

ARTICLE

DOI: 10.1038/s42003-018-0196-2

OPEN

# An asparagine/glycine switch governs product specificity of human N-terminal methyltransferase NTMT2

Cheng Dong<sup>1</sup>, Guangping Dong<sup>2</sup>, Li Li<sup>1</sup>, Licheng Zhu<sup>1,3</sup>, Wolfram Tempel<sup>1</sup>, Yanli Liu<sup>1</sup>, Rong Huang<sup>2</sup> & Jinrong Min<sup>1,4</sup>

$\alpha$ -N-terminal methylation of proteins is an important post-translational modification that is catalyzed by two different N-terminal methyltransferases, namely NTMT1 and NTMT2. Previous studies have suggested that NTMT1 is a tri-methyltransferase, whereas NTMT2 is a mono-methyltransferase. Here, we report the first crystal structures, to our knowledge, of NTMT2 in binary complex with S-adenosyl-L-methionine as well as in ternary complex with S-adenosyl-L-homocysteine and a substrate peptide. Our structural observations combined with biochemical studies reveal that NTMT2 is also able to di-/tri-methylate the GPKRIA peptide and di-methylate the PPKRIA peptide, otherwise it is predominantly a mono-methyltransferase. The residue N89 of NTMT2 serves as a gatekeeper residue that regulates the binding of unmethylated versus monomethylated substrate peptide. Structural comparison of NTMT1 and NTMT2 prompts us to design a N89G mutant of NTMT2 that can profoundly alter its catalytic activities and product specificities.

<sup>1</sup>Structural Genomics Consortium, University of Toronto, Toronto M5G1L7 ON, Canada. <sup>2</sup>Department of Medicinal Chemistry and Molecular Pharmacology, Center for Cancer Research, Institute for Drug Discovery, Purdue University, West Lafayette, IN 47907, USA. <sup>3</sup>School of Life Sciences, Jingtangshan University, 343009 Ji'an, Jiangxi, China. <sup>4</sup>Department of Physiology, University of Toronto, Toronto M5S 1A8 ON, Canada. These authors contributed equally: Cheng Dong, Guangping Dong. Correspondence and requests for materials should be addressed to R.H. (email: [huang-r@purdue.edu](mailto:huang-r@purdue.edu)) or to J.M. (email: [jr.min@utoronto.ca](mailto:jr.min@utoronto.ca))

Protein methylation participates in regulation of a broad spectrum of cellular processes. Besides the extensively studied protein lysine/arginine methylation<sup>1,2</sup>, the addition of a methyl group at the free  $\alpha$ -N-termini of proteins represents a unique mode of post-translational modification and remains underexplored, though its discovery dates back to 1976<sup>3</sup>.  $\alpha$ -N-terminal methylation is conserved from prokaryotes to humans, and a variety of N-terminally methylated proteins have been identified in ribosomal and histone proteins<sup>4,5</sup>. Recent studies have shed some light on the functions of N-terminal methylation. For instance, in yeast, loss of N-terminal methylation of the ribosomal protein Rpt1 leads to impaired cell growth and hypersensitivity to stress<sup>6</sup>. In *Drosophila*,  $\alpha$ -N-terminal methylation level of histone H2B increases during development<sup>7</sup>. In human, loss of N-terminal methylation of regulator of chromatin condensation 1 (RCC1) diminishes its binding affinity for DNA, and results in defects of spindle assembly and chromosome segregation<sup>8</sup>. The absence of N-terminal methylation of DNA damage-binding protein 2 (DDB2) decreases the localization of DDB2 to UV-induced DNA damage foci and hinders nucleotide excision repair<sup>9</sup>. The N-terminal methylation of CENP-B enhances its binding to centromeric DNA in cells<sup>10</sup>. The N-terminal methylation of CENP-A is not only required for cell survival, recruitment of CENP-T/I, and proper chromosome segregation, but may also accelerate tumorigenesis in p53-deficient background<sup>11</sup>.

The first  $\alpha$ -N-terminal methyltransferase (NTMT), human NTMT1 (also known as METTL11A/NRMT1) and its yeast ortholog, had just been functionally characterized in 2010<sup>5,12</sup>, although the first crystal structure of NTMT1 had been determined in 2005 (PDB ID: 2EX4). NTMT1 is an S-adenosyl-L-methionine (SAM)-dependent methyltransferase. During the enzymatic reaction, NTMT1 transfers a methyl group from SAM to the  $\alpha$ -amino group of the protein substrates, resulting in the production of S-adenosyl-L-homocysteine (SAH) and  $\alpha$ -N-methylated proteins. NTMT1 recognizes proteins bearing an N-terminal X-P-K/R consensus sequence, including RCC1, RB1, DDB2, CENP-A/B, PARP3, etc.<sup>9–15</sup>. Knockdown of NTMT1 results in hypersensitivity of breast cancer cell lines to double-stranded DNA breaks (DSBs) and increased proliferation of estrogen receptor positive breast cancer cells MCF-7 and LCC9<sup>16</sup>; NTMT1 knockout mice are phenotypically defective in DNA repair and exhibit premature aging<sup>17</sup>. In 2013, another human NTMT, NTMT2/METTL11B/NRMT2, was described as a mono-methyltransferase<sup>18</sup>, although NTMT1 is able to catalyze trimethylation<sup>12,19,20</sup>. We and another group previously solved crystal structures of NTMT1 in ternary complex with SAH and substrate peptides, and proposed a catalytic mechanism<sup>13,21</sup>. However, the molecular mechanism of methylation by NTMT2 remains elusive.

In order to unravel the molecular basis of the substrate and product specificity of NTMT2, we determine the X-ray crystal structures of NTMT2 in binary complex with SAM, as well as in ternary complex with SAH and an RCC1-derived peptide (SPKRIA). We also perform mutational analysis and comprehensively investigate the substrate specificity and product methylation states of NTMT1 and NTMT2 for a panel of 20 peptides. Our results manifest that NTMT2 is not a sole mono-methyltransferase, but is also able to fully methylate both GPKRIA and PPKRIA peptides. Furthermore, we identify N89 as a key residue for product specificity of NTMT2. The N89G mutant of NTMT2 is more active than the wild type NTMT2, which is able to convert S/APKRIA peptides from the mono-methylation state to di-/tri-methylation states in our *in vitro* enzymatic assays.

**Table 1** Data collection and refinement statistics

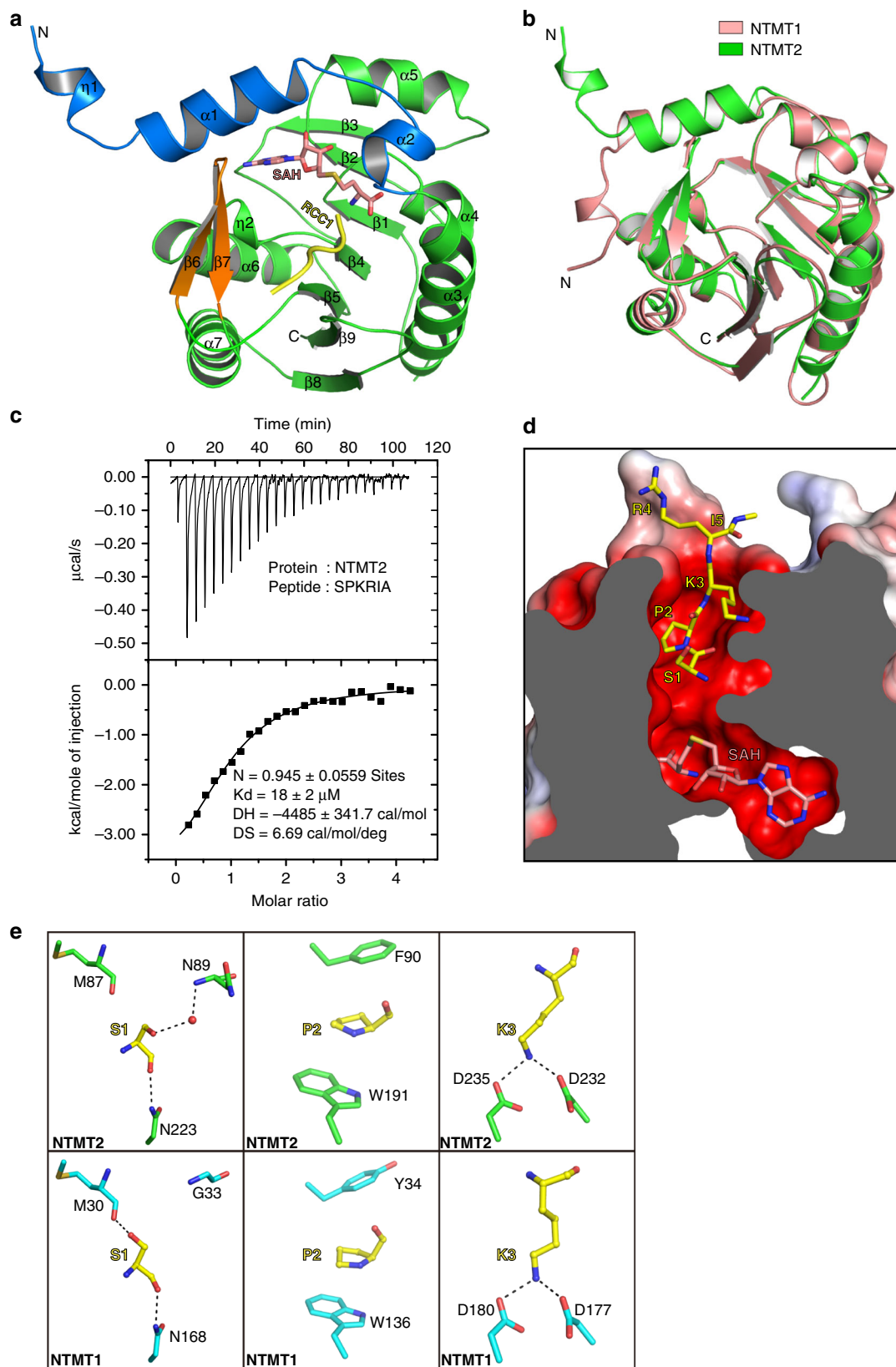
	NTMT2-SAM	NTMT2-SAH-SPKRIA
<i>Data collection</i>		
Space group	$P 2_1 2_1 2$	$P 4_3$
Cell dimensions		
$a, b, c$ (Å)	45.59, 132.96, 42.29	44.36, 44.36, 262.06
$\alpha, \beta, \gamma$ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Resolution (Å)	45.60–2.00 (2.05–2.00)	31.14–1.20 (1.22–1.20)
$R_{\text{sym}}$ or $R_{\text{merge}}$	0.086 (0.700)	0.074 (0.973)
$I/\sigma I$	10.2 (1.7)	11.0 (2.0)
Completeness (%)	94.4 (91.7)	96.0 (93.2)
Redundancy	2.6 (2.6)	6.1 (6.6)
<i>Refinement</i>		
Resolution (Å)	18.98–2.00	31.10–1.20
No. of reflections	16,002	145,394
$R_{\text{work}}/R_{\text{free}}$	0.208/0.250	0.154/0.178
No. of atoms	1780	4160
Protein	1691	3690
Ligand/ion	–	72
Water	63	314
<i>B</i> -factors	27.4	17.0
Protein	27.3	16.1
Ligand/ion	–	17.6
Water	28.9	25.7
R.m.s. deviations		
Bond lengths (Å)	0.014	0.02
Bond angles (°)	1.7	1.9

The values in parentheses are for highest-resolution shell

## Results

**Overall structure of NTMT2.** So far, two NTMTs that methylate X-P-K/R are identified in mammals. NTMT1 has been classified as a tri-methyltransferase<sup>12,19,20</sup>, whereas NTMT2 was reported as a mono-methyltransferase<sup>18</sup>. To determine the molecular basis of different product specificities between NTMT2 and NTMT1, we solved crystal structures of the NTMT2-SAM binary complex and the NTMT2-SAH-SPKRIA peptide ternary complex. Despite that an unmethylated peptide was used for crystallization, a methylated  $\alpha$ -amino group in the N-terminus of the peptide was traced in the crystal structure. Crystal diffraction data and model refinement statistics are summarized in Table 1. Based on the folding pattern, NTMT2 is a SAM-dependent class I methyltransferase<sup>22</sup>, which consists of a central seven-stranded  $\beta$  sheet ( $\beta 1$ – $\beta 5$  and  $\beta 8$ – $\beta 9$ ), flanked by three  $\alpha$ -helices ( $\alpha 3$ – $\alpha 5$ ) and two  $\alpha$ -helices ( $\alpha 6$ – $\alpha 7$ ) on each side, respectively (Fig. 1a). In addition, NTMT2 contains two auxiliary regions: an N-terminal  $\alpha$ -lid ( $\eta 1$ ,  $\alpha 1$ – $\alpha 2$ ), and a  $\beta$ -lid ( $\beta 6$ – $\beta 7$ ) inserted between  $\beta 5$  and  $\alpha 8$ . These two lids cover the core domain that contributes to the substrate recognition (Fig. 1a, Supplementary Fig. 1a). The overall architecture of NTMT2 closely resembles that of NTMT1 with a backbone RMSD below 1 Å (Fig. 1b).

The NTMT2-SAM binary complex subtly differs from the NTMT2-SAH-SPKRIA ternary complex in that the turn of the  $\beta$ -lid of the ternary complex is shifted closer to the catalytic core, which facilitates hydrogen bonding between substrate K3 and both D232 and D235 of NTMT2 (Supplementary Fig. 1b). The co-factor SAH/SAM in the ternary complex is nearly superimposable with that in the substrate-free binary complex. Based on sequence alignments, the N-terminal methyltransferases have a conserved D<sub>x</sub>G<sub>x</sub>G<sub>x</sub>GR motif located on the  $\beta 1$ – $\alpha 4$  loop, which is involved in the SAH/SAM binding (Supplementary Fig. 1c). Indeed, the carboxylate moiety of SAH forms a salt bridge with



**Fig. 1** Overall structure of NTMT2. **a** Ribbon diagram of NTMT2 in ternary complex with SAH and a SPKRIA peptide derived from N-terminal RCC1. **b** Structural alignment of NTMT1 (PDB: 5E1B) and NTMT2. **c** Isothermal titration calorimetry measurement of the interaction between NTMT2 and the SPKRIA peptide. **d** Cross-section view of the SAH-SPKRIA-binding pocket. The electrostatic potential surface of NTMT2 plotted at  $\pm 5$  kT/e (red, negative; blue, positive). **e** Close-up views of the S-P-K motif recognition by NTMT2 and NTMT1 (PDB: 5E1B). Residues of NTMT2 and NTMT1 involved in the interactions are labeled and shown in green and cyan, respectively. Hydrogen bonds are indicated as dashed lines, and water molecule is shown as red sphere

R129, and its amino group is held by hydrogen bonds with the G124 and Q190 carbonyls, respectively (Supplementary Fig. 1d). Its ribosyl moiety is stacked against Y76 with its 2',3'-diol group coordinated by D146. Its adenine group forms hydrogen bonds with the Q175 side-chain carboxamide and the L174 main-chain amide group, respectively. M147, corresponding to NTMT1 I92, forms a methionine- $\pi$  interaction with the adenine group as well (Supplementary Fig. 1d). As a result, the SAH molecule is deeply buried by the  $\alpha$ -lid and the substrate is inserted into the conserved binding pocket (Supplementary Fig. 2a).

**NTMT1 and NTMT2 employ a similar substrate recognition mode.** RCC1 is a physiological substrate of NTMT1<sup>12</sup>, but no physiological substrates have been reported for NTMT2. To compare substrate recognition by NTMT2 and NTMT1, we first examined whether NTMT2 could bind to any known NTMT1 substrates like RCC1. Isothermal titration calorimetry (ITC) shows that NTMT2 is able to bind to the N-terminal hexapeptide of RCC1 (SPKRIA) with a  $K_d$  value of 18  $\mu$ M (Fig. 1c), which is  $\sim$ 20-fold weaker than that of NTMT1 ( $K_d$  value of 0.8  $\mu$ M)<sup>13</sup>. Our ternary complex structure shows that, like in NTMT1, the substrate peptide is deeply inserted into the binding pocket through hydrophilic and hydrophobic interactions, and the  $\alpha$ -amine of the first residue (S1) points towards the sulfur atom of SAH (Fig. 1d). In the NTMT2 complex, S1 fits snugly into the channel, where the main-chain O atom is anchored by N223 in NTMT2 (N168 in NTMT1) through hydrogen bonding (Fig. 1e, Supplementary 2b). The N223A mutant abrogates substrate binding affinity and has no enzymatic activity (Supplementary Fig. 2c, d). The side chain of S1 has adopted the p conformation<sup>23</sup> and interacts, mediated by solvent, with the main chain of N89 in NTMT2, whereas in the NTMT1 complex, S1 present in the m conformer hydrogen bonds with the NTMT1 main chain at M87 (Fig. 1e).

The substrate residue P2 is sandwiched by the aromatic side chains of W191 and F90 (Fig. 1e). W191 also appears to play an important role in NTMT2 structural stability, as mutations of W191 to A or even hydrophobic amino acids Y, L or I, resulted in insoluble proteins. The  $\epsilon$ -amine of substrate K3 interacts with two aspartates in both NTMT1 and NTMT2 structures (Fig. 1e). Neither mutant D232A nor D235A of NTMT2 exhibits any measurable interaction with the substrate or the enzymatic activity (Supplementary Fig. 2c, e, f). In NTMT2, the main-chain N and O atoms of K3 are involved in a water-mediated hydrogen bond with the side chain of N223 and a direct hydrogen bond with the main chain of I270, respectively (Supplementary Fig. 2b).

Substrate residues downstream of the S-P-K motif protrude from the pocket on the surface of NTMT2 (Fig. 1d), indicating that they are not essential for the recognition. Indeed, the fourth residue R4 is not involved in direct interaction and the main chain of the fifth residue I5 is accommodated by Q268 (Supplementary Fig. 2b). In summary, aside from minor differences, the sets of residues responsible for substrate binding in NTMT1 and NTMT2 overlap (Fig. 1e), which explains why these two enzymes share an X-P-K/R substrate recognition motif.

**NTMT2 also acts as a di-/tri-methyltransferase.** Previous studies have indicated that NTMT1 can catalyze mono-methylation, di-methylation, and tri-methylation, but NTMT2 was reported as a mono-methyltransferase<sup>12,18–20</sup>. Comparison of the substrate-binding sites (Fig. 1e) does not provide a straightforward explanation. In order to gain a full understanding of the methylation states of the products, we synthesized 20 XPKRIA peptides with X being any of the 20 standard amino acids, and characterized the

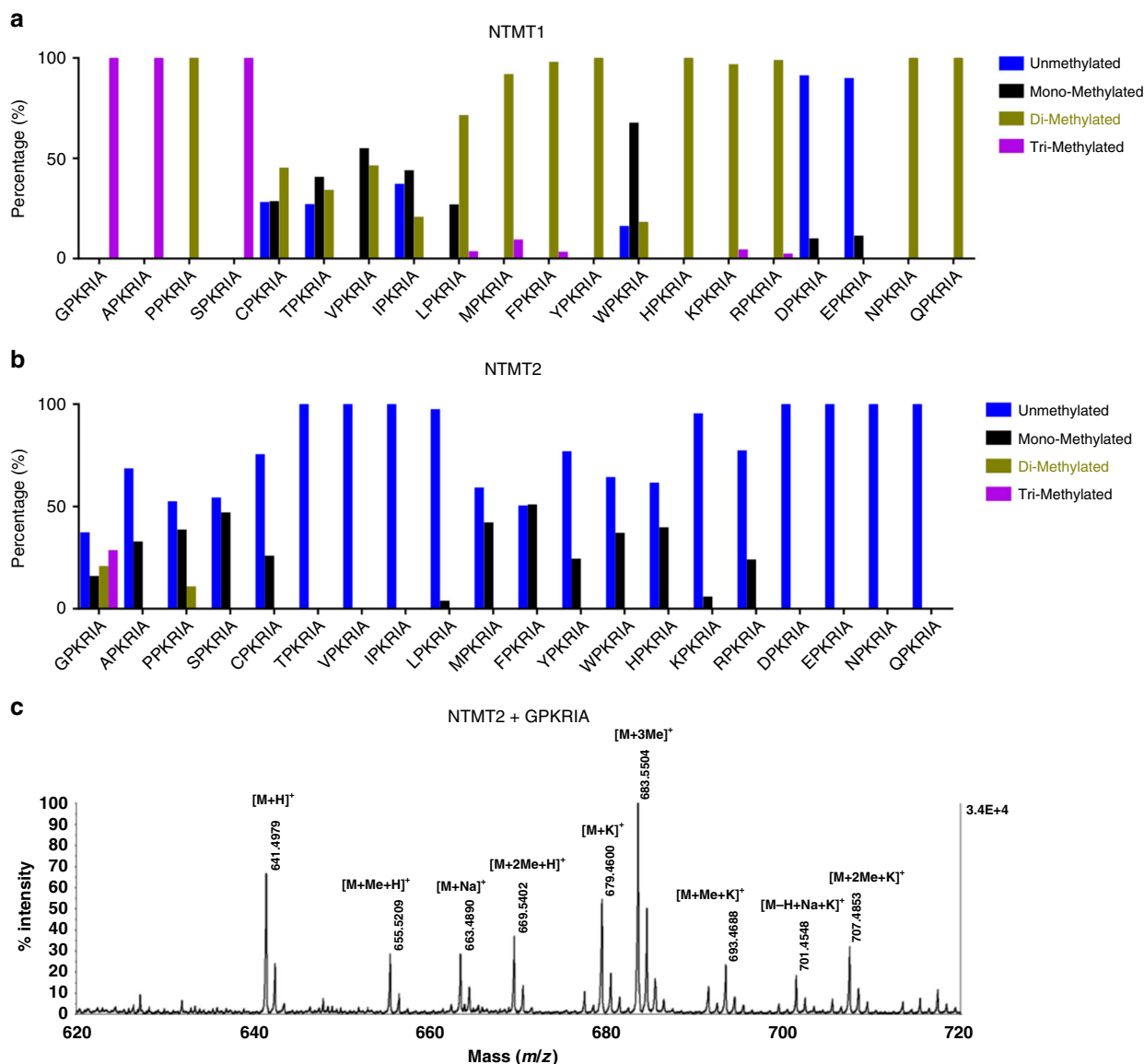
methylation progression of these peptides by both NTMT1 and NTMT2 using matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) technique under the same conditions<sup>19</sup> (Fig. 2a, b, Supplementary Data 1). Since the enzymatic activity of NTMT2 is much lower than that of NTMT1, we had to use higher concentration of NTMT2 (2  $\mu$ M) in our assay. We also increased the NTMT1 concentration from previously used 0.2 to 2  $\mu$ M<sup>13</sup> for comparison. Our results revealed that NTMT1 could exhaustively methylate [A/G/P/S]-PKRIA peptides to full degree of methylation states, and mainly di-methylate the other peptides. Of note, NTMT1 even methylated [D/E]-PKRIA, albeit the catalytic activities for those two peptides were very low (Fig. 2a), as we expected that the negatively charged pocket would repel these substrates<sup>13</sup>. Compared with NTMT1, NTMT2 has a weaker catalytic capability, which only mono-methylated most XPKRIA peptides. Surprisingly, NTMT2 could convert GPKRIA peptide into mono-methylation, di-methylation, and tri-methylation states (Fig. 2b, c). Moreover, NTMT2 was able to di-methylate the PPKRIA peptide despite low yield (Fig. 2b). Taken together, NTMT2 is more than a mono-methyltransferase, but also functions as di-/tri-methyltransferase, depending on the identity of the first amino acid of the substrates.

**N89 of NTMT2 serves as a gatekeeper for the catalytic activity and alters its product specificity.** NTMT2, unlike NTMT1, failed to di-/tri-methylate SPKRIA (Fig. 2a, b). Comparison of the substrate-binding sites in NTMT1 (M30, N168, W136, D177, and D180) and NTMT2 (M87, N223, W191, D232, and D235) revealed that the residues that engaged the direct substrate binding were identical (Fig. 1e). Hence, these direct substrate-binding residues are not responsible for such different methylation states of products generated by NTMT1 and NTMT2.

In order to understand the molecular basis for product specificities between NTMT1 and NTMT2, we compared the residues near the substrate-binding sites and found that two tandem aromatic residues near the  $\alpha$ -amino group of the peptide are different between NTMT1 and NTMT2 (FY75-76 in NTMT2 and YW19-20 in NTMT1) (Fig. 3a, Supplementary Fig. 3a, b). Previous studies have shown that the aromatic residue mutant Y305F of SET7/9, or Y334F of SET8 could alter the specificity of the SET domain histone lysine methyltransferases from a mono-methyltransferase to a di-methyltransferase<sup>24–26</sup>. Therefore, we investigated both single and double mutants of NTMT2 including F75Y, Y76W, and FY75-76YW, but none of them altered the methylation state of NTMT2 (Supplementary Fig. 3c–e), which is consistent with a recent report<sup>27</sup>. One reason is that the Y76 of NTMT2 or Y19 of NTMT1 is 5.4 and 6.8 Å away from the  $\alpha$ -amino group, respectively (Supplementary Fig. 3a, b). Such long distances lack the ability to form a direct CH $\cdots$ O hydrogen bond with a methylated  $\alpha$ -amino group as observed in the SET domain methyltransferases<sup>26</sup>. Nevertheless, the wild-type NTMT2 can carry out tri-methylation of G-PKRIA peptide and di-methylation of P-PKRIA peptide, indicating that  $\alpha$ -amino groups of some substrates are able to be further deprotonated and fully methylated by NTMT2. Therefore, we hypothesized that other residue(s) near the substrate-binding site may regulate such product specificity.

We then scrutinized the overall structures of NTMT1 and NTMT2 and identified a small conformational change in the  $\alpha$ 2- $\alpha$ 3 loop, referred to as  $\Omega$  loop here. NTMT1 adopts an open conformation, while NTMT2 displays a relatively closed conformation (Fig. 3b). Close examination identified that NTMT1 carries a G33, whereas NTMT2 harbors an N89 at the same location (Fig. 3a). The comparison of the surface representation of NTMT2 and NTMT1 indicates that the N89 in NTMT2 creates



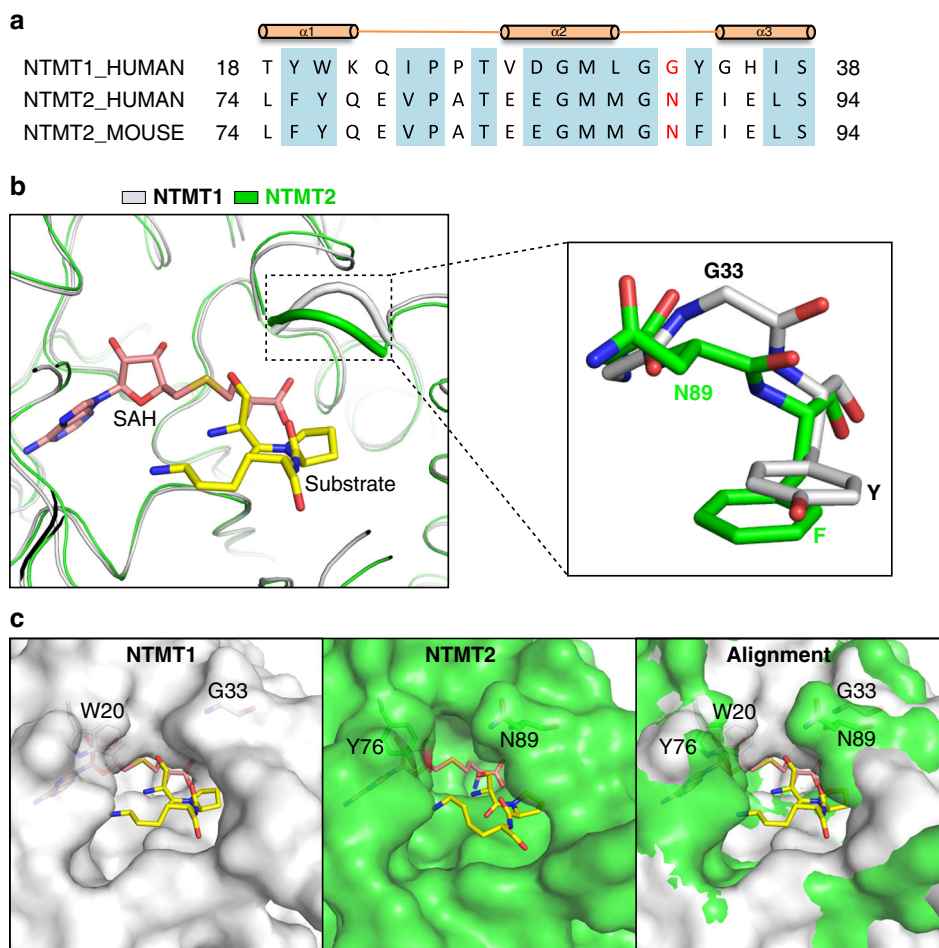


**Fig. 2** NTMT1 and NTMT2 exhibit different methylation patterns. MALDI-MS analysis of 20 peptides catalyzed by NTMT1 (**a**) and NTMT2 (**b**). Synthetic peptides of XPKRIA with the first position (X) being any of the 20 standard amino acids are used in this study. **c** MALDI-MS analysis of GPKRIA peptide catalyzed by NTMT2

a narrower substrate-binding channel compared to G33 in NTMT1 (Fig. 3c). Based on this observation, we generated the NTMT2 N89G mutant and analyzed the methylation products by the MALDI-MS assay. Our results reveal that N89G mutant display enhanced activity than wild-type NTMT2, and has a higher rate of mono-methylation (Fig. 4a, Supplementary Data 1). Remarkably, it completely methylates the SPKRIA peptide to mono-methylation and di-methylation states. Moreover, it also efficiently converts the APKRIA peptide to mono-methylation and di-methylation states, as well as to a tri-methylation state (Fig. 4a). Therefore, N89 functions as a gatekeeper for the substrate binding and product specificity of NTMT2.

**The N89G mutant has a higher affinity for mono-methylated peptide than wild type.** We hypothesized that the long side chain of N89 in NTMT2 might cause steric hindrance for binding the methylated SPKRIA (Sme1PKRIA) peptide. Indeed, wild-type NTMT2 barely demonstrated any binding affinity, while the

single point mutation of N89G in NTMT2 bound to Sme1PKRIA peptide with a  $K_d$  value of 36  $\mu\text{M}$  (Fig. 4b). Likewise, wild-type NTMT2 had no detectable binding to Ame1PKRIA, but the N89G mutant exhibited a  $K_d$  value of 32  $\mu\text{M}$  towards Ame1PKRIA (Fig. 4c). Consistent with these binding data, our MALDI-MS results showed that wild-type NTMT2 did not produce any di-methylation or tri-methylation products (Fig. 4d), but N89G mutant has robust enzymatic activity towards Sme1PKRIA (Fig. 4e). These observations also explain why NTMT2 is only capable of transferring one methyl group to the  $\alpha$ -amine of serine, whereas N89G could further methylate mono-methylated serine. Taken together, the smallest glycine, even when it was mono-methylated or di-methylated (Gme1/2-PKRIA) could still insert into the binding pocket. But the mono-methylation of any amino acid other than glycine or proline would result in products that were not able to pass through the gatekeeper N89 and thereby would be detached from the NTMT2 catalytic site. Nevertheless, the N89G mutant would expand the substrate-binding channel



**Fig. 3** N89 of NTMT2 serves as a gatekeeper residue to regulate the substrate entry into the active pocket. **a** Sequence alignment of human NTMT1, human NTMT2, and mouse NTMT2. The G33 of NTMT1 and N89 of NTMT2 are colored in red. **b** Structural alignment of NTMT1-SPKRIA (PDB: 5E1B) (gray) and NTMT2-SPKRIA (green) complexes. The  $\Omega$  loops are indicated by dashed lines and the residues of the  $\Omega$  loops are labeled. The  $\Omega$  loop of NTMT1 adopts an open conformation, whereas NTMT2 acts as a relatively closed conformation. **c** Surface representations of the substrate active pocket from NTMT1 (gray) and NTMT2 (green). The Y76 of NTMT2 generates a large empty space above the S atom of SAH compared to W20 of NTMT1. The N89 of NTMT2 creates a narrow substrate-binding channel compared to G33 of NTMT1

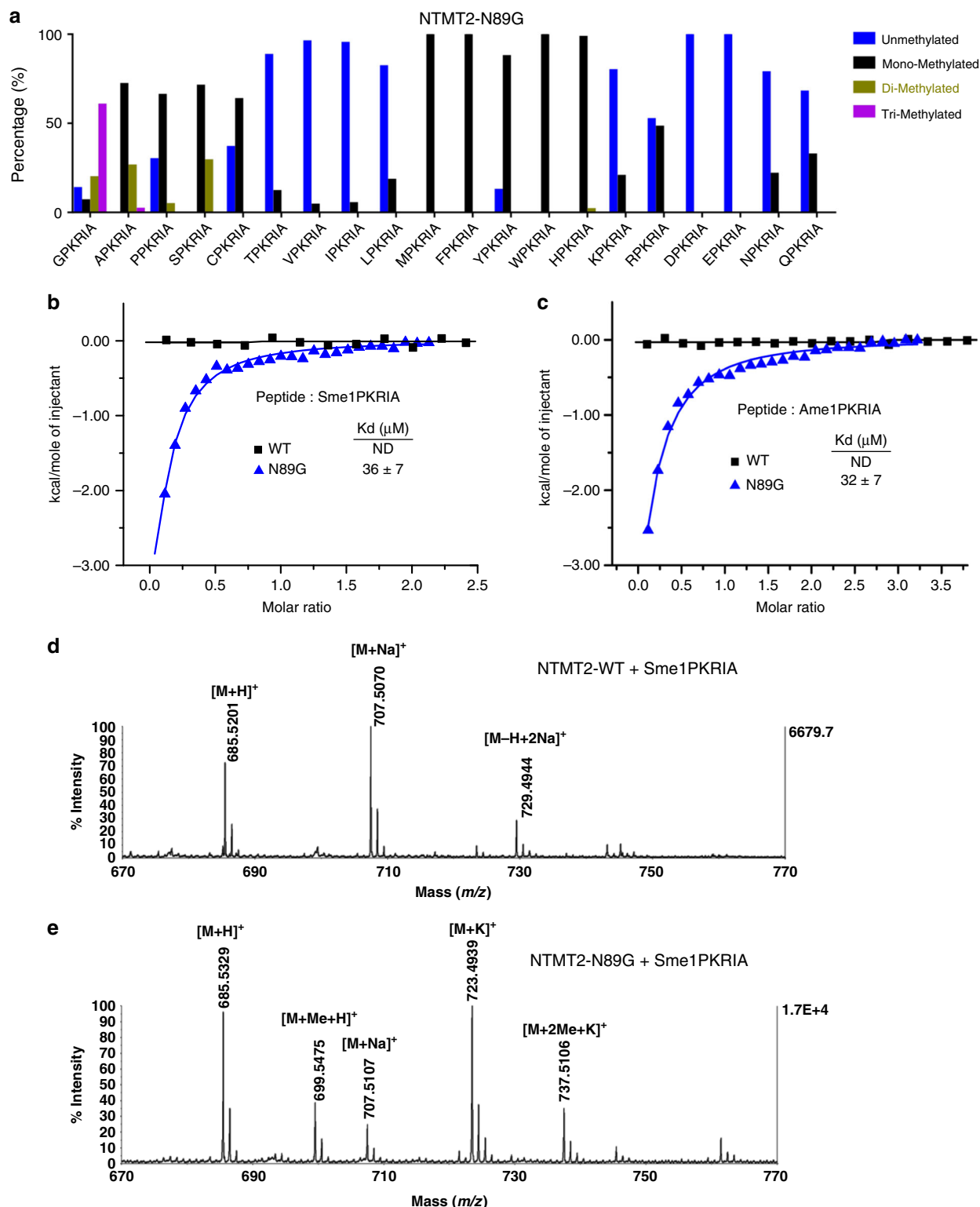
and facilitate the substrates access to the active site, resulting in enhanced enzymatic activity, even alteration of the product specificity from mono-methylation to tri-methylation.

**Proposed catalytic mechanism of NTMT2.** Although there is neither direct hydrogen bond formed between NTMT2 and the substrate  $\alpha$ -amino group, nor a general base that is responsible for deprotonating  $\alpha$ -amino group, there are multiple water molecules present at the active channel that mediated the interaction between  $\alpha$ -amino group of substrates and the active site (Supplementary Fig. 4a). We proposed that the  $\alpha$ -amino group of the first residue is further stabilized by multiple water molecules-mediated hydrogen bonds, orienting the acceptor  $\alpha$ -amino group toward the sulfonium ion of SAM in close proximity. Y76, H195, D222, and D235 serve as general bases for proton abstraction from the substrate  $\alpha$ -amino group through multiple water-mediated hydrogen-bonding networks (Supplementary Fig. 4b, c). The deprotonated  $\alpha$ -amine is a better nucleophile to attack the methyl group of SAM to accomplish the methyl transfer.

## Discussion

We determined the crystal structure of NTMT2 in ternary complex with cofactor SAH and a substrate peptide SPKIRA,

which is consistent with its classification as a SAM-dependent class I methyltransferase<sup>22</sup>. It has been proposed that NTMT2 acts as a mono-methyltransferase<sup>18</sup>. In this study, we comprehensively investigated the substrate recognitions of NTMT1 and NTMT2, and the methylation state specificities of their products. To our knowledge, this is the first study to indicate that NTMT2 can also achieve tri-methylation or di-methylation when the glycine or proline is present at first residue of substrates, respectively. Our data also suggest that NTMT2 employs an  $S_N2$  nucleophilic attack mechanism by multiple water-mediated deprotonation of  $\alpha$ -amino group. Like NTMT1, NTMT2 processes the methylation reaction by a dynamic distributive mechanism (Fig. 2c), in which NTMT2 releases the product after each methyl addition, and then rebind for further methylation<sup>20</sup>. Notably, NTMT2 does not bind to Sme1PKRIA or Ame1PKRIA peptide *in vitro*, resulting in loss of the di-methylation activity. Although both human N-terminal methyltransferases share a high degree of similarity in terms of global structures, NTMT2's product specificity and enzymatic activity clearly distinguishes itself as a unique cellular methyltransferase. We found that the  $\Omega$  loop of NTMT2 guards the substrate-binding pocket in a relatively stricter manner compared to that of NTMT1. We also identify that N89 serves as a gatekeeper residue to control the



**Fig. 4** The NTMT2 N89G mutant exhibits increased methylation activity. **a** MALDI-MS analysis of 20X-PKRIA peptides catalyzed by NTMT2-N89G mutant. ITC measurements of the interactions between the wild-type (WT) or N89G mutant NTMT2 and the Sme1PKRIA (**b**) or Ame1PKRIA peptide (**c**). ND indicates no detectable binding. **d** MALDI-MS analysis of Sme1PKRIA peptide catalyzed by the WT NTMT2. **e** MALDI-MS analysis of Sme1PKRIA peptide catalyzed by the NTMT2 N89G mutant

entry of substrate into the active pocket. Accordingly, we generated a N89G mutant of NTMT2 that exhibited higher enzymatic efficiency towards the substrates. In addition, this mutant is sufficient to bind to Ame1PKRIA or Sme1PKRIA peptide, and converts them to tri-methylation or di-methylation, respectively.

NTMT1 and NTMT2 exhibit different expression patterns. For instance, the NTMT1 mRNA is abundantly expressed in mouse brain and ubiquitously expressed in human tissues, whereas the NTMT2 mRNA is strongly expressed in mouse and human muscles<sup>18,28</sup>. Therefore, NTMT1 and NTMT2 may contribute to

different biological activities in a cellular context-dependent manner. After the initiator methionine excision, the proteins bearing a G-P-K/R consensus sequence at the N-terminus (for instance, CENP-A/B) could be converted to tri-methylation states by both NTMT1 and NTMT2. For the other types of substrates, NTMT2 primarily carries out mono-methylation, while NTMT1 is responsible for further di-/tri-methylation.

## Methods

**Protein expression and purification.** DNA encoding human NTMT2 (residues 58–278) was amplified by PCR and cloned into pET28-MKH8SUMO vector using the InFusion™ cloning kit (ClonTech) employing the manufacturer's instructions. This recombinant plasmid contains a SUMO tag and TEV cleavage site at the N-terminus. The protein was overexpressed in *Escherichia coli* BL21 (DE3) competent cells after induction with 0.2 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) at 16 °C overnight. Cells were lysed in 20 mM Tris-HCl pH 7.5, 400 mM NaCl, 5% glycerol and 2 mM beta-mercaptoethanol buffer and purified by Ni-NTA agarose chromatography. The SUMO tag was cleaved by TEV protease at 4 °C overnight and removed by reloading onto the Ni-NTA. The protein was diluted and applied onto HiTrap Q HP anion exchange chromatography column (GE Healthcare) equilibrated with 20 mM Tris-HCl pH 7.5, 25 mM NaCl and 0.5 mM tris (2-carboxyethyl) phosphine (TCEP). The proteins were eluted with a linear gradient of 0–50% elution buffer (20 mM Tris-HCl pH 7.5, 1 M NaCl and 0.5 mM TCEP). The proteins were further purified by gel filtration Superdex 200 10/300 (GE Healthcare). The gel filtration buffer contains 20 mM Tris-HCl pH 7.5, 150 mM NaCl and 0.5 mM TCEP. The purified protein was concentrated to 20 mg mL<sup>-1</sup> for crystallization. Mutants were created by QuickChange PrimeSTAR Mutagenesis Basal Kit and verified by DNA sequencing, the primers used in this study are listed in Supplementary Table 1. Mutant proteins were purified with the same protocol as wild type. The expression and purification of NTMT1 were performed as described previously<sup>13</sup>. Briefly, the corresponding gene encoding NTMT1 (residues 2–223) was subcloned into pET28a-LIC expression vector and then expressed at 16 °C overnight in the Terrific Broth medium. The protein was purified by Ni-NTA affinity column. The eluted protein was further purified by gel filtration Superdex 200 10/300 (GE Healthcare) pre-equilibrated with 20 mM Tris-HCl pH 7.5, 150 mM NaCl and 0.5 mM TCEP.

**Crystallization and structure determination.** Since NTMT2 purified from *E. coli* contains endogenous methyl donor SAM/SAH, so we did not add any additional SAM/SAH during the crystallization. The NTMT2-SAM complex was crystallized in 20% (w/v) PEG3350 and 0.2 M sodium acetate via sitting drop vapor diffusion by mixing 1 μL protein and 1 μL reservoir solution at 4 °C. To get the ternary complex crystal, the protein was incubated with SPKRIA peptide (from 100 mM stock) at a molar ratio of 1:1.5 for 1 h on ice before setting up the crystallization trial. The crystals of NTMT2 in complex with SPKRIA were obtained in 30% PEG2000 (w/v) and 0.1 M potassium thiocyanate at 18 °C. The crystals were cryo-protected in the reservoir solution supplemented with 20% (v/v) glycerol and flash-frozen in liquid nitrogen.

Diffraction data were collected under cooling in a stream of cold nitrogen gas. Symmetry-related reflection intensities were merged with AIMLESS<sup>29</sup> software. The program PHASER<sup>30</sup> was used for molecular replacement. Atomic models of the crystal structures were interactively rebuilt with COOT. Model geometry was validated with PHENIX.MOLPROBITY<sup>31</sup>. Details for individual crystal structures follow.

Diffraction data for the ternary NTMT2-SAH-SPKRIA complex were collected first on a rotating copper anode source and processed with XDS<sup>32</sup>. The structure was solved by molecular replacement with coordinates from PDB entry 5E1D<sup>13</sup> and automatically rebuilt with ARP/wARP<sup>33</sup>. Further restrained refinement in REFMAC5<sup>34</sup> and interactive rebuilding of the model were performed against another data set that was collected at the Canadian Light Source beamline 08ID<sup>35</sup> and processed with HKL-3000<sup>36</sup>. Atomic anisotropic displacement parameters were analyzed on the PAVARTI<sup>37</sup> server.

Diffraction data for the SAM-NTMT2 complex were collected on a rotating copper anode and processed with XDS. Coordinates from the ternary complex model were used for molecular replacement. Restrained refinement of the atomic model was performed with REFMAC5 and BUSTER<sup>38</sup>.

**Isothermal titration calorimetry.** ITC measurements were performed in 20 mM Tris-HCl pH 7.5, 150 mM NaCl and 0.5 mM TCEP at 25 °C using MicroCal VP-ITC instrument. The final concentrations of protein and peptide were 50–80 μM and 0.6–1.2 mM, respectively. The peptide was titrated into the protein solution with 26 injections of 10 μL each. Injections were spaced 180 s with a reference power of 15 μcal s<sup>-1</sup>. The ITC data were processed using Origin software.

**Methyltransferase activity assays.** A mixture (36 μL) of 2 μM enzyme (NTMT1 or NTMT2), 25 mM Tris-HCl pH 7.5, 50 mM KCl and 200 μM SAM was incubated at 37 °C for 5 min. 4 μL of 100 μM peptide substrate (final concentration 10

μM) was then added to initiate the reaction. The reaction was incubated overnight and then stopped with quenching solution (20 mM NH<sub>4</sub>H<sub>2</sub>PO<sub>4</sub>, 0.4% (v/v) TFA in 1:1 acetonitrile/water). Reaction mixtures were analyzed with an Applied Biosystems Voyager matrix-assisted laser desorption/ionization time-of-flight mass spectrometer. Processing of all spectra results in Data Explorer included application of a noise filter (correlation factor of 1.0) and a baseline correction. Areas of the monoisotopic peaks for all relevant species were combined to obtain a total area for each sample. The fraction of each methylation state was calculated by summing the areas of all the monoisotopic peaks for that state ([M + H]<sup>+</sup>, [M + Na]<sup>+</sup>, [M + K]<sup>+</sup>, [M - H + 2Na]<sup>+</sup>, [M - H + 2K]<sup>+</sup>, and [M - H + Na + K]<sup>+</sup>) and then dividing by the total area.

## Data availability

The datasets generated during the current study are available from the corresponding author on reasonable request. The atomic coordinates and structure factors have been deposited in the RCSB Protein Data Bank. The accession codes for the NTMT2-SAM and NTMT2-SAH-SPKRIA are 5UBB and 6DUB, respectively.

Received: 30 May 2018 Accepted: 15 October 2018

Published online: 02 November 2018

## References

- Zhang, Y. & Reinberg, D. Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails. *Genes Dev.* **15**, 2343–2360 (2001).
- Greer, E. L. & Shi, Y. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* **13**, 343–357 (2012).
- Brosius, J. & Chen, R. The primary structure of protein L16 located at the peptidyltransferase center of *Escherichia coli* ribosomes. *FEBS Lett.* **68**, 105–109 (1976).
- Stock, A., Clarke, S., Clarke, C. & Stock, J. N-terminal methylation of proteins: structure, function and specificity. *FEBS Lett.* **220**, 8–14 (1987).
- Webb, K. J., Lipson, R. S., Al-Hadid, Q., Whitelegge, J. P. & Clarke, S. G. Identification of protein N-terminal methyltransferases in yeast and humans. *Biochemistry* **49**, 5225–5235 (2010).
- Kimura, Y. et al. N-Terminal methylation of proteasome subunit Rpt1 in yeast. *Proteomics* **13**, 3167–3174 (2013).
- Villar-Garea, A. et al. Developmental regulation of N-terminal H2B methylation in *Drosophila melanogaster*. *Nucleic Acids Res.* **40**, 1536–1549 (2012).
- Chen, T. et al. N-terminal alpha-methylation of RCC1 is necessary for stable chromatin association and normal mitosis. *Nat. Cell Biol.* **9**, 596–603 (2007).
- Cai, Q. et al. alpha-N-methylation of damaged DNA-binding protein 2 (DDB2) and its function in nucleotide excision repair. *J. Biol. Chem.* **289**, 16046–16056 (2014).
- Dai, X. et al. Identification of novel alpha-n-methylation of CENP-B that regulates its binding to the centromeric DNA. *J. Proteome Res.* **12**, 4167–4175 (2013).
- Sathyan, K. M., Fachinetti, D. & Foltz, D. R. alpha-amino trimethylation of CENP-A by NRMT is required for full recruitment of the centromere. *Nat. Commun.* **8**, 14678 (2017).
- Tooley, C. E. et al. NRMT is an alpha-N-methyltransferase that methylates RCC1 and retinoblastoma protein. *Nature* **466**, 1125–1128 (2010).
- Dong, C. et al. Structural basis for substrate recognition by the human N-terminal methyltransferase 1. *Genes Dev.* **29**, 2343–2348 (2015).
- Dai, X., Rulten, S. L., You, C., Caldecott, K. W. & Wang, Y. Identification and functional characterizations of N-terminal alpha-N-methylation and phosphorylation of serine 461 in human poly(ADP-ribose) polymerase 3. *J. Proteome Res.* **14**, 2575–2582 (2015).
- Bailey, A. O. et al. Posttranslational modification of CENP-A influences the conformation of centromeric chromatin. *Proc. Natl Acad. Sci. USA* **110**, 11827–11832 (2013).
- Bonsignore, L. A., Butler, J. S., Klinge, C. M. & Schaner Tooley, C. E. Loss of the N-terminal methyltransferase NRMT1 increases sensitivity to DNA damage and promotes mammary oncogenesis. *Oncotarget* **6**, 12248–12263 (2015).
- Bonsignore, L. A. et al. NRMT1 knockout mice exhibit phenotypes associated with impaired DNA repair and premature aging. *Mech. Ageing Dev.* **146–148**, 42–52 (2015).
- Petkowski, J. J. et al. NRMT2 is an N-terminal monomethylase that primes for its homologue NRMT1. *Biochem. J.* **456**, 453–462 (2013).
- Richardson, S. L. et al. A direct, ratiometric, and quantitative MALDI-MS assay for protein methyltransferases and acetyltransferases. *Anal. Biochem.* **478**, 59–64 (2015).



20. Richardson, S. L. et al. Kinetic mechanism of protein N-terminal methyltransferase 1. *J. Biol. Chem.* <https://doi.org/10.1074/jbc.M114.626846> (2015).
21. Wu, R., Yue, Y., Zheng, X. & Li, H. Molecular basis for histone N-terminal methylation by NRMT1. *Genes Dev.* **29**, 2337–2342 (2015).
22. Schubert, H. L., Blumenthal, R. M. & Cheng, X. Many paths to methyltransferase: a chronicle of convergence. *Trends Biochem. Sci.* **28**, 329–335 (2003).
23. Lovell, S. C., Word, J. M., Richardson, J. S. & Richardson, D. C. The penultimate rotamer library. *Proteins* **40**, 389–408 (2000).
24. Xiao, B. et al. Structure and catalytic mechanism of the human histone methyltransferase SET7/9. *Nature* **421**, 652–656 (2003).
25. Zhang, X. et al. Structural basis for the product specificity of histone lysine methyltransferases. *Mol. Cell* **12**, 177–185 (2003).
26. Couture, J. F., Dirk, L. M., Brunzelle, J. S., Houtz, R. L. & Trievel, R. C. Structural origins for the product specificity of SET domain protein methyltransferases. *Proc. Natl Acad. Sci. USA* **105**, 20659–20664 (2008).
27. Shields, K. M. et al. Select human cancer mutants of NRMT1 alter its catalytic activity and decrease N-terminal trimethylation. *Protein Sci.* **26**, 1639–1652 (2017).
28. Fagerberg, L. et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell. Proteom.* **13**, 397–406 (2014).
29. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D* **69**, 1204–1214 (2013).
30. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
31. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
32. Kabsch, W. Xds. *Acta Crystallogr. D* **66**, 125–132 (2010).
33. Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3**, 1171–1179 (2008).
34. Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **67**, 355–367 (2011).
35. Grochulski, P., Fodje, M. N., Gorin, J., Labiuk, S. L. & Berg, R. Beamline 08ID-1, the prime beamline of the Canadian Macromolecular Crystallography Facility. *J. Synchrotron Radiat.* **18**, 681–684 (2011).
36. Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr. D* **62**, 859–866 (2006).
37. Zucker, F., Champ, P. C. & Merritt, E. A. Validation of crystallographic models containing TLS or other descriptions of anisotropy. *Acta Crystallogr. D* **66**, 889–900 (2010).
38. Smart, O. S. et al. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr. D* **68**, 368–380 (2012).

## Acknowledgements

The SGC is a registered charity (number 1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman

Institute for Innovation, Genome Canada through Ontario Genomics Institute [OGI-055], Innovative Medicines Initiative (EU/EFPIA) [ULTRA-DD grant no. 115766], Janssen, Merck KGaA, Darmstadt, Germany, MSD, Novartis Pharma AG, Ontario Ministry of Research, Innovation and Science (MRIS), Pfizer, São Paulo Research Foundation-FAPESP, Takeda, and Wellcome. This work is also supported by a NSERC grant RGPIN-2016-06300 (J.M.). The authors acknowledge the support from NIH grants R01GM117275 (R.H.), U01CA214649 (R.H.), and P30 CA023168 (Purdue University Center for Cancer Research). Research described in this paper was performed using beamline 08ID-1 at the Canadian Light Source, which is supported by the Canada Foundation for Innovation, Natural Sciences and Engineering Research Council of Canada, the University of Saskatchewan, the Government of Saskatchewan, Western Economic Diversification Canada, the National Research Council Canada, and the Canadian Institutes of Health Research.

## Author contributions

C.D., R.H., J.M. conceived the project. C.D. performed protein purification and crystallization experiments. C.D. and G.D. performed enzyme assays. G.D. carried out MS methylation analysis. L.Z. and Y.L. assisted with protein purification. C.D., W.T. determined crystal structures. C.D., G.D., L.L., R.H., and J.M. wrote the manuscript. All authors discussed the results and contributed to the final manuscript.

## Additional information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s42003-018-0196-2>.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018