



Catching Nucleosome by Its Decorated Tails Determines Its Functional States

Parveen Sehrawat[†], Rahul Shobhawat[†] and Ashutosh Kumar*

Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai, India

OPEN ACCESS

Edited by:

Dileep Vasudevan,
Institute of Life Sciences (ILS), India

Reviewed by:

Sivaraman Padavattan,
National Institute of Mental Health and
Neurosciences (NIMHANS), India

Lars Nordenskiöld,
Nanyang Technological University,
Singapore

Alexey V. Onufriev,
Virginia Tech, United States

*Correspondence:

Ashutosh Kumar
ashutoshk@iitb.ac.in

[†]These authors have contributed
equally to this work and share first
authorship

Specialty section:

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

Received: 24 March 2022

Accepted: 07 June 2022

Published: 14 July 2022

Citation:

Sehrawat P, Shobhawat R and
Kumar A (2022) Catching Nucleosome
by Its Decorated Tails Determines Its
Functional States.
Front. Genet. 13:903923.
doi: 10.3389/fgene.2022.903923

The fundamental packaging unit of chromatin, i.e., nucleosome, consists of ~147 bp of DNA wrapped around a histone octamer composed of the core histones, H2A, H2B, H3, and H4, in two copies each. DNA packaged in nucleosomes must be accessible to various machineries, including replication, transcription, and DNA damage repair, implicating the dynamic nature of chromatin even in its compact state. As the tails protrude out of the nucleosome, they are easily accessible to various chromatin-modifying machineries and undergo post-translational modifications (PTMs), thus playing a critical role in epigenetic regulation. PTMs can regulate chromatin states *via* charge modulation on histones, affecting interaction with various chromatin-associated proteins (CAPs) and DNA. With technological advancement, the list of PTMs is ever-growing along with their writers, readers, and erasers, expanding the complexity of an already intricate epigenetic field. In this review, we discuss how some of the specific PTMs on flexible histone tails affect the nucleosomal structure and regulate the accessibility of chromatin from a mechanistic standpoint and provide structural insights into some newly identified PTM–reader interaction.

Keywords: nucleosome, PTMs, histone tails, acetylation, acylation, methylation

1 INTRODUCTION

In the eukaryote's nucleus, DNA is packaged into the macromolecular “beads on a string”-like structure called chromatin using highly basic histone proteins. A nucleosome is the basic and efficient unit of this organization in which 145–147 bp of DNA are wrapped around a histone octamer (two molecules of each histone H2A, H2B, H3, and H4). Two pairs of H3–H4 dimer form a tetramer stabilized by a characteristic hydrophobic four-helix bundle structure between H3 and H3', and then two dimers of H2A–H2B interact with H3–H4 tetramer on each side through a second homologous hydrophobic four-helix bundle structure between H2B and H4, forming a globular octamer from which disordered tails protrude out. Through extensive hydrogen-bonding and electrostatic interactions, histones coordinate with DNA *via* conserved histone fold domains, resulting in the bending of negatively charged DNA over a positively charged octamer surface. This bent conformation of DNA brings the phosphate backbone of the two strands closer, and this energetically constrained conformation is maintained by neutralizing negative charges by positively charged lysine and arginine side chains (Luger et al., 1997).

These strong and extensive interactions render the nucleosome a stable disc that can sterically inhibit the binding of chromatin-associated proteins (CAPs). Virtually all eukaryotic organisms use the inhibitory nature of this packaging to regulate access to DNA. However, the information encoded inside the DNA must be retrieved at appropriate times. Although DNA is very tightly compacted, it still remains accessible to many enzyme machineries that replicate it, repair it, and use it to produce RNA

molecules and proteins. For doing this, chromatin and nucleosomes must be inherently dynamic and highly malleable. Numerous biochemical and structural studies established the dynamic nature of the nucleosome in terms of its conformation and composition. The disordered N-terminal tail of histones have an affinity to DNA, forming a dynamic complex with DNA termed as “fuzzy conformational ensembles,” which regulate the chromatin structure and dynamics (Ghoneim, Fuchs, and Musselman 2021; Peng et al., 2021; Shukla, Agarwal, and Kumar 2022). Polach and Widom, (1995) demonstrated the phenomenon of intrinsic structural dynamics of nucleosome known as “DNA breathing,” i.e., partially unwrapping and rewinding of DNA spontaneously. This dynamic unwrapping/rewinding phenomenon is exploited by several DNA-binding proteins like transcription factors in a tunable and analogous fashion. Using FRET experiments, the Langowski group showed that disassembly of nucleosome is initiated by DNA breathing resulting in a dynamic “octasome,” which opens on a 50 μ s time scale at an angle of $\approx 20^\circ$. This results in disruption of dimer tetramer interface with H2A–H2B dimer evicting first followed by H3–H4 tetramer removal (Böhm et al., 2011; Gansen et al., 2018).

In addition to DNA breathing, cells have also evolved various other mechanisms to make nucleosomal DNA more accessible: histones posttranslational modifications, histone chaperones, histone variants, and chromatin remodelers. These regulatory mechanisms control the genome function without changing the nucleotide sequence, also referred to as “epigenetic” marks. Histone post-translational modification is the process of covalently attaching adducts like methyl group, acetyl group, phosphate group, and ubiquitin group. These modifications present on free N-terminal tails or inside the histone fold domain affect the structure and dynamics of nucleosomes locally and chromatin globally and provide the binding platform for different groups of proteins like transcription factors, chromatin remodelers, histone-/DNA-modifying enzymes, and chaperones, especially the charge-altering PTMs inside the globular histone octamer core can modify the electrostatic interaction of histone–histone or histone–DNA, thereby altering the structure and dynamics of nucleosomes (Fenley et al., 2018). For instance, phosphorylation in combination with acetylation inside the nucleosomal DNA entry–exit site modulates DNA accessibility by transcription complexes (Brehove et al., 2015). Misregulation of these modifications can cause many diseases like cancer; therefore, regulating this epigenetic mark is necessary for proper functioning. In this review, we have discussed the role of four PTMs (acetylation, acylation, seronylation, and methylation) present on flexible and intrinsically disordered histone tails in regulating chromatin accessibility and function. Several excellent reviews on other modifications like phosphorylation (Sawicka and Seiser 2014; Treviño, Wang, and Walker 2015), SUMOylation (Ryu and Hochstrasser 2021), and ubiquitination (Mattiroli and Penengo 2021) are good read to get a better understanding.

2 HISTONE TAILS AND PTMS

The nucleosome is a globular structure, but the unstructured N-terminal tail of each histone protrudes out from its core. The

pioneering work of Vincent Allfrey in the 1960s and subsequent studies revealed that these tails are subjected to many post-translational modifications like acetylation, methylation, and phosphorylation, thus acting as a hub of chromatin signaling (Millán-Zambrano et al., 2022). Covalent modifications of histone tails can alter the chromatin structure *via* cis-effects or trans-effects. Cis-effects are employed by changing the biophysical properties of modified histone chains, like altering the electrostatic charge or structure of the tail, which in turn affects internucleosomal contacts. For example, histone acetylation on lysine residue exerts its effect by neutralizing the positive charge of histone tails. Charge-neutralized tails generate a localized decondensation of the chromatin fiber, resulting in better availability of DNA double helix to the transcription machinery. Acetylation at H4K16 inhibits the packaging of a nucleosomal array in a compact 30-nm chromatin fiber *in vitro* and further abolishes cross-fiber interactions (Michael et al., 2006). In fact, out of four acetylations possible in the H4 tail, K16 acetylation is unique as only this modification reduces the cation-induced folding of the 12-mer nucleosome array implicating cis-effect of acetyl mark (Allahverdi et al., 2011). Multiscale computational studies supported by NMR experiments revealed that acetylated H4 tails lose local contacts and reduced tail availability for forming critical internucleosomal interactions resulting in the unfolding of chromatin fiber (Collepardo-Guevara et al., 2015; Bascom and Schlick 2018). Similarly, phosphorylation adds a net negative charge generating “charge patches,” which result in alteration of nucleosome packaging (Dou and Gorovsky 2000). Bulky groups, such as ADP-ribose and ubiquitin, also affect the arrangements of the histone tails and open up nucleosome arrays.

Histone modifications also act *via* trans-effects, where modification-binding partners are recruited to the chromatin. This is similar to “reading” a specific covalent histone mark by modification reader proteins. For example, the acetylation mark is read by proteins having “bromodomains” (Jacobson et al., 2000). Similarly, methylated lysine or arginine residues are read by chromodomains or similar domains (e.g., MBT and Tudor) to facilitate the modulation of chromatin (Maurer-Stroh et al., 2003). Acetylation, methylation, and phosphorylation were the initially detected and extensively studied modifications. Recent advancement in the high-sensitive mass spectrometry technique has played a pivotal role in revealing a wide array of new modifications, including ubiquitylation, SUMOylation, ADP-ribosylation, a dozen of various acyl groups, seronylation, and lactylation (Zhao and Garcia 2015). Based on diversity and biological specificity of distinct modifications, Strahl and Allis, (2000) proposed the “histone code” hypothesis, which states that “multiple post-translational modifications form a specific pattern either in combination or sequential fashion on same or different histone tail, to perform a specific downstream function.” The key players involved in defining the histone code are the enzymes or proteins that write, read, and then erase these marks in a specific sequence or modification. This fine-tuned action is critical for regulating most nuclear processes, including replication, recombination, DNA damage and repair, transcription, and differentiation.

The crosstalk and specific recognition of histone PTM by its cognate reader define the temporal and spatial modulation of the genome. After the initial discovery of the bromodomain as an acetylation mark reader and chromodomain as a methylation mark reader, several epigenetic studies have identified a diverse repertoire of “readers” regulating the dynamic nature of the chromatin landscape. For example, RAG2 protein of RAG1/2 V(D)J recombinase reads H3K4me3 modification and induces V(D)J recombination at the T- and B-cell receptor gene locus. Mutations in the reader motif of RAG2 impair V(D)J recombination and can result in immunodeficiency syndromes (Matthews et al., 2007). Many chromatin-associating multi-subunit enzymatic complexes contain a set of multiple readers within one or different subunits, and these readers having specificities for different marks can be in close proximities. These complexes can be “writers” or “erasers” [histone acetyltransferases and histone deacetylases] that can redefine the epigenetic landscape by adding or removing modifications at different sites or chromatin remodelers that can alter the structure and dynamics of chromatin. Combinatorial readout of multivalent histone PTMs on the same tail or at a different tail can provide a lock and key type mechanism to carry out a specific biological function at targeted genomic loci. Owing to their fundamental role, any misreading of these epigenetic modifications has been shown to contribute to many human diseases, including cancer and developmental and autoimmune disorders (Chi, Allis, and Wang 2010; Shen and Laird 2013).

In some cases of acute myeloid leukemia, the reader module of H3K9 trimethylation (PHD motif) is found to be fused with nuclear pore protein (NUP98). This fusion protein remains bound to H3K9me3, interfering with the removal of this modification and the addition of H3K27me3, thereby affecting the normal differentiation of progenitor and hematopoietic cells (Wang et al., 2009). In fact, misinterpretation of acetyl marks by their respective reader domains has been implicated in uterine, bladder, cervical, and other tumors (Zhao et al. 2021). Targeting one of the bromodomain protein families BET (bromodomain and extraterminal domain) by small molecules resulted in the reversal of a cancer cell phenotype in the patient-derived NUT midline carcinoma cell line (Filippakopoulos et al., 2010). Understanding the basic aspects of epigenetic control and the genesis of epimutation-induced human disorders requires an understanding of the molecular mechanism and functional importance of PTM–reader interactions. In the following sections, we discuss the molecular mechanism of PTM readout by different reader modules and the functional significance of these newly identified PTM–reader interactions.

3 ACETYLATION

Acetylation of the lysine residue at the ϵ -amino group was the first PTM discovered in thymus histones by Philips in 1961 (Allfrey et al., 1964). The negative charge on the acetyl group neutralizes the positive charge of the lysine side chain, thereby altering the electrostatic properties of histone proteins. This modification is generally correlated with a transcriptionally active state, and the

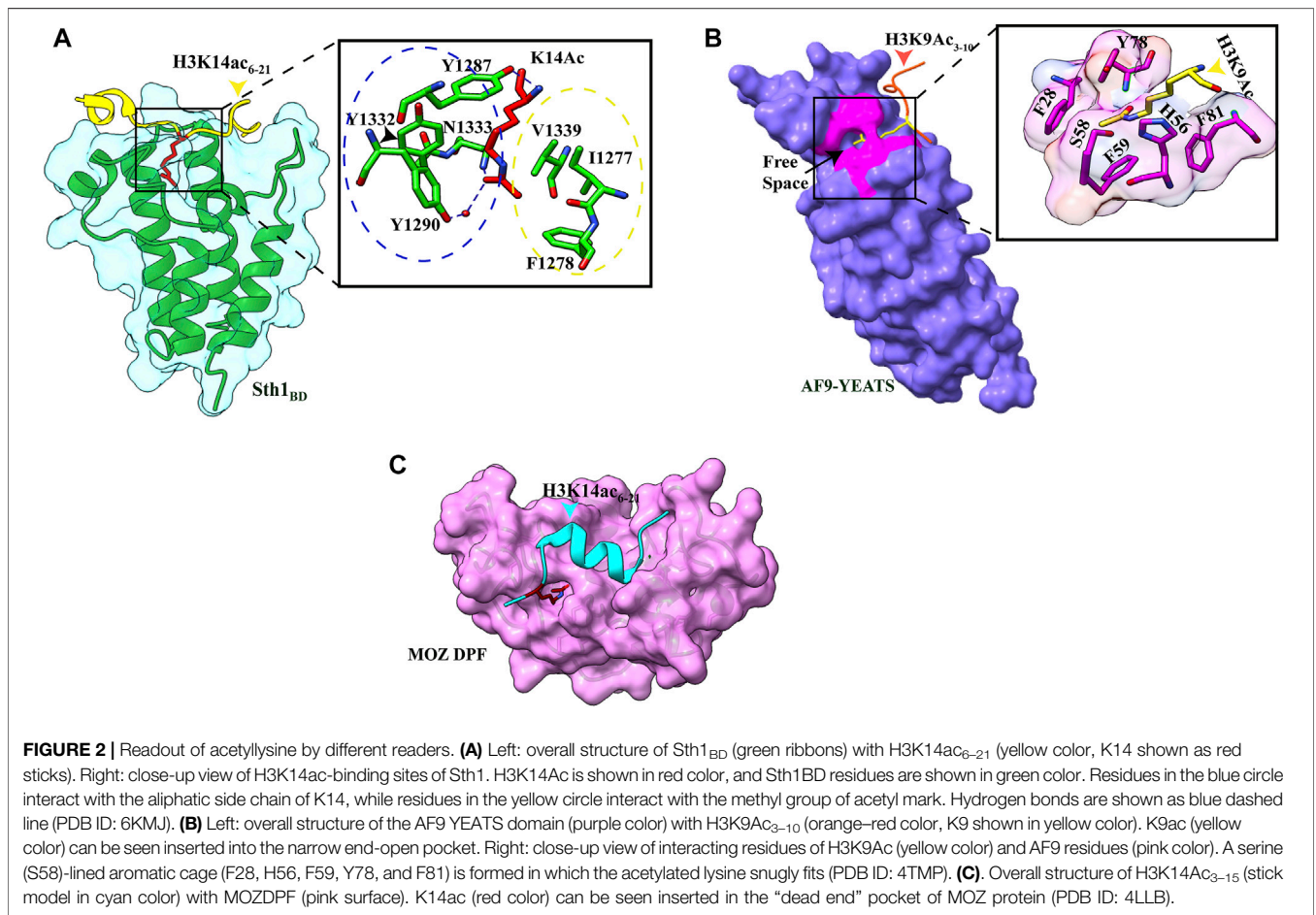
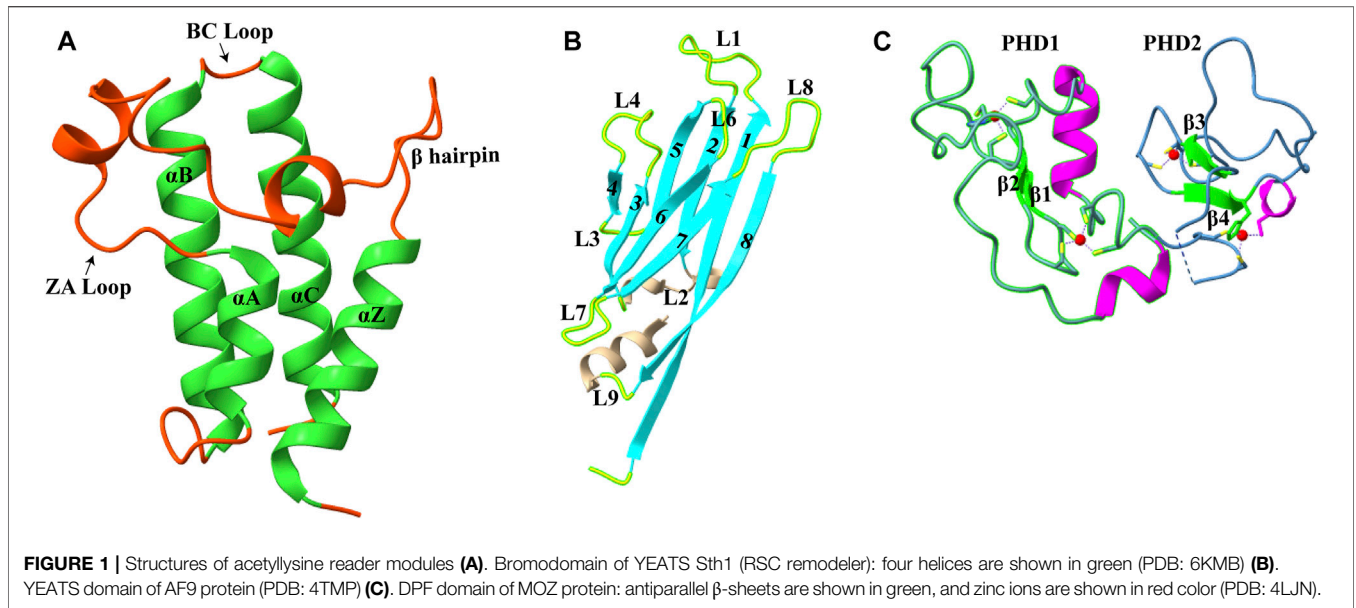
turnover of this modification is controlled by two groups of enzymes: histone acetyltransferases (HATs) and histone deacetylases (HDACs) (discussed in detail in Marmorstein and Zhou, (2014) and Xia et al. (2020)). This mark is found on all histone tails H2A (K5 and K9), H2B (K5, K12, K15, K16, K20, and K120), H3 (K4, K9, K14, K18, K23, K27, K36, and K56), and H4 (K5, K8, K12, K16, K20, and K91) (Musselman et al., 2012). Although acetylation was discovered about 60 years ago, the first reader of acetylated lysine, a bromodomain, was discovered only in 1999 (Dhalluin et al., 1999). Till now, three types of protein domains able to “read” acetyllysine marks have been identified: bromodomains, DPF domain, and YEATS domain (Figure 1). Here, we will discuss these newly identified reader proteins of acetylation mark.

3.1 Recognition of H3K14Ac by the Bromodomain Module of RSC Chromatin Remodeler

Bromodomains are evolutionarily conserved domains that act as histone lysine acetylation readers. In humans, 61 bromodomains in 46 different proteins have been identified, and these proteins are part of transcription-regulating complexes, chromatin remodelers, and PTM writers. Based on the structure and sequence, bromodomains are divided into eight subfamilies (I–VIII) (Filippakopoulos et al., 2012). Even with little sequence homology, all bromodomains have a conserved structural fold consisting of four α -helices (α Z, α A, α B, and α C) (Figure 1A). The two highly variable loops, ZA and BC, joining these helices, form a deep hydrophobic acetyllysine binding pocket (Sanchez and Zhou 2009).

One of the yeast chromatin remodelers, the RSC complex, consists of seven bromodomains. Acetylation of histone H3 lysine at the 14th position enhanced RSC binding to nucleosomes and augmented the RSC remodeling activity (Duan and Smerdon 2014; Lorch, Maier-Davis, and Kornberg 2018). Recently, Chen et al. (2020) showed that out of seven bromodomains present in RSC (one in the Sth1 subunit and two each in Rsc1, Rsc2, and Rsc4 subunits), the C-terminal bromodomain of Sth1 is the primary domain responsible for recognizing H3K14Ac. ITC experiments using H3K14Ac containing H3_{6–21} peptide revealed that Rsc1 and Rsc2 have no significant interaction, while Rsc4 (dissociation constant of 263 μ M) has a 16-fold weaker interaction than Sth1 bromodomain (dissociation constant of 16 μ M). Further ITC results with an array of histone peptides containing different acetylation sites demonstrated that the Sth1 bromodomain could also strongly bind to H3K20Ac with similar K_D as that of H3K14Ac. Sequence analysis of these peptides revealed a conserved feature in both H3K14Ac and H4K20Ac peptides: the following two residues after lysine are hydrophobic, and the third one is a conserved arginine (K(Ac) $\Phi\Phi$ R motif, where Φ represents any hydrophobic amino acid). Mutation at the +1 and +2 position with neutral or polar amino acid and mutation of +3 arginine abolish the interaction between the peptide and Sth1 bromodomain.

Like other bromodomain-containing proteins, Sth1 has a hydrophobic pocket formed by four amphipathic α -helices in



which K14Ac is inserted. In addition to hydrophobic contacts, interactions between the aliphatic side chain of K14 with three aromatic amino acids (Y1287, Y1332, and Y1290), the methyl

group of acetyl mark with V1339, I1277, and F1278, and the hydrogen bond between the carbonyl oxygen of the acetyl group and N1333 of Sth1 are responsible for strong affinity (**Figure 2A**).

The other three residues of the K(Ac)ΦΦR motif also make extensive hydrogen bonds and hydrophobic contacts with the Sth1 bromodomain, explaining the specific recognition of this motif by Sth1_{BD}. Critical residues of hydrophobic pockets (Y1332 and N1333) are not present in Rsc1 and Rsc2 subunit bromodomains which explains their insignificant interaction with the H3K14Ac peptide (Chen et al., 2020).

More recently, the *C. elegans* homolog of Sth1, SMARCA4, was also shown to be highly selective for H3K14Ac and showed a similar binding affinity. The hydrogen-bonding and hydrophobic interactions are very well conserved in *C. elegans* SMARCA4 bromodomain and H3_{7–20}K14Ac-modified peptide complex. In this case also, the K(Ac)ΦΦR motif is involved in the extensive electrostatic, hydrophobic, and hydrogen-bonding interactions that ensure specific and robust binding between SMARCA4 and H3K14Ac-containing peptides (Enriquez et al., 2021).

The RSC complex is the most abundant and well-characterized chromatin remodeler of the SWI/SNF family, comprising about 17 subunits in *Saccharomyces cerevisiae*. Like other remodelers, using its main catalytic subunit Sth1, the RSC complex catalyzes the ATP hydrolysis reaction and uses this energy to evict or slide the histone octamer to expose DNA-binding sites on chromatin. At the H3K14Ac-enriched transcription start sites (TSSs), the RSC complex is recruited, which generates a nucleosome-free region enabling RNA Pol II to initiate transcription (Carey, Li, and Workman 2006; Lorch et al., 2011). Also, H3K14Ac is found at UV-irradiated DNA sites, which recruit the RSC complex to facilitate DNA repair by chromatin remodeling (Yu et al., 2005; Duan and Smerdon 2014).

3.2 Recognition of H3K9ac by the YEATS Domain

A study published in 2014 showed YEATS domain as a novel reader of histone acetylation marks. It is an evolutionarily conserved protein module from yeast to humans and is named after its founding domain-containing proteins, Yaf9, ENL, AF9, Taf14, and Sas5 (Masson et al., 2003). Three YEATS domain-containing proteins in *S. cerevisiae* and four proteins in humans are associated with transcription-regulating complexes, chromatin-remodeling complexes, and HAT complexes (Schulze et al., 2009). ITC and pull-down assay revealed that binding of the YEATS domain of AF9 protein to histone H3 tail is acetylation dependent, and the AF9 YEATS domain binds strongly to H3K9Ac (K_D of 3.7 μ M) as well as to H3K27Ac (K_D of 7.0 μ M) and H3K18Ac (K_D of 11.0 μ M), however, to a lesser extent (Li et al., 2014). The crystal structure of YEATS domain with different acetylated histone peptides uncovered a unique serine-lined aromatic sandwich pocket for specific acetyllysine readout. The AF9 YEATS domain adopts an immunoglobulin fold in which eight antiparallel β strands form a two-layer β sandwich, and H3K9Ac long side chain is inserted into a serine-lined aromatic cage formed in the cleft of loops L4 and L6 (Figures 1B, 2B). In the AF9 YEATS–H3K9Ac complex, the YEATS domain uses strands β 2 and β 7 and loops L1, L4, L6, and L8 to form extensive hydrogen bonds and hydrophobic interactions with the T3–S10 segment of H3. In addition to hydrogen and hydrophobic interactions, multiple aromatic residues in the acetyllysine binding

pocket are involved in multiple sets of CH– π interactions, which collectively contribute to the stable binding (Figure 2B). Key residues involved in the generation of the aromatic cage are highly conserved among different YEATS domain-containing proteins from yeast to humans. The interaction of the YEATS domain and H3 tail is also highly dependent on amino acids flanking the K9, especially arginine, at the eighth position, as mutation at this site resulted in a 200-fold binding decline (Li et al., 2014).

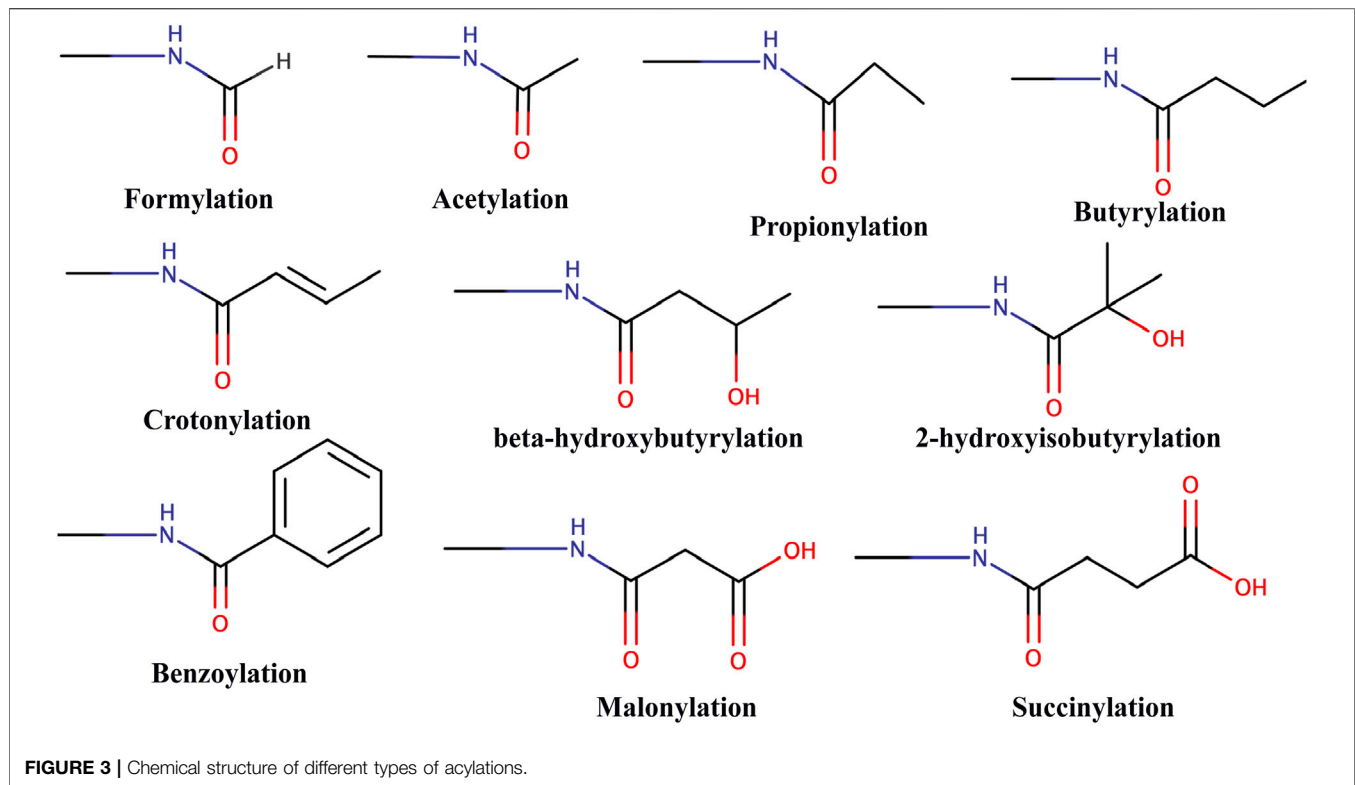
AF9 is subunit of a large protein complex, Super Elongation Complex (SEC), which has been shown to mediate enhanced transcription of several loci in MLL-rearranged leukemias and developmental genes by releasing paused Pol II (Smith, Lin, and Shilatifard 2011). ChIP-seq and CoIP experiments suggested that the YEATS domain (N-terminal part of AF9) is critical for the recruitment of AF9 at the H3K9ac mark around the transcription start sites and C-terminal of AF9 is required for the interaction with other proteins of SEC complex *in vivo*. One of the critical interacting partners of AF9 is DOT1L, an H3K79 methyltransferase, and H3K79me3 mark is associated with active transcription. Several *in vivo* experiments revealed that AF9 is required for DOT1L recruitment at targeted genes and subsequent deposition of H3K79me3 to promote active transcription (Li et al., 2014).

3.3 Recognition of H3K14ac by the DPF Domain

The double PHD finger (DPF) domain, a subgroup of PHD (plant homeodomain) fingers, is a tandem of PHD fingers with a face-to-back orientation where two domains form a single structure. This domain has been found in two protein families, histone acetyltransferase MYST family proteins (MOZ or KAT6A and MORF or KAT6B) and subfamilies of SWI/SNF chromatin remodeler (BAF and PBAF complex). The DPF domain from all these proteins is homologous, and all the key residues are conserved, forming a highly similar secondary structure consisting of two antiparallel β -sheets followed by a C-terminal α -helix which is coordinated by two zinc atoms *via* Cys4–His–Cys3 motif in a cross-brace topology (Figure 1C). Two PHD fingers are linked with one another in a face-to-back orientation mediated by the interaction between glutamic acid and arginine in the α -helix of the first PHD finger. Although the PHD finger was originally recognized as a methylation mark reader, the DPF domain of DPF3b was shown to bind H3K14 acetylation mark. The structural aspects of acetyllysine–H3K14Ac interaction are discussed in the next section.

4 ACYLATIONS

The latest advancements in mass spectrometry revealed a wide array of acylation marks in histones apart from classical acetylation modification. These acyl marks include butyrylation (Kbu) (Chen et al., 2007), propionylation (Kpr) (Chen et al., 2007), crotonylation (Kcr) (Tan et al., 2011), succinylation (Ksucc) (Xie et al., 2012), malonylation (Kma) (Xie et al., 2012), 2-hydroxyisobutyrylation (Khib) (Dai et al., 2014), β -hydroxybutyrylation (Kbhb) (Xie et al., 2016),



benzoylation (Kbz) (Huang et al., 2018), lactylation (Kla) (Zhang et al., 2019), glutarylation (Kglu) (Bao et al., 2019), and isobutyrylation (Kibu) (Zhu et al., 2021) (Figure 3). These modifications are derived from their respective acyl-CoAs, a product of different metabolic pathways. Therefore, these specific marks can identify the metabolic state of the cell and regulate chromatin dynamics and gene expression according to the need of the cell (Nitsch, Shahidian, and Schneider 2021). Also, various new studies suggest that these different acylation marks are important for eliciting specific epigenetic responses (Dutta, Abmayr, and Workman 2016).

Till date, no specific or selective writer, reader, or eraser for non-acetyl acylation modification has been identified so far. A recent study with high-throughput profiling of an acyl-CoA/protein using CoA/AcetylTransferase Interaction Profiling (CANTIP) revealed only known acetyl mark-interacting proteins (Levy et al., 2020). In fact, p300 lysine acetyltransferase (also known as KAT3B) has been shown to be able to catalyze the transfer of all types of different acyl groups (Nitsch, Shahidian, and Schneider 2021).

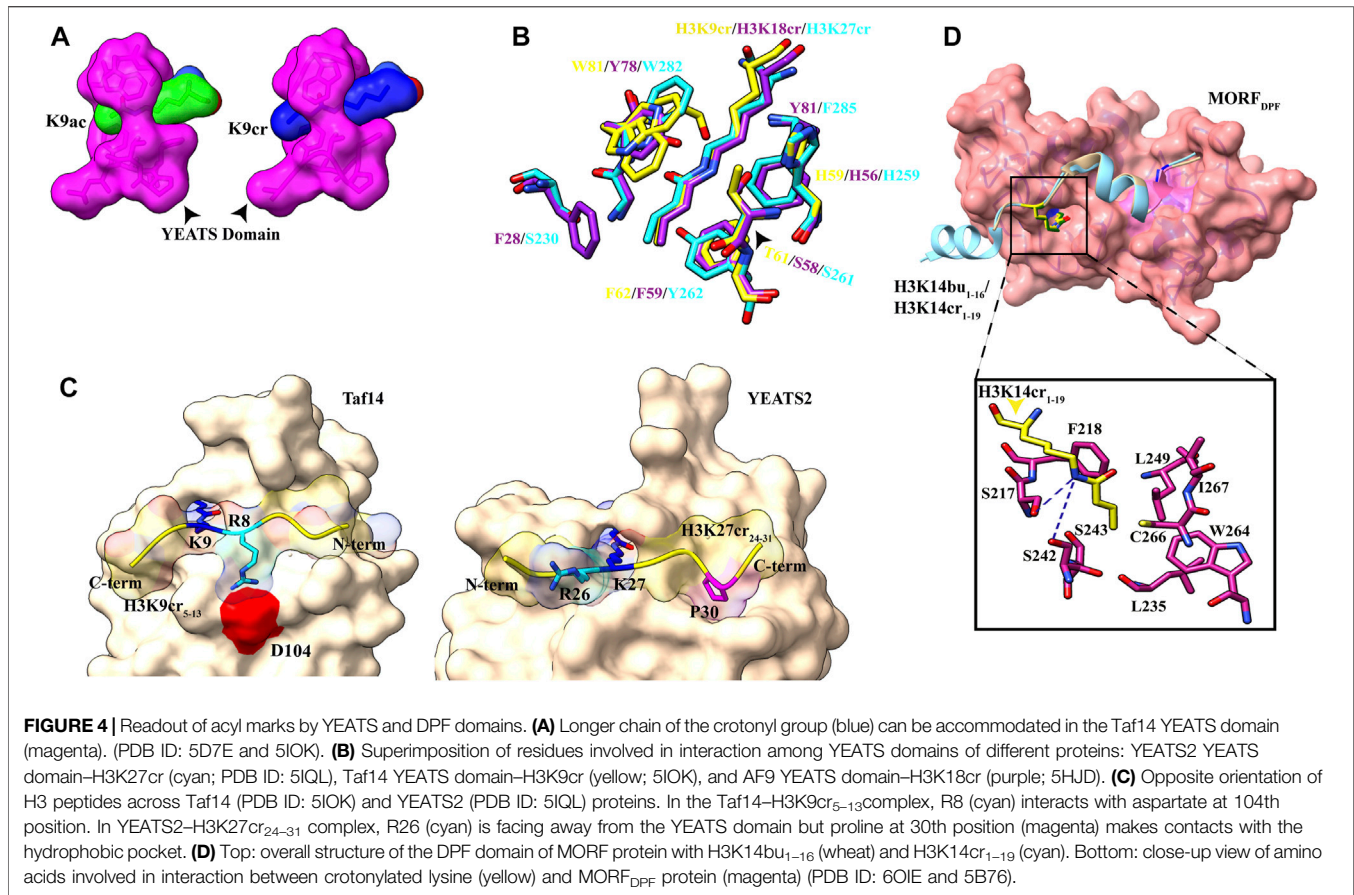
4.1 Recognition of Acyl Marks by Bromodomains

Given that bromodomain is a major protein module that reads lysine acetylation, an initial study found out that the bromodomain of bromodomain-containing protein (BRD4) was able to bind Kbu and Kpr but with very less affinity than Kac (Vollmuth and Geyer 2010). A more comprehensive study

where 49 bromodomains were assayed for their binding affinity to different acyl-modified H3 peptides revealed that only bromodomains having larger binding pockets such as CECR2 and BRD9 were able to bind long-chain Kbu modification, and the second bromodomain of TAF1 was able to interact with Kcr, albeit with reduced affinity compared with Kac (Flynn et al., 2015). All these studies implied that bromodomains could read a few acyl marks, but these interactions are not strong and significant compared to acetyl modification.

4.2 Recognition of H3K9acyl/H3K18acyl/H3K27cr by the YEATS Domain

In the crystal structure of the AF9 YEATS domain-H3K9Ac complex, a clear open space at the end of the aromatic sandwich cage led to the hypothesis that this open space can accommodate a large chain of bulkier acyl marks (Li et al., 2014) (Figure 1B). Further calorimetric titrations and NMR 2D ^{15}N - ^1H heteronuclear single quantum coherence (HSQC) spectra revealed that, indeed, the AF9 YEATS domain could bind to the H3 tail peptides, which has crotonylation (cr), propionylation (pr), butyrylation (bu), and formylation (fo) modifications at K9, K18, and K27 positions with no significant binding for the H3K14 site (Li et al., 2016; Zhang et al., 2016). An increase in the hydrocarbon chain beyond the acetyl group resulted in 2.4-, 1.9-, and 1.4-fold binding enhancement for Kcr-, Kpr-, and Kbu-modified peptides, respectively. Similar studies with Taf14 and YEATS2 (subunit of ATAC histone acetyltransferase complex)



established that the YEATS domain is a Kac reader module with the highest affinity to crotonylation mark (Andrews et al., 2016; D.; Zhao et al., 2016). The crystal structure of all three proteins in complex with H3 tail peptide having crotonylation mark showed a highly similar structure where the extended side chain of Kcr fits comfortably into the narrow end-open pocket of the YEATS domain (Figure 4A).

Two conserved aromatic residues in different YEATS domains (F62 and W81 for Taf14, F59 and Y78 for AF9, and Y268 and W282 for YEATS2) make a sandwich arrangement with the planar crotonylamide group, which crosses the β -sandwich cage at a 90° angle in a corkscrew-like manner (Figure 4B). This arrangement favors a novel “aromatic– π –aromatic” stacking (also called “ π – π – π ” stacking) (Klein et al., 2018). Additionally, extensive hydrophobic interactions, the amide– π interactions, the CH– π interactions, and electrostatic interactions (mainly of Ca and C β of the alkene moiety with the carbonyl oxygen of Q79) between the side chain of crotonylated lysine and pocket residues significantly contribute to specific recognition of Kcr by the YEATS domain (Krone et al., 2020).

Flanking residues in all three H3K9, H3K18, and H3K27 are conserved, sharing a common motif “A_(–2)R_(–1)KS₍₊₁₎.” An acidic aspartate residue of Taf14 and AF9 YEATS domains forms charge-stabilized hydrogen-bonding interaction with ($n-1$) arginine of H3K9cr peptide. Interestingly, in the YEATS domain of YEATS2 protein, this acidic aspartate residue is replaced by the

neutral asparagine residue, which does not recognize “R_(–1).” While Taf14 and AF9 YEATS domain prefer H3K9cr and H3 N-terminal residues “K4–Q5–T6–A7–R8” have extensive interactions with loops L6 and L8 surface residues, YEATS2 binds H3K27cr more strongly and oppositely oriented C-terminal residues “S28–A29–P30–A31” fits nicely on the surface of L6 and L8 loops (Figure 4C). In the crystal structure of YEATS2, the YEATS domain in complex with H3K27cr revealed a hydrophobic pocket in YEATS2 in which H3P30 fits snugly and facilitates the correct positioning of H3K27cr, explaining the site specificity of YEATS2 (D. Zhao et al., 2016).

A recent study showed the role of histone crotonylation and Taf14 in the yeast metabolic cycle. In the yeast metabolic cycle, acetylation increases in high oxygen consumption state, followed by generation of crotonylation intermediates (Gowans et al., 2019). As the cells shift to a low oxygen consumption state, acetylation mark is replaced by crotonylation marks and in this LOC state H3K9cr and Taf14 repress the pro-growth genes, contrary to earlier studies showing their role in gene expression (Sabari et al., 2015).

4.3 Recognition of H3K14acyl by the DPF Domain

Pull-down assays and ITC experiments using an array of H3K14 bearing different acylations revealed that the DPF domain of DPF2 and MOZ HAT displays more affinity for

H3K14cr, H3K14bu, and H3K14pr than H3K14Ac with crotonylation being the most favored (Xiong et al., 2016). On a similar line, a combination of fluorescence spectroscopy, NMR, and histone peptide pull-down assay established the specificity of the DPF domain of MORF HAT for H3K14cr and H3K14bu (Klein et al., 2017). In all the crystal structures solved for the DPF domain in complex with H3K14cr or H3K14bu, H3 tail peptide has extensive contacts with double PHD finger of the DPF module with segments H3_{4–11} and H3_{17–25} (in case of full length taken) adopting α -helical conformation (Xiong et al., 2016; Klein et al., 2017, 2019). The overall structural analysis revealed binding of the DPF module to acyllysine in a ping-pong-like manner with three characteristic interactions (Klein et al., 2019).

The first PHD domain of the DPF module forms a unique zinc-finger domain in which a hydrophobic pocket is formed at the β -sheet-2 surface. In the MORF (and MOZ) proteins, the hydrophobic pocket is formed by the amino acid residues I228–C230 of β -1, N235–G237 of β -2, and amino acid residues involved in zinc ion coordination S210 (S217), F211 (F218), L242 (L249), W257 (W264), C259 (C266), I260 (I267), and E261 (Figure 4D). The planar crotonylation group of H3K14 is inserted snugly into this hydrophobic reader pocket and stably positioned with the help of four water-mediated hydrogen-bonding interactions and four pairs of hydrophobic contacts. The structural and sequence alignment analysis of DPF domains showed that glycine residue G237 of β -strand-2 is a critical component of the hydrophobic pocket due to its free side chain. In classical PHD fingers, this glycine residue is replaced by bulky amino acids like phenylalanine or tyrosine, which fill the pocket and block large chain acylation mark insertion. One of the phenylalanines (F211 in MOZ and F218 in MORF) is in close proximity to the inserted crotonyl group and forms a π - π interaction between the C=C double bond of the crotonyl group and the aromatic ring of phenylalanine. This additional interaction is responsible for selectivity and strengthens the interaction between DPF domain and H3K14cr (Klein et al., 2019). Additionally, H3 residues R2 and K4 are inserted into two “acidic” pockets formed at the surface of β -1 sheet of the second PHD domain and held by hydrogen bonding and electrostatic interactions.

5 SEROTONYLATION

Serotonylation is the attachment of the serotonin molecule to the glutamine residue of proteins. Serotonin [or 5-HT (hydroxy tryptamine)] is a monoamine with an abundant presence and diverse functions varying from neurotransmitter to hormone release and gastrointestinal motility. Additionally, serotonin has been shown to have the ability to covalently modify several proteins, including RacI, small guanosine triphosphatase, and fibronectin and thereby regulate their functions (Walther et al., 2003; Watts, Priestley, and Thompson 2009). The tissue transglutaminase 2 (TGM2) enzyme is responsible for conjugating serotonin to cytosolic

proteins *via* the transamidation reaction (Hummerich et al., 2012).

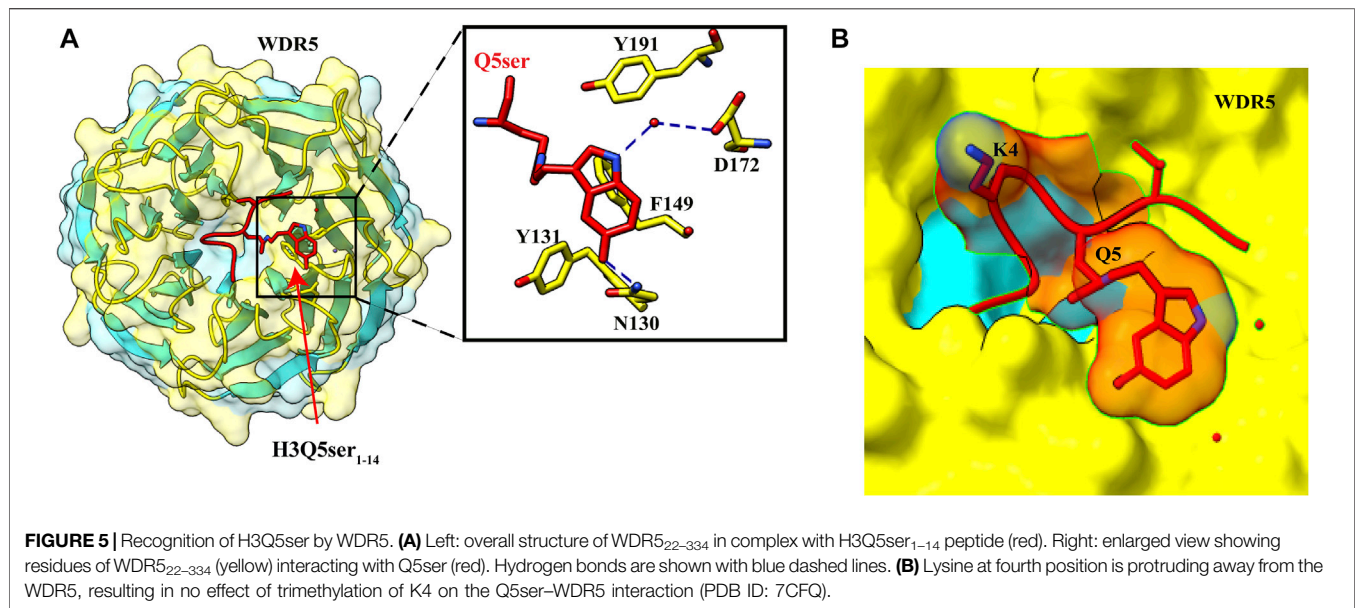
It was previously known that TGMs could modify histones *in vitro* and do so very fast compared to some of the known native substrates (Abad, and Franco 1996). But recently, using the bio-orthogonal metabolic-labeling approach, Farrelly et al. (2019) showed that TGM2 can catalyze serotonylation of glutamine at the fifth position of histone H3 trimethylated lysine 4 (H3K4me3)-marked nucleosomes, resulting in the presence of combinatorial H3K4me3Q5ser *in vivo*.

5.1 Recognition of H3Q5ser by WDR5

In a pull-down assay, WDR5 was captured using H3 peptide with H3K4me3Q5ser dual marks as the bait, suggesting that WDR5 could be a potential reader of this modification (Farrelly et al., 2019). WDR5 is a core subunit of a histone methyltransferase enzyme; the MLL (mixed-lineage leukemia) complex is responsible for trimethylation of H3K4. Further pull-down assays and ITC experiment established that serotonylated H3Q5 enhances the binding of WDR5 by at least two-fold than that of unmodified H3, and H3K4 trimethylation mark has no significant effect on this binding, also supported by the observation that there was no electron density for the trimethyl group of K4 in the crystal structure of WDR5–H3K4me3Q5ser complex (Jie et al., 2022) (Figure 5B).

The crystal structure of WDR5–H3Q5ser complex revealed that the serotonyl group is placed in a shallow hydrophobic surface pocket of WDR5. The WDR5–Q5ser interaction is stabilized *via* a network of hydrogen bonds (one between OH group of serotonin with amide group N130 residue of WDR5 and another between amide group of serotonin with WDR5 D172 side chain) and van der Waals contacts (between hydrophobic moiety of serotonyl group and aromatic side chains of Y131, F149, and Y191 of WDR5) (Figure 5A). Additionally, R2 of H3 peptide also participates in this complex formation as it is anchored into a negatively charged central channel and interacts with WDR5's F133 and F263 through cation- π interactions (Jie et al., 2022).

In neuroblastoma cells, upon recognition of H3Q5ser modification, WDR5 is recruited to promoter regions of oncogenic genes *GPX1*, *C-MYC*, and *PDCD6* that can promote tumor formation (Jie et al., 2022). Knockdown studies implied that serotonylation of H3Q5 is not dependent on H3K4me3; instead, there was a decrease in the level of H3K4me3 upon WDR5 knockout, also seen in the case of TGM2 knockdown, which may be due to less recruitment of MLL1 complex. H3K4me3Q5ser displays a ubiquitous pattern of tissue expression in mammals, with enrichment observed in the brain and gut, two organ systems responsible for the bulk of 5-HT production. Genome-wide analyses of human serotonergic neurons, developing mouse brain, and cultured serotonergic cells indicate that H3K4me3Q5ser nucleosomes are enriched in euchromatin, are sensitive to cellular differentiation, and correlate with permissive gene expression—phenomena that are linked to the enhanced function of TFIID (Farrelly et al., 2019).



6 METHYLATION

Histone methylation and its importance in transcription were first observed in the 1960s. There are three lysine methylation states: -mono-, -di-, and -tri (me1, me2, and me3); since methylation does not change histone's charge configuration, the primary function of these methylations is to interact with effector molecules that specifically recognize these modifications. Generally, all other histone modifications are specific for the active or repressed state, while in methylated chromatin it depends on its methylation state and the modification position. For example, H3K4, H3K36, and H3K79 methylations are considered to mark active transcription (Heintzman et al., 2007), whereas H3K9, H3K27, and H4K20 methylations are associated with silenced chromatin states (Bernstein et al., 2005; Barski et al., 2007).

Methyllysine-specific readers have a peculiar characteristic: they recognize histone modification by an aromatic cage that comprises two to four aromatic amino acids. These aromatic amino acids in some complexes are perpendicular to each other, which helps encircle the entire lysine methylation. The compartment of the aromatic cage defines whether mono-, di-, or tri-methylation state interacts with it. Therefore a small compartment limits its interaction with a higher methylation state because of steric hindrance, while a large compartment favors the interaction with a higher methylation state. Interaction in the compartment between the methylammonium group and the aromatic cage is stabilized by cation- π interactions and the hydrophobic and van der Waal interactions. Amino acids surrounding methyllysine play a vital role in the reader's specificity for a particular methylated lysine. Some readers show very low specificity, while others are specific for a specific methylated state. Beyond caging of the methyllysine, the mechanism of recognition of surrounding residues varies among readers. A number of evolutionarily conserved domains

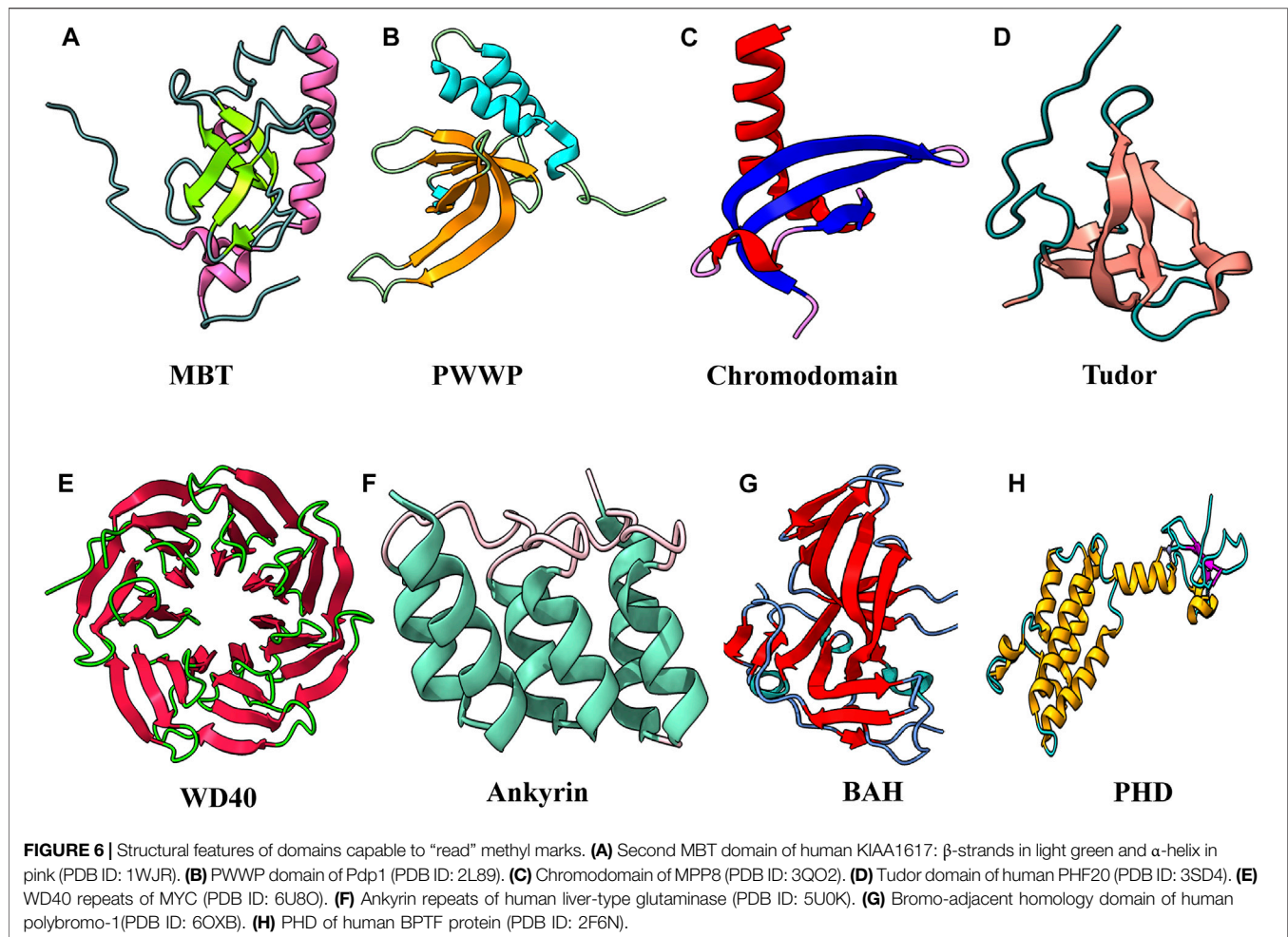
were discovered that interact specifically with the methylated histone. These “reader” proteins contain methyl-lysine-binding motifs, including PHD, chromo, Tudor, PWWP, WD40, BAH, ADD, ankyrin repeat, and MBT domains (Figure 6). These readers can distinguish target methyllysine based on their methylation state and surrounding amino acid sequence (Musselman et al., 2012).

6.1 Royal Superfamily

In this superfamily, domains are structurally related and have β -barrel topology. It is believed that they come from a common ancestor, which has the conserved binding ability with the methylated substrate. All the family members consist of a slightly curved β -barrel with three β -strands followed by a short 3_{10} helix, and different members are distinguished based on additional strands or helices. This family includes MBT, Tudor, chromodomain, and PWWP (Yap and Zhou 2010).

6.1.1 Recognition of H4K20me1/me2 by the MBT Domain of L3MBTL1

Isothermal titration calorimetry (ITC) assay using different methylation states of H4K20 peptide revealed that the MBT domain of human L3MBTL1 displays more affinity toward mono- and dimethylation states and does not bind to unmodified and trimethylated histone peptide. However, the binding is relatively of low affinity ($K_D = 5\text{--}40\ \mu\text{M}$) and promiscuous (Min et al., 2007). The crystal structure of L3MBTL1 with three MBT domains bound to histone peptide of 11 residues (H4 residue 15–25 with H4K20me2) shows the similar structure of all MBT domains assembled in a triangular shape. The structure is consistent with previously known structures of the MBT domain. It consists of four β -strands that form β -barrel and other extended arms of helices and a shorter strand (Wang et al., 2003). L3MBTL1 has three repeats of the MBT domain; however, only the second MBT domain binds



to methyllysine (Santiveri et al., 2008; Eryilmaz et al., 2009; Li et al., 2007; Min et al., 2007). Superimposition of all three MBT domains revealed that the shorter side chain of cysteine 363 in the aromatic cage of the second MBT was replaced by bulky amino acids in MBT1 and MBT3. Thus, steric hindrance prevents the binding of MBT1 and MBT3 with methylated lysine. It also preferentially reads mono- and dimethylation by the cavity insertion mode, and specificity toward a lower methylation state is because of the aspartate residue present in the aromatic cage. Typical interactions in the aromatic cage are cation- π between aromatic residues and positively charged methylammonium of methyllysine. Additionally, aspartate binds to methylammonium *via* hydrogen bonding in the MBT domain.

In contrast, the flanking residues of the peptide substrate show little interaction with the protein. Only two water-mediated hydrogen bonds are present, first between the backbone carbonyl group of H18 and Y386 of the protein and a second between the carbonyl NH group of K20 and N358 of the protein.

H4K20 methylation was previously linked to chromatin compaction. However, H4K20me1 was found in the actively transcribing genes, contradicting the previously suggested role of methylation in chromatin compaction. A recent study revealed

that H4K20me1 nucleosomal arrays were less compacted than H4K20me0 and H4K20me3 nucleosomal arrays, and the H4 tail was more dynamic in K20 mono-methylation. This study suggests that mono-methylation of H4K20 facilitates the opening and accessibility of chromatin (Shoab et al., 2021).

6.1.2 PWWP Domain

PWWP domain was first identified in the WHSC1 protein that contains the 100–130 amino acid structural motif (Stec et al., 1998). It also has a conserved Pro-Trp-Trp-Pro motif, and it consists of five β -strand barrels packed against the helical bundle. Despite the sequence conservation in different proteins, some variation in the PWWP motif can occur. For example, methyltransferase DNMT3a/b has SWWP (Qiu et al., 2002), and hepatoma-derived growth factor (HDGF) has a PHWP motif instead of a PWWP motif (Sue et al., 2004). It was initially identified as the DNA-binding protein; however, its similarity to the Tudor and chromodomain suggests that it might have the ability to bind methylated lysine. DNMT3a protein responsible for DNA methylation contains the PWWP domain. This domain is also known to interact with methylated histone tails, which led to assumptions that it might have dual binding to histone tails and the dsDNA.

Pdp1 protein, which contains PWWP domains, binds to methylated lysine, and dsDNA was seen by fluorescence polarization assay (FPA) (Qiu et al., 2012). The binding studies showed that PWWP domains of Pdp1 bind to the H3K20 trimethylation. After that, many other PWWP domains were shown to exhibit the binding with methylated lysine. Except for Pdp1, all the proteins containing PWWP domains bind specifically to the H3K36 methylation, suggesting its role as the H3K36 methylation sensor (Vezzoli et al., 2010; van Nuland et al., 2013; Wang et al., 2020). As it binds to the trimethylation state of lysine, it suggests that the binding cavity of the PWWP domain is wider to accommodate the bulkier me3 group than the MBT domain, which can only interact with mono- and dimethylated states. Therefore, this domain shows less specificity for the degree of methylation state.

Structural analysis of Pdp1 revealed that the aromatic cage is formed by Y63, W66, and F94. Cation- π interactions are used by the Pdp1 PWWP domain to recognize the trimethylated lysine at 20th position, and two residues (D97 and N99) from the loop between β 3 and β 4 form an extensive network of hydrogen bonds with the histone H4 tail residues (R19, K20, and V21) (Qiu et al., 2012). Y63, W66, and F94 amino acid side chains are perpendicularly oriented, forming an aromatic cage accommodating the trimethylammonium group. Mutations of the residues that compose the aromatic cage abolish methylated histone peptide binding.

6.1.3 Recognition of H3K9me3 Marks by HP1 Chromodomain

The chromatin organization modifier domain (chromodomain) is the smallest member of this superfamily. The structural motif is based on the HP1 fold, consisting of three curved antiparallel β -sheets followed by the α -helix (Ball et al., 1997). There are approximately 55 proteins identified which contain chromodomain. These chromodomain-containing proteins were associated with chromatin silencing. These proteins are divided into two groups: canonical (based on HP1 structure) group includes polycomb proteins Cbx1-9, CMT1-3, and CYD; and noncanonical, including CHD1-8, RBBP1, and HRP1.

HP1 and polycomb proteins recognize H3K9me3 (Jacobs and Khorasanizadeh 2002) and H3K27me3 (Min, Zhang, and Xu 2003), respectively, through their ARKS/T motif. The histone tail inserted between two strands forms the complete β -barrel in both proteins. This insertion of the histone tail is stabilized by the electrostatic interactions and the hydrogen bond between the backbone. This interaction involves seven amino acids preceding methyllysine and one following amino acid. This recognition method prefers the recognition of trimethylation over mono- and dimethylation. Recent cryo-EM structure of H3K9me3 dinucleosome with HP1 α , HP1 β , and HP1 γ revealed how heterochromatin is organized. In this structure, dinucleosomes trimethylated at K9 are bridged by two symmetric molecules of HP1. Linker DNA between the nucleosome is not interacting with the HP1, which leaves linker DNA to interact with ACF (ATP-utilizing chromatin assembly and remodeling factor) (Machida et al., 2018).

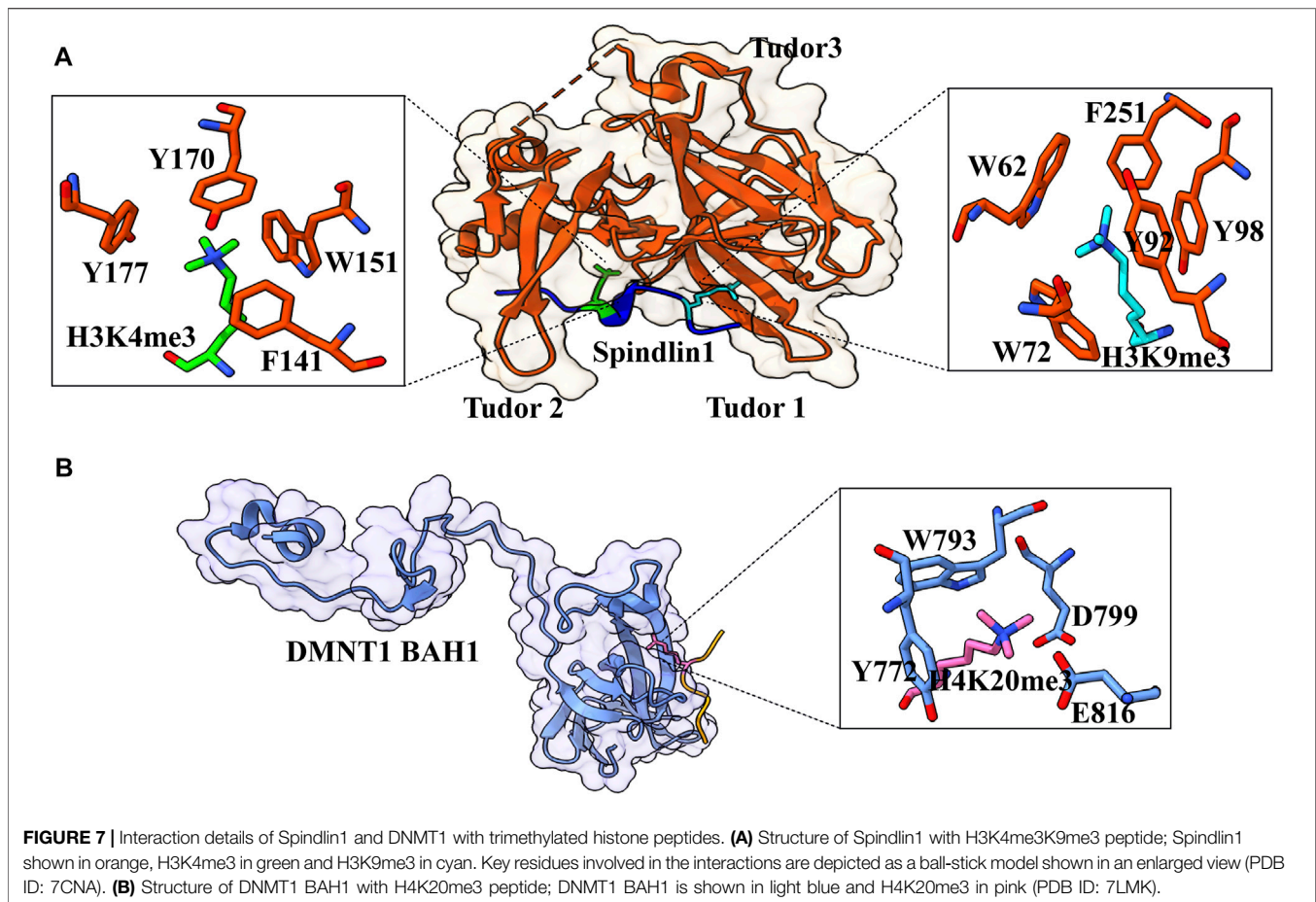
The noncanonical chromodomain proteins are based on the chromo ATPase/helicase-DNA-binding (CHD) protein. They contain two chromodomains, both at the N-terminal, for example, SNF2-type helicase, which is involved in chromatin remodeling. CHD7 specifically recognizes H3K1me1 as the enhancer for the gene.

6.1.4 Recognition of H3K4me3K9me3 Bivalent Mark by the Tudor Domain of Spindlin1

Tudor domains are structurally diverse and mediate protein-protein interactions. Tudor domains interact with all methylation states. This domain consists of approximately 60 amino acids of four or five β -strands which form a β -barrel structure followed by one or two helices (Selenko et al., 2001). Tudor domain-containing protein interacts with the H3K4me3 (Wang et al., 2011; Yang et al., 2012), H3K9me2 (Arita et al., 2012), H3K36me3 (Cai et al., 2013), and H4K20me3 (Hirano et al., 2012). Almost 30 known proteins have this Tudor domain, including JMJD2, 53BP1, SGF29, Spindlin1, UHRF1, PHF1, OHF19, LBR, and TDRD3 (which recognizes methylated arginine residues). Proteins in this family are involved in various biochemical processes like DNA methylation, nonhomologous end joining, DNA damage and repair, transcription activation and repression, and rRNA expression.

To meet the growing demand for ribosomes in rapidly growing cells, more copies of rRNA are produced at a greater transcription rate. The repressive histone methylation marks present on H3K9/K27 and H4K20 in the heterochromatin region are linked to rRNA transcription suppression. Because it has been established that H3K4 methylation is required for active gene expression, the cell must establish H3K4 methylation and H3K9 demethylation to convert the suppressed rRNA expression to active expression. Although H3K4 and H3K9 trimethylation are mutually exclusive, bivalent H3K4me3 and H3K9me3 have been documented in specific cell types (Mikkelsen et al., 2007; Bilodeau et al., 2009; Rugg-Gunn et al., 2010; Matsumura et al., 2015). It was previously reported that KDM4A and KDM4C recognize the H3K4me3 via their Tudor domain and help in the demethylation of H3K9me3. This study suggests that these methylation marks can coexist on the same H3 N-terminal tail and functionally crosstalk (Huang et al., 2006; Yamamoto, and Fujimori 2016). It was also previously known that euchromatin rRNA genes contain bivalent mark H3K4me3K9me3 (Murayama et al., 2008). So for rRNA synthesis, the bivalent mark of H3K4 and H3K9 trimethylation is needed.

The Tudor domain is found in Spindlin1, a protein that aids in rRNA expression. Spindlin1 creates a complex with C11orf84 that recognizes the bivalent mark on the histone H3 tail. Tudor 2 domain residues F141, W151, Y170, and Y177 form an aromatic pocket for trimethylated K4, whereas Tudor 1 domain residues W62, W72, Y91, and Y98, as well as Tudor 3 domain residue F251, recognize trimethylated K9 (Figure 7A) (Du et al., 2021). In addition to cation- π interactions formed by dual methylated lysine, the N-terminal amino group of H3A1 forms a hydrogen bond with the side-chain carboxylate group of D189, and guanidino moiety of H3R2 is ion-paired with



the side-chain carboxylate group of D184 (Du et al., 2021). This bivalent recognition further helps in the dislocation of HP1 from the rRNA chromatin, which relaxes the chromatin and helps in the recruitment of RNA polymerase I, which leads to the expression of rRNA. Spindlin1 Tudor 3 domain is responsible for the binding to C11orf84, while the other two Tudor domains act in concert to recognize a noncanonical bivalent histone mark H3K4me3K9me3.

6.2 The WD40 Repeats

WD40 repeats are present in proteins showing very diverse protein–protein interactions. WD repeats usually have 40–60 amino acids and conserved tryptophan–aspartate (WD) residue. This motif can be found in each blade of a WDR domain. The structural plasticity of WDR domains allows them to keep their β -propeller fold even after deletion of WD repeats, which can range from five to eight but are generally seven (Böcskei, and Polgár 1998; Juhász et al., 2005). Each of these repeats folds into a four-stranded β -sheet, and these propellers are large, usually containing ~300 amino acids.

SET1 methyltransferase (catalyze methylation of H3K4) subunit WDR5 contains seven WD repeats to form the β -propeller structure. SET1 needs WDR5 for its assembly and activity. First, it was found that WDR5 binds to

H3K4me2 and me3. However, the crystal structure of WDR5 with the unmodified, mono-, di-, and trimethylated H3K4 reveals that WDR5 interacts with H3R2 and acts as an arginine reader. WDR5 interacts with lysine only through E332, present at the protein's surface. WDR5 anchor as the arginine pocket binds to the unmodified and demethylated arginine residue (Dharmarajan et al., 2012). It was proposed to act as a histone modification intermediate that binds to arginine and presents lysine for methylation.

6.3 Recognition of H3K9me1/me2 by Ankyrin Repeat Domain of G9a

It is a ~33 amino acid repeat, and each repeat consists of helix–loop–helix structure with a β -hairpin/loop region; and in most proteins, 4–7 repeats are present, which stack onto each other to form the right-handed solenoid-like structure that helps in the protein–protein interaction (Sedgwick and Smerdon 1999). A short-distance interaction between inter α -helices helps form the solenoid, forming a typical globular protein shape. The structure of the AR domain is stabilized through inter- and intra-helical hydrophobic and hydrogen bonding *via* polar residues at the N-terminal. Hydrophobic interactions between P5 and H7 form the L-shape of the domain, and hydrogen bond

between T4 and H7 facilitates the formation of β -hairpin with the adjacent loop (Yuan et al., 2004).

The H3K9me1 and H3K9me2 peptides bind to G9a ankyrin repeats with low affinity ($K_D = 14$ and 6 mM, respectively), with a crystal structure of ankyrin repeats–H3K9me2 peptide complex was solved. The dimethylammonium group of K9me2 is placed into an aromatic pocket lined by three tryptophan residues and glutamate while the H3 peptide is sandwiched between β -turns and the fourth and fifth helices ankyrin repeats (Collins et al., 2008). Peptide residues 9–11, which comprise K9me2, are involved in intermolecular recognition. The formation of complexes is hampered by mutations of certain peptide or aromatic cage residues. G9a's SET domain-mediated methyltransferase activity is unaffected by changes in methyllysine recognition by the ankyrin repeat, indicating that the reading and writing domains work independently.

6.4 The Bromo Adjacent Homology Domain

Bromo adjacent homology (BAH) domain-containing proteins are frequently associated with chromatin functions, and mutations in these domains can result in diseases. BAH has been found to have various roles in chromatin biology, including protein–protein interaction, identification of methylated lysine, and DNA methylation. They have six tandem repeats of the bromodomain at the N-terminus, followed by the repetitive sequence motif of an unknown function; therefore, these are referred to as bromo adjacent homology domains. The majority of BAH domain proteins that have been identified have specific chromatin links, such as nucleosome remodeling and histone and DNA modifications.

BAH has an essential role in DNA methylation, which is an epigenetic mark that catalyzes the addition of the methyl group to the fifth position of cytosine (5-methylcytosine) (Suzuki and Bird 2008). Each round of DNA replication produces hemimethylated DNA, which must be converted to fully methylated DNA before the next round of replication, or the methylation marks will be lost. The family of DNA methyltransferases (Dnmts) mediates the transfer of the methyl group from S-adenyl methionine to the fifth position of cytosine at the CpG dinucleotide (Moore, Le, and Fan 2013). DNMT1 has N-terminal regulatory and C-terminal catalytic domains. Human Dnmt1 (1,616 amino acids) has two regions: N-terminal regulatory domain (1–1,139) and C-terminal catalytic domain (1,140–1,616). It also has several regions in N-terminal like CXXC (Bestor 1992), replication focus targeting sequence (RFTS) (Leonhardt et al., 1992), and BAH domain (Callebaut, Courvalin, and Mornon 1999).

H4K20 trimethylation is a significant heterochromatin mark that suppresses repetitive sequences in the human genome. By recognizing H4K20me3 *via* its first Bromo adjacent homology domain, DNMT1 promotes DNA methylation. The structure of DNMT1 with H4K20me3_{14–25} peptide shows that the side chain is inserted into a pocket created by DNMT1 BAH1 Y772, W793, D799, and E816 (**Figure 7B**) (Ren et al., 2021). The side chains of DNMT1 BAH1 D765 and E818 are bidentate hydrogen-bonded to the backbone amides of H4. BAH1 binds to H4K20me3, generating a change in the structure of DNMT1 that allows the autoinhibition linker to be displaced. DNMT1 can break

away from the linker's autoinhibition and get activated. H4K20me3 mark is put just after the S-phase of the cell cycle; therefore, H4K20me3 patterns may differ in new histones and parental histone. For a fact, recycled histone H4K20 is extensively methylated throughout replication, but fresh histones are only methylated during the G2/M phase (Ren et al., 2021). In this context, only the parental histones probably facilitate the DNMT1 activation, which lasts beyond the S-phase. This mechanism occurs alongside the UHRF-1-specific S-phase maintenance of DNA methylation (Rothbart et al., 2012).

6.5 PHD Domain

The PHD finger is a short zinc-binding module with a high cysteine content but few secondary structural elements that are now being characterized as a protein–protein or protein–phospholipid interaction domain. The PHD finger is a 50–80 amino acid motif composed of a two-strand antiparallel β -sheet and α -helix linked to the Cys4-His-Cys3 motif in a cross-brace shape by two zinc atoms (Bottomley et al., 2005; Elkin et al., 2005). The structure and function studies on PHD domain-containing proteins like BPFT (Li et al., 2006), IGN2 (Peña et al., 2006), and YNG1 (Taverna et al., 2006), which bind to H3K4 higher methylation state, provide the sequence and methylation state-specific recognition mechanism. Because of their ability to read numerous different post-translational modifications simultaneously, PHD proteins are some of the best examples of the “reader” class of proteins in combinatorial control of transcription. PHD fingers are found on a number of proteins involved in chromatin remodeling. Due to their frequent occurrence near other known chromatin interaction domains (bromodomains, PWWP domains), it was thought that PHD fingers could recognize histone modifications. In the various structures of PHD fingers that have been solved, the H3K4me3 mark can be found on most peptides. The modified histone peptide forms a strand that integrates into the PHD finger's existing antiparallel sheet, and the majority of the complex structures have a similar topology.

7 CONCLUSION

The advancement in mass spectrometry and structural biology field in the past few years has revolutionized the field of epigenetics, resulting in an ever-growing list of PTMs. At the same time, the list of proteins having the capability to write, read, and erase these epigenetic marks is also growing, further expanding the complexity of epigenetics. For example, the identification of several short-chain acylation marks and their readers are critical in linking the metabolic state of the cell to gene regulation as most of the substrates for these PTMs are derived from different metabolic reactions. In the past 20 years, a lot of efforts have been put into understanding the molecular and structural mechanism of PTM readouts by their cognate reader proteins. These studies have identified a wide range of interacting modes, including cation– π and π – π – π stacking interactions (Klein et al., 2018; Du et al., 2021). Selective recognition of PTM by its cognate reader or multivalent

readout of multiple PTMs on the same or different tails by a group of readers targets the protein or protein complex at targeted genomic sites and performs the downstream function. A deep understanding of these PTMs and their recognition by reader proteins is critical as any misregulation of these recognition mechanisms can lead to several human disorders like cancer. In case of misregulation, in-depth characterization of these binding mechanisms can aid in developing specific epigenetic-driven therapeutic targets. From the structural point of view, we have a solid knowledge base of different PTM–reader interactions. However, there are some questions that remain to be explicitly answered, concerning selectivity and specificity, like how the same reader can differentiate between acetyl and acyl modifications *in vivo*. Although a large number of PTMs have been identified and characterized, their reader partner is still to be

characterized. Maximum structural studies on reader–PTM interactions till date used only short peptides. Therefore, structural elucidation of reader–PTM interaction at the nucleosomal level and further at the nucleosomal array level remains a major challenge. With improving structural biology tools and computational methods in tandem, it will be easier to overcome the existing limitations and answer questions associated with the molecular mechanism of “decorated tails” recognition by the epigenetic machinery.

AUTHOR CONTRIBUTIONS

AK, PS, and RS: conceptualization, review, and editing; PS and RS: original draft.

REFERENCES

- Allahverdi, A., Yang, R., Korolev, N., Fan, Y., Davey, C. A., Liu, C.-F., et al. (2011). The Effects of Histone H4 Tail Acetylations on Cation-Induced Chromatin Folding and Self-Association. *Nucleic Acids Res.* 39 (5), 1680–1691. doi:10.1093/nar/gkq900
- Allfrey, V. G., Faulkner, R., and Mirsky, A. E. (1964). ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS. *Proc. Natl. Acad. Sci. U.S.A.* 51 (5), 786–794. doi:10.1073/pnas.51.5.786
- Andrews, F. H., Shinsky, S. A., ShanleShanle, E. K., Bridgers, J. B., Gest, A., TsunTsun, I. K., et al. (2016). The Taf14 YEATS Domain Is a Reader of Histone Crotonylation. *Nat. Chem. Biol.* 12 (6), 396–398. doi:10.1038/nchembio.2065
- Arita, K., Isogai, S., Oda, T., Unoki, M., Sugita, K., Sekiyama, N., et al. (2012). Recognition of Modification Status on a Histone H3 Tail by Linked Histone Reader Modules of the Epigenetic Regulator UHRF1. *Proc. Natl. Acad. Sci. U.S.A.* 109 (32), 12950–12955. doi:10.1073/pnas.1203701109
- Ball, L. J., MurzinaMurzina, N. V., Broadhurst, R. W., ArcherRaineSingh, S. J., Stott, F. J., Murzin, A. G., et al. (1997). Structure of the Chromatin Binding (Chromo) Domain from Mouse Modifier Protein 1. *EMBO J.* 16 (9), 2473–2481. doi:10.1093/emboj/16.9.2473
- Ballestar, E., Abad, C., and Franco, L. (1996). Core Histones Are Glutaminyl Substrates for Tissue Transglutaminase. *J. Biol. Chem.* 271 (31), 18817–18824. doi:10.1074/JBC.271.31.18817
- Bao, X., Liu, Z., Zhang, W., Gladysk, K., Fung, Y. M. E., Tian, G., XiongLi, Y., Wong, J. W. H., Yuen, K. W. Y., and Li, X. D. Yi Man Eva Fung; Jason Wing Hon Wong; Karen Wing Yee Yuen (2019). Glutarylation of Histone H4 Lysine 91 Regulates Chromatin Dynamics. *Mol. Cell.* 76 (4), 660–675. doi:10.1016/J.MOLCEL.2019.08.018
- Barski, A., Cuddapah, S., Cui, K., Roh, T.-Y., Schones, D. E., Wang, Z., et al. (2007). High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell.* 129 (4), 823–837. doi:10.1016/j.cell.2007.05.009
- Bascom, G. D., and Schlick, T. (2018). Chromatin Fiber Folding Directed by Cooperative Histone Tail Acetylation and Linker Histone Binding. *Biophysical J.* 114 (10), 2376–2385. doi:10.1016/j.bpj.2018.03.008
- Bernstein, B. E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., BaileyBailey, D. K., Huebert, D. J., et al. (2005). Genomic Maps and Comparative Analysis of Histone Modifications in Human and Mouse. *Cell.* 120 (2), 169–181. doi:10.1016/j.cell.2005.01.001
- Bestor, T. H. (1992). Activation of Mammalian DNA Methyltransferase by Cleavage of a Zn Binding Regulatory Domain. *EMBO J.* 11 (7), 2611–2617. doi:10.1002/j.1460-2075.1992.tb05326.x
- Bilodeau, S., Kagey, M. H., Frampton, G. M., RahlRahl, P. B., and Young, R. A. (2009). SetDB1 Contributes to Repression of Genes Encoding Developmental Regulators and Maintenance of ES Cell State. *Genes. Dev.* 23 (21), 2484–2489. doi:10.1101/gad.1837309
- Böhm, V., Hieb, A. R., Andrews, A. J., Gansen, A., Rocker, A., Tóth, K., et al. (2011). Nucleosome Accessibility Governed by the Dimer/Tetramer Interface. *Nucleic Acids Res.* 39 (8), 3093–3102. doi:10.1093/nar/gkq1279
- Bottomley, M. J., Stier, G., Pennacchini, D., Legube, G., Simon, B., Akhtar, A., et al. (2005). NMR Structure of the First PHD Finger of Autoimmune Regulator Protein (AIRE1). *J. Biol. Chem.* 280 (12), 11505–11512. doi:10.1074/jbc.M413959200
- Brehove, M., Wang, T., North, J., Luo, Y., Dreher, S. J., Shimko, J. C., et al. (2015). Histone Core Phosphorylation Regulates DNA Accessibility. *J. Biol. Chem.* 290 (37), 22612–22621. doi:10.1074/jbc.M115.661363
- Cai, L., RothbartRothbart, S. B., Lu, R., Xu, B., Chen, W.-Y., Tripathy, A., et al. (2013). An H3K36 Methylation-Engaging Tudor Motif of Polycomb-like Proteins Mediates PRC2 Complex Targeting. *Mol. Cell.* 49 (3), 571–582. doi:10.1016/j.molcel.2012.11.026
- Callebaut, I., Courvalin, J.-C., and Mornon, J.-P. (1999). The BAH (Bromo-Adjacent Homology) Domain: A Link between DNA Methylation, Replication and Transcriptional Regulation. *FEBS Lett.* 446 (1), 189–193. doi:10.1016/s0014-5793(99)00132-5
- Carey, M., Li, B., and Workman, J. L. (2006). RSC Exploits Histone Acetylation to Abrogate the Nucleosomal Block to RNA Polymerase II Elongation. *Mol. Cell.* 24 (3), 481–487. doi:10.1016/j.molcel.2006.09.012
- Chen, G., Li, W., Yan, F., Wang, D., and Chen, Y. (2020). The Structural Basis for Specific Recognition of H3K14 Acetylation by Sth1 in the RSC Chromatin Remodeling Complex. *Structure* 28 (1), 111–118. doi:10.1016/j.str.2019.10.015
- Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S. C., et al. (2007). Lysine Propionylation and Butyrylation Are Novel Post-Translational Modifications in Histones. *Mol. Cell. Proteomics* 6 (5), 812–819. doi:10.1074/MCP.M700021-MCP200
- Chi, P., Allis, C. D., and Wang, G. G. (2010). Covalent Histone Modifications - Miswritten, Misinterpreted and Mis-Erased in Human Cancers. *Nat. Rev. Cancer* 10 (7), 457–469. doi:10.1038/nrc2876
- Colleparado-Guevara, R., Portella, G., Vendruscolo, M., Frenkel, D., Schlick, T., and Orozco, M. (2015). Chromatin Unfolding by Epigenetic Modifications Explained by Dramatic Impairment of Internucleosome Interactions: A Multiscale Computational Study. *J. Am. Chem. Soc.* 137 (32), 10205–10215. doi:10.1021/jacs.5b04086
- Collins, R. E., Northrop, J. P., Horton, J. R., Lee, D. Y., Zhang, X., Stallcup, M. R., et al. (2008). The Ankyrin Repeats of G9a and GLP Histone Methyltransferases Are Mono- and Dimethyllysine Binding Modules. *Nat. Struct. Mol. Biol.* 15 (3), 245–250. doi:10.1038/nsmb.1384
- Dai, L., Peng, C., Montellier, E., Lu, Z., Chen, Y., Ishii, H., et al. (2014). Lysine 2-Hydroxyisobutyrylation Is a Widely Distributed Active Histone Mark. *Nat. Chem. Biol.* 10 (5), 365–370. doi:10.1038/NCHEMBIO.1497
- Dhalluin, C., CarlsonL Zeng, C He, J. E. A. K. Aggarwal, Zeng, L., He, C., Aggarwal, A. K., Zhou, M.-M., et al. (1999). Structure and Ligand of a Histone Acetyltransferase Bromodomain. *Nature* 399 (6735), 491–496. doi:10.1038/20974

- Dharmarajan, V., Lee, J.-H., Patel, A., Skalnik, D. G., and Cosgrove, M. S. (2012). Structural Basis for WDR5 Interaction (Win) Motif Recognition in Human SET1 Family Histone Methyltransferases. *J. Biol. Chem.* 287 (33), 27275–27289. doi:10.1074/jbc.M112.364125
- Dou, Y., and Gorovsky, M. A. (2000). Phosphorylation of Linker Histone H1 Regulates Gene Expression *In Vivo* by Creating a Charge Patch. *Mol. Cell.* 6 (2), 225–231. doi:10.1016/s1097-2765(00)00024-1
- Du, Y., Yan, Y., Xie, S., Huang, H., Wang, X., NgNg, R. K., et al. (2021). Structural Mechanism of Bivalent Histone H3K4me3K9me3 Recognition by the Spindlin1/C11orf84 Complex in RRNA Transcription Activation. *Nat. Commun.* 12 (1), 949. doi:10.1038/s41467-021-21236-x
- Duan, M.-R., and Smerdon, M. J. (2014). 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 (12), 8353–8363. doi:10.1074/jbc.M113.540732
- Dutta, A., Abmayr, S. M., and Workman, J. L. (2016). Diverse Activities of Histone Acylations Connect Metabolism to Chromatin Function. *Mol. Cell.* 63 (4), 547–552. doi:10.1016/j.molcel.2016.06.038
- Elkin, S. K., Ivanov, D., Ewalt, M., Ferguson, C. G., Hyberts, S. G., SunSun, Z.-Y. J. J., et al. (2005). A PHD Finger Motif in the C Terminus of RAG2 Modulates Recombination Activity. *J. Biol. Chem.* 280 (31), 28701–28710. doi:10.1074/jbc.M504731200
- Enríquez, P., Krajewski, K., Strahl, B. D., RothbartRothbart, S. B., Downen, R. H., and Rose, R. B. (2021). Binding Specificity and Function of the SWI/SNF Subunit SMARCA4 Bromodomain Interaction with Acetylated Histone H3K14. *J. Biol. Chem.* 297 (4), 101145. doi:10.1016/j.jbc.2021.101145
- Eryilmaz, J., Pan, P., Amaya, A., Dong, A., Adams-Cioaba, M. A., et al. (2009). Structural Studies of a Four-MBT Repeat Protein MBTD1. *PLOS ONE* 4 (10), e7274–7. doi:10.1371/journal.pone.0007274
- Farrelly, L. A., Thompson, R. E., Zhao, S., Lepack, A. E., Lyu, Y., BhanuBhanu, N. V., et al. (2019). Histone Serotonylation Is a Permissive Modification that Enhances TFIID Binding to H3K4me3. *Nature* 567 (7749), 535–539. doi:10.1038/s41586-019-1024-7
- Fenley, A. T., Anandakrishnan, R., Kidane, Y. H., and OnufrievOnufriev, A. V. (2018). Modulation of Nucleosomal DNA Accessibility via Charge-Altering Post-Translational Modifications in Histone Core. *Epigenetics Chromatin* 11 (1), 11. doi:10.1186/s13072-018-0181-5
- Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.-P., Barsyte-Lovejoy, D., et al. (2012). Histone Recognition and Large-Scale Structural Analysis of the Human Bromodomain Family. *Cell.* 149 (1), 214–231. doi:10.1016/j.cell.2012.02.013
- Filippakopoulos, P., Qi, J., Picaud, S., Shen, Y., Smith, W. B., Fedorov, O., et al. (2010). Selective Inhibition of BET Bromodomains. *Nature* 468 (7327), 1067–1073. doi:10.1038/nature09504
- Flynn, E. M., Huang, O. W., Poy, F., Oppikofer, M., BellonBellon, S. F., Tang, Y., et al. (2015). A Subset of Human Bromodomains Recognizes Butyryllysine and Crotonyllysine Histone Peptide Modifications. *Structure* 23 (10), 1801–1814. doi:10.1016/j.str.2015.08.004
- Fülöp, V., Böcskei, Z., and Polgár, L. (1998). Prolyl Oligopeptidase. *Cell.* 94 (2), 161–170. doi:10.1016/s0092-8674(00)81416-6
- Gansen, A., Felekyan, S., Kühnemuth, R., Lehmann, K., Tóth, K., Seidel, C. A. M., et al. (2018). High Precision FRET Studies Reveal Reversible Transitions in Nucleosomes between Microseconds and Minutes. *Nat. Commun.* 9 (1), 4628. doi:10.1038/s41467-018-06758-1
- Ghoneim, M., Fuchs, H. A., and Musselman, C. A. (2021). Histone Tail Conformations: A Fuzzy Affair with DNA. *Trends Biochem. Sci.* 46 (7), 564–578. doi:10.1016/j.tibs.2020.12.012
- Gowans, G. J., Bridgers, J. B., Zhang, J., Dronamraju, R., Burnetti, A., King, D. A., et al. (2019). Recognition of Histone Crotonylation by Taf14 Links Metabolic State to Gene Expression. *Mol. Cell.* 76 (6), 909–921. e3. doi:10.1016/j.molcel.2019.09.029
- Heintzman, N. D., StuartStuart, R. K., Hon, G., Fu, Y., ChingHawkins, C. W., HawkinsBarrera, R. D., et al. (2007). Distinct and Predictive Chromatin Signatures of Transcriptional Promoters and Enhancers in the Human Genome. *Nat. Genet.* 39 (3), 311–318. doi:10.1038/ng1966
- Hirano, Y., Hizume, K., Kimura, H., Takeyasu, K., Haraguchi, T., and Hiraoka, Y. (2012). Lamin B Receptor Recognizes Specific Modifications of Histone H4 in Heterochromatin Formation. *J. Biol. Chem.* 287 (51), 42654–42663. doi:10.1074/jbc.M112.397950
- Huang, H., Zhang, D., Wang, Y., Perez-Neut, M., Han, Z., Zheng, Y. G., et al. (2018). Lysine Benzoylation Is a Histone Mark Regulated by SIRT2. *Nat. Commun.* 9 (1). doi:10.1038/s41467-018-05567-W
- Huang, Y., Fang, J., Bedford, M. T., Zhang, Y., and Xu, R.-M. (2006). Recognition of Histone H3 Lysine-4 Methylation by the Double Tudor Domain of JMJD2A. *Science* 312 (5774), 748–751. doi:10.1126/science.1125162
- Hummerich, R., Thumfart, J.-O., Findeisen, P., Bartsch, D., and Schloss, P.J.; rg-Oliver Thumfart (2012). Transglutaminase-Mediated Transamidation of Serotonin, Dopamine and Noradrenaline to Fibronectin: Evidence for a General Mechanism of Monoamination. *FEBS Lett.* 586 (19), 3421–3428. doi:10.1016/j.febslet.2012.07.062
- Jacobs, S. A., and Khorasanizadeh, S. (2002). Structure of HP1 Chromodomain Bound to a Lysine 9-Methylated Histone H3 Tail. *Science* 295 (5562), 2080–2083. doi:10.1126/science.1069473
- Jacobson, R. H., Ladurner, A. G., King, D. S., and Tjian, R. (2000). Structure and Function of a Human TAF II 250 Double Bromodomain Module. *Science* 288 (5470), 1422–1425. doi:10.1126/science.288.5470.1422
- Juhász, T., Szeltner, Z., Fülöp, V., and Polgár, L. (2005). Unclosed β -Propellers Display Stable Structures: Implications for Substrate Access to the Active Site of Prolyl Oligopeptidase. *J. Mol. Biol.* 346 (3), 907–917. doi:10.1016/j.jmb.2004.12.014
- Klein, B. J., Jang, S. M., Lachance, C., Mi, W., Lyu, J., Sakuraba, S., et al. (2019). Histone H3K23-specific Acetylation by MORF Is Coupled to H3K14 Acylation. *Nat. Commun.* 10 (1), 4724. doi:10.1038/s41467-019-12551-5
- Klein, B. J., Smithy, J., Wang, X., Ahn, J., Andrews, F. H., Zhang, Y., et al. (2017). Recognition of Histone H3K14 Acylation by MORF. *Structure* 25 (4), 650–654. e2. doi:10.1016/j.str.2017.02.003
- Klein, B. J., Vann, K. R., Andrews, F. H., Wang, W. W., Zhang, J., Zhang, Y., et al. (2018). Structural Insights into the π - π - π Stacking Mechanism and DNA-Binding Activity of the YEATS Domain. *Nat. Commun.* 9 (1), 4574. doi:10.1038/s41467-018-07072-6
- Krone, M. W., Travis, C. R., LeeHannah J Eckvahl, G. Y., Eckvahl, H. J., Houk, K. N., and Waters, M. L. (2020). More Than π - π - π Stacking: Contribution of Amide- π and CH- π Interactions to Crotonyllysine Binding by the AF9 YEATS Domain. *J. Am. Chem. Soc.* 142 (40), 17048–17056. doi:10.1021/jacs.0c06568
- Le Masson, I., Yu, D. Y., Jensen, K., Chevalier, A., Courbeyrette, R., Boulard, Y., et al. (2003). Yaf9, a Novel NuA4 Histone Acetyltransferase Subunit, Is Required for the Cellular Response to Spindle Stress in Yeast. *Mol. Cell Biol.* 23 (17), 6086–6102. doi:10.1128/mcb.23.17.6086-6102.2003
- Leonhardt, H., Page, A. W., Weier, H.-U., and Bestor, T. H. (1992). A Targeting Sequence Directs DNA Methyltransferase to Sites of DNA Replication in Mammalian Nuclei. *Cell.* 71 (5), 865–873. doi:10.1016/0092-8674(92)90561-p
- Levy, M. J., Montgomery, D. C., Sardi, M. E., Sardi, M. E., Montano, J. L., Bergholtz, S. E., et al. (2020). A Systems Chemoproteomic Analysis of Acyl-CoA/Protein Interaction Networks. *Cell. Chem. Biol.* 27 (3), 322–333. e5. doi:10.1016/j.CHEMBIOL.2019.11.011
- Li, H., Fischle, W., Wang, W., Duncan, E. M., Liang, L., Murakami-Ishibe, S., et al. (2007). Structural Basis for Lower Lysine Methylation State-specific Readout by MBT Repeats of L3MBTL1 and an Engineered PHD Finger. *Mol. Cell.* 28 (4), 677–691. doi:10.1016/j.molcel.2007.10.023
- Li, H., Ilin, S., Wang, W., Duncan, E. M., Wysocka, J., Allis, C. D., et al. (2006). Molecular Basis for Site-specific Read-Out of Histone H3K4me3 by the BPTF PHD Finger of NURF. *Nature* 442 (7098), 91–95. doi:10.1038/nature04802
- Li, Y., Sabari, B. R., Panchenko, T., Wen, H., Zhao, D., Guan, H., et al. (2016). Molecular Coupling of Histone Crotonylation and Active Transcription by AF9 YEATS Domain. *Mol. Cell.* 62 (2), 181–193. doi:10.1016/j.molcel.2016.03.028
- Li, Y., Wen, H., Xi, Y., Tanaka, K., Wang, H., Peng, D., et al. (2014). AF9 YEATS Domain Links Histone Acetylation to DOT1L-Mediated H3K79 Methylation. *Cell.* 159 (3), 558–571. doi:10.1016/j.cell.2014.09.049
- Lorch, Y., Griesenbeck, J., Boeger, H., Maier-Davis, B., and Kornberg, R. D. (2011). Selective Removal of Promoter Nucleosomes by the RSC Chromatin-Remodeling Complex. *Nat. Struct. Mol. Biol.* 18 (8), 881–885. doi:10.1038/nsmb.2072
- Lorch, Y., Maier-Davis, B., and Kornberg, R. D. (2018). Histone Acetylation Inhibits RSC and Stabilizes the +1 Nucleosome. *Mol. Cell.* 72 (3), 594–600. doi:10.1016/j.molcel.2018.09.030

- Luger, K., MäderMäder, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997). Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution. *Nature* 389 (6648), 251–260. doi:10.1038/38444
- Machida, S., Takizawa, Y., Ishimaru, M., Sugita, Y., Sekine, S., Nakayama, J.-i., et al. (2018). Structural Basis of Heterochromatin Formation by Human HP1. *Mol. Cell* 69 (3), 385–397. e8. doi:10.1016/j.molcel.2017.12.011
- Marmorstein, R., and Zhou, M.-M. (2014). Writers and Readers of Histone Acetylation: Structure, Mechanism, and Inhibition. *Cold Spring Harb. Perspect. Biol.* 6 (7), a018762, a018762. doi:10.1101/cshperspect.a018762
- Matsumura, Y., Nakaki, R., Inagaki, T., Yoshida, A., Kano, Y., Kimura, H., et al. (2015). H3K4/H3K9me3 Bivalent Chromatin Domains Targeted by Lineage-specific DNA Methylation Pauses Adipocyte Differentiation. *Mol. Cell* 60 (4), 584–596. doi:10.1016/j.molcel.2015.10.025
- Matthews, A. G. W., Kuo, A. J., Ramón-Maiques, S., Han, S., Champagne, K. S., Ivanov, D., et al. (2007). RAG2 PHD Finger Couples Histone H3 Lysine 4 Trimethylation with V(D)J Recombination. *Nature* 450 (7172), 1106–1110. doi:10.1038/nature06431
- Mattiroli, F., and Penengo, L. (2021). Histone Ubiquitination: An Integrative Signaling Platform in Genome Stability. *Trends Genet.* 37 (6), 566–581. doi:10.1016/j.tig.2020.12.005
- Maurer-Stroh, S., Nicholas J. Dickens, Dickens, N. J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F., and Ponting, C. P. (2003). The Tudor Domain 'Royal Family': Tudor, Plant Agenet, Chromo, PWWP and MBT Domains. *Trends Biochem. Sci.* 28 (2), 69–74. doi:10.1016/S0968-0004(03)00004-5
- Mikkelsen, T. S., Ku, M., Jaffe, D. B., Issac, B., Lieberman, E., Giannoukos, G., et al. (2007). Genome-Wide Maps of Chromatin State in Pluripotent and Lineage-Committed Cells. *Nature* 448 (7153), 553–560. doi:10.1038/nature06008
- Millán-Zambrano, G., Burton, A., Bannister, A. J., and Schneider, R. (2022). Histone Post-Translational Modifications - Cause and Consequence of Genome Function. *Nat. Rev. Genet.* doi:10.1038/s41576-022-00468-7
- Min, J., Allali-Hassani, A., Nady, N., Qi, C., Ouyang, H., Liu, Y., et al. (2007). L3MBTL1 Recognition of Mono- and Dimethylated Histones. *Nat. Struct. Mol. Biol.* 14 (12), 1229–1230. doi:10.1038/nsmb1340
- Min, J., Zhang, Y., and Xu, R.-M. (2003). Structural Basis for Specific Binding of Polycomb Chromodomain to Histone H3 Methylated at Lys 27. *Genes. Dev.* 17 (15), 1823–1828. doi:10.1101/gad.269603
- Moore, L. D., Le, T., and Fan, G. (2013). DNA Methylation and its Basic Function. *Neuropsychopharmacol* 38 (1), 23–38. doi:10.1038/npp.2012.112
- Murayama, A., Ohmori, K., Fujimura, A., Minami, H., Yasuzawa-Tanaka, K., Kuroda, T., et al. (2008). Epigenetic Control of RDNA Loci in Response to Intracellular Energy Status. *Cell* 133 (4), 627–639. doi:10.1016/j.cell.2008.03.030
- Musselman, C. A., Lalonde, M.-E., Côté, J., and Kutateladze, T. G. (2012). Perceiving the Epigenetic Landscape through Histone Readers. *Nat. Struct. Mol. Biol.* 19 (12), 1218–1227. doi:10.1038/nsmb.2436
- Nitsch, S., Zorro Shahidian, L., and Schneider, R. (2021). Histone Acylations and Chromatin Dynamics: Concepts, Challenges, and Links to Metabolism. *EMBO Rep.* 22 (7), e52774. doi:10.15252/EMBR.202152774
- Pack, L. R., Yamamoto, K. R., and Fujimori, D. G. (2016). Opposing Chromatin Signals Direct and Regulate the Activity of Lysine Demethylase 4C (KDM4C). *J. Biol. Chem.* 291 (12), 6060–6070. doi:10.1074/jbc.M115.696864
- Peña, P. V., Davrazou, F., Shi, X., Walter, K. L., VerkhushaVerkhusha, V. V., Gozani, O., et al. (2006). Molecular Mechanism of Histone H3K4me3 Recognition by Plant Homeodomain of ING2. *Nature* 442 (7098), 100–103. doi:10.1038/nature04814
- Peng, Y., Li, S., Onufriev, A., Landsman, D., and Panchenko, A. R. (2021). Binding of Regulatory Proteins to Nucleosomes Is Modulated by Dynamic Histone Tails. *Nat. Commun.* 12 (1), 5280. doi:10.1038/s41467-021-25568-6
- Polach, K. J., and Widom, J. (1995). Mechanism of Protein Access to Specific DNA Sequences in Chromatin: A Dynamic Equilibrium Model for Gene Regulation. *J. Mol. Biol.* 254 (2), 130–149. doi:10.1006/jmbi.1995.0606
- Qiu, C., Sawada, K., Zhang, X., and Cheng, X. (2002). The PWWP Domain of Mammalian DNA Methyltransferase Dnmt3b Defines a New Family of DNA-Binding Folds. *Nat. Struct. Mol. Biol.* 9 (3), 217–224. doi:10.1038/nsb759
- Qiu, Y., Zhang, W., Zhao, C., Wang, Y., Wang, W., Zhang, J., et al. (2012). Solution Structure of the Pdp1 PWWP Domain Reveals its Unique Binding Sites for Methylated H4K20 and DNA. *Biochem. J.* 442 (3), 527–538. doi:10.1042/BJ20111885
- Ren, W., Fan, H., Grimm, S. A., Kim, J. J., Li, L., Guo, Y., et al. (2021). DNMT1 Reads Heterochromatic H4K20me3 to Reinforce LINE-1 DNA Methylation. *Nat. Commun.* 12 (1), 2490. doi:10.1038/s41467-021-22665-4
- Rothbart, S. B., Krajewski, K., Nady, N., Tempel, W., XueBadeaux, S., Badeaux, A. I., et al. Wolfram Tempel (2012). Association of UHRF1 with Methylated H3K9 Directs the Maintenance of DNA Methylation. *Nat. Struct. Mol. Biol.* 19 (11), 1155–1160. doi:10.1038/nsmb.2391
- Rugg-Gunn, P. J., Cox, B. J., Ralston, A., Rossant, J., and Rossant, Janet (2010). Distinct Histone Modifications in Stem Cell Lines and Tissue Lineages from the Early Mouse Embryo. *Proc. Natl. Acad. Sci. U.S.A.* 107 (24), 10783–10790. doi:10.1073/pnas.0914507107
- Ryu, H.-Y., and Hochstrasser, M. (2021). Histone Sumoylation and Chromatin Dynamics. *Nucleic Acids Res.* 49 (11), 6043–6052. doi:10.1093/nar/gkab280
- Sabari, B. R., Tang, Z., Huang, H., Yong-Gonzalez, V., Molina, H., KongKong, H. E., et al. (2015). Intracellular Crotonyl-CoA Stimulates Transcription through P300-Catalyzed Histone Crotonylation. *Mol. Cell* 58 (2), 203–215. doi:10.1016/j.molcel.2015.02.029
- Sanchez, R., and Zhou, M.-M. (2009). The Role of Human Bromodomains in Chromatin Biology and Gene Transcription. *Curr. Opin. Drug Discov. Dev.* 12 (5), 659–665. Available at: <https://pubmed.ncbi.nlm.nih.gov/19736624>.
- Santiveri, C. M., Lechtenberg, B. C., Allen, M. D., Sathyamurthy, A., JaulentJaulent, A. M., Freund, S. M. V., et al. (2008). The Malignant Brain Tumor Repeats of Human SCML2 Bind to Peptides Containing Monomethylated Lysine. *J. Mol. Biol.* 382 (5), 1107–1112. doi:10.1016/j.jmb.2008.07.081
- Sattler, M., Selenko, P., Sprangers, R., Stier, G., Bühler, D., and Fischer, U. (2001). SMN Tudor Domain Structure and Its Interaction with the Sm Proteins. *Nat. Struct. Mol. Biol.* 8 (1), 27–31. doi:10.1038/83014
- Sawicka, A., and Seiser, C. (2014). Sensing Core Histone Phosphorylation - a Matter of Perfect Timing. *Biochimica Biophysica Acta (BBA) - Gene Regul. Mech.* 1839 (8), 711–718. doi:10.1016/j.bbaggm.2014.04.013
- Schulze, J. M., Wang, A. Y., and Kobor, M. S. (2009). YEATS Domain Proteins: a Diverse Family with Many Links to Chromatin Modification and transcription. This Paper Is One of a Selection of Papers Published in This Special Issue, Entitled CSBMCB's 51st Annual Meeting - Epigenetics and Chromatin Dynamics, and Has Undergone the Journal's Usual Peer Review Process. *Biochem. Cell. Biol.* 87 (1), 65–75. doi:10.1139/o08-111
- Sedgwick, S. G., and Smerdon, S. J. (1999). The Ankyrin Repeat: A Diversity of Interactions on a Common Structural Framework. *Trends Biochem. Sci.* 24 (8), 311–316. doi:10.1016/S0968-0004(99)01426-7
- Shen, H., and Laird, P. W. (2013). Interplay between the Cancer Genome and Epigenome. *Cell* 153 (1), 38–55. doi:10.1016/j.cell.2013.03.008
- Shoib, M., Chen, Q., Shi, X., Nair, N., Prasanna, C., Yang, R., et al. (2021). Histone H4 Lysine 20 Mono-Methylation Directly Facilitates Chromatin Openness and Promotes Transcription of Housekeeping Genes. *Nat. Commun.* 12 (1), 4800. doi:10.1038/s41467-021-25051-2
- Shogren-KnaakShogren-Knaak, M., Ishii, H., Sun, J.-M., Pazin, M. J., Davie, J. R., and Peterson, C. L. (2006). Histone H4-K16 Acetylation Controls Chromatin Structure and Protein Interactions. *Science* 311 (5762), 844–847. doi:10.1126/science.1124000
- Shukla, S., Agarwal, P., and Kumar, A. (2022). Disordered Regions Tune Order in Chromatin Organization and Function. *Biophys. Chem.* 281 (February), 106716. doi:10.1016/j.bpc.2021.106716
- Smith, E., Lin, C., and Shilatifard, A. (2011). The Super Elongation Complex (SEC) and MLL in Development and Disease. *Genes. Dev.* 25 (7), 661–672. doi:10.1101/gad.2015411
- Stec, I., Wright, T. J., van Ommende Boer, G.-J. B., de Boer, P. A. J., van Haeringen, A., Moorman, A. F. M., et al. (1998). WHSC1, a 90 Kb SET Domain-Containing Gene, Expressed in Early Development and Homologous to a Drosophila Dysmorphia Gene Maps in the Wolf-Hirschhorn Syndrome Critical Region and Is Fused to IgH in T(1;14) Multiple Myeloma. *Hum. Mol. Genet.* 7 (7), 1071–1082. doi:10.1093/hmg/7.7.1071
- Strahl, B. D., and Allis, C. D. (2000). The Language of Covalent Histone Modifications. *Nature* 403 (6765), 41–45. doi:10.1038/47412
- Sue, S.-C., Chen, J.-Y., Lee, S.-C., Wu, W.-g., and Huang, T.-h. (2004). Solution Structure and Heparin Interaction of Human Hepatoma-Derived Growth Factor. *J. Mol. Biol.* 343 (5), 1365–1377. doi:10.1016/j.jmb.2004.09.014

- Suzuki, M. M., and Bird, A. (2008). DNA Methylation Landscapes: Provocative Insights from Epigenomics. *Nat. Rev. Genet.* 9 (6), 465–476. doi:10.1038/nrg2341
- Tan, M., Luo, H., Lee, S., Jin, F., Yang, J. S., Montellier, E., et al. (2011). Identification of 67 Histone Marks and Histone Lysine Crotonylation as a New Type of Histone Modification. *Cell*. 146 (6), 1016–1028. doi:10.1016/j.CELL.2011.08.008
- Taverna, S. D., Ilin, S., Rogers, R. S., Tanny, J. C., Lavender, H., Li, H., et al. (2006). Yng1 PHD Finger Binding to H3 Trimethylated at K4 Promotes NuA3 HAT Activity at K14 of H3 and Transcription at a Subset of Targeted ORFs. *Mol. Cell*. 24 (5), 785–796. doi:10.1016/j.molcel.2006.10.026
- Treviño, L. S., Wang, Q., and Walker, C. L. (2015). Phosphorylation of Epigenetic "readers, Writers and Erasers": Implications for Developmental Reprogramming and the Epigenetic Basis for Health and Disease. *Prog. Biophys Mol. Biol.* 118 (1–2), 8–13. doi:10.1016/j.pbiomolbio.2015.02.013
- van Nuland, R., van Schaik, F. M., Simonis, M., van Heesch, S., Cuppen, E., Boelens, R., et al. (2013). Nucleosomal DNA Binding Drives the Recognition of H3K36-Methylated Nucleosomes by the PSIP1-PWWP Domain. *Epigenetics Chromatin* 6 (1), 12. doi:10.1186/1756-8935-6-12
- Vezzoli, A., Bonadies, N., Allen, M. D., Freund, S. M. V., Santiveri, C. M., KvinlaugKvinlaug, B. T., et al. (2010). Molecular Basis of Histone H3K36me3 Recognition by the PWWP Domain of Brpf1. *Nat. Struct. Mol. Biol.* 17 (5), 617–619. doi:10.1038/nsmb.1797
- Vollmuth, F., and Geyer, M. (2010). Interaction of Propionylated and Butyrylated Histone H3 Lysine Marks with Brd4 Bromodomains. *Angew. Chem. Int. Ed.* 49 (38), 6768–6772. doi:10.1002/ANIE.201002724
- Walther, D. J., Peter, J.-U., Winter, S., Hölte, M., Paulmann, N., Grohmann, M., et al. (2003). Serotonylation of Small GTPases Is a Signal Transduction Pathway that Triggers Platelet α -Granule Release. *Cell*. 115 (7), 851–862. doi:10.1016/s0092-8674(03)01014-6
- Wang, G. G., Song, J., Wang, Z., Dormann, H. L., Casadio, F., Li, H., et al. (2009). Haematopoietic Malignancies Caused by Dysregulation of a Chromatin-Binding PHD Finger. *Nature* 459 (7248), 847–851. doi:10.1038/nature08036
- Wang, H., Farnung, L., Dienemann, C., and Cramer, P. (2020). Structure of H3K36-Methylated Nucleosome-PWWP Complex Reveals Multivalent Cross-Gyre Binding. *Nat. Struct. Mol. Biol.* 27 (1), 8–13. doi:10.1038/s41594-019-0345-4
- Wang, W., Chen, Z., Mao, Z., Zhang, H., Ding, X., Chen, S., et al. (2011). Nucleolar Protein Spindlin1 Recognizes H3K4 Methylation and Stimulates the Expression of RRNA Genes. *EMBO Rep.* 12 (11), 1160–1166. doi:10.1038/embor.2011.184
- Wang, W. K., Tereshko, V., Bocconi, P., MacGrogan, D., Nimer, S. D., and Patel, D. J. (2003). Malignant Brain Tumor Repeats. *Structure* 11 (7), 775–789. doi:10.1016/S0969-2126(03)00127-8
- Watts, S. W., PriestleyPriestley, J. R. C., and ThompsonThompson, J. M. (2009). Serotonylation of Vascular Proteins Important to Contraction. *PLoS One* 4 (5), e5682. doi:10.1371/journal.pone.0005682
- Xia, C., Tao, Y., Li, M., Che, T., and Qu, J. (2020). Protein Acetylation and Deacetylation: An Important Regulatory Modification in Gene Transcription (Review). *Exp. Ther. Med.* 20 (4), 2923–2940. doi:10.3892/etm.2020.9073
- Xie, Z., Dai, J., Dai, L., Tan, M., Cheng, Z., Wu, Y., et al. (2012). Lysine Succinylation and Lysine Malonylation in Histones. *Mol. Cell. Proteomics* 11 (5), 100–107. doi:10.1074/MCP.M111.015875
- Xie, Z., Zhang, D., Chung, D., Tang, Z., Huang, H., Dai, L., et al. (2016). Metabolic Regulation of Gene Expression by Histone Lysine β -Hydroxybutyrylation. *Mol. Cell*. 62 (2), 194–206. doi:10.1016/j.MOLCEL.2016.03.036
- Xiong, X., Panchenko, T., Yang, S., Zhao, S., Yan, P., Zhang, W., et al. (2016). Selective Recognition of Histone Crotonylation by Double PHD Fingers of MOZ and DP2. *Nat. Chem. Biol.* 12 (12), 1111–1118. doi:10.1038/nchembio.2218
- Yang, N., Wang, W., Wang, Y., Wang, M., Zhao, Q., Rao, Z., et al. (2012). Distinct Mode of Methylated Lysine-4 of Histone H3 Recognition by Tandem Tudor-like Domains of Spindlin1. *Proc. Natl. Acad. Sci. U.S.A.* 109 (44), 17954–17959. doi:10.1073/pnas.1208517109
- Yap, K. L., and Zhou, M.-M. (2010). Keeping it in the Family: Diverse Histone Recognition by Conserved Structural Folds. *Crit. Rev. Biochem. Mol. Biol.* 45 (6), 488–505. doi:10.3109/10409238.2010.512001
- Yu, Y., Teng, Y., Liu, H., Reed, S. H., and Waters, R. (2005). UV Irradiation Stimulates Histone Acetylation and Chromatin Remodeling at a Repressed Yeast Locus. *Proc. Natl. Acad. Sci. U.S.A.* 102 (24), 8650–8655. doi:10.1073/pnas.0501458102
- Yuan, C., Li, J., Mahajan, A., Poi, M. J., Byeon, I.-J. L., and Tsai, M.-D. (2004). Solution Structure of the Human Oncogenic Protein Gankyrin Containing Seven Ankyrin Repeats and Analysis of its Structure–Function Relationship. *Biochemistry* 43 (38), 12152–12161. doi:10.1021/bi049116o
- Zhang, D., Tang, Z., Huang, H., Zhou, G., Cui, C., Weng, Y., et al. (2019). Metabolic Regulation of Gene Expression by Histone Lactylation. *Nature* 574 (7779), 575–580. doi:10.1038/S41586-019-1678-1
- Zhang, Q., Zeng, L., Zhao, C., Ju, Y., Konuma, T., and Zhou, M.-M. (2016). Structural Insights into Histone Crotonyl-Lysine Recognition by the AF9 YEATS Domain. *Structure* 24, 1606–1612. doi:10.1016/j.str.2016.05.023
- Zhao, D., Guan, H., Zhao, S., Mi, W., Wen, H., Li, Y., et al. (2016). YEATS2 Is a Selective Histone Crotonylation Reader. *Cell. Res.* 26, 629–632. doi:10.1038/cr.2016.49
- Zhao, S., Allis, C. D., and Wang, G. G. (2021). The Language of Chromatin Modification in Human Cancers. *Nat. Rev. Cancer* 21 (7), 413–430. doi:10.1038/s41568-021-00357-x
- Zhao, Y., and Garcia, B. A. (2015). Comprehensive Catalog of Currently Documented Histone Modifications. *Cold Spring Harb. Perspect. Biol.* 7 (9), a025064. doi:10.1101/cshperspect.a025064
- ZhaoZhao, J., Chen, W., Pan, Y., Zhang, Y., Sun, H., Wang, H., et al. (2021). Structural Insights into the Recognition of Histone H3Q5 Serotonylation by WDR5. *Sci. Adv.* 7 (25), eabf4291. doi:10.1126/sciadv.abf4291
- Zhu, Z., Han, Z., Halabelian, L., Yang, X., Ding, J., Zhang, N., et al. (2021). Identification of Lysine Isobutyrylation as a New Histone Modification Mark. *Nucleic Acids Res.* 49 (1), 177–189. doi:10.1093/NAR/GKAA1176

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

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

Copyright © 2022 Sehrawat, Shobhawat and Kumar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.