Structural insights into the mechanisms controlling the modification and the readout of histone H3.1

Hossein Davarinejad

Thesis submitted to the
Faculty of Medicine
in partial fulfillment of the requirements
for the Doctorate in Philosophy degree in Biochemistry

Department of Biochemistry, Microbiology and Immunology
Faculty of Medicine
University of Ottawa

Abstract

The histone H3.1 is the canonical histone H3 inserted into new nucleosomes during replication. The unstructured N-terminal region (also referred to as the tail) of histone H3.1 is decorated by many post-translational modifications (PTMs). These PTMs serve as epigenetic signals coordinating essential nuclear processes such as transcription, replication and DNA damage repair mechanisms. These marks can be reversibly deposited by specialized "writer" domains in chromatin remodelling enzymes. Similarly, histone PTMs are recognized by "reader" domains which aid in coordinating appropriate responses and recruitment of relevant factors. The contribution of PTMs deposited on histone H3.1 to nuclear transactions has been extensively described in chromatin structure, regulation of gene expression, cell cycle progression, apoptosis, and DNA damage repair. However, with several DNA damage repair pathways occurring during different cell cycle phases, the contribution of histone H3 variants and PTMs is not fully understood. In this work, we characterized the tetratricopeptide repeat (TPR) domain of TONSOKU (TSK), a DNA damage response protein, as a variant-specific reader for histone H3.1 and established a role for H3.1 in replication-coupled DNA repair via the homologous recombination pathway. Using X-ray crystallography, I solved the structure of TPR^{TSK} and characterized a large network of molecular interactions within many binding pockets on TPR^{TSK}, which stabilize the histone H3.1 tail, including the H3.1 alanine 31 variant-selectivity pocket. Using biochemical assays, I provide an extensive PTM reading profile for TPR^{TSK} on histone H3.1 showing TPR^{TSK} is permissive to H3.1 modifications related to active genes and relaxed chromatin but sensitive to those associated with chromatin condensation, including lysine 27 mono-methylation deposited by ATXR5/6 H3.1-specific methyltransferases. Finally, I characterized novel regulatory mechanisms that inhibit ATXR5/6 activity behind the replication fork. This work

provides structural and biochemical evidence for understating mechanisms coordinating the DNA damage response during replication that safeguard the genome.

Acknowledgements

I am forever grateful and indebted to Dr. Jean-Francois Couture, my PhD thesis supervisor, who gave me everything I needed and more for my PhD studies. His support and guidance made it possible for me to remain in my program during hard times and finish my PhD very proudly. Under his supervision, I secured many scholarships, won many awards, published in a journal I couldn't imagine achieving, learnt many techniques, learnt how to better think, present, and write, He is truly inspiring and the success he helped me build will pave my future, Thank you, Jeff!

I am thankful to my colleagues and friends Dr. Monika Joshi and Dr. Sabina Sarvan whose expertise and time were invaluable to my learning experience. I am very thankful to Nidhi, Yidai, Ashley, Sabrina, Alexis, Janelle, Siyona, and other members of the Couture lab for the pleasure of working with them and all the memories.

I thank Drs. Jyh-Huan (Eric) Lee, John Baenziger, David Picketts, and Marjorie Brand, who were members of my Thesis Advisory and Comprehensive Exam Committees, for their input and guidance in my projects.

A special thanks to Dr. Yannick Jacob and his lab at Yale University for the amazing collaboration.

I would like to thank members of Alarcon, Downey, Baetz, Cote, Lee, Figeys, Giguère, Baenziger, and other labs at the Roger Guindon building for the help and fun times.

I thank the Government of Ontario and University of Ottawa for the scholarships granted to me.

A special thanks to the administration, staff, and members of the uOttawa Faculty of Medicine, Department of BMI, and BMIGSA for creating a great learning environment.

Thank you, Agnes, for being the reason why I ended up in Ottawa. A blessing in disguise indeed.

I would like to thank my father Masoud and uncle Saiid who helped me financially during my studies.

A big thank you to my dear friends Chela, Natasha, Véronique, Sabina, Ioana, Ron, Benny, Shiva, Marija, David, Katherine, Erin, Lisa, Andrée-Anne, Brian, and Amirhossein who kept me in check mentally during these challenging years. And thank you Mooshmoosh and Kuchulu for your companionship.

Table of Contents

Abstract ii

Acknowledgements iv

Table of Contents v

List of Figures viii

List of Tables ix

List of Abbreviations x

Chapter 1 | Introduction 1

1.1 The chromatin landscape 1
1.2 Chromatin structure is important for gene regulation 1
1.3 Histones can be post-translationally modified 2
1.4 Histones and their variants are crucial epigenetic signalling platforms 3
1.5 Epigenetic signalling at the cross-roads of many important biological pathways 4
1.6 DNA damage can occur spontaneously and at a high rate 5
1.7 Unrepaired DNA damage pose high-risk problems 6
1.8 Context-specific DNA damage repair mechanisms 6
1.9 NHEJ and HR have common sensors but diverge in mediators and effectors 7
1.10 Contribution of NHEJ to DSB 7
1.11 HR is available during the S-phase 8
1.12 Histone readers recognize PTMs related to DNA damage repair 9
1.12.1 Alpha-helical histone readers 10
1.12.2 14-3-3 proteins read phosphorylated S10 and S28 on histone H3. 10
1.12.3 Ankyrin repeat domains read H3K9me1/2 and H4K20me0 16
1.12.4 Tetratricopeptide repeats (TPRs) bind long unstructured peptides 21
1.12.5 Bromodomains are readers for acetyl-lysine residues on histones H2A, H2B, H3, and H4 22
1.13 Rationale and Hypothesis 25

Chapter 2 | The histone H3.1 variant regulates TONSOKU-mediated DNA repair during replication 26

2.1 Authors and Affiliations 26
2.2 My Contributions 26
2.3 Abstract 27
2.4 Introduction 28
2.5 Results and Discussion 28
2.6 Supplementary Materials 42
2.7 Acknowledgments 42
2.7.1 Author Contributions 42
2.7.2 Funding 43
2.7.3 Competing interests 43
2.7.4 Data and materials availability 43

Chapter 3 | Structural and Biochemical Analysis of histone H3.1 recognition mechanisms by TONSOKU 44

3.1 Authors and Affiliations 44
3.2 My Contributions 44
3.3 Manuscript Submission 45
3.4 Abstract 46
3.5 Introduction 47
3.6 Results and Discussion 50
3.6.1 A systemic approach to measure the binding footprint of TPR^{TSK} to histone H3.1 50
3.6.2 Organisation of TPR^{TSK} poly-TPR domain 50
3.6.3 Histone H3.1 residues 31-40 bind the N-terminal Lobe (N-lobe) of TPR^{TSK} 50
3.6.4 The Center-Lobe of TPR^{TSK} encircles residues 24-30 of histone H3.1 53
3.6.5 The K27 sensing pocket of TPR^{TSK} 55
3.6.6 The wide C-terminal Lobe of TPR^{TSK} binds two regions of histone H3.1 57
3.6.7 Sponastrime dysplasia mutations in TONSL affect H3.1-binding. 60
3.7 Materials and Methods 75
3.7.1 Site-directed mutagenesis and peptide synthesis. 75
3.7.2 Protein Expression and Purification 75
3.7.3 ITC assay 76
3.7.4 Peptide synthesis, purification, and characterization 76
3.7.5 Crystallography 77
3.7.6 EMSA 77
3.8 Authorship Contribution Statement 78
3.9 Author Contributions 78
3.10 Data Availability 78
3.11 Acknowledgements 78

Chapter 4 | ATXR5/6 Forms Alternative Protein Complexes with PCNA and the Nucleosome Core Particle 79

4.1 Authors and Affiliations 79
4.2 My Contributions 79
4.3 Abstract 80
4.4 Introduction 81
4.5 Results and Discussion 82
4.5.1 An evolutionary conserved motif in ATXR6 mediates its interaction with PCNA 82
4.5.2 Crystal structure of PCNA in complex with ATXR6 PIP motif 86
4.5.3 PCNA inhibits nucleosome methylation by ATXR6 92
4.6 Materials and Methods 95
4.6.1 ATXR6 expression and purification 95
4.6.2 PCNA expression and purification 95
4.6.3 ITC 96
4.6.4 Gel filtration and analytical ultracentrifugation 97
4.6.5 Methyltransferase assays 97
4.6.6 EMSAs 98
4.7 Accession numbers 98
4.8 Acknowledgments 98
4.9 Supplementary data 98
4.10 Acknowledgments 98

Chapter 5 | Discussion and Perspectives 99

5.1 Summary 99
5.2 Solenoid poly-TPRs have unique features 100
5.3 Tandem alpha-helical domains – a growing histone reader class 103
5.4 PTM reading profile of TSK complements its cell-cycle-specific role 105
5.5 A refined role for H3K36me3 in HR 106
5.6 The TPR^{TSK} C-lobe: An add-on compatible binding model 107
5.7 TPR and ARD anchor on different surfaces of the nucleosome 111
5.8 Limitations and Future Directions 115

References 117

Appendix – Supplementary Materials for Chapter 2 129

List of Figures

Figure 1.1 – Comparison of the structure of α-helical histone reader domains. 14
Figure 1.2 – Histone H3 binding pockets of 14-3-3 and BRD domain. 15
Figure 1.3 – ARD domains of GLP9 and TONSL bind H3 and H4 peptides with different mechanisms. 20
Figure 2.1 – The TPR domain of TSK specifically interacts with the N-terminal tail of the H3.1 variant. 35
Figure 2.2 – Crystal structure of plant TPR^{TSK} bound to the H3.1 tail. 37
Figure 2.3 – Mutations in TSK suppress heterochromatin amplification of atxr5/6 mutants. 39
Figure 2.4 – H3.1 is required to mediate genomic instability in atxr5/6 mutants. 40
Figure 3.1 – Histone H3.1 residues 31-40 bind the N-terminal lobe (N-Lobe) of TPR^{TSK} 63
Figure 3.2 – The Center-Lobe of TPR^{TSK} encircles residues 24-30 of histone H3.1. 65
Figure 3.3 – The wide C-terminal Lobe (C-lobe) of TPR^{TSK} binds two regions of histone H3.1 66
Figure 3.4 – SD mutations negatively impact histone H3.1 binding. 67
Figure 4.1 – ATXR6 directly associates with PCNA. 85
Figure 4.2 – Crystal structure of ATXR6 PIP motif in complex with PCNA 89
Figure 4.3 – Comparative analysis of PIP:PCNA complexes 91
Figure 4.4 – ATXR6 PIP motif binds with high affinity to AtPCNA. 91
Figure 4.5 – PCNA inhibits ATXR6 methylation on NCP but not H3.1 peptide. 94
Figure 5.1 – TPRs bind long peptides in various configurations. 102
Figure 5.2 – TPR^{TSK} binds H3.1 during the S-phase of the cell cycle. 105
Figure 5.3 – A model for simultaneous reading of H3 by TPR and BRD 110
Figure 5.4 – ARD and TPR bind the NCP on nearby surfaces. 113
Figure 5.5 – H3.1 residues 39-45 are entrapped by the nucleosome disc. 114

List of Tables

Table 3-1 – Thermodynamic parameters for titration of TPR^{TSK} with post-translationally modified H3.1 peptides. 68
Table 3-2 – Thermodynamic parameters for titration of mutants TPR^{TSK} with H3.1 unmodified peptides. 70
Table 3-3: Peptide synthesis chemical structures based on the main backbone peptide, and post-translational modifications at the different described amino acid positions. 71
Table 3-4 – Data collection and refinement statistics for TPR^{TSK} and H3.1 K27ac complex. 74
Table 4-1 – Data collection and refinement statistics for the ATXR6 PIP/PCNA complex. 90
Table 5-1 – Comparison of the characteristics of tube-forming solenoid tetratricopeptide domains. 101
Table 5-2 – Targets and PTM-reading profiles of α-helical histone reader domains. 104

List of Abbreviations

A

ARD	ankyrin repeat domain
At	*Arabidopsis thaliana*
ATXR5/6	*Arabidposis* Trithorax-related proteins 5/6

B

BARD1	BRCA1-associated RING domain 1
BER	base-excision repair
βME	β-mercaptoethanol
BRCA1	breast cancer gene 1
BRCT	BRACA1 C-terminal domain
BRD	bromodomain

C

Ch.	chromosome
Cu	*Citrus unshiu*
CAF-1	chromatin assembly factor-1

D

DDR	DNA damage response
DNA	deoxyribonucleotide
DSB	double-strand break
dsDNA	double-stranded DNA
ssDNA	single-stranded DNA

E

E[z]	enhancer of zeste
EMSA	Electromobility shift assay

G

Gm	*Glycine max*

H

HP1	heterochromatin-associated protein 1
HR	homologous recombination
Hs	*Homo sapiens*

I

ITC	isothermal calorimetry

J

JMJD2A/B	Jumonji domain 2A/B

K

KMD4A/B	lysine-specific demethylase 4A/B

L

LRR	leucine-rich repeat

M

MLL	Mixed-Lineage Leukemia
MMS22L	methyl methanesulfonate-sensitivity protein 22-like
Mn	*Manihot notabilis*
MRE11	meiotic recombination 11
MRN	MRE11-RAD50-NBS1

N

NCP	nucleosome core particle
NER	nucleotide excision repair
NHEJ	non-homologous end joining
NBS1	Nijmegen breakage syndrome 1

P

PCNA	Proliferating-Cell Nuclear Antigen
PHD	Plant homeodomain
PIP	PCNA-interacting peptide

R

RING	Really Interesting New Gene
ROS	reactive oxygen species
RPA	Replication Protein A

S

SAM	*S*-adenosyl methionine
SCC	sister-chromatid cohesin
SEC	size-exclusion chromatography
SET	Su[var]-E[z]-TRX
SSB	single-strand break
Su[var]	suppressor of variegation

T

TONSL	TONSOKU-like
TPR	Tetratricopeptide repeat
TSK	TONSOKU
TRX	Trithorax
TSS	transcription start site

X

Xl	*Xenopus laevis*

Chapter 1 | Introduction

1.1 The chromatin landscape

The human genome is a massive library of inheritable information and collectively amounts to ~2 m of DNA in length (~6 Gb) which is disbursed in forty-six long molecules. DNA includes regions that code for instructions known as genes, and non-coding sequences that have regulatory roles (*1*). To confine the genome within the nucleus with an average radius of ~15 μm, cells compact DNA by wrapping the molecule around scaffolds made from eight histone proteins known as the histone octamer. The octamer is composed of two copies of each core histone protein types H2A, H2B, H3, and H4, and together with 1.8 turns of DNA that is bound to them (per octamer) they form disc like structures known as the nucleosome core particle (NCP) (*2*). Nucleosomes are the most fundamental level of DNA compaction and the most relaxed form of chromatin, that is, DNA in association with all of its packaging proteins. Chromatin can assume many forms; the most widely recognized form of chromatin, which if stained, is visible under light microscopes, is the chromosome (*3*). These structures represent the highest order of DNA compaction or the most condensed form of chromatin with a length of ~10 μm for the largest chromosome (Ch.1).

1.2 Chromatin structure is important for gene regulation

Biological pathways, such as those that replicate DNA or respond to stress, help sustain life and maintain the health of cells, and ultimately depend on processes that occur in the nucleus. These pathways rely on communicating information about the environment. They also rely on coordinating the timing and place of their actions and responses, and many require access to genes. Chromatin condensation and decondensation is one of the mechanisms that facilitate access to DNA (*4*). Regions of chromatin that house physically accessible genes are decondensed to the nucleosome level and form the relaxed chromatin structure that is termed the 11 nm fibre after the

diameter of the nucleosome (*5, 6*). These regions are also known as euchromatin and are gene-rich (*7*). Conversely, heterochromatic regions are formed by tighter packaging of nucleosomes and are less accessible for transactions requiring access to genes. These regions can be made further distinct as facultative heterochromatin (fHC) and constitutive heterochromatin (cHC) (*8*). fHC involves contextually condensed chromatin and formation of the 30nm fiber with the assistance of the linker histone H1 (*8*) while cHC perpetuates and maintains its condensed structure in higher order packaging (≥30 nm) and is mostly found in telomeric and pericentromeric regions (*5, 6, 8*). Spatial access to genes, however, is not the only hurdle for processes that require engagement with DNA. Several processes further serve as gatekeepers in regulating what and when access is granted and much of these processes are regulated by epigenetic signalling.

1.3 Histones can be post-translationally modified

Other than forming the 11 nm fiber, nucleosomes serve another important function. The histone proteins offer many sites that play the role of flag holders for conveying vastly different messages such as signals for changes to chromatin structure, state of the cell cycle, or changes to gene expression. Specifically, these sites are post-translationally modifiable residues such as lysine, arginine, serine, and threonine that can covalently accept chemical moieties such as methyl (me), acetyl (ac), ubiquityl (ub), and phosphoryl (ph) groups (*9*). Addition of these functional groups, which can covey different information, is catalyzed by specialized protein domains that can be shared among a protein family. For example, SET domain proteins and PRMTs methylate lysine and arginine residues, respectively (*10, 11*). Acetylation of lysine residues is catalyzed by HATs (*12*), and kinases phosphorylate serine and threonine residues (*13*). These enzymes deposit functional groups (marks) on histones and are referred to as histone “writers”. Similarly, post-

translational modifications (PTMs) of histones can be removed by "eraser" such as demethylases (*14*), deacetylases (*15*), ubiquitin-specific proteases (*16*), and phosphatases (*17*).

1.4 Histones and their variants are crucial epigenetic signalling platforms

Histone variants, a nomenclature assigned because they have been identified after the canonical histone proteins, are incorporated in nucleosomes and differ from their canonical forms by amino acid sequence substitutions or deletion while still remaining vastly homologous. For instance, in humans, the histone H3.3 variant is different by four residues compared to the canonical histone H3.1 to which it is ~97% identical. Histone variants are typically linked to a specific biological context; for instance, histone H2A variants H2AX (*18*) and H2AZ (*19*) are found in regions of chromatin where stress response is initiated. While canonical histone H3.1 is incorporated in newly assembled nucleosomes during replication and is widely found in euchromatin, its variants H3.3 (*20*) and CENP-A (*21*) are utilized in heterochromatin and centromeric regions, respectively.

When embedded in a nucleosome, much of the surfaces of histones are covered by the clustering of their α-helically structured histone fold domains which hold the histone octamer together. Areas of the octamer that are enwrapped by DNA are also shielded from the nucleoplasm (*2*). However, several histone residues positioned at the top and bottom of the nucleosome disc are exposed and are known to be post-translationally modified by several writers. Furthermore, histones feature unstructured N- and C-terminal regions, frequently referred to as histone "tails", that extend out from the nucleosome, exposed to the nucleoplasm and known to be hotspots for PTMs. The identity of the histone with respect to its type and variants, the abundance of modifiable residues on histones, the diversity of the functional groups, and the sequence-specific manner by which histones acquire PTMs provide histones with a sophisticated array of employable signals. These

PTMs serve as signalling hubs that provide key signals peripheral to gene sequences, leaving DNA sequences unaltered. Collectively, these signals are also referred to as epigenetic signalling.

1.5 Epigenetic signalling at the cross-roads of many important biological pathways

Virtually all essential processes in cells that transact with DNA are regulated by histone PTMs directly, indirectly via adapter proteins, or are ultimately downstream of these signalling cascades. For example, the process of gene expression is tightly regulated by activating or repressive histone marks. In humans, polycomb repressive complexes 2 and 1 (PRC2/1) tri-methylate lysine 27 on the tail of histone H3 and mono-ubiquitinate lysine 119 on the exposed surface of histone H2A (H2AK119ub), respectively. Recruitment of PRC1/2 to genes marks as transcriptionally silent by facilitating inhibition of the transcriptional machinery (*22*) and promoting the formation of heterochromatin by a few mechanisms including recruitment of chaperones that insert histone variant H3.3 into nucleosomes (*23*). Conversely, methylation of histone H3 lysine 4 (H3K4me) has been widely linked to actively transcribed genes. Mono-methylated H3K4 (H3K4me1) deposited by MLL1/2 are markers of active enhancer regions which play an important role in the stimulation of transcription, and tri-methylated H3K4 (H3K4me3) is enriched at promoters and near transcription start sites (TSS)s. In plants, ATXR5/6 proteins monomethylate histone H3.1 on lysine 27 (H3.1K27me1), and this mark is associated with chromatin condensation in pericentromeric regions known as chromocenters (*24, 25*). Several histone PTMs such as H3K115ac and K122ac occur near the nucleosome dyad (central axis that separates the histone copies) and loosen the interaction of DNA with the histone octamer which is thought to be important for repositioning of the nucleosome (*26*). Histone PTM patterns also serve as markers of the cell-cycle. For instance, histone H3 is distinctly phosphorylated during mitosis and meiosis; these include H3S10ph and H3S28ph that are also associated with condensed chromatin (*27*).

Histone PTMs also regulate their own inheritance patterns; for example, phosphorylation of threonine 3 on H3 (H3T3ph) plays a role in asymmetric distribution during recycling of pre-replication histones onto sister chromatids (*28*). Akin to their links to gene regulation, histone PTMs are crucial markers for signalling and coordinating DNA damage response.

1.6 DNA damage can occur spontaneously and at a high rate

DNA damage is a phenomenon that endangers the integrity of the genome. Several factors can contribute to DNA damage that can manifest itself in different phenotypes. Environmental factors are one category and include high-energy radiation of various wavelengths such as γ-rays, X-rays, ultraviolet (UV), and infrared (*29*). These types of radiations can excite electrons leading to the ionization of atoms which affects their chemistry. In addition, by-products of endogenous oxidative metabolism or exogenous factors such as heavy metals and genotoxins contribute to the production of free radicals known as reactive oxygen species (ROS) that electron deficiencies prime them to "snatch" electrons from other atoms that they encounter (*30*). Such phenomena can contribute to single or double stranded breaks in DNA or fusion of molecules with the nucleotides. Natural processes such as DNA replication and mitosis can also jeopardize DNA integrity as they are error-prone. For example, during replication, an incorrect nucleotide can be inserted into the DNA, and the mismatch can result in local bulging of the DNA double-helix (*31*), and unfaithful or aborted activity of topoisomerases can induce single-strand breaks (SSBs) (*32*). During chromosome segregation, errors in the coordination of the mitotic spindle can shatter chromosomes and form many double-strand breaks (DSBs) (*33*). DSBs can also be caused by the collapse of stalled replication forks (*34*).

1.7 Unrepaired DNA damage pose high-risk problems

Unrepaired DNA poses great risks to genomic integrity. For example, unrepaired mismatched DNA can lead to permanent mutation of the genetic code in subsequent cellular divisions. Single-stranded DNA lesions that occur at a high rate can lead to stalling of RNA-polymerase and DSBs can lead to chromosomal rearrangements such as fusion or translocation of genes (*35*). Many diseases are the consequence of DNA damage and defects in their repair mechanism or their epigenetic regulation. These include developmental defects, neurodegenerative disorders, premature aging, and several cancers (*36*). For example, deficiency of DNA damage repair pathways are prevalent in prostate cancers (*37*). Inheritance of germline mutations in proteins that recruit DNA repair machinery are found in children suffering from sponastrime dysplasia who develop abnormal bones and stature around the spine, neck, and face (*38, 39*). Translocation of chromosome 21 as a result of breakage and fusion to other chromosomes has been linked to a rare kind of Down syndrome and is related to increased binding of DNA repair proteins even in the absence of DNA damage (*40*). Individuals with the rare autosomal recessive ataxia talangiectasia disease, which affects DNA damage sensors, develop weakened immunity, lung disease, neurological problems, and an increased risk of cancer (*41*).

1.8 Context-specific DNA damage repair mechanisms

To combat the deleterious effects of DNA damage, cells respond to DNA damage in a context-specific manner. These processes require sensing of the type of DNA damage and mounting the appropriate DNA damage response (DDR).

For example, DNA glycosylases detect and remove alkylated or deaminated nucleotides via the base-excision repair (BER) pathway, which also involves the replacement of the excised base and ligation of new insertions (*42*). Similarly, the nucleotide-excision repair (NER) pathway removes

and restores bulky DNA lesions caused by cross-linked DNA (*43*). The most dangerous type of DNA damage are DSBs as they have the potential of large-scale changes via chromosomal rearrangements and cells repair DSBs via two major mechanisms: non-homologous end joining (NHEJ) and homologous recombination (HR).

1.9 NHEJ and HR have common sensors but diverge in mediators and effectors

In the early stages of DDR, the MRN (MRE11, RAD50, NBS1) complex plays an important role in marking damaged sites and spreading the stress signal by recruiting ATM kinases (*44*) that in turn signal for cell-cycle arrest via phosphorylation of cell-cycle regulators such as p53 (*45*) and by phosphorylating H2AX on serine 139, a variant-PTM known as γH2AX, to amplify the stress signal by marking the damaged region (*44, 46*).

Recruitment of BRCA1 is also involved in the sensing stage of the DDR. Specifically, via recognition of γH2AX, MDC1 recruits the RNF8 ubiquitin-ligase assembly and ubiquitinates H2A (*47*) and γH2AX near the site of damage (*48*). This is important for translocation of BRCA1 and its binding partner BARD1 which have ubiquitin-ligase activity of their own. Via their RING domains, BRCA1-BARD1 ubiquitinate H2A K127 and K129 on nucleosomes and further the stress signalling (*49*). Interestingly, both MRN and BRCA1 play additional roles in mediating HR-specific responses.

1.10 Contribution of NHEJ to DSB

The NHEJ pathway involves aligning and rejoining of DNA at broken ends and is the first line of response for DSBs, accounts for the majority of DSB repairs, and can occur during the G1, S, and G2 phases of the cell cycle. Although the involvement of NHEJ is not fully ruled out during mitosis (M), its participation is somewhat of a question as key factors that promote this pathway are

suppressed. Studies comparing the activity of NHEJ during different cell cycles have shown that its contribution to DSB repair peaks in the G1 phase (*50-52*). While NHEJ is considered the default DSB response, it can be overridden by an alternative method of repair if the repair attempt by NHEJ is unsuccessful or a higher fidelity choice of repair such as HR is available. Additionally, the structure of a DSB break can influence choice of repair (*50, 51*).

Downstream of the initial sensors, the main NHEJ effectors form foci to restore DNA integrity. NHEJ is driven by the localization of Ku70/80 DNA-binding complexes (*53*) and DNA-Pk (*54*) proteins to the DSB site and 53BP1 plays an important role in promoting this pathway by protecting the broken ends and preventing resection of DNA by MRN or limiting it to short segments (4-5 nucleotides) (*45, 47, 55-59*). 53BP1's localization is promoted by binding of its Tudor domain to H4K20me2, a mark associated with pre-replicative chromatin (*60*), and also by binding of its BRCT domain to γH2AX (*59*). A double-stranded DNA break poses a ligation challenge as two ends of the breakage do not necessarily remain compatible for rejoining by NHEJ. Assessment of the broken ends with respect to overhang homology or presence of blunt ends determines how and which polymerases, ligases, and nucleases are recruited during NHEJ (*50*).

1.11 HR is available during the S-phase

The homologous recombination DNA repair pathway is primarily active during the S-phase of the cell cycle and is shut off leading to mitosis (*61*). This pathway relies on pairing the damaged DNA with another DNA molecule to which it has homology as a template for restoring DNA sequences in the break region. Accordingly, during the S-phase, DNA is replicated, and sister-chromatids provide this template. In this manner, HR is equipped with an advantage with minimal risk of loss of DNA sequences compared to NHEJ and is the preferred method of DSB repair. Thus, HR signals its availability by counteracting the promoters of NHEJ.

A key signal that indicates sister chromatid availability is H4K20 in unmethylated form (H4K20me0) which indicates a post-replicative and premature chromatin status (*62*). This mark, promotes HR in two ways; first by deterring 53BP1 from binding the histone and negatively impacting is localization but this alone is not sufficient for overriding NHEJ (*58*); and second, by recruiting BRCA1. The activation of BRCA1 on the pathway-specific mediation front is via the ankyrin domain of its binding partner BARD1 which specifically recognizes H4K20me0 (*58*). Critically, BRCA1 recruitment to the site of damage is required for depletion of 53BP1 and this promotes the formation of resected DNA by the MRN complex to produce long ssDNA overhangs (*63*).

Another key step in HR involves coating of the ssDNA by RPA proteins that protect it from nucleases but also help its recognition as an invading strand that requires the recombinase protein RAD51 for its insertion into homologous sequences of the sister chromatid (*64*). However, recruitment of RAD51 is dependant on its interaction with two binding partners TONSL and MMS22L. Similar to BRCA1, the ankyrin repeat domain of TONSL specifically binds H4K20me0 (*62*) and MMS22L directly interacts with RPA proteins where RAD51 is directed via the TONSL-MMS22L heterodimer (*65, 66*).

HR, similar to NHEJ and other mechanism that rely on identifying marked sections of DNA, to a great extent depend on reading, writing, and erasing PTMs on histones.

1.12 Histone readers recognize PTMs related to DNA damage repair

Post-translational modifications (PTMs) of histone proteins provide a complex signalling circuitry. They are integral to regulating important cellular functions and can act as epigenetic memory for coordinating processes that cross paths by tracking and triggering them in a spatiotemporal manner. For example, many essential processes such as DNA replication, gene expression, or DNA

damage repair, which are facilitated by histone PTMs, rely on access to genes in chromatin. In turn, regulation of changes to chromatin structure that promote its relaxation or condensation for meeting these requirements are also controlled by histone PTMs. Therefore, such processes must sense their environment for cues and recruit their effectors accordingly. A prominent way by which these processes check for their pre-requisites is by employing specialized protein domains that bind histone sequences and "read" the status of one or more sequences for the presence or the type of PTMs.

1.12.1 Alpha-helical histone readers

Over the past few decades, many histone readers have been discovered to preferentially bind to the methylated, acetylated, or phosphorylated forms of histone proteins. The majority of these domains are composed of a mixture of secondary structures such as plant homeodomains (PHD), chromodomains, and Tudor domains, but few are unique in that they feature only a single structural motif such as WD40 domains that form propeller-like bladed β-sheets (β-propellers). In the following subsections I review the three histone readers 14-3-3s, ARDs, and BRDs and compare how these domains use various structural mechanisms to recognize unmodified or modified histones while all are composed of an α-helical fold yet very diverse in their tertiary structures, giving rise to intriguing peptide binding mechanisms resulting in vastly different footprints of their targets.

1.12.2 14-3-3 proteins read phosphorylated S10 and S28 on histone H3.

The first report of these proteins described them in the 1960s as a ~30kDa protein isolated from bovine brain tissue that was named after its elution fraction [14] during chromatography and electrophoresis banding patterns [3-3] after gel filtration (*67*). So far, within the 14-3-3 family, fifteen members have been found in plants (*68*), seven in humans (*69*) (plus two phosphorylated

isoforms) (*70*), two in yeast (*71*), and one in protozoans such as amoeba and intestinal parasites (*72, 73*).

Early studies of 14-3-3s reported that these proteins are abundant in bovine brain (*67*) and rabbit optic nerve tissues (*74*) and in the spinal fluid of humans with neurogenerative disorders (*75*). The first characterization of the biological function of 14-3-3s accordingly linked these proteins to biosynthesis of essential neurotransmitters via hydroxylase activation (*76*) and later, 14-3-3s were more widely linked to the functions of the nervous system and several related diseases such as Miller-Dieker syndrome (12796778), ALS, Alzheimer's, Parkinson's (*77*), and Cruezfeldt-Jacob's disease (9450766). These proteins are now known to bind to more than 300 targets (*78*) with some of them linked to apoptosis and cell proliferation (*79*).

The 14-3-3 proteins have also been implicated in the regulation of transcription via epigenetic signalling by reading the phosphorylated state of S10 and S28 of histone H3. For instance, 14-3-3 binds H3S10ph deposited by PIM1 and recruits the histone acetyltransferase MOF that deposits the H4K16ac mark in the enhancer region of the transcription-factor-related gene FOS1L and promotes phosphorylation of RNA polymerase II and transcription elongation (*80, 81*). The same mark (H3S10ph) is also deposited at promoter regions by MSK1/2 kinases and is recognized by 14-3-3, and can occur in combination with K14ac or K9ac on the same histone (*68, 80*).

Assays comparing the binding affinities using libraries of 14-3-3-binding peptides have suggested the two motifs RSX[S^{ph}/T^{ph}]XP and RXX[Y/F][S^{ph}/T^{ph}]XP as consensus binding sequences (*82, 83*). However, in some cases the target sequences differ such as the sequences in the H3S10 region (ARKSTG) that only have a single amino acid (a lysine) between the arginine at the beginning of the motif and the phosphoserine. The same pattern is duplicated for S28 region (ARKSAP) and similarly S28ph is recognized by 14-3-3. Studies monitoring the response to gluconeogenesis show

that in mice, during fasting, H3S28 is phosphorylated and this mark recruits 14-3-3 and stimulates polymerase activity on gluconeogenic genes that are turned off after feeding by upregulation of the PP2A dephosphorylase (*84*). Interestingly, the presence of the P at the +2 position near S28 on H3 conforms to the proposed motifs and assays comparing the binding of various sequences at this position show a preference for the proline (*82*). Accordingly, an H3 peptide of the S28ph motif binds 14-3-3ζ ~10-fold more strongly than an S10ph peptide, but interestingly with a similar K_d if K14 is acetylated on the same peptide [S10ph+K14ac] and with a 2-fold difference with K9ac (*85*).

An individual 14-3-3 is composed of 9 α-helices arranged in a helix-turn-helix pattern. The first two α-helices form a pair that are shortest in length and are followed by a pair of α-helices that are distinctly long. The remaining five α-helices, which are of similar lengths, compose the rest of the protein (Figure 1.1A). The protein folds in a shape highly resembling of a "thumbs up" hand gesture, with the long and short pairs of helices forming the thumb and its base, respectively, while the rest of the helices form the remaining fingers curved halfway to the palm (Figure 1.1A). In this manner, the 14-3-3 fold creates a half-pipe with a handle with a concave groove suitable for peptide binding (Figure 1.2A). Studies comparing two isoforms of 14-3-3 report that many regions on their outer surface have non-conserved residues and vary substantially in amino acid properties (*71*), consistent with many binding partners interacting with 14-3-3s. However, the residues lining the amphipathic groove are nearly identical (*71*), suggesting that across several isoforms and homologs, these proteins have evolutionarily remained specialized Sph/Tph readers. Within the conserved groove, electrostatic mapping of the 14-3-3 surface reveals a distinct basic patch (Figure 1.2A) that is also conserved and plays an important role in the binding of the phosphorylated form of serine and threonine residues.

The 14-3-3 proteins feature an extended binding mode facilitated by homo- or heterodimerization. This is mediated by the pairs of short and long α-helices on both monomers. The two pairs place the backs of the “thumbs” together in opposite directions and bring their half-pipes close to each, transforming the U-shape channel of a monomer into a W-shape peptide binding cleft in the dimer. (Figure 1.1A). In this manner, these proteins can serve as dual readers with the flexibility to read two phosphorylated serine/threonine residues located on two different targets or at various distances apart on the same target.

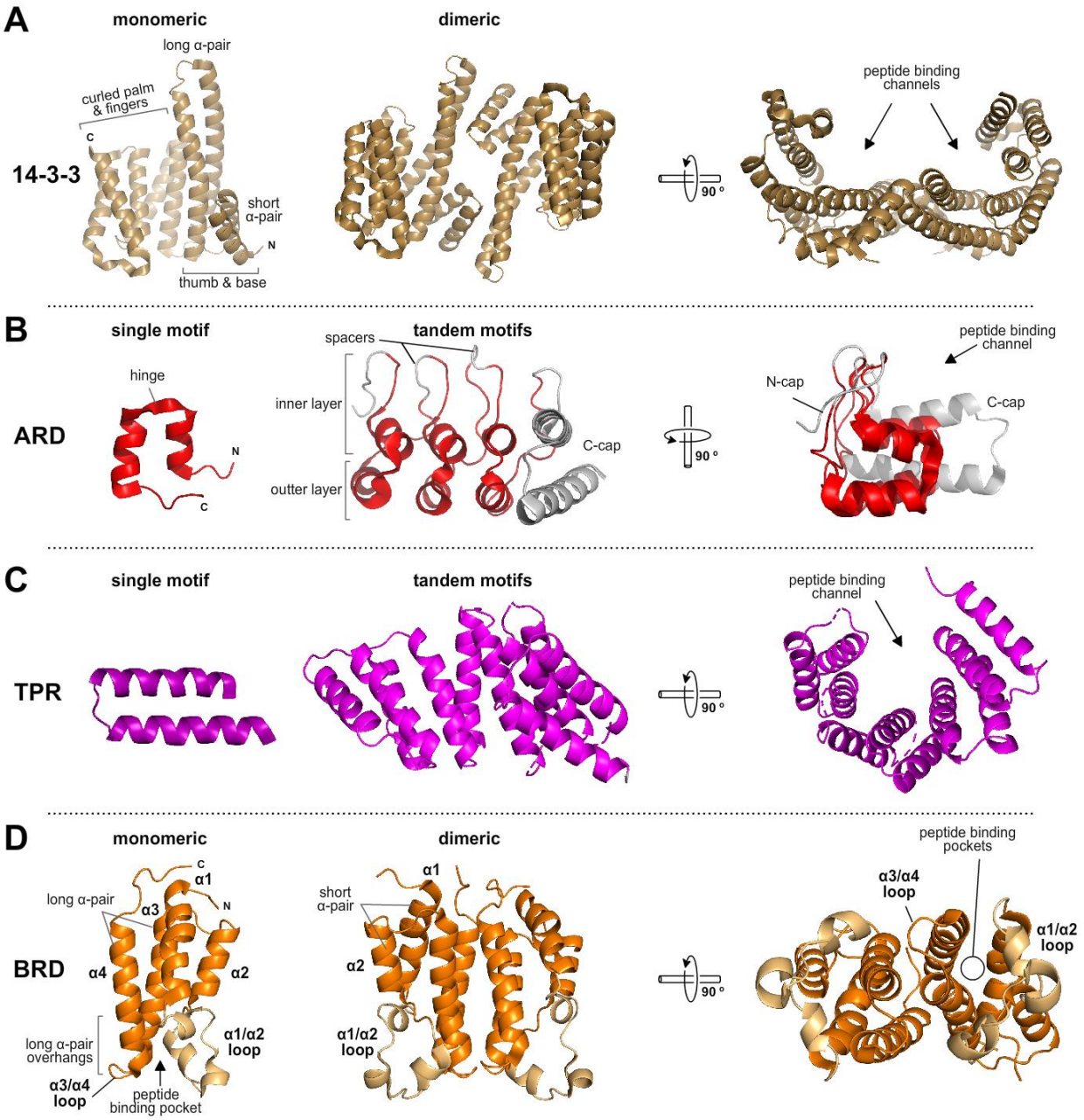


Figure 1.1 – Comparison of the structure of α-helical histone reader domains. The arrangement of α-helices are shown as cartoons for 14-3-3 [PDB ID:2C1J (ref. *80*)], ARD [PDB ID: 5JA4 (ref. *86*)], TPR [PDB ID: 7T7T (ref. *87*)], and BRD [PDB ID: 1X0J (ref. *88*)]. The monomeric/single motif forms of the reader domain are shown on left, and examples of higher order copies are shown in the centre. Panels on the right show a rotation of the higher-order structures at the indicated axis and angles. Panel B shows canonical ARD sequences in red and spacers/non-canonical ARD-like caps in grey. In panel D, the α1/α2 loop is emphasized in light orange and the rest of the motif is shown in darker orange.

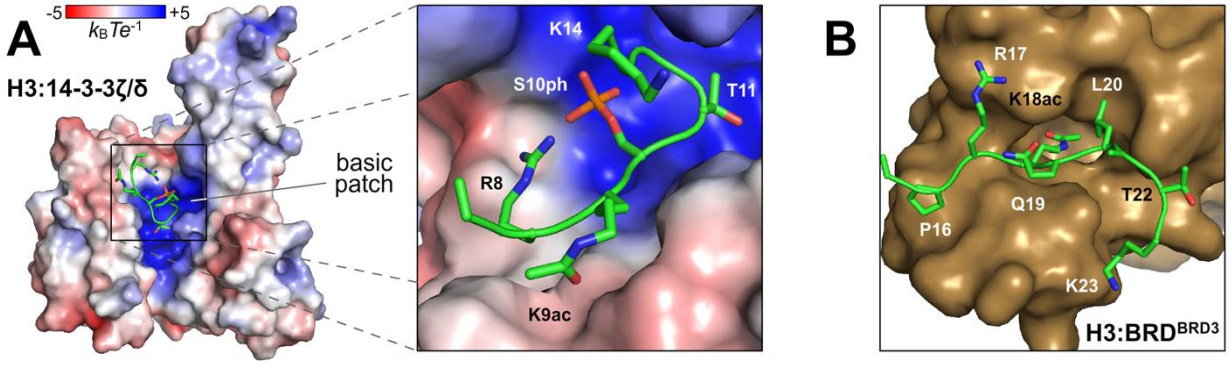


Figure 1.2 – Histone H3 binding pockets of 14-3-3 and BRD domain. (A) Surface representation of 14-3-3ζ or its phosphorylated form δ in electrostatic potential gradient shown in red (negative), white (neutral), or blue (positive) in k_BTe^{-1} units where e is the electron, T is absolute temperature, and k_B is the Boltzmann constant. A closeup of the H3-binding region rotated 90º counter clockwise is shown in the window on the right. H3 residues are shown as green cartoons (backbone) or sticks (side chains) [PDB ID:2C1J (ref. *80*)]. **(B)** Surface representation of the H3K18ac binding pocket from the second bromodomain of BRD3 protein is shown in brown. Histone H3 backbone is rendered in a green cartoon or sticks (side chains) [PDB ID: 5HJC (ref. *89*)].

1.12.3 Ankyrin repeat domains read H3K9me1/2 and H4K20me0

Initially identified in *Drosophila melanogaster* Notch/LIN-12 and *Saccharomyces cerevisiae* Swi6/Cdc10 proteins (*90*), the Ankyrin repeat domain (ARD) was later named due to its presence in the cytoskeletal ankyrin protein (*91*). One of the most abundant domains, an individual ankyrin comprises of approximately 33 residues that can repeat in tandem forming ARDs of various repeat numbers. At the N-terminal region of a single ARD, its amino acids form a three-residue short loop leading to a pair of short antiparallel α-helices connected, in most cases, by a conserved glycine that acts as a hinge that enables the formation of a very short turn between the two helices (Figure 1.1B). The C-terminal region is formed by a 10-mer loop and, like the N-terminal loop, projects outward perpendicular to the preceding α-helix (*92*), shaping a single ankyrin into a structure resembling two "L" letters attached at the top with a tight hinge (Figure 1.1B). In an ankyrin tandem repeat, these loops connect the copies and together form hairpins with β-hairpin turns (4-residue turns) between the α-helical regions (Figure 1.1B).

The number of repeats between ARDs varies considerably (*92*). For example, as many as 24 repeats have been identified in the poly(ADP-ribose) polymerase Tankyrase (*93*), while the majority of proteins that feature them have less than 10 and most frequently 3 or 4 repeats, such as in the DNA repair protein TONSOKU-Like (TONSL) (*86*) and tumour suppressor-related 53BP2 (*94*), respectively. Each repeat adds about 13º of a left-handed twist, and in tandem, the α-helices stack side-by-side (Figure 1.1B, Figure 1.3), forming a double-layered curve as the ARD structure grows with additional repeats (*95*). In this arrangement, the first α-helix of the individual ankyrin repeats makes the outer layer while the inner layer is composed of the second α-helices and the juxtaposed β-hairpins between the repeats creating many finger-like projections lining the edge of the inner layer. In such confirmation, the inner layer creates a groove where peptide binding

occurs (Figure 1.1B, Figure 1.3) (*92, 96*). ARDs can also differ by a few features that aid in diversifying target recognition. For instance, ARDs can include up to 4-5 spacer residues between each repeat that adjusts the length of the hairpins in the corresponding regions (Figure 1.1B). Additionally, ARD can be C- or N-terminally "capped" by non-canonical ankyrins (Figure 1.1B, Figure 1.3).

Early studies established that ARDs play important functions in mediating protein-protein interactions enabling their involvement in signal transduction, inflammatory response, cytoskeletal integrity, transcription, and development (*92*). Interestingly, besides proteins, ARDs have been shown to bind sugars and lipids and in biotechnology, synthetic ARDs named Designed Ankyrin Repeat Proteins (DARPins) are developed as an alternative to antibodies for protein detection as they offer the advantages of being more resilient as they lack disulfide bonds and being smaller and easier to express while retaining high specificity (*97*).

ARDs also play a significant role in epigenetic signalling. For example, ARDs are found in the euchromatic histone-lysine N-methyltransferase and G9a-like protein complex (G9A-GLP). These methyltransferases modify and read the mono- and di-methylated form of H3K9 (*98-100*). While their SET domains carry out the methylation reaction, G9A-GLP H3K9me1/me2 reading modules map to their ARDs. Fluorescence polarization assays show that while G9a has a stronger preference for H3K9me2 and GLP for H3K9me1, binding is reduced by 10- to 100-fold between GLP/G9a and H3K9me0 or K9me3, respectively (*98*).

In GLP, the ARD comprises 6 ankyrin repeats with a C-terminal cap composed of a non-canonical ankyrin repeat (Figure 1.3A). A crystal structure of ARD^{GLP} bound to H3 residues 1-15 with a K9me2 modification (*98*) reveals three distinct pockets aligned linearly on the surface of the ARD groove and arranged in order according to their pocket size (Figure 1.3A). While the small pocket

is vacant, the medium pocket is occupied by K9me2. This pocket comprises three tryptophan residues (Trp881, Trp843, Trp848) that form a partial hydrophobic cage surrounding H3K9me2. A glutamic acid (Glu851) residue completes the cage, and its carboxylate closely interacts with H3K9me2 ε-amine (Figure 1.3A). Mutations of any residues forming the cage negatively impact histone H3 binding (*101*). The neighbouring large pocket accommodates histone H3 residues S10-G13. In the large pocket, additional contacts are made between the ARD^{GLP} and the side chains of S10 and T11, and while not tested by the authors, the architecture of this spatially restrictive pocket suggests that any PTMs on these two residues will negatively impact binding. Around the same pocket, Asp905 and Arg913 from ARD form a wall that, together with H3 G12 and G13, facilitate a sharp turn of the H3 backbone out of the pocket (*98*). Interestingly, the residues forming the K9me2 and S10-G13 binding pockets between ARD^{G9a} and ARD^{GLP} are conserved (*98*).

Unlike G9a/GLP ARDs, which prefer methyl-lysine residues (me1/2), the ankyrin repeat domains of TONSL and BRCA1-associated RING Domain 1 (BARD1) prefer unmethylated lysine residues. TONSL and its heterodimerizing partner MMS22L participate in the recruitment of RAD51 to repair damaged DNA at sites of double-strand breaks, and their activity is very important for maintaining genetic integrity during replication (*66, 102-104*). TONSL ARD recognizes histone H4 on K20, and its binding is inhibited by the methylation of this residue (H4K20me1) (*86*). The structure of $TONSL^{ARD}$ bound to histone H4 (*86*) reveals several interesting structural features (Figure 1.3B). The C-terminal ARD-like cap structure binds K12-G14 residues, while the three canonical ARDs bind A15-K20 of histone H4. The hairpin-shaped N-terminal extension binds histone H4 V21-Q27 and contacts the histone H4 core domain (Figure 1.3B). The binding pocket for H4K20 is located near the N-terminal region of the ARD^{TONSL} and employs both aliphatic and polar interactions to stabilize the unmodified lysine. In this pocket,

conserved residues Asp559, Glu530 and Glu568 of ARD^{TONSL} make hydrogen bonds with the H4K20 ε-amine while His534 and Trp563 restrict the space around the lysine side chain (Figure 1.3B) and further make this pocket unsuitable for any modifications on H4K20.

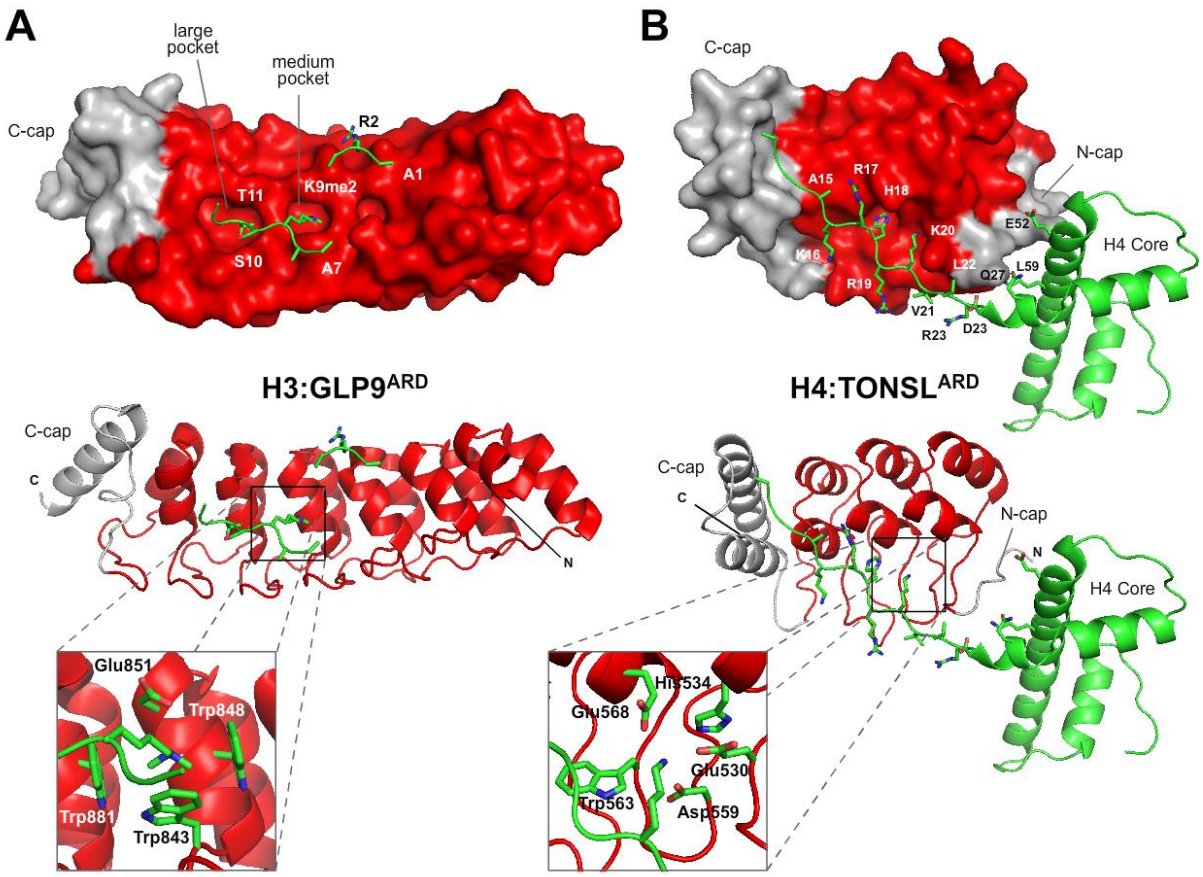


Figure 1.3 – ARD domains of GLP9 and TONSL bind H3 and H4 peptides with different mechanisms. Ankyrin repeat domains of GLP9 [PDB ID: 3B95 (ref. *98*)] and TONSL [PDB ID: 5JA4 (ref. *86*)] are compared in surface representation [top] and in cartoon depiction [center]. In both structures, canonical ARD sequences are shown in red and ARD-like caps at C- or N-terminals are shown in grey. The windows traced with dashes at the bottom show a close-up view of the binding pocket in the corresponding structure. Histone proteins are shown as cartoons (backbone - green) or sticks (side chains).

1.12.4 Tetratricopeptide repeats (TPRs) bind long unstructured peptides

Analysis of the *CDC23* gene in 1990 led to identifying a 34 amino-acid motif with repeating patterns. (*105*). Other studies found the same motif via computational studies, and sequence alignment of proteins with this motif lead to the characterization of the tetratricopeptide repeat (TPRs). The first structure of a TPR (from Hsp70-binding protein Hop) demonstrated that the TPR folds in an anti-parallel helix-turn-helix structure (Figure 1.1C) (*106*) with short turns of 4-5 amino acids, which frequently include glycine or proline residues. TPRs can form super helices when placed in tandems (Figure 1.1C, Figure 5.1) and bear a close resemblance to PPR (pentatricopeptide repeats) [35aa] and SLR (suppressor-enhancer of lin repeats) [36-44 aa] motifs , which similar to TPRs can form a solenoid structure (*107*). Based on the position and types of amino acids, a motif was suggested for identifying TPRs. However, recent structural studies showed that certain TPRs can deviate from these motifs. Moreover, several poly-TPR domains (tandems) have been identified that include the TPR fold but are composed of non-canonical residues.

Individual TPR domains form a concave surface in the space between their two helices, resulting in a cleft suitable to bind a ligand (*108*). Tandems of 3 to 16 TPRs have been reported to form right-handed super helices, which extend the ligand binding area to longer unstructured peptides where the amphipathic channel created by the TPR array can fine-tune binding to polar and hydrophobic contacts (Figure 1.1C, Figure 5.1). Furthermore, while TPR tandems can adopt the same structural fold, they bind different peptides in different proteins suggesting that sequence variability in TPR residues determines the target specificity. This is well demonstrated in the binding preference of TPR1 and TPR2A of Hop which distinguish Hsp70 and Hsp90, respectively, even though the target peptides of these protein share a common and conserved EEVD motif (*106*).

Protein engagement can also be tuned by the ability of some poly-TPR proteins to oligomerize. This is the case of O-linked GlcNAc transferase (OGT) (*109*) poly-TPR domain that can readily oligomerize in solution.

TPRs control diverse biological processes such as protein transport, enzymatic inhibition, and transcription and are found in both prokaryotes and eukaryotes and in proteins localizing to the nucleus, cytosol, membrane, and mitochondria (*110*).

In several proteins, TPRs are known to bind long and unstructured peptides. For instance, the TPR domain of SCC4, a subunit of cohesin complexes that aid in sister chromatid segregation during anaphase, binds over a hundred amino acids from the N-terminal region of another cohesin subunit SCC2 (*111, 112*). Similarly, within the transcription machinery, the RNA-polymerase associated protein Ctr9 binds an extensive unstructured region of the elongation factor Paf1 (*113*).

1.12.5 Bromodomains are readers for acetyl-lysine residues on histones H2A, H2B, H3, and H4

This class of α-helical reader domains derives their name from the *Drosophila* Trithorax Group protein Brahma (BRM), a homolog of the yeast SWI/SNF chromatin remodellers (*114, 115*). Bromodomains (BRDs) [brahmadomains] are found in over 60 variations and in over 40 proteins, many of which are chromatin remodellers (*116*). These domains commonly recognize acetyl-lysine residues (*117*) and are frequently found among other reader and writer domains on epigenetic signalling proteins such as histone methyltransferases (HMTs) and histone acetyltransferases (HATs) (*116*).

Similar to 14-3-3s, but unlike ARDs and TPRs, bromodomains do not occur in tandem repeats; although a protein may have more than one bromodomain that may be near or at a distance from each other such as in BRD4 [two spaced out BRDs] (*118*) and poly-BRD1 [PB1] [six BRDs within

same region] (*119*). Bromodomains can form left- or right-handed structures based on the twist of their helices. These structures are canonically composed of 4 anti-parallel α-helices which are designated as αZ, αA, αB, αC (N- to C-terminal) (Figure 1.1D) in left-handed BRDs, and as α1-4 in right-handed bromodomains (*120*). The first two helices in BRDs are typically shorter (3 to 4 turns), while the last two C-terminal helices are longer by one or two turns. Short loops (2 to 4 a.a.) connect α2 to α3 and α4 to α3. However, a distinctly long loop (~32 a.a.) connects the two short N-terminal α-helices (α1 to α2) (Figure 1.1D). In some BRDs, the long loop may form short (1 or 2 turns) α-helical structures such as in the N-terminal bromodomain of the protein kinase BRD2 (*88*) (Figure 1.1D), the bromodomain of CBP (*121*), and the second bromodomain of BRD3 (*89*). Similar to ARD domains, some BRDs have N- or C-terminal caps creating appendages partially analogous to the rest of BRD's fold. For example, the sixth bromodomain in PB1 features a fifth and long α-helix (*122*).

The structure and mode of peptide binding in bromodomains is quite different than 14-3-3s, ARDs, and TPRs, in that they do not form long concave grooves for peptide binding along the length of their tertiary structure. BRDs, form elongated structures by arranging their four α-helices in a bundle like a gesture using four fingers in which the thumb, index, middle and ring finger are brought together in parallel. On the top side of this α-helical bundle, the helices almost align and fill the space inside of the α-helical bundle; however, at the opposite side (bottom), the two longer helices produce overhangs relative to the short pair of α-helices. This gap is filled with residues from the long α1/2 loop which points away from the α-bundle. With contribution from sequences in the α3/4 loop, this region forms the peptide binding cleft (Figure 1.1D). The arrangement of the a-helices and the loops creates a deep pocket (Figure 1.2B) composed of hydrophobic residues important for binding the acetyl-lysine residue. Additionally, several conserved hydrophobic

residues aid in the clustering and arrangement of the BRD α-helical bundle. Furthermore, the participation of water molecules around the acetyl moiety, an asparagine residue within the α3/α4 loop, and a PxY motif within the α3/α4 loop are widely conserved in BRDs (*116*). Interestingly, a crystal structure of BR1 (domain) from the BRD2 protein captured this domain in monomeric and dimeric forms (*88*). Reciprocal co-immuniprecipitation with *in vivo* co-expression revealed that the homodimer has biological relevance, and binding assays showed that mutation of several residues identified at the interface of the BR1 homodimer drastically reduced binding to a histone H4K12ac peptide (*88*). In the homodimer of $BR1^{BRD2}$, the double bromodomain is formed by back-to-back placement of their long α-helices so that the peptide binding sites face the same way (Figure 1.1D).

1.13 Rationale and Hypothesis

Studies have shown that the α-helical ARD domain in TONSL is a reader of histone H4K20me0 (*62*). The recognition of the unmethylated histone leads to the accumulation of TONSL at DSB sites and promotes homologous recombination during replication. Furthermore, methylation of H4K20 by SET8 prevents spurious TONSL recruitment. In plants, no SET8 homolog has been identified, and the ARD is not conserved in the plant ortholog of TONSL, TSK. Interestingly, a TPR domain is conserved on the N-terminal regions of both TSK and TONSL. TPR domains have been reported to assume alpha-helical folds and bind long, unstructured peptides. Binding assays show that TONSL binds histone H3 via an unknown region. Additionally, in plants, mutation of TSK rescues defects linked to the loss of histone H3.1 H3K27 mono-methyltransferases ATXR5/6.

I hypothesize that in plants, TSK recruitment to chromatin depends on its TPR domain in (TPR^{TSK}) binding to histone H3 and that ATXR5/6 play regulatory roles for operating the HR switch in plants by regulating TSK-binding via methylation of H3K27.

To address this hypothesis, I used structural biology and biochemical assays to characterize these proteins and their complexes.

In Chapter 2 and Chapter 3 I focus on the role of TPR^{TSK} and its structural and biochemical characterization, and in Chapter 4, I explore regulation of ATXR5/6 activity.

Chapter 2 | The histone H3.1 variant regulates TONSOKU-mediated DNA repair during replication

Chapter 2 is a manuscript published on 17 March 2022 by the American Association for Advancement of Science (AAAS) in the Science journal. The publisher permits use of this publication in an academic thesis.

Science | Volume 375 | Issue Number 6586: 1281-1286.
https://www.science.org/doi/10.1126/science.abm5320

2.1 Authors and Affiliations

Hossein Davarinejad[1†], Yi-Chun Huang[2†], Benoit Mermaz[2], Chantal LeBlanc[2], Axel Poulet[2], Geoffrey Thomson[2], Valentin Joly[2], Marcelo Muñoz[3,4], Alexis Arvanitis-Vigneault[1], Devisree alsakumar[5,6], Gonzalo Villarino[2], Alex Ross[3,4], Benjamin H. Rotstein[4,7], Emilio I. Alarcon[3,4], Joseph S. Brunzelle[8], Philipp Voigt[5,6], Jie Dong[2,9], Jean-François Couture[1*], Yannick Jacob[2*]

1 Ottawa Institute of Systems Biology, Department of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, Ontario K1H 8M5, Canada
2 Department of Molecular, Cellular and Developmental Biology, Yale University, New Haven, CT 06511, USA.
3 BEaTS Research Laboratory, Division of Cardiac Surgery, University of Ottawa Heart Institute, Ottawa, Ontario K1Y 4W7, Canada
4 Department of Biochemistry, Microbiology, and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, Ontario K1H 8M5, Canada.
5 Wellcome Centre for Cell Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3BF, UK.
6 Epigenetics Programme, Babraham Institute, Cambridge CB22 3AT, UK
7 University of Ottawa Heart Institute, Ottawa, Ontario K1Y 4W7, Canada.
8 Department of Molecular Pharmacology and Biological Chemistry, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA.
9 Institute of Crop Science, Zhejiang University, Hangzhou 310058, China.

* Corresponding author. Email: yannick.jacob@yale.edu (Y.J.); jean-francois.couture@uottawa.ca (J.-F.C.)
† These authors contributed equally to this work.

2.2 My Contributions

For this publication, I confirmed the histone H3.1-TSK interactions by performing ITC assays. I solved the structure of TPR^{TSK} bound to the tail of histone H3.1. I performed the *in vitro* pulldown assays and ITCs with histone H3.1 PTMs and performed the *in vitro* H3 binding studies with TPR^{TSK} mutants. I generated constructs and purified protein for all of my experiments. I generated the figures corresponding to my work (indicated at the end of each figure caption) and contributed to the writing of the manuscript.

2.3 Abstract

The tail of replication-dependent histone H3.1 varies from that of replication-independent H3.3 at the amino acid located at position 31 in plants and animals, but no function has been assigned to this residue to demonstrate a unique and conserved role for H3.1 during replication. We found that TONSOKU (TSK/TONSL), which rescues broken replication forks, specifically interacts with H3.1 via recognition of alanine 31 by its tetratricopeptide repeat domain. Our results indicate that genomic instability in the absence of ATXR5/ATXR6-catalyzed histone H3 lysine 27 monomethylation in plants depends on H3.1, TSK, and DNA polymerase theta (Pol θ). This work reveals an H3.1-specific function during replication and a common strategy used in multicellular eukaryotes for regulating post-replicative chromatin maturation and TSK, which relies on histone monomethyltransferases and reading of the H3.1 variant.

2.4 Introduction

Chromatin replication requires multiple regulatory mechanisms to ensure the maintenance of genome integrity. One of these mechanisms relies on TONSOKU-LIKE (TONSL), a key player in initiating homologous recombination (HR) when replication forks encounter double stranded DNA breaks (DSBs) (*65, 102-104, 123-125*). In animals, TONSL is recruited to chromatin via its ankyrin repeat domain (ARD), which specifically interacts with unmethylated histone H4 Lys20 (H4K20me0) (*62, 102*). Post-replicative maturation of chromatin is accomplished via SET8/PR-Set7/SETD8 (*126-128*), which monomethylates histone H4 Lys20 (H4K20me1) and thus prevents TONSL from binding chromatin and initiating HR-based DNA repair outside of DNA replication and the G2 phase of the cell cycle (*62*). Comparative analysis shows that plants contain a TONSL ortholog (TSK/BRUSHY1/ MGOUN3) (*124, 125, 129*) but lack SET8. In addition, the ARD of TONSL in animals is not conserved in TSK orthologs (Figure 2.1A) (*102*), indicating that post-replicative chromatin maturation in plants is unlikely to depend on the methylated state of H4 Lys20.

We reasoned that TSK might interact with histones in plants through a different domain. Sequence alignment of TSK orthologs shows extensive similarity in the N-terminal tetratricopeptide repeat (TPR) domain (fig. S1), which is conserved in animals (Figure 2.1A) (*66, 102*). Many TPR domains bind long peptides (>20 amino acids), adopting an extended conformation (*130*). We therefore hypothesized that one of the N-terminal unstructured tails of histones could specifically interact with the TPR domain of TSK (TPR^{TSK}).

2.5 Results and Discussion

To assess our hypothesis, we performed in vitro binding assays with *Arabidopsis thaliana* TPR^{TSK} and the tails of different histones. We detected binding of TPR^{TSK} with H3 variants, with stronger

binding for H3.1 relative to H3.3 (Figure 2.1B). A preference for TPR^{TSK} to bind H3.1 over H3.3 was also observed using nucleosomes and in *A. thaliana* protoplasts (fig. S2, A to C). In vascular plants, amino acids 31 and 41 vary between the N-terminal tails of H3.1 and H3.3 (Figure 2.1C) (*131*). We created hybridH3.1/H3.3 tail- GST fusion proteins based on these differences and determined that only alanine at position 31 of H3.1 (H3.1 Ala31) is required for the H3.1- binding specificity of TPR^{TSK} (Figure 2.1D). Variation at position 31 of H3 is also observed between replication-dependentH3.1/H3.2 variants and the replication-independent H3.3 in mammals (Figure 2.1C). Similarly to plant TSK orthologs, the TPR domain of mouse TONSL also interacts preferentially with H3.1 relative to H3.3 (Figure 2.1E). We then assessed the impact of TPR^{TSK} binding to H3.1 in the context of methylation at different lysine residues in the N-terminal tail of H3.1.We found that increasing levels of methylation at Lys4, Lys9, Lys27, and Lys36 have a negative impact on the interaction of TPR^{TSK} and H3.1, with binding being most sensitive to methylation at Lys27 (Figure 2.1,F and G, and fig. S2D). The binding profile of TSK on histone H3 suggests a preference for binding the newly synthesized H3.1 variant. To gain mechanistic insights into how TSK discriminates H3.1 from H3.3, we solved the crystal structure of the TPR^{TSK}-$H3.1^{1-45}$ complex at 3.17 Å resolution by using the TSK ortholog from *Citrus unshiu* (*Cu*TSK) (fig. S1 and table S1). TPR^{TSK} folds as 11 TPR motifs placed in tandem, which collectively form a hollow solenoid tube (Figure 2.2, A and B, and fig. S3A). The C-terminal lobe of the tube is composed of TPRs 6 to 11 and generates a wide channel in which two segments of H3.1 (Lys4 to Lys9 and Lys18 to Ala24) are found along opposite sides of its wall (Figure 2.2, A and C). In the center lobe, TPRs 3 to 7 form a narrow tunnel that encircles Ala25 to Pro30 ofH3.1 (Figure 2.2,A and C). H3.1 Lys27 is located inside a polar pocket, where its ε-amine is surrounded by the side chains of Asp234, Cys238, and Ser208 and the backbone carbonyl groups of Asp234

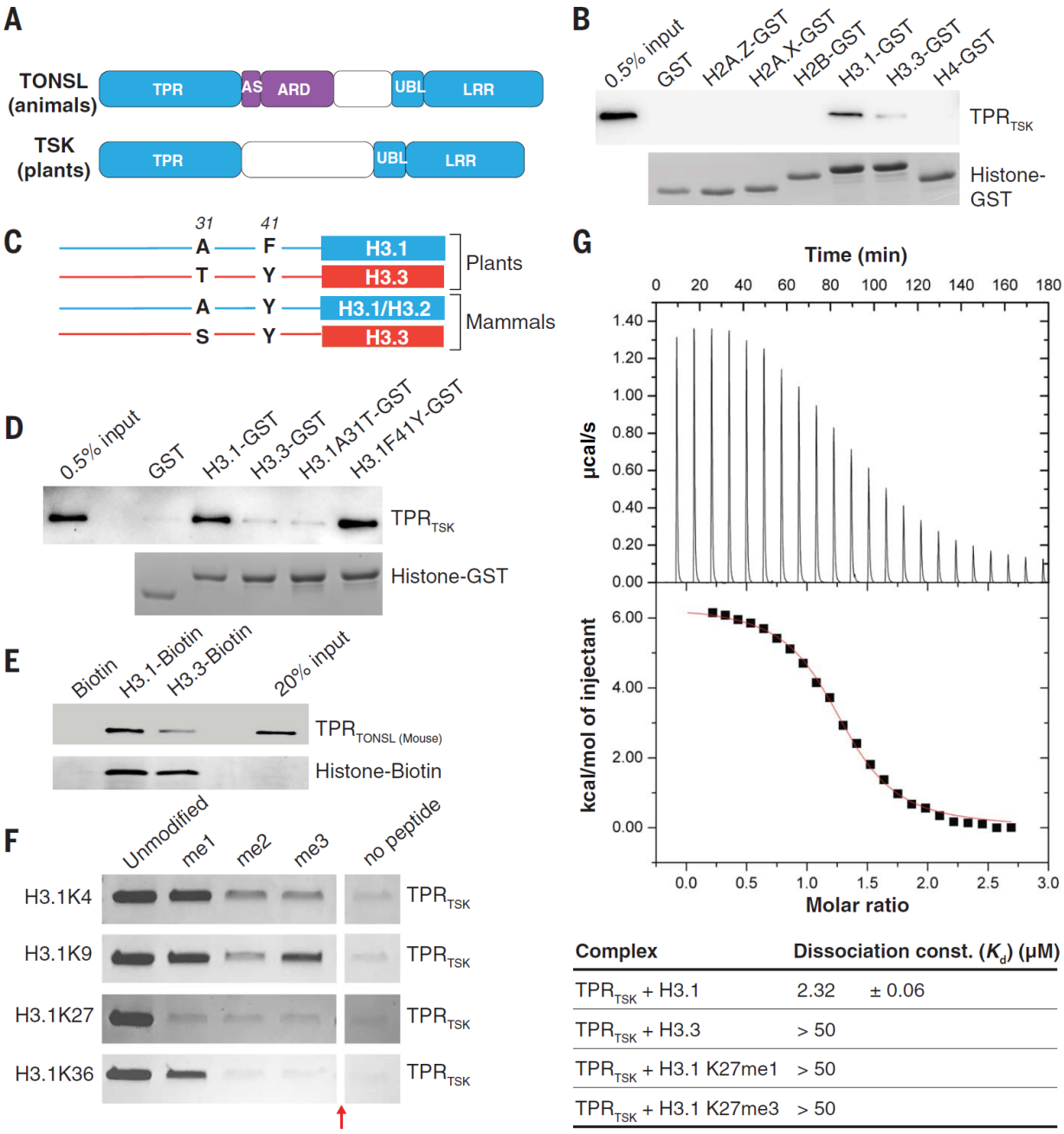
and Gly246 (Figure 2.2D). The polarity of this pocket makes it non conducive for the binding of hydrophobic moieties such as methyl groups, thus explaining the large decrease in binding affinity of TPR^{TSK} to H3.1 when Lys27 is mono- or trimethylated (Figure 2.1G and fig. S2D). TPRs 1 to 3 make up the N-terminal lobe of TPR^{TSK}, which forms an open channel that accommodates Pro30 to Arg40 of H3.1 (Figure 2.2, A and C). A deep pocket formed between α-helices 2 to 4 (TPRs 1 and 2) is occupied by the side chain of H3.1 K36, where its e-amine is in close proximity to the carboxyl group of Asp54 (Figure 2.2, A, C, and E). The side chain of H3.1 A31 is oriented toward the aliphatic portion of three residues (Arg109, Gln113, and Gln72) strictly conserved among plant TSK orthologs (Figure 2.2F and fig. S3A). These residues form a shallow pocket in which Gln113 and Gln72 also likely interact with the H3.1 backbone via hydrogen bonds with the carbonyl group of Gly34 and the amide group of Ala31, respectively (Figure 2.2, F and G). Consistent with our binding assays (Figure 2.1, B, D, and G, and fig. S2D), modeling of an A31T substitution (corresponding to H3.3) in H3.1 generates van der Waals clashes between the Cγ methyl group of Thr31 and the aliphatic chain of Gln113, and similarly between the hydroxyl group of Thr31 and Arg109 (fig. S3B). We mutated various amino acids of TPR^{TSK} from different H3.1 binding pockets and validated that they contribute to the TPR^{TSK}-H3.1 interaction (fig. S3C). Overall, the structure of the TPR^{TSK}-H3.1 complex supports our finding that TSK preferentially binds the replication dependent H3.1 variant. In plants, the histone H3 K27 monomethyltransferases ATXR5 and ATXR6 (ATXR5/6) maintain genome stability by specifically methylating the H3.1 variant (H3.1K27me1) during DNA replication (*24, 25, 132*). Loss of H3.1K27me1 in atxr5/6 double mutants results in genomic amplification of heterochromatin, transposable element derepression, and disruption of heterochromatin structure (*24, 133*). Additional work has shown that heterochromatin amplification in atxr5/6mutants is dependent on DNA repair (*134*). Therefore,

ATXR5/6 may play a role analogous to the mammalian H4K20 monomethyltransferase SET8 in regulating TONSL/TSK activity, with the difference that H3.1K27me1, not H4K20me1, is the key histone modification used in plants to prevent TSK from interacting with chromatin and initiating DNA repair. To validate this model, we generated an atxr5/6 tsk triple mutant in *A. thaliana* (fig. S4A). Flow cytometry analyses of atxr5/6 tsk mutants showed suppression of heterochromatin amplification induced by the absence of H3.1K27me1, as represented by the loss of the broad peaks corresponding to 8C and 16C endoreduplicated nuclei in atxr5/6 mutants (Figure 2.3A and fig. S4B). This result was confirmed by genome sequencing of 16C nuclei from leaf tissue (Figure 2.3B). We also observed transcriptional suppression of the genome instability marker BRCA1, which is highly expressed in atxr5/6 but not in atxr5/6 tsk (fig. S4C) (*135*). In addition, the number of chromocenters adopting a hollowed sphere conformation characteristic of atxr5/6 mutants was decreased when TSK was inactivated (Figure 2.3, C and D) (*134*). Similarly, transcriptional derepression in heterochromatin of atxr5/6 mutants was reduced in the absence of TSK (Figure 2.3E, fig. S4, D to F, and table S2). These results indicate that the heterochromatic defects caused by the loss of H3.1K27me1 in plants are dependent on TSK. In mammals, TONSL is recruited to newly replicated chromatin and promotes DNA repair via HR at broken replication forks (*65, 66, 102-104*). Cell cycle expression analysis in synchronized tobacco cells indicates that TSK is specifically expressed in S phase (*136*), which supports a conserved role for TSK during replication. To assess whether H3.1K27me1 suppresses HR activity in plants, we used a reporter system for HR based on intrachromosomal recombination restoring activity at a colorimetric GUS transgene (*137*). Our results show that GUS activity is much stronger in atxr5/6 mutants than in wild-type plants, but not in atxr5/6 tsk mutants (Figure 2.3F and fig. S5), thus indicating a role for H3.1K27me1 in preventing TSK-mediated HR in plants. The protein kinases ATM and ATR, which

participate in the early signaling steps leading to HR-mediated DNA repair (*138*), were previously shown to be required for inducing heterochromatin amplification in atxr5/6 mutants (*134*). We therefore tested the contributions of different DNA repair pathways to the phenotypes observed in atxr5/6mutants. Mutating nonhomologous end-joining (Ku70, Ku80, and LIG4) or HR (RAD51, RAD54, and BRCA2A/ BRCA2B) genes did not have a major effect on heterochromatin amplification in atxr5/6 mutants (Figure 2.3G), although eliminating the HR recombinase RAD51 enhanced the morphological phenotypes of atxr5/6 mutants (fig. S6A). In contrast, a mutation in RAD17 suppressed heterochromatin amplification, loss of chromatin structure, and transcriptional derepression in atxr5/6 mutants (Figure 2.3G and fig. S6, B and C). RAD17 is responsible for loading the MRE11-RAD50-NBS1 complex that mediates DNA resection, one of the initial steps of HR (*139*). In animals, DNA resection can also lead to substrates that are repaired in an error-prone manner by Pol θ via polymerase theta–mediated end joining (TMEJ), which can create large tandem duplications of 1 kb to 1Mb (*140*).We introduced a mutant allele of the A. thaliana POLQ/TEBICHI gene coding for Pol θ in the atxr5/6 background and observed strong suppression of heterochromatin amplification and related phenotypes (Figure 2.3G and fig. S7, A to D). Taken together, these results show that genomic instability in atxr5/6 mutants is caused by a TSK-dependent pathway involving TMEJ. The specificity of ATXR5/6 and TSK for replication-dependent H3.1 led us to hypothesize that this H3 variant is responsible for inducing TSK-mediated genomic instability in atxr5/6 mutants. To test whether A31 of H3.1 is required for heterochromatin amplification in the absence of H3.1K27me1, we used a genetic system based on expression of the H3.1 point mutant H3.1S28A that mimics the phenotypes of atxr5/6 mutants (Figure 2.4 A to F, and fig. S8A) (*141*). The S28A point mutation prevents H3.1 K27 monomethylation by ATXR5/6 (*141*) but does not affect the binding of TPR^{TSK} to H3.1 (fig. S8B),

thus supporting a role for H3.1K27me1 in preventing the interaction of TSK with H3.1 in vivo. We then transformed A. thaliana with a transgene expressing H3.1S28A A31T (Ala31 replaced with Thr, as in plant H3.3 variants) and observed suppression of the heterochromatin phenotypes (Figure 2.4, A to E, and fig. S8A), which demonstrates the importance of H3.1 A31 in regulating TSK activity in plants. The dependence of TSK on H3.1 explains why plants expressing H3.1A31T do not induce heterochromatin amplification despite losing ATXR5/6-catalyzed H3.1K27me1 (Figure 2.4A) (*25*). A role for H3.1 A31 in mediating TSK activity is also supported by the observation that plants expressing H3.1A31T are hypersensitive to genotoxic stress, similarly to tsk and h3.1 mutants (fig. S9, A to F). We also used the H3.1S28A genetic system to assess the role of K4, K9, and K36 of H3.1 in contributing to the interaction with the TPR domain of TSK. Our *in vivo* results show that alanine replacement at K4 and K36 almost completely suppresses genomic instability and transcriptional derepression mediated by expression of H3.1S28A (Figure 2.4F and fig. S10, A to D). These results are in line with in vitro experiments showing thatH3.1K4A and H3.1K36A, but not H3.1K9A, strongly disrupt binding of TSK to H3.1 (fig. S10E). Finally, we used CRISPR to create a septuple mutant background, where all five H3.1 genes are inactivated, in addition to mutations in ATXR5/6 (atxr5/6 h3.1) (fig. S11, A and B). In atxr5/6 h3.1 septuple mutants, both heterochromatin amplification and transcriptional derepression are suppressed (fig. S11, C and D), thus confirming that the H3.1 variant is required to induce these phenotypes. These results support the idea that TSK makes specific interactions with the N-terminal tail of the H3.1 variant in vivo to disrupt heterochromatin stability and silencing when H3.1K27me1 deposition is impaired. Overall, this work uncovers a role for the TPR domain of TSK in selectively interacting with the H3.1 variant. Previous work in human cell lines has shown that the TSK ortholog TONSL copurifies with H3.1 in affinity purification/biochemical

fractionation assays (*142*), and that TONSL-mediated double-stranded DNA break repair depends on the H3.1 chaperone CAF-1 (2). These findings, combined with our identification of the TPR domain of TSK/TONSL acting as an H3.1 reader, point to a model where post-replicative chromatin maturation in plants and animals relies on similar mechanisms involving H3.1 and clade-specific enzymes that monomethylate histones to prevent TSK/TONSL binding (Figure 2.4G). In plants, monomethylation occurs at H3.1 K27 via ATXR5/6 and prevents binding of TSK through the TPR domain. In animals, SET8-mediated monomethylation at H4 K20 interferes with TONSL binding via the ARD domain (*62*). However, in both plants and animals, recruitment of TSK/TONSL to chromatin likely relies on the ability of the conserved TPR domain to preferentially interact with the H3.1 variant. Thus, our work reveals the importance of selectively incorporating H3.1 variants during DNA replication, as it confers a window of opportunity during the cell cycle for the TSK/TONSL DNA repair pathway to resolve broken replication forks.



Complex	Dissociation const. (K_d) (µM)
TPR_{TSK} + H3.1	2.32 ± 0.06
TPR_{TSK} + H3.3	> 50
TPR_{TSK} + H3.1 K27me1	> 50
TPR_{TSK} + H3.1 K27me3	> 50

Figure 2.1 – The TPR domain of TSK specifically interacts with the N-terminal tail of the H3.1 variant. (A) Domain architecture of animal and plant TONSL/TSK. TPR, tetratricopeptide repeats; AS, acidic sequence; ARD, ankyrin repeat domain; UBL, ubiquitin-like; LRR, leucine-rich repeats. Conserved domains are shown in blue. **(B)** Pull-down assay using TPR^{TSK} and GST (glutathione S-transferase) tagged with the N-terminal tails of histones H2A.Z, H2A.X, H2B, H3.1, H3.3, and H4 from plants. **(C)** Representation of plant and mammalian H3.1/H3.2 (blue) and H3.3 (red) H3 variants. Thin lines and blocks represent the histone tails and cores, respectively; numbers indicate amino acid positions in H3. **(D)** Peptide pull-down assay using plant TPR^{TSK} and GST tagged with the tails of histones H3.1, H3.3, H3.1A31T, and H3.1F41Y. **(E)** Peptide pull-down assay using mouse TPR^{TONSL} and biotin-tagged histones H3.1 and H3.3 (full-length proteins)

from mammals. **(F)** Peptide pull-down assay using plant TPR^{TSK} and methylated peptides at Lys4, Lys9, Lys27, and Lys36 of H3.1 (amino acids 1 to 45). The red arrow indicates a gel lane that was removed. **(G)** Isothermal titration calorimetry assay using plant TPR^{TSK} and different H3 peptides. **[The experiments and figures corresponding to panels B,D,E were performed by the Jacob group and collaborators. Data and figures for panels F and G were produced by me.]**

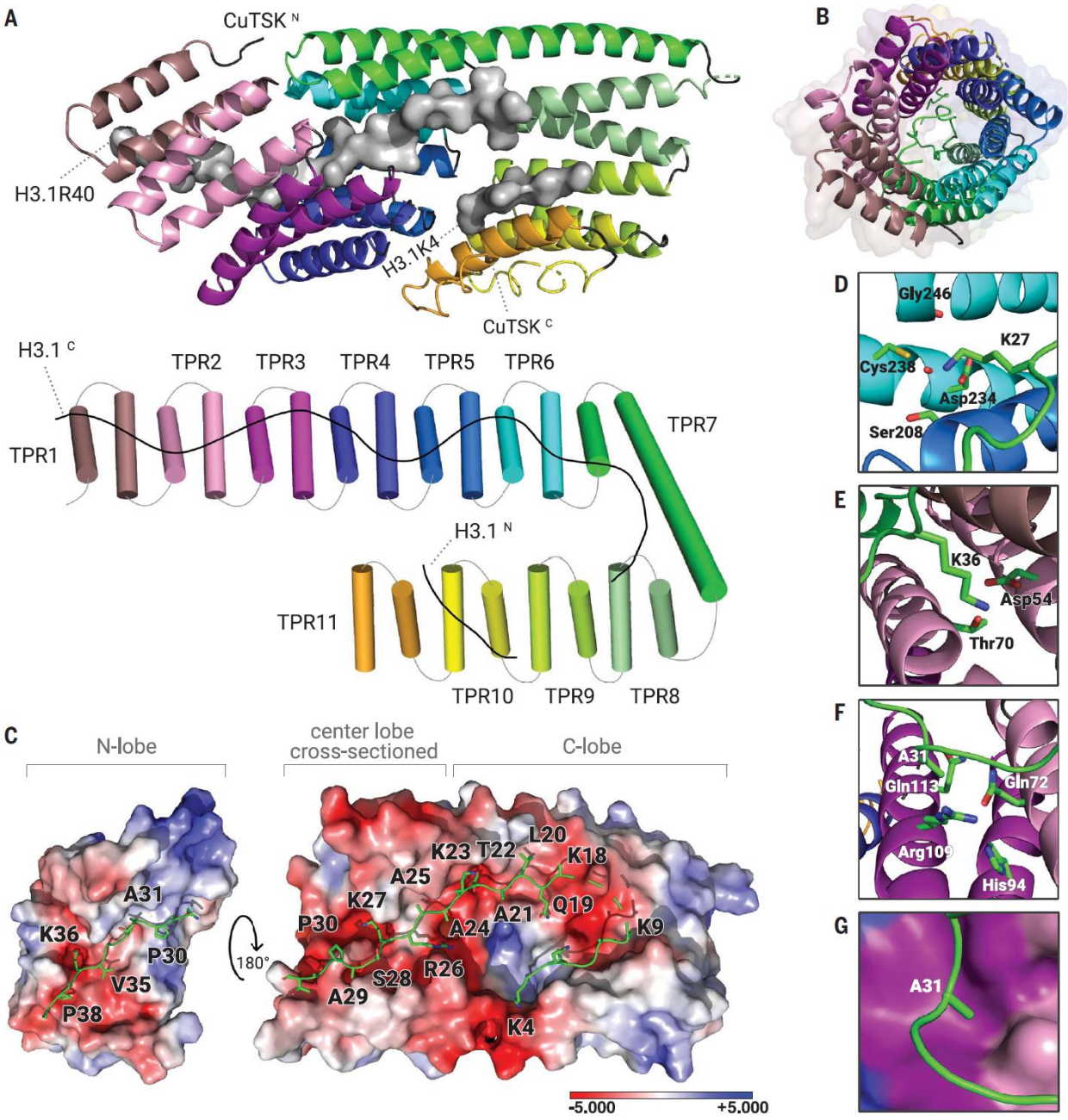


Figure 2.2 – Crystal structure of plant TPR^{TSK} bound to the H3.1 tail. (A) The TPR domain is depicted as a cartoon (top) or a cylinder (bottom) with individual TPR motifs as distinct colors. H3.1 is shown as surfaces (top) or line (bottom). **(B)** Channel view of the TPR solenoid tube showing the space inside the tube where H3.1 is extended (represented as a green line). **(C)** Surface representation of the TPR domain shown as electrostatic potential gradients contoured from +5.000 $k_BT\ e^{-1}$ (blue) to –5.000 $k_BT\ e^{-1}$ (red), where e is the electron, T is absolute temperature, and k_B is the Boltzmann constant. H3.1 is depicted as sticks. The N-terminal lobe (N-lobe) is rotated 180° along the horizontal axis relative to the center lobe and the C-terminal lobe (C-lobe). The surface of the center lobe is sectioned off to reveal the underlying segment of H3.1. **(D** to **F)** Amino acid residues from TPR^{TSK} (3-letter code) interacting with H3.1 residues (1-letter code) in their binding pockets are shown for Lys27 (D), Lys36 (E), and Ala31 (F). **(G)** Surface

representation of the H3.1A31 binding pocket. Surface colors correspond to those of TPR helices shown in (F). **[X-ray data collection at the synchrotron was facilitated by Dr. Brunzelle. Additionally, he provided an initial model. Construct engineering, protein purification, crystallization, preliminary data collection (screening) using in-house x-ray source, modelling, refinement, and the final structure were obtained by myself; additionally I generated this figure.]**

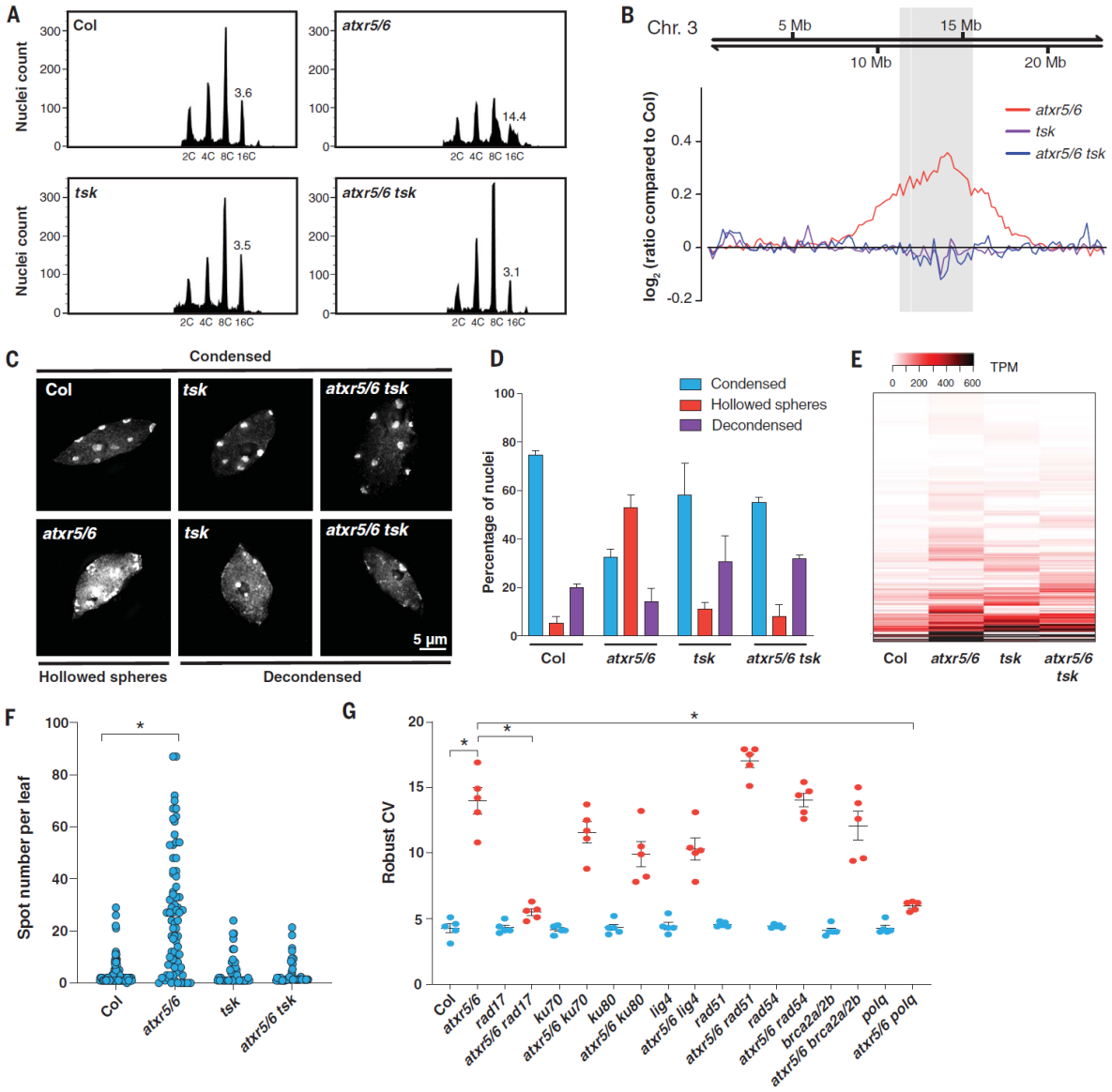


Figure 2.3 – Mutations in TSK suppress heterochromatin amplification of atxr5/6 mutants. **(A)** Flow cytometry profiles of Col, atxr5/6, tsk, and atxr5/6 tsk leaf nuclei. The numbers below the peaks indicate ploidy levels of the nuclei. The numbers above the 16C peaks indicate the robust coefficient of variation (rCV). **(B)** Chromosomal view (chromosome 3 of A. thaliana) of DNA sequencing reads from sorted 16C nuclei. The pericentromeric region is highlighted in gray. **(C)** Leaf interphase nuclei of Col, atxr5/6, tsk, and atxr5/6 tsk stained with DAPI. **(D)** Quantification of nuclei from the experiment shown in (C). Error bars indicate SEM. **(E)** Heatmap showing the relative expression levels of atxr5/6-induced transposable elements as measured by TPM (transcripts per million). **(F)** Average number of blue spots per leaf in Col, atxr5/6, tsk, and atxr5/6 tsk as determined using a GUS reporter for homologous recombination. *P < 0.0001 [Welch's analysis of variance (ANOVA) followed by Dunnett's T3 test]. **(G)** rCV values for 16C nuclei obtained by flow cytometry analyses. Each dot represents an independent biological replicate. Horizontal bars indicate means. Error bars represent SEM. *P < 0.05 (Welch's ANOVA followed by Dunnett's T3 test) **[These experiments and the figure were produced by the Jacob group and collaborators]**.

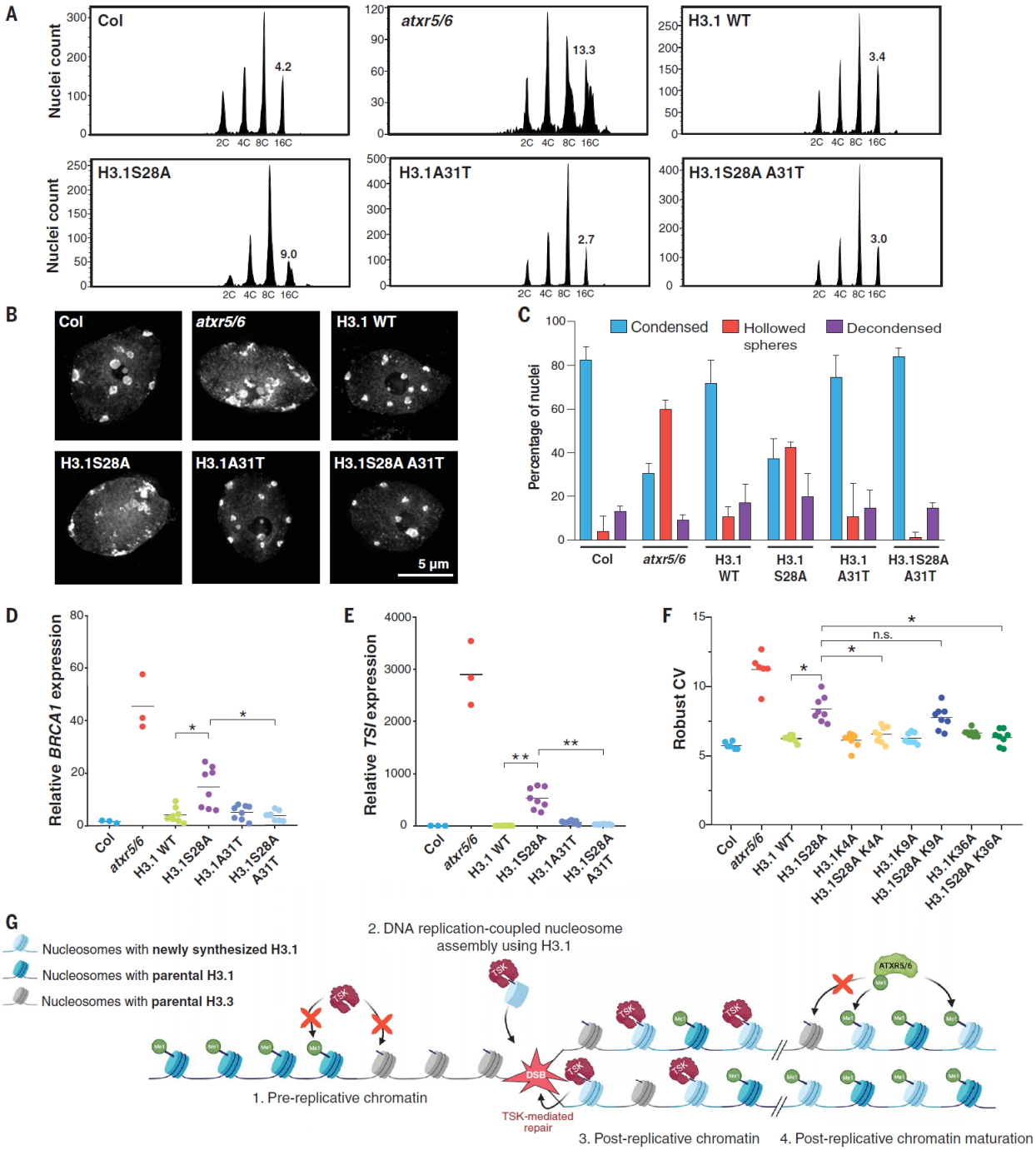


Figure 2.4 – H3.1 is required to mediate genomic instability in atxr5/6 mutants. (A) Flow cytometry of leaf nuclei. Numbers below the peaks indicate ploidy, and those above indicate rCV. **(B)** Leaf nuclei of Col, atxr5/6, and first-generation (T1) H3.1 lines stained with DAPI. **(C)** Quantification from nuclei in (B). Error bars indicate SEM. **(D** and **E)** Reverse transcription quantitative polymerase chain reaction of BRCA1 and TSI. Horizontal bars indicate means. *P < 0.05, **P < 0.001 (Welch's ANOVA followed by Dunnett's T3 test). **(F)** rCV for 16C nuclei obtained by flow cytometry. For Col and atxr5/6, each dot represents a biological replicate. For the H3.1 lines, each dot represents one T1 plant. Horizontal bars indicate means. *P < 0.05 (Welch's ANOVA followed by Dunnett's T3 test); n.s., not significantly different. **(G)** Model

depicting the interplay among H3.1, TSK, and ATXR5/6 during replication. Step 1: TSK cannot interact with chromatin containing H3.3K27me0 or H3.1K27me1. Step 2: Newly synthesized copies of H3.1 (H3.1K27me0) in complex with TSK are inserted at replication forks. Step 3: DSBs caused by broken replication forks are repaired by TSK. Step 4: Monomethylation of newly inserted H3.1 (but not H3.3) at Lys27 by ATXR5/6 prevents binding of TSK. **[These experiments and the figure were produced by the Jacob group and collaborators]**.

2.6 Supplementary Materials

The supplementary Materials and Methods, Figures (figS1-S12), Tables (TableS1-S13), and references corresponding to these materials are in the Appendix.

2.7 Acknowledgments

We thank X. Dong (Duke University) for providing us the seeds for the brca2a mutant (13F-1) and the recombination reporter line Col 1445; C. Bolick, E. Williams, and N. Guzzo (Yale) for help with plant growth and maintenance; K. Nelson for technical help with flow cytometry; members of the Yale Center for Genome Analysis; the Edinburgh Protein Production Facility (EPPF) for their support; and E. Campos (University of Toronto), Yale scientists W. Liu, J. Nijjer, and W. Yuan, and members of the Jacob and Couture labs for discussions, advice, and/or materials that contributed to this work.

2.7.1 Author Contributions

Y.J. and J.-F.C. designed the research and supervised the study. Y.J., Y.-C.H., and J.D. designed and performed the initial experiments that identified the TPR domain of TSK/TONSL as an H3.1 reader. H.D. confirmed the H3.1-TSK interaction after doing the ITC assays, solving the crystal structure, and performing the in vitro structure-activity relationship study. Y.-C.H. performed the HR assays, and the analyses of the tsk and other DNA repair mutants, the H3.1S28A lines, and the H3.1 CRISPR mutants. J.D., H.D., and Y.-C.H. performed the in vitro binding assays, and A.A.-V. and J.D. contributed to the ITC assays. B.M. generated DNA repair mutants and contributed to their analysis. C.L. generated the H3.1S28A lines, and contributed with G.V. to in vivo experiments using these lines. A.P. performed the RNA-seq and DNA-seq analyses. G.T., V.J., and Y.-C.H. generated and validated the CRISPR mutants. M.M., A.R., B.H.R., and E.I.A. generated histone peptides. D.V. and P.V. designed and performed nucleosome pulldown experiments. J.S.B.

collected structural data and generated a preliminary model. Y.J. and J.-F.C. wrote the manuscript with contributions from C.L., J.D., H.D., Y.-C.H., and P.V.

2.7.2 Funding

Supported by NIH grant R35GM128661 (Y.J.), grants from the Natural Science Engineering Research Council and the Canadian Institutes of Health Research (PJG-47391S) (J.-F.C.), an Ontario Graduate Scholarship and a University of Ottawa Excellence Scholarship (H.D.), a Yale University Brown Fellowship (B.M.), Fonds de Recherche du Québec–Nature et Technologies (FRQNT) grant 272565 (V.J.), and a Strategic Research Postdoctoral Fellowship from the University of Ottawa Heart Institute and the Strategic Research Endowed Funds (M.M.). Work in the Voigt lab was supported by Wellcome Trust grant 104175/Z/14/Z Sir Henry Dale Fellowship (P.V.) and UK Biotechnology and Biological Sciences Research Council grant BBS/E/B/000C0421. The Wellcome Centre for Cell Biology received core funding from Wellcome Trust grant 203149. The EPPF was supported by the Wellcome Trust through Multi-User Equipment grant 101527/Z/13/Z. B.H.R. and E.I.A. thank NSERCRTI (RTI-2019-00009).

2.7.3 Competing interests

The authors declare that they have no competing interests.

2.7.4 Data and materials availability

Sequencing data (DNA-seq and RNA-seq datasets) generated for this study are available at the Gene Expression Omnibus (GEO) under accession code GSE184738. The Protein Data Bank accession number for the TPR^{TSK}-$H3.1^{(1-45)}$ structure is 7T7T. All data are available in the main text or the supplementary materials.

Chapter 3 | Structural and Biochemical Analysis of histone H3.1 recognition mechanisms by TONSOKU

Chapter 3 is a manuscript submitted on 4 July 2023 to Elsevier (Manuscript Number: JMB-D-23-00539) for consideration in the Journal of Molecular Biology.

3.1 Authors and Affiliations

Hossein Davarinejad[1], Marcelo Muñoz[2], Monika Joshi[1], Sabina Sarvan[1], Alex Ross[2], Joseph S. Brunzelle[3], Emilio I. Alarcon[2], Jean-François Couture[1*]

1 Ottawa Institute of Systems Biology, Department of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, Ontario K1H 8M5, Canada
2 BEaTS Research Laboratory, Division of Cardiac Surgery, University of Ottawa Heart Institute, Ottawa, Ontario K1Y 4W7, Canada
3 Department of Molecular Pharmacology and Biological Chemistry, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA.

* Corresponding author: jean-francois.couture@uottawa.ca (J.-F.C.)

3.2 My Contributions

For this manuscript, I performed all of the ITC assays for TPR^{TSK} with H3.1 PTMs. I also performed the majority of the ITC assays for TPR^{TSK} mutants with H3.1. I solved the crystal structure of TPR^{TSK} bound to histone H3.1K27ac. I performed the EMSA assays and the ITC assays for SD mutants. I purified the protein for all experiments performed by myself and generated all figures for the manuscript. Lastly, I wrote the manuscript.

3.3 Manuscript Submission

From: em.jmb.0.84752f.13b98bf4@editorialmanager.com on behalf of JMB (ELS) <em@editorialmanager.com>
Date: Tuesday, July 4, 2023 at 4:41 PM
To: Jean-François Couture <jean-francois.couture@uottawa.ca>
Subject: JMB-D-23-00539 - Confirming your submission to Journal of Molecular Biology
Journal: Journal of Molecular Biology
Title: Structural and Biochemical Analysis of Histone H3.1 recognition mechanisms by TONSOKU
Corresponding Author: Dr. Jean-Francois Couture
Co-Authors: Hossein Davarinejad; Marcelo Munoz; Monika Joshi; Sabina Sarvan; Alex Ross; Joseph S Brunzelle; Emilio I Alarcon
Manuscript Number: JMB-D-23-00539

Dear Dr. Couture,

Your submission entitled "Structural and Biochemical Analysis of Histone H3.1 recognition mechanisms by TONSOKU" has been received by journal Journal of Molecular Biology. It has been assigned the following manuscript number: JMB-D-23-00539.

You will be able to check on the progress of your paper by logging on to the Editorial Manager as an Author. The URL is https://www.editorialmanager.com/jmb/.

EDITORIAL NOTE: Please note that during summertime there may be a delay in processing your manuscript. It can be challenging during this time, as reviewers and editors generally take a well-deserved break. We apologize in advance if our editorial assessment of your paper takes longer than normal to complete. Our editorial office may be reached at jmb@elsevier.com. We are doing our best to give you a timely decision.

Thank you for submitting your work to this journal.

Kind regards,

Journal of Molecular Biology

This journal uses the Elsevier Article Transfer Service. This means that if an editor feels your manuscript is more suitable for an alternative journal, then you might be asked to consider transferring the manuscript to such a journal. The recommendation might be provided by a Journal Editor, a dedicated Scientific Managing Editor, a tool assisted recommendation, or a combination. For more details see the journal guide for authors.

3.4 Abstract

TONSOKU (TSK) recognizes histone H3.1 and regulates heterochromatic DNA replication and genome stability. We have previously shown how the tetratricopeptide repeat (poly-TPR) of TSK discriminates between histone H3 variants and preferentially binds H3.1. In this study, we show that TPR^{TSK} employs several binding pockets to bind the N-terminus of histone H3.1 either in solution or when embedded in the nucleosome. Consistent with the ability of TSK to repair DNA damages during replication, several post-translational modifications (PTMs) of histone H3.1 deposited during mitosis and linked to chromatin condensation negatively impact TPR^{TSK} binding to histone H3.1. Conversely, H3.1 PTMs enriched in gene regions undergoing active DNA transactions are neutral to TPR^{TSK} binding. Finally, sponastrime dysplasia mutations mapping to the tetratricopeptide repeat of TSK-LIKE (TONSL) lose binding to histone H3.1. Overall, our findings reveal a large network of molecular interactions between histone H3.1 and the poly-TPR domain of TSK, altogether establishing this domain as a multi-PTM sensor involved in DNA damage repair.

3.5 Introduction

Epigenetic regulatory mechanisms take many forms and are essential to many nuclear transactions, including gene expression, DNA damage repair, and replication (*143*). At the centre of these processes is the formation of a nucleoprotein complex composed of DNA and histone proteins, referred to as the nucleosome. Nucleosomes are disk-like structures composed of two copies of the canonical histones H2A, H2B, H3, and H4 and a DNA fragment of approximately 150bp. Dynamic in nature, the nucleosome constantly undergoes rounds of displacement and structural reorganization. Moreover, the nucleosome structure can be altered by the replacement of a canonical histone with a histone variant and by post-transcriptional modifications (PTMs) on the histones (*143*).

Another key component that contributes to imparting epigenetic events is the participation of proteins that recognize a given histone with a specific post-translational modification or lack thereof. For example, bromodomain-containing proteins SWI/SNF bind acetylated lysine residues to contribute to chromatin remodelling in promoter regions (*144, 145*). 14-3-3 proteins and BIR domains generally recognize phosphorylated serine and threonine residues; for example, phosphorylated S10 on H3 is recognized by 14-3-3, and this interaction promotes transcription elongation (*146*). Moreover, the BIR domain of Survivin reads the first four N-terminal residues of H3 when T3 is phosphorylated, which is important for the progression of mitosis (*147*). Plant homeodomain (PHD), chromatin organization modifier (CM), malignant brain tumour (MBT), and Tudor domains generally bind methyl-lysine peptide and are linked to a myriad of nuclear events (*148*). More recently, YEATS domain (named after the enzymes harboring this domain; Yaf9, ENL, AF9, Taf14, and Sas5) in AF9 has been associated with controlling transcription rates by associating with crotonylated lysine residues in active promoter or enhancer regions (*149*).

Interestingly, some reader domains prefer binding the N-terminal region of histone proteins in their unmodified form. For example, the PHD domains of ATXR5/6 (*150*) and UHRF1 (*151*) preferentially bind the first ten residues of histone H3 in their unmodified forms. Similarly, the tetratricopeptide (TPR) domain of TONSOKU (TSK) binds the N-terminal tail of histone H3.1 and distinguishes it from H3.3 (*87*).

Initially identified in a plant mutant giving rise to roots with the shape of pig feet (Tonsoku in Japanese) (*124*), TONSOKU -LIKE (TSK-LIKE – also referred to as TONSL) was identified in RNAi screens as a protein involved in HR-mediated DDR for stalled or collapsed replication forks in humans (*65, 102, 104, 152*). Expectedly, TONSL co-localizes with DNA damage markers at sites of DNA damage post-UV exposure. Interestingly, TONSL interacts with members of chromatin remodelling complexes facilitating transcription elongation (*153*) and is found in the S phase of the cell cycle, suggesting that it responds to errors in DNA processing during replication (*154*). Thus, in the presence of the DNA damage agent camptothecin, TONSL rapidly accumulates to the nucleus in punctate foci where DDR takes place (*66*). TONSL helps the recruitment of Rad51 and promotes the displacement of histones from nucleosomes (*154*). In mammals, the Ankyrin Repeat Domain (ARD) of TONSL binds the N-terminal tail of histone H4. Methylation on histone H4 K20 (H4K20) inhibits TONSL accumulation to damaged sites and homologous recombination (*62*). Similarly, the TPR domain of TSK selectively binds the N-terminal region of replication-coupled histone H3.1, and its methylation on K27 blocks the binding of TSK (*87*).

Recent structural studies showed that the TPR domain of TSK (TPR^{TSK}) is formed by 11 consecutive TPR motifs that collectively fold as a solenoid tube. Two segments (R2 to T11 and K18 to R40) of histone H3.1 bind in the hollow center of the tube, where it engages in several interactions with TPR^{TSK}. TPR^{TSK} uses multiple binding pockets to stabilize histone H3.1, yet their

contributions to histone H3.1 stabilization remain elusive. Moreover, given that other readers are heavily impacted by PTMs deposited on histone H3.1 (*155*), we hypothesize that PTMs might impact TPR^{TSK} binding to histone H3.1. Using a structure-based mutational approach, we show that several pockets contribute to TPR^{TSK} binding to histone H3.1. Several TONSL mutations identified in patients with sponastrime dysplasia, an autosomal recessive disease causing spine and facio-cranial deformities (*38, 39*), map to amino acids adjacent or within the center of the peptide-binding tube of the TPR domain and correspond to loss of histone H3.1 binding. Moreover, PTMs on residues found in the narrowest region of TSK in the center of the solenoid tube have the most detrimental impact on its binding to histone H3.1. The crystal structure of H3.1K27ac bound to TPR shows how this modification induces local conformational changes creating a new binding pocket to accommodate this PTM. These observations shed some light on PTMs that will maintain a chromatin template that is conducive to HR-mediated DNA damage repair.

3.6 Results and Discussion

3.6.1 A systemic approach to measure the binding footprint of TPR^{TSK} to histone H3.1

We employed two complementary approaches to gain further insights into the mechanisms underlying the interactions between TPR^{TSK} and H3.1. First, we mutated many TPR^{TSK} residues found near histone H3.1 and measured their binding affinities for a peptide corresponding to the first 45 residues of histone H3.1. In parallel, we used a library of +40 peptides, see Table 1, bearing known histone H3.1 post-translational modifications and measured their binding affinity to TPR^{TSK}.

3.6.2 Organisation of TPR^{TSK} poly-TPR domain

The poly-TPR domain of TSK comprises 11-tandem TPR motifs collectively forming a solenoid that folds into a tube-like structure around the N-terminal region of histone H3.1 (Figure 3.1A) (*87*). The tube can be organized in three topologically distinct regions, consisting of a narrow center-lobe formed by the C-terminal α-helix of TPR motif 3, TPR 4-6, and the N-terminal region of TPR7 (Figure 3.2A). On each side of the center-lobe, the N- and C-terminal lobes form pseudo-tubes mimicking laterally cross-sectioned pipes that expose histone H3.1 peptide to solvent. The N-terminal lobe comprises TPR motifs 1 (TPR1), TPR2, and the N-terminal α-helix of TPR3 and forms a narrow channel for binding histone H3.1 (Figure 3.1A). Wider than the N-terminal lobe, the C-terminal lobe comprises the C-terminal portion of TPR7 and TPRs 8-11 and binds two regions of histone H3.1 simultaneously (Figure 3.3A).

3.6.3 Histone H3.1 residues 31-40 bind the N-terminal Lobe (N-lobe) of TPR^{TSK}

Histone H3.1 residues 31-40 bind in a solvent-exposed concave surface of the TPR^{TSK} N-lobe and the residues near the C-terminal region of the histone H3.1 peptide are located near the first TPR motif (Figure 3.1A and 1B) [henceforth single-letter codes refer to H3.1 residues and three-letter

codes denote TPR^{TSK} amino acids]. In this region, Glu26 carboxylate group is found near H39 backbone amide. In addition, a cluster of residues that collectively use the aliphatic portion of their side chains (Arg29, Trp30, and Val33) surround P38 (Figure 3.1C). Intriguingly, placing a polar group in this region (Val33Thr) does not impact H3.1 binding (Figure 3.1D). This observation is consistent with the geometry of the pocket as all TPR^{TSK} residues point away from P38. In the same region, the guanidinium group of Arg29 is in proximity to K37 backbone carbonyl (Figure 3.1C). Surprisingly, the substitution of Arg29 for a glutamate residue increases H3.1 binding by 5-fold (Figure 3.1D). A closer inspection of the P38/K37 region suggests that the newly introduced carboxylate group could contact either P38 or K37 backbone amide groups or contribute to the adjacent K36 binding pocket (Figure 3.1C). Consistent with the lack of electron density beyond K37 Cβ (Figure 3.1C), methylation of its e-amine does not change H3.1 binding (Figure 3.1E).

Previous structural studies showed that residues T45 to P38 of histone H3.1 locate between two DNA gyres of the nucleosome core particle (NCP) (*2*). Specifically, T45 and G44 contribute to the first turn of histone H3.1 α-helix 1 while P43 is adjacent to one of the DNA strands. Furthermore, R42, Y41 (F41 in plants), R40, and H39 are sandwiched between two DNA gyres near the entry/exit points while P38 is the first solvent-exposed residue. Consistent with this model, asymmetric di-methylation of R42 does not impact TPR^{TSK} binding (Figure 3.1E) but it destabilizes the NCP by loosening the interaction between histone H3 and DNA, a modification that correlates with increased transcriptional activity (*156*). Interestingly, the structure of TPR^{TSK} bound to histone H3.1 shows that P38 is the last residue for which discernible electronic density can be observed (Figure 3.1C). The potentially non-overlapping footprint of TPR and the nucleosome disc on histone H3.1 suggests that TPR^{TSK} can recognize the N-terminus of histone H3.1 even when embedded within the NCP. Notably, TPR^{TSK} readily binds histone H3.1 when

embedded in octamers, free in solution, or its H3.1 tail peptide alone, using pull-down assays (*87*). To evaluate whether TPR^{TSK} binding to histone H3.1 induces destabilization of the nucleosome, namely causing the DNA to dissociate, we performed EMSA assays and tracked the migration pattern of the nucleosome on native gels. Our data shows that increasing concentrations of TPR^{TSK} twice caused molecular weight shifts to the NCP specie representing nucleosomes that are bound to one or two TPR^{TSK} molecules (Figure 3.1F). Consistent with our proposed TPR^{TSK}: NCP complex model, these shifts do not lead to an accumulation of free DNA suggesting that binding of TSK to NCP via TPR does not impact the structural integrity of the nucleosome (Figure 3.1F).

In the same lobe, the hydroxyl groups of Thr70 and Ser58, the carboxylate, and the backbone carbonyl of Asp54 form the histone H3.1 K36 binding pocket (Figure 3.1C). Mutation of Thr70 and Asp54 (Thr70Leu and Asp54Val) lead to a 21-fold loss in H3.1 binding (Figure 3.1D). Accordingly, increasing the K36 methylation states leads to a progressive loss of H3.1 binding (Figure 3.1E). Similarly, introducing an acetyl group on K36 leads to a 15-fold loss of binding (Figure 3.1E). Histone H3 K36me3 and SETD2 are required for RPA and RAD51 foci formation (*157*). Mechanistically, histone H3 K36me3 is recognized by the PWWP domain of lens epithelium–derived growth factor (LEDGF/p75) and its accumulation recruits C-terminal binding protein interacting protein (CtIP) to promote dsDNA resection at break sites (*158*). The antagonistic impact of K36me3 on TSK binding suggests that the DNA damage repair mechanisms linked to TONSL and histone H3K36 methyltransferases differ, occur at different time-window during cell cycle progression, or that this PTM temporarily blocks TSK binding to prevent spurious HR.

In contrast to K36, G34 does not contribute to TPR^{TSK} binding, while V35 is found near Asp36 side chain (data not shown). The next residues, G33 and T32, are found near TPR^{TSK}. The

backbone amide of G33 is near the carbonyl moiety of the Gln72 side chain (Figure 3.1C). Nearby, the backbone carbonyl of T32 is within range to make a hydrogen bond with Arg109 guanidinium group. Accordingly, the substitution of Gln72 or Arg109 leads to an important loss of binding (>21-fold) to histone H3.1 (Figure 3.1D). As previously reported, Arg109 and Gln72 cooperate with Gln113 (Figure 3.1C) to form the pocket contributing to the selectivity of TPR^{TSK} for the histone H3.1 variant (*87*). The three residues form a tight enclosure around A31 that is not permissive to larger side chains such as T31 of histone H3.3. Finally, T32 hydroxyl group is in the proximity of Asp163 and Asp166 side chains and the backbone carbonyl of Lys162, which form a relatively narrow pocket surrounding T32 (Figure 3.1C). Accordingly, phosphorylation of this residue leads to an 8.5-fold decrease in TPR^{TSK} binding (Figure 3.1D). These observations are consistent with studies showing that T32ph levels peak during mitosis (*159*), a phase of the cell cycle during which TSK is presumed not to be expressed.

3.6.4 The Center-Lobe of TPR^{TSK} encircles residues 24-30 of histone H3.1

The Center-Lobe of TPR^{TSK} forms a tight channel coordinating the binding of residues 24-30 of histone H3.1 (Figure 3.2A). Within this channel, P30 binds a surface formed by Arg116 guanidinium group, and A29 side chain binds in a shallow groove formed by the aliphatic portions of Asp166 and Asn169 (Figure 3.2C). A29 amide group also makes a hydrogen bond with the Asn170 side chain carbonyl's group. Accordingly, Asn170Ala and Asp166Arg negatively impact binding to histone H3.1 (Figure 3.2D). Close to the mid-section of the channel, S28 lies atop a long groove (Figure 3.2B and C) formed by residues Arg116 and Glu120 where its hydroxyl group makes a hydrogen bond with the ε-amine of Arg116. Phosphorylation of S28 by mitogen-activated pathway kinases (MAPK) MSK1/2 are linked to rapid but transient expression of repressed genes during mitosis (*160, 161*) and shows significant enrichment in promoters of half of all stress-

related genes (*162*). Modelling a phosphorylated S28 in this pocket introduces multiple important clashes with both Arg116 and Glu120, and even though TSK and S28ph may not overlap temporarily, regions marked with this PTM would be protected from HR-mediated DNA damage repair by the TONSL-MMS22L complex.

Near S28, the R26 side chain lays over the surface of TPR5, making hydrophobic contacts with Met176 (Figure 3.2C) via their aliphatic portions. Accordingly, substituting this residue for an arginine lowers H3.1 binding by 6-fold (Figure 3.2D). In this orientation, the R26 guanidinium side chain can also make hydrogen bonds with the carboxylate group of Glu177 and the carbonyl moiety of the Asn216 side chain. These interactions are important as Glu177His and Asn216His mutants reduce binding by >20- and 2-fold, respectively (Figure 3.2D). Consistently, post-translational modification of R26 impairs TPR^{TSK} binding. Although this loss appears negligible when it is mono- and asymmetrically di-methylated, citrullination of the same residue leads to an 11-fold loss of binding (Figure 3.2E). It is interesting to note that while R26me1/me2 binding affinities are similar to unmodified peptide; the binding mechanism appears dissimilar as a loss in enthalpy is compensated by a gain in entropy (Figure 3.2E). Methylation of R26 is catalyzed by coactivator-associated arginine methyltransferase 1 (CARM1) and protein arginine methyltransferase 4 (PRMT4), while peptidyl arginine deiminase 4 converts the R26 guanidinium group into citrulline. PAD4 is the main deaminating enzyme in humans that localizes in the nucleus and targets histones, including H3 (*163, 164*). Interestingly, R26me2a inhibits citrullination of R26 (*164*), suggesting that R26 methylation may sustain TPR^{TSK} binding in protecting the residue from PAD4 activity.

In neutrophils, hyper-citrullination of R26 is correlated with widespread decondensation of chromatin and is thought to contribute to the formation of neutrophil extracellular traps (NETs)

made from the loose chromatin fibre and other proteins and cytotoxic reagents which are extruded from the cells during NETosis (a form of triggered cell death) as a defence mechanism (*165, 166*). Our binding assays show that R26ci reduces H3.1 binding by 11-fold (Figure 3.2E) and in agreement with the biological role of R26ci the large reduction in TPR^{TSK} binding suggest a mechanism of avoiding recruitment of DNA-repair proteins in cells destined to cell death.

3.6.5 The K27 sensing pocket of TPR^{TSK}

In plants, ATXR5/6 mono-methyltransferases deposit K27me1 behind the replication fork and prevent reduplication of DNA via ectopic TSK-initiated homologous recombination (*87*). CLF, SWN, and MEA are plant homologs of EZH2 and deposit K27me2/me3, which have repressive roles for transcription and are important for silencing transposable elements (*167, 168*). The binding pocket of K27 is in the TPR^{TSK} Center-lobe (Figure 3.2A, B, and C) and is formed by the side chains of Cys238, Asp234, Ser208 and Gly246 backbone carbonyl. The thiol group of Cys238 and the carboxylate moiety of Asp234 are within hydrogen bonding distances of K27 ε-amine. Consistently, the substitution of Ser208, Cys238 and Asp234 by asparagine, arginine, and lysine residues, respectively, severely impair TPR^{TSK} binding to H3.1 (Figure 3.2D). Because the K27 binding pocket is relatively narrow, we surmised that PTMs deposited on H3.1K27 would prevent the binding of TPR^{TSK}. Accordingly, the mono-, di- and tri-methylation of K27, PTMs predominantly deposited by PRC2 complexes and ATXR5/6 (*169*), negatively impact TPR^{TSK} binding to histone H3.1 (Figure 3.2E). Surprisingly, we observed that acetylation and crotonylation of the same residue do not impact TPR^{TSK} binding (Figure 3.2E).

To understand the structural underpinnings controlling TPR^{TSK} ability to accommodate H3.1K27ac binding, we co-crystalized TPR^{TSK} in complex with an H3.1 K27ac peptide and solved its structure by molecular replacement (**Figure *3.2*F**). TPR^{TSK}/H3.1K27ac structure aligns well

(r.m.s.d = 0.326Å^2) with TPR^{TSK}/H3.1 suggesting that K27 acetylation does not trigger large structural reorganization of the poly-TPR domain. The K27ac side chain extends longitudinally in parallel with one of the TPR5 α-helix, with the acetyl moiety pointing toward the TPR. In this orientation, the carbonyl group of the acetyl moiety makes hydrogen bonds with the imidazole and guanidinium groups of His243 and Arg209, respectively. The orientation of K27ac contrasts with the position of K27 as the latter point directly toward TPR6 a-helices. Reorientation of His243 and Arg209 side chains enables the opening of that region of TPR^{TSK}, which likely accommodates K27ac binding (Figure 3.2F).

Acetylation of K27 is mainly catalyzed by the histone acetyltransferase (HAT) enzymes CBP/p300 (*170, 171*) and, to a lesser extent, by GCN5/PCAF (*172*). This PTM is in the active chromatin region and is promoted by acetyl-CoA availability in the cytoplasmic environment, linking this epigenetic mark with sensing the metabolic environment (*173*). GCN5 also catalyzes crotonylation of the same residue (K27cr), and both K27ac and K27cr are enriched in enhancers and transcription start sites and are generally linked to active transcription (*174*). Akin to its propensity to be modified post-transcriptionally, K27 is an important platform for recruiting histone readers. While BRDs are the main reader of the acetylated form of K27, YEATS domains, found in transcription-related proteins, can read both K27ac and K27cr with more affinity for K27cr (*89*). Our data show that TPR^{TSK} binding is largely inhibited by PTMs typically linked to transcriptional repression (K27me1/2/3). However, the poly-TPR-containing protein is fully compatible with PTMs found in regions typically undergoing active nuclear transactions (ex. K27ac and K27cr) (Figure 3.2E). These observations suggest that TPR^{TSK}, in addition to binding nucleosomes with newly synthesized H3.1, can bind nucleosomes reconstituted with recycled parental H3-H4 tetramers with these pre-existing marks.

3.6.6 The wide C-terminal Lobe of TPR^{TSK} binds two regions of histone H3.1

The C-terminal lobe coordinates the binding of two regions of histone H3.1, including K23-K18 and A1-S10 (Figure 3.3A and B). These two segments are separated by a region of histone H3.1 for which no discernable electronic density could be observed and is highly likely unstructured (T11-R17). In the first segment, histone H3.1 K23 is located at the junction between TPR^{TSK} Center and C-terminal lobes in a pocket predominantly formed by residues located in TPR7. In this region, Gln293, Gln290, and Asn294 cluster the aliphatic portion of their side chains to maintain K23 (Figure 3.3C). The side chain of Asn294 also contacts the backbone of K23, and its ε-amine is found near the hydroxyl group of Thr297 and the carbonyl of the Gln293 side chain. Unexpectedly, Gln293Lys binds histone H3.1 with similar binding affinities as for the wild-type and substitution of Gln293 by phenylalanine results in a 5-fold gain of binding compared to wild-type TPR^{TSK} (Figure 3.3D). This observation is likely due to the burying of the phenylalanine side chain between the aliphatic portions of TPR^{TSK} Gln290 and H3.1 K23. Additionally, methylation, acetylation, and crotonylation of K23 result in a negligible impact on histone H3.1 binding by TPR^{TSK} (Figure 3.3E).

The binding pocket of T22 is located between TPR6 and TPR7 and formed by Asn94 and Tyr259 side chains and buttressed by Val298 (Figure 3.3C). Histone H3.1 T22 hydroxyl group is also near Glu256 and Tyr271 carboxylate and hydroxyl groups, respectively. In agreement with the spatial restrictiveness and shielding of the T22 sidechain from solvent, T22 phosphorylation prevents TPR^{TSK} binding to histone H3.1 (Figure 3.3E). This modification (T22ph) is among several other histone PTMs identified as a marker of mitosis (*175*) during which HR is suppressed as activation of this DDR pathway poses a risk of telomer fusion for cells undergoing mitosis (*176*).

Unlike A21, L20 binds in a well-defined hydrophobic cleft formed by the side chain of Val304 and the aliphatic potions of Lys300, Thr297, and Mse351 (Figure 3.3C). Consistent with the crystal structure, replacing Met351 with a histidine leads to a 5-fold loss in binding affinity (Figure 3.3D). The side chain of Q19 points towards the TPR surface above the joining loop of TPR8 α-helices and is surrounded by the aliphatic portions of Phe353 and Trp355. Additionally, the side chain amide of Q19 contacts the backbone carbonyl of Ser350 (Figure 3.3C). Consistent with the lack of electron density for the K18 side chain, its methylation (me1/me2/me3) or acetylation minimally impacts histone H3.1 binding (Figure 3.3E).

Similar to K18, no electron density is observed for residues corresponding to G12-R17, suggesting that this region of histone H3.1 does not interact with TPR^{TSK}. Accordingly, binding studies with peptides harboring PTMs in that region do not result in noticeable differences in binding affinity compared to wild-type histone H3.1 (Figure 3.3E). Interestingly, H3.1K14 methylation is associated with DNA-damage responses via recruitment of the ATR (ataxia telangiectasia) kinase and Rad3, which interact with RPA proteins and are checkpoint activators (*177*), suggesting that the binding footprint between readers differ between the type of DNA damage responses.

Following the unstructured region of histone H3.1, A7 is positioned above a cleft where its amide and carbonyl groups make hydrogen bonds with Asn424 side chain (Figure 3.3C). Three helices from TPR10 and TPR11 form a channel beneath H3.1 where Lys391 and Glu387 form a wall stabilizing T6 (Figure 3.3C). In the same pocket, Lys391 ε-amine contacts T6 hydroxyl group while T6 Cγ interacts with the aliphatic portion of Glu387. Intriguingly, phosphorylation of T6 does not lead to a major loss of binding (Figure 3.3E). This may be explained by the fact that this region is exposed to solvent. Notably, this channel opens to a large pocket immediately next to

Lys391 (Figure 3.3C). Residues Asn464, Glu463, and Ser460 form a three-finger-like projection on the surface of TPR11 where histone H3.1 Q5 is positioned (Figure 3.3C). In this cleft, the side chain of Asn464 makes a hydrogen bond with the Q5 backbone amide group. The side chain of Q5 extends towards the space between Ser460 and Glu463, with its amide group close to the hydroxyl of Ser460. Near the loop that joins the α-helices of TPR10, a deep pocket is formed by multiple TPR residues to accommodate the K4 side chain (Figure 3.3C). In this pocket, the aliphatic portion of K4 side chain is sandwiched between the aliphatic portions of Asn427 and Tyr467. At the bottom of this pocket, K4 ε-amine is near Tyr467 hydroxyl group and the side chains of Asp430 and Asp179. Asp430His or Asp179Asn mutations lead to a 5- or 3-fold loss in H3.1 binding (Figure 3.3D), suggesting that hydrogen bonding between Asp430 or Asp179 and H3.1 K4 contributes to the stability of the peptide. Interestingly, while di- and tri-methylation of K4 negatively impacts H3.1 binding, its monomethylation enhances TPR^{TSK} binding (Figure 3.3E).

H3K4 methylation is a PTM predominantly deposited by the complex of proteins associated with Set1 (COMPASS). Among COMPASS members, MLL3/4 predominantly deposits H3K4me1 at enhancer regions. Interestingly, knockdown of the histone H3K4 mono-methyltransferase Mixed Lineage Leukemia (MLL3) (*178*) results in a loss of recruitment of proteins linked to DDR, such as RAD51 and p53BP1. Deposited by p300/CBP (*179*) and antagonistic to the repressive H3K27me3 mark, H3K27ac is also enriched at active enhancers (*180*). Given that TPR^{TSK} preferentially binds to H3K4me1 (Figure 3.3E) and can accommodate K27ac (Figure 3.3F), it is tempting to speculate that MLL3/4 complexes and TONSL coordinate their actions to contribute to DDR.

The sidechain of histone H3.1 T3 faces away from the surface of the TPR^{TSK} and points towards the open cavity at the centre of the C-lobe where it is exposed to solvent, suggesting that this residue can tolerate PTMs. Intriguingly, T3ph shows a 4-fold decrease in TPR^{TSK} binding (Figure 3.3E). Coincidentally, the first two residues of the histone introduce a short turn in this region bringing the backbones of A1 and R2 near T3 sidechain (Figure 3.3C). Such orientation likely contributes to the instability of the peptide when T3 is phosphorylated. Catalyzed by the kinase Haspin, and similar to T22ph, T3ph is deposited during mitosis and is important for the recruitment of the chromosome passenger complex and the interactions between the kinetochore and the mitotic spindle (*181*).

3.6.7 Sponastrime dysplasia mutations in TONSL affect H3.1-binding.

We used the structure of TPR^{TSK} as a basis to investigate mechanisms contributing to human disease associated with TSK. Sequence alignment of TPR^{TONSL} and structural mapping revealed that several mutations (Q41H, Q173N, and D363H) observed in patients suffering from sponastrime dysplasia impact residues located in the inner part of the TPR^{TSK} (Figure 3.4) solenoid and more precisely cluster in and adjacent to the center-lobe of the TPR^{TSK}. Binding assays show that these mutations negatively impact TPR^{TONSL} binding to histone H3.1 (Figure 3.4). Spontaneous chromosome breaks are observed in TONSL variants causing sponastrime dysplasia (*182*). We surmise that failure to bind histone H3.1 by TONSL variants impairs DDR, ultimately leading to genetic defects observed in these patients.

3.6.8 Summary and Conclusions

Over the years, many PTMs deposited on the nucleosome have been identified and biologically characterized. Akin to DNA methylation, histone modifications serve as signals without affecting the identity of the underlying sequences, albeit histone PTMs involve a far more diverse set of chemical groups. To equal the plethora of histone PTMs, eukaryotes have developed a range of specialized histone reader domains which recognize motifs with specific modifications. These readers typically bind five to fifteen residues, and within their small footprint, they may read more than one PTM simultaneously. Moreover, chromatin remodelling complexes can carry several readers, writers, and erasers to expand their readout mechanisms to coordinate nuclear signalling events precisely. Here, we characterized and validated the footprint of TPR^{TSK} on histone H3.1 and revealed that this reader contacts 35 amino acids on the tail of H3.1, which span from its N-terminus to the residues neighbouring the edges of the DNA gyres when embedded in a nucleosome. We further performed studies to characterize the binding preference of TPR^{TSK} for H3.1. From the library of PTMs tested in this study, we found no significant modifications (>2-fold change) that promote TSK binding beyond that of the unmodified H3.1 (Figure 3.1E, Figure 3.2E and Figure 3.2E). The extensiveness of the contact between TSK and H3.1 and TPR's ability to bind H3.1 in multiple forms (NCP, octamer, histone, peptide) suggest that the protein may act as a histone H3.1 chaperone beyond its role during HR-mediating DDR. In such cases, in collaboration with Asf1, TPR may act as a structural scaffold in shielding histone H3.1 N-terminal region from spurious modifications before its nuclear localization and nucleosome assembly. This idea is consistent with our model that the ideal binding epitope for TPR^{TSK} is the unmodified form of histone H3.1. Further consistent with this chaperoning activity of TSK, the human TSK, TONSL, carries a second histone reader, the ankyrin repeat domain (ARD), which binds half of the N-terminal region of histone H4 in its unmodified form (*62*) further suggesting that TONSL is

an H3-H4 chaperone. Moreover, the plant H3K27me0/me1 switch for HR that relies on TPR^{TSK} reading the H3K27me status, appears to have evolutionarily diverged to another switch in humans based an ARD^{TONSL} recognizing H4K20me0/me1 status, but interestingly TPR in TONSL remains conserved and its affinity for unmodified H3.1 is also retained (*142*) suggesting that H3.1 recognition by TPR^{TONSL} is crucial to TONSL activity.

Our binding assays further show that K4, R26, K27, and K36 on H3.1 are PTM-sensitive spots for TPR^{TSK} on histone H3.1 which severely impacts TPR^{TSK} binding for most PTMs tested (Figure 3.1E, Figure 3.2E and Figure 3.2E). Interestingly though, these sites feature the bypass switches K4me1, R26me1, K27ac/cr, and to an extent, K36me1, which appear to serve as coordination mechanisms to time HR with their corresponding pathways or play a role in favouring or deterring HR over other DNA damage repair pathways. Similarly, TPR^{TSK} binding is inhibited by histone H3.1 PTMs marking cell-cycle progression, including marks deposited during mitosis such as T32ph, T22ph, T3ph, and likely S28ph further suggesting that TSK predominantly functions as an S-phase specific DDR protein. Finally, our binding studies on SD mutations suggest that loss of histone H3.1 binding leads to developmental defects and deficiencies in chromosome integrity.

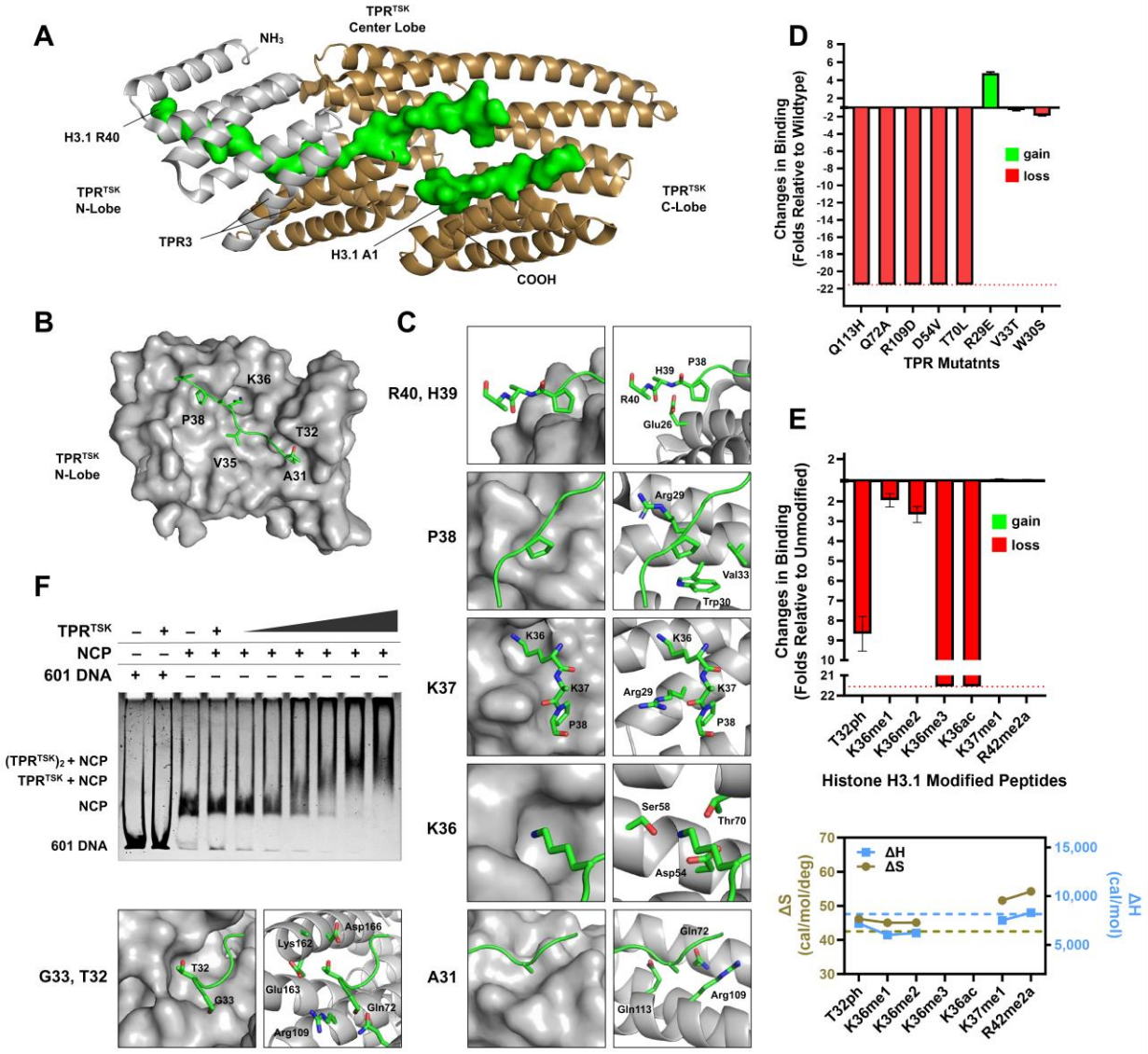


Figure 3.1 – Histone H3.1 residues 31-40 bind the N-terminal lobe (N-Lobe) of TPR^{TSK}
(A) Arrangement of TPR motifs in TSK (cartoon) and histone H3.1 peptide (green surface). TPR helices forming the N-lobe are coloured grey. **(B)** Surface depiction of TPR^{TSK} N-lobe (grey) showing H3.1 side chains as sticks (green). **(C)** Each panel shows a close-up view of the binding pocket of histone H3.1 residues. For each pocket, the left panel shows a surface representation of TPR^{TSK} in grey and the H3.1 residue(s) side chain (green) rendered as sticks; the right panel shows the same pocket as a cartoon with the addition of TPR^{TSK} residues from the pocket as sticks. Oxygen and nitrogen atoms are rendered in red and blue, respectively. TPR^{TSK} and histone H3.1 carbon atoms are coloured green. **(D, E)** ITC assays comparing binding affinities between different mutations in TPR^{TSK} (panel D) or different PTMs on histone H3.1(panel E). The bars represent changes in Kd values, in folds, proportional to the wildtype protein (D) or unmodified H3.1 peptide I. Green bars represent stronger binding (lower Kd), and red bars correspond to loss of binding (higher Kd). Red bars reaching the dotted red lines

near the bottom of the graphs (~21-fold loss) indicate Kd values higher than 50 μM. Error bars shown are standard errors of the mean (SDM, n=2). The bottom portion of panel E demonstrates changes in entropy (ΔS) on the left y-axis (gold) and changes in enthalpy (ΔH) on the right y-axis (light blue) for the corresponding histone H3.1 PTMs. Dashed lines mark the ΔS/ΔH corresponding to the unmodified H3.1 peptide. Kd, ΔS and ΔH values are included in Table 3-1 and Table 3-2. **(F)** Electrophoretic mobility shift assays. Symbols at the top of the gel indicate the presence (+) or absence (-) of corresponding components. Lanes 4 and onward include an increasing concentration of TPR^{TSK}. The band labelled as "601 DNA" represents free DNA or DNA dissociated from the nucleosome.

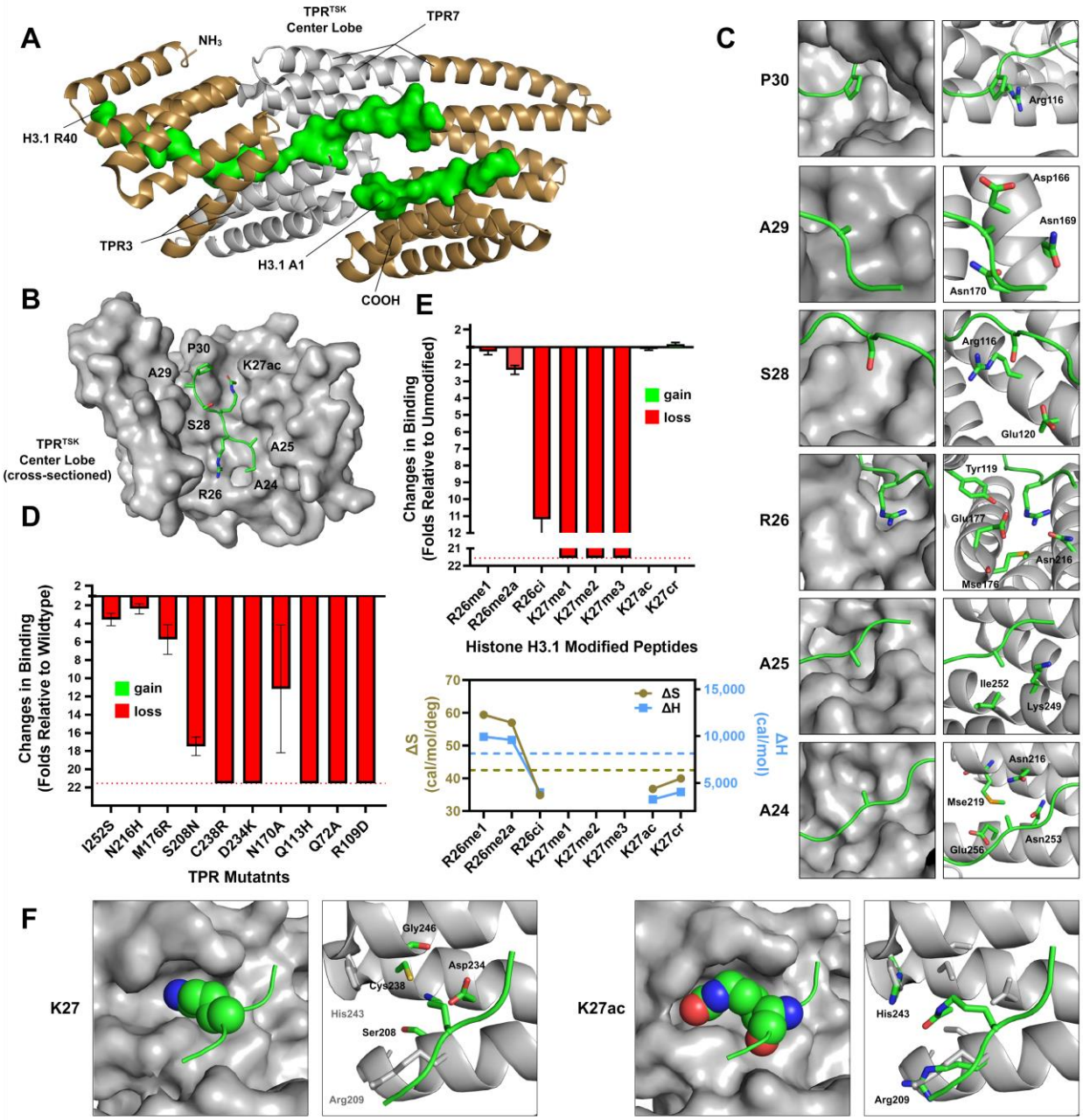


Figure 3.2 – The Center-Lobe of TPR^{TSK} encircles residues 24-30 of histone H3.1. (A) Arrangement of TPR motifs in TSK (cartoon) and histone H3.1 peptide (green surface). TPR helices forming the center-lobe are coloured grey. **(B)** Surface depiction of TPR^{TSK} center lobe (grey) showing H3.1 side chains as sticks (green). **(C)**, **(D)** and **(E)** panels are rendered as in Figure 3.1C, D and E, respectively. Acetylation of H3.1 K27 induces a new side-chain conformation. K_d, ΔS and ΔH values are included in Table 3-1 and Table 3-2. **(F)** Surface and cartoon representation of H3.1K27 and H3.1K27ac binding pockets. Acetylated K27 binds in an alternative binding pocket (black labels). Grey sticks represent the orientation of corresponding amino acids in the TPR^{TSK} structure bound to the unmodified histone H3.1 peptide

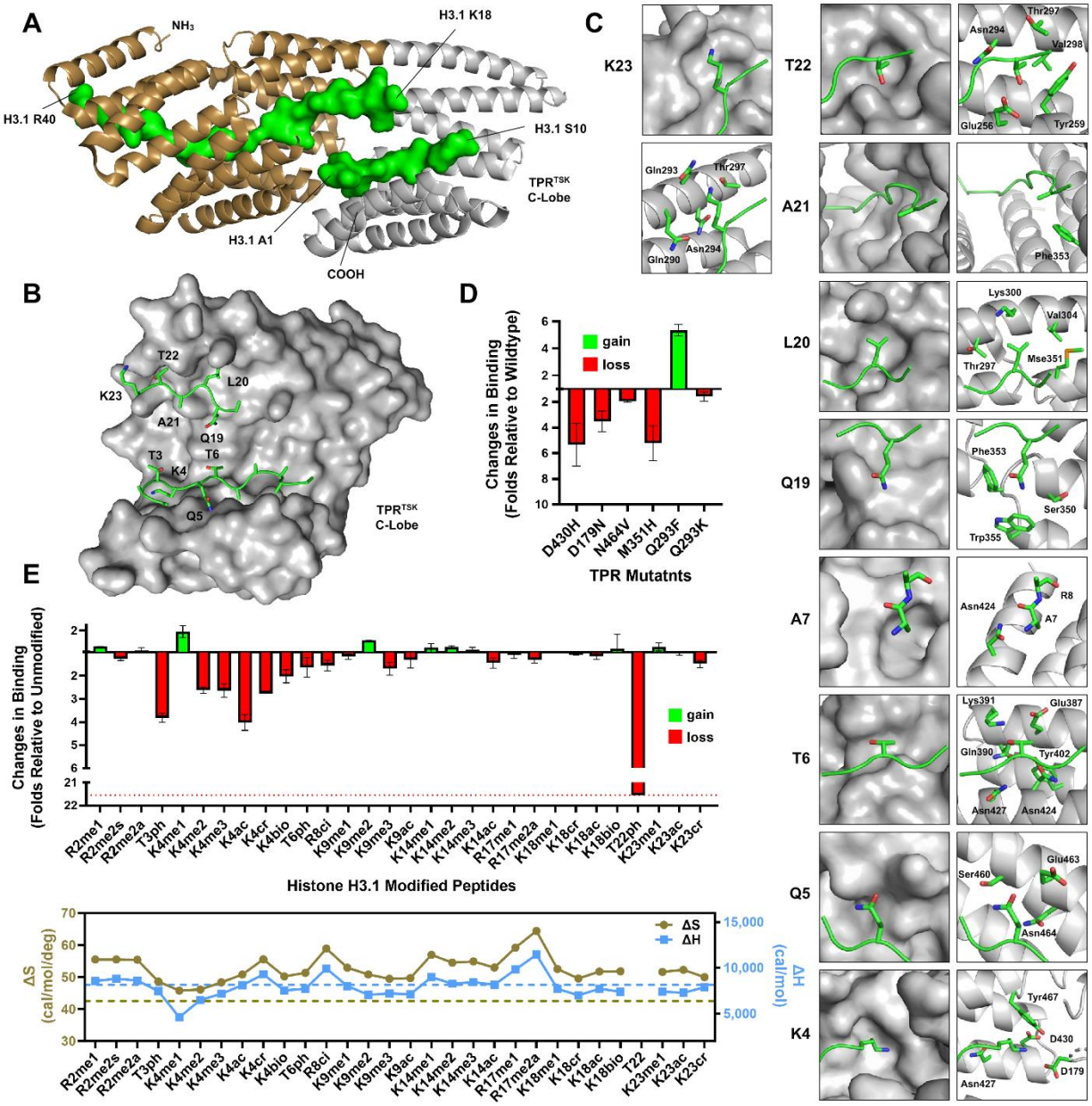
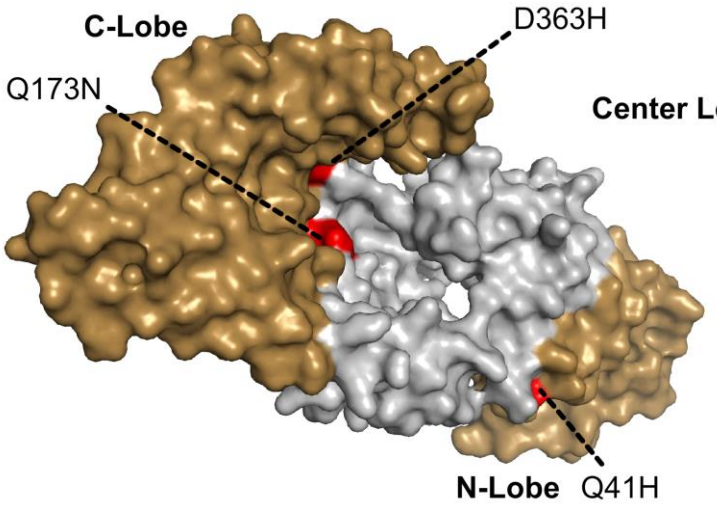


Figure 3.3 – The wide C-terminal Lobe (C-lobe) of TPR^{TSK} binds two regions of histone H3.1 (A) Arrangement of TPR motifs in TSK (cartoon) and histone H3.1 peptide (green surface). TPR helices forming the C-lobe are coloured grey. (B) Surface depiction of TPR^{TSK} C-lobe (grey) showing H3.1 side chains as sticks (green). ***(*C), (D)*** and **(E)** panels are rendered as in Figure 3.1 C, 1D and 1E, respectively. K_d, ΔS and ΔH values are included in Table 3-1 and Table 3-2.



TPR^{TONSL}	Kd (μM)
wildtype	8.55 ± 0.15
Q41H	no binding
Q173N	no binding
D363H	no binding

Figure 3.4 – SD mutations negatively impact histone H3.1 binding. (Top) Mapping of TONSL SD mutants on the TPR^{TSK}. TPR^{TSK} is shown as surface with the SD mutants coloured in red. Equilibrium dissociation constants of TPR^{TONSL} (wild-type and SD mutants) for histone H3.1.

Table 3-1 – Thermodynamic parameters for titration of TPR^{TSK} with post-translationally modified H3.1 peptides.

H3.1 Peptide	K_d (μM)	N (Sites)	ΔS (cal/mol/deg)	ΔH (cal/mol)
unmodified	2.32 ± 0.06	1.26 ± 0.01	47.60 ± 0.00	6373 ± 75
R2me1	1.87 ± 0.13	0.92 ± 0.14	55.50 ± 0.80	8558 ± 269
R2me2s	2.94 ± 0.18	0.98 ± 0.08	55.50 ± 0.80	8814 ± 195
R2me2a	2.20 ± 0.28	0.94 ± 0.00	55.40 ± 0.20	8616 ± 137
T3ph	9.06 ± 0.21	0.97 ± 0.11	48.60 ± 0.20	7453 ± 36
K4me1	1.20 ± 0.27	0.93 ± 0.10	45.75 ± 3.95	4603 ± 446
K4me2	6.10 ± 0.32	1.12 ± 0.06	46.10 ± 0.30	6487 ± 60
K4me3	6.13 ± 0.67	1.20 ± 0.01	48.35 ± 0.75	7161 ± 158
K4ac	9.34 ± 0.78	1.20 ± 0.17	50.80 ± 0.10	8112 ± 28
K4cr	6.41 ± 0.00	1.04 ± 0.03	55.55 ± 0.15	9288 ± 38
K4bio	4.69 ± 0.66	1.22 ± 0.01	50.20 ± 0.40	7525 ± 198
T6ph	3.81 ± 0.97	1.03 ± 0.17	51.30 ± 0.10	7722 ± 114
R8ci	2.07 ± 1.01	1.03 ± 0.04	58.90 ± 1.90	9931 ± 460
K9me1	2.73 ± 0.30	1.12 ± 0.08	52.90 ± 1.70	8010 ± 432
K9me2	1.52 ± 0.01	1.16 ± 0.01	50.80 ± 0.90	7064 ± 263
K9me3	3.93 ± 0.64	1.10 ± 0.08	49.45 ± 0.75	7215 ± 323
K9ac	3.06 ± 0.81	1.11 ± 0.01	49.65 ± 1.85	7125 ± 384
K14me1	1.95 ± 0.33	0.96 ± 0.03	56.95 ± 2.85	8991 ± 938
K14me2	1.91 ± 0.11	0.99 ± 0.01	54.50 ± 3.40	8267 ± 1026
K14me3	2.15 ± 0.26	0.94 ± 0.05	54.85 ± 1.15	8449 ± 408
K14ac	3.38 ± 0.52	1.14 ± 0.01	52.95 ± 0.05	8154 ± 98
R17me1	2.57 ± 0.32	0.96 ± 0.03	59.20 ± 2.70	9814 ± 867
R17me2a	3.08 ± 0.26	0.84 ± 0.12	64.40 ± 6.80	11452 ± 2038
K18me1	2.04 ± 0.26	1.02 ± 0.01	52.50 ± 1.80	7721 ± 596
K18cr	2.54 ± 0.01	1.14 ± 0.10	49.55 ± 1.55	7003 ± 456
K18ac	2.70 ± 0.30	1.03 ± 0.09	51.70 ± 0.60	7722 ± 34
K18bio	2.02 ± 1.17	1.07 ± 0.05	51.80 ± 0.60	7400 ± 212
K23me1	--	--	--	--
K23ac	1.90 ± 0.29	1.11 ± 0.03	51.60 ± 1.50	7419 ± 521
K23cr	2.42 ± 0.19	1.00 ± 0.04	52.25 ± 2.15	7286 ± 30
R26me1	3.45 ± 0.37	1.16 ± 0.01	49.95 ± 0.15	7873 ± 558
R26me2a	2.89 ± 0.42	1.04 ± 0.01	59.40 ± 0.30	9950 ± 2
R26ci	5.38 ± 0.62	1.09 ± 0.12	57.00 ± 1.80	9596 ± 595
K27me1	25.95 ± 11.50	0.77 ± 0.00	34.90 ± 1.40	3998 ± 694
K27me2	--	--	--	--
K27me3	--	--	--	--
K27ac	--	--	--	--
K27cr	2.50 ± 0.18	1.18 ± 0.22	36.75 ± 0.25	3241 ± 118
T32ph	2.01 ± 0.17	1.15 ± 0.12	39.95 ± 1.35	4052 ± 336
K36me1	4.51 ± 0.78	1.24 ± 0.08	45.05 ± 1.65	6010 ± 596

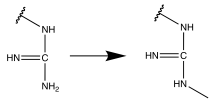
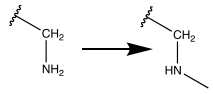
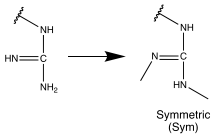
K36me2	6.17 ± 0.92	1.13 ± 0.01	45.10 ± 0.10	6218 ± 107
K36me3	20.11 ± 2.02	0.97 ± 0.00	46.15 ± 2.85	7200 ± 882
K36ac	--	--	--	--
K37me1	--	--	--	--
R42me2a	2.20 ± 0.02	0.97 ± 0.21	51.60 ± 2.90	7515 ± 836

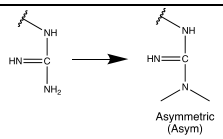
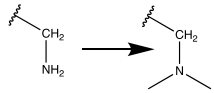
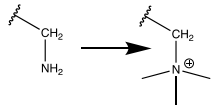
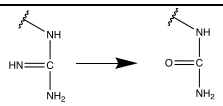
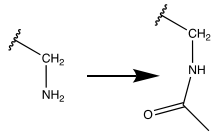
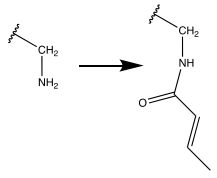
Table 3-2 – Thermodynamic parameters for titration of mutants TPR^{TSK} with H3.1 unmodified peptides.

TPR^{TSK} mutant	K_d (μM)	N (Sites)	ΔS (cal/mol/deg)	ΔH (cal/mol)
wildtype	2.32 ± 0.06	1.26 ± 0.01	47.60 ± 0.00	6373 ± 75
D430H	12.37 ± 3.87	1.34 ± 0.17	56.20 ± 4.20	9824 ± 1406
D179N	8.15 ± 1.86	1.09 ± 0.07	52.05 ± 2.45	8396 ± 844
N464V	4.55 ± 0.14	1.08 ± 0.01	58.10 ± 0.40	9829 ± 143
M351H	12.07 ± 3.14	1.01 ± 0.08	54.95 ± 3.15	9452 ± 1069
Q293F	0.44 ± 0.04	1.04 ± 0.05	53.20 ± 1.00	7032 ± 337
Q293K	3.67 ± 0.84	1.06 ± 0.08	55.30 ± 3.10	8874 ± 1031
I252S	8.21 ± 1.59	1.14 ± 0.09	48.55 ± 3.05	7599 ± 1239
N216H	5.51 ± 1.29	0.95 ± 0.07	50.00 ± 1.30	7570 ± 523
M176R	13.31 ± 3.79	1.08 ± 0.06	44.45 ± 3.85	6447 ± 1297
S208N	40.55 ± 2.35	41.55 ± 1.35	41.55 ± 1.35	6288 ± 444
C238R	--	--	--	--
D234K	--	--	--	--
N170A	25.90 ± 16.30	0.80 ± 0.11	14.02 ± 4.68	2178 ± 941
Q113H	--	--	--	--
Q72A	--	--	--	--
R109D	--	--	--	--
D54V	--	--	--	--
T70L	--	--	--	--
R29E	0.49 ± 0.02	1.00 ± 0.08	57.45 ± 1.75	8344 ± 535
V33T	2.95 ± 0.11	0.97 ± 0.01	54.55 ± 0.05	8547 ± 31
W30S	4.42 ± 0.03	0.97 ± 0.00	56.75 ± 0.05	9423 ± 16

Table 3-3: Peptide synthesis chemical structures based on the main backbone peptide, and post-translational modifications at the different described amino acid positions.

Main Peptides Backbone										
	1	6	10	16	20	26	30	36	40	46
H3.1→	ARTKQ-TARKS-TGGKA-PRKQL-ATKAA-RKSAP-ATGGV-KKPHR-FRPGT-Y									

Name	Structure modification	Molecular Ions found m/z $[M+nH]^{n+}$					
		4+	5+	6+	7+	8+	9+
H3.1 WT	N/A	1231.0	985.0	820.8	703.7	615.9	547.7
H3.1 me1	 or 	1234.2	987.6	823.1	705.8	617.6	549.1
H3.1 me2	 Symmetric (Sym) or	1237.7	990.6	825.5	707.7	619.4	550.8

	 or 						
H3.1 me3		1241.2	993.2	827.8	709.8	621.2	552.2
H3.1 ci		1231.0	985.0	821.1	704.0	616.0	547.7
H3.1 ac		1241.5	993.4	827.8	709.7	621.1	552.3
H3.1 cr		1247.7	998.4	832.2	713.6	624.5	555.1

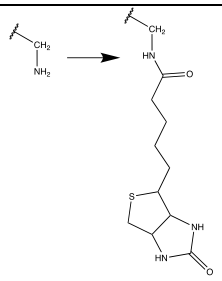
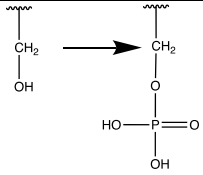
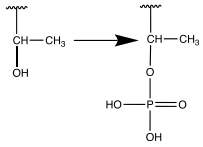
H3.1 bio		1287.5	1030.0	858.5	736.0	644.2	572.7
H3.1 ph	 or 	1250.7	1000.8	834.1	715.3	626.0	556.4

Table 3-4 – Data collection and refinement statistics for TPR^{TSK} and H3.1 K27ac complex.

PDB accession number	
Data Collection	
Space group	$P\ 2_1 2_1 2_1$
Cell dimensions	
a, b, c (Å)	82.93, 92.36, 207.49
α, β, γ (°)	90.0, 90.0, 90.0
Resolution	40.35 - 2.79 (2.89 - 2.79)
R_{meas}	4.7 (47.4)
$I / \sigma I$	
Completeness (%)	97.9 (97.6)
Redundancy	4.5 (4.2)
Refinement	
Resolution (Å)	40.35 - 2.79
Num. reflections	39609 (3782)
R_{work} / R_{free}	0.2603 (0.3819) / 0.2999 (0.4741)
Num. atoms	
TPR^{TSK}	6821
$H3.1^{1-45}$	379
B-factors (Å^2)	
Protein	110
Ligand	111
R.m.s. deviations	
Bond lengths (Å)	0.011
Bond angles (°)	1.39
Molprobity scores	2.64
Clashscore	19.49
Ramachandran favored	95.88
Ramachandran allowed	3.80

* Highest-resolution shell is shown in parentheses.

3.7 Materials and Methods

3.7.1 Site-directed mutagenesis and peptide synthesis.

All the TSK or TONSL mutants were generated using a site-directed mutagenesis kit (Stratagene) and a gene encoding residues 1-530 of *Citrus unshiu* TSK (*Cu*TSK$^{TPR1-530}$) cloned in a pET22a(+) vector. A gene encoding residues 1-461 of *Xenopus laevis* TONSL (*Xl*TONSL$^{TPR1-461}$) were cloned in a pSMT3 vector. DNA Sequencing confirmed that only the desired mutation was introduced. Peptide synthesis was performed as previously described, and lyophilized peptides were resuspended in water and kept at -20°C.

3.7.2 Protein Expression and Purification

*Cu*TSK$^{TPR1-530}$ or *Cu*TSK$^{TPR1-490}$ was overexpressed as a TEV cleavable His-tagged protein and purified as previously described (*87*) Following TEV cleavage in NaP_i buffer (50 mM NaPi pH 7.5, 0.5 M NaCl, 10% glycerol, 5 mM βME), *Cu*TSK$^{TPR1-530}$ was concentrated and further purified by size exclusion chromatography (SEC) using a Superdex 75 (S75) pre-equilibrated in the NaPi buffer supplemented with 350mM NaCl. For crystallography purposes, purification of CuTSK fragment corresponding to residues 1-490 or Tris buffer (20 mM Tris pH 7.5, 200 mM NaCl, 5% glycerol, 5 mM βME) for crystallography. *Xl*TONSL$^{TPR1-461}$ was overexpressed as His-SUMO tagged protein and purified in HEPES buffer (50mM HEPES pH 7.0, 1 M NaCl, 10% glycerol, 5mM βME, supplemented with salt-active nuclease and lysozyme). Following ULP1 cleavage in HEPES buffer (50mM HEPES pH 7.0, 500 mM NaCl, 10% glycerol, 5mM βME) the construct was purified with SEC using an S75 pre-equilibrated with HEPES buffer (50mM HEPES pH 7.0, 200 mM NaCl, 10% glycerol, 5mM βME).

3.7.3 ITC assay

ITC experiments were performed using a VP-ITC calorimeter (MicroCal, Northampton, MA) by titrating *Cu*TSK$^{TPR1-530}$ or XlTONSL$^{TPR1-461}$ (50 μM) with injections of histone H3.1$^{1-45}$ peptides (750 μM or 1250 μM, respectively) in 50 mM NaP_i pH 7.5, 350 mM NaCl, 10% glycerol, and 5 mM βME (*Cu*TSK) or 50mM HEPES pH 7.0, 1 M NaCl, 10% glycerol, 5mM βME (*Xl*TONSL). The experiments were performed at 19°C, and the titration data were analyzed using Origin software (OriginLab Corporation, Northampton, MA). ITC experiments were repeated two times using independent preparations of protein. Equilibrium dissociation constants are indicated in Table S1 and S2.

3.7.4 Peptide synthesis, purification, and characterization

Fmoc protected amino acids, and low-loading Wang resin were purchased from CEM. All peptides were synthesized using microwave-assisted Fmoc solid phase peptide synthesis in a Liberty Blue automated system. Briefly, the required amount of resin was swelled in DMF for 5 min. Next, Fmoc deprotection was carried out with 20% piperidine at 90°C for 60s. Standard coupling cycles using DIC/Oxyma Pure were run at 90°C for 240s in each amino acid. Peptides were cleaved from the resin and deprotected with TFA/TIS/EDT/H_2O (92.5/2.5/2.5/2.5 %v/v) at 42°C for 30 min, and then precipitated in -20°C diethyl ether. Peptide crude products were then dried under vacuum overnight and purified by RP-HLPC in a Waters 1525EF semi-preparative system with a 21.6 x 250 mm C18 column at 20 mL/min. Peptide purity and identity was confirmed via RP-UPLC-UV/MS in a Waters Acquity UPLC Xevo TQD using a 2.1x100 mm UPLC BEH C8 column. A purity of >95% was determined through HPLC peak analysis. Table S3 includes the chemical structures of post-translational modifications synthesized for each peptide described in this work.

3.7.5 Crystallography

*Cu*TSK$^{TPR1-490}$ (20 mg/ml) was incubated with an H3.1K27ac peptide (1-45) (5:1 peptide: *Cu*TSK$^{TPR1-490}$ molar ratio), and the complex was crystalized by vapour diffusion using a mother liquor composed of 25% 1,2-propanediol, 5% glycerol, 0.1 M Na/K phosphate pH 6.0. A full data set was collected at the 21-ID-D beamline of the Life Science-Collaborative Access Team at the Advanced Photon Source Synchrotron. The structure of *Cu*TSK$^{TPR1-490}$ / H3.1K27ac was determined by molecular replacement using *Cu*TSK$^{TPR1-490}$ (7T7T) as a search model. Using Phaser, two molecules of *Cu*TSK$^{TPR1-490}$ were placed in the asymmetric unit. Missing residues were modelled in the calculated phases using Coot (42), and the structure was further refined using phenix.refine (43). Clash scores were determined by MolProbity (44). H3.1 residues 1-11 and 18-40 were modelled for the two H3.1 chains in the asymmetric unit. Refinement statistics are reported in table S4.

3.7.6 EMSA

Electromobility shift assays (EMSA) were performed by loading free DNA 601 element or nucleosomes, alone or pre-incubated (30 min) with TPR^{TSK}, on 4% acrylamide (60:1) gels in TBE (Tris, boric acid, EDTA) buffer at 4 ºC (equilibrated by 1-hour pre-run). NCP and TPR^{TSK} were incubated using the following ratios (1 to 0.5, 1, 2, 4, 8, 16, or 32 (NCP: TPR^{TSK}).

3.8 Authorship Contribution Statement

Hossein Davarinejad: Conceptualization, Methodology, Software, Investigation, Formal analysis, Visualization and Writing–original draft. **Marcelo Munoz and Alex Ross**: Methodology, Formal analysis, Visualization. **Monika Joshi, Sabina Sarvan** and **Joseph S. Brunzelle**: Methodology, and Formal analysis. **Emilio Alarcon**: Conceptualization, Methodology, Investigation and Supervision. **Jean-Francois Couture**: Conceptualization, Supervision, Project administration, funding acquisition and writing an original draft.

3.9 Author Contributions

JFC conceived the project in consultation with HD. HD purified TPR^{TSK} proteins and performed the binding and structural studies. MJ and SS helped to complete the binding studies with the TPR^{TSK} mutants. MM and AR synthesized and validated the quality of the peptides. JSB collected a full dataset of the TPR^{TSK} bound to an H3.1K27ac peptide. MM provided guidance and the infrastructure to synthesize the peptides. JFC supervised the project, acquired funding, and provided resources. HD and JFC wrote the manuscript with contributions from the others.

3.10 Data Availability

TPR^{TSK}/H3K27ac structure has been deposited in the protein databank with the accession number XXXX.pdb

3.11 Acknowledgements

This work was supported by a Canadian Institutes of Health Research (CIHR) project grant (PJG-473915) (awarded to JFC). Hossein Davarinejad and Dr. Marcelo Munoz receive a doctoral and a post-doctoral scholarship from Ontario Graduate Studies and CIHR, respectively.

Chapter 4 | ATXR5/6 Forms Alternative Protein Complexes with PCNA and the Nucleosome Core Particle

Chapter 4 is a manuscript published on 29 March 2019 by Elsevier in the Journal of Molecular Biology. The publisher permits use of this publication in an academic thesis.

JMB | Volume 431 | Issue Number 7: 1370-1379
https://www.sciencedirect.com/science/article/abs/pii/S002228361930097X

4.1 Authors and Affiliations

Hossein Davarinejad[1], Monika Joshi[1], Narimane Ait-Hamou[1], Kim Munro[2] and Jean-François Couture[1*]

1 Ottawa Institute of Systems Biology, Department of Biochemistry, Microbiology and Immunology, University of Ottawa, 451 Smyth Road, Ottawa, ON K1H 8M5, Canada
2 Protein Function Discovery Facility, Queen's University, Kingston, ON K7L 3N6, Canada

* Correspondence to Jean-François Couture:
University of Ottawa, Ottawa Institute of Systems Biology
451 Smyth Road, Roger Guindon Hall, Ottawa, ON K1H 8M5, Canada
jean-francois.couture@uottawa.ca

4.2 My Contributions

For this publication, I purified the protein for all of the experiments. I crystalized and solved the structure of PCNA:PIP. I performed the ITC and binding assays (SEC). I isolated the PCNA:PIP complex for the analytical ultracentrifugation experiment. Additionally, I performed the EMSA assay and the structural comparative analyses. Lastly, I wrote the manuscript.

4.3 Abstract

The proliferating cell nuclear antigen (PCNA) is a sliding clamp associated with DNA polymerases and serves as a binding platform for the recruitment of regulatory proteins linked to DNA damage repair, cell cycle regulation, and epigenetic signaling. The histone H3 lysine-27 (H3K27) mono-methyltransferase Arabidopsis trithorax-related protein 5/6 (ATXR5/6) associates with PCNA, and this interaction has been proposed to act as a key determinant controlling the reestablishment of H3K27 mono-methylation following replication. In this study, we provide biochemical evidence showing that PCNA inhibits ATXR6 enzymatic activity. The structure of the ATXR6 PCNA-interacting peptide (PIP) in complex with PCNA indicates that a trio of hydrophobic residues contributes to the binding of the enzyme to the sliding clamp. Finally, despite the presence of three PIP binding clefts, only two molecules of ATXR6 bind to PCNA likely enabling the recruitment of a third protein to the sliding clamp. Collectively, these results rule out the model wherein PCNA-bound ATXR6 actively reestablishes H3K27 mono-methylation following DNA replication and provides insights into the role of ATXR6 PIP motif in its interaction with PCNA.

4.4 Introduction

Methylation of histone proteins is a widely observed post-translational modification linked to many nuclear transactions (*183*). Among those, methylation of lysine residues plays important roles in chromatin remodeling and epigenetic signaling (*184*). This posttranslational modification occurs mostly on histone tails and is catalyzed by histone lysine methyltransferases, a family of enzymes typically characterized by the presence of a catalytic SET (Su[var], E[z], Trithorax) domain (*185*). Methylation of K27 in histone H3 (H3K27) is predominantly associated with heterochromatin formation and found in poorly transcribed regions (*8*). In mammals, H3K27 methylation is deposited by Drosophila E(z) homologs, EZH1/2, and regulate crucial processes such as the expression of Hox genes (*186*), which control homeotic development, X chromosome inactivation (*187*), and differentiation of embryonic stem cells to mesodermal stem cell lineage (*188*)[7]. In plants, H3K27 methylation is catalyzed by several enzymes including MEA, CLF, SWN, and Arabidopsis trithorax-related proteins (ATXR) 5/6. ATXR5/6 are SET domain proteins (*189*) important for heterochromatin condensation (*24*) and regulate ribosomal RNAs [10] as well as myoinositol phosphate synthase expression (*190*). In heterochromatin, ATXR5/6 also regulate DNA replication as double mutants result in over-replicated heterochromatin (*133*). Mechanistically, ATXR5/6 specifically monomethylate H3.1 (*25*), and post-translational modifications deposited in the vicinity of the methylation site inhibit the enzymes (*25*). The PHD domain of ATXR5/6 preferentially binds the N-terminus of histone H3, and these enzymes mono-methylate H3K27 when incorporated in a nucleosome core particle (NCP) (*155*). An evolutionary conserved region connecting the SET and PHD domains in ATXR5/6 includes a consensus sequence showing high homology to the proliferating cell nuclear antigen (PCNA)-interacting peptide (PIP) motif. PCNA belongs to the DNA sliding clamp family of proteins and assembles as

a homo-trimeric ring-like structure, which circles around DNA. Initially identified as a protein associating with DNA polymerase and an important regulatory factor controlling DNA replication, PCNA is now known to associate with several other proteins including epigenetic signaling and chromatin remodeling enzymes such as histone acetyltransferase (HAT) p300 (*191*), DNA methyltransferase 1 (DNMT1) (*192*), chromatin assembly factor 1 (CAF1) (*193*), Williams syndrome transcription factor:imitation switch/sucrose non-fermentable 2H (WSTF:ISWI/SNF2H) (*194*), and EZH2 (*195*). Yeast two-hybrid and bi-molecular fluorescence complementation assays showed that ATXR5/6 colocalize with PCNA (*132*). Studies on genomic integrity in plants also revealed that in addition to mutations in the PHD and SET domains, mutation in ATXR5/6 PIP motif disrupts chromatin organization and H3K27 methylation (*133*); yet, little is known about the underlying mechanisms of ATXR5/6 PIP contribution in maintaining this important epigenetic mark. In the present study, we have determined the structure of ATXR6 PIP motif in complex with PCNA. We find that despite three occupied sites on the PCNA trimer by a peptide corresponding to ATXR6 PIP motif, only two molecules of the full-length enzyme can bind to the sliding clamp. The structure and biochemical analysis reveal how a central hydrophobic residue in the ATXR6 PIP motif confers high-affinity binding to PCNA. We also find that methylation of the NCP by ATXR6 is blocked by PCNA, suggesting that ATXR6 must dissociate from PCNA to re-establish H3K27 methylation.

4.5 Results and Discussion

4.5.1 An evolutionary conserved motif in ATXR6 mediates its interaction with PCNA

Alignment of the region separating the PHD and SET domains of ATXR6 between 25 plant species (Figure 4.1a) shows that a region comprising residues QTKIIDFF has high-sequence similarity to a peptide motif (Q-x-x-V/L/I-x-x-F/Y-F/Y) found in proteins known to interact with PCNA. To

confirm that this region directly mediates an interaction between PCNA and ATXR6, each protein was homogeneously purified, co-incubated, and separated by size exclusion chromatography. As shown in Figure 4.1b, gel filtration profile shows that the protein complex elutes as a single peak with an apparent molecular weight of ~170 kDa. However, incubation of a PIP-motif mutant, MnATXR6 I103R, with PCNA yields two protein species eluting at molecular weights of ~90 and ~40 kDa. Residues Leu126/ Ile128 in PCNA are critical for PIP binding but dispensable for trimerization of the protein (*196*). The double mutant was purified, incubated with wild-type ATXR6, and separated by gel filtration. As demonstrated in Figure 4.1b and similar to the observation made with ATXR6 PIP mutant, two peaks corresponding to the trimeric form of PCNA (L126A/I128A) and ATXR6 were observed, suggesting that both ATXR6 PIP motif and the PIP binding cleft of PCNA are important for the formation of the complex. To define the stoichiometry of the complex, ATXR6 wild-type or mutant was incubated with PCNA and separated using analytical ultracentrifugation (Figure 4.1c). Incubation of wild-type ATXR6 with PCNA shows a predominant peak at an s-value of 6, which corresponds to an approximated molecular weight of 174 kDa. Conversely, incubation of ATXR6 I103R with PCNA shows two predominant peaks at s-values corresponding to 40.1 and 90.8 kDa, indicating that the mutation impairs the association of ATXR6 to PCNA. Collectively, these results suggest that ATXR6 PIP motif participates in the binding of PCNA and that two molecules of ATXR6 bind to a trimer of PCNA. The homotrimeric nature of PCNA provides the mechanism and possibility for more than one partner to bind simultaneously. In such model previously coined as sliding clamp “tool belt” (*197*), a variety of PCNA partners are in the bound state and ready to engage at the site of action. Notably, studies using mutant/wild-type heterotrimers of PCNA have shown that in cases such as the Okazaki fragment maturation, inactivation of two of the three PCNA binding sites has no effect on the

biological process (*198*), suggesting that site anchoring may not play a significant role in sequential binding of all PCNA partners. The vast majority of available stoichiometric and structural data investigating the interaction of PIP-motif proteins with trimeric PCNA are limited to analyses using a short peptide and report a 3(PIPmotif): 1(PCNA trimer) ratio. Only few studies provide insight into the stoichiometry of PCNA bound to full-length proteins. For example, one molecule of RNAse HII occupies each of the three PCNA monomers in distinct orientations, but collectively, the three molecules come in close proximity to block the aperture of the DNA clamp in the absence of DNA (*199*). Similarly, the crystal structure of flap endonuclease 1 (FEN-1) bound to PCNA shows the same stoichiometry of binding and captures FEN-1 in different orientations, in one of which, FEN-1 is pulled away from the PCNA central hole and supposedly inactive, and another orientation in which FEN-1 is close to the central hole where interaction with DNA is possible for endonuclease activity. Similar to ATXR6, growth arrest and DNA damage- inducible protein 45 (Gadd45), a DNA damage response protein, binds PCNA in a 2:1 ratio [25]; however, Gadd45 does not contain a PIP-motif and no structural data are available for the full-length protein in complex with PCNA. Interestingly, unlike RNSase HII and FEN-1, ATXR6 and Gadd45 have no known activity on DNA; rather, they serve as signaling proteins. PCNA was reported to regulate the turnover of SET8 by serving as a co-recruitment hub for both SET8 and its associated E3 ligase, $CRL4^{Cdt2}$, which targets SET8 for ubiquitin mediated proteasomal degradation (*200, 201*). Akin to SET8, ATXR5/6 expression is regulated during cell cycle progression (*132*). Therefore, it is possible that in addition to its regulation at the expression levels, the remaining binding pocket on PCNA is employed to recruit the ubiquitination machinery to regulate the turnover of the methyltransferase.

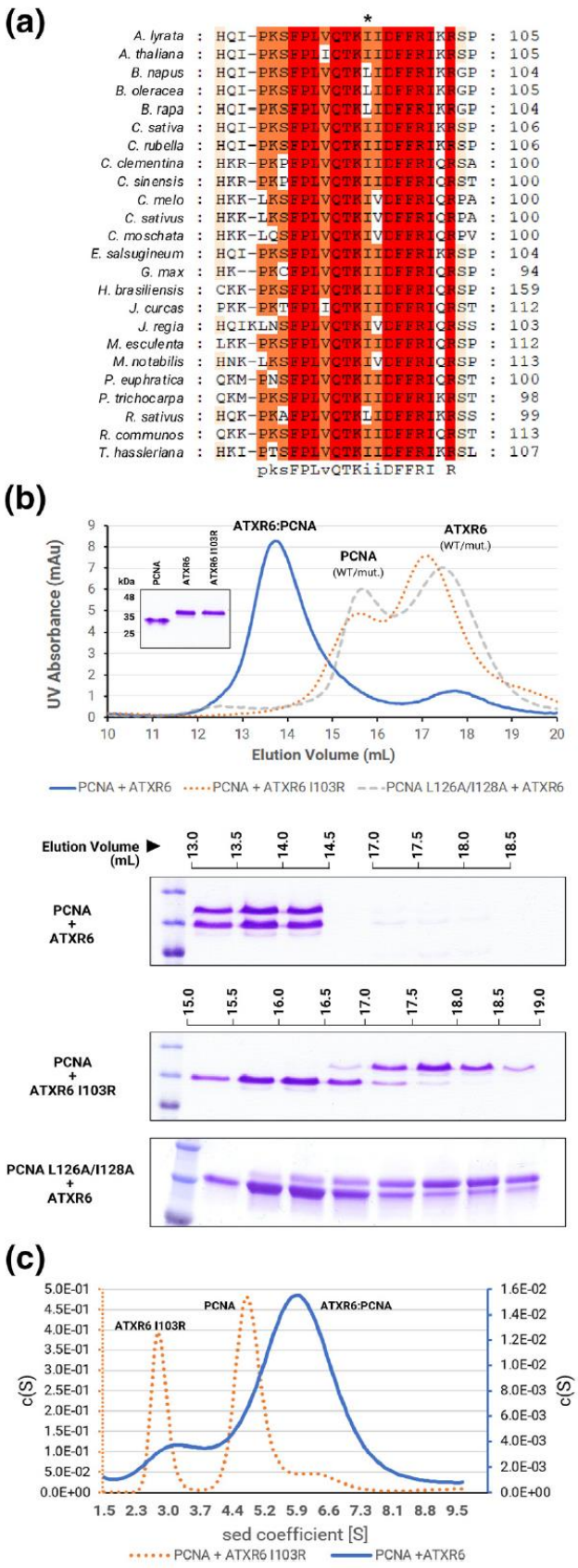


Figure 4.1 – ATXR6 directly associates with PCNA. (a) Sequence alignment of ATXR6 PIP motifs across multiple plant species. The asterisk indicates the isoleucine residue, or its equivalent in other ATXR6 homologs, mutated in this study. **(b)** Gel filtration profiles of PCNA: ATXR6 complex, PCNA + ATXR6 I103R, and PCNA L126A/I128A + ATXR6. Coomassie-stained SDS-PAGE corresponding to each peak fraction is shown below the curves. Individual proteins used in these experiments are shown as inlet overlaid on the gel filtration profiles. **(c)** Analytical ultracentrifugation sedimentation patterns for PCNA:ATXR6 and PCNA + ATXR6 I103R.

4.5.2 Crystal structure of PCNA in complex with ATXR6 PIP motif

To understand the molecular interactions mediating the ATXR6:PCNA complex, we solved the crystal structure of *At*PCNA in complex with *Gm*ATXR6 PIP motif at 2.06 Å using molecular replacement (Table 4-1). The crystal lattice has a P1 space group arrangement, and the asymmetric unit contains six polypeptide chains of PCNA and six chains corresponding to the ATXR6 PIP motif. Of the 19 amino acids belonging to the ATXR6 peptide [residues 78– 96], we unambiguously modeled 13 residues [78– 90] in the electron density map, whereas residues 91 to 96 had no electron density and were omitted from the structure. The structure of the PCNA trimer shows a multifaceted three-lobed disc with a circular hole at the center (Figure 4.2a and b). The monomeric unit has distinct topologies and is connected to other monomers in an N- to C-terminal fashion (Figure 4.2a and c). Each monomer houses an arc-shaped region formed by four slanted α-helices, which line the central hole ring of the trimeric unit where DNA interaction is possible. The back of each inner ring arch is supported by a second and opposite-facing arch formed by a twisted nine-stranded antiparallel blade composed of β2–4, β6, β10, β14, and β16–18, which runs the length of the monomer. Two additional antiparallel blades encompass the arches on the sides of the molecule: a 5-stranded β sheet (β1, β5, β7–9) located at the N-terminal side of the monomer, and at the opposite side, near the C-terminus, a four-stranded β-sheet (β11–13, β15). In the trimeric unit, at the junction of the monomers, the N-terminal β-sheet of one monomer joins the C-terminal β-sheet of the neighboring monomer to form a larger and continuous antiparallel nine stranded blade (Figure 4.2a) with a similar twist as the central blade (Figure 4.2c). These interactions are held in place by backbone hydrogen bonds between β8 (N-terminal sheet) residues Glu115, Asp113 and Ile111, and β13 (C-terminal sheet) residues Thr177, Asn179, and Val181 (Figure 4.2a). The central (monomers) and the junctional nine-stranded blades form the outer ring of the PCNA

trimer and shield the inner ring composed of 12 α-helices. The outer ring is also made distinct by protrusion of the N/C-terminal loops and an inter-blade connector loop that extends the length of the monomer in front of the central blade connecting β9 and β10 (Figure 4.2a). The crystal structure reveals that one molecule of the ATXR6 PIP motif occupies a binding site located on each PCNA monomer (Figure 4.2a). On PCNA, the binding site spans the space that is surrounded by and in close proximity to the long inter-blade loop, the C-terminal loop, and the loops connecting β3–β4 and β16–β17 of the central blade (Figure 4.2a). The PCNA trimer disk can be distinguished by two faces: a top surface on which the PIP-motif binding sites appear (Figure 4.2a and b) and a reverse side opposite to the PIP binding motif. The peptide predominantly adopts an extended conformation, with a 1-turn 3_{10} helix spanning residues I85, D86, and F87 (single-letter residues denote the peptide) facilitated by intra-molecular interactions in which the amide groups of both F87/88 engage in a hydrogen bond with the carbonyl group of I85, and hydrogen bond with K83 and D85 carbonyl and amine groups, respectively (Figure 4.2c and d). Within the PIP binding pocket of PCNA, two pockets and a shallow groove appear as key interaction sites. The large PCNA pocket is formed by several residues located in the β3–β4 connecting loop and the inter-blade loop (Figure 4.2a). The large pocket and its adjacent shallow groove, which lies atop the β16–β17 loop (Figure 4.2a), are lined with multiple hydrophobic residues including Leu126 and Ile128 (Figure 4.2c and d). In the PCNA:ATXR6 PIP complex, ATXR6 I84 and F87 are buried in the large pocket and F88's phenyl ring lays flat in the shallow groove. Comparative analysis of PIP-bound PCNA structures of ATXR6, FEN-1 (*202*), DNA Polymerase δ p66 (*203*), and spartan (SPRTN/DVC-1) (*204*) (Figure 4.3a and b) reveals a shared mode of binding in which the shallow groove is occupied by a bulky hydrophobic residue, while the large pocket accommodates one bulky and one non-bulky hydrophobic residues. Moreover, the short 3_{10} helix is common to other

PIP motifs (not shown) and appears to orient the hydrophobic residues of the motif in such a way that they contact the hydrophobic patch of PCNA (Figure 4.3a and b). The smaller binding pocket is positioned between the β3–β4 connecting loop and the C-terminal loop. In the small aperture of this pocket, ATXR6 Q81 makes (Figure 4.2c and d) key hydrogen bonds between its side chain's amide group and the carbonyl of Ala-252 (Figure 4.2d). The binding mode and the residue, a glutamine, interacting with PCNA small binding pocket are conserved, as similar interactions are observed in the PIP motif of FEN-1 (*202*), DNA polymerase δ p66 (*203*), and SPRTN/DVC-1 (*204*) (Figure 4.3a and b). Complementary to this conserved interaction, PCNA C-terminal loop engages three additional backbone to backbone hydrogen bonds with the N-terminal residues of the PIP motif. These interactions hold the two un-ordered loops together in a parallel β-sheet-like manner and involve one hydrogen bond between Pro253 carbonyl group and the amide of T82 while V80 and Ile255 main chains share two hydrogen bonds. The PIP motif is further stabilized by hydrogen bonds between K83 ε-amine and the hydroxyl group of Ser43, located in the β3–β4 connecting loop, as well as R89 carbonyl group of the amide of Gly127. Consistent with the large set of interactions, isothermal calorimetry (ITC) shows that ATXR6 PIP motif binds to PCNA (Figure 4.4) with a dissociation constant of ~1.1 μM, binding affinities that are comparable to the PIP motif of PARG (*205*) and Pol η (*206*) within 2- to 3-fold differences.

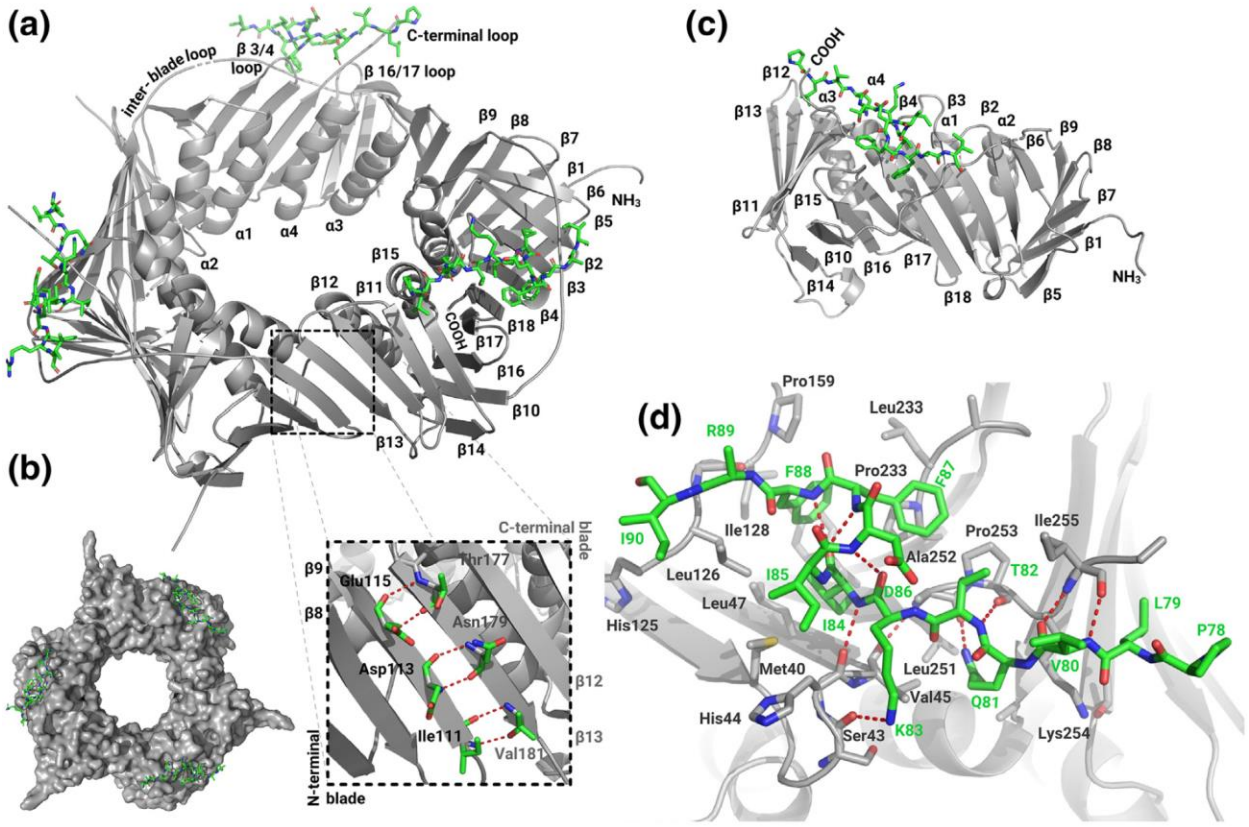


Figure 4.2 – Crystal structure of ATXR6 PIP motif in complex with PCNA (a) Cartoon representation of a PCNA trimer bound to three ATXR6 PIP motifs. Labeled β-sheets are indicated on a single PCNA monomer, while loops and α-helices are labeled on a neighboring monomer. ATXR6 PIP motif is highlighted as stick. Carbon, oxygen, and nitrogen atoms are colored in green, red, and blue respectively. A close-up view of the junction between two PCNA monomers is indicated as a panel showing the N-terminal blade of one monomer and the C-terminal blade of the neighboring monomer. Hydrogen bonds are rendered as red dash lines. **(b)** Surface representation of the structure shown in panel A. **(c)** Cartoon representation of a single PCNA monomer. **(d)** Zoomed view of *At*PCNA PIP binding pocket (gray) and ATXR6 PIP motif (green) highlighting the molecular interactions.

Table 4-1 – Data collection and refinement statistics for the ATXR6 PIP/PCNA complex.

PDB accession number	6O09
Data collection	
Space group	*P*1
Cell dimensions	
a, *b*, *c* (Å)	72.9, 90.5, 90.5
α, *β*, *γ* (°)	60.0, 73.4, 73.5
Resolution	29.65–2.06 (2.13–2.06)[a]
R_{meas}	0.06 (0.36)
$I/\sigma I$	9.1 (2.1)
Completeness (%)	94.4 (91.1)
Redundancy	1.9 (1.8)
CC1/2	0.99 (0.59)
Refinement	
Resolution (Å)	29.65–2.06
No. reflections	219,871
R_{work}/R_{free}	0.1970/0.2329
No. atoms	
PCNA	11,252
$ATXR6^{PIP}$	659
Water	406
B-factors (Å^2)	
Protein	46.7
Ligands	52.8
Water	49.2
R.m.s. deviations	
Bond lengths (Å)	0.007
Bond angles (°)	0.919
Molprobity scores	1.40
Ramachandran favored (%)	2.2
Ramachandran allowed (%)	97.5

[a] Highest-resolution shell is shown in parentheses.

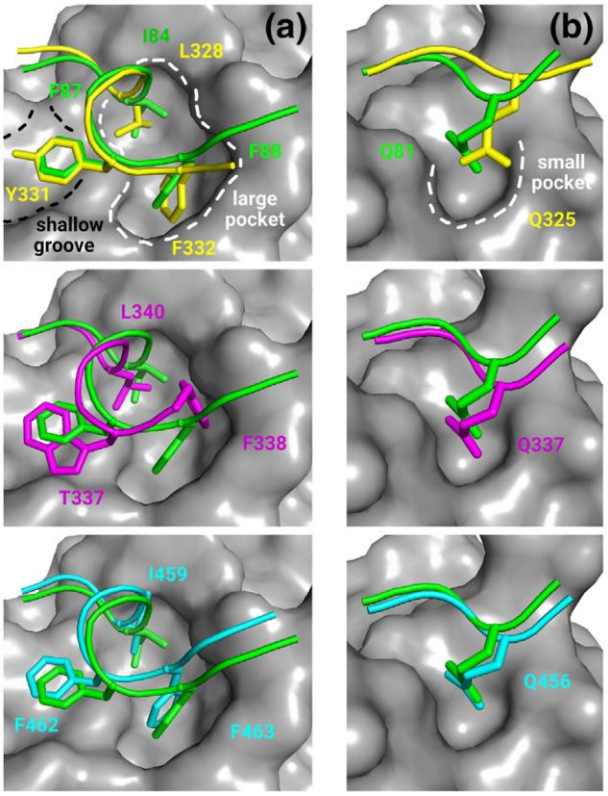


Figure 4.3 – Comparative analysis of PIP:PCNA complexes (a) Close-up view of (a) the large binding pocket and the shallow groove as well as the (b) small binding pocket of PCNA. Panels compare the PIP motifs of ATXR6 (green), FEN-1 (5DAI; magenta), DNA polymerase δ p66 (1U76; cyan), and SPRTN/DVC1 (5IY4; yellow). Images were produced using PyMOL by retrieving available structures fetching PDB IDs and aligning associated PCNA molecules with the PCNA structure from this work.

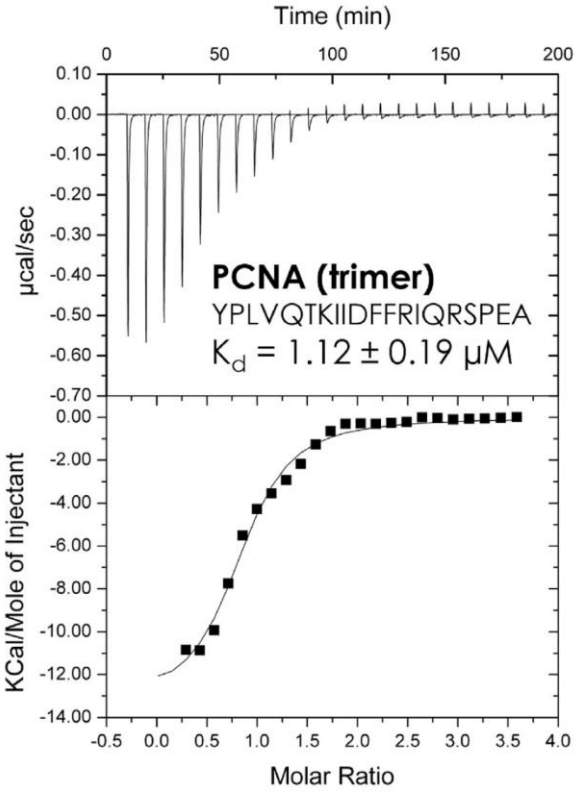


Figure 4.4 – ATXR6 PIP motif binds with high affinity to *At*PCNA. (a) Isothermal titration calorimetry curve of ATXR6 PIP titrated into PCNA. The upper and lower panels show the heat of binding for peptide binding and the integrated data, respectively. The calculated N value is ~0.9. Calculation was performed assuming that one ATXR6 PIP motif binds to one protomer of PCNA.

4.5.3 PCNA inhibits nucleosome methylation by ATXR6

Mutation of AtATXR5 PIP motif disrupts H3K27 methylation in plants (*133*) and the enzyme preferentially mono-methylates histone H3 when incorporated in the NCP over its free form (*155*). To test the interplay between the binding of ATXR6 to PCNA and NCP mono-methylation by the enzyme, we performed histone lysine methyltransferase assays either in the absence or in the presence of the sliding clamp. As shown in Figure 4.5a, in the absence of PCNA, ATXR6 methylates the NCP. However, increasing concentration of PCNA results in a progressive loss of NCP methylation by ATXR6. Accordingly, the inhibitory effect of PCNA on ATXR6 I103R enzymatic activity is lower when compared to wild-type (Figure 4.5b), suggesting that the interaction of PCNA with ATXR6 negatively regulates histone H3 methylation. To test whether PCNA prevents the binding of the cofactor or impairs the formation of the methyltransfer pore, we performed methyltransferase assays using a peptide corresponding to the ATXR6 methylation site [histone H3.1 18–35]. As shown in Figure 4.5c, increasing concentration of PCNA does not prevent the methylation of the peptide, suggesting that the binding of PCNA to ATXR6 does not directly affect its methyltransferase activity but rather prevents its association to the nucleosome. To further demonstrate this idea, we performed electromobility shift assays (EMSAs) with ATXR6 and nucleosome either in presence or absence of PCNA. Our results show that the addition of PCNA prevents the association of ATXR6 to the nucleosome as illustrated by the similarity in the migration profile between the sample devoid of the methyltransferase and the excess of the sliding clamp (Figure 4.5d). Previous studies revealed that mutation of ATXR5/6 PIP motif leads to a complete loss of methylation in plant, and this effect was initially attributed to the inability of the mutant to associate with PCNA and methylate histone H3.1 during the passage of the replication machinery. However, our studies show that PCNA blocks the methyltransferase activity of

ATXR5/6 on the NCP. These observations are reminiscent to the inhibition of the H4K20 methyltransferase SET8 by the sliding clamp (*207*) and may point to a model wherein PIP containing histone methyltransferase must disengage from PCNA to re-establish epigenetic marks on histones following replication. In such model, PCNA would serve as a carrier that ensures a local concentration of these enzymes at the replication fork. Alternatively, as proposed for SET8 (*207*), PCNA may outcompete histone methylation at stalled replication forks as a mean to clear the methyltransferase from the locus via the ubiquitin pathway. Mutations in the PIP-motif of ATXR5/6 disrupt the genomic integrity and H3K27 distribution in plants (*133*). In this work, we demonstrate that this conserved motif is essential for ATXR6 interaction with PCNA. Our data also provide a map of the molecular interactions between the motif and PCNA as well as show that PCNA-bound ATXR6 is unable to methylate histone H3.1 K27 in nucleosomes highlighting a potential shared binding region on ATXR5/6 for PCNA and the nucleosome.

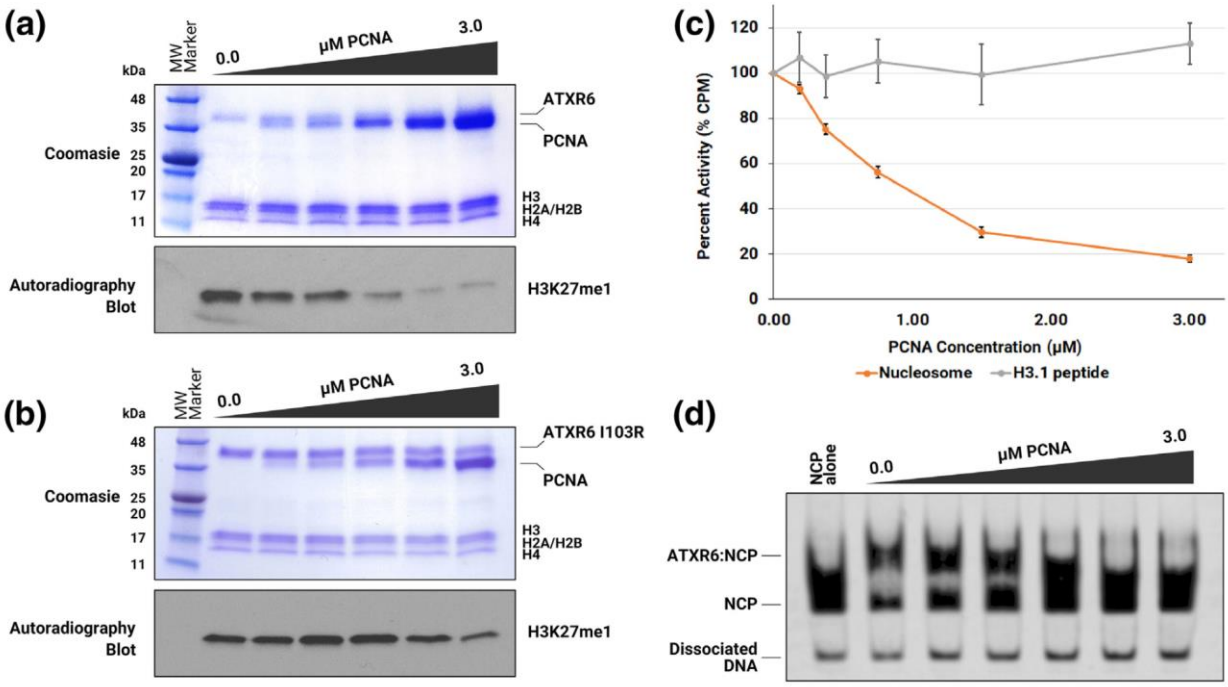


Figure 4.5 – PCNA inhibits ATXR6 methylation on NCP but not H3.1 peptide. (a and **b)** Methyltransferase assays were performed in the absence (0.0 μM) or with increasing concentrations of PCNA (≤3.0 μM) with either wild-type (a) or I103R (b) ATXR6. The top and bottom panels show Coomassie-stained SDS-PAGE gels of each sample and detection of NCP methylation by autoradiography, respectively. **(c)** Quantitative analysis of nucleosome (orange) or H3.1 18–25 peptide (gray) methylation by scintillation counting either in the absence or in the presence of PCNA. Activity is represented as a percentage of the reactions performed in the absence of PCNA. Error bars indicate the standard deviation of each reaction, which were performed at least three times in triplicate. **(d)** EMSAs comparing the migration profiles of ATXR6-bound NCP in the absence or presence of an increasing concentration of PCNA.

4.6 Materials and Methods

4.6.1 ATXR6 expression and purification

A codon-optimized construct corresponding to *Morus notabilis (Mn)* ATXR6 was cloned into pGST-4 T-1, and the plasmid was transformed in Escherichia coli BL21 DE3 Rosetta cells. The following day, several colonies were transferred into 1 L of Luria–Bertani supplemented with 100 μM $ZnCl_2$ and incubated at 37 °C. Once the OD_{600nm} reached 0.7, protein expression was induced using 0.2 mM IPTG and incubated at 18 °C for 16 h. Following centrifugation, cells were collected in phosphate-buffered saline supplemented with 5 mM βME and cells were lysed by sonication. Following centrifugation of the lysate at 31,000g and filtration of the supernatant, the GST–ATXR6 fusion protein was applied onto glutathione Sepharose (Pierce). Following extensive washing steps, the protein-bound resin was TEV cleaved for 16 h at 4 °C in 20 mM Tris (pH 8.0), 500 mM NaCl, and 5 mM βME and collected as flowthrough. MnATXR6 was concentrated and further purified by size exclusion chromatography using a Superdex™ 75 pre-equilibrated in gel filtration buffer [20 mM Tris (pH 8.0), 200 mM NaCl, 5 mMβME]. The protein elutes as a monomer at an apparent molecular weight of 40.1 kDa.

4.6.2 PCNA expression and purification

A codon-optimized construct for bacterial expression corresponding to *Arabidopsis thaliana* PCNA was cloned into pET29(+)a. PCNA was overexpressed for *Mn*ATXR6, and bacterial cell pellets were resuspended and lysed in 50 mM sodium phosphate (pH 7.0), 500 mM NaCl, and 5 mM βME by sonication. Following centrifugation of the lysate at 31,000g and filtration, the supernatant was applied on cobalt-agarose (TaKaRa) resin preequilibrated in lysis buffer. The beads were washed with 10 column volume of lysis buffer and eluted with the same buffer supplemented with 500 mM imidazole. The eluted protein was concentrated and applied onto a

Superdex™ 200 equilibrated with gel filtration buffer. Consistent with its ability to form a trimer, the elution profile shows that PCNA elutes at an apparent molecular weight of ~90 kDa. Crystallization A synthetic peptide corresponding to the PIP motif (residues 78–96) of *Glycine max (Gm)* ATXR6 was ordered from GeneScript with an extra tyrosine located at the N-terminus to facilitate quantification. The lyophilized PIP peptide was initially suspended in water and subsequently introduced to PCNA in a 2:1 molar ratio in gel filtration buffer. The complex was incubated for 2 h at 4 °C. Crystals were grown in 0.2 M sodium malonate and 20% PEG 3350. Harvested crystals were cryoprotected using Fromblin®, and a full data set was collected using a Rigaku MicroMax- 007HF. Images were collected using an R-AXIS IV++ detector (Rigaku), and the data set was processed and scaled using Structure Studio (Rigaku). The structure was solved by molecular replacement using Phenix (*208*) and a search model corresponding to A. thaliana PCNA (*At*PCNA; 2ZVV—PDB code) (*209*). Six PCNA molecules were modeled in the asymmetric unit. The model was built, refined, and validated using Coot (*210*), Phenix, and Molprobity (*211*), respectively. Figures of the PCNA/PIP complex were generated using PyMOL Molecular Graphics System (Schrödinger, LLC).

4.6.3 ITC

ITC experiments were performed using a VP-ITC calorimeter (MicroCal) by injecting the peptide (0.22 mM) into a solution of PCNA (10 μM) in gel filtration buffer. The experiment was performed at 20 °C, and the titration data were analyzed using Origin software (OriginLab Corp.). The *Gm*ATXR6 PIP peptide binds to *At*PCNA with a binding stoichiometry (N values) of ~0.9 molecule of peptide per monomer of PCNA.

4.6.4 Gel filtration and analytical ultracentrifugation

ATXR6 (WT/mutant) was co-incubated at a 4:1 molar ratio with a PCNA (WT/mutant) trimer in gel filtration buffer. Protein mixtures were applied to a Superdex™ 200 pre-equilibrated with gel filtration buffer. Following size exclusion chromatography, the main peak was concentrated, and sedimentation velocity measurements were performed at 20 °C using a Beckman Coulter Optima XL-I analytical ultracentrifuge. Samples were housed in a centrifuge cell equipped with double-sector charcoal-Epon centerpiece. Sedimentation was carried out at 30,000 rpm in an An60-Ti rotor, during which 200 concentration distributions were determined with Abs280 optics 180-s intervals. The sedimentation coefficient distribution was determined using the program SEDFIT (*212*).

4.6.5 Methyltransferase assays

Methyltransferase assays were performed in 25 μL reaction volumes as previously documented (*213*). For each assay, ATXR6 (1 μg) was incubated with either NCPs (~0.3 μM) or an H3.1 peptide corresponding to residues 18–36 (20 μM), and tritium-labeled S-adenosyl-methionine [3H-SAM] (2 μCi) in a methyltransferase assay buffer composed of 50 mM Tris (pH 8.0), 20 mM KCl, 10 mMMgCl2, 10 mMβME, and 10% glycerol. Methyltransferase assays were performed in the absence or increasing concentration of PCNA (0.0 to 3.0 μM) and incubated at 30 °C for 1 h for NCPs or 20 h for reactions performed with peptides. Reactions were stopped by adding Laemmli buffer and boiling. Quantitative analysis of the methyltransferase assays was performed by spotting the reactions on P81 cellulose filter paper (Reaction Biology Corp.). After drying, filters were washed four times in a 50-mM sodium bicarbonate solution and dried again. Filter papers were sectioned and soaked in 5 mL of ScintiSafe™ Econo1 (Fisher) scintillation liquid prior analysis using a Tri-Carb2910TR (PerkinElmer).

4.6.6 EMSAs

EMSAs were performed as previously described (*155*). Briefly, increasing concentration of PCNA (0.0 to 3.0 μM) was added to a binding reaction composed of 1 μg ATXR6, 10 μM Cy5-labeled NCP harboring a K27M mutation, 1 mM AdoHcy, 1% bovine serum albumin, 0.1% Triton-X100, 5 ng/μl poly deoxyinosinic–deoxycytidylic [dI–dC], 50 mM Tris (pH 8.0), 20 mM KCl, 10 mM MgCl2, 10 mM βME, and 10% glycerol. The samples were applied to a 6% native poly-acrylamide gel (60:1).

4.7 Accession numbers

Crystallographic data for the ATXR6 PIP/PCNA complex were deposited with the PDB accession number 6O09.

4.8 Acknowledgments

This work was supported by a Natural Sciences and Engineering Research Council discovery grant awarded to J.-F.C (RGPIN-2016-04977). We would like to thank Dr. Sabina Sarvan for her technical assistance.

4.9 Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2019.02.020.

4.10 Acknowledgments

This work was supported by a Natural Sciences and Engineering Research Council discovery grant awarded to J.-F.C (RGPIN-2016-04977). We would like to thank Dr. Sabina Sarvan for her technical assistance.

Chapter 5 | Discussion and Perspectives

5.1 Summary

In mammals, TONSL contributes to the recruitment of the DNA repair machinery specific to HR. This activity is mediated by ARD^{TONSL}, which recognizes and binds H4K20me0. However, this process is bypassed by the enzyme H4K20 mono-methyltransferase SET8. It is worth noting that this model does not apply to plants, as neither ARD nor SET8 are conserved.

During my PhD studies, I made a significant contribution to understanding HR in plants with the help of our collaborators. In Chapter 2, I identified the TPR domain in TSK as a replacement for this HR mechanism mediating the pathway by recognizing H3.1K27me0. I demonstrated that methylation of this site (H3.1K27me1) by ATXR5/6 methyltransferases post-replication serves as an "off" switch for the mechanism. Furthermore, I solved the crystal structure of TPR^{TSK}, revealing the solenoid folding of its α-helical motifs and the mechanism by which it distinguishes histone H3 variants H3.1 and H3.3 from each other.

In Chapter 3, I conducted a structural and biochemical characterization of all TPR^{TSK} binding pockets for histone H3.1. This investigation revealed that a large number of contacts in the complex stabilize H3.1 on TPR^{TSK}. I validated these binding pockets and demonstrated that sponastrime dysplasia mutations in and near the enclosed central lobe of the TPR lead to a loss of TPR's ability to bind histone H3.1. Additionally, I provided a tail-wide binding profile for histone H3.1 PTMs, in which most PTMs were shown to inhibit TPR^{TSK} binding. I also solved the structure of H3.1K27ac bound to TPR^{TSK}, revealing how the rearrangement of the K27 pocket accommodates the binding of the modified residue. Furthermore, I showed that TPR^{TSK} binding to the nucleosome does not displace the DNA.

In Chapter 4, I delved into the mechanisms that regulate the activity of ATXR5/6 behind the replication fork. I demonstrated that the well-conserved PIP motif in these proteins binds PCNA, and this event inhibits their methyltransferase activity by physically outcompeting the binding of the nucleosome. Finally, I solved the crystal structure of PCNA bound to three molecules of PIP, revealing their mechanism of binding. Moreover, I showed that only two full-length ATXR5/6 proteins can be simultaneously bound to PCNA.

5.2 Solenoid poly-TPRs have unique features

The Protein Data Bank's repository of tandem repeat proteins (TPRs) is primarily populated by structures of short [2-4] tandem repeats. Consequently, only a limited number of structures are currently available for tube-forming solenoid TPRs bound to long peptides. Among these structures, those from TSK, Ctr9, and Scc4 exhibit different domain organizations and modes of peptide binding, making them interesting for comparison. Table 5-1 summarizes the characteristics of these structures, as shown in Figure 5.1.

The presence and positioning of a significantly long α-helix near the center of certain TPRs can play a crucial role in the overall folding of the solenoid. For instance, in TSK, the long helix provides rigidity, aligning the center and the C-terminal lobes of the tube. Interestingly, it also reverses the direction of the solenoid back towards the center lobe (Figure 5.1A).). In Scc4, the long helix is positioned more centrally and appears to contribute to the straightness of the TPR tube and peptide in the central region (Figure 5.1A) unlike Ctr9, where the absence of this helix coincides with a bending of the tube and peptide (not shown).

Furthermore, the longer loops between TPR motifs in Ctr9 allow for overall stretching of the tube and create multiple surface openings along its length (Figure 5.1C).). This is in contrast to TSK or Scc4, which expose their underlying peptides only at the terminal regions. In the structure of

Scc4, these openings can serve as binding grooves for additional peptides that can interact with both the TPR and the tube-binding peptide. Although the exact role of the β-sheets in these structures is unclear, in Scc4, two small β-sheets form on the TPR with the participation of the peptide, contributing to regional stability. Moreover, a third small β-sheet connects the C-terminal region of the peptide to another segment of the same peptide in Scc4 (Figure 5.1B). Lastly, the N-terminal regions of Scc4 and Ctr9 can bind peptides on both their inner and outer surfaces, extending the range of peptide recognition by 25-80 amino acids.

Table 5-1 – Comparison of the characteristics of tube-forming solenoid tetratricopeptide domains.

TPR	Solenoid Arrangement			Topology		
module	Repeats	Handedness	Packing	Special Features	Tube Structure	Target Binding Area
TSK (H3.1)	11	right-handed	condensed	extra-long TPR7 α-helix solenoid reversed at TPR8	covered at centre, open on N- and C-terminal regions on opposite faces	inner tube surface
Ctr9 (Paf1)	21	right-handed	stretched	Absence of long TPR, bent at the centre of the N-terminal β-sheet	Alternation of covered (2) and open (3) surfaces in 5 regions	inner tube + outer tube surface at N-terminal region secondary peptide binding
Scc4 (Scc2)	13	left-handed	condensed	long TPR8 α-helix β-sheet between TPR and peptide in TPR6 /7 region β-sheet between TPR N-t loop and peptide	covered at centre, open on N- and C-terminal regions on opposite faces	inner tube + outer tube surface at N-terminal region

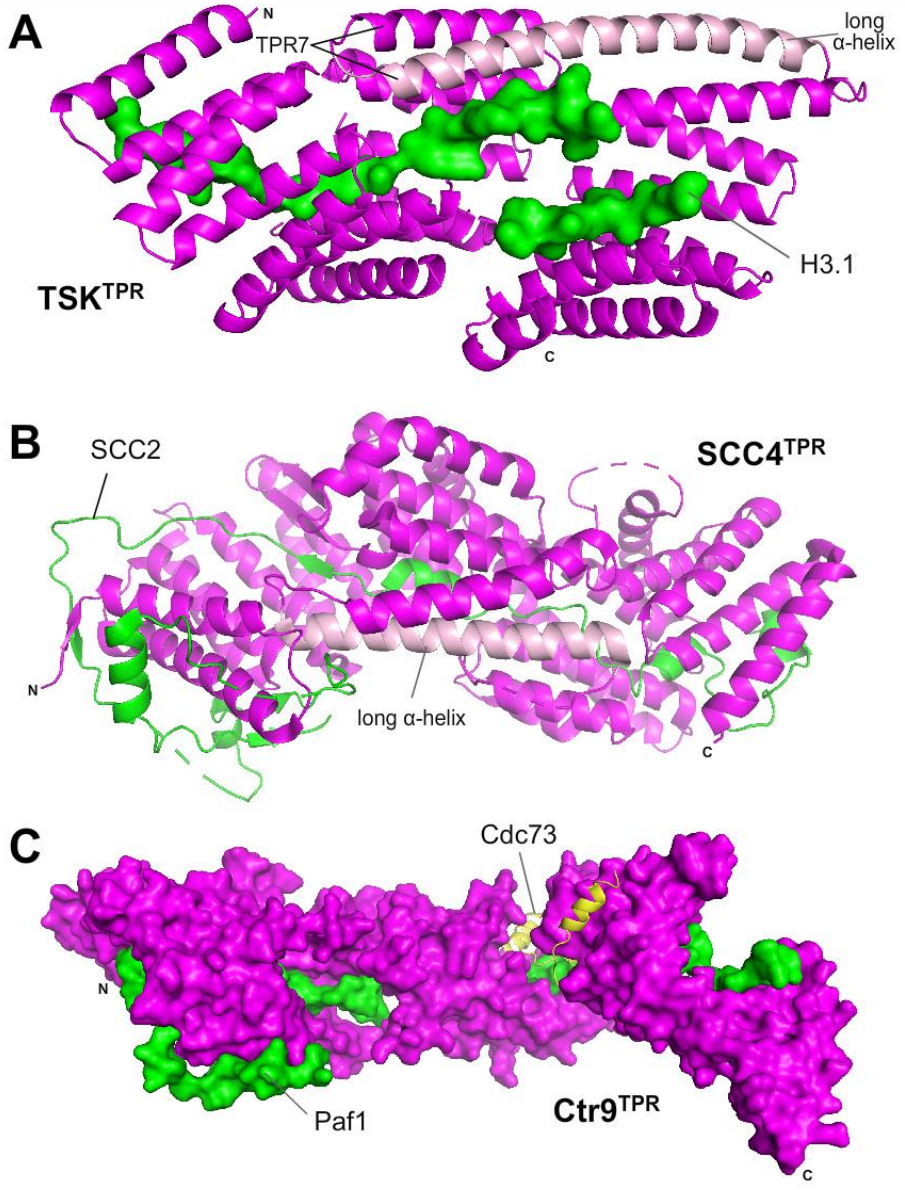


Figure 5.1 – TPRs bind long peptides in various configurations. The TPR domains of TSK [PDB ID: 7T7T (ref. *87*)], SCC4 [PDB ID: 4XDN (ref. *112*)] and Ctr9 [PDB ID: 6AF0 (ref. *113*)] are shown in magenta and their binding peptides are shown in green. For TSK and SCC4 the TPR domains are shown as cartoons revealing the orientation of the H3.1 (surface) or SCC2 (cartoon) peptides. The SCC2 peptide is shown in cartoon revealing secondary structures formed by this peptide. The long α-helices in TSK and SCC4 are shown distinctly in light pink. Ctr9 is shown as surface representation revealing openings in within its surface. The Cdc73 peptide is shown as yellow cartoon revealing the orientation of this secondary peptide relative to the opening in SCC4 surface and its underlying Paf1 peptide.

5.3 Tandem alpha-helical domains – a growing histone reader class

Histone reader domains play crucial roles in epigenetic signalling. So far, only 14-3-3, ARD, TPR, and BRD domains have been identified histone PTM readers featuring α-helical motifs as their sole secondary structure. These domains exhibit diverse footprints and strategies in binding peptides and their associated PTMs (Table 5-2). The extensive footprints of TPR domains broaden their ability to detect multiple PTMs on the same target, while concurrently narrowing the range of target peptides recognized by these proteins. This is in contrast to domains that require smaller footprints (e.g., ARDs and BRDs). However, the ability of these domains to utilize additional appendages or form dimers may enhance binding diversity. Furthermore, the arrangement of these domains alongside other reader/writer/eraser domains on a given chromatin remodeler likely determines their specific ability to read modified residues. Furthermore, the nuclear localization may play a role in determining their targets.

To our knowledge, we have discovered the only poly-TPR domain (TSK) that functions as a histone reader. The extensive footprint of TSK in reading the entire histone H3 N-terminal region limits its target to a single histone thus far. Moreover, the architecture of TSK's binding pockets further restricts its ability to recognize the histone H3.1 variant specifically. However, our research has demonstrated that this long footprint can simultaneously recognize multiple PTMs deposited on the N-terminus of histone H3.1. This unique mechanism is observed in tandem repeat domains, where forming a continuous tube-like binding channel allows for coverage of a large surface area. Compared to ARD domains, the solenoid structure of TPRs protects larger peptide regions, although the significance of this observation is not yet fully understood. Chromatin remodelling processes involving histone movements, such as incorporation or eviction, will likely benefit from comprehensive contact with a histone peptide in one or more regions, as several sidechains

pointing in opposite directions can engage the reader. Additionally, this type of contact may serve as a mechanism to inhibit histone modification. Since the identification of HP1 as a reader of H3K9me3 (*214*) over 20 years ago, the recent discovery of a poly-TPR binding to histone H3.1 serves as a reminder that there may be other yet-to-be-identified poly-TPR domains or novel folds that can function as histone readers.

Table 5-2 – Targets and PTM-reading profiles of α-helical histone reader domains.

Reader	Arrangement	Footprint Region	Target histones	PTM-reading profile
14-3-3*	single protein (can form dimer)	small (1-10)	H3 H4	S10ph, (S10ph+K9ac/me2), (S10ph+K14ac), S28ph (ref. *85*)
ARD^{TSK}	tandem repeated motifs 5 repeats	medium (1-20)	H4	(K20† + K16†/ac) (ref. *86*)
TPR^{TSK}	tandem repeated motifs 11 repeats	extensive (20+)	H3.1	K36†, T32†, K27†/ac/cr, R26me1, K23†/me1/ac, T22†, T6† K4†/me1, T3†
BRD*	single protein (can form dimer)	small	H2A H2B H3 H4 H1	acetylated K5/9/13/119/125/127/129 acetylated K5/11/12/15/16/20//116/120/125 acetylated K4/9/14/18/23/27/36/37/56/79/122 acetylated K58/12/16/20/91 acetylated K16/20/21/22/25/31/33/45/205/211/212/216/217/218 (reviewed in ref.*117*)

* refers to a protein family.
() indicates a combination of PTMs in which the bolded PTM is required for binding.
† indicates an unmodified amino acid.

5.4 PTM reading profile of TSK complements its cell-cycle-specific role

Our studies on the impact of PTMs deposited on histone H3.1 on TPR^{TSK} binding revealed that several of these marks have inhibitory effects while some are permissive. These marks are related to different processes, although from the categorization of these signals (Figure 5.2) emerges a profile for the binding of TPR^{TSK} concerning chromatin structure and cell cycle. H3.1 PTMs negatively impacting TPR^{TSK} binding predominantly reside in condensed chromatin and peak during mitosis. Conversely, PTMs that are compatible with or promote the binding of TPR^{TSK} are linked to chromatin restoration, opening and predominantly deposited in regions of the genome undergoing active DNA transactions such as gene expression and DDR.

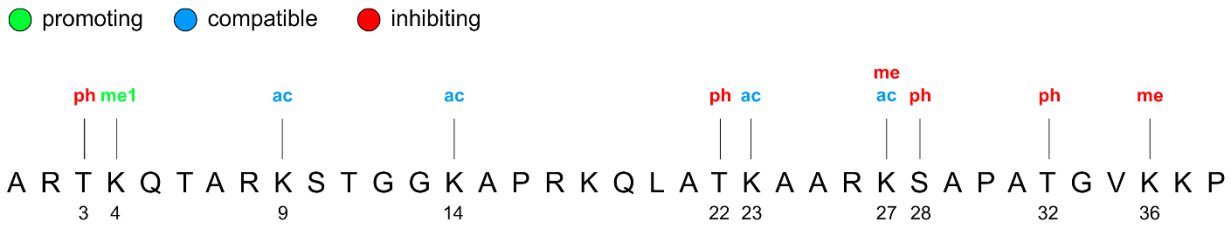


Figure 5.2 – TPR^{TSK} binds H3.1 during the S-phase of the cell cycle. The sequences of the N-terminal region of histone H3.1 are shown. PTMs related to DDR are marked above their corresponding residues and are colour coded according to the indicated impact on TPR^{TSK} binding.

5.5 A temporal role for H3K36me3 in HR

H3K36me3 is important for the recruitment of the mismatch repair (MMR) and NHEJ machineries via 53BP1 (*215*). This PTM is also required for the proper formation of RPA and RAD51 foci that are essential steps in HR progression (*157*). This observation posits that TSK and K36me3 must mediate HR within the same time window; however, our data shows that H3.1K36me3 is detrimental to TPR^{TSK} binding (Figure 3.1E). Therefore, it is unclear in what order these non-overlapping events occur. Specifically, whether K36 trimethylation occurs after TSK binds and eventually releases H3.1, or if K36me3 occurs before TSK-binding but is later removed to allow the interaction. Considering the contribution of H3K36me3 to several DDR pathways, the latter scenario is more likely. Such model suggests that H3K36me3 plays a role in the sensing DDR steps, and one could postulate that K36me3 temporarily blocks TSK binding to coordinate other epigenetic events such as pre-recruitment of RPA to which TONSL:MMS22L eventually bind and recruit RAD51 (*65*).

Few studies have looked at the expression patterns of K36me3 demethylases during DNA damage. KDM4A/JMJD2A levels decrease post-damage induction (*57, 216*) while KDM4B/JMJD2B increase instantly or show a delayed pattern of upregulation during DNA-damage response (*217*). These observations suggest that K36me3 demethylases may be key in the proper progression of HR by temporarily removing K36me3, serving as a signal for TONSL recruitment. Finally, methylation of K36 contributes to HR by coordinating cell-cycle progression by delaying (K36me3) or promoting (K36me1) firing at replication origins (*218*).

5.6 The TPR^{TSK} C-lobe: An add-on compatible binding model

The poly-TPR domain of TSK exhibits an intriguing tube-like structure, resembling a "penne" pasta shape. It features a fully closed mid-section with two scoop-like overhangs at the opposing ends. The C-terminal lobe of the TPR tube is considerably longer and wider than the N-terminal lobe. Within this lobe, two segments of H3.1 bind to the opposite walls (Figure 3.3A and B). Although crystal structures of the complex have revealed the A1-S10 and Q19-H39 residues of H3.1, residues 11-18 have not been observed through x-ray crystallography. Additionally, our binding studies, where we manipulated this region with PTMs, showed no difference in binding (Figure 3.3E). This suggests that the TPR C-lobe does not bind to this particular region.

Apart from allowing the H3.1 segment to remain flexible, the open surface and the width of the channel imply that the TPR C-lobe might have evolved to expose a selectively chosen sequence of modifiable histone residues to other readers if other domains or segments of TSK do not fold back onto this opening. These readers could potentially dock on its surface, form a roof over its capless channel, and engage with the H3.1 residues underneath. In this model, the additional histone reader or writer is likely to be a small domain, such as a plant homeodomain (PHD), or a domain with the majority of its surface away from the histone reading site. This arrangement would enable it to fit into the TPR C-lobe channel effectively.

The crystal structure of the BRPF1 bromodomain bound to an H3.1K14ac peptide (PDB ID 5FFV) shows that its bromodomain binds H3.1 10-17 and reads K14ac. Studies investigating the role of H3K14ac in yeast show that this mark, together with K9ac, is enriched post UV irradiation and that manipulating the histone acetylase GCN5 results in inefficient nucleotide excision repair (NER) and related mechanism (*219*). Furthermore, a yeast ATP-dependant RSC (Remodels the Structure of Chromatin) complex, which has an important role in nucleosome repositioning and

histone eviction (*220*), was shown to have increased stability on nucleosomes with H3K14ac PTMs (*221*). Accordingly, the ATPase component of RSC, Sth1, harbours a C-terminal bromodomain (*222*). In contrast, other K14 readers, such as the double PHD finger of MORF (K14ac/cr reader) contact a much larger sequence of H3.1 (a.a. 1-17) with conflicting footprint and interestingly induce these residues to form an α-helix of (*223, 224*). The accessibility of the K14 region further suggests that larger PTMs such as ubiquitin (ub) may be deposited in this region without impacting TPR^{TSK} binding. K14ub is catalyzed by the Cullin4 Ub-ligase component of the Clr4 methyltransferase complex (CLRC) complex and is a requirement for methylation of K9 by Clr4, and this combination PTM is important for heterochromatin formation (*225*).

While investigating readers of H3.1 in the 10-18 region through structural analysis, including BRD^{BRPF1}, we discovered that bromodomains are well-suited for binding TPR^{TSK}:H3.1 due to their elongated structures, with the majority of their surface perpendicular and pointing away from the H3 reading site (Figure 1.1D). Our attempts to model the bromodomain of BRPF1 revealed that while this domain can insert its H3 binding site into the TPR C-lobe channel and reach H3 without clashes, the curvature of the BRD-bound peptide prevents proper alignment between H3 peptides of the two structures on one end. The presence of proline at position 16 (H3:BRD) introduces a turn that facilitates the connection between R17 (H3:BRD structure) and K18 backbones (H3:TPR structure). However, on the opposite end, the residues near T11 on H3:BRD would require a turn in the opposite direction. Interestingly, two consecutive glycine residues at positions 12 and 13 would provide the necessary flexibility to reorientate histone H3.1.

By aligning and crystal structures using common H3.1 sequences, we also modelled a second bromodomain, BRD^{Brtd}, which reads K18ac, on top of the TPR C-lobe (Figure 5.3). The structure of this BRD included H3 residues A15 to T22, resulting in an overlap of five amino acids (K18 to

T22) on one end while leaving the T11-K14 region unmodeled between both structures. In this model, BRD^{Brtd} forms an arched cap over the open channel of the TPR, contacting the TPR on both sides of the arc while two segments of H3 pass underneath it. In this manner, BRD^{Brtd} covers nearly all H3 residues in the TPR C-lobe channel, and its αZ-αA loop appears to come into close contact with the sidechain of H3 Q5 and the backbone of R8. Notably, in this model, K18ac points upward away from the TPR and into the deep hydrophobic cleft of the bromodomain. Although we did not observe significant clashes in this model, further robust evaluation using computational techniques like molecular dynamics and modelling of additional readers could provide insights into whether this "add-on" reading is an intended feature for the open surface of TPR^{TSK}.

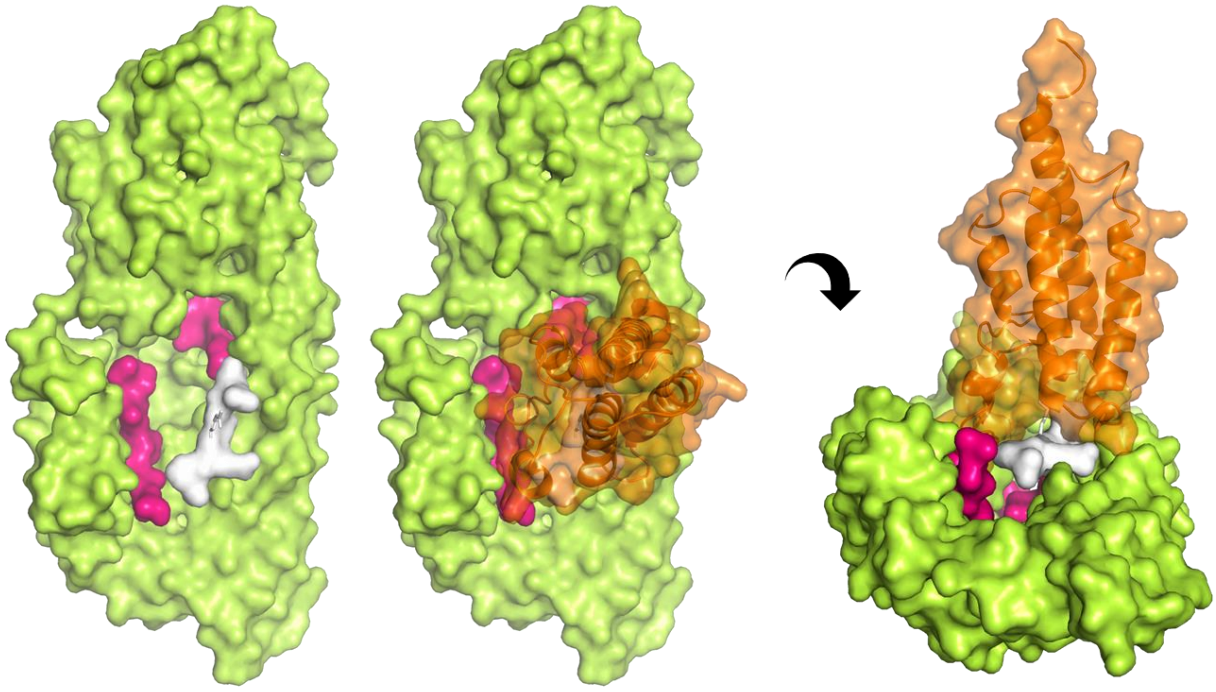


Figure 5.3 – A model for simultaneous reading of H3 by TPR and BRD. **(left)** top view of the structure of TPR^{TSK} [green] showing H3 N-terminal region [magenta and white] within the channel of its C-terminal TPR lobe. The purple H3 segments correspond to residues co-crystalized with TPR^{TSK} from our studies while the white H3 segments were co-crystalized with BRD^{Brdt} (PDB ID 2WP1). **(centre)** the same view as the left panel but with the addition of BRD^{Brdt} in cartoon and transparent surface. **(right)** side view of the complex from the TPR C-lobe channel viewpoint.

5.7 TPR and ARD anchor on different surfaces of the nucleosome

Saredi et al. solved the crystal structure of the TONSL ankyrin repeat domain in complex with a histone H3-H4 dimer (*86*) (Figure 1.3B). This model can be aligned with H3-H4 within nucleosome structures to provide insight into the NCP:ARD complex. In the ARD crystal structure, H4 residues 12-100 were solved and reveal the orientation of the H4 tail (residues12-23) with respect to the structured H4 core. Overlaying the NCP with this structure using H3-H4, without any adjustments, results in a perfect alignment of the histones and produces no clashes with between ARD^{TONSL} and the nucleosome. The model suggests that the ARD lands on the NCP perpendicular to its top surface where the histone cores are exposed and is not found in the vicinity of DNA (Figure 5.4). As such, the ARD helps extend the H4 tail upward and away from the NCP surface. Furthermore, in this model, Lys521 from $TONSL^{ARD}$ appears to make aliphatic contacts with H2A residue Glu91.

The H3.1:TPR^{TSK} structure from our studies shows that H3.1 P38 binds the surface of TPR while the backbone of H39 hangs over the TPR surface with no electron density for its sidechain (Figure 3.1B and C). The crystal structure of the nucleosome show that the H3 H39 sidechain is embedded in the NCP disc while P38 sticks out from between the DNA gyres at the edge of the NCP disc and is exposed to the nucleoplasm (Figure 5.5). This observation begged us to model the TPR domain on the NCP to gain insight into the possible structure of this complex and its relationship with the ARD on the same nucleosome (Figure 5.4). Aligning the NCP with TPR by overlaying H3 P38 and the backbone of H39 produces surface clashes between the DNA molecule and the first α-helix of TPR1 as several atoms between the structures co-occupy the same space (<2 Å); however, tilting the TPR tube, which produces a minor 2.0 Å movement of the structure with respect to H39/P38,

resolves these clashes. This model, interestingly, places the N-terminal lobe of the TPR tube in such a way that the top DNA gyre occupies the open space above it while the front of N-lobe sits on the bottom gyre (Figure 5.4). Furthermore, within the surface of the TPR N-lobe, several arginine side chains face the DNA backbone. Future cryo-EM studies should provide insights in the assembly of the TONSL/NCP complex.

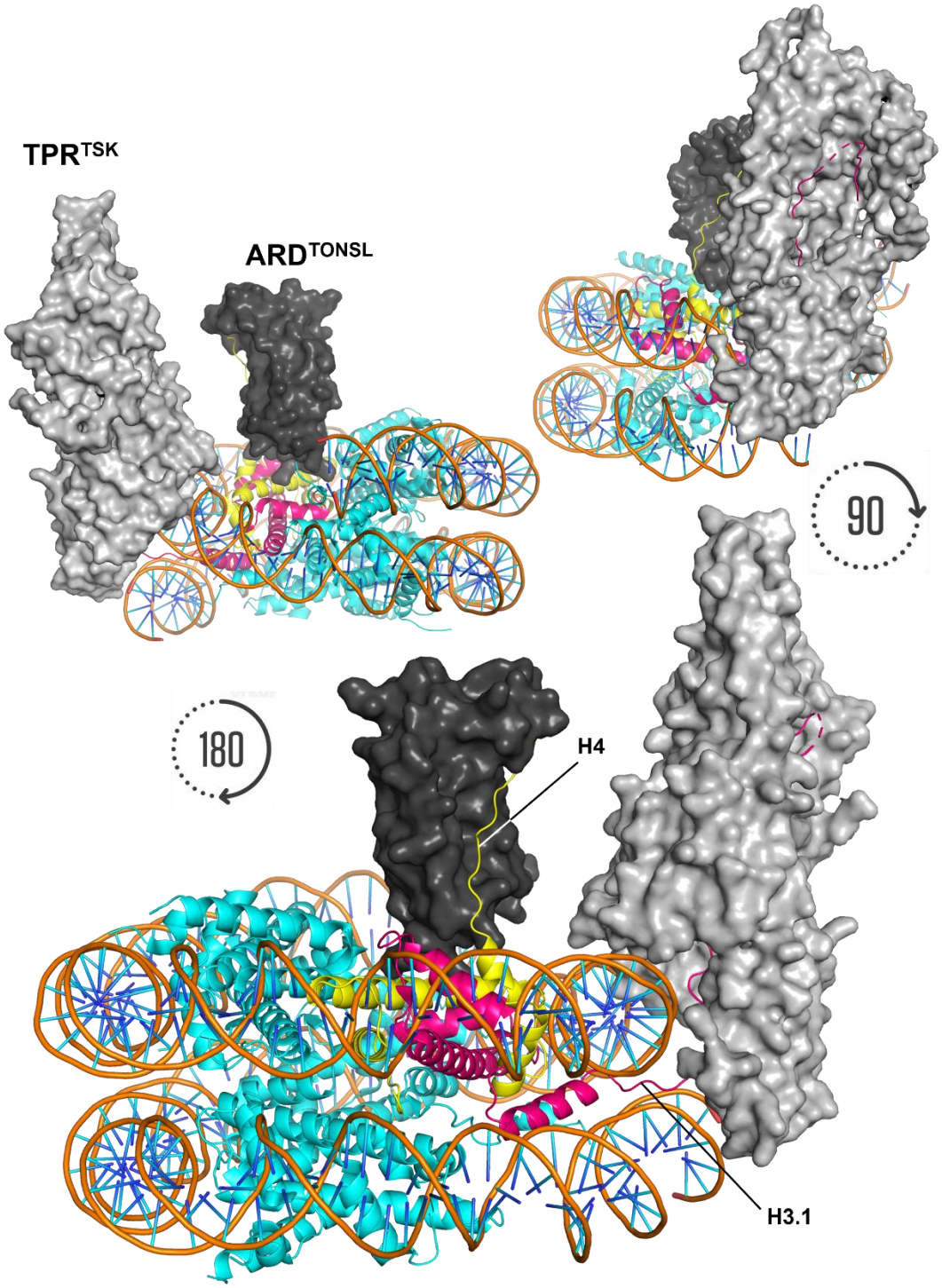


Figure 5.4 – ARD and TPR bind the NCP on nearby surfaces. Model was generated by aligning crystal structures of the NCP (PDB ID: 3KWQ), ARD^{TONSL}:H4 (5JA4), and TPR^{TSK}:H3.1 (7T7T). **[bottom]** A view of the complex showing TPR in grey, ARD in black, H3 in magenta, H4 in yellow, and H2A/H2B in cyan. The same complex is shown rotated around the axis in the centre of the NCP **[top right**] at 90º or **[top left]** at 180º.

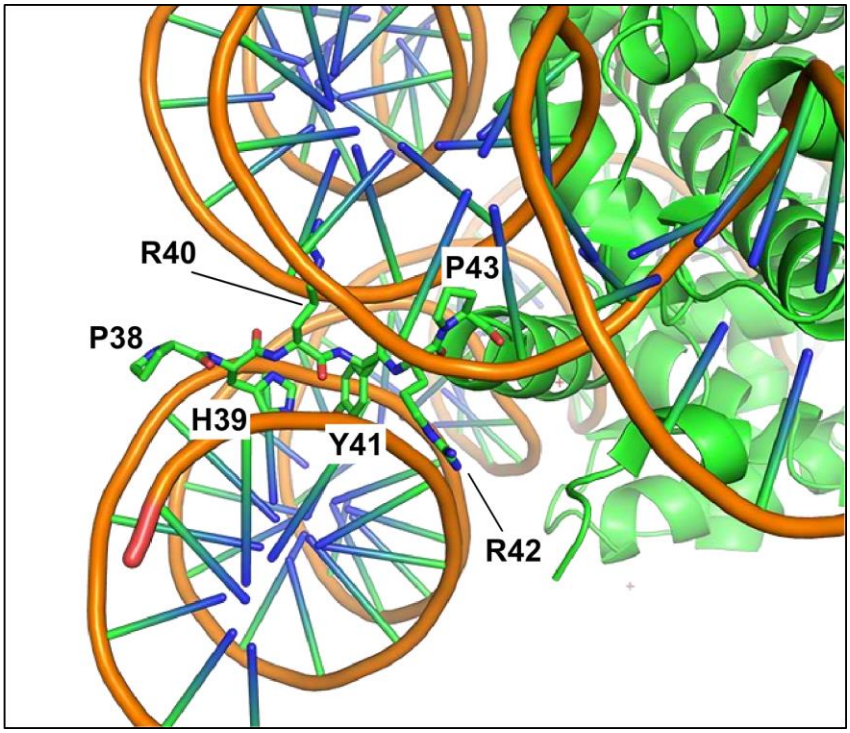


Figure 5.5 – H3.1 residues 39-45 are entrapped by the nucleosome disc. Cartoon demonstration of the nucleosome core particle from PDB ID: 3KWQ. The backbone of the DNA molecule is shown in orange and histone folds are shown as green α-helices. The region before (N-terminal to) the first α-helix of histone H3.1 leading to the unstructured histone tail is shown as sticks and labelled with corresponding H3.1 amino acids. The lower DNA pass is exiting the nucleosome disc near H39-R42 region. Figure generated by Pymol software (Schrödinger).

5.8 Limitations and Future Directions

My studies on the PCNA: ATXR5/6 complex revealed different stoichiometry between complexes formed using only the PIP motif or the full-length ATXR5/6 protein. The crystal structure clearly shows that three molecules of PIP can bind PCNA simultaneously, but analytical ultracentrifugation studies show that only two molecules of the full-length enzymes bind to a PCNA trimer. Therefore, our studies were limited in identifying the mechanisms that explain this difference. This study could benefit from additional structural studies, such as crystal structures that include the PHD, SET, or both domains. Alternatively, cryo-EM studies of the full-length ATXR5/6:PCNA complex could provide a more comprehensive understanding of the structural underpinning controlling the stoichiometry of the complex.

Furthermore, while I demonstrated that H3.1K27 monomethylation by ATXR5/6 is regulated by PCNA-binding, the chronological significance of this finding remains unresolved. It is known that ATXR5/6 restore H3.1K27me1 marks behind the replication fork, and this activity is important for the “chromatin memory” post-replication. We also showed that this mark deposited by ATXR5/6 is crucial for inhibiting TSK binding and preventing spurious HR. However, the mechanism by which ATXR5/6 senses the need to withhold H3.1K27 methylation upon DNA damage remains to be elucidated. Given our findings regarding PCNA's inhibition of these enzymes, it is tempting to suggest that PCNA-binding serves as this signal. Yet, the next line of inquiry should focus on determining the mechanism that triggers PCNA-binding. Alternatively, demethylation of H3.1K27me1 resulting from DDR could be the missing link and should be investigated.

Solving the structure of my TONSL$^{TPR\text{-}ARD}$:NCP complex model would be invaluable in understanding the molecular details of how TONSL engages with the nucleosome, particularly regarding the contribution of DNA to the binding of the complex. Additionally, while I

characterized significant details within the TPR footprint on histone H3.1 using X-ray crystallography, these findings could be put into a broader context by studying the participation of the rest of TONSL using cryo-EM. Such studies performed on the complex of full-length TONSL:NCP could inform us whether ARD and TPR interact with each other or with TONSL's LRR domain, which binds MMS22L. These interactions may serve as a sensing mechanism that signals to TONSL-MMS22L that the complex has properly engaged with the nucleosome via both histones H3.1 and H4 before subsequent recruitment of RAD51.

References

1. C. P. Bird, B. E. Stranger, E. T. Dermitzakis, Functional variation and evolution of non-coding DNA. *Current Opinion in Genetics & Development* **16**, 559-564 (2006).
2. K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, T. J. Richmond, Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-260 (1997).
3. L. Shemilt *et al.*, Karyotyping Human Chromosomes by Optical and X-Ray Ptychography Methods. *Biophysical Journal* **108**, 706-713 (2015).
4. R. M. Martin, M. C. Cardoso, Chromatin condensation modulates access and binding of nuclear proteins. *Faseb j* **24**, 1066-1072 (2010).
5. G. Li, D. Reinberg, Chromatin higher-order structures and gene regulation. *Current Opinion in Genetics & Development* **21**, 175-186 (2011).
6. G. Li, P. Zhu, Structure and organization of chromatin fiber in the nucleus. *FEBS Letters* **589**, 2893-2904 (2015).
7. N. Gilbert *et al.*, Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. *Cell* **118**, 555-566 (2004).
8. N. Saksouk, E. Simboeck, J. Déjardin, Constitutive heterochromatin formation and transcription in mammals. *Epigenetics & Chromatin* **8**, 3 (2015).
9. Y. Zhao, B. A. Garcia, Comprehensive Catalog of Currently Documented Histone Modifications. *Cold Spring Harbor Perspectives in Biology* **7**, (2015).
10. H.-M. Herz, A. Garruss, A. Shilatifard, SET for life: biochemical activities and biological functions of SET domain-containing proteins. *Trends in Biochemical Sciences* **38**, 621-639 (2013).
11. J. Xu, S. Richard, Cellular pathways influenced by protein arginine methylation: Implications for cancer. *Molecular Cell* **81**, 4357-4368 (2021).
12. S. D. Furdas, S. Kannan, W. Sippl, M. Jung, Small Molecule Inhibitors of Histone Acetyltransferases as Epigenetic Tools and Drug Candidates. *Archiv der Pharmazie* **345**, 7-21 (2012).
13. D. Rossetto, N. Avvakumov, J. Côté, Histone phosphorylation: a chromatin modification involved in diverse nuclear events. *Epigenetics* **7**, 1098-1108 (2012).
14. E. Dimitrova, A. H. Turberfield, R. J. Klose, Histone demethylases in chromatin biology and beyond. *EMBO Rep* **16**, 1620-1639 (2015).
15. E. Seto, M. Yoshida, Erasers of Histone Acetylation: The Histone Deacetylase Enzymes. *Cold Spring Harbor Perspectives in Biology* **6**, (2014).
16. T. Yuan *et al.*, Inhibition of Ubiquitin-Specific Proteases as a Novel Anticancer Therapeutic Strategy. *Frontiers in Pharmacology* **9**, (2018).
17. R. S. Gil, P. Vagnarelli, Protein phosphatases in chromatin structure and function. *Biochim Biophys Acta Mol Cell Res* **1866**, 90-101 (2019).
18. M. Podhorecka, A. Skladanowski, P. Bozko, H2AX Phosphorylation: Its Role in DNA Damage Response and Cancer Therapy. *J Nucleic Acids* **2010**, (2010).
19. Y. Xu *et al.*, Histone H2A.Z controls a critical chromatin remodeling step required for DNA double-strand break repair. *Mol Cell* **48**, 723-733 (2012).
20. H. Stroud *et al.*, Genome-wide analysis of histone H3.1 and H3.3 variants in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences* **109**, 5370-5375 (2012).

21. K. Yoda *et al.*, Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro. *Proceedings of the National Academy of Sciences* **97**, 7266-7271 (2000).
22. A. Laugesen, J. W. Højfeldt, K. Helin, Role of the Polycomb Repressive Complex 2 (PRC2) in Transcriptional Regulation and Cancer. *Cold Spring Harb Perspect Med* **6**, (2016).
23. Z. Liu *et al.*, SUMOylated PRC1 controls histone H3.3 deposition and genome integrity of embryonic heterochromatin. *The EMBO Journal* **39**, e103697 (2020).
24. Y. Jacob *et al.*, ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. *Nature Structural & Molecular Biology* **16**, 763 (2009).
25. Y. Jacob *et al.*, Selective methylation of histone H3 variant H3.1 regulates heterochromatin replication. *Science* **343**, 1249-1253 (2014).
26. M. Manohar *et al.*, Acetylation of histone H3 at the nucleosome dyad alters DNA-histone binding. *J Biol Chem* **284**, 23312-23321 (2009).
27. F. Hans, S. Dimitrov, Histone H3 phosphorylation and cell division. *Oncogene* **20**, 3021-3027 (2001).
28. J. Xie *et al.*, Histone H3 Threonine Phosphorylation Regulates Asymmetric Histone Inheritance in the Drosophila Male Germline. *Cell* **163**, 920-933 (2015).
29. I. V. Mavragani, Z. Nikitaki, S. A. Kalospyros, A. G. Georgakilas, Ionizing Radiation and Complex DNA Damage: From Prediction to Detection Challenges and Biological Significance. *Cancers (Basel)* **11**, (2019).
30. M. S. Cooke, M. D. Evans, M. Dizdaroglu, J. Lunec, Oxidative DNA damage: mechanisms, mutation, and disease. *The FASEB Journal* **17**, 1195-1214 (2003).
31. R. Hakem, DNA-damage repair; the good, the bad, and the ugly. *Embo j* **27**, 589-605 (2008).
32. M. Crewe, R. Madabhushi, Topoisomerase-Mediated DNA Damage in Neurological Disorders. *Frontiers in Aging Neuroscience* **13**, (2021).
33. A. A. Guerrero *et al.*, Centromere-localized breaks indicate the generation of DNA damage by the mitotic spindle. *Proceedings of the National Academy of Sciences* **107**, 4159-4164 (2010).
34. Y. C. Huang, W. Yuan, Y. Jacob, The Role of the TSK/TONSL-H3.1 Pathway in Maintaining Genome Stability in Multicellular Eukaryotes. *Int J Mol Sci* **23**, (2022).
35. T. E. Wilson, S. Sunder, Double-strand breaks in motion: implications for chromosomal rearrangement. *Curr Genet* **66**, 1-6 (2020).
36. V. Tiwari, D. M. Wilson, 3rd, DNA Damage and Associated DNA Repair Defects in Disease and Premature Aging. *Am J Hum Genet* **105**, 237-257 (2019).
37. S. Burdak-Rothkamm, W. Y. Mansour, K. Rothkamm, DNA Damage Repair Deficiency in Prostate Cancer. *Trends in Cancer* **6**, 974-984 (2020).
38. L. C. Burrage *et al.*, Bi-allelic Variants in TONSL Cause SPONASTRIME Dysplasia and a Spectrum of Skeletal Dysplasia Phenotypes. *The American Journal of Human Genetics* **104**, 422-438 (2019).
39. H. R. Chang *et al.*, Hypomorphic Mutations in TONSL Cause SPONASTRIME Dysplasia. *The American Journal of Human Genetics* **104**, 439-453 (2019).
40. D. Necchi *et al.*, Defective DNA repair and increased chromatin binding of DNA repair factors in Down syndrome fibroblasts. *Mutat Res* **780**, 15-23 (2015).

41. M. S. Meyn, Ataxia-telangiectasia, cancer and the pathobiology of the ATM gene. *Clinical Genetics* **55**, 289-304 (1999).
42. H. E. Krokan, M. Bjørås, Base excision repair. *Cold Spring Harb Perspect Biol* **5**, a012583 (2013).
43. J. A. Marteijn, H. Lans, W. Vermeulen, J. H. J. Hoeijmakers, Understanding nucleotide excision repair and its roles in cancer and ageing. *Nature Reviews Molecular Cell Biology* **15**, 465-481 (2014).
44. S. Qiu, J. Huang, MRN complex is an essential effector of DNA damage repair. *J Zhejiang Univ Sci B* **22**, 31-37 (2021).
45. J. H. Lee, A. A. Goodarzi, P. A. Jeggo, T. T. Paull, 53BP1 promotes ATM activity through direct interactions with the MRN complex. *Embo j* **29**, 574-585 (2010).
46. J. Kobayashi *et al.*, Histone H2AX participates the DNA damage-induced ATM activation through interaction with NBS1. *Biochem Biophys Res Commun* **380**, 752-757 (2009).
47. N. Mailand *et al.*, RNF8 ubiquitylates histones at DNA double-strand breaks and promotes assembly of repair proteins. *Cell* **131**, 887-900 (2007).
48. M. S. Huen *et al.*, RNF8 transduces the DNA-damage signal via histone ubiquitylation and checkpoint protein assembly. *Cell* **131**, 901-914 (2007).
49. R. Kalb, D. L. Mallery, C. Larkin, J. T. Huang, K. Hiom, BRCA1 is a histone-H2A-specific ubiquitin ligase. *Cell Rep* **8**, 999-1005 (2014).
50. H. H. Y. Chang, N. R. Pannunzio, N. Adachi, M. R. Lieber, Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nature Reviews Molecular Cell Biology* **18**, 495-506 (2017).
51. J. Her, S. F. Bunting, How cells ensure correct repair of DNA double-strand breaks. *Journal of Biological Chemistry* **293**, 10502-10511 (2018).
52. A. N. Blackford, M. Stucki, How Cells Respond to DNA Breaks in Mitosis. *Trends in Biochemical Sciences* **45**, 321-331 (2020).
53. Z. Shao *et al.*, Persistently bound Ku at DNA ends attenuates DNA end resection and homologous recombination. *DNA Repair (Amst)* **11**, 310-316 (2012).
54. Q. Wu *et al.*, Understanding the structure and role of DNA-PK in NHEJ: How X-ray diffraction and cryo-EM contribute in complementary ways. *Prog Biophys Mol Biol* **147**, 26-32 (2019).
55. T. Lei, S. Du, Z. Peng, L. Chen, Multifaceted regulation and functions of 53BP1 in NHEJ-mediated DSB repair (Review). *Int J Mol Med* **50**, 90 (2022).
56. J. M. Daley, P. Sung, 53BP1, BRCA1, and the choice between recombination and end joining at DNA double-strand breaks. *Mol Cell Biol* **34**, 1380-1388 (2014).
57. F. A. Mallette *et al.*, RNF8- and RNF168-dependent degradation of KDM4A/JMJD2A triggers 53BP1 recruitment to DNA damage sites. *Embo j* **31**, 1865-1878 (2012).
58. K. Nakamura *et al.*, H4K20me0 recognition by BRCA1–BARD1 directs homologous recombination to sister chromatids. *Nature Cell Biology* **21**, 311-318 (2019).
59. E. Rass, S. Willaume, P. Bertrand, 53BP1: Keeping It under Control, Even at a Distance from DNA Damage. *Genes (Basel)* **13**, (2022).
60. S. Pellegrino, J. Michelena, F. Teloni, R. Imhof, M. Altmeyer, Replication-Coupled Dilution of H4K20me2 Guides 53BP1 to Pre-replicative Chromatin. *Cell Rep* **19**, 1819-1831 (2017).
61. X. Li, W.-D. Heyer, Homologous recombination in DNA repair and DNA damage tolerance. *Cell Research* **18**, 99-113 (2008).

62. G. Saredi *et al.*, H4K20me0 marks post-replicative chromatin and recruits the TONSL-MMS22L DNA repair complex. *Nature* **534**, 714-718 (2016).
63. M. Isono *et al.*, BRCA1 Directs the Repair Pathway to Homologous Recombination by Promoting 53BP1 Dephosphorylation. *Cell Rep* **18**, 520-532 (2017).
64. P. Baumann, S. C. West, Role of the human RAD51 protein in homologous recombination and double-stranded-break repair. *Trends Biochem Sci* **23**, 247-251 (1998).
65. W. Piwko *et al.*, The MMS22L-TONSL heterodimer directly promotes RAD51-dependent recombination upon replication stress. *Embo j* **35**, 2584-2601 (2016).
66. L. O'Donnell *et al.*, The MMS22L-TONSL complex mediates recovery from replication stress and homologous recombination. *Mol Cell* **40**, 619-631 (2010).
67. B. Moore, V. Perez, F. Carlson, Physiological and biochemical aspects of nervous integration. *Englewood Cliffs, NJ: Prentice-Hall*, 343-359 (1967).
68. S. Winter, W. Fischle, C. Seiser, Modulation of 14-3-3 interaction with phosphorylated histone H3 by combinatorial modification patterns. *Cell Cycle* **7**, 1336-1342 (2008).
69. A. Aitken, 14-3-3 proteins: a historic overview. *Semin Cancer Biol* **16**, 162-172 (2006).
70. A. Aitken, S. Howell, D. Jones, J. Madrazo, Y. Patel, 14-3-3 alpha and delta are the phosphorylated forms of raf-activating 14-3-3 beta and zeta. In vivo stoichiometric phosphorylation in brain at a Ser-Pro-Glu-Lys MOTIF. *J Biol Chem* **270**, 5706-5709 (1995).
71. K. L. Pennington, T. Y. Chan, M. P. Torres, J. L. Andersen, The dynamic and stress-adaptive signaling hub of 14-3-3: emerging mechanisms of regulation and context-dependent protein-protein interactions. *Oncogene* **37**, 5587-5604 (2018).
72. M. Lalle, A. M. Salzano, M. Crescenzi, E. Pozio, The Giardia duodenalis 14-3-3 protein is post-translationally modified by phosphorylation and polyglycylation of the C-terminal tail. *J Biol Chem* **281**, 5137-5148 (2006).
73. M. L. Knetsch *et al.*, Isolation of a Dictyostelium discoideum 14-3-3 homologue. *Biochim Biophys Acta* **1357**, 243-248 (1997).
74. P. F. Erickson, B. W. Moore, Investigation of the axonal transport of three acidic, soluble proteins (14-3-2, 14-3-3, and S-100) in the rabbit visual system. *J Neurochem* **35**, 232-241 (1980).
75. P. F. Boston, P. Jackson, R. J. Thompson, Human 14-3-3 protein: radioimmunoassay, tissue distribution, and cerebrospinal fluid levels in patients with neurological disorders. *J Neurochem* **38**, 1475-1482 (1982).
76. T. Ichimura, T. Isobe, T. Okuyama, T. Yamauchi, H. Fujisawa, Brain 14-3-3 protein is an activator protein that activates tryptophan 5-monooxygenase and tyrosine 3-monooxygenase in the presence of Ca2+,calmodulin-dependent protein kinase II. *FEBS Lett* **219**, 79-82 (1987).
77. D. Berg, C. Holzmann, O. Riess, 14-3-3 proteins in the nervous system. *Nat Rev Neurosci* **4**, 752-762 (2003).
78. C. MacKintosh, Dynamic interactions between 14-3-3 proteins and phosphoproteins regulate diverse cellular processes. *Biochemical Journal* **381**, 329-342 (2004).
79. V. Obsilova, T. Obsil, The 14-3-3 Proteins as Important Allosteric Regulators of Protein Kinases. *Int J Mol Sci* **21**, (2020).
80. N. Macdonald *et al.*, Molecular basis for the recognition of phosphorylated and phosphoacetylated histone h3 by 14-3-3. *Mol Cell* **20**, 199-211 (2005).

81. A. Zippo, A. De Robertis, R. Serafini, S. Oliviero, PIM1-dependent phosphorylation of histone H3 at serine 10 is required for MYC-dependent transcriptional activation and oncogenic transformation. *Nat Cell Biol* **9**, 932-944 (2007).
82. M. B. Yaffe *et al.*, The Structural Basis for 14-3-3:Phosphopeptide Binding Specificity. *Cell* **91**, 961-971 (1997).
83. A. J. Muslin, J. W. Tanner, P. M. Allen, A. S. Shaw, Interaction of 14-3-3 with Signaling Proteins Is Mediated by the Recognition of Phosphoserine. *Cell* **84**, 889-897 (1996).
84. Y. Zhao *et al.*, Histone phosphorylation integrates the hepatic glucagon-PKA-CREB gluconeogenesis program in response to fasting. *Mol Cell* **83**, 1093-1108.e1098 (2023).
85. S. Winter *et al.*, 14-3-3 proteins recognize a histone code at histone H3 and are required for transcriptional activation. *Embo j* **27**, 88-99 (2008).
86. G. Saredi *et al.*, H4K20me0 marks post-replicative chromatin and recruits the TONSL–MMS22L DNA repair complex. *Nature* **534**, 714-718 (2016).
87. H. Davarinejad *et al.*, The histone H3.1 variant regulates TONSOKU-mediated DNA repair during replication. *Science* **375**, 1281-1286 (2022).
88. Y. Nakamura *et al.*, Crystal Structure of the Human BRD2 Bromodomain: INSIGHTS INTO DIMERIZATION AND RECOGNITION OF ACETYLATED HISTONE H4*. *Journal of Biological Chemistry* **282**, 4193-4201 (2007).
89. Y. Li *et al.*, Molecular Coupling of Histone Crotonylation and Active Transcription by AF9 YEATS Domain. *Mol Cell* **62**, 181-193 (2016).
90. L. Breeden, K. Nasmyth, Similarity between cell-cycle genes of budding yeast and fission yeast and the Notch gene of Drosophila. *Nature* **329**, 651-654 (1987).
91. S. E. Lux, K. M. John, V. Bennett, Analysis of cDNA for human erythrocyte ankyrin indicates a repeated structure with homology to tissue-differentiation and cell-cycle control proteins. *Nature* **344**, 36-42 (1990).
92. L. K. Mosavi, T. J. Cammett, D. C. Desrosiers, Z.-y. Peng, The ankyrin repeat as molecular architecture for protein recognition. *Protein Science* **13**, 1435-1448 (2004).
93. S. Smith, I. Giriat, A. Schmitt, T. de Lange, Tankyrase, a Poly(ADP-Ribose) Polymerase at Human Telomeres. *Science* **282**, 1484-1487 (1998).
94. S. Gorina, N. P. Pavletich, Structure of the p53 Tumor Suppressor Bound to the Ankyrin and SH3 Domains of 53BP2. *Science* **274**, 1001-1005 (1996).
95. P. Michaely, D. R. Tomchick, M. Machius, R. G. W. Anderson, Crystal structure of a 12 ANK repeat stack from human ankyrinR. *The EMBO Journal* **21**, 6387-6396 (2002).
96. T. J. Magliery, L. Regan, Sequence variation in ligand binding sites in proteins. *BMC Bioinformatics* **6**, 240 (2005).
97. Z. Islam, R. S. K. Nagampalli, M. T. Fatima, G. M. Ashraf, New paradigm in ankyrin repeats: Beyond protein-protein interaction module. *International Journal of Biological Macromolecules* **109**, 1164-1173 (2018).
98. R. E. Collins *et al.*, The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. *Nature Structural & Molecular Biology* **15**, 245-250 (2008).
99. M. Tachibana *et al.*, G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev* **16**, 1779-1791 (2002).

100. M. Tachibana *et al.*, Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev* **19**, 815-826 (2005).
101. B. E. Collins, C. B. Greer, B. C. Coleman, J. D. Sweatt, Histone H3 lysine K4 methylation and its role in learning and memory. *Epigenetics & Chromatin* **12**, 7 (2019).
102. E. Duro *et al.*, Identification of the MMS22L-TONSL complex that promotes homologous recombination. *Mol Cell* **40**, 632-644 (2010).
103. B. C. O'Connell *et al.*, A genome-wide camptothecin sensitivity screen identifies a mammalian MMS22L-NFKBIL2 complex required for genomic stability. *Mol Cell* **40**, 645-657 (2010).
104. W. Piwko *et al.*, RNAi-based screening identifies the Mms22L-Nfkbil2 complex as a novel regulator of DNA replication in human cells. *Embo j* **29**, 4210-4222 (2010).
105. R. S. Sikorski, M. S. Boguski, M. Goebl, P. Hieter, A repeating amino acid motif in CDC23 defines a family of proteins and a new relationship among genes required for mitosis and RNA synthesis. *Cell* **60**, 307-317 (1990).
106. C. Scheufler *et al.*, Structure of TPR Domain-Peptide Complexes: Critical Elements in the Assembly of the Hsp70-Hsp90 Multichaperone Machine. *Cell* **101**, 199-210 (2000).
107. P. R. E. Mittl, W. Schneider-Brachert, Sel1-like repeat proteins in signal transduction. *Cellular Signalling* **19**, 20-31 (2007).
108. L. Cerveny *et al.*, Tetratricopeptide repeat motifs in the world of bacterial pathogens: role in virulence mechanisms. *Infect Immun* **81**, 629-635 (2013).
109. M. Jínek *et al.*, The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin alpha. *Nat Struct Mol Biol* **11**, 1001-1007 (2004).
110. N. Zeytuni, R. Zarivach, Structural and Functional Discussion of the Tetra-Trico-Peptide Repeat, a Protein Interaction Module. *Structure* **20**, 397-405 (2012).
111. W. C. Chao *et al.*, Structural Studies Reveal the Functional Modularity of the Scc2-Scc4 Cohesin Loader. *Cell Rep* **12**, 719-725 (2015).
112. S. M. Hinshaw, V. Makrantoni, A. Kerr, A. L. Marston, S. C. Harrison, Structural evidence for Scc4-dependent localization of cohesin loading. *Elife* **4**, e06057 (2015).
113. P. Deng *et al.*, Transcriptional elongation factor Paf1 core complex adopts a spirally wrapped solenoidal topology. *Proc Natl Acad Sci U S A* **115**, 9998-10003 (2018).
114. J. A. Kennison, J. W. Tamkun, Dosage-dependent modifiers of polycomb and antennapedia mutations in Drosophila. *Proc Natl Acad Sci U S A* **85**, 8136-8140 (1988).
115. J. W. Tamkun *et al.*, brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell* **68**, 561-572 (1992).
116. T. Fujisawa, P. Filippakopoulos, Functions of bromodomain-containing proteins and their roles in homeostasis and cancer. *Nature Reviews Molecular Cell Biology* **18**, 246-262 (2017).
117. G. A. Josling, S. A. Selvarajah, M. Petter, M. F. Duffy, The role of bromodomain proteins in regulating gene expression. *Genes (Basel)* **3**, 320-343 (2012).
118. Z. Yang, N. He, Q. Zhou, Brd4 recruits P-TEFb to chromosomes at late mitosis to promote G1 gene expression and cell cycle progression. *Mol Cell Biol* **28**, 967-976 (2008).
119. Y. Xue *et al.*, The human SWI/SNF-B chromatin-remodeling complex is related to yeast rsc and localizes at kinetochores of mitotic chromosomes. *Proc Natl Acad Sci U S A* **97**, 13015-13020 (2000).

120. C. Dhalluin *et al.*, Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491-496 (1999).
121. L. Xu *et al.*, Structural insight into the recognition of acetylated histone H3K56ac mediated by the bromodomain of CREB-binding protein. *Febs j* **284**, 3422-3436 (2017).
122. P. Filippakopoulos *et al.*, Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell* **149**, 214-231 (2012).
123. T.-H. Huang *et al.*, The Histone Chaperones ASF1 and CAF-1 Promote MMS22L-TONSL-Mediated Rad51 Loading onto ssDNA during Homologous Recombination in Human Cells. *Molecular Cell* **69**, 879-892.e875 (2018).
124. T. Suzuki *et al.*, A novel Arabidopsis gene TONSOKU is required for proper cell arrangement in root and shoot apical meristems. *Plant J* **38**, 673-684 (2004).
125. S. Takeda *et al.*, BRU1, a novel link between responses to DNA damage and epigenetic gene silencing in Arabidopsis. *Genes & development* **18**, 782-793 (2004).
126. J. Fang *et al.*, Purification and functional characterization of SET8, a nucleosomal histone H4-lysine 20-specific methyltransferase. *Curr Biol* **12**, 1086-1099 (2002).
127. K. Nishioka *et al.*, PR-Set7 Is a Nucleosome-Specific Methyltransferase that Modifies Lysine 20 of Histone H4 and Is Associated with Silent Chromatin. *Molecular Cell* **9**, 1201-1213 (2002).
128. J. C. Rice *et al.*, Mitotic-specific methylation of histone H4 Lys 20 follows increased PR-Set7 expression and its localization to mitotic chromosomes. *Genes & Development* **16**, 2225-2230 (2002).
129. S. Guyomarc'h, T. Vernoux, J. Traas, D. X. Zhou, M. Delarue, MGOUN3, an Arabidopsis gene with TetratricoPeptide-Repeat-related motifs, regulates meristem cellular organization. *J Exp Bot* **55**, 673-684 (2004).
130. A. Perez-Riba, L. S. Itzhaki, The tetratricopeptide-repeat motif is a versatile platform that enables diverse modes of molecular recognition. *Curr Opin Struct Biol* **54**, 43-49 (2019).
131. L. Lu, X. Chen, S. Qian, X. Zhong, The plant-specific histone residue Phe41 is important for genome-wide H3.1 distribution. *Nature Communications* **9**, 630 (2018).
132. C. Raynaud *et al.*, Two cell-cycle regulated SET-domain proteins interact with proliferating cell nuclear antigen (PCNA) in Arabidopsis. *The Plant Journal* **47**, 395-407 (2006).
133. Y. Jacob *et al.*, Regulation of heterochromatic DNA replication by histone H3 lysine 27 methyltransferases. *Nature* **466**, 987 (2010).
134. W. Feng *et al.*, Large-scale heterochromatin remodeling linked to overreplication-associated DNA damage. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 406-411 (2017).
135. H. Stroud *et al.*, DNA Methyltransferases Are Required to Induce Heterochromatic Re-Replication in Arabidopsis. *PLOS Genetics* **8**, e1002808 (2012).
136. T. Suzuki *et al.*, TONSOKU is Expressed in S Phase of the Cell Cycle and its Defect Delays Cell Cycle Progression in Arabidopsis. *Plant and Cell Physiology* **46**, 736-742 (2005).
137. J. M. Lucht *et al.*, Pathogen stress increases somatic recombination frequency in Arabidopsis. *Nature Genetics* **30**, 311-314 (2002).
138. R. M. Williams, X. Zhang, Roles of ATM and ATR in DNA double strand breaks and replication stress. *Progress in Biophysics and Molecular Biology* **163**, 109-119 (2021).
139. Q. Wang *et al.*, Rad17 recruits the MRE11-RAD50-NBS1 complex to regulate the cellular response to DNA double-strand breaks. *Embo j* **33**, 862-877 (2014).

140. J. A. Kamp, R. van Schendel, I. W. Dilweg, M. Tijsterman, BRCA1-associated structural variations are a consequence of polymerase theta-mediated end-joining. *Nature Communications* **11**, 3615 (2020).
141. J. Dong *et al.*, H3.1K27me1 maintains transcriptional silencing and genome stability by preventing GCN5-mediated histone acetylation. *The Plant Cell* **33**, 961-979 (2021).
142. E. I. Campos *et al.*, Analysis of the Histone H3.1 Interactome: A Suitable Chaperone for the Right Event. *Mol Cell* **60**, 697-709 (2015).
143. G. Millan-Zambrano, A. Burton, A. J. Bannister, R. Schneider, Histone post-translational modifications - cause and consequence of genome function. *Nat Rev Genet* **23**, 563-580 (2022).
144. S. Awad, A. H. Hassan, The Swi2/Snf2 bromodomain is important for the full binding and remodeling activity of the SWI/SNF complex on H3- and H4-acetylated nucleosomes. *Ann N Y Acad Sci* **1138**, 366-375 (2008).
145. L. Zeng, M.-M. Zhou, Bromodomain: an acetyl-lysine binding domain. *FEBS Letters* **513**, 124-128 (2002).
146. C. S. Karam, W. A. Kellner, N. Takenaka, A. W. Clemmons, V. G. Corces, 14-3-3 Mediates Histone Cross-Talk during Transcription Elongation in Drosophila. *PLOS Genetics* **6**, e1000975 (2010).
147. A. A. Jeyaprakash, C. Basquin, U. Jayachandran, E. Conti, Structural Basis for the Recognition of Phosphorylated Histone H3 by the Survivin Subunit of the Chromosomal Passenger Complex. *Structure* **19**, 1625-1634 (2011).
148. H. Greschik, R. Schule, T. Gunther, Selective targeting of epigenetic reader domains. *Expert Opin Drug Discov* **12**, 449-463 (2017).
149. D. Zhao, Y. Li, X. Xiong, Z. Chen, H. Li, YEATS Domain-A Histone Acylation Reader in Health and Disease. *J Mol Biol* **429**, 1994-2002 (2017).
150. E. Bergamin *et al.*, Molecular basis for the methylation specificity of ATXR5 for histone H3. *Nucleic Acids Res* **45**, 6375-6387 (2017).
151. L. Hu, Z. Li, P. Wang, Y. Lin, Y. Xu, Crystal structure of PHD domain of UHRF1 and insights into recognition of unmodified histone H3 arginine residue 2. *Cell Res* **21**, 1374-1378 (2011).
152. L. O'Donnell *et al.*, The MMS22L-TONSL Complex Mediates Recovery from Replication Stress and Homologous Recombination. *Molecular Cell* **40**, 619-631 (2010).
153. D. Reinberg, R. J. Sims, 3rd, de FACTo nucleosome dynamics. *J Biol Chem* **281**, 23297-23301 (2006).
154. S. J. Hill *et al.*, Systematic screening reveals a role for BRCA1 in the response to transcription-associated DNA damage. *Genes & Development* **28**, 1957-1975 (2014).
155. E. Bergamin *et al.*, Molecular basis for the methylation specificity of ATXR5 for histone H3. *Nucleic Acids Research* **45**, 6375-6387 (2017).
156. F. Casadio *et al.*, H3R42me2a is a histone modification with positive transcriptional effects. *Proc Natl Acad Sci U S A* **110**, 14894-14899 (2013).
157. S. X. Pfister *et al.*, SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Rep* **7**, 2006-2018 (2014).
158. M. Daugaard *et al.*, LEDGF (p75) promotes DNA-end resection and homologous recombination. *Nature Structural & Molecular Biology* **19**, 803-810 (2012).
159. A. D. Caperta *et al.*, Distribution patterns of phosphorylated Thr 3 and Thr 32 of histone H3 in plant mitosis and meiosis. *Cytogenet Genome Res* **122**, 73-79 (2008).

160. Z. Dong, A. M. Bode, The role of histone H3 phosphorylation (Ser10 and Ser28) in cell growth and cell transformation. *Mol Carcinog* **45**, 416-421 (2006).
161. P. N. Lau, P. Cheung, Histone code pathway involving H3 S28 phosphorylation and K27 acetylation activates transcription and antagonizes polycomb silencing. *Proc Natl Acad Sci U S A* **108**, 2801-2806 (2011).
162. A. Sawicka *et al.*, H3S28 phosphorylation is a hallmark of the transcriptional response to cellular stress. *Genome Res* **24**, 1808-1820 (2014).
163. E. R. Vossenaar, A. J. Zendman, W. J. van Venrooij, G. J. Pruijn, PAD, a growing family of citrullinating enzymes: genes, features and involvement in disease. *Bioessays* **25**, 1106-1118 (2003).
164. Y. Wang *et al.*, Human PAD4 regulates histone arginine methylation levels via demethylimination. *Science* **306**, 279-283 (2004).
165. Y. Wang *et al.*, Histone hypercitrullination mediates chromatin decondensation and neutrophil extracellular trap formation. *Journal of Cell Biology* **184**, 205-213 (2009).
166. V. Papayannopoulos, Neutrophil extracellular traps in immunity and disease. *Nature Reviews Immunology* **18**, 134-147 (2018).
167. Y. Liu *et al.*, Histone H3K27 dimethylation landscapes contribute to genome stability and genetic recombination during wheat polyploidization. *Plant J* **105**, 678-690 (2021).
168. Q. Shen, Y. Lin, Y. Li, G. Wang, Dynamics of H3K27me3 Modification on Plant Adaptation to Environmental Cues. *Plants (Basel)* **10**, (2021).
169. P. Chammas, I. Mocavini, L. Di Croce, Engaging chromatin: PRC2 structure meets function. *Br J Cancer* **122**, 315-328 (2020).
170. B. M. Dancy, P. A. Cole, Protein lysine acetylation by p300/CBP. *Chem Rev* **115**, 2419-2452 (2015).
171. Q. Jin *et al.*, Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *Embo j* **30**, 249-262 (2011).
172. R. J. Burgess, H. Zhou, J. Han, Z. Zhang, A role for Gcn5 in replication-coupled nucleosome assembly. *Mol Cell* **37**, 469-480 (2010).
173. C. Chen *et al.*, Cytosolic acetyl-CoA promotes histone acetylation predominantly at H3K27 in Arabidopsis. *Nat Plants* **3**, 814-824 (2017).
174. K. Li, Z. Wang, Histone crotonylation-centric gene regulation. *Epigenetics Chromatin* **14**, 10 (2021).
175. C. Lu, M. Coradin, K. A. Janssen, S. Sidoli, B. A. Garcia, Combinatorial Histone H3 Modifications Are Dynamically Altered in Distinct Cell Cycle Phases. *J Am Soc Mass Spectrom* **32**, 1300-1311 (2021).
176. A. Orthwein *et al.*, Mitosis Inhibits DNA Double-Strand Break Repair to Guard Against Telomere Fusions. *Science* **344**, 189-193 (2014).
177. Q. Zhu *et al.*, SETD2-mediated H3K14 trimethylation promotes ATR activation and stalled replication fork restart in response to DNA replication stress. *Proceedings of the National Academy of Sciences* **118**, e2011278118 (2021).
178. T. Rampias *et al.*, The lysine-specific methyltransferase KMT2C/MLL3 regulates DNA repair components in cancer. *EMBO Rep* **20**, (2019).
179. L. H. Wang, M. A. E. Aberin, S. Wu, S. P. Wang, The MLL3/4 H3K4 methyltransferase complex in establishing an active enhancer landscape. *Biochem Soc Trans* **49**, 1041-1054 (2021).

180. T. Zhang, Z. Zhang, Q. Dong, J. Xiong, B. Zhu, Histone H3K27 acetylation is dispensable for enhancer activity in mouse embryonic stem cells. *Genome Biol* **21**, 45 (2020).
181. F. Wang *et al.*, Histone H3 Thr-3 phosphorylation by Haspin positions Aurora B at centromeres in mitosis. *Science* **330**, 231-235 (2010).
182. L. Micale *et al.*, Novel TONSL variants cause SPONASTRIME dysplasia and associate with spontaneous chromosome breaks, defective cell proliferation and apoptosis. *Hum Mol Genet* **29**, 3122-3131 (2020).
183. K. Hyun, J. Jeon, K. Park, J. Kim, Writing, erasing and reading histone lysine methylations. *Experimental & Molecular Medicine* **49**, e324 (2017).
184. E. L. Greer, Y. Shi, Histone methylation: a dynamic mark in health, disease and inheritance. *Nature Reviews Genetics* **13**, 343 (2012).
185. S. Lanouette, V. Mongeon, D. Figeys, J. F. Couture, The functional diversity of protein lysine methylation. *Molecular Systems Biology* **10**, 724 (2014).
186. J. Müller *et al.*, Histone Methyltransferase Activity of a Drosophila Polycomb Group Repressor Complex. *Cell* **111**, 197-208 (2002).
187. K. Plath *et al.*, Role of Histone H3 Lysine 27 Methylation in X Inactivation. **300**, 131-135 (2003).
188. Y. Yu *et al.*, Inhibition of EZH2 Promotes Human Embryonic Stem Cell Differentiation into Mesoderm by Reducing H3K27me3. *Stem cell reports* **9**, 752-761 (2017).
189. L. O. Baumbusch *et al.*, The Arabidopsis thaliana genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic Acids Research* **29**, 4319-4333 (2001).
190. D. Latrasse *et al.*, Dual function of MIPS1 as a metabolic enzyme and transcriptional regulator. *Nucleic Acids Research* **41**, 2907-2917 (2013).
191. S. Hasan, P. O. Hassa, R. Imhof, M. O. Hottiger, Transcription coactivator p300 binds PCNA and may have a role in DNA repair synthesis. *Nature* **410**, 387 (2001).
192. L. S.-H. Chuang *et al.*, Human DNA-(Cytosine-5) Methyltransferase-PCNA Complex as a Target for $p21^{WAF1}$. **277**, 1996-2000 (1997).
193. T. R. Ben-Shahar *et al.*, Two Fundamentally Distinct PCNA Interaction Peptides Contribute to Chromatin Assembly Factor 1 Function. *Molecular and Cellular Biology* **29**, 6353-6365 (2009).
194. R. A. Poot *et al.*, The Williams syndrome transcription factor interacts with PCNA to target chromatin remodelling by ISWI to replication foci. *Nature Cell Biology* **6**, 1236 (2004).
195. P. A *et al.*, EZH2 promotes DNA replication by stabilizing interaction of POLδ and PCNA via methylation-mediated PCNA trimerization. *Epigenetics & Chromatin* **11**, 44 (2018).
196. C. M. Kondratick, J. M. Litman, K. V. Shaffer, M. T. Washington, L. M. Dieckman, Crystal structures of PCNA mutant proteins defective in gene silencing suggest a novel interaction site on the front face of the PCNA ring. *PloS one*. 2018 (10.1371/journal.pone.0193333).
197. C. Indiani, P. McInerney, R. Georgescu, M. F. Goodman, M. O'Donnell, A Sliding-Clamp Toolbelt Binds High- and Low-Fidelity DNA Polymerases Simultaneously. *Molecular Cell* **19**, 805-815 (2005).
198. D. Dovrat, J. L. Stodola, P. M. J. Burgers, A. Aharoni, Sequential switching of binding partners on PCNA during in vitro Okazaki fragment maturation. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 14118-14123 (2014).
199. D. Bubeck *et al.*, PCNA directs type 2 RNase H activity on DNA replication and repair substrates. *Nucleic Acids Research* **39**, 3652-3666 (2011).

200. S. Jørgensen *et al.*, SET8 is degraded via PCNA-coupled CRL4(CDT2) ubiquitylation in S phase and after UV irradiation. *Journal of Cell Biology* **192**, 43-54 (2011).
201. T. Abbas *et al.*, $CRL4^{Cdt2}$ Regulates Cell Proliferation and Histone Gene Expression by Targeting PR-Set7/Set8 for Degradation. *Molecular Cell* **40**, 9-21 (2010).
202. A. S. Altieri *et al.*, A small protein inhibits proliferating cell nuclear antigen by breaking the DNA clamp. *Nucleic Acids Research* **44**, 6232-6241 (2016).
203. J. B. Bruning, Y. Shamoo, Structural and Thermodynamic Analysis of Human PCNA with Peptides Derived from DNA Polymerase-δ p66 Subunit and Flap Endonuclease-1. *Structure* **12**, 2209-2219 (2004).
204. Y. Wang, M. Xu, T. Jiang, Crystal structure of human PCNA in complex with the PIP box of DVC1. *Biochemical and Biophysical Research Communications* **474**, 264-270 (2016).
205. T. Kaufmann *et al.*, A novel non-canonical PIP-box mediates PARG interaction with PCNA. *Nucleic Acids Research* **45**, 9741-9759 (2017).
206. A. Hishiki *et al.*, Structural Basis for Novel Interactions between Human Translesion Synthesis Polymerases and Proliferating Cell Nuclear Antigen. **284**, 10552-10560 (2009).
207. T. S. Girish, R. K. McGinty, S. Tan, Multivalent Interactions by the Set8 Histone Methyltransferase With Its Nucleosome Substrate. *Journal of Molecular Biology* **428**, 1531-1543 (2016).
208. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D: Biological Crystallography* **66**, 213-221 (2010).
209. W. Strzalka, T. Oyama, K. Tori, K. Morikawa, Crystal structures of the Arabidopsis thaliana proliferating cell nuclear antigen 1 and 2 proteins complexed with the human p21 C-terminal segment. *Protein Science : A Publication of the Protein Society* **18**, 1072-1080 (2009).
210. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography* **66**, 486-501 (2010).
211. V. B. Chen *et al.*, MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography* **66**, 12-21 (2010).
212. P. Schuck, Size-Distribution Analysis of Macromolecules by Sedimentation Velocity Ultracentrifugation and Lamm Equation Modeling. *Biophysical Journal* **78**, 1606-1619 (2000).
213. E. Bergamin, J. F. Couture, in *Methods in Enzymology,* R. Marmorstein, Ed. (Academic Press, 2016), vol. 573, pp. 209-240.
214. M. Lachner, D. O'Carroll, S. Rea, K. Mechtler, T. Jenuwein, Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116-120 (2001).
215. Z. Sun *et al.*, H3K36me3, message from chromatin to DNA damage repair. *Cell Biosci* **10**, 9 (2020).
216. N. L. Young, R. Dere, Mechanistic insights into KDM4A driven genomic instability. *Biochem Soc Trans* **49**, 93-105 (2021).
217. L. C. Young, D. W. McDonald, M. J. Hendzel, Kdm4b histone demethylase is a DNA damage response protein and confers a survival advantage following γ-irradiation. *J Biol Chem* **288**, 21376-21388 (2013).
218. F. Pryde *et al.*, H3 k36 methylation helps determine the timing of cdc45 association with replication origins. *PLoS One* **4**, e5882 (2009).

219. Y. Yu, Y. Teng, H. Liu, S. H. Reed, R. Waters, UV irradiation stimulates histone acetylation and chromatin remodeling at a repressed yeast locus. *Proc Natl Acad Sci U S A* **102**, 8650-8655 (2005).
220. T. J. Parnell, J. T. Huff, B. R. Cairns, RSC regulates nucleosome positioning at Pol II genes and density at Pol III genes. *EMBO J* **27**, 100-110 (2008).
221. M. R. Duan, M. J. Smerdon, Histone H3 lysine 14 (H3K14) acetylation facilitates DNA repair in a positioned nucleosome by stabilizing the binding of the chromatin Remodeler RSC (Remodels Structure of Chromatin). *J Biol Chem* **289**, 8353-8363 (2014).
222. R. W. Baker *et al.*, Structural insights into assembly and function of the RSC chromatin remodeling complex. *Nat Struct Mol Biol* **28**, 71-80 (2021).
223. B. J. Klein *et al.*, Histone H3K23-specific acetylation by MORF is coupled to H3K14 acylation. *Nat Commun* **10**, 4724 (2019).
224. I. Dreveny *et al.*, The double PHD finger domain of MOZ/MYST3 induces alpha-helical structure of the histone H3 tail to facilitate acetylation and methylation sampling and modification. *Nucleic Acids Res* **42**, 822-835 (2014).
225. S. Adhikary *et al.*, Atypical plant homeodomain of UBR7 functions as an H2BK120Ub ligase and breast tumor suppressor. *Nature Communications* **10**, 1398 (2019).

Appendix – Supplementary Materials for Chapter 2

The references made in the materials below correspond to reference numbers within the Reference section of this appendix.

The data, figures, and tables in the following material of this appendix were produced by me. The remainder of the materials were produced by our collaborators.

Figure S2-Panel D

Figure S3

Table S1

Science
AAAS

Supplementary Materials for

The histone H3.1 variant regulates TONSOKU-mediated DNA repair during replication

Hossein Davarinejad *et al.*

Corresponding authors: Jean-François Couture, jean-francois.couture@uottawa.ca; Yannick Jacob, yannick.jacob@yale.edu

Science **375**, 1281 (2022)
DOI: 10.1126/science.abm5320

The PDF file includes:

Materials and Methods
Figs. S1 to S12
Tables S1 and S3
References

Other Supplementary Material for this manuscript includes the following:

Tables S2, S4, and S5
MDAR Reproducibility Checklist

Materials and Methods

Plant materials

A. thaliana plants were grown under cool-white fluorescent lights (~100 μmol m^{-2} s^{-1}) in long-day conditions (16 h light/8 h dark). The *atxr5/6* double mutant was described previously (*17*). *tsk/bru1-4* (*At3g18730*, SALK_034207 (*29*)), *ku70-2* (*At1g16970*, SALK_123114c (*30*)), *ku80-7* (*At1g48050*, SALK_112921 (*30*)), *lig4-4* (*At5g57160*, SALK_044027 (*31*)), *rad17-2* (*At5g66130*, SALK_009384 (*32*)), *brca2a* (*At4g00020*, 13F-1 allele (*33*)), *brca2b* (*At5g01630*, SALK_037617 (*33*)), *rad54-1* (*At3g19210*, SALK_038057 (*34*)), and *teb-5* (*At4g32700*, SALK_018851 (*35*)) are in the Col-0 genetic background. They were obtained from the Arabidopsis Biological Resource Center (Columbus, Ohio), and described in previous publications. CRISPR/Cas9 was used to mutate *RAD51* (*At5g20850*) in Col-0 and in the *atxr5/6* mutant background. The *h3.1* quadruple mutant (*htr1 htr2 htr3 htr9*) used for transformation of the *H3.1* transgenes (WT and point mutants) was described previously (*16*). The *h3.1* pentuple mutant (*htr1 htr2 htr3 htr9 htr13*) was created by performing temperature-optimized CRISPR/Cas9 at the *HTR1*, *HTR2* and *HTR13* genes in the *h3.1* quadruple mutant (*36*). *htr1* and *htr2* have T-DNA insertions just outside the coding sequence in the *h3.1* quadruple mutant (*16*). Transgenic plants expressing H3.1 WT or H3.1A31T (in the *h3.1* quadruple mutant) used in the MMS genotoxic assay were described previously (*16*). The septuple mutant *atxr5 atxr6 htr1 htr2 htr3 htr9 htr13* was generated by crossing the *atxr5/6* double mutant with the *h3.1* pentuple mutant, followed by multiplex CRISPR/Cas9 editing at *HTR9* and *HTR13*.

Plasmid constructs

The coding sequences of the TPR domains of *A. thaliana* TSK (AtTSK; a.a. 1-525 followed by a stop codon) and of mouse TONSL (a.a. 1-515 followed by a stop codon) were cloned into the pET32a vector using BamHI and SalI, yielding pET32a-TSK and pET32a-TONSL, respectively. For the *C. unshiu* TSK (CuTSK) construct, a.a. sequences ENLYFQG (TEV cleavage site) followed by CuTSK a.a. 1-530 (accession ID: GAY58445.1) were cloned into the pET22a(+) vector using BamHI and XhoI sites (GeneScript, Piscataway, NJ). CuTSK 1-490 construct was generated by placing a stop codon after residue 490 using site-directed mutagenesis (QuickChange, Agilent Technologies, Santa Clara, CA).

The coding sequences of the N-terminal tails of *A. thaliana* H2A.Z (a.a. 1-16), H2A.X (a.a. 1-16), H2B (a.a. 1-35), H3.1 (a.a. 1-58), H3.3 (a.a. 1-58), and H4 (a.a. 1-30) were fused with a C-terminal GST tag by cloning into pET28-Mff(1-61)-PP-GST (Addgene plasmid #73042; gifted by D. Chan) using the NdeI and BamHI sites (*37*). The H3.1A31T and H3.1F41Y mutations were engineered by site-directed mutagenesis (QuikChange II XL, Agilent Technologies).

HTR1 (H3.1, *At5g65360*) and its promoter (1167 bp upstream of the start codon) were cloned into pENTR/D-TOPO (Thermo Fisher Scientific, Waltham, MA), subcloned using Gateway Technology into pB7WG (*38*), and modified using site-directed mutagenesis to generate the following H3.1 point mutant constructs: H3.1S28A, H3.1K4A, H3.1S28A K4A, H3.1K9A, H3.1S28A K9A, H3.1K36A, H3.1S28A K36A, H3.1A31T and H3.1S28A A31T.

Protein expression

A Rosetta (DE3) *E. coli* strain (#70954, Sigma, St. Louis, MO) was used for the expression of the following recombinant proteins: AtTSK, mouse TONSL and the histone-GST fusion proteins. The bacteria were cultured in LB, and 1 mM IPTG was used to induce protein expression. For Selenium Methionine (SelMet)-CuTSK$_{(1-490)}$, the plasmid was transformed into B843 *E. coli* and grown in M9 minimal medium supplemented with SelMet (Complete kit MD12-500, Molecular Dimensions, Holland, OH). CuTSK$_{(1-490)}$ construct was expressed at 18°C and induced with IPTG (0.2 mM) at OD_{600nm} = 0.6.

For purification of AtTSK and TONSL (containing an N-terminal Trx-His-S tag), the cell pellets were resuspended in NPI-10 buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole, pH 8) containing 1 mM PMSF and sonicated. After centrifugation, the supernatant was passed through a Ni-NTA agarose column. The column was then washed with NPI-20 buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole, pH 8), and the recombinant proteins were eluted with NPI-250 buffer (50 mM NaH_2PO_4, 300 mM NaCl, 250 mM imidazole, pH 8). The eluted proteins were further purified by size exclusion chromatography (SEC), aliquoted, and stored at -80°C.

For purification of the histone-GST fusion proteins, the cell pellets were resuspended in 1X PBS (137 mM NaCl, 10 mM phosphate, 2.7 mM KCl, pH 7.4) containing 1 mM PMSF before sonification and centrifugation. The supernatant was passed through a Glutathione Sepharose 4B column, and bound proteins were washed with 1× PBS and eluted using EB buffer (50 mM Tris, 50 mM NaCl, 30 mM reduced L-Glutathione, 10% glycerol, pH 8.0). Proteins were aliquoted and stored at -80°C.

For purification of CuTSK (fragments 1-490 and 1-530), cell pellets were resuspended in NaP_i buffer (50 mM NaP_i pH 7.5, 1 M NaCl, 10% glycerol, 5 mM βME) and purified at 4°C using cobalt resin (Talon) (TaKaRa, Japan). Proteins were TEV-cleaved on beads and purified using SEC (Superdex75, GE Healthcare, Chicago, IL) columns equilibrated with NaP_i buffer (250 mM NaCl for pulldowns; 350 mM NaCl for ITC) or Tris buffer (20 mM Tris pH 7.5, 200 mM NaCl, 5% glycerol, 5 mM βME) for crystallography.

In vitro binding assays

For the binding assays involving AtTSK and the GST-tagged histone N-terminal tail proteins, 2 μg of AtTSK was mixed with 2 μg of GST or GST-tagged histone proteins in 400 μl of binding buffer (25 mM Tris, 250 mM NaCl, 0.05% NP-40, pH 8.0), and the mixture was incubated at 4°C overnight. 15 μl pre-washed Glutathione Sepharose 4B agarose beads were added to each tube and incubated for 30 min to pull down the GST-tagged histone proteins. The beads were washed four times with 1 mL of binding buffer, with each wash performed by rotating for 5 min at 4°C. After the final wash, 15 μl of 2× SDS loading buffer was added to each tube, and the proteins were eluted by boiling at 95°C for 5 min. The samples were separated on a 10% SDS-PAGE gel. The lower part of the gel was subjected to Coomassie staining to visualize the GST or GST-tagged histone N-terminal tail proteins, and the upper part of the gel was subjected to Western blot analyses using an anti-His antibody (H1029) (Sigma). Each pulldown assay was performed at least three times. For the binding assays using TONSL and the biotin-tagged full-length histone proteins (1-135 aa, Active Motif, Carlsbad, CA), 1 μg of TONSL was mixed with 1 μg of biotin or biotin-tagged

histone proteins in 400 µl of binding buffer (25 mM Tris, 450 mM NaCl, 0.05% NP-40, pH 8.0), and incubated at 4°C for 30 min. Pre-washed MyOne Streptavidin beads (20 µl) (Invitrogen, Waltham, MA) were added to each tube and incubated for 30 min to pull down the biotin-tagged histone proteins. The beads were washed four times with 1 mL of binding buffer for 5 min at 4°C. After the final wash, 15 µl of 2× SDS loading buffer was added to each tube, and the proteins were eluted by boiling at 95°C for 5 min. The samples were separated on a Bio-Rad 4–20% Mini-PROTEAN® TGX™ Precast Protein Gel. The lower part of the gel was subjected to Coomassie staining to visualize the biotin or biotin-tagged histone proteins, and the upper part of the gel was subjected to Western blot analyses using an anti-His antibody (H1029) (Sigma). Pulldown assays were performed three times.

For binding assays involving $CuTSK_{(1-530)}$ and biotinylated H3.1 peptides, 50 µL of streptavidin agarose resin (Millipore, Burlington, MA, #69203-3) was washed with pulldown buffer (200 mM NaCl, 150 mM NaP_i pH 7.5, 10% glycerol, 5 mM βME) and saturated with $H3.1_{(1-45)}$ peptides (unmodified/modified) while incubating at 4°C for 30 min. Peptide-bound resin was incubated with 50 µg $CuTSK_{(1-530)}$ for 30 min in 200 µL of pulldown buffer, washed, boiled, and loaded onto SDS 4-20% acrylamide gels. Pulldown assays were performed three times with two separate purifications of $CuTSK_{(1-530)}$.

For nucleosome pulldown assays, histone octamers were reconstituted using *Xenopus laevis* H2A, H2B, H4, and *A. thaliana* H3.1 or H3.3 and purified by gel filtration on a S200 size exclusion column (GE Healthcare). A biotinylated 209-bp DNA fragment containing the 601 nucleosome positioning sequence was generated by PCR with a biotinylated forward primer and purified by ion exchange chromatography on a HiTrap Q column followed by ethanol precipitation. Mononucleosomes were assembled from histone octamers and 601 DNA by gradient dialysis as described previously (*39*). Nucleosome assembly was verified by native gel electrophoresis on 6% acrylamide gels in 0.5× TGE buffer (12.5 mM Tris pH 8.0, 95 mM glycine, 0.5 mM EDTA). GST-TPR was expressed as GST-$CuTSK_{(1-530)}$ in pGST vector and purified as described for CuTSK above. To carry out the binding assays, all incubation and wash steps were performed at 4°C with end-over-end rotation. Centrifugation of beads before washes was done at 1,500 *g* for 2 min at 4°C. 3 µg (23 pmol) of assembled nucleosomes were immobilized on streptavidin sepharose high performance beads (GE Healthcare) that were blocked with 1 mg/ml BSA in pulldown buffer (20 mM HEPES pH 7.9, 175 mM NaCl, 10% glycerol, 1 mM EDTA, 1 mM DTT, 0.1% NP-40, 0.1 mg/ml BSA) by overnight incubation. After three 5-min washes with pulldown buffer, nucleosome-bound beads were incubated with GST-CuTSK for 2 h. Beads were then washed 5 times for 10 min each with high salt pulldown buffer (as above but with 350 mM NaCl) before elution of bound proteins by boiling in 1.5× SDS sample buffer (95 mM Tris HCl pH 6.8, 15% glycerol, 3% SDS, 75 mM DTT, 0.15% bromophenol blue). Protein binding was analyzed by Western Blotting with an anti-GST-HRP antibody (ab3416, Abcam, Waltham, MA). Nucleosome loading was confirmed by Western Blotting with anti-H3 antibody (ab176842, Abcam). The nucleosome pulldown assays were repeated three times, using three different preparations of nucleosomes.

ITC assay

ITC experiments were performed using a VP-ITC calorimeter (MicroCal, Northampton, MA) by injecting peptides (750 μM) into a solution of $CuTSK_{(1-530)}$ (50 μM) in 50 mM NaP_i pH 7.5, 350 mM NaCl, 10% glycerol, and 5 mM βME. The experiment was performed at 19°C, and the titration data were analyzed using Origin software (OriginLab Corporation, Northampton, MA). ITC experiments were replicated three times using two separate batches of synthesized peptides and three separate purifications of $CuTSK_{(1-530)}$.

Histone H3 peptide synthesis

Fmoc-protected amino acids and rink amide low loading resin were purchased from CEM (Matthews, NC). Fmoc-Lysine(Boc)(Me)-OH, Fmoc-Lysine$(Me)_2$-OH, and Fmoc-Lysine$(Me)_3$-OH building blocks were purchased from Bachem (Bubendorf, Switzerland). All peptides were synthesized using microwave-assisted Fmoc solid phase peptide synthesis in a Liberty Blue automated system. Briefly, the required amount of resin was swelled in DMF for 5 min. Next, Fmoc deprotection was carried out with 20% piperidine at 90°C for 60s. Standard coupling cycles using DIC/Oxyma Pure were run at 90°C for 240s in each amino acid. Peptides were cleaved from the resin and deprotected with TFA/TIS/EDT/H_2O (92.5/2.5/2.5/2.5 % v/v) at 42°C for 30 min, and then precipitated in -20°C diethyl ether. Peptide crude products were then dried under vacuum overnight and purified by RP-HLPC in a Waters 1525EF semi-preparative system with a 21.6 × 250 mm C18 column at 20 mL/min. Peptide purity and identity were confirmed via RP-UPLC-UV/MS in a Waters Acquity UPLC Xevo TQD using a 2.1 × 100 mm UPLC BEH C8 column. A purity of >95% was determined through HPLC peak analysis. The molecular ions identified for each peptide are described in Table S3. Lyophilized peptides were resuspended in water.

Crystallography

$CuTSK_{(1-490)}$ (20 mg/ml) was incubated with $H3.1_{(1-45)}$ (5:1 peptide:TSK molar ratio) and crystalized in 25% 1,2-propanediol, 5% glycerol, 0.1 M Na/K phosphate pH 6.0. A single-wavelength anomalous dispersion (SAD) data set was collected at the 21-ID-D beamline of the Life Science-Collaborative Access Team at the Advanced Photon Source Synchrotron. The structure of $CuTSK_{(1-490)}$ was determined by SAD at the selenium peak wavelength. The reflections were processed and scaled using HKL2000 (*40*) and 23 selenium atoms were identified and refined using the SHELX C/D programs (*41*). Phases were calculated using SHELX-E and the Arp/Warp program was used to generate the initial model. One chain was traced and used as a search model for molecular replacement of two TSK chains in the asymmetric unit using Phaser. Missing residues were modeled in the calculated phases using Coot (*42*) and the structure was further refined using phenix.refine (*43*). Clash scores were determined by MolProbity (*44*). The final model includes 17 selenium atoms, one water molecule, CuTSK residues 1-124, 128-150 and 159-483, and H3.1 residues 4-9 and 18-40. Missing residues were not modeled due to lack of electron density.

BiFC and confocal microscopy

The TPR domain (a.a. 1-524) of AtTSK with a nuclear localization signal (NLS) was cloned into the Gateway destination vector pUC-DEST-VYCE®GW, and histones H3.1 and H3.3 (a.a. 1-136)

were cloned into the Gateway destination vector pUC-DEST-VYNE®GW (*45*). pSAT6-mCherry-VirD2NLS was used as a nuclear marker. The protoplasts were isolated from 3- to 4-week-old *A. thaliana* plants (*atxr5/6)* and transfected following the tape-Arabidopsis sandwich method (*46*). After 14–18 h incubation in low-light conditions, protoplast images were acquired using a confocal spinning disk unit (Yokogawa CSU-W1), mounted on a Nikon Eclipse Ti2 microscope body (Nikon, Minato City, Tokyo, Japan). A 60× water objective (N.A. = 1.2) and a 1.5× post magnification along with 514 nm and 561 nm lasers were used for imaging as described (*47*). The images were processed with FIJI (*48*). Assays were repeated three times with similar results.

Plant nuclei microscopy

Leaves from four-week-old plants were fixed in 3.7% formaldehyde in cold Tris buffer (10 mM Tris-HCl pH 7.5, 10 mM NaEDTA, 100 mM NaCl) for 20 min, then washed for 10 min in Tris buffer. The leaves were finely chopped with a razor blade in 500 µl LB01 buffer (15 mM Tris-HCl pH 7.5, 2 mM NaEDTA, 0.5 mM spermine-4HCl, 80 mM KCl, 20 mM NaCl and 0.1% Triton X-100) and filtered through a 30 µm mesh (Sysmex Partec, Gorlitz, Germany). 10 µl of lysate was mixed to 10 µl of sorting buffer (100 mM Tris-HCl pH 7.5, 50 mM KCl, 2mM $MgCl_2$, 0.05% Tween-20 and 5% sucrose) and spread onto a coverslip. After drying for 30 min, cold methanol was added onto each coverslip for 3 min. Methanol was removed and TBS-Tx (20 mM Tris pH 7.5, 100 mM NaCl, 0.1% Triton X-100) was added for 5 min. The coverslips were mounted onto slides with Vectashield mounting medium containing DAPI (Vector Laboratories, Burlingame, CA). Imaging was done using a Nikon Eclipse Ni-E microscope with a 100× CFI PlanApo Lamda objective (Nikon) and equipped with an Andor Clara camera. Z-series optical sections of each nucleus were obtained at 0.3-µm steps. Images were deconvolved by ImageJ using the deconvolution plugin. Three biological samples per genotype were assessed for each experiment. Twenty-five nuclei were analyzed for each sample.

RT-qPCR

RNA extraction from three-week-old leaf tissue was performed using TRIzol (Invitrogen). RNA samples were treated with RQ1 RNase-free DNase (Promega, Madison, WI) at 37°C for 30 min. 1 µg of total RNA was used to produce cDNA with SuperScript II Reverse Transcriptase (Invitrogen) using oligo-dT primers. Real-time PCR was done using a CFX96 Real-Time PCR Detection System (Bio-Rad, Hercules, CA) with KAPA SYBR FAST qPCR Master Mix (2×) Kit (Kapa Biosystems, Wilmington, MA). Relative quantities were determined by the C_t method (*49*) with *ACTIN* as the normalizer. At least three biological samples were used for each experiment.

Primer name	Sequence
ACTIN-F	TCGTGGTGGTGAGTTTGTTAC
ACTIN-R	CAGCATCATCACAAGCATCC
TSI-F	ATCCAGTCCGAAGAACGCGAACTA
TSI-R	TCACTTGTGAGTGTTCGTGAGGTC
BRCA1-F	CATGTGCCTTTTGTCAGTGTTC
BRCA1-R	TGGAGCCCATTCAGCACAGTTT
H3.1 transgene-F	GCAGCGCCGTCGCAGCACTTCAGG
H3.1 transgene-R	ACTCTAGCATGGCCGCGGGATATC

Flow cytometry

To generate flow cytometry profiles, rosette leaves from three-week-old plants were finely chopped in 0.5 ml Galbraith buffer (45 mM $MgCl_2$, 20 mM MOPS, 30 mM sodium citrate, 0.1% Triton X-100, 40 µg/µl RNase A) and filtered through a 30 µm mesh (Sysmex Partec). Nuclei were stained by adding 20 µg/ml propidium iodide (Sigma) to each sample, followed by vortexing. The samples were analyzed using a BD FACS LSR Fortessa X20 or BD FACSAria II sorter (Becton Dickinson, Franklin Lakes, NJ). FlowJo 10.0.6 (Tree Star, Ashland, Oregon) was used to generate profiles and for quantification (nuclei counts and rCV values). Each biological replicate consisted of a leaf from one plant. To sort 16C nuclei for DNA sequencing, samples were prepared by chopping rosette leaves from four-week-old plants as described above. 100,000 nuclei for each sample (two biological replicates per genotype) were sorted using a BD FACSAria II sorter with a 100-µm nozzle.

DNA extraction and sequencing

Genomic DNA was extracted from sorted 16C nuclei using the Arcturus™ PicoPure™ DNA Extraction Kit (ThermoFisher Scientific, Waltham, MA). Samples were incubated at 65°C overnight, and then at 95°C for 10 min. The DNA samples were purified using a genomic DNA Clean and Concentrator kit (Zymo Research, Irvine, CA). Sequencing libraries were generated at the Yale Center for Genome Analysis (YCGA) using the xGen Prism DNA library prep kit for NGS (Integrated DNA Technologies, Coralville, IA). Sequencing was performed on an Illumina NovaSeq 6000 using the S4 XP workflow (Illumina, San Diego, CA). Paired-end reads were filtered and trimmed using fastp (version 0.21.0 with default parameters) (*50*). Reads with quality inferior to 20 were removed from the datasets (Table S4). Sequencing datasets were aligned against the *A. thaliana* genome (TAIR10) using bowtie2 with default parameters (*51*). Duplicate reads were removed using the Picard toolkit (https//broadinstitute.github.io/picard) (MarkDuplicates with *REMOVE_DUPLICATES=true*). The mapped reads were filtered based on mapping quality using samtools (-q 30) (*52*) (Table S4). Biological replicates were analyzed for consistency with deepTools2 (fig. S12A) (*53*). For generating the chromosomal representations, the program featureCounts (version 1.6.4 (*54*)) was used to count the paired-end fragments present in each 200-kb region of the *A. thaliana* genome. As previously described (*55*), the log2 ratio was centered on the average ratio of any two compared libraries (i.e. mutant vs Col) on the first 5 Mbp of chromosome 1 for normalization. Plot profiles were done using R (version 3.6.2) (*56*) and Gviz (*57*).

RNA sequencing

For each biological replicate, leaves from three individual plants growing in the same flat were pooled. Two biological replicates per genotype were sequenced. RNA was extracted from three-week-old leaf tissue using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany). RNA quality was verified using the Agilent 2100 Bioanalyzer Nano RNA Assay. Libraries were prepared at the YCGA with 1 µg of total RNA using Illumina's TruSeq Stranded Total RNA with Ribo-Zero Plant (Illumina). The libraries were amplified with eight PCR cycles, validated using Agilent Bioanalyzer 2100 High sensitivity DNA assay and quantified using the KAPA Library Quantification Kit (Illumina® Platforms). Sequencing was done on an Illumina NovaSeq 6000 using the S4 XP workflow. Paired-end reads were filtered and trimmed using fastp (version 0.21.0

with default parameters) (*50*). Reads with quality inferior to 20 were removed from the data sets (Table S5). Biological replicates were analyzed for consistency with deepTools2 (fig. S12B) (*53*). Data sets were aligned against the *A. thaliana* genome (TAIR10) using STAR (version 2.7.2a) allowing two mismatches (--outFilterMismatchNmax 2) (*58*). Transposable elements (TEs) were defined according to Panda *et al*, 2016 (*59*). featureCounts (version 1.6.4) (*23*) was used to count the paired-end fragments overlapping with TEs. TPM (transcripts per million) values were calculated for TEs. TEs were considered to be upregulated in mutant lines if they showed a ≥2-fold up-regulation as compared to Col in both biological replicates, and had a value of TPM ≥ 5. The heatmap was drawn with the R built-in function (version 3.6.2) (*25*).

Somatic recombination assay

The inverted repeat GUS reporter line used in this study was described previously (*23*). This reporter line was crossed with the following mutants: *atxr5 atxr6, tsk,* and *atxr5 atxr6 tsk*. Plants in the F3 generation homozygous for the GUS reporter gene, and either WT (control) or homozygous mutant for *atxr5/6, tsk*, or *atxr5/6 tsk* were identified. Experiments were performed at least three times in four-week-old F4 plants as previously described (*23*).

MMS genotoxic assay

Seeds were germinated and grown on ½ MS plates with or without 100 µg/ml methyl methanesulfonate (MMS) (Thermo Fisher Scientific) under cool-white fluorescent lights (~100 µmol m^{-2} s^{-1}) in long-day conditions (16 h light/8 h dark). Seedlings were grown on vertically oriented plates for root length measurements. Measurements were done 14 days after germination. The experiments were repeated three times with similar results.

CRISPR

The *rad51* mutant was obtained by multiplex CRISPR/Cas9-mediated deletion of the full *RAD51* gene. Two guide RNAs (*RAD51-F:* GTAGTGTTGTATAAACCACG and *RAD51-R:* AACACCTAGGTATCACTCGG) were designed with CHOPCHOP v3 (*24*) and cloned into an entry vector as described previously (*10*). The resulting AtU6.26:gRNA cassettes were amplified using the PhusionFlash polymerase (ThermoFisher Scientific). A modular cloning (MoClo) reaction (*25*) was then used to clone the F and R amplicons at positions #3 and #4, respectively, of the pAGM65879 acceptor vector (Addgene plasmid #153214; gifted by S. Marillonnet), which provides an *RPS5a*-driven, intron-optimized SpCas9 variant at position #2 (*26*). An *OLE1p:OLE1-RFP* reporter construct was used at position #5 for selection of transformants. Final constructs were agroinfiltrated into Col and *atxr5/6* T0 plants. In both backgrounds, one transgene-free T2 plant heterozygous for the *rad51* deletion was selected by PCR and selfed. The resulting T3 populations were screened for homozygous *rad51* mutants.

CRISPR/Cas9 editing of *HTR9* and *HTR13* used a Level2 MoClo vector. This was constructed using the pAGM4723 acceptor vector containing a YAOp:SaCas9:E9 cassette at position #2 (reverse orientation), AtU6.26:gRNA cassettes targeting *HTR9* (CTCAACGCCACCGTTCCTGG A) and *HTR13* (CTCAAGGCAACAGTTCCTGGA) at positions #3 and #4, an *OLE1p:OLE-RFP* cassette at position #5 and a *Nos*:*Hyg*:*Ocs* at position #6. The YAOp:SaCas9:E9 cassette was

cloned from plasmids LBJJ491 (Addgene plasmid #117513), pEPOR0SP0020 (Addgene plasmid #117531) and pICSL60004 (Addgene plasmid #117519), which were gifts from Jonathan D. Jones (*60*) and Nicola Patron (*61*). The AtU6.26:gRNA cassettes used pICSL90002 (Addgene plasmid #68261), also a gift from Nicola Patron.

All CRISPR/Cas9 editing vectors used components from The MoClo Toolkit (Addgene kit #1000000044) (*62*) and The MoClo Plant Parts Kit (Addgene kit #1000000047) (*63*). Transformations were done by floral dip using *Agrobacterium tumefaciens* GV3101. Transformant T1 seeds were selected using the *OLE1p:OLE-RFP* reporter.

Amplicon sequencing

DNA extracted from T1 plants was amplified and pooled using Custom rhAmpSeq Panels and the rhAmpSeq Library Kit (Integrated DNA Technologies, Coralville, IA). Multiplexed libraries were then sequenced on a NovaSeq 6000 (Illumina) producing paired-end 2×150bp reads. Reads were analyzed using the CRISPResso2 pipeline (*64*).

rhAMpSeq primers	Sequence
HTR9-F	AACTCCTAAAATGGCTCGTAC**rC**AAGC
HTR9-R	AAGCTCAGTACTCTTCTGATACTT**rC**CTGA
HTR13-F	GTTTGATTTCGAAATGGCTCGTA**rC**TAAG
HTR13-R	CAGTGCTCTTCTGATACTTCCT**rG**ATCT

* r indicates RNA bases

Graphic design

The model depicted in Fig. 4G was created with BioRender.com.

Figure S1: Alignment of plant TSK proteins. The alignment was generated with Clustal Omega and represented using

Jalview. NCBI reference sequences: GAY58445.1 (*Citrus unshiu*), NP_188503.2 (*A. thaliana*), XP_006585323.1 (*Glycine max*), XP_015624059.1 (*Oryza sativa*), XP_024378964.1 (*Physcomitrella patens*), and XP_024518191.1 (*Selaginella moellendorffii*). Dark blue, blue and light blue residues indicate 100%, 80%, and 60% identity, respectively, across all six protein sequences. The sequences corresponding to the TPR domain are indicated.

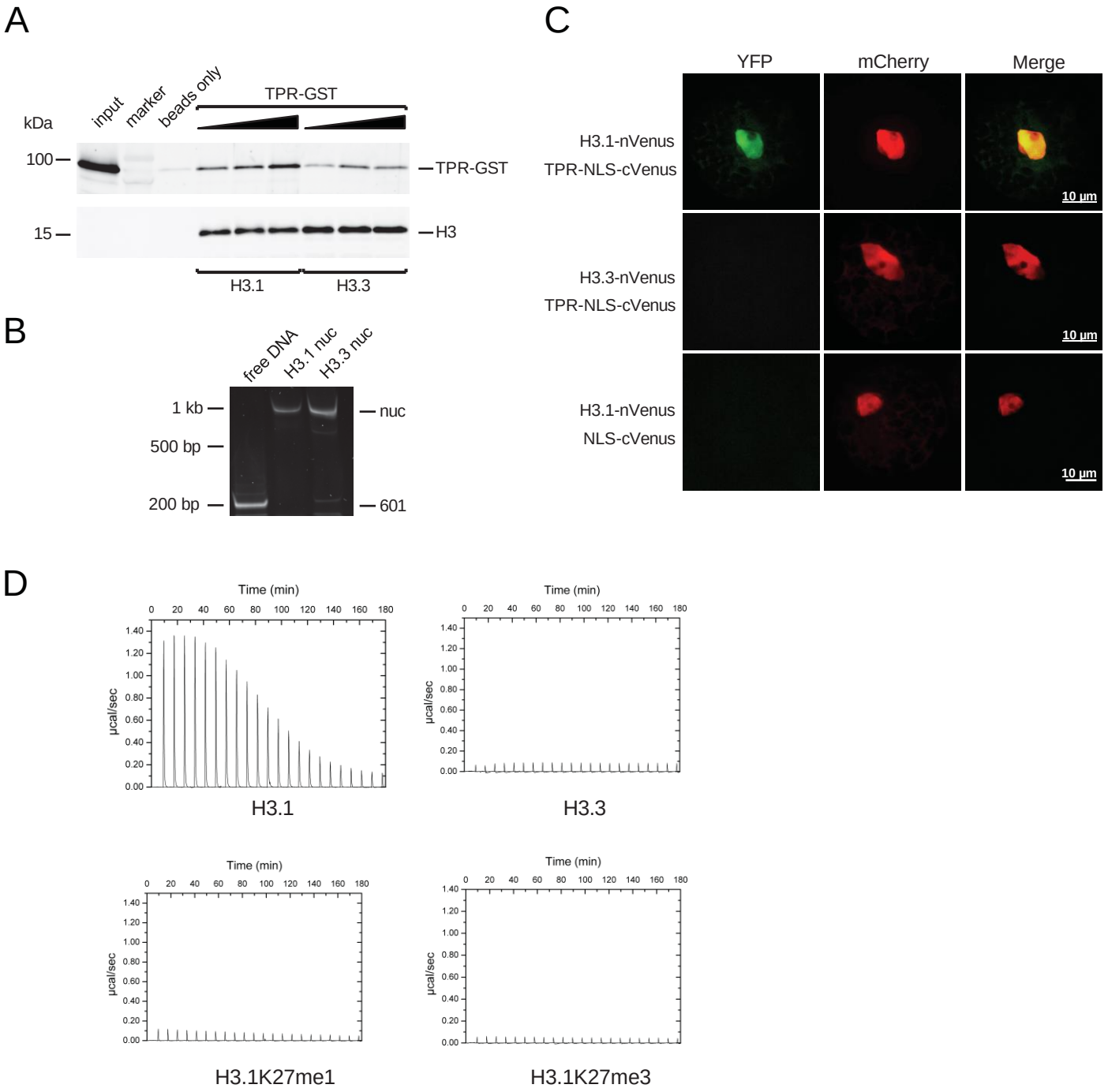


Figure S2: The TPR domain of TSK specifically interacts with H3.1 in a nucleosomal context and in protoplasts. (A) Nucleosome pulldown assays with recombinant mononucleosomes containing either *A. thaliana* H3.1 or H3.3 and increasing amounts of GST-tagged TPR_{TSK} (0.5, 1, and 2-fold molar ratio to nucleosomes). Nucleosome loading and binding of TPR were assessed by Western Blot with anti-H3 and anti-GST antibodies, respectively. (B) Native gel electrophoresis was performed to verify assembly of recombinant nucleosomes used for pulldown experiments shown in panel A. Mononucleosomes (nuc) were separated from free 601 DNA (601) on a native 6% polyacrylamide gel and visualized with SYBR safe stain. (C) Bimolecular fluorescence complementation (BiFC) assay in *A. thaliana* protoplasts. H3.1 or H3.3 fused to the N-terminus of Venus (YFP) and $AtTPR_{TSK}$-NLS (nuclear localization signal) fused to the C-terminus of Venus were co-transformed into protoplasts. mCherry-VirD2NLS was co-expressed as a nuclear marker. H3.1-nVenus and

NLS-cVenus were co-transformed into protoplasts as a negative control. (D) Thermograms of ITC assays using $CuTSK_{(1-530)}$ and H3.1, H3.3, H3.1K27me1, and H3.1K27me3 peptides.

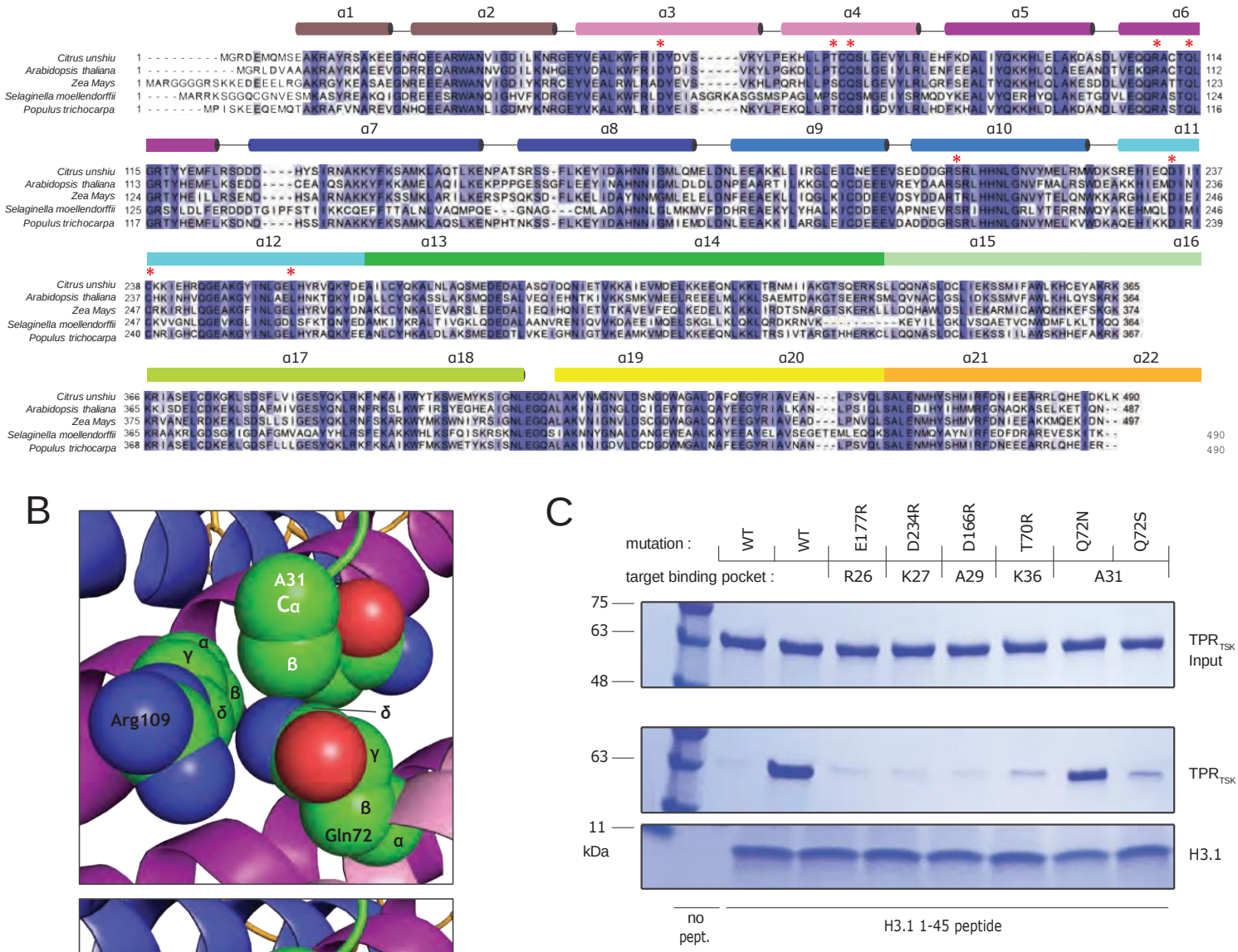


Figure S3: Analysis of amino acid residues of TPR$_{TSK}$ interacting with H3.1 (A) Sequence alignment of the TPR domain of TSK from multiple plant species. Red asterisks indicate the CuTPR$_{TSK}$ residues which interact with H3.1 K27, K36, and A31. Cylinders atop the sequences mark boundaries of CuTPR$_{TSK}$ sequences of individual TPR folds. (B) Spherical representation of the atomic radii of H3.1A31 in its binding pocket on CuTPR$_{TSK}$. (Top panel) A31 is shown with its Cβ bordering the amine group of Gln72. (Bottom panel) Threonine replacing alanine at position 31 (as in H3.3) produces van der Waals clashes between T31-C γ and Gln72-Cγ/Cδ/amine group, between T31-OH and Arg109-Cβ/Cγ, and between T31-OH and Gln72-amine group. (C) Streptavidin pulldowns with biotin-H3.1$_{(1-45)}$ and

various CuTSK$_{(1-530)}$ point mutants targeting different H3.1 binding pockets.

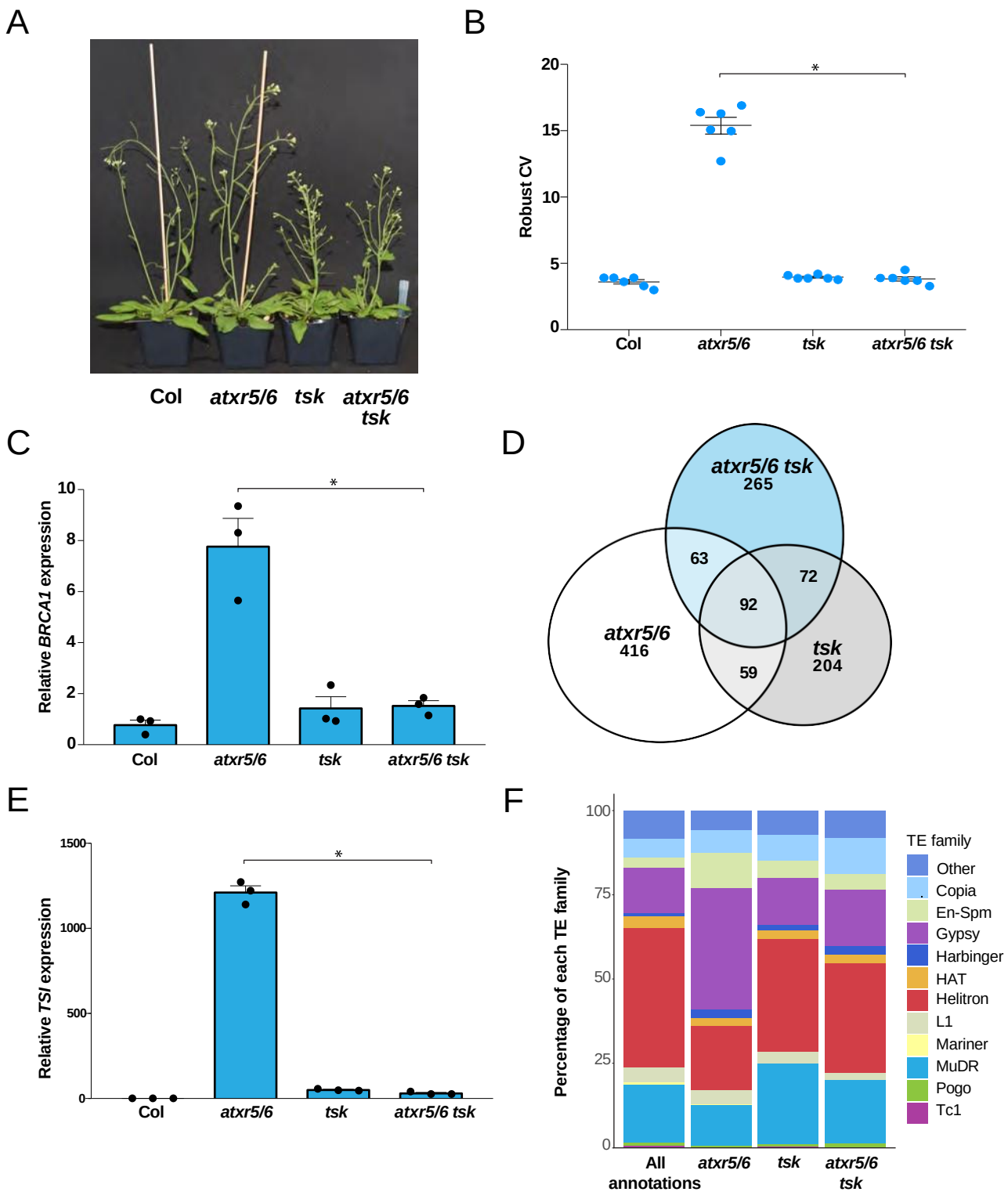


Figure S4: Effect of *TSK* on genome stability and transcriptional de-repression in *atxr5/6* mutants. (A) Morphological phenotypes of *atxr5/6*, *tsk* and *atxr5/6 tsk*. (B) Robust CV values for 16C nuclei obtained by flow cytometry analysis. Each dot represents a biological replicate. Horizontal bars indicate the mean. SEM are shown. The asterisk indicates a significant difference as determined by a Brown-Forsythe and Welch ANOVA test followed by the Dunnett T3 test for multiple comparisons: * $p < 0.0001$ (C) RT-qPCR analysis of the genome stability marker *BRCA1* in Col, *atxr5/6*, *tsk* and *atxr5/6 tsk*. The average of three biological replicates and SEM are shown. Unpaired *t*-test: * $p < 0.01$. (D) Venn diagram showing the number of upregulated TEs (≥2-fold change) in *atxr5/6*, *tsk* and *atxr5/6 tsk* compared to Col plants ($p_{adj} < 0.05$). (E) RT-qPCR analysis of the DNA repeat *TSI* in Col,

atxr5/6, *tsk* and *atxr5/6 tsk*. The average of three biological replicates and SEM are shown. Unpaired *t*-test: * $p <$ 0.0001. (F) Distribution of reactivated TEs in *atxr5/6*, *tsk*, and *atxr5/6 tsk* among the different TE families.

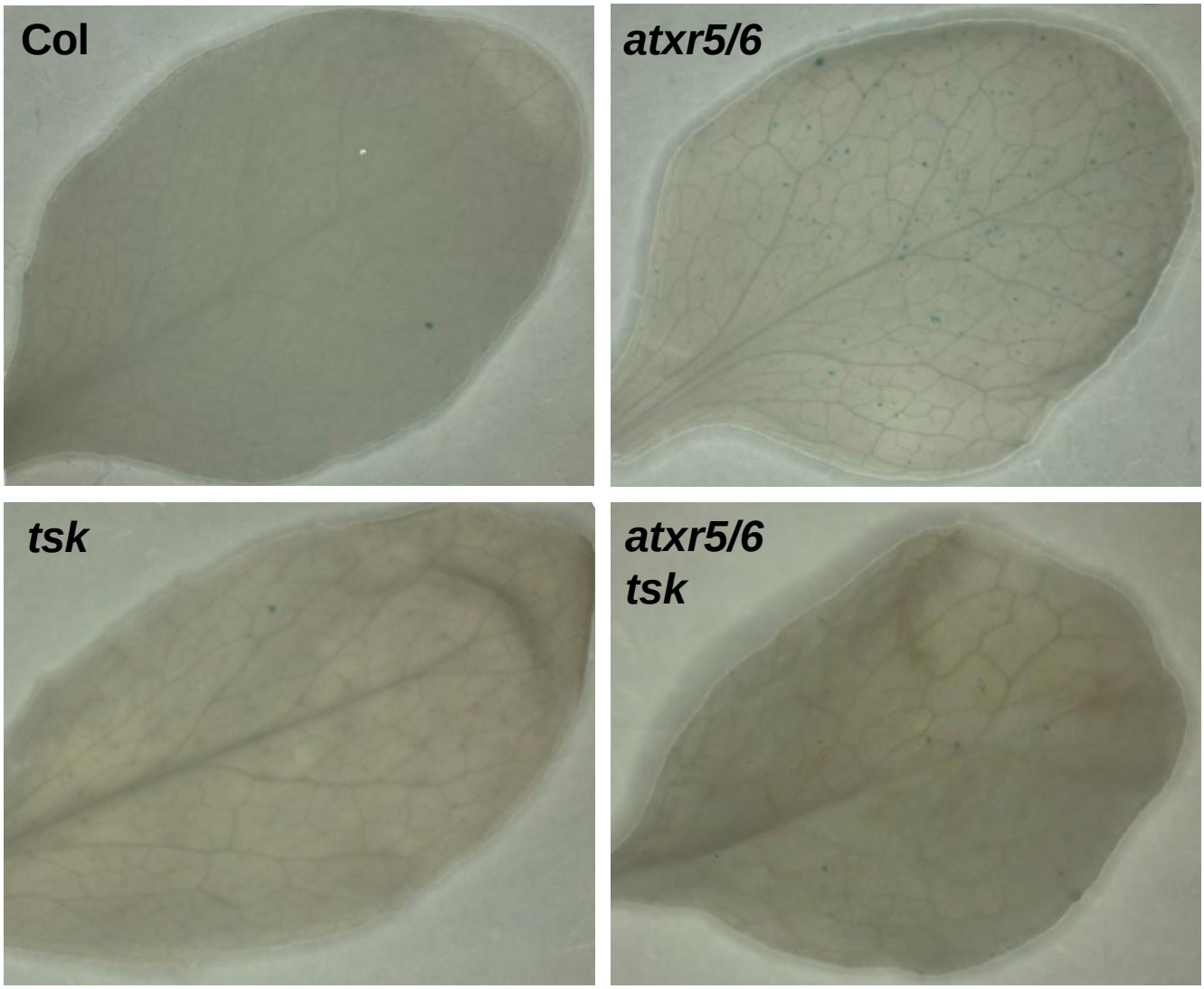


Figure S5: Increased levels of homologous recombination in *atxr5/6* mutants. Representative images of somatic recombination events (GUS activity) detected in the leaves of Col, *atxr5/6*, *tsk* and *atxr5/6 tsk*.

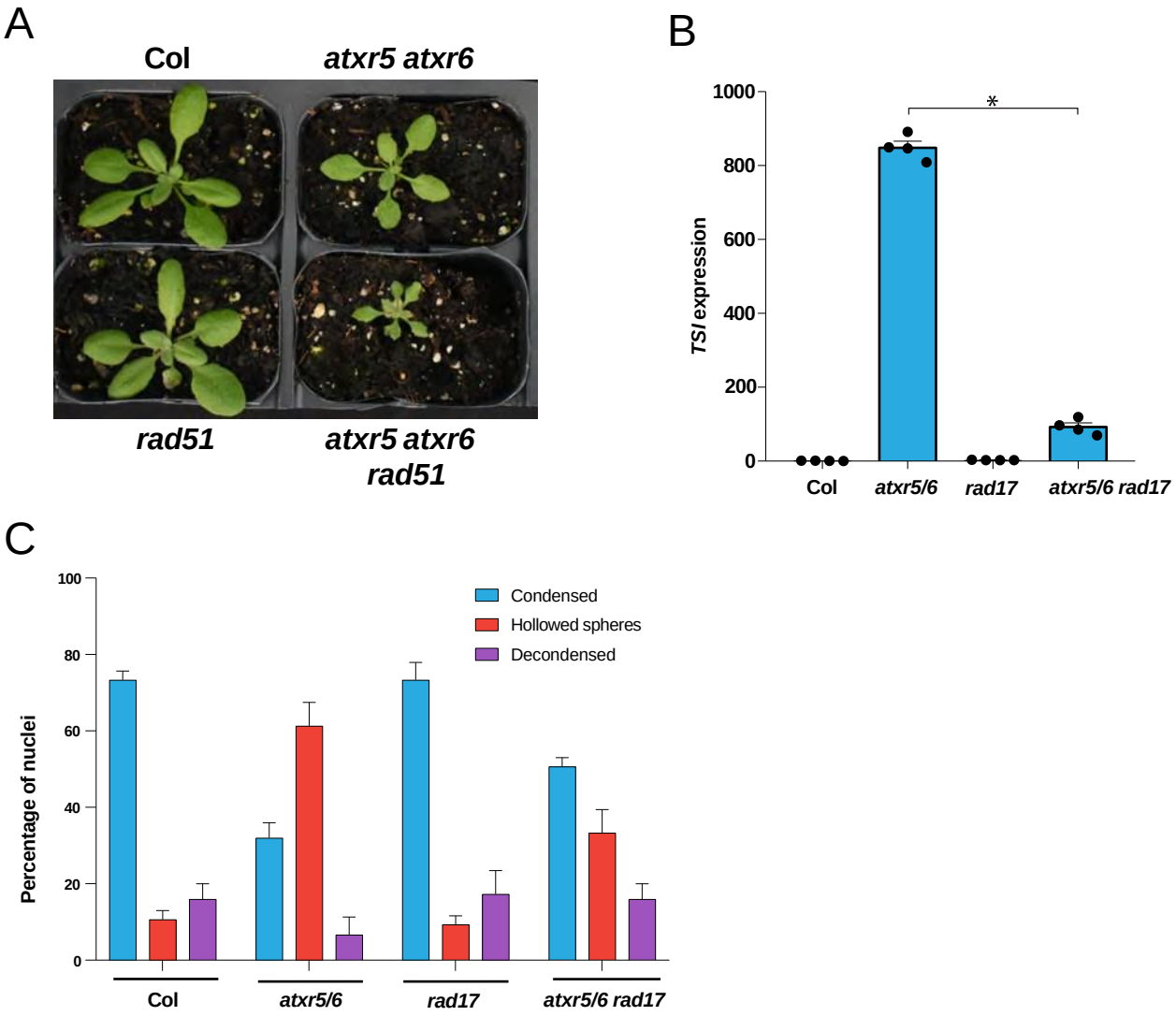


Figure S6: Inactivating *RAD17* in *atxr5/6* mutants suppresses heterochromatic defects. (A) Morphological phenotypes of *atxr5/6*, *rad51* and *atxr5/6 rad51*. (B) RT-qPCR analyses of the repetitive element *TSI*. Each dot represents an independent biological replicate. The average of four biological replicates and SEM are shown. Unpaired *t*-test: * $p < 0.0001$. (C) Quantification of chromocenter appearance from DAPI-stained nuclei. Shown is the percentage of nuclei that are fully condensed, displaying a hollowed sphere conformation and fully/partially decondensed. Twenty-five nuclei for three biological replicates of each genotype were assessed. Error bars indicate SEM.

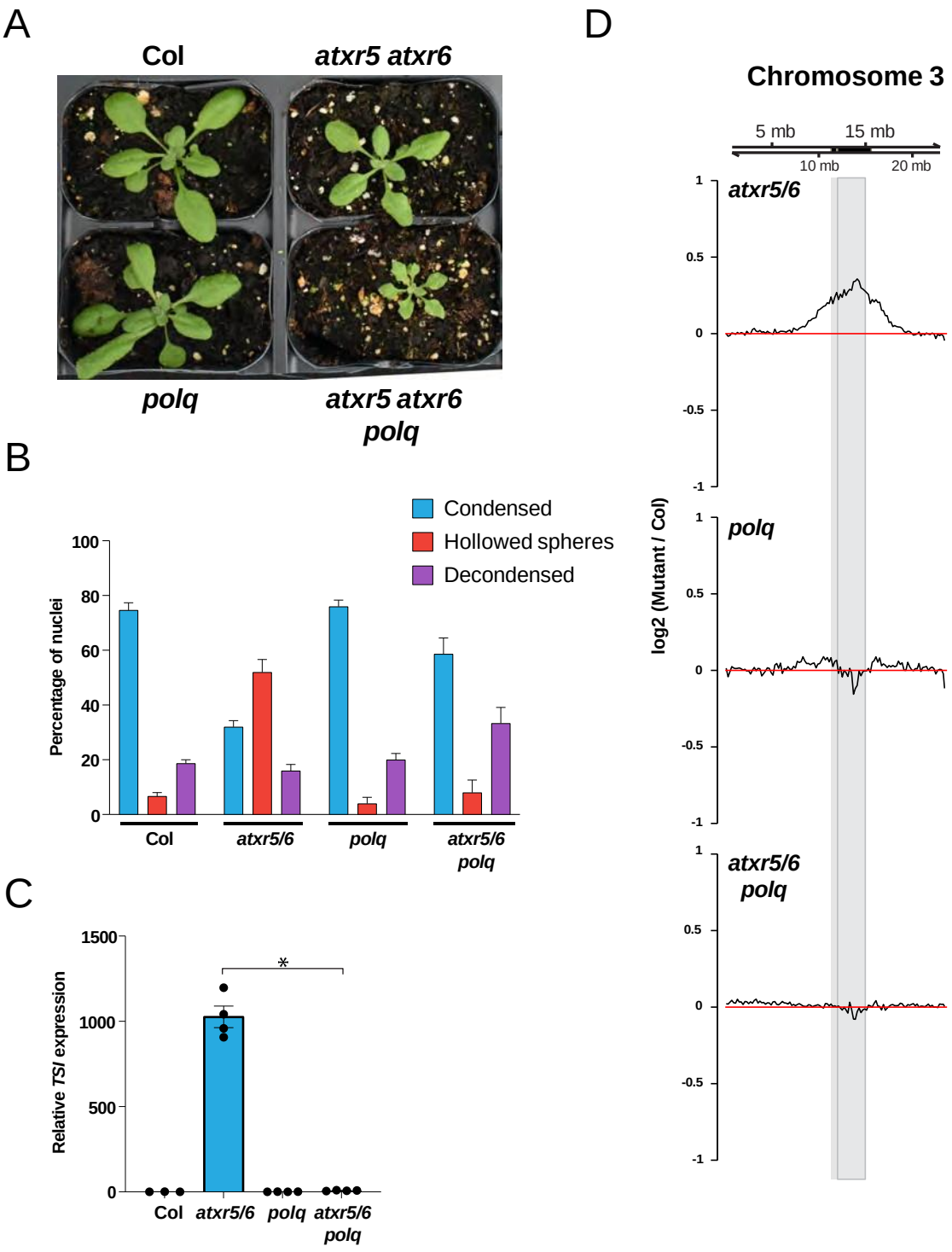


Figure S7: A mutation in *POLQ* suppresses the heterochromatin phenotypes of *atxr5/6* mutants. (A) Morphological phenotypes of *atxr5/6*, *polq* and *atxr5/6 polq*. (B) Quantification of chromocenter appearance from DAPI-stained nuclei. Shown is the percentage of nuclei that are fully condensed, displaying a hollowed sphere conformation and fully/partially deconsensed. Twenty-five nuclei for three biological replicates of each genotype were assessed. Error bars indicate SEM. (C) RT-qPCR analyses of the repetitive element *TSI*. Each dot represents an independent biological replicate. The average of at least three biological replicates and SEM are shown. Unpaired *t*-test: * $p < 0.0001$. (D). Chromosomal view

(chromosome 3 of *A. thaliana*) of DNA sequencing reads from *atxr5/6*, *polq* and *atxr5/6 polq* sorted nuclei, normalized to reads from Col. The centromeric region is highlighted in gray.

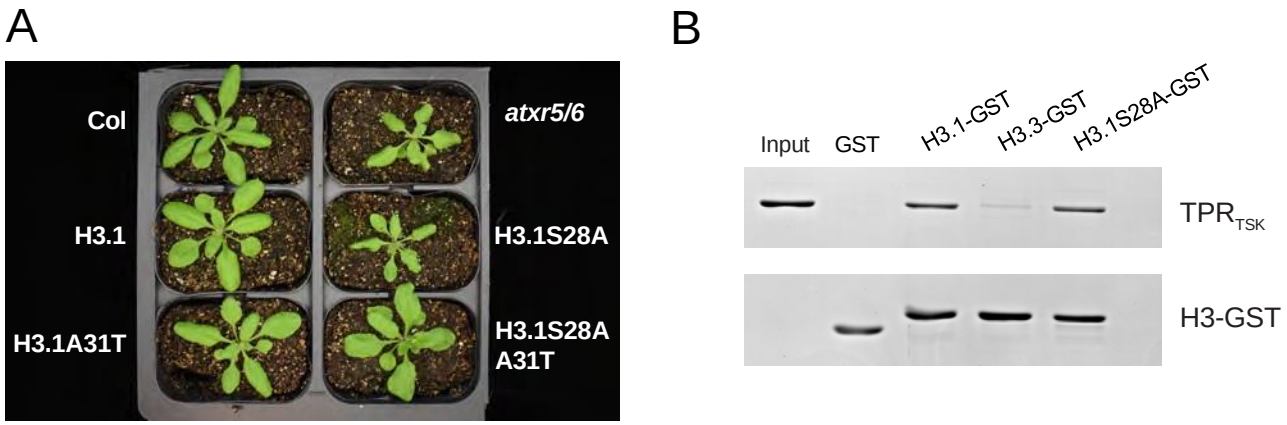


Figure S8: Expression of the histone point mutant H3.1S28A in *A. thaliana* does not interfere with binding of TSK to H3.1. (A) Morphological phenotypes of T1 plants expressing different H3.1 transgenes. (B) Pulldown assay using the TPR domain of TSK and GST-tagged histones H3.1, H3.3 and H3.1S28A.

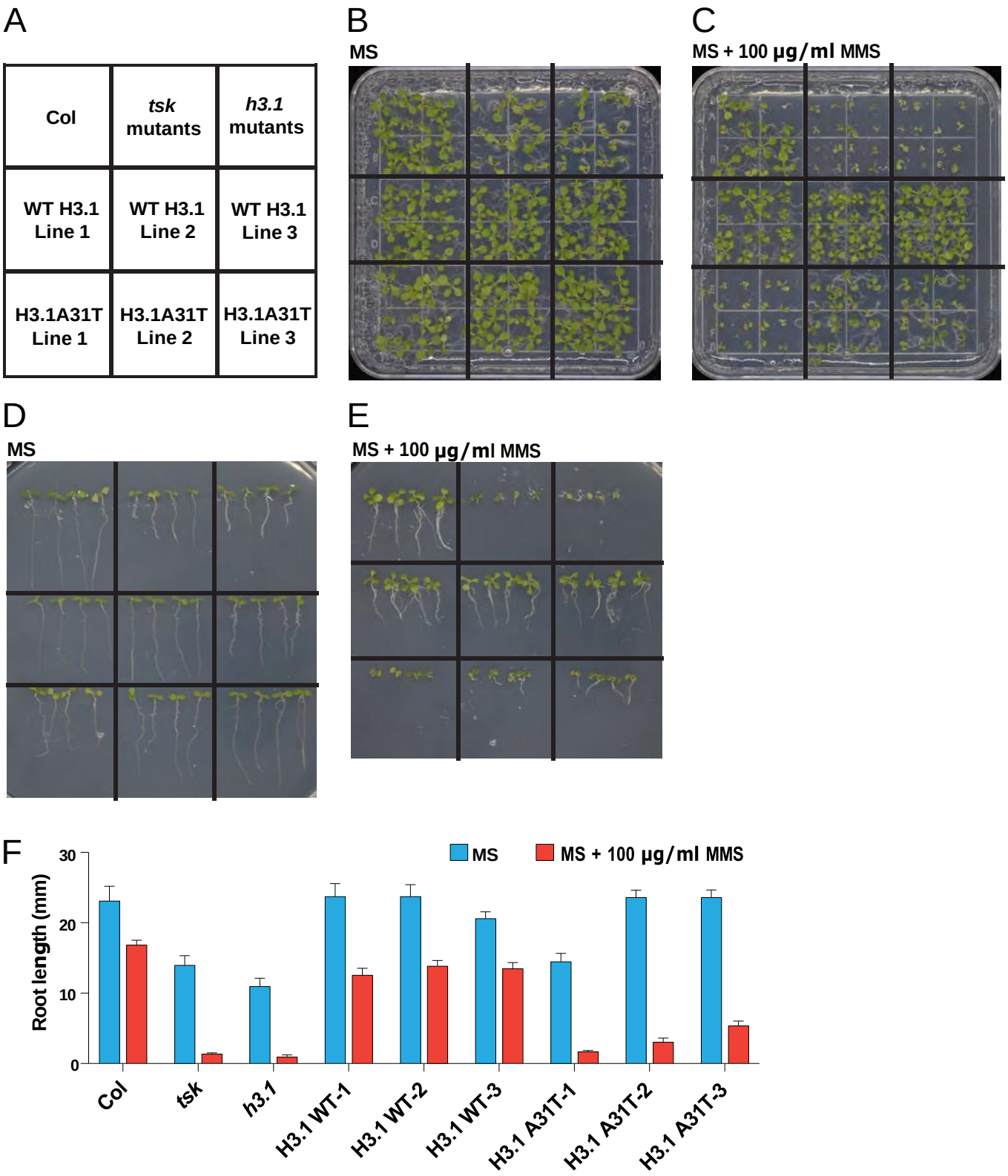


Figure S9: Sensitivity of *tsk* and *h3.1* mutants, and H3.1A31T-expressing plants, to genotoxic stress. (A) Layout of plant genotypes grown on plates shown in B-E. H3.1 WT lines 1-3 and H3.1A31T lines 1-3 are T4 transgenic plants from independent T1 parents. (B-C) Representative seedlings grown on horizontally-oriented ½ MS plates in the absence (B) or the presence (C) of 100 µg/ml MMS. (D-E) Representative seedlings grown on vertically-oriented plates in the absence (D) or the presence (E) of

MMS. (F) Root length of seedlings grown with or without MMS. SEM is shown. Eight seedlings were measured for each genotype.

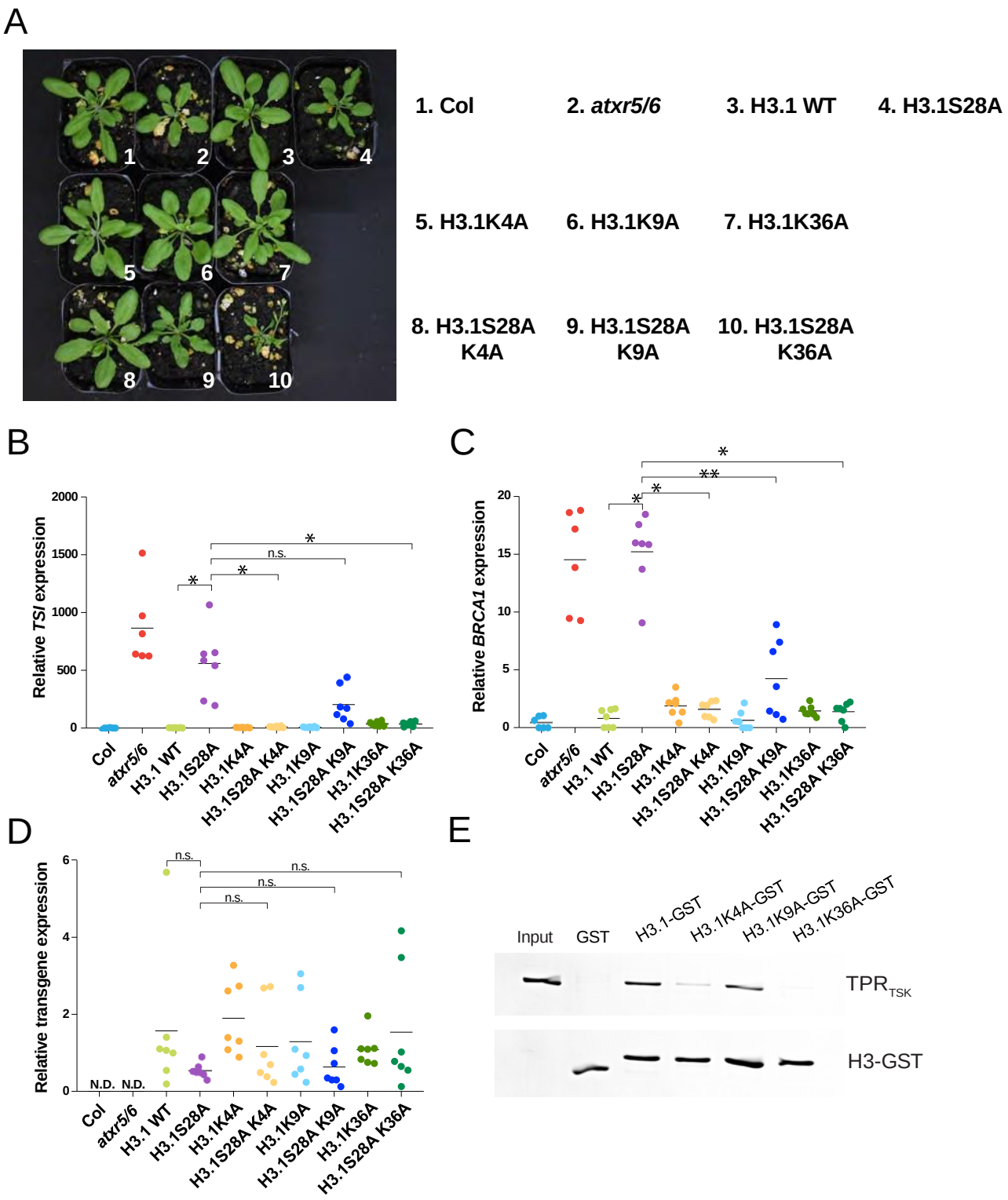


Figure S10: Effects of alanine replacement at K4, K9 and K36 of H3.1 on genomic stability. (A) Morphological phenotypes of T1 plants expressing different *H3.1* transgenes. (B-D) RT-qPCR analyses of the repetitive element *TSI* (B), *BRCA1* (C) and the *H3.1* transgene (D) in Col, *atxr5/6* mutants and the *H3.1* replacement lines. For Col and *atxr5/6*, each dot represents an independent biological replicate. For the *H3.1* lines, each dot represents one T1 plant. Horizontal bars indicate the mean. N.D. = not detected. The asterisks indicate a significant difference as determined by a Brown-Forsythe and Welch ANOVA test followed by the Dunnett T3 test for multiple comparisons: * $p < 0.05$, ** $p < 0.001$,

*** $p < 0.0005$ and n.s. = not significantly different. (E) Pulldown assay using the TPR domain of TSK and GST-tagged histones H3.1, H3.1K4A, H3.1K9A and H3.1K36A.

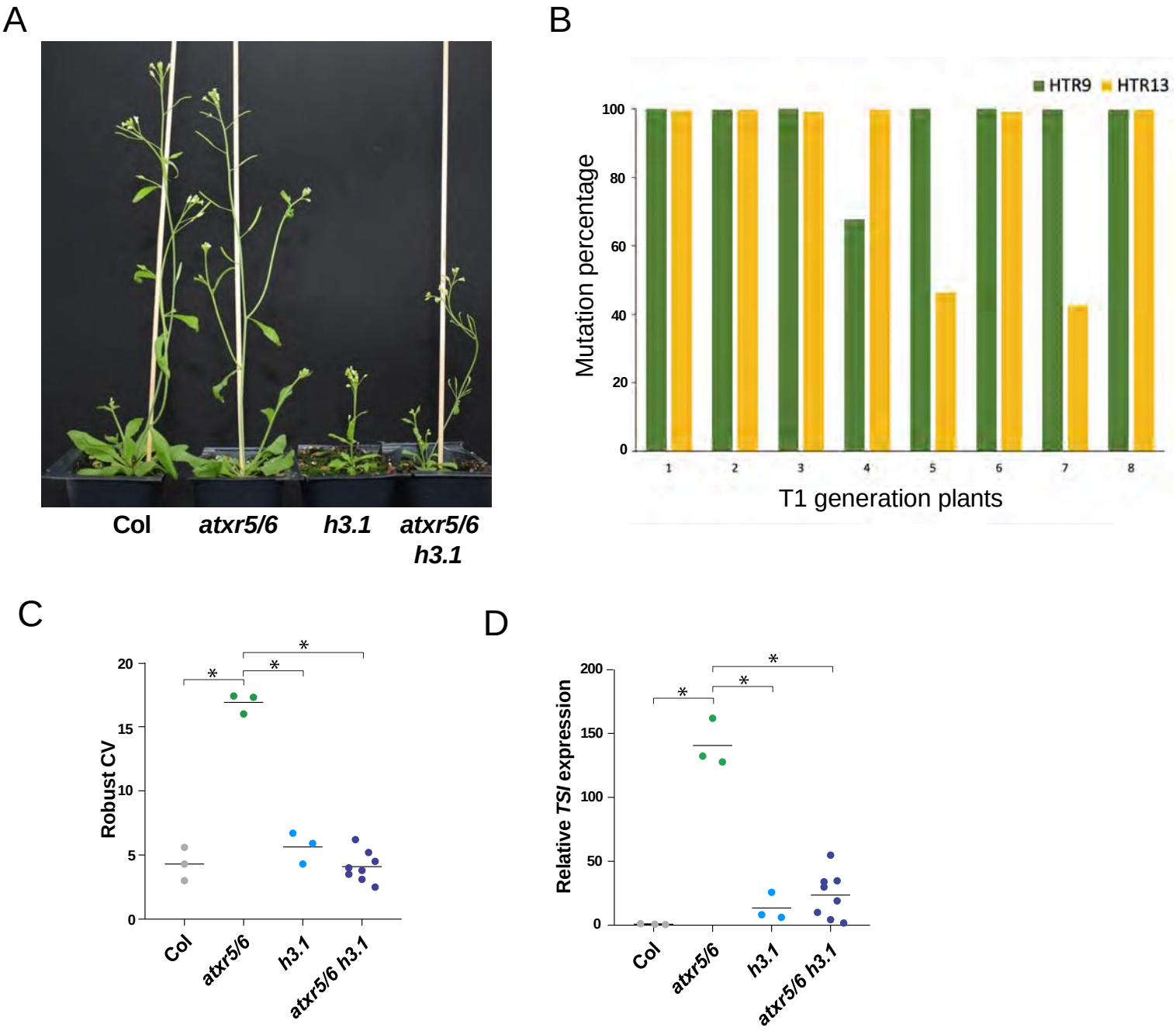


Figure S11: H3.1 is required to induce heterochromatin amplification and transcriptional reactivation in the absence of ATXR5/6. (A) Morphological phenotypes of *atxr5/6*, *h3.1* and *atxr5/6 h3.1* mutants. (B) Percentage of mutated *HTR9* and *HTR13* alleles in individual T1 plants (*atxr5/6 h3.1*) as determined by amplicon sequencing. (C) rCV values for 16C nuclei obtained by flow cytometry analysis. For Col, *atxr5/6*, and *h3.1*, each dot represents a biological replicate. For the *atxr5/6 h3.1* CRISPR lines, each dot represents one T1 plant. Horizontal bars indicate the mean. The asterisks indicate a significant difference as determined by a Brown-Forsythe and Welch ANOVA test followed by the Dunnett T3 test for multiple comparisons: * $p < 0.005$. (D) RT-qPCR analyses of the repetitive element *TSI*. Each dot represents an independent biological replicate. The average of at least three biological replicates and SEM are shown.The asterisks indicate a significant difference as determined by a Brown-Forsythe and Welch

ANOVA test followed by the Dunnett T3 test for multiple comparisons: * $p < 0.05$.

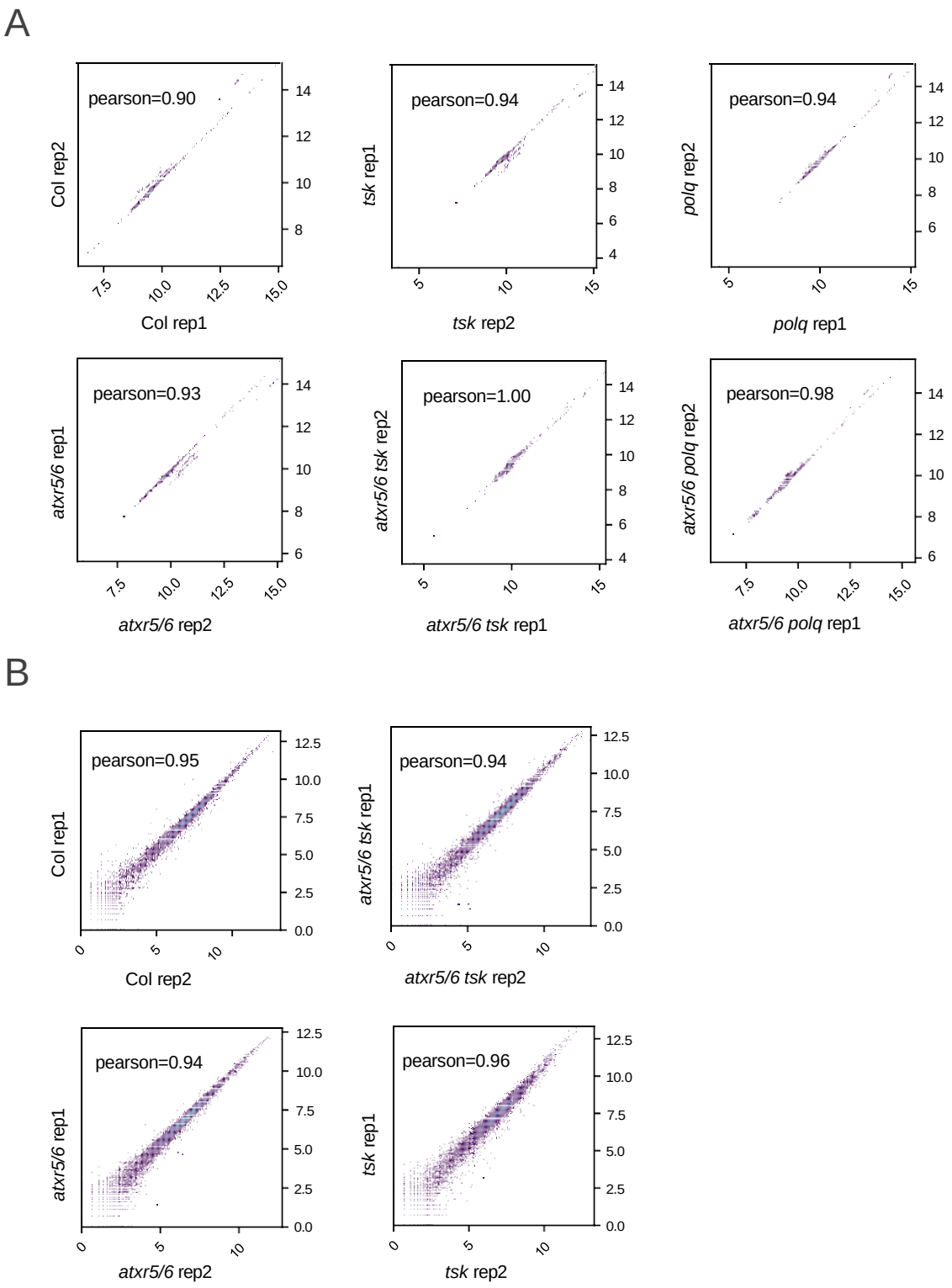


Figure S12: Scatterplots and Pearson correlation coefficients for the sequencing biological replicates. (A) DNA-seq replicates of Col, *atxr5/6*, *tsk*, *atxr5/6 tsk*, *polq* and *atxr5/6 polq*. (B) RNA-seq replicates of Col, *atxr5/6*, *tsk* and *atxr5/6 tsk*.

Table S1: Data collection and refinement statistics for TPR_{TSK} and H3.1 complex.

PDB accession number	7T7T
Data Collection	
Space group	P 2_1 2_1 2_1
Cell dimensions	
a, b, c (Å)	82.89, 92.33, 209.91
α, β, γ (°)	90, 90, 90
Resolution	37.21 - 3.17 (3.28 - 3.17)
R_{meas}	0.05 (0.33)
R_{pim}	0.01
I / σI	19.5 (2.6)
No. unique reflections	27369 (2515)
Completeness (%)	99.9 (99.9)
Redundancy	26.1 (26.1)
CC1/2	0.99 (0.92)
Wilson *B*-Factor	31
Refinement	
Resolution (Å)	37.21 - 3.17
No. reflections	27365 (2515)
R_{work} / R_{free}	0.27 / 0.31
Num. atoms	
TPR^{TSK}	6832
$H3.1^{1-45}$	343
Water	1
B-factors (Å^2)	
TPR^{TSK}	29
$H3.1^{1-45}$	28
R.m.s. deviations	
Bond lengths (Å)	0.004
Bond angles (°)	0.74
Molprobity score	1.76
Clashscore	8.67
Ramachandran favored (%)	95.74
Ramachandran allowed (%)	4.26
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	0.00

* Highest-resolution shell is shown in parentheses.

Table S3: Summary of synthesized histone H3 peptides. Peptide synthesis based on the main backbone peptide, and further modifications with different amino acids in positions "1", "2", and "3". Most relevant mass spectra molecular ions identified are presented.

Main Peptide Backbone					
ARTKQTARKSTGGKAPRKQLATKAAR-**1**-SAP-**2**-TGGVKKPHR-**3**-RPGTY-NH_2					
Peptide	**Amino Acid Modification**			**Theorical Molecular Ions**	**Identified Molecular Ions**
	1	**2**	**3**		
H3.1	Lys	Ala	Phe	[M+7H]7+: 703.7 [M+8H]8+: 615.7 [M+9H]9+: 547.4	[M+7H]7+: 703.6 [M+8H]8+: 615.7 [M+9H]9+: 547.5
H3.3	Lys	Thr	Tyr	[M+7H]7+: 710.1 [M+8H]8+: 621.5 [M+9H]9+: 552.7	[M+7H]7+: 710.1 [M+8H]8+: 621.5 [M+9H]9+: 552.5
H3.1-K27Me1	Lys(Me)	Ala	Phe	[M+7H]7+: 705.6 [M+8H]8+: 617.5 [M+9H]9+: 549.0	[M+7H]7+: 705.7 [M+8H]8+: 617.5 [M+9H]9+: 549.1
H3.1-K27Me3	$Lys(Me)_3$	Ala	Phe	[M+7H]7+: 709.6 [M+8H]8+: 621.0 [M+9H]9+: 552.1	[M+7H]7+: 709.8 [M+8H]8+: 621.1 [M+9H]9+: 552.3

References and Notes

1. E. Duro, C. Lundin, K. Ask, L. Sanchez-Pulido, T. J. MacArtney, R. Toth, C. P. Ponting, A. Groth, T. Helleday, J. Rouse, Identification of the MMS22L-TONSL complex that promotes homologous recombination. *Mol. Cell* **40**, 632–644 (2010).

2. T. H. Huang, F. Fowler, C. C. Chen, Z. J. Shen, B. Sleckman, J. K. Tyler, The Histone Chaperones ASF1 and CAF-1 Promote MMS22L-TONSL-Mediated Rad51 Loading onto ssDNA during Homologous Recombination in Human Cells. *Mol. Cell* **69**, 879–892.e5 (2018).

3. B. C. O'Connell, B. Adamson, J. R. Lydeard, M. E. Sowa, A. Ciccia, A. L. Bredemeyer, M. Schlabach, S. P. Gygi, S. J. Elledge, J. W. Harper, A genome-wide camptothecin sensitivity screen identifies a mammalian MMS22L-NFKBIL2 complex required for genomic stability. *Mol. Cell* **40**, 645–657 (2010).

4. L. O'Donnell, S. Panier, J. Wildenhain, J. M. Tkach, A. Al-Hakim, M. C. Landry, C. Escribano-Diaz, R. K. Szilard, J. T. Young, M. Munro, M. D. Canny, N. K. Kolas, W. Zhang, S. M. Harding, J. Ylanko, M. Mendez, M. Mullin, T. Sun, B. Habermann, A. Datti, R. G. Bristow, A. C. Gingras, M. D. Tyers, G. W. Brown, D. Durocher, The MMS22L-TONSL complex mediates recovery from replication stress and homologous recombination. *Mol. Cell* **40**, 619–631 (2010).

5. W. Piwko, L. J. Mlejnkova, K. Mutreja, L. Ranjha, D. Stafa, A. Smirnov, M. M. Brodersen, R. Zellweger, A. Sturzenegger, P. Janscak, M. Lopes, M. Peter, P. Cejka, The MMS22L-TONSL heterodimer directly promotes RAD51-dependent recombination upon replication stress. *EMBO J.* **35**, 2584–2601 (2016).

6. W. Piwko, M. H. Olma, M. Held, J. N. Bianco, P. G. Pedrioli, K. Hofmann, P. Pasero, D. W. Gerlich, M. Peter, RNAi-based screening identifies the Mms22L-Nfkbil2 complex as a novel regulator of DNA replication in human cells. *EMBO J.* **29**, 4210–4222 (2010).

7. T. Suzuki, S. Inagaki, S. Nakajima, T. Akashi, M. A. Ohto, M. Kobayashi, M. Seki, K. Shinozaki, T. Kato, S. Tabata, K. Nakamura, A. Morikami, A novel Arabidopsis gene TONSOKU is required for proper cell arrangement in root and shoot apical meristems. *Plant J.* **38**, 673–684 (2004).

8. S. Takeda, Z. Tadele, I. Hofmann, A. V. Probst, K. J. Angelis, H. Kaya, T. Araki, T. Mengiste, O. Mittelsten Scheid, K. Shibahara, D. Scheel, J. Paszkowski, BRU1, a novel link between responses to DNA damage and epigenetic gene silencing in Arabidopsis. *Genes Dev.* **18**, 782–793 (2004).

9. G. Saredi, H. Huang, C. M. Hammond, C. Alabert, S. Bekker-Jensen, I. Forne, N. Reverón-Gómez, B. M. Foster, L. Mlejnkova, T. Bartke, P. Cejka, N. Mailand, A. Imhof, D. J. Patel, A. Groth, H4K20me0 marks post-replicative chromatin and recruits the TONSL–MMS22L DNA repair complex. *Nature* **534**, 714–718 (2016).

10. J. Fang, Q. Feng, C. S. Ketel, H. Wang, R. Cao, L. Xia, H. Erdjument-Bromage, P. Tempst, J. A. Simon, Y. Zhang, Purification and functional characterization of SET8, a nucleosomal histone H4-lysine 20-specific methyltransferase. *Curr. Biol.* **12**, 1086–1099 (2002).

11. K. Nishioka, J. C. Rice, K. Sarma, H. Erdjument-Bromage, J. Werner, Y. Wang, S. Chuikov, P. Valenzuela, P. Tempst, R. Steward, J. T. Lis, C. D. Allis, D. Reinberg, 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 (2002).

12. J. C. Rice, K. Nishioka, K. Sarma, R. Steward, D. Reinberg, C. D. Allis, Mitotic-specific methylation of histone H4 Lys 20 follows increased PR-Set7 expression and its localization to mitotic chromosomes. *Genes Dev.* **16**, 2225–2230 (2002).

13. S. Guyomarc'h, T. Vernoux, J. Traas, D. X. Zhou, M. Delarue, MGOUN3, an Arabidopsis gene with TetratricoPeptide-Repeat-related motifs, regulates meristem cellular organization. *J. Exp. Bot.* **55**, 673–684 (2004).

14. A. Perez-Riba, L. S. Itzhaki, The tetratricopeptide-repeat motif is a versatile platform that enables diverse modes of molecular recognition. *Curr. Opin. Struct. Biol.* **54**, 43–49 (2019).

15. L. Lu, X. Chen, S. Qian, X. Zhong, The plant-specific histone residue Phe41 is important for genome-wide H3.1 distribution. *Nat. Commun.* **9**, 630 (2018).

16. Y. Jacob, E. Bergamin, M. T. Donoghue, V. Mongeon, C. LeBlanc, P. Voigt, C. J. Underwood, J. S. Brunzelle, S. D. Michaels, D. Reinberg, J. F. Couture, R. A.

Martienssen, Selective methylation of histone H3 variant H3.1 regulates heterochromatin replication. *Science* **343**, 1249–1253 (2014).

17. Y. Jacob, S. Feng, C. A. LeBlanc, Y. V. Bernatavichute, H. Stroud, S. Cokus, L. M. Johnson, M. Pellegrini, S. E. Jacobsen, S. D. Michaels, ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. *Nat. Struct. Mol. Biol.* **16**, 763–768 (2009).

18. C. Raynaud, R. Sozzani, N. Glab, S. Domenichini, C. Perennes, R. Cella, E. Kondorosi, C. Bergounioux, Two cell-cycle regulated SET-domain proteins interact with proliferating cell nuclear antigen (PCNA) in Arabidopsis. *Plant J.* **47**, 395–407 (2006).

19. Y. Jacob, H. Stroud, C. Leblanc, S. Feng, L. Zhuo, E. Caro, C. Hassel, C. Gutierrez, S. D. Michaels, S. E. Jacobsen, Regulation of heterochromatic DNA replication by histone H3 lysine 27 methyltransferases. *Nature* **466**, 987–991 (2010).

20. W. Feng, C. J. Hale, R. S. Over, S. J. Cokus, S. E. Jacobsen, S. D. Michaels, Large-scale heterochromatin remodeling linked to overreplication-associated DNA damage. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 406–411 (2017).

21. H. Stroud, C. J. Hale, S. Feng, E. Caro, Y. Jacob, S. D. Michaels, S. E. Jacobsen, DNA methyltransferases are required to induce heterochromatic re-replication in Arabidopsis. *PLOS Genet.* **8**, e1002808 (2012).

22. T. Suzuki, S. Nakajima, S. Inagaki, M. Hirano-Nakakita, K. Matsuoka, T. Demura, H. Fukuda, A. Morikami, K. Nakamura, TONSOKU is expressed in S phase of the cell cycle and its defect delays cell cycle progression in Arabidopsis. *Plant Cell Physiol.* **46**, 736–742 (2005).

23. J. M. Lucht, B. Mauch-Mani, H. Y. Steiner, J. P. Metraux, J. Ryals, B. Hohn, Pathogen stress increases somatic recombination frequency in Arabidopsis. *Nat. Genet.* **30**, 311–314 (2002).

24. R. M. Williams, X. Zhang, Roles of ATM and ATR in DNA double strand breaks and replication stress. *Prog. Biophys. Mol. Biol.* **163**, 109–119 (2021).

25. Q. Wang, M. Goldstein, P. Alexander, T. P. Wakeman, T. Sun, J. Feng, Z. Lou, M. B. Kastan, X. F. Wang, Rad17 recruits the MRE11-RAD50-NBS1 complex to regulate the cellular response to DNA double-strand breaks. *EMBO J.* **33**, 862–877 (2014).

26. J. A. Kamp, R. van Schendel, I. W. Dilweg, M. Tijsterman, BRCA1-associated structural variations are a consequence of polymerase theta-mediated end-joining. *Nat. Commun.* **11**, 3615 (2020).

27. J. Dong, C. LeBlanc, A. Poulet, B. Mermaz, G. Villarino, K. M. Webb, V. Joly, J. Mendez, P. Voigt, Y. Jacob, H3.1K27me1 maintains transcriptional silencing and genome stability by preventing GCN5-mediated histone acetylation. *Plant Cell* **33**, 961–979 (2021).

28. E. I. Campos, A. H. Smits, Y. H. Kang, S. Landry, T. M. Escobar, S. Nayak, B. M. Ueberheide, D. Durocher, M. Vermeulen, J. Hurwitz, D. Reinberg, Analysis of the Histone H3.1 Interactome: A Suitable Chaperone for the Right Event. *Mol. Cell* **60**, 697–709 (2015).

29. K. Brzezinka, S. Altmann, I. Bäurle, BRUSHY1/TONSOKU/MGOUN3 is required for heat stress memory. *Plant Cell Environ.* **42**, 771–781 (2019).

30. S. Valuchova, J. Fulnecek, Z. Prokop, P. Stolt-Bergner, E. Janouskova, C. Hofr, K. Riha, Protection of Arabidopsis Blunt-Ended Telomeres Is Mediated by a Physical Association with the Ku Heterodimer. *Plant Cell* **29**, 1533–1545 (2017).

31. M. L. Heacock, R. A. Idol, J. D. Friesner, A. B. Britt, D. E. Shippen, Telomere dynamics and fusion of critically shortened telomeres in plants lacking DNA ligase IV. *Nucleic Acids Res.* **35**, 6490–6500 (2007).

32. F. Heitzeberg, I. P. Chen, F. Hartung, N. Orel, K. J. Angelis, H. Puchta, The Rad17 homologue of Arabidopsis is involved in the regulation of DNA damage repair and homologous recombination. *Plant J.* **38**, 954–968 (2004).

33. S. Wang, W. E. Durrant, J. Song, N. W. Spivey, X. Dong, Arabidopsis BRCA2 and RAD51 proteins are specifically involved in defense gene transcription during plant immune responses. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 22716–22721 (2010).

34. K. Osakabe, K. Abe, T. Yoshioka, Y. Osakabe, S. Todoriki, H. Ichikawa, B. Hohn, S. Toki, Isolation and characterization of the RAD54 gene from Arabidopsis thaliana. *Plant J.* **48**, 827–842 (2006).

35. S. Inagaki, T. Suzuki, M. A. Ohto, H. Urawa, T. Horiuchi, K. Nakamura, A. Morikami, Arabidopsis TEBICHI, with helicase and DNA polymerase domains, is required for regulated cell division and differentiation in meristems. *Plant Cell* **18**, 879–892 (2006).

36. C. LeBlanc, F. Zhang, J. Mendez, Y. Lozano, K. Chatpar, V. F. Irish, Y. Jacob, Increased efficiency of targeted mutagenesis by CRISPR/Cas9 in plants using heat stress. *Plant J.* **93**, 377–386 (2018).

37. R. Liu, D. C. Chan, The mitochondrial fission receptor Mff selectively recruits oligomerized Drp1. *Mol. Biol. Cell* **26**, 4466–4477 (2015).

38. M. Karimi, D. Inzé, A. Depicker, GATEWAY vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci.* **7**, 193–195 (2002).

39. P. Voigt, G. LeRoy, W. J. Drury 3rd, B. M. Zee, J. Son, D. B. Beck, N. L. Young, B. A. Garcia, D. Reinberg, Asymmetrically modified nucleosomes. *Cell* **151**, 181–193 (2012).

40. Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).

41. G. M. Sheldrick, Experimental phasing with SHELXC/D/E: Combining chain tracing with density modification. *Acta Crystallogr. D* **66**, 479–485 (2010).

42. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).

43. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).

44. V. B. Chen, W. B. Arendall 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).

45. C. M. Lee, M. W. Li, A. Feke, W. Liu, A. M. Saffer, J. M. Gendron, GIGANTEA recruits the UBP12 and UBP13 deubiquitylases to regulate accumulation of the ZTL photoreceptor complex. *Nat. Commun.* **10**, 3750 (2019).

46. F. H. Wu, S. C. Shen, L. Y. Lee, S. H. Lee, M. T. Chan, C. S. Lin, Tape-Arabidopsis Sandwich - a simpler Arabidopsis protoplast isolation method. *Plant Methods* **5**, 16 (2009).

47. J. Nijjer, C. Li, Q. Zhang, H. Lu, S. Zhang, J. Yan, Mechanical forces drive a reorientation cascade leading to biofilm self-patterning. *Nat. Commun.* **12**, 6632 (2021).

48. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

49. K. J. Livak, T. D. Schmittgen, Analysis of relative gene expression data using real-time quantitative PCR and the 2(-ΔΔC(T)) Method. *Methods* **25**, 402–408 (2001).

50. S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).

51. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

52. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

53. F. Ramírez, D. P. Ryan, B. Grüning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dündar, T. Manke, deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).

54. Y. Liao, G. K. Smyth, W. Shi, featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

55. C. J. Hale, M. E. Potok, J. Lopez, T. Do, A. Liu, J. Gallego-Bartolome, S. D. Michaels, S. E. Jacobsen, Identification of Multiple Proteins Coupling Transcriptional Gene Silencing to Genome Stability in Arabidopsis thaliana. *PLOS Genet.* **12**, e1006092 (2016).

56. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2018).

57. F. Hahne, R. Ivanek, Visualizing Genomic Data Using Gviz and Bioconductor. *Methods Mol. Biol.* **1418**, 335–351 (2016).

58. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

59. K. Panda, L. Ji, D. A. Neumann, J. Daron, R. J. Schmitz, R. K. Slotkin, Full-length autonomous transposable elements are preferentially targeted by expression-dependent forms of RNA-directed DNA methylation. *Genome Biol.* **17**, 170 (2016).

60. B. Castel, L. Tomlinson, F. Locci, Y. Yang, J. D. G. Jones, Optimization of T-DNA architecture for Cas9-mediated mutagenesis in Arabidopsis. *PLOS ONE* **14**, e0204778 (2019).

61. O. Raitskin, C. Schudoma, A. West, N. J. Patron, Comparison of efficiency and specificity of CRISPR-associated (Cas) nucleases in plants: An expanded toolkit for precision genome engineering. *PLOS ONE* **14**, e0211598 (2019).

62. E. Weber, C. Engler, R. Gruetzner, S. Werner, S. Marillonnet, A modular cloning system for standardized assembly of multigene constructs. *PLOS ONE* **6**, e16765 (2011).

63. C. Engler, M. Youles, R. Gruetzner, T. M. Ehnert, S. Werner, J. D. Jones, N. J. Patron, S. Marillonnet, A golden gate modular cloning toolbox for plants. *ACS Synth. Biol.* **3**, 839–843 (2014).

64. K. Clement, H. Rees, M. C. Canver, J. M. Gehrke, R. Farouni, J. Y. Hsu, M. A. Cole, D. R. Liu, J. K. Joung, D. E. Bauer, L. Pinello, CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **37**, 224–226 (2019).