Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Towards the structural characterization of the human methyltransferome

Michael K Fenwick and Steven E Ealick

Approximately 2000 structures of methyltransferases (MTases) are currently available, displaying fifteen different folds for binding a methyl donor and providing molecular level insight into nearly half the human methyltransferome. Several MTases involved in gene expression and regulation are catalytically inefficient when isolated, and their catalytic domains often show inhibitory active site architectures. Recently reported structures of complexes that more closely reflect biological context have begun to reveal the structural basis of activation. DNA and particular histone MTases are allosterically activated by binding histone modifications using reader domains or separate reader proteins, and some MTases operating beyond chromatin are activated by binding an activator protein. In this review, we describe the structural status of the human methyltransferome and then discuss newly revealed structural mechanisms of MTase activation.

Address
Department of Chemistry and Chemical Biology, Cornell University, Ithaca, NY 14853, United States

Corresponding authors: Fenwick, Michael K (mkf7@cornell.edu), Ealick, Steven E (see3@cornell.edu)

Current Opinion in Structural Biology 2018, **53**:12–21

This review comes from a themed issue on **Catalysis and regulation**

Edited by **Hazel Holden** and **Alice Vrielink**

https://doi.org/10.1016/j.sbi.2018.03.007

Structural classification and distribution

The Protein Data Bank (PDB) currently contains about 2000 structures of methyltransferases (MTases), which can be classified according to the structural fold used to bind the methyl donor. Greater than 90% of the more than 300 methylation reactions classified under EC 2.1.1 use *S*-adenosylmethionine (SAM) as the methyl donor. SAM-dependent methylation reactions can be further subdivided into nucleophilic substitution or radical reactions. Structural alignment and clustering of MTase catalytic domains shows eleven SAM-dependent [1–9] and four SAM-independent [10–19] folds used for binding the methyl donor, including one radical SAM MTase fold [1,20••]. Additional classes of radical SAM MTases have been identified but are structurally uncharacterized. The gene names and PDB codes of the founding MTase structures of the various fold classes, the fold distribution in the PDB, common substrates, and methyl donors of SAM-independent folds are given in Figure 1. The reader is also referred to earlier reviews for structural descriptions of the major fold classes [21].

The human methyltransferome comprises about 200 MTases and contains greater numbers of 7-β-strand (Rossmann-like) and SET class MTases in comparison to the yeast methyltransferome [22]. This is partly due to the greater complexity of gene regulation in humans and is reflected by the larger number of chromatin-modifying 7-β-strand class DNA and protein arginine MTases (DNMTs and PRMTs) and SET class protein lysine MTases (PKMTs). Many human PKMTs and PRMTs add methyl groups to lysines and arginines positioned at the N-terminal tails of histones. These and other histone post-translational modifications provide a molecular code that histone readers, writers, and erasers use to regulate DNA methylation and gene expression. In this review, the available human MTase structures are summarized and then focus is shifted to a functionally important and structurally interesting subset of MTases whose activation requires structural changes induced by binding histones or binding a protein activator.

Structural characterization of the human methyltransferome

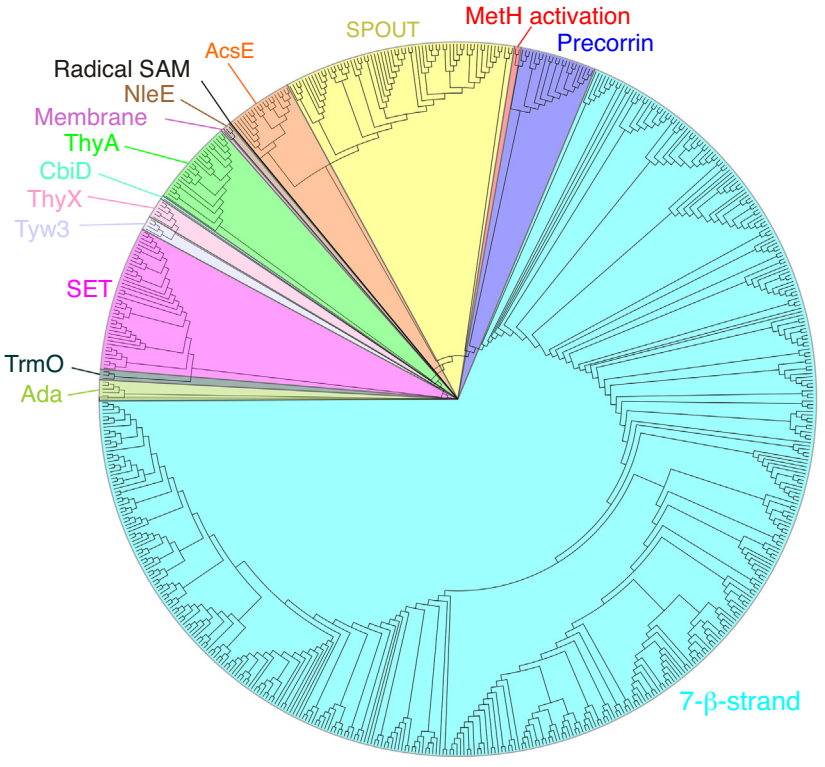
Structural coordinates for nearly half of the MTases of the putative human methyltransferome [22] are currently available (Figure 2). A total of 201 known or predicted human MTases use eight SAM-dependent and three SAM-independent fold classes. The human 7-β-strand, SET, and SPOUT classes have 125, 56, and 8 members, respectively, of which 51, 57, and 50%, respectively, have human or orthologous eukaryotic structures available for their SAM-binding folds.

Several recently reported human MTase structures with bound substrates or substrate analogs have revealed active site architectures of different MTase families. These include structures of DNA cytosine-C5 MTase DNMT3A [23••], histone H3K27 MTase EZH2 [24••,25••], H3K4 MTases KMT2A and KMT2C [26••], C72 C5 tRNA MTase NSUN6 [27•], m1A58 tRNA MTase TRMT61A [28•], cap-specific mRNA nucleoside 2′-O MTase CMTR1 [29], mono and asymmetric di-methylarginine MTase CARM1 [30•], monomethylarginine MTase

Figure 1

SAM-dependent class		Typical methyl acceptors	PDB (year)
I	***Hp*HhaIM (7-β-strand)**	Various	1HMY (1993)
II	***Ec*MetH (MetH activation)**	MetH B12-Co(I)	1MSK (1997)
III	***Bm*CbiF (Precorrin)**	Precorrins	1CBF (1999)
IV	***Ec*RlmB (SPOUT)**	RNA	1GZ0 (2002)
V	***Nc*Dim5 (SET)**	Protein lysine	1ML9 (2002)
VI	***Af*CbiD (CbiD)**	Precorrin-5B	1SR8 (2004)
VII	***Ss*Taw3 (Taw3/Tyw3)**	yW-86	1TLJ (2004)
VIII	***Hi*TrmO (TrmO/β-barrel)**	tRNA	1XQB (2004)
IX	***Ec*RlmN (Radical SAM/$(\beta\alpha)_6$ barrel)**	Various	3RFA (2011)
X	***Ma*2698 (ICMT/membrane)**	Various	4A2N (2012)
XI	***Ec*NleE (NleE)**	NZF cysteine	4R29 (2014)

SAM-independent class		Methyl donor	Methyl acceptor	Methyl acceptor	PDB (year)
I	***Ec*ThyA (ThyA)**	MTHF	dUMP		2TSC (1991)
II	***Ec*Ada (Ada)**	O6-mG-DNA	MGMT		1SFE (1996)
III	***Mt*AcsE (AcsE/TIM barrel)**	5MeTHF	CoFeSP	ACS	1F6Y (2000)
	***Hs*BHMT**	Betaine	Hcy-Zn		1LT8 (2002)
	***Tm*0268/MetH**	5MeTHF	B12-Co(I)	Hcy-Zn	1Q8J (2004)
	***Mb*MtmB1**	Methylamine	Corrinoid	CoM	1TV2 (2004)
	***Tm*MetE**	5MeTHF	Hcy-Zn		1XPG (2005)
	***Mb*MtaB**	Methanol-Zn	FIII	CoM	2I2X (2006)
	***Rm*MttB**	Trimethylamine	Corrinoid	CoM	4YYC (2015)
	***Ec*MmuM**	SMM, (*R,S*)-SAM	Hcy-Zn		5DMM (2015)
IV	***Tm*ThyX (ThyX)**	MTHF	dUMP		1KQ4 (2002)



Fold	% in PDB
7-β-strand	68.6
SPOUT	10.1
SET	6.6
ThyA	3.9
Precorrin	3.7
AcsE	3.0
Ada	0.9
ThyX	0.9
Tyw3	0.6
TrmO	0.5
MetH activation	0.3
NleE	0.3
CbiD	0.2
Membrane	0.2
Radical SAM	0.2

Current Opinion in Structural Biology

MTase structural fold classification and distribution in the PDB, as of December 2017. Fold classes represent topologically similar structures used to bind a methyl donor and are listed by the gene name(s) of the founding structure(s) with alternative designations given in parentheses, and the PDB code of the first released structure. The phylogenetic tree was produced by structural alignment (using DALI [76]) and UPGMA clustering of 641 MTase catalytic domains having less than 95% sequence similarity. The pie chart gives the percentages of the 641 MTases represented by the different fold classes. Abbreviations are *Hp*, *Haemophilus parahaemolyticus*; *Ec*, *Escherichia coli*; *Bm*, *Bacillus megaterium*; *Nc*, *Neurospora crassa*; *Af*, *Archaeoglobus fulgidus*; *Ss*, *Sulfolobus solfataricus*; *Hi*, *Haemophilus influenzae*; *Ma*, *Methanosarcina acetivorans*; *Mt*, *Moorella thermoacetica*; *Hs*, *Homo sapiens*; *Tm*, *Thermotoga maritima*; *Mb*, *Methanosarcina barkeri*; *Rm*, *Rhizobium meliloti*; B12, vitamin B12; SPOUT, *E. coli* SpoU-TrmD; SET, *Drosophila melanogaster* suppressor of variegation 3–9, enhancer of zeste, and trithorax; yW, wybutosine; ICMT, isoprenylcysteine carboxyl MTase; NZF, Npl4-like Zinc Finger; MTHF, 5,10-methylenetetrahydrofolate; dUMP, deoxyuridine monophosphate; O6-mG, O6 methylguanine; MGMT, O6-mG DNA MTase; TIM, triosephosphate isomerase; 5MeTHF, methyltetrahydrofolate; CoFeSP, corrinoid/iron-sulfur protein; ACS, acetyl-CoA synthase; Hcy, homocysteine; CoM, coenzyme M; FIII, 5-hydroxybenzimidazolyl cob(I)amide; and SMM, S-methylmethionine.