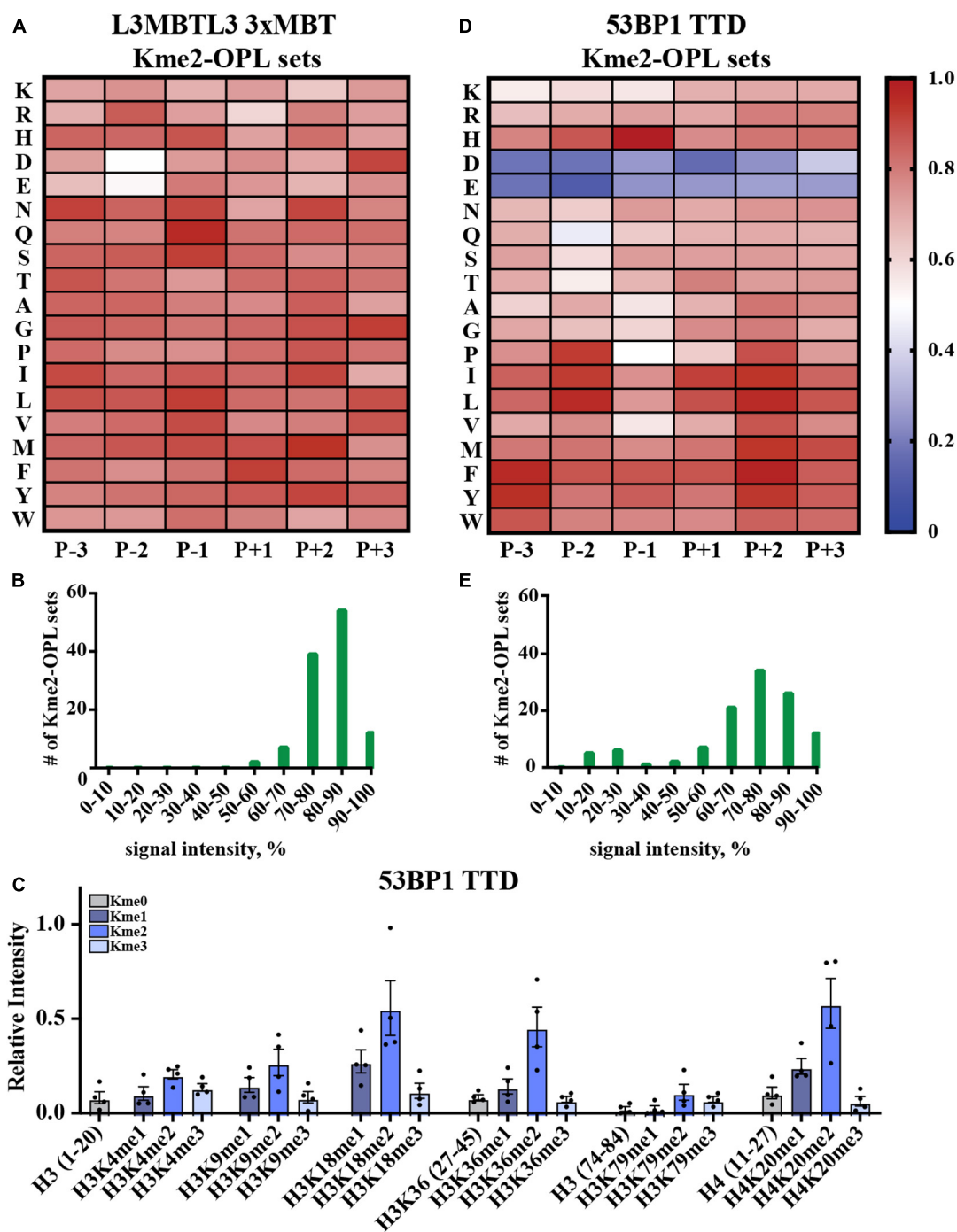
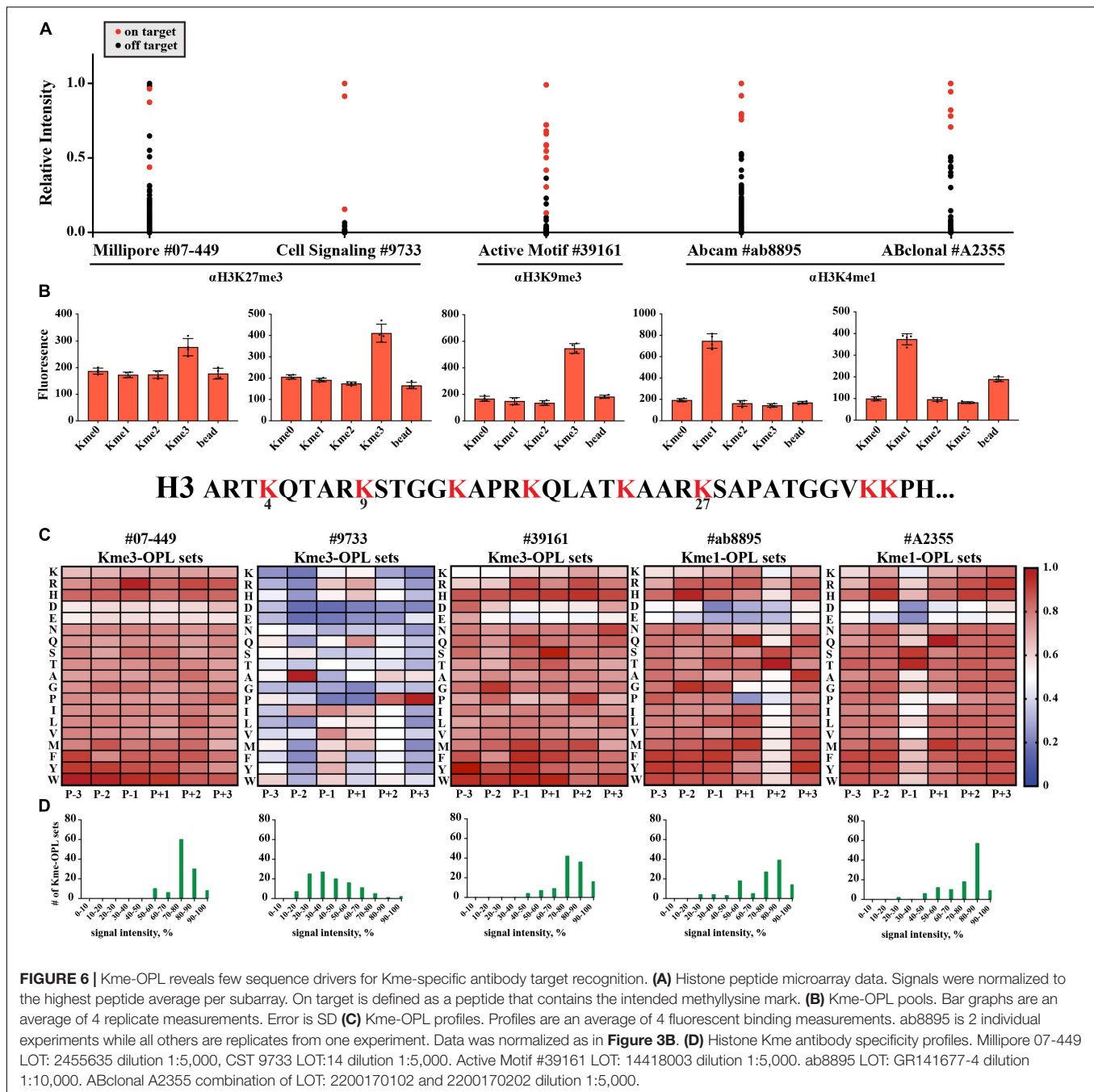


FIGURE 5 | Promiscuity in Kme binding is recognized by Kme-OPL. **(A)** L3MBTL3 3xMBT Kme2-OPL profile. Two independent binding measurements with 2–3 technical replicate measures for each Kme2-OPL set were averaged. Data was GST background subtracted and normalized as in **Figure 3B**. **(B)** L3MBTL3 3xMBT specificity profile. **(C)** 53BP1 TTD histone peptide microarray data. Data points were normalized to the highest peptide signal, and error is SEM **(D)** 53BP1 TTD Kme2-OPL profile. Two independent binding measurements with 2 technical replicate measures for each Kme2-OPL set were averaged. Data was GST subtracted and normalized as in **Figure 3B**. **(E)** 53BP1 TTD specificity profile.

Kme2-OPL profile belonged to acidic residues in the P-2 position (**Figure 5A**).

Another promiscuous Kme reader is 53BP1. 53BP1 TTD is reported to bind p53K370me2, p53K382me2 (Roy et al., 2010),

H4K20me1/2 (Botuyan et al., 2006; Hartlerode et al., 2012), H3K18me2 (Shanle et al., 2017) and H3K36me2 (Tong et al., 2015) (**Table 1** and **Supplementary Figure S7A**). Our peptide microarray data confirmed the ability of 53BP1 TTD to



recognize these histone Kmes (**Figure 5C**). The sequences surrounding these Kme sites are not conserved (**Supplementary Figure S7B**), signifying 53BP1 TTD is a non-specific Kme reader. Kme-OPL pool screening was consistent with prior reports showing 53BP1 TTD preferred Kme2 (**Figure 2A**). The Kme2-OPL profile showed 53BP1 TTD interactions were largely non-specific, binding to Kme2 in almost all sequences, with the exception of acidic residues (**Figures 5D,E** and **Supplementary Figures S7C,D**).

From our collective analyses of Kme readers, it was apparent that sequence-specific Kme readers had lower Kme-OPL pool

signals than non-specific readers (**Figure 2**). The Kme3-OPL pool signal average for CDYL2, the most sequence-specific reader in our screen, was only ~400 RFU. The low signal was explained by CDYL2 chromo tolerating few residues (**Figure 4B**), leaving CDYL2 only able to bind to a small number of peptides in each pool. The opposite was shown for L3MBTL3 3xMBT, the most sequence promiscuous reader. L3MBTL3 3xMBT had a high signal, ~1,600 RFU, for the Kme2-OPL pool. L3MBTL3 bound to di-methylated peptides regardless of amino acid sequence (**Figure 5B**). Consequently, L3MBTL3 was able to bind to most peptides in the Kme2 pool, resulting in a high signal.

Kme-OPL Revealed Few Sequence Determinants for Histone Kme-Specific Antibody Target Recognition

We wondered what our assay would report as the sequence determinants of histone Kme-specific antibodies, as antibodies are not generally characterized in an unbiased way with regard to the sequences they are presented. We assayed five antibodies: Millipore #07-449, anti-H3K27me3; Cell Signaling Technologies #9733, anti-H3K27me3; Active Motif #39161, anti-H3K9me3; Abcam #ab8895, anti-H3K4me1; and ABClonal #A2355, anti-H3K4me1. First, we used histone peptide microarrays to establish whether an antibody was specific for its target histone Kme (**Figure 6A**). Next, we determined methyl order specificity of each antibody by proxy of the Kme-OPL pools (**Figure 6B**). Finally, we queried sequence preferences by assaying each antibody on Kme-OPL sets of the preferred methyl order (**Figure 6C** and **Supplementary Figure S8**). Each antibody was specific for its intended methyl order. Surprisingly, antibodies were sequence tolerant (**Figure 6D**), and the determinants of antibody sequence specificity were dictated by, at most, two amino acid positions (**Figure 6C**). Millipore #07-449 lacked target specificity on both peptide microarray and Kme3-OPL sets. Cell Signaling Technologies #9733 was the most specific antibody on microarray, but Kme3-OPL profiling showed the antibody's recognition of H3K27 likely came from selectivity in only two positions, P-2 A and P+3 P. Active Motif #39161 was not specific on microarray and did not have a specific Kme3-OPL profile. The two H3K4me1 antibodies, Abcam #ab8895 and ABClonal #A2355, had similar microarray and Kme1-OPL readouts. Both antibodies recognized H3K4me1 on microarray and were sequence selective in one position in the Kme1-OPL. #ab8895 was specific for P+2 S/T, and #A2355 was specific for P-1 S/T. Both antibodies slightly preferred P+1 Q, and the combination of P+1 Q and P+2 S/T or P-1 S/T is unique to H3K4, the intended antibody target. The most specific antibodies, #9733, #ab8895 and #A2355, recognized their target Kme by only two selective positions.

LoB scoring of histone antibody Kme-OPL profiles revealed that, from a strictly sequence selectivity perspective, these antibodies were unlikely to recognize their intended target. For example, in the LoB scoring of ab8895 anti-H3K4me1, 364 lysine-centered seven-mer motifs in the proteome had a LoB score equal to or higher than that of H3K4. ab8895 is the most cited H3K4me1 antibody (Shah et al., 2018). Undoubtedly, ab8895 recognizes H3K4me1, but we cannot explain this solely by high-affinity interactions, as we predict 364 sequences to bind equally or better than H3K4. Likely, the community does not robustly detect off target proteins in techniques like western blot (He et al., 2019) or ChIP (Mohaghegh et al., 2019) because histones are so much more abundant, as reported by Wisniewski et al. (2014). ab8895 anti-H3K4me1 LoB scores ranged from 0 to 801.5. The highest ab8895 LoB scoring protein present in Wisniewski's mass spectrometry data was Sideroflexin-4, which had a LoB score of 784.25. Comparatively, H3K4 had a LoB score of 738.5. Averaging three measurements and reporting the standard deviation, the

copy number of Sideroflexin-4 was $29 \pm 6.5 \times 10^3$ particles/cell, and the copy number of H3.1 was $33 \pm 4.3 \times 10^6$ particles/cell; H3.1 was 1,100 times more abundant than Sideroflexin-4. Since sequences with higher LoB scores exhibited higher binding affinity (**Figure 3F**), Sideroflexin-4 likely binds to ab8895 with higher or at least the same affinity as H3K4. However, binding affinity does not equate to antibody recognition when amounts of targets are so different. High abundance of histones has worked in our favor in the chromatin field. Conversely, studying methylated proteins that are not histones using antibodies will be challenging. Even if an antibody is specific for a protein, the abundance of histones or other competing proteins may obscure detection.

DISCUSSION

Here we report on the development of a Kme-OPL platform that is able to capture the optimal binding sequence P-3/+3 for Kme reader domains and Kme antibodies. Kme-OPL profiles were validated with structural and quantitative binding data, corroborating that Kme-OPL profiles were accurate and could be used to predict optimal binding sequences. Additionally, LoB scoring utilized the surrounding amino acid sequence information to relate Kme-OPL findings to the proteome, toward the goal of identifying potential protein-protein interactions. In the future, this assay can be used for *de novo* characterization of putative Kme readers with no known activity and in drug discovery pipelines, for the identification of high affinity ligands for screening assays, and for the design of peptidomimetic Kme reader antagonists.

The Kme-OPL approach for identifying Kme driven interactions is complimentary to other available tools such as mass spectrometry and SPOT array. When a Kme reader is pulled down from cells and analyzed by mass spectrometry, the nature of identified interactions is unknown. LoB scoring of a Kme reader can be used to prioritize potential direct binding partners. This assay can also inform on optimal protein binding sequences and can be used in SPOT array construction, which relies on prior knowledge of protein binding motifs.

The Kme-OPL reader assay has a few limitations. One limitation of the assay is the use of Kme-OPL pools. Currently, we use the Kme-OPL pools to determine a preferred methyl order for a Kme reader, and the Kme-OPL pool with the highest signal dictates the methyl order used for measurements with Kme-OPL sets. A protein that selectively binds to only few residues may give no detectable signal in Kme-OPL pool screening and would not be continued in our current workflow. The Kme-OPL pool step therefore may report false negatives for highly sequence-specific Kme readers. Another limitation of this assay is the peptide library being constructed with residues P-3/+3 of the Kme. Although the best MPP8 chromo peptide had a K_d of $0.84 \pm 0.067 \mu\text{M}$, making it ten-fold tighter than the H3K9me3 peptide of the same length (**Figure 3C**), the H3K9me3 (1-20) peptide had a K_d of $0.23 \pm 0.03 \mu\text{M}$ (Rothbart et al., 2012). For the CDYL2 inhibitor, specific amino acids in the P-4 position

also increased binding affinity (Barnash et al., 2016). Conversely, residues P+2 and P+3 were not informative for CDYL2 binding and were not included in inhibitor design. Here, we inform on residues most likely to interact with a Kme reader, but P-3/+3 of the Kme will not be the binding footprint for every protein or antibody.

DATA AVAILABILITY STATEMENT

Datasets generated for this study can be found at GitHub (https://github.com/ariana-kupai/LoB_scores) and the Histone Antibody Specificity Database (www.histoneantibodies.com).

AUTHOR CONTRIBUTIONS

AK designed and performed experiments and analyzed data. RV designed experiments and analyzed data. BD developed code, designed experiments, and analyzed data. SR designed and performed experiments and analyzed data. All authors contributed to the writing of the manuscript.

REFERENCES

- Alix, J.-H., Hayes, D., Lontie, J.-F., Colson, C., Glatigny, A., and Lederer, F. (1979). Methylated amino acids in ribosomal proteins from *Escherichia coli* treated with ethionine and from a mutant lacking methylation of protein L11. *Biochimie* 61, 671–679. doi: 10.1016/s0300-9084(79)80165-0
- Ambler, R. P., and Rees, M. W. (1959). ϵ -N-Methyl-lysine in Bacterial Flagellar Protein. *Nature* 184, 56–57. doi: 10.1038/184056b0
- Bannister, A., Zegerman, P., Allshire, R., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1chromo domain. *Nature* 410, 120–124. doi: 10.1038/35065138
- Barnash, K. D., Lamb, K. N., Stuckey, J. I., Norris, J. L., Cholensky, S. H., Kireev, D. B., et al. (2016). Chromodomain Ligand Optimization via Target-Class Directed Combinatorial Repurposing. *ACS Chem. Biol.* 11, 2475–2483. doi: 10.1021/acscchembio.6b00415
- Botuyan, M. V., Lee, J., Ward, I. M., Kim, J.-E., Thompson, J. R., Chen, J., et al. (2006). Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA Repair. *Cell* 127, 1361–1373. doi: 10.1016/j.cell.2006.10.043
- Cai, L., Rothbart, S. B., Lu, R., Xu, B., Chen, W. Y., Tripathy, A., et al. (2013). An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. *Mol. Cell* 49, 571–582. doi: 10.1016/j.molcel.2012.11.026
- Chang, Y., Sun, L., Kokura, K., Horton, J. R., Fukuda, M., Espejo, A., et al. (2011). MPP8 mediates the interactions between DNA methyltransferase Dnmt3a and H3K9 methyltransferase GLP/G9a. *Nat. Commun.* 2:533. doi: 10.1038/ncomms1549
- Cornett, E. M., Dickson, B. M., Krajewski, K., Spellmon, N., Umstead, A., Vaughan, R. M., et al. (2018). A functional proteomics platform to reveal the sequence determinants of lysine methyltransferase substrate selectivity. *Sci. Adv.* 4:eav2623. doi: 10.1126/sciadv.aav2623
- Cornett, E. M., Dickson, B. M., and Rothbart, S. B. (2017). Analysis of histone antibody specificity with peptide microarrays. *J. Vis. Exp.* 126:55912. doi: 10.3791/55912
- Cornett, E. M., Ferry, L., Defossez, P. A., and Rothbart, S. B. (2019). Lysine Methylation Regulators Moonlighting outside the Epigenome. *Mol. Cell* 75, 1092–1101. doi: 10.1016/j.molcel.2019.08.026
- Creixell, P., Palmeri, A., Miller, C. J., Lou, H. J., Santini, C. C., Nielsen, M., et al. (2015). Unmasking determinants of specificity in the human kinome. *Cell* 163, 187–201. doi: 10.1016/j.cell.2015.08.057

FUNDING

We acknowledge support from the National Institutes of Health to SR (R35GM124736) and RV (F99CA245821). AK is supported by an institutional fellowship from Van Andel Institute Graduate School.

ACKNOWLEDGMENTS

We thank colleagues for sharing GST-fusion plasmids (see **Supplementary Table S1**). We thank Evan Cornett for his contributions to the establishment of K-OPL screening technology in our laboratory. We also thank Xiaobing Shi and members of the Rothbart lab for their insightful comments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2020.00241/full#supplementary-material>

- Cui, G., Park, S., Badeaux, A. I., Kim, D., Lee, J., Thompson, J. R., et al. (2012). PHF20 is an effector protein of p53 double lysine methylation that stabilizes and activates p53. *Nat. Struct. Mol. Biol.* 19, 916–924. doi: 10.1038/nsmb.2353
- Dickson, B. M., Cornett, E. M., Ramjan, Z., and Rothbart, S. B. (2016). ArrayNinja: an open source platform for unified planning and analysis of microarray experiments. *Methods Enzym.* 574, 53–77. doi: 10.1016/bs.mie.2016.02.002
- Escamilla-Del-Arenal, M., da Rocha, S. T., Spruijt, C. G., Masui, O., Renaud, O., Smits, A. H., et al. (2013). Cdy1, a new partner of the inactive X chromosome and potential reader of H3K27me3 and H3K9me2. *Mol. Cell Biol.* 33, 5005–5020. doi: 10.1128/MCB.00866-13
- Ferry, L., Fournier, A., Tsusaka, T., Adelmant, G., Shimazu, T., Matano, S., et al. (2017). Methylation of DNA Ligase 1 by G9a/GLP recruits UHRF1 to replicating DNA and regulates DNA methylation. *Mol. Cell* 67:550–565 e5. doi: 10.1016/j.molcel.2017.07.012
- Fischle, W., Franz, H., Jacobs, S. A., Allis, C. D., and Khorasanizadeh, S. (2008). Specificity of the chromodomain Y chromosome family of chromodomains for lysine-methylated ARK(S/T) motifs. *J. Biol. Chem.* 283, 19626–19635. doi: 10.1074/jbc.M802655200
- Franz, H., Mosch, K., Soeroes, S., Urlaub, H., and Fischle, W. (2009). Multimerization and H3K9me3 binding are required for CDYL1b heterochromatin association. *J. Biol. Chem.* 284, 35049–35059. doi: 10.1074/jbc.M109.052332
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31, 3784–3788. doi: 10.1093/nar/gkg563
- Gatchalian, J., Gallardo, C. M., Shinsky, S. A., Ospina, R. R., Liendo, A. M., Krajewski, K., et al. (2016). Chromatin condensation and recruitment of PHD finger proteins to histone H3K4me3 are mutually exclusive. *Nucleic Acids Res.* 44, 6102–6112. doi: 10.1093/nar/gkw193
- Gayatri, S., Cowles, M. W., Vemulapalli, V., Cheng, D., Sun, Z.-W., and Bedford, M. T. (2016). Using oriented peptide array libraries to evaluate methylarginine-specific antibodies and arginine methyltransferase substrate motifs. *Sci. Rep.* 6:28718. doi: 10.1038/srep28718
- Guo, A., Gu, H., Zhou, J., Mulhern, D., Wang, Y., Lee, K. A., et al. (2014). Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. *Mol. Cell. Proteomics* 13, 372–387. doi: 10.1074/mcp.O113.027870
- Hartlerode, A. J., Guan, Y., Rajendran, A., Ura, K., Schotta, G., Xie, A., et al. (2012). Impact of histone H4 lysine 20 methylation on 53BP1 responses to chromosomal double strand breaks. *PLoS One* 7:e49211. doi: 10.1371/journal.pone.0049211

- He, J., Fu, X., Zhang, M., He, F., Li, W., Abdul, M. M., et al. (2019). Transposable elements are regulated by context-specific patterns of chromatin marks in mouse embryonic stem cells. *Nat. Commun.* 10:34. doi: 10.1038/s41467-018-08006-y
- Hornbeck, P. V., Zhang, B., Murray, B., Kornhauser, J. M., Latham, V., and Skrzypek, E. (2015). PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res.* 43, D512–D520. doi: 10.1093/nar/gku1267
- Houghten, R. A., Pinilla, C., Blondelle, S. E., Appel, J. R., Dooley, C. T., and Cuervo, J. H. (1991). Generation and use of synthetic peptide combinatorial libraries for basic research and drug discovery. *Nature* 354, 84–86. doi: 10.1038/354084a0
- James, L. I., Barsyte-Lovejoy, D., Zhong, N., Krichevsky, L., Korboukh, V. K., Herold, J. M., et al. (2013). Discovery of a chemical probe for the L3MBTL3 methyllysine reader domain. *Nat. Chem. Biol.* 9, 184–191. doi: 10.1038/nchembio.1157
- Kaustov, L., Ouyang, H., Amaya, M., Lemak, A., Nady, N., Duan, S., et al. (2011). Recognition and specificity determinants of the human cbx chromodomains. *J. Biol. Chem.* 286, 521–529. doi: 10.1074/jbc.M110.191411
- Klein, B. J., Wang, X., Cui, G., Yuan, C., Botuyan, M. V., Lin, K., et al. (2016). PHF20 readers link methylation of Histone H3K4 and p53 with H4K16 ACETYLATION. *Cell Rep.* 17, 1158–1170. doi: 10.1016/j.celrep.2016.09.056
- Kokura, K., Sun, L., Bedford, M. T., and Fang, J. (2010). Methyl-H3K9-binding protein MPP8 mediates E-cadherin gene silencing and promotes tumour cell motility and invasion. *EMBO J.* 29, 3673–3687. doi: 10.1038/emboj.2010.239
- Kuo, A. J., Song, J., Cheung, P., Ishibe-Murakami, S., Yamazoe, S., Chen, J. K., et al. (2012). The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. *Nature* 484, 115–119. doi: 10.1038/nature10956
- Li, H., Fischle, W., Wang, W., Duncan, E. M., Liang, L., Murakami-Ishibe, S., et al. (2007). Structural basis for lower lysine methylation state-specific readout by MBT repeats of L3MBTL1 and an engineered PHD finger. *Mol. Cell* 28, 677–691. doi: 10.1016/j.molcel.2007.10.023
- Li, J., Li, Z., Ruan, J., Xu, C., Tong, Y., Pan, P. W., et al. (2011). Structural Basis for specific binding of human MPP8 Chromodomain to Histone H3 Methylated at Lysine 9. *PLoS One* 6:e0025104. doi: 10.1371/journal.pone.0025104
- Liu, H., Galka, M., Mori, E., Liu, X., Lin, Y. F., Wei, R., et al. (2013). A method for systematic mapping of protein lysine methylation identifies functions for HP1beta in DNA damage response. *Mol. Cell.* 50, 723–735. doi: 10.1016/j.molcel.2013.04.025
- Liu, L., Zhen, X. T., Denton, E., Marsden, B. D., and Schapira, M. (2012). ChromoHub: a data hub for navigators of chromatin-mediated signalling. *Bioinformatics* 28, 2205–2206. doi: 10.1093/bioinformatics/bts340
- Min, J., Allali-Hassani, A., Nady, N., Qi, C., Ouyang, H., Liu, Y., et al. (2007). L3MBTL1 recognition of mono- and dimethylated histones. *Nat. Struct. Mol. Biol.* 14, 1229–1230. doi: 10.1038/nsmb1340
- Mohaghegh, N., Bray, D., Keenan, J., Penvose, A., Andrienas, K. K., Ramlall, V., et al. (2019). NextPBM: a platform to study cell-specific transcription factor binding and cooperativity. *Nucleic Acids Res.* 47:e31. doi: 10.1093/nar/gkz020
- Nady, N., Krichevsky, L., Zhong, N., Duan, S., Tempel, W., Amaya, M. F., et al. (2012). Histone recognition by human malignant brain tumor domains. *J. Mol. Biol.* 423, 702–718. doi: 10.1016/j.jmb.2012.08.022
- Ong, S.-E., and Mann, M. (2006). Identifying and quantifying sites of protein methylation by heavy Methyl SILAC. *Curr. Protoc. Protein Sci.* Chapter 14:Unit 14.9.
- Rothbart, S. B., Krajewski, K., Nady, N., Tempel, W., Xue, S., Badeaux, A. I., et al. (2012). Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat. Struct. Mol. Biol.* 19, 1155–1160. doi: 10.1038/nsmb.2391
- Roy, S., Musselman, C. A., Kachirskaja, I., Hayashi, R., Glass, K. C., Nix, J. C., et al. (2010). Structural insight into p53 recognition by the 53BP1 tandem Tudor domain. *J. Mol. Biol.* 398, 489–496. doi: 10.1016/j.jmb.2010.03.024
- Shah, R. N., Grzybowski, A. T., Cornett, E. M., Johnstone, A. L., Dickson, B. M., Boone, B. A., et al. (2018). Examining the roles of H3K4 methylation states with systematically characterized antibodies. *Mol. Cell.* 72:162–177.e7. doi: 10.1016/j.molcel.2018.08.015
- Shanle, E. K., Shinsky, S. A., Bridgers, J. B., Bae, N., Sagum, C., Krajewski, K., et al. (2017). Histone peptide microarray screen of chromo and Tudor domains defines new histone lysine methylation interactions. *Epigenetics Chromatin* 10:12. doi: 10.1186/s13072-017-0117-5
- Simhadri, C., Daze, K. D., Douglas, S. F., Quon, T. T., Dev, A., Gignac, M. C., et al. (2014). Chromodomain antagonists that target the polycomb-group methyllysine reader protein chromobox homolog 7 (CBX7). *J. Med. Chem.* 57, 2874–2883. doi: 10.1021/jm401487x
- Stuckey, J. I., Dickson, B. M., Cheng, N., Liu, Y., Norris, J. L., Cholensky, S. H., et al. (2016). A cellular chemical probe targeting the chromodomains of Polycomb repressive complex 1. *Nat. Chem. Biol.* 12, 180–187. doi: 10.1038/nchembio.2007
- Tong, Q., Cui, G., Botuyan, M. V., Rothbart, S. B., Hayashi, R., Musselman, C. A., et al. (2015). Structural plasticity of methyllysine recognition by the tandem tudor domain of 53BP1. *Structure* 23, 312–321. doi: 10.1016/j.str.2014.11.013
- Trojer, P., Li, G., Sims, R. J. III, Vaquero, A., Kalakonda, N., Boccuni, P., et al. (2007). L3MBTL1, a histone-methylation-dependent chromatin lock. *Cell* 129, 915–928. doi: 10.1016/j.cell.2007.03.048
- Tsukasa, T., Kikuchi, M., Shimazu, T., Suzuki, T., Sohtome, Y., Akakabe, M., et al. (2018). Tri-methylation of ATF7IP by G9a/GLP recruits the chromodomain protein MPP8. *Epigenetics Chromatin* 11:56. doi: 10.1186/s13072-018-0231-z
- Vermeulen, M., Eberl, H. C., Matarese, F., Marks, H., Denissov, S., Butter, F., et al. (2010). Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142, 967–980. doi: 10.1016/j.cell.2010.08.020
- Wisniewski, J. R., Hein, M. Y., Cox, J., and Mann, M. (2014). A “proteomic ruler” for protein copy number and concentration estimation without spike-in standards. *Mol. Cell. Proteomics* 13, 3497–3506. doi: 10.1074/mcp.M113.037309
- Wozniak, G. G., and Strahl, B. D. (2014). Hitting the ‘mark’: interpreting lysine methylation in the context of active transcription. *Biochim. Biophys. Acta BBA* 1839, 1353–1361. doi: 10.1016/j.bbarm.2014.03.002

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Kupai, Vaughan, Dickson and Rothbart. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



H3K4me1 Distribution Predicts Transcription State and Poising at Promoters

Sunhee Bae¹ and Bluma J. Lesch^{1,2*}

¹ Department of Genetics, Yale School of Medicine, New Haven, CT, United States, ² Yale Cancer Center, Yale School of Medicine, New Haven, CT, United States

OPEN ACCESS

Edited by:

Steve Bilodeau,
Laval University, Canada

Reviewed by:

Richard Alan Katz,
Fox Chase Cancer Center,
United States
Naoko Hattori,
National Cancer Center Research
Institute, Japan

*Correspondence:

Bluma Lesch
bluma.lesch@yale.edu

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 21 February 2020

Accepted: 03 April 2020

Published: 05 May 2020

Citation:

Bae S and Lesch BJ (2020)
H3K4me1 Distribution Predicts
Transcription State and Poising at
Promoters.
Front. Cell Dev. Biol. 8:289.
doi: 10.3389/fcell.2020.00289

Monomethylation on lysine 4 of histone H3 (H3K4me1) is commonly associated with distal enhancers, but H3K4me1 is also present at promoter regions proximal to transcription start sites. To assess a possible role for H3K4me1 in dictating gene regulatory states at promoters, we examined H3K4me1 peak density around promoters in human and mouse germ cells using an analytic strategy that allowed us to assess relationships between different epigenetic marks on a promoter-by-promoter basis. We found that H3K4me1 exhibits either a bimodal pattern at active promoters, where it flanks H3K4me3, or a unimodal pattern at poised promoters, where it coincides with both H3K4me3 and H3K27me3. This pattern is correlated with gene expression level, but is more strongly linked to a poised chromatin state, defined by the simultaneous presence of H3K4me3 and H3K27me3, than to transcriptional activity. The pattern is especially prominent in germ cells, but is also present in other cell types, including embryonic stem cells and differentiated somatic cells. We propose that H3K4me1 is a key feature of the poised epigenetic state, and suggest possible roles for this mark in epigenetic memory.

Keywords: histone, bivalent, poised, promoter, stem cell, germ cell, spermatogenesis, pluripotency

INTRODUCTION

Post-translational modifications on histone tails are closely correlated to transcriptional states. For example, trimethylation of histone H3 lysine 4 (H3K4me3) marks active gene promoters, while H3K9me3 marks regions subject to long-term repression (Barski et al., 2007). One such modification is monomethylation on lysine 4 of histone 3 (H3K4me1), a mark that has been linked to enhancers. Identifying regions enriched for H3K4me1 and depleted in H3K4me3, or regions enriched for both H3K4me1 and H3K27ac, have proven to be feasible methods for enhancer discovery (Heintzman et al., 2007; Creyghton et al., 2010). Mechanistic studies on the regulation of H3K4me1 marks have focused on the roles of MLL3 and MLL4, enzymes responsible for placing monomethylation on unmethylated lysine 4 which act primarily at enhancer regions (Guo et al., 2013; Hu et al., 2013; Dorighi et al., 2017).

At the same time, not all H3K4me1-enriched regions correspond to enhancers. H3K4me1 marks also exist at promoters, which implies that the H3K4me1 modification may have a context-dependent role in regulating transcription. Several studies have addressed questions about the role of H3K4me1 at promoters and highlighted the transcriptional features that may distinguish the functions of H3K4me1 at promoters compared to enhancers. H3K4me1 at promoters in the

absence of H3K4me3 was associated with gene repression, while at active gene promoters H3K4me1 appeared to flank H3K4me3 in one study examining skeletal muscle cells (Cheng et al., 2014). Another study of mouse liver and pancreas likewise found that H3K4me1 signal flanks active promoters, although the relationship to H3K4me3 was not examined (Hoffman et al., 2010). An alternative view suggests that promoters and enhancers are a single class of transcriptional elements distinguished by different levels of transcription, and the varying H3K4me3 to H3K4me1 signal ratio at promoters compared to enhancers reflects the rates at which the elements recruit RNA Polymerase II (Core et al., 2014; Andersson et al., 2015).

Cell type is another aspect of regulatory context that may play a role in H3K4me1 function. Different modes of epigenetic regulation can take on larger or smaller roles in certain cell types. For example, embryonic stem cells (ESCs) and germ cells exhibit specialized regulatory mechanisms to maintain pluripotency and reprogramming potential. An epigenetic feature that is especially prominent in ESCs and germ cells is poised chromatin, which is defined by the simultaneous presence of two histone modifications, H3K4me3 and H3K27me3, at transcriptionally repressed promoters (Azuara et al., 2006; Bernstein et al., 2006; Mikkelsen et al., 2007). H3K4me3 at promoter regions is usually associated with active transcription, while H3K27me3 is associated with repression (Barski et al., 2007). The co-occurrence of these two opposing modifications at the same locus is thought to serve regulatory roles in ES cells and germ cells by preventing DNA methylation and preparing for resolution to active or more fully repressed states as cells differentiate (Azuara et al., 2006; Bernstein et al., 2006; Mikkelsen et al., 2007; Lesch and Page, 2014).

Here, we report distinct patterns of H3K4me1 that predict transcriptional regulatory states at promoters in germ cells and ESCs. We used an alternative approach to ChIP-seq data analysis in which we summarized chromatin signal around promoters based on ChIP-seq peak density instead of signal density. This approach allowed us to quantitatively interrogate histone modification patterns at a promoter-by-promoter level, in contrast to more commonly-used approaches that require pooling of multiple promoters to obtain average ChIP signals for a given mark. We used our approach to ask if patterns of H3K4me1 deposition near promoters could convey information about the regulatory state of the promoter. We examined ChIP-seq data for H3K4me1, H3K4me3, H3K27me3, and H3K27ac in mouse and human male germ cells, and found that H3K4me1 peak density around the transcription start sites (TSS) exhibits either a broad bimodal profile or a narrower unimodal profile centered at the TSS. We then examined the position of the H3K4me1 marks relative to H3K4me3, and found that unimodal H3K4me1 directly at the TSS predicts a poised state of chromatin, while bimodal H3K4me1 flanking the TSS predicts an active state. The bimodal distribution is not explained by nucleosome clearing, and the unimodal distribution is not explained only by low expression at these promoters. We conclude that unimodal H3K4me1 centered on the TSS is a characteristic feature of the poised epigenetic state in ESCs and germ cells.

MATERIALS AND METHODS

Human Subjects

These studies were carried out in accordance with the Declaration of Helsinki. The protocol was approved by Yale University's Human Subjects Institutional Review Board. Written informed consent was obtained from all subjects.

Mouse Experiments

This study was carried out in accordance with the principles of the Basel Declaration and recommendations of Yale University's Institutional Animal Care and Use Committee (IACUC). All procedures involving mice were approved by the Yale IACUC.

Sample Collection and Sorting

Human testis samples were obtained from adult male patients undergoing vasectomy reversals at the Infertility Clinic of St. Louis. All men whose tissue was used in this study had a prior history of fertility demonstrated by at least one living child. Epididymal sperm quality and abundance proximal to the vasectomy site was assessed at the time of biopsy, and abundant, motile, morphologically normal sperm were confirmed for each patient. Testis biopsy samples were minced, dissociated using collagenase and trypsin, and then filtered to obtain a single-cell suspension as described (Bellvé, 1993). Mouse testes were isolated from adult CD1 males (Charles River Laboratories), and tissue from several mice was pooled before cell separation. Pachytene spermatocyte and round spermatid fractions were collected by StaPut (Shepherd et al., 1981; Bellvé, 1993; Liu et al., 2015), and pooled fractions were counted on a hemocytometer. Purity was >95% for each human sample and >90% for each mouse sample, as assessed by counts of 100 cells from each fraction under phase optics. Cells were washed once in PBS, fixed in 1% formaldehyde for 8 min at room temperature, and then quenched with 2.5 M glycine for 5 min at room temperature. Fixed cells were snap frozen in liquid nitrogen, then stored at -80°C .

Chromatin Immunoprecipitation and Library Preparation

H3K4me3 and H3K27me3 data was previously published and is available on GEO (GSE68507). For H3K4me1 and H3K27ac ChIP, between 5×10^4 and 5×10^6 cells were used as starting material, depending on the number obtained from sample isolation and sorting. Pachytene spermatocytes and round spermatids were treated identically. For human spermatogenic cells, fixed cells frozen in lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl [pH 8]) were thawed on ice. For mouse spermatogenic cells, fixed cells frozen in PBS were thawed on ice, then washed once in cold PBS and resuspended in 100 μl lysis buffer. Once in lysis buffer, cells were incubated on ice for 5 min. 200 μl ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl [pH 8], 167 mM NaCl) was then added to each sample. Samples were sonicated in aliquots of 150 μl in 0.5 ml Eppendorf tubes at 4°C using a BioRuptor (Diagenode) for 35 cycles on High setting, 30 s on/30 s off. Aliquots of the same sample were then re-pooled and spun

down at $12,000 \times g$ for 5 min, and the supernatant moved to a fresh tube. Chromatin from each sample was then split into two separate tubes (150 μ l in each), and 700 μ l dilution buffer, 50 μ l lysis buffer, and 100 proteinase inhibitor cocktail (Complete Mini tablets, Roche #11836153001) were added to each tube. 50 μ l of each sample was set aside as input. The remainder of the ChIP was performed as previously described (Lesch et al., 2013), except that the second wash for H3K27ac samples was performed in high-salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl (pH 8.1), 500 mM NaCl) instead of low-salt immune complex wash buffer. Immunoprecipitation was performed using 1.0 μ g of antibody to H3K4me1 (Abcam #ab8895, RRID:AB_306847) or 1.0 μ g of antibody to H3K27ac (Abcam #ab4729, RRID:AB_2118291).

Sequencing Library Preparation and Sequencing

ChIP libraries were prepared using a TruSeq ChIP sample prep kit (Illumina), according to the manufacturer's instructions, except that size selection was performed after (instead of before) PCR amplification. All libraries were sequenced on an Illumina HiSeq2500 with 40-base-pair single-end reads.

ChIP-seq Data Processing

Image analysis and base calling were done with the standard Illumina pipeline for HiSeq2500. Data was quality-filtered using `fastq_quality_filter` from the FASTX toolkit (RRID:SCR_005534) with the following parameters: `-q 20 -p 80`. ChIP-seq data was aligned to either the mouse (mm10) or human (hg19) genome using Bowtie2 in `--end-to-end --fast` mode with default settings (Langmead and Salzberg, 2012, RRID:SCR_005476). Peaks were called using MACS2 with the following parameters: `narrowPeak, q = 0.1` (H3K4me3); `broadPeak, q = 0.2` (H3K4me1 from mouse PS and RS); `broadPeak, q = 0.1` (all other data) (Zhang et al., 2008, RRID:SCR_013291). *q*-values for peak calling were selected based on reconciliation of peak boundaries with ChIP signal visualized on the UCSC genome browser.

RNA-seq Data Processing

RNA-seq data was processed using kallisto (Bray et al., 2016, RRID:SCR_016582) with the following parameters: `--bias -b 40 -t 8`. Ensembl transcripts (cDNA and ncDNA) from mm10 or hg19 were used to create reference indexes for each species (Cunningham et al., 2019).

Data Analysis

Density plots were generated using custom R scripts. First, the center of each peak ascertained by MACS2 was obtained and the distance from the center of the peak to the nearest TSS was calculated. The distance values were used as input to calculate a density distribution using a Gaussian smoothing function with bandwidth of 15, and the resulting probability function was plotted in R. Heatmaps were generated using the `ggplot2` package in R (Wickham, 2011, RRID:SCR_014601) and custom R scripts. For each TSS, its distance to the nearest H3K4me1 and H3K4me3 was calculated, and the two distance values per TSS

were used to generate a heatmap based on 2d bin counts using `geom_bin2d()` in `ggplot2`.

To obtain a list of poised promoter regions, a list of H3K4me3 peaks that overlap H3K27me3 peaks were generated with BEDTools using `-wa` option (Quinlan and Hall, 2010, RRID:SCR_006646). To obtain H3K4me1 peaks at poised promoters, a list of H3K4me1 peaks with overlaps with poised H3K4me3 was generated again with `-wa` option on BEDTools.

Violin plots were generated using `ggplot2` and custom R scripts. TPM values obtained through kallisto were used to categorize the TSS based on expression level, using the following categories: $TPM \leq 1$, $1 < TPM \leq 5$, $5 < TPM \leq 10$, $TPM > 10$. For each TSS, a category based on TPM value was assigned, and its distance to the center of the nearest H3K4me1 peak was calculated. The distances in each TPM category were used to create violin plots using `geom_violin()` and `geom_boxplot()` in `ggplot2`.

Statistical comparisons of multi-group, non-parametric data were performed using the Kruskal-Wallis rank sum test with the `kruskal.test()` function in R.

Data Availability

All germ cell datasets are available on GEO under accession numbers GSE68507 (H3K4me3 ChIP-seq, H3K27me3 ChIP-seq, and RNA-seq) and GSE145225 (H3K4me1 and H3K27ac ChIP-seq). Other public datasets used in this study are from mouse myeloid cells (GSE85072), fetal liver (GSE119201), and adult kidney (GSE31039).

Code Availability

Custom R and Python scripts used for this study are available on Github¹.

RESULTS

Distribution of H3K4me1 Peaks Near Promoters Differs From Other Histone Marks

We set out to examine H3K4me1 distribution in germ cells and embryonic stem cells (ESCs), cell types that share many regulatory mechanisms related to reprogramming and pluripotency. To evaluate the distribution of H3K4me1 at promoters relative to other histone modifications, we called peaks from ChIP-seq data for H3K4me1 and H3K27ac (this study), and from H3K4me3 and H3K27me3 (Lesch et al., 2016) from mouse and human spermatogenic cells at two developmental stages, pachytene spermatocytes and round spermatids. Pachytene spermatocytes are undergoing the first meiotic division, while round spermatids are haploid germ cells that have completed meiosis but have not yet differentiated into sperm. These two cell types represent two very different cellular states, but share gene regulatory features characteristic of germ and stem cells (Lesch et al., 2013; Hammoud et al., 2014). We calculated the base pair

¹<https://github.com/Lesch-Lab/H3K4me1-profiles>

unit distance from each peak to the nearest TSS and visualized the distribution of peak distances with a density plot, where the *y* axis shows the probability that the nearest peak is centered a given distance away from the TSS (**Figure 1A**). H3K4me3, H3K27me3 and H3K27ac peak distances were all centered within 1kb of the TSS, and the distributions of the three marks were unimodal, as expected for these promoter-centric modifications. In contrast, H3K4me1 exhibited a mixed unimodal and bimodal pattern in which a large fraction of peaks were centered 300–1000 bp away from the nearest promoter region. The bimodal pattern was unique to H3K4me1 and was present across all cell types examined, although it appeared less pronounced in human compared to mouse cells (**Figure 1B**).

We considered several possible explanations for the unique distribution of H3K4me1 peaks around promoters. First, most H3K4me1 signal might be coming from enhancers, as has been previously described. If so, peak distribution would be centered away from the TSS, but we would expect to see a broad distribution of H3K4me1 peaks extending away from the TSS in both directions, rather than the accumulation in density we observe in the 300–1000 bp range. Second, the bimodal profile could result from nucleosome clearing at active promoters, similar to the pattern commonly seen in metagene plots of H3K4me3 signal (Barski et al., 2007). This explanation has been previously proposed in the context of bimodal H3K4me1 profiles (Hoffman et al., 2010). However, our method does not detect nucleosome clearing for H3K4me3 at active TSS, where this phenomenon is known to occur (**Figure 1C**), indicating that our method of analysis does not reveal nucleosome-free regions. Finally, the unimodal distributions we observed for other marks indicate that a bimodal pattern is not simply an artifact of our analysis method. We conclude that H3K4me1 occupies a distinctive distribution at promoters in addition to its well-studied role at distal enhancers (Rickels et al., 2017; Local et al., 2018).

The Bimodal Distribution Is Characteristic of H3K4me1 Across Cell Types

To determine whether the pattern of H3K4me1 distribution around the TSS was unique to germ cells, we applied the same data-processing protocol to published ChIP-seq data from mouse embryonic stem cells (mESCs) and human embryonic stem cells (hESCs) (Rada-Iglesias et al., 2011; Garland et al., 2019). The mixed unimodal/bimodal pattern of H3K4me1 density around the TSS was recapitulated in mESCs, and to a lesser degree in hESCs (**Figure 2A**). We then examined somatic cells across several lineages using public datasets, and found that similar mixed unimodal/bimodal H3K4me1 patterns are present in mouse myeloid cells, fetal liver, and adult kidney cells (**Figure 2B**; Encode Project Consortium, 2012; Kotzin et al., 2016). Whereas both the unimodal and bimodal H3K4me1 patterns in fetal liver closely recapitulated the pattern observed in germ cells, the unimodal TSS probability density was dampened relative to the bimodal density in myeloid and kidney cells. We postulate that the bimodal H3K4me1 pattern is shared among all cell types,

while the unimodal pattern is more accentuated in cell types that have pluripotent or multipotent potential.

H3K4me1 Distribution Pattern Correlates to Expression Level

We next asked whether the H3K4me1 distribution around the TSS was related to gene expression level. For each TSS in the mouse genome, we found the distance to the nearest H3K4me1 peak, then grouped the TSS into four categories based on expression level as measured by RNA-seq (Lesch et al., 2016) (see section “Materials and Methods”). We observed a consistent trend in which H3K4me1 peaks overlap the TSS with a unimodal distribution at genes with lower expression levels, and flank the TSS with a bimodal distribution at genes with higher expression (**Figure 3**). The difference in distribution among expression levels was statistically significant for all cell types tested in mouse and for human pachytene spermatocytes ($p < 0.05$, Kruskal-Wallis test). The same trend is present in human round spermatids and hESCs, although the distribution differences were not statistically significant. Lack of statistical significance in the latter two cell types may be due to increased noise in the datasets, although we cannot exclude a true biological difference in histone mark distributions.

H3K4me1 Distribution Defines Two Categories of H3K4me3-Positive Promoters

Because H3K4me1 is an intermediate molecular state between unmethylated H3K4 and H3K4me3, we considered the possibility that H3K4me1 distribution at promoters is purely a byproduct of mechanisms that directly regulate H3K4me3. In this scenario, H3K4me1 would be expected to mark genes in transition between expression states: either gain of H3K4me3 along with gain of transcription, or loss of H3K4me3 along with loss of transcription.

We therefore examined how the H3K4me1 pattern relates to the H3K4me3 pattern around promoter regions. We calculated the distance to both the nearest H3K4me1 peak and the nearest H3K4me3 peak for each TSS (**Figure 4A**). We visualized the distribution of the TSS with respect to surrounding H3K4me1 and H3K4me3 marks as a heatmap, where the total number of data points represented in the heatmap is equal to the number of TSS that neighbor a H3K4me1 and a H3K4me3 region within 2000bp. A clear pattern emerged from this analysis, in which the TSS either grouped at the center of the heatmap or into “wings” on both sides of the center cluster (**Figure 4B**). The TSS occupying the center of the plot represent promoters with both H3K4me1 and H3K4me3 centered on the TSS, while the group of TSS that form the “wings” correspond to promoters that have only H3K4me3 directly at the TSS and are flanked by H3K4me1. This finding indicates that TSS marked by H3K4me3 can be classified into two groups based on the distribution of nearby H3K4me1 signal, and implies that H3K4me1 distribution is not purely dictated by the distribution of H3K4me3 at a given promoter.

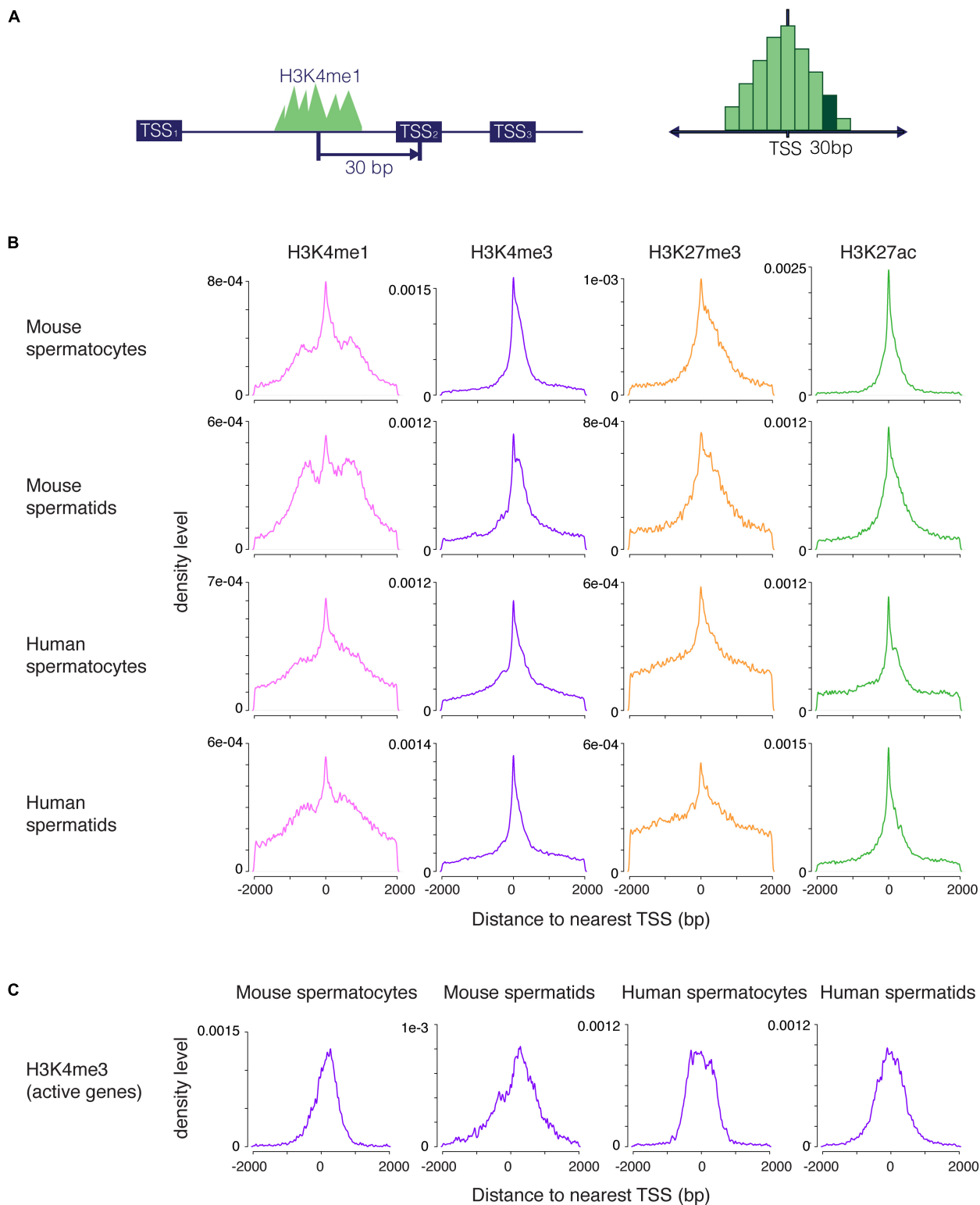


FIGURE 1 | Distribution of H3K4me1, H3K4me3, H3K27me3, and H3K27ac peaks near promoters. **(A)** Scheme for generating density plots. For each peak, the distance from the peak center to its nearest transcription start site (TSS) was obtained, and a density distribution was calculated from the distance values. **(B)** Density plots for four histone modifications in mouse and human male germ cells at two stages of spermatogenesis. All four marks have a unimodal set of peaks centered at the TSS, but only H3K4me1 has an additional bimodal peak density displaced from the TSS. **(C)** Density distribution of H3K4me3 peaks at highly transcribed (tpm > 10) TSS. The density plots do not exhibit profiles that indicate nucleosome clearing. bp, base pairs.

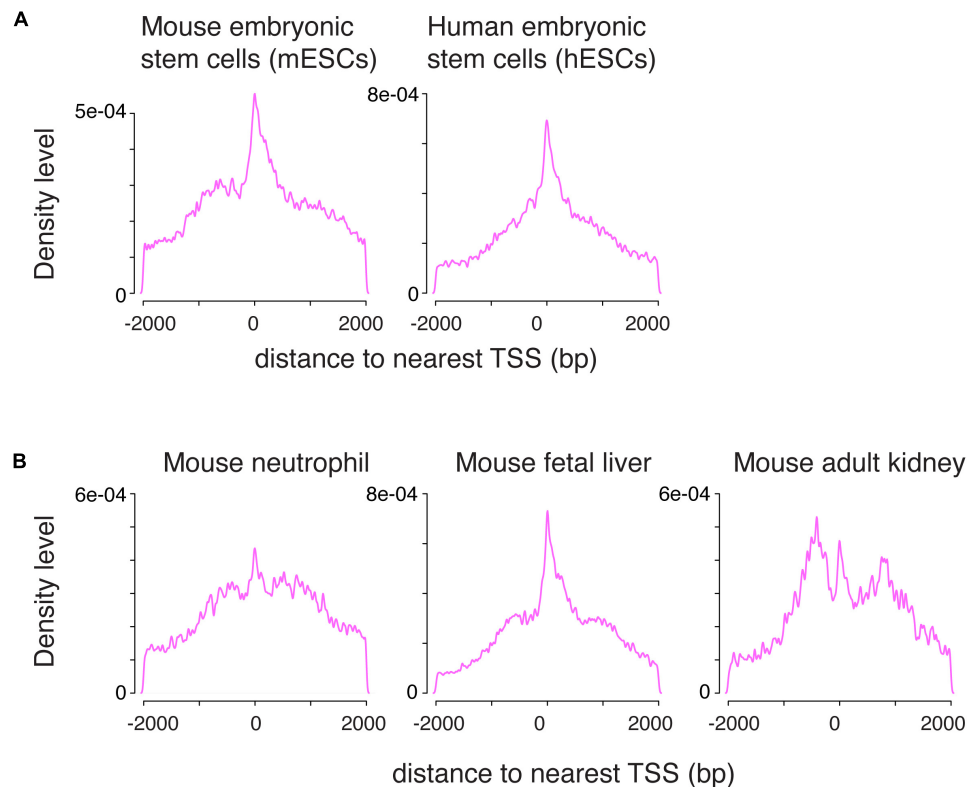


FIGURE 2 | The combination of unimodal and bimodal H3K4me1 distributions is observed across various cell types. **(A)** Density distribution in mouse and human embryonic stem cells. A mixed unimodal and bimodal pattern of H3K4me1 is present, similar to germ cells. **(B)** Density distribution in mouse neutrophils, fetal liver, and adult kidney. The mix of unimodal and bimodal distributions is observed in all three cell types, but relative prominence of the unimodal H3K4me1 distribution varies. bp, base pairs.

H3K4me1 Pattern Separates Poised From Active Promoters

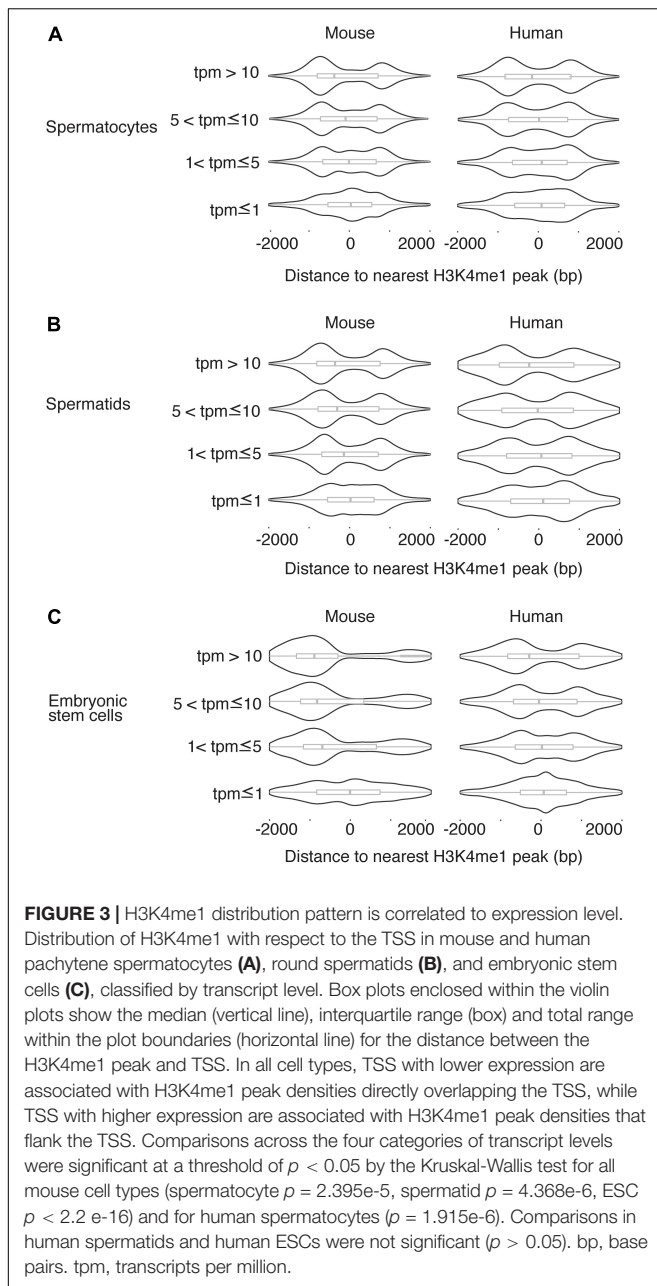
Based on our finding that transcription is related to the shape of the H3K4me1 distribution, we hypothesized that the two patterns observed in our joint analysis might represent poised and active promoters. Active promoters are H3K4me3 positive, H3K27me3 negative, and transcriptionally active, while poised promoters are marked by both H3K4me3 and H3K27me3 and are transcriptionally repressed. We sorted the TSS into “poised” and “active” groups based on whether or not the H3K4me3 region nearest a given TSS overlapped with H3K27me3 signals. When poised and active TSS were plotted separately, we found that the pattern observed in our first heatmap resolved into separate bimodal and unimodal patterns based on whether the TSS is active or poised, respectively (**Figures 4B–D**). These two patterns are identifiable in the ChIP-seq signal tracks: H3K4me1 is centered at the TSS at poised, H3K4me3/H3K27me3-positive promoters, while at active gene promoters H3K4me1 “clears” from TSS bearing high H3K4me3 signals to flank the H3K4me3 marks (**Figure 4E**). The H3K4me1 signal pattern at active promoters helps to explain the asymmetry we observe in the bimodal distribution of H3K4me1 in mouse ESCs, in which a higher proportion of H3K4me1 regions occur upstream of the TSS (**Figure 3C**). H3K4me3 peaks are often centered downstream

of the TSS at active genes, and when bimodal H3K4me1 regions symmetrically flank a H3K4me3 region at a TSS, the upstream H3K4me1 peak will therefore be positioned closer to the TSS than the downstream peak. Since our method of analysis records the single closest H3K4me1 peak for each TSS, the H3K4me1 peak upstream of the TSS is more likely to be recorded. This asymmetry is particularly pronounced in mouse ESC, possibly due to higher resolution of the sequencing data, but is also present in the H3K4me1 distribution for other cell types.

In all three cell types from mouse, heatmaps of TSS reflected a distribution in which H3K4me1 directly occupied promoter regions of poised genes and separated to border the H3K4me3 signal in active promoters. A similar phenomenon was observed for human germ cells and ESCs. While H3K4me1 marks did not always directly overlap the TSS of poised genes in human round spermatids and ESCs, the marks localized closer to the TSS for poised genes compared to active genes (**Figures 4C,D**).

Poising Is Dominant Over Expression Level in Determining H3K4me1 Peak Profile

The correlations between H3K4me1 distribution, poised chromatin state, and gene expression level raised two possibilities for the role of H3K4me1: H3K4me1 could be functionally related



to expression level and correlated with poised chromatin, or H3K4me1 could be functionally related to the poised histone modification state and correlated with gene repression. To discriminate between the two possibilities, we first obtained a list of H3K4me3 peaks that overlap H3K27me3 peaks (“poised” peaks). Then we calculated the distances between the TSS and the nearest H3K4me1 peak and grouped the TSS based on whether the nearest H3K4me1 peaks overlapped with poised peaks. TSS where the nearest H3K4me1 peak overlapped with a poised peak were classified as poised TSS. We then defined epigenetically ‘active’ TSS as TSS where the nearest H3K4me1 peak did not overlap a poised peak, and the nearest H3K4me3 peak was within 500bp of the TSS (see section “Materials

and Methods”). Then, each of the two categories of TSS were further classified into four groups based on expression levels, using the same thresholds for transcription level that we used in our previous analysis. As expected, poised promoters were more likely to be silent ($\text{tpm} = 0$) compared to epigenetically active genes, and when expressed their transcript levels were significantly lower than expressed epigenetically active genes (Supplementary Figure S1). However, there were enough poised genes with high expression levels and epigenetically active genes with low expression levels to test for statistical significance ($n \geq 167$ for all categories and datasets). We found that the unimodal, promoter-centered H3K4me1 distribution marked poised promoters regardless of expression level, while a bimodal H3K4me1 distribution marked epigenetically ‘active’ (H3K4me3-only) promoters regardless of expression level (Figures 5A–C). We conclude that the promoter-centered, unimodal distribution of H3K4me1 is a feature of the poised state.

DISCUSSION

We identified two distinct H3K4me1 distribution patterns at promoters in mouse and human germ cells and ESCs by examining the density of H3K4me1 peaks around transcription start sites. We found a unimodal distribution of peaks at one set of promoters, corresponding to H3K4me1 occupying the promoter region directly at the TSS. A different set of promoters is associated with a bimodal peak density that corresponds to H3K4me1-enriched regions flanking the TSS. These distributions corresponded to alternative regulatory states. Unimodal H3K4me1 signal was correlated with lower transcript expression and the presence of poised (H3K4me3/H3K27me3 bivalent) chromatin, while the bimodal H3K4me1 pattern was linked to high expression level and the absence of poising. By examining promoters where chromatin state and expression level were discordant, we found that H3K4me1 patterns were more strongly associated with the chromatin state dictated by H3K4me3 and H3K27me3 than with the expression level of the promoter.

The observation that H3K4me1 signal exhibits two different patterns around promoters has been previously reported in mouse muscle cells *in vitro* and in adult mouse pancreatic islet and liver cells, as well as a handful of other cell types (Hoffman et al., 2010; Cheng et al., 2014). We found that these two promoter-associated H3K4me1 patterns are also present in germ cells. In contrast to these previous studies, we found that the bimodal pattern of H3K4me1 is independent of nucleosome clearing, and instead may signify a more direct regulation of the balance between H3K4me1 and H3K4me3. In addition, we associate the unimodal H3K4me1 distribution directly with the bivalent H3K4me3/H3K27me3 histone modification state. In muscle cells, H3K4me1 was found to occupy repressed H3K27me3-positive promoters, and similar findings have been reported for brown adipocytes and hESCs (Cheng et al., 2014; Dozmorov, 2015; Brunmeir et al., 2016). While these studies found that H3K4me1 correlates with H3K27me3 alone and acts as a repressive

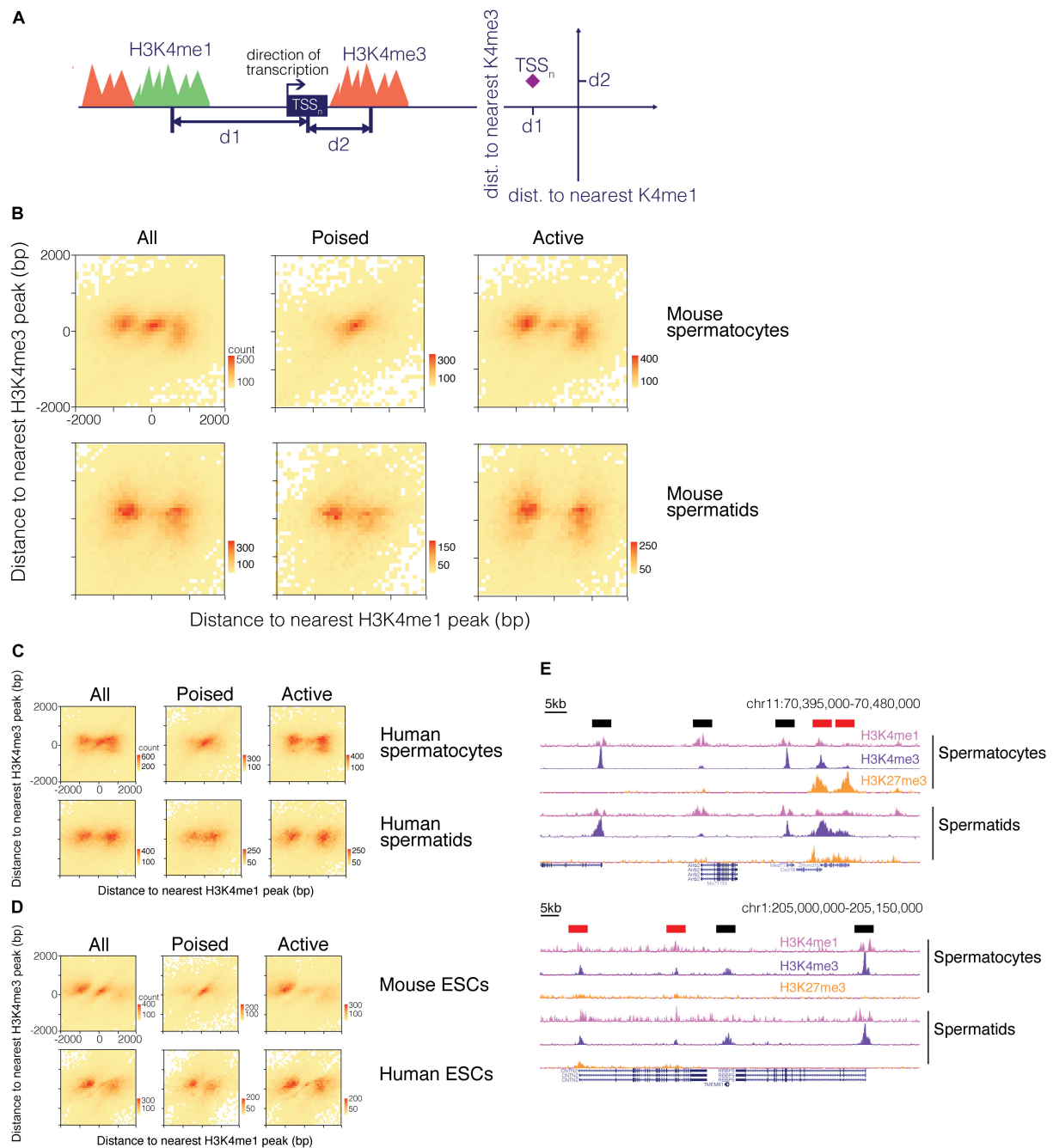


FIGURE 4 | H3K4me1 distribution around H3K4me3 predicts the epigenetic state of promoters. **(A)** Scheme for mapping TSS with respect to both H3K4me1 and H3K4me3 marks. For each TSS, distances to its nearest H3K4me1 peak and its nearest H3K4me3 peak were calculated, and the TSS were classified based on distances to the two histone modifications. **(B)** Heatmap of TSS distribution in mouse pachytene spermatocytes and round spermatids. Most TSS were concentrated at either the center of the heatmap or both sides of the center group to form “wings”. The center cluster and the wings separate when the TSS are grouped based on poising. **(C)** Heatmap of TSS distribution in human germ cells. **(D)** Heatmap of TSS distribution in mouse and human embryonic stem cells. **(E)** ChIP-seq signal tracks at poised and active promoters in mouse (top) and human (bottom) germ cells. H3K4me1 signal at active promoters (black bars above signal tracks) flanks H3K4me3 signal while H3K4me1 signal at poised promoters (red bars above signal tracks) exhibits a unimodal profile. bp, base pairs.

mark, we find that unimodal H3K4me1 patterns correlate strongly with poised promoters marked by both H3K4me3 and H3K27me3. This finding suggests that H3K4me1 is an essential feature of poised chromatin in cells with pluripotent

and reprogramming potential, such as ESCs and germ cells. Alternatively or in addition, H3K4me1 may act as either a “memory” or precursor of poising at transiently-poised promoters in these populations.

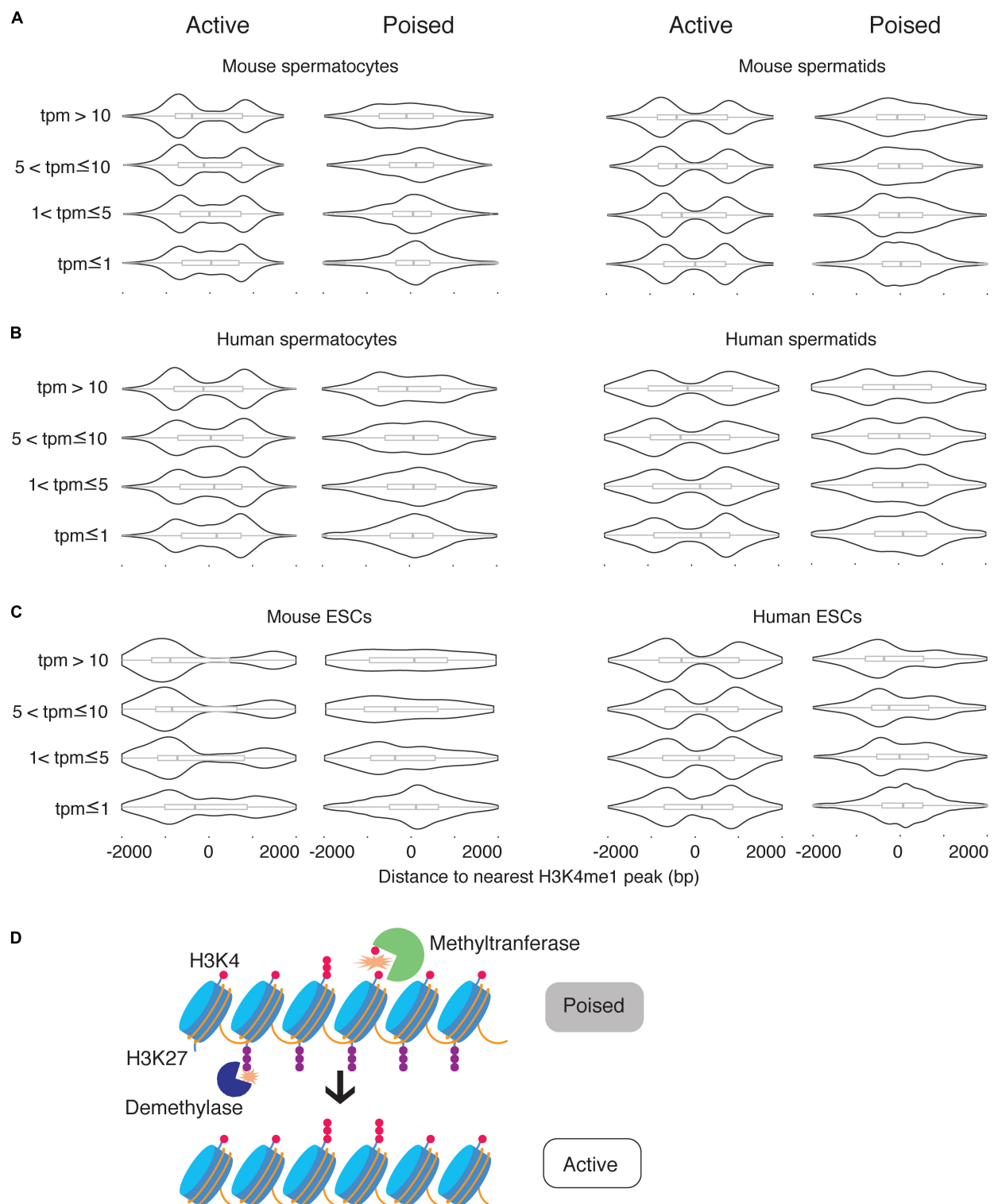


FIGURE 5 | H3K4me1 peak profile correlates best with epigenetic poising at promoters. **(A)** Distribution of H3K4me1 with respect to the TSS at active promoters and poised promoters in mouse spermatocytes and spermatids, classified by expression level. **(B)** Distribution of H3K4me1 in human spermatocytes and spermatids, classified by expression levels. **(C)** Distribution of H3K4me1 in mouse and human ESCs. Box plots enclosed within the violin plots show the median (vertical line), interquartile range (box) and total range within the plot boundaries (horizontal line) for the distance between the H3K4me1 peak and TSS. All comparisons between active promoters and poised promoters were significant ($p < 0.05$) by the Kruskal-Wallis test. **(D)** Model for the role of H3K4 monomethylation in activation of poised promoters. Preexisting H3K4me1 is further methylated to H3K4me3, while H3K27me3 is removed from the promoter region.

The latter possibility has been explored through the study of “Placeholder” nucleosomes, which are characterized by H3K4me1 and H2A.Z in zebrafish sperm. These specialized nucleosomes deter DNA methylation at genomic regions they occupy until the regions acquire poising or activating marks at zygotic genome activation (Murphy et al., 2018). H3K4me1 and H2A.Z serve as a transient chromatin state that passively drives demethylation of DNA without activating transcription while DNA methylation patterns are reprogrammed. We speculate that the H3K4me1 we observe at poised promoters in human and mouse germ cells serves a similar regulatory function: it maintains an epigenetically neutral state at promoters that are important for development and differentiation. Interestingly, H2A.Z is also enriched at poised promoters in mESCs and hESCs (Ku et al., 2012).

Why do pluripotent systems incorporate H3K4me1 at poised promoters? Structure studies have shown that DNMT3L directs the DNA methyltransferases DNMT3a and DNMT3b to nucleosomes bearing unmethylated H3K4, and that monomethylation is sufficient to deter binding of histone-interacting domain of DNMT3L (Ooi et al., 2007). H3K4me1 could serve to broadly demarcate regions of the DNA that must be kept hypomethylated during spermatogenesis and early embryogenesis. H3K4me1 can prevent DNA methylation without recruiting transcription-activating regulators, keeping these regions in a transcriptionally neutral state. In this respect, H3K4me1 is an ideal molecular marker to carry “memories” of transcription. This view of H3K4me1 may also help explain the bimodal H3K4me1 distribution we observed at promoters of active genes. H3K4me1 would appear to flank H3K4me3 peaks at promoter regions if the H3K4me1 marks are replaced with H3K4me3 proximal to the TSS as transcription is activated (**Figure 5D**).

Interpretation of H3K4me1 as transcriptional memory also has interesting implications for the mechanisms underlying epigenetic inheritance. When DNA is replicated, parent histones are thought to be distributed to one of the daughter strands, which results in a “dilution” of histone modifications (Alabert et al., 2015). At the same time, H3K4me3 regions must be kept narrowly focused at the TSS: spreading of H3K4me3 is associated with defects in transcriptional activation (Kidder et al., 2014). Narrowly H3K4-trimethylated regions may therefore be necessary for accurate control of transcription, but also endanger inheritance of transcriptional memory. Broad H3K4me1 regions could overcome this dilemma: a broad stretch of H3K4me1 is more likely to be “remembered” by both strands of the replicated DNA and would also maintain a chromatin state that is permissive for recruitment of the transcriptional machinery. While additional experiments

need to be done to validate the role of H3K4me1 in regulating H3K4me3 deposition, our findings provide a new perspective in interpreting the role and regulation of H3K4me1 marks at promoters.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the NCBI Gene Expression Omnibus (GEO) under accession number GSE145225.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Yale University Human Subjects Institutional Review Board. The patients/participants provided their written informed consent to participate in this study. The animal study was reviewed and approved by the Yale University Institutional Animal Care and Use Committee.

AUTHOR CONTRIBUTIONS

SB wrote the code, designed and performed the analysis, interpreted the data, and wrote the manuscript. BL performed the ChIP experiments, contributed to data interpretation, and contributed to writing the manuscript.

FUNDING

This work was supported by a Burroughs Wellcome Career Award for Medical Scientists, a Francis G. Kingsley Memorial Fellowship, and a Searle Scholar Award to BL.

ACKNOWLEDGMENTS

We thank J. McCarrey and S. Silber for contribution of mouse and human samples, respectively, and J. McCarrey for performing cell isolations. We are grateful to members of the Lesch laboratory for helpful comments and discussions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2020.00289/full#supplementary-material>

REFERENCES

- Alabert, C., Barth, T. K., Reverón-Gómez, N., Sidoli, S., Schmidt, A., Jensen, O. N., et al. (2015). Two distinct modes for propagation of histone PTMs across the cell cycle. *Genes Dev.* 29, 585–590. doi: 10.1101/gad.256354.114
- Andersson, R., Sandelin, A., and Danko, C. G. (2015). A unified architecture of transcriptional regulatory elements. *Trends Genet.* 31, 426–433. doi: 10.1016/j.tig.2015.05.007
- Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jørgensen, H. F., John, R. M., et al. (2006). Chromatin signatures of pluripotent cell lines. *Nat. Cell Biol.* 8, 532–538. doi: 10.1038/ncb1403

- Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D. E., Wang, Z., et al. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823–837. doi: 10.1016/j.cell.2007.05.009
- Bellvé, A. R. (1993). Purification, culture, and fractionation of spermatogenic cells. *Methods Enzymol.* 225, 84–113. doi: 10.1016/0076-6879(93)25009-q
- Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., et al. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315–326. doi: 10.1016/j.cell.2006.02.041
- Bray, N. L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Erratum: near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* 34:888. doi: 10.1038/nbt0816-888d
- Brunmeir, R., Wu, J., Peng, X., Kim, S.-Y., Julien, S. G., Zhang, Q., et al. (2016). Comparative transcriptomic and epigenomic analyses reveal new regulators of murine brown adipogenesis. *PLoS Genet.* 12:e1006474. doi: 10.1371/journal.pgen.1006474
- Cheng, J., Blum, R., Bowman, C., Hu, D., Shilatifard, A., Shen, S., et al. (2014). A role for H3K4 monomethylation in gene repression and partitioning of chromatin readers. *Mol. Cell* 53, 979–992. doi: 10.1016/j.molcel.2014.02.032
- Core, L. J., Martins, A. L., Danko, C. G., Waters, C. T., Siepel, A., and Lis, J. T. (2014). Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat. Genet.* 46, 1311–1320. doi: 10.1038/ng.3142
- Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U.S.A.* 107, 21931–21936. doi: 10.1073/pnas.1016071107
- Cunningham, F., Achuthan, P., Akanni, W., Allen, J., Amode, M. R., Armean, I., et al. (2019). *Ensembl. Nucleic Acids Res.* 47, D745–751. doi: 10.1093/nar/gky1113
- Dorigi, K. M., Swigut, T., Henriques, T., Bhanu, N. V., Scruggs, B. S., Nady, N., et al. (2017). Mll3 and Mll4 facilitate enhancer RNA synthesis and transcription from promoters independently of H3K4 monomethylation. *Mol. Cell* 66, 568–576. e4.
- Dozmorov, M. G. (2015). Polycomb repressive complex 2 epigenomic signature defines age-associated hypermethylation and gene expression changes. *Epigenetics* 10, 484–495. doi: 10.1080/15592294.2015.1040619
- Encode Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74. doi: 10.1038/nature11247
- Garland, W., Comet, I., Wu, M., Radzisheuskaya, A., Rib, L., Vitting-Seerup, K., et al. (2019). A functional link between NUCLEAR RNA decay and transcriptional control mediated by the polycomb repressive complex 2. *Cell Rep.* 29, 1800–1811.e6. doi: 10.1016/j.celrep.2019.10.011
- Guo, C., Chen, L. H., Huang, Y., Chang, C.-C., Wang, P., Pirozzi, C. J., et al. (2013). KMT2D maintains neoplastic cell proliferation and global histone H3 lysine 4 monomethylation. *Oncotarget* 4, 2144–2153.
- Hammoud, S. S., Low, D. H. P., Yi, C., Carrell, D. T., Guccione, E., and Cairns, B. R. (2014). Chromatin and transcription transitions of mammalian adult germline stem cells and spermatogenesis. *Cell Stem Cell* 15, 239–253. doi: 10.1016/j.stem.2014.04.006
- Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., David Hawkins, R., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat. Genet.* 39, 311–318. doi: 10.1038/ng1966
- Hoffman, B. G., Robertson, G., Zavaglia, B., Beach, M., Cullum, R., Lee, S., et al. (2010). Locus co-occupancy, nucleosome positioning, and H3K4me1 regulate the functionality of FOXA2-, HNF4A-, and PDX1-bound loci in islets and liver. *Genome Res.* 20, 1037–1051. doi: 10.1101/gr.104356.109
- Hu, D., Gao, X., Morgan, M. A., Herz, H.-M., Smith, E. R., and Shilatifard, A. (2013). The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. *Mol. Cell. Biol.* 33, 4745–4754. doi: 10.1128/mcb.01181-13
- Kidder, B. L., Hu, G., and Zhao, K. (2014). KDM5B focuses H3K4 methylation near promoters and enhancers during embryonic stem cell self-renewal and differentiation. *Genome Biol.* 15:R32.
- Kotzin, J. J., Spencer, S. P., McCright, S. J., Kumar, D. B. U., Collet, M. A., Mowle, W. K., et al. (2016). The long non-coding RNA Morrbid regulates Bim and short-lived myeloid cell lifespan. *Nature* 537, 239–243. doi: 10.1038/nature19346
- Ku, M., Jaffe, J. D., Koche, R. P., Rheinbay, E., Endoh, M., Koseki, H., et al. (2012). H2A.Z landscapes and dual modifications in pluripotent and multipotent stem cells underlie complex genome regulatory functions. *Genome Biol.* 13:R85.
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Lesch, B. J., Dokshin, G. A., Young, R. A., McCarrey, J. R., and Page, D. C. (2013). A set of genes critical to development is epigenetically poised in mouse germ cells from fetal stages through completion of meiosis. *Proc. Natl. Acad. Sci. U.S.A.* 110, 16061–16066. doi: 10.1073/pnas.1315204110
- Lesch, B. J., and Page, D. C. (2014). Poised chromatin in the mammalian germ line. *Development* 141, 3619–3626. doi: 10.1242/dev.113027
- Lesch, B. J., Silber, S. J., McCarrey, J. R., and Page, D. C. (2016). Parallel evolution of male germline epigenetic poising and somatic development in animals. *Nat. Genet.* 48, 888–894. doi: 10.1038/ng.3591
- Liu, Y., Niu, M., Yao, C., Hai, Y., Yuan, Q., Liu, Y., et al. (2015). Fractionation of human spermatogenic cells using STA-PUT gravity sedimentation and their miRNA profiling. *Sci. Rep.* 5:8084.
- Local, A., Huang, H., Albuquerque, C. P., Singh, N., Lee, A. Y., Wang, W., et al. (2018). Identification of H3K4me1-associated proteins at mammalian enhancers. *Nat. Genet.* 50, 73–82. doi: 10.1038/s41588-017-0015-6
- Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., et al. (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448, 553–560. doi: 10.1038/nature06008
- Murphy, P. J., Wu, S. F., James, C. R., Wike, C. L., and Cairns, B. R. (2018). Placeholder nucleosomes underlie germline-to-embryo DNA methylation reprogramming. *Cell* 172, 993–1006. e13.
- Ooi, S. K. T., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., et al. (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* 448, 714–717. doi: 10.1038/nature05987
- Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842. doi: 10.1093/bioinformatics/btq033
- Rada-Iglesias, A., Bajpai, R., Swigut, T., Brugmann, S. A., Flynn, R. A., and Wysocka, J. (2011). A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 470, 279–283. doi: 10.1038/nature09692
- Rickels, R., Herz, H.-M., Sze, C. C., Cao, K., Morgan, M. A., Collings, C. K., et al. (2017). Histone H3K4 monomethylation catalyzed by Trx and mammalian COMPASS-like proteins at enhancers is dispensable for development and viability. *Nat. Genet.* 49, 1647–1653. doi: 10.1038/ng.3965
- Shepherd, R. W., Millette, C. F., and DeWolf, W. C. (1981). Enrichment of primary pachytene spermatocytes from the human testis. *Gamete Res.* 4, 487–498. doi: 10.1002/mrd.1120040602
- Wickham, H. (2011). ggplot2. *Wiley Interdiscipl. Rev. Comput. Statist.* 3, 180–185. doi: 10.1002/wics.147
- Zhang, Y., Liu, T., Meyer, C. A., Eeckhoutte, J., Johnson, D. S., Bernstein, B. E., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9:R137.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Bae and Lesch. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Use of Mononucleosome Immunoprecipitation for Analysis of Combinatorial Histone Post-translational Modifications and Purification of Nucleosome-Interacting Proteins

Kashif Aziz Khan, Marlee K. Ng and Peter Cheung*

Department of Biology, York University, Toronto, ON, Canada

OPEN ACCESS

Edited by:

Jean-Philippe Lambert,
Laval University, Canada

Reviewed by:

Nicoletta Landsberger,
University of Milan, Italy

Michael Bulger,
University of Rochester, United States

*Correspondence:

Peter Cheung
pmcheung@yorku.ca

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 14 February 2020

Accepted: 16 April 2020

Published: 08 May 2020

Citation:

Khan KA, Ng MK and Cheung P
(2020) The Use of Mononucleosome
Immunoprecipitation for Analysis
of Combinatorial Histone
Post-translational Modifications
and Purification
of Nucleosome-Interacting Proteins.
Front. Cell Dev. Biol. 8:331.
doi: 10.3389/fcell.2020.00331

The nucleosome is the principal structural unit of chromatin. Although many studies focus on individual histone post-translational modifications (PTMs) in isolation, it is important to recognize that multiple histone PTMs can function together or cross-regulate one another within the nucleosome context. In addition, different modifications or histone-binding surfaces can synergize to stabilize the binding of nuclear factors to nucleosomes. To facilitate these types of studies, we present here a step-by-step protocol for isolating high yields of mononucleosomes for biochemical analyses. Furthermore, we discuss differences and variations of the basic protocol used in different publications and characterize the relative abundance of selected histone PTMs and chromatin-binding proteins in the different chromatin fractions obtained by this method.

Keywords: nucleosome, immunoprecipitation, mononucleosome IP, MNase, combinatorial histone modifications, chromatin-binding proteins, histone variant

INTRODUCTION

The nucleosome is the fundamental repeating unit of chromatin in eukaryotic cells and is the main physiological state by which the functional genome engages the nuclear environment. A nucleosome typically consists of 147 bp of DNA wrapped around a histone octamer comprising two copies each of the core histones H2A, H2B, H3, and H4 (Luger et al., 1997). The composition and characteristics of nucleosomes can vary within the genome through the incorporation of histone variants or post-translational modifications (PTMs) of core and variant histones (Talbert and Henikoff, 2017). Moreover, the positioning and organization of nucleosomes over different parts of the genome can be further modulated by chromatin-remodeling complexes and chromatin-binding proteins (Lai and Pugh, 2017).

Core and linker histones are the main protein components of nucleosomes. They are highly and specifically expressed during S phase to cope with the demands of DNA replication-coupled chromatin assembly. Core histones are also ubiquitously distributed across the genome to form the general scaffold of genomic chromatin. Unlike core histones, histone variants are expressed and deposited into chromatin in a replication-independent manner (Henikoff and Smith, 2015). The distribution of histone variants can also be more targeted and localized such as the restriction

of CENPA to centromeres. To date, variants of all 4 core histone types have been identified. The variant family of H2A is the most diverse and includes several members such as H2A.Z-1, H2A.Z-2, H2A.X, macroH2A1, macroH2A2, and H2A.Bbd. Other well-studied variants include H3.3, CENP-A/cenH3, H3.X, H3.Y of the H3 family, and H2BE, TSH2B, H2BFWT of the H2B family (Law and Cheung, 2013; Maze et al., 2014). More recently, an H4 variant, H4G, has also been identified in human cells (Long et al., 2019). Similar to core histones, histone variants are post-translationally modified at amino acids conserved between the variant and its core histone counterpart, or at variant-specific sequences.

The regulation and functions of histone PTMs have been heavily studied in the past 25 years (Zhao and Garcia, 2015; Stillman, 2018). Histones are modified by a variety of modifying enzymes and the modified histones, in turn, can elicit or facilitate specific downstream events. In the natural context, many combinations of histone PTMs co-exist on the same histone molecule and also on different histones within the same nucleosome. Some histone modifications are functionally coupled, such as the requirement of H2B mono-ubiquitylation for H3K4 methylation in yeast and human cells, or the coupling of H3 phosphorylation and acetylation during activation of immediate-early genes (Lee et al., 2010; Ng and Cheung, 2016). These examples illustrate that some histone PTMs cross-talk and cross-regulate one another as part of their regulatory mechanisms (Strahl and Allis, 2000; Lee et al., 2010). In addition, there are also combinations of histone PTMs that co-exist to mark distinct chromatin states. For example, although H3K4- and H3K27-methylation are respectively linked to gene activation and repression, these two histone modifications can also exist together to mark “bivalent” domains that correspond to poised but transcriptionally silenced genes in the stem cell genome. More interestingly, biochemical characterization of bivalent nucleosomes showed that the respective H3 methylation marks are located on the different H3 molecules within the same nucleosome. Therefore, nucleosomes can be heterotypically modified on the histone dimer pairs, leading to more complex combinatorial patterns of histone PTMs within chromatin (Voigt et al., 2012; Sen et al., 2016; Shema et al., 2016).

One of the functions mediated by histone PTMs is the recruitment of effector proteins via PTM-dependent interactions. The discovery of bromodomain-containing proteins binding to acetylated histones, and some chromodomain-containing proteins binding to methylated histones, led to the concept that there are families of “reader” proteins that bind specific histone modifications and are recruited to target sites in a PTM-dependent manner to execute downstream functions (Taverna et al., 2007). There is also accumulating evidence that recruitment of effector proteins can occur across multiple core and variant histones by binding to multiple epitopes within the nucleosomes. This type of “multivalent” interaction is not easily deciphered using individual histones or histone peptides alone since they could involve interactions with physically distal modifications or epitopes found on different histones. Indeed, one study that directly compared the PTM-reader interactions between peptide versus nucleosome contexts found only limited

overlap of the co-purified reader proteins using the respective peptide/nucleosome baits (Nikolov et al., 2011). Similarly, many studies used affinity purification to identify histone-binding proteins and they generally do not distinguish between soluble versus nucleosomal histones. For example, in an effort to identify H2A.Z interacting proteins, many studies used whole cell extracts as a source of histones and predominantly identified chaperones and remodeling complexes that bind free H2A.Z (reviewed in Ng and Cheung, 2016). However, additional unique proteins were identified when H2A.Z in the nucleosome context was used as a bait to pull down H2A.Z interacting proteins (Fujimoto et al., 2012). Therefore, there is a need, as well as a growing interest, in studying histone-nuclear factor interactions at the nucleosomal level.

One of the approaches used to study histones in the nucleosome context is the reconstitution of nucleosomes *in vitro*. The nucleosome core particle or a nucleosome array can be reconstituted under low salt conditions using recombinant histones and recombinant DNA containing multiple repeats of a nucleosome positioning sequence such as the “601 sequence” (Lowary and Widom, 1998; Dyer et al., 2004). Alternatively, *in vitro* nucleosome assembly can also be done using ATP-dependent assembly factors such as recombinant ACF and RSF1 (Lusser and Kadonaga, 2004). In addition, chemically modified or peptide-ligated recombinant histones carrying specific PTMs have been generated that are in turn assembled into “designer” nucleosomes (Muller and Muir, 2015; Nadal et al., 2018). These approaches allow better control over the composition of the nucleosomes and produce a homogenous sample that is suitable for *in vitro* biochemical assays. However, such nucleosomes lack the complex range of PTMs normally seen in endogenous nucleosomes and may not fully replicate physiological chromatin.

Endogenous nucleosomes are historically obtained by treatment of chromatin with micrococcal nuclease (MNase), which preferentially cuts the linker DNA to generate single nucleosomes (reviewed in Kornberg, 1977), followed by immunoprecipitation (IP) of core/variant histones or histones modified by specific PTMs. Mononucleosome IP has been used by us and others to demonstrate preferential combinations of histone PTMs or histone variants that co-exist within individual nucleosomes (Sarcinella et al., 2007; Ku et al., 2012; Voigt et al., 2012; Chen et al., 2014; Lacoste et al., 2014; Wang et al., 2014, 2018; Won et al., 2015; Surface et al., 2016), or to identify proteins interacting with histone PTMs or histone variants in the nucleosome context (Draker et al., 2012; Kim et al., 2013; Sansoni et al., 2014; Vardabasso et al., 2015; Li et al., 2016; Punzeler et al., 2017; Zhang et al., 2017; Zink et al., 2017; Sun et al., 2018). In addition, the same method has been used to show incorporation of specific core/variant histone in the chromatin (Kanda et al., 1998; Wiedemann et al., 2010; Lau et al., 2011; Ruiz and Gamble, 2018), and to demonstrate effects of oncohistones on chromatin (Bender et al., 2013; Chan et al., 2013; Lewis et al., 2013; Herz et al., 2014; Fang et al., 2016; Lu et al., 2016; Piunti et al., 2017). However, there are subtle to considerable differences among the protocols used in different studies, which may lead to variations in findings, such as some differences in the H2A.Z nucleosome-interacting proteins found in different

studies. We, therefore, review here the differences and variations among the protocols used by different publications to generate and immunoprecipitate mononucleosomes in order to provide direct comparisons for the readers. In addition, we also describe a mononucleosome purification and IP protocol used in our lab as a starting point for readers to test and optimize. This protocol describes a step-by-step procedure to obtain a high yield of mononucleosomes using MNase followed by IP of histone variant containing mononucleosomes. This protocol can be used to identify co-existing PTMs on histone variants and partnered core histones within the nucleosome, as well as nucleosome-interacting proteins. The schematic representation of mononucleosome IP protocol is shown in **Figure 1**.

VARIATIONS AND OPTIMIZATION OF THE MONONUCLEOSOME IP PROTOCOL

Studies of histones at the nucleosomal level require a good yield of mononucleosomes that is typically obtained by *in nucleo* digestion of nuclei by MNase. Nuclei are isolated by swelling of cells in a hypotonic solution followed by the addition of a detergent to disrupt the cellular membrane (Mendez and Stillman, 2000). Pure nuclei are recovered by centrifugation and then digested with MNase in a CaCl_2 -containing buffer to cut the linker region, followed by centrifugation to recover the mononucleosome containing supernatant (S1). There are generally only minor differences amongst protocols used by different studies in terms of the composition of hypotonic solution or CaCl_2 -containing buffer for the digestion of nuclei by MNase to extract S1; however, there are significant differences in the approaches used to recover remaining mononucleosomes from the pellet as the 2nd supernatant (S2) (**Figure 2**).

A number of studies used MNase-digested supernatant (S1) only for IP, leaving out the remaining chromatin and insoluble material after MNase digestion (Foltz et al., 2006; Wiedemann et al., 2010; Kim et al., 2013; Chen et al., 2014; Vardabasso et al., 2015; Won et al., 2015; Li et al., 2016; Punzeler et al., 2017; Zink et al., 2017; Kim et al., 2018; Ruiz and Gamble, 2018; Sun et al., 2018). Other studies included additional steps to obtain an S2 fraction for the maximum recovery of mononucleosomes from MNase-digested chromatin. A variety of different approaches have been used for this step; one approach is to sonicate the insoluble material to increase the yield of mononucleosomes (Viens et al., 2006; Fang et al., 2015) while another is to use a high salt-containing buffer based on the classical Dignam nuclear extraction protocol (Dignam et al., 1983) to extract residual chromatin and proteins. The latter method involves incubating the MNase-digested nuclei in 420–500 mM NaCl-containing buffer followed by centrifugation to collect S2 (Kanda et al., 1998; Sarcinella et al., 2007; Lau et al., 2011; Ku et al., 2012; Voigt et al., 2012; Surface et al., 2016; Zhang et al., 2017). Although nucleosome integrity is maintained under high salt concentration, these conditions can disrupt the interaction of chromatin-binding proteins. Moreover, we have observed significant precipitation of insoluble chromatin during the salt

extraction, likely due to the mobilization of linker histone H1 under these conditions (Clark and Thomas, 1986; Al-Natour and Hassan, 2007). Therefore, for our studies of combinatorial histone PTMs and nucleosome-interacting proteins, we prefer to isolate mononucleosomes from MNase-digested nuclei under low or no salt conditions. We, like other studies, collect the mononucleosome containing supernatant S1 by centrifugation after digestion with MNase, but we then resuspend the MNase-digested nuclei pellet in salt-free TE buffer to lyse the nuclei for maximum recovery of mononucleosomes (Draker et al., 2012). As demonstrated in **Figure 3**, suspending intact nuclei in the no salt buffer leads to rupture of the nuclear membrane and further liberation of free intact nucleosomes.

Although our S1 showed a good yield of mononucleosomes, our S2 fraction showed even higher yield of mononucleosomes, as evident in the side-by-side comparison of the amount of core histone proteins in these fractions by Coomassie gel staining or by the levels of histone H3 and H2A.Z in immunoblot analysis (**Figure 4A**). Densitometry analysis of the H4 band from Coomassie gel and H3 blot using ImageJ (Schneider et al., 2012) showed that around 75% of total histones in all 3 fractions were recovered in the combined S1 and S2 fractions (**Figure 4B**). Blotting of FLAG-tagged histones showed similar results (data not shown). Although some histones are still left in the insoluble pellet, this method allows us to pool both S1 and S2 fractions before IP to ensure that results are obtained using the majority of released nucleosomes. Slightly different from our method, other studies isolate mononucleosomes after a second longer incubation of the post-S1 extracted nuclei in physiological salt concentration. For example, Imhof and colleagues incubated the MNase-digested nuclei in 150 mM NaCl containing buffer with end-to-end rotation overnight and centrifuged to recover S2. Nonetheless, in both cases, the S1 and S2 fractions were pooled before IP (Sansoni et al., 2014).

To analyze the differences and similarities between the nucleosomes in the S1, S2, and pellet fractions in our protocol, we blotted for different histone PTMs in these fractions after normalization to the histone H3 levels in the samples (**Figure 5A**). For most PTMs, such as H3K4me3, H3K9me3, H3K27me3, K3K36me3, H3K79me2, and ub-H2A, their levels are quite comparable in the nucleosomes found in all three fractions. This suggests that the bulk of general chromatin is similarly distributed among the different fractions, and raises the confidence that the S1 + S2 fractions are representative of the bulk of genomic chromatin. However, we did observe that the majority of ubiquitylated H2B (detected by the mono-clonal H2Bub antibody) was found in the insoluble pellet fraction. We currently do not have an explanation for this observation but it is possible that H2Bub is enriched at chromatin domains or sub-nuclear compartments that are more resistant to MNase digestion. On the other hand, H2Bub is often associated with active chromatin in yeast and mammalian cells and; therefore, is not expected to be associated with repressive or compacted chromatin. When we examined the distribution of nuclear factors across the different fractions (see next section), we also find proteins associated with active transcription, such as Brd2 and RNA polymerase II (RNAPII) in the pellet fraction too,

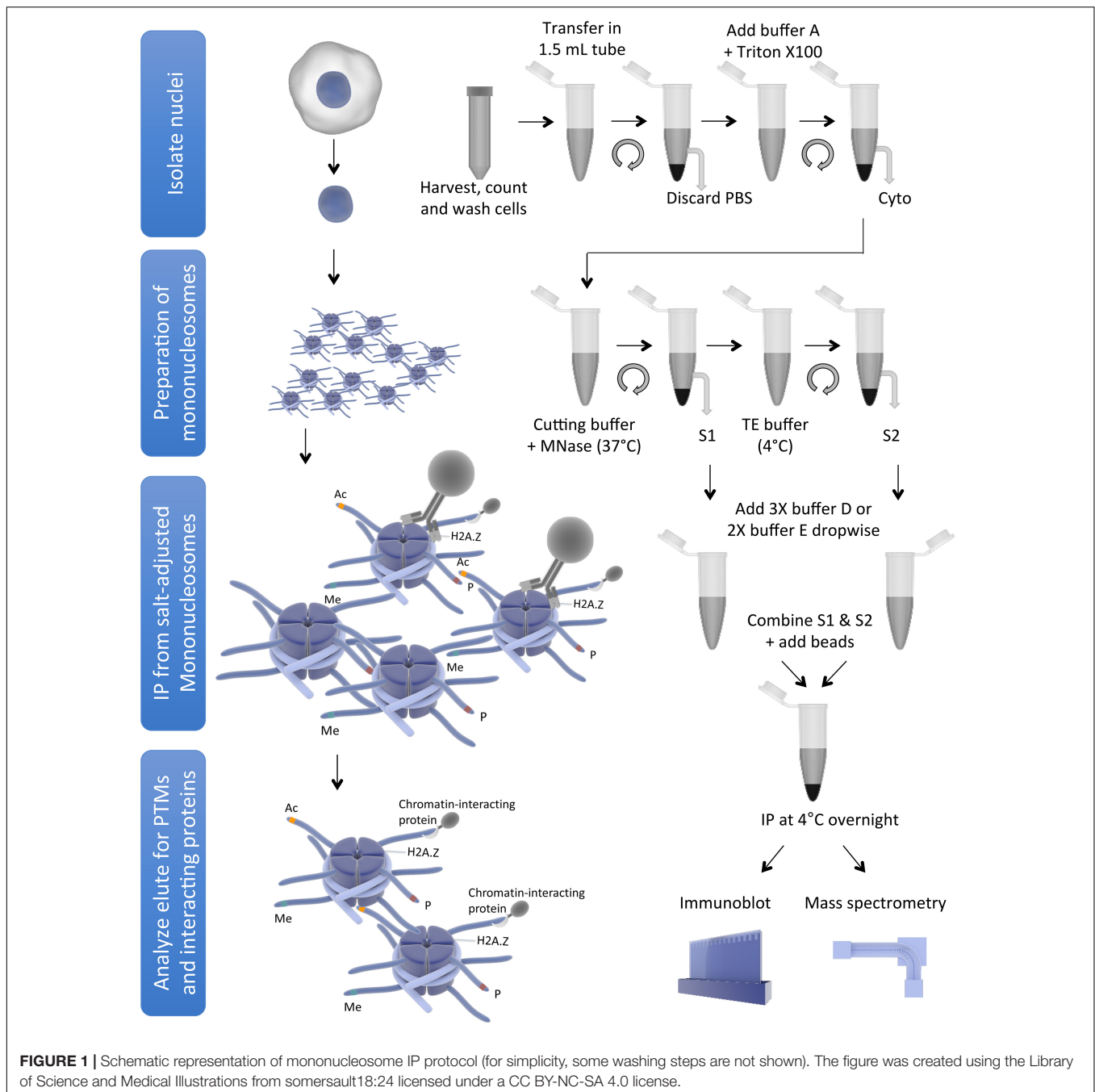


FIGURE 1 | Schematic representation of mononucleosome IP protocol (for simplicity, some washing steps are not shown). The figure was created using the Library of Science and Medical Illustrations from somersault18:24 licensed under a CC BY-NC-SA 4.0 license.

suggesting that the pellet still contains significant amounts of chromatin associated with active transcription. H2Bub may be associated with an active chromatin compartment that remains insoluble in this fractionation protocol, and further investigation will be needed to fully understand the molecular basis of this observation.

For additional characterization of the separated fractions, we also blotted for selected chromatin-interacting proteins including Brd2, USP39, USP7, PHF6, and RNAPII in these fractions. In this case, we did the comparisons using an equivalent number of cells instead of H3 normalization since

different amounts of histones were found in these fractions. We found that chromatin-interacting proteins were mostly found in the S1 fraction as well as in the insoluble-pellet fraction (**Figure 5B**). Despite the fact that we are able to recover the majority of nucleosomes in the combined S1 and S2 fractions, much more of the chromatin-interacting proteins were found in the insoluble pellet compared to the S1/S2 fractions. Of interest, the heterochromatin marker, HP1 α , and nucleolar marker, Fibrillarin, were only detected in the insoluble-pellet fraction by our assay, suggesting that heterochromatin and nucleoli are more resistant to MNase digestion (**Figure 5B**). As

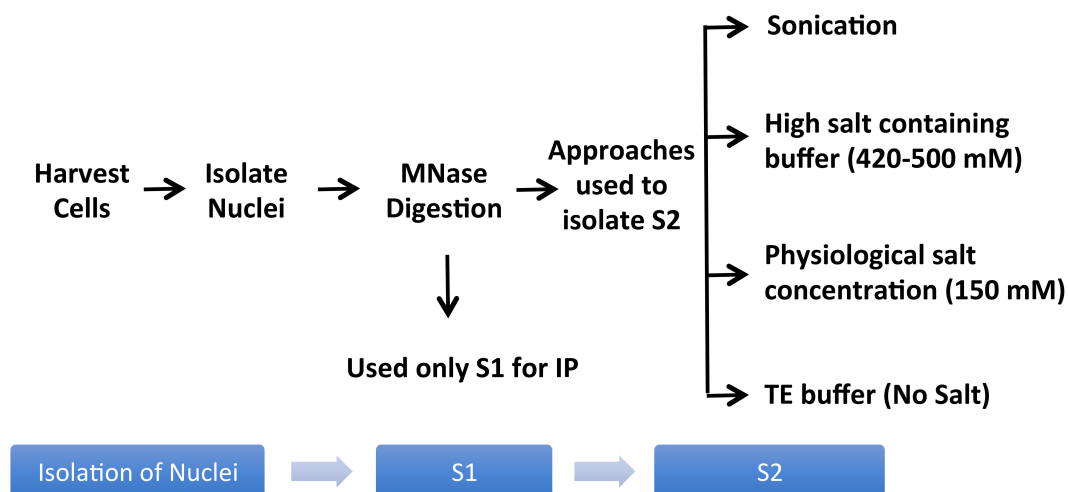


FIGURE 2 | Variations of the mononucleosome IP protocol used in different publications. Pure nuclei are digested with MNase to cut the linker region followed by centrifugation to recover the MNase-digested supernatant (S1). Several studies used S1 only for IP, leaving out the insoluble material altogether (Foltz et al., 2006; Wiedemann et al., 2010; Kim et al., 2013; Chen et al., 2014; Vardabasso et al., 2015; Won et al., 2015; Li et al., 2016; Punzeler et al., 2017; Zink et al., 2017; Kim et al., 2018; Ruiz and Gamble, 2018; Sun et al., 2018). Additional steps used in the literature to obtain an S2 fraction for the maximum recovery of mononucleosomes from MNase-digested chromatin include sonication (Viens et al., 2006; Fang et al., 2015), extraction using high salt containing buffer (Kanda et al., 1998; Sarcinella et al., 2007; Lau et al., 2011; Ku et al., 2012; Voigt et al., 2012; Surface et al., 2016; Zhang et al., 2017), buffer with physiological salt concentration (Sansoni et al., 2014), or no-salt TE buffer (Draker et al., 2012).

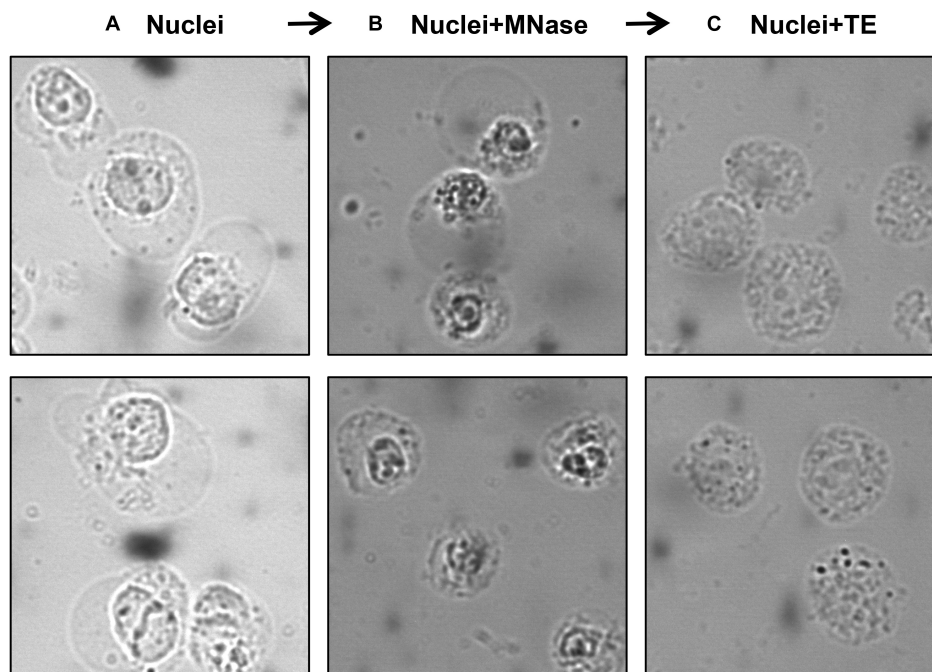


FIGURE 3 | Visualization of nuclei at different steps of the protocol using a phase-contrast microscope. Isolated nuclei at Step 6 (A) MNase-digested nuclei at Step 9 (B) TE-resuspended nuclei at Step 11 (C).

the S1 fraction contained more chromatin-interacting proteins compared to S2, our results suggest that one may opt to use S1 fraction only to assay chromatin-nuclear protein interactions, especially if the sensitivity of the detection of the nuclear proteins is an issue. However, we would recommend incubating

the nuclei in S1 for a couple of hours after stopping the MNase reaction with EGTA (see step 9 below), as it helps to increase the yield of mononucleosomes in S1 versus S2 with no effect on the yield of mononucleosomes in the insoluble pellet (Figure 4C).

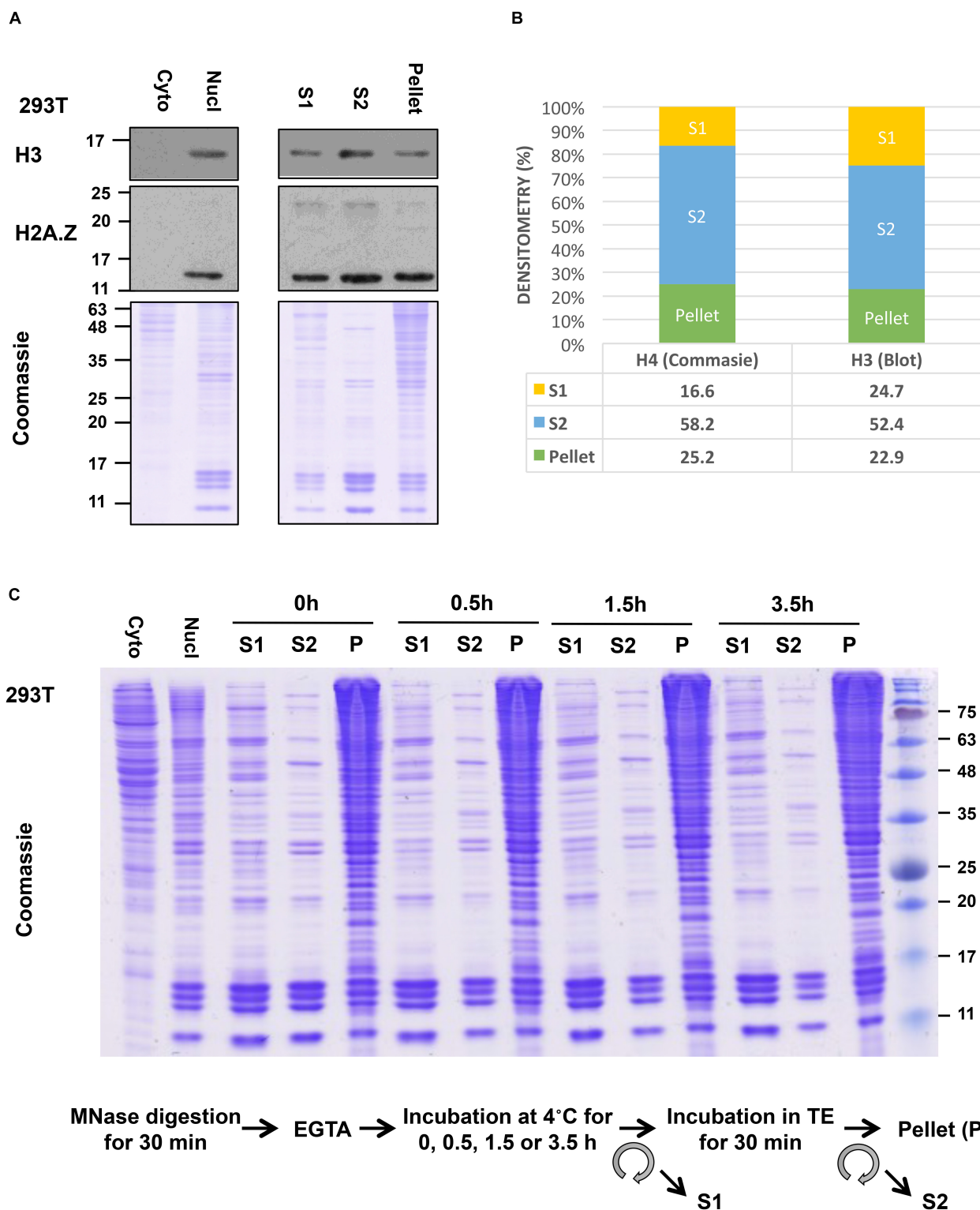
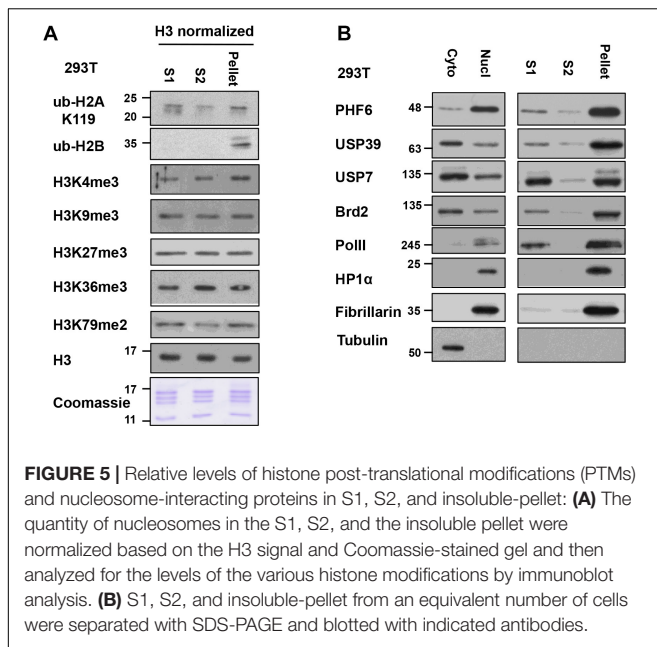


FIGURE 4 | Yield of mononucleosomes in the MNase-digested (S1), the TE-soluble (S2), and the insoluble-pellet fractions. **(A)** Nuclei from 293T cells were treated with 1.0 U/10⁶ cells and different fractions from an equivalent number of cells were separated with 10–15% SDS-PAGE and blotted with indicated antibodies or stained with Coomassie stain. Cytoplasmic and nuclear extract from an equivalent number of cells was included in parallel. **(B)** Densitometry analysis of H4 band from Coomassie gel and H3 blot using ImageJ. **(C)** The nuclei were incubated for an extended period after stopping the MNase reaction with EGTA (step 9) followed by centrifugation to recover S1 and then incubated in TE buffer to isolate S2 as described in the protocol.



As a last note about this method, we have mostly expressed epitope-tagged histones in mammalian cells for purification of mono-nucleosomes because of the excellent immunoprecipitation efficiencies and capacities of the variety of antibodies against different epitope tags. Many studies have used histones that are tagged either at the C- or N-termini for detection and analysis, and expression of these histones, particularly in mammalian cells, generally does not affect their incorporation into chromatin, nor result in any deleterious effects. Although we have not validated the protocol for animal tissues, the technique should be adaptable for tissue samples. Once nuclei have been isolated from tissue samples using standard methods (Zaret, 2005; Krishnaswami et al., 2016), they can be treated with MNase for isolation of mononucleosomes using this protocol. However, it is important to note that, compared to tissue culture cells, the amount of tissues available for analyses will likely be limiting, which will impact on the final yield of nucleosomes. In addition, it should be possible to use various antibodies raised against endogenous histones or histone PTMs to immunoprecipitate endogenous nucleosomes akin to the native ChIP method. However, capturing sufficient nucleosomes for biochemical analyses will require antibodies that have excellent immunoprecipitation efficiencies and in larger scale compared to ChIP. The suitability of different histone antibodies for this nucleosome immunoprecipitation technique will have to be empirically determined in each case.

STEP-BY-STEP PROTOCOL

Isolation of Nuclei

1. Harvest 293T cells expressing FLAG-tagged histone of interest via trypsinization, resuspend in complete media and count cells (see Note 1 and 2).

2. Pellet cells via centrifugation (5 min, at $300 \times g$, 4°C), wash 1–2 times in cold 1X PBS and transfer cells in a 1.5 mL microcentrifuge tube. An aliquot can be taken out to prepare whole cell lysate to monitor the expression of FLAG-tagged histone.

Like any other protein analysis, samples and buffers should be kept on ice all the time unless otherwise stated.

3. Wash cells by thoroughly resuspending in 1 mL of Buffer A (for recipe see **Table 1**) containing protease inhibitors (see Note 3 and 4).
4. Pellet cells via centrifugation: $300 \times g$, 5 min, 4°C .
5. Remove and discard the supernatant. Thoroughly resuspend cells in 1 mL Buffer A as above and add Triton-X 100 to a final concentration of 0.2%. Mix and incubate on ice for 5 min.

Make 10% Triton-X 100 in sterile water and use 20 μL /1 mL of Buffer A to get a final concentration of 0.2%

6. Pellet nuclei: $600 \times g$, 5 min, 4°C . Keep and store the supernatant at -80°C , containing the cytoplasmic fraction, if desired.

MNase Digestion

7. Wash the nuclei in 1 mL of Buffer A containing protease inhibitors (see Note 3 and 4).
8. Remove and discard the supernatant and resuspend the pellet in 500 μL of Cutting Buffer (for recipe see **Table 2**). Add MNase (Worthington) to the nuclei suspension at 1 U per 10^6 cells and digest nuclei at 37°C for 30 min (see Note 5 and 6). Conditions for MNase digestion must be determined empirically for each cell line used (Note 6; **Figure 6**).
9. Stop the MNase reaction by adding EGTA to a final concentration of 20 mM, mix and place on ice.

E.g., add 20 μL of 0.5 M EGTA stock/500 μL sample to get 20 mM final.

10. Pellet nuclei: $1300 \times g$, 5 min, 4°C . Transfer the supernatant to a fresh tube and store on ice (referred

TABLE 1 | Buffer A.

Stock	The amount for 40 mL	Final concentration
1M HEPES (pH = 8) (pH = 7.35 if using NEM)	400 μL	10 mM
1M KCl	400 μL	10 mM
1M MgCl_2	60 μL	1.5 mM
1M Sucrose	13.6 mL	340 mM
50% Glycerol	8 mL	10%
1M DTT ¹	40 μL	1 mM
Autoclaved MilliQ H_2O	Till 40 mL	

¹Add fresh.

TABLE 2 | Cutting buffer.

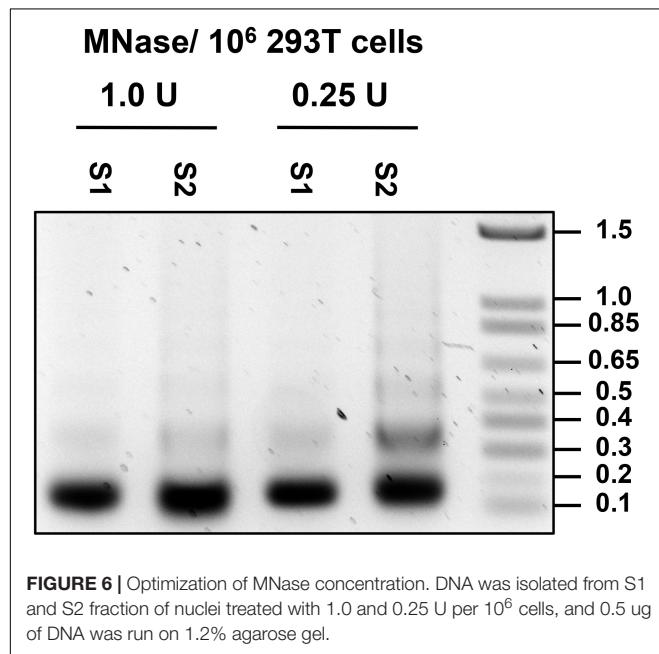
Stock	The amount for 40 mL	Final concentration
1M NaCl	600 μ L	15 mM
1M KCl	2.4 mL	60 mM
1M Tris (pH = 7.5)	400 μ L	10 mM
1M CaCl ₂	80 μ L	2 mM
Autoclaved MilliQ H ₂ O	Till 40 mL	

TABLE 3 | TE buffer.

Stock	The amount for 40 mL	Final concentration
1M Tris (pH 8.0)	400 μ L	10 mM
0.5M EDTA (pH 8.0)	80 μ L	1 mM
Autoclaved MilliQ H ₂ O	Till 40 mL	

TABLE 4 | 2X SDS sample buffer.

Stock	The amount for 40 mL	Final concentration
1M Tris (pH = 7.4)	0.8 mL	20 mM
0.5M EDTA (pH 8.0)	1.6 mL	20 mM
10% SDS	8 mL	2%
50% Glycerol	16 mL	20%
Autoclaved MilliQ H ₂ O	Till 40 mL	



hereafter as the MNase-digested S1 fraction). A portion can be stored at -80°C for further analysis.

- Resuspend the pellet with 500 μ L of TE buffer (for recipe see **Table 3**) containing protease inhibitors (see Note 3 and 4). Incubate on ice for 30 min, mixing with a pipette every 10–15 min. Alternatively, samples can be rotated constantly using an end-over-end rotator at 4°C .
- Spin samples 5 min at $13,000 \times g$, 4°C to pellet cell debris and insoluble material. Transfer supernatant to a new tube (referred here as the TE-soluble S2 fraction). A portion can be stored at -80°C for further analysis.
- If desired, resuspend pellet in 250 μ L of PBS and add an equal amount of hot 2X sample buffer (for recipe see **Table 4**), boil for 5 min to denature the proteins, and sonicate to resuspend proteins. Centrifuge at $> 13,000 \times g$ to remove insoluble material, if any. Transfer supernatant to a new tube (referred here as an insoluble-pellet fraction).

Immunoprecipitation (IP) of Nucleosomes

- Transfer *exactly* 475 μ L of S1 fraction from step 10 to a new tube. To re-adjust salt concentration, add 475 μ L of 2X Buffer E *dropwise* (1 drop every 2–3 s) with constant

mixing using the vortex at low speed (see Note 7 and 8; for recipe see **Table 5**).

- Transfer *exactly* 475 μ L of S2 fraction from step 12 to a new tube. To re-adjust salt concentration of the S2 fraction, add 237 μ L of 3X Buffer D *dropwise* (1 drop every 2–3 s) with constant mixing using the vortex at low speed (Note 7 and 8; for recipe see **Table 6**).
- Centrifuge tubes from Steps 14 and 15 (5 min at $13,000 \times g$, 4°C) and collect supernatant (salt-adjusted S1 and S2 fractions).
- Pool the salt-adjusted S1 and S2 fractions in a 15 mL propylene tube to use as input for IP.
- Transfer 50–100 μ L from each pooled sample to a new tube and store at -80°C (save as input); use the remainder of the sample for IP.
- Transfer 50% slurry of FLAG/M2 resin in a microcentrifuge tube using a wide bore tip. We typically use 20 μ L per IP condition and wash a batch of beads enough for all IP conditions in an experiment (see Notes 9 and 10).

The current protocol uses a FLAG-tagged histone as an example. See Notes 11 and 12 for details on other antibodies and types of beaded support for purification.

- Pellet the beads by centrifuge ($700 \times g$ for 15 s at 4°C) and discard the supernatant.

TABLE 5 | 2X buffer E.

Stock	The amount for 40 mL	Final concentration
1M HEPES (pH = 7.5)	1.2 mL	30 mM
4M NaCl	2.25 mL	225 mM
1M MgCl ₂	120 μ L	3 mM
10% Triton X-100	1.6 mL	0.4%
50% Glycerol	16 mL	20%
Autoclaved MilliQ H ₂ O	Till 40 mL	

TABLE 6 | 3X buffer D.

Stock	The amount for 40 mL	Final concentration
1M HEPES (pH = 7.5)	2.4 mL	60 mM
4M NaCl	4.5 mL	450 mM
1M MgCl ₂	180 μ L	4.5 mM
0.5M EGTA (pH = 8)	48 μ L	0.6 mM
10% Triton X-100	2.4 mL	0.6%
50% Glycerol	24 mL	30%
Autoclaved MilliQ H ₂ O	Till 40 mL	

21. Wash the beads twice in 1X buffer D by pelleting ($700 \times g$ for 15 s at 4°C) and gently resuspending with the buffer without pipetting up and down (see Note 13).
22. After the final wash, resuspend the beads with 1X buffer D to make 50% slurry.
23. Incubate 20 μ L of 50% slurry with the IP sample from step 17 overnight at 4°C in 15-mL propylene tubes. Keep samples rotating constantly using an end-over-end rotator.
24. After overnight incubation, pellet the beads by centrifugation ($700 \times g$ for 15 s at 4°C). Save supernatant, if desired (= UNBOUND).
25. Resuspend beads in 500 μ L 1X buffer D and transfer the beads to a microcentrifuge tube using a wide-bore 1000 μ L tip for subsequent washing steps.
26. Repeat the step above using 500 μ L 1X buffer D to transfer leftover beads from 15 mL tube to the same microcentrifuge.
27. Wash pelleted down beads 2–4 times in 1X Buffer D by pelleting ($700 \times g$ for 15 s at 4°C) and gently resuspending with the buffer without pipetting up and down (see Note 13).
28. Wash the beads another 2–4 times in 1X Buffer D + 0.5% Triton X-100 without pipetting up and down (see Note 13).
29. The immunoprecipitated nucleosomes and proteins can be processed for mass spectrometry or immunoblot analysis at this point.

SDS-PAGE and Immunoblot Analysis

30. Gently resuspend the beads collected from step 28 in 20 μ L 2X SDS sample buffer and boil for 5 min to denature the protein followed by centrifugation to pellet the beads: $700 \times g$ for 15 s at room temperature (see Note 14). Also, boil the input from step 18 in an equal volume of 2X SDS sample buffer.
31. Run input and IPs on 15% SDS-PAGE and stain with Coomassie stain for 1 h followed by destaining overnight to confirm the presence of mononucleosomes.
32. Normalize the IP'd nucleosomes from different samples by the total histones IP'd (e.g., based on Coomassie-stained histone bands) or by H3 (by immunoblot analysis). Once the amount of H3 IP'd across different samples has been normalized, perform immunoblot analysis using antibodies against Flag, different histone PTMs, or other proteins of interest to ascertain the relative amounts of

proteins/histone PTMs that are associated with or co-IP with the immunoprecipitated nucleosomes.

Extraction of DNA From MNase-Digested S1 and TE-Soluble S2 Fractions

1. Aliquot 100 μ L each of S1 and S2 fractions in 1.5 mL microtubes. Add 1 μ L of 10 mg/ml Proteinase K and incubate at 37°C for 3 h to overnight.
2. Add an equal volume of Phenol: Chloroform/Isoamyl alcohol (24:1), mix well by rocking and centrifuge at high speed for 5 min.
3. Collect the top aqueous phase of each tube and transfer to a new 1.5 mL microtube with an equal volume of Chloroform/Isoamyl alcohol 24:1. Vortex and centrifuge at high speed for 5 min.
4. Collect the top aqueous phase of each tube and transfer to a new 1.5 mL microtube. Centrifuge at high speed for 5 min to separate aqueous phase from carried over organic phase, if any.
5. Collect the top aqueous phase in a new 1.5 mL microtube and precipitate DNA by adding 1/10 volume of 3M Sodium acetate pH 5.2 and twice the volume of ice-cold ethanol.
6. Incubate at –20°C for 2 h or overnight and centrifuge at high speed at 10 min to pellet the DNA.
7. Discard the supernatant and wash the pellet with 500 μ L of 70% ethanol. Centrifuge at high speed for 5 min.
8. Discard the supernatant and air-dry the pellet for 15 min.
9. Resuspend the pellet in 100 μ L of TE buffer. If necessary, allow the DNA to dissolve at 4°C for a couple of hours and quantify DNA using nanodrop.
10. Run 0.5 μ g of DNA per condition in 1.2% Agarose gel to access the level of shearing of chromatin.

Notes

1. Cell culture: Avoid cells that have become over-confluent. Passage cells while still sub-confluent. We typically passage 293T cells twice a week.
2. The required number of cells depends on the target application. We typically IP from a confluent 15 cm plate of 293T cells (roughly 40–50 million cells) and 1/20 of eluate is used for Coomassie gel and blotting PTMs while 1/4 of eluate is used for blotting nucleosome-interacting proteins.
3. To avoid protein degradation, add the following protease inhibitors to all the buffers at a stated concentration just prior to use; PMSF 200 μ M, Aprotinin 1 μ g/ml, Leupeptin 10 μ M, Pepstatin 1 μ M.
4. To prevent loss of PTMs, add the following additives to all the buffers, depending on the experiment; an inhibitor of deubiquitinases *N*-ethylmaleimide (NEM) at 10 mM final concentration, HDAC inhibitor Sodium Butyrate at 5 mM final concentration, or phosphatase inhibitor Microcystin-LR at 0.1 μ M final concentration.
5. MNase (from Worthington) is resuspended in 5 mM Tris-HCl (pH 7.5)/10 μ M CaCl₂ at 50 U/ μ L, aliquoted, and

stored at -20°C . Freeze-thaw cycles should be kept to a minimum. As different manufacturers define units of MNase differently, special attention should be paid to the unit definition and conversion formula if using a different source.

6. Optimal MNase amount needed for chromatin digestion must be determined empirically for each cell line used, source of MNase and specific batch of MNase. This is done by treating cells with varying amounts of MNase and DNA extracted from the S1 and S2 fractions for each condition are run on a 1.2% Agarose gel (see section “Extraction of DNA From MNase-Digested S1 and TE-Soluble S2 Fraction”). Mononucleosome preparation typically should consist of mononucleosomes with a small amount of dinucleosomes still present (**Figure 6**). Special attention should be paid to avoid over-shearing as MNase is able to cleave DNA internally at ~ 10 bp interval where it is exposed to the outer surface of the core nucleosome (Clark, 2010).
7. If 3X Buffer D and 2X Buffer E are prepared ahead of time, store at 4°C ; incubation on ice may cause salts to precipitate.
8. Vortexing the samples in a microcentrifuge tube with the lid open can lead to a sample spill. So the optimal vortex speed can be first determined using a “dummy” sample.
9. Mix the beads well by rotating the original container. For ease of handling and to avoid mechanical damage, pipette the beads with wide bore tips that are available commercially or can simply be generated by cutting off the end of the regular pipette tip.
10. We typically use 20 μL of beads that may be increased depending on the expression level of the target protein. Using < 20 μL beads can present visualization difficulties. The pellet of beads can be better visualized against a light lamp or dark background.
11. For IP of biotinylated proteins, use Streptavidin-Agarose (Sigma, cat. no. S1638) or Streptavidin Sepharose High Performance (GE, cat. no. 17-5113-01). For IP of Strep-tag tagged Histones, Strep-Tactin XT Superflow (IBA, cat. no. 2-4010) can be used. Alternatively, Sepharose or Magnetic Protein G/A beads along with specific antibodies against a histone protein or PTM may be used (see Note 12 for further details). The choice of Protein G or A beads depends on the affinity of the beaded support with the isotype of the antibody being used. For further information refer to Zhang and Chen (2001) and Bonifacino et al. (2016).
12. If histone- or histone PTM-specific antibodies are used, particular attention should be paid to the specificity of each antibody used and they should all be carefully validated (Bordeaux et al., 2010; Laflamme et al., 2019). Histone PTM antibodies are generally raised using modified peptides but the presence of other PTMs on endogenous histones may interfere with the detection by the antibody (Egelhofer et al., 2011; Rothbart et al., 2015). Consultation with histone antibody databases such as <http://www.histoneantibodies.com/>

or <https://www.encodeproject.org/antibodies/> can offer insights into possible occlusions of the targeted epitopes by surrounding modifications for tested commercial antibodies. Another issue to be considered is that antibodies need to be checked for cross-reactivity with other proteins or histones that have identical or similar motifs (for example, the ARKS amino acid motif flanking K9 is identical to the sequence around K27 of histone H3). Testing the loss of reactivity of histone PTM antibodies toward histones where the specific sites of modification are mutated could be useful for confirming their specificities. Lastly, as this protocol is similar to native-ChIP, antibodies suitable for native ChIP/ChIP-seq may be a good starting choice for selecting antibodies.

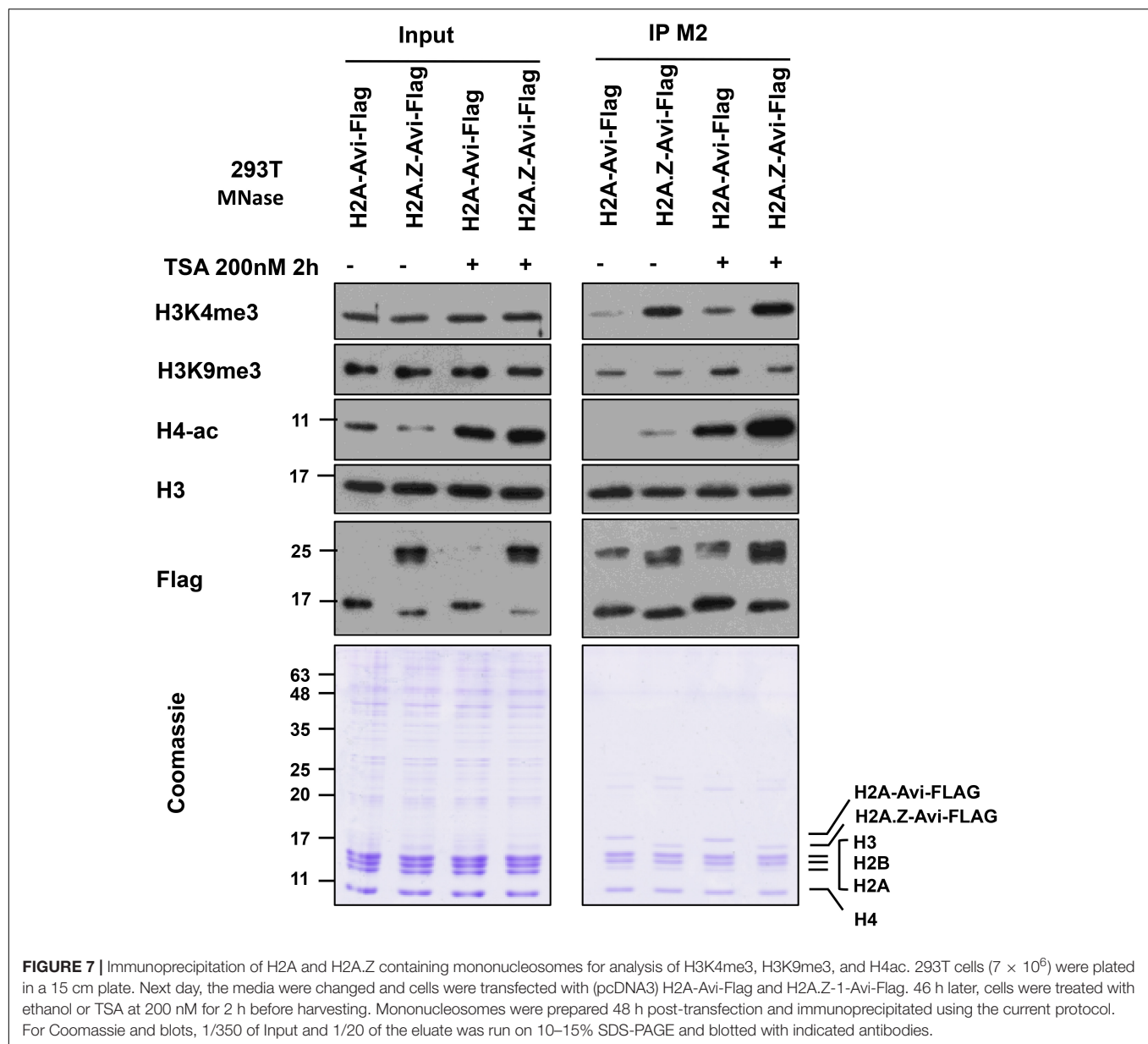
13. Avoid losing any beads during washing step by careful aspiration and by not disturbing the pelleted beads. The supernatant can be aspirated with 200 μL micropipette tips attached to a vacuum line of an aspirator. The tip should be advanced along the side of the tube until it reaches just above the beads. After the last wash, pipette the last few μL of wash buffer using a narrow-bore flat tip, if available.
14. Mix the beads with elution buffer by gently tapping the tube and not by pipetting up and down to avoid losing the beads. It is unnecessary to remove the eluate from the beads and can be stored along with beads at -20°C .

TYPICAL RESULTS

An example of mononucleosome IP analysis for H3 PTMs co-existing with histone variant H2A.Z is shown in **Figure 7**. We expressed H2A-Avi-Flag or H2A.Z-Avi-Flag in transfected 293T cells followed by mononucleosome preparation. H2A and H2A.Z containing mononucleosomes were immunoprecipitated with FLAG-M2 resin using the current protocol. The eluate was normalized by the amount of H3 and blotted for different histone PTMs. As seen before (Draker et al., 2012), we observed that H3K4me3 and H4-ac were enriched on H2A.Z containing mononucleosomes while H3K9me3 was enriched on H2A containing mononucleosomes. This difference was also observed in hyperacetylated histone condition following treatment of cells with the histone deacetylase inhibitor trichostatin A (TSA). As TSA results in the overall acetylation of histones, we observed an increase in acetylation of histone H4 following TSA treatment (**Figure 7**).

TIME CONSIDERATIONS

Once the cells expressing FLAG-tagged histone are ready, isolation of nuclei takes 1 h followed by the preparation of mononucleosomes in 1.5 h. Generally, we perform the IP the same day, which can be incubated for 1 h or until the next day (i.e., overnight). Washing of beads and elution may take 1 h the next day. So mononucleosome IP is typically completed over 2 days. The eluate may be analyzed immediately (e.g., resolved by SDS-PAGE) or stored at -20°C .



DISCUSSION

We have used this mononucleosome IP protocol to isolate the histone variant H2A.Z-containing nucleosomes from human cells and performed proteomic screen with these nucleosomes to identify proteins that co-purify with H2A.Z-1 nucleosomes (Draker et al., 2012). Some of our validated hits (e.g., Brd2, PWWP2A, PHF14) were also identified in other independent mononucleosome IP screens (Kim et al., 2013; Vardabasso et al., 2015; Surface et al., 2016; Punzeler et al., 2017). However, we note that there are variations in the interactors identified by these different studies, possibly due to the differences in the protocol and/or cell line used. For that reason, we characterized the relative abundance of selected histone PTMs and chromatin-binding proteins and presented the data as well as a detailed

protocol in this article. The use of mononucleosome IP has been gaining popularity and has been used to identify proteins interacting with other histone variant-containing nucleosomes such as macroH2A and H2A.Bbd (Sansoni et al., 2014; Sun et al., 2018) as well as CENP-A/cenH3 and H3.Y (Fang et al., 2015; Zink et al., 2017). Apart from histone variants, this approach was also used to demonstrate the YEATS domain as a crotonyllysine reader by showing a preference of AF9 YEATS domain for crotonylated over acetylated lysines in histone H3 (Li et al., 2016). In addition, we and others have used this method to examine the relative abundance of histone PTMs on core or variant histones in the purified mononucleosomes. This approach allows the identification of PTMs on different histones that are co-enriched within the nucleosome context as well as possible cross-talks amongst different PTM combinations.

For example, it has been demonstrated that H2A.Z and H3.3 containing mononucleosomes are enriched for activating histone PTMs such as K4 methylated H3 (Viens et al., 2006; Sarcinella et al., 2007; Won et al., 2015), whereas macroH2A containing mononucleosomes are enriched for repressive histone PTMs like K9 methylated H3 (Won et al., 2015). In addition, this method has also been used to dissect the cross-talk between macroH2A1 and H2B acetylation (Chen et al., 2014; Ruiz and Gamble, 2018), and H2A.Z ubiquitylation with H3K27me3 (Ku et al., 2012; Surface et al., 2016; Ng et al., 2019). Lastly, the incorporation of specific core or variant histones into chromatin has been demonstrated by the nucleosome-immunoprecipitation method (Kanda et al., 1998; Wiedemann et al., 2010; Lau et al., 2011; Ruiz and Gamble, 2018). Altogether, these studies have shown a general utility of this method.

One other area in which mononucleosome immunoprecipitation has been particularly useful is the study of oncohistones i.e., mutations at K27, and K36 in histone H3 and its variants that have been linked to oncogenesis. In fact, expression of H3.3 K27M, and K36M mutants results in global loss of methylation status of endogenous H3 at the corresponding lysine (Mohammad and Helin, 2017). Additional studies used mononucleosome immunoprecipitation to demonstrate that H3.3K27M mutant containing mononucleosomes have reduced H3K27me3 and increased H3K27ac on the wild type endogenous H3 partner, and are enriched for the H3K27 methyltransferase EZH2 (Bender et al., 2013; Chan et al., 2013; Lewis et al., 2013; Fang et al., 2017). Indeed, loss of H3 methylation results in the formation of H3K27M-K27ac heterotypic nucleosomes that are enriched with acetyl-binding bromodomain-containing protein 1 (BRD1) and BRD4 proteins (Herz et al., 2014; Piunti et al., 2017). Similarly, H3.3K9M mutant containing mononucleosomes were observed to have decreased K9 methylation at endogenous H3, and increased association with the H3K9 demethylase KDM3B and the H3K9/K56 deacetylase SIRT6 (Herz et al., 2014). Subsequently, it was also demonstrated that H3.3K36M mutant containing mononucleosomes displayed reduced H3K36me2/3 on endogenous H3 and are enriched for H3K36 methyltransferases SetD2, Nsd1, Nsd2, and MMSET (Fang et al., 2016; Lu et al., 2016).

An advantage of this protocol is that it can potentially identify proteins that interact with nucleosomes through multiple PTMs on different histones of the same nucleosome. These proteins may not otherwise be identified using the standard co-IP procedure. Brd2 has been shown to interact with H2A.Z-1 and H2A.Z-2 nucleosomes through a combinatorial interaction with H2A.Z and acetylated H4 (Draker et al., 2012; Vardabasso et al., 2015). Conversely, ubiquitylation of H2A.Z was found to antagonize Brd2 binding to H2A.Z nucleosomes (Surface et al., 2016). Such an approach has also been used to confirm the multivalent interaction of PWWP2A with H2A.Z nucleosomes (Punzeler et al., 2017). Lastly, IP of mononucleosomes was used to demonstrate the existence of asymmetrically modified nucleosomes in addition to symmetrically modified nucleosomes, with respect to H3K27me2/3 and H4K20me1 (Voigt et al., 2012).

As with any experimental approaches, there are limitations and drawbacks to this mononucleosome IP method. First,

we rely on the expression of tagged histones (e.g., Flag-tagged H2A.Z) in transfected cells in order to use the highly efficient Flag M2 beads to capture the tagged histone-containing nucleosomes. Immunoprecipitation of endogenous histones is always desirable, and in theory this protocol should also work with antibodies against specific histones or histone PTMs. However, as mentioned previously, it depends on the immunoprecipitation efficiencies of the antibodies against the intended targets, which will have to be individually tested and optimized. For our purposes, we chose to express tagged histones because of the advantage of well-characterized and highly efficient antibodies against epitope tags. In our hands, the expression levels of the tagged H2A.Z is fairly comparable to the endogenous H2A.Z protein levels (data not shown), and we have not observed any interference of histone incorporation, nor growth defects or deleterious effects in these tagged H2A.Z-expressing cells. We also typically choose to place the epitope tag on the C-terminus of histones, but it is possible that, in some cases, the presence of a tag may interfere with normal PTM profile and thus the association of histone binding proteins. If necessary, one could test for potential interference effects by comparing results using N-terminus vs. C-terminus tagged histones. All these parameters should be carefully monitored for different histones, histone variants, and tags. Second, as noted earlier, it would be highly desirable to immunoprecipitate nucleosomes using histone-PTM-specific antibodies; however, these are not always available (see note 12 for considerations for choosing the right antibody for the assay). To overcome this limitation, we have previously developed a technique named BICON (biotinylation-assisted isolation of co-modified nucleosomes) whereby we co-modify targeted histone substrates with a fusion of a histone-modifying enzyme and the *E. coli* biotin ligase BirA, and then use streptavidin-couple reagents to pull down the biotinylated and enzyme-modified histone/nucleosome [see (Lau and Cheung, 2013) for specific details]. As proof of principle, we co-expressed H3.3 with a 15 amino acid tag called Avi-tag that is the specific biotinylation target of BirA, along with an MSK1-BirA fusion (MSK1 is a well-characterized H3 kinase) in human cells, and demonstrated efficient purification of nucleosomes that contain H3.3 that are both biotinylated and phosphorylated.

In this protocol, we rely on micrococcal nuclease digestion to release mono-nucleosomes from bulk chromatin. We have chosen an MNase concentration that generates mostly mono-nucleosomes without excessive digestion, but we have not tested whether higher amounts of MNase could release more nucleosomes and chromatin-binding proteins from the pellet fraction. Indeed, chromatin is sensitive to MNase at varying degrees. Although the precise reasons for differential sensitivity of individual nucleosome to MNase are unknown, preference of MNase toward adenine/thymine (A/T) nucleotides may be a contributing factor. It has been reported that over-digestion can result in artificial depletion of A/T rich genomic regions (Kensche et al., 2016). As shown in **Figure 5** of this article, there are significant amounts of chromatin-bound proteins, as well as some nucleosomes, that remain in the insoluble pellet after MNase digestion and hence are not used for immunoprecipitation. Moreover, some proteins

(e.g., HP1 α , Fibrillarin) and modified histones (e.g., H2Bub) appear to be almost exclusively found in the insoluble pellet. Therefore, this mono-nucleosome immunoprecipitation approach will not be suitable for examining nuclear factors or histones that are exclusively retained in the insoluble pellet. The precise explanation of these observations is not known, but there could be MNase-resistant or insoluble fractions of chromatin that are enriched with specific types of chromatin. Along that line of thinking, a previous study discovered and characterized a sonication-resistant fraction of chromatin in their ChIP-seq method that is enriched for proteins/genomic sequences associated with unique subtypes of heterochromatin, and is refractory to cell reprogramming (Becker et al., 2017). Therefore, similar genomic/proteomic analyses of the pellet fraction from our fractionation protocol may reveal new insights into subtypes of chromatin domains and compartments as well.

Materials

Reagents

1 kb Plus DNA ladder (Thermo Fisher Scientific, cat. no. 10787-018)
 293T cells cultured on 145-mm plates (expressing FLAG-tagged Histone protein)
 6X Gel Loading Dye, Purple (NEB, cat. no. B7024)
 Agarose (Fisher BioReagents, cat. no. BP160)
 Anti-Flag M2 Affinity Gel (Sigma-Aldrich, cat. no. A2220)
 Aprotinin (Sigma-Aldrich, cat. no. A1153)
 CaCl₂ solution (Sigma-Aldrich, cat. no. 442909)
 Chloroform (Caledon, cat. no. 3001-2)
 Dithiothreitol (DTT; Fisher BioReagents, cat. no. BP172)
 DMEM Media, high glucose (HyClone, cat. no. SH30243FS)
 EDTA (Sigma-Aldrich, cat. no. EDS)
 EGTA (Sigma-Aldrich, cat. no. E4378)
 Equilibrated Phenol solution (Sigma-Aldrich, cat. no. P4557)
 Ethanol (Commercial Alcohols, cat. no. P006EAAN)
 Fetal Bovine Serum (Sigma-Aldrich, cat. no. F1051)
 Glycerol (Fisher BioReagents, cat. no. BP229)
 HEPES (Fisher BioReagents, cat. no. BP310)
 Isoamyl alcohol (Fisher BioReagents, cat. no. BP1150)
 KCl (Sigma-Aldrich, cat. no. P9541)
 Leupeptin (Sigma-Aldrich, cat. no. L2884)
 MgCl₂ (Sigma-Aldrich, cat. no. M9272)
 Micrococcal Nuclease (Worthington Biochem, cat. no. LS004798)
 Microcystin-LR (Cayman, cat. no. 10007188)
 NaCl (Fisher BioReagents, cat. no. BP358)
 N-ethylmaleimide (NEM; Sigma-Aldrich, cat. no. E3876)
 Pepstatin (Sigma-Aldrich, cat. no. P5318)
 Phenylmethylsulfonyl fluoride (PMSF; Sigma-Aldrich, cat. no. P7626)

Phosphate buffered saline (PBS; Wisent, cat. no. 311-010-CL)
 Phenol:chloroform:isoamyl alcohol (UltraPuro, Thermo Fisher Scientific, cat. no. 15593-031)
 Chloroform/Isoamyl alcohol 24:1 (CI)
 Proteinase K (Sigma-Aldrich, cat. no. P2308)
 RedSafe Nucleic Acid Staining Solution (iNtRON Biotechnology, cat. no. 21141)
 Sodium acetate (Sigma-Aldrich, cat. no. S2889)
 Sodium Butyrate (Sigma-Aldrich, cat. no. B5887)
 Sucrose (Bioshop, cat. no. SUC700)
 Tris (Fisher BioReagents, cat. no. BP152)
 Triton X100 (Sigma-Aldrich, cat. no. T9284)
 Trypsin 0.05%, EDTA 0.53MM in HBSS, 1X (Wisent, cat. no. 325-542)

Antibodies

Anti-FLAG-M2 affinity gel (Sigma-Aldrich, cat. no. A2220)
 Anti-PHF6 (Bethyl, cat. no. A301-450A)
 Anti-USP39 (Abcam, cat. no. ab131332)
 Anti-USP7 (Bethyl, cat. no. A300-033A)
 Anti-Brd2 (Abcam, cat. no. ab3718)
 Anti-RNA PolII (Covance, cat. no. MMS-126-R)
 Anti-HP1 α (Santa-Cruz, cat. no. sc-47701)
 Anti-Fibrillarin (Abcam, cat. no. ab5821)
 Anti-Tubulin (Sigma-Aldrich, cat. no. T6074)
 Anti-H3 (Abcam, cat. no. ab1791)
 Anti-H2A.Z (Active Motif, cat. no. 19113)
 Anti-H3K4me3 (Millipore, cat. no. 07-030)
 Anti-H3K9me3 (Millipore, cat. no. 07-442)
 Anti-H3K27me3 (Millipore, cat. no. 07-449)
 Anti-H3K36me3 (Abcam, cat. no. ab9050)
 Anti-H3K79me2 (Millipore, cat. no. ABE459)
 Anti-ub-H2A K119 (Cell Signaling, cat. no. 8240)
 Anti-ub-H2B (Medi Mabs, cat. no. MM-0029)
 Anti-H4-ac (Millipore, cat. no. 06-946)

Equipment

Vacuum aspirator
 Ice and ice bucket
 End-to-end rotator
 1.5-ml microcentrifuge tubes
 15-ml and 50-ml propylene tubes
 Agarose gel apparatus
 Electrophoresis power supply
 Micropipettes (e.g., Pipetman, Gilson)
 Refrigerated microcentrifuge (e.g., Eppendorf 5415R)
 SDS-PAGE running and transfer apparatus
 Tissue culture dish, 145 mm (Grenier Bio-One, cat. no. 639160)
 UV Spectrophotometer (e.g., NanoDrop 2000, Thermo Scientific)
 Water bath set at 37°C.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

PC conceived the original idea and supervised the project. MN performed initial optimization of the protocol and KK performed the experimental work presented in this

manuscript. KK wrote the first draft followed by revision with contributions from MN and PC.

FUNDING

This research was funded in part by Canadian Cancer Society Research Institute (CCSRI Innovation Grant number 704315) and the Canadian Institutes of Health Research (CIHR Operating Grant number RN227427 – 324983) awarded to PC.

REFERENCES

- Al-Natour, Z., and Hassan, A. H. (2007). Effect of salt on the binding of the linker histone H1 to DNA and nucleosomes. *DNA Cell Biol.* 26, 445–452. doi: 10.1089/dna.2006.0512
- Becker, J. S., McCarthy, R. L., Sidoli, S., Donahue, G., Kaeding, K. E., He, Z., et al. (2017). Genomic and proteomic resolution of heterochromatin and its restriction of alternate fate genes. *Mol. Cell* 68, 1023–1037.e15. doi: 10.1016/j.molcel.2017.11.030
- Bender, S., Tang, Y., Lindroth, A. M., Hovestadt, V., Jones, D. T., Kool, M., et al. (2013). Reduced H3K27me3 and DNA hypomethylation are major drivers of gene expression in K27M mutant pediatric high-grade gliomas. *Cancer Cell* 24, 660–672. doi: 10.1016/j.ccr.2013.10.006
- Bonifacino, J. S., Gershlick, D. C., and Dell'Angelica, E. C. (2016). Immunoprecipitation. *Curr. Protoc. Cell Biol.* 71, 7.2.1–7.2.24. doi: 10.1002/cpcb.3
- Bordeaux, J., Welsh, A., Agarwal, S., Killiam, E., Baquero, M., Hanna, J., et al. (2010). Antibody validation. *Biotechniques* 48, 197–209. doi: 10.2144/000113382
- Chan, K. M., Fang, D., Gan, H., Hashizume, R., Yu, C., Schroeder, M., et al. (2013). The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. *Genes Dev.* 27, 985–990. doi: 10.1101/gad.217778.113
- Chen, H., Ruiz, P. D., Novikov, L., Casill, A. D., Park, J. W., and Gamble, M. J. (2014). MacroH2A1.1 and PARP-1 cooperate to regulate transcription by promoting CBP-mediated H2B acetylation. *Nat. Struct. Mol. Biol.* 21, 981–989. doi: 10.1038/nsmb.2903
- Clark, D. J. (2010). Nucleosome positioning, nucleosome spacing and the nucleosome code. *J. Biomol. Struct. Dyn.* 27, 781–793. doi: 10.1080/073911010010524945
- Clark, D. J., and Thomas, J. O. (1986). Salt-dependent co-operative interaction of histone H1 with linear DNA. *J. Mol. Biol.* 187, 569–580. doi: 10.1016/0022-2836(86)90335-9
- Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* 11, 1475–1489. doi: 10.1093/nar/11.5.1475
- Draker, R., Ng, M. K., Sarcinella, E., Ignatchenko, V., Kislinger, T., and Cheung, P. (2012). A combination of H2A.Z and H4 acetylation recruits Brd2 to chromatin during transcriptional activation. *PLoS Genet.* 8:e1003047. doi: 10.1371/journal.pgen.1003047
- Dyer, P. N., Edayathumangalam, R. S., White, C. L., Bao, Y., Chakravarthy, S., Muthurajan, U. M., et al. (2004). Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* 375, 23–44. doi: 10.1016/s0076-6879(03)75002-2
- Egelhofer, T. A., Minoda, A., Klugman, S., Lee, K., Kolasinska-Zwierz, P., Alekseyenko, A. A., et al. (2011). An assessment of histone-modification antibody quality. *Nat. Struct. Mol. Biol.* 18, 91–93. doi: 10.1038/nsmb.1972
- Fang, D., Gan, H., Cheng, L., Lee, J.-H., Zhou, H., Sarkaria, J. N., et al. (2017). H3K27me3-mediated silencing of Wilms Tumor 1 supports the proliferation of brain tumor cells harboring the H3.3K27M mutation. *bioRxiv* [Preprint]. doi: 10.1101/114082
- Fang, D., Gan, H., Lee, J. H., Han, J., Wang, Z., Riester, S. M., et al. (2016). The histone H3.3K36M mutation reprograms the epigenome of chondroblastomas. *Science* 352, 1344–1348. doi: 10.1126/science.aae0065
- Fang, J., Liu, Y., Wei, Y., Deng, W., Yu, Z., Huang, L., et al. (2015). Structural transitions of centromeric chromatin regulate the cell cycle-dependent recruitment of CENP-N. *Genes Dev.* 29, 1058–1073. doi: 10.1101/gad.259432.115
- Foltz, D. R., Jansen, L. E., Black, B. E., Bailey, A. O., Yates, J. R. III, and Cleveland, D. W. (2006). The human CENP-A centromeric nucleosome-associated complex. *Nat. Cell Biol.* 8, 458–469. doi: 10.1038/ncb1397
- Fujimoto, S., Seebart, C., Guastafierro, T., Prenni, J., Caiafa, P., and Zlatanova, J. (2012). Proteome analysis of protein partners to nucleosomes containing canonical H2A or the variant histones H2A.Z or H2A.X. *Biol. Chem.* 393, 47–61. doi: 10.1515/BC-2011-216
- Henikoff, S., and Smith, M. M. (2015). Histone variants and epigenetics. *Cold Spring Harb. Perspect. Biol.* 7:a019364. doi: 10.1101/cshperspect.a019364
- Herz, H. M., Morgan, M., Gao, X., Jackson, J., Rickels, R., Swanson, S. K., et al. (2014). Histone H3 lysine-to-methionine mutants as a paradigm to study chromatin signaling. *Science* 345, 1065–1070. doi: 10.1126/science.1255104
- Kanda, T., Sullivan, K. F., and Wahl, G. M. (1998). Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. *Curr. Biol.* 8, 377–385. doi: 10.1016/s0960-9822(98)70156-3
- Kensche, P. R., Hoeijmakers, W. A., Toenhake, C. G., Bras, M., Chappell, L., Berriman, M., et al. (2016). The nucleosome landscape of Plasmodium falciparum reveals chromatin architecture and dynamics of regulatory sequences. *Nucleic Acids Res.* 44, 2110–2124. doi: 10.1093/nar/gkv1214
- Kim, K., Punj, V., Choi, J., Heo, K., Kim, J. M., Laird, P. W., et al. (2013). Gene dysregulation by histone variant H2A.Z in bladder cancer. *Epigenetics Chromatin* 6:34. doi: 10.1186/1756-8935-6-34
- Kim, K., Shin, Y., Kim, J., Ulmer, T. S., and An, W. (2018). H3K27me1 is essential for MMP-9-dependent H3N-terminal tail proteolysis during osteoclastogenesis. *Epigenetics Chromatin* 11:23. doi: 10.1186/s13072-018-0193-1
- Kornberg, R. D. (1977). Structure of chromatin. *Annu. Rev. Biochem.* 46, 931–954. doi: 10.1146/annurev.bi.46.070177.004435
- Krishnaswami, S. R., Grindberg, R. V., Novotny, M., Venepally, P., Lacar, B., Bhutani, K., et al. (2016). Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat. Protoc.* 11, 499–524. doi: 10.1038/nprot.2016.015
- Ku, M., Jaffe, J. D., Koche, R. P., Rheinbay, E., Endoh, M., Koseki, H., et al. (2012). H2A.Z landscapes and dual modifications in pluripotent and multipotent stem cells underlie complex genome regulatory functions. *Genome Biol.* 13:R85. doi: 10.1186/gb-2012-13-10-r85
- Lacoste, N., Woolfe, A., Tachiwana, H., Gareia, A. V., Barth, T., Cantaloube, S., et al. (2014). Mislocalization of the centromeric histone variant CenH3/CENP-A in human cells depends on the chaperone DAXX. *Mol. Cell* 53, 631–644. doi: 10.1016/j.molcel.2014.01.018
- Laflamme, C., McKeever, P. M., Kumar, R., Schwartz, J., Kolahdouzan, M., Chen, C. X., et al. (2019). Implementation of an antibody characterization procedure and application to the major ALS/FTD disease gene C9ORF72. *eLife* 8:e48363. doi: 10.7554/eLife.48363
- Lai, W. K. M., and Pugh, B. F. (2017). Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat. Rev. Mol. Cell Biol.* 18, 548–562. doi: 10.1038/nrm.2017.47

- Lau, A. T., Lee, S. Y., Xu, Y. M., Zheng, D., Cho, Y. Y., Zhu, F., et al. (2011). Phosphorylation of histone H2B serine 32 is linked to cell transformation. *J. Biol. Chem.* 286, 26628–26637. doi: 10.1074/jbc.M110.215590
- Lau, P. N., and Cheung, P. (2013). Elucidating combinatorial histone modifications and crosstalks by coupling histone-modifying enzyme with biotin ligase activity. *Nucleic Acids Res.* 41:e49. doi: 10.1093/nar/gks1247
- Law, C., and Cheung, P. (2013). Histone variants and transcription regulation. *Subcell. Biochem.* 61, 319–341. doi: 10.1007/978-94-007-4525-4_14
- Lee, J. S., Smith, E., and Shilatifard, A. (2010). The language of histone crosstalk. *Cell* 142, 682–685. doi: 10.1016/j.cell.2010.08.011
- Lewis, P. W., Muller, M. M., Koletsky, M. S., Cordero, F., Lin, S., Banaszynski, L. A., et al. (2013). Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science* 340, 857–861. doi: 10.1126/science.1232245
- Li, Y., Sabari, B. R., Panchenko, T., Wen, H., Zhao, D., Guan, H., et al. (2016). Molecular coupling of histone crotonylation and active transcription by AF9 YEATS domain. *Mol. Cell* 62, 181–193. doi: 10.1016/j.molcel.2016.03.028
- Long, M., Sun, X., Shi, W., Yanru, A., Leung, S. T. C., Ding, D., et al. (2019). A novel histone H4 variant H4G regulates rDNA transcription in breast cancer. *Nucleic Acids Res.* 47, 8399–8409. doi: 10.1093/nar/gkz547
- Lowary, P. T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* 276, 19–42. doi: 10.1006/jmbi.1997.1494
- Lu, C., Jain, S. U., Hoelper, D., Bechet, D., Molden, R. C., Ran, L., et al. (2016). Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. *Science* 352, 844–849. doi: 10.1126/science.aac7272
- Luger, K., Mader, 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. doi: 10.1038/38444
- Lusser, A., and Kadonaga, J. T. (2004). Strategies for the reconstitution of chromatin. *Nat. Methods* 1, 19–26. doi: 10.1038/nmeth709
- Maze, I., Noh, K. M., Soshnev, A. A., and Allis, C. D. (2014). Every amino acid matters: essential contributions of histone variants to mammalian development and disease. *Nat. Rev. Genet.* 15, 259–271. doi: 10.1038/nrg3673
- Mendez, J., and Stillman, B. (2000). Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. *Mol. Cell Biol.* 20, 8602–8612. doi: 10.1128/mcb.20.22.8602-8612.2000
- Mohammad, F., and Helin, K. (2017). Oncohistones: drivers of pediatric cancers. *Genes Dev.* 31, 2313–2324. doi: 10.1101/gad.309013.117
- Muller, M. M., and Muir, T. W. (2015). Histones: at the crossroads of peptide and protein chemistry. *Chem. Rev.* 115, 2296–2349. doi: 10.1021/cr5003529
- Nadal, S., Raj, R., Mohammed, S., and Davis, B. G. (2018). Synthetic post-translational modification of histones. *Curr. Opin. Chem. Biol.* 45, 35–47. doi: 10.1016/j.cbpa.2018.02.004
- Ng, M. K., Braunschweig, U., Blencowe, B. J., and Cheung, P. (2019). Ubiquitylated H2A.Z nucleosomes are associated with nuclear architectural proteins and global transcriptional silencing. *bioRxiv* [Preprint]. doi: 10.1101/759852
- Ng, M. K., and Cheung, P. (2016). A brief histone in time: understanding the combinatorial functions of histone PTMs in the nucleosome context. *Biochem. Cell Biol.* 94, 33–42. doi: 10.1139/bcb-2015-0031
- Nikolov, M., Stutzer, A., Mosch, K., Krasauskas, A., Soeroes, S., Stark, H., et al. (2011). Chromatin affinity purification and quantitative mass spectrometry defining the interactome of histone modification patterns. *Mol. Cell. Proteomics* 10:M110005371. doi: 10.1074/mcp.M110.005371
- Piunti, A., Hashizume, R., Morgan, M. A., Bartom, E. T., Horbinski, C. M., Marshall, S. A., et al. (2017). Therapeutic targeting of polycomb and BET bromodomain proteins in diffuse intrinsic pontine gliomas. *Nat. Med.* 23, 493–500. doi: 10.1038/nm.4296
- Punzeler, S., Link, S., Wagner, G., Keilhauer, E. C., Kronbeck, N., Spitzer, R. M., et al. (2017). Multivalent binding of PWWP2A to H2A.Z regulates mitosis and neural crest differentiation. *EMBO J.* 36, 2263–2279. doi: 10.15252/embj.201695757
- Rothbart, S. B., Dickson, B. M., Raab, J. R., Grzybowski, A. T., Krajewski, K., Guo, A. H., et al. (2015). An interactive database for the assessment of histone antibody specificity. *Mol. Cell* 59, 502–511. doi: 10.1016/j.molcel.2015.06.022
- Ruiz, P. D., and Gamble, M. J. (2018). MacroH2A1 chromatin specification requires its docking domain and acetylation of H2B lysine 20. *Nat. Commun.* 9:5143. doi: 10.1038/s41467-018-07189-8
- Sanson, V., Casas-Delucchi, C. S., Rajan, M., Schmidt, A., Bonisch, C., Thomae, A. W., et al. (2014). The histone variant H2A.Bbd is enriched at sites of DNA synthesis. *Nucleic Acids Res.* 42, 6405–6420. doi: 10.1093/nar/gku303
- Sarcinella, E., Zuzarte, P. C., Lau, P. N., Draker, R., and Cheung, P. (2007). Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin. *Mol. Cell Biol.* 27, 6457–6468. doi: 10.1128/MCB.00241-07
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675.
- Sen, S., Block, K. F., Pasini, A., Baylin, S. B., and Easwaran, H. (2016). Genome-wide positioning of bivalent mononucleosomes. *BMC Med. Genomics* 9:60. doi: 10.1186/s12920-016-0221-6
- Shema, E., Jones, D., Shores, N., Donohue, L., Ram, O., and Bernstein, B. E. (2016). Single-molecule decoding of combinatorially modified nucleosomes. *Science* 352, 717–721. doi: 10.1126/science.aad7701
- Stillman, B. (2018). Histone modifications: insights into their influence on gene expression. *Cell* 175, 6–9. doi: 10.1016/j.cell.2018.08.032
- Strahl, B. D., and Allis, C. D. (2000). The language of covalent histone modifications. *Nature* 403, 41–45. doi: 10.1038/47412
- Sun, Z., Filipescu, D., Andrade, J., Gaspar-Maia, A., Ueberheide, B., and Bernstein, E. (2018). Transcription-associated histone pruning demarcates macroH2A chromatin domains. *Nat. Struct. Mol. Biol.* 25, 958–970. doi: 10.1038/s41594-018-0134-5
- Surface, L. E., Fields, P. A., Subramanian, V., Behmer, R., Udeshi, N., Peach, S. E., et al. (2016). H2A.Z.1 monoubiquitylation antagonizes BRD2 to maintain poised chromatin in ESCs. *Cell Rep.* 14, 1142–1155. doi: 10.1016/j.celrep.2015.12.100
- Talbert, P. B., and Henikoff, S. (2017). Histone variants on the move: substrates for chromatin dynamics. *Nat. Rev. Mol. Cell Biol.* 18, 115–126. doi: 10.1038/nrm.2016.148
- Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D., and Patel, D. J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* 14, 1025–1040. doi: 10.1038/nsmb1338
- Vardabasso, C., Gaspar-Maia, A., Hasson, D., Punzeler, S., Valle-Garcia, D., Straub, T., et al. (2015). Histone variant H2A.Z.2 mediates proliferation and drug sensitivity of malignant melanoma. *Mol. Cell* 59, 75–88. doi: 10.1016/j.molcel.2015.05.009
- Viens, A., Mechold, U., Brouillard, F., Gilbert, C., Leclerc, P., and Ogryzko, V. (2006). Analysis of human histone H2AZ deposition in vivo argues against its direct role in epigenetic templating mechanisms. *Mol. Cell Biol.* 26, 5325–5335. doi: 10.1128/MCB.00584-06
- Voigt, P., LeRoy, G., Drury, W. J. III, Zee, B. M., Son, J., Beck, D. B., et al. (2012). Asymmetrically modified nucleosomes. *Cell* 151, 181–193. doi: 10.1016/j.cell.2012.09.002
- Wang, W. L., Anderson, L. C., Nicklay, J. J., Chen, H., Gamble, M. J., Shabanowitz, J., et al. (2014). Phosphorylation and arginine methylation mark histone H2A prior to deposition during *Xenopus laevis* development. *Epigenetics Chromatin* 7:22. doi: 10.1186/1756-8935-7-22
- Wang, Y., Long, H., Yu, J., Dong, L., Wassef, M., Zhuo, B., et al. (2018). Histone variants H2A.Z and H3.3 coordinately regulate PRC2-dependent H3K27me3 deposition and gene expression regulation in mES cells. *BMC Biol.* 16:107. doi: 10.1186/s12915-018-0568-6
- Wiedemann, S. M., Mildner, S. N., Bonisch, C., Israel, L., Maiser, A., Matheisl, S., et al. (2010). Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y. *J. Cell Biol.* 190, 777–791. doi: 10.1083/jcb.201002043
- Won, K. J., Choi, I., LeRoy, G., Zee, B. M., Sidoli, S., Gonzales-Cope, M., et al. (2015). Proteogenomics analysis reveals specific genomic orientations of distal regulatory regions composed by non-canonical histone variants. *Epigenetics Chromatin* 8:13. doi: 10.1186/s13072-015-0005-9
- Zaret, K. (2005). Micrococcal nuclease analysis of chromatin structure. *Curr. Protoc. Mol. Biol.* 69, 21.1.1–21.1.17. doi: 10.1002/0471142727.mb2101s69

- Zhang, N., and Chen, J. L. (2001). Purification of recombinant proteins and study of protein interaction by epitope tagging. *Curr. Protoc. Mol. Biol.* 41, 10.15.1–10.15.9. doi: 10.1002/0471142727.mb1015s41
- Zhang, Z., Jones, A. E., Wu, W., Kim, J., Kang, Y., Bi, X., et al. (2017). Role of remodeling and spacing factor 1 in histone H2A ubiquitination-mediated gene silencing. *Proc. Natl. Acad. Sci. U.S.A.* 114, E7949–E7958. doi: 10.1073/pnas.1711158114
- Zhao, Y., and Garcia, B. A. (2015). Comprehensive catalog of currently documented histone modifications. *Cold Spring Harb. Perspect. Biol.* 7:a025064. doi: 10.1101/cshperspect.a025064
- Zink, L. M., Delbarre, E., Eberl, H. C., Keilhauer, E. C., Bonisch, C., Punzeler, S., et al. (2017). H3.Y discriminates between HIRA and DAXX chaperone complexes and reveals unexpected insights into human DAXX-H3.3-H4 binding and deposition requirements. *Nucleic Acids Res.* 45, 5691–5706. doi: 10.1093/nar/gkx131
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Copyright © 2020 Khan, Ng and Cheung. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Proximity Labeling Techniques to Study Chromatin

Henning Ummethum and Stephan Hamperl*

Chromosome Dynamics and Genome Stability, Institute of Epigenetics and Stem Cells, Helmholtz Zentrum München, Munich, Germany

OPEN ACCESS

Edited by:

Jean-Philippe Lambert,
Laval University, Canada

Reviewed by:

Samuel A. Myers,
Broad Institute, United States
Etienne Coyaude,
Université de Lille, France

*Correspondence:

Stephan Hamperl
stephan.hamperl@
helmholtz-muenchen.de

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Genetics

Received: 21 February 2020

Accepted: 14 April 2020

Published: 12 May 2020

Citation:

Ummethum H and Hamperl S
(2020) Proximity Labeling Techniques
to Study Chromatin.
Front. Genet. 11:450.
doi: 10.3389/fgene.2020.00450

Mammals contain over 200 different cell types, yet nearly all have the same genomic DNA sequence. It is a key question in biology how the genetic instructions in DNA are selectively interpreted by cells to specify various transcriptional programs and therefore cellular identity. The structural and functional organization of chromatin governs the transcriptional state of individual genes. To understand how genomic loci adopt different levels of gene expression, it is critical to characterize all local chromatin factors as well as long-range interactions in the 3D nuclear compartment. Much of our current knowledge regarding protein interactions in a chromatin context is based on affinity purification of chromatin components coupled to mass spectrometry (AP-MS). AP-MS has been invaluable to map strong protein-protein interactions in the nucleus. However, the interaction is detected after cell lysis and biochemical enrichment, allowing for loss or gain of false positive or negative interaction partners. Recently, proximity-dependent labeling methods have emerged as powerful tools for studying chromatin in its native context. These methods take advantage of engineered enzymes that are fused to a chromatin factor of interest and can directly label all factors in proximity. Subsequent pull-down assays followed by mass spectrometry or sequencing approaches provide a comprehensive snapshot of the proximal chromatin interactome. By combining this method with dCas9, this approach can also be extended to study chromatin at specific genomic loci. Here, we review and compare current proximity-labeling approaches available for studying chromatin, with a particular focus on new emerging technologies that can provide important insights into the transcriptional and chromatin interaction networks essential for cellular identity.

Keywords: protein-protein interactions, proxisome, BioID, APEX2, dCas9, ChIP, affinity purification, mass spectrometry

INTRODUCTION

A long-standing question in cell biology is how the same genome can lead to different cell types. A major driving force that determines cellular identity is their underlying gene expression landscape. Whether a gene is turned on or off depends mostly on its physical accessibility, governed by the local chromatin context (Klemm et al., 2019). The basic unit of chromatin is the nucleosome core particle, a protein-DNA complex consisting of 146 bp of DNA wrapped around a histone octamer. Histones play a central role in DNA accessibility, due to histone variants and a multitude of post-translational modifications (PTMs) that influence binding of secondary chromatin factors. This can lead to further compaction and heterochromatin formation, restricting or completely

blocking access for the transcription machinery. Other factors influencing chromatin structure are DNA methylation, long non-coding RNAs and chromatin remodelers. However, the full extent of chromatin modifications and complex interactions of a given gene in the complex nuclear environment are poorly understood. Therefore, it is crucial to identify all factors that are part of this process by studying protein-protein interactions of known chromatin factors.

The most widely applied methods to study protein-protein interactions in a chromatin context are affinity purification or immunoprecipitation followed by mass spectrometry (AP-MS/IP-MS). After cell lysis, soluble proteins are captured and enriched by a ligand (bait) coupled to a solid support. The most commonly used ligands are antibodies targeting epitope-tagged (AP) or endogenous (IP) proteins (prey). These ligands are attached to a solid support, in most cases agarose, sepharose or magnetic beads. After enrichment, proteins are analyzed by mass spectrometry to identify proteins interacting with the prey. Consequently, successful application of IP-MS depends on the availability of an antibody or known interactor of the protein of interest. For AP-MS, common antibodies can be used, because an epitope tag is fused to the protein of interest. While these methods can efficiently identify strong protein-protein interactions that are not disrupted after cell lysis and solubilization, transient interactors with lower affinity can be lost during the purification steps – typically performed under high-salt and detergent conditions. In addition, interactions are only detected after lysis and enrichment and thus no longer in its native environment of living cells.

Proximity labeling followed by mass spectrometry analysis can address these key limitations of AP- and IP-MS. The basic principle of all proximity labeling methods is to introduce a covalent biotin tag to proteins in the neighborhood of a selected target in living cells. To this end, enzymes convert a supplemented substrate into a highly reactive biotinylated intermediate that then transfers biotin to amino acid side chains in proximity. Spatial restriction of labeling is achieved by fusing the enzyme to the target protein as well as reducing the labeling time. Currently, three major enzymes are used for proximity labeling: biotin ligase (BioID, BioID2, TurboID, miniTurbo), horse radish peroxidase (HRP), and engineered ascorbate peroxidase (APEX, APEX2). After the labeling reaction, cells are lysed and the biotinylated proteins are extracted with streptavidin beads and subjected to mass spectrometry. Identified candidates in proximity with the bait protein can be summarized as the “proxisome” (Roux et al., 2012).

The advantages of proximity labeling in comparison to conventional methods to study chromatin are manifold. One major benefit is the ability to analyze protein interactions in a native context, because covalent biotinylation occurs before cell lysis and solubilization. As streptavidin-biotin is one of the strongest non-covalent interactions found in nature, harsh conditions can be used to force insoluble proteins into solution without the constraint of maintaining protein-protein interactions during the purification process. Therefore, proximity labeling enables the study of proxisomes even in insoluble cell compartments like the nuclear matrix, nucleoli and other

nuclear structures – difficult to study with conventional methods. Additionally, *in vivo* covalent biotinylation enables the detection of transient interactions and low abundance proteins. Finally, biotinylation is an infrequent protein modification in many organisms, thus no additional endogenous proteins are part of the background in mass spectrometry analysis (de Boer et al., 2003).

Here, we will review and compare current proximity labeling approaches available for studying chromatin, with a particular focus on new emerging technologies that can provide important insights into the transcriptional and chromatin interaction networks from specific gene loci to whole genome interactions in nuclear compartments.

PROXIMITY LABELING METHODS

Biotin Ligase (BioID)

The *Escherichia coli* BirA biotin ligase converts biotin and ATP into biotinoyl-5'-adenylate (bioAMP) (Barker and Campbell, 1981a,b; Eisenberg et al., 1982). One of the physiological roles of the BirA-bioAMP complex is to target the only biotinylation site in *E. coli*, a lysine residue in the biotin carboxyl carrier protein (BCCP) subunit of acetyl-CoA carboxylase. To take advantage of this highly specific reaction, an unnatural substrate mimicking a short peptide sequence was created (Schatz, 1993; Beckett et al., 1999). This biotin acceptor peptide (BAP) can be fused to proteins of interest (POI) and co-expressed with BirA, which in turn recognizes and conjugates biotin on the lysine of BAP (Smith et al., 1998). The newly biotinylated protein can be efficiently purified by streptavidin pull-down (de Boer et al., 2003). In a different approach, this system was used to study protein-protein interactions by fusing BirA and BAP to two interacting proteins (Fernández-Suárez et al., 2008). However, interacting protein pairs must be known *a priori*.

A mutated BirA* (R118G) from *E. coli* made an unbiased approach possible by disrupting binding of bioAMP to BirA (Kwon and Beckett, 2000; Kwon et al., 2000). Consequently, bioAMP diffuses from the enzyme and can readily react with lysine residues of any protein. Interestingly, *in vitro* experiments showed that biotinylation efficiency is proximity-dependent, meaning that substrates closer to BirA* were more readily biotinylated (Choi-Rhee et al., 2004; Cronan, 2005). To promiscuously biotinylate proteins in mammalian cells, a codon-optimized BirA* was designed and fused to the protein of interest (Roux et al., 2012). With this approach, termed BioID, it was now possible to identify the proximal proteome of in theory any protein of interest. By switching from the *E. coli* to the *Aquifex aeolicus* biotin ligase, the size of the BioID moiety was reduced from 35 to 28 kDa (Kim et al., 2016). Later, it was possible to reduce the labeling time from a minimum of 6 h to 10 min with an *E. coli* biotin ligase mutated at 14 amino acids, namely TurboID (Branon et al., 2018). In parallel, a mutated and truncated biotin ligase from *Bacillus subtilis* (BASU) was developed and achieved efficient labeling for subsequent LC-MS/MS analysis in 30 min (Ramanathan et al., 2018). However, this improved activity was only demonstrated in a very specific context in which BirA* is fused to a small peptide that recognizes RNA motifs.

Furthermore, during the development of TurboID/miniTurboID, BASU showed kinetics comparable with BioID and BioID2 (Branon et al., 2018; **Figure 1A**; and **Table 1**).

Engineered Ascorbate Peroxidase (APEX)

Peroxidases can oxidize various chromogenic substrates in the presence of H_2O_2 , making it a versatile tool for biochemistry applications. For example, horseradish peroxidase (HRP) has been used to enhance contrast for electron microscopy by polymerizing 3,3'-diaminobenzidine after OsO_4 treatment (Graham and Karnovsky, 1966; Li et al., 2010). Peroxidases can also catalyze the oxidation of phenol derivatives to phenoxyl radicals (Gross and Sizer, 1959). This chemistry is the basis of tyramide signal amplification, a widely used technique for immunostainings (Mayer and Bendayan, 1997). Phenoxyl radicals can also react with electron-rich amino acids, predominantly tyrosine (>95%), but also tryptophan and cysteine (Udeshi et al., 2017). Because these radicals are very short lived (<1 ms), they can only react with amino acid residues in close proximity of the peroxidase (Mortensen and Skibsted, 1997). The first biotin-based proximity labeling study was done with HRP and aryl azide-biotin as substrate (Kotani et al., 2008). However, HRP is not active in the mammalian cytosol, because two essential disulfide bridges cannot form in the reducing environment (Martell et al., 2012). Introduction of an engineered ascorbate peroxidase (APEX) derived from pea overcame this caveat (Rhee et al., 2013). It is active in all cell compartments and can label surrounding proteins through incubation with H_2O_2 and biotin-phenol. The biotin-phenoxyl radicals primarily target tyrosine residues at surface-exposed sites of proteins. Furthermore, APEX with 28 kDa has a lower molecular weight opposed to the 44 kDa HRP, making the fusion protein less likely to compromise the native structure or function of the protein. The low catalytic activity of this first APEX version prompted a directed evolution approach and the development of the more active A134P mutated version of the enzyme named APEX2 (Lam et al., 2015; **Figure 1B**; and **Table 1**).

General Considerations for BioID and APEX Experiments

In summary, biotin ligases and ascorbate peroxidases provide a powerful tool to investigate the proximity of a protein of interest, giving insight into potential interaction partners. Nevertheless, fusing a relatively large 27–28 kDa enzyme to the bait protein may influence its function and/or localization (Roux et al., 2012; Roux, 2013; Kim and Roux, 2016). The moiety has a similar size as other common tags, e.g., green fluorescence protein (GFP). Consequently, a good practice might be to fuse the proximity labeling enzyme to N- or C-termini of target proteins that have already been successfully tagged with GFP or another moiety in the same size range. In general, the concept of proximity labeling does not allow direct testing for interaction partners, but rather provide a candidate list of possible interactors (Roux et al., 2012). The functional relevance of these candidates should then be validated by further experimentation.

Additionally, the labeling radius is not clear, especially for biotin ligases. The reactive bioAMP has a half-life of minutes, potentially enabling it to diffuse away from the biotin ligase (Rhee et al., 2013). However, a BioID study of the nuclear pore complex reported an effective labeling radius of only ~10 nm (Kim et al., 2014). Interestingly, the insertion of a flexible linker into the fusion protein can increase the labeling radius (Kim et al., 2016). The APEX2 generated biotin-phenoxyl radicals are very short lived (<1 ms), which leads to a decreasing degree of biotinylation with increasing physical radius from the peroxidase (Hung et al., 2016). When combining APEX2 biotinylation with the ratiometric Stable Isotope Labeling with Amino Acids in Cell Culture (SILAC) approach, it is possible to achieve high spatial resolution, especially in non-membrane enclosed compartments (Hung et al., 2016). Electron microscopy images suggest the labeling radius of biotin-phenoxyl radicals to be ~10–20 nm (Mayer and Bendayan, 1997). Another drawback is that the strong biotin-streptavidin bond does not allow for efficient elution of biotinylated proteins from the beads. This is usually circumvented by on-bead digestion, but interactions of non-biotinylated proteins with the beads can introduce many false positives. Additionally, the biotinylated peptides that are cleaved off the beads still containing part of streptavidin are too complex to be analyzed by mass spectrometry, leading to a loss of important peptides for later analysis. New methods, e.g., Biotin Site Identification Technology (BioSITE) and Direct Detection of Biotin-containing Tags (DiDBit) aim at addressing these issues by first digesting the proteins and subsequently enriching with biotin nano- or antibodies (Schiapparelli et al., 2014; Kim et al., 2018). Using antibodies does not lead to complex undetectable peptides. Also, this approach can potentially increase sensitivity, because enrichment on the peptide level greatly reduces the background of non-biotinylated peptides (Udeshi et al., 2017; Kim et al., 2018). Additionally, this approach allowed the identification of the preferential biotinylation sites on proteins (Udeshi et al., 2017).

When designing a proximity labeling experiment, an important point to consider is that large amounts of false positives can be generated due to random spatial association with the protein of interest. Consequently, negative controls are mandatory and should always be included in the experimental setup (Lobingier et al., 2017). In general, two types of controls are recommended – a technical control without the proximity labeling reaction and importantly, a spatial control mimicking the reaction at specific subcellular locations. Technical controls give insight into contaminants arising through the enrichment strategy, whereas the spatial control expresses the enzyme alone or fused to a localization tag, e.g., NLS-BirA*, and provides information of common contaminants of the labeling reaction itself. Furthermore, it is crucial to limit and achieve similar expression levels of bait and control fusion proteins, otherwise different levels of background can mask bona fide interactions. For BioID, cells with no BirA*, BirA* alone, or BirA* fused to a localization tag are the three most common controls. This is transferrable to APEX experiments, but it is also possible to omit H_2O_2 or biotin phenol instead of using no APEX. Furthermore, a database named CRAPome

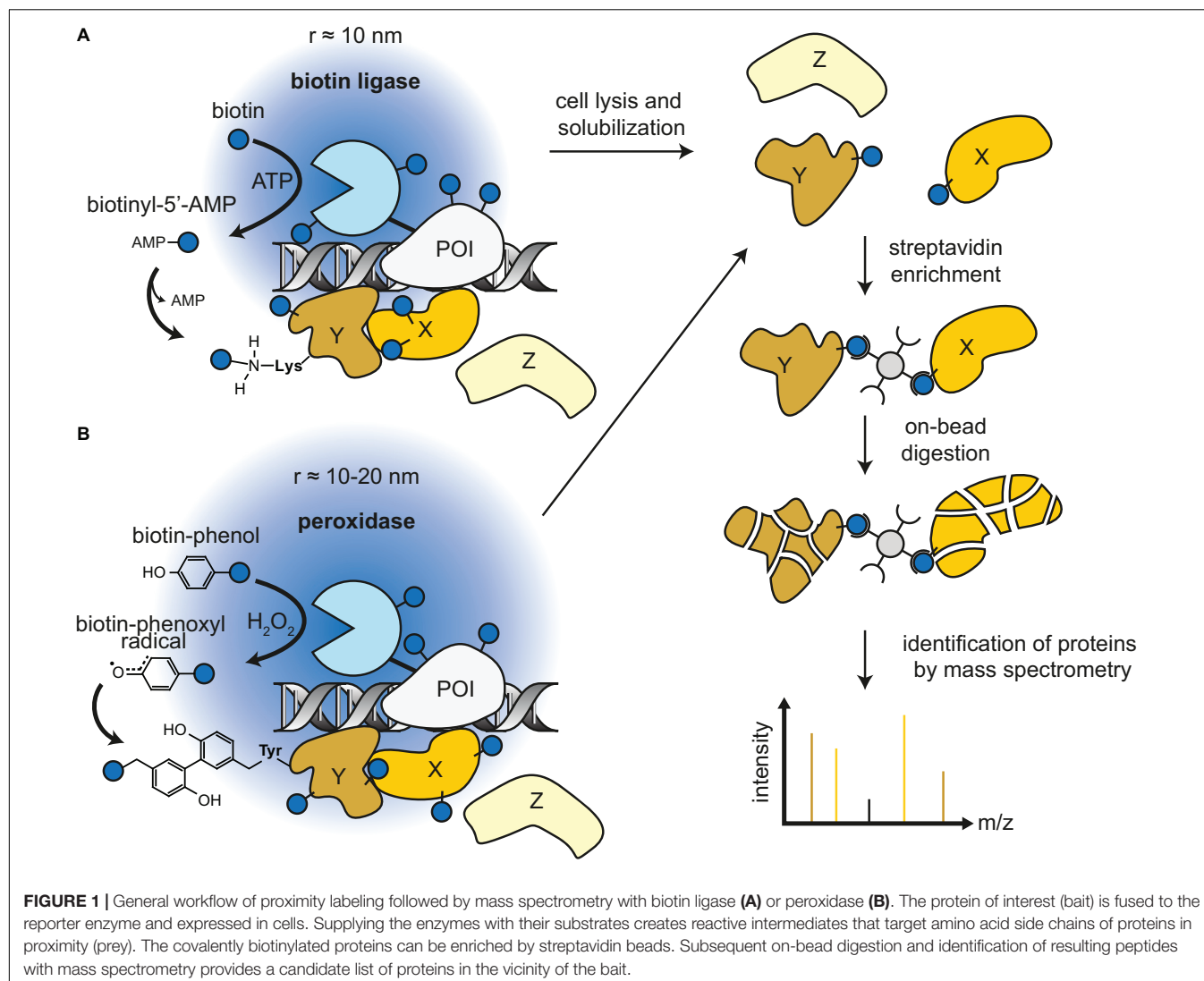


FIGURE 1 | General workflow of proximity labeling followed by mass spectrometry with biotin ligase (A) or peroxidase (B). The protein of interest (bait) is fused to the reporter enzyme and expressed in cells. Supplying the enzymes with their substrates creates reactive intermediates that target amino acid side chains of proteins in proximity (prey). The covalently biotinylated proteins can be enriched by streptavidin beads. Subsequent on-bead digestion and identification of resulting peptides with mass spectrometry provides a candidate list of proteins in the vicinity of the bait.

TABLE 1 | Overview of available proximity labeling enzymes and their characteristics.

Enzyme	Type	Source organism	Ami no acid mutations	Size in kDa	Labeling time	Substrate incubation time	Substrates	Labeling targets	Reference
BiolD	Biotin ligase	<i>E. coli</i>	R118G	35	6–24 h	6–24 h	Biotin	Lys	Roux et al., 2012
BiolD2	Biotin ligase	<i>A. aeolicus</i>	R40G	27	6–24 h	6–24 h	Biotin	Lys	Kim et al., 2016
BASU	Biotin ligase	<i>B. subtilis</i>	13 mut., Δ N-term	29	30 min–12 h	30 min–12 h	Biotin	Lys	Ramanathan et al., 2018
miniTurbo	Biotin ligase	<i>E. coli</i>	12 mut., Δ N-term	28	10–60 min	10–60 min	Biotin	Lys	Branon et al., 2018
TurboID	Biotin ligase	<i>E. coli</i>	14 mut., Δ N-term	35	10–60 min	10–60 min	Biotin	Lys	Branon et al., 2018
HRP	Peroxidase	Horseradish	–	44	5–10 min	5–10 min	Biotin-phenol, Fluorescein-aryl azide	Tyr, Trp, Cys, His	Kotani et al., 2008; Li et al., 2014; Rees et al., 2015
APEX	Peroxidase	Pea	K14D, E112K, W41F	28	1 min	30–60 min	Biotin-phenol	Tyr, Trp, Cys, His	Martell et al., 2012; Rhee et al., 2013
APEX2	Peroxidase	Soybean	K14D, E112K, W41F, A134P	28	1 min	30–60 min	Biotin-phenol, -aniline, -naphthylamine	Tyr, Trp, Cys, His	Lam et al., 2015

for known contaminants in immunoprecipitation and BioID experiments has been established (Mellacheruvu et al., 2013). It is possible to select specific negative controls from other studies, e.g., NLS-BirA* if the experimental designs are highly similar. However, if the cell type or the enrichment strategy of the control differs significantly, it is always recommended to include an internal experimental control, rather than solely relying on the CRAPome database. Furthermore, experimental design also entails whether to use a qualitative or one of the many quantitative mass spectrometry approaches. There does not seem to be a preferential method for proximity labeling, so it comes down to technical considerations (see Santin, 2019 for details and Ankney et al., 2016 for a comprehensive summary of quantitative approaches).

Another point of consideration for proximity labeling is that the amount of biotinylation does not necessarily reflect the strength of association. In fact, biotinylation relies on the number and accessibility of the targeted amino acid residues, mostly lysine or tyrosine. This also means that intrinsically disordered regions of proteins, which are very sensitive to changes in pH, salt concentration and PTMs, can introduce biases in proximity labeling studies (Minde et al., 2020). On average, the preferentially targeted lysines in BioID are a lot more abundant in intrinsically disordered regions than tyrosines preferred by APEX. This could also explain the fact that a biotinylation gradient is observed with APEX, but not with BioID.

CHROMATIN FACTORS TARGETED BY PROXIMITY LABELING

Proximity labeling has been used to study chromatin factors in many different nuclear compartments (Figure 2). In the following sections, selected studies are presented covering a wide range of proximity labeling techniques. A more complete list can be found in **Supplementary Table S1**.

Histone Variants and Post-translational Modifications

One of the first approaches to study chromatin with biotin ligase was to identify histone modifications in the vicinity of RAD18 (Shoaib et al., 2013). The authors fused BirA to RAD18 and BAP to histones H3.1 and H2A. By combining this approach with native Chromatin Immunoprecipitation (NChIP), it was found that the H4 histones in proximity to RAD18 are hyperacetylated compared to bulk histones. Importantly, this study proved the feasibility to fuse BirA to any nuclear protein of interest and determine features of the surrounding histones. However, the extensive MNase digestion only creates biotinylated mononucleosomes, therefore excluding the analysis of non-histone protein interactions.

The development of BioID has overcome this limitation. Additionally, BioID applications are technically less challenging, as only one genetic fusion of BirA* to the protein of interest is required without the counterpart BAP. By fusing BirA* to the histone H3 like nucleoprotein CENP-A, HJURP was identified as a centromere associating protein in S phase (Zasadzińska et al.,

2018). This interaction was later confirmed by *in vitro* BioID (ivBioID) (Remnant et al., 2019). In this variation of the assay, the biotin substrate is only added after a brief pre-extraction period and therefore allows quick substrate penetration and biotinylation in a timescale of minutes. This addresses the shortcomings of the regular BioID approach, which needs a biotin incubation time of at least 6 h. However, it is less suited for soluble proteins, because they are washed from the cells after permeabilization. Furthermore, it does not require treatment of cells with H₂O₂, a potentially oxidative damage-inducing agent. However, the use of H₂O₂ in the regular APEX2 protocol at low concentrations and short time periods of 60 s may not severely impact signaling pathways (Veal et al., 2007). Also, the development of TurboID reduced the biotin labeling time to 10 min, addressing the same issue of the standard BioID. Nevertheless, ivBioID seems to provide lower background and can resolve even finer time intervals, providing a snapshot of the proxisome at the time of lysis. In addition, ivBioID can be used in any genetically modifiable organism regardless of difficulties with endogenous biotin levels or biotin delivery.

In contrast, the APEX2 approach has been developed mainly in mammalian cells and is not easily transferable to other organisms. The main concern is the delivery of the substrate biotin phenol into the cell or nucleus (Hung et al., 2016). For yeast, removing the cell wall by zymolyase or osmotic shock allows the entry of biotin phenol (Hwang and Espenshade, 2016). Further optimization of the protocol by a different group enabled the proteomic mapping of the mitochondrial matrix and the nucleus (Singer-Krüger et al., 2019). As an example, fusion of APEX2 to the core H2B histone Htb1 identified Yer156c, a nuclear protein with unknown function previously not detected with traditional IP-MS approaches.

Recently, a method named ChromID to study the proxisome of specific histone PTMs was published (Villaseñor et al., 2020). In this approach, engineered chromatin readers (eCRs) are fused to the biotin ligase BASU. In this study, eCRs for histone tri-methylated H3K4, H3K9, and H3K27 have been developed and successfully used with proximity labeling. Additionally, the authors were able to employ a bivalent eCR to study the proxisome at H3K4me3 and H3K27me3 marked sites. This method has very promising potential for studying associating factors of histone modifications in different conditions. It might also be useful for tracking histone mark proxisomes during developmental changes. However, with a labeling time of 12 h, ChromID might be less suitable to study dynamic cellular processes.

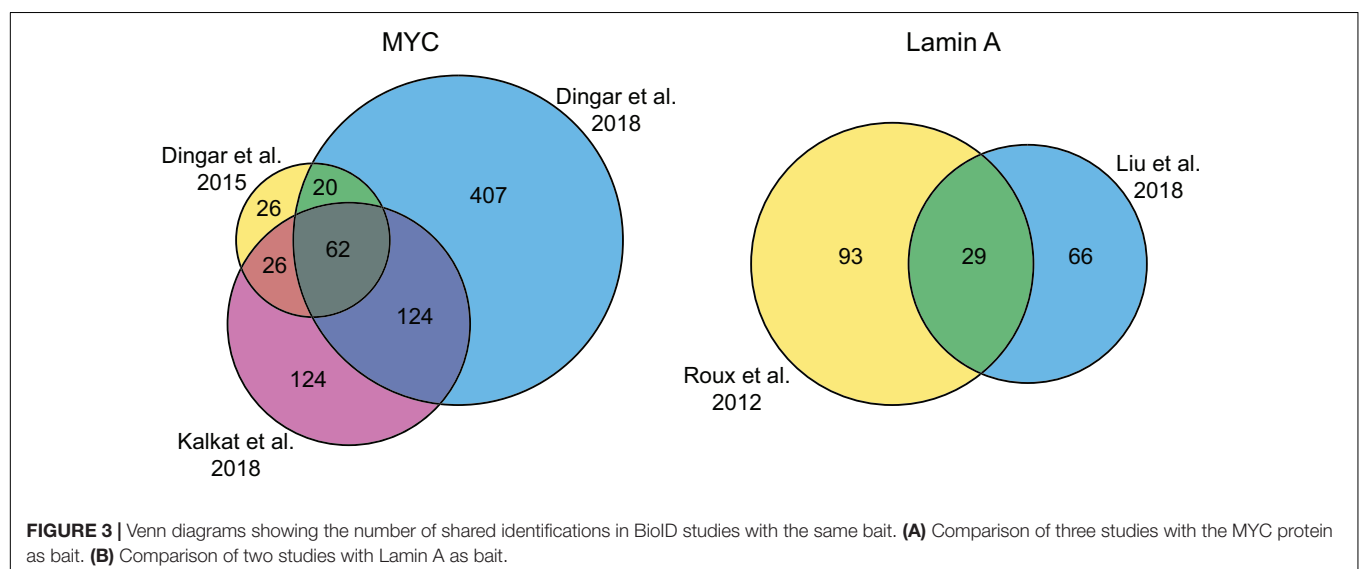
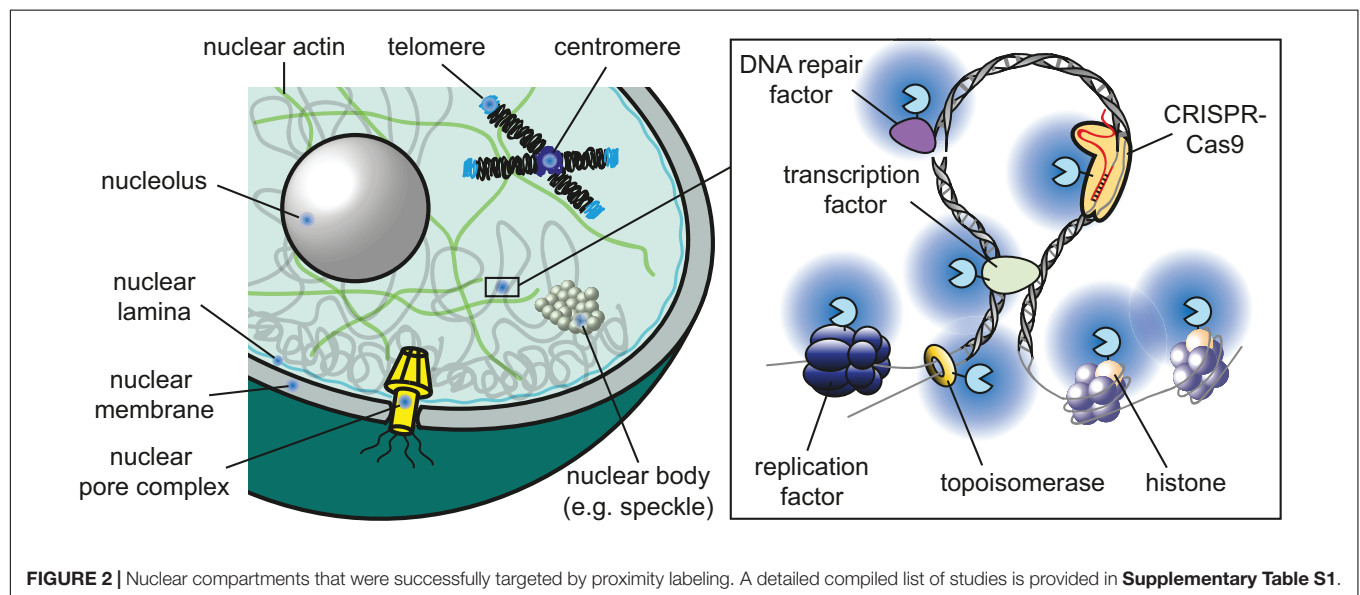
Transcription Factors

Multiple proxisomes of transcription factors have been uncovered with the help of proximity labeling. Fusion of BirA* to the MYC oncoprotein in cultured HEK293 and tumor xenografts confirmed known and identified over 70 new potential interaction partners, ranging from chromatin remodelers to transcription factors (Dingar et al., 2015). Therefore, proximity labeling significantly improved our knowledge of potential MYC interactors, which has been difficult to study with classical

IP/AP-MS due to difficult to solubilize chromatin-bound complexes containing MYC. Later, the same group identified protein phosphatase 1 (PP1) and its regulatory subunit protein phosphatase-1 nuclear-targeting subunit (PNUTS) as MYC-interactors in HeLa cells (Dingar et al., 2018). In an additional study, the six highly conserved MYC homology boxes (MBs) were individually deleted and the mutants were fused to BirA* in HEK293 cells (Kalkat et al., 2018). Some of these MBs are crucial for MYC-dependent malignant transformation. The resulting six BioID proxisomes were compared to the wild type proxisome and gave important insights into the binding targets of the individual MYC homology boxes. Interestingly, when comparing these three MYC proxisome studies done in HeLa cells and HEK293 cells/tumor xenografts, a large overlap of 62 candidates can be observed despite the disparity of cellular systems (**Figure 3**). This suggests that BioID can efficiently

and reproducibly detect specific interactions and these can be considered as the “core” high-confidence hits, whereas the other candidates only detected in one or two of the studies may contain more bona fide targets in a rather cell- and context-specific manner. In general, this example also illustrates that proximity labeling techniques can potentially discriminate between such a hierarchy of different interaction levels.

GFI1B is a master regulator of developmental hematopoiesis, which can also play both oncogenic and oncosuppressor roles in hematologic malignancies (Anguita et al., 2017). To study GFI1B's proxisome, a GFI1B-BioID2 fusion protein was used (McClellan et al., 2019). Besides many known interactors and members of other transcription complexes, the H3K4me1/2 and H3K9me1/2 specific lysine demethylase LSD1 was identified. To repress transcription, GFI1B needs to bind to LSD1 via its SNAG domain. To identify LSD1-dependent transcriptional



regulatory complexes, BioID2 fusions with a wildtype SNAG domain or different mutant alleles were created. Importantly, specific enrichment of the BRAF-HDAC complex (BHC) was only detected in the proxisome of the constructs with intact SNAG domain. Consequently, proximity labeling was able to identify LSD1 dependence of the BHC complex. Thus, the two examples of GFI1B and MYC show that combining proximity labeling data sets of different deletion mutants of the same protein can even identify the interactome of specific functional domains.

Another transcription factor fused to BirA* is SOX2 (Kim et al., 2017). Copy number gains of SOX2 arise in almost all squamous cell carcinomas (SQCC) of the lung, suggesting a functional role in disease progression (Weina and Utikal, 2014). Similar to many other transcription factors, direct targeting of SOX2 by small molecule inhibitors was not successful. The first SOX2 proxisome analysis by BioID confirmed the association with histone acetyltransferase EP300 in HEK293 cells, an interaction that was not clear due to conflicting AP-MS studies (Kim et al., 2017). This approach illustrates how proximity labeling can be used to screen for interaction partners of “non-druggable” oncoproteins that can then be targeted for improved therapeutic control of transcription factor oncogenic functions. Interestingly, of the 82 candidates, 46 were also found in at least one of eight AP-MS SOX2 interactome studies, suggesting BioID is able to identify new and verify many known interactions.

The fusion protein EWS-Fli-1 can be generated after chromosomal translocations and is present in most cases of Ewing sarcoma, an aggressive form of bone cancer (Li et al., 2015). To assess the interactome of EWS-Fli-1, a tandem affinity purification approach was first applied (Elzi et al., 2014). However, the majority of the expressed and tagged fusion protein was not effectively solubilized under non-denaturing conditions. Again, this limitation was overcome by BioID, where they could detect and subsequently verify a connection between the lysosome and EWS-Fli-1 protein turnover. Interestingly, this could be achieved with ~10 times less cell material than in the tandem affinity purification approach (Elzi et al., 2014).

ZEB1 is a transcription factor mediating epithelial-to-mesenchymal transition during development, but also in tumor progression (Zhang et al., 2015). To identify potential co-repressors of ZEB1, BioID was employed and allowed the identification of every core member of the nucleosome remodeling and deacetylase (NuRD) complex (Manshouri et al., 2019). Of note, the authors fused BirA* to either the N- and C-terminus of ZEB1 and only considered candidates present in both proxisomes and not present in the control. With this approach, they could identify 68 potential interactors of ZEB1. Subsequent experiments revealed the Rab22 GTPase-activating protein TBC1D2b gene as a ZEB1/NuRD complex target. TBC1D2b is crucial for suppressing E-cadherin internalization, a process that promotes the epithelial-to-mesenchymal transition.

Chromatin Remodelers and Topoisomerases

Chromatin remodelers play a key role in reshaping the chromatin landscape to grant access to transcription,

replication or DNA-repair factors. They can either influence the DNA-binding properties of histones through N-terminal modifications or directly move, evict or restructure nucleosomes in an ATP-dependent manner. For example, the histone methyltransferase NSD2 specifically dimethylates histone H3 lysine 36 (H3K36me2). This modification is associated with gene activation and overexpression of NSD2 has been linked to some forms of cancer (Kuo et al., 2011). To uncover potential interaction partners of NSD2, a BirA* fusion protein was overexpressed in NSD2 stable knock-out cells (Huang et al., 2019). The authors pursued both a qualitative and quantitative approach. First, they analyzed six biological replicates of NSD2-BirA* expressing cells against wild type cells, resulting in 63 candidates. In the second approach, label-free quantitative mass spectrometry analysis of NSD2 BioID with NLS-BirA* as the control resulted in 24 nuclear candidates. The overlap between the qualitative and quantitative approaches provided 16 potential high-confidence interactors. Further characterization of hits identified PARP1 as a regulator of NSD2 upon DNA-damage.

Topoisomerases are necessary for all biological processes that require DNA topology changes, including transcription, replication or chromatin remodeling (Chen et al., 2013). Thus, protein-protein interaction maps of topoisomerases would be particularly important to understand their essential functions in the cell, but difficult to achieve due to their insoluble properties in biochemical assays. BioID with topoisomerase II β as bait and no BirA*, GFP-BirA* or NLS-BirA* as controls could identify 25 proximal proteins, of which 4 were known and 21 unknown (Uusküla-Reimand et al., 2016). Here, the usage of three distinct kinds of controls increased the stringency of analysis. The authors could subsequently confirm TOP2II β associations with CTCF and cohesin subunits at the boundaries of topologically associating domains.

DNA Repair and Replication Factors

The MCM2-7 complex is known for its helicase activity during replication in S-phase, but has also been associated with DNA repair, chromatin organization and cell cycle regulation (Bailis and Forsburg, 2004). In an attempt to identify a more complete interaction map, Dubois et al. employed affinity or proximity purification followed by LC-MS/MS in a side-by-side comparison (Dubois et al., 2016). To this end, the authors fused either GFP or BirA* to each of the six MCM2-7 subunits and using the SILAC method subsequently pulled down with GFP nanobodies or streptavidin beads. The BioID approach generated roughly the same number of potential interactors as AP-MS. Interestingly, in this case the two approaches only shared ~15–20% of candidate hits, but it is not clear if this could originate from high background of both methods due to the endogenously high expression level of MCM complexes in cells (Dubois et al., 2016). Following etoposide treatment, they could identify DNA damage specific MCM interactors including the DDB1-CUL4 complex involved in nucleotide excision repair. Unfortunately, no BirA* reference (e.g., NLS-BirA*) was used in this study, increasing potential false-negatives. However, they could still generate high confidence hits due to very stringent cut-offs and

to the improved statistical power of 12 data sets merged from the two purification approaches.

In a similar approach, the same group probed the interactome of the master regulator HNF4 α , which plays a crucial role in development and tumorigenesis (Babeu et al., 2019). Increased expression of the isoform P2-HNF4 α has recently been implicated in colorectal cancer. BioID with P2-HNF4 α -BirA* and immunoprecipitation with P2-HNF4 α -GFP in HEK293T or HCT116 (colorectal cancer cell line without HNF4 α expression) revealed an association of P2-HNF4 α with DNA repair factors including PARP1, RAD50, and PRKDC. They confirmed these interactions by co-immunoprecipitation with endogenous HNF4 α in colorectal cancer cell lines. Here, BioID generated about four times more candidates than AP-MS, but also had a higher background. Interestingly, the DNA repair factors were found in the relatively small overlap of both approaches. This suggests that using both methods simultaneously can potentially provide biologically relevant candidate hits.

Another example of coupling proximity and affinity purification with mass spectrometry is a study investigating the interactome of the DNA repair factor Ku70 (Abbasi and Schild-Poulter, 2019). Besides its well-known role in non-homologous end-joining, Ku70 is also implicated in other chromatin processes, e.g., transcriptional regulation or DNA replication (Mo and Dynan, 2002; Abdelbaqi et al., 2013). BioID identified a total of 501 candidates across three biological replicates, while AP-MS detected 282. Interestingly, on average, ~55% of BioID candidates of a biological replicate were present in all three biological replicates, whereas this proportion was only ~18% for AP-MS. This indicates that the AP-MS method is not as reproducible as BioID for probing the interactome of Ku70.

Together, these studies indicate that proximity labeling is able to discover physical interactors that are also found in AP-MS experiments. However, the overlap during side-by-side experiments is usually small. Interestingly, ~50% of candidates of the SOX2 BioID proxisome could be found in at least one of eight different SOX2 AP-MS interactomes (Kim et al., 2017). The individual overlaps of the AP-MS interactomes with the BioID proxisome range from ~0 to 40% (Supplementary Table 5 in Kim et al., 2017). This indicates a large variation in the AP-MS interactomes and is most likely due to experimental design factors in the AP-MS experiments, e.g., cell type, crosslinking conditions, enrichment strategy and analysis parameters. It will be interesting to see if proximity labeling is less susceptible to variations resulting from different experimental designs. The three MYC BioID studies described above had a large overlap, but the experimental parameters were very similar (see section Transcription Factors and **Figure 3A**). When comparing two studies with Lamin A as bait in the same cell type, but with different enrichment and mass spectrometry analysis strategies, there is still a decent overlap of candidates (**Figure 3B**). Based on these examples, the studies suggest that the generated candidate lists of BioID proximity labeling experiments are less susceptible to variations caused by experimental design factors than AP-MS, but more comparative studies on distinct targets will be needed to verify this speculation.

Locus Specific

All methods described until now give insight into chromatin interactions that can occur genome-wide without any spatial information. However, it would also be interesting to investigate protein-protein interactions at specific DNA loci, especially in the context of oncogenes. The most commonly used method is a special form of IP-MS called chromatin immunoprecipitation (ChIP). There are two types of ChIP: cross-linked (XChIP) and native ChIP (NChIP). In XChIP, the chromatin is reversibly cross-linked with associating proteins and subsequently sheared by sonication. For NChIP, the native chromatin is extensively digested by micrococcal nuclease (MNase). To immunoprecipitate the local chromatin environment, antibodies targeting histone posttranslational modifications or chromatin factors of interest are used. The isolated and purified DNA is then sequenced to allocate genomic locations of the protein-protein or protein-DNA interaction. NChIP is mostly limited to histone proteins due to their high abundance and stable interaction with DNA, whereas other proteins are lost without crosslinking during the stringent IP conditions. However, XChIP can generate false positives by crosslinking randomly associating proteins or after cell lysis by non-specific binding of factors to the sheared chromatin or bead material. Furthermore, crosslinking agents distort the native environment of chromatin before analysis (Beneke et al., 2012; Gavrilov et al., 2015).

Another method is the Proteomics of isolated chromatin segments (PICH), which deploys complementary DNA probes after chemical crosslinking to capture the local chromatin composition (Déjardin and Kingston, 2009). A different approach targets a specific genomic region with a site-specific recombinase that can then be purified by affinity purification (Griesenbeck et al., 2003; Hamperl et al., 2014). Recently, the CUT&RUN method was introduced as an alternative to ChIP for genome wide profiling of the local chromatin environment of a chromatin factor of interest (Skene and Henikoff, 2017). In this *in situ* approach, protein A-fused MNase is directed to a specific antibody against the chromatin target of interest and leads to the release of protein-DNA complexes into solution without the requirement of crosslinking agents (Skene and Henikoff, 2017). As this technique basically represents a proximity-based reaction in close to native conditions, it will be interesting to pursue how the CUT&RUN method could complement BioID and APEX studies.

The development of a catalytically dead dCas9-BirA* fusion protein has laid the foundation for an *in vivo* approach using proximity labeling (Schmidtman et al., 2016). In principle, cells expressing dCas9-BirA* in combination with a single guide RNA can be targeted to any genomic locus of interest. Incubation with biotin should then allow to label the locus-proximal proteins *in vivo*. This original approach, termed CasID, was validated by targeting the repetitive sequences of telomeres, major satellite and minor satellite DNA (Schmidtman et al., 2016). The authors could identify known interactors but also validated zinc-finger protein 512 as a new major satellite repeat associating protein. However, the generated telomere protein list was rather short, with only seven significantly enriched proteins. It was possible to

increase BirA* activity and thereby gain more protein enrichment by designing a longer flexible glycine-serine linker between the dCas9 protein and BirA* (Li et al., 2019). In this study, they were able to generate a telomere associating protein list of over 300. Although increasing the chance for false positives with this extended linker approach, the authors could identify and validate desmoplakin as a telomere associating protein. To target single copy loci in the future, critical steps to optimize may include using multiple sgRNAs targeting the same locus, increasing cell numbers or optimizing the streptavidin pulldown (Schmidtman et al., 2016).

The next advance tried to address the slow reaction dynamics of BirA* by fusing APEX2 to dCas9. Similar to the studies with dCas9-BirA*, in this approach termed dCas9-APEX2 biotinylation at genomic elements by restricted spatial tagging (C-BERST), first the telomeres and centromeres were targeted, which allowed specific profiling of their subnuclear proteomes (Gao et al., 2018). Simultaneously, an approach to study non-repetitive single loci, termed genomic locus proteomics (GLOPro) was developed (Myers et al., 2018). The authors used five different sgRNAs targeting and tiling the same locus. These sgRNAs were expressed in separate HEK293T cell lines. Consequently, they were able to overlap the data sets and eliminate common noise. To limit artifacts from constitutive overexpression of dCas9-APEX2, expression was fine-tuned by an inducible promoter. With this approach, a snapshot of the proximal proteome of the *hTERT* and *c-MYC* promoters were obtained. In general, a benefit of these approaches is the possibility of using a simple and highly effective control without sgRNA or a non-specific sgRNA.

Another method was not only able to identify locus-specific proximal proteins, but also RNA and long range DNA-interactions by subsequent chemical crosslinking and high-throughput sequencing (Qiu et al., 2019). The authors did not fuse APEX2 directly to dCas9, but expressed a sgRNA that contains MS2 stem loops. This secondary structure is then specifically recognized by the MS2 coat protein (MCP) fused to APEX2. A major drawback of this approach could be non-bound MCP-APEX2 fusion proteins that generate false-positives. In agreement, the authors show that low expression of MCP-APEX2 is necessary for successful enrichment.

Protein-protein interactions at telomeres are of broad interest, because telomere length plays an important role in tumorigenesis. Telomerase is reactivated in most cancers, but there are cancers in which telomerase is suppressed and telomeres are maintained by alternative lengthening of telomeres (ALT). To identify proximal factors of ALT cell telomeres, BioID proximity labeling with TRF1-BirA* was used (Garcia-Exposito et al., 2016). By comparing the proxisome of ALT-positive U2OS with telomerase-positive HeLa cells, they were able to identify a role of translesion DNA synthesis in the ALT mechanism. Since biotin labeling occurs over all the different cell cycle states of telomeres, the BioID approach has the advantage over the previously used PiCH method (Déjardin and Kingston, 2009) to give a more comprehensive overview of protein interactions at telomeres. However, PiCH or APEX2 provide a “snapshot” and would therefore be superior when combined with cell synchronization if the goal is to analyze different time points during the cell cycle.

In summary, it is possible to use proximity labeling to identify the proxisomes of specific loci by employing dCas9 fusion proteins. However, targeting non-repetitive single loci is challenging, due to the low number of bound proximity labeling enzymes and resulting low biotinylation levels. In most studies, repetitive DNA was targeted greatly increasing the signal-to-noise ratio. Nevertheless, this limitation could be partially overcome by tiling the locus with multiple sgRNAs, as performed with the *hTERT* and *c-MYC* promoters (Myers et al., 2018). Thus, it could become feasible that the proxisomes are determined even at single copy gene loci, but that remains to be seen.

Nuclear Compartments

Instead of focusing on specific factors or genomic loci, another emerging application of proximity labeling is the analysis of whole compartments or difficult to isolate/purify components of the nucleus, such as the nuclear envelope, centromeres, or nuclear bodies. As a proof-of-principle, the nuclear lamina was targeted by BioID of Lamin A, a major component of this nuclear compartment (Roux et al., 2012). Lamin A was also studied in the context of Hutchinson-Gilford progeria, a premature aging syndrome. Here, BirA* was fused to normal Lamin A or the truncated form characteristic to this disease, called progerin (Chojnowski et al., 2015). By comparing the differential abundance of proximal proteins, they could detect reduced association of LAP2 α with progerin compared to Lamin A. In another approach, the proxisome of Lamin B1 was explored with a lenti-virus-delivered LMNB1-BirA* fusion protein (Fu et al., 2015). This mode of delivery to perform BioID may be of advantage in cells that are difficult to transfect. A different study tried to address some key issues of the BioID approach while probing the proxisome of LAP2 β , another inner nuclear membrane component. The conventional fusion protein LAP2 β -BirA* is too large for correct localization, because it cannot travel through the nuclear pore complex from the outer to the inner nuclear membrane. To circumvent this restriction, a method with the rapamycin inducible dimerization between FK506 binding protein (FKBP) and FKBP-rapamycin binding (FRB) combined with BioID was developed (Chojnowski et al., 2018). In short, FRB (~10 kDa) was fused to LAP2 β and FKBP to BirA*. This smaller fusion protein was able to pass the nuclear pore complex and localize correctly. Subsequently, the rapamycin induced dimerization allowed FKBP-BirA* to bind FRB-LAP2 β . Importantly, this system is internally controlled without addition of rapamycin and seems to reduce false-positive identifications (Chojnowski et al., 2018). However, false-positives are still conceivable when the dimerization occurs in the cytoplasm before LAP2 β is relocated to the inner nuclear membrane. In a similar approach, the nuclear vicinity of vesicle-associated membrane protein-associated protein B (VABP) was explored with rapamycin directed APEX2 (James et al., 2019). VABP localizes primarily to the ER, but also to the inner nuclear membrane. In this study, the APEX2-FKBP fusion protein was additionally tagged with a nuclear localization signal. In combination with FRB-VABP, this allowed the specific enrichment of the nuclear proxisome of VABP.

Besides the nuclear lamina, certain nuclear bodies were analyzed by proximity labeling. SUP-46 is a *Caenorhabditis elegans* RNA binding protein with an essential role in sustaining transgenerational germline immortality. Proxisome analysis of the human homologs MYEF2 and HNRNPM with BioID revealed robust associations with paraspeckles, nuclear stress granules and the nucleolus (Johnston et al., 2017). Interestingly, a large overlap of ~60% was observed among the 133 and 110 candidates in the MYEF2 and HNRNPM BioID assays, respectively.

CONCLUSION AND PERSPECTIVES

Recent developments in proximity labeling techniques have provided a valuable platform to study chromatin in new ways. BioID, APEX and their successors have become a valuable complementation to classical nuclear protein-protein interaction studies like AP/MS and ChIP. Different variations of these assays have started to shed light on the native environment of specific chromatin factors, specific gene loci or even whole nuclear compartments.

Interestingly, peroxidases can also directly label RNA and potentially DNA with biotin (Fazal et al., 2019; Zhou et al., 2019). With APEX-seq, it is possible to probe the vicinity not only for proteins, but also for various forms of RNAs. It has been used in parallel to APEX-MS to study the organization of the translation initiation complex and repressive RNA granules (Padrón et al., 2019). In the future, this approach can potentially uncover the localization of RNAs, e.g., lncRNAs, in the native vicinity of proteins or specific loci.

Proximity labeling also has the capability of studying microprotein-protein interactions (Chu et al., 2017). Small open reading frames (smORFs) encode hundreds of thousands of microproteins and small peptides, of which only few have been characterized. However, some of these microproteins have important biological functions and uncovering their native context in the cell can give clues regarding function.

Recently, combinations of proximity labeling and protein-fragment complementation assays (PCA) were developed. In a PCA, two POI are fused to either half of a split reporter protein (enzyme or fluorescent protein). The reporter protein is reconstituted only upon interaction of the POI. However, the exact interaction dynamics between the two split fragments remain unknown, e.g., if the reconstitution is reversible. For proximity labeling, split-BioID and split-APEX2 have now been reported (Munter et al., 2017; Schopp et al., 2017; Xue et al., 2017; Han et al., 2019). As the biotinylation is dependent on the correct localization of both targeted factors, this approach can significantly reduce the number of false positives (Munter

et al., 2017). This approach is specifically interesting for transient protein interactions, where labeling only occurs at the right time and the right site when a protein complex is formed or a biological process has been initiated. Furthermore, splitting the reporter enzyme results in smaller tags for the POI and therefore potentially less functional impact.

A recent paper indicated that biotin ligase-based proximity labeling may potentially allow the study of intrinsically unstructured regions. These flexible lysine-rich protein domains are more accessible and show faster biotinylation kinetics than structured, less exposed regions (Minde et al., 2020). Consequently, a time course biotin “painting” approach could even give insight into differences of secondary or tertiary protein structures. In conclusion, proximity labeling is emerging as a powerful complementary tool to study the local environment of chromatin factors that can significantly improve our understanding of the complex interaction networks in the nucleus.

AUTHOR CONTRIBUTIONS

HU searched the literature, created the figures, and wrote and edited the manuscript. SH provided the guidance and edited the manuscript.

FUNDING

SH was supported by the Helmholtz Association and by a Starting Grant from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program (Grant Agreement No. 852798).

ACKNOWLEDGMENTS

We would like to thank all the scientists whose results we discussed in this review and apologize to all scientists whose work we could not cite due to space limitations. We would also like to thank all members of the Hamperl lab for their inspiring input, helpful discussions, and continuous support. We are especially grateful to Dr. Adam Burton for critical reading and helpful comments on the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.00450/full#supplementary-material>

REFERENCES

- Abbasi, S., and Schild-Poulter, C. (2019). Mapping the ku interactome using proximity-dependent biotin identification in human cells. *J. Proteome Res.* 18, 1064–1077. doi: 10.1021/acs.jproteome.8b00771
- Abdelbaqi, K., Di Paola, D., Rampakakis, E., and Zannis-Hadjopoulos, M. (2013). Ku protein levels, localization and association to replication origins in different stages of breast tumor progression. *J. Cancer* 4, 358–370. doi: 10.7150/jca.6289
- Agircan, F. G., Hata, S., Nussbaum-Krammer, C., Atorino, E., and Schiebel, E. (2016). Proximity mapping of human separase by the BioID approach.

- Biochem. Biophys. Res. Commun. 478, 656–662. doi: 10.1016/j.bbrc.2016.08.002
- Anguita, E., Candel, F. J., Chaparro, A., and Roldán-Etcheverry, J. J. (2017). Transcription factor GFI1B in health and disease. *Front. Oncol.* 7:54. doi: 10.3389/fonc.2017.00054
- Ankney, J. A., Muneer, A., and Chen, X. (2016). Relative and absolute quantitation in mass spectrometry-based proteomics. *Annu. Rev. Anal. Chem.* 11, 49–77. doi: 10.1146/annurev-anchem-061516-45357
- Babeu, J.-P., Wilson, S. D., Lambert, É., Lévesque, D., Boisvert, F.-M., and Boudreau, F. (2019). Quantitative proteomics identifies DNA repair as a novel biological function for hepatocyte nuclear factor 4 α in colorectal cancer cells. *Cancers* 11:626. doi: 10.3390/cancers11050626
- Bailis, J. M., and Forsburg, S. L. (2004). MCM proteins: DNA damage, mutagenesis and repair. *Curr. Opin. Genet. Dev.* 14, 17–21. doi: 10.1016/j.gde.2003.11.002
- Barker, D. F., and Campbell, A. M. (1981a). Genetic and biochemical characterization of the birA gene and its product: evidence for a direct role of biotin holoenzyme synthetase in repression of the biotin operon in *Escherichia coli*. *J. Mol. Biol.* 146, 469–492. doi: 10.1016/0022-2836(81)90043-90047
- Barker, D. F., and Campbell, A. M. (1981b). The birA gene of *Escherichia coli* encodes a biotin holoenzyme synthetase. *J. Mol. Biol.* 146, 451–467. doi: 10.1016/0022-2836(81)90042-90045
- Beckett, D., Kovaleva, E., and Schatz, P. J. (1999). A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci. Publ. Protein Soc.* 8, 921–929. doi: 10.1110/ps.8.4.921
- Beneke, S., Meyer, K., Holtz, A., Hüttner, K., and Bürkle, A. (2012). Chromatin composition is changed by poly(ADP-ribosylation) during chromatin immunoprecipitation. *PLoS One* 7:e32914. doi: 10.1371/journal.pone.0032914
- Branon, T. C., Bosch, J. A., Sanchez, A. D., Udeshi, N. D., Svinkina, T., Carr, S. A., et al. (2018). Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* 36, 880–887. doi: 10.1038/nbt.4201
- Brecht, R. M., Liu, C. C., Beilinson, H. A., Khitun, A., Slavoff, S. A., and Schatz, D. G. (2020). Nucleolar localization of RAG1 modulates V(D)J recombination activity. *Proc. Natl. Acad. Sci. U.S.A.* 117, 4300–4309. doi: 10.1073/pnas.1920021117
- Carneseccchi, J., Sigismondo, G., Domsch, K., Baader, C. E. P., Rafiee, M.-R., Krijgsvel, J., et al. (2020). Multi-level and lineage-specific interactomes of the Hox transcription factor Ubx contribute to its functional specificity. *Nat. Commun.* 11:1388. doi: 10.1038/s41467-020-15223-x
- Chen, S. H., Chan, N.-L., and Hsieh, T. (2013). New mechanistic and functional insights into DNA topoisomerases. *Annu. Rev. Biochem.* 82, 139–170. doi: 10.1146/annurev-biochem-061809-100002
- Choi-Rhee, E., Schulman, H., and Cronan, J. E. (2004). Promiscuous protein biotinylation by *Escherichia coli* biotin protein ligase. *Protein Sci. Publ. Protein Soc.* 13, 3043–3050. doi: 10.1110/ps.04911804
- Chojnowski, A., Ong, P. F., Wong, E. S. M., Lim, J. S. Y., Mutalif, R. A., Navasankari, R., et al. (2015). Progerin reduces LAP2 α -telomere association in Hutchinson-Gilford progeria. *eLife* 4:e07759. doi: 10.7554/eLife.07759
- Chojnowski, A., Sobota, R. M., Ong, P. F., Xie, W., Wong, X., Dreesen, O., et al. (2018). 2C-BioID: an advanced two component BioID system for precision mapping of protein interactomes. *iScience* 10, 40–52. doi: 10.1016/j.isci.2018.11.023
- Chou, C.-C., Zhang, Y., Umoh, M. E., Vaughan, S. W., Lorenzini, I., Liu, F., et al. (2018). TDP-43 pathology disrupts nuclear pore complexes and nucleocytoplasmic transport in ALS/FTD. *Nat. Neurosci.* 21, 228–239. doi: 10.1038/s41593-017-0047-43
- Chu, Q., Rathore, A., Diedrich, J. K., Donaldson, C. J., Yates, J. R., and Saghatelian, A. (2017). Identification of microprotein-protein interactions via APEX tagging. *Biochemistry* 56, 3299–3306. doi: 10.1021/acs.biochem.7b00265
- Cloer, E. W., Siesser, P. F., Cousins, E. M., Goldfarb, D., Mowrey, D. D., Harrison, J. S., et al. (2018). p62-Dependent phase separation of patient-derived KEAP1 mutations and NRF2. *Mol. Cell. Biol.* 38:e644-17. doi: 10.1128/MCB.00644-17
- Cronan, J. E. (2005). Targeted and proximity-dependent promiscuous protein biotinylation by a mutant *Escherichia coli* biotin protein ligase. *J. Nutr. Biochem.* 16, 416–418. doi: 10.1016/j.jnutbio.2005.03.017
- de Boer, E., Rodriguez, P., Bonte, E., Krijgsvel, J., Katsantoni, E., Heck, A., et al. (2003). Efficient biotinylation and single-step purification of tagged transcription factors in mammalian cells and transgenic mice. *Proc. Natl. Acad. Sci. U.S.A.* 100, 7480–7485. doi: 10.1073/pnas.1332608100
- Déjardin, J., and Kingston, R. E. (2009). Purification of proteins associated with specific genomic loci. *Cell* 136, 175–186. doi: 10.1016/j.cell.2008.11.045
- Dingar, D., Kalkat, M., Chan, P.-K., Srikumar, T., Bailey, S. D., Tu, W. B., et al. (2015). BioID identifies novel c-MYC interacting partners in cultured cells and xenograft tumors. *J. Proteomics* 118, 95–111. doi: 10.1016/j.jprot.2014.09.029
- Dingar, D., Tu, W. B., Resette, D., Lourenco, C., Tamachi, A., De Melo, J., et al. (2018). MYC dephosphorylation by the PP1/PNUTS phosphatase complex regulates chromatin binding and protein stability. *Nat. Commun.* 9:3502. doi: 10.1038/s41467-018-05660-5660
- Dubois, M.-L., Bastin, C., Lévesque, D., and Boisvert, F.-M. (2016). Comprehensive characterization of minichromosome maintenance complex (MCM) protein interactions using affinity and proximity purifications coupled to mass spectrometry. *J. Proteome Res.* 15, 2924–2934. doi: 10.1021/acs.jproteome.5b01081
- Eisenberg, M. A., Prakash, O., and Hsiung, S. C. (1982). Purification and properties of the biotin repressor. A bifunctional protein. *J. Biol. Chem.* 257, 15167–15173.
- Elzi, D. J., Song, M., Hakala, K., Weintraub, S. T., and Shio, Y. (2014). Proteomic analysis of the EWS-Fli-1 interactome reveals the role of the lysosome in EWS-Fli-1 turnover. *J. Proteome Res.* 13, 3783–3791. doi: 10.1021/pr500387m
- Fazal, F. M., Han, S., Parker, K. R., Kaewsapsak, P., Xu, J., Boettiger, A. N., et al. (2019). Atlas of subcellular RNA localization revealed by APEX-Seq. *Cell* 178, 473–490.e26. doi: 10.1016/j.cell.2019.05.027
- Fernández-Suárez, M., Chen, T. S., and Ting, A. Y. (2008). Protein-protein interaction detection in vitro and in cells by proximity biotinylation. *J. Am. Chem. Soc.* 130, 9251–9253. doi: 10.1021/ja801445p
- Fu, Y., Lv, P., Yan, G., Fan, H., Cheng, L., Zhang, F., et al. (2015). MacroH2A1 associates with nuclear lamina and maintains chromatin architecture in mouse liver cells. *Sci. Rep.* 5:17186. doi: 10.1038/srep17186
- Gao, X. D., Tu, L.-C., Mir, A., Rodriguez, T., Ding, Y., Leszyk, J., et al. (2018). C-BERST: defining subnuclear proteomic landscapes at genomic elements with dCas9-APEX2. *Nat. Methods* 15, 433–436. doi: 10.1038/s41592-018-0006-2
- García-Exposito, L., Bournique, E., Bergoglio, V., Bose, A., Barroso-Gonzalez, J., Zhang, S., et al. (2016). Proteomic profiling reveals a specific role for telomerase DNA polymerase η in the alternative lengthening of telomeres. *Cell Rep.* 17, 1858–1871. doi: 10.1016/j.celrep.2016.10.048
- Gavrilov, A., Razin, S. V., and Cavalli, G. (2015). In vivo formaldehyde cross-linking: it is time for black box analysis. *Brief. Funct. Genomics* 14, 163–165. doi: 10.1093/bfpg/elu037
- Graham, R. C., and Karnovsky, M. J. (1966). The early stages of absorption of injected horseradish peroxidase in the proximal tubules of mouse kidney: ultrastructural cytochemistry by a new technique. *J. Histochem. Cytochem. Off. J. Histochem. Soc.* 14, 291–302. doi: 10.1177/14.4.291
- Griesenbeck, J., Boeger, H., Strattan, J. S., and Kornberg, R. D. (2003). Affinity purification of specific chromatin segments from chromosomal loci in yeast. *Mol. Cell. Biol.* 23, 9275–9282. doi: 10.1128/MCB.23.24.9275-9282.2003
- Gross, A. J., and Sizer, I. W. (1959). The oxidation of tyramine, tyrosine, and related compounds by peroxidase. *J. Biol. Chem.* 234, 1611–1614.
- Hamperl, S., Brown, C. R., Gareia, A. V., Perez-Fernandez, J., Bruckmann, A., Huber, K., et al. (2014). Compositional and structural analysis of selected chromosomal domains from *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 42:e2. doi: 10.1093/nar/gkt891
- Han, Y., Branon, T. C., Martell, J. D., Boassa, D., Shechner, D., Ellisman, M. H., et al. (2019). Directed evolution of split APEX2 peroxidase. *ACS Chem. Biol.* 14, 619–635. doi: 10.1021/acscchembio.8b00919
- Huang, X., LeDuc, R. D., Fornelli, L., Schunter, A. J., Bennett, R. L., Kelleher, N. L., et al. (2019). Defining the NSD2 interactome: PARP1 PARylation reduces NSD2 histone methyltransferase activity and impedes chromatin binding. *J. Biol. Chem.* 294, 12459–12471. doi: 10.1074/jbc.RA118.006159
- Hung, V., Udeshi, N. D., Lam, S. S., Loh, K. H., Cox, K. J., Pedram, K., et al. (2016). Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat. Protoc.* 11, 456–475. doi: 10.1038/nprot.2016.018
- Hwang, J., and Espenshade, P. J. (2016). Proximity-dependent biotin labelling in yeast using the engineered ascorbate peroxidase APEX2. *Biochem. J.* 473, 2463–2469. doi: 10.1042/BCJ20160106
- James, C., Müller, M., Goldberg, M. W., Lenz, C., Urlaub, H., and Kehlenbach, R. H. (2019). Proteomic mapping by rapamycin-dependent targeting of APEX2 identifies binding partners of VAPB at the inner nuclear membrane. *J. Biol. Chem.* 294, 16241–16254. doi: 10.1074/jbc.RA118.007283
- Johnston, W. L., Krizus, A., Ramani, A. K., Dunham, W., Youn, J. Y., Fraser, A. G., et al. (2017). *C. elegans* SUP-46, an HNRNPM family RNA-binding

- protein that prevents paternally-mediated epigenetic sterility. *BMC Biol.* 15:61. doi: 10.1186/s12915-017-0398-y
- Kalkat, M., Resetcu, D., Lourenco, C., Chan, P.-K., Wei, Y., Shiah, Y.-J., et al. (2018). MYC protein interactome profiling reveals functionally distinct regions that cooperate to drive tumorigenesis. *Mol. Cell* 72, 836–848.e7. doi: 10.1016/j.molcel.2018.09.031
- Kim, B. R., Coyaudo, E., Laurent, E. M. N., St-Germain, J., Van de Laar, E., Tsao, M.-S., et al. (2017). Identification of the SOX2 interactome by BioID reveals EP300 as a mediator of SOX2-dependent squamous differentiation and lung squamous cell carcinoma growth. *Mol. Cell. Proteomics* 16, 1864–1888. doi: 10.1074/mcp.M116.064451
- Kim, D. I., Cutler, J. A., Na, C. H., Reckel, S., Renuse, S., Madugundu, A. K., et al. (2018). BioSITE: a method for direct detection and quantitation of site-specific biotinylation. *J. Proteome Res.* 17, 759–769. doi: 10.1021/acs.jproteome.7b00775
- Kim, D. I., Jensen, S. C., Noble, K. A., Kc, B., Roux, K. H., Motamedchaboki, K., et al. (2016). An improved smaller biotin ligase for BioID proximity labeling. *Mol. Biol. Cell* 27, 1188–1196. doi: 10.1091/mbc.E15-12-0844
- Kim, D. I., Kc, B., Zhu, W., Motamedchaboki, K., Doye, V., and Roux, K. J. (2014). Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2453–E2461. doi: 10.1073/pnas.1406459111
- Kim, D. I., and Roux, K. J. (2016). Filling the void: proximity-based labeling of proteins in living cells. *Trends Cell Biol.* 26, 804–817. doi: 10.1016/j.tcb.2016.09.004
- Klemm, S. L., Shipony, Z., and Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* 20, 207–220. doi: 10.1038/s41576-018-0089-88
- Kochanova, N. Y., Schauer, T., Mathias, G. P., Lukacs, A., Schmidt, A., Flatley, A., et al. (2020). A multi-layered structure of the interphase chromocenter revealed by proximity-based biotinylation. *Nucleic Acids Res.* 48, 1405–1415. doi: 10.1093/nar/gkaa145
- Kotani, N., Gu, J., Isaji, T., Uda, K., Taniguchi, N., and Honke, K. (2008). Biochemical visualization of cell surface molecular clustering in living cells. *Proc. Natl. Acad. Sci. U.S.A.* 105, 7405–7409. doi: 10.1073/pnas.0710346105
- Kuo, A. J., Cheung, P., Chen, K., Zee, B. M., Kioi, M., Luring, J., et al. (2011). NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. *Mol. Cell* 44, 609–620. doi: 10.1016/j.molcel.2011.08.042
- Kwon, K., and Beckett, D. (2000). Function of a conserved sequence motif in biotin holoenzyme synthetases. *Protein Sci. Publ. Protein Soc.* 9, 1530–1539. doi: 10.1110/ps.9.8.1530
- Kwon, K., Streaker, E. D., Ruparelia, S., and Beckett, D. (2000). Multiple disordered loops function in corepressor-induced dimerization of the biotin repressor. *J. Mol. Biol.* 304, 821–833. doi: 10.1006/jmbi.2000.4249
- Lam, S. S., Martell, J. D., Kamer, K. J., Deerinck, T. J., Ellisman, M. H., Mootha, V. K., et al. (2015). Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat. Methods* 12, 51–54. doi: 10.1038/nmeth.3179
- Lambert, J.-P., Picard, S., Fujisawa, T., Hou, H., Savitsky, P., Uuskula-Reimand, L., et al. (2019). Interactome rewiring following pharmacological targeting of BET bromodomains. *Mol. Cell* 73:621. doi: 10.1016/j.molcel.2018.11.006
- Lambert, J.-P., Tucholska, M., Go, C., Knight, J. D. R., and Gingras, A.-C. (2015). Proximity biotinylation and affinity purification are complementary approaches for the interactome mapping of chromatin-associated protein complexes. *J. Proteomics* 118, 81–94. doi: 10.1016/j.jprot.2014.09.011
- Li, J., Wang, Y., Chiu, S.-L., and Cline, H. T. (2010). Membrane targeted horseradish peroxidase as a marker for correlative fluorescence and electron microscopy studies. *Front. Neural Circ.* 4:6. doi: 10.3389/neuro.04.006.2010
- Li, P., Meng, Y., Wang, Y., Li, J., Lam, M., Wang, L., et al. (2019). Nuclear localization of desmoplakin and its involvement in telomere maintenance. *Int. J. Biol. Sci.* 15, 2350–2362. doi: 10.7150/ijbs.34450
- Li, X.-W., Rees, J. S., Xue, P., Zhang, H., Hamaia, S. W., Sanderson, B., et al. (2014). New insights into the DT40 B cell receptor cluster using a proteomic proximity labeling assay. *J. Biol. Chem.* 289, 14434–14447. doi: 10.1074/jbc.M113.529578
- Li, Y., Luo, H., Liu, T., Zacksenhaus, E., and Ben-David, Y. (2015). The ets transcription factor Fli-1 in development, cancer and disease. *Oncogene* 34, 2022–2031. doi: 10.1038/ncr.2014.162
- Liu, C.-H., Chien, M.-J., Chang, Y.-C., Cheng, Y.-H., Li, F.-A., and Mou, K. Y. (2020). Combining proximity labeling and cross-linking mass spectrometry for proteomic dissection of nuclear envelope interactome. *J. Proteome Res.* 19, 1109–1118. doi: 10.1021/acs.jproteome.9b00609
- Liu, X., Salokas, K., Tamene, F., Jiu, Y., Weldatsadik, R. G., Öhman, T., et al. (2018). An AP-MS- and BioID-compatible MAC-tag enables comprehensive mapping of protein interactions and subcellular localizations. *Nat. Commun.* 9:1188. doi: 10.1038/s41467-018-03523-3522
- Lobinger, B. T., Hüttenhain, R., Eichel, K., Miller, K. B., Ting, A. Y., von Zastrow, M., et al. (2017). An approach to spatiotemporally resolve protein interaction networks in living cells. *Cell* 169, 350–360.e12. doi: 10.1016/j.cell.2017.03.022
- López-Sooop, G., Rønningen, T., Rogala, A., Richartz, N., Blomhoff, H. K., Thiede, B., et al. (2017). AKAP95 interacts with nucleoporin TPR in mitosis and is important for the spindle assembly checkpoint. *Cell Cycle* 16, 947–956. doi: 10.1080/15384101.2017.1310350
- Manshour, R., Coyaudo, E., Kundu, S. T., Peng, D. H., Stratton, S. A., Alton, K., et al. (2019). ZEB1/NuRD complex suppresses TBC1D2b to stimulate E-cadherin internalization and promote metastasis in lung cancer. *Nat. Commun.* 10:5125. doi: 10.1038/s41467-019-12832-z
- Martell, J. D., Deerinck, T. J., Sancak, Y., Poulos, T. L., Mootha, V. K., Sosinsky, G. E., et al. (2012). Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nat. Biotechnol.* 30, 1143–1148. doi: 10.1038/nbt.2375
- Martin, A. P., Jacquemyn, M., Lipecka, J., Chhuon, C., Aushev, V. N., Meunier, B., et al. (2019). STK38 kinase acts as XPO1 gatekeeper regulating the nuclear export of autophagy proteins and other cargoes. *EMBO Rep.* 20:e48150. doi: 10.15252/embr.201948150
- Martino, J., Brunette, G. J., Barroso-González, J., Moiseeva, T. N., Smith, C. M., Bakkenist, C. J., et al. (2019). The human Shu complex functions with PDS5B and SPIDR to promote homologous recombination. *Nucleic Acids Res.* 47, 10151–10165. doi: 10.1093/nar/gkz738
- Mayer, G., and Bendayan, M. (1997). Biotinyl-Tyramide: a novel approach for electron microscopic immunocytochemistry. *J. Histochem. Cytochem.* 45, 1449–1454. doi: 10.1177/002215549704501101
- Mazina, M. Y., Ziganshin, R. H., Magnitov, M. D., Golovnin, A. K., and Vorobyeva, N. E. (2020). Proximity-dependent biotin labelling reveals CP190 as an Ecr/Usp molecular partner. *Sci. Rep.* 10:4793. doi: 10.1038/s41598-020-61514-61510
- McClellan, D., Casey, M. J., Bareyan, D., Lucente, H., Ours, C., Velinder, M., et al. (2019). Growth factor independence 1B-Mediated transcriptional repression and lineage allocation require lysine-specific demethylase 1-Dependent recruitment of the BHC complex. *Mol. Cell. Biol.* 39:e20-19. doi: 10.1128/MCB.00020-19
- Mellacheruvu, D., Wright, Z., Couzens, A. L., Lambert, J.-P., St-Denis, N. A., Li, T., et al. (2013). The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat. Methods* 10, 730–736. doi: 10.1038/nmeth.2557
- Minde, D.-P., Ramakrishna, M., and Lilley, K. S. (2020). Biotin proximity tagging favours unfolded proteins and enables the study of intrinsically disordered regions. *Commun. Biol.* 3, 1–13. doi: 10.1038/s42003-020-0758-y
- Minderjahn, J., Schmidt, A., Fuchs, A., Schill, R., Raithel, J., Babina, M., et al. (2020). Mechanisms governing the pioneering and redistribution capabilities of the non-classical pioneer PU.1. *Nat. Commun.* 11:402. doi: 10.1038/s41467-019-13960-13962
- Mirza, A. N., McKellar, S. A., Urman, N. M., Brown, A. S., Hollmig, T., Aasi, S. Z., et al. (2019). LAP2 proteins chaperone GLI1 movement between the lamina and chromatin to regulate transcription. *Cell* 176, 198–212.e15. doi: 10.1016/j.cell.2018.10.054
- Mo, X., and Dynan, W. S. (2002). Subnuclear localization of ku protein: functional association with RNA Polymerase II elongation sites. *Mol. Cell. Biol.* 22, 8088–8099. doi: 10.1128/MCB.22.22.8088-8099.2002
- Mortensen, A., and Skibsted, L. H. (1997). Importance of carotenoid structure in radical-scavenging reactions. *J. Agric. Food Chem.* 45, 2970–2977. doi: 10.1021/jf970010s
- Moser, B., Basilio, J., Gotzmann, J., Brachner, A., and Foisner, R. (2020). Comparative interactome analysis of emerin, MAN1 and LEM2 reveals a unique role for LEM2 in nucleotide excision repair. *Cells* 9:463. doi: 10.3390/cells9020463
- Müller, M., James, C., Lenz, C., Urlaub, H., and Kehlenbach, R. H. (2020). Probing the environment of emerin by enhanced ascorbate Peroxidase 2 (APEX2)-Mediated proximity labeling. *Cells* 9:605. doi: 10.3390/cells9030605
- Munter, S. D., Görnemann, J., Derua, R., Lesage, B., Qian, J., Heroes, E., et al. (2017). Split-BioID: a proximity biotinylation assay for dimerization-dependent protein interactions. *FEBS Lett.* 591, 415–424. doi: 10.1002/1873-3468.12548

- Myers, S. A., Wright, J., Peckner, R., Kalish, B. T., Zhang, F., and Carr, S. A. (2018). Discovery of proteins associated with a predefined genomic locus via dCas9-APEX-mediated proximity labeling. *Nat. Methods* 15, 437–439. doi: 10.1038/s41592-018-0007-1
- Okuyama, K., Strid, T., Kuruvilla, J., Somasundaram, R., Cristobal, S., Smith, E., et al. (2019). PAX5 is part of a functional transcription factor network targeted in lymphoid leukemia. *PLoS Genet.* 15:e1008280. doi: 10.1371/journal.pgen.1008280
- Padrón, A., Iwasaki, S., and Ingolia, N. T. (2019). Proximity RNA labeling by APEX-Seq reveals the organization of translation initiation complexes and repressive RNA granules. *Mol. Cell* 75, 875–887.e5. doi: 10.1016/j.molcel.2019.07.030
- Qiu, W., Xu, Z., Zhang, M., Zhang, D., Fan, H., Li, T., et al. (2019). Determination of local chromatin interactions using a combined CRISPR and peroxidase APEX2 system. *Nucleic Acids Res.* 47:e52. doi: 10.1093/nar/gkz134
- Ramanathan, M., Majzoub, K., Rao, D. S., Neela, P. H., Zarnegar, B. J., Mondal, S., et al. (2018). RNA–protein interaction detection in living cells. *Nat. Methods* 15, 207–212. doi: 10.1038/nmeth.4601
- Rees, J. S., Li, X.-W., Perrett, S., Lilley, K. S., and Jackson, A. P. (2015). Selective proteomic proximity labeling assay using tyramide (SPPLAT): a quantitative method for the proteomic analysis of localized membrane-bound protein clusters. *Curr. Protoc. Protein Sci.* 80, 19.27.1–19.27.18. doi: 10.1002/0471140864.ps1927s80
- Remnant, L., Booth, D. G., Vargiu, G., Spanos, C., Kerr, A. R. W., and Earnshaw, W. C. (2019). In vitro BioID: mapping the CENP-A microenvironment with high temporal and spatial resolution. *Mol. Biol. Cell* 30, 1314–1325. doi: 10.1091/mbc.E18-12-0799
- Rhee, H.-W., Zou, P., Udeshi, N. D., Martell, J. D., Mootha, V. K., Carr, S. A., et al. (2013). Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* 339, 1328–1331. doi: 10.1126/science.1230593
- Roux, K. J. (2013). Marked by association: techniques for proximity-dependent labeling of proteins in eukaryotic cells. *Cell. Mol. Life Sci.* 70, 3657–3664. doi: 10.1007/s00018-013-1287-1283
- Roux, K. J., Kim, D. I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* 196, 801–810. doi: 10.1083/jcb.201112098
- Santin, Y. G. (2019). Uncovering the in vivo proxisome using proximity-tagging methods. *BioEssays* 41:900131. doi: 10.1002/bies.201900131
- Savitsky, P., Krojer, T., Fujisawa, T., Lambert, J.-P., Picaud, S., Wang, C.-Y., et al. (2016). Multivalent histone and DNA engagement by a PHD/BRD/PWWP triple reader cassette recruits ZMYND8 to K14ac-Rich chromatin. *Cell Rep.* 17:2724. doi: 10.1016/j.celrep.2016.11.014
- Schatz, P. J. (1993). Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in *Escherichia coli*. *Biotechnology* 11, 1138–1143. doi: 10.1038/nbt1093-1138
- Schiapparelli, L. M., McClatchy, D. B., Liu, H.-H., Sharma, P., Yates, J. R., and Cline, H. T. (2014). Direct detection of biotinylated proteins by mass spectrometry. *J. Proteome Res.* 13, 3966–3978. doi: 10.1021/pr5002862
- Schmidtman, E., Anton, T., Rombaut, P., Herzog, F., and Leonhardt, H. (2016). Determination of local chromatin composition by CasID. *Nucleus* 7, 476–484. doi: 10.1080/19491034.2016.1239000
- Schopp, I. M., Amaya Ramirez, C. C., Debeljak, J., Kreibich, E., Skribbe, M., Wild, K., et al. (2017). Split-BioID a conditional proteomics approach to monitor the composition of spatiotemporally defined protein complexes. *Nat. Commun.* 8:15690. doi: 10.1038/ncomms15690
- Shoaib, M., Kulyassov, A., Robin, C., Winczura, K., Tarlykov, P., Despas, E., et al. (2013). PUB-NChIP—“in vivo biotinylation” approach to study chromatin in proximity to a protein of interest. *Genome Res.* 23, 331–340. doi: 10.1101/gr.134874.111
- Singer-Krüger, B., Fröhlich, T., Franz-Wachtel, M., Nalpas, N., Macek, B., and Jansen, R.-P. (2019). APEX2-mediated proximity labeling resolves protein networks in *Saccharomyces cerevisiae* cells. *FEBS J.* 287, 325–344. doi: 10.1111/febs.15007
- Skene, P. J., and Henikoff, S. (2017). An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* 6:e21856. doi: 10.7554/eLife.21856
- Smith, P. A., Tripp, B. C., DiBlasio-Smith, E. A., Lu, Z., LaVallie, E. R., and McCoy, J. M. (1998). A plasmid expression system for quantitative in vivo biotinylation of thioredoxin fusion proteins in *Escherichia coli*. *Nucleic Acids Res.* 26, 1414–1420. doi: 10.1093/nar/26.6.1414
- Udeshi, N. D., Pedram, K., Svinkina, T., Fereshtehian, S., Myers, S. A., Aygun, O., et al. (2017). Antibodies to biotin enable large-scale detection of biotinylation sites on proteins. *Nat. Methods* 14, 1167–1170. doi: 10.1038/nmeth.4465
- Uusküla-Reimand, L., Hou, H., Samavarchi-Tehrani, P., Rudan, M. V., Liang, M., Medina-Rivera, A., et al. (2016). Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders. *Genome Biol.* 17:182. doi: 10.1186/s13059-016-1043-1048
- van Tienen, L. M., Mieszczanek, J., Fiedler, M., Rutherford, T. J., and Bienz, M. (2017). Constitutive scaffolding of multiple Wnt enhanceosome components by Legless/BCL9. *eLife* 6:e20882. doi: 10.7554/eLife.20882
- Veal, E. A., Day, A. M., and Morgan, B. A. (2007). Hydrogen peroxide sensing and signaling. *Mol. Cell* 26, 1–14. doi: 10.1016/j.molcel.2007.03.016
- Viita, T., Kyheröinen, S., Prajapati, B., Virtanen, J., Frilander, M. J., Varjosalo, M., et al. (2019). Nuclear actin interactome analysis links actin to KAT14 histone acetyl transferase and mRNA splicing. *J. Cell Sci.* 132:jcs226852. doi: 10.1242/jcs.226852
- Villaseñor, R., Pfaendler, R., Ambrosi, C., Butz, S., Giuliani, S., Bryan, E., et al. (2020). ChromID identifies the protein interactome at chromatin marks. *Nat. Biotechnol.* [Epub ahead of print]. doi: 10.1038/s41587-020-0434-432
- Vishnoi, N., Dhanasekaran, K., Chalfant, M., Surovstev, I., Khokha, M. K., and Lusk, C. P. (2020). Differential turnover of Nup188 controls its levels at centrosomes and role in centriole duplication. *J. Cell Biol.* 219:e201906031. doi: 10.1083/jcb.201906031
- Weina, K., and Utikal, J. (2014). SOX2 and cancer: current research and its implications in the clinic. *Clin. Transl. Med.* 3:19. doi: 10.1186/2001-1326-3-19
- Xie, W., Chojnowski, A., Boudier, T., Lim, J. S. Y., Ahmed, S., Ser, Z., et al. (2016). A-type lamins form distinct filamentous networks with differential nuclear pore complex associations. *Curr. Biol.* 26, 2651–2658. doi: 10.1016/j.cub.2016.07.049
- Xue, M., Hou, J., Wang, L., Cheng, D., Lu, J., Zheng, L., et al. (2017). Optimizing the fragment complementation of APEX2 for detection of specific protein-protein interactions in live cells. *Sci. Rep.* 7:12039. doi: 10.1038/s41598-017-12365-12369
- Youn, J.-Y., Dunham, W. H., Hong, S. J., Knight, J. D. R., Bashkurov, M., Chen, G. I., et al. (2018). High-Density proximity mapping reveals the subcellular organization of mRNA-Associated granules and bodies. *Mol. Cell* 69, 517–532.e11. doi: 10.1016/j.molcel.2017.12.020
- Zasadzińska, E., Huang, J., Bailey, A. O., Guo, L. Y., Lee, N. S., Srivastava, S., et al. (2018). Inheritance of CENP-A nucleosomes during DNA replication requires HJURP. *Dev. Cell* 47, 348–362.e7. doi: 10.1016/j.devcel.2018.09.003
- Zhang, P., Sun, Y., and Ma, L. (2015). ZEB1: at the crossroads of epithelial-mesenchymal transition, metastasis and therapy resistance. *Cell Cycle* 14, 481–487. doi: 10.1080/15384101.2015.1006048
- Zhang, S., Williamson, N. A., and Bogoyevitch, M. A. (2018). Complementary proteomics strategies capture an ataxin-1 interactome in Neuro-2a cells. *Sci. Data* 5, 1–6. doi: 10.1038/sdata.2018.262
- Zhou, Y., Wang, G., Wang, P., Li, Z., Yue, T., Wang, J., et al. (2019). Expanding APEX2 substrates for spatial-specific labeling of nucleic acids and proteins in living cells. *Angew. Chem. Int. Ed.* 58, 11763–11767. doi: 10.1002/anie.201905949
- Zhu, C., Li, L., Zhang, Z., Bi, M., Wang, H., Su, W., et al. (2019). A non-canonical role of YAP/TEAD is required for activation of estrogen-regulated enhancers in breast cancer. *Mol. Cell* 75, 791–806.e8. doi: 10.1016/j.molcel.2019.06.010

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Ummethum and Hamperl. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Exploring the Histone Acetylation Cycle in the Protozoan Model *Tetrahymena thermophila*

Suzanne Wahab, Alejandro Saettone, Syed Nabeel-Shah[†], Nora Dannah and Jeffrey Fillingham^{*}

Department of Chemistry and Biology, Ryerson University, Toronto, ON, Canada

OPEN ACCESS

Edited by:

Christoph Franz Kurat,
Ludwig Maximilian University
of Munich, Germany

Reviewed by:

Stephan Hamperl,
Helmholtz Zentrum München,
Germany
Yifan Liu,
University of Michigan, United States

*Correspondence:

Jeffrey Fillingham
jeffrey.fillingham@ryerson.ca

[†] Present address:

Syed Nabeel-Shah,
Donnelly Centre, University
of Toronto, Toronto, ON, Canada;
Department of Molecular Genetics,
University of Toronto, Toronto, ON,
Canada

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 20 April 2020

Accepted: 28 May 2020

Published: 30 June 2020

Citation:

Wahab S, Saettone A,
Nabeel-Shah S, Dannah N and
Fillingham J (2020) Exploring
the Histone Acetylation Cycle
in the Protozoan Model *Tetrahymena*
thermophila.
Front. Cell Dev. Biol. 8:509.
doi: 10.3389/fcell.2020.00509

The eukaryotic histone acetylation cycle is composed of three classes of proteins, histone acetyltransferases (HATs) that add acetyl groups to lysine amino acids, bromodomain (BRD) containing proteins that are one of the most characterized of several protein domains that recognize acetyl-lysine (Kac) and effect downstream function, and histone deacetylases (HDACs) that catalyze the reverse reaction. Dysfunction of selected proteins of these three classes is associated with human disease such as cancer. Additionally, the HATs, BRDs, and HDACs of fungi and parasitic protozoa present potential drug targets. Despite their importance, the function and mechanisms of HATs, BRDs, and HDACs and how they relate to chromatin remodeling (CR) remain incompletely understood. *Tetrahymena thermophila* (Tt) provides a highly tractable single-celled free-living protozoan model for studying histone acetylation, featuring a massively acetylated somatic genome, a property that was exploited in the identification of the first nuclear/type A HAT Gcn5 in the 1990s. Since then, *Tetrahymena* remains an under-explored model for the molecular analysis of HATs, BRDs, and HDACs. Studies of HATs, BRDs, and HDACs in *Tetrahymena* have the potential to reveal the function of HATs and BRDs relevant to both fundamental eukaryotic biology and to the study of disease mechanisms in parasitic protozoa.

Keywords: histone acetyltransferase, bromodomain, chromatin remodeling, histone deacetylase *Tetrahymena thermophila*, *Tetrahymena*

INTRODUCTION/BACKGROUND

Chromatin Remodeling

Eukaryotic cells package their genomic DNA into chromatin. The basic unit of chromatin, the nucleosome, consists of a histone octamer of four core histones: H2A, H2B, H3, and H4 (Luger et al., 1997). Histone variants are able to substitute for the core canonical histones within the nucleosomes and often confer specific structural and functional features (Henikoff and Smith, 2015). Additional factors, such as linker histones, further organize nucleosomes into higher-order chromatin structures (Fyodorov et al., 2018). Chromatin ultimately needs to be remodeled for DNA transactions such as transcription to occur (Clapier and Cairns, 2009). Mechanisms of chromatin remodeling (CR) involve ATP-dependent histone sliding [e.g., SWI/SNF (Johnson et al., 2005), INO80 (Poli et al., 2017), and ISWI (Yan et al., 2016)] and the selective insertion/removal of histone variants [SWR (Morrison and Shen, 2009) and INO80 (Brahma et al., 2017)], as well as the post-translational modification (Gardner et al., 2011) (PTM) of specific amino acids within histones including lysine acetylation and methylation.

Histone PTMs can lead to downstream events via recruitment of proteins with specific PTM-interacting domains including the bromodomain (BRD) that interacts with acetylated lysines (Filippakopoulos et al., 2012).

The Histone Acetylation Cycle and Its Relevance to Human Disease

The histone acetylation cycle begins with the selective addition of an acetyl group to a specific lysine residue in histones, a process known as histone acetylation, and is catalyzed, or “written,” by histone acetyltransferases [HATs (Berndsen and Denu, 2008)]. HATs can be guided to their specific histone targets by physically interacting with proteins containing histone-interacting domains (Lalonde et al., 2014). Bromodomain proteins are able to interact, or “read” acetyl-lysine (Kac), and the cycle is complete when the acetyl is removed, or “erased,” by histone deacetylases (HDACs) (Kuo and Allis, 1998). Histone acetylation occurs either at the nucleosomal level by type A HATs [SAGA (Baker and Grant, 2007) and NuA4 (Doyon and Cote, 2004) complexes] or on histones prior to their deposition into chromatin by type B HATs [Hat1 (Parthun, 2007), Rtt109 (Fillingham et al., 2008)]. Although the focus of this review is histone acetylation, it is important to note that proteomic studies have identified hundreds to thousands of acetylated proteins in a variety of model systems from parasitic protozoa to mammalian cells (Zhao et al., 2010; Jeffers and Sullivan, 2012; Miao et al., 2013; Li et al., 2019). To reflect this, HATs and HDACs that acetylate/deacetylate the group of lysine residues on non-histone substrates are also be referred to as lysine acetyltransferases/deacetylases (KATs/KDACs) (Allis et al., 2007).

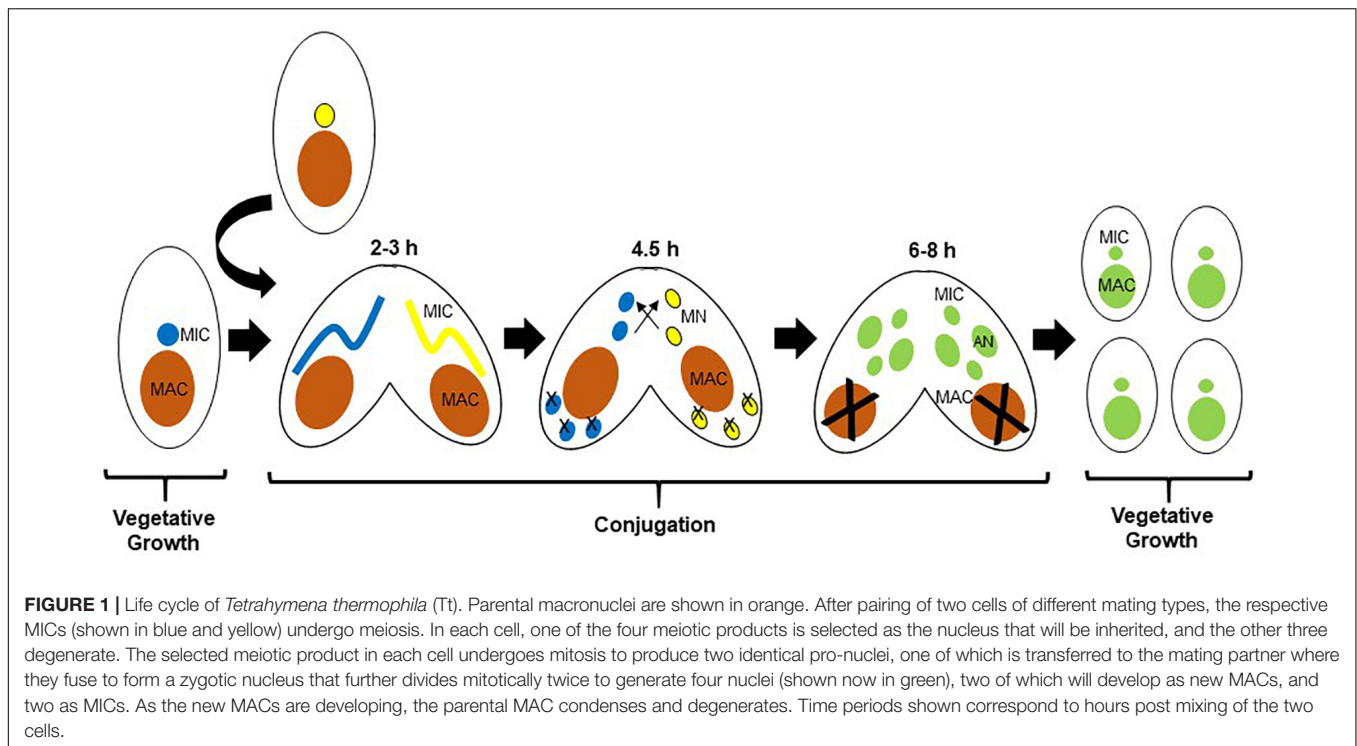
The proteins of the histone acetylation cycle have clinical significance in cancer (Somech et al., 2004; Jain and Barton, 2017; Richters and Koehler, 2017) and have attracted interest as potential druggable targets. For example, translocation of BRD-containing BRD4 to NUTM1 in human cells generates an oncoprotein that drives a rare and aggressive form of squamous cell carcinoma, NUT midline carcinoma (NMC) (French et al., 2004). Multiple small molecules have been developed that disrupt BRD-Kac interactions (Cochran et al., 2019) and are subject of investigation for their efficacy in treating cancers such as NMC. HATs, BRDs, and HDACs are often required for the viability and pathogenesis of parasitic protozoa (Jeffers et al., 2017) which are among the Neglected Tropical Diseases (NTD) common in regions of Africa, Asia, and the Americas (World Health Organization, 2010). Anti-protozoa drugs exist but their efficacy is being compromised by rising resistance (de Koning, 2017). The availability of small molecule inhibitors to proteins of the histone acetylation cycle has driven interest in their use to treat these diseases. The Alveolate lineage of eukaryotes includes the parasitic apicomplexa, with *Plasmodium* and *Toxoplasma* species. Gcn5 was shown to be essential for *Toxoplasma* DNA replication, prompting a search for drugs that target this HAT (Vanagas et al., 2012; Jeffers et al., 2016). In *Plasmodium*, PfBDP1 containing a C-terminal BRD and an N-terminal ankyrin repeat is required for penetration of red blood cells (Josling et al., 2015),

potentially providing a drug target. *Plasmodium* HDACs are also being investigated as possible drug targets (Andrews et al., 2012).

Tetrahymena thermophila as a Model for the Study of the Histone Acetylation Cycle

The ciliate *Tetrahymena thermophila* (Tt), a free-living genetically tractable Alveolate, has been a beneficial model for early studies of the fundamental biology of histone acetylation due to its unique nuclear biology (Grunstein and Allis, 2018). Tt is suitable to study apicomplexan biology as well as that of other protozoan parasites due to their evolutionary relationship. Tt is also a proven model for discovery-based chromatin biology, based in part on the biology of the ciliates, which segregate germline-specific silent, and somatic transcriptionally active chromatin into two distinct nuclei (Orias et al., 2011). The micronucleus (MIC) is diploid, divides by mitosis, and is not transcribed during growth. The MIC undergoes meiosis during conjugation, the sexual phase of the life cycle, and is analogous to a germline nucleus. The macronucleus (MAC) is highly polyploid with ~45 copies of each MAC chromosome, divides amitotically without functional centromeres, is transcriptionally active, and is not inherited sexually; analogous to a somatic nucleus. During amitosis, multiple copies of each macronuclear chromosome are randomly partitioned between the two daughter cells. As a result, alleles segregate randomly, and therefore vegetative progeny of a cell initially heterozygous after conjugation become homozygous for one of the alleles after a number of cell divisions in a process called phenotypic assortment (Orias and Flacks, 1975). The two nuclei are related to each other through the sexual life cycle, conjugation (Martindale et al., 1982), the milestones of which are shown in **Figure 1**. MAC development includes a variety of programmed DNA rearrangements that includes chromosome fragmentation, and programmed deletion of DNA sequence called internal eliminated sequences (IESs) which is epigenetically regulated in a process that is initiated by genome-wide transcription of non-coding RNAs (Chalker and Yao, 2001) (ncRNAs) from the normally silent MIC. The ncRNAs then direct RNAi-dependent assembly of distinct chromatin domains in the new MAC (Taverna et al., 2002; Liu et al., 2007), a prelude to DNA deletion (Yao et al., 2015) thought to be similar to the chromatin diminution seen in the parasitic nematode *Ascaris* (Wang et al., 2017).

Tetrahymena thermophila has features of a model genetic organism including fast growth in axenic culture, and the ability to undergo large scale and synchronous matings. Both the MAC (Eisen et al., 2006) and MIC (Hamilton et al., 2016) genomes have been sequenced and annotated (Stover et al., 2006). A well-developed functional proteomic pipeline exists for the study of epigenetic regulators in particular for affinity purification coupled to mass spectrometry (AP-MS) to effectively solubilize native (i.e., un-crosslinked) protein complexes both bound and unbound to chromatin (Xiong et al., 2011; Garg et al., 2013, 2019; Saettone et al., 2018, 2019a; Ashraf et al., 2019; Nabeel-Shah et al., 2020). A critical feature of this approach is the expression of the epitope tagged proteins



at levels closely approximating that of the endogenous protein. In *Tetrahymena* this is achieved by exact gene replacement mediated by homologous recombination, adding the epitope tag in-frame at the C-terminus of the endogenous protein, such that the epitope tagged polypeptide is expressed under its own promoter, as in budding yeast. Physical Interactome mapping experiments are performed using a minimum of two biological replicates in parallel to control experiments using untagged parental strains, facilitating the identification of interaction partners significantly over-represented in the samples, a process aided by use of algorithms such as SAINTexpress (Teo et al., 2014). Functional genomic approaches that have been developed for Tt include definition of the transcriptome through microarray (Miao et al., 2009) and RNA-Seq analysis (Xiong et al., 2011). Because Tt gene expression can be assessed in a variety of developmental stages, network analysis of transcriptome data can be used to predict functional relationships between genes (Xiong et al., 2011). In addition, genomic localization of proteins involved in the acetylation cycle can be tracked using ChIP-Seq (chromatin immunoprecipitation followed by next generation sequencing) (Saettone et al., 2018, 2019b).

HISTONE ACETYLATION IN *Tetrahymena thermophila*

In addition to their distinct morphologies and functions, differences exist in the complement of chromatin proteins as well as the degree of histone acetylation in the MAC and MIC. Although the same core histones (H2A, H2B, H3, and H4) are

present in both nuclei, the MIC possesses two versions of histone H3 (Allis et al., 1979). One of which is indistinguishable from that found in the MAC (named H3S for slow), and the other is unique to MIC, and has a faster mobility in SDS-PAGE (and thus named H3F for fast) as a consequence of a regulated proteolytic event where six amino acids are removed from the N-terminus (Allis et al., 1980a; **Figure 2A**); the underlying enzymology of which is unknown. Nucleus-specific linker histones exist, Hho1 for the MAC and Mlh1 (Micronuclear linker histone 1) for the MIC. Mlh1 in particular is proteolyzed (Allis et al., 1984) into several smaller polypeptides (alpha, beta, gamma, and delta). Although HHO and MLH are non-essential genes, DAPI staining combined with knockout analysis showed that both function in the condensation of chromatin in their respective nuclei (Shen et al., 1995). The MIC also features the exclusive localization of CNA1 (Cervantes et al., 2006; Cui and Gorovsky, 2006), an ortholog of the centromeric-specific H3 variant CENPA, consistent with apparent lack of centromeres in the MAC. Transcription associated histone variants Hv1 (H2A.Z) and Hv2 (H3.3) are MAC-specific (Allis et al., 1980b) during growth and starvation, appearing in the MIC only during selected times in conjugation (Stargell et al., 1993; Cui et al., 2006). In addition to being widely distributed in the MAC by immunofluorescence (IF), Hv1 localizes to the nucleoli indicating that it may be involved in rDNA transcription (Allis et al., 1982). The use of specific anti-Hv1 antibodies in indirect IF indicates that Hv1 may present in the crescent MIC, a time when it is transcriptionally active (Stargell et al., 1993). Hv2 differs in 16 amino acids from the major, abundant H3 proteins and is expressed constitutively (Thatcher et al., 1994). Genetic analysis suggests that the primary importance of HHT3 in growing cells is a consequence of

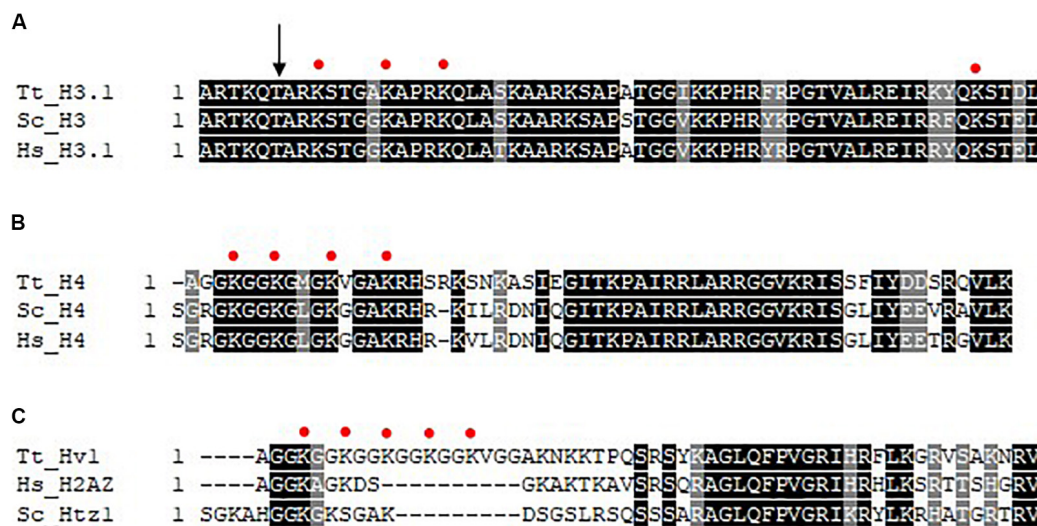


FIGURE 2 | Multiple sequence alignment of histone N-termini. Clustal Omega was used to perform multiple sequence alignment (EMBL-EBI) and shaded using BoxShade (ExPASy). Absolutely conserved residues are represented by a black shade and conserved residues are represented by a gray shade. Red circles represent sites of acetylation as indicated in the text. **(A)** N-termini of histone H3. The arrow indicates site of N-terminal proteolysis. **(B)** N-termini of histone H4. **(C)** N-termini of histone H2A variant Hv1/H2AZ. Tt, *Tetrahymena thermophila*; Hs, *homo sapiens*; Sc, *Saccharomyces cerevisiae*.

its constitutive expression rather than its primary sequence (Yu and Gorovsky, 1997).

Some of the earliest observations linking histone acetylation and transcription occurred using the Tt model. Pulse-labeling studies using radiolabeled acetate was used to show that histones in general are highly acetylated in the MAC but not MIC in growing cells (Vavra et al., 1982a,b), a finding that provided a key functional link of histone acetylation to gene expression. Initial studies on histone deposition-associated acetylation focused on pulse-labeling cells with radiolabeled acetate early in conjugation where MICs replicate rapidly, but are transcriptionally inactive, thus any MIC-specific histone acetylation at this time is related functionally to histone deposition and chromatin assembly. Initial studies showed that newly synthesized histone H3 and histone H4 were deposited into MIC in mono- and di-acetylated (H3) and di-acetylated (H4) forms (Allis et al., 1985). Later work identified these deposition-associated acetylation sites on K9/14 on H3 (Sobel et al., 1995; **Figure 2A**) and H4 K4/11 (Chicoine et al., 1986; **Figure 2B**, note corresponds to H4K5/12 in yeast and human cells: Tt H4 contains a single deletion at amino acid 3, so each acetylation site is numbered one less than that observed in most other H4s). Similar studies on pulse-labeled MACs isolated during growth revealed H4K7 (**Figure 2B**, corresponds to H4K8) and H4K4 to be the principle sites of acetylation correlated with transcription on H4, with H4K11 and H4K15 also acetylated at lower levels (Chicoine et al., 1986; **Figure 2A**). Transcription-associated sites of acetylation on H3 were revealed to be H3K9 and H3K14 and to a lesser extent H3K18 and H3K23 (Chicoine et al., 1986; **Figure 2B**). Subsequent studies found Hv1 to be acetylated at positions 4, 7, 10, 13, 16 in its N-terminus (Allis et al., 1986; **Figure 2C**). When indirect IF was performed with newly developed anti-tetra-acetylated H4 and anti-penta acetylated Hv1 antibodies, MAC localization

was observed (Lin et al., 1989). When studying the extent of histone acetylation during conjugation, only a slight increase in acetylation was shown in the early developing new MACs compared to the developing micronuclei at the same stage. A greater amount of acetylation is detected in the advanced stage of anlagen development comparable to the parental macronuclei (Chicoine and Allis, 1986) suggesting that there is modulation of the histone acetylation cycle during nuclear development.

HISTONE ACETYL-TRANSFERASES IN *Tetrahymena thermophila*

Histone acetyltransferases are responsible for catalyzing the transfer of acetyl groups from acetyl-CoA onto lysine residues on the amino-termini of histones. There are five genes that encode clear orthologs of HATs [GCN5/HAT2, HAT1, and 3 MYST-family HATs (MYST1-3)].

Gcn5/SAGA

The search for the gene encoding a HATs was hampered by relatively low amounts of the enzyme that made it difficult to obtain peptide sequence. For example a HAT activity was identified and characterized in yeast in the early 1980s (Travis et al., 1984), but the gene responsible was not cloned (No Author, 2018). The massive amount of acetylated chromatin in the Tt MAC was key to finding the first gene encoding a HATs. A novel SDS-PAGE acetyltransferase activity assay was used to show that the Tt MAC possess a 55 kDa protein (p55) able to incorporate [3H] acetate from [3H] acetyl-coA into a histone H3 substrate (Brownell and Allis, 1995). After partial purification using the in gel-assay to monitor the activity, Edman degradation and subsequent molecular

cloning followed by comparative sequence analysis determined that p55 is orthologous to Gcn5 (Brownell et al., 1996), a transcriptional adaptor, or co-activator previously described as necessary for activity of transcriptional activators in yeast (Georgakopoulos and Thireos, 1992). This finding established the mechanistic link between chromatin structure and gene expression, and reinforced the idea of histone acetylation as a mark of transcriptionally active chromatin. Gcn5 is broadly conserved in eukaryotes present in most if not all sequenced eukaryotes including *Toxoplasma gondii* (Wang et al., 2014) and *Plasmodium falciparum* (Fan et al., 2004). In yeast and human cells, Gcn5 is found in the multi-subunit complex SAGA (Grant et al., 1997), a transcriptional co-activator complex containing ~19 subunits (Allard et al., 1999; Daniel and Grant, 2007) that, as a type A HAT, targets the nucleosomal N-terminal tail of H3 (Grant et al., 1997, 1998, 1999). The original type A HAT, Gcn5p, may also possess type B HAT activity in *Saccharomyces cerevisiae* as it is involved in the acetylation of the NH2-terminal tail of newly synthesized histone H3 (Burgess et al., 2010).

Details on the function and mechanism of Tt Gcn5 have lagged since its initial discovery. Recombinant Tt Gcn5 acetylates free histone H3 including on H3K13 and H3K18 (Garg et al., 2013; **Figure 2A**). The questions as to the nature of Tt SAGA was recently addressed (Saettone et al., 2018) by AP-MS of Ibd1 (see section “Group II BRD Proteins,” below) that revealed p55/Gcn5 to be present in a complex with a clear ortholog of Ada2, as well as the BRD-containing Ibd1, and PHD-domain containing protein *Ada2-Associated Protein 1* (AAP1). One function of PHD domains is to interact with methyl-lysine PTMs (Arrowsmith and Schapira, 2019). Subsequent AP-MS of Ada2 reciprocally identified Gcn5, Ibd1, Aap1 and three additional PHD domain-containing proteins (Aap2-4) that were not observed in repeated Ibd1 AP-MS. Consistent with this, a SAGA-like complex consisting of orthologs of Gcn5, Ada2, and a PHD domain protein (PHD1) in addition to a protein of unknown function, was purified from Apicomplexan *P. falciparum* by incubation of extracts with a biotinylated H3K4me3 peptide (Hoeijmakers et al., 2019). The same complex minus the protein of unknown function co-purified when extracts were incubated with an H4K5/8/12ac peptide. Additionally, AP-MS of PfGcn5 revealed an interaction with an additional PHD-domain-containing protein PHD2, purification of which enriched Gcn5 and Ada2 but not PHD1. The similarities between Gcn5 membership in multiple protein complexes in Tt and Pf suggest that multiple SAGA-like complexes exist in Alveolates that are composed of a “core” of Ada2 and Gcn5 with different epigenetic readers that in Tt and Pf are composed of PHD and BRD proteins. To demonstrate this conclusively, AP-MS of Tt Aap2-4 will be required, the prediction being that each will co-purify with Ada2 and Gcn5 but none of the other readers. If this is the case, it will be important to determine chromatin-binding specificity of the respective PHD fingers of the AAPs. BLASTP analysis indicates that AAP1 is the highest match in the Tt genome of Pf PHD1 with high homology between the two in their fourth PHD domain of PHD1 that matches consensus for H3K4me2/3-binding.

Hat1

Hat1 was originally purified from yeast cytoplasmic extracts in a complex with Hat2 (Rbap46 in mammalian cells). Hat1 is a type B HAT, highly specific for H4K5 and H4K12 on free histones in yeast and humans (Parthun, 2007). It encodes a clear ortholog of HAT1, but to date it has not been subjected to molecular analysis. Experiments performed by Allis and colleagues showed MIC and cytoplasmic extracts of growing Tt to possess a HAT activity on Tt H4 (Richman et al., 1988) that had specificity for position 4 or positions 4 and 11 (Richman et al., 1988; **Figure 2B**, the sites that correspond to H4K5/12 in Tt). The same activity did not acetylate mononucleosomes, consistent with Hat1 activity in yeast and human cells (Parthun et al., 1996; Verreault et al., 1998). Based on work performed in yeast and human cells, it would make sense that this activity was performed by the Hat1 complex (Parthun et al., 1996), composed of the Hat1 HAT bound to the Hat2/RbAp46 histone chaperone (Parthun, 2013). The expression of HAT1 is essential in human cells (Nagarajan et al., 2013). Interestingly, when the cytoplasmic activity was heated to 45°C, the Hat1-like activity on H4 was retained, but now the activity also mono-acetylated H3 at an unknown lysine residue (Richman et al., 1988). It will be informative to characterize the Hat1 complex in Tt and to determine whether it has activity on H3 as well as H4.

NuA4/Esa1

NuA4/TIP60 are multi-subunit type A HAT complexes in yeast and human cells responsible for acetylation of nucleosomal histone H4 and H2A with catalytic subunit Esa1/Tip60 (Doyon and Cote, 2004; Jacquet et al., 2016). Analogous H2A/H4-specific Hat A activities are more poorly characterized outside the Opisthokonts. In Tt, a H2A/H4-specific nucleosomal HAT activity was previously partially purified from MAC DNase-treated extracts, and labeled NuA4-like (Ohba et al., 1999). Yeast/human NuA4/TIP60 complexes and the Tt NuA4-like activity both have specificity for lysines 5, 8, 12, and 16 of H4 and lysines 5 and 9 of H2A on nucleosomes (Ohba et al., 1999), suggesting the activity could be catalyzed by an analogous Tt complex. However, the Tt NuA4-like activity co-purified on a sucrose gradient with a predicted size of ~80 kD which is much smaller than that of NuA4/TIP60 complex, a 1.0–1.5 MDa multi-protein platform of at least 13–16 subunits. Comparative sequence analysis of the MAC genome suggests that there are three potential Tt genes encoding an ortholog of Esa1/Tip60, named MYST1-3, with MYST1 and MYST2 situated side by side in the MAC genome, possibly the result of a tandem duplication. Yeast also encode three MYST family HATs, each of which nucleates a distinct HAT complex with non-overlapping functions (Esa1-NuA4, Sas2 – SAS complex and Sas3 of the NuA3 complex). The parasitic protozoa *Trypanosoma brucei* (*T. brucei*) encodes three MYST family proteins named HAT1-3 that all localize to its nucleus (Kawahara et al., 2008). *P. falciparum* encodes a single gene encoding a MYST family HAT (Miao et al., 2010). Clear orthologs of genes encoding core NuA4/Tip60 proteins Eaf1 and Epl1 are not present in the Tt MAC genome or that of other Alveolates, consistent with the idea that a NuA4 complex is not well conserved outside the Opisthokonts.

Identification of the Tt MYST HAT underlying the NuA4-like activity, and identification of co-purifying proteins, is likely to inform NuA4/TIP60 characterization in the parasitic protozoa.

H3K56 Acetylation in *Tetrahymena thermophila*

H3K56ac is associated with DNA replication associated chromatin assembly, gene expression, and maintenance of genome stability in yeast and human cells (Masumoto et al., 2005; Das et al., 2009). H3K56 is conserved in Tt H3 (**Figure 2A**), and the MAC possesses robust levels of H3K56ac during growth (Garcia et al., 2007) and early nuclear development (Akematsu et al., 2017). Although not widely studied outside of Fungi and human cells, H3K56ac has been reported to be present in parasitic protozoa such as *P. falciparum* (Gupta et al., 2016). H3K56ac is catalyzed by Rtt109 in Fungi (Collins et al., 2007; Driscoll et al., 2007; Fillingham et al., 2008) and p300/CBP in humans (Das et al., 2009). In Fungi and human cells, histone H3/H4 chaperone ASF1 is also required to catalyze H3K56ac (Recht et al., 2006; Das et al., 2009) by the respective HAT. Tt does not encode a clear ortholog of either Rtt109 or p300/CBP but does encode a single copy of ASF1 (Garg et al., 2013). Rtt109 has drawn recent interest as a possible drug target to combat pathogenic fungal infection due to its fungal-specific nature and importance to viability (Wurtele et al., 2010). Despite their non-homologous primary amino acid sequence, Rtt109 and p300/CBP have structural similarity of their catalytic core (Bazan, 2008). Human Gcn5 has been reported to acetylate H3K56 in human cells (Tjeertes et al., 2009), but not in yeast. Recombinant Tt p55/Gcn5 does not possess H3K56ac activity *in vitro* on core histone substrates in the presence or absence of recombinant yeast or Tt Asf1 (Garg et al., 2013) which argues against Gcn5 being the H3K56-specific HAT in Tt. It should be noted that the HAT assay was performed with chicken and not Tt histones, and that in Tt Gcn5 exists in a protein complex with an Ada2 ortholog (see section “Gcn5/SAGA”), so its behavior *in vitro* may reflect absence of key components. Functionally, the importance of the modification to growth/genome stability in Tt has yet to be determined, or even if it is linked to chromatin assembly as in yeast. Arabidopsis have also been demonstrated to encode orthologs of p300/CBP suggesting that p300/CBP was present in last common ancestor of plants and Opisthokonts, which should also include protist lineages. If H3K56ac is important to protozoan viability, the H3K56-specific HAT (particularly if novel) may have potential as a drug target for treatment of parasitic protozoa infection.

BROMODOMAIN PROTEINS IN *Tetrahymena thermophila*

There are several protein domains that selectively recognize and bind to acetylated Lysine (Kac) residues in histones (Jain and Barton, 2017) including PHD (Zeng et al., 2010), YEATS (Li et al., 2014) and the BRD that is the focus of this section. BRD-containing proteins are frequently dysregulated

in cancer (Jain and Barton, 2017) and their expression has been demonstrated to be important for pathogenesis of several parasitic protozoa (Schulz et al., 2015). Importantly, BRDs can be targeted by small-molecule inhibitors leading to the idea they can be targeted to control cancer and/or infection by parasitic protozoa (Jeffers et al., 2017; Kougnassoukou Tchare et al., 2019; Hanquier et al., 2020; Nguyen et al., 2020). The polyploid MAC of Tt is massively enriched for acetylated chromatin which makes it a potentially useful system for discovering new BRD functions. A BLASTP-based survey of the Tt MAC genome revealed 14 potential BRD-containing gene products (Saettone et al., 2018). Tt BRD proteins present one BRD per protein which is different to other described eukaryotes where BRD proteins often are present in tandem. The BRD is composed of four helices with the ZA and BC loops connecting helices αZ to αA , and αB to αC . The ZA and BC loops form a hydrophobic pocket that functions in Kac binding (Dhalluin et al., 1999). The ZA and BC loops interact with residues flanking the Kac and are somewhat variable in sequence, reflecting the fact that different BRDs have distinct lysine acetylation sites in histones and non-histone proteins (Zaware and Zhou, 2019). A structure-based multiple sequence alignment of the 14 Tt BRDs (**Figure 3**) shows conservation of BRD secondary structure. BRD-based Kac recognition involves hydrogen bond formation with a conserved Asparagine in the BC loop (Zaware and Zhou, 2019) which is conserved in all 14 Tt BRD proteins (**Figure 3**). Phylogenetic analysis was performed using the 14 Tt BRD sequences (Saettone et al., 2018) permitting their delineation into three main groups which is shown along with domain analysis of the respective full length protein in **Figure 4**.

Group I BRD Proteins

Group I is composed of two proteins, the first a clear ortholog of Mixed-Lineage Leukemia 1 (MLL1), and BroP3 (**Figure 4**). Mammalian MLL1 is a histone methyl-transferase (KMT2A) that positively regulates transcription by tri-methylation of H3K4 (H3K4me3) and is considered an ortholog of yeast Set1. Both mammalian MLL1 and yeast Set1 co-purify with a related protein complex, COMPASS (Shilatifard, 2012). A previous study demonstrated that hyper-acetylated MAC-specific histone H3 tends to be rich in H3K4me3 (Taverna et al., 2007). It seems possible that in Tt, highly expressed genes are rich in both modifications. Our sequence alignment indicates that the BRD of Tt MLL1 [unlike that of human (Filippakopoulos et al., 2012)] is predicted to have Kac-binding activity (**Figure 3**) but its relatively divergent ZA loop sequence (**Figure 3**) raises the possibility of a non-histone target. It is tempting to speculate that the BRD of MLL1 recruits histone methyltransferases (HMTs) to add the H3K4me3 modification to acetylated chromatin, promoting high levels of transcription. Interestingly, GroupII BRD protein Ibd1 co-purifies with another putative H3K4me3 HMT that does not itself possess a BRD (see below Section “Group II BRD Proteins”). BroP3, the other group I member, has a simple domain structure similar to Ibd1 (Section “Group II BRD Proteins”) with a single N-terminal BRD.

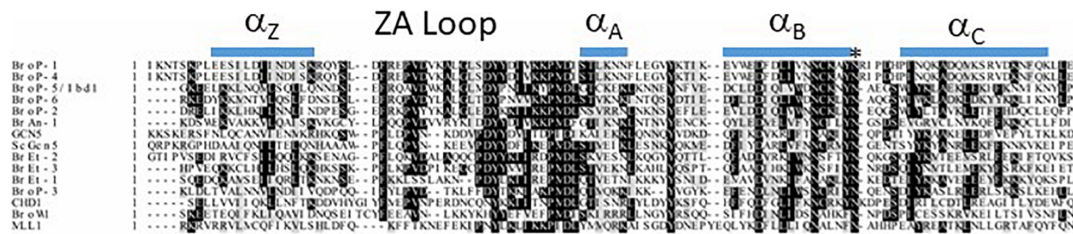


FIGURE 3 | Multiple sequence alignment of 14 Tt bromodomains (BRD). BRD sequence was extracted from the respective full length sequence using SMART. Clustal Omega was used to perform multiple sequence alignment (EMBL-EBI) and shaded using BoxShade (ExPASy) with absolutely conserved residues indicated by a black shade and conserved residues by a gray shade. The asterisk represents the highly conserved Asparagine (N) that makes contact with acetyl-lysine. Predicted secondary structure (JPRED) is shown along the top of the alignment.

Group II BRD Proteins

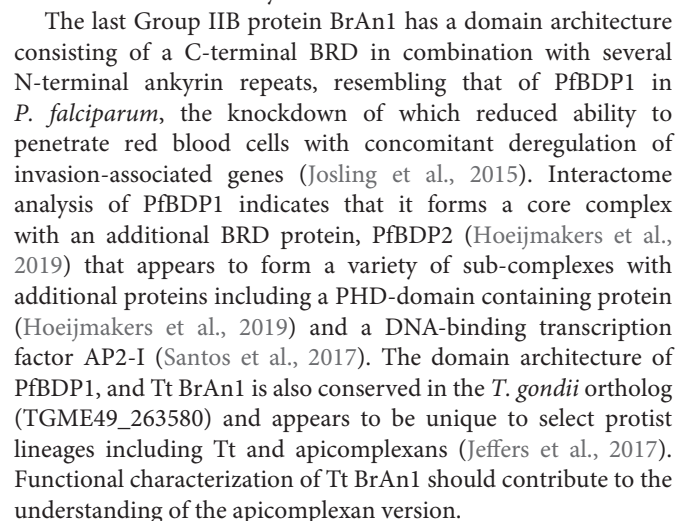
Group II is subdivided into two principal divisions (**Figure 4**). Group IIA includes chromodomain helicase protein (CHD), with BRD situated between two N-terminal chromodomains, in addition to BroW1, possessing several WD-repeats and a C-terminal BRD. The MIC copy of the gene encoding BroW1 includes an IES (Hamilton et al., 2016). IES are rarely found in MIC sequence corresponding to coding sequence in the MAC due to their imprecise mechanism of elimination.

Group IIB includes BrAn1, in addition to several proteins related by their simple domain organization, BroP1, BroP2, BroP4, BroP5, and BroP6 (**Figure 4**). BroP5 (also known as Ibd1 -Interactive BRD protein 1) is the best characterized BRD protein to date in Tt, initially identified as a member of the Tt SWI/SNF CR complex via AP-MS of the conserved Snf5 subunit (Saettone et al., 2018). Reciprocal AP-MS experiments revealed Ibd1 to also be a member of SWR, a SAGA-like and an HMT-containing complex (Saettone et al., 2018). Incubation of recombinant Ibd1 with a histone post-translation modification peptide array (Saettone et al., 2018) suggest that its BRD recognizes specific histone acetylation associated with the transcriptionally active MAC, H3K9/14ac, and H4K8ac (Chicoine et al., 1987; Taverna et al., 2007). Indirect IF of Ibd1 showed localization to the MAC throughout the Tt life cycle consistent with a role for Ibd1-containing CR complexes in the regulation of transcription (Saettone et al., 2018). Support for this hypothesis was provided by ChIP-Seq analysis of Ibd1 that revealed enrichment in the coding regions of highly expressed genes during growth leading to the proposal of the “pile-on” model where histone acetylation leads to Ibd1-dependent recruitment (or “piling on”) of multiple CR activities that each individually contribute to high levels of transcription. To test the model, it will be important to assess the degree to which Ibd1 recruits each of the respective CR complex to the set of Ibd1-bound highly expressed genes. Also, although the model proposes an additive effect for each CR complex, it is possible that redundancies exist among them. The CR complexes hypothesized to be recruited by the Kac-binding of Ibd1 BRD include SWR, SWI/SNF, SAGA and a putative HMT.

The putative HMT that co-purifies with Ibd1 is uncharacterized, but sequence analysis indicates that it has homology to HMTs that have specificity for H3K4 or H3K36,

both modifications associated with transcription. As discussed above for MLL, the association of Ibd1 and the HMT potentially could link transcription-associated histone acetylation to H3K4me3, providing a molecular mechanism behind the observation that hyper-acetylated MAC-specific histone H3 tends to be rich in H3K4me3 (Taverna et al., 2007). The SWI/SNF complex is an ATP-dependent CR complex nucleated by SNF2/Brg1 ATPase subunit. The human Brd9 protein, similar to Ibd1 with only one BRD, was recently shown to be a member of human SWI/SNF complex (Wang et al., 2019). Unlike yeast and human cells, Tt Brg1/SNF2 does not possess a C-terminal BRD (Fillingham et al., 2006). The specificity of the Ibd1 BRD appears similar to that of the human and yeast Brg1/Snf2 subunit. The biochemical function of SWI/SNF is ATP-dependent remodeling of nucleosome structure by mobilizing nucleosomes by sliding and/or ejection of histone octamers (Saha et al., 2006). We speculate that the physical interaction of Ibd1 with SWI/SNF links histone acetylation to nucleosome sliding or ejection. Interestingly the association between Ibd1 and SWI/SNF appears much less stable during conjugation (Saettone et al., 2018) indicating that the physical interaction between the two may be subject to regulation. The function of SWR-C in yeast (SRCAP in humans) is the deposition of Htz1/H2AZ (Kobor et al., 2004; Ruhl et al., 2006). Previously the Allis lab demonstrated that Hv1, like Ibd1, localizes to the MAC of growing cells implicating Hv1 in transcription (Wenkert and Allis, 1984). Tt SWR-C was defined by the intersection of common proteins in Ibd1 and Swc4 AP-MS (Saettone et al., 2018; Ashraf et al., 2019) with clear orthologs identified for Swr1, Yaf9, Rvb1, RvB2, Swc2, Swc4, and Swc5. AP-MS of Hv1 co-purified Swr1, Swc2, Arp5, and Rvb1 (Ashraf et al., 2019), consistent with Tt SWR-C function in the deposition of Hv1. In yeast, the NuA4 HAT acetylates H4 which recruits SWR via the BRD-containing Bdf1 subunit to deposit Htz1 (Durant and Pugh, 2007; Altaf et al., 2010). It is tempting to speculate that the Tt NuA4-like activity (discussed above) acetylates H4, proving a platform for Ibd1 recruitment and subsequent Hv1 deposition. Support for this model rests in the fact that Ibd1 appears to recognize H4K8ac (Saettone et al., 2018) [analogous to H4K7ac (**Figure 2B**)] that is characteristic of NuA4 activity (Allard et al., 1999).

Group IIB proteins also include BroP6, BroP2, and BroP4, that are similar to BroP5/Ibd1 in that they are small (~400 aa), with



Group III BRD Proteins

Group III (Figure 4) includes p55/Gcn5 and three proteins with a single BRD in combination with an Extra-Terminal (ET) Domain. Tt p55/Gcn5 possesses a C-terminal BRD, as does Gcn5 in other organisms. A two-step model was proposed by Taverna and colleagues for yeast Gcn5 which first acetylates H3K14 to provide a platform for binding by its BRD which stimulates its HAT activity on H3K18 (Cieniewicz et al., 2014). By analogy, the role of Tt Gcn5 BRD could be similar, stimulating Gcn5 activity after initial recruitment. In this case, the role of the Ibd1 BRD within SAGA would be to recruit SAGA to a region of chromatin perhaps acetylated by the NuA4 activity described above where it would acetylate H3, stimulating transcription. Bret-1 Bret-2 and Bret-3 are predicted to have an ET domain. They are similar in domain structure to the BET sub-family of human BRD proteins (BRD2, BRD3, BRD4, and the testis-specific BRDT) that harbor at their amino-termini two BRD followed by an ET domain that mediates protein-protein interactions. BET protein are intense subjects of research in human cells where they are implicated in cancer and are targets for molecules such as JQ1. Differently to human (and yeast) BET that possess two BRDs, Tt BETs only possess one BRD.

HISTONE DEACETYLATION IN *Tetrahymena thermophila*

Histone deacetylases remove acetyl groups from lysine residues. Inhibitors that target HDACs have been used to target human diseases such as cancer (Jain and Zain, 2011; Lee et al., 2015). HDAC inhibitors are also being investigated in the treatment of parasitic diseases (Vanagas et al., 2012; Carrillo et al., 2015; Chua et al., 2017). Wiley and colleagues used a bioinformatic query of the Tt MAC genome (Smith et al., 2008) to predict the existence of 18 HDACs that are named THDs (*Tetrahymena* Histone Deacetylase) and classified according to their similarity to yeast HDACs Rpd3 (class I, 3 members including THD1), Hda1 (class II, 2 members including THD2), and Sir2 (class III, 11 members) with an additional 2 classified as HDAC-like, one of which (Thd5) is predicted to be an ortholog of HDAC11, the smallest HDAC and it is the sole member of HDAC IV family (Gregorette et al., 2004) implicated in mitosis and meiosis (Sui et al., 2020). Detailed molecular analysis has been performed on Tt HDACs, THD1 (Wiley et al., 2000, 2005; Parker et al., 2007), and THD2 (Smith et al., 2008).

The Class I HDAC THD1

Class I THD1 was shown to be recruited to developing new macronuclei (Wiley et al., 2000) and to be important for the integrity of macronuclear chromatin in logarithmically dividing cells (Wiley et al., 2005). Cells knocked down for *THD1* contain higher amounts of MAC DNA, large extrusion bodies, and enlarged nucleoli (Wiley et al., 2005). It was further shown (Parker et al., 2007) that MAC chromatin in *THD1* knockdowns failed to condense during starvation, which was correlated

with aberrant hyper-phosphorylation of histone H1 and the overexpression of CDC2, encoding the major histone H1 kinase. Class I HDACs such as Rpd3 are conserved among eukaryotes and are frequently found in corepressor complexes, where they mediate repression by a variety of transcription factors. In humans, the SIN3/RPD3 complex that also contains RbAp46/48 in addition to several other proteins, is targeted to specific genes through protein-protein interactions between SIN3 and either DNA-binding repressors or corepressors (Lewis et al., 2004; Keogh et al., 2005). The metazoan DREAM complex is responsible for the transcriptional regulation of cell cycle-related genes (Sadasivam and DeCaprio, 2013). Recent findings hint at the existence of a DREAM complex in Tt (Zhang et al., 2018). In human cells a Sin3B/HDAC complex robustly interacts with the DREAM complex in a cell-cycle-dependent manner (Bainor et al., 2018). It remains to be determined whether THD1 exerts its phenotype through the Tt DREAM complex.

The Class II HDAC THD2

Based on work in the 1980s, it was known that although transcription-related acetylation was never observed in the MIC, deposition-related patterns could be observed in the presence of HDAC inhibitors such as sodium butyrate (Allis et al., 1985) indicating that H3 and H4 assembled into MIC chromatin were subject to deposition-related acetylation but that was quickly removed post-assembly. GFP-tagging was used to demonstrate that the class II HDAC named Thd2 (Tt histone deacetylase 2) localized specifically to the MIC. A complete deletion of THD2 showed ectopic H3 and H4 acetylation in the MIC indicating that WT Thd2 function is to remove deposition related acetylation (Smith et al., 2008). Interestingly, the THD2 KO also displayed a defect in MIC morphology as well as the regulated proteolytic processing of its histone H3, specifically deficient in producing the fast form of H3 that in the MIC is phosphorylated on Ser10 (Allis and Gorovsky, 1981; Allis and Wiggins, 1984), a mitotic PTM necessary for chromosome condensation and segregation (Wei et al., 1998, 1999) suggesting that Thd2 functions upstream of the proteolytic cleavage and subsequent phosphorylation of Ser10 on histone H3. It will be interesting to determine if AP-MS of Thd2 can help identify the elusive protease responsible for this enigmatic process.

Class III HDAC

Class III histone deacetylases, known as sirtuins, couple the deacetylation of lysine with the hydrolysis of NAD⁺ by transferring the acetyl group to the ADP-ribose moiety to form O-acetyl-ADP-ribose, releasing free nicotinamide (Dang, 2014). Nicotinamide can thus be used as an inhibitor of sirtuin class HDACs and was used to demonstrate a possible role for the sirtuins in meiotic prophase as well as the degradation of the parental MAC during conjugation (Slade et al., 2011).

More could be learned through the molecular analysis of Tt HDACs. The use of Trichostatin A, a selective inhibitor of class I and II HDACs, resulted in defects in the progression through meiosis and also affected the deletion of IESs (Duharcourt and Yao, 2002). The identity of relevant HDAC(s) that function in these processes remains unknown.

CONCLUSION AND PERSPECTIVES

Tetrahymena thermophila offers a powerful model system with well-developed functional genomics with which to explore and understand the components of the histone acetylation cycle. The unique nuclear biology of Tt has been extremely useful in the past in the development of the histone acetylation field. Despite its efficacy the model has been underexplored. A complete understanding of the function and mechanism of the Tt histone acetylation cycle, in particular the role of histone acetylation in the regulation of H2AZ dynamics, should yield fundamental knowledge on the mechanism of transcription. In addition, the position of Tt on the evolutionary tree will permit insight into Alveolate-specific biology such as the composition of NuA4 and the identity of the protist H3K56-specific HAT.

REFERENCES

- Akematsu, T., Fukuda, Y., Garg, J., Fillingham, J. S., Pearlman, R. E., and Loidl, J. (2017). Post-meiotic DNA double-strand breaks occur in *Tetrahymena*, and require Topoisomerase II and Spo11. *eLife* 6:e26176.
- Allard, S., Utley, R. T., Savard, J., Clarke, A., Grant, P., Brandl, C. J., et al. (1999). NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. *EMBO J.* 18, 5108–5119. doi: 10.1093/emboj/18.18.5108
- Allis, C. D., Allen, R. L., Wiggins, J. C., Chicoine, L. G., and Richman, R. (1984). Proteolytic processing of h1-like histones in chromatin: a physiologically and developmentally regulated event in *Tetrahymena micronuclei*. *J. Cell Biol.* 99, 1669–1677. doi: 10.1083/jcb.99.5.1669
- Allis, C. D., Berger, S. L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., et al. (2007). New nomenclature for chromatin-modifying enzymes. *Cell* 131, 633–636.
- Allis, C. D., Bowen, J. K., Abraham, G. N., Glover, C. V., and Gorovsky, M. A. (1980a). Proteolytic processing of histone H3 in chromatin: a physiologically regulated event in *Tetrahymena micronuclei*. *Cell* 20, 55–64. doi: 10.1016/0092-8674(80)90234-2
- Allis, C. D., Chicoine, L. G., Richman, R., and Schulman, I. G. (1985). Deposition-related histone acetylation in micronuclei of conjugating *Tetrahymena*. *Proc. Natl. Acad. Sci. U.S.A.* 82, 8048–8052. doi: 10.1073/pnas.82.23.8048
- Allis, C. D., Glover, C. V., Bowen, J. K., and Gorovsky, M. A. (1980b). Histone variants specific to the transcriptionally active, amitotically dividing macronucleus of the unicellular eucaryote, *Tetrahymena thermophila*. *Cell* 20, 609–617. doi: 10.1016/0092-8674(80)90307-4
- Allis, C. D., Glover, C. V., and Gorovsky, M. A. (1979). Micronuclei of *Tetrahymena* contain two types of histone H3. *Proc. Natl. Acad. Sci. U.S.A.* 76, 4857–4861. doi: 10.1073/pnas.76.10.4857
- Allis, C. D., and Gorovsky, M. A. (1981). Histone phosphorylation in macro- and micronuclei of *Tetrahymena thermophila*. *Biochemistry* 20, 3828–3833. doi: 10.1021/bi00516a025
- Allis, C. D., Richman, R., Gorovsky, M. A., Ziegler, Y. S., Touchstone, B., Bradley, W. A., et al. (1986). hvl1 is an evolutionarily conserved H2A variant that is preferentially associated with active genes. *J. Biol. Chem.* 261, 1941–1948.
- Allis, C. D., and Wiggins, J. C. (1984). Proteolytic processing of micronuclear H3 and histone phosphorylation during conjugation in *Tetrahymena thermophila*. *Exp. Cell Res.* 153, 287–298. doi: 10.1016/0014-4827(84)90601-3
- Allis, C. D., Ziegler, Y. S., Gorovsky, M. A., and Olmsted, J. B. (1982). A conserved histone variant enriched in nucleoli of mammalian cells. *Cell* 31, 131–136. doi: 10.1016/0092-8674(82)90412-3
- Altav, M., Auger, A., Monnet-Saksouk, J., Brodeur, J., Piquet, S., Cramet, M., et al. (2010). NuA4-dependent acetylation of nucleosomal histones H4 and H2A directly stimulates incorporation of H2A.Z by the SWR1 complex. *J. Biol. Chem.* 285, 15966–15977. doi: 10.1074/jbc.m110.117069
- Andrews, K. T., Tran, T. N., and Fairlie, D. P. (2012). Towards histone deacetylase inhibitors as new antimalarial drugs. *Curr. Pharm. Des.* 18, 3467–3479.

AUTHOR CONTRIBUTIONS

SW wrote the manuscript, edited and prepared Figure 3. AS edited the manuscript and prepared Figures 1, 4, SN-S and ND edited the manuscript. JF conceived, wrote, prepared Figure 2, and edited the manuscript. All authors read and approved the final manuscript.

FUNDING

Research in the Fillingham lab was supported by the Faculty of Science, Ryerson University, in addition to a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

- Arrowsmith, C. H., and Schapira, M. (2019). Targeting non-bromodomain chromatin readers. *Nat. Struct. Mol. Biol.* 26, 863–869. doi: 10.1038/s41594-019-0290-2
- Ashraf, K., Nabeel-Shah, S., Garg, J., Saettone, A., Derynck, J., Gingras, A. C., et al. (2019). Proteomic analysis of histones H2A/H2B and variant Hv1 in *Tetrahymena thermophila* reveals an ancient network of chaperones. *Mol. Biol. Evol.* 36, 1037–1055. doi: 10.1093/molbev/msz039
- Bainor, A. J., Saini, S., Calderon, A., Casado-Polanco, R., Giner-Ramirez, B., Moncada, C., et al. (2018). The HDAC-associated Sin3B protein represses DREAM complex targets and cooperates with APC/C to promote quiescence. *Cell. Rep.* 25, 2797.e8–2807.e8.
- Baker, S. P., and Grant, P. A. (2007). The SAGA continues: expanding the cellular role of a transcriptional co-activator complex. *Oncogene* 26, 5329–5340. doi: 10.1038/sj.onc.1210603
- Bazan, J. F. (2008). An old HAT in human p300/CBP and yeast Rtt109. *Cell Cycle* 7, 1884–1886. doi: 10.4161/cc.7.12.6074
- Berndsen, C. E., and Denu, J. M. (2008). Catalysis and substrate selection by histone/protein lysine acetyltransferases. *Curr. Opin. Struct. Biol.* 18, 682–689. doi: 10.1016/j.sbi.2008.11.004
- Brahma, S., Udugama, M. I., Kim, J., Hada, A., Bhardwaj, S. K., Hailu, S. G., et al. (2017). INO80 exchanges H2A.Z for H2A by translocating on DNA proximal to histone dimers. *Nat. Commun.* 8:15616.
- Brownell, J. E., and Allis, C. D. (1995). An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in *Tetrahymena macronuclei*. *Proc. Natl. Acad. Sci. U.S.A.* 92, 6364–6368. doi: 10.1073/pnas.92.14.6364
- Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y., et al. (1996). *Tetrahymena* histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84, 843–851. doi: 10.1016/s0092-8674(00)81063-6
- Burgess, R. J., Zhou, H., Han, J., and Zhang, Z. (2010). A role for Gcn5 in replication-coupled nucleosome assembly. *Mol. Cell.* 37, 469–480. doi: 10.1016/j.molcel.2010.01.020
- Carrillo, A. K., Guiguemde, W. A., and Guy, R. K. (2015). Evaluation of histone deacetylase inhibitors (HDACi) as therapeutic leads for human African trypanosomiasis (HAT). *Bioorg. Med. Chem.* 23, 5151–5155. doi: 10.1016/j.bmc.2014.12.066
- Cervantes, M. D., Xi, X., Vermaak, D., Yao, M. C., and Malik, H. S. (2006). The CNA1 histone of the ciliate *Tetrahymena thermophila* is essential for chromosome segregation in the germline micronucleus. *Mol. Biol. Cell* 17, 485–497. doi: 10.1091/mbc.e05-07-0698
- Chalker, D. L., and Yao, M. C. (2001). Nongenetic, bidirectional transcription precedes and may promote developmental DNA deletion in *Tetrahymena thermophila*. *Genes Dev.* 15, 1287–1298. doi: 10.1101/gad.884601
- Chicoine, L. G., and Allis, C. D. (1986). Regulation of histone acetylation during macronuclear differentiation in *Tetrahymena*: evidence for control at the level

- of acetylation and deacetylation. *Dev. Biol.* 116, 477–485. doi: 10.1016/0012-1606(86)90148-x
- Chicoine, L. G., Richman, R., Cook, R. G., Gorovsky, M. A., and Allis, C. D. (1987). A single histone acetyltransferase from *Tetrahymena macronuclei* catalyzes deposition-related acetylation of free histones and transcription-related acetylation of nucleosomal histones. *J. Cell Biol.* 105, 127–135. doi: 10.1083/jcb.105.1.127
- Chicoine, L. G., Schulman, I. G., Richman, R., Cook, R. G., and Allis, C. D. (1986). Nonrandom utilization of acetylation sites in histones isolated from *Tetrahymena*. Evidence for functionally distinct H4 acetylation sites. *J. Biol. Chem.* 261, 1071–1076.
- Chua, M. J., Arnold, M. S., Xu, W., Lancelot, J., Lamotte, S., Spath, G. F., et al. (2017). Effect of clinically approved HDAC inhibitors on Plasmodium, Leishmania and Schistosoma parasite growth. *Int. J. Parasitol. Drugs Drug Resist.* 7, 42–50. doi: 10.1016/j.ijpddr.2016.12.005
- Cieniewicz, A. M., Moreland, L., Ringel, A. E., Mackintosh, S. G., Raman, A., Gilbert, T. M., et al. (2014). The bromodomain of Gcn5 regulates site specificity of lysine acetylation on histone H3. *Mol. Cell. Proteomics* 13, 2896–2910. doi: 10.1074/mcp.m114.038174
- Clapier, C. R., and Cairns, B. R. (2009). The biology of chromatin remodeling complexes. *Annu. Rev. Biochem.* 78, 273–304.
- Cochran, A. G., Conery, A. R., and Sims, R. J. III (2019). Bromodomains: a new target class for drug development. *Nat. Rev. Drug Discov.* 18, 609–628. doi: 10.1038/s41573-019-0030-7
- Collins, S. R., Miller, K. M., Maas, N. L., Roguev, A., Fillingham, J., Chu, C. S., et al. (2007). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* 446, 806–810. doi: 10.1038/nature05649
- Cui, B., and Gorovsky, M. A. (2006). Centromeric histone H3 is essential for vegetative cell division and for DNA elimination during conjugation in *Tetrahymena thermophila*. *Mol. Cell. Biol.* 26, 4499–4510. doi: 10.1128/mcb.00079-06
- Cui, B., Liu, Y., and Gorovsky, M. A. (2006). Deposition and function of histone H3 variants in *Tetrahymena thermophila*. *Mol. Cell. Biol.* 26, 7719–7730. doi: 10.1128/mcb.01139-06
- Dang, W. (2014). The controversial world of sirtuins. *Drug Discov. Today Technol.* 12, e9–e17. doi: 10.1016/j.ddtec.2012.08.003
- Daniel, J. A., and Grant, P. A. (2007). Multi-tasking on chromatin with the SAGA coactivator complexes. *Mutat. Res.* 618, 135–148. doi: 10.1016/j.mrfmmm.2006.09.008
- Das, C., Lucia, M. S., Hansen, K. C., and Tyler, J. K. (2009). CBP/p300-mediated acetylation of histone H3 on lysine 56. *Nature* 459, 113–117. doi: 10.1038/nature07861
- de Koning, H. (2017). Drug resistance in protozoan parasites. *Emerg. Top. Life Sci.* 1, 627–632. doi: 10.1042/etls20170113
- Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., and Zhou, M. M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. *Nature* 399, 491–496. doi: 10.1038/20974
- Doyon, Y., and Cote, J. (2004). The highly conserved and multifunctional NuA4 HAT complex. *Curr. Opin. Genet. Dev.* 14, 147–154. doi: 10.1016/j.gde.2004.02.009
- Draker, R., Ng, M. K., Sarcinella, E., Ignatchenko, V., Kislinger, T., and Cheung, P. (2012). A combination of H2A.Z and H4 acetylation recruits Brd2 to chromatin during transcriptional activation. *PLoS Genet.* 8:e1003047. doi: 10.1371/journal.pgen.1003047
- Driscoll, R., Hudson, A., and Jackson, S. P. (2007). Yeast Rtt109 promotes genome stability by acetylating histone H3 on lysine 56. *Science* 315, 649–652. doi: 10.1126/science.1135862
- Duharcourt, S., and Yao, M. C. (2002). Role of histone deacetylation in developmentally programmed DNA rearrangements in *Tetrahymena thermophila*. *Eukaryot. Cell* 1, 293–303. doi: 10.1128/ec.1.2.293-303.2002
- Durant, M., and Pugh, B. F. (2007). NuA4-directed chromatin transactions throughout the *Saccharomyces cerevisiae* genome. *Mol. Cell. Biol.* 27, 5327–5335. doi: 10.1128/mcb.00468-07
- Eisen, J. A., Coyne, R. S., Wu, M., Wu, D., Thiagarajan, M., Wortman, J. R., et al. (2006). Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. *PLoS Biol.* 4:e286. doi: 10.1371/journal.pbio.0040286
- Fan, Q., An, L., and Cui, L. (2004). Plasmodium falciparum histone acetyltransferase, a yeast GCN5 homologue involved in chromatin remodeling. *Eukaryot. Cell* 3, 264–276. doi: 10.1128/ec.3.2.264-276.2004
- Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J. P., Barsyte-Lovejoy, D., et al. (2012). Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* 149, 214–231. doi: 10.1016/j.cell.2012.02.013
- Fillingham, J., Recht, J., Silva, A. C., Suter, B., Emili, A., Stagljar, I., et al. (2008). Chaperone control of the activity and specificity of the histone H3 acetyltransferase Rtt109. *Mol. Cell. Biol.* 28, 4342–4353. doi: 10.1128/mcb.00182-08
- Fillingham, J. S., Garg, J., Tsao, N., Vythilingum, N., Nishikawa, T., and Pearlman, R. E. (2006). Molecular genetic analysis of an SNF2/brahma-related gene in *Tetrahymena thermophila* suggests roles in growth and nuclear development. *Eukaryot. Cell* 5, 1347–1359. doi: 10.1128/ec.00149-06
- French, C. A., Kutok, J. L., Faquin, W. C., Toretsky, J. A., Antonescu, C. R., Griffin, C. A., et al. (2004). Midline carcinoma of children and young adults with NUT rearrangement. *J. Clin. Oncol.* 22, 4135–4139. doi: 10.1200/jco.2004.02.107
- Fyodorov, D. V., Zhou, B. R., Skoultschi, A. I., and Bai, Y. (2018). Emerging roles of linker histones in regulating chromatin structure and function. *Nat. Rev. Mol. Cell Biol.* 19, 192–206. doi: 10.1038/nrm.2017.94
- Garcia, B. A., Hake, S. B., Diaz, R. L., Kauer, M., Morris, S. A., Recht, J., et al. (2007). Organismal differences in post-translational modifications in histones H3 and H4. *J. Biol. Chem.* 282, 7641–7655. doi: 10.1074/jbc.m607900200
- Gardner, K. E., Allis, C. D., and Strahl, B. D. (2011). Operating on chromatin, a colorful language where context matters. *J. Mol. Biol.* 409, 36–46. doi: 10.1016/j.jmb.2011.01.040
- Garg, J., Lambert, J. P., Karsou, A., Marquez, S., Nabeel-Shah, S., Bertucci, V., et al. (2013). Conserved Asf1-importin beta physical interaction in growth and sexual development in the ciliate *Tetrahymena thermophila*. *J. Proteomics* 94, 311–326. doi: 10.1016/j.jpropt.2013.09.018
- Garg, J., Saettone, A., Nabeel-Shah, S., Cadorin, M., Ponce, M., Marquez, S., et al. (2019). The Med31 conserved component of the divergent mediator complex in *Tetrahymena thermophila* participates in developmental regulation. *Curr. Biol.* 29, 2371.e–2379.e.
- Georgakopoulos, T., and Thireos, G. (1992). Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. *EMBO J.* 11, 4145–4152. doi: 10.1002/j.1460-2075.1992.tb05507.x
- Grant, P. A., Duggan, L., Cote, J., Roberts, S. M., Brownell, J. E., Candau, R., et al. (1997). Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev.* 11, 1640–1650. doi: 10.1101/gad.11.13.1640
- Grant, P. A., Eberharther, A., John, S., Cook, R. G., Turner, B. M., and Workman, J. L. (1999). Expanded lysine acetylation specificity of Gcn5 in native complexes. *J. Biol. Chem.* 274, 5895–5900. doi: 10.1074/jbc.274.9.5895
- Grant, P. A., Schieltz, D., Pray-Grant, M. G., Steger, D. J., Reese, J. C., Yates, J. R. III, et al. (1998). A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation. *Cell* 94, 45–53. doi: 10.1016/s0092-8674(00)81220-9
- Gregoret, I. V., Lee, Y. M., and Goodson, H. V. (2004). Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis. *J. Mol. Biol.* 338, 17–31. doi: 10.1016/j.jmb.2004.02.006
- Grunstein, M., and Allis, C. D. (2018). Genetics, biochemistry, and “Simple” organisms converge to unlock secrets in histone biology: the 2018 albert lasker basic medical research award. *JAMA* 320, 1233–1234.
- Gupta, D. K., Patra, A. T., Zhu, L., Gupta, A. P., and Bozdech, Z. (2016). DNA damage regulation and its role in drug-related phenotypes in the malaria parasites. *Sci. Rep.* 6:23603.
- Hamilton, E. P., Kapusta, A., Huvos, P. E., Bidwell, S. L., Zafar, N., Tang, H., et al. (2016). Structure of the germline genome of *Tetrahymena thermophila* and relationship to the massively rearranged somatic genome. *eLife* 5:e19090.
- Hanquier, J., Gimeno, T., Jeffers, V., and Sullivan, W. J. Jr. (2020). Evaluating the GCN5b bromodomain as a novel therapeutic target against the parasite *Toxoplasma gondii*. *Exp. Parasitol.* 211:107868. doi: 10.1016/j.exppara.2020.107868
- Henikoff, S., and Smith, M. M. (2015). Histone variants and epigenetics. *Cold Spring Harb. Perspect. Biol.* 7:a019364. doi: 10.1101/cshperspect.a019364

- Hoeijmakers, W. A. M., Miao, J., Schmidt, S., Toenhake, C. G., Shrestha, S., Venhuizen, J., et al. (2019). Epigenetic reader complexes of the human malaria parasite, *Plasmodium falciparum*. *Nucleic Acids Res.* 47, 11574–11588. doi: 10.1093/nar/gkz1044
- Jacquet, K., Fradet-Turcotte, A., Avvakumov, N., Lambert, J. P., Roques, C., Pandita, R. K., et al. (2016). The TIP60 complex regulates bivalent chromatin recognition by 53BP1 through Direct H4K20me binding and H2AK15 acetylation. *Mol. Cell.* 62, 409–421. doi: 10.1016/j.molcel.2016.03.031
- Jain, A. K., and Barton, M. C. (2017). Bromodomain histone readers and cancer. *J. Mol. Biol.* 429, 2003–2010. doi: 10.1016/j.jmb.2016.11.020
- Jain, S., and Zain, J. (2011). Romidepsin in the treatment of cutaneous T-cell lymphoma. *J. Blood Med.* 2, 37–47.
- Jeffers, V., Gao, H., Checkley, L. A., Liu, Y., Ferdig, M. T., and Sullivan, W. J. Jr. (2016). Garcinol Inhibits GCN5-mediated lysine acetyltransferase activity and prevents replication of the parasite *Toxoplasma gondii*. *Antimicrob. Agents Chemother.* 60, 2164–2170. doi: 10.1128/aac.03059-15
- Jeffers, V., and Sullivan, W. J. Jr. (2012). Lysine acetylation is widespread on proteins of diverse function and localization in the protozoan parasite *Toxoplasma gondii*. *Eukaryot. Cell* 11, 735–742. doi: 10.1128/ec.00088-12
- Jeffers, V., Yang, C., Huang, S., and Sullivan, W. J. Jr. (2017). Bromodomains in protozoan parasites: evolution, function, and opportunities for drug development. *Microbiol. Mol. Biol. Rev.* 81:e00047-16.
- Johnson, C. N., Adkins, N. L., and Georgel, P. (2005). Chromatin remodeling complexes: ATP-dependent machines in action. *Biochem. Cell Biol.* 83, 405–417. doi: 10.1139/o05-115
- Josling, G. A., Petter, M., Oehring, S. C., Gupta, A. P., Dietz, O., Wilson, D. W., et al. (2015). A plasmodium falciparum bromodomain protein regulates invasion gene expression. *Cell Host Microbe* 17, 741–751. doi: 10.1016/j.chom.2015.05.009
- Kawahara, T., Siegel, T. N., Ingram, A. K., Alsford, S., Cross, G. A., and Horn, D. (2008). Two essential MYST-family proteins display distinct roles in histone H4K10 acetylation and telomeric silencing in trypanosomes. *Mol. Microbiol.* 69, 1054–1068. doi: 10.1111/j.1365-2958.2008.06346.x
- Keogh, M. C., Kurdistani, S. K., Morris, S. A., Ahn, S. H., Podolny, V., Collins, S. R., et al. (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* 123, 593–605. doi: 10.1016/j.cell.2005.10.025
- Kobor, M. S., Venkatasubrahmanyam, S., Meneghini, M. D., Gin, J. W., Jennings, J. L., Link, A. J., et al. (2004). A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS Biol.* 2:E131. doi: 10.1371/journal.pbio.0020131
- Kougnassoukou Tchira, P. E., Filipakopoulos, P., and Lambert, J. P. (2019). Emerging tools to investigate bromodomain functions. *Methods* doi: 10.1016/j.jmeth.2019.11.003 [Epub ahead of print].
- Kuo, M. H., and Allis, C. D. (1998). Roles of histone acetyltransferases and deacetylases in gene regulation. *Bioessays* 20, 615–626. doi: 10.1002/(sici)1521-1878(199808)20:8<615::aid-bies4>3.0.co;2-h
- Lalonde, M. E., Cheng, X., and Cote, J. (2014). Histone target selection within chromatin: an exemplary case of teamwork. *Genes Dev.* 28, 1029–1041. doi: 10.1101/gad.236331.113
- Lee, H. Z., Kwitkowski, V. E., Del Valle, P. L., Ricci, M. S., Saber, H., Habtemariam, B. A., et al. (2015). FDA approval: belinostat for the treatment of patients with relapsed or refractory peripheral T-cell lymphoma. *Clin. Cancer Res.* 21, 2666–2670. doi: 10.1158/1078-0432.ccr-14-3119
- Lewis, P. W., Beall, E. L., Fleischer, T. C., Georlette, D., Link, A. J., and Botchan, M. R. (2004). Identification of a *Drosophila* Myb-E2F2/RBF transcriptional repressor complex. *Genes Dev.* 18, 2929–2940. doi: 10.1101/gad.1255204
- Li, Y., Li, H., Sui, M., Li, M., Wang, J., Meng, Y., et al. (2019). Fungal acetylome comparative analysis identifies an essential role of acetylation in human fungal pathogen virulence. *Commun. Biol.* 2:154.
- Li, Y., Wen, H., Xi, Y., Tanaka, K., Wang, H., Peng, D., et al. (2014). AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation. *Cell* 159, 558–571. doi: 10.1016/j.cell.2014.09.049
- Lin, R., Leone, J. W., Cook, R. G., and Allis, C. D. (1989). Antibodies specific to acetylated histones document the existence of deposition- and transcription-related histone acetylation in *Tetrahymena*. *J. Cell Biol.* 108, 1577–1588. doi: 10.1083/jcb.108.5.1577
- Liu, Y., Taverna, S. D., Muratore, T. L., Shabanowitz, J., Hunt, D. F., and Allis, C. D. (2007). RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in *Tetrahymena*. *Genes Dev.* 21, 1530–1545. doi: 10.1101/gad.1544207
- Luger, K., Mader, 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.
- Martindale, D. W., Allis, C. D., and Bruns, P. J. (1982). Conjugation in *Tetrahymena thermophila*. A temporal analysis of cytological stages. *Exp. Cell Res.* 140, 227–236. doi: 10.1016/0014-4827(82)90172-0
- Martindale, D. W., Allis, C. D., and Bruns, P. J. (1985). RNA and protein synthesis during meiotic prophase in *Tetrahymena thermophila*. *J. Protozool.* 32, 644–649. doi: 10.1111/j.1550-7408.1985.tb03094.x
- Masumoto, H., Hawke, D., Kobayashi, R., and Verreault, A. (2005). A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature* 436, 294–298. doi: 10.1038/nature03714
- Miao, J., Fan, Q., Cui, L., Li, X., Wang, H., Ning, G., et al. (2010). The MYST family histone acetyltransferase regulates gene expression and cell cycle in malaria parasite *Plasmodium falciparum*. *Mol. Microbiol.* 78, 883–902. doi: 10.1111/j.1365-2958.2010.07371.x
- Miao, J., Lawrence, M., Jeffers, V., Zhao, F., Parker, D., Ge, Y., et al. (2013). Extensive lysine acetylation occurs in evolutionarily conserved metabolic pathways and parasite-specific functions during *Plasmodium falciparum* intraerythrocytic development. *Mol. Microbiol.* 89, 660–675. doi: 10.1111/mmi.12303
- Miao, W., Xiong, J., Bowen, J., Wang, W., Liu, Y., Braguinets, O., et al. (2009). Microarray analyses of gene expression during the *Tetrahymena thermophila* life cycle. *PLoS One* 4:e4429. doi: 10.1371/journal.pone.0004429
- Morrison, A. J., and Shen, X. (2009). Chromatin remodelling beyond transcription: the INO80 and SWR1 complexes. *Nat. Rev. Mol. Cell Biol.* 10, 373–384. doi: 10.1038/nrm2693
- Nabeel-Shah, S., Ashraf, K., Saettone, A., Garg, J., Derynck, J., Lambert, J. P., et al. (2020). Nucleus-specific linker histones Hho1 and Mlh1 form distinct protein interactions during growth, starvation and development in *Tetrahymena thermophila*. *Sci. Rep.* 10:168.
- Nagarajan, P., Ge, Z., Sirbu, B., Doughty, C., Agudelo Garcia, P. A., Schleiderer, M., et al. (2013). Histone acetyl transferase 1 is essential for mammalian development, genome stability, and the processing of newly synthesized histones H3 and H4. *PLoS Genet.* 9:e1003518. doi: 10.1371/journal.pgen.1003518
- Nguyen, H. H. T., Yeoh, L. M., Chisholm, S. A., and Duffy, M. F. (2020). Developments in drug design strategies for bromodomain protein inhibitors to target *Plasmodium falciparum* parasites. *Expert Opin. Drug Discov.* 15, 415–425. doi: 10.1080/17460441.2020.1704251
- No Author (2018). Chasing histone biology from sea urchins to yeast. *Cell* 175, 27–29. doi: 10.1016/j.cell.2018.08.008
- Ohba, R., Steger, D. J., Brownell, J. E., Mizzen, C. A., Cook, R. G., Cote, J., et al. (1999). A novel H2A/H4 nucleosomal histone acetyltransferase in *Tetrahymena thermophila*. *Mol. Cell. Biol.* 19, 2061–2068. doi: 10.1128/mcb.19.3.2061
- Orias, E., Cervantes, M. D., and Hamilton, E. P. (2011). *Tetrahymena thermophila*, a unicellular eukaryote with separate germline and somatic genomes. *Res. Microbiol.* 162, 578–586. doi: 10.1016/j.resmic.2011.05.001
- Orias, E., and Flacks, M. (1975). Macronuclear genetics of *Tetrahymena*. I. Random distribution of macronuclear genecopies in pyriformis, T., synngen 1. *Genetics* 79, 187–206.
- Papamichos-Chronakis, M., Watanabe, S., Rando, O. J., and Peterson, C. L. (2011). Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell* 144, 200–213. doi: 10.1016/j.cell.2010.12.021
- Parker, K., Maxson, J., Mooney, A., and Wiley, E. A. (2007). Class I histone deacetylase Thd1p promotes global chromatin condensation in *Tetrahymena thermophila*. *Eukaryot. Cell* 6, 1913–1924. doi: 10.1128/ec.00217-07
- Parthun, M. R. (2007). Hat1: the emerging cellular roles of a type B histone acetyltransferase. *Oncogene* 26, 5319–5328. doi: 10.1038/sj.onc.1210602
- Parthun, M. R. (2013). Histone acetyltransferase 1: more than just an enzyme? *Biochim. Biophys. Acta* 1819, 256–263. doi: 10.1016/j.bbaggm.2011.07.006
- Parthun, M. R., Widom, J., and Gottschling, D. E. (1996). The major cytoplasmic histone acetyltransferase in yeast: links to chromatin replication and histone metabolism. *Cell* 87, 85–94. doi: 10.1016/s0092-8674(00)81325-2

- Poli, J., Gasser, S. M., and Papamichos-Chronakis, M. (2017). The INO80 remodeller in transcription, replication and repair. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 372:20160290. doi: 10.1098/rstb.2016.0290
- Recht, J., Tsubota, T., Tanny, J. C., Diaz, R. L., Berger, J. M., Zhang, X., et al. (2006). Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6988–6993. doi: 10.1073/pnas.0601676103
- Ren, Q., and Gorovsky, M. A. (2001). Histone H2A.Z acetylation modulates an essential charge patch. *Mol. Cell.* 7, 1329–1335. doi: 10.1016/s1097-2765(01)00269-6
- Richman, R., Chicoine, L. G., Collini, M. P., Cook, R. G., and Allis, C. D. (1988). Micronuclei and the cytoplasm of growing *Tetrahymena* contain a histone acetylase activity which is highly specific for free histone H4. *J. Cell Biol.* 106, 1017–1026. doi: 10.1083/jcb.106.4.1017
- Richters, A., and Koehler, A. N. (2017). Epigenetic modulation using small molecules - targeting histone acetyltransferases in disease. *Curr. Med. Chem.* 24, 4121–4150.
- Ruhl, D. D., Jin, J., Cai, Y., Swanson, S., Florens, L., Washburn, M. P., et al. (2006). Purification of a human SRCAP complex that remodels chromatin by incorporating the histone variant H2A.Z into nucleosomes. *Biochemistry* 45, 5671–5677. doi: 10.1021/bi060043d
- Sadasivam, S., and DeCaprio, J. A. (2013). The DREAM complex: master coordinator of cell cycle-dependent gene expression. *Nat. Rev. Cancer* 13, 585–595. doi: 10.1038/nrc3556
- Saettone, A., Garg, J., Lambert, J. P., Nabeel-Shah, S., Ponce, M., Burtch, A., et al. (2018). The bromodomain-containing protein Ibd1 links multiple chromatin-related protein complexes to highly expressed genes in *Tetrahymena thermophila*. *Epigenet. Chromatin* 11:10.
- Saettone, A., Nabeel-Shah, S., Garg, J., Lambert, J. P., Pearlman, R. E., and Fillingham, J. (2019a). Functional proteomics of nuclear proteins in *Tetrahymena thermophila*: a review. *Genes* 10:333. doi: 10.3390/genes10050333
- Saettone, A., Ponce, M., Nabeel-Shah, S., and Fillingham, J. (2019b). RACS: rapid analysis of ChIP-Seq data for contig based genomes. *BMC Bioinformatics* 20:533. doi: 10.1186/s12859-019-3100-2
- Saha, A., Wittmeyer, J., and Cairns, B. R. (2006). Chromatin remodelling: the industrial revolution of DNA around histones. *Nat. Rev. Mol. Cell Biol.* 7, 437–447. doi: 10.1038/nrm1945
- Santos, J. M., Josling, G., Ross, P., Joshi, P., Orchard, L., Campbell, T., et al. (2017). Red blood cell invasion by the malaria parasite is coordinated by the PfAP2-I transcription factor. *Cell Host Microbe* 21:e10.
- Schulz, D., Mugnier, M. R., Paulsen, E. M., Kim, H. S., Chung, C. W., Tough, D. F., et al. (2015). Bromodomain proteins contribute to maintenance of bloodstream form stage identity in the African Trypanosome. *PLoS Biol.* 13:e1002316. doi: 10.1371/journal.pbio.1002316
- Shen, X., Yu, L., Weir, J. W., and Gorovsky, M. A. (1995). Linker histones are not essential and affect chromatin condensation *in vivo*. *Cell* 82, 47–56. doi: 10.1016/0092-8674(95)90051-9
- Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu. Rev. Biochem.* 81, 65–95. doi: 10.1146/annurev-biochem-051710-134100
- Slade, K. M., Freggiaro, S., Cottrell, K. A., Smith, J. J., and Wiley, E. A. (2011). Sirtuin-mediated nuclear differentiation and programmed degradation in *Tetrahymena*. *BMC Cell Biol.* 12:40. doi: 10.1186/1471-2121-12-40
- Smith, J. J., Torigoe, S. E., Maxson, J., Fish, L. C., and Wiley, E. A. (2008). A class II histone deacetylase acts on newly synthesized histones in *Tetrahymena*. *Eukaryot. Cell* 7, 471–482. doi: 10.1128/ec.00409-07
- Sobel, R. E., Cook, R. G., Perry, C. A., Annunziato, A. T., and Allis, C. D. (1995). Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proc. Natl. Acad. Sci. U.S.A.* 92, 1237–1241. doi: 10.1073/pnas.92.4.1237
- Somech, R., Izraeli, S., and Simson, A. J. (2004). Histone deacetylase inhibitors—a new tool to treat cancer. *Cancer Treat. Rev.* 30, 461–472. doi: 10.1016/j.ctrv.2004.04.006
- Stargell, L. A., Bowen, J., Dadd, C. A., Dedon, P. C., Davis, M., Cook, R. G., et al. (1993). Temporal and spatial association of histone H2A variant hv1 with transcriptionally competent chromatin during nuclear development in *Tetrahymena thermophila*. *Genes Dev.* 7, 2641–2651. doi: 10.1101/gad.7.12b.2641
- Stover, N. A., Krieger, C. J., Binkley, G., Dong, Q., Fisk, D. G., Nash, R., et al. (2006). *Tetrahymena* genome database (TGD): a new genomic resource for *Tetrahymena thermophila* research. *Nucleic Acids Res.* 34, D500–D503.
- Sui, L., Zhang, S., Huang, R., and Li, Z. (2020). HDAC11 promotes meiotic apparatus assembly during mouse oocyte maturation via decreasing H4K16 and alpha-tubulin acetylation. *Cell Cycle* 19, 354–362. doi: 10.1080/15384101.2019.1711315
- Taverna, S. D., Coyne, R. S., and Allis, C. D. (2002). Methylation of histone h3 at lysine 9 targets programmed DNA elimination in *Tetrahymena*. *Cell* 110, 701–711. doi: 10.1016/s0092-8674(02)00941-8
- Taverna, S. D., Ueberheide, B. M., Liu, Y., Tackett, A. J., Diaz, R. L., Shabanowitz, J., et al. (2007). Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. *Proc. Natl. Acad. Sci. U.S.A.* 104, 2086–2091. doi: 10.1073/pnas.0610993104
- Teo, G., Liu, G., Zhang, J., Nesvizhskii, A. I., Gingras, A. C., and Choi, H. (2014). SAINTexpress: improvements and additional features in Significance Analysis of INteractome software. *J. Proteomics* 100, 37–43. doi: 10.1016/j.jprot.2013.10.023
- Thatcher, T. H., MacGaffey, J., Bowen, J., Horowitz, S., Shapiro, D. L., and Gorovsky, M. A. (1994). Independent evolutionary origin of histone H3.3-like variants of animals and *Tetrahymena*. *Nucleic Acids Res.* 22, 180–186. doi: 10.1093/nar/22.2.180
- Tjeertes, J. V., Miller, K. M., and Jackson, S. P. (2009). Screen for DNA-damage-responsive histone modifications identifies H3K9Ac and H3K56Ac in human cells. *EMBO J.* 28, 1878–1889. doi: 10.1038/emboj.2009.119
- Travis, G. H., Colavito-Shepanski, M., and Grunstein, M. (1984). Extensive purification and characterization of chromatin-bound histone acetyltransferase from *Saccharomyces cerevisiae*. *J. Biol. Chem.* 259, 14406–14412.
- Vanagas, L., Jeffers, V., Bogado, S. S., Dalmasso, M. C., Sullivan, W. J. Jr., and Angel, S. O. (2012). Toxoplasma histone acetylation remodelers as novel drug targets. *Expert Rev. Anti Infect. Ther.* 10, 1189–1201. doi: 10.1586/eri.12.100
- Vavra, K. J., Allis, C. D., and Gorovsky, M. A. (1982a). Regulation of histone acetylation in *Tetrahymena* macro- and micronuclei. *J. Biol. Chem.* 257, 2591–2598.
- Vavra, K. J., Colavito-Shepanski, M., and Gorovsky, M. A. (1982b). Histone acetylation and the deoxyribonuclease I sensitivity of the *Tetrahymena* ribosomal gene. *Biochemistry* 21, 1772–1781. doi: 10.1021/bi00537a012
- Verreault, A., Kaufman, P. D., Kobayashi, R., and Stillman, B. (1998). Nucleosomal DNA regulates the core-histone-binding subunit of the human Hat1 acetyltransferase. *Curr. Biol.* 8, 96–108. doi: 10.1016/s0960-9822(98)70040-5
- Wai, D. C. C., Szyska, T. N., Campbell, A. E., Kwong, C., Wilkinson-White, L. E., Silva, A. P. G., et al. (2018). The BRD3 ET domain recognizes a short peptide motif through a mechanism that is conserved across chromatin remodelers and transcriptional regulators. *J. Biol. Chem.* 293, 7160–7175. doi: 10.1074/jbc.ra117.000678
- Wang, J., Dixon, S. E., Ting, L. M., Liu, T. K., Jeffers, V., Croken, M. M., et al. (2014). Lysine acetyltransferase GCN5b interacts with AP2 factors and is required for *Toxoplasma gondii* proliferation. *PLoS Pathog.* 10:e1003830. doi: 10.1371/journal.ppat.1003830
- Wang, J., Gao, S., Mostovoy, Y., Kang, Y., Zagorskin, M., Sun, Y., et al. (2017). Comparative genome analysis of programmed DNA elimination in nematodes. *Genome Res.* 27, 2001–2014. doi: 10.1101/gr.225730.117
- Wang, X., Wang, S., Troisi, E. C., Howard, T. P., Haswell, J. R., Wolf, B. K., et al. (2019). BRD9 defines a SWI/SNF sub-complex and constitutes a specific vulnerability in malignant rhabdoid tumors. *Nat. Commun.* 10:1881.
- Wei, Y., Mizzen, C. A., Cook, R. G., Gorovsky, M. A., and Allis, C. D. (1998). Phosphorylation of histone H3 at serine 10 is correlated with chromosome condensation during mitosis and meiosis in *Tetrahymena*. *Proc. Natl. Acad. Sci. U.S.A.* 95, 7480–7484. doi: 10.1073/pnas.95.13.7480
- Wei, Y., Yu, L., Bowen, J., Gorovsky, M. A., and Allis, C. D. (1999). Phosphorylation of histone H3 is required for proper chromosome condensation and segregation. *Cell* 97, 99–109. doi: 10.1016/s0092-8674(00)80718-7
- Wenkert, D., and Allis, C. D. (1984). Timing of the appearance of macronuclear-specific histone variant hv1 and gene expression in developing new macronuclei of *Tetrahymena thermophila*. *J. Cell Biol.* 98, 2107–2117. doi: 10.1083/jcb.98.6.2107

- Wiley, E. A., Myers, T., Parker, K., Braun, T., and Yao, M. C. (2005). Class I histone deacetylase Thd1p affects nuclear integrity in *Tetrahymena thermophila*. *Eukaryot. Cell* 4, 981–990. doi: 10.1128/ec.4.5.981-990.2005
- Wiley, E. A., Ohba, R., Yao, M. C., and Allis, C. D. (2000). Developmentally regulated rpd3p homolog specific to the transcriptionally active macronucleus of vegetative *Tetrahymena thermophila*. *Mol. Cell. Biol.* 20, 8319–8328. doi: 10.1128/mcb.20.22.8319-8328.2000
- World Health Organization [WHO] (2010). *Working to Overcome the Global Impact of Neglected Tropical Diseases: First WHO Report on Neglected Tropical Diseases*. Geneva: World Health Organization.
- Wurtele, H., Tsao, S., Lepine, G., Mullick, A., Tremblay, J., Drogaris, P., et al. (2010). Modulation of histone H3 lysine 56 acetylation as an antifungal therapeutic strategy. *Nat. Med.* 16, 774–780. doi: 10.1038/nm.2175
- Xiong, J., Yuan, D., Fillingham, J. S., Garg, J., Lu, X., Chang, Y., et al. (2011). Gene network landscape of the ciliate *Tetrahymena thermophila*. *PLoS One* 6:e20124. doi: 10.1371/journal.pone.0020124
- Yan, L., Wang, L., Tian, Y., Xia, X., and Chen, Z. (2016). Structure and regulation of the chromatin remodeller ISWI. *Nature* 540, 466–469. doi: 10.1038/nature20590
- Yang, X., Wu, X., Zhang, J., Zhang, X., Xu, C., Liao, S., et al. (2017). Recognition of hyperacetylated N-terminus of H2AZ by TbBDF2 from *Trypanosoma brucei*. *Biochem. J.* 474, 3817–3830. doi: 10.1042/bcj20170619
- Yao, M., Chao, J., and Cheng, C. (2015). “Programmed genome rearrangements in *Tetrahymena*,” in *Mobile DNA III*, eds N. Craig, M. Chandler, M. Gellert, A. Lambowitz, P. Rice, and S. Sandmeyer (Washington, DC: ASM Press), 349–367. doi: 10.1128/microbiolspec.MDNA3-0012-2014
- Yu, L., and Gorovsky, M. A. (1997). Constitutive expression, not a particular primary sequence, is the important feature of the H3 replacement variant hv2 in *Tetrahymena thermophila*. *Mol. Cell. Biol.* 17, 6303–6310. doi: 10.1128/mcb.17.11.6303
- Zaware, N., and Zhou, M. M. (2019). Bromodomain biology and drug discovery. *Nat. Struct. Mol. Biol.* 26, 870–879. doi: 10.1038/s41594-019-0309-8
- Zeng, L., Zhang, Q., Li, S., Plotnikov, A. N., Walsh, M. J., and Zhou, M. M. (2010). Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. *Nature* 466, 258–262. doi: 10.1038/nature09139
- Zhang, J., Yan, G., Tian, M., Ma, Y., Xiong, J., and Miao, W. (2018). A DP-like transcription factor protein interacts with E2f1 to regulate meiosis in *Tetrahymena thermophila*. *Cell Cycle* 17, 634–642. doi: 10.1080/15384101.2018.1431595
- Zhao, S., Xu, W., Jiang, W., Yu, W., Lin, Y., Zhang, T., et al. (2010). Regulation of cellular metabolism by protein lysine acetylation. *Science* 327, 1000–1004.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Wahab, Saettone, Nabeel-Shah, Dannah and Fillingham. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



A New View of Genome Organization Through RNA Directed Interactions

Gabriel Khelifi^{1,2,3} and Samer M. I. Hussein^{1,2,3*}

¹ Department of Molecular Biology, Medical Biochemistry and Pathology, Université Laval, Québec, QC, Canada, ² Université Laval Cancer Research Center, Université Laval, Québec, QC, Canada, ³ Oncology Division, Centre Hospitalier Universitaire (CHU) de Québec-Université Laval Research Center, Québec, QC, Canada

Keywords: RNA-DNA interactions, long non-coding RNAs, chromatin modifying complex, chromatin structure, genome organization, RNA-RNA interactions

INTRODUCTION

Over the past few decades, we have come to appreciate the complexity of processes regulating chromatin architecture. Ranging from chromatin accessibility (Klemm et al., 2019) to long-range genome organization (Cremer and Cremer, 2001), a wide array of mechanisms is used by the cell to control gene expression. We have recently begun to understand the primordial role of non-coding RNAs (ncRNAs) in such processes and can now declare that several types of RNAs are essential to the regulation of gene expression (Cech and Steitz, 2014). In the nucleus, long non-coding RNAs (lncRNAs) (Bonasio and Shiekhattar, 2014), together with enhancer RNAs (eRNAs) (Li et al., 2016a), stable intronic sequence RNAs (sisRNAs) (Chan and Pek, 2019) and various other classes of transcripts (reviewed in (Li and Fu, 2019)) come together to ensure tight regulation of the chromatin. lncRNAs represent transcripts of more than 200 nucleotides that do not contain any apparent open reading frame (Marchese et al., 2017). While some lncRNAs are localized and active in the cytosol (Noh et al., 2018), many are nuclear and implicated in transcriptional regulation (Vance and Ponting, 2014). These nuclear lncRNAs can modulate the expression of genes through interactions with DNA or chromatin-associated proteins (Bonasio and Shiekhattar, 2014; **Figure 1A**). eRNAs are abundantly transcribed RNAs generated from enhancer regions (Li et al., 2016a). They modulate enhancer activity potentially through interactions with the mediator complex, transcription factors or chromosomal looping factors. Lastly, sisRNAs represent RNAs containing intronic sequences, and increasing evidence shows that several of them act on chromatin regulation (Chan and Pek, 2019). All of these various classes of chromatin-associated RNAs are essential to the regulation of gene expression (Li and Fu, 2019). Some transcripts appear to be “cis-acting,” influencing the expression of genes within their own chromosome, while others control transcriptional processes on other chromosomes in *trans*. Some *cis*-acting RNAs have been shown to function through the formation of R-loops with the complementary sequence from their transcribed loci and affect local gene expression, as is the case with *GATA3-AS1* and *VIM-AS1* (Boque-Sastre et al., 2015; Gibbons et al., 2018). Both *cis*- and *trans*-acting RNAs can affect gene expression through direct RNA-DNA contacts (i.e., RNA-DNA triplexes), as employed by *MEG3* and *KHPS1* (Mondal et al., 2015; Blank-Giwojna et al., 2019), or indirectly through protein intermediates such as the interaction of *FIRRE* lncRNA with the SAF-A protein (Hacisuleyman et al., 2014; **Figure 1A**). These interactions often require specific RNA “domains,” such as motifs recognized by chromatin bound proteins, or RNA-DNA triplex-forming motifs such as polypurine tracts (Li et al., 2016b; Li and Fu, 2019). RNAs exhibiting interactions in both *cis* and *trans* have also

OPEN ACCESS

Edited by:

Jean-Philippe Lambert,
Laval University, Canada

Reviewed by:

Alessandro Bonetti,
Karolinska Institutet (KI), Sweden
Ahmad M. Khalil,
Case Western Reserve University,
United States

*Correspondence:

Samer M. I. Hussein
samer.hussein@fmed.ulaval.ca

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 11 April 2020

Accepted: 02 June 2020

Published: 14 July 2020

Citation:

Khelifi G and Hussein SMI (2020) A
New View of Genome Organization
Through RNA Directed Interactions.
Front. Cell Dev. Biol. 8:517.
doi: 10.3389/fcell.2020.00517

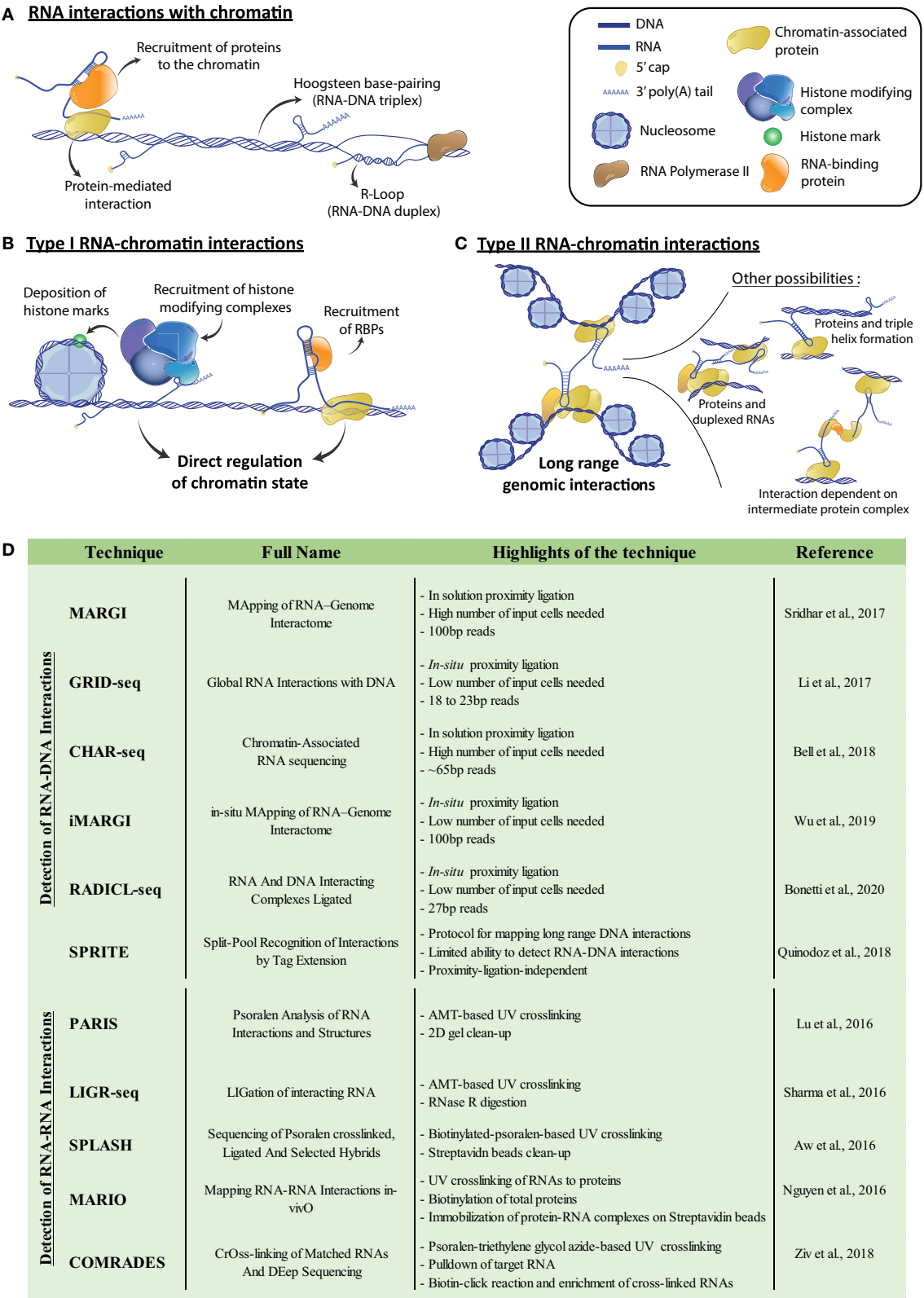


FIGURE 1 | (A) Potential means of chromatin-RNA association. Interactions can be mediated by protein intermediates or rely on either Hoogsteen base-pairing to form RNA-DNA triplexes, or Watson-Crick base-pairing to form RNA-DNA duplexes such as R-loops. **(B)** Representation of potential type I RNA-chromatin (Continued)

FIGURE 1 | interactions, where RNAs recruit associated proteins to the chromatin, enabling the modulation of genomic accessibility. For example, recruited proteins can be part of chromatin modifying complexes or act as transcription factors. **(C)** Representation of potential type II RNA-chromatin interactions, where RNAs interact with genetically distant loci and modulate genome architecture, potentially bringing them physically closer. Interactions can be direct associations of the RNA with the DNA, or be mediated either by chromatin-bound proteins, or RNA duplexing. RNA bound protein complex intermediates can also generate indirect interactions. **(D)** Table of referenced techniques for identification of genome and transcriptome-wide DNA-RNA or RNA-RNA interactions, respectively. For more details, we refer the reader to the corresponding publications within the table. AMT, 4'-aminomethyl-4,5',8-trimethylpsoralen.

been described, for example: the lncRNAs *FIRRE* (Hacisuleyman et al., 2014; Lewandowski et al., 2019) and *ANRIL* (Kong et al., 2018). This hints at the great complexity of processes regulated by chromatin-associated RNAs.

We postulate that RNAs acting on chromatin can be functionally separated in two groups, following how they affect the transcriptional landscape of the cell. First, RNAs can act locally to where they bind (i.e., short-range) on the structure of the chromatin itself, for example, by modifying its accessibility through the recruitment of structural protein factors or protein complexes that establish chromatin marks (Type I) (**Figure 1B**). Various Type I RNAs have been characterized. For example, the lncRNAs *HOTAIR* and *FENDERR* and the sisRNA generated from intronic sequences of *SMYD3*, do so in part by recruiting the PRC2 repressive complex to their genomic binding sites (Guil et al., 2012; Grote et al., 2013; Mozdarani et al., 2020). Second, RNAs can control the organization of the genome, which we define as Type II interactions. They do so by promoting long-range chromatin interactions and the bridging of distant genomic loci (**Figure 1C**). Type II RNAs include examples such as the lncRNAs *LUNARI* or *Kancr*, both involved in chromosomal looping and activation of genes near the loop anchor points (Trimarchi et al., 2014; Li et al., 2020). Additionally, some lncRNAs can potentially exhibit both types of interactions. For example, the lncRNA *XIST* is transcribed from the X-chromosome and is directly implicated in its inactivation (XCI) (Loda and Heard, 2019). During this process, *XIST* transcripts recruit protein factors necessary for XCI, highlighting a Type I mechanism (Chu et al., 2015). Upon XCI, progressive compaction of the X-chromosome occurs and allows *XIST* to spread on the whole X-chromosome, which results in global heterochromatinization (Engreitz et al., 2013). Moreover, the multiple domains which enable *XIST* to attach to X-chromosome-bound proteins and the *XIST*-dependent X-chromosome conformation, hint at a potential Type II mechanism for this lncRNA in XCI.

While a growing number of RNAs acting on gene expression are being characterized, we still lack a bigger picture on how prevalent these interactions are, and on the importance of the proposed mechanisms. The first major obstacle in delineating which RNAs act on chromatin regulation resides in the immense number of existing transcripts. For example, estimates place the number of potentially functional lncRNAs in the tens of thousands (Marchese et al., 2017), and characterization of functional eRNAs or sisRNAs is still too early to grant accurate estimates (Li et al., 2016a; Chan and Pek, 2019). A tremendous amount of work lies ahead to fully understand the above-mentioned processes. Understanding where and how

RNAs interact with chromatin, and the resulting effect on gene regulation, therefore remains an upcoming challenge in the characterization of the global transcriptional landscape.

GENOME-WIDE CATALOGING OF RNA-CHROMATIN INTERACTIONS

The past 20 years have been marked by the development of several techniques aiming to map long-range chromatin interactions and to decipher genome architecture. Efficient techniques now include direct ligation of proximal DNA fragments (3C) (Kempfer and Pombo, 2020), ligation of barcodes to interacting DNA fragments (SPRITE) (Quinodoz et al., 2018), or physical isolation of thin nuclear sections and analysis of the DNA contained within (GAM) (Beagrie et al., 2017). The 3C-based methodologies have inspired new protocols to resolve RNA-DNA interactions, such as MARGI (Sridhar et al., 2017), GRID-seq (Li et al., 2017), CHAR-seq (Bell et al., 2018), iMARGI (Wu et al., 2019), and RADICL-seq (Bonetti et al., 2020). Moreover, SPRITE was also adapted to reveal RNA-DNA interactions (Quinodoz et al., 2018; **Figure 1D**). Therefore, a vast number of datasets showing RNA-chromatin interactions now exists for diverse cellular contexts. MARGI has already enabled the identification of Type I mechanisms implicating RNA-chromatin interactions in the establishment of various chromatin marks (Sridhar et al., 2017). Additionally, GRID-seq showed that an enrichment of RNA-chromatin interactions is implicated in the role of super-enhancers (Li et al., 2017), representing potential Types I and II RNA-chromatin interactions. GRID-seq was also used to detect the diverse genomic binding sites of *XIST*, both locally to its transcription site, but also throughout the X chromosome. It also highlighted the sites of initiation of XCI via *XIST*. These findings establish the major importance of such high-throughput techniques in characterizing different mechanisms responsible for the tight control of chromatin by RNAs.

LIMITATIONS OF CURRENT METHODOLOGIES FOR PROBING RNA-CHROMATIN INTERACTIONS

However, limitations are present in the developed techniques, and many challenges still lie ahead to fully understand how chromatin is regulated by RNAs. First, the developed techniques generally exhibit a high prevalence for RNA reads that map to introns, revealing a widespread capture of nascent RNAs from loci undergoing transcription (actively transcribed

RNAs) (Li et al., 2017; Bonetti et al., 2020). During their transcription process, RNAs can be captured via fixation to their genomic loci, therefore appearing as mapped interactions. These highly abundant interactions represent a major contaminant of such experiments, potentially overshadowing actual functional interactions. Their presence should be accounted for to prevent them from affecting the sequencing depth achieved for functional interactions and to allow for the detection of lower-abundance RNA-chromatin interactions. To reduce the effect of nascent RNA-bias, RADICL-seq uses a controlled RNase H digestion step (Bonetti et al., 2020). Although this step does indeed reduce this bias, a significant portion of nascent RNA still remains, and it would be necessary to explore additional ways to remove these unwanted RNAs.

Another main limitation in the current protocols and available datasets resides in the relatively small size of the sequenced tags corresponding to the DNA and its interacting RNA (Li et al., 2017; Sridhar et al., 2017; Bell et al., 2018; Quinodoz et al., 2018; Wu et al., 2019; Bonetti et al., 2020; **Figure 1D**). Indeed, their short size results in poor mapping of the obtained DNA-RNA pairs to the genome and transcriptome. This problem is further amplified when working with transcripts containing repeated sequences. For example, lncRNAs generally possess a large number of transposable elements (TEs) and other repeated sequences (Johnson and Guigó, 2014), and 83% of lncRNAs contain one or more TEs, compared to only 6% of mRNAs (Kelley and Rinn, 2012). Interestingly, some TEs are implicated in localization of RNA transcripts to the nucleus (Lubelsky and Ulitsky, 2018), which further highlights a role in genome regulation. These TEs have also emerged as important domains in lncRNA function, with several cases demonstrating that an embedded TE acts as the RNA's functional motif (Johnson and Guigó, 2014). These TEs can enable interactions with complementary sequences in DNA or other RNAs, as seems to be the case with the lncRNAs *ANRIL* (Holdt et al., 2013) and *LEADER* (Profumo et al., 2019). Therefore, a loss in reads corresponding to repeat elements due to an inability to map them represents a major potential hurdle in the current protocols. This is evident in the relatively low percentage of uniquely mapped reads for protocols such as GRID-seq or RADICL-seq (14 and 45%, respectively) (Li et al., 2017; Bonetti et al., 2020). The higher read mapping observed with RADICL-seq compared to GRID-seq is due to an additional 7 base-pairs in the final read length (Bonetti et al., 2020). These protocols use restriction enzymes to generate sequence lengths of around 20 base-pairs for GRID-seq and 27 for RADICL-seq for both RNA and DNA tags (**Figure 1D**). The increase in read length with RADICL-seq allows mapping of some reads corresponding to TE-containing RNAs. This enabled the authors to determine that transcripts containing TEs are indeed differentially engaged in interactions with chromatin, once again hinting at the importance of TEs in chromatin regulation. Other methodologies, namely MARGI (Sridhar et al., 2017) and iMARGI (Wu et al., 2019), circumvent this limitation through a protocol that preserves the full length of the respective RNA and DNA tags. This enables the generation of libraries containing longer fragments and results in higher mapping of reads.

Thus far, these techniques have relied on “short-read” sequencing technologies, which do not fully overcome the challenge of mapping repeated and complex sequences within RNAs (Dijk et al., 2018). To remediate this problem, one option would be to incorporate “long-read” sequencing technologies to these protocols. Indeed, throughout the last 10 years, the rise in the availability of such sequencing techniques has meant that more and more laboratories can get access to this technology. While it is still in its early stages compared to “short-read” sequencing approaches, the sequencing of long DNA or RNA fragments, ranging from a few hundred nucleotides to tens of kilobases, results in a more accurate alignment of repeat sequences (Dijk et al., 2018). A higher percentage of DNA and RNA pairs should therefore be uniquely mapped, even in the presence of interactions dependent on the complementarity of repeat elements. Additionally, long-read sequencing could provide more detailed information on the RNAs interacting with chromatin. For one, the specific isoforms of RNAs which interact with chromatin could be detected. Also, detection of nascent RNAs would be more precise, as current analysis only take intronic reads into account when counting for nascent RNAs, whereas long-read sequencing will reveal the whole transcript. Therefore, long-read sequencing represents a very promising new tool for further iterations of such protocols.

RNA-RNA INTERACTIONS TO HELP UNDERSTAND LNCRNA-CHROMATIN INTERACTIONS

While it is clear now that chromatin-associated RNAs affect the structure and regulation of chromatin, the role of RNA-RNA interactions (RRIs) is not as well-explored in this context. Due to RNA's inherent ability to base pair and form complex higher order structures, it can simultaneously interact with DNA and multiple RNA and protein molecules (Lu and Chang, 2018). Inter-RNA interactions could enable Type II interactions, through base-pairing between two chromatin-associated RNAs (**Figure 1C**). Meanwhile, intra-RNA interactions, through the RNA's secondary structure, may help in identifying sites available for binding to chromatin or RNA-binding proteins. With these features in mind, we expect that the associations revealed by the techniques probing RNA-chromatin interactions will serve as starting points for integration of RRI networks involved in chromatin regulation. Various protocols recently aimed to investigate RRIs on a transcriptome-wide scale, such as PARIS (Lu et al., 2016), LIGR-seq (Sharma et al., 2016), SPLASH (Aw et al., 2016), and MARIO (Nguyen et al., 2016). Additionally, another technique, COMRADES (Ziv et al., 2018), initially intended to probe RRIs for a single RNA, could also be used for genome-wide RRIs (**Figure 1D**). These protocols crosslink RNA duplexes, to reveal both RRIs and, to some extent, the secondary structure of every RNA. However, the limitations highlighted for RNA-DNA probing techniques, such as read length and unique mapping of repeated sequences, still apply to RRI-probing protocols. These techniques have nevertheless proven to be instrumental in revealing several cellular processes dependent on

RRIs. For example, PARIS highlighted structural folding patterns in the *XIST* A-repeats, which is necessary for binding of SPEN (Lu et al., 2016), a transcriptional repressor involved in X-inactivation (Chu et al., 2015). Analysis of RRIs in combination with RNA localization of *XIST* on the chromatin by GRID-seq (Li et al., 2017) consequently highlights the potential Type I and II mechanisms of *XIST* in XCI. This example demonstrates how combining RRI data to an integrative map of the genome organization extracted from 3C or other related techniques, coupled with existing RNA-chromatin interaction information, will provide a better understanding of the complex mechanisms behind chromatin regulation. Long-range, indirect chromatin interactions mediated by several duplexed RNAs, or by protein complexes exhibiting RNA-binding functions (Figure 1C) will only then become more apparent. Overall, these types of studies provide a more complete view on the complexity of genome organization and chromatin structure.

CONCLUDING REMARKS

Taken together, the methods described here represent useful tools for elucidating the role of RNA-DNA and RNA-RNA interactions in gene expression regulation. While various improvements are still needed, the existing datasets represent a comprehensive look of how a genome could be organized through RNA interactions. Several RNA-DNA interactions have now been cataloged either through these genome-wide techniques or through specific RNA

directed techniques. Of these, lncRNAs seem to represent an important fraction of the factors that regulate gene expression, chromatin accessibility and genome organization. In addition, these RNAs are not limited only to chromatin binding but may act as conduits to bring in other types of interactors, such as other RNAs, RNA-binding proteins, and transcriptional complexes. All these elements combined together help forge the transcriptional landscape necessary to maintain and transition between defined cell states.

AUTHOR CONTRIBUTIONS

GK and SH conceived and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by a grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) (grant # 2016-05847). SH was a Junior 1 Research Scholar of the Fonds de Recherche du Québec-Santé (FRQ-S).

ACKNOWLEDGMENTS

We would like to thank Victoire Fort for her input and help in proofreading the manuscript.

REFERENCES

- Aw, J. G. A., Shen, Y., Wilm, A., Sun, M., Lim, X. N., Boon, K.-L., et al. (2016). *In vivo* mapping of eukaryotic RNA interactomes reveals principles of higher-order organization and regulation. *Mol. Cell* 62, 603–617. doi: 10.1016/j.molcel.2016.04.028
- Beagrie, R. A., Scialdone, A., Schueler, M., Kraemer, D. C. A., Chotalia, M., Xie, S. Q., et al. (2017). Complex multi-enhancer contacts captured by genome architecture mapping. *Nature* 543, 519–524. doi: 10.1038/nature21411
- Bell, J. C., Jukam, D., Teran, N. A., Risca, V. I., Smith, O. K., Johnson, W. L., et al. (2018). Chromatin-associated RNA sequencing (ChAR-seq) maps genome-wide RNA-to-DNA contacts. *Elife* 7:e27024. doi: 10.7554/eLife.27024.037
- Blank-Giwojna, A., Postepska-Igielska, A., and Grummt, I. (2019). lncRNA KHPS1 activates a poised enhancer by triplex-dependent recruitment of epigenomic regulators. *Cell Rep.* 26, 2904.e4–2915.e4. doi: 10.1016/j.celrep.2019.02.059
- Bonasio, R., and Shiekhattar, R. (2014). Regulation of Transcription by Long Noncoding RNAs. *Annu. Rev. Genet.* 48, 1–23. doi: 10.1146/annurev-genet-120213-092323
- Bonetti, A., Agostini, F., Suzuki, A. M., Hashimoto, K., Pascarella, G., Gimenez, J., et al. (2020). RADICL-seq identifies general and cell type-specific principles of genome-wide RNA-chromatin interactions. *Nat. Commun.* 11:1018. doi: 10.1038/s41467-020-14337-6
- Boque-Sastre, R., Soler, M., Oliveira-Mateos, C., Portela, A., Moutinho, C., Sayols, S., et al. (2015). Head-to-head antisense transcription and R-loop formation promotes transcriptional activation. *Proc. Natl. Acad. Sci. U.S.A.* 112, 5785–5790. doi: 10.1073/pnas.1421197112
- Cech, T. R., and Steitz, J. A. (2014). The noncoding RNA revolution—trashing old rules to forge new ones. *Cell* 157, 77–94. doi: 10.1016/j.cell.2014.03.008
- Chan, S. N., and Pek, J. W. (2019). Stable intronic sequence RNAs (sisRNAs): an expanding universe. *Trends Biochem. Sci.* 44, 258–272. doi: 10.1016/j.tibs.2018.09.016
- Chu, C., Zhang, Q. C., da Rocha, S. T., Flynn, R. A., Bharadwaj, M., Calabrese, J. M., et al. (2015). Systematic discovery of xist RNA binding proteins. *Cell* 161, 404–416. doi: 10.1016/j.cell.2015.03.025
- Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat. Rev. Genet.* 2, 292–301. doi: 10.1038/35066075
- Dijk, E. L., van, Jaszczyszyn, Y., Naquin, D., and Thermes, C. (2018). The third revolution in sequencing technology. *Trends Genet.* 34, 666–681. doi: 10.1016/j.tig.2018.05.008
- Engreitz, J. M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., et al. (2013). The xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* 341:1237973. doi: 10.1126/science.1237973
- Gibbons, H. R., Shaginurova, G., Kim, L. C., Chapman, N., Spurlock, C. F., and Aune, T. M. (2018). Divergent lncRNA GATA3-AS1 regulates GATA3 transcription in T-Helper 2 cells. *Front. Immunol.* 9:2512. doi: 10.3389/fimmu.2018.02512
- Grote, P., Witter, L., Hendrix, D., Koch, F., Währisch, S., Beisaw, A., et al. (2013). The tissue-specific lncRNA fendrr is an essential regulator of heart and body wall development in the mouse. *Dev. Cell* 24, 206–214. doi: 10.1016/j.devcel.2012.12.012
- Guil, S., Soler, M., Portela, A., Carrère, J., Fonalleras, E., Gómez, A., et al. (2012). Intronic RNAs mediate EZH2 regulation of epigenetic targets. *Nat. Struct. Mol. Biol.* 19, 664–670. doi: 10.1038/nsmb.2315
- Hacisuleyman, E., Goff, L. A., Trapnell, C., Williams, A., Henao-Mejia, J., Sun, L., et al. (2014). Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat. Struct. Mol. Biol.* 21, 198–206. doi: 10.1038/nsmb.2764
- Holdt, L. M., Hoffmann, S., Sass, K., Langenberger, D., Scholz, M., Krohn, K., et al. (2013). Alu Elements in ANRIL non-coding RNA at chromosome 9p21 modulate atherogenic cell functions through trans-regulation of gene networks. *PLoS Genet.* 9:e1003588. doi: 10.1371/journal.pgen.1003588

- Johnson, R., and Guigó, R. (2014). The RIDL hypothesis: transposable elements as functional domains of long noncoding RNAs. *RNA* 20, 959–976. doi: 10.1261/rna.044560.114
- Kelley, D., and Rinn, J. (2012). Transposable elements reveal a stem cell-specific class of long noncoding RNAs. *Genome Biol.* 13:R107. doi: 10.1186/gb-2012-13-11-r107
- Kempfer, R., and Pombo, A. (2020). Methods for mapping 3D chromosome architecture. *Nat. Rev. Genet.* 21, 207–226. doi: 10.1038/s41576-019-0195-2
- Klemm, S. L., Shipony, Z., and Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* 20, 207–220. doi: 10.1038/s41576-018-0089-8
- Kong, Y., Hsieh, C.-H., and Alonso, L. C. (2018). ANRIL: a lncRNA at the CDKN2A/B locus with roles in cancer and metabolic disease. *Front. Endocrinol.* 9:405. doi: 10.3389/fendo.2018.00405
- Lewandowski, J. P., Lee, J. C., Hwang, T., Sunwoo, H., Goldstein, J. M., Groff, A. F., et al. (2019). The firre locus produces a trans-acting RNA molecule that functions in hematopoiesis. *Nat. Commun.* 10:5137. doi: 10.1038/s41467-019-12970-4
- Li, W., Notani, D., and Rosenfeld, M. G. (2016a). Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat. Rev. Genet.* 17, 207–223. doi: 10.1038/nrg.2016.4
- Li, W., Shen, W., Zhang, B., Tian, K., Li, Y., Mu, L., et al. (2020). Long non-coding RNA LncKdm2b regulates cortical neuronal differentiation by cis-activating Kdm2b. *Protein Cell* 11, 161–186. doi: 10.1007/s13238-019-0650-z
- Li, X., and Fu, X.-D. (2019). Chromatin-associated RNAs as facilitators of functional genomic interactions. *Nat. Rev. Genet.* 20, 503–519. doi: 10.1038/s41576-019-0135-1
- Li, X., Zhou, B., Chen, L., Gou, L.-T., Li, H., and Fu, X.-D. (2017). GRID-seq reveals the global RNA–chromatin interactome. *Nat. Biotechnol.* 35:nbt3968. doi: 10.1038/nbt.3968
- Li, Y., Syed, J., and Sugiyama, H. (2016b). RNA-DNA Triplex Formation by Long Noncoding RNAs. *Cell Chem. Biol.* 23, 1325–1333. doi: 10.1016/j.chembiol.2016.09.011
- Loda, A., and Heard, E. (2019). Xist RNA in action: past, present, and future. *PLoS Genet.* 15:e1008333. doi: 10.1371/journal.pgen.1008333
- Lu, Z., and Chang, H. Y. (2018). The RNA base-pairing problem and base-pairing solutions. *Cold Perspect Biol.* 10:a034926. doi: 10.1101/cshperspect.a034926
- Lu, Z., Zhang, Q. C., Lee, B., Flynn, R. A., Smith, M. A., Robinson, J. T., et al. (2016). RNA duplex map in living cells reveals higher-order transcriptome structure. *Cell* 165, 1267–1279. doi: 10.1016/j.cell.2016.04.028
- Lubelsky, Y., and Ulitsky, I. (2018). Sequences enriched in Alu repeats drive nuclear localization of long RNAs in human cells. *Nature* 555, 107–111. doi: 10.1038/nature25757
- Marchese, F. P., Raimondi, I., and Huarte, M. (2017). The multidimensional mechanisms of long noncoding RNA function. *Genome Biol.* 18:206. doi: 10.1186/s13059-017-1348-2
- Mondal, T., Subhash, S., Vaid, R., Enroth, S., Uday, S., Reinius, B., et al. (2015). MEG3 long noncoding RNA regulates the TGF- β pathway genes through formation of RNA-DNA triplex structures. *Nat. Commun.* 6:7743. doi: 10.1038/ncomms8743
- Mozdarani, H., Ezzatizadeh, V., and Parvaneh, R. R. (2020). The emerging role of the long non-coding RNA HOTAIR in breast cancer development and treatment. *J. Transl. Med.* 18:152. doi: 10.1186/s12967-020-02320-0
- Nguyen, T. C., Cao, X., Yu, P., Xiao, S., Lu, J., Biase, F. H., et al. (2016). Mapping RNA–RNA interactome and RNA structure *in vivo* by MARIO. *Nat. Commun.* 7:12023. doi: 10.1038/ncomms12023
- Noh, J. H., Kim, K. M., McClusky, W. G., Abdelmohsen, K., and Gorospe, M. (2018). Cytoplasmic functions of long noncoding RNAs. *Wiley Interdiscip. Rev. RNA* 9:e1471. doi: 10.1002/wrna.1471
- Profumo, V., Forte, B., Percio, S., Rotundo, F., Doldi, V., Ferrari, E., et al. (2019). LEADeR role of miR-205 host gene as long noncoding RNA in prostate basal cell differentiation. *Nat. Commun.* 10:307. doi: 10.1038/s41467-018-08153-2
- Quinodoz, S. A., Ollikainen, N., Tabak, B., Palla, A., Schmidt, J. M., Detmar, E., et al. (2018). Higher-order inter-chromosomal hubs shape 3D genome organization in the nucleus. *Cell* 174, 744.e24–757.e24. doi: 10.1016/j.cell.2018.05.024
- Sharma, E., Sterne-Weiler, T., O'Hanlon, D., and Blencowe, B. J. (2016). Global mapping of human RNA–RNA interactions. *Mol. Cell* 62, 618–626. doi: 10.1016/j.molcel.2016.04.030
- Sridhar, B., Rivas-Astroza, M., Nguyen, T. C., Chen, W., Yan, Z., Cao, X., et al. (2017). Systematic mapping of RNA–chromatin interactions *In vivo*. *Curr. Biol.* 27, 602–609. doi: 10.1016/j.cub.2017.01.011
- Trimarchi, T., Bilal, E., Ntziachristos, P., Fabbri, G., Dalla-Favera, R., Tsigos, A., et al. (2014). Genome-wide mapping and characterization of Notch-regulated long noncoding RNAs in acute leukemia. *Cell* 158, 593–606. doi: 10.1016/j.cell.2014.05.049
- Vance, K. W., and Ponting, C. P. (2014). Transcriptional regulatory functions of nuclear long noncoding RNAs. *Trends Genet.* 30, 348–355. doi: 10.1016/j.tig.2014.06.001
- Wu, W., Yan, Z., Nguyen, T. C., Chen, Z. B., Chien, S., and Zhong, S. (2019). Mapping RNA–chromatin interactions by sequencing with iMARGI. *Nat. Protoc.* 14, 3243–3272. doi: 10.1038/s41596-019-0229-4
- Ziv, O., Gabryelska, M. M., Lun, A. T. L., Gebert, L. F. R., Sheu-Gruttadauria, J., Meredith, L. W., et al. (2018). COMRADES determines *in vivo* RNA structures and interactions. *Nat. Methods* 15, 785–788. doi: 10.1038/s41592-018-0121-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a shared affiliation with the authors though no other collaboration at time of review.

Copyright © 2020 Khelifi and Hussein. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Interactions With Histone H3 & Tools to Study Them

William A. Scott^{1,2} and Eric I. Campos^{1,2*}

¹ Genetics & Genome Biology Program, The Hospital for Sick Children, Toronto, ON, Canada, ² Department of Molecular Genetics, University of Toronto, Toronto, ON, Canada

OPEN ACCESS

Edited by:

Jean-Philippe Lambert,
Laval University, Canada

Reviewed by:

Alejandra Loyola,
Fundación Ciencia & Vida, Chile
Nicole Francis,
Institute of Clinical Research De
Montreal (IRCM), Canada

*Correspondence:

Eric I. Campos
eric.campos@sickkids.ca

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 15 May 2020

Accepted: 10 July 2020

Published: 31 July 2020

Citation:

Scott WA and Campos EI (2020)
Interactions With Histone H3 & Tools
to Study Them.
Front. Cell Dev. Biol. 8:701.
doi: 10.3389/fcell.2020.00701

Histones are an integral part of chromatin and thereby influence its structure, dynamics, and functions. The effects of histone variants, posttranslational modifications, and binding proteins is therefore of great interest. From the moment that they are deposited on chromatin, nucleosomal histones undergo dynamic changes in function of the cell cycle, and as DNA is transcribed and replicated. In the process, histones are not only modified and bound by various proteins, but also shuffled, evicted, or replaced. Technologies and tools to study such dynamic events continue to evolve and better our understanding of chromatin and of histone proteins proper. Here, we provide an overview of H3.1 and H3.3 histone dynamics throughout the cell cycle, while highlighting some of the tools used to study their protein–protein interactions. We specifically discuss how histones are chaperoned, modified, and bound by various proteins at different stages of the cell cycle. Established and select emerging technologies that furthered (or have a high potential of furthering) our understanding of the dynamic histone–protein interactions are emphasized. This includes experimental tools to investigate spatiotemporal changes on chromatin, the role of histone chaperones, histone posttranslational modifications, and histone-binding effector proteins.

Keywords: H3.1, H3.3, histone, nucleosome, chromatin, epigenetic, proteomic

INTRODUCTION

Chromatin is composed of DNA and associated proteins, of which histones are prominent. Histones and DNA assemble to form repetitive units known as nucleosomes. Each nucleosome organizes a stretch of ~147 bp of DNA wrapped around a histone octamer (Luger et al., 1997). The octamer is, in turn, composed of a central (H3-H4)₂ tetramer, flanked by two H2A-H2B dimers. Nucleosomal arrays give rise to an 11 nm fiber that resembles “beads on a string,” as seen in early micrographs of chromatin (Olins et al., 1976). Histone proteins are heavily modified through combinatorial posttranslational modifications (PTMs), especially over their N-terminal tails that protrude from the nucleosomal core (Huang et al., 2015a; Andrews et al., 2016). These modifications influence local protein–protein interactions (PPIs) and chromatin structures, and consequently have important implications on DNA accessibility, transcription, repair, and replication. A large number of histone PTMs thereby correlate, or anti-correlate, with various biological outputs (Campos and Reinberg, 2009; Allis and Jenuwein, 2016).

This review highlights current models of H3.1 and H3.3 dynamics, namely: the histones, their PTMs, deposition pathways, and cell cycle dynamics. Each section provides an overview of the techniques used to formulate the models.

HISTONE VARIANTS AND POST-TRANSLATIONAL MODIFICATIONS

Certain histone variants are believed to influence the biophysical characteristics of nucleosomes (Campos and Reinberg, 2009), thereby relaying functional consequences on chromatin. Histone H2A has a relatively high number of variants, while histones H2B and H4 have undergone little evolutionary divergence—likely reflecting their positions within the nucleosome and their roles in stabilization of the nucleosome core particle (Henikoff and Smith, 2015). There are, however, a large number of histone H3 variants in humans, namely H3.1, H3.2, H3.3, H3t/H3.4, H3.5, H3.Y, H3.X, CENP-A, the more recently proposed H3.3-like H3.6 and H3.8, as well as the H3.1-like H3.7 (Franklin and Zweidler, 1977; Earnshaw and Rothfield, 1985; Albig et al., 1996; Wiedemann et al., 2010; Schenk et al., 2011; Taguchi et al., 2017). Of these, the replication-coupled H3.1, and replication-independent H3.3 variants are arguably some of the better-studied histone proteins and hence the focus herein. A broader overview of known histone variants is available elsewhere (Talbert et al., 2012; Biterge and Schneider, 2014; Henikoff and Smith, 2015).

H3.1 and H3.2 differ by a single amino acid at residue 96 (**Figure 1**; Hake and Allis, 2006). Both are expressed in S-phase (Wu et al., 1982; Mendiratta et al., 2019) and deposited on replicating DNA (Tagami et al., 2004). They are, therefore, considered to be replication-coupled (RC) histones. Reflecting the need for considerable histone production during DNA replication, RC histones are expressed from histone gene clusters in S-phase (Wu and Bonner, 1981). As such, H3.1 predominates in cycling cells (Wu et al., 1982; Marzluff et al., 2002).

Conversely, the H3.3 variant (encoded by *H3F3A* and *H3F3B*) is expressed at low levels throughout interphase (Wu and Bonner, 1981; Mendiratta et al., 2019) to maintain proper nucleosome density as histones turnover. It is thus referred to as replication-independent (RI). H3.3 is particularly enriched over actively transcribed genes, but is also deposited over repetitive DNA elements, such as pericentromeric regions and telomeres (Ahmad and Henikoff, 2002; Drane et al., 2010; Goldberg et al., 2010). It accumulates in terminally differentiated cells (Grove and Zweidler, 1984), and is also the only non-centromeric H3 variant in some species (e.g., yeast within the Ascomycota phylum) (Postberg et al., 2010; Talbert and Henikoff, 2010). Astonishingly, the H3.3 RI histone differs from H3.1 by only five residues (Hake and Allis, 2006). Yet, these minute differences are sufficient to confer specificity to distinct interacting proteins, such as histone chaperones (Elsasser et al., 2012; Ricketts et al., 2015).

Histone pools are exquisitely regulated at the transcriptional and posttranscriptional levels. Reduced histone transcription disturbs the cell cycle (Nelson et al., 2002) and excessive production of histones outside of S-phase leads to chromosomal

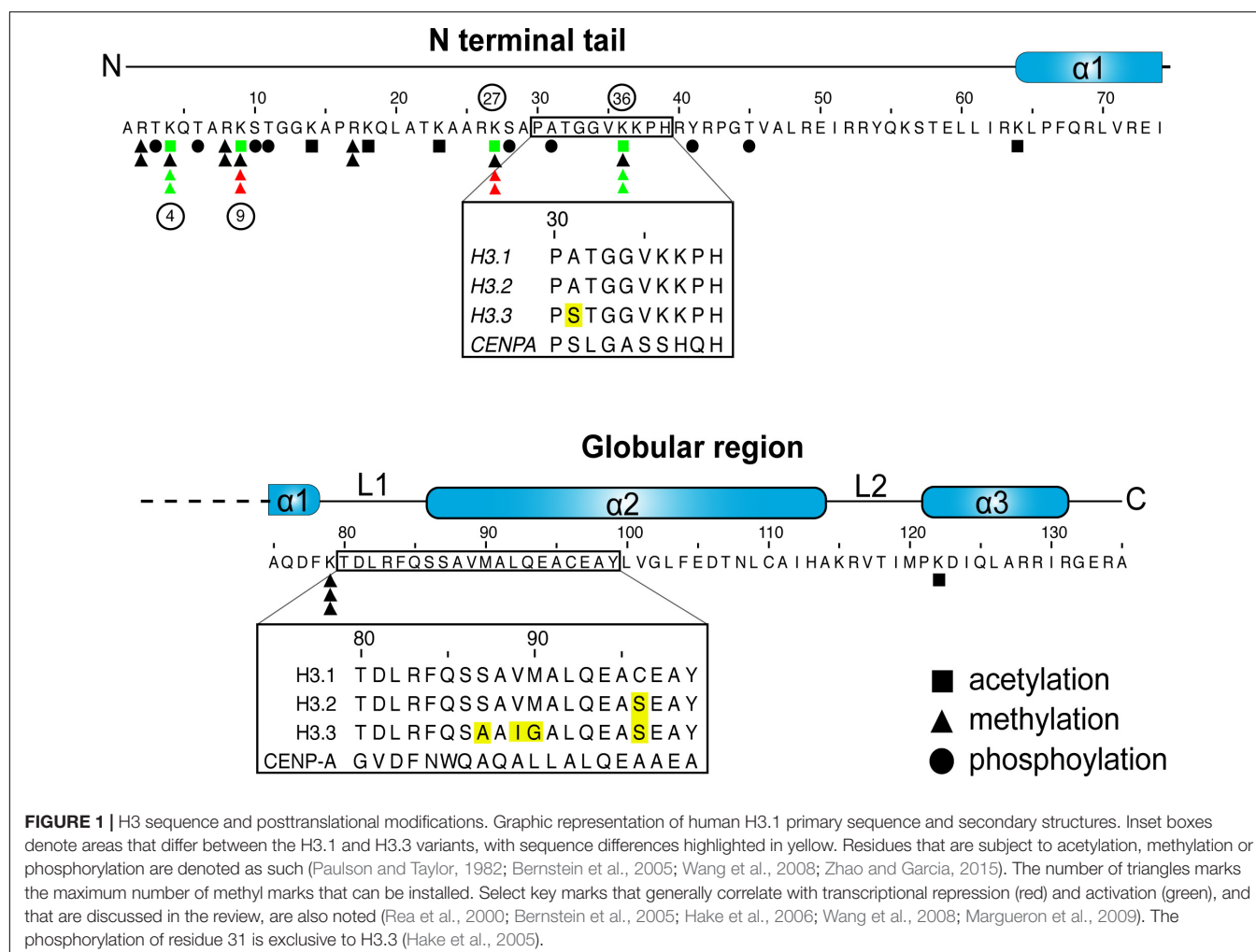
instability (Gunjan and Verreault, 2003), partly through a stoichiometric imbalance (Meeks-Wagner and Hartwell, 1986). Soluble histones pools are also kept in check by certain histone chaperones (Groth et al., 2005; Cook et al., 2011). Histones are then deposited onto DNA to form nucleosomes and specialize local chromatin regions (Campos and Reinberg, 2009). Once on chromatin, histones remain highly dynamic, even when deposited in heterochromatic regions (Consortium et al., 2007; Deal et al., 2010). Such dynamics are particularly evident as DNA is transcribed, repaired, replicated, and condensed.

A large number of histone residues are subject to various PTMs, including methylation, acetylation, and phosphorylation, to name but a few [see Zhao and Garcia (2015) for a comprehensive list]. Combinatorial histone PTMs, particularly over the lysine-rich N-terminal histone tails influence local chromatin structures and dynamics, and often correlate with transcriptional status (Zhao and Garcia, 2015; Allis and Jenuwein, 2016; **Figure 1**).

Histone acetylation has long been shown to correlate with active gene transcription (Allfrey et al., 1964; Gorovsky et al., 1973; Davie and Candido, 1978; Chahal et al., 1980). The PTM neutralizes the positive charge on the ϵ -amino group of lysine residues, leading to numerous downstream events. For one, lysine acetylation on histone proteins is believed to counteract chromatin compaction (Wong and Marushige, 1976; Wallace et al., 1977; Nelson et al., 1978; Simpson, 1978; Vidali et al., 1978; Annunziato et al., 1988; Tse et al., 1998). For example, the acetylation of lysine 122 on H3 (H3K122ac) has been found to destabilize nucleosomes by disrupting histone–DNA interactions (Tropberger et al., 2013). Beyond direct biophysical effects, the binding of numerous effector proteins that “read” modified histones further influences chromatin structures. For example, acetyl marks are recognized by the bromodomain, YEATS, or plant homeodomain (PHD) of some chromatin-associated proteins. Similarly, methylated lysines are recognized by a “Royal Family” (tudor, MBT, chromodomain, PWWP), as well as numerous other domains, including PHD, WD40, ankyrin repeats, BAH, and ADD (see Patel and Wang, 2013; Andrews et al., 2016).

“Reader proteins” exert further effects on chromatin. The H3K9me2/3 and H3K27me2/3 marks, for example, correlate with chromatin compaction and transcriptional repression as a result of some proteins that bind these marks. H3K9me3 enriches at constitutive heterochromatic regions, such as pericentric chromatin and telomeres (Talbert and Henikoff, 2006). The mark is catalyzed by the SUV39H1/2 histone methyltransferases in humans (Rea et al., 2000), which is, in turn, recognized by the heterochromatin protein 1 (HP1) via its chromodomain (Smothers and Henikoff, 2000; Lachner et al., 2001). This process is sustained through a positive feedback loop, where HP1 re-recruits SUV39H1/2 to propagate the mark (Talbert and Henikoff, 2006). The HP1 protein further phase separates—that is, adopts liquid-like properties to form a membraneless compartment—thereby driving chromatin compaction (Larson et al., 2017; Strom et al., 2017).

Similarly, the H3K27me3 mark—which is particularly enriched over facultative heterochromatin—also spreads



through a positive feedback loop driven by the polycomb repressive complex 2 (PRC2). The EED subunit of this complex binds the H3K27me3 mark to allosterically activate the EZH2 catalytic subunit, thereby propagating the mark to neighboring nucleosomes (Margueron et al., 2009; Oksuz et al., 2018). Our current understanding of polycomb proteins is evolving at a fast pace, and is well discussed in recent publications (Cheutin and Cavalli, 2019; Laugesen et al., 2019; van Mierlo et al., 2019; Yu et al., 2019; Chammas et al., 2020).

Not all histones marks are believed to alter chromatin structures but can still influence biological events by preventing or promoting interactions with other proteins. For example, H3K4me2/3 enriches near the transcriptional start site (TSS) of actively transcribed genes (Bernstein et al., 2005; Kim et al., 2005; Roh et al., 2006). The mark alone fails to stimulate transcription *in vitro* (Pavri et al., 2006), but does prevent the installment of repressive H3K9 and H3K27 methyl marks on the same histone tail (Binda et al., 2010; Schmitges et al., 2011; Voigt et al., 2012). Just as importantly, the mark facilitates a number of events, including transcriptional initiation, splicing, and even termination (Sims et al., 2007; Vermeulen et al., 2007; Terzi et al., 2011).

H3K36me2/3 is another important mark that correlates with gene expression, but enriches over transcribed gene bodies (Bannister et al., 2005). Like H3K4me3, it is also inhibitory toward PRC2 activity on the same histone tail (Schmitges et al., 2011; Yuan et al., 2011; Voigt et al., 2012). The mark has been notably associated with transcriptional elongation, splicing, and the inhibition of cryptic transcription (Carrozza et al., 2005; Keogh et al., 2005; Luco et al., 2010).

Beyond the examples listed above, innumerable combinatorial PTMs coupled to dynamic effects of histone “writers,” “erasers,” and “readers” add further complexity. Therefore, a single histone mark may be impactful in numerous ways and the study of such plasticity requires suitable technologies. New sequencing-based techniques have had an immense impact on that front (Dirks et al., 2016; Nakato and Sakata, 2020; Stewart-Morgan et al., 2020), as have recent developments in mass spectrometry (MS)-based pipelines (Gingras et al., 2007; Eubanks et al., 2017; Simithy et al., 2018; Sequeira and Vermeulen, 2019; Samavarchi-Tehrani et al., 2020). There is, however, also great excitement on ever evolving molecular and biochemical techniques to study histone dynamics, their marks, and PPIs.





TOOLS TO STUDY INTERACTIONS WITH HISTONE PTMs

Numerous techniques that are used to study histone PTMs rely on antibodies to recognize (**Figure 2A**) or isolate associated proteins or to map their genomic location (**Figure 3**). Alternatives, such as recombinant antibodies and purified histone modification interaction domains (HMIDs), are being developed and show promise (Kungulovski et al., 2014; Hattori and Koide, 2018). For example, a recently engineered HP1 chromodomain is reported to surpass antibodies in avidity when binding H3K9me3, without losing specificity (Albanese et al., 2020). Once validated, HMIDs can be expressed and purified at the required scale, eliminating lot variations associated with polyclonal antibodies. Until these alternatives become commonplace, antibody use remains the standard. It is thereby critical to emphasize the need to thoroughly validate the specificity of antibody or binding module (Rothbart et al., 2015).

Of the various antibody-based techniques used to study histone PTMs, chromatin immunoprecipitation (ChIP), has been

particularly powerful (Gilmour and Lis, 1984; Hebbes et al., 1988; Solomon et al., 1988). Coupled to microarray (ChIP-chip), sequencing (e.g., ChIP-SAGE), and later next-generation sequencing (ChIP-seq), ChIP-based techniques facilitated the genomic mapping of histone PTMs, and studies that correlated histone PTMs with various chromatin states or biological effects (Robyr et al., 2002; Roh et al., 2004; Barski et al., 2007). In addition to mapping histones, PTMs, and histone binding-proteins, an increasing number of variations on the ChIP technique are used to investigate histone dynamics [e.g., ChOR-seq, SCAR-seq – see Stewart-Morgan et al. (2020)]. Furthermore, direct analysis of ChIP material (e.g., ChIP-western or ChIP-MS) is possible, and can inform on protein associations within chromatin fragments containing specific histone PTMs (Ji et al., 2015). As with all ChIP-based experiments, the technique is limited by antibody specificity, the abundance of the epitope, and downstream detection (e.g., MS). It, however, is a relatively accessible technique that is applicable toward different ends.

Histone peptides or nucleosomal particles containing specific PTM(s) are perhaps more commonly immobilized to isolate

A PTM Detection		Advantages	Disadvantages	References
Polyclonal Antibodies		<ul style="list-style-type: none"> Widely available 	<ul style="list-style-type: none"> Can lack specificity 	Rothbart et al., 2015
Monoclonal & Recombinant Antibodies		<ul style="list-style-type: none"> Mostly available Improves reproducibility 	<ul style="list-style-type: none"> Labor-intensive Costly 	Hattori & Koide, 2018
Histone PTM Interaction Domains		<ul style="list-style-type: none"> Improves reproducibility Easily expressed Cost-effective 	<ul style="list-style-type: none"> Requires engineering & careful validation 	Kungulovski et al., 2014 Villasenor et al., 2020
Mass Spectrometry		<ul style="list-style-type: none"> Unbiased 	<ul style="list-style-type: none"> Costly Requires more input material 	Simithy et al., 2018

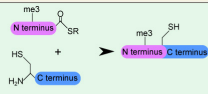
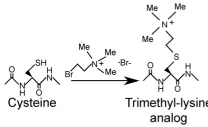
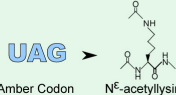
B PTM Installment		Advantages	Disadvantages	References
Native Chemical Ligation		<ul style="list-style-type: none"> Installs specific & combinatorial PTMs 	<ul style="list-style-type: none"> Costly (peptides) 	He et al., 2003
PTM analogs		<ul style="list-style-type: none"> Cost-effective 	<ul style="list-style-type: none"> May not perfectly recapitulate PTM 	Simon et al., 2007
Evolved tRNA/tRNA synthetase pair		<ul style="list-style-type: none"> Scarless 	<ul style="list-style-type: none"> Labor-intensive 	Neumann et al., 2008

FIGURE 2 | Detecting and installing histone posttranslational modifications (PTMs). **(A)** Common tools used to detect histone PTMs. Antibodies remain standard; however, they can lack specificity and require proper validation. Alternatives, such as recombinant antibodies and histone PTM interaction domains are increasingly available. Mass spectrometry can also provide an unbiased detection. **(B)** Techniques used to install and study specific histone marks. Modified histones are important for the study of histone protein-protein interactions that are modulated by PTMs.

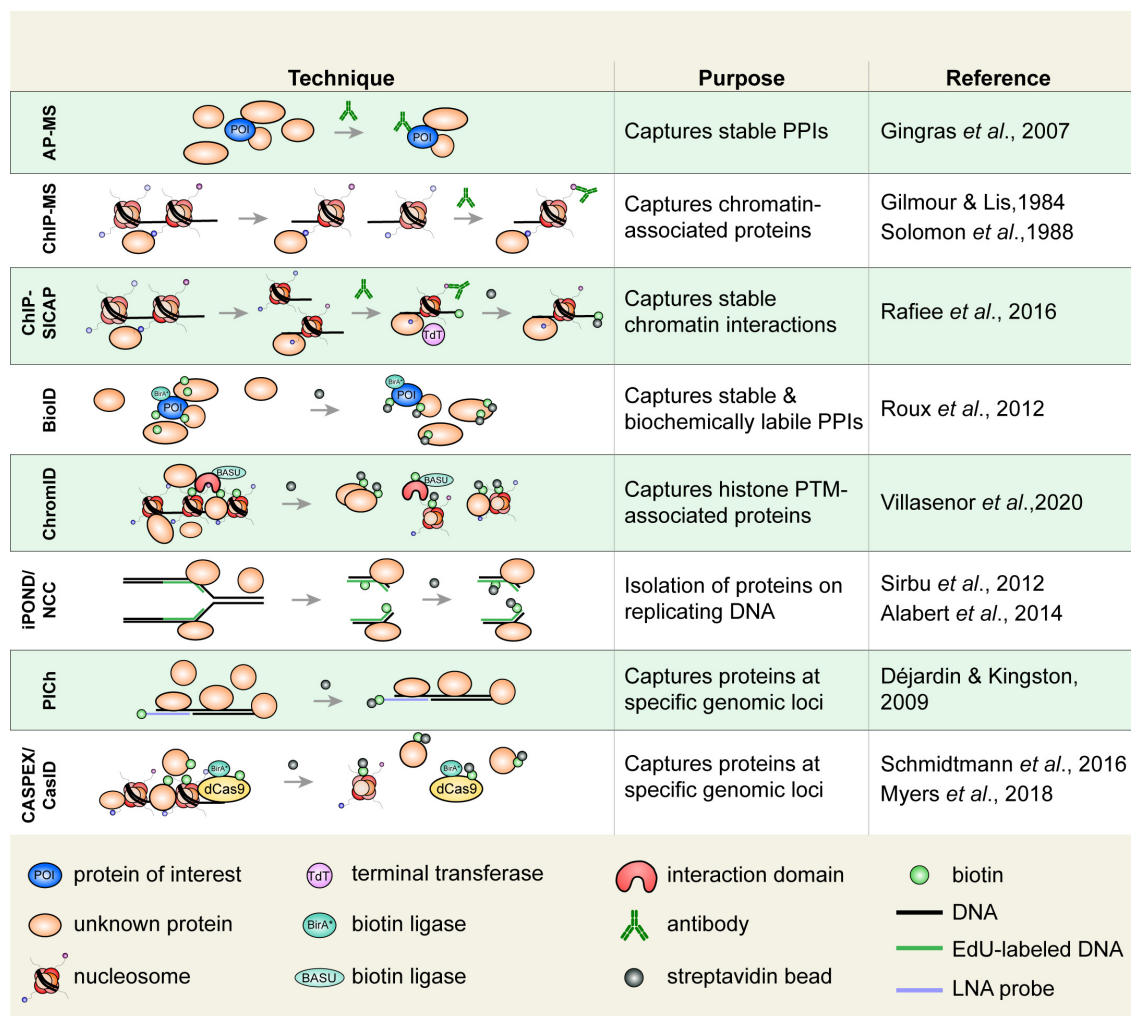


FIGURE 3 | Select techniques used to study histone or chromatin-associated proteins. Affinity purification-mass spectrometry (AP-MS) is used to isolate biochemically stable protein-protein interactions. Chromatin immunoprecipitation (ChIP), in which an epitope of interest is isolated from sheared chromatin fragments, can be coupled to MS to identify associated proteins. ChIP-SICAP uses an additional DNA biotinylation step to wash proteins not directly bound to chromatin. Proximity-dependent labeling techniques are increasingly used to capture stable and biochemically labile protein interactions. Biotin identification (BioID) uses a biotin ligase fused to a protein of interest to biotinylate proximal proteins. Biotinylated proteins are captured on streptavidin beads. ChromID is similar to BioID in that a biotin ligase, BASU, is fused to a histone-binding domain to biotinylate proteins near a PTM of interest. Isolation of proteins on nascent DNA (iPOND)/nascent chromatin capture (NCC), are used to isolate proteins associated with replicating DNA. Cells are pulsed with a thymidine analog (e.g., EdU), which is incorporated on replicated DNA enabling the isolation of replicated chromatin fragments. Proteomics of isolated chromatin (PiCh) is used to identify proteins that are bound to a specific genomic region. A biotin-tagged locked nucleic acid (LNA) probe is used to isolate chromatin fragments with DNA complementary to the probe. In CASPEX/CasID, catalytically dead Cas9 (dCas9) is fused to APEX or BirA*, respectively, to perform biotin labeling of a specific genomic locus. PPIs, protein-protein interactions.

and identify histone readers. There are different approaches that are used to install PTMs on histone proteins *in vitro* (Figure 2B). Histones can, of course, be enzymatically modified *in vitro*, but that risks yielding a mixture of modified and unmodified histones, modifying more than one residue, or results in multiple states (e.g., mono-, di-, tri- lysine methylation). Small histone peptides containing specific marks can also be chemically synthesized, but they then lack the important nucleosomal context.

It is possible to obtain homogeneously modified histones by chemically ligating synthetic histone tails (containing specific

marks) to tailless recombinant histones (He *et al.*, 2003; Shogren-Knaak *et al.*, 2003). In this system, a synthesized C-terminal histone globular region containing an N-terminal cysteine and a synthesized N-terminal histone tail containing a C-terminal thioester are spontaneously ligated to produce full-length protein. The technique has helped elucidate mechanisms of chromatin readers, such as that of the BPTF protein, which simultaneously binds the H3K4me2/3 and H4K16ac marks (Ruthenburg *et al.*, 2011). There are also strategies to install site-specific PTM analogs. In this approach, the residue of interest is mutated to a cysteine and an aminoethylation reaction ligates a PTM

analog. This semisynthetic method has generated numerous PTM mimics, including methyl lysine analogs (often referred to as MLAs) and acetyl lysine mimics (Simon et al., 2007; Huang et al., 2010). Though the PTM analogs allow for a qualitative analysis of binding proteins, the sulfide substitution on the side chain has been suggested to affect binding strength of certain interactions (Seeliger et al., 2012; Chen Z. et al., 2018). Nevertheless, the semisynthetic strategies remain a formidable tool (reviewed by Holt and Muir, 2015).

Biorthogonal systems allow for the expression of “scarless” recombinant histones. They have been developed to express dedicated tRNA/tRNA-synthetase pairs to expand the genetic code and incorporate modified amino acids into recombinant proteins expressed in bacteria (Neumann et al., 2008), or even exogenous histones expressed in mammalian cells (Elsasser et al., 2016). While an effective system, the evolution of tRNA/tRNA-synthetase pairs is labor-intensive. Regardless of the process used, the modified histones can be assembled into nucleosomes, immobilized to capture interacting proteins, or used to study histone PTM-protein interactions.

Candidate proteins can also be tested against multiple histone PTMs to survey the marks they recognize. Peptide arrays contain short, synthetic histone tails etched on a solid surface. By containing different PTMs, each polypeptide allows researchers to discern the ability of a reader to bind to specific marks or combinations thereof (Mauser and Jeltsch, 2019). However, because these arrays only contain a portion of the histone tail, they may skew binding. As with all screening techniques, a careful validation is required. The technique, however, is often used and has generated insightful data pertaining to histone PTM-binding proteins. A refined quantitative approach using immobilized peptides, combined with stable isotope labeling by amino acids in cell culture (SILAC) was used to distinguish specific binders over background binders (Vermeulen et al., 2007, 2010). By incubating modified peptide with lysates from cells grown in the presence of heavy isotopes, and unmodified peptides with lysates from cells grown in the presence of light isotopes, the general transcription factor TFIID was found to associate with the H3K4me3 mark using mass spectrometry (Vermeulen et al., 2007).

Newer technologies also allow for the identification of histone PTM binding proteins. Proximity-dependent labeling approaches (Martell et al., 2012; Roux et al., 2012)—discussed in greater detail below—have been revolutionizing the proteomics field. A clever take on the technique, called ChromID uses engineered protein modules to bind histone PTMs (Villasenor et al., 2020). Further coupled to a promiscuous biotin ligase enzyme, the module biotinylates proteins that are directly or indirectly associated with the associated histone PTM. The biotinylated proteins are then captured on streptavidin beads under denaturing conditions, and identified by MS. To further capitalize on the technique, a larger collection of interacting modules will now need to be developed.

While several histone marks are relatively well studied, the majority arguably remain poorly characterized. Screening tools like ChromID will surely prove immensely beneficial in dissecting the roles of histone PTMs, and perhaps even combinations thereof.

Once the interacting proteins have been identified, a battery of biophysical tools are used to validate the interaction and characterize its thermodynamic properties. This establishes binding affinities and can provide additional information, such as protein stoichiometry within a complex. Common techniques include isothermal titration calorimetry (ITC), dynamic light scattering, analytical ultracentrifugation (AUC), bio-layer interferometry (BLI), and microscale thermophoresis (MTS), to name a few (see Ausio, 2000; Lewis and Murphy, 2005; Concepcion et al., 2009; Stetefeld et al., 2016; Asmari et al., 2018).

It is important to emphasize that histone PTMs are found in thousands of combinations throughout the genome. To add further complexity, these marks are constantly added, removed, and bound by other proteins. Antibody-based techniques are instrumental to their study, but an increasing number of innovative technologies also continue to facilitate research through such plasticity.

HISTONE PROTEIN-PROTEIN INTERACTIONS: FROM PROTEIN TRANSLATION TO NUCLEAR IMPORT

Guided by biochemical and cellular approaches, two different models (which may not necessarily be mutually exclusive) have been proposed to explain the pathway by which newly synthesized histones are folded, processed, and imported into the nucleus [discussed in Grover et al. (2018) and Pardal et al. (2019)]. In both models, new histones are folded by molecular chaperones and imported into the nucleus while bound by importin-4. The two models mainly differ in the order of events, and their subcellular localization.

In the first model, affinity purification of epitope tagged cytoplasmic H3 followed by column chromatography resolved a number of core histone protein complexes that hint toward an organized processing and nuclear import of newly synthesized histones (Campos et al., 2010; Alvarez et al., 2011). Newly translated H3 and H4 monomeric units are first folded by molecular chaperones (complex Ia: H3-HSC70; complex Ib: H4-HSP90-HSC70) before their assembly into H3-H4 dimers by HSP90 and tNASP (complex II). Complex III involves processing by sNASP and the HAT1 holoenzyme. NASP proteins help fold the H3-H4 histones (Bowman et al., 2017), facilitate H4 acetylation by HAT1 (Campos et al., 2010), and regulate soluble H3-H4 levels (Cook et al., 2011). Once processed, the H3-H4 dimers are transferred to the ASF1 histone chaperone for import into the nucleus in complex IV (H3-H4-ASF1-importin 4). ASF1 then transfers the histones to other histone chaperones that deposit them onto DNA.

There are a number of important intricacies regarding protein isoforms and the sequential addition of histone PTMs (Alvarez et al., 2011). Some were captured by alternative means since protein chromatography resolves biochemically stable protein complexes. For example, the isolation and proteomic analysis of polysomes containing histone polypeptides undergoing translation allowed for the identification of

SETDB1 (KMT1E/ESET) as the enzyme responsible for the monomethylation of lysine 9 on a subset of new H3 proteins (Rivera et al., 2015). This nuance is important since the mark is believed to prime replicating chromatin for heterochromatin formation (Loyola et al., 2006).

The second histone nuclear import model is based on cellular technique known as rapamycin-activated protease through induced dimerization and release of tethered cargo (RAPID-release) (Apta-Smith et al., 2018). In this technique, new histones are tethered to the cytosolic side of the outer mitochondrial membrane, released through rapamycin-activated cleavage of the tethering moiety, and tracked using a fluorescent tag. While the effect of tethering histones is unclear, the study clearly illustrates flexibility in the histone processing pathway. The experimental pipeline shows that H3 and H4 monomers can be imported into the nucleus while directly associated with importin-4. H3-H4 dimer assembly and histone chaperoning then occurs in the nucleus. The two models are not necessarily mutually exclusive, but the reports emphasize the need to further explore the pre-deposition histone pathways and highlights the benefits of considering multiple experimental approaches.

HISTONE DEPOSITION ON CHROMATIN: VIVE LA DIVERSITÉ!

Two of the three main H3-H4 histone deposition pathways were elucidated long before the pre-deposition processing pathways above. Biochemical approaches and *in vitro* systems were critical toward the identification of variant-specific histone chaperones that deposit histones on DNA. ASF1-bound H3-H4 histones can be transferred to the CAF-1, HIRA, and DAXX histone chaperones for deposition on DNA.

Chromatography-based fractionation of HEK293 nuclear extracts enabled the isolation of histone deposition activity that occurred on replicating DNA. Histone deposition was tested from fractions, using an *in vitro* replication system in the presence of soluble histones. This led to the identification of the replication-coupled Chromatin Assembly Factor 1 (CAF-1) (Smith and Stillman, 1989). CAF-1 is a three-subunit H3-H4 histone chaperone. It is coupled to DNA replication because of its interaction with the proliferating cell nuclear antigen (PCNA) processivity ring, via a PCNA interacting peptide (PIP) motif (Rolef Ben-Shahar et al., 2009). ASF1-bound H3-H4 dimers are first transferred to CAF-1 through direct interactions between ASF1 and the p60 subunit of CAF-1 (Tyler et al., 2001; Mello et al., 2002). Histone deposition is then a matter of thermodynamics (Das et al., 2010; Liu et al., 2012; Sauer et al., 2018).

Gel filtration, relative amino acid composition, and AUC analyses showed that a single ASF1 binds an H3-H4 dimer, occluding histone tetramerization prior to deposition (English et al., 2005), something that was further confirmed by X-ray crystallography (English et al., 2006; Natsume et al., 2007). The CAF-1 winged helix domain binds DNA and promotes the tetramerization of two H3-H4 dimers while forming nucleosomes by depositing the (H3-H4)₂ tetramer on DNA (Liu et al., 2012; Mattioli et al., 2017; Sauer et al., 2018). This stepwise

transfer of H3-H4 from ASF1 to CAF-1 and, ultimately, DNA is explained by a “nucleosome assembly funnel” (Das et al., 2010). In this model, free histones have high free energy and are handed from chaperone to chaperone to be assembled into stable intermediates until ultimately being transferred to DNA, the state with the lowest free energy. The dissociation constant of histones bound to either ASF1 or CAF-1 is in the low nanomolar range (Donham et al., 2011; Liu et al., 2012). However, fluorescent anisotropy and electrophoretic mobility shift assays (EMSAs) determined that yeast Asf1 and CAF-1 preferentially associate when Asf1 is pre-bound to H3-H4 (Liu et al., 2012). Förster resonance energy transfer (FRET) experiments, in which equimolar amounts of donor- and acceptor-labeled H4 were mixed to measure tetramer formation, resulted in nearly identical fluorescent emission spectra for (H3-H4)₂ bound to CAF-1 or DNA, suggesting that CAF-1 primes H3-H4 for deposition on DNA (Liu et al., 2012; Sauer et al., 2018). The resulting tetrasome folds some 80 bp of DNA until H2A-H2B dimers complete the nucleosome (Brower-Toland et al., 2002; Sauer et al., 2018).

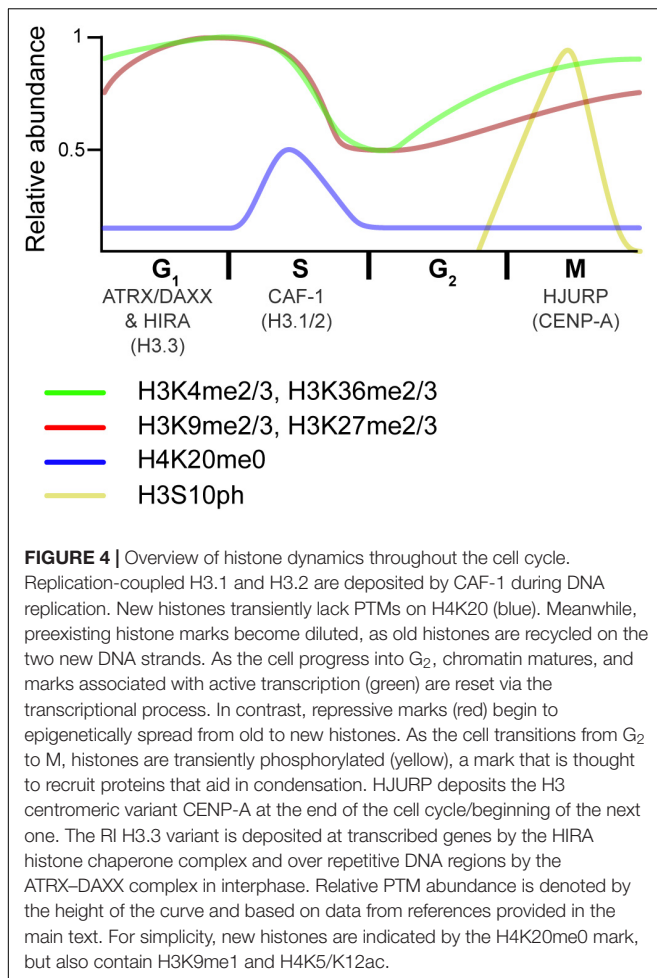
Unlike CAF-1, the HIRA histone chaperone easily deposits histones on static DNA templates *in vitro* (Ray-Gallet et al., 2002). Affinity purification of epitope-tagged H3.1 and H3.3 from mammalian cells found the CAF-1 and HIRA protein complexes to act as their respective histone chaperones (Tagami et al., 2004). In mammals, it is the ubinuclein-1/2 subunit that confers specificity toward H3.3, as shown via biophysical and structural analyses (Ricketts et al., 2015). HIRA notably operates over transcribed genes, where nucleosomal histones are disrupted (Goldberg et al., 2010; Sarai et al., 2013).

Biochemical fractionation of affinity purified H3.3-associated proteins also led to the identification of yet another H3.3 histone chaperone, DAXX, and its binding partner ATRX (Drane et al., 2010; Lewis et al., 2010). Protein crystallization found the H3.3-specific G90 residue to be key for the interaction with DAXX (Elsasser et al., 2012). Unlike HIRA, DAXX-ATRX mainly deposits H3.3 over repressed, repetitive DNA elements, where nucleosomes are important for stability.

Biochemical fractionations also helped identify the histone chaperone for the centromeric H3 variant, CENP-A: HJURP; while IF-based studies showed that CENP-A deposition occurs in late M/early G₁ (Jansen et al., 2007; Dunleavy et al., 2009). As such, CAF-1, HIRA, DAXX, and HJURP deposit different H3 variants, at specific times, and over different genomic regions (Figure 4). While these deposition pathways are well established, new ones are seeing light (see section “Histone Recycling and Tools to Study Interactions at Replication Forks”).

TECHNIQUES TO STUDY HISTONE OCCUPANCY AND DEPOSITION

To test for histone chaperone activity, proteins must be shown to specifically bind histones, promote histone deposition on DNA without the use of ATP, and dissociate from the final product (*i.e.*, nucleosome) (Gurard-Levin et al., 2014; Hammond et al., 2017; Grover et al., 2018; Ricketts et al., 2019). *In vitro* and *in vivo* techniques are available for this purpose.



In vitro methods used to assess histone deposition on naked DNA are well established (Germond et al., 1975; Noll et al., 1975; Loyola et al., 2004). These minimally contain free histones, a histone chaperone, and a DNA template, which are incubated for a fixed period of time near physiological salt concentrations. If using a short piece of linear DNA (i.e., accommodates a single nucleosome), the assembly can be observed by techniques as simple as EMSA (Laskey et al., 1978). If using longer DNA templates (e.g., a plasmid), other techniques are better suited. Micrococcal nuclease (MNase) preferentially cleaves internucleosomal DNA, and a limited digest releases mononucleosomes and multimers thereof (Noll, 1974). Deproteinized DNA can then be separated by gel electrophoresis for assessment. MNase digests are particularly informative on the quality of nucleosome assembly and nucleosome spacing (Lusser and Kadonaga, 2004), but other techniques better assess the proportion of nucleosome assembly. The DNA supercoiling assay, in turn, measures topological changes due to the formation of nucleosomes on a closed circular DNA template. This assay directly quantifies the extent of nucleosome assembly because each nucleosome adds a superhelical turn (Germond et al., 1975). The assay demonstrated the histone chaperone activity of the

first histone chaperone to be isolated from extracts (Laskey et al., 1978), and remains a gold standard in the field.

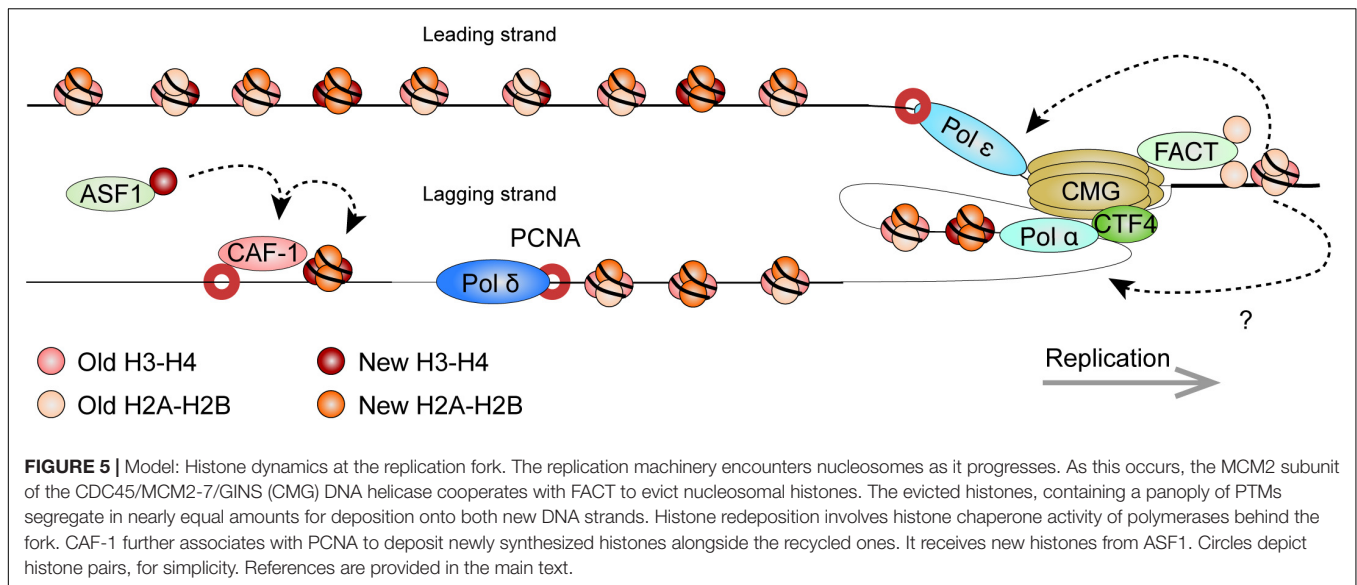
It is also possible to assess histone occupancy *in vivo*, notably by ChIP. At a more global level, MNase digestion can be performed on genomic chromatin from intact nuclei. It is even possible to probe specific genomic regions to assess their relative accessibility (Wu et al., 1979). A newer technique is, however, more commonly used to assess chromatin accessibility at the genome-wide level. The assay for transposase-accessible chromatin coupled to sequencing (ATAC-seq), is a method that involves the fragmentation and tagging (tagmentation) of the genome with sequencing adaptors using the Tn5 transposase. Sequencing reads then reveal genomic regions that are highly represented and, thus, within more accessible chromatin (Buenrostro et al., 2013).

There are, of course, several other specialized techniques to quantify and qualify nucleosome assembly, such as electron microscopy (EM) and single-molecule techniques (see Duzdevich and Greene, 2013; Schwartzman and Tanay, 2015; Senapati et al., 2015).

HISTONE RECYCLING AND TOOLS TO STUDY INTERACTIONS AT REPLICATION FORKS

While the deposition of new histones via CAF-1 is well-established, exciting findings are beginning to shed light on the eviction, segregation, and redeposition of pre-existing nucleosomal histones [see Grover et al. (2018); Sauer et al. (2018), and Stewart-Morgan et al. (2020) for detailed updates on the topic]. When DNA is copied, preexisting nucleosomal histones dissociate from the replicating DNA strand (eviction) to redistribute on both nascent DNA strands (segregation) and form new nucleosomes (redeposition) (Figure 5). This histone recycling process is extremely complex and requires precise steps to ensure that epigenetic information is maintained while faithfully replicating DNA. Studying histone recycling at replication forks is challenging because of the dynamic nature of the process. There is also a need to uncouple the deposition of recycled, pre-existing nucleosomal histones from that of their newly synthesized counterparts. While biochemical approaches, such as those discussed above, continue to better our understanding of histone recycling, so are new emerging technologies.

The eukaryotic replisome is spearheaded by the CDC45/MCM2-7/GINS (CMG) DNA helicase that unwinds the DNA double helix (Gambus et al., 2006; Moyer et al., 2006). As the helicase tracks along chromatin, it also makes contact with nucleosomes ahead of the replication machinery and plays an important role in histone eviction. The minichromosome maintenance 2 (MCM2) subunit harbors a conserved N-terminal region that imparts histone chaperone activity (Ishimi et al., 2001; Foltman et al., 2013; Huang et al., 2015b). The eviction process requires the facilitates chromatin transcription (FACT) histone chaperone. Pull-down experiments in yeast showed that FACT and Mcm2 cooperatively bind histones (Foltman et al., 2013).



Strains with mutations affecting Mcm2 histone-binding residues suffered from a loss of heterochromatin, highlighting its importance in histone recycling. Crystallography and single molecule assays using optical tweezers show that the SUPT16H subunit of FACT (Spt16 in yeast) can displace and tether nucleosomal H2A-H2B dimers while stabilizing the (H3-H4)₂ tetramer (Chen P. et al., 2018; Wang et al., 2018; Liu et al., 2020). Further structural work showed that MCM2 in turn associates with (H3-H4)₂ surfaces that normally interact with DNA, thereby shielding the tetramer from aberrant interactions (Huang et al., 2015b). Altogether, this suggests a synergistic role for MCM2 and FACT in the disassembly of nucleosomes on replicating DNA.

By labeling newly replicated DNA, and separating replicated and non-replicated chromatin on density gradients, early studies showed that pre-existing nucleosomal histones segregate to both leading and lagging strands (Jackson et al., 1975). Elegant EM visualization of replicating minichromosomes further found that nucleosomes rapidly reform on nascent chromatin near the replication fork (Sogo et al., 1986). More recently, SILAC, coupled to a controlled pulse-chasing of pre-existing or new histones, further demonstrated that histones are predominantly recycled as (H3-H4)₂ tetrameric and H2A-H2B dimeric units (Xu et al., 2010). Curiously, H3.3-containing tetramers were more apt to dissociate into dimers thereby allowing intermixing with new histones. Meanwhile, H2A-H2B dimers readily re-associated with new and old (H3-H4)₂ tetramers.

There is now exciting new data regarding the mechanisms by which nucleosomal histones are evicted, segregate, and reassemble on the newly replicated DNA. Sister chromatids after replication by DNA sequencing (SCAR-seq) in mouse ESCs (mESCs) demonstrated parental histones segregated with a slight preference for the leading strand (Petryk et al., 2018). In the technique, replicating DNA is labeled with a thymidine analog (e.g., 5-ethynyl-2'-deoxyuridine, EdU), while DNA fragments containing either new or old histones are immunoprecipitated

(via H4K5ac and H4K20me2 marks, respectively). Nascent DNA is subsequently captured via the EdU thymidine analog (see below) and subject to alkaline denaturation to isolate and sequence the newly synthesized strand.

New and old (recycled) histone deposition was also followed in yeast strains expressing Mcm2 mutants that do not bind histones (Gan et al., 2018). Enrichment and sequencing of protein-associated nascent DNA (eSPAN), showed enrichment of old histones on the leading strand as a result of the deficiency. Therefore, Mcm2 is critical for proper histone recycling. In eSPAN, another thymidine analog (5-bromo-2'-deoxyuridine, BrdU) is used to label replicating DNA. This is followed by MNase digestion and ChIP of H3K4me3 (enriched on parental nucleosomes) and of H3K56ac (enriched on new histones in yeast), and followed by strand-specific sequencing (ChIP-ssSeq), as in SCAR-seq. Nascent DNA is mapped back in relation to origins of replication to determine strand identity. Coupled to yeast strains defective for other replication proteins, eSPAN further identified Ctf4 (which links the CMG complex to the lagging strand) and Pol α (which initiates lagging strand synthesis) as additional components that are required for histone recycling on the lagging strand (Gan et al., 2018). Curiously, biophysical analyses indicate that Pol α preferentially binds H2A-H2B (Evrin et al., 2018), highlighting the need for further investigation on mechanistic details of the (H3-H4)₂ and H2A-H2B recycling at replication forks. Like for the Mcm2 mutant yeast strains, the impairment of this pathway also disrupted gene silencing in yeast (Evrin et al., 2018).

Similarly, recent data implicate the leading strand polymerase ε subunits POLE3-POLE4 in nucleosome assembly. Gel filtration chromatography (which resolves protein complexes based on size) demonstrated H3-H4 binding by human POLE3-POLE4 *in vitro*, while immunoprecipitations confirmed the *in vivo* interaction (Bellelli et al., 2018). Supercoiling assays also demonstrated their ability to promote tetrasome formation *in vitro*, showing *bona fide* histone chaperone activity. This

was confirmed *in vivo*, through pulse-chase experiments using fluorescently labeled, SNAP-tagged H3. The SNAP-tag reacts with benzylguanine derivatives and is an efficient way to label and follow proteins. RNAi depletion of POLE3 or POLE4 resulted in a reduced deposition of the tagged histone (Bellelli et al., 2018). Whether additional factors other than MCM2, Pol α , and Pol ϵ further participate in histone recycling at replication forks remains to be seen.

MINIMAL HISTONE SHUFFLING AT THE FORK

Data suggest that recycled histones remain coupled to the replication machinery, and a recently developed *in vitro* system nicely demonstrates this. Biotinylated histones were assembled on a nucleosome positioning sequence at a specific location on a circular DNA template. Their position was then mapped before and after DNA replication (Madamba et al., 2017). Nucleosome positioning was relatively well-preserved when the reaction was driven by eukaryotic (*Xenopus* egg extracts) but not viral (SV40 T-antigen) replication machineries. Techniques have also been developed to track the accuracy of histone redeposition in replicating cells.

Chromatin occupancy after DNA replication by next-generation sequencing [ChOR-seq (Reveron-Gomez et al., 2018)] maps histone distribution on replicated DNA. Akin to iPOND/NCC (described below), cells are briefly pulsed with a recoverable thymidine analog (e.g., EdU) that is incorporated into replicating DNA. ChIP is then performed to isolate chromatin fragments with a specific histone PTM. This allows for the recognition of histone PTMs that are enriched on old, pre-existing histones at either repressed or transcribed regions of the genome. Replicated DNA fragments are further selected and sequenced. ChOR-seq data suggest that old, recycled histones re-incorporate on newly synthesized DNA with surprising fidelity—within some 250 bp of their pre-replication position. Such a tight coupling between recycled histones and the replication machinery is in line with early EM micrographs and biochemical analyses of replicating minichromosomes, showing nucleosomes reassembling some 225–285 bp behind the fork (Herman et al., 1981; Sogo et al., 1986), though the latter could not differentiate between new and recycled histones.

To further follow specific histones through cell division, a biotinylation system akin to that of the *Xenopus* egg extract system (above) was also established in an ESC model (Escobar et al., 2019). Endogenous replication-coupled histone variants were fused to a biotin acceptor peptide (BAP) after which dCas9-BirA was transiently recruited to a specific genomic locus. BirA is a bacterial biotin ligase that recognizes and biotinylates the BAP. A controlled, local biotinylation of histones was therefore achieved. Expression of the dCas9-BirA fusion was tightly regulated and restricted to late G₁. The biotinylated histones were then followed by ChIP-seq following DNA replication. As per the other studies, the recycled histones remained near their original position after DNA replication. Interestingly, while this was especially true for repressed chromatin regions, histones

found on transcribed regions were more apt to disperse in this system. A similar system was also established in yeast, but with BirA fused to the tetracycline repressor TetR (Schlissel and Rine, 2019). Local histone biotinylation was thereby driven by BirA recruitment to an intergenic, single-copy tetracycline operator. ChIP-seq again demonstrated faithful nucleosome redeposition following DNA replication. The biotin ChIP-seq peak remained after rounds of replication, only diminishing in intensity because of the dilution of old biotinylated histones with new non-biotinylated ones. This positional memory was disrupted by the mutation of Mcm2 or depletion of a Pol ϵ subunit, further highlighting the importance of histone chaperone pathways.

POST-REPLICATIVE RE-ESTABLISHMENT OF HISTONE MARKS AND CHROMATIN MATURATION

Local chromatin structures need to be re-established following the co-deposition of new and recycled histones on new DNA strands (Figure 4). This maturation process requires, in part, the spreading of at least certain repressive histone marks (e.g., H3K9me2/3 and H3K27me2/3), as well as transcriptional restart (Reinberg and Vales, 2018; Stewart-Morgan et al., 2019).

To further unravel the molecular details by which post-replicative chromatin matures, a number of exciting new proteomic techniques were used. The isolation of proteins on nascent DNA (iPOND) allows researchers to probe the proteome of replicating DNA and post-replicative maturing chromatin (Sirbu et al., 2012). In iPOND, cells are pulsed with EdU, which is incorporated into replicating DNA. When followed by a thymidine chase, it is possible to distinguish newly replicated DNA from maturing chromatin. Proteins that associate with replicating or maturing DNA are then isolated from sheared DNA fragments through the EdU label. EdU contains an alkyne group that is covalently linked to azide coupled moiety (i.e., biotin) *in vitro*, via a copper-catalyzed cycloaddition reaction [“click chemistry” (Gierlich et al., 2006)]. Associated proteins are finally analyzed by western blotting or MS. Different versions of the technique exist; nascent chromatin capture (NCC) coupled to SILAC uses a similar protocol, with a direct comparison of proteins that are associated on nascent versus maturing chromatin (Alabert et al., 2014). As expected, proteins, such as DNA polymerases and the CAF-1 histone chaperone, enriched on nascent chromatin. This elegant technique also led to some intriguing observations. For example, various histone “writers” differently enriched on nascent and mature chromatin. This offers mechanistic insights on the post-replicative re-establishment of histone PTMs.

The analysis of histone PTMs using the NCC-SILAC pipeline showed that there are different propagation modes for different histone PTMs (Alabert et al., 2015). As progression through S-phase caused a twofold dilution of marks, some marks were quickly re-established through G₂ (e.g., H3K4me3), whereas others took the remainder of the cell cycle or longer (e.g., H3K9me3 and H3K27me3). A progressive, coordinated

restoration of histone PTMs was also seen when comparing bulk histone PTM levels by SILAC-MS as synchronized cells progressed through the cell cycle (Zee et al., 2012). The analysis showed that methyl states are not equally reestablished. For example, H3K9me2 resulted from H3K9me1 acquisition of a second methyl group in late-G₁/S as well as from the acquisition of 2 methyl groups on newly synthesized H3 in G₂/M. H3K9me3 was in turn established through the addition of a third methyl group to pre-existing H3K9me2 in G₁/S, and from newly synthesized H3 acquiring 3 methyl groups in G₂/M. In contrast, H3K27me2 was largely the result of unmodified residues acquiring 2 methyl groups in G₂/M, while H3K27me3 re-establishment patterns were similar to H3K9me3, but clearly antagonized by the H3K36me3 mark (Zee et al., 2012; Alabert et al., 2020).

In contrast to repressive H3K9me3 and H3K27me3 marks that rely on spreading (Lachner et al., 2001; Margueron et al., 2009), the restoration of the transcription-associated H3K4me3 PTM occurred shortly after DNA replication, with a faster restoration over genomic regions with high transcriptional levels (Reveron-Gomez et al., 2018). Transcription-dependent reestablishment of post-replicative chromatin was further supported by repli-ATAC-seq (Stewart-Morgan et al., 2019), where EdU-labeled replicating DNA is subject to tagmentation, but click chemistry is used to isolate newly replicated fragments. Importantly, the technique showed that, in mESCs, post-replicative chromatin is largely inaccessible until transcription resumes.

MITOTIC BOOKMARKING: A BRIEF PRIMER

Following DNA replication and chromatin maturation, chromatin is condensed in preparation for mitotic segregation of sister chromatids. These changes are accompanied by distinct chromatin alterations that are still under investigation. Mitotic entry is notably characterized by high levels of phosphorylated histones H1 and H3 (Lake et al., 1972; Gurley et al., 1974). The latter is, however, believed to be particularly critical for mitosis (Ohsumi et al., 1993). H3 is specifically phosphorylated at serine 10 (H3S10ph) (Paulson and Taylor, 1982), by the Aurora B (Goto et al., 2002) and VRK1 kinases in mammals (Kang et al., 2007). Phosphorylation begins at the centromere in G₂, spreads throughout the genome during G₂/M, and is mostly lost as cells enter telophase (Hendzel et al., 1997). H3 is also phosphorylated at S28 by Aurora B (Goto et al., 2002) beginning in prophase and until anaphase (Goto et al., 1999). The exact roles of these modifications require further study, but they do influence PPIs, and likely assist in a number of mitotic events. For example, HP1 occupancy on chromatin diminishes during mitosis and data indicate H3S10ph prevents HP1 binding (Fischle et al., 2005; Hirota et al., 2005). The condensin I and II complexes, which have been implicated in proper chromatin compaction and mitotic progression, are also thought to be recruited by histone H3 PTMs (Hirano and Mitchison, 1994; Giet and Glover, 2001; Takemoto et al., 2007).

Although it lacks sequence similarity to H3 at its N-terminus, the centromeric H3 variant CENP-A, shares a small stretch of sequence similarity to the H3S10 region and is also phosphorylated after the onset of H3S10ph and through anaphase (Zeitlin et al., 2001). CENP-A is important for mitotic division, and is deposited by the histone chaperone HJURP in late telophase/early G₁ (Jansen et al., 2007; Dunleavy et al., 2009). The exact roles of its phosphorylation remain elusive, however, unphosphorylated CENP-A was shown to cause improper microtubule attachment at the kinetochore (Kunitoku et al., 2003). Readers are referred to recent CENP-A reviews for further information on this histone variant (Catania and Allshire, 2014; Muller and Almouzni, 2017).

Data suggest that mitotic “bookmarking” via histone PTMs and chromatin-bound proteins maintains epigenetic information on chromatin as it undergoes profound changes (Kouskouti and Talianidis, 2005; Valls et al., 2005; Young et al., 2007). For example, ChIP-seq in mitotically arrested mESCs demonstrated that H3K27ac was retained at housekeeping gene promoters and stem-cell associated enhancers during mitosis, suggesting that the mark primes transcriptional activation in G₀/G₁. Recent data demonstrated H3K4me3 remained associated with most promoters, while H3K27ac was maintained at only a subset of enhancers and promoters, but quickly reestablished at the anaphase/telophase transition (Kang et al., 2020). This was determined using a combinatorial approach with ChIP-seq and EU-RNA-seq, a technique where cells are treated with ethynyl uridine to label newly synthesized RNA, which can then be isolated and sequenced to generate a transcriptome of newly synthesized transcripts. A number of transcription factors, including stem cell regulators, such as SOX2 and OCT4, also remain bound to chromatin in mitosis, which likely contributes to the restoration of transcriptional programs upon mitotic exit (Chen D. et al., 2002; Deluz et al., 2016; Liu et al., 2017). Readers are referred to the following reviews for more information on mitotic bookmarking (Festuccia et al., 2017; Palozola et al., 2019).

HISTONE DYNAMICS DURING GENE TRANSCRIPTION

Early studies showed that nucleosomes hinder transcriptional initiation when using *in vitro* transcription systems (Knezetic and Luse, 1986; Lorch et al., 1987). The influence of histones, and their PTMs, toward gene transcription *in vivo* is now amply evident. Experimental models allowing transient histone depletion or the loss of the PTM-rich N-terminal histone tails affect gene expression (Han and Grunstein, 1988; Mann and Grunstein, 1992; Lenfant et al., 1996; Bintu et al., 2012). Over the years, a large number of histone PTMs have been correlated, or anti-correlated, with gene transcription (Barski et al., 2007; Wang et al., 2009). As discussed, H3K4me3 is particularly enriched over the promoter and transcriptional start site (TSS) of active genes, whereas transcribed gene bodies are enriched for H3K36me3 (Bernstein et al., 2005; Talbert and Henikoff, 2006; Wang et al., 2008, 2009; Figure 6).

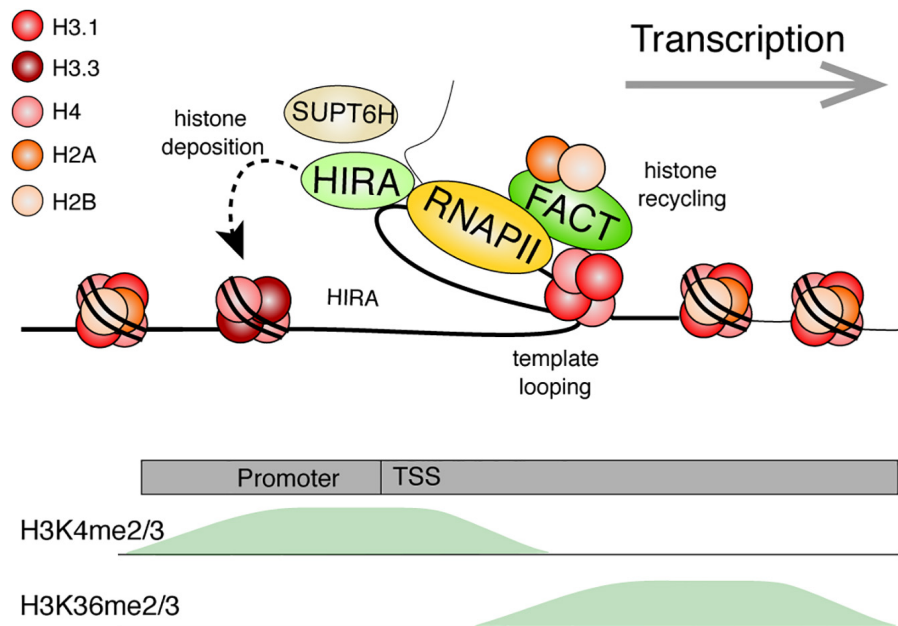


FIGURE 6 | Model: Histone dynamics during transcription. The RNA Polymerase II machinery maneuvers through nucleosomal DNA as it transcribes genes. Histone PTMs are believed to facilitate numerous processes (e.g., splicing or preventing cryptic transcription). Transcribed genes are typically enriched for H3K4me2/3 near the transcriptional start site (TSS) and with H3K36me2/3 within the gene body. The histone chaperone, FACT, facilitates transcription by disrupting histone-DNA contacts and helps preserve nucleosomes by tethering the H2A-H2B dimer while stabilizing the (H3-H4)₂ tetramer. Template looping is hypothesized to aid in transfer of the tetramer behind the transcriptional machinery. FACT and other histone chaperones including HIRA and SUPT6H further ensure that proper nucleosome density is maintained behind the polymerase through deposition of new histones and aiding in nucleosome reassembly, respectively. References are provided in the main text.

Chromatin immunoprecipitation nicely captures a nucleosome-free region near the TSS of transcribed genes, and nucleosome phasing at flanking positions (Yuan et al., 2005; Schones et al., 2008). At the promoter, nucleosomes hinder protein binding and chromatin remodeling is often required for efficient transcriptional initiation (Klemm et al., 2019; Brahma and Henikoff, 2020). Once initiated, the elongating transcriptional machinery must then navigate through chromatinized DNA (Kujirai and Kurumizaka, 2020). Nucleosomal histones are disrupted in the process, and mechanisms exist to facilitate transcription while preserving the chromatin environment.

An early *in vitro* study, in which a nucleosome was assembled onto a plasmid, showed that the histones were displaced by the viral SP6 RNA polymerase to reassemble at a different location on the plasmid (Clark and Felsenfeld, 1992). It was suggested that such nucleosome displacement likely involved a direct transfer mechanism without complete histone dissociation from DNA, since competitor DNA had little quenching effect on histones under specific conditions (Studitsky et al., 1994). In addition to histone exchange in the absence of transcription, ChIP, fluorescence recovery after photobleaching (FRAP), and MNase-seq (which, like ATAC-seq, assesses chromatin accessibility) experiments show that H2A-H2B dimer loss and exchange—and to a lesser degree (H3-H4)₂ exchanges—correlate with active transcription (Kimura and Cook, 2001; Jamai et al., 2007; Cole et al., 2014).

Recent cryo-EM studies revealed that RNA Polymerase II pauses at specific sites within the nucleosome while preserving the histone octamer (Kujirai et al., 2018). Single-molecule force spectroscopy techniques (Gosse et al., 2019), such as optical and magnetic tweezers, enabled controlled biophysical analyses on polymerases and nucleosomes. Using a dual-trap optical tweezer experiment to follow single RNA polymerases on a DNA template containing a single nucleosome, RNA Polymerase II was shown to pause, especially before reaching the dyad axis on the nucleosome (Hodges et al., 2009). DNA looping then facilitated the transfer of histones behind the polymerase.

However, high gene activity appears to be particularly disruptive for nucleosomes (Kulaeva et al., 2010). An *in vitro* system demonstrated this by stalling one or two elongating *Escherichia coli* RNA polymerases ahead of a nucleosome. Stalling was achieved by depleting UTP or UTP and CTP on DNA templates with C and U tracks upstream of the nucleosome. Addition of all nucleotides then allowed transcriptional elongation to occur. The passage of the first RNA polymerase tended to evict an H2A-H2B dimer from some nucleosomes, leaving behind a histone hexamer; whereas the second polymerase could displace the remaining histones (Kulaeva et al., 2010). Nucleosomal density and histone PTMs are, however, maintained over the transcribed gene to facilitate various co-occurring events, and prevent aberrant transcription (Smolle and Workman, 2013). Studies in *Drosophila* indicated

that the H3.3 RI histone variant enriches over actively transcribed genes (Ahmad and Henikoff, 2002; Schwartz and Ahmad, 2005), an observation that was then extended to mammals (Ray-Gallet et al., 2011). As mentioned above, nucleosomal density is indeed maintained through transcription-coupled H3.3 deposition, via the HIRA histone chaperone (Goldberg et al., 2010; Sarai et al., 2013).

Additional histone chaperones facilitate transcription while promoting histone recycling at transcribed genes. The histone chaperone FACT is one of the better studied components involved in the disassembly and re-assembly of nucleosomes during transcription. The FACT heterodimer, composed of the SUPT16H and SSRP1 subunits in humans, was first purified based on its ability to facilitate transcription of chromatinized DNA templates *in vitro* (Orphanides et al., 1998). It is, however, important to emphasize that the histone chaperone has since been implicated in numerous other biological events (Gurova et al., 2018). Although first described as an H2A-H2B histone chaperone, it can also bind histones H3 and H4 (Martin et al., 2018; Wang et al., 2018; Mayanagi et al., 2019). FACT allows a controlled assembly and disassembly of nucleosomes *in vitro*. Specifically, recent AUC, optical tweezer, and structural data demonstrate that SUPT16H binds nucleosomal DNA, stabilizes the central (H3-H4)₂ tetramer, and tethers an H2A-H2B dimer at its DNA binding surface. Meanwhile, SSRP1 facilitates the redeposition of the tethered H2A-H2B dimer, while further maintaining (H3-H4)₂ on DNA (Chen P. et al., 2018; Wang et al., 2018; Liu et al., 2020). The two subunits synergize to allow for proper disassembly, stabilization, preservation, and reassembly of the nucleosome. Using magnetic tweezers to immobilize and stretch a DNA template containing a single nucleosome, FACT was beautifully shown to promote an orderly histone eviction and reassembly on DNA (Chen P. et al., 2018). Nucleosomes disassembled at lower forces in the presence of FACT. Moreover, the nucleosome did not properly reassemble upon repeated DNA stretching experiments unless FACT was present in the reaction.

Curiously, FACT is not equally expressed across tissues (Garcia et al., 2011), and its loss results in the misregulation of only a subset of transcribed genes (Li et al., 2007). This may, perhaps, be explained by functional redundancy with other histone chaperones. By fractionating nuclear extracts based on the presence of FACT-like activities using an *in vitro* transcription system on chromatinized templates, a recent study identified LEDGF and HDGF2 as novel transcription-coupled histone chaperones (LeRoy et al., 2019). Behind the polymerase, the yeast Spt6 H3-H4 chaperone (SUPT6H in humans), also assists with nucleosome reassembly at highly transcribed genes (Bortvin and Winston, 1996; Ivanovska et al., 2011). ChIP-chip studies in yeast demonstrated that loss of FACT or Spt6 resulted in a transcription-dependent shuffling of evicted histones (Jeronimo et al., 2019). Using the same local histone biotinylation system as with DNA replication, gene induction also led to a gradual loss of histones over the transcribed gene (Schlissel and Rine, 2019). While the ChIP-seq biotin peak diminished in intensity during transcription, its position was largely maintained. Together, this

further highlights the important role of histone chaperones in chromatin maintenance.

LOCUS-SPECIFIC CHROMATIN ANALYSIS

Rapid advancements in proteomic screening tools now allow us to identify proteins that are bound at specific locations in the genome (**Figure 3**). ChIP-based experiments have proven instrumental to our understanding of the epigenome, and have also evolved to study PPIs on chromatin. ChIP coupled to western blotting or even MS (ChIP-MS) can identify proteins that are enriched on a chromatin fragment containing an epitope of interest (e.g., a histone mark) (Ji et al., 2015). A more stringent variation of the approach, ChIP-SICAP (Rafiee et al., 2016) identifies chromatin-bound proteins through an initial ChIP, followed by DNA biotinylation and washing, which releases proteins that are not directly bound to chromatin. Though readily applicable, ChIP-based approaches are limited by the quality of the antibodies that are used, and require sufficient material when adapted toward proteomic analyses. The aforementioned ChromID technique (Villasenor et al., 2020) addresses these concerns but will require the design of additional histone-binding modules in order to be tailored for a broader range of experiments investigating protein-protein associations with diverse histone marks.

Equally exciting techniques now allow for locus-specific analysis of chromatin dynamics. To probe sequence-specific chromatin-associated proteins, proteomics of isolated chromatin segments (PICH) was developed (Déjardin and Kingston, 2009). Using locked nucleic acids to hybridize to and isolate sequence specific chromatin fragments and subsequent MS analysis, the authors compared telomeric chromatin composition in cells that use different telomere maintenance mechanisms. The technique has since been adapted to study rDNA (Ide and Déjardin, 2015). While this tour de force proved insightful, it requires a large amount of input material. Protein abundance and the signal-to-noise ratio may also confound results (Gauchier et al., 2020).

Proximity-dependent labeling techniques are yet another promising tool that does not require large amounts of input material, as tags amplify the signal of associated proteins. When fused to a protein of interest (bait), an engineered ascorbate peroxidase (APEX) enzyme (Martell et al., 2012), or promiscuous biotin ligase (e.g., BirA*; a mutant form of the BirA enzyme), labels proximal proteins (preys). Biotinylated proteins are then captured on streptavidin beads and identified by mass spectrometry—a technique known as proximity-dependent biotin identification, or BioID (Roux et al., 2012).

The revolutionary take on prior tagging methodologies (Chen and Ting, 2005) proved immensely powerful. Long labeling times allow the tagging of thousands of interactors per cell compared to AP-MS, which can only capture a snapshot of interactors at a specific time. Therefore, these techniques allow for the amplification and identification of less abundant interactions that traditional pulldowns may miss using comparable input

material. While APEX offers rapid labeling, it requires the addition of hydrogen peroxide, which is toxic to cells at high concentrations (Rhee et al., 2013). Biotin ligases such as BirA* also provide greater signal compared to APEX-based techniques because they target abundant lysine residues, while APEX biotinylates electron-rich residues of lower abundance (e.g., tyrosine). Nevertheless, BioID, and clever derivatives using APEX or biotin ligases, are increasingly utilized to capture stable and biochemically labile interactions, including that of histone proteins (Lambert et al., 2015; Zasadzinska et al., 2018). Interestingly, there is an ongoing effort to use BioID to map the cell where 192 baits from 32 cellular compartments identified over 35,000 unique proximal associations (Go et al., 2019). While insightful, it is also important to recognize that the fusion of two proteins can skew results. Prey proteins also represent proximal and not necessarily direct physical interactions, and BioID data analysis is further influenced by the negative controls to which the data are compared.

Nevertheless, ingenious BioID variations (such as ChromID), continue to be designed [readers are referred to other recent reviews on proximity-based labeling (Kim and Roux, 2016; Samavarchi-Tehrani et al., 2020)]. Of particular interest are those that fuse APEX (CASPEX, C-BERST, and CAPLOCUS) or biotin ligases (CasID) to dCas9, to label proteins at a specific genomic locus, via sgRNA targeting (Schmidtman et al., 2016; Gao et al., 2018; Myers et al., 2018; Qiu et al., 2019). In initial experiments, CASPEX was localized to the hTERT promoter and identified known interactors, such as TP53 and MAZ (Myers et al., 2018). CasID also proved effective and successfully identified telomeric-bound proteins, such as components of the shelterin complex (Schmidtman et al., 2016).

CONCLUDING REMARKS

Chromatin is a dynamic structure; nucleosomes can lead to chromatin compaction and occlude accessibility, but also facilitate events, such as gene transcription. In addition to their dynamic nature in G₁ and G₂, H3.1, H3.3, and other nucleosomal histones are disassembled, reassembled, or altogether replaced during transcription and DNA replication. However, mechanisms exist to maintain epigenetic features, such

as histone PTMs, at precise locations. The intricacies of these processes are still under investigation, but our understanding of them has only been made possible by ever-evolving technologies and the ingenuity of the researchers that develop them.

Here, we highlighted H3 and its processing, PTMs, and interactions. However, other histones, histone variants, and their PTMs also have important biological consequences and these techniques can easily be applied to better understand other histone proteins. The experiments that are showcased are meant to provide a brief overview of the techniques used to study histones and their interactions, and covers but a snippet of available tools. As the chromatin biology field evolves, so will the technology, each time furthering our understanding of histone deposition, modification, binding, and eviction. The growing arsenal of techniques will allow us to continue to dissect histone dynamics through development, transcription, DNA replication & repair, mitosis, and countless other biological events in health and disease.

AUTHOR CONTRIBUTIONS

WS wrote the manuscript under the supervision and guidance of EC. All authors contributed to the article and approved the submitted version.

FUNDING

Financial support is provided by the Canadian Institutes of Health Research (WS and EC), the Ontario Graduate Scholarship program (WS), the National Sciences and Engineering Research Council (EC), the Cancer Research Society (EC), and the Garron Family Cancer Centre (EC).

ACKNOWLEDGMENTS

The authors would like to thank Stephanie Tran and Catherine Deane for insightful discussions on the article. Additionally, the authors would like to thank Erum Dhanji and Hwa Young Yun for assistance proof-reading the manuscript.

REFERENCES

- Ahmad, K., and Henikoff, S. (2002). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* 9, 1191–1200. doi: 10.1016/s1097-2765(02)00542-7
- Alabert, C., Barth, T. K., Reverón-Gómez, N., Sidoli, S., Schmidt, A., Jensen, O. N., et al. (2015). Two distinct modes for propagation of histone PTMs across the cell cycle. *Genes Dev.* 29, 585–590. doi: 10.1101/gad.256354.114
- Alabert, C., Bukowski-Wills, J. C., Lee, S. B., Kustatscher, G., Nakamura, K., De Lima Alves, F., et al. (2014). Nascent chromatin capture proteomics determines chromatin dynamics during DNA replication and identifies unknown fork components. *Nat. Cell Biol.* 16, 281–293.
- Alabert, C., Loos, C., Voelker-Albert, M., Graziano, S., Forne, I., Reveron-Gomez, N., et al. (2020). Domain Model Explains Propagation Dynamics and Stability of Histone H3K27 and H3K36 Methylation Landscapes. *Cell Rep.* 30:1223–1234 e1228.
- Albanese, K. I., Krone, M. W., Petell, C. J., Parker, M. M., Strahl, B. D., Brustad, E. M., et al. (2020). Engineered reader proteins for enhanced detection of methylated lysine on histones. *ACS Chem. Biol.* 15, 103–111. doi: 10.1021/acscchembio.9b00651
- Albig, W., Ebentheuer, J., Klobeck, G., Kunz, J., and Doenecke, D. (1996). A solitary human H3 histone gene on chromosome 1. *Hum. Genet.* 97, 486–491. doi: 10.1007/s004390050078
- Allfrey, V. G., Faulkner, R., and Mirsky, A. E. (1964). Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 51, 786–794. doi: 10.1073/pnas.51.5.786
- Allis, C. D., and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* 17, 487–500. doi: 10.1038/nrg.2016.59
- Alvarez, F., Munoz, F., Schilcher, P., Imhof, A., Almouzni, G., and Loyola, A. (2011). Sequential establishment of marks on soluble histones H3 and H4. *J. Biol. Chem.* 286, 17714–17721. doi: 10.1074/jbc.m111.223453

- Andrews, F. H., Strahl, B. D., and Kutateladze, T. G. (2016). Insights into newly discovered marks and readers of epigenetic information. *Nat. Chem. Biol.* 12, 662–668. doi: 10.1038/nchembio.2149
- Annunziato, A. T., Frado, L. L., Seale, R. L., and Woodcock, C. L. (1988). Treatment with sodium butyrate inhibits the complete condensation of interphase chromatin. *Chromosoma* 96, 132–138. doi: 10.1007/bf00331045
- Apta-Smith, M. J., Hernandez-Fernaund, J. R., and Bowman, A. J. (2018). Evidence for the nuclear import of histones H3.1 and H4 as monomers. *EMBO J.* 37:e98714.
- Asmari, M., Rati, R., Alhazmi, H. A., and El Deeb, S. (2018). Thermophoresis for characterizing biomolecular interaction. *Methods* 146, 107–119. doi: 10.1016/j.ymeth.2018.02.003
- Ausio, J. (2000). Analytical ultracentrifugation and the characterization of chromatin structure. *Biophys. Chem.* 86, 141–153. doi: 10.1016/s0301-4622(00)00144-7
- Bannister, A. J., Schneider, R., Myers, F. A., Thorne, A. W., Crane-Robinson, C., and Kouzarides, T. (2005). Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *J. Biol. Chem.* 280, 17732–17736. doi: 10.1074/jbc.m500796200
- Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., et al. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823–837. doi: 10.1016/j.cell.2007.05.009
- Bellelli, R., Belan, O., Pye, V. E., Clement, C., Maslen, S. L., Skehel, J. M., et al. (2018). POLE3-POLE4 Is a Histone H3-H4 chaperone that maintains chromatin integrity during DNA replication. *Mol. Cell* 72:112–126 e115.
- Bernstein, B. E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D. K., Huebert, D. J., et al. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* 120, 169–181. doi: 10.1016/j.cell.2005.01.001
- Binda, O., Leroy, G., Bua, D. J., Garcia, B. A., Gozani, O., and Richard, S. (2010). Trimethylation of histone H3 lysine 4 impairs methylation of histone H3 lysine 9: regulation of lysine methyltransferases by physical interaction with their substrates. *Epigenetics* 5, 767–775. doi: 10.4161/epi.5.8.13278
- Bintu, L., Ishibashi, T., Dangkulwanich, M., Wu, Y. Y., Lubkowska, L., Kashlev, M., et al. (2012). Nucleosomal elements that control the topography of the barrier to transcription. *Cell* 151, 738–749. doi: 10.1016/j.cell.2012.10.009
- Bitterge, B., and Schneider, R. (2014). Histone variants: key players of chromatin. *Cell Tissue Res.* 356, 457–466. doi: 10.1007/s00441-014-1862-4
- Bortvin, A., and Winston, F. (1996). Evidence That Spt6p controls chromatin structure by a direct interaction with histones. *Science* 272, 1473–1476. doi: 10.1126/science.272.5267.1473
- Bowman, A., Koide, A., Goodman, J. S., Colling, M. E., Zinne, D., Koide, S., et al. (2017). sNASP and ASF1A function through both competitive and compatible modes of histone binding. *Nucleic Acids Res.* 45, 643–656. doi: 10.1093/nar/gkw892
- Brahma, S., and Henikoff, S. (2020). Epigenome regulation by dynamic nucleosome unwrapping. *Trends Biochem. Sci.* 45, 13–26. doi: 10.1016/j.tibs.2019.09.003
- Brower-Toland, B. D., Smith, C. L., Yeh, R. C., Lis, J. T., Peterson, C. L., and Wang, M. D. (2002). Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA. *PNAS* 99, 1960–1965. doi: 10.1073/pnas.022638399
- Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y., and Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* 10, 1213–1218. doi: 10.1038/nmeth.2688
- Campos, E. I., Fillingham, J., Li, G., Zheng, H., Voigt, P., Kuo, W. H., et al. (2010). The program for processing newly synthesized histones H3.1 and H4. *Nat. Struct. Mol. Biol.* 17, 1343–1351.
- Campos, E. I., and Reinberg, D. (2009). Histones: annotating chromatin. *Annu. Rev. Genet.* 43, 559–599. doi: 10.1146/annurev.genet.032608.103928
- Carrozza, M. J., Li, B., Florens, L., Suganuma, T., Swanson, S. K., Lee, K. K., et al. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* 123, 581–592. doi: 10.1016/j.cell.2005.10.023
- Catania, S., and Allshire, R. C. (2014). Anarchic centromeres: deciphering order from apparent chaos. *Curr. Opin. Cell Biol.* 26, 41–50. doi: 10.1016/j.cob.2013.09.004
- Chahal, S. S., Matthews, H. R., and Bradbury, E. M. (1980). Acetylation of histone H4 and its role in chromatin structure and function. *Nature* 287, 76–79. doi: 10.1038/287076a0
- Chammas, P., Mocavini, I., and Di Croce, L. (2020). Engaging chromatin: PRC2 structure meets function. *Br. J. Cancer* 122, 315–328. doi: 10.1038/s41416-019-0615-2
- Chen, D., Hinkley, C. S., Henry, R. W., and Huang, S. (2002). TBP dynamics in living human cells: constitutive association of TBP with mitotic chromosomes. *Mol. Biol. Cell* 13, 276–284. doi: 10.1091/mbc.01-10-0523
- Chen, I., and Ting, A. Y. (2005). Site-specific labeling of proteins with small molecules in live cells. *Curr. Opin. Biotechnol.* 16, 35–40. doi: 10.1016/j.copbio.2004.12.003
- Chen, P., Dong, L., Hu, M., Wang, Y. Z., Xiao, X., Zhao, Z., et al. (2018). Functions of FACT in breaking the nucleosome and maintaining its integrity at the single-nucleosome level. *Mol. Cell* 71:284–293 e284.
- Chen, Z., Notti, R. Q., Ueberheide, B., and Ruthenburg, A. J. (2018). Quantitative and structural assessment of histone methyllysine analogue engagement by cognate binding proteins reveals affinity decrements relative to those of native counterparts. *Biochemistry* 57, 300–304. doi: 10.1021/acs.biochem.7b00926
- Cheutin, T., and Cavalli, G. (2019). The multiscale effects of polycomb mechanisms on 3D chromatin folding. *Crit. Rev. Biochem. Mol. Biol.* 54, 399–417. doi: 10.1080/10409238.2019.1679082
- Clark, D. J., and Felsenfeld, G. (1992). A nucleosome core is transferred out of the path of a transcribing polymerase. *Cell* 71, 11–22. doi: 10.1016/0092-8674(92)90262-b
- Cole, H. A., Ocampo, J., Iben, J. R., Chereji, R. V., and Clark, D. J. (2014). Heavy transcription of yeast genes correlates with differential loss of histone H2B relative to H4 and queued RNA polymerases. *Nucleic Acids Res.* 42, 12512–12522. doi: 10.1093/nar/gku1013
- Concepcion, J., Witte, K., Wartchow, C., Choo, S., Yao, D., Persson, H., et al. (2009). Label-free detection of biomolecular interactions using biolayer interferometry for kinetic characterization. *Comb. Chem. High Throughput Screen.* 12, 791–800. doi: 10.2174/138620709789104915
- Consortium, E. P., Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigo, R., Gingeras, T. R., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447, 799–816. doi: 10.1038/nature05874
- Cook, A. J., Gurard-Levin, Z. A., Vassias, I., and Almouzni, G. (2011). A specific function for the histone chaperone NASP to fine-tune a reservoir of soluble H3-H4 in the histone supply chain. *Mol. Cell* 44, 918–927. doi: 10.1016/j.molcel.2011.11.021
- Das, C., Tyler, J. K., and Churchill, M. E. (2010). The histone shuffle: histone chaperones in an energetic dance. *Trends Biochem. Sci.* 35, 476–489. doi: 10.1016/j.tibs.2010.04.001
- Davie, J. R., and Candido, E. P. (1978). Acetylated histone H4 is preferentially associated with template-active chromatin. *Proc. Natl. Acad. Sci. U.S.A.* 75, 3574–3577. doi: 10.1073/pnas.75.8.3574
- Deal, R. B., Henikoff, J. G., and Henikoff, S. (2010). Genome-Wide Kinetics of Nucleosome Turnover Determined by Metabolic Labeling of Histones. *Science* 328, 1161–1164. doi: 10.1126/science.1186777
- Déjardin, J., and Kingston, R. E. (2009). Purification of proteins associated with specific genomic Loci. *Cell* 136, 175–186. doi: 10.1016/j.cell.2008.11.045
- Deluz, C., Friman, E. T., Streibinger, D., Benke, A., Raccaud, M., Callegari, A., et al. (2016). A role for mitotic bookmarking of SOX2 in pluripotency and differentiation. *Genes Dev.* 30, 2538–2550. doi: 10.1101/gad.289256.116
- Dirks, R. A., Stunnenberg, H. G., and Marks, H. (2016). Genome-wide epigenomic profiling for biomarker discovery. *Clin. Epigenet.* 8:122.
- Donham, D. C. II, Scorgie, J. K., and Churchill, M. E. (2011). The activity of the histone chaperone yeast Asf1 in the assembly and disassembly of histone H3/H4-DNA complexes. *Nucleic Acids Res.* 39, 5449–5458. doi: 10.1093/nar/gkr097
- Drane, P., Ouararhni, K., Depaux, A., Shuaib, M., and Hamiche, A. (2010). The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev.* 24, 1253–1265. doi: 10.1101/gad.566910
- Dunleavy, E. M., Roche, D., Tagami, H., Lacoste, N., Ray-Gallet, D., Nakamura, Y., et al. (2009). HJURP is a cell-cycle-dependent maintenance and deposition

- factor of CENP-A at centromeres. *Cell* 137, 485–497. doi: 10.1016/j.cell.2009.02.040
- Duzdevich, D., and Greene, E. C. (2013). Towards physiological complexity with in vitro single-molecule biophysics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 368:20120271. doi: 10.1098/rstb.2012.0271
- Earnshaw, W. C., and Rothfield, N. (1985). Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma* 91, 313–321. doi: 10.1007/bf00328227
- Elsasser, S. J., Ernst, R. J., Walker, O. S., and Chin, J. W. (2016). Genetic code expansion in stable cell lines enables encoded chromatin modification. *Nat. Methods* 13, 158–164. doi: 10.1038/nmeth.3701
- Elsasser, S. J., Huang, H., Lewis, P. W., Chin, J. W., Allis, C. D., and Patel, D. J. (2012). DAXX envelops a histone H3.3–H4 dimer for H3.3-specific recognition. *Nature* 491, 560–565. doi: 10.1038/nature11608
- English, C. M., Adkins, M. W., Carson, J. J., Churchill, M. E., and Tyler, J. K. (2006). Structural basis for the histone chaperone activity of Asf1. *Cell* 127, 495–508. doi: 10.1016/j.cell.2006.08.047
- English, C. M., Maluf, N. K., Tripet, B., Churchill, M. E. A., and Tyler, J. K. (2005). ASF1 Binds to a Heterodimer of Histones H3 and H4: A Two-Step Mechanism for the Assembly of the H3–H4 Heterotetramer on DNA. *Biochemistry* 44, 13673–13682. doi: 10.1021/bi051333h
- Escobar, T. M., Oksuz, O., Saldana-Meyer, R., Descostes, N., Bonasio, R., and Reinberg, D. (2019). Active and Repressed Chromatin Domains Exhibit Distinct Nucleosome Segregation during DNA Replication. *Cell* 179:953–963 e911.
- Eubanks, C. G., Dayebgadah, G., Liu, X., and Washburn, M. P. (2017). Unravelling the biology of chromatin in health and cancer using proteomic approaches. *Expert. Rev. Proteom.* 14, 905–915. doi: 10.1080/14789450.2017.1374860
- Evrin, C., Maman, J. D., Diamante, A., Pellegrini, L., and Labib, K. (2018). Histone H2A–H2B binding by Pol alpha in the eukaryotic replisome contributes to the maintenance of repressive chromatin. *EMBO J.* 37:e99021.
- Festuccia, N., Gonzalez, I., Owens, N., and Navarro, P. (2017). Mitotic bookmarking in development and stem cells. *Development* 144, 3633–3645. doi: 10.1242/dev.146522
- Fischle, W., Tseng, B. S., Dormann, H. L., Ueberheide, B. M., Garcia, B. A., Shabanowitz, J., et al. (2005). Regulation of HP1–chromatin binding by histone H3 methylation and phosphorylation. *Nature* 438, 1116–1122. doi: 10.1038/nature04219
- Foltman, M., Evrin, C., De Piccoli, G., Jones, R. C., Edmondson, R. D., Katou, Y., et al. (2013). Eukaryotic replisome components cooperate to process histones during chromosome replication. *Cell Rep.* 3, 892–904. doi: 10.1016/j.celrep.2013.02.028
- Franklin, S. G., and Zweidler, A. (1977). Non-allelic variants of histones 2a, 2b and 3 in mammals. *Nature* 266, 273–275. doi: 10.1038/266273a0
- Gambus, A., Jones, R. C., Sanchez-Diaz, A., Kanemaki, M., Van Deursen, F., Edmondson, R. D., et al. (2006). GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nat. Cell Biol.* 8, 358–366. doi: 10.1038/ncb1382
- Gan, H., Serra-Cardona, A., Hua, X., Zhou, H., Labib, K., Yu, C., et al. (2018). The Mcm2–Ctf4–Polalpha Axis Facilitates Parental Histone H3–H4 Transfer to Lagging Strands. *Mol. Cell* 72:140–151 e143.
- Gao, X. D., Tu, L. C., Mir, A., Rodriguez, T., Ding, Y., Leszyk, J., et al. (2018). C-BERST: defining subnuclear proteomic landscapes at genomic elements with dCas9–APEX2. *Nat. Methods* 15, 433–436. doi: 10.1038/s41592-018-0006-2
- Garcia, H., Fleyshman, D., Kolesnikova, K., Safina, A., Commane, M., Paszkiewicz, G., et al. (2011). Expression of FACT in mammalian tissues suggests its role in maintaining of undifferentiated state of cells. *Oncotarget* 2, 783–796. doi: 10.18632/oncotarget.340
- Gauchier, M., Van Mierlo, G., Vermeulen, M., and Dejardin, J. (2020). Purification and enrichment of specific chromatin loci. *Nat. Methods* 17, 380–389. doi: 10.1038/s41592-020-0765-4
- Germond, J. E., Hirt, B., Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). Folding of the DNA double helix in chromatin-like structures from simian virus 40. *Proc. Natl. Acad. Sci. U.S.A.* 72, 1843–1847. doi: 10.1073/pnas.72.5.1843
- Gierlich, J., Burley, G. A., Gramlich, P. M., Hammond, D. M., and Carell, T. (2006). Click chemistry as a reliable method for the high-density postsynthetic functionalization of alkyne-modified DNA. *Org. Lett.* 8, 3639–3642. doi: 10.1021/ol0610946
- Giet, R., and Glover, D. M. (2001). Drosophila Aurora B Kinase Is Required for Histone H3 phosphorylation and condensin recruitment during chromosome condensation and to organize the central spindle during cytokinesis. *J. Cell Biol.* 152, 669–681.
- Gilmour, D. S., and Lis, J. T. (1984). Detecting protein–DNA interactions in vivo Distribution of RNA polymerase on specific bacterial genes. *Biochemistry* 81, 4275–4279. doi: 10.1073/pnas.81.14.4275
- Gingras, A. C., Gstaiger, M., Raught, B., and Aebersold, R. (2007). Analysis of protein complexes using mass spectrometry. *Nat. Rev. Mol. Cell Biol.* 8, 645–654.
- Go, C. D., Knight, J. D. R., Rajasekharan, A., Rathod, B., Hesketh, G. G., Abe, K. T., et al. (2019). A proximity biotinylation map of a human cell. *bioRxiv* [Preprint]. doi: 10.1101/796391
- Goldberg, A. D., Banaszynski, L. A., Noh, K. M., Lewis, P. W., Elsasser, S. J., Stadler, S., et al. (2010). Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* 140, 678–691.
- Gorovsky, M. A., Pleger, G. L., Keevert, J. B., and Johann, C. A. (1973). Studies on histone fraction F2A1 in macro- and micronuclei of Tetrahymena pyriformis. *J. Cell Biol.* 57, 773–781. doi: 10.1083/jcb.57.3.773
- Gosse, C., Strick, T. R., and Kostrz, D. (2019). Molecular scaffolds: when DNA becomes the hardware for single-molecule investigations. *Curr. Opin. Chem. Biol.* 53, 192–203. doi: 10.1016/j.cbpa.2019.09.006
- Goto, H., Tomono, Y., Ajiro, K., Kosako, H., Fujita, M., Sakurai, M., et al. (1999). Identification of a Novel Phosphorylation Site on Histone H3 coupled with mitotic chromosome condensation. *J. Biol. Chem.* 274, 25543–25549. doi: 10.1074/jbc.274.36.25543
- Goto, H., Yasui, Y., Nigg, E. A., and Inagaki, M. (2002). Aurora-B phosphorylates Histone H3 at serine28 with regard to the mitotic chromosome condensation. *Genes Cell* 7, 11–17. doi: 10.1046/j.1356-9597.2001.00498.x
- Groth, A., Ray-Gallet, D., Quivy, J. P., Lukas, J., Bartek, J., and Almouzni, G. (2005). Human Asf1 regulates the flow of S phase histones during replicational stress. *Mol. Cell* 17, 301–311. doi: 10.1016/j.molcel.2004.12.018
- Grove, G. W., and Zweidler, A. (1984). Regulation of nucleosomal core histone variant levels in differentiating murine erythroleukemia cells. *Biochemistry* 23, 4436–4443. doi: 10.1021/bi00314a030
- Grover, P., Asa, J. S., and Campos, E. I. (2018). H3–H4 histone chaperone pathways. *Annu. Rev. Genet.* 52, 109–130. doi: 10.1146/annurev-genet-120417-031547
- Gunjan, A., and Verreault, A. (2003). A Rad53 Kinase-Dependent Surveillance Mechanism that Regulates Histone Protein Levels in *S. cerevisiae*. *Cell* 115, 537–549. doi: 10.1016/s0092-8674(03)00896-1
- Gurard-Levin, Z. A., Quivy, J. P., and Almouzni, G. (2014). Histone chaperones: assisting histone traffic and nucleosome dynamics. *Annu. Rev. Biochem.* 83, 487–517. doi: 10.1146/annurev-biochem-060713-035536
- Gurley, L. R., Walters, R. A., and Tobey, R. A. (1974). Cell cycle-specific changes in histone phosphorylation associated with cell proliferation and chromosome condensation. *J. Cell Biol.* 60, 356–364. doi: 10.1083/jcb.60.2.356
- Gurova, K., Chang, H. W., Valieva, M. E., Sandlesh, P., and Studitsky, V. M. (2018). Structure and function of the histone chaperone FACT - Resolving FACTual issues. *Biochim. Biophys. Acta Gene Regul. Mech.* 1861, 892–904. doi: 10.1016/j.bbaggm.2018.07.008
- Hake, S. B., and Allis, C. D. (2006). Histone H3 variants and their potential role in indexing mammalian genomes: the "H3 barcode hypothesis". *Proc. Natl. Acad. Sci. U.S.A.* 103, 6428–6435. doi: 10.1073/pnas.0600803103
- Hake, S. B., Garcia, B. A., Duncan, E. M., Kauer, M., Dellaire, G., Shabanowitz, J., et al. (2006). Expression patterns and post-translational modifications associated with mammalian histone H3 variants. *J. Biol. Chem.* 281, 559–568. doi: 10.1074/jbc.m509266200
- Hake, S. B., Garcia, B. A., Kauer, M., Baker, S. P., Shabanowitz, J., Hunt, D. F., et al. (2005). Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes. *Proc. Natl. Acad. Sci. U.S.A.* 102, 6344–6349. doi: 10.1073/pnas.0502413102
- Hammond, C. M., Stromme, C. B., Huang, H., Patel, D. J., and Groth, A. (2017). Histone chaperone networks shaping chromatin function. *Nat. Rev. Mol. Cell Biol.* 18, 141–158. doi: 10.1038/nrm.2016.159
- Han, M., and Grunstein, M. (1988). Nucleosome loss activates yeast downstream promoters in Vivo. *Cell* 55:1137–1145.

- Hattori, T., and Koide, S. (2018). Next-generation antibodies for post-translational modifications. *Curr. Opin. Struct. Biol.* 51, 141–148. doi: 10.1016/j.sbi.2018.04.006
- He, S., Bauman, D., Davis, J. S., Loyola, A., Nishioka, K. †\$, Gronlund, J. L., et al. (2003). Facile synthesis of site-specifically acetylated and methylated histone proteins: Reagents for evaluation of the histone code hypothesis. *PNAS* 100, 12033–12038. doi: 10.1073/pnas.2035256100
- Hebbes, T. R., Thorne, A. W., and Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J.* 7, 1395–1402. doi: 10.1002/j.1460-2075.1988.tb02956.x
- Hendzel, M. J., Wei, Y., Mancini, M. A., Van Hooser, A., Ranalli, T., Brinkley, B. R., et al. (1997). Mitosis-specific phosphorylation of histone H3 initiates primarily within pericentromeric heterochromatin during G2 and spreads in an ordered fashion coincident with mitotic chromosome condensation. *Chromosoma* 106, 348–360. doi: 10.1007/s004120050256
- Henikoff, S., and Smith, M. M. (2015). Histone variants and epigenetics. *Cold Spring Harb. Perspect. Biol.* 7, a019364. doi: 10.1101/cshperspect.a019364
- Herman, T. M., Depamphilis, M. L., and Wassarman, P. M. (1981). Structure of chromatin at deoxyribonucleic acid replication forks: location of the first nucleosomes on newly synthesized simian virus 40 deoxyribonucleic acid. *Biochemistry* 20, 621–630. doi: 10.1021/bi00506a027
- Hirano, T., and Mitchison, T. J. (1994). A heterodimeric coiled-coil protein required for mitotic chromosome condensation in vitro. *Cell* 79, 449–458. doi: 10.1016/0092-8674(94)90254-2
- Hirota, T., Lipp, J. J., Toh, B. H., and Peters, J. M. (2005). Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* 438, 1176–1180. doi: 10.1038/nature04254
- Hodges, C., Bintu, L., Lubkowska, L., Kashlev, M., and Bustamante, C. (2009). Nucleosomal Fluctuations Govern the Transcription Dynamics of RNA Polymerase II. *Science* 325, 626–628. doi: 10.1126/science.1172926
- Holt, M., and Muir, T. (2015). Application of the protein semisynthesis strategy to the generation of modified chromatin. *Annu. Rev. Biochem.* 84, 265–290. doi: 10.1146/annurev-biochem-060614-034429
- Huang, H., Lin, S., Garcia, B. A., and Zhao, Y. (2015a). Quantitative proteomic analysis of histone modifications. *Chem. Rev.* 115, 2376–2418. doi: 10.1021/cr500491u
- Huang, H., Stromme, C. B., Saredi, G., Hodl, M., Strandsby, A., Gonzalez-Aguilera, C., et al. (2015b). A unique binding mode enables MCM2 to chaperone histones H3-H4 at replication forks. *Nat. Struct. Mol. Biol.* 22, 618–626. doi: 10.1038/nsmb.3055
- Huang, R., Holbert, M. A., Tarrant, M. K., Curtet, S., Colquhoun, D. R., Dancy, B. M., et al. (2010). Site-specific introduction of an Acetyl-Lysine mimic into peptides and proteins by cysteine alkylation. *J. Am. Chem. Soc.* 132, 9986–9987. doi: 10.1021/ja103954u
- Ide, S., and DeJardin, J. (2015). End-targeting proteomics of isolated chromatin segments of a mammalian ribosomal RNA gene promoter. *Nat. Commun.* 6:6674.
- Ishimi, Y., Komamura-Kohno, Y., Arai, K., and Masai, H. (2001). Biochemical activities associated with mouse MCM2 protein. *J. Biol. Chem.* 276, 42744–42752. doi: 10.1074/jbc.m106861200
- Ivanovska, I., Jacques, P. E., Rando, O. J., Robert, F., and Winston, F. (2011). Control of chromatin structure by spt6: different consequences in coding and regulatory regions. *Mol. Cell Biol.* 31, 531–541. doi: 10.1128/mcb.01068-10
- Jackson, V., Granner, D. K., and Chalkley, R. (1975). Deposition of histones onto replicating chromosomes. *Proc. Natl. Acad. Sci. U.S.A.* 72, 4440–4444. doi: 10.1073/pnas.72.11.4440
- Jamai, A., Imoberdorf, R. M., and Strubin, M. (2007). Continuous histone H2B and transcription-dependent histone H3 exchange in yeast cells outside of replication. *Mol. Cell* 25, 345–355. doi: 10.1016/j.molcel.2007.01.019
- Jansen, L. E., Black, B. E., Foltz, D. R., and Cleveland, D. W. (2007). Propagation of centromeric chromatin requires exit from mitosis. *J. Cell Biol.* 176, 795–805. doi: 10.1083/jcb.200701066
- Jeronimo, C., Poitras, C., and Robert, F. (2019). Histone Recycling by FACT and Spt6 during transcription prevents the scrambling of histone modifications. *Cell Rep.* 28:1206–1218 e1208.
- Ji, X., Dadon, D. B., Abraham, B. J., Lee, T. I., Jaenisch, R., Bradner, J. E., et al. (2015). Chromatin proteomic profiling reveals novel proteins associated with histone-marked genomic regions. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3841–3846.
- Kang, H., Shokhirev, M. N., Xu, Z., Chandran, S., Dixon, J. R., and Hetzer, M. W. (2020). Dynamic regulation of histone modifications and long-range chromosomal interactions during postmitotic transcriptional reactivation. *Genes Dev.* 34, 913–930. doi: 10.1101/gad.335794.119
- Kang, T. H., Park, D. Y., Choi, Y. H., Kim, K. J., Yoon, H. S., and Kim, K. T. (2007). Mitotic histone H3 phosphorylation by vaccinia-related kinase 1 in mammalian cells. *Mol. Cell Biol.* 27, 8533–8546. doi: 10.1128/mcb.00018-07
- Keogh, M. C., Kurdastani, S. K., Morris, S. A., Ahn, S. H., Podolny, V., Collins, S. R., et al. (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* 123, 593–605. doi: 10.1016/j.cell.2005.10.025
- Kim, D. I., and Roux, K. J. (2016). Filling the void: proximity-based labeling of proteins in living cells. *Trends Cell Biol.* 26, 804–817. doi: 10.1016/j.tcb.2016.09.004
- Kim, T. H., Barrera, L. O., Zheng, M., Qu, C., Singer, M. A., Richmond, T. A., et al. (2005). A high-resolution map of active promoters in the human genome. *Nature* 436, 876–880. doi: 10.1038/nature03877
- Kimura, H., and Cook, P. R. (2001). Kinetics of Core Histones in Living Human Cells: Little Exchange of H3 and H4 and Some Rapid Exchange of H2B. *J. Cell Biol.* 153, 1341–1353.
- Klemm, S. L., Shipony, Z., and Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* 20, 207–220. doi: 10.1038/s41576-018-0089-8
- Knezetic, J. A., and Luse, D. S. (1986). The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro. *Cell* 45, 95–104. doi: 10.1016/0092-8674(86)90541-6
- Kouskouti, A., and Talianidis, I. (2005). Histone modifications defining active genes persist after transcriptional and mitotic inactivation. *EMBO J.* 24, 347–357. doi: 10.1038/sj.emboj.7600516
- Kujirai, T., Ehara, H., Fujino, Y., Shirouzu, M., Sekine, S., and Kurumizaka, H. (2018). Structural basis of the nucleosome transition during RNA polymerase II passage. *Science* 362, 595–598. doi: 10.1126/science.aau9904
- Kujirai, T., and Kurumizaka, H. (2020). Transcription through the nucleosome. *Curr. Opin. Struct. Biol.* 61, 42–49.
- Kulaeva, O. I., Hsieh, F. K., and Studitsky, V. M. (2010). RNA polymerase complexes cooperate to relieve the nucleosomal barrier and evict histones. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11325–11330. doi: 10.1073/pnas.1001148107
- Kungulovski, G., Kycia, I., Tamas, R., Jurkowska, R. Z., Kudithipudi, S., Henry, C., et al. (2014). Application of histone modification-specific interaction domains as an alternative to antibodies. *Genome Res.* 24, 1842–1853. doi: 10.1101/gr.170985.113
- Kunitoku, N., Sasayama, T., Marumoto, T., Zhang, D., Honda, S., Kobayashi, O., et al. (2003). CENP-a phosphorylation by aurora-a in prophase is required for enrichment of aurora-b at inner centromeres and for kinetochore function. *Dev. Cell* 5, 853–864. doi: 10.1016/s1534-5807(03)00364-2
- Lachner, M., O'carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116–120. doi: 10.1038/35065132
- Lake, R. S., Goidl, J. A., and Salzman, N. P. (1972). F1-histone modification at metaphase in chinese hamster cells. *Exp. Cell Res.* 73, 113–121. doi: 10.1016/0014-4827(72)90108-5
- Lambert, J. P., Tucholska, M., Go, C., Knight, J. D., and Gingras, A. C. (2015). Proximity biotinylation and affinity purification are complementary approaches for the interactome mapping of chromatin-associated protein complexes. *J. Proteom.* 118, 81–94. doi: 10.1016/j.jprot.2014.09.011
- Larson, A. G., Elnatan, D., Keenen, M. M., Trnka, M. J., Johnston, J. B., Burlingame, A. L., et al. (2017). Liquid droplet formation by HP1alpha suggests a role for phase separation in heterochromatin. *Nature* 547, 236–240. doi: 10.1038/nature22822
- Laskey, R. A., Honda, B. M., Mills, A. D., and Finch, J. T. (1978). Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. *Nature* 275, 416–420. doi: 10.1038/275416a0
- Laugesen, A., Hojfeldt, J. W., and Helin, K. (2019). Molecular Mechanisms Directing PRC2 Recruitment and H3K27 Methylation. *Mol. Cell* 74, 8–18. doi: 10.1016/j.molcel.2019.03.011
- Lenfant, F., Mann, R. K., Thomsen, B., Ling, X., and Grunstein, M. (1996). All four core histone N-termini contain sequences required for the repression of basal transcription in yeast. *EMBO J.* 15, 3974–3985. doi: 10.1002/j.1460-2075.1996.tb00771.x

- LeRoy, G., Oksuz, O., Descostes, N., Aoi, Y., Ganai, R. A., Kara, H. O., et al. (2019). LEDGF and HDGF2 relieve the nucleosome-induced barrier to transcription in differentiated cells. *Sci. Adv.* 5:eay3068. doi: 10.1126/sciadv.aay3068
- Lewis, E. A., and Murphy, K. P. (2005). Isothermal titration calorimetry. *Methods Mol. Biol.* 305, 1–16.
- Lewis, P. W., Elsaesser, S. J., Noh, K. M., Stadler, S. C., and Allis, C. D. (2010). Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14075–14080. doi: 10.1073/pnas.1008850107
- Li, Y., Zeng, S. X., Landais, I., and Lu, H. (2007). Human SSRP1 has Spt16-dependent and -independent roles in gene transcription. *J. Biol. Chem.* 282, 6936–6945. doi: 10.1074/jbc.m603822200
- Liu, W. H., Roemer, S. C., Port, A. M., and Churchill, M. E. (2012). CAF-1-induced oligomerization of histones H3/H4 and mutually exclusive interactions with Asf1 guide H3/H4 transitions among histone chaperones and DNA. *Nucleic Acids Res.* 40, 11229–11239. doi: 10.1093/nar/gks906
- Liu, Y., Pelham-Webb, B., Di Giammartino, D. C., Li, J., Kim, D., Kita, K., et al. (2017). Widespread mitotic bookmarking by histone marks and transcription factors in pluripotent stem cells. *Cell Rep.* 19, 1283–1293. doi: 10.1016/j.celrep.2017.04.067
- Liu, Y., Zhou, K., Zhang, N., Wei, H., Tan, Y. Z., Zhang, Z., et al. (2020). FACT caught in the act of manipulating the nucleosome. *Nature* 577, 426–431. doi: 10.1038/s41586-019-1820-0
- Lorch, Y., Lapointe, J. W., and Kornberg, R. D. (1987). Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell* 49, 203–210. doi: 10.1016/0092-8674(87)90561-7
- Loyola, A., Bonaldi, T., Roche, D., Imhof, A., and Almouzni, G. (2006). PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Mol. Cell* 24, 309–316. doi: 10.1016/j.molcel.2006.08.019
- Loyola, A., He, S., Oh, S., McCafferty, D. G., and Reinberg, D. (2004). Techniques used to study transcription on chromatin templates. *Methods Enzymol.* 377, 474–499. doi: 10.1016/s0076-6879(03)77031-1
- Luco, R. F., Pan, Q., Tominaga, K., Blencowe, B. J., Pereira-Smith, O. M., and Misteli, T. (2010). Regulation of alternative splicing by histone modifications. *Science* 327, 996–1000. doi: 10.1126/science.1184208
- Luger, K., Mader, 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. doi: 10.1038/38444
- Lusser, A., and Kadonaga, J. T. (2004). Strategies for the reconstitution of chromatin. *Nat. Methods* 1, 19–26. doi: 10.1038/nmeth709
- Madamba, E. V., Berthet, E. B., and Francis, N. J. (2017). Inheritance of Histones H3 and H4 during DNA Replication In Vitro. *Cell Rep.* 21, 1361–1374. doi: 10.1016/j.celrep.2017.10.033
- Mann, R. K., and Grunstein, M. (1992). Histone H3 N-terminal mutations allow hyperactivation of the yeast GAL 1 gene in vivo. *EMBO J.* 11, 3297–3306. doi: 10.1002/j.1460-2075.1992.tb05408.x
- Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J., Drury, W. J., et al. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* 461, 762–767. doi: 10.1038/nature08398
- Martell, J. D., Deerinck, T. J., Sancar, Y., Poulos, T. L., Mootha, V. K., Sosinsky, G. E., et al. (2012). Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nat. Biotechnol.* 30, 1143–1148. doi: 10.1038/nbt.2375
- Martin, B. J. E., Chruscicki, A. T., and Howe, L. J. (2018). Transcription Promotes the Interaction of the FAcilitates Chromatin Transactions (FACT) complex with nucleosomes in *Saccharomyces cerevisiae*. *Genetics* 210, 869–881. doi: 10.1534/genetics.118.301349
- Marzluff, W. F., Gongidi, P., Woods, K. R., Jin, J., and Maltais, L. J. (2002). The human and mouse replication-dependent histone genes. *Genomics* 80, 487–498. doi: 10.1006/geno.2002.6850
- Mattioli, F., Gu, Y., Yadav, T., Balsbaugh, J. L., Harris, M. R., Findlay, E. S., et al. (2017). DNA-mediated association of two histone-bound complexes of yeast Chromatin Assembly Factor-1 (CAF-1) drives tetrasome assembly in the wake of DNA replication. *eLife* 6:e22799.
- Mausser, R., and Jeltsch, A. (2019). Application of modified histone peptide arrays in chromatin research. *Arch. Biochem. Biophys.* 661, 31–38. doi: 10.1016/j.ab.2018.10.019
- Mayanagi, K., Saikusa, K., Miyazaki, N., Akashi, S., Iwasaki, K., Nishimura, Y., et al. (2019). Structural visualization of key steps in nucleosome reorganization by human FACT. *Sci. Rep.* 9:10183.
- Meeks-Wagner, D., and Hartwell, L. H. (1986). Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. *Cell* 44, 43–52. doi: 10.1016/0092-8674(86)90483-6
- Mello, J. A., Sillje, H. H., Roche, D. M., Kirschner, D. B., Nigg, E. A., and Almouzni, G. (2002). Human Asf1 and CAF-1 interact and synergize in a repair-coupled nucleosome assembly pathway. *EMBO Rep.* 3, 329–334. doi: 10.1093/embo-reports/kvf068
- Mendiratta, S., Gatto, A., and Almouzni, G. (2019). Histone supply: Multitiered regulation ensures chromatin dynamics throughout the cell cycle. *J. Cell Biol.* 218, 39–54. doi: 10.1083/jcb.201807179
- Moyer, S. E., Lewis, P. W., and Botchan, M. R. (2006). Isolation of the Cdc45/Mcm2–7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *PNAS* 103, 10236–10241. doi: 10.1073/pnas.0602400103
- Muller, S., and Almouzni, G. (2017). Chromatin dynamics during the cell cycle at centromeres. *Nat. Rev. Genet.* 18, 192–208. doi: 10.1038/nrg.2016.157
- Myers, S. A., Wright, J., Peckner, R., Kalish, B. T., Zhang, F., and Carr, S. A. (2018). Discovery of proteins associated with a predefined genomic locus via dCas9-APEX-mediated proximity labeling. *Nat. Methods* 15, 437–439. doi: 10.1038/s41592-018-0007-1
- Nakato, R., and Sakata, T. (2020). Methods for ChIP-seq analysis: A practical workflow and advanced applications. *Methods* doi: 10.1016/j.jymeth.2020.03.005 [Online ahead of print]
- Natsume, R., Eitoku, M., Akai, Y., Sano, N., Horikoshi, M., and Senda, T. (2007). Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4. *Nature* 446, 338–341. doi: 10.1038/nature05613
- Nelson, D. A., Perry, W. M., and Chalkley, R. (1978). Sensitivity of regions of chromatin containing hyperacetylated histones to DNase I. *Biochem. Biophys. Res. Commun.* 82, 365–363.
- Nelson, D. M., Ye, X., Hall, C., Santos, H., Ma, T., Kao, G. D., et al. (2002). Coupling of DNA synthesis and histone synthesis in S phase independent of cyclin/cdk2 activity. *Mol. Cell Biol.* 22, 7459–7472. doi: 10.1128/mcb.22.21.7459-7472.2002
- Neumann, H., Peak-Chew, S. Y., and Chin, J. W. (2008). Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. *Nat. Chem. Biol.* 4, 232–234. doi: 10.1038/nchembio.73
- Noll, M. (1974). Subunit structure of chromatin. *Nature* 251, 249–251. doi: 10.1038/251249a0
- Noll, M., Thomas, J. O., and Kornberg, R. D. (1975). Preparation of Native Chromatin and Damage Caused by Shearing. *Science* 187, 1203–1206. doi: 10.1126/science.187.4182.1203
- Ohsumi, K., Katagiri, C., and Kishimoto, T. (1993). Chromosome Condensation in *Xenopus* Mitotic Extracts Without Histone H1. *Science* 262, 2033–2035. doi: 10.1126/science.8266099
- Oksuz, O., Narendra, V., Lee, C. H., Descostes, N., Leroy, G., Raviram, R., et al. (2018). Capturing the Onset of PRC2-mediated repressive domain formation. *Mol. Cell* 70:1149–1162 e1145.
- Olins, A. L., Senior, M. B., and Olins, D. E. (1976). Ultrastructural features of chromatin nu bodies. *J. Cell Biol.* 68, 787–793. doi: 10.1083/jcb.68.3.787
- Orphanides, G., Leroy, G., Chang, C., Luse, D. S., and Reinberg, D. (1998). FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* 92, 105–116. doi: 10.1016/s0092-8674(00)80903-4
- Palozola, K. C., Lerner, J., and Zaret, K. S. (2019). A changing paradigm of transcriptional memory propagation through mitosis. *Nat. Rev. Mol. Cell Biol.* 20, 55–64. doi: 10.1038/s41580-018-0077-z
- Pardal, A. J., Fernandes-Duarte, F., and Bowman, A. J. (2019). The histone chaperoning pathway: from ribosome to nucleosome. *Essays Biochem.* 63, 29–43. doi: 10.1042/ebc20180055
- Patel, D. J., and Wang, Z. (2013). Readout of epigenetic modifications. *Annu. Rev. Biochem.* 82, 81–118. doi: 10.1146/annurev-biochem-072711-165700
- Paulson, J. R., and Taylor, S. S. (1982). Phosphorylation of Histones 1 and 3 and Nonhistone High Mobility Group 14 by an Endogenous Kinase in HeLa metaphase chromosome. *J. Biol. Chem.* 257, 6064–6072.
- Pavri, R., Zhu, B., Li, G., Trojer, P., Mandal, S., Shilatifard, A., et al. (2006). Histone H2B monoubiquitination functions cooperatively with FACT to regulate

- elongation by RNA polymerase II. *Cell* 125, 703–717. doi: 10.1016/j.cell.2006.04.029
- Petryk, N., Dalby, M., Wenger, A., Stromme, C. B., Strandsby, A., Andersson, R., et al. (2018). MCM2 promotes symmetric inheritance of modified histones during DNA replication. *Science* 361, 1389–1392. doi: 10.1126/science.aau0294
- Postberg, J., Forcob, S., Chang, W. J., and Lipps, H. J. (2010). The evolutionary history of histone H3 suggests a deep eukaryotic root of chromatin modifying mechanisms. *BMC Evol. Biol.* 10:259. doi: 10.1186/1471-2148-10-259
- Qiu, W., Xu, Z., Zhang, M., Zhang, D., Fan, H., Li, T., et al. (2019). Determination of local chromatin interactions using a combined CRISPR and peroxidase APEX2 system. *Nucleic Acids Res.* 47:e52. doi: 10.1093/nar/gkz134
- Rafiee, M. R., Girardot, C., Sigismondo, G., and Krijgsvel, J. (2016). Expanding the circuitry of pluripotency by selective isolation of chromatin-associated proteins. *Mol. Cell* 64, 624–635. doi: 10.1016/j.molcel.2016.09.019
- Ray-Gallet, D., Quivy, J. P., Scamps, C., Martini, E. M., Lipinski, M., and Almouzni, G. (2002). HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. *Mol. Cell* 9, 1091–1100. doi: 10.1016/s1097-2765(02)00526-9
- Ray-Gallet, D., Woolfe, A., Vassias, I., Pellentz, C., Lacoste, N., Puri, A., et al. (2011). Dynamics of histone H3 deposition in vivo reveal a nucleosome gap-filling mechanism for H3.3 to maintain chromatin integrity. *Mol. Cell* 44, 928–941. doi: 10.1016/j.molcel.2011.12.006
- Rea, S., Eisenhaber, F., O'carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593–599. doi: 10.1038/35020506
- Reinberg, D., and Vales, L. D. (2018). Chromatin domains rich in inheritance. *Science* 361, 33–34. doi: 10.1126/science.aat7871
- Reveron-Gomez, N., Gonzalez-Aguilera, C., Stewart-Morgan, K. R., Petryk, N., Flury, V., Graziano, S., et al. (2018). Accurate Recycling of Parental Histones Reproduces the Histone Modification Landscape during DNA Replication. *Mol. Cell* 72:239–249 e235.
- Rhee, H. W., Zou, P., Udeshi, N. D., Martell, J. D., Mootha, V. K., Carr, S. A., et al. (2013). Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* 339, 1328–1331. doi: 10.1126/science.1230593
- Ricketts, M. D., Frederick, B., Hoff, H., Tang, Y., Schultz, D. C., Singh Rai, T., et al. (2015). Ubinuclein-1 confers histone H3.3-specific-binding by the HIRA histone chaperone complex. *Nat. Commun.* 6:7711.
- Ricketts, M. D., Han, J., Szurgot, M. R., and Marmorstein, R. (2019). Molecular basis for chromatin assembly and modification by multiprotein complexes. *Protein Sci.* 28, 329–343. doi: 10.1002/pro.3535
- Rivera, C., Saavedra, F., Alvarez, F., Diaz-Celis, C., Ugalde, V., Li, J., et al. (2015). Methylation of histone H3 lysine 9 occurs during translation. *Nucleic Acids Res.* 43, 9097–9106. doi: 10.1093/nar/gkv929
- Robyr, D., Suka, Y., Xenarios, I., Kurdistani, S. K., Wang, A., Suka, N., et al. (2002). Microarray deacetylation maps determine genome-wide functions for yeast histone deacetylases. *Cell* 109, 437–446. doi: 10.1016/s0092-8674(02)00746-8
- Roh, T. Y., Cuddapah, S., Cui, K., and Zhao, K. (2006). The genomic landscape of histone modifications in human T cells. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15782–15787. doi: 10.1073/pnas.0607617103
- Roh, T. Y., Ngau, W. C., Cui, K., Landsman, D., and Zhao, K. (2004). High-resolution genome-wide mapping of histone modifications. *Nat. Biotechnol.* 22, 1013–1016. doi: 10.1038/nbt990
- Rolf Ben-Shahar, T., Castillo, A. G., Osborne, M. J., Borden, K. L., Kornblatt, J., and Verreault, A. (2009). Two fundamentally distinct PCNA interaction peptides contribute to chromatin assembly factor 1 function. *Mol. Cell Biol.* 29, 6353–6365. doi: 10.1128/mcb.01051-09
- Rothbart, S. B., Dickson, B. M., Raab, J. R., Grzybowski, A. T., Krajewski, K., Guo, A. H., et al. (2015). An interactive database for the assessment of histone antibody specificity. *Mol. Cell* 59, 502–511. doi: 10.1016/j.molcel.2015.06.022
- Roux, K. J., Kim, D. I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* 196, 801–810. doi: 10.1083/jcb.201112098
- Ruthenburg, A. J., Li, H., Milne, T. A., Dewell, S., McGinty, R. K., Yuen, M., et al. (2011). Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell* 145, 692–706. doi: 10.1016/j.cell.2011.03.053
- Samavarchi-Tehrani, P., Samson, R., and Gingras, A. C. (2020). Proximity Dependent Biotinylation: Key Enzymes and Adaptation to Proteomics Approaches. *Mol. Cell Proteomics* 19, 757–773. doi: 10.1074/mcp.r120.001941
- Sarai, N., Nimura, K., Tamura, T., Kanno, T., Patel, M. C., Heightman, T. D., et al. (2013). WHSC1 links transcription elongation to HIRA-mediated histone H3.3 deposition. *EMBO J.* 32, 2392–2406. doi: 10.1038/emboj.2013.176
- Sauer, P. V., Gu, Y., Liu, W. H., Mattioli, F., Panne, D., Luger, K., et al. (2018). Mechanistic insights into histone deposition and nucleosome assembly by the chromatin assembly factor-1. *Nucleic Acids Res.* 46, 9907–9917. doi: 10.1093/nar/gky823
- Schenk, R., Jenke, A., Zilbauer, M., Wirth, S., and Postberg, J. (2011). H3.5 is a novel hominid-specific histone H3 variant that is specifically expressed in the seminiferous tubules of human testes. *Chromosoma* 120, 275–285. doi: 10.1007/s00412-011-0310-4
- Schlissel, G., and Rine, J. (2019). The nucleosome core particle remembers its position through DNA replication and RNA transcription. *Proc. Natl. Acad. Sci. U.S.A.* 116, 20605–20611. doi: 10.1073/pnas.1911943116
- Schmidtman, E., Anton, T., Rombaut, P., Herzog, F., and Leonhardt, H. (2016). Determination of local chromatin composition by CasID. *Nucleus* 7, 476–484. doi: 10.1080/19491034.2016.1239000
- Schmitges, F. W., Prusty, A. B., Faty, M., Stutzer, A., Lingaraju, G. M., Aiwanian, J., et al. (2011). Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol. Cell* 42, 330–341. doi: 10.1016/j.molcel.2011.03.025
- Schones, D. E., Cui, K., Cuddapah, S., Roh, T. Y., Barski, A., Wang, Z., et al. (2008). Dynamic regulation of nucleosome positioning in the human genome. *Cell* 132, 887–898. doi: 10.1016/j.cell.2008.02.022
- Schwartz, B. E., and Ahmad, K. (2005). Transcriptional activation triggers deposition and removal of the histone variant H3.3. *Genes Dev.* 19, 804–814. doi: 10.1101/gad.1259805
- Schwartzman, O., and Tanay, A. (2015). Single-cell epigenomics: techniques and emerging applications. *Nat. Rev. Genet.* 16, 716–726. doi: 10.1038/nrg3980
- Seeliger, D., Soeroes, S., Klingberg, R., Schwarzer, D., Grubmüller, H., and Fischle, W. (2012). Quantitative assessment of protein interaction with methyl-lysine analogues by hybrid computational and experimental approaches. *ACS Chem. Biol.* 7, 150–154. doi: 10.1021/cb200363r
- Senapati, P., Sudarshan, D., Gadad, S. S., Shandilya, J., Swaminathan, V., and Kundu, T. K. (2015). Methods to study histone chaperone function in nucleosome assembly and chromatin transcription. *Methods Mol. Biol.* 1288, 375–394. doi: 10.1007/978-1-4939-2474-5_22
- Sequeira, V. M., and Vermeulen, M. (2019). Identifying Readers for (hydroxy)methylated DNA Using Quantitative Interaction Proteomics: Advances and Challenges Ahead. *J. Mol. Biol.* doi: 10.1016/j.jmb.2019.12.014 [Online ahead of print]
- Shogren-Knaak, M. A., Fry, C. J., and Peterson, C. L. (2003). A native peptide ligation strategy for deciphering nucleosomal histone modifications. *J. Biol. Chem.* 278, 15744–15748. doi: 10.1074/jbc.m301445200
- Simithy, J., Sidoli, S., and Garcia, B. A. (2018). Integrating proteomics and targeted metabolomics to understand global changes in histone modifications. *Proteomics* 18:e1700309.
- Simon, M. D., Chu, F., Racki, L. R., De La Cruz, C. C., Burlingame, A. L., Panning, B., et al. (2007). The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell* 128, 1003–1012. doi: 10.1016/j.cell.2006.12.041
- Simpson, R. T. (1978). Structure of chromatin containing extensively acetylated H3 and H4. *Cell* 13, 691–699. doi: 10.1016/0092-8674(78)90219-2
- Sims, R. J. III, Millhouse, S., Chen, C. F., Lewis, B. A., Erdjument-Bromage, H., Tempst, P., et al. (2007). Recognition of trimethylated histone H3 lysine 4 facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol. Cell* 28, 665–676. doi: 10.1016/j.molcel.2007.11.010
- Sirbu, B. M., Couch, F. B., and Cortez, D. (2012). Monitoring the spatiotemporal dynamics of proteins at replication forks and in assembled chromatin using isolation of proteins on nascent DNA. *Nat. Protoc.* 7, 594–605. doi: 10.1038/nprot.2012.010
- Smith, S., and Stillman, B. (1989). Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. *Cell* 58, 15–25. doi: 10.1016/0092-8674(89)90398-x
- Smolle, M., and Workman, J. L. (2013). Transcription-associated histone modifications and cryptic transcription. *Biochim. Biophys. Acta* 1829, 84–97. doi: 10.1016/j.bbagr.2012.08.008

- Smothers, J. F., and Henikoff, S. (2000). The HP1 chromo shadow domain binds a consensus peptide pentamer. *Curr. Biol.* 10, 27–30. doi: 10.1016/s0960-9822(99)00260-2
- Sogo, J. M., Stahl, H., Koller, T., and Knippers, R. (1986). Structure of replicating simian virus 40 minichromosomes. The replication fork, core histone segregation and terminal structures. *J. Mol. Biol.* 189, 189–204. doi: 10.1016/0022-2836(86)90390-6
- Solomon, M. J., Larsen, P. L., and Varshavsky, A. (1988). Mapping Protein-DNA Interactions In Vivo with Formaldehyde Evidence That Histone H4 is retained on a highly transcribed gene. *Cell* 53, 937–947. doi: 10.1016/s0092-8674(88)90469-2
- Stetefeld, J., McKenna, S. A., and Patel, T. R. (2016). Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophys. Rev.* 8, 409–427. doi: 10.1007/s12551-016-0218-6
- Stewart-Morgan, K. R., Petryk, N., and Groth, A. (2020). Chromatin replication and epigenetic cell memory. *Nat. Cell Biol.* 22, 361–371. doi: 10.1038/s41556-020-0487-y
- Stewart-Morgan, K. R., Reveron-Gomez, N., and Groth, A. (2019). Transcription restart establishes chromatin accessibility after DNA replication. *Mol. Cell* 75:284–297 e286.
- Strom, A. R., Emelyanov, A. V., Mir, M., Fyodorov, D. V., Darzacq, X., and Karpen, G. H. (2017). Phase separation drives heterochromatin domain formation. *Nature* 547, 241–245. doi: 10.1038/nature22989
- Studitsky, V. M., Clark, D. J., and Felsenfeld, G. (1994). A histone octamer can step around a transcribing polymerase without leaving the template. *Cell* 76, 371–382. doi: 10.1016/0092-8674(94)90343-3
- Tagami, H., Almouzni, G., Ray-Gallet, D., and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* 116, 51–61. doi: 10.1016/s0092-8674(03)01064-x
- Taguchi, H., Xie, Y., Horikoshi, N., Maehara, K., Harada, A., Nogami, J., et al. (2017). Crystal Structure and Characterization of Novel Human Histone H3 Variants. H3.6, H3.7, and H3.8. *Biochemistry* 56, 2184–2196.
- Takemoto, A., Murayama, A., Katano, M., Urano, T., Furukawa, K., Yokoyama, S., et al. (2007). Analysis of the role of Aurora B on the chromosomal targeting of condensin I. *Nucleic Acids Res.* 35, 2403–2412. doi: 10.1093/nar/gkm157
- Talbert, P. B., Ahmad, K., Almouzni, G., Ausio, J., Berger, F., Bhalla, P. L., et al. (2012). A unified phylogeny-based nomenclature for histone variants. *Epigenet. Chromatin* 5:7.
- Talbert, P. B., and Henikoff, S. (2006). Spreading of silent chromatin: inaction at a distance. *Nat. Rev. Genet.* 7, 793–803. doi: 10.1038/nrg1920
- Talbert, P. B., and Henikoff, S. (2010). Histone variants-ancient wrap artists of the epigenome. *Nat. Rev. Mol. Cell Biol.* 11, 264–275. doi: 10.1038/nrm2861
- Terzi, N., Churchman, L. S., Vasiljeva, L., Weissman, J., and Buratowski, S. (2011). H3K4 trimethylation by Set1 promotes efficient termination by the Nrd1-Nab3-Sen1 pathway. *Mol. Cell Biol.* 31, 3569–3583. doi: 10.1128/mcb.05590-11
- Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., et al. (2013). Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. *Cell* 152, 859–872. doi: 10.1016/j.cell.2013.01.032
- Tse, C., Sera, T., Wolffe, A. P., and Hansen, J. C. (1998). Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. *Mol. Cell Biol.* 18, 4629–4638. doi: 10.1128/mcb.18.8.4629
- Tyler, J. K., Collins, K. A., Prasad-Sinha, J., Amiot, E., Bulger, M., Harte, P. J., et al. (2001). Interaction between the Drosophila CAF-1 and ASF1 chromatin assembly factors. *Mol. Cell Biol.* 21, 6574–6584. doi: 10.1128/mcb.21.19.6574-6584.2001
- Valls, E., Sanchez-Molina, S., and Martinez-Balbas, M. A. (2005). Role of histone modifications in marking and activating genes through mitosis. *J. Biol. Chem.* 280, 42592–42600. doi: 10.1074/jbc.m507407200
- van Mierlo, G., Veenstra, G. J. C., Vermeulen, M., and Marks, H. (2019). The Complexity of PRC2 Subcomplexes. *Trends Cell Biol.* 29, 660–671. doi: 10.1016/j.tcb.2019.05.004
- Vermeulen, M., Eberl, H. C., Matarese, F., Marks, H., Denisov, S., Butter, F., et al. (2010). Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142, 967–980. doi: 10.1016/j.cell.2010.08.020
- Vermeulen, M., Mulder, K. W., Denisov, S., Pijnappel, W. W., Van Schaik, F. M., Varier, R. A., et al. (2007). Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* 131, 58–69. doi: 10.1016/j.cell.2007.08.016
- Vidali, G., Boffa, L. C., Bradbury, E. M., and Allfrey, V. G. (1978). Butyrate suppression of histone deacetylation leads to accumulation of multiacetylated forms of histones H3 and H4 and increased DNase I sensitivity of the associated DNA sequences. *Proc. Natl. Acad. Sci. U.S.A.* 75, 2239–2243. doi: 10.1073/pnas.75.5.2239
- Villasenor, R., Pfandler, R., Ambrosi, C., Butz, S., Giuliani, S., Bryan, E., et al. (2020). ChromID identifies the protein interactome at chromatin marks. *Nat. Biotechnol.* 38, 728–736. doi: 10.1038/s41587-020-0434-2
- Voigt, P., Leroy, G., Drury, W. J. III, Zee, B. M., Son, J., Beck, D. B., et al. (2012). Asymmetrically modified nucleosomes. *Cell* 151, 181–193. doi: 10.1016/j.cell.2012.09.002
- Wallace, R. B., Sargent, T. D., Murphy, R. F., and Bonner, J. (1977). Physical properties of chemically acetylated rat liver chromatin. *Proc. Natl. Acad. Sci. U.S.A.* 74, 3244–3248. doi: 10.1073/pnas.74.8.3244
- Wang, T., Liu, Y., Edwards, G., Krzizike, D., Scherman, H., and Luger, K. (2018). The histone chaperone FACT modulates nucleosome structure by tethering its components. *Life Sci Alliance* 1:e201800107. doi: 10.26508/lsa.201800107
- Wang, Z., Schones, D. E., and Zhao, K. (2009). Characterization of human epigenomes. *Curr. Opin. Genet. Dev.* 19, 127–134.
- Wang, Z., Zang, C., Rosenfeld, J. A., Schones, D. E., Barski, A., Cuddapah, S., et al. (2008). Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat. Genet.* 40, 897–903. doi: 10.1038/ng.154
- Wiedemann, S. M., Mildner, S. N., Bonisch, C., Israel, L., Maiser, A., Matheis, S., et al. (2010). Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y. *J. Cell Biol.* 190, 777–791. doi: 10.1083/jcb.201002043
- Wong, T. K., and Marushige, K. (1976). Modification of histone binding in calf thymus chromatin and in the chromatin-protamine complex by acetic anhydride. *Biochemistry* 15, 2041–2046. doi: 10.1021/bi00655a003
- Wu, C., Bingham, P. M., Livak, K. J., Holmgren, R., and Elgin, S. C. R. (1979). The Chromatin Structure of Specific Genes: I. Evidence for Higher Order Domains of Defined DNA sequence. *Cell* 16, 797–806. doi: 10.1016/0092-8674(79)90095-3
- Wu, R. S., and Bonner, W. M. (1981). Separation of Basal Histone Synthesis from S-Phase Histone Synthesis in Dividing Cells. *Cell* 27, 321–330. doi: 10.1016/0092-8674(81)90415-3
- Wu, R. S., Tsai, S., and Bonner, W. M. (1982). Patterns of histone variant synthesis can distinguish G0 from G1 cells. *Cell* 31, 367–374. doi: 10.1016/0092-8674(82)90130-1
- Xu, M., Long, C., Chen, X., Huang, C., Chen, S., and Zhu, B. (2010). Partitioning of Histone H3-H4 Tetramers during DNA Replication-Dependent Chromatin Assembly. *Science* 328, 94–98. doi: 10.1126/science.1178994
- Young, D. W., Hassan, M. Q., Pratap, J., Galindo, M., Zaidi, S. K., Lee, S. H., et al. (2007). Mitotic occupancy and lineage-specific transcriptional control of rRNA genes by Runx2. *Nature* 445, 442–446. doi: 10.1038/nature05473
- Yu, J. R., Lee, C. H., Oksuz, O., Stafford, J. M., and Reinberg, D. (2019). PRC2 is high maintenance. *Genes Dev.* 33, 903–935. doi: 10.1101/gad.325050.119
- Yuan, G. C., Liu, Y. J., Dion, M. F., Slack, M. D., Wu, L. F., Altschuler, S. J., et al. (2005). Genome-Scale Identification of Nucleosome Positions in *S. cerevisiae*. *Science* 309, 626–630. doi: 10.1126/science.1112178
- Yuan, W., Xu, M., Huang, C., Liu, N., Chen, S., and Zhu, B. (2011). H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J. Biol. Chem.* 286, 7983–7989. doi: 10.1074/jbc.m110.194027
- Zasadzinska, E., Huang, J., Bailey, A. O., Guo, L. Y., Lee, N. S., Srivastava, S., et al. (2018). Inheritance of CENP-A Nucleosomes during DNA Replication Requires HJURP. *Dev. Cell* 47:348–362 e347.

- Zee, B. M., Britton, L. M., Wolle, D., Haberman, D. M., and Garcia, B. A. (2012). Origins and formation of histone methylation across the human cell cycle. *Mol. Cell Biol.* 32, 2503–2514. doi: 10.1128/mcb.06673-11
- Zeitlin, S. G., Barber, C. M., Allis, C. D., and Sullivan, K. E. (2001). Differential regulation of CENP-A and histone H3 phosphorylation in G2/M. *J. Cell Sci.* 114, 653–661.
- Zhao, Y., and Garcia, B. A. (2015). Comprehensive catalog of currently documented histone modifications. *Cold Spring Harb. Perspect. Biol.* 7:a025064. doi: 10.1101/cshperspect.a025064

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Scott and Campos. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Structural Paradigms in the Recognition of the Nucleosome Core Particle by Histone Lysine Methyltransferases

Ashley Janna[†], Hossein Davarinejad[†], Monika Joshi[†] and Jean-Francois Couture^{*}

Ottawa Institute of Systems Biology, Shanghai Institute of Materia Medica-University of Ottawa Research Center in Systems and Personalized Pharmacology, Department of Biochemistry, Microbiology and Immunology, University of Ottawa, Ottawa, ON, Canada

OPEN ACCESS

Edited by:

Christoph Franz Kurat,
Ludwig Maximilian University
of Munich, Germany

Reviewed by:

Axel Imhof,
Ludwig Maximilian University
of Munich, Germany
James Ronald Davie,
University of Manitoba, Canada

*Correspondence:

Jean-Francois Couture
jean-francois.couture@uottawa.ca

[†]These authors share first authorship

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 08 April 2020

Accepted: 19 June 2020

Published: 31 July 2020

Citation:

Janna A, Davarinejad H, Joshi M
and Couture J-F (2020) Structural
Paradigms in the Recognition of the
Nucleosome Core Particle by Histone
Lysine Methyltransferases.
Front. Cell Dev. Biol. 8:600.
doi: 10.3389/fcell.2020.00600

Post-translational modifications (PTMs) of histone proteins play essential functions in shaping chromatin environment. Alone or in combination, these PTMs create templates recognized by dedicated proteins or change the chemistry of chromatin, enabling a myriad of nuclear processes to occur. Referred to as cross-talk, the positive or negative impact of a PTM on another PTM has rapidly emerged as a mechanism controlling nuclear transactions. One of those includes the stimulatory functions of histone H2B ubiquitylation on the methylation of histone H3 on K79 and K4 by Dot1L and COMPASS, respectively. While these findings were established early on, the structural determinants underlying the positive impact of H2B ubiquitylation on H3K79 and H3K4 methylation were resolved only recently. We will also review the molecular features controlling these cross-talks and the impact of H3K27 tri-methylation on EZH2 activity when embedded in the PRC2 complex.

Keywords: histone, epigenetics, methylation, ubiquitylation, chromatin

INTRODUCTION—THE NUCLEOSOME

The genetic material of a typical eukaryotic cell approximately measures 2 meters and must be restricted to the confines of the nucleus. The cell employs four α -helical basic proteins to create a scaffold around which DNA can be compacted: histones H2A, H2B, H3, and H4. First, two histone H3–H4 heterodimers dimerize to form a heterotetramer, upon which two H2A–H2B heterodimers will bind. The H2A protomers contact H3 and H4 at the extremities of the heterotetramer; meanwhile, the H2B protomers form an extensive dimerization interface (Arents et al., 1991) to create a symmetrical disk-shaped histone octamer. A DNA fragment of approximately 150 bp will then wrap twice around the histone octamer of basic histone proteins to form a repetitive structure known as the nucleosome (Noll, 1977; Luger et al., 1997) [referred therein as nucleosome core particle (NCP)]. However, in recent years, incorporation of histone variants in nucleosomes brought diversity to that model (Koyama and Kurumizaka, 2018; Talbert et al., 2019).

LYSINE METHYLATION

Protein lysine methylation involves the transfer of up to three methyl groups to the ϵ -amine of a lysine residue. To this day, lysine methylation has been observed in both nuclear and cytoplasmic proteins and is now considered a prevalent modification in eukaryotes, prokaryotes, and archaea (Iwabata et al., 2005; Jung et al., 2008; Botting et al., 2010; Pang et al., 2010). Methylation of a lysine residue was first reported by Ambler and Rees (1959) in the flagellin protein of *Salmonella typhimurium*. These findings, further led by additional studies on histone H1, H3, and H4 lysine methylation (Couture and Trievel, 2006; Lee et al., 2010), unveiled that this post-translational modification (PTM) fine-tunes the activity of transcription factors (Yang et al., 2009), participates in the assembly of multi-subunit complexes (Zhang et al., 2005; Donlin et al., 2012), and contributes to the structural organization of chromosomes (Lanouette et al., 2014).

HISTONE LYSINE METHYLATION; WHEN PLANTS PROVIDE THE FIRST HINT

Initially reported by Allfrey et al. (1964), the field of histone lysine methylation grew exponentially in the early 2000 after the identification that the Large Subunit MethylTransferase (LSMT) can methylate lysine 14 of Ribulose-1,5-bisphosphate carboxylase/oxygenase (Ying et al., 1999). Following this seminal discovery, the group of Thomas Jenuwein reported the methylation of Lys-9 on histone H3 by the SUV3/9 family of methyltransferases (MTs; Rea et al., 2000). During the same period, using basic alignment tools, several groups identified evolutionary conserved motifs (GXG, YXG, NHXCXPN) found in a wide range of evolutionary conserved proteins (Jenuwein, 2001). Given the enrichment of these motifs in proteins including Suppressor of variegation, Enhancer of zeste, and Trithorax (SET) (Jenuwein, 2001), these enzymes were coined as SET domain lysine MTs. However, over the years, few notable cases of histone lysine MTs, such as Dot1 and PR domain MTs (PRDM), were reported to lack a SET domain. Therefore, the nomenclature for these enzymes was changed to lysine (K) MT (Allis et al., 2007). Since their discoveries, these enzymes have been shown to site-specifically methylate histone and non-histone substrates and are now recognized as critical regulators of chromatin structure and other cellular functions (Lanouette et al., 2014). They are extremely specific and, in most cases, have the ability to recognize a single lysine side chain on a single protein (Lanouette et al., 2014).

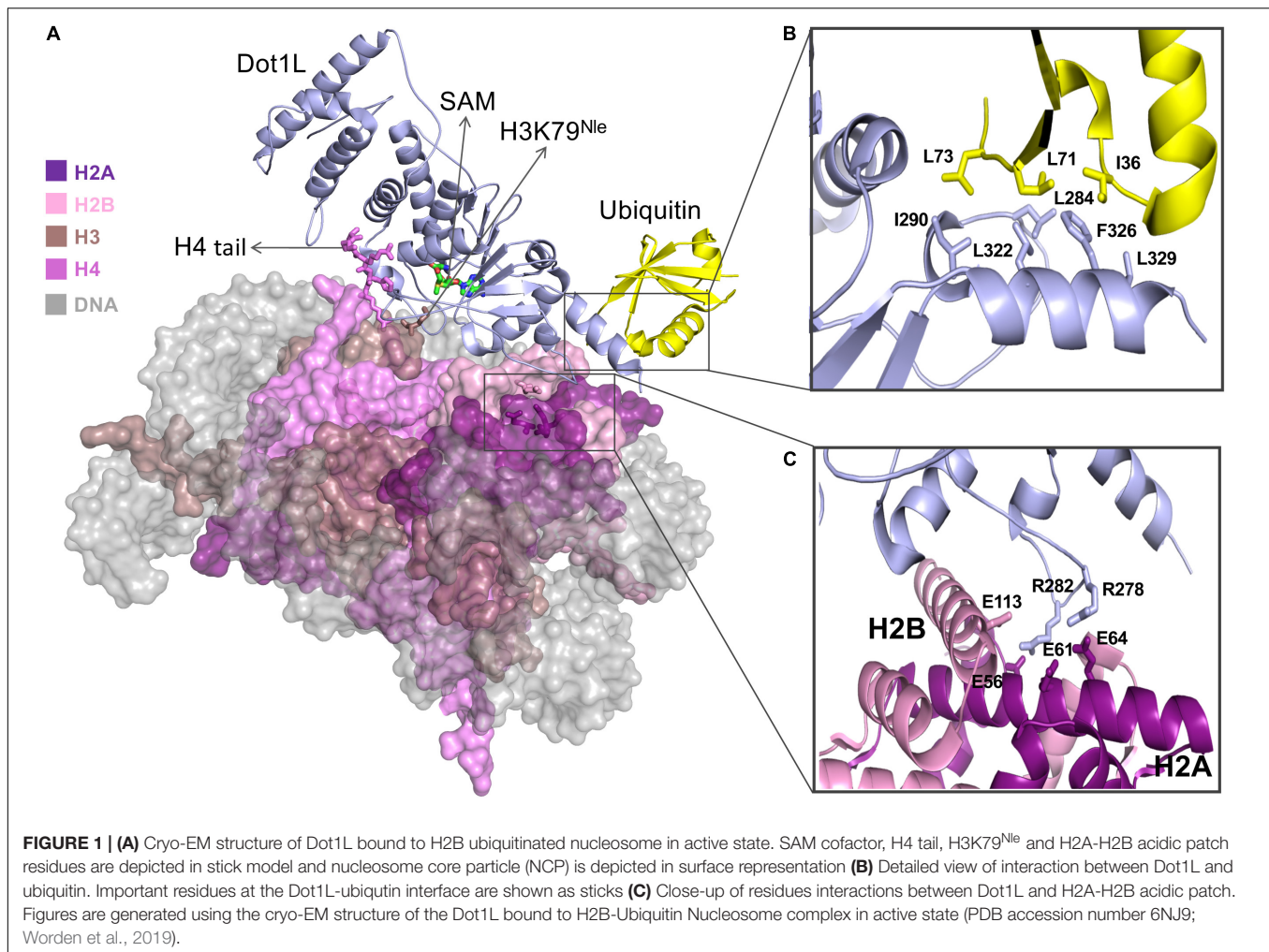
DIFFERENT MECHANISMS OF HISTONE RECOGNITION AND METHYLATION BY SET DOMAIN HKMTs

Despite being evolutionary conserved, SET domain HKMTs can be separated into at least two different categories.

This classification arises from many studies showing that HKMTs display divergence in their catalytic properties when homogeneously purified. For example, the histone H3 K36 MT SETD2 methylates, with the same catalytic efficiency (Eram et al., 2015), a peptide, the full-length histone H3 or the NCP. Conversely, other HKMTs such as ATXR5/6, EZH2, and SET8 preferentially methylate the NCP (Nishioka et al., 2002; Kirmizis et al., 2004; Margueron et al., 2008; Qiao et al., 2011). These observations suggest that this subgroup of SET domain HKMTs harbor unique structural determinants able to bind DNA. Moreover, the ubiquitination of the nucleosome or chromatin template creates better substrates for Dot1 and SET1 enzymes, respectively. Recently, several cryo-EM structures unraveled the intricacies underlying the recognition of the nucleosome by the EZH2 complex and the ubiquitinated form of the nucleosome by Dot1 and members of the SET1 family of MTs. Below, we will review the critical observations reported in these papers.

STRUCTURAL INSIGHTS INTO THE RECOGNITION OF H2BUB NUCLEOSOME BY DOT1L

Initially identified in a genetic screen to discover genes conferring defects in telomeric silencing (Singer et al., 1998; Nguyen and Zhang, 2011), disruptor of telomeric silencing-1 (Dot1) remained, for several years, the only non-SET domain histone lysine MTs. Biochemical characterization of Dot1 revealed that the enzyme mono-, di-, or tri-methylate H3K79, a modification initially linked to transcriptional regulation and DNA damage response (Nguyen and Zhang, 2011). Evolutionary conserved (Feng et al., 2002; Vlaming and van Leeuwen, 2016), human Dot1L is composed of 1537 highly conserved residues. The catalytic site is located on the N-terminus of the protein while its C-terminal extension interacts with proteins that direct Dot1L to specific genomic loci (Kuntimaddi et al., 2015; Worden et al., 2019). Initial biochemical characterization of Dot1L revealed that the MT activity of Dot1L depends on two critical factors. First, Dot1L prefers to methylate H3K79 in the context of the nucleosome (Feng et al., 2002; McGinty et al., 2008). Second, mono-ubiquitination of histone H2B on lysine 120 (Briggs et al., 2002; Ng et al., 2002; McGinty et al., 2008) (H2BK120ub) greatly enhances H3K79 methylation. Initial model showing that H2BK120 and H3K79 are closely juxtaposed on the same solvent-exposed surface of the mono-nucleosome (McGinty et al., 2008; Wood et al., 2018; Zhang and Kutateladze, 2019) lend further credence to that model. However, despite important structural insights provided by the crystal structure of Dot1L catalytic domain (Min et al., 2003), the molecular underpinnings underlying the positive impact of H2B ubiquitination on K79 methylation by Dot1L remained unexplained. Recently, several structures provided insights into the various steps linked to Dot1L binding to (Anderson et al., 2019; Worden et al., 2019; Yao et al., 2019), methylation of (Worden et al., 2019), and disengagement from (Valencia-Sanchez et al., 2019) the



nucleosomes. Three steps referred to as poised, active, and post-catalysis states.

DOT1L RECOGNIZES H2A-H2B ACIDIC PATCH IN THE NUCLEOSOME VIA AN ARGININE ANCHOR

Initial biochemical studies revealed that Dot1L preferentially methylates K79 when histone H3 is embedded in the NCP. The cryo-EM structures of Dot1L show that the C-terminal region of Dot1L contacts ubiquitin and the acidic patch of H2A-H2B (**Figure 1**). In the C-terminal region of Dot1L, a long nucleosome-interacting loop, which connects two parallel β -strands, makes contacts with the acidic patch on the nucleosome (Valencia-Sanchez et al., 2019; Worden et al., 2019). More specifically, this loop contains two evolutionary conserved arginine residues (Arg278 and Arg282) that recognize the H2A-H2B acidic patch on the nucleosome (Anderson et al., 2019; Jang et al., 2019; Valencia-Sanchez et al., 2019; Worden et al., 2019; Yao et al., 2019) (**Figure 1**). Interestingly, these structures show that, akin to SIR3 (Armache et al.,

2011), latency-associated nuclear antigen (Barbera et al., 2006), RCC1 (Makde et al., 2010), PRC1 Ubiquitylation Module (McGinty et al., 2014), Dot1L uses arginine anchors to engage the H2A-H2B acidic patch located on the surface of the nucleosomal disk.

The active site of Dot1L, consisting of an S-adenosyl-L-methionine (SAM) binding pocket and a lysine-binding channel, is positioned above H3K79 in the cryo-EM structure of the DOT1L-H2BK120Ub nucleosome complex (poised state—see below). Three loops of Dot1L form the lysine-binding channel that connects the side chain of H3K79 to the methyl donor SAM. Within these loops, several aromatic and hydrophobic residues surround the entrance of the channel and make direct contact with H3 residues adjacent to K79 (Yao et al., 2019). Within the complex, the histone H4 tail sits on $\alpha 2$ helix of histone H3 and extends to the N-terminal region of Dot1L and its active site to mediate extensive electrostatic and hydrophobic interactions with the MT. The importance of this network of interaction is underscored by mutational studies showing that substitution of histone H4 residues negatively impact the methylation of K79 by Dot1L (Yao et al., 2019).

DOT1L RECOGNIZES H2B UBIQUITIN VIA ITS HYDROPHOBIC C-TERMINAL HELIX

The cryo-EM structure of Dot1L–H2BK120Ub nucleosome complex reveals that Dot1L extensively interacts with core histones on the disk-face of nucleosome with its C-terminal region sandwiched between ubiquitin and the histone H2A–H2B dimer (**Figure 1**). The direct association of Dot1L with the H2BK120-conjugated ubiquitin extends the recognition interface between Dot1L and the surface of the NCP. Docking of the Dot1L–H2BK120Ub nucleosome complex cryo-EM structure with the cryo-EM structure of Dot1L in complex with an unmodified nucleosome complex shows a good fit of the Dot1L–H2BK120Ub nucleosome complex structure with the Dot1L-unmodified nucleosome complex, indicating that mono-ubiquitination of H2BK120 does not change the overall location of Dot1L on the surface of the nucleosome (Yao et al., 2019). The structures show the proximity of H2B-ubiquitin and the C-terminal helix of the Dot1L catalytic domain. A hydrophobic patch on ubiquitin lies near several hydrophobic residues located on an alpha helical region of Dot1L. More specifically, an area surrounding Ile36 on ubiquitin stacks on a hydrophobic patch surrounding Phe326 on Dot1L (**Figure 1**). The importance of these interactions was confirmed by mutational studies followed by histone MT assays which showed that substitution of these hydrophobic residues impairs H3K79 methylation activity of Dot1L toward ubiquitinated nucleosome but has a minor impact on the ability of Dot1L to methylate the unmodified nucleosome (Anderson et al., 2019; Jang et al., 2019; Valencia-Sanchez et al., 2019; Worden et al., 2019; Yao et al., 2019).

CRYO-EM STUDIES OF DOT1L UNRAVELS THREE STATES

Comparative analysis of Dot1L structures bound to the ubiquitinated form of the nucleosome revealed three structurally distinct forms of the complex. In the first form, also referred to as the poised state, Dot1L is positioned above histone H3K79. In this conformation, Dot1L makes contacts with ubiquitin and adjacent regions of H3K79 (Yao et al., 2019) as well as uses its arginine residues to bind to the NCP acidic patch. The observation that the catalytic site of Dot1L is separated from H3K79 indicates that Dot1L and/or the nucleosome must undergo conformational rearrangement from a poised to an active state to enable methylation (Anderson et al., 2019; Jang et al., 2019; Valencia-Sanchez et al., 2019; Worden et al., 2019; Yao et al., 2019). To trap the active state, the Cryo-EM structure of Dot1L was solved in complex with a modified ubNCP wherein K79 on histone H3 is replaced by Norleucine (Nle) (**Figure 1A**) (Worden et al., 2019; Zhang and Kutateladze, 2019); a non-native amino acid that increases the affinity of lysine MTs for their substrates in a cofactor-dependent manner (Brown et al., 2014; Jayaram et al., 2016).

Trapping the active state of the complex enabled the following observations. First, ubiquitin on H2BK120 notably restricts the orientation of Dot1L in the complex, forcing the active site of Dot1L to face the nucleosome. The contact between Dot1L and the H2A–H2B acidic patch further limits Dot1L's motion, positioning Dot1L in a catalytically competent orientation. In both active and poised state complexes, Dot1L C-terminus contacts ubiquitin and the nucleosome acidic patch, anchoring Dot1L to one edge of the nucleosome and therefore providing a pivot point about which Dot1L can rotate. The active state is further stabilized by an interaction between the histone H4 tail and a groove located in the N-terminal region of Dot1L, a region situated ~ 5 Å away from the pivot contact point, but brings another N-terminal part of Dot1L closer to the nucleosome surface. Compared to the poised state, the active state of Dot1L is rotated clockwise by $\sim 20^\circ$ around the ubiquitin and pivots down toward the nucleosome face by 25 Å (Worden et al., 2019; Zhang and Kutateladze, 2019). Interestingly, the side chain of K79 of histone H3 in the poised state complex is inaccessible for catalysis, lying parallel to the lateral surface of the nucleosomal histone core. However, in the active state, a conformational change of K79^{Nle} and its neighboring residues reorients both its backbone and side chain by $\sim 90^\circ$. This movement exposes K79 ϵ -amine to the solvent and enables its insertion into Dot1L active site (Worden et al., 2019; Zhang and Kutateladze, 2019).

The post-catalysis state was determined in the presence of S-adenosyl homocysteine and, as evidenced by mass spectrometry, H3K79 mono- and di-methylated NCP (Valencia-Sanchez et al., 2019). In this conformation, the distance between Dot1L active site and H3K79 is approximately 22 Å and unlike the poised state structure (Anderson et al., 2019; Jang et al., 2019; Yao et al., 2019), the post-catalysis state of Dot1L maintains interactions with the histone H4 tail (Valencia-Sanchez et al., 2019). Overall, the post-catalysis structure shows that Dot1L establishes multivalent interactions on the surface of the nucleosome including histone H4 tail and H2A–H2B acidic patch in addition to ubiquitin.

However, distortion of the cryo-EM density map of Dot1L's C-terminal helix suggests that motion at this site and near the acidic patch is reduced by ubiquitin. This facilitates Dot1L to carry mono-methylation, and even di- and tri-methylation of H3K79 irrespective of H2B ubiquitination. Collectively, cryo-EM structures of Dot1L in complex with ubiquitinated nucleosome complemented with biochemical experiments provided critical insights into the molecular mechanism of Dot1L-mediated methylation of lysine 79 in histone H3 and explained its crosstalk with histone H2B ubiquitination.

RECOGNITION OF H2B UBIQUITINATED NCP BY COMPASS

Initially identified in yeast, the complex associated with SET1 (COMPASS) is formed of several regulatory subunits including WDR5, RbBP5, Ash2L, DPY-30, CFP1, BIG1, as well as the catalytic unit SET1 (Miller et al., 2001). Each subunit plays important roles in the biology of SET1 and contributes, to

various extents, to the H3K4 MT activity of the complex. Owing to its link to various aggressive forms of cancers (Ford and Dingwall, 2015; Rao and Dou, 2015), several groups elucidated the crystal structure of several subunits including WDR5 (Patel et al., 2008; Dharmarajan et al., 2012; Zhang et al., 2012), RbBP5 (Mittal et al., 2018; Han et al., 2019), Ash2L (Chen et al., 2011, 2012; Sarvan et al., 2011; Zhang et al., 2015), the catalytic domain of SET1 (or its homologs) (Malumbres et al., 1997; Li et al., 2016), as well as Cfp1 (Xu et al., 2011; He et al., 2019; Yang et al., 2020). While these structures provided critical insights into the molecular underpinnings controlling the formation of COMPASS, they did not capture the entire spectrum of interactions contributing to the assembly of COMPASS. The first glimpse at COMPASS assembly was unraveled by the cryo-EM structure of budding yeast COMPASS (Qu et al., 2018) and the crystal structure of the SET1 catalytic module (Hsu et al., 2018). The cryo-EM structure shows that COMPASS assembles in a Y-shaped conformation in which WDR5 and RbBP5 (Cps30 and Cps50) β -propeller domains form the upper tips of COMPASS. Cfp1 (Cps40) connects these propellers, while Ash2L (Cps60) and Dpy-30 (Cps25) form the base of the complex. The catalytic domain of SET1 is found at the junction of the Y-shaped complex and makes contacts with every subunit, except for Dpy-30 (Qu et al., 2018) (**Figure 2A**). Interestingly, the cryo-EM structure nicely explains the modest stimulatory functions of Dpy-30 on the MT activity of SET1 on peptides when the complex is assembled with purified components (Haddad et al., 2018). Clustering of the particles revealed two conformationally distinct complexes, suggesting that COMPASS is a structurally dynamic complex that can exist in at least two conformers likely helping COMPASS to adapt to the structurally dynamic environment of chromatin (Maeshima et al., 2019).

H3K4 methylation by COMPASS is stimulated when the nucleosome is ubiquitinated on H2B (Sun and Allis, 2002; Kim et al., 2013; Holt et al., 2015). Recently, several papers documented the structural details controlling the recognition of the ubiquitinated form of the nucleosome. These structures show that COMPASS recognizes two parts of the nucleosome. On the one hand, COMPASS binds the surface of the NCP disk and the first eight residues of histone H3. Except for WDR5 and Dpy-30, all the other subunits directly contact the histone proteins, ubiquitin, and/or the nucleosomal DNA (Hsu et al., 2019). COMPASS engages both the ubiquitinated and non-ubiquitinated nucleosomes in similar fashions. However, in the presence of histone H2B ubiquitination, RbBP5 and SET1 make additional contacts with the ubiquitin moiety (**Figures 2B–G**). The SET1 catalytic domain packs against the H2A α 2 helix using two points of contact. First, a region preceding the SET1 catalytic domain contacts three residues on H2A. The same protein also surrounds the C-terminus of the same helix on H2A using a cluster of five evolutionarily conserved hydrophobic residues. The presence of this cluster in other members of the KMT2 family of enzymes points to a model wherein the catalytic domain of these enzymes may bind similarly to the surface of the nucleosome. Correlatively,

mutations of these residues result in a loss of H3K4 di- and tri-methylation (Nakanishi et al., 2008). In the presence of ubiquitinated H2B, a region immediately preceding the catalytic domain of SET1, which includes an Arginine Rich Motif (ARM), and a fraction of its SET domain form a coil binding to a pocket formed by Ile36, the β 1– β 2 loop, and the tail of ubiquitin (**Figure 2B**). The ARM motif, which is sandwiched between COMPASS subunits and uNCP, is located near the acidic patch created by the H2A–H2B interface (Nakanishi et al., 2008; Kim et al., 2013) (**Figures 2D,E**). These observations are supported by biochemical and *in vivo* data showing that mutation of the residues forming the ARM motif negatively impacts H3K4 methylation (Kim et al., 2013). Altogether, these observations indicate that this motif serves as an important link between H2B ubiquitination and H3K4 methylation.

Several hydrophobic residues located on both the N- and C-termini of RbBP5 interact with a hydrophobic patch on ubiquitin (**Figure 2F**). The β -propeller domain of RbBP5 also makes polar contacts with the C-terminus of ubiquitin (**Figure 2G**). In addition to binding to ubiquitin, RbBP5 directly interacts with a cleft formed by α 3 and α C of histone H2B, α 2 of H2A, as well as DNA. Mutation of the residues forming this RbBP5–NCP interface impair H3K4 methylation by COMPASS, underscoring the importance of these interactions (Hsu et al., 2019). Located in the same region of COMPASS and directly interacting with RbBP5 (Yang et al., 2020), weak but discernable structural information can be detected in a region of Cfp1 composed of positively charged residues. Based on the predicted location of these residues near the nucleosomal DNA, the cryo-EM structure suggests that Cfp1 directly binds DNA. Similarly, the Ash2L (Cps60) SPRY domain directly interacts with the phosphate backbone of the nucleosomal DNA (Hsu et al., 2019).

The cryo-EM structures of COMPASS in complex with the ubiquitinated and non-ubiquitinated nucleosomes have provided important information regarding how COMPASS engages its substrate and the structural underpinnings mediating its enzymatic activity. The findings suggest that the presence of ubiquitin may alter the dynamics of the catalytic subunit in alleviating an auto-inhibitory function of the SET1 ARM motif (Hsu et al., 2019). Furthermore, the interactions between COMPASS and uNCP appear to stabilize further the N-terminus of histone H3 in the catalytic domain. In the absence of ubiquitin, the structure presents only three H3 residues (T3, K4, and Q5) interacting with SET1, while in the presence of ubiquitin, A1 to R8 are distinguishable in the SET1 catalytic domain. This suggests that the interactions between COMPASS and ubiquitin induce conformational changes that increase the interface between the catalytic domain of SET1 and the residues flanking H3K4. Altogether, these results show that cross-talk between protein complex subunits and pre-existing modifications on the nucleosome represents a way to control H3K4 methylation (Jeon et al., 2018). Interestingly, such cross-talk has also been proposed as a mode of activation for other histone MTs, such as EZH2 (Margueron et al., 2009; Jiao and Liu, 2015; Brooun et al., 2016).

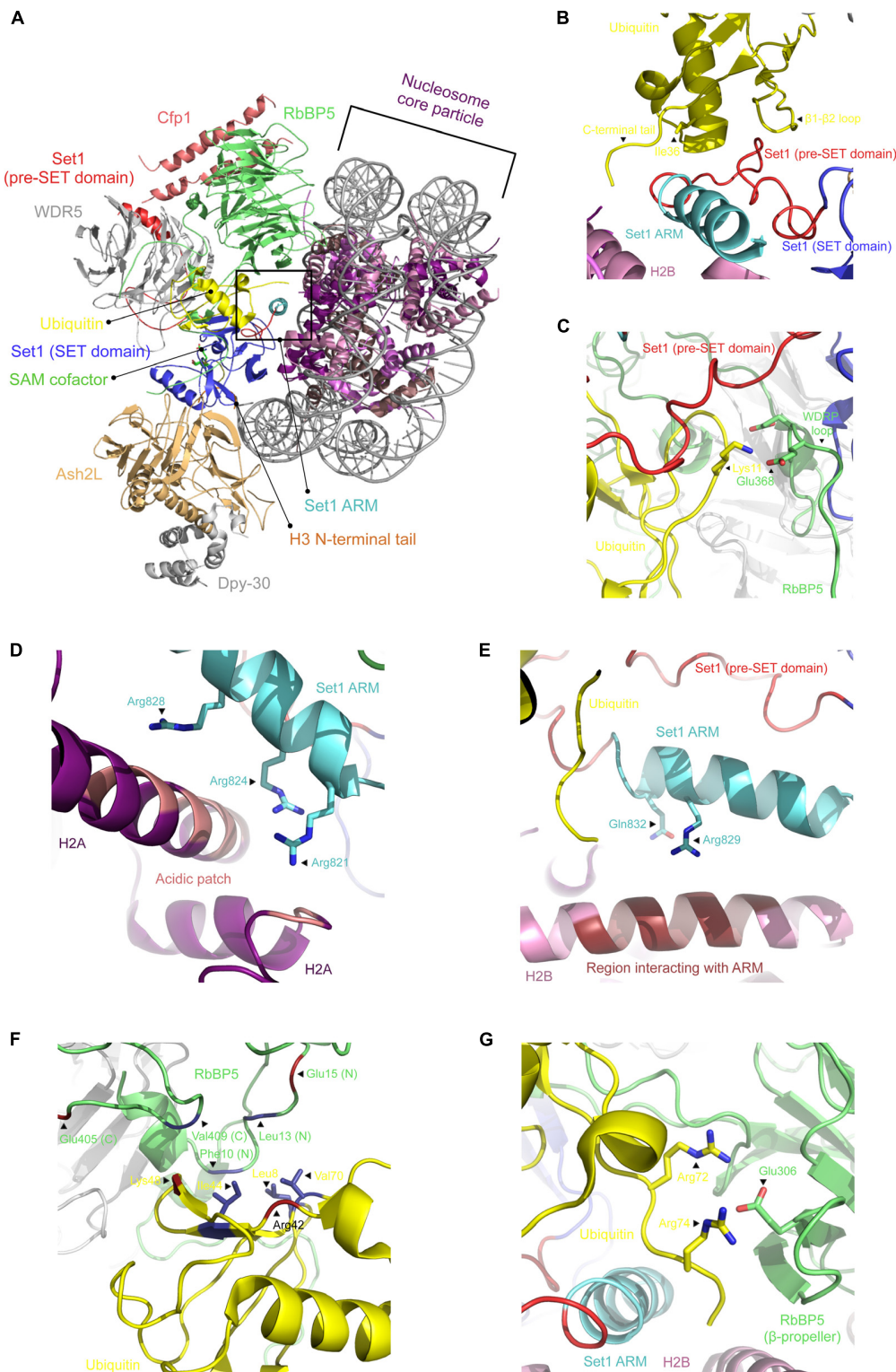


FIGURE 2 | Cryo-EM structure of COMPASS bound to the ubiquitinated nucleosome core particle. **(A)** Cartoon representation of COMPASS cryo-EM structure bound to the ubiquitinated nucleosome in which each subunit is indicated. **(B)** Zoomed view on the interactions between ubiquitin and the pre-SET domain of SET1. Shown are the polar contacts between RbBP5 WDRP loop and ubiquitin **(C)** and the cluster of positively charged residues of the SET1 ARM motif interacting with the H2A acidic patch **(D)** and histone H2B **(E)**. Cartoon representation of RbBP5 N- and C-termini that make contacts with ubiquitin's hydrophobic patch (depicted as sticks) **(F)**, as well as the contacts made between its β -propeller domain and ubiquitin C-terminal end **(G)**. All figures were prepared using the cryo-EM structure of the COMPASS catalytic module in complex with the ubiquitinated nucleosome (PDB accession number 6UH5).

EZH2 AND H3K27ME3

The Enhancer of zeste E(z) gene was discovered as an important regulatory element in maintaining suppression of homeotic gene expression such as those determining pigmentation in *Drosophila melanogaster* (Kalisch and Rasmuson, 1974; Wu et al., 1989). A subsequent study revealed that the C-terminal region of E(z) gene product, now known as the SET domain, shares homology with regions of the Trithorax (Trx) (Jones and Gelbart, 1993) and Suppressor of variegation [Su(var)] proteins. In humans, EZH2 is one of the two homologs of the fruit fly's E(z) enzyme which trimethylates H3K27 and preferentially methylates dinucleosome substrates over mononucleosomes, and the MT activity is further stimulated by the linker histone H1 (Martin et al., 2006). Local H3K27me3 is linked to suppression of targeted gene expression while this mark can spread to regulate processes such as cell differentiation and X-chromosome inactivation by negatively regulating gene expression. EZH2 SET domain is the catalytic component of Polycomb repressive complex 2 (PRC2) which also includes embryonic ectoderm development (EED), suppressor of zeste 12 (SUZ12), and Retinoblastoma Binding Protein 4 (RbBP4) as core components. In contrast to SET domain lysine MTs such as ATXR5/6 (Jacob et al., 2009), EZH2 alone is not catalytically active and minimally requires EED and the VEFS [Vrn2-Emf2-Fis2-Su(z)12] box of SUZ12 to methylate H3K27 (Cao and Zhang, 2004). Other components, namely, jumonji AT-rich interactive domain 2 (JARID2), Adipocyte Enhancer-Binding Protein 2 (AEBP2), and polycomb-like (PCL) proteins associate with and modulate PRC2 activity or its recruitment to chromatin. These include interaction with unmethylated CpG islands (Li et al., 2017), activation at *de novo* H3K27me3 nucleation sites (Oksuz et al., 2018), or determining exclusivity of PRC2 subcomplexes (Grijzenhout et al., 2016). Although earlier understanding of the relationship between PRC2 and PRC1, which monoubiquitinates K119 on H2A (Wang et al., 2004), suggested that cooperative repression by these complexes is mediated by the detection of H3K27me3 via Cbx in PRC1 (Senthilkumar and Mishra, 2009), recent evidence suggest that JARID2 also binds the H2A-K119ubiquitinated form of the NCP suggesting that cross-talk between PRC1 and PRC2 involves more than H3K27me3 and that it may not be unidirectional or in the chronological order previously described.

STRUCTURAL ANALYSIS OF PRC2 AND EZH2 ACTIVATION BY THE HOLOENZYME

The absence of EZH2 activity was elegantly explained by the crystal structure of EZH2 CXC-SET domains alone (Wu et al., 2013). The structure shows that the EZH2 substrate-binding groove is in a closed state as a result of hydrogen bonds between residues in the I-SET and post-SET regions of EZH2 likely barring the H3K27 to enter the channel. The CXC domain also appears to play an autoinhibitory role in EZH2 by pulling away from the post-SET domain, which contributes

to the formation of the cofactor binding site rendering this pocket structurally incomplete. The crystal structure of the minimal PRC2 complex revealed an extensive network of inter-domain interactions involving all domains of EZH2, EED, and VEFS(SUZ12) in such a way that EZH2 wraps around both VEFS and EED and overall holds the entire complex together while connecting the insertion domain of EED to the N-terminal region of VEFS near its SET domain (Jiao and Liu, 2015). Comparison of EZH2 and the minimal PRC2 structures reveal that interaction with EED/SUZ12 rotates the post-SET in such a way that the catalytic channel opens, the cofactor binding site formation is completed, and EZH2 is catalytically competent.

READING AND WRITING H3K27ME3 BY EZH2

A structure of the minimal PRC2 complex shows that the complex binds both a stimulating (K27me3) and a [pseudo]-substrate (K27M) H3 peptide simultaneously (Jiao and Liu, 2015). The structure shows that while the substrate H3K27M peptide interacts with the SET domain, the stimulating H3K27me3 peptide binds the β -propeller domain of EED and interacts with the SRM domain of EZH2 (Jiao and Liu, 2015). Structural analysis and enzymatic assays revealed that K27M, observed in glioblastomas, stalls PRC2 activity (Lewis et al., 2013) and spreading of K27 trimethylation due to positioning of arginine 26 in the active site which makes stronger contacts than the wildtype lysine while addition of an R26A mutation to the K27M peptide restores PRC2 MT activity. Interestingly, MT activity using wildtype substrate (H3K27) is increased by over fivefold in the presence of the H3K27me3 peptide which binds EED/SRM(EZH2) (Jani et al., 2019). Accordingly, PRC2 shows lower MT activity on mononucleosome substrates compared to di- or oligonucleosome substrates (Yuan et al., 2012). Oligo-nucleosomes reconstituted with short DNA linker (20 vs. 46 and 66 bp) are more robustly methylated by PRC2 indicating that the length of the linker DNA further controls H3K27 methylation. Incubation of PRC2 with an array of peptides collectively spanning H3 1–42 shows enhanced methylation when H3 35–42 peptide was added to the reaction including cases were oligonucleosomes were dispersed (> 20 bp linkers). Furthermore, MT assays show that presence of histone H1 positively stimulates PRC2 activity in dinucleosomes (Yuan et al., 2012) suggesting that H1-mediated chromatin compaction stimulates PRC2 activity by providing access to a stimulating H3 from a neighboring nucleosomes.

A cryo-EM structure capturing PRC2 bound to a 35 bp linked dinucleosome provides unique insights into the enzyme complex simultaneously engaging with a pseudo-substrate (K27M) nucleosome and a stimulating (K27me3) neighbor nucleosome (Poepsel et al., 2018). Interestingly, the EZH2 CXC domain makes several contacts with nucleosomal DNA, where the H3 tail extends out of the nucleosome disc (**Figure 3**). EZH2 SBD also binds DNA at the exit site of the H3 tail but

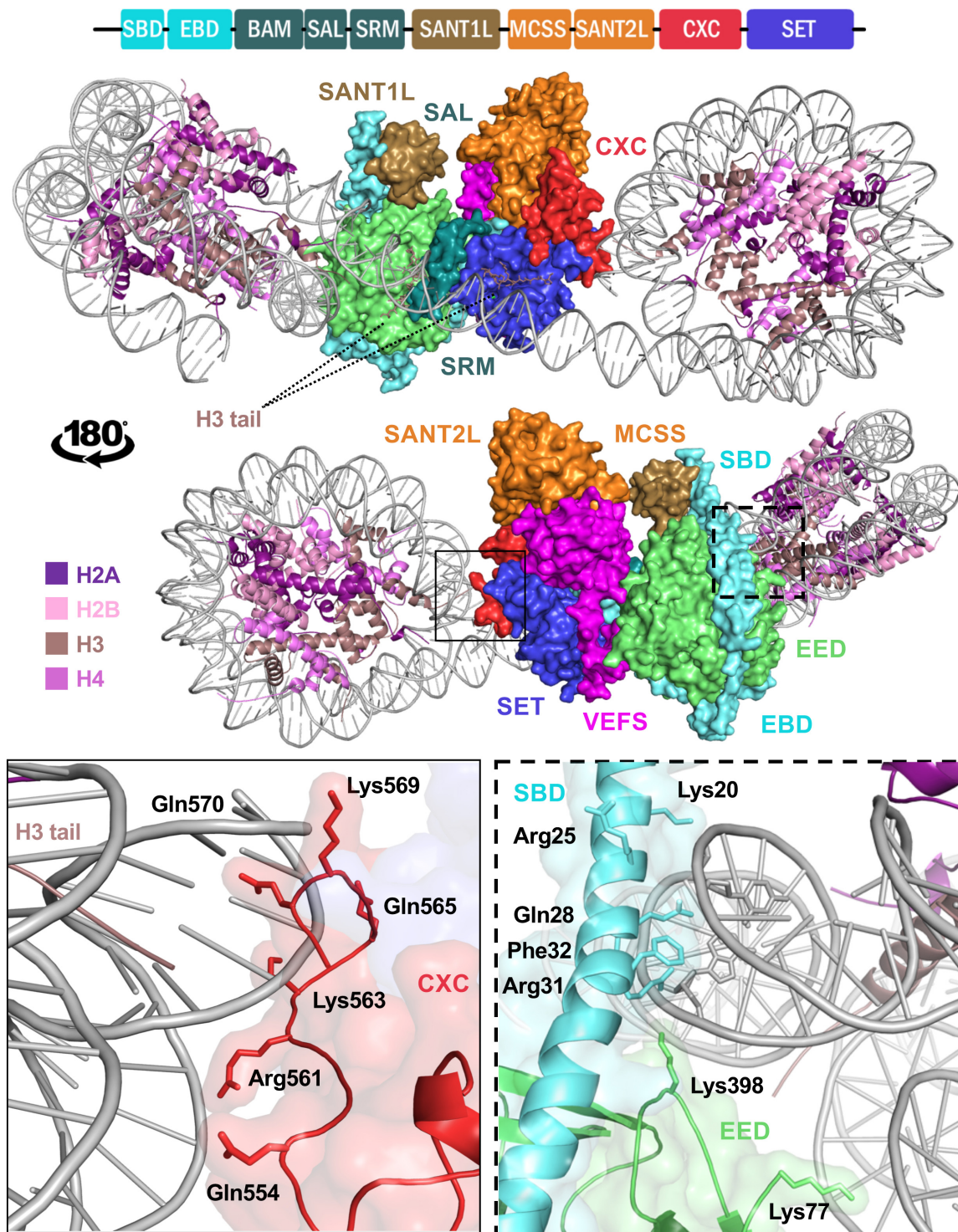


FIGURE 3 | Structure of EZH2 in a minimal PRC2 assembly in complex with an asymmetric di-nucleosome. The schematic at the top represents domain configuration of EZH2. The demonstrations show the same mPRC2:NCP complex from a front and back view. EZH2 domains in the structure are colored according to the linear schematic depiction. The VEFS domain of SUZ12 is colored as magenta and EED is represented as light green. H3 tail with K27M substitution of the substrate nucleosome is shown bound to the substrate groove of EZH2 SET domain (purple blue). The neighboring nucleosome with a modified H3 tail bearing a trimethylated lysine at the position of K27 (K27me3) is shown in a groove between EED and EZH2 SRM domain (teal). Zoomed demonstration of the framed areas on structure is shown at the bottom of the figure. Positively charged and polar residues of EZH2 CXC domain (solid frame) and EZH2 SBD/EED (dashed frame) within proximity to DNA back bone are labeled with their corresponding residue numbers. EZH2 SBD hydrophobic residue is shown in close proximity of DNA.

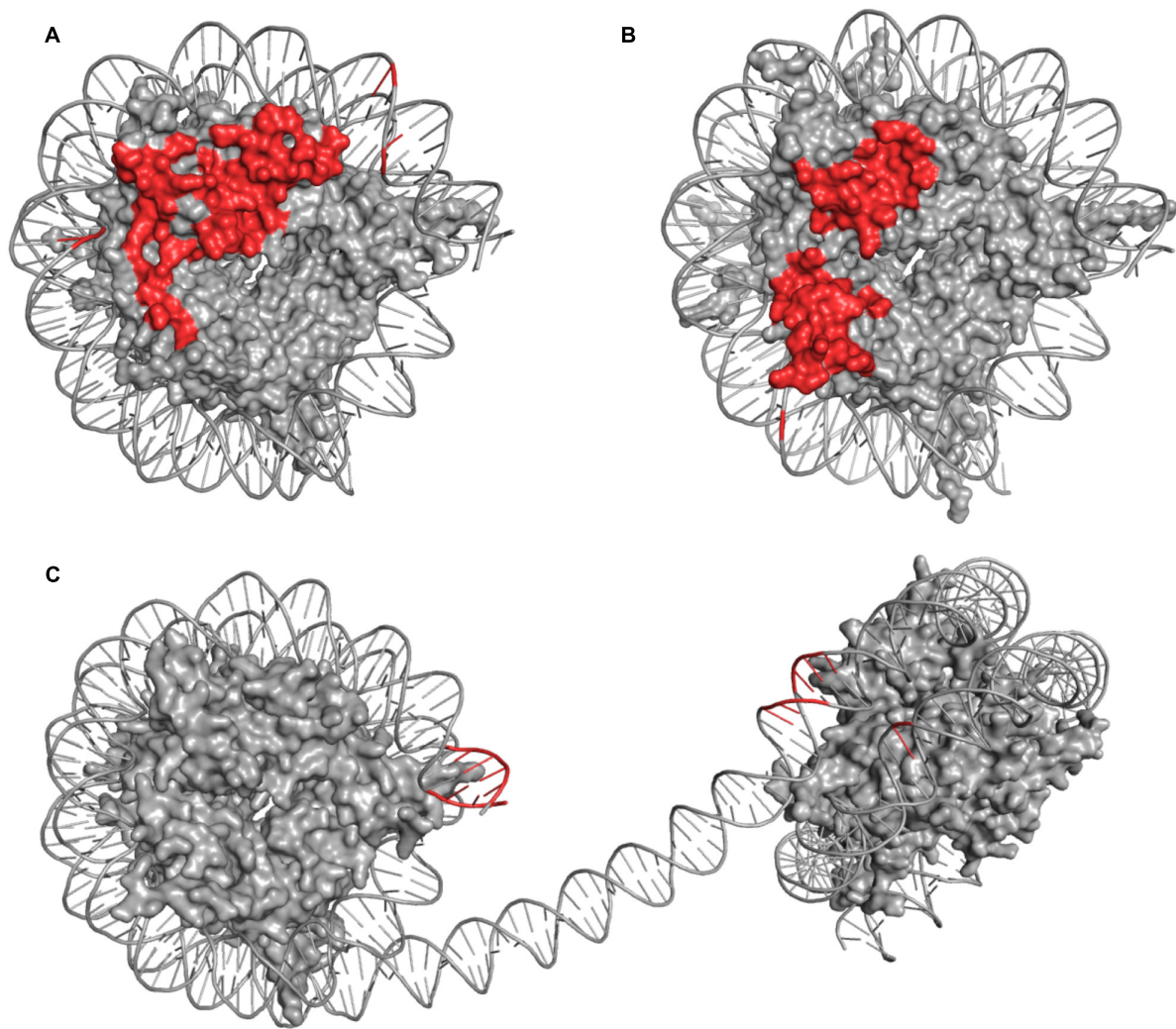


FIGURE 4 | Comparison of DOT1L (A), COMPASS (B), and mPRC2 (C) Modes of Engagement with the Nucleosome. The nucleosome discs represent NCP structures captured in the same orientations. The side-by-side comparison shows surface representation of histones inside the nucleosome disc (gray) and cartoons depiction of DNA (gray) highlighting amino acids or DNA bases which come to close contact (≤ 5 Å) with residues of the associated enzyme or enzyme complex (red). Contact points involving histone tails were omitted due to lack of structural continuity.

on the neighboring nucleosome (Figure 3). Positively charged and polar residues on the surface of CXC are nearby the DNA backbone. An additional bi-partite sequence, ⁴⁹¹RKKRKHR⁴⁹⁷, and ⁵⁰⁴RKIQLKK⁵¹⁰ in CXC are candidates for DNA interaction; however, these residues could not be modeled in the structure. Similarly, a cluster of polar residues spanning a region of the SBD likely interact with the DNA backbone while an aromatic residue in this region is oriented suitably for intercalating with DNA bases (Figure 3). The area corresponding to EED residues 70 KGKWKSKKCK79 can also potentially bind DNA; however, only residues 77–79 were resolved in the structure of which K79 comes to close contact with DNA backbone. Compared to the crystal structure of PRC2 in the absence of nucleosome, the SET, CXC, and SBD domains are the main components that undergo displacement/tilt after nucleosome binding.

DISCUSSION

Comparative analysis between COMPASS and Dot1L (Figures 4A,B) binding modes of the ubiquitinated nucleosome reveals notable similarities. Both make significant contacts with the surface of the NCP disk and touch each histone protein. Similarly, both make a limited number of contacts with DNA, with, however, differences in the location of these interactions. Dot1L binds DNA adjacent to H3/H2A near histone H3 tail exit site while COMPASS binds DNA near the exit site of histone H4 N-terminus. Also, Dot1L interacts with two distinct regions on the face of the NCP disk, while COMPASS binds a continuous surface. In stark contrast to COMPASS and Dot1L, PRC2 does not contact the surface of the NCP (Figure 4C) but makes several interactions with DNA located at the exit site of the H3 tail on the di-nucleosome.

Together, the cryo-EM structures of Dot1L in complex with the ubiquitinated nucleosome provided pivotal insights into the molecular mechanism underlying Dot1L-mediated methylation H3K79 by histone H2B ubiquitination (Jang et al., 2019; Valencia-Sanchez et al., 2019; Worden et al., 2019; Yao et al., 2019). The cryo-EM structures of COMPASS bound to H2B ubiquitinated NCP uncovered the crucial functions of COMPASS subunits in recognizing different parts of the nucleosome and further the essential functions of SET1 ARM motif in linking H2B ubiquitination and H3K4 methylation. Considering that Dot1L and COMPASS complexes are linked to leukemia, these findings may help in the design of inhibitors that could serve as effective therapeutic agents.

REFERENCES

- Allfrey, V. G., Faulkner, R., and Mirsky, A. E. (1964). Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci. U.S.A.* 51, 786–794. doi: 10.1073/pnas.51.5.786
- Allis, C. D., Berger, S. L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., et al. (2007). New nomenclature for chromatin-modifying enzymes. *Cell* 131, 633–636.
- Ambler, R. P., and Rees, M. W. (1959). Epsilon-N-Methyl-lysine in bacterial flagellar protein. *Nature* 184, 56–57. doi: 10.1038/184056b0
- Anderson, C. J., Baird, M. R., Hsu, A., Barbour, E. H., Koyama, Y., Borgnia, M. J., et al. (2019). Structural basis for recognition of ubiquitylated nucleosome by Dot1L methyltransferase. *Cell Rep.* 26, 1681–1690.e5. doi: 10.1016/j.celrep.2019.01.058
- Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E., and Moudrianakis, E. N. (1991). The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proc. Natl. Acad. Sci. U.S.A.* 88, 10148–10152. doi: 10.1073/pnas.88.22.10148
- Armache, K. J., Garlick, J. D., Canzio, D., Narlikar, G. J., and Kingston, R. E. (2011). Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. *Science* 334, 977–982. doi: 10.1126/science.1210915
- Barbera, A. J., Chodaparambil, J. V., Kelley-Clarke, B., Joukov, V., Walter, J. C., Luger, K., et al. (2006). The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science* 311, 856–861. doi: 10.1126/science.1120541
- Botting, C. H., Talbot, P., Paytubi, S., and White, M. F. (2010). Extensive lysine methylation in hyperthermophilic crenarchaea: potential implications for protein stability and recombinant enzymes. *Archaea* 2010:106341.
- Briggs, S. D., Xiao, T., Sun, Z. W., Caldwell, J. A., Shabanowitz, J., Hunt, D. F., et al. (2002). Gene silencing: trans-histone regulatory pathway in chromatin. *Nature* 418:498. doi: 10.1038/nature00970
- Brooun, A., Gajiwala, K. S., Deng, Y. L., Liu, W., Bolanos, B., Bingham, P., et al. (2016). Polycomb repressive complex 2 structure with inhibitor reveals a mechanism of activation and drug resistance. *Nat. Commun.* 7:11384.
- Brown, Z. Z., Muller, M. M., Jain, S. U., Allis, C. D., Lewis, P. W., and Muir, T. W. (2014). Strategy for “detoxification” of a cancer-derived histone mutant based on mapping its interaction with the methyltransferase PRC2. *J. Am. Chem. Soc.* 136, 13498–13501. doi: 10.1021/ja5060934
- Cao, R., and Zhang, Y. (2004). SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Mol. Cell* 15, 57–67. doi: 10.1016/j.molcel.2004.06.020
- Chen, Y., Cao, F., Wan, B., Dou, Y., and Lei, M. (2012). Structure of the SPRY domain of human Ash2L and its interactions with RbBP5 and DPY30. *Cell Res.* 22, 598–602. doi: 10.1038/cr.2012.9
- Chen, Y., Wan, B., Wang, K. C., Cao, F., Yang, Y., Protacio, A., et al. (2011). Crystal structure of the N-terminal region of human Ash2L shows a winged-helix motif involved in DNA binding. *EMBO Rep.* 12, 797–803. doi: 10.1038/embor.2011.101
- Couture, J. F., and Trievel, R. C. (2006). Histone-modifying enzymes: encrypting an enigmatic epigenetic code. *Curr. Opin. Struct. Biol.* 16, 753–760. doi: 10.1016/j.sbi.2006.10.002
- Dharmarajan, V., Lee, J. H., Patel, A., Skalnik, D. G., and Cosgrove, M. S. (2012). Structural basis for WDR5 interaction (Win) motif recognition in human SET1 family histone methyltransferases. *J. Biol. Chem.* 287, 27275–27289. doi: 10.1074/jbc.M112.364125
- Donlin, L. T., Andresen, C., Just, S., Rudensky, E., Pappas, C. T., Kruger, M., et al. (2012). Smyd2 controls cytoplasmic lysine methylation of Hsp90 and myofilament organization. *Genes Dev.* 26, 114–119. doi: 10.1101/gad.177758.111
- Eram, M. S., Kuznetsova, E., Li, F., Lima-Fernandes, E., Kennedy, S., Chau, I., et al. (2015). Kinetic characterization of human histone H3 lysine 36 methyltransferases, ASH1L and SETD2. *Biochim. Biophys. Acta* 1850, 1842–1848. doi: 10.1016/j.bbagen.2015.05.013
- Feng, Q., Wang, H., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Struhl, K., et al. (2002). Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Curr. Biol.* 12, 1052–1058. doi: 10.1016/s0960-9822(02)00901-6
- Ford, D. J., and Dingwall, A. K. (2015). The cancer COMPASS: navigating the functions of MLL complexes in cancer. *Cancer Genet.* 208, 178–191. doi: 10.1016/j.cancergen.2015.01.005
- Grijzenhout, A., Godwin, J., Koseki, H., Gdula, M. R., Szumska, D., McGouran, J. F., et al. (2016). Functional analysis of AEBP2, a PRC2 Polycomb protein, reveals a Trithorax phenotype in embryonic development and in ESCs. *Development* 143, 2716–2723. doi: 10.1242/dev.123935
- Haddad, J. F., Yang, Y., Takahashi, Y. H., Joshi, M., Chaudhary, N., Woodfin, A. R., et al. (2018). Structural analysis of the Ash2L/Dpy-30 complex reveals a heterogeneity in H3K4 methylation. *Structure* 26, 1594–1603.e4. doi: 10.1016/j.str.2018.08.004
- Han, J., Li, T., Li, Y., Li, M., Wang, X., Peng, C., et al. (2019). The internal interaction in RBBP5 regulates assembly and activity of MLL1 methyltransferase complex. *Nucleic Acids Res.* 47, 10426–10438. doi: 10.1093/nar/gkz819
- He, C., Liu, N., Xie, D., Liu, Y., Xiao, Y., and Li, F. (2019). Structural basis for histone H3K4me3 recognition by the N-terminal domain of the PHD finger protein Spp1. *Biochem. J.* 476, 1957–1973. doi: 10.1042/bcj20190091
- Holt, M. T., David, Y., Pollock, S., Tang, Z., Jeon, J., Kim, J., et al. (2015). Identification of a functional hotspot on ubiquitin required for stimulation of methyltransferase activity on chromatin. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10365–10370. doi: 10.1073/pnas.1504483112
- Hsu, P. L., Li, H., Lau, H. T., Leonen, C., Dhall, A., Ong, S. E., et al. (2018). Crystal structure of the COMPASS H3K4 methyltransferase catalytic module. *Cell* 174, 1106–1116.e9. doi: 10.1016/j.cell.2018.06.038
- Hsu, P. L., Shi, H., Leonen, C., Kang, J., Chatterjee, C., and Zheng, N. (2019). Structural basis of H2B ubiquitination-dependent H3K4 methylation by COMPASS. *Mol. Cell* 76, 712–723.e4. doi: 10.1016/j.molcel.2019.10.013
- Iwabata, H., Yoshida, M., and Komatsu, Y. (2005). Proteomic analysis of organ-specific post-translational lysine-acetylation and -methylation in mice by use of anti-acetyllysine and -methyllysine mouse monoclonal antibodies. *Proteomics* 5, 4653–4664. doi: 10.1002/pmic.200500042
- Jacob, Y., Feng, S., LeBlanc, C. A., Bernatavichute, Y. V., Stroud, H., Cokus, S., et al. (2009). ATXR5 and ATXR6 are H3K27 monomethyltransferases required

AUTHOR CONTRIBUTIONS

All authors equally contributed to the preparation of the manuscript. The sections on Dot1L, COMPASS, and EZH2 were written by MJ, AJ, and HD, respectively. HD also prepared Figure 4.

FUNDING

This work was supported by Canadian Institutes of Health Research grants (PJT-148869 and PJT-148533).

- for chromatin structure and gene silencing. *Nat. Struct. Mol. Biol.* 16, 763–768. doi: 10.1038/nsmb.1611
- Jang, S., Kang, C., Yang, H. S., Jung, T., Hebert, H., Chung, K. Y., et al. (2019). Structural basis of recognition and destabilization of the histone H2B ubiquitinated nucleosome by the DOT1L histone H3 Lys79 methyltransferase. *Genes Dev.* 33, 620–625. doi: 10.1101/gad.323790.118
- Jani, K. S., Jain, S. U., Ge, E. J., Diehl, K. L., Lundgren, S. M., Müller, M. M., et al. (2019). Histone H3 tail binds a unique sensing pocket in EZH2 to activate the PRC2 methyltransferase. *Proc. Natl. Acad. Sci. U.S.A.* 116, 8295–8300. doi: 10.1073/pnas.1819029116
- Jayaram, H., Hoelper, D., Jain, S. U., Cantone, N., Lundgren, S. M., Poy, F., et al. (2016). S-adenosyl methionine is necessary for inhibition of the methyltransferase G9a by the lysine 9 to methionine mutation on histone H3. *Proc. Natl. Acad. Sci. U.S.A.* 113, 6182–6187. doi: 10.1073/pnas.1605523113
- Jenuwein, T. (2001). Re-SET-ting heterochromatin by histone methyltransferases. *Trends Cell Biol.* 11, 266–273. doi: 10.1016/s0962-8924(01)02001-3
- Jeon, J., McGinty, R. K., Muir, T. W., Kim, J. A., and Kim, J. (2018). Crosstalk among Set1 complex subunits involved in H2B ubiquitylation-dependent H3K4 methylation. *Nucleic Acids Res.* 46, 11129–11143. doi: 10.1093/nar/gky920
- Jiao, L., and Liu, X. (2015). Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. *Science* 350:aac4383. doi: 10.1126/science.aac4383
- Jones, R. S., and Gelbart, W. M. (1993). The Drosophila Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax. *Mol. Cell Biol.* 13, 6357–6366. doi: 10.1128/mcb.13.10.6357
- Jung, S. Y., Li, Y., Wang, Y., Chen, Y., Zhao, Y., and Qin, J. (2008). Complications in the assignment of 14 and 28 Da mass shift detected by mass spectrometry as in vivo methylation from endogenous proteins. *Anal. Chem.* 80, 1721–1729. doi: 10.1021/ac7021025
- Kalisch, W.-E., and Rasmuson, B. (1974). Changes of zeste phenotype induced by autosomal mutations in *Drosophila melanogaster*. *Hereditas* 78, 97–103. doi: 10.1111/j.1601-5223.1974.tb01432.x
- Kim, J., Kim, J. A., McGinty, R. K., Nguyen, U. T., Muir, T. W., Allis, C. D., et al. (2013). The n-SET domain of Set1 regulates H2B ubiquitylation-dependent H3K4 methylation. *Mol. Cell* 49, 1121–1133. doi: 10.1016/j.molcel.2013.01.034
- Kirmizis, A., Bartley, S. M., Kuzmichev, A., Margueron, R., Reinberg, D., Green, R., et al. (2004). Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. *Genes Dev.* 18, 1592–1605. doi: 10.1101/gad.1200204
- Koyama, M., and Kurumizaka, H. (2018). Structural diversity of the nucleosome. *J. Biochem.* 163, 85–95. doi: 10.1093/jb/mvx081
- Kuntimaddi, A., Achille, N. J., Thorpe, J., Lokken, A. A., Singh, R., Hemenway, C. S., et al. (2015). Degree of recruitment of DOT1L to MLL-AF9 defines level of H3K79 Di- and tri-methylation on target genes and transformation potential. *Cell Rep.* 11, 808–820. doi: 10.1016/j.celrep.2015.04.004
- Lanouette, S., Mongeon, V., Figeys, D., and Couture, J. F. (2014). The functional diversity of protein lysine methylation. *Mol. Syst. Biol.* 10:724. doi: 10.1002/msb.134974
- Lee, J. S., Smith, E., and Shilatifard, A. (2010). The language of histone crosstalk. *Cell* 142, 682–685. doi: 10.1016/j.cell.2010.08.011
- Lewis, P. W., Müller, M. M., Koletsky, M. S., Cordero, F., Lin, S., Banaszyński, L. A., et al. (2013). Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science* 340, 857–861. doi: 10.1126/science.1232245
- Li, H., Liefke, R., Jiang, J., Kurland, J. V., Tian, W., Deng, P., et al. (2017). Polycomb-like proteins link the PRC2 complex to CpG islands. *Nature* 549, 287–291. doi: 10.1038/nature23881
- Li, Y., Han, J., Zhang, Y., Cao, F., Liu, Z., Li, S., et al. (2016). Structural basis for activity regulation of MLL family methyltransferases. *Nature* 530, 447–452. doi: 10.1038/nature16952
- Luger, K., Mader, 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.
- Maeshima, K., Ide, S., and Babokhov, M. (2019). Dynamic chromatin organization without the 30-nm fiber. *Curr. Opin. Cell Biol.* 58, 95–104. doi: 10.1016/j.ccb.2019.02.003
- Makde, R. D., England, J. R., Yennawar, H. P., and Tan, S. (2010). Structure of RCC1 chromatin factor bound to the nucleosome core particle. *Nature* 467, 562–566. doi: 10.1038/nature09321
- Malumbres, M., Mangues, R., Ferrer, N., Lu, S., and Pellicer, A. (1997). Isolation of high molecular weight DNA for reliable genotyping of transgenic mice. *Biotechniques* 22, 1114–1119. doi: 10.2144/97226st03
- Margueron, R., Justin, N., Ohno, K., Sharpe, M. L., Son, J., Drury, W. J., et al. (2009). Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* 461, 762–767. doi: 10.1038/nature08398
- Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C. L., et al. (2008). Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Mol. Cell* 32, 503–518. doi: 10.1016/j.molcel.2008.11.004
- Martin, C., Cao, R., and Zhang, Y. (2006). Substrate preferences of the EZH2 histone methyltransferase complex. *J. Biol. Chem.* 281, 8365–8370. doi: 10.1074/jbc.m513425200
- McGinty, R. K., Henrici, R. C., and Tan, S. (2014). Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. *Nature* 514, 591–596. doi: 10.1038/nature13890
- McGinty, R. K., Kim, J., Chatterjee, C., Roeder, R. G., and Muir, T. W. (2008). Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* 453, 812–816. doi: 10.1038/nature06906
- Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., et al. (2001). COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl. Acad. Sci. U.S.A.* 98, 12902–12907. doi: 10.1073/pnas.231473398
- Min, J., Feng, Q., Li, Z., Zhang, Y., and Xu, R. M. (2003). Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. *Cell* 112, 711–723. doi: 10.1016/s0092-8674(03)00114-4
- Mittal, A., Hobor, F., Zhang, Y., Martin, S. R., Gamblin, S. J., Ramos, A., et al. (2018). The structure of the RbBP5 beta-propeller domain reveals a surface with potential nucleic acid binding sites. *Nucleic Acids Res.* 46, 3802–3812. doi: 10.1093/nar/gky199
- Nakanishi, S., Sanderson, B. W., Delventhal, K. M., Bradford, W. D., Staehling-Hampton, K., and Shilatifard, A. (2008). A comprehensive library of histone mutants identifies nucleosomal residues required for H3K4 methylation. *Nat. Struct. Mol. Biol.* 15, 881–888. doi: 10.1038/nsmb.1454
- Ng, H. H., Feng, Q., Wang, H., Erdjument-Bromage, H., Tempst, P., Zhang, Y., et al. (2002). Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. *Genes Dev.* 16, 1518–1527. doi: 10.1101/gad.1001502
- Nguyen, A. T., and Zhang, Y. (2011). The diverse functions of Dot1 and H3K79 methylation. *Genes Dev.* 25, 1345–1358. doi: 10.1101/gad.205781
- Nishioka, K., Rice, J. C., Sarma, K., Erdjument-Bromage, H., Werner, J., Wang, Y., et al. (2002). PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. *Mol. Cell* 9, 1201–1213. doi: 10.1016/s1097-2765(02)00548-8
- Noll, M. (1977). DNA folding in the nucleosome. *J. Mol. Biol.* 116, 49–71. doi: 10.1016/0022-2836(77)90118-8
- Oksuz, O., Narendra, V., Lee, C.-H., Descostes, N., LeRoy, G., Raviram, R., et al. (2018). Capturing the onset of PRC2-mediated repressive domain formation. *Mol. Cell* 70, 1149–1162.e5. doi: 10.1016/j.molcel.2018.05.023
- Pang, C. N. I., Gasteiger, E., and Wilkins, M. R. (2010). Identification of arginine- and lysine-methylation in the proteome of *Saccharomyces cerevisiae* and its functional implications. *BMC Genomics* 11:92. doi: 10.1186/1471-2164-11-92
- Patel, A., Vought, V. E., Dharmarajan, V., and Cosgrove, M. S. (2008). A conserved arginine-containing motif crucial for the assembly and enzymatic activity of the mixed lineage leukemia protein-1 core complex. *J. Biol. Chem.* 283, 32162–32175. doi: 10.1074/jbc.m806317200
- Poepsel, S., Kasinath, V., and Nogales, E. (2018). Cryo-EM structures of PRC2 simultaneously engaged with two functionally distinct nucleosomes. *Nat. Struct. Mol. Biol.* 25, 154–162. doi: 10.1038/s41594-018-0023-y
- Qiao, Q., Li, Y., Chen, Z., Wang, M., Reinberg, D., and Xu, R. M. (2011). The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation. *J. Biol. Chem.* 286, 8361–8368. doi: 10.1074/jbc.m110.204115
- Qu, Q., Takahashi, Y. H., Yang, Y., Hu, H., Zhang, Y., Brunzelle, J. S., et al. (2018). Structure and conformational dynamics of a COMPASS Histone H3K4

- methyltransferase complex. *Cell* 174, 1117–1126.e12. doi: 10.1016/j.cell.2018.07.020
- Rao, R. C., and Dou, Y. (2015). Hijacked in cancer: the KMT2 (MLL) family of methyltransferases. *Nat. Rev. Cancer* 15, 334–346. doi: 10.1038/nrc3929
- Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593–599. doi: 10.1038/35020506
- Sarvan, S., Avdic, V., Tremblay, V., Chaturvedi, C. P., Zhang, P., Lanouette, S., et al. (2011). Crystal structure of the trithorax group protein ASH2L reveals a forkhead-like DNA binding domain. *Nat. Struct. Mol. Biol.* 18, 857–859. doi: 10.1038/nsmb.2093
- Senthilkumar, R., and Mishra, R. K. (2009). Novel motifs distinguish multiple homologues of Polycomb in vertebrates: expansion and diversification of the epigenetic toolkit. *BMC Genomics* 10:549. doi: 10.1186/1471-2164-10-549
- Singer, M. S., Kahana, A., Wolf, A. J., Meisinger, L. L., Peterson, S. E., Goggin, C., et al. (1998). Identification of high-copy disruptors of telomeric silencing in *Saccharomyces cerevisiae*. *Genetics* 150, 613–632.
- Sun, Z. W., and Allis, C. D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* 418, 104–108. doi: 10.1038/nature00883
- Talbert, P. B., Meers, M. P., and Henikoff, S. (2019). Old cogs, new tricks: the evolution of gene expression in a chromatin context. *Nat. Rev. Genet.* 20, 283–297. doi: 10.1038/s41576-019-0105-7
- Valencia-Sanchez, M. I., De Ioannes, P., Wang, M., Vasilyev, N., Chen, R., Nudler, E., et al. (2019). Structural basis of Dot1L stimulation by histone H2B Lysine 120 ubiquitination. *Mol. Cell* 74, 1010–1019.e6. doi: 10.1016/j.molcel.2019.03.029
- Vlaming, H., and van Leeuwen, F. (2016). The upstreams and downstreams of H3K79 methylation by DOT1L. *Chromosoma* 125, 593–605. doi: 10.1007/s00412-015-0570-5
- Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R. S., et al. (2004). Role of histone H2A ubiquitination in Polycomb silencing. *Nature* 431, 873–878. doi: 10.1038/nature02985
- Wood, K., Tellier, M., and Murphy, S. (2018). DOT1L and H3K79 methylation in transcription and genomic stability. *Biomolecules* 8:11. doi: 10.3390/biom8010011
- Worden, E. J., Hoffmann, N. A., Hicks, C. W., and Wolberger, C. (2019). Mechanism of Cross-talk between H2B Ubiquitination and H3 Methylation by Dot1L. *Cell* 176, 1490–1501.e12. doi: 10.1016/j.cell.2019.02.002
- Wu, C. T., Jones, R. S., Lasko, P. F., and Gelbart, W. M. (1989). Homeosis and the interaction of zeste and white in *Drosophila*. *Mol. Gen. Genet.* 218, 559–564. doi: 10.1007/bf00332424
- Wu, H., Zeng, H., Dong, A., Li, F., He, H., Senisterra, G., et al. (2013). Structure of the catalytic domain of EZH2 reveals conformational plasticity in cofactor and substrate binding sites and explains oncogenic mutations. *PLoS One* 8:e83737. doi: 10.1371/journal.pone.0083737
- Xu, C., Bian, C., Lam, R., Dong, A., and Min, J. (2011). The structural basis for selective binding of non-methylated CpG islands by the CFP1 CXXC domain. *Nat. Commun.* 2:227.
- Yang, X. D., Huang, B., Li, M., Lamb, A., Kelleher, N. L., and Chen, L. F. (2009). Negative regulation of NF-kappaB action by Set9-mediated lysine methylation of the RelA subunit. *EMBO J.* 28, 1055–1066. doi: 10.1038/emboj.2009.55
- Yang, Y., Joshi, M., Takahashi, Y. H., Ning, Z., Qu, Q., Brunzelle, J. S., et al. (2020). A non-canonical monovalent zinc finger stabilizes the integration of Cfp1 into the H3K4 methyltransferase complex COMPASS. *Nucleic Acids Res.* 48, 421–431.
- Yao, T., Jing, W., Hu, Z., Tan, M., Cao, M., Wang, Q., et al. (2019). Structural basis of the crosstalk between histone H2B monoubiquitination and H3 lysine 79 methylation on nucleosome. *Cell Res.* 29, 330–333. doi: 10.1038/s41422-019-0146-7
- Ying, Z., Mulligan, R. M., Janney, N., and Houtz, R. L. (1999). Rubisco small and large subunit N-methyltransferases. Bi- and mono-functional methyltransferases that methylate the small and large subunits of Rubisco. *J. Biol. Chem.* 274, 36750–36756. doi: 10.1074/jbc.274.51.36750
- Yuan, W., Wu, T., Fu, H., Dai, C., Wu, H., Liu, N., et al. (2012). Dense chromatin activates polycomb repressive complex 2 to regulate H3 Lysine 27 methylation. *Science* 337, 971–975. doi: 10.1126/science.1225237
- Zhang, K., Lin, W., Latham, J. A., Riefler, G. M., Schumacher, J. M., Chan, C., et al. (2005). The Set1 methyltransferase opposes Ipl1 aurora kinase functions in chromosome segregation. *Cell* 122, 723–734. doi: 10.1016/j.cell.2005.06.021
- Zhang, P., Chaturvedi, C. P., Tremblay, V., Cramet, M., Brunzelle, J. S., Skiniotis, G., et al. (2015). A phosphorylation switch on RbBP5 regulates histone H3 Lys4 methylation. *Genes Dev.* 29, 123–128. doi: 10.1101/gad.254870.114
- Zhang, P., Lee, H., Brunzelle, J. S., and Couture, J. F. (2012). The plasticity of WDR5 peptide-binding cleft enables the binding of the SET1 family of histone methyltransferases. *Nucleic Acids Res.* 40, 4237–4246. doi: 10.1093/nar/gkr1235
- Zhang, Y., and Kutateladze, T. G. (2019). Methylation of histone H3K79 by Dot1L requires multiple contacts with the ubiquitinated nucleosome. *Mol. Cell* 74, 862–863. doi: 10.1016/j.molcel.2019.05.013

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Janna, Davarinejad, Joshi and Couture. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Methods to Study Intracellular Movement and Localization of the Nucleotide Excision Repair Proteins at the DNA Lesions in Mammalian Cells

Mihaela Robu[†], Rashmi G. Shah[†] and Girish M. Shah^{*}

CHU de Québec Université Laval Research Centre (site CHUL), Laboratory for Skin Cancer Research and Axe Neuroscience, Québec, QC, Canada

OPEN ACCESS

Edited by:

Jean-Philippe Lambert,
Laval University, Canada

Reviewed by:

Wim Vermeulen,
Erasmus University Rotterdam,
Netherlands
Darel Hunting,
Université de Sherbrooke, Canada
Antonio Conconi,
Université de Sherbrooke, Canada

*Correspondence:

Girish M. Shah
girish.shah@crchul.ulaval.ca

[†] These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 31 July 2020

Accepted: 27 October 2020

Published: 17 November 2020

Citation:

Robu M, Shah RG and Shah GM
(2020) Methods to Study Intracellular
Movement and Localization of the
Nucleotide Excision Repair Proteins
at the DNA Lesions in Mammalian
Cells. *Front. Cell Dev. Biol.* 8:590242.
doi: 10.3389/fcell.2020.590242

Nucleotide excision repair (NER) is the most versatile DNA repair pathway that removes a wide variety of DNA lesions caused by different types of physical and chemical agents, such as ultraviolet radiation (UV), environmental carcinogen benzo[a]pyrene and anti-cancer drug carboplatin. The mammalian NER utilizes more than 30 proteins, in a multi-step process that begins with the lesion recognition within seconds of DNA damage to completion of repair after few hours to several days. The core proteins and their biochemical reactions are known from in vitro DNA repair assays using purified proteins, but challenge was to understand the dynamics of their rapid recruitment and departure from the lesion site and their coordination with other proteins and post-translational modifications to execute the sequential steps of repair. Here, we provide a brief overview of various techniques developed by different groups over last 20 years to overcome these challenges. However, more work is needed for a comprehensive knowledge of all aspects of mammalian NER. With this aim, here we provide detailed protocols of three simple yet innovative methods developed by many teams that range from local UVC irradiation to in situ extraction and sub-cellular fractionation that will permit study of endogenous as well as exogenous NER proteins in any cellular model. These methods do not require unique reagents or specialized instruments, and will allow many more laboratories to explore this repair pathway in different models. These techniques would reveal intracellular movement of these proteins to the DNA lesion site, their interactions with other proteins during repair and the effect of post-translational modifications on their functions. We also describe how these methods led us to identify hitherto unexpected role of poly(ADP-ribose) polymerase-1 (PARP1) in NER. Collectively these three simple techniques can provide an initial assessment of the functions of known and unknown proteins in the core or auxiliary events associated with mammalian NER. The results from these techniques could serve as a solid foundation and a justification for more detailed studies in NER using specialized reagents and more sophisticated tools. They can also be suitably modified to study other cellular processes beyond DNA repair.

Keywords: nucleotide excision repair (NER), NER proteins, poly(ADPR-ribose) polymerase-1 (PARP1), localization at DNA damage, local irradiation, *in situ* extraction, sub-cellular fractionation, intracellular movement

INTRODUCTION

The nucleotide excision repair (NER) is the most versatile DNA repair pathway that eliminates a wide variety of DNA lesions caused by different types of physical and chemical agents, such as ultraviolet radiation (UV), environmental carcinogen benzo[a]pyrene and anti-cancer drug carboplatin. It is the only repair pathway in mammalian cell that removes UV-induced DNA damage, such as cyclobutane pyrimidine dimers (CPD) and pyrimidine (6-4) pyrimidone (6-4PP) photoproducts (Scharer, 2013). The photosensitivity and susceptibility to develop sunlight-induced skin cancers in the individuals carrying mutations in NER genes, such as Xeroderma pigmentosum (XP) strongly indicates the importance of NER in the repair of UV damaged DNA (Giordano et al., 2016). The NER pathway uses more than 30 proteins to recognize the lesion, remove 24-32 nucleotides from the strand containing the damage, synthesize a new strand using the undamaged strand as a template and fill the gap (Marteijn et al., 2014). NER is divided into two sub-pathways: transcription-coupled repair (TCR) and global genome repair (GGR). TCR rapidly removes lesions that efficiently block the elongating RNA polymerase II complex during transcription, while GGR occurs in the whole genome (Vermeulen and Foustieri, 2013). These two sub-pathways differ in the lesion recognitions step, and subsequently converge to complete the repair process, as briefly summarized below.

The Xeroderma pigmentosum C (XPC) protein starts the GGR sub-pathway of NER by recognizing the distortion of the DNA double helix and binding to the unpaired nucleotides facing the damaged nucleotide. Its arrival at the damage site is a prerequisite for the recruitment of downstream proteins and the repair of the lesion (Puumalainen et al., 2014). In the mammalian cells, XPC's task of rapidly finding and localizing at the lesion site in the chromatin context is helped by UV-damaged DNA binding (UV-DDB) complex (DDB1 and DDB2) and poly(ADP-ribose) polymerase-1 (PARP1) (Pines et al., 2013). Following UV irradiation, PARP1 and DDB2 arrive rapidly and independently at the lesion, and influence each other's functions. PARP1 stabilizes DDB2 at the lesion and DDB2 stimulates PARP1's catalytic activity (Luijsterburg et al., 2012; Pines et al., 2012; Robu et al., 2013). PARP1 cleaves the substrate nicotinamide adenine dinucleotide (NAD⁺) and transfers the ADP-ribose moieties to poly(ADP-ribosyl)ate or PARylate itself and DDB2 as well as other acceptors proteins (Barkauskaite et al., 2015; Ray Chaudhuri and Nussenzweig, 2017). DDB2 activates the Cul4A-RBX1 ubiquitin ligase complex, containing DDB2, DDB1, Cul4A, and Rbx1 (Groisman et al., 2003) to modify histones, Cul4A and DDB2 itself. Together, these post-translational modifications (PTM) and the chromatin remodeling around the lesion facilitate recruitment and stabilization of XPC at the site. In addition, XPC forms a complex with PARP1 in the nucleoplasm prior to irradiation and is escorted rapidly by PARP1 to the lesion site after UV damage, thus improving the efficiency of initiation of GGR (Robu et al., 2017). The TCR pathway, on the other hand, is initiated when the elongating RNA polymerase II stalls at the lesion site and recruits Cockayne syndrome B (CSB) that is involved with other partners in recognition of the damage and

remodeling of chromatin (Van Der Weegen et al., 2020). Once the damage is recognized, the GGR and TCR sub-pathways converge with the recruitment of the RPA, XPA, and the basal transcription factor TFIIH to verify the damage. The dual incision of damaged DNA by endonucleases XPF-ERCC1 and XPG, followed by gap filling by different polymerases (polymerases δ , ϵ , and κ) and DNA ligation by ligase I or XRCC1-ligase III complex complete the repair (Mullenders, 2018).

The progress in understanding mammalian NER was relatively slower than other DNA repair pathways, largely because of a conceptual and some technical hurdles. In the early stage, the specific biochemical reactions carried out by a majority of the mammalian NER proteins were identified using the *in vitro* DNA repair assays with purified proteins and from their equivalent proteins in bacteria and yeast (Aboussekhra et al., 1995; Sugawara et al., 1998; Ticli and Prosperi, 2019). However, these techniques lack the spatio-temporal properties (Ticli and Prosperi, 2019). Hence, the bigger challenge was to understand the dynamics of their rapid recruitment to the lesion site and coordination of sequential steps of NER along with other proteins in mammalian cells. Based on yeast models of NER (Svejstrup et al., 1995) and some mammalian studies (He and Ingles, 1997), it was proposed that human cells carry out NER by "repairosome," a multi-protein complex containing most of the NER proteins. Considering that mammalian NER is initiated within seconds after DNA damage and continues for several hours, this concept posed a logistical challenge of keeping many of these multi-functional NER proteins engaged in a repairosome for the entire period of repair before and after their task is required. This model also hindered the discovery of new mammalian NER proteins, if they were not previously identified as a member of the mammalian repairosome complex. Despite accumulating evidence to the contrary, as described below, the concept of repairosome carrying out mammalian NER in human cells prevailed until 2003 Friedberg (2003).

A series of innovative methods developed over the last two decades by many groups allowed a rapid gain in our understanding of mammalian NER and challenged the repairosome concept. The development of green fluorescent protein (GFP) technology and photobleaching procedures permitted the visualization and quantification of the mobility of GFP-tagged NER factors (Vermeulen, 2011). Based on the speed of the recovery of the fluorescence of GFP-XPF-ERCC1 complex in the bleached area, Houtsmuller et al. (1999) concluded that ERCC1/XPF was not part of a large NER holocomplex or "repairosome." The same study also revealed the second hurdle that NER is not spatially constrained to sub-nuclear structures or "foci." This was because the uniform distribution pattern of the GFP-tagged NER proteins, seen in the unirradiated nuclei remained unchanged after global UVC-irradiation. In contrast, after the global exposure of cells with ionizing irradiation, etoposide or topoisomerase inhibitors, the double strand break (DSB) repair proteins accumulate in specific sub-nuclear structures called ionizing radiation induced foci (Maser et al., 1997; Pryde et al., 2005; Polo and Jackson, 2011). These foci therefore serve as an excellent physical location in the nucleus to examine roles of different proteins involved in DSB

repair. However, this fortuitous natural event of formation of foci that allowed rapid progress in understanding of DSB repair does not occur in UVC-irradiated cells, which prevented progress in study of NER.

The most important breakthrough in understanding mammalian NER came with the simultaneous development of local irradiation technique by two independent groups (Katsumi et al., 2001; Moné et al., 2001). In this method, the cell monolayer is covered with a UVC opaque polycarbonate filter with 3–8 μm pores, permitting irradiation of a small defined area within the nucleus, which can be readily detected by signal for DNA damage in the form of CPD. This technique revealed the order of assembly of NER proteins, ending the long-standing debate that recruitment of all NER proteins to the lesion site depends on the presence of XPC and not XPA (Volker et al., 2001). This technique also laid to rest the hypothesis that mammalian NER is carried out by a repairosome complex since the arrival and departure of GFP-tagged NER proteins, ERCC1-GFP and TFIIH-GFP at the lesion site, occurred as and when they were required in the sequential process of repair (Moné et al., 2004).

Despite its numerous advantages, including the low cost and use of basic equipment, the local irradiation technique needed optimization for the visualization of different repair proteins at damage site in different cell lines (Ticli and Prosperi, 2019). For instance, we observed that the uniform distribution of signal for PARP1 throughout the nucleus before irradiation did not change after local UVC irradiation. This was not due to lack of accumulation of PARP1 at the local site but due to the noise from the strong signal of PARP1 in rest of the nucleus drowning out the minor change in the signal intensity of PARP1 at the lesion site. To circumvent this problem, our team used two independent approaches. First, we used PAR formation as a proxy for recruitment of PARP1 to the lesion site, because binding of PARP1 to any type of DNA lesion results in its catalytic activation and formation of PAR (Pascal and Ellenberger, 2015). Using local UVC irradiation, we showed concurrent signals for CPD and PAR at the site of local irradiation (Vodenicharov et al., 2005). Thus, monitoring the outcome of recruitment of a protein at the local DNA lesion site offered us a good alternative to identify PARP1 as a new player in NER. The second approach was to deplete free PARP1 from nuclei using *in situ* high salt extraction as an additional step after fixation, to significantly reduce the noise from rest of the PARP1 while retaining the DNA damage-bound PARP1 (Purohit et al., 2016). Similar protocols for selective depletion of unrelated protein molecules using a mild treatment with DNase or RNase have been used to improve the detection of the XPG, DDB2, XPC (Dutto et al., 2017), and Ku80 (Britton et al., 2013) at the lesion sites. Along with this, the live-cell imaging revealed that NER is highly dynamic process with a continuous exchange of the repair factors during the repair reaction (Vermeulen, 2011; Ticli and Prosperi, 2019).

Thus, the local irradiation techniques and its improvements have revealed roles of many proteins in mammalian NER, including the unsuspected implication of the abundant protein PARP1 (Luijsterburg et al., 2012; Pines et al., 2012; Robu et al., 2013). However, local irradiation technique does not work for studying some proteins. For example, it is difficult to

visualize accumulation of the GFP-tagged TCR protein CSB at locally induced UVC spots since only a small fraction (15%) of CSB is recruited to the damage site (Van Den Boom et al., 2004; Aydin et al., 2014; Wienholz et al., 2019). While this situation is not different from the abundant protein PARP1, binding of CSB to DNA damage does not result in distinct functional product which could serve as a proxy for CSB recruitment. The second limitation of local irradiation technique with UVC opaque filter is that it allows study of NER only after UVC and not after treatment with other agents, such as cisplatin or Illudin S (Marteijn et al., 2014). Lastly, local UVC irradiation mediated visualization of a given NER protein is not amenable to study how chromatin marks and various PTMs (e.g., ubiquitination, phosphorylation, sumoylation, acetylation, and PARylation) occurring in the vicinity of the damage, affect the speed and accuracy of recruitment of core NER factors (Dantuma and Van Attikum, 2016). These modifications regulate the higher-order structure of chromatin to facilitate the sequential traffic of NER proteins at the lesion site through control over their recruitment and departure as well as their degradation. However, PTMs occur not only during DNA damage response but also for housekeeping functions, and identical PTM occurs on multiple proteins at the same site; hence immunodetection of PTM at local irradiation fails to identify uniquely repair related PTM of a single protein. The immunodetection of proteins after local irradiation would also not discriminate between unmodified and PTM-altered proteins. In this context, the use of sub-cellular and sub-nuclear fractionation linked with immunoprecipitation and immunoblotting approaches modified for a specific protein or its PTM can permit study of the dynamic response of the protein in NER of the DNA damage. The isolation of subcellular fractions (cytoplasmic, nucleoplasmic, and chromatin bound proteins) from mammalian cells has the advantage of revealing the changes in the intracellular redistribution of the proteins of interest after DNA damage. For instance, if PTM of a protein occurs only when it is bound to DNA lesion, then cell fractionation can reveal the enrichment of PTM-modified protein in chromatin-bound fraction. In addition, the co-immunoprecipitation studies of a protein in each fraction can reveal DNA-lesion specific interacting partners of that protein, which could be different from those in other sub-cellular fractions.

Despite the tremendous progress made in the study of mammalian NER, there is need for more studies on multiple fronts in NER. The studies on NER of chemotherapeutic drug-induced DNA damage could reveal clinically exploitable knowledge to improve therapeutic efficacy of these drugs. There could be many other unanticipated proteins, like PARP1, playing different auxiliary roles in controlling the functions of core NER proteins. There are significant gaps in our knowledge of early steps of mammalian TCR sub-pathway, such as roles of other transcription initiation and elongation factors who happen to be present in the vicinity of stalled RNA polymerase II. Lastly, we have just begun to understand TCR of nucleolar DNA at stalled RNA polymerase I, but more studies could reveal an attractive target to control cancer cells which are highly dependent on rDNA transcription to meet demands of fast-growing cells. To stimulate more studies in NER, there is a need for widespread

accessibility to techniques that will allow many more laboratories to explore this repair pathway in different models. In this context, some of the leading techniques in the field, such as live cell microscopy imaging using fluorescent tagged proteins (Vermeulen, 2011; Ticli and Prosperi, 2019) or UVC laser with quartz optics to cause damage in defined sub-nuclear zones (Dinant et al., 2007) have produced excellent data and will continue to be useful in future. However, these techniques need specialized reagents, engineered cell lines and expensive instruments which may not be available in most laboratories.

Therefore, here we describe detailed protocols for three relatively simple yet powerful techniques: local irradiation, in situ extraction and subcellular fractionation. These techniques use reagents and tools that are readily available in most laboratories and will allow identification of potential role of both endogenous or exogenous tagged proteins in mammalian NER in most cellular models. These techniques could be readily modified to study NER after treatment with other DNA damaging agents beyond UVC or to specifically study TCR. In addition, the cellular fractionation protocol provides an enriched nucleoplasmic or chromatin-bound protein fraction that can be useful for many downstream applications, such as immunoprecipitation and proteomics to identify the components of repair complexes or partners of the target proteins. Collectively these three simple techniques can provide an initial assessment of the functions of known and unknown proteins in the core or auxiliary events associated with the efficiency of mammalian NER. The results from these techniques could serve as a solid foundation and a justification for more detailed studies in NER using specialized reagents and more sophisticated tools, as required.

MATERIALS AND EQUIPMENT

Reagents for Cell Culture

- Human skin fibroblasts (GM00637-Coriell Institute) or any other cell line of interest.
- Appropriate medium for the cell culture, e.g., Minimum essential medium (α MEM) (Gibco, cat. no. 12561056, store at 4°C).
- Penicillin-streptomycin solution (10,000 U/mL penicillin and 10,000 μ g/mL streptomycin, Hyclone, cat. no. SV30010, store at -20°C).
- Bovine growth serum (replacement for fetal bovine serum, consists of bovine calf serum supplemented with chemically defined components, Hyclone, cat. no. SH30541.03, store at -20°C).
- Solution of 0.25% Trypsin and 2.12 mM EDTA (Wisent Inc., cat. no. 325-043-EL, store at 4°C).
- 1X Phosphate-buffer saline (PBS), pH 7.4. Prepare 10X buffer containing 1.37 M sodium chloride (NaCl) (Sigma-Aldrich, cat. no. S7653), 27 mM potassium chloride (KCl) (Sigma-Aldrich, cat. no. P9541), 100 mM disodium hydrogen phosphate (Na_2HPO_4) (Sigma-Aldrich, cat. no. S0876), and 20 mM potassium dihydrogen phosphate (KH_2PO_4) (Sigma-Aldrich, cat. no. P5379) in distilled water. Do not adjust the pH. Store at RT.
- Trypan Blue Stain (0.4%) (Gibco, cat. no. 15250-061).
- PJ34 hydrochloride, PARP1 inhibitor (Abcam, cat.no. ab120981). Prepare a 30 mM stock solution (10 mg/mL) in distilled water. Store the solution as aliquots at -20°C .

Reagents for Local Irradiation Protocol

- 3% (w/v) Paraformaldehyde solution (PFA) (Sigma-Aldrich, cat. no. P6148; store at 4°C).

Note 1. The PAF solution is prepared fresh prior to use. To prepare 100 mL solution, add 3 g of PFA powder in 80 mL water. To dissolve the PFA, add 100 μ l of 1N NaOH and heat the solution in a 60°C water bath. Vortex to mix. Cool the PAF solution, add 10 mL of 10X PBS and 100 μ l of 1N HCl solution to adjust the pH at 7.4. Add water to make the final volume to 100 mL and filter it to avoid particles. Keep the solution at RT until use.

Caution: PFA is toxic and corrosive. Avoid any direct contact and wear appropriate personal protective equipment.

- C (CSK) buffer. It contains 100 mM NaCl (Sigma-Aldrich, cat. no. S7653), 300 mM sucrose (Sigma-Aldrich, cat. no. 84097), 10 mM PIPES pH 6.8 (Sigma-Aldrich, cat. no. P1851), 3 mM MgCl_2 (Sigma-Aldrich, cat. no. M2670), and 1 mM EGTA [ethylene glycol-bis (β -aminoethyl ether)-N,N,N',N'-tetraacetic acid] (Sigma-Aldrich, cat. no. E8145). Prepare the buffer in distilled water. The C buffer can be stored at 4°C for several months.
- 100% Methanol (Fisher Chemical, cat.no. A452-44). **Caution:** Methanol is toxic if inhaled and in contact with skin. Avoid any direct contact and wear appropriate personal protective equipment.
- 100% Trichloroacetic Acid (TCA) (Sigma-Aldrich, cat. no. 490-10). Store at 4°C. **Caution:** TCA is corrosive. Avoid any direct contact and wear appropriate personal protective equipment.
- 100% Ethanol.
- Blocking buffer: PBS containing 5% (w/v) albumin from bovine serum (Sigma-Aldrich, cat. no. A9647) and 0.1% Triton X-100.
- Blocking buffer for PAR detection: PBS containing 5% (w/v) milk powder and 0.1% Tween 20 (Sigma-Aldrich, cat. no. P2287-500 mL).
- 12 N HCl (Sigma-Aldrich, cat. no. 320331-500 mL).
- Wash buffer: 0.1% Tween 20 in PBS.
- DAPI (Sigma-Aldrich, cat. no. D9542) stock solution (5 mg/mL) in distilled water. For long-term storage, the solution can be aliquoted and stored at -20°C . The solution can be kept at 4°C, protected from light, for short term storage. **Caution:** DAPI is mutagen. Avoid any direct contact and wear appropriate personal protective equipment.
- 1X Phosphate-buffer saline (PBS), pH 7.4.
- Mounting solution (Prolong Gold antifade reagent, Invitrogen, cat. no. P36934).
- Primary antibodies: mouse anti-CPD (clone TDM-2, Cosmo Bio, cat. no. NMDND001 used at 1/1000 dilution),

mouse anti thymine dimers (T-T) (clone KTM53, Kamiya Biomedical Company, cat.no. MC-062, 1/2000), goat anti-DDB2 (R&D, cat. no. AF3297, 1/500), and rabbit anti-PAR (LP-96-10, Aparptosis, 1/250).

- Secondary antibody conjugated to fluorescent dye: Alexa Fluor 488 and 594 goat anti-rabbit and anti-mouse IgG (Molecular Probe, Invitrogen, cat. no. A11029, A11034, A11012, and A11005) and Alexa Fluor 488 donkey anti-goat IgG (Molecular Probe, Invitrogen, cat. no. A11055) are used at 1/500 dilution.

Reagents for *in situ* Extraction Protocol

- 3% (w/v) Paraformaldehyde solution (PFA).
- C (CSK) buffer.
- C+T buffer (CSK buffer + 0.5% Triton): 100 mM NaCl, 300 mM sucrose, 10 mM PIPES pH6.8, 3 mM MgCl₂, 1 mM EGTA, 0.5% Triton X-100 (Sigma-Aldrich, cat. no. T8787). C+T buffer can be stored at 4°C for several months.
- C+T+S buffer (High salt CSK buffer): 420 mM NaCl, 300 mM sucrose, 10 mM PIPES pH 6.8, 3 mM MgCl₂, 1 mM EGTA, 0.5% Triton X-100. C+T+S buffer can be stored at 4°C for several months.
- 100% Ethanol.
- Blocking buffer: PBS containing 5% (w/v) BSA and 0.1% Triton X-100.
- 12 N HCl (Sigma-Aldrich, cat. no. 320331-500 mL).
- Wash buffer: PBS-0.1% Tween 20.
- 5 mg/mL DAPI stock solution.
- 1X PBS, pH 7.4.
- Mounting solution.
- Primary antibodies: mouse anti-CPD (1/1000), mouse anti thymine dimers (T-T) (1/2000), goat anti-DDB2 (1/500), rabbit anti-XPC (Gene Tex, cat. no. GTX70309, 1/1000 dilution), mouse anti-PARP1 (clone F123, Alexis, cat. no. ALX804211, 1/500 dilution), rabbit anti-XPA (Santa Cruz, cat. no. sc-853, used at 1/500 dilution).
- Secondary antibody conjugated to fluorescent dye: Alexa Fluor 488 and 594 goat anti-rabbit and anti-mouse IgG and Alexa Fluor 488 donkey anti-goat IgG (1/500).

Cell Fractionation Reagents

Note 2. Prepare the buffers in advance without protease or phosphatase inhibitors and PMSF and store it at 4°C and add the inhibitors on the day of the experiment. Alternatively, prepare the complete buffer and freeze it at −20°C.

Note 3. Prepare in advance the protease and phosphatase inhibitors and store in aliquots at −20°C. Thaw them on the day of experiment.

- 1X PBS, pH 7.4. Store at 4°C.
- 100 mM phenylmethylsulfonyl fluoride (PMSF) (Sigma-Aldrich, cat. no. P7626): 17.42 mg/mL in isopropanol.
- 100 mM β-glycerophosphate disodium (Sigma-Aldrich, cat. no. G9422), Ser/Thr phosphatases inhibitor: 21.604 mg/mL (solubility limit 50 mg/mL) in distilled water.

- 1 M sodium fluoride (Sigma-Aldrich, cat. no. S6521), Ser/Thr, and acidic phosphatases inhibitor: 41.99 mg/mL in distilled water.
- 100 mM sodium orthovanadate (Sigma-Aldrich, cat. no. S6508), Tyr and alkaline phosphatases inhibitor: 18.391 mg/mL in distilled water. Set pH to 9.0 with 1N HCl and boil until colorless. Cool to room temperature. Repeat this cycle until the solution remains at pH 9.0 after boiling and cooling. Bring up to the initial volume with distilled water.
- 10X protease inhibitors (Roche, cat. No. 1836170): dissolve one tablet in 1 mL distilled water.
- Buffer A: 10 mM Hepes, pH 7.8 (Sigma-Aldrich, cat. no. H4034), 10 mM KCl, 1.5 mM MgCl₂, 0.34 M sucrose, 10% glycerol (Sigma-Aldrich, cat. no. G5516), 0.1 % Triton X-100, 1 mM PMSF, 1X protease inhibitor, 10 mM β-glycerophosphate, 10 mM sodium fluoride, 1 mM sodium orthovanadate in distilled water.
- Buffer B: 50 mM Tris pH 7.8 (Sigma-Aldrich, cat. no. T6066), 420 mM NaCl, 1 mM EDTA (Ethylenediaminetetraacetic acid, disodium salt) (Thermo Fisher Scientific, cat. no. BP120-500), 0.5 % IGEPAL (Sigma-Aldrich, cat. no. I3021), 0.34 M sucrose, 10 % glycerol, 1 mM PMSF, 1X protease inhibitor, 10 mM β-glycerophosphate, 10 mM sodium fluoride, 1 mM sodium orthovanadate in distilled water.
- Chromatin extraction buffer: 20 mM Tris-HCl pH 7.5, 100 mM KCl, 2 mM MgCl₂, 1 mM CaCl₂ (Sigma-Aldrich, cat. no. C3306), 0.3 M sucrose, 0.1% Triton X-100, 1X protease inhibitor, 1 mM PMSF, 10 mM β-glycerophosphate, 10 mM sodium fluoride, 1 mM sodium orthovanadate in distilled water.
- Micrococcal nuclease solution > 100U/μl (Thermo Fisher Scientific, cat. no. 88216).
- Benzonase solution 379U/μl (Sigma-Aldrich, cat. no. E8263-5Ku).
- 10 mM MG132 (Abcam, cat. no. ab141003) stock solution. Dissolve at 4.76 mg/mL in DMSO. Solution can be aliquoted and store at −20°C.
- 10% TCA solution.
- Solubilisation solution: 0.25N NaOH (Sigma-Aldrich, cat. no. 415413-1L), 0.025% Triton in water.
- 22% Bradford assay dye solution (Biorad, cat. no. 5000006) in distilled water. Store at 4°C.
- Protein (BSA) standard solution (2 mg/mL) (Thermo Fisher Scientific, cat. no. PI23209).
- Antibodies: goat anti-DDB2 (1/1000), rabbit anti-XPC (1/1000), rabbit anti-XPA (used for IP at 1/100 dilution), mouse anti-PARP1 (clone F123, 1/5000), rabbit anti-PARP1 (1/5000), rabbit anti-beclin (Cell Signaling, cat. no. 3495, used at 1/1000), and rabbit anti-histone H3 (Abcam, cat. no. ab1791, used at 1/2000).

Equipment

- Sterile 35- and 100-mm culture dishes (or other size) (Corning, cat. no. 353001).
- Cell culture incubator (set at 5% CO₂ and 37°C).

- Hemocytometer (Thermo Fisher Scientific, cat. no. 0267110).
- Microscope cover glass (Thermo Fisher Scientific, cat. no. 12545102).
- Isopore membrane filters (3–8 μm pores, Millipore, cat. no. TMTP02500).
- Forceps.
- Ultraviolet hand lamp EF-140 with UVC lamp (BLE-2537S) (Thermo Fisher Scientific, cat. no. 119921200).
- UVX digital radiometer (UVP, cat. no. 534-243534-89).
- Parafilm (Sigma-Aldrich, cat. no. P76680).
- Inverted microscope (Axiovert 200, Zeiss) with AxioCam MRm camera.
- Microscope slides (Thermo Fisher Scientific, cat. no. 12-550-123).
- 1.5- and 15-mL tubes.
- Pipettes, tips, and scrapers.
- Refrigerated centrifuges.
- Sonic Dismembrator Model 500 (Thermo Fisher Scientific).

STEPWISE PROCEDURES

Overview

Protocols described here have been widely used in NER field for the last 20 years (Ticli and Prosperi, 2019). They have been adapted in our laboratory for analyses of PARP1 accumulation and its activation to form PAR at the DNA lesion site and its influence on the movement of NER proteins after induction of DNA damage. A scheme for each of these protocols is shown in **Figures 1, 2**. Cells are processed in monolayer after local irradiation with UVC (**Figure 1**) to visualize the movement of target proteins using immunofluorescent labeling. To improve the detection signal of the repair proteins at the lesion site, the unbound or “free” cytoplasmic and nucleoplasmic proteins are extracted by submerging the coverslips in buffers with increasing detergent and salt concentrations. Globally irradiated cells are either scrapped off the dishes or trypsinized to form single cell suspension (**Figure 2A**) and centrifuged followed by sequential extraction of the cell pellet in different buffers. The liberated cellular fractions are collected by centrifugation at each of the steps (**Figure 2A**). After removal of the nucleoplasm, the chromatin pellet is digested with high concentration of nucleases (MNase or benzonase) to liberate chromatin-bound protein fraction. The movement of repair proteins is analyzed by migrating these fractions on SDS-PAGE gel followed by immunoblotting of various proteins (**Figure 2B**).

A. Stepwise Procedure for Basic Local Irradiation Protocol

A.1 Cell culture [approximately ~ 24 h before irradiation]

A.1.1 Sterilize the glass coverslips. With sterile forceps, dip the glass coverslips in 100% ethanol. Air dry by place them in an angle into a 15 cm plate. Once they are dry, they are placed in 35 mm sterile dishes.

Note 4. Prepare one coverslip for each experimental time point and controls. To determine the number of samples, one needs to understand that the normal NER kinetics at damage site (15 s to 3 h) is affected by gene manipulation. For example, the XPC half-life at damage site is 1 h, however, in XPA deficient cells, the level of XPC was stable from 30 min to the end of protocol (3.3 h) (Luijsterburg et al., 2010). Hence, it is very important to run a pilot experiment in which early and late time points are included.

Note 5. The buffers volumes are for one 35 mm dish.

Note 6. Certain controls need to be run with each experiment, so plan to include these conditions when planning to seed the cells: IgG in place of primary antibody to check the specificity of the primary antibody, secondary antibody alone without primary antibody to verify lack of non-specific binding and no antibodies to verify autofluorescence.

A.1.2 In order to prepare cells on coverslips, calculate the number of coverslips needed for the experiment and the total number of cells required. Scale up the cells to have 30–40% more cells than that required on the day of seeding the coverslips. Trypsinize the cells and take a cell count to ensure that there are enough cells to carry out the experiment and that the number of cells seeded are same in each experiment. In our experiments, we add 300,000 cells per dish in 1.5 ml medium. If we need to seed 10 dishes, we make a master mix for about 13 dishes, by suspending $300,000 \times 13$ (3.9×10^6) cells in total of 1.5×13 ml (19.5 ml) medium and we add 1.5 ml of this slowly on the glass coverslips. Before each pipetting, mix the medium well. Cells should be at 70–80% confluency at the time of treatment, however this can vary depending on the experimental setting.

A.1.3 Grow the cells on the coverslip for 24 h in appropriate condition (5% CO₂, 37°C).

A.2 UVC irradiation [timing ~ 2 min per plate]

A.2.1 Pre-warm the UVC lamp for at least 10 min to avoid dose variability between the samples.

A.2.2 Measure the lamp flux using the UV meter across the surface of the irradiation. Determine the surface area required to irradiate based on the size of the dish used for the experiment. Measure the doses in five spatial points (the four corners and the middle of the irradiation surface) of this region and calculate the average to determinate the time needed to obtain certain dose of UVC, knowing that Dose = Intensity of the lamp (mw/cm^2) \times Time of exposure (sec). Place the dish within this region to accurately irradiate the cells.

A.2.3 Aspirate the media or remove it and store it for reuse. Wash the cells one time with 2 mL of 1X PBS. Aspirate the PBS and add 500 μl 1X PBS. For global irradiation, go to step A.2.4. For local irradiation, gently place the polycarbonate filter having 5 μm pores over the cell monolayer. Slowly aspirate the PBS from the dish. This will leave just enough PBS to form a thin layer between the cells and the filter.

Note 7. Do not move the filter once placed because of the risk to detach the cell-layer.

Note 8. The UVC rays do not pass through plastic, hence the lid of the dish is removed, but only once the dish is placed inside the UV chamber, to avoid cell contamination. To irradiate the cells,

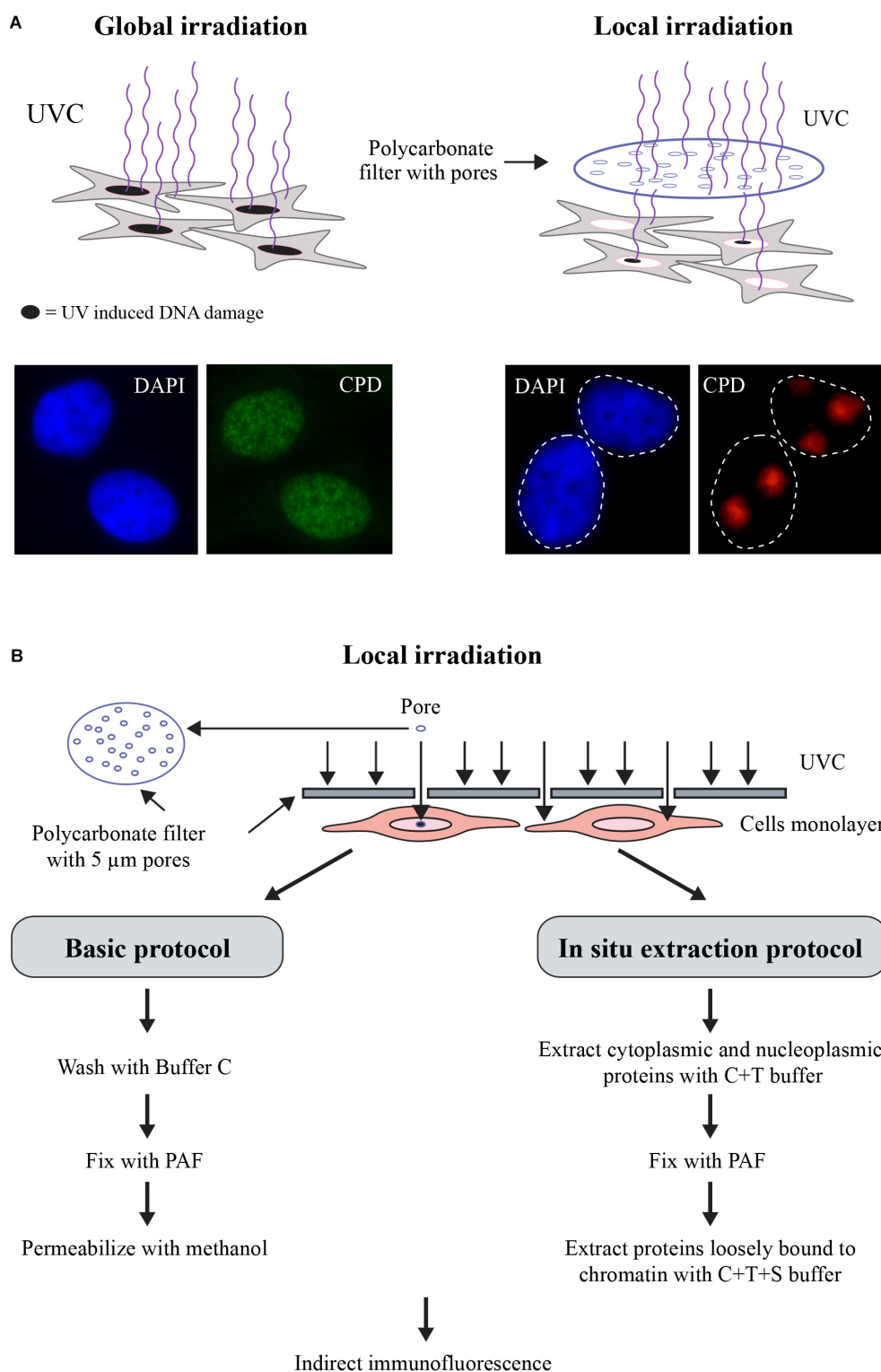
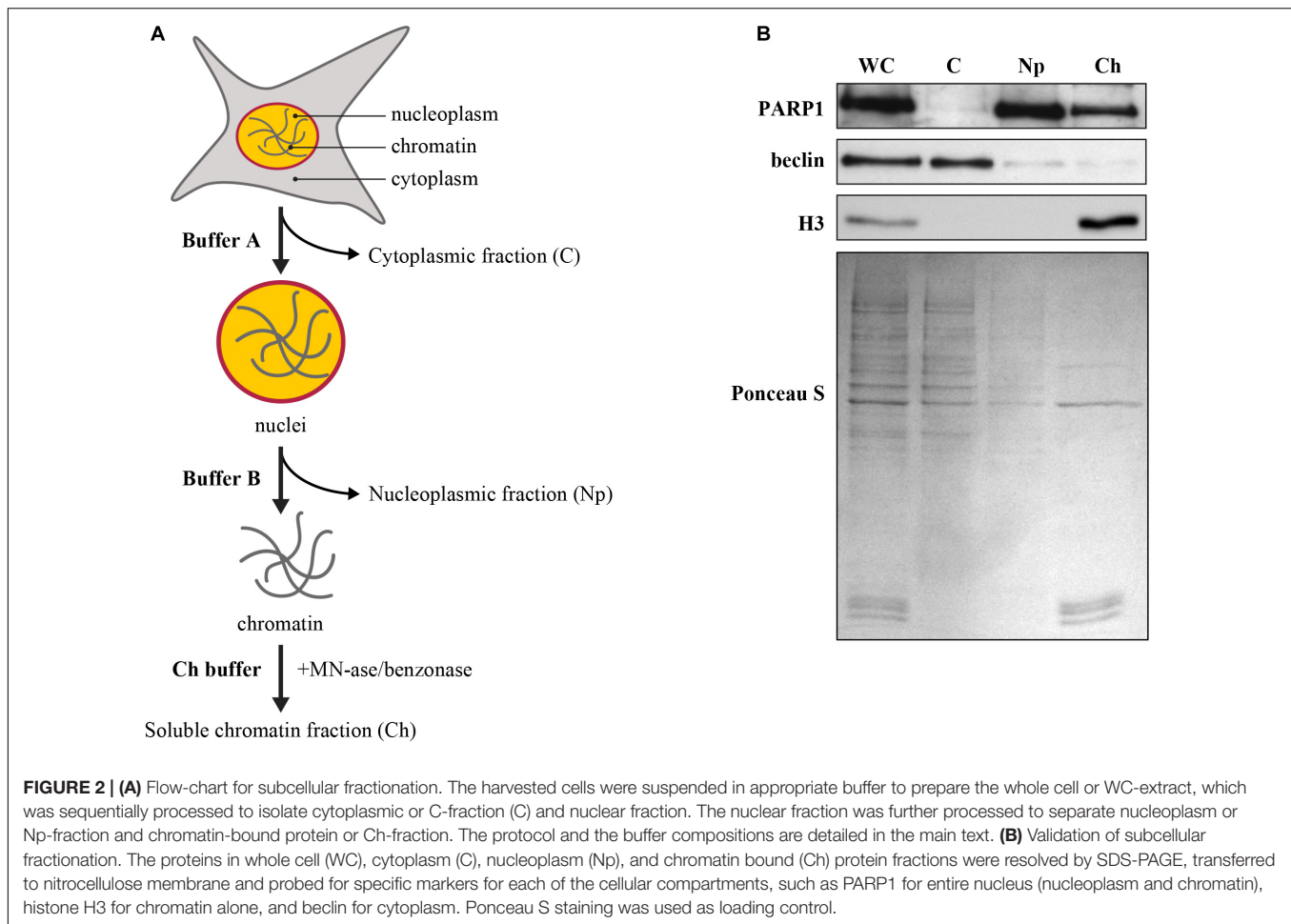


FIGURE 1 | (A) Scheme for Global or Local UVC-irradiation. The cells are exposed to UVC either unfiltered for global irradiation (left panel) or through an isopore polycarbonate filter with 3, 5, or 8 μm holes (right panel). DNA damage induced by either protocol of irradiation was detected by indirect immunofluorescence using an antibody specific for UV-induced photolesions. DAPI staining is carried out to define the nuclei. **(B)** Flow chart for the two versions of the local irradiation protocol. In the basic version (left), locally irradiated samples are fixed and processed for immunofluorescent detection, whereas in the *in situ* extraction protocol (right), the locally irradiated samples are processed for removal of non-DNA bound proteins prior to and after fixation step followed by immunofluorescent detection DNA-bound proteins. The steps and the buffer compositions are described in the main text.



remove the lid of the dish and start the timer simultaneously. After the calculated time, place the lid back to stop irradiation.

A.2.4 Irradiate the cells with a dose of 100 J/m^2 UVC for local irradiation or 10 J/m^2 (or specified dose) for global irradiation. Doses are applied by increasing the time of exposure.

A.2.5 Add media back to the dish and incubate for an appropriate time depending on the experiment setup.

A.3 Cell fixation and permeabilization [timing ~ 1 h]

Note 9. Do not allow the cells to dry out during any steps, since drying will increase the background fluorescence. When a buffer is aspirated, the new buffer should be added quickly.

Note 10. To stain the polymers of ADP-ribose, follow the steps A.3.6 to A.3.9.

A.3.1 Aspirate the media and wash 2 times with 1.5 mL of buffer C.

A.3.2 Fix the cells in 1.5 mL of 3% paraformaldehyde for 10 min at RT.

A.3.3 Wash 3 times with 2 mL 1X PBS.

A.3.4 Permeabilize in 1.5 mL of cold 100% methanol for 30 min at -20°C .

A.3.5 Wash 2 times with 2 mL 1X PBS.

A.3.6 Aspirate the media and wash with 1.5 mL of cold PBS.

A.3.7 Fix the cells in 1.5 mL of 10% TCA in PBS for 15 min on ice.

A.3.8 Aspirate the TCA and add 1.5 mL of cold 70% ethanol; incubate for 3 min at RT.

A.3.9 Repeat the step A.3.8 using sequentially 90 and 100% cold ethanol.

A.4 Indirect immunofluorescence [timing ~ 4 h]

A.4.1 Incubate the coverslips in 1.5 mL blocking buffer for 30 min at RT or overnight at 4°C . These steps will prevent non-specific antibody binding. *Safe stop!* The coverslips can be stored at 4°C for 3-4 days in the blocking solution. In this case, seal the plate with parafilm strip around the lids to avoid evaporation and sample drying.

A.4.2 Optional! Immunofluorescent labeling of CPD or 6-4PP in UV-damaged DNA requires denaturing of cellular DNA. To do this, after blocking, wash the coverslips 5 times with PBS and incubate them for 5 min at RT in 1.5 mL of 2 N HCl.

Note 11. During this incubation time, prepare a humidified chamber by placing a moist paper towels at the bottom of a box or dish. Cover the paper with parafilm, so that the coverslips are not in direct contact with the paper towels. Since the coverslips will be

placed on it, information necessary to identify the coverslip can be written on the parafilm using a fine permanent marker (do this at regular spacing, taking into account the size of the coverslip, such that the coverslip can be placed without touching each other).

A.4.3 Wash 5 times with 2 mL 1X PBS.

Note 12. Do not aspirate the last PBS wash. Removing coverslip from a bufferless dish results in their breaking due to surface tension. After removing the coverslips, do not discard the dishes. The coverslips are placed in their respective dishes for washings between the antibody incubations.

A.4.4 Dilute the first antibodies in blocking buffer (40 μ l of diluted antibody per 25 mm coverslip). Place 40 μ l drops of the diluted antibody near each identified spot on the prepared parafilm. Remove the coverslip from the dish using pointed forceps and blot the excess buffer by touching the edges of the coverslips on a paper towel. Gently invert the coverslip over the antibody, placing it “cell side facing down” over the drop. Cover the box with a lid or an aluminum foil to keep the moisture in. Incubate for 1 h at RT or overnight at 4°C.

A.4.5 Using the same forceps, gently remove the coverslip from the humid chamber and place it “cell side facing up” in its original dish containing 2 mL of wash buffer.

Note 13. Be very careful to not mix up the side on which the cells are grown.

A.4.6 Wash with 2 mL of wash buffer 3 \times 5 min each.

Note 14. Subsequent steps must be done in the dark to minimize the photobleaching of the fluorophore. Cover the dishes containing the coverslips and the humidified chamber with aluminum foil.

A.4.7 Dilute the secondary antibody at 1/500 dilution in blocking buffer. Proceed as describe in Note 11, 12 and step A.4.4. Incubate the coverslips in a humidified chamber for 30 min at RT.

A.4.8 Transfer the coverslips back to the dishes and wash with 2 mL of wash buffer 3 \times 5 min each.

A.4.9 Prepare 1.5 mL of DAPI solution (0.25 μ g/mL) in 1X PBS per 35 mm dish. Add over the coverslip in the plates and incubate for 5-10 min at RT. DAPI helps in microscopic identification of the nuclei by staining the DNA.

A.4.10 Wash 2 times with 2 mL 1X PBS to remove excess DAPI.

A.4.11 Wash 2 times with 2 mL distilled water to remove salts.

A.4.12 Add 2 mL of distilled water in the dish and remove the coverslips with forceps. Remove the excess water by dabbing the edges of the coverslips over a paper towel. Let the coverslips air dry into a box on a paper towel.

A.4.13 Mount the coverslips on a microscopic slide by inverting them on a drop of mounting media. The 20 μ l drop of mounting media is placed on a slide using a 200 μ l tip with a cut end. Take care to not introduce air bubbles while placing the drop and while placing the coverslip over it. Allow the antifade to slowly spread and cover the entire coverslip with minimal leakage outside the border of the coverslip. Seal the coverslips

edges with clear nail polish to prevent drying and movement of the coverslips during image acquisition.

Safe stop! The slides can be viewed immediately or stored at -20°C , in dark, for months.

A.5 Image acquisition and analysis [timing \sim 2 days]

Note 15. An important issue in interpreting the results of the recruitment of protein in the damaged region of the nucleus is the quantification of the signal. In the local irradiation technique, since most of the NER proteins would be present in the nucleus even before irradiation, the extent of enrichment of the protein at the site of DNA lesion after irradiation is derived as fold increase in signal intensity at the lesion site that is identified by CPD signal over the signal for the same protein in equivalent area of interest in the portion of the nucleus that is not irradiated. To analyze hundreds of images of this type and obtain a very robust statistical data in shorter time, it is advisable to develop macros either for the microscopy software or freely available Image J program.

A.5.1 Take the images of a least 100 nuclei at 40x or 63x magnification. We use an inverted fluorescence microscope Axiovert 200 (Carl Zeiss) equipped with AxioCam MRm.

Safe stop! Data processing can be conducted as per convenience of the user.

A.5.2 Delineate manually all the CPD positive spots using AxioVision 4.7 or others type of image analysis software.

A.5.3 To quantify the level of protein at lesion site, measure its fluorescence at the CPD spots.

A.5.4 Delineate and measure the intensity of a similar area in the unirradiated zone of nucleus to obtain a background corrected signal for the protein of interest.

A.5.5 Subject the data for the intensity of at least 100 spots from three independent experiments to statistical analysis to determinate the significance of difference.

B. Stepwise Procedure for *in situ* Extraction Protocol

Note 16. Cells are seeded and irradiated as described in local irradiation protocol; follow the steps A.1 and A.2.

B.1 Cell fixation and permeabilization [timing \sim 1 h]

Note 17. Do not allow the cells to dry out during any steps, since drying will increase the background fluorescence. When a buffer is aspirated, the new buffer should be added quickly.

B.1.1 Aspirate the media and wash 2 times with 1.5 mL of buffer C.

B.1.2 Permeabilize in 1.5 mL of buffer C+T for 8 min at RT. This step will remove soluble cytoplasmic and nucleoplasmic proteins.

B.1.3 Wash 2 times with 2 mL 1X PBS.

B.1.4 Fix the cells in 1.5 mL of 3% paraformaldehyde for 10 min at RT.

B.1.5 Wash 3 times with 2 mL 1X PBS.

B.1.6 Extract with C+T+S buffer for 20 min on ice to remove the protein bound to chromatin under unchallenged condition. This step is *optional* and recommended to extract the abundant proteins bound to undamaged chromatin.

B.1.7 Wash 3 times with 2 mL 1X PBS.

B.2 Indirect immunofluorescence [timing ~ 4 h]

Note 18. Follow the step A.4 described in local irradiation protocol.

B.3 Image acquisition and analysis [timing ~ 2 days]

Note 19. Follow the step A.5 described in local irradiation protocol.

C. Stepwise Procedure for Subcellular Fractionation Protocol

Note 20. Before starting the procedure, prepare all your solutions and keep them refrigerated. Switch on the refrigerated centrifuges. Perform all steps on ice to avoid proteins degradation. When scraping the cells from the plates, place them on ice (make a thin layer of ice in trays and place aluminum foil on it. Place the plates on these ice trays). Identify and prepare all the tubes for the cell collection and place them on ice.

Note 21. Although the volumes are small, use 15 mL tubes and a centrifuge with a swing out rotor to obtain good pellets.

Note 22. Seed at least 10 cm dish per condition. Cell should be at 80% confluency the day of irradiation. The buffer volumes are for one 10 cm petri dish.

Note 23. Irradiate the cells globally with 10 J/m² UVC by following the steps described in A.2, except step A.2.3 (without polycarbonate filter).

Note 24. If using proteasome inhibitors to study protein degradation due to ubiquitination, the proteasome inhibitor MG132 is added at 5-10 μ M to the cells, 1 h before treatment.

C.1 Extraction of soluble protein fractions [timing ~ 3 h]

C.1.1 Aspirate media and wash the plate once with 3 mL cold 1X PBS. Scrape the cells in 1 mL cold PBS and transfer them to a 15 mL tube placed on ice. Add another 1 mL cold PBS to the dish, scrape and pool with the previously scraped cells.

C.1.2 Centrifuge at 600g for 10 min at 4°C.

C.1.3 Aspirate the PBS and keep the tube on ice.

C.1.4 Add 190 μ l of cold buffer A to the cell pellet. Mix with P200 pipette several times and leave the tube on ice for 7 min. Remove an aliquot as whole cell extract (WC) (about 30 μ l to which add 30 μ l of 2X loading buffer).

C.1.5 Centrifuge the extract at 1000g for 5 min at 4°C.

C.1.6 Transfer the supernatant using a P200 to labeled Eppendorf tube. This is the cytoplasmic extract. Take an aliquot for Western (30 μ l).

C.1.7 Wash the pellet once with 190 μ l of buffer A; mix using P200.

C.1.8 Centrifuge the extract at 1000g for 5 min at 4°C.

C.1.9 Discard the washing (remove with P200, taking care not to touch the pellet).

C.1.10 Add 190 μ l of buffer B to each tube, mix with P200 pipette and transfer the cells to Eppendorf tubes. The extracts should be viscous due to nuclear breakage and release of DNA.

C.1.11 Incubate the tube on ice for 30 min and centrifuge at maximum speed (16,000g) at 4°C for 30 min.

C.1.12 Transfer the supernatant into a new tube; this is the nucleoplasmic extract. Take a fraction for western (30 μ l). Keep the pellet to extract chromatin.

Safe stop! You can either extract the chromatin on the same day or freeze the pellet and extract it later. Freeze all others fractions at -20°C.

C.2 Extraction of chromatin bound protein fraction [timing ~ 2 h]

Note 25. The chromatin bound proteins are removed by digestion with nucleases in chromatin buffer. It is important to optimize the removal of proteins bound to chromatin in the cell extracts. This can be done by resolving the DNA released after nuclease digestion on agarose gel.

Note 26. To prepare the chromatin-bound protein fraction, we use 100 U/mL MNase or 25 U/mL benzonase. Higher amount of MNase (4000 U/mL) or benzonase (50 U/mL) can be used to digest the DNA down to mononucleosomal level.

C.2.1 Suspend the chromatin pellet in 50 μ l of chromatin buffer.

C.2.2 The tight chromatin pellet is opened by mild sonication at lowest setting (setting 11 in Sonic Dismembrator Model 500 from Thermo Fisher Scientific) for 10 s. For higher volumes, time might be more, up to 15-20 s. If the clumps of chromatin persist, put the tube on ice after the first sonication, and then sonicate again, for 10 s. Usually one sonication is enough to take care of the clumps.

C.2.3 Add 100 U/mL MNase to each tube (0.25 μ l of 20 U/ μ l stock). Mix by vortexing. Incubate the tubes at RT for 40 min. Vortex intermittently.

C.2.4 Stop the reaction with 5 mM EDTA (1 μ l of 500 mM stock solution) and 5 mM EGTA (5 μ l of 100 mM). Spin at maximum speed for 10 min and collect the supernatant into a new tube. This is the chromatin extract.

Safe stop! You can either freeze the fraction at -30°C or proceed to the Western blot analysis.

C.3 Western blots analysis of cellular fractions [timing ~ 2 days]

Note 27. Always validate the protocol by verifying the purity of different subcellular compartments before analyzing the movement of repair proteins, since many of them are present in more than one cellular compartment at the time.

Note 28. Cell fractions are run on Western blots either based on their cell numbers or protein content. Protein estimation is done by using Bradford assay in a 96 well plate.

C.3.1 Whole cell extract in C.1.4: Sonicate the whole cell extract fraction at setting 45 for 20 seconds to reduce viscosity. This can be either frozen for later use or used to estimate proteins followed by Western blot analysis right away.

C.3.2 For protein estimation in WC, precipitate the proteins by adding 1-2 μ l of the above prepared extract to 100 μ l of cold 10% TCA on ice for at least 30 min. Centrifuge at 16000 g for 5 min. Wash the obtained protein pellet with ethanol to remove traces of TCA and dissolve it in 50 μ l solution of 0.25N NaOH-0.025% Triton X-100.

C.3.3 To estimate the proteins in nucleoplasmic and chromatin extract obtained in C.1.12 and C.2.4 dilute the fractions 1: 20 and 1:10 in NaOH-Triton X-100 buffer, respectively. This will prevent the interference of salt and detergents on protein estimation and also bring the concentration of the sample in the linear range of standard curve for protein estimation (5–150 $\mu\text{g/mL}$).

C.3.4 Quantify the protein concentration using a Bradford assay or similar assay.

C.3.5 Separate 5–10 μg proteins of each cellular compartment on 10 and 12% SDS-PAGE, transfer it in wet condition (100V for 90 min or 35V for 16 h) in Tris-glycine transfer buffer without SDS to a nitrocellulose membrane (GE Healthcare Life Science) and probe with primary antibody against known cytoplasmic (beclin 1/1000), nuclear (polyclonal PARP1 1/1000), and chromatin bound protein (H3 1/2000) markers (**Figure 2B**). We observe the localisation of beclin in cytoplasm, that of histone H3 on the chromatin, whereas PARP1 is present in both, nucleoplasm and chromatin-bound fraction.

RESULTS AND DISCUSSION

The three simple methods described above do not require specialized reagents and equipment and can be readily performed in most laboratories. These methods will permit analyses of the fate and functions of various endogenous or exogenous proteins during NER. These functions range from recruitment and persistence or departure from the lesion, their interactions with other partners and role of PTMs in these processes. If one uses exogenous tagged NER proteins, note that PTM may be affected by the presence of tag, as observed for DDB2 and XPC (Puumalainen et al., 2014; Matsumoto et al., 2015), which will require validation with untagged or endogenous proteins. The fluorescent signal from tags, such as GFP may be quenched by some solvents or chemicals used in these protocols, therefore, an antibody-based detection of the tag or the protein may be required. Here, we provide examples of how these protocols could specifically reveal the NER-related functions of four key proteins DDB2, PARP1, XPC and XPA in GM human skin fibroblasts.

Local UVC Irradiation With Immunofluorescence Detection as a Primary Screen to Identify Proteins Recruited to DNA Lesion During NER

The local irradiation of cells at 100 J/m^2 UVC through an isopore polycarbonate filter with irregularly distributed 5 μm holes (**Figure 1**) results in DNA damage in defined subnuclear regions (Katsumi et al., 2001; Moné et al., 2001). Note that local irradiation requires higher dose than global irradiation to achieve comparable levels of DNA damage because most of the incident UVC-irradiation reaching the cells from different angles of the source lamp is blocked by UVC-opaque filter except the radiation coming directly from above the pores. The cells can be harvested at various time points after irradiation to follow the time course

of recruitment of proteins at damage site. After harvesting, cells are fixed with formaldehyde to preserve cellular morphology, immobilize cellular components in their original locations and prevent their degradation. Since antibody molecules are too large to enter the cells, the fixed cells are permeabilized by treatment with non-ionic detergents such as Triton X-100 (T) followed by direct or indirect immunocytochemistry protocols to detect the proteins of interest in the locally irradiated zone versus their native distribution in the nucleus and/or in the cytoplasm. Using this pioneering technique, Volker et al. (2001) clearly showed the sequential accumulation of key endogenous NER proteins XPC, XPB, and XPA to the local UV-irradiated sub-nuclear zones. Here, as an example, we show the enrichment of the signal for endogenous DDB2 (green signal), a protein arriving rapidly at the subnuclear UV-lesion site containing CPD (red signal), within 10 min after local UVC irradiation (**Figure 3A**). DDB2 has been shown to facilitate recruitment and stabilization of XPC at the lesions site (Puumalainen et al., 2016). The basic local UVC irradiation readily serves as a low cost and relatively simple technique for primary screening of recruitment of various proteins to the DNA lesion site. Once the recruitment of any protein to lesion site is confirmed, one could perform more detailed studies to understand its significance in NER.

However, one of the limitations of local irradiation technique is that it does not work for proteins that are abundant in the nucleus, such as PARP1. In an unchallenged condition, PARP1 is free in the nucleoplasm and also bound to the chromatin for its housekeeping functions (Kim et al., 2005), hence it exhibits a homogenous distribution throughout the nucleus. Since a very small fraction of PARP1 localizes at DNA damage after local UVC-irradiation, it is challenging to discern this minor enrichment of the signal of PARP1 at the DDB2 spots (which serve as proxy for DNA lesion site) from the overwhelming background signal from rest of the nuclear PARP1 (**Figure 3B**). To circumvent this problem, we exploited the fact that after binding to any type of DNA damage, PARP1 is strongly catalytically activated to form PAR (Pascal and Ellenberger, 2015). Hence, we used PAR detection as a proxy for recruitment and binding of PARP1 at the site of DNA lesion (Vodenicharov et al., 2005). Using local UVC irradiation, we showed that signal for PAR colocalizes with CPD within seconds at the site of UV-lesion (**Figure 3C**). It is important to note that strong reagents used for fixation or permeabilization of cells, such as use of methanol with formaldehyde (for DDB2 and PARP1) and TCA (for PAR), can affect the integrity of nuclear structure thereby releasing the nuclear content in the cytoplasm (Hoetelmans et al., 2001), which is visible in local irradiation protocol (**Figure 3C**). Nonetheless the fractions of DDB2 or PAR linked to DNA remain clearly visible in this protocol. The rapid formation of PAR within seconds after UVC irradiation (Vodenicharov et al., 2005) when CPD are just formed and DNA strand breaks by NER have not yet been created, challenged the prevalent opinion that PARP1 could not play any role in NER, which was based on two assumptions, namely PARP1 activation requires DNA strand breaks and PARP1 is not a member of the repairosome complex. Subsequent studies from many teams, including ours, showed that rapid arrival of PARP1 and PAR formation at the CPD lesion

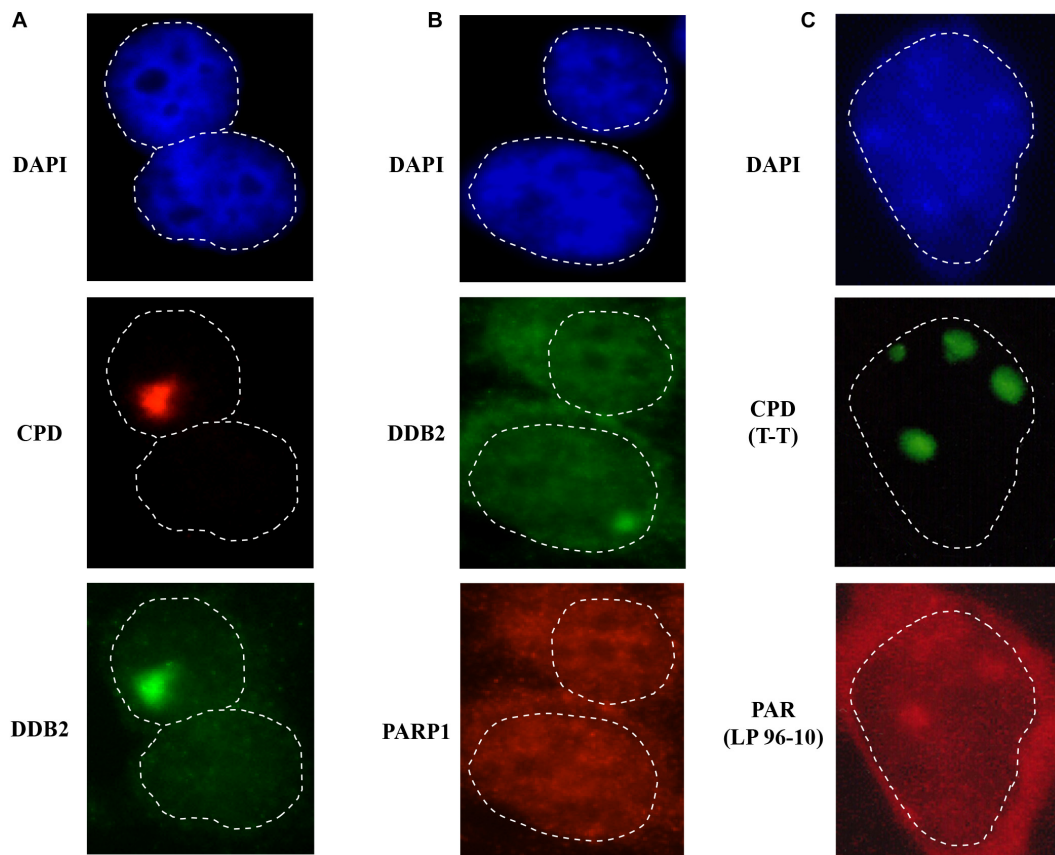


FIGURE 3 | Detection of early NER proteins at the site of damage. GMU6 human skin fibroblasts were locally irradiated with 100 J/m² through 5 μ m pores filter, fixed after 10 min and processed for immunofluorescence labeling for CPD and DDB2 (panel **A**) or DDB2 and PARP1 (panel **B**). **(C)** PAR formation at local damage site. GMU6 were fixed with TCA-ethanol at 15 sec after local irradiation and probed for PAR and CPD using the specific antibodies. DAPI staining was carried out to define the nuclei. Note that the cytoplasmic leakage of signal for proteins or PAR (panels **B,C**) is caused by use of the harsh chemicals, such as TCA for PAR-precipitation or methanol for membrane permeabilization. Nonetheless, the signal for these targets immobilized at lesion site remains clearly visible in the nucleus.

site play key roles with DDB2 to initiate GGR and to recruit XPC to the lesion site (Luijsterburg et al., 2012; Pines et al., 2012; Robu et al., 2013, 2017). The time course of local irradiation also showed that PARP1 activation at the lesion site is a transient process because the signal for PAR disappears within one hour after irradiation (Vodenicharov et al., 2005). Thus, local UVC irradiation is a simple technique that can be used in any cellular model for primary screening of the direct or indirect role of any protein or a process in NER of UVC-induced DNA damage.

Local Irradiation With *in situ* Cell Fractionation as a Sensitive and Quantitative Method to Assess Enrichment of Proteins to DNA Lesion During NER

The basic local UVC irradiation technique does not work for direct detection and quantification of the recruitment of the abundant nuclear proteins at the damage site, because the strong signal from rest of the protein would mask the minor extent of protein enrichment at the lesion site. In case of PARP1, we used

proxy signal of PAR formation at the lesion site to circumvent this problem, but this proxy option is not feasible for most of the NER proteins. Moreover, PAR can be made by other PARPs too, and therefore, it would be ideal to directly demonstrate the recruitment of PARP1 to the lesion site. Hence, here we describe a method that offers a more general solution by combining local irradiation with the selective depletion of the unbound or “free” proteins from the nuclei (noise) to reveal the signal for lesion-recruited proteins (Figure 4). This method allowed us to detect and quantify the extent of enrichment of PARP1 at the lesion site (Purohit et al., 2016). Since Triton removes free proteins not bound to DNA, the CSK buffer containing 100 mM NaCl (buffer C) with Triton X-100 (C+T) has been used prior to fixation with formaldehyde in the immunofluorescence detection of repair proteins recruited to damaged DNA (Balajee and Geard, 2001; Mirzoeva and Petrini, 2001). Almost all the DDB2 signal outside the local irradiation zone was depleted in C+T sample, demonstrating a clear improvement of signal to noise ratio for DDB2 in C+T buffer over C buffer after local irradiation (Figure 4A, DDB2 panel). However, C+T did not significantly deplete PARP1 protein outside the local

irradiation zone. It was earlier shown that 1.6 M NaCl could deplete all the PARP1 from cells (Kaufmann et al., 1991). Therefore, we titrated the salt concentration to determine that 0.42 M salt added to C+T buffer to create (C+T+S (salt)) buffer can extract most of the unbound PARP1 as well as DDB2 from unirradiated cells or from the unirradiated zones of the nuclei. With C+T+S protocol, we could clearly visualize recruitment of PARP1 to the local irradiation zone and also improve the signal to noise ratio for DDB2 (**Figure 4A**, DDB2 and PARP1 panels).

The improved protocol not only permitted quantification of PARP1 enrichment at the UVC-lesion of irradiated cells but also identified its localization to DNA for housekeeping functions in the unirradiated control cells (Purohit et al., 2016). We also used this improved technique to obtain kinetics of recruitment and departure of XPC protein over 3 h after local UVC irradiation (Robu et al., 2017). This was achieved by quantification of the extent of enrichment of XPC at the lesion site over the background signal for XPC in non-irradiated zone of the nucleus (**Figure 4B**). Here, we noted a strong accumulation of XPC at CPD lesions at 10 min after irradiation followed by a steady decline by 180 min (**Figure 4B** and chart), confirming the previously published data (Wang et al., 2003; Puumalainen et al., 2014). Together, the basic local irradiation combined with in situ extraction by C+T or C+T+S buffers can readily permit immunofluorescent visualization and quantification of the signal of a much larger group of NER proteins at the site of local UVC-damaged DNA.

Sub-cellular Fractionation After Global UV Irradiation to Study Diverse Functional Aspects of NER Proteins

The basic and improved local UVC irradiation techniques allow visualization of recruitment of NER proteins to the lesion site. Here, we describe an alternative approach of the global UVC-irradiation followed by sub-cellular fractionation and immunoblotting to examine biochemical and mechanistic studies related to recruitment, tracking intracellular movement of NER proteins and their interaction with other partners in different cellular compartments. This technique is suitable for study of NER of DNA damage caused by global UVC irradiation, as described here, but also by chemical agents or drugs in cellular models. In this protocol, it is important to incorporate multiple validation steps during the experiments to minimize the variability in data and to generate the quantitative data from Western blot analysis. These validation steps include quantification of protein extracts, validation of antibody and internal controls, detection of the combined linear range for the target protein for the detection methods and internal loading control (Taylor and Posch, 2014). We have already used this technique to identify the role of PARP1 in the lesion recognition step of NER pathways (Vodenicharov et al., 2005; Robu et al., 2013, 2017; Purohit et al., 2016). Here, we describe various uses of sub-cellular fractionation technique to study different functional aspects of NER proteins.

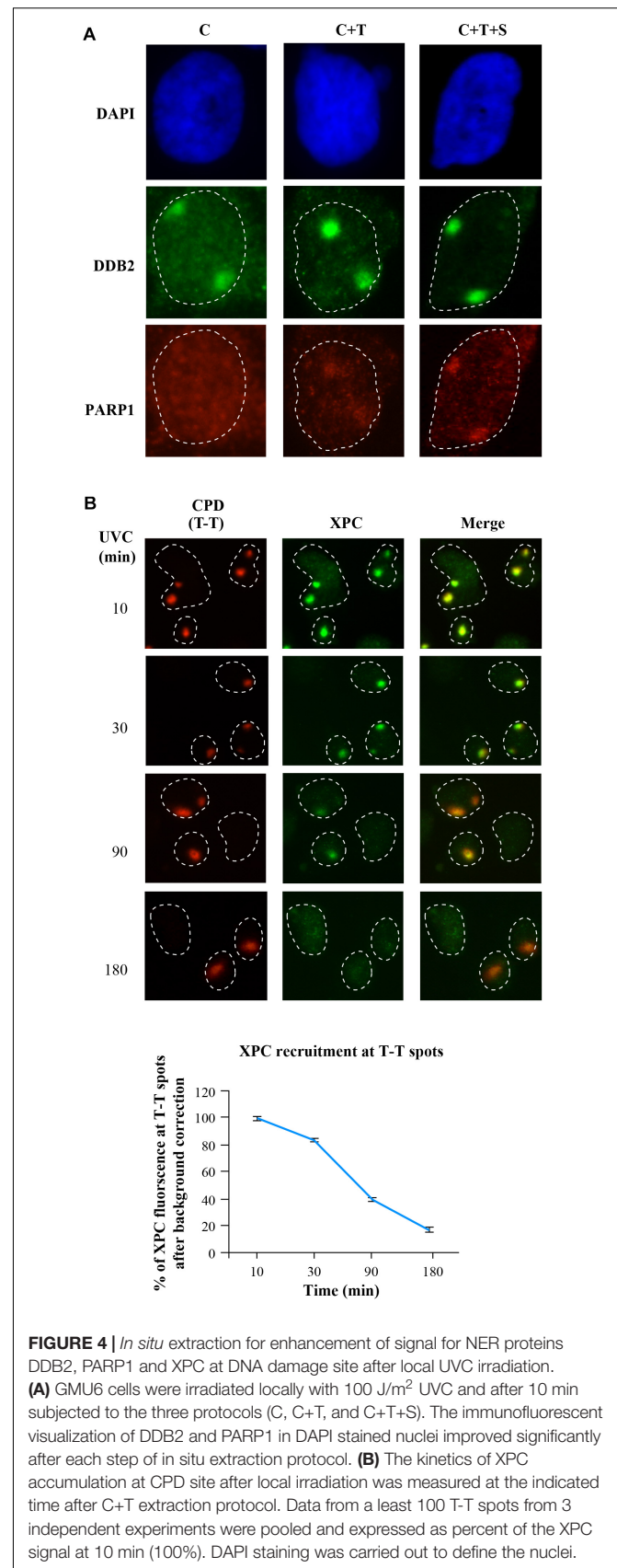
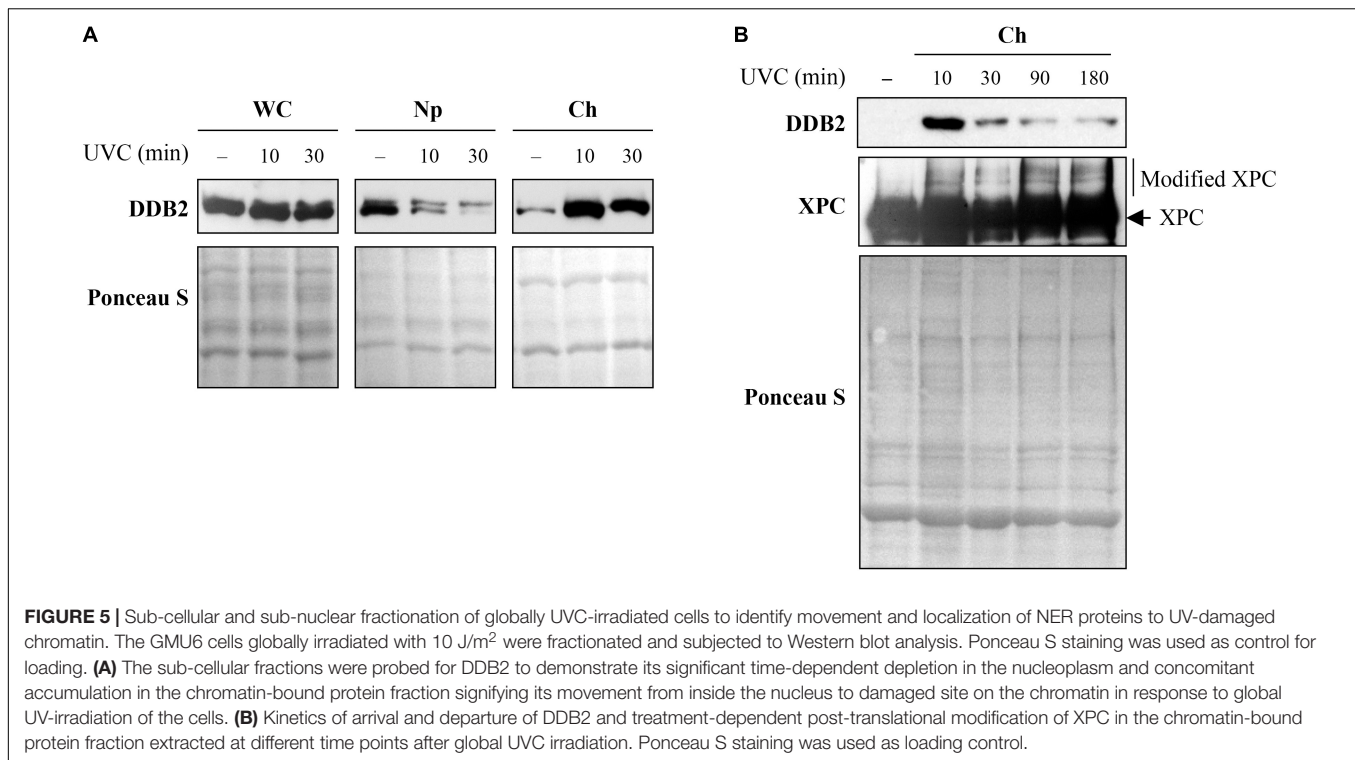


FIGURE 4 | *In situ* extraction for enhancement of signal for NER proteins DDB2, PARP1 and XPC at DNA damage site after local UVC irradiation. **(A)** GMU6 cells were irradiated locally with 100 J/m² UVC and after 10 min subjected to the three protocols (C, C+T, and C+T+S). The immunofluorescent visualization of DDB2 and PARP1 in DAPI stained nuclei improved significantly after each step of in situ extraction protocol. **(B)** The kinetics of XPC accumulation at CPD site after local irradiation was measured at the indicated time after C+T extraction protocol. Data from a least 100 T-T spots from 3 independent experiments were pooled and expressed as percent of the XPC signal at 10 min (100%). DAPI staining was carried out to define the nuclei.



Kinetics of Recruitment and Departure of Repair Proteins During NER

To assess the timing of initial recruitment and accumulation of early NER proteins at the UV damage, cells were harvested before or 10 and 30 min after global irradiation with 10 J/m² UVC to obtain whole cell (WC), nucleoplasm (Np), and chromatin-bound (Ch) protein fractions. Each of these subcellular fractions were separated on SDS-PAGE and immunoblotted for DDB2 (**Figure 5A**). As shown in earlier studies (Matsumoto et al., 2015), a significant amount of the cellular DDB2 (WC fraction) was present in the nucleoplasm (Np) but not in the chromatin-bound fraction (Ch) prior to irradiation. However, 10 min after irradiation, there was a massive accumulation of DDB2 in the chromatin fraction accompanied by its depletion from the nucleoplasmic fraction, confirming this observation made earlier (Radic-Otrin et al., 2002). Note that the Western blot of whole cell extract (WC) reveals no change in the signal for DDB2 after irradiation. Thus, the sub-nuclear fractionation of cells at several time points after global UVC irradiation can serve as a relatively simple primary screen to characterize the movement of some NER protein from different cellular compartments (cytoplasm to nucleoplasm) to the UV-induced DNA lesions in the chromatin.

However, this simple sub-nuclear fractionation approach does not work for all NER proteins. For example, when we tracked chromatin-bound DDB2 and XPC for up to 3 h after global UVC irradiation, we observed a strong accumulation of DDB2 at 10 min and a significant decrease from 30 min to 3 h (**Figure 5B**, DDB2 panel). In contrast, we could not identify any significant change in XPC levels from 10 min to 3 h after irradiation (**Figure 5B**, XPC panel), which has been reported

earlier (Fei et al., 2011). However, it was evident from the local UVC irradiation linked to situ fractionation studies that XPC was indeed recruited to the lesion site 10 min after irradiation and departed by 3 h (**Figure 4B**). This difference in results between two techniques could be because the lesion recognition proteins, such as XPC, constantly scan the DNA for the presence of damage by association and dissociation from chromatin, which explains large signal for chromatin-bound XPC in control cells. Moreover, upon UV irradiation, while most of the cellular DDB2 molecules (85%) localize to the UV-damaged chromatin, only 25% of the XPC molecules are immobilized to the damaged DNA (Luijsterburg et al., 2007). Thus, it is important to examine different techniques to probe recruitment and departure of NER proteins. Interestingly, although simple immunoblotting of chromatin extract of post-irradiated samples did not reveal changes in accumulation kinetics of XPC, it revealed a clear shift in XPC mobility (**Figure 5B**), indicating a post-translational modification of XPC at the lesion site, which has often been suggested to be ubiquitination (Sugasawa et al., 2005; Wang et al., 2005). Thus, although inconclusive, such initial results from Western blot of sub-cellular fraction can provide a lead to the post-translational modification of XPC and similar proteins. This can then be further characterized by immunoprecipitation and mass spectrometry to identify specific sites of modification and study their impact on the NER functions of the protein.

Implication of Post-Translational Modifications in Recruitment and Departure of NER Proteins in NER

Different PTM regulate the arrival and/or departure of repair proteins during NER (Dantuma and Van Attikum, 2016).

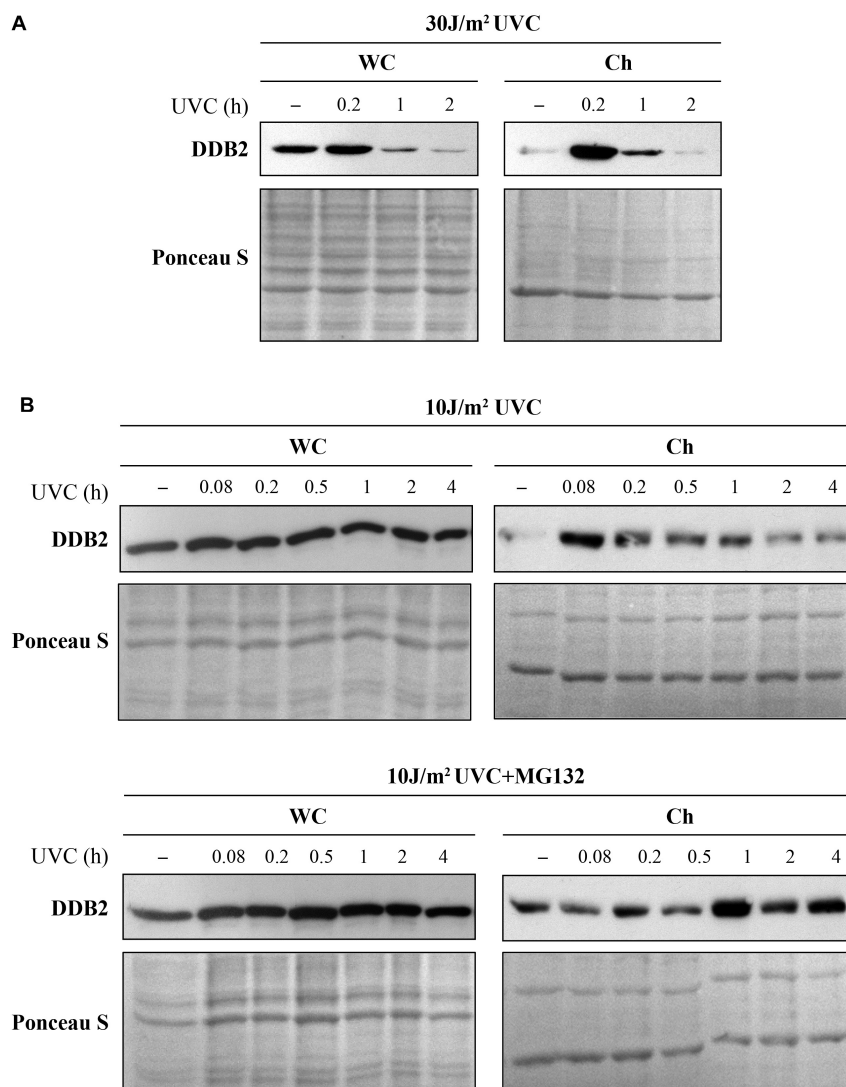
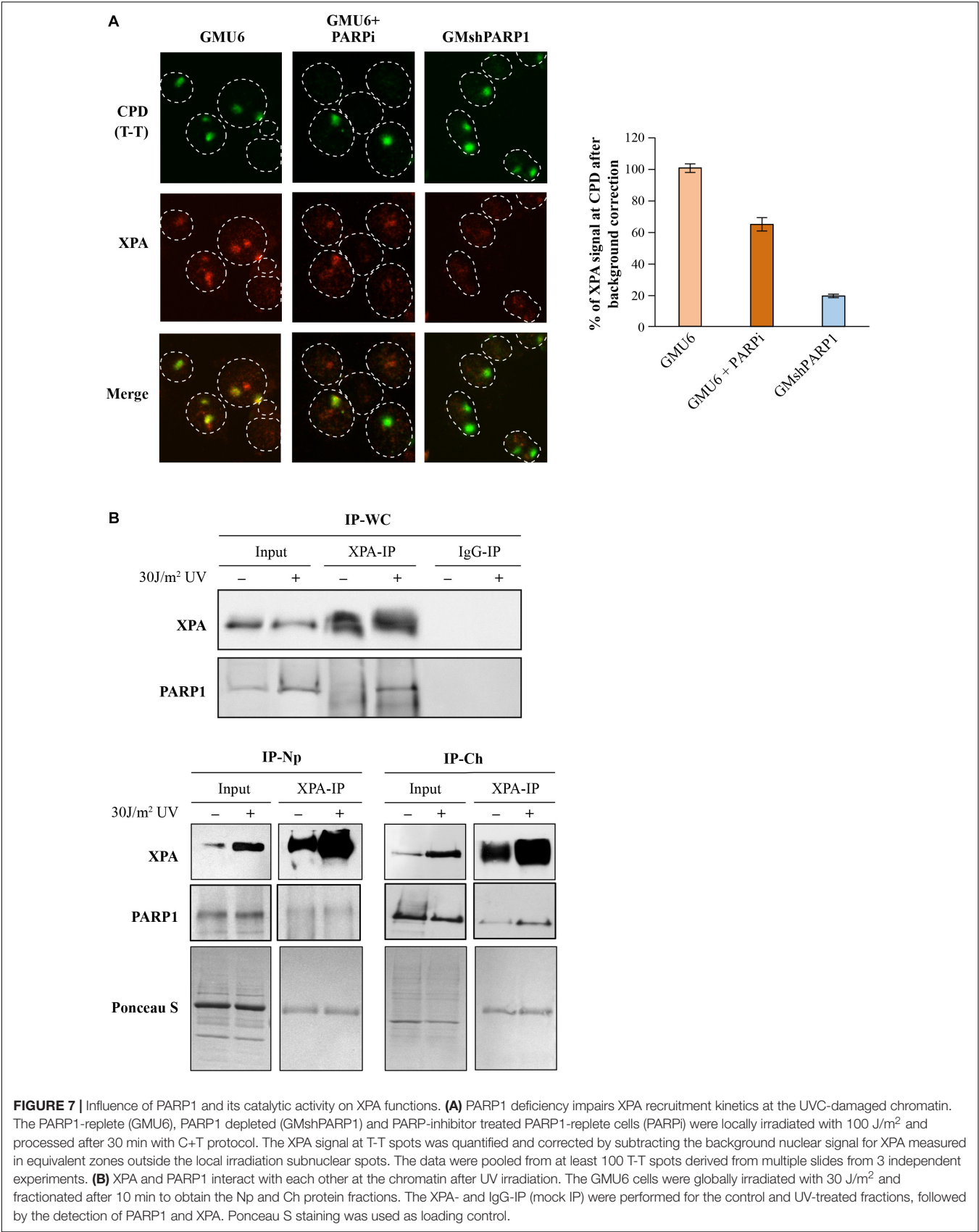


FIGURE 6 | Ultraviolet radiation irradiation induces the DDB2 ubiquitination and degradation. The GMU6 cells were globally irradiated with UVC at 30 **(A)** or 10 J/m² **(B)**, and fractionated at the indicated time. The whole cell extracts and chromatin-bound protein fractions were separated on SDS-PAGE and immunoblotted for DDB2. Where specified in panel **(B)**, 10 μ M of the proteasome inhibitor MG132 was added 1 h before irradiation with 10 J/m² UVC to show that the time-dependent depletion of DDB2 in Ch fraction is a result of proteasomal degradation of PTM-modified DDB2. Ponceau S staining was used as loading control.

Coupling cell fractionation with Western blot analysis offers complementary information on how the PTM affect the protein functions, whereas this information is not available if one uses total cell extracts for immunoblot analyses. Using proteasome and protein synthesis inhibitors in time course studies or modifying the acceptor sites to prevent a specific PTM, one can shed light on the role of a specific PTM in the repair process. For example, it is known that UVC (10 J/m²) induced binding of DDB2 to DNA lesions results in its auto-ubiquitination and proteasomal degradation by 3–6 h, followed by a complete recovery in 24 h (Rapic-Otrin et al., 2002; Puumalainen et al., 2014; Matsumoto et al., 2015). However, the extent of DDB2 degradation depends on the dose of UV. In cells irradiated with 30 J/m² UVC, we noted reduction in DDB2 levels from

1 h after irradiation in both, whole cell extract and chromatin bound fraction of GMU6 cells (**Figure 6A**). However, in the cells exposed to lower doses of UVC (10 J/m²), the reduction in DDB2 levels from 0.2 to 4 h was visible in the chromatin-bound protein fractions but not in the whole cell extracts (**Figure 6B**, top panel) indicating that the DDB2 reduction is closely associated with its interaction with damaged DNA (Rapic-Otrin et al., 2002). To determine whether the reduction in DDB2 level at chromatin is due to just its departure from the lesion site or if there is a concomitant proteasomal degradation of the protein, we added the proteasome inhibitor MG132 to the medium 1 h before irradiation with UVC (10 J/m²). We noted that addition of proteasome inhibitor made no difference to the levels of DDB2 in the whole cell extracts



up to 4 h, but suppressed its departure from the lesion site in the chromatin-bound protein fraction (**Figure 6B**, bottom panel). These results indicate that the reduction in the DDB2 signal at chromatin from 1 to 4 h is caused by proteasome-mediated DDB2 degradation, which confirms similar conclusion drawn in the previous studies (Rapic-Otrin et al., 2002; El-Mahdy et al., 2006).

Combination of Above Protocols to Study the Influence of PARP1 on Function of XPA in NER

Influence of PARP1 on the kinetics of recruitment of XPA by improved local irradiation technique

XPA was the first NER protein shown to interact with PAR (Pleschke et al., 2000) and the biochemical studies revealed the presence of a PAR specific binding domain in its C-terminal region (Fischer et al., 2014). To test the influence of the PARP1 and PAR on XPA kinetics, GMU6 cells were locally irradiated with UVC 30 min after treatment with PARP inhibitor, PJ-34, followed by the *in situ* extraction and immunofluorescence labeling. The accumulation of XPA at the local irradiation site was significantly reduced in the presence of PARP inhibitor and abrogated when PARP1 was depleted by shRNA, demonstrating that the recruitment of XPA to DNA lesion in NER is dependent on PARP1 and its catalytic activity (**Figure 7A** and chart).

Sub-cellular Fractionation as a Method to Detect UVC-induced Interaction of XPA with PARP1 in Chromatin-bound Fraction

The pulldown of a known NER protein followed by mass spectrometry identification of the interactome of this protein is the most exploited technique for identification of new players in NER. Many repair proteins form complexes with other proteins at the lesion site after induction of DNA damage (Boeing et al., 2016; Zhen and Yu, 2018), but some of them also interact with other proteins in the absence of any DNA damage (Isabelle et al., 2010). Therefore, immunoblotting and proteomic analyses of co-IP's protein partners in each of the sub-cellular fractions (and not in the total cellular extracts) can provide information about the protein-protein interactions, such as when and where these interactions occur inside the cell, whether they do so before or after irradiation, whether they occur at the lesion site on the chromatin or outside in the nucleoplasm.

To exemplify, we examined the mode of interaction of XPA with PARP1 in different sub-nuclear compartments of GM skin fibroblasts before and after DNA damage (**Figure 7B**). The XPA-IP of the whole cell extracts revealed a significant UV-induced increase in the interaction between XPA and PARP1 (**Figure 7B**, top panel), which has also been reported earlier (King et al., 2012). PARP1 is a nuclear protein, therefore to identify the sub-nuclear compartment in which this complex was formed, we performed IP of nucleoplasmic and chromatin bound fraction. The IP of nucleoplasmic fraction of control and UV-irradiated GMU6 cells with XPA did not pull down PARP1, whereas the

same IP performed in chromatin fraction revealed the UV-induced association of XPA with PARP1 (**Figure 7B**, bottom panels). Thus, the interaction of XPA with PARP1 in the whole cell extracts takes place only at the chromatin level. This limited interaction of PARP1 with XPA is in contrast with our previously reported interaction between PARP1 and XPC, which is not only at the chromatin after irradiation, but also before irradiation in both the nucleoplasm and chromatin extracts (Robu et al., 2017). In summary, the study of interaction of an NER protein with other protein partners, when carried out in the appropriate sub-cellular or sub-nuclear fraction would be far more informative in deciphering the role of their interactions on each other's function in NER.

The repair of damaged DNA is crucial for maintaining genome integrity. A plethora of proteins are recruited to the lesion site in an orchestrated fashion to detect and remove the damage from chromatin. Although we have a good knowledge of minimal set of protein required for GGR, it is not yet clear how the NER recognition proteins can find the lesions so quickly in the compact structure of chromatin (Kusakabe et al., 2019). Moreover, it is becoming increasingly evident from the current research in the field of DNA repair that “accessory” proteins, i.e., those not involved in carrying out the core biochemical reactions of the repair, play multiples roles at the damage site, including modification of the chromatin environment to make it more accessible to the core repair proteins (Puumalainen et al., 2016; Gsell et al., 2020). Thus, the local UVC irradiation and *in situ* extraction coupled with indirect immunofluorescence and sub-cellular fractionation techniques allow us to gain more insights into the factors influencing the trafficking of NER proteins to and from damage site, identification of new factors and/or PTM involved in NER and understanding of the molecular mechanisms of this repair pathway. The protocols described here are centered on NER pathway, but they can be suitably modified to study a variety of biological processes, ranging from other DNA repair pathways to cell cycle and cell death.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

MR and RS performed the experiments and analyzed the data. All authors conceived and designed the project, wrote and revised the manuscript, and read and approved the submitted version.

FUNDING

This work was supported by the Discovery Grant (RGPIN 2016-05868) and the Discovery Accelerator Supplement Grant (RGPA-492875-2016) from the Natural Sciences and Engineering Research Council of Canada to GS.

REFERENCES

- Aboussekhra, A., Biggerstaff, M., Shivji, M. K., Vilpo, J. A., Moncollin, V., Podust, V. N., et al. (1995). Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell* 80, 859–868. doi: 10.1016/0092-8674(95)90289-9
- Aydin, O. Z., Marteiijn, J. A., Ribeiro-Silva, C., Rodriguez Lopez, A., Wijgers, N., Smeenk, G., et al. (2014). Human ISWI complexes are targeted by SMARCA5 ATPase and SLIDE domains to help resolve lesion-stalled transcription. *Nucleic Acids Res.* 42, 8473–8485. doi: 10.1093/nar/gku565
- Balajee, A. S., and Geard, C. R. (2001). Chromatin-bound PCNA complex formation triggered by DNA damage occurs independent of the ATM gene product in human cells. *Nucleic Acids Res.* 29, 1341–1351. doi: 10.1093/nar/29.6.1341
- Barkauskaite, E., Jankevicius, G., and Ahel, I. (2015). Structures and mechanisms of enzymes employed in the synthesis and degradation of PARP-dependent protein ADP-ribosylation. *Mol. Cell* 58, 935–946. doi: 10.1016/j.molcel.2015.05.007
- Boeing, S., Williamson, L., Encheva, V., Gori, I., Saunders, R. E., Instrell, R., et al. (2016). Multiomic analysis of the UV-induced DNA damage response. *Cell Rep.* 15, 1597–1610. doi: 10.1016/j.celrep.2016.04.047
- Britton, S., Coates, J., and Jackson, S. P. (2013). A new method for high-resolution imaging of Ku foci to decipher mechanisms of DNA double-strand break repair. *J. Cell Biol.* 202, 579–595. doi: 10.1083/jcb.201303073
- Dantuma, N. P., and Van Attikum, H. (2016). Spatiotemporal regulation of posttranslational modifications in the DNA damage response. *EMBO J.* 35, 6–23. doi: 10.15252/embj.201592595
- Dinant, C., De Jager, M., Essers, J., Van Cappellen, W. A., Kanaar, R., Houtsmuller, A. B., et al. (2007). Activation of multiple DNA repair pathways by sub-nuclear damage induction methods. *J. Cell Sci.* 120, 2731–2740. doi: 10.1242/jcs.004523
- Dutto, I., Cazzalini, O., Stivala, L. A., and Prosperi, E. (2017). An improved method for the detection of nucleotide excision repair factors at local UV DNA damage sites. *DNA Repair.* 51, 79–84. doi: 10.1016/j.dnarep.2017.01.005
- El-Mahdy, M. A., Zhu, Q., Wang, Q. E., Wani, G., Praetorius-Ibba, M., and Wani, A. A. (2006). Cullin 4A-mediated proteolysis of DDB2 protein at DNA damage sites regulates in vivo lesion recognition by XPC. *J. Biol. Chem.* 281, 13404–13411. doi: 10.1074/jbc.m511834200
- Fei, J., Kaczmarek, N., Luch, A., Glas, A., Carell, T., and Naegeli, H. (2011). Regulation of nucleotide excision repair by UV-DDB: prioritization of damage recognition to internucleosomal DNA. *PLoS Biol.* 9:e1001183. doi: 10.1371/journal.pbio.1001183
- Fischer, J. M., Popp, O., Gebhard, D., Veith, S., Fischbach, A., Beneke, S., et al. (2014). Poly(ADP-ribose)-mediated interplay of XPA and PARP1 leads to reciprocal regulation of protein function. *FEBS J.* 281, 3625–3641. doi: 10.1111/febs.12885
- Friedberg, E. C. (2003). DNA damage and repair. *Nature* 421, 436–440.
- Giordano, C. N., Yew, Y. W., Spivak, G., and Lim, H. W. (2016). Understanding photodermatoses associated with defective DNA repair: syndromes with cancer predisposition. *J. Am. Acad. Dermatol.* 75, 855–870. doi: 10.1016/j.jaad.2016.03.045
- Groisman, R., Polanowska, J., Kuraoka, I., Sawada, J., Saijo, M., Drapkin, R., et al. (2003). The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage. *Cell* 113, 357–367. doi: 10.1016/s0092-8674(03)00316-7
- Gsell, C., Richly, H., Coin, F., and Naegeli, H. (2020). A chromatin scaffold for DNA damage recognition: how histone methyltransferases prime nucleosomes for repair of ultraviolet light-induced lesions. *Nucleic Acids Res.* 48, 1652–1668. doi: 10.1093/nar/gkz1229
- He, Z., and Ingles, C. J. (1997). Isolation of human complexes proficient in nucleotide excision repair. *Nucleic Acids Res.* 25, 1136–1141. doi: 10.1093/nar/25.6.1136
- Hoetelmans, R. W., Prins, F. A., Cornelese-Ten Velde, I., Van Der Meer, J., Van De Velde, C. J., and Van Dierendonck, J. H. (2001). Effects of acetone, methanol, or paraformaldehyde on cellular structure, visualized by reflection contrast microscopy and transmission and scanning electron microscopy. *Appl. Immunohistochem. Mol. Morphol.* 9, 346–351. doi: 10.1097/00129039-200112000-00010
- Houtsmuller, A. B., Rademakers, S., Nigg, A. L., Hoogstraten, D., Hoeijmakers, J. H., and Vermeulen, W. (1999). Action of DNA repair endonuclease ERCC1/XPF in living cells. *Science* 284, 958–961. doi: 10.1126/science.284.5416.958
- Isabelle, M., Moreel, X., Gagne, J. P., Rouleau, M., Ethier, C., Gagne, P., et al. (2010). Investigation of PARP-1, PARP-2, and PARG interactomes by affinity-purification mass spectrometry. *Proteome Sci.* 8:22. doi: 10.1186/1477-5956-8-22
- Katsumi, S., Kobayashi, N., Imoto, K., Nakagawa, A., Yamashina, Y., Muramatsu, T., et al. (2001). In situ visualization of ultraviolet-light-induced DNA damage repair in locally irradiated human fibroblasts. *J. Invest. Dermatol.* 117, 1156–1161. doi: 10.1046/j.0022-202x.2001.01540.x
- Kaufmann, S. H., Brunet, G., Talbot, B., Lamarr, D., Dumas, C., Shaper, J. H., et al. (1991). Association of poly(ADP-ribose) polymerase with the nuclear matrix: the role of intermolecular disulfide bond formation, RNA retention, and cell type. *Exp. Cell Res.* 192, 524–535. doi: 10.1016/0014-4827(91)90072-3
- Kim, M. Y., Zhang, T., and Kraus, W. L. (2005). Poly(ADP-ribosylation) by PARP-1: 'PAR-laying' NAD⁺ into a nuclear signal. *Genes Dev.* 19, 1951–1967. doi: 10.1101/gad.1331805
- King, B. S., Cooper, K. L., Liu, K. J., and Hudson, L. G. (2012). Poly(ADP-ribose) contributes to an association between Poly(ADP-ribose)polymerase-1 and *Xeroderma pigmentosum* complementation group A in nucleotide excision repair. *J. Biol. Chem.* 287, 39824–39833. doi: 10.1074/jbc.m112.393504
- Kusakabe, M., Onishi, Y., Tada, H., Kurihara, F., Kusao, K., Furukawa, M., et al. (2019). Mechanism and regulation of DNA damage recognition in nucleotide excision repair. *Genes Environ.* 41:2.
- Luijsterburg, M. S., Goedhart, J., Moser, J., Kool, H., Geverts, B., Houtsmuller, A. B., et al. (2007). Dynamic in vivo interaction of DDB2 E3 ubiquitin ligase with UV-damaged DNA is independent of damage-recognition protein XPC. *J. Cell Sci.* 120, 2706–2716. doi: 10.1242/jcs.008367
- Luijsterburg, M. S., Lindh, M., Acs, K., Vrouwe, M. G., Pines, A., Van Attikum, H., et al. (2012). DDB2 promotes chromatin decondensation at UV-induced DNA damage. *J. Cell Biol.* 197, 267–281. doi: 10.1083/jcb.201106074
- Luijsterburg, M. S., Von Bornstaedt, G., Gourdin, A. M., Politi, A. Z., Mone, M. J., Warmerdam, D. O., et al. (2010). Stochastic and reversible assembly of a multiprotein DNA repair complex ensures accurate target site recognition and efficient repair. *J. Cell Biol.* 189, 445–463. doi: 10.1083/jcb.200909175
- Marteijn, J. A., Lans, H., Vermeulen, W., and Hoeijmakers, J. H. (2014). Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat. Rev. Mol. Cell Biol.* 15, 465–481. doi: 10.1038/nrm3822
- Maser, R. S., Monsen, K. J., Nelms, B. E., and Petrini, J. H. (1997). hMre11 and hRad50 nuclear foci are induced during the normal cellular response to DNA double-strand breaks. *Mol. Cell Biol.* 17, 6087–6096. doi: 10.1128/mcb.17.10.6087
- Matsumoto, S., Fischer, E. S., Yasuda, T., Dohmae, N., Iwai, S., Mori, T., et al. (2015). Functional regulation of the DNA damage-recognition factor DDB2 by ubiquitination and interaction with *xeroderma pigmentosum* group C protein. *Nucleic Acids Res.* 43, 1700–1713. doi: 10.1093/nar/gkv038
- Mirzoeva, O. K., and Petrini, J. H. (2001). DNA damage-dependent nuclear dynamics of the Mre11 complex. *Mol. Cell Biol.* 21, 281–288. doi: 10.1128/mcb.21.1.281-288.2001
- Moné, M. J., Bernas, T., Dinant, C., Goedvree, F. A., Manders, E. M., Volker, M., et al. (2004). In vivo dynamics of chromatin-associated complex formation in mammalian nucleotide excision repair. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15933–15937. doi: 10.1073/pnas.0403664101
- Moné, M. J., Volker, M., Nikaido, O., Mullenders, L. H., Van Zeeland, A. A., Verschure, P. J., et al. (2001). Local UV-induced DNA damage in cell nuclei results in local transcription inhibition. *EMBO Rep.* 2, 1013–1017. doi: 10.1093/embo-reports/kve224
- Mullenders, L. H. F. (2018). Solar UV damage to cellular DNA: from mechanisms to biological effects. *Photochem. Photobiol. Sci.* 17, 1842–1852. doi: 10.1039/c8pp00182k
- Pascal, J. M., and Ellenberger, T. (2015). The rise and fall of poly(ADP-ribose): an enzymatic perspective. *DNA Repair.* 32, 10–16. doi: 10.1016/j.dnarep.2015.04.008

- Pines, A., Mullenders, L. H., Van Attikum, H., and Luijsterburg, M. S. (2013). Touching base with PARPs: moonlighting in the repair of UV lesions and double-strand breaks. *Trends Biochem. Sci.* 38, 321–330. doi: 10.1016/j.tibs.2013.03.002
- Pines, A., Vrouwe, M. G., Martijn, J. A., Typas, D., Luijsterburg, M. S., Cansoy, M., et al. (2012). PARP1 promotes nucleotide excision repair through DDB2 stabilization and recruitment of ALC1. *J. Cell Biol.* 199, 235–249. doi: 10.1083/jcb.201112132
- Pleschke, J. M., Kleczkowska, H. E., Strohm, M., and Althaus, F. R. (2000). Poly(ADP-ribose) binds to specific domains in DNA checkpoint proteins. *J. Biol. Chem.* 275, 40974–40980. doi: 10.1074/jbc.m006520200
- Polo, S. E., and Jackson, S. P. (2011). Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications. *Genes Dev.* 25, 409–433. doi: 10.1101/gad.2021311
- Pryde, F., Khalili, S., Robertson, K., Selfridge, J., Ritchie, A. M., Melton, D. W., et al. (2005). 53BP1 exchanges slowly at the sites of DNA damage and appears to require RNA for its association with chromatin. *J. Cell Sci.* 118, 2043–2055. doi: 10.1242/jcs.02336
- Purohit, N. K., Robu, M., Shah, R. G., Geacintov, N. E., and Shah, G. M. (2016). Characterization of the interactions of PARP-1 with UV-damaged DNA in vivo and in vitro. *Sci. Rep.* 6:19020.
- Puumalainen, M. R., Lessel, D., Ruthemann, P., Kaczmarek, N., Bachmann, K., Ramadan, K., et al. (2014). Chromatin retention of DNA damage sensors DDB2 and XPC through loss of p97 segregase causes genotoxicity. *Nat. Commun.* 5:3695.
- Puumalainen, M. R., Ruthemann, P., Min, J. H., and Naegeli, H. (2016). *Xeroderma pigmentosum* group C sensor: unprecedented recognition strategy and tight spatiotemporal regulation. *Cell Mol. Life Sci.* 73, 547–566. doi: 10.1007/s00018-015-2075-z
- Rapic-Otrin, V., Mclenigan, M. P., Bisi, D. C., Gonzalez, M., and Levine, A. S. (2002). Sequential binding of UV DNA damage binding factor and degradation of the p48 subunit as early events after UV irradiation. *Nucleic Acids Res.* 30, 2588–2598. doi: 10.1093/nar/30.11.2588
- Ray Chaudhuri, A., and Nussenzweig, A. (2017). The multifaceted roles of PARP1 in DNA repair and chromatin remodelling. *Nat. Rev. Mol. Cell Biol.* 18, 610–621. doi: 10.1038/nrm.2017.53
- Robu, M., Shah, R. G., Petitclerc, N., Brind'amour, J., Kandan-Kulangara, F., and Shah, G. M. (2013). Role of poly(ADP-ribose) polymerase-1 in the removal of UV-induced DNA lesions by nucleotide excision repair. *Proc. Natl. Acad. Sci. U.S.A.* 110, 1658–1663. doi: 10.1073/pnas.1209507110
- Robu, M., Shah, R. G., Purohit, N. K., Zhou, P., Naegeli, H., and Shah, G. M. (2017). Poly(ADP-ribose) polymerase 1 escorts XPC to UV-induced DNA lesions during nucleotide excision repair. *Proc. Natl. Acad. Sci. U.S.A.* 114, E6847–E6856. doi: 10.1073/pnas.1209507110
- Scharer, O. D. (2013). Nucleotide excision repair in eukaryotes. *Cold Spring Harb. Perspect. Biol.* 5:a012609. doi: 10.1101/cshperspect.a012609
- Sugasawa, K., Ng, J. M., Masutani, C., Iwai, S., Van Der Spek, P. J., Eker, A. P., et al. (1998). *Xeroderma pigmentosum* group C protein complex is the initiator of global genome nucleotide excision repair. *Mol. Cell* 2, 223–232. doi: 10.1016/s1097-2765(00)80132-x
- Sugasawa, K., Okuda, Y., Saijo, M., Nishi, R., Matsuda, N., Chu, G., et al. (2005). UV-induced ubiquitylation of XPC protein mediated by UV-DDB-ubiquitin ligase complex. *Cell* 121, 387–400. doi: 10.1016/j.cell.2005.02.035
- Svejstrup, J. Q., Wang, Z., Feaver, W. J., Wu, X., Bushnell, D. A., Donahue, T. F., et al. (1995). Different forms of TFIIH for transcription and DNA repair: holo-TFIIH and a nucleotide excision repairosome. *Cell* 80, 21–28. doi: 10.1016/0092-8674(95)90447-6
- Taylor, S. C., and Posch, A. (2014). The design of a quantitative western blot experiment. *Biomed. Res. Int.* 2014:361590.
- Ticli, G., and Prosperi, E. (2019). In situ analysis of DNA-Protein complex formation upon radiation-induced DNA damage. *Int. J. Mol. Sci.* 20:5736. doi: 10.3390/ijms20225736
- Van Den Boom, V., Citterio, E., Hoogstraten, D., Zotter, A., Egly, J. M., Van Cappellen, W. A., et al. (2004). DNA damage stabilizes interaction of CSB with the transcription elongation machinery. *J. Cell Biol.* 166, 27–36. doi: 10.1083/jcb.200401056
- Van Der Weegen, Y., Golan-Berman, H., Mevissen, T. E. T., Apelt, K., Gonzalez-Prieto, R., Goedhart, J., et al. (2020). The cooperative action of CSB, CSA, and UVSSA target TFIIH to DNA damage-stalled RNA polymerase II. *Nat. Commun.* 11:2104.
- Vermeulen, W. (2011). Dynamics of mammalian NER proteins. *DNA Repair.* 10, 760–771. doi: 10.1016/j.dnarep.2011.04.015
- Vermeulen, W., and Foustier, M. (2013). Mammalian transcription-coupled excision repair. *Cold Spring Harb. Perspect. Biol.* 5:a012625. doi: 10.1101/cshperspect.a012625
- Vodenicharov, M. D., Ghodgaonkar, M. M., Halappanavar, S. S., Shah, R. G., and Shah, G. M. (2005). Mechanism of early biphasic activation of poly(ADP-ribose) polymerase-1 in response to ultraviolet B radiation. *J. Cell Sci.* 118, 589–599. doi: 10.1242/jcs.01636
- Volker, M., Mone, M. J., Karmakar, P., Van Hoffen, A., Schul, W., Vermeulen, W., et al. (2001). Sequential assembly of the nucleotide excision repair factors in vivo. *Mol. Cell* 8, 213–224. doi: 10.1016/s1097-2765(01)00281-7
- Wang, Q., Zhu, Q., Wani, M. A., Wani, G., Chen, J., and Wani, A. A. (2003). Tumor suppressor p53 dependent recruitment of nucleotide excision repair factors XPC and TFIIH to DNA damage. *DNA Repair.* 2, 483–499. doi: 10.1016/s1568-7864(03)00002-8
- Wang, Q. E., Zhu, Q., Wani, G., El-Mahdy, M. A., Li, J., and Wani, A. A. (2005). DNA repair factor XPC is modified by SUMO-1 and ubiquitin following UV irradiation. *Nucleic Acids Res.* 33, 4023–4034. doi: 10.1093/nar/gki684
- Wienholz, F., Zhou, D., Turkyilmaz, Y., Schwertman, P., Tresini, M., Pines, A., et al. (2019). FACT subunit Spt16 controls UVSSA recruitment to lesion-stalled RNA Pol II and stimulates TC-NER. *Nucleic Acids Res.* 47, 4011–4025. doi: 10.1093/nar/gkz055
- Zhen, Y., and Yu, Y. (2018). Proteomic analysis of the downstream signaling network of PARP1. *Biochemistry* 57, 429–440. doi: 10.1021/acs.biochem.7b01022

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a shared affiliation with the authors at the time of review.

Copyright © 2020 Robu, Shah and Shah. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Approaches to Study Native Chromatin-Modifying Complex Activities and Functions

Maxime Galloy^{1,2†}, Catherine Lachance^{1,2†}, Xue Cheng^{1,2}, Félix Distéfano-Gagné^{1,2}, Jacques Côté^{1,2*‡} and Amelie Fradet-Turcotte^{1,2*‡}

OPEN ACCESS

Edited by:

Christoph Franz Kurat,
Ludwig Maximilian University
of Munich, Germany

Reviewed by:

Thomas M. Vondriska,
University of California, Los Angeles,
United States
Tuncay Baubec,
University of Zurich, Switzerland

*Correspondence:

Amelie Fradet-Turcotte
amelie.fradet-turcotte@
crchudequebec.ulaval.ca
Jacques Côté
jacques.cote@
crchudequebec.ulaval.ca

[†]These authors have contributed
equally to this work and share first
authorship

[‡]These authors have contributed
equally to this work and share last
authorship

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 23 June 2021

Accepted: 24 August 2021

Published: 16 September 2021

Citation:

Galloy M, Lachance C, Cheng X,
Distéfano-Gagné F, Côté J and
Fradet-Turcotte A (2021) Approaches
to Study Native Chromatin-Modifying
Complex Activities and Functions.
Front. Cell Dev. Biol. 9:729338.
doi: 10.3389/fcell.2021.729338

¹ St-Patrick Research Group in Basic Oncology, Oncology Division, Centre Hospitalier Universitaire (CHU)
de Québec-Université Laval Research Center, Québec, QC, Canada, ² Department of Molecular Biology, Medical
Biochemistry and Pathology, Laval University Cancer Research Center, Université Laval, Québec, QC, Canada

The modification of histones—the structural components of chromatin—is a central topic in research efforts to understand the mechanisms regulating genome expression and stability. These modifications frequently occur through associations with multisubunit complexes, which contain active enzymes and additional components that orient their specificity and read the histone modifications that comprise epigenetic signatures. To understand the functions of these modifications it is critical to study the enzymes and substrates involved in their native contexts. Here, we describe experimental approaches to purify native chromatin modifiers complexes from mammalian cells and to produce recombinant nucleosomes that are used as substrates to determine the activity of the complex. In addition, we present a novel approach, similar to the yeast anchor-away system, to study the functions of essential chromatin modifiers by quickly inducing their depletion from the nucleus. The step-by-step protocols included will help standardize these approaches in the research community, enabling convincing conclusions about the specificities and functions of these crucial regulators of the eukaryotic genome.

Keywords: chromatin, histone modification, native protein complexes, recombinant nucleosomes, nuclear protein depletion

INTRODUCTION

Understanding chromatin structure and function has been a focus of intense research for decades. It is now well established that the chromatin plays primary active roles regulating genome-related processes, including gene-specific expression, DNA damage repair, and DNA replication during cell division (Allis and Jenuwein, 2016). Every few years, breakthrough discoveries propel the field into new territories, expanding our understanding of the molecular mechanisms involved and providing crucial insights on human diseases like cancer, an illness with combined genetic and epigenetic origins (Shen and Laird, 2013). This has led to an impressive surge in the development of epigenetic approaches to treat cancer (Brien et al., 2016). The research breakthroughs that have allowed these leaps have been conceptual, such as the identification of histone writers/readers/erasers, as well as technological, including approaches to study genome organization and modifications with increasing precision.

The epigenetic language written on the chromatin has been studied for over 20 years, and new mechanistic understandings continuously emerge. The post-translational modifications (PTMs) of conserved histone residues are diverse, and their combinations form specific signatures that can be read by effectors. As the same PTM can have opposing functions when added on different residues/histones, it is crucial to clearly identify and characterize the chromatin modifiers regulating each specific modified histone residue, as well as the reader proteins that recognize them. Historically, many studies have used histone peptides or free histones to identify the enzymes responsible for specific PTMs, as well as recombinant proteins instead of enzymes in their native forms. However, the specificities identified in these studies are often significantly different from those observed *in vivo* or when assayed *in vitro* on chromatin substrates (reviewed in Lalonde et al., 2014). In fact, true native specificity can only be reproduced in a test tube using native substrates, nucleosomal histones/chromatin, and enzymes in their physiological contexts, including any associated factors. It is clearly established that the histone residue specificity of chromatin modifiers can be determined by associated factors within large enzyme-containing protein complexes (Lalonde et al., 2014). In addition, associated reader modules within these large complexes further regulate the specificity of the modifiers by mediating crosstalk between different histone modifications (Lalonde et al., 2014).

In parallel, determining chromatin modifier specificity *in vivo* can be difficult because of their possible indirect effects on the modifications of other specific residues. Indeed, many histone PTMs are regulated by crosstalk between histone modifications, a phenomenon that can induce secondary effects when histone modifiers are depleted or knocked out (McGinty et al., 2008; Lalonde et al., 2014; Jacquet et al., 2016; Wojcik et al., 2018; Zhang and Kutateladze, 2019). This is a major reason why both *in vivo* and *in vitro* experiments are required to truly understand the intricate molecular mechanisms regulating chromatin modifiers.

APPROACHES TO CHARACTERIZE CHROMATIN-MODIFYING ENZYME ACTIVITIES AND FUNCTIONS

Over the years, multiple approaches to purify native protein complexes from mammalian nuclear extracts (NEs) have been developed. Several chromatin remodeling and modifying complexes have been efficiently purified to near homogeneity by introducing transgenes encoding tagged components. Tandem affinity purification (TAP) was developed more than 20 years ago in yeast and quickly transferred to higher eukaryotes (Rigaut et al., 1999; Puig et al., 2001). Different tag combinations have been tested, some of which provide high specificity and efficient elutions/high yields in native conditions. For example, the FLAG epitope has remained front and center for many years (Hopp et al., 1988; Einhauer and Jungbauer, 2001). However, one of the main issues with this approach is the formation of artefactual associations due to overexpression of the transgene compared to the physiological protein level (Gibson et al., 2013;

Lalonde et al., 2014). This often occurs when proteins with significant homology (paralogs) to specific subunits or even the tagged protein are expressed. Keeping expression at a near-physiological level is possible, by aiming for single-copy genome integration and using different promoter strengths [e.g., using retroviral vectors at low multiplicity of infection or the Flp-InTM system (Thermo Fisher Scientific)]. Random genome integration can also create problems due to well-established positional effects on gene expression depending on the local chromatin state. Thus, achieving near-physiological expression can require isolating and characterizing multiple clones. The development of safe harbors for transgene integration has great advantages because the different cell lines created are isogenic, as in lower eukaryotes. The AAVS1 and ROSA26 loci have been frequently used for this purpose, and recent genome editing tools have made this endeavor quite straightforward (DeKolver et al., 2010; Sadelain et al., 2012; Dalvai et al., 2015). Of course, the gold standard of reproducing physiological conditions is to tag the endogenous gene and use it to purify the native chromatin modifying complex. This can now also be achieved, thanks to the power of clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9-based genome editing (Dalvai et al., 2015).

To obtain native chromatin substrates, several protocols are available to purify endogenous chromatin from human cell lines (Côté et al., 1995; Utley et al., 1996; Schnitzler, 2000). Substrates are produced by micrococcal nuclease digestion and can be fractionated in different lengths, from long oligonucleosomes to mononucleosomes. They have several advantages: they already contain the vast majority of known histone marks, reproduce the natural substrate, and enable binding specificity analysis. However, these marks are present at low stoichiometries on each residue, and the DNA sequences associated with the nucleosomes are heterogeneous. Using recombinant histones to assemble nucleosome core particles (NCPs) is a powerful alternative to producing chromatin substrates with defined DNA and histone status. The “Widom” sequence is extremely efficient for nucleosome assembly/positioning using recombinant histone octamers (Lowary and Widom, 1998; Thåström et al., 1999). The resulting recombinant nucleosomes are homogeneous, and their main advantage is that they can be enzymatically or chemically modified to introduce a specific histone PTM at a given residue with high stoichiometry. This allows very clear analysis of crosstalk occurring between different histone marks, via specific reader proteins, or within chromatin remodeling/modifying complexes. Over the past few years, the use of purified native chromatin modifying complexes and recombinant nucleosomes has produced exquisite high-resolution 3D structures of chromatin-bound complexes, which have provided crucial mechanistic insights (Poepsel et al., 2018; Patel et al., 2019; Farnung et al., 2020; Wagner et al., 2020; Wang et al., 2020).

In vivo loss of function analysis of chromatin modifiers has relied mostly on RNA interference-mediated depletion and mouse gene knockouts. However, these can create indirect and/or downstream effects that mask the primary role of the enzyme being studied (discussed in Lalonde et al., 2014). For example, depletion of chromatin modifiers such as the HBOI

acetyltransferase, impact cell cycle progression into S phase (Doyon et al., 2006; Miotto and Struhl, 2010; Havasi et al., 2013; Feng et al., 2016). As many chromatin modifiers are essential for cell viability or normal growth and because several histone marks are regulated during the cell cycle (Ma et al., 2015), changes due to of knockdown of such factors can mislead the investigator into linking a chromatin modification to a specific enzyme (Lalonde et al., 2014). The development of rapid depletion approaches, as used in lower eukaryotes, can bypass these problems. The auxin-inducible degron has proved popular; however, the tag is known to affect protein stability even in the absence of auxin, which can be alleviated by expressing transport inhibitor response 1 (Nishimura et al., 2020; Yesbolatova et al., 2020). The recently described degradation TAG does not seem to have the same problem, and efficiently and suddenly targets proteins to the proteasome (Nabet et al., 2018).

In this method article, we present up-to-date detailed protocols related to the approaches discussed above. First, we describe the use of the $3 \times \text{FLAG-2} \times \text{Strep}$ tag fused to a gene of interest (either expressed from *AAVS1* or endogenously tagged) to purify native chromatin modifying complexes for biochemical/enzymatic assays. Second, we provide a step-by-step protocol for the production of recombinant mono- and di-nucleosomes, the latter having particular potential since recent studies have highlighted the importance of linker DNA in the mechanisms of many remodelers and modifiers (Poepsel et al., 2018; Bhardwaj et al., 2020). Finally, we present a new rapid depletion approach for mammalian cells, inspired by the yeast anchor-away system (Haruki et al., 2008). In this system, a nuclear protein is tagged with a peptide that can be induced to bind a tagged endogenous ribosomal protein, which acts as an anchor to rapidly export it to the cytosol, crippling its nuclear function.

STEP-BY-STEP PROCEDURES

Purification of Endogenously Tagged NuA4/TIP60 Complex (Figure 1)

Tagging an endogenous gene to purify its encoded protein has the advantage of reflecting its physiological expression and regulation. Insertion of the small $3 \times \text{FLAG-2} \times \text{Strep}$ tag using CRISPR/Cas9 and donor DNA is efficient, does not require a selection marker, and creates very limited sequence perturbation of the 5' or 3' untranslated regions. NuA4/TIP60 complex is a highly conserved multisubunit complex essential for cell cycle progression, gene-specific regulation, and DNA repair (Doyon et al., 2004). Here, the E1A binding protein p400 (EP400) NuA4/TIP60 was selected for endogenous tagging as it was successfully reported to be efficient to purify the native complex (Dalvai et al., 2015). However, the tagging of many other subunits of this complex has been successfully used to purified native complex (Ikura et al., 2000; Doyon et al., 2004, 2006; Dalvai et al., 2015; Jacquet et al., 2016). The generation of endogenously tagged EP400 in K562 cells was performed essentially as described in the third protocol presented here, and was reported in Dalvai et al. (2015). In this section, we describe the protocol for TAP

of $3 \times \text{FLAG-2} \times \text{Strep}$ -tagged native complexes from nuclear extracts (NEs) (Figure 1A).

1. NE preparation from K562 cells expressing endogenous $3 \times \text{FLAG-2} \times \text{Strep-EP400}$

A. Large-scale expansion of K562 cells

- i. Grow 1–3 L of K562 cells in spinner flasks with gentle agitation in basal RPMI medium supplemented with 25 mM HEPES-NaOH pH 7.4.
- ii. Monitor cell growth and viability daily, carefully maintaining cultures between 2×10^5 and 8×10^5 cells/mL.
- iii. Harvest cells at or slightly below 8×10^5 cells/mL, before they reach confluency.
- iv. Pellet the cells in a preparative centrifuge ($700 \times g$, 10 min, 4°C). Resuspend and pool the pellets in 50 mL of cold phosphate-buffered saline (PBS). Centrifuge ($700 \times g$, 10 min, 4°C), place the pellet on ice, and immediately prepare the NE.

B. NE preparation (as previously described in Abmayr et al., 2006)

- i. Wash the cell pellets by adding four packed cell volumes of hypotonic buffer (10 mM HEPES pH 7.9, 1.5 mM MgCl_2 , and 10 mM KCl, with 0.2 mM phenylmethylsulfonyl fluoride (PMSF) and 0.5 M dithiothreitol (DTT) added just before use). Centrifuge ($1,900 \times g$, 5 min, 4°C), remove the supernatant quickly, and resuspend well with three packed cell volumes of hypotonic buffer. Incubate on ice 10 min.
- ii. Transfer the cells to a glass Dounce homogenizer with a type B pestle. Homogenize by douncing 15 times, then centrifuge ($3,500 \times g$, 10 min, 4°C).
- iii. Collect the supernatant (i.e., S-100 cytoplasmic extract) and estimate the packed nuclear volume of the pellet using the gradations on the conical tube.
- iv. Add half the packed nuclear volume of low salt buffer (10 mM HEPES pH 7.9, 1.5 mM MgCl_2 , 20 mM KCl, 25% glycerol, 0.2 mM ethylenediaminetetraacetic acid (EDTA), with 0.2 mM PMSF and 0.5 M DTT added fresh) and resuspend well with gentle vortexing. Then, extract the soluble proteins by adding half the packed nuclear volume of high salt buffer (10 mM HEPES pH 7.9, 1.5 mM MgCl_2 , 1.2 M KCl, 25% glycerol, and 0.2 mM EDTA, with 0.2 mM PMSF and 0.5 M DTT added fresh) dropwise with gentle vortexing.
- v. Dounce twice using a Dounce homogenizer with a type B pestle and incubate on a nutator for 30 min at 4°C .
- vi. Pellet the extracted nuclei by ultracentrifugation ($100,000 \times g$, 1 h, 4°C). Quickly transfer the supernatant (i.e., the NE) to a new Falcon tube.
- vii. Snap freeze the NE in liquid nitrogen and store at -80°C .

2. TAP of EP400 (as described in Doyon and Côté, 2016)

- i. Thaw the NE* on ice, adjust to 0.1% Tween-20 (using a 10% stock), and centrifuge ($40,000 \times g$, 1 h, 4°C).

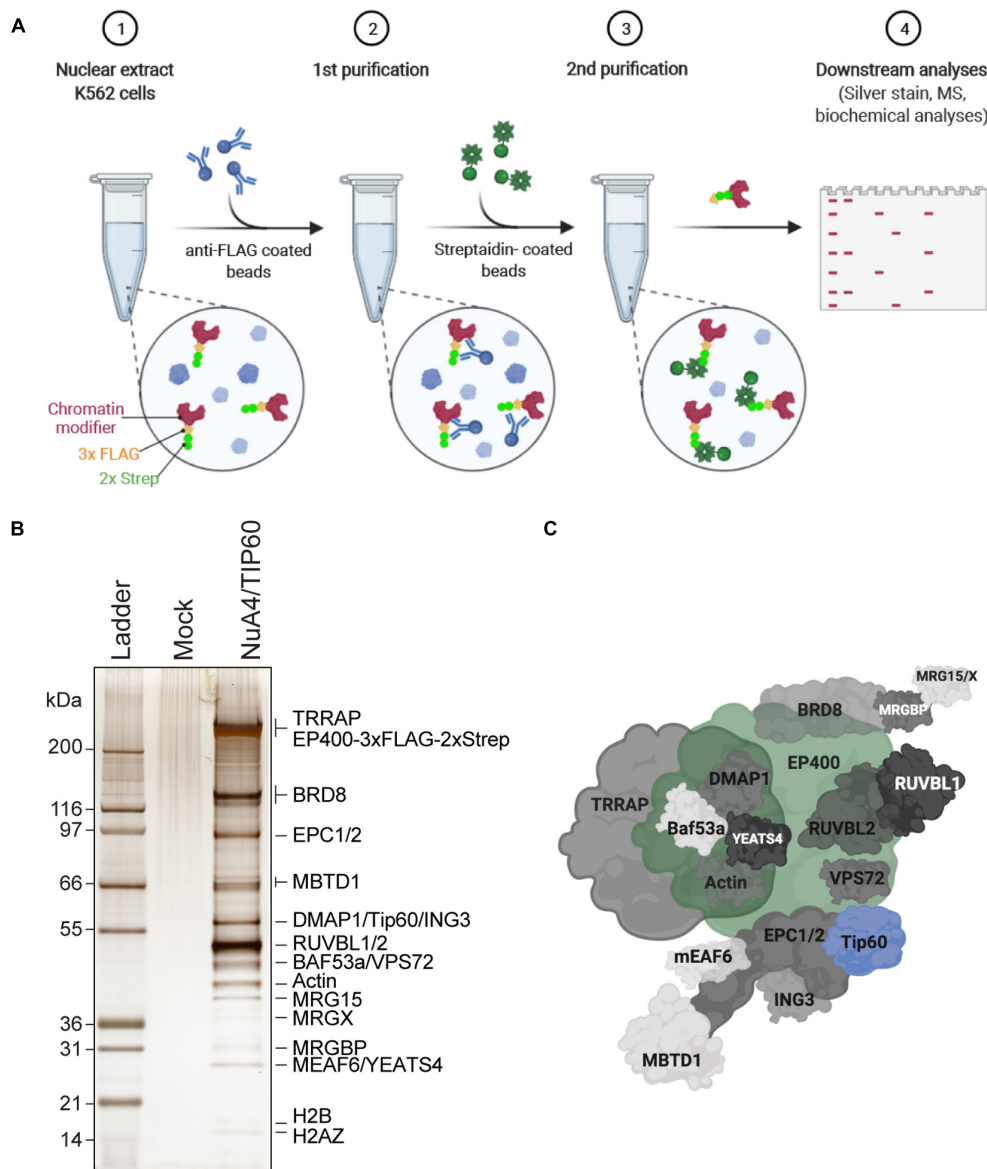


FIGURE 1 | Purification of endogenously tagged native chromatin modifying enzymes **(A)** TAP. Step 1: Nuclear extracts are isolated from engineered K562 cell lines (here, C-terminally 3 × FLAG-2 × Strep-tagged endogenous EP400 was used as an example). Step 2: FLAG-Strep-tagged proteins are immunoprecipitated using anti-FLAG-coated beads and eluted with 3 × FLAG peptides. Step 3: Eluted FLAG-Strep-tagged proteins are isolated using streptavidin-coated beads and eluted with biotin. Step 4: Purified FLAG-Strep-tagged proteins are either analyzed by silver staining and/or MS, or directly used in biochemical assays. Created with BioRender.com. **(B)** Silver-stained purified NuA4/TIP60 native complex after TAP of 3 × FLAG-2 × Strep-EP400. K562 cells expressing the 3 × FLAG-2 × Strep tag alone (mock) were used as a control purification. **(C)** Schematic representation of the NuA4/TIP60 complex. The TIP60 subunit, which encompasses the acetyltransferase activity of the complex, is colored in blue and the tagged subunit EP400 is colored in green. Created with BioRender.com.

- ii. Preclear the NE using 250 μ L Sepharose CL-6B resin prewashed with PBS and equilibrated with TAP buffer (20 mM HEPES-KOH pH 7.9, 300 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, and 10% glycerol, with 10 mM sodium butyrate, 10 mM β -glycerophosphate, 1 mM PMSF, 5 mM NaF, 100 μ M orthovanadate, 2 μ g/mL leupeptin, 2 μ g/mL pepstatin, and 5 μ g/mL aprotinin added fresh) in a 10 mL Poly-Prep chromatography column. Collect the precleared NE in a 15 mL tube.
- iii. Add 250 μ L anti-FLAG M2 affinity beads to the precleared NE and incubate for 2 h at 4°C with rotation.
- iv. Transfer to a 10 mL Poly-Prep chromatography column, harvest the flowthrough (FLAG-FT), and pass it through the column again. Wash the beads with 40 column volumes (CVs) of TAP buffer, then 40 CVs of TAP wash buffer #1 (20 mM HEPES-KOH pH 7.9, 300 mM KCl, 0.1% Tween-20, and 10% glycerol, with 1 mM DTT, 10 mM sodium butyrate, 10 mM β -glycerophosphate, 1 mM PMSF, 5 mM NaF, 100 μ M

- orthovanadate, 2 $\mu\text{g/mL}$ leupeptin, 2 $\mu\text{g/mL}$ pepstatin, and 5 $\mu\text{g/mL}$ aprotinin added fresh), followed by 40 CVs of TAP wash buffer #2 (20 mM HEPES-KOH pH 7.9, 150 mM KCl, 0.1% Tween-20, and 10% glycerol, with 1 mM DTT, 10 mM sodium butyrate, 10 mM β -glycerophosphate, 1 mM PMSF, 5 mM NaF, 100 μM orthovanadate, 2 $\mu\text{g/mL}$ leupeptin, 2 $\mu\text{g/mL}$ pepstatin, and 5 $\mu\text{g/mL}$ aprotinin added fresh).
- v. Transfer the beads in a 1.5 mL Eppendorf tube. Use TAP wash buffer #2 to rinse the column and collect all the beads. Centrifuge ($239 \times g$, 5 min, 4°C) and carefully remove the supernatant.
 - vi. Elute the complex with 2.5 CVs of TAP wash buffer #2 supplemented with 200 $\mu\text{g/mL}$ 3 \times FLAG peptide for 1 h at 4°C on a rotator.
 - vii. Centrifuge ($250 \times g$, 5 min, 4°C) and carefully transfer the supernatant into a Micro Bio-Spin column placed in a 2 mL microcentrifuge tube. Centrifuge ($250 \times g$, 1 min, 4°C) to collect the eluate. Collect a 15 μL sample to resolve by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE; FLAG first elution).
 - viii. Repeat steps vi and vii.
 - ix. Pool the FLAG elutions and add 125 μL Strep-Tactin Superflow Sepharose affinity matrix prewashed with 1 mL PBS followed by 1 mL TAP wash buffer #2. Incubate for 1 h at 4°C on a rotator.
 - x. Centrifuge ($250 \times g$, 5 min, 4°C) and remove the flowthrough.
 - xi. Wash the beads three times with 1 mL TAP wash buffer #2.
 - xii. Elute the complex with 1 CV of TAP wash buffer #2 supplemented with 5 mM D-biotin for 1 h at 4°C on a rotator.
 - xiii. Centrifuge ($250 \times g$, 5 min, 4°C) and carefully transfer the supernatant into a Micro Bio-Spin column placed in a 2 mL microcentrifuge tube. Centrifuge ($250 \times g$, 1 min, 4°C) to collect the eluate. Aliquot a 15 μL sample for SDS-PAGE (biotin elution).
 - xiv. Repeat steps xiii and xiv.
 - xv. Aliquot the purified complex. Snap freeze in liquid nitrogen and keep at -80°C .

* NEs should always be kept on ice, and all purification steps should be performed at 4°C in a cold room.

3. Analysis of EP400 complex subunits by SDS-PAGE and silver staining (Figures 1B,C)

- i. Load 15 μL of the FLAG and biotin elution fractions on a NuPage/Bolt 4–12% Bis-Tris gel and run for 42 min at 200 V in MOPS SDS running buffer (50 mM MOPS, 50 mM Tris, 0.1% SDS, and 1 mM EDTA).
- ii. Incubate the gel for 1 h in 50% methanol, then for 30 min in 10% MeOH/7% acetic acid, and finally for 30 min in 10% glutaraldehyde.
- iii. Wash the gel at least four times for 30 min with ultrapure water and let it soak in water overnight.
- iv. The next day, change the ultrapure water 2–3 times before staining.
- v. Incubate the gel with 5 $\mu\text{g/mL}$ DTT for 30 min.

- vi. Stain the gel with 0.1% (w/v) silver nitrate in ultrapure water for 30 min.
 - vii. Rinse twice with ultrapure water.
 - viii. Condition the gel with two rapid washes in carbonate developing solution (283 mM sodium carbonate, 0.0185% formaldehyde), then incubate in exactly 160 mL of solution.
 - ix. Stop the reaction when proper staining is attained by adding 8 mL of 2.3 M citric acid.
- ### 4. Sample preparation for mass spectrometry (MS) analysis
- i. Load 50 μL samples of the biotin elutions onto a Bolt 12% Bis-Tris gel and let run for >1 cm (approximately 4 min) at 200V to retain all the proteins in one band.
 - ii. Wash the gel briefly with ultrapure water and incubate it for 30 min in 10% MeOH/7% acetic acid.
 - iii. Incubate the gel O/N in Sypro Ruby gel stain.
 - iv. Rinse twice with ultrapure water and cut out the protein bands under UV light.
 - v. Rinse the bands twice with 70% acetonitrile and store the samples at -80°C .

CHARACTERIZATION OF *in vitro* HISTONE ACETYLTRANSFERASE (HAT) ACTIVITY (Figure 2)

Chromatin modifiers like NuA4/TIP60 efficiently acetylate histone tails as well as purified histones; however, the use of oligonucleosomes purified from cell nuclei (Côté et al., 1995; Utley et al., 1996) or reconstituted from recombinant histones revealed different specificities toward these substrates (Lalonde et al., 2014). In this section, we describe a robust method used to reconstitute mono- and dinucleosomes from recombinant histones purified from *Escherichia coli*. The method is adapted (Dyer et al., 2003) and can be used to assemble recombinant nucleosome core particles (rNCPs) containing different types of histones. Here, rNCPs were reconstituted using untagged human H2A/H2B and *Xenopus laevis* H3/H4. Cysteine 110 of histone H3 was replaced with an alanine to block undesired cysteine alkylation in assays where analogs are used to label other residues (e.g., H4 K20Cme) (Simon, 2010). Histones tagged on their N-termini with tags such as histidine (His) or FLAG can also be used to assemble rNCPs; however, we have noted that on H2A, these tags interfere with NuA4/TIP60 activity *in vitro* (data not shown). These observations are in line with previous studies showing that NuA4 binds the N-terminal tail of histone H2A (Huang and Tan, 2013). Here, two DNA fragments were used for reconstitution: a 151 bp fragment that contains a single copy of the 601 DNA used to assemble mononucleosomes (Lowary and Widom, 1998; Thåström et al., 1999; Dyer et al., 2003) and a 388 bp fragment containing two copies of 601 separated by a 48-bp linker DNA to assemble dinucleosomes (Kato et al., 2017; Machida et al., 2018). Note that an array of DNA fragments can be used to reconstitute rNCPs, facilitating structural studies of nucleosome assembly (Engelholm et al., 2009; Muthurajan et al., 2016). In this section, we also describe the main steps required to assess the activity of the purified NuA4/TIP60 complex *in vitro*.

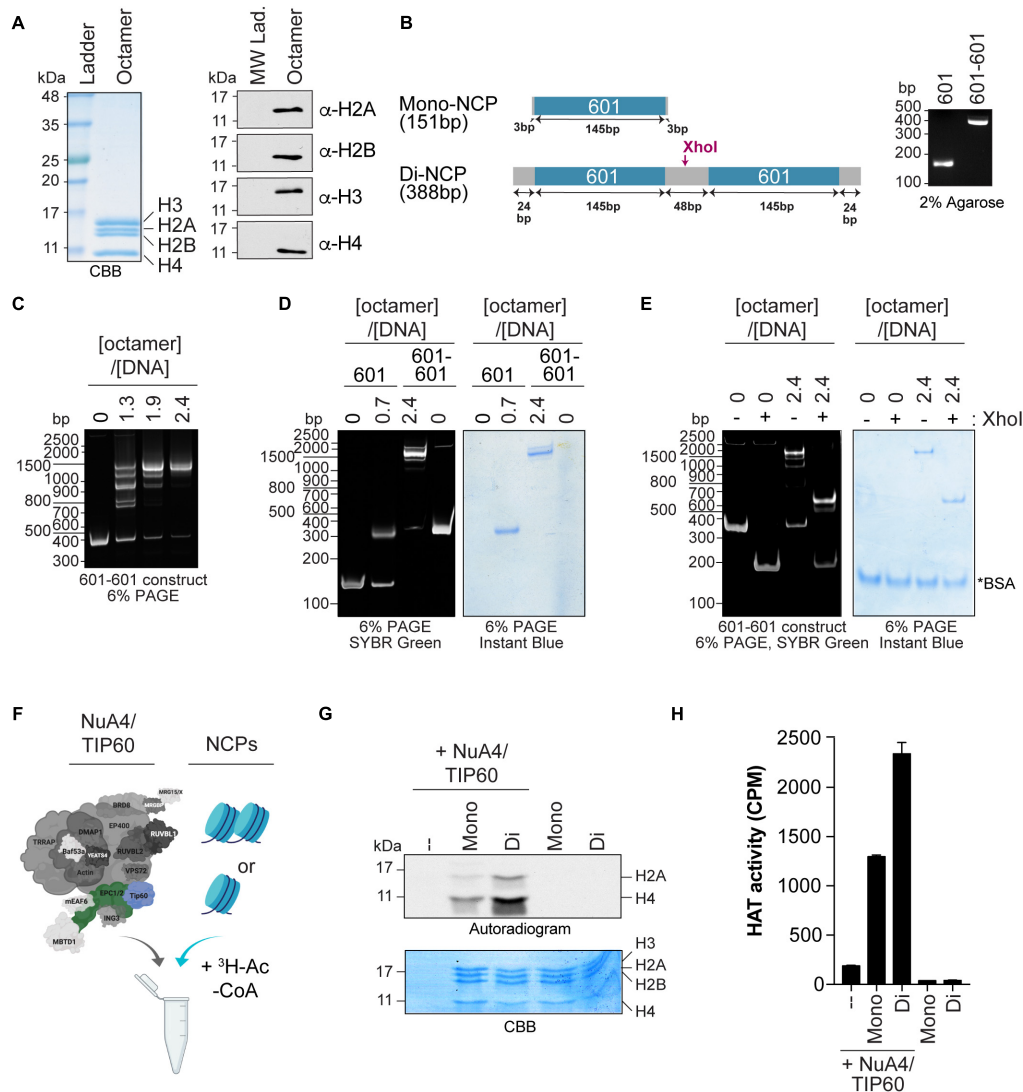


FIGURE 2 | Quantifying the HAT activity of native complexes on recombinant nucleosomes (**A,B**). Analysis of the octamers and 601 DNA used to reconstitute NCPs. (**A**) Coomassie Brilliant Blue (CBB) staining (left panel) and Western blot analyses (right panel) of refolded octamers. (**B**) Schematic representation of the Widom 601 DNA sequence used to wrap mono- and di-NCPs (left panel). The position of the XhoI cleavage site used to validate di-NCP assembly is indicated. Purified DNA obtained from a large prep purification was resolved on a 2% agarose gel and stained with ethidium bromide (right panel). (**C-E**) Analysis of 200 ng of NCP reconstitution on native gels stained with either SYBR green or InstantBlue Coomassie Stain. 200 ng of DNA alone (0) is used as a control. (**C**) Refolding of di-NCPs using the indicated octamer:DNA ratios. (**D**) Mono- and di-NCPs refolded at their optimal octamer:DNA ratios. (**E**) Di-NCPs treated with XhoI for 1 h. The cleaved form of the 601-601 DNA is observed at approximately 200 bp. *BSA used in the reaction. (**F-H**) HAT assay using NuA4/TIP60 complex purified using EPC1-tagged subunit and reconstituted NCPs. The TIP60 subunit, which encompasses the acetyltransferase activity of the complex, is colored in blue and the tagged subunit EPC1 is colored in green. Created with BioRender.com. (**F**) Schematic representation of the HAT assay. (**G**) Autoradiogram showing the HAT activity of 1 μ L of purified native NuA4/TIP60 complex on histone H2A and H4. The gel was kept with the X-ray film at -80°C for 4 days. Coomassie Brilliant Blue (CBB) staining was used to show equal loading. (**H**) Quantification of a representative experiment ($n = 3$). HAT reactions were spotted on P81 filters and analyzed with a scintillation counter. Error bars indicate the range between technical duplicate samples.

using gel- and liquid-based assays (Côté et al., 1995; Utley et al., 1996).

1. Reconstitution of mono- and dinucleosome-containing rNCPs

Histone purification (A), refolding of core histones into octamers (B), large-scale purification of 601 and 601-601

DNA (C), and nucleosome reconstitution were performed as previously described (Dyer et al., 2003).

A. Histone purification

- i. On day 1, transform bacterial expression vectors for histones (human H2A and H2B in pET15b, *X. laevis* H3_{C110A} in pET3d and *X. laevis* H4 in pET3a) into BL21

- (DE3) competent cells and plate on lysogeny broth (LB) plates with ampicillin.
- ii. On day 2, resuspend colonies and inoculate 1 L of LB plus 100 µg/mL of ampicillin. When cells reach an optical density at 600 nm of 0.5–0.8, induce with 0.4 mM isopropyl β-D-thiogalactoside (IPTG) at 37°C. *An aliquot can be taken prior to addition of IPTG as a negative control for histone induction.*
 - iii. Pellet the induced bacteria via centrifugation (6,000 × g, 15 min, 4°C) 3 h (H2A, H2B, and H3C110A) or 1.5 h (H4) post-induction, transfer the bacteria to a 50 mL centrifuge tube in 35 mL of histone wash buffer (50 mM Tris pH 7.5, 100 mM NaCl, and 1 mM EDTA, with 1 mM benzamidine and 5 mM β-mercaptoethanol added fresh), and snap freeze in liquid nitrogen. *An aliquot can be taken prior to centrifugation to monitor histone induction.*
 - iv. To prepare inclusion bodies, thaw the bacteria in warm water and perform two rounds of freeze-thaw lysis. Add 1 mg/mL lysozyme, nutate 30 min at 4°C, and sonicate until the lysate loses its viscosity. Add histone wash buffer to a total volume of 100 mL, centrifuge (12,000 × g, 20 min, 4°C), resuspend the pellet in 75 mL histone wash buffer + 1% Triton X-100, centrifuge (12,000 × g, 20 min, 4°C), resuspend the pellet in 75 mL histone wash buffer, and centrifuge again (12,000 × g, 20 min, 4°C). Drain the pellet well for the next step. *Cell lysates and purified proteins should be kept on ice at all times unless stated otherwise.*
 - v. To unfold inclusion bodies, transfer the pellet to a centrifugation tube and dissolve in 260 µL dimethyl sulfoxide for 30 min at room temperature. Mince with a spatula twice during this time. Add 8 mL unfolding buffer (6 M guanidinium HCl and 20 mM Tris pH 7.5, with 5 mM DTT added fresh) and nutate for 1 h at room temperature. Centrifuge (23,000 × g, 10 min, room temperature) and retain the supernatant. Rinse the pellet with 1 mL unfolding buffer and centrifuge again (23,000 × g, 10 min, room temperature). Pool the supernatants and dialyze them in urea dialysis buffer (7 M urea, 10 mM Tris pH 8, 1 mM EDTA, and 100 mM NaCl, with 5 mM β-mercaptoethanol added fresh, 2 × 1 L for 3–4 h and 1 × in 1 L overnight*) using dialysis tubing with a cutoff of 3.5 kDa.
 - vi. To purify the histones, cation exchange chromatography is used. Rinse a clean HiTrap SP HP Sepharose column with 10 mL of water and equilibrate with 20 mL 0.22-µm filtered Buffer A (7 M urea and 20 mM Tris pH 8, with 5 mM β-mercaptoethanol added fresh) at 2 mL/min using a peristaltic pump. Load the dialyzed protein sample and wash with 30 mL Buffer A at 1.5 mL/min. Connect the SP column to an NGC Quest 10 Plus Chromatography System, ensuring that no air bubbles enter the system, and run a linear gradient over 25 CV of 0–100% 0.22-µm filtered Buffer B with a flow rate of 1.5 mL/min and a back-pressure limit of 0.28 MPa. Collect 1.5 mL fractions and monitor the absorbance at 280 nm (A_{280}) and conductivity (mS/cm) during elution. Proteins will elute according to their charges, with histones usually eluting at approximately 36 mL and 10 mS/cm.
 - vii. Resolve eluted fractions from the peaks by 15% SDS-PAGE and pool fractions containing purified histones (hH2A: 14.09 kDa, hH2B: 13.97 kDa, xH3: 15.4 kDa, xH4: 11.37 kDa). Dialyze in 2 mM β-mercaptoethanol in water (3 × 4 L for 3–4 h and 1 × in 4 L overnight) using dialysis tubing with a cutoff of 3.5 kDa.
 - viii. Centrifuge any precipitate, use A_{280} values to determine histone concentrations using extinction coefficients (hH2A: 4,470 M⁻¹ cm⁻¹, hH2B: 7,450 M⁻¹ cm⁻¹, xH3: 4,470 M⁻¹ cm⁻¹, xH4: 5,960 M⁻¹ cm⁻¹) and lyophilize 5 mg of dry histone per 15 mL centrifuge tube using a lyophilizer. Store lyophilized histones at –80°C.
* To reduce protein carbamylation by cyanate present in old urea, do not leave your protein in the urea buffer for more than 24 h and deionize the 7 M urea solution for 1 h using MB AG 501-X8 (D) resin prior to adding Tris, EDTA, and NaCl.
- B. Refolding of core histones into octamers**
- i. Unfold lyophilized histones by nutating them for 30 min at room temperature in 0.22-µm filtered unfolding buffer (20 mM Tris pH 7.5, 7 M guanidine-HCl, and 10 mM DTT added fresh) to a final concentration of ~2 mg/mL.
 - ii. Combine histones in equimolar ratios in unfolding buffer to a final concentration of ~1 mg/mL and dialyze in 650 mL refolding buffer (10 mM Tris pH 7.5, 2 M NaCl, 1 mM EDTA, and 5 mM β-mercaptoethanol added fresh) in dialysis tubing with a cutoff of 3.5 kDa, 2 × 4 h and 1 × overnight.
 - iii. Collect octamers in 15 mL centrifuge tubes, centrifuge to remove any precipitate (4,000 × g, 10 min, 4°C), and concentrate samples to <1 mL using Amicon Ultra-0.5 centrifugal filter units with a cutoff of 30 kDa (4,000 × g, 10 min, 4°C).
 - iv. To purify refolded octamers, load the sample on a S200 Superdex 16/60 FPLC column pre-equilibrated with 1.25 CV of 0.22-µm filtered refolding buffer using a 1 mL sample loop. Elute protein complexes at a flow rate of 1 mL/min and a back-pressure limit of 0.5 MPa, and collect 2.5 mL fractions. Monitor the A_{280} during elution. Octamers elute first, at approximately 62.5 mL, then H3-H4 tetramers and H2A-H2B dimers at approximately 70 and 82 mL, respectively.
 - v. Resolve eluted fractions from the peaks by 15% SDS-PAGE, pool fractions containing reconstituted octamers (the four histones should be present in equimolar ratios; **Figure 2A**), and dialyze in 2 mM β-mercaptoethanol in water using dialysis tubing with a cutoff of 3.5 kDa, 3 × 4 L for 3–4 h and 1 × in 4 L overnight.
 - vi. Concentrate to ≤15 mg/mL with an Amicon Ultra-0.5 centrifugal filter with a cutoff of 30 kDa (4,000 × g, 10 min, 4°C), use the A_{280} to determine the concentration using an extinction coefficient (octamer: 44,700 M⁻¹ cm⁻¹), and store at 4°C for 2–3 months or at –20°C in 50% v/v glycerol for years. Octamers stored in glycerol need to be dialyzed against fresh refolding buffer, and their

concentration should be re-evaluated prior to use in NCP reconstitution.

** Histones are difficult to quantify accurately. To attain a better idea of the yield, resolve them by SDS-PAGE and stain with Coomassie Brilliant Blue to see if the unfolded histones look equal. Unpaired histones will precipitate during dialysis and reduce the final yield.*

C. Large scale purification of 601 and 601-601 DNA (**Figure 2B**)

The 601-DNA (Widom DNA) (Lowary and Widom, 1998) and 601-601 DNA are obtained by digesting a 32×601 DNA plasmid (pGEM-3z/601) or E23-L48-E23 plasmid (Kato et al., 2017), which contain a repeated 147-bp nucleosome positioning sequence (Dyer et al., 2003) with EcoRV.

- i. Grow DH10 β cells transformed with a 601 or 601-601 DNA plasmid in 3× 800 mL of LB plus 100 μ g/mL of ampicillin at 37°C and purify the plasmids using a Qiagen GigaPrep kit according to the manufacturer's instructions.
- ii. Digest 3 mg of purified DNA at 37°C with 1,500 units of EcoRV in 15 mL of 1× NEB3 buffer [100 mM NaCl, 50 mM Tris-HCl pH 7.9, 10 mM MgCl₂, and 100 μ g/mL bovine serum albumin (BSA)] for 16 h to release the positioning sequence DNA. Verify complete digestion by resolving samples on a 1.5% agarose gel.
- iii. Transfer the reaction in a 50 mL conical tube. Precipitate backbone DNA (2.5 kbp) by adding 5.1 mL of 40% PEG6000, 2.5 mL of 5M NaCl and 0.15 mL of ultrapure water to reach final concentrations of 9% PEG6000 and 500 mM NaCl in a final volume of 22.50 mL. Incubate 4 h on ice, centrifuge (15,000 × g, 30 min, 4°C), carefully decant the supernatant, which contains the smaller DNA fragments (151 bp), and repeat the PEG precipitation for an additional 2 h. Collect the supernatant after centrifugation (15,000 × g, 30 min, 4°C). Verify the purity of the DNA fragments by resolving them on a 2% agarose gel.
- iv. For 11 mL of supernatant, precipitate the small DNA fragments by adding 27.5 mL of ice-cold 100% EtOH and 0.5 mL of 5M NaCl to reach final concentrations of 70% EtOH and 200 mM NaCl in a final volume of 39 mL. Incubate overnight at 4°C, centrifuge (15,000 × g, 30 min, 4°C), transfer the pellet to a 1.5 mL Eppendorf tube in 1 mL of ice-cold 70% EtOH, centrifuge (15,000 × g, 5 min, 4°C), remove the supernatant, dry the pellet for 15 min, and resuspend the pellet in 0.2 mL TE buffer (10 mM Tris pH 8.0, 1 mM EDTA). This method usually yields approximately 1 mg of purified nucleosome positioning sequence DNA (**Figure 2B**).

D. Nucleosome reconstitution

- i. Define the molar ratios required for reconstitution to minimize the presence of free DNA. Determine empirically the optimal molar ratios for mono- and dinucleosomes (here, a ratio of 0.7 and 2.4 octamers per DNA were used, respectively) (**Figures 2C,D**).

- ii. Combine 50 μ g of octamers (117,760 g/mol) with either 60 μ g of 601 DNA (151 bp, 99,660 g/mol) or 45 μ g of 601-601 DNA (386 bp, 254,760 g/mol). Final reagent concentrations should be 2 M KCl, ~ 0.7 mg/mL DNA, 10 mM Tris pH 7.5, 1 mM EDTA, and 1 mM DTT. Incubate for 30 min at 4°C, transfer to a 0.5 mL Slide-a-Lyzer with a cutoff of 10 kDa and dialyze against 2 L high salt reconstitution buffer (2 M KCl, 10 mM Tris pH 7.5, and 1 mM EDTA, with 1 mM DTT added fresh) with a decreasing salt gradient over 18 h at 4°C. The gradient is created by constantly pumping low salt reconstitution buffer (0.2 M KCl, 10 mM Tris pH 7.5, and 1 mM EDTA, with 1 mM DTT added fresh) into the 2 L beaker as described (Dyer et al., 2003). Transfer the samples to a fresh beaker containing 400 mL low salt buffer and dialyze for an additional 3 h.
- iii. Concentrate to \leq 1 mg/mL with an Amicon Ultra-0.5 centrifugal filter with a cutoff of 100 kDa (4,000 × g, 10 min, 4°C), determine the concentration using the A₂₈₀, and store at 4°C for 1–2 months. Recombinant nucleosomes can be stored at –80°C in 5–10% v/v glycerol for years.
- iv. Resolve 0.2 μ g of recombinant nucleosomes on 6% polyacrylamide gels (DNA retardation gels) in 0.2× TBE (18 mM Tris, 18 mM boric acid, and 0.4 mM EDTA, pH 8.0) according to the manufacturer's instructions (**Figure 2D**). Samples are prepared in 1× nucleosome loading buffer (8×: 40% sucrose, 0.1% bromophenol blue) and 0.2× TBE. DNA fragments and proteins are visualized by incubating the gel in 1× SYBR green in PBS and in InstantBlue™ Coomassie Stain, respectively, according to the manufacturers' instructions. Free 601 DNA, reconstituted mono-NCPs, 601-601 DNA, and reconstituted di-NCPs migrate at 151, 350, 350, and 1,500 bp, respectively.
- v. Digest 0.2 μ g of recombinant nucleosome with 5 units of XhoI (**Figure 2E**) in 10 μ L of 1× CutSmart buffer (50 mM potassium acetate, 20 mM Tris-acetate pH 7.9, 10 mM MgCl₂-acetate, and 100 μ g/mL BSA) at 37°C for 1 h. Resolve on 6% polyacrylamide as above. Cleaved di-NCPs migrate at 650 bp.

2. HAT assays

The activity of purified NuA4/TIP60 complex varies between preps. Thus, the amount of NuA4/TIP60 used in HAT assays need to be determined for each preps while doing liquid assays to obtain counts that are in the linear range of the scintillation counter.

E. Reactions (**Figure 2F**)

- i. Perform HAT reactions in a final volume of 15 μ L in a 1.5 mL Eppendorf tube. First, combine 0.5 μ g of reconstituted NCPs with 1 μ L of purified NuA4/TIP60 complex in HAT buffer (50 mM Tris-HCl pH 8, 10 mM sodium butyrate, 5% glycerol, and 0.1 mM EDTA, with 1 mM DTT added fresh). Calculate the KCl molarity in the reaction based on the amounts in the NCP and NuA4/TIP60 complex buffers, and add to a final

concentration of 50 mM if necessary. At this point, the reaction volume will be 13.75 μ L.

- ii. Incubate on ice 10 min.
- iii. Add 1.25 μ L (0.125 μ Ci) of H^3 -labeled acetyl-CoA (2.1 Ci/mmol) and incubate at 30°C for 30 min in a water bath.
- iv. At this point, the reaction can either be stopped to detect acetylated histones by SDS-PAGE (step 2B) or used in liquid HAT assays to quantify total HAT activity (step 2C).

F. Detection of acetylated histones (Figure 2G)

- i. To perform SDS-PAGE, quench the HAT reaction with 5 μ L of 4 \times Laemmli sample buffer and denature the sample for 5 min at 95°C.
- ii. Load samples on 15 or 18% SDS-PAGE and migrate for 75–200 min at 160 V, depending on the desired resolution.
- iii. Stain the gel with Coomassie Brilliant Blue for 30 min, destain four times for 20 min in 30% methanol and 10% acetic acid, and take an image of the gel.
- iv. Destain overnight, incubate the gel for 60 min in EN³HANCE, quickly rinse twice in 10% glycerol in water, and incubate for 30 min at room temperature. Dry the gel for 2 h 30 min at 60°C in a gel dryer, place it in an autoradiography cassette with an X-ray film, and store it at –80°C for 1 d to several weeks before developing.

G. Liquid HAT assays (Figure 2H)

- i. Microcentrifuge the reactions, then pipette them onto individual P81 phosphocellulose filter papers and air-dry for 30 min.
- ii. Wash away free H^3 -labeled acetyl-CoA with 50 mM carbonate buffer pH 9.2 (3.3 mM Na₂CO₃ and 47.7 mM NaHCO₃) three times for 5 min each, then perform an extra quick rinse with acetone. Air-dry for at least 10 min.
- iii. Place each P81 paper into a scintillation vial and add 5 mL of EcoLite (+) Liquid Scintillation Cocktail. Measure the H^3 counts [in counts per minute (CPM)] for 30 min with a scintillation counter.

ESTABLISHMENT OF A MAMMALIAN (m)ANCHOR-AWAY SYSTEM (Figure 3)

Functional studies of chromatin modifiers using knockdown/knockout approaches are well known to be often associated with undesired secondary effects on important biological processes, such as gene-specific transcription and the cell cycle progression (reviewed in Lalonde et al., 2014). To bypass this issue, we have developed a cellular system to rapidly remove a targeted protein from its usual cellular compartment based on chemically induced proximity. This system was first described as the anchor-away system in yeast (Haruki et al., 2008); however, the use of rapamycin in the system has limited its application in mammalian cells because of its toxicity, instability, and slow clearance. To adapt this method, we took advantage of the S-(+)-abscisic acid (ABA) plant stress pathway, in which the phytohormone ABA binds to pyrabactin resistance

(PYR)/PYR1-like (PYL)/regulatory component of ABA receptor family members (Figure 3A). This allows us to use ABA to induce proximity between the interacting complementary surface of PYL (PYLcs), fused to a protein of interest, and proteins fused to the complementary surface of ABA insensitive 1 (ABI1; ABI1cs) (Liang et al., 2011). We chose to fuse the ribosomal protein L13 (RPL13) to ABI1cs to take advantage of its shuttling from the nucleus to the cytoplasm like in the yeast anchor-away system. This would specifically induce the removal from the nucleus of a PYLcs-tagged protein upon ABA addition. CRISPR-Cas9 technology was used to endogenously tag the RPL13 C-terminus with ABI1cs in U2OS cells. Next, a transgene expressing PYL1cs fused to enhanced green fluorescent protein (eGFP) was integrated into the AAVS1 safe harbor locus (Hockemeyer et al., 2009; DeKolver et al., 2010; Lombardo et al., 2011), and depletion of PYLcs-eGFP from the nucleus was validated by immunofluorescence upon ABA treatment.

1. Endogenous tagging of the RPL13 C-terminus using CRISPR/Cas9 (as described in Dalvai et al., 2015)

A. Generation of single guide (sg)RNA and donor plasmid targeting RPL13 (Figure 3B)

- i. Endogenous tagging was performed essentially as described (Doyon and Côté, 2016). Scan your sequence using a web-based CRISPR design tool¹ to identify sgRNAs that cleave no more than 30 bp away from the stop codon. For RPL13, the target site was 5'-CTGATTCCAAGTCCCCAGGA-3' (Figure 3B). Generate an sgRNA containing an extra G at its 5' end to accommodate the transcription initiation requirements of the human U6 promoter. A BbsI restriction enzyme site is also added on each end to enable cloning into the pSpCas9(BB)-2A-Puro (pX459) V2.0 vector. Order non-phosphorylated oligonucleotides. Clone the annealed sgRNA into BbsI-digested pX459 V2.0 vector.
- ii. For the RPL13 donor plasmid, amplify the homology arms (left: 600 bp and right: 1,000 bp) by polymerase chain reaction (PCR) using genomic DNA isolated from K562 cells. Be sure to mutate the PAM sequence of the sgRNA target sites if it occurs within the homology arms. Assemble using the Zero Blunt TOPO cloning kit. Introduce the ABI1cs sequence (SV-ABAactDA plasmid) (Liang et al., 2011) with a C-terminal V5 tag into the pTOPO-RPL13 donor plasmid via PCR extension overlap using a Gibson Assembly Cloning Kit.

B. Generation of isogenic U2OS cell lines expressing endogenous RPL13-ABI1cs-V5 (Figures 3C,D).

- i. Maintain U2OS cells in McCoy's modified medium supplemented with 10% fetal bovine serum and penicillin-streptomycin at 37°C under 5% CO₂.
- ii. Electroporate 1×10^6 U2OS cells with 2 μ g of the donor plasmid, 1 μ g of the pX459 V2.0 plasmid expressing eSpCas9, and the gRNA using an Amaxa 4D-Nucleofector

¹<http://crispr.mit.edu/>

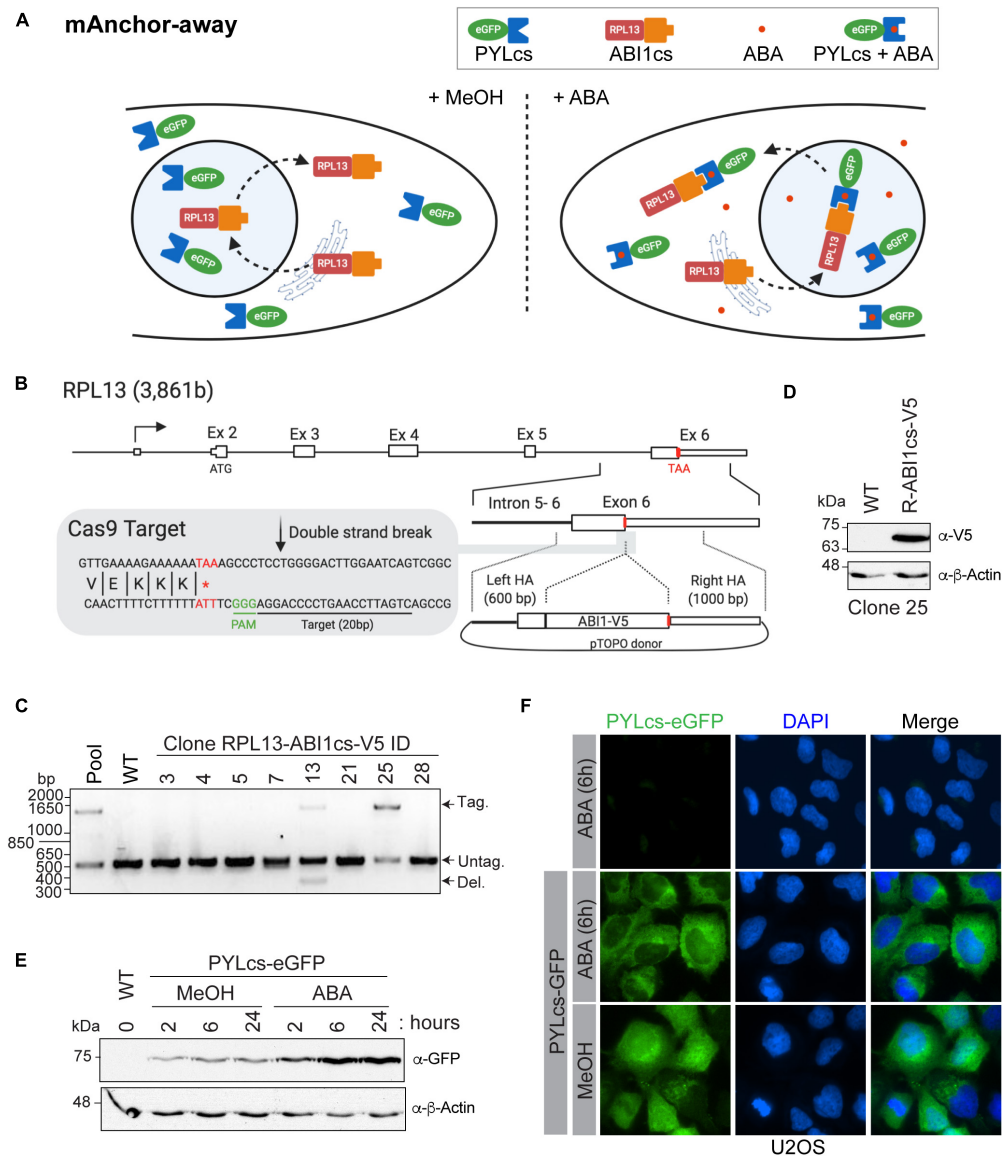


FIGURE 3 | A mammalian anchor-away system to quickly and efficiently deplete chromatin-modifying enzymes from the nucleus **(A)**. Schematic of ABA-induced translocation of a nuclear protein to the cytoplasm. In the absence (or presence) of ABA, the ABI1cs-RPL13 fusion protein constantly shuttles between the cytoplasm and the nucleus like most ribosomal proteins, transiting to the nucleolus to assemble ribosome particles with the ribosomal RNA and then going back to the cytoplasm. Upon the addition of ABA, the dimerization of PYLcs-tagged protein with the ABI1cs-RPL13 fusion protein triggers its rapid depletion from the nucleus to the cytoplasm, as “anchored away” by the RPL13 ribosomal protein. **(B)** Strategy to establish U2OS cell lines expressing RPL13-ABI1cs-V5 from the endogenous locus. The *RPL13* locus, the Cas9 cleavage site, and the donor construct are indicated. The sequence of the sgRNA targeting *RPL13* is underlined, and the stop codon is indicated in red. HA: homology arm. **(C,D)** Characterization of RPL13-ABI1cs-V5 isogenic cell lines. **(C)** Results of an out-out PCR-base assay conducted on U2OS genomic DNA to detect ABI1cs-V5 integration at the *RPL13* locus. The primers are located in the homology arms and yield a longer PCR product if ABI1cs-V5 is integrated (1,452 bp vs 510 bp for the wild-type allele). In panel **(D)**, whole-cell extracts of wild type U2OS cells and those expressing RPL13-ABI1cs-V5 (clone #25) were western blotted with an anti-V5 antibody to confirm RPL13-ABI1cs-V5 expression (top panel). β-actin was used as a loading control (bottom panel). **(E)** Whole-cell extracts of U2OS RPL13-ABI1cs-V5 cells stably expressing PYLcs-eGFP from the *A4VS1* locus. Samples were collected at the indicated times following treatment with 100 μM ABA (0.1% MeOH was used as a negative control). An anti-GFP antibody was used to detect PYLcs-eGFP (top panel) and β-actin was used as a loading control (bottom panel). **(F)** Immunofluorescence of U2OS cells expressing either RPL13-ABI1cs-V5 alone (clone #25, as a negative control) or RPL13-ABI1cs-V5 and PYLcs-eGFP. Cells were treated with either 100 μM ABA or 0.1% MeOH. DNA was stained with DAPI.

and an SE XL Kit, according to the manufacturer’s recommendations.

- iii. Expand and select clones by limiting dilution starting 3-d post-transfection in 96-well plates.

- iv. Extract genomic DNA with QuickExtract DNA extraction solution and amplify by PCR using the primers 5’-ACTTATGGCAGCGAACCTGA -3’ and 5’-ACCTCCCCACAAGAAAACCG -3’. Resolve on a 1%

agarose gel in TAE buffer (40 mM Tris, 1 mM EDTA, and 20 mM glacial acetic acid) to identify the ABI1cs-V5 insertion (**Figure 3C**). Sequence both alleles to validate accurate gene modification and confirm the absence of indels in the non-targeted allele.

- v. Confirm expression of RPL13-ABI1cs-V5 in selected clones by western blotting with an anti-V5 antibody (**Figure 3D**).

2. Generation of U2OS RPL13-ABI1cs-V5 cells expressing PYLcs-eGFP from the AAVS1 safe harbor locus

- i. Amplify PYLcs (SV-ABAactDA plasmid) (Liang et al., 2011) by PCR and clone it in the AAVS1 Puro PGK1 3 × FLAG Twin Strep plasmid using Gibson Assembly Cloning Kit. Amplify eGFP by PCR and replace the 3 × FLAG Twin Strep tag with it via Gibson cloning.
- ii. To introduce PYLcs-eGFP at the AAVS1 safe harbor locus via nuclease-driven targeted integration (Dalvai et al., 2015), electroporate 1×10^6 U2OS RPL13-ABI1cs-V5 cells with 1 µg of zinc finger nuclease expression vector (Hockemeyer et al., 2009) and 4 µg of the AAVS1 PYLcs-eGFP donor construct.
- iii. Subject cells to puromycin selection for 7 d, starting 3 d post-transfection.
- iv. Confirm PYLcs-eGFP expression in the enriched pool following a time-course with 100 µM ABA by western blotting with an anti-GFP antibody (**Figure 3E**).

3. Imaging of ABA-treated RPL13-ABI1cs-V5 U2OS cells

- i. Seed enriched pools of RPL13-ABI1cs U2OS cells with AAVS1-integrated PYLcs-eGFP in 6-well plates containing autoclaved coverslips.
- ii. At 60% confluency, treat cells with 100 µM ABA (Liang et al., 2011) dissolved in 0.1% methanol (or 0.1% methanol as a control) for 2–24 h.
- iii. Wash cells with PBS and fix them with 4% methanol-free formaldehyde for 15 min.
- iv. Wash four times with PBS and stain the nuclei using 4',6-diamidino-2-phenylindole (DAPI; 2 µg/mL). Mount the coverslips on microscope slides using Fluoromount G mounting medium.
- v. Capture images with a Nikon Ti Eclipse inverted fluorescence microscope equipped with a Hamamatsu Orca ER camera using 40× or 60× objectives (**Figure 3F**).

RESULTS AND DISCUSSION

TAP of Endogenously Tagged NuA4/TIP60 Complex

Studies of native chromatin modifying activities require good yields of purified intact complexes to be obtained by TAP. K562 cells are an excellent model line to purify endogenously tagged chromatin remodelers, as they are permissive to genome editing and tolerate the conditions required to isolate clones. Importantly, large volumes can be cultivated as suspension cultures, which is essential to purify the yields of chromatin

modifiers required to perform *in vitro* biochemical analyses and potentially even structural studies. Following the establishment of isogenic K562 cell lines expressing an endogenously TAP-tagged component of the NuA4/TIP60 complex, the FLAG and streptavidin portions of the TAP-tag are used to purify the entire complex in a stepwise manner (**Figure 1A**). In this specific example, the E1A binding protein p400 (EP400) subunit was used. Nuclear extracts were prepared from K562 cells expressing tagged EP400 as well as a line expressing the 3 × FLAG-2 × Strep tag only (mock). Copurifying proteins were resolved by SDS-PAGE and analyzed by silver staining (**Figure 1B**). The high sensitivity of silver-stained gels enabled the unambiguous identification of complex subunits, which were not observed in the mock sample. Mass spectrometry analysis of the purified complex confirmed the purification of the entire NuA4/TIP60 complex (**Figure 1C**).

TAP-based approaches have been used to efficiently purify NuA4/TIP60 complexes using an array of tagged subunits (Ikura et al., 2000; Doyon et al., 2004, 2006; Dalvai et al., 2015; Jacquet et al., 2016). In addition to endogenously tagged proteins, tagged chromatin modifier cDNAs integrated into a safe harbor genomic locus like AAVS1 can also be used for complex purification (Dalvai et al., 2015). The latter approach is a great alternative, as it allows the expression of near-physiological levels of the tagged protein, is straightforward, and offers high flexibility to study proteins that are difficult to tag at their endogenous locus. It can also be highly useful to compare panels of truncations and mutant proteins in an isogenic setting.

Characterizing Native Chromatin-Modifying Activities *in vitro*

Chromatin-modifying activities such as acetylation and methylation can be recapitulated *in vitro* using substrates such as peptides mimicking histone tails, purified histones (recombinant or native), rNCPs, and oligonucleosomes isolated from NEs. Of these, rNCPs provide a unique tool to study how nucleosome assembly and specific histone marks affect enzyme activity. Recently, chromatin-remodeling/modifying complexes were found to exhibit different specificity toward mononucleosomes and dinucleosomes, highlighting their higher-order structural specificity (Poepsel et al., 2018; Bhardwaj et al., 2020). We thus used rNCPs to assess whether the activity of the native NuA4/TIP60 complex is affected by the structural differences between the two types of rNCPs.

Mono- and dinucleosomes were reconstituted with octamers of core histone proteins assembled from purified human H2A/H2B and *X. laevis* H3/H4 recombinant histones (**Figure 2A**), and with short DNA fragments containing either one or two 601 nucleosome positioning sequences (**Figure 2B**) (Lowary and Widom, 1998; Kato et al., 2017). These sequences have a high affinity for histone octamers and direct nucleosome assembly with high efficiency (Lowary and Widom, 1998; Thåström et al., 1999). Mononucleosome assembly was performed as previously described (Dyer et al., 2003). For the assembly of dinucleosomes, we used the sequence designed by Kato et al. (2017), which contains an internucleosomal

48-bp spacer that accommodates the efficient assembly of two nucleosomes on the donor DNA (Engelholm et al., 2009; Machida et al., 2018). Consistent with previous reports, we observed the formation of a predominant slower-migrating species that corresponds to dinucleosomes with an octamer:DNA molar ratio of 2.4. Nucleosome assembly was monitored by native gel electrophoresis to reveal both the DNA and protein content (Figures 2C,D). Smaller migrating species observed at lower octamer:DNA ratios with the 601-601 DNA were previously found to be due to the assembly of one nucleosome on either one of the positioning elements (Engelholm et al., 2009). The presence of two nucleosomes on the 601-601 DNA in the slower-migrating species observed was further validated by cleaving the linker DNA with XhoI (Figures 2B,E; Machida et al., 2018).

The acetyltransferase activity of purified native NuA4/TIP60 was assessed by *in vitro* HAT assays, in which mono- or dinucleosomes were mixed with the purified complex and H³-labeled acetyl-CoA (Figure 2F). The reactions were analyzed by both SDS-PAGE and liquid HAT assays. The autoradiogram revealed that both histones H2A and H4 are more efficiently acetylated when recombinant dinucleosomes were used as substrate (Figure 2G). Coomassie Brilliant Blue staining confirmed that similar amounts of histones were used in each reaction. Liquid HAT assays of the same samples corroborated these results (Figure 2H), which are consistent with a recent observation for the chromatin-modifier polycomb repressive complex 2, which is more active on dinucleosomes than on H3 tails or single nucleosomes (Poepsel et al., 2018). The recent development of a method to assemble rNCPs with various DNA fragments and histone variants further highlights the potential of this approach to characterize the activity of native complexes (Changolkar and Pehrson, 2003; Muthurajan et al., 2016; Sekulic and Black, 2016). For example, rNCPs can be reconstituted with different DNA fragments to study the impacts of different linker DNAs between nucleosomes and in the 5' and 3' ends of the template DNA (Engelholm et al., 2009). In addition, specific mutations and PTMs on residues can be engineered on canonical or variant histones prior to reconstitution, allowing very precise questions to be tested experimentally.

Establishment of an Anchor-Away System for Nuclear Depletion in Mammalian Cells

The mammalian anchor-away system allows the rapid and robust depletion of a PYLcs-tagged protein from the nucleus (Figure 3A). As long-term depletion of NuA4/TIP60 complex subunits leads to cell toxicity (Gorini et al., 2007; Fazzio et al., 2008; Hu et al., 2009; Steunou et al., 2014; Numata et al., 2020), this system provides a great alternative to study its roles in biological processes, including DNA repair (Jacquet et al., 2016). In this example, modifier depletion immediately prior to DNA break induction would be useful to separate an acetyltransferase's transcriptional functions from its role acetylating histones surrounding the break. To establish this system in mammalian cells, we fused an ABI1cs-V5 tag to the C-terminal domain of the ribosomal protein RPL13 (homologous to the anchor used

in the yeast system) using a CRISPR/Cas9-driven approach. A sgRNA targeting the 3'-end of exon 6 was designed to target the nuclease near the stop codon (Figure 3B). A donor molecule containing the TAP tag and homology arms for RPL13 was used to integrate the tag and delete the endogenous stop codon. Following transfection into U2OS cells, tag incorporation was detected in the pool and in two isolated clones by PCR (Figure 3C). Accurate gene modification was also confirmed by western blot analysis (Figure 3D) and by sequencing, as previously described (Dalvai et al., 2015). The PYLcs-eGFP fusion protein was integrated at the AAVS1 locus in U2OS-RPL13-ABI1cs-V5 clone 25, and its expression was confirmed by western blot analysis following the addition of ABA (or MeOH as a negative control; Figure 3E). Using immunofluorescence, we observed that adding ABA triggered the nuclear exclusion of eGFP 6 h after treatment (Figure 3F). Interestingly, the PYLcs-eGFP fusion was less abundant by western blot before ABA addition, suggesting that the fusion is expressed at low levels and stabilized upon complexing with RPL13-ABI1cs-V5 (Figure 3D). Although these results are promising, further experiments will be required to validate that the fusion between PYLc and a ribosomal protein can efficiently deplete subunits of chromatin-modifying enzymes from the nucleus, to establish the kinetics of this process and to determine if the fusion impact the endogenous function of RPL13. The fact that we were unable to isolate a clone with homozygous tagged RPL13 (Figure 3C) raises a concern about the impact of the fusion on the cellular function of the protein. The main advantage of this system compared to other systems based on proteasome degradation (AID/Tir1 and dTAG approaches (Nabet et al., 2018; Nishimura et al., 2020; Yesbolatova et al., 2020)) is that the recovery of essential nuclear protein should be faster in the absence of the drug as no degradation is involved. It is thus expected that it will have less impact on cellular fitness during experiments. Nonetheless, our results provide proof of concept of the mammalian anchor-away system's great potential to enable temporal examination of the specific functions of essential chromatin modifiers.

CONCLUSION

The detailed step-by-step protocols provided here will be helpful to researchers interested in rapidly characterizing native chromatin modifying complexes. Of course, the enzymatic assays used will differ depending on the PTM deposited, and NCP composition should be modified depending on the presumed target or hypothesis being tested (e.g., H3.3 variants vs H3.1, H2A.Z/H2A.X vs H2A). Streamlining these approaches within research teams will greatly expand the experimental angles available to address scientific questions about chromatin-based molecular mechanisms.

MATERIALS AND EQUIPMENT

Reagents

E23-L48-E23 plasmid (Kato et al., 2017)

- 3 × FLAG peptide (Sigma, Cat. No. F4799)
 AAVS1 Puro PGK1 3 × FLAG Twin Strep Plasmid (Addgene, Cat. No. 68375)
 Absciscic acid (Sigma, Cat. No. 5.30339)
 Acetic acid (Anachemia, Cat. No. 000598-460)
 Acetonitrile (Sigma, Cat. No. 271004)
 Agar A (BioBasic, Cat. No. FB0010)
 Ampicillin (Bioshop, cat. No. AMP201.100)
 Anti-FLAG M2 affinity beads (Sigma, Cat. No. F1804)
 Anti-GFP (Roche, Cat. No. 11814460001)
 Anti-V5 antibody (Abcam, Cat. No. SV5-Pk1)
 Aprotinin (Sigma, Cat. No. A3886)
 BbsI (NEB, Cat. No. R0539)
 Benzamidinium hydrochlorate hydrate (Sigma, Cat. No. B6506)
 β-mercaptoethanol (Sigma, Cat. No. M3148)
 β-glycerophosphate (Sigma, Cat. No. G9422)
 Bolt 4–12% Bis-Tris gels (Thermo Fisher Scientific, Cat. No. NW04120BOX)
 Bolt 12% Bis-Tris gels (Thermo Fisher Scientific, Cat. No. NW00122BOX)
 Boric acid (Sigma, Cat. No. B6768)
 Bromophenol blue (Bioshop, Cat. No. BRO222)
 Citric acid (Sigma, Cat. No. 251275)
 Coomassie Brilliant Blue (Bioshop, Cat. No. CBB555)
 Coverslips (FisherBrand, Cat. No. 12541B)
 CutSmart Buffer (NEB, Cat. No. B7204S)
 D-Biotin (ThermoFisher, Cat. No. B20656)
 DH10β (ThermoFisher, Cat. No. EC0113)
 Dimethyl sulfoxide (Sigma, Cat. No. D8418)
 DNA retardation gel (Life Technology, Cat. No. EC6365BOX)
 DTT (Bioshop, Cat. No. DTT002)
 EcoRV (NEB, Cat. No. R3195L)
 EDTA (Sigma, Cat. No. E5134)
 EN3HANCE solution (PerkinElmer, Cat. No. 6NE9701)
 Ethanol, 100% (Commercial Alcohols, Cat. No. P016EAAAN)
 Fetal bovine serum (ThermoFisher, Cat. No. 12483020)
 Filter paper – P81 (Sigma, Cat. No. Z742570)
 Fluoromount G (eBioscience, Cat. No. 00-4958-02)
 Formaldehyde (Sigma, Cat. No. 252549)
 Gibson Assembly kit (NEB, Cat. No. E5510S)
 Glutaraldehyde (Sigma, Cat. No. G6403)
 Glycerol (Bioshop, Cat. No. GLY001)
 Guanidinium-HCl (BioBasic, Cat. No. GB0242)
 H³-labeled acetyl-CoA (PerkinElmer Life Sciences, cat. No. NET290050UC)
 HEPES (Bioshop, Cat. No. HEP001.1)
 HEPES, 1 M, for cell culture (Life Technologies, Cat. No. 15630080)
 HiTrap SP HP Sepharose columns (GE Healthcare, Cat. No. 45-100-294)
 InstantBlue™ Coomassie Stain (Bioshop, Cat. No. CBB555.25)
 Isopropyl β-D-thiogalactoside (Sigma, Cat. No. I6758-10G)
 K562 cells (ATCC, Cat. No. CCL-243)
 KCl, reagent grade (Bioshop, Cat. No. POC308)
 Leupeptin (Sigma, Cat. No. 78435)
 EcoLite (+) Liquid Scintillation Cocktail (MP Biomedicals, Cat. No. 0188247504)
 Lysozyme from chicken egg (Sigma, Cat. No. L6876)
 MB AG 501-X8 (D) resin (BioRad, Cat. No. 1436425)
 McCoy's Modified Medium (ThermoFisher, Cat. No. 16600108)
 Methanol (Fisher Chemical, Cat. No. A412)
 Methanol Free 16% Formaldehyde (ThermoFisher, Cat. No. 28908)
 MgCl₂ ACS reagent grade (Bioshop, Cat. No. MAG510)
 Microscope slides (FisherBrand, Cat. No. 22-034-486)
 MOPS (Sigma, Cat. No. M3183)
 Na₂CO₃ (Sigma, Cat. No. 451614)
 NaCl (Bioshop, Cat. No. SOD001.5)
 NaF (Sigma, Cat. No. 201154)
 NaHCO₃ (Sigma, Cat. No. S6014)
 Orthovanadate (Sigma, Cat. No. S6508)
 P81 phosphocellulose filter paper (Sigma, Cat. No. Z742570)
 pGEM-3z/601 (Addgene, Cat. No. 26656)
 PEG6000 (Cederlane, Cat. No. 8.07491.1000)
 Penicillin-Streptomycin (ThermoFisher, Cat. No. 15140122)
 Pepstatin A (Sigma, Cat. No. P5318)
 Pet15b vector (EDM Millipore, Cat. No. 69661)
 Pet28a Synthetic Human H2A.1 (Addgene, Cat. No. 42634)
 Pet28a Human H2B.1 (Addgene, Cat. No. 42630)
 Pet3d Xenopus H3_{C110A} and Pet3a Xenopus H4 plasmids (a kind gift from Cheryl Arrowsmith)
 PMSF (Thermo Fisher Scientific, Cat. No. 36978)
 pSpCas9(BB)-2A-Puro (pX459) V2.0 vector (Addgene, Cat. No. 62988)
 Puromycin (Sigma, Cat. No. P8833)
 Qiagen GigaPrep kit (Qiagen, Cat. No. 12191)
 QuickExtract DNA extraction solution (Epicentre, Cat. No. QE09050)
 RPMI medium (Life Technologies, Cat. No. 21870-092)
 Sepharose CL-6B resin (Sigma, Cat. No. CL-6B-200)
 SE XL Kit, nucleofection (Lonza, Cat. No. V4LC-2020)
 Silver nitrate (Sigma, Cat. No. S8157)
 Sodium butyrate (Sigma, Cat. No. B5887)
 Sodium carbonate (Sigma, Cat. No. 451614)
 Strep-Tactin Superflow Sepharose affinity matrix (IBA, Cat. No. 2-1206-010)
 Sucrose, biotechnology grade (Bioshop, Cat. No. SUC700)
 SV-ABAactDA plasmid (Addgene, Cat. No. 38247)
 SYBR green (ThermoFisher, Cat. No. S7563)
 Sypro™ Ruby protein gel stain (Bio-Rad, Cat. No. 1703125)
 Tris base (Bioshop, Cat. No. TRS001.5)
 Triton X-100 (BioBasic, Cat. No. TB0198)
 Tryptone powder (BioBasic, Cat. No. G211)
 Tween-20 (Bioshop, Cat. No. TWN508)
 U2OS cells (ATCC, HTB-96)
 Urea, reagent grade (Bioshop, Cat. No. URE002)

Yeast extract (BioBasic, Cat. No. G0961)
Zero Blunt pTOPO cloning kit (Life Technology, Cat. No. 450245)

Equipment

Amicon Ultra-0.5 centrifugal filter units (Millipore, Cat. No. UFC503024)
Bioruptor (Diagenode)
Dounce homogenizer with a type B pestle (Thomas Scientific, Cat. No. 1229H80)
Dialysis tubing with a cut off of 3.5 kDa (Thermo Scientific, Cat. No. REF68035)
Electrophoresis and blotting apparatus (Biorad, Cat. Nos. 1658001FC and 1703930)
Labconco FreeZone 1 Liter Benchtop Freeze Dry System (Cat. No. 7740020)
LS 6500 Multi-purpose Scintillation Counter (Beckman Coulter)
Nikon Ti Eclipse fluorescence microscope with a Hamamatsu Orca ER camera
NGC Scout 10 Plus and fractionator (BioRad)
Micro Bio-Spin columns (BioRad, Cat. No. 7326204)
Nutator
Optima LE-80K Ultracentrifuge (Beckman-Coulter)
Peristaltic pump (Buchler Instruments Polystaltic Pump)
Poly-Prep chromatography columns (Bio-Rad, Cat. No. 7311550)
S200 HiLoad 16/600 Superdex FPLC columns (GE Healthcare, Cat. No. 28-9893-35)
Slide-A-Lyzer 0.5 mL with a cut off of 10 kDa (Life Technology, Cat. No. 66383)
Scintillation vials (Fisher, Cat. No. 03-337-20)
Sorvall LYNX 4000 Superspeed Centrifuge and tubes (Thermo Scientific)
Spinner Flasks, 3L (Fisher, Cat. No. 10203E)
UV spectrophotometer (Nanodrop One^c, Thermo Scientific)
Water bath

Solutions

Buffer A	7 M urea, 20 mM Tris pH 8, add 5 mM β -mercaptoethanol fresh (just before use)
Buffer B	7 M urea, 20 mM Tris pH 8, 1 M NaCl, add 5 mM β -mercaptoethanol fresh
Carbonate developing solution	283 mM sodium carbonate, 0.0185% formaldehyde
CutSmart buffer	50 mM potassium acetate, 20 mM Tris-acetate pH 7.9, 10 mM $MgCl_2$ -acetate, 100 μ g/mL BSA
HAT buffer	50 mM Tris-HCl pH 8, 10 mM sodium butyrate, 5% glycerol, 0.1 mM EDTA, add 1 mM DTT fresh
High salt buffer	10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 1.2 M KCl, 25% glycerol, 0.2 mM EDTA, add 0.2 mM PMSF and 0.5 mM DTT fresh

High salt reconstitution buffer	2 M KCl, 10 mM Tris pH 7.5, 1 mM EDTA, add 1 mM DTT fresh
Histone wash buffer	50 mM Tris pH 7.5, 100 mM NaCl, 1 mM EDTA, add 1 mM benzamidine and 5 mM β -mercaptoethanol fresh
Hypotonic buffer	10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, add 0.2 mM PMSF and 0.5 mM DTT fresh
Low salt buffer	10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 20 mM KCl, 25% glycerol, 0.2 mM EDTA, add 0.2 mM PMSF and 0.5 M DTT fresh
Low salt reconstitution buffer	0.2 M KCl, 10 mM Tris pH 7.5, 1 mM EDTA, add 1 mM DTT fresh
MOPS SDS running buffer	50 mM MOPS, 50 mM Tris, 0.1% SDS, 1 mM EDTA
NEB3 buffer	100 mM NaCl, 50 mM Tris-HCl pH 7.9, 10 mM $MgCl_2$, 100 μ g/mL BSA
Nucleosome loading buffer (8 \times)	40% sucrose, 0.1% bromophenol blue
Refolding buffer	10 mM Tris pH 7.5, 2 M NaCl, 1 mM EDTA, add 5 mM β -mercaptoethanol fresh
Running buffer	25 mM Tris, 50 mM glycine, 0.1% SDS
TAE buffer	40 mM Tris, 1 mM EDTA, 20 mM glacial acetic acid
TAP buffer	20 mM HEPES-KOH pH 7.9, 300 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, 10% glycerol, add 10 mM sodium butyrate, 10 mM β -glycerophosphate, 1 mM PMSF, 5 mM NaF, 100 μ M orthovanadate, 2 μ g/mL leupeptin, 2 μ g/mL pepstatin, and 5 μ g/mL aprotinin fresh
TAP wash buffer #1	20 mM HEPES-KOH pH 7.9, 300 mM KCl, 0.1% Tween-20, 10% glycerol, add 1 mM DTT, 10 mM sodium butyrate, 10 mM β -glycerophosphate, 1 mM PMSF, 5 mM NaF, 100 μ M orthovanadate, 2 μ g/mL leupeptin, 2 μ g/mL pepstatin, and 5 μ g/mL aprotinin fresh
TAP wash buffer #2	20 mM HEPES-KOH pH 7.9, 150 mM KCl, 0.1% Tween-20, 10% glycerol, add 1 mM DTT, 10 mM sodium butyrate, 10 mM β -glycerophosphate, 1 mM PMSF, 5 mM NaF, 100 μ M orthovanadate, 2 μ g/mL leupeptin, 2 μ g/mL pepstatin, and 5 μ g/mL aprotinin fresh
TBE (0.2 \times)	18 mM Tris, 18 mM boric acid, 0.4 mM EDTA, pH 8.0
Unfolding buffer	20 mM Tris pH 7.5, 7 M guanidine-HCL, add 10 mM DTT fresh
Urea dialysis buffer	7 M urea, 10 mM Tris pH 8, 1 mM EDTA, 100 mM NaCl, add 5 mM β -mercaptoethanol fresh.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

AF-T and JC conceived the original idea and supervised the project. MG, CL, XC, and FD-G performed initial protocol optimization. MG and CL performed the experimental work presented in this manuscript. AF-T and JC wrote the manuscript

with contributions from MG and CL. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded in part by the Natural Sciences and Engineering Research Council of Canada (NSERC

Grant Number RGPIN-2016-05844, awarded to AF-T), and the Canadian Institutes of Health Research (CIHR Operating Grant Number FDN-143314, awarded to JC). JC is a tier 1 Canada Research Chair in Chromatin Biology and Molecular Epigenetics. AF-T is a tier 2 Canada Research Chair in Molecular Virology and Genomic Instability and is supported by the Foundation J.-Louis Lévesque.

REFERENCES

- Abmayr, S. M., Yao, T., Parmely, T., and Workman, J. L. (2006). Preparation of nuclear and cytoplasmic extracts from mammalian cells. *Curr. Protocols Mol. Biol.* Chapter 12:Unit12.1. doi: 10.1002/0471142727.mb1201s75
- Allis, C. D., and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* 17:487. doi: 10.1038/nrg.2016.59
- Bhardwaj, S. K., Hailu, S. G., Olufemi, L., Brahma, S., Kundu, S., Hota, S. K., et al. (2020). Dinucleosome specificity and allosteric switch of the ISW1a ATP-dependent chromatin remodeler in transcription regulation. *Nat. Commun.* 11:5913. doi: 10.1038/s41467-020-19700-1
- Brien, G. L., Valerio, D. G., and Armstrong, S. A. (2016). Exploiting the epigenome to control cancer-promoting gene-expression programs. *Cancer Cell* 29, 464–476. doi: 10.1016/j.ccell.2016.03.007
- Changolkar, L. N., and Pehrson, J. R. (2003). Histone MacroH2A purification and nucleosome reconstitution. *Methods Enzymol.* 375, 228–238. doi: 10.1016/S0076-6879(03)75015-0
- Côté, J., Utley, R. T., and Workman, J. L. (1995). Basic analysis of transcription factor binding to nucleosomes. *Methods Mol. Genet.* 6, 108–128. doi: 10.1016/S1067-2389(06)80009-9
- Dalvai, M., Loehr, J., Jacquet, K., Huard, C. C., Roques, C., Herst, P., et al. (2015). A scalable genome-editing-based approach for mapping multiprotein complexes in human cells. *Cell Rep.* 13, 621–633. doi: 10.1016/j.celrep.2015.09.009
- DeKelver, R. C., Choi, V. M., Moehle, E. A., Paschon, D. E., Hockemeyer, D., Meijsing, S. H., et al. (2010). Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Res.* 20, 1133–1142. doi: 10.1101/gr.106773.110
- Doyon, Y., and Côté, J. (2016). Preparation and analysis of native chromatin-modifying complexes. *Methods Enzymol.* 573, 303–318. doi: 10.1016/bs.mie.2016.01.017
- Doyon, Y., Cayrou, C., Ullah, M., Landry, A. J., Côté, V., Selleck, W., et al. (2006). ING tumor suppressor proteins are critical regulators of chromatin acetylation required for genome expression and perpetuation. *Mol. Cell* 21, 51–64. doi: 10.1016/j.molcel.2005.12.007
- Doyon, Y., Selleck, W., Lane, W. S., Tan, S., and Côté, J. (2004). Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans. *Mol. Cell. Biol.* 24, 1884–1896. doi: 10.1128/mcb.24.5.1884-1896.2004
- Dyer, P. N., Edayathumangalam, R. S., White, C. L., Bao, Y., Chakravarthy, S., Muthurajan, U. M., et al. (2003). Reconstitution of nucleosome core particles from recombinant histones and DNA. *Methods Enzymol.* 375, 23–44. doi: 10.1016/S0076-6879(03)75002-2
- Einhauer, A., and Jungbauer, A. (2001). The FLAGTM peptide, a versatile fusion tag for the purification of recombinant proteins. *J. Biochem. Biophys. Methods* 49, 455–465. doi: 10.1016/S0165-022X(01)00213-5
- Engelholm, M., de Jager, M., Flaus, A., Brenk, R., van Noort, J., and Owen-Hughes, T. (2009). Nucleosomes can invade DNA territories occupied by their neighbors. *Nat. Struct. Mol. Biol.* 16, 151–158. doi: 10.1038/nsmb.1551
- Farnung, L., Ochmann, M., and Cramer, P. (2020). Nucleosome-CHD4 chromatin remodeler structure maps human disease mutations. *eLife* 9:e56178. doi: 10.7554/eLife.56178
- Fazio, T. G., Huff, J. T., and Panning, B. (2008). An RNAi screen of chromatin proteins identifies Tip60-p400 as a regulator of embryonic stem cell identity. *Cell* 134, 162–174. doi: 10.1016/j.cell.2008.05.031
- Feng, Y., Vlassis, A., Roques, C., Lalonde, M., González-Aguilera, C., Lambert, J., et al. (2016). BRPF 3- HBO 1 regulates replication origin activation and histone H3K14 acetylation. *EMBO J.* 35, 176–192. doi: 10.15252/embj.201591293
- Gibson, T. J., Seiler, M., and Veitia, R. A. (2013). The transience of transient overexpression. *Nat. Methods* 10, 715–721. doi: 10.1038/nmeth.2534
- Gorrini, C., Squatrito, M., Luise, C., Syed, N., Perna, D., Wark, L., et al. (2007). Tip60 is a haplo-insufficient tumour suppressor required for an oncogene-induced DNA damage response. *Nature* 448, 1063–1067. doi: 10.1038/nature06055
- Haruki, H., Nishikawa, J., and Laemmli, U. K. (2008). The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. *Mol. Cell* 31, 925–932. doi: 10.1016/j.molcel.2008.07.020
- Havasi, A., Haeghele, J. A., Gall, J. M., Blackmon, S., Ichimura, T., Bonegio, R. G., et al. (2013). Histone acetyl transferase (HAT) HBO1 and JADE1 in epithelial cell regeneration. *Am. J. Pathol.* 182, 152–162. doi: 10.1016/j.ajpath.2012.09.017
- Hockemeyer, D., Soldner, F., Beard, C., Gao, Q., Mitalipova, M., Dekelver, R. C., et al. (2009). Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. *Nat. Biotechnol.* 27, 851–857. doi: 10.1038/nbt.1562
- Hopp, T. P., Prickett, K. S., Price, V. L., Libby, R. T., March, C. J., Cerretti, D. P., et al. (1988). A short polypeptide marker sequence useful for recombinant protein identification and purification. *Bio/Technology* 6, 1204–1210. doi: 10.1038/nbt1088-1204
- Hu, Y., Fisher, J. B., Koprowski, S., McAllister, D., Kim, M. S., and Lough, J. (2009). Homozygous disruption of the Tip60 gene causes early embryonic lethality. *Dev. Dyn.* 238, 2912–2921. doi: 10.1002/dvdy.22110
- Huang, J., and Tan, S. (2013). Piccolo NuA4-catalyzed acetylation of nucleosomal histones: critical roles of an Esa1 tudor/chromo barrel loop and an epl1 enhancer of polycomb A (EPcA) basic region. *Mol. Cell. Biol.* 33, 159–169. doi: 10.1128/mcb.01131-12
- Ikura, T., Ogryzko, V. V., Grigoriev, M., Groisman, R., Wang, J., Horikoshi, M., et al. (2000). Involvement of the TIP60 histone acetylase complex in DNA repair and apoptosis. *Cell* 102, 463–473. doi: 10.1016/S0092-8674(00)00051-9
- Jacquet, K., Fradet-Turcotte, A., Avvakumov, N., Lambert, J. P., Roques, C., Pandita, R. K., et al. (2016). The TIP60 complex regulates bivalent chromatin recognition by 53BP1 through Direct H4K20me binding and H2AK15 acetylation. *Mol. Cell* 62, 409–421. doi: 10.1016/j.molcel.2016.03.031
- Kato, D., Osakabe, A., Arimura, Y., Mizukami, Y., Horikoshi, N., Saikusa, K., et al. (2017). Crystal structure of the overlapping dinucleosome composed of hexasome and octasome. *Science* 356, 205–208. doi: 10.1126/science.aak9867
- Lalonde, M. E., Cheng, X., and Côté, J. (2014). Histone target selection within chromatin: an exemplary case of teamwork. *Genes Dev.* 28, 1029–1041. doi: 10.1101/gad.236331.113
- Liang, F. S., Ho, W. Q., and Crabtree, G. R. (2011). Engineering the ABA Plant stress pathway for regulation of induced proximity. *Sci. Signal.* 4:rs2. doi: 10.1126/scisignal.2001449
- Lombardo, A., Cesana, D., Genovese, P., di Stefano, B., Provati, E., Colombo, D. F., et al. (2011). Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat. Methods* 8, 861–869. doi: 10.1038/nmeth.1674
- Lowary, P. T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* 276, 19–42. doi: 10.1006/jmbi.1997.1494
- Ma, Y., Kanakousaki, K., and Buttitta, L. (2015). How the cell cycle impacts chromatin architecture and influences cell fate. *Front. Genet.* 6:19. doi: 10.3389/fgene.2015.00019

- Machida, S., Takizawa, Y., Ishimaru, M., Sugita, Y., Sekine, S., Nakayama, J. I., et al. (2018). Structural basis of heterochromatin formation by human HP1. *Mol. Cell* 69, 385.e–397.e. doi: 10.1016/j.molcel.2017.12.011
- McGinty, R. K., Kim, J., Chatterjee, C., Roeder, R. G., and Muir, T. W. (2008). Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* 453, 812–816. doi: 10.1038/nature06906
- Miotto, B., and Struhl, K. (2010). HBO1 histone acetylase activity is essential for DNA replication licensing and inhibited by geminin. *Mol. Cell* 37, 57–66. doi: 10.1016/j.molcel.2009.12.012
- Muthurajan, U., Mattioli, F., Bergeron, S., Zhou, K., Gu, Y., Chakravarthy, S., et al. (2016). In vitro chromatin assembly: strategies and quality control. *Methods Enzymol.* 573, 3–41. doi: 10.1016/bs.mie.2016.01.002
- Nabet, B., Roberts, J. M., Buckley, D. L., Paulk, J., Dastjerdi, S., Yang, A., et al. (2018). The dTAG system for immediate and target-specific protein degradation. *Nat. Chem. Biol.* 14, 431–441. doi: 10.1038/s41589-018-0021-8
- Nishimura, K., Yamada, R., Hagihara, S., Iwasaki, R., Uchida, N., Kamura, T., et al. (2020). A super-sensitive auxin-inducible degron system with an engineered auxin-TIR1 pair. *Nucleic Acids Res.* 48:e108. doi: 10.1093/nar/gkaa748
- Numata, A., Kwok, H. S., Zhou, Q. L., Li, J., Tirado-Magallanes, R., Angarica, V. E., et al. (2020). Lysine acetyltransferase Tip60 is required for hematopoietic stem cell maintenance. *Blood* 136, 1735–1747. doi: 10.1182/blood.2019001279
- Patel, A. B., Moore, C. M., Greber, B. J., Luo, J., Zukin, S., Ranish, J., et al. (2019). Architecture of the chromatin remodeler RSC and insights into its nucleosome engagement. *eLife* 8:e54449. doi: 10.7554/eLife.54449
- Poepsel, S., Kasinath, V., and Nogales, E. (2018). Cryo-EM structures of PRC2 simultaneously engaged with two functionally distinct nucleosomes. *Nat. Struct. Mol. Biol.* 25, 154–162. doi: 10.1038/s41594-018-0023-y
- Puig, O., Caspary, F., Rigaut, G., Rutz, B., Bouveret, E., Bragado-Nilsson, E., et al. (2001). The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* 24, 218–229. doi: 10.1006/meth.2001.1183
- Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 17, 1030–1032. doi: 10.1038/13732
- Sadelain, M., Papapetrou, E. P., and Bushman, F. D. (2012). Safe harbours for the integration of new DNA in the human genome. *Nat. Rev. Cancer* 12, 51–58. doi: 10.1038/nrc3179
- Schnitzler, G. R. (2000). Isolation of histones and nucleosome cores from mammalian cells. *Curr. Protocols Mol. Biol.* Chapter 21:Unit21.5. doi: 10.1002/0471142727.mb2105s50
- Sekulic, N., and Black, B. E. (2016). Preparation of recombinant centromeric nucleosomes and formation of complexes with nonhistone centromere proteins. *Methods Enzymol.* 573, 67–96. doi: 10.1016/bs.mie.2016.01.014
- Shen, H., and Laird, P. W. (2013). Interplay between the cancer genome and epigenome. *Cell* 153, 38–55. doi: 10.1016/j.cell.2013.03.008
- Simon, M. D. (2010). Installation of site-specific methylation into histones using methyl lysine analogs. *Curr. Protocols Mol. Biol.* Chapter 21, Unit 21.18.1–10. doi: 10.1002/0471142727.mb2118s90
- Steunou, A. L., Rossetto, D., and Côté, J. (2014). “Regulating chromatin by histone acetylation,” in *Fundamentals of Chromatin*, eds J. Workman and S. Abmayr (New York, NY: Springer), doi: 10.1007/978-1-4614-8624-4_4
- Thåström, A., Lowary, P. T., Widlund, H. R., Cao, H., Kubista, M., and Widom, J. (1999). Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J. Mol. Biol.* 288, 213–229. doi: 10.1006/jmbi.1999.2686
- Utley, R. T., Owen-Hughes, T. A., Juan, L. J., Côté, J., Adams, C. C., and Workman, J. L. (1996). In vitro analysis of transcription factor binding to nucleosomes and nucleosome disruption/displacement. *Methods Enzymol.* 274, 276–291. doi: 10.1016/S0076-6879(96)74024-7
- Wagner, F. R., Dienemann, C., Wang, H., Stützer, A., Tegunov, D., Urlaub, H., et al. (2020). Structure of SWI/SNF chromatin remodeller RSC bound to a nucleosome. *Nature* 579, 448–451. doi: 10.1038/s41586-020-2088-0
- Wang, H., Farnung, L., Dienemann, C., and Cramer, P. (2020). Structure of H3K36-methylated nucleosome–PWWP complex reveals multivalent cross-gyre binding. *Nat. Struct. Mol. Biol.* 27, 1–6. doi: 10.1038/s41594-019-0345-4
- Wojcik, F., Dann, G. P., Beh, L. Y., Debelouchina, G. T., Hofmann, R., and Muir, T. W. (2018). Functional crosstalk between histone H2B ubiquitylation and H2A modifications and variants. *Nat. Commun.* 9:1394. doi: 10.1038/s41467-018-03895-5
- Yesbolatova, A., Saito, Y., Kitamoto, N., Makino-Itou, H., Ajima, R., Nakano, R., et al. (2020). The auxin-inducible degron 2 technology provides sharp degradation control in yeast, mammalian cells, and mice. *Nat. Commun.* 11:5701. doi: 10.1038/s41467-020-19532-z
- Zhang, Y., and Kutateladze, T. G. (2019). Methylation of histone H3K79 by dot1l requires multiple contacts with the ubiquitinated nucleosome. *Mol. Cell* 74, 862–863. doi: 10.1016/j.molcel.2019.05.013

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Galloy, Lachance, Cheng, Distéfano-Gagné, Côté and Fradet-Turcotte. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: frontiersin.org/about/contact



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

@frontiersin



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership