HHS Public Access

Author manuscript

Chem Rev. Author manuscript; available in PMC 2019 July 31.

Published in final edited form as:

Chem Rev. 2018 July 25; 118(14): 6656–6705. doi:10.1021/acs.chemrev.8b00008.

Chemical and Biochemical Perspectives of Protein Lysine Methylation

Minkui Luo

Chemical Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA; Program of Pharmacology, Weill Graduate School of Medical Science, Cornell University, New York, NY 10021, USA

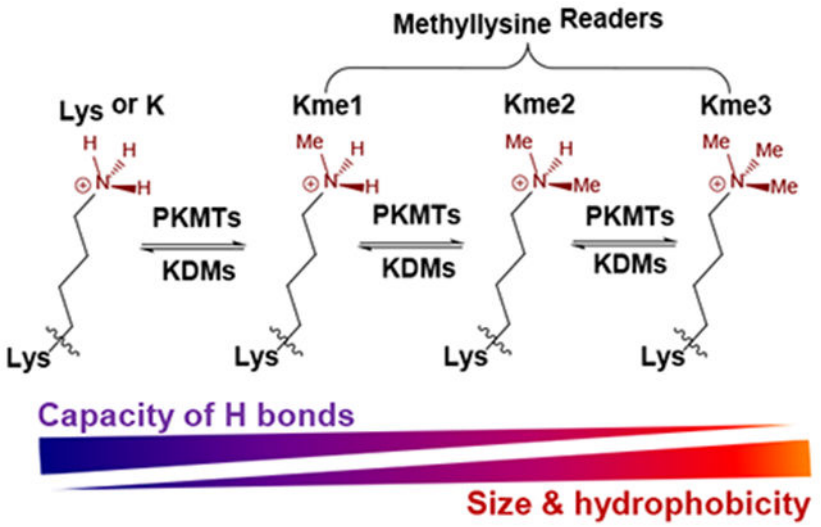
Abstract

Protein lysine methylation is a distinct posttranslational modification that causes minimal changes in the size and electrostatic status of lysine residues. Lysine methylation plays essential roles in regulating fates and functions of target proteins in an epigenetic manner. As a result, substrates and degrees (free versus mono/di/tri) of protein lysine methylation are orchestrated within cells by balanced activities of protein lysine methyltransferases (PKMTs) and demethylases (KDMs). Their dysregulation is often associated with neurological disorders, developmental abnormalities or cancer. Methyllysine-containing proteins can be recognized by downstream effector proteins, which contain methyllysine reader domains, to relay their biological functions. While numerous efforts have been made to annotate biological roles of protein lysine methylation, limited work has been done to uncover mechanisms associated with this modification at a molecular or atomic level. Given distinct biophysical and biochemical properties of methyllysine, this review will focus on chemical and biochemical aspects in addition, recognition, and removal of this posttranslational mark. Chemical and biophysical methods to profile PKMT substrates will be discussed along with classification of PKMT inhibitors for accurate perturbation of methyltransferase activities. Semisynthesis of methyllysine-containing proteins will also be covered given the critical need for these reagents to unambiguously define functional roles of protein lysine methylation.

Graphical Abstract

Corresponding Author: luom@mskcc.org; 646-888-3066.

The author declares no competing financial interest.



1. Introduction

The transfer of genetic information has traditionally been described as a forward flow from DNA to RNA to proteins. However, this classical definition does not cover biological complexity, in particular, how downstream products such as proteins and metabolites act on DNA, RNA and proteins in a heritable manner.(1, 2) Methylation of biological products such as DNA and proteins is arguably the most important biochemical reaction that challenges the central dogma of the unidirectional flow of genetic information.(3–6) For instance, cytosine in DNA can be methylated by DNMTs with *S*-adenosyl-L-methionine (SAM) as a cofactor (or co-substrate) (Figure 1).(1, 7) This event can determine the biological outcomes associated with the methylated DNA without altering its genetic code. SAM-dependent methylation in proteins can be found at side chains of lysine (Lys or K), arginine (Arg or R), aspartate (Asp or D), glutamate (Glu or E), histidine (His or H), asparagine (Asn or N), glutamine (Gln or Q) and cysteine (Cys or C), as well as at N-terminal α-amino and C-terminal carboxylate residues.(8–14) In terms of biological functions, protein methylation has been shown to affect cellular fates of proteins by modulating their stability, localization and interaction with their binding partners.(15) Well-characterized histone methylation marks have attracted great attention for two decades because of their tight association with epigenetic modulation of transcription.(16) SAM-dependent RNA methylation has also been observed in numerous types of RNA entities, including mRNAs, tRNAs, and non-coding RNAs with their associated functional roles revealed gradually.(17)

Most biological methylation reactions are catalyzed by methyltransferases with SAM as the methyl donor (Figure 1).(18, 19) Second only to adenosine triphosphate (ATP), SAM is one of the most commonly used enzyme cofactors.(18, 20) Across living organisms, SAM is produced by SAM synthases (also named as *S*-methionine adenosyltransferases, MATs) with L-methionine and ATP as substrates (Figure 1).(19, 21) Occasionally, SAM can be produced from 5′-halogen-5′-deoxyadenosine and L-methionine through the reverse reactions of SAM-consuming enzymes such as SalL and FDAS.(22) SAM's rich biochemical reactivity

is largely embedded around its sulfonium center (Figure 1). For instance, the two alkyl sulfonium bonds in SAM can be subjected to enzymatic homolytic cleavage to generate a 5′-deoxyladenosyl radical or a 3-amino-3-carboxylpropyl radical.(23, 24) The sulfonium bond in SAM's homocystine moiety can also undergo intramolecular heterolytic cleavage to generate homoserine lactone, which is a key precursor in biosynthesis of acyl-homoserine lactones (AHLs) for bacterial quorum sensing.(25) Alternatively, this sulfonium bond can undergo intermolecular heterolytic cleavage to present the 1-aminopropane moiety in SAM as a key building block for polyamine biosynthesis (Figure 1).(19, 26) More often, SAM is used by methyltransferases as the methyl donor of diverse biological substrates including DNA, RNA, proteins and small-molecule metabolites.(18–21)

SAM-dependent methyltransferases are classified in terms of their substrates. Protein lysine methyltransferases (PKMTs) are defined by their ability to transfer up to three methyl groups from the cofactor SAM to the ε-amine of a lysine side chain of their protein substrates.(27) Among around 100 putative PKMT candidates encoded by the human genome, more than 60% of them have been characterized with lysine methyltransferase activities on diverse histone and nonhistone substrates.(28, 29) The proteins containing this posttranslational mark can then be recognized by downstream effectors through their "reader" domains.(30–32) Lysine methylation can also be removed through an oxidative demethylation reaction conducted by >30 protein lysine demethylases (KDMs) encoded by the human genome.(33–35) The resulting dynamic methylation states control functional roles of protein lysine methylation and are often dysregulated in disease states.(33, 35) Enormous efforts have been made to understand physiological and pathogenic functions of PKMTs.(36–38) For instance, histone methylation marks such as the trimethylation of histone H3 lysine 4, 9 and 27 (H3K4me3, H3K9me3 and H3K27me3) have been tightly linked with transcriptional regulation of neighboring genes.(6, 39) Additionally, the roles of nonhistone methylation have gained more attention because the biology of histone methylation is insufficient to rationalize diverse functions of PKMTs.(12) Many small-molecule inhibitors have also been developed against specific PKMTs with some in various states of advancement as drug candidates.(33, 38, 40)

In contrast to the extensive work in characterizing biological functions of PKMTs and developing PKMT-specific inhibitors,(12, 30, 31, 33, 38, 40) limited effort has been made to understand protein lysine methylation from chemical and biochemical perspectives. For instance, while the challenge in developing specific antibodies to recognize the proteins containing methyllysine is broadly appreciated,(41, 42) little work has been conducted to examine molecular origins of this challenge and formulate innovative strategies to overcome it. Given diverse substrate profiles of PKMTs and lack of well-defined sequence specificities of these substrates,(21, 27, 37) it is of great interest to explore chemical origins of the substrate diversity and examine how proteins have been tuned to add, recognize and remove methyllysine marks in a highly context-dependent manner. With regard to small-molecule inhibitors, there has been more focus on developing high-quality PKMT inhibitors than utilizing these compounds in correct contexts.(33, 38, 40) Additionally, new inhibitors and chemical tools for PKMTs are still in great demand to perturb and dissect complicated biology of protein lysine methylation.(27) One goal of this review is to guide readers in appreciating elegant biophysical and biochemical properties of protein lysine methylation

and applying them for functional annotation and perturbation. With the main focus on the underlying chemical mechanisms, this review starts with a basic introduction to specific biophysical properties of lysine and methyllysine, then dives into chemical mechanisms of addition, recognition, and removal of this modification, and ends with an overview of techniques and reagents for perturbation and functional annotation of methyllysine. While this review intends to cover the relevant literature to a maximal degree, the author apologizes for the omission of many high-quality studies due to space limitation.

2. General Properties of Lysine and Methyllysine as Protein Building Blocks

2.1. L-Lysine.

L-Lysine (Lys or K) is an essential α-amino acid building block of proteins with distinct biochemical and biophysical properties. As one of the three basic natural amino acids (Lys, Arg and His), Lys contains a primary amine (ε-amine) appended to its α-carbon via a linear four-carbon linker (Figure 2). The lone-pair electrons of the ε-amine are protonated at physiological pH of 7.4 to a degree of > 99.9%, as estimated by 3 orders of magnitude higher pK_a of the ε-amine (~10.5 versus 7.4).(43, 44) The protonated ε-amine of Lys thus carries a formal charge of +1.(44) For Lys residues in proteins, this polar and positively-charged ε-amine can be positioned either within catalytic sites or more often on protein surface with exposure to bulk solvent. The C4 hydrocarbon chain of Lys, unlike its ε-amine, maintains hydrophobic properties. This amphiphilic character thus enables Lys to engage in molecular recognition through diverse modes such as cation-π interaction (a columbic attraction between a protonated Lys cation and π-electron-rich surface of an aromatic amino acid residue), an ion-pairing salt bridge, hydrogen bonds (the ε-amine as a donor and an acceptor), and the hydrophobic interaction via entropy-driven desolvation of its C4 hydrocarbon. In aqueous media, the cation-π interaction of Lys is energetically preferred in comparison with the formation of a salt bridge with a carboxylate moiety.(44, 45) Such preference is attributed to higher free-energy penalty of desolvation upon the formation of a salt bridge, which involves a negatively charge carboxylate ion such as Glu and Asp. In contrast, the cation–π complex between a protonated Lys and an aromatic residue gains comparable electrostatic attraction but pays less energy penalty of desolvation for a hydrophobic aromatic residue.

To illustrate Lys-participating interactions, the ε-amine moiety of Lys (its protonated state) is often shown as a positively charged nitrogen surrounded by a methylene and three neutral hydrogen atoms.(27) However, it has been less appreciated that the +1 charge on the ε-nitrogen is merely a formal charge, which should not be viewed as the positive electrostatic density localized on an electronegative nitrogen atom (Figure 2).(44, 46) Instead, the nitrogen still carries significant negative charge with the net +1 charge delocalized around nearby atoms (three hydrogen atoms and the ε-amine-adjacent methylene group) (Figure 2). Whereas it is convenient to depict the +1 formal charge of a protonated amine moiety for its formation of hydrogen bonds and a salt bridge, it is more relevant to envision the actual electronic density on the ε-nitrogen and its neighboring atoms for cation-π interactions (see discussion later). Additionally, cation-π interactions are strongest when a cation is situated

perpendicular to the aromatic plane of a Phe, Tyr or Trp residue.(46) This spatial orientation is preferred to engage in the maximal cation-π interaction between an ammonium cation and the side chain of an aromatic amino acid.(46)

2.2. Epsilon-*N*-monomethyl lysine (Kme1).

The ε-amine moiety of Lys can be methylated up to three times from unmodified lysine to mono-, di- and tri-methylated Lys (Kme1, Kme2 and Kme3) (Figure 2). The progression of Lys methylation gradually alters biophysical properties of Lys, which determine how these Lys modifications (Kme1/2/3) engage in specific interactions and are selectively recognized by their biological effectors.(30, 31) The pK_a value of the secondary amine of Kme1 is ~ 10.7 (estimated on the basis of pK_a of dimethylamine), which is comparable with the pK_a 10.5 of free Lys (Figure 2). At physiological pH of 7.4, the secondary amine of Kme1 is protonated to a degree of > 99.9% with a formal charge of +1 on the electronegative ε-nitrogen (Figure 2). As a more relevant view, the +1 charge is spread around adjacent hydrogen and hydrocarbon moieties (two hydrogen atoms, one methyl and one methylene) (Figure 2). Despite dispersion of the +1 charge into two neighboring hydrocarbons, positively charged Kme1 maintains the ability to form a salt bridge with negatively charged amino acids such as Glu and Asp (Figure 2). In contrast to a protonated Lys as the donor and acceptor of three hydrogen bonds, Kme1 can only serve as the donor and acceptor of two hydrogen bonds (Figure 2). Addition of one methyl group onto free Lys also increases the overall size and hydrophobicity of its side chain (Figure 2). Such a difference can play an important role in distinguishing Lys monomethylation from other states of Lys methylation by structurally matched biological effectors as detailed later.(30, 31)

2.3. Epsilon-*N*-dimethyl lysine (Kme2).

Similar to Lys and Kme1, Kme2 contains two lone-pair electrons on its ε-amine. The pK_a value of the tertiary ε-amine of Kme2 is 10.2 (estimated on the basis of *N*,*N*-dimethylethylamine), which is comparable to those of the ε-amines of Lys and Kme1 (Figure 2). Given its pK_a of 10.2, the ε-nitrogen of Kme2 is largely protonated at physiological pH (>99%). In this protonation state, Kme2 can serve as the donor or acceptor of one hydrogen bond, and form a salt bridge. The positively charged Kme2 after protonation can also engage in cation-π interactions (Figure 2). In comparison with the net +1 formal charge of Lys and Kme1, the positive charge of Kme2 is more dispersed among three neighboring hydrocarbon moieties (two methyl groups and one methylene moiety) and one hydrogen atom (Figure 2). Addition of two methyl groups significantly increases the overall size and hydrophobicity of the side chain of Lys (Figure 2). The combined effects are associated with the higher spatial requirement for Kme2 to engage in cation-π and hydrogen bond interactions for its molecular recognition.(30, 31)

2.4. Epsilon-*N*-trimethyl lysine (Kme3).

Kme3 is the extreme state of lysine methylation (Figure 2). The side chain of Kme3 is an obligatory cation with a permanent +1 formal charge at physiological pH (Figure 2). Like Lys, Kme1 and Kme2, the positively charged Kme3 can form a salt bridge with negatively charged amino acid residues such as Asp and Glu (Figure 2). Unlike free Lys, Kme1 and Kme2, Kme3 cannot form a hydrogen bond either as a donor or acceptor because of the

quaternary ammonium moiety (Figure 2). Here the permanent +1 formal charge of Kme3 is broadly dispersed on the four neighboring hydrocarbon moieties (three methyl groups and one methylene moiety) (Figure 2). Addition of three methyl groups substantially enhances the overall size and hydrophobicity of the side chain of Lys (Figure 2). Compared with free Lys, Kme1 and Kme2, Kme3 is characterized with the largest van der Waals radius and thus the highest spatial demand upon engaging in cation-π interactions. Collectively, it is more accurate to depict the quaternary ε-ammonium of Kme3 as a hydrophobic cloud with an electronegative nitrogen atom enshrouded centrally and the overall +1 charge evenly dispersed around the four surrounding hydrocarbons (Figure 2). This description will become more apparent upon elucidating the underlying mechanism for recognition of Kme3 in a biological setting.

3. Classification of Protein Lysine Methyltransferases (PKMTs)

3.1. Overview of protein methylation.

Polar lysine residues are often positioned at solvent-exposed protein surface regions. The lone-pair electrons of the ε-amine of Lys, together with its preferred localization on protein surface, make lysine susceptible to diverse posttranslational modifications (Figure 3). Among characterized lysine modifications in eukaryotic cells are methylation, acetylation, propionylation, malonylation, butyrylation, succinylation, glutarylation, myristoylation, biotinylation, ubiquitination, SUMOylation, and neddylation (Figure 3).(27, 47–52) There are at least two factors that make protein lysine methylation distinct from other lysine posttranslational modifications (Figure 3): (a) each addition of a methyl group onto Lys does not alter the +1 formal charge of the ε-amine moiety at physiological pH, whereas other Lys modifications are acylation and convert the protonated, positively charged ε-amine into a neutral amide; (b) lysine methylation is the smallest posttranslational modification and thus minimally alters overall size of the side chain in comparison with other lysine modifications. (21)

Lys, Kme1 and Kme2 contain lone-pair electrons on their ε-amine moieties and are expected to be readily alkylated (*e.g.* methylated) under basic conditions. However, high pK_a of Lys, Kme1 and Kme2 (10.2 ~ 10.7) makes their ε-amines mainly protonated under physiological conditions (pH = 7.4) and thus inert as nucleophiles. To circumvent the activation barrier of deprotonation for the subsequent nucleophilic substitution reaction, protein lysine methylation is catalyzed by various PKMT enzymes in biological settings (Figure 4).(53–57) The human genome encodes more than 60 characterized PKMTs, which can be classified as SET domain-containing PKMTs (Class V methyltransferases, SET for *S*uppressor of variegation 3-9, *E*nhancer of zeste and *T*rithorax, three genetic phenotypes of Drosophila) and non-SET-domain PKMTs (Class I methyltransferases) (Figure 4, Figure 5 and Table 1). (27–29, 40) More than 90% of PKMTs belong to the family of SET domain-containing PKMTs (Table 1 and Figure 5). DOT1L,(54) METTL10,(58) METTL20,(59, 60) METTL21A,(61) METTL21B,(62, 63) METTL21C,(64) METTL21D (VCP-KMT),(65, 66) METTL22,(67) eEF1A-KMT1 (N6AMT2 as named previously, the homolog of yeast Efm5),(68) eEF2-KMT,(69) and CaM-KMT(70) are among well-characterized members of non-SET domain human PKMTs (Figure 5 and Table 1). Although > 60 human proteins

have been characterized as PKMTs, they are often multiple-domain macromolecules consisting of hundreds to thousands of amino acids embedding small methyltransferase domains (130-aa SET or non-SET domain).(55, 71, 72) Their methyltransferase activities may only account for part of their biological functions. Caution should thus be taken to dissect the functions of the methyltransferase domains of PKMTs from their other roles as full-length proteins (*e.g.* as effectors for recognition of other posttranslational modifications or as structural scaffolds for formation of protein complexes).

3.2. SET domain-containing PKMTs.

The SET domain of PKMTs (Class V methyltransferases) consists of approximately 130 amino acids, often flanked by pre-SET and post-SET domains (Figure 4).(71) Its structural topology is distinct from other types of methyltransferases by a characteristic 'pseudoknot' fold of the SET domain (Figure 4).(71) On the basis of phylogenetic sequences of SET domains, human PKMTs can be further divided into subfamilies with respective structural topology (Table 1 and Figure 5).(73) For instance, G9a and GLP1 belong to a subfamily of classical PKMTs in which their SET domains alone are sufficient for enzyme catalysis (Table 1 and Figure 5).(74) In contrast, ASH1L, SETD2 and NSD1/2/3 are within a subfamily of PKMTs containing an auto-inhibitory SET domain, whose apo-isomer is expected to be catalytically inactive and must go through dramatic conformational changes for substrate binding and enzyme catalysis (Table 1 and Figure 5).(55, 75–78) The SET domains of the MLL subfamily alone are inert but catalytically active in the presence of their binding partners such as WDR5, RbBP5, Ash2L and Dpy30 (referred as WRAD) (Table 1, Figure 5 and Figure 6).(79–81) A similar requirement for the formation of a protein complex for enzyme catalysis has also been documented for EZH1 and EZH2 (EZH1/2, EED and Suz12 referred as PRC2 complexes) (Table 1, Figure 5 and Figure 6).(82, 83) A 5-member SMYD subfamily and a 17-member PRDM subfamily of human PKMTs are characterized by insertion of the MYND domain (myeloid translocation protein 8, Nervy, and DEAF-1) (84–88) and the PR (PRDI-BF1 and RIZ) domain within their SET domains,(89) respectively (Table 1 and Figure 5).

3.3. Non-SET domain PKMTs.

In comparison with SET domain-containing PKMTs harboring well-defined methyltransferase activities, DOT1L,(54) METTL10,(58) METTL20 (ETFβ-KMT),(59, 60) METTL21A,(61) METTL21B (eEF1A-KMT3),(62, 63) METTL21C,(64) METTL21D (VCP-KMT),(65, 66) METTL22,(67) eEF1A-KMT1 (N6AMT2 as named previously, the homolog of yeast Efm5),(68) eEF2-KMT,(69) and CaM-KMT(70) are among human non-SET domain PKMTs demonstrated with methyltransferase activities (Figure 5). These PKMTs are structurally homologous to protein arginine methyltransferases (PRMTs) (Figure 5).(28) Non-SET-domain PKMTs and PRMTs belong to the canonical Rossmann-fold-like family of methyltransferases (Class I) with their structural topology featuring a seven-stranded β-sheet connected by α-helices. Besides protein Lys and Arg, small-molecule, DNA and RNA are among the substrates of the class-I methyltransferases (Figures 4,5). The human genome encodes 200 putative methyltransferases, most of which have yet been characterized.(28) It will be intriguing to explore whether some of these putative methyltransferases are PKMTs. Recently, after examining a catalytically-dead SET-domain

mutant of MLL1, the Cosgrove group reported that the MLL1-WRAD (WDR5, RbBP5, Ash2L and Dpy30) complex contains a noncanonical site located at its Ash2L subunit to catalyze H3K4 methylation.(95–97) Because Ash2L has no homology with SET domain-containing Class-V methyltransferases and non-SET-domain Class-I methyltransferases, it is of great interest to explore the structural features of the cryptic PKMT site, which can be informative to identify novel PKMTs.

4. PKMT-catalyzed Methylation: Cofactor, Substrates and Enzyme Catalysis

4.1. Overview of PKMT-catalyzed methylation.

PKMTs are classified as transferase enzymes (EC 2 in enzyme nomenclature), which catalyze the transfer of specific moieties (*e.g.,* a methyl group, EC 2.1.1 subclass) from one molecule (co-substrate or cofactor such as *S*-adenosyl-L-methionine or SAM) to a substrate (*e.g.,* a lysine or arginine residue for protein methylation). PKMT-catalyzed methylation reactions are expected to go through an approximately linear S_N2 transition state with the ε-nitrogen of Lys, Kme1 and Kme2 as a nucleophile and SAM's sulfonium as a leaving group (Figure 7).(98) In order to assemble S_N2 transition states for effective catalysis, PKMTs, in particular those with promiscuous substrate profiles (see discussion below), need to bind various substrates with their target Lys(Kme1/2) residues positioned within their catalytic pockets; align the sulfonium methyl group of the SAM cofactor on a nearly linear path required for a S_N2 reaction; deprotonate the ε-nitrogen of Lys(Kme1/2) to free its lone pair of electrons (Figure 7).(98, 99) Upon poising the resulting activated lysine substrate and the SAM cofactor along a nearly linear reaction trajectory, PKMTs may engage in other dynamic motions and interactions to further stabilize the S_N2 transition state and facilitate the subsequent chemical step of methylation (Figure 7).(98, 99) Upon the completion of the transfer reaction, PKMTs release the methylated product and a byproduct *S*-adenosylhomocysteine (SAH) for another catalytic cycle.(19, 100) As detailed below, the structures of many PKMTs have been elegantly tuned to promote the stepwise catalysis.

4.2. SAM-binding modes of PKMTs.

3-D structures of PKMTs reveal that these enzymes bind the SAM cofactor in two distinct modes (Figure 4).(101, 102) The SAM cofactor in SET domain-containing PKMTs is oriented in a relatively packed configuration with the Cβ-Cγ-Sδ-C5′ dihedral angle around 120° (*e.g.,* 114° for SET7/9 and 131° for SETD8).(103–109) In contrast, the SAM cofactor in non-SET-domain PKMTs, as exemplified with DOT1L,(54) adopts an extended conformation with the Cβ-Cγ-Sδ-C5′ dihedral angle around 180° (Figure 8). Upon further analyzing the modes of interaction (MOI) of the SAM cofactor in SET domain-containing PKMTs and the non-SET-domain DOT1L, the Schapira group generalized that PKMTs interact with the α-amino acid moiety of SAM through salt bridges, hydrogen bonds or their combination; the adenine moiety of SAM through combined hydrogen bond and hydrophobic interactions (Figure 8).(102) However, MOIs of SET domain-containing PKMTs and the non-SET-domain DOT1L are different at least in two aspects.(102) First, the 2′, 3′-hydroxyl moiety of SAM, as well as its mimics SAH and sinefungin, in SET domain-

containing PKMTs are solvent exposed and does not involve any essential interaction with neighboring residues, while the comparable 2′,3′-hydroxyl moiety in DOT1L is deeply embedded in a cofactor-binding pocket and forms two strong hydrogen bonds with the carboxylic side chain of the nearby E186 residue (Figure 8).(54, 110) Second, SAM's sulfonium methyl group in SET domain-containing PKMTs resides in a highly compacted, electronegative environment, whereas the comparable sulfonium methyl moiety in DOT1L is localized in a relatively vacant space (Figure 8). (54, 71, 110) These SAM-interacting networks are expected to be essential for PKMTs to position the SAM cofactor in a ready orientation for enzyme catalysis.

A combination of experimental and computational methods is often required to elucidate molecular details of how PKMTs interact with the SAM cofactor.(98, 111, 112) In the case of SET7/9, the characteristic ^{1}H-NMR chemical shifts, quantum mechanical (QM) calculation, crystallographic data and kinetics of SET7/9 mutants implicated that the sulfonium methyl moiety of the cofactor SAM engages in noncanonical carbon-oxygen (CH•••O) hydrogen-bonding interactions with the hydroxyl group of Y335 as well as the two amide oxygen atoms of H293 and G264 (Figure 9).(113–116) These interactions were shown to be stronger than typical CH•••O hydrogen-bonding interactions, likely because the positively charged sulfonium acidifies the protons on the adjacent carbon atoms (Figure 9). For instance, the ^{1}H chemical shift of the methyl group of SET7/9-bound SAM is 3.8 ppm, which is 0.8 ppm downfield relative to that of SAM in solution and thus reflects a more electron-withdrawing environment surrounding SAM's methyl group in SET7/9 (Figure 9). (113) In comparison with Y335 in native SET7/9, the Y335F substitution, which does not significantly alter the overall structure and stability of SET7/9, diminishes SAM's affinity by 3 orders of magnitude.(113) Consistent with a less electron-withdrawing environment, the Y335F substitution shifts the ^{1}H-NMR signal of SAM's methyl group to upfield.(113) These observations can be partially rationalized by the lost CH•••O interaction of the methyl group with SET7/9's Y335 residue.

Binding modes of PKMTs for the SAM cofactor in PKMTs can also be probed experimentally with binding isotope effects (BIEs).(98, 111, 112) In this approach, the changes of bond orders and vibrational modes of the SAM cofactor upon binding PKMTs can be reported by relevant BIEs when the corresponding atoms of SAM are replaced with heavy isotopes (Figure 10).(98, 112) Upon measuring the BIE of [CT_3]-SAM in complex with NSD2, the Schramm laboratory observed an inverse BIE of 0.65.(112) This large inverse BIE strongly argues that the SAM cofactor, from a solution state to the NSD2-bound state, experiences extreme steric impingement and thus restricts the stretching vibrational modes of the BIE-associated C-H bonds.(112) Interestingly, upon the formation of the pseudo ternary Michaelis-Menten complex of NSD2 with the SAM cofactor and H3K36M nucleosome (H3K36MNuc, an inactive substrate mimic), the steric impingement is relaxed and the BIE of [CT_3]-SAM was reported to be 0.990.(111, 112) While this [CT_3]-BIE is significantly less inverse than that of the binary SAM-NSD2 complex (0.990 of the ternary SAM-NSD2-H3K36MNuc complex versus 0.65 of the binary SAM-NSD2 complex), this small inverse BIE of the ternary complex is comparable with [CD_3] and [CT_3] BIEs of SAM in complex with SETD8 (0.990 versus 0.959 and 0.979).(98, 111, 112) As further supported by computational modeling, these slightly inverse BIEs are consistent with the noncanonical

carbon-oxygen (CH•••O) hydrogen-binding interactions of the methyl group of SAM with the PKMTs through a conversed phenolic hydroxyl group and amide oxygen atoms as implicated in several PKMTs (Y1179, R1135 and F1117 of NSD2; Y336, R295 and C270 of SETD8; Y335, H293 and G264 of SET7/9) (Figures 9, 10).(98, 111, 112) In comparison, the binary SAM-NSD2 complex and the ternary SAM-NSD2-H3K36MNuc complex show comparable [5′-$^{3}H_{2}$]-SAM BIEs of 0.973 and 0.982, respectively.(111, 112) These different BIEs argue that SAM binding is different for the two PKMTs (Figures 9, 10).

Upon probing molecular rationales for PKMTs to recognize the SAM cofactor, the Trievel laboratory also examined potential noncovalent interactions between SAM's sulfonium and its adjacent oxygen atoms (S•••O chalcogen bonds) around the active site of SET7/9.(117) The sulfur-oxygen (S•••O) chalcogen interaction originates from partial bond order through a σ-antibonding orbital of SAM's sulfur atom and a lone pair of electrons of the backbone amide oxygen of the nearby N265 residue (Figure 11). In comparison with N265 in native SET7/9, the N265A mutation does not alter the overall structure and stability of SET7/9 but decreases its affinity to SAM by 8-fold.(117) In contrast, the N265A variant diminishes the affinity to SAH, a SAM mimic with SAM's sulfonium replaced by the cognate neutral thioether, only by 2-fold. Consistent with this observation, quantum mechanical (QM) calculation suggests that sulfonium cations form a stronger S•••O chalcogen bond than cognate neutral thioethers. The noncanonical sulfur-oxygen (S•••O) chalcogen interaction is thus essential for SET7/9 to engage in the maximal binding to the SAM cofactor.(117) It remains to be investigated experimentally whether unconventional carbon-oxygen (CH•••O) and sulfur-oxygen (S•••O) chalcogen interactions play roles in recognizing the SAM cofactor by other SET domain-containing PKMTs in a broad manner.

4.3. Histone and nonhistone targets of PKMTs.

SET domain-containing PKMTs were first characterized to methylate histones and originally named as histone methyltransferases.(27) However, recent advancement in identifying PKMT substrates indicated that SET domain-containing PKMTs can act on diverse nonhistone substrates (*e.g.,* full-length G9a and GLP1 can methylate hundreds to thousands of nonhistone substrates, see discussion later).(21) Using the nomenclature of PKMTs rather than histone methyltransferases is more consistent with diverse substrate profiles of these enzymes. In contrast, only a small set of proteins were identified as the substrates of non-SET domain PKMTs as exemplified by DOT1L for K79 of histone H3,(54, 118) METTL10 for K318 of eEF1A,(58) METTL20 for K199 and K202 of the β-subunit of the electron transfer flavoprotein (ETFβ),(59, 60) METTL21A for HSP70s (K561 in HSPA1, K585 in HSPA5, and K561 in HSPA8),(61) METTL21B for K165 of eukaryotic elongation factor 1 alpha (eEF1A),(62, 63) METTL21C for K390 of eEF1A as well as K35 automethylation, (64, 119) METTL21D (VCP-KMT) for K315 of valosin-containing protein (VCP),(65) METTL22 for K135 of Kin17,(67) eEF1A-KMT1 for K79 of eEF1A,(68) eEF2-KMT for K525 of eukaryotic elongation factor 2 (eEF2),(69) and CaM-KMT for K115 of calmodulin. (70) Given that many known Lys methylation sites have not been assigned to specific PKMT activities, it remains to be determined whether non-SET domain PKMTs, like SET domain-containing PKMTs, can act on diverse substrates.

4.4. Sequence motifs and binding modules of PKMT substrates.

While some PKMTs recognize substrates promiscuously and others are more specific, there are no general rules to define unambiguously PKMT substrates through their primary sequences. For instance, SET7/9 was shown to methylate histone H3, p53 and TAF10 by recognizing a common peptide motif [R/K][S/T]K (K as the modification site).(106, 120) After further structural and biochemical characterization of the enzyme-substrate complex, the consensus sequence of potential SET7/9 substrates was extended to [R/K][S/T/A]K[DNQK].(121) Upon searching SET7/9 targets with the sequence lead, several new substrates including TAF7 and E2F1 were then identified.(106, 122) Nevertheless, the consensus sequence of SET7/9 substrates was then expanded more systematically with arrayed peptides as substrate candidates and led to identification of 91 potential substrates.(123) However, no consensus sequence was defined for many newly-identified SET7/9 substrates including STAT3,(124) androgen receptor,(125, 126) Yap,(127) SIRT1,(128) FOXO3,(129, 130) DNMT1,(121) Rb(131) and NF-kappa B.(132) A similar situation was also encountered upon identifying substrates of G9a and GLP1, which were first characterized to methylate histone H3 and DNMT3A through a consensus sequence TARK (K as the modification site).(133–135) The structural characterization of GLP1-substrate complexes revealed that Asp1135 and Asp1145 of GLP1 directly interact with R at the −1 position and T at the −3 position through salt bridges and hydrogen bonds, respectively.(135) However, this consensus sequence cannot be found in newly-identified G9a/GLP1 substrates such as p53,(136) C/EBP-β,(137) Reptin,(138) and MyoD.(139) More importantly, recently advanced technologies allowed revealing hundreds of substrates of G9a and GLP1,(99, 140–144) which do not share well-defined consensus sequences. Similarly, no consensus sequence can be defined upon examining SETD8 substrates histone H4 (RHRK),(107) PCNA (GHIK)(145) and Numb (LERK);(146) EZH2 substrates histone H3 (AARKS),(147) RORα (SARKS)(148) and STAT3 (KTLKS).(124, 149, 150) Notably, citrullination of H3R26 in H3 (AARKS) suppresses the EZH2-mediated methylation of H3K27 (AARKS) by 30,000 fold. This result suggests that the presence of a citrulline residue rather than the absence of an arginine residue at the “-1” position diminishes the reactivity of H3K27 as EZH2’s substrates.(151)

Accumulated evidence strongly argues that many SET domain-containing PKMTs such as SET7/9,(152) G9a,(153) GLP1,(99, 140–143) SETD8(146, 154) and EZH2(124, 149, 150) can bind multiple substrates in a promiscuous manner. A possible working model is that their substrate-binding pockets are structurally flexible and thus can adopt multiple configurations to accommodate different substrates. An alternative but nonexclusive explanation for substrate promiscuity is that a binding pocket recognizes many substrates mainly through their backbones rather than side chains. In addition, structurally flexible substrates may also adopt multiple conformations for optimal binding by PKMTs. Computational modeling showed that SMYD3 can engage in substrate interactions through multiple conformations of substrate peptides.(155, 156) On the basis of this model, a few proteins were identified as novel substrates of SMYD3.(155, 156) While these hypotheses could be readily tested by examining 3-D structures of the corresponding PKMT-substrate complexes, the so-far solved structures of PKMT-substrate complexes do not cover a broad scope of PKMT substrates. As indirect evidence, CARM1, an arginine protein

methyltransferase with a structural topology similar to that of DOT1L, recognizes its substrates mainly through amide backbones and with minimal interactions with the side chains of these substrates.(157) In contrast to SET domain-containing PKMTs, no 3-D structure of an enzyme-substrate complex has been solved for non-SET domain PKMTs. Many of non-SET domain PKMTs (*e.g.,* DOT1L)(54, 118, 158, 159) recognize full-length proteins rather than truncated peptides as active substrates.

4.5. Lysine binding pockets of PKMTs.

Although it is difficult to predict consensus sequences and binding modes for PKMTs to recognize their substrates, SET domain-containing PKMTs do follow some general rules to position the side chains of their substrates (Lys, Kme1 and Kme2) into catalytic sites. All SET domain-containing PKMTs bind the side chains of Lys, Kme1 and Kme2 through a hydrophobic channel mainly consisting of aromatic amino acids (Tyr, Phe and occasionally Trp).(71) These structural motifs are expected to achieve maximal engagement with lysine, Kme1 and Kme2 through both desolvation of their C4 hydrocarbon (hydrophobic effect) and cation-π interactions through their positively charged ε-amine moieties. The hydroxyl residue of a highly-conserved Tyr (*e.g.*, Y1154 in G9a, see Figure 12), which is located at the interface of the SAM cofactor and the ε-amine substrate, also forms a strong hydrogen bond with the ε-amine nitrogen. Such an interaction locks this ε-amine moiety and positions its lone-pair electrons in a ready configuration for subsequent methylation. Lysine-binding pockets of SET domain-containing PKMTs are thus well tuned to engage substrates for catalysis. It has also been noted that certain PKMTs such as SUV420H1/2(160) and EZH2 Y641F/N/S/H/C mutants(161–165) prefer pre-methylated lysine residues as substrates. Given that no similar aromatic-rich pocket can be mapped for Class-I PKMTs such as DOT1L, METTL10, METTL20, METTL21A, METTL21B, METTL21C, METTL21D (VCP-KMT), METTL22, eEF1A-KMT1, eEF2-KMT, and CaM-KMT, these PKMTs likely adopt different strategies to engage their substrates for catalysis.

4.6. Lysine deprotonation in PKMT-catalyzed methylation reaction.

Besides efficient engagement of substrates, catalytic sites of SET domain-containing PKMTs are also elegantly structured to remove one proton from Lys, Kme1 and Kme2 substrates and thus activate them for subsequent methylation. With SET7/9 as an example, a rudimentary mechanism of substrate deprotonation was speculated to involve SET7/9's Y335 as a general base.(166) However, this mechanism was soon ruled out because of the comparable pK_a between the phenolic hydroxyl group of Tyr (pKa = 10.5) and the ε-amines of Lys, Kme1 and Kme2 (pK_a = 10.2 ~ 10.7) (Figure 2). Cation-π interactions between the ε-amine of Lys (free or Kme1/2) and nearby aromatic residues in lysine-binding pockets further increases the p*K*a of Lys, Kme1 and Kme2 to favor their protonated states. The phenolic hydroxyl group with p*K*a around 10 is thus not sufficient to remove the proton from the ε-amine group at physiological pH of 7.4. Despite the lack of a general base at the catalytic site of SET7/9 (other SET-domain PKMTs as well), computational modeling revealed that deprotonation of the Lys ε-amine can occur through transient formation of dynamic water channels around SET7/9's catalytic site.(167) Molecular dynamics simulations and quantum mechanical/molecular mechanics (QM/MM) calculations further suggest that the SAM binding by SET7/9 at its catalytic site induces electronic repulsion

between the positive sulfonium moiety of the SAM cofactor and the protonated ε-amine of Lys, and thus further decreases the pK_a of the latter to 8.2.(167) The formation of the dynamic water channels and the proximity between SAM's sulfonium and the substrate lysine were proposed to be essential for efficient catalytic turnover of SET7/9. Computational modeling also showed that similar deprotonation activation also occurs for SETD8 to methylate its H4K20 substrate,(168) likely for other SET domain-containing PKMTs in general. It remains to be investigated whether the formation of dynamic water channels is also required for non-SET-domain PKMTs to catalyze their methylation reactions, given similar lack of a general base for lysine deprotonation at catalytic sites.

4.7. Transition state stabilization of SET domain-containing PKMTs.

Stepwise progression of protein lysine methylation is accompanied by changes of bond order and vibrational modes of the sulfonium methyl moiety of the SAM cofactor and the ε-amine moiety of the substrate Lys.(98) Multiple computational and experimental methods have been implemented to elucidate molecular features of the transition states of PKMT-catalyzed methylation reactions.(98, 111, 167–175) The structures of the transition states of PKMT-catalyzed methylation can be probed experimentally with kinetic isotope effects (KIEs) in combination with computational modeling.(98, 111) Computational modeling can afford an unbiased set of candidate structures of transition states. KIEs (*e.g.,* the methyl group of SAM and the ε-amine of the substrate Lys) of individual PKMTs are then used as electrostatic and geometric constraints to define the best matched transition state structure(s) with atomic resolution. While KIEs are conventionally determined by ratios of steady-state kinetic parameters (k_{cat} and K_m) with a pair of isotopic substrates, more precise approaches with a mixed isotopic pair in a competition format are required for PKMT-catalyzed methylation, for which KIEs are often within a range of a few percent from unity.(98, 111, 176) Among successful approaches to determine precisely KIEs of PKMT-catalyzed methylation are remote radioactive labeling and isotope ratio mass spectrometry (MS).(98, 111)

For NSD2-catalyzed H3K36 methylation, the Schramm laboratory relied on the remote radioactive labeling method to determine intrinsic [$^{14}CH_3$]-SAM, [^{36}S]-SAM, [CT_3]-SAM and [CD_3]-SAM KIEs (Figure 13).(111) A large primary [$^{14}CH_3$] KIE of 1.113 and a nearly extreme [^{36}S] KIE of 1.018 are characteristic for a S_N2 transition state with the methyl transfer reaction as a rate-limiting step.(111) [CT_3]-SAM and [CD_3]-SAM of NSD2-catalyzed H3K36 methylation show inverse KIEs of 0.77 and 0.83, respectively (Figure 13). To measure KIEs of SETD8-catalyzed H4K20 methylation, we developed a set of mathematic algorithm, in combination with matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (MS), to determine isotopic ratios of monomethylated H4K20.(98) SETD8 showed a characteristic primary intrinsic [$^{13}CH_3$]-SAM KIE of 1.04 for a S_N2 transition state and an inverse intrinsic α-secondary [CD_3]-SAM KIE of 0.90 (Figure 13).(98) Here the different KIEs between SETD8 and NSD2 argue two distinct transition states with certain commonly shared S_N2 characters.(98) Using these KIEs as experimental constraints, in combination with computational modeling, NSD2 was revealed to adopt a late, asymmetric S_N2 transition state with its N-C distance of 2.10 Å and S-C distance of 2.53 Å (Figure 13).(111) In comparison, the transition state of SETD8

was characterized with an early, asymmetric S_N2 character with its N-C distance of 2.38 Å and S-C distance of 2.05 Å (Figure 13).(98)

Besides the transition state structures of NSD2 and SETD8 solved with KIEs as experimental constraints, unconstrained computational modeling also revealed the transition state structures of several SET domain-containing PKMTs including SETD8 and SET7/9 (Figure 14).(167–175) The computationally modeled transition state structure of SETD8 is consistent with an early, asymmetric S_N2 transition state with the KIEs as experimental constraints.(168) Interestingly, the modeled transition state structures of SET7/9 are substrate-dependent with symmetric S_N2 characters for histone 3 lysine 4 peptide substrate (comparable N-C and S-C distances of 2.2~2.4 Å) and with late, asymmetric S_N2 characters for p53 lysine 372 peptide substrate (Figure 14).(167, 170) Remarkably, all of the known transition state structures of PKMTs feature a relatively fixed distance of 4.4~4.6 Å between the nitrogen nucleophile and the sulfonium leaving group but differ in the position of the transferred methyl group (Figure 14).

Computational modeling further revealed that the reaction path of SET7/9-catalyzed lysine methylation toward its transition state features less degree of electrostatic change in comparison with that of a solution-phase S_N2 methylation reaction.(171, 173) SET7/9's methylation reaction is accelerated by pre-organization of an electrostatic environment at its catalytic site.(171, 173) Several pieces of nonexclusive evidence support the preorganized electrostatic environment in SET7/9 and likely many other SET domain-containing PKMTs for enzyme catalysis. SET7/9 binds SAM in the ground state partially through leveraging three essential non-canonical CH•••O hydrogen-bonding interactions of SAM's methyl group with the two amide oxygens of H293 and G264 and the phenolic oxygen of Tyr336 (Figure 9).(113–116) On the path toward the transition state, these CH•••O hydrogen-bonding interactions are expected to be maintained and thus construct an electrostatic pore to confine the motion of the sulfonium methyl at a catalytically ready conformation. In the course of examining methylation reactions catalyzed by catechol-O-methyltransferase, glycine-N-methyltransferase and their mutants, the Klinman group obtained inverse 2° kinetic isotope effects (KIEs) of [CT_3]-SAM and found that the magnitude of these inverse KIEs positively correlates with their catalytic efficiency.(177, 178) The origin of these inverse KIEs was attributed to the compaction of active site residues toward SAM's methyl group at the transition states.(179) The compaction model proposed for catechol-O-methyltransferase and glycine-N-methyltransferase can be also applied to rationalizing how SET7/9 and other PKMTs pre-organize their electrostatic environment to align SAM's methyl group along the nearly linear trajector of a S_N2 transition state for enzyme catalysis. This rationale is consistent with the inverse α-secondary [CD_3]/[CT_3]-SAM KIEs of 0.77 and 0.83 for NSD2 and 0.90 for SETD8 (Figure 13), whose magnitude is too large to be rationalized solely by noncanonical carbon-oxygen (CH•••O) hydrogen-binding interactions of SAM's methyl group upon binding PKMTs.(98, 111) Whereas the large inverse KIEs can be readily attributed to the axial compression of the N-S distance along the linear trajectory of the transition state, computational modeling suggests that other modes such as equatorial compression can also lead to comparable KIEs.(179) It is likely that the equatorial compression through non-canonical CH•••O hydrogen-bonding interactions and the N-S axial compression act together for assembling transition states of diverse PKMTs.(98, 179)

4.8. Product specificity of PKMTs.

A lysine side chain of PKMT substrates can be subjected to methylation for up to three times to generate mono-, di- and tri-methylated products (Kme1/2/3), respectively. This product specificity is controlled by multiple factors including the topology of catalytic sites of PKMTs, amino acid sequences and prior methylation states of substrates, as well as their biological contexts such as the presence of coactivators. The molecular mechanism underlying the product specificity can be summarized by the following general principle: a methylation reaction can only occur if PKMTs can assemble a lysine substrate (Lys, Kme1 and Kme2) and the SAM cofactor into an active transition-state configuration. In the context of the reaction path discussed above (Figure 7), the whole process of catalysis demands PKMT to bind the Lys(Kme1/2) substrate and SAM in a manner so that a water channel for deprotonation can be formed at the catalytic site and thus free the lone pair of electrons of the ε-amine as a nucleophile; the activated ε-amine needs to be aligned in proximity with SAM's sulfonium methyl group and in a nearly linear trajectory for a S_N2 reaction. This general principle can then be applied to rationalize the product specificity of many PKMTs.

It has been well characterized that most SET domain-containing PKMTs can leverage a Phe/Tyr switch to control product specificity with a smaller Phe residue preferred for higher methylation states (di, tri) and a bulkier Try residue for lower methylation states (mono, di) (Figure 15).(180, 181) For instance, human G9a, which has F1152 as the Phe/Tyr switch site, can di- and tri-methylate histone H3K9, whereas its F1152Y mutant can only monomethylate this substrate (Figure 15).(182) In contrast, SET7/9 and SETD8, which have a Tyr residue (SETD8's Y335 and SET7/9's Y305) as their Phe/Tyr switch, are in general characterized as the PKMTs for monomethylaitoin (Figure 15).(181, 183) Structural characterization of SET7/9, SETD8 and their Tyr-to-Phe switch mutants indicates that native SET7/9 and SETD8 rely on the phenolic oxygen of the Try residue to immobilize a water molecule at their catalytic sites (Figure 15).(181, 183) As further revealed by the structures of the Y335F mutant of SETD8 and the Y305F mutant of SET7/9, the absence of the phenolic oxygen will liberate the water molecule and thus spare a cavity to accommodate the methyl group of methyllysine substrates. This mode of interaction thus allows positing the lone-pair electrons of methyllysine's ε-amine adjacent to SAM's methyl moiety for catalysis. (183)

It is worth noting the exception for the rule of the Phe/Tyr switch. For instance, although SET7/9 and SETD8 have been classified as the PKMTs harboring monomethylation activity because of their characteristic Tyr switch, several exceptions have been reported that the two PKMTs can also carry out dimethylation reactions for certain substrates *in vitro* and under cellular settings. Among the characterized dimethylated products are K2076 of MINT(123) and K140 of STAT3 for SET7/9(124) and K158 of Numb for SETD8.(146) Although alternative interpretation of these observations could be that SET7/9 and SETD8 carried out monomethylation and other PKMTs added the second methyl group under a cellular setting, computational modeling indeed suggests that SET7/9 can carry out dimethylation by releasing the Tyr-bound water molecule via a channel formed by the G292, A295, Y305 and Y335 residues around the catalytic site of SET7/9.(184)

Set1/COMPASS complex is the yeast homolog of human MLL and SET1 complexes.(185) Although yeast Set1, human MLL and human SET1 all contain a Tyr residue (Y1052 for yeast Set1) at the Phe/Tyr switch site, they can mono, di-, and trimethylate H3K4 *in vivo*. (81, 186) With the yeast Set1 as a model, Cps40/Ypl138 was shown to be essential for the activity of H3K4 trimethylation and thus proposed to induce the conformational change of Y1052 upon interacting with Set1 for a higher degree of methylation.(187–189) This observation thus presents the feasibility to modulate product specificity of SET domain-containing PKMTs via their binding partners. Besides modulating the conformation of the Phe/Tyr switch, other regions of SET domain-containing PKMTs can also be tailored to alter product specificity. For instance, the Y641 residue of EZH2, which is located in the opposite side of the switch residue F724 of EZH2, is often mutated to F/N/S/H/C in human B-cell lymphomas.(161–165) The Y641F/N/S/H/C mutants elevate the catalytic activity of EZH2 toward trimethylation products likely through expanding its catalytic site to accommodate bulkier mono/di-methylated histone H3K27 substrates. For non-SET-domain PKMTs, the full-length DOT1L mainly monomethylates histone H3K79 in the absence of its binding partner AF10, whereas the presence of AF10 alters the specificity of DOT1L toward the dimethylated product.(190) Collectively, product specificity of PKMTs can be determined not only by their primary sequences such as the canonical Phe/Tyr switch but also by the features of substrates and biological contexts of individual PKMTs.

4.9. Potential product inhibition of PKMTs.

Methylated proteins and *S*-adenosylhomocysteine (SAH) are products and the byproduct of PKMT-catalyzed methylation, respectively. While only limited kinetic studies were carried out to explore potential product inhibition for PKMT-catalyzed methylation, the inhibitory effect of methylated products was shown to be minimal. An unmethylated substrate and the methylated product are structurally similar except at the site of lysine modification. Methylated products at catalytic sites of PKMTs can often be competed readily by substrates for multiple turnovers.

In contrast, the reaction byproduct SAH shows a broad range of inhibition against PKMTs (*e.g.,* K_d = 23 μM for SET7/9 versus K_d = 0.36 μM for SYMD2).(28) Here no general rule can be found to predict K_d values of SAH against even closely-related PKMTs. For PKMTs, SAM generally shows higher affinity than SAH, likely because of the formation of stronger S•••O chalcogen bonds via SAM's sulfonium moiety.(117) In mammalian cells, SAH-mediated byproduct inhibition of PKMTs is less concerned because intracellular SAH can be efficiently degraded into adenine and homocysteine by SAH hydrolase (SAHH).(191, 192) Indeed, a common practice to increase intracellular concentrations of SAH is to treat cells with SAHH inhibitors such as adenosine dialdehyde.(193) As an extreme case, the level of intracellular SAH can be elevated by dysregulating a SAM-SAH-associated metabolic pathway. For instance, certain cancers (*e.g.,* lung, liver, kidney, bladder and colon cancers) can upregulate intracellular SAH by rapid consumption of SAM by overexpressing nicotinamide *N*-methyltransferase.(194, 195) The resulting elevation of intracellular SAH and thus inhibition of a panel of PKMTs have been linked to an altered epigenetic state and cancer malignancy.(195) Recent studies also showed that intracellular concentrations of SAM are modulated in a context-dependent manner.(1) For instance, depletion of

methionine (a precursor for SAM's biosynthesis) via diet restriction can rapidly reduce the level of intracellular SAM.(196) Meanwhile, the amount of intracellular SAM can be indirectly affected by signaling enzymes.(1, 196) It is thus interesting to investigate how SAH-mediated inhibition affects the activities of PKMTs in the context of varied concentrations of intracellular SAM.

5. Molecular Recognition of Methyllysine-containing Proteins

5.1. Classification of reader domains that recognize protein methyllysine(s).

As protein lysine methylation progresses for mono-, di-, and to tri-methylated products, the size of lysine side chain gradually increases without altering the overall +1 formal charge on the ε-amine at physiological pH (Figure 2). However, these methylation events can slightly alter the ability of the target Lys to engage in cation-π interactions (increased dispersion of +1 charge around neighboring hydrocarbons and decreased desolvation energy penalty upon binding reader motifs). The progressive lysine methylation also alters the ability to form hydrogen bonds as a donor and acceptor (Figure 2). The collective difference between Lys, Kme1, Kme2 and Kme3, though subtle, can be distinguished by the proteins containing methyllysine reader motifs (Table 2). A wealth of Lys(Kme1/2/3)-specific reader modules have been discovered over the past decade and can be classified into at least 15 classes (Table 2): ADD (*e.g.,* a domain in DNMT3L),(197) ankyrin (*e.g.,* domains in G9a and GLP1),(198) BAH (*e.g.,* a domain in ORC1),(199) chromo barrel (*e.g.,* a domain in MSL3), (200) chromodomain (*e.g.,* a domain in HP1),(201) double chromodomain (*e.g.,* a domain in CHD1),(202) HEAT (*e.g.,* a domain in Condensin II),(203) MBT (*e.g.,* domain in L3MBT1), (204, 205) PHD (*e.g.,* domains in TAF3, BPTF, ING2 and BHC80),(206–209) PWWP (*e.g.,* domains in hMSH6 and NSD2: PDB 2GFU, 5VC8), SAWADEE (*e.g.,* a domain in SHH1), (210) tandem tudor domain (*e.g.,* a domain in 53BP1),(205) Tudor (*e.g.,* a domain in PHF1), (211) WD40 (*e.g.,* a domain in EED)(212, 213) and zf-CW (*e.g.,* a domain in ZCWPW1). (214, 215) It is worth noting that, while the list above contains all methyllysine reader modules characterized so far, it does not mean that a protein containing such module(s) can always interact with a methyllysine-containing protein. For instance, ADD (*e.g.,* DNMT3a), (216) Tudor (*e.g.,* TDRD3, SMN and SPF30)(217–220) and WD40 (*e.g.,* WDR5),(221) though generally classified as methyllysine readers on the basis of sequence homology, can also be reader proteins of methylarginine. In addition, cognate targets of many methyllysine-specific reader modules remain to be uncovered; no homology model can predict methyllysine-flanking sequences recognized by potential readers. For instance, L3MBTL1 harbors three MBT repeats with each containing a characteristic aromatic pocket (see discussion later) for potential binding of methyllysine.(204, 222–224) However, the binding assays and crystallographic data of native L3MBTL1 and its binding pocket mutants revealed that only the second MBT repeat involves the interaction with its binding partner H4K20me1/2.(204, 222, 223) Given that most well-characterized binding partners of methyllysine reader modules are restricted to histones and that nonhistone substrates of PKMTs are rapidly emerging, it is of great interest to examine whether methyllysine reader modules such as the first and third MBT repeats of L3MBTL1 can recognize nonhistone targets. It is likely that methyllysine reader modules can recognize cognate histone and nonhistone peptides containing similar but not identical methyllysine sequence motifs.

5.2. Mechanisms for reader modules to recognize methyllysine(s).

Selective recognition of Lys(Kme1/2/3) by their matched reader modules is essential for downstream functions of protein lysine methylation. Although it remains challenging to predict binding partners of potential readers solely on the basis of sequence homology, the accumulated structures of reader domains in complex with their methyllysine ligands have shed light on several general mechanistic rules underlying such recognition. One common trait for most methyllysine-specific readers is to recognize methyllysine(s) through a hydrophobic pocket containing aromatic residues (*e.g.,* Phe, Tyr and Trp). Given that Lys, Kme1, Kme2 and Kme3 all carry the overall +1 formal charge at physiological pH, the aromatic pocket serves as a preferred docking site for cation-π interactions. For an optimal cation-π interaction, an ammonium cation (Lys, Kme1, Kme2 and Kme3) is expected to be poised perpendicular to an aromatic ring within its van der Waals contact.(44, 45) In contrast ot the cation-π interaction, the hydrophobic effect of the aromatic pocket only plays a secondary role in recognizing methyllysine.(225) It has been estimated that a favorable cation-π interaction between an ammonium ion and an aromatic ring can gain 2.6 kcal/mol net binding free energy.(44) This value is even higher than that gained via the formation of an ammonium-carboxylate salt bridge, because of higher desolvation penalty for the coulomb interaction, or via a modest hydrogen bond.(44)

While free energy gained through cation-π interactions is the main driver in binding methyllysine targets, the structures of these reader modules have also been tuned to discriminate distinct states of lysine methylation by exploiting subtle differences of the biophysical properties of Lys, Kme1, Kme2 and Kme3. As a general rule with few exceptions, the reader domains that prefer Kme3 over Lys, Kme1, Kme2 often have their binding pockets purely consisting of aromatic residues.(30, 31, 215) The size and geometry of these aromatic pockets are well suited to maximize cation-π interactions with partially charged δ-methylene and three ε-N methyl groups. For the reader domains with the preference of Kme3, replacing any of the three methyl groups with a hydrogen atom causes the loss of the corresponding cation-π and van der Waals interactions. Such a loss is gradually magnified when the second and third methyl groups are removed. Among the known examples (Figure 16) are the interactions of the PHD finger domain of BPTF with H3K4me3 (PDB: 2F6J);(207) the chromodomain of Polycomb and the WD40 domain of EED with H3K27me3 (PDB: 1PFB; 3IIW);(212, 213) the Tudor domain of PHF and the PWWP domain of ZMYND11 with H3K36me3 (PDB: 4HCZ; 4N4H);(211, 226) the tandem chromodomains of human CHD1 (PDB: 2B2W).(202) One exception is ATRX's ADD motif, whose methyllysine-recognizing pocket consists of only one aromatic residue and otherwise polar residues (PDB: 3QLA) (Figure 16).(227) The ATRX ADD domain slightly prefers H3K9me3 over H3K9me2/1 through forming nonconventional carbon-oxygen (CH•••O) hydrogen-bonding interactions between the pocket residues Y203, Q219 and A224 and the ε-N methyl groups (Figure 16).(227)

Unlike Kme3, Kme2 can form a hydrogen bond or salt bridge with an acidic residue such as Glu and Asp. This difference is often exploited by reader modules that prefer Kme2 over Kme3. The Kme2-binding pockets of these reader domains are mainly composed of aromatic amino acids in combination with an acidic amino acid such as Glu or Asp (Figure

17).(198, 199, 204, 205) The aromatic residues engage in cation-π interactions through the partially charged δ-methylene and two ε-N methyl groups of Kme2, while Glu or Asp forms a salt bridge or hydrogen bond with the tertiary amine of Kme2 (Figure 17). Among these examples are the interactions of the BAH domain of ORC1, the second MBT domain of L3MBTL1, the tandem Tudor domain of 53BP1 with H4K20me2 (E93, D355, D1521 in PDB 4DOW, 2PQW and 3LGF, respectively),(199, 204, 205, 224) and the ankyrin repeat of GLP1 with H3K9me2 (E851 in PDB 3B95) (Figure 17).(198) Here it is worth noting that the different modes of interaction between Kme2 and Kme3 often cause only a modest difference in K_d (within 10-fold), likely because the loss of a salt bridge or hydrogen bond for Kme3 can be compensated by hydrophobic and cation-π interactions of the extra ε-amine methyl group in Kme3. Similarly, the slight selectivity of the reader modules for Kme2 over Kme1 largely arises from the former's ability to engage in cation-π and hydrophobic interactions with aromatic residues. The complimentary contributions of cation-π and hydrophobic interactions and the formation of a salt bridge or hydrogen bond may rationalize the promiscuous recognition of H3K4me3/2/1 by the double chromodomain of CHD1,(202) the PHD finger domain of Pygo-HD1-BCL9 complex (PDB 2B2W; 2VPE);(202) H3K4me2/3 by the tandem tudor domain of SGF29 (PDB 3MEA);(228) H3K9me3/2/1 by the PHD-tandem Tudor domain of UHRF1 (PDB 3ASK);(229) H3K9me3/2 by the chromodomain of HP1 (PDB 1KNE);(201) and likely H3K36me3/2 by the PWWP domain of BRPF1 (PDB 2X4W) (Figure 18).(230) One exception is the recognition mode of H3K9me2 by the SAWADEE domain of SHH1 (PDB 4IUT), for which no obvious polar interaction can be identified within its methyllysine-binding pocket (Figure 18).(210)

To solely recognize lower methylation states such as free Lys and Kme1, reader modules engage in more salt bridges and hydrogen bonds rather than cation-π interactions. For instance, the ADD domain of DNMT3L recognizes unmodified H3K4 through two salt bridges and one hydrogen bond between its Glu88/Glu90/Asn93 residues and the ε-amine of H3K4 (PDB 2PVC) (Figure 19).(197) Given the limited space of this polar pocket, adding one methyl group onto the free Lys was shown to significantly decrease the affinity to the ADD domain and adding two or three methyl groups is expected to completely abolish this interaction. As another example, the PHD finger domain of BHC80 recognizes unmodified H3K4 by forming a salt bridge with D489, and hydrogen bonds with the backbone carbonyl group of E488 and one water molecule (PDB 2PUY) (Figure 19).(209) The small polar pocket likely can only accommodate one extra methyl group (Kme1). Interestingly, the chromo barrel domain of MSL3 was uncovered as a methyllysine reader module to recognize H4K20me through an aromatic-rich pocket (PDB 3OA6) (Figure 19), whose spatial arrangement is similar to those of the reader domains of Kme3.(200) It remains to be determined how the chromo barrel of MSL3 distinguishes H4K20me from H4K20me2/3.

5.3. Strategies to amplify affinity and selectivity of individual methyllysine reader modules.

While many methyllysine reader domains have been documented to recognize methyllysine flanked by specific neighboring residues, the affinity and selectivity of a single reader domain for its target is often modest. The majority of methyllysine reader domains interact

with methyllysine peptides with dissociation constants (K_d) in a range of high nanomolar to low micromolar affinity.(30, 31, 215, 231) Additionally, many of these domains can only modestly distinguish the targets that differ by one methyl group (*e.g.,* Kme3 versus Kme2). For instance, the PHD finger of human BPTF (Figure 20), the largest subunit of the ATP-dependent chromatin-remodelling complex NURF (nucleosome remodelling factor), bind H3K4me2/3 peptides with comparable K_d values of 5.0 μM and 2.7 μM (a difference of 0.4 kcal/mol of binding free energy).(207) The extremely modest gain (~2-fold) of the affinity from H3K4me2 to H3K4me3 likely arises from the additional cation-π interaction of Kme3 with aromatic residue(s) (Tyr10, Tyr17, Tyr23 and Trp32) in the methyllysine binding pocket of BPTF (Figure 16).

Given limited affinity and selectivity of a single methyllysine reader module for its target, one can envision that there are additional mechanisms to enhance recognition and binding to these Lys marks in biological settings. One of these mechanisms is to increase local concentrations of methyllysine proteins above threshold K_d values. If methyllysine targets can be clustered to reach an extremely high local concentration (>10-fold above K_d), it is likely that such an effect will facilitate recruitment of methyllysine readers. For instance, HP1 (heterochromatin-associated protein, a small protein containing around 200 amino acids) harbors a chromodomain to recognize H3K9me2/3 with modest K_d values of 2.5~7 μM (Figure 18).(201) In a cellular context, the efficient recruitment of HP1 to specific chromatin loci is partially benefited from high local concentrations of H3Kme3 (Figure 21).(232)

An alternative way to enhance target recognition of methyllysine reader proteins is to cluster multiple cognate reader modules for a specific methyllysine motif within a single protein or protein complex (Figure 21). This microenvironment is also expected to lower apparent K_d values for target binding. Among many examples are the ankryin repeats of G9a and GLP1, which recognize H3K9me1/2.(198) G9a and GLP1 have been shown to form a heteromeric complex and depend on each other to execute their full functions.(198) Both of the proteins harbor an ankryin repeat with the K_d values of 14 μM and 5 μM for H3K9me, and 6 μM and 7 μM for H3K9me2, respectively.(198) The local-concentration effect can be further enforced by multiple-module reader proteins (or protein complexes) in combination with regionally clustered cognate methyllysine proteins (a multivalent binding mode) (Figure 21).(233) It appears that ZMET2 and likely its homologs leverage such a mode to recognize H3K9me2 marks and then deposit DNA methylation.(234) ZMET2 contains BAH and chromo domains (PDB 4FT2), both of which interact with H3K9me2 (Figure 22).(234) The proposed model of H3K9me2-directed DNA methylation is that ZMET2 recognizes H3K9me2-containing nucleosomes through simultaneous binding of two H3K9me2-containing tails with the BAH domain and chromo domain (Figure 22).(234) This multivalent interaction then positions the DNA methyltransferase domain of ZMET2 for subsequent methylation of nearby nucleosome DNA.

The multivalent binding mode can also rationalize enhanced affinity and selectivity of a multiple-module reader protein (or protein complex) against a methyllysine in the context of other matched binding motifs (Figure 21). In an ideal setting, the apparent K_d value of a multiple-module reader protein in a prefect complex with its binding partners is expected to

be the product of the individual K_d values.(235) The selectivity and affinity of a multiple-module reader protein (protein complex) to a specific methyllysine can thus be fulfilled in the presence of other matched binding partners. One example is the PHD-bromodomain cassette of BPTF (Figure 20).(207) Here the PHD domain and the bromodomain are separated by an α-helical linker, and recognize H3K4me3 and acetylated lysine, respectively. The recognition of H3K4me3 is enhanced in the presence of H4K16 acetylation (H4K16ac) but not other H4 acetylation marks such as H4K12ac and H3K20ac because of the preferred distance between H3K4me3 and H4K16ac.(207)

On the other hand, the affinity for a specific methyllysine can also be suppressed in the presence of an unmatched binding motif (Figure 21). For instance, the PHD finger of ING2 (PDB: 2G6Q) preferentially recognizes H3K4me3 with the K4me3 and the H3R2 separated by Thr3 (Figure 23).(208) This was shown to form a preferred hydrogen bond between Thr3 and ING2's K249. In contrast, the same PHD finger disfavors H3K9me3 and H3K27me3, which contain an Arg (R8/R26) instead of Thr adjacent to Kme3 and likely abolish the T3-involved hydrogen bond (Figure 23). Collectively, even though individual methyllysine reader modules may recognize their binding partners with only modest affinity and specificity, these effects can be significantly amplified via multivalent interactions associated with a target methyllysine and its adjacent binding motifs (Figure 21). Interestingly, many PKMTs or their complexes (*e.g.* NSD1-3 and MLL1-4) contain an array of reader modules for methyllysine(s), as well as other posttranslational modification marks.(55, 72, 236) Such an observation strongly argues that the enzymatic activities of PKMTs can be regulated reciprocally by multivalent interactions of methyllysine reader domains.

6. Erasing Methyl Marks by Protein Lysine Demethylases (KDMs).

6.1. Overview of human KDMs.

The chemically inert nature of methyllysine allows few options for enzymes to remove a methyl group from the ε-amine of a protein lysine residue. The two so-far well-characterized mechanisms of enzymatic demethylation of protein methyllysine involve amino-oxidation and hydroxylation.(33–35) The former is catalyzed by lysine-specific demethylases (LSD).(237, 238) The human genome encodes 2 LSD enzymes (LSD1 and LSD2), which are characterized by the presence of an amine oxidase-like (AOL) domain, known for several metabolic enzymes, and a SWIRM domain, which is characteristic for chromatin-associated proteins (Figure 24).(237, 238) The hydroxylation reaction of protein methyllysine is carried out by KDMs bearing characteristic JmjC domains (Figure 25).(239) These JmjC domains are expected to adopt a double-stranded β-helix topology, as observed for a superfamily of 2-oxoglutarate (2-OG) dioxygenases, with the catalytic sites of KDMs harbored at one end (see KDM4A as an example, PDF 2OQ6, Figure 26).(240) The human genome encodes around 30 JmjC-domain-containing proteins, > 50% of which have been demonstrated as active KDMs.(239) Human KDMs are further diverged into 7 subfamilies on the basis of the sequence homology of their JmjC domains (Figure 25).(239)

6.2. Catalytic motifs and reaction mechanisms of KDMs.

Although the net reaction of all KDMs is to remove a methyl group from the ε-amine of a protein lysine residue, they belong to two families of oxidoreductases (EC 1 in enzyme nomenclature) according to their different catalytic mechanisms. LDS1 and LDS2 are classified as EC 1.5.8 subfamily oxidoreductases with the CH-NH moiety of substrates as an electron donor and a flavin cofactor as an electron acceptor (Figure 27). The crystal structures of LDS1 and LSD2 revealed that their catalytic AOL domains fold into two well-separated, intercalated subdomains with one binding to the FAD cofactor in an extended conformation and the other binding to a methyllysine substrate on the opposite side (Figure 24).(237, 238) The center at the interface of the two domains harbors a spacious cavity as the enzyme active site (Figure 24). This binding mode and the conformation of the FAD cofactor in LDS1 and LSD2 closely resemble those of FAD-dependent oxidases.(241) After recruiting their substrates, the side chain of methyllysine is poised adjacent to the flavin moiety of the cofactor (Figure 27).(237, 238) Such proximity together with matched redox potential promotes oxidation of the ε-amine methyl group, coupled with reduction of FAD into FADH2, to generate the corresponding imine intermediate (Figure 27).(241) The catalytic cycle is then completed upon hydrolysis of the imine intermediate into formaldehyde as a byproduct. Dioxygen-mediated oxidation of FADH2 restores FAD with H_2O_2 as a byproduct (Figure 27). The involvement of the key imine intermediate prevents LDS1 and LDS2 from oxidizing Kme3 as a substrate because of the incapability of Kme3 to be oxidized into an imine.(241)

JmjC-domain-containing KDMs, classified as EC 1.13 subfamily oxidoreductases, are dioxygenases with an ε-amine methyl moiety and 2-OG as electron donors. Structural characterization of the JmjC catalytic core of KDM4A provided the first mechanistic insight of KDMs (Figure 26, PDB 2OQ6).(240) Sequence alignment and structural comparison revealed that JmjC-domain-containing KDMs resemble cupin metalloenzymes in relation to their double-stranded β-helix and jellyroll-like topology (Figure 26).(243) KDM-catalyzed lysine demethylation involves two cofactors, Fe(II) and 2-oxoglutarate, which are located at one end of the jellyroll fold (Figure 28). In general, the catalytic Fe is coordinated with two histidines and one acidic residue (HXX[D/E]···H) via a highly-conserved 2-His-1-carboxylate facial triad (Figure 28).(239, 244) Other binding sites of the Fe cofactor are either vacant or occupied by waters in the absence of 2-OG. The 2-OG cofactor chelates the Fe cofactor through its 2-oxo carboxylate moiety (Figure 28). Interestingly, not all JmjC KDMs bear the characteristic HXX[D/E]···H motif. One exception is PHF2 (Figure 29), which was characterized as an H3K9me2 demethylase.(245) This enzyme appears to use a Tyr (Y321) instead of the conserved His of the facial triad in other KDMs to coordinate with the Fe cofactor (Figure 29).(245)

While a stepwise mechanism of KDM-catalyze demethylation remains to be elaborated, dioxygen is anticipated to enter the catalytic cycle of KDMs by occupying a vacant Fe-coordinate site at the 2-OG-bound stage (Figure 28).(239) Subsequent formation of an Fe(IV) oxo species is facilitated by the electron transfer from Fe coupled with the oxidation of 2OG into succinate and CO_2. In the proximity of the highly reactive Fe(IV)=O species is positioned the methyl group of a methyllysine substrate so that the Fe(IV)=O intermediate

can extract a hydrogen atom from the methyl group to yield Fe(III)-OH and a N-methylene radical. The following rebound of the N-methylene radical with Fe(III)-OH leads to a hemi-aminal lysine intermediate (Figure 28). Alternatively, the hemi-aminal intermediate can be the product of Fe(IV)=O-mediated oxygen insertion (Figure 28). Between the two mechanisms, the oxygen-radical rebound has been favored given a similar observation in a model system.(242) The catalytic cycle is then completed upon hydrolysis of the hemi-aminal lysine intermediate to release the byproduct formaldehyde (Figure 28). In comparison with LDS1/2, which rely on the production of an imine intermediate and thus can only act on Kme1/2 as substrates,(241) JmjC-domain-containing KDMs catalyze a direct hydroxylation of methyllysine and thus can act on Kme1/2/3 as substrates. Recently, an unconventional mechanism was reported for lysyl oxidase-like 2 protein (LOXL2) to remove H3K4me3 through an oxidative deamination at the δ-carbon of methyllysine (Figure 30).(246) For this deamination reaction, though removing a methyllysine mark, installs a novel allysine residue rather than restores an intact lysine residue as the final product.

6.3. Substrate specificity of KDMs.

KDMs recognize substrates through their methylation states and the sequences flanking methyllysine residues. While JmjC-domain-containing KDMs can remove methyl groups from Kme1/2/3 and LSD1/2 can only act on Kme1/2 as substrates,(239, 241) it remains to be determined whether LSD1/2 indeed could not bind Kme3 substrates or can bind them but fail to promote catalysis. A general observation is that KDMs, harboring demethylase activities towards the substrates with high degrees of methylation (*e.g.* Kme3/2), can act on the cognate methyllysine substrates with lower degrees of methylation (*e.g.* Kme1/2). The structures of KDMs in complex with substrates reveal that these enzymes contain a cavity to accommodate the methyllysine side chain of substrates (Figure 26,31).(237, 238, 240) The methyllysine side chain, which is buried deeply near the Fe-(II) cofactor, engages in few specific interactions with KDMs (Figure 26).(240) It is thus expected that Kme2-binding pockets of some KDMs provide certain steric constraints and thus prevent them from binding Kme3-containing substrates. In contrast, it is ready for KDMs to accommodate the cognate substrates with lower degrees of methylation.(244) Collectively, at least two factors, the demethylation mechanism (JmjC enzymes versus LSD1/2) and the spatial restriction of methyllysine binding pockets, determine whether a KDM can act on Kme3 as a substrate. In contrast, demethylase activities of most KDMs have been demonstrated with Kme1/2 as substrates.

Besides a specific methylation state of lysine substrates, sequences flanking a methyllysine residue also play roles on substrate recognition of KDMs. The structures of KDMs (LSD1/2 and JmjC enzymes) in complex with substrates reveal long-range interactions for substrate recognition (Figure 31).(237, 238, 240) For instance, the side chains of Arg2, Gln5, Thr6, Arg8 and Ser10 of histone H3 involve intermolecular or intramolecular interactions for recognition of LSD1's substrate (Figure 31).(237) So far, most histone lysine methylation marks are shown to be substrates of KDMs.(247) Similar to PKMTs that can act on diverse nonhistone substrates, KDMs have been shown to demethylate methyllysine in nonhistone proteins such as E2F1,(122) DNMT1,(248) STAT3,(124) ERα(249) and HSP90α.(250) Most of these nonhistone targets were characterized as LSD1's substrates without well-

defined sequence homology at their demethylation sites (Figure 31). This observation is remarkable because LSD1 was shown to bind its H3K4me2 substrate in a highly sequence-specific manner (Figure 31). The paradox between the promiscuous sequences and the specific recognition of LSD1 substrates can be rationalized by at least two nonexclusive mechanisms. First, the substrate-recognizing pocket of LSD1 can be flexible and thus adopt multiple conformations to accommodate different substrates. Second, LSD1's substrate specificity is modulated through recruitment of other regulatory partners. For instance, the SWIRM domain of LSD1 can associate with androgen receptors and this interaction alters LSD1's substrate preference from H3K4me2 to H3K9me2.(251) It is also of interest to explore whether other KDMs like LSD1 act on diverse nonhistone substrates.

7. Profiling Protein Lysine Methylation

Given that PKMTs can function via methylating diverse histone and nonhistone targets, many efforts have been made to interrogate global methylation profiles (methylomes) of individual PKMTs.(21, 27, 141) However, it is not trivial to uncover PKMT-associated methylomes with conventional methods for at least two reasons. First, distinct from other lysine posttranslational modifications, protein lysine methylation does not alter the overall +1 form charge of a lysine side chain (Figure 3). Secondly, adding a small methyl group does not affect the overall size of a lysine residue (Figure 2). The similar electrostatic and steric properties between lysine and methyllysine thus make it challenging to recognize methyllysine-containing proteins in an efficient and selective manner. Such situations can be further complicated if individual PKMTs act in a temporal and context-dependent manner or a set of PKMTs show redundant methyltransferase activities. As a result, many biologically relevant protein lysine methylation events can be invisible in complicated cellular contexts. To address these challenges, the most commonly used strategy relies on broadly specific anti-methyllysine antibodies to enrich methylomes.(252–254) In addition, sequence-promiscuous methyllysine reader domains can be used as an antibody equivalent to enrich methylomes.(140, 142, 255) Recent efforts in the Luo laboratory as well as others also allow the development of SAM analog cofactors for chemical labeling of methylomes for target enrichment and identification.(21, 27)

7.1. Uncovering lysine methylomes through chemical labeling.

PKMT-catalyzed protein lysine methylation is characterized by the transfer of SAM's sulfonium methyl group to protein lysine substrates through a S_N2 transition state.(98) Given the challenge to profile proteome-wide lysine methylation events, an attractive alternative is to uncover lysine methylomes through chemical labeling.(21) In this approach, the sulfonium methyl group of the SAM cofactor can be replaced with its isotopic or chemical isosteres (*e.g.,* CD_3, $^{13}CD_3$, CT_3, $^{14}CH_3$, allyl and propargyl), which are transferred by native or engineered PKMTs to their substrates.(141, 256–260) The distinct properties of these methyl isosteres then allow more robust signal readouts or target enrichment.

7.2. Enzymatic labeling lysine methylomes with isotopic SAM cofactor.

For chemical labeling of a methylome with isotopic isosteres of the SAM cofactor (heavy methyl labeling), *S*-[methyl-^{13}C/^{2}H]-SAM and radioactive *S*-[methyl-^{14}C/^{3}H]-SAM have been broadly used.(257, 258) Because of SAM's poor membrane permeability, the direct use of the isotopic SAM cofactors is largely restricted for its use in *in vitro* settings such as cell lysates.(21) SAM is synthesized within living cells by highly-conserved methionine adenosyltransferases (MATs) with endogenous ATP and methionine as co-substrates (Figure 32).(19, 21) To label PKMT targets inside living cells with isotopically labeled SAM cofactors, *S*-[methyl-^{2}H/^{3}H/^{13}C/^{14}C]-methionine can be used as the precursors for the MAT-catalyzed biosynthesis (Figure 32). The isotopically labeled methionine can be readily internalized by living cells, likely through amino acid transporter(s), and processed into the corresponding SAM cofactors by MATs for target labeling (Figure 32).(257, 258) In this labeling process, protein translation inhibitors such as cycloheximide are added to prevent direct incorporation of the isotopic methionine into protein through RNA translation machinery.(258)

For *S*-[methyl-^{2}H/^{13}C]-SAM, which contain stable isotopes, [CH_3]/[methyl-^{2}H/^{13}C]-SAM are often used in a combination (Figure 32). The resultant [CH_3]/[methyl-^{2}H/^{13}C]-labeled PKMT targets can then be characterized unambiguously upon detecting the pair of [CH_3]/[methyl-^{2}H/^{13}C]-labeled products or in a more quantitaive manner by analyzing their isotopic ratios with MS (Figure 32).(257) Ong and Mann pioneered chemical labeling of a methylome with *S*-[$^{13}CD_3$]-methionine as the biosynthetic precursor of *S*-[$^{13}CD_3$]-SAM and uncovered around 100 methylation events from the proteome of HeLa cells. The existence of a pair of mass shifts of 14 Da for [CH_3]-labeling and 18 Da for [$^{13}CD_3$]-labeling is characteristic upon revealing bona fide methylation events (Figure 32).(257) With the aid of a recently developed mathematic algorithm, the isotopic ratios can be determined precisely by MS with statistical errors within only a few percent.(98)

For chemical labeling of a methylome with *S*-[methyl-^{3}H/^{14}C]-SAM, these radioactive cofactors are often used as tracers in a mixture containing a large amount of unlabeled SAM as a cold carrier.(258) The limited use of radioactive materials can be due to the presence of unlabeled cognates in available radioactive materials (*e.g.,* for a high percentage of ^{12}C-SAM in ^{14}C-containing SAM), cost of excessive radioactive materials and potential environmental contamination. Therefore, using *S*-[methyl-^{3}H/^{14}C]-SAM as PKMT cofactors is often restricted in well-defined *in vitro* settings because of their poor membrane permeability.(27) In contrast, *S*-[methyl-^{3}H/^{14}C]-methionine can be readily uptaken by cells and processed by endogenous MATs into *S*-[methyl-^{3}H/^{14}C]-SAM for target labeling inside living cells. The resultant [methyl-^{3}H/^{14}C]-labeled targets can be probed by autoradiography for characteristic radio emission. Detection thresholds and signal-to-noise ratios of *S*-[methyl-^{3}H/^{14}C]-SAM-labeled targets can be further improved in combination with advanced sample enrichment strategies as discussed later.(141)

Recent work on KIEs of PKMTs with [CD_3]-SAM cofactors and [CT_3]-SAM further revealed that [CD_3]-SAM and [CT_3]-SAM are more reactive than SAM on the basis of their inverse KIEs and 20~30% larger k_{cat}/K_m values in comparison with [CH_3]-SAM.(98, 111) This difference can originate from the increased steric impingement of vibrational modes of

the methyl group of the SAM cofactor at the transition state in comparison with those in the ground state (Figures 9, 10). Such an effect allows heavy SAM cofactors such as [CD_3]-SAM and [CT_3]-SAM to react faster than native SAM. In contrast, [$^{14}CH_3$]-SAM is slightly less reactive than SAM as reflected by its normal KIEs and 3~5% smaller k_{cat}/K_m values.(98, 111) This rate decrease is due to the loss of the overall bond order of the transferred methyl group at the transition state. Given low radiospecificity of ^{14}C, [CT_3]-SAM is preferred over [$^{14}CH_3$]-SAM as a cofactor surrogate for target labeling.

7.3. Further chemical derivatization of lysine methylomes for MS analysis.

Besides enzymatic labeling of methylomes with isotopic SAM cofactors, proteome-wide Kme1/2 sites as well as unmodified Lys can be subjected to further chemical derivatization (Figure 33).(256, 261–263) For instance, Lys and Kme1 can be modified by anhydrides (*e.g.,* propionic anhydride) to generate the corresponding acylated products (Figure 33).(256, 261) Free Lys and Kme1/2 can be modified by formaldehyde followed by $NaBH_4$-mediated reduction to yield the corresponding Kme3 products (Figure 33).(263) These chemical modifications make the processed protein samples more homogenous for trypsin digestion and MS analysis. In the process of methylome profiling, methyllysine-containing proteins are often subjected to proteolytic digestion by trypsin alone or in combination with other proteases followed by MS analysis.(263) While trypsin primarily cleaves the amide bonds after basic amino acid residues such as Lys and Arg, my laboratory noted that there is residual activity of trypsin to cleave the amide bonds after a Kme1 site. In contrast, the amide bonds of acylated Lys and Kme1/2 are inert for trypsin cleavage (an observation in the Luo laboratory). As a result, the pattern of MS-revealed peptide sequences after trypsin digestion can alter according to the sites and degrees (Lys, Kme1, Kme2 and Kme3) of protein lysine methylation. This chemical derivatization step essentially spares all Lys sites from trypsin digestion (Figure 33) and thus allow a homogenous, ArgC-like proteolytic digestion pattern regardless of the prior status of Lys methylation. For the formaldehyde-involving reductive amination, isotopically labeled formaldehyde can be used to further distinguish the methylation events associated with chemical derivatization from those catalyzed by PKMTs.(263) In addition, the processed peptides have Kme3 at all Lys sites. The prior states of lysine methylation thus have no effect on their ionization efficiency for quantitive MS analysis. Collectively, the distinct chemical properties and thus selective derivatization of Lys, Kme1, Kme2 and Kme3 allow the production of more homogenous lysine methylome samples for trypsin digestion and MS quantification (Figure 33).

7.4. Chemical labeling by native PKMTs with clickable SAM analog cofactors.

Given the merit of terminal-alkyne/azide-containing moieties for copper-catalyzed cyclization addition (click reaction), terminal-alkyne/azide-containing (clickable), activity-based probes have been developed to examine protein posttranslational modifications.(27) While the utility of clickable SAM analogs was documented for methyltransferases in 2006, it was until 2010 that there was the first report of an active clickable SAM analog to label PKMT targets (Figure 34).(27) The Weinhold group first showed that (*E*)-pent-2-en-4-ynyl SAM analog (EnYn-SAM) has the detectable activity as a SAM surrogate toward a fungal PKMT (Dim-5) and two human PKMTs (MLL1 and MLL4).(264) Here EnYn-SAM features SAM's sulfonium methyl group replaced by a sterically bulky, clickable (*E*)-pent-2-

en-4-ynyl moiety (Figure 34). To explore general applicability of sulfonium-alkyl SAM analogs as cofactors of wild-type PKMTs, the Luo laboratory evaluated five SAM analogs (allyl-SAM, propargyl-SAM, EnYn-SAM, (*E*)-hex-2-en-5-ynyl-SAM (Hey-SAM) and 4-propargyloxy-but-2-enyl-SAM (Pob-SAM) (Figure 34) against a panel of human PKMTs.(191) Among the examined 5×5 pairs of PKMTs and SAM analogs, only native SUV39H2, G9a and GLP1 are active towards allyl-SAM. In contrast, the bulky SAM analogs such as EnYn-SAM, Hey-SAM and Pob-SAM showed undetectable activities as cofactors of the 5 examined native PKMTs.

The lack of activity of propargyl-SAM particularly caught our attention given that the overall structure of this SAM analog is comparable to that of allyl-SAM, which is an active cofactor for several PKMTs.(21, 27, 143, 191, 259, 265) Propargyl-SAM was envisioned as the smallest SAM surrogate containing a clickable moiety for the transfer reaction and subsequent target characterization with the well-established alkyne-azide click chemistry (Figure 34).(259) Unfortunately, progargyl-SAM, though was characterized to be stable at an acidic pH, is unstable at physiological pH with a half-life time shorter than 1 min, a time scale that is too short to label PKMT targets in an efficient manner.(259) At physiological pH, progargyl-SAM can be subjected to deprotonation at the double-activated alkyne-sulfonium carbon and decompose into keto-SAM via a putative allene intermediate (Figure 34).(259) Our laboratory as well as the Weinhold laboratory independently overcome the stability issue by replacing the sulfonium with selenium to lower the acidicity at the double-activated carbon center (Figure 34).(266) The resultant ProSeAM (propargylic *Se*-adenosyl-L-selenomethionine) has a half-life time of 1–2 hours at physiological pH with its decomposition mechanism different from that of propargyl-SAM.(259) Consistent with the activity profile of allyl-SAM, ProSeAM is active toward human SUV39H2, G9a and GLP1, which harbor di-/tri-methylation activities (Figure 34).(259, 260) Under a steady-state, GLP1 processes ProSeAM as a cofactor with k_{cat} of 0.375 min^{-1} and K_m of 45.4 μM, which are only 5- to 15-fold different from k_{cat} of 1.97 min^{-1} and K_m of 3.1 μM of the native cofactor SAM. [247] Although the work of the Weinhold group showed that ProSeAM is also active toward SET7/9, which was known for its monomethylation activity,(266) the weak activity of SET7/9 on ProSeAM likely arises from a highly sensitive assay used there. Collectively, the activity profile of ProSeAM towards PKMTs strongly argues that ProSeAM is an active SAM surrogate cofactor for the PKMTs (*e.g.,* SUV39H2, G9a and GLP1) whose enzymatic active sites are spacious enough for di-/tri-methylation.

7.5. Chemical labeling with allele-specific PKMT-cofactor pairs.

Although ProSeAM and EnYn-SAM are active cofactors for native PKMTs and have demonstrated their use for target characterization through their clickable moieties, the labeling efficiency of these SAM analogs is modest and their general applicability towards a broad family of PKMTs remains to be examined.(143, 265) More importantly, it is challenging to correlate the revealed targets to the activities of specific PKMTs because ProSeAM and EnYn-SAM are active towards multiple endogenous PKMTs.(259, 260, 264) It is also challenging to associate an altered lysine methylome of a designated PKMT activity through specific perturbation, given the potential cascade effect of such perturbation. To address these challenges as well as low activities of wild-type PKMTs toward sulfonium-

alky SAM analogs, we envisioned a Bioorthogonal Profiling Protein Methylation (BPPM) technology (Figure 35).(27, 99, 143, 265, 267) In BPPM, highly conserved SAM-binding pockets of PKMTs are engineered to process bulky sulfonium-alky SAM analogs (*e.g.,* Hey-SAM or Ab-SAM), which are otherwise too steric to be accommodated by wild-type PKMTs.(99, 143)

Upon searching engineered PKMT-cofactor pairs for BPPM, we first focused on G9a and GLP1 given their well-characterized *in vitro* methyltransferase activities.(99, 143) The Y1154A mutant of G9a and its equivalent Y1211A mutant of GLP1 but not wild-type G9a and GLP1 were shown to be active towards Hey-SAM and 4-azidobut-2-enyl SAM (Ab-SAM). Remarkably, the apparent $k_{cat}/K_{m,cofactor}$ values of the two PKMT variants with Hey-SAM as a cofactor are comparable to those of wild-type G9a and GLP1 with the native cofactor SAM.(99, 143) Benefited from the clickable terminal-alkyne/azido moieties of these SAM analog cofactors, the resultant labeled substrates can be readily visualized through conjugation with fluorescent dyes or enriched with biotin reporters (Figure 36).(268, 269) Since these engineered PKMTs contain a single-site mutation, which is remote from their substrate-binding pockets, these mutations are not expected to affect substrate recognition. This argument is further supported by ready validation of many BPPM-revealed substrate candidates.(268, 269)

Allyl-SAM, ProSeAM, EnYn-SAM, Hey-SAM, Pob-SAM and Ad-SAM have demonstrated cofactor activities towards native or engineered PKMTs.(99, 143) One common structural feature of these active SAM analog cofactors is that the sp^3 carbon center is double-activated by its immediate adjacent sulfonium and sulfonium-β sp^1/sp^2 carbons for a S_N2 transfer reaction (Figure 34).(99, 143) The sulfonium-β sp^1/sp^2 carbons are indispensable for *S*-alkyl SAM analogs to be active towards PKMTs (native or engineered), whereas the equivalent *S*-alkyl SAM analogs lacking the sulfonium-β sp^1/sp^2 carbons are typically inert as methyltransferase cofactor surrogates.(99) A direct rationale for this observation is that the sulfonium-β sp^1/sp^2 carbons conjugate and thus stabilize S_N2 transition states of the PKMT-catalyzed transalkylation reaction (Figure 37).(99) Consistent with partial S-C bond breaking at S_N2 transition states of PKMTs revealed by KIEs, the replacement of the sulfonium in *S*-alkyl SAM analogs with selenium only slightly accelerates the rates of the transalkylation reaction of a small set of PKMTs (native or engineered).(260)

Structural analysis on native PKMTs suggests that the highly conserved Tyr residue in SET domain-containing PKMTs (*e.g.,* Y1154 of G9a and Y1211 of GLP1) is essential for their transmethylation reactions.(99) Small but significant inverse BIEs upon binding $[CD_3]$-SAM and $[CT_3]$-SAM by PKMTs suggest the steric impingement of the surrounding environment of SAM's methyl group (Figure 13).(98, 111, 112) The observed inverse BIEs could partially originate from noncanonical CH•••O hydrogen-bonding interaction.(98, 111, 112) Such interaction is expected to construct an electrostatic pore to confine the motion of the sulfonium methyl group to assemble a S_N2 transition state (Figure 7).(98, 111, 112) Disruption of this interaction with Ala mutation in G9a and GLP1 leads to a 300-fold decrease of the catalytic efficiency with the SAM cofactor (Figure 37).(99) We envision that the loss of the noncanonical CH•••O hydrogen-bonding interaction in G9a Y1154 mutant and GLP1 Y1211 mutant with the native SAM cofactor can be partially compensated by the

transition-state stabilization through the sulfonium-β sp^2 carbons of these SAM analogs. This model is consistent with the 10-fold higher k_{cat} value of the allyl-SAM than that of the native SAM as the cofactors of G9a Y1154 mutant and GLP1 Y1211 mutant (Figure 37).(99) Remarkably, double-activated sulfonium-β-alkenyl SAM analogs show comparable affinity to the Y1211A/Y1154A mutants as reflected by less than 5-fold difference of their K_m values (12–80 μM for Y1154 and 8–40 μM for Y1211A), despite the dramatic difference in the size of their sulfonium-δ-substituents. In contrast, the corresponding k_{cat} values can alter by 100-fold (Figure 37).(99) This observation suggests that the two PKMT mutants are spacious and flexible enough to accommodate structurally diverse *S*-alkyl SAM analogs.(99) However, the optimal binding of structurally matched *S*-alkyl SAM analogs is essential to correctly position the double-activated sp^3 carbon center for a linear S_N2 transition state (Figure 37). Collectively, two essential structural features for allele specific *S*-alkyl SAM analog cofactors are (i) sterically matched δ-substituents for optimal cofactor binding to assemble S_N2 transition states; (ii) the sulfonium-β-sp^2 moiety to lower the energy barrier of S_N2 transition states.

7.6. Bioorthogonal profiling of protein lysine methylation inside living cells.

Similar to the native SAM cofactor, *S*-alkyl SAM analog cofactors show poor membrane permeability and thus are not suitable for target labeling in living cells.(21, 270) Inspired by the success of BPPM in uncovering the methylome of designated PKMTs with cell lysates, we advanced the BPPM technology for living cells (Figure 38).(21, 270) The key step of implementing BPPM inside living cells is to hijack the biosynthetic pathway of SAM with engineered MATs and membrane-permeable *S*-alkyl methionine analogs for *in situ* production of the corresponding *S*-alkyl SAM analogs. The three-step BPPM within living cells consist of the biosynthesis of SAM analogs from methionine analog precursors by MAT2A I117A mutant, *in situ* target labeling by engineered PKMTs, and subsequent enrichment of the distinct modified targets via the click chemistry (Figure 38).(21, 270) Human MAT2A I117A variant was shown to be the so-far most efficient enzyme to process bulky *S*-alkyl methionine analogs into the corresponding SAM analogs.(19, 270) With G9a and GLP1 as examples, we showed that (*E*)-hex-2-en-5-ynyl homocysteine (Hey-methionine analog) can be processed by the I117A variant of human MAT2A into the corresponding Hey-SAM.(270) In the presence of the BPPP-feasible G9a and GLP1 mutants (Y1154A of G9a and Y1211A of GLP1), histone H3 and other chromatin targets of G9a and GLP1 can be efficiently labeled. The subsequent chromatin enrichment with a biotin-azide probe followed by genome-wide sequencing revealed the preferred chromatin loci harboring the methyltransferase activities of G9a and GLP1.(270) Besides robust enrichment through the clickable moiety and unambiguous assignment of the targets to specific PKMTs, the BPPM method in living cells has the merit to capture dynamic methylation events that are subjected to rapid demethylation by KDMs. Bulky lysine modifications by alkyl-SAM cofactors are expected to be more inert for the removal by KDMs.

7.7. Recognizing lysine methylomes through physical interaction.

Because methyllysine-containing proteins often account for a small fraction of a cellular proteome, efficient enrichment of methyllysine-containing proteins is essential to increase signal-to-noise ratios for subsequent characterization.(21, 27) Given the similarity of the

overall physical properties between free lysine and methyllysine, a current hurdle for methylome profiling is to access reagents that can distinguish methyllysine versus free lysine of a broad spectrum of otherwise cognate proteins. While many high-quality anti-methyllysine antibodies were developed to recognize specific methyllysine modifications within well-defined peptide sequences, broad recognition of these antibodies for methyllysine-containing proteomes remains to be improved.(41) In parallel with the effort to develop high-quality pan-anti-methyllysine antibodies, methyllysine reader domains can also be utilized to recognize and enrich lysine methylomes.(140–142, 255) These reagents, in combination with advanced MS technologies, have become valuable additions to elucidating methylomes.

7.7.1. Antibody-based immuno-enrichment of lysine methylomes.—The Mann group documented an early effort to enrich a lysine methylome with broadly specific anti-Kme antibodies.(257) In this case, likely because of low quality of the anti-Kme antibody as a single enrichment reagent, only H3K27 and H4K20 were identified to contain methyllysine modifications.(257) Given potential limitation for individual anti-methyllysine antibodies to recognize a broad spectrum of methyllysine-containing peptide epitopes, recent efforts often rely on a pool of anti-Kme antibodies to recognize methyllysine-containing proteins. For instance, the Bonaldi group combined a broad collection of commercially available anti-Kme antibodies against 5 methyllysine epitopes.(252) In combination with the heavy methyl labeling, fractionation and high-resolution LC-MS analysis, 74 lysine modification sites were revealed in HeLaS3 cells.(252) With three sets of in-house-developed broadly specific anti-Kme antibodies as enrichment reagents, coupled with the heavy methyl labeling, fractionation and LC-MS analysis, the Garcia group identified 413 methyllysine proteins and 552 methyllysine sites in HeLa cells.(271) With a pool of broadly specific anti-Kme polyclonal antibodies, the Garcia group identified 1032 Kme1 sites in esophageal squamous cell carcinoma (ESCC) cells and 1861 Kme1 sites in ESCC cells overexpressing SMYD2.(272) Among the revealed 1861 Kme1 sites, the levels of 35 Kme1 sites showed the negative correlation with both shRNA-mediated knockdown of SMYD2 and the selective inhibition of SMYD2 by LLY-507.(272) In a similar manner, Guo *et al.* developed three broadly specific anti-Kme antibodies as enrichment reagents and applied them to uncover 130 methyllysine proteins and 165 methyllysine sites in HCT116 cells.(254) Some of these anti-methyllysine antibodies can also be used as far Western Blotting reagents in recognition of methyllysine-containing proteins. Interestingly, there is no significant overlap among the methyllysine sites revealed in these different settings, likely due to the lack of pan-anti-methyllysine property of these antibodies even used in combination. It will be useful to examine the structures of these antibodies in complex with methyllysine epitopes of different peptide sequences and evaluate the feasibility to further improve the quality of broadly specific anti-methyllysine antibodies.

7.7.2. Recognition and enrichment of methylomes with methyllysine reader domains.—While methylation only slightly alters physical properties of a lysine residue, such difference can be distinguished to a certain degree by proteins containing methyllysine reader domains. Given that some readers recognize methyllysine with no significant involvement of neighboring residues, these reader domains can be utilized in a similar

manner as pan-anti-methyllysine antibodies. For instance, as revealed by the structure of the three Malignant Brain Tumor domains (3×MBT) in complex with the target peptide, this methyllysine reader motif preferentially recognizes Kme1/2 over Lys and Kme3 independent of methyllysine-flanking sequences (Figure 17).(140–142) The selectivity of Kme1/2 over Lys and Kme3 can attribute to the preferred cation-π and hydrogen bond interaction between the methyl ammonium of Kme1/2 and the D355 residue of 3×MBT in its geometrically matched binding pocket (Figure 17).(204, 205) Promiscuity of 3×MBT toward methyllysine-flanking sequences was further supported by its comparable K_d values upon binding H3K4me1/2, H3K9me1/2, H4K20me1/2, and p53 K382me1/2 peptides.(204, 205) The Gozani group first advanced this observation to utilizing 3×MBT as an enrichment reagent for Kme1/2-containing peptides. With a GST-tagged 3×MBT protein, they uncovered several hundreds of Kme1/2-containing protein candidates from the nuclear extract of HEK293T cells.(140–142) Additionally, the well-defined structure of the 3×MBT-peptide complex allowed the use of the D355N mutant of the GST-tagged 3×MBT protein as a negative control.(140–142)

The Li group presented another example using methyllysine reader domains to enrich a methylome.(255) While the chromo domain of HP1β has been well characterized to recognize H3K9me2/3 peptides (Figure 18), Li and colleagues envisioned the utility of this chromo domain to recognize other Kme2/3-containing proteins.(201) With the immobilized HP1β chromo domain as an enrichment reagent, they identified 109 candidates.(255) Because only the native HP1β chromodomain but no inactive variant of HP1β chromodomain was used as the enrichment reagent in this work, it remains to be determined whether these candidate proteins contain Kme2/3 sites for the recognition by HP1β chromo domain or simply interact with HP1β chromodomain in a Kme2/3-independent manner. With the 3×MBT domain and the HP1β chromodomain as two examples, it is interesting to explore the potential utility of other methyllysine reader domains as methylome-enriching reagents. While the structures of many reader domains in complex with methyllysine-containing peptides have been reported,(30, 31, 215) there is no structural report about how anti-methyllysine antibodies preferentially recognize peptides containing methyllysine epitopes. Such structural information can be valuable to develop new methyllysine-recognizing entities for methylome enrichment.

8. Small-molecule Inhibition of Methyltransferase Activities of PKMTs

PKMTs have received much attention in the past decade given their emerging roles as epigenetic modulators. PKMTs are multifunctional proteins containing a methyltransferase domain for catalysis as well as other motifs to interact with diverse binding partners.(33, 34, 38, 40) The methyltransferase domain of PKMTs functions through methylating diverse histone and nonhistone substrates in a highly context-specific manner.(38) Among the most common biological methods for perturbation of individual PKMTs is shRNA/siRNA-mediated silencing of PKMT transcripts. However, this approach generally targets full-length PKMTs rather than their catalytic domains and itself is not sufficient to unambiguously define the functional roles of the methyltransferase activities of PKMTs. The phenotypes associated with the methyltransferase activities of PKMTs, after shRNA/siRNA-mediated silencing, need to be validated by successful rescue with wild-type PKMTs but not

catalytically dead mutants. Another emerging approach to perturb PKMTs is to rely on CRISPR/Cas9-mediated genome editing, which can target specific domains as well as full-length PKMTs. However, this approach is irreversible and lack of temporal control. In contrast, small-molecule inhibitors of PKMTs can be developed with the specificity against their methyltransferase activities.(27, 38) Like other chemical tools, small-molecule inhibitors have the merit to perturb PKMTs in a temporal (accurate timing), spatial (defined location) and dose-dependent manner.(27, 38) Certain PKMT inhibitors can also be further developed as therapeutic reagents.(33) In contrast, among potential concerns of using PKMT inhibitors are their unpredictable off-target effects and lack of general applicability across cell lines or species.(38) Many prior efforts have been made in academia and industry to develop PKMT inhibitors and led to access to hundreds of inhibitors against human PKMTs.(33) However, the overall quality of these PKMT inhibitors can greatly vary from a small set of compounds that have been vigorously characterized *in vitro*, in cellular contexts and in animals to many others that were simply examined with *in vitro* biochemical assays.(38) The lack of vigorous characterization of many PKMT inhibitors makes it challenging to select suitable compounds for biological studies. There is always a risk to use well-characterized PKMT inhibitors incorrectly and thus misinterpret their biological outcomes.(38) It is thus essential to establish the criteria of high-quality PKMT inhibitors and define relevant contexts for their correct use.

PKMT inhibitors can be classified according to their mechanisms of action (MOA).(38) Defining MOAs is not only informative to optimize potency and selectivity of PKMT inhibitors but also essential to interpret the biological outcomes of specific inhibitors. Along the reaction path of a lysine methylation reaction, PKMTs first recruit the SAM cofactor and substrates to two adjacent but relatively independent binding sites. PKMTs then promote deprotonation of lysine nucleophile through dynamic water channels and catalyze the transmethylation reaction through a S_N2 transition state (Figure 7).(98) For many PKMTs, efficient catalysis also requires the participation of remote residues or the presence of other regulatory PKMT-binding partners (*e.g.,* WDR5 and RBBP5 for MLL1; EED and SUZ12 for EZH2).(79–83) Methyltransferase activities can be inhibited by competitive occupancy of small molecules in the binding sites of the SAM cofactor, substrate peptides or regulatory binding partners of PKMTs. On the basis of these MOAs, representative PKMT inhibitors will be discussed with a main focus on their correct use as chemical tools to interrogate functions of PKMTs.

8.1. General criteria of PKMT inhibitors as chemical probes.

PKMT inhibitors can be developed either as chemical probes to perturb the methyltransferase activities of specific PKMT(s) or as drug candidates for therapeutic use.(38, 273) On the basis of these respective uses, different principles should be applied upon evaluating overall quality of PKMT inhibitors.(274) To develop chemical probes, Frye introduced five general principles in his article of “The Art of the Chemical Probe”.(274) Upon rephrasing these principles in the current setting, high-quality PKMT chemical probes should: (1) show sufficient *in vitro* potency and selectivity against one or a set of designated PKMT(s); (2) show decent *in vivo* or at least cellular potency and selectivity, which correlate with the corresponding *in vitro* data (*in vivo* or cellular EC_{50} versus *in vitro* IC_{50}); (3) be

well characterized in term of MOA *in vitro* and *in vivo* or in a cellular context (*e.g.,* SAM-competitive, substrate-competitive, allosteric or covalent inhibitors); (4) demonstrate at least one utilization (*e.g.,* the treatment of a PKMT chemical probe recapitulates certain biological readouts implicated by genetic perturbation of the PKMT); (5) be accessible through well-described chemical methods or from commercial vendors. Additional criteria required for drug candidates include proper pharmacokinetic, pharmacodynamic and toxicity profiles.

While a well-defined spectrum of potential off-target effects is an essential parameter for a PKMT chemical probe, a therapeutic index of the dose of efficacy versus toxicity is more concerned for a drug candidate. Great caution should be made when PKMT inhibitors are used under unprecedented biological settings (*e.g.,* across different cell types, cellular versus *in vivo* use, oral, intraperitoneal injection versus intravenous dose). Even for the best characterized PKMT inhibitors, it is essential to reevaluate vigorously their target engagement and inhibitory efficiency under different biological settings. Current efforts have been mainly focused on developing small-molecule inhibitors against human PKMTs. While the sequences of human PKMTs can be homologous across species, certain caution should be made by assuming that the potency and selectivity of a PKMT inhibitor could be fully maintained for sequence-related PKMT homologues. There are not short of examples that single-point mutations are sufficient to alter potent inhibitors into completely inactive compounds.(93) A PKMT inhibitor that is characterized for its use in human cell lines needs to be reevaluated upon perturbing related PKMTs across species.

8.2. Inhibition of PKMTs through SAM-competitive MOA.

A common strategy for developing PKMT inhibitors is to engage small molecules to occupy SAM-binding pockets of PKMTs and thus inhibit their catalysis.(38) Among well-characterized SAM-competitive SAM analog inhibitors are EPZ004777 against DOT1L, (275) Pr-SNF/Bn-SNF against SETD2 and Pr-SNF against NSD2 (Figure 39).(75, 77) A key character of these SAM-competitive inhibitors is that their IC_{50} values show a linear increase in the presence of increased concentrations of the SAM cofactor. A SAM-competitive MOA can be further supported by structural data of these PKMTs in complex with corresponding inhibitors, for which SAM-binding pockets are occupied by the inhibitors.(75, 77, 276) As revealed by structural data, most of these PKMTs show dramatic conformational changes upon binding these SAM-competitive inhibitors. For instance, several SAM-binding loop regions of DOT1L, which used to interact with the α-amino carboxylic acid moiety of the SAM cofactor, open up to accommodate the bulkier phenyl urea moiety of EPZ004777; the autoinhibitory loop of SETD2, which used to be in a close configuration upon binding SAM or SAM analogs, makes nearby a 180-degree flipping to accommodate the bulky benzyl moiety of Bn-SNF.(75, 276) While Bn-SNF was characterized as a SAM-competitive inhibitor, its parent IC_{50} also depends on the presence of a H3K36 peptide substrate (Figure 39).(75) This inhibitor shows higher affinity toward SETD2 by forming a ternary inhibitor-SETD2-substrate complex. As a result, Bn-SNF is better defined as a SAM-competitive, substrate-dependent inhibitor (Figure 39).(75)

The pyridone-based GSK126(277, 278) and UNC1999 (EZH2 or dual EZH1/2 inhibitors), (279) and several fragment-based DOT1L inhibitors recently developed by Novartis(280–282) are the SAM-competitive PKMT inhibitors with no structural similarity to the SAM cofactor (Figure 39). Like canonical SAM-competitive cofactor analog inhibitors (Figure 39), these non-nucleotide-based compounds also show linearly increased IC_{50} values in the presence of increased concentrations of the SAM cofactor. The SAM-competitive MOA is also supported by their structural data with the SAM-binding pockets fully or partially occupied by these compounds.(93, 280–282) Interestingly, upon binding most of these compounds, both EZH2 and DOT1L show dramatic conformational changes around their SAM-binding sites.(93, 276, 280–282) For instance, several newly released structures of EZH2 or its homolog enzyme in complex with pyridone-based inhibitors showed that a loop region adjacent to the adenine-binding pocket adopts a new conformation upon replacing SAH with these inhibitors as ligands.(93) In the case of fragment-based DOT1L inhibitors, the binding of these compounds by DOT1L (Figure 39) also involves a distinct pocket adjacent to its SAM-binding site and then induces the conformational changes that are unfavorable for SAM binding.(280–282) The inhibitor-induced conformational changes could also rationalize how these inhibitors selectively interact with specific PKMTs but not others even if they all share highly-conversed domains for SAM recognition. In most of these scenarios, PKMTs are subjected to dramatic conformational changes upon binding these SAM-competitive inhibitors. Similar to the situation of protein kinases,(283) the strong preference of certain PKMTs through the exploration of their structurally matched alternative conformations could be essential for the high potency and selectivity of these PKMT inhibitors. Exploring and targeting structurally distinct conformers of PKMTs may present a general strategy to developing potent and selective SAM-competitive inhibitors.

Multiple factors can affect the potency of SAM-competitive, substrate-noncompetitive PKMT inhibitors inside cells. In general, canonical SAM-competitive PKMT inhibitors act better in a cellular setting with a low concentration of endogenous SAM (Figure 39). Nevertheless, as long as concentrations of these PKMT inhibitors is sufficient to compete with SAM, the methyltransferase activities of the engaged PKMTs should be inhibited regardless of the presence of peptide substrates (Figure 39). Given that intracellular concentrations of SAM may be dramatically affected by the level of other metabolites such as methionine and vary by several folds across cell or tissue types,(1, 196) it is thus important to evaluate the efficiency of SAM-competitive inhibitors in different cellular contexts. The situation can be further complicated in the presence of PKMT complexes. For instance, EZH2 requires the binding of EED and SUZ12 to form a core complex to efficiently catalyze H3K27 methylation *in vivo*.(82, 213, 284) The methyltransferase activity of the core complex can be further enhanced allosterically through interaction of the EED subunit with an H3K27me3 peptide.(82, 213, 284) Here, the SAM-competitive EZH2 inhibitor GSK126 showed around 10-fold increase of its affinity to the EED-SUZ12-EZH2 complex with the additional presence of an H3K27me3 peptide.(277) Given that DOT1L-catalyzed H3K79 methylation can be accelerated upon forming the DOT1L-AF10 complex, (190) it will be interesting to explore whether the SAM-competitive DOT1L inhibitor EPZ004777 shows different affinity to DOT1L versus the DOT1L-AF10 complex. Target engagement can be distinct for SAM-competitive, substrate-dependent inhibitors such as Bn-

SNF for SETD2.(75) In these cases, the optimal cellular setting for the target engagement is the presence of an substrate and the absence of the SAM cofactor to facilitate the maximal formation of the ternary inhibitor-PKMT-substrate complex (Figure 39).(75)

8.3. Inhibition of PKMTs through substrate-competitive MOA.

Another MOA of PKMT inhibitors is to engage small molecules to occupy peptide-binding pockets of PKMTs (substrate-competitive inhibitors) (Figure 40).(33, 38) Given different PKMTs may contain distinct peptide-binding pockets to recognize their substrates, the substrate-competitive MOA is expected to be more feasible for developing selective PKMT inhibitors. UNC0638/UNC0642 (G9a/GLP1 inhibitors),(285) (*R*)-PFI-2 (an inhibitor of SETD7/9),(127) LLY-507(272) and A-893 (two inhibitors of SMYD2),(286) and A-196 (an inhibitor of SUV4-20H1/2)(287) are examples of well-characterized substrate-competitive PKMT inhibitors with cellular activities. Target inhibition of these compounds is characterized by a linear increase of their apparent IC_{50} values with increased concentrations of substrates. The substrate-competitive MOA of these inhibitors is also consistent with the binary PKMT structures in which the small-molecule ligands occupy the substrate-binding pockets of PKMTs (Figure 40). Substrate-competitive PKMT inhibitors can be further classified according to whether the SAM cofactor can be involved to form more stable inhibitor-PKMT-SAM ternary complexes (Figure 40). The G9a/GLP1 inhibitors UNC0638 and UNC0642 and the SUV4-20H1/2 inhibitor A-196 are canonical substrate-competitive inhibitors with their apparent IC_{50} independent upon the presence of the SAM cofactor.(285, 287) In contrast, (*R*)-PFI-2 is a substrate-competitive, SAM-dependent inhibitor with the strong preference to form the inhibitor-SET7/9-SAM ternary complex and thus strengthen the interaction of (*R*)-PFI-2 with SET7/9 in the presence of the SAM cofactor (decreased apparent K_d values with increased concentrations of SAM).(127) AZ505, an analog of the SMYD2 inhibitor A-893, also showed a similar substrate-competitive, SAM-dependent character.(86) It remains to be determined whether the presence of the SAM cofactor increases the affinity of substrate-competitive inhibitors A-893 and LLY-507 to SMYD2.

Cellular target engagement of canonical substrate-competitive PKMT inhibitors negatively correlates with the affinity of PKMTs to inhibitor-competitive substrates as well as their intracellular concentrations. Because PKMTs may methylate diverse histone and nonhistone targets with a broad range of $K_{m,substrate}$ values and intracellular concentrations, EC_{50} values of substrate-competitive PKMT inhibitors will likely vary on different substrates. It is thus likely to apply modest doses of substrate-competitive PKMT inhibitors to perturb the methylation events associated with a subset of substrates with low affinity (high $K_{m,substrate}$) and intracellular concentrations, but spare those substrates with high affinity (low $K_{m,substrate}$) and high intracellular concentrations from inhibition.(38)

It is worth noting that some PKMT inhibitors may partially occupy substrate-binding pockets but do not show the expected substrate-competitive character (the increase of IC_{50} versus increased concentrations of substrates). For instance, EPZ030456 is a selective and potent inhibitor of SMYD3 with a non-competitive character for its substrate MAP3K2.(288) However, the structure of SMYD3 in complex with EPZ030456 shows that the inhibitor residues in the pocket that is otherwise occupied by MAP3K2, a SMYD3 substrate,

in the structure of the SMYD3-MAP3K2 complex (Figure 41).(288) The Lys-binding site of the MAP3K2 substrate can also be occupied by the propyl dimethylamino moiety of the SMYD3 inhibitor GSK2807 (Figure 41).(289) However, GSK2807 also does not show the noncompetitive character against the substrate MAP3K2. To rationalize these paradoxical observations, it was proposed that the optimal binding of MAP3K2 to SMYD3 may mainly occur outside of the substrate pocket with minimal engagement of the lysine substrate and its neighboring residue(s). As a result, the presence of the MAP3K2 substrate has no inference in SMYD3's binding to EPZ030456 and GSK2807. It therefore should be cautious to conclude substrate-competitive MOA solely on the basis of potential steric clash between inhibitors and substrates revealed by their PKMT complexes. The clashed occupancy of inhibitors and substrates does not always grant a substrate-competitive MOA because they may accommodate each other through alternative conformations without significant energy penalty. In contrast, because of the well-defined binding mode of SAM in PKMTs, it is more confident to conclude a SAM-competitive MOA upon observing even partial occupancy of inhibitors in SAM-binding pockets of PKMTs.

8.4. Inhibition of PKMTs by allosteric MOA.

Besides SAM-competitive and substrate-competitive MOAs, a distinct set of PKMT inhibitors were developed by targeting the allosteric sites essential for catalysis (Figure 42). It has been documented that the methyltransferase activities of certain PKMTs depend upon the formation of higher-order complexes and thus participation of the residues remote from their catalytic sites. For instance, the H3K4 methyltransferase activity of MLL1 depends on the formation of a core complex with at least three binding partners WDR5, ASH2L and RbBP5 with the MLL1-WDR5 interaction essential for the catalysis.(79–81) MM-401(81) and OICR-9429(79) were developed as the allosteric inhibitors of MLL1 by their competitive occupancy of the central channel of WDR5, which is otherwise occupied by MLL1 to form the core complex for catalysis (Figure 6 and Figure 41). Similar strategies have also been applied to develop MI-2/3/463(80, 290) and SAH-EZH2(83) as disruptors of MLL1-Menin and EZH2-EED-SUZ12 complexes for allosteric inhibition of MLL1 and EZH2, respectively (Figure 6 and Figure 41). The methyltransferase activity of the core EED-SUZ12-EZH2 complex can be alloterically modulated through the interaction of the aromatic cage of EED subunit consisting of F97, Y148 and Y365 with an H3K27me3 peptide. EED226, A-395 and their derivatives were developed by occupying this aromatic cage.(82, 284) The induced conformational arrangement of the aromatic cage upon binding these inhibitors is expected to be rendered to remote site(s) to perturb catalysis (Figure 6 and Figure 41). The PKMT inhibitors characterized as allosteric MOA often engage their targets at the sites remote from the catalytic pockets and thus show noncompetitive characters towards the SAM cofactor and substrates (EC_{50} independent upon the presence of substrate and SAM) (Figure 42). PKMT inhibitors with the allosteric MOA are therefore complimentary with SAM-competitive and substrate-competitive PKMT inhibitors for engagement of PKMTs in a cellular context.

8.5. Inhibition of PKMTs through covalent or suicide MOA.

While the development of cysteine-targeted covalent inhibitors has been well documented for kinases, few PKMTs inhibitors were developed through covalent MOA. With substrate-

competitive SETD8 inhibitors UNC0379 and MS2177 as lead scaffolds, the Jin laboratory developed its analog MS453 by installing an electrophilic acrylamide group to selectively target the C311 residue of SETD8 (Figure 43).(291, 292) MS453 showed time-dependent target engagement and thus decreased IC_{50} values as extending preincubation time with SETD8. This covalent MOA is also supported by the structure of SETD8 in complex with MS453.(292) Upon screening > 5,000 commercial compounds, SPS8I1 (NSC663284), SPS8I2 (ryuvidine) and SPS8I3 (BVT948) were identified as SETD8 inhibitors with modest selectivity (Figure 43).(293, 294) These compounds as well as their derivatives such as SGSS05-NS and SPECS21(295) feature a common quinone motif with anticipated covalent modification of a Cys residue. SPS8I1 (NSC663284) and SPS8I2 (ryuvidine) were characterized to covalently modify C311 of SETD8, while SPS8I3 (BVT948) may target multiple Cys residues of SETD8.(293) To develop inhibitors against DOT1L, Yao and coworkers described SAM-based suicide inhibitors (Figure 43).(110) These compounds are expected to undergo an intracellular cyclization to generate a highly reactive aziridinium (Figure 43). This intermediate then reacts with the K79 of the substrate H3 to generate a binary substrate-inhibitor adduct for DOT1L inhibition (Figure 43).(110) A similar suicide strategy was also reported to develop potent inhibitors of protein arginine methyltransferases.(296) It remains to investigate whether covalent or suicide MOAs can be generally applicable for PKMT inhibition.

8.6. Evaluation of target engagement of PKMT inhibitors.

Because complicated factors contribute to the potency of PKMT inhibitors, their cellular target engagement should be evaluated. PKMT inhibitors are often appended with a functional anchor (*e.g.*, a terminal alkyne, azide or biotin moiety) at an inert position of inhibition. This anchor can then be used to pull down engaged PKMT(s) from cell lysates.(38) Alternatively, a cellular thermal shift assay (CTSA) can be implemented.(297) Here the binding of inhibitors to PKMTs is expected to increase their thermal stability. Target engagement of PKMT inhibitors can also be evaluated indirectly through their efficiency to block relevant methylation marks.(38) For canonical SAM-competitive and allosteric PKMT inhibitors, complete target engagement simply requires to show a dose-dependent decrease of relevant methylation marks (*e.g.* H3K27me3 for EZH1/2).(38) For canonical SAM-competitive or allosteric MOAs, lack of PKMT-dependent methylation often indicates that the current dose is sufficient for these inhibitors to fully occupy the SAM-binding pocket or allosteric site of the targeted PKMT to prevent catalysis. In contrast, complete target engagement of a substrate-competitive PKMT inhibitor should be evaluated with the most robust substrate (the highest intracellular concentration and the lowest $K_{m,substrate}$). In addition, the depletion of specific methylation marks also depends on the lifetime of the examined protein substrates and antagonistic activities of demethylases. A longer period of target engagement is required for more stable methylation marks to be suppressed.(38)

It is worth noting that, besides the inhibition of methyltransferase activities, the engagement of PKMT inhibitors with their targets can lead to conformational changes of PKMTs and thus alter other binding events. While substrate-competitive PKMT inhibitors perturb the catalysis through preventing PKMTs from binding their substrates, such perturbation is also expected to alter the integrity of the whole PKMT complex (*e.g.* MM-401 disrupts the

interaction of WDR5 with MLL1 in the MLL1–WDR5–ASH2L–RbBP5 complex).(81) Therefore, the treatment of PKMT inhibitors can alter components of PKMT complexes and have profound effects beyond the only inhibition of substrate methylation.

9. Semisynthesis of Proteins Containing Methyllysine or Methyllysine Mimics

Given diverse functional roles of methyllysine-containing proteins, it is of critical need to access high-quality methyllysine-containing proteins. Homogenous proteins containing methyllysine modifications cannot be readily prepared through direct enzymatic labeling because of low efficiency of PKMT-catalyzed methylation *in vitro* and in living cells.(27) In contrast, semisynthesis of methyllysine-containing proteins or their precursors have been well documented.(27) These semisynthethic methods can be classified into three categories in general: (a) cysteine-specific chemical conjugation, (b) nonsense-suppression-mediated mutagenesis and (c) chemical ligation as will be detailed below.

9.1. Cysteine-specific chemical conjugation.

The free-thiol of cysteine in cysteine-containing proteins is often explored for site-specific chemical installation of lysine, methyllysine or their analogs. With the free-thiol of cysteine as a nucleophile, the Shokat laboratory first reported a semisynthetic approach to conjugate an *N*-methyl aminoethyl moiety to proteins (Figure 44).(298) The resulting *N*-methylated aminoethylcysteines are generated as methyllysine analogs (MLA) (Figure 44). The overall similarity between methyllysine and MLA has been confirmed by their equivalent recognition as epitopes by anti-methyllysine antibodies and methyllysine binders, and as substrates of multiple PKMTs.(298) For instance, Margueron *et. al.* relied on this MLA approach to prepare a series of MLA-containing histones and used them as substrates to examine the crosstalk between the PRC2 complex (EED-Suz12-EZH2) and various histone methylation marks (*e.g.,* H3K27, H3K36, and H3K9).(213) This work demonstrated that the EED subunit of the PRC2 complex enhances EZH2's catalysis by interacting with nucleosomes containing the MLAs of H3K27me3 and H3K9me1/2/3 but not H3K36me1/2/3.(213) In a more systemic manner, the Zhu laboratory evaluated biochemical compatibility of MLA-containing histones including their interactions with the reader proteins 53BP1 and the G9a ankyrin repeats, and their reactivities as the substrates of 6 PKMTs and one KDM.(299) In comparison with natural methyllysine, the replacement of the γ methylene moiety with the sulfide results in a slight increase of the corresponding bond length by 0.28 Å and a small decrease of p*K*a by 1.1 unit.(299) Such difference seems to have more impact upon quantitative evaluation of MLAs as methyllysine surrogates, in which natural methyllysine is often better than MLAs to engage binding partners.(300)

The Davis laboratory recently developed a cysteine-specific, radical-based semisynthetic method to incorporate methyllysine as well as other side chains into proteins.(301) On the basis of the prior work of 2,5-dibromohexanediamide (DBHDA) as a double-activated electrophile, the coworkers in the Davis laboratory converted a cysteine residue in proteins into dehydroalanine (Dha) (Figure 44).(301, 302) The resultant Dha-containing protein was then subjected to $NaBH_4$-mediated radical coupling with organic iodides to afford the

corresponding methyllysine-containing products (Figure 44). With the newly-developed semisynthetic method, they generated a wide range of methyllysine-containing proteins. With H3K9me3-containing H3 as a demonstration, they showed that the semisynthetic histone can serve as a substrate of KDM4A. This semisynthetic method to access diverse methyllysine-containing proteins features its readiness, robustness and scalability. Yang and coworkers independently reported a method to convert Dha into methyllysine through a Zn/Cu-mediated radical coupling reaction (Figure 44).(303) Here a Dha residue is generated from Cys by *O*-mesitylenesulfonylhydroxylamine (MSH), which can be accessed chemically or introduced genetically in a site-specific manner through unnatural amino acid precursors (see discussion below) (Figures 44,45).(303) In comparison with the DBHDA/MSH-mediated production of Dha from a natural cysteine, site-specific incorporation of unnatural amino acid precursors allows installing methyllysine without influence of other existing cysteine residues.(303) One limitation for this radical-based semisynthetic method lies in the introduction of a *D,L*-methyllysine mixture rather than naturally occurring *L*-methyllysine. In addition, this method has not demonstrated its utility in living cells given relatively harsh conditions for the radical reactions.

9.2. Incorporation of methyllysine precursors through nonsense suppressors.

Unnatural amino acids can be incorporated into proteins in a site-specific manner with orthogonally engineered tRNA/tRNA synthetase pairs.(27) This method has been well documented for its utility in a cell-free translational system, within bacterial, fungal and mammalian cells, and in living animals.(304, 305) Given the challenge of recognizing the small difference between methyllysine and free lysine by cellular translational machinery, site-specific incorporation of methyllysine into proteins is often achieved by installing methyllysine precursors or caged methyllysine, followed by additional chemical modifications to yield methyllysine-containing proteins (Figure 45). For instance, the Schultz laboratory reported the work to prepare proteins containing MLA through site-specific phenylselenocysteine (SecPh) chemistry (Figure 45).(306) Here SecPh was incorporated into proteins through a nonsense suppressor coupled with orthogonal tRNA/tRNA synthetase pairs. The SecPh was then subjected to H_2O_2-mediated oxidation to afford Dha.(306) In this situation, the authors reacted the Dha-containing protein with *N*-methylated 2-mercaptoethylamine to generate the corresponding MLA-containing product (Figure 45). Yang and coworkers recently reported a new synthetic path to access Dha-containing proteins.(303) Distinct from the DBHDA-mediated production of Dha from cysteine (the Davis laboratory)(301) and the H_2O_2-mediated production of Dha from SecPh (the Schultz laboratory),(306) Yang and coworkers site-specifically introduced an *O*-phosphoserine (Sep) with orthogonally engineered tRNA/tRNA synthetase pairs (Figure 45).(303) The Sep residue was then chemically eliminated to afford Dha under basic conditions. Theoretically, these Dha-containing protein intermediates can be subjected to either *N*-methylated 2-mercaptoethylamine to afford MLA-containing protein products or the transition-metal-mediated radical coupling with the corresponding organic iodides to afford methyllysine-containing protein products (Figures 44,45). It is worth noting that the Dha-mediated chemistry always affords the *D,L* isomers of methyllysine or MLAs and needs to be performed *in vitro* with denatured proteins.

To access proteins containing enantiomerically pure methylysine, the Liu laboratory recently reported the amber-suppressor-mediated incorporation of N^{ε}-(4-azidobenzoxycarbonyl)-δ,ε-dehydrolysine (AcdK) as a methyllysine precursor.(307) AcdK contains a liable δ,ε-dehydrolysine side chain, which is protected by an azidobenzoxycarbonyl moiety (Figure 45). This moiety is readily subjected to phosphine-triggered self-cleavage to afford allysine (AlK) through an unstable intermediate δ,ε-dehydrolysine (Figure 45).(307) The subsequent $NaCNBH_3$-mediated reductive amination of AlK with monomethylamine or dimethylamine yields corresponding methyllysine-containing proteins. In comparison with the Dha-mediated incorporation of methyllysine and MLAs, the AcdK-AlK-based method is distinct for its feature to install enantiomerically pure methyllysine (Figure 45). Given the requirement of the reductive amination upon the conversion of AlK into methyllysine, this approach can only be implemented for installation of Kme1 and Kme2 but not Kme3 into proteins (Figure 45).(307) Because of involvement of multiple chemical steps under biologically incompatible conditions, the AcdK-AlK-based semisynthetic method, similar to Dha-mediated incorporation of methyllysine, was developed for *in vitro* biochemical settings.

Another broadly used method to access methyllysine-containing proteins relies on amber-suppressor-mediated incorporation of caged methyllysine precursors, followed by chemical removal of the caging moieties (Figure 45). To prepare recombinant proteins containing site-specific mononmethyllysine, the Schultz laboratory and the Liu laboratory documented *N*-allyloxycarbonyl(308) and *N*-(*ortho*-nitro)benzyl methyllysine,(309) respectively, as unnatural amino acid building blocks (Figure 45). With the aid of a pyrrolysine(Pyl)-based genetic code expansion system, the two caged methyllysine building blocks can be incorporated into proteins through site-specific nonsense suppressors. Methyllysine can also be masked by a carboxybenzyl (Cbz)(310) or a *tert*-butyloxycarbonyl (Boc) group.(311, 312) The resultant methyllysine is readily unmasked upon transition-metal or acid-triggered decaging reactions (Figure 45). In conjunction of developing natural amino acid caging strategies, propargyloxycarboxyl,(313) derivatized (*ortho*-nitro)benzyl,(314) ortho-azidobenzyloxycarbonyl(315) and 1-(*trans*-cyclooctene)oxycarbonyl(316) moieties have been documented as lysine caging reagents.(317) The resultant lysine derivatives can be readily recognized by corresponding engineered Pyl-tRNA synthetase (PylRS)-tRNA systems as natural amino acids for protein biosynthesis. The lysine decaging can be achieved through transition-metal-, UV light-, phosphine- and tetrazine-based bioorthogonal cleavage chemistry, respectively. These lysine-caging strategies are expected to be transferrable for the site-specific incorporation of caged monomethyllysine (Kme1) into proteins (Figure 45). More importantly, the metal-, photo- and small-molecule-induced lysine decaging chemistry is bioorthogonal and biocompatible, and can be combined with amber-suppressor-mediated translation to afford Kme1-containing proteins inside living cells.(317) The bioorthogonal cleavage chemistry also provides the feasibility to release the caged lysine residues of functional importance and interrogate the associated biology with a temporal control. However, the lysine caging strategies can only be applied to install Kme1. To prepare proteins containing dimethyllysine with the amber-suppressor system, the site-specific dimethylation of target proteins has to be carried out in the context of globally protected lysine residues.(312) Recently, the Li laboratory reported a lysine analog with a photo-labile

diazirine appended at the γ-position of lysine (Figure 45). Given the similarity between the diazirine lysine analog and free lysine, this lysine analog is caged as an unnatural amino acid and then can be processed by orthogonally engineered tRNA/tRNA synthetase pairs for amber-suppressor-mediated protein synthesis.(318) The peptide containing the diazirine methyllysine analog can render photo-induced cross-linking reactivity to covalently capture their binding partners *in vitro* and inside living cells.(318, 319)

9.3. Chemical ligation.

In comparison with cysteine-specific chemical conjugation and nonsense suppressor-mediated protein biosynthesis, chemical ligation features its ability to assemble a target protein containing multiple types of lysine methylation.(27) Native chemical ligation (NCL) and expressed protein ligation (EPL) are two chemical ligation methods to study protein post-translational modifications including methylation (Figure 46).(320) A key transformation of chemical ligation involves the coupling reaction between the C-terminal thioester of a peptide and an N-terminal cysteine-containing peptide (Figure 46). The residual cysteine, if unnecessary, can be converted into alanine through desulfurization.(320) A key difference between NCL and EPL is that the latter involves an intein to generate a recombinant thioester intermediate and then catalyze the subsequent ligation reaction.

With the aid of NCL, the Muir laboratory is able to semi-synthesize nucleosomes containing more than 50 combinations of methylation with other posttranslational modifications.(321, 322) These nucleosomes can be individually DNA-barcoded and used as baits to dissect their recognition partners by their effector proteins. With NCL strategies, the Danishefsky laboratory chemically synthesized several biologically important glycosylated proteins including erythropoietin and granulocyte-macrophage colony-stimulating factor.(323, 324) Such efforts can be also readily transferred to chemically synthesize methyllysine-containing proteins with modest complexity. To apply the expressed protein ligation (EPL) inside living cells, the Muir laboratory recently reported protein semisynthesis using ultrafast *trans*-splicing inteins (Figure 46).(325–327) Here they genetically fused the carboxylate end of the N-terminal fragment of a target protein product with the N-terminal fragment of ultrafast split intein. The complementary C-terminal fragment of the ultrafast split intein can be chemically synthesized and fused with the C-terminal fragment of the target protein product as well as a cell-penetrating peptide. The cell-penetrating peptide drives the cellular uptake of the fused C-terminal intein and is cleaved intracellularly. The N-terminal intein and the C-terminal intein then form the functional splicing complex to ligate the N-terminal and C-terminal fragment to afford the target protein within living cells. Similar to NCL, installation of methyllysine can be envisioned at the N-terminal fragment through a nonsense suppressor or at the C-terminal fragment through chemical synthesis. One limitation of NCL, EPL and *trans*-splicing intein ligation is that these reactions occur at a cysteine site or at sites with a cysteine as the precursor for desulfurization. Such a requirement may not be fulfilled for all protein targets but can be avoided potentially with subtiligase as reported recently by the Cole laboratory (Figure 46).(328) Subtiligase is an engineered peptide ligase derived from the protease subtilisin through mutagenesis, which renders peptidase activity toward aminolysis and thus facilitates the coupling of a peptide containing activated C-terminal esters with a peptide containing an N-terminal α-amine moiety. In addition, it is more

convenient to implement NCL, EPL and *trans*-splicing intein ligation to install methyllysine at N-terminal or C-terminal region rather than in the central region of target proteins.

10. Summary and Perspective

In this review, lysine and three types of methyllysine residues were discussed in great details along with their distinct use as protein building blocks. Conversion of lysine into methyllysine in proteins is catalyzed by PKMTs via a stepwise process consisting of recognition of the SAM cofactor and diverse substrates, Lys deprotonation, and then assembling of a S_N2 transition state. PKMTs have been well tuned to catalyze the methylation reaction in a context dependent manner. One of the key characters of methyllysine, unlike other types of posttranslationally modified lysine, is the minimal change of charge and size in contrast to an unmodified lysine. Therefore, methyllysine reader proteins rely on multiple mechanisms to exploit the difference between free lysine and Kme1/2/3. This ability to distinguish lysine and Kme1/2/3, though modest, can be amplified by multivalent interactions between reader proteins and multiple posttranslational modifications. Like many other posttranslational modifications, protein lysine methylation is reversible and can be removed mainly by two families of demethylases via oxidative mechanisms. However, our knowledge of biological roles of protein lysine methylation is largely limited by the difficulty in determining unambiguously context-dependent substrate profiles and downstream functions of PKMTs. This review covers the recent progress of developing chemical tools to address some of these challenges. However, there is still a great demand for developing innovative and integrated approaches to examine protein lysine methylation. While chemical properties of protein lysine methylation make it technically challenging to probe methyllysine-associated biology, these distinct chemical properties allow the avenues to examine protein methylation that may not be feasible for other types of posttranslational modifications. It is well accepted that a protein does not exist in a rigid or static conformation under a native setting but rather as an ensemble of enormous conformational states in a dynamic equilibrium. Most of our understanding of protein lysine methylation was derived from static structures of its writers, readers and erasers. We anticipated more relevant understanding of protein lysine methylation in the context of dynamic conformational landscapes of PKMTs. This review aims to highlight these opportunities as well as the associated challenges. Importantly, these insights also rely on the basic understanding of protein lysine methylation from a chemical and biochemical perspective. As an emerging field with important biological relevance, our understanding of protein methylation is expected to be facilitated by integrated approaches including chemical methods.

ACKNOWLEDGEMENTS

The author thanks members in the Luo laboratory for their helpful comments on this manuscript, Miss Minaya for editing references, Dr. Zhen Wang for providing the electrostatic potential surface images in Figure 2, the National Institute of General Medical Sciences of the National Institutes of Health of USA (R01GM096056, R01GM120570, R01GM109760), National Cancer Institute (5P30 CA008748-44), Mr. William H. Goodwin and Mrs. Alice Goodwin Commonwealth Foundation for Cancer Research, and the Experimental Therapeutics Center of Memorial Sloan Kettering Cancer Center for financial supports.

Biography

Minkui Luo received his B.S. degree in organic chemistry at Fudan University. He then worked with Dr. John T. Groves at Princeton University and received his Ph.D. in bioorganic chemistry and chemical biology. Thereafter, Dr. Luo worked as a postdoctoral fellow with Dr. Vern Schramm at Albert Einstein College of Medicine. In 2008, Dr. Luo established his independent laboratory at Memorial Sloan Kettering Cancer Center. There he leveraged his background at the interface of chemistry and biology and mainly focuses on combining chemical and biological tools to define, perturb and manipulate the functions of protein methylation. His contribution in the field of chemical biology and protein methylation has been recognized by multiple awards including NIH New Innovator Award, V Scholar Award, Alfred W. Bressler Scholar Award, Basil O'Connor Starter Scholar Award, and Eli Lilly Award.

ABBREVIATIONS

2-OG
2-oxoglutarate

53BP1
p53-binding protein 1

Ab-SAM
4-azidobut-2-enyl SAM analog

AcdK
N^{ε}-(4-azidobenzoxycarbonyl)-δ,ε-dehydrolysine

AF10
ALL1-fused gene from chromosome 10 protein

AHL(s)
acyl-homoserine lactone(s)

Ala or A
L-alanine

AlK
L-allysine

AOL
amine oxidase like

Arg or R
L-arginine

ASH1L
scAsh1 like histone lysine methyltransferase

Ash2L
Ash2-like protein

Asn or N
L-asparagine

Asp or D
L-aspartic acid

ATP
adenosine triphosphate

ATRX
α-thalassemia mental retardation X-linked protein

BAH
bromo adjacent homology

BHC80
BRAF35-HDAC complex protein

BIE(s)
binding isotope effect(s)

Boc
tert-butyloxycarbonyl

BPPM
Bioorthogonal Profiling Protein Methylation

BPTF
bromodomain and PHD finger-containing transcription factor

BRPF1
bromodomain and PHD finger-containing protein 1

C/EBP-β
CCAAT-enhancer-binding protein beta

CaM-KMT
calmodulin-lysine *N*-methyltransferase

CARM1
coactivator associated arginine methyltransferase 1

Cbz
carboxybenzyl

CHD1
chromodomain helicase DNA binding protein 1

COMPASS
complex associated with Set1

Cps40/Ypl138
complex proteins associated with SET1 protein SPP1/Ypl138 gene product

CRISPR
Clustered Regularly Interspaced Short Palindromic Repeats

CTSA
cellular thermal shift assay

Cys or C
L-cysteine

DBHDA
2,5-dibromohexanediamide

Dha
dehydroalanine

DNA
2′-deoxyribonucleic acid

DNMT(s)
DNA methyltransferase(s)

DNMT1/3A/3B/3L
DNA methyltransferase 1/3A/3B/3L

DOT1L
disruptor of telomeric silencing 1-like

EPL
expressed protein ligation

Dpy30
protein dpy-30 homolog

E2F1
E2F transcription factor 1

EED
embryonic ectoderm development

eEF1A
eukaryotic translation elongation factor 1 alpha

eEF1A-KMT1
eEF1A lysine methyltransferase 1 alpha

eEF2-KMT
eukaryotic elongation factor 2 lysine methyltransferase

Efm5
PKMT that trimethylates elongation factor 1-alpha at Lys79

EnYn-SAM
(*E*)-pent-2-en-4-ynyl SAM analog

ERα
estrogen receptor alpha

ETFβ
electron transfer flavoprotein β

EZH1
enhancer of zeste homologue 1

EZH2
enhancer of zeste 2 polycomb repressive complex 2

F or Phe
L-phenylalanine

FAD
flavin adenine dinucleotide

FDAS
5′-fluorodeoxyadenosine synthase

FOXO3
forkhead box protein O3

GLP1
G9a-like protein 1

Gln or Q
L-glutamine

Glu or E
L-glutamic acid

Gly or G
L-glycine

H3
Histone 3

H3K4
Histone 3 lysine 4

H3K4me3
H3K4 trimethylation

H3K9
Histone 3 lysine 9

H3K9me3
H3K9 trimethylation

H3K27
Histone 3 lysine 27

H3K27me3
H3K27 trimethylation

H3K36
histone 3 lysine 36

H3K36M
histone 3 K36M mutant

H3K36MNuc
nucleosome containing H3K36MNuc

H4
histone 4

H4K20
histone 4 lysine 20

H4K20me1/2/3
H4K20 mono/di-tri-methylation

Hey-SAM
(*E*)-hex-2-en-5-ynyl-SAM

His or H
L-histidine

hMSH6
human MutS protein homolog 6

^{1}H-NMR
proton nuclear magnetic resonance

HP1
heterochromatin protein 1

HSP70/90/A1/A5/A8
heat shock protein 70/90/A1/A5/A8

ING2
inhibitor of growth 2

JmjC
Jumonji C

JMJD2A
Jumonji domain-containing protein 2A

K
L-Lysine

Kme1/2/3
mono/di/tri-methylated lysine

KDM2A/4A
lysine demethylase 2A/4A

KDM(s)
histone demethylase(s)

KIE(s)
kinetic isotope effect(s)

L3MBT1/3
lethal(3) malignant brain tumor 1/3

LOXL2
lysyl oxidase-like 2

LRWD1
leucine-rich repeat and WD repeat-containing protein 1

LSD1/2
lysine specific demethylase 1/2

Lys
L-Lysine

MALDI-TOF
matrix-assisted laser desorption ionization time-of-flight

MAT(s)
S-methionine adenosyltransferase(s)

MAP3K2
mitogen-activated protein kinase kinase kinase 2

MBT
malignant brain tumor

METTL10/20/21A/21B/21C/21D/22
methyltransferase-like 10/20/21A/21B/21C/21D/22

MINT
Msx2-interacting protein

MLA(s)
methyllysine analogs

MLL1/2/3/4
mixed lineage leukemia protein 1/2/3/4

MOA
mode of action

MOI
mode of interaction

mRNA
messager RNA

MSH
O-mesitylenesulfonylhydroxylamine

MSL3
male-specific lethal 3

MS
mass spectrometry

MYND
myeloid translocation protein-8, Nervy, and DEAF-1

N6AMT2
N-6 adenine-specific DNA methyltransferase 2 (putative)

$NaBH_4$
sodium borohydride

$NaCNBH_3$
sodium cyanoborohydride

NCL
native chemical ligation

NSD1/2/3
nuclear receptor-binding SET domain-containing protein 1

NSM
Nonsense-suppression mutagenesis

NURF
nucleosome remodelling factor

ORC1
origin recognition complex subunit 1

PCNA
proliferating cell nuclear antigen

PHD
plant homeodomain

Phe or F
L-phenylalanine

PHF1/2
PHD finger protein 1/2

PKMT(s)
protein lysine methyltransferase(s)

Pob-SAM
4-propargyloxy-but-2-enyl-SAM

ProSeAM
propargylic *Se*-adenosyl-L-selenomethionine

ppm
parts-per-million (10^{-6})

PRMT(s)
protein arginine methyltransferase(s)

PRMT1-11
protein arginine methyltransferase 1-11

PRC2
Polycomb repressive complex 2

PRDM
PR (PRDI-BF1 and RIZ) domain containing

Pygo-HD1-BCL9
pygo/histone deacetylase 1/B-cell lymphoma 9 protein complex

Pyl
L-pyrrolysine

PylRS
Pyl-tRNA synthetase

QM
quantum mechanical

QM/MM
quantum mechanical/molecular mechanics

Rb
retinoblastoma protein

RbBP5
RB binding protein 5

RNA
ribonucleic acid

RORα
RAR-related orphan receptor alpha

SAM
S-5′-adenosyl-L-methionine

SAH
S-5′-adenosyl-L-homocysteine

SAHH
SAH hydrolase

SalL
gene L for the biosynthesis of salinosporamide A

SecPh
L-phenylselenocysteine

Sep
O-phosphoserine

Ser or S
L-serine

Set1
SET domain-containing protein 1

SET
Su(var)3-9, Enhancer-of-zeste and Trithorax

SETD2
SET domain containing 2

SETD8
SET domain containing (lysine methyltransferase) 8

SET7/9
SET domain containing (lysine methyltransferase) 7

SETDB1
SET domain, bifurcated 1

SGF29
SAGA-associated factor 29

SHH1
SAWADEE homeodomain homolog-1

SIRT1
sirtuin (silent mating type information regulation 2 homolog) 1

SMN
survival of motor neuron

SMYD
SET and MYND domain containing

SMYD1-5
SET and MYND domain containing 1-5

SNF
sinefungin

SPF30
splicing factor spf30

STAT3
signal transducer and activator of transcription 3

Stk31
Serine/threonine-protein kinase 31

SUV39H1
suppressor of variegation 3–9 homologue 1

SUV420H1/2
suppressor of variegation 4-20 homolog 1/2

SUZ12
suppressor of zeste 12

SWI/SNF
switch/sucrose nonfermentable

SWIRM
small α-helical domain

TAF3/7/10
TATA-box binding protein associated factor 3/7/10

TDRD3
tudor domain-containing protein 3

Thr or T
L-threonine

tRNA
transfer RNA

Trp or W
L-tryptophan

Tyr or Y
L-tyrosine

UHRF1
Ubiquitin-like-containing PHD and RING finger domains protein 1

UTX
ubiquitously transcribed tetratricopeptide repeat X-chromosome protein

VCP-KMT
valosin-containing protein lysine methyltransferase

WDR5
WD repeat-containing protein 5

WRAD
WDR5-RbBP5-Ash2L-Dpy30 complex

Yap
Yes-associated protein

zf-CW
zinc finger CW

ZCWPW1
zinc finger CW type with PWWP domain 1

ZMET2
zea methyltransferase2

ZMYND11
zinc finger MYND domain-containing protein 11

REFERENCES

(1). Reid MA, Dai Z, and Locasale JW The Impact of Cellular Metabolism on Chromatin Dynamics and Epigenetics. Nat. Cell Biol. 2017, 19, 1298–1306. [PubMed: 29058720]

(2). Allis CD, and Jenuwein T The Molecular Hallmarks of Epigenetic Control. Nat. Rev. Genet 2016, 17, 487–500. [PubMed: 27346641]

(3). Jakovcevski M, and Akbarian S Epigenetic Mechanisms in Neurological Disease. Nat. Med 2012, 18, 1194–1204. [PubMed: 22869198]

(4). Han S, and Brunet A Histone Methylation Makes its Mark on Longevity. Trends Cell Biol. 2012, 22, 42–49. [PubMed: 22177962]

(5). Flavahan WA, Gaskell E, and Bernstein BE Epigenetic Plasticity and the Hallmarks of Cancer. Science 2017, 357, eaal2380. [PubMed: 28729483]

(6). Kouzarides T Chromatin Modifications and their Function. Cell 2007, 128, 693–705. [PubMed: 17320507]

(7). Lyko F The DNA Methyltransferase Family: a Versatile Toolkit for Epigenetic Regulation. Nat. Rev. Genet 2017, 19, 81–92. [PubMed: 29033456]

(8). Tolstykh T, Lee J, Vafai S, and Stock JB Carboxyl Methylation Regulates Phosphoprotein Phosphatase 2A by Controlling the Association of Regulatory B Subunits. EMBO J. 2000, 19, 5682–5691. [PubMed: 11060019]

(9). Clarke S Aging as War between Chemical and Biochemical Processes: Protein Methylation and the Recognition of Age-damaged Proteins for Repair. Ageing Res. Rev. 2003, 2, 263–285. [PubMed: 12726775]

(10). Winter-Vann AM, and Casey PJ Post-prenylation-processing Enzymes as New Targets in Oncogenesis. Nat. Rev. Cancer 2005, 5, 405–412. [PubMed: 15864282]

(11). Clarke SG Protein Methylation at the Surface and Buried Deep: Thinking Outside the Histone Box. Trends Biochem. Sci. 2013, 38, 243–252. [PubMed: 23490039]

(12). Biggar KK, and Li SS Non-histone Protein Methylation as a Regulator of Cellular Signalling and Function. Nat. Rev. Mol. Cell Biol. 2015, 16, 5–17. [PubMed: 25491103]

(13). Klotz AV, Thomas BA, Glazer AN, and Blacher RW Detection of Methylated Asparagine and Glutamine Residues in Polypeptides. Anal. Biochem 1990, 186, 95–100. [PubMed: 2356973]

(14). Sprung R, Chen Y, Zhang K, Cheng D, Zhang T, Peng J, and Zhao Y Identification and Validation of Eukaryotic Aspartate and Glutamate Methylation in Proteins. J. Proteome Res. 2008, 7, 1001–1006. [PubMed: 18220335]

(15). Blanc RS, and Richard S Arginine Methylation: the Coming of Age. Mol. Cell 2017, 65, 8–24. [PubMed: 28061334]

(16). Nimura K, Ura K, and Kaneda Y Histone Methyltransferases: Regulation of Transcription and Contribution to Human Disease. J. Mol. Med 2010, 88, 1213–1220. [PubMed: 20714703]

(17). Zhao BS, Roundtree IA, and He C Post-transcriptional Gene Regulation by mRNA Modifications. Nat. Rev. Mol. Cell Biol. 2017, 18, 31–42. [PubMed: 27808276]

(18). Fontecave M, Atta M, and Mulliez E *S*-Adenosylmethionine: Nothing Goes to Waste. Trends Biochem. Sci. 2004, 29, 243–249. [PubMed: 15130560]

(19). Wang R, Zheng W, and Luo M A Sensitive Mass Spectrum Assay to Characterize Engineered Methionine Adenosyltransferases with *S*-Alkyl Methionine Analogues as Substrates. Anal. Biochem 2014, 450, 11–19. [PubMed: 24374249]

(20). Klimasauskas S, and Weinhold E A New Tool for Biotechnology: AdoMet-dependent Methyltransferases. Trends Biotechnol. 2007, 25, 99–104. [PubMed: 17254657]

(21). Wang R, and Luo M A Journey toward Bioorthogonal Profiling of Protein Methylation inside Living Cells. Curr. Opin. Chem. Biol 2013, 17, 729–737. [PubMed: 24035694]

(22). Thomsen M, Vogensen SB, Buchardt J, Burkart MD, and Clausen RP Chemoenzymatic Synthesis and *in situ* Application of *S*-Adenosyl-L-Methionine Analogs. Org. Biomol. Chem 2013, 11, 7606–7610. [PubMed: 24100405]

(23). Booker SJ Radical SAM Enzymes and Radical Enzymology. Biochim. Biophys. Acta 2012, 1824, 1151–1153. [PubMed: 22850428]

(24). Lin H *S*3-Adenosylmethionine-Dependent Alkylation Reactions: When are Radical Reactions Used? Bioorg. Chem 2011, 39, 161–170. [PubMed: 21762947]

(25). Geske GD, O'Neill JC, and Blackwell HE Expanding Dialogues: From Natural Autoinducers to Non-Natural Analogues that Modulate Quorum Sensing in Gram-negative Bacteria. Chem. Soc. Rev 2008, 37, 1432–1447. [PubMed: 18568169]

(26). Gutierrez JA, Luo M, Singh V, Li L, Brown RL, Norris GE, Evans GB, Furneaux RH, Tyler PC, Painter GF, et al. Picomolar Inhibitors as Transition-State Probes Of 5′-Methylthioadenosine Nucleosidases. ACS Chem. Biol. 2007, 2, 725–734. [PubMed: 18030989]

(27). Luo M Current Chemical Biology Approaches to Interrogate Protein Methyltransferases. ACS Chem. Biol. 2012, 7, 443–463. [PubMed: 22220966]

(28). Richon VM, Johnston D, Sneeringer CJ, Jin L, Majer CR, Elliston K, Jerva LF, Scott MP, and Copeland RA Chemogenetic Analysis of Human Protein Methyltransferases. Chem. Biol. Drug. Des 2011, 78, 199–210. [PubMed: 21564555]

(29). Copeland RA, Solomon ME, and Richon VM Protein Methyltransferases as a Target Class for Drug Discovery. Nat. Rev. Drug Discov. 2009, 8, 724–732. [PubMed: 19721445]

(30). Andrews FH, Strahl BD, and Kutateladze TG Insights into Newly Discovered Marks and Readers of Epigenetic Information. Nat. Chem. Biol 2016, 12, 662–668. [PubMed: 27538025]

(31). Wagner T, Robaa D, Sippl W, and Jung M Mind the Methyl: Methyllysine Binding Proteins in Epigenetic Regulation. Chemmedchem 2014, 9, 466–483. [PubMed: 24449612]

(32). Musselman CA, Lalonde ME, Cote J, and Kutateladze TG Perceiving the Epigenetic Landscape through Histone Readers. Nat. Struct. Mol. Biol 2012, 19, 1218–1227. [PubMed: 23211769]

(33). Kaniskan HU, Martini ML, and Jin J Inhibitors of Protein Methyltransferases and Demethylases. Chem. Rev 2018, 118, 989–1068. [PubMed: 28338320]

(34). Mcallister TE, England KS, Hopkinson RJ, Brennan PE, Kawamura A, and Schofield CJ Recent Progress in Histone Demethylase Inhibitors. J. Med. Chem 2016, 59, 1308–13029. [PubMed: 26710088]

(35). Thinnes CC, England KS, Kawamura A, Chowdhury R, Schofield CJ, and Hopkinson RJ Targeting Histone Lysine Demethylases - Progress, Challenges, and the Future. Biochim. Biophys. Acta 2014, 1839, 1416–14132. [PubMed: 24859458]

(36). Helin K, And Dhanak D Chromatin Proteins and Modifications as Drug Targets. Nature 2013, 502, 480–488. [PubMed: 24153301]

(37). Kudithipudi S, and Jeltsch A Approaches and Guidelines for the Identification of Novel Substrates of Protein Lysine Methyltransferases. Cell Chem. Biol. 2016, 23, 1049–1055. [PubMed: 27569752]

(38). Luo M Inhibitors of Protein Methyltransferases as Chemical Tools. Epigenomics 2015, 7, 1327–1338. [PubMed: 26646500]

(39). Schneider R, Bannister AJ, and Kouzarides T Unsafe Sets: Histone Lysine Methyltransferases and Cancer. Trends Biochem. Sci. 2002, 27, 396–402. [PubMed: 12151224]

(40). Kaniskan HU, and Jin J Recent Progress in Developing Selective Inhibitors of Protein Methyltransferases. Curr. Opin. Chem. Biol 2017, 39, 100–108. [PubMed: 28662389]

(41). Egelhofer TA, Minoda A, Klugman S, Lee K, Kolasinska-Zwierz P, Alekseyenko AA, Cheung M-S, Day DS, Gadel S, Gorchakov AA, et al. An Assessment of Histone-modification Antibody Quality. Nat. Struct. Mol. Biol 2011, 18, 91–94. [PubMed: 21131980]

(42). Rothbart SB, Dickson BM, Raab JR, Grzybowski AT, Krajewski K, Guo AH, Shanle EK, Josefowicz SZ, Fuchs SM, Allis CD, et al. An Interactive Database for the Assessment of Histone Antibody Specificity. Mol. Cell 2015, 59, 502–511. [PubMed: 26212453]

(43). Andre I, Linse S, and Mulder FA Residue-specific pKa Determination of Lysine and Arginine Side Chains by Indirect 15N and 13C NMR Spectroscopy: Application to apo Calmodulin. J. Am. Chem. Soc. 2007, 129, 15805–15813. [PubMed: 18044888]

(44). Beaver JE, and Waters ML Molecular Recognition of Lys and Arg Methylation. ACS Chem. Biol. 2016, 11, 643–653. [PubMed: 26759915]

(45). Gallivan JP, and Dougherty DA A Computational Study of Cation-pi Interactions vs Salt Bridges in Aqueous Media: Implications for Protein Engineering. J. Am. Chem. Soc 2000, 122, 870–874.

(46). Gallivan JP, and Dougherty DA Cation-pi Interactions in Structural Biology. Proc. Natl. Acad. Sci. USA 1999, 96, 9459–9464. [PubMed: 10449714]

(47). Seeler JS, and Dejean A SUMO and the Robustness of Cancer. Nat. Rev. Cancer 2017, 17, 184–197. [PubMed: 28134258]

(48). Sabari BR, Zhang D, Allis CD, and Zhao Y Metabolic Regulation of Gene Expression Through Histone Acylations. Nat. Rev. Mol. Cell Biol. 2017, 18, 90–101. [PubMed: 27924077]

(49). Enchev RI, Schulman BA, and Peter M Protein Neddylation: Beyond Cullin-RING Ligases. Nat. Rev. Mol. Cell Biol. 2015, 16, 30–44. [PubMed: 25531226]

(50). Mattiroli F, and Sixma TK Lysine-targeting Specificity in Ubiquitin and Ubiquitin-like Modification Pathways. Nat. Struct. Mol. Biol 2014, 21, 308–316. [PubMed: 24699079]

(51). Choudhary C, Weinert BT, Nishida Y, Verdin E, and Mann M The Growing Landscape of Lysine Acetylation Links Metabolism and Cell Signalling. Nat. Rev. Mol. Cell Biol. 2014, 15, 536–550. [PubMed: 25053359]

(52). Kothapalli N, Camporeale G, Kueh A, Chew YC, Oommen AM, Griffin JB, and Zempleni J Biological Functions of Biotinylated Histones. J. Nutr. Biochem 2005, 16, 446–448. [PubMed: 15992689]

(53). Melcher M, Schmid M, Aagaard L, Selenko P, Laible G, and Jenuwein T Structure-function Analysis of SUV39H1 Reveals a Dominant Role in Heterochromatin Organization, Chromosome Segregation, and Mitotic Progression. Mol. Cell. Biol 2000, 20, 3728–3741. [PubMed: 10779362]

(54). Min J, Feng Q, Li Z, Zhang Y, and Xu RM Structure of the Catalytic Domain of Human DOT1L, a Non-SET Domain Nucleosomal Histone Methyltransferase. Cell 2003, 112, 711–723. [PubMed: 12628190]

(55). Bennett RL, Swaroop A, Troche C, and Licht JD The Role of Nuclear Receptor-binding SET Domain Family Histone Lysine Methyltransferases in Cancer. Cold Spring Harb. Perspect Med. 2017, 7.

(56). Milite C, Feoli A, Viviano M, Rescigno D, Cianciulli A, Balzano AL, Mai A, Castellano S, and Sbardella G The Emerging Role of Lysine Methyltransferase SETD8 in Human Diseases. Clin. Epigenetics 2016, 8, 102. [PubMed: 27688818]

(57). Liu Y, Liu K, Qin S, Xu C, and Min J Epigenetic Targets and Drug Discovery: Part 1: Histone Methylation. Pharmacol. Ther 2014, 143, 275–294. [PubMed: 24704322]

(58). Shimazu T, Barjau J, Sohtome Y, Sodeoka M, and Shinkai Y Selenium-based S-Adenosylmethionine Analog Reveals the Mammalian Seven-β-strand Methyltransferase METTL10 to be an EF1A1 Lysine Methyltransferase. Plos One 2014, 9, E105394. [PubMed: 25144183]

(59). Malecki J, Ho AY, Moen A, Dahl HA, and Falnes PO Human METTL20 is a Mitochondrial Lysine Methyltransferase that Targets the β-Subunit of Electron Transfer Flavoprotein (ETFβ) and Modulates its Activity. J. Biol. Chem 2015, 290, 423–434. [PubMed: 25416781]

(60). Rhein VF, Carroll J, He J, Ding S, Fearnley IM, and Walker JE Human METTL20 Methylates Lysine Residues adjacent to the Recognition Loop of the Electron Transfer Flavoprotein in Mitochondria. J. Biol. Chem 2014, 289, 24640–24651. [PubMed: 25023281]

(61). Jakobsson ME, Moen A, Bousset L, Egge-Jacobsen W, Kernstock S, Melki R, and Falnes PO Identification and Characterization of a Novel Human Methyltransferase Modulating Hsp70 Protein Function through Lysine Methylation. J. Biol. Chem 2013, 288, 27752–27763. [PubMed: 23921388]

(62). Malecki J, Aileni VK, Ho AYY, Schwarz J, Moen A, Sorensen V, Nilges BS, Jakobsson ME, Leidel SA, and Falnes PO The Novel Lysine Specific Methyltransferase METTL21B Affects mRNA Translation through Inducible and Dynamic Methylation of Lys-165 in Human Eukaryotic Elongation Factor 1 Alpha (Eef1a). Nucleic Acids Res. 2017, 45, 4370–4389. [PubMed: 28108655]

(63). Hamey JJ, Wienert B, Quinlan KGR, and Wilkins MR METTL21B is a Novel Human Lysine Methyltransferase of Translation Elongation Factor 1A: Discovery by CRISPR/Cas9 Knockout. Mol. Cell. Proteomics 2017, 16, 2229–2242. [PubMed: 28663172]

(64). Jakobsson ME, Davydova E, Malecki J, Moen A, and Falnes PO Saccharomyces Cerevisiae Eukaryotic Elongation Factor 1A (Eef1a) is Methylated at Lys-390 by a METTL21-Like Methyltransferase. Plos One 2015, 10, E0131426. [PubMed: 26115316]

(65). Kernstock S, Davydova E, Jakobsson M, Moen A, Pettersen S, Maelandsmo GM, Egge-Jacobsen W, and Falnes PO Lysine Methylation of VCP by a Member of a Novel Human Protein Methyltransferase Family. Nat. Commun 2012, 3, 1038. [PubMed: 22948820]

(66). Cloutier P, Lavallee-Adam M, Faubert D, Blanchette M, and Coulombe B A Newly Uncovered Group of Distantly Related Lysine Methyltransferases Preferentially Interact with Molecular Chaperones to Regulate their Activity. Plos Genet. 2013, 9, E1003210. [PubMed: 23349634]

(67). Cloutier P, Lavallee-Adam M, Faubert D, Blanchette M, and Coulombe B Methylation of the DNA/RNA-binding Protein Kin17 By METTL22 Affects its Association with Chromatin. J. Proteomics 2014, 100, 115–124. [PubMed: 24140279]

(68). Hamey JJ, Winter DL, Yagoub D, Overall CM, Hart-Smith G, and Wilkins MR Novel N-terminal and Lysine Methyltransferases that Target Translation Elongation Factor 1A in Yeast and Human. Mol. Cell. Proteomics 2016, 15, 164–176. [PubMed: 26545399]

(69). Davydova E, Ho AY, Malecki J, Moen A, Enserink JM, Jakobsson ME, Loenarz C, and Falnes PO Identification and Characterization of a Novel Evolutionarily Conserved Lysine-specific Methyltransferase Targeting Eukaryotic Translation Elongation Factor 2 (Eef2). J. Biol. Chem 2014, 289, 30499–30510. [PubMed: 25231979]

(70). Magen S, Magnani R, Haziza S, Hershkovitz E, Houtz R, Cambi F, and Parvari R Human Calmodulin Methyltransferase: Expression, Activity on Calmodulin, and Hsp90 Dependence. Plos. One 2012, 7, E52425. [PubMed: 23285036]

(71). Qian C, and Zhou MM SET Domain Protein Lysine Methyltransferases: Structure, Specificity and Catalysis. Cell. Mol. Life Sci. 2006, 63, 2755–2763. [PubMed: 17013555]

(72). Vougiouklakis T, Hamamoto R, Nakamura Y, and Saloura V The NSD Family of Protein Methyltransferases in Human Cancer. Epigenomics 2015, 7, 863–874. [PubMed: 25942451]

(73). Wu H, Min JR, Lunin VV, Antoshenko T, Dombrovski L, Zeng H, Allali-Hassani A, Campagna-Slater V, Vedadi M, Arrowsmith CH, et al. Structural Biology of Human H3K9 Methyltransferases. Plos One 2010, 5, E8570. [PubMed: 20084102]

(74). Tachibana M, Sugimoto K, Fukushima T, and Shinkai Y SET Domain-containing Protein, G9a, is a Novel Lysine-preferring Mammalian Histone Methyltransferase with Hyperactivity and Specific Selectivity to Lysines 9 and 27 of Histone H3. J. Biol. Chem. 2001, 276, 25309–25317. [PubMed: 11316813]

(75). Zheng W, Ibanez G, Wu H, Blum G, Zeng H, Dong A, Li F, Hajian T, Allali-Hassani A, Amaya MF, et al. Sinefungin Derivatives as Inhibitors and Structure Probes of Protein Lysine Methyltransferase SETD2. J. Am. Chem. Soc 2012, 134, 18004–18014. [PubMed: 23043551]

(76). An S, Yeo KJ, Jeon YH, and Song JJ Crystal Structure of the Human Histone Methyltransferase ASH1L Catalytic Domain and its Implications for the Regulatory Mechanism. J. Biol. Chem 2011, 286, 8369–8374. [PubMed: 21239497]

(77). Tisi D, Chiarparin E, Tamanini E, Pathuri P, Coyle JE, Hold A, Holding FP, Amin N, Martin AC, Rich SJ, et al. Structure of the Epigenetic Oncogene MMSET and Inhibition by N-Alkyl Sinefungin Derivatives. ACS Chem. Biol. 2016, 11, 3093–3105. [PubMed: 27571355]

(78). Yang S, Zheng X, Lu C, Li GM, Allis CD, and Li H Molecular Basis for Oncohistone H3 Recognition by SETD2 Methyltransferase. Genes Dev. 2016, 30, 1611–1616. [PubMed: 27474439]

(79). Grebien F, Vedadi M, Getlik M, Giambruno R, Grover A, Avellino R, Skucha A, Vittori S, Kuznetsova E, Smil D, et al. Pharmacological Targeting of the WDR5-MLL Interaction in C/EBPα N-terminal Leukemia. Nat. Chem. Biol 2015, 11, 571–578. [PubMed: 26167872]

(80). Borkin D, He S, Miao H, Kempinska K, Pollock J, Chase J, Purohit T, Malik B, Zhao T, Wang J, et al. Pharmacologic Inhibition of the Menin-MLL Interaction Blocks Progression of MLL Leukemia in vivo. Cancer Cell 2015, 27, 589–602. [PubMed: 25817203]

(81). Cao F, Townsend EC, Karatas H, Xu J, Li L, Lee S, Liu L, Chen Y, Ouillette P, Zhu J, et al. Targeting MLL1 H3K4 Methyltransferase Activity in Mixed-Lineage Leukemia. Mol. Cell 2014, 53, 247–261. [PubMed: 24389101]

(82). He Y, Selvaraju S, Curtin ML, Jakob CG, Zhu H, Comess KM, Shaw B, The J, Lima-Fernandes E, Szewczyk MM, et al. The EED Protein-protein Interaction Inhibitor A-395 Inactivates the PRC2 Complex. Nat. Chem. Biol 2017, 13, 389–395. [PubMed: 28135237]

(83). Kim W, Bird GH, Neff T, Guo G, Kerenyi MA, Walensky LD, And Orkin SH Targeted Disruption of the EZH2-EED Complex Inhibits EZH2-dependent Cancer. Nat. Chem. Biol 2013, 9, 643–650. [PubMed: 23974116]

(84). Mazur PK, Reynoird N, Khatri P, Jansen PW, Wilkinson AW, Liu S, Barbash O, Van Aller GS, Huddleston M, Dhanak D, et al. SMYD3 Links Lysine Methylation of MAP3K2 to Ras-driven Cancer. Nature 2014, 510, 283–287. [PubMed: 24847881]

(85). Sirinupong N, Brunzelle J, Doko E, and Yang Z Structural Insights into the Autoinhibition and Posttranslational Activation of Histone Methyltransferase Smyd3. J. Mol. Biol 2011, 406, 149–159. [PubMed: 21167177]

(86). Ferguson AD, Larsen NA, Howard T, Pollard H, Green I, Grande C, Cheung T, Garcia-Arenas R, Cowen S, Wu J, et al. Structural Basis of Substrate Methylation and Inhibition of SMYD2. Structure 2011, 19, 1262–1273. [PubMed: 21782458]

(87). Sirinupong N, Brunzelle J, Ye J, Pirzada A, Nico L, and Yang Z Crystal Structure of Cardiac-specific Histone Methyltransferase SMYD1 Reveals Unusual Active Site Architecture. J. Biol. Chem 2010, 285.

(88). Saddic LA, West LE, Aslanian A, Yates JR, Rubin SM, Gozani O, and Sage J Methylation of the Retinoblastoma Tumor Suppressor by SMYD2. J. Biol. Chem 2010, 285, 37733–37740. [PubMed: 20870719]

(89). Mzoughi S, Tan YX, Low D, and Guccione E The Role of PRDMs in Cancer: One Family, Two Sides. Curr. Opin. Genet. Dev 2016, 36, 83–91. [PubMed: 27153352]

(90). Justin N, Zhang Y, Tarricone C, Martin SR, Chen S, Underwood E, De Marco V, Haire LF, Walker PA, Reinberg D, et al. Structural Basis of Oncogenic Histone H3K27M Inhibition of Human Polycomb Repressive Complex 2. Nat. Commun 2016, 7, 11316. [PubMed: 27121947]

(91). Li Y, Han J, Zhang Y, Cao F, Liu Z, Li S, Wu J, Hu C, Wang Y, Shuai J, et al. Structural Basis for Activity Regulation of MLL Family Methyltransferases. Nature 2016, 530, 447–452. [PubMed: 26886794]

(92). Hsieh JJ, Cheng EH, and Korsmeyer SJ Taspase1: A Threonine Aspartase Required for Cleavage of MLL and Proper HOX Gene Expression. Cell 2003, 115, 293–303. [PubMed: 14636557]

(93). Brooun A, Gajiwala KS, Deng YL, Liu W, Bolanos B, Bingham P, He YA, Diehl W, Grable N, Kung PP, et al. Polycomb Repressive Complex 2 Structure with Inhibitor Reveals a Mechanism of Activation and Drug Resistance. Nat. Commun 2016, 7, 11384. [PubMed: 27122193]

(94). Jiao L, and Liu X Structural Basis of Histone H3K27 Trimethylation by an Active Polycomb Repressive Complex 2. Science 2015, 350, Aac4383.

(95). Shinsky SA, Hu M, Vought VE, Ng SB, Bamshad MJ, Shendure J, and Cosgrove MS A Non-active-site SET Domain Surface Crucial for the Interaction of MLL1 and the Rbbp5/Ash2L Heterodimer within MLL Family Core Complexes. J. Mol. Biol 2014, 426, 2283–2299. [PubMed: 24680668]

(96). Patel A, Vought VE, Swatkoski S, Viggiano S, Howard B, Dharmarajan V, Monteith KE, Kupakuwana G, Namitz KE, Shinsky SA, et al. Automethylation Activities within the Mixed Lineage Leukemia-1 (MLL1) Core Complex Reveal Evidence Supporting a “Two-Active Site” Model for Multiple Histone H3 Lysine 4 Methylation. J. Biol. Chem. 2014, 289, 868–884. [PubMed: 24235145]

(97). Patel A, Vought VE, Dharmarajan V, and Cosgrove MS A Novel Non-SET Domain Multi-subunit Methyltransferase Required for Sequential Nucleosomal Histone H3 Methylation by the Mixed Lineage Leukemia Protein-1 (MLL1) Core Complex. J. Biol. Chem 2011, 286, 3359–3369. [PubMed: 21106533]

(98). Linscott JA, Kapilashrami K, Wang Z, Senevirathne C, Bothwell IR, Blum G, And Luo M Kinetic Isotope Effects Reveal Early Transition State of Protein Lysine Methyltransferase SET8. Proc. Natl. Acad. Sci. USA 2016, 113, 8369–8378.

(99). Islam K, Chen Y, Wu H, Bothwell IR, Blum GJ, Zeng H, Dong A, Zheng W, Min J, Deng H, et al. Defining Efficient Enzyme-cofactor Pairs for Bioorthogonal Profiling of Protein Methylation. Proc. Natl. Acad. Sci. USA 2013, 110, 16778–16783. [PubMed: 24082136]

(100). Ibáñez G, Mcbean JL, Astudillo YM, and Luo M An Enzyme-coupled Ultrasensitive Luminescence Assay for Protein Methyltransferases. Anal. Biochem. 2010, 401, 203–210. [PubMed: 20227379]

(101). Schapira M Structural Chemistry of Human SET Domain Protein Methyltransferases. Curr. Chem. Genomics 2011, 5, 85–94. [PubMed: 21966348]

(102). Campagna-Slater V, Mok MW, Nguyen KT, Feher M, Najmanovich R, and Schapira M Structural Chemistry of the Histone Methyltransferases Cofactor Binding Site. J. Chem. Inf. Model 2011, 51, 612–623. [PubMed: 21366357]

(103). Wilson JR, Jing C, Walker PA, Martin SR, Howell SA, Blackburn GM, Gamblin SJ, and Xiao B Crystal Structure and Functional Analysis of the Histone Methyltransferase SET7/9. Cell 2002, 111, 105–115. [PubMed: 12372304]

(104). Kwon T, Chang JH, Kwak E, Lee CW, Joachimiak A, Kim YC, Lee JW, and Cho YJ Mechanism of Histone Lysine Methyl Transfer Revealed by the Structure of SET7/9-AdoMet. EMBO J. 2003, 22, 292–303. [PubMed: 12514135]

(105). Xiao B, Jing C, Wilson JR, Walker PA, Vasisht N, Kelly G, Howell S, Taylor IA, Blackburn GM, and Gamblin SJ Structure and Catalytic Mechanism of the Human Histone Methyltransferase SET7/9. Nature 2003, 421, 652–656. [PubMed: 12540855]

(106). Couture JF, Collazo E, Hauk G, and Trievel RC Structural Basis for the Methylation Site Specificity of SET7/9. Nat. Struct. Mol. Biol 2006, 13, 140–146. [PubMed: 16415881]

(107). Fang J, Feng Q, Ketel CS, Wang HB, Cao R, Xia L, Erdjument-Bromage H, Tempst P, Simon JA, and Zhang Y Purification and Functional Characterization of SET8, a Nucleosomal Histone H4-Lysine 20-specific Methyltransferase. Curr. Biol 2002, 12, 1086–1099. [PubMed: 12121615]

(108). Couture JF, Collazo E, Brunzelle JS, and Trievel RC Structural and Functional Analysis of SET8, a Histone H4 Lys-20 Methyltransferase. Genes Dev. 2005, 19, 1455–1465. [PubMed: 15933070]

(109). Xiao B, Jing C, Kelly G, Walker PA, Muskett FW, Frenkiel TA, Martin SR, Sarma K, Reinberg D, Gamblin SJ, et al. Specificity and Mechanism of the Histone Methyltransferase Pr-Set7. Genes Dev. 2005, 19, 1444–1454. [PubMed: 15933069]

(110). Yao Y, Chen P, Diao J, Cheng G, Deng L, Anglin JL, Prasad BVV, and Song Y Selective Inhibitors of Histone Methyltransferase DOT1L: Design, Synthesis, and Crystallographic Studies. J. Am. Chem. Soc 2011, 133, 16746–16749. [PubMed: 21936531]

(111). Poulin MB, Schneck JL, Matico RE, Mcdevitt PJ, Huddleston MJ, Hou W, Johnson NW, Thrall SH, Meek TD, and Schramm VL Transition State for the NSD2-catalyzed Methylation of Histone H3 Lysine 36. Proc. Natl. Acad. Sci. USA 2016, 113, 1197–1201. [PubMed: 26787850]

(112). Poulin MB, Schneck JL, Matico RE, Hou W, Mcdevitt PJ, Holbert M, and Schramm VL Nucleosome Binding Alters the Substrate Bonding Environment of Histone H3 Lysine 36 Methyltransferase NSD2. J. Am. Chem. Soc 2016, 138, 6699–6702. [PubMed: 27183271]

(113). Horowitz S, Yesselman JD, Al-Hashimi HM, and Trievel RC Direct Evidence for Methyl Group Coordination by Carbon-oxygen Hydrogen Bonds in the Lysine Methyltransferase SET7/9. J. Biol. Chem 2011, 286, 18658–18663. [PubMed: 21454678]

(114). Horowitz S, and Trievel RC Carbon-oxygen Hydrogen Bonding in Biological Structure and Function. J. Biol. Chem 2012, 287, 41576–41582. [PubMed: 23048026]

(115). Horowitz S, Dirk LM, Yesselman JD, Nimtz JS, Adhikari U, Mehl RA, Scheiner S, Houtz RL, Al-Hashimi HM, and Trievel RC Conservation and Functional Importance of Carbon-oxygen Hydrogen Bonding in AdoMet-dependent Methyltransferases. J. Am. Chem. Soc 2013, 135, 15536–15548. [PubMed: 24093804]

(116). Horowitz S, Adhikari U, Dirk LM, Del Rizzo PA, Mehl RA, Houtz RL, Al-Hashimi HM, Scheiner S, and Trievel RC Manipulating Unconventional CH-based Hydrogen Bonding in a Methyltransferase via Noncanonical Amino Acid Mutagenesis. ACS Chem. Biol 2014, 9, 1692–1697. [PubMed: 24914947]

(117). Fick RJ, Kroner GM, Nepal B, Magnani R, Horowitz S, Houtz RL, Scheiner S, and Trievel RC Sulfur-oxygen Chalcogen Bonding Mediates AdoMet Recognition in the Lysine Methyltransferase SET7/9. ACS Chem. Biol 2016, 11, 748–754. [PubMed: 26713889]

(118). Mcginty RK, Kohn M, Chatterjee C, Chiang KP, Pratt MR, and Muir TW Structure-activity Analysis of Semisynthetic Nucleosomes: Mechanistic Insights into the Stimulation of Dot1L by Ubiquitylated Histone H2B. ACS Chem. Biol 2009, 4, 958–968. [PubMed: 19799466]

(119). Huang J, Hsu YH, Mo C, Abreu E, Kiel DP, Bonewald LF, Brotto M, and Karasik D METTL21C is a Potential Pleiotropic Gene for Osteoporosis and Sarcopenia Acting through the Modulation of the NFκb Signaling Pathway. J. Bone Miner Res 2014, 29, 1531–1540. [PubMed: 24677265]

(120). Kurash JK, Lei H, Shen Q, Marston WL, Granda BW, Fan H, Wall D, Li E, And Gaudet F Methylation of p53 by Set7/9 Mediates p53 Acetylation and Activity in vivo. Mol. Cell 2008, 29, 392–400. [PubMed: 18280244]

(121). Pradhan S, Chin HG, Esteve PO, and Jacobsen SE SET7/9 Mediated Methylation of Non-histone Proteins in Mammalian Cells. Epigenetics 2009, 4, 282–285.

(122). Kontaki H, and Talianidis I Lysine Methylation Regulates E2F1-induced Cell Death. Mol Cell 2010, 39, 152–160. [PubMed: 20603083]

(123). Dhayalan A, Kudithipudi S, Rathert P, and Jeltsch A Specificity Analysis-based Identification of New Methylation Targets of the SET7/9 Protein Lysine Methyltransferase. Chem. Biol 2011, 18, 111–120. [PubMed: 21276944]

(124). Yang J, Huang J, Dasgupta M, Sears N, Miyagi M, Wang B, Chance MR, Chen X, Du Y, Wang Y, et al. Reversible Methylation of Promoter-bound STAT3 by Histone-modifying Enzymes. Proc. Natl. Acad. Sci. USA 2010, 107, 21499–21504. [PubMed: 21098664]

(125). Gaughan L, Stockley J, Wang N, Mccracken SR, Treumann A, Armstrong K, Shaheen F, Watt K, Mcewan IJ, Wang C, et al. Regulation of the Androgen Receptor by SET9-mediated Methylation. Nucleic Acids Res. 2011, 39, 1266–1279. [PubMed: 20959290]

(126). Ko S, Ahn J, Song CS, Kim S, Knapczyk-Stwora K, and Chatterjee B Lysine Methylation and Functional Modulation of Androgen Receptor by Set9 Methyltransferase. Mol. Endocrinol 2011, 25, 433–444. [PubMed: 21273441]

(127). Barsyte-Lovejoy D, Li F, Oudhoff MJ, Tatlock JH, Dong A, Zeng H, Wu H, Freeman SA, Schapira M, Senisterra GA, et al. (*R*)-PFI-2 is a Potent and Selective Inhibitor of SETD7 Methyltransferase Activity in Cells. Proc. Natl. Acad. Sci. USA 2014, 111, 12853–12858. [PubMed: 25136132]

(128). Liu X, Wang D, Zhao Y, Tu B, Zheng Z, Wang L, Wang H, Gu W, Roeder RG, and Zhu WG Methyltransferase Set7/9 Regulates p53 Activity by Interacting with Sirtuin 1 (SIRT1). Proc. Natl. Acad. Sci. USA 2011, 108, 1925–1930. [PubMed: 21245319]

(129). Xie Q, Hao Y, Tao L, Peng S, Rao C, Chen H, You H, Dong MQ, and Yuan Z Lysine Methylation of FOXO3 Regulates Oxidative Stress-Induced Neuronal Cell Death. EMBO Rep. 2012, 13, 371–377. [PubMed: 22402663]

(130). Calnan DR, Webb AE, White JL, Stowe TR, Goswami T, Shi X, Espejo A, Bedford MT, Gozani O, Gygi SP, et al. Methylation by Set9 Modulates Foxo3 Stability and Transcriptional Activity. Aging 2012, 4, 462–479. [PubMed: 22820736]

(131). Munro S, Khaire N, Inche A, Carr S, and La Thangue NB Lysine Methylation Regulates the pRb Tumour Suppressor Protein. Oncogene 2010, 29, 2357–2367. [PubMed: 20140018]

(132). Yang XD, Huang B, Li MX, Lamb A, Kelleher NL, and Chen LF Negative Regulation of NF-Kappa B Action by Set9-mediated Lysine Methylation of the RelA Subunit. EMBO. J 2009, 28, 1055–1066. [PubMed: 19262565]

(133). Patnaik D, Chin HG, Esteve PO, Benner J, Jacobsen SE, and Pradhan S Substrate Specificity and Kinetic Mechanism of Mammalian G9a Histone H3 Methyltransferase. J. Biol. Chem 2004, 279, 53248–53258. [PubMed: 15485804]

(134). Chin HG, Pradhan M, Esteve PO, Patnaik D, Evans TC Jr., and Pradhan S Sequence Specificity and Role of Proximal Amino Acids of the Histone H3 Tail on Catalysis of Murine G9a Lysine 9 Histone H3 Methyltransferase. Biochemistry 2005, 44, 12998–13006. [PubMed: 16185068]

(135). Chang Y, Sun L, Kokura K, Horton JR, Fukuda M, Espejo A, Izumi V, Koomen JM, Bedford MT, Zhang X, et al. MPP8 Mediates the Interactions Between DNA Methyltransferase Dnmt3a and H3K9 Methyltransferase GLP/G9a. Nat. Commun 2011, 2, 533. [PubMed: 22086334]

(136). Huang J, Dorsey J, Chuikov S, Zhang XY, Jenuwein T, Reinberg D, and Berger SL G9a and Glp Methylate Lysine 373 in the Tumor Suppressor P53. J. Biol. Chem 2010, 285, 9636–9641. [PubMed: 20118233]

(137). Pless O, Kowenz-Leutz E, Knoblich M, Lausen J, Beyermann M, Walsh MJ, and Leutz A G9a-mediated Lysine Methylation Alters the Function of CCAAT/Enhancer-Binding Protein β. J. Biol. Chem 2008, 283, 26357–26363. [PubMed: 18647749]

(138). Lee JS, Kim Y, Kim IS, Kim B, Choi HJ, Lee JM, Shin HJR, Kim JH, Kim JY, Seo SB, et al. Negative Regulation of Hypoxic Responses via Induced Reptin Methylation. Mol. Cell 2010, 39, 71–85. [PubMed: 20603076]

(139). Ling BM, Bharathy N, Chung TK, Kok WK, Li S, Tan YH, Rao VK, Gopinadhan S, Sartorelli V, Walsh MJ, et al. Lysine Methyltransferase G9a Methylates the Transcription Factor MyoD and Regulates Skeletal Muscle Differentiation. Proc. Natl. Acad. Sci. USA 2012, 109, 841–846. [PubMed: 22215600]

(140). Carlson SM, Moore KE, Green EM, Martin GM, and Gozani O Proteome-wide Enrichment of Proteins Modified by Lysine Methylation. Nat. Protoc 2014, 9, 37–50. [PubMed: 24309976]

(141). Carlson SM, and Gozani O Emerging Technologies to Map the Protein Methylome. J. Mol. Biol 2014, 426, 3350–3362. [PubMed: 24805349]

(142). Moore KE, Carlson SM, Camp ND, Cheung P, James RG, Chua KF, Wolf-Yadlin A, and Gozani O A General Molecular Affinity Strategy for Global Detection and Proteomic Analysis of Lysine Methylation. Mol Cell. 2013, 50, 444–456. [PubMed: 23583077]

(143). Islam K, Bothwell I, Chen Y, Sengelaub C, Wang R, Deng H, and Luo M Bioorthogonal Profiling of Protein Methylation Using Azido Derivative of S-Adenosyl-L-methionine. J. Am. Chem. Soc 2012, 134, 5909–5915. [PubMed: 22404544]

(144). Rathert P, Dhayalan A, Murakami M, Zhang X, Tamas R, Jurkowska R, Komatsu Y, Shinkai Y, Cheng XD, and Jeltsch A Protein Lysine Methyltransferase G9a Acts on Non-histone Targets. Nat. Chem. Biol 2008, 4, 344–346. [PubMed: 18438403]

(145). Takawa M, Cho HS, Hayami S, Toyokawa G, Kogure M, Yamane Y, Iwai Y, Maejima K, Ueda K, Masuda A, et al. Histone Lysine Methyltransferase SETD8 Promotes Carcinogenesis by Deregulating PCNA Expression. Cancer Res. 2012, 72, 3217–3227. [PubMed: 22556262]

(146). Dhami GK, Liu H, Galka M, Voss C, Wei R, Muranko K, Kaneko T, Cregan SP, Li L, and Li SS Dynamic Methylation of Numb by Set8 Regulates its Binding to p53 and Apoptosis. Mol. Cell 2013, 50, 565–576. [PubMed: 23706821]

(147). Knutson SK, Wigle TJ, Warholic NM, Sneeringer CJ, Allain CJ, Klaus CR, Sacks JD, Raimondi A, Majer CR, Song J, et al. A Selective Inhibitor of EZH2 Blocks H3K27 Methylation and Kills Mutant Lymphoma Cells. Nat. Chem. Biol 2012, 8, 890–896. [PubMed: 23023262]

(148). Lee JM, Lee JS, Kim H, Kim K, Park H, Kim JY, Lee SH, Kim IS, Kim J, Lee M, et al. EZH2 Generates a Methyl Degron that is Recognized by the DCAF1/DDB1/CUL4 E3 Ubiquitin Ligase Complex. Mol. Cell 2012, 48, 572–786. [PubMed: 23063525]

(149). Kim E, Kim M, Woo DH, Shin Y, Shin J, Chang N, Oh YT, Kim H, Rheey J, Nakano I, et al. Phosphorylation of EZH2 Activates STAT3 Signaling via STAT3 Methylation and Promotes Tumorigenicity of Glioblastoma Stem-like Cells. Cancer Cell 2013, 23, 839–852. [PubMed: 23684459]

(150). Dasgupta M, Dermawan JK, Willard B, and Stark GR STAT3-driven Transcription Depends upon the Dimethylation of K49 by EZH2. Proc. Natl. Acad. Sci. USA 2015, 112, 3985–3990. [PubMed: 25767098]

(151). Clancy KW, Russell AM, Subramanian V, Nguyen H, Qian Y, Campbell RM, and Thompson PR Citrullination/Methylation Crosstalk on Histone H3 Regulates ER-target Gene Transcription. ACS Chem. Biol 2017, 12, 1691–1702. [PubMed: 28485572]

(152). Del Rizzo PA, and Trievel RC Substrate and Product Specificities of SET Domain Methyltransferases. Epigenetics 2011, 6, 1059–1067. [PubMed: 21847010]

(153). Zhang X, Huang Y, and Shi X Emerging Roles of Lysine Methylation on Non-histone Proteins. Cell Mol. Life Sci 2015, 72, 4257–4272. [PubMed: 26227335]

(154). Shi XB, Kachirskaia L, Yamaguchi H, West LE, Wen H, Wang EW, Dutta S, Appella E, and Gozani O Modulation of p53 Function by SET8-mediated Methylation at Lysine 382. Mol. Cell 2007, 27, 636–646. [PubMed: 17707234]

(155). Fu W, Liu N, Qiao Q, Wang M, Min J, Zhu B, Xu RM, and Yang N Structural Basis for Substrate Preference of SMYD3, a SET Domain-containing Protein Lysine Methyltransferase. J. Biol. Chem 2016, 291, 9173–9180. [PubMed: 26929412]

(156). Lanouette S, Davey JA, Elisma F, Ning Z, Figeys D, Chica RA, and Couture JF Discovery of Substrates for a SET Domain Lysine Methyltransferase Predicted by Multistate Computational Protein Design. Structure 2015, 23, 206–215. [PubMed: 25533488]

(157). Boriack-Sjodin PA, Jin L, Jacques SL, Drew A, Sneeringer C, Scott MP, Moyer MP, Ribich S, Moradei O, and Copeland RA Structural Insights into Ternary Complex Formation of Human CARM1 with Various Substrates. ACS Chem. Biol 2016, 11, 763–771. [PubMed: 26551522]

(158). Chatterjee C, Mcginty RK, Fierz B, and Muir TW Disulfide-directed Histone Ubiquitylation Reveals Plasticity in hDOT1L Activation. Nat. Chem. Biol 2010, 6, 267–269. [PubMed: 20208522]

(159). Mcginty RK, Kim J, Chatterjee C, Roeder RG, and Muir TW Chemically Ubiquitylated Histone H2B Stimulates hDOT1L-mediated Intranucleosomal Methylation. Nature 2008, 453, 812–816. [PubMed: 18449190]

(160). Wu H, Siarheyeva A, Zeng H, Lam R, Dong A, Wu XH, Li Y, Schapira M, Vedadi M, and Min J Crystal Structures of the Human Histone H4K20 Methyltransferases SUV420H1 and SUV420H2. FEBS Lett. 2013, 587, 3859–3868. [PubMed: 24396869]

(161). Sneeringer CJ, Scott MP, Kuntz KW, Knutson SK, Pollock RM, Richon VM, and Copeland RA Coordinated Activities of Wild-type plus Mutant EZH2 Drive Tumor-associated Hypertrimethylation of Lysine 27 on Histone H3 (H3K27) in Human B-Cell Lymphomas. Proc. Natl. Acad. Sci. USA 2010, 107, 20980–20985. [PubMed: 21078963]

(162). Wigle TJ, Knutson SK, Jin L, Kuntz KW, Pollock RM, Richon VM, Copeland RA, and Scott MP The Y641C Mutation of EZH2 Alters Substrate Specificity for Histone H3 Lysine 27 Methylation States. FEBS Lett. 2011, 585, 3011–3014. [PubMed: 21856302]

(163). Majer CR, Jin L, Scott MP, Knutson SK, Kuntz KW, Keilhack H, Smith JJ, Moyer MP, Richon VM, Copeland RA, et al. A687V EZH2 is a Gain-of-function Mutation Found in Lymphoma Patients. FEBS Lett. 2012, 586, 3448–3451. [PubMed: 22850114]

(164). Mccabe MT, Graves AP, Ganji G, Diaz E, Halsey WS, Jiang Y, Smitheman KN, Ott HM, Pappalardi MB, Allen KE,et al. Mutation of A677 in Histone Methyltransferase EZH2 in Human B-Cell Lymphoma Promotes Hypertrimethylation of Histone H3 on Lysine 27 (H3K27). Proc. Natl. Acad. Sci. USA 2012, 109, 2989–2994. [PubMed: 22323599]

(165). Ott HM, Graves AP, Pappalardi MB, Huddleston M, Halsey WS, Hughes AM, Groy A, Dul E, Jiang Y, Bai Y, et al. A687V EZH2 is a Driver of Histone H3 Lysine 27 (H3K27) Hypertrimethylation. Mol. Cancer Ther 2014, 13, 3062–3073. [PubMed: 25253781]

(166). Trievel RC, Beach BM, Dirk LMA, Houtz RL, and Hurley JH Structure and Catalytic Mechanism of a SET Domain Protein Methyltransferase. Cell 2002, 111, 91–103. [PubMed: 12372303]

(167). Zhang X, and Bruice TC Histone Lysine Methyltransferase SET7/9: Formation of a Water Channel Precedes Each Methyl Transfer. Biochemistry 2007, 46, 14838–14844. [PubMed: 18044969]

(168). Zhang X, and Bruice TC Product Specificity and Mechanism of Protein Lysine Methyltransferases: Insights from the Histone Lysine Methyltransferase SET8. Biochemistry 2008, 47, 6671–6677. [PubMed: 18512960]

(169). Zhang X, and Bruice TC Enzymatic Mechanism and Product Specificity of SET-domain Protein Lysine Methyltransferases. Proc. Natl. Acad. Sci. USA 2008, 105, 5728–5732. [PubMed: 18391193]

(170). Zhang X, and Bruice TC Mechanism of Product Specificity of Adomet Methylation Catalyzed by Lysine Methyltransferases: Transcriptional Factor p53 Methylation by Histone Lysine Methyltransferase SET7/9. Biochemistry 2008, 47, 2743–2748. [PubMed: 18260647]

(171). Hu P, and Zhang Y Catalytic Mechanism and Product Specificity of the Histone Lysine Methyltransferase SET7/9: an ab initio QM/MM-FE Study with Multiple Initial Structures. J. Am. Chem. Soc 2006, 128, 1272–1278. [PubMed: 16433545]

(172). Guo HB, and Guo H Mechanism of Histone Methylation Catalyzed by Protein Lysine Methyltransferase SET7/9 and Origin of Product Specificity. Proc. Natl. Acad. Sci. USA 2007, 104, 8797–8802. [PubMed: 17517655]

(173). Wang S, Hu P, and Zhang Y Ab initio Quantum Mechanical/Molecular Mechanical Molecular Dynamics Simulation of Enzyme Catalysis: the Case of Histone Lysine Methyltransferase SET7/9. J. Phys. Chem. B 2007, 111, 3758–3764. [PubMed: 17388541]

(174). Hu P, Wang S, and Zhang Y How do SET-Domain Protein Lysine Methyltransferases Achieve the Methylation State Specificity? Revisited by ab initio QM/MM Molecular Dynamics Simulations. J. Am. Chem. Soc 2008, 130, 3806–3813. [PubMed: 18311969]

(175). Yao J, Chu Y, An R, and Guo H Understanding Product Specificity of Protein Lysine Methyltransferases from QM/MM Molecular Dynamics and Free Energy Simulations: the Effects of Mutation on SET7/9 beyond the Tyr/Phe Switch. J. Chem. Inf. Model 2012, 52, 449–456. [PubMed: 22242964]

(176). Schramm VL Enzymatic Transition State Theory and Transition State Analogue Design. J. Biol. Chem 2007, 282, 28297–28300. [PubMed: 17690091]

(177). Zhang J, and Klinman JP Enzymatic Methyl Transfer: Role of an Active Site Residue in Generating Active Site Compaction that Correlates with Catalytic Efficiency. J. Am. Chem. Soc 2011, 133, 17134–17137. [PubMed: 21958159]

(178). Zhang J, and Klinman JP Convergent Mechanistic Features Between the Structurally Diverse N- and O-Methyltransferases: Glycine N-Methyltransferase and Catechol O-Methyltransferase. J. Am. Chem. Soc 2016, 138, 9158–9165. [PubMed: 27355841]

(179). Wilson PB, and Williams IH Influence of Equatorial CHO Interactions on Secondary Kinetic Isotope Effects for Methyl Transfer. Angew. Chem. Int. Ed 2016, 55, 3192–3195.

(180). Schotta G, Sengupta R, Kubicek S, Malin S, Kauer M, Callen E, Celeste A, Pagani M, Opravil S, De La Rosa-Velazquez IA, et al. A Chromatin-wide Transition to H4K20 Monomethylation Impairs Genome Integrity and Programmed DNA Rearrangements in the Mouse. Genes Dev. 2008, 22, 2048–2061. [PubMed: 18676810]

(181). Couture JF, Dirk LMA, Brunzelle JS, Houtz RL, and Trievel RC Structural Origins for the Product Specificity of SET Domain Protein Methyltransferases. Proc. Natl. Acad. Sci. USA 2008, 105, 20659–20664. [PubMed: 19088188]

(182). Collins RE, Tachibana M, Tamaru H, Smith KM, Jia D, Zhang X, Selker EU, Shinkai Y, and Cheng XD In vitro and in vivo Analyses of a Phe/Tyr Switch Controlling Product Specificity of Histone Lysine Methyltransferases. J. Biol. Chem 2005, 280, 5563–5570. [PubMed: 15590646]

(183). Del Rizzo PA, Couture JF, Dirk LMA, Strunk BS, Roiko MS, Brunzelle JS, Houtz RL, and Trievel RC SET7/9 Catalytic Mutants Reveal the Role of Active Site Water Molecules in Lysine Multiple Methylation. J. Biol. Chem 2010, 285, 31849–31858. [PubMed: 20675860]

(184). Bai Q, Shen Y, Yao X, Wang F, Du Y, Wang Q, Jin N, Hai J, Hu T, and Yang J Modeling a New Water Channel that Allows SET9 to Dimethylate p53. Plos One 2011, 6, E19856. [PubMed: 21625555]

(185). Herz HM, Hu D, and Shilatifard A Enhancer Malfunction in Cancer. Mol Cell 2014, 53, 859–666. [PubMed: 24656127]

(186). Southall SM, Wong PS, Odho Z, Roe SM, and Wilson JR Structural Basis for the Requirement of Additional Factors for MLL1 SET Domain Activity and Recognition of Epigenetic Marks. Mol. Cell 2009, 33, 181–191. [PubMed: 19187761]

(187). Thornton JL, Westfield GH, Takahashi YH, Cook M, Gao X, Woodfin AR, Lee JS, Morgan MA, Jackson J, Smith ER, et al. Context Dependency of Set1/COMPASS-mediated Histone H3 Lys4 Trimethylation. Genes Dev. 2014, 28, 115–120. [PubMed: 24402317]

(188). Takahashi YH, Lee JS, Swanson SK, Saraf A, Florens L, Washburn MP, Trievel RC, and Shilatifard A Regulation of H3K4 Trimethylation via Cps40 (Spp1) of COMPASS is Monoubiquitination Independent: Implication for a Phe/Tyr Switch by the Catalytic Domain of Set1. Mol. Cell Biol. 2009, 29, 3478–3486. [PubMed: 19398585]

(189). Schneider J, Wood A, Lee JS, Schuster R, Dueker J, Maguire C, Swanson SK, Florens L, Washburn MP, and Shilatifard A Molecular Regulation of Histone H3 Trimethylation by COMPASS and the Regulation of Gene Expression. Mol. Cell 2005, 19, 849–856. [PubMed: 16168379]

(190). Deshpande AJ, Deshpande A, Sinha AU, Chen L, Chang J, Cihan A, Fazio M, Chen CW, Zhu N, Koche R, et al. AF10 Regulates Progressive H3K79 Methylation and HOX Gene Expression in Diverse AML Subtypes. Cancer Cell 2014, 26, 896–908. [PubMed: 25464900]

(191). Wang R, Ibanez G, Islam K, Zheng W, Blum G, Sengelaub C, and Luo M Formulating a Fluorogenic Assay to Evaluate S-Adenosyl-L-Methionine Analogues as Protein Methyltransferase Cofactors. Mol. Biosyst 2011, 7, 2970–2981. [PubMed: 21866297]

(192). Collazo E, Couture JF, Bulfer S, and Trievel RC A Coupled Fluorescent Assay for Histone Methyltransferases. Anal. Biochem 2005, 342, 86–92. [PubMed: 15958184]

(193). Chen DH, Wu KT, Hung CJ, Hsieh M, and Li C Effects of Adenosine Dialdehyde Treatment on in vitro and in vivo Stable Protein Methylation in Hela Cells. J. Biochem 2004, 136, 371–376. [PubMed: 15598895]

(194). Kraus D, Yang Q, Kong D, Banks AS, Zhang L, Rodgers JT, Pirinen E, Pulinilkunnil TC, Gong F, Wang YC, et al. Nicotinamide N-Methyltransferase Knockdown Protects against Diet-Induced Obesity. Nature 2014, 508, 258–262. [PubMed: 24717514]

(195). Ulanovskaya OA, Zuhl AM, and Cravatt BF NNMT Promotes Epigenetic Remodeling in Cancer by Creating a Metabolic Methylation Sink. Nat. Chem. Biol 2013, 9, 300–306. [PubMed: 23455543]

(196). Mentch SJ, Mehrmohamadi M, Huang L, Liu X, Gupta D, Mattocks D, Gomez Padilla P, Ables G, Bamman MM, Thalacker-Mercer AE, et al. Histone Methylation Dynamics and Gene Regulation Occur through the Sensing of One-carbon Metabolism. Cell Metab. 2015, 22, 861–873. [PubMed: 26411344]

(197). Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, et al. DNMT3L Connects Unmethylated Lysine 4 of Histone H3 to de novo Methylation of DNA. Nature 2007, 448, 714–717. [PubMed: 17687327]

(198). Collins RE, Northrop JP, Horton JR, Lee DY, Zhang X, Stallcup MR, and Cheng X The Ankyrin Repeats of G9a and GLP Histone Methyltransferases are Mono- and Dimethyllysine Binding Modules. Nat. Struct. Mol. Biol 2008, 15, 245–250. [PubMed: 18264113]

(199). Kuo AJ, Song J, Cheung P, Ishibe-Murakami S, Yamazoe S, Chen JK, Patel DJ, and Gozani O The BAH Domain of ORC1 Links H4K20me2 to DNA Replication Licensing and Meier-Gorlin Syndrome. Nature 2012, 484, 115–119. [PubMed: 22398447]

(200). Kim D, Blus BJ, Chandra V, Huang P, Rastinejad F, and Khorasanizadeh S Corecognition of DNA and a Methylated Histone Tail by the MSL3 Chromodomain. Nat. Struct. Mol. Biol 2010, 17, 1027–1029. [PubMed: 20657587]

(201). Jacobs SA, and Khorasanizadeh S Structure of HP1 Chromodomain Bound to a Lysine 9-methylated Histone H3 Tail. Science 2002, 295, 2080–2083. [PubMed: 11859155]

(202). Flanagan JF, Mi LZ, Chruszcz M, Cymborowski M, Clines KL, Kim Y, Minor W, Rastinejad F, and Khorasanizadeh S Double Chromodomains Cooperate to Recognize the Methylated Histone H3 Tail. Nature 2005, 438, 1181–1185. [PubMed: 16372014]

(203). Piazza I, Rutkowska A, Ori A, Walczak M, Metz J, Pelechano V, Beck M, and Haering CH Association of Condensin with Chromosomes Depends on DNA Binding by its HEAT-repeat Subunits. Nat. Struct. Mol. Biol 2014, 21, 560–568. [PubMed: 24837193]

(204). Min J, Allali-Hassani A, Nady N, Qi C, Ouyang H, Liu Y, Mackenzie F, Vedadi M, and Arrowsmith CH L3MBTL1 Recognition of Mono-and Dimethylated Histones. Nat. Struct. Mol. Biol 2007, 14, 1229–1230. [PubMed: 18026117]

(205). Roy S, Musselman CA, Kachirskaia I, Hayashi R, Glass KC, Nix JC, Gozani O, Appella E, and Kutateladze TG Structural Insight into p53 Recognition by the 53BP1 Tandem Tudor Domain. J. Mol. Biol 2010, 398, 489–496. [PubMed: 20307547]

(206). Kungulovski G, Mauser R, Reinhardt R, and Jeltsch A Application of Recombinant TAF3 PHD Domain Instead of anti-H3K4me3 Antibody. Epigenetics Chromatin 2016, 9, 11. [PubMed: 27006701]

(207). Li H, Ilin S, Wang W, Duncan EM, Wysocka J, Allis CD, and Patel DJ Molecular Basis for Site-specific Readout of Histone H3K4me3 by the BPTF PHD Finger of NURF. Nature 2006, 442, 91–95. [PubMed: 16728978]

(208). Pena PV, Davrazou F, Shi X, Walter KL, Verkhusha VV, Gozani O, Zhao R, and Kutateladze TG Molecular Mechanism of Histone H3K4me3 Recognition by Plant Homeodomain of ING2. Nature 2006, 442, 100–103. [PubMed: 16728977]

(209). Lan F, Collins RE, De Cegli R, Alpatov R, Horton JR, Shi X, Gozani O, Cheng X, and Shi Y Recognition of Unmethylated Histone H3 Lysine 4 Links BHC80 to LSD1-mediated Gene Repression. Nature 2007, 448, 718–722. [PubMed: 17687328]

(210). Law JA, Du J, Hale CJ, Feng S, Krajewski K, Palanca AM, Strahl BD, Patel DJ, and Jacobsen SE Polymerase IV Occupancy at RNA-directed DNA Methylation Sites Requires SHH1. Nature 2013, 498, 385–389. [PubMed: 23636332]

(211). Musselman CA, Avvakumov N, Watanabe R, Abraham CG, Lalonde ME, Hong Z, Allen C, Roy S, Nunez JK, Nickoloff J, et al. Molecular Basis for H3K36me3 Recognition by the Tudor Domain of PHF1. Nat. Struct. Mol. Biol 2012, 19, 1266–1272. [PubMed: 23142980]

(212). Min J, Zhang Y, and Xu RM Structural Basis for Specific Binding of Polycomb Chromodomain to Histone H3 Methylated at Lys 27. Genes Dev. 2003, 17, 1823–1828. [PubMed: 12897052]

(213). Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ 3rd, Voigt P, Martin SR, Taylor WR, De Marco V, et al. Role of the Polycomb Protein EED in the Propagation of Repressive Histone Marks. Nature 2009, 461, 762–767. [PubMed: 19767730]

(214). He F, Umehara T, Saito K, Harada T, Watanabe S, Yabuki T, Kigawa T, Takahashi M, Kuwasako K, Tsuda K, et al. Structural Insight into the Zinc Finger CW Domain as a Histone Modification Reader. Structure 2010, 18, 1127–11239. [PubMed: 20826339]

(215). Musselman CA, Khorasanizadeh S, and Kutateladze TG Towards Understanding Methyllysine Readout. Biochim. Biophys. Acta 2014, 1839, 686–693. [PubMed: 24727128]

(216). Zhao Q, Rank G, Tan YT, Li H, Moritz RL, Simpson RJ, Cerruti L, Curtis DJ, Patel DJ, Allis CD, et al. PRMT5-mediated Methylation of Histone H4R3 Recruits DNMT3A, Coupling Histone and DNA Methylation in Gene Silencing. Nat. Struct. Mol. Biol 2009, 16, 304–311. [PubMed: 19234465]

(217). Cote J, and Richard S Tudor Domains Bind Symmetrical Dimethylated Arginines. J. Biol. Chem 2005, 280, 28476–28483. [PubMed: 15955813]

(218). Liu K, Guo Y, Liu H, Bian C, Lam R, Liu Y, Mackenzie F, Rojas LA, Reinberg D, Bedford MT, et al. Crystal Structure of TDRD3 and Methyl-Arginine Binding Characterization of TDRD3, SMN and SPF30. Plos One 2012, 7, E30375. [PubMed: 22363433]

(219). Sikorsky T, Hobor F, Krizanova E, Pasulka J, Kubicek K, and Stefl R Recognition of Asymmetrically Dimethylated Arginine by TDRD3. Nucleic Acids Res. 2012, 40, 11748–11755. [PubMed: 23066109]

(220). Yang Y, Lu Y, Espejo A, Wu J, Xu W, Liang S, and Bedford MT TDRD3 is an Effector Molecule for Arginine-methylated Histone Marks. Mol. Cell 2010, 40, 1016–1123. [PubMed: 21172665]

(221). Iberg AN, Espejo A, Cheng D, Kim D, Michaud-Levesque J, Richard S, and Bedford MT Arginine Methylation of the Histone H3 Tail Impedes Effector Binding. J. Biol. Chem 2008, 283, 3006–3610. [PubMed: 18077460]

(222). Li H, Fischle W, Wang W, Duncan EM, Liang L, Murakami-Ishibe S, Allis CD, and Patel DJ Structural Basis for Lower Lysine Methylation State-specific Readout by MBT Repeats of L3MBTL1 and an Engineered PHD Finger. Mol. Cell 2007, 28, 677–691. [PubMed: 18042461]

(223). Trojer P, Li G, Sims RJ, Vaquero A, Kalakonda N, Boccuni P, Lee D, Erdjument-Bromage H, Tempst P, Nimer SD, et al. L3MBTL1, a Histone-methylation-dependent Chromatin Lock. Cell 2007, 129, 915–928. [PubMed: 17540172]

(224). Gao C, Herold JM, Kireev D, Wigle T, Norris JL, and Frye S Biophysical Probes Reveal a "Compromise" Nature of the Methyl-Lysine Binding Pocket in L3MBTL1. J. Am. Chem. Soc 2011, 133, 5357–5362. [PubMed: 21428286]

(225). Kamps JJ, Huang J, Poater J, Xu C, Pieters BJ, Dong A, Min J, Sherman W, Beuming T, Matthias Bickelhaupt F, et al. Chemical Basis for the Recognition of Trimethyllysine by Epigenetic Reader Proteins. Nat. Commun 2015, 6, 8911. [PubMed: 26578293]

(226). Wen H, Li Y, Xi Y, Jiang S, Stratton S, Peng D, Tanaka K, Ren Y, Xia Z, Wu J, et al. ZMYND11 Links Histone H3.3K36me3 to Transcription Elongation and Tumour Suppression. Nature 2014, 508, 263–268. [PubMed: 24590075]

(227). Iwase S, Xiang B, Ghosh S, Ren T, Lewis PW, Cochrane JC, Allis CD, Picketts DJ, Patel DJ, Li H, et al. ATRX ADD Domain Links an Atypical Histone Methylation Recognition Mechanism to Human Mental-retardation Syndrome. Nat. Struct. Mol. Biol 2011, 18, 769–776. [PubMed: 21666679]

(228). Bian C, Xu C, Ruan J, Lee KK, Burke TL, Tempel W, Barsyte D, Li J, Wu M, Zhou BO, et al. Sgf29 Binds Histone H3K4me2/3 and is Required for SAGA Complex Recruitment and Histone H3 Acetylation. EMBO J. 2011, 30, 2829–28242. [PubMed: 21685874]

(229). Arita K, Isogai S, Oda T, Unoki M, Sugita K, Sekiyama N, Kuwata K, Hamamoto R, Tochio H, Sato M, et al. Recognition of Modification Status on a Histone H3 Tail by Linked Histone Reader Modules of the Epigenetic Regulator UHRF1. Proc. Natl. Acad. Sci. USA 2012, 109, 12950–12955. [PubMed: 22837395]

(230). Vezzoli A, Bonadies N, Allen MD, Freund SM, Santiveri CM, Kvinlaug BT, Huntly BJ, Gottgens B, and Bycroft M Molecular Basis of Histone H3K36me3 Recognition by the PWWP Domain of Brpf1. Nat. Struct. Mol. Biol 2010, 17, 617–619. [PubMed: 20400950]

(231). Patel DJ, and Wang Z Readout of Epigenetic Modifications. Annu. Rev. Biochem 2013, 82, 81–118. [PubMed: 23642229]

(232). Kilic S, Bachmann AL, Bryan LC, and Fierz B Multivalency Governs HP1α Association Dynamics with the Silent Chromatin State. Nat. Commun 2015, 6, 7313. [PubMed: 26084584]

(233). Ruthenburg AJ, Li H, Patel DJ, and Allis CD Multivalent Engagement of Chromatin Modifications by Linked Binding Modules. Nat. Rev. Mol. Cell Biol 2007, 8, 983–994. [PubMed: 18037899]

(234). Du J, Zhong X, Bernatavichute YV, Stroud H, Feng S, Caro E, Vashisht AA, Terragni J, Chin HG, Tu A, et al. Dual Binding of Chromomethylase Domains to H3K9me2-containing Nucleosomes Directs DNA Methylation in Plants. Cell 2012, 151, 167–180. [PubMed: 23021223]

(235). Fasting C, Schalley CA, Weber M, Seitz O, Hecht S, Koksch B, Dernedde J, Graf C, Knapp EW, and Haag R Multivalency as a Chemical Organization and Action Principle. Angew. Chem. Int. Ed 2012, 51, 10472–10498.

(236). Rao RC, and Dou Y Hijacked in Cancer: the KMT2 (MLL) Family of Methyltransferases. Nat. Rev. Cancer 2015, 15, 334–346. [PubMed: 25998713]

(237). Stavropoulos P, Blobel G, and Hoelz A Crystal Structure and Mechanism of Human Lysine-specific Demethylase-1. Nat. Struct. Mol. Biol 2006, 13, 626–632. [PubMed: 16799558]

(238). Chen F, Yang H, Dong Z, Fang J, Wang P, Zhu T, Gong W, Fang R, Shi YG, Li Z, et al. Structural Insight into Substrate Recognition by Histone Demethylase LSD2/KDM1b. Cell Res. 2013, 23, 306–309. [PubMed: 23357850]

(239). Nowak RP, Tumber A, Johansson C, Che KH, Brennan P, Owen D, and Oppermann U Advances and Challenges in Understanding Histone Demethylase Biology. Curr. Opin. Chem. Biol 2016, 33, 151–159. [PubMed: 27371875]

(240). Ng SS, Kavanagh KL, Mcdonough MA, Butler D, Pilka ES, Lienard BM, Bray JE, Savitsky P, Gileadi O, Von Delft F, et al. Crystal Structures of Histone Demethylase JMJD2A Reveal Basis for Substrate Specificity. Nature 2007, 448, 87–91. [PubMed: 17589501]

(241). Cole PA Chemical Probes for Histone-modifying Enzymes. Nat. Chem. Biol 2008, 4, 590–597. [PubMed: 18800048]

(242). Zaragoza JPT, Yosca TH, Siegler MA, Moenne-Loccoz P, Green MT, and Goldberg DP Direct Observation of Oxygen Rebound with an Iron-hydroxide Complex. J. Am. Chem. Soc 2017, 139, 13640–13643. [PubMed: 28930448]

(243). Black JC, Van Rechem C, and Whetstine JR Histone Lysine Methylation Dynamics: Establishment, Regulation, and Biological Impact. Mol. Cell 2012, 48, 491–507. [PubMed: 23200123]

(244). Krishnan S, Horowitz S, and Trievel RC Structure and Function of Histone H3 Lysine 9 Methyltransferases and Demethylases. Chembiochem 2011, 12, 254–263. [PubMed: 21243713]

(245). Horton JR, Upadhyay AK, Hashimoto H, Zhang X, and Cheng X Structural Basis for Human PHF2 Jumonji Domain Interaction with Metal Ions. J. Mol. Biol 2011, 406, 1–8. [PubMed: 21167174]

(246). Herranz N, Dave N, Millanes-Romero A, Pascual-Reguant L, Morey L, Diaz VM, Lorenz-Fonfria V, Gutierrez-Gallego R, Jeronimo C, Iturbide A, et al. Lysyl Oxidase-like 2 (LOXL2) Oxidizes Trimethylated Lysine 4 in Histone H3. FEBS J. 2016, 283, 4263–4273. [PubMed: 27735137]

(247). Accari SL, and Fisher PR Emerging Roles of Jmjc Domain-containing Proteins. Int. Rev. Cell Mol. Biol 2015, 319, 165–220. [PubMed: 26404469]

(248). Nicholson TB, and Chen T LSD1 Demethylates Histone and Non-histone Proteins. Epigenetics 2009, 4, 129–132. [PubMed: 19395867]

(249). Zhang X, Tanaka K, Yan J, Li J, Peng D, Jiang Y, Yang Z, Barton MC, Wen H, and Shi X Regulation of Estrogen Receptor α by Histone Methyltransferase SMYD2-Mediated Protein Methylation. Proc. Natl. Acad. Sci. USA 2013, 110, 17284–17289. [PubMed: 24101509]

(250). Abu-Farha M, Lanouette S, Elisma F, Tremblay V, Butson J, Figeys D, and Couture JF Proteomic Analyses of the SMYD Family Interactomes Identify HSP90 as a Novel Target for SMYD2. J. Mol. Cell Biol 2011, 3, 301–308. [PubMed: 22028380]

(251). Wu CY, Hsieh CY, Huang KE, Chang C, and Kang HY Cryptotanshinone Down-regulates Androgen Receptor Signaling by Modulating Lysine-specific Demethylase 1 Function. Int. J. Cancer 2012, 131, 1423–1434. [PubMed: 22052438]

(252). Bremang M, Cuomo A, Agresta AM, Stugiewicz M, Spadotto V, and Bonaldi T Mass Spectrometry-based Identification and Characterisation of Lysine and Arginine Methylation in the Human Proteome. Mol. Biosyst 2013, 9, 2231–2247. [PubMed: 23748837]

(253). Lott K, Li J, Fisk JC, Wang H, Aletta JM, Qu J, and Read LK Global Proteomic Analysis in Trypanosomes Reveals Unique Proteins and Conserved Cellular Processes Impacted by Arginine Methylation. J. Proteomics 2013, 91, 210–225. [PubMed: 23872088]

(254). Guo A, Gu H, Zhou J, Mulhern D, Wang Y, Lee KA, Yang V, Aguiar M, Kornhauser J, Jia X, et al. Immunoaffinity Enrichment and Mass Spectrometry Analysis of Protein Methylation. Mol. Cell Proteomics 2014, 13, 372–387. [PubMed: 24129315]

(255). Liu H, Galka M, Mori E, Liu X, Lin YF, Wei R, Pittock P, Voss C, Dhami G, Li X, et al. A Method for Systematic Mapping of Protein Lysine Methylation Identifies Functions for HP1β in DNA Damage Response. Mol. Cell 2013, 50, 723–735. [PubMed: 23707759]

(256). Afjehi-Sadat L, and Garcia BA Comprehending Dynamic Protein Methylation with Mass Spectrometry. Curr. Opin. Chem. Biol 2013, 17, 12–19. [PubMed: 23333572]

(257). Ong SE, Mittler G, and Mann M Identifying and Quantifying in vivo Methylation Sites by Heavy Methyl SILAC. Nat. Methods 2004, 1, 119–126. [PubMed: 15782174]

(258). Cheng D, Vemulapalli V, and Bedford MT Methods Applied to the Study of Protein Arginine Methylation. Methods Enzymol. 2012, 512, 71–92. [PubMed: 22910203]

(259). Bothwell IR, Islam K, Chen Y, Zheng W, Blum G, Deng H, and Luo M Se-Adenosyl-L-selenomethionine Cofactor Analogue as a Reporter of Protein Methylation. J. Am. Chem. Soc 2012, 134, 14905–14912. [PubMed: 22917021]

(260). Bothwell IR, and Luo M Large-Scale, Protection-Free Synthesis of Se-Adenosyl-L-Selenomethionine Analogues and their Application as Cofactor Surrogates of Methyltransferases. Org. Lett 2014, 16, 3056–3059. [PubMed: 24852128]

(261). Garcia BA, Mollah S, Ueberheide BM, Busby SA, Muratore TL, Shabanowitz J, and Hunt DF Chemical Derivatization of Histones for Facilitated Analysis by Mass Spectrometry. Nat. Protoc 2007, 2, 933–938. [PubMed: 17446892]

(262). Plazas-Mayorca MD, Zee BM, Young NL, Fingerman IM, Leroy G, Briggs SD, and Garcia BA One-pot Shotgun Quantitative Mass Spectrometry Characterization of Histones. J. Proteome Res 2009, 8, 5367–5374. [PubMed: 19764812]

(263). Blair LP, Avaritt NL, Huang R, Cole PA, Taverna SD, and Tackett AJ Masssquirm: An Assay for Quantitative Measurement of Lysine Demethylase Activity. Epigenetics 2011, 6, 490–469. [PubMed: 21273814]

(264). Peters W, Willnow S, Duisken M, Kleine H, Macherey T, Duncan KE, Litchfield DW, Luscher B, and Weinhold E Enzymatic Site-specific Functionalization of Protein Methyltransferase Substrates with Alkynes for Click Labeling. Angew. Chem. Int. Ed 2010, 49, 5170–5173.

(265). Wang R, Zheng W, Yu H, Deng H, and Luo M Labeling Substrates of Protein Arginine Methyltransferase with Engineered Enzymes and Matched S-Adenosyl-L-Methionine Analogues. J. Am. Chem. Soc 2011, 133, 7648–7651. [PubMed: 21539310]

(266). Willnow S, Martin M, Luscher B, and Weinhold E A Selenium-based Click AdoMet Analogue for Versatile Substrate Labeling with Wild-type Protein Methyltransferases. Chembiochem 2012, 13, 1167–1173. [PubMed: 22549896]

(267). Guo H, Wang R, Zheng W, Chen Y, Blum G, Deng H, and Luo M Profiling Substrates of Protein Arginine N-Methyltransferase 3 with S-Adenosyl-L-Methionine Analogues. ACS Chem. Biol 2014, 9, 476–484. [PubMed: 24320160]

(268). Blum G, Islam K, and Luo M Bioorthogonal Profiling of Protein Methylation (BPPM) Using an Azido Analog of S-Adenosyl-L-Methionine. Curr. Protoc. Chem. Biol 2013, 5, 45–66. [PubMed: 23667794]

(269). Blum G, Bothwell IR, Islam K, and Luo M Profiling Protein Methylation with Cofactor Analog Containing Terminal Alkyne Functionality. Curr. Protoc. Chem. Biol 2013, 5, 67–88. [PubMed: 23788324]

(270). Wang R, Islam K, Liu Y, Zheng W, Tang H, Lailler N, Blum G, Deng H, and Luo M Profiling Genome-wide Chromatin Methylation with Engineered Posttranslation Apparatus within Living Cells. J. Am. Chem. Soc 2013, 135, 1048–1056. [PubMed: 23244065]

(271). Cao XJ, Arnaudo AM, and Garcia BA Large-scale Global Identification of Protein Lysine Methylation in vivo. Epigenetics 2013, 8, 477–485. [PubMed: 23644510]

(272). Nguyen H, Allali-Hassani A, Antonysamy S, Chang S, Chen LH, Curtis C, Emtage S, Fan L, Gheyi T, Li F, et al. LLY-507, a Cell-Active, Potent, and Selective Inhibitor of Protein Lysine Methyltransferase SMYD2. J. Biol. Chem 2015, 290, 13641–13653. [PubMed: 25825497]

(273). Kaniskan HU, and Jin J Chemical Probes of Histone Lysine Methyltransferases. ACS Chem. Biol 2014, 10, 40–50. [PubMed: 25423077]

(274). Frye SV The Art of the Chemical Probe. Nat. Chem. Biol 2010, 6, 159–161. [PubMed: 20154659]

(275). Daigle SR, Olhava EJ, Therkelsen CA, Majer CR, Sneeringer CJ, Song J, Johnston LD, Scott MP, Smith JJ, Xiao Y, et al. Selective Killing of Mixed Lineage Leukemia Cells by a Potent Small-Molecule DOT1L Inhibitor. Cancer Cell 2011, 20, 53–65. [PubMed: 21741596]

(276). Basavapathruni A, Jin L, Daigle SR, Majer CR, Therkelsen CA, Wigle TJ, Kuntz KW, Chesworth R, Pollock RM, Scott MP, et al. Conformational Adaptation Drives Potent, Selective and Durable Inhibition of the Human Protein Methyltransferase DOT1L. Chem. Biol. Drug. Des 2012, 80, 971–980. [PubMed: 22978415]

(277). Van Aller GS, Pappalardi MB, Ott HM, Diaz E, Brandt M, Schwartz BJ, Miller WH, Dhanak D, Mccabe MT, Verma SK, et al. Long Residence Time Inhibition of EZH2 in Activated Polycomb Repressive Complex 2. ACS Chem. Biol 2013, 9, 622–629. [PubMed: 24304166]

(278). Mccabe MT, Ott HM, Ganji G, Korenchuk S, Thompson C, Van Aller GS, Liu Y, Graves AP, Della Pietra A III, Diaz E, et al. EZH2 Inhibition as a Therapeutic Strategy for Lymphoma with EZH2-Activating Mutations. Nature 2012, 492, 108–112. [PubMed: 23051747]

(279). Konze KD, Ma A, Li F, Barsyte-Lovejoy D, Parton T, Macnevin CJ, Liu F, Gao C, Huang XP, Kuznetsova E, et al. An Orally Bioavailable Chemical Probe of the Lysine Methyltransferases EZH2 and EZH1. ACS Chem. Biol 2013, 8, 1324–1334. [PubMed: 23614352]

(280). Chen C, Zhu H, Stauffer F, Caravatti G, Vollmer S, Machauer R, Holzer P, Mobitz H, Scheufler C, Klumpp M, Tiedt R, et al. Discovery of Novel DOT1L Inhibitors Through a Structure-based Fragmentation Approach. ACS Med. Chem. Lett 2016, 7, 735–740. [PubMed: 27563395]

(281). Scheufler C, Mobitz H, Gaul C, Ragot C, Be C, Fernandez C, Beyer KS, Tiedt R, and Stauffer F Optimization of a Fragment-based Screening Hit toward Potent DOT1L Inhibitors Interacting in an Induced Binding Pocket. ACS Med. Chem. Lett 2016, 7, 730–734. [PubMed: 27563394]

(282). Mobitz H, Machauer R, Holzer P, Vaupel A, Stauffer F, Ragot C, Caravatti G, Scheufler C, Fernandez C, Hommel U, et al. Discovery of Potent, Selective, and Structurally Novel DOT1L Inhibitors by a Fragment Linking Approach. ACS Med. Chem. Lett 2017, 8, 338–343. [PubMed: 28337327]

(283). Lee GM, and Craik CS Trapping Moving Targets with Small Molecules. Science 2009, 324, 213–215. [PubMed: 19359579]

(284). Qi W, Zhao K, Gu J, Huang Y, Wang Y, Zhang H, Zhang M, Zhang J, Yu Z, Li L, et al. An Allosteric PRC2 Inhibitor Targeting the H3K27me3 Binding Pocket of EED. Nat. Chem. Biol 2017, 13, 381–388. [PubMed: 28135235]

(285). Liu F, Barsyte-Lovejoy D, Li F, Xiong Y, Korboukh V, Huang XP, Allali-Hassani A, Janzen WP, Roth BL, Frye SV, et al. Discovery of an in vivo Chemical Probe of the Lysine Methyltransferases G9a and GLP. J. Med. Chem 2013, 56, 8931–8942. [PubMed: 24102134]

(286). Sweis RF, Wang Z, Algire M, Arrowsmith CH, Brown PJ, Chiang GG, Guo J, Jakob CG, Kennedy S, Li F, et al. Discovery of A-893, a New Cell-Active Benzoxazinone Inhibitor of Lysine Methyltransferase SMYD2. ACS Med. Chem. Lett 2015, 6, 695–700. [PubMed: 26101576]

(287). Bromberg KD, Mitchell TR, Upadhyay AK, Jakob CG, Jhala MA, Comess KM, Lasko LM, Li C, Tuzon CT, Dai Y, et al. The SUV4-20 Inhibitor A-196 Verifies a Role for Epigenetics in Genomic Integrity. Nat. Chem. Biol 2017, 13, 317–324. [PubMed: 28114273]

(288). Mitchell LH, Boriack-Sjodin PA, Smith S, Thomenius M, Rioux N, Munchhof M, Mills JE, Klaus C, Totman J, Riera TV, et al. Novel Oxindole Sulfonamides and Sulfamides: EPZ031686, the First Orally Bioavailable Small Molecule SMYD3 Inhibitor. ACS Med. Chem. Lett 2016, 7, 134–138. [PubMed: 26985287]

(289). Van Aller GS, Graves AP, Elkins PA, Bonnette WG, Mcdevitt PJ, Zappacosta F, Annan RS, Dean TW, Su DS, Carpenter CL, et al. Structure-based Design of a Novel SMYD3 Inhibitor that Bridges the SAM- and MEKK2-binding Pockets. Structure 2016, 24, 774–781. [PubMed: 27066749]

(290). Grembecka J, He S, Shi A, Purohit T, Muntean AG, Sorenson RJ, Showalter HD, Murai MJ, Belcher AM, Hartley T, et al. Menin-MLL Inhibitors Reverse Oncogenic Activity of MLL Fusion Proteins in Leukemia. Nat. Chem. Biol 2012, 8, 277–284. [PubMed: 22286128]

(291). Ma A, Yu W, Xiong Y, Butler KV, Brown PJ, and Jin J Structure-activity relationship Studies of SETD8 Inhibitors. Medchemcomm 2014, 5, 1892–1898. [PubMed: 25554733]

(292). Butler KV, Ma A, Yu W, Li F, Tempel W, Babault N, Pittella-Silva F, Shao J, Wang J, Luo M, et al. Structure-based Design of a Covalent Inhibitor of the SET Domain-containing Protein Lysine Methyltransferase SETD8. J. Med. Chem 2016, 59, 9881–9889. [PubMed: 27804297]

(293). Blum G, Ibanez G, Rao X, Shum D, Radu C, Djaballah H, Rice JC, and Luo M Small-molecule Inhibitors of SETD8 with Cellular Activity. ACS Chem. Biol 2014, 9, 2471–2478. [PubMed: 25137013]

(294). Lafave LM, Beguelin W, Koche R, Teater M, Spitzer B, Chramiec A, Papalexi E, Keller MD, Hricik T, Konstantinoff K, et al. Loss of BAP1 Function Leads to EZH2-dependent Transformation. Nat. Med 2015, 21, 1344–1349. [PubMed: 26437366]

(295). Milite C, Feoli A, Viviano M, Rescigno D, Mai A, Castellano S, and Sbardella G Progress in the Development of Lysine Methyltransferase SETD8 Inhibitors. Chemmedchem 2016, 11, 1680–1685. [PubMed: 27411844]

(296). Osborne T, Roska RLW, Rajski SR, and Thompson PR In situ Generation of a Bisubstrate Analogue for Protein Arginine Methyltransferase 1. J. Am. Chem. Soc 2008, 130, 4574–4575. [PubMed: 18338885]

(297). Jafari R, Almqvist H, Axelsson H, Ignatushchenko M, Lundback T, Nordlund P, and Martinez Molina D The Cellular Thermal Shift Assay for Evaluating Drug Target Interactions in Cells. Nat. Protoc 2014, 9, 2100–2122. [PubMed: 25101824]

(298). Simon MD, Chu F, Racki LR, De La Cruz CC, Burlingame AL, Panning B, Narlikar GJ, and Shokat KM The Site-specific Installation of Methyl-Lysine Analogs into Recombinant Histones. Cell 2007, 128, 1003–1012. [PubMed: 17350582]

(299). Jia G, Wang W, Li H, Mao Z, Cai G, Sun J, Wu H, Xu M, Yang P, Yuan W, et al. A Systematic Evaluation of the Compatibility of Histones Containing Methyl-Lysine Analogues with Biochemical Reactions. Cell Res. 2009, 19, 1217–1220. [PubMed: 19770845]

(300). Chen Z, Notti RQ, Ueberheide B, and Ruthenburg AJ Quantitative and Structural Assessment of Histone Methyllysine Analogue Engagement by Cognate Binding Proteins Reveals Affinity Decrements Relative to those of Native Counterparts. Biochemistry 2018, 57, 300–304. [PubMed: 29111671]

(301). Wright TH, Bower BJ, Chalker JM, Bernardes GJ, Wiewiora R, Ng WL, Raj R, Faulkner S, Vallee MR, Phanumartwiwath A, et al. Posttranslational Mutagenesis: a Chemical Strategy for Exploring Protein Side-chain Diversity. Science 2016, 354, Aag1465. [PubMed: 27708059]

(302). Chalker JM, Lercher L, Rose NR, Schofield CJ, and Davis BG Conversion of Cysteine into Dehydroalanine Enables Access to Synthetic Histones Bearing Diverse Post-Translational Modifications. Angew. Chem. Int. Ed 2012, 51, 1835–1839.

(303). Yang A, Ha S, Ahn J, Kim R, Kim S, Lee Y, Kim J, Soll D, Lee HY, and Park HS A Chemical Biology Route to Site-specific Authentic Protein Modifications. Science 2016, 354, 623–626. [PubMed: 27708052]

(304). Terasaka N, Iwane Y, Geiermann AS, Goto Y, and Suga H Recent Developments of Engineered Translational Machineries for the Incorporation of Non-canonical Amino Acids into Polypeptides. Int. J. Mol. Sci 2015, 16, 6513–6531. [PubMed: 25803109]

(305). Chin JW Expanding and Reprogramming the Genetic Code. Nature 2017, 550, 53–60. [PubMed: 28980641]

(306). Guo J, Wang J, Lee JS, and Schultz PG Site-specific Incorporation of Methyl- and Acetyl-Lysine Analogues into Recombinant Proteins. Angew. Chem. Int. Ed 2008, 47, 6399–6401.

(307). Wang ZA, Zeng Y, Kurra Y, Wang X, Tharp JM, Vatansever EC, Hsu WW, Dai S, Fang X, and Liu WR A Genetically Encoded Allysine for the Synthesis of Proteins with Site-Specific Lysine Dimethylation. Angew. Chem. Int. Ed 2017, 56, 212–216.

(308). Ai HW, Lee JW, and Schultz PG A Method to Site-specifically Introduce Methyllysine into Proteins in E. coli. Chem. Commun 2010, 46, 5506–5508.

(309). Wang Y-S, Wu B, Wang Z, Huang Y, Wan W, Russell WK, Pai P-J, Moe YN, Russell DH, and Liu WR A Genetically Encoded Photocaged Nε-Methyl-L-Lysine. Mol. Biosyst 2010, 6, 1575–1578.

(310). Wang ZA, and Liu WR Proteins with Site-Specific Lysine Methylation. Chemistry 2017, 23, 11732–11737. [PubMed: 28500859]

(311). Nguyen DP, Alai MMG, Kapadnis PB, Neumann H, and Chin JW Genetically Encoding Nε-Methyl-L-Lysine in Recombinant Histones. J. Am. Chem. Soc 2009, 131, 14194–14195. [PubMed: 19772323]

(312). Nguyen DP, Alai MMG, Virdee S, and Chin JW Genetically Directing ε-N, N-Dimethyl-L-Lysine in Recombinant Histones. Chem. Biol 2010, 17, 1072–1076. [PubMed: 21035729]

(313). Li J, Yu J, Zhao J, Wang J, Zheng S, Lin S, Chen L, Yang M, Jia S, Zhang X, et al. Palladium-triggered Deprotection Chemistry for Protein Activation in Living Cells. Nat. Chem 2014, 6, 352–361. [PubMed: 24651204]

(314). Klan P, Solomek T, Bochet CG, Blanc A, Givens R, Rubina M, Popik V, Kostikov A, and Wirz J Photoremovable Protecting Groups in Chemistry and Biology: Reaction Mechanisms and Efficacy. Chem. Rev 2013, 113, 119–191. [PubMed: 23256727]

(315). Yanagisawa T, Ishii R, Fukunaga R, Kobayashi T, Sakamoto K, and Yokoyama S Multistep Engineering of Pyrrolysyl-tRNA Synthetase to Genetically Encode Nε-(O-Azidobenzyloxycarbonyl) Lysine for Site-Specific Protein Modification. Chem. Biol 2008, 15, 1187–1197. [PubMed: 19022179]

(316). Li J, Jia S, and Chen PR Diels-Alder Reaction-triggered Bioorthogonal Protein Decaging in Living Cells. Nat. Chem. Biol 2014, 10, 1003–1005. [PubMed: 25362360]

(317). Li J, and Chen PR Development and Application of Bond Cleavage Reactions in Bioorthogonal Chemistry. Nat. Chem. Biol 2016, 12, 129–137. [PubMed: 26881764]

(318). Xie X, Li XM, Qin F, Lin J, Zhang G, Zhao J, Bao X, Zhu R, Song H, Li XD, et al. Genetically Encoded Photoaffinity Histone Marks. J. Am. Chem. Soc 2017, 139, 6522–6525. [PubMed: 28459554]

(319). Yang T, Li XM, Bao X, Fung YM, and Li XD Photo-Lysine Captures Proteins that Bind Lysine Post-translational Modifications. Nat. Chem. Biol 2016, 12, 70–72. [PubMed: 26689789]

(320). Allis CD, and Muir TW Spreading Chromatin into Chemical Biology. Chembiochem 2011, 12, 264–279. [PubMed: 21243714]

(321). Dann GP, Liszczak GP, Bagert JD, Muller MM, Nguyen UTT, Wojcik F, Brown ZZ, Bos J, Panchenko T, Pihl R, et al. ISWI Chromatin Remodellers Sense Nucleosome Modifications to Determine Substrate Preference. Nature 2017, 548, 607–611. [PubMed: 28767641]

(322). Nguyen UT, Bittova L, Muller MM, Fierz B, David Y, Houck-Loomis B, Feng V, Dann GP, and Muir TW Accelerated Chromatin Biochemistry Using DNA-barcoded Nucleosome Libraries. Nat. Methods 2014, 11, 834–840. [PubMed: 24997861]

(323). Wang P, Dong S, Shieh JH, Peguero E, Hendrickson R, Moore MAS, and Danishefsky SJ Erythropoietin Derived by Chemical Synthesis. Science 2013, 342, 1357–1360. [PubMed: 24337294]

(324). Zhang Q, Johnston EV, Shieh JH, Moore MA, and Danishefsky SJ Synthesis of Granulocyte-macrophage Colony-stimulating Factor as Homogeneous Glycoforms and Early Comparisons with Yeast Cell-derived Material. Proc. Natl. Acad. Sci. USA 2014, 111, 2885–2890. [PubMed: 24516138]

(325). Shah NH, Dann GP, Vila-Perello M, Liu Z, and Muir TW Ultrafast Protein Splicing is Common among Cyanobacterial Split Inteins: Implications for Protein Engineering. J. Am. Chem. Soc 2012, 134, 11338–11341. [PubMed: 22734434]

(326). David Y, Vila-Perello M, Verma S, and Muir TW Chemical Tagging and Customizing of Cellular Chromatin States Using Ultrafast Trans-splicing Inteins. Nat. Chem 2015, 7, 394–402. [PubMed: 25901817]

(327). Stevens AJ, Sekar G, Shah NH, Mostafavi AZ, Cowburn D, and Muir TW A Promiscuous Split Intein with Expanded Protein Engineering Applications. Proc. Natl. Acad. Sci. USA 2017, 114, 8538–8543. [PubMed: 28739907]

(328). Henager SH, Chu N, Chen Z, Bolduc D, Dempsey DR, Hwang Y, Wells J, and Cole PA Enzyme-catalyzed Expressed Protein Ligation. Nat. Methods 2016, 13, 925–927. [PubMed: 27669326]

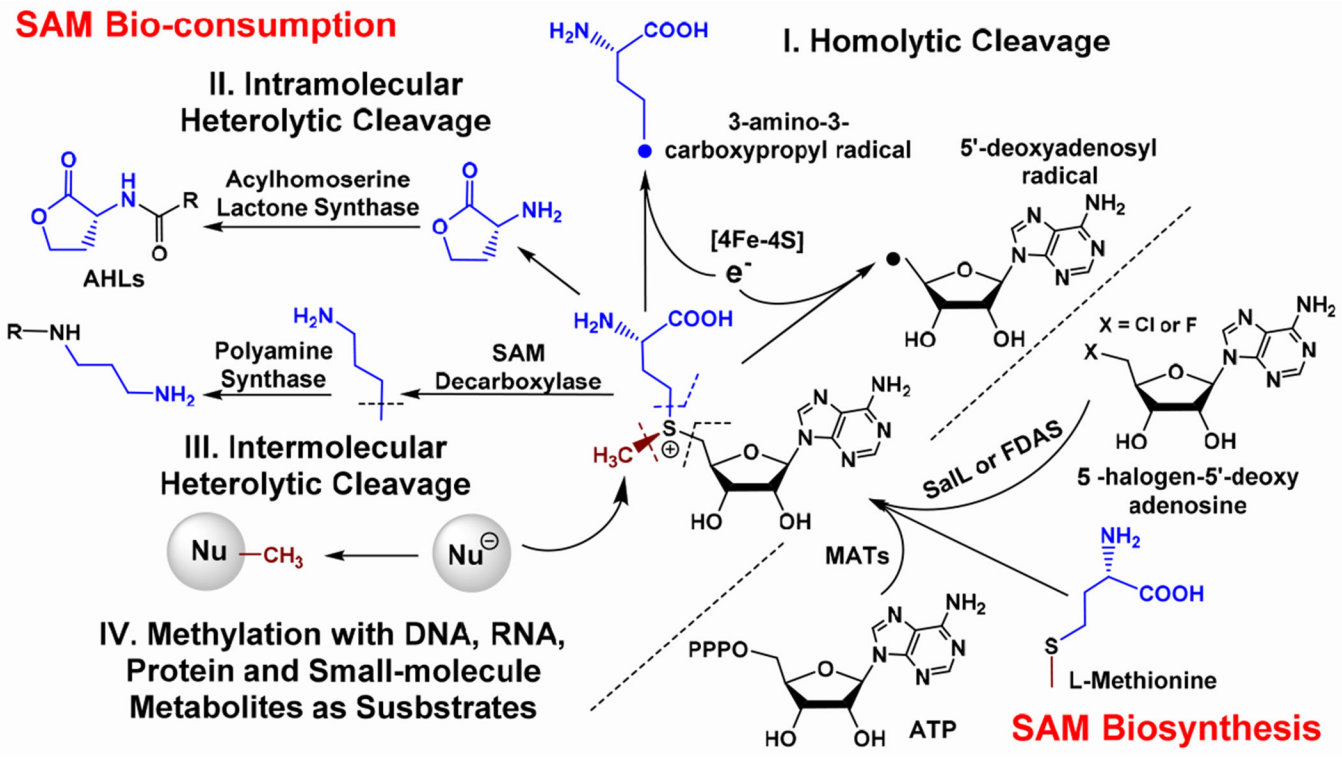


Figure 1.

Biosynthesis and bio-consumption of SAM. In general, SAM's biosynthesis is carried out by L-methionine adenosyltransferases (MATs or SAM synthetases) with L-methionine and ATP as substrates. Occasionally, bacterial enzymes such as SalL and FDAS can generate SAM with 5′-Cl/F-5′-deoxyadenosine and L-methionine as substrates. Bio-consumption of SAM can be classified as (i) homolytic, (ii) intramolecular heterolytic, (iii) intermolecular heterolytic cleavage of two CH_2-sulfonium bonds, and (iv) intermolecular heterolytic cleavage of the methyl-sulfonium bond. The latter accounts for methylation with DNA, RNA, proteins and small-molecule metabolites as substrates.

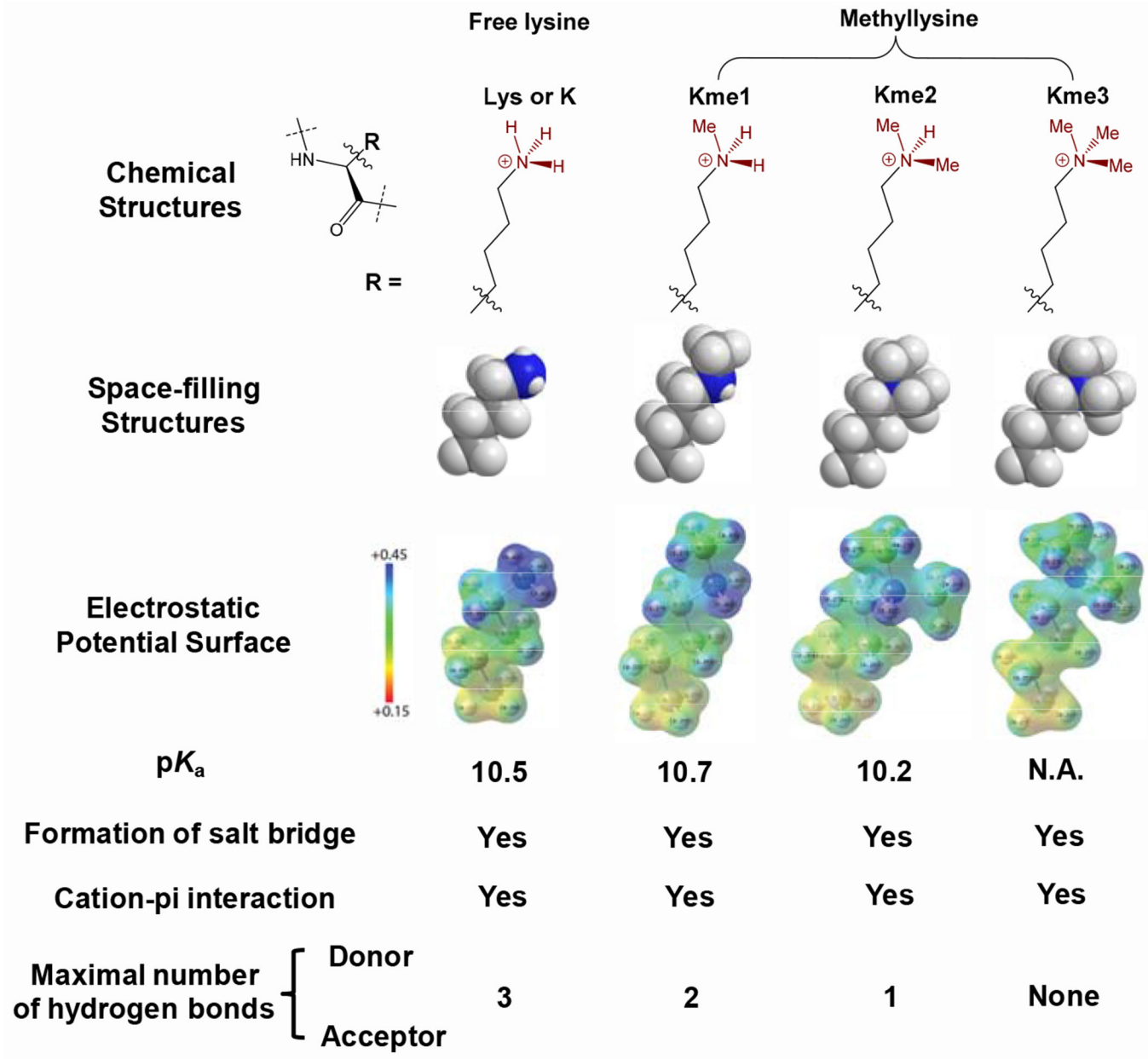


Figure 2.

Biophysical and biochemical properties of free lysine (Lys) and methyllysine (Kme1/2/3). The side chains of Lys and Kme1/2/3 are shown in a space-filling model and with electrostatic potential surface. Their p*K*a values and capability to form a salt bridge and hydrogen bonds at physiological pH of 7.4 are compared.

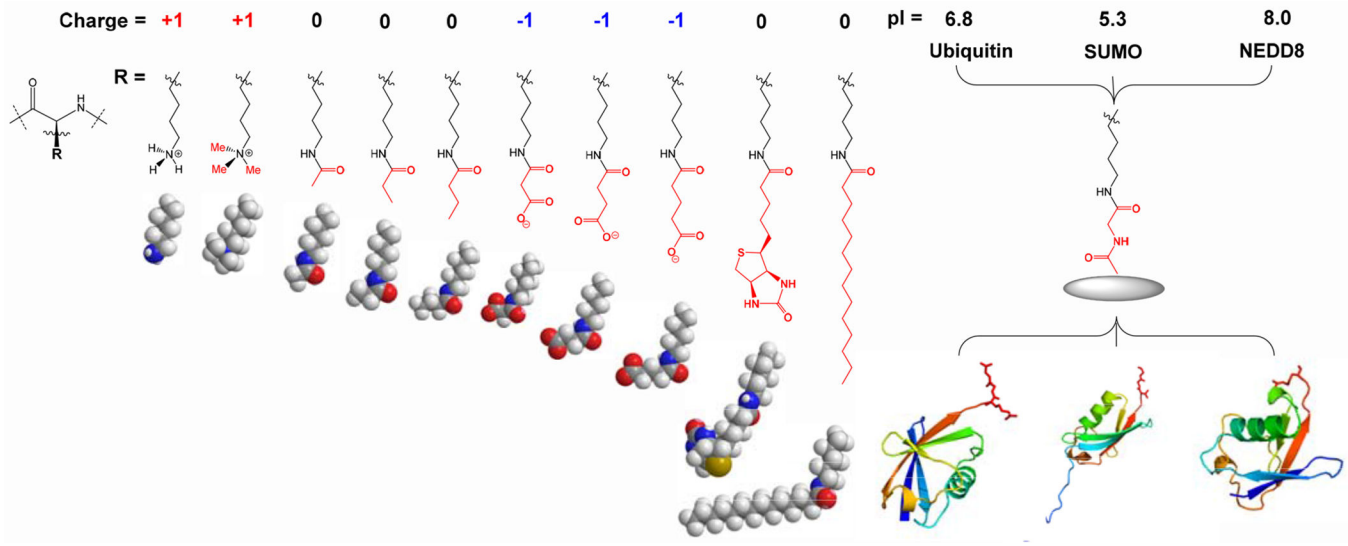


Figure 3.

Structures and formal charges of posttranslationally modified Lys residues. Different from other posttranslational modifications, lysine methylation is characterized by unaltered charge (+1) at physiological pH and a minimal change of size relative to an unmodified lysine. Relative sizes of the posttranslational modifications except ubiquitin, SUMO and Nedd8 are compared in a space-filling model.

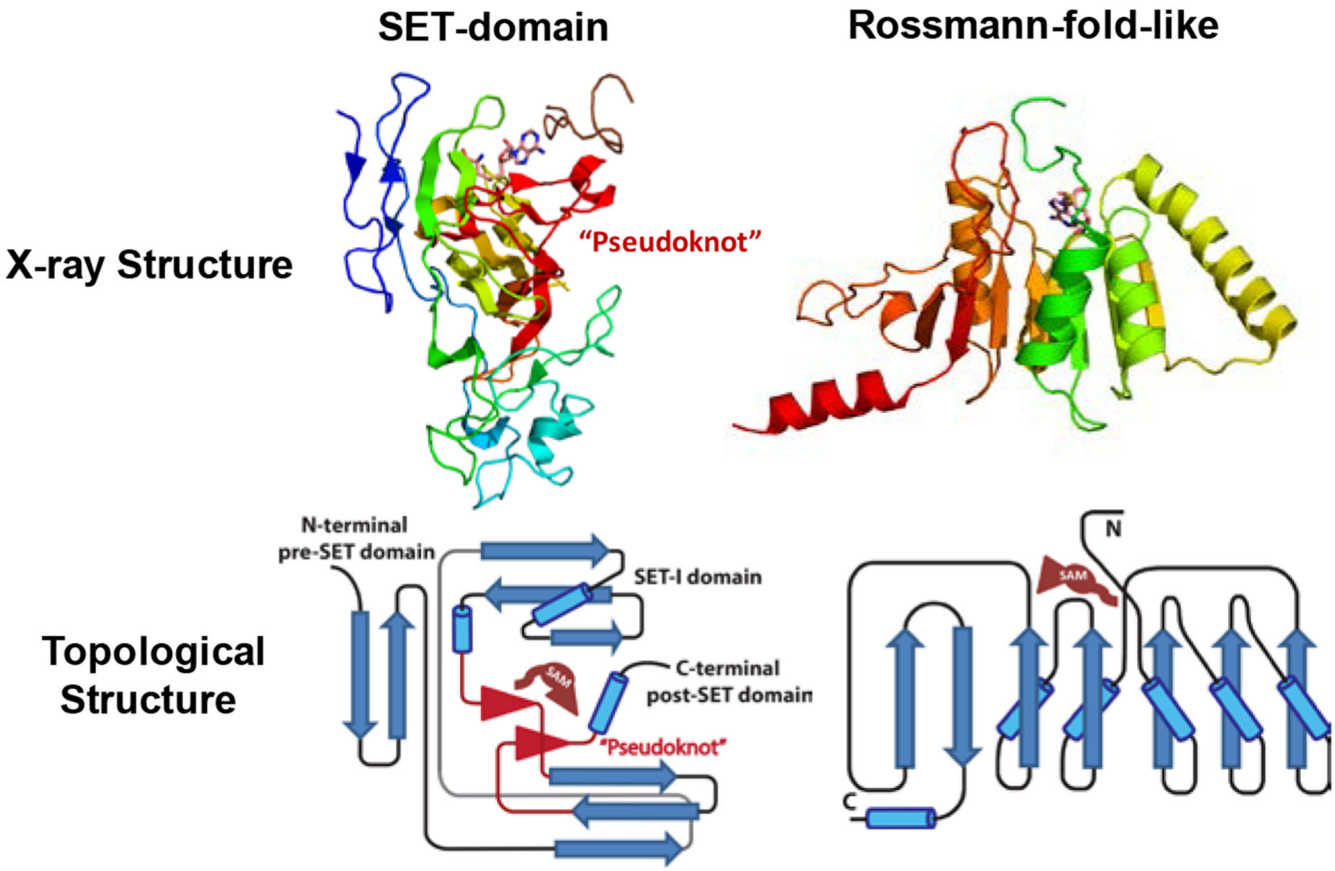


Figure 4.

Structures and topology of PKMTs. The cartoon 3-D structures and 2–D topology of G9a (PDB 2O8J) and DOT1L (PDB 1NW3) are shown as representative examples of SET domain-containing PKMTs and Rossmann-fold-like (non-SET-domain) PKMTs, respectively. The SAM binding of the two PKMTs and the “pseudoknot” fold of G9a in red are highlighted.

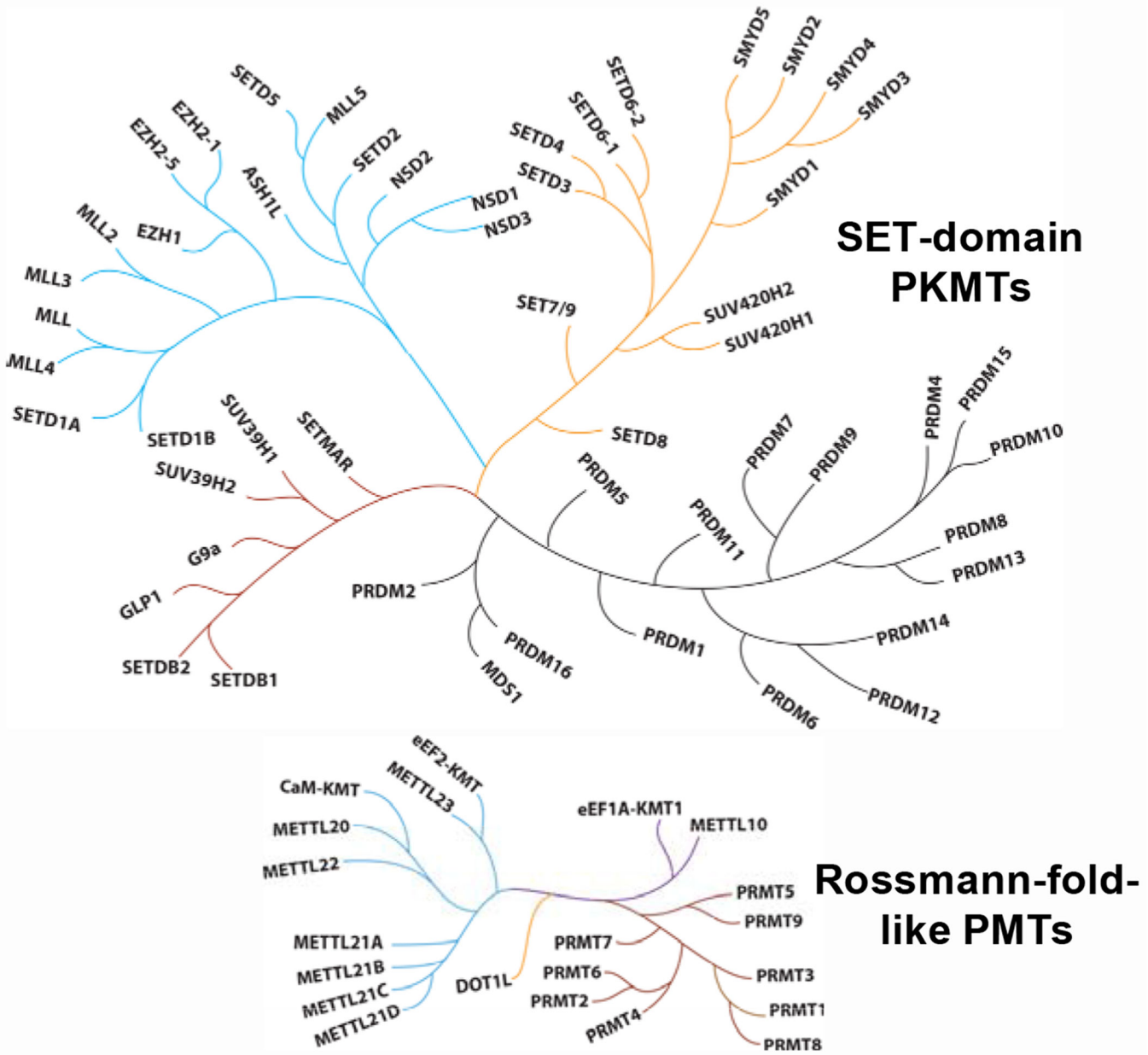


Figure 5.

Phylogenic trees of SET domain-containing PKMTs and Rossmann-fold-like PKMTs. Classification and relative positions of PKMTs were referred in a previous report.(28, 73)

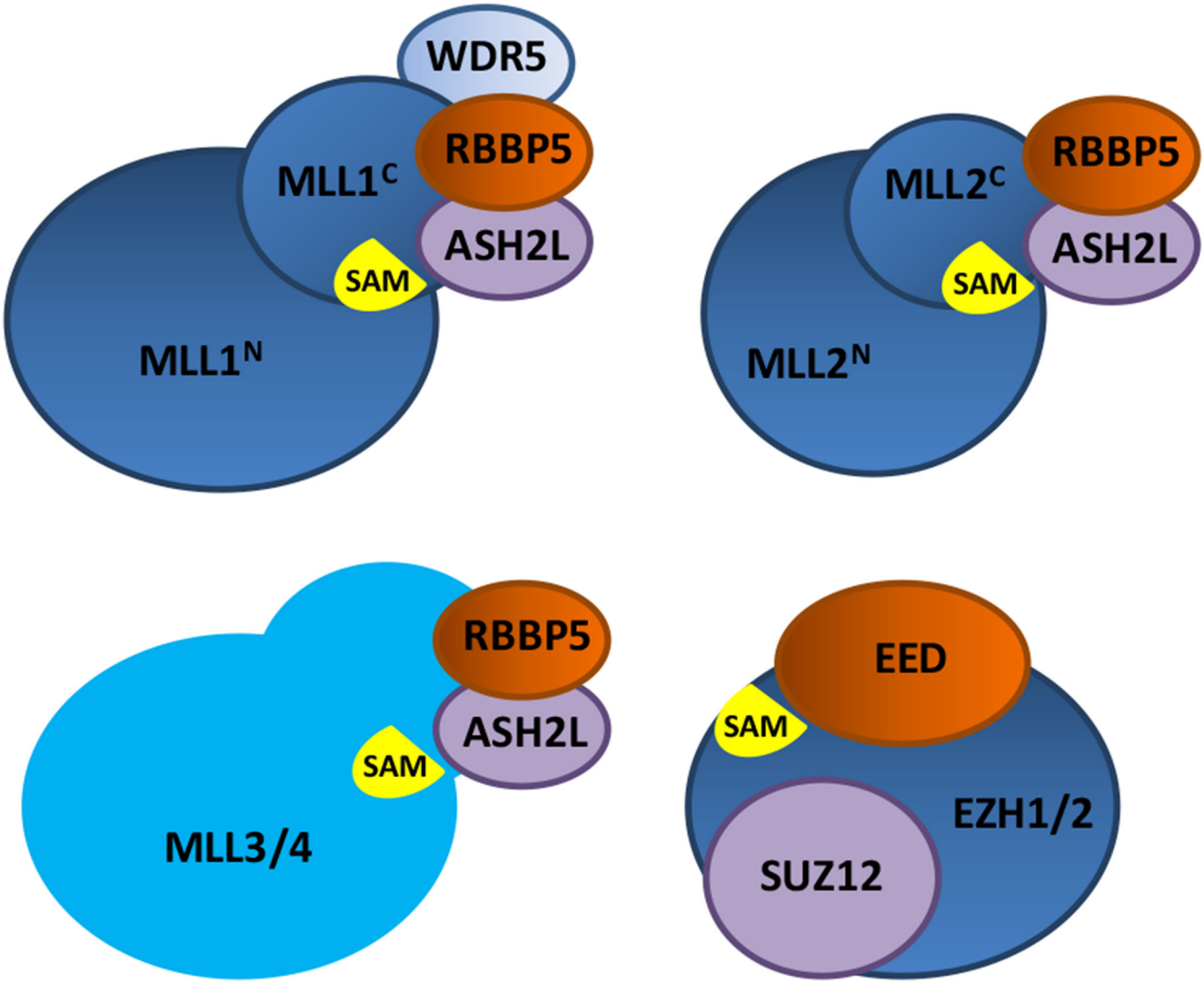


Figure 6.

Representative PKMT complexes. Here shown are key components of catalytically active MLL1-4 and EZH1/2 complexes.(90, 91) MLL1/2 contain cleavage sites for the threonine aspartase Taspase1. (92) The relative topology of individual subunits was presented on the basis of homogenous X-ray structures as reported.(90, 91, 93, 94)

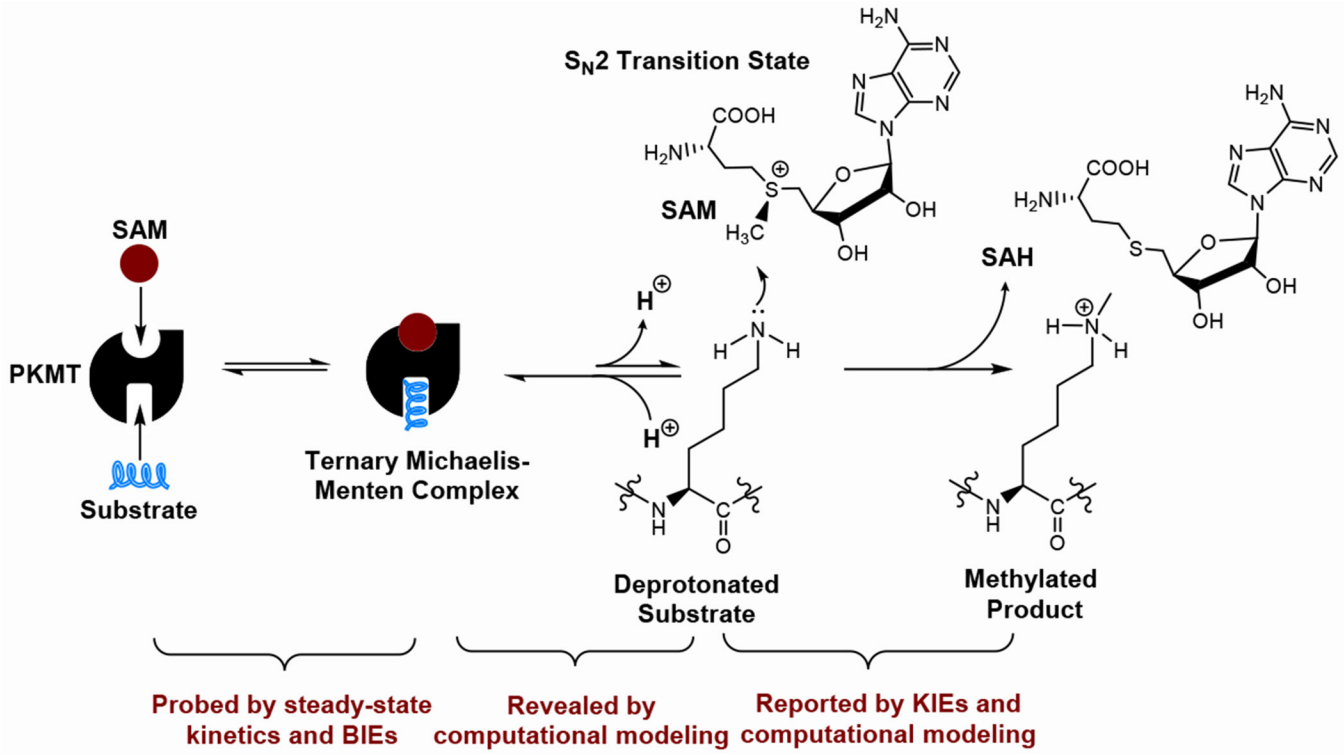


Figure 7.

Reaction path of PKMT-catalyzed lysine monomethylation and relevant biochemical methods to examine this process. PKMT first recognizes the SAM cofactor and its substrate. The ε-amine of lysine is then subjected to enzyme-mediated deprotonation. The overall rate-limitation step is expected to be the assembling of a S_N2 transition state, followed by releasing methylated lysine and SAH as a product and a byproduct, respectively. Reproduced with permission from Ref. (98) Copyright 2016 PNAS.

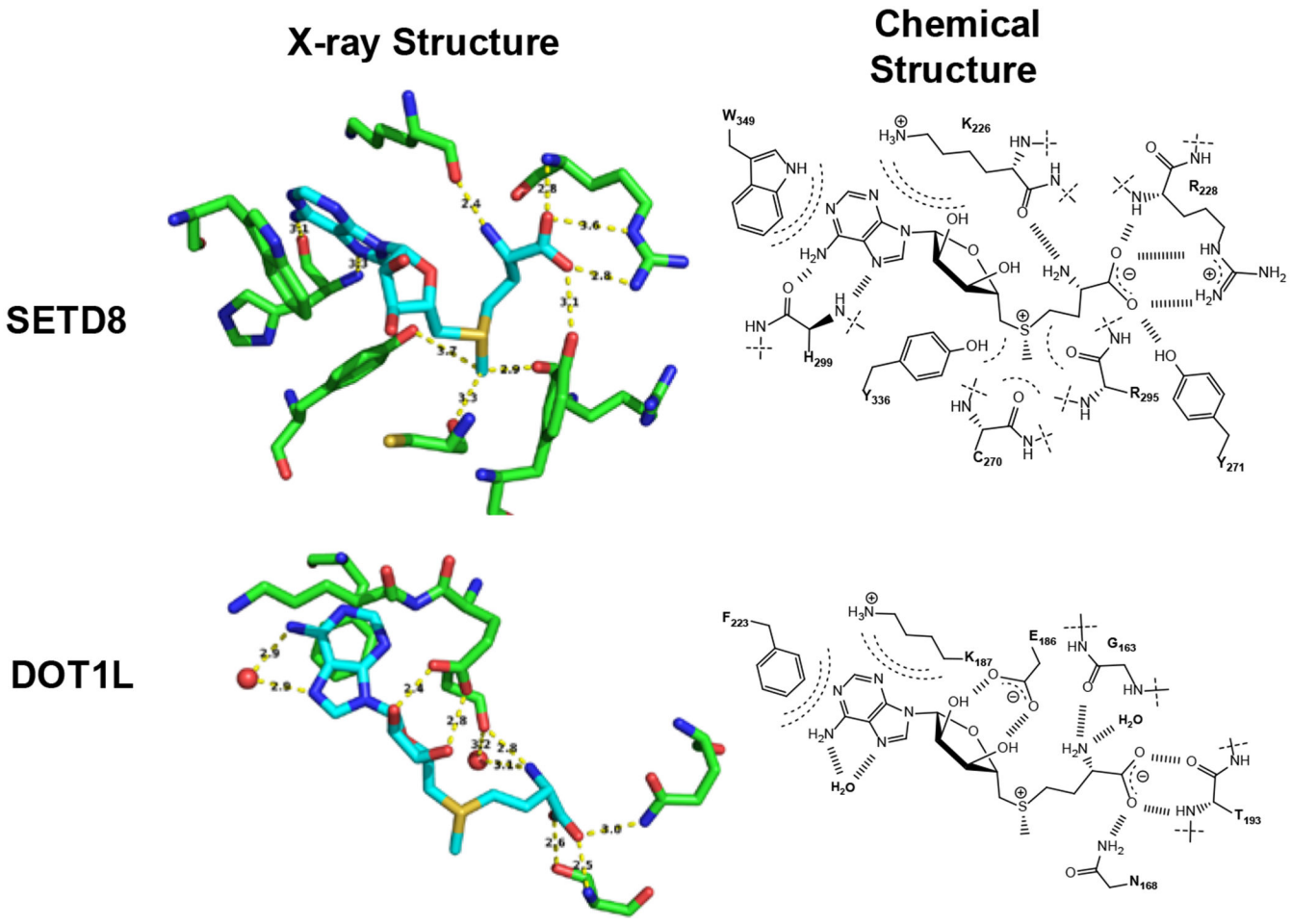


Figure 8.

SAM-binding modes of PKMTs. The interacting networks of SAM in SETD8 (PDB 3F9W) and DOT1L (PDB 1NW3) are shown as the representative examples of SET domain-containing PKMTs and Rossmann-fold-like (non-SET-domain) PKMTs, respectively. Both the stick mode generated with PyMOL and structural details of SAM binding are shown for clarity.

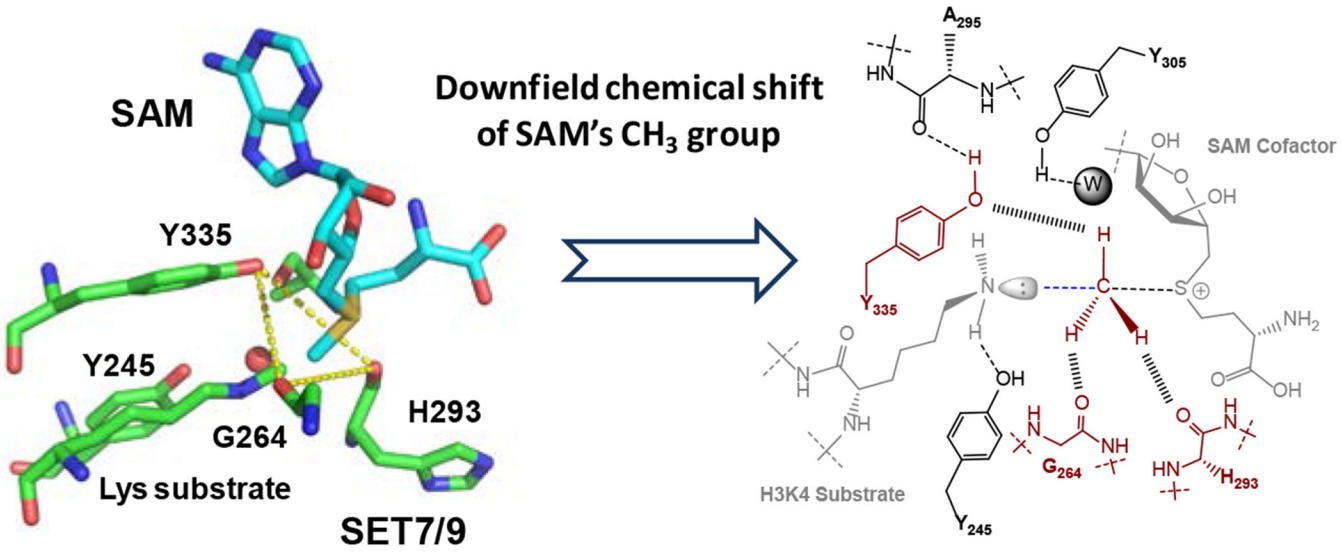


Figure 9.

Noncanonical carbon-oxygen (CH•••O) interaction for SAM engagement. SET7/9 was used as an example to illustrate the noncanonical CH•••O interaction. The hypothetic transition state structure was generated upon aligning two PDB files 1XQH and 1N6C. Both the stick mode generated with PyMOL and a chemical structure model are shown for clarity.

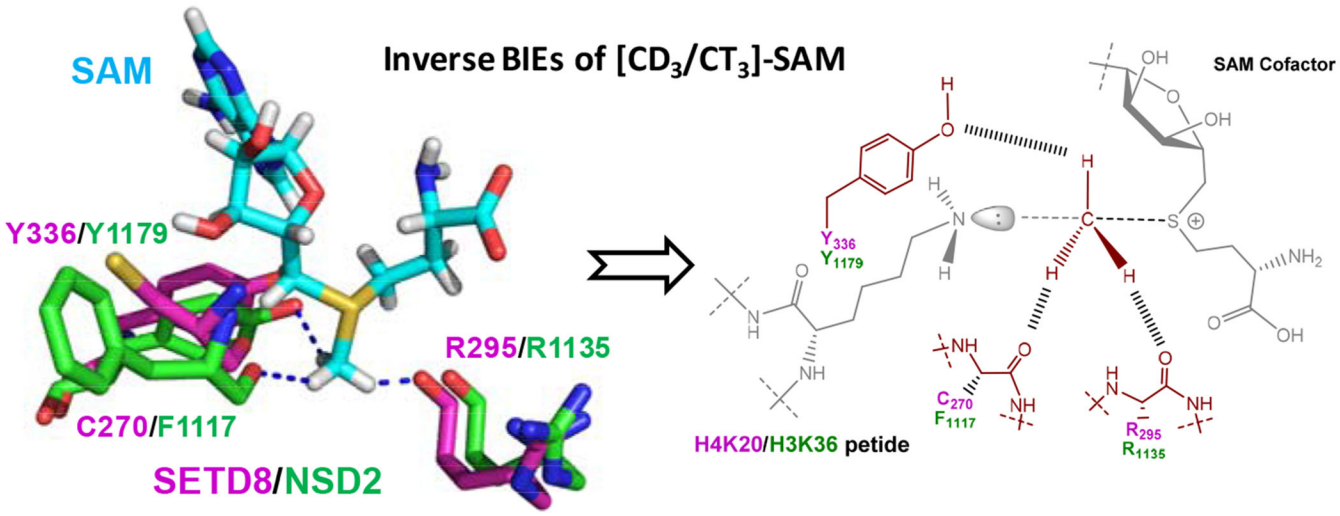


Figure 10.

Inverse BIEs of CD_3-SAM associated with SAM binding. Inverse BIEs of CD_3-SAM were reported for SETD8 and NSD2. The noncanonical CH…O interactions were highlighted with the stick mode generated with PyMOL (PDB 3F9W for SETD8 and 5LSU for NSD2) and with the chemical structures at proposed transition states. SAM in the NSD2-SAM complex (PDB 5LSU) was used to depict the interactions of SETD8 and NSD2 at the catalytic sites.

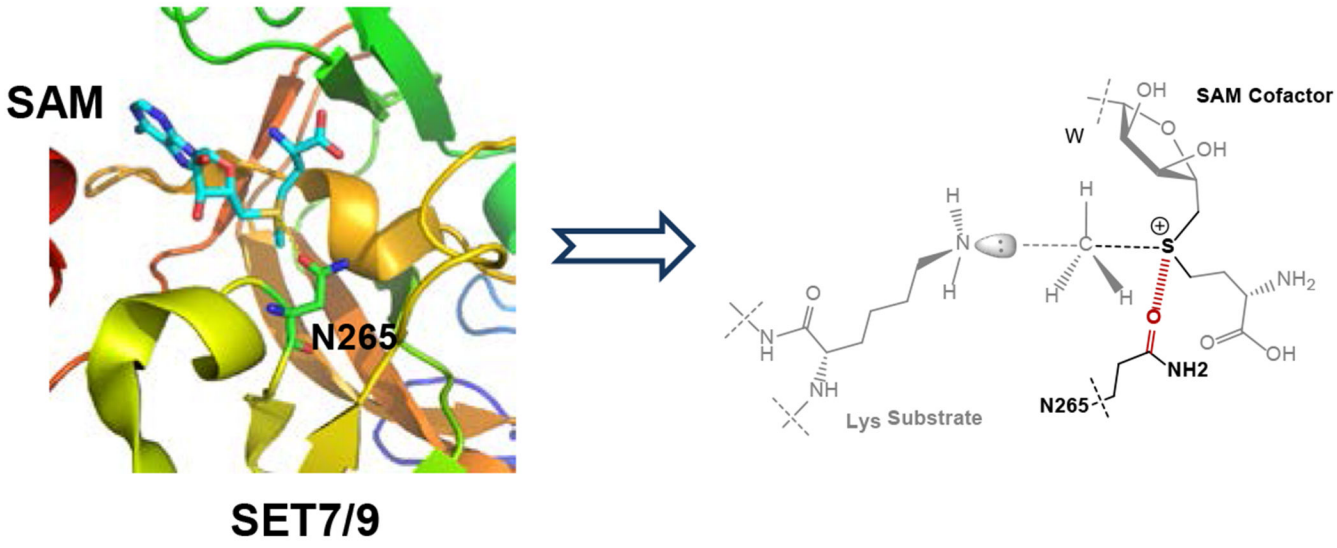


Figure 11.

Noncanonical sulfur-oxygen (S•••O) interaction for SAM engagement. SET7/9 was used as an example to illustrate the noncanonical chalcogen-oxygen interaction. The hypothetic transition state structure was generated upon aligning two PDB files 1XQH and 1N6C. Both the stick mode generated with PyMOL and the chemical structure model at a proposed transition state are shown for clarity.

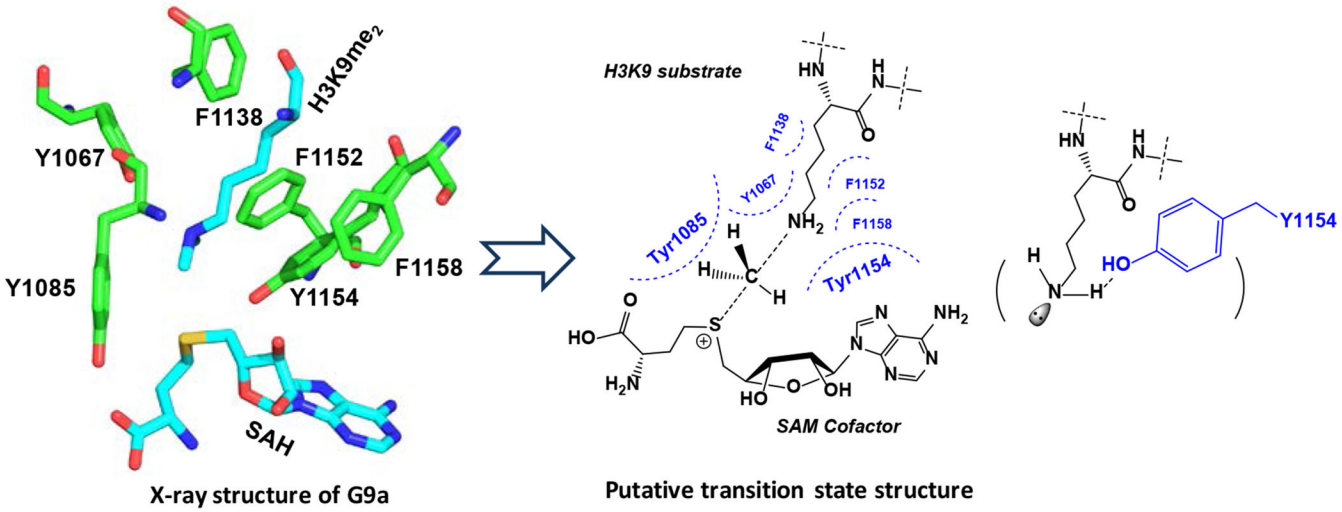


Figure 12.

Substrate-binding mode of SET domain-containing PKMTs. With G9a as an example (PDB 2O8J), its lysine binding pocket consists of multiple aromatic residues to engage in cation-π interactions with positively charged lysine or methyllysine residues. The hydrophobic pocket positions the side chains of these residues in a ready linear trajectory for a S_N2 transition state, partially through a hydrogen bond between their ε-amine moieties and G9a's Y1154 residue.

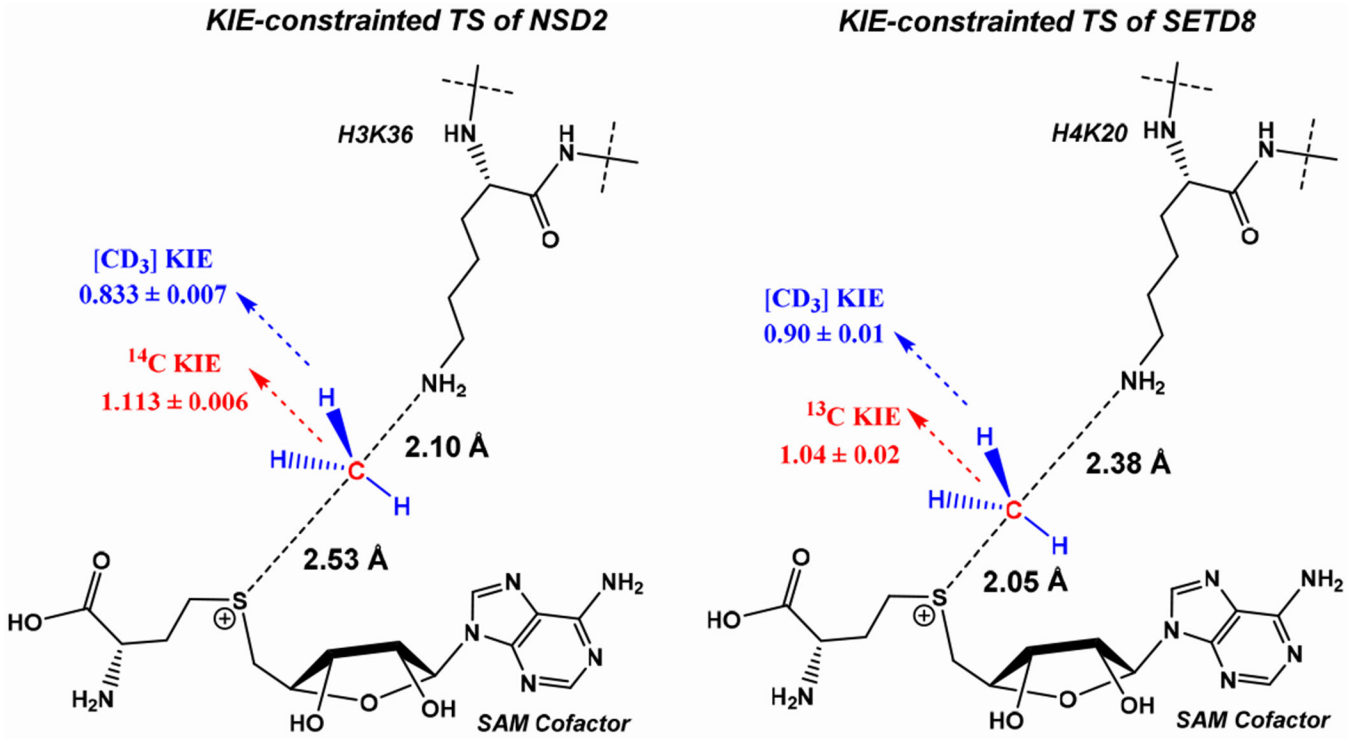


Figure 13.

Transition states of NSD2 and SETD8. A late S_N2 transition state of NSD2 and an early S_N2 transition state of SETD8 were solved with their KIEs as geometrical constraints.

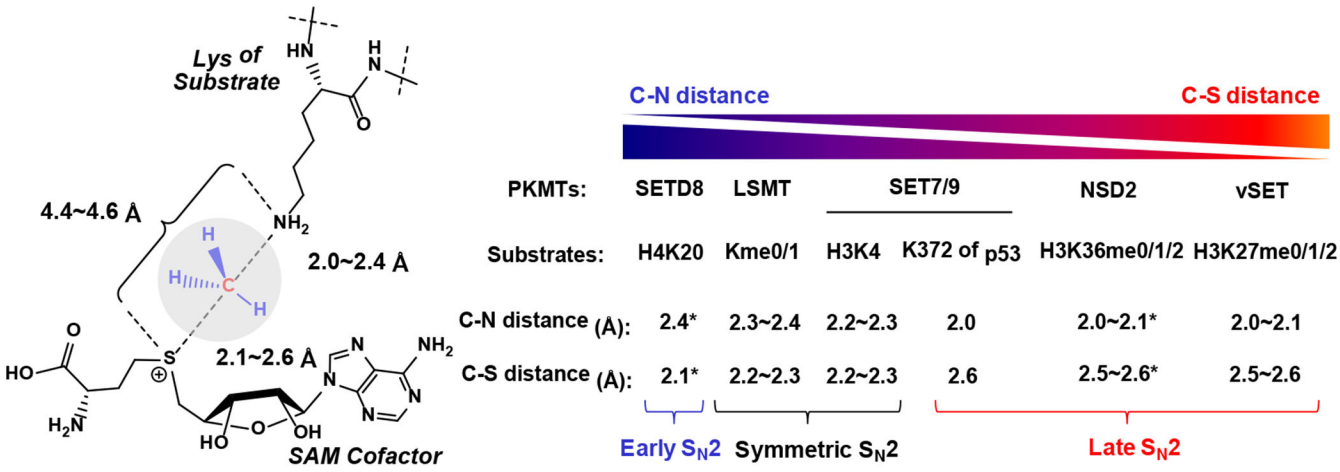


Figure 14.

Summary of the reported transition states of PKMTs. Described here are the PKMTs, their substrates, and characteristic C-S and C-N distances at the solved transition states.

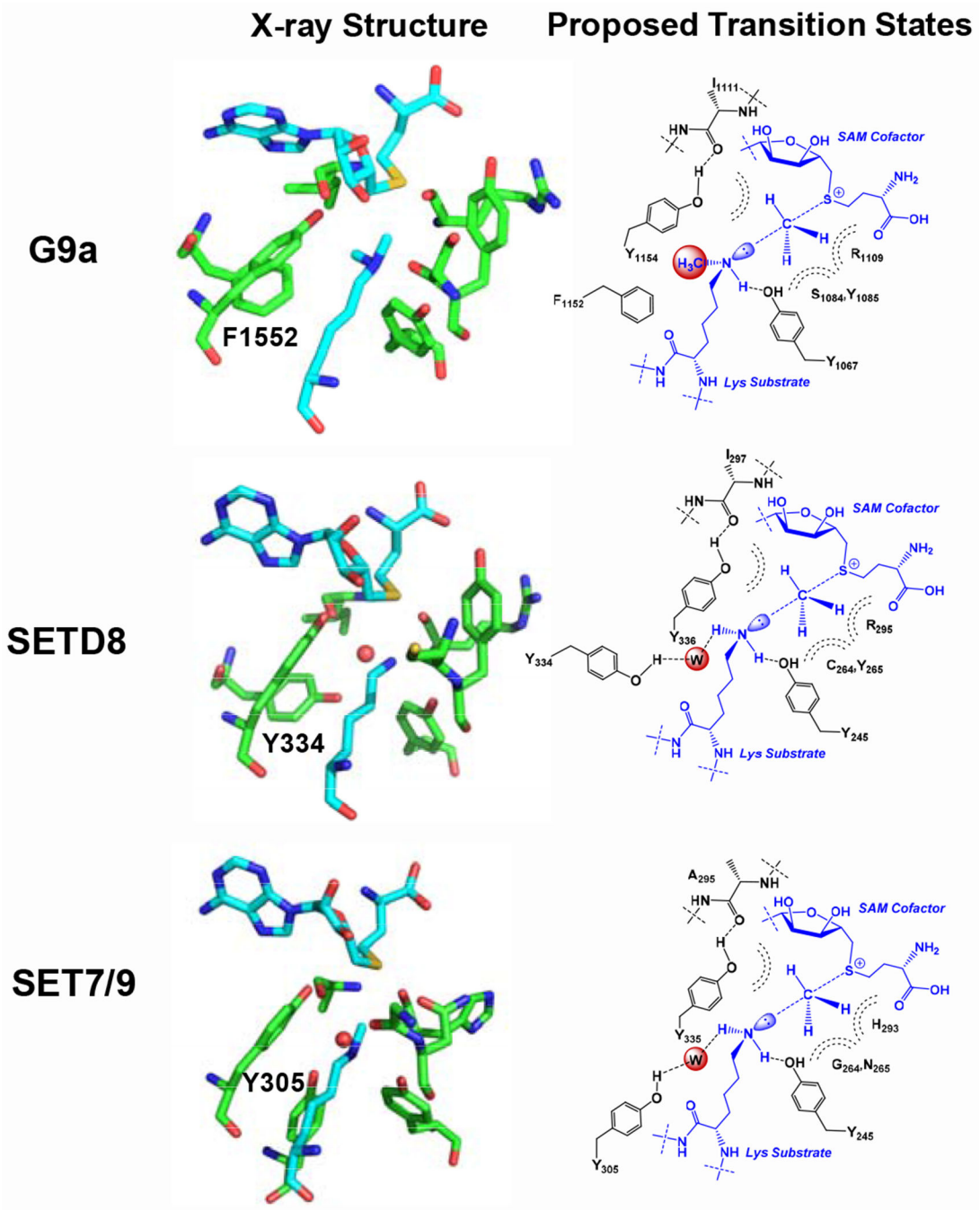


Figure 15.

Phe-Tyr switch associated with product specificity of PKMTs. G9a can carry out dimethylation. In contrast, SETD8 and SET7/9 mainly monomethylate their substrates. Such product specificity is expected to be controlled by a characteristic Phe-Tyr switch and the associated water molecule. The Phe residue allows a vacant space for dimethylation. The Tyr residue binds a water molecule, which occupies the space otherwise for a monomethylated substrate, and thus prevents further methylation at their transition states. PDB files of 2O8J, 2F9W and 1XQH for G9a, SETD8 and SET7/9, respectively.

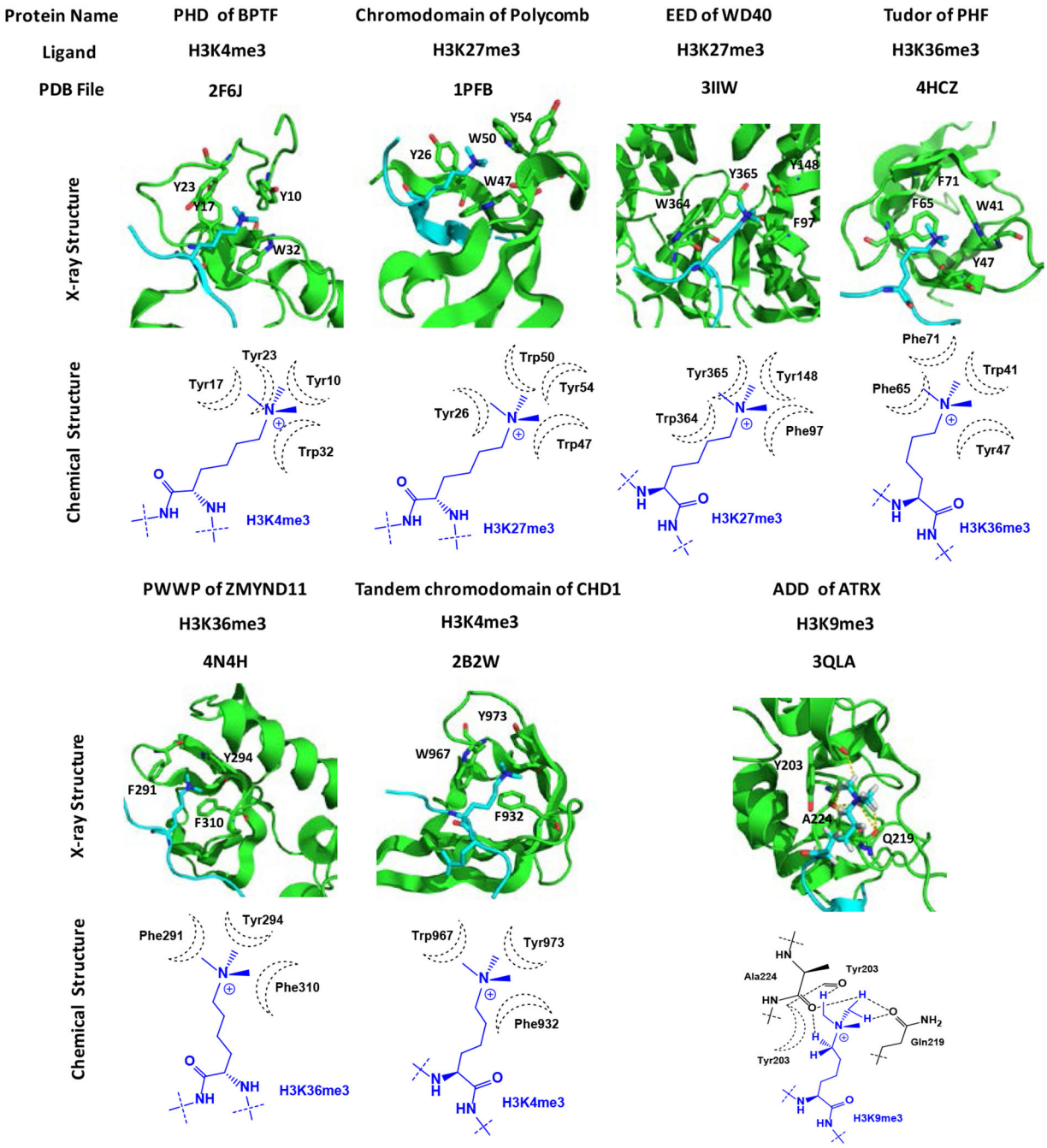


Figure 16.

Recognition of Kme3 by diverse reader domains. Shown here are a collection of reader domains, their preferential Kme3 ligands and the associated PDB files. Both the stick mode generated by PyMOL and chemical structures are shown for clarity.

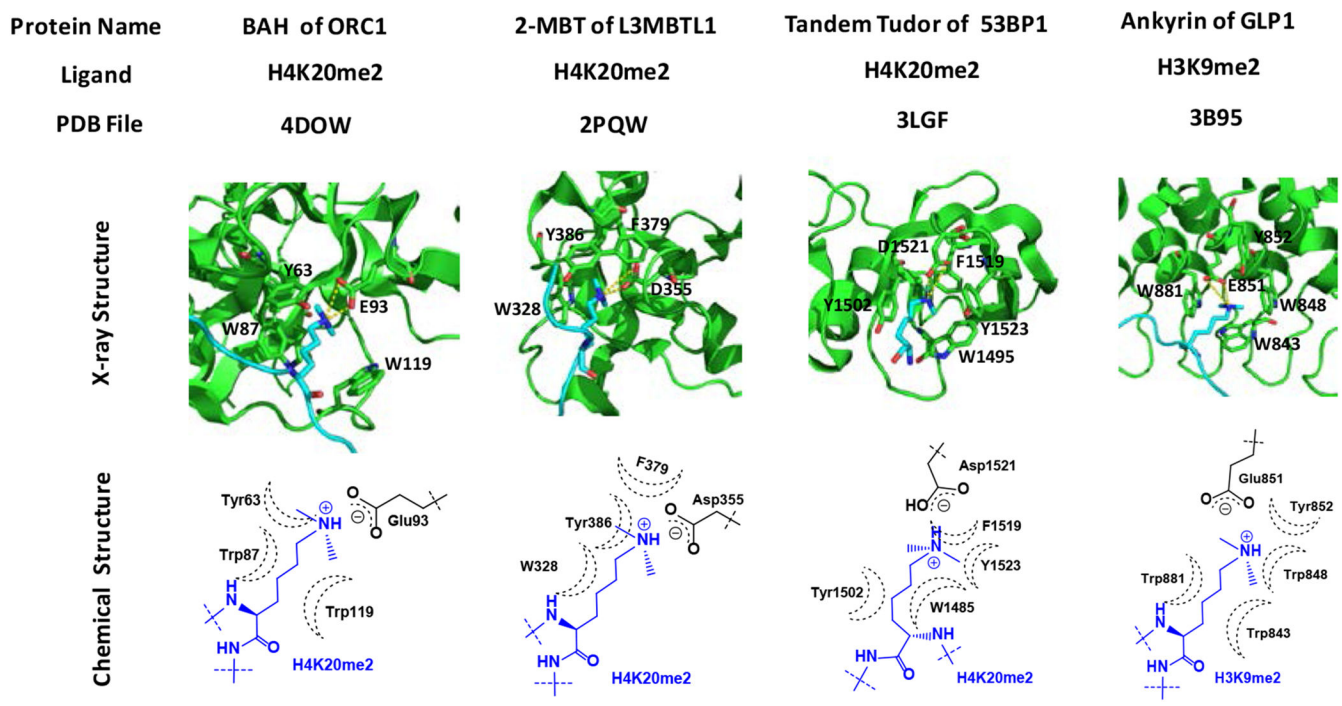


Figure 17.

Recognition of Kme2 by diverse reader domains. Shown here are a collection of reader domains, their preferential Kme2 ligands and the associated PDB files. Both the stick mode generated by PyMOL and chemical structures are shown for clarity.

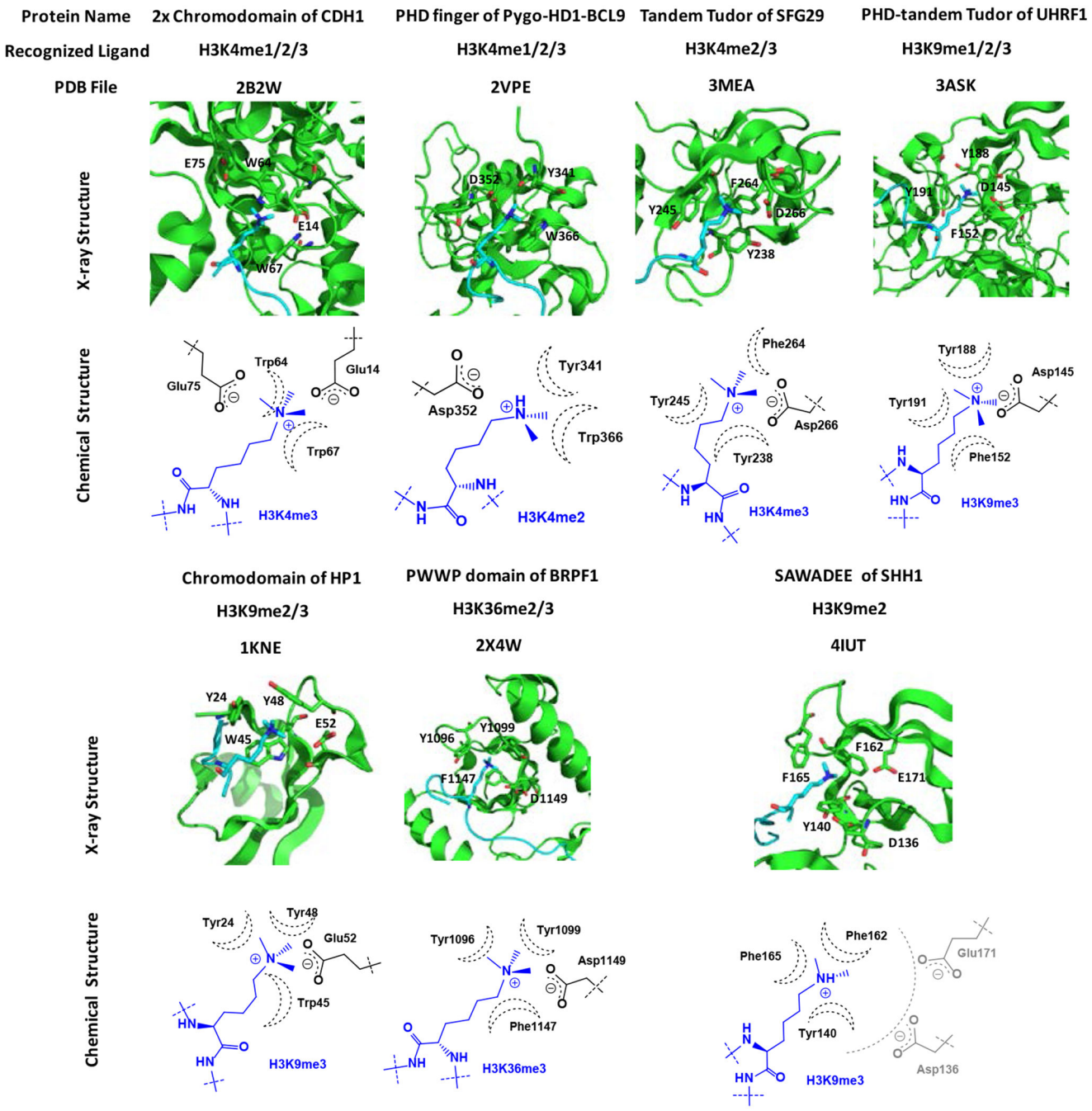


Figure 18.

Promiscuous recognition of methyllysine by diverse reader domains. Shown here are a collection of reader domains, their promiscuous methyllysine ligands and PDB files. Both the stick mode generated with PyMOL and chemical structures are shown for clarity.

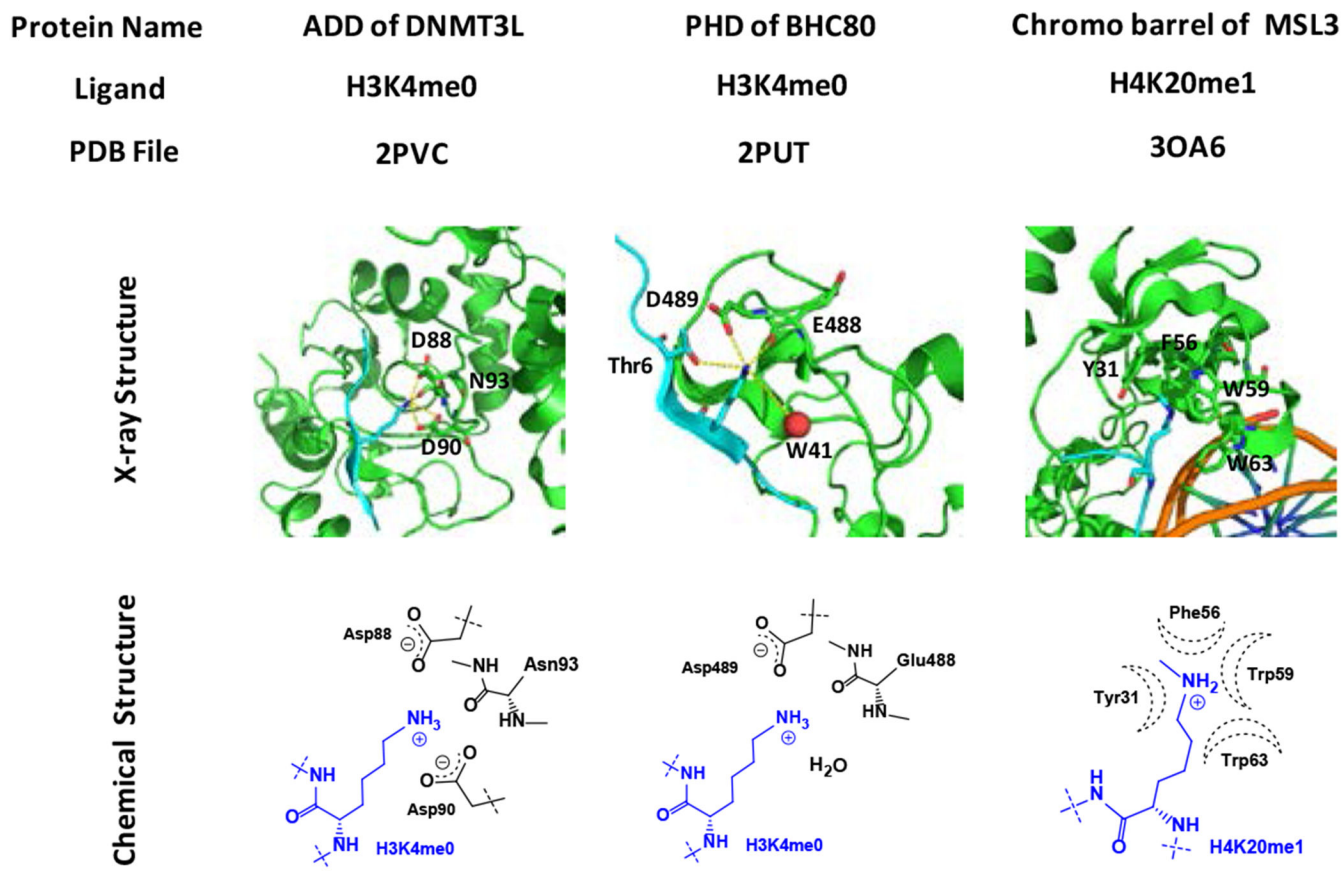


Figure 19.

Recognition of unmodified Lys and Kme1 by reader domains. Shown here are a collection of reader domains, their preferential low states of lysine methylation and the associated PDB files. Both the stick mode generated with PyMOL and chemical structures are shown for clarity.

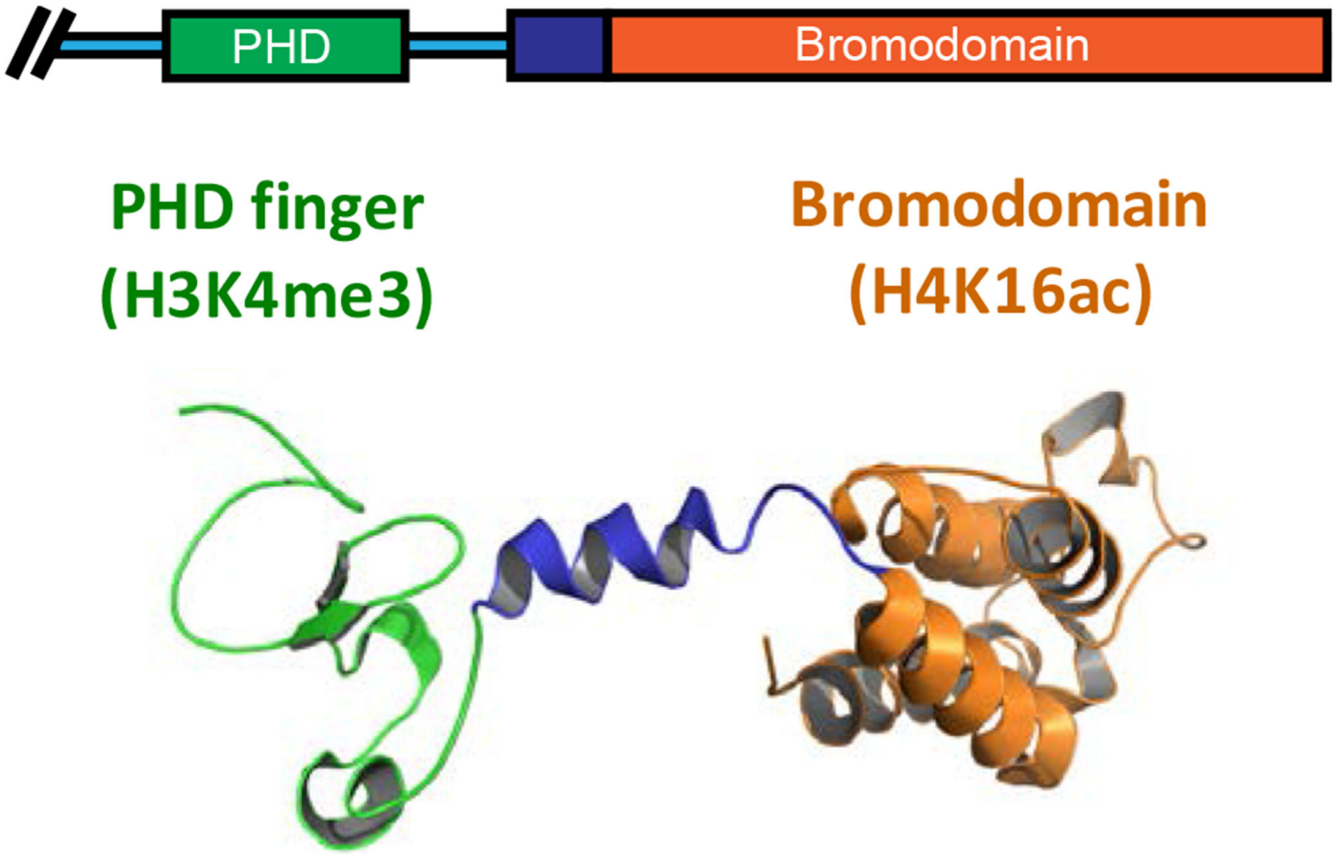


Figure 20.
Structure and topology of human BPTF. Human BPTF (PDB: 2F6J) contains two reader domains to recognize H3K4me3 and H4K16ac in a synergistic manner. The 2–D topology and 3-D cartoon structure of human BPTF are shown.

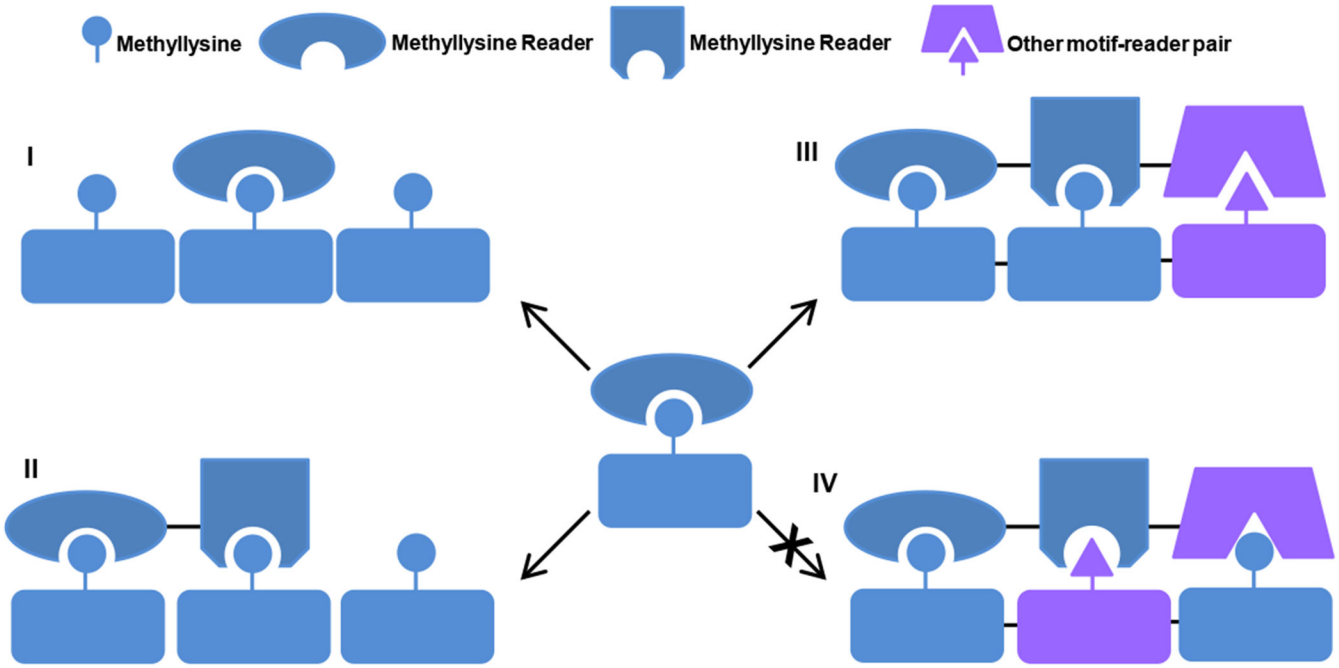


Figure 21.

Modules to agonize or antagonize methyllysine reader domains. The interaction of methyllysine reader domains and their ligands can be strengthened in the presence of locally high concentrations of these ligands or through additional interactions of reader domains with other posttranslational marks. In contrast, such interactions can be antagonized in the presence of unmatched ligands.

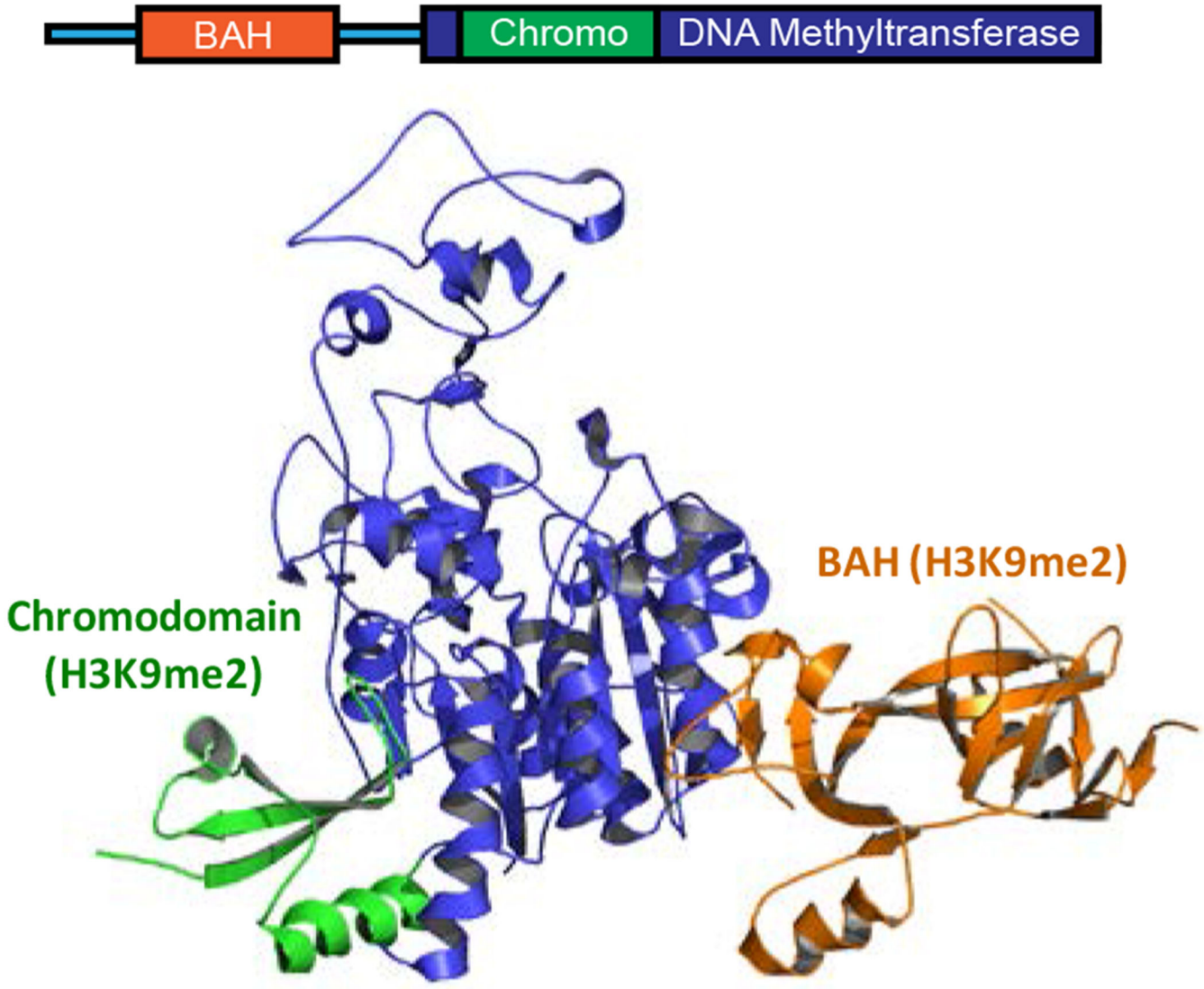


Figure 22.

Structure and topology of human ZMET2. ZMET2 (PDB 4FT2) contains BAH and chromo domains, and is expected to recognize H3K9me2-containing nucleosomes through binding two H3K9me2 simultaneously. Such multivalent interactions positions the DNA methyltransferase domain of ZMET2 to methylate nearby nucleosome DNA. The 2–D topology and 3-D cartoon structure of ZMET2 are shown.

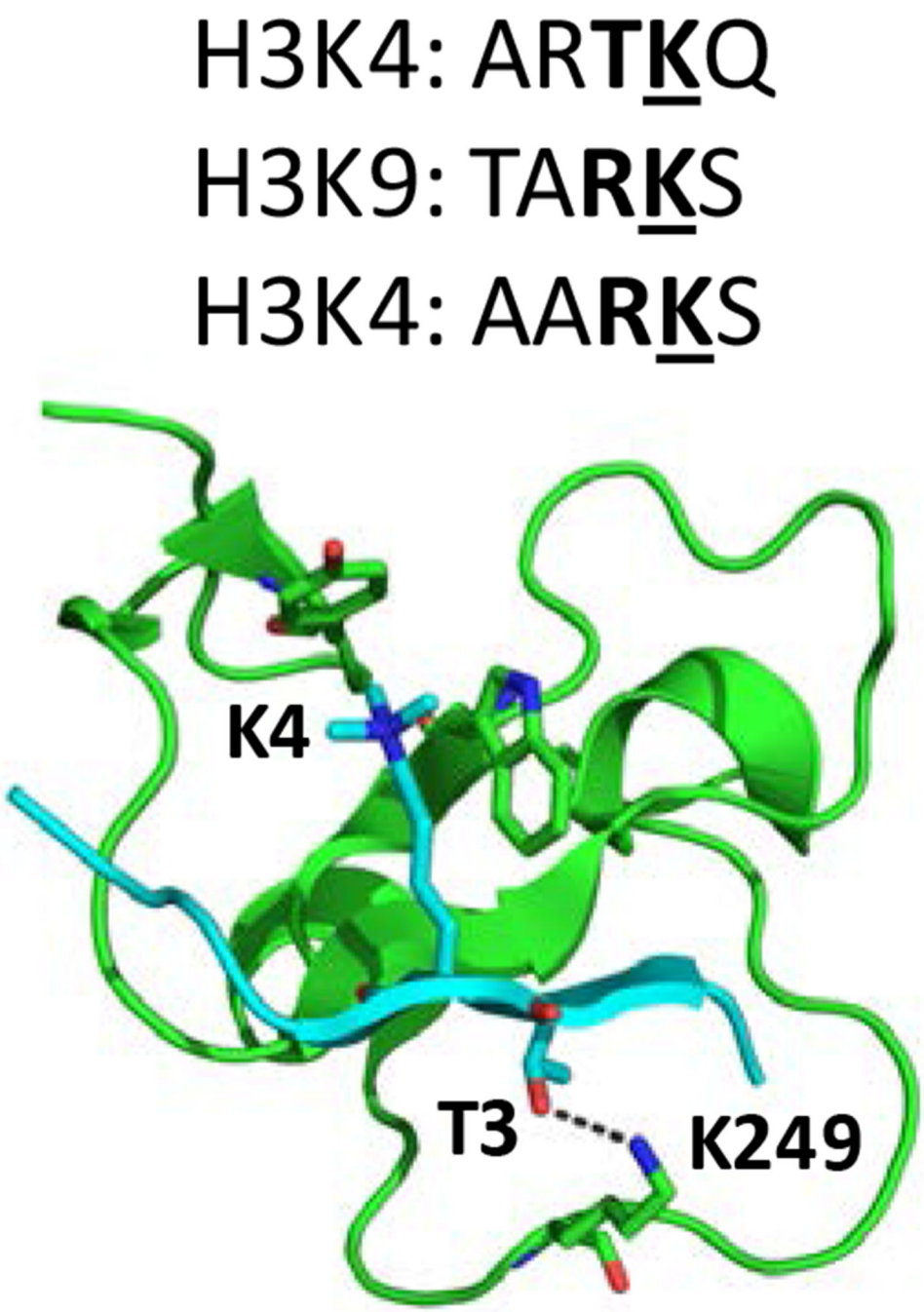


Figure 23.
Structure of the PHD finger of ING2 and its preferential ligand. The PHD finger of ING2 (PDB 2G6Q) recognizes H3K4me3, which is separated from H3R2 by Thr3. In contrast, this PHD finger disfavors H3K9me3 and H3K27me3, which are adjacent to an Arg residue (R8/R26).

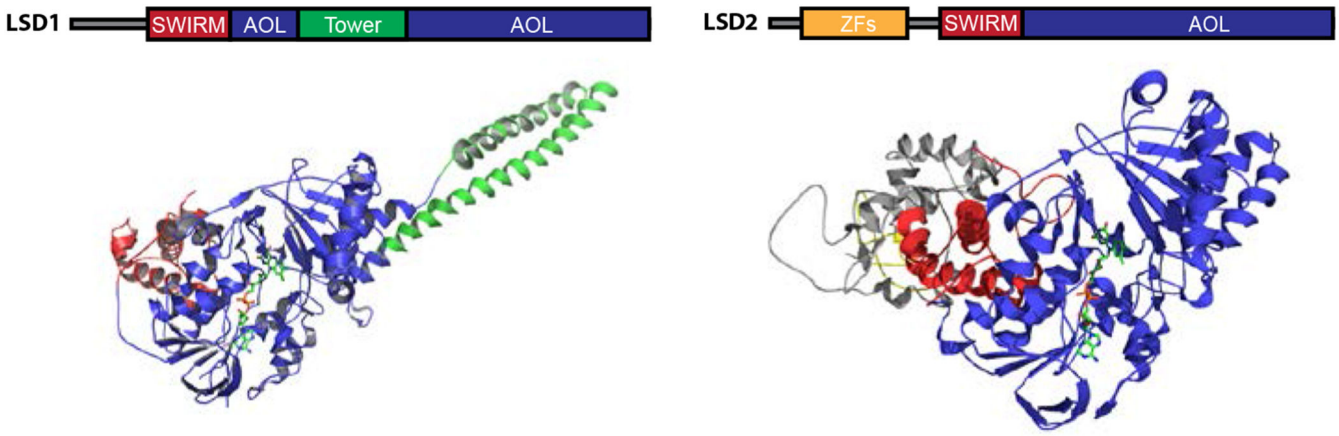


Figure 24.

Topology and 3-D structures of LSD1 and LSD2. The structures of LSD1/2 were generated on the basis of PDB files 2H94 and 4GU0 with PyMOL, respectively.(237, 238)

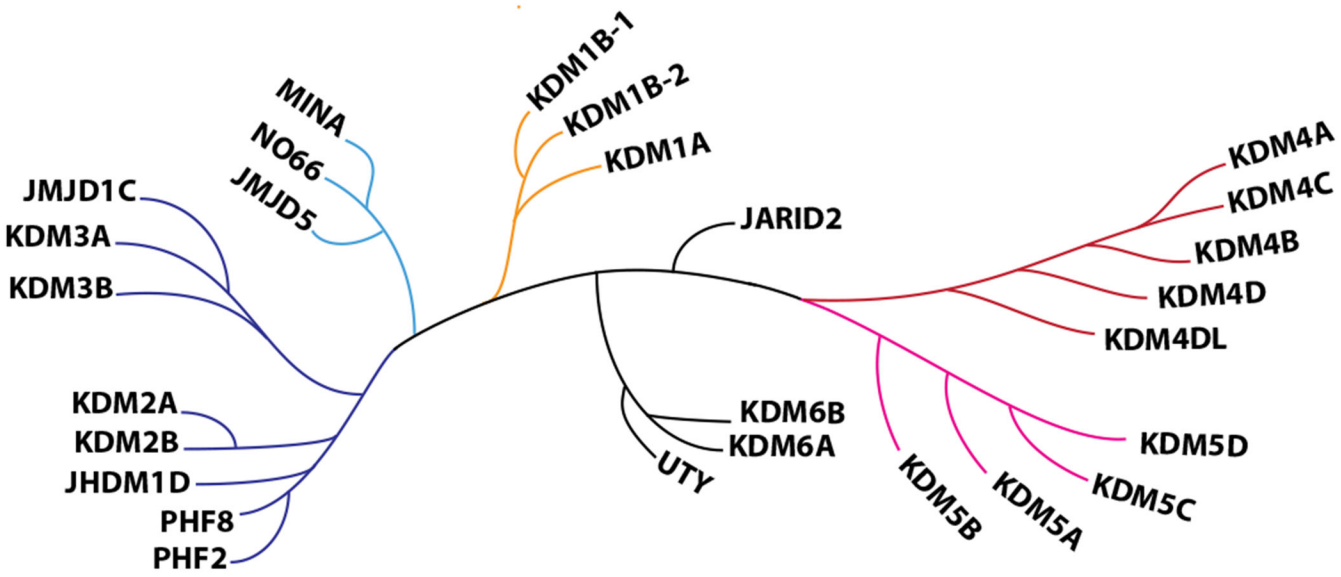


Figure 25.
Phylogenic tree of JmjC-domain-containing human KDMs. Classification and relative positions of KDMs were referred in a previous report.(33–35)

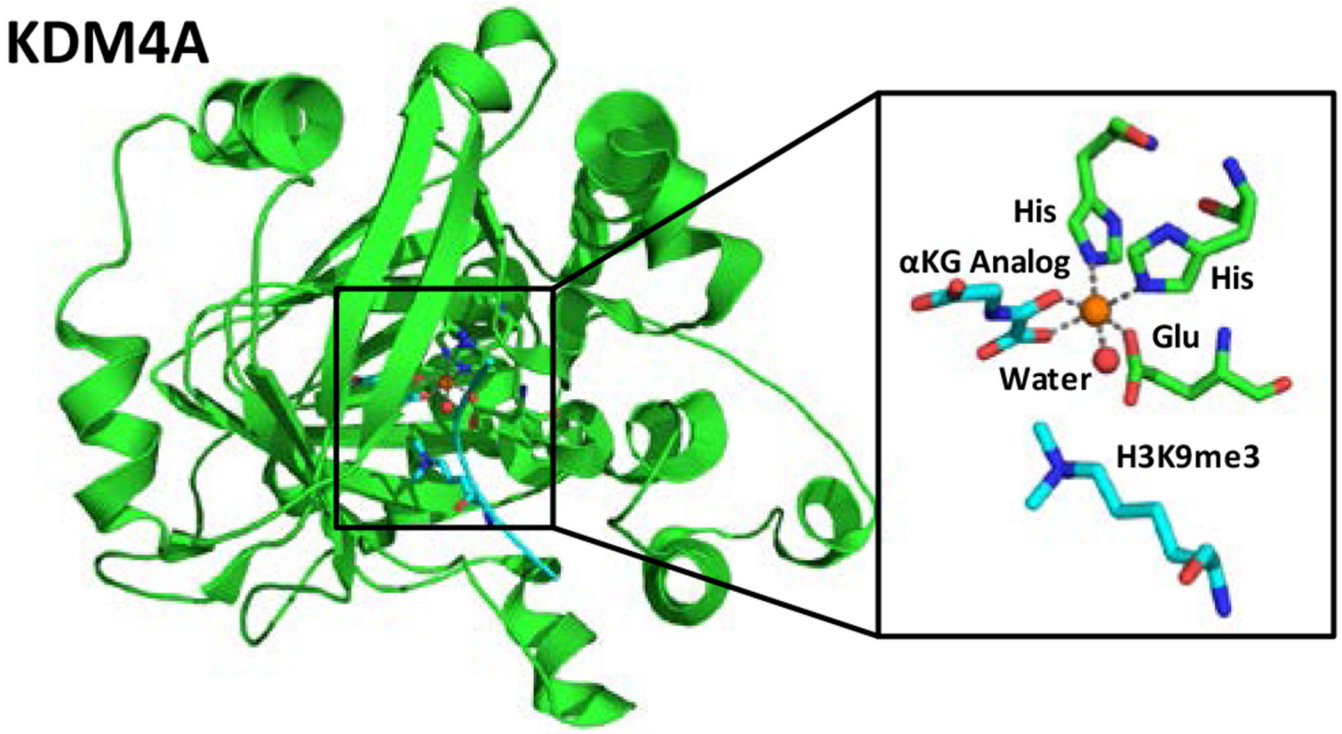


Figure 26.
Representative structure and catalytic site of JmjC domain-containing KDMs. KDM4A (PDB 2OQ6) is shown here as an example.

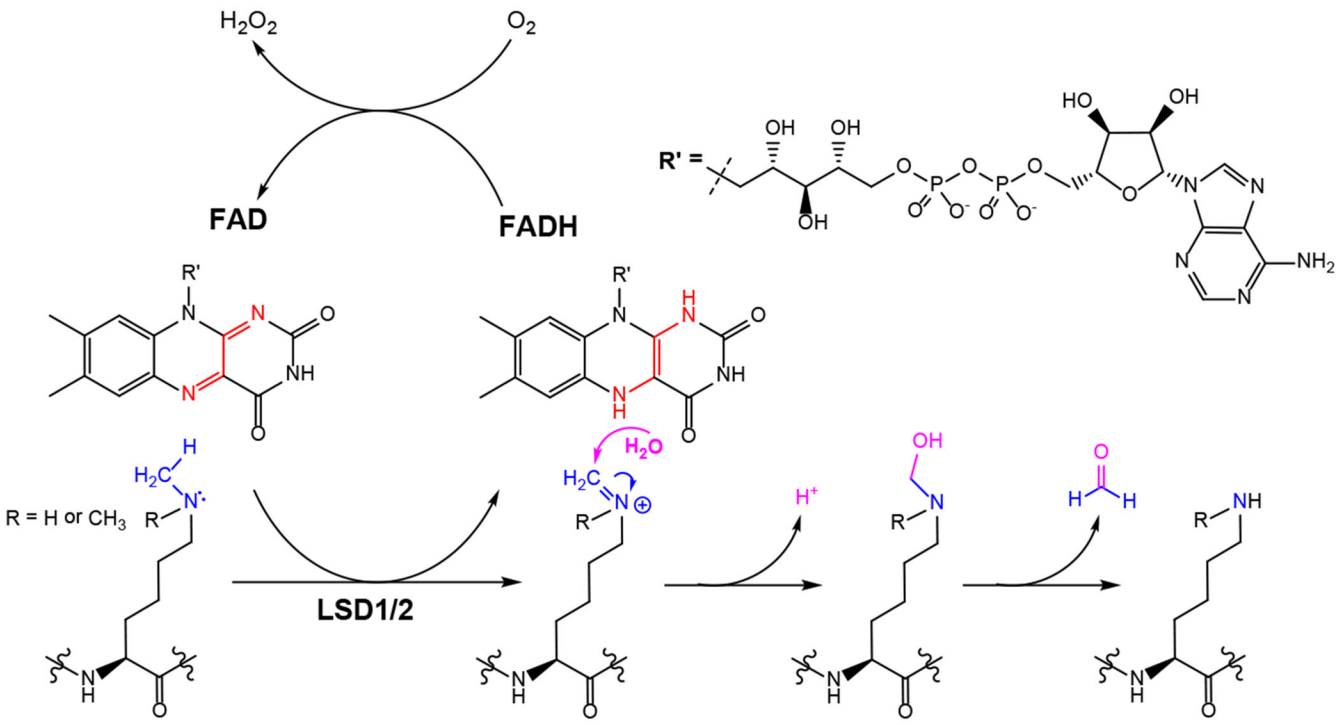


Figure 27.
Chemical mechanism of demethylation reaction catalyzed by LSDs.

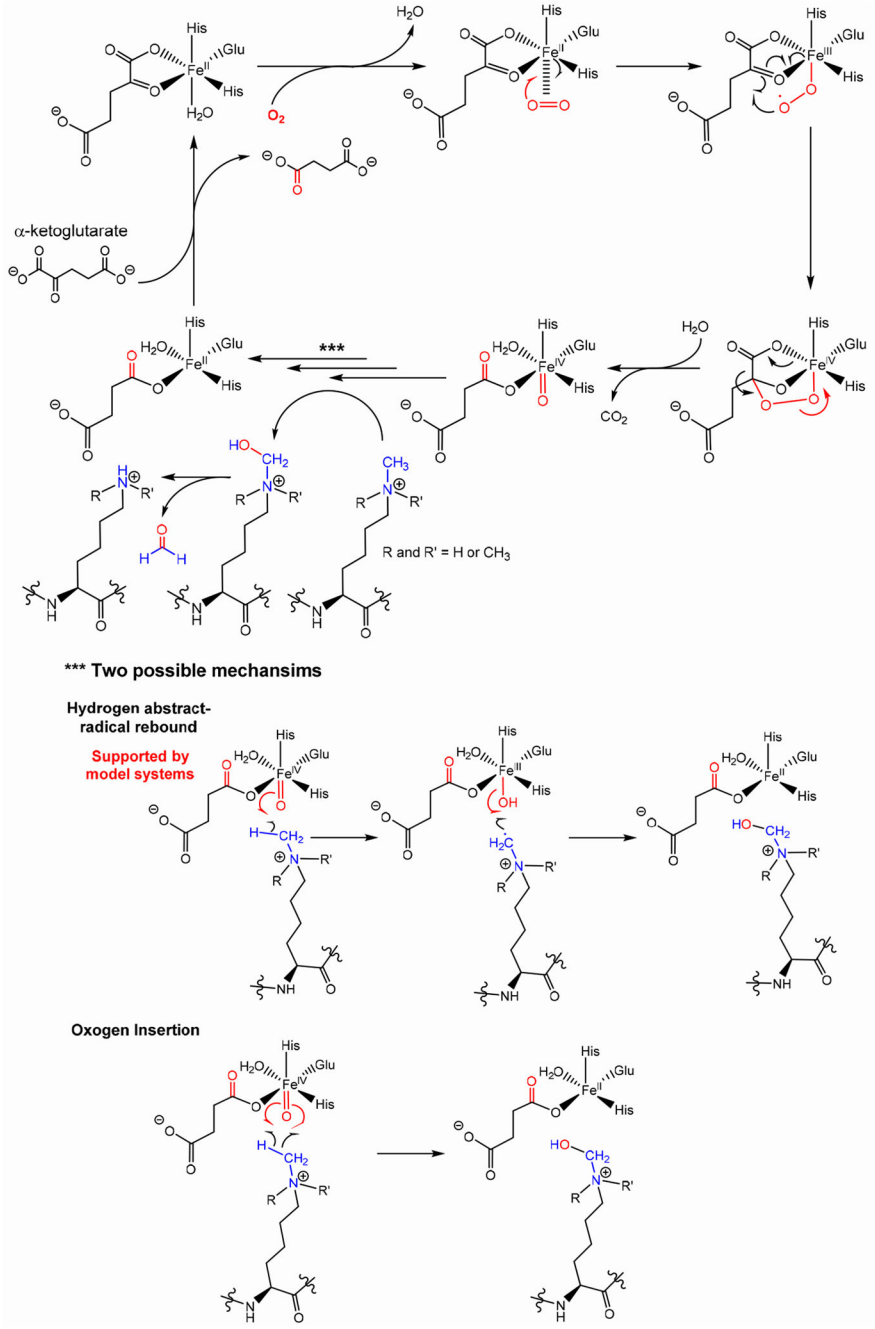


Figure 28.

Proposed chemical mechanism of a demethylation reaction catalyzed by JmjC-domain-containing KDMs. Two possible mechanisms: a stepwise radical reaction versus a concerted reaction. The stepwise radical mechanism is consistent with a model system.(242)

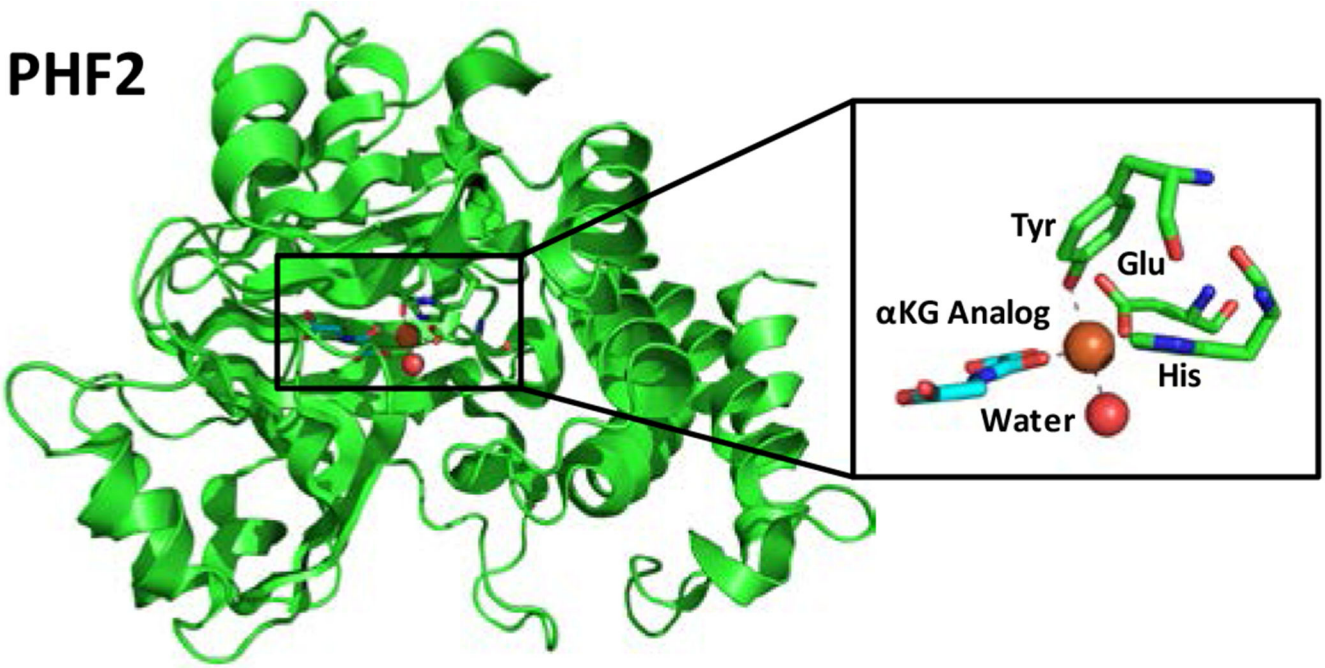


Figure 29.

Structure and distinct catalytic site of the H3K9me2 demethylase PHF2. PHF2 appears to use a Tyr (Y321) instead of the conserved His of the facial triad in other KDMs to coordinate the Fe cofactor.

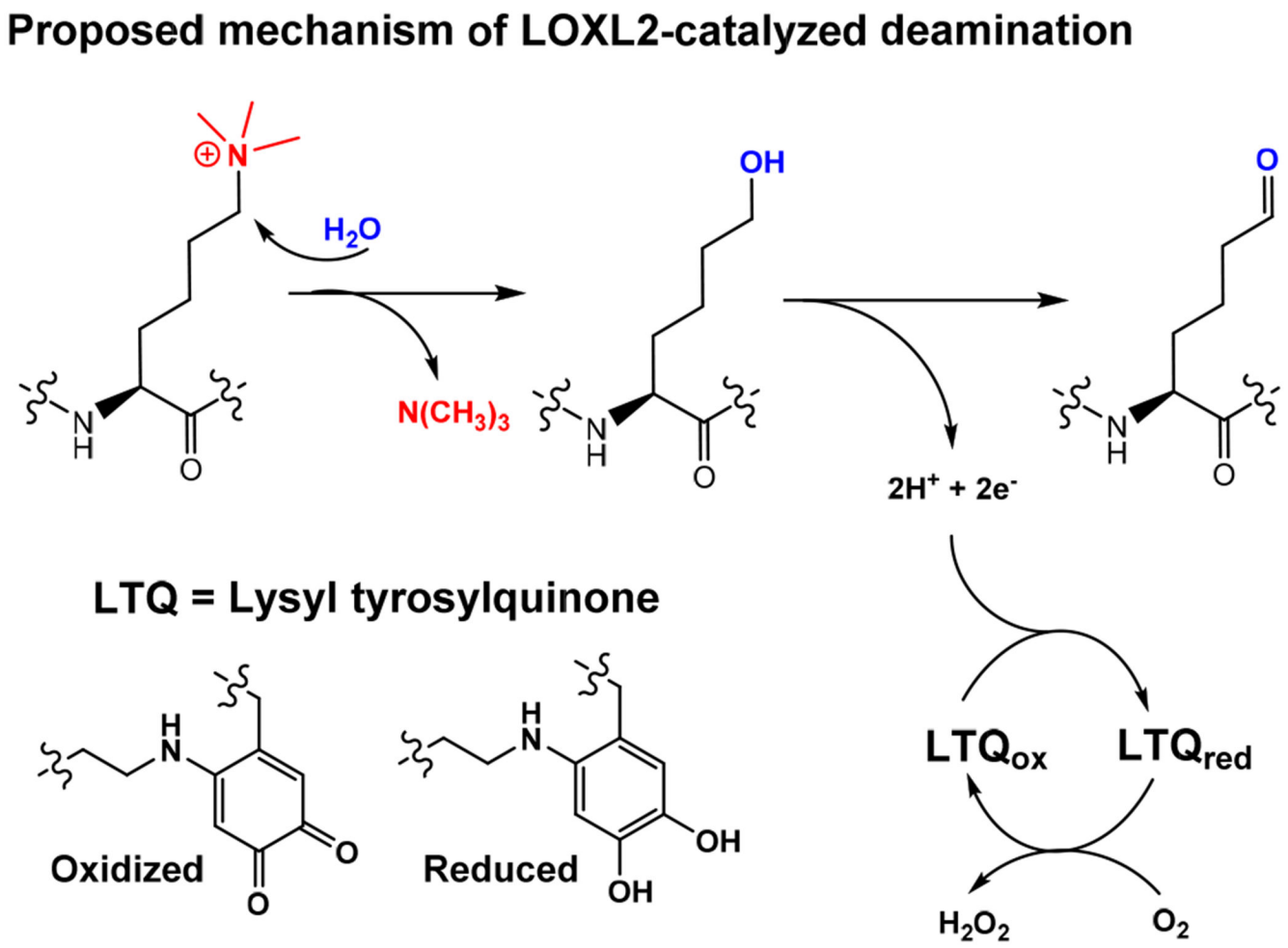


Figure 30.
Proposed catalytic cycle of a LOXL2-catayzled deamination reaction.

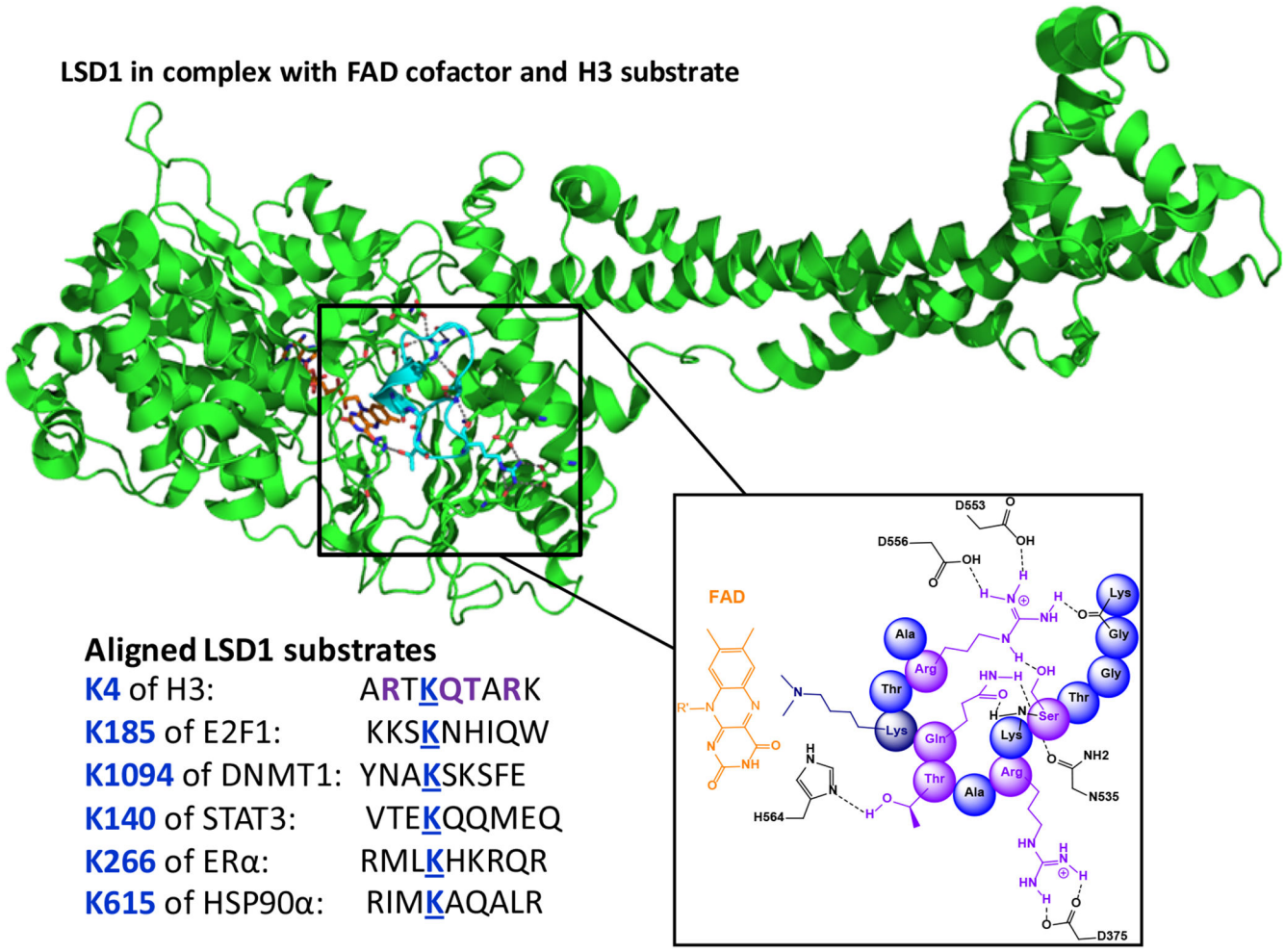


Figure 31.

Structure, specific long-range substrate recognition and promiscuous substrate sequences of LSD1.

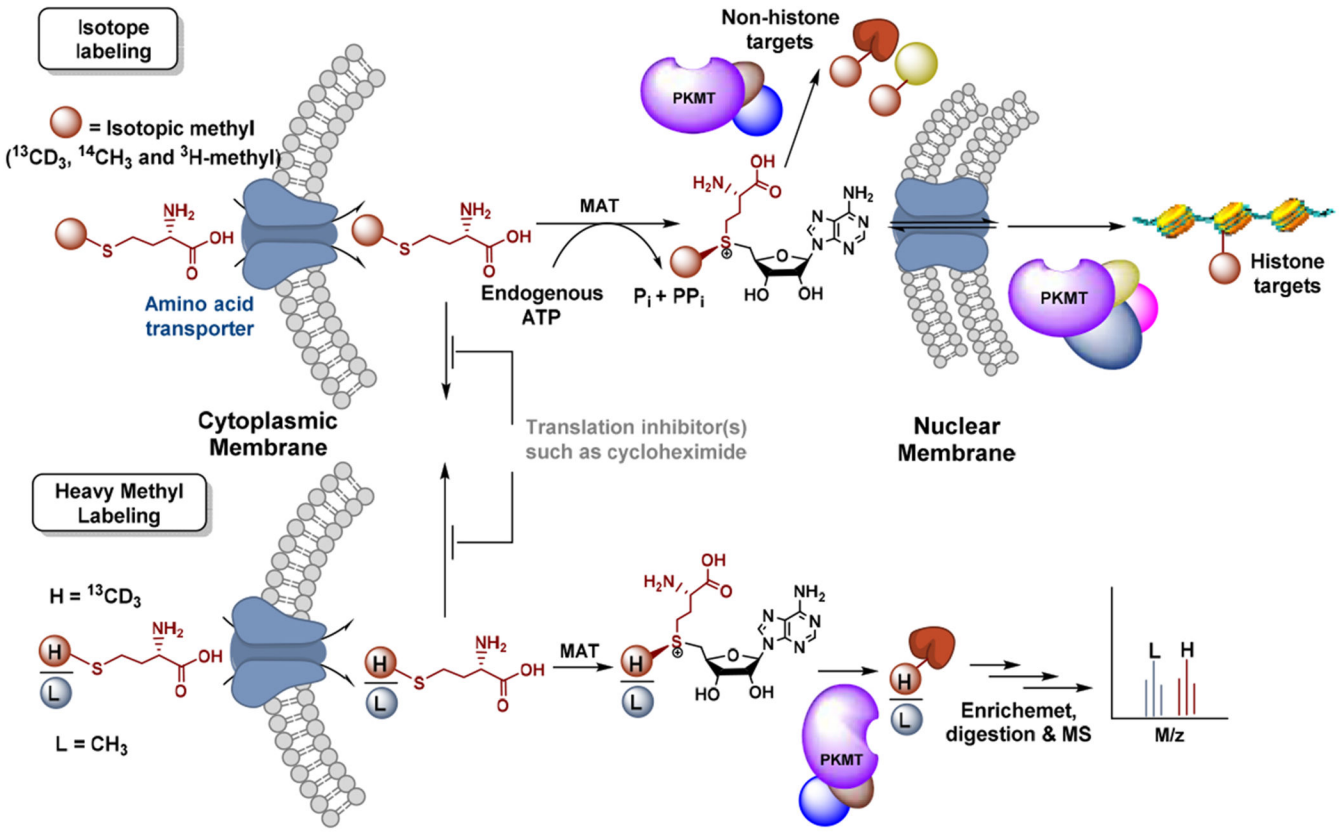


Figure 32.

Isotopical labeling of methylomes within living cells. To label PKMT targets with isotopic SAM cofactors in living cells, *S*-[heavy-methyl] methionine can be used as the substrates of MAT. The resultant biosynthesized SAM will be used as cofactors by native PMTs to label their substrates. Alternatively, a mixture of *S*-[heavy-methyl] L-methionine and unlabeled L-methionine with known ratios can be used. The resultant labeled PKMT targets can be identified upon detecting the light/heavy pair of labeled products in a quantitaive manner. Adapted from *Epigenetic Technological Applications*, Luo M., "Chapter 10: Current Methods for Methylome Profiling" 187-212, Copyright 2015 with permission from Elsevier.

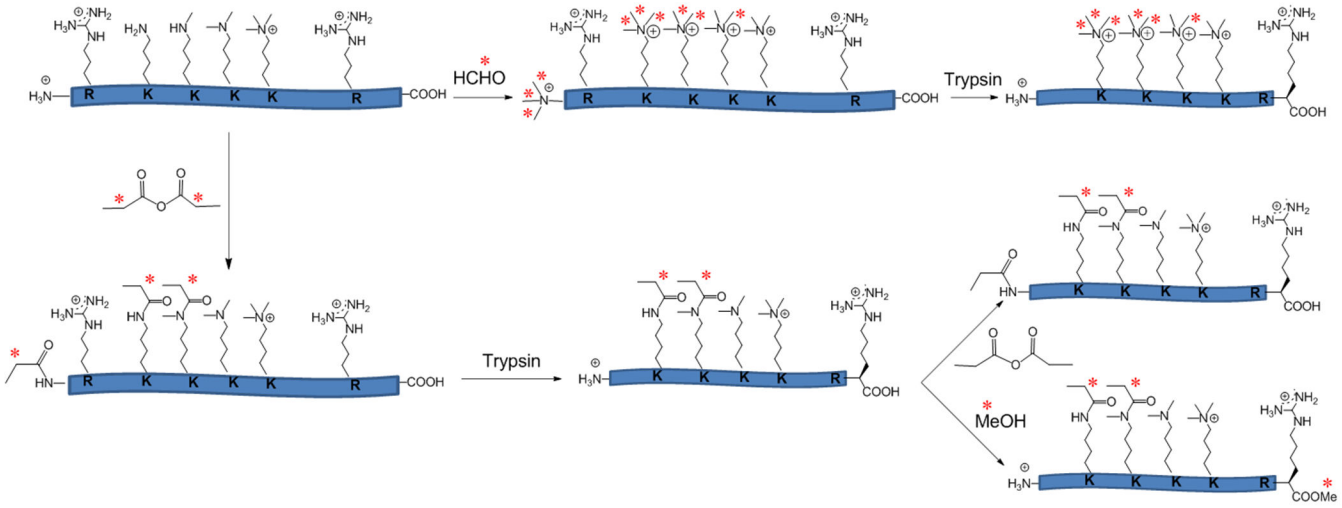


Figure 33.

Methylome derivatization with isotopic bar codes. Characteristic light-to-heavy isotopic mass shifts can be introduced at various stages of sample processing. Their relative MS ratios will be used for MS quantification. Free amine group can be labeled with isotopic propionic anhydride or formaldehyde. Free carboxylic group can be labeled with isotopic methanol. *Positions for isotopic labeling. Adapted from *Epigenetic Technological Applications*, Luo M., "Chapter 10: Current Methods for Methylome Profiling" 187-212, Copyright 2015 with permission from Elsevier.

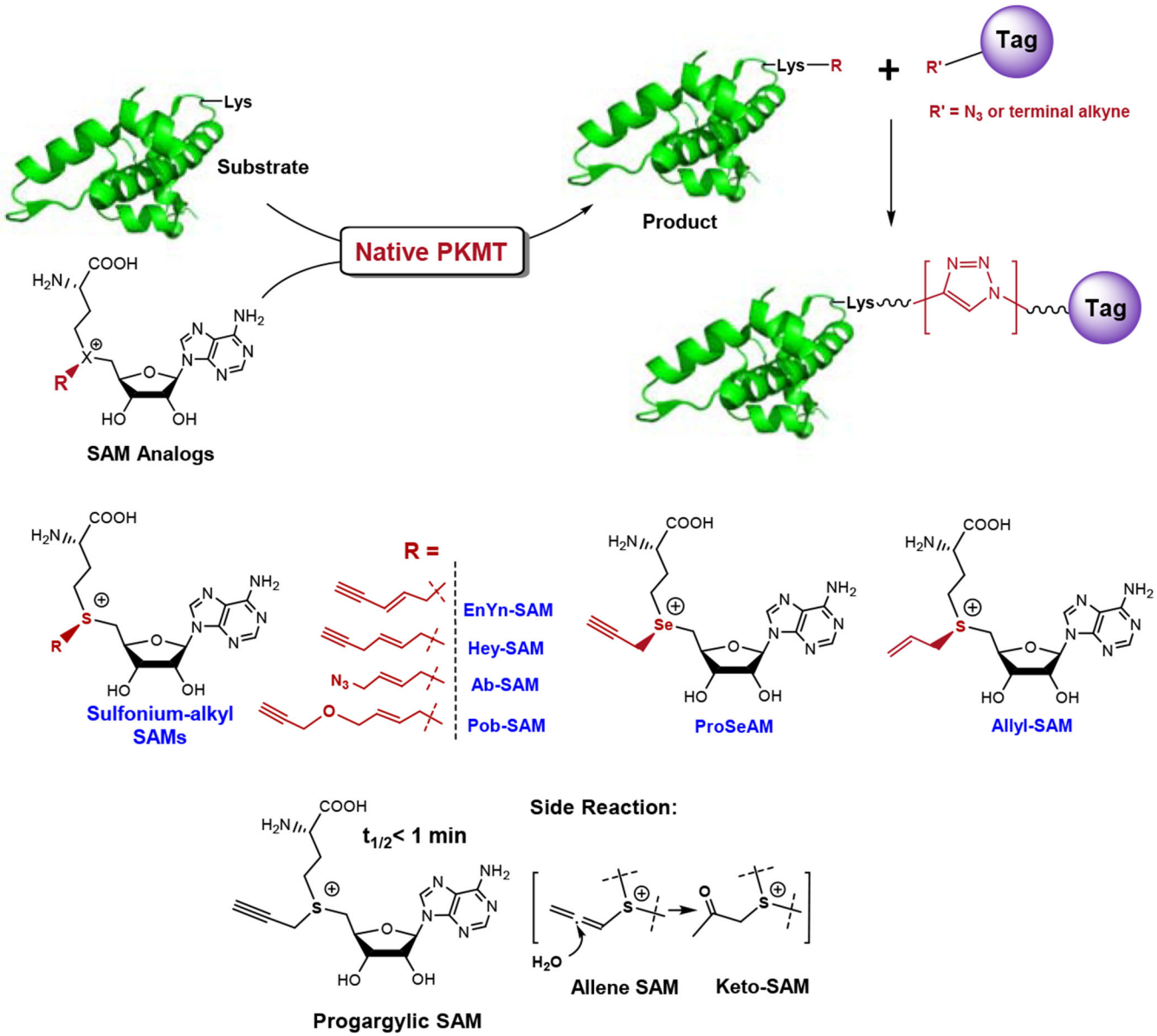


Figure 34.

Methylome chemical labeling with SAM analogs. SAM analog cofactors containing clickable sulfonium-alkyl moieties such as terminal alkyne or azide groups can be processed by some native PKMTs for substrate labeling. The terminal alkyne or azide groups feature their ready conjugation with other probes (*e. g.* dye and biotin) via the well-established Huisgen cycloaddition reaction (the click chemistry).

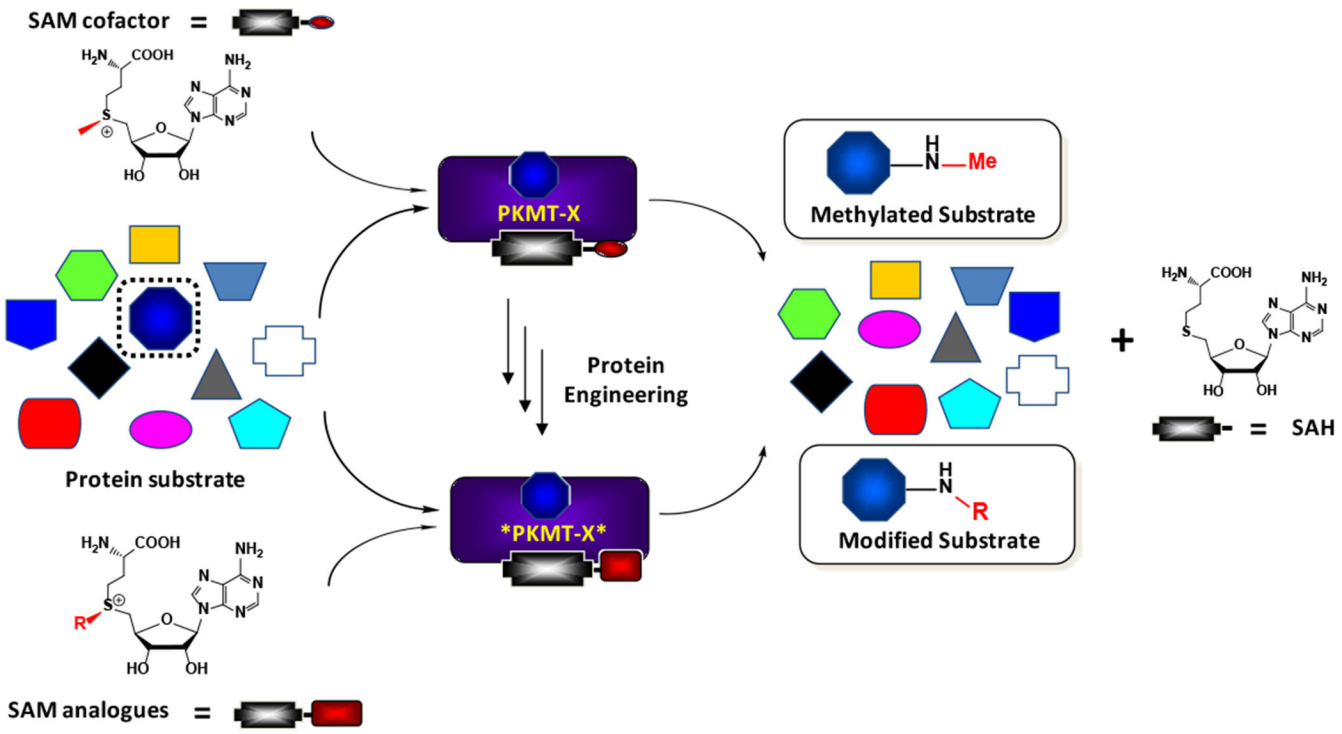


Figure 35.

Bioorthogonal Profiling of Protein Methylation (BPPM). For BPPM, SAM-binding pockets of PKMTs are engineered to accommodate *S*-alkyl SAM analogs, which are otherwise too bulky to serve as cofactors of native PKMTs. The engineered PKMTs can then transfer the distinct alkyl moieties to the substrates of native PKMTs. The BPPM approach allows the distinctly-labeled targets to be assigned to individual (engineered) PKMTs in an unambiguous manner. Reproduced from *ACS Chem. Biol.* **2012**, *7*, 443-463, Copyright 2012 American Chemical Society.

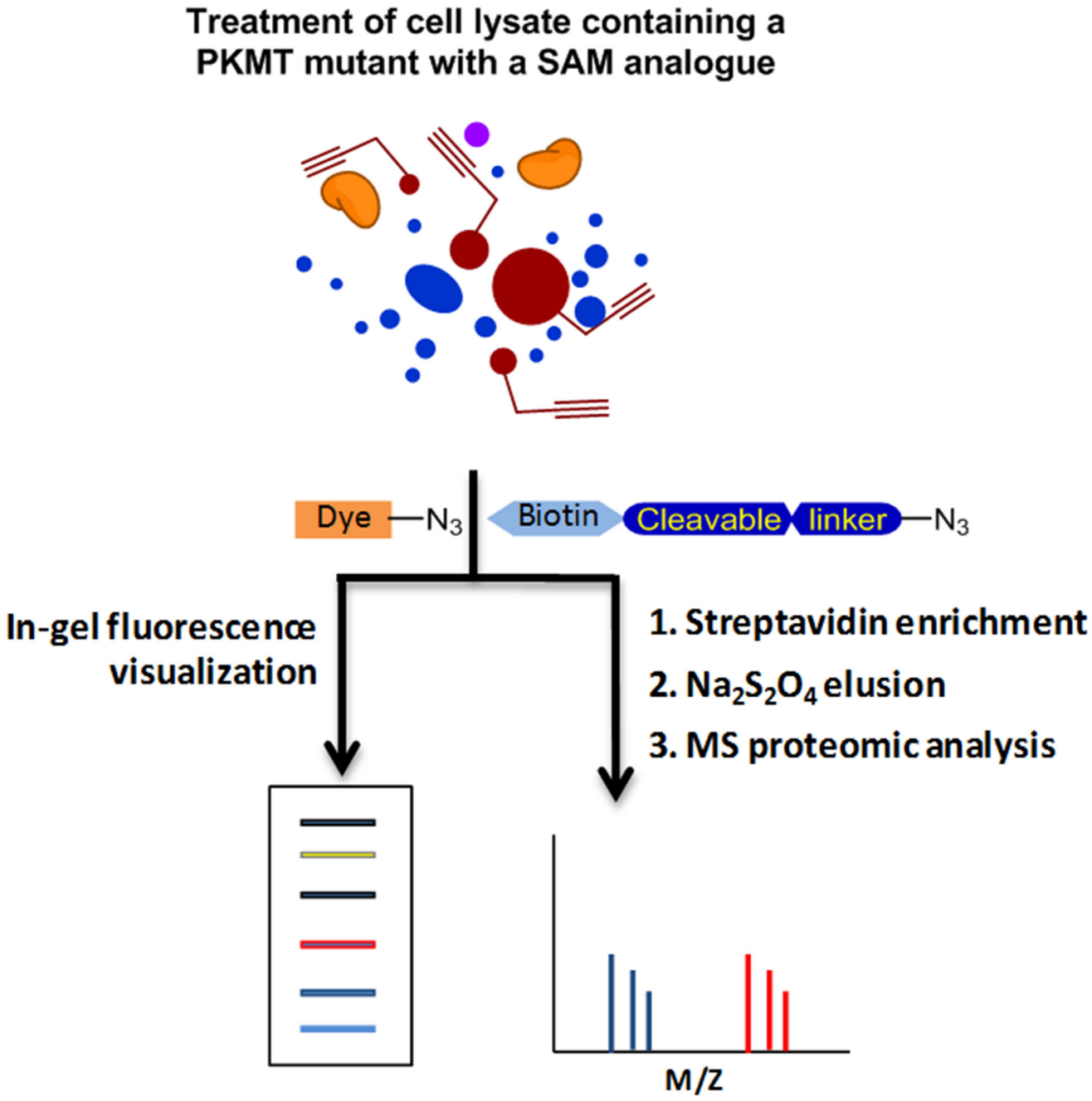


Figure 36.

Workflow of methylome profiling by BPPM. Cells were transfected with a PKMT mutant plasmid and then lysed, followed by treatment with SAM analog cofactors. BPPM-labeled targets were then conjugated with fluorescent dyes for in-gel fluorescence or with cleavable azido-azo-biotin probes for target enrichment. Adapted from *Epigenetic Technological Applications*, Luo M., "Chapter 10: Current Methods for Methylome Profiling" 187-212, Copyright 2015 with permission from Elsevier.

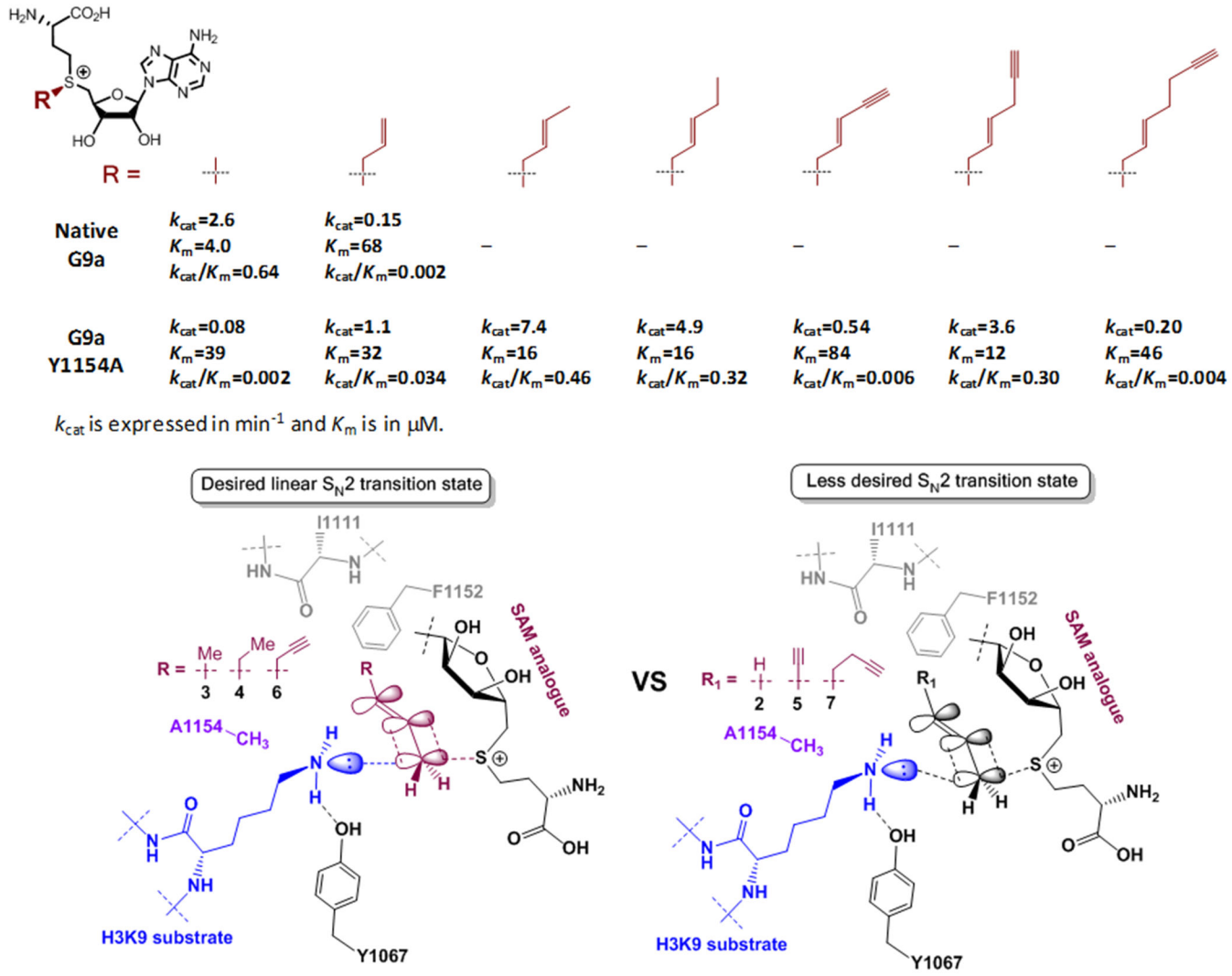


Figure 37.

Structure-activity-relationship for efficient BPPM. Engineered G9a was used as an example to rationalize how structurally matched G9a variants and SAM analog cofactors facilitate target labeling by increasing k_{cat} rather than decreasing K_m. Such an observation is expected to associate with more readily assembled transition states for structurally matched enzyme-cofactor pairs. Adapted with permission from Ref. (99) Copyright 2013 PNAS.

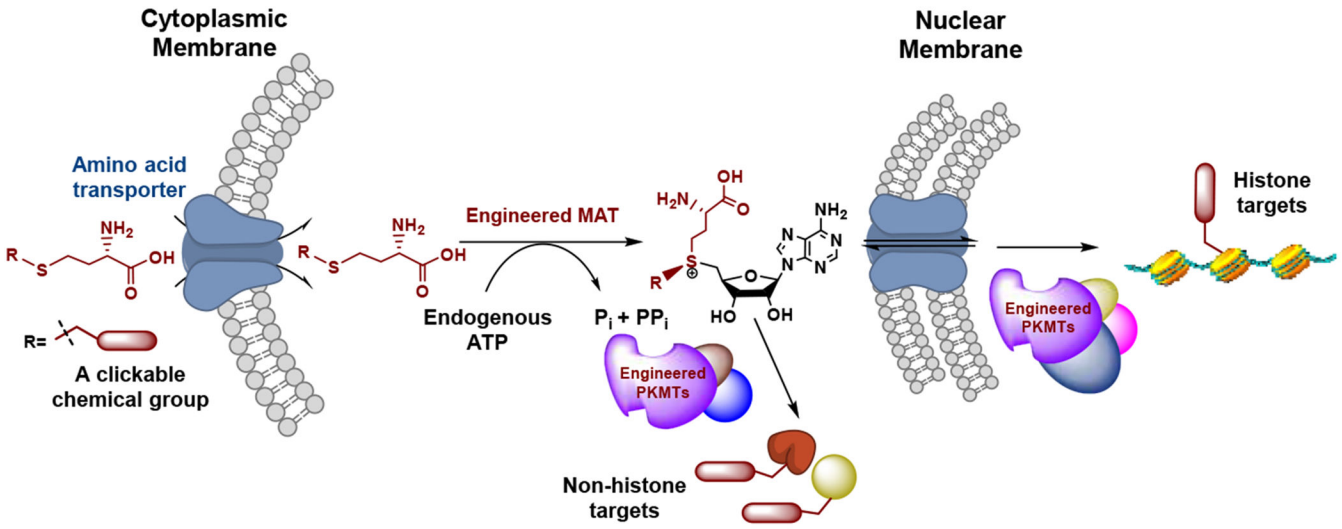


Figure 38.

BPPM in living cells. SAM biosynthetic pathway is hijacked by engineered MATs to process membrane-permeable *S*-alkyl methionine analogs for *in situ* production of the corresponding *S*-alkyl SAM analogs. The three-step BPPM within living cells consist of the biosynthesis of SAM analogs from methionine analog precursors by engineered MATs, *in situ* target labeling by engineered PKMTs, and subsequent enrichment of the distinct modified targets via the click chemistry. Adapted from *Epigenetic Technological Applications*, Luo M., “Chapter 10: Current Methods for Methylome Profiling” 187-212, Copyright 2015 with permission from Elsevier.

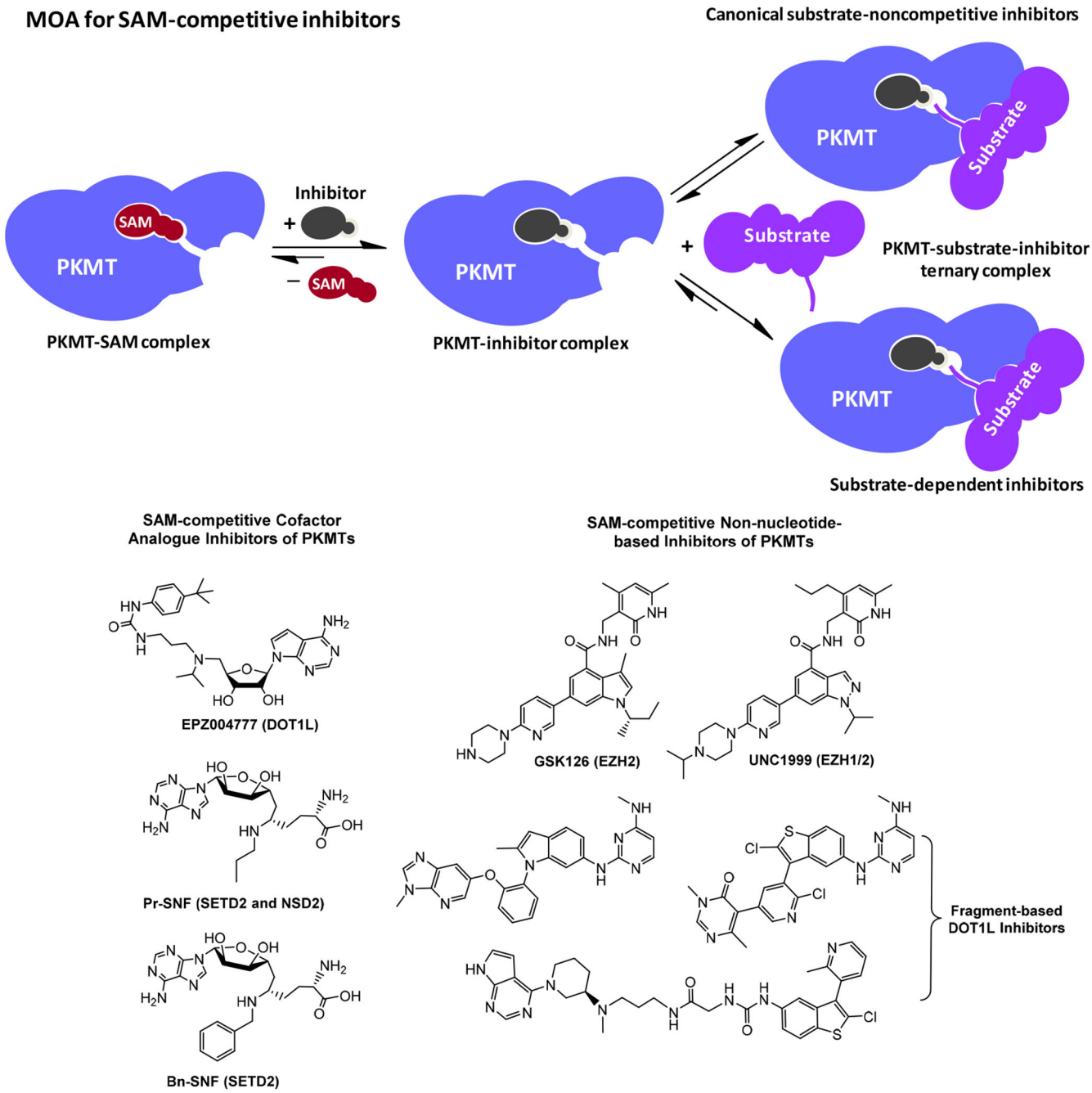


Figure 39.

MOA and representative structures of SAM-competitive PKMT inhibitors. The PKMT targets of these inhibitor are shown in parenthesis. Partially adapted with permission from Ref. (38) Copyright 2015 Future Medicine Ltd.

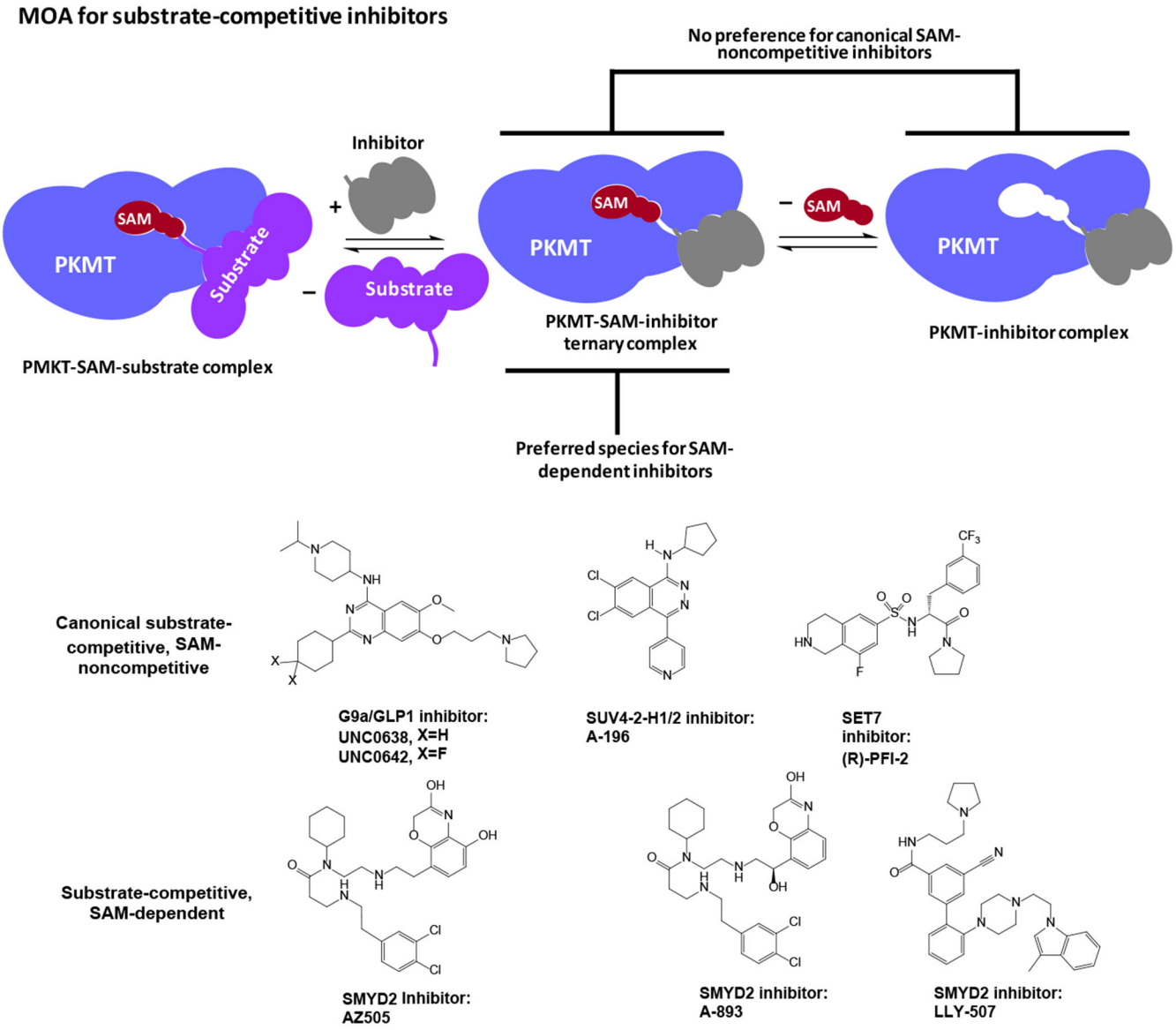


Figure 40.

MOA and representative structures of substrate-competitive inhibitors. These inhibitors can be further classified as SAM-noncompetitive or SAM-dependent inhibitors on the basis of their ability to form more stable PKMT-SAM-inhibitor ternary complexes. Partially adapted with permission from Ref. (38) Copyright 2015 Future Medicine Ltd.

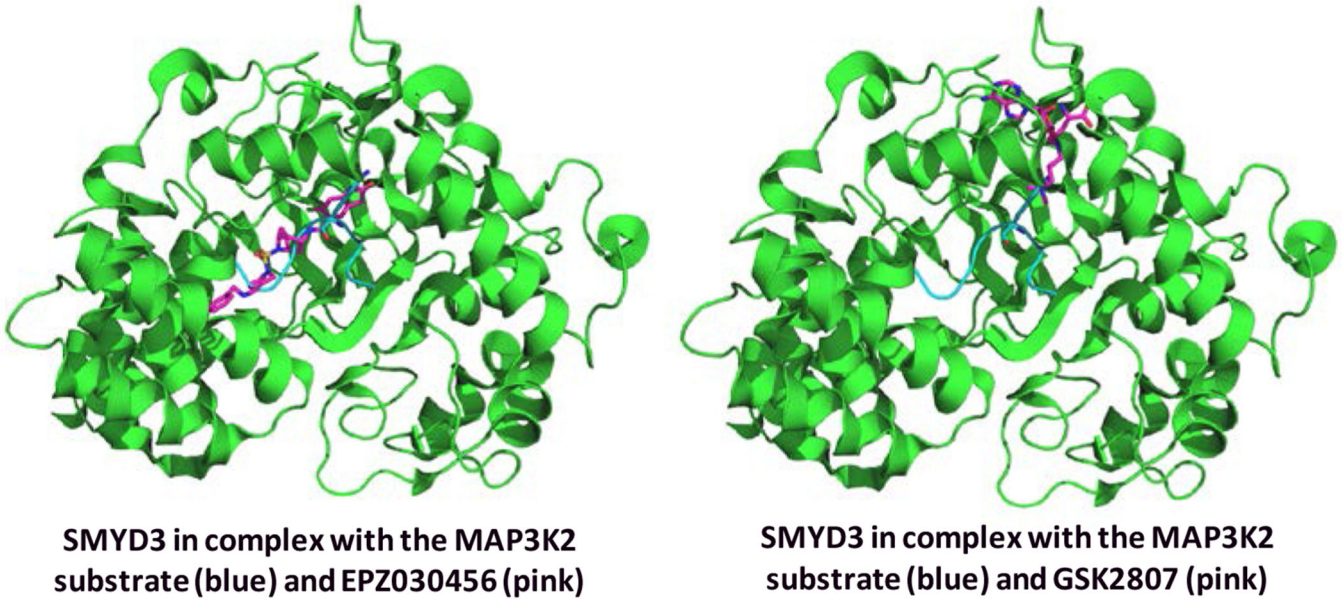


Figure 41.

Structures of SMYD3 in complex with its substrate and inhibitors. The overall structure of SMYD3 (green) is displayed with its substrate MAP3K2 (blue) and the SMYD3 inhibitors EPZ030456 (pink in left) and GSK2807 (pink in right). The two inhibitors show potential steric clashes with the substrate.

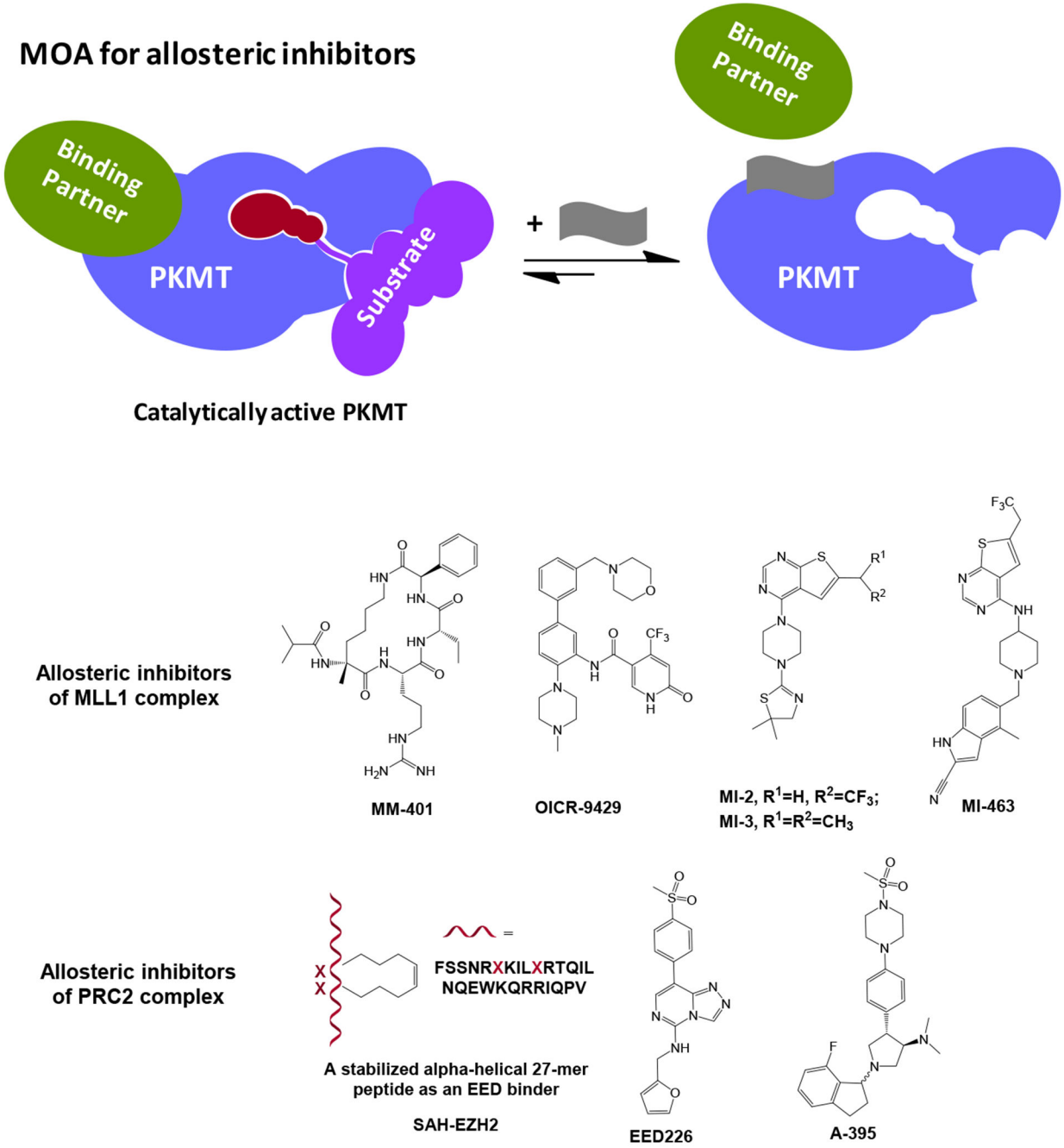


Figure 42.

MOA of allosteric inhibitors of PKMTs and their representative structures. Partially adapted with permission from Ref. (38) Copyright 2015 Future Medicine Ltd.

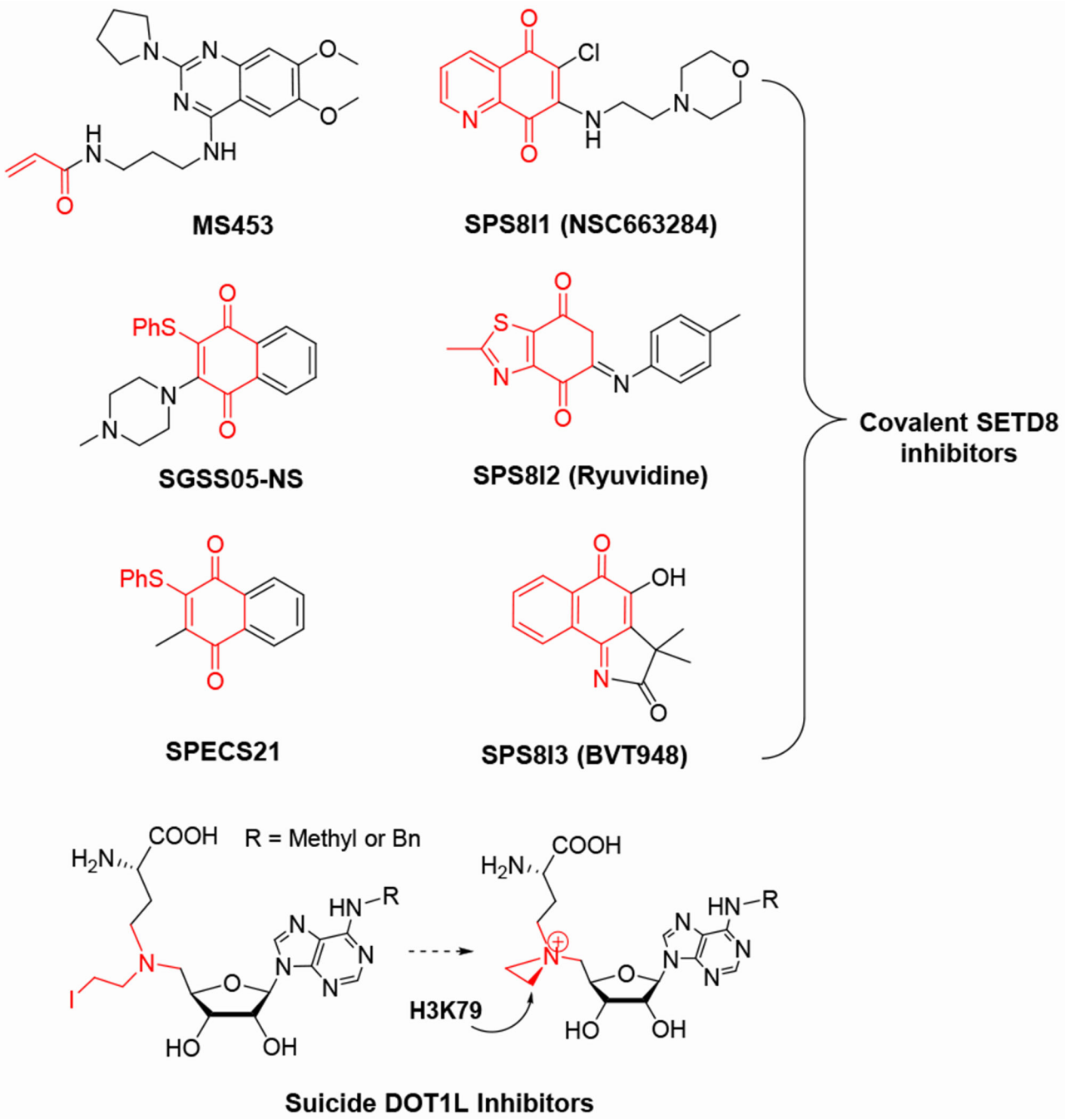


Figure 43.
Representative structures of covalent inhibitors of SETD8 and a suicide inhibitor of DOT1L.

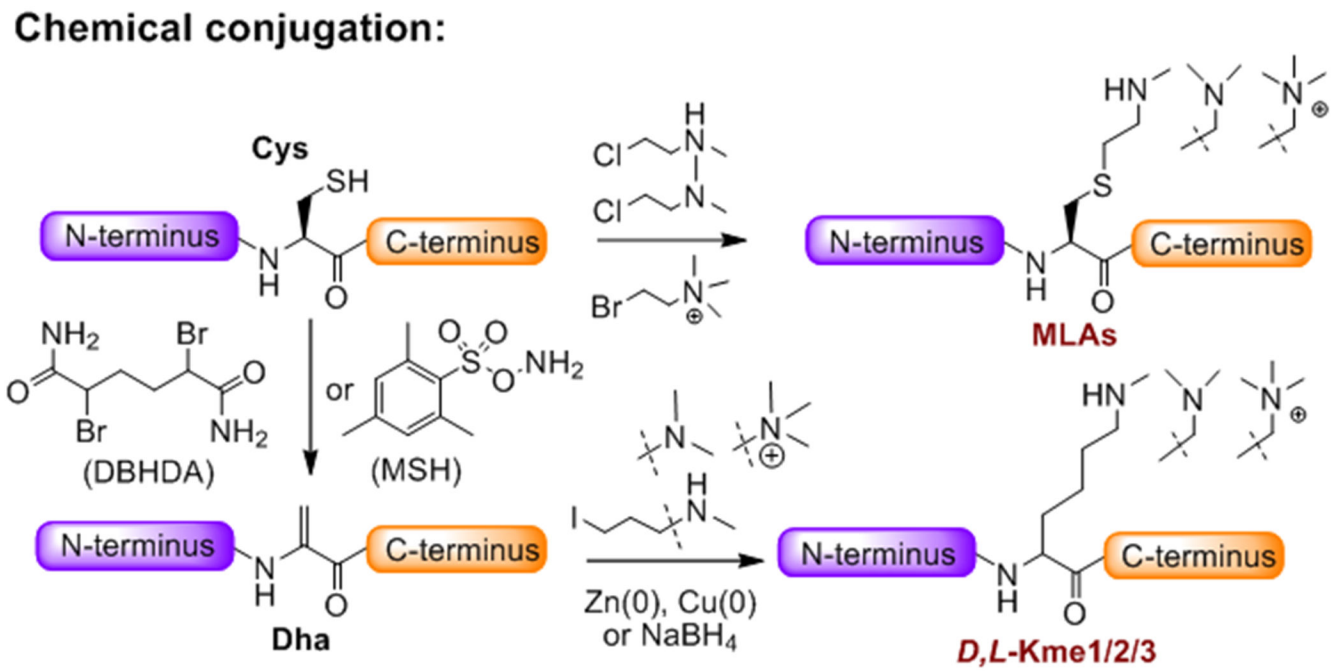


Figure 44.

Semi-synthesis of proteins containing site-specific methyllysine analogs via direct chemical conjugation. Cys and Dha are the sites allowing chemical incorporation of methyllysine side chains into proteins.

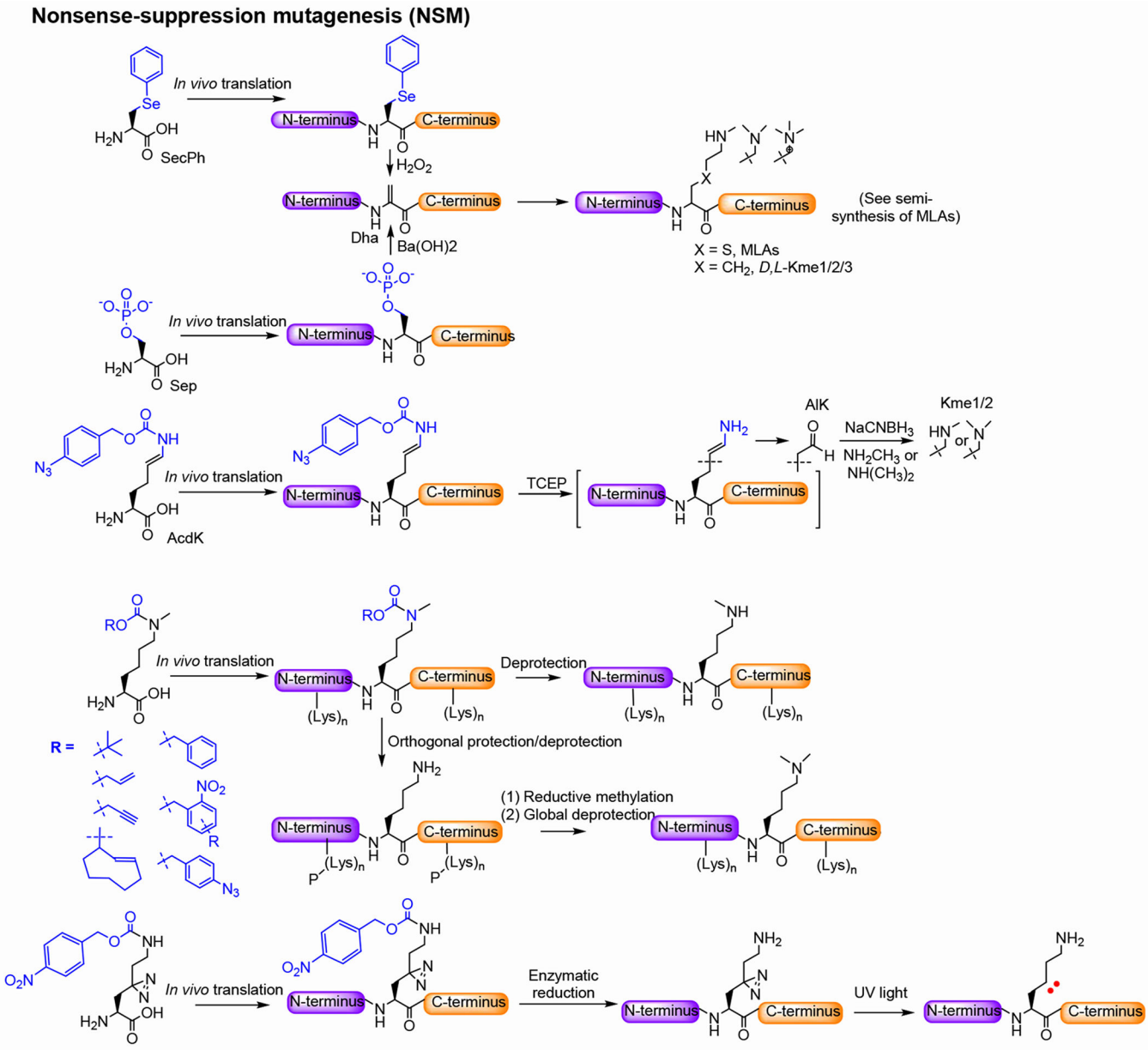


Figure 45.

Biosynthesis of proteins containing site-specific methyllysine residues or methyllysine analogs via nonsense-suppression mutagenesis. A dozen of nonnatural methyllysine precursors can be incorporated into proteins through nonsense-suppression mutagenesis and then converted into corresponding methyllysine or methyllysine analogs.

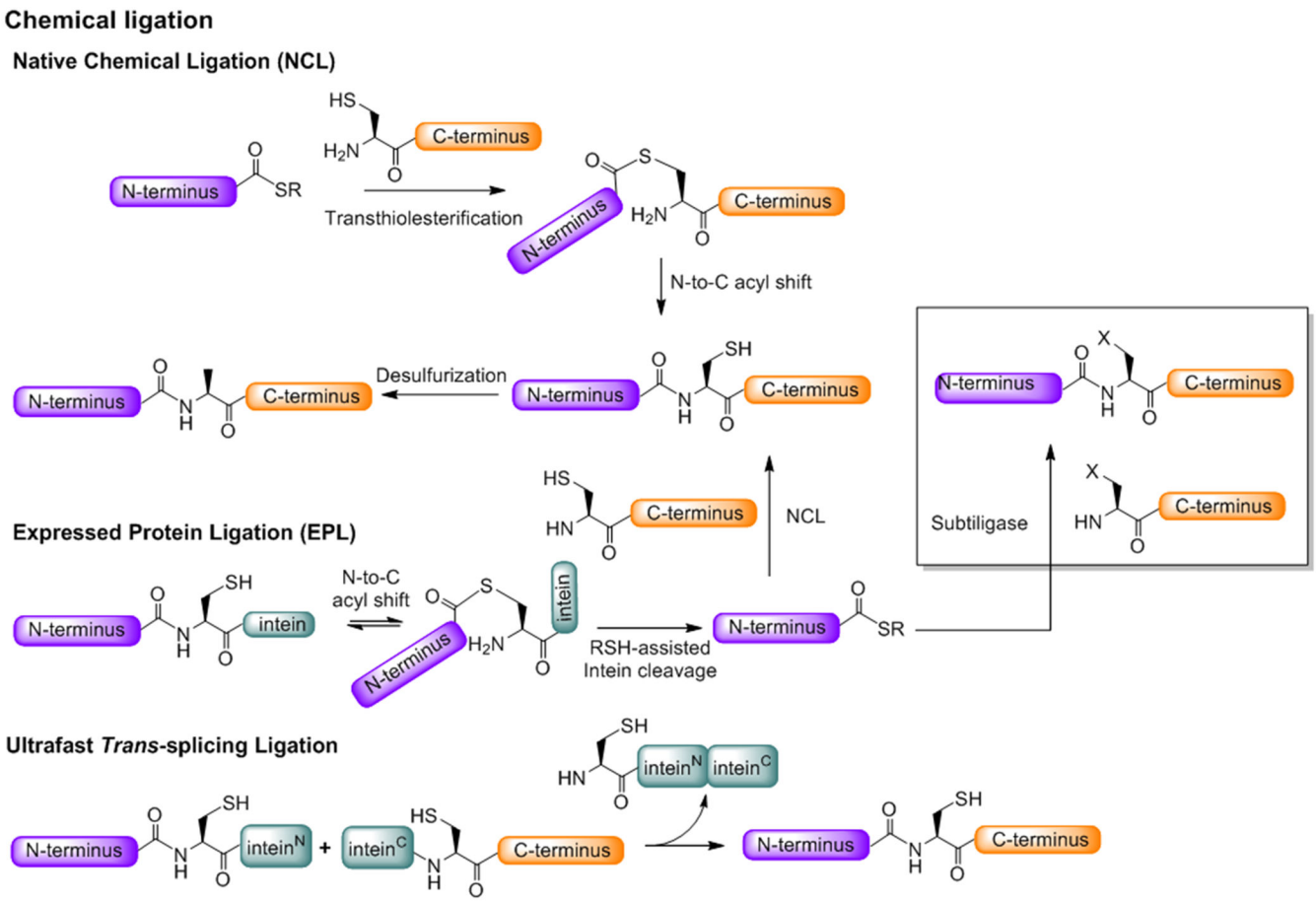


Figure 46.

Biosynthesis of proteins containing site-specific methyllysine residues via diverse chemical ligation strategies. Native chemical ligation, expressed protein ligation and *ultrafast* trans-splicing ligation can be implemented to incorporate methyllysine-containing truncated peptides into full-length products.

Table 1.

Nomenclature of human PKMTs.

Classification		Common Abbreviation	Full Name	Standardized Name	Common Alternatives
SET domain PKMTs	Canonical PKMTs	SUV39H1	Suppressor of variegation 3-9 homolog 1	KMT1A	
		SUV39H2	Suppressor of variegation 3-9 homolog 2	KMT1B	
		G9a	N.A.	KMT1C	EuHMTase2
		GLP1	G9a-like protein 1	KMT1D	EuHMTase1
		SETDB1	SET domain bifurcated 1	KMT1E	ESET
		SETDB2	SET domain bifurcated 2	KMT1F	CLLL8
		SETD8	SET domain-containing protein 8	KMT5A	Pr-SET7
		SUV4-20H1	Suppressor of variegation 4-20 homolog 1 (Su(var)4-20 homolog 1)	KMT5B	
		SUV4-20H2	Suppressor of variegation 4-20 homolog 2 (Su(var)4-20 homolog 2)	KMT5C	
		SET7	SET domain-containing protein 7	KMT6	SET7/9
		SETD3	SET domain-containing protein 3	SETD3	
		SETD4	SET domain-containing protein 4	SETD4	
		SETD5	SET domain-containing protein 5	SETD5	
		SETD6	SET domain-containing protein 6	SETD6	
	Auto-inhibitory SET domain	ASH1L	Absent small and homeotic disks protein 1 homolog	KMT2F	
		SETMAR	SET domain and mariner transposase fusion gene-containing protein	SETMAR	Metnase
		SETD2	SET domain-containing protein 2	KMT3A	HYPB
		NSD1	Nuclear receptor-binding SET domain-containing protein 1	KMT3B	
		NSD2	Nuclear SET domain-containing protein 2	NSD2	WHSC1/MMSET
		NSD3	Nuclear SET domain-containing protein 3	NSD3	
	PKMTs active in complexes	MLL	Myeloid/lymphoid or mixed-lineage leukemia protein	KMT2A	ALL1/HRX
		MLL2	Myeloid/lymphoid or mixed-lineage leukemia protein 2	KMT2B	ALR
		MLL3	Myeloid/lymphoid or mixed-lineage leukemia protein 3	KMT2C	
		MLL4	Myeloid/lymphoid or mixed-lineage leukemia protein 4	KMT2D	
		MLL5	Myeloid/lymphoid or mixed-lineage leukemia protein 5	KMT2E	
		SET1A	SET domain-containing protein 1A	KMT2F	
		SET1B	SET domain-containing protein 1B	KMT2G	
		EZH2	Enhancer of zeste homolog 2	KMT6	
		EZH1	Enhancer of zeste homolog 1	EZH1	
	MYND-inserted	SMYD1	SET and MYND domain containing protein 1	KMT3C	
		SMYD2	SET and MYND domain containing protein 2	SMYD2	
		SMYD3	SET and MYND domain containing protein 3	SMYD3	

Classification		Common Abbreviation	Full Name	Standardized Name	Common Alternatives
		SMYD4	SET and MYND domain containing protein 4	SMYD4	
		SMYD5	SET and MYND domain containing protein 5	SMYD5	
	PR-inserted	PRDM1	PR (PRDI-BF1 and RIZ) domain proteins 1	KMT8	BLIMP1
		PRDM2	PR (PRDI-BF1 and RIZ) domain proteins 2	PRDM2	RIZ
		PRDM3	PR (PRDI-BF1 and RIZ) domain proteins 3	PRDM3	MDS1-EV1/MSP1
		PRDM4	PR (PRDI-BF1 and RIZ) domain proteins 4	PRDM4	
		PRDM5	PR (PRDI-BF1 and RIZ) domain proteins 5	PRDM5	
		PRDM6	PR (PRDI-BF1 and RIZ) domain proteins 6	PRDM6	
		PRDM7	PR (PRDI-BF1 and RIZ) domain proteins 7	PRDM7	
		PRDM8	PR (PRDI-BF1 and RIZ) domain proteins 8	PRDM8	
		PRDM9	PR (PRDI-BF1 and RIZ) domain proteins 9	PRDM9	
		PRDM10	PR (PRDI-BF1 and RIZ) domain proteins 10	PRDM10	
		PRDM11	PR (PRDI-BF1 and RIZ) domain proteins 11	PRDM11	
		PRDM12	PR (PRDI-BF1 and RIZ) domain proteins 12	PRDM12	
		PRDM13	PR (PRDI-BF1 and RIZ) domain proteins 13	PRDM13	
		PRDM14	PR (PRDI-BF1 and RIZ) domain proteins 14	PRDM14	
		PRDM15	PR (PRDI-BF1 and RIZ) domain proteins 15	PRDM15	
		PRDM16	PR (PRDI-BF1 and RIZ) domain proteins 16	PRDM16	
		PRDM17	PR (PRDI-BF1 and RIZ) domain proteins 17	PRDM17	
Non-SET domain PKMTs		DOT1L	Disruptor of telomeric silencing-1-like	KMT4	
		METTL10	Methyltransferase-like protein 10	METTL10	
		ETFP-KMT	Lysine methyltransferase of the P-subunit of the electron transfer flavoprotein	METTL20	
		METTL21A	Methyltransferase-like protein 21A	METTL21A	
		eEF1A-KMT3	Lysine methyltransferase 3 of human eukaryotic elongation factor 1 alpha	METTL21B	
		METTL21C	Methyltransferase-like protein 21C	METTL21C	
		VCP-KMT	Valosin-containing protein lysine methyltransferase	METTL21D	VCP-KMT
		METTL22	Methyltransferase-like protein 22	METTL22	
		eEF1A-KMT1	Lysine methyltransferase 1 of human eukaryotic elongation factor 1 alpha	eEF1A-KMT1	
		eEF2-KMT	Lysine methyltransferase of human eukaryotic elongation factor 2	eEF2-KMT1	FAM86A
		CaM-KMT	Calmodulin-lysine N-methyltransferase	CaM-KMT	

Table 2.

Nomenclature of 15 classes of methyllysine reader domains and their representative examples.

Classification	Example(s)
ADD	DNMT3L
Ankyrin	G9a/GLP1
BAH	ORC1
Chromo barrel	MSL3
Chromodomain	HP1
Double chromodomain	CHD1
HEAT	Condensin II
MBT	L3MBT1
PHD	TAF3/BPTF/ING2/BHC80
PWWP	hMSH6/NSD2
SAWADEE	SHH1
Tandem tudor domain	53BP1
Tudor	PHF1
WD40	EED
Zf-CW	ZCWPW1