Structure of Human SMYD2 Protein Reveals the Basis of p53 Tumor Suppressor Methylation

Li Wang, Ling Li, Hailong Zhang, Xiao Luo, Jingquan Dai, Shaolian Zhou, Justin Gu, Jidong Zhu, Peter Atadja, Chris Lu, En Li, and Kehao Zhao

From the Novartis Institutes for BioMedical Research, Building 8, Lane 898, Halei Road, Pudong, Shanghai 201203, China

Background: SMYD2 is a lysine methyltransferase that mediates functions of target protein by specific site methylation.

Results: SMYD2 prefers to monomethylate Lys-370 of p53, and the specificity is explained by high resolution structure of the enzyme bound to p53.

Conclusion: CTD domain and a unique EDEE motif play critical roles in p53 Lys-370 methylation by SMYD2.

Significance: The findings provide molecular insights into the mechanism of p53 recognition by SMYD2.

SMYD2 belongs to a subfamily of histone lysine methyltransferase and was recently identified to methylate tumor suppressor p53 and Rb. Here we report that SMYD2 prefers to methylate p53 Lys-370 over histone substrates in vitro. Consistently, the level of endogenous p53 Lys-370 monomethylation is significantly elevated when SMYD2 is overexpressed in vivo. We have solved the high resolution crystal structures of the full-length SMYD2 protein in binary complex with its cofactor S-adenosylmethionine and in ternary complex with cofactor product S-adenosylhomocysteine and p53 substrate peptide (residues 368–375), respectively. p53 peptide binds to a deep pocket of the interface between catalytic SET(1–282) and C-terminal domain (CTD) with an unprecedented U-shaped conformation. Subtle conformational change exists around the p53 binding site between the binary and ternary structures, in particular the tetratricopeptide repeat motif of the CTD. In addition, a unique EDEE motif between the loop of anti-parallel β7 and β8 sheets of the SET core not only interacts with p53 substrate but also forms a hydrogen bond network with residues from CTD. These observations suggest that the tetratricopeptide repeat and EDEE motif may play an important role in determining p53 substrate binding specificity. This is further verified by the findings that deletion of the CTD domain drastically reduces the methylation activity of SMYD2 to p53 protein. Meanwhile, mutation of EDEE residues impairs both the binding and the enzymatic activity of SMYD2 to p53 Lys-370. These data together reveal the molecular basis of SMYD2 in specifically recognizing and regulating functions of p53 tumor suppressor through Lys-370 monomethylation.

SMYD2 belongs to a five member SET domain-dependent methyltransferase subfamily (SMYD1 to -5) with a characteristic MYND (myeloid, Nervy, and DEAF-1) zinc ion binding motif that splits the SET domain (1). SMYD2 was identified as a histone H3 lysine 36 (H3K36)2 methyltransferase (1), and recent study has shown that it also methylates histone H3K4 upon interacting with chaperone protein Hsp90 (2). Despite being highly expressed in heart and brain, however, SMYD2 has not been found to associate with heart development in mice (3). In contrast, its highly related paralogue SMYD1 has been demonstrated to be critical in regulating heart and skeletal muscle development through the histone H3K4 methylation activity (4–6). On the other hand, although SMYD3 also methylates histone H3K4, the biological effects are largely distinct from that of SMYD1. This is supported by evidence that SMYD3 is frequently overexpressed in cancer cells and may play an essential role in regulating development and progression of tumor cells (7). Less is known about the histone methyltransferase activities and functions of both SMYD4 and SMYD5. However, the former was recently identified as a tumor suppressor gene to regulate the expression of platelet-derived growth factor receptor α polypeptide in breast carcinogenesis (8, 9). Together, the published data indicate that this subfamily of enzymes plays diverse roles through methyltransferase activities in regulating cellular functions.

In addition to histone methylation, SMYD family enzymes also methylate lysine residues on non-histone protein targets. Indeed, SMYD3 was reported to regulate the angiogenesis function through methylating lysine 831 of vascular endothelial growth factor receptor 1 (VEGFR1) (10). SMYD2 methylates the lysine 370 of tumor suppressor p53 and results in the repression of p53-mediated transcription activation (11). Importantly, down-regulation of SMYD2 in cells by short interfering RNA promotes p53 dependent apoptosis, underlining the potential role of the protein in tumorigenesis (11). This is further supported by the findings that patients with high level of SMYD2 in esophageal squamous cell carcinoma have a much lower survival rate than those with normal SMYD2 level (12). Intriguingly, SMYD2 was recently found to methylate another...
tumor suppressor protein, Rb, at its Lys-860 (13). The biological function of Rb Lys-860 methylation by SMYD2 remains unclear; however, it was suggested that Lys-860 methylation may trigger the interaction with L3MBTL1, a chromatin compaction factor containing a triple malignant brain tumor, for subsequent transcription repression of Rb target genes (13). These findings illustrate the importance of SMYD2 in regulating cell proliferation and tumor progression via its non-histone protein methylation activities.

The tumor suppressor p53 regulates cellular responses to various forms of genotoxic stresses. A complex network of post-translational modifications, including lysine methylation at its C terminus region, is of paramount importance in regulating transcription activity of the protein (14). Besides Lys-370 methylation of p53 by SMYD2, three other lysine methyltransferases, SET7, SET8, and G9a/GLP, are known to specifically methylate p53 in addition to their histone or other protein targets (14, 15). For SET8, it acts exclusively for monomethylation of lysine 382 of p53 to suppress transactivation of target genes, including p21 (16, 17), whereas G9a/GLP has been demonstrated to dimethylate Lys-373 of p53 and negatively regulate p53-mediated apoptosis (18). Methylation at Lys-372 by SET7 results in increased nucleosome localization and stability of p53 that ultimately leads to transcription activation of target genes and triggers p53-dependent apoptosis (11, 19). Interestingly, the repressive Lys-370 methylation by SMYD2 is inhibited by SET7-mediated Lys-372 methylation, suggesting a cross-talk between the two sites that may implicate an important mechanism in regulating p53 functions (11).

The crystal structures of SMYD1 and SMYD3 have been reported recently. Both structures revealed that SMYD proteins share a conserved catalytic core SET domain, with the unique MYND and C-terminal domain arranged around the core to process enzymatic activity (20–22). However, the biochemical and structural mechanisms underlying the basis of the SMYD2 in recognizing histone and non-histone substrates are largely unknown. To this end, here we report that SMYD2 prefers to methylate p53 Lys-370 over histone substrates. We then determined the high resolution crystal structures of the full-length SMYD2 protein in binary complex with its cofactor AdoMet and in ternary complex with cofactor product S-adenosylhomocysteine (AdoHcy) and p53 substrate peptide, respectively. We further investigated regions in SMYD2 that are important in determining p53 substrate specificity by mutagenesis studies. Our results demonstrated that a C-terminal three-helix-turn-helix motif containing a tetratricopeptide repeat (TPR) and a unique EDEE motif within the catalytic core of SET are important for p53 methylation. Together, our structures provide insights into the molecular basis of p53 recognition by SMYD2.

**EXPERIMENTAL PROCEDURES**

**Protein Expression and Purification**—To prepare protein for crystallization and ITC assay, the full-length gene of SMYD2 was cloned into the pFastBac vector with an N-terminal His<sub>6</sub> tag. The protein was expressed in the sF9 cell line using the baculovirus expression system following a standard protocol. Cells were suspended in buffer containing 50 mM Tris-HCl (pH 8.0), 200 mM NaCl and lysed by sonication. The soluble SMYD2 protein in the centrifuged supernatant was purified by a Ni<sup>2+</sup>-NTA (Qiagen) affinity chromatographic column. Tobacco etch virus protease was added to the SMYD2 solution and incubated overnight at 4 °C to remove the His<sub>6</sub> tag. The tag-removed SMYD2 was then passed through the Ni<sup>2+</sup>-NTA column again and further purified by a Superdex-200 gel filtration column. The purified protein was concentrated to 15 mg/ml and stored in 20 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 5 mM β-mercaptoethanol.

To prepare protein for biochemical assay, the full-length genes of SMYD2, SMYD1, and SMYD3 fused with an N-terminal FLAG tag were separately subcloned into the pcDNA3.1 vector. The protein was transiently expressed in 293T cell line and purified by one-step affinity chromatography using anti-FLAG M2 affinity gel (Sigma). The target protein was eluted with 0.1 mg/ml FLAG peptide (Sigma) and dialyzed against buffer (50 mM Tris-HCl, pH 8.0, and 100 mM NaCl). The protein was finally concentrated to 2 mg/ml in 50 mM Tris–HCl, pH 8.0, 100 mM NaCl, and 5 mM β-mercaptoethanol (supplementary Fig. S1).

**Crystallization, Data Collection, and Structure Determination**—SMYD2 and AdoMet co-crystals were obtained by the hanging drop vapor diffusion method. Briefly, 10 mg/ml SMYD2 was mixed with a 3-fold molar excess of AdoMet on ice for at least 1 h, and the complex was then mixed in equal volume with the reservoir solution containing 0.8 m lithium chloride, 0.1 m Tris-HCl (pH 8.5), and 32% PEG 4000 at 20 °C. For the SMYD2 ternary complex, SMYD2 was mixed with a 5-fold molar excess of AdoHcy and an 11-mer p53 peptide (HSSHLK-SKKGQ) on ice for at least 1 h, and the complex was then mixed in an equal volume with the reservoir solution containing 0.1 m Heps (pH 7.5) and 25% PEG 3350 at 20 °C. Crystals were obtained in a few days and were flash frozen in liquid nitrogen by the same reservoir solution containing additional 20% glycerol. Diffraction data were collected at the Shanghai Synchrotron Radiation Facilities, and data were processed using HKL2000 (23).

The binary complex crystal belongs to the P<sub>2</sub>₁<sub>1</sub>₂<sub>1</sub>₁ space group with unit cell dimensions of a = 52.48 Å, b = 67.48 Å, and c = 141.97 Å, respectively. The homologous protein structure of the SMYD3-AdoMet complex (Protein Data Bank code 3MEK), which shows 32% sequence identity, was used to create a search model for molecular replacement. The initial trial of molecular replacement using the full-length model failed to find SMYD2 solution. The N-domain (aa 6–269) of SMYD3 successfully resulted in the correct solution for SMYD2. The structure was further refined by CNS Solve (24) using annealing, energy minimization, and B individual refinement. The rest of the missed region was then built into the model using ARP/WARP (25) followed by iterative manual building in COOT (26) and refinement in Refmac5 (27). The ternary complex crystal was in the same space group with unit cell dimensions of a = 53.40 Å, b = 71.78.48 Å, and c = 121.29 Å, respectively. The structure was solved by molecular replacement with Molrep. Subsequent refinement was carried out using Refmac5 and
manual model building in COOT. The statistics of the structure refinement and the quality of the final model are summarized in Table 1. All of the figures were made with PyMOL.

**Peptides and Antibodies**—The following peptides were chemically synthesized (GL Biochem) for histone lysine methyltransferase activity and pull-down assay: H3K4(1–20), ART-K(me0, me1, me2)QTKARTSGGKPKRQPL; H3K9(1–20), QTARAMELA(me0, me1, me2)QTKARTSGGKPKRQPL; H3K14(27–36), QTLAKRAMELA(me0, me1, me2)QTKARTSGGKPKRQPL; H3K27(19–36), QLAQAKKAR(me0, me1, me2)QTKARTSGGKPKRQPL; H3K36(26–46), RKSAPATGGVK(me0, me1, me2)KPHRYRP-GTV; H4K20(13–29), GGAHKRQ(me0, me1, me2)VLKDNI-QGL; p53(361–380), GSRAHSSHLKKKGQSTSRH; p53(361–380)–biotin, GSRAHSSHLKKKGQSTSRH–K(Biotin)–OH; p53(363–393), Biotin–RAHSSHLKKKGQSTSRH–KKLMTFTEGPDSD; H3(1–20)–biotin, ARTKQTKARTSGGKPKRQL–K(Biotin)–OH; H3(28–48)–biotin, SAPATGGVKKPHRYRPV–A–K(Biotin)–OH; H4(12–32)–biotin, KGGAKHRKVLNQNGN–K(Biotin)–OH.

Anti-p53 Lys-370 monomethylation (p53 K370me1) rabbit polyclonal antibody was generated using peptide GSRAHSSHLKKKGQSTSRH as immunogen. Briefly, rabbit was immunized and boosted by a keyhole limpet hemocyanin-conjugated peptide. The antibody was purified by an immunoaffinity column. The antibody selectively binds to p53 K370me1 in subsequent dot blot and ELISAs (supplemental Fig. S5). Anti-SMYD2 (Abcam) and p53 (CST, catalog no. 9282) antibodies were purchased from their respective commercial sources.

In **Vitro Histone Lysine Methyltransferase Activity Assay**—Histone H3 was expressed, refolded, and purified as described (28). The full-length human p53 protein was expressed in *Escherichia coli* with a C-terminal His<sub>6</sub> tag and purified to near homogeneity by Ni<sup>2+</sup>–NTA affinity and Superdex-200 gel filtration columns. For the methyltransferase activity assay, 200 mM SMYD protein was incubated with 25 μM peptide substrate together with 25 μM AdoMet (Sigma) for 3 h or incubated with 4 μM protein substrate and 5 μM AdoMet (Sigma) for 2 h at 25 °C in the reaction buffer (20 mM Tris-HCl (pH 8.0), 10 mM MgCl<sub>2</sub>, 0.01% Tween 20, and 1 mM DTT). The reaction was then quenched by 5% TFA. The methyltransferase activity was measured by the concentration of cofactor product AdoHcy using LC-MS methods with the d4-AdoHcy as the internal standard (29). All of the experiments were performed on an AB Sciex (Foster City, CA) API 4000 triple quadrupole mass spectrometer configured with a Shimadzu (Nagakyo-ku, Kyoto, Japan) LC-20AD liquid chromatograph.

**Isothermal Titration Calorimetry**—ITC was performed at 25 °C using a MicroCal Auto-ITC200 instrument (GE Healthcare). The protein was dialyzed overnight against ITC buffer (50 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 1 mM DTT) before the titration. The powder of peptide was dissolved to a final concentration of 1 mM in the same buffer. During titration, 40-μl aliquots of p53 peptide (aa 361–380) were injected into a solution of 100 μM SMYD2 protein. ITC data were corrected for the heat of dilution by subtracting the mixing enthalpies from titrant solution injections into protein-free ITC buffer. ITC data were analyzed using Origin 7.0 software (Origin Lab Corp.), and ΔH was calculated using an one-site binding model.

**Peptide Pull-down Assay**—The assay was carried out as described previously (30) in the binding buffer of 50 mM Tris-HCl (pH 8.0), 100 mM NaCl, and 0.01% Tween 20. 1 μg of C-terminal biotinylated peptide substrate was incubated with 1 μg of protein for 2 h at 4 °C. The sample was then mixed with streptavidin beads (Pierce) for 1 h with gentle shaking. After washing, the bound complex was eluted by SDS-PAGE loading buffer and subjected to Coomassie Blue staining or immunoblot analysis.

**Cell Culture and p53 Methylation in Vivo**—The U2OS cell line was cultured at 37 °C in DMEM supplemented with 10% FBS. The wild type SMYD2 in pCDNA3.1 was transiently transfected into cells using Lipofectamine 2000. The empty vector was used as a control. 24 h after transfection, the cells were harvested and subjected for immunoprecipitation with anti-p53 antibody using a standard protocol. Following extensive washes, the immunoprecipitated material was eluted with SDS-loading buffer for immunoblot analysis. Methylated p53 was detected by Western blot using an anti-p53 K370me1 antibody. For co-transfection experiments, 293T cells were plated at a density of 0.5 × 10<sup>6</sup> cells/well to a 6-well plate 24 h prior to transfection. 4 μg each of SMYD2 and p53 plasmid in pCDNA3.1 and 20 μl of Lipofectamine 2000 reagent were used for each well. Cells were harvested 48 h after transfection. p53 Lys-370 methylation was then analyzed by the same anti-p53 K370me1 antibody.

**RESULTS**

SMYD2 Prefers to Methylate p53 Lys-370 in Vitro, and Over-expression of SMYD2 Promotes p53 Lys-370 Methylation in Vivo—SMYD2 was originally identified as a H3K36 dimethylation enzyme (1) and subsequently has been shown to also possess H3K4 methylation activity when interacting with HSP90α (2). SMYD2 protein has been reported to methylate non-histone proteins p53 (11) and Rb (13). The accumulating evidence that SMYD2 methylates multiple targets raises the question of substrate specificity. To gain insight into substrate recognition by SMYD2, we tested the methyltransferase activity of SMYD2 to a set of peptide substrates. To do this, FLAG-SMYD2 was transiently expressed in 293T cells and purified to near homogeneity by a one-step affinity chromatographic column (supplemental Fig. S1). The protein was then mixed with the cofactor AdoMet and peptide substrate and reacted for 3 h at 25 °C. The activity of the SMYD2 was calculated based on measuring the amount of co-factor product AdoHcy by LC-MS as reported previously (29). SMYD2 shows very weak activity to H3K36 (aa 26–46) but possesses observable activity to H3K4, H3K36me1, and H3K36me2 peptides (Fig. 1, upper left). Interestingly, SMYD2 also weakly methylates a few other histone peptides, including H3K27 (aa 19–36) and H4K20 (aa 13–29) with activities comparable with that of histone H3K4. Interestingly, similar to that of SMYD2, we also found that both SMYD1 and SMYD3 only show very weak or negligible activities to histone substrates (5, 7) (supplemental Fig. S2). Surprisingly, SMYD2 shows more than 10-fold higher activity to p53 peptide (residues 361–380) than all histone peptide substrates. In agreement with the published results (11), by using mass spectrometry methylation mapping analysis, we further con-
FIGURE 1. SMYD2 methylates and binds p53 K370. Upper, SMYD2 specifically methylates p53 Lys-370 peptide in vitro. Activities of SMYD2 to un-, mono-, and dimethylated histone and p53 K370 peptides (left), recombinant histone H3, and mononucleosome and p53 proteins (right). The asterisk indicates the same length (aa 1–20) of unmethylated H3K9 peptide because H3K4 was used in the experiments. The experiments were done in triplicate; average values ± S.D. are shown. Middle, SMYD2 interacts with p53 Lys-370 peptide substrate in vitro. Left, results of a biotinylated peptide pull-down assay using the indicated biotinylated peptides with the respective purified FLAG-SMYD2 and SMYD3. The biotin peptide–captured protein analyzed by Western blot using anti-FLAG tag antibody (Sigma) is shown in the top left panel. The total input of SMYD protein is shown in a Coomassie Blue staining SDS-PAGE in the bottom left panel. Right, isothermal titration calorimetry analysis result of the binding affinity of p53 Lys-370 peptide substrate to SMYD2. Top, calorimetric response as successive injections of the peptide are added to the reaction cell. Bottom, binding isotherm of the calorimetric titration shown in the top. A single binding model was applied for curve fitting. Lower, SMYD2 methylates p53 Lys-370 in vivo. Left, Western blot results for endogenous p53 Lys-370 monomethylation when SMYD2 was overexpressed in U2OS cells. The empty vector was used as a control. Right, Western blot analysis of p53 Lys-370 monomethylation from 293T cells when both p53 and SMYD2 were overexpressed. p53 K370 methylation was analyzed by a potent and selective anti-p53 K370me1 rabbit polyclonal antibody (supplemental Fig. S5). GAPDH was used as a control to normalize the total protein level in the samples. Error bars, S.D.
firmed that SMYD2 specifically monomethylates Lys-370 on a longer peptide (aa 363–393) (supplemental Fig. S3).

We next asked whether a similar conclusion could be drawn when protein substrates are applied to the study. We produced recombinant full-length human p53 protein, the histone H3, and the mononucleosome core particle to near homogeneity. We next compared the enzymatic activity of SMYD2 with these protein substrates (Fig. 1, upper right). Consistently, the activity of SMYD2 toward p53 protein was about 3- and 6-fold higher than the histone H3 and nucleosome substrates, respectively.

To understand whether the difference of SMYD2 activity is correlated with substrate binding, we measured the binding affinity of SMYD2 to substrates in peptide pull-down assays. Consistent with the results from a biochemical assay, the bio-
tin-p53 peptide immobilized on streptavidin beads captured a significantly higher amount of SMYD2 compared with a much lower level by H3K4 and no detectable signal of the enzyme by H3K36 and H4K20 (Fig. 1, middle left). Meanwhile, SMYD3 showed no binding activity to any peptide substrates, including H3K36. ITC analysis further confirmed the binding of the p53 peptide (aa 361–380) to SMYD2 with a dissociation constant ($K_d$) of about 20 μM, whereas the binding of histone H3K4(1–20) peptide to SMYD2 was about 35-fold weaker (Fig. 1, middle right and supplemental Fig. S4). Together, our results demonstrate that SMYD2 prefers to bind and monomethylate p53 Lys-370 in vitro.

Previous studies by Huang et al. (11) reported that SMYD2 methylates p53 Lys-370 when both proteins are co-expressed in cells. However, it is not clear if the exogenous or endogenous p53 is mainly methylated by SMYD2. To address this question and further verify the physiological relevance of SMYD2 and p53 Lys-370 methylation, we conducted an in vivo p53 Lys-370 methylation activity test. U2OS cells were transfected with wild type FLAG-SMYD2, whereas the empty vector was used as control. 24 h after transfection, the cells were lysed, and the endogenous p53 protein was then immunoprecipitated by a p53-spe-
cific antibody. The level of p53 Lys-370 monomethylation was then detected by a highly selective anti-p53 K370me1 antibody (supplemental Fig. S5). Compared with the control cells, more than 4-fold endogenous p53 Lys-370 methylation was detected in SMYD2-transfected cells (Fig. 1, lower left). In addition, p53 Lys-370 methylation was significantly increased when both p53 and SYMD2 were overexpressed in 293T cells (Fig. 1, lower right), similar to the published report except where H1299 cells were applied for the experiments (11). Taken together, these results demonstrate the in vivo monomethylation of endoge-

### TABLE 1

|                | SMYD2-AdoMet | SMYD2-AdoHcy[p53] |
|----------------|-------------|------------------|
| Space group    | P2$_1$/a 2, 2, 2 | P2$_1$/a 2, 2, 2 |
| Resolution (Å) | 2.1         | 2.3              |
| Cell dimensions |             |                  |
| a, b, c (Å)    | 52.48, 67.48, 141.97 | 53.40, 71.78, 121.29 |
| R$_{merge}$ (°) | 8.9 (44.3) | 8.2 (35.9) |
| completeness (%)| 11.1 (3.9) | 8.5 (2.2) |
| Redundancy     | 99.9 (99.8) | 95.9 (91.2) |
| Unique reflections | 33,004 | 20,928 |
| R$_{free}$ (%) | 19.9 (22.8) | 24.1 (26.7) |
| R$_{plot}$ (%) | 22.4 (27.9) | 26.3 (31.4) |
| t.r.m.s deviations |
| Bond lengths (Å) | 0.01 | 0.01 |
| Bond angles (degrees) | 1.179 | 1.00 |

and $F_o - F_c$ electron density maps (Fig. 2a, left), except for His-365, Ser-366, and Ser-367 in p53, which are omitted from the final model because of their disorder. The overall structures of SMYD2 in two complexes are highly similar, with a root mean square (r.m.s.) deviation of about 0.4 Å for all Cα atoms. However, conformational changes around the p53 substrate binding site have been clearly observed and will be discussed in detail in this report.

The overall structure of SMYD2 is composed of an N-ter-

### Overall Structure of SMYD2—To gain insights into the struc-
tural basis and the molecular mechanism of p53 recognition by SMYD2, we determined the structures of a binary complex of the full-length SMYD2 with the cofactor AdoMet at 2.1 Å and a ternary complex containing the cofactor product AdoHcy and p53 11-mer peptide substrate (aa 365–375) at 2.3 Å resolution with excellent crystallographic statistics (Table 1), respectively. The entire AdoMet molecule, including its methyl donor, is unambiguously defined in the binary structure. The cofactor product AdoHcy and most of the p53 peptide residues in the ternary complex are also well defined in both unbiased $2F_o - F_c$ and $F_o - F_c$ electron density maps (Fig. 2a, left), except for His-365, Ser-366, and Ser-367 in p53, which are omitted from the final model because of their disorder. The overall structures of SMYD2 in two complexes are highly similar, with a root mean square (r.m.s.) deviation of about 0.4 Å for all Cα atoms. However, conformational changes around the p53 substrate binding site have been clearly observed and will be discussed in detail in this report.

The overall structure of SMYD2 is composed of an N-ter-

### Overall Structure of SMYD2—To gain insights into the struc-
tural basis and the molecular mechanism of p53 recognition by SMYD2, we determined the structures of a binary complex of the full-length SMYD2 with the cofactor AdoMet at 2.1 Å and a ternary complex containing the cofactor product AdoHcy and p53 11-mer peptide substrate (aa 365–375) at 2.3 Å resolution with excellent crystallographic statistics (Table 1), respectively. The entire AdoMet molecule, including its methyl donor, is unambiguously defined in the binary structure. The cofactor product AdoHcy and most of the p53 peptide residues in the ternary complex are also well defined in both unbiased $2F_o - F_c$ and $F_o - F_c$ electron density maps (Fig. 2a, left), except for His-365, Ser-366, and Ser-367 in p53, which are omitted from the final model because of their disorder. The overall structures of SMYD2 in two complexes are highly similar, with a root mean square (r.m.s.) deviation of about 0.4 Å for all Cα atoms. However, conformational changes around the p53 substrate binding site have been clearly observed and will be discussed in detail in this report.

The overall structure of SMYD2 is composed of an N-ter-

### Overall Structure of SMYD2—To gain insights into the struc-
tural basis and the molecular mechanism of p53 recognition by SMYD2, we determined the structures of a binary complex of the full-length SMYD2 with the cofactor AdoMet at 2.1 Å and a ternary complex containing the cofactor product AdoHcy and p53 11-mer peptide substrate (aa 365–375) at 2.3 Å resolution with excellent crystallographic statistics (Table 1), respectively. The entire AdoMet molecule, including its methyl donor, is unambiguously defined in the binary structure. The cofactor product AdoHcy and most of the p53 peptide residues in the ternary complex are also well defined in both unbiased $2F_o - F_c$ and $F_o - F_c$ electron density maps (Fig. 2a, left), except for His-365, Ser-366, and Ser-367 in p53, which are omitted from the final model because of their disorder. The overall structures of SMYD2 in two complexes are highly similar, with a root mean square (r.m.s.) deviation of about 0.4 Å for all Cα atoms. However, conformational changes around the p53 substrate binding site have been clearly observed and will be discussed in detail in this report.

The overall structure of SMYD2 is composed of an N-ter-

### Overall Structure of SMYD2—To gain insights into the struc-
tural basis and the molecular mechanism of p53 recognition by SMYD2, we determined the structures of a binary complex of the full-length SMYD2 with the cofactor AdoMet at 2.1 Å and a ternary complex containing the cofactor product AdoHcy and p53 11-mer peptide substrate (aa 365–375) at 2.3 Å resolution with excellent crystallographic statistics (Table 1), respectively. The entire AdoMet molecule, including its methyl donor, is unambiguously defined in the binary structure. The cofactor product AdoHcy and most of the p53 peptide residues in the ternary complex are also well defined in both unbiased $2F_o - F_c$ and $F_o - F_c$ electron density maps (Fig. 2a, left), except for His-365, Ser-366, and Ser-367 in p53, which are omitted from the final model because of their disorder. The overall structures of SMYD2 in two complexes are highly similar, with a root mean square (r.m.s.) deviation of about 0.4 Å for all Cα atoms. However, conformational changes around the p53 substrate binding site have been clearly observed and will be discussed in detail in this report.

The overall structure of SMYD2 is composed of an N-ter-

### Overall Structure of SMYD2—To gain insights into the struc-
tural basis and the molecular mechanism of p53 recognition by SMYD2, we determined the structures of a binary complex of the full-length SMYD2 with the cofactor AdoMet at 2.1 Å and a ternary complex containing the cofactor product AdoHcy and p53 11-mer peptide substrate (aa 365–375) at 2.3 Å resolution with excellent crystallographic statistics (Table 1), respectively. The entire AdoMet molecule, including its methyl donor, is unambiguously defined in the binary structure. The cofactor product AdoHcy and most of the p53 peptide residues in the ternary complex are also well defined in both unbiased $2F_o - F_c$ and $F_o - F_c$ electron density maps (Fig. 2a, left), except for His-365, Ser-366, and Ser-367 in p53, which are omitted from the final model because of their disorder. The overall structures of SMYD2 in two complexes are highly similar, with a root mean square (r.m.s.) deviation of about 0.4 Å for all Cα atoms. However, conformational changes around the p53 substrate binding site have been clearly observed and will be discussed in detail in this report.
of MYND, the loop (aa 96–103) connecting the long bent helix 
\( \alpha \)A of MYND and \( \alpha \)B of SET-I, the loop (aa 239–245) connecting the post-SET and SET core, and the loop (aa 187–194) between \( \beta \)7 and \( \beta \)8. Interestingly, although the CTD shares the similar structure, the relative orientation of this domain against the SET domain among SMYD proteins is drastically different,
as demonstrated by the large r.m.s. deviation of 5.51 and 3.17 Å for all Ca atoms of SMYD2 (aa 298–432) against the corresponding segment of SMYD1 and SMYD3, respectively (Fig. 2c). Importantly, compared with SMYD2, the CTD of SMYD1 rotated about 22° outward against SET, creating a more open and spacious interface between its SET and the CTD, which is not arranged for proper substrate binding (20). On the other hand, the CTD in the SMYD3 crystal structure rotates about 13° toward its SET domain compared with SMYD2. These two domains in SMYD3 are closely connected with each other through the interactions between the loop of β7-β8 and concave surface of αH-αL, forming a closed interface between SET and CTD. Such an arrangement may represent an autoinhibitory configuration of the structure (21, 22).

Cofactor Site of SMYD2—Similar to other SET domain structures, both AdoMet and AdoHcy adopt a condensed U-type conformation and are well superimposed with each other (Fig. 2d). AdoMet and AdoHcy share similar interactions with SMYD2. Therefore, we will focus on discussing the cofactor AdoMet binding to SMYD2. AdoMet binds to the SMYD2 in a pocket formed by the N-terminal β1 and β2 sheets, the 3α- containing loop connecting αD and αE helices in SET-I, and the post-SET domain. The interactions between AdoMet and SMYD proteins in the current structure are very similar to the structures reported, which mainly involve the consensus sequences of NHXC and YXF from the SET core (20–22). The adenine moiety is located between the aliphatic side chain of Lys-17 and the phenyl ring of Phe-260 and forms conserved hydrogen bonds between its purine N6 atom and the carbonyl atom of Ser-207. The O2’ and O3’ atoms of the ribose ring form hydrogen bonds with the NE2 atom of the His-137, and its O4’ atom also forms a hydrogen bond with the amide ND2 of the Asn-206. These interactions, together with the additional van der Waals contacts between the ribose ring and side chains of Tyr-245 and Tyr-258, are important for maintaining the stable conformation of the ribose ring. The α-amide group of the AdoMet also forms hydrogen bonds with the respective carbonyl atom of Lys-17 and Arg-19. In combination with the salt bridge between the carboxylate group and the side chain of Arg-19, these interactions together position the methyl donor to meet the substrate lysine for methyl transfer reaction.

Although the cofactor AdoMet or its analogs share conserved interactions with SMYD proteins, it is also noteworthy to mention the subtle differences in the binding pocket. In SMYD1-sinefungine, the NE2 atom of the Gln-133 (corresponding to Glu-133 in SMYD2 and Glu-130 in SMYD3; supplemental Fig. S6) forms a hydrogen bond with the O2’ of the ribose ring, whereas in both SMYD2 and SMYD3, the side chain of the residue Glu-133 rotates away from the ribose ring and does not form such interactions. In addition, the surface potential of the AdoMet pocket in SMYD3 is more negatively charged when compared with that of SMYD2 (data not shown). The conformation variation among SMYD family proteins may implicate the diversity of the AdoMet binding pocket and the drugability of the site (31).

Substrate Binding Site of SMYD2—p53 peptide binds to a wide open and deep pocket of SMYD2 (Fig. 3a). The deep center of the pocket is largely defined by the αE of the post-SET and loop (aa 239–245) of the SET core on one side and by the anti-parallel β7 and β8 sheets of the SET core on the other side. Outside of the core, the αC (aa 143–160) and the perpendicular C terminus of αD (aa 178–182) of SET-I, together with the region connecting the αB of SET-I and long bent αA of MYND, sit on top of the pocket. The bottom of the pocket is mainly composed of the TPR1 (αH and αI) and TPR2 (αA and αK) of the CTD, in which the TPR1 rotates about 45° toward the post-SET with respect to the plate of the TPR2 and results in its helix-turn-helix shielding the visibility of C terminus of the p53 substrate.

Upon binding of the p53 peptide, conformational change exists around the substrate binding pocket when compared with the cofactor-bound binary structure in SMYD2. The most obvious difference is that the loop connecting the two α helices of TPR1 together with the C terminus of the first α-helix (αH) shifts about 2 Å toward the p53 binding pocket, resulting in a more compact substrate binding interface compared with that in the SMYD2-AdoMet complex structure. This may be attributable to the interactions of the p53 residues with the TPR domain. Accordingly, residues in this helix also show a conformational difference, as illustrated in Fig. 3b. The side chain of Arg-253 forming a positively charged patch with Arg-306 in the binary structure moves away in the ternary complex structure. The side chains of Glu-303, Glu-304, Arg-307, His-310, and Tyr-311 also show a difference in both structures in accommodating the TPR1 helix movement.

The lysine binding tunnel is composed of residues Cys-181, Asn-182, Gly-183, Phe-184, Tyr-240, and Tyr-258 that are strictly conserved in the SMYD family with the exception of Gly-183, which is substituted by serine in SMYD3. Those residues interact with the aliphatic side chain of Lys-370 mostly through van der Waals contacts and adopt a conserved conformation similar to the SMYD1 and SMYD3 structures, except for the side chain of SMYD1 Phe-182 (equivalent to Phe-184 in SMYD2; supplemental Fig S6), which rotates away from the substrate lysine and leads to a more spacious tunnel than those in SMYD2 and SMYD3 (20). The difference of the lysine tunnel of SMYD2 is even more obvious compared with that of SET7, which contains a much narrower tunnel mostly encompassed by aromatic residues (32). Importantly, Tyr-305, crucial in switching the SET7 from mono- to dimethylase, is replaced by Val-215 in SMYD2. Another residue, Tyr-245 of SET7, also important for product specificity is replaced by Cys-36, which is far from the tunnel and active site in SMYD2 (32, 33). This suggests that these residues are unlikely to be playing a role in determining product specificity of SMYD2. Overall, the lysine binding tunnel in SMYD2 shares a relatively conserved but subtly different configuration with SMYD1 and SMYD3. Such a difference is much more pronounced when compared with that of SET7.

SMYD2 exhibits a large configuration difference at the substrate pocket when compared with other substrate-bound SET domain structures (Fig. 3c). For instance, the loop (aa 187–194) and β7 and β8 sheets in SMYD2 are correspondingly represented by a more stretched loop but a much shorter anti-parallel β-strand in SET7, in which those two regions together with the SET-1 form a rather shallow groove to sandwich the p53
peptide substrate (19). Additionally, a long loop connecting the a-helix of the post-SET and SET core b-strand, which is important for p53 interactions in SET7, is substituted by a much shorter loop of residues 239–245 in SMYD2. On the other side, the core of the substrate binding pocket in the SMYD2 is highly conserved with SMYD1 and SMYD3. However, substantial conformational changes within the core are still observed, particular for the loop (aa 239–245) connecting the post-SET and SET core and the loop (aa 187–194) between the b7 and b8 sheets. Together, the p53-bound SMYD2 complex structure reveals a unique p53 substrate binding pocket among the SMYD family and other SET domain structures. This may implicate the specificity of p53 substrate recognition by SMYD2.

Recognition of p53 by SMYD2—The p53 peptide binds to SMYD2 with an unprecedented U-shaped conformation. The N-terminal residues His-368 and Leu-369 of the peptide chain pack perpendicular to the top of the post-SET to form one arc of the shape. The peptide sharply turns about 90° at the lysine binding tunnel and extends alongside the deep center pocket between the loop (aa 239–245) and the anti-parallel b7 and b8 sheets of the SET core to form the central bottom of the U-shape. It then makes another 90° turn at its Lys-372 and exits from the deep pocket toward the TPR1 of CTD to complete the U-shape (Fig. 3a). The residues (aa 370–372) in the bottom of the U-shaped p53 peptide make extensive contacts with residues mainly from the SET core through main-chain-to-main-chain hydrogen bonds to anchor the peptide to SMYD2, a common feature in SET domain structures (34–36). In contrast, the N terminus His-368 and Leu-369 interact with residues from post-SET and SET-I with mixed main-chain-to-side-chain hydrogen bonds and van der Waals contacts, whereas its C terminus (aa 373–375) interacts with the TPR domains mainly through side-chain hydrogen bonds (Fig. 3d and Fig. 3e). Together, these observations highlight the significant accountability of residues from SET core, SET-I, post-SET, and CTD in determining SMYD2 substrate specificity.

The main-chain NH of His-368 forms a hydrogen bond with the side-chain atom OG of the Ser-257 in post-SET, and its main-chain carbonyl oxygen forms hydrogen bonds with the hydroxyl group of Tyr-258 and the side-chain amine of the Arg-253, respectively. In addition, the side chain of p53 Leu-369 extends to a shallow hydrophobic groove on the top of the lysine funnel and forms van der Waals contacts with residues Gly-183, Phe-184, and Ser-196 of the SET core. These residues together with Thr-105, Leu-108, Ile-112, and Val-179 of SET-I and Ile-198 form the groove and share sequence conservation with both SMYD1 and SMYD3 (supplemental Fig. 5e). This indicates that the hydrophobic groove in SMYD1 and SMYD3 may also be responsible for interacting with the residue prior to lysine substrate.

The Lys-370 residue of p53 substrate is connected to the catalytic site by the lysine binding tunnel from the opposite side of the cofactor AdoMet binding pocket (Fig. 2d). The e-amine group of the Lys-370 is positioned toward the methyl group of AdoMet with a distance of 2.6 Å between the two groups. Outside of the catalytic center, its main-chain atom NH forms a hydrogen bond with the carbonyl oxygen of Gly-183 and contacts with the carbonyl group of Leu-180 through a well ordered water molecule. Its carbonyl oxygen forms hydrogen bonds with main-chain amine and side-chain OG1 atoms of Thr-185, respectively.

The NH of Ser-371 forms a hydrogen bond with the main-chain carbonyl oxygen of Tyr-240. Its carbonyl oxygen also interacts with the carbonyl oxygen of Tyr-240 and NH of Asn-242. The OG1 atom of the Ser-371 forms another hydrogen bond with the carbonyl oxygen of Lys-373 within the substrate, which may stabilize the sharp bent conformation of the peptide. The NH and carbonyl oxygen of Lys-372 form hydrogen bonds with the side-chain atoms of Glu-187 and Asn-380 of SMYD2, respectively. The aliphatic side chain is deeply inserted into the pocket and forms hydrophobic interactions with residues Leu-191 and Ile-214 of the SET core. Its e-amine group is further stabilized by hydrogen bonding to the carbonyl of Val-215, and water molecules mediated the hydrogen bond network with the side chain of Glu-187, carbonyl oxygens of Thr-185 and Ser-239, and OH of Tyr-217. It was reported that methylation of p53 Lys-372 by SET7 inhibited Lys-370 methylation by SMYD2. Blockage of SMYD2 activity to p53 is mostly due to loss of the binding to Lys-372-methylated substrate (11). Our analysis suggests that this could be due to the steric hindrance between the methyl-lysine and carbonyl oxygen of Val-215 (the distance between the e-amine of Lys-372 and carbonyl oxygen of Val-215 in SMYD2 is 2.8 Å) and the potential side chain of Glu-187 and also possibly through disruption of the water molecule-mediated hydrogen bond network.

The side chains of both Lys-373 and Gln-375 in p53 extend to a shallow inner groove jointly formed by residues from CTD and SET (Fig. 3f). The aliphatic chain of p53 Lys-373 mainly positions toward an aromatic cage formed by Tyr-245, Tyr-344, Tyr-370, and Tyr-374 and interacts directly with the side-chain of Tyr-344 through van der Waal interactions, whereas its e-a-

---

**FIGURE 3.** SMYD2 and p53 peptide substrate interactions. a, surface representation of SMYD2 ternary complex structure with the same color as in Fig. 2b (left). The detailed p53 binding pocket is highlighted in the enlarged view on the right. b, comparison of the p53 binding sites between SMYD2-AdoMet and its ternary complex structures. The p53 peptide is depicted in a ribbon diagram. Side chains of residues in the ternary complex that show conformational changes are highlighted with the same color as in Fig. 2b, and the corresponding residues in the binary complex are shown by a purple line. c, superposition of SMYD2-p53 complex structure and SET7 in complex with mono-methylated p53 Lys-370 peptide (Protein Data Bank code 1XQH). SMYD2 is colored in cyan, and the p53 peptide is colored in yellow. SET7 is colored in salmon, and its p53 peptide is colored in purple. The red arrows indicate the b7 and b8 sheet loop of the SET core and the loop connecting SET core b10 and e1 of post-SET in SMYD2, respectively. d, stereoview of the p53 substrate interactions with SMYD2 within the ternary complex. Hydrogen bonds are indicated with yellow dashed lines. Water molecules are represented by small magenta cross-line marks. e, schematic representation of the interaction between SMYD2 and p53. Hydrogen bonds are denoted by yellow dashed lines. The peptide main-chain interactions are shown above the peptide trace (yellow), and side chain interactions are shown below. For clarity of the figure, residues that engaged in van der Waal interactions are not included. f, schematic diagram highlighting the interactions between p53 peptide and CTD of SMYD2. Hydrogen bonds are depicted by yellow dashed lines. Water molecules are represented by small magenta cross-line marks. SMYD2 unique residues are in cyan, and SMYD family conserved ones are in purple with ball-and-stick representation.
mine group forms hydrogen bonds with OD1 of Asn-380, OD2 of Asp-242, and OH atoms of Tyr-374 and Tyr-370, respectively (Fig. 3, d and e). The side chain of p53 Gln-375 is surrounded by residues His-341, Tyr-344, and Gln-345 and forms hydrogen bonds with side chains of Tyr-370, Tyr-245, and Gln-345, respectively. Although the TPRs among SMYD structures are well superimposed, this groove shows significant differences. Four of the nine residues (Tyr-370, Tyr-374, Asp-242, and Leu-244) are conserved in the SMYD family, and the rest (His-341, Tyr-344, Gln-345, Asn-380, and Tyr-245) are unique to SMYD2 (supplemental Fig. S6). Compared with a mutually well-positioned groove in SMYD2 that fits perfectly for both Lys-373 and Gln-375 of the p53 peptide, substitution of the corresponding variable residues in SMYD1 results in a shallower and more open groove accommodated with neutral charged surface potential and a deeper, narrower, and negatively charged one in SMYD3 (Fig. 3f) (data not shown). The observations collectively indicate that the same groove in SMYD1 and SMYD3 participates in substrate binding; however, the binding conformation of their corresponding peptide substrate presumably would be quite different from that of p53 in SMYD2, which may reflect the diversity of the SMYD family proteins in recognizing their substrates and, more importantly, the CTD domain in modulating the substrate specificity.

**CTD Domain and EDEE Motif Are Important for p53 Lys-370 Methylation by SMYD2**—Given the importance of the CTD domain in substrate binding pocket formation and, more specifically, the extensive interactions of the domain with the C-terminal residues of the p53 peptide, we reason that CTD is one of the key regions in determining the substrate specificity of SMYD2. To further elucidate the biochemical role of the CTD, we carried out mutagenesis studies to analyze the effect of the CTD to SMYD2 p53 methylation activity.

We first expressed and purified a CTD-deleted fragment of SMYD2 (aa 1–276; hereafter referred to as SMYD2-ΔCTD) to near homogeneity in *E. coli.* We then examined whether this deletion mutation alters the p53 Lys-370 methylation activity and affects substrate preference. Interestingly, SMYD2-ΔCTD shows about 5-fold increased activity to histone substrate H3K4 but not to H3K36, whereas the CTD deletion does not impair the p53 methylation (Fig. 4a, left). Nevertheless, the SMYD2-ΔCTD still maintains the preference for the p53 peptide substrate despite a higher activity to H3K4. To gain more insight into the role of the SMYD2 CTD in protein substrate methylation, we compared the WT and SMYD2-ΔCTD activities to the recombinant histone H3 and p53 protein. Strikingly, as shown in Fig. 4a (right), deletion of the CTD in SMYD2 leads to a 5-fold reduction of p53 protein methylation activity. However, no histone H3 methylation activity impairment was observed.

The residues forming an aromatic cage within the interface of the SET and CTD are crucial in interacting with the C terminus of the p53 peptide, so we then asked whether mutations of these residues would affect p53 methylation activity. Substitution of Tyr-245 and Tyr-374 to phenylalanine and alanine residues, respectively, leads to almost complete loss of the p53 methylation activity of SMYD2 (Fig. 4b). The hydroxyl group of Tyr-245 interacts with the ND1 atom of the His-341 in CTD. This group also contributes to stabilizing the side chain of the Gln-373 in p53 substrate through direct hydrogen bonding (Fig. 3e). Substitution of Tyr-245 to phenylalanine not only disrupts the hydrogen bond with His-341 of the enzyme but also abrogates the interactions with the substrate peptide, therefore leading to severe impairment of the enzymatic activity. On the other hand, in addition to the direct hydrogen bond between the OH group of Tyr-245 and OD1 of Asp-242, Tyr-374 contacts with side chains of Asp-242 and Tyr-370 through water molecule-mediated hydrogen bonds in the binary complex. This water molecule is replaced by the ϵ-amino group of Lys-373 in p53 substrate. However, the hydrogen bond network still remains. Replacement of the Tyr-374 by an alanine residue would abrogate the hydrogen bond network, thus destabilizing the CTD and SET domain interface conformation, resulting in the loss of the p53 substrate binding to the enzyme.

In addition to the aromatic cage, we then asked whether other regions involved in intramolecular interactions between the SET and CTD contribute to p53 binding. We focus on the interactions that are unique to SMYD2 with special attention to residues that are not conserved in other SMYD family proteins (supplemental Fig. S6). First, the residues of the EDEE motif at the SET core form a hydrogen bond network with Arg-390 and Glu-429 from CTD, and the side chain of Glu-189 interacts with residues Tyr-422 and Glu-425 of the CTD as well as Arg-58 of MYND through a series of water molecule-mediated hydrogen bond networks in both binary and ternary complex structures (Fig. 4c). These contacts are clearly absent in both SMYD1 and SMYD3 structures. Next, on the far left side of the SET and CTD interface, the side chain of the residue Asp-252 forms hydrogen bonds with the residue Arg-299 at the CTD. Adjacent to this, residues Arg-253 and Arg-306 form a positively charged surface that might also be involved in interacting with residues from the extended N terminus of p53, which, however, are disordered in the present ternary structure (Fig. 4d). To assess whether those two regions have a role in conferring p53 substrate specificity, we mutated those residues

---

**FIGURE 4. The EDEE motif of SET core and CTD domain are important for p53 recognition by SMYD2.** a, comparison of methyltransferase activities of the full-length and C-terminal domain-deleted SMYD2 to peptide (left) and protein substrates (right). Full-length SMYD2 is shown as a red column, whereas the SMYD2-ΔCTD is shown as a blank column. The experiments were done in triplicate; average values ± S.D. are shown. b, comparison of p53 peptide methyltransferase activities of WT SMYD2 and its CTD and SET interface hydrophobic groove mutations at different reaction times. Activity is represented as the mean of triplicate determinations, with error bars indicating one S.D. c, a close-up view showing the detailed interactions among EDEE motif, C terminus of the CTD domain, and the MYND domain. The corresponding residues in SMYD1 and SMYD3 are shown in yellow and purple lines, respectively. d, detailed view of intermolecular interactions between post-SET and the N terminus of the CTD domain. SMYD2 residues are represented by ball-and-stick models with their carbon atoms colored the same as in Fig. 2b. The corresponding residues in SMYD1 and SMYD3 are represented by a line with colors the same as in Fig. 2b. Hydrogen bonds are indicated by black dashed lines. Water molecules are represented by spheres with the oxygen atom colored in magenta. e, pull-down analysis of SMYD2 mutations by biotin-p53 peptides (left). The experiments were similar to that in Fig. 1 (middle). Methyltransferase activity comparison for mutant and WT SMYD2 to p53 peptide substrate in averaged triplicate is shown on the right.
Structure Basis of p53 Recognition by SMYD2

oppositely charged ones and compared their binding and methylation activities against p53 with the wild type SMYD2. We also have verified that such mutations did not change the biophysical properties of SMYD2 (data not shown). Substitution of Glu-187 and Glu-190, respectively, by a lysine not only severely reduces the binding but also significantly impairs the methylation activity to p53 by the enzyme (Fig. 4e). In addition, mutation of Asp-252 to an arginine residue results in a modest decrease of both p53 peptide binding and enzymatic activities. R253Q has a minor effect on p53 binding and enzymatic activity (Fig. 4e). We reasoned that the reduced activity of D252R may own to the abrogation of the hydrogen bonds with Arg-299 in the D252R mutation, which in turn may slightly alter the relative orientation between CTD and SET domain. The R253Q mutation may lead to the formation of the hydrogen bonds with Arg-306, which may still be feasible for p53 binding.

DISCUSSION

In this study, the biochemical and structural characterization of SMYD2 yields insights into the molecular basis of p53 substrate specificity by the enzyme. We have demonstrated that SMYD2 alone sufficiently monomethylates p53 in vitro and in vivo. The observation that SMYD2 shows very weak methylation activity to H3K36 and observable activity to H3K4 peptide is consistent with published data indicating that SMYD2 only methylated histone H3 to a very limited extent (1, 2). The preferential recognition of p53 Lys-370 substrate by SMYD2 highlights the mechanistic difference of the enzyme to non-histone and histone substrate methylation at the molecular level. Conversely, the evidence that only subtle conformational changes exist within the p53 binding site between binary and ternary complexes indicates that there is no need of significant structural rearrangement of SMYD2 for p53 substrate binding. However, this might not be the case for the binding of the enzyme to histone substrate. SMYD2 showing higher activity to p53 substrate prompted us to ask whether the enzyme will behave similarly to another tumor suppressor protein, Rb. Indeed, SMYD2 shows comparable activity to p53 Lys-370 and Rb Lys-860 (supplemental Fig. S7). This further suggests that the molecular basis underlying histone and non-histone methylation recognition by SMYD2 could be through distinct mechanisms.

The center of the SMYD2 substrate binding pocket seems to be more rigid because it shares high conformation conservation between the binary and ternary complexes, with only one noticeable difference from the side chain of Glu-187 (Fig. 3b). The observable conformational change in CTD between the two structures suggests that regions outside of this central pocket must be crucial in determining the specificity of the substrate binding. Of particular interest are residues from a hydrophobic groove within the bottom interface of SET core and CTD. The extensive interactions of those residues with the C terminus residues of p53 peptide substrate implicate the importance of this groove in mediating p53 substrate specificity (Fig. 3f). Indeed, Y374A and Y245F mutants almost completely lose the p53 methylation activities by SMYD2 (Fig. 4b), suggesting that maintaining both the conformation of this hydrophobic core and the corresponding interactions with p53 is crucial in proper recognition of the substrate. Additionally, the finding that deletion of CTD causes no activity change to the p53 peptide substrate but a great reduction of Lys-370 methylation to full-length p53 protein by SMYD2 indicates that (Fig. 4a, right) 1) the remaining SET domain substrate binding pocket is sufficient for p53 peptide binding; 2) CTD does not affect the catalytic activity of SMYD2; 3) the p53 peptide substrate may maintain its binding conformation, whereas the p53 protein may not be available to the CTD-deleted SMYD2; and 4) the CTD domain may also participate in interacting with distal regions of p53. The CTD of SMYD2 is composed of TPR repeats that have been demonstrated to be involved in protein-protein interactions (37). Interestingly, a positively charged surface formed by residues in the first helix of the CTD, Arg-292, Arg-296, and Arg-299, may be important to form electrostatic interactions with negatively charged residues, such as acidic residues of p53, to facilitate the methylation activity of the SMYD2. Investigation of the p53 protein and SMYD2 complex structure will further reveal the potential molecular basis of the CTD in p53 protein recognition.

In contrast to that in SMYD3 (38), deletion of CTD in SMYD2 leads to 5-fold increased methylation activity to histone H3K4. The exact mechanism of how SMYD2−ΔCTD shows higher activity to the histone H3K4 peptide is not clear in the present report. The CTD domain may play a role in negatively regulating efficient binding of the H3K4 peptide to full-length SMYD2.

The EDEE motif residues within the anti-parallel loop between β7 and β8 interact extensively with CTD through the hydrogen bond network to bridge the two domains together. This loop does not confer a conformational change in p53 substrate-bound ternary complex compared with that of the AdoMet-bound binary structure, suggesting the rigidity of the conformation and importance of the interactions in stabilizing the relative orientation between the SET and CTD for p53 substrate binding. Thus, we speculate that the EDEE motif is important in maintaining the relative conformation between CTD and SET domains for proper substrate binding. Indeed, E189K and E190K mutations lead to impaired binding and methylation activities to p53 by SMYD2 (Fig. 4e). These two mutations would attenuate the hydrogen bond network with Arg-390 and Glu-429 and thereby no longer allow the loop remaining in contact with the CTD. On the other side, substitution of Glu-187 by a lysine residue leads to dramatically decreased activity of the enzyme (Fig. 4e). The side chain of Glu-187 is involved in stabilizing p53 peptide conformation through hydrogen bonding to main-chain atoms of the substrate (Fig. 3, d and e). E187K would result in abrogation of such interactions with p53. Together, our data indicate that this unique EDEE motif of SMYD2 is crucial for maintaining the relative orientation of CTD and SET as well as for p53 substrate binding.

In conclusion, our data reported here demonstrate the physiological relevance of p53 Lys-370 methylation by SMYD2 and a structural explanation of the molecular basis in p53 recognition. In addition to p53 methylation, SMYD2 also methylates another tumor suppressor, Rb, implicating the importance of the SMYD2 activity in cellular functions through methylation
non-histone protein targets (13). Additional structural studies will aid in further understanding the substrate specificity and diversity of SMYD2 that in turn undoubtedly will help in the development of a selective and potent small molecule drug to treat SMYD2-caused diseases, such as esophageal squamous cell carcinoma (12).

Addendum—While this manuscript was under revision, two groups separately reported SMYD2 bound to AdoMet cofactor product AdoHcy and AdoMet analog sinefugine structures, respectively (38, 39). Another group published a crystal structure of SMYD2 in a ternary complex with p53 peptide and AdoHcy (40).

Acknowledgments—We thank Homan Chan, Maya Zhang, Jianyong Shou, Zhui Chen, and Wen Xiong for helpful discussions; the staff at Shanghai Synchrotron Radiation Facilities for data collection; and Yi Wei for mass spectrometry support.

REFERENCES
1. Brown, M. A., Sims, R. J., 3rd, Gottlieb, P. D., and Tucker, P. W. (2006) Mol. Cancer 5, 26
2. Abu-Farha, M., Lambert, J. P., Al-Madhoun, A. S., Elisma, F., Skerjanc, I. S., and Figys, D. (2008) Mol. Cell Proteomics 7, 560–572
3. Diehl, F., Brown, M. A., van Amerongen, M. J., Novoyatleva, T., Wietelmann, A., Harriss, J., Ferrazzi, F., Böttger, T., Harvey, R. P., Tucker, P. W., and Engel, F. B. (2010) PLoS One 5, e9748
4. Kawamura, S., Yoshigai, E., Kuhara, S., and Tashiro, K. (2008) Proc. Natl. Acad. Sci. U.S.A. 105, 161–168
5. Tan, X., Rotllant, J., Li, H., DeDeyne, P., and Du, S. J. (2006) Proc. Natl. Acad. Sci. U.S.A. 103, 2713–2718
6. Park, C. Y., Pierce, S. A., von Drehle, M., Ivey, K. N., Morgan, J. A., Blau, H. M., and Srivastava, D. (2010) Proc. Natl. Acad. Sci. U.S.A. 107, 20750–20755
7. Hamamoto, R., Furukawa, Y., Morita, M., Iimura, Y., Silva, F. P., Li, M., Yagyu, R., and Nakamura, Y. (2004) Nat. Cell Biol. 6, 731–740
8. Ha, L., Zhu, Y. T., Qi, C., and Zhu, Y. J. (2009) Cancer Res. 69, 4067–4072
9. Thompson, E. C., and Travers, A. A. (2008) PLoS One 3, e3008
10. Komatsui, M., Hamamoto, R., Silva, F. P., Yamaguchi, K., Nagayasu, T., Shibuya, M., Nakamura, Y., and Furukawa, Y. (2007) Cancer Res. 67, 10759–10765
11. Huang, J., Perez-Burgos, L., Placek, B. J., Kozaki, K. I., Muramatsu, T., Shimada, Y., Aiko, S., Yoshizumi, Y., Ichikawa, D., Otsuji, E., and Inazawa, J. (2009) Nature 460, 629–632
12. Komatsu, S., Imoto, I., Tsuda, H., Kozaki, K. I., Muramatsu, T., Shimada, Y., Aiko, S., Yoshizumi, Y., Ichikawa, D., Otsuji, E., and Inazawa, J. (2009) Carcinogenesis 30, 1139–1146
13. Saddic, I. A., West, L. E., Aslanian, A., Yates, J. R., 3rd, Rubin, S. M., Gozani, O., and Sage, J. (2010) J. Biol. Chem. 285, 37733–37740
14. Scoumanne, A., and Chen, X. (2008) Histol. Histopathol. 23, 1143–1149
15. Dai, C., and Gu, W. (2010) Trends Mol. Med. 16, 528–536
16. Shi, X., Kachirskia, I., Yamaguchi, H., West, L. E., Wen, H., Wang, E. W., Dutta, S., Appella, E., and Gozani, O. (2007) Mol. Cell 27, 636–646
17. West, L. E., Roy, S., Lachmi-Weiner, K., Hayashi, R., Shi, X., Appella, E., Kutateladze, T. G., and Gozani, O. (2010) J. Biol. Chem. 285, 37725–37732
18. Huang, J., Dorsey, J., Chuikov, S., Pérez-Burgos, L., Zhang, X., Jenuwein, T., Reingberg, D., and Berger, S. L. (2010) J. Biol. Chem. 285, 9636–9641
19. 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 Reingberg, D. (2004) Nature 432, 353–360
20. Sirinupong, N., Brunzelle, J., Ye, J., Pirzada, A., Nico, L., and Yang, Z. (2010) J. Biol. Chem. 285, 40635–40644
21. Sirinupong, N., Brunzelle, J., Doko, E., and Yang, Z. (2011) J. Mol. Biol. 406, 149–159
22. Xu, S., Wu, J., Sun, B., Zhong, C., and Ding, J. (2011) Nucleic Acids Res. 39, 4438–4449
23. Otwinowski, Z., and Minor, W. (1997) Methods Enzymol. 276, 307–326
24. Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ivey, K. N., Mc- Coy, A. J., Moriarty, N. W., Read, R. I., Sacchettini, J. C., Sauter, N. K., and Terwilliger, T. C. (2002) Acta Crystallogr. D Biol. Crystallogr. 58, 1948–1954
25. CCP4 (1994) Acta Crystallogr D Biol. Crystallogr. 50, 760–763
26. Emsley, P., and Cowtan, K. (2004) Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132
27. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. D Biol. Crystallogr. 53, 240–255
28. Luger, K., Rechsteiner, T. J., and Richmond, T. J. (1999) Methods Enzymol. 304, 3–19
29. Schmitges, F. W., Prusty, A. B., Faty, M., Stützer, A., Lingaraju, G. M., Aiwaszian, J., Sax, J., Hess, D., Li, L., Zhou, S., Bunker, R. D., Wirth, U., Bovwmeester, T., Bauer, A., Ly-Hartig, N., Zhao, K., Chan, H., Gu, J., Gut, H., Fischle, W., Müller, J., and Thomä, N. H. (2011) Mol. Cell 42, 330–341
30. Shi, X., Hong, T., Walter, K. L., Ewalt, M., Michishita, E., Hung, T., Carney, D., Peña, P., Lan, F., Kaadige, M. R., Lacoste, N., Cayrou, C., Davrazou, F., Saha, A., Cairns, B. R., Ayer, D. E., Kutateladze, T. G., Shi, Y., Côté, J., Chua, K. F., and Gozani, O. (2006) Nature 442, 96–99
31. Campagna-Slater, V., Mok, M. W., Nguyen, K. T., Feher, M., Najmanov- rich, R., and Schapira, M. (2011) J. Chem. Inf. Model 51, 612–623
32. 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
33. Couture, J. F., Collazo, E., Hauk, G., and Trievel, R. C. (2006) Nat. Struct. Mol. Biol. 13, 140–146
34. Qian, C., and Zhou, M. M. (2006) Cell Mol. Life Sci. 63, 2755–2763
35. Cheng, X., Collins, R. E., and Zhang, X. (2005) Annu. Rev. Biophys. Biomol. Struct. 34, 267–294
36. Del Rizzo, P. A., Couture, J. F., Dirks, M. L., Strunk, B. S., Roiko, M. S., and Aiko, S., Yoshizumi, Y., Ichikawa, D., Otsuji, E., and Inazawa, J. (2009) Mol. Cell 34, 636–646
37. D’Andrea, L. D., and Regan, L. (2003) Trends Biochem. Sci. 28, 655–662
38. Xu, S., Zhong, C., Zhang, T., and Ding, J. (2011) J. Mol. Biol. In press
39. Shi, X., Sirinupong, N., Brunzelle, J., and Yang, Z. (2011) PLoS One 6, e21640
40. Ferguson, A. D., Larsen, N. A., Howard, T., Pollard, H., Green, I., Grande, C., Cheung, T., Garcia-Arenas, R., Cowen, S., Wu, J., Godin, R., Chen, H., and Keen, N. (2011) Structure 19, 1262–1273