Structural basis for substrate recognition by the human N-terminal methyltransferase 1

Cheng Dong,[1,5] Yunfei Mao,[2,3,5] Wolfram Tempel,[1] Su Qin,[1] Li Li,[1] Peter Loppnau,[1] Rong Huang,[2,3] and Jinrong Min[1,4]

[1]Structural Genomics Consortium, University of Toronto, Toronto, Ontaria M5G 1L7, Canada; [2]Department of Medicinal Chemistry, [3]The Institute for Structural Biology, Drug Discovery and Development, Virginia Commonwealth University, Richmond, Virginia 23219, USA; [4]Department of Physiology, University of Toronto, Toronto, Ontario M5S 1A8, Canada

α-N-terminal methylation represents a highly conserved and prevalent post-translational modification, yet its biological function has remained largely speculative. The recent discovery of α-N-terminal methyltransferase 1 (NTMT1) and its physiological substrates propels the elucidation of a general role of α-N-terminal methylation in mediating DNA-binding ability of the modified proteins. The phenotypes, observed from both NTMT1 knockdown in breast cancer cell lines and knockout mouse models, suggest the potential involvement of α-N-terminal methylation in DNA damage response and cancer development. In this study, we report the first crystal structures of human NTMT1 in complex with cofactor S-adenosyl-L-homocysteine (SAH) and six substrate peptides, respectively, and reveal that NTMT1 contains two characteristic structural elements (a β hairpin and an N-terminal extension) that contribute to its substrate specificity. Our complex structures, coupled with mutagenesis, binding, and enzymatic studies, also present the key elements involved in locking the consensus substrate motif XPK (X indicates any residue type other than D/E) into the catalytic pocket for α-N-terminal methylation and explain why NTMT1 prefers an XPK sequence motif. We propose a catalytic mechanism for α-N-terminal methylation. Overall, this study gives us the first glimpse of the molecular mechanism of α-N-terminal methylation and potentially contributes to the advent of therapeutic agents for human diseases associated with deregulated α-N-terminal methylation.

Supplemental material is available for this article.

Received August 18, 2015; revised version accepted October 14, 2015.

Post-translational modifications (PTMs) of proteins during or after their synthesis contribute to the functional diversity and dynamics of the human proteome through regulating protein three-dimensional (3D) structure, stability, activity, localization, and interaction with other cellular molecules, including proteins and nucleic acids. PTMs can occur on both the amino acid side chains and backbones of proteins, often by covalent attachment of functional groups such as the methyl group. α-N-terminal methylation is a ubiquitously present PTM that involves the methylation of the α-amino group of the newly exposed N-terminal amino acid following cleavage of the initiating methionine. Although α-N-terminal methylation was first identified almost 40 years ago (Brosius and Chen 1976; Wittmann-Liebold and Pannenbecker 1976; Stock et al. 1987) and is conserved from prokaryotes to humans, its function remains elusive and comparatively underexplored. α-N-terminal methylation has been hypothesized to regulate protein stability via N-end rule pathways or mediate protein–protein interactions (Tooley and Schaner Tooley 2014). Intriguingly, α-N-terminal methylation was recently found to mediate protein–DNA interactions between chromatin and regulator of chromatin condensation (RCC1) (Chen et al. 2007; Hitakomate et al. 2010), potentially through the positively charged and exhaustively methylated N terminus of RCC1. RCC1 is the only known guanine nucleotide exchange factor for the Ran GTPase, which plays indispensable roles in nucleocytoplasmic transport, mitosis, and nuclear envelope assembly (Chen et al. 2007). Notably, α-N-terminal methyltransferase 1 (NTMT1/NRMT1/METTL11A) catalyzes the N-terminal methylation of RCC1, representing a novel regulatory mechanism for the Ran GTPase activity (Tooley et al. 2010; Webb et al. 2010). Abrogation of N-terminal methylation of RCC1 by mutations or knockdown of NTMT1 significantly diminishes the binding ability and association of RCC1 to chromatin and causes spindle pole defects and aberrant mitosis, underlining the functional significance of NTMT1-mediated α-N-methylation in mitotic spindle assembly and chromosome segregation (Chen et al. 2007).

In addition to RCC1, several other proteins are subject to α-N-terminal methylation (Webb et al. 2010), including retinoblastoma protein RB1 and SET (Tooley et al. 2010), damaged DNA-binding protein 2 (DDB2) (Cai et al. 2014), centromere H3 variants CENP-A/B (Bailey et al. 2013; Dai et al. 2015), *Drosophila* H2B (Villar-Garea et al. 2012), and poly(ADP-ribose) polymerase 3 (Dai et al. 2015), while data bank analysis of NTMT1/2's consensus sequence predicts the existence of potentially >300 targets for α-N-terminal methylation (Tooley and Schaner Tooley 2014). NTMT1 is responsible for the α-N-terminal methylation of DDB2 in response to the generation of UV-induced cyclobutane pyrimidine dimers (Cai et al. 2014). This methylation of DDB2 promotes its recruitment to form foci at the sites of DNA damage and facilitates nucleotide excision repair, possibly indicating a role of NTMT1 in the DNA damage response (DDR) network (Cai et al. 2014). Moreover, knockdown of NTMT1 leads

[*Keywords*: α-N-terminal methylation; methyltransferase; NTMT1; crystal structure]
[5]**These authors contributed equally to this work.**
Corresponding authors: jr.min@utoronto.ca, rhuang@vcu.edu
Article published online ahead of print. Article and publication date are online at http://www.genesdev.org/cgi/doi/10.1101/gad.270611.115.

to hypersensitivity of breast cancer cell lines to both etoposide and γ-irradiation treatments, further suggesting NTMT1 as a component of DDR (Bonsignore et al. 2015a). Interestingly, NTMT1 knockout mice suffer a high mortality rate shortly after birth and exhibit premature aging and phenotypes characteristic of mouse models deficient for DDR molecules, indicating the biological significance of NTMT1 in vivo (Bonsignore et al. 2015b). Additionally, NTMT1-mediated α-N-methylation of CENP-B promotes its binding to centromeric DNA (Dai et al. 2015). Overall, the above findings highlight a general and important role of NTMT1-mediated α-N-methylation in facilitating interactions between methylated target proteins and DNA.

Although some progress has been made in the field of α-N-terminal methylation, many questions remain to be answered. For instance, why does NTMT1 specifically carry out the α-N-terminal methylation, and what is the catalytic mechanism? Given that the majority of known physiological substrates of NTMT1 contains an XPK (X = S/P/A/G) N-terminal sequence, what is the structural basis for the requirement of this consensus sequence, and is there any residue tolerance along the N-terminal consensus sequence? In an effort to address these questions, we determined the X-ray crystal structures of human NTMT1 in ternary complexes with its cofactor product (S-adenosyl-L-homocysteine [SAH]) and six different hexapeptides as substrates, including the very N-terminal fragment of RCC1 and its mutant peptides. We deduced the molecular mechanism of NTMT1-mediated α-N-methylation on its physiological substrate, RCC1, based on data obtained from structure-based mutagenesis as well as enzymatic characterizations.

Results and Discussion

Overall structure of the NTMT1 ternary complexes

NTMT1 is an α-N-terminal methyltransferase, highly conserved from yeast to humans (Fig. 1; Webb et al. 2010). So far, all known substrates of NTMT1 contain the N-terminal consensus sequence XPK (X = S/P/A/G) (Fig. 2A), although NTMT1 can also methylate peptides with X being F, Y, C, M, K, R, N, Q, or H in vitro (Petkowski et al. 2012). In order to understand the substrate specificity of NTMT1, we determined the crystal structures of the full-length human NTMT1 in complex with its cofactor (SAH) and a peptide derived from either human (sequence: SPKRIA) or mouse (sequence: PPKRIA) RCC1. Furthermore, we also generated crystals of NTMT1 in complex with SAH and either the RPK or YPK peptide, which can be efficiently methylated by NTMT1 in vitro (Tooley et al. 2010). These crystal structures reveal that the human NTMT1 folds as a single domain (Fig. 2B), and the root mean square deviation (RMSD) values between these structures and the NTMT1-SAH binary structure (Protein Data Bank [PDB]: 2EX4) range from 0.2 to 0.4 Å for the Cα atoms of NTMT1. The crystal diffraction data and refinement statistics for all of these structures are shown in Supplemental Table S1.

NTMT1 includes a typical methyltransferase Rossmann fold that consists of a seven-strand β sheet and five α helixes, of which two α helixes (α6 and α7) pack on one side of the β sheet, and the other three α helixes (α3, α4, and α5) pack on the other side of the β sheet (Fig. 2B). In addition to the highly conserved Rossmann fold, NTMT1 contains two unique structural elements distinct from other methyltransferases: a β hairpin inserted between strand β5 and helix α7 and an N-terminal extension consisting of two α helixes (α1 and α2) and one 3_{10} helix (η1) (Figs. 1, 2B). These two unique structural elements are extensively involved in substrate binding (Fig. 2B; Supplemental Fig. S1), suggesting their potential contributions to substrate specificity.

SAH is bound to NTMT1 in an extended conformation (Fig. 2C), similar to SAH binding in other Rossmann fold methyltransferases such as arginine methyltransferases and DOT1 lysine methyltransferase (Min et al. 2003; Sawada et al. 2004). Likewise, NTMT1 uses a conserved DxGxGxG motif to surround the ribosyl and methionyl moieties of SAH. Furthermore, the carboxylate moiety of SAH forms a salt bridge interaction with the highly conserved Arg74, and the ribosyl group stacks with the indole ring of Trp20. In addition, the adenine moiety of SAH is flanked by the hydrophobic side chains of Ile92 and Val137 and interacts with the main chain amide group of Leu119 and the side chain of Gln120 through hydrogen bonding (Fig. 2C).

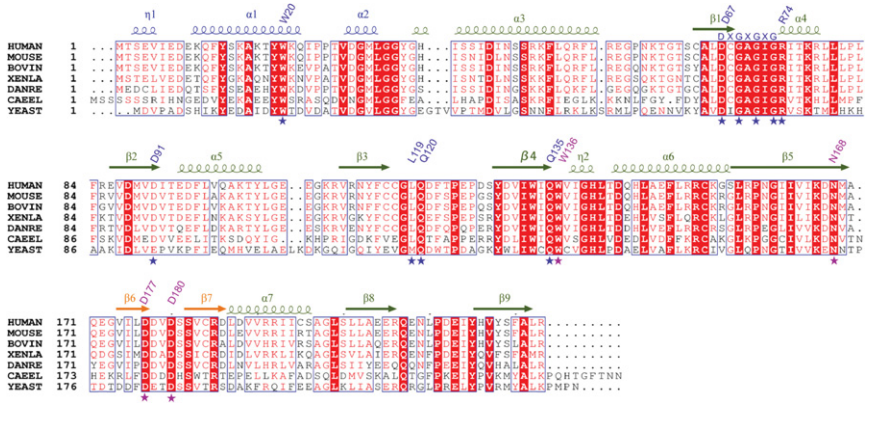

Figure 1. Sequence alignment of NTMT1 homologs from different species. (Human) *Homo sapiens* (UniProt: Q9BV86); (mouse) *Mus musculus* (UniProt: Q8R2U4); (bovin) *Bos taurus* (UniProt: Q2T9N3); (xenla) *Xenopus laevis* (UniProt: Q4KLE6); (danre) *Danio rerio* (UniProt: Q6NWX7); (caeel) *Caenorhabditis elegans* (UniProt: Q9N4D9); (yeast) *Saccharomyces cerevisiae* (UniProt: P38340). The secondary structure elements of human NTMT1 are indicated *above* with colors corresponding to Figure 2B. Identical residues are marked by red background, and conserved residues are colored in red. Residues involved in the interactions with SAH and substrate are numbered and indicated with blue and purple stars, respectively.

Structural basis for the specific N-terminal methylation of a consensus motif, XPK, by NTMT1

Since hexapeptides composed of the first six residues of RCC1 are recognized by NTMT1 for α-N-terminal methylation in a fashion similar that of to longer peptides (Richardson et al. 2015), we used the RCC1-based hexapeptides for crystallization and resolved all of their residues in our complex structures. Our structural

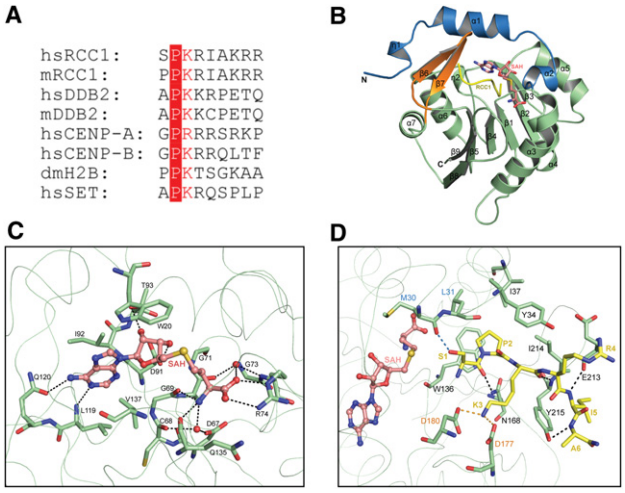

Figure 2. Crystal structure of NTMT1 in complex with SAH and hRCC1-6 (SPKRIA). (*A*) Sequence alignment of known NTMT1 substrates. (*B*) Ribbon diagram of NTMT1 in complex with cofactor SAH and substrate peptide hRCC1-6 (SPKRIA). The Rossmann fold is shown as a pale green cartoon. The N-terminal extended helixes and the inserted hairpin are colored in blue and orange, respectively. SAH is shown in stick mode with carbon atoms in salmon, and hRCC1-6 is shown as a yellow cartoon. (*C*) Close-up view of the coordination of the cofactor SAH in the active site of NTMT1. SAH is shown as ball and stick in salmon. Residues contributing to hydrogen bonding and van der Waals interactions are labeled; the hydrogen bonds are displayed as dashed lines, and mediating water molecules are shown as red spheres. (*D*) Close-up view of substrate binding (SPKRIA) in the active site. The cofactor SAH is shown in salmon ball and stick mode, and the interacting residues of NTMT1 and the substrate peptides are shown as pale-green and yellow sticks, respectively. The hydrogen bonds are displayed as dashed lines, and the residues of NTMT1 are labeled in the same color scheme as in *B*. Molecular representations were prepared with PyMOL (https://www.pymol.org).

data reveal that all of the peptides bind in a similar mode to an extensively negatively charged channel of NTMT1, comprised of highly conserved residues among different species (Fig. 3A,B; Supplemental Fig. S2A). Furthermore, this negatively charged channel is connected to the cofactor SAH-binding pocket. Moreover, the substrate peptide is inserted into this channel with the α-N-amino group of the substrate pointing toward the putative methyl group of SAM to accept the methyl transfer. Interestingly, the cofactor SAH/SAM is completely buried inside the NTMT1 ternary complex. These unique features of NTMT1 are in striking contrast to lysine/arginine methyltransferases, such as PRMT5, in which a substrate-binding groove is formed on the surface of these enzymes, and the target histone arginine H4R3 is inserted into a narrow channel to reach the SAM-binding pocket (Fig. 3C; Antonysamy et al. 2012). Therefore, our crystal structures of these NTMT1 ternary complexes described above explain the specificity of NTMT1 in catalyzing α-N-terminal methylation.

In the NTMT1–SAH–hRCC1 (SPKRIA) complex, the carbonyl oxygen of the first residue serine (S1) of hRCC1 forms a main chain hydrogen bond with the side chain carboxamide of Asn168 (Fig. 2D; Supplemental Fig. S1). This hydrogen bond is conserved in other peptide complex structures (Supplemental Figs. S2, S3), accompanied by the conservation of this Asn168 residue of NTMT1 in other species (Fig. 1A). Given that mutations of Asn168 to Lys or Ala were proposed to cause steric clashes or disrupt the hydrogen bond (Tooley et al. 2010), we used site-directed mutagenesis to change the Asn168 residue of NTMT1 to Lys. As expected, the N168K mutant of NTMT1 exhibited ~36-fold reduction in K_m value as well as about two-fold reduction in K_{cat} value compared with those of wild-type NTMT1 (Table 1), highlighting the importance of this hydrogen bond in the interaction of NTMT1 with the peptide substrate.

On the other hand, the side chain of S1 is also able to form a hydrogen bond with the main chain of Met30, but this interaction is not conserved in other complex structures. Interestingly, our structures reveal a spacious binding pocket for the side chain of S1 (Fig. 3A), providing an explanation for the observed tolerance of NTMT1 to a variety of residue types at the first position (Fig. 2A; Table 2; Tooley et al. 2010; Webb et al. 2010). Tooley et al. (2010) previously reported that the first residue of the substrates could be replaced by other residues types, including the positively charged residue R and the bulky residue Y but not the negatively charged residues D/E or hydrophobic residues L/I/W (Tooley et al. 2010).

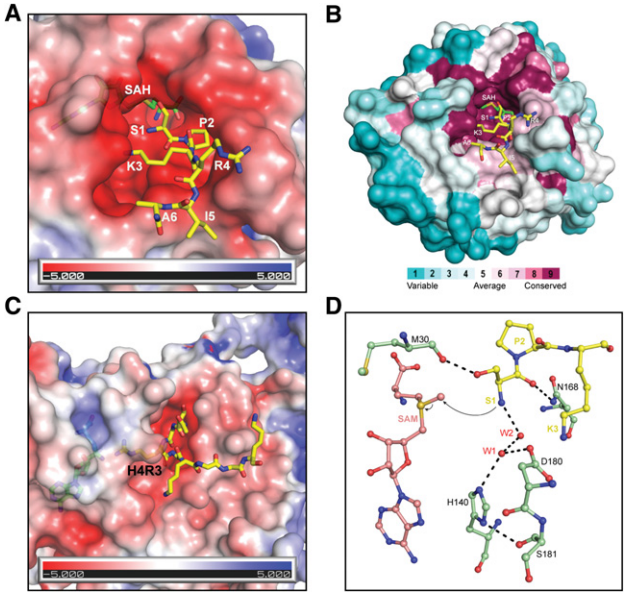

Figure 3. NTMT1 harbors a conserved active site for α-N-terminal methylation. (*A*) NTMT1 is shown in electrostatic potential surface representation, and the substrate (NTMT1–SAH–hRCC1 [SPKRIA]) is shown as sticks. Adaptive Poisson-Boltzmann Solver (APBS) (Baker et al. 2001) and PyMOL were used for coordinate processing, charge calculation, and visualization, respectively. (*B*) Conservation analysis of NTMT1 using the Web server ConSurf (Ashkenazy et al. 2010). The conservation scores are mapped to the molecular surface of NTMT1. The highly conserved residues are colored in red, and the most variable residues are colored in cyan. The cofactor SAH and substrate RCC1 are shown in green and yellow sticks, respectively. (*C*) The catalytic pocket of PRMT5. PRMT5 (PDB: 4GQB) (Antonysamy et al. 2012) is shown in electrostatic surface representation, and the cofactor SAM analog A9145C and the histone H4R3 peptide are shown as green and yellow sticks, respectively. (*D*) Catalytic mechanism for the α-N-terminal methylation. The methyl group of SAM was computationally modeled in the active site and is shown in salmon ball and stick. For clarity, only the first three residues of substrates are displayed. Putative catalytic water molecules are shown as red spheres and labeled W1 and W2. In the crystallographic models, corresponding positions are occupied by glycerol, other cosolutes, or water.

Table 1. *Catalytic activity of human NTMT1 and its mutants*

Enzyme	K_m	k_{cat}	k_{cat}/K_m
Wild type	7.3 μM ± 0.7 μM	0.10 min^{-1} ± 0.01 min^{-1}	1.4×10^{-2} μM^{-1} min^{-1}
N168K	263 μM ± 141 μM	0.05 min^{-1} ± 0.01 min^{-1}	1.9×10^{-4} μM^{-1} min^{-1}
W136F	>200 μM	>0.04 min^{-1} ± 0.01 min^{-1}	—
W136I	ND	ND	ND
D180K	ND	ND	ND
D180Y	ND	ND	ND
H140K	ND	ND	ND

(ND) No detectable activity at 250 μM peptide and a saturating amount of SAM.

Accordingly, our structures show that the S1-binding pocket is large enough to accommodate the aromatic residue Y in the NTMT1–SAH–YPK structure (Supplemental Fig. S2A,C), which is further confirmed by our binding and enzymatic data (Table 2; Supplemental Figs. S4, S5). Our kinetic and binding results also indicated that R could be recognized and methylated by NTMT1 (Table 2). The negative charge inside the substrate-binding channel provides an intrinsic barrier for the S1 pocket against either the negatively charged residues D/E or hydrophobic residues L/I/W. We failed to detect any activity of NTMT1 to DPKRIA (Table 2). Although NTMT1 could still bind to WPKRIA and LPKRIA, their binding and enzymatic activities were much weaker (Table 2). Interestingly, we found that the isothermal titration calorimetry (ITC) curves for some peptides exhibit apparent bimodal binding. It was revealed by further characterization with mass spectrometry experiments and ITC assays that NTMT1 purified from *Escherichia coli* contains endogenous methyl donor SAM, which could methylate the substrate peptides during ITC experiments even without adding SAM (Supplemental Figs. S5, S6). It has been well documented that some methytransferases could copurify with endogenous SAM, such as in the case of Dot1L (Min et al. 2003); therefore, one possible explanation is that the bimodal ITC binding curves arise from a mixture of the peptide substrate and the corresponding methylated products, both of which could bind to NTMT1 with respective affinities. The PPK peptide is the most optimal substrate that can be efficiently methylated by NTMT1, which therefore has the most obvious bimodal binding curve. Taken together, the first residue within the consensus sequence of the NTMT1 substrates is anchored through a hydrogen bond with the conserved Asn168 of NTMT1 in a spacious binding pocket, which exposes the substrate's reactive α-amino group to the putative methyl donor SAM in the complex structures, and this very N-terminal residue can tolerate most residue substitutions except the negatively charged residues D and E.

The second residue (P2) of hRCC1 is accommodated in a pocket formed by a few hydrophobic residues of NTMT1, such as Leu31, Ile37, and Ile214 (Fig. 2D). Interestingly, the P2 residue also forms a stacking interaction with the indole of Trp136, reminiscent of the proline-rich sequence binding to the WW and SH3 domains (Fig. 2D; Macias et al. 2002). When we substituted Trp136 with either Phe or Ile, the enzymatic activities of these mutants were either significantly diminished or undetectable (Table 1). Our enzymatic studies also indicate that the peptides with the second residue mutated to either Q, I, E, or S could not be methylated and were incapable of binding to NTMT1 (Table 2). The above results also support the previously reported finding that P2 cannot tolerate any substitutions (Webb et al. 2010).

The third residue (K3) of hRCC1 forms two key hydrogen bonds with side chains of Asp177 and Asp180 of NTMT1 (Fig. 2D). When we mutated these two residues, D180A/K did not exhibit any detectable binding, and D177A displayed reduced binding affinity in our ITC experiments (Table 1; Supplemental Fig. S7), which further confirm the importance of these hydrogen bonds and are in concordance with previous mutation studies (Petkowski et al. 2012). It has also been reported that the residue in the third position does not tolerate substitution to other residues except arginine (Tooley et al. 2010). Based on our complex structures, substitution of the lysine with

Table 2. *Effect of substrate residue substitution on enzymatic activity and binding affinity of NTMT1*

Peptide	Sequence	K_m	k_{cat}	k_{cat}/K_m	ITC/$K_{d1/d2}$
RCC1-6-1S	SPKRIA	7.9 μM ± 0.7 μM	0.07 min^{-1} ± 0.001 min^{-1}	9.0×10^{-3} μM^{-1} min^{-1}	14.3 μM ± 1.3 μM/0.8 μM + 0.1 μM
RCC1-6-1P	PPKRIA	0.3 μM ± 0.04 μM	0.11 min^{-1} ± 0.006 min^{-1}	3.8×10^{-1} μM^{-1} min^{-1}	189 nM ± 85 nM/2.8 nM ± 1.3 nM
RCC1-6-1Y	YPKRIA	1.6 μM ± 0.3 μM	0.04 min^{-1} ± 0.003 min^{-1}	2.5×10^{-2} μM^{-1} min^{-1}	4.1 μM ± 0.6 μM/0.1 μM ± 0.08 μM
RCC1-6-1R	RPKRIA	4.0 μM ± 0.5 μM	0.09 min^{-1} ± 0.003 min^{-1}	2.3×10^{-2} μM^{-1} min^{-1}	3.7 μM ± 0.5/0.3 ± 0.05 μM
RCC1-6-1W	WPKRIA	126 μM ± 7 μM	0.10 min^{-1} ± 0.002 min^{-1}	7.9×10^{-4} μM^{-1} min^{-1}	47 μM ± 3 μM
RCC1-6-1L	LPKRIA	54 μM ± 6 μM	0.11 min^{-1} ± 0.004 min^{-1}	2.0×10^{-3} μM^{-1} min^{-1}	48 μM ± 5 μM
RCC1-6-1D	DPKRIA	ND	ND	ND	NB
RCC1-6-2I	SIKRIA	ND	ND	ND	NB
RCC1-6-2Q	SQKRIA	ND	ND	ND	NB
RCC1-6-2E	SEKRIA	ND	ND	ND	NB
RCC1-6-2S	SSKRIA	ND	ND	ND	NB

(ND) No detectable activity at 250 μM peptide and a saturating amount of SAM; (NB) no detectable binding at 50 μM NTMT1 and 1 mM peptide by ITC.

any other residues except arginine would disrupt these two hydrogen bonds. Consistently, all known physiologically α-N-terminally methylated proteins contain the XPK motif except the human centromeric H3 variant CENP-A, which harbors an XPR motif.

The fourth residue of hRCC1 is located adjacent to the entrance of the negatively charged substrate-binding channel (Fig. 3A). Either an arginine or a lysine residue is present at this position in most known substrates of NTMT1 (Fig. 2A), presumably because a positively charged residue would be favored in this negatively charged environment, although the arginine residue does not appear to form strong salt bridge interactions with NTMT1 in our crystal structures. The fifth and sixth residues, such as I5 and A6 in our case, only form some main chain hydrogen bonds with NTMT1, which explains the observed diversity of residue types at these positions (Fig. 2A).

Catalytic mechanism

The structure of the catalytic pocket of NTMT1 suggests an S_N2 mechanism for the α-amino methylation (Coward and Sweet 1971), as the methyl-accepting nitrogen of S1 or P1 and the putative methyl group would come within a proximity of ~4 Å (Fig. 3D; Supplemental Fig. S8). This result is in agreement with our previous work showing that NTMT1 catalyzes methylation via a random-ordered Bi–Bi mechanism, which involves the formation of a ternary complex of both substrates bound to the NTMT1, with either substrate binding first to NTMT1 (Richardson et al. 2015). In addition, the distance between the sulfonium ion and the α-nitrogen atom of the substrate is ~4.7 Å in our ternary complexes, which is longer than the distance of 3.6 Å reported in a previous docking study (Tooley et al. 2010). This longer distance may support the feasibility of using a triazole group to produce a potent and selective NTMT1 inhibitor with a K_i of 203 nM via click chemistry (Zhang et al. 2015). Our structures show that the cofactor product SAH is buried deeply in the SAM-binding site, participating in hydrogen-bonding networks and hydrophobic interactions (Fig. 3A). That might also explain why SAH is a much more potent inhibitor (IC_{50} = 3.3 μM) for NTMT1 than me3-RCC1-6 (IC_{50} > 25 μM), as determined in previous studies (Richardson et al. 2015). Two highly conserved residues, Asp180 and His140, are present near the α-amino group (Figs. 1, 3D) and could mediate methyl transfer via one or more water molecules. Mechanistically, Asp180 and His140 can act as bases to facilitate deprotonation of the target α-N-terminal amino group. Specifically, the Nπ atom of His140, anchored by the carbonyl oxygen atom of Ser181, would ensure a favorable orientation for Nτ to accept a hydrogen bond from a water molecule, which in turn could deprotonate the substrate ammonium group (Fig. 3D). The deprotonated amino group could then attack the methyl group of SAM. In line with the catalytic mechanism proposed above, our enzymatic data show that both the D180K and H140K mutants of NTMT1 are incapable of catalyzing methylation (Table 1).

To provide further insights into the catalytic mechanism of NTMT1, our complex structures with me1-SPKRIA and me1-PPKRIA indicate that monomethylated substrate peptides have essentially the same orientation as unmethylated substrate peptides (Supplemental Fig. S3), suggesting that NTMT1 does not have a significant preference in binding nonmethylated or monomethylated substrates. These results are consistent with the observation that NTMT1-mediated methylation proceeds in a distributive mechanism with comparable catalytic efficiencies for both nonmethylated and monomethylated peptide substrates, as reported previously (Richardson et al. 2015).

In summary, we propose that NTMT1-mediated methylation of the α-amino group of its physiological substrates occurs through a S_N2 reaction mechanism. In particular, the highly conserved Asp180 and His140 of NTMT1 that are located in close proximity to the α-amino group of the substrate would act as general bases to facilitate the deprotonation of the α-amino group and then mediate the methyl transfer from SAM. This study of NTMT1 in complex with its cofactor (SAH) and physiological substrate (RCC1) provides a structural foundation on which the field of α-N methylation can expand. Given that retinoblastoma protein RB1 and SET (Tooley et al. 2010), DDB2 (Cai et al. 2014), centromere H3 variants CENP-A/B (Bailey et al. 2013; Dai et al. 2015), *Drosophila* H2B (Villar-Garea et al. 2012), and poly(ADP-ribose) polymerase 3 (Dai et al. 2015) have been recently reported to be modified through α-N-terminal methylation and that there exists potentially >300 putative substrates harboring the NTMT1/2's consensus sequence (Tooley and Schaner Tooley 2014), the biological functions of α-N-terminal methylation are likely pleiotropic and may not be limited to only the modulation of DNA-binding abilities of the modified proteins. Thus, our proposed model of the recognition of a consensus XPK substrate motif by NTMT1 may shed light on the understanding of NTMT1's substrate specificity and catalytic mechanism in general. While the participation of α-N-terminal methylation in the DDR network and cancer development begins to unfold, our structural analysis of NTMT1 in complex with its cofactor and physiological substrate provides valuable information to promote the elucidation of roles of α-N-terminal methylation in human physiology and pathogenesis as well as potentially aid the development of therapeutic agents against human diseases associated with deregulated α-N-terminal methylation.

Materials and methods

Protein expression and purification

The gene of human NTMT1 (2–223) was amplified and cloned into a modified pET28a-LIC vector to express NTMT1 with a 6xHis tag and a thrombin cleavage site at the N terminus. The recombinant NTMT1 plasmid was transformed into an *E. coli* BL21 (DE3) codon plus RIL strain for induced expression with 0.2 mM IPTG overnight at 16°C. NTMT1 was purified by Ni^{2+} affinity and anion exchange chromatography followed by further purification through Superdex 200 10/300 (GE Healthcare). The buffer for gel filtration contained 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, and 0.5 mM TECP. The peak fractions were collected and concentrated to 37 mg/mL for crystallization assay. The mutant proteins were purified using the same procedure as described above.

Crystallization and structure determination

To crystallize the complex, NTMT1 was first incubated with substrate peptides at a molar ratio of 1:1.5 for 1 h on ice, and the crystals were grown at 277K using the sitting drop vapor diffusion method by mixing 1 μL of complex solution with 1 μL of reservoir consisting of 23%–28% PET3350 and 14%–18% Tacsimate (pH 6.0). The crystals were flash-frozen in liquid nitrogen after soaking in cryo-protectant solutions containing the reservoir solution supplemented with 20% (v/v) glycerol.

Structures were determined by molecule replacement, and the detailed protocol is shown in the Supplemental Material.

ITC

ITC measurements were performed by VP-ITC instrument (MicroCal) at 25°C. The purified NTMT1 were diluted to 50–100 µM using the ITC buffer containing 20 mM Tris-HCl (pH7.5) and 150 mM NaCl. The peptides were dissolved into ITC buffer to yield 100 mM, and the pH value was adjusted to 7.5 with NaOH. The peptides were diluted to 0.75–2 mM with the ITC buffer. Twenty-five to 28 injections of 10 µL of peptide were carried out with a spacing of 240 sec and a reference power of 15 µcal/sec, and the ITC data were analyzed using Origin software. The ITC titration curves are shown in Supplemental Figures S5 and S7.

Methyltransferase activity assays

Rate measurements were performed using a SAH hydrolase-coupled fluorescence assay, which monitors the production of SAH as previously described (Richardson et al. 2015). To determine NTMT1 activity, reactions were initiated by the addition of 10 µL of a peptide in various concentrations to 90 µL of reaction mixtures consisting of 25 mM Tris (pH 7.5), 50 mM KCl, 0.2 µM NTMT1, 10 µM SAH hydrolase, 100 µM SAM, and 15 µM ThioGlo1. Fluorescence changes were monitored using a FlexStation 3 microplate reader (Ex = 370 nm, Em = 500 nm) for 15 min at 37°C. Product formation was calculated as previously described (Richardson et al. 2015). All reactions were performed in triplicates. Initial velocity measurements for all mutants of NTMT1 were performed using hRCC1-6 (SPKRIA) as the peptide substrate. Data for the initial rates were fit to the Michaelis-Menten model using least squares nonlinear regression (curve fit) through GraphPad Prism 5 software (version 5.04).

Data access

The atomic coordinates and structure factors have been deposited in the PDB. The accession codes for the NTMT1–SAH–peptide complexes are 5E1B, 5E1M, 5E1O, 5E1D, 5E2B, and 5E2A, respectively.

Acknowledgments

We thank Hong Wu for providing the original NTMT1 clone, and Brianna D. Mackie for some peptides. The Structural Genomics Consortium is a registered charity (no. 1097737) that receives funds from AbbVie; Bayer Pharma AG; Boehringer Ingelheim; Canada Foundation for Innovation; Eshelman Institute for Innovation; Genome Canada; Innovative Medicines Initiative (EU/EFPIA) (Unrestricted Leveraging of Targets for Research Advancement and Drug Discovery grant no. 115766); Janssen, Merck, and Co.; Novartis Pharma AG; Ontario Ministry of Economic Development and Innovation; Pfizer; São Paulo Research Foundation-FAPESP; Takeda; and the Wellcome Trust. Some diffraction experiments were performed at the Structural Biology Center and Northeastern Collaborative Access Team beam lines at the Advanced Photon Source at Argonne National Laboratory (ANL). ANL is operated by the University of Chicago Argonne, LLC, for the U.S. Department of Energy Office of Biological and Environmental Research under contract DE-AC02-06CH11357. We also thank the Virginia Commonwealth University Presidential Research Quest Fund for financial support to R.H.

References

Antonysamy S, Bonday Z, Campbell RM, Doyle B, Druzina Z, Gheyi T, Han B, Jungheim LN, Qian Y, Rauch C, et al. 2012. Crystal structure of the human PRMT5:MEP50 complex. *Proc Natl Acad Sci* **109:** 17960–17965.

Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N. 2010. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* **38:** W529–W533.

Bailey AO, Panchenko T, Sathyan KM, Petkowski JJ, Pai PJ, Bai DL, Russell DH, Macara IG, Shabanowitz J, Hunt DF, et al. 2013. Posttranslational modification of CENP-A influences the conformation of centromeric chromatin. *Proc Natl Acad Sci* **110:** 11827–11832.

Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci* **98:** 10037–10041.

Bonsignore LA, Butler JS, Klinge CM, Schaner Tooley CE. 2015a. Loss of the N-terminal methyltransferase NRMT1 increases sensitivity to DNA damage and promotes mammary oncogenesis. *Oncotarget* **6:** 12248–12263.

Bonsignore LA, Tooley JG, Van Hoose PM, Wang E, Cheng A, Cole MP, Schaner Tooley CE. 2015b. NRMT1 knockout mice exhibit phenotypes associated with impaired DNA repair and premature aging. *Mech Ageing Dev* **146–148:** 42–52.

Brosius J, Chen R. 1976. The primary structure of protein L16 located at the peptidyltransferase center of *Escherichia coli* ribosomes. *FEBS Lett* **68:** 105–109.

Cai Q, Fu L, Wang Z, Gan N, Dai X, Wang Y. 2014. α-N-methylation of damaged DNA-binding protein 2 (DDB2) and its function in nucleotide excision repair. *J Biol Chem* **289:** 16046–16056.

Chen T, Muratore TL, Schaner-Tooley CE, Shabanowitz J, Hunt DF, Macara IG. 2007. N-terminal α-methylation of RCC1 is necessary for stable chromatin association and normal mitosis. *Nat Cell Biol* **9:** 596–603.

Coward JK, Sweet WD. 1971. Kinetics and mechanism of methyl transfer from sulfonium compounds to various nucleophiles. *J Org Chem* **36:** 2337–2346.

Dai X, Rulten SL, You C, Caldecott KW, Wang Y. 2015. Identification and functional characterizations of N-terminal α-N-methylation and phosphorylation of serine 461 in human poly(ADP-ribose) polymerase 3. *J Proteome Res* **14:** 2575–2582.

Hitakomate E, Hood FE, Sanderson HS, Clarke PR. 2010. The methylated N-terminal tail of RCC1 is required for stabilisation of its interaction with chromatin by Ran in live cells. *BMC Cell Biol* **11:** 43.

Macias MJ, Wiesner S, Sudol M. 2002. WW and SH3 domains, two different scaffolds to recognize proline-rich ligands. *FEBS Lett* **513:** 30–37.

Min J, Feng Q, Li Z, Zhang Y, Xu RM. 2003. Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. *Cell* **112:** 711–723.

Petkowski JJ, Schaner Tooley CE, Anderson LC, Shumilin IA, Balsbaugh JL, Shabanowitz J, Hunt DF, Minor W, Macara IG. 2012. Substrate specificity of mammalian N-terminal α-amino methyltransferase NRMT. *Biochemistry* **51:** 5942–5950.

Richardson SL, Mao Y, Zhang G, Hanjra P, Peterson DL, Huang R. 2015. Kinetic mechanism of protein N-terminal methyltransferase 1. *J Biol Chem* **290:** 11601–11610.

Sawada K, Yang Z, Horton JR, Collins RE, Zhang X, Cheng X. 2004. Structure of the conserved core of the yeast Dot1p, a nucleosomal histone H3 lysine 79 methyltransferase. *J Biol Chem* **279:** 43296–43306.

Stock A, Clarke S, Clarke C, Stock J. 1987. N-terminal methylation of proteins: structure, function and specificity. *FEBS Lett* **220:** 8–14.

Tooley CE, Petkowski JJ, Muratore-Schroeder TL, Balsbaugh JL, Shabanowitz J, Sabat M, Minor W, Hunt DF, Macara IG. 2010. NRMT is an α-N-methyltransferase that methylates RCC1 and retinoblastoma protein. *Nature* **466:** 1125–1128.

Tooley JG, Schaner Tooley CE. 2014. New roles for old modifications: emerging roles of N-terminal post-translational modifications in development and disease. *Protein Sci* **23:** 1641–1649.

Villar-Garea A, Forne I, Vetter I, Kremmer E, Thomae A, Imhof A. 2012. Developmental regulation of N-terminal H2B methylation in *Drosophila melanogaster*. *Nucleic Acids Res* **40:** 1536–1549.

Webb KJ, Lipson RS, Al-Hadid Q, Whitelegge JP, Clarke SG. 2010. Identification of protein N-terminal methyltransferases in yeast and humans. *Biochemistry* **49:** 5225–5235.

Wittmann-Liebold B, Pannenbecker R. 1976. Primary structure of protein L33 from the large subunit of the *Escherichia coli* ribosome. *FEBS Lett* **68:** 115–118.

Zhang G, Richardson SL, Mao Y, Huang R. 2015. Design, synthesis, and kinetic analysis of potent protein N-terminal methyltransferase 1 inhibitors. *Org Biomol Chem* **13:** 4149–4154.