Catalytic Roles for Carbon-Oxygen Hydrogen Bonding in SET Domain Lysine Methyltransferases

Jean-François Couture1, Glenn Hauk1,2, Mark J. Thompson3, G. Michael Blackburn3, and Raymond C. Trievel1

From the (1) Department of Biological Chemistry, University of Michigan, Ann Arbor, MI USA; (2) Current Address: Novartis Institutes for BioMedical Research, Cambridge, MA USA; (3) Department of Chemistry, University of Sheffield, Sheffield, UK.

Running Title: Carbon-Oxygen Hydrogen Bonding in Lysine Methyltransferases

Address correspondence to: Raymond C. Trievel, University of Michigan, Department of Biological Chemistry, 1301 Catherine Road, 4412 Medical Science Building I, Ann Arbor, MI 48109-0606; Tel. 734-647-0889; Fax. 734-764-3509; E-mail: rtrievel@umich.edu

SET domain enzymes represent a distinct family of protein lysine methyltransferases in eukaryotes. Recent studies have yielded significant insights into the structural basis of substrate recognition and the product specificities of these enzymes. However, the mechanism by which SET domain methyltransferases catalyze the transfer of the methyl group from S-adenosyl-L-methionine to the lysine ε-amine has remained unresolved. To elucidate this mechanism, we have determined the structures of the plant SET domain enzyme, pea Rubisco large subunit methyltransferase, bound to S-adenosyl-L-methionine and its non-reactive analogs L-AzaAdoMet and Sinefungin and characterized the binding of these ligands to a homolog of the enzyme. The structural and biochemical data collectively reveal that S-adenosyl-L-methionine is selectively recognized through carbon-oxygen hydrogen bonds between the cofactor’s methyl group and an array of structurally conserved oxygens that comprise the methyltransfer pore in the active site. Furthermore, the structure of the enzyme co-crystallized with the product ε-N-trimethyllysine reveals a trigonal array of carbon-oxygen hydrogen bonding between the ε-ammonium methyl groups and the oxygens in the pore. Taken together, these results establish a central role for carbon-oxygen hydrogen bonding in aligning the cofactor’s methyl group for transfer to the lysine ε-amine and in coordinating the methyl groups after transfer to facilitate multiple rounds of lysine methylation.

Protein lysine methylation has emerged as a prominent post-translational modification in gene regulatory and intracellular signaling pathways. In the nucleus, site-specific methylation of lysines within histones, transcription factors, and mitotic proteins governs a diverse array of processes within the nucleus including gene expression, DNA damage checkpoint control, cell cycle progression, and mitosis (1). In 2000, a breakthrough in our understanding of this modification occurred with the discovery of the first histone-specific protein lysine methyltransferases (PKMTs) (2). These enzymes possess a conserved catalytic SET domain, a 120 residue motif that was first identified in three Drosophila gene regulatory factors, SU(VAR)3-9, E(Z), and TRX (3). This domain shares no apparent sequence or structural homology with other S-adenosyl-L-methionine (AdoMet)-dependent enzymes, establishing the SET domain family as a novel class of methyltransferases. This seminal discovery heralded the identification of a multitude of SET domain PKMTs which methylate histone and non-histone substrates (1). Since their identification, crystal structures of several SET domain PKMTs in complex with various substrates and products have been reported, yielding insights into the catalytic mechanism and protein substrate specificity of this methyltransferase family. These structures include the histone methyltransferases SET7/9 (4-8), SET8 (also known as PR-SET7) (9,10), and Neurospora DIM-5 (11) as well as pea Rubisco large subunit methyltransferase (pLSMT) (12,13), a non-histone specific plant enzyme which trimethylates Lys-14 in the large subunits of the Rubisco holoenzyme (14). Together, these structures have elucidated the
SET domain’s novel β-sheet architecture which harbors the AdoMet and protein substrate binding sites. A distinctive feature of this fold is the arrangement of the cofactor and substrate binding clefts which are located on opposing faces of the SET domain. This configuration differs significantly from other methyltransferases which generally bind AdoMet and the methyl-acceptor in adjoining clefts (15). Structures of several PKMT ternary complexes have elucidated the path of methyltransfer between the two substrate binding sites (5,8-11,13). The protein substrate intercalates in the substrate binding cleft as parallel β-strand, depositing the lysyl side chain into a narrow hydrophobic channel which traverses the core of the SET domain. This channel terminates in a narrow aperture, which we refer to as the methyltransfer pore, that opens into the AdoMet binding cleft. During methyl transfer, the cofactor’s methyl group is positioned into the pore for the SN2-based transfer reaction with the ε-amine group. Thus, the arrangement of the substrate binding sites promotes highly specific protein substrate recognition while permitting mono-, di-, or trimethylation of the lysine ε-amine group through iterative rounds of catalysis.

Despite the insights obtained from these structural studies, the mechanism by which the methyl group is transferred between the cofactor and lysine binding sites is poorly understood. In prior structural studies of pLSMT, we reported carbon-oxygen (CH···O) hydrogen bonding between the methyl group of N-ε-monomethyllysine (MeLys) and a set of structurally conserved oxygen atoms that comprise the methyltransfer pore (13). This category of hydrogen bonds can occur when an aliphatic or aromatic carbon atom becomes polarized by an adjacent covalently-bonded heteroatom, acidifying the CH hydrogen atoms (16). CH···O hydrogen bonds are estimated to possess half the energy of conventional hydrogen bonds (17), although recent ab initio calculations suggest that CH groups polarized by neighboring cations can engage hydrogen bonding that is energetically comparable to OH···O interactions (18). There is a growing body of literature describing CH···O hydrogen bonding in macromolecular structure, emphasizing its contributions to protein folding (16,19-26), protein:protein interfaces (25), protein:nucleotide recognition (27,28), protein:ligand interactions (29-33), and enzyme catalysis (34-38). Based on the structure of the pLSMT:MeLys complex, we proposed that CH···O hydrogen bonding between the methyltransfer pore and mono- and dimethyllysine facilitates subsequent rounds of methylation by orienting the ε-amine toward the methyltransfer pore while rotating the existing methyl groups out of the transfer path (13).

Our initial observation of CH···O hydrogen bonding in the active site of pLSMT prompted us to further examine methyl group interactions in the SET domain, particularly within the methyltransfer pore. To this end, we have determined the structures of pLSMT in complex with AdoMet and its non-reactive analogs Sinefungin (also known as Adenosyl-L-ornithine) and AzaAdoMet (Fig. 1) and quantified the binding of these ligands to a homolog of the enzyme. In addition, we report the structure of pLSMT bound to ε-N-trimethyllysine (MeLys) and S-adenosyl-L-homocysteine (AdoHey), representing the terminal product complex in lysine trimethylation. Taken together, the biochemical and structural data furnish direct evidence for CH···O hydrogen bonding to methyl groups bound within the SET domain and yield mechanistic insights into the catalytic functions of these interactions in PKMTs.

**EXPERIMENTAL PROCEDURES**

**Chemicals and Reagents—** S-adenosyl-L-methionine, S-adenosyl-L-homocysteine, L-lysine acetate, and Sinefungin were purchased from Sigma, while L-ε-monomethyllysine and L-ε-trimethyllysine were obtained from Bachem. AdoMet was further purified by anion-exchange chromatography as previously reported (39). L-AzaAdoMet was synthesized and purified according to the method of Thompson et al. (40).

**Protein cloning, expression, and purification—** The DNA encoding residues 43-477 of Arabidopsis thaliana Rubisco LSMT (aLSMT) was amplified from an Arabidopsis cDNA library (Research Genetics) and cloned into the GATEWAY vector pDEST14 (Invitrogen) with a C-terminal hexahistidine-tag and a TEV protease cleavage site. Both aLSMT and pLSMT were overexpressed and purified as previously reported (12). After gel filtration purification in 20 mM
TRIS pH 8.0, 150 mM NaCl, and 10 mM β-mercaptoethanol, the proteins were concentrated to approximately 50 mg/ml, flash-frozen in liquid nitrogen, and stored at –80 °C.

Isothermal titration calorimetry—Cofactor binding was measured using isothermal titration calorimetry (ITC) with a VP-ITC calorimeter (Microcal, LLC). Experiments were performed in 20 mM sodium phosphate pH 7.5 and 100 mM NaCl at 20 °C. Due to the aggregation of pLSMT during ITC, titration experiments were performed with aLSMT, a closely related ortholog of pLSMT which shares an overall sequence identity of 64% with the pea enzyme and is 100% identical in its active site. Titrations were performed with various concentrations of ligand (1.0-5.0 mM) and enzyme (0.1-0.4 μM). Data were then analyzed using Origin 7.0 (OriginLab Corp.) with blank injections of the ligand into the buffer subtracted from the titrations prior to data analysis. All of the calculated binding curves have binding stoichiometries (N-values) between 0.8 and 1.0. Curve fitting errors for each titration experiment are reported in Table 1.

Crystallization, structure determination, and modeling—Crystals of pLSMT in complex with AdoMet, Sinefungin:MeLys, AzaAdoMet:Lysine, and Me₃Lys:AdoHcy were obtained in 0.9-1.7 M NaAcetate at pH 6.8 as previously reported (13). Crystals were serially dehydrated by increasing the NaAcetate concentration step-wise to 3.0 M, harvested in mother liquor supplemented with 15% 1,2-propanediol, and flash-frozen in liquid nitrogen. Oscillation images were collected on a Mar165 CCD detector at the COM-CAT 32-ID beamline at the Advanced Photon Source Synchrotron, Argonne, IL. Data were integrated and scaled using d*TREK (41). The structures of the various pLSMT complexes were solved using the pLSMT:AdoHcy:lysine coordinates (1P0Y.pdb) as a starting model (13). Manual model building was carried out using O (42), and structures were refined with CNS using the default stereochemical restraint parameters (43). Simulated annealing Fobs−Fc omit maps were calculated in the absence of ligands in order to accurately model them into the active site in an unbiased manner. Crystallographic data and refinement statistics for each structure are summarized in Table 2. For evaluating CH···O hydrogen bonding parameters, hydrogen atoms were modeled onto AdoMet using PRODRG2 (44), and CH···O hydrogen bond distances and angles within the methyltransfer pore were measured with Swiss-PDB Viewer (45) (Table 3).

RESULTS

Calorimetric analysis of AdoMet, AzaAdoMet, and Sinefungin binding to aLSMT—To investigate the interactions of different chemical moieties within the methyltransfer pore of pLSMT, we determined the thermodynamic parameters for the binding of AdoMet, Sinefungin, and AzaAdoMet to a homolog of the enzyme using isothermal titration calorimetry (ITC). The latter two cofactor analogs were selected for analysis because they possess chemical moieties that differ with respect to the methyl group and sulfonium cation of AdoMet, but are otherwise isosteric (Fig. 1). In AzaAdoMet, a tertiary nitrogen is substituted for the AdoMet sulfonium cation, rendering the methyl group inert for transfer. In Sinefungin, a δ-amine and a methylene moiety are incorporated in place of the cofactor’s methyl group and sulfonium cation, respectively. The structural variations in these cofactor analogs provide a direct means for thermodynamically quantifying the interactions of distinct chemical groups within the enzyme’s methyltransfer pore. Unfortunately, attempts to measure the binding of these ligands to pLSMT were unsuccessful due to its propensity to aggregate during calorimetry experiments (Fig. 2).

Based on our knowledge of the active site, we postulated that Sinefungin would exhibit the greatest affinity for aLSMT due to conventional NH···O hydrogen bonding between the ligand’s primary δ-amine group and the oxygens comprising the methyltransfer pore. Surprisingly, AdoMet binds to aLSMT with a 14-fold higher affinity than Sinefungin and exhibits a correspondingly higher enthalpy of binding to the enzyme (Table 1). Due to the isosteric nature of the ligands, these data suggest that the
methyltransfer pore preferentially interacts with cofactor’s methyl group in comparison to the δ-amine of Sinefungin. Conversely, AzaAdoMet binds over 3000-fold more weakly to aLSMT than AdoMet, indicating that the cofactor’s sulfonium cation is essential for high affinity binding to the enzyme. The impaired binding of AzaAdoMet to aLSMT is most likely due to the replacement of the sulfonium group by a tertiary amine (Fig. 1), which exhibits a pH-dependent ionization state (see below). Finally, we measured the binding of AdoHcy to aLSMT and observed that it exhibits an approximately 70-fold lower affinity for aLSMT compared to AdoMet, despite the greater binding enthalpy of the product versus the substrate. These differences imply that an entropic penalty is involved in the association of AdoHcy with the enzyme, which may derive from losses in the rotational freedom along the Cγ–Sδ and Sδ–C5’ bonds in the product upon binding. AdoMet does not incur this entropic penalty to the same degree owing to the presence of its Cε-methyl group which limits rotation around these bond axes. In summary, these data reveal that AdoMet is selectively recognized by aLSMT through specific interactions with the methyl and sulfonium groups of the cofactor.

Structures of pLSMT bound to AdoMet, AzaAdoMet, and Sinefungin—To elucidate the determinants of cofactor recognition by pLSMT, we have solved the crystal structures of the enzyme bound to AdoMet and its analogs. AdoMet was crystallized in a binary complex with pLSMT, whereas AzaAdoMet and Sinefungin were co-crystallized in ternary complexes with lysine and MeLys, respectively. These free amino acids serve as minimally competent substrates for pLSMT (13) and facilitated crystallization of the Sinefungin and AzaAdoMet complexes. The pLSMT binary and ternary complexes crystallized in the same space group with similar unit cell dimensions and three molecules in the asymmetric unit, as previously reported (12,13).

In each of the pLSMT complexes, the cofactor adopts a horseshoe-shaped conformation within the AdoMet binding pocket, inserting either an amine or methyl group into the methyltransfer pore (Fig. 3). This binding mode is structurally conserved in AdoHcy complexes of the enzyme (12,13) and other cofactor-bound SET domain structures (46), illustrating that the variations in affinity exhibited by Sinefungin, AzaAdoMet, and AdoMet are not due to overt differences in their respective binding conformations. In the pLSMT:AdoMet complex, the cofactor’s methyl group is coordinated in the methyltransfer pore through short distance interactions with the carbonyl oxygens of Ser221 (3.3 Å) and Asp-239 (3.0 Å) and the hydroxyl group of Tyr287 (3.3 Å), which is an invariant residue in SET domain PKMTs (Fig. 3A). Similarly, the AzaAdoMet methyl group engages in short range contacts with the carbonyl oxygen of Ser221 (3.0 Å) and the hydroxyl group of Tyr287 (3.0 Å) (Fig. 3B). In the Sinefungin complex, the ligand’s δ-amine group is bound in the methyltransfer pore through NH···O hydrogen bonds with the hydroxyl group of Tyr287 and the carbonyl oxygen of Asp239, whereas the carbonyl oxygen of Ser221 is too distant (3.6 Å) to engage conventional hydrogen bonding (Fig. 3C). To ascertain whether short-range carbon-oxygen interactions are present in the active sites of other SET domain PKMTs, we examined the 1.7 Å resolution structure of the histone H3 Lys4-specific methyltransferase SET7/9 bound to AdoMet (6) and found that the methyl group interacts with three oxygens in the pore: the carbonyl oxygens of Gly264 (3.7 Å) and His293 (3.0 Å) and the hydroxyl group of the invariant Tyr335 (3.5 Å) (Fig. 3D). The agreement in the oxygen interaction distances in the methyltransfer pores of the pLSMT and SET7/9 structures indicates the distances measured in the pLSMT complexes are accurate despite their lower resolution compared to the SET/9:AdoMet complex.

The carbon-oxygen distances (dCO) observed in the structures of the pLSMT and SET7/9 AdoMet complexes are too short to be classified as conventional van der Waals interactions, which are greater than or equal to 3.7 Å as measured between an oxygen and an aliphatic carbon along the C-H bond axis (16). In the absence of hydrogen atoms, the closest non-overlapping value of dCO is the sum of carbon and oxygen van der Waals radii (1.75 Å and 1.5 Å, respectively) which is 3.25 Å. The observed dCO values in the pLSMT and SET7/9 complexes (Fig. 3A,B,D) concur with previously published distances for CH···O hydrogen bonding (16,30). Furthermore, several of these interactions are
shorter than the minimum carbon-oxygen van der Waals contact distance, suggesting strong CH···O hydrogen bonding between the methyl group and the oxygens that comprise the methyltransfer pore. These short range contacts may derive from the strong polarization of the methyl group by the adjacent sulfonium cation, which would confer a partial positive charge on the carbon atom, enhancing CH···O hydrogen bonding in the methyltransfer pore. Finally, there is a remarkable degree of structural homology in the alignment and interactions of AdoMet methyl group within the respective methyltransfer pores of pLSMT and SET7/9, despite the lack of sequence homology between these enzymes (Fig. 3A,D). These similarities imply that the cofactor binding mode is structurally conserved in order to precisely align the cofactor’s methyl group in the transfer pore during catalysis.

To further evaluate CH···O hydrogen bonding within the methyltransfer pore, hydrogen atoms were modeled onto AdoMet in order to measure the mean hydrogen bond angles and lengths in the three pLSMT:AdoMet complexes in the asymmetric unit. The methyl hydrogens were arranged in a staggered conformation along the Cε-Sδ bond axis to minimize steric hindrance with the Cγ and C5′ methylene groups in the cofactor (Fig. 4A). Intriguingly, this conformation aligns the methyl hydrogens in a nearly perfect trigonal geometry with the three hydrogen bond acceptors in the methyltransfer pore. Measurements of the mean hydrogen-oxygen distances (dH) are less than the sum of the van der Waals radii for these atoms (2.7 Å), which is indicative of CH···O hydrogen bonding (Table 3). Moreover, the dH values and the Cε-H···O hydrogen bond angles (θ) (Fig. 4B) agree with statistical analyses of CH···O hydrogen bonding in protein:ligand interactions (32). The standard deviations in dH and θ vary less than 10% and 3%, respectively, among the three pLSMT:AdoMet complexes in the asymmetric unit, indicating a preferred geometric alignment of the methyl group within the methyltransfer pore, as previously discussed. The correlations in these values are not due to non-crystallographic symmetry averaging as no symmetry restraints were applied to the SET domain during refinement.

The propensity of the oxygens in the methyltransfer pore to engage in CH···O hydrogen bonding provides a molecular basis for the AdoMet binding selectivity of the SET domain. AdoMet binds with the highest affinity to aLSMT (Table 1) owing to the three CH···O hydrogen bonds that coordinate the methyl group in the methyltransfer pore (Fig. 3A,4). The weaker binding observed for Sinefungin may result from fewer conventional NH···O hydrogen bonds between its δ-amine and the pore oxygens (Fig. 3C). Alternatively, the difference in the binding enthalpies of the ligands may be due to variations in their respective desolvation energies or to changes in the protonation state of Sinefungin’s δ-amine during binding. Nonetheless, the structural and biochemical data furnish evidence for strong CH···O hydrogen bonding with the AdoMet methyl group in the methyltransfer pore. The short dCδ and dH values of these hydrogen bonds (Table 3) reflect the relative strength of these interactions, which is induced by the polarization of the neighboring sulfonium cation in the cofactor. These distances concur with short-range CH···O hydrogen bonding observed with aromatic CH groups adjacent to positively charged nitrogens in heterocycles (30), illustrating that neighboring cations can significantly augment CH···O interactions.

In contrast to AdoMet, the relatively weak binding of AzaAdoMet indicates that its methyl group participates in weaker CH···O interactions within the active site. In this analog, the polarization of the methyl group is dependent on the protonation state of its tertiary γ-amine (pKₐ ~7.1), which mimics the AdoMet sulfonium cation (40,47). Under basic conditions, the methyl group is relatively unpolarized due to the amine’s deprotonation. This charge neutralization is reflected in the substantially impaired binding of AzaAdoMet to aLSMT at pH 7.5 (Table 1). However, CH···O hydrogen bonding is observed in the structure of pLSMT bound to AzaAdoMet (Fig. 3B) which was crystallized under slightly acidic conditions (pH 6.8) in which a sufficient concentration of the γ-ammonium form of the ligand would be present in solution for crystallization with the enzyme. Unfortunately, we were not able to determine the pH-dependent binding of AzaAdoMet to aLSMT due to the aggregation of the enzyme at pH values less than 7.0. Nonetheless, our data are consistent with surface plasmon resonance studies of the E coli
methionine repressor MetJ and AzaAdoMet in which ligand binding was only observed under acidic conditions, eliciting co-repression comparable to AdoMet (47). In summary, these results demonstrate that the strength of the CH···O interactions are strongly dependent on the degree of polarization of the heteroatom to which the methyl group is covalently bonded.

Structure of pLSMT:Me₃Lys:AdoHcy ternary complex—In addition to characterizing cofactor interactions with pLSMT, we sought to understand how the methyl groups are accommodated within the active site after transfer to the lysine ε-amine, expanding upon our initial structural analysis of lysine and MeLys recognition (13). To examine these interactions, we determined the structure of pLSMT bound to AdoHcy and the free amino acid trimethyllysine (Me₃Lys). This structure represents the terminal product complex in lysine ε-amine trimethylation and is the first Me₃Lys-bound structure reported for a SET domain PKMT. To insure that the trimethylated ε-ammonium group was properly modeled into the structure, simulated annealing $F_o - F_c$ omit maps were calculated for the active site in the absence of the products (Fig. 5A). These maps revealed tri-lobe shaped electron density for the ε-ammonium cation, permitting the unambiguous assignment of the methyl groups in the experimental map (Fig. 5B). The overall binding mode of Me₃Lys in the lysine binding channel is structurally analogous to those reported for lysine and MeLys in complex with pLSMT (13).

An examination of the carbon-oxygen interactions between Me₃Lys and the oxygens in the methyltransfer pore yields an explanation for the orientation of the trimethylated ε-ammonium group within the active site. The quaternary ε-ammonium cation is poised in a striking conformation in which its three methyl groups are directly aligned with three oxygens in the methyltransfer pore (Fig. 5B). This arrangement of the methyl group and ε-amine is critical for the prerequisite linear geometry of the $S_N2$ methyltransfer reaction, as recently shown in quantum mechanical/molecular mechanical free energy modeling of catalysis in SET7/9 (48).

DISCUSSION

The results presented herein furnish compelling evidence for CH···O hydrogen bonding in the SET domain active site, corroborating our initial observation of these interactions in the structure of pLSMT bound to MeLys (13). The presence of CH···O hydrogen bonding has several implications in regards to the catalytic mechanism of SET domain PKMTs. Based on our biochemical and structural data, we propose that the CH···O interactions in the methyltransfer pore serve two distinct functions in catalysis: 1) to align the AdoMet methyl group for the $S_N2$ transfer to the lysine ε-amine and 2) to orient the mono- and dimethylated ε-amine into favorable orientations for subsequent methylation.

The analysis of cofactor binding by pLSMT suggests that CH···O hydrogen bonding has dual roles in facilitating methyltransfer with AdoMet. First, CH···O hydrogen bonding in the methyltransfer pore orients the methyl group in the correct position for transfer to the lysine ε-amine (Fig. 6). The virtually identical alignment of the AdoMet methyl group within the pores of pLSMT and SET7/9 (Fig. 3A,D) implies that its orientation is structurally conserved in the active sites of all SET domain PKMTs. The arrangement of the methyl group and ε-amine is critical for the prerequisite linear geometry of the $S_N2$ methyltransfer reaction, as recently shown in quantum mechanical/molecular mechanical free energy modeling of catalysis in SET7/9 (48). Secondly, CH···O hydrogen bonding within the methyltransfer pore (Fig. 4,6) provides the necessary environment for an $S_N2$ linear methyltransfer reaction. The fact that the final product of pLSMT catalyzed lysine methylation is the ε-N-trimethylammonium cation establishes that the reaction mechanism does not use general base catalysis and that therefore the transition state has cationic charge dispersal from sulfur through carbon to nitrogen, S···C···N(+). This will benefit from the proximity of the three oxygens (Ser221,
Asp239, and Tyr287) via hydrogen bonding and an electronegative field effect while retaining the essentially $S_N2$ character of the methyltransfer. Finally, mutagenesis of the invariant tyrosine in SET domain PKMTs corroborates the proposed catalytic roles for the oxygens in the methyltransfer pore. Mutations of this residue in SET7/9 (12,49) and in DIM-5 (50) severely compromise enzymatic activity, demonstrating that the pore oxygens are critical for catalysis. In both pLSMT and SET7/9, the phenol side chain of the invariant tyrosine engages in CH⋯O hydrogen bonding with the AdoMet methyl group (Fig. 3A,D). Loss of the hydroxyl group through mutation abrogates this key interaction, impairing the alignment of the methyl group during catalysis.

In addition to aligning the AdoMet methyl group for transfer, the oxygens in the methyltransfer pore also engage in CH⋯O hydrogen bonding with the methylated lysine $\varepsilon$-amino group. Our findings reveal that the methyl groups are bound predominantly through CH⋯O hydrogen bonds, whereas the aliphatic portion of the side chain engages in hydrophobic interactions with the aromatic residues lining the lysine binding channel (Fig. 5). This methyllysine binding mode differs significantly from proteins bearing chromodomains, in which the di- and trimethylated lysine $\varepsilon$-amine is bound within an aromatic cage motif through cation-π interactions (51,52). The distinct mechanisms by which chromo and SET domains recognize methyllysine likely reflects their respective biological functions. Whereas the aromatic cage serves as a general methyllysine binding motif, the stereoregular constraints inherent in CH⋯O hydrogen bonding are suited to precisely aligning the mono- and dimethylated $\varepsilon$-amine in a linear geometry with the AdoMet methyl group for subsequent rounds of $S_N2$ methyltransfer. Additional studies of SET domain PKMTs in complex with mono-, di-, and trimethyllysine are required to address how CH⋯O hydrogen bonding specifically orients the methylated $\varepsilon$-amine during each stage of catalysis.

REFERENCES

1. Dillon, S. C., Zhang, X., Trievel, R. C., and Cheng, X. (2005) *Genome Biol* **6**, 227
2. Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., and Jenuwein, T. (2000) *Nature* **406**, 593-599.
3. Tschiersch, B., Hofmann, A., Krauss, V., Dorn, R., Korge, G., and Reuter, G. (1994) *Embo J* **13**, 3822-3831
4. Jacobs, S. A., Harp, J. M., Devarakonda, S., Kim, Y., Rastinejad, F., and Khorasanizadeh, S. (2002) *Nat Struct Biol* **9**, 833-838.
5. Xiao, B., Jing, C., Wilson, J. R., Walker, P. A., Vasisht, N., Kelly, G., Howell, S., Taylor, I. A., Blackburn, G. M., and Gamblin, S. J. (2003) *Nature* **421**, 652-656.
6. Kwon, T., Chang, J. H., Kwak, E., Lee, C. W., Joachimiak, A., Kim, Y. C., Lee, J., and Cho, Y. (2003) *Embo J* **22**, 292-303.
7. Chuikov, S., Kurash, J. K., Wilson, J. R., Xiao, B., Justin, N., Ivanov, G. S., McKinney, K., Tempst, P., Prives, C., Gamblin, S. J., Barlev, N. A., and Reinberg, D. (2004) *Nature* **432**, 353-360
8. Couture, J. F., Collazo, E., Hauk, G., and Trievel, R. C. (2006) *Nat Struct Mol Biol* **13**, 140-146
9. Couture, J. F., Collazo, E., Brunzelle, J. S., and Trievel, R. C. (2005) *Genes Dev* **19**, 1455-1465
10. Xiao, B., Jing, C., Kelly, G., Walker, P. A., Muskett, F. W., Frenkiel, T. A., Martin, S. R., Sarma, K., Reinberg, D., Gamblin, S. J., and Wilson, J. R. (2005) *Genes Dev* **19**, 1444-1454
11. Zhang, X., Yang, Z., Khan, S. I., Horton, J. R., Tamaru, H., Selker, E. U., and Cheng, X. (2003) *Mol Cell* **12**, 177-185.
12. Trievel, R. C., Beach, B. M., Dirk, L. M., Houtz, R. L., and Hurley, J. H. (2002) *Cell* **111**, 91-103.
13. Trievel, R. C., Flynn, E. M., Houtz, R. L., and Hurley, J. H. (2003) *Nat Struct Biol* **10**, 545-552.
14. Houtz, R. L., Royer, M., and E., S. M. (1991) *Plant Physiol.* **97**, 913-920.
15. Yeates, T. O. (2002) *Cell* **111**, 5-7.
16. Derewenda, Z. S., Lee, L., and Derewenda, U. (1995) *J Mol Biol* **252**, 248-262.
17. Vargas, R., Garza, J., Dixon, D. A., and Hay, B. P. (2000) *J Am Chem Soc* **122**, 4750-4755.
18. Scheiner, S., Kar, T., and Gu, Y. (2001) *J Biol Chem* **276**, 9832-9837.
19. Bella, J., and Berman, H. M. (1996) *J Mol Biol* **264**, 734-742.
20. Fabiola, G. F., Krishnaswamy, S., Nagarajan, V., and Pattabhi, V. (1997) *Acta Crystallogr. D Biol. Crystallogr.* **53**, 316-320.
21. Chakrabarti, P., and Chakrabarti, S. (1998) *J Mol Biol* **284**, 867-873.
22. Senes, A., Ubarretxena-Belandia, I., and Engelman, D. M. (2001) *Proc Natl Acad Sci U S A* **98**, 9056-9061.
23. Chamberlain, A. K., and Bowie, J. U. (2002) *J Mol Biol* **322**, 497-503.
24. Babu, M. M., Singh, S. K., and Balaram, P. (2002) *J Mol Biol* **322**, 871-880.
25. Jiang, L., and Lai, L. (2002) *J Biol Chem* **277**, 37732-37740.
26. Manikandan, K., and Ramakumar, S. (2004) *Proteins* **56**, 768-781.
27. Mandel-Gutfreund, Y., Margalit, H., Jernigan, R. L., and Zhurkin, V. B. (1998) *J Mol Biol* **277**, 1129-1140.
28. Treger, M., and Westhof, E. (2001) *J Mol Recognit* **14**, 199-214.
29. Engh, R. A., Brandstetter, H., Sucher, G., Eichinger, A., Baumann, U., Bode, W., Huber, R., Poll, T., Rudolph, R., and von der Saal, W. (1996) *Structure* **4**, 1353-1362.
30. Pierce, A. C., Sandretto, K. L., and Bemis, G. W. (2002) *Proteins* **49**, 567-576.
31. Klaholz, B., and Moras, D. (2002) *Structure (Camb)* **10**, 1197-1204.
32. Sarkhel, S., and Desiraju, G. R. (2004) *Proteins* **54**, 247-259.
33. Pierce, A. C., ter Haar, E., Binch, H. M., Kay, D. P., Patel, S. R., and Li, P. (2005) *J Med Chem* **48**, 1278-1281.
34. Rice, J. C., and Allis, C. D. (2001) *Curr Opin Cell Biol* **13**, 263-273.
35. Derewenda, Z. S., Derewenda, U., and Kobos, P. M. (1994) *J Mol Biol* **241**, 83-93.
36. Bond, C. S., Zhang, Y., Berriman, M., Cunningham, M. L., Fairlamb, A. H., and Hunter, W. N. (1999) *Structure Fold Des* **7**, 81-89.
37. Meyer, P., Schneider, B., Sarfati, S., Deville-Bonne, D., Guerreiro, C., Boretto, J., Janin, J., Veron, M., and Canard, B. (2000) *Embo J* **19**, 3520-3529.
38. Bach, R. D., Thorpe, C., and Dmitrenko, O. (2002) *J Phys Chem* **106**, 4325-4335.
39. Chirpich, T. P., Zappia, V., Costilow, R. N., and Barker, H. A. (1970) *J Biol Chem* **245**, 1778-1789.
40. Thompson, M. J., Mekhalfia, A., Hornby, D. P., and Blackburn, G. M. (1999) *J Org Chem* **64**, 7467-7473.
41. Messerschmidt, A., and Pflugrath, J. W. (1987) *J Appl Cryst* **20**, 306-315.
42. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A* **47**, 110-119.
43. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M.,
Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr D Biol Crystallogr* **54**, 905-921.

44. Schuttelkopf, A. W., and van Aalten, D. M. (2004) *Acta Crystallogr D Biol Crystallogr* **60**, 1355-1363.

45. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* **18**, 2714-2723.

46. Xiao, B., Wilson, J. R., and Gamblin, S. J. (2003) *Curr Opin Struct Biol* **13**, 699-705.

47. Parsons, I. D., Persson, B., Mekhalfia, A., Blackburn, G. M., and Stockley, P. G. (1995) *Nucleic Acids Res* **23**, 211-216.

48. Hu, P., and Zhang, Y. (2006) *J Am Chem Soc* **128**, 1272-1278.

49. Wilson, J. R., Jing, C., Walker, P. A., Martin, S. R., Howell, S. A., Blackburn, G. M., Gamblin, S. J., and Xiao, B. (2002) *Cell* **111**, 105-115.

50. Zhang, X., Tamaru, H., Khan, S. I., Horton, J. R., Keefe, L. J., Selker, E. U., and Cheng, X. (2002) *Cell* **111**, 117-127.

51. Nielsen, P. R., Nietlispach, D., Mott, H. R., Callaghan, J., Bannister, A., Kouzarides, T., Murzin, A. G., Murzina, N. V., and Laue, E. D. (2002) *Nature* **416**, 103-107.

52. Jacobs, S. A., and Khorasanizadeh, S. (2002) *Science* **295**, 2080-2083.

FOOTNOTES

We thank Joseph Brunzelle for his assistance in X-ray data collection, and Robert Houtz for his gift of purified AdoMet. We also wish to acknowledge Martha Ludwig for insightful discussions and Daniel Peisach and Robert Houtz for reading the manuscript and providing useful comments. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Sciences, under Contract No W-31-109-Eng-38. The authors gratefully acknowledge the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor for the support of this research program (Grant 085P1000817). J.-F.C. is the recipient of a Canadian Institutes of Health Research (CIHR) Postdoctoral Fellowship. This work was supported by a grant from the National Institutes of Health (GM073839 to R.C.T.).

The coordinates and structure factor files for the pLSMT complexes have been deposited in the Protein Databank.

The abbreviations used are: AdoHcy, S-adenosyl-L-homocysteine; AdoMet, S-adenosyl-L-methionine; CH···O, carbon (hydrogen)-oxygen; \(d_{CO}\), carbon-oxygen distance; \(d_{OH}\), hydrogen-oxygen distance; ITC, isothermal titration calorimetry; MeLys, \(\epsilon\)-N-monomethyllysine; Me3Lys, \(\epsilon\)-N-trimethyllysine; protein lysine methyltransferase, PKMT; Rubisco LSMT, Ribulose-1,5 bisphosphate carboxylase/oxygenase large subunit N-methyltransferase.

FIGURE LEGENDS

**Fig. 1.** Structures of AdoMet, AdoHcy, and the cofactor analogs Sinefungin (adenosyl-L-ornithine) and AzaAdoMet.

**Fig. 2.** ITC analysis of AdoMet binding to aLSMT. The top panel represents the titration of AdoMet into an aLSMT solution with the heat evolved (\(\mu\)cal/sec) plotted versus time (min). The bottom panel illustrates the binding isotherm with the fitted binding curve.

**Fig. 3.** Crystal structures of pLSMT and SET7/9 bound to AdoMet and cofactor analogs. A, Structure of the pLSMT methyltransfer pore in complex with AdoMet (B molecule in the asymmetric unit shown).
The carbon atoms of AdoMet and the enzyme are rendered in green and gray, respectively. Conventional hydrogen bonds are illustrated in cyan, while CH···O hydrogen bonds and their respective $d_{CO}$ values are highlighted in orange. Residues which comprise the methyltransfer pore are labeled in red for clarity. B, Structure of AzaAdoMet bound within the methyltransfer pore of pLSMT, rendered as in A. C, Structure of pLSMT in complex with Sinefungin, illustrated as in A. D, Crystal structure of SET7/9 bound to AdoMet (1N6A.pdb) determined at 1.7 Å resolution, depicted as in A. The cutoff distances for CH···O and conventional NH···O hydrogen bonds in the methyltransfer pores are 3.7 Å and 3.3 Å, respectively (16).

Fig. 4. Stereochemical parameters of CH···O hydrogen bonding within the methyltransfer pore of pLSMT. A, C$\varepsilon$H···O hydrogen-oxygen distances ($d_H$) with the carbonyl oxygens of Ser221 and Asp239 and the hydroxyl group of Tyr287 in the pore (B molecule in the asymmetric unit illustrated as in Fig. 3A). The carbon atoms of AdoMet and the protein are rendered in green and gold, respectively, while CH···O hydrogen bonds and associated $d_H$ values are shown in orange. The view along the C$\varepsilon$-S$\delta$ bond axis of AdoMet illustrates the staggered orientation of the modeled hydrogens in relation to the C$\gamma$ and C5’ methylene groups of the cofactor. B, C$\varepsilon$H···O hydrogen bond angles ($\theta$) in the methyltransfer pore. The structure of the pLSMT:AdoMet complex is rotated relative to Fig. 4A to illustrate the favorable $\theta$ values with the acceptor oxygens. The mean values of $d_H$ and $\theta$ for the three pLSMT molecules in the asymmetric unit are reported in Table 3.

Fig. 5. Crystal structure of ternary complex of pLSMT bound to Me$_3$Lys and AdoHcy. A, Simulated annealing $F_\sigma$-$F_c$ map of the pLSMT active site (gray carbons) calculated in the absence of the products AdoHcy (green carbons) and Me$_3$Lys (gold carbons). The map is contoured around Me$_3$Lys at 2.0 $\sigma$ in the active site of the C molecule in the asymmetric unit. B, The simulated annealing map viewed along the C$\varepsilon$-N$\zeta$ bond axis of Me$_3$Lys, illustrating the trilobal electron density of the trimethylated $\varepsilon$-ammonium group. CH···O hydrogen bonding between the methyl groups and acceptor oxygens in the methyltransfer pore and their associated $d_{CO}$ values are illustrated in orange.

Fig. 6. Proposed catalytic role of the methyltransfer pore in the SET domain active site. The carbonyl oxygens of Ser221 and Asp239 and the hydroxyl group of the invariant Tyr287 engage in CH···O hydrogen bonding with the AdoMet methyl group, aligning it for the $S_N$2 transfer to the lysine $\varepsilon$-amine.

| Ligand        | $K_D$ (µM) | $\Delta H$ (kcal/mol) |
|---------------|------------|-----------------------|
| AdoMet        | 0.29 ± 0.02* | -11.0 ± 0.1*          |
| Sinefungin    | 4.2 ± 0.1   | -8.3 ± 0.1            |
| AzaAdoMet     | > 1000 ± 100 | < -3.9 ± 0.8          |
| AdoHcy        | 21.2 ± 1.6  | -14.7 ± 0.3           |

* Deviations represent curve-fitting errors calculated from the binding isotherms.
Table 2: Crystallographic data and refinement statistics for pLSMT complexes

| Crystal Parameters | AdoMet | Sinefungin:MeLys | AzaAdoMet:Lys | AdoHcy:Me3Lys |
|-------------------|--------|------------------|---------------|---------------|
| Space group       |        | I222             |               |               |
| Unit Cell         | a (Å)  | 131.45           | 132.37        | 132.63        | 131.86        |
|                   | b      | 155.86           | 156.50        | 159.49        | 157.55        |
|                   | c      | 263.61           | 267.59        | 268.51        | 267.60        |
|                   | α, β, γ (°) | α = β = γ = 90 |               |               |
| Asymmetric Unit   |        | 3 molecules      |               |               |

Data Collection statistics

| Resolution Range (Å) | 13.0 – 2.45 | 28.0 – 2.35 | 16.8 – 2.60 | 15.0 – 2.44 |
|----------------------|-------------|-------------|-------------|-------------|
| Total Reflections    | 517485      | 356765      | 314052      | 521939      |
| Unique Reflections   | 97807       | 108311      | 85610       | 102787      |
| Rsym (%)             | 6.5         | 4.8         | 10.3        | 5.8         |
| I / σ(I)             | 12.7        | 13.7        | 7.2         | 13          |
| Completeness (%)     | 99.3        | 97.6        | 98.6        | 99.1        |

Refinement Statistics

| Resolution Range (Å) | 12.9 – 2.45 | 13.0 – 2.35 | 16.7 – 2.60 | 15.0 – 2.44 |
|----------------------|-------------|-------------|-------------|-------------|
| Reflections (Fo > 2σ) | 84670       | 96094       | 85270       | 99453       |
| Final Model           |             |             |             |             |
| Protein Atoms         | 10483       | 10483       | 10483       | 10483       |
| Ligands               | 81          | 114         | 111         | 117         |
| Water                 | 559         | 450         | 387         | 390         |
| R-Factors             |             |             |             |             |
| Rworking              | 24.8        | 25.6        | 26.2        | 25.2        |
| Rfree                 | 28.8        | 29.5        | 29.8        | 28.7        |
| Luzzati Coordinate Error (Å) | 0.42 | 0.49 | 0.47 | 0.42 |
| Cross-validated Luzzati Coordinate Error (Å) | 0.49 | 0.50 | 0.51 | 0.48 |
| R.M.S. Bond Length (Å) | 0.007       | 0.007       | 0.008       | 0.007       |
| Bond Angles (°)       | 1.2         | 1.2         | 1.3         | 1.2         |
| Average B –factors (Å²) | 68.1        | 70.8        | 71.4        | 66.6        |
| Protein               |             |             |             |             |
| Ligands               | 48.3        | 58.7        | 92.8        | 80.2        |
| Water                 | 61.9        | 64.0        | 56.8        | 59.1        |

R-factor: \( R_{\text{working}} = \frac{\sum | | F_o | - | F_c | |}{\Sigma | F_o |}; R_{\text{free}} = \frac{\Sigma_T | | F_o | - | F_c | |}{\Sigma_T | F_o |}, \) where T is a test data set of 4.5 % of the total reflections randomly chosen and set aside before refinement.

Table 3: CH···O hydrogen bond parameters for the AdoMet methyl group within the three pLSMT molecules in the asymmetric unit.

| CH···O H-Bond Acceptor | Mean \( d_{CO} \) (Å) | Mean \( d_H \) (Å) | Mean \( θ \) (°) |
|------------------------|------------------------|-------------------|------------------|
| Ser221 Carbonyl Oxygen | 3.2 ± 0.2              | 2.3 ± 0.2         | 147 ± 3          |
| Asp239 Carbonyl Oxygen | 3.2 ± 0.2              | 2.4 ± 0.2         | 139 ± 4          |
| Tyr287 Hydroxyl Group  | 3.3 ± 0.1              | 2.3 ± 0.1         | 155 ± 4          |
FIGURE 1

S-adenosyl-L-methionine (AdoMet)

S-adenosyl-L-homocysteine (AdoHcy)

L-AzaAdoMet

Adenosyl-L-ornithine (Sinefungin)
FIGURE 2

Time (hrs)

µcal/sec

Molar Ratio

kcal/mole of injectant
FIGURE 4

A

B

AdoMet
Y287
S221
D239

AdoMet
Y287
D239
S221
Catalytic roles for carbon-oxygen hydrogen bonding in set domain lysine methyltransferases

Jean-Francois Couture, Glenn Hauk, Mark J. Thompson, G. Michael Blackburn and Raymond C. Trievel

J. Biol. Chem. published online May 8, 2006

Access the most updated version of this article at doi: 10.1074/jbc.M602257200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts