Cell, Vol. 112, 711–723, March 7, 2003, Copyright ©2003 by Cell Press

Structure of the Catalytic Domain of Human DOT1L, a Non-SET Domain Nucleosomal Histone Methyltransferase

Jinrong Min,[1,3] Qin Feng,[2,3] Zhizhong Li,[1] Yi Zhang,[2] and Rui-Ming Xu[1,*]

[1]W.M. Keck Structural Biology Laboratory
Cold Spring Harbor Laboratory
Cold Spring Harbor, New York 11724

[2]Department of Biochemistry and Biophysics
Lineberger Comprehensive Cancer Center
University of North Carolina at Chapel Hill
Chapel Hill, North Carolina 27599

Summary

Dot1 is an evolutionarily conserved histone methyltransferase that methylates lysine-79 of histone H3 in the core domain. Unlike other histone methyltransferases, Dot1 does not contain a SET domain, and it specifically methylates nucleosomal histone H3. We have solved a 2.5 Å resolution structure of the catalytic domain of human Dot1, hDOT1L, in complex with S-adenosyl-L-methionine (SAM). The structure reveals a unique organization of a mainly α-helical N-terminal domain and a central open α/β structure, an active site consisting of a SAM binding pocket, and a potential lysine binding channel. We also show that a flexible, positively charged region at the C terminus of the catalytic domain is critical for nucleosome binding and enzymatic activity. These structural and biochemical analyses, combined with molecular modeling, provide mechanistic insights into the catalytic mechanism and nucleosomal specificity of Dot1 proteins.

Introduction

Higher-order chromatin structures are of profound importance in gene regulation and epigenetic inheritance (Wu and Grunstein, 2000). Posttranslational modifications of core histones critically influence the establishment and maintenance of higher-order chromatin structures. The unstructured tails of certain core histones are extensively modified by acetylation, methylation, phosphorylation, ribosylation, and ubiquitination. A "histone code" hypothesis, linking histone modifications to chromatin structures, has been the focus of intensive recent studies (Strahl and Allis, 2000; Turner, 2000). Histone methylation has emerged as a major form of histone modification. (Strahl and Allis, 2000; Zhang and Reinberg, 2001). In particular, a large family of SET domain-containing histone methyltransferases (HMTases) has recently been identified (Lachner and Jenuwein, 2002). To date, SET domain proteins have been shown to methylate various N-terminal lysine residues of histones H3 and H4. Histone lysine methylation has been associated with diverse biological processes, ranging from transcriptional regulation to the faithful transmission of chromosomes during the cell division (Grewal and Elgin, 2002).

A recent addition to the histone modification repertoire is histone H3 lysine-79 (Lys79) methylation (Feng et al., 2002; Lacoste et al., 2002; Ng et al., 2002a; van Leeuwen et al., 2002). In contrast to all other known methylation sites, Lys79 occupies a position in the ordered core domain of histone H3 and resides in a short turn connecting the first and second helices of the conserved histone fold. In the nucleosome core particle structure (Luger et al., 1997; White et al., 2001), both histone H3 Lys79 residues in the histone octamer are exposed and do not interact with DNA or other histones. Dot1 is an evolutionarily conserved protein that was originally identified in *S. cerevisiae* as a *d*isruptor *o*f *t*elomeric silencing (Singer et al., 1998). It also functions at the pachytene checkpoint during the meiotic cell cycle (San-Segundo and Roeder, 2000). A careful sequence analysis revealed that it possesses certain characteristic SAM binding motifs, similar to the ones in protein arginine methyltransferases (Dlakic, 2001). Surprisingly, we, as well as others, found that yeast Dot1 and its human counterpart hDOT1L exhibit intrinsic histone methyltransferase activity toward Lys79 of histone H3 in the nucleosome core (Feng et al., 2002; Lacoste et al., 2002; Ng et al., 2002a; van Leeuwen et al., 2002). Dot1 proteins do not contain a SET domain, a conserved sequence motif found in all previously characterized histone lysine methyltransferases. This unexpected finding identifies the Dot1 family of proteins as the first non-SET domain-containing histone lysine methyltransferases. Furthermore, Dot1 and hDOT1L methylate nucleosomal substrates, but not free histone H3. Several SET domain proteins also require nucleosome substrates for efficient methylation; they include yeast Set2, which methylates Lys36 of histone H3 (Strahl et al., 2002), and human SET8/PR-Set7, which methylates Lys20 of histone H4 (Fang et al., 2002; Nishioka et al., 2002). Lys79 of H3 is globally methylated, and Dot1 is solely responsible for this methylation in *S. cerevisiae* (Ng et al., 2002a; van Leeuwen et al., 2002). In addition, Lys79 is mono-, di- and trimethylated by Dot1 in vitro and in vivo.

Methylation of H3 Lys79 is important for gene silencing and the proper localization of silent information regulators (SIR) in *S. cerevisiae* (Ng et al., 2002a; van Leeuwen et al., 2002). Interestingly, both deletion and overexpression of Dot1 results in loss of silencing. Mutation of H3 Lys79 and surrounding residues on the core nucleosome surface also abolishes silencing in *S. cerevisiae* (Park et al., 2002). Since ~90% of the chromatin is methylated by Dot1, it has been proposed that the principal function of H3 Lys79 methylation is to mark open, transcriptionally active, chromatin domains, which effectively limits the binding of SIR proteins to discrete loci in the genome wherein genes are silenced. In light of a recent finding that di- and trimethylated H3 Lys4 in *S. cerevisiae* have distinct roles in gene activation (Santos-Rosa et al., 2002), it is also possible that mono-, di-, and trimethylated H3 Lys79 have different functions

*Correspondence: xur@cshl.org

[3]These authors contributed equally to this work.

in marking distinct chromatin domains. Remarkably, ubiquitination of Lys123 of histone H2B is required for methylation of Lys79 of H3 in *S. cerevisiae* (Briggs et al., 2002; Ng et al., 2002b), a situation that bears striking similarity to the methylation of Lys4 of H3 by Set1 (Dover et al., 2002; Sun and Allis, 2002). It has been suggested that an ubiquitinated histone H2B might serve as a "spacer" between adjacent nucleosomes to allow free access by histone methyltransferases. The precise function of H3 Lys79 in higher eukaryotes is not yet clear. However, a functional role for H3 Lys79 methylation in differentiating chromatin domains is likely, considering its conservation across diverse species.

In view of the great impact of histone methylation in a wide range of biological processes, it is of fundamental importance to understand the molecular mechanisms of histone methylation. The crystal structures of several SET domain histone lysine methyltransferases have been determined recently (Jacobs et al., 2002; Min et al., 2002; Trievel et al., 2002; Wilson et al., 2002; Zhang et al., 2002), and they should facilitate further structure-and-function studies of SET domain HMTases. In an effort to determine the structural basis of histone core domain lysine methylation and nucleosome specificity of the novel Dot1 family of histone methyltransferases, we have determined a 2.5 Å resolution structure of the catalytic domain of the human homolog of Dot1, hDOT1L, which reveals important insights into the structural basis of nucleosome binding and the catalytic mechanism of this important group of enzymes.

Results

Overview

The full-length hDOT1L consists of 1537 amino acids (Feng et al., 2002), but only the N-terminal ~360 amino acids share significant sequence similarity with yeast Dot1 (Figure 1A). We determined that amino acids 1–416 of hDOT1L, hDOT1L(1–416), contain the active HMTase catalytic domain (Figure 1B). We have solved the crystal structure of hDOT1L(1–416) complexed with SAM to 2.5 Å resolution, using multiwavelength anomalous diffraction (MAD) of seleno-methionyl (SeMet) protein crystals. Detailed statistics of structure determination and refinement are shown in Table 1. The ordered structure contains residues 4 to 332. Although not observed in the structure, the C-terminal tail is present in the crystallized protein, as indicated by mass spectrometry and SDS-PAGE analyses of the protein recovered from hDOT1L(1–416) crystals (data not shown). The absence of a long C-terminal tail in the ordered structure is likely due to its structural flexibility in the absence of substrates (discussed below). The refined model has good stereochemistry: 89% of the main chain ϕ/ψ angles are in the most favored region and none in the disallowed region of the Ramachandran plot, calculated using PROCHECK (Laskowski et al., 1993).

The catalytic core has an elongated structure (Figure 2A) with linear dimensions of approximately 80 Å × 40 Å × 40 Å. The overall fold of the structure is shown in a topological diagram (Figure 2B). The structure consists of an N-terminal domain with five α helices (αA–αE) and two pairs of short β strand hairpins (β1, β2 and β3, β4). The C-terminal domain is an open α/β structure with a seven-strand central β sheet (β5–β11) and five α helices (αF–αJ). Helices αF–αH are packed onto one side (front) of the β sheet, and the other two helices (αI, αJ) are on the opposite side (Figure 2A). In the β sheet, β5–β9 and β10 are parallel to each other, but β11, which is between β9 and β10, is antiparallel to the others (Figure 2). β11 is followed by helix αK, which points away from the center of the structure. Residues C-terminal to αK (aa 333–416) cannot be modeled reliably because of weak electron density. The N-terminal domain, which is located on the same side of the β sheet with helices αF–αH, and the open α/β structure domain are in close contact. Major interactions between the two domains are mediated through the long αH helix and residues immediately C-terminal to αH. They interact extensively with residues located in the β1-β2 hairpin, helices αD and αE, and regions immediately adjacent to the two helices in the N-terminal domain. A SAM molecule is bound in a pocket formed between loop L-EF, which covalently links the two domains in the structure, and the C-terminal ends of β5 and β8 (Figure 2) in the open α/β structure.

SAM Binding Pocket

The SAM molecule is well ordered in the crystal structure (Figure 3A). Interestingly, purified recombinant protein produced in *E. coli* contains endogenously bound SAM. This was revealed by the electron density for SAM using data collected from crystals grown in the absence of SAM. In addition, the anomalous difference electron density map of SeMet hDOT1L shows that the bound SAM contains selenium (Figure 3A). The *E. coli* strain producing SeMet proteins presumably used the supplied seleno-methionine to make Se-adenosyl-L-methionine. The active site consists of the SAM binding pocket and a narrow channel leading to the methyl group of SAM (Figures 3B and 3C). The SAM binding pocket is wider at the entrance but becomes narrower toward the center of the protein. The methionyl moiety of SAM inserts deep into the pocket with the adenine group toward the outside. The inside of the SAM binding pocket is negatively charged, while the adenine binding region is more hydrophobic. Detailed protein-SAM interactions are shown schematically in Figure 3D. Five segments of the protein form the SAM binding pocket: D1 region (aa 133–139 located in loop L-EF) forms the right front of the pocket; region I (aa 161–169 located in β5-loop-αG) and region I′ (aa 186–189 and 191 in β6-loop-αH, also known as post-I region) together form the backside of the pocket; D2 (aa 221–224 in loop L-7I) forms the backside at the entrance of the pocket; region II (aa 239–245 in loop L-I8) forms the left side of the pocket. The I, post-I, and II consensus sequence motifs are also found in other types of SAM-utilizing methyltransferases (Dlakic, 2001). D1 and D2 do not contain sequence similarity with other methyltransferases.

While SAM interacts with a great number of neighboring residues via hydrophobic and van der Waals interactions, only five amino acids make direct hydrogen bonds with SAM via side-chain atoms (Figure 3D). Three of the residues, namely Asp161, Gln168, and Glu186, reside in the I and post-I motifs (Figure 1A). Asp161 is

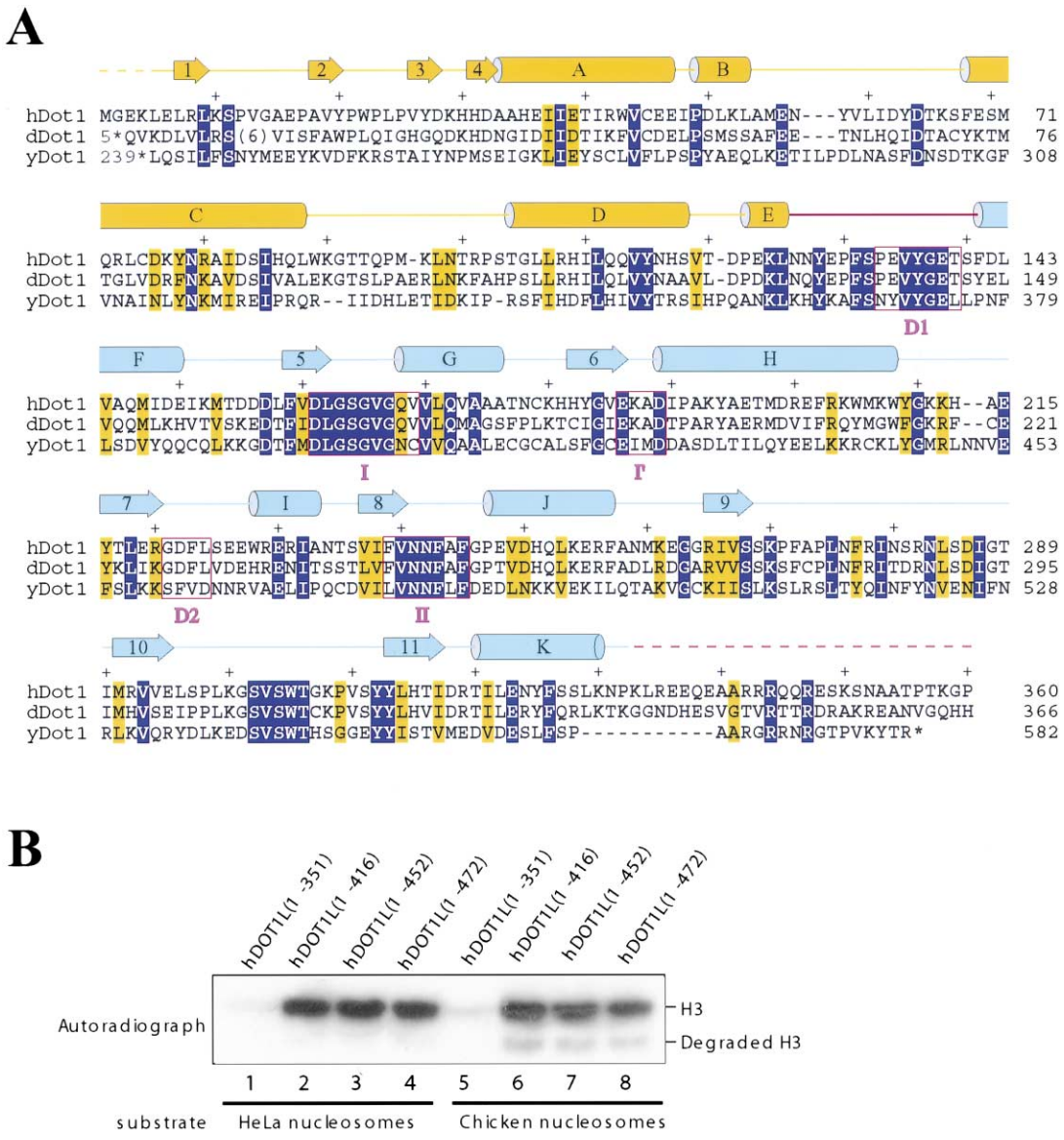

Figure 1. The Catalytic Domain of Dot1-like Proteins

(A) Approximately 360 amino acids at the N terminus of hDOT1L share significant sequence homology with the C-terminal region of yeast Dot1 (yDot1) and the N terminus region of a *Drosophila* homolog (dDot1). Identical residues between the three proteins are shown in white letters over a blue background, and similar amino acids are highlighted in yellow. Amino acids enclosed in red rectangles interact with SAM. Every ten amino acids are marked with a "+" sign, and secondary structure elements are shown above the sequence. Yellow, cyan, and red indicate different regions of the structure as shown in Figure 2A. Disordered regions are indicated by dashed lines.

(B) An autoradiograph showing that hDOT1L(1–416) is an active histone methyltransferase. Purified HeLa and chicken nucleosomes were used in the assay.

part of the highly conserved DxGxGxG signature motif found in SAM binding proteins (Dlakic, 2001). In the hDOT1L structure, the DLGSGVGQ sequence of amino acids forms the backside of the surface surrounding the sugar and methionyl moieties. Mutating the glycine residues abolishes the enzymatic activity of both yeast and human Dot1s, thus highlighting the critical importance of maintaining an intact SAM binding pocket for the function of Dot1 (Feng et al., 2002; Lacoste et al., 2002; Ng et al., 2002a; van Leeuwen et al., 2002). Gln168 is not conserved in other types of SAM-utilizing methyltransferases, but this position is always occupied by a polar residue in the Dot1 family of proteins, suggesting that the hydrogen bond interaction with SAM is important for the Dot1 family of enzymes. The carboxylate group of Glu186 interacts with the ribose hydroxyl groups of SAM via two hydrogen bonds. The carboxylate-ribose interaction is likely to be conserved because either a glutamate or an aspartate is found in the post-I position of other SAM-dependent methyltransferases. Two additional hydrogen bonds with SAM involve Thr139 and Asp222, located in D1 and D2 regions, respectively. However, neither residue is conserved even among Dot1 family members. None of the amino acids

Table 1. Summary of Crystallographic Analysis

Data Sets	hDOT1L(1–416)	hDOT1L(1–351) λ1	λ2	λ3
Diffraction data				
Resolution (Å)	2.5	2.2	2.2	2.2
Measured reflections	175,389	287,450	301,630	287,730
Unique reflections	23,800	39,911	39,974	39,918
Completeness (%; $I/\sigma \geq 0$)	96.3	95.2	95.2	94.7
Average I/σ	18.3	15.4	12.8	14.2
R_{merge} (%)[A]	5.9	6.4	7.9	6.8
Phasing				
Phasing power[B] at 2.5 Å				
Anomalous/isomorphous		3.37/–	4.94/1.46	3.39/2.56
Overall figure of merit		0.76		
Refinement				
Resolution range (Å)	50.0–2.5			
R factor[C]/R_{free} (%)	20.4/24.6			
Number of protein/SAM atoms	2672/27			
Number of ions	1 acetate, 2 sulfate ions			
Number of water molecules	67			
Rms deviations				
Bond lengths	0.008 Å			
Bond angles	1.38°			
Dihedrals	21.66°			
Improper	0.92°			

[A] $R_{merge} = \Sigma|I - \langle I \rangle|/\Sigma\langle I \rangle$, where I and $\langle I \rangle$ are the measured and averaged intensities of multiple measurements of the same reflection. The summation is over all the observed reflections.
[B] Phasing power = rms. ($\langle F_H \rangle/E$), where F_H is the calculated structure factor of the heavy atoms and E is the residual lack of closure.
[C] R factor $= \Sigma||F_O| - |F_C||/\Sigma|F_O|$, where F_O denotes the observed structure factor amplitude and F_C denotes the structure factor calculated from the model.

in the D2 region appear to be conserved (Figure 1A). In contrast, four amino acids in D1, Val135, Tyr136, Gly137, and Glu138, are conserved among Dot1 family members, yet their side chains point away from SAM, and only van der Waals interactions with SAM are involved (Figure 4A). The D1 region resides in the loop connecting the N- and C-terminal domains.

A Possible Lysine Binding Channel

A close examination of the conserved residues in D1 reveals that they are involved in a network of conserved intramolecular interactions. Glu138 makes one hydrogen bond with Tyr115 of helix αD in the N-terminal domain. Glu138 also makes one hydrogen bond to the neighboring Tyr136, which interacts with Gln168 via a buried water molecule (Figure 4A). These interactions stabilize the positioning of loop-LEF and result in the formation of a channel leading to the inside of the SAM binding pocket (Figure 3B). The small size of the conserved glycine Gly137 between the two residues is critical to avoid blocking the channel. Across the channel from Gly137 is Phe243, which is located in region II. Phe243 does not contact SAM directly, and its side chain points outward and is solvent accessible. Interestingly, the phenyl ring interacts with Trp305 (3.9 Å apart), located at the tip of the loop connecting β10 and β11 (Figure 4A). Trp305 is also solvent exposed. Trp305 and neighboring Val303, Ser304, and Thr306 are conserved among Dot1 family members (Figure 1A). The high temperature factors of the amino acids located in the loop suggest that the loop could be dynamic in solution. Both Phe243 and Trp305 are exposed, and their interaction may stabilize the conformation of both residues in the nonhydrophobic environment. In the absence of such an interaction, Phe243 could adopt alternative conformations from an energetic standpoint. One possible alternative conformation would completely block the entrance of the channel. In fact, rudimentary electron density suggests that such a situation may exist as a minor conformation of Phe243. Thus, the Phe243-Trp305 interaction might gate or regulate the opening and closing of the putative lysine binding channel.

The narrowest dimension of the channel is approximately 4 Å, which can accommodate the passage of a mono-, di-, or trimethylated lysine. The methyl group of SAM is lined up with the channel inside the SAM binding pocket (Figure 3B). These observations suggest that Lys79 of histone H3 to access the active site through this channel. Four amino acids, Thr139, Asn241, Ser269, and Tyr312, form a significant portion of the channel (Figure 4A). Two of the residues, Asn241 and Tyr312, are highly conserved (Figure 1A), and we examined their effects in the HMTase activity of hDOT1L by mutagenesis. Changing Asn241 to an aspartic acid or alanine abolishes the HMTase activity (Figure 4B, lanes 2 and 3). Substituting Tyr312 with a phenylalanine has little effect. However, changing Tyr312 to an alanine abolishes the HMTase activity (Figure 4B, lanes 4 and 5). The activity loss of the three mutants is not due to protein misfolding, as circular dichroism reveals that the inactive N241A, N241A, and Y312A mutants have very similar spectra compared to that of the active Y321F mutant (Figure 4B). Thus, a subtle change around the channel identified in the structure has a dramatic effect on the

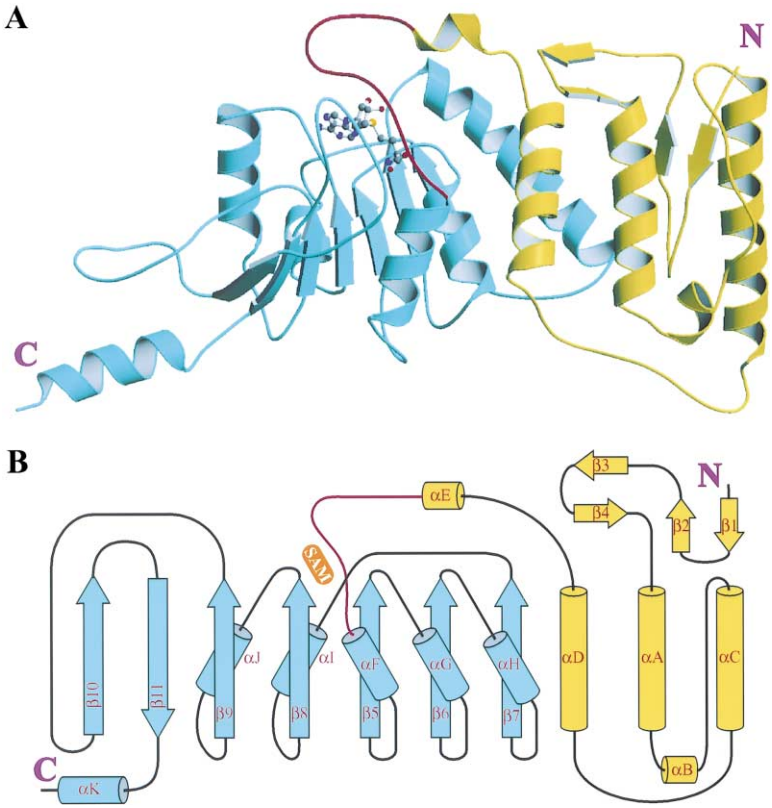

Figure 2. The Overall Structure of hDOT1L(1–416)

(A) Ribbon diagram of the hDOT1L(1–416) structure. The N-terminal region (aa 1–126) is colored yellow, the open α/β structure (aa 141–332) is shown in cyan, and the loop L-EF connecting the two regions is shown in red. The bound SAM molecule is shown in a ball-and-stick model (carbon, gray; nitrogen, blue; oxygen, red; sulfur, yellow).

(B) Topological diagram of hDOT1L(1–416). Secondary structure elements are colored using the same scheme as in (A). Helices are labeled alphabetically and strands are numbered numerically in an order from the N- to the C terminus. The position of the SAM binding site is also indicated.

HMTase activity, supporting the proposal that the channel is likely to bind lysine.

Disordered C-Terminal Region Is Important for Enzyme Activity and Nucleosome Binding

Amino acids C-terminal to αK (aa 333–416) are disordered in the structure. We showed earlier that certain amino acids located between residue 351 and 416 are important for the enzymatic activity of hDOT1L, as hDOT1L(1–416) can methylate oligonucleosomes efficiently but hDot1(1–351) cannot under the same condition (Figure 1B). Since the C-terminal end of hDOT1L(1–416) is spatially distant from the SAM binding site and because access to the active site is rather restricted, we reasoned that the C terminus might be important in substrate binding rather than having a direct involvement in catalysis. To evaluate this structure-guided prediction, we tested the enzymatic activity (Figure 5A) and nucleosome binding capability (Figure 5B) of a series of hDOT1L variants truncated at the disordered C-terminal region.

To identify the elements within the region between amino acids 351 and 416 that are critical for the HTMase activity of hDOT1L, we added back 19, 24, and 34 amino acids to the C terminus of hDOT1L(1–351), termed hDOT1L(1–370), hDOT1L(1–375), and hDOT1L(1–385), respectively, to see if the HMTase activity of hDOT1L could be restored. Figure 5A (lanes 2, 3, and 4) shows that all three constructs failed to restore the HMTase activity to a level comparable to that of hDOT1L(1–416). We then fused the last 26 amino acids of hDOT1L(1–416) to the inactive hDOT1L(1–375), creating an internal deletion mutant Δ(376–390). Figure 5A (lane 6) shows that this mutant is as active as the wild-type hDOT1L(1–416), suggesting that aa 391–416 are crucial for the HMTase activity. This region of hDOT1L is enriched with positively charged residues (Figure 5C), and we reasoned that these residues might be important for the HMTase activity. In support of this hypothesis, removing a stretch of amino acids containing nine of the 15 positively charged residues in this region, Δ(390–407), significantly reduced the enzymatic activity (Figure 5A, lane 7).

We next examined mononucleosome binding of hDot1(1–416) and several of its truncation variants by gel mobility shift assays (GMSA). Figure 5B shows that enzymatically active hDOT1L(1–416) and Δ(376–390) cause slower migration of mononucleosomes (lanes 2–4 and 5–7), while weakly active Δ(390–407) and inactive hDOT1L(1–385) show no detectable gel shifts within the sensitivity limit of our assay. The addition of SAM in the GMSA does not change the gel shift pattern in Figure 5B (data not shown). The above observation qualitatively correlates the nucleosome binding ability with the HMTase activity of hDOT1L proteins and implicates the importance of the positively charged C-terminal region in nucleosome binding. Our modeling of hDOT1L-nucleosome interaction suggests that the C-terminal region is near DNA (see Discussion section), which led us to hypothesize that the disordered C-terminal positively charged region might be involved in substrate binding by interacting with the negatively charged nucleosomal DNA. Consistent with this hypothesis, we can detect strong, but apparently non-sequence-specific, binding of hDOT1L(1–416) with DNA extracted from mono-

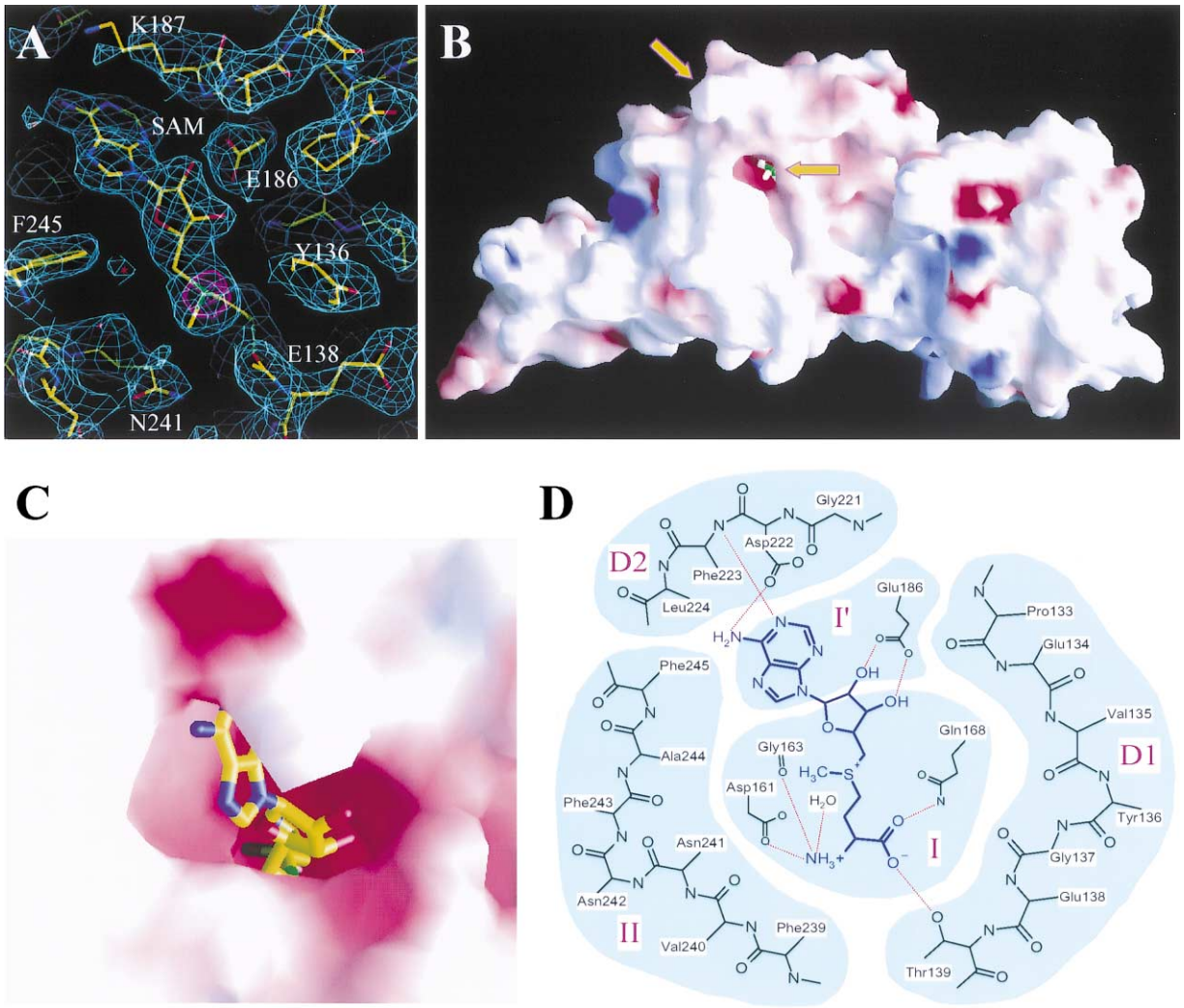

Figure 3. SAM Binding Pocket

(A) The 2.5 Å resolution MAD-phased electron density map around the SAM molecule is shown in meshed blue lines. The map starts at 1.5 σ level and is contoured at 1.0 σ interval. The density shown in magenta is a selenium anomalous difference map indicating selenium-substituted sulfur positions. The map was calculated using data collected at the peak wavelength of Se anomalous scattering, and the map is contoured at 1.0 σ interval starting at 10.0 σ level. The refined model is superimposed and several SAM-interacting residues are labeled.

(B) A surface representation of the hDOT1L(1–416) structure viewed from a direction similar to that in Figure 2A. Red, blue, and white indicate negatively charged, positively charged, and neutral protein surface regions, respectively. The methyl group and sulfur atom of SAM (carbon, white; sulfur, green) can be seen through the channel in the front indicated by a horizontal yellow arrow. The yellow arrow at the top indicates the entrance to the SAM binding pocket.

(C) The SAM binding pocket viewed from a direction indicated by the arrow at the top of (B). The inside of the SAM binding pocket is negatively charged. SAM is shown in a ball-and-stick model (carbon, yellow; nitrogen, blue; oxygen, red; sulfur, green).

(D) A schematic diagram showing SAM-protein interactions. Amino acids in five regions are involved in interaction with SAM (see Figure 1A), and these regions are shaded cyan. Most amino acids are involved in hydrophobic or van der Waals interactions. Specific interactions via hydrogen bonds are shown with red dotted lines.

nucleosomes (data not shown). To further probe the nature of hDOT1L-nucleosome interaction, we used a 21-residue highly charged peptide, C-RS$_{10}$, containing ten tandem repeats of arginine and serine as a competitor in our nucleosome binding and HMTase assays. Adding the C-RS$_{10}$ peptide in GMSA prevented mononucleosomes from entering the gel well (data not shown). The HMTase activities of hDOT1L(1–416) and Δ(376–390), shown in lanes 1–4 and 5–8 of Figure 5D, respectively, are effectively competed with increasing concentrations of the peptide. These results clearly demonstrate the functional importance of the C-terminal charged region of hDOT1L(1–416).

Discussion

The overall structure of hDOT1L reveals a novel N-terminal domain and a C-terminal domain having a familiar open α/β structure similar to a number of SAM-dependent methyltransferases (Cheng and Roberts, 2001).

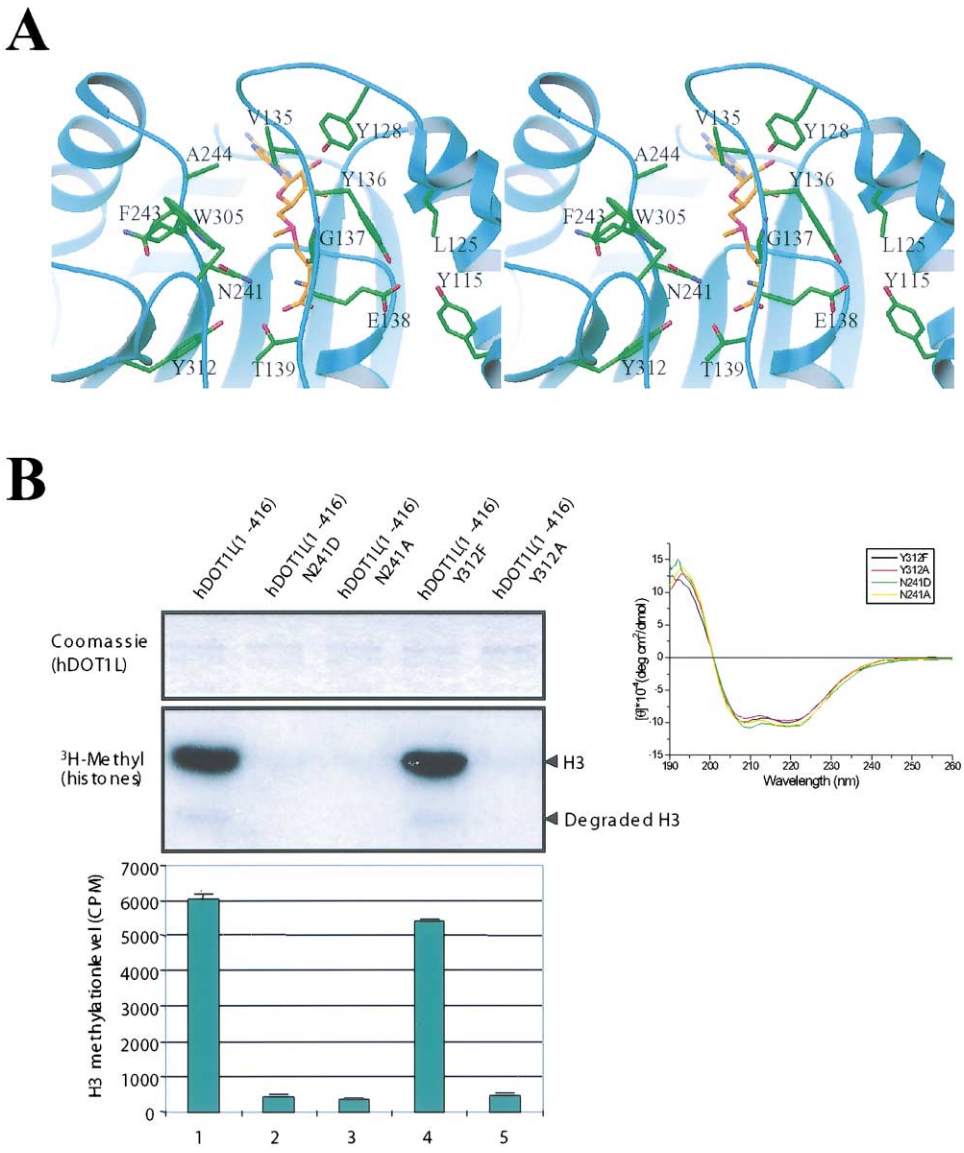

Figure 4. A Possible Lysine Binding Channel

(A) A stereo diagram showing the side chains of residues in the D1 region and residues that are important for the formation of a putative lysine binding channel (carbon, green; nitrogen, blue; oxygen, red). A ribbon diagram showing the overall structure of hDOT1L(1–416) and a stick-and-ball model of SAM (carbon, brown; sulfur, magenta) are superimposed.

(B) HMTase assay of hDOT1L point mutants. Top left: equal amount of mutant proteins used for HMTase assays. Middle left: an autoradiograph showing the HMTase activity of hDOT1L mutants. Bottom: a histogram representation showing the quantitation of the HMTase activity of these mutants. Top right: superimposed UV circular dichroism spectra of the hDOT1L mutants. The spectrum for each mutant is colored according to the scheme indicated in the enclosed box.

This is distinctly different from the other class of histone lysine methyltransferases known to date, namely, SET domain proteins (Jacobs et al., 2002; Min et al., 2002; Trievel et al., 2002; Wilson et al., 2002; Zhang et al., 2002). Our structural and biochemical data presented here provide significant insights into the answers to a number of important questions pertaining to Dot1 proteins, including: (1) what are the determinants for lysine specificity?, (2) how does hDOT1L interact with the nucleosome?, and (3) what is the catalytic mechanism of the enzymes?

Determinants of Lysine Specificity and Catalytic Mechanism

Dot1 proteins are the only non-SET domain histone lysine methyltransferase identified to date. The 2.5 Å hDOT1L structure reveals that it shares a similar open α/β structure with other classical methyltransferases (Cheng and Roberts, 2001), yet none of these other methyltransferases is known to methylate lysine. A comparison of these structures should enable us to detect unique features of Dot1 proteins that may account for their activities toward lysine.

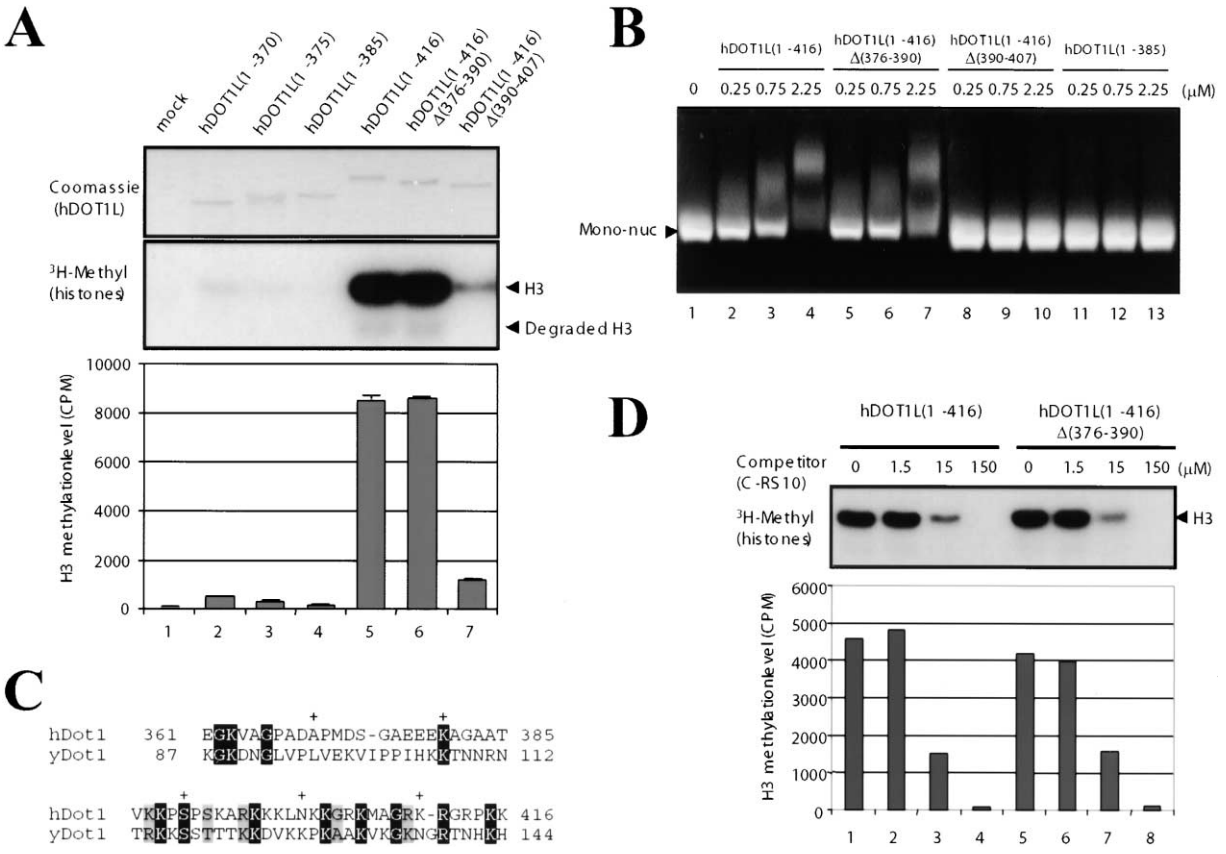

Figure 5. HMTase Activity and Mononucleosome Binding of hDOT1L Derivatives

(A) The positively charged region (aa 390–407) is important for the HMTase activity of hDOT1L. Equal amounts of hDOT1L deletion mutants (top), indicated on top of the panel, were analyzed for HMTase activity (middle) and the quantitation is presented in the bottom panel.

(B) The positively charged region required for HMTase activity is also required for mononucleosome binding. Mononucleosome binding of the hDOT1L deletions mutants was analyzed by gel mobility shift assays at three different protein concentrations.

(C) Amino acid sequence (aa 361–416) of the C-terminal of hDOT1L(1–416) encompassing the charged region required for the HMTase activity. A stretch of similar amino acids located at the N-terminal of yeast Dot1 is also shown. Identical amino acids between the yeast and human proteins are shown with white letters over a dark background, and similar amino acids are highlighted with gray shading.

(D) The HMTase activities of hDOT1L(1–416) and Δ(376–390) in the absence and presence of increasing concentrations of a positively charged competitor peptide were analyzed by autoradiograph (top) and the quantitation is shown at the bottom panel.

A search of the protein databank using the DALI server revealed clear structural resemblance of the catalytic domain of hDOT1L with a variety of classical methyltransferases. The closest match of any characterized methyltransferases with hDOT1L is the catechol O-methyltransferase (COMPT; pdb code 1vid; Z score 12.9) (Vidgren et al., 1994). COMPT catalyzes the transfer of the methyl group from SAM to one hydroxyl group of catechols, and the structure of COMPT in complex with SAM and a competitive inhibitor (3,5-dinitrocatechol), which mimics the structure of catechol substrates, is shown in Figure 6A. It is also informative to compare the hDOT1L structure with the structure of a peptide and cofactor bound ternary complex of the *P. furiosus* L-isoaspartyl methyltransferase (IAMT) (Figure 6B; pdb code 1jg3; Griffith et al., 2001). Although the connectivity of the sixth and seventh β strands of IAMT differs from that in others, it is clear that hDOT1L shares a conserved open α/β core with COMPT and IAMT (Figures 6A–6C). Both COMPT and IAMT also have a helical N-terminal domain similar to hDOT1L, though the size and specific configuration of the N-terminal domains differ among these proteins. The linker connecting the two domains (colored red in Figures 6A–6C) plays a crucial role in cofactor and substrate bindings in each of these proteins. COMPT has a short inter-domain loop, which results in a relatively shallow and wide binding site for the relatively bulky substrate-like inhibitor. The substrate binding site is formed by amino acids located in the inter-domain linker and the C-terminal ends of the fourth, fifth, and sixth β strands. Amazingly, when the two structures are superimposed, the COMPT substrate binding site coincides with the putative lysine binding channel we identified in hDOT1L (Figures 3B and 4A). The methyl-accepting L-isoaspartate of the IAMT substrate polypeptide is bound in a channel leading to the mostly buried SAM binding pocket, much the same as the proposed lysine binding channel of hDOT1L (Figure 3B). The L-isoaspartate binding channel is formed between the inter-domain linker, which includes a short β strand, and a region immediately C-terminal to the seventh β strand of the open α/β core. As pointed out earlier, IAMT has an un-

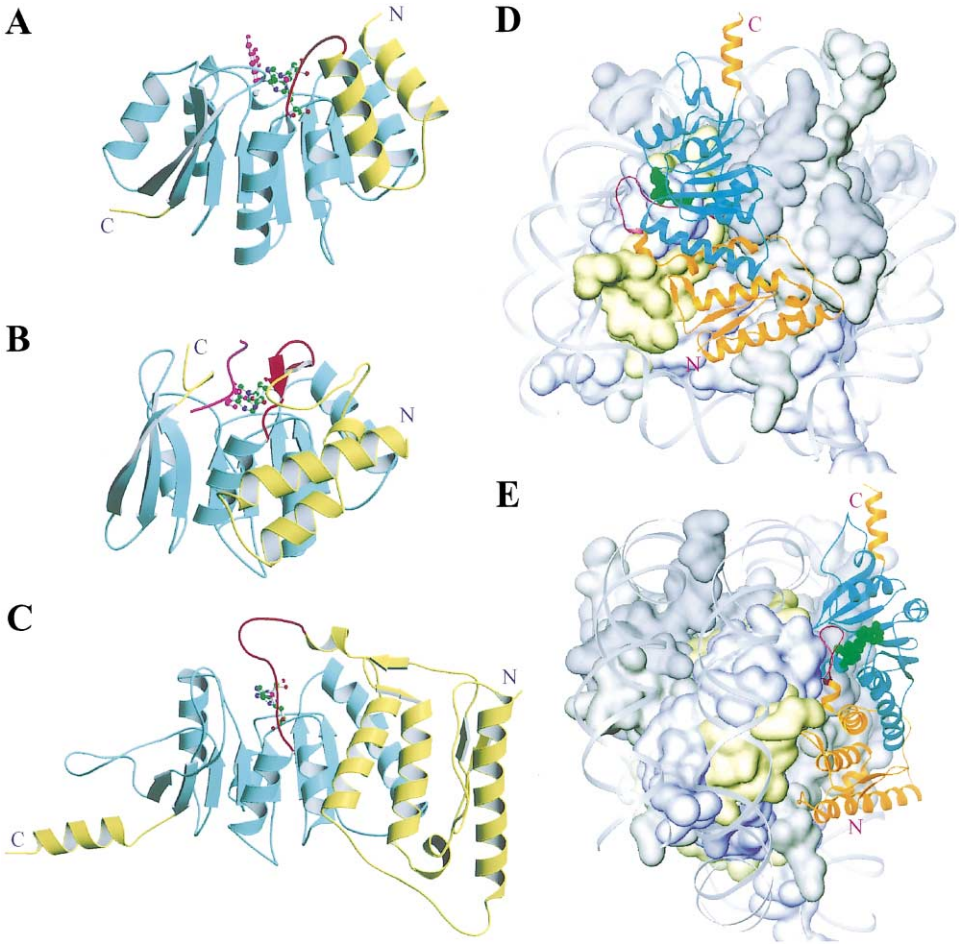

Figure 6. Structure Comparison and Modeling of hDOT1L-Nucleosome Core Particle Interaction

(A) The structure of catechol O-methyltransferase is shown in a ribbon diagram. The open α/β structure is colored cyan, the inter-domain linker is colored red, and other regions are colored yellow. SAM and a catechol-like inhibitor are shown as stick-and-ball models. The inhibitor is colored magenta, and the coloring scheme for SAM is: carbon, green; nitrogen, blue; oxygen, red; sulfur, magenta.

(B) The structure of *P. furiosus* L-isoaspartyl methyltransferase complexed with adenosine and a methyl accepting peptide. The same coloring scheme for the protein and adenosine (as for SAM), as in (A), is used. The peptide main chain and the side chain of the methyl acceptor L-isoaspartate (ball-and-stick) are shown in magenta.

(C) The structure of hDOT1L(1–416) in a ribbon representation, colored as in (A). All three structures shown in (A)–(C) are orientated similarly.

(D) The modeled hDOT1L(1–416)-nucleosome core particle complex viewed along the superhelical axis. DNA is shown in a ribbon representation (gray) and hDOT1L(1–416) is colored as in (C), except that yellow is changed to orange. SAM is shown as a green CPK model. Histones are shown in surface models (H3, blue; H4, gold; H2A, light gray; H2B, dark gray).

(E) A side view of the same model, as shown in (D).

usual connectivity between the last two β strands; thus, the seventh β strand in IAMT is equivalent to the sixth strand in the open α/β structure of COMPT and hDOT1L. Most significantly, the L-isoaspartate binding channel lines up with the proposed lysine binding channel when the IAMT and hDOT1L structures are superimposed.

The above observations strongly support the notion that the inter-domain linker and the region C-terminal to $\beta6$ are important for determining methyltransferase types and substrate specificity. We pointed out earlier that the D1 region in loop L-EF is likely to be crucial for conveying the lysine specificity of hDOT1L. This region interacts with the N-terminal domain to stabilize the conformation of the inter-domain linker and directly contributes to the formation of the putative lysine binding channel likely to be used to access the active site by Lys79 of histone H3. In addition to loop L-EF, the loop connecting $\beta6$ and $\beta7$ and C-terminal end of $\beta4$ also form part of the proposed lysine binding channel. Asn241 is located at the end of $\beta4$, and Tyr312 is located in the $\beta6$-$\beta7$ loop. Disruption of the hDOT1L HMTase activity by N241D, N241A, and Y312A mutations suggests that a precise architecture of the proposed lysine binding channel is required for the function of hDOT1L.

The methyl group of SAM is aligned with the channel, which is ideally suited for an "in-line" methyl transfer reaction (Figures 3B and 4A). The size of the channel can accommodate the binding of a lysine; the narrowest dimension of the channel in hDOT1L is approximately 4 Å, which can accommodate the passage of a mono-, di-, or trimethylated lysine. Although the channel is not very deep (approximately 4 Å from the entrance of the

channel to the methyl group of SAM), sufficient proximity to the donor methyl group for methyl transfer requires passage through the ~4 Å "bottle neck." In methyl transfer reactions, the methyl acceptor lysine is deprotonated, and the lone electron pair of the amine group carries out nucleophilic attack of the C-S bond between the methyl group and the sulfur atom in SAM. It appears that different methyltransferases use distinct mechanisms to achieve deprotonation. For example, it has been suggested that two conserved glutamate residues in the active site of arginine methyltransferase PRMT3 can effectively deprotonate an arginine (Zhang et al., 2000). This process is still considerably unclear in lysine methyltransferases, as no negatively charged residues capable of carrying out the deprotonation step are present in the active sites (Jacobs et al., 2002; Min et al., 2002; Trievel et al., 2002; Wilson et al., 2002; Zhang et al., 2002). It has been suggested that the overall negatively charged environment within the active site may facilitate deprotonation (Zhang et al., 2002). An alternative scenario involving a particular tyrosine residue has also been proposed (Trievel et al., 2002; Zhang et al., 2002). In this regard, the catalytic mechanism of hDOT1L appears to resemble that of SET domain proteins. First, there are no negatively charged residues in the vicinity of the active site. Second, the SAM binding pocket has a negatively charged environment. Third, Tyr312 is situated near the entry of the putative lysine binding channel less than 7.0 Å away from the SAM methyl group. However, the hydroxyl group of Tyr312 appears not to be important in the deprotonation process, as the Y312F mutant of hDOT1L is fully functional. Interestingly, the position corresponding to Tyr312 of hDOT1L in a putative *C. elegans* homolog is a phenylalanine. Thus, hDOT1L may rely on the negatively charged environment surrounding the SAM binding site to achieve lysine deprotonation. It is interesting to note that the negatively charged carboxylate group of SAM is compensated by a conserved arginine in PRMT3. In SET domain proteins, an arginine also interacts with the SAM carboxylate, although the arginine may be substituted with a histidine or with some other residues in rare cases. In hDOT1L, the SAM carboxylate group does not interact with a positively charged residue, and the negative charge of the carboxylate group may directly participate in the deprotonation of the methyl acceptor lysine.

The putative lysine binding channel and the entry of the SAM binding pocket is separate in the hDOT1L structure. Thus, it appears that substrate binding is independent of SAM binding. As a consequence, it is possible that Lys79 of histone H3 can undergo multiple rounds of methylation without being released from the enzyme. A similar suggestion has also been proposed for certain SET domain proteins (Trievel et al., 2002). Another interesting feature of hDOT1L is that the C-terminal charged region is likely to bind to DNA, which may provide an additional, more flexible foothold on the substrate while SAM is exchanged prior to an additional round of catalysis. However, one has to keep in mind that, unlike the situation in arginine dimethylation, a mixture of mono-, di-, and trimethylated H3 Lys79 coexist, and ~50% of bulk histone H3 is trimethylated in yeast (van Leeuwen et al., 2002). A complex spectrum of differential methylation states are also found for certain lysine residues methylated by SET domain proteins. A mechanistic understanding of histone lysine mono-, di-, and trimethylation is lacking at present.

Modeling hDOT1L-Nucleosome Interaction

Because the crystal structure of the nucleosome core particle is known (Luger et al., 1997), hDOT1L-nucleosome interactions can be explored by molecular modeling using the two known structures. The following criteria were used as guidelines for modeling. (1) Lys79 of histone H3 has to be capable of reaching the methyl group of SAM. Consistent with our structure-based predictions, we find that the proposed lysine binding channel is the only way Lys79 can access the donor methyl group in the hDOT1L(1–416) structure. The side-chain conformation of Lys79 is allowed to vary, because it is not engaged in interactions with DNA, other histones, or stabilization of the histone H3 structure. (2) No steric clashes between structurally rigid secondary structure elements of hDOT1L(1–416) and histones are allowed. (3) Clashes with DNA are disallowed. Certain clashes via potentially flexible loops are tolerated as protein-protein or protein-DNA interactions may alter the conformation of these loops.

Although the aforementioned constraints do not give rise to a unique model, only a limited number of models are possible because aligning Lys79 with the methyl group of SAM effectively reduces the degree of freedom to rotations around an axis approximately parallel to the superhelical axis of the nucleosome and passing through the putative lysine binding channel. A representative model of hDOT1L-nucleosome complex satisfying the above criteria is shown in Figures 6D and 6E. A number of features of hDOT1L-nucleosome interactions are clear from the modeling. First, a segment of the loop connecting β10 and β11 always clashes with core histones in all the possible models. The loop segment is located in front of the potential lysine binding channel and obstructs the histone from getting close to the channel. The loop segment has high temperature factors, indicating that the loop is capable of undergoing large conformational changes. Furthermore, the observed conformation of this loop segment is partially stabilized by crystal packing, which suggests that the loop may have a different conformation in solution. A conserved stretch of amino acids Val303, Ser304, Trp305, and Thr306 are located at the tip of this loop. We described earlier that Trp305 interacts with Phe243 situated at the entry of the potential lysine binding channel. Our modeling suggests that this loop connecting β10 and β11 has to undergo a conformational change for lysine 79 of histone H3 to reach the active site, thus the Trp305-Phe243 interaction will most likely be disrupted. Interestingly, the residue preceding Lys79 is a phenylalanine, which is located in the same turn connecting two histone-fold helices and on the nucleosome surface. It is possible that Phe243 of hDOT1L can pack with Phe78 of histone H3.

Second, hDOT1L interacts extensively with the nucleosome. In most of the possible models, hDOT1L interacts with all four histones on the same side of the nucleosome disk surface. The contiguous nucleosome surface surrounding Lys79 of H3 is important for silenc-

ing (Park et al., 2002) and is within the area of interactions with hDOT1L in all of the possible models. This extensive interaction area provides an explanation as to why Dot1 proteins methylate nucleosomal substrates rather than free histones. It also provides another possible explanation, in addition to the ones already proposed (Park et al., 2002), for the silencing function of the important area surrounding histone H3 Lys79; i.e., mutation of histones in this area may alter the Dot1 recognition surface and result in the loss of histone H3 Lys79 methylation.

Third, the C-terminal helix αK is near DNA. We have shown that the disordered region C-terminal to αK of hDOT1L is important for binding to DNA and nucleosomes. The proximity of αK to DNA shows that the disordered C-terminal fragment is in a position to contact DNA directly. Yeast Dot1 lacks a C-terminal extension that could potentially serve the same function as the charged C-terminal region of hDOT1L (Figure 1A). Interestingly, a stretch of amino acids N-terminal to the conserved region is enriched with positively residues (Figure 5C), which may function equivalently as the charged region in hDOT1L. An extension at the N terminus of the catalytic core can also reach DNA, as the N terminus of the conserved catalytic core is located in the opposite direction of the elongated structure (Figures 2A, 6D, and 6E).

Fourth, the C-terminal half of the structure, i.e., the open α/β structure, and regions adjacent to helix αK in particular, are near the C-terminal helix of histone H2B. It has been shown previously that ubiquitination of Lys123 of histone H2B in yeast is required for the methylation of H3 Lys79 (Briggs et al., 2002; Ng et al., 2002b). Our modeling suggests that the catalytic core of hDOT1L may be able to interact with the ubiquitin covalently attached to the C-terminal helix of histone H2B. We have previously noticed that hDOT1L was unable to methylate H3 on oligonucleosomes assembled from recombinant core histones in vitro, while oligonucleosomes reconstituted from HeLa core histones are excellent substrates (Ng et al., 2002b). This observation is consistent with the proposal that ubiquitination of histone H2B may serve as a spacer between adjacent nucleosomes to allow access to histone modification enzymes (Sun and Allis, 2002). It would be interesting to know whether hDOT1L methylates recombinant mononucleosomes assembled from histones produced in *E. coli* and whether Dot1 proteins interact with ubiquitinated histone H2B. Such an interaction could be physiologically significant in recruiting Dot1 proteins to specific chromatin domains.

Conclusions

The structure of the catalytic domain of hDOT1L in complex with SAM reveals a unique organization of the N-terminal helical domain and the C-terminal open α/β structure that resembles certain SAM-dependent methyltransferases. SAM is bound between the loop connecting the two domains and the open α/β structure. We have identified a potential lysine binding channel in the structure and shown that the C-terminal disordered region is important for nucleosome binding and for the activity of hDOT1L. Our modeling provides important insights into nucleosome-hDOT1L interactions and the methylation of nucleosomal histones. Future biochemical and structural studies of the hDOT1L-nucleosome complex are needed for understanding the exact mode of hDOT1L-nucleosome interactions and the molecular mechanism of nucleosomal histone methylation.

Experimental Procedures

Protein Methods

To produce recombinant hDOT1L fragments containing N-terminal 351, 370, 375, 385, 416, 452, and 472 amino acids in *E. coli*, corresponding cDNA fragments were amplified by PCR and cloned into a pGEX-KG vector (between EcoRI-XhoI sites). The presence of correct inserts were verified by DNA sequencing. The plasmid expressing the first 416 residue of hDOT1L, pGEX-KG/hDOT1L(1–416), was used for subsequent Δ(376–390) and Δ(390–407) deletions and N241D, N241A, Y312F, and Y312A point mutations. The deletion and point mutations were carried out by PCR and confirmed by sequencing. GST-fused hDOT1L fragments and mutants were first purified on a glutathione-sepharose column. The GST-tag was then removed by thrombin digestion in the column. The eluted proteins were further purified on HiTrap-SP and Superdex-75 columns (AP Biotech).

Two fragments, hDOT1L(1–351) and hDOT1L(1–416), were used for crystallization. Purified proteins were concentrated to 25–30 mg/ml for crystallization in a buffer containing 20 mM Tris (pH 8.0), 200 mM NaCl, 1 mM EDTA, and 1 mM DTT. Crystals of native hDOT1L(1–351), hDOT1L(1–351), and hDOT1L(1–416) complexed with SAM (3 mM final concentration) were all grown in conditions containing 0.1 M sodium acetate, 1.3–1.9 M ammonium sulfate, and 1–3 mM Tris (2-carboxyethyl) phosphine (TCEP). SeMet substituted hDOT1L(1–351) was produced using *E. coli* strain DL41 (DE3) in a defined medium containing 30 mg/L of SeMet (Hendrickson et al., 1990).

Circular dichroism measurements of the point mutants were carried out on a Jasco J-710 spectropolarimeter at room temperature. Mutant hDOT1L proteins at ~2 mg/ml concentrations were used for measurements, and multiple scans of each sample were averaged.

Crystallographic Analysis

All diffraction data were collected at 95° K using a CCD detector (ADSC) at beamline X26C of National Synchrotron Light Source (NSLS), Brookhaven National Laboratory. Raw data were processed using the HKL software package (Otwinowski and Minor, 1997). Crystals of hDOT1L(1–351), hDOT1L(1–351)-SAM, and hDOT1L(1–416)-SAM complexes all belong to the $P6_5$ spacegroup. The cell dimensions for these three different crystals are 152.9 Å × 152.9 Å × 51.1 Å, 152.9 Å × 152.9 Å × 50.8 Å, and 152.8 Å × 152.8 Å × 50.8 Å, respectively. We confirmed that the crystallized proteins contain intact hDOT1L(1–351) and hDOT1L(1–416) fragments by recovering the protein from the crystals and analyzed them by SDS-PAGE and mass spectrometry. The crystal structure was solved by the method of SeMet multiple wavelength anomalous diffraction (MAD). Three 2.2 Å resolution MAD data sets (inflection point, $\lambda_1 = 0.9806$ Å; peak, $\lambda_2 = 0.9797$ Å; remote, $\lambda_3 = 0.93$ Å) of a SeMet hDOT1L(1–351)-SAM complex crystal were used. Ten Se positions were identified (nine from SeMet, one from Se-substituted SAM) and used for MAD phasing to 2.5 Å resolution. The PHASES (Furey and Swaminathan, 1997) and SOLVE (Terwilliger and Berendzen, 1999) software packages were used for phasing and solvent flattening. Detailed phasing statistics are shown in Table 1. The 2.5 Å resolution SeMet MAD electron density map was used for model building. Model building and refitting was performed using the graphics program O (Jones et al., 1991). An unrefined hDOT1L(1–351) model was then used as initial model for refinement against the hDOT1L(1–416) data. The hDOT1L(1–416) structure was refined using CNS (Brunger et al., 1998). 8% of the data were used for crossvalidation purposes and were not included in the refinement. Refinement statistics can also be found in Table 1. Modeling of hDOT1L(1–416)-nucleosome core particle (pdb code: 1AOI) interaction was carried out manually using O. Figures were prepared using the Molscript (Kraulis, 1991), Raster3D (Merritt and Bacon, 1997), Ribbons (Carson, 1997), and Grasp (Nicholls et al., 1991) software packages.

Histone Methyltransferase Assay

Histone methyltransferase assay was performed as previously described (Wang et al., 2001). Oligonucleosomes purified from HeLa cells or chicken blood cells were quantified and equal concentration was used in each assay. Briefly, recombinant hDOT1L proteins (0.75 μM) were incubated with oligonucleosomes (1.5 μM) in reactions containing 20 mM Tris-HCl (pH 8.0), 4 mM EDTA, 1 mM PMSF, 0.5 mM DTT, and 0.5 μl ^{3}H-SAM (15 Ci/mM; NEN Life Science Products) for 1 hr at 30°C. For peptide competition, 1.5 μM, 15 μM, or 150 μM of positively charged C-RS10 peptide (C-RSRSRSRSRSRSRS RSRSRS) was added to reactions. Reactions were stopped by addition of 7 μl of 5 × SDS loading buffer, and proteins were separated in an 18% SDS-PAGE. Following stain and destain of Coomassie blue, gels were treated with Rapid Autoradiography Enhancer (NEN Life Science Products) for 40 min, dried, and exposed to X-ray films. For quantitations, gel slices were excised and counted with scintillation counting. Two independent experiments were performed at the same time and average ^{3}H (CPM) values were used in the figures.

Gel Shift Assay

Different amount of recombinant hDOT1L proteins were incubated with HeLa mononucleosomes (1.5 μM) under the same conditions as those in the histone methyltransferase assay. Following incubation for 1 hr at 30°C, samples were resolved in a 1.8% agarose gel and visualized under UV by ethidium bromide staining.

Acknowledgments

We thank Dieter Schneider, Annie Heroux, Hang Shi, and Ying Huang for technical assistance; Adrian Krainer for C-RS$_{10}$ peptide; Jia Fang for purified nucleosomes; Michael Myers for mass analyses; and Arne Stenlund, Lee Henry, Ira Hall, and Michael Myers for comments on the manuscript. Y.Z. is a Kimmel Scholar and is supported by grants from NIH (GM63070) and The American Cancer Society (RSG-00-351-01-GMC). R.M.-X. is supported by the W.M. Keck foundation and a grant from NIH (GM63716).

Received: October 15, 2002
Revised: January 14, 2003

References

Briggs, S.D., Xiao, T., Sun, Z.W., Caldwell, J.A., Shabanowitz, J., Hunt, D.F., Allis, C.D., and Strahl, B.D. (2002). Gene silencing: trans-histone regulatory pathway in chromatin. Nature *418*, 498.

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., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D *54*, 905–921.

Carson, M. (1997). Ribbons. Methods Enzymol. *277*, 493–505.

Cheng, X., and Roberts, R.J. (2001). AdoMet-dependent methylation, DNA methyltransferases and base flipping. Nucleic Acids Res. *29*, 3784–3795.

Dlakic, M. (2001). Chromatin silencing protein and pachytene checkpoint regulator Dot1p has a methyltransferase fold. Trends Biochem. Sci. *26*, 405–407.

Dover, J., Schneider, J., Tawiah-Boateng, M.A., Wood, A., Dean, K., Johnston, M., and Shilatifard, A. (2002). Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. J. Biol. Chem. *277*, 28368–28371.

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

Feng, Q., Wang, H., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Struhl, K., and Zhang, Y. (2002). Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. Curr. Biol. *12*, 1052–1058.

Furey, W., and Swaminathan, S. (1997). PHASES-95: a program package for processing and analyzing diffraction data from macromolecules. Methods Enzymol. *277*, 590–620.

Grewal, S.I., and Elgin, S.C. (2002). Heterochromatin: new possibilities for the inheritance of structure. Curr. Opin. Genet. Dev. *12*, 178–187.

Griffith, S.C., Sawaya, M.R., Boutz, D.R., Thapar, N., Katz, J.E., Clarke, S., and Yeates, T.O. (2001). Crystal structure of a protein repair methyltransferase from Pyrococcus furiosus with its L-isoaspartyl peptide substrate. J. Mol. Biol. *313*, 1103–1116.

Hendrickson, W.A., Horton, J.R., and LeMaster, D.M. (1990). Selenomethionyl proteins produced for analysis by multiwavelength anomalous diffraction (MAD): a vehicle for direct determination of three-dimensional structure. EMBO J. *9*, 1665–1672.

Jacobs, S., Harp, J.M., Devarakonda, S., Kim, Y., Rastinejad, F., and Khorasanizadeh, S. (2002). The active site of the SET domain is constructed on a knot. Nat. Struct. Biol. *9*, 833–838.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Kraulis, P.J. (1991). Molscript–a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

Lachner, M., and Jenuwein, T. (2002). The many faces of histone lysine methylation. Curr. Opin. Cell Biol. *14*, 286–298.

Lacoste, N., Utley, R.T., Hunter, J.M., Poirier, G.G., and Cote, J. (2002). Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. J. Biol. Chem. *277*, 30421–30424.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to produce both detailed and schematic plots of proteins. J. Appl. Crystallogr. *24*, 946–956.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature *389*, 251–260.

Merritt, E.A., and Bacon, D.J. (1997). Raster3D: photorealistic molecular graphics. Methods Enzymol. *277*, 505–524.

Min, J., Zhang, X., Cheng, X., Grewal, S.I., and Xu, R.M. (2002). Structure of the SET domain histone lysine methyltransferase Clr4. Nat. Struct. Biol. *9*, 828–833.

Ng, H.H., Feng, Q., Wang, H., Erdjument-Bromage, H., Tempst, P., Zhang, Y., and Struhl, K. (2002a). Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. Genes Dev. *16*, 1518–1527.

Ng, H.H., Xu, R.M., Zhang, Y., and Struhl, K. (2002b). Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3-lysine 79. J. Biol. Chem. *277*, 34655–34657.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

Nishioka, K., Rice, J.C., Sarma, K., Erdjument-Bromage, H., Werner, J., Wang, Y., Chuikov, S., Valenzuela, P., Tempst, P., Steward, R., et al. (2002). PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. Mol. Cell *9*, 1201–1213.

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

Park, J.H., Cosgrove, M.S., Youngman, E., Wolberger, C., and Boeke, J.D. (2002). A core nucleosome surface crucial for transcriptional silencing. Nat. Genet. *32*, 273–279.

San-Segundo, P.A., and Roeder, G.S. (2000). Role for the silencing protein Dot1 in meiotic checkpoint control. Mol. Biol. Cell *11*, 3601–3615.

Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. Nature *419*, 407–411.

Singer, M.S., Kahana, A., Wolf, A.J., Meisinger, L.L., Peterson, S.E., Goggin, C., Mahowald, M., and Gottschling, D.E. (1998). Identifica-

tion of high-copy disruptors of telomeric silencing in Saccharomyces cerevisiae. Genetics *150*, 613–632.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. Nature *403*, 41–45.

Strahl, B.D., Grant, P.A., Briggs, S.D., Sun, Z.W., Bone, J.R., Caldwell, J.A., Mollah, S., Cook, R.G., Shabanowitz, J., Hunt, D.F., and Allis, C.D. (2002). Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. Mol. Cell. Biol. *22*, 1298–1306.

Sun, Z.W., and Allis, C.D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature *418*, 104–108.

Terwilliger, T.C., and Berendzen, J. (1999). Automated MAD and MIR structure solution. Acta Crystallogr. D *55*, 849–861.

Trievel, R.C., Beach, B.M., Dirk, L.M.A., Houtz, R.L., and Hurley, J.H. (2002). Structure and catalytic mechanism of a SET domain protein methyltransferase. Cell *111*, 91–103.

Turner, B.M. (2000). Histone acetylation and an epigenetic code. Bioessays *22*, 836–845.

van Leeuwen, F., Gafken, P.R., and Gottschling, D.E. (2002). Dot1p modulates silencing in yeast by methylation of the nucleosome core. Cell *109*, 745–756.

Vidgren, J., Svensson, L.A., and Liljas, A. (1994). Crystal structure of catechol O-methyltransferase. Nature *368*, 354–358.

Wang, H., Huang, Z.Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B.D., Briggs, S.D., Allis, C.D., Wong, J., Tempst, P., and Zhang, Y. (2001). Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. Science *293*, 853–857.

White, C.L., Suto, R.K., and Luger, K. (2001). Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. EMBO J. *20*, 5207–5218.

Wilson, J.R., Jing, C., Walker, P.A., Martin, S.R., Howell, S.A., Blackburn, G.M., Gamblin, S.J., and Xiao, B. (2002). Crystal structure and functional analysis of the histone methyltransferase SET7/9. Cell *111*, 105–115.

Wu, J., and Grunstein, M. (2000). 25 years after the nucleosome model: chromatin modifications. Trends Biochem. Sci. *25*, 619–623.

Zhang, X., Zhou, L., and Cheng, X. (2000). Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. EMBO J. *19*, 3509–3519.

Zhang, X., Tamaru, H., Khan, H.I., Horton, J.R., Keefe, L.J., Selker, E.U., and Cheng, X. (2002). Structure of the *Neurospora* SET domain protein DIM-5, a histone H3 lysine methyltransferase. Cell *111*, 117–127.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. Genes Dev. *15*, 2343–2360.

Accession Numbers

Atomic coordinates for the hDOT1L(1-416) structure have been deposited in the RCSB database under ID code 1NW3.