



HAL
open science

Étude structurale des protéine arginine méthyltransférases : reconnaissance des substrats et conception rationnelle de modulateurs

Nils Marechal

► **To cite this version:**

Nils Marechal. Étude structurale des protéine arginine méthyltransférases : reconnaissance des substrats et conception rationnelle de modulateurs. Biochimie, Biologie Moléculaire. Université de Strasbourg, 2018. Français. NNT : 2018STRAJ048 . tel-03763164

HAL Id: tel-03763164

<https://theses.hal.science/tel-03763164>

Submitted on 29 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



UNIVERSITÉ DE STRASBOURG

École Doctorale des Sciences de la Vie et de la Santé



Thèse Présentée par

Nils MARECHAL

Soutenue le 14 septembre 2018

Pour l'obtention du grade de :

DOCTEUR DE L'UNIVERSITÉ DE STRASBOURG

Discipline : Sciences du Vivant

Spécialité : Aspects moléculaires et cellulaires de la biologie

Étude structurale des Protéine Arginine Méthyltransférases

Reconnaissance des substrats et conception rationnelle de modulateurs

Thèse dirigée par le Pr. Jean CAVARELLI

Dr. Françoise OCHSENBEIN	Rapporteuse externe
Dr. Lionel MOUREY	Rapporteur externe
Dr. Valérie LAMOUR	Examinatrice

Remerciements

Je souhaite tout d'abord remercier Jean Cavarelli, mon directeur de thèse, pour m'avoir accueilli au sein de son équipe et pour m'avoir accompagné durant ces quatre années de thèses, dans ma construction scientifique et dans mon souhait d'enseigner. Merci de m'avoir permis de travailler sur de nombreux projets passionnants, de m'avoir donné l'occasion de me former à la microscopie électronique et d'avoir laissé s'exprimer mes fantaisies expérimentales. Surtout, merci pour votre humanité et votre patience.

J'aimerais également remercier les membres du jury de thèse pour avoir accepté de juger et sanctionner mes travaux : Françoise Ochsenein, responsable du laboratoire « Assemblage moléculaire et intégrité du génome » à l'institut de biologie intégrative de la cellule, Saclay ; Lionel Mourey, directeur de recherche dans l'équipe « Biophysique structurale » à l'institut de pharmacologie et de biologie structurale, Toulouse ; Valérie Lamour, directrice de recherche au laboratoire « Stabilité de la chromatine et mobilité de l'ADN » à l'institut de génétique et de biologie moléculaire et cellulaire, Illkirch.

Merci à l'IGBMC et au CERBM-GIE pour m'avoir accordé un financement LabEx lors de mes trois premières années de thèse, et à la ligue contre le cancer et à l'université de Strasbourg pour m'avoir donné l'opportunité de prolonger d'un an mes travaux.

Merci à l'infrastructure française de biologie structurale intégrative (FRISBI), aux différentes plateformes et services communs de l'IGBMC pour m'avoir permis de réaliser mes travaux dans les meilleures conditions de ressources scientifiques et technologiques.

Merci à tous nos collaborateurs : Alain Wagner et Julie Schmitt, Matthijs van Haren et Nathaniel Martin, Paola Arimondo et Ludovic Halby, Gianluca Sbardella pour les projets de développement d'inhibiteurs et de dérivés peptidiques. Sarah Cianferani-Sanglier et Oscar Hernandez pour les analyses en spectrométrie de masse.

Un grand merci aux membres de l'équipe. À Vincent, pour la formation en cristallographie, collecte et traitement des données, et en biochimie que vous m'avez dispensé. À Nathalie et Luc, pour vos précieux conseils au quotidien. À vous trois, avec qui j'ai toujours eu plaisir à échanger au coin de la paillasse ou autour d'un café, sur des sujets professionnels et personnels.

Merci à vous, Julio, Corinne et Jean-François, pour avoir commencé ma formation à l'utilisation de ces incroyables outils que sont les microscopes électroniques.

Merci à toi Rachel, pour tous les moments que nous avons partagés. Je n'oublierais pas ton naturel et ton humour lors des bons moments comme des instants plus compliqués durant ces quatre années. Mumin, notre rencontre restera gravée dans ma mémoire. Merci à vous, Brice et Arnaud,

pour votre bonne camaraderie et tous vos conseils en microscopie électronique. Merci Pernelle et Moamen pour votre agréable compagnie et votre bonne humeur au quotidien.

Merci à vous, Geoffrey, Jérôme et Nicolas pour votre amitié. Sans vous je n'aurais su avoir la confiance qui m'a aidé à me hisser là où j'en suis aujourd'hui. Merci à Cyrielle, Alice, Astrid, Joanna, Pascaline Morgane et Céline, mes camarades de l'ESBS avec qui j'ai passé trois années extraordinaires. Merci à Memo qui m'a donné l'occasion de finir mes études par un stage incroyable aux États-Unis et m'a aidé à mûrir.

Je voudrais adresser des remerciements tout particuliers à ma mère et à mes deux grands-parents, Vonnette et Papychef, qui malgré mes difficultés m'ont toujours encouragé et ont su me guider vers des voies alternatives pour que je réussisse. Merci aussi à mon frère Ewen, pour tous ses conseils avant, pendant et probablement après ce doctorat. Merci aussi à toi Nolwenn pour ta générosité, car bien que je sois loin, je suis rassuré de savoir que tu es là pour prendre soin de notre famille.

Je tiens à terminer ces remerciements par Marilyne et Nicolas. Marilyne, merci de m'avoir soutenu et accompagné durant ces treize années. Je te remercie de m'avoir suivi à Strasbourg, loin de tes racines pour que je puisse y faire ma thèse et de continuer à me suivre. Pardon pour ces quatre années durant lesquelles j'avais souvent l'esprit ailleurs... Mon petit Nicolas. Merci pour le bonheur que tu nous procures chaque jour et pour la distraction que tu m'as apportée lors de la rédaction de ce manuscrit.

À Papychef et Vonnette

TABLE DES MATIERES

I - Introduction.....	8
A - Épigenétique - État de l'art.....	8
1 - La molécule d'ADN : support de l'information génétique.....	8
2 - Le génome : partition orchestrée par l'épigenétique.....	8
3 - États de la chromatine.....	10
4 - La méthylation de l'ADN.....	12
5 - Modifications post-traductionnelles des histones.....	12
6 - Les variants d'histones.....	15
7 - Les complexes de remodelage ATP-dépendant de la chromatine.....	15
8 - La cis-trans isomérisation des prolines.....	16
B - Les Protéine-αRginine MéthylTransférases.....	17
1 - Généralités.....	17
2 - Les PRMT : fonctions biologiques.....	21
3 - Régulation de l'activité des PRMT.....	40
4 - Structure des PRMT.....	41
5 - Reconnaissance des méthyl-arginines et déméthylation.....	57
6 - Inhibiteurs des PRMT.....	60
II - Étude structurale et fonctionnelle de SmPRMT3.....	79
A - Introduction.....	79
B - Résultats et discussion.....	81
1 - Choix des constructions utilisées.....	81
2 - Étude de SmPRMT3 par cryo-microscopie électronique.....	100
3 - Étude préliminaire du complexe SmPRMT3-SmrpS2.....	104
C - Matériel et méthodes.....	112
1 - Modélisation par homologie.....	112
2 - Clonage des constructions de SmPRMT3 et de SmrpS2.....	113
3 - Expression et purification des constructions SmPRMT3.....	116
4 - Expression et purification des complexes SmPRMT3-SmrpS2.....	119
5 - Caractérisation biophysique de SmPRMT3-fl et -C3.....	120
6 - Essais de cristallisation de SmPRMT3.....	121
7 - Étude de SmPRMT3 par cryo-microscopie électronique.....	123
III - Étude structurale et fonctionnelle de PRMT2.....	125
A - PRMT2.....	125
B - Résultats et discussion.....	125
1 - Étude structurale et fonctionnelle de PRMT2.....	125
2 - Reconstitution et caractérisation du complexe PRMT2-rsf1.....	127
3 - Étude cristallographique du complexe PRMT2-rsf1.....	128
4 - Étude du complexe PRMT2-rsf1 par cryo-microscopie électronique.....	129
C - Matériel et méthodes.....	135
1 - Expression et purification des protéines.....	135
2 - Reconstruction et purification du complexe PRMT2-rsf1.....	135
3 - Spectrométrie de masse.....	136
4 - Cristallisation du complexe PRMT2-rsf1.....	136
5 - Étude du complexe PRMT2-rsf1 par cryo-microscopie électronique.....	136
IV - Étude structurale et fonctionnelle de complexes PRMT-peptides substrats.....	139
A - Introduction.....	139
B - Résultats et discussion.....	141
1 - Les peptides analogues de l'état de transition.....	141
2 - Complexes CARM1-dérivés de l'histone H3.....	144
3 - Généralisation des dérivés peptidiques à d'autres PRMT.....	155
C - Perspectives.....	155
1 - Dérivés peptidiques de l'histone H3.....	155

2 - Application à d'autre peptides et substrats complets.....	156
D - Matériel et méthodes.....	157
1 - Préparation des peptides.....	157
2 - Expression et purification des constructions de mCARM1.....	157
3 - Cristallisation des complexes mCARM1-dérivés peptidiques.....	157
4 - Collecte et traitement des données.....	158
V - Étude structure-fonction du mutant CARM1 P409A.....	159
A - Résultats.....	162
1 - Analyse structurale du mutant CARM1 P409A.....	162
2 - Caractérisation fonctionnelle du mutant P409A de CARM1.....	164
B - Discussion.....	165
C - Matériel et méthodes.....	167
1 - Mutagenèse dirigée.....	167
2 - Expression et purification du tétramère d'histones (H3H4) ₂	167
3 - Expression et purification de GST-PABP1 (437-488).....	168
4 - Fluorimétrie par balayage différentiel (DSF).....	169
5 - Détermination de l'activité enzymatique de CARM1 WT et P409A.....	169
6 - Détermination de la constante de dissociation CARM1-substrat.....	169
7 - Étude structurale de mCARM1 (130-487) P409A.....	170
VI - Conception d'inhibiteurs des PRMT.....	172
A - Introduction.....	172
B - Résultats.....	173
1 - Les composés LH.....	173
2 - Les composés U.....	176
3 - Les composés EML.....	184
4 - Fixation et inhibition des PRMT par les composés SATT.....	190
5 - Perspectives.....	196
C - Matériel et méthodes.....	199
1 - Expression et purification des PRMT.....	199
2 - Synthèse des composés SATT.....	200
3 - Préparation des composés.....	200
4 - Co-cristallisation des complexes PRMT-inhibiteurs.....	200
5 - Mesure de la stabilité thermique.....	202
6 - Mesure des IC ₅₀	202
7 - Préparation des récepteurs et ligands pour le « docking » <i>in silico</i>	203
VII - Conclusions et perspectives.....	204
A - Étude structurale de SmPRMT3.....	204
1 - Étude de SmPRMT3 par bio-cristallographie.....	204
2 - Étude structurale de SmPRMT3 par cryo-microscopie électronique.....	205
3 - Travaux préliminaires sur le complexe SmPRMT3-Smrs2.....	205
B - Étude structurale et fonctionnelle de PRMT2.....	206
C - Étude structurale des complexes CARM1-peptides.....	207
D - Analyse structurale et fonctionnelle du mutant P409A.....	209
E - Conception d'inhibiteurs sélectifs des PRMT.....	210
1 - Les composés LH.....	211
2 - Les composés U.....	211
3 - Les composés EML.....	211
4 - Développement des composés SATT.....	211
VIII - Annexes.....	216
Articles joints et communications scientifiques.....	233
Bibliographie.....	234

Abréviations

aDMA	asymétrique diméthyl Arginine
ADN	Acide DésoxyriboNucléique
ARG	Arginine
ARN	Acide RiboNucléique
BTP	Bis Tris Propane
CARM1	« Coactivator-Associated aRginine méthyltransferase 1 »
Cryo-EM	« Cryo-Electron Microscopy »
DMSO	DiMéthylSulfOxyde
DO600	Densité Optique à 600 nm
DSF	« Differential Scanning Fluorimetry »
DTT	DithioThréiTol
EDTA	ÉthylèneDiamineTétraAcétique
ESI-TOF	« Electro-Spray Ionisation-Time Of Flight
F1	« Full length »
GST	Glutathion S-Transférase
H1/H2/H3	Histone 1/2/3
H2A/H2B	Histone 2A/2B
IC ₅₀	« Inhibitory Concentration 50 % »
IMAC	« Immobilized Metal ion Affinity Chromatography »
IPTG	IsoPropyl β-D-1-ThioGalactopyranoside
LB	Lysogénie-Broth
mCARM1	CARM1 de souris
MES	Acide MorpholinoEthanoSulfonique
MMA	Mono Méthyl Arginine
mPRMT2	PRMT2 de souris
PABP1	PolyA-Binding Protein 1
PEG	PolyEthylène Glycol
PH	« Pleckstrin Homology domain »
PMSF	Fluorure de PhénylMéthylSulfonyle
PRMT	Protéine aRginine MéthylTransférase
RMSD	« Root Mean Square Deviation »
RpS2	« Ribosomal protein S2 »
Rsf1	« Repressor splicing factor 1 »
SAH	S-Adénosyl-L-Homocystéine / AdoCy
SAM	S-Adénosyl-L-Méthionine / AdoMet
SAR	Structure-Activity Relashionship
sDMA	symétrique diméthyl Arginine
SDS-PAGE	« Sodium Dodecyl Sulfate-PolyAcrylamide Gel Electrophoresis »
SFG	Sinefungine
SH3	« Src-Homology 3 »
SMN	« Survival of Motor Neuron »
SmPRMT3	PRMT3 de <i>Schistosoma mansoni</i>
SmrpS2	rpS2 de <i>Schistosoma mansoni</i>
TCEP	Tris(2-CarboxyEthyl)Phosphine
TCEP	Tris(2-CarboxyEthyl)Phosphine
TEV	« Tobacco etch Virus »
TIM	« Triose-Phosphate Isomérase »
UA280	Unité d'Absorbance à 280 nm
WT	« Wild Type »
zPRMT2	PRMT2 de poisson zèbre
βME	β-MercaptoEthanol

Étude structurale des Protéine Arginine Méthyltransférases
Reconnaissance des substrats et conception rationnelle de modulateurs

Préambule

Les protéines-arginine méthyltransférases (PRMT) sont impliquées dans de nombreux processus cellulaires comme la transcription, la maturation et le transport des ARN, la traduction, la transduction du signal, la réparation de l'ADN et l'apoptose. Au moins neuf PRMT (notées PRMT1 à PRMT9) ont été identifiées chez l'homme. Les PRMT sont principalement divisées en trois grandes classes (type I, II et III) selon la nature chimique des arginines modifiées (monométhylation, diméthylation symétrique ou asymétrique), le donneur de méthyl étant la S- adénosyl-méthionine (SAM). Les PRMT sont des protéines modulaires, généralement associées sous forme d'homodimères, chaque monomère contenant de 320 à 970 acides aminés. Un monomère de PRMT est construit autour d'un module catalytique qui est commun à toutes les PRMT et auquel des domaines additionnels spécifiques à chaque PRMT sont éventuellement ajoutés. L'unité fonctionnelle des PRMT nécessite la formation d'un dimère de deux modules catalytiques. Plus de 130 protéines substrats de PRMT ont été identifiées à ce jour, chacune des 9 PRMT ayant une spécificité structurale et fonctionnelle qui lui est propre, sans qu'il y ait redondance entre PRMT à l'intérieur d'un organisme donné.

De nombreux travaux ont montré que la méthylation des arginines est un évènement important pour la régulation de l'expression des gènes et que des dérégulations de cette modification épigénétique sont impliquées dans le développement, la progression ou l'agressivité de certains cancers dont par exemple les cancers du sein, du poumon ou de la prostate. Les PRMT constituent ainsi une des nouvelles cibles potentielles en chimiothérapie. La découverte d'inhibiteurs spécifiques de PRMT pourrait donc ouvrir des voies thérapeutiques nouvelles et de nombreux travaux et développements se sont produits ces dernières années.

En combinant des techniques biochimiques et structurales, l'équipe « Biologie structurale des cibles épigénétiques » dirigée par le Professeur Jean Cavarelli cherche à comprendre à l'échelle atomique le mode d'action des PRMT et de leurs complexes et à développer des approches rationnelles basées sur les structures tridimensionnelles pour caractériser des médiateurs moléculaires des activités biologiques de ces cibles. Pour stabiliser et capturer *in vitro* des complexes PRMT-substrat, la stratégie initialement développée au sein de l'équipe s'appuyait sur des composés réactifs analogues du SAM ou mimes de l'état de transition de la réaction de méthylation, permettant de former une liaison covalente avec une arginine substrat dans le site de méthylation des PRMT. Il résulte de cette approche que ces analogues ou mines pourraient servir de base à la conception de nouveaux inhibiteurs spécifiques ou sélectifs d'une famille de PRMT donnée. L'originalité de notre approche était donc de résoudre deux défis en utilisant une même classe de composés (mimes de l'état de transition). Nous utilisons en effet ces analogues ou mines

pour comprendre les modes de reconnaissance protéine-protéine lors de la réaction de méthylation et pour servir de base à la conception de nouveaux inhibiteurs spécifiques ou sélectifs d'une PRMT donnée.

À la suite des travaux effectués dans l'équipe sur PRMT2, PRMT4, PRMT6, PRMT7, les travaux présentés dans cette thèse ont comporté trois axes principaux :

- l'étude structurale et fonctionnelle de complexes PRMT4-inhibiteurs, PRMT4-peptides,
- l'étude structurale de la protéine PRMT2 et de ses complexes,
- l'étude structurale de la protéine SmPRMT3 et de ses complexes.

Ce manuscrit comporte sept chapitres dont cinq présentent les travaux effectués et les résultats obtenus au cours de cette thèse. Il débute par un chapitre de synthèse qui présente l'état des connaissances fonctionnelles et structurales de la famille des PRMT. Le second chapitre présente les travaux réalisés sur SmPRMT3 et sur son complexe avec rpS2. Le chapitre III détaille les résultats et travaux sur PRMT2. L'étude structurale et fonctionnelle de plusieurs complexes PRMT-peptides substrats est regroupée dans un même chapitre (Chapitre IV). Un chapitre particulier présente les travaux effectués sur un mutant ponctuel de CARM1 (Chapitre V). Le dernier chapitre dédié à la présentation des résultats obtenus durant cette thèse porte sur les travaux et résultats obtenus dans la conception rationnelle d'inhibiteurs des PRMT cibles de l'équipe (Chapitre VI), suivi d'un dernier chapitre (Chapitre VII) dégageant les axes principaux de développement qui feront suite à ce travail de thèse.

I - Introduction

A - Épigénétique - État de l'art

1 - [La molécule d'ADN : support de l'information génétique](#)

En 1900, la redécouverte par Carl Correns, Hugo de Vries et Erich von Tschermak-Seysenegg des lois de la génétique (Simunek et al., 2011), décrites pour la première fois par le moine Autrichien Gregor Mendel, fut le point de départ de toutes les recherches sur l'hérédité et sur la génétique moderne. Chacune de nos cellules contient l'information, qualifiée d'information génétique, nécessaire à sa survie et gouvernant sa capacité de s'adapter à son environnement. C'est en 1928 que le microbiologiste Frederick Griffith démontrait que l'information génétique est portée par une molécule (Petsko, 2006), identifiée seize ans plus tard comme l'acide désoxyribonucléique ou ADN (Avery et al., 1944). Faisant suite à l'identification de la structure tridimensionnelle de l'ADN par Watson et Crick en 1953, Nirenberg et Matthaei montrèrent en 1961 que la plus petite unité des gènes codant l'information génétique est constituée d'un triplet de nucléotides. Le développement de méthodes efficaces de séquençage de l'ADN (Maxam et al., 1977 ; Sanger et al., 1977) permit par la suite d'acquérir de nombreuses connaissances sur la structure des gènes procaryotes et eucaryotes. À la fin du 20^e siècle, le projet de séquençage de la portion euchromatique du génome humain était initié dans le but de faire avancer la recherche médicale, et s'acheva par la publication des séquences au tout début du 21^e siècle (Human Genome Sequencing Consortium, 2004). Les génomes de nombreux organismes ont depuis été séquencés, dévoilant la complexité de leur architecture, l'organisation des gènes et leur fonctionnement.

2 - [Le génome : partition orchestrée par l'épigénétique](#)

Chez l'homme, le noyau de chaque cellule contient 23 paires de chromosomes représentant un total de 3,2 milliards de paires de bases. L'expression des gènes dans les cellules qui composent nos organismes diffère selon le tissu qu'elles constituent et leurs besoins à différents moments de leur cycle de vie. Cependant, les molécules d'ADN sont les mêmes dans toutes les cellules somatiques chez un individu, si bien que la seule organisation de l'information portée par cette molécule ne permet pas d'expliquer les différences d'expression des gènes observées dans différentes cellules, tissus et organes, ni la dynamique au cours du temps. Nous savons aujourd'hui que les différences observées dans l'expression de l'information génétique sont principalement contrôlées par l'épigénétique.

L'idée d'épigénétique est apparue précocement avec les travaux sur l'hérédité et le développement. Sa définition a évolué avec les avancées scientifiques dans le domaine de la

génétique jusqu'à la définition de l'épigénétique moderne acceptée aujourd'hui. Le terme d'épigénotype a pour la première fois été employé en 1942 (Waddington, 2012) pour décrire l'ensemble des processus faisant le lien entre le génotype, défini comme les constituants héréditaires d'un organisme, et le phénotype, défini comme les caractéristiques visibles de l'organisme adulte. Waddington définissait alors l'épigénétique comme « la branche de la biologie qui étudie les interactions entre les gènes, et leurs produits qui établissent le phénotype » (Deans et al., 2015 ; Waddington, 2012). Jusque dans les années 1960, le terme épigénétique était donc employé pour décrire les processus évolutifs et de différenciation menant au développement d'un organisme à partir d'un ovule fécondé (Felsenfeld, 2014).

Aujourd'hui, bien qu'il soit de mieux en mieux défini, le terme épigénétique est encore massivement employé par les acteurs de la recherche en biologie pour décrire le découplage entre l'information génétique, stable et identique dans toutes les cellules, et les nombreux phénotypes observés (Deans et al., 2015). Les progrès de notre compréhension du cycle cellulaire, de l'expression des gènes et de la réplication de l'ADN, liés aux découvertes des modifications post-traductionnelles des histones, de la méthylation de l'ADN et de leurs effets sur l'expression des gènes, ont cependant fait évoluer notre définition de l'épigénétique vers une vision moléculaire (Compere et al., 1981 ; Gershey et al., 1968 ; Holliday et al., 1975 ; Murray, 1964 ; Ord et al., 1967).

Dorénavant, la définition la plus adaptée de l'épigénétique est la modification de la structure des régions chromosomiques permettant à une cellule d'adapter les niveaux d'expression de ses gènes indépendamment les uns des autres afin de répondre de manière appropriée aux contraintes de son environnement. Ces changements doivent être stables, réversibles et héréditaires, et sont associés à divers mécanismes moléculaires impliquant les modifications covalentes des cytosines et des histones et le remodelage de la chromatine. L'altération de cette information épigénétique est associée au développement de nombreuses pathologies, notamment aux cancers (Portela et al., 2010). Les mécanismes épigénétiques sont donc des pivots dans les processus de développement et d'homéostasie des organismes et font l'objet d'un nombre grandissant d'études comme l'atteste la quantité d'articles sur publiés ces 20 dernières années, tel qu'illustré sur la figure 1.

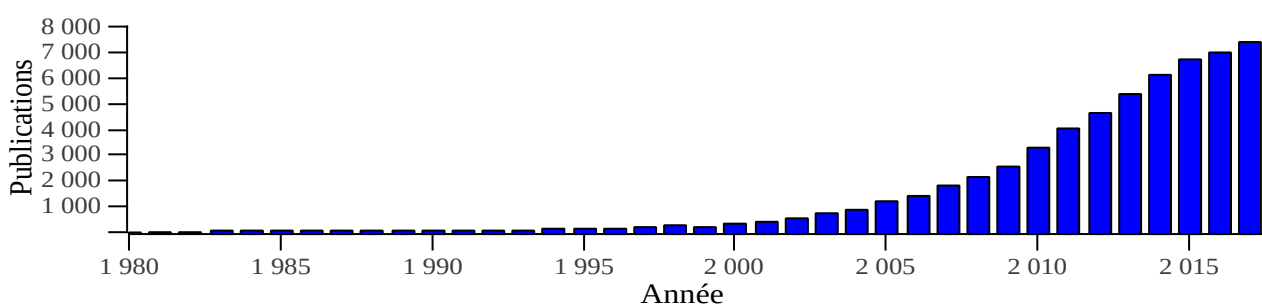


Figure 1 - Evolution du nombre de publications sur l'épigénétique depuis 1980.
Données extraites d'une recherche dans PubMed réalisée avec le mot clé « epigenetic ».

3 - États de la chromatine

Walter Flemming, père de la notion de chromatine, avait observé dans des noyaux cellulaires une substance s'organisant en structures filamenteuses lors de la division cellulaire. Ces structures sont les chromosomes, constitués de chromatine. La chromatine est observée sous deux formes dès 1928 (Heitz, 1928). La première forme est hautement condensée et le reste au cours du cycle cellulaire. Elle est appelée hétérochromatine. La seconde forme, l'euchromatine, est plus relâchée. Toutes deux sont composées de nucléosomes, qui en constituent l'unité de base (Kornberg, 1974). Comme l'illustre la figure 2, les nucléosomes sont constitués d'un octamère d'histones H2a, H2b, H3 et H4 (Kornberg et al., 1974) qui s'assemblent par paires pour former le cœur protéique du nucléosome, autour duquel s'enroule la double hélice d'ADN sur une longueur d'environ 147 paires de bases (Luger et al., 1997).

Les extrémités amino-terminales des histones, appelées queues, demeurent flexibles et s'étendent de part et d'autre du nucléosome. Une neuvième histone, H1, interagit avec le point d'entrée/sortie de l'ADN pour stabiliser le nucléosome (Sivolob et al., 2003). Le complexe nucléoprotéique ainsi constitué de l'ADN, de l'octamère d'histones et de l'histone H1 s'appelle le chromatosome. La formation du chromatosome induit un niveau de condensation de l'ADN supplémentaire, appelé fibre de diamètre 30 nm, qui adopte une structure solénoïdale (Marsden et al., 1979 ; Robinson et al., 2006 ; Thoma, 1979). Le modèle de condensation généralement accepté stipule que la fibre de 30 nm s'enroulerait sur elle-même pour former un solénoïde de diamètre 300 nm, qui lui-même se condenserait en solénoïde de diamètre 700 nm, se condensant pour former le chromatide (Banfalvi, 2008 ; Li et al., 2015).

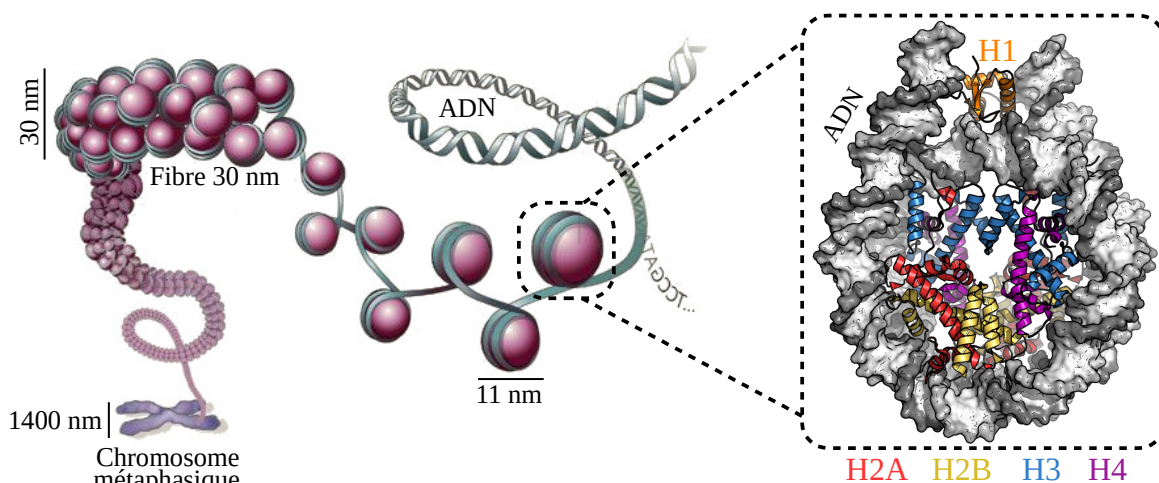


Figure 2 - Représentation schématique de la compaction de l'ADN.

Figure adaptée de la communication scientifique (Baylin et al., 2007). A gauche, représentation schématique de la compaction de l'ADN. A droite, structure cristallographique du chromatosome à 3,5 Å (PDB 4QLC ; Zhou et al. 2015).

L'euchromatine et l'hétérochromatine sont respectivement les lieux d'une activité transcriptionnelle intense et faible. Ces différences sont dues au niveau d'accessibilité des gènes par

les complexes enzymatiques responsables de leur régulation. Des expériences de digestion de la chromatine par des DNAses ont en effet montré que les régions hétérochromatiques, très compactes, sont résistantes à ces enzymes (Razin et al., 1977). L'euchromatine et l'hétérochromatine sont caractérisées par des compositions en protéines ainsi que des modifications chimiques de l'ADN et des histones qui leur sont spécifiques. Par exemple, les nucléosomes dans les régions hétérochromatiques sont enrichis en histone H1 et en 5-méthylcytosine (Ball et al., 1983), tandis qu'on dénombre de nombreuses acétylations des histones H3 et H4 dans les nucléosomes des régions euchromatiques (Hebbes et al., 1988). Ces modifications sont en fait les interrupteurs épigénétiques qui gouvernent la dynamique de la chromatine, c'est-à-dire ses états de compaction et les mécanismes qui leur sont associés : l'activation et la répression de la transcription, la réparation et la réplication de l'ADN. Comme schématisé sur la figure 3, ces interrupteurs épigénétiques sont actionnés selon les besoins de l'organisme par des machineries protéiques incluant des complexes de remodelage de la chromatine, des modifications covalentes de l'ADN, des modifications post-traductionnelles des histones et l'insertion de variants d'histones par des chaperons spécifiques. Tous ces mécanismes sont interconnectés.

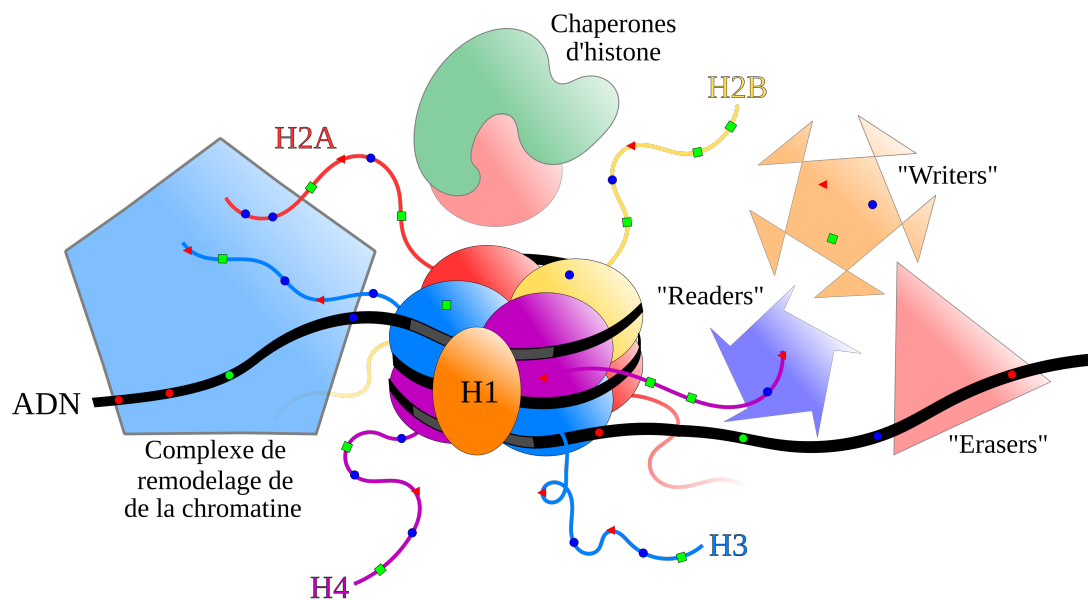


Figure 3 - Schéma des familles de complexes impliqués dans la régulation épigénétique.

Figure inspirée de la publication « Understanding the Complexity of Epigenetic Target Space: Miniperspective » (Pande, 2016). Les petits carrés, triangles et ronds colorés schématisent des modifications chimiques de l'ADN et des protéines.

En effet, les modifications de l'ADN et des histones servent de points d'ancrage à des domaines de protéines spécifiques, qualifiés de « readers », qui permettront le recrutement d'autres enzymes de modifications de l'ADN et des histones, qualifiées de « writers », ou inversement possédant la capacité d'effacer ces modifications. Ces dernières sont qualifiées « d'erasers ». Ces fonctions « readers », « writers » et « erasers » sont rassemblées au sein de complexes protéiques impliqués dans l'activation et la répression de la transcription, notamment des coactivateurs et

corépresseurs de la transcription et des remodeleurs de la chromatine, qui modulent l'activité des gènes en modulant l'accessibilité de l'ADN. Les paragraphes suivants traiteront synthétiquement de la méthylation de l'ADN, des modifications covalentes des histones et des variants d'histones. Les effets biologiques de ces modifications épigénétiques seront présentés ainsi que les protéines responsables de leur synthèse, de leur reconnaissance et de leur effacement.

4 - La méthylation de l'ADN

Les îlots CpG sont des séquences d'ADN riches en succession de CG. Chez les animaux, 72 % de ces îlots sont méthylés sur les cytosines (Saxonov et al., 2006). La méthylation des cytosines est la seule méthylation de l'ADN identifiée dans le monde animal et se localise principalement dans les régions organisées en hétérochromatine (Razin et al., 1977), ce qui associe cette modification à un état transcriptionnel inactif. La méthylation de l'ADN est catalysée par les enzymes de la famille des méthyltransférases de l'ADN (DNMTs). Ces enzymes utilisent le cofacteur S-adénosyl-L-méthionine (SAM) comme donneur de méthyle, et transfèrent ce groupement sur le carbone 5 de la cytosine substrat (Malygin et al., 2012). Les méthylations de l'ADN sont reconnues par des protéines possédant des domaines de fixation du méthyle-CpG (les domaines MBD). Ces domaines sont retrouvés dans des protéines recrutant des complexes impliqués dans la répression de la transcription, tels que des complexes de remodelage de la chromatine et des désacétylases d'histones (Du et al., 2015). Chez les femmes, la copie surnuméraire du chromosome X est totalement inactivée par méthylation de l'ADN. L'altération de la méthylation de l'ADN est associée à plusieurs pathologies dont des cancers (Chow et al., 2005 ; Eden et al., 1994 ; Miranda et al., 2007). Comme les modifications épigénétiques sont généralement réversibles, la méthylation de l'ADN peut être enlevée par les enzymes de la famille TET, par oxydation de la 5-méthylcytosine en 5-hydroxyméthylcytosine. La restauration de la cytosine non modifiée se termine par excision de la base et réparation de l'ADN (Wu et al., 2017).

5 - Modifications post-traductionnelles des histones

Un grand nombre de modifications post-traductionnelles des histones prend en général place sur leur extrémité amino-terminale. Ces modifications sont pour la plupart covalentes et impliquent le transfert de petits groupes chimiques ou de petites protéines telle l'ubiquitine. Elles agissent en modifiant les interactions entre nucléosomes (Luger et al., 1997) ou avec l'ADN, mais surtout en recrutant des effecteurs au niveau des promoteurs ou inversement en inhibant leurs interactions avec la chromatine. Comme l'illustre la figure 4, ces modifications post-traductionnelles concernent principalement les lysines, arginines, sérine, thréonines et tyrosines, et seuls les résidus alanine, valine, isoleucine et phénylalanine ne sont pas répertoriés comme modifiés. Les effets des modifications des histones peuvent être isolés, synergiques ou antagonistes selon l'histone

concernée, le site modifié et le type de modification apposé. Le lien entre ces modifications, leur impact fonctionnel, ainsi que le dialogue entre ces marques épigénétiques, est qualifié de « code histone » et représente un terrain de recherche majeur en biologie moderne.

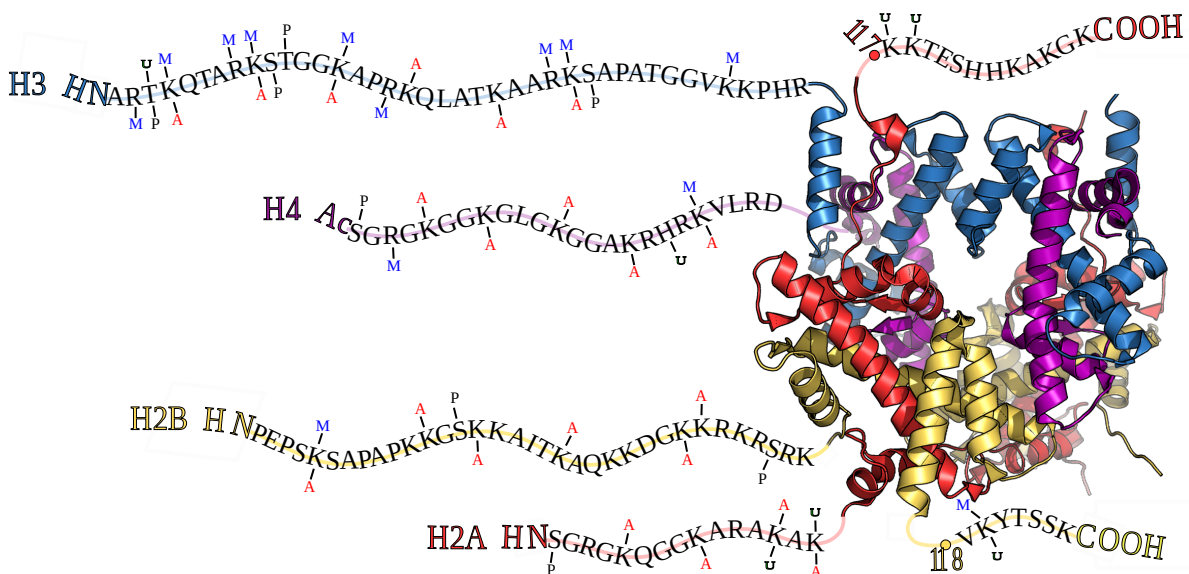


Figure 4 - Représentation schématique des modifications des histones.

Cette figure a été adaptée de la publication « The nucleosome: From genomic organization to genomic regulation » (Khorasanizadeh, 2004), et mise à jour avec l'article suivant : (Li et al., 2017). Les méthylation, acétylation, phosphorylation et ubiquitinylation sont annotées respectivement M (bleu), A (rouge), P (noir) et U (vert).

a - Phosphorylation des histones

La phosphorylation des histones a lieu sur les résidus sérine, thréonine et tyrosine. Elle est catalysée par les protéines kinases et ajoute une charge négative sur les résidus, tout en leur permettant d'établir des liaisons hydrogène supplémentaires et de s'associer à des ions (Taverna et al., 2007). L'effet le plus documenté concernant cette modification est la phosphorylation du variant d'histone H2A.X par ATM (Ataxia-Telangiectasia Mutated) et ATR (ATM- et Rad3-related) sur la sérine 129 au niveau des sites de cassure double-brin de l'ADN, qui permet le recrutement de la machinerie de réparation de l'ADN (Morrison et al., 2005). La phosphorylation des histones est aussi associée à la régulation de la transcription, en particulier de gènes prolifératifs, et à la compaction de la chromatine (Rossetto et al., 2012). Notre connaissance des éléments reconnaissant spécifiquement les histones méthylées est limitée : à ce jour, seul les membres de la famille 14-3-3 et les protéines à domaine BCRT ont été identifiés (Yun et al., 2011). Les phosphorylations peuvent être effacées par des « erasers » de la famille des protéines phosphatases.

b - La méthylation de lysines

La méthylation des lysines est catalysée par les histones-lysines méthyltransférases (HKMTs). Ces enzymes utilisent le SAM comme donneur de méthyle et peuvent catalyser la synthèse de mono-, di- et triméthyllysine. La méthylation des lysines préserve la charge de cet acide

aminé mais y ajoute des groupements supplémentaires qui modifient sa géométrie et augmentent son encombrement stérique. Cela module la spécificité de cet acide aminé pour différents « readers », regroupant un ensemble de protéines arborant notamment des chromo-domaines, des homéodomains de plantes (PHD), des domaines Tudor et WD40 (Teske et al., 2017). De fait, selon le site de méthylation et le nombre de méthyles ajoutés, la méthylation des lysines détermine des fonctions différentes : par exemple, les méthylation des résidus H3K9, H3K27 et H4K20 sont associées à la répression de la transcription tandis que les méthylation de H3K4, H3K36 et H3K79 sont des signaux d'activation transcriptionnelle (Martin et al., 2005 ; Sims et al., 2003). Ainsi, étant donnée l'étendue et l'importance des fonctions biologiques associées aux méthylation des lysines, les HKMTs sont la cible de nombreux efforts de développement d'inhibiteurs (Teske et al., 2017). Les méthylation des lysines sont effacées par les lysines déméthylases (KDMT), qui se classifient en deux groupes. Le groupe principal est constitué de la famille des déméthylases à domaine Jumonji (jmjC), qui utilisent l'ion Fe^{2+} et le 2-oxoglutarate comme cofacteurs. La réaction de déméthylation catalysée par cette famille de KDMT fait intervenir une réaction d'hydroxylation de l'atome d'azote ζ de la lysine (Langley et al., 2016). Le second groupe correspond aux déméthylases à flavines. Elles utilisent la flavine adénine dinucléotide (FAD) comme cofacteur et catalysent la déméthylation des lysines par oxydation (Burg et al., 2015).

c - L'acétylation des lysines

L'acétylation des lysines est catalysée par les enzymes de la famille des histone acétyltransférases (HAT). Ces enzymes utilisent le cofacteur acétyl-coenzyme A (acétyl-coA) comme donneur de groupement acétyle et catalysent une réaction de substitution nucléophile (Marmorstein et al., 2001 ; Zhang et al., 2014). L'acétylation des lysines a pour particularité d'ajouter un groupement amide susceptible de créer de nouvelles interactions, tout en neutralisant la charge positive de la lysine. À l'inverse de certaines méthylation des lysines, l'acétylation des histones est associée à un état d'activation transcriptionnel et au décompactage de la chromatine (Bannister et al., 2011 ; Eberharter et al., 2002). Le mécanisme d'action admis est que la neutralisation de la charge de la lysine diminue l'interaction entre les nucléosomes et l'ADN, ce qui a pour conséquence une décompaction de la chromatine autour de la zone acétylée suivi d'une accessibilité facilitée pour les facteurs de transcription (Seto et al., 2014). Les acétyllysines sont également des points d'ancrage pour les protéines arborant des bromo-domaines (Josling et al., 2012). Ces protéines participent ensuite au recrutement de facteurs de transcription généraux et spécifiques pour faciliter l'activation de la transcription. Les « erasers » responsables de la désacétylation des lysines sont les histones désacétylases, ou HDAC.

d - Les modifications des arginines

Les arginines peuvent subir plusieurs types de modifications dont la citrullination et l'ADP-ribosylation. La plus répandue est la méthylation des arginines, catalysée par les enzymes de la famille des protéine arginine méthyltransférases (PRMT). Cette famille d'enzyme est le sujet principal de cette thèse et cette partie sera détaillée dans le chapitre I, partie B.

6 - Les variants d'histones

Lors de la réplication de l'ADN en phase S du cycle cellulaire, une quantité importante d'histones sont synthétisées en vue d'être incorporées à la molécule d'ADN fille au niveau de la fourche de réplication, pour former le nouveau nucléosome (Marzluff et al., 2002). Ce mécanisme est global et qualifié de « réplication-dépendant ». Des histones ou nucléosomes entiers peuvent également être localement remplacés de manière « réplication-indépendante » (Ahmad et al., 2002). Dans ce cas, une histone classique peut être remplacée par une histone vierge de toute modification ou par un variant d'histone. Les variants d'histones se différencient des histones conventionnelles (H1, H2A, H2B, H3 et H4) par seulement quelques résidus. Sur un gène, le remplacement d'une histone par une histone vierge ou un variant a pour effet de profondément remanier ou totalement effacer l'état épigénétique antérieur et peut ainsi influencer le niveau d'expression du gène concerné. Actuellement, les variants d'histones les plus connus sont H3.3, CENP-A et H2A.X. Les régions transcriptionnellement actives sont enrichies en variant H3.3 (Filipescu et al., 2013). Ce variant est très proche de H3 et ne diffère que de quatre acides aminés dont trois sont situés dans le cœur de l'histone. Ces trois acides aminés restreindraient l'incorporation de H3.3 au mécanisme indépendant de la réplication. Bien que ce variant soit très répandu, sa fonction demeure mal comprise. Un autre variant de H3, CENP-A, est un marqueur moléculaire des centromères dans les cellules animales et végétales (Palmer et al., 1991). Le variant de H2, H2A.X, diffère de H2 par la présence d'une sérine faisant office de site de phosphorylation. Cette histone est incorporée aux sites de cassure double-brins de l'ADN et serait impliquée dans le recrutement de la machinerie de réparation (Morrison et al., 2005). D'autres variants d'histones existent, mais leurs fonctions demeurent mal caractérisées.

7 - Les complexes de remodelage ATP-dépendant de la chromatine

Les complexes de remodelage de la chromatine sont de gros complexes protéiques combinant les « readers » et « erasers » détaillés ci-dessus et assurant la régulation de l'expression des gènes. Ils utilisent l'adénosine triphosphate (ATP) comme source d'énergie pour restructurer les nucléosomes et modifier le niveau d'accessibilité de la chromatine (Vignali et al., 2000). Le complexe de remodelage et de désacétylation des nucléosomes (NurD) est l'un des principaux remodeleurs de l'ADN. NurD possède des domaines MBD2 et MBD3 capables de reconnaître les

îlots CpG et les méthylations de l'ADN. Il possède également une activité hélicase ATP-dépendante portée par les domaines ChD2/3/4, qui lui permet de repositionner les nucléosomes. NurD possède également une activité HDAC, ce qui a longtemps classé ce complexe parmi les répresseurs transcriptionnels, mais de récentes études montrent que NurD pourrait également induire l'expression des gènes (Basta et al., 2015). Une autre famille de remodeleurs de la chromatine est constituée par les complexes SWI/SNF. Les SWI/SNF possèdent un bromo-domaine leur permettant d'être recrutés sur la chromatine acétylée et sont pourvus de l'activité hélicase ATP-dépendante nécessaire à la réorganisation des nucléosomes. Les SWI/SNF sont des stimulateurs de l'activité transcriptionnelle (Tang et al., 2010). Finalement les complexes de remodelage de la chromatine appartenant à la famille des Chromo-domaine-helicase DNA-binding (ChD) possèdent deux chromo-domaines permettant de se fixer aux histones dont les lysines sont méthylées, et bénéficient de l'activité hélicase ATP-dépendante caractéristique de ces complexes. Les ChD sont à la fois des activateurs et répresseurs transcriptionnels, à l'instar du marquage épigénétique qu'ils reconnaissent (Marfella et al., 2007).

8 - [La cis-trans isomérisation des prolines](#)

La cis-trans isomérisation des prolines est une modification de la conformation du résidu proline. La structure particulière de cet acide aminé rend son isomérisation énergétiquement défavorable, particulièrement en configuration cis, si bien qu'une famille de chaperons existe pour assister les protéines lors de cette étape de repliement : les peptidyl-proline cis-trans isomérases ou PPIs (Wedemeyer et al., 2002). Cette modification de la structure spatiale des prolines n'est pas reconnue comme une modification épigénétique par un « reader » et n'est généralement pas considérée comme une modification épigénétique. Elle permet néanmoins d'influencer la synthèse d'autres marquages épigénétiques. Sur l'histone H3, l'acétylation de la lysine 14 favorise la conformation trans de la proline 16 ce qui limite la triméthylation de K14 (Howe et al., 2014). Dans *Saccharomyces cerevisiae*, la PPI Fpr4 catalyse l'isomérisation de H3P30 et H3P38, ce qui antagonise la méthylation de K36 *in vitro* et *in vivo* (Nelson et al., 2006).

B - Les Protéine-arginine MéthylTransférases

1 - Généralités

La mono- ou diméthylation des arginines est catalysée par les membres de la famille des Protéine-arginine MéthylTransférases (PRMT). Dans l'état actuel des connaissances, les PRMT ne sont trouvées que chez les organismes eucaryotes, des champignons aux métazoaires en passant par les végétaux. Cependant, des motifs caractéristiques des PRMT, décrits plus en détails dans le paragraphe « Structure du module catalytique », sont retrouvés dans des protéines prédites d'archaea, telle que la méthyltransférase de la protéine ribosomique L11 de *Methanosarcina mazei* (UniProt A0A0E3RB29). Les PRMT sont des enzymes (E.C. 2.1.1.319-322) qui utilisent le donneur de méthyle universel S-5'-adénosyl-L-méthionine (SAM/AdoMet) comme cofacteur afin de catalyser la mono- ou diméthylation d'une peptidyl-arginine (figure 5).

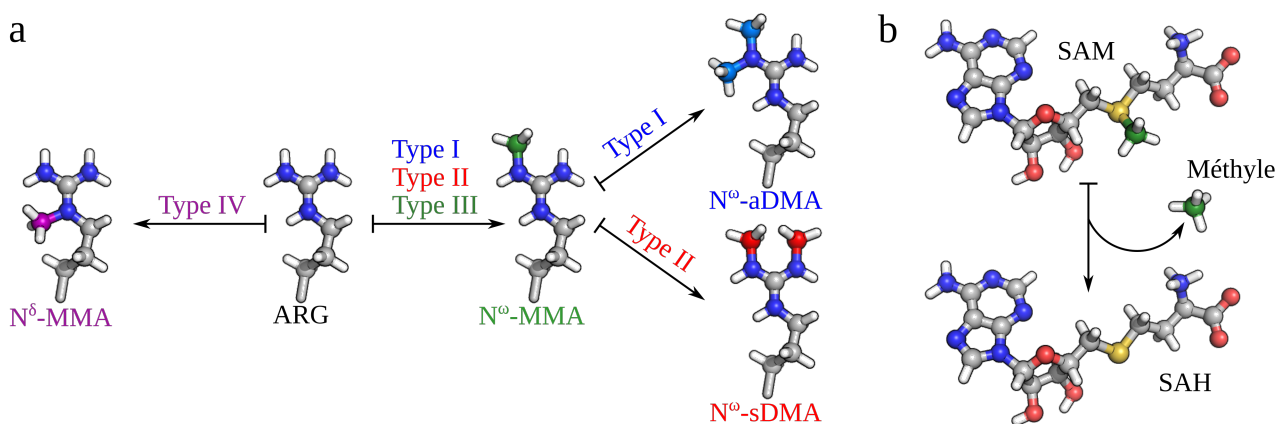


Figure 5 - Méthylation des arginines par les PRMT.

(a) mono- et diméthylation des arginines par les PRMT de type I (bleu), II (rouge), III (vert) et IV (violet). (b) Le S-adenosyl-L-méthionine (SAM) est dégradé en S-adenosyl-L-homocystéine (SAH) avec libération d'un groupe méthyle lors de la méthylation par les enzymes SAM-dépendante.

Neuf PRMT ont été identifiées chez les mammifères se répartissant en trois types, tel qu'illustré figures 5 : les PRMT de type I (PRMT1, PRMT2, PRMT3, PRMT4/CARM1, PRMT6 et PRMT8) et II (PRMT5, PRMT9) catalysent la formation de N^ω-monométhylarginine (MMA) puis respectivement de N^ω-diméthylarginine asymétrique et symétrique (aDMA & sDMA). Les PRMT de type III (PRMT7) ne produisent quant à elles que du MMA. Un quatrième type de PRMT (PRMT2) a été caractérisé chez la levure (Niewmierzycka et al., 1999) catalysant la synthèse de N^δ-monométhylarginine. Le type IV n'a actuellement pas encore été observé chez les autres eucaryotes. Les PRMT de type I ciblent généralement des motifs qualifiés de GAR pour « Glycine-Arginine Rich », de séquence G/F GGRGG G/F, à l'exception de PRMT4 qui ne possède pas de séquence consensus, mais méthylerait préférentiellement des motifs kAxRk dans lesquels une alanine est positionnée en n-2 du site de méthylation ainsi que des séquences contenant des prolines à proximité du site (Schurter et al., 2001 ; Shishkova et al., 2017).

Comme le montre la figure 6, les PRMT sont des protéines modulaires construites autour d'un module catalytique très conservé en structure et en séquence (en orange, bleu et violet, figure 6), pouvant présenter des extensions amino- et carboxy-terminales (en vert, figure 6) parfois organisées en domaines et leur conférant spécificité et sélectivité.

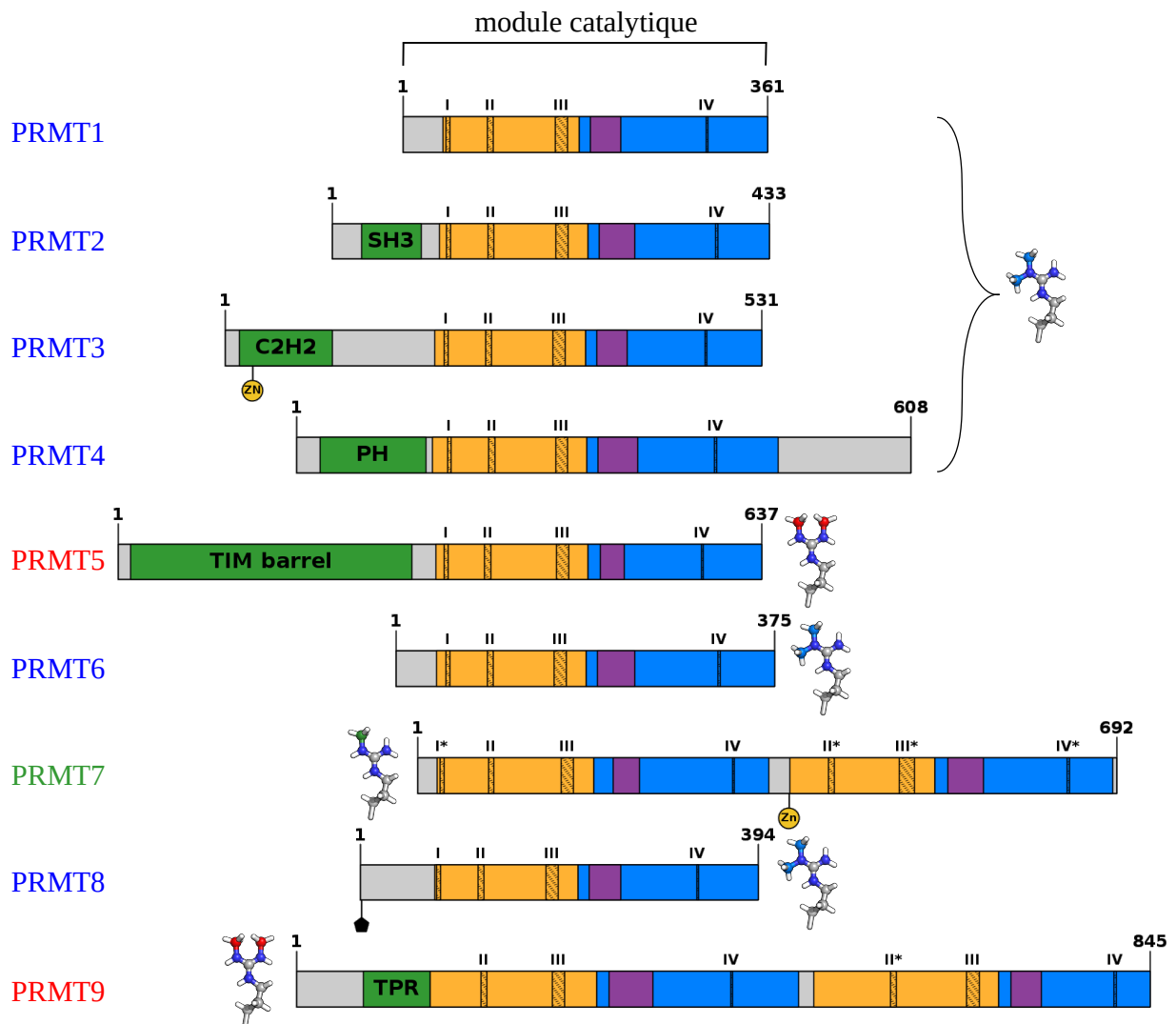
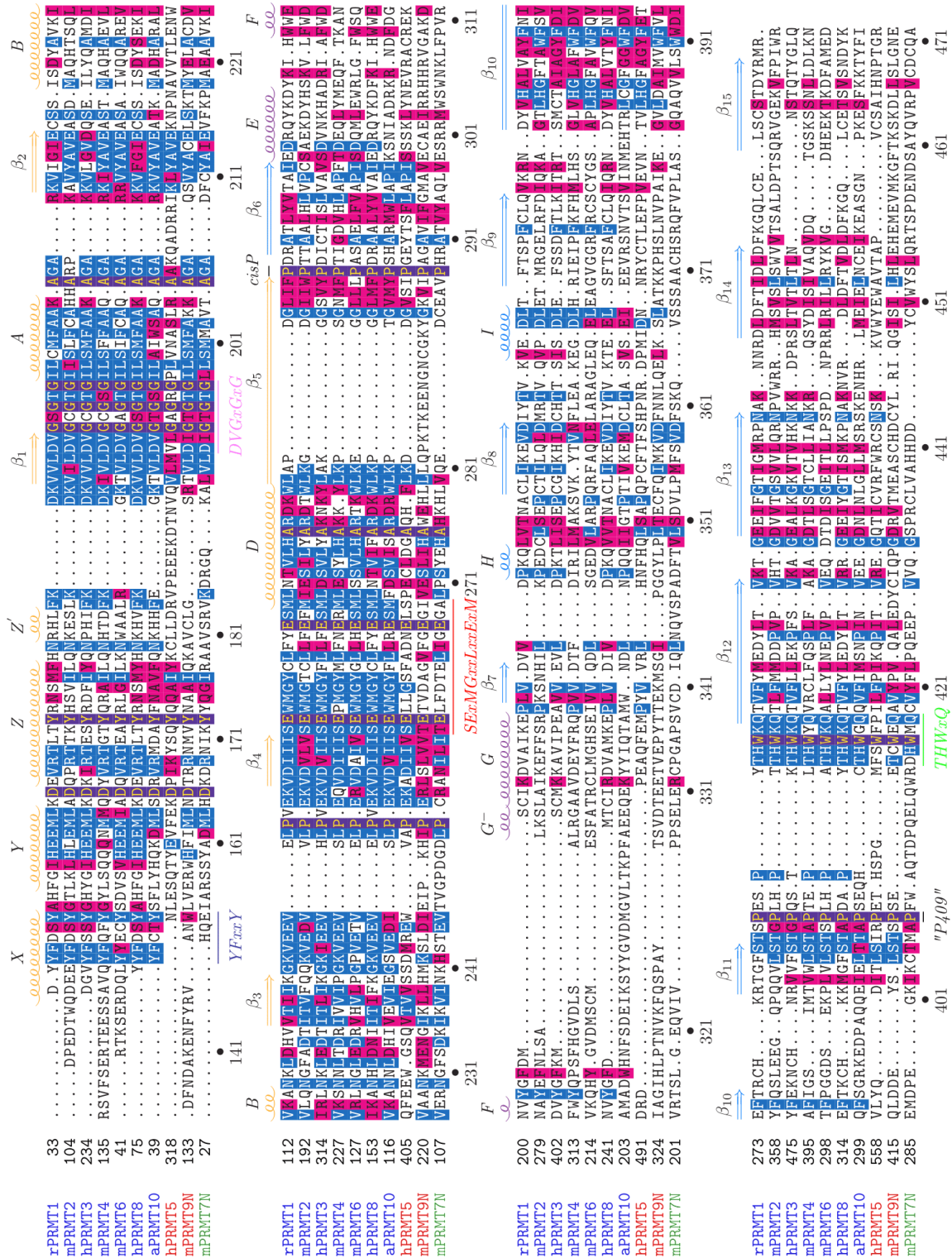


Figure 6 - Représentation schématique des neuf PRMT caractérisées chez les mammifères.

La numérotation indiquée correspond aux PRMT humaines. La délimitation des différents domaines a été déterminée d'après les structures disponibles, sinon par alignement de séquences. Les domaines additionnels sont représentés en vert. Les repliements de Rossmann, tonneau β et bras de dimérisation sont représentés respectivement en orange, bleu et violet. Les séquences consensus indiquées pour les 4 motifs sont représentatives des PRMT de type I, prédominantes. Les PRMT de type I, II et III sont annotées respectivement en bleu, rouge et vert.

Le tableau 1 présente les identités de séquences calculées entre les modules catalytiques à partir de l'alignement 1, supplémenté des séquences de 11 organismes différents. Les modules catalytiques de l'ensemble des PRMT partagent une identité de séquence de 27 % environ, avec des identités de séquences de 36 %, 32 % et 40 % respectivement au sein des PRMT de type I, II et III. Parmi les membres de cette famille d'enzymes, les identités s'échelonnent de 37 % pour le membre le plus faiblement conservé (PRMT6) à 67 % pour le plus conservé (CARM1).



Alignment 1 - Alignement de séquences des modules catalytiques des PRMT.

Les structures de toutes les PRMT disponibles dans la PDB ont été utilisées pour construire l'alignement. Seul un représentant par PRMT est présenté. La structure de PRMT9 n'a pas été résolue et sa séquence a été alignée sur l'alignement. Pour PRMT7 et PRMT9, seuls les modules amino-terminaux ont été représentés. Les structures secondaires de mCARM1 sont montrées en haut de l'alignement. Le domaine de fixation du SAM, le tonneau β et le bras de dimérisation sont colorés en orange, bleu clair et violet. La numérotation de la dernière ligne correspond à mCARM1. Les motifs I, II III et IV sont indiqués en bleu foncé, rose, rouge et vert. Les acides aminés sont colorés en fonction de la similarité avec la séquence consensus calculée à partir de cet alignement: invariant (violet) ou similaire (bleu). Les abréviations utilisées sont : m/*Mus musculus*, h/*Homo sapiens*, r/*Rattus norvegicus* et a/*Arabidopsis thaliana*. Cette figure a été générée par le programme TeXshade (Beitz, 2000).

	Type I						Type II		Type III
	PRMT1	PRMT2	PRMT3	CARM1	PRMT6	PRMT8	PRMT5	PRMT9	PRMT7
Identité (%)	27 ()								
moyenne (min - max)	36 (15 - 96)						32 (12 - 87)		40 (22 - 75)
	60 (48 - 96)	38 (23 - 60)	39 (15 - 82)	67 (47 - 94)	37 (24 - 60)	57 (33 - 95)	51 (30 - 87)	44 (32 - 62)	
Nombre d'organismes	10	5	10	6	7	6	8	4	8

Tableau 1 - Identités de séquence entre les modules catalytiques des PRMT.

Les identités de séquences ont été calculées avec l'outil en ligne SIAS (Université de Madrid) avec une matrice de score BLOSUM62, à partir d'un alignement de modules catalytiques de PRMT d'organismes se répartissant sur plusieurs embranchements, des mammifères aux levures. Pour PRMT7 et PRMT9, seuls les modules amino-terminaux ont été considérés. Ligne « Nombre d'organismes » : nombre d'organismes utilisé pour les calculs d'identité de séquence.

L'ensemble des eucaryotes ne possèdent pas le même éventail de PRMT, comme synthétisé dans le tableau 2, qui résume les PRMT présentes dans quelques organismes modèles. Seules PRMT1 et PRMT5 sont présentes chez tous les organismes, et les PRMT3 et PRMT7 sont également assez largement répandues. À l'inverse, les plantes possèdent une PRMT de type I nommée PRMT10 qui leur est spécifique, tandis que les levures sont les seuls organismes dans lesquels l'expression de la PRMT de type IV, RMT2, a pu être caractérisée.

	Type I							Type II		Type III	Type IV
	PRMT 1	PRMT 2	PRMT 3	CARM 1	PRMT 6	PRMT 8	PRMT 10	PRMT 5	PRMT 9	PRMT 7	RMT2
Mammifères	+	+	+	+	+	+		+	+	+	
Oiseaux	+		+			+		+	+	+	
Reptiles	+		+	+		+		+	+	+	
Amphibiens	+	+	+	+	+	+		+	+	+	
Poissons	+	+	+	+	+	+		+	+	+	
Urochordata	+		+	+	+			+	+	+	
Cephalochordata	+	+	+	+	+			+	+	+	
Echinodermes	+	+	+	+	+			+		+	
Arthropodes	+		+	+				+		+	
Nématodes	+							+	+	+	
Cnidaires	+	+	+					+	+		
Levures à fission	+		+					+			+
Levures bourgeonantes	+							+			+
Plantes	+		+	+		+	+	+		+	

Tableau 2 - Répartition des PRMT chez les eucaryotes.

Tableau adapté de l'arbre phylogénétique publié par Wang et Li (Wang et al., 2012). Les organismes sont classés par éloignement phylogénétique avec les mammifères. Les plantes possèdent une PRMT de type I unique parmi les autres eucaryotes.

2 - Les PRMT : fonctions biologiques

Un nombre croissant d'études impliquent les PRMT dans la plupart des processus cellulaires, de la transduction du signal à la régulation de l'expression des gènes, en passant par la biogenèse du ribosome. Ces recherches dévoilent une complexité croissante du mode d'action de ces enzymes, qui ont pour cible aussi bien les histones que les protéines non-histones. Le tableau 3 résume les séquences ciblées par les PRMT, ainsi que les fonctions et les localisations cellulaires de ces enzymes, qui sont détaillées dans les paragraphes ci-dessous.

Type	PRMT	Localisations	Fonctions	Substrats
TYPE I	PRMT1	cytoplasmique nucléaire	régulation de la transcription	H4R3 (+) ChTOP (+) NIP45 (+) STP5 (-) TAF15 (+) AE9a (?) HNF4 (+) Ash2L (?)
			transduction du signal	MRE11 (+) 53BP1 (+)
			intégrité du génome	ER α hnRNP-Q
	PRMT2	cytoplasmique nucléaire	régulation de la transcription	indéterminé
			régulation de l'épissage	Rsf1 (?)
	PRMT3	cytoplasmique	assemblage du ribosome	RpS2 (+)
	PRMT4 CARM1	nucléaire	régulation de la transcription	H3R17 (+) P300 (-) H3R26 (+) -
régulation de l'épissage et maturation des ARN			CA150 PABP1 SAP49 SmB	
PRMT6	nucléaire	régulation de la transcription	H3R2 (-) CRTC2 (+)	
		intégrité du génome	ADN polymérase β	
PRMT8	nucléaire (controversée)	régulation de la transcription (mais peu caractérisée)	Protéine du sarcome d'Ewing	
TYPE II	PRMT5	cytoplasmique nucléaire	régulation de la transcription	H3R8 (-) H3R2 (+) H4R3 (-) STP5 (+) H2AR3 (-) HOXA9 (+)
			assemblage du spliceosome	SmD1 (+) SmD3 (+)
			architecture du ribosome et de l'appareil de Golgi	S10 GM130 (ribosome) (Golgi)
PRMT9	nucléaire (une seule étude)	régulation de l'épissage (mais peu caractérisée)	SAP145	
TYPE III	PRMT7	cytoplasmique nucléaire	régulation de la transcription	H4R3 (-) H3R2 (+) H2AR3 (-) -
			assemblage du spliceosome	SmD1 (+) SmD3 (+)

Tableau 3 - Fonctions, localisations cellulaires et substrats des PRMT.

Ce tableau a été construit à partir des articles cités dans le paragraphe I-B-2 et sur les articles suivants : Shishkova et al. 2017; Gao, Dhar, et Bedford 2017; Herrmann et al. 2005; Fujiwara et al. 2006; Stopa, Krebs, et Shechter 2015. Colonne « localisations » la forme prédominante est représentée en gras. Colonne « substrats », les histones sont représentées en gras et les annotations (+), (-) et (?) indiquent une stimulation, une répression ou un effet encore indéfini.

α - PRMT1

➤ PRMT1 et la régulation de la transcription

Le gène codant pour PRMT1 fut identifié pour la première fois chez le rat (Lin et al., 1996) puis l'humain (Katsanis et al., 1997). PRMT1 est le membre prédominant des PRMT de type I dans les cellules mammifères où elle catalyse la majorité des aDMA (Tang et al., 2000). Cette PRMT est impliquée dans la régulation de la transcription, pour laquelle elle posséderait principalement un rôle d'activateur. Le mode d'action de PRMT1 dépend principalement de son activité de méthylation et s'applique à des partenaires histones comme non-histones.

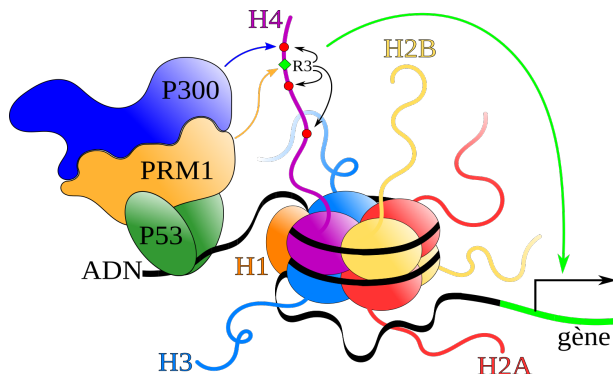


Figure 7 - Activation de la transcription par PRMT1 en synergie avec p300.

PRMT1 est recrutée sur le promoteur par des facteurs de transcription. PRMT1 méthyle H4R3 ce qui stimule l'acétylation de H4 par les membres de la famille p300 et induit l'activation de la transcription. Les flèches représentent la synthèse d'une modification post-traductionnelle ou un effet de stimulation.

La méthylation du résidu H4R3 par PRMT1 est le principal mécanisme d'activation de la transcription par cette PRMT (Strahl et al., 2001 ; Wang et al., 2001), et permet de stimuler l'acétylation des lysines de l'histone H4 par la famille des coactivateurs de la famille p300. La méthylation de H4R3 est nucléaire, mais PRMT1 a été décrite comme principalement cytoplasmique.

La translocation de PRMT1 dans le noyau et son recrutement au niveau des promoteurs sont donc des événements préliminaires à son activité de régulateur transcriptionnel, et nécessitent différents

partenaires dont la nature dépend des fonctions des gènes à réguler. Par exemple, l'expression des gènes cibles du facteur de transcription et répresseur de tumeurs p53 est induite par le recrutement de PRMT1 par p53 (An et al., 2004), tel qu'illustré en figure 7. Le méthylosome, un complexe formé par plusieurs protéines dont PRMT1 et PRMT5 (Friesen et al., 2001a), est recruté sur la chromatine enrichie en 5-hydroxyméthylcytosine (5hmC) par FoP/ChTOP (Friend of Prmt1/Chromatin Target Of Prmt1), ce qui induit la méthylation de H4R3 donc l'activation de la transcription (Takai et al., 2014). Ce mécanisme serait contrôlé par une boucle de régulation puisque la reconnaissance du 5hmC par ChTOP est dépendante de sa méthylation sur un motif GAR par PRMT1 et PRMT5 (van Dijk et al., 2010). Les gènes régulés par le complexe PRMT1-ChTOP sont sous contrôle des récepteurs aux œstrogènes, et la méthylation de ChTOP serait nécessaire à l'activation œstrogènes-dépendante de leur transcription (Takai et al., 2014). De plus, certains gènes associés au développement de glioblastomes sont enrichis en 5hmC (Takai et al., 2014), ce qui indique que l'inhibition de l'activité de PRMT1 pourrait être une approche thérapeutique contre ces cancers. Dans la leucémie myéloïde aiguë, PRMT1 est recrutée au niveau de promoteurs cibles par l'isoforme AE9a (Aml1-Eto 9a) de l'onco-protéine AML1-ETO (Shia et al., 2012). AML1-ETO est une protéine chimérique issue de la fusion des gènes codant l'isoforme AML-1L du facteur de

transcription hématopoïétique RUNX1 (runt-related transcription factor 1) et l'onco-protéine ETO (Eight-Twenty One). Le recrutement de PRMT1 par AML1-ETO stimule l'acétylation de H3K9 et H3K14, des marques épigénétiques associées à cette pathologie, et AE9a est elle-même faiblement méthylée par PRMT1 en R142. De plus, AE9a possède la capacité d'induire l'auto-renouvellement des cellules tumorales (Wang et al., 2011), mais cette propriété est bloquée dans des modèles murins déficients (KO) pour PRMT1 (Shia et al., 2012) si bien que le développement d'inhibiteurs ciblant PRMT1 pourrait être une stratégie de lutte contre cette leucémie.

La régulation de la transcription par PRMT1 n'implique pas nécessairement la diméthylation asymétrique de H4R3, mais peut intervenir en modulant les interactions protéine-protéine et protéine-acide nucléique de régulateurs de la transcription ou de partenaires qui leur sont associés, ou en modifiant leur localisation subcellulaire. Par exemple, PRMT1 interfère avec l'avancement de la transcription en méthylant le facteur d'élongation SPT5 (SuPressor of Ty 5), ce qui réduit son association avec l'ARN polymérase (Kwak et al., 2003). Inversement, la méthylation de TAF15 (Tata-binding protein Associated Factor 15) induit la relocalisation de TAF15 vers le noyau et stimule l'expression de ses gènes cibles en agissant sur l'étape d'activation de la transcription (Jobert et al., 2009). PRMT1 régule à la hausse l'expression de plusieurs gènes de cytokines en facilitant l'interaction entre la famille des facteurs de transcription du système immunitaire NFAT (Nuclear Factor of Activated T-cells) avec NIP45 (Nfat-Interacting Protein 45). Cette interaction est médiée par la méthylation de l'extrémité amino-terminale de NIP45, et le niveau de PRMT1 est lui-même augmenté par la stimulation de la voie de signalisation du récepteur des lymphocytes T (Mowen et al., 2004). PRMT1 est également recrutée par le récepteur nucléaire orphelin HNF4 (Hepatocyte Nuclear Factor 4) et méthyle son domaine de fixation à l'ADN. Cela augmenterait l'affinité de HNF4 pour ses éléments de réponse et stimulerait la transcription de gènes impliqués dans plusieurs voies métaboliques (Barrero et al., 2006). PRMT1 pourrait aussi moduler la transcription des gènes en régulant l'activité de protéines de la famille des « writers ». En effet, PRMT1 méthyle Ash2L (Absent, small or homeotic discs 2 Like), un membre commun des complexes de triméthylation de H3K4 dans les cellules de mammifères (Butler et al., 2011). La marque épigénétique H3K4^{me3} est associée à un état actif de la chromatine mais l'implication de PRMT1 dans la régulation de cette modification post-traductionnelle n'est pas encore élucidée. Néanmoins, il s'agit de la première modification post-traductionnelles observée sur cette protéine. Finalement, une fonction d'activation indépendante de l'activité enzymatique de PRMT1 a été reportée (Lei et al., 2009) : le récepteur nucléaire orphelin TR3 (Thyroid hormone Receptor) interagit avec PRMT1, ce qui a pour conséquences d'inhiber la méthylation de H4R3 dans des cellules HEK293T et de protéger TR3 contre la dégradation. Le récepteur nucléaire est ainsi disponible pour se fixer à ses éléments de réponse sur l'ADN et active la transcription en synergie avec le coactivateur des récepteurs nucléaires SRC-2 (Steroid Receptor Coactivator 2).

➤ **PRMT1 et l'intégrité du génome**

Le maintien de l'intégrité du génome est un mécanisme crucial pour la croissance cellulaire et pour la prévention du développement de cellules tumorales. Lors des différentes étapes du cycle cellulaire, des « checkpoints » doivent être validés pour permettre à la cellule de poursuivre son cycle ou bien d'entrer en apoptose. Un de ces « checkpoints » consiste à contrôler les dommages sur l'ADN. MRE11 (meiotic recombination 11) est un constituant du complexe MRN (Mre11-Rad50-Nbs1), impliqué dans la réparation des cassures double-brin de l'ADN. L'activité 3'-5' exonucléase ADN-double-brin dépendant de MRE11 fait intervenir un motif GAR et est assujettie à la méthylation de ce motif par PRMT1 (Boisvert et al., 2005a). De plus, l'hypométhylation du motif GAR perturbe le « checkpoint » des dommages sur l'ADN lors de la phase S, ce qui signifie que PRMT1 peut intervenir dans la validation ou le rejet du « checkpoint » et donc influencer le devenir cellulaire. PRMT1 méthyle également le motif GAR de 53BP1 (p53 Binding Protein 1), une protéine fixant le suppresseur de tumeur p53, ce qui induit le recrutement de 53BP1 sur les sites de cassures double-brin de l'ADN (Boisvert, Rhie, et al. 2005). 53BP1 joue un rôle dans la sélection du mécanisme de réparation des cassures double-brin par jonction des extrémités, rapide, ou par recombinaison homologue, plus lent mais plus efficace (Daley et al., 2014). En intervenant sur ces deux mécanismes, PRMT1 serait donc impliquée à la fois dans le choix de la voie de réparation et dans la modulation du mécanisme lui-même.

➤ **PRMT1 et la transduction du signal**

L'expression de nombreux gènes impliqués dans l'homéostasie et le cycle cellulaire est sous le contrôle de stimuli externes distants, tels que les hormones, ou de contacts, comme dans les jonctions intercellulaires ou les interactions avec la matrice. Ces messages sont transmis au cytoplasme puis au noyau par le biais de récepteurs transmembranaires ou nucléaires, dans le cas où l'hormone peut franchir la membrane plasmique. Ces signaux primaires déclenchent une cascade de signaux secondaires menant à l'activation ou à la répression de l'expression de gènes. Ces mécanismes sont appelés transduction du signal et un nombre croissant d'études impliquent PRMT1 dans la régulation de ces voies de signalisation.

Dans le cytoplasme, PRMT1 méthyle le domaine de fixation à l'ADN (DNA Binding Domain, DBD) du dimère de récepteur nucléaire ER α (Estrogen Receptor α), activé par la fixation de l'hormone œstrogène (Le Romancer et al., 2008). La méthylation du DBD de ER α permet son association avec le coactivateur transcriptionnel Src (Steroid receptor coactivator), appartenant à la famille des p160, ainsi qu'avec la sous-unité régulatrice p85 de PI3K (PhosphoInositide-3-Kinase), l'enzyme responsable de la synthèse du messenger cytoplasmique Phosphatidylinositol TriPhosphate (PIP3). Le complexe ER α ^{me2}-Src-PI3K active la voie de signalisation Akt, qui régule différents processus cellulaires (Song et al., 2005) dont la prolifération, la survie et la migration, et recrute la kinase d'adhésion focale FAK. En méthylant le DBD de ER α , PRMT1 permet donc de contrôler la

survie et la prolifération cellulaire en intervenant sur le mécanisme de signalisation rapide des œstrogènes. La dérégulation de ces mécanismes est corrélée au développement de tumeurs, et ER α est hyper-méthylé dans certains cancers du sein (Le Romancer et al., 2008). Dans les cellules leucémiques, PRMT1 stimule également l'activation de la MAP (Mitogene-Activated Protein) kinase p38 (Chang et al., 2010), décrite comme un probable répresseur de la différenciation, et permettrait donc de maintenir les cellules dans un état indifférencié. Le rôle joué par PRMT1 dans les voies de signalisations MAP et Akt souligne une fois de plus l'importance du développement d'inhibiteurs de PRMT1 dans le cadre de thérapies contre divers cancers.

b - PRMT2

➤ PRMT2 et la régulation de la transcription

PRMT2 diméthyle asymétriquement mais faiblement l'histone H4 *in vitro*, ce qui classe cette enzyme parmi les PRMT de type I, mais suggère que H4 n'est pas son substrat préférentiel (Lakowski et al., 2009). PRMT2 est impliquée dans la régulation de l'expression des gènes en tant que coactivateur des récepteurs nucléaires. En effet, l'interaction entre PRMT2 et le récepteur nucléaire de l'hormone œstrogène ER α stimule les fonctions d'activations (AF) ligand-indépendante AF1 et ligand-dépendante AF2 (Qi et al., 2002) portées respectivement par les domaines amino- et carboxy-terminaux de ER α (Wärnmark et al., 2003). L'interaction entre PRMT2 et ER α jouerait un rôle dans le développement du cancer du sein : dans les tissus concernés la sur-expression de variants d'épissages de PRMT2 module l'activation par ER α de l'activité de plusieurs promoteurs, dont celui du facteur de transcription SnaiI, activé, et celui de la E-cadhérine, réprimé (Zhong et al., 2011). Le facteur de transcription SnaiI réprime lui-même l'expression de la E-cadhérine (Batlle et al., 2000), une molécule d'adhésion cellule-cellule dont la perte est associée avec la fin de l'inhibition de contact et l'apparition des métastases (Jeanes et al., 2008). PRMT2 interagit également avec le récepteur nucléaire de l'androgène (AR) activé par l'hormone et tous deux sont transloqués du cytoplasme vers le noyau, où PRMT2 coactive les gènes cibles de AR (Meyer et al., 2007).

À l'inverse de son rôle de coactivateur de récepteurs nucléaires, des études suggèrent que PRMT2 participe à l'inhibition de la transcription par le biais d'interactions avec les répresseurs I κ B α et RB. Le facteur de transcription des gènes anti-apoptotiques NF κ B (Nuclear Factor kappa B) est transloqué et séquestré dans le cytoplasme par son interaction avec I κ B α (Inhibitor of nf- κ B α), qui masque son signal de localisation nucléaire (Rice et al., 1993). Par ce mécanisme et par inhibition de la fixation de NF κ B à ses éléments de réponse sur l'ADN (Baeuerle et al., 1988), I κ B α réprime l'expression des gènes anti-apoptotiques. Dans des cellules HEK293 stimulées par l'hormone TNF α (Tumor Necrosis Factor α), l'expression de PRMT2 empêche l'export de I κ B α vers le cytoplasme par l'exportine CMR1 (Chromosomal Maintenance 1), ce qui induit

l'accumulation de I κ B α dans le noyau, empêche la fixation de NF κ B à ses éléments de réponse et facilite l'apoptose. Le mécanisme par lequel PRMT2 bloque l'export de I κ B α n'a pas encore été identifié. PRMT2 réprime également l'expression de gènes via son interaction avec RB : le facteur de transcription E2F-1 (E2 Factor 1) stimule l'activité des gènes d'entrée en phase S (Bertoli et al., 2013) et peut être inactivé par la fixation du répresseur du cycle cellulaire RB (RetinoBlasoma). L'interaction entre PRMT2 et RB induit la répression de l'activité transcriptionnelle de E2F1, ce qui se traduit par l'entrée précoce des cellules en phase S dans un modèle de fibroblaste embryonnaire murin PRMT2^{-/-} (Yoshimoto et al., 2006).

PRMT2 jouerait également un rôle dans la régulation de la transcription en aval de plusieurs voies de signalisations. Dans le mécanisme de contrôle des réserves de graisse, la fixation de l'hormone leptine à son récepteur transmembranaire active une voie de signalisation JAK/STAT : STAT3 (Signal Transducer and Activator of Transcription 3) est activé par phosphorylation par une enzyme JAK (Janus Kinase) associée au récepteur de la leptine, ce qui induit sa translocation dans le noyau suivi de sa fixation à ses éléments de réponse sur l'ADN (Matsuoka et al., 1999 ; Rawlings, 2004). PRMT2 méthyle STAT3 sur R31, ce qui réduirait la durée de son activation par phosphorylation (Iwasaki et al., 2010) et donc réduirait l'amplitude de l'activation des gènes cibles de STAT3. Les modèles murins KO pour PRMT2 sont maigres, se nourrissent insuffisamment et présentent de faibles taux sanguins en leptine, ce qui suggère que la conception d'inhibiteurs de PRMT2 pourrait être une stratégie pour le développement de thérapies contre l'obésité. De plus, la réponse cellulaire anti- ou pro-inflammatoire aux interleukines dépendrait aussi de la durée de l'activation de STAT3 (Braun et al., 2013), ce qui suggère que PRMT2 pourrait jouer un rôle dans la régulation de la réponse inflammatoire à la fois par son interaction avec I κ B α et par la méthylation de STAT3. PRMT2 serait également un effecteur de la voie de signalisation Wnt/ β -caténine. En effet, chez le xénope, la β -caténine, coactivateur des facteurs de transcription de la famille TCF/LEF et constituant de la voie de signalisation wnt (Cadigan, 2012 ; Clevers et al., 1997 ; Hsu et al., 1998), recruterait PRMT2 au niveau de ses promoteurs cibles. À ces sites, la chromatine est enrichie en marquage H3R8, nécessaire à l'établissement de l'architecture dorsale lors du développement embryonnaire (Blythe et al., 2010).

➤ **PRMT2 et la régulation de l'épissage**

Lors de cette thèse, nous avons identifié le facteur d'épissage rsf1 comme premier substrat non-histone de PRMT2, en lien avec son rôle dans la régulation de l'épissage (Cura et al., 2017). L'analyse structurale de PRMT2 et de son complexe formé avec rsf1 sera abordée dans le chapitre III. L'implication de PRMT2 dans le mécanisme d'épissage a été renforcée par la caractérisation de son interaction avec le facteur d'épissage SAM68 (Src substrate Associated during Mitosis 68). SAM68 module l'épissage alternatif des isoformes Longues (L) et courtes (s) du régulateur de l'apoptose BCL-X (B Cell Lymphoma eXtra Large/small). Les BCL-X L/s sont deux variants d'une

protéine membranaire mitochondriale et son respectivement pro- ou anti-apoptotiques (Paronetto et al., 2007). PRMT2 interagit par l'intermédiaire de son domaine SH3 avec SAM68 ce qui module la localisation subcellulaire du facteur d'épissage. De cette manière, l'expression de PRMT2, dans des cellules stimulées par TNF α ou par un lipopolysaccharide, induit une réponse anti-apoptotique en favorisant l'augmentation du rapport BCL-X(L)/BCL-X(s) (Vhuiyan et al., 2017).

c - PRMT3

PRMT3 est une PRMT de type I principalement cytoplasmique (Tang et al. 1998), généralement considérée comme une PRMT associée au ribosome dont le seul substrat identifié est la protéine de la petite sous-unité du ribosome, rpS2 (ribosomal protein S2). Dans *S. pombe* et chez les mammifères, PRMT3 interagit avec la sous-unité 40S du ribosome par l'intermédiaire de l'interaction entre son domaine amino-terminal et la protéine ribosomique S2 (rpS2), et il a été établi que PRMT3 méthyle l'extrémité amino-terminale de rpS2 sur un motif GAR (Bachand et al., 2004 ; Swiercz et al., 2005). L'interaction entre PRMT3 et rpS2 serait conservée chez la plupart des eucaryotes, à l'exception de *S. cerevisiae* qui ne possède pas d'homologue de PRMT3 et où la méthylation de rpS2 est effectuée par RMT1, homologue de PRMT1 (Lipson et al., 2010). L'interaction PRMT3-rpS2 participe à la biogenèse de la sous-unité 40S du ribosome sans que l'activité méthyltransférase de PRMT3 n'intervienne, si bien que l'importance de la méthylation sur rpS2 n'est pas encore comprise (Bachand et al., 2004 ; Perreault et al., 2009). L'interaction PRMT3-rpS2 protégerait également la protéine ribosomique contre l'ubiquitylation suivie de la dégradation par le protéasome et rpS2 augmenterait modestement l'activité enzymatique de PRMT3 (Choi et al., 2008 ; Miyata et al., 2010). PRMT3 influencerait également la biogenèse du ribosome en intervenant sur les voies de maturation du pre-ARN ribosomique 18S, tel que démontré chez les plantes (Hang et al., 2014). RpS2 est hypométhylée dans des embryons murins PRMT3^{-/-}. Ces embryons sont petits comparés aux contrôles, mais viables (Swiercz et al., 2007). Dans des neurones déficients en PRMT3, les épines dendritiques sont difformes et on observe une diminution de la synthèse d' α CAMKII, une protéine kinase contrôlant la plasticité cérébrale, dont l'expression est induite par BDNF (Brain-Derived Neurotrophic Factor) (Miyata et al., 2010). C'est une indication supplémentaire de l'implication de PRMT3 dans le fonctionnement du ribosome.

Dans *S. pombe*, rpS2 est nécessaire à l'export du pre-40S, le complexe ribonucléoprotéique précurseur de la petite sous-unité 40S, des nucléoles vers le cytoplasme (Perreault et al., 2008). PDCD2 et son paralogue PDCD2L (Programmed Cell Death 2/Like) ont récemment été décrits comme s'associant avec les pre-40S par le biais d'interactions avec rpS2 (Landry-Voyer et al., 2016 ; Minakhina et al., 2016), et participent à la maturation tardive du ribosome dans le cytoplasme. De plus, PDCD2L fait la navette entre le cytoplasme et le noyau, par l'intermédiaire de l'exportine CMR1, et forme un complexe ternaire extra-ribosomique avec rpS2 et PRMT3 (Landry-Voyer et al.,

2016). Ces éléments suggèrent que rpS2 est exportée seule ou avec le pre-40S des nucléoles vers le cytoplasme, avec l'intervention de PDCD2 et PDCD2L, servant de chaperon pour la translocation à travers la paroi nucléaire par l'exportine CMR1. Dans le cytoplasme, PRMT3 interagit avec le complexe rpS2-PDCE2L et avec la petite sous-unité du ribosome en fixant rpS2, qu'il stabilise et méthyle sur le motif GAR. La suite est plus mystérieuse : comme PRMT3 est nucléaire mais agirait sur la maturation des ARN ribosomiques dans le noyau, on peut émettre deux hypothèses : (I) le complexe PRMT3-rpS2-PDCD2L est transloqué temporairement dans le noyau où PRMT3 agit sur la maturation des ARNr, (II) c'est un substrat modifié par PRMT3 qui est transloqué dans le noyau et module la maturation des ARNr. Cette seconde hypothèse est plus vraisemblable puisque PRMT3 n'est pas observée dans le noyau. Le seul substrat identifié de PRMT3 étant rpS2, il est tentant de proposer que ce soit rpS2 qui jouerait le rôle d'intermédiaire entre PRMT3 et la maturation des ARNr. Cette idée est renforcée par le fait que dans *S. cerevisiae*, l'homologue de PDCD2L, Tsr4P, est nécessaire pour la maturation du pre-ARNr 20S en ARNr 18S (Li et al. 2009).

Finalement, rpS2 est sur-exprimée dans des tumeurs humaines (Chiao et al., 1992) et dans un modèle murin de tumeur des abajoues (Shin et al., 1993). De plus, PRMT3 interagit *in vitro* et *in situ* avec le répresseur de tumeur DAL-1 (Differentially expressed in Adenocarcinoma of the Lung protein 1), ce qui inhibe l'activité enzymatique de PRMT3, et DAL-1 est perdu dans plus de la moitié des tumeurs des cancers du poumon hors tumeurs à petites cellules (Singh et al., 2004). Ces éléments désignent PRMT3 comme une cible potentielle pour le développement de nouveaux traitements contre le cancer (Wang, Hu, et Stearns 2009). PRMT3 stabilisant rpS2, cibler cette interaction pourrait représenter une approche thérapeutique.

d - PRMT4/CARM1

PRMT4 est de loin l'un des membres de la famille des PRMT les plus étudiés du fait de sa fonction de coactivateur des récepteurs nucléaires (Chen, 1999), et son association avec des coactivateurs lui a valu d'être nommée CARM1 pour Coactivator Associated aRginine Méthyltransferase 1. CARM1 diméthyle asymétriquement H3R17 et H3R26 et se classe ainsi parmi les PRMT de type I (Schurter et al., 2001). Les fonctions biologiques de CARM1 sont associées à son activité enzymatique et ne peuvent pas être restituées par ses orthologues. Le remplacement de CARM1 par une version mutée inactive dans des lignées cellulaires ou un modèle murin mène en effet aux mêmes conséquences qu'un knock-out (Kim et al., 2010a).

➤ CARM1 et la régulation de la transcription

CARM1 a été identifiée à la fin des années 90 et a rapidement été caractérisée comme coactivateur des récepteurs nucléaires. La famille des coactivateurs des récepteurs nucléaires p160/SRC (Steroid Receptor Coactivator) est composée de trois membres (SRC-1, SRC-2 et SRC-3) qui interagissent principalement avec les récepteurs nucléaires mais également avec d'autres facteurs de

transcription. Les SRC coactivent la transcription des gènes en recrutant d'autres coactivateurs arborant une activité histone-acétyltransférase ou -méthyltransférase (Xu et al., 2003). Dans les cellules de mammifère, CARM1 et l'acétyltransférase d'histone p300/CBP (Creb Binding Protein) sont recrutées par les p160/SRC au niveau de récepteurs nucléaires activés par la fixation d'hormone stéroïdienne (Chen, 1999 ; Chen et al., 2000). p300 acétyle H3 en plusieurs positions dont K18, ce qui stimule la méthylation de H3R17 par CARM1 et conduit à l'augmentation de l'activation transcriptionnelle par les récepteurs nucléaires (Bauer, 2002 ; Daujat et al., 2002). Le recrutement de CARM1 et p300/CBP par p160 est médié par des interactions entre les domaines AD1 (Activation Domain) et AD2 de p160 avec respectivement p300/CBP et le module catalytique de CARM1 (Teyssier et al., 2002). Un modèle à basse résolution du complexe ER α -SRC-3-p300/CBP-CARM1 obtenu par cryo-microscopie électronique (Yi et al., 2017) a récemment permis de mieux comprendre le mécanisme structural de la coactivation par CARM1 (figure 8) : en réponse à la fixation d'hormone, le dimère d'ER α fixé à son élément de réponse recrute deux p160/SRC-3, qui eux-mêmes font office de plateforme pour la fixation de p300/CBP. p300 catalyse alors la production d'un niveau basal d'acétylation de lysines sur H3. CARM1 est alors recruté par le complexe (ER α)₂-(p160/SRC-3)₂-(p300/CBP)₁ et déplace un des SRC-3. Cela entraîne un réarrangement structural de p300 et stimule son activité d'acétylation. L'augmentation de H3K18^{Ac} stimule en retour la méthylation de H3R17 par CARM1 et l'activation de la transcription.

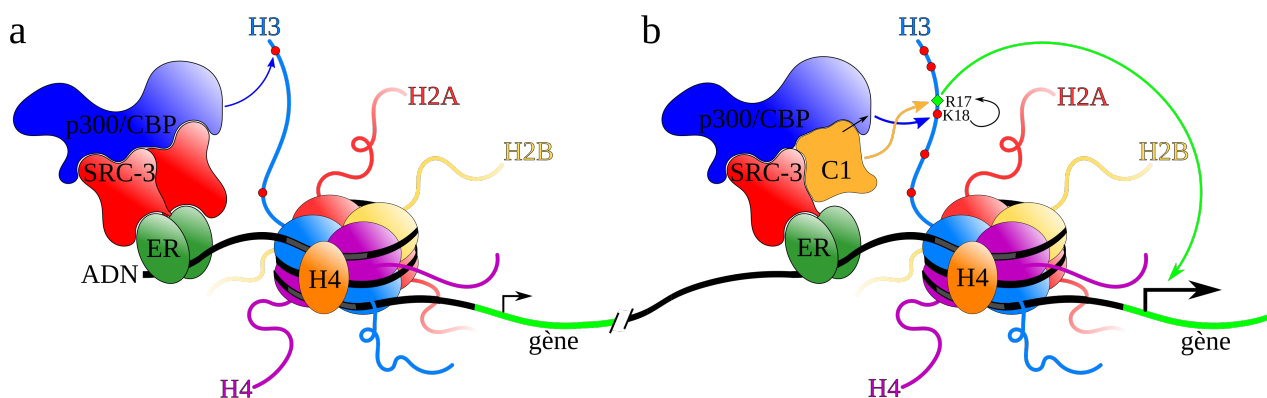


Figure 8 - Coactivation de la transcription par les récepteurs nucléaires par CARM1 et p300/CBP.

Figure adaptée de (Yi et al., 2017). (a) Le complexe p160/SRC-3-p300/CBP est recruté au niveau du promoteur par les récepteurs nucléaires ayant fixés une hormone stéroïdienne. Un niveau basal d'acétylation de H3 est déposé par p300. (b) CARM1 (C1) remplace un des SRC-3, induisant la stimulation de l'acétylation de H3 par p300. L'acétylation de H3K18 stimule alors la méthylation de H3R17 par CARM1, ce qui permet l'activation de la transcription.

Le mécanisme coopératif d'activation de la transcription par (p160/SRC-3)-p300/CBP-CARM1 est étendu à d'autres complexes de régulation de la transcription, où les p160/SRC-1, -2 et -3 joueraient le rôle de plateforme pour adapter CARM1 et p300 à d'autres récepteurs nucléaires ou facteurs de transcription, modulant ainsi la spécificité de CARM1 pour une variété de processus cellulaires. En effet, CARM1 est également recrutée par SRC-2/GRIP1 et agit en synergie avec p300/CBP pour coactiver l'activité transcriptionnelle des récepteurs aux œstrogènes (Chen et al.,

2000). Lors de la différenciation des adipocytes et dans la régulation de l'équilibre énergétique, l'activité transcriptionnelle du récepteur nucléaire PPAR γ serait en partie coactivée par un complexe p160/SRC-p300/CBP-CARM1 (Yadav et al., 2008). Cette hypothèse est appuyée par le fait que les embryons murins KO pour cette PRMT ont moins de réserves graisseuses, sont petits et meurent précocement avant leur naissance (Yadav et al., 2008). Dans leurs tissus, l'expression des gènes sous contrôle des œstrogènes est perturbée et les protéines p300/CBP et PABP1 (PolyA Binding Protein 1) sont hypométhylées (Yadav et al., 2003), ce qui indique que ces deux protéines sont également des substrats de CARM1. Alors que CARM1 est essentiellement un coactivateur transcriptionnel, la méthylation de p300/CBP par CARM1 inhibe le recrutement de CREB, un facteur de transcription cAMP-dépendant, par p300. Cet antagonisme caractérise CARM1 comme un corépresseur de la voie de signalisation par le cAMP (Xu, 2001). CARM1 serait également impliquée dans le contrôle du cycle cellulaire et le maintien de l'intégrité du génome par le biais de deux mécanismes. Dans le premier mécanisme, CARM1 méthyle H3R17 et H3R26 au niveau du promoteur du gène de la cycline E1 (CE1), en présence du facteur de transcription E2F et du coactivateur p160/SRC-3/ACTR (ACTivator of Thyroid hormone Receptor) (El Messaoudi et al., 2006). Cela stimule l'expression de CE1, ce qui est associé à une transition de la phase G1 du cycle cellulaire vers la phase S (Ohtsubo et al., 1995). Dans des tumeurs mammaires de grade 3, une expression des transcrits de CARM1, CE1 et ACTR est observée (El Messaoudi et al., 2006). Lors du second mécanisme, CARM1, p300, et parfois PRMT1 agissent en synergie avec le facteur de transcription et suppresseur de tumeur p53 pour méthyle et acétyler H3 au niveau des promoteurs des gènes cibles de p53, stimulant ainsi leur expression (An et al., 2004). Le rôle des p160 dans l'interaction avec p53 n'est aujourd'hui pas défini. CARM1 jouerait aussi un rôle dans la réponse inflammatoire : un complexe formé par CARM1, p300/CBP et p160/SRC-1 est recruté par le facteur de transcription des gènes pro-inflammatoires NF κ B au niveau de ses éléments de réponse sur la chromatine, où le complexe CARM1-CBP-SRC-1 dépose les trois marques H3R17^{me2}, H3K9^{Ac}, H3K14^{Ac}, ce qui se traduit par l'activation de la transcription des gènes cibles de ce facteur de transcription. Une diminution de la citrullination de H3 est également observée sur ces sites. NF κ B interagit directement avec CARM1 par l'intermédiaire de sa sous-unité p65. Ces modifications sont également observées au niveau des promoteurs de gènes contrôlant l'inflammation dans des monocytes (Covic et al., 2005 ; Lawrence, 2009 ; Miao et al., 2006).

CARM1 faciliterait l'expression des gènes en aidant à maintenir l'accessibilité de la chromatine aux facteurs de transcription : la désacétylation des histones par le complexe NurD et le corépresseur transcriptionnel TIF1 inhibe l'expression des gènes en remodelant la chromatine d'un état décondensé et accessible vers une forme condensée. La méthylation de H3 par CARM1, *in vitro* comme *in situ*, empêche la fixation de ces deux complexes à la chromatine protégeant les histones de la désacétylation et facilitant la transcription des gènes (Wu et al., 2012).

Plus récemment, une étude a lié CARM1 au maintien en latence du VIH-1. Le complexe de super-élongation (SEC) permet la transcription des gènes du VIH-1, nécessaire à sa sortie de latence. Le SEC est recruté par l'acétylation de H3K27 par p300/CBP. Parallèlement, cette modification post-traductionnelle stimule la diméthylation de H3R26 par CARM1, ce qui a pour effet inverse de prévenir la fixation du SEC. Une étude a montré que l'inhibition de CARM1 rétablissait le recrutement du SEC et donc permettait une sortie de latence du VIH, ce qui ouvre la voie à des thérapies de type « shock-and-kill » ou des inhibiteurs de latence permettraient d'induire l'expression du VIH (Virus de l'Immunodéficience Humaine) afin de détruire ses cellules réservoirs dans des patients stables sous trithérapie (Zhang et al., 2017b).

➤ **CARM1 dans la régulation de l'épissage alternatif et de la maturation des ARNs**

Outre l'activation de la transcription, CARM1 est associée à la régulation de l'épissage et la maturation des ARN messagers. Un des substrats non-histone bien caractérisé de CARM1 est la Polyadenylate-binding protein 1 (PABP1), avec deux sites de méthylation en R455 et R460 (Lee, 2002). PABP1 se fixe à la queue polyA des ARN messagers et permet leur circularisation en interagissant avec le complexe d'initiation de la traduction eIF4F (Imataka et al., 1998). Cette méthylation a été identifiée *in vitro* puis confirmée *in vivo* mais son importance n'a pas encore été expliquée.

Une étude par criblage a permis d'identifier d'autres substrats de CARM1 engagés dans la régulation de l'épissage (Cheng et al., 2007) : l'inhibiteur de l'élongation CA150, SAP49 un constituant du facteur d'épissage SF3b et SmB.

Le spliceosome, un complexe ribonucléoprotéique nucléaire responsable de l'épissage des ARN messagers, s'assemble suivant une succession d'étapes nucléaires et cytoplasmiques (Fischer et al., 1997 ; Liu et al., 1997 ; Urlaub et al., 2001). Dans le cytoplasme, sept protéines Sm (SmB, SmD1, SmD2, SmD3, SmE, SmF et SmG) sont assemblées par le complexe SMN (Survival of Motor Neuron) sous forme d'un disque autour des snRNA (small nuclear RNA) U1, U2, U4 ou U5, pour former les snRNP (small nuclear RiboNucleoProtein) U1 à U5. Les snRNP-U sont alors transloqués dans le noyau où ils s'assemblent sur l'ARN pré-messager autour des sites d'épissages, pour former le spliceosome (Chiou et al., 2014 ; Will et al., 2011). La méthylation des Sm par PRMT5 et PRMT7 est nécessaire pour l'assemblage du spliceosome, mais le rôle de CARM1 n'est pas encore élucidé. L'assemblage du spliceosome fait intervenir le complexe SMN, qui interagit via son domaine TUDOR avec CA150, un autre substrat de CARM1 (Cheng et al., 2007). Cette interaction semble nécessiter l'activité catalytique de CARM1 puisque SMN et CA150 ne s'associent plus *in vitro* lorsque CA150 est issu de cellules déficientes en CARM1.

Outre la modulation de l'assemblage du spliceosome, CARM1 sembler jouer un rôle directement dans la régulation du processus d'épissage : dans certains tissus de rats, l'isoforme V3

de CARM1, ne possédant pas de domaine amino-terminal mais des séquences spécifiques carboxy-terminales, interagit avec U1C (Ohkura et al., 2005). U1C est une protéine reconnaissant le site d'épissage en 5' des ARN prés-messagers et nécessaire à l'assemblage du spliceosome (Heinrichs et al., 1990 ; Rossi et al., 1996). La fixation de CARM1-V3 à U1C affecte la sélection du site d'épissage (Ohkura et al., 2005). De plus, CARM1 promeut le saut d'exon de manière activité-dépendant dans des fibroblastes embryonnaires murins (Cheng et al., 2007).

➤ **CARM1 et le développement de cancers**

CARM1 participe au contrôle de la croissance et de la différenciation cellulaire : cette PRMT est nécessaire à la différenciation des cellules musculaires (Chen et al., 2002). CARM1 stimule également la pluripotence dans des blastomères murins (Torres-Padilla et al., 2007), et dans des cellules souches embryonnaires de souris, la méthylation de H3R17 et H3R26 retarde la différenciation (Wu et al., 2009). La dérégulation de ces processus est associée au développement de tumeurs et de nombreuses études ont mis en évidence l'implication de CARM1 dans le développement de certains cancers. CARM1 agit notamment comme coactivateur en aval de la voie de signalisation wnt/ β -caténine. Mécanistiquement, CARM1 fixe la β -caténine et, en synergie avec p300, stimule l'activation du récepteur aux androgènes (Koh et al., 2002). La β -caténine est capable d'interagir directement avec AR et augmente son activation transcriptionnelle ligand-dépendante dans des cellules du cancer de la prostate (Yang et al., 2002). CARM1 est également recrutée par la β -caténine au niveau des promoteurs des gènes Wnt-spécifique où elle diméthyle H3R17 et régule à la hausse l'expression des gènes cibles de la β -caténine. Dans des cellules tumorales du cancer colorectal, la suppression de CARM1 inhibe la prolifération cellulaire ancrage-indépendante et la survie cellulaire, ce qui impliquerait CARM1 dans le développement de métastases (Ou et al., 2011). CARM1 est également sur-exprimée dans le cancer colorectal et est déterminante pour la différenciation et la prolifération des cellules cancéreuses du cancer du sein ER α -dépendant (Al-Dhaheri et al., 2011 ; Kim et al., 2010b). L'implication de CARM1 dans le développement des cancers en fait une cible pour le développement d'inhibiteurs.

e - PRMT6

PRMT6 a été identifiée par homologie de séquence sur le chromosome I humain en 2002. Il s'agit d'une PRMT de type I principalement nucléaire (Frankel et al., 2002).

➤ **Régulation de la transcription par PRMT6**

PRMT6 est largement impliquée dans la régulation de l'expression des gènes et fonctionne principalement comme un répresseur transcriptionnel. PRMT6 méthyle l'histone H3 où elle appose le marquage asymétrique H3R2^{me2a}, ce qui est associé à une faible activité transcriptionnelle et classe donc PRMT6 parmi les corépresseurs de la transcription (Hyllus et al., 2007).

Mécanistiquement, le marquage H3R2^{me2a} antagonise la synthèse de la modification H3K4^{me3}, associée à un état actif de la chromatine, par le complexe de méthylation de H3K4 Ash2/WDR5/MLL (Guccione et al., 2007). WDR5 possède un domaine WD40 dont la fixation sur H3 permet l'ancrage du complexe de méthylation de H3K4 sur la chromatine. Les structures cristallographiques du complexe WDR5-histone H3 (PDB 2H13) montrent que cette interaction fait intervenir le résidu H3R2 non méthylé, et que la poche où se fixe l'arginine ne permet pas d'accommoder une arginine diméthylée asymétriquement (Couture et al., 2006). La modification de H3R2 par PRMT6 inhibe la fixation à la chromatine d'autres régulateurs transcriptionnels possédant des modules d'interaction à H3, incluant des domaines TUDOR et PHD (Iberg et al., 2008). Finalement, PRMT6 coopérerait aussi avec le complexe Polycomb Repressive 2 (PCR2), responsable du marquage H3K27^{me3}, dans la répression de la transcription des gènes selon un mécanisme non encore élucidé (Stein et al., 2016). H2AR29 est également une cible de PRMT6 *in vivo*. Ce marquage est lui aussi associé à un état inactif des gènes concernés (Waldmann et al., 2011).

PRMT6 aurait également un rôle de coactivateur transcriptionnel. Dans les cellules hépatiques, PRMT6 méthyle sur plusieurs sites CRTC2 (CREB Regulated Transcription Coactivator 2), coactivateur du facteur de transcription CREB. Cela stimule l'interaction entre CRTC2 et CREB au niveau des promoteurs de gènes métaboliques et augmente leur activité transcriptionnelle (Han et al., 2014). PRMT6 participerait donc à la régulation du métabolisme énergétique, une hypothèse consolidée par l'observation que PRMT6 est enrichie dans les cellules hépatiques de modèles murins de l'obésité (Han et al., 2014).

➤ **PRMT6 et l'intégrité du génome**

PRMT6 est impliquée dans le maintien de l'intégrité du génome via son interaction avec l'ADN polymérase β (El-Andaloussi et al., 2006). Cette polymérase est responsable de la réparation de l'ADN par excision de base azotée (Beard et al., 2006). PRMT6 fixe et méthyle l'ADN polymérase β sur R83 et R152 *in vitro* et *in vivo*. Ces modifications augmentent la fixation de la polymérase à l'ADN et améliore sa processivité (El-Andaloussi et al., 2006), facilitant le mécanisme de réparation de l'ADN.

➤ **PRMT6 et le développement de cancers**

PRMT6 est sur-exprimée dans de nombreuses cellules cancéreuses (Yang et al., 2013) induisant un enrichissement en H3R2^{me2a}. Ce marquage inhibe la fixation du domaine PHD de UHRF1 (Ubiquitin-like, containing PHD and RING Finger domains), qui dans les cellules saines recrute la méthyltransférase de l'ADN DNMT1 (DNA méthylTransférase 1). Cela induit une hypométhylation de l'ADN environnant les histones modifiées H3R2^{me2a} (Veland et al., 2017), une caractéristique retrouvée dans de nombreux cancers (Ehrlich, 2009). De plus, la méthylation de

H3R2 par PRMT6 réprime l'expression de l'inhibiteur de l'angiogenèse thrombospondin-1 (Michaud-Levesque et al., 2009), or la vascularisation est un facteur déterminant de la croissance des tumeurs. La conception d'inhibiteurs sélectifs de PRMT6 représente ainsi un enjeu pour le développement de thérapies anti-cancéreuses.

➤ **PRMT6 et la résistance aux rétrovirus**

Finalement, un nombre croissant d'études lient l'activité de PRMT6 à la résistance aux rétrovirus. PRMT6 diméthyle le coactivateur transcriptionnel du VIH, TAT, *in vitro* et *in vivo* sur les résidus R52 et R53. Cela exclu TAT hors des nucléoles (Fulcher et al., 2016) et de ce fait perturbe son interaction avec la région de trans-activation TAT-spécifique, et interfère avec la formation du complexe TAT-cycline T1, elle-même activatrice de TAT. Dans des cellules en culture déficientes pour PRMT6, la réplication virale est plus rapide et la quantité de virion produite plus grande (Xie et al., 2007). L'auto-méthylation de PRMT6 en R35 semble nécessaire à son activité anti-HIV ainsi qu'à sa stabilité (Singhroy et al., 2013).

f - **PRMT8**

PRMT8 est une des PRMT de type I les moins bien caractérisées. Elle est un paralogue de PRMT1 et a été identifiée en 2005 par homologie de séquence sur le chromosome 12 humain (Lee et al., 2005a). PRMT8 possède deux propriétés uniques dans cette famille d'enzymes. Premièrement, son extrémité amino-terminale serait myristoylée, ce qui induirait l'ancrage de PRMT8 à la membrane plasmique. Ensuite, cette enzyme est exprimée principalement dans les tissus cérébraux (Lee et al., 2005a). La myristoylation de PRMT8 et sa localisation sur la membrane plasmique sont toutefois contestées : alors que la première caractérisation de PRMT8 avait été effectuée sur une enzyme produite de manière exogène à partir du premier codon initiateur possible dans la séquence (Lee et al., 2005a), faisant apparaître un site de myristoylation, une seconde étude *in situ* sur une PRMT8 endogène a démontré à l'inverse une localisation nucléaire dans les neurones du cervelet, du cortex et de l'hippocampe (Kousaka et al., 2009). Une localisation nucléaire et non membranaire n'est pas compatible avec la présence d'une myristoylation.

➤ **PRMT8 et le développement du système nerveux**

Plusieurs études impliquent PRMT8 dans le développement et l'intégrité du système nerveux central. PRMT8 est exprimée dans les cellules souches pluripotentes de souris sous contrôle de facteurs de transcription de la pluripotence, où elle régule la différenciation des précurseurs des neurones (Solari et al., 2016). PRMT8 serait également enrichie au niveau du promoteur du gène de la tenascin-R (Lee et al., 2017), une protéine de la matrice extracellulaire cérébrale impliquée dans l'adhésion, la différenciation et la migration cellulaire (Anlar et al., 2012). La fixation de PRMT8 à ce promoteur participe à l'intégrité des connections synaptiques, ce qui suggère que PRMT8 est, à l'image d'autre PRMT, un régulateur épigénétique (Lee et al., 2017). La

dérégulation de l'expression de la tenascin-R est associée à des maladies neurodégénératives et au développement de tumeurs. Dans des souris déficientes en PRMT8, on observe de plus un déséquilibre des niveaux en phospholipides membranaires, des altérations des structures cérébrales, du développement des neurones et du fonctionnement et de la composition des synapses. Ces modèles murins souffrent d'hyperactivité et de comportement moteurs anormaux (Kim et al., 2015 ; Penney et al., 2017).

➤ **PRMT8 est sur-exprimée dans plusieurs cancers**

Le seul substrat connu de PRMT8 est la protéine du sarcome d'Ewing (EWS), identifiée par « pull-down ». PRMT8 se fixe sur cette protéine au niveau de sa région RGGb box 3 amino-terminal et méthyle *in vitro* cette même région, sans que leur interaction n'en soit altérée (Pahlich et al., 2008). L'influence de l'interaction entre PRMT8 et EWS sur l'apparition de ce cancer n'est cependant aujourd'hui pas connue. Bien que cerveau-spécifique, PRMT8 est aussi sur-exprimée dans les cancers du poumon, du tube digestif et des ovaires. On la retrouve également sur-exprimée dans le cancer cervical et la sur-expression de PRMT8 dans le cancer du tube digestif indique un pronostic vital défavorable (Hernandez et al., 2017).

g - PRMT5

PRMT5 a été identifiée en 1999 et initialement appelée JBP1 (Janus kinase Binding Protein 1) du fait de son interaction avec la tyrosine kinase JAK2 (Pollack et al., 1999a). Deux ans plus tard, deux études permettaient de mieux caractériser PRMT5 en montrant que cette enzyme est principalement cytoplasmique, forme des homo-oligomères *in vivo* (Rho et al., 2001), et catalyse la formation d'arginines symétriquement diméthylées, ce qui classe cette enzyme comme la première PRMT de type II identifiée chez les mammifères (Branscombe et al., 2001).

➤ **PRMT5 et la formation des snRNPs**

PRMT5 joue un rôle dans la régulation de l'assemblage assisté des snRNPs lors du processus de biogenèse du spliceosome : dans les cellules, PRMT5 est retrouvé dans un complexe appelé méthylosome associé aux protéines MEP50/WDR77/WD45 et pICln (Westley J. Friesen et al. 2002; W. J. Friesen et al. 2001), une protéine chaperon impliquée dans l'assemblage des snRNPs (Chari et al., 2008 ; Pu et al., 1999) et qui empêche l'assemblage spontané des protéines Sm sur les snRNA-U *in vitro*. pICln recrute les protéines Smd1 et Smd3 sur le méthylosome où celles-ci se fixent à MEP50 (Friesen et al., 2002) puis sont méthylées par PRMT5 (Friesen et al., 2001a). Leur méthylation permet aux protéines Smd1 et Smd3 d'être reconnues puis recrutées par le domaine TUDOR du complexe SMN (Brahms et al., 2001 ; Friesen et al., 2001b), ce qui stimule l'assemblage du noyau Sm des snRNPs par SMN (Meister, 2002) et permet l'assemblage subséquent des snRNPs (Gonsalvez et al., 2007).

➤ Répression de la transcription par PRMT5

PRMT5 agirait comme corépresseur transcriptionnel principalement en méthylant des cibles histones, en particulier l'histone H4 sur R3 et l'histone H3 sur R8 (Pollack et al., 1999b ; Zhao et al., 2009). Dans le noyau des cellules de mammifères, PRMT5 fixe COPR5, ce qui permet le recrutement de ce complexe au niveau des nucléosomes et stimule la méthylation spécifique de H4R3, mais pas de H3R8 (Lacroix et al., 2008). À l'instar de MEP50 qui permet la fixation des protéines Sm, COPR5 jouerait donc un rôle d'adaptateur et de régulateur de la spécificité de substrat de PRMT5. La méthylation de H4R3 est reconnue par le domaine PHD de l'ADN méthyltransférase DNMT3A, ce qui induit la méthylation de l'ADN environnant et la répression de la transcription (Zhao et al., 2009). Les bases moléculaires de la répression de la transcription associée à la méthylation de H3R8 par PRMT5 ne sont pas bien identifiées, mais impliqueraient le facteur de remodelage de l'ADN ATP-dépendant SWI/SNF (Pal et al., 2004). PRMT5 s'associe effectivement avec des complexes SWI/SNF et méthyle H3R8 et H4R3.

De par sa fonction de coactivateur de la transcription, PRMT5 est impliquée dans la régulation de plusieurs processus biologiques. PRMT5 jouerait un rôle dans le contrôle du déroulement du cycle cellulaire (Fabrizio et al., 2002). Cette PRMT est en effet retrouvée associée au complexe CERC (Cyclin E repression Complex) sur le site d'initiation de la transcription du promoteur du gène de la cycline E, une région riche en histones H4 méthylés sur les arginines. La sur-expression de PRMT5 inhibe l'activité du promoteur de CE1 et bloque la prolifération cellulaire. Plusieurs études associent aussi PRMT5 à la régulation du cycle circadien. Les protéines CRY1 (Cryptochrome 1) et PER1 (Periode 1) sont deux constituants d'un large complexe multi-protéique qui, en réprimant l'expression des gènes codant pour ses propres sous-unités, permet d'établir un oscillateur biologique d'une période d'environ vingt-quatre heures (Aryal et al., 2017). Dans les cellules humaines, CRY1 interagit avec PRMT5 ce qui permet son recrutement au niveau du promoteur du gène codant la protéine PER1 (Na et al., 2012). Les nucléosomes sont enrichis en marquage H4R3^{me2s} sur ces sites, régulant à la baisse l'expression de PER1. Chez *A. thaliana*, PRMT5 serait aussi associée au contrôle du cycle circadien par le contrôle de l'épissage alternatif de la protéine du complexe circadien PRR9 (Pseudo-Response Regulator 9) (Sanchez et al., 2010).

Le rôle de la méthylation de H4R3 et H3R8 dans la répression transcriptionnelle associe PRMT5 au développement de plusieurs cancers. Dans des cellules tumorales sur-exprimant PRMT5, ces méthylations sont enrichies sur les régions promotrices des suppresseurs de tumeurs ST7 et NM23, ce qui induit la diminution de l'acétylation de H3K9 et l'extinction de ces gènes (Pal et al., 2004).

PRMT5 méthyle aussi des histones avant leur incorporation dans les nucléosomes, ce qui pourrait être associé au maintien de la répression transcriptionnelle lors du développement

embryonnaire : durant les étapes initiales du développement dans les œufs de xénope, la chromatine se trouve sous forme pluripotente et est peu complexée avec des histones. Les histones « libres » sont associées à des chaperons, dont la nucléoplasmine. PRMT5 méthyle les histones libres H2A, des variants de H2A ainsi que H4 (Wilczek et al., 2011).

PRMT5 peut également agir sur l'expression des gènes en méthylant des protéines non-histones. Tout comme PRMT1, PRMT5 dans le contexte du méthylosome méthyle le facteur d'élongation de la transcription SPT5 sur son domaine ROW, qui régule son interaction avec l'ARN polymérase II. Cela diminuerait l'association avec STP5-ARN polymérase II et donc régulerait à la baisse l'élongation de la transcription (Kwak et al., 2003).

➤ **Activation de la transcription par PRMT5**

L'activation transcriptionnelle par PRMT5 s'exprime par l'apposition du marquage épigénétique H3R2^{me2s}. À l'inverse de H3R2^{me2a}, catalysé par PRMT6, H3R2^{me2s} stimule l'interaction H3-WDR5, et donc favorise la synthèse de H3K4^{me3} par Ash2/WDR5/MLL, ce qui est associé à l'activation de la transcription (Migliori et al., 2012). De plus, H3R2^{me2s} inhibe le recrutement des complexes corépresseurs Sin3a, NurD et PRC2 en interférant avec la fixation de leur constituant commun RBBP7 à H3 (Migliori et al., 2012).

PRMT5 peut également coactiver l'expression des gènes en méthylant des protéines non-histones telles que HOXA9, un facteur de transcription homéobox-spécifique associé à l'expression de gènes hématopoïétiques et dérégulé dans des cas de leucémies (Gao et al., 2016 ; Huang et al., 2011). En réponse à un stimulus par TNF α , PRMT5 est recrutée par HOXA9 au niveau du promoteur de la molécule d'adhérence E-selectine et méthyle HOXA9 R140. Cette modification est nécessaire à l'expression de la E-selectine (Bandyopadhyay et al., 2012), une protéine nécessaire à l'ancrage des leucocytes lors de la réponse inflammatoire (Granger et al., 2010).

➤ **Autres implications de PRMT5**

PRMT5 pourrait être impliquée dans la régulation de la localisation cellulaire des PIWI (P-element Induced WImpy testis), des protéines liant les ARNs nucléaires piRNA et qui jouent un rôle dans la différenciation des cellules germinales, l'ARN interférence et la suppression des transposons (Thomson et al., 2009). Le complexe PRMT5-MEP50 méthyle les protéines PIWI sur des séquences amino-terminales conservées, ce qui module leur interaction avec des protéines arborant des domaines TUDOR et pourrait être nécessaire à la localisation des PIWI dans des structures nommées « nuages » dans les cellules germinales (Vagin et al., 2009). PRMT5 méthyle également la protéine ribosomique S10 et la protéine de l'appareil de Golgi GM130, ce qui associerait PRMT5 dans les mécanismes d'assemblage et de maintien de ces deux constituants cellulaires associés à la synthèse et la maturation des protéines (Ren et al., 2010 ; Zhou et al., 2010).

h - PRMT9

PRMT9 est la moins étudiée des PRMT. Elle a été identifiée en même temps que PRMT8 par homologie de séquence (Jaeho Lee et al. 2005, 8) et diméthyle symétriquement plusieurs substrats *in vitro*, ce qui la classe comme la seconde PRMT de type II identifiée chez les mammifères (Y. Yang et al. 2015, 9). PRMT9, comme PRMT7, est constituée de deux modules catalytiques en tandem mais se différencie de celle-ci par la présence d'un domaine de trois répétitions de tétratricopeptides (TPR) amino-terminal (Yang et al., 2015b). Ces domaines sont connus pour être impliqués dans les interactions protéine-protéine, mais la fonction du domaine TPR de PRMT9 n'a pas encore été identifiée.

La seule cible identifiée de PRMT9 l'associe au mécanisme d'épissage de l'ARN. PRMT9 forme un complexe avec les facteurs d'épissage SAP145 et SAP49 et méthyle *in vitro* et *in vivo* SAP145 sur l'arginine 508 (Yang et al., 2015b). SAP145 et SAP49 sont des constituants des snRNPs et la méthylation de SAP145 forme un site de reconnaissance par le domaine TUDOR de SMN, facilitant l'assemblage du spliceosome. L'atténuation de l'expression de PRMT9 influence l'épissage des ARN.

Aujourd'hui, plusieurs études pourraient associer PRMT9 au développement de cancers : PRMT9 est en effet sur-exprimée dans plusieurs cancers, en particulier les lymphomes, les mélanomes et les cancers des testicules et du pancréas (Uhlen et al., 2015). De plus, la sur-expression de PRMT9 dans le carcinome hépatocellulaire (Jiang et al., 2018) stimule la prolifération de cellules cancéreuses et le développement de métastases, tandis que l'inhibition de son expression bloque leur prolifération. À l'inverse, le miRNA-543 inhibe l'expression de PRMT9 chez les patients atteints d'ostéosarcome (Zhang et al., 2017a), ce qui indique que PRMT9 pourrait jouer un rôle de suppresseur de tumeurs dans le cadre de cette pathologie. Le rôle joué par PRMT9 dans le développement de ces cancers fait de cette PRMT une nouvelle cible pour le développement d'inhibiteurs.

i - PRMT7

Une unique PRMT de type III a été caractérisée chez les mammifères : PRMT7, responsable du marquage MMA. PRMT7 a été découverte en 2003 dans le cadre d'un criblage ayant pour but d'identifier les gènes de réponse à des composés cytotoxiques (Gros et al., 2003). Comme PRMT9, PRMT7 a pour particularité d'être constituée de deux modules catalytiques en tandem. La structure cristallographique de PRMT7 de souris résolue au sein de l'équipe (PDB 4C4A) montre que le second module catalytique possède un site catalytique dégénéré et inactif (Cura et al., 2014). De plus, PRMT7 catalyserait uniquement la synthèse d'arginine monométhylée ce qui la classe ainsi comme la première PRMT de type III caractérisée (Miranda et al., 2004). La spécificité de produit de PRMT7 a longtemps été débattue entre les défenseurs de la classification en type III et ceux

montrant que PRMT7 produit des arginines diméthylées symétriquement (Lee et al., 2005b ; Wang et al., 2009a). La question a été tranchée à partir de 2012 dans deux études portant sur plusieurs substrats peptidiques et entiers de PRMT7, confirmant l'appartenance de cette enzyme au troisième type et montrant une sélectivité de PRMT7 pour les séquences basiques (Feng et al., 2013 ; Zurita-Lopez et al., 2012).

➤ **Coopération entre PRMT5 et PRMT7**

Des recherches montrent que PRMT5 et PRMT7 possèdent des substrats communs et les méthylent sur des sites identiques ou différents. PRMT7 agit synergiquement avec PRMT5, mais sans redondance, lors de la méthylation des protéines Smd1/D3 dans le processus d'assemblage des snRNPs (Gonsalvez et al., 2007). En effet, les deux protéines sont nécessaires à ce processus puisque la délétion de l'une ou de l'autre perturbe l'assemblage des snRNPs. Cette synergie est nécessaire à l'assemblage du complexe ribonucléoprotéique suivi de la formation du spliceosome. PRMT7 et PRMT5 coopéreraient également pour la méthylation du facteur d'élongation de la traduction eEF2 en réponse à un stimulus par le facteur de croissance bFGF (Jung et al., 2011). Finalement, comme PRMT5, PRMT7 méthyle H3R2, ce qui est associé avec un état actif de la chromatine (Migliori et al., 2012).

Beaucoup de ces études se sont focalisées sur la détection des diméthylations symétriques des arginines, corrélée à des KO de PRMT5 et PRMT7, afin d'étudier les fonctions de PRMT7, jusqu'à ce que la classification de PRMT7 comme type III soit mieux établie. Il est donc possible que la perte des sDMA observée pour les KO de PRMT7 soit due à l'absence d'un marquage MMA déposé par PRMT7 qui pourrait être nécessaire à la méthylation subséquente par PRMT5.

➤ **PRMT7 et la régulation de la pluripotence**

Plusieurs études indiquent une implication cruciale de PRMT7 dans la régulation du cycle cellulaire et de la différenciation. PRMT7 est en effet exprimée dans les cellules souches et germinales embryonnaires indifférenciées mais n'est plus détectée après différenciation, ce qui suggère que PRMT7 est associée au maintien de la pluripotence (Buhr et al., 2008). De plus, PRMT7 est exprimée dans les cellules souches satellites dont le rôle est de régénérer le tissu musculaire suite à une blessure. PRMT7 y réprime l'expression de p21 (Blanc et al., 2016), un inhibiteur des kinases du cycle cellulaire CDK (Saha et al., 1997), et stimule l'expression de l'ADN méthyltransférase 3b, maintenant quiescentes les cellules souches satellites (Blanc et al., 2016). PRMT7 est également exprimée durant le développement des testicules au stade embryonnaire, dans lesquelles elle méthyle H4R3. Cette modification s'accumule dans les cellules germinales et est particulièrement enrichie au niveau des séquences sous empreinte parentale, c'est-à-dire des gènes « éteints » lors du transfert héréditaire de l'information épigénétique (Jelinic et al., 2006).

➤ PRMT7 et l'intégrité du génome

Comme présenté précédemment, PRMT7 a été initialement identifiée dans le cadre d'une campagne de criblage des gènes de réponse à des composés cytotoxiques, dont des molécules modifiant la structure chimique de l'ADN. Dans ce contexte, PRMT7 co-localise avec le complexe de remodelage de la chromatine SWI/SNF au niveau des promoteurs de gènes impliqués dans la réparation de l'ADN. En synergie avec l'activité de SWI/SNF et par la méthylation de H2AR3 et H4R3, PRMT7 participe à la répression de ces gènes (Karkhanis et al., 2012). Les défauts de réparation de l'ADN sont une des causes primordiales au développement de cancers. De surcroît, la surexpression de PRMT7 serait associée au cancer du sein. En effet, PRMT7 réprime l'expression des protéines d'adhésion cellulaire E-cadhérines ce qui stimule le développement de métastases (Yao et al., 2014).

Finalement, PRMT7 interagit avec l'argininosuccinate synthase (ASS1) (Verma et al., 2017). ASS1 est impliquée dans une des étapes de la synthèse de l'arginine, ce qui a poussé Verma et ses collaborateurs à proposer que la synthèse et la méthylation des arginines pourraient être des processus coordonnés. Dans la citrullinémie de type I, ASS1 est mutée sur plusieurs résidus. Ces mutations interfèrent avec la formation du complexe PRMT7-ASS1, ce qui représente une première piste dans l'exploration des mécanismes conduisant à cette pathologie (Verma et al., 2017).

3 - Régulation de l'activité des PRMT

Les mécanismes de régulation de l'activité des PRMT demeurent aujourd'hui encore mal compris. Deux études ont lié le suppresseur de tumeur DAL-1 à la régulation des PRMT. DAL-1 interagit avec le module catalytique de PRMT3 et inhibe son activité enzymatique (Singh et al., 2004). DAL-1 interagit également avec PRMT5 et inhibe la méthylation des protéines Sm, ou inversement stimule celle de MBP (Jiang et al., 2005). Le mécanisme de la modulation de ces deux PRMT par DAL-1 n'est toutefois pas connu. Une étude a montré que chez CARM1, la sérine 217 peut être phosphorylée *in vivo*, ce qui inhibe l'activité de l'enzyme (Feng et al., 2009). S217 se localise dans la poche de fixation du SAM et participe à la fixation du cofacteur. La phosphorylation de ce résidu introduit un groupement encombrant et fortement chargé négativement pouvant interférer avec la fixation du SAM. Ce résidu est conservé dans les PRMT de type I ce qui suggère que la phosphorylation de cette position pourrait être un mécanisme de régulation commun. Selon un autre mode de régulation, CARM1 serait auto-méthylée au niveau de son domaine carboxy-terminal *in vivo* et *in vitro*, sans que son activité enzymatique n'en soit perturbée. Cela perturberait la modulation de la transcription et de l'épissage des prés-ARNm par CARM1 (Kuhn et al., 2011).

4 - Structure des PRMT

Actuellement, les structures de huit des neuf PRMT caractérisées chez les mammifères ont été résolues, à l'exception de PRMT9. Au total, 57 structures de PRMT de mammifères ont été déposées dans la base de données RCSB-PDB, mais aucune ne correspond à une PRMT complète, à l'exception de PRMT5 (PDB 3UA3 et 4GQB) (Antonysamy et al., 2012 ; Sun et al., 2011). Toutes les PRMT sont en effet des protéines modulaires construites autour d'un module catalytique commun. Celui-ci est complété par des extensions amino-terminales de courte longueur (PRMT1, PRMT6, PRMT7 et PRMT8) prédites comme non organisées, ou de longueur plus importante et organisées en domaines (PRMT2, PRMT3, PRMT4/CARM1, PRMT5 et PRMT9) parfois séparés du module catalytique par un segment non structuré.

a - Structure du module catalytique

Le module catalytique des PRMT forme une structure hautement conservée constituée de deux principaux domaines (figure 9) : un domaine de Rossmann et un tonneau β . Le domaine de Rossmann est un repliement permettant la fixation des cofacteurs dérivés de nucléotides (Gherardini et al., 2010 ; Rossmann et al., 1974), et dans le cas des PRMT et de nombreuses autres méthyltransférases, le domaine de Rossmann fixe le cofacteur S-adenosyl-méthionine (Gana et al., 2013).

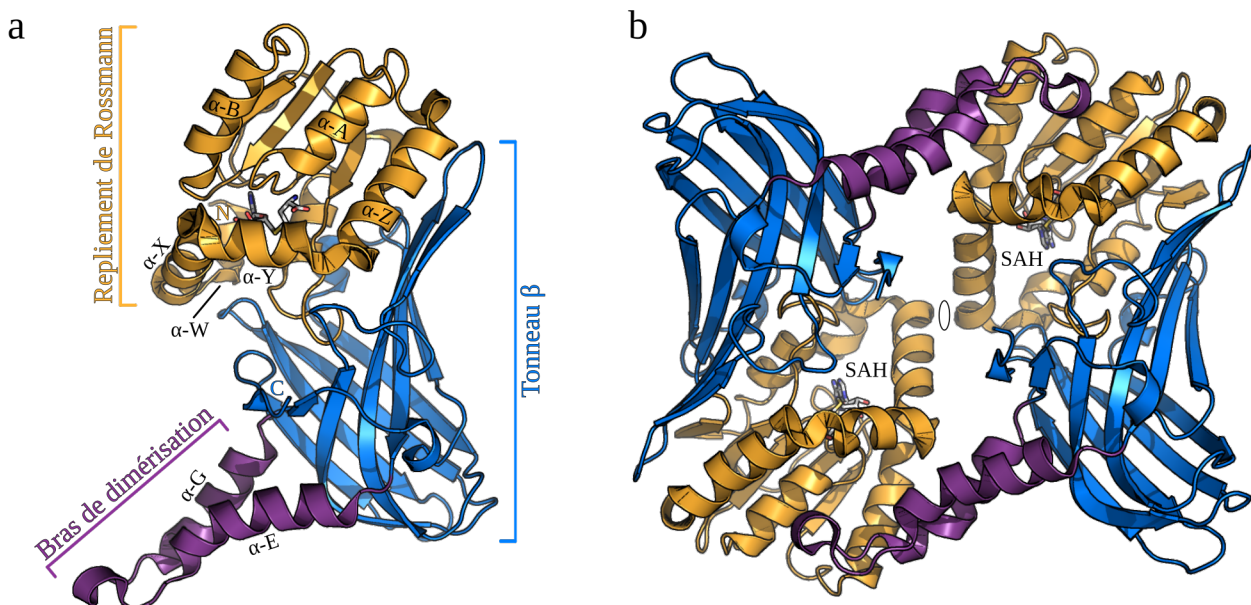


Figure 9 - Module catalytique des PRMT.

(a) Représentation en ruban d'un module catalytique de CARM1 de souris (PDB 5IH3). (b) Homo-dimère formé de deux modules catalytiques de CARM1 de souris agencés tête-bêche. Le domaine en repliement de Rossmann est coloré en orange. Le tonneau β est coloré en bleu. Le bras de dimérisation est coloré en violet. Le ligand SAH est représenté sous forme de bâtonnets.

Ce domaine est constitué d'un repliement de type beta-alpha-beta, classiquement composé d'un feuillet β constitué de cinq brins parallèles entourés de part et d'autre de quatre hélices α (Rossmann et al., 1974). Chez les PRMT, la chaîne polypeptidique en amont du domaine de Rossmann se replie autour du cofacteur SAM afin de l'enfourer au coeur de la structure (figure 9a), et forme une à deux hélices notées α -X et α -Y (Troffer-Charlier et al., 2007). Le repliement de ce segment établit, à l'interface avec le tonneau β , le site actif de l'enzyme. Bien qu'elle ne soit pas visible dans toutes les structures cristallographiques (PRMT1 et PRMT8), l'hélice α -X est présente a priori chez toutes les PRMT de type I. Dans le cas de la PRMT de type II, PRMT5, le segment équivalent à α -X est organisé en boucle et cette même région n'est visible dans aucune structure cristallographique des PRMT de type III (PRMT7).

Le second domaine naît de l'interruption du domaine de Rossmann par une cis-proline strictement conservée (P288 chez mCARM1), et prend la forme d'un tonneau β constitué de huit brins placés de manière antiparallèle. La jonction entre les brins β 6 et β 7 (alignement 1, page 19) est interrompue par deux hélices α , de longueurs variables selon la PRMT, qui forment le bras de dimérisation (figure 9 et figure 6). Le bras de dimérisation d'un premier monomère interagit majoritairement avec le repliement de Rossmann du module catalytique d'un second monomère de manière à former un homo-dimère tête-bêche adoptant une symétrie C_2 (figure 9b). Ce dimère constitue l'unité fonctionnelle des PRMT (Cheng et al., 2011b ; Troffer-Charlier et al., 2007 ; Zhang et al., 2003, 2000), à l'exception de PRMT7 et PRMT9. En effet, ces deux enzymes sont constituées chez la plupart des organismes d'une chaîne polypeptidique arborant deux modules catalytiques en tandem qui s'associent pour mimer le dimère caractéristique des PRMT (Cura et al., 2014 ; Hasegawa et al., 2014a). Le cofacteur et l'arginine substrat sont fixés dans une poche divisée en deux compartiments : le site de fixation du SAM et le site de fixation de l'arginine, qui sont illustrés sur la figure 10.

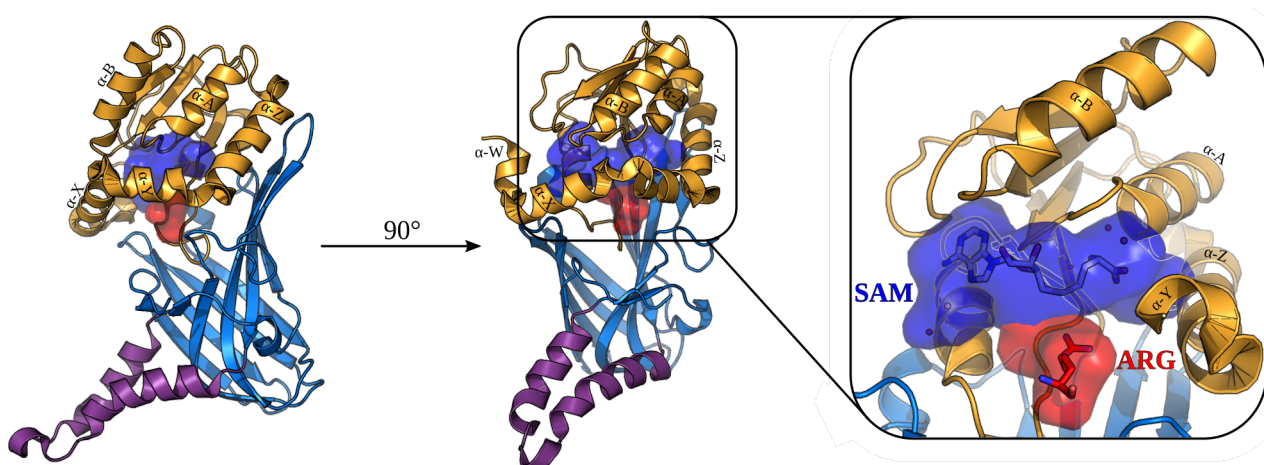


Figure 10 - Poches de fixation des substrats des PRMT.

Représentation en ruban d'un module catalytique de CARM1 de souris (5IH3) et des poches de fixation du cofacteur SAM (bleu) et de l'arginine substrat (rouge). Les poches ont été calculées à l'aide d'un programme « maison » PyPDB2.

Le site de fixation du SAM (en bleu) est formé au cœur du repliement de Rossmann entre les extrémités carboxy-terminales des brins $\beta 1$, $\beta 2$ et $\beta 4$ et les hélices α -X, -Y, -Z, -A et -B. Le site de fixation de l'arginine (en rouge) se situe à l'interface entre les deux domaines. Il est façonné lors du repliement des régions α -X et α -Y autour du cofacteur et permet la création d'un environnement dont le solvant est exclu, propice au déroulement de la réaction de méthylation qui a lieu à l'interface entre les deux poches.

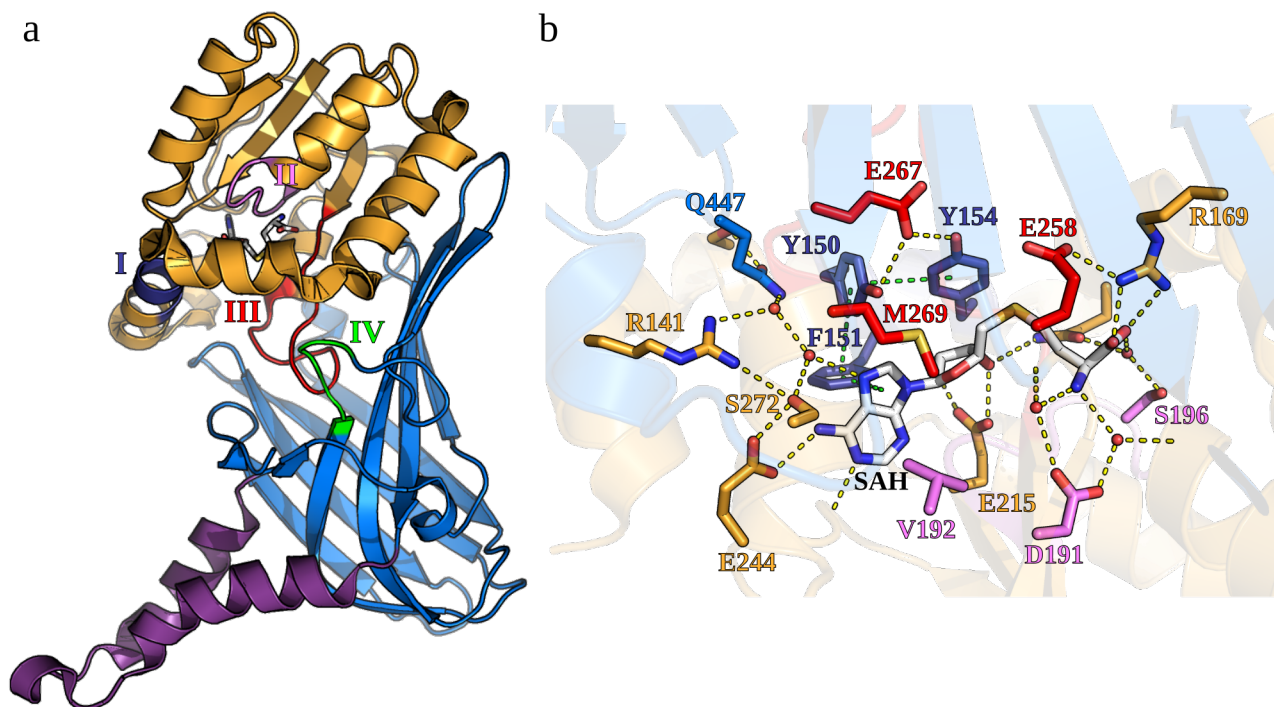


Figure 11 - Motifs caractéristiques du module catalytique des PRMT.

(a) Représentation ruban d'un module catalytique de CARM1 de souris (PDB 5IH3). Les motifs I, II, III et IV sont représentés respectivement en bleu foncé, rose, rouge et vert. (b) Site de fixation du SAH dans CARM1 de souris. Le code couleur précédent est repris. Les lignes pointillées jaunes et vertes représentent respectivement des liaisons hydrogène et interactions aromatiques.

Quatre motifs structurellement conservés se répartissent sur le repliement de Rossmann et le tonneau β , et sont essentiels à la structure et à l'activité de méthylation des PRMT. Le motif I (figures 11, 6 et alignement 1) est présent principalement chez les PRMT de type I, où il se structure pour former l'hélice α -X (Troffer-Charlier et al., 2007). De séquence conservée YFxxY, le motif I participe à la fixation du cofacteur et à la formation du site de méthylation : les trois résidus aromatiques forment ensemble un réseau d'interactions π - π orthogonales afin de structurer l'hélice α -X (Y150, F151 et Y154 chez mCARM1, figures 11b et 12a) et le résidu aromatique central (F151) établit une interaction π - π orthogonale avec la partie adénine du cofacteur pour venir refermer α -X sur le site de fixation du SAM. Les résidus aromatiques des extrémités du motif I (Y150 et Y154) créent également deux liaisons hydrogène avec un résidu catalytique acide du motif III (E267 dans mCARM1, figures 11b et 12a/b) afin de l'orienter vers l'intérieur de la poche de méthylation nouvellement formée.

L'hélice α -X arborant le motif I est totalement ou partiellement visible dans les structures de PRMT2, PRMT3 de rat, CARM1, PRMT6 et PRMT10 de plante (Bonfond et al., 2015 ; Cheng et al., 2011b ; Cura et al., 2017 ; Troffer-Charlier et al., 2007 ; Zhang et al., 2000), mais est désordonnée dans les structures cristallographiques de PRMT1 et PRMT8 (Lee et al., 2015 ; Zhang et al., 2003). Pour les PRMT de type II (PRMT5 et PRMT9) et la PRMT de type III (PRMT7), le motif I est fortement dégénéré. Le segment ne forme pas de structures secondaires dans PRMT5 (Antonysamy et al., 2012) et n'est pas visible dans les structures cristallographiques de PRMT7 si bien que sa structure secondaire n'est pas connue (Cura et al., 2014 ; Debler et al., 2016 ; Hasegawa et al., 2014a ; Wang et al., 2014).

Le motif II, de séquence DVGxGxG, est commun à toutes les PRMT et forme une boucle qui se localise entre le brin β 1 et l'hélice α -A du domaine de Rossmann (figure 11a/b). Les deux premiers résidus amino-terminaux varient selon le type de PRMT I, II ou III mais la répétition GxGxG est strictement conservée. GxGxG est un motif consensus à toutes les méthyltransférases utilisant un repliement de Rossmann pour fixer le SAM (Kozbial et al., 2005). Les glycines, flexibles et dépourvues de chaînes latérales, permettent de créer un espace suffisant autorisant la fixation de la partie adénosine du cofacteur.

Le motif III arbore les résidus catalytiques (E258 et E267 chez mCARM1) qualifiés de résidus « pompe » et « ancre » (Mailliot, 2013 ; Troffer-Charlier et al., 2007), et est de ce fait également conservé chez toutes les PRMT. Il forme une longue boucle entre le brin β 4 et la dernière hélice du repliement de Rossmann (α -D), façonnant la base de la poche de fixation de l'arginine (figures 11 et 12). Le résidu carboxy-terminal de ce motif (M269 dans mCARM1, figure 11) est un des discriminants du type de PRMT (Gui et al., 2014 ; Mailliot, 2013). Il s'agit d'une méthionine pour les types I, d'une sérine pour les PRMT5 ou d'un résidu hydrophobe pour les PRMT7 et PRMT9. Ce résidu se positionne à la verticale de la partie ribose du cofacteur.

Finalement, le motif IV forme une boucle insérée entre les brins β 11 et β 12 du domaine organisé en tonneau β (figures 11 et 12). La séquence THWxQ est hautement conservée chez les PRMT de type I tandis que pour les PRMT de type II, elle est remplacée par FSWFP (PRMT5) ou TCWxQ (PRMT9), et par DHWxQ chez les PRMT de type III. Seul le résidu tryptophane, impliqué dans l'architecture du site actif de l'enzyme, est totalement conservé. Les structures de PRMT disponibles montrent que les autres résidus du motif IV ont principalement un rôle dans la stabilisation de la boucle par le biais de liaisons hydrogène et d'interactions hydrophobes avec un acide aspartique à l'extrémité amino-terminale de α -Y, avec le brin β 7 du tonneau β en aval du motif IV et avec la boucle qui succède au bras de dimérisation (figure 12).

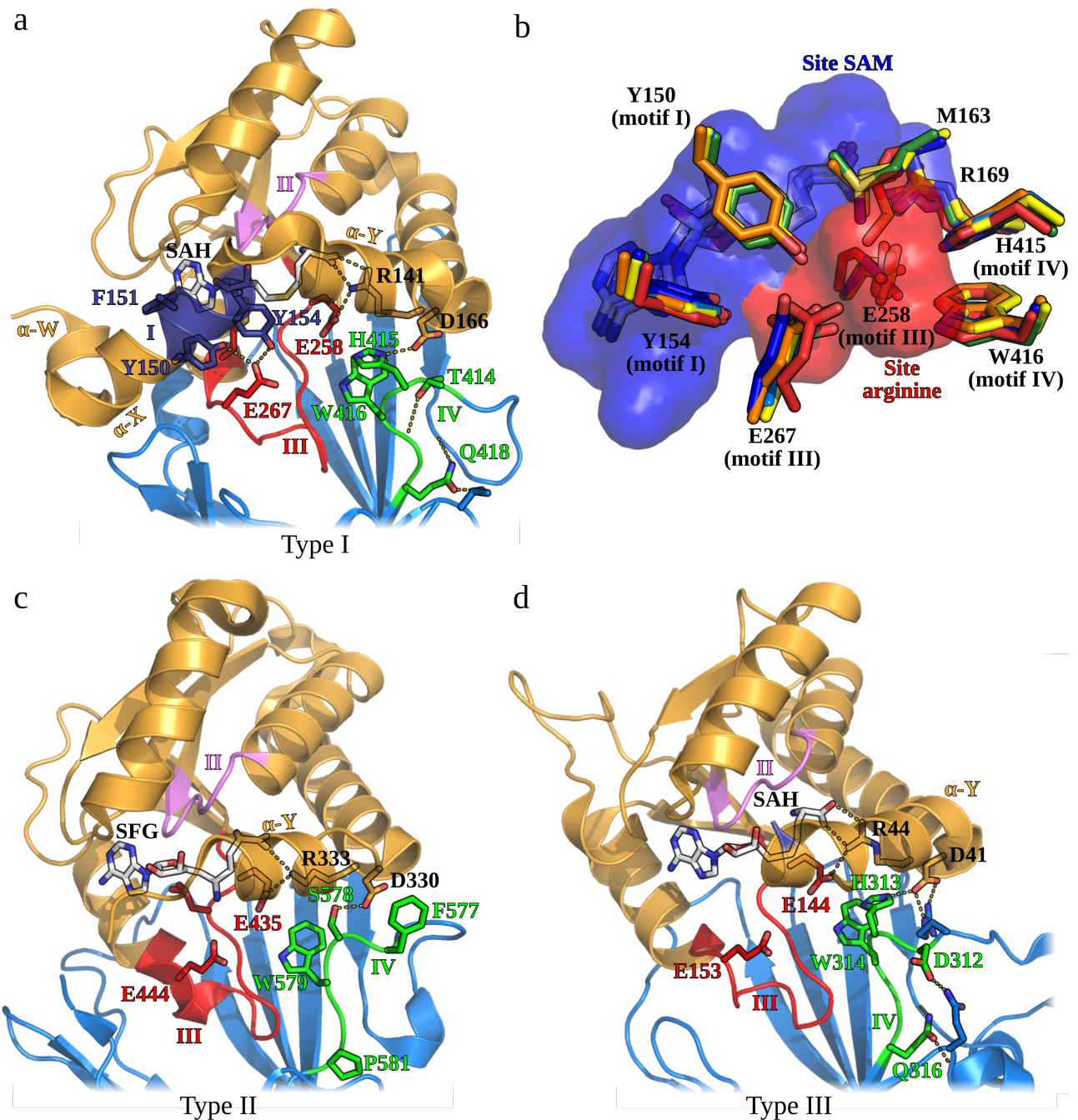


Figure 12 - Structure des motifs conservés dans les PRMT de types I, II et III.

Le repliement de Rossmann et le tonneau β de CARM1 (a, PDB 5IH3), PRMT5 (c, PDB 4GQB) et PRMT7 (d, PDB 4C4A) sont représentés respectivement en orange et bleu. Les motifs I, II, III et IV sont représentés respectivement en bleu foncé, rose, rouge et vert. Le motif I est absent chez les PRMT de type II tandis qu'il est dégénéré chez les PRMT de type III. Les liaisons hydrogène sont représentées par des tirets. (b) Superposition des sites de méthylation de PRMT1 (1OR8, rouge), PRMT2 (5JMQ, orange), PRMT3 (1F3L, jaune), CARM1 (5IH3, vert), PRMT6 (4C03, bleu clair) et PRMT8 (5DST, bleu foncé) illustrant la conservation. Les motifs I de PRMT1 et PRMT8 ont été reconstruits par homologie. Les volumes correspondant aux poches de fixation du SAM (bleu) et de l'arginine (rouge) ont été obtenus comme indiqué figure 10. L'annotation des résidus correspond à CARM1 de souris.

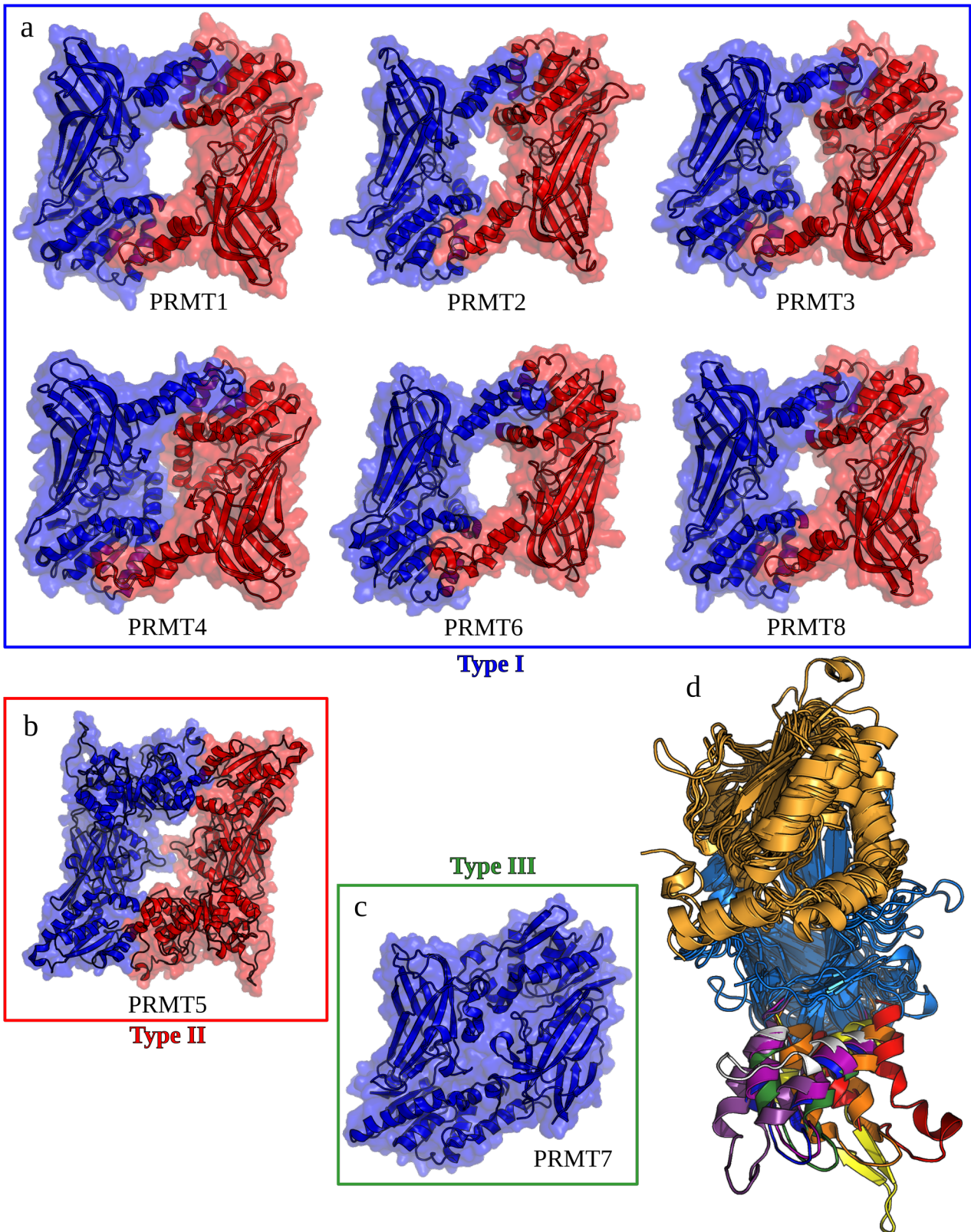


Figure 13 - Comparaison de l'organisation spatiale des PRMT.

(a) Structure du dimère des PRMT de type I. (b) Structure du dimère de la PRMT de type II, PRMT5. (c) Structure mimant le dimère de la PRMT de type II, PRMT7 (PDB 4C4A). (d) Superposition des modules catalytiques de PRMT1, PRMT2, PRMT3, PRMT4, PRMT5, PRMT6, PRMT7 (N-ter) et PRMT8.

b - Caractéristiques spécifiques des PRMT

Les structures des PRMT se différencient par la géométrie du dimère de modules catalytiques et par les éventuels domaines amino- et carboxy-terminaux qui les complètent. Il a été démontré que les domaines amino-terminaux participent à la fixation des substrats et partenaires et sont probablement structurellement stabilisés lors de ces interactions. La figure 13 présente les dimères catalytiques des PRMT de type I (figure 13a), II (figure 13b) et III (figure 13c). La superposition des modules catalytiques présentée sur la figure 13d dévoile une très forte similarité structurale entre les différentes PRMT, mais la comparaison des dimères formés par ces modules catalytiques montre que ceux-ci se différencient par leur géométrie, variant de l'ovale au rectangle, et par la taille de la cavité centrale. Chez CARM1, la cavité est obstruée sur une face dans les structures cristallographiques. C'est également le cas de PRMT7. La cavité est ouverte sur les deux faces et plus ou moins ample chez les autres PRMT. La géométrie de cette cavité est déterminée par l'orientation relative et l'espacement entre les deux dimères, qui sont gouvernés par la taille du bras de dimérisation, de la boucle à son extrémité, et de l'angle formé par le bras de dimérisation avec le tonneau β (figure 13d). Les différences de géométrie entre les dimères de modules catalytiques pourraient influencer l'accessibilité aux substrats, et donc agir sur la spécificité des PRMT pour leurs cibles.

➤ PRMT2

Comme l'illustre la figure 14, PRMT2 de souris est constituée de 448 acides aminés et

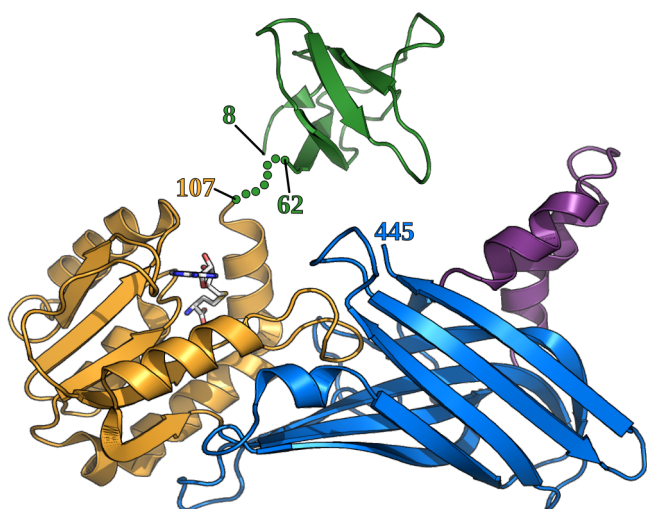


Figure 14 - Modèle de PRMT2.

Structure RMN du domaine amino-terminal SH3 de PRMT2 de souris (en vert, PDB 1X2P) et structure cristallographique du module catalytique de PRMT2 de souris (en orange, bleu et violet, PDB 5JMQ).

arbore un domaine amino-terminal formant un repliement de type SRC-homology 3 (SH3, résidus 8 à 62), dont la structure a été résolue par RMN (PDB 1X2P) (Chikayama et al., 2005). Ces domaines sont constitués d'une soixantaine d'acides aminés formant un tonneau de 5 brins β antiparallèles. Les domaines SH3 sont connus pour être impliqués dans les interactions protéine-protéine (Kurochkina et al., 2012 ; Sparks et al., 1996), mais la fonction précise du domaine SH3 de PRMT2 n'a pas encore été identifiée. PRMT2 est un des sujets étudiés dans cette thèse, et la présentation de sa structure sera abordée au chapitre III.

➤ PRMT3

Comme l'illustre la figure 15, PRMT3 de souris est constituée de 532 acides aminés et

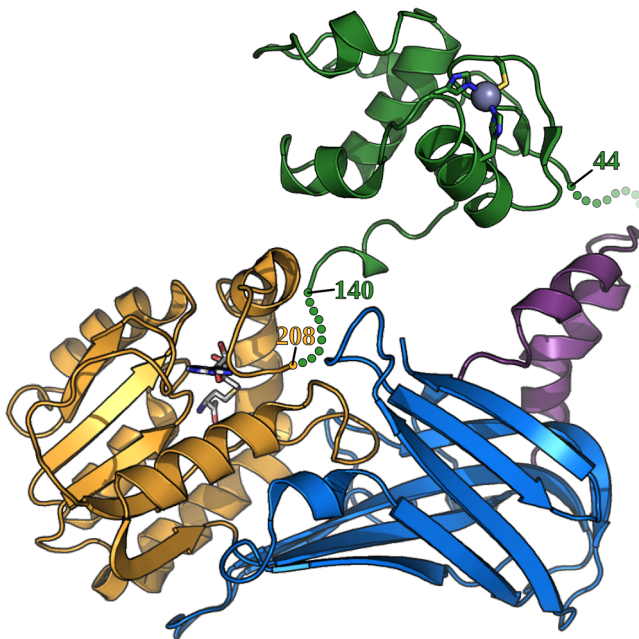


Figure 15 - Modèle de PRMT3.

Structure RMN du domaine amino-terminal de PRMT3 de souris (en vert, PDB 1WIR) et structure cristallographique du module catalytique de PRMT3 de rat (en orange, bleu et violet, PDB 1F3L).

possède un domaine amino-terminal globulaire d'environ une centaine d'acides aminés (résidus 44 à 140), organisés en 5 hélices α et deux feuillets β stabilisés par un motif de fixation du zinc similaire au doigt de zinc C2H2. Ce domaine est nécessaire à la fixation à son substrat et partenaire rpS2. Sur ce domaine, la tyrosine 87 (numérotation de la PRMT3 humaine) serait essentielle à cette interaction (Handrkova et al., 2011). Ce résidu se localise sur la face du domaine amino-terminal opposée au C2H2, riche en résidus acides, ce qui suggère que cette région pourrait être impliquée dans la fixation de ce partenaire plutôt basique. La structure du domaine amino-terminal de PRMT3 a été résolue par RMN (PDB 1WIR) (Miyamoto et al., 2004).

➤ CARM1

Comme le montre la figure 16, CARM1 (Troffer-Charlier et al., 2007) est une protéine de 608 acides aminés (chez la souris) qui possède un domaine amino-terminal dont la structure cristallographique a été résolue au sein de l'équipe (Troffer-Charlier et al., 2007). Ce domaine est séparé du module catalytique par un segment de cinq acides aminés et adopte un repliement de type pleckstrin homology (PH) constitué d'une centaine d'acides aminés organisés en deux feuillets antiparallèles placés face à face, avec à leur base une hélice alpha. Les domaines PH sont décrits comme un élément de reconnaissance protéine-protéine. Bien que sa fonction exacte soit inconnue, ce domaine est nécessaire à la stimulation de l'activité transcriptionnelle des récepteurs nucléaires (Teyssier et al., 2002). Le module catalytique de CARM1 (Troffer-Charlier et al., 2007) possède quelques particularités par rapport à celui des autres PRMT de type I. Le domaine de fixation du SAM débute par une courte hélice supplémentaire, notée α -W (en cyan sur la figure 16), formant un coude avec l'hélice α -X et nécessaire à l'activité de CARM1 (Troffer-Charlier et al., 2007). Le motif III contient une proline conservée (P259 dans mCARM1) qui pourrait ajouter de la rigidité dans ce motif et donc modifier la dynamique du module catalytique et de la réaction de méthylation. D'autre part, le tonneau β de CARM1 est pourvu d'un brin additionnel formé par l'extension carboxy-terminale de la protéine, représenté en marron sur la figure 16b (résidus 471 à 478 dans

mCARM1). Ce brin se situe à proximité du site de fixation de l'arginine, dans la zone qui interagit avec le substrat et pourrait être essentiel à la sélectivité de CARM1. De surcroît, CARM1 est la seule PRMT à posséder à la fois un domaine amino-terminal et un domaine carboxy-terminal additionnel. Ce dernier (résidus 507-608 dans mCARM1) est supposé flexible à l'état libre car il est sujet à la dégradation lorsque CARM1 est purifiée seule (Troffer-Charlier et al., 2007).

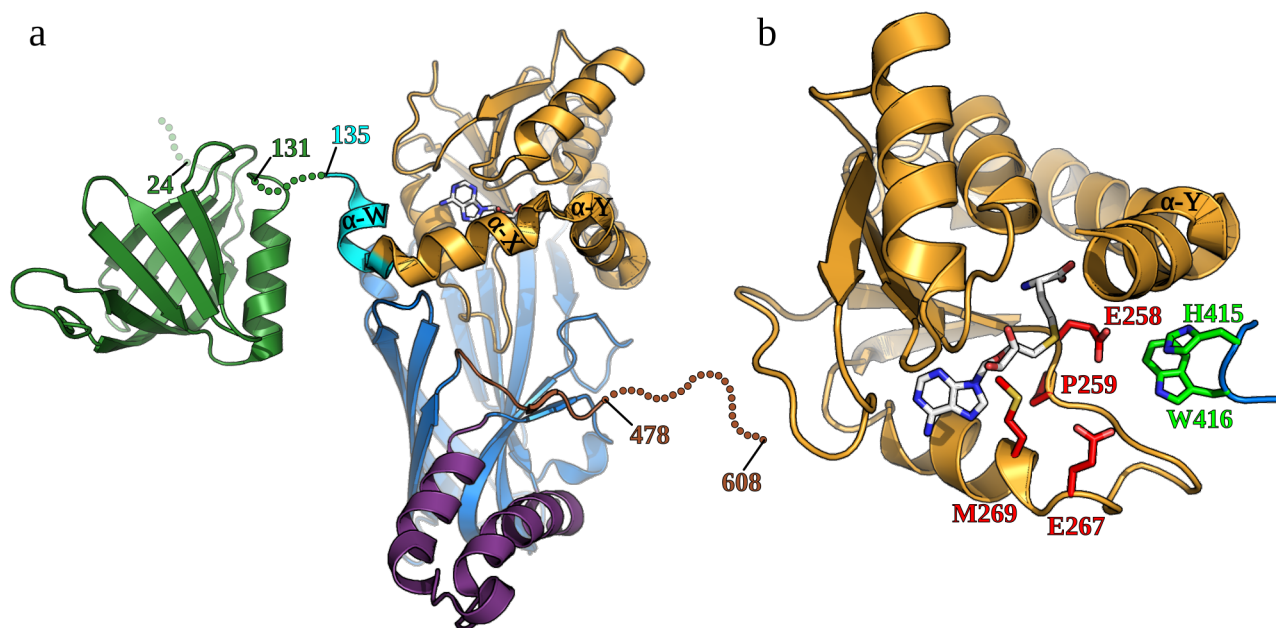


Figure 16 - Particularités de CARM1.

(a) Structures cristallographiques du domaine amino-terminal (vert, PDB 4OQB) et du module catalytique de CARM1 (PDB 5IH3). (b) Domaine de Rossmann de mCARM1 fixé à du SAH. Pour des raisons de visibilité, les hélices α -X et α -W n'ont pas été représentées.

➤ PRMT5

La première PRMT de type II identifiée chez les mammifères, PRMT5 (Branscombe et al., 2001 ; Pollack et al., 1999a ; Rho et al., 2001), possède un mode de dimérisation singulier différent de toutes les autres PRMT, tel qu'illustré en figure 17. En effet, les structures cristallographiques des PRMT5 de nématode (PDB 3UA3), humaine (PDB 4GQB) et de xénope (PDB 4G56) montrent que l'oligomérisation est assurée par l'interaction entre le domaine de Rossmann du premier monomère et le domaine amino-terminal du second monomère. À l'inverse, la dimérisation est assurée par l'interaction entre le bras de dimérisation et le domaine de Rossmann chez les autres PRMT. Chez PRMT5, les bras de dimérisation se rejoignent au cœur du dimère pour compléter l'interaction (Antonysamy et al., 2012 ; Ho et al., 2013 ; Sun et al., 2011). Le domaine amino-terminal adopte une structure de tonneau de type TIM. Les tonneaux de type TIM sont des repliements très répandus, adoptés par environ 10 % des enzymes (Goldman et al., 2016). Ils se caractérisent par un cœur cylindrique formé de 8 brins parallèles entourés d'un anneau de 8 hélices α . Le domaine amino-terminal de PRMT5 est donc impliqué dans la dimérisation de l'enzyme mais aussi dans le recrutement de son partenaire MEP50. Il est étroitement fixé au module

catalytique et est le seul domaine amino-terminal des PRMT ayant été observé dans les structures cristallographiques de PRMT complètes.

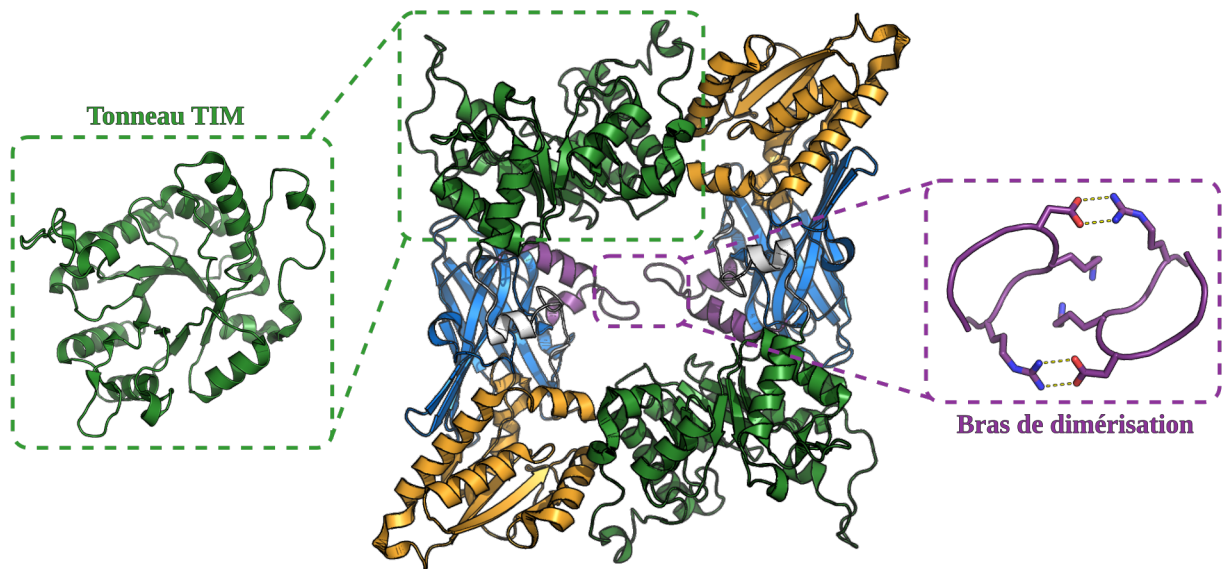
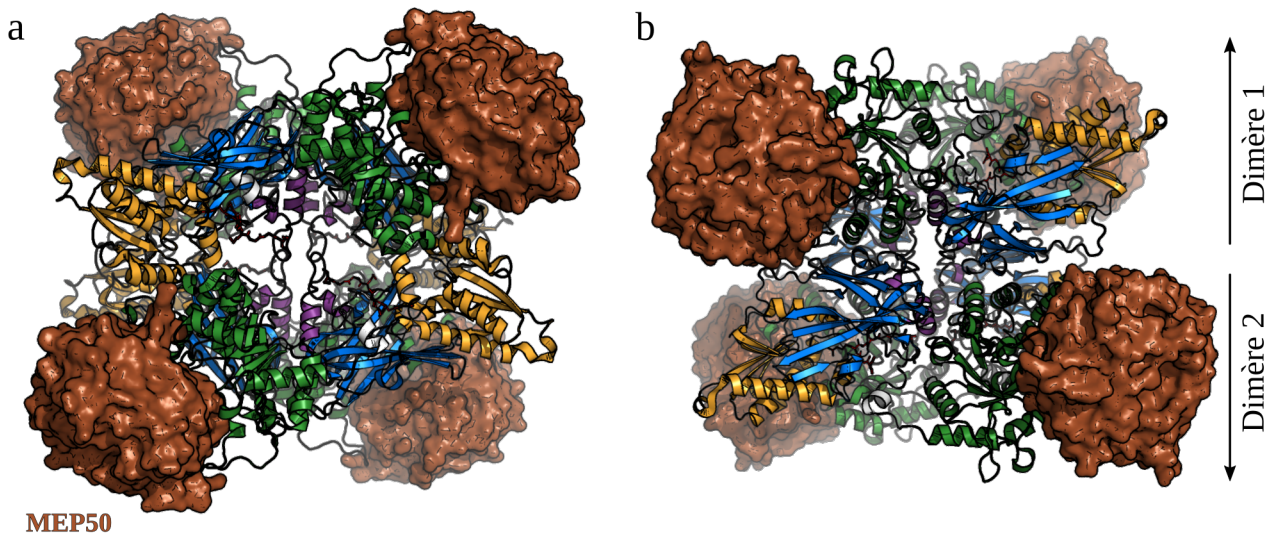


Figure 17 - Structure cristallographique du dimère de PRMT5.

À gauche : structure cristallographique du dimère de PRMT5 humain (PDB 4GQB). Les domaines de Rossmann, tonneau β et le bras de dimérisation constituant le module catalytique sont représentés respectivement en rubans orange, bleu et violet. Le domaine amino-terminal adoptant un repliement de type tonneau de type TIM est représenté en rubans vert. À droite : détails de l'interface entre les deux bras de dimérisation.

Les structures cristallographiques du complexe formé par PRMT5 avec son partenaire MEP50 et un peptide H4 ont également été résolues (Antonyamy et al., 2012 ; Ho et al., 2013). Ces structures dévoilent une organisation tétramérique de PRMT5 dans laquelle deux dimères interagissent par leurs faces équivalentes, ce qui conduit à ce que les cavités centrales des deux dimères soient orientées vers des directions opposées, comme l'indiquent les flèches en figure 18b.



MEP50
Figure 18 - Structure du complexe PRMT5-MEP50.

Vues du complexe constituant du méthylosome PRMT5 – MEP50 (PDB 4GQB) de face (a) et de profil (b). MEP50 est représenté en surface marron.

Chaque monomère reçoit un peptide de H4 et fixe un partenaire MEP50 en périphérie du cœur de PRMT5, par le biais d'interactions avec le domaine amino-terminal de PRMT5. Cette organisation spatiale suggère que PRMT5 constitue le cœur structurant du méthylosome, et qu'il permet un ancrage simultané de plusieurs substrats. La structure non cristallographique de ce complexe a également été résolue par cryo-microscopie électronique (Timm et al., 2018) et montre qu'en absence de cofacteur, les complexes PRMT5-MEP50 interagissent entre eux pour former des réseaux d'aspect fibreux tandis qu'en présence de sinefungine (un inhibiteur naturel des méthyltransférases, de structure proche du cofacteur SAM) le complexe forme une structure similaire à celle observée dans les structures cristallographiques. Cette organisation en « fibres » pourrait avoir une signification biologique.

➤ PRMT7

Les structures des PRMT de type III, PRMT7, de plusieurs organismes ont été résolues en 2014, celle de mammifère l'ayant été au sein de l'équipe (Cura et al., 2014 ; Hasegawa et al., 2014b ; Wang et al., 2014). Comme illustré sur les figures 6 et 13c, PRMT7 a pour particularité d'être constituée d'une répétition en tandem de deux modules catalytiques, peut être issue de la duplication du gène au cours de l'évolution. De manière assez surprenante, la PRMT7 de trypanosome (PDB 4M36) ne possède pas cette particularité puisqu'elle est constituée d'un unique module catalytique qui dimérise de façon classique. Les structures cristallographiques des PRMT7 de souris (PDB 4C4A, figure 19) et de nématode (PDB 3WST) révèlent plusieurs caractéristiques inédites chez cette enzyme comparées aux autres PRMT.

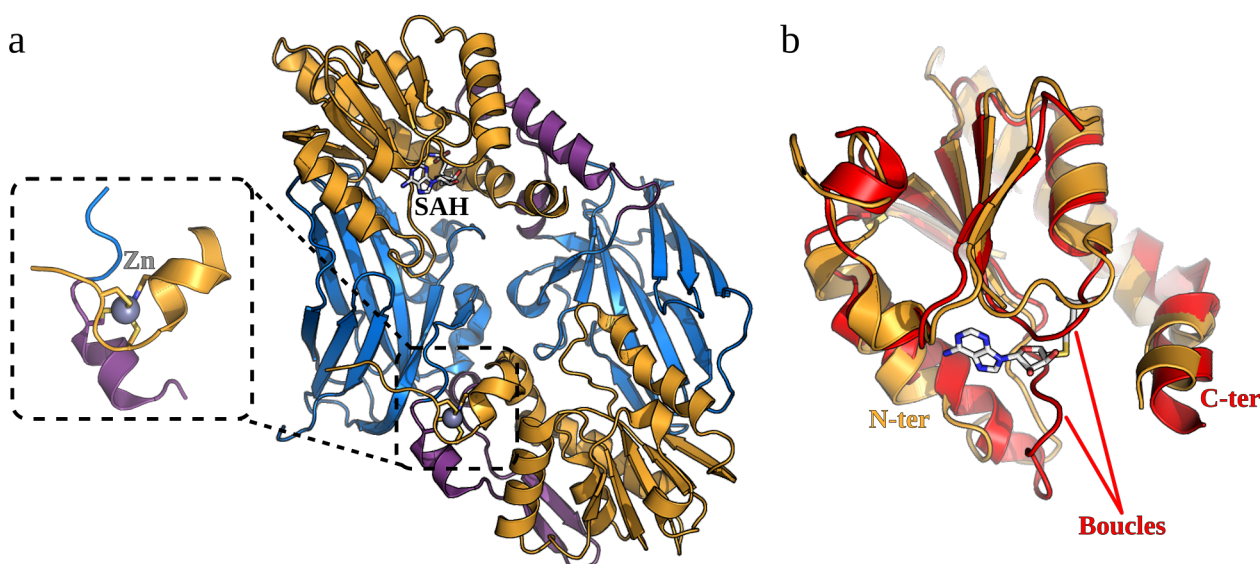


Figure 19 - Structure et spécificités de la PRMT7 de souris.

(a) Structure de la PRMT7 de souris (PDB 4C4A). À gauche : détails du site de fixation du zinc. Au milieu : monomère de PRMT7 mimant la structure d'un dimère. (b) Superposition des domaines de Rossmann des modules catalytiques amino- (en orange) et carboxy-terminaux (en rouge). La PRMT est représentée en ruban, suivant le code couleur employé dans ce manuscrit. La molécule de SAH est représentée en bâtonnets. L'ion zinc est représenté par une sphère.

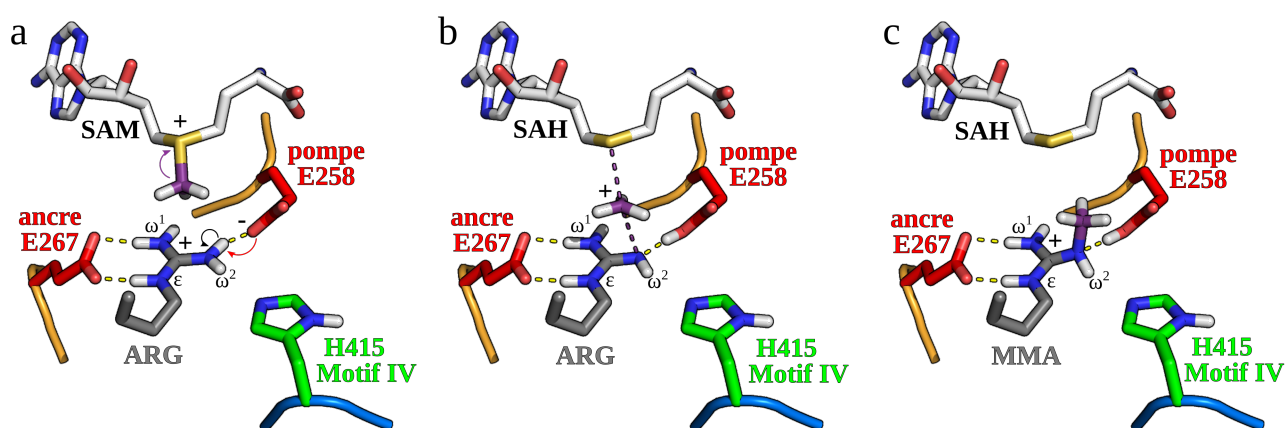
Les deux modules catalytiques en tandem (15-353 et 361-692 chez mPRMT7) s'associent de façon à mimer la forme dimérique classique des PRMT. Cette configuration est stabilisée par des interactions entre les bras de dimérisation d'un module catalytique avec le domaine de Rossmann du module catalytique adjacent (figure 19a). Comme pour CARM1, la cavité centrale est obstruée sur une face. Cependant, dans le cas de PRMT7 c'est la face opposée qui permet l'accès du substrat. Cette configuration spatiale est stabilisée par un motif de fixation du zinc faisant intervenir deux cystéines du bras de dimérisation du module catalytique amino-terminal, et un couple cystéine-histidine appartenant au domaine de Rossmann du second module catalytique (figure 19a). Les structures cristallographiques dévoilent également que le module catalytique carboxy-terminal ne fixe pas de SAH et est donc non productif, car les sites de fixation du SAM et de l'arginine substrat sont en partie obstrués par deux boucles dans les motifs II et III, tel qu'illustré figure 19b.

➤ **PRMT9**

La seconde PRMT de type II, PRMT9, a comme PRMT7 pour particularité d'être constituée de deux modules catalytiques en tandem. La structure de cette PRMT n'a pas encore été déterminée, mais le second domaine catalytique pourrait être non-productif car il possède plusieurs mutations dans des séquences conservées, dont l'insertion de deux sérines au lieu de deux glycines dans le très conservé motif II de séquence consensus DVGxGxG, nécessaire à la fixation du cofacteur. La nature dégénérée du second module catalytique de PRMT9 demeure néanmoins controversée. Il n'a en effet pas été démontré que celui-ci ne catalyse pas la réaction de méthylation, et pourrait être un intermédiaire évolutif entre une forme active et une forme inactive du module catalytique. Le chemin évolutif des PRMT est aujourd'hui incompris : les PRMT pourraient avoir initialement existé sous forme de monomères mimant un dimère, puis avoir convergé vers une forme dimérique avec perte du second module catalytique, ou inversement évoluer vers une forme monomérique avec deux modules catalytiques en tandem. L'étude structurale de cette PRMT représente donc à la fois un intérêt thérapeutique et un intérêt fondamental dans la compréhension de l'histoire et du fonctionnement de cette famille de protéines, et nous tentons actuellement d'en résoudre la structure.

c - Réaction de méthylation

Le mécanisme chimique de méthylation des PRMT a été proposé pour la première fois lors de l'étude structurale de PRMT3 (Zhang et al., 2000) et correspondrait à un mécanisme de substitution nucléophile de type 2 (SN2) comparable à celui décrit pour la plupart des méthyltransférases (Coward et al., 1977). Dans le mécanisme proposé, le second glutamate du motif III (l'ancre, E267 dans mCARM1) crée deux liaisons hydrogène avec les azotes ω^1 et ϵ du noyau guanidinium de l'arginine tandis que le premier glutamate (la pompe, E258 dans mCARM1) forme une unique liaison hydrogène avec ω^2 (figure 20a) (Mailliot, 2013). Ces interactions permettent de stabiliser la charge négative sur ω^1 et inversement de stabiliser un doublet non-liant sur ω^2 , ce qui lui confère des propriétés nucléophiles (Zhang et al., 2000). De plus, la charge négative de la pompe est neutralisée par son interaction avec un résidu basique conservé (R169 dans mCARM1). La pompe agit comme une base générale et déprotone l'azote de ω^2 (figure 20b). Cela permet l'attaque nucléophile sur le groupement méthyle du SAM, avec formation d'un intermédiaire réactionnel chargé positivement. Ce mécanisme a pu être modélisé lors de trois simulations par dynamique moléculaire hybride MM/MQ sur les types I PRMT1 (Zhang et al., 2013) et PRMT3 (Chu et al., 2013) et sur la type II PRMT5 (Yue et al., 2015).



➤ Déterminants structuraux de la diméthylation asymétrique

Le site de fixation de l'arginine et l'interface entre les poches SAM et arginine, lieu de la réaction de méthylation, sont très conservés parmi les PRMT de type I, tel qu'illustré figure 12b. Cela suggère que les PRMT de type I utilisent un mécanisme commun de diméthylation asymétrique des arginines. L'arginine substrat est stabilisée et orientée principalement par l'action de deux résidus : l'ancre et la pompe tel que décrit ci-dessus (E258 et E267 du motif III, notation de mCARM1) qui forment trois liaisons hydrogène avec les azotes ω^1 , ω^2 et ϵ du noyau guanidinium

de l'arginine (Bonfond et al., 2015 ; Mailliot, 2013 ; Troffer-Charlier et al., 2007). Les géométries des poches de fixation du SAM et de l'arginine positionneraient l'azote ω^1 dans l'axe du méthyle réactif du SAM. L'histidine du motif IV (H415 dans mCARM1) interviendrait dans la stabilisation et le positionnement du résidu monométhylarginine en formant une liaison hydrogène avec le premier groupement méthyle. Du côté opposé, la méthionine conservée (figure 21a/b, M269 dans mCARM1) localisée à la verticale du SAM imposerait par contrainte stérique au premier groupement méthyle de se positionner à proximité de H415 (Mailliot, 2013). Le résidu monométhylarginine se trouve alors dans une position favorisant la seconde méthylation sur le même azote. Un tel mécanisme a pu être « observé » *in silico* lors de deux simulations de dynamique moléculaire hybride MM/MQ sur PRMT1 et PRMT3 (Chu et al., 2013 ; Zhang et al., 2013).

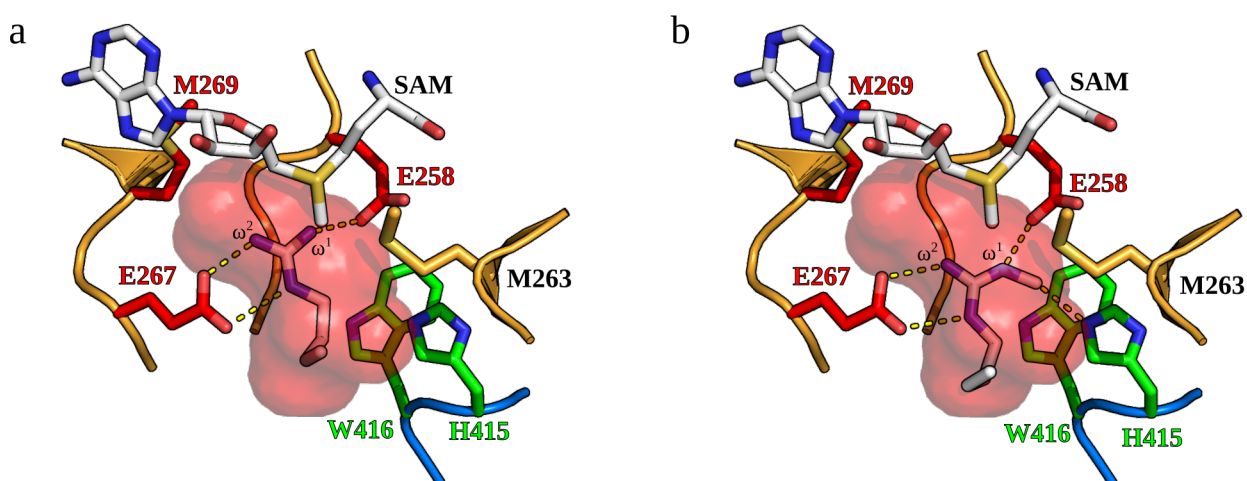


Figure 21 - Déterminants structuraux de la diméthylation asymétrique.

(a) Modélisation d'un résidu arginine dans le site actif de mCARM1 (PB 5IH3). (b) Modélisation d'un résidu monométhylarginine dans le site actif de mCARM1 (PDB 5IH3). Les liaisons hydrogène sont représentées en lignes hachurées. Le code couleur reprend celui utilisé dans la figure 11.

➤ Déterminants structuraux de la diméthylation symétrique

Dans le cas des PRMT de type II, seules les structures des PRMT5 de plusieurs organismes ont été résolues (Antonysamy et al., 2012 ; Ho et al., 2013). Comme pour les types I, l'arginine substrat est stabilisée et orientée dans la poche par les deux résidus acides glutamiques catalytiques par le biais de liaisons hydrogène. L'orientation du noyau guanidinium par rapport au groupement méthyl du SAM est différent des PRMT de type I et favorise la première méthylation sur l'azote ω^2 (figure 22a). De surcroît, l'histidine du motif IV (H415 chez mCARM1) est remplacée par une sérine (S578) chez PRMT5 et une cystéine chez PRMT9. L'espace ainsi dégagé et les contraintes stériques côté E444 (E267 chez mCARM1) imposeraient au groupement méthyle sur ω^2 de se positionner au-dessus de W579 (motif IV), facilitant la seconde méthylation sur ω^1 (figure 22b). Ce mécanisme a pu être « observé » lors d'une simulation de dynamique moléculaire hybride MM/MQ de la méthylation par PRMT5 (Yue et al., 2015). Il est intéressant de mentionner que le second

module catalytique de PRMT9 possède un motif IV SHW ressemblant au motif THW des PRMT de type I. Cependant, le motif II GxGxG est dégénéré en SxGxS indiquant que ce second module pourrait être inactif, à l'instar de PRMT7.

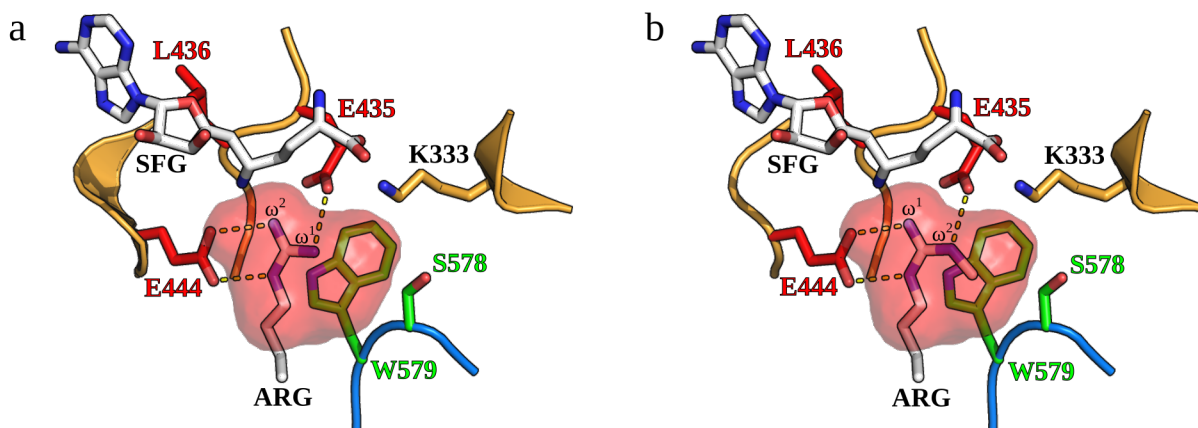


Figure 22 - Déterminants structuraux de la diméthylation symétrique.

(a) Fixation de l'arginine dans le site de méthylation de PRMT5 (PDB 4GQP), en présence de sinefungine. (b) Modélisation d'une monométhylarginine dans la poche de méthylation de PRMT5 (PDB 4GQP). Le site actif de PRMT5 est représenté en ruban et bâtonnets. Le code couleur reprend celui utilisé dans la figure 11. Les substrats arginine, monométhylarginine et l'inhibiteur SFG mimant le SAM sont représentés en bâtonnets blancs. Le volume rouge transparent correspond à la poche de fixation de l'arginine. L'annotation correspond à la PRMT5 humaine.

➤ Déterminants structuraux de la monométhylation

Dans les cellules mammifères, la seule PRMT de type III caractérisée est PRMT7, dont la structure cristallographique a été déterminée au sein de l'équipe (Cura et al., 2014). Dans la structure cristallographique de mPRMT7, le segment peptidique situé en amont du repliement de Rossmann, correspondant à la région qui forme l'hélice α -X dans les PRMT de type I, n'est pas structuré. Cela laisse entrevoir une cavité en lieu et place des sites de fixation du SAM et de l'arginine. Il en va de même pour les PRMT7 de *Caenorhabditis elegans* (Hasegawa et al., 2014a) et *Trypanosoma brucei* (Debler et al., 2016). À l'exception de *T. brucei* pour laquelle l'unité fonctionnelle de PRMT7 est formée du dimère classique, les PRMT7 de la plupart des organismes sont constituées d'un monomère comprenant deux modules catalytiques en tandem mimant une structure pseudo-dimérique. Les structures obtenues dévoilent que le second module catalytique de PRMT7 est figé dans une configuration inactive par un site de fixation du zinc stabilisant l'interaction entre les deux modules en tandem, et par deux boucles dans le site de fixation du SAM du second module catalytique (Cura et al., 2014).

L'hypothèse initiale développée pour expliquer la diméthylation des arginines par les PRMT de type I et II était que la forme dimérique permet de catalyser les deux méthylations successivement dans un module catalytique, puis dans l'autre sans que le substrat ne soit relargué par la PRMT (Gary et al., 1998 ; Lee et al., 1977). En effet, les sites de méthylation des arginines sont en général localisés dans des régions flexibles, si bien que seul un faible réarrangement structural du substrat pourrait être nécessaire pour la transition de l'arginine d'un site actif vers

l'autre. Par extrapolation, il a ainsi été suggéré au sein de l'équipe que la présence d'un unique module catalytique productif chez PRMT7 pourrait expliquer que cette enzyme ne produise que des arginines monométhylées (Cura et al., 2014). En effet après la première méthylation, le relargage du SAH et la fixation d'un nouveau SAM nécessiterait la libération du substrat monométhylé qui, s'il a moins d'affinité pour PRMT7 que le substrat non modifié, ne pourra pas subir la seconde méthylation. Cette hypothèse est néanmoins remise en question par le fait que la PRMT7 de *T. brucei*, formée d'un dimère de deux modules catalytiques productifs, catalyse la production de monométhylarginine. De plus comme PRMT7, PRMT9 possède un second module catalytique en tandem, qui pourrait être inactif, bien que cette enzyme ai été récemment classée parmi les PRMT de type II (Yang et al., 2015b). Cela suggère que la restriction à la monométhylation dépendrait plutôt de contraintes stériques autour du site de méthylation.

➤ Cinétique de la méthylation

Lors d'une réaction enzymatique, l'ordre dans lequel se fixent, sont modifiés puis relargués les substrats et produits, caractérise la cinétique de la réaction catalysée. Deux types de cinétiques existent : les réactions séquentielles et les réactions de type ping-pong. Lors d'un mécanisme de type ping-pong, un premier substrat doit se fixer à l'enzyme et une réaction intermédiaire a lieu, lors de laquelle l'enzyme est modifiée. Cette première réaction permet de catalyser la réaction enzymatique suivante sur un second substrat puis les produits sont relargués les uns à la suite des autres. Dans une réaction séquentielle, tous les substrats doivent être fixés à l'enzyme pour que la réaction puisse avoir lieu. Les réactions séquentielles peuvent être aléatoires et l'ordre de fixation des substrats n'a pas d'importance, ou ordonnée et à l'inverse les substrats doivent fixer l'enzyme dans un ordre défini afin que la réaction chimique puisse avoir lieu. Pour les PRMT, la réaction de substitution nucléophile présentée au-dessus n'est pas compatible avec une cinétique de type ping-pong car le groupement méthyle est transféré directement sur l'arginine, sans que la PRMT ne soit modifiée. Le mécanisme catalysé par les PRMT serait donc de type séquentiel. Cependant, l'ordre dans lequel le substrat et le SAM se fixent à l'enzyme, puis la séquence de relargage du produit méthylé et du SAH sont controversés. Une première étude portant sur l'analyse cinétique de PRMT1 a suggéré un mécanisme dans lequel l'ordre de fixation du SAM et de l'arginine ou de la monométhylarginine serait aléatoire (Obiany et al., 2008). La même année, une autre étude menée sur PRMT6 (Lakowski et al., 2008) suggère au contraire un mécanisme séquentiel lors duquel le SAH se fixerait en premier, suivi du substrat. Le produit serait ensuite libéré le premier. Ce mécanisme est conforté par une étude parue sur PRMT1 en 2011, parvenant aux mêmes conclusions (Feng et al., 2011). De plus, les données structurales obtenues en 2007 sur CARM1 (Troffer-Charlier et al., 2007) ont montré une transition désordre-ordre suivant la fixation du SAM, ce qui suggère que la fixation du cofacteur est nécessaire à l'établissement de la poche de fixation de l'arginine substrat. Les structures cristallographiques de PRMT6 (Bonnefond et al., 2015) montrent

également que la fixation du seul cofacteur ne nécessite pas la fixation de l'arginine substrat et permet le repliement de l'hélice α -X donc l'établissement d'un site catalytique fonctionnel. Larowski et Frankel ont de surcroît démontré dans leur étude de 2008 sur PRMT6 une préférence de l'enzyme pour le substrat monométhylé plutôt que non méthylé. Cela suggère un mécanisme distributif plutôt que processif. En effet, lors d'un mécanisme distributif, il n'y a pas d'accumulation de produit monométhylé puisque celui-ci est diméthylé directement après la première modification (Lakowski et al., 2008). Un mécanisme distributif a également été observé pour PRMT1, PRMT2 et PRMT3 (Kölbl et al., 2009 ; Lakowski et al., 2009). Un mécanisme distributif pourrait expliquer pourquoi la plus petite unité fonctionnelle des PRMT est constituée d'un dimère : la forme dimérique permettrait de procéder à la première méthylation dans le premier module catalytique, puis la seconde méthylation dans le deuxième module catalytique, sans relargage du substrat.

5 - Reconnaissance des méthyl-arginines et déméthylation

a - Reconnaissance des diméthylarginines

Les lysines acétylées sont reconnues par les bromo-domaines et les lysines méthylées sont reconnues par une large variété d'éléments dont les chromo-domaines, les domaines PHD, Tudor, WD40, etc. Le domaine Tudor a également été identifié comme un élément de reconnaissance des diméthylarginines. L'interaction entre ce domaine et les arginines méthylées symétriquement a été pour la première fois démontrée en 2001, par la fixation sDMA-dépendante de SMN aux protéines Sm (Brahms et al., 2001). Plus tard, les domaines Tudor de TDRD1, TDRKH, TDRD6 et SPF30 sont venus enrichir la famille des protéines reconnaissant les sDMA (Chen et al., 2009 ; Côté et al., 2005 ; Vagin et al., 2009). Les Tudors de SMN et SF30B fixent également, mais dans une moindre mesure les sDMA. Les aDMA sont, elles, reconnues spécifiquement par le domaine Tudor de TDRD3 (Yang et al., 2010).

Les structures de plusieurs domaines Tudor en interaction avec des diméthylarginines ont été résolues ces dernières années, dévoilant le mécanisme de reconnaissance de cette modification post-traductionnelle (Liu et al., 2010 ; Sikorsky et al., 2012 ; Tripsianes et al., 2011). Les domaines Tudors sont des structures globulaires conservées d'une soixantaine d'acides aminés, organisées en quatre brins β antiparallèles formant un feuillet. La reconnaissance des diméthylarginines a lieu dans une « cage » cubique formée par quatre résidus aromatiques tyrosine et phénylalanine placés de manière orthogonale les uns par rapport aux autres, tel qu'illustré sur la figure 23. Les diméthylarginines sont stabilisées par des interactions cation- π établies entre la charge positive répartie sur le noyau guanidinium de l'arginine et les deux résidus aromatiques placés en parallèle

au-dessus et au-dessous du guanidinium. Les groupements méthyles forment des interactions CH- π avec les deux résidus aromatiques positionnés de manière orthogonale au plan du guanidinium.

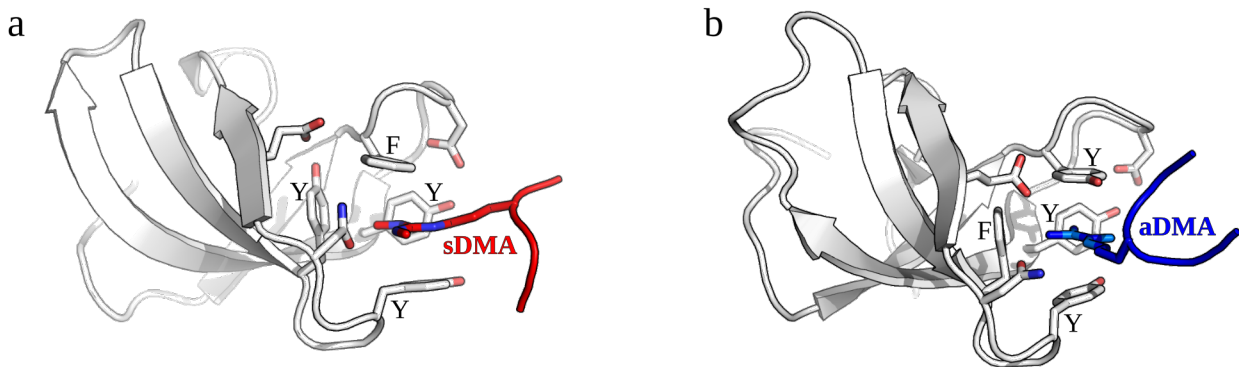


Figure 23 - Reconnaissance des diméthylarginines par les domaines Tudor.

(a) Structure cristallographique du domaine Tudor de SND1 (PDB 3OMC) en complexe avec un peptide de PIWIL1 diméthylé symétriquement sur l'arginine R4. (b) Structure RMN du domaine Tudor de TDRD3 (PDB 2LTO) en complexe avec un peptide de l'ADN polymérase II méthylé asymétriquement sur une arginine. Les domaines Tudor sont représentés en rubans blancs. Les résidus impliqués dans la reconnaissance de la diméthylarginine sont représentés en bâtonnets. Le peptide arborant l'arginine diméthylée est représenté en ruban et bâtonnets rouge et bleu.

De nombreux complexes impliqués dans l'expression des gènes, tels que NurD et MLL, possèdent un domaine WD40 qui interagit étroitement avec H3R2 et leur permet de s'ancrer au nucléosome. Comme indiqué dans le paragraphe I.B.2, le marquage H3R2^{me2s} stimule la fixation de WD40 à H3 ce qui indique une reconnaissance sélective pour cette modification par WD40. Aucune structure du complexe WD40-H3R2^{me2s} n'est cependant disponible pour expliquer les bases structurales de cette interaction.

b - Déméthylation des arginines

La réversibilité est une propriété cruciale des modifications épigénétiques car elle permet d'adapter le niveau d'expression des gènes en fonction des besoins de la cellule. Ainsi, les modifications sont apposées, interprétées puis retirées de la chromatine par des effecteurs auxquels on se réfère souvent par les termes de « writers », « readers » et « erasers ». Les lysines méthylées et acétylées sur les histones sont restaurées par les histones déméthylases (KDM) et désacétylases (HDAC), la phosphorylation des sérines, thréonines et tyrosines est retirée par les histones déphosphorylases et les variants d'histones sont extraits des nucléosomes par leurs chaperons respectifs. La méthylation des arginines est à l'heure actuelle la seule modification épigénétique pour laquelle aucun « eraser » n'a été identifié, si bien que d'autres voies permettant la réversibilité de cette modification ont été explorées.

➤ **Hypothèse de la déimination**

Les peptidylarginine déiminases sont des enzymes catalysant la synthèse de citrulline à partir de l'arginine (McGraw et al. 1999). Il a un temps été considéré que cette enzyme pourrait être responsable du recyclage des monométhylarginines en arginines par déimination (Cuthbert et al.,

2004 ; Wang, 2004). Les peptidylarginine déiminases ont toutefois démontré leur incapacité à déméthyliser les monométhylarginines *in vitro* (Raijmakers et al., 2007). De plus, la déimination des arginines, particulièrement sur les histones, a un rôle biologique spécifique qui diffère des méthylations des arginines (Fuhrmann et al., 2016), si bien que cette modification post-traductionnelle est aujourd'hui davantage considérée comme une marque épigénétique à part entière que comme un intermédiaire de la déméthylation des arginines.

➤ Clivage des queues d'histones

En 2007, la famille des déméthylases d'histones Jumonji, initialement identifiées comme des déméthylases des lysines (Trewick et al., 2005), a été associée à la déméthylation des arginines. Ces enzymes possèdent un domaine Jumonji (jmjC) très conservé, possédant une activité catalytique oxygénase dépendante des cofacteurs Fe^{2+} et 2-oxoglutarate (Pilka et al., 2015). Chang et ses collaborateurs ont suggéré que JMJD6 possède un domaine jmjC et déméthyle *in vitro* et *in situ* H3R2^{me2} et H4R3^{me2} (Chang et al., 2007). Cette hypothèse a été consolidée par des travaux *in vitro* ayant observé des activités arginine déméthylases pour les lysines déméthylases à domaine jmjC KDM3A, KDM4E, KDM5C et KDM6B. Cette observation s'accompagne du relargage de formaldéhyde comme c'est le cas lors de la déméthylation des lysines par cette classe d'enzyme (Walport et al., 2016).

La déméthylation des arginines par les protéines comportant un domaine jmjC demeure toutefois contestée et un autre mécanisme faisant intervenir ce domaine a récemment été proposé. Une étude publiée en 2017 (Liu et al. 2017) a montré que deux protéines à domaines jmjC orphelines, JMJD5 et JMJD7, possèdent une activité endopeptidase et clivent les queues des histones H3 et H4 en aval des méthylarginines, sans les déméthyliser. Cette coupure entraîne la transition de l'ARN polymérase II d'un état de pause vers une élongation de la transcription. En 2018, les structures de JMJD5 en complexe avec des méthylarginines et des peptides ont été résolues par deux équipes (Liu et al., 2018 ; Wilkins et al., 2018). La structure de complexes JMJD5-sDMA obtenue par Liu et ses collaborateurs montre que le groupement peptidique de l'arginine diméthylée se positionne dans le site catalytique (figure 24a), ce qui est compatible avec leur modèle de coupure de la chaîne polypeptidique en aval d'une sDMA.

À l'inverse, Wilkins et ses collaborateurs ont émis l'hypothèse que JMJD5 catalyserait l'hydroxylation du carbone C3 des peptidyl-arginines, et leurs structures du complexe JMJD5-peptide avec et sans inhibiteur montrent une fixation du peptide à JMJD5 incompatible avec une fonction protéolytique mais confortant leur hypothèse de l'hydroxylation (figure 24b/c). L'implication des domaines jmjC dans la déméthylation des arginines demeure donc controversée.

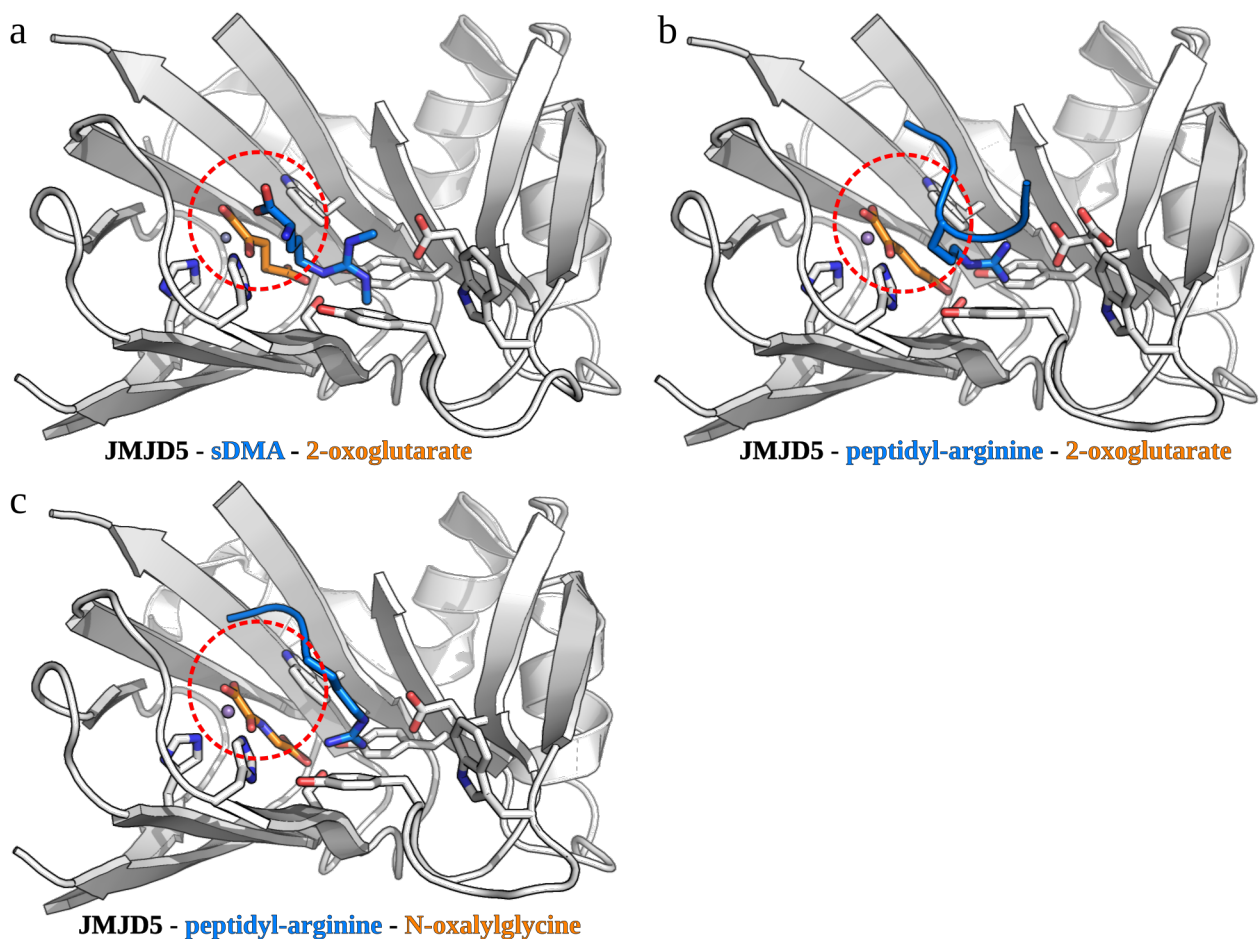


Figure 24 - Mode de fixation du domaine *jmjC* de JMJD5 à ses substrats.

Domaines *jmjC* de JMJD5 en complexe avec une arginine symétriquement diméthylée (a, PDB 6AX3) et un peptide arborant une arginine (b et c, PDBs 6F4S et 6F4T). Le domaine *jmjC* est représentés en rubans blanc. Les résidus du site catalytique sont représentés en bâtonnets. Les peptides substrats et la sDMA sont représentés en rubans et bâtonnets bleus. L'ion métallique catalytique est représenté sous forme de sphère grise. Le cofacteur 2-oxoglutarate et l'inhibiteur N-oxalylglycine sont représentés en bâtonnets oranges. Le site catalytique est encerclé de rouge.

6 - Inhibiteurs des PRMT

Comme présenté au paragraphe B.2 dans ce chapitre, les PRMT sont associées à une grande variété de fonctions biologiques, en particulier dans les processus de régulation de la transcription et de l'épissage, de maintien de l'intégrité du génome et de transduction du signal, mécanismes piliers du contrôle du cycle cellulaire, de la différenciation et de la prolifération cellulaire. La dérégulation de ces processus biologiques est associée au développement de pathologies et de nombreuses études associent la dérégulation de l'expression ou de l'activité des PRMT à différentes formes de cancer, ainsi qu'avec d'autres maladies chroniques et virales. Dans ces pathologies, le niveau d'expression ou d'activité des PRMT est généralement élevé si bien que les PRMT sont considérées aujourd'hui comme des cibles émergentes pour le développement d'inhibiteurs à visée pharmacologique, et sont le fruit d'importants efforts de laboratoires académiques et d'industriels (Cha et al., 2012 ; Morettin et al., 2015 ; Yang et al., 2013). Ce dernier chapitre introductif vise à présenter de façon aussi exhaustive que possible les inhibiteurs développés pour cibler les PRMT. Un premier paragraphe

traitera succinctement de la sinefungine (SFG), un inhibiteur naturel des méthyltransférases. Le second paragraphe présentera les inhibiteurs non naturels des PRMT et sera divisé en deux parties : les inhibiteurs analogues du cofacteur SAM, dont le développement a débuté à la fin des années 2000 et se poursuit aujourd'hui, et les inhibiteurs non analogues du SAM, dont le développement a commencé en 2004. Ces derniers seront classés selon la famille de structure chimique auxquels ils appartiennent.

a - La sinefungine, un inhibiteur naturel des méthyltransférases

Les PRMT fixent les dérivés naturels du SAM tel que le SAH (figure 25a), produit lors de la réaction de méthylation par déméthylation du SAM, et la sinefungine (SFG, figure 25b). Cette molécule est un analogue structural du SAM et est un inhibiteur naturel des méthyltransférases SAM-dépendante, produit par la bactérie *Streptomyces griseolus* et jouant le rôle d'antibiotique antifongique (Berry et al., 1978). L'utilisation de cette molécule à des fins thérapeutiques est toutefois limitée par sa très forte toxicité chez les mammifères (Zweygarth et al., 1986), mais elle est utilisée comme inhibiteur des méthyltransférases dans de nombreux travaux de recherche. Cet inhibiteur est compétitif du cofacteur SAM et inhibe les méthyltransférases de façon non sélective.

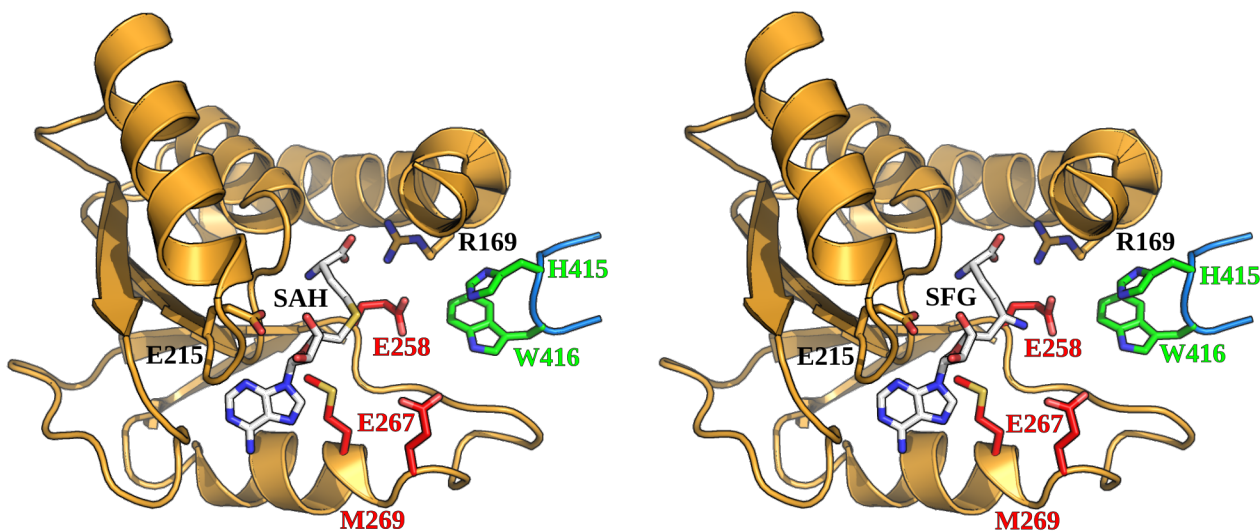


Figure 25 - Inhibiteurs compétitifs naturels des PRMT.

Représentation schématique du site catalytique de mCARM1 en complexe avec du S-adénosylméthionine (a, PDB 5IH3) ou de la sinefungine (b, PDB 5IS6). Le code couleur reprend celui utilisé pour la figure 11. Pour améliorer la lisibilité, les hélices α -X et α -W n'ont pas été représentées.

b - Inhibiteurs artificiels des PRMT

De nombreux inhibiteurs non naturels des PRMT ont été décrits dans la littérature. Les plus efficaces seront présentés plus en détail ci-dessous. Parmi ces inhibiteurs, seule une partie a pu être co-cristallisée avec sa cible. La figure 26 présente la projection des sites occupés par les inhibiteurs des différentes PRMT sur la structure du module catalytique de mCARM1.

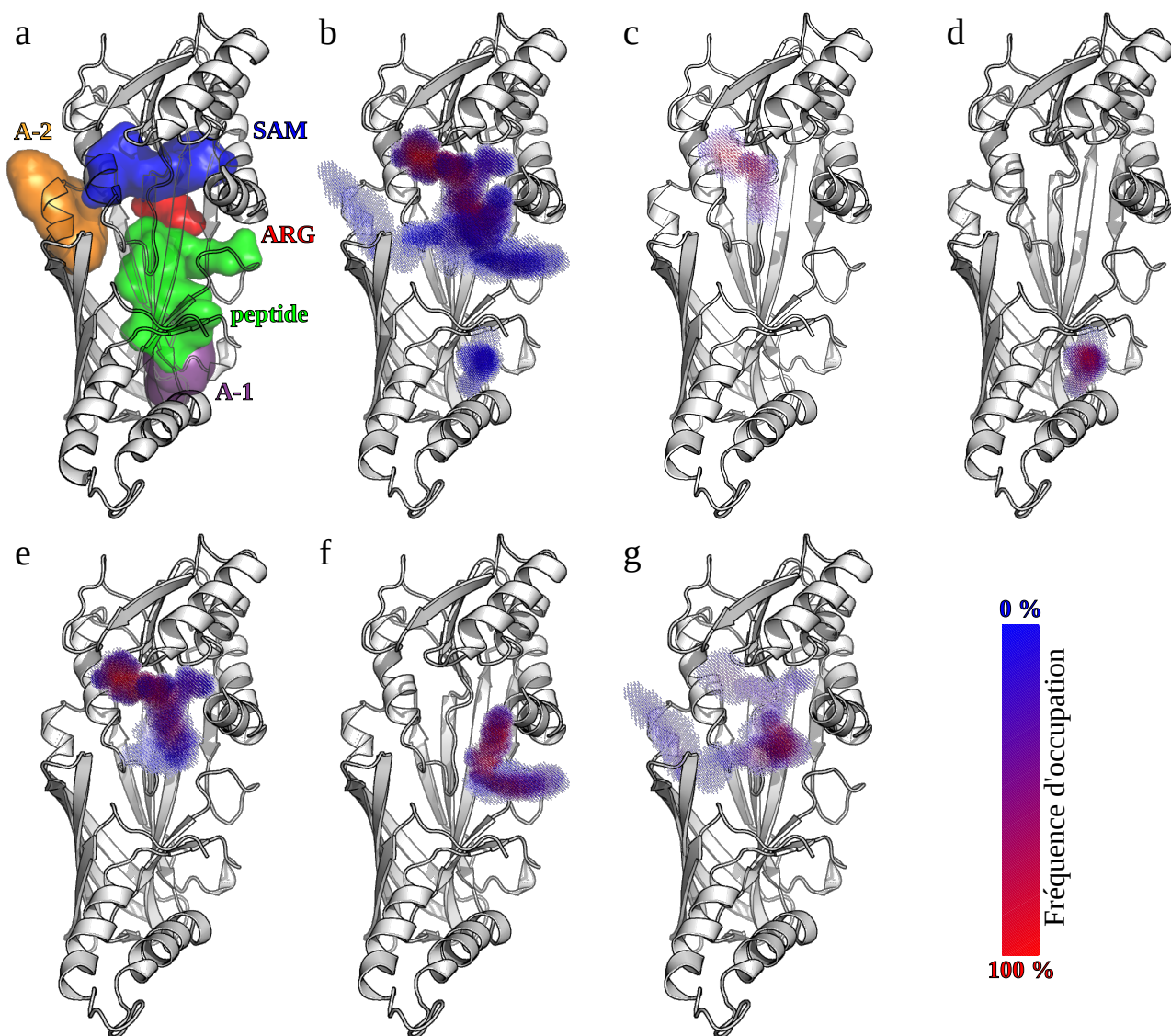


Figure 26 - Sites de fixation des inhibiteurs des PRMT.

(a) Sites de fixation répertoriés sur les PRMT, représentés sur le module catalytique de mCARM1 (PDB 5IH3). Les poches de fixation du SAM et de l'arginine sont représentées respectivement par les volumes bleu et rouge. Le site de fixation du peptide PABP1 (PDB 5LGP) est représenté par le volume vert. Les sites de fixation allostériques A-1 et A-2 sont représentés par les volumes respectivement violet et orange. (b) Projection sur mCARM1 de tous les sites occupés par les inhibiteurs de (c) PRMT2, (d) PRMT3, (e) CARM1, (f) PRMT5 et (g) PRMT6. Le gradient du rouge au bleu représente le taux d'occupation du site, calculé par comptage des atomes des inhibiteurs présents à proximité de chaque point. Pour améliorer la visibilité, les hélices α -W et α -X ont été effacées.

Les inhibiteurs des PRMT peuvent être classés en deux catégories : les composés analogues du SAM et les composés non analogues du SAM. Les composés analogues du SAM sont des inhibiteurs compétitifs du cofacteur possédant des attributs structuraux du SAM, en particulier le groupement adénosine. Cette structure chimique leur permet d'occuper le site de fixation du cofacteur, ce qui empêche de fait le déroulement de la réaction de méthylation (figure 26a). Les développements décrits ci-dessous montrent que ces composés possèdent un important potentiel d'inhibition. De plus, malgré leur similarité avec ce cofacteur, la conception de ces composés a permis d'obtenir une forte sélectivité pour les PRMT comparé à d'autres méthyltransférases. Une

des difficultés avec ce type de composés pourrait provenir de la faible capacité du SAM et de ses analogues à traverser les membranes cellulaires (Halby et al., 2018 ; Zhang et al., 2016), mais ils demeurent néanmoins de bons outils pour l'exploration *in vitro* des interactions PRMT-composés pouvant gouverner la spécificité et la sélectivité de nouveaux inhibiteurs.

La seconde catégorie englobe elle-même plusieurs types d'inhibiteurs. Les composés non-analogues du SAM peuvent être compétitifs du cofacteur en occupant la poche de fixation du SAM, et inhibent l'activité des PRMT par un mécanisme similaire aux analogues du SAM. Les composés compétitifs du substrat occupent la poche de fixation de l'arginine ou l'entrée de cette poche, ce qui empêche les substrats de se fixer aux PRMT. Certains composés peuvent être à la fois compétitifs de la fixation du SAM et du substrat et occupent totalement ou partiellement les deux poches. Finalement, des inhibiteurs allostériques ont été décrits pour PRMT3. Ces inhibiteurs se fixent à distance des poches de fixation du SAM et de l'arginine et provoquent des changements conformationnels pouvant empêcher la fixation de substrats.

Comme l'indique le tableau 4, la majorité des inhibiteurs publiés ont pour cibles PRMT1 et PRMT4/CARM1. Quelques inhibiteurs existent également pour PRMT2, PRMT3, PRMT6, PRMT8 et PRMT5 mais aucun pour PRMT7 et PRMT9. De plus, seule une minorité de tous ces composés a été cristallisée avec leur cible spécifique.

PRMTx	PRMT1	PRMT2	PRMT3	PRMT4	PRMT6	PRMT8	PRMT5	PRMT9	PRMT7
Inhibiteurs publiés	15	1	6	20	4	1	5	0	0
Inhibiteurs cristallisés	0	3	5	9	2	0	4	0	0

Tableau 4 - Inhibiteurs de PRMT publiés et cristallisés avec leurs cibles spécifiques.

Dans la ligne « Inhibiteurs publiés », seuls les meilleurs composés, présentés dans les paragraphes qui suivent, ont été comptabilisés. La ligne « Inhibiteurs cristallisés » décompte pour chaque PRMT le nombre d'inhibiteurs non naturels présents dans les structures cristallographiques. Les PRMT de type I, II et III sont colorées respectivement en bleu, rouge et vert.

Les inhibiteurs des PRMT ciblent principalement 5 poches : la poche de fixation du SAM (bleu), la poche de fixation de l'arginine (rouge), le site de fixation du peptide (vert) et deux régions distantes du site actif notées A-1 et A2 (violet et orange). Comme l'illustre la figure 26b, la majorité des inhibiteurs ayant été co-cristallisés avec leur cible se fixent au site actif des PRMT, en particulier dans les régions occupées par le cofacteur SAM et l'arginine substrat. Pour PRMT2, les inhibiteurs publiés ne ciblent que ces deux sites (figure 26c). En revanche aucun inhibiteur de PRMT3 ne se fixe au site actif, pas plus qu'au site de reconnaissance du peptide. Tous les inhibiteurs connus de cette PRMT sont décrits comme allostérique et fixent le site A-1. Les inhibiteurs de CARM1 ciblent principalement les poches de fixation du SAM et de l'arginine, et une minorité d'entre eux occupent une partie du site peptide à l'entrée de la poche de fixation de l'arginine. Les inhibiteurs de PRMT5 se fixent à la poche de fixation de l'arginine et au site peptide

à proximité de l'hélice α -Y. Les inhibiteurs de PRMT6 sont les plus diversifiés. La majorité d'entre eux occupe le site de fixation de l'arginine, mais on en retrouve également dans la poche de fixation du SAM et sur le site peptide. Quelques inhibiteurs de PRMT6 occupent le second site distant du site actif, A-2.

➤ **Les inhibiteurs analogues du SAM (Fin des années 2000 à aujourd'hui)**

Les inhibiteurs analogues du SAM ont été décrits pour PRMT1, PRMT2, CARM1 et PRMT6. Les premiers inhibiteurs de cette catégorie ont été conçus au sein de l'équipe entre 2008 et 2013 dans le cadre d'une collaboration avec l'équipe du docteur Alain Wagner, directeur du laboratoire de chimie biofonctionnelle de l'université de Strasbourg (Ajebbar, 2012 ; Mailliot, 2013). Le projet avait pour objectif initial d'explorer l'espace chimique dans les poches de fixation du SAM et de l'arginine de CARM1 en vue de déterminer les bases structurales de la spécificité de produit de cette enzyme. Ces inhibiteurs, nommés SA (figure 27) sont divisés en trois groupes : les composés analogues du SAM (SA920 et SA765), les analogues SAM-arginine (SA678, 707, 401, 684, 435 et 375) et les analogues SAM-peptide. Tous ces composés sont constitués d'une partie SAM à laquelle sont ajoutés différents groupements chimiques interagissant avec le site de méthylation et la poche de fixation de l'arginine, comme le montre la figure 28. Ces composés ont démontré une capacité à inhiber sélectivement CARM1 comparé à PRMT1, avec des IC_{50} entre 2,5 μ M et 70 μ M contre 1 μ M pour le SAH.

Van Haren et ses collaborateurs (Université d'Utrecht) ont développé des composés, notés U (U1 à U3), similaire aux inhibiteurs SA, conçus pour mimer l'état de transition des réactions de méthylation des arginines par les PRMT (Cura et al., 2017 ; van Haren et al., 2015). Ces composés sont des inhibiteurs de PRMT2, CARM1 et PRMT6, et en collaboration avec cette équipe de chimistes, nous avons caractérisé les bases structurales de la spécificité et de la sélectivité des inhibiteurs UX pour les PRMT (chapitre VI). Halby et ses collaborateurs ont également développé des composés analogues d'état de transition, notés LH, mais cette fois-ci de la méthylation des acides nucléiques par les méthyltransférases de l'ADN et de l'ARN (DNMTs / RNMTs). Ces composés n'inhibent ni les DNMTs, ni les RNMTs, mais nous avons montré dans le cadre d'une collaboration que ces composés peuvent être utilisés comme structures de départ pour la conception rationnelle d'inhibiteurs des PRMT (chapitre VI) (Halby et al., 2018).

Des composés de type SAM-arginine ont également été synthétisés par Ward, Pike et leurs collaborateurs (Dowden et al., 2010, 2011). Quatre de ces composés (composés 1 à 3, Figure 27) inhibent PRMT1 avec des IC_{50} respectives de 2,9, 3,9, 5,6 et 6,2 μ M comparées à une IC_{50} sub-micromolaire pour la SFG et sont sélectifs de PRMT1 comparé à CARM1 et à la lysine méthyl transférase SET7.

ANALOGUES DU SAM

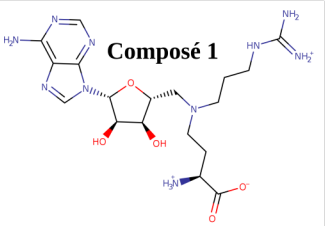
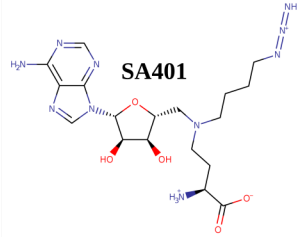
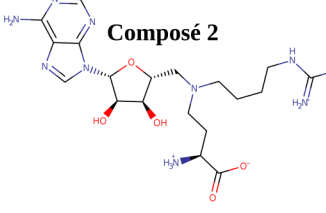
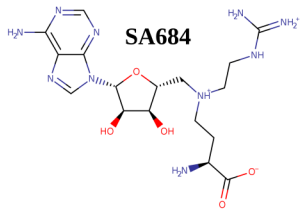
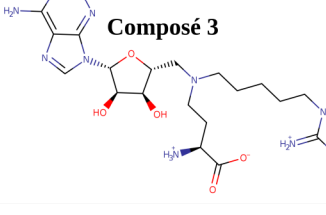
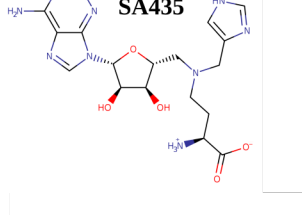

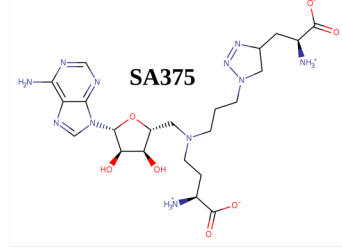
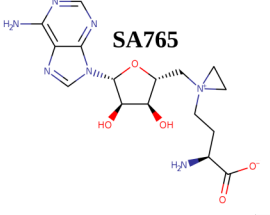
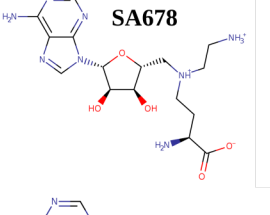
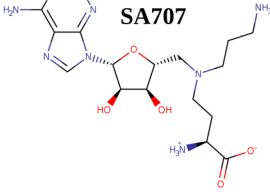
Composé	IC ₅₀	Composé	IC ₅₀
 <p>Composé 1</p>	PRMT1 SET7 (6,2 µM) (>100 µM)	 <p>SA401</p>	CARM1 PRMT1 (14 µM) (16 µM)
 <p>Composé 2</p>	PRMT1 SET7 (2,9 µM) (>50 µM)	 <p>SA684</p>	CARM1 PRMT1 (2,5 µM) (113 µM)
 <p>Composé 3</p>	PRMT1 SET7 (5,6 µM) (>100 µM)	 <p>SA435</p>	CARM1 PRMT1 (43 µM) (295 µM)
 <p>SA920</p>	CARM1 PRMT1 (16 µM) (380 µM)	 <p>SA375</p>	CARM1 PRMT1 (70 µM) (107 µM)
 <p>SA765</p>	CARM1 PRMT1 (13 µM) (- µM)		
 <p>SA678</p>	CARM1 PRMT1 (5 µM) (23 µM)		
 <p>SA707</p>	CARM1 PRMT1 (30 µM) (32 µM)		

Figure 27 - Inhibiteurs analogues du SAM.

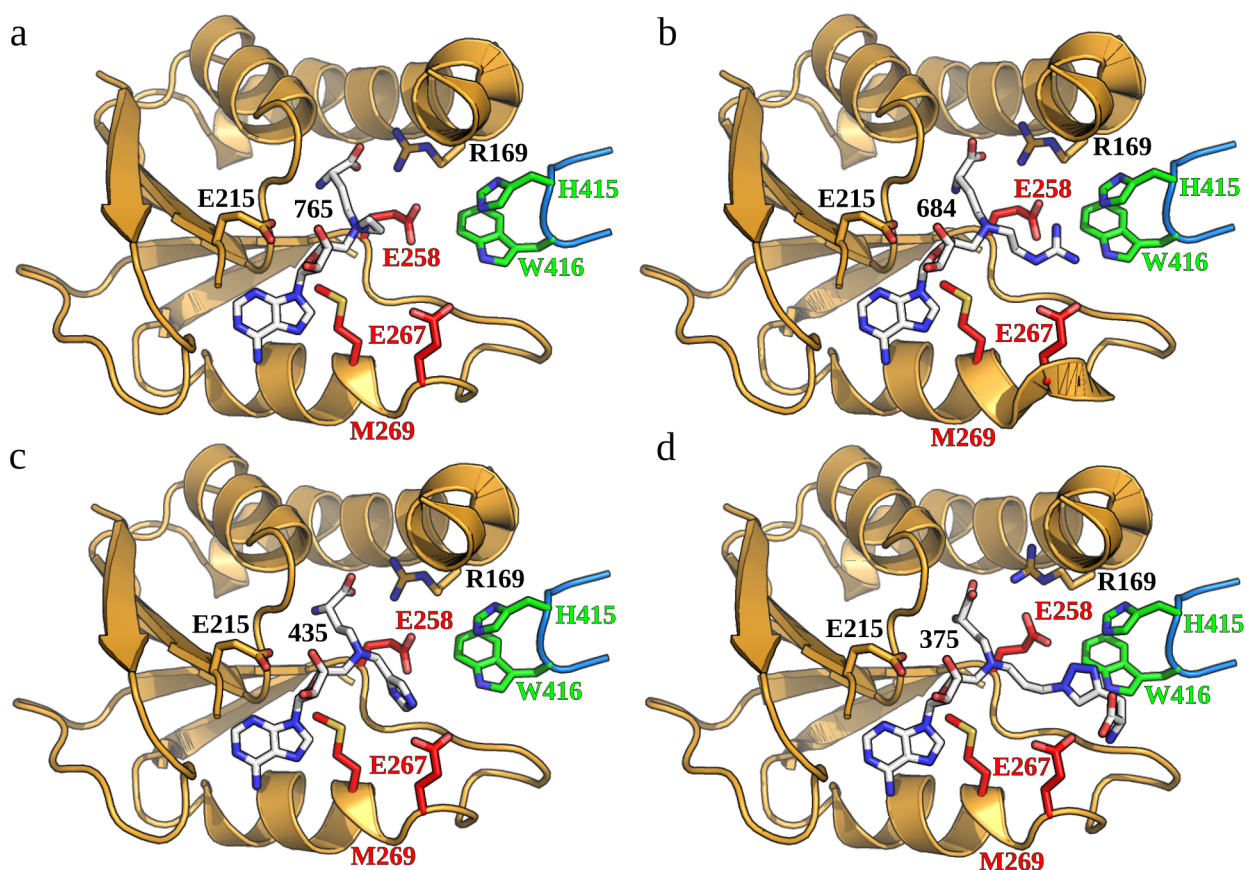


Figure 28 - Exemples de composés SA dans le site catalytique de CARM1 de souris.

Sites catalytiques de mCARM1 en complexe avec l'analogue du SAM SA765 (a, PDB 5ISH), les mimes SAM-arginine SA684 et SA435 (b, PDB 5LKJ & c, PDB 5ISB) et le composé intermédiaire entre un mime SAM-arginine et SAM-peptide SA375 (d, PDB 5IS9). La protéine est représentée en ruban et bâtonnets. Le code couleur reprend celui utilisé pour la figure 11. Pour améliorer la lisibilité, les hélices α -X et α -W n'ont pas été représentées.

➤ Les inhibiteurs non analogues du SAM (2004-Aujourd'hui)

◆ Les composés AMI et AMA (2004-2007)

En 2004, Bedford et ses collaborateurs (Cheng et al., 2004) ont mis en évidence les premiers modulateurs de PRMT1 par criblage d'une banque chimique (DIVETSet, Chembridge ; San Diego, CA) en utilisant des anticorps reconnaissant les diméthylarginines asymétriques. Parmi 9 « hits », ils ont identifié deux molécules (AMI-1 et AMI-6, figure 29) inhibant sélectivement les PRMT, en particulier PRMT1, mais pas les autres méthyltransférases. Ces deux composés ne sont pas compétitifs de la fixation du SAM et ont des IC_{50} respectives de 8,81 et 5,11 μ M pour PRMT1, comparé à 1,63 μ M pour la sinefungine. La structure chimique de certains de ces composés est néanmoins peu adaptée à un usage comme molécule thérapeutique car ils contiennent de nombreux groupements chargés et donneur/accepteurs d'hydrogène. Ces composés inhibent la méthylation des arginines dans des lignées cellulaires (Cheng et al., 2004). Cependant, il a été montré par la suite que les inhibiteurs des PRMT constitués de groupements naphthalène et sulfonyl, tel qu'AMI-1, inhiberaient la méthylation des arginines en interagissant avec le substrat plutôt qu'avec la PRMT (Feng et al., 2010). Dans la même étude, Bedford et ses collaborateurs ont également mis en

évidence deux activateurs de PRMT1 (AMA-1 et AMA-2, figure 29) qui augmentent la méthylation jusqu'à trois fois et pourraient servir à explorer les mécanismes de régulation de l'activité des PRMT.

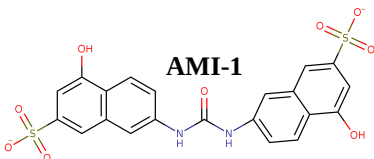
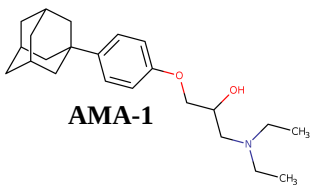
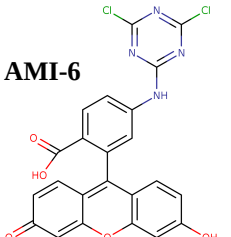
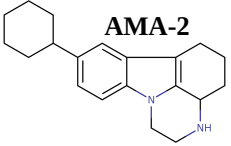
COMPOSES AMI			
Composé	IC ₅₀	Composé	IC ₅₀
 <p>AMI-1</p>	PRMT1 (8,8 μM)	 <p>AMA-1</p>	activateur de PRMT1
 <p>AMI-6</p>	PRMT1 (5,1 μM)	 <p>AMA-2</p>	activateur de PRMT1

Figure 29 - Inhibiteurs non analogues du SAM : les AMI.

◆ *Les dérivés de l'allantodapsonne (2007-2011)*

En utilisant une approche par criblage virtuel d'une banque chimique (Cambridge Crystallographic Data Centre) puis en validant les « hits » par tests d'inhibition biochimique et *in situ* sur lignées cancéreuses, Jung, Sippl et leurs collaborateurs ont pu identifier un nouvel inhibiteur de PRMT1, l'allantodapsonne, ayant une IC₅₀ de 1,7 μM (figure 30) (Spannhoff et al., 2007). Contrairement à AMI-1, sa structure chimique possède des propriétés plus adaptées à l'utilisation comme molécule thérapeutique. Cette molécule est capable notamment de pénétrer dans les cellules. Les résultats de docking suggèrent que l'allantodapsonne est un inhibiteur compétitif de l'arginine substrat.

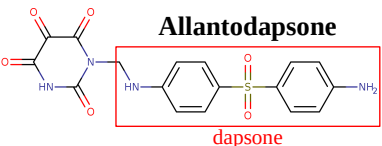
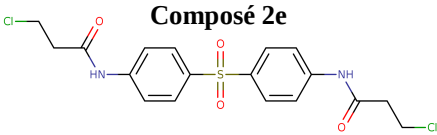
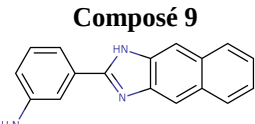
ALLANTODAPSONNE ET DERIVES			
Composé	IC ₅₀	Composé	IC ₅₀
 <p>Allantodapsonne</p> <p>dapsonne</p>	PRMT1 (1,7 μM)	 <p>Composé 2e</p>	PRMT1 (1,5 μM)
 <p>Composé 9</p>	PRMT1 (14,3 μM)		

Figure 30 - Inhibiteurs non analogues du SAM : allantodapsonne et ses dérivés.

Un pharmacophore a été généré à partir de l'allantodapsonne, et utilisé pour le criblage de la base chimique virtuelle « Chembridge » (Heinke et al., 2009). Neuf nouveaux inhibiteurs ont été identifiés, mais leurs IC_{50} étaient significativement moins bonnes que pour le composé précédent : le meilleur, le composé 9, ayant une IC_{50} de 14,34 μM . L'optimisation par SAR (Structure-Activity Relationship) du dapsonne (figure 30), une sous-structure de l'allantodapsonne, a mené à l'identification du composé 2e ayant une IC_{50} de 1,5 μM pour PRMT1 (figure 30), comparable à l'allantodapsonne (Bissinger et al., 2011). Ce composé est cependant fortement cytotoxique, probablement du fait de ses propriétés alkylantes.

◆ *Les amidines (2014-2015)*

Yang, Li et leurs collaborateurs (Yu et al., 2015) se sont basés sur la structure de la stilbamidine (figure 31), un inhibiteur faible de PRMT1 mis en évidence lors du criblage virtuel ayant mené à l'identification de l'allantodapsonne, afin de développer de nouveaux composés. La stilbamidine a pour particularité d'être symétrique et de posséder deux groupements benzamidine (figure 31), qui consiste en un groupement benzène lié à un groupement amidine : deux azotes liés à un même carbone, avec les électrons des deux liaisons covalentes délocalisés sur les trois atomes. Le tout forme une structure chimique plane ressemblant au noyau guanidinium des arginines.

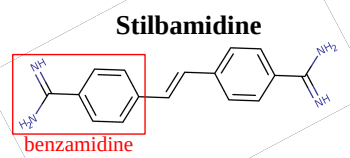
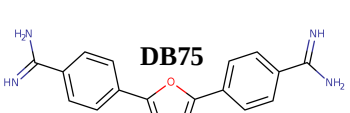
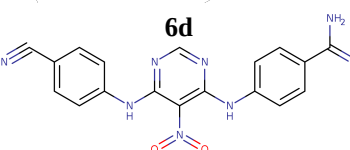
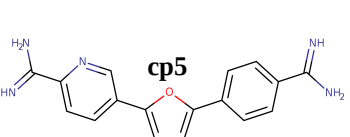
DERIVES DES AMIDINES			
Composé	IC_{50}	Composé	IC_{50}
 <p>Stilbamidine</p>	PRMT1 (15 μM) CARM1 (>400 μM) PRMT5 (44 μM) PRMT6 (173 μM)	 <p>DB75</p>	PRMT1 (9 μM) CARM1 (>400 μM) PRMT5 (166 μM) PRMT6 (283 μM)
 <p>6d</p>	PRMT1 (2 μM) CARM1 (10 μM) PRMT5 (>100 μM) PRMT6 (>100 μM)	 <p>cp5</p>	PRMT1 (7 μM) CARM1 (>400 μM) PRMT5 (186 μM) PRMT6 (211 μM)

Figure 31 - Inhibiteurs non analogues du SAM : les amidines.

L'optimisation par SAR de la benzamidine a produit un composé synthétisé (6d, figure 31) inhibant sélectivement PRMT1 comparé à CARM1, PRMT5 et PRMT6. Dans le test d'activité, son IC_{50} est 10 fois inférieure à celle d'AMI-1 mais 4 fois supérieure au SAH. Ce composé est prédit par docking comme interagissant via le groupement amidine avec le site de fixation de l'arginine ce qui le classerait comme inhibiteur compétitif du substrat. Le composé 6d inhibe la prolifération cellulaire de trois lignées tumorales. Au sein de l'équipe, la détermination de la structure cristallographique de mCARM1 (28-507, PDB 3B3J) avait nécessité l'utilisation de benzamidine comme additif (Troffer-Charlier et al., 2007), et on observe la fixation du composé en plusieurs endroits du module catalytique.

Zheng, Ivanov et leurs collaborateurs ont également exploré la piste des amidines (Yan et al., 2014). Ils ont synthétisé trente composés symétriques qualifiés de di-amidines. Deux de ces composés (DB75 et le composé 5, figure 31) ont démontré leur capacité à inhiber PRMT1 (IC_{50} respectivement 9,4 μM et 7,2 μM) comparé à la stilbamidine (15,2 μM dans l'évaluation) et sont sélectifs de PRMT1 par rapport à CARM1, PRMT5 et PRMT6. Selon leur étude, le DB75 occuperait le site de fixation de l'arginine et partiellement le site du cofacteur et serait donc compétitif des deux substrats. Ce composé peut traverser les membranes cellulaires et inhibe la prolifération cellulaire dans plusieurs lignées leucémiques (Yan et al., 2014).

◆ *Les cyanines (2012-2015)*

Afin de développer des ligands fluorescents de PRMT1 permettant de traquer cette enzyme dans les cellules, Zheng, Henary et leurs collaborateurs ont synthétisé des dérivés des triméthyle cyanines (Sinha et al., 2012). Les cyanines sont des colorants utilisés en biologie du fait de leurs propriétés fluorescentes et la possibilité de les modifier chimiquement pour qu'elles interagissent avec divers partenaires.

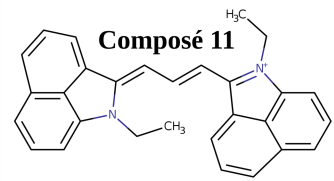
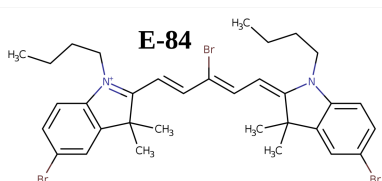
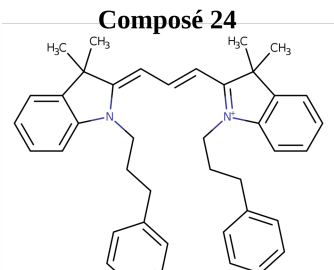
DERIVES DES CYANINES			
Composé	IC_{50}	Composé	IC_{50}
 <p>Composé 11</p>	PRMT1 (4,1 μM) CARM1 (100 μM) PRMT3 (8,3 μM) PRMT5 (13 μM)	 <p>E-84</p>	PRMT1 (3,4 μM) PRMT5 (25 μM) CARM1 (25 μM) PRMT6 (3,0 μM)
 <p>Composé 24</p>	PRMT1 (6,3 μM)		

Figure 32 - Inhibiteurs non analogues du SAM : les cyanines

Les auteurs ont pu identifier 13 composés dont les composés « 11 » et « 24 », présentés en figure 32, ayant des IC_{50} respectives de 4,1 et 6,3 μM pour PRMT1. Le composé 11 est sélectif pour PRMT1, PRMT3, PRMT5 et PRMT6 comparé à CARM1. Pour rendre ces composés plus sélectifs, les auteurs ont procédé à une optimisation par SAR et ont pu mettre en évidence le composé E-84 (figure 32, sélectif pour PRMT1 et PRMT6 comparé à CARM1, PRMT5 et PRMT8 avec des IC_{50} proches de 3 μM pour PRMT1 et PRMT6. Les membranes cellulaires sont perméables à E-84 et ce composé inhibe la méthylation *in vivo* des arginines et la prolifération de lignées cellulaires leucémiques (Hu et al., 2015). Bien que les composés 11 et 24 aient été décrits comme non-

compétitifs par analyse enzymologique, les simulations de docking suggèrent que E-84 est compétitif de l'arginine substrat.

◆ *Les dérivés de la curcumine (2008-2013)*

La curcumine, colorant du curcuma, a été initialement caractérisée comme inhibiteur des acétyltransférases des histones (Balasubramanyam et al., 2004).

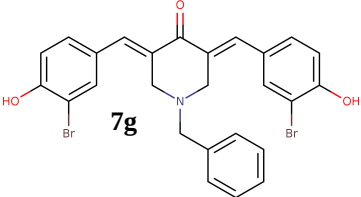
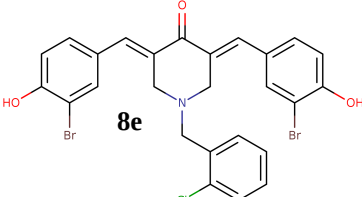
DERIVES DE LA CURCUMINE			
Composé	IC ₅₀	Composé	IC ₅₀
 <p>7g</p>	PRMT1 (>500 μM) CARM1 (8,6 μM) SET7 (>500 μM) -	 <p>8e</p>	PRMT1 (>500 μM) CARM1 (8,1 μM) SET7 (174 μM) -

Figure 33 - Inhibiteurs non analogues du SAM : les dérivés de la curcumine.

La curcumine a été identifiée comme inhibiteur de CARM1 par criblage radiométrique, lors d'un effort de recherche d'inhibiteurs de plusieurs enzymes épigénétiques (Mai et al., 2008). L'optimisation par SAR de dérivés de la curcumine (Cheng et al., 2011a) a mené au développement des composés 8e et 7g, (figure 33). Ces inhibiteurs ont des IC₅₀ de 8,1 et 8,6 μM (contre 18 μM pour AMI-5 dans le même test d'activité) et sont sélectifs de CARM1 contre PRMT1 et SET7. Le composé 7g a démontré une inhibition dose-dépendante sur l'activité du promoteur de l'antigène prostatique spécifique sur des lignées cellulaires humaines de l'adénocarcinome de la prostate. Le taux de surexpression de cet antigène est mesuré dans le dépistage du cancer de la prostate (Atan et al., 2013). Le mode de fixation de ces inhibiteurs n'est actuellement pas connu.

◆ *Les dérivés du pyrazole et de l'imidazole (2008-Aujourd'hui)*

Des composés dérivés des hétérocycles azotés pyrazole et imidazole ont été décrits comme inhibiteurs de PRMT1, CARM1, PRMT5 et PRMT6. Comme l'illustre la figure 34, la majorité de ces composés sont sélectifs de CARM1 puisque 8 d'entre eux sur les 11 présentés inhibent CARM1 comparé à PRMT1, PRMT3 et les méthyltransférases SET7/9 avec des IC₅₀ de l'ordre du micromolaire ou submicromolaire. Par criblage haut débit d'une banque chimique, Purandare et ses collaborateurs (Purandare et al., 2008) ont identifié un dérivé aminé du pyrazole ayant une IC₅₀ de 1,8 μM pour CARM1 (composé 1, figure 34). L'optimisation par SAR du noyau 1-phényl-3-trifluorométhyle-pyrazol de ce composé a mené à l'identification du composé 7b (figure 34) ayant une IC₅₀ de 80 nM pour CARM1 et significativement sélectif de CARM1 par rapport à PRMT1 et PRMT3. En se basant sur cette dernière molécule, Allan et ses collaborateurs (MethylGene Inc) ont mis au point les composés 7a et 1b (figure 34) possédant tous deux des IC₅₀ de 60 nM pour CARM1

(Allan et al., 2009). 7a est fortement sélectif de CARM1 comparé à PRMT1 et SET7/9 mais aucun des deux composés n'a démontré d'activité sur lignée cellulaire. Par la suite, il a été démontré que la substitution du groupement alaninamide par une fonction diamine augmentait la stabilité du composé et facilitait son élimination (Therrien et al., 2009). L'analyse par docking a suggéré que le composé se fixe dans le site de fixation du SAM.

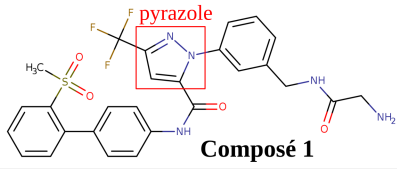
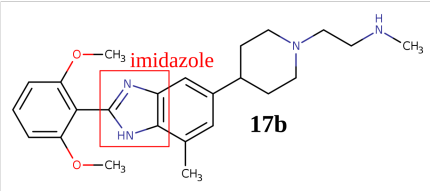
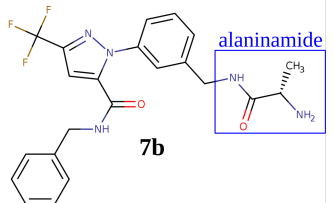
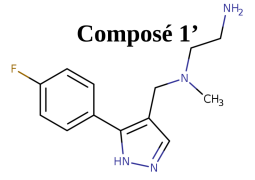
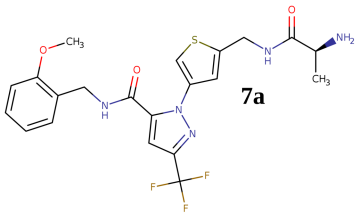
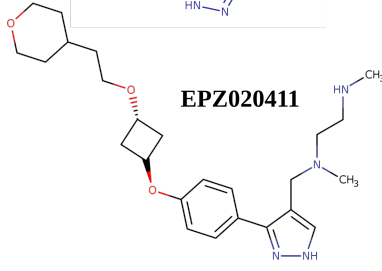
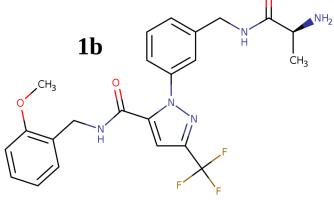
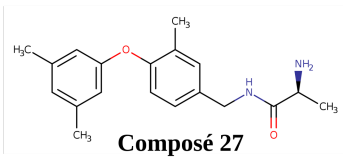
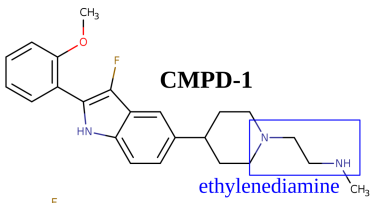
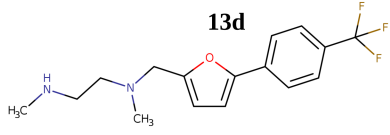
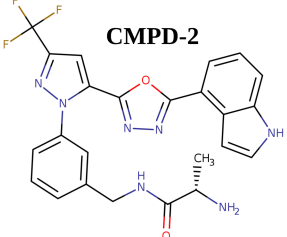
DERIVES DU PYRAZOLE ET DE L'IMIDAZOLE			
Composé	IC ₅₀	Composé	IC ₅₀
 Composé 1	CARM1 (1,8 µM)	 17b	PRMT1 PRMT3 (>25 µM) (>25 µM) CARM1 - (0,07 µM)
 7b	PRMT1 PRMT3 (>25 µM) (>25 µM) CARM1 - (0,08 µM)	 Composé 1'	PRMT1 PRMT6 (0,02 µM) (0,01 µM) PRMT8 - (0,07 µM)
 7a	PRMT1 SET7/9 (>100 µM) (>100 µM) CARM1 - (0,06 µM)	 EPZ020411	PRMT1 PRMT6 (0,19 µM) (0,01 µM) PRMT8 - (0,22 µM)
 1b	CARM1 (0,06 µM)	 Composé 27	PRMT1 CARM1 (>100 µM) (0,05 µM) PRMT5 PRMT6 (>100 µM) (5,2 µM)
 CMPD-1	PRMT1 PRMT3 (>10 µM) (>10 µM) CARM1 - (0,03 µM)	 13d	PRMT1 (8,2 µM)
 CMPD-2	PRMT1 PRMT3 (>10 µM) (>10 µM) CARM1 - (0,03 µM)		

Figure 34 - Inhibiteurs non analogues du SAM : les dérivés du pyrazole et de l'imidazole.

Sack et ses collaborateurs ont synthétisé deux composés, l'un dérivé de l'indole (CMPD-1, figure 34), le second dérivé du pyrazole (CMPD-2, figure 34) dont la structure chimique est proche de 7a (Sack et al., 2011). Ces deux composés ont des IC₅₀ respectives de 30 et 27 nM pour CARM1

et sont sélectifs de cette enzyme comparé à PRMT1 et PRMT3. Leur fixation est dépendante du SAH et les structures cristallographiques de ces composés en complexe avec CARM1 et du SFG (figure 35) montrent que les groupements piperidin-1-yl-N-méthylethanamine et benzyl-L-alaninamide respectifs de CMPD-1 et CMPD-2 occupent le site de fixation de l'arginine tandis que leurs noyaux indole et pyrazole occupent l'entrée de la poche de l'arginine (PDB 2Y1W et 2Y1X) (Sack et al., 2011). Wan et ses collaborateurs (Bristol-Myers Squibb) ont par la suite mis au point le composé 17b (Wan et al., 2009) dérivé de l'imidazole présentant une IC_{50} de 70 nM pour CARM1 et sélectif comparé à PRMT1 et PRMT3 (figure 34).

C'est également dans cette famille de composé qu'a été identifié le premier inhibiteur de PRMT6, développé par la firme Epizyme (Mitchell et al., 2015). Le composé « 1' » (figure 34) a été obtenue par criblage biochimique de leur banque chimique interne. Ce composé possède une IC_{50} de 11 nM pour PRMT6 mais n'est pas sélectif contre PRMT1 (20 nM) et PRMT8 (70 nM). Le composé a été optimisé par SAR jusqu'à obtenir l'inhibiteur EPZ020411 (figure 34) sélectif de PRMT6 contre PRMT1 et PRMT8 tout en conservant une IC_{50} de 10 nM. Les structures cristallographiques de PRMT6 en complexe avec les composés « 1' » (PDB 4Y2H) et EPZ020411 (PDB 4Y30) montrent, comme observé dans les structures cristallographiques du complexe CARM1 – CMPD-1/2, que le groupement éthylènediamine occupe le site de fixation de l'arginine. La molécule possède une biodisponibilité compatible avec son utilisation *in vivo*.

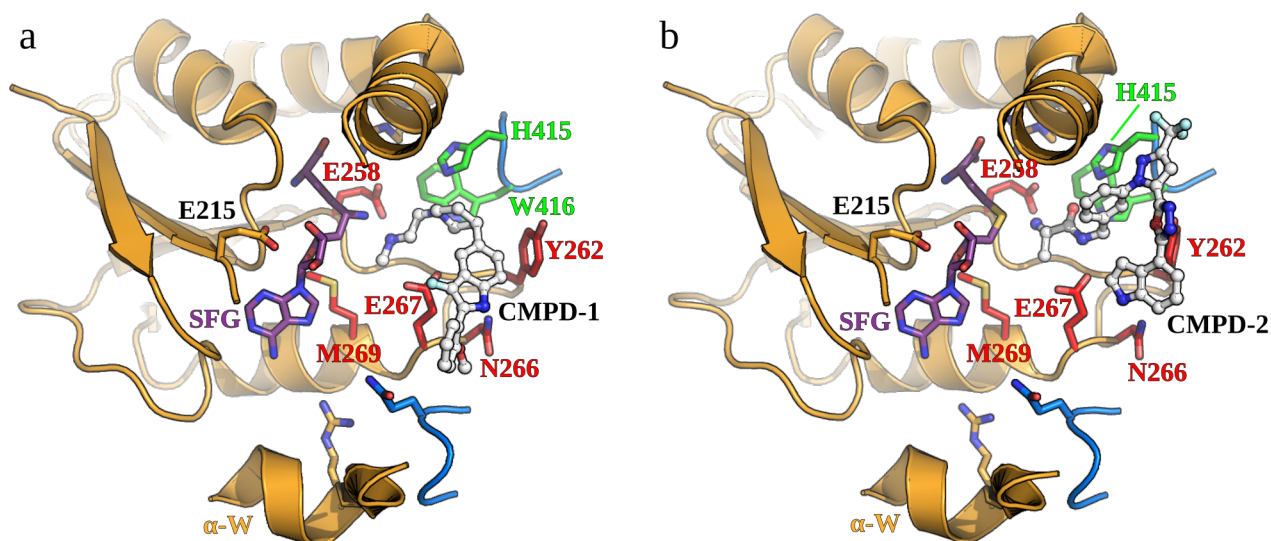


Figure 35 - Modes de fixation des composés CMPD-1 et CMPD-2 à CARM1.

Fixations des composés CMPD-1 (a, PDB 2Y1W) et CMPD-2 (b PDB 2Y1X) au module catalytique de CARM1 de souris. Les inhibiteurs sont représentés en modèles boules-bâtons blancs. Le code couleur reprend celui de la figure 11.

L'une des caractéristiques communes aux composés cités ci-dessus est qu'ils possèdent tous un groupement éthylènediamine, glycinamide ou alaninamide interagissant avec le site de fixation de l'arginine substrat de CARM1. Un criblage virtuel des composés incorporant des sous-structures éthylènediamine et alaninamide dans la banque de données ZINC a été réalisé (Ferreira de Freitas et al., 2016). Les meilleurs « hits » ont été validés par des tests d'activité biochimique puis leurs

structures ont été optimisées en prenant en compte les modes de fixation observés lors du docking *in silico*, jusqu'à obtention du composé 27 (figure 34). Cet inhibiteur est sélectif pour CARM1 comparé à PRMT1, PRMT3, PRMT5, PRMT6, PRMT7 et PRMT8, possède une IC₅₀ de 50 nM pour CARM1 et inhibe la méthylation par cette enzyme dans des cellules HEK293.

En 2014, Hong et ses collaborateurs ont montré par modélisation et dynamique moléculaire que le résidu D84 de PRMT1 humaine, interagissant avec le groupement NH₃⁺ de la partie méthionine du SAM, est un site cible pour le développement d'inhibiteurs (Hong et al., 2014). Récemment, ils ont conçu, synthétisé et testé l'inhibition de 28 composés dérivés d'un cycle tétrazole ou furane, et auxquels peuvent être agencés d'autres groupements, afin de cibler le résidu D84 (Yang et al., 2017). Le meilleur composé (13d, figure 34) possède une IC₅₀ de 8,2 μM pour PRMT1 contre 0,55 μM pour le SAH, et inhibe efficacement la prolifération d'une lignée cancéreuse hépatique. À nouveau, ce composé possède un groupement ethylènediamine ce qui suggère que ce type de structure chimique possède un fort potentiel pour le développement d'inhibiteurs des PRMT.

◆ Les inhibiteurs allostériques (2012-2015)

Des inhibiteurs non compétitifs du SAM et du substrat ont été décrits uniquement pour PRMT3, et tous appartiennent à une même classe de molécules dérivés d'un « hit » obtenu par criblage biochimique d'une banque de 16000 composés (Siarheyeva et al., 2012).

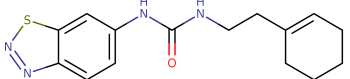
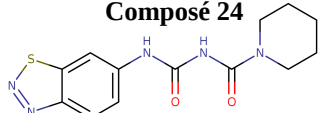
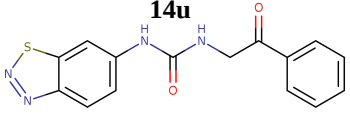
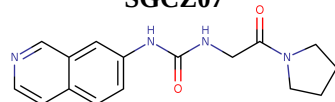
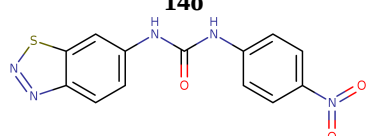
INHIBITEURS ALLOSTERIQUES			
Composé	IC ₅₀	Composé	IC ₅₀
<p>Composé 1</p> 	<p>PRMT3 (2,50 μM)</p>	<p>Composé 24</p> 	<p>PRMT3 (0,23 μM)</p>
<p>14u</p> 	<p>PRMT3 (0,48 μM)</p>	<p>SGCZ07</p> 	<p>PRMT3 (0,03 μM)</p>
<p>14o</p> 	<p>PRMT3 (0,30 μM)</p>		

Figure 36 - inhibiteurs non analogues du SAM : les allostériques.

Siarheyeva et ses collaborateurs ont ainsi identifié un inhibiteur (composé 1, figure 36) ayant une IC₅₀ de 2,5 μM pour PRMT3 et sélectif de l'enzyme contre PRMT1, CARM1, PRMT5 et PRMT8 ainsi que contre les méthyltransférases G9a, EHMT1, SUV31H1, SETD7 et SETD8

(Siarheyeva et al., 2012). Des analyses enzymatiques ont montré que cet inhibiteur possède un mode d'action allostérique.

La structure cristallographique du composé 1 en complexe avec PRMT3 (PDB 3SMQ) montre que cette molécule se fixe dans une cavité formée par la jonction entre le bras de dimérisation et le tonneau β d'un monomère de PRMT3 et interagit avec l'hélice α -Y du monomère adjacent. Cela empêche le repliement de l'hélice α -X (figure 9, page 41) et donc la formation du site catalytique, ce qui suggère que l'inhibiteur empêche le repliement de la PRMT dans une configuration productive. Cela bloque la progression de la réaction à la différence d'un mode d'action par allostérie. Le composé traverse efficacement les membranes cellulaires mais est fortement métabolisé. L'affinité et la stabilité de ce composé ont été optimisées par Liu et ses collaborateurs (Liu et al., 2013) jusqu'à obtention d'inhibiteurs possédant des IC_{50} sub-micromolaires pour PRMT3. Les trois meilleurs (composés 14u, 14o et 24, figure 36) ont des IC_{50} de 480, 300 et 230 nM. La co-cristallisation des composés 14u, 14o et 24 avec PRMT3 a permis de concevoir le composé SGCZ07 (PDB 4RYL, figures 36 et 37) sélectif pour PRMT3 et actif au niveau cellulaire avec une IC_{50} de 31 nM (Kaniskan et al., 2015).

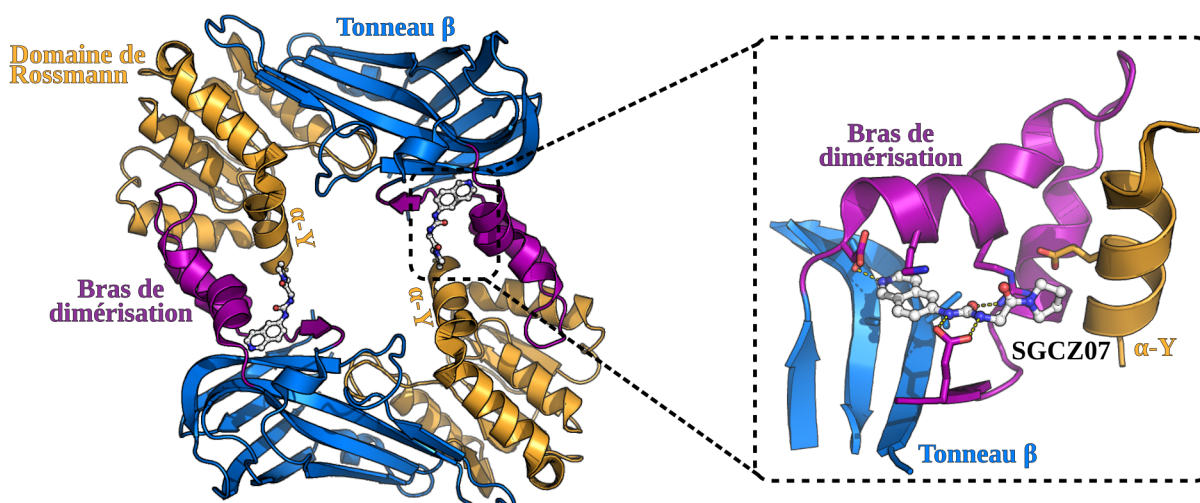


Figure 37 - Mode de fixation de SGCZ07 à la PRMT3 humaine.

Détails de la fixation du composé SGCZ07 (PDB 4RYL) au dimère de modules catalytiques de PRMT3 humaine. Les inhibiteurs sont représentés en modèles boules-bâtons blancs. Le code couleur reprend celui utilisé dans la figure 11.

◆ *Autres types d'inhibiteurs*

Récemment, une approche par pharmacophore suivie de tests d'activité a permis d'identifier 5 inhibiteurs de PRMT5 (Mao et al., 2017). L'optimisation de ces molécules a mené au développement du composé « 17 » (figure 38), ayant une IC_{50} de 330 nM et inhibant sélectivement la prolifération d'une lignée cellulaire leucémique. Des analyses enzymologiques suggèrent une compétition avec le SAM, confirmant des poses obtenues par docking. Une seconde classe d'inhibiteurs de PRMT5 a été identifiée par criblage biochimique haut débit dans le cadre d'une

collaboration entre plusieurs laboratoires pharmaceutiques (Chan-Penebre et al., 2015 ; Duncan et al., 2015 ; Rioux et al., 2016). Un premier « hit » (EPZ007345, figure 38 présentant une IC₅₀ de 326 nM contre PRMT5 a été obtenu et une campagne d'optimisation par SAR a permis d'identifier deux inhibiteurs (EPZ015666 et EPZ015866, figure 38) hautement sélectifs de PRMT5 (IC₅₀ de 22 et 4 nM) contre une large gamme d'autres enzymes produisant des modifications post-traductionnelles. De plus ces deux composés possèdent une faible vitesse d'élimination dans le plasma et inhibent la prolifération de lignées cellulaires de lymphome. Des essais sur des microsomes humains et murins démontrent que ces composés ont une biodisponibilité compatible avec un usage *in vivo*. Les structures cristallographiques de PRMT5 co-cristallisées avec ces composés montrent qu'ils occupent le site de fixation de l'arginine et l'entrée du site catalytique en présence de SFG (figure 39a), et sont donc des inhibiteurs compétitifs de la fixation du substrat. En 2016, l'inhibiteur de PRMT5 EPZ015938, issu de la même famille de composés que EPZ015666 et EPZ015866, a entamé des tests cliniques de phase I pour le développement de thérapies contre les lymphomes non hodgkinien et les tumeurs solides (Copeland, 2018).

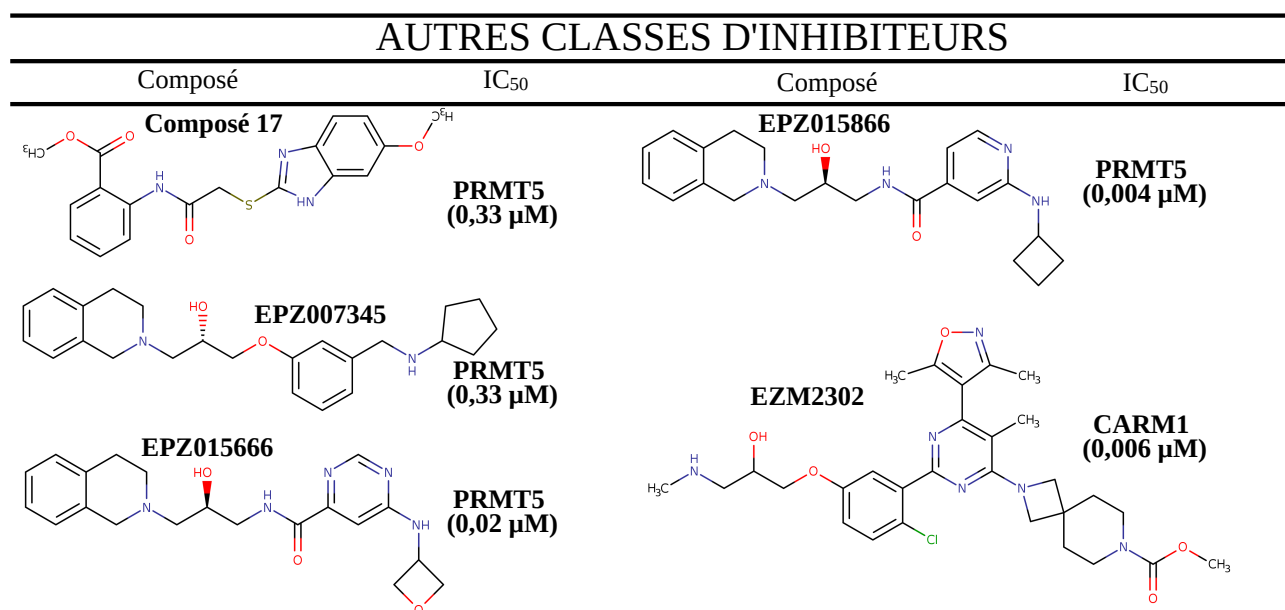


Figure 38 - Inhibiteurs non analogues du SAM : autres cas.

Une collaboration entre Epizyme et GSK (Drew et al., 2017) a également permis de mettre au point le composé EZM2302 (figure 38) par conception rationnelle basée sur les structures cristallographiques puis l'optimisation du ligand. Comme précédemment, les structures cristallographiques montrent que EZM2302 occupe le site de fixation de l'arginine et l'entrée du site catalytique de CARM1 en présence de SAH (figure 39b). Ce composé a une IC₅₀ *in vitro* de 6 nM, et 11 nM *in situ*, et est sélectif pour CARM1 contre une large gamme de méthyltransférases, et a démontré un effet anti-prolifératif sur plusieurs lignées cancéreuses.

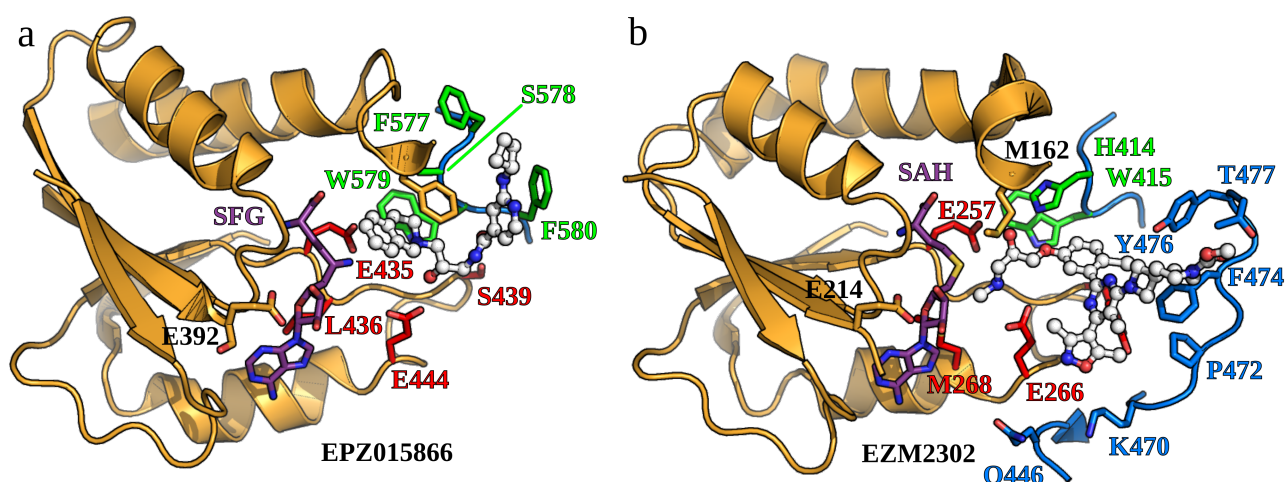


Figure 39 - Modes de fixation des composés EPZ015866 et EZM2302 à PRMT5 et CARM1.

Fixations du composé EPZ015866 dans PRMT5 humaine (a, PDB 5C9Z) et EZM2302 dans CARM1 humaine (b, PDB 6ARJ). Les inhibiteurs sont représentés en modèles boules-bâtons blancs. Le code couleur reprend celui de la figure 11.

Aujourd'hui aucun inhibiteur sélectif de PRMT7, PRMT8 et PRMT9 n'a été décrit et il sera nécessaire de résoudre plus de structures de PRMT en complexes avec ces inhibiteurs pour pouvoir les optimiser de manière rationnelle et obtenir des inhibitions et sélectivités optimales.

c - Regard critique sur le développement d'inhibiteurs des PRMT

À l'heure actuelle, de nombreux d'inhibiteurs des PRMT efficaces ont été développés mais la plupart d'entre eux ne ciblent que PRMT1, CARM1 et PRMT5 et aucun n'est encore utilisé dans un cadre thérapeutique. Plusieurs obstacles restent à franchir dans le développement d'inhibiteurs sélectifs de ces enzymes. D'abord, la détermination du pouvoir inhibiteur des composés est très dépendante du protocole expérimental utilisé : par exemple AMI-1, régulièrement utilisé comme référence, possède des IC_{50} de 5,1, 1,2 et 55 μ M selon que les inhibitions ont été mesurées par méthode colorimétrique (Cheng et al., 2004), par fluorimétrie à résolution temporelle (Spannhoff et al., 2007) ou par FRET à résolution temporelle (Xie et al., 2014). Il est nécessaire que les divers acteurs impliqués dans le développement d'inhibiteurs s'accordent sur un protocole commun, fiable et reproductible afin de pouvoir comparer efficacement les divers inhibiteurs entre eux et choisir les composés les plus prometteurs à optimiser. Par exemple, le SAH et la SFG devraient être utilisés systématiquement comme référence pour la détermination des IC_{50} en plus des inhibiteurs déjà décrits car la fixation aux PRMT et la capacité inhibitrice de ces ligands a été prouvée, tandis que les inhibiteurs cités ci-dessus, rarement cristallisés avec les PRMT, pourraient agir par effet dénaturant ou en précipitant la protéine. Une fois cette harmonisation des méthodes réalisée, il faudra aussi standardiser les substrats utilisés pour les tests d'inhibition *in vitro*. En effet certaines mesures d' IC_{50} ont été réalisées sur des substrats peptidiques et d'autres sur des substrats entiers, ce qui peut affecter la stabilité du complexe PRMT-substrat et de ce fait influencer le résultat de

l'expérience. Afin de se rapprocher des conditions *in vivo*, des substrats entiers devraient préférentiellement être employés. La nature des substrats doit également être ajustée afin que les substrats préférentiels pour chaque PRMT soient employés. L'utilisation de substrats « par défaut », par exemple un peptide GAR fusionné à une étiquette GST (Glutathion S-Transférase), pourrait ne pas mener au même résultat qu'un peptide spécifique à la PRMT étudiée. En effet, dans le cas des inhibiteurs compétitifs du substrat, l'IC₅₀ mesurée résulte directement du différentiel d'affinité entre les complexes PRMT-inhibiteur et PRMT-substrat, si bien que l'utilisation d'un substrat peu spécifique peut mener à la mesure d'une inhibition sur-évaluée.

Pour conclure, il faudrait systématiquement mesurer l'effet des inhibiteurs des PRMT sur des lignées cellulaires « saines » afin de démontrer l'innocuité de ces composés sur l'homéostasie cellulaire en général. Ces enzymes sont en effet impliquées dans une très grande diversité de processus cellulaires vitaux.

Travaux de thèse

II - Étude structurale et fonctionnelle de SmPRMT3

A - Introduction

Notre équipe est depuis plusieurs années impliquée dans le développement d'inhibiteurs des PRMT. Dans le cadre du projet A-ParaDDisE, nous avons entrepris de déterminer la structure de la PRMT3 de *Schistosoma mansoni* (SmPRMT3) par diffraction des rayons-X et cryo-microscopie électronique afin d'identifier les caractéristiques structurales qui différencient la PRMT3 humaine de celle de schistosome et qui pourrait permettre de développer des inhibiteurs spécifiques de SmPRMT3. De plus, nous nous sommes également intéressés au complexe formé entre SmPRMT3 et son substrat potentiel la protéine ribosomique S2 de *S. mansoni*. Les PRMT3 eucaryotes établissent une interaction conservée avec rpS2. Cette interaction et la méthylation de rpS2 pourraient être nécessaires à la maturation du ribosome, et comprendre le mode de fixation entre ces deux protéines représenterait une avancée à la fois pour la compréhension mécanistique de la reconnaissance PRMT-partenaire et pour le développement de composés inhibant cette interaction.

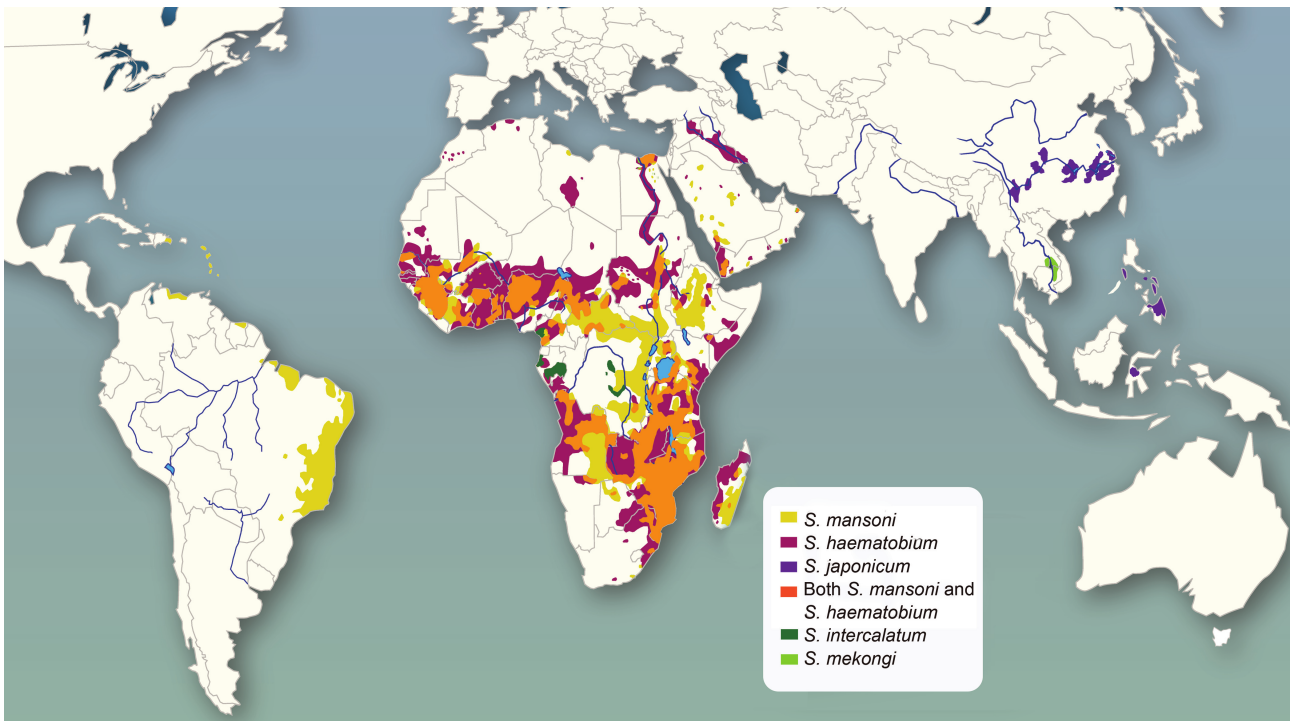


Figure 40 - Carte de répartition de la schistosomiase.

Le code couleur représente les différentes espèces impliquées dans les schistosomiases, par zone (Weerakoon et al., 2015).

La schistosomiase est une infection tropicale présente dans 78 pays en voie de développement, particulièrement d'Afrique subsaharienne et d'Amérique du Sud (figure 40). Cette pathologie atteint une population estimée à au moins 240 millions de personnes, en particulier des enfants et jeunes adultes. La population à haut risque de contamination approcherait les 850

millions d'individus (Tchuem Tchuente et al., 2017 ; Weerakoon et al., 2015). Les schistosomiasés sont provoquées par 6 espèces de schistosomes (figure 40) et se manifestent par des symptômes très variés allant de l'anémie chez les jeunes enfants jusqu'au développement de tumeurs, en passant par des dysfonctions du foie, des reins, de la circulation sanguine et s'accompagnant souvent d'une perturbation des fonctions cognitives (Organisation Mondiale de la Santé). Rien qu'en Afrique subsaharienne, les cas de décès attribués aux schistosomiasés dépasseraient les 200000 annuel et s'y ajoutent les répercussions économiques dues aux difficultés d'apprentissage des enfants touchés et à la diminution de la participation des malades au développement économique (Organisation Mondiale de la Santé).

Actuellement, les infections par schistosomes sont majoritairement traitées par le praziquantel (Cabezas-Cruz et al., 2014). Cette molécule est active contre tous les schistosomes, particulièrement leur forme mature. Elle est peu toxique et on ne déplore que des effets secondaires négligeables, ce qui permet son utilisation massive (Vale et al., 2017). Cette utilisation a néanmoins favorisé l'apparition de souches de schistosomes résistantes en Afrique subsaharienne. Des efforts sont réalisés afin de synthétiser des dérivés du praziquantel permettant de surmonter l'apparition des résistances, mais la cible moléculaire du praziquantel n'est pas identifiée, pas plus que son mécanisme d'action, ce qui limite la possibilité d'optimiser rationnellement cette molécule.

La complexité du cycle de vie de ces trématodes, qui infectent deux hôtes gastéropode et mammifère, et traversent plusieurs cycles de développement (Weerakoon et al., 2015), implique une régulation subtile de l'expression de leurs gènes par un contrôle épigénétique (Cabezas-Cruz et al., 2014). Par conséquent les développements thérapeutiques se concentrent aujourd'hui sur la conception d'inhibiteurs des enzymes de modification des histones.

Dans une étude sous financement Européen (SETTREND ID 241865, projet A-ParaDDisE) impliquant plusieurs instituts de recherche, des enzymes de modification des histones ont été identifiées par ARN interférence dont la désacétylase d'histones SmHDAC8, et les méthyltransférases SmKDM1 et SmPRMT3. L'extinction de ces enzymes par ARN interférence dans les schistosomes bloque le cycle de vie de ces parasites et peut être létale pour ces trématodes, ce qui justifie la conception d'inhibiteurs contre ces enzymes. SmHDAC8 est actuellement la cible la plus étudiée car sa structure est la plus éloignée de son homologue chez l'humain, comparée aux autres HDACs (Cabezas-Cruz et al., 2014). Le mécanisme structural de l'inhibition et de la sélectivité des inhibiteurs de SmHDAC8 a été décrit par une équipe de l'IGBMC (Marek et al., 2013).

B - Résultats et discussion

1 - Choix des constructions utilisées

La séquence de SmPRMT3, longue de 553 acides aminés, a dans un premier temps été comparée à celles de ses homologues dans d'autres organismes afin de mettre en évidence les similarités et différences avec les autres PRMT3. La figure 41a schématise l'organisation de SmPRMT3. Comme ses homologues, SmPRMT3 est organisée en deux régions ordonnées : un module catalytique classique additionné d'un domaine amino-terminal caractéristique des PRMT3. Afin de mieux visualiser le positionnement de ces insertions et la conservation structurale du site actif, un modèle par homologie a été construit avec le programme Modeller (Sali et al., 1993) à partir de l'alignement 2. Selon cet alignement et le modèle présenté figure 41b/c/d, le module catalytique et le site actif sont structurellement très conservés et les quatre motifs caractéristiques sont préservés. En effet, comme le montre la figure 41c/d, les résidus de SmPRMT3 impliqués dans la fixation du cofacteur et de l'arginine substrat sont similaires à ceux de la protéine humaine.

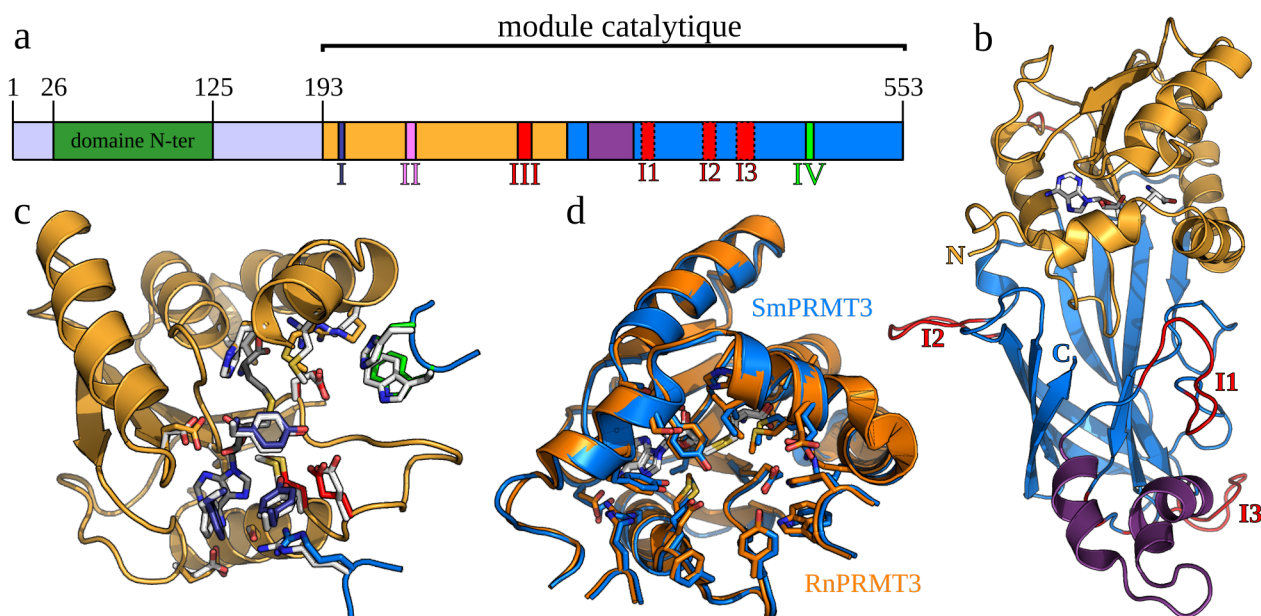


Figure 41 - Modèle par homologie de SmPRMT3.

(a) Représentation schématique de la séquence complète de SmPRMT3 mettant en évidence le domaine amino-terminal. Le code couleur correspond à celui utilisé en figure 11. Les insertions sont représentées en rouge. (b / c) Modèle par homologie du domaine catalytique de SmPRMT3. (c) Sites de fixation du SAM et de l'arginine dans le modèle de SmPRMT3. Les résidus correspondant de la PRMT3 de rat (PDB 1F3L) ont été représentés en bâtonnets blancs. Le SAH est représenté en bâtonnets gris. Les résidus de SmPRMT3 appartenant aux motifs I, III et IV ont été représentés respectivement en bleu foncé, rouge et vert. (d) Alignement des sites de fixation du SAM et de l'arginine substrat du modèle de SmPRMT3 et de la structure de PRMT3 de rat, mettant en évidence la conservation structurale de SmPRMT3.

Le module catalytique de SmPRMT3 se caractérise par la présence de trois grandes insertions dans le domaine organisé en tonneau β . Celles-ci sont représentées en rouge sur la figure 41. La première insertion, I1, longue de 9 résidus, se situerait dans la séquence en aval du bras de dimérisation et structurellement à proximité du motif IV. La seconde insertion notée I2, longue de 9

résidus également, serait insérée entre l'hélice α I et le brin β 9. La dernière insertion, I3, est longue de 11 résidus formés d'une répétition de sérine et asparagine, et se placerait entre les brins β 9 et β 10. Son positionnement à l'extrémité du tonneau β pourrait indiquer qu'elle prolonge les deux brins. Ces trois séquences sont insérées dans une zone de faible conservation donc leurs localisations exactes ne sont pas connues.

a - Clonage et expression de plusieurs constructions de SmPRMT3

Lors de mon arrivée au laboratoire nous disposions de deux clones de PRMT3 de *Schistosoma mansoni* correspondant aux fragments 1-553 (complet/fl) et 141-553 (domaine amino-terminal excisé/C1) clonés respectivement dans les plasmides pnEAthX et pnEAvhX (Diebold et al., 2011).

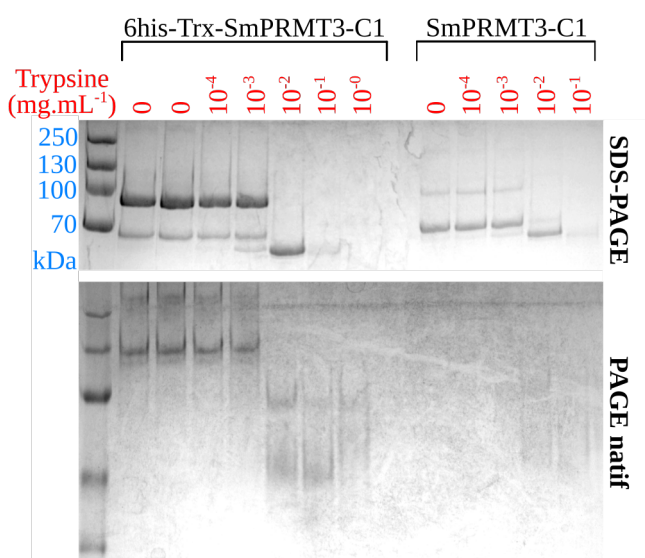


Figure 42 - Trypsinolyse de SmPRMT3-C1.

Des échantillons de 6his-Trx-TEV-SmPRMT3-C1 et son produit de digestion par la protéase TEV ont été digérés par des concentrations croissantes en trypsine (rouge) puis analysés en parallèle par SDS-PAGE 12 % et native-PAGE 12 %. Lorsque la protéine est agrégée, elle pénètre peu ou pas dans le gel de concentration natif, permettant de discriminer entre les fragments résistants présentant ou non un intérêt à être clonés puis produits.

Une des premières étapes à toute étude structurale est la délimitation de fragments stables et solubles de l'objet étudié. La pureté des macromolécules, leur homogénéité en solution sont des prérequis à leur cristallisation. Afin de délimiter les régions ordonnées de SmPRMT3, des digestions ménagées ont été pratiquées. La digestion ménagée consiste à digérer un échantillon en présence de concentrations croissantes en diverses protéases. Les parties de l'échantillon non structurées sont plus sensible aux protéases et seront protéolysées, tandis que les parties repliées résistent à de plus grandes concentrations en enzyme. En identifiant les sites de clivage, on peut ensuite cloner les constructions plus stables et homogènes de l'échantillon.

La trypsinolyse de SmPRMT3-C1, présentée en figure 42, a permis de mettre en évidence une bande résistante à 10^{-2} mg.mL⁻¹ de trypsine, qui n'est de surcroît pas sujette à l'agrégation. L'agrégation est repérable lors d'une électrophorèse en condition native lorsque l'échantillon reste concentré sur le fond du puits de dépôt. La bande résistante a été analysée par spectrométrie de masse en tandem par la plateforme de protéomique de l'IGBMC, mais la couverture de la séquence n'a pas permis d'identifier précisément le site de clivage. En nous basant sur la masse moléculaire apparente sur gel et sachant que la trypsine digère la chaîne polypeptidique en aval des résidus

lysine et arginine, nous avons estimé que les fragments résistant les plus probables sont les fragments 171-553 (C2) et 184-553 (C3). Le fragment C3 a par la suite été réutilisé pour des digestions supplémentaires par les protéases α -chymotrypsine, papaïne, subtilysine et endoprotéinase afin d'en affiner les limites. À nouveau, la couverture obtenue par spectrométrie de masse n'a pas permis d'identifier précisément les sites de coupure, mais de la même manière que précédemment nous avons estimé que les bandes correspondraient aux fragments 199-553 (C4) et 210-553 (C5). La localisation de ces fragments sur SmPRMT3 est représentée dans la figure 43.

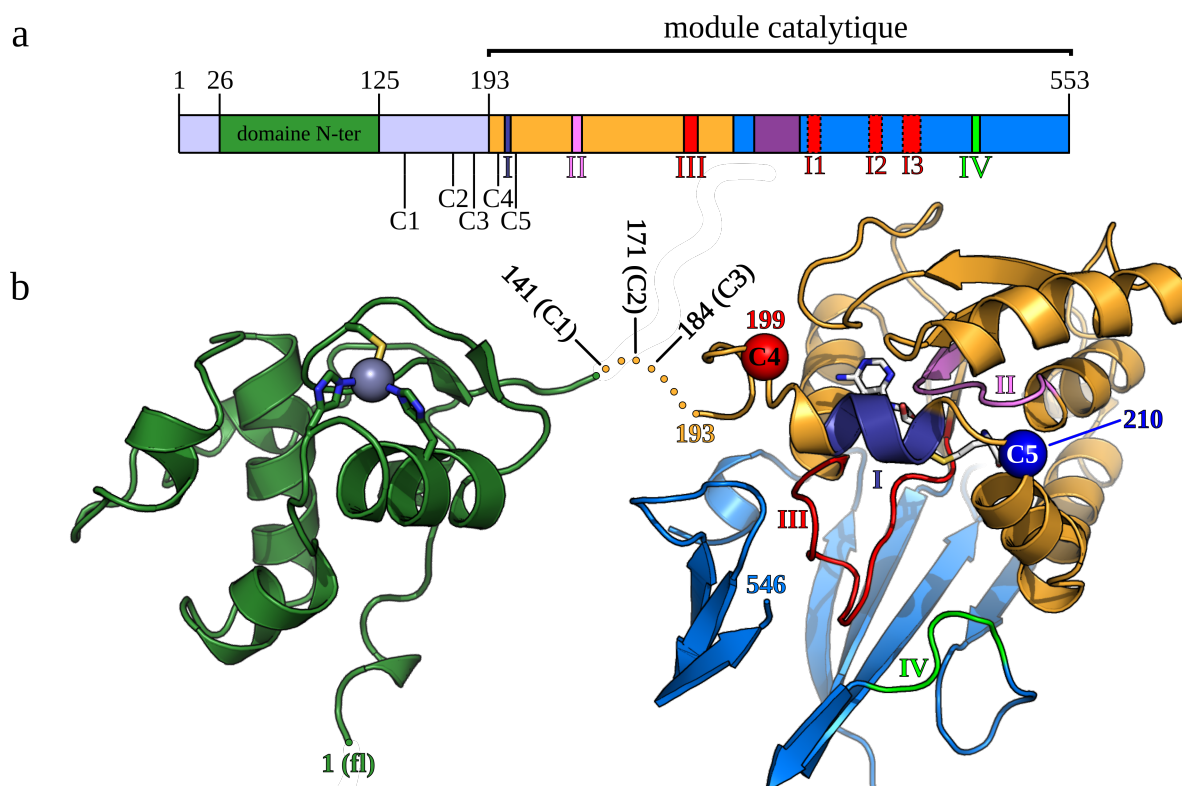


Figure 43 - Modèle par homologie de SmPRMT3 avec indication des constructions utilisées.

(a) Représentation en ruban de la séquence de PRMT3 de *S. mansoni*. Le domaine amino-terminal est représenté en vert. Les insertions sont représentées en rouge et sont numérotées I1 à I3. Les motifs sont numérotés de I à IV en dessous du ruban. Les constructions des fragments de SmPRMT3 sont numérotées C1 à C5. (b) Indication des constructions de SmPRMT3 clonées sur le modèle calculé. Le domaine amino-terminal correspond à la structure RMN du domaine N-ter de mPRMT3 (PDB 1WIR). Les sphères rouges et bleue indiquent les limites des fragments C4 et C5 sur le module catalytique. Le code couleur de la PRMT reprend celui utilisé en figure 11.

Selon l'alignement 2 page 82, et comme illustré sur la figure 43, les fragments C1 à C4 sont localisés en amont du motif I, qui délimite le début du module catalytique. Le fragment C5 se situe juste en aval du motif I et correspond donc à un module catalytique dont l'hélice α -X, participant à la formation des sites de méthylation et de fixation du cofacteur, est absente. Ainsi C5 présente moins d'intérêt pour des études structurales en vue d'un développement de composés à action pharmacologique. La structure cristallographique de la PRMT3 de rat (PDB 1F3L) (Zhang et al., 2000) a été obtenue avec le fragment 184-528, correspondant pour SmPRMT3 au résidu 168, proche des fragments C2 et C3. Le succès obtenu pour la construction de PRMT3 de rat conforte donc notre choix de fragments pour la cristallisation de PRMT3 de *S. mansoni*.

Les constructions suivantes ont été réalisées : SmPRMT3-fl (1-553) avec deux étiquettes 6his-Trx-(thr) et 6his-(nc), SmPRMT3-C1 (141-553) et -C2 (171-553) avec le tag 6his-Trx-(TEV), SmPRMT3-C3 (184-553) avec trois tags 6his-Trx-(TEV)/-(thr)/-(3C), 6his-(TEV) et 6his-(nc), SmPRMT3-C4 (199-553) avec deux tags 6his-(nc) et (nc)-6his, et SmPRMT3-C5 (210-553) avec le tag 6his-(nc).

6his et 6his-Trx correspondent à une étiquette hexahistidine seul ou couplé avec une protéine de solubilisation, la thiorédoxine, et (thr), (TEV), (3C) ou (nc) correspondent respectivement à des sites de clivage par la thrombine, la protéase du virus de la mosaïque du tabac, la protéase 3c, ou à aucun site de clivage. Ces constructions ont été réalisées séquentiellement en fonction des résultats obtenus lors des purifications et essais de cristallisation. Les séquences, poids moléculaire, points isoélectriques et coefficients d'extinction moléculaire théoriques des constructions fl à C5 sont disponibles en annexes 1 et 2.

b - Purification des fragments de SmPRMT3

Les premiers efforts se sont portés sur les constructions de SmPRMT3-fl et C1 qui étaient déjà disponibles lors de mon arrivée. Un grand nombre d'essais de purification ont été réalisés, avec les tags 6his-Trx- et 6his-(thr/TEV/nc)- et différentes étapes de purification. La présence de tags sur les protéines peut interférer avec l'établissement des interactions nécessaires à la formation d'un cristal. Par conséquent, il est souvent recommandé de cliver ce tag en amont des essais de cristallisation. À l'inverse, il arrive que les tags créent des interactions fortuites favorisant la formation des cristaux, si bien que des essais avec et sans coupure du tag ont été réalisés. Le protocole de purification de l'ensemble des constructions de SmPRMT3 peut être divisé en trois phases. La protéine est d'abord extraite du lysat par une chromatographie d'affinité pour le cobalt (Imobilized Metal ion Affinity Chromatography, IMAC) grâce au tag 6his présent dans chacune des constructions. Ensuite, l'échantillon subit une étape de purification par chromatographie échangeuse d'anions, qui permet de séparer SmPRMT3 de plusieurs contaminants et particulièrement des acides nucléiques qui fixent plus fortement la résine cationique. Une résine anionique n'était pas utilisable avec SmPRMT3 car le pI théorique de cette protéine se situe autour de 5,5, si bien qu'un pH inférieur à 5, probablement dénaturant, aurait été nécessaire pour charger positivement SmPRMT3. Finalement, SmPRMT3 est purifiée par chromatographie d'exclusion de taille (ou filtration sur gel), permettant d'isoler la protéine des derniers contaminants et de corps agrégés. Une étape supplémentaire optionnelle de clivage des tags a été ajoutée lors de la chromatographie IMAC, ou avant la filtration sur gel. L'étape de concentration en amont de la filtration sur gel est réalisée par précipitation au sulfate d'ammonium pour les fragments C3 et C4, et sur membrane centrifugeable pour les autres.

Le tableau 5 résume les constructions testées en purification et présente brièvement leurs conditions d'expression et leurs comportements durant la purification.

Construction	Expression dans <i>E. coli</i>	Solubilité avant clivage	Solubilité après clivage	Stabilité avant clivage	Stabilité après clivage
6his-Trx-thr-SmPRMT3-fl	30 °C, 6 h	+++	+++	+++	++
6his-SmPRMT3-fl	30 °C, 6 h	+++	non clivable	+++	non clivable
6his-Trx-TEV-SmPRMT3-C1	20 °C, 12 h	+++	+/-	+++	++
6his-Trx-TEV-SmPRMT3-C2	20 °C, 12 h	non testé	+	non testé	++
6his-Trx-TEV-SmPRMT3-C3	20 °C, 12 h	+++	+++	+++	++
6his-TEV-SmPRMT3-C3	20 °C, 12 h	+++	+++	+++	++
6his-SmPRMT3-C3	20 °C, 12 h	+++	non clivable	+++	non clivable
6his-SmPRMT3-C4	20 °C, 12 h	+++	non clivable	+++	non clivable
SmPRMT3-C4-6his	20 °C, 12 h	+	non clivable	+++	non clivable

Tableau 5 - Conditions d'expression et comportements en purification des constructions utilisées.

La figure 44 montre les chromatogrammes des étapes d'échangeuse d'anions (a) et de filtration sur gel (b) pour la construction complète de SmPRMT3 supplémenté d'un simple tag 6his, et l'analyse des fractions par SDS-PAGE (c).

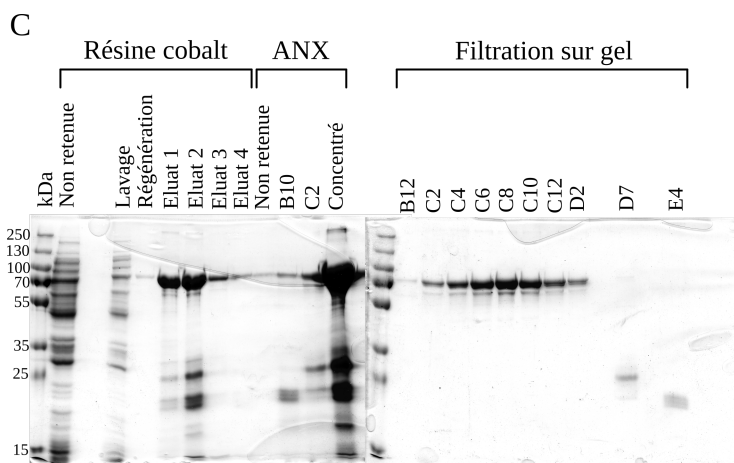
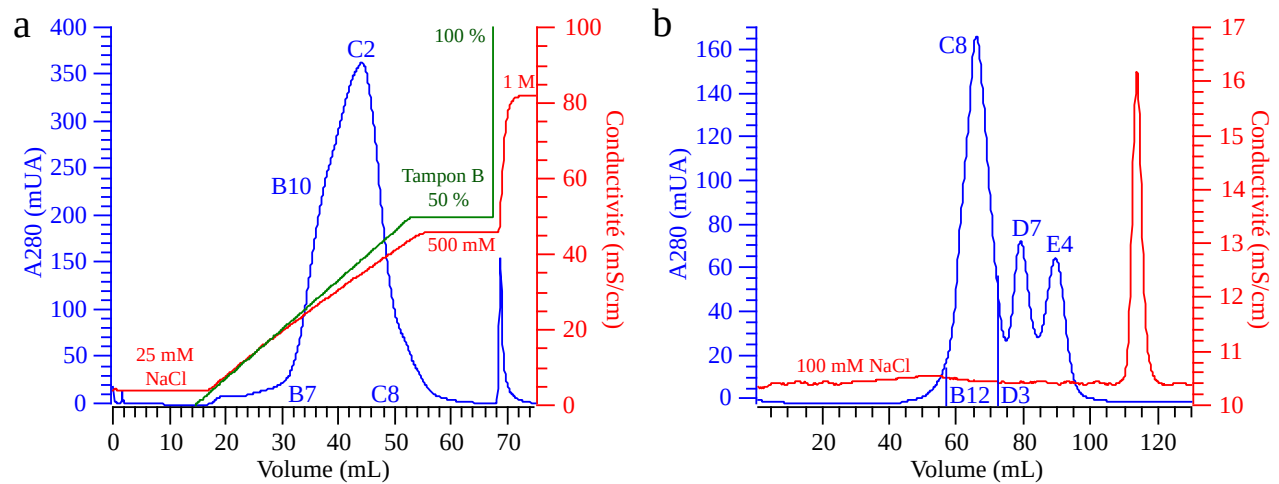


Figure 44 - Purification de 6his-SmPRMT3-fl.

(a) Chromatographie échangeuse d'anions HiTrap ANX FF. Les fractions B7 à C8 ont été réunies, concentrées puis purifiées par filtration sur gel. (b) Chromatographie par filtration sur gel Superdex 200 (16/60). SmPRMT3 est efficacement séparée de deux contaminants présents dans les pics D7 et E4. (c) Analyse par SDS-PAGE sur gel de polyacrylamide à 12 % des étapes de purification. Coloration au bleu de Coomassie. 6his-SmPRMT3-fl est efficacement exprimé dans *E. coli* puis purifié depuis l'extrait brut. L'analyse par électrophorèse montre que cette construction est stable et le profil d'éluion suggère qu'elle est homogène et non agrégée.

La construction complète de SmPRMT3 est efficacement séparée de ses contaminants et l'aspect symétrique de son profil d'élution lors de l'étape de filtration sur gel suggère que l'échantillon est homogène. De plus, aucune protéine n'est éluée dans le volume mort (à environ 50 mL sur la figure 44b) ce qui indique que 6his-SmPRMT3-fl n'a pas de tendance à s'agréger. Finalement, l'analyse par SDS-PAGE montre que cette construction est stable. Dans ce chapitre, la dénomination « stable » désignera la résistance de la protéine à la dégradation par des protéases naturellement présentes et actives dans l'échantillon malgré la présence d'antiprotéases.

À l'inverse de la protéine complète, la construction C1 doit être exprimée avec une étiquette 6his-Trx permettant de la solubiliser, sinon la protéine est principalement exprimée dans des corps d'inclusions. Le clivage du tag 6his-Trx pendant la purification entraîne l'agrégation d'une partie de la protéine, comme en témoigne l'élution d'une partie de l'échantillon dans le volume mort lors de l'étape de filtration sur gel, présenté en figure 45b. Lors de cette même étape, la présence de deux pics d'élution contenant C1 indique que l'échantillon est hétérogène.

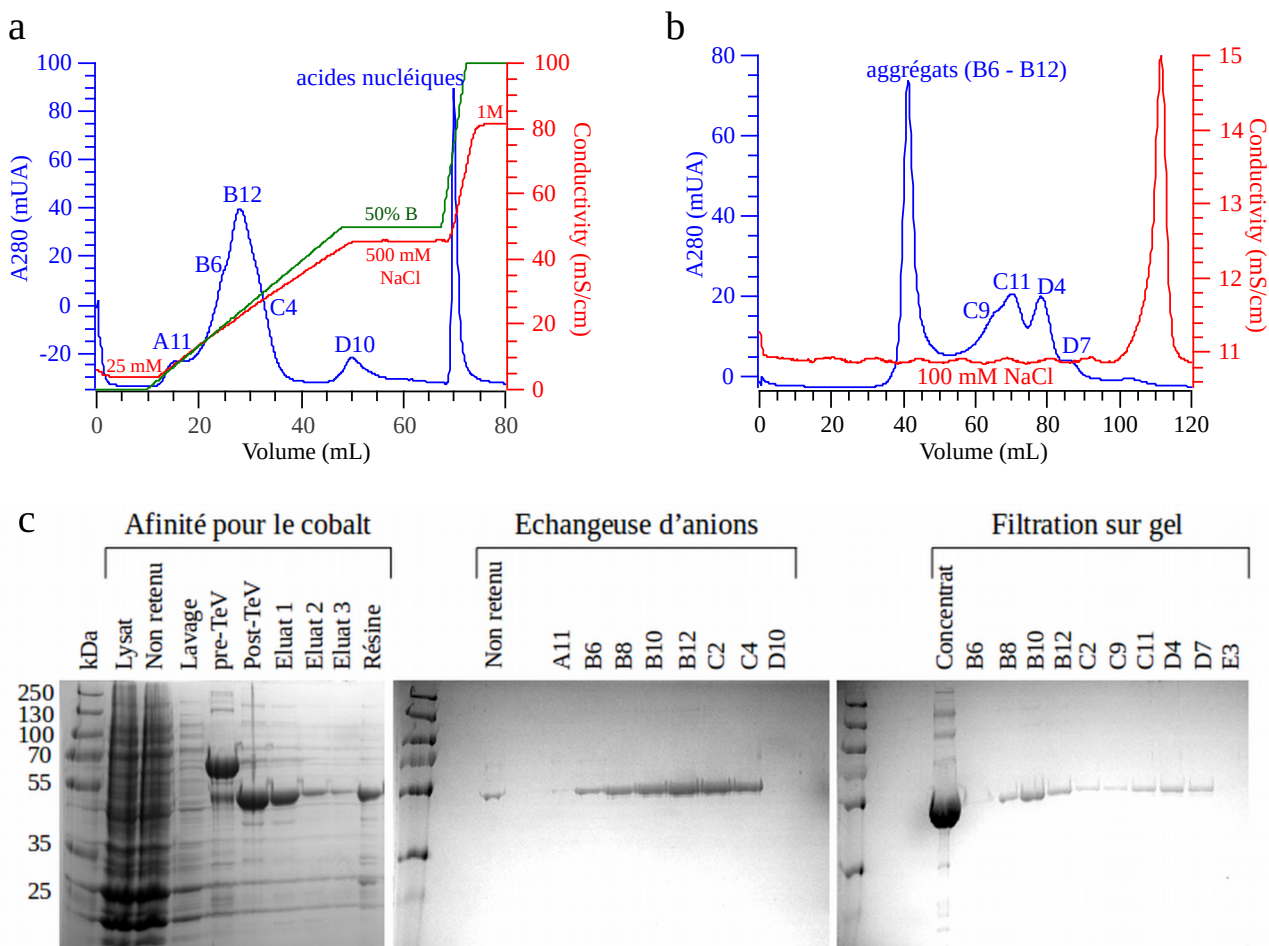


Figure 45 - Purification de SmPRMT3-C1.

(a) Chromatographie échangeuse d'anions HiTrap ANX FF. Les fractions B5 à C5 sont réunies, concentrées puis déposées sur résine de filtration sur gel. (b) Chromatographie par filtration sur gel Superdex 200 (16/60). Les fractions B6 à B12 contiennent la PRMT agrégée. Les fractions C8 à D7 contiennent la forme soluble de SmPRMT3-C1. (c) Analyse SDS-PAGE sur gel de polyacrylamide 12 % des étapes de purification. Coloration au bleu de Coomassie. 6his-SmPRMT3-C1 est efficacement exprimé dans *E. coli*. La construction est stable mais s'agrége et est très hétérogène.

Comme le fragment C1 s'est montré instable et peu soluble, nous nous sommes ensuite concentrés sur les fragments C2 et C3, dont les limites sont proches de celles de la PRMT3 de rat ayant été cristallisée par Zhang et ses collaborateurs. Comme nous ne savions pas comment se comporteraient ces fragments, les purifications ont été initialement réalisées avec des constructions arborant une étiquette 6his-Trx. La figure 46 présente les purifications réalisées pour les fragments C2 et C3. Le chromatogramme de l'étape de filtration sur gel et son analyse par électrophorèse montrent que la construction C2 s'agrège en partie après clivage du tag 6his-Trx, bien qu'en proportion moindre comparée à C1. À l'inverse, la construction 6his-Trx-TEV-C3 clivée ou non-clivée est parfaitement soluble.

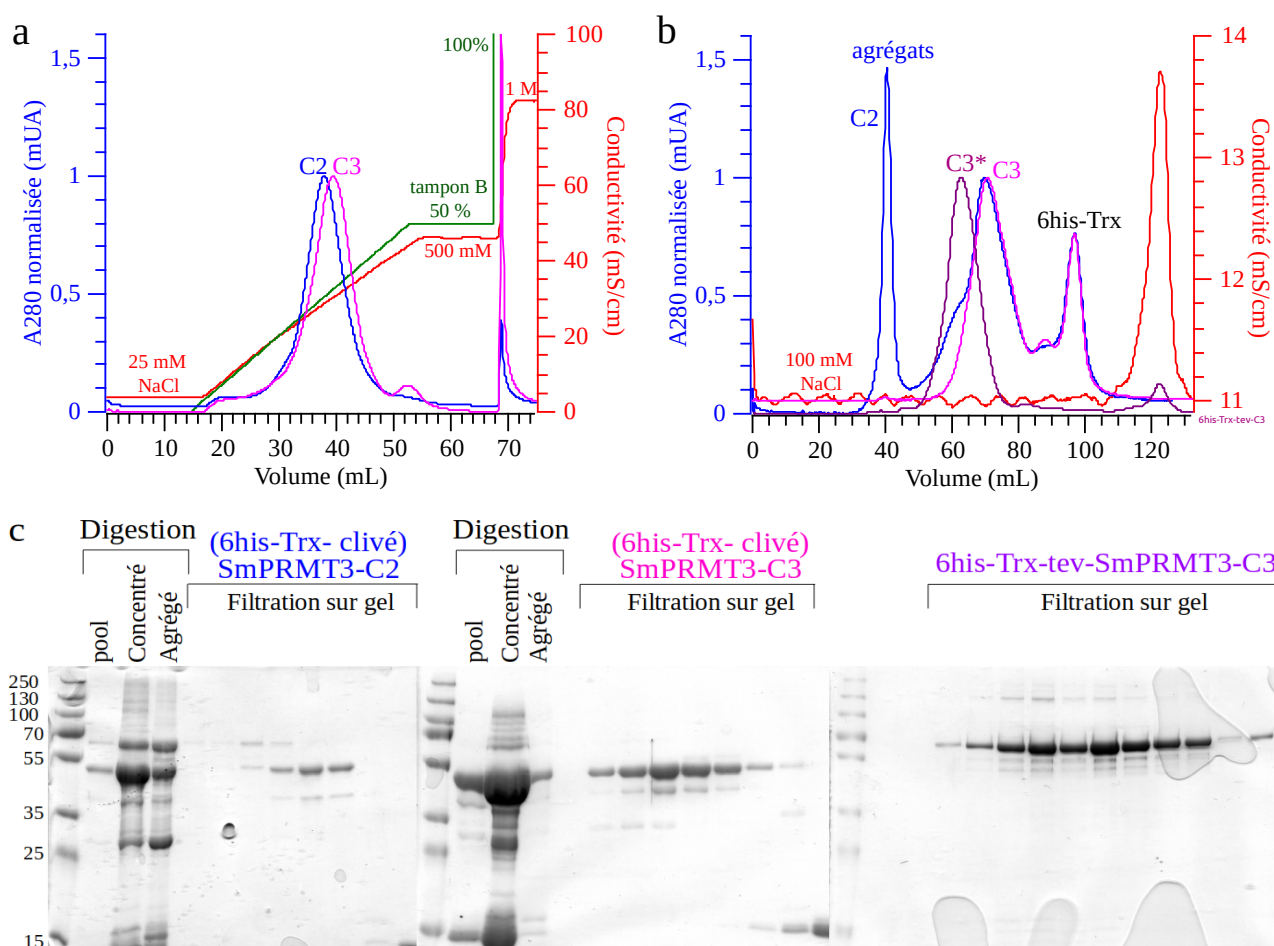


Figure 46 - Purification des constructions SmPRMT3-C2/-C3 et 6his-Trx-TEV-SmPRMT3-C3.

Les absorbances à 280 nm (A280) ont été normalisées afin de pouvoir mieux comparer les chromatogrammes. (a) Chromatographie échangeuse d'anion HiTrap ANX FF. Les profils d'éluion des constructions C2 et C3 sont comparables. (b) Chromatographie par filtration sur gel Superdex 200 (16/60). Les constructions C2 et C3 s'éluent de manière semblable mais C2 possède une propension à l'agrégation contrairement à C3. 6His-Trx-TEV-C3 (C3*) apparaît plus volumineuse du fait de la présence de l'étiquette hexahistidine et du site de clivage. (c) Analyse SDS-PAGE sur gel de polyacrylamide 12 % des étapes de purification. Coloration au bleu de Coomassie. Le clivage de SmPRMT3-C2 n'est pas totalement efficace. Les constructions C2 et C ont toutes deux tendance à se dégrader. C3 est efficacement exprimée dans *E. coli* puis purifié depuis l'extrait brut. La construction est stable et très soluble.

Les constructions 6his-Trx-TEV-C2 et -C3 ont toutefois tendance à se dégrader après coupure du tag 6his-Trx, comme l'illustrent les gels d'électrophorèse présentés en figure 46c. Sans

clivage, 6his-Trx-TEV-SmPRMT3-C3 se dégrade peu, ce qui suggère que le tag apporte un effet protecteur. Suite à ces résultats, le fragment C3 a été sélectionné pour sa qualité de solubilité comparé à C2. Des constructions de 6his-Trx-SmPRMT3-C3 incorporant un site thrombine ou 3C plutôt que TEV ont été clonés puis exprimés afin de vérifier que la dégradation n'est pas due à une digestion aspécifique par la TEV, lorsque celle-ci est utilisée en excès. Aucune de ces constructions n'a cependant permis de limiter la dégradation de SmPRMT3, ce qui suggère que la dégradation a été provoquée par des protéases provenant du système d'expression et actives malgré la présence d'antiprotéases (PMSF, Complete® EDTA-free). Cela a également démontré l'effet protecteur du tag lors de la purification, puisqu'il limite la dégradation tant qu'il n'est pas clivé.

Afin de résoudre les problèmes de dégradation, un criblage de différentes solutions tampons et de pH a été réalisé. La construction 6his-Trx-TEV-SmPRMT3-C3 a été incubée à température ambiante dans différents tampons sur une gamme de pH allant de 5,5 à 9,5, afin d'observer quelles conditions permettraient de limiter la dégradation de la protéine (figure 47).

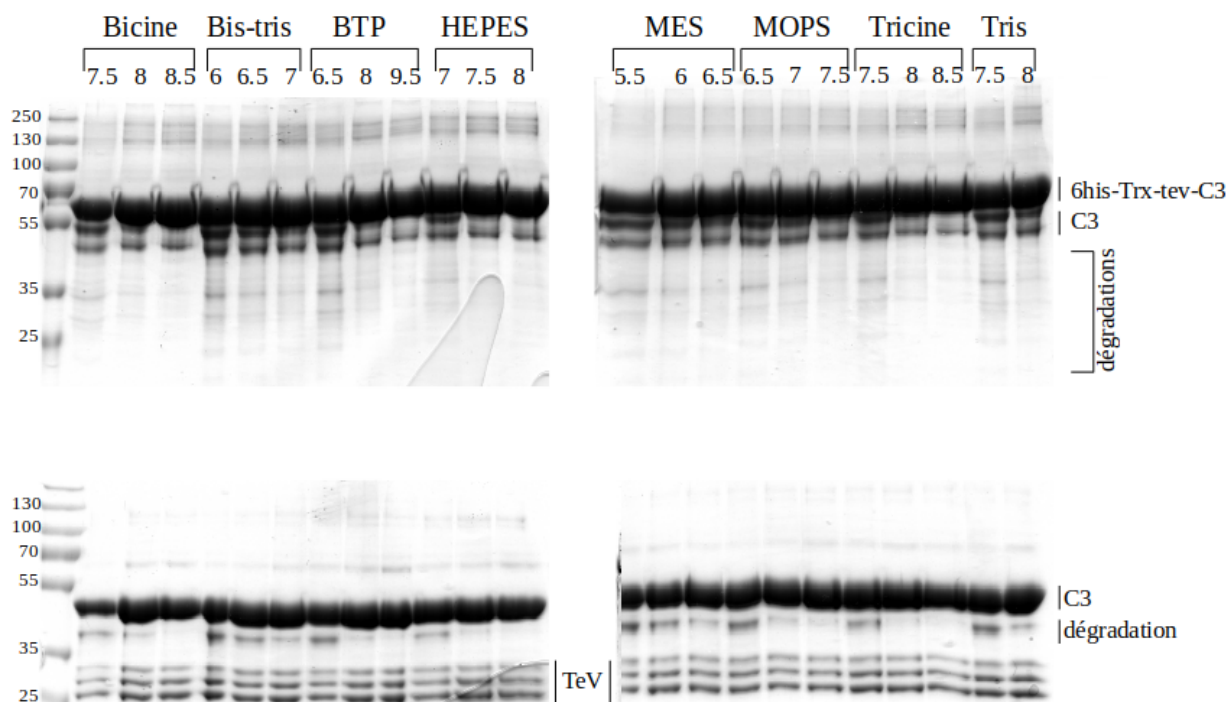


Figure 47 - Optimisation du tampon de purification pour la construction SmPRMT3-C3.

Analyse SDS-PAGE sur gel de polyacrylamide 12 %. 6his-Trx-TEV-SmPRMT3-C3 incubé à température ambiante 12 heures dans différents tampons et à divers pH, avec (en haut) ou sans (en bas) protéase TEV. Moins de bandes de dégradation sont visibles entre pH 8,5 et 9,5, en particulier dans le BTP.

Ce criblage a été réalisé en présence ou en absence de protéase TEV afin de caractériser le ou les tampons qui permettraient de stabiliser l'échantillon avant et après clivage du tag. Les résultats de ce criblage, présentés sur un gel d'électrophorèse en figure 47, montrent que les bandes de dégradation visibles sous la bande principale disparaissent fortement voire totalement pour des pH supérieurs à 8. Dans le Bis-Tris propane (BTP) entre les pH 8 et 9,5, les bandes de dégradations disparaissent à la fois lorsque le tag est laissé ou lorsqu'il est clivé par la TEV. Le BTP à pH 9 a été

par la suite utilisé comme tampon de purification des constructions tronquées de SmPRMT3. Une dégradation résiduelle a cependant persisté après clivage des tags lors d'essais de purifications supplémentaires.

Comme les constructions SmPRMT3-fl et -C3 sont très solubles, le tag 6his-Trx a été remplacé par une étiquette hexahistidine simple (6his), néanmoins nécessaire pour l'étape de purification par IMAC. Nous avons pu constater que ce tag à lui seul stabilisait SmPRMT3-fl et -C3 (données non présentées), ce qui suggère que la dégradation débute par l'extrémité amino-terminale de SmPRMT3 et non par son extrémité carboxy-terminale. De premiers succès en cristallisation avec la construction 6his-TEV-SmPRMT3-C3, qui seront présentés ci-dessous, ont montré que le tag hexahistidine n'empêche pas la formation de cristaux. Les sites de clivages ont donc été retirés de nos constructions afin de limiter la présence de parties non structurellement organisées tout en maintenant l'effet protecteur du tag 6his. La figure 48 montre les purifications des constructions 6his-SmPRMT3-C3 et -C4 (avec tags hexahistidine clivables ou non clivables).

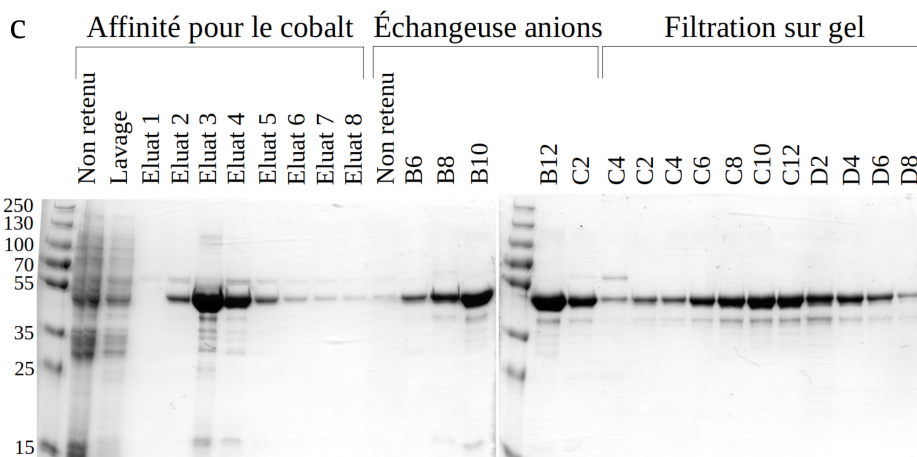
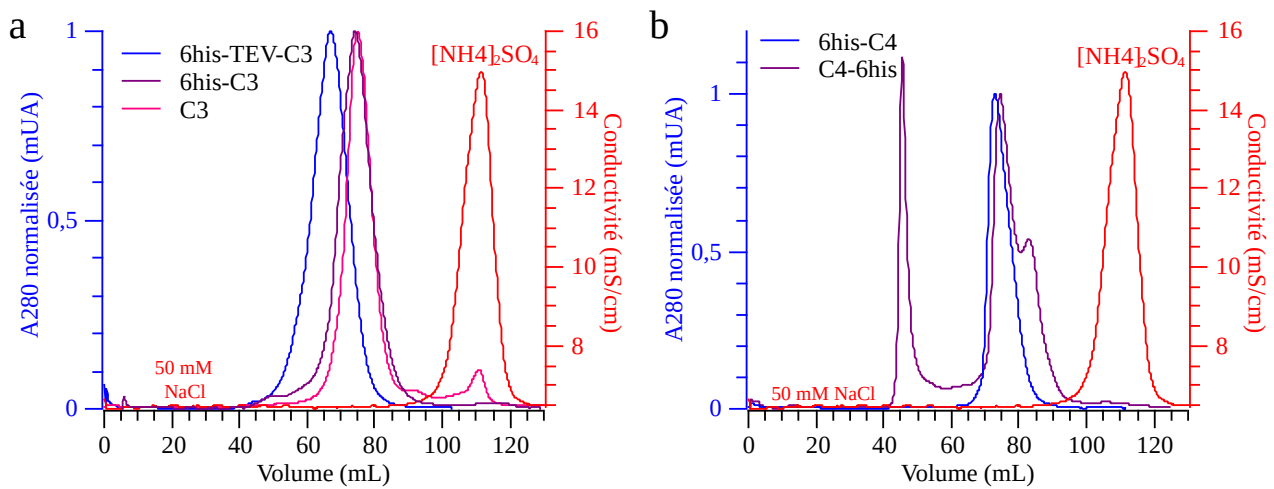


Figure 48 - Purification de 6his-C3 et -C4.

(a) Étape de filtration sur gel pour les constructions de SmPRMT3-C3 étiquetées hexahistidine. La construction 6his-TEV-SmPRMT3 (bleu) paraît plus volumineuse que sa forme clivée (rose) ainsi que la construction sans site de clivage par la protéase TEV (violet). (b) Chromatographie par filtration sur gel Superdex S200 pour les deux constructions de SmPRMT3-C4. L'introduction du tag

hexahistidine à l'extrémité carboxy-terminale (violet) détériore la solubilité de SmPRMT3-C4 comparé au tag amino-terminal (bleu). (c) Analyse par SDS-PAGE sur gel de polyacrylamide 12 % des étapes de purification de 6his-SmPRMT3-C3. Coloration au bleu de Coomassie. 6his-SmPRMT3-C3 est efficacement exprimé dans *E. coli* puis purifié depuis l'extrait brut. La construction est stable et très soluble.

Comme illustré précédemment en figure 43, le fragment C4 commence très près du motif I si bien que le tag pourrait interférer avec le repliement du domaine de fixation du SAM. Une construction du fragment C4 arborant une étiquette 6his carboxy-terminal a donc été produite. Comme le montre les résultats de purification, les constructions 6his-C3 et 6his-C4 sont stables, solubles et l'aspect symétrique de leurs profils d'éluion indiquent qu'elles sont homogènes (figure 48a/b). De plus, la construction 6his-SmPRMT3-C3 dépourvue du site TEV est éluée plus tardivement que la construction avec le site TEV, ce qui indique un réel bénéfice de la suppression du site de clivage sur la compacité de SmPRMT3 (figure 48a). L'introduction du tag hexahistidine à l'extrémité carboxy-terminale pour la seconde construction du fragment C4 induit une agrégation partielle, comme le montre l'éluion d'une partie de l'échantillon dans le volume mort (figure 48b). Cette solution a donc été abandonnée.

Afin de vérifier que SmPRMT3 exprimée dans *E. coli* est fonctionnelle, donc correctement structurée, l'activité de SmPRMT3-fl a été testée. Comme nous ne disposions pas encore de son substrat rpS2, nous avons utilisé rsf1, un substrat identifié au sein de l'équipe (chapitre III) et capable d'être méthylé par plusieurs autres PRMT. Comme le montre la figure 49, SmPRMT3 est active sur rsf1.

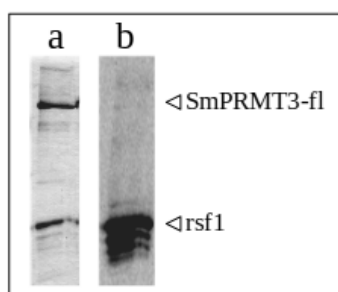


Figure 49 - Méthylation de rsf1 de *S. frugiperda* par SmPRMT3-fl.
 1 μ M de SmPRMT3-fl a été incubé avec 10 μ M de rsf1 en présence de 20 μ M de SAM 14 C, dans 20 mM Tris pH 8, 100 mM NaCl, 0,5 mM DTT, pendant 4 h à 23°C. (a) Séparation par SDS-PAGE Tris-glycine sur gel de polyacrylamide 15% des constituants du milieu réactionnel de la réaction de méthylation de rsf1 par SmPRMT3-fl. (b) Autoradiographie de (a) réalisée avec un film de stockage au phosphore (Fujifilm[®]) pendant 72 heures. Le film a été révélé à l'aide d'un scanner Typhoon (Molecular Dynamics[®]). Bande haute : SmPRMT3 (1-553). Bande basse : rsf1 de *Spodoptera frugiperda* et dégradations de rsf1.

Pour résumer, la protéine SmPRMT3 complète et cinq fragments limités au module catalytique ont été clonés, exprimés et purifiés. Deux types de tag ont été testés : le tag 6his-Trx permettant l'étape d'affinité pour le cobalt et améliorant la solubilité de la protéine exprimée, et le tag hexahistidine simple nécessaire à l'étape d'affinité. Les tags étaient initialement clivables par les protéases TEV, 3C ainsi que la thrombine mais l'expérience a montré que ces tags améliorent la résistance de SmPRMT3 aux protéases provenant du système d'expression et actives malgré l'ajout d'antiprotéases. Les sites de clivage ont été par la suite supprimés des constructions. Parmi les constructions de SmPRMT3 testées, les fragments C3 et C4 et la protéine complète ont démontré un bon comportement lors de la purification, tandis que les fragments C1 et C2 sont sujets à l'agrégation. Les travaux sur le fragment C5 n'ont pas été poussés, car ce fragment débute après le motif I donc présente moins d'intérêt dans le cadre d'études structurales. Finalement, le test d'activité réalisé sur le substrat rsf1 a montré que SmPRMT3 peut être produite sous forme fonctionnelle dans *E. coli*.

c - Caractérisation biophysique de SmPRMT3-fl et SmPRMT3-C3

Les deux principales constructions englobant la SmPRMT3 complète (SmPRMT3-fl/1-553) et le module catalytique (SmPRMT3-C3/184-553) ont été caractérisées par le biais de plusieurs approches expérimentales biophysiques complémentaires : la diffusion dynamique de la lumière (DLS) pour mesurer l'homogénéité des échantillons, la diffusion des rayons X aux petits angles (SAXS) pour mesurer l'homogénéité et étudier le comportement en solution, la spectrométrie de masse (MS) pour contrôler l'intégrité des protéines et mesurer leur état oligomérique, et la fluorimétrie par balayage différentiel (DSF) pour mesurer leur stabilité. La figure 50 présente les résultats de DLS obtenus avec cinq réplicats techniques, pour les constructions 6his-SmPRMT3-fl et -C3. Les données sont résumées dans le tableau 6 et ont été calculées à partir des courbes d'auto-corrélation du signal de DLS (figure 50).

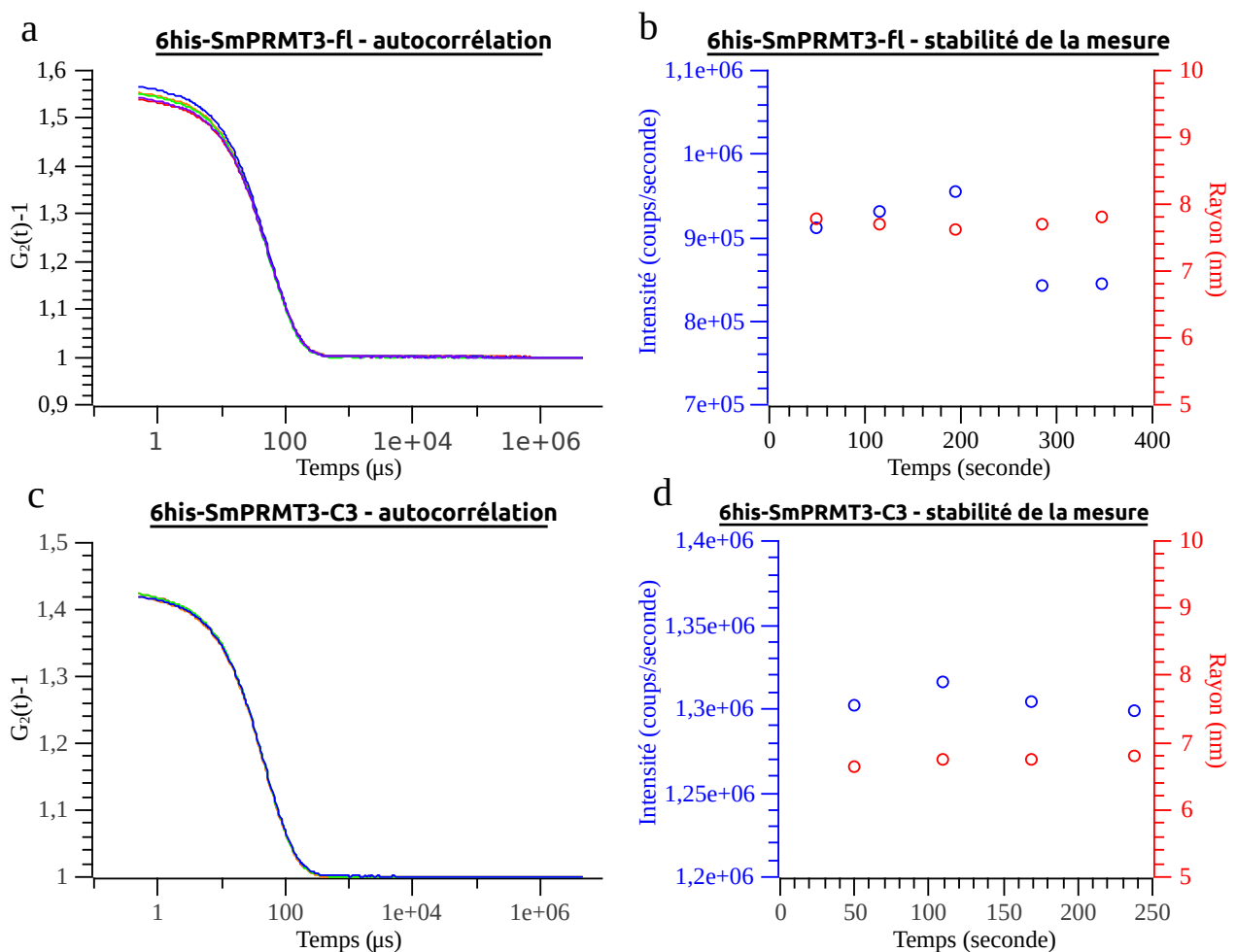


Figure 50 - Caractérisation de SmPRMT3-fl et -C3 : diffusion dynamique de la lumière.

(a/c) Courbes d'auto-corrélation respectivement pour les constructions 6his-SmPRMT3-fl et -C3. (b/d) intensité du signal et rayons hydrodynamiques évalués sur 4 à 5 réplicats techniques respectivement pour les constructions 6his-SmPRMT3-fl et -C3. Les rayons hydrodynamiques estimés sont constants.

La superposition des courbes d'auto-corrélation et les graphiques présentant les variations d'intensité du signal de DLS et des rayons hydrodynamiques calculés, montrent que le signal est

stable durant les cinq mesures. Les échantillons de SmPRMT3-fl et SmPRMT3-C3 produisent un signal monomodal (ligne de base comprise entre 0,997 et 1,002) et les échantillons sont faiblement polydisperses (pourcentage de polydispersité respectivement de 15 % et 17,5 %). Les rayons hydrodynamiques calculés (respectivement 7,7 et 6,8 nm), ont permis d'estimer des masses théoriques respectives d'environ 300 et 400 kDa. Les dimères de SmPRMT3-fl et -C3 ont une masse moléculaire d'environ 85 et 127 kDa ce qui implique que ces protéines formeraient des assemblages d'environ trois dimères. Néanmoins, les masses déterminées par DLS sont déterminées à partir de la masse volumique moyenne d'une protéine extrapolée à un modèle sphérique, et sont donc très biaisées.

Construction	Intensité (coups.s ⁻¹)	Coefficient de diffusion (cm ² .s ⁻¹)	Rayon hydrodynamique (nm)	Polydispersité (%)	Poids moléculaire (kDa)	Ligne de base	SOS
Fl (1-553)	~ 9.10 ⁵	2,725.10 ⁻⁷ ± 0,025.10 ⁻⁷	7,73 ± 0,07	14,99 ± 1,45	~ 400	1,00095 ± 0,00058	0,769 ± 0,192
C3 (184-553)	~ 1.10 ⁶	3,134.10 ⁻⁷ ± 0,045.10 ⁻⁷	6,79 ± 0,10	17,49 ± 1,18	~ 300	1,00035 ± 0,00054	0,313 ± 0,137

Tableau 6 - Caractérisation de SmPRMT3-fl et -C3 : diffusion dynamique de la lumière.

Ce tableau résume les valeurs moyennes obtenues à partir des réplicats de mesures DLS présentés en figure 50.

Pour caractériser le comportement en solution de SmPRMT3, des expériences de diffusion des rayons X aux petits angles (SAXS) ont été réalisées. Les résultats sont présentés en figure 51. Le tableau 7 résume les valeurs mesurées et calculées à partir des données de SAXS. Les courbes de Guinier obtenues (figure 51c/d) sur ces deux constructions sont linéaires, ce qui traduit l'homogénéité des échantillons. L'estimation par SAXS de la plus grande dimension des particules observées, permettant de tracer les fonctions de distribution des distances (51e/f), mesure 89 Å pour les deux constructions de SmPRMT3. Cette valeur est en adéquation avec le dimère cristallographique de PRMT3 de rat, dont la plus grande dimension est d'environ 9 nm. Les rayons de giration ont été estimés à 34 Å et 33 Å pour respectivement SmPRMT3-fl et C3. Ces rayons de giration sont du même ordre de grandeur que celui calculé à l'aide de crysol (suite ATSAS) sur le dimère cristallographique de la PRMT3 humaine (1F3L), mesurant environ 28 Å. Par comparaison avec un standard de BSA, la masse moléculaire en solution de SmPRMT3-C3 a été estimée à 87 kDa, ce qui suggère que SmPRMT3-C3 se comporte comme un dimère en solution puisque la masse moléculaire d'un dimère de cette construction est de 84 kDa environ.

Construction	I ₀	R _g (Å)	Distance max (Å)	Poids moléculaire
Fl (1-553)	2,318 ± 0,026	32,6 ± 0,5	89	Non calculée
C3 (184-553)	0,452 ± 0,007	33,6 ± 0,6	89	~ 87 kDa

Tableau 7 - Caractérisation de SmPRMT3-fl et -C3 : Diffusion des rayons-X aux petits angles.

Les I₀ sont extrapolés depuis la figure 51. Étonnamment, le R_g de SmPRMT3-fl est inférieur au R_g de SmPRMT3-C3. Le poids moléculaire (~ 87 kDa) de C3 est compatible avec l'existence d'un dimère en solution (~2x43 kDa). Le poids moléculaire de la construction SmPRMT3-fl n'a pas pu être calculé car l'étalon de BSA s'est agrégé lors de plusieurs mesures consécutives.

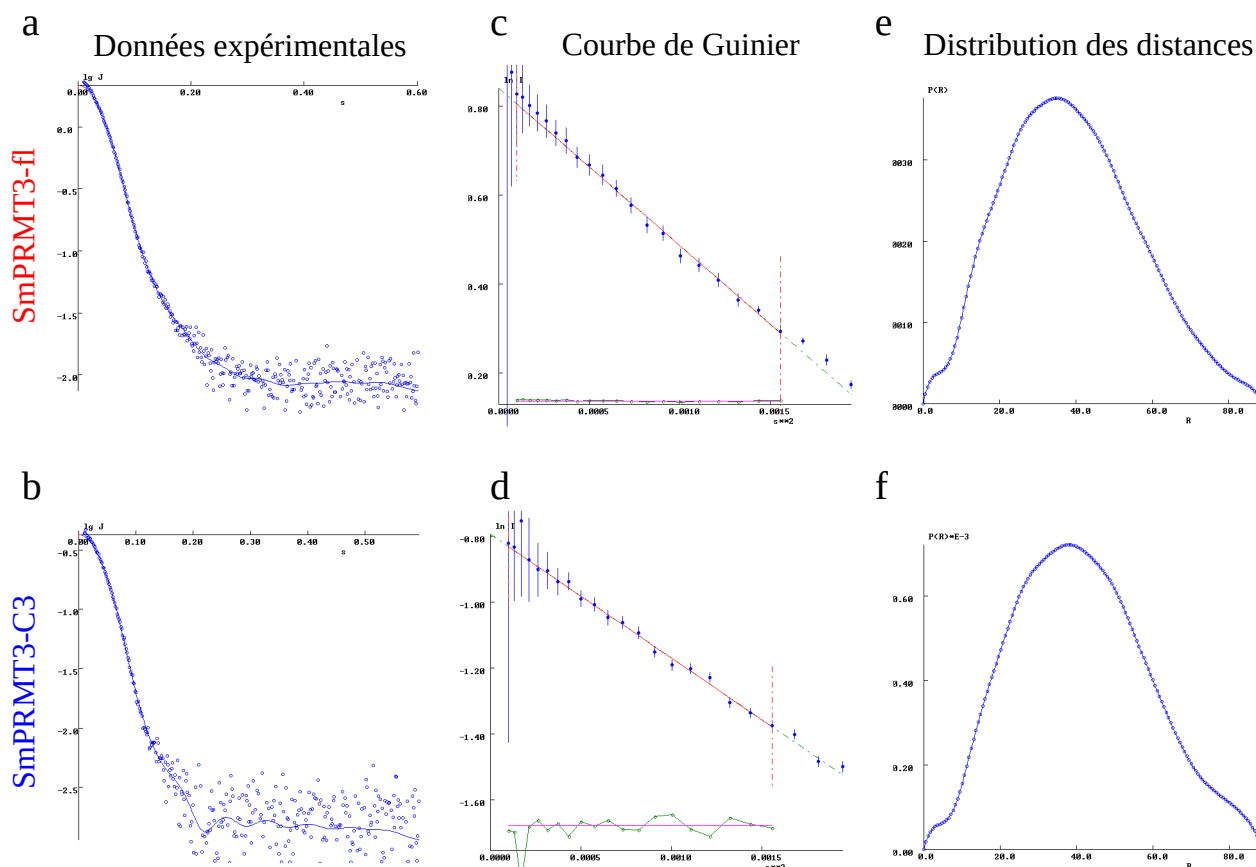


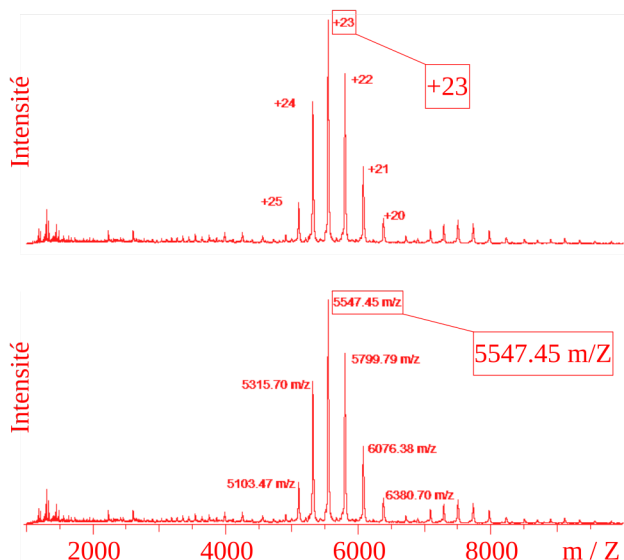
Figure 51 - Caractérisation de SmPRMT3-fl et -C3 : diffusion des rayons-X aux petits angles.

Les trois séries de graphiques présentent les valeurs d'intensité de diffusion des rayons X (a/b), les courbes de Guinier, permettant l'évaluation du rayon de giration R_g (c/d), et les fonctions de distribution des distances (e/f) pour les constructions SmPRMT3-fl (en rouge) et SmPRMT3-C3 (en bleu).

Des mesures de spectrométrie de masse (MS) ont également été réalisées afin de vérifier l'intégrité des constructions et de confirmer l'état oligomérique (figure 52). Le tableau 8 résume les résultats obtenus en spectrométrie de masse. Une masse moléculaire de 42268,6 kDa a été mesurée pour le monomère de SmPRMT3-C3, ce qui est conforme à la masse attendue. Cela confirme que la protéine n'est pas dégradée lors de la purification, et qu'elle n'a subi aucune modification post-traductionnelle artefactuelle provenant de *E. coli*, ni ne présente de sites d'auto-méthylation, comme c'est parfois observé pour d'autres PRMT (Bonfond et al., 2015).

L'existence de dimères en solution a été confirmée par spectrométrie de masse en conditions natives sur SmPRMT3-fl. En effet, la masse moléculaire mesurée est de 127591,4 kDa (figure 52), ce qui correspond à un dimère de SmPRMT3-fl fixant deux atomes de zincs (masse théorique attendue de 127458,1 kDa). Les atomes de zinc sont fixés par le motif de type doigt de zinc C2H2 sur le domaine amino-terminal des PRMT3, ce qui suggère que ce domaine est correctement replié lorsque SmPRMT3-fl est exprimée dans *E. coli*. À noter que le différentiel de masse observé en faveur de la masse mesurée, d'environ 133 Da, pourrait correspondre à des modifications post-traductionnelles. La phosphorylation des domaines amino-terminaux de PRMT3 a notamment été décrite (Handrkova et al. 2011) et est nécessaire pour sa fixation à son substrat.

a SmPRMT3-fl - conditions non dénaturantes



b SmPRMT3-C3 - conditions dénaturantes

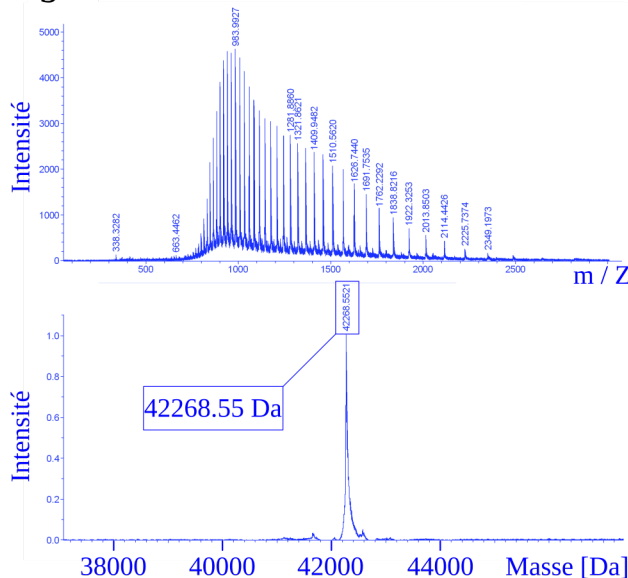


Figure 52 - Caractérisation de SmPRMT3-fl / C3 : spectrométrie de masse.

(a) Analyse en conditions natives de SmPRMT3-fl. Le produit des charges (en haut) par les rapports masse/charge (en bas) indique un poids moléculaire de 127591,4 Da. (b) Analyse en conditions dénaturantes de SmPRMT3-C3. Le poids moléculaire mesuré (en bas) depuis le spectre Intensité = f(masse/charge) (en bas) est de 42268,6 Da.

Construction	Masse mesurée (Da)	Masse attendue (Da)	Différentiel de masse (Da)
SmPRMT3-Fl (1-553)	127591,4 ± 13	127458,1	+ 133
SmPRMT3-C3 (184-553)	42268,6	42268,7	- 0,1

Tableau 8 - Caractérisation de SmPRMT3-fl / C3 : spectrométrie de masse.

Ce tableau récapitule les résultats de spectrométrie de masse : la masse mesurée pour SmPRMT3-C3 correspond à la masse attendue, confirmant l'intégrité de la PRMT à l'issue de la purification. La masse calculée pour SmPRMT3-fl confirme que la PRMT est un dimère en solution.

Finalement, la stabilité thermique du module catalytique de SmPRMT3-C3 a été mesurée par DSF. En DSF, on suit la dénaturation d'une protéine en mesurant à température croissante sa fluorescence à 330 nm et 350 nm, les longueurs d'ondes d'émission des résidus tyrosine et tryptophane respectivement en milieu hydrophobe et en milieu hydraté. Comme le montre la figure 53, la forme apo de SmPRMT3 possède un T_m de 41,5 °C contre respectivement 49,7 °C et 52,1 °C pour les formes de SmPRMT3 en complexe avec le cofacteur SAH ou l'inhibiteur naturel des méthyltransférases SFG, soit un effet stabilisateur jusqu'à +10 °C environ. Chez la plupart des PRMT, ces deux ligands sont ajoutés dans l'échantillon en amont de la cristallisation lorsque aucun autre ligand de travail n'est utilisé. Le SAH ou la SFG sont généralement nécessaire au repliement de l'hélice α -X sur le site catalytique et à la stabilisation de la structure. Le SAH est d'ailleurs intrinsèquement présent dans les structures cristallographiques de CARM1 même lorsque aucun ligand de travail n'est ajouté. Les ligands SAH et SFG ont donc été utilisés afin de stabiliser les constructions de SmPRMT3 en vue des essais de cristallisation.

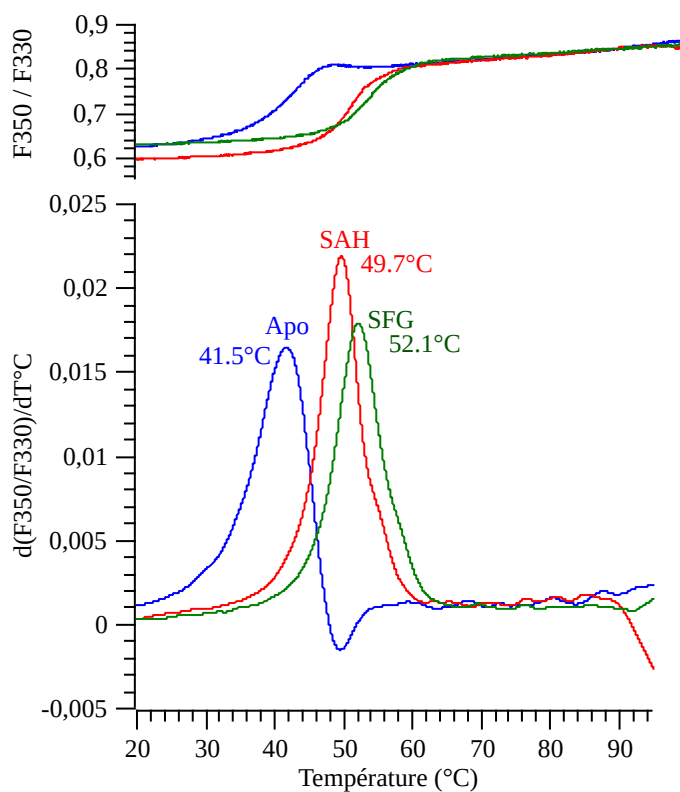


Figure 53 - Caractérisation de SmPRMT3-C3 : stabilité thermique.

En haut : Rapport de la fluorescence à 350 nm (résidus aromatiques exposés) et à 330 nm (résidus aromatiques enfouis) en fonction de la température, pour la construction SmPRMT3-C3 sous forme apo (bleu) ou en présence de SAH (rouge) ou de SFG (vert). En bas : Dérivée première des courbes (haut), permettant de déterminer la température de fusion T_m . La présence de ces deux ligands stabilise le module catalytique de SmPRMT3.

Pour résumer, la caractérisation biophysique de SmPRMT3 complète et du fragment C3 montre que les deux échantillons ne sont pas dégradés et sont homogènes. Le SAXS ainsi que la spectrométrie de masse indiquent également que la forme entière et le fragment C3 de SmPRMT3 se comportent comme des dimères en solution. Aucune forme oligomérique de plus haut poids moléculaire n'a été détectée. La spectrométrie de masse a également montré que le domaine amino-terminal de SmPRMT3 fixe un ion zinc, probablement grâce au motif C2H2 spécifique des PRMT3, ce qui suggérerait que ce domaine est correctement replié. Finalement, les mesures de stabilité thermique par DSF ont montré que la fixation d'un ligand dans le site actif stabilise fortement la protéine, ce qui est également observé pour les autres PRMT et ces ligands seront ajoutés lors des essais de cristallisation.

d - Cristallisation de SmPRMT3-fl et SmPRMT3-C3

Des essais de cristallisation sur des cribles commerciaux ont été réalisés pour la PRMT3 complète et les fragments C1, C2 et C3, mais aucun cristal n'a été obtenu à partir de ces cribles. Comme indiqué dans la partie purification, les constructions C1 et C2 ont tendance à s'agréger, et les essais réalisés sur ces deux fragments n'ont été effectués qu'à partir de la fraction encore soluble, si bien que peu d'échantillon était disponible. La majorité des essais se sont donc concentrés sur la PRMT3 entière et le fragment C3. Le détail des conditions testées est résumé dans le tableau 9.

Un criblage systématique de 96 conditions PEG-pH mis au point au laboratoire a été testé sur les constructions étiquetées histidine de SmPRMT3-C3, ce qui a permis d'obtenir des cristaux dans du PEG20000 avec du BTP pH 9 (figure 54b, en haut). Il est intéressant de noter que le pH de cristallisation correspond à celui déterminé lors du criblage de tampons réalisé pour optimiser la purification de SmPRMT3-C3. Ce crible a été testé sur les constructions de SmPRMT3-fl et a

permis d'obtenir de premiers cristaux dans du PEG20000 avec de l'HEPES pH 7,5 et pH 8, et 100 mM NaCl. Les cristaux obtenus par ce crible ont pu être reproduits manuellement dans des plaques de cristallisation VDX à 24 puits.

Construction	Échantillon			T°C	Crible	Cristal ?
	Protéine (mg.mL ⁻¹)	ligand	protéase			
fl (1-553)	6	SAH & SFG	-	20	Qiagen JCSG+ Qiagen PEG suite I Hampton PACT PEG-pH_Screen_I	∅ ∅ ∅ ✓
6his-fl (1-553)	6	SAH & SFG	-	20	Mol. Dim. BCS Mol. Dim. PACT PEG-pH_Screen_I	∅ ∅ ✓
6his-Trx-TEV / -thr-C1 (141-553)	3	SAH	TEV & thr	20	Qiagen JCSG+ Qiagen PEG suite I	∅ ∅
C2 (171-553)	6	SAH	-	20	Qiagen JCSG+ Qiagen PEG suite I	∅ ∅
6his-Trx-TEV-C3 (184-553)	2	SAH	TEV	20	Qiagen JCSG+ Qiagen PEG suite I Qagen AmSO4 suite	∅ ∅ ∅
SmPRMT3-C3 (184-553)	2	SAH	-	20	Qiagen JCSG+ Qiagen PEG suite I Qiagen AmSO4 suite	∅ ∅ ∅
6his-TEV-C3 (184-553)	5 & 10	SAH	-	20	Qiagen JCSG+ Qiagen PEG suite I Qiagen AmSO4 suite PEG-pH_Screen_I	∅ ∅ ∅ ✓
6his-C3 (184-553)	6 ± ensemencement	SFG	-	20	Qiagen JCSG+ Qiagen PEG suite I Qiagen Anions suite Qiagen Cations suite Hampton Index	✓ ∅ ∅ ∅ ∅
6his-C4 (199-553)	7,5 & 6	SFG	-	20	PEG-Malo_Screen_I PEG-Malo_Screen_II	∅ ∅

Tableau 9 - Résumé des conditions de cristallisation testés pour SmPRMT3.

(Colonne 1) Nom de la construction testées. Lorsque aucune étiquette n'est indiqué, c'est que celui-ci a été clivé lors de la purification. (Colonne 2) Composition de l'échantillon : concentration en protéine, ligand ajouté à 1 ou 2 mM, protéase ajoutée dans la goutte de cristallisation. (Colonne 3) Température utilisée lors des essais de cristallisation. (Colonne 4) Cribles commerciaux ou mis au point au laboratoire. (Colonne 5) ∅ pour aucun cristal obtenu, ✓ pour au moins un cristal obtenu.

Comme illustré sur la figure 54a, les cristaux de SmPRMT3-fl forment des disques à un pH proche de 7,5 et un mélange disques/baguettes pour un pH proche de 8. Ce changement de morphologie est également observé à pH 8 pour des concentrations croissantes en NaCl. Ces cristaux ne montrent malheureusement pas de qualité de diffraction des rayons X.

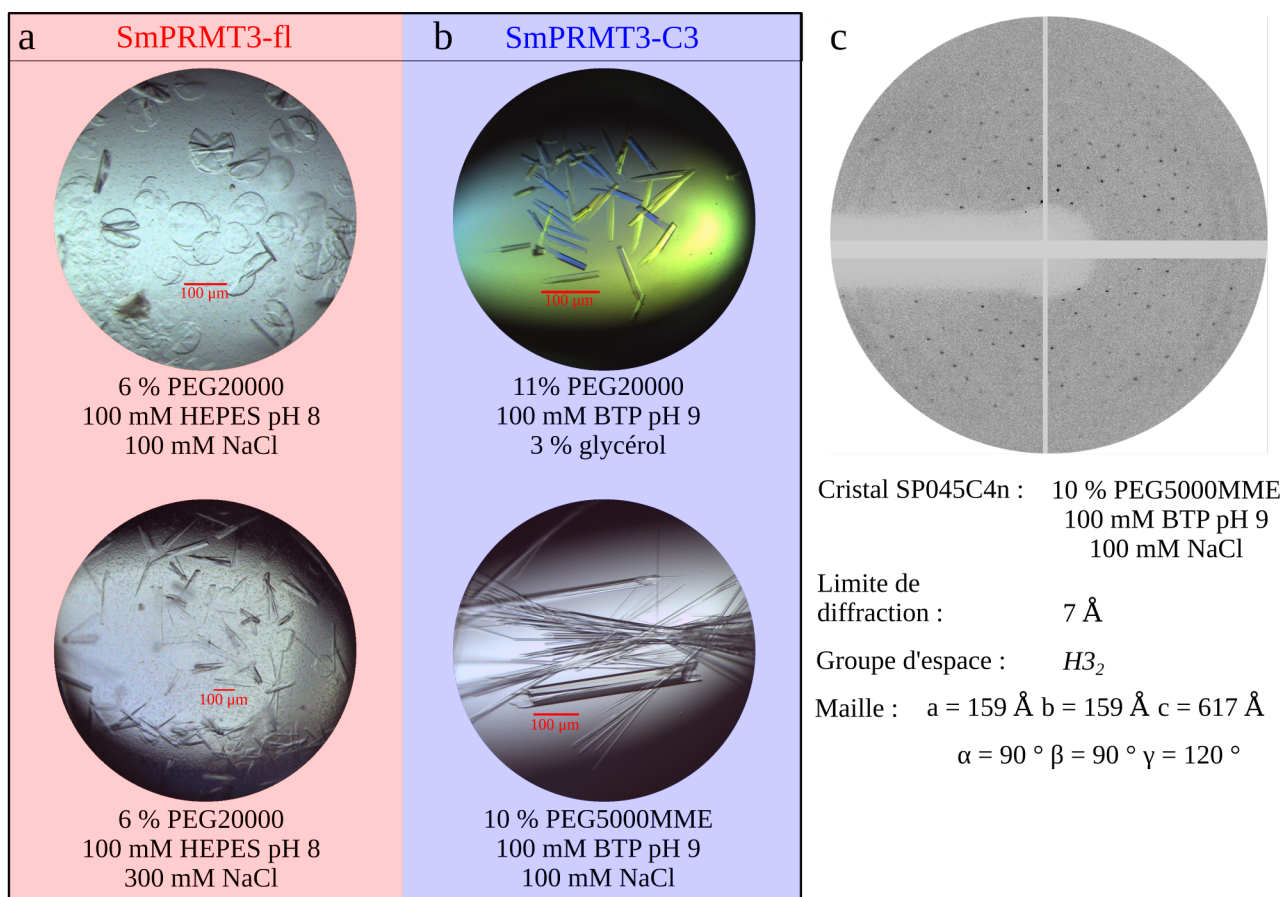


Figure 54 - Cristallisation de SmPRMT3-fl et -C3.

(a) Cristaux de SmPRMT3-fl obtenus à 20°C. (b) Cristaux de SmPRMT3-C3 (haut) et 6his-SmPRMT3-C3 (bas) obtenus à 20°C. (c) Cliché de diffraction obtenu au synchrotron Soleil (ligne Proxima 1) à partir du cristal (b)-bas et paramètres de maille associés.

Les cristaux de SmPRMT3-C3 forment des baguettes fines et irrégulières apparaissant en seulement trente minutes après formation de la goutte de cristallisation, et ne diffractent pas. À 4 °C, la croissance de ces aiguilles n'est pas ralentie mais s'accompagne d'une augmentation de la nucléation et de la formation d'une multitude de microcristaux. De nombreux essais en plaque de 24 puits incluant diverses constructions, agents cristallins, pH, molécules tampons et sels ont été réalisés jusqu'à obtenir les conditions suivantes : PEG5000MME ou PEG6000 concentrés entre 7 % et 12 %, avec 100 mM de BTP pH 9 et 100 mM de NaCl. Les cristaux qui se développent dans ces conditions, présentés en figure 54b (en bas), sont plus grands et croissent moins vite, mais sont obtenus en petite quantité. En effet, le rendement n'est que d'un à deux cristaux exploitables par plaque de 24 puits, contre une grande proportion d'aiguilles très fines inutilisables.

De plus, une seconde phase de nucléation a régulièrement lieu induisant le développement d'aiguilles à la surface des cristaux déjà formés. Les constructions C1 et C2, moins solubles lors de la purification, n'ont pas été testées dans ces dernières conditions tandis que la construction C4 a donné lieu à la formation de très petites aiguilles.

Quatre-vingt-seize différents additifs commercialisés par la société Hampton® Research ont été indépendamment ajoutés au précipitant, mais aucun n'a permis d'améliorer significativement la cristallisation du fragment C3. Pour tenter de trouver des conditions différentes de celles déjà établies et qui permettraient peut-être d'obtenir des cristaux exploitables, des essais de cristallisation de 6his-SmPRMT3-C3 par ensemencement ont été réalisés sur les cribles commerciaux Qiagen® JCSG+, PEG suite I, Anions, cations et Hampton® Research Index, ainsi que dans les conditions citées au-dessus en présence des additifs Hampton® Research. Aucune amélioration significative de l'aspect et de l'homogénéité des cristaux n'a été obtenue avec les additifs. Des petites aiguilles se sont développées dans les cribles JCSG+ et Index, mais dans des conditions proches de celles déjà déterminées, c'est-à-dire dans des PEG à pH 9.

Le dernier jeu de données, collectés en octobre 2017 à l'ESRF, n'a permis d'atteindre qu'une résolution de 7Å environ, ne permettant pas de résoudre la structure de cette PRMT. Les cristaux appartiennent au groupe d'espace trigonal $H3_2$ et ont pour paramètres de maille $a = 159 \text{ \AA}$ $b = 159 \text{ \AA}$ $c = 617 \text{ \AA}$ $\alpha = 90^\circ$ $\beta = 90^\circ$ $\gamma = 120^\circ$, comme illustré en figure 54c (statistiques données à l'annexe 3). Ces paramètres sont compatibles avec un contenu de six à huit monomères par unité asymétrique soit un coefficient de Matthews compris entre $2,90 \text{ \AA}^3 \cdot \text{Da}^{-1}$ et $2,18 \text{ \AA}^3 \cdot \text{Da}^{-1}$.

Au total pour le fragment C3, 60 boîtes de cristallisation ont été réalisées manuellement et 50 boîtes ont été produites à l'aide d'un robot de cristallisation (Mosquito, TPP Labtech), soit plus de 6200 gouttes. Quelques pistes demeurent à explorer afin d'améliorer la qualité de diffraction des cristaux de SmPRMT3. En premier lieu, il serait nécessaire de tester une construction du fragment C2 dans les dernières conditions de cristallisation. En effet, bien que celui-ci soit moins soluble et homogène que le fragment C3, le fragment C2 est le plus proche structuralement des fragments de PRMT3 de mammifère dont les structures cristallographiques ont été publiées. Le segment supplémentaire de C2 comparé à C3 pourrait amener les contacts manquants pour permettre un empilement cristallin de meilleure qualité.

Plusieurs pistes pourraient être explorées pour améliorer les conditions de cristallisation actuelles ou trouver de nouvelles conditions de cristallisation. La déshydratation des cristaux de SmPRMT3 pourrait être envisagée. Cette méthode consiste à diminuer la teneur en solvant dans des cristaux protéiques, ce qui peut amener à des réarrangements de la structure du cristal allant d'une amélioration de l'empilement cristallin jusqu'à un changement complet de groupe d'espace (Klingl et al., 2015 ; Krauss et al., 2012). Cette méthode nécessite cependant beaucoup de cristaux afin d'ajuster le degré de déshydratation, et il faudra trouver un moyen de rendre la cristallisation de SmPRMT3-C3 plus reproductible afin d'atteindre cet objectif. La méthylation des lysines devra également être essayée. Cette méthode permet de modifier les propriétés physico-chimiques des protéines et a dans certains cas permis d'améliorer, voire de rendre possible la cristallisation de certaines protéines récalcitrantes (Walter et al., 2006).

Finalement, nous pourrions utiliser des cristallophores (Engilberge et al. 2017). Cette famille de complexes des lanthanides, récemment développée dans le cadre d'un partenariat entre l'IBS à Grenoble et l'ENS de Lyon, possède de fortes propriétés de nucléation grâce à leur capacité à se positionner aux interfaces protéine-protéine dans les cristaux.

2 - Etude de SmPRMT3 par cryo-microscopie électronique

Au regard des difficultés à obtenir des cristaux de qualité suffisante pour résoudre une structure cristallographique, nous avons tenté une autre approche utilisant la cryo-microscopie électronique (cryo-EM) sur particules individuelles. Malgré les innovations technologiques apportées à cette technique ces dernières années, la microscopie électronique s'applique en général à des objets nettement plus gros. Avec les PRMT, nous travaillons à proximité de la limite basse de cette technologie. Néanmoins, les développeurs du plateau de phase volta ont récemment publié la structure de l'hémoglobine humaine résolue par cryo-microscopie électronique à 3,2 Å (Khoshouei et al. 2017). Cette protéine forme un tétramère de 64 kDa et est fortement symétrique. SmPRMT3-fl forme un dimère d'environ 130 kDa adoptant théoriquement une symétrie C2, mais la symétrie peut être brisée par la possible flexibilité des domaines amino-terminaux.

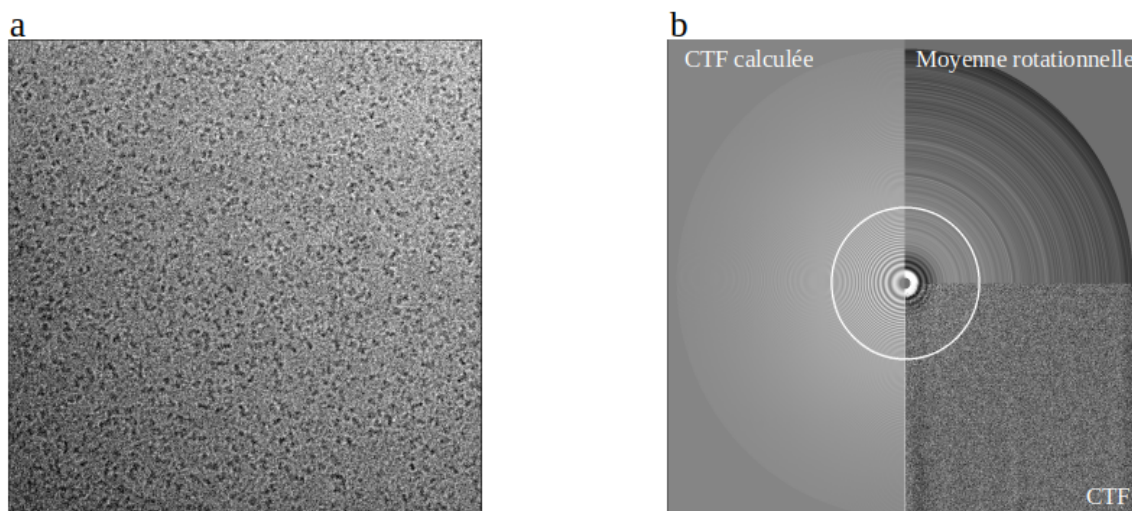


Figure 55 - Observation de 6his-SmPRMT3-fl en cryo-EM en transmission.

(a) Exemple d'une micrographie électronique de l'échantillon de 6his-SmPRMT3-fl à 7 μM , enregistrée à 100 keV et à une défocalisation de 3,8 μm . (b) Calcul des paramètres de la CTF. En bas à droite, spectre de puissance de la micrographie ci-contre. En haut à droite : moyenne rotationnelle du spectre de puissance. À gauche, le modèle de CTF (Contrast Transfer Function) calculé pour cette image avec Gctf, permettant la détermination des paramètres de la CTF.

Un premier échantillon de 6his-SmPRMT3-fl a été préparé sur une grille de type Quantifoil® Cuivre/Rhodium (Cu/Rh) R22 couverte d'un film de carbone perforé tel que décrit dans la section « matériel et méthodes ». La grille a été observée à un grossissement de 93000, une défocalisation de 2 μm à 3,8 μm sur un microscope électronique en transmission FEI Tecnai Polara F30 utilisé à 100 keV et équipé d'une caméra Falcon I (FEI®). La figure 55 présente un exemple de l'échantillon observé. 6his-SmPRMT3-fl se répartie de façon homogène dans le tampon et on n'observe pas

d'agrégation. La concentration (7 μM) semble cependant trop élevée puisque les particules sont presque toutes en contact (figure 55a). De plus, malgré la haute concentration en particules, une accélération des électrons à 100 keV et une défocalisation importante, seuls trois anneaux de Thon sont distinctement visibles sur les spectres de puissance des micrographies. Le nombre et l'intensité des anneaux de Thon renseignent sur le rapport signal/bruit et sur la résolution au-delà de laquelle l'information ne pourra plus être extraite des micrographies. Dans le cas présenté, le faible nombre d'anneaux de Thon visibles indique que la glace est trop épaisse, ce qui ajoute du bruit au signal diffusé par les particules.

De nouveaux échantillons ont été préparés pour une collecte de données. La concentration de l'échantillon a été abaissée à 0,7 μM afin d'éviter que les particules ne se superposent ou ne se touchent dans le tampon. L'acquisition des données a été réalisée à un grossissement de 100000, une défocalisation de 0,3 μm à 0,5 μm sur un microscope électronique en transmission FEI Titan Krios utilisé à 300 keV et équipé d'un plateau de phase Volta, d'un système de correction des aberrations sphériques et d'une caméra K2 (Gatan[®]). 2855 micrographies ont été collectées puis pré-traitées de la manière suivante : les micrographies ont été corrigées du gain de la caméra K2 avec la commande CLIP du programme IMOD (Kremer et al., 1996), les mouvements de particules et dommages radiatifs ont été corrigés avec le programme motioncorr2 (Zheng et al., 2017) et les paramètres de la fonction de transfert de contraste (CTF) ont été évalués avec le programme Gctf-1.06 (Zhang, 2016). Les données ont ensuite été traitées avec Gautomatch (Zhang, non publié) et dans Relion2 (Kimanius et al., 2016), EMAN2 (Tang et al., 2007) et cryoSPARC (Punjani et al., 2017).

La figure 56 présente les résultats obtenus pour la collecte sur Titan de l'échantillon de 6his-SmPRMT3-fl à 0,7 μM . Comme le montre l'exemple de micrographie, les particules se répartissent à nouveau de façon homogène dans le tampon et ne semblent pas agrégées. Bien que la concentration ait été abaissée d'un facteur 10 par rapport au premier échantillon, la concentration apparente semble inchangée. Cela suggère que le carbone sur la grille pourrait être saturé en protéines, si bien qu'à 7 μM comme à 0,7 μM l'excès de protéine se retrouve dans les trous. Un ajustement supplémentaire de la concentration en protéine devra être réalisé. Le peu d'anneaux de Thon visibles sur le spectre de puissance montre qu'à nouveau, la glace est trop épaisse et dégrade donc le rapport signal/bruit. Ce paramètre nécessitera également un ajustement. Le traitement des données a néanmoins permis d'obtenir des classes 2D correctes, présentées sur la figure 56c.

Sur ces classes, on peut apercevoir des particules dont la plus longue dimension mesure environ 90 Å, ce qui est compatible avec la taille du dimère de SmPRMT3. Les classes dévoilent le détail de sous-structures à environ 6-10 Å, mais on n'observe pas la symétrie caractéristique produite par le positionnement tête-bêche des deux modules catalytiques. De ce fait, il est difficile d'identifier visuellement une PRMT. De plus, toutes les classes semblent représenter des particules

vues sous un axe à peu près similaire, ce qui indique que l'échantillon souffre d'une très forte orientation préférentielle des particules. Un modèle *ab initio* a tout de même été généré dans cryoSPARC à partir de 239000 particules issues des meilleures classes 2D. Ce modèle est représenté sur la figure 56d.

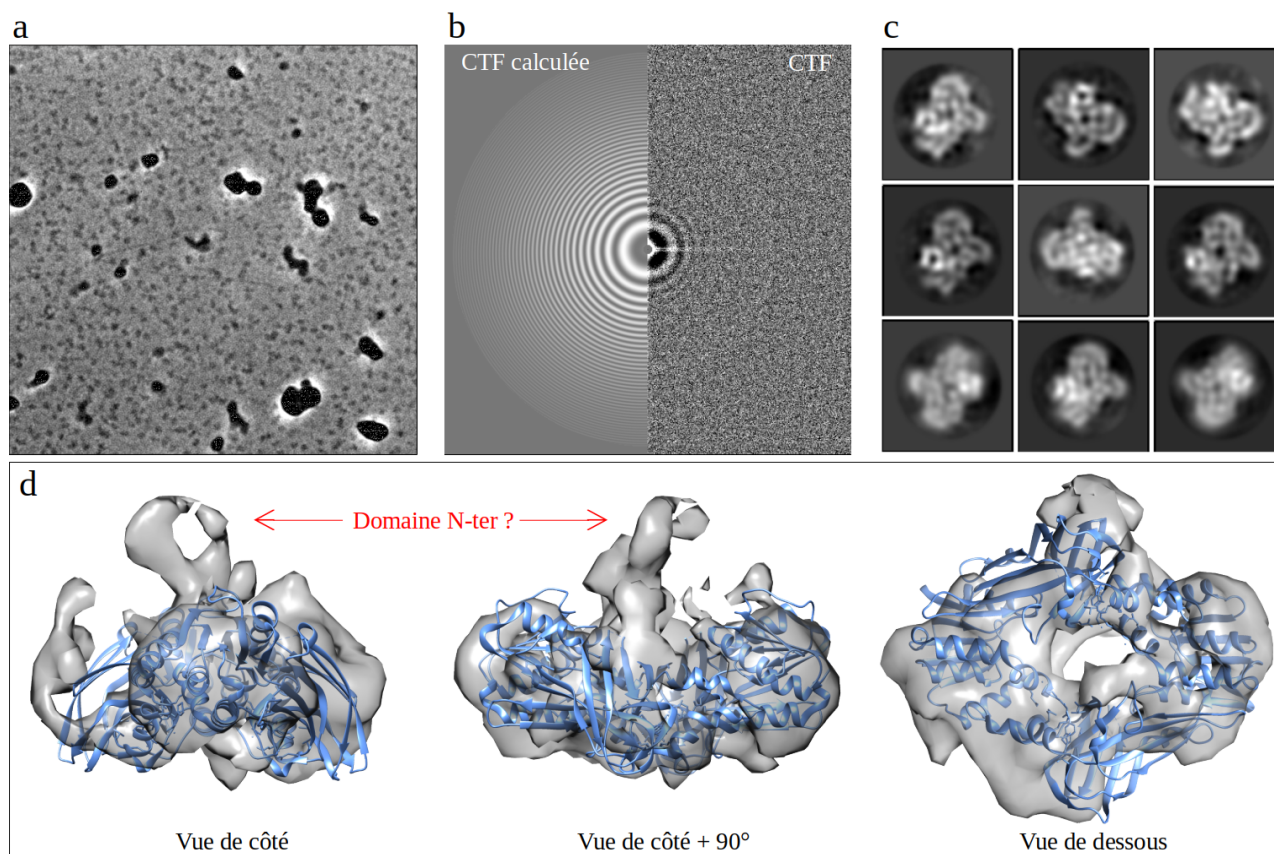


Figure 56 - Traitement des données collectées sur le microscope Titan Krios.

(a) Exemple de micrographie électronique pour l'échantillon collecté. (b) Calcul des paramètres de la CTF. À droite, spectre de puissance de la micrographie ci-contre. À gauche, le modèle de CTF calculé pour cette image avec Gctf. (c) Exemple des classes obtenues dans Relion2 après traitement du jeu de données. (d) modèle *ab initio* de SmPRMT3-fl calculé dans cryoSPARC. La structure cristallographique de la PRMT3 humaine (PDB 1F3L), représentée en rubans, a été placée dans la carte de densité.

Le logiciel EMAN2 a également été utilisé pour générer un modèle, mais sans succès. La forme générale du modèle calculé avec le programme cryoSPARC rappelle un dimère de module catalytique de PRMT, avec un diamètre d'environ 9 nm, une cavité centrale et une apparente symétrie C_2 . Le modèle montre également des densités supplémentaires à la verticale de l'orifice central, qui pourraient correspondre aux domaines amino-terminaux. Cependant, le modèle est de piètre qualité, vraisemblablement du fait de l'orientation préférentielle des particules.

Au regard de ces premiers résultats, les prochaines étapes seront d'améliorer la qualité de l'échantillon. La concentration en protéine devra être optimisée afin d'éviter la saturation en particules. Le traitement de la grille devra être adaptée pour affiner la glace. Pour cela, les paramètres de décharge de plasma et de buvardage pourront être ajustés. La décharge de plasma

permet de charger la pellicule de carbone sur la grille, qui est initialement hydrophobe. Dans l'air raréfié, des radicaux oxygène réagissent avec le carbone et induisent une charge négative de la grille. D'autres traitements permettent d'obtenir une charge positive ou d'augmenter son hydrophobie. Le buvardage est l'étape durant laquelle le surplus d'échantillon est absorbé par application de deux papiers absorbants. Durant le buvardage, la température et l'humidité de la chambre peuvent être ajustées, ainsi que la pression appliquée sur l'échantillon par les papiers absorbants. L'utilisation de grilles arborant des trous de plus petit diamètre pourrait permettre d'obtenir une glace fine plus aisément (Drulyte et al., 2018).

Deux facteurs peuvent être responsables de l'apparition d'orientations préférentielles. Premièrement, une glace trop fine peut forcer les particules à adopter une même conformation ou bien la présence d'un ménisque peut repousser les particules vers la périphérie des trous. Ce n'est pas le cas pour l'échantillon de SmPRMT3, pour lequel la glace est au contraire trop épaisse. Le second facteur peut être le déplacement des particules vers l'interface air-tampon juste avant que l'échantillon ne soit vitrifié. Entre le moment où l'échantillon est déposé sur la grille et la vitrification, les mouvements browniens peuvent engendrer le regroupement des protéines vers cette interface. De plus, le déplacement des particules est d'autant plus rapide que celles-ci sont de faible poids moléculaire. Une fois à l'interface, les particules ont tendance à s'orienter toutes de la même façon. C'est probablement le problème rencontré dans le cas de SmPRMT3, qui pour rappel n'a une masse moléculaire que de 130 kDa environ. Pour résoudre le problème prédominant d'orientation préférentielle des particules, on peut jouer sur la tension superficielle dans le tampon en utilisant des détergents, ou masquer l'interface air-tampon en utilisant des agents surfactants (Drulyte et al., 2018). Cela pourrait permettre d'augmenter l'éventail des orientations adoptées par SmPRMT3. Les détergents CHAPS, le DDM et NP-40 ainsi que le surfactant dodécylmaltoside ont démontré leur efficacité et seront testés en priorité.

3 - Étude préliminaire du complexe SmPRMT3-SmRpS2

La protéine ribosomique S2 est un substrat et partenaire bien caractérisé de PRMT3 chez les mammifères, les plantes et les levures à fission. Bien qu'aucune étude n'ait été réalisée sur *S. mansoni*, la conservation de l'interaction PRMT3-rpS2 dans le règne animal suggère qu'elle existe également chez ce parasite. L'expression du complexe SmPRMT3-SmRpS2 répond à deux objectifs. Le premier est de comprendre comment PRMT3 reconnaît, fixe et méthyle son partenaire dans le but de mieux comprendre les bases structurales de la méthylation des arginines par les PRMT. Ensuite, la co-cristallisation du complexe SmPRMT3-SmRPS2 plutôt que SmPRMT3 seule pourrait permettre d'obtenir de meilleurs cristaux. Pour finir, l'interaction entre PRMT3 et rpS2 est nécessaire au bon déroulement de la maturation des ribosomes. Comprendre cette interaction pourrait permettre de concevoir des composés ciblant un autre site que le site actif conservé de SmPRMT3 et déstabilisant le complexe SmPRMT3-SmRpS2. Chez l'homme, rpS2 est sur-exprimée dans plusieurs cancers et l'interaction PRMT3-rpS2 a été décrite comme stabilisant la protéine ribosomique. La déstabilisation de ce complexe pourrait donc être une stratégie de développement de thérapies contre le cancer.

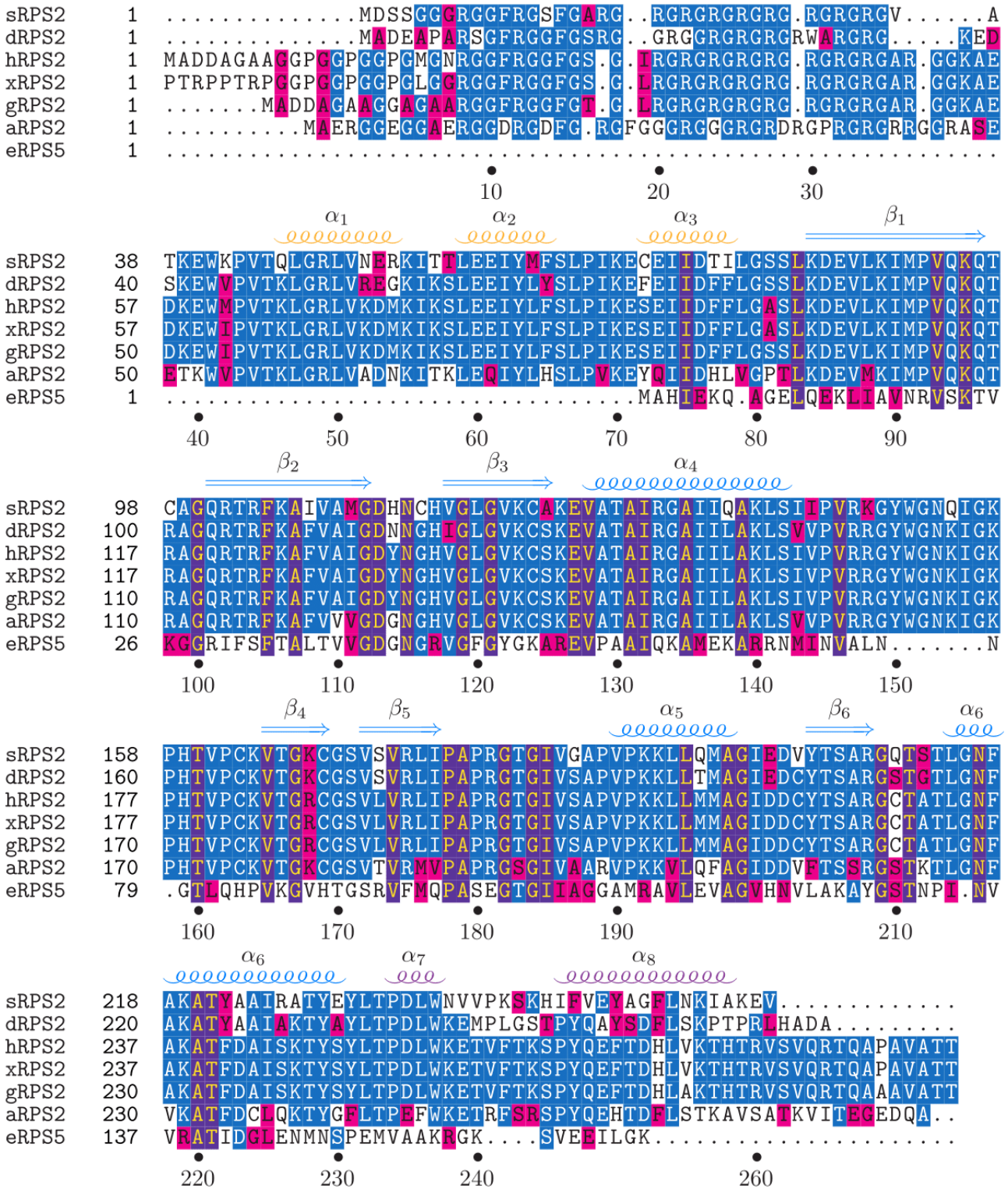
a - Comparaison des séquences de SmRpS2 et de ses homologues

L'alignement des séquences de SmRpS2 avec d'autres rpS2 et ses homologues rpS5 chez les bactéries et les archaea (alignement 3) révèle une importante homologie de séquence, particulièrement entre SmRpS2 et les rpS2 eucaryotes et rpS5 d'archaea. SmRpS2 possède à l'instar des autres rpS2 eucaryotes une région amino-terminale riche en répétition de RG, connue pour être méthylée par PRMT3. Cette région est inexistante chez les procaryotes, ce qui corrèle avec l'absence de PRMT3 dans ces organismes. Cela suggère que rpS2 est bien un substrat et partenaire de PRMT3 dans *S. mansoni*. Un modèle par homologie de SmRpS2 a été généré à l'aide de Modeller à partir des rpS2 humaine et de levure. Ce modèle est présenté sur la figure 57d.

b - Expression et purification de rpS2 et des complexes PRMT3-rpS2

Choi, Jung et leurs collaborateurs (Choi et al., 2008) ont montré lors de tests de rétention de PRMT3 par des constructions de GST-rpS2 que l'interaction entre PRMT3 et rpS2 de souris fait intervenir le domaine amino-terminal de PRMT3 et les résidus 100 à 293 de rpS2, représenté sur la figure 57c. Cependant leur étude ne permet pas de discriminer précisément quels sont les éléments du fragment 100-293 (81-261 chez *S. mansoni*) nécessaires à la fixation de PRMT3. Dans le contexte ribosomique (structure du ribosome humain, PDB 5AJ0) (Natchiar et al., 2017) présenté sur la figure 57a/b, rpS2 se localise au cœur de la petite sous-unité du ribosome et interagit avec des ARNr et des protéines. RpS2 se décompose en trois sections : un segment amino-terminal non structuré arborant le motif GAR (résidus 1-36 chez SmRpS2, en rouge sur la figure), un domaine de

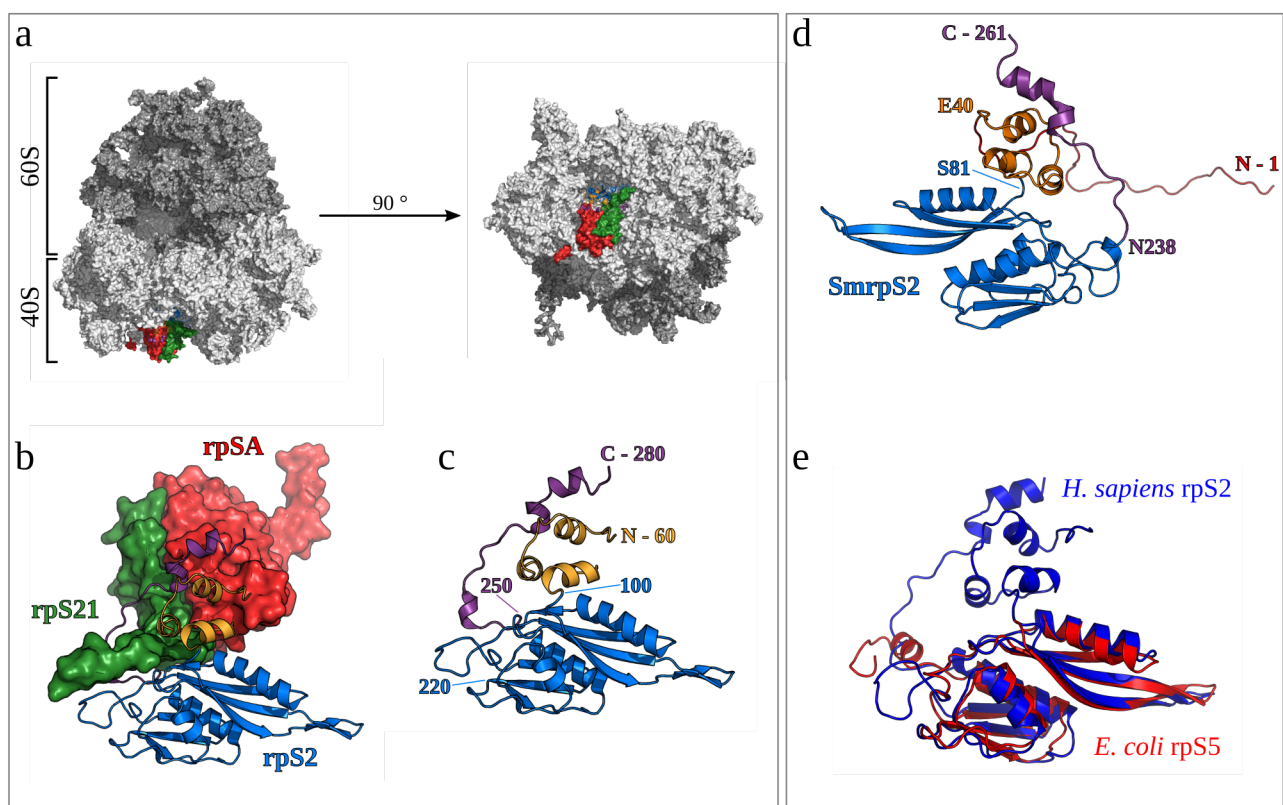
Motif GAR



Alignement 3 - Alignement de séquences de rpS2 eucaryotes et rpS5 bactérien.

L'alignement présente les séquences complètes de rpS2 eucaryotes et rpS5 bactérien. Les structures secondaires de SmpS2 sont montrées au-dessus de l'alignement. Le motif GAR, les trois hélices amino-terminales, le domaine de fixation aux ARN et le segment carboxy-terminal sont annotés en rouge, orange, bleu clair et violet. La numérotation de la dernière ligne correspond à SmpS2. Les acides aminés sont colorés en fonction de la similarité avec la séquence consensus : invariant (violet) ou similaire (bleu). Les abréviations utilisées sont : s/*Schistosoma mansoni*, d/*Danio rerio*, h/*Homo sapiens*, g/*Gallus gallus*, a/*Arabidopsis thaliana* et e/*Escherichia coli*. Toutes les séquences utilisées pour construire l'alignement ne sont pas présentées afin d'améliorer la lisibilité de la figure.

fixation à l'ARN (résidus 81-237 chez SmrpS2, en bleu sur la figure) et un domaine formé de quatre hélices α (résidus 40-80 et 238-261 dans SmrpS2, en orange et violet dans la figure). Le segment amino-terminal non structuré est méthylé par PRMT3. Sa fonction est inconnue mais ce type de séquences est décrit comme interagissant avec les protéines et particulièrement avec les acides nucléiques (Thandapani et al., 2013). Le domaine formé de quatre hélices α interagit très peu avec le reste de rpS2, mais semble stabilisé dans le ribosome par des interactions avec ses partenaires rpS21 et rpSA, respectivement en vert et rouge sur la figure 57b. Plusieurs des constructions (60-180, 60-220 et 180-293) utilisées par Choi, Jung et leurs collaborateurs pour borner les zones d'interactions entre PRMT3 et rpS2 coupent le domaine de fixation de l'ARN (en bleu) en plusieurs points et donc menacent l'intégrité structurale de rpS2. Cela peut avoir diminué la précision de leur détermination de la zone reconnue par PRMT3.



Comme la délimitation des régions de SmrpS2 intervenant dans l'interaction avec SmPRMT3 est incertaine, nous avons cloné et exprimé dans *E. coli* la protéine rpS2 complète (fl) et le fragment tronqué du segment non structuré arborant le motif GAR (résidus 40-261). Ces deux

constructions sont étiquetées avec la GST afin d'améliorer leur solubilité. La figure 58 présente les résultats d'un test d'expression de ces constructions de SmrpS2 dans *E. coli*. Le test consiste à exprimer les constructions de GST-rpS2 puis à les isoler du lysat cellulaire par adsorption sur une résine d'affinité pour le glutathion. Les fractions insolubles (les débris cellulaires sédimentés) et retenues (la résine d'affinité pour le glutathion) ont été analysées par électrophorèse. Le gel montre que les constructions fl et 40-261 sont exprimées mais totalement insolubles, malgré le tag GST.

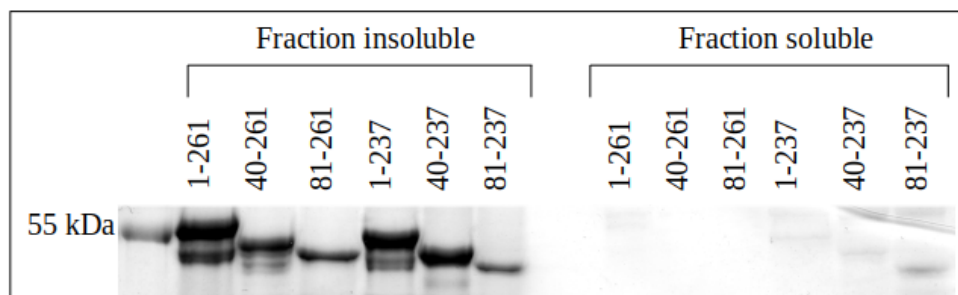


Figure 58 - Expression et solubilité de SmrpS2 dans *E. coli*.

Analyse par SDS-PAGE sur gel de polyacrylamide 12 % des tests de solubilité des fragments de SmrpS2. SmrpS2 et ses fragments sont exprimés dans *E. coli* mais principalement insolubles. La délétion de l'extrémité amino-terminale (1-40 et 1-81) ne rétablit pas la solubilité. Inversement, la délétion de l'extrémité carboxy-terminale (237-261) permet de rétablir partiellement la solubilité. La délétion des extrémités 1-81 et 237-261 engendre la solubilité maximale.

La co-expression de SmrpS2 avec les constructions SmPRMT3-fl et -C3 a été envisagée afin d'aider à solubiliser SmrpS2 durant son expression. De façon surprenante, l'expression de GST-rpS2 (fl et 40-261) dans *E. coli* inhibe la production de 6his-SmPRMT3-fl mais pas de 6his-SmPRMT3-C3 (données non présentées). SmPRMT3-C3 ne possède pas le domaine amino-terminal et ne peut donc pas interagir spécifiquement avec SmrpS2, et aucune solubilisation n'a donc été observée.

L'expression des deux constructions de SmrpS2, étiquetées GST ou hexahistidine à l'extrémité amino-terminale, a ensuite été testée en cellules d'insecte chez *S. frugiperda*, grâce au système d'expression baculovirus. Dans ces cellules, les systèmes ribosomiques et les chaperons eucaryotes pourraient être plus adaptés à l'expression de SmrpS2 que ceux de *E. coli*. Seules les constructions étiquetées GST ont pu être exprimées en partie soluble (données non présentées), ce qui est tout de même un progrès comparé à l'expression de SmrpS2 en système bactérien. Une interaction avec ses partenaires ribosomiques ou avec SmPRMT3 pourrait être nécessaire afin d'exprimer SmrpS2 sous forme totalement soluble.

Nous avons donc tenté de co-exprimer SmrpS2 et SmPRMT3 dans des cellules d'insecte. La figure 59 présente le résultat des tests d'expression et d'interaction de la construction 6his-3c-SmPRMT3-fl avec des constructions de GST-3c-SmrpS2 co-exprimées dans des cellules d'insectes. Dans ce test, similaire au test d'expression présenté ci-dessus, l'existence d'une interaction entre les

fragments de SmrpS2 et SmPRMT3 se traduit par la co-purification des deux protéines sur la résine d'affinité pour la GST. À l'inverse, si SmPRMT3 n'interagit pas avec le fragment de SmrpS2, elle n'est pas retenue lors de cette purification.

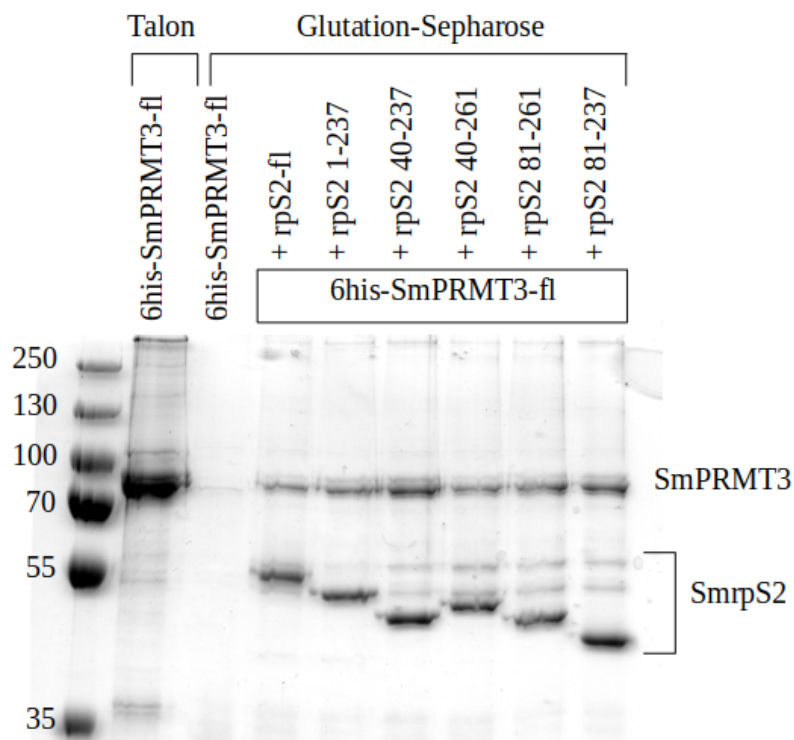


Figure 59 - Interaction des complexes SmPRMT3 – SmrpS2.

Analyse SDS-PAGE (12 %) de mini-purifications du complexe SmPRMT3 - SmrpS2 par affinité pour le glutathion. 6his-SmPRMT3-fl n'interagit pas seule avec la résine GSH, à moins d'être étiquetée GST. 6his-SmPRMT3-fl est retenue par l'ensemble des constructions de SmrpS2 étiquetées GST, même les plus courtes.

L'approche par co-expression des deux partenaires a permis la production d'un complexe soluble dans *S. frugiperda*, pour les constructions complètes et 40-261 de SmrpS2. En effet, on observe sur le gel (figure 59) les bandes qui correspondent aux deux partenaires. Cela confirme aussi l'existence de l'interaction supposée entre rpS2 et SmPRMT3 dans *S. mansoni*.

Une première purification du complexe SmPRMT3-SmrpS2 (fl/fl) co-exprimé en cellules d'insecte a été réalisée. La purification s'est déroulée en deux étapes : le complexe a été isolé par chromatographie d'affinité pour le glutathion à haute force ionique

(500 mM de NaCl) et en présence de nucléases afin de se débarrasser des acides nucléiques interagissant avec SmrpS2, tout en maintenant le complexe soluble. Le complexe a ensuite été élué par clivage du tag GST par la protéase 3C puis purifié par filtration sur gel en diminuant la concentration en NaCl à 100 mM. Durant toutes les étapes, l'agent réducteur TCEP a été ajouté à une concentration de 5 mM dans les tampons pour éviter l'agrégation potentielle du complexe qui pourrait être provoquée par la présence de trois résidus cystéines exposées à la surface de SmrpS2.

La figure 60a/c présente les résultats obtenus lors de cette purification. Sur le chromatogramme, le profil d'éluion (en violet), correspondant au complexe SmPRMT3-fl-SmrpS2-fl, montre que la majorité de l'échantillon contenant le complexe est élué dans le volume mort de la colonne. Cela signifie que le complexe est principalement agrégé. Lors des travaux précédents, nous avons observé que SmPRMT3 est une protéine très soluble. Le fait que le complexe soit agrégé signifie que SmPRMT3 n'est pas suffisante pour solubiliser SmrpS2 hors de son contexte ribosomique.

Comme indiqué ultérieurement, le domaine de rpS2 formé de quatre hélices est peu solidaire du reste de la protéine et interagit, dans un contexte ribosomique, avec les protéines SA et S21. Cette interaction pourrait être nécessaire à son organisation et à sa stabilisation. Ainsi lorsque rpS2 est exprimée en dehors du ribosome, les interactions manquantes avec ses partenaires pourraient engendrer le mauvais repliement de ce domaine et conduire à l'agrégation de SmrpS2. Comme illustré en figure 57d, cette hypothèse est d'autant plus plausible que trois des quatre hélices constituant ce domaine sont placées en tandem dans la structure primaire (résidus 40-80 dans SmrpS2, en orange sur la figure) tandis que la quatrième hélice se localise à l'extrémité carboxy-terminale (résidus 245-261 dans SmrpS2, en violet sur la figure). La quatrième hélice est reliée à l'extrémité carboxy-terminale du domaine de liaison à l'ARN (résidus 81-237 dans SmrpS2) par un segment non structuré interagissant, dans le ribosome, avec rp21 (figure 57b). L'interaction entre l'hélice carboxy-terminale et le trio d'hélices amino-terminales est médiée par plusieurs résidus hydrophobes, qui s'ils sont exposés, pourraient expliquer l'agrégation de rpS2 observée.

Afin de vérifier cette hypothèse, de nouvelles constructions de SmrpS2 ont été clonées, avec une étiquette GST, puis exprimées dans *E. coli*. Les limites de ces constructions respectent les domaines de SmrpS2 afin de ne pas menacer l'intégrité structurale des différents fragments, présentés en figure 57d. Le fragment 1-237 (segments rouge+orange+bleu) comprend la région non structurée amino-terminale arborant le motif GAR, le trio d'hélices du domaine amino-terminal, le domaine de liaison à l'ARN, mais pas la quatrième hélice carboxy-terminale. Le fragment 40-237 (segments orange+bleu) comprend le trio d'hélices du domaine amino-terminal, le domaine de liaison à l'ARN, mais pas la quatrième hélice carboxy-terminale. Le fragment 81-261 (segments bleu+violet) comprend uniquement le domaine de liaison à l'ARN et l'hélice carboxy-terminale. Le fragment 81-237 (segment bleu) n'est constitué que du domaine de liaison à l'ARN. Comme illustré en figure 57e, ce dernier correspond structurellement à l'homologue de rpS2 dans *E. coli*, rpS5, et pourrait donc être plus soluble dans *E. coli* car plus aisément pris en charge par le système de chaperons bactérien.

La figure 58 page 107 présente le résultat du test d'expression et de solubilité dans *E. coli* de ces constructions. Les fragments issus de la suppression de segments amino-terminaux, notés N-261 (pour 1-261, 40-261 et 81-261), sont exprimés mais totalement insolubles. Les fragments N-237 sont principalement insolubles mais retrouvent une légère solubilité, en particulier 81-237, qui est l'homologue structurale de rpS5 de *E. coli*. Bien que ces fragments ne soient toujours pas suffisamment solubles pour être exprimés dans *E. coli*, ce test semble néanmoins confirmer que le segment 237-261 (l'hélice carboxy-terminale, en violet sur la figure 57) pourrait être responsable de l'agrégation observée lors de la purification du complexe.

Comme Choi, Jung et leurs collaborateurs ont indiqué que le segment 100-293 de rpS2 de souris (le fragment 81-261 chez SmrpS2) est nécessaire à la fixation de PRMT3, nous avons souhaité vérifier que l'ensemble des fragments de SmrpS2 interagissent toujours avec SmPRMT3. Comme précédemment, les fragments ont été co-exprimés avec SmPRMT3-fl dans des cellules d'insectes et les potentiels complexes ont été purifiés par affinité des constructions GST-SmrpS2 pour la résine glutathion. La figure 59 page 108 montre que l'interaction SmPRMT3-SmrpS2 est maintenue pour l'ensemble des fragments de SmrpS2, ce qui prouve que le segment 81-237 (100-250 chez la souris) est suffisant pour la fixation à SmPRMT3. SmPRMT3 fixe donc SmrpS2 principalement par son domaine de fixation à l'ARN, et ne nécessite pas l'hélice carboxy-terminale constituant le domaine de quatre hélices, à l'inverse de ce qui a été proposé par Choi, Jung et leurs collaborateurs. Le complexe SmPRMT3-SmrpS2(81-237) a été purifié selon le protocole indiqué précédemment. Les résultats de cette purification sont présentés sur la figure 60a/b.

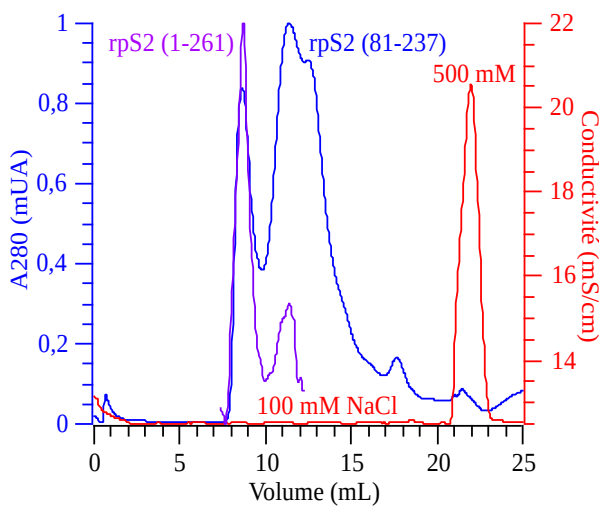
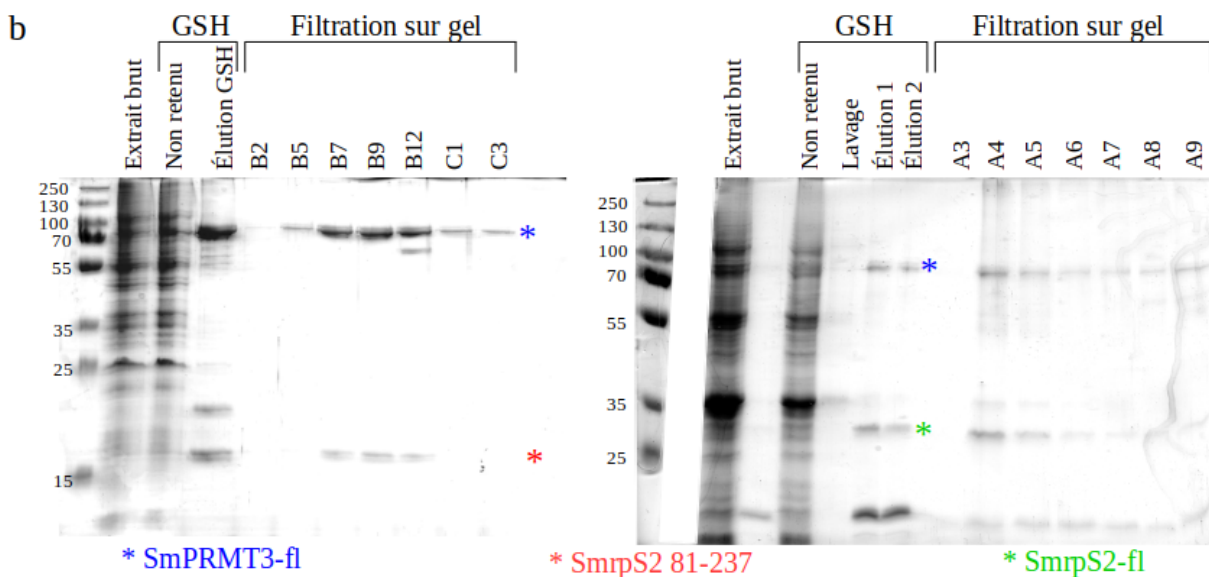


Figure 60 - Purification des complexes PRMT3-rpS2.

(a) Filtration sur gel (Superdex 200 10/300 GL) pour les complexes SmPRMT3-SmrpS2 fl/fl (en violet) et fl/81-237 (en bleu). (b) Analyse par SDS-PAGE sur gel de polyacrylamide 12% des étapes de purification des complexes SmPRMT3-SmrpS2 fl/81-237 (à gauche) et fl/fl (à droite). Les complexes sont efficacement élués par clivage par la thrombine sur résine d'affinité pour le glutathion. Le chromatogramme de filtration sur gel (a) montre que le complexe formé entre SmPRMT3-fl et SmrpS2-fl s'agrège en grande partie, tandis que le complexe formé avec la forme tronquée en SmrpS2 demeure totalement soluble. Pour ce dernier, un contaminant est présent dans la fraction B12. Il n'a pour l'heure pu être identifié.



Comme le montre le profil d'élution bleu, une minorité d'échantillon est cette fois élue dans le volume mort (fraction A3) tandis que la majeure partie est incluse dans la résine et s'élue selon un double pic d'élution se chevauchant. L'analyse par électrophorèse indique que la fraction exclue de la colonne ne contient pas le complexe, tandis que le pic double correspond au complexe SmPRMT3-SmrpS2(81-237) soluble.

Dans cette étude préliminaire, nous avons pu montrer que PRMT3 et rpS2 de *S. mansoni* forment un complexe stable *in vitro*. Nous avons également montré que ce complexe est maintenu par l'interaction entre PRMT3 et le domaine de fixation à l'ARN de rpS2, et ne nécessite pas d'interaction avec le domaine formé de quatre hélices α . Nous avons pu purifier sous forme soluble un complexe constitué de la SmPRMT3 complète et du domaine de fixation à l'ARN (81-237) de SmrpS2, ce qui ouvre la voie aux étapes suivantes de caractérisation du complexe et de son étude structurale.

Plusieurs phases suivront ce travail initial. Les complexes formés par SmPRMT3 avec les constructions 1-237 et 40-237 de SmrpS2 devront être exprimés et purifiés afin de déterminer quel est le plus grand fragment permettant d'obtenir un complexe stable et soluble. Les homogénéités structurale et stœchiométrique des complexes devront être évalués par DLS, SAXS et spectrométrie de masse. Une fois caractérisés, un ou plusieurs de ces complexes pourront donner lieu à une étude structurale par cristallographie et/ou cryo-microscopie électronique en transmission. Finalement, un complexe ternaire impliquant PRMT3, rpS2 et PDCD2L a récemment été identifié chez les mammifères (Landry-Voyer et al., 2016). Ce complexe existe probablement aussi chez *S. mansoni* et devrait faire l'objet d'analyses, car il présente un intérêt majeur pour la compréhension du mécanisme de maturation du ribosome.

C - Matériel et méthodes

Cette section « Matériel et méthodes comporte sept parties :

- 1 Modélisation par homologie
- 2 Clonage des constructions de SmPRMT3 et de SmrpS2
- 3 Expression et purification des constructions SmPRMT3
- 4 Expression et purification des complexes SmPRMT3-SmrpS2
- 5 Caractérisation biophysique de SmPRMT3-fl et -C3
- 6 Cristallisation de SmPRMT3
- 7 Étude de SmPRMT3 par cryo-microscopie électronique

1 - [Modélisation par homologie](#)

➤ **SmPRMT3**

Les séquences issues des banques de séquences Uniprot et NCBI des PRMT3 de 10 organismes dont SmPRMT3 ont été ajoutées à l'alignement 1 présenté en introduction à l'aide du programme d'alignement MAFFT (Kato et al., 2013). L'alignement 2 a été affiné dans le programme Jalview (Waterhouse et al., 2009). Les séquences alignées des PRMT3 ont été extraites. Un modèle du module catalytique de SmPRMT3 a été généré par homologie avec la structure de la PRMT3 de rat (1F3L), à l'aide du programme Modeller (Sali et al., 1993). Dans l'alignement présenté, toutes les séquences ne sont pas présentées afin d'améliorer la lisibilité.

➤ **SmrpS2**

Les séquences de rpS2 et rpS5 de 15 organismes eucaryotes, archaea et bactériens dont *S. mansoni* ont été récupérées sur Uniprot puis alignées à l'aide le programme Clustal Omega (Sievers et al., 2014). L'alignement 3 a été corrigé à la main dans le programme Jalview. Un modèle de SmrpS2 a été généré par homologie aux structures des rpS2 de levure, drosophile et humaine extraites de structures du ribosome (respectivement les PDB 4U4R, 4V6W et 5AJ0), à l'aide du programme Modeller. Dans l'alignement présenté, toutes les séquences ne sont pas présentées afin d'améliorer la lisibilité.

2 - Clonage des constructions de SmPRMT3 et de SmrpS2

a - Protéolyse ménagée

➤ **Par la trypsine**

Les échantillons des fragments 6his-Trx-TEV-SmPRMT3(141-553) et son produit de digestion par la TEV ont été incubés avec une concentration croissante en trypsine pancréatique bovine (Sigma) établie entre 10^{-4} mg.mL⁻¹ et 1 mg.mL⁻¹. La PRMT a été utilisée à une concentration finale de 0,25 mg.mL⁻¹. Après 30 minutes à 20 °C dans 20 mM Tris-HCl pH 8, 100 mM NaCl, 1 mM TCEP, la réaction a été arrêtée sur glace. Chaque échantillon a été divisé en deux puis séparé par SDS-PAGE et PAGE natif.

➤ **Par d'autres protéases**

Le fragment 184-553 (C3) déterminé par trypsinolyse a été digéré par les enzymes α -chymotrypsine, papaïne, subtilysine et endoprotéinase du kit ProtiAceI selon le protocole indiqué (Hampton).

b - Les plasmides pnEA

Les plasmides pnEA (Diebold et al., 2011) sont des dérivés de pET15 permettant la sur-expression de protéines de fusion dans des cellules bactérienne disposant d'un système d'expression T7 et sous pression de sélection par l'ampicilline. Les tags peuvent être positionnés aux extrémités 5' ou 3' du gène inséré et permettent l'expression de protéines étiquetées histidine avec ou sans thioredoxine (6his- et 6his-TRX-) et Gluthation-S-transférase (GST). Ces tags arborent également un site de coupure par les protéases TEV, thrombine (thr) ou 3C permettant de cliver la protéine de fusion.

c - Sous-clonage des fragments de SmPRMT3 - expression en bactéries

Les constructions pnEA_{vHX}-SmPRMT3-C2 et -C3 ont été obtenus par délétion par PCR (ADN polymérase Phusion, Thermo™) depuis le plasmide pnEA_{vHX}-SmPRMT3-C1 avec les amorces ci-dessous :

- SM3-170_sens : 5' -AATAAAACATATGTTACTTTCAACTATGTCTGCCC-3'
- SM3-184_sens : 5' -AATAAAACATATGTCACCGTGTGATCAATCGTC-3'
- SM3_anti : 5' -AATAAAACATATGGCCCTGGAAGTAAAGG-3'

Les amplicons ont été incubés 1 heure à 37 °C en présence de *DpnI* pour dégrader les plasmides matrice puis purifiés à l'aide du kit « Gel and PCR clean-up » (NucleoSpin®). La digestion des amplicons par *NdeI* pour générer les extrémités cohésives a échoué, probablement du fait de la proximité des sites de restrictions des extrémités des amplicons. Pour palier ce problème, les amplicons ont été phosphorylés par la polynucléotide kinase du phage T4 (Thermo™) puis re-

circularisés par ligation avec la ligase du phage T4 (Thermo™) selon les protocoles recommandés par le fournisseur. Les extrémités cohésives ont été générées par digestion avec *NdeI* sur ces plasmides après amplification dans *E. coli* DH5α puis les plasmides ont été re-circularisés comme décrit précédemment.

Les fragments C2 et C3 ont été sous-clonés dans les plasmides pnEAv-H et pnEAv/H par digestion-insertion à l'aide des enzymes de restrictions *BamHI* et *NdeI* (NEB™), de la phosphatase alcaline bovine (NEB™) et de la ligase du phage T4 (Thermo™) dans les conditions recommandées par le fournisseur. Les constructions dépourvues du site de clivage par la protéase TEV, pnEAH-SmPRMT3-fl et -C3 ont été obtenues par délétion par PCR depuis les plasmides pnEAHX-SmPRMT3-fl et pnEAvH-SmPRMT3-C3 avec les amorces ci-dessous :

- pnEAH_SmPRMT3fl_F : 5' -ATGACCGATTCAAACCCAATGG-3'
- pnEAH_SmPRMT3-184_F : 5' -TCACCGTGTGATCAATCGTC-3'
- pnEAH_R : 5' -GTGATGATGATGATGATGGCTG-3'

Les fragments C4 et C5 ont été obtenus par délétion par PCR depuis le plasmide pnEAH-SmPRMT3-C3 avec les amorces ci-dessous :

- SmPRMT3_199-553_f : 5' -AATGATTCTTCATATTTTGGATC-3'
- SmPRMT3_210-553_f : 5' -TTTGAAATACACGGTGAAATG-3'
- pnEA_R : 5' -CATGGTATATCTCCTTC-3'

Les amplicons ont été purifiés, phosphorylés puis re-circularisés tel que décrit plus haut.

d - Sous-clonage de SmrpS2 - expression en bactérie et cellules d'insecte

Le clone de rpS2 a été fourni par Raymond Pierce de l'Institut Pasteur de Lille.

➤ **Système d'expression bactérien**

Les fragments SmrpS2 1-261 et 40-261 ont été amplifiés par PCR depuis le plasmide

pCR2.1_SmrpS2 contenant le transcrit codant pour la protéine avec les amorces ci-dessous :

- pnEAv_SmRPS2_1-261_f : 5' -CCTTTACTTCCAGGGCCATATGGATTTCGTCAGGTGGG-3'
- pnEAv_SmRPS2_40-261_f : 5' -CCTTTACTTCCAGGGCCATGAATGGAAGCCTGTTAC-3'
- pnEAv_SmRPS2_r : 5' -TGCTTAAGCTCGAGCATATCAAACCTTCTTTAGCAATTTTGG-3'

pnEAvG a été digéré par *NdeI* (NEB™) en présence de CIP (Thermo™). Les amplicons et plasmides ont été purifiés tel que décrit plus haut puis introduits dans le plasmide pnEAvG et par la méthode du clonage séquence/ligation-indépendant (SLIC) avec l'ADN polymérase du phage T4 (NEB™) en présence de tampon NEB2.1. Les plasmides pnEAvG_SmrpS2_81-261, 1-237 et 40-

237 ont été générés par délétion par PCR à partir des plasmides pnEAvG_SmrsS2_1-261 et 40-261 avec les couples d'amorces ci-contre :

Délétions carboxy-terminales :

- pnEAv_F : 5' -TGATATGCTCGAGCTTAAGCAATTG-3'
- SmRPS2-N-237_R : 5' -CCATAAATCAGGTGTTAGATACTC-3'

Délétions amino-terminales :

- pnEAv_R : 5' -CATATGGCCCTGGAAGTAAAGG-3'
- SmRPS2-81_F : 5' -TCAAGTTTAAAAGATGAAGTAC-3'

Le plasmide pnEAvG_SmrsS2_81-237 a été obtenu par délétion par PCR à partir du plasmide pnEAvG_SmrsS2_1-237 avec les amorces ci-dessus. SmrsS2_1-261 a également été sous-cloné dans pMMSchX afin d'être co-exprimé avec SmPRMT3-fl dans *E. coli*. Le plasmide pMMSchX est un dérivé des pnEA incluant un gène de résistance au chloramphenicol et une origine de réplication différente de celle des pnEA. Ils ont été développés pour des productions en co-expression et mis à disposition par l'équipe du Dr. Romier à l'IGBMC. Les amplicons ont été purifiés, phosphorylés puis re-circularisés tel que décrit pour le clonage des fragments de SmPRMT3.

➤ **Système d'expression baculovirus-cellules d'insecte**

Les clonages, création et amplification des virus, l'infection des cellules d'insecte et l'expression des protéines ont été réalisés par Nathalie Troffer-Charlier au sein de l'équipe et du service baculovirus de l'IGBMC.

Les fragments SmrsS2 1-261 (complète/fl) et 36-261 ont été sous clonés dans les vecteurs pAc8-GST et pAc8-His (Abdulrahman et al., 2009) par SLIC après amplification avec les amorces ci-dessous :

- forward : 5' -CTGTTCCAGGGGCCCCATATGGATTCGTCAGGTGGG-3'
- forward36 : 5' -CTGTTCCAGGGGCCCCATATGGTTGCAACAAAAGAATGGAAGC-3'
- reverse : 5' -CGGATCCGTTTAAACCATATGTCAAACTTCTTTAGCAATTTTG-3'

Les vecteurs exprimant des fragments de SmrsS2 ont été obtenus suivant la même procédure que pour les plasmides pnEAvG_SmrsS2. Pour la co-exprimer avec rpS2, SmPRMT3-f a également été sous clonée dans les plasmides pAc8-GST et pAc8-His par SLIC après amplification avec les amorces ci-dessous :

- forward : 5' -CTGTTCCAGGGGCCCCATATGACCGATTCAAACCC-3'
- reverse : 5' -CGGATCCGTTTAAACCATATGTCACCGATTAAATCAAAG-3'

e - Amplification et séquençage des plasmides

Les plasmides ont été introduits dans les bactéries *E. coli* DH5 α par choc thermique puis amplifié dans ces cellules. Les plasmides ont été extraits des cellules à l'aide du kit NucleoSpin® Plasmid (Macherey-Nagel). Les différentes constructions ont été séquençées par la méthode de Sanger par le service de séquençage de GATC Biotech/Eurofins et vérifiées.

3 - Expression et purification des constructions SmPRMT3

a - Expression des constructions courtes SmPRMT3-C1 à -C4

Les cellules *E. coli* BL21(DE3) ont été transformées par choc thermique avec les plasmides pnEA_xSmPRMT3-C1 à -C4, cultivées 60 minutes à 37 °C dans du milieu LB sans antibiotique puis étalées sur LB agar supplémenté avec 100 $\mu\text{g.mL}^{-1}$ d'ampicilline. Les transformants ont été cultivés 12 heures à 37 °C puis resuspendus dans du milieu LB quelques minutes avant d'être introduits dans les fioles de culture, constituées de milieu LB deux fois concentré supplémenté avec 200 $\mu\text{g.mL}^{-1}$ d'ampicilline. La culture s'est poursuivie à 37 °C, 210 rotations par minutes. Lorsqu'une densité optique à 600 nm de 0,4 a été atteinte, la température a été réglée sur 20 °C, puis l'expression des protéines a été induite par l'addition de 0,4 M d'IPTG lorsque la densité optique à 600 nm a atteint 0,7. L'induction s'est poursuivie pendant une durée de 12 à 15 heures. Les cellules ont été récoltées par centrifugation à 1000 g pendant 30 minutes, puis directement utilisées pour la purification ou stockées à -20 °C dans leur tampon de lyse.

b - Expression des constructions complète SmPRMT3-fl

Les cellules et les milieux de cultures ont préparés de manière analogue aux constructions courtes. Il en va de même pour le déroulement initial de la culture. Les protéines ont été exprimées à 30 °C pendant 6 heures.

c - Lyse bactérienne et clarification de l'extrait brut

Le volume de lyse à prévoir est de 30 mL par litre de culture. Les cellules ont été lysées par sonication sur glace trois fois 5 minutes pour un volume compris entre 100 et 200 mL, 5 fois 3 minutes pour un volume compris entre 50 et 100 mL, avec une amplitude d'onde de 60 % et une pulsation de 0,5 seconde (Labsonic Sartorius P). Entre chaque cycle de sonication, les suspensions cellulaires ont été homogénéisées puis laissées au repos 30 secondes sur glace afin de faire baisser la température. Le lysat cellulaire a été centrifugé pendant 45 minutes à 42000 g afin de séparer la fraction soluble (extrait cellulaire) des débris cellulaires.

d - Purification de 6his-Trx-TEV- et 6his-SmPRMT3-fl

Les cellules ont été lysées dans 250 mM NaCl, 50 mM Tris-HCl pH 8, 1 mM TCEP, 0,01 % NP40, cocktail d'antiprotéases Complete® EDTA-free (une pastille par 50 mL de lysat), et le lysat a été traité tel qu'indiqué dans la section « lyse bactérienne et clarification de l'extrait brut ». L'extrait cellulaire a été supplémenté avec 10 mM d'imidazole puis incubé pendant 2 heures avec la résine de chromatographie d'affinité sur ions métalliques immobilisés (Talon, Clontech) à 4 °C sous agitation douce. La résine a été déposée dans une colonne et lavée avec 10 volumes de tampon de lavage (50 mM NaCl, 50 mM Tris-HCl pH 8, 1 mM TCEP). L'échantillon a été élué 4 fois dans 1 volume de tampon de lavage supplémenté par 150 mM d'imidazole.

Les éluats contenant la protéine ont été regroupés puis dilués dix fois dans 20 mM Tris-HCl pH 8, 1 mM TCEP, 10 µM EDTA afin de faire baisser la concentration en sel à 25 mM. La protéine a été ensuite purifiée par chromatographie échangeuse d'anions (ANX FF, HiTrap) avec un gradient de NaCl de 25 à 500 mM, dans 20 mM Tris-HCl pH 8, 1 mM TCEP, 10 µM EDTA, à 1 mL.min⁻¹.

Les fractions contenant la protéine ont été rassemblées et concentrées sur membrane (Amicon centrifugal filter, 50 kDa) puis chargées sur une colonne de filtration sur gel (Superdex 200, 120 mL) et éluées dans 100 mM NaCl, 20 mM HEPES pH8, 1 mM TCEP. Les fractions contenant la protéine ont été rassemblées et concentrées sur membrane. Pour la construction étiquetée 6his-Trx-thr, la fusion peut être clivée dans le concentrat avant la filtration sur gel, avec 0,02 U d'enzyme par microgramme de substrat, pendant 1 heure à 37 °C.

e - Purification de SmPRMT3-C1

Les cellules ont été lysées dans 250 mM NaCl, 50 mM Tris-HCl pH 8, 5 % glycérol, 1 mM TCEP, 0,01 % NP40, antiprotéases sans EDTA, et le lysat a été traité tel qu'indiqué dans la section « lyse bactérienne et clarification de l'extrait brut ». L'incubation sur résine d'affinité s'est déroulée tel que décrit plus haut. Avant de procéder à une digestion par la protéase TEV, la résine a été lavée et l'imidazole éliminé avec 10 volumes de tampon de lavage (Tampon de lyse sans NP40 ni antiprotéases), par gravité. La digestion par la TEV a eu lieu sur la résine durant 12 heures à 4 °C, sous agitation douce. L'échantillon a été ensuite élué par gravité dans 3 volumes de tampon de lavage.

L'éluat a été dilué et purifié par chromatographie échangeuse d'anion comme indiqué au-dessus, dans 20 mM Tris-HCl pH 8, 1 mM TCEP mais sans EDTA.

Les fractions contenant la protéine ont été rassemblées, concentrées à 3 mg.mL⁻¹ sur membrane (Amicon centrifugal filter, 30 kDa) puis purifiées par filtration sur gel tel qu'indiqué pour la construction complète, dans 100 mM NaCl, 20 mM Tris-HCl pH8, 1 mM TCEP.

f - Purification de SmPRMT3-C2 et -C3

Les cellules ont été lysées dans 250 mM NaCl, 50 mM Tris-HCl pH 8, 1 mM TCEP, 0,01 % NP40, en présence d'antiprotéases (Complete® EDTA-free), et le lysat a été traité tel qu'indiqué dans la section « lyse bactérienne et clarification de l'extrait brut ». La première étape de purification par affinité a été réalisée selon le même protocole que décrit précédemment, avec une élution par l'imidazole. Les éluats ont été dialysés contre le tampon d'échangeuse d'anions afin de diminuer les concentrations en sel et imidazole puis l'échantillon a été purifié par échangeuse d'anion comme indiqué ci-dessus. Les fractions contenant la protéine ont été concentrées sur membrane puis digérées par la protéase TEV, pendant 12 heures à 4 °C sous agitation douce. La dernière étape de filtration sur gel est analogue à celle effectuée pour la construction C1.

g - Criblage de tampons pour la stabilisation de SmPRMT3

1 µg de 6his-Trx-TEV-SmPRMT3-C3 ont été incubés pendant 12 heures, à 20 °C dans 100 mM NaCl, 1 mM TCEP avec 50 mM en différents tampons couvrant une large gamme de pH, en présence ou en absence de protéase TEV sous un volume final de 10 µL :

- MES pH 5,5, 6,0 et 6,5
- Bis-Tris pH 6,0, 6,5 et 7,0
- MOPS pH 6,5, 7,0 et 7,5
- HEPES pH 7,0, 7,5 et 8,0
- Bicine pH 7,5, 8,0 et 8,5
- Tricine pH 7,5, 8,0 et 8,5
- Tris-HCl pH 7,5 et 8,0
- Bis-Tris-Propane (BTP) pH 6,5, 8,0 et 9,5

La réaction a été arrêtée par l'ajout de tampon de charge pour électrophorèse et les échantillons ont été analysés par SDS-PAGE.

h - Purification de 6his-TEV-SmPRMT3-C3, 6his-SmPRMT3-C3/-C4 et SmPRMT3-C4-6his

Les cellules ont été lysées dans 250 mM NaCl, 50 mM BTP pH 9, 1 mM TCEP, 0,01 % NP40, antiprotéases sans EDTA, et le lysat a été traité tel qu'indiqué dans la section « lyse bactérienne et clarification de l'extrait brut ». La première étape de purification par affinité a été réalisée selon le même protocole que décrit précédemment, avec une élution, en présence de PMSF à 100 µM, par l'imidazole ou par digestion sur colonne par la protéase TEV, selon que l'on souhaite cliver le tag hexahistidine ou non (uniquement pour la construction 6his-TEV-SmPRMT3-C3). Les fractions contenant la protéine ont été rassemblées et diluées 10 fois dans 20 mM BTP pH 9, 1 mM TCEP, 1 mM EDTA puis purifiées par chromatographie échangeuse d'anion dans 20 mM BTP pH 9, 1 mM TCEP, 1 mM EDTA avec un gradient de 25 mM à 500 mM NaCl. Les fractions contenant la protéine ont été rassemblées et précipitées avec 1 volume de sulfate d'ammonium à

4 M, puis les protéines ont été récoltées par centrifugation à 15000 g, 15 minutes à 4 °C. Les culots protéiques ont été repris dans 1 mL de tampon de filtration sur gel (50 mM NaCl, 20 mM BTP pH 9, 1 mM TCEP, 1 mM EDTA) puis purifiés par chromatographie d'exclusion. La protéine a ensuite été concentrée sur membrane Amicon® jusqu'à environ 15 mg.mL⁻¹.

4 - Expression et purification des complexes SmPRMT3-SmrpS2

a - Tests d'expression et de solubilité pour l'expression en bactérie

Les cellules *E. coli* BL21(DE3) ont été transformées par choc thermique avec les plasmides pNEAvG_SmrpS2 ou pMMSchX_SmrpS2_1-261 et pNEAH_SmPRMT3-fl/-C3. Une colonie de chaque transformant a été utilisée pour inoculer 5 mL de milieu de culture LB supplémenté avec les antibiotiques correspondants et les cultures ont été cultivées à 37 °C, avec 200 rpm. Lorsque les cultures ont atteint une densité optique à 600 nm de 0,6, l'expression des protéines a été induite par addition de 0,5 mM d'IPTG final pendant 3 heures à 37 °C ou 10 heures à 20 °C. Les cellules ont été récoltées par centrifugation puis lysées par sonication dans un tampon constitué de 50 mM Tris-HCl pH 8, 200 mM NaCl, 1 mM TCEP, supplémenté par les antiprotéases PMSF (Sigma) et Complete EDTA-free (sigma). Les lysats ont été clarifiés par centrifugation puis les surnageants ont été incubés 30 minutes avec 25 µL de résine glutathion-sépharose 4B (GE Healthcare®), à 4 °C sous agitation douce. La résine est lavée au tampon de lyse puis est analysée par SDS-PAGE avec les fractions insolubles.

b - Tests d'interaction en système baculovirus-cellules d'insecte

1,5.10⁶ cellules sf9 dans 3 ml de milieu de Grace (Sigma) supplémenté avec 10 % de serum de veau fœtal (Gibco) ont été infectées par les baculovirus codant pour les différentes constructions de 6his-SmPRMT3-fl et GST-SmrpS2. Les cellules infectées ont été cultivées pendant 48 heures à 27 °C dans des plaques 6 puits. Les cellules ont été récoltées par centrifugation puis lavées au PBS + 10 % glycérol, avant d'être lysées par sonication dans 500 µL de 50 mM Tris-HCl pH 8, 250 mM NaCl, 5 % glycérol, 1 mM TCEP, 0,01 % NP40. Les lysats ont été clarifiés par une centrifugation de 10 minutes à 20000 g à 4 °C, puis incubés avec 40 µL de résine glutathion-sépharose 4B (GE Healthcare®) sous agitation douce à 4 °C. La résine a été lavée avec 1 mL de tampon de lyse puis analysée par SDS-PAGE.

c - Expression en système baculovirus-cellules d'insecte

800.10⁶ cellules sf21 ont été infectées dans 1 litre de milieu SF900II (Gibco) et incubées à 27 °C, 110 rpm pendant 48 heures. Les cellules ont été sédimentées par centrifugation 20 minutes à 360 g puis lavées dans 30 mL de PBS+10 % glycérol froid. Les culots ont été stockés à -20 °C.

d - Purification du complexe SmPRMT3-fl-SmrpS2 (fl/81-237)

Les cellules sf21 co-exprimant les protéines 6his-SmPRMT3-fl et GST-SmrpS2 (81-237) ont été lysées par sonication dans 50 mM Tris-HCl pH 8, 500 mM NaCl, 5 % glycérol, 5 mM TCEP, 0,01 % NP40 supplémenté par des antiprotéases et de la RNase A (10 µg.mL⁻¹). Le lysat a été clarifié par centrifugation puis mis en contact avec de la résine d'affinité pour la glutathion-S-transférase (GSH-sépharose 4B, GE Healthcare®) sous agitation douce à 4 °C. La résine est lavée avec 30 volumes de tampon de lyse puis l'échantillon est élué par digestion à la protéase 3C, 15 heures à 4 °C sous agitation douce. Les éluats contenant la protéine ont été concentrés sur membrane puis (Amicon® 30 kDa) puis purifiés par filtration sur gel (colonne Superdex® 200 10/300 GL) dans 20 mM Tris pH 8, 100 mM NaCl, 5 mM TCEP.

5 - Caractérisation biophysique de SmPRMT3-fl et -C3

Nous avons réalisé les expériences de DLS, SAXS et DSF sur les équipements mis à disposition par la plateforme de bio-cristallographie de l'IGBMC, rattachée à l'infrastructure INSTRUCT. Les expériences de spectrométrie de masse ont été réalisées par la plateforme de protéomique de l'IGBMC et par le laboratoire de spectrométrie de masse bio-organique de Cronenbourg (LSMBO).

a - Diffusion dynamique de la lumière (DLS)

La DLS est une méthode permettant de mesurer l'homogénéité (la polydispersité) d'un échantillon à partir de son coefficient de diffusion. La diffusion dynamique de la lumière induite par les constructions SmPRMT3-fl et -C3 a été mesurée sur une solution de protéine à 5 mg.mL⁻¹ diluée dans du tampon de filtration sur gel, à 20 °C dans des Uvettes en plastique. Avant dilution, les échantillons ont été centrifugés pendant 20 minutes à 15000 g à 4 °C afin d'éliminer les éventuels agrégats. Les échantillons ont été dilués dans du tampon de filtration sur gel filtré sur membrane (22 µm). Les mesures ont été effectuées à l'aide d'un appareil à DLS DynaPro NanoStar (Wyatt).

b - Diffusion des rayons-X aux petits angles (SAXS)

Les expériences de diffusion des rayons-X aux petits angles (SAXS) ont été effectuées sur un BioSAXS-1000 (Rigaku), équipé d'un détecteur par comptage de photons Pilatus 100k (Dectris). Les mesures ont été effectuées à 4 °C, 2 heures contre le tampon puis sur les échantillons de SmPRMT3 et BSA (Sigma) respectivement dilués à environ 1 et 4 mg.mL⁻¹. Avant dilution, les échantillons ont été centrifugés pendant 20 minutes à 15000 g à 4 °C afin d'éliminer les éventuels agrégats. Les échantillons ont été dilués dans du tampon de filtration sur gel filtré sur membrane (22 µm).

Les courbes échantillons et tampons ont été soustraites à l'aide du logiciel SAXSLab (Rigaku®) puis les données ont été traitées avec la suite Atsas (primus, gnom, dammin, dammif et dammaver, (Franke et al., 2009 ; Konarev et al., 2003 ; Svergun, 1992, 1999). La masse moléculaire de SmPRMT3-C3 a été estimée à partir d'un standard de BSA, et calculée à partir de la formule suivante :

$$M = \frac{I_0^{C3}}{[C3]} \cdot M^{BSA} \cdot \frac{I_0^{BSA}}{[BSA]}$$

Avec M^{BSA}=66,767 kDa, M^{C3}=42,269 kDa, I^{BSA}=1,168, I^{C3}=0,452, [BSA]=3,85 mg.mL⁻¹, [C3]=1,14 mg.L⁻¹.

c - Spectrométrie de masse

➤ **Module catalytique de SmPRMT3**

L'analyse par spectrométrie de masse en conditions dénaturantes du module catalytique de SmPRMT3 a été réalisée par la plateforme de protéomique de l'IGBMC. La mesure a été effectuée à l'aide d'un spectromètre de type ESI-TOF sur un échantillon de 6his-TEV-SmPRMT3-C3 digéré par la protéase TEV puis dilué à 100 µM et dialysé contre 100 µM d'acétate d'ammonium.

➤ **SmPRMT3 complète**

L'analyse par spectrométrie de masse en conditions natives de la forme complète de SmPRMT3 a été réalisée par le LSMBO. La mesure a été effectuée à l'aide d'un spectromètre de type ESI-TOF sur un échantillon de 6his-Trx-thr-SmPRMT3-fl digéré par la thrombine puis dilué à 10 µM et dialysé contre 100 µM d'acétate d'ammonium.

d - Fluorimétrie par balayage différentiel (DSF)

Les mesures de DSF ont été effectuées sur un Prometheus NT.48 (Nanotemper). SmPRMT3-C3 a été diluée à 0,5 mg.mL⁻¹ dans le tampon de filtration sur gel correspondant, avec ou sans ligand à 1 mM. La puissance du laser d'excitation a été ajustée de manière à obtenir entre 5000 et 20000 coups. Après équilibration à 20 °C, la température a été augmentée jusqu'à 95 °C, à 1 °C par minute. Les intensités de fluorescence à 330 et 350 nm ont été mesurées en temps réel.

6 - Essais de cristallisation de SmPRMT3

Les essais de cristallisation ont été réalisés sur les constructions de SmPRMT3 complète et tronquées (C1 à C4), avec ou sans les tags 6his et 6his-Trx. Lorsque la protéine utilisée pour les essais arborait encore une étiquette en fin de purification, la protéase correspondante a été ajoutée dans les gouttes de cristallisation.

a - Criblage de conditions de cristallisation en plaque de 96 puits

Le criblage de conditions de cristallisation a été effectué à l'aide du robot de cristallisation Mosquito (Tpp Labtech®). Les gouttes de cristallisation ont été préparées sous atmosphère humide (~80 %) selon la méthode de la goutte assise, dans des plaques de type MRC 96 puits. Les gouttes ont été formées par le mélange 1:1 de 100 nl de solution protéique et 100 nl de solution de précipitant. Le réservoir de précipitant contient un volume de 50 µL. La solution protéique standard est constituée de SmPRMT3 diluée dans le tampon de filtration sur gel correspondant, supplémenté par 1 mM de SAH ou SFG. Les plaques de cristallisation ont été stockées à 20 °C dans des armoires RockImager® (Formulatrix) opérant des prises de vues régulières, toutes les 24 heures jusqu'à 3 mois. Des cribles commerciaux et maison ont également été utilisés.

b - Additifs

Le criblage d'additifs a été réalisé à l'aide du robot de cristallisation Mosquito (Tpp Labtech®) en plaques de type MRC 96 puits. Les additifs ont été ajoutés dans les puits à un ratio 1:10 par rapport aux solutions de précipitant. Les gouttes ont été formées tel qu'indiqué ci-dessus.

c - Cristallisation en plaque de 24 puits

SmPRMT3 a été cristallisé en complexe avec du SAH ou de la SFG selon la méthode de diffusion de vapeur en goutte suspendue, dans une plaque de cristallisation VDX 24 puits (Hampton™ research). Les gouttes de cristallisation ont été constituées sur des lamelles de verre 22 mm traitées à la silicone (HR3-231, Hampton® research) par le mélange de 1 µL de solution de protéine à 6 mg.mL⁻¹, avec une concentration en ligand de 1 mM à 2 mM, et 1 µL de précipitant, constituée classiquement d'un gradient de 6 % à 11 % en PEG5000MME, 100 mM BTP pH 9, 100 mM de NaCl. Les réservoirs de la plaque de cristallisation ont été remplis avec 0,5 mL de précipitant puis les lamelles ont été scellées sur la plaque à l'aide d'huile de paraffine (Merk™). Les cristaux ont été obtenus après une semaine à l'obscurité, 22 °C. Les précipitant, les ratios de conception des gouttes et la température de croissance ont pu varier lors des essais d'optimisation.

d - Ensemencement

Les germes ont été produits à partir de petits cristaux de 6his-SmPRMT3-C3 obtenus dans 9 % PEG5000MME, 100 mM BTP pH 9, 100 mM NaCl. Les cristaux ont été broyés dans la goutte à l'aide d'une tige en verre puis suspendus dans 25 volumes de solution du réservoir. Les fragments ont été à nouveau broyés 4 fois 30 secondes dans un tube à essai contenant des billes métalliques par l'action d'un vortex. Entre chaque action, la suspension de germes est préservée sur glace.

L'ensemencement sur plaque de 96 puits a été effectuée par l'ajout de 30 nL de suspension de germes à 70 nL de solution de précipitant avant mélange précipitant/protéine. L'ensemencement

sur plaque 96 puits a été réalisé par trempage successif d'un vibrisse de chat dans la suspension de germes puis dans la goutte de cristallisation.

e - Cryo-protection des cristaux

Les cristaux ont été trempés quelques secondes dans une solution similaire au réservoir, de concentration en agent cristallisant 2 % supérieure, et supplémentée avec 15 % PEG400 final. Après trempage, les cristaux ont été rapidement congelés par transfert dans l'azote liquide ou passage sous cryo-jet.

f - Collecte et traitement des données

Les données ont été collectées sur les lignes de rayonnement synchrotron ID23 de l'ESRF (Grenoble), Proxima 1 de SOLEIL (Saclay), équipées d'un détecteur Pilatus 6M (Dectris), ou sur l'installation de la plateforme de cristallographie de l'IGBMC. Le diffractomètre de la plateforme est équipé d'une source Rigaku FR-X à anode tournante et d'un détecteur Eiger 4M (Dectris). Les données de diffraction ont été traitées à l'aide du programme XDS (Kabsch, 2010) et de la suite de programmes CCP4 (Collaborative Computational Project, Number 4, 1994). Les cristaux appartiennent au groupe d'espace $H3_2$.

7 - Étude de SmPRMT3 par cryo-microscopie électronique

Les criblages et collectes en cryo-microscopie électronique ont été accomplis sur les Polara F30 et Titan Krios de la plateforme de biologie structurale de l'IGBMC (FRISBI). Le traitement des données a été effectué sur les armoires de calculs de la plateforme de microscopie électronique moléculaire de l'IGBMC.

a - Préparation des échantillons pour la microscopie électronique

➤ **Solution protéique**

La protéine 6his-SmPRMT3-fl a été diluée à 0,5 et 0,05 UA280 soit environ 7 et 0,7 μM (selon le coefficient d'extinction moléculaire théorique $69790 \text{ M}^{-1}\cdot\text{cm}^{-1}$) dans du tampon de filtration sur gel préalablement filtré (membrane 22 μm) et supplémenté avec 200 μM SFG final. Les échantillons ont été centrifugés 20 minutes à 20000 g, à 4 °C afin de sédimenter les agrégats insolubles. Le surnageant a été récupéré pour être utilisé lors de la préparation des grilles.

➤ **Préparation et stockage des grilles**

Les grilles Quantifoil® R2/2 cuivre/rhodium couvertes d'un film de carbone perforé ont été traitées par décharge de plasma dans l'air résiduel à $2\cdot 10^{-1}$ mbar, avec une charge électrique de 2 mA pendant 15 secondes à l'aide d'un système ELMO (Cordouan®). 3 μL d'échantillon ont été déposés sur les grilles dans l'enceinte d'un système de cryo-préparation Vitrobot® Mark IV, sous atmosphère

saturée en eau à 4 °C. Le surplus d'échantillon a été absorbé par buvardage, réglé sur force 10 pendant une seconde. L'échantillon a été congelé par immersion dans un bain d'éthane liquide. Les grilles ont été transférées dans des portoirs EM-tec® GB-4 et stockées dans l'azote liquide.

➤ Collectes et traitement des données

◆ *Criblage à 100 keV sur microscope électronique en transmission Tecnai Polara F30*

La qualité des grilles a été contrôlée sur le cryo-microscope électronique en transmission Tecnai Polara F30 équipé d'un canon à émission de champs et d'une caméra par comptage direct d'électrons Falcon I (FEI®). Les collectes destinées à contrôler la qualité de l'échantillon ont été réalisées avec une accélération de 100 keV et un grossissement de 93000, une dose de $10 \text{ e}^- \cdot \text{Å}^{-2} \cdot \text{s}^{-1}$ et une gamme de défocalisation de 2 μm à 3,8 μm .

◆ *Collecte à 300 keV sur microscope électronique en transmission FEI Titan Krios*

La collecte a été réalisée sur le cryo-microscope électronique en transmission Titan Krios équipé d'un canon à émission de champs, d'une caméra par comptage direct d'électrons K2 (Gatan®), d'un système de correction des aberrations sphériques et d'un plateau de phase Volta (VPP). Le microscope a été utilisé à une accélération de 300 keV et un grossissement de 100000. La K2 a été utilisée en mode super-résolution. Une dose de $1,28 \text{ e}^- \cdot \text{Å}^{-2} \cdot \text{s}^{-1}$ soit environ $50 \text{ e}^- \cdot \text{Å}^{-2}$ a été appliquée et une gamme de défocalisations de 0,3 μm à 0,5 μm a été utilisée. La collecte a été réalisée avec VPP (Volta Phase Plate).

◆ *Traitement des données*

Les micrographies « vides » et fortement contaminées ont été éliminées manuellement. Pour les « films » issus de l'enregistrement avec la K2, la normalisation des micrographies contre le gain des pixels de la caméra a été effectuée avec la commande CLIP du programme IMOD (Kremer et al., 1996) et les mouvements et dommages radiatifs de l'échantillon ont été corrigés avec MotionCorr2 (Zheng et al., 2017). Les paramètres de la fonction de transfert de contraste (CTF) ont été déterminés à l'aide du logiciel Gctf-1.06 (Zhang, 2016). Les micrographies possédant des paramètres de CTF aberrants ont été éliminées grâce au script python StarFileToolbox. Les particules ont été sélectionnées à l'aide du programme Gautomatch par correspondance avec des classes 2D générées à partir de particules sélectionnées manuellement. Les étapes d'extraction et classification ont été effectuées dans Relion-2.1 (Kimanius et al., 2016). Le script « StarFileToolbox » permet de lire les fichiers de métadonnées star utilisés en microscopie électronique, d'afficher des graphiques de différentes statistiques et paramètres associés au jeu de données, de classer et trier les micrographies/particules, de modifier et générer des fichiers de métadonnées star et de déplacer les fichiers images selon des paramètres recherchés.

III - Étude structurale et fonctionnelle de PRMT2

A - PRMT2

PRMT2 est une des PRMT de type I les moins caractérisées. Elle est connue pour son rôle de coactivateur des récepteurs nucléaires et son implication dans la régulation de l'épissage, mais jusqu'en 2017 sa structure atomique n'était pas connue et aucun substrat n'avait été identifié. Au sein de l'équipe, au cours de cette thèse, nous avons résolu les structures cristallographiques à haute résolution des PRMT2 de souris et de poisson zèbre. Ce travail a donné lieu à une publication (Cura et al., 2017), jointe dans la section « Articles joints et communications scientifiques ». Ce chapitre présentera des résultats complémentaires, en particulier sur l'étude du complexe PRMT2-rsf1, pour lesquels une publication est en préparation.

B - Résultats et discussion

1 - Étude structurale et fonctionnelle de PRMT2

Dans le but de maximiser nos chances de résoudre la structure cristallographique de PRMT2, nous avons réalisé le clonage, l'expression et la purification de PRMT2 de souris (mPRMT2) et de son homologue chez *Danio rerio* (zPRMT2). En effet, quelques variations dans la séquence d'acides aminés entre deux homologues peuvent avoir une influence considérable et non prédictible sur son comportement lors des étapes d'expression, de purification et de cristallisation. mPRMT2 et zPRMT2 sont constitués respectivement de 445 et 408 acides aminés et partagent une identité de séquence de 55 % environ, et 59 % en restreignant l'analyse au module catalytique.

Des constructions entières de mPRMT2 et zPRMT2, ainsi qu'une construction de zPRMT2 débarrassée du domaine SH3 (73-408) ont été clonées. Après près de 2800 essais de cristallisation sur les trois constructions, nous sommes parvenus à résoudre quatre structures cristallographiques à haute résolution de mPRMT2 (PDB 5FLU 1,89 Å et 5FWA 1,8 Å) en complexe avec du SAH et avec le premier inhibiteur décrit de PRMT2 (le composé U1/CP1), ainsi que du domaine catalytique de zPRMT2 (PDB 5G02 2,45 Å et 5FUB 2,00 Å) en complexe avec du SAH et de la SFG. Les cristaux de zPRMT2 appartiennent au groupe d'espace $H3_2$ et ceux de mPRMT2 au groupe d'espace $C222_1$.

Les structures cristallographiques montrent que mPRMT2 et zPRMT2 forment un dimère classique de PRMT de type I. La structure du module catalytique est très conservée et les modules catalytiques de mPRMT2 et zPRMT2 se superposent avec une RMSD (Root Mean Square Deviation) de seulement 0,68 Å. Dans mPRMT2, le site de fixation de l'arginine substrat est occupé par un ion calcium hydraté par six molécules d'eau qui forment des liaisons hydrogène avec les

deux acides glutamiques catalytiques du motif III et avec l'histidine conservée du motif IV. Le calcium est ici un additif nécessaire à la cristallisation de mPRMT2 et nous avons montré par des tests d'activités que cet ion inhibait la méthylation par PRMT2. L'activité enzymatique est totalement restaurée par chélation du calcium par de l'EDTA. Dans le site actif de zPRMT2, le site calcium est occupée par une molécule de MES (acide MorpholinoEthanoSulfonique), utilisée comme tampon dans la cristallisation de l'enzyme.

Les structures cristallographiques de mPRMT2 ont été obtenues à partir de l'enzyme entière. Pourtant, le domaine amino-terminal SH3 n'est pas observé dans la carte de densité électronique, mais les expériences sur les cristaux obtenus ont montré que ce domaine était présent dans la structure cristallisée. Cela suggère que ce domaine est flexible avec un comportement identique déjà observé pour le domaine N-terminal de CARM1 (Troffer-Charlier et al., 2007). La fonction exacte de ce domaine n'est pas bien caractérisée, mais nous avons démontré que sa présence stimule l'activité enzymatique de PRMT2, suggérant un rôle dans la fixation du substrat.

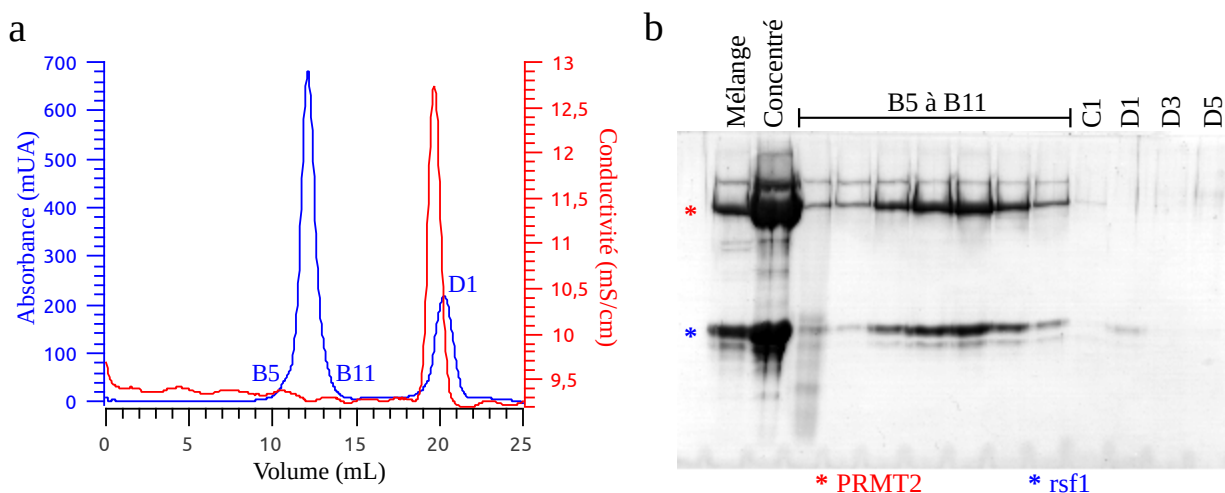
Lors de cette étude, nous avons également identifié le facteur d'épissage rsf1 comme un partenaire de PRMT2. PRMT2 et rsf1 forment un complexe stable et homogène en solution, et le facteur d'épissage est méthylé par PRMT2 sur plusieurs sites.

Tous ces résultats sont détaillés dans l'article que nous avons publié dans Febs J. en 2017 « Structural studies of protein arginine methyltransferase 2 reveal its interactions with potential substrates and inhibitors ». Cet article est joint à ce manuscrit de thèse. En complément de ce travail, nous avons résolu la structure de PRMT2 avec 2 autres inhibiteurs (composés U2 et U3, PDB 5FWD et 5JMQ) et dont les résultats seront présentés dans le chapitre VI.

2 - Reconstitution et caractérisation du complexe PRMT2-rsf1

Peu d'informations biochimiques et structurales sont disponibles sur la stœchiométrie des complexes incorporant des PRMT. Nous avons montré par spectrométrie de masse en conditions natives que PRMT6 fixe deux cofacteurs mais un seul peptide de H4 en solution (Bonfond et al., 2015). À l'inverse, nos structures cristallographiques de mCARM1 en complexe avec des dérivés de peptides PABP1 (van Haren et al., 2017) et les structures PRMT5 en complexe avec MEP50 et un peptide de H4 (Antonysamy et al., 2012) montrent que chaque monomère fixe un substrat. Pour mCARM1, nous suspectons néanmoins que l'occupation d'un dimère par deux peptides puisse être un artefact cristallographique. Afin de mieux caractériser le complexe formé entre PRMT2 et rsf1, nous avons souhaité déterminer sa stœchiométrie et son homogénéité. Rsf1 a été ajouté en excès à un échantillon de PRMT2 en présence de SFG, puis le mélange a été analysé par chromatographie de filtration sur gel.

La figure 61 présente les résultats de purification du complexe PRMT2-rsf1 reconstitué. Sur le chromatogramme, on observe une élution dans deux pics : un premier composant du mélange est élué sous un volume correspondant au domaine de fractionnement de la résine (fractions B5 à B11) puis un second constituant, de petite masse moléculaire, est élué en fin de chromatographie. L'analyse par électrophorèse de ces deux pics d'élutions montre que l'échantillon élué entre les fractions B5 et B11 est constitué de complexe PRMT2-rsf1 co-purifiant de façon stable. La symétrie du profil d'élution suggère que le complexe est homogène. L'échantillon qui élue en fin de chromatographie (fraction D1) correspond à l'excès de rsf1 non fixé par PRMT2, efficacement séparé du complexe par la colonne de filtration sur gel. Les échantillons de complexe ont été analysés par spectrométrie de masse native par les Dr. Sarah Cianferani et Oscar Hernandez Alba du LSMBO.



Le spectre de masse présenté en figure 62 indique la présence d'un mélange constitué de rsf1 méthylée libre, de monomère et de dimère de PRMT2 et de complexe PRMT2-rsf1. La masse mesurée pour ce complexe (environ 120 kDa) correspond à la masse de deux monomères de PRMT2 (environ 51 kDa) et d'un monomère de rsf1 (environ 18 kDa). Dans les conditions expérimentales utilisées, le complexe PRMT2-rsf1 se comporte donc comme un complexe hétérotrimérique en solution, tel qu'observé pour PRMT6. L'observation de monomère de PRMT2 est un artefact expérimental toujours observé en spectrométrie de masse. Les masses mesurées indiquent également que les dimères de PRMT2 et le complexe ne contiennent pas de SFG bien que celle-ci ai été ajoutée lors de la mise en contact des deux partenaires. Cela indique que l'inhibiteur a été éliminé durant l'étape de purification du complexe et que le domaine de fixation du SAM pourrait ne pas être correctement replié. Cette analyse semble donc indiquer que l'échantillon est moins homogène qu'il n'apparaît sur le chromatographe. La formation et la dissociation du complexe PRMT2-rsf1 pourrait être très dynamique, si bien que les deux partenaires co-purifie car ils sont majoritairement associés mais que des formes dissociées soient détectables en spectrométrie de masse.

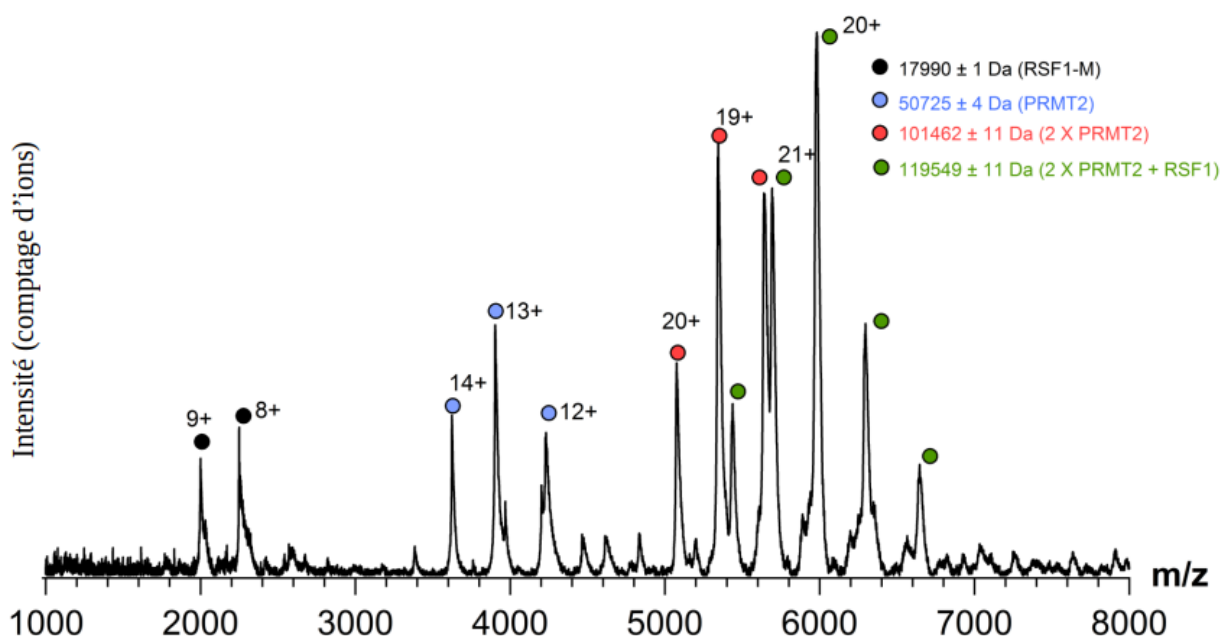


Figure 62 - Analyse par spectrométrie de masse du complexe PRMT2 – rsf1.

Analyse par spectrométrie de masse ESI-TOF en conditions non dénaturantes de l'échantillon de complexe PRMT2-rsf1 reconstitué. Les cercles noirs correspondent à rsf1 méthylé, les cercles bleus correspondent au monomère de PRMT2, les cercles rouges sont des dimères de PRMT2 et les cercles verts représente les complexes PRMT2-rsf1.

3 - Étude cristallographique du complexe PRMT2-rsf1

Nous avons procédé à des essais de cristallisation du complexe PRMT2-rsf1, dans les conditions usuelles de cristallisation de PRMT2 de souris détaillées dans l'article ci-joint et par recherche *ab initio* des conditions en utilisant des cribles commerciaux. Malheureusement, les seuls cristaux obtenus appartiennent au groupe d'espace $C222_1$ de paramètres de maille $a = 66\text{\AA}$ $b = 115\text{\AA}$

$c = 133\text{\AA}$ $\alpha = 90^\circ$ $\beta = 90^\circ$ $\gamma = 90^\circ$, correspondant aux cristaux de mPRMT2 seul. La densité électronique montre que rsf1 est absent des cristaux. Les conditions de cristallisations contiennent 100 mM de CaCl_2 . L'ion Ca^{2+} est observé fixé en plusieurs points de la structure, et notamment dans le site de fixation de l'arginine où il est complexé par les deux acides glutamiques catalytiques et l'histidine du motif IV. Cet ion pourrait nuire à la fixation d'une arginine de rsf1 dans la poche et donc à l'incorporation de rsf1 lors de la formation des cristaux. Il semble donc que dans toutes les conditions que nous avons testées, c'est la cristallisation de mPRMT2 seul qui a dominé, le partenaire étant chassé lors du processus de cristallisation. Pour pallier ces problèmes, de nouvelles conditions de cristallisation ont été explorées en changeant de molécules tampon et sans utiliser le CaCl_2 . Des cristaux ont été obtenus mais à nouveaux ceux-ci ne contenaient pas le partenaire rsf1.

4 - Étude du complexe PRMT2-rsf1 par cryo-microscopie électronique

Devant la difficulté de cristalliser le complexe PRMT2-rsf1, nous nous sommes tournés vers la cryo-microscopie électronique. Les progrès de la technologie et les travaux menés sur SmPRMT3 ont démontré que l'étude par cryo-microscopie électronique des PRMT est techniquement envisageable. La masse moléculaire du dimère de SmPRMT3 est d'environ 130 kDa contre à peu près 120 kDa pour le complexe PRMT2-rsf1. Les étapes de décharge de plasma et de vitrification des premiers échantillons ont été réalisées de manière similaire à celles utilisées pour SmPRMT3. Une grille de type Quantifoil R22 Cu/Rh couverte d'un film perforé de carbone a été soumise à une décharge de plasma pendant 15 secondes, à une charge électrique de 2 mAh puis l'échantillon a été déposé sur grille. L'excès d'échantillon a été enlevé par buvardage sous une atmosphère saturée en eau, à 4 °C. La protéine directement issue d'une fraction de purification par filtration sur gel a été utilisée, sans étape finale de concentration qui pourrait engendrer la formation d'agrégats.

Les premières visualisations de l'échantillon, dont un exemple est présenté en figure 63a, réalisées sur un microscope Tecnai Polara F30 utilisé à 100 keV et un grossissement de 59000, ont montré que le bruit de fond était trop élevé pour permettre la visualisation de la protéine. De nouveaux échantillons ont été préparés pour lesquels la concentration en NaCl a été abaissée à 75 mM comparé aux 100 mM initiaux, et la molécule tampon Tris a été remplacée par l'HEPES. L'analyse des nouvelles grilles, dont un exemple est présenté en figure 63b, a montré une amélioration significative du bruit de fond. Les particules sont réparties sur l'ensemble du champ d'observation et ne semblent pas s'agréger. Les objets observés mesurent 9 nm à 12 nm et certaines vues montrent un objet pouvant ressembler à un dimère de PRMT, comme illustré sur la figure 63b dans l'encadré rouge.

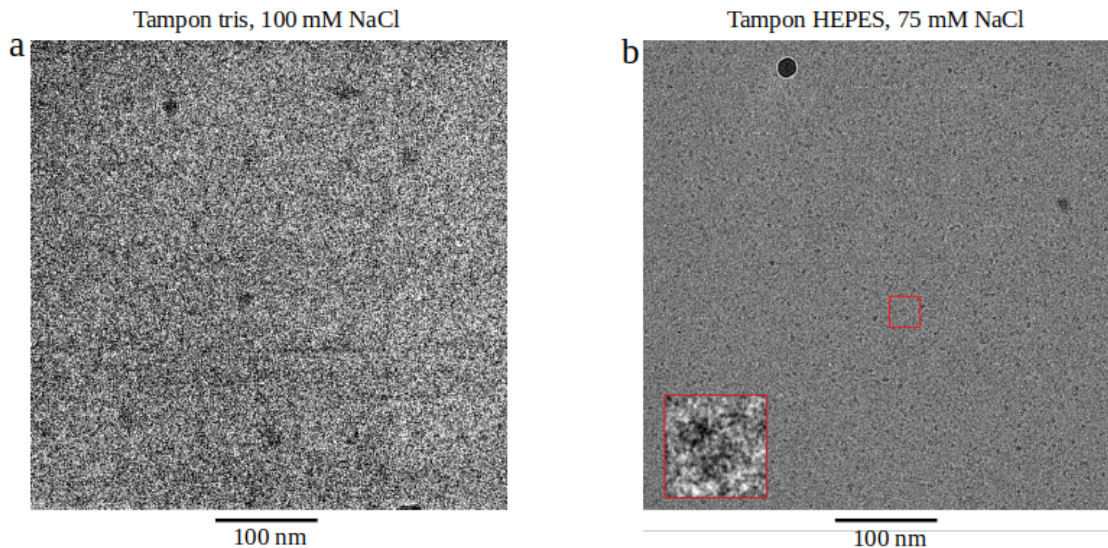


Figure 63 - Criblage des conditions de préparation des grilles pour le complexe PRMT2-rsf1.

(a) Micrographie de la première grille de complexe PRMT2-rsf1 préparée à partir de l'échantillon contenu dans un tampon 20 mM Tris-HCl pH 8, 100 mM NaCl. (b) Micrographie d'une grille de complexe PRMT2-rsf1 préparée à partir de l'échantillon contenu dans un tampon 20 mM HEPES pH 7,5, 75 mM NaCl. Les acquisitions ont été faites à 100 keV et un grossissement de 59000. L'encadré rouge dans (b) présente une particule de diamètre 9 nm ressemblant à un dimère de PRMT2 vu de face.

Une première collecte a été réalisée sur un cryo-microscope en transmission FEI Titan Krios. Étant donné la faible masse moléculaire de la particule, un plateau de phase volta a été utilisé pour amplifier le contraste à basse résolution et à faible défocalisation. Les données ont été enregistrées à une accélération de 300 keV, une dose de $50 \text{ e}^- \cdot \text{Å}^{-2}$, un grossissement de 130000 x et une gamme de défocalisations comprise entre $0,3 \mu\text{m}$ et $0,5 \mu\text{m}$. Un jeu de données de 2395 micrographies a été collecté. La figure 64 présente un exemple caractéristique de micrographie et du calcul de sa CTF. On peut observer que l'échantillon est homogène, sans agrégats protéiques et à une concentration optimale où les particules occupent la totalité du champ d'observation mais ne se chevauchent pas. Cependant comme le montre le spectre de puissance, les données sont un peu astigmatiques et les anneaux de Thon sont pratiquement invisibles, même aux petites fréquences spatiales donc à basse résolution. Cela provient probablement de la préparation de l'échantillon.

Lorsque nous avons observé la grille à un plus faible grossissement afin de choisir les régions à sélectionner pour la collecte, nous avons observé la formation de gouttes d'échantillon au centre des carreaux, comparable à l'illustration sur la figure 65b. Ce phénomène est caractéristique d'une étape de décharge de plasma et/ou de buvardage insuffisant. Le résultat est une glace trop épaisse qui a dégradé le rapport signal/bruit, donc le contraste, ce qui explique que l'on n'observe peu ou pas d'anneaux de Thon dans les spectres de puissance du jeu de données. Ce faible niveau de contraste a eu deux effets lors du traitement du jeu de données. Premièrement, les détails des particules à haute résolution n'ont pas pu être extraits. Ensuite, les paramètres de la CTF n'ont pas pu être déterminés avec précision, si bien que les images n'ont pas pu être corrigées efficacement.

Nous avons toutefois réalisé le traitement de ces données dans Relion2 et cryoSPARC jusqu'à obtention des classes 2D et du modèle présenté en figure 64.

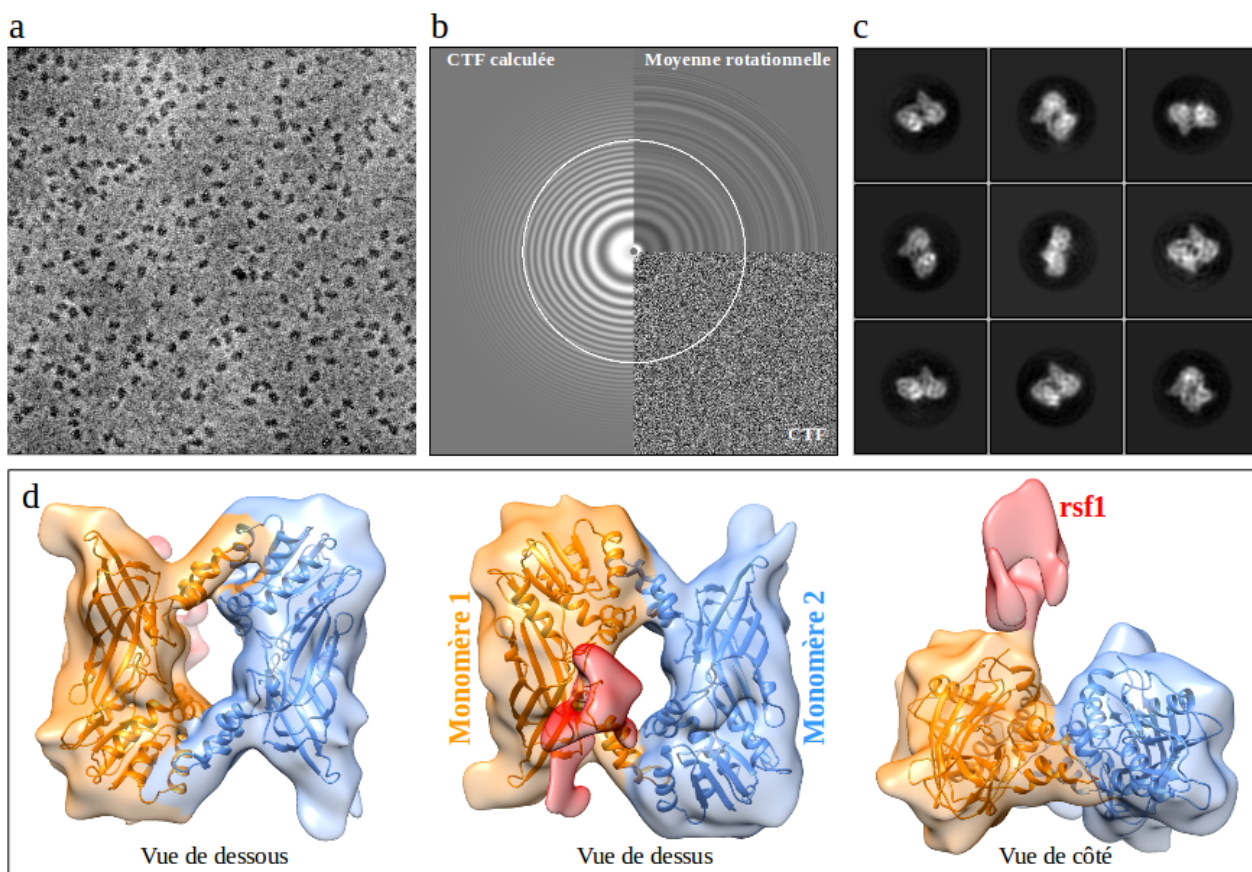


Figure 64 - Structure à basse résolution du complexe PRMT2-rsf1.

(a) Micrographie électronique du complexe PRMT2-rsf1. (b) Calcul des paramètres de la CTF par Gctf 1.06. À gauche : modèle de CTF calculé à partir des données expérimentales. En haut à droite : moyenne rotationnelle calculée par la rotation pondérée du spectre de puissance d'une micrographie autour de son centre. En bas à droite : spectre de puissance d'une micrographie. (c) Exemple des classes obtenues pour le premier jeu de données. (d) Modèle du complexe PRMT2-rsf1 calculé et affiné dans cryoSPARC jusqu'à 7,5 Å selon la « Gold-standard Fourier-Shell correlation ». Les densités correspondant aux monomères de PRMT2 sont représentées en bleu et orange. La densité rouge correspond à rsf1 et au domaine SH3. La structure cristallographique de mPRMT2 (PDB 5FUL), représentée en rubans, a été placée dans la carte de densité.

Dans Relion, les classes obtenues montrent un dimère de PRMT2 semblable au dimère cristallographique, et contrairement à SmPRMT3, l'échantillon de PRMT2 ne semble pas adopter d'orientations préférentielles. À ce stade, les classes laissent difficilement entrevoir la présence ou l'absence de rsf1. Deux modèles *ab initio* ont été calculés dans cryoSPARC à partir de 196000 particules extraites des meilleures classes. Une étape de classification 3D a été réalisée afin de trier les particules en fonction de leur ressemblance à l'un ou l'autre des deux modèles *ab initio*. Enfin, le meilleur modèle a été affiné avec les particules qui lui ont été attribuées par l'étape de classification 3D. Le modèle (figure 64) a été affiné jusqu'à une résolution maximale de 7,5 Å, probablement limitée par le faible contraste précédemment décrit.

Le modèle est constitué d'un dimère de PRMT2 adoptant la forme classique également observée dans les structures cristallographiques. L'affinement a fait apparaître une densité

supplémentaire à la verticale d'un seul des monomères, représentée en rouge sur la figure. Cette densité se situe dans la continuité de l'hélice α -X de PRMT2. Elle est trop volumineuse pour correspondre au seul domaine amino-terminal SH3 de PRMT2 et elle pourrait correspondre au partenaire rsf1 interagissant avec le domaine SH3. Ce premier modèle, bien que limité en résolution, confirme l'existence du complexe PRMT2-rsf1 en solution et montre, par son asymétrie, que le dimère de PRMT2 ne fixe qu'un seul rsf1. De plus, le positionnement de rsf1 dans la continuité de l'hélice α -X de PRMT2 montre que l'interaction entre PRMT2 et rsf1 est médiée par le domaine SH3 de PRMT2. Néanmoins, la densité correspondant au rsf1 est incomplète et semble former plusieurs lobes, ce qui pourrait indiquer une certaine flexibilité de cette région.

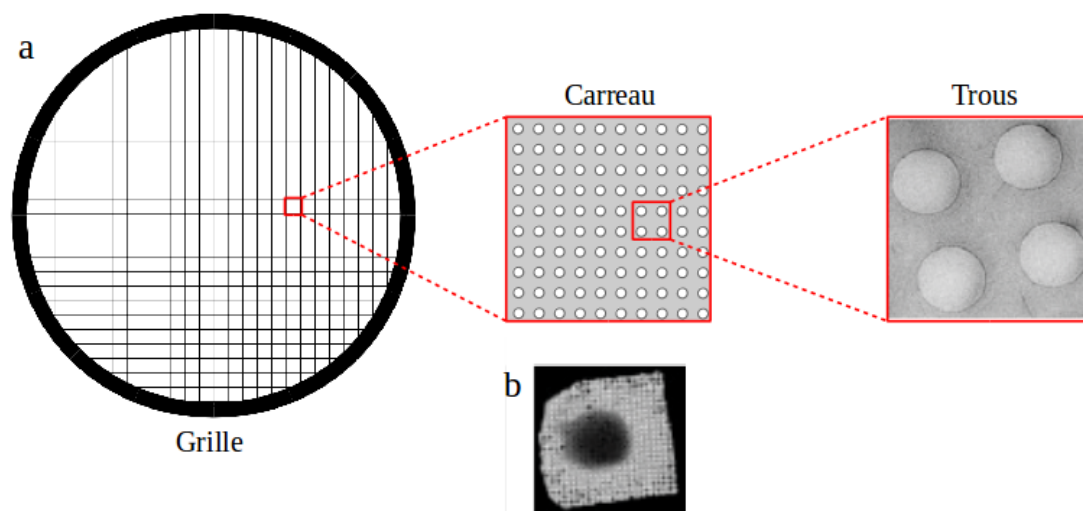


Figure 65 - Schéma d'organisation d'une grille de cryo-EM perforée de type R22.

(a) Schéma d'une grille de microscopie électronique. La grille est constituée d'une couche de cuivre et d'une couche de rhodium. La couche de cuivre forme un maillage de carreaux (300 carreaux pour les grilles Quantifoil Cu/Rh R22 utilisées). Le maillage sert de support à un mince film de carbone perforé (trous de 2 μ m espacés de 2 μ m pour les grilles utilisées). (b) Micrographie d'un carreau représentant le phénomène de goutte observé lors de la première collecte.

À partir des leçons tirées de ce premier essai, nous avons décidé d'agir sur les étapes de préparation de l'échantillon afin d'améliorer le contraste lors de la prochaine collecte de données. Plusieurs paramètres ont été modifiés. Premièrement, les échantillons sont encore issus d'une fraction de filtration sur gel, mais une fraction légèrement plus concentrée a été choisie. En augmentant la concentration en particules, nous espérons augmenter la part des électrons diffusés, donc le signal. Ensuite, l'intensité et la durée de la décharge de plasma ont été augmentées de 15 secondes et 2 mAh à 20 secondes et 2,5 mAh, afin d'accroître l'hydrophilie de la grille et donc d'améliorer la répartition de l'échantillon sur celle-ci. Finalement, la force appliquée sur la grille par les papiers filtres lors de l'étape de buvardage a également été augmentée, et cette étape est réalisée dans une atmosphère saturée en vapeur d'eau. De plus la quantité d'échantillon déposé sur la grille a été augmentée de 2,5 μ L à 3 μ L. Ces changements dans le mode opératoire de préparation des échantillons ont pour objectif d'amincir l'épaisseur de glace afin de diminuer le bruit de fond.

Ainsi, en augmentant la concentration en particules et en réduisant la participation du milieu à la diffusion des électrons, nous espérons améliorer le contraste.

Une nouvelle acquisition a été réalisée très récemment sur les nouvelles grilles, avec le Titan Krios utilisé à 300 keV pour un grossissement de 130000 et avec plateau de phase volta. Cette fois, la gamme de défocalisation ciblée a été augmentée à 0,4-0,6 μm dans le but d'améliorer encore le contraste. Un jeu de données de 8789 micrographies a été collecté. La qualité de la glace a été significativement améliorée puisque les gouttes observées lors de la première acquisition ne sont plus présentes, et qu'une glace fine a pu être obtenue sur de nombreux carreaux.

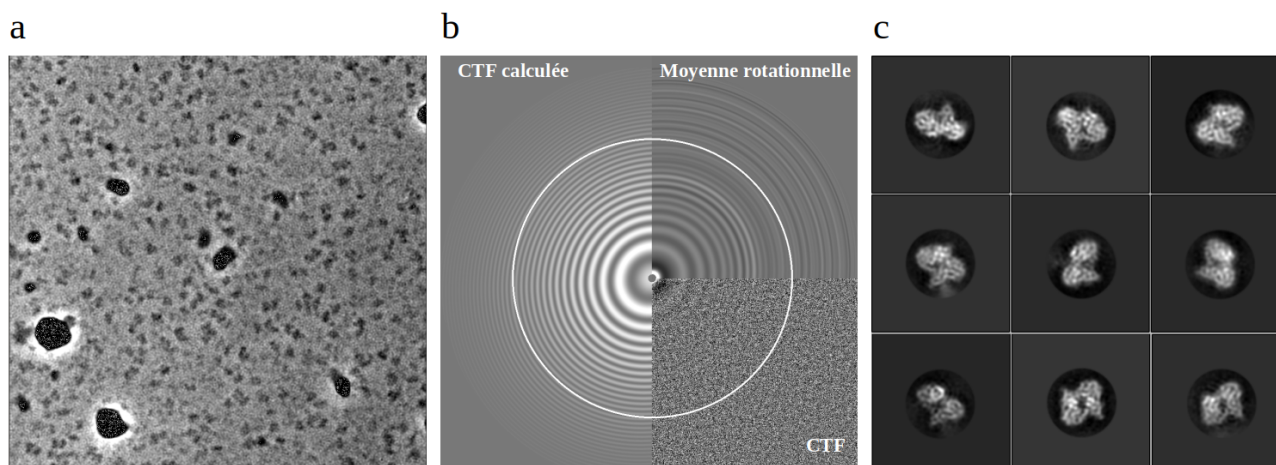


Figure 66 - Second échantillon du complexe PRMT2 – rsf1.

(a) Micrographie électronique du complexe PRMT2-rsf1. (b) Calcul des paramètres de la CTF par Gctf 1.06. A gauche : modèle de CTF calculé à partir des données expérimentales. En haut à droite : moyenne rotationnelle calculée par la rotation pondérée du spectre de puissance d'une micrographie autour de son centre. En bas à droite : spectre de puissance d'une micrographie. (c) Exemple de classes obtenues pour le second jeu de données.

La figure 66 présente les premiers résultats de collecte et de traitement des données pour cette seconde série d'acquisitions. Comme le montre la micrographie, la concentration en protéine est plus élevée sans que la protéine ne s'agrège. Toutefois, beaucoup de particules sont adjacentes ce qui peut compliquer les étapes de sélection et de classification. Une concentration intermédiaire entre cet échantillon et le précédent devrait convenir à l'avenir. L'amélioration de la glace est visible par son aspect moins granulaire comparée à celle présentée sur la figure 64 pour la première collecte, et on peut cette fois observer plus nettement les anneaux de Thon sur le spectre de puissance. Cela démontre que les changements dans le protocole de préparation des grilles ont permis d'améliorer le rapport signal/bruit comparé à la première collecte. L'intensité des anneaux demeurent néanmoins assez faible ce qui suggère que l'échantillon peut encore être amélioré. Les premières classes 2D présentées sur la figure 66 ont été calculées à partir de 357000 particules et laissent cette fois entrevoir des détails des structures secondaires. Ce plus haut niveau de détails est permis à la fois par l'utilisation d'un plus grand nombre de particules par classe et par une correction plus efficace des images permise par la détermination des paramètres de la CTF. Nous

espérons avec ce dernier jeu de données repousser la limite de résolution de 8 Å du premier modèle, et faire apparaître des détails de l'interaction établie entre PRMT2 et rsf1.

Des étapes supplémentaires seront certainement nécessaires avant d'obtenir une structure à haute résolution du complexe PRMT2-rsf1. Les anneaux de Thon sont discernables sur le nouveau jeu de données mais leurs amplitudes sont toujours faibles. Il serait risqué d'augmenter encore la concentration en complexe car les particules sont déjà pratiquement en contact et de nombreux oligomères de PRMT2 ont été observés dans le dernier jeu de données. Il sera donc nécessaire de travailler encore à l'amélioration de la glace, par exemple en utilisant d'autres types de grilles, en changeant le traitement des grilles ou en modifiant la composition du tampon, par exemple en réduisant le sel ou en changeant de type de sel. Le premier modèle a mis en évidence une certaine flexibilité du complexe PRMT2-rsf1. Cette hétérogénéité conformationnelle est très probablement une propriété intrinsèque du complexe mais devra être réduite afin de pouvoir augmenter la résolution locale dans les zones d'interaction entre PRMT2 et rsf1.

Plusieurs facteurs peuvent expliquer la flexibilité conformationnelle du complexe. Premièrement, dans la structure cristallographique de PRMT2 de souris on n'observe pas le domaine SH3, ce qui indique que celui-ci est flexible ou désordonné en absence d'un partenaire. Pour les PRMT, la fixation de SAM/SAH ou de SFG dans la poche de fixation du SAM est généralement un prérequis au repliement de l'hélice α -X et donc à la structuration locale du site catalytique. Les structures cristallographiques des modules catalytiques de CARM1 et PRMT5 en complexe avec des peptides et du SFG suggèrent également que ce ligand stabiliserait la fixation de l'arginine substrat dans le site de fixation de l'arginine par le biais d'interactions directes entre l'azote de la sinefungine et le groupement guanidinium de l'arginine substrat (Antonysamy et al., 2012 ; Boriack-Sjodin et al., 2016). L'analyse par spectrométrie de masse en conditions natives du complexe mPRMT2-rsf1 semble montrer que la sinefungine est éliminée du site de fixation du SAM lors de la purification du complexe. Ainsi, un doute subsiste sur la présence de sinefungine dans les échantillons utilisés pour la cryo-microscopie électronique et il est possible que les hélices α -X et α -Y ne soit par parfaitement structurées. Le ligand sera ajouté dans les prochains échantillons après purification. De plus, l'ajout de sinefungine pourrait aider à stabiliser une arginine de rsf1 dans la poche de fixation comme cela a été observé pour CARM1 et PRMT5. Si cela n'est pas suffisant, nous pourrions tenter de lier rsf1 et PRMT2 à l'aide d'un agent chimique de couplage tel que le glutaraldéhyde. Cela pourrait permettre de piéger le complexe dans ses configurations spatiales préférentielles et donc de réduire l'hétérogénéité conformationnelle de l'échantillon.

C - Matériel et méthodes

L'expression de PRMT2 dans le système d'expression baculovirus-cellules d'insecte a été réalisé par Nathalie Troffer-Charlier. La purification de PRMT2 a été réalisée par Nathalie Troffer-Charlier au laboratoire. L'expression et la purification de rsf1, et l'assemblage du complexe ont été réalisés par Vincent Cura. Les analyses par spectrométrie de masse ont été réalisées par le docteur Sarah Cianferani et son post-doctorant Oscar Hernandez Alba du Laboratoire de Spectrométrie de Masse Bio-Organique (LSMBO). Les cristallisations de PRMT2 et la résolution des structures des PRMT2, les essais en cristallisation du complexe PRMT2-rsf1 ont été réalisés au sein de l'équipe.

La préparation des grilles de microscopie électronique et les collectes de données ont été réalisées sur le matériel mis à disposition par l'infrastructure FRISBI, au sein du département de biologie structurale intégrative de l'IGBMC. Le traitement des données de cryo-microscopie électronique a été effectué sur les rames de calculs de la plateforme de microscopie électronique de l'IGBMC. Sur cette problématique, j'ai particulièrement développé les travaux relatifs à la préparation, l'acquisition et le traitement des données de cryo-microscopie électronique.

1 - Expression et purification des protéines

Les parties expression et purification de PRMT2 et rsf1 sont présentées dans l'article joint en annexe (Cura et al., 2017) et ne seront pas détaillées ici.

2 - Reconstruction et purification du complexe PRMT2-rsf1

PRMT2 et rsf1 ont été mélangés à un ratio molaire 1:1 ou 1:2 en présence de 0,5 mM de SAH ou SFG. Le mélange a été incubé 12 heures à 4 °C. Pour les études structurales du complexe, le tag hexahistidine amino-terminal a été clivé par une incubation supplémentaire de 90 minutes en présence de thrombine, à température ambiante. La digestion a été stoppée sur glace. Le mélange a été concentré à 150 µL par centrifugation sur filtre centrifugeable Amicon® 30 kDa (Merk Millipore) en présence de 100 µM PMSF, puis purifié par chromatographie de filtration sur gel sur résine Superdex S200 10/300, dans 20 mM Tris-HCl pH 7,5, 100 mM NaCl, 0,5 mM DTT, 100 µM PMSF. Pour les échantillons destinés à la caractérisation et à la cristallisation, une étape supplémentaire de concentration sur membrane a été effectuée. Le complexe a été concentré à environ 3 mg.mL⁻¹. Pour les échantillons destinés à la microscopie électronique, le tampon Tris a été remplacé par de l'HEPES et la concentration du NaCl a été baissée à 75 mM.

3 - Spectrométrie de masse

Le complexe PRMT2-rsf1 a été dilué à 14 μM dans 20 mM Tris-HCl pH 7,5, 100 mM NaCl, 0,5 mM DTT puis le tampon a été échangé contre une solution de 400 mM $\text{CH}_3\text{CO}_2\text{NH}_4$ à l'aide d'une colonne de dessalage Zeba™ 7K (Thermo). Les masses des espèces présentes en solution ont été mesurées par spectrométrie de masse en conditions non dénaturantes sur un spectromètre de masse ESI-TOF.

4 - Cristallisation du complexe PRMT2-rsf1

Les essais de co-cristallisation du complexe PRMT2-rsf1 ont été réalisés dans les conditions décrites dans l'article joint en annexe (Cura et al., 2017). D'autres essais de cristallisation incluant des cribles commerciaux ont été effectués. Ils incluent les cribles 96 puits JCSG+ (Quiagen), les cribles systématiques PEG-pH, Pro Complex (Quiagen), The PEGs suite I (Quiagen).

5 - Étude du complexe PRMT2-rsf1 par cryo-microscopie électronique

a - Préparation de l'échantillon

Afin de garantir son homogénéité, le complexe PRMT2-rsf1 a été prélevé directement dans une fraction d'élué de chromatographie par filtration sur gel. L'échantillon a été centrifugé 20 minutes à 20000 g à 4 °C afin de sédimenter les possibles agrégats ou contaminants. Le surnageant a été utilisé pour préparer les grilles.

b - Préparation des grilles

➤ **Collecte 1**

Des grilles Quantifoil® R2/2 cuivre/rhodium couvertes d'un film de carbone perforé ont été traitées par décharge de plasma dans l'air résiduel à $1,8 \cdot 10^{-1}$ mbar, avec une charge électrique de 2 mA pendant 15 secondes à l'aide d'un système ELMO (Cordouan®). 2,5 μL d'échantillon ont été déposés sur les grilles dans l'enceinte d'un système de cryo-préparation Vitrobot® Mark IV, sous atmosphère saturée en eau à 4 °C. Le surplus d'échantillon a été absorbé par buvardage avec une force d'application réglée sur force 10 pendant une seconde. L'échantillon a été congelé par immersion dans un bain d'éthane liquide. Les grilles ont été transférées dans des portoirs EM-tec® GB-4 et stockées dans l'azote liquide.

➤ **Collecte 2**

Des grilles Quantifoil® R2/2 cuivre/rhodium couvertes d'un film de carbone perforé ont été traitées par décharge de plasma dans l'air résiduel à $1,8 \cdot 10^{-1}$ mbar, avec une charge électrique de 2,5 mA pendant 20 secondes à l'aide d'un système ELMO (Cordouan®). 3 μL d'échantillon ont été

déposés sur les grilles dans l'enceinte d'un système de cryo-préparation Vitrobot® Mark IV, sous atmosphère saturée en eau à 4 °C. Le surplus d'échantillon a été absorbé par buvardage avec une force d'application réglée sur force 6 pendant une seconde. L'échantillon a été congelé par immersion dans un bain d'éthane liquide. Les grilles ont été transférées dans des portoirs EM-tec® GB-4 et stockées dans l'azote liquide.

c - Criblage des grilles

L'évaluation des conditions de préparation des grilles a été réalisé sur un cryo-microscope électronique en transmission Tecnai Polara F30 équipé d'un canon à émission de champs et d'une caméra par comptage direct d'électrons Falcon I (FEI®).

d - Collecte des données

Les films ont été collectés sur un cryo-microscope électronique en transmission Titan Krios équipé d'une source d'électrons à émission de champs, d'une caméra par comptage direct d'électrons K2 (Gatan®), d'un système de correction des aberrations sphériques, d'un plateau de phase Volta (VPP) et d'un filtre à énergie Quantum GIG (Gatan). Les films ont été enregistrés à une accélération de 300 keV, avec la caméra K2 utilisée en mode super-résolution (images au format tif, taille du pixel = 0,44 Å), à un grossissement de 130000 et une dose de 50 e⁻.Å⁻². Les ouvertures des condenseurs 1 et 2 ont été réglés respectivement sur 2000 et 50. Le microscope est contrôlé par le logiciel SerialEM (Mastrorade, 2005). La première collecte a été effectuée à des défocalisations cibles de 0,3 µm et 0,5 µm. Pour la seconde collecte, les défocalisations visées sont de 0,4 µm et 0,6 µm. La position du plateau de phase a été modifiée toutes les 60 minutes environ pour éviter des décalages de phases supérieures à 0,8 π.

e - Traitement des données de microscopie électronique

➤ **Pré-traitement des données**

La normalisation des micrographies contre le gain des pixels de la caméra a été effectuée avec le programme CLIP de la suite IMOD (Kremer et al., 1996). Les mouvements et dommages radiatifs de l'échantillon ont été corrigés avec MotionCorr2 (Zheng et al., 2017). Dans MotionCorr2, les films ont été compressés deux fois (taille du pixel = 0,88 Å) et convertis au format mrc. Les paramètres de la fonction de transfert de contraste (CTF) ont été déterminés à l'aide du logiciel Gctf-1.06 (Zhang, 2016).

➤ **Tri des micrographies et traitement des données**

Les micrographies ont été triées l'aide du script python « StarFileToolbox » afin d'en séparer celles qui pourraient dégrader la qualité du jeu de données. Pour cela, les micrographies dont les scores de FOM (« figure of merit ») calculées par Gctf diffèrent significativement du score moyen du jeu de données ont été écartées. La visualisation de ces micrographies a permis de vérifier

quel la plupart contiennent des enregistrements pris sur le carbone, tandis que les quelques micrographies restantes ont subi des erreurs de communication de la camera lors de la collecte. Ensuite, les micrographies dont l'astigmatisme diverge fortement de l'astigmatisme moyen (\pm écart type) du jeu de données ont également été écartées. Les micrographies retenues ont ensuite été passées en revue une par une et celles contenant de forte contaminations, une glace de mauvaise qualité ou aucune particule, ont été rejetées. Les micrographies dont les paramètres de CTF ont été mal déterminés ont également été écartées.

À l'aide du programme Relion2 (Kimanius et al., 2016) et StarFileToolbox, le fichier star contenant les métadonnées des micrographies collectées a été classé par valeurs de défocalisations croissantes, puis un sous-ensemble de micrographies représentatives de la gamme de défocalisations a été généré pour la sélection manuelle des particules. 1000 à 2000 particules ont été sélectionnées et extraites dans des boites de deux fois la plus grande dimension de la particule, soit 180 Å pour un dimère de PRMT. Dix classes de références ont été produites. Ces références ont été utilisées pour la sélection semi-automatique des particules sur l'ensemble des jeux de données avec Gautomatch (Zhang, 2016, non publié). Les particules ont à nouveau été extraites dans des boites de 180 Å à partir des coordonnées déterminées semi-automatiquement. La taille du pixel des images a été compressée deux fois. Plusieurs cycles de classifications ont été réalisés, avec 100 classes à aligner. Afin d'accélérer les calculs lors des premiers cycles, la première itération a été réalisée sur un sous-ensemble de 100 particules par classes soit 10000 particules. Lors des premiers cycles également, la résolution et le rayon du masque ont été limités respectivement à 10 Å et 150 Å. Entre chaque classification 2D, les classes de mauvaise qualité ou correspondant à des contaminations ont été écartées et les particules peuplant les bonnes classes ont été ré-extraites et recentrées. Après la dernière classification, les particules ont été ré-extraites dans une boite de 2,5 fois la plus grande dimension de la particule, avec la taille du pixel initiale. Les modèles ont été générés ab-initio puis affinés dans cryoSPARC (Punjani et al., 2017).

IV - Étude structurale et fonctionnelle de complexes PRMT-peptides substrats

A - Introduction

Les bases structurales des interactions PRMT–protéine substrat qui gouvernent la spécificité de la réaction de méthylation demeurent aujourd’hui encore mal comprises. Il s’agit de comprendre comment une PRMT donnée reconnaît et méthyle un ou plusieurs substrats protéiques. Ces enzymes sont en effet capables de méthyler plusieurs substrats, parfois sur plusieurs sites (par exemple, CARM1 méthyle PABP1 en R455 et R460 et H3 en R17 et R26, sur des séquences très différentes).

Cela est principalement dû au manque de données structurales sur les complexes PRMT-substrats entiers. En effet, malgré un nombre croissant de structures déposées (environ 80 dont 42 résolues par notre équipe), la majorité d’entre elles concernent le module catalytique seul ou en complexe avec du cofacteur ou des inhibiteurs.

La difficulté à obtenir des informations structurales sur la reconnaissance entre les PRMT et leurs partenaires substrats comme non-substrats provient de causes multiples. Les PRMT agissent probablement dans le contexte de complexes plus larges impliquant des partenaires nécessaires aux recrutements sur les sites de méthylation et pour l’interaction avec leurs substrats. Ces complexes sont le plus souvent peu caractérisés structuralement. L’exemple le plus étudié concerne PRMT5, qui dans le contexte du méthylosome nécessite son partenaire MEP50 pour pouvoir fixer et méthyler les protéines Sm (Friesen et al., 2001a). Un autre exemple concerne la protéine p160 qui recrute CARM1 et p300 pour former un complexe de coactivation de la transcription par les récepteurs nucléaires, par méthylation des arginines et acétylation des lysines sur l’histone H3 (Chen, 1999 ; Chen et al., 2000).

Quelques modules catalytiques ont été co-cristallisés avec des petits peptides, dont une structure cristallographique de PRMT1 de rat dans laquelle seul un résidu arginine est visible (PDB 1OR8, (Xing Zhang et Cheng 2003)), une structure de PRMT5 humaine en complexe avec un peptide de l’histone H4 long de 8 résidus (PDB 4GQB, Antonysamy et al. 2012) et cinq structures de CARM1 humaine en complexe avec des peptides de PABP1 correspondant à deux sites de méthylation préférentiels (R455 et R460, 6 et 9 résidus) et de l’histone H3 (peptide long de 9 résidus autour de R17) avec divers degrés de méthylation des arginines (PDB 5DX1, 5DX8, 5DXA, 5DWQ et 5DX0, Boriack-Sjodin et al. 2016). De plus, seul PRMT5 a pu être co-cristallisée entière en complexe avec son partenaire du méthylosome MEP50 (PDB 4GQB, Antonysamy et al. 2012 ; 4G56, Ho et al. 2013).

Les travaux de l'équipe ont montré que la reconnaissance d'un peptide substrat impliquait des interactions transitoires dites à « courtes distances » (situées au voisinage du centre catalytique) et des interactions dites à « longues distances » impliquant des zones éloignées du site de méthylation (Mailliot, 2013). Dans sa thèse, Justine Mailliot a montré que le peptide naturel H3 1-41 possède une affinité ($K_d = 4 \mu\text{M}$) significativement plus élevée pour mCARM1 130-507 que le peptide naturel H3 1-31 ($K_d = 40 \mu\text{M}$).

La stratégie développée dans l'équipe pour stabiliser les complexes PRMT-partenaires a été d'utiliser une approche visant à lier de manière covalente l'arginine d'un substrat donné à une molécule dérivée du cofacteur SAM (Ajebbar, 2012 ; Mailliot, 2013), constituant ainsi des mimes de l'état de transition de la réaction de méthylation (composés appelés analogues « SAM-peptide »). Ce type de pontage covalent s'est révélé être une stratégie efficace pour capturer et caractériser des complexes enzyme-substrats peu stables.

Deux approches ont été développées, soit un pontage *in situ* soit un pontage *in vitro* (figure 67). Les intermédiaires chimiques développés lors des étapes initiales de ces travaux ont été appelés analogues du SAM et analogues « SAM-arginine ».

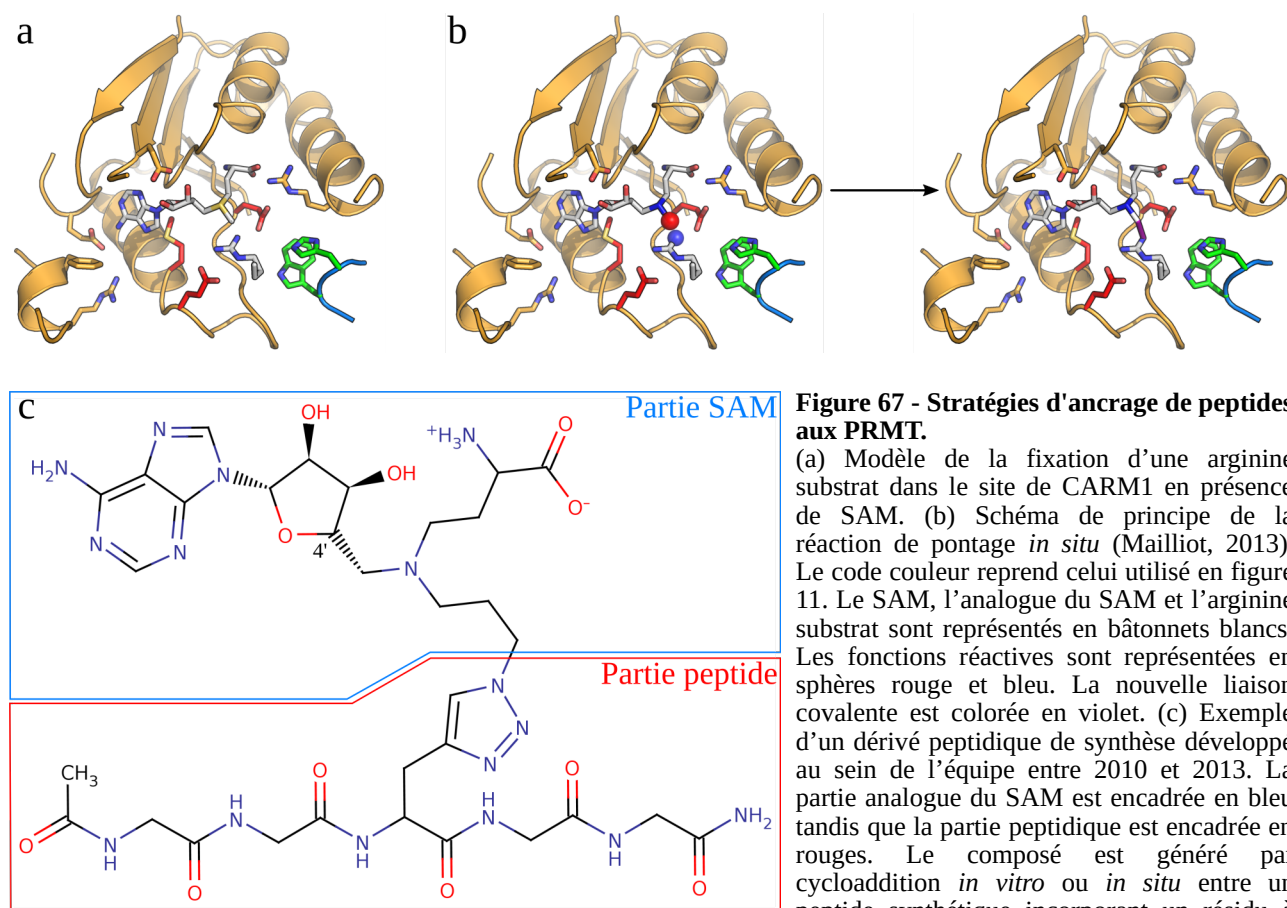


Figure 67 - Stratégies d'ancrage de peptides aux PRMT.

(a) Modèle de la fixation d'une arginine substrat dans le site de CARM1 en présence de SAM. (b) Schéma de principe de la réaction de pontage *in situ* (Mailliot, 2013). Le code couleur reprend celui utilisé en figure 11. Le SAM, l'analogue du SAM et l'arginine substrat sont représentés en bâtonnets blancs. Les fonctions réactives sont représentées en sphères rouge et bleu. La nouvelle liaison covalente est colorée en violet. (c) Exemple d'un dérivé peptidique de synthèse développé au sein de l'équipe entre 2010 et 2013. La partie analogue du SAM est encadrée en bleu tandis que la partie peptidique est encadrée en rouges. Le composé est généré par cycloaddition *in vitro* ou *in situ* entre un peptide synthétique incorporant un résidu à

fonction alcyne avec un analogue du SAM à fonction azoture.

Avant cette thèse, de nombreux travaux avaient été réalisés en particulier en utilisant mCARM1 comme cible. Malheureusement, aucune structure cristallographique de complexes CARM1-analogue « SAM-peptide » n'avait été obtenue.

Dans la continuité de cette approche, en collaboration avec les équipes de chimistes du Dr. N. I. Martin (Université d'Utrecht, Pays-Bas) et du Dr. G. Sbardella (Université de Salernes, Italie), nous avons mis au point durant cette thèse de nouveaux types d'analogues « SAM-peptide » qui ont permis de capturer et de stabiliser des complexes PRMT-peptides. Ce travail, résumé dans le premier paragraphe de ce chapitre, a donné lieu en 2017 à la publication « Transition state mimics are valuable mechanistic probes for structural studies with the arginine methyltransferase CARM1 », publiée dans la revue PNAS (van Haren et al., 2017) et jointe dans la section « Articles joints et communications scientifiques ». Au vu des résultats obtenus pendant cette thèse, l'analyse des structures cristallographiques suggère que les natures chimiques des analogues « SAM-peptides » utilisés précédemment (par exemple celui présenté en figure 67c) n'étaient probablement pas les mieux adaptées.

Un second paragraphe de ce chapitre présentera la généralisation de l'approche des dérivés peptidiques à d'autres substrats de CARM1 et en particulier l'histone H3. Un dernier paragraphe traitera de la généralisation de cette approche à d'autres PRMT et leurs substrats respectifs.

B - Résultats et discussion

1 - Les peptides analogues de l'état de transition

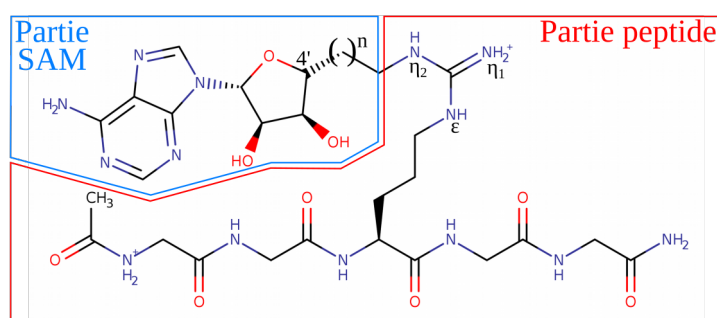


Figure 68 - Analogues « SAM-peptide ».

Formule chimique d'un analogue « SAM-peptide » quelconque tel que ceux utilisés pendant cette thèse. Plusieurs longueurs d'espaceur ($n = 1$ ou $n = 2$) ont été utilisées, avec deux types de liaisons : saturée ou insaturée.

Lors de cette thèse et dans le cadre de nos collaborations présentées ci-dessus, nous avons mis au point de nouveaux analogues « SAM-peptides ». Comme schématisé en figure 68, le résidu modifié est constitué d'un groupement adénosine ponté par le carbone 4' à l'azote η_2 du noyau guanidinium de l'arginine par un espaceur carboné long de deux à trois carbones. Cet espaceur peut de plus

posséder une liaison saturée ou insaturée. La structure de ces composés est très semblable à celle des dérivés peptidiques initialement développés au sein de l'équipe et présentés en figure 67c. Ces composés ne possèdent néanmoins pas le groupement méthionine du SAM.

Afin de démontrer la faisabilité de cette approche, de premiers composés ont été conçus en utilisant des peptides dérivés de PABP1, centrés sur les deux sites préférentiels de méthylation par CARM1 : R455 (peptide 447-459) ou R460 (peptide 456-466). Des dérivés peptidiques incorporant des espaceurs comportant deux à trois carbones, avec ou sans insaturation, ont été synthétisés et leur capacité inhibitrice a été mesurée sur CARM1 puis PRMT1 (dont PABP1 n'est pas un substrat naturel). Les expériences d'inhibition ont montré que les dérivés peptidiques PABP1 avec un espaceur de trois carbones inhibent fortement CARM1 avec une IC_{50} de l'ordre de 90 nM, et sont sélectifs de CARM1 comparé à PRMT1 d'un facteur 5 à 300. Les composés incorporant un espaceur de deux carbones inhibent significativement moins bien CARM1, avec des IC_{50} de l'ordre de 2 à 6 μ M. La forme saturée ou insaturée de l'espaceur ne semble pas influencer l'affinité de ces composés pour CARM1 et l'ancrage par l'arginine méthylée permet d'obtenir une inhibition supérieure d'un facteur 600 comparée à un peptide équivalent incorporant une arginine diméthylée asymétriquement à la place de l'arginine modifiée. Cela témoigne d'un important gain d'affinité apporté par la stabilisation dans le module catalytique due au groupement adénosine.

Pour démontrer formellement l'existence d'une interaction CARM1-dérivé peptidique de PABP1, et afin de déterminer les bases structurales de la fixation de ces composés par CARM1, des dérivés de PABP1 (R455 et R460) avec un espaceur de deux carbones, saturés et insaturés ont été co-cristallisés avec la construction du module catalytique de mCARM1 (130-487). Quatre structures cristallographiques à haute résolution de mCARM1 en complexe avec les composés PABP1 R455-saturé (PDB 5LGP, 2,04 Å) / -insaturé (PDB 5LGR, 2,00 Å) et PABP1 R460-saturé (PDB 5LGQ, 2,11 Å) / insaturé (PDB 5LGS, 2,10 Å) ont été résolues (statistiques de collecte et de traitement en annexes 6 et 7). Les structures appartiennent au groupe d'espace $P2_12_12$ caractéristique de mCARM1 avec pour paramètres de maille $a = 75 \text{ \AA}$ $b = 99 \text{ \AA}$ $c = 208 \text{ \AA}$. L'unité asymétrique contient un tétramère de mCARM1. Dans tous les cas, on observe un peptide fixé par monomère. Les cartes de densité électroniques révèlent l'ensemble des résidus pour les peptides centrés sur R455, et seuls les deux acides aminés carboxy-terminaux ne sont pas visibles pour les peptides centrés sur R460. Dans le site actif, les interactions établies entre CARM1 et la partie adénosine de l'arginine modifiée sont similaires à celles attendues avec le cofacteur SAM, et le groupement guanidinium forme trois liaisons hydrogène avec les acides glutamiques catalytiques E258 et E267 et H415 du motif IV, comme le ferait l'arginine substrat. La différence entre les composés saturés et insaturés se limite au déplacement de la méthionine conservée dans le motif III des PRMT de type I, et a une modification du schéma de solvatation dans la région qui normalement accueille la partie méthionine du SAM.

Nos structures ont été alignées sur les structures cristallographiques de CARM1 humaine en complexe avec des peptides PABP1 naturels et du SFG, obtenues quelques mois plus tôt par Copeland et ses collaborateurs (Boriack-Sjodin et al., 2016). La comparaison de ces structures

soulève plusieurs points. Tout d'abord, la conformation des dérivés peptidiques est rigoureusement la même que celle observée par Copeland et ses collaborateurs pour des peptides non ancrés (PDB 5DX1), ce qui démontre que l'arginine modifiée n'introduit aucun biais dans le positionnement du peptide. Ensuite, nos structures montrent que la présence de SFG, utilisé par nos compétiteurs pour stabiliser le peptide dans les structures du complexe PABP1-CARM1 humaine, introduit des distorsions du site catalytique, qui ne sont pas présentes avec l'utilisation de nos composés. Cela signifie que l'utilisation des dérivés peptidiques maintient le site actif de CARM1 dans une configuration productive. De surcroît, l'espaceur carboné entre le groupement adénosine et le groupement guanidinium se positionne exactement comme le méthyle arboré par l'arginine dans les structures des complexes PABP1 R455^{me1}/R460^{me1}-CARM1 humaine (PDB 5DXA), ce qui indique que nos composés sont des analogues de l'état de transition de la réaction de méthylation par les PRMT. Finalement, la comparaison des acides aminés de PABP1 visibles dans les cartes de densité électronique montre une meilleure couverture de quelques résidus pour les dérivés peptidiques comparée aux peptides non ancrés, probablement due au gain d'affinité apporté par l'ancrage des dérivés peptidiques au module catalytique de CARM1.

Avec ces travaux publiés en 2017, nous avons validé l'utilisation des dérivés peptidiques pour la stabilisation de complexes CARM1-substrat en vue d'études structurales. Les dérivés peptidiques permettent donc d'ancrer un analogue du substrat dans le site catalytique de CARM1, en mimant l'état de transition de la réaction de méthylation et sans perturber la conformation du peptide ni l'état du site actif. Fort de ce premier succès, nous travaillons désormais à généraliser cette approche à d'autres substrats de mCARM1 ainsi qu'à d'autres PRMT et en particulier PRMT1, PRMT2 et PRMT6 et PRMT7. Nous espérons ainsi établir des règles décrivant un mécanisme commun de reconnaissance « courtes et longues distances » entre les PRMT et leurs substrats, ou à l'inverse, démontrer des stratégies de fixations qui pourraient être spécifiques à une PRMT donnée.

2 - Complexes CARM1-dérivés de l'histone H3

Les travaux de l'équipe ont montré que la reconnaissance d'un peptide substrat par CARM1 impliquait des interactions transitoires dites à « courtes distances » et des interactions dites à « longues distances » (Mailliot, 2013). Comme l'illustre la figure 69, la reconnaissance à « courtes distances » fait intervenir les résidus du substrat situés à proximité du site de méthylation et les régions de la PRMT qui bordent le site catalytique.

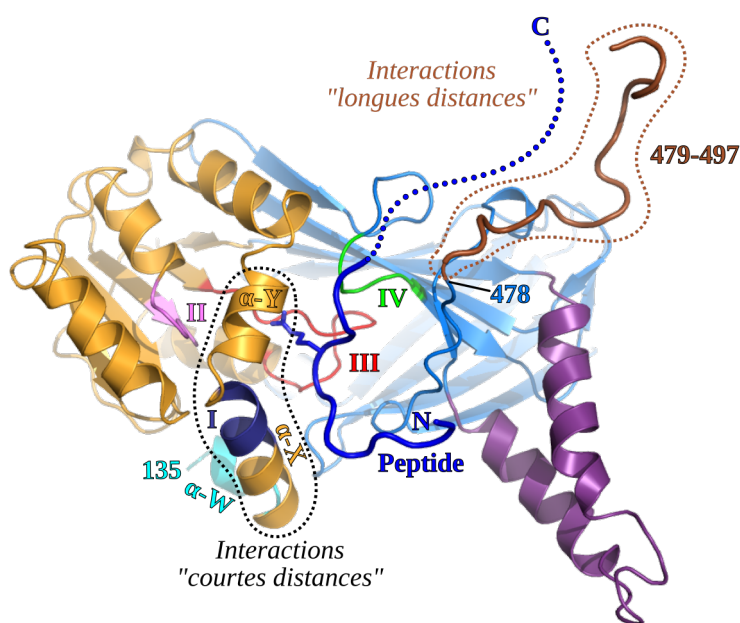


Figure 69 - Schéma de la reconnaissance CARM1 peptide.

Le code couleur reprend celui utilisé pour la figure 11. Le peptide est représenté en bleu vif. L'extrémité carboxy-terminale de mCARM1 est colorée en marron.

À l'inverse, les interactions à « longues distances » (schématisées en lignes hachurées marron) mettent en jeu des régions potentielles de reconnaissance protéine-substrats situées dans des zones plus éloignées du site de méthylation. Comme schématisé en figure 69, les reconnaissances à « longues distances » font probablement intervenir les extensions carboxy-terminales du peptide et nécessitent des peptides plus longs que ceux utilisés dans l'étude sur les complexes CARM1-peptides PABP1. La reconnaissance à « longues distances » implique aussi des

régions carboxy-terminales de CARM1 (résidus 479-497), comme cela a été montré par Justine Mailliot durant sa thèse.

Pour la suite des travaux, nous avons utilisé deux peptides H3 « courts » (13-31) et « longs » (1-41) centrés sur le site de méthylation H3R17, et trois constructions de mCARM1 : 130-487, 130-497 et 130-507. Ces trois constructions cristallisent dans des conditions identiques avec la même maille cristalline, contenant un tétramère dans l'unité asymétrique. Dans les structures cristallographiques, la construction 130-487 est ordonnée du résidu 135 au résidu 478 (figure 69) tandis que les constructions 130-497 et 130-507 sont ordonnées entre les résidus 135 et 497 (figure 69). Le segment 479-497, présent dans les constructions 130-497 et 130-507 de mCARM1, est ordonné dans les monomères A et C et interagit avec le monomère B ou D de l'unité asymétrique adjacente, dans un sillon qui pourrait être une région de fixation du substrat. Comme indiqué ci-dessus, la région 479-497 de CARM1 (en marron, figure 69) pourrait interagir avec l'extrémité carboxy-terminale de peptides plus longs que ceux utilisés précédemment.

Les peptides dérivés de l'histone H3 13-31 (19 résidus) et 1-41 (41 résidus) couvrent respectivement une zone allant de -4 à +14 et de -16 à +24 par rapport au site de méthylation H3R17. Les dérivés de PABP1 (447-459 & 456-466) utilisés précédemment couvraient une zone de -8 à +4 et de -4 à +5 par rapport aux sites de méthylation R455 et R460.

L'utilisation de peptides plus longs que ceux dérivés de PABP1 a une seconde fonction. Lors du travail réalisé avec les dérivés peptidiques de PABP1, nous avons observé la fixation de deux peptides par dimère. Ceci est en contradiction avec les résultats obtenus par l'équipe sur PRMT6 (Bonfond et al., 2015) et PRMT2 (Chapitre III), qui ont montré qu'un dimère de PRMT ne fixe qu'un unique substrat. Nous pensons donc que la fixation de deux peptides par dimère n'est probablement pas biologiquement significative et pourrait être un biais des structures cristallographiques utilisant des peptides courts.

Nous avons aussi utilisé des peptides dérivés de l'histone H3 13-31 et 1-41 comportant un résidu K18 acétylé (K18^{Ac}). Des corrélations ont en effet été mises en évidence, liant la méthylation de H3R17 et l'acétylation de H3K18, dont les origines restent à expliquer. Comme indiqué au chapitre I, CARM1 est principalement connu pour son rôle de coactivateur des récepteurs nucléaires et agit en méthylant H3R17 et H3R26 *in vivo* (figure 70), en synergie avec les protéines des familles p160 et p300. L'acétylation de H3K18 par p300 est nécessaire à la méthylation de H3R17 par CARM1. Justine Mailliot a démontré dans sa thèse que CARM1 a une plus grande affinité pour H3K18^{Ac} que pour H3K18 ce qui expliquerait pourquoi le peptide méthylé en R17 est préférentiellement d'abord acétylé en K18. À l'inverse, Yue et ses collaborateurs (Yue et al., 2007) suggèrent que la méthylation en R17 sur les peptides préalablement marqués K18^{Ac} n'est pas due à un gain d'affinité mais à la stabilisation de l'intermédiaire réactionnel chargé positivement lors de la réaction de méthylation de l'arginine, facilité par la neutralisation de la charge en K18. Pour cela, ils s'appuient sur l'analyse des séquences de substrats de CARM1 montrant qu'aucune cible de CARM1, excepté H3, ne contient de résidu chargé en position n+1 de l'arginine méthylée.

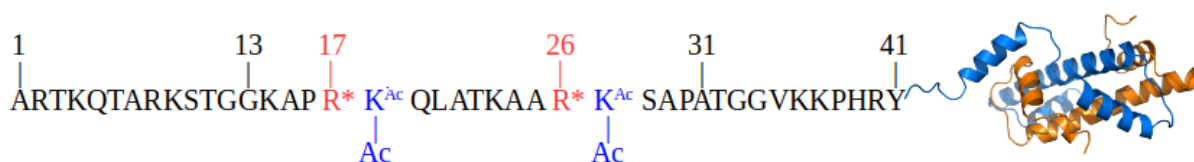


Figure 70 - Sites de méthylation par CARM1 et d'acétylation par p300 sur l'histone H3.

Le dimère d'histone H3-H4 est représenté respectivement en rubans bleu et orange. Une partie de la séquence de la queue de l'histone H3 de souris est détaillée. Les arginines colorées en rouge correspondent aux sites de méthylation par CARM1 et les lysines colorées en bleu représentent les sites d'acétylation par p300 en n+1 des arginines cibles. Les astérisques correspondent aux arginines modifiées dans les dérivés peptidiques, mais celles-ci ne sont pas modifiées en même temps sur un même dérivé.

Lors de la réalisation de cette étude, les complexes CARM1-dérivés de H3 ont initialement été co-cristallisés dans les mêmes conditions qu'utilisées pour les dérivés de PABP1 avec un mélange PEG/sulfate d'ammonium comme agent cristallisant. Bien qu'une grande quantité de

cristaux de complexe ont été obtenus et que le traitement des données a confirmé que les dérivés de peptides sont fixés à mCARM1, les densités électroniques obtenues n'étaient pas de bonne qualité pour les peptides substrats. Nous avons suspecté que le PEG pourrait entrer en compétition avec la partie peptidique des dérivés de H3 et avons cherché de nouvelles conditions de cristallisation. L'ensemble des résultats sont présentés dans les paragraphes suivants.

a - Reconnaissance à « courtes » et « longues » distances

Les trois constructions de mCARM1 sont purifiées en routine au laboratoire depuis plusieurs années (Hassenboehler, 2010 ; Mailliot, 2013 ; Troffer-Charlier et al., 2007). Les constructions de mCARM1 étiquetées GST ont été extraites du lysat cellulaire par affinité pour une résine au glutathion. Les protéines sont ensuite éluées par clivage de la GST par la protéase TEV puis sont une fois de plus purifiées par chromatographie de filtration sur gel.

➤ **Structures des complexes CARM1-peptides H3 cristallisés avec du PEG**

Les dérivés de peptides H3 courts (13-31) et longs (1-41) ont été co-cristallisés avec les constructions 130-487, 130-497 et 130-507 de mCARM1 avec le PEG3350 comme agent cristallisant principal, dans les mêmes conditions ayant permis d'obtenir des cristaux de complexes CARM1-peptides dérivés de PABP1. Des cristaux ont été obtenus pour toutes les constructions de mCARM1 en complexe avec des peptides courts, tandis que des cristaux ont été obtenus pour les peptides longs uniquement pour la construction mCARM1 (130-487). Les données de diffractions de trente-trois cristaux ont été collectées sur les lignes de lumière synchrotron ID30b et Proxima 1 de l'ESRF et de SOLEIL, et traitées avec le programme XDS et la suite de programmes CCP4. Les structures ont été résolues par remplacement moléculaire puis affinées dans Phenix. La construction et la correction des modèles a été réalisée avec le programme Coot. Le tableau 10 résume les résultats obtenus à partir des jeux de diffraction collectés.

	H3 13-31	H3 13-31 K18 ^{Ac}	H3 1-41	H3 1-41 K18 ^{Ac}
mCARM1 130-487	4 jeux	5 jeux	10 jeux	10 jeux
mCARM1 130-497	1 jeu *	Pas de cristaux	Pas de cristaux	Pas de cristaux
mCARM1 130-507	Pas de cristaux	3 jeux **	Pas de cristaux	Pas de cristaux
Peptide ordonné	13-20	13-22	15-20	15-22
Résolutions (Å)	2,12-2,32	2,19-2,38	1,88-2,42	1,97-2,37

Tableau 10 - Résumé des cristaux de complexes mCARM1-dérivés de H3 collectés.

Ce tableau présente les données structurales extraites des trente-trois cristaux de complexes mCARM1-dérivés de H3 collectés. Les astérisques désignent les structures pour lesquelles les cartes de densité électronique correspondant aux peptides sont de bonne qualité. Un astérisque indique une structure. La couverture est une indication du nombre de résidus de peptides observés. La ligne « Résolutions » répertorie l'intervalle de résolutions maximales obtenues avec ces complexes. Seules les statistiques de collecte et de traitement des deux meilleurs structures sont présentées en annexes 7 et 8.

Comme indiqué dans le tableau 10, les peptides H3 n'ont pu être visualisés que pour trois (annexes 7 et 8) des trente-trois structures (représentées par un astérisque dans le tableau 10). Ces

structures concernent uniquement les peptides H3 13-31 co-cristallisés avec les constructions 130-497 et 130-507 de mCARM1. Comme cela avait été observé avec les dérivés de PABP1, un dérivé de H3 est fixé par monomère de CARM1. Pour les trente autres structures, concernant principalement des peptides H3 1-41, le traitement a mis en évidence une faible fixation des dérivés de H3 sur les modules catalytiques de mCARM1, aussi bien pour la partie mimant le SAM que pour la partie peptidique. Ceci est dû à la présence du cofacteur SAH et montre donc l'existence d'un mélange peptide-SAH dans les cristaux.

Dans nos cristaux de CARM1, pour les trois constructions utilisées, le SAH est présent de manière intrinsèque alors qu'il n'a été introduit ni dans les étapes de purification, ni dans les solutions utilisées pour la cristallisation. Un nouveau ligand de travail doit donc chasser la molécule de SAH présente sur chaque monomère. Le chapitre VI montrera que l'affinité des composés pour CARM1 se traduit dans les structures cristallographiques par leur capacité/incapacité à remplacer complètement ou partiellement les molécules de SAH d'un ou de plusieurs monomères du tétramère de CARM1. En plus de la faible fixation des dérivés de H3 à mCARM1, une couverture incomplète du peptide utilisé a été observée pour l'ensemble des structures. En effet, au mieux seuls les résidus 13 à 22 sont observés pour les peptides courts, et 15 à 22 dans le cas des peptides longs. On n'observe que les résidus 15 à 18 dans la plupart des structures. En tout, vingt-et-une des trente structures ne contiennent pas de peptide ordonné ou contiennent un mélange peptide-SAH, avec une couverture peptidique souvent limitée aux résidus 16 à 18. Ces premiers résultats nous ont amené à suspecter que les conditions de cristallisation utilisées pourraient nuire à la fixation des peptides, comme cela sera développé dans le paragraphe suivant.

➤ **Influence du PEG sur la stabilisation des complexes PRMT-peptides**

La résolution par notre équipe de très nombreuses structures cristallographiques de mCARM1 en complexe avec du SAH ou des inhibiteurs, nous a conduit à constater la présence de nombreuses molécules de PEG ordonnées à la surface de la protéine. En effet, le PEG est l'agent cristallisant principal utilisé lors de la cristallisation de CARM1. Les molécules de PEG forment des chaînes plus ou moins longues qui, de part leur nature chimique et leurs charges, peuvent mimer les propriétés d'une chaîne polypeptidique et se fixer dans des zones d'interactions protéine-protéine.

Comme l'illustre la figure 71a, les molécules de PEG (représentés par le volume rouge) se localisent dans un renforcement du module catalytique de mCARM1 correspondant au site de fixation du peptide situé à l'entrée du site actif, dans la poche de fixation de l'arginine substrat et sur le tonneau β coté extérieur du dimère.

Le regroupement des molécules de PEG trouvées dans toutes nos structures de mCARM1 puis leur superposition sur une structure du complexe CARM1-PABP1 (figure 71b) montre une colocalisation du peptide et des molécules de PEG. Cela suggère que la présence de PEG pourrait

interférer avec la formation du complexe CARM1-dérivé peptidique, d'autant plus que dans les conditions permettant la cristallisation, l'agent cristallisant est présent à forte concentration comparé aux peptides.

La présence de PEG sur ces points particuliers de la structure pourrait néanmoins fournir des indications concernant les zones de fixation encore inconnues des peptides. En effet, les zones auxquelles se fixent les PEG, près du motif IV et sur l'extérieur du tonneau β , pourraient être des sites d'interaction des substrats de CARM1. C'est d'autant plus vraisemblable que cette région interagit avec le segment 479-497 dans les structures de mCARM1 130-497 et 130-507, ce qui démontre un potentiel de fixation de chaînes polypeptidiques. Des sites de molécules de PEG sont également observés sur PRMT2, mais moins de structures sont disponibles pour définir des zones d'interaction de manière fiable.

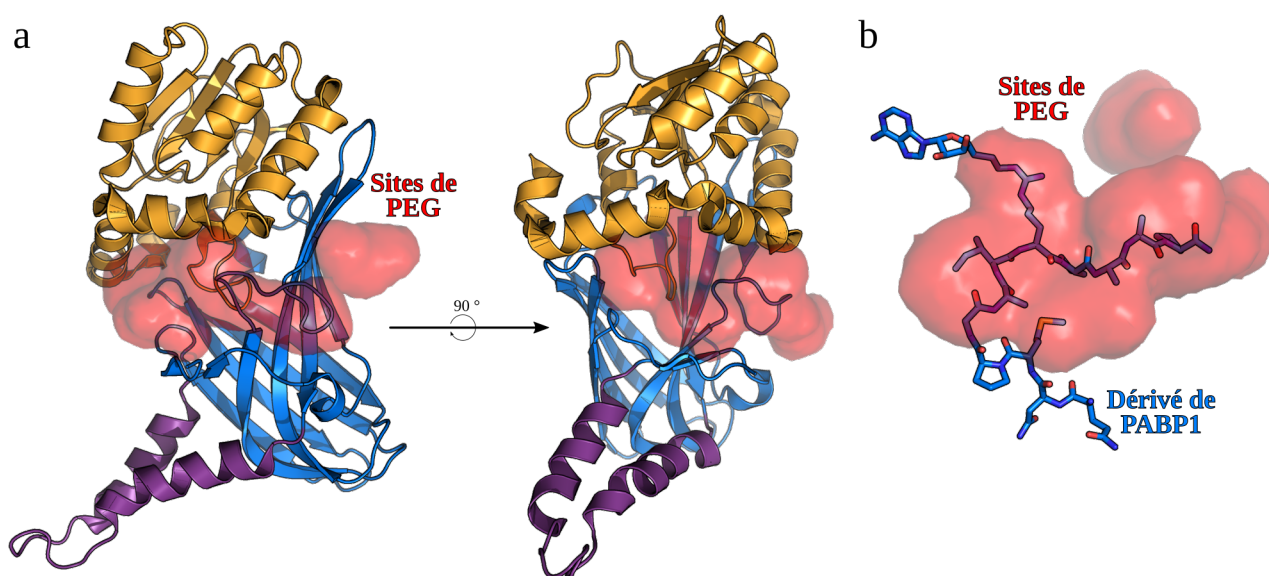


Figure 71 - Sites de fixation des molécules de PEG sur CARM1.

Afin de délimiter les zones où se fixent l'agent cristallisant, 20 structures cristallographiques de mCARM1 ont été alignées et les résidus correspondant à des molécules de PEG ont été sauvegardés dans un unique fichier PDB. A l'aide du programme déjà utilisé pour visualiser les poches de fixation du SAM et de l'arginine, des volumes ont été calculés pour les positions des molécules de PEG. (a) Positions occupées par des molécules de PEG (densité rouge) sur un module catalytique de mCARM1. (b) Superposition d'une structure du complexe mCARM1-dérivé de PABP1 (PDB 5LGP) et des sites de fixation du PEG (densité rouge). Pour améliorer la visibilité, CARM1 n'a pas été représentée.

Afin de pallier le problème potentiel que pourrait représenter la compétition entre les molécules de PEG et les peptides, de nouvelles conditions de cristallisation excluant les PEG ont été recherchées. Dans une étude (McPherson, 2001), McPherson a comparé l'utilisation de divers sels comme agents cristallinisants pour la cristallisation des macromolécules et a montré que le malonate de sodium permet d'obtenir des cristaux pour un plus grand nombre de protéines comparé aux autres sels, y compris le sulfate d'ammonium. Cette efficacité de malonate serait due à sa forte force ionique et à ses propriétés cosmotropes. Des conditions de cristallisation de mCARM1 130-487 et 130-497 avec le malonate comme agent cristallisant ont été mises au point en criblant ces trois constructions contre une gamme de concentration en divers sels ajustés à une large gamme de pH.

Ces trois constructions cristallisent dans 1 M à 1,5 M de malonate pour des pH compris entre 7 et 8,5.

➤ Structures des complexes CARM1-peptides H3 cristallisés avec du malonate

La construction mCARM1 130-497 a été co-cristallisée avec les dérivés de H3 en utilisant le malonate comme agent cristallisant. Des cristaux ont été obtenus pour tous les dérivés de H3, même pour les complexes avec les peptides H3 1-41, bien que peu de cristaux soient disponibles pour ces derniers. Quinze jeux de données ont été collectés sur la ligne de rayonnement synchrotron ID30b de l'ESRF ou sur l'équipement de l'IGBMC, puis traités tel qu'indiqué dans le paragraphe précédent.

Comme le montrent les résultats de traitement résumés dans le tableau 11, des densités électroniques de bonne qualité pour les dérivés de H3 courts et longs sont visible dans six des douze structures, ce qui est significativement plus fréquent que pour les structures obtenues à partir des conditions de cristallisation en PEG (3 pour 30). De plus, les résidus 12-22 (peptides 13-31) ou 15-21 (peptides 1-41) sont visible dans douze des quinze structures et aucun mélange peptide-SAH n'est observé. La résolution des structures est néanmoins moins bonne que pour les cristaux obtenus dans le PEG. Les premières structures de mCARM1 résolues par l'équipe provenaient de cristaux obtenus dans le sulfate d'ammonium. Ces structures avaient des résolutions équivalentes à celles observées pour les structures issues des cristaux obtenus en malonate, c'est pour cela que les PEG avaient été historiquement conservés comme agent cristallisant par notre équipe.

	H3 13-31	H3 13-31 K18 ^{Ac}	H3 1-41	H3 1-41 K18 ^{Ac}
mCARM1 130-497	4 jeux *	4 jeux ****	2 jeux	2 jeux *
Peptide visible	13-22	13-22	15-21	15-21
Résolutions (Å)	2,60-2,75	2,20-3,01	2,23-2,65	2,40-2,80

Tableau 11 - Résumé des structures obtenues à partir des conditions en malonate.

Les astérisques désignent les structures pour lesquelles les cartes de densité électronique correspondant aux peptides sont de bonne qualité. Un astérisque indique une structure. La couverture est une indication du nombre de résidus de peptides observés. La ligne « Résolutions » répertorie l'intervalle de résolutions maximales obtenues avec ces complexes.

Parmi les douze structures résolues à partir des cristaux obtenus en malonate, quatre correspondent au peptide 13-31, quatre au peptide 13-31 K18^{Ac}, deux au peptide 1-41 et les deux dernières correspondent au peptide 1-41 K18^{Ac}. Trois structures pour lesquelles la qualité des densités électroniques des peptides étaient très bonne ont été finalement affinées. Ces structures correspondent aux complexes mCARM1 (130-497) avec les dérivés de H3 13-31 (2,20 Å), H3 13-31 K18^{Ac} (2,54 Å) et H3 1-41 K18^{Ac} (2,45 Å).

Aucune structure n'a été totalement affinée pour les complexes avec H3 1-41 car les densités électroniques correspondant aux peptides sont de très mauvaise qualité. Les statistiques de collecte et de traitement des données pour ces structures sont réunies dans l'annexe 8. La figure 72

présente les structures cristallographiques de mCARM1 (130-497) en complexe avec les peptides dérivés de H3. le peptide H3 se positionne tel qu’observé dans les structures obtenues par Copeland et ses collaborateurs (PDB 5DWQ) (Boriack-Sjodin et al., 2016), ce qui suggère une fois de plus, sur un autre substrat que PABP1, que notre approche permet de produire des informations biologiquement pertinentes.

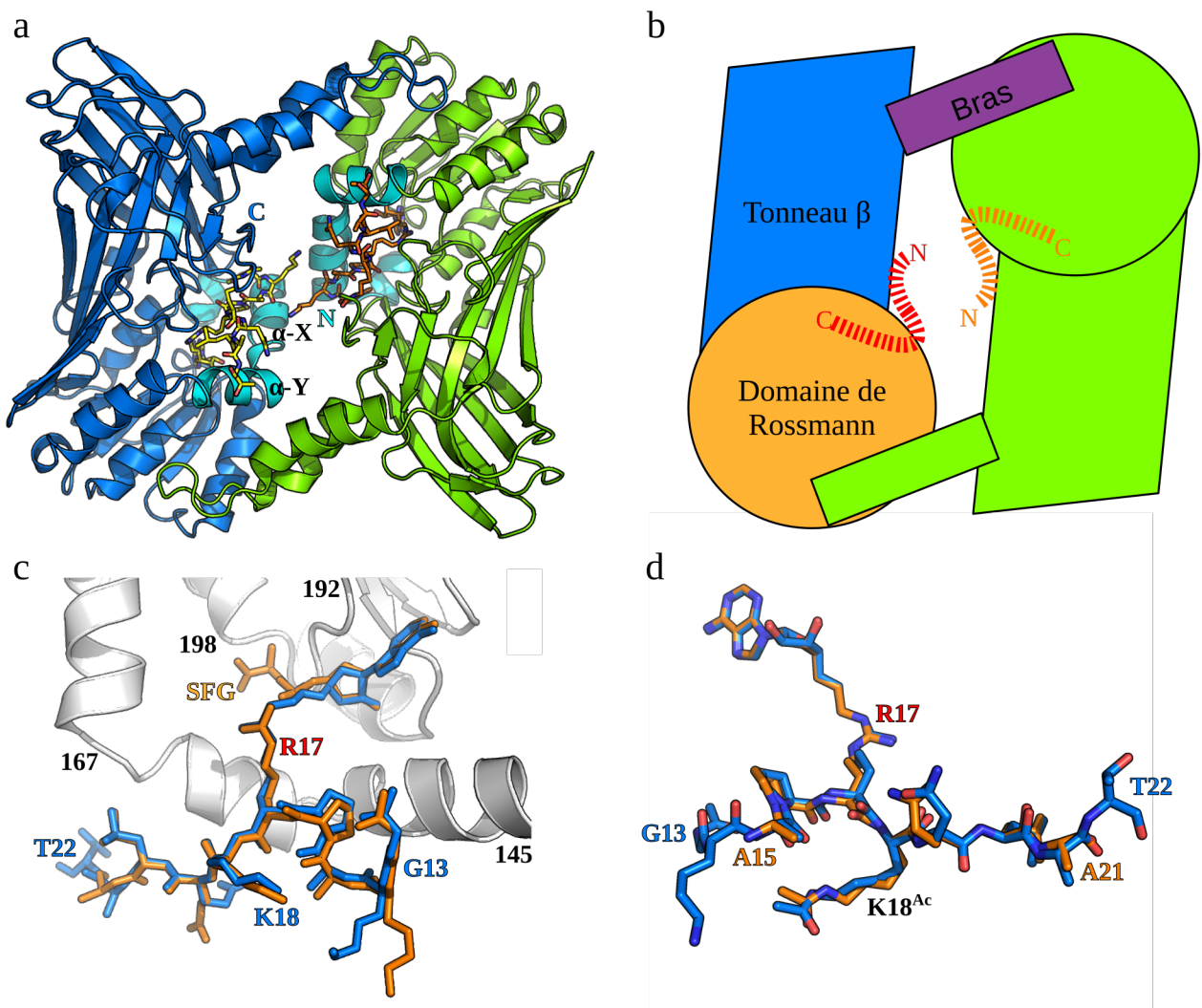


Figure 72 - Structures cristallographiques de mCARM1 en complexe avec les dérivés de H3.

(a) Dimère de mCARM1 en complexe avec le dérivé H3 (13-31) K18^{Ac}. Le dimère est représenté en rubans bleu et vert. Les peptides sont représentés en bâtonnets jaune et orange. Les hélices amino-terminales de chaque monomère sont coloriées en bleu clair. (b) Représentation schématique de 72a. Le monomère de droite est représenté en vert, et son peptide associé en orange comme sur la figure 72a. le monomère de gauche est coloré en orange, bleu et violet pour délimiter le domaine de Rossmann, le tonneau β et le bras de dimérisation. Le peptide qui lui est associé est coloré en rouge. (c) Comparaison du mode de fixation du dérivé de peptide H3 (13-31) K18 (bleu) avec le peptide équivalent (orange) co-cristallisé avec CARM1 humaine et du SFG (Boriack-Sjodin et al., 2016). CARM1 est représenté en ruban blanc. (d) Comparaison de la couverture des dérivés de peptides H3 (13-31) K18^{Ac} (bleu) et H3 (1-41) K18^{Ac} (orange).

Il ressort de l’observation de ces structures que l’utilisation des peptides H3 1-41 ne permet pas d’augmenter le nombre d’acides aminés visibles comparé aux peptides H3 13-31. Au contraire, le nombre de résidus visibles pour les peptides longs est inférieur à celui observé pour les peptides courts à leur extrémité amino-terminale. On peut imaginer que la forme cristalline obtenue n’est pas

compatible avec les interactions « longues distances » CARM1-peptides d'où peut être l'absence de cristaux obtenus avec les peptides H3 1-41 et les constructions CARM1 130-497 et 130-507.

L'utilisation de malonate comme agent cristallisant n'augmente donc pas le nombre de résidus visible dans les structures. Néanmoins, l'usage de ce sel permet d'améliorer la reproductibilité des cristaux contenant un peptide ordonné. Aussi, seule l'utilisation de malonate comme agent cristallisant a permis d'obtenir des cristaux de mCARM1 130-497 avec les dérivés de peptides H3 1-41.

Plusieurs éléments sont susceptibles d'expliquer pourquoi les peptides H3 1-41 semblent moins bien fixer mCARM1 que les peptides H3 13-31 :

- les constructions de CARM1 utilisées (130-487 et 130-497), plus courtes que celles utilisées par Justine Mailliot pour les mesures d'affinité (130-507), pourraient perdre des interactions « longues distances » stabilisatrices entre le peptide et CARM1. La construction 130-507 avait été utilisée initialement en co-cristallisation avec les peptides H3 courts, mais avant d'avoir mis au point de nouvelles conditions de cristallisation basées sur le malonate,
- l'empilement des cristaux de mCARM1 fait intervenir des interactions inter-dimères à proximité de surfaces potentiellement impliquées dans la fixation du substrat. Ces contacts cristallins pourrait interférer avec la fixation du peptide, ce qui signifie que cette forme cristalline pourrait ne pas être compatible avec des peptides plus longs qui nécessiteraient une restructuration du dimère de CARM1,
- la cavité souvent observée au centre des dimères des PRMT est obstruée par les hélices α -X et α -Y dans les structures de CARM1 (figures 13a et 72a), peut être du fait des contraintes cristallographiques. Cette configuration force l'extrémité amino-terminale des dérivés longs de H3 à former une boucle pour s'extraire de la cavité, ce qui a également été observé pour les dérivés de PABP1. Néanmoins, dans les structures des complexes mCARM1-PABP1, les dérivés de peptides pourraient être stabilisés par des interactions entre le peptide fixé à un premier monomère et le peptide ancré au second monomère, ce qui expliquerait qu'ils soient entièrement visibles dans la carte de densité électronique. Aucune interaction de ce type n'est observée pour les dérivés longs (1-41) de H3. Dans le cas des dérivés courts 13-31, l'acétylation à l'extrémité G13 amino-terminale, ajoutée lors de la synthèse pour neutraliser les peptides, forme deux liaisons hydrogène avec Y262 et N266 dans une cavité sur la surface de mCARM1. Ceci stabilise les résidus 13 à 15 dans une configuration probablement non biologiquement significative. Dans le milieu intracellulaire, les peptides pourraient avoir à traverser l'orifice entre les deux monomères afin d'être correctement fixé.

Afin de faciliter les réarrangements conformationnels qui pourraient être nécessaire au passage des peptides à travers le dimère (Troffer-Charlier et al., 2007), les échantillons ont été incubés 12 heures, à température ambiante ou à 37 °C, en amont de la cristallisation. Les changements conformationnels subis par les protéines s'étalent sur des durées pouvant aller de la microseconde à plusieurs minutes. Une incubation de plusieurs heures devrait donc être suffisante. Cependant, cela n'a pas permis d'améliorer la fixation des dérivés de H3 à mCARM1 car les mêmes structures cristallographiques ont été obtenues.

b - Dialogue entre l'acétylation de H3K18 et la méthylation de H3R17

En utilisant des dérivés de peptides H3 acétylés et non acétylés sur K18, nous espérons établir les éléments structuraux favorisant la méthylation de H3R17 par CARM1 suite à l'acétylation sur K18 par p300, et ainsi expliquer le dialogue entre ces deux modifications post-traductionnelles des queues d'histones.

Les données concernant les structures des complexes mCARM1-peptides H3 acétylés sont indiquées dans le paragraphe précédent. Sur l'ensemble des structures traitées (douze structures obtenues en utilisant le malonate comme agent cristallisant et trente-trois en utilisant le PEG3350), les densités électroniques pour les peptides K18^{Ac} est en moyenne meilleure que pour les peptides K18, ce qui témoigne directement du gain d'affinité apporté par l'acétylation (Mailliot, 2013). Parmi les trois meilleures structures cristallographiques de complexes mCARM1-H3 présentées dans le paragraphe précédent, deux complexes incorporent un résidu K18^{Ac} (un avec peptide 1-41 et un avec peptide 13-31) et un complexe contient un peptide court incorporant un résidu K18 non acétylé.

La figure 73 met en évidence les interactions entre les dérivés de peptide H3 et mCARM1. Les densités électroniques dans les chaînes A et C permettent de construire les structures atomiques des peptides H3 13-31 K18 (figure 73a) et K18^{Ac} (figure 73b) et de discriminer entre les positions occupées par les différents atomes de l'acétylation de K18. Étonnamment, l'acétylation en K18 n'a pratiquement aucun effet sur la structure locale de mCARM1 et ne modifie ni la configuration du peptide, ni la position de la lysine comparativement à la forme non acétylée. La modification post-traductionnelle permet cependant d'établir des interactions supplémentaires permettant d'expliquer le gain d'affinité mesuré.

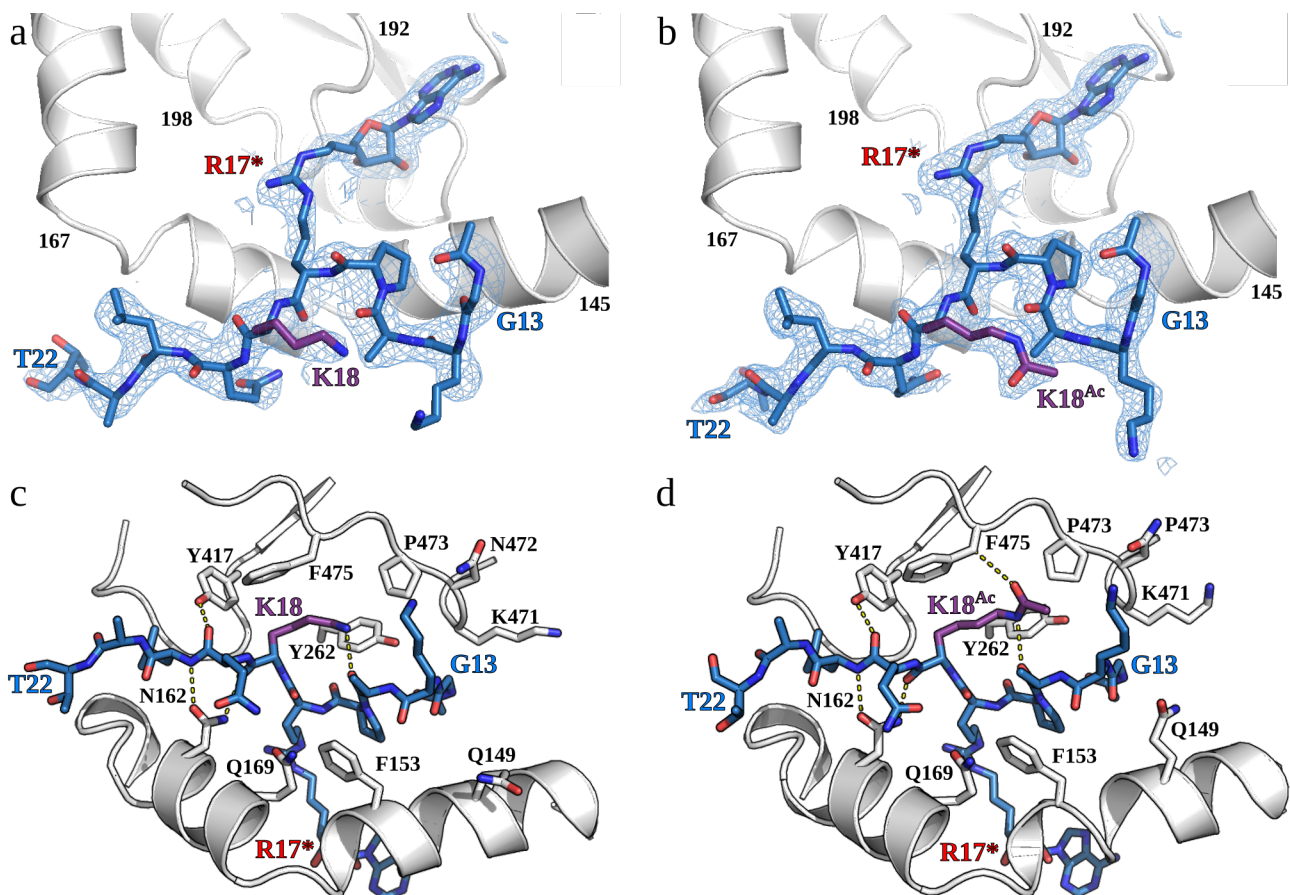


Figure 73 - Modes de fixation des peptides H3 (13-31) K18 et K18^{Ac} à mCARM1.

Les peptides H3 sont représentés bâtonnets bleus, excepté la lysine 18 qui est colorée en violet. mCARM1 est représenté en rubans blanc. Les liaisons hydrogène sont représentées en lignes hachurées jaunes. L'arginine modifiée est annotée en rouge avec une astérisque. (a / b) mise en évidence de la densité électronique pour les dérivés de peptides H3 (13-31) K18 (a) et K18^{Ac} (b). (c / d) Mode de fixation des dérivés de peptides H3 (13-31) K18 (c) et K18^{Ac} (d). Les cartes de densités électroniques 2F_o-F_c sont représentées par un maillage bleu, avec un niveau de contourage de 1 σ .

Le gain d'affinité mesuré pour le peptide K18^{Ac} pourrait être expliqué par l'ensemble des observations indiquées ci-dessous :

- l'environnement à proximité de l'azote chargé positivement ou de l'acétylation du résidu K18 ne contient pas de résidus formellement chargés. L'acétylation, en neutralisant la charge de la lysine, pourrait permettre de minimiser la répulsion entre l'azote positivement chargé de K18 et les noyaux environnants. Cependant, le fait que K18 se positionne de manière similaire à K18^{Ac} suggère que cet effet répulsif est négligeable. Ceci est conforté par l'analyse du potentiel électrostatique à l'aide du logiciel APBS (Baker et al., 2001), montrant que la zone de fixation de K18 et K18^{Ac} est globalement chargée négativement (non présenté),

- L'acétylation de K18 introduit des interactions supplémentaires sans détruire celles existant dans le cas du peptide non modifiée. La liaison hydrogène intra-peptide, établie entre l'azote de la chaîne latérale de K18 et l'oxygène de la chaîne principale de l'alanine 15 en position n-3, est maintenue malgré l'acétylation. Il en est de même avec l'ensemble des interactions de type « van der Waals » entre la chaîne aliphatique de K18 et les groupements phénol et phényle de Y417, F475

et Y262 de mCARM1. L'acétylation ajoute une liaison hydrogène faible de type C-H...O entre l'oxygène du groupement acétyle et le carbone aliphatique β de F475. Les deux carbones du groupement acétyle permettent également d'établir des interactions de « van der Waals » supplémentaires avec le méthyle de A15 du peptide et le cycle de P473 de mCARM1,

- L'énergie de désolvatation de la lysine acétylée est probablement moindre que celle du résidu natif, d'où un gain sur l'énergie de fixation protéine-peptide.

Comme indiqué dans le paragraphe introductif, Yue et ses collaborateurs ont proposé que l'acétylation de K18 par p300 préalable à la méthylation de R17 par CARM1 n'est pas due à un gain d'affinité mais à la stabilisation de l'intermédiaire réactionnel chargé positivement lors de la réaction de méthylation de l'arginine, facilité par la neutralisation de la charge en K18. Cependant, nos structures montrent que l'azote de la chaîne latérale du résidu K18 se situe à une distance de 12 Å du site de méthylation. Une telle distance ne semble pas appuyer l'hypothèse émise par Yue et ses collaborateurs. Justine Mailliot (Mailliot, 2013) a proposé dans sa thèse un modèle selon lequel l'acétylation de K18 favoriserait son interaction avec N266 par établissement de deux liaisons hydrogène entre l'azote de la chaîne latérale de K18 et l'oxygène de la chaîne latérale de N266, et l'oxygène de l'acétyle sur K18 et l'azote de la chaîne latérale de N266. Le positionnement du peptide dans nos structures invalide également ce modèle. Nos structures suggèrent que le gain d'affinité de H3 K18Ac pour CARM1 serait issu de l'établissement de quelques interactions de « van der Waals » supplémentaires, d'une liaison hydrogène faible et d'un probable « gain » dans l'énergie de désolvatation du peptide. Il n'est pas non plus à exclure que la méthylation de H3R17 K18Ac-dépendante par CARM1 ne soit pas due au gain d'affinité apporté par l'acétylation de K18, mais que CARM1 soit recruté au niveau des promoteurs en cours d'activation suite à la reconnaissance de l'acétylation en K18 par le bromodomaine de p300/CBP. Cette dernière hypothèse est cependant remise en question par des travaux récents ayant montré, par l'étude biochimique et structurale par cryo-EM du complexe ER-CBP/p300-SRC/p160-CARM1, que la fixation de CARM1 à ce complexe stimule l'activité d'acétylation de p300/CBP, ce qui en retour stimule la méthylation par CARM1 de manière K18_{Ac} dépendante.

3 - Généralisation des dérivés peptidiques à d'autres PRMT

Suite aux travaux précédents, nous travaillons à généraliser l'utilisation des dérivés peptidiques à d'autre PRMT et en particulier à PRMT1, PRMT6 et PRMT7. Ces dernières méthylent l'histone H4 en position 3 (R3, figure 74) (Jelinic et al., 2006 ; Lakowski et al., 2009 ; Strahl et al., 2001).

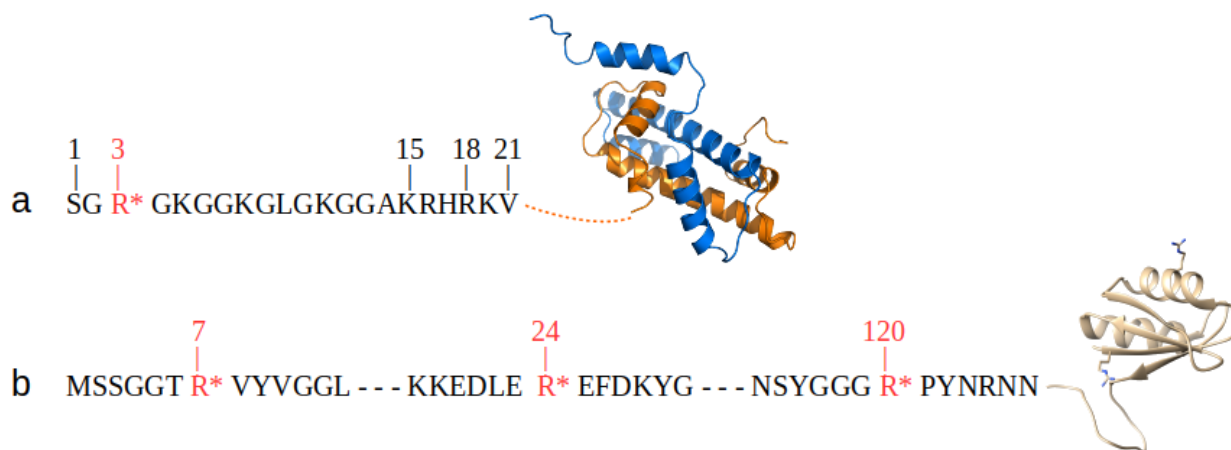


Figure 74 - Sites de méthylation sur H4 et rsf1.

(a) Site de méthylation sur H4. Le dimère d'histone H3 H4 est représenté respectivement en rubans bleu et orange. Une partie de la séquence de la queue de l'histone H4 de souris est détaillée. (b) Sites de méthylation par PRMT2 sur rsf1. Rsf1 est représenté en ruban beige. Une partie de la séquence de la queue de rsf1 de *S. frugiperda* est détaillée. Les arginines représentées en rouge correspondent aux sites méthylés. Les astérisques correspondent aux arginines modifiées dans les dérivés peptidiques, mais celles-ci ne sont pas modifiées en même temps sur un même dérivé.

Des dérivés de peptide H4 (1 à 15, 1 à 18 et 1 à 21) incorporant l'arginine modifié en R3 ont été synthétisés. Nous avons récemment caractérisé rsf1 de *S. frugiperda*, un facteur d'épissage méthylé par PRMT2 sur trois sites (Cura et al., 2017). Trois dérivés peptidiques de rsf1 incorporant le résidu arginine modifié en R7, R24 ou R120 (1 à 13, 18 à 30 et 114 à 126) ont également été synthétisés. La figure 74 présente les séquences utilisées pour la conception de ces peptides. Des essais de cristallisation sont aujourd'hui en cours sur ces dérivés peptidiques.

C - Perspectives

1 - Dérivés peptidiques de l'histone H3

L'utilisation de peptides H3 longs n'a pas permis d'améliorer la couverture des peptides dans les conditions expérimentales utilisées. En nous basant sur le nombre d'acides aminés des dérivés de H3 visibles dans nos structures, de nouveaux peptides, plus courts, ont été synthétisés : les peptides H3 (10-25) K18 et K18^{Ac}, centrés sur R17 et les peptides H3 (22-31) K27 et K27^{Ac} centrés sur R26, un autre site de méthylation par CARM1. Avec les peptides 10-25, nous espérons obtenir une meilleure couverture peptidique en éliminant les éventuelles interactions non spécifiques inter-peptides et en limitant la possible déstabilisation de peptides longs due à

l'empilement cristallin et aux extrémités carboxy-terminales « flottantes ». Les peptides centrés sur le site de méthylation R26 (22-31) pourraient permettre de comprendre comment CARM1 reconnaît et méthyle spécifiquement un second site sur un même substrat. La co-cristallisation du complexe CARM1 130-507 avec des dérivés de H3 1-41 dans les conditions malonate n'a pas encore été testée. La combinaison de la cristallisation en malonate et de l'utilisation de la construction mCARM1 130-507, ayant une plus forte affinité pour H3 1-41 (Mailliot, 2013), pourrait permettre d'améliorer la stabilité du complexe CARM1-dérivé de H3.

2 - [Application à d'autres peptides et substrats complets](#)

Plusieurs autres pistes de développement sont en cours.

1. Les analogues SAM-peptides utilisés actuellement sont des mimes de la première étape de la réaction de méthylation. De nouveaux composés vont être synthétisés pour incorporer un second groupement méthyl dans le but d'obtenir des mimes de l'état de transition de la seconde étape de la réaction de méthylation, à la fois pour les PRMT de type I et pour les PRMT de type II.

2. En continuité des travaux présentés dans ce chapitre, nous proposons d'utiliser des substrats entiers possédant un site de méthylation où l'arginine est modifiée par un analogue similaire à celui utilisé pour les dérivés peptidiques. Cette approche structurale pourrait permettre de capturer et de stabiliser un complexe PRMT-substrat. Pour atteindre cet objectif, deux approches pourraient être envisagées en fonction de la position des sites de méthylation :

- par ligation chimique en conditions natives (Johnson et al., 2006). Cette approche serait utilisée pour les sites de méthylation situés dans les extrémités amino- ou carboxy-terminales de la protéine substrat. Cette méthode consiste à faire réagir un peptide arborant une fonction chimique électrophile sur sa chaîne principale avec un second peptide arborant une fonction chimique nucléophile, par exemple la chaîne latérale d'une cystéine. La réaction engendre la formation d'une liaison peptidique. Nous souhaiterions utiliser cette technologie pour relier un peptide incorporant l'arginine modifiée au reste du substrat correspondant,

- par incorporation *in vivo* de résidus non naturels lors de la synthèse des protéines (Wals et al., 2014). Cette approche serait utilisée pour les sites de méthylation situés dans un cœur structuré d'une protéine substrat. Cette méthode élégante nécessite l'ingénierie génétique d'un organisme pour lui ajouter la capacité à incorporer un acide aminé non naturel ajouté au milieu de culture, dans notre cas une arginine arborant un groupement adénosine, lors de la synthèse de la protéine. Le signal d'incorporation étant un codon stop détourné. Cette méthode nécessite cependant la conception d'un couple aminoacyl-ARNt-synthétase/ARNt adapté.

En parallèle de nos travaux avec les dérivés peptidiques, nous travaillons depuis un an à l'analyse du complexe formé entre PRMT2 et son substrat entier rsf1 par cryo-microscopie électronique (travaux détaillés au chapitre III).

D - Matériel et méthodes

1 - Préparation des peptides

Les peptides ont été synthétisés par l'équipe du Dr. N. I. Martin et fournis sous forme de poudre. Les échantillons ont été dissous dans de l'eau de manière à atteindre une concentration proche de 20 mM. La concentration a été estimée par absorbance de l'adénine à 260 nm (coefficient d'absorbance moléculaire du SAH = $15400 \text{ M}^{-1} \cdot \text{cm}^{-1}$). La synthèse des dérivés de peptides H3 mène à une solution de pH très acide. Les solutions de dérivés d'H3 ont été neutralisées par l'ajout de quelques microlitres d'ammoniac à 10 %.

2 - Expression et purification des constructions de mCARM1

L'expression, la purification et la cristallisation de CARM1 sont déjà bien maîtrisée (Hassenboehler, 2010 ; Mailliot, 2013 ; Troffer-Charlier et al., 2007). Les protocoles d'expression et de purification sont indiqués dans la publication « Transition state mimics are valuable probes for structural studies with the arginine méthyltransferase CARM1 » (van Haren et al., 2017). Pour résumer, les constructions de CARM1 130-487, 130-497 et 130-507 étiquetées GST sont exprimées à l'aide du système baculovirus-cellules d'insecte dans des cultures en suspensions de cellules de *S. frugiperda*. Les protéines ont été purifiées en deux étapes de chromatographie d'affinité pour le glutathion puis de filtration sur gel.

3 - Cristallisation des complexes mCARM1-dérivés peptidiques

mCARM1 a été cristallisé en complexe avec du SAH selon la méthode de diffusion de vapeur en goutte suspendue, dans une plaque de cristallisation VDX 24 puits (Hampton™ research). Les gouttes de cristallisation ont été constituées sur des lamelles de verre 22 mm traitées à la silicone (HR3-231, Hampton® research) par le mélange de 2 μL de solution de protéine à $2 \text{ mg} \cdot \text{mL}^{-1}$, avec une concentration en peptide comprise entre 500 μM et 2 mM, et 1 μL de précipitant, constituée d'un gradient de 12 % à 22 % en PEG3350, 100 mM Tris-HCl pH 8,5, 100-300 mM de sulfate d'ammonium, ou d'une solution de 1,0 – 1,5 M de malonate di-sodique, 200 mM NaCl, 100 mM MES pH 5,5–7,0 / Tris-HCl pH 7,5. Les réservoirs de la plaque de cristallisation ont été remplis avec 0,5 mL de précipitant puis les lamelles ont été scellées sur la plaque à l'aide d'huile de paraffine (Merk™). Les cristaux ont été obtenus après une semaine à l'obscurité, 22 °C.

4 - Collecte et traitement des données

Les cristaux ont été cryo-protégés par trempage dans 5 μL de solution de réservoir supplémenté par 15 % de PEG400 puis congelés par immersion dans l'azote liquide. Les données ont été collectées sur les lignes ID30b de l'ESRF (Grenoble), Proxima 1 de SOLEIL (Saclay), équipées d'un détecteur Pilatus 6M (Dectris), ou sur l'installation de la plateforme de cristallographie de l'IGBMC. Le diffractomètre de la plateforme est équipé d'une source Rigaku FR-X à anode tournante et d'un détecteur Eiger 4M (Dectris). Les données de diffraction ont été traitées à l'aide du programme XDS et de la suite de programmes CCP4. Les cristaux appartiennent au groupe d'espace $P2_12_12$ avec un tétramère de mCARM1 par unité asymétrique. Les structures ont été résolues par remplacement moléculaire et les modèles ont été affinés avec Phenix. Les constructions graphiques ont été réalisées dans le programme Coot. L'affinement a été effectué avec utilisation de trois groupes TLS par monomère. Les figures ont été réalisées avec PyMol (Schrodinger®).

V - Étude structure-fonction du mutant CARM1 P409A

Les résidus strictement conservés sont très peu nombreux chez les PRMT, et la plupart sont directement impliqués dans le mécanisme catalytique et appartiennent aux motifs caractérisant cette famille d'enzymes (voir Alignement 4, page 160). La figure 75 représente la projection de la conservation de séquence sur le module catalytique de mCARM1. Sur cette figure, les positions les moins conservées sont colorés en bleus tandis que l'augmentation de la conservation se traduit par un décalage vers le rouge.

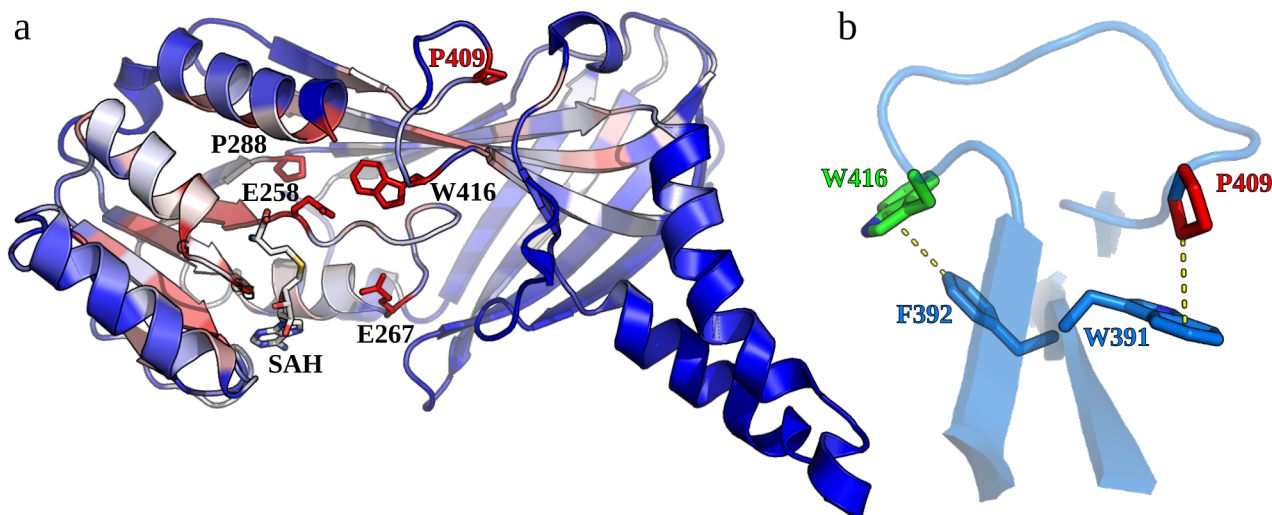
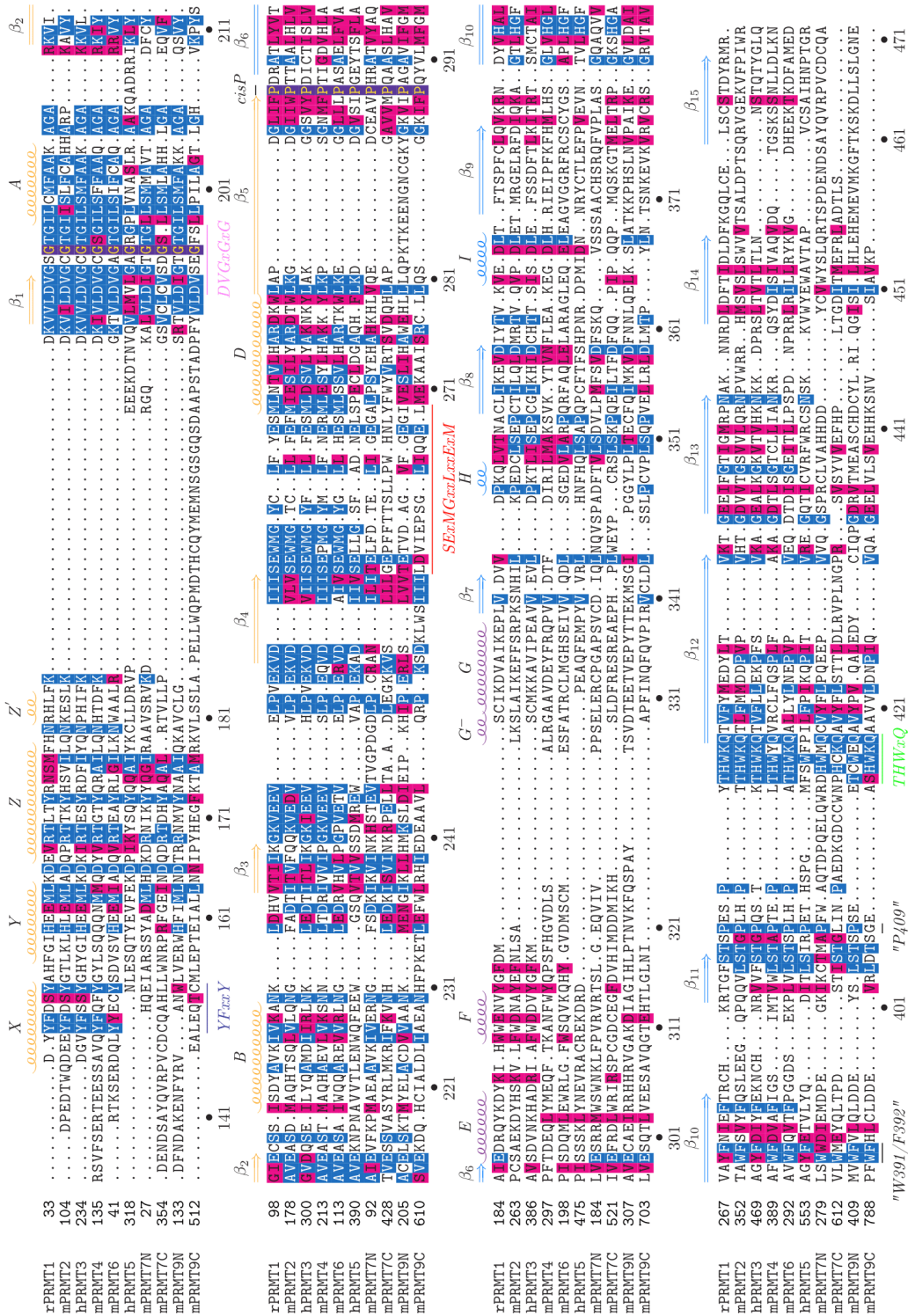


Figure 75 - Conservation de la structure primaire du module catalytique des PRMT.

(a) Module catalytique de CARM1 de souris (PDB 5IH3) coloré selon la conservation de la séquence du moins (bleu) vers le plus conservé (rouge). Une conservation moyenne correspond à une teinte blanche. Pour faciliter la visualisation, les hélices α -X et α -Y ont été masquées. Les résidus strictement conservés sont représentés en bâtonnets rouges. Les annotations correspondent à mCARM1. (b) Détails du contexte structural dans lequel se situe la proline conservée (P409 dans mCARM1). La proline est représentée en rouge. Le tryptophane du motif IV est représenté en vert. Les deux résidus aromatiques conservés sont représentés en bleu.

Comme illustré ci-dessus et dans la figure 13 page 46 de l'introduction, le bras de dimérisation est une des structures les moins conservées chez les PRMT, puisqu'elle varie fortement en taille et en structure parmi les différents membres de cette famille. Les régions les plus conservées (en blanc et nuances de rouge) correspondent aux motifs II, III et IV, tandis que le motif I n'est caractéristique que des PRMT de type I. L'interface de dimérisation sur le domaine de Rossmann, et les résidus enfouis impliqués dans le maintien du domaine de Rossmann et du tonneau β sont également plus conservés que les résidus exposés. Chez les PRMT, les résidus strictement conservés sont très rares. En effet, dans les modules catalytiques productifs, seuls huit résidus sont quasi-strictement conservés parmi l'ensemble des PRMT, dont six appartiennent aux motifs II, III et IV : les trois glycines du motif II (G193, G195 et G197 dans mCARM1) permettant de créer l'espace indispensable à la fixation du SAM, les deux acides glutamiques catalytiques (E258 et E267 dans mCARM1) du motif III et le tryptophane (W416 dans mCARM1) du motif IV, participant à l'architecture de la poche de fixation de l'arginine. Seuls deux résidus sont strictement



Alignement 4 - Alignement de séquences de quelques PRMT de type I, II et III.

Les structures de toutes les PRMT disponibles ont été utilisées, sans redondance, pour construire l'alignement. Seul un représentant par PRMT est présenté. La structure de PRMT9 n'a pas été résolue et sa séquence a été alignée sur l'alignement. L'alignement est restreint au module catalytique et les modules carboxy-terminaux de PRMT7 et PRMT9 ont été réalignés sur les domaines amino-terminaux. Les structures secondaires de mCARM1 sont montrées en haut de l'alignement. Le domaine de fixation du SAM, le tonneau β et le bras de dimérisation sont coloriés en orange, bleu clair et violet. La numérotation de la dernière ligne correspond à mCARM1. Les motifs I, II III et IV sont indiqués en bleu foncé, rose, rouge et vert. Les acides aminés sont coloriés en fonction de la similarité avec la séquence consensus : invariant (violet) ou similaire (bleu). Les abréviations utilisées sont : m/*Mus musculus*, h/*Homo sapiens*, r/*Rattus norvegicus*. Pour améliorer la lisibilité, aPRMT10 n'est pas présentée.

conservés en dehors de ceux des motifs II, III et IV : deux prolines, l'une cis et l'autre trans. La cis-proline (P288 dans mCARM1) est située à la jonction entre le domaine de Rossmann et le tonneau β . Cette proline crée un angle à environ 95° permettant de terminer le domaine de fixation du SAM et d'initier le domaine organisé en tonneau β du module catalytique. La seconde proline (P409 dans mCARM1), en configuration trans, n'a jamais fait l'objet d'études. En général, les prolines conservées adoptent le plus majoritairement une conformation cis (Lorenzen et al., 2004), comme c'est le cas de P288, et il est assez rare de trouver une proline conservée en conformation trans.

Cette proline se situe dans un tournant β de type II, juste en amont du motif IV et paraît former dans les structures tridimensionnelles le point d'ancrage d'une bouche dont la formation interrompt le brin β_{11} , donnant naissance au motif IV. Ce qui est surprenant, c'est que la présence de ce résidu semble être en corrélation avec un état productif du module catalytique. En effet, comme le montre l'alignement 4, cette proline est absente dans les modules catalytiques carboxy-terminaux des PRMT7 et PRMT9 de nombreux organismes, respectivement non-productif et supposé non-productif. Dans le cas de la PRMT9 de *H. vulgaris*, la proline est présente sur le module catalytique « non-productif » tandis que c'est le tryptophane strictement conservé du motif IV qui est muté en phénylalanine, ce qui suggère qu'il pourrait exister un lien évolutif entre ces deux résidus.

Ce résidu se différencie également des autres résidus conservés par plusieurs aspects. Tout d'abord, il est plus distant du site de méthylation que les autres résidus (plus de 18 Å dans mCARM1) et n'intervient pas directement dans la fixation du SAM et de l'arginine, contrairement à ceux des motifs I, II, III et IV. Ensuite, il se localise dans une région moins conservée du module catalytique, plutôt située sur sa surface externe, tandis que les autres résidus conservés sont enfouis ou situés dans la cavité formée entre les dimères des PRMT. Finalement, comme l'illustre la figure 75b, cette proline est systématiquement placée à la verticale du plan d'un résidu aromatique (W391 dans mCARM1) du brin β_{10} , lui-même positionné en n-1 d'un second résidu aromatique conservé (F392 dans mCARM1) impliqué dans une interaction aromatique orthogonale avec le résidu tryptophane strictement conservé du motif IV (W416) dans la poche de fixation de l'arginine.

Bien que l'ensemble des éléments présentés ci-dessus semblent indiquer une corrélation entre la présence du résidu P409 et l'intégrité fonctionnelle ou structurale du module catalytique, aucune étude n'a été menée afin de caractériser son rôle et d'expliquer sa conservation. Nous avons estimé que ce résidu pourrait être important pour l'activité des PRMT et nous avons réalisé une étude biochimique, biophysique et structurale du mutant ponctuel de P409A de cette proline chez mCARM1.

A - Résultats

1 - Analyse structurale du mutant CARM1 P409A

Le mutant P409A (130-487) a été purifié suivant le protocole mis au point pour la construction équivalente de mCARM1 non mutée (WT). Aucune altération de son comportement n'a été décelée lors des étapes d'expression et de purification, ce qui suggère que la mutation n'affecte ni le repliement de la protéine lors de sa synthèse, ni sa solubilité.

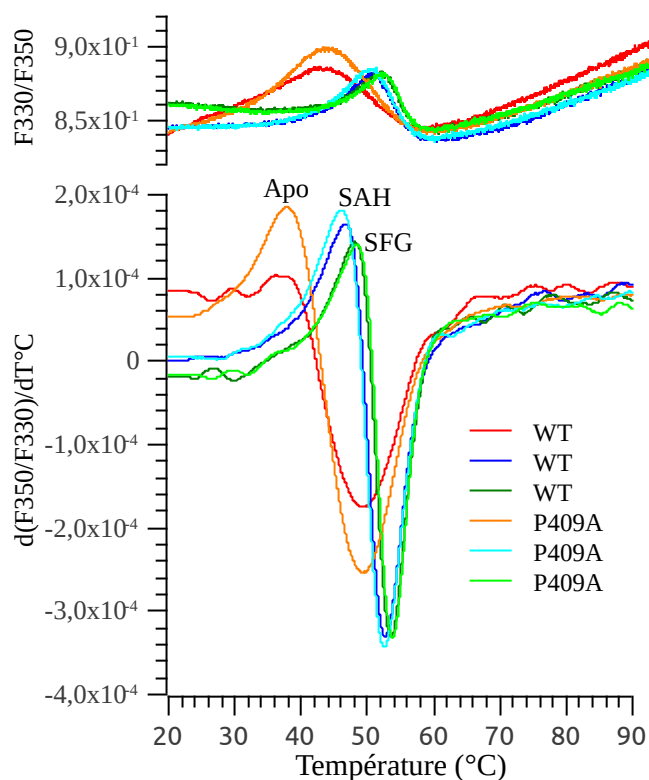


Figure 76 - Stabilité thermique de mCARM1 P409A vs WT. Détermination de la stabilité thermique de mCARM1 (130-487) WT (rouge, bleu et vert) et P409A (orange, cyan et vert pâle) sous forme apo ou fixé à du SAH ou de la SFG par fluorimétrie par balayage différentiel. En haut : rapport des fluorescences à 350 nm (aromatiques en milieu aqueux) et 330 nm (aromatiques en milieu hydrophobe). En bas : dérivée première des courbes ci-dessus.

une étude cristallographique du mutant mCARM1 (130-487) P409A. Les cristaux de mCARM1 P409A ont été obtenus dans des conditions similaires à l'enzyme WT et les données ont été collectées sur la source de rayonnement synchrotron ID30b à l'ESRF. Les statistiques de collecte et le traitement des données relatives à cette structure sont donnés à l'annexe 4. La structure cristallographique de mCARM1 (130-487) P409A, présentée sur la figure 77, a été résolue à une résolution de 2,04 Å. Comme mCARM1 WT, le mutant P409A cristallise dans le groupe d'espace $P2_12_12$ avec les paramètres de maille $a = 75 \text{ \AA}$ $b = 98 \text{ \AA}$ $c = 207 \text{ \AA}$ identiques à ceux de l'enzyme WT.

Afin de déterminer si la mutation P409A a un impact sur la stabilité de mCARM1 bien qu'elle ne semble pas avoir d'effet sur son repliement, sa résistance à la dénaturation thermique a été mesurée par DSF entre 20 °C et 90 °C. Les données obtenues, présentées sur la figure 76, montrent que mCARM1 WT et P409A ont des températures de fusion comparables à 0,2 °C près, aussi bien sous forme apo (~38 °C, orange et rouge) qu'en présence de SAH (~46 °C, bleu et cyan) ou en présence de SFG (~48 °C, vert foncé et vert clair). Cela indique que la mutation P409A a un effet négligeable, voire nul sur la stabilité de mCARM1.

Afin d'observer l'effet de la substitution de la proline 409 par une alanine sur la structure globale et locale du module catalytique de CARM1, nous avons réalisé

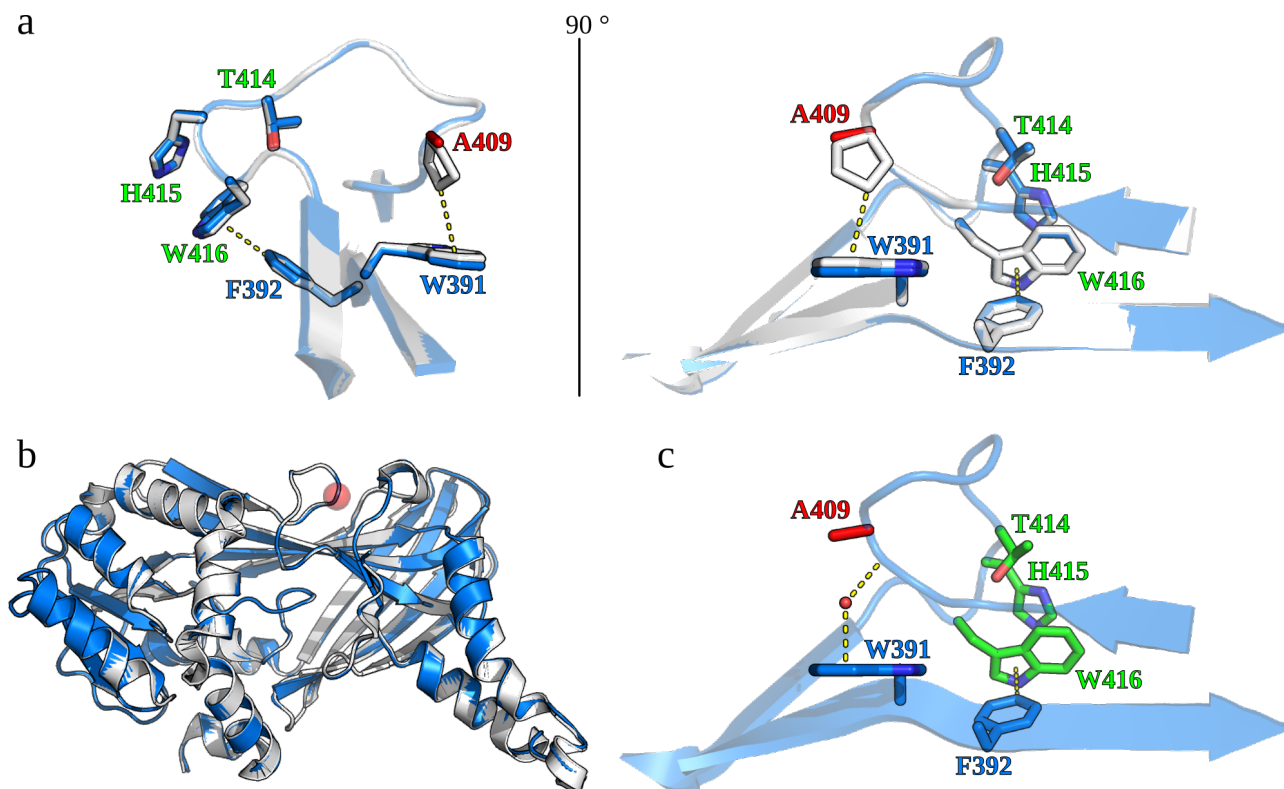


Figure 77 - Structure cristallographique du mutant mCARM1 P409A.

(a) Alignement structural local de mCARM1 WT (en blanc, PDB 5IH3) et P409A (en bleu) autour du motif IV, du résidu P/A409 et de son contexte structural. Le résidu A409 est représenté en rouge et les résidus du motif IV sont annotés en vert. (b) Alignement structural global des modules catalytiques de mCARM1 de souris WT et P409A. La sphère rouge transparente indique la position du résidu 409. Le code couleur est identique à (a). (c) Vue de profil du motif IV du mutant P409A. Une molécule d'eau (petite sphère rouge) vient compenser l'absence d'interaction P409-W391 en formant une liaison hydrogène avec l'azote de la chaîne principale de A409 et une interaction π -dipôle avec W391.

La superposition des structures des modules catalytiques de mCARM1 P409A et WT, illustrée en figure 77, montre que la mutation ne provoque aucun effet global sur la structure de mCARM1 ($\text{RMSD}_{\text{global}} = 0,20 \text{ \AA}$; $\text{RMSD}^{\text{C}\alpha}_{\text{global}} = 0,17 \text{ \AA}$). La mutation n'a pas non plus d'impact local sur la conformation de la chaîne principale autour de la position mutée ($\text{RMSD}_{405-420} = 0,12 \text{ \AA}$; $\text{RMSD}^{\text{C}\alpha}_{405-420} = 0,12 \text{ \AA}$), pas plus que sur les chaînes latérales des résidus environnants et sur le motif IV. En effet, le carbone $\text{C}\beta$ de A409 occupe pratiquement la même position que le $\text{C}\beta$ de P409, à $0,47 \text{ \AA}$ près. La perte des carbones $\text{C}\gamma$ et $\text{C}\delta$ est compensée par la présence d'une molécule d'eau qui établit les contacts manquants avec les résidus environnants.

Cette molécule d'eau, visible sur la figure 77, établit une liaison hydrogène avec l'azote de la chaîne principale du résidu 409, et fait une interaction dipolaire OH- π avec W391, ce qui remplace l'interaction CH- π établie entre les résidus P409 et W391 dans le cas de l'enzyme sauvage.

2 - Caractérisation fonctionnelle du mutant P409A de CARM1

a - Activité enzymatique sur le tétramère d'histone H3-H4

Dans le but de caractériser l'effet de la mutation P409A sur l'activité de mCARM1, les activités enzymatiques des enzymes WT et mutante ont été évaluées en utilisant le tétramère (H3-H4)₂ comme substrat. Les activités, présentées en figure 78, ont été déterminées à partir de quatre mesures de vitesses initiales (78a/b) pour des concentrations croissantes en enzyme.

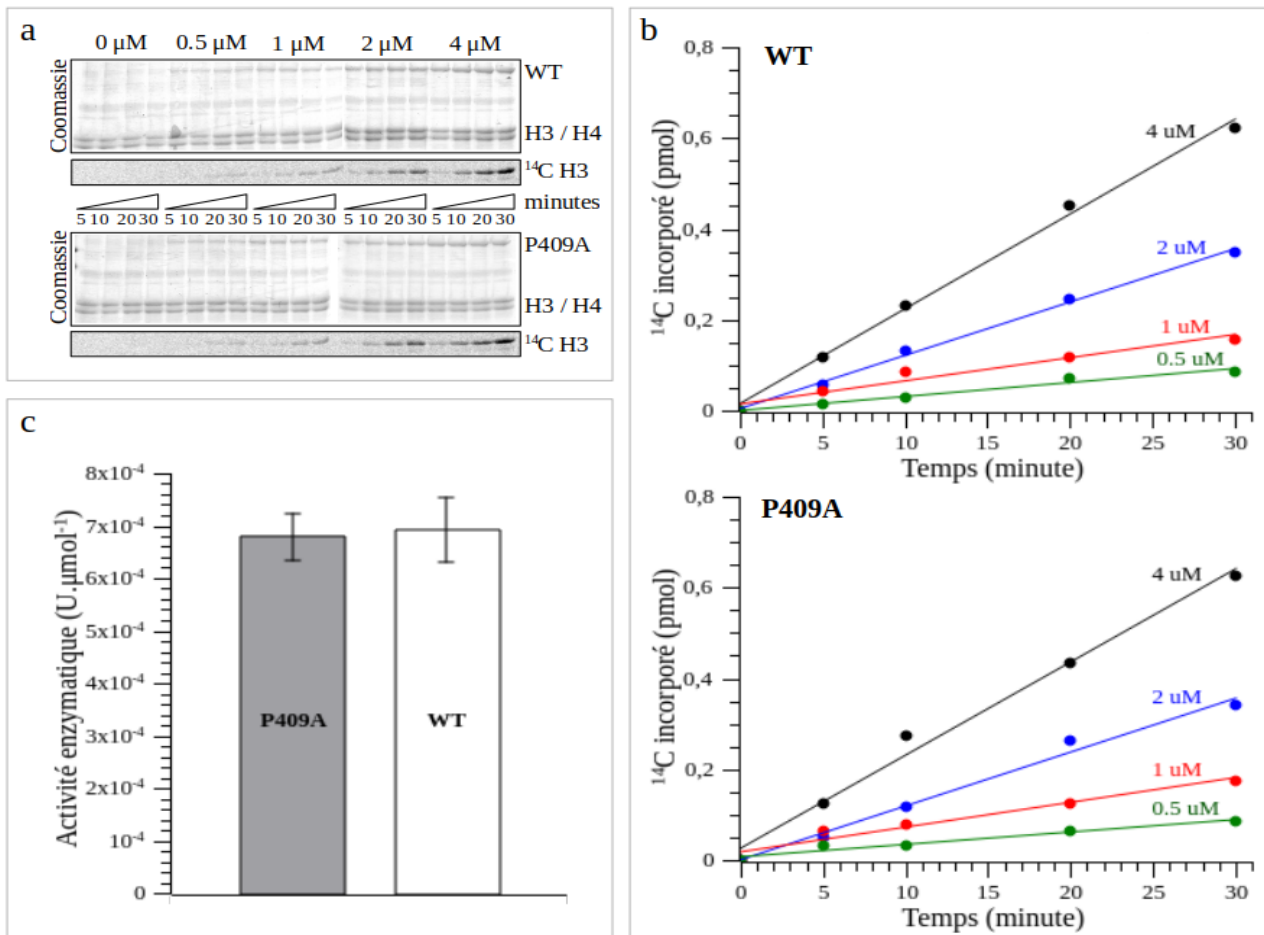


Figure 78 - Activité enzymatique de CARM1.

(a) Séparation par SDS-PAGE sur gel de polyacrylamide 16 % de mCARM1 (130-487) WT et P409A. Le gel a été coloré au bleu de Coomassie afin de révéler la présence de l'enzyme et du substrat. La radioactivité a été révélée par exposition sur un film de stockage au phosphore. (b) Détermination de la vitesse initiale de transfert de méthyle depuis le cofacteur SAM vers l'histone H3 par mCARM1 WT (haut) et P409A (bas) à différentes concentrations en enzyme. (c) Activités enzymatique de mCARM1 (130-487) WT (en blanc) et P409A (en gris) déterminées à partir des vitesses initiales.

Les résultats montrent que mCARM1 P409A est active et possède une activité enzymatique ($7,8 \cdot 10^{-4} \text{ U.}\mu\text{mol}^{-1}$) similaire à WT ($7,9 \cdot 10^{-4} \text{ U.}\mu\text{mol}^{-1}$) dans les conditions expérimentales utilisées.

b - Affinité pour le peptide PABP1 (437-488)

Afin d'étudier l'importance du résidu proline sur la fixation du substrat, les constantes de dissociation du complexe enzyme-substrat ont été déterminées par thermophorèse sur mCARM1 (130-487) WT et P409A, en utilisant comme partenaire un peptide du substrat PABP1 (résidus 437-488), fusionné à la GST. Ce fragment de PABP1 possède deux sites de méthylation

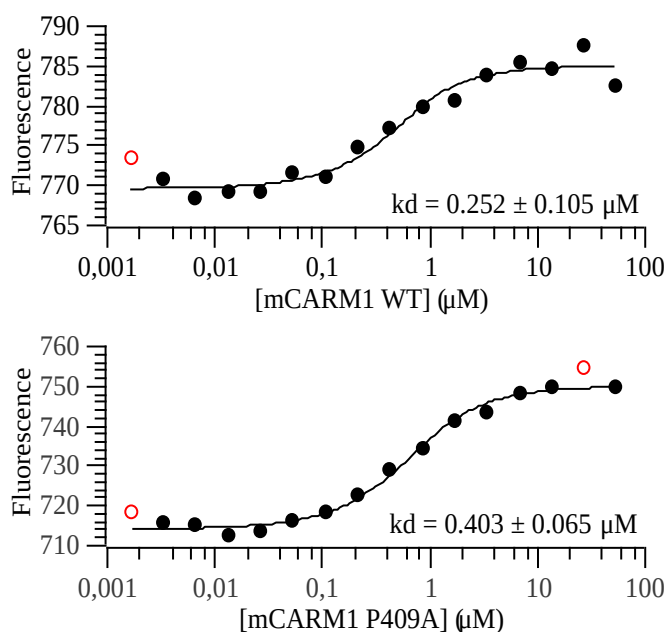


Figure 79 - Affinité de mCARM1 pour PABP1 (437-488). Détermination de la constante de dissociation (k_d) pour mCARM1 (130-487) WT (en haut) et P409A (en bas) par thermophorèse. Les cercles noirs représentent la fluorescence mesurée pour différentes concentrations en mCARM1. Les cercles rouges correspondent aux valeurs aberrantes non prises en compte pour les calculs de k_d . Les courbes correspondent au « fit » permettant la détermination du k_d et obéissant à la loi d'action de masse.

préférentiels aux positions 455 et 460 et ne possède aucune lysine susceptible de lier le marqueur fluorescent. Le marquage ne risquait donc pas d'interférer avec la reconnaissance enzyme-substrat. À l'inverse, l'histone H3 possède des lysines à proximité des sites de méthylation. Le tag GST contient lui plusieurs lysines exposées ainsi que l'extrémité amino-terminale, ciblées par le marquage. Comme le montrent les courbes de titrage du peptide GST-PABP1 par mCARM1, présentées sur figure 79, les k_d déterminés pour l'enzyme WT (en haut) et le mutant P409A (en bas) sont respectivement d'environ $0,25 \pm 0,11 \mu\text{M}$ et $0,40 \pm 0,07 \mu\text{M}$. Ces k_d sont donc du même ordre de grandeur dans les conditions expérimentales, ce qui montre que la mutation P409A n'impacte pas la fixation de CARM1 à ce peptide.

B - Discussion

À ce stade, nous avons pu montrer que la mutation du résidu strictement conservé P409 en alanine sur mCARM1 ne perturbe pas la structure tridimensionnelle de l'enzyme, tant localement autour de la mutation que globalement. La substitution d'une proline par une alanine est une mutation permise par l'évolution (avec un score de +1 sur la matrice de substitution PAM250) si bien que la mutation ponctuelle P409A est assez neutre, ce qui a permis à l'enzyme de compenser les pertes d'interactions en accommodant une molécule d'eau. Le parfait alignement local et global des structures WT et P409A montrent qu'une proline n'est pas à elle seule nécessaire à la formation du tournant β de type II donnant naissance à la boucle du motif IV, et les résidus avoisinant la position 409 pourraient également participer à stabiliser cette structure. La récurrence d'une proline à cette position pourrait simplement s'expliquer par l'avantage énergétique apporté par les résidus

prolines dans la formation des tournants β de type II (Gibbs et al., 2002). Cela n'explique cependant pas l'apparente corrélation entre la présence d'une proline et l'état productif ou non productif du module catalytique. De surcroît, la boucle post-P409 varie fortement en taille et en structure selon la PRMT, tandis que le résidu proline reste invariant. La capacité du module catalytique à absorber l'effet de la mutation P409A et la compensation par la molécule d'eau expliquent les stabilités thermiques similaires mesurées pour WT et P409A. Afin de mieux comprendre l'éventuelle importance structurale de P409, il sera nécessaire de réaliser des mutations moins neutres et probablement de muter d'autres positions ponctuellement et en synergie les unes avec les autres.

Le fait que le mutant de mCARM1 soit actif montre que P409 n'est pas strictement nécessaire pour un module catalytique productif. Cependant, la similarité des activités enzymatiques mesurées pour WT et P409A ne permet pas à ce stade de démontrer que la mutation n'a pas d'effet sur la réaction de méthylation. En effet, des variations des paramètres K_m et V_{max} pourraient expliquer une activité enzymatique similaire, et il sera nécessaire de réaliser une étude cinétique du mutant P409A pour comprendre son rôle éventuel dans la dynamique de l'enzyme. La mesure des constantes de dissociation suggère que P409 n'intervient pas non plus dans la fixation de CARM1 à son substrat. Les activités enzymatiques et les mesures des k_d ne permettent pas d'obtenir une image des mécanismes sous-jacents d'association et de dissociation du complexe enzyme-substrat, et le k_d apparent pourrait cacher un changement compensé des k_{on} et k_{off} du complexe.

De plus, ces mesures ont été réalisées sur un peptide et non sur un substrat entier, ce qui ne permet pas d'exclure que P409 pourrait intervenir dans les interactions à plus longue distance du site de méthylation du substrat. Aussi, une seule mesure a été effectuée et des répliquats devraient être réalisés afin de conforter ce résultat. Pour aller plus loin, des analyses par dynamique moléculaire pourraient être effectuées afin d'observer l'effet de diverses mutations à cette position sur la dynamique locale de la boucle et sur le site actif durant la réaction de méthylation, et le travail décrit ci-dessus devrait être élargi à d'autres PRMT appartenant aux quatre types. N.J. Zondlo a montré que les interactions proline-aromatique dans les protéines sont majoritairement formées par des liaisons électrostatiques entre le nuage d'électrons π sur les faces des résidus aromatique et les hydrogènes électropositifs de la chaîne latérale de la proline (Zondlo, 2013), ce qui est le cas de P409. Zondlo montre que ces interactions permettent de moduler l'isomérisation cis-trans de la proline par le simple ajout de 1 kcal.mol⁻¹. Dans le cas des PRMT, une telle isomérisation impliquerait un changement structural du motif IV et pourrait décrire un mécanisme commun de régulation ou de la dynamique de la méthylation par les PRMT. Pour finir, il n'est pas exclu que la région où se localise P409 et ses résidus associés W391 et F392 soit un site de fixation de partenaires des PRMT, impliquant un mécanisme commun à toutes les PRMT. L'identification des partenaires serait alors nécessaire à l'identification du rôle de P409.

C - Matériel et méthodes

L'expression, la purification et la cristallisation de CARM1 ont été décrits précédemment (Hassenboehler, 2010 ; Mailliot, 2013 ; Troffer-Charlier et al., 2007). Dans ce paragraphe, on ne détaillera que les procédures spécifiques du mutant P409A de mCARM1 (130-487). Les protocoles d'expression et de purification sont identiques à ceux indiqués dans la publication « Transition state mimics are valuable probes for structural studies with the arginine méthyltransferase CARM1 » jointe en annexe (van Haren, Marechal et al. 2017). Pour résumer, les constructions de CARM1 130-487 WT et P409A étiquetées GST sont exprimées à l'aide du système baculovirus-cellules d'insecte dans des cultures en suspensions de cellules de *S. frugiperda*. Les protéines sont purifiées avec un protocole comportant deux étapes de chromatographie d'affinité pour le glutathion puis de filtration sur gel.

1 - Mutagenèse dirigée

La mutation P409A a été introduite sur le gène de mCARM1 (130-487) par ré-amplification du plasmide complet selon le protocole suivant : 200 ng de plasmide pDEST20-Carm1-130-487 ont été amplifiés par PCR à l'aide de 0,05 U de polymérase Pfu (ThermoFisher) en présence de 10 % DMSO avec les amorces ci-dessous à 1 μ M, dans le tampon recommandé par le fournisseur de l'enzyme dans un volume final de 50 μ L.

- Sens : 5' -TGGCTATCCACAGCCGCAACAGAGCCCCTGAC-3'
- Anti-sens : 5' -GTCAGGGGCTCTGTTGCGGCTGTGGATAGCCA-3'

La PCR s'est déroulée sur 20 cycles dans les conditions suivantes : dénaturation 30 secondes à 94 °C, hybridation des amorces 30 secondes à 55 °C, polymérisation 20 secondes à 68 °C. Après PCR, le milieu réactionnel a été incubé 2 heures à 37 °C en présence de 80 U d'enzyme de restriction *DpnI* afin de digérer la matrice. Les plasmides mutés ont été transformés dans des cellules de *E. coli* DH5 α par choc thermique, puis été ré-extraits par mini-prep à l'aide du kit NucleoSpin Plasmid (Macherey-Nagel) et séquencés par méthode de Sanger par l'intermédiaire du service de séquençage GATC-Biotech/Eurofin.

2 - Expression et purification du tétramère d'histones (H3H4)₂

Des cellules *E. coli* BL21(DE3) ont été transformées par choc thermique avec un plasmide permettant la co-expression inductible d'un tétramère d'histone H3-H4 étiqueté histidine sur H4 (Diebold et al., 2011). Les cellules ont été cultivées à 37 °C dans 3 fois 1 L de milieu de culture Terrific-Broth (TB) supplémenté de 100 μ g.mL⁻¹ d'ampicilline jusqu'à une densité optique de 0,6 à 600 nm. L'expression des protéines a été induite par l'ajout de 0,2 mM d'IPTG pendant 2 heures à

37 °C. Les bactéries ont été récoltées par centrifugation pendant 15 minutes à 1000 g puis les cellules ont été resuspendues dans 100 mL de tampon de lyse (50 mM Tris-HCl pH 8, 2 M NaCl, 0,5 mM DTT) et congelées à -20 °C. La suspension cellulaire a été décongelée et les bactéries lysées par sonication (Labsonic® P Sartorius, amplitude 60, cycle 0,5, deux fois 5 minutes) à 4 °C en présence d'un cocktail d'antiprotéases (une pastille de Complete® EDTA-free par 50 mL). Le lysat a été clarifié par centrifugation 45 minutes à 42000 g puis incubé avec 3 mL de résine Ni-Sepharose 6 FF (GE Healthcare®) pendant 2 heures à 4 °C sous agitation douce, puis la résine a été lavée avec 10 volumes de NaCl 2 M, 50 mM Tris HCl pH 8, 10 mM imidazole. H3H4 a été élué dans deux fois 5 mL de tampon contenant 250 mM d'imidazole puis concentré à 1 mL sur membrane Amicon® Ultra (10 kDa). La protéine purifiée et concentrée a été clarifiée par centrifugation à 25000 g, pendant 20 minutes à 4 °C puis purifiée par chromatographie d'exclusion de taille (Superdex® 200 16/60) à 1 mL.min⁻¹ dans 2 M NaCl, 50 mM Tris-HCl pH 8, 0,5 mM EDTA, 0,5 mM DTT. Les fractions contenant H3H4 ont été rassemblées et concentrées sur membrane, puis conservées à -20 °C. La concentration en protéine a été mesurée par absorbance à 280 nm avec un coefficient d'extinction moléculaire théorique de 29750 M⁻¹.cm⁻¹ (pour un dimère de H3H4) et évaluée à 68 µM de H3, soit une absorbance de 2,03 à 280 nm.

[3 - Expression et purification de GST-PABP1 \(437-488\)](#)

Des cellules *E. coli* BL21(DE3) ont été transformées par choc thermique avec un plasmide pGEX permettant l'expression inductible du fragment de PABP1 (437-488) étiqueté à la glutathion-S-transférase (GST). Les cellules ont été cultivées à 37 °C dans 500 mL de milieu de culture LB supplémenté de 100 µg.mL⁻¹ d'ampicilline jusqu'à atteinte d'une densité optique à 600 nm (DO600) d'environ 0,6. L'expression de GST-PABP1 (437-488) a été induite par 0,5 mM d'IPTG pendant 20 heures à 20 °C. Les bactéries ont été récoltées par centrifugation 15 minutes à 1000 g puis les cellules ont été resuspendues dans 50 mL de tampon de lyse (50 mM Tris-HCl pH 8,5, 200 mM NaCl, 0,5 mM DTT) et congelées à -20 °C.

La suspension cellulaire a été décongelée et les bactéries lysées par sonication (Labsonic® P Sartorius, amplitude 60, cycle 0,5, deux fois 5 minutes), à 4 °C en présence d'un cocktail d'antiprotéases (une pilule de Complete® EDTA-free par 50 mL). Le lysat a été clarifié par centrifugation 45 minutes à 42000 g puis incubé avec 2 mL de résine Gluthatione-sepharose 4B (GSH, GE Healthcare®) pendant 4 heures à 4 °C sous agitation douce, puis la résine a été lavée avec 10 volumes de tampon de lyse. GST-PABP1 a été élué dans 5 volumes de tampon de lyse supplémenté de 50 mM de glutathion réduit puis concentré sur membrane Amicon® Ultra (10 kDa). Pendant la concentration, le tampon a été changé contre un tampon similaire contenant 50 % de glycérol. La concentration en GST-PABP1 a été mesurée par absorbance à 280 nm avec un

coefficient d'extinction moléculaire théorique de $46370 \text{ M}^{-1}.\text{cm}^{-1}$ et évaluée à $162 \mu\text{M}$, soit une absorbance de 7,52 à 280 nm. La solution protéique est stockée à $-20 \text{ }^\circ\text{C}$.

4 - [Fluorimétrie par balayage différentiel \(DSF\)](#)

Les mesures de DSF ont été effectuées sur un prometheus NT.48 (Nanotemper[®]). mCARM1 (130-487) « sauvage » (WT) et le mutant P409A ont été dilués à une concentration de $0,3 \text{ mg.mL}^{-1}$ dans 20 mM Tris-HCl pH 7,5, 50 mM NaCl, 1 mM TCEP, supplémenté avec 2 mM de SAH ou de SFG, ou un volume équivalent en tampon pour l'échantillon « apo ». Les échantillons ont été incubés à température ambiante pendant 10 minutes puis chargés dans des capillaires standard pour Prometheus NT.48 (NanoTemper[®]). La puissance du laser d'excitation du NT.48 a été ajustée de manière à obtenir entre 5000 et 20000 coups. Après équilibration 15 minutes à $20 \text{ }^\circ\text{C}$, la température a été augmentée jusqu'à $95 \text{ }^\circ\text{C}$, à $1 \text{ }^\circ\text{C}$ par minute. Les intensités de fluorescences à 330 et 350 nm ont été mesurées en temps réel.

5 - [Détermination de l'activité enzymatique de CARM1 WT et P409A](#)

Les échantillons de mCARM1 (130-487) WT et P409A ont été dilués à $0,5$, 1 , 2 et $4 \mu\text{M}$ dans 50 mM Tris-HCl pH 8, 200 mM NaCl, $0,25 \text{ mM}$ EDTA, 3 mM βME en présence de 10 mg.mL^{-1} de BSA et de $50 \mu\text{M}$ de son substrat le tétramère $(\text{H3H4})_2$. $50 \mu\text{L}$ de chaque échantillon ont été préparés. Un échantillon sans enzyme a également été préparé.

Les échantillons ont été incubés à température ambiante pendant 10 minutes puis la réaction enzymatique a été initiée par l'ajout de $1 \mu\text{L}$ de SAM radioactive (^{14}C) à $10 \mu\text{M}$ ($\mu\text{Cu}.\mu\text{mol}^{-1}$). $8 \mu\text{L}$ d'échantillons ont été prélevés aux temps 5, 10, 20 et 30 minutes et leurs réactions ont été stoppées par mélange à $3 \mu\text{L}$ de tampon de charge pour SDS-PAGE trois fois concentré.

Les échantillons ont été séparés par SDS-PAGE sur gel de polyacrylamide à 16 % afin de séparer l'enzyme, la BSA, le SAM et les histones. Le gel a été coloré au bleu de Coomassie puis séché sur papier Wattman à l'aide d'une plaque chauffante sous vide. La radioactivité a été détectée par exposition d'un film de stockage au phosphore (Fujifilm[®]) pendant 72 heures. Le film a été révélé à l'aide d'un scanner Typhoon (Molecular Dynamics[®]).

6 - [Détermination de la constante de dissociation CARM1-substrat](#)

Le tampon de GST-PABP1 a été échangé contre du tampon PBS (Phosphate Buffer Saline) à l'aide d'une colonne de dessalage Zeba 7K (Thermo[™]) afin de se débarrasser du tampon Tris et de l'agent réducteur, pouvant interférer avec la réaction de marquage chimique. Les concentrations en GST-PABP1 et en marqueur Red-NHS (réagit avec les amines primaires, NanoTemper[™]) ont été ajustées à respectivement à $20 \mu\text{M}$ et $40 \mu\text{M}$ dans du PBS. GST-PABP1 et le marqueur ont été

incubé avec un ratio 1:1 pendant 30 minutes à température ambiante à l'obscurité. GST-PABP1 et le marqueur non fixé ont été séparés par filtration sur gel sur la colonne B (NanoTemper™) équilibré avec du tampon d'analyse (50 mM Tris-HCl pH 8, 200 mM NaCl, 3 mM β ME, 0,25 mM EDTA, 0,05 % Tween-20), puis les fractions contenant la protéine ont été clarifiées par centrifugation pendant 5 minutes à 25000 g à 4 °C. La concentration a été estimée à 3,9 μ M par mesure de l'absorbance à 280 nm.

Le tampon des échantillons de mCARM1 (130-487) WT et P409A a été échangé contre le tampon d'analyse à l'aide d'une colonne de dessalage Zeba® tel qu'indiqué ci-dessus puis les échantillons ont été concentrés sur membrane Amicon® 30 kDa, 0,5 mL (Merk™) et clarifiés par centrifugation 10 minutes à 25000 g à 4 °C. Les concentrations ont été évaluées à 68 μ M (P409A) et 151 μ M (WT) tel qu'indiqué pour GST-PABP1.

Une solution de mCARM1 WT/P409A, GAT-PABP1 et de SFG a été préparée pour obtenir les conditions suivantes : 0,5 mM SFG, 0,5 μ M GST-PABP1, mCARM1 de 54 μ M à 1,65 nM avec une dilution d'un facteur 2, pour un volume final de 20 μ L par échantillon. Les échantillons ont été chargés par capillarité dans des capillaires Premium (Cat# MI-K005) pour Monolith® NT-115 (Nanotemper®). La thermophorèse a été effectuée sur un Monolith® NT-115 à 22 °C, avec une intensité du laser d'excitation de 25 % et des intensités du laser de thermophorèse à 30 %, 50 % et 80 % pour P409A et 25 %, 50 % et 65 % pour WT. Les durées de l'état initial, la diffusion et la contre-diffusion ont été paramétrées respectivement à 5, 30 et 5 secondes. Les k_d ont été évalués sur l'étape de diffusion.

7 - Étude structurale de mCARM1 (130-487) P409A

a - Cristallisation

mCARM1 (P409A) a été cristallisé en complexe avec du SAH selon la méthode de diffusion de vapeur en goutte suspendue, dans une plaque de cristallisation VDX 24 puits (Hampton™ research). Les gouttes de cristallisation ont été constituées sur des lamelles de verre 22 mm traitées à la silicone (HR3-231, Hampton® research) par le mélange de 2 μ L de solution de protéine à 2 mg.mL⁻¹, SAH à 2 mM et 1 μ L de précipitant, constituée d'un gradient de 12 % à 22 % en PEG3350, 100 mM Tris-HCl pH 8, 100 mM NaCl. Les réservoirs de la plaque de cristallisation ont été remplis avec 0,5 mL de précipitant puis les lamelles ont été scellées sur la plaque à l'aide d'huile de paraffine (Merk™). Les cristaux ont été obtenus après une semaine à l'obscurité, 22 °C.

b - Collecte et traitement des données

Les cristaux ont été cryo-protégé par trempage dans 5 μ L de solution de réservoir supplémenté par 15 % de PEG400 puis congelés par immersion dans l'azote liquide. Les données ont été collectées sur la ligne ID30b de l'ESRF (Grenoble), équipée d'un détecteur Pilatus 6M (Dectris). Les données de diffraction ont été traitées à l'aide du programme XDS et la suite de programmes CCP4. Les cristaux appartiennent au groupe d'espace $P2_12_12$ avec deux dimères de mCARM1 par unité asymétrique. La structure a été résolue par remplacement moléculaire et le modèle a été construit et affiné avec le programme Phenix. Les constructions manuelles ont été réalisées avec le programme Coot. L'affinement a été effectué avec utilisation de trois groupes TLS par monomère. Les figures ont été réalisées avec PyMol (Schrodinger[®]). Le tableau résumant les statistiques de collecte et le traitement des données se trouve à l'annexe 4.

VI - Conception d'inhibiteurs des PRMT

A - Introduction

Un historique des inhibiteurs des PRMT1, PRMT2, PRMT3, CARM1, PRMT5 et PRMT6 a été présenté dans le chapitre I de cette thèse. Il n'existe à l'heure actuelle aucun composé modulateur de PRMT7, PRMT8 et PRMT9. Malgré un site actif très conservé parmi les PRMT, des variations subtiles de la géométrie et des propriétés physico-chimiques des poches de fixation du SAM et de l'arginine ont permis de concevoir des composés qui inhibent spécifiquement et sélectivement ces enzymes *in vitro*. Le principal obstacle au développement d'inhibiteurs des PRMT, et en particulier des composés dérivants du SAM provient de leur difficulté à franchir les membranes cellulaires pour agir sur leurs cibles, et un nombre limité de ces inhibiteurs ont été testés sur des lignées cellulaires saines et cancéreuses afin de démontrer leur efficacité *in vivo*.

Les travaux initiaux de l'équipe sur le développement d'inhibiteurs des PRMT ont débutés en 2008 lors d'une collaboration avec l'équipe du Dr. Alain Wagner. L'idée originale n'était alors pas de développer des composés à action pharmacologique mais de concevoir des outils moléculaires permettant d'ancrer des dérivés peptidiques aux PRMT, dans le but d'étudier les bases structurales de la reconnaissance PRMT-substrat. Pour déterminer les fonctions chimiques qui permettraient de relier un analogue du SAM, ancré à son site de fixation, à une partie peptide, censée interagir avec le module catalytique, divers dérivés du SAM arborant des structures chimiques mimant l'arginine substrat ont été développés (Ajebbar, 2012 ; Mailliot, 2013). Ces composés intermédiaires ont démontré une capacité à inhiber sélectivement CARM1 et de nombreux inhibiteurs ont été conçus entre 2009 et 2013. Notre démarche actuelle s'inscrit dans la continuité de ce projet.

Nous sommes aujourd'hui impliqués dans l'effort collectif de développement de composés à action pharmacologique actifs sélectivement sur quatre cibles : PRMT2, PRMT4/CARM1, PRMT6 et PRMT7. Dans ce but, nous collaborons avec quatre équipes de chimistes européennes : les équipes du Dr. Alain Wagner (faculté de pharmacie de Strasbourg), du Dr. Paola Arimondo (ETaC Toulouse, puis Institut Pasteur à Paris), du Dr. Nathaniel I. Martin de (université d'Utrecht, Pays-Bas) et du Pr. Gianluca Sbardella, (université de Salernes, Italie). Nous utilisons une approche par conception rationnelle d'inhibiteurs à partir des structures cristallographiques, en utilisant des molécules dérivées du SAM comme sous-structures initiales.

Au cours de cette thèse, nous avons travaillé sur 4 familles de dérivés du SAM : Les composés LH, U, EML et SATT. Le travail sur les composés LH sera traité dans un premier chapitre. Les résultats ont donné lieu à une publication en 2018, « Hijacking DNA

méthyltransferase transition state analogues to produce chemical scaffolds for PRMT inhibitors » (Halby et al., 2018). Cet article est joint à ce manuscrit dans la section « Articles joints et communications scientifiques ».

Dans un deuxième paragraphe, les bases structurales de l'inhibition de PRMT2, CARM1 et PRMT6 par les composés U et EML seront présentées. Le dernier paragraphe de ce chapitre traitera du développement d'une famille d'inhibiteurs, nommés SATT, qui explore un site inédit du module catalytique des PRMT. Ces composés SATT s'inscrivent dans la continuité de la collaboration existante depuis de nombreuses années avec l'équipe du Dr. Alain Wagner à Strasbourg.

B - Résultats

1 - [Les composés LH](#)

Ce travail ayant déjà donné lieu à la publication jointe en annexe (Halby et al., 2018), seuls quelques points principaux seront développés ici. Les composés LH ont été initialement développés comme inhibiteurs des méthyltransférases de l'ADN par l'équipe du Dr. Paola Arimondo. Ces composés ont été conçus pour mimer l'état de transition établi lors de la réaction de méthylation de la cytosine par les méthyltransférases SAM-dépendante, et sont constitués d'une sous-structure analogue du SAM reliée à une fonction chimique mimant une base cytosine. Sept de ces molécules ont été synthétisées, deux d'entre elles possèdent la partie méthionine du SAM, les cinq autres ne disposant que de la partie adénosine. Bien que la conception de ces molécules les prédisposait à être des inhibiteurs des méthyltransférases de l'ADN et de l'ARN, ces composés se sont révélés totalement inactifs ou très faiblement actifs lors de tests d'inhibition réalisés *in vitro* sur les méthyltransférases de l'ADN DNMT1 et DNMT3A, et sur un jeu de huit méthyltransférases de l'ARN humaines et virales.

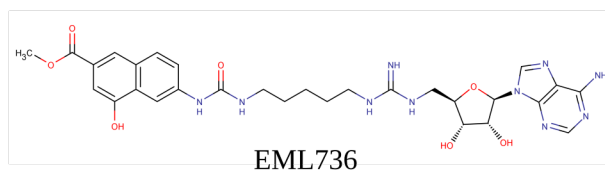
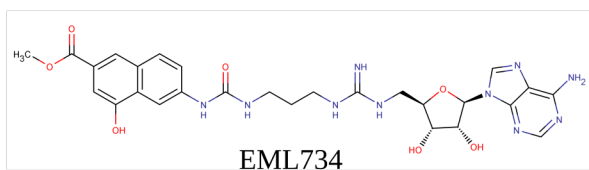
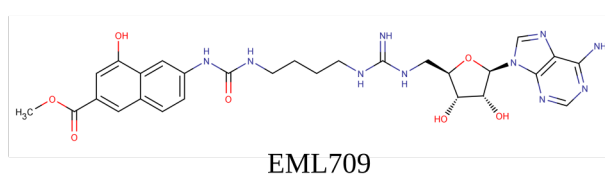
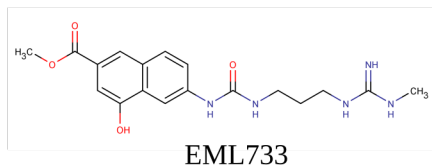
La ressemblance de la structure de ces molécules avec les inhibiteurs des PRMT préalablement développés au sein de notre équipe nous a poussés à proposer qu'ils pourraient inhiber les PRMT. L'inhibition de l'activité enzymatique de PRMT1, PRMT4/CARM1, PRMT6 et PRMT7 par les sept composés LH a été testée *in vitro* en premier lieu au sein de l'équipe puis auprès d'un prestataire sur un jeu plus large d'enzymes. Ces tests ont démontré que les composés LH sont de potentiels inhibiteurs sélectifs de CARM1 et PRMT7, et dans une moindre mesure de PRMT6. Notamment les composés 2 (LH1458) et 4 (LH1452) inhibent CARM1 avec des IC_{50} respectives de 14 μ M et 1,5 μ M. CARM1 est impliquée dans la propagation de plusieurs cancers, dont les cancers du sein, de la prostate et du colon. Afin de mesurer l'efficacité des composés 2 et 4 *in situ*, des lignées cellulaires cancéreuses modèles de ces trois cancers ont été cultivées en présence de ces inhibiteurs. Aucune altération de la prolifération cellulaire n'a pu être observée avec ces deux inhibiteurs, pas plus que l'inhibition de la méthylation de H3R17, substrat de CARM1, en présence

du composé 4. L'absence d'effet cellulaire est probablement le fait de la faible capacité de ces composés à traverser les membranes cellulaires, tel que démontré par des tests de perméabilité cellulaire sur le composé 4.

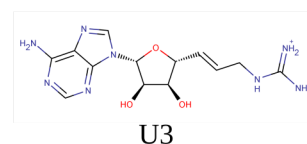
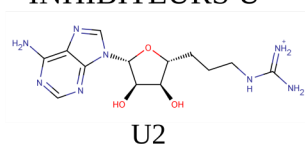
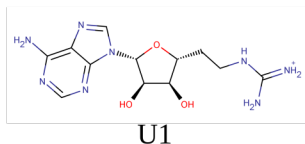
Dans le but de comprendre les bases structurales de l'inhibition des PRMT par les inhibiteurs LH, et donc de permettre leur optimisation, des essais en co-cristallisation avec CARM1, PRMT6 et PRMT7 de souris ont été réalisés. Quatre structures de mCARM1 (130-487) en complexe avec les composés 1 (LH1236), 3 (LH1246), 4 et 7 (LH1427) ont été résolues (PDB : 5TBH, 2,34 Å ; 5LV2, 2,29 Å ; 5TBJ, 2,32 Å ; 5TBI, 2,29 Å) ainsi que deux structures de mPRMT6 en complexe avec les composés 1 et 2 (PDB : 5LV4, 1,65 Å ; 5LV5, 1,80 Å). En revanche, CARM1 n'a pas pu être co-cristallisée avec le composé 2, et PRMT7 n'a pu être co-cristallisée avec aucun composé LH. Les structures des complexes PRMT-LH montrent que les composés se fixent selon le mode de fixation attendu, c'est-à-dire avec la partie analogue du SAM occupant le site de fixation du cofacteur et la structure mimant la cytosine qui se localise dans le site de fixation de l'arginine. Pour CARM1, les composés peuvent être classés dans deux catégories : ceux déplaçant totalement le SAM (composé 4) qui occupent la totalité des sites actifs dans les structures cristallographiques, et les composés déplaçant partiellement le SAM (composés 1, 3 et 7) qui occupent la moitié des sites actifs. Le composé 4 est celui possédant la plus forte inhibition pour CARM1 ce qui suggère une corrélation entre le taux d'occupation des sites de fixation et l' IC_{50} des inhibiteurs. Pour ce composé, l'inhibition est également corrélée avec le déplacement de cinq résidus conservés chez CARM1 : N159, M163, R169, E258 du motif III et H415 du motif IV. Ces déplacements ne sont pas ou peu présents avec les composés 1, 3 et 7.

En l'état, bien que les composés LH n'ont pas été optimisés pour les PRMT, ils permettent d'atteindre des IC_{50} de l'ordre du micromolaire pour CARM1. Les données structurales permettront donc d'améliorer significativement leurs affinités et sélectivités pour ces enzymes. Aussi, la biodisponibilité de ces composés devra être optimisée pour qu'ils soient actifs sur les cellules cancéreuses. Cette étude démontre la faisabilité de détourner des inhibiteurs des méthyltransférases des acides nucléiques pour la conception d'inhibiteurs des PRMT.

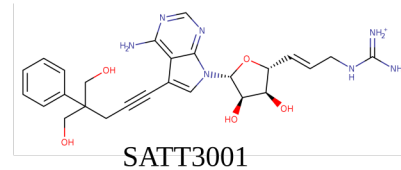
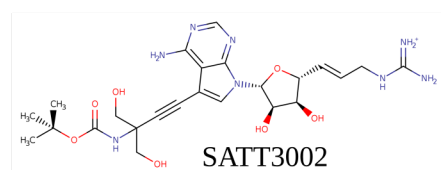
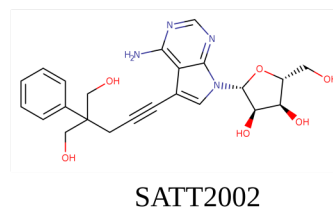
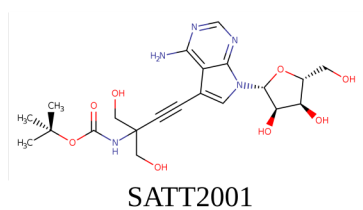
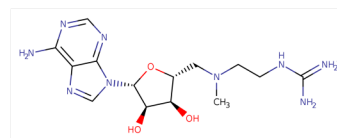
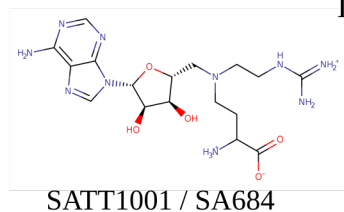
INHIBITEURS EML



INHIBITEURS U



INHIBITEURS SATT



Composés SATT testés *in silico*

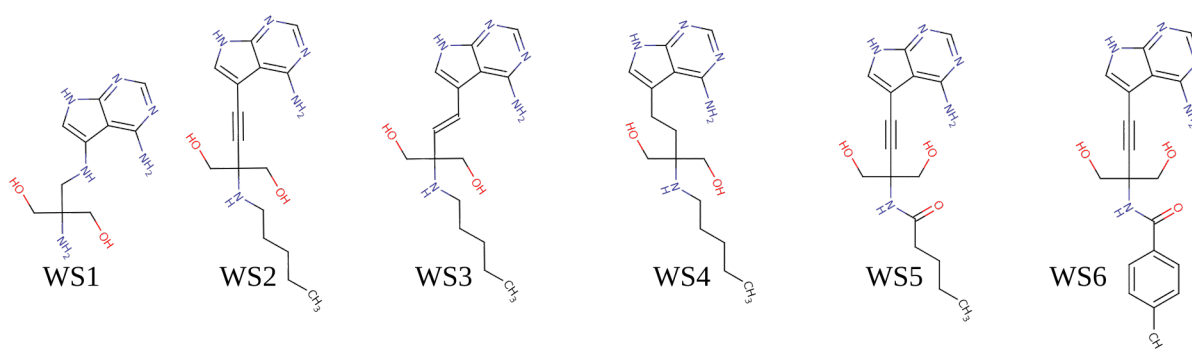


Figure 80 - Représentation schématique des composés EML, U et SATT.

2 - Les composés U

Les composés notés U (U1, U2 et U3) ont été conçus par le Dr. Nathaniel I. Martin et ses collaborateurs pour inhiber l'activité enzymatique des PRMT par compétition avec à la fois le cofacteur SAM et l'arginine substrat (van Haren et al., 2015). Pour cela ils ont synthétisé des ligands analogues du SAM, mimant l'état de transition de la réaction de méthylation et ainsi constitués d'une partie adénosine, liée par le carbone 4' à une extension aliphatique arborant un groupement guanidinium. Les structures chimiques de ces inhibiteurs sont représentées sur la figure 80. Des composés de même type avaient été présentés dans la thèse de Justine Mailliot, mais ceux-ci étaient pourvus du groupement méthionine du SAM. Le tableau 12 résume les IC_{50} mesurées par N. I. Martin et ses collaborateurs pour ces composés et le SAH contre PRMT1, CARM1, PRMT6 et la méthyltransférase G9a.

PRMT	SAH	U1	U2	U3
PRMT1	6,21 ± 0,56	11,09 ± 2,77	1,30 ± 0,38	16,96 ± 3,73
CARM1	0,67 ± 0,19	0,12 ± 0,02	0,56 ± 0,25	0,15 ± 0,05
PRMT6	0,20 ± 0,25	20,23 ± 8,67	0,72 ± 0,33	5,15 ± 1,27
G9a	16,64 ± 6,43	> 50	> 50	> 50

Tableau 12 - Inhibition (IC_{50}) des PRMT par les composés U.

Le tableau de données est issu de l'article de M. van Haren (van Haren et al., 2015). Les IC_{50} sont indiquées en μM et ont été mesurées par la méthode basée sur des anticorps anti-diméthylarginine chimioluminescents (BPS Bioscience).

Ces composés montrent une spécificité beaucoup plus forte pour les PRMT que pour la méthyltransférase G9a et possèdent des IC_{50} s'échelonnant de 20 à 0,12 μM pour PRMT1, CARM1 et PRMT6. De plus, U1 est environ 100 et 200 fois plus sélectif pour CARM1 que pour respectivement PRMT1 et PRMT6. U3 est environ 110 et 35 fois plus sélectif pour CARM1 que pour respectivement PRMT1 et PRMT6. U2 est assez peu sélectif entre ces trois enzymes.

Dans le cadre de notre collaboration avec l'équipe de N. I. Martin, nous avons approfondi la caractérisation de l'interaction entre ces composés et PRMT2, CARM1, PRMT6 et PRMT7 par des méthodes biophysiques et structurales. Afin de mesurer directement la fixation de ces composés aux PRMT, des mesures de stabilisation thermique ont préalablement été réalisées. Le tableau 13 résume les variations de T_m des PRMT en présence des inhibiteurs U, de SAH ou sous forme apo. Une variation significative du T_m comparé à la forme apo montre la présence d'interactions stabilisatrices ou déstabilisatrices. Deux méthodes ont été utilisées : la DSF et le thermofluor. La mesure de stabilité thermique était historiquement réalisée par thermofluor au sein de l'équipe. Avec cette méthode, la dénaturation thermique des protéines est mesurée par le gain de fluorescence d'un fluorophore interagissant avec les parties hydrophobes de la protéine au cours de sa dénaturation. Les mesures d'interaction PRMT-ligand ont donc été réalisées avec l'une ou l'autre méthode selon

que l'expérience ait permis d'obtenir un résultat interprétable ou non. En effet, la mesure de stabilité par thermofluor ou DSF peut fonctionner ou échouer selon le couple PRMT-inhibiteur.

	Méthode	Ligand	ΔT_m (°C) vs apo	ΔT_m (°C) vs SAH
rPRMT1	Thermofluor	U1	+ 8,0	- 2,5 (SAH)
		U2	données non disponibles	
		U3	données non disponibles	
mPRMT2	Thermofluor	U1	+ 15,0	+ 9,6 (SAH)
		U2	+ 12,2	+ 6,8 (SAH)
		U3	+ 9,7	+ 4,3 (SAH)
SmPRMT3	DSF	U1	+14,3	+ 6,0 (SAH)
		U2	+ 17,2	+ 8,9 (SAH)
		U3	+ 12,7	+ 4,4 (SAH)
mCARM1	Thermofluor	U1	+ 7,0	+ 2,5 (SAH)
		U2	+ 6,5	+ 1,0 (SAH)
		U3	+ 6,0	+ 0,5 (SAH)
mPRMT6	DSF	U1	+ 19,5	+ 8,2 (SAH)
		U2	précipité	
		U3	+ 1,0	- 13,1 (SAH)

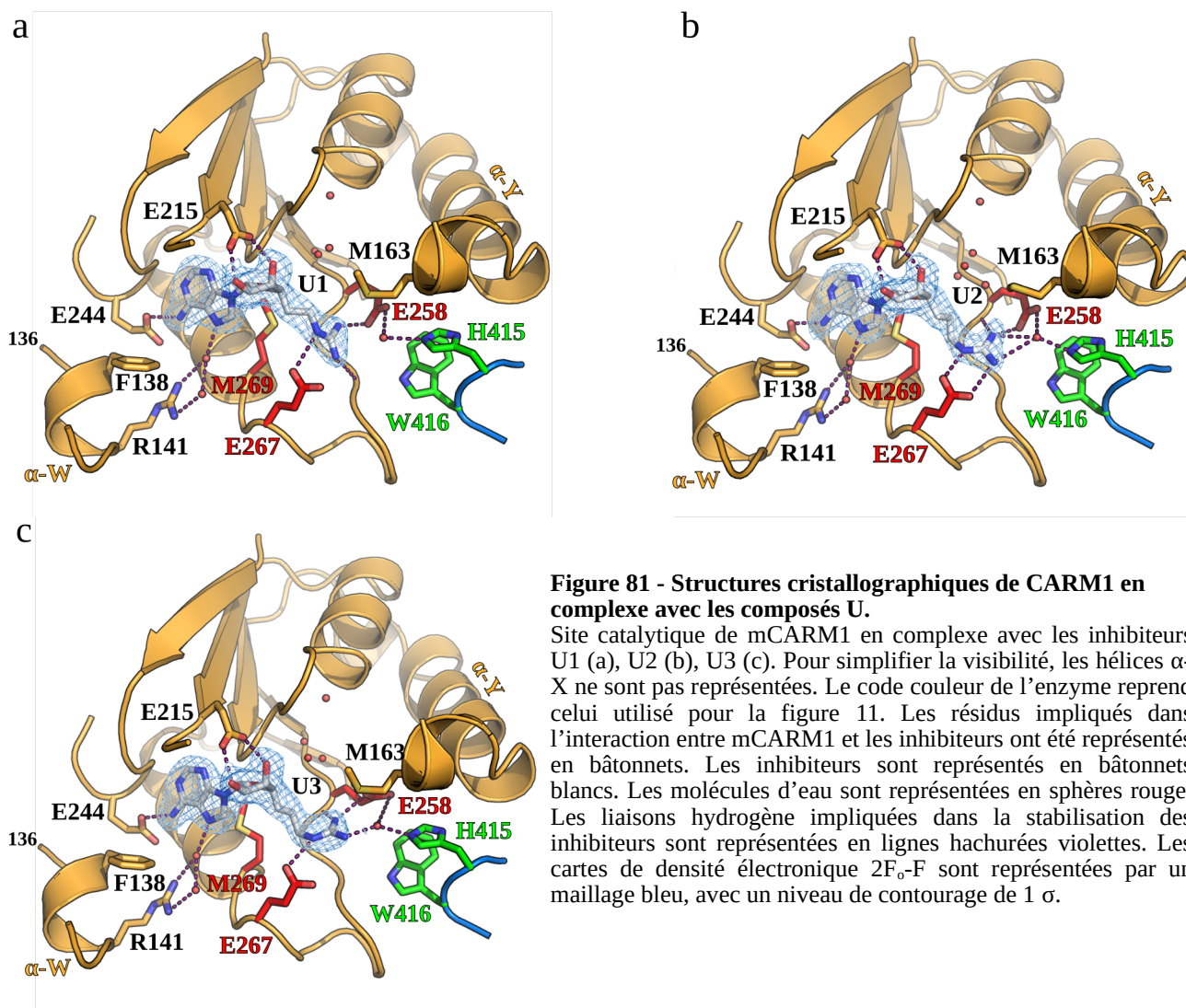
Tableau 13 - Stabilisation thermique des PRMT par les composés U.

Pour presque tous les composés, un effet plus ou moins important de stabilisation/déstabilisation thermique a pu être mesuré ce qui indique qu'ils interagissent directement avec les enzymes testées. Pour PRMT1, seul le composé U1 a pu être testé et un effet stabilisateur a été mesuré, néanmoins inférieur à celui induit par le SAH. Pour PRMT2, un fort effet stabilisateur a été mesuré avec les trois U comparé à la forme apo ou fixée à SAH, dans l'ordre U1>U2>U3. Les trois inhibiteurs stabilisent fortement SmPRMT3 comparé à la forme apo et au SAH, avec U2 ayant l'effet le plus important. Pour mCARM1 les trois inhibiteurs ont un effet stabilisateur comparable par rapport à la forme apo et fixée au SAH, avec un ordre de stabilisation U1>U2>U3, en quasi-corrélation avec les IC₅₀ publiées. Finalement pour PRMT6, U1 a un très fort effet stabilisateur comparé à la forme apo et au SAH, tandis que U3 n'a pratiquement aucun effet et U2 entraîne la précipitation de l'enzyme. Ceci n'est pas en corrélation avec les IC₅₀ publiées puisque U3 (20 µM) a une IC₅₀ quatre fois inférieure à U1 (5 µM). U2 inhibe PRMT6 (0,7 µM), ce qui dans nos mains s'explique surtout par la précipitation de la protéine lors de ce test.

Afin de déterminer les bases structurales de la sélectivité des composés U pour les PRMT, des essais de co-cristallisation avec mPRMT2 entier (fl), mCARM1 (130-487), mPRMT6 (fl et 34-378) et mPRMT7 (fl) ont été réalisés. Les structures de CARM1 et PRMT2 en complexe avec les trois U, et de PRMT6 en complexe avec U1 et U2 ont pu être résolues. Les données collectées avec PRMT7 ont révélé que les structures contiennent un SAH intrinsèque mais pas les inhibiteurs. Les statistiques de collecte et de traitement des données sont disponibles aux annexes 12, 13 et 14.

α - CARM1

La figure 81 présente les structures cristallographiques des complexes CARM1-composés U et dévoile les interactions établies entre les inhibiteurs et le site actif de la PRMT. Le détail de ces interactions, les identifiants PDB des structures et leurs résolutions sont résumés dans le tableau 14.



Dans le tableau 14 ainsi que dans les tableaux équivalents qui suivront, les interactions de types « van der Waals », aromatiques et les liaisons hydrogène établies entre les inhibiteurs et le site actif des PRMT ont été détaillées. Les scores indiqués dans les colonnes « contacts » ont été calculés à l'aide de la formule du potentiel de Lennard-Jones, et donnent une indication de l'enfouissement de ces inhibiteurs. Ces scores prennent en compte les molécules d'eau. Dans la colonne « liaisons hydrogène », les astérisques indiquent le nombre de liaisons établies. Les liaisons hydrogène dites fortes, pour lesquelles l'atome donneur, l'hydrogène et l'accepteur sont alignés, sont indiquées par un astérisque bleu. Comme la partie adénosine de ces inhibiteurs fixe les PRMT

de manière similaire au cofacteur SAM, les interactions présentées dans les tableaux ne concernent que les structures non-analogues du SAM des inhibiteurs.

mCARM1 (130-487)-U				
Inhibiteur	Structure	contacts	aromatiques	Liaisons H
U1	5K8V / 2,25 Å	Y150 M269 P259 G261 + 3 x HOH score : -0,332	∅	E267** E258** M260*
U2	5K8W / 2,10 Å	P259 M269 Y150 + 2 x HOH score : -0,197	∅	E258** E267** M260*
U3	5K8X / 1,99 Å	M163 M260 M269 Y150 score : -0,215	∅	Y154* E267* E258** HOH*

Tableau 14 - Structures cristallographiques et interactions de mCARM1 avec les composés U.

Dans la colonne « contacts », les scores ont été calculés avec la formule du potentiel de Lennard-Jones. Dans la colonne « liaisons H », les astérisques noirs et bleus correspondent respectivement à des liaisons hydrogène faibles et fortes.

Les composés U se fixent dans les quatre modules catalytiques de l'unité asymétrique dans les structures cristallographiques de mCARM1, ce qui est associé à une forte inhibition. En effet, mCARM1 contient intrinsèquement le cofacteur SAH, qui doit donc être chassé par les inhibiteurs pour que ceux-ci soient visibles dans les cartes de densité électronique. Comme l'illustre la figure 81, la partie adénosine des U occupe le site de fixation du SAM tandis que l'extension aliphatique-guanidinium se fixe à la poche de fixation de l'arginine. Les interactions établies entre la partie adénosine et le site actif de CARM1 sont similaires à celles établies avec le SAM et ne seront donc pas détaillées. Les structures montrent que la longueur et l'état de saturation du « bras » aliphatique séparant la partie adénosine du guanidinium a un impact sur les interactions établies entre le guanidinium et le site de fixation de l'arginine, ce qui module la sélectivité de ces inhibiteurs. Les densités électroniques couvrent la totalité des inhibiteurs ce qui témoignent d'une fixation efficace de ces composés. Ceci est en accord avec les IC₅₀ sub-micromolaires mesurées. U1 et U3 ont des IC₅₀ similaires de l'ordre de 150 nM, tandis que U2 est environ quatre fois moins actif.

La structure de U2 lui permet de positionner le groupement guanidinium de manière similaire à l'arginine substrat, par établissement de cinq liaisons hydrogène, dont deux liaisons fortes avec le résidu « ancre » (E267), deux liaisons hydrogène faibles avec la « pompe » (E258) et une liaison hydrogène faible avec l'oxygène de la chaîne principale de M260. Les composés U1 et U3 sont également capables d'établir cinq liaisons hydrogène dont trois fortes et possèdent des « scores » de plus faible énergie que U2, comme indiqué dans le tableau 14. Cela témoigne d'interactions plus étroites entre CARM1 et U1/U3 comparé à U2. U1 et U3 sont donc de meilleurs inhibiteurs de mCARM1 que U2 car ils établissent des liaisons hydrogène de plus faible énergie et sont plus étroitement liés à la poche de fixation de l'arginine. Cependant, comme le guanidinium de U2 se positionne de manière plus similaire à l'arginine substrat dans mCARM1 comparé à U1 et

U3, la structure de ce composé a servi de base à la conception des composés présentés dans le chapitre traitant des dérivés peptidiques.

L'analyse ci-dessus montre que le calcul de scores énergétiques de contacts et l'évaluation qualitative des liaisons hydrogène établies entre les PRMT et les inhibiteurs permet de corréliser aisément les données structurales aux IC_{50} mesurées.

b - PRMT2

La structure de mPRMT2 a été déterminée au cours de cette thèse (Cura et al., 2017). Cette PRMT est peu caractérisée et les études identifiant les cibles de PRMT2 et les pathologies dans lesquelles cette protéine est impliquée sont récentes. Jusqu'à présent aucun inhibiteur ciblant cette PRMT n'avait été décrit.

Comme aucune IC_{50} n'avait été déterminée pour les composés U sur PRMT2 par N. I. Martin et ses collaborateurs, les IC_{50} ont été mesurées par radiométrie au laboratoire. Pour cela, nous avons utilisé rsf1, une protéine d'insecte issue du système d'expression de PRMT2, et efficacement méthylée par PRMT2 (Cura et al., 2017). L' IC_{50} de U1 a été estimée à 16,3 μ M contre 18,3 μ M pour le SAH (Cura et al., 2017), ce qui fait de ce composé le tout premier inhibiteur décrit pour PRMT2. Les IC_{50} de U2 et U3 pour mPRMT2 n'ont pas pu être estimées avec une confiance suffisante car des valeurs significativement différentes ont été mesurées lors de trois tests d'activité indépendants. Cependant, les données de stabilité thermique semblent indiquer que ces deux composés fixent efficacement mPRMT2, d'un ordre de grandeur presque comparable à U1.

La figure 82 présente les structures cristallographiques des complexes PRMT2-composés U et dévoile les interactions établies entre les inhibiteurs et le site actif de la PRMT. Le détail de ces interactions, les identifiants PDB des structures et leurs résolutions sont résumés dans le tableau 15.

À nouveau les structures cristallographiques illustrées sur la figure 82 montrent que la partie adénosine des U occupe le site de fixation du SAM tandis que l'extension aliphatique-guanidinium se fixe dans la poche de fixation de l'arginine. Les interactions établies entre la partie adénosine et le site actif de PRMT2 sont similaires à celles établies avec le SAM et ne seront donc pas détaillées. Les densités électroniques sont visibles pour l'ensemble de la structure des composés U1 et U2 ce qui témoigne d'une fixation efficace de ces inhibiteurs à l'enzyme. À l'inverse, la densité électronique autour du noyau guanidinium de U3 est diffuse, ce qui suggère que ce groupement est mobile.

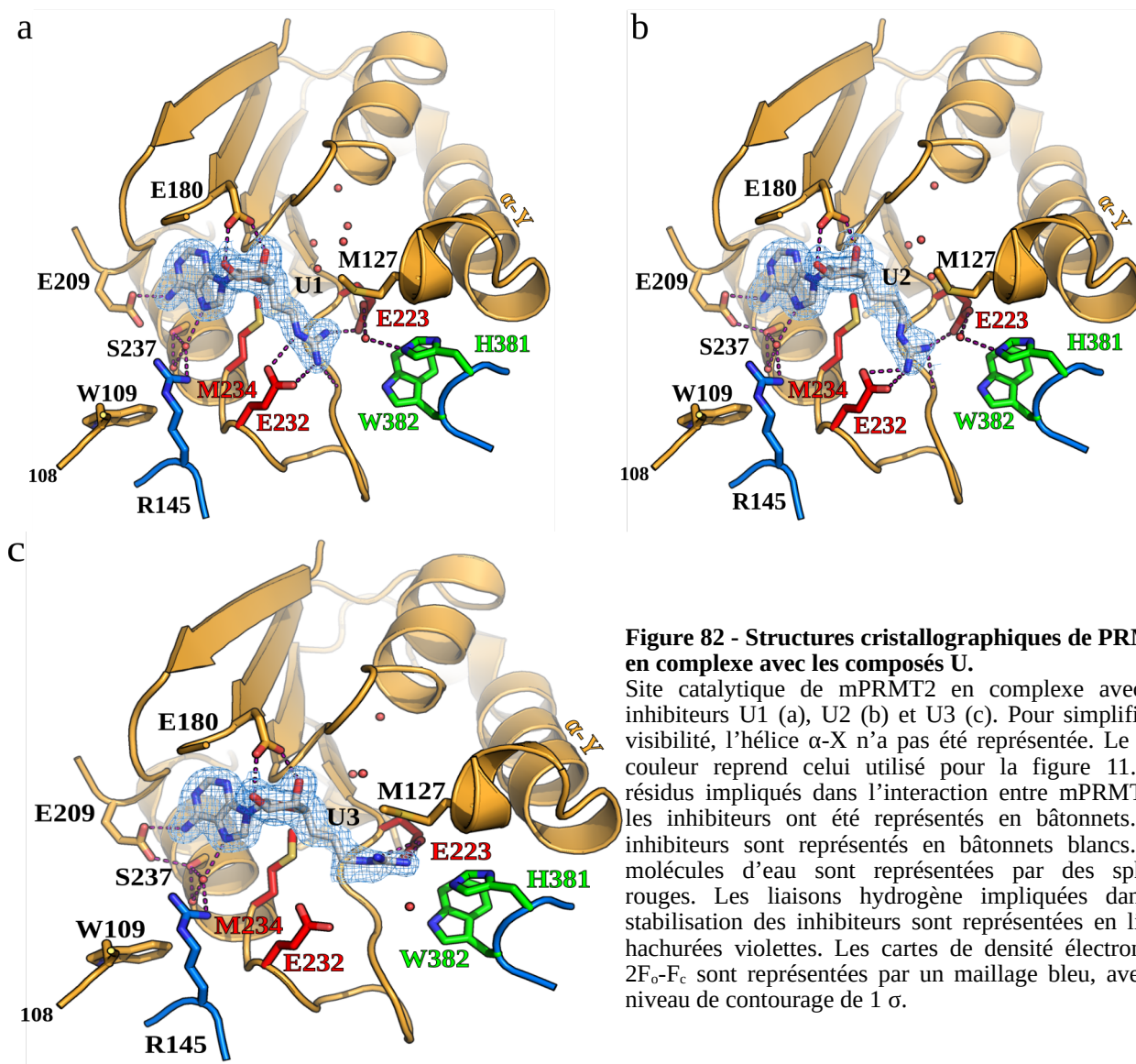


Figure 82 - Structures cristallographiques de PRMT2 en complexe avec les composés U.

Site catalytique de mPRMT2 en complexe avec les inhibiteurs U1 (a), U2 (b) et U3 (c). Pour simplifier la visibilité, l'hélice α -X n'a pas été représentée. Le code couleur reprend celui utilisé pour la figure 11. Les résidus impliqués dans l'interaction entre mPRMT2 et les inhibiteurs ont été représentés en bâtonnets. Les inhibiteurs sont représentés en bâtonnets blancs. Les molécules d'eau sont représentées par des sphères rouges. Les liaisons hydrogène impliquées dans la stabilisation des inhibiteurs sont représentées en lignes hachurées violettes. Les cartes de densité électronique $2F_o - F_c$ sont représentées par un maillage bleu, avec un niveau de contourage de 1σ .

mPRMT2 (fl)-U				
Inhibiteur	Structure	contacts	aromatiques	Liaisons H
U1	5FWA / 1,80 Å	W224 Y118 M234 Y114 G226 + 3 x HOH score : -0,290	∅	E223** E232** M225*
U2	5FWD / 2,00 Å	W224 Y118 M234 M127 G226 + 4 x HOH score : -0,329	∅	E223* E232* M225*
U3	5JMQ / 1,80 Å	Y118 M127 M225 M234 W382 + 1 x HOH score : -0,284	∅	Y118* E223** H381*

Tableau 15 - Structures cristallographiques et interactions de mPRMT2 avec les composés U.

Dans la colonne « contacts », les scores ont été calculés avec la formule du potentiel de Lennard-Jones. Dans la colonne « liaisons H », les astérisques noirs et bleus correspondent à des liaisons hydrogène respectivement faibles et fortes.

U1 semble être le meilleur inhibiteur avec un score de contact de -0,290 et l'établissement de cinq liaisons hydrogène dont trois fortes avec l'ancre (E232), la pompe (E223) et la chaîne principale de M225 (M260 dans mCARM1). Bien que U2 semble établir avec mPRMT2 des interactions interatomiques plus favorables que U1 (score = -0,329), probablement du fait de son bras aliphatique plus long, son groupement guanidinium n'établit que trois liaisons hydrogène, dont une seule forte avec l'ancre, la pompe et la chaîne principale de M225. Par extrapolation avec ce qui a été observé sur mCARM1, cela suggère que U2 est un moins bon inhibiteur de PRMT2 que U1, ce qui est également en corrélation avec le plus faible effet stabilisateur mesuré. L'insaturation sur le bras aliphatique de U3 induit un positionnement orthogonal du noyau guanidinium par rapport à U1 et U2. Cela détériore légèrement les interactions interatomiques (score = -0,284) et diminue l'habilité du composé à établir des liaisons hydrogène, puisque parmi les quatre liaisons formées une seule est considérée comme forte. Ces données structurales suggèrent que U1 est le meilleur inhibiteur de PRMT2, suivi de U2 puis U3, ce qui est en totale corrélation avec les valeurs de stabilisation mesurées en thermofluor.

c - PRMT6

Seuls les composés U1 et U2 ont pu être co-cristallisés avec mPRMT6. La figure 83 présente les structures cristallographiques des complexes de PRMT6 avec U1 et U2 et dévoile les interactions établies entre ces inhibiteurs et le site actif de la PRMT. Le détail de ces interactions et les résolutions des structures obtenues sont résumés dans le tableau 16.

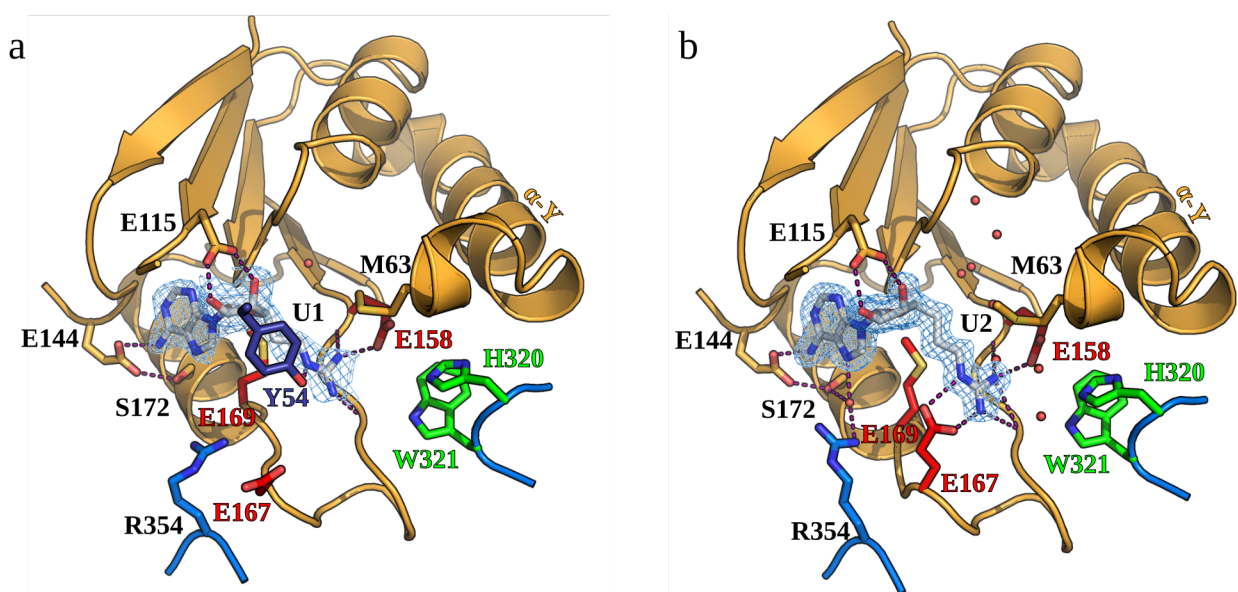


Figure 83 - Structures cristallographiques de mPRMT6 en complexe avec les inhibiteurs U.

Site catalytique de mPRMT6 en complexe avec les inhibiteurs U1 (a) et U2 (b). Pour simplifier la visibilité, l'hélice α -X n'a pas été représentée. Le code couleur reprend celui utilisé pour la figure 11. Les résidus impliqués dans l'interaction entre mPRMT6 et les inhibiteurs ont été représentés en bâtonnets blancs. Les molécules d'eau sont représentées par des sphères rouges. Les liaisons hydrogène impliquées dans la stabilisation des inhibiteurs sont représentées en lignes hachurées violettes. Les cartes de densité électronique $2F_o - F_c$ sont représentées par un maillage bleu, avec un niveau de contourage de 1σ .

mPRMT6 (fl)-U				
Inhibiteur	Structure	contacts	aromatiques	Liaisons H
U1	N.D / 1,70 Å	E158 W159 M160 Y54 M169 G161 score : -0,237	∅	E158** Y54* M160*
U2	N.D / 2,16 Å	Y54 M160 G161 E167 E158 W159 M169 G93 + 4 x HOH score : -0,309	∅	E158** E167** M160**
U3	SAH	∅		

Tableau 16 - Structures cristallographiques et interactions de mPRMT6 avec les composés U.

Dans la colonne « contacts », les scores ont été calculés avec la formule du potentiel de Lennard-Jones. Dans la colonne « liaisons H », les astérisques noirs et bleus correspondent à des liaisons hydrogène respectivement faibles et fortes. N.D indique que les structures n'ont pas été déposées.

À nouveau les structures cristallographiques illustrées sur la figure 83 montrent que la partie adénosine des U occupe le site de fixation du SAM tandis que l'extension aliphatique-guanidinium se fixe dans la poche de fixation de l'arginine. Les interactions établies entre la partie adénosine et le site actif de PRMT6 sont similaires à celles établies avec le SAM et ne seront donc pas détaillées. Les densités électroniques sont visibles pour l'ensemble de la structure des composés U1 et U2, ce qui témoigne d'une fixation efficace de ces inhibiteurs à l'enzyme.

Pour rappel les IC₅₀ de U1, U2 et U3 contre PRMT6 mesurées par le Dr. N. I. Martin et ses collaborateurs sont respectivement de 20,2, 0,7 et 5,2 µM. Dans mPRMT6, U2 forme les mêmes interactions que celles décrites précédemment pour le complexe mPRMT2-U1. Le score de U2 est de -0,309 contre -0,237 pour U1, ce qui traduit des contacts interatomiques entre le composé et la PRMT significativement plus favorables envers U2, probablement dus au plus long bras aliphatique. Le groupement guanidinium de U2 est lui capable d'établir six liaisons hydrogène (un des hydrogènes est partagé entre deux atomes accepteurs) dont quatre fortes avec la « pompe » (E158), l'« ancre » (E167) et la chaîne principale du résidu M160 (M260 dans mCARM1). U1 n'en établit que quatre dont deux fortes avec E158, E167 et Y54. Ces différences expliquent la plus forte inhibition de PRMT6 par U2 comparé à U1.

d - Généralisation

Pour conclure, les structures cristallographiques montrent que les inhibiteurs U se positionnent dans le site actif des PRMT selon le mode de fixation pour lequel ils ont été conçus : La partie adénosine se fixe de manière similaire au cofacteur SAM et la partie guanidinium se localise dans le site de fixation de l'arginine substrat, où elle interagit avec les acides glutamiques catalytiques du motif III. La sélectivité des inhibiteurs U pour PRMT2, CARM1 ou PRMT6 semble provenir de deux facteurs : l'amplitude des contacts interatomiques favorables, qui dépend de la topologie de la poche en relation avec la taille du bras aliphatique, et leur capacité à former des

liaisons hydrogène fortes avec l'ancre, la pompe et plus particulièrement l'oxygène de la chaîne principale d'un résidu méthionine (M260 / M225 / M160). Le meilleur inhibiteur U de chacune des PRMT testées (CARM1 : U1 ; PRMT2 : U1 ; PRMT6 : U2) établit en effet une liaison hydrogène forte avec cette méthionine. Cette interaction empêche pourtant que le guanidinium ne se positionne de manière similaire à l'arginine substrat, en désaxant le groupement guanidinium du plan formé par l'ancre, la pompe et l'histidine du motif IV. Cela suggère que tenter d'imiter exactement la conformation adoptée par l'arginine substrat dans le site actif n'est pas gage d'une inhibition optimale. Dans CARM1, PRMT2 et PRMT6, la liaison hydrogène établie avec la tyrosine du motif I (Y154 / Y118 / Y54) se forme lorsqu'une liaison hydrogène forte avec l'ancre n'est pas possible. Finalement, aucun des inhibiteurs U n'établit d'interactions de type aromatiques avec mCARM1, mPRMT2 et mPRMT6.

3 - [Les composés EML](#)

La troisième famille de composés a été récemment synthétisée par l'équipe du Pr. Gianluca Sbardella. Nommés EML (EML709, EML733, EML734 et EML736), ces composés sont, comme les U, des dérivés du SAM arborant un groupement guanidinium, à l'exception d'EML733 qui ne possède pas de partie adénosine. Ces composés ont été conçus pour se fixer dans les sites de fixation du SAM et de l'arginine, puis explorer l'orifice d'entrée du site actif. Pour cela, ils possèdent en plus une extension aliphatique saturée (nommée espaceur) liée à un groupement urée-aromatique. La différence entre les composés EML709, EML734 et EML736 réside dans la longueur de l'espaceur séparant le groupement guanidinium de la structure urée-aromatique, et constitué respectivement de 4, 3 et 5 atomes de carbone. Les structures chimiques de ces composés sont représentées sur la figure 80.

PRMT	Substrat	Méthode	EML709	EML734	EML736	EML733
PRMT1	H4	AlphaLISA	0,23	0,51	0,32	0,42
PRMT3	H4	Radiométrique	43,4	57,2	21,6	-
CARM1	H3		2,14	13,9	0,42	-
PRMT5	H2A		76,7	52,1	51,4	-
PRMT6	GST-GAR		>100	72,8	>100	-
PRMT7	GST-GAR		0,55	0,31	0,22	-
PRMT8	H4		18,7	8,29	5,87	-

Tableau 17 - Inhibition des PRMT par les composés EML.

Le tableau de données a été produit par Gianluca Sbardella et collaborateurs. Les IC₅₀ indiquées, en µM, ne sont pas encore publiées. L'inhibition de PRMT1 a été mesurée par la méthode AlphaLISA (Bielefeld-Sevigny, 2009) tandis qu'une méthode radiométrique a été appliquée aux autres PRMT.

Le tableau 17 résume les IC₅₀ (en µM) mesurées pour ces composés contre les PRMT par Gianluca Sbardella et ses collaborateurs. Comme le montre les données, les EML sont sélectifs de PRMT1 et PRMT7, et sont les premiers inhibiteurs sélectifs de ces deux enzymes. Les EML sont

également sélectifs de PRMT4/CARM1 mais dans une moindre mesure, en particulier EML709 et EML736. Dans le cadre de notre collaboration avec l'équipe de G. Sbardella, nous avons approfondi la caractérisation de l'interaction entre ces composés et PRMT2, CARM1, PRMT6 et PRMT7 par DSF/thermofluor et par bio-cristallographie. À nouveau, des mesures de stabilisation thermique ont préalablement été réalisées sur PRMT1 de rat, CARM1 de souris et PRMT7 de souris afin de détecter directement l'existence d'interactions. Le tableau 18 résume les variations de T_m des PRMT en présence des inhibiteurs EML, de SAH ou sous forme apo. Comme le montre les valeurs mesurées, un effet stabilisateur des ligands EML par rapport à la forme apo a été mesuré pour l'ensemble des PRMT testées. Tous ces composés semblent donc stabiliser ces trois PRMT dans les conditions expérimentales. L'effet est cependant moins stabilisant qu'avec du SAH pour rPRMT1 et est équivalent au SAH pour mCARM1. Pour mPRMT7, un effet plus stabilisateur que le SAH a été mesuré. Par PRMT, les effets stabilisateurs induits par les trois EML sont similaires.

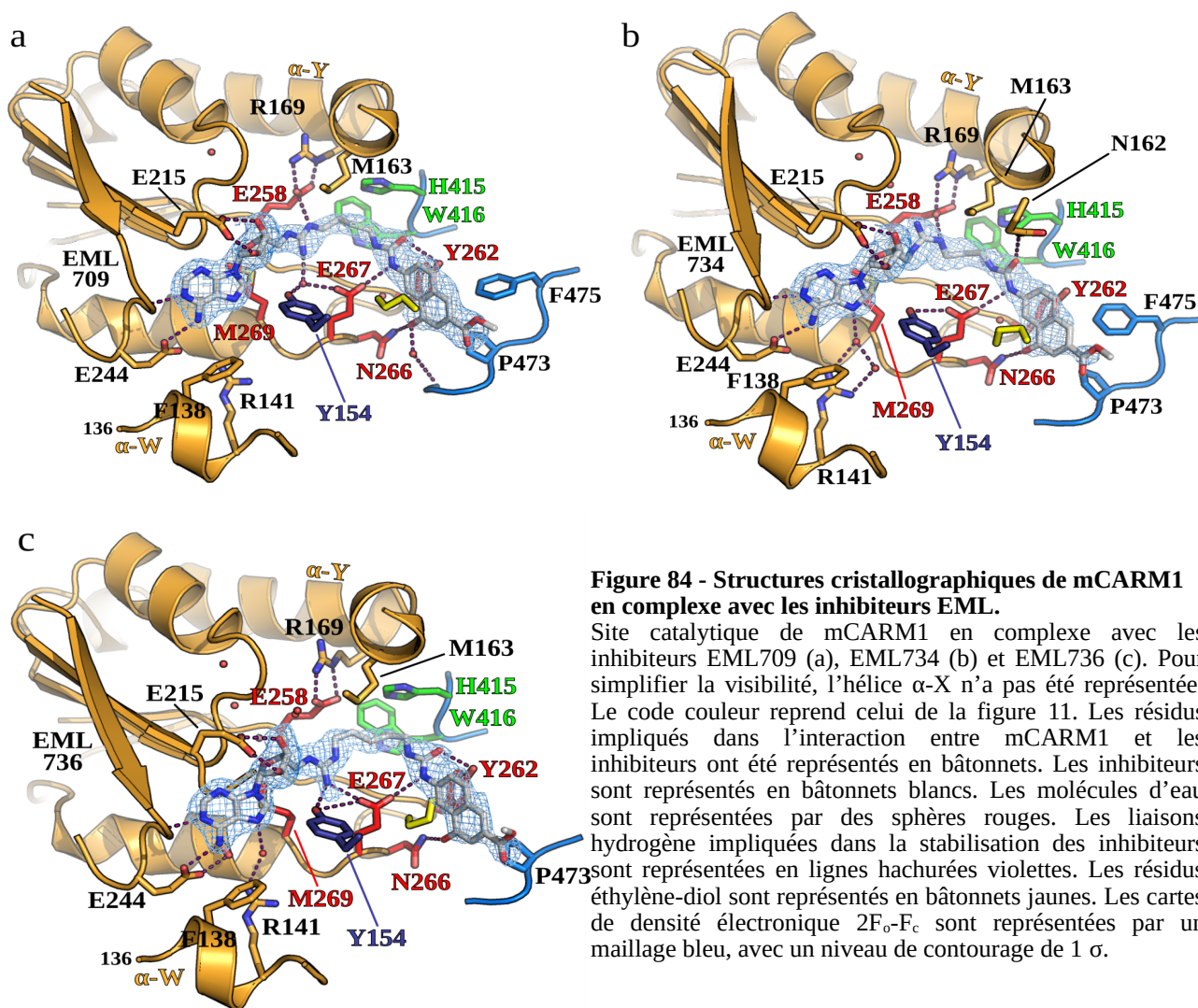
	Méthode	Ligand	ΔT_m (°C) vs apo	ΔT_m (°C) vs SAH
rPRMT1	Thermofluor	EML709	+ 5,5	- 5,0 (SAH)
		EML734	+ 5,0	- 5,5 (SAH)
		EML736	+ 5,5	- 5,0 (SAH)
mCARM1	DSF	EML709	+ 4,2	+ 0,1 (SAH)
		EML734	+ 3,7	- 0,4 (SAH)
		EML736	+ 3,4	- 0,7 (SAH)
mPRMT7	Thermofluor	EML709	+ 4,7	+ 2,2 (SAH)
		EML734	+ 4,0	+ 2,5 (SAH)
		EML736	+4,2	+ 2,7 (SAH)

Tableau 18 - Stabilisation thermique des PRMT par les composés EML.

Afin de déterminer les bases structurales de la sélectivité des composés EML pour PRMT7 et CARM1 comparé à PRMT2 et PRMT6, des essais de co-cristallisation avec mPRMT2 entier (fl), mCARM1 (130-487), mPRMT6 (fl et 34-378) et mPRMT7 (fl) ont été réalisés. Les structures cristallographiques de CARM1 en complexe avec EML709, EML734 et EML736, et de PRMT6 en complexe avec EML734 et EML736 ont été résolues. Comme ce fut le cas avec les composés U, du SAH a été observé dans les cristaux de PRMT7 et ceux-ci ne contenaient pas les inhibiteurs. Les densités électroniques correspondant aux ligands dans PRMT2 n'étaient visible que pour la partie adénosine, ce qui montre que ces composés sont mal fixés dans les structures. Aucune structure obtenue par cristallisation en présence de EML733 ne contenait le composé. Les statistiques de collecte et de traitement des données sont présentées aux annexes 15 et 16. Comme pour les composés U, les interactions formées par la partie adénosine sont similaires à celles établies avec le SAM et ne seront donc pas détaillées.

α - CARM1

La figure 84 présente les structures cristallographiques des complexes de CARM1-EML et dévoile les interactions établies entre ces inhibiteurs et le site actif de la PRMT. Le détail de ces interactions et les résolutions des structures obtenues sont résumés dans le tableau 19.



Les EML occupent tous les modules catalytiques de l'unité asymétrique dans les structures cristallographiques de mCARM1. Dans les structures cristallographiques présentées en figure 84, les densités électroniques sont visibles pour l'ensemble de la structure des composés EML, ce qui témoigne d'une fixation efficace de ces inhibiteurs à l'enzyme. Ceci est en accord avec les IC_{50} mesurées. Comme les U, les EML disposent d'une sous-structure guanidinium, liée à l'adénine de manière similaire aux U mais avec un espaceur carboné de seulement un atome de carbone, ce qui impose au guanidinium de se localiser à l'interface entre les sites de fixation du SAM et de l'arginine. La structure urée-aromatique se positionne quant à elle à la sortie de la poche de fixation de l'arginine.

mCARM1 (130-487) – EML				
Inhibiteur	Structure	contacts	aromatiques	Liaisons H
EML709	N.D / 2,30 Å	F475 Y154 Y150 E258 P473 Y262 M163 E258 M269 F153 H415 M260 W416 score : -0,382	F153 Y262 F475	E267* N266* E258* HOH**
EML733	SAH	Ø		
EML734	N.D / 2,25 Å	F475 Y154 K471 P473 Y262 M163 M269 F153 H415 M260 Y150 score : -0,309	F153 Y262 F475	E267* N266 * E258** Q159*
EML736	N.D / 2,05 Å	F475 K471 P473 Y154 E258 M260 M269 F153 H415 W416 M163 + 2 x HOH score : -0,399	F153 Y262 F475	E267** N266* Y262* Y150*

Tableau 19 - Structures cristallographiques et interactions des complexes mCARM1-EML.

Dans la colonne « contacts », les scores ont été calculés avec la formule du potentiel de Lennard-Jones. Dans la colonne « liaisons H », les astérisques noirs et bleus correspondent à des liaisons hydrogène respectivement faibles et fortes. N.D indique que les structures n'ont pas été déposées.

Les structures cristallographiques montrent que l'espaceur carboné entre les groupements guanidinium et la structure urée-aromatique n'influencent pas le positionnement de la partie urée-aromatique. En effet, les groupements adénosine et urée-aromatique forment deux points d'ancrage dans le site actif de CARM1, et ce sont les conformations du groupement guanidinium, de l'espaceur et des résidus alentours qui sont modifiés. En effet, les espaceurs de EML709 et EML736 (longs respectivement de 4 et 5 carbones) induisent un déplacement de l'histidine du motif IV (H415). Le déplacement de ce résidu a été associé à une forte inhibition lors de l'étude sur les composés LH présentés ci-dessus, et EML709 et EML736 sont justement les composés pour lesquels les plus petites IC₅₀ sur CARM1 ont été mesurées. Pour EML734 (espaceur de 3 carbones), le guanidinium adopte une configuration inversée de 180° par rapport aux EML709 et EML736. Ce positionnement à 180° est moins favorable pour établir des liaisons hydrogène avec les résidus environnants.

Les trois EML établissent chacun cinq liaisons hydrogène par l'intermédiaire des groupements guanidinium, urée et d'une fonction alcool arborée par un des cycles aromatiques. La différence entre ces composés provient des résidus impliqués et la force de ces liaisons. EML736 forme 5 liaisons hydrogène fortes contre 2 pour EML709 et une seule pour EML734. La pompe (E267) est toujours impliquée dans une liaison hydrogène forte avec le guanidinium (EML709 et EML734) ou avec le guanidinium et l'urée (EML736). Le résidu N266 est nécessaire à la stabilisation des groupements aromatiques en établissant une liaison hydrogène de force variable avec la fonction alcool arborée par un des cycles aromatiques.

Les trois composés établissent des interactions aromatiques-aromatiques diagonales similaires entre leurs groupements aromatiques et les résidus F153, Y262 et F475. Les scores pour

EML709 (-0,382), EML734 (-0,309) et EML736 (-0,399) traduisent des interactions de van der Waals plus favorable pour EML736 et EML709 qu'avec EML734.

Comme pour les composés U, on observe pour les composés EML dans mCARM1 une concordance entre le nombre de liaisons hydrogène fortes, les valeurs des scores, et les inhibitions mesurées. En effet, EML736 forme cinq liaisons hydrogène fortes, possède le plus petit score et à la plus faible IC_{50} (0,42 μ M), suivi de EML709 avec deux liaisons hydrogène fortes, un score équivalent à EML736 et une IC_{50} intermédiaire (2,14 μ M) puis EML734 avec une seule liaison hydrogène forte, le plus petit score et la plus grande IC_{50} (13,9 μ M).

Pour les trois inhibiteurs, on observe qu'un fragment du polyéthylène glycol, utilisé comme agent cristallisant dans les conditions de cristallisation, se positionne parallèlement aux groupements aromatiques. Nous pensons que cette molécule d'éthylène glycol se fixe fortuitement à cet emplacement. Il est toutefois possible qu'en absence d'éthylène glycol les contacts entre le groupement aromatique des EML et CARM1 soient différents. Dans ce cas, les conclusions tirées de l'observation des structures avec éthylène glycol sont biaisées car l'agent cristallisant induit artificiellement la stabilisation d'une partie de l'inhibiteur. L'utilisation de sels comme agent cristallisant lors de la cristallisation pourrait permettre d'éviter que les polyéthylènes glycols n'interviennent dans la fixation des ligands ou substrats dans le site catalytique. Les composés EML pourraient être co-cristallisés avec CARM1 et le malonate comme agent cristallisant, puisque nous avons déjà montré dans le chapitre sur les dérivés peptidiques que la cristallisation avec du malonate plutôt que du PEG permettait d'améliorer la fixation des dérivés peptidiques à l'enzyme.

b - PRMT6

La figure 85 présente les structures cristallographiques des complexes de PRMT6-EML et dévoile les interactions établies entre ces inhibiteurs et le site actif de la PRMT. Le détail de ces interactions et les résolutions des structures déterminées sont résumés dans le tableau 20.

La fixation des groupements non-SAM à PRMT6 sont significativement différentes de ce qui a été observé sur CARM1. En effet, les deux structures cristallographiques présentées en figure 85 dévoilent que les composés forment un coude qui positionne les groupements urée-aromatiques à l'horizontal du groupement adénine, au lieu d'occuper le site de fixation de l'arginine et l'entrée du site actif. Cette conformation empêche le repliement de l'hélice α -X autour du site de fixation du cofacteur et de ce fait empêche l'établissement d'un site actif productif. Dans le cas de EML736 dont l' IC_{50} pour PRMT6 est supérieure à 0,1 mM, la densité électronique est peu visible pour tous les atomes au-delà du guanidinium, ce qui suggère que le composé est très flexible. À l'inverse, la densité électronique autour de EML734 recouvre tout le composé ce qui montre qu'il est efficacement fixé.

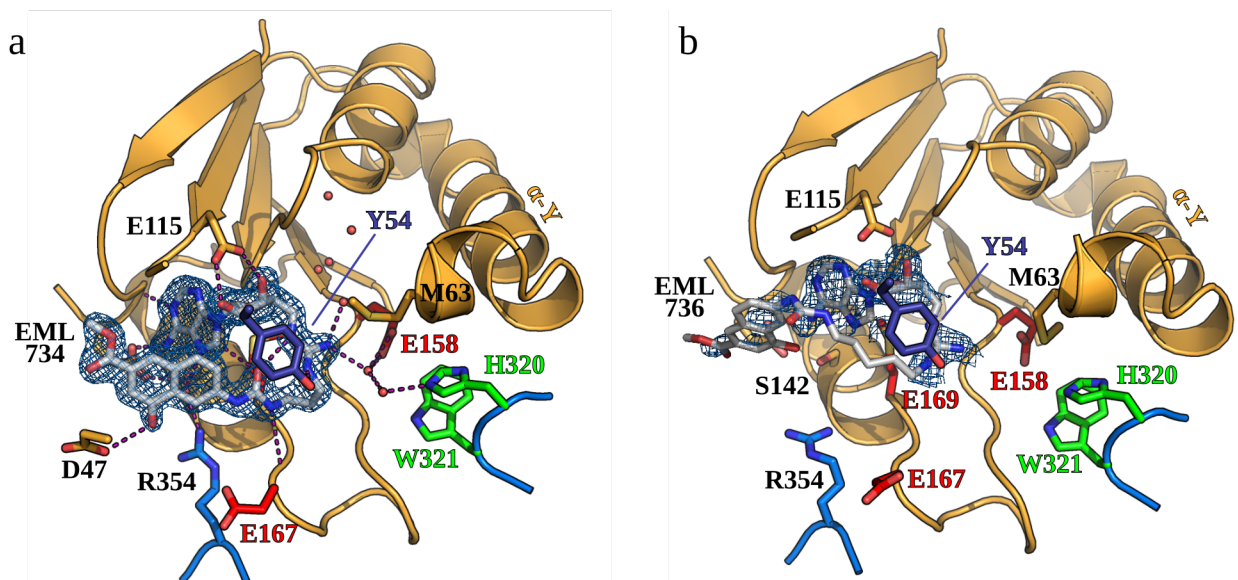


Figure 85 - Structure cristallographique de mPRMT6 en complexe avec les inhibiteurs EML.

Site catalytique de mPRMT6 en complexe avec les inhibiteurs EML734 (a) et EML736 (b). Pour simplifier la visibilité, l'hélice α -X n'a pas été représentée. Le code couleur reprend celui de la figure 11. Les résidus impliqués dans l'interaction entre mPRMT6 et les inhibiteurs ont été représentés en bâtonnets. Les inhibiteurs sont représentés en bâtonnets blancs. Les molécules d'eau sont représentées par des sphères rouges. Les liaisons hydrogène impliquées dans la stabilisation des inhibiteurs sont représentées en lignes hachurées violettes. Les cartes de densité électronique $2F_o - F_c$ sont représentées par un maillage bleu, avec un niveau de contourage de 1σ .

Comme indiqué dans le tableau 20, EML734 établit quatre liaisons hydrogène fortes avec son environnement, mais seulement deux avec des résidus de PRMT6, les deux autres étant établies avec des molécules d'eau. Le score (-0,402) estimé pour ce composé semble du même ordre de grandeur que pour le complexe CARM1-EML736 (-0,399), mais sept molécules d'eau sont prises en compte dans le calcul de ces scores.

mPRMT6 (fl) – EML				
Inhibiteur	Structure	contacts	aromatiques	Liaisons H
EML709	SAH		\emptyset	
EML733	SAH		\emptyset	
EML734	N.D / 1,70 Å	Y50 Y51 Y54 M169 M63 E158 W159 S168 M160 G93 + 7 HOH score : -0,402	Y50 Y54	D47* E167* HOH**
EML736	N.D / 2,00 Å	faible occupation		

Tableau 20 - Structures cristallographiques et interactions des complexes mPRMT6-EML.

Dans la colonne « contacts », les scores ont été calculés avec la formule du potentiel de Lennard-Jones. Dans la colonne « liaisons H », les astérisques noirs et bleus correspondent à des liaisons hydrogène respectivement faibles et fortes. N.D indique que les structures n'ont pas été déposées.

C'est significativement plus que pour tous les autres composés décrits ci-dessus, et notamment que le complexe CARM1-EML736 (deux molécules d'eau). Cela indique que la désolvatation de ce composé est moins efficace lors de sa fixation à PRMT6 qu'à mCARM1. De plus, la conformation coudée de EML736 est stabilisée par deux liaisons hydrogène internes établies par le groupement urée avec l'adénine et le guanidinium. Ces éléments structuraux

pourraient expliquer les mauvaises IC_{50} mesurées pour ces composés sur PRMT6 par G. Sbardella et ses collaborateurs.

c - Généralisation

Les composés EML709, EML734 et EML736 fixent très différemment PRMT2, CARM1 et PRMT6 ce qui est en adéquation avec la sélectivité de ces inhibiteurs pour CARM1. Le composé EML733, dépourvu de la partie adénosine, ne s'est fixé à aucune des PRMT testées.

Dans PRMT2, l'absence de densité électronique au-delà de la partie adénosine montre que ces composés ne sont pas stabilisés par des interactions spécifiques. Dans PRMT6, seuls les composés EML734 et EML736 se sont fixé au site actif, mais selon une conformation coudée impliquant des interactions internes. Les structures non-analogues du SAM de ces composés ne se localisent pas dans le site de fixation de l'arginine substrat de PRMT6 et sont mal désolvatés, ce qui pourrait expliquer les mauvaises IC_{50} mesurées pour ces composés contre PRMT6.

Les trois composés ont été co-cristallisés dans CARM1. Contrairement à ce qui a été observé avec PRMT6, les parties non-adénosine des composés EML occupent le site de fixation de l'arginine et l'orifice de la poche du site actif. Les densités électroniques sont visibles pour EML709, EML734 et EML736 ce qui indique qu'ils sont fortement stabilisés dans le site actif de mCARM1. Les composés occupent tous les modules catalytiques dans l'unité asymétrique des structures cristallographiques ce qui est associé à un fort potentiel inhibiteur. Dans CARM1, les composés EML sont stabilisés par deux points d'ancrage : la partie adénosine qui se lie au site de fixation du cofacteur, et le groupement urée-aromatique qui se fixe à l'entrée du site de fixation de l'arginine et interagit étroitement avec N266, Y153, Y262 et F475. De ce fait l'espaceur, de longueur variable, entre ce groupement et la partie adénosine s'adapte au site de fixation de l'arginine ce qui module la sélectivité des différents EML pour CARM1. La fixation de EML709 et EML736, les deux meilleurs EML pour CARM1, engendre le déplacement de l'histidine du motif IV, ce qui serait associé à un fort potentiel inhibiteur, tel qu'observé dans l'étude sur les composés LH.

4 - Fixation et inhibition des PRMT par les composés SATT

La recherche de « hot spots », c'est-à-dire de sites sur les surfaces protéiques ayant des propriétés de fixation des petites molécules, est une méthode utilisée pour concevoir des modulateurs très affins et sélectifs envers des cibles thérapeutiques. Dans les structures cristallographiques de macromolécules biologiques, les molécules utilisées pour la cristallisation telles que les sels, les tampons, les agents cristallisants et les additifs peuvent être stabilisées en certains points de la structure par des interactions avec les macromolécules cristallisées. Les sites où se localisent ces interactions constituent des « hot spots ».

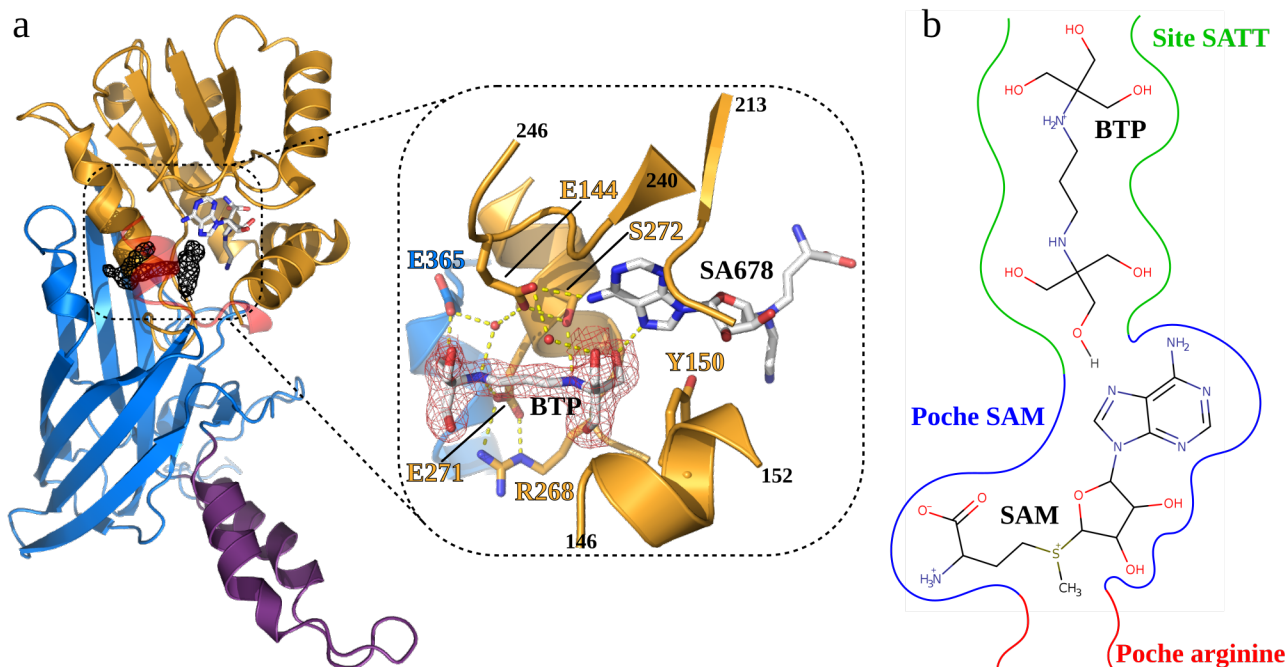


Figure 86 - Identification d'une nouvelle poche à la surface de CARM1.

(a) A gauche : structure cristallographique du module de CARM1 de souris en complexe avec l'inhibiteur SA678. Le code couleur reprend celui de la figure 11. SA678 est représenté en bâtonnets blancs. La densité électronique du BTP est représentée en maillage noir. Le ruban rouge transparent correspond à l'hélice α -W repliée en absence de BTP. Au milieu : interactions formées entre le BTP et le module catalytique de mCARM1. Le BTP est représenté en bâtonnets blancs entouré de sa densité électronique $2F_o - F_c$ en maillage rouge avec un niveau de contourage de 1σ . Les molécules d'eau sont représentées par des sphères rouges. Les liaisons hydrogène sont représentées en lignes hachurées jaunes. (b) Schéma du positionnement de la molécule de BTP et de son interaction avec le SAM.

Lors de l'affinement d'une structure de CARM1 co-cristallisé avec l'inhibiteur dérivé du SAM SA678 (PDB 5NTC), la carte de densité électronique a montré la présence de BTP (Bis-Tris Propane, figure 86) provenant du milieu de cristallisation, fortement stabilisé sur le module catalytique. Comme le montre la figure 86a, la molécule de BTP se localise à l'interface formée entre le domaine de Rossmann et le tonneau β dans la région où se replie en temps normal l'hélice amino-terminale α -W, représentée en ruban rouge transparent sur la figure. En occupant cette position, la molécule de BTP empêche le repliement de α -W et dévoile la présence d'une nouvelle poche en surface du module catalytique. Le BTP y est stabilisé par un réseau de liaisons hydrogène établies avec des résidus du domaine de Rossmann et du tonneau β , ainsi qu'avec des molécules d'eau. De surcroît, le BTP forme une liaison hydrogène avec l'azote 7 du groupement adénine de SA684, tel que schématisé sur la figure 86b. Nous avons ainsi considéré que la zone de fixation du BTP, notée site SATT, pourrait être une région à cibler pour le développement de nouveaux inhibiteurs des PRMT. De plus, le site SATT présente une plus faible conservation de séquence entre les PRMT par rapport aux poches de fixation du SAM et de l'arginine, ce qui pourrait permettre d'obtenir des composés plus sélectifs. La localisation de ce nouveau site par rapport aux poches de fixation de l'arginine substrat et du cofacteur est présenté sur la figure 87.

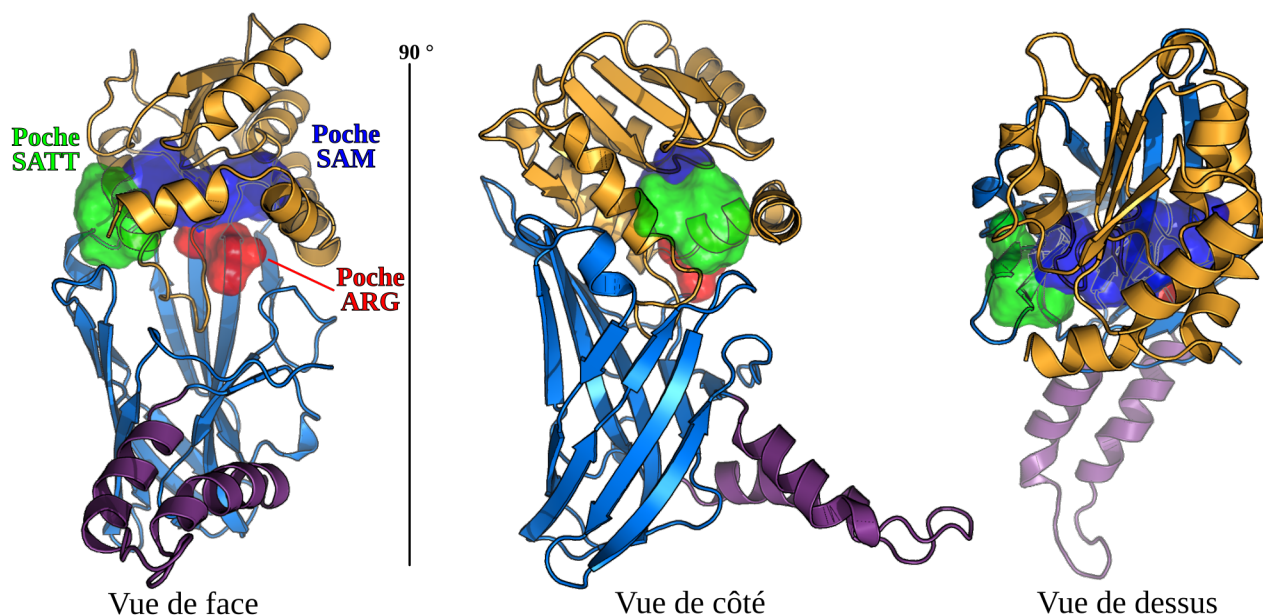


Figure 87 - Poches de fixation de l'arginine substrat, du SAM et du site SATT.

Représentation des sites de fixation de l'arginine et du cofacteur SAM, ainsi que du site SATT sur la structure de mCARM1 (PDB 5NTC). La protéine est représentée en ruban. Le domaine de Rossmann, le tonneau β et le bras de dimérisation sont respectivement colorés en orange, en bleu et en violet. Les poches de fixations de l'arginine, du SAM et la poche SATT sont représentés en volumes transparents et sont respectivement colorées en rouge, bleu et vert.

a - Conception des composés SATT

Dans le but de déterminer les structures chimiques qui seraient les plus adaptées pour faire la liaison entre le site de fixation de l'adénine et le site SATT, nous avons dans un premier temps conçu, en concertation avec les Dr. Alain Wagner et Julie Schmitt, six composés dérivés de l'adénine et du BTP (WS1 à WS6). Leurs structures sont présentées en figure 80. Pour ces composés, la liaison hydrogène entre le BTP et l'adénine a été remplacée par un espaceur carboné. Celui-ci est constitué de deux carbones reliés par une liaison covalente simple, double ou triple. Au préalable de toute synthèse chimique, des simulations de docking dans la structure de mCARM1 ont été réalisées dans AutoDock (Morris et al., 2009) afin de choisir le type d'espaceur qui convient le mieux.

La figure 88 présente les poses optimales pour les deux meilleurs composés, tel que déterminés par docking *in silico*. Les poses de docking les plus similaires aux positionnements adoptés par la partie adénine de SA678 et par le BTP ont été obtenues avec les composés possédant une liaison covalente triple dans l'espaceur, tel qu'illustré figure 88a. Cette liaison permet d'obtenir une jonction linéaire entre la partie adénine et la partie mimant le BTP, qui ne perturberait pas les résidus alentours et permet de positionner le groupement mimant le BTP de manière très similaire à la conformation prise par le BTP dans la structure cristallographique. La liaison covalente double semble interférer à la fois avec le positionnement de la partie adénine et avec la configuration de la partie mimant le BTP, comme le montre la pose présentée sur la figure 88b. Finalement, la liaison

saturée (pose de docking non présentée) entraîne le mauvais positionnement à la fois du groupement mimant l'adénine et de la structure mimant le BTP.

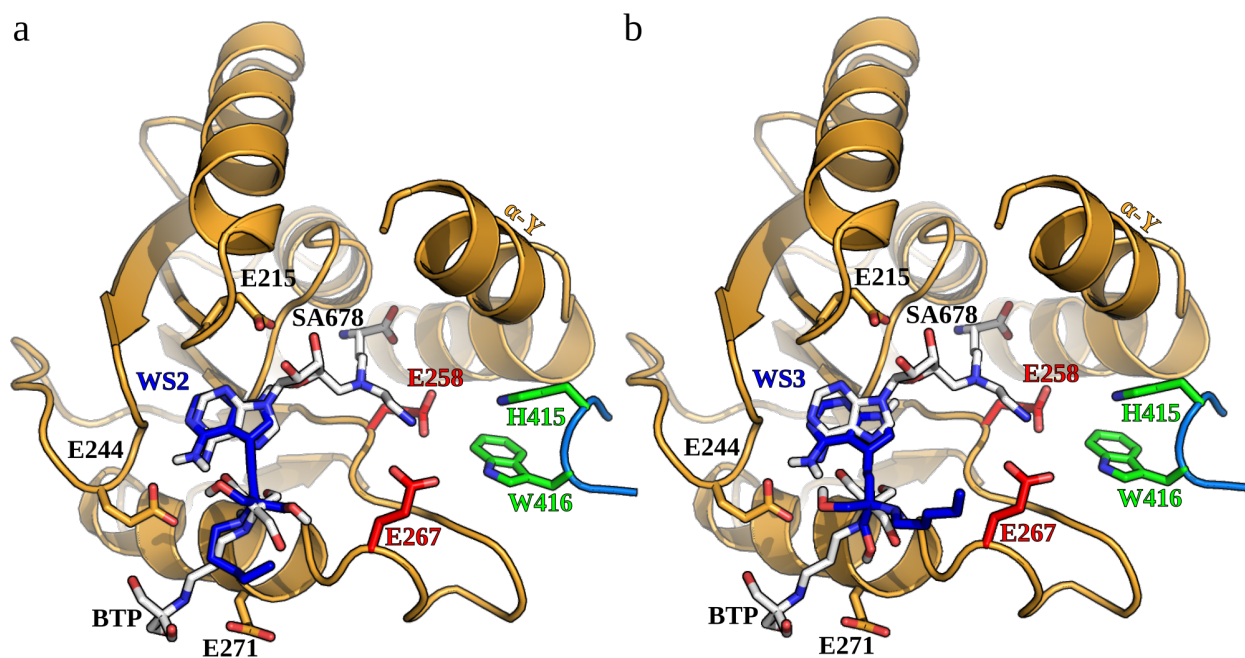


Figure 88 - Résultats de docking des composés WS2 et WS3 dans la structure de mCARM1.

(a) Meilleure pose du docking du composé WS2 dans la structure de mCARM1. (b) Meilleure pose du docking du composé WS3 dans la structure de mCARM1. Le code couleur reprend celui de la figure 11. SA678 et le BTP sont représentés en bâtonnets blancs. Les composés WS2 et WS3 sont représentés en bâtonnets bleus.

En nous basant sur ces résultats *in silico*, nous avons imaginé les composés SATT dont l'architecture est divisée en trois sous-structures : un groupement mimant le SAM constitué d'un dérivé de l'adénosine, un groupement mimant l'arginine similaire à celui de U3, et un groupement occupant le site SATT. Ces composés sont donc conçus pour occuper simultanément le site de fixation du cofacteur, le site de fixation de l'arginine substrat et la poche SATT, représentées respectivement en bleu, rouge et vert sur la figure 87.

Les composés SATT3001 et SATT3002 et leurs intermédiaires SATT1001/SA684, SATT1002, SATT2001 et SATT2002 ont été synthétisés par Julie Schmitt. Leurs structures sont représentées en figure 80 page 175. SATT3001 et SATT3002 consistent en un dérivé du composé U2 dont l'azote 7 de l'adénine est substitué par un carbone, ce qui permet d'y attacher l'espaceur incorporant la liaison triple. L'espaceur est lui-même lié à un groupement aromatique phénylpropanediol, (SATT3001) ou aliphatique boc-sérinol (SATT3002). Ces deux groupements permettront dans un premier temps d'explorer l'espace chimique dans le site SATT.

b - Inhibition des PRMT par les composés SATT

Contrairement aux composés U et EML, l'interaction directe de ces composés avec les PRMT n'a pas été mesurée par stabilisation thermique des PRMT car les quantités d'inhibiteurs disponibles étaient trop faibles, et nous avons choisi de mesurer directement le potentiel inhibiteur de ces composés sur les PRMT. Les inhibitions de PRMT1, PRMT3, CARM1, PRMT5 et PRMT8 par les inhibiteurs SATT3001 et SATT3002 ont été mesurées à l'aide des kits PRMT Chemiluminescent Assay (BPS Bioscience™). Nous avons choisi ces kits plutôt que la méthode radiométrique utilisée généralement au laboratoire afin d'utiliser une méthode standardisée déjà utilisée dans plusieurs études de développement d'inhibiteurs des PRMT (van Haren et al., 2015, 2017 ; Zhang et al., 2015). La méthode radiométrique a tout de même été utilisée pour PRMT2, PRMT6 et PRMT7 car aucun kit n'était disponible pour ces enzymes.

	PRMT1	PRMT2	PRMT3	CARM1	PRMT6	PRMT8	PRMT5	PRMT7
Méthode	kit	radiométrie	kit	kit	radiométrie	kit	kit	radiométrie
SAH / SFG	-	18,3 ± 2,0	-	-	1,90 ± 0,26	-	-	27,7 ± 8,5
SATT3001	0,74 ± 0,60	8,0 ± 0,9	3,1 ± 0,6	0,13 ± 0,01	20,9 ± 2,8	11,4 ± 1,9	14,8 ± 2,2	41,5 ± 14,5
SATT3002	0,95 ± 0,14	130 ± 56	12,2 ± 1,5	12,3 ± 1,8	37,5 ± 4,0	17,6 ± 4,8	43,4 ± 5,9	> 100

Tableau 21 - Inhibition des PRMT par les composés SATT3001 et SATT3002.

Ce tableau résume les IC₅₀ en µM pour les composés SATT3001 et SATT3002 contre les PRMT1 à 8. Les PRMT de type I, II et III sont représentées respectivement en bleu, rouge et vert. Les composés sélectifs d'une PRMT sont indiqués en gras.

Le tableau 21 résume les IC₅₀ (en µM) mesurées pour les inhibitions des PRMT1 à PRMT7 par les inhibiteurs SATT3001 et SATT3002. Les données d'inhibition indiquent des premiers résultats encourageants. Tout d'abord, le composé SATT3001 inhibe PRMT1, 2, 3 et 4 à des concentrations inférieures à 10 µM, particulièrement PRMT4/CARM1 avec une IC₅₀ de l'ordre de 130 nM. C'est également le cas pour SATT3002 sur PRMT1 (IC₅₀ de l'ordre de 1 µM). Les inhibitions contre les PRMT de type II et III sont significativement moins efficaces, ce qui suggère que cibler le site SATT induit une sélectivité pour les PRMT de type I. SATT3001 est plus sélectif de CARM1 d'environ un facteur 6, 24, 62 et 88 comparé à PRMT1, PRMT3, PRMT2 et PRMT8, et plus de 100 fois plus sélectif comparé aux autres PRMT. SATT3002 est lui plus sélectif de PRMT1 d'environ un facteur 13 comparé à PRMT3 et CARM1, d'un facteur 19, 39 et 46 comparé à PRMT8, PRMT6 et PRMT5 et supérieure à 100 fois plus sélectif comparé aux autres PRMT. Les IC₅₀ de SATT3001 et SATT3002 sont toutefois encore élevées si on les compare à certains inhibiteurs de CARM1 affichant des IC₅₀ de l'ordre de la dizaine de nanomolaire (Drew et al., 2017). Il s'agit cependant d'un début encourageant pour une première phase et ces composés devraient pouvoir être optimisés.

c - Analyse structurale de la fixation des composés SATT à CARM1

Des essais en cristallisation de mCARM1 (130-487) et mPRMT2 (fl) en présence de SATT3001 et SATT3002 ont été réalisés. Des cristaux ont été obtenus pour les deux protéines en présence des deux inhibiteurs, et les données ont été enregistrées en utilisant la source de rayonnement synchrotron SOLEIL. Les statistiques de collecte et de traitement des données pour les complexes CARM1-SATT sont disponibles en annexe 17. Les cristaux de mCARM1 et mPRMT2 appartiennent aux groupes d'espace $P2_12_12$ et $C222_1$ habituels. Les cartes de densités électroniques dévoilent que les inhibiteurs SATT3001 et SATT3002 ne sont fixés que sur CARM1.

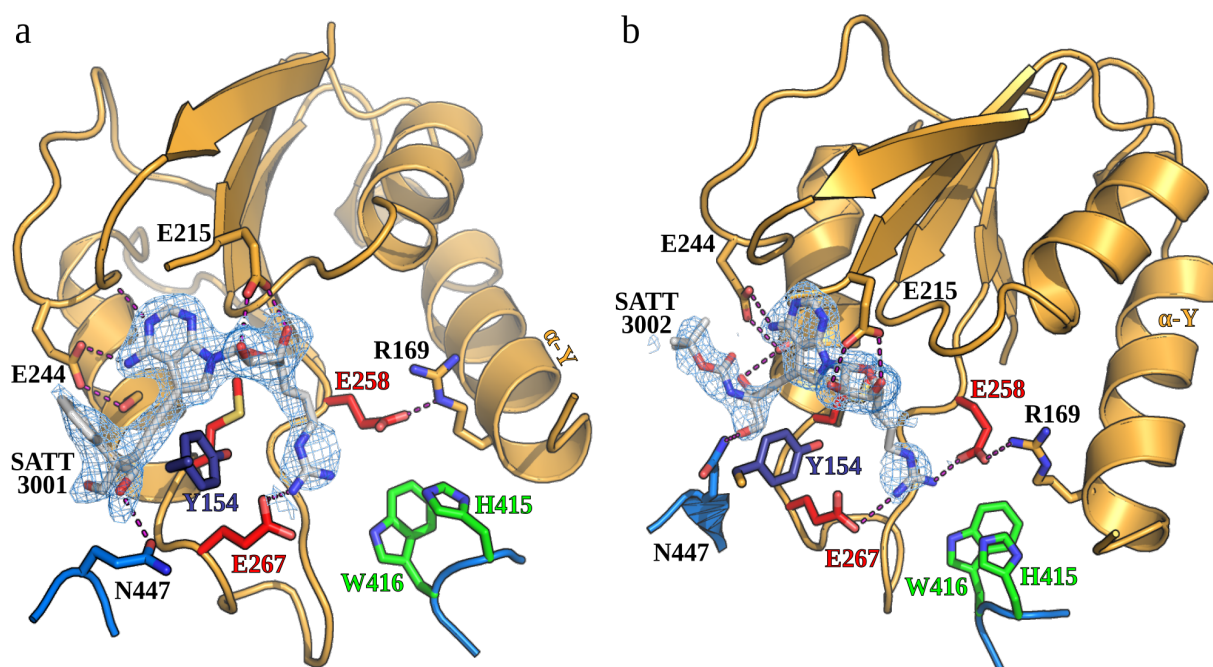


Figure 89 - Structure de mCARM1 en complexe avec les inhibiteurs SATT300X.

Domaine de Rossmann de mCARM1 en complexe avec les inhibiteurs SATT3001 (a) et SATT3002 (b). Les résidus interagissant avec les inhibiteurs sont représentés en bâtonnets. Les inhibiteurs sont représentés en bâtonnets blancs. Le code couleur reprend celui utilisé en figure 11. Les cartes de densité électronique $2F_o - F_c$ sont représentées par un maillage bleu, avec un niveau de contourage de 1σ . Pour des questions de visibilité, l'hélice α -X n'a pas été représentée.

SATT3001 occupe les quatre modules catalytiques dans l'unité asymétrique des cristaux de CARM1, tandis que SATT3002 n'en occupe que la moitié, ce qui est cohérent avec leurs IC_{50} respectives de 0,13 et 12,3 μ M. La figure 89 présente le mode de fixation de SATT3001 (a) et SATT3002 (b) au module catalytique de CARM1 et confirme que les deux composés se fixent selon le mode de fixation attendu : la partie adénosine interagit avec le site de fixation de manière similaire au cofacteur, l'extension analogue au composé U3 se localise dans la poche de fixation de l'arginine où le guanidinium interagit avec les résidus catalytiques, et pour finir les groupements phénylpropanediol (SATT3001) et boc-sérinol (SATT3002) occupent le site SATT. Ce mode de fixation induit le déplacement de l'hélice amino-terminale α -W, tel qu'observé lors de la fixation de la molécule de BTP.

Dans le cas du SATT3002, dont l' IC_{50} pour CARM1 est plus élevée, les densités électroniques ne sont pas de bonne qualité pour l'extension guanidinium et pour l'espaceur linéaire faisant le lien entre l'adénine et le groupement boc-sérinol. Cela suggère que ces parties sont flexibles ou qu'il y a une occupation partielle du composé, ce qui est vraisemblablement le cas puisque SATT3002 n'occupe qu'un site sur deux dans la structure cristallographique. À l'inverse, la densité électronique autour du groupement boc-sérinol est bien définie, ce qui indique que ce groupement est fixé efficacement dans le site SATT.

Pour SATT3001, la densité électronique est de bonne qualité sur l'ensemble du composé. Cependant, les groupements phénylpropanediol du composé semblent adopter des configurations différentes dans les quatre sous unités de la structure cristallographique de mCARM1.

Un des points soulevés par les structures cristallographiques des complexes mCARM1-SATT3001/2 est que les groupements occupant le site SATT forment relativement peu d'interactions avec la protéine comparés aux groupements adénosine et guanidinium. C'est particulièrement le cas pour le SATT3001 dont le groupement phényle ne semble former aucune liaison avec CARM1, ce qui pourrait expliquer la mobilité apparente de ces régions des inhibiteurs. Ces composés modifient cependant vraisemblablement le schéma d'hydratation local du module catalytique en déplaçant l'hélice α -W et en introduisant les groupements boc-sérinol et phénylpropanediol à fort encombrement stérique. Malgré cela, nous avons pu mesurer une IC_{50} sub-millimolaire pour SATT3001 sur CARM1 et proche 10 μ M dans le cas de SATT3002, ce qui signifie qu'en optimisant les groupements chimiques occupant le site SATT pour qu'ils établissent plus d'interactions, nous devrions pouvoir atteindre une plus forte inhibition et induire davantage de sélectivité.

5 - [Perspectives](#)

a - Inhibiteurs de PRMT1 et PRMT7

PRMT1 est la PRMT prédominante et sa dérégulation est impliquée dans le développement de nombreux cancers. De nombreuses études ont mené au développement de composés inhibant cette enzyme mais la plupart sont peu sélectifs et un seul d'entre eux a permis d'atteindre une IC_{50} sub-micromolaire (Mitchell et al., 2015). Ce dernier est toutefois plus sélectif pour PRMT6. Le développement d'inhibiteurs spécifiques et sélectifs de PRMT1 souffre de l'impossibilité actuelle de co-cristalliser PRMT1 avec des ligands ciblant son site actif. En effet, bien que quatre structures cristallographiques de PRMT1 de rat aient été publiées et déposées à la PDB par deux équipes (Gui et al., 2011 ; Zhang et al., 2003), la forme cristallographique obtenue est incompatible avec l'étude de complexes PRMT-inhibiteur puisque l'hélice α -X n'y est pas visible, donc le site catalytique

productif n'est pas formé. Cela pourrait être le résultat des conditions dans laquelle rPRMT1 a cristallisé, à un pH non physiologique de 4,7.

Le constat est le même pour PRMT7, une autre cible émergente pour le développement d'inhibiteurs. Dans les structures cristallographiques de PRMT7 de souris (Cura et al., 2014), on n'observe pas non plus le repliement du segment correspondant à l'hélice α -X sur le site catalytique. De plus, les structures de PRMT7 de *T. brucei* (Debler et al., 2016) et *C. elegans* (Hasegawa et al., 2014b) présentent cette même caractéristique, bien que les conditions de cristallisation soient différentes. Dans ces dernières structures, on observe que le segment équivalent à α -X forme un brin β en interaction avec le tonneau β d'une molécule voisine, ce qui suggère que l'empilement cristallin empêche le repliement de ce segment permettant de former un site catalytique productif. Ainsi à l'heure actuelle, nous ne sommes parvenus à cristalliser aucun complexe PRMT7-inhibiteur et cela pourrait être dû à la forme cristallographique obtenue. Aujourd'hui, nous travaillons à déterminer des constructions de ces protéines, et des conditions de cristallisation qui nous permettraient de co-cristalliser des inhibiteurs avec PRMT1 et PRMT7.

b - Composés EML

Les composés EML présentés dans ce chapitre sont les premiers composés inhibants sélectivement PRMT1 et PRMT7, et dans une moindre mesure CARM1. Les essais de co-cristallisation des EML avec PRMT7 n'ont pour l'heure pas permis d'obtenir de données structurales, pour les raisons évoquées au paragraphe précédent, et aucun essai sur PRMT1 n'a été réalisé car nous ne maîtrisons pas encore sa cristallisation. Des structures cristallographiques des EML fixés à CARM1 et PRMT6 ont toutefois été résolues. Le mode de fixation dans PRMT6 explique les mauvaises inhibitions mesurées sur cette enzyme, et à l'inverse les structures cristallographiques des complexes CARM1-EML dévoilent les bases structurales de l'inhibition de CARM1 par ces composés : les groupements adénosine et urée-aromatique établissent deux points d'ancrage dans la poche de fixation du SAM et à l'entrée du site catalytique, et la longueur de l'espaceur qui fait le lien entre ces deux groupements module l'affinité de l'inhibiteur pour CARM1. Pour rappel, l'espaceur arbore un groupement guanidinium qui se localise à l'interface entre la poche de fixation du SAM et le site de fixation de l'arginine. À partir de ces données, la spécificité et la sélectivité des EML pour CARM1 pourraient être améliorées en agissant sur l'espaceur et sur la partie aromatique des composés. Il pourrait être envisagé de rallonger le bras aliphatique reliant la partie adénosine au guanidinium, afin que ce groupement se positionne dans le site de fixation de l'arginine d'une manière analogue à ce qui est observé pour les composés U. Il faudra en même temps raccourcir l'espaceur entre le guanidinium et le groupement urée-aromatique afin de ne pas perturber le positionnement de ce dernier. Le groupement urée aromatique pourra également subir des modifications afin d'optimiser les interactions établies avec CARM1. Des

simulations de docking pourront être menées pour guider ces modifications chimiques de la structure des composés EML.

c - Composés SATT

Comme les composés LH, U et EML, Les inhibiteurs SATT sont des dérivés du SAM. Leur structure en partie semblable à une molécule d'origine biologique pourrait les exposer à la métabolisation par l'organisme. De plus, ces composés sont de grande taille et sont assez polaires, ce qui pourrait limiter leur pénétration dans la cellule par diffusion à travers les membranes, comme c'est le cas pour les composés LH (Halby et al., 2018 ; Yang et al., 2015a). Pour surmonter ces obstacles, nous tentons actuellement de remplacer les groupements adénosine et guanidinium par des structures moins chargées. Des composés, notés SATT400X et non présentés dans ce manuscrit, ont été conçus et testés *in silico* par docking et dynamique moléculaire, et sont actuellement en cours de synthèse. Les groupements occupant le site SATT devront également être optimisés afin d'augmenter l'affinité de ces composés pour les PRMT, et afin de moduler leur sélectivité pour les différents membres de cette famille d'enzymes. Le remplacement des groupements adénosines et la conception de dérivés moins polaires devra également être appliquée aux inhibiteurs LH, U et EML.

Les composés EML et SATT restent à être testés sur des lignées cellulaires saines et cancéreuses pour déterminer leur activité *in vivo* et contrôler leur innocuité envers les cellules saines. Les composés SATT en particulier, devraient être testés *in vitro* sur d'autres méthyltransférases afin de confirmer l'absence d'effets « off-target ».

C - Matériel et méthodes

1 - Expression et purification des PRMT

a - PRMT1

Les clones de PRMT1 (15-353) de rat ont été donnés par le Dr. J.M Hevel de l'université de l'Utah. La protéine a été exprimée dans la souche *E. coli* BL21(DE3) transformée avec le plasmide pET28_rPRMT1-S14 permettant l'expression inductible de la PRMT1 de rat tronquée au résidu 13 et arborant une étiquette hexahistidine non clivable à l'extrémité amino-terminale. Les cellules ont été cultivées dans du milieu LB à 37 °C jusqu'à une DO600 d'environ 0,4, puis la température a été diminuée à 22 °C. L'expression de PRMT1 a été induite par 0,5 mM d'IPTG à une DO600 d'environ 0,7 puis l'expression s'est poursuivie durant 12 heures. Les cellules ont été récoltées par centrifugation puis lysées par sonication dans 250 mM NaCl, 50 mM Tris-HCl pH 8, 1 mM TCEP, 0,01 % NP40, 1X Complete® EDTA-free protease inhibitor cocktail, à 4 °C. Le lysat a été clarifié puis 6his-rPRMT1-S14 a été incubé avec de la résine Ni-Sepharose™ 6 Fast Flow (GE Healthcare) dans le milieu de lyse supplémenté de 20 mM d'imidazole. La résine a été lavée avec 250 mM NaCl, 50 mM Tris-HCl pH 8, 20 mM imidazole, 1 mM TCEP puis les protéines ont été éluées avec 250 mM imidazole. La concentration en protéine dans les fractions a été estimée par mesure de l'absorbance à 280 nm. Les éluas contenant la protéine ont été dilués dix fois avec 25 mM NaCl, 20 mM Tris-HCl pH 8, 0,5 mM DTT, 0,5 mM EDTA puis chargés sur résine échangeuse d'anions HiTrap™ ANX Fast Flow (GE Healthcare). Les protéines ont été éluées avec 500 mM NaCl. Les fractions contenant la protéine ont été rassemblées puis précipitées par 2 M de sulfate d'ammonium. Les protéines ont été récoltées par centrifugation à 15000 g, 15 minutes à 4 °C puis resuspendues dans 1 mL de tampon de filtration sur gel (100 mM NaCl, 20 mM Tris-HCl pH 8, 0,5 mM DTT, 0,5 mM EDTA) supplémenté de 1 mM SAH. Une dernière étape de purification a consisté en une filtration sur gel sur résine Superdex™ S200 à 1 mL.min⁻¹. Les fractions contenant la protéine ont été rassemblées puis concentrées sur Amicon 30 kDa (Merk Millipore), à 525 g à 4 °C. D'éventuels agrégats ont finalement été éliminés par centrifugation 20 minutes à 20000 g à 4 °C.

b - PRMT6

La purification de PRMT6 de souris (construction 1-378 ou 34-378) a été mise au point au sein de l'équipe par le Dr. Luc Bonnefond (Bonnefond et al., 2015). La protéine a été exprimée dans la souche *E. coli* BL21(DE3) transformée avec le plasmide pnEA/vH-mPRMT6_34-378. Les cellules ont été cultivées dans du milieu LB à 37 °C jusqu'à une DO600 d'environ 0,4, puis la température a été baissée à 30 °C. L'expression de PRMT6 a été induite par 0,1 mM d'IPTG à une DO600 d'environ 0,5 puis l'expression s'est poursuivie durant 4 heures. Les cellules ont été récoltées par centrifugation puis lysées à froid par sonication dans 300 mM NaCl, 50 mM Tris-HCl

pH 8, 7 mM β ME, 5 % glycérol, 20 mM imidazole, 1 mM PMSF. Le lysat a été clarifié par centrifugation 1 heure à 40000 g puis l'échantillon a été incubé avec de la résine NiNTA[®] (Quiagen) pendant 1 heure à 4 °C. La résine a été lavée avec le tampon de lyse puis les protéines ont été éluées par 500 mM imidazole. Les éluas contenant la protéine (mesurés par Bradford) ont été dialysés contre 150 mM NaCl, 50 mM Tris HCl pH 8, 7 mM β ME, 1 mM EDTA pendant la nuit en présence de protéase TEV. L'échantillon a finalement été purifié par filtration sur gel sur résine Superdex[™] S200 16/60 (GE Healthcare) à 1 mL.min⁻¹ dans 20 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM DTT. Les fractions contenant la protéine ont été rassemblées puis concentrées sur Amicon 10 kDa (Merk Millipore).

Les protéines mPRMT2 et mCARM1 (130-487) ont été purifiées tel que présenté dans les chapitres III et IV.

2 - Synthèse des composés SATT

Les composés SATT ont été synthétisés par Julie Schmitt sous la direction d'Alain Wagner.

3 - Préparation des composés

Tous les composés ont été initialement fournis sous forme de poudre. La sinefungine (Sigma[™]) et la S-adenosylhomocystéine (Sigma[™]) ont été dissous dans l'eau ultra-pure à une concentration d'environ 10 mM. Les composés EML709, EML734 et EML736 ont été dissous dans l'eau ultra-pure à une concentration d'environ 20 mM. EML733 a été dissout dans une solution aqueuse de 20 % DMSO à une concentration d'environ 16 mM. Les composés SATT1002, SATT3001 et SATT3002 ont été dissous dans l'eau ultra-pure à des concentrations respectives de 20, 22 et 50 mM. Les composés SATT2001 et SATT2002 ont été dissous dans une solution aqueuse de 25 % DMSO à une concentration d'environ 50 mM. Les concentrations exactes en SAH, SFG, EML709, EML734, EML736 et SATT1002 ont été mesurées par absorbance à 260 nm avec le coefficient d'absorbance molaire de l'adénine de 15400 M⁻¹.cm⁻¹. Les concentrations en EML733, SATT2001, SATT2002, SATT3001 et SATT3002 ont été évaluées par mesure du volume final en fonction de la masse de composé dissout.

4 - Co-cristallisation des complexes PRMT-inhibiteurs

a - CARM1

mCARM1 (130-487) a été cristallisé en complexe avec des inhibiteurs selon la méthode de diffusion de vapeur en goutte suspendue, dans une plaque de cristallisation VDX 24 puits (Hampton[™] research). Les gouttes de cristallisation ont été constituées sur des lamelles de verre 22 mm traitées à la silicone (HR3-231, Hampton[®] research) par le mélange de 2 μ L de solution de

protéine à 2 mg.mL^{-1} , une concentration en inhibiteurs de $0,5 \text{ mM}$ à 2 mM (selon solubilité) et $1 \mu\text{L}$ de précipitant, constituée d'un gradient de 12% à 22% en PEG3350, 100 mM Tris-HCl pH 8, 100 mM NaCl. Les réservoirs de la plaque de cristallisation ont été remplis avec $0,5 \text{ mL}$ de précipitant puis les lamelles ont été scellées sur la plaque à l'aide d'huile de paraffine (Merk™). Les cristaux ont été obtenus après une semaine à l'obscurité, à $22 \text{ }^\circ\text{C}$.

b - PRMT2

mPRMT2 (entière) a été cristallisé en complexe avec des inhibiteurs selon la méthode de diffusion de vapeur en goutte suspendue, dans une plaque de cristallisation VDX 24 puits (Hampton™ research). Les gouttes de cristallisation ont été constituées sur des lamelles de verre 22 mm traitées à la silicone (HR3-231, Hampton® research) par le mélange de $2 \mu\text{L}$ de solution de protéine à $3,5 \text{ mg.mL}^{-1}$, une concentration en inhibiteurs de $0,5 \text{ mM}$ à 1 mM (selon solubilité) et $1 \mu\text{L}$ de précipitant, constituée de 16% de PEG 6000, 100 mM HEPES-NaOH pH 7, 100 mM CaCl_2 . Les réservoirs de la plaque de cristallisation ont été remplis avec $0,5 \text{ mL}$ de précipitant puis les lamelles ont été scellées sur la plaque à l'aide d'huile de paraffine (Merk™). Les cristaux ont été obtenus après une semaine à l'obscurité, à $22 \text{ }^\circ\text{C}$.

c - PRMT6

MPRMT6 (1-378 ou 34-378) a été cristallisé en complexe avec des inhibiteurs selon la méthode de diffusion de vapeur en goutte suspendue, dans une plaque de cristallisation VDX 24 puits (Hampton™ research). Les gouttes de cristallisation ont été constituées sur des lamelles de verre 22 mm traitées à la silicone (HR3-231, Hampton® research) par le mélange de $1 \mu\text{L}$ de solution de protéine à 6 mg.mL^{-1} , une concentration en inhibiteurs de 1 mM et $1 \mu\text{L}$ de précipitant, constituée de 10% de PEG 6000, 100 mM Tris-HCl ou HEPES-NaOH pH 8, 200 mM NaCl. Les réservoirs de la plaque de cristallisation ont été remplis avec 1 mL de précipitant puis les lamelles ont été scellées sur la plaque à l'aide d'huile de paraffine (Merk™). Les cristaux ont été obtenus après une semaine à l'obscurité, à $22 \text{ }^\circ\text{C}$.

d - Collecte et traitement des données

Les cristaux ont été cryo-protégé par immersion dans une solution de réservoir supplémentée par 15% de PEG400, puis les cristaux ont été congelés par immersion dans l'azote liquide. Les données ont été collectées aux lignes Proxima 1 du synchrotron SOLEIL et ID23-2 du synchrotron ESRF, équipées de détecteurs Pilatus 6M, Pilatus 2MF ou Q315, ou sur le diffractomètre de la plateforme de l'IGBMC, équipé d'une source Rigaku FR-X à anode tournante et d'un détecteur Eiger 4M (Dectris). Les données ont été traitées avec le programme XDS et la suite de programmes CCP4, puis les structures ont été affinées avec le programme Phenix. Les constructions manuelles ont été réalisées avec le programme coot. Les cristaux de CARM1

appartiennent aux groupes d'espace $P_{2_12_12}$ ou P_{2_1} avec respectivement quatre ou huit monomères par unité asymétrique. Les cristaux de PRMT2 et PRMT6 appartiennent respectivement aux groupes d'espaces C_{222_1} et $I4_2$, avec un et deux monomères par unité asymétrique.

5 - [Mesure de la stabilité thermique](#)

a - Thermofluor (PRMT1, PRMT2, CARM1 et PRMT7)

Les échantillons sont chauffés en présence d'un composé qui devient fluorescent au contact de zones hydrophobes. On peut de cette manière suivre la dénaturation thermique des protéines par mesure de la fluorescence en fonction de la température. La PRMT a été diluée à $0,25 \text{ mg.mL}^{-1}$ dans une solution contenant 0,2 à 5 mM de ligand (selon le composé) et 6X de Sypro Orange 25 X (Thermo Fisher Scientific), 20 mM HEPES pH 7,5, 50 mM NaCl (PRMT1) ou 50 mM Tris-HCl pH 8,0, 200 mM NaCl (CARM1 et PRMT7). Lorsque le composé a été dilué dans une solution aqueuse de DMSO, un contrôle DMSO a été effectué. Les échantillons sont chauffés progressivement de $20 \text{ }^\circ\text{C}$ à $95 \text{ }^\circ\text{C}$ dans un cycleur thermique pour PCR quantitative en temps réel (MiniOpticon, BioRad™). La fluorescence du Sypro a été mesurée à 570 nm après excitation à 470 nm. Les T_m sont mesurés au niveau des points d'inflexion des courbes de dénaturation.

b - Fluorimétrie par balayage différentiel (PRMT3, CARM1 et PRMT6)

La mesure du T_m par DSF a été effectuée sur un Prometheus NT.48 tel qu'indiqué dans les chapitres II et V. Pour ces mesures, PRMT6 a été diluée à $0,5 \text{ mg.mL}^{-1}$ dans 20 mM Tris-HCl pH7,5, 150 mM NaCl, en présence du ligand à 1 mM et avec ou sans DTT à 48 mM.

6 - [Mesure des IC₅₀](#)

a - Mesure radiométrique directe (PRMT2, PRMT6 et PRMT7)

Cette expérience permet de quantifier par SDS-PAGE l'incorporation d'un méthyl radioactif issu du SAM par le substrat. Il s'agit donc d'une observation directe de la formation de produit.

Pour PRMT2, la réaction de méthylation en présence d'inhibiteurs a été mise en œuvre selon le protocole décrit dans l'article « Structural studies of protein arginine methyltransferase 2 reveal its interactions with potential substrates and inhibitors » joint en annexe, avec rsf1 entier utilisé comme substrat. Pour PRMT7, le mélange contenant $1,9 \text{ } \mu\text{M}$ de mPRMT7, $2,2 \text{ } \mu\text{M}$ du substrat entier rsf1 dilués dans 50 mM Tris-HCl pH 8,0, 100 mM NaCl, 0,5 mM DTT, $0,02 \text{ mg.mL}^{-1}$ BSA et $100 \text{ } \mu\text{M}$ PMSF a été incubé à température ambiante en présence de concentrations croissantes en inhibiteurs. La réaction a été initiée par ajout de $20 \text{ } \mu\text{M}$ de cofacteur SAM ^{14}C puis arrêtée après 3 heures à température ambiante par ajout à 1X finale de tampon de charge SDS-PAGE 4X. Pour PRMT6, la réaction a été initiée par l'ajout du mélange SAM $^{14}\text{C}/(\text{H}_3/\text{H}_4)_2$ à la solution d'enzyme

préalablement incubée avec des concentrations croissantes en inhibiteurs, 20 minutes à température ambiante. La réaction s'est déroulée dans les conditions suivantes : 286 nM de mPRMT6, 0,22 unités d'absorbance à 280 nm de (H3/H4)₂ dilués dans 50 mM Tris-HCl pH 8,0, 100 mM NaCl, 1 mM DTT, 0,04 mg.mL⁻¹ BSA et 1 mM EDTA. La réaction a été arrêtée après une incubation de 3 heures à température ambiante par ajout à la concentration d'usage de tampon de charge SDS-PAGE quatre fois concentré. Les mesures ont été réalisées en triplicats. Les IC₅₀ ont été déterminées par ajustement d'une fonction sigmoïde dose-réponse aux valeurs expérimentales à l'aide du logiciel statistique R-2.13.

b - Mesure par chimio-luminescence (PRMT3, CARM1, PRMT5 et PRMT8)

Le principe de ce test est la reconnaissance d'un substrat méthylé par un anticorps spécifique primaire, puis l'amplification du signal par reconnaissance de l'anticorps primaire par un anticorps secondaire fusionné à la peroxydase, catalysant la synthèse d'un produit luminescent. Il s'agit donc d'une observation indirecte de la formation de produit. Pour réaliser ces évaluations d'IC₅₀, les kits « PRMTx Chemiluminescent Assay » (BPS Bioscience™) ont été utilisés selon le protocole recommandé par le constructeur. Les mesures ont été réalisées en triplicats.

7 - Préparation des récepteurs et ligands pour le « docking » *in silico*

Les ligands ont été modélisés dans le logiciel Maestro (suite Schrodinger™) puis exportés au format mol2. Dans AutoDockTools (suite MGLTools), le récepteur, c'est-à-dire la macromolécule sur laquelle vont être effectués les calculs de docking, a été préparé à partir de la structure obtenue par co-cristallisation de mCARM1 avec le composé SA678 (PDB 5NTC). Seule la chaîne B a été préservée. Après suppression des ligands et molécules d'eau, la structure a été protonée, les types d'atomes assignés selon la topologie AutoDock puis les charges (Kollman) ajoutées. Les hydrogènes « non polaires », c'est-à-dire non associés à un atome électronégatif ont été supprimés. Le récepteur a été sauvegardé au format pdbqt. Les ligands ont été préparés dans AutoDockTools de manière identique à la structure. Les liaisons covalentes pouvant subir des rotations ont été assignées. Dans le cas des ligands, les charges ont été assignées selon la méthode de Gasteiger. Avec AutoGrid4, une grille a été générée autour des sites de fixation du SAM et du BTP. Les calculs de docking ont été effectués avec AutoDock4 selon un algorithme génétique de paramètre suivants : 50 calculs comprenant des populations de 150 conformations avec un maximum de 27000 générations et 2500000 évaluations d'énergie. Les poses ont été triées par niveau d'énergie croissant.

VII - Conclusions et perspectives

Les travaux développés au cours de cette thèse avaient deux objectifs principaux : (I) comprendre au niveau atomique la régulation de la réaction de méthylation des protéines (reconnaissance protéine-protéine et interactions entre modifications post-traductionnelles); (II) concevoir des inhibiteurs spécifiques et à fort potentiel de plusieurs PRMT cibles par une approche rationnelle fondée sur les structures tridimensionnelles.

Pour atteindre ces objectifs nous avons fait appel à différentes méthodes biochimiques et biophysiques et deux méthodes structurales ont été utilisées pour résoudre les structures tridimensionnelles : la diffraction des rayons X sur des cristaux et la cryo-microscopie électronique. Pendant cette thèse, trois PRMT ont été étudiées : PRMT2, PRMT3 et PRMT4/CARM1.

A - Étude structurale de SmPRMT3

Cette étude a été initiée dans un contexte plus large de développement de nouveaux médicaments épigénétiques contre certaines maladies tropicales négligées dont les schistosomias, avec pour objectif de découvrir des inhibiteurs spécifiques de la PRMT3 de *Schistosoma mansoni* (SmPRMT3, 553 acides aminés), une cible prometteuse identifiée dans une première campagne de criblage.

1 - Étude de SmPRMT3 par bio-cristallographie

Durant cette thèse, six constructions de SmPRMT3 ont été clonées : la PRMT3 entière et cinq fragments restreints au module catalytique. Ces fragments ont été exprimés dans *E. coli* avec ou sans « tag » permettant d'améliorer leur solubilité. Les purifications de ces constructions ont été optimisées et trois d'entre elles, dont la PRMT complète, ont été retenues. Les caractérisations biochimiques et biophysiques de ces constructions de SmPRMT3 ont montré que les échantillons sont stables et homogènes, et que cette protéine se comporte comme un dimère en solution.

Des conditions de cristallisation ont été établies pour deux des constructions de SmPRMT3, mais la cristallisation de cette enzyme s'est révélée peu reproductible. Les cristaux obtenus avec la construction complète ont une qualité de diffraction médiocre et une seule des constructions du domaine catalytique de SmPRMT3 (184-553/C3) a permis d'obtenir des cristaux diffractant à une limite de résolution de 7 Å, qui malgré de nombreuses tentatives d'optimisation n'a pas pu être améliorée. Cette construction cristallise dans le groupe d'espace H3₂ avec comme paramètres de maille (a = 159 Å, b = 159 Å, c = 617 Å). Ces grands paramètres suggèrent que les cristaux de

SmPRMT3 possèdent probablement un désordre important, ce qui expliquerait leur faible qualité de diffraction.

La déshydratation des cristaux sera une piste à explorer afin d'améliorer la qualité de la diffraction, mais cette méthode requiert de disposer d'un grand nombre de cristaux et donc une meilleure maîtrise des conditions de cristallisation.

2 - [Étude structurale de SmPRMT3 par cryo-microscopie électronique](#)

Face à la difficulté à résoudre la structure de SmPRMT3 par diffraction des rayons X, nous avons abordé une étude structurale par cryo-microscopie électronique en transmission. Les premières observations ont montré que la construction complète de SmPRMT3 est observable en cryo-microscopie électronique et forme un échantillon homogène après vitrification. Des premières données ont été collectées. Leur analyse a mené à l'obtention de classes 2D de bonne qualité, ce qui est prometteur quant à l'utilisation de cette technologie pour l'étude structurale des PRMT. Toutefois cette expérience a mis en évidence un important phénomène d'orientation préférentielle de SmPRMT3. Pour y remédier, des détergents et des agents tensio-actifs devront être ajoutés à l'échantillon afin de diversifier l'espace des orientations adoptées par SmPRMT3. Des premières tentatives en ce sens ont eu lieu récemment afin de mettre au point de nouvelles conditions de préparation de l'échantillon en présence des détergents CHAPS, DDM et NP40.

3 - [Travaux préliminaires sur le complexe SmPRMT3–SmrpS2](#)

Dans la continuité de l'étude structurale de SmPRMT3, nous avons initié le projet de caractérisation du complexe formé par PRMT3 avec son substrat rpS2. RpS2 est une protéine ribosomique de 28 kDa conservée chez les organismes eucaryotes et impliquée dans la maturation du ribosome. Établir les bases structurales de l'interaction entre PRMT3 et rpS2 représenterait une avancée dans la compréhension des mécanismes moléculaires de reconnaissance PRMT–partenaire.

Pour étudier ce complexe, nous avons choisi d'utiliser comme modèle SmPRMT3, dont l'expression et la purification étaient déjà bien maîtrisées. SmrpS2 s'est révélé peu soluble et son expression seule ou en complexe avec SmPRMT3 dans *E. coli* a été un échec. Cette difficulté a été contournée en co-exprimant SmPRMT3 avec différents fragments de SmrpS2 à l'aide d'un système d'expression baculovirus-cellules d'insecte : plusieurs constructions de la protéine SmrpS2 entière et sous forme de domaines isolés ont été clonées et sur-exprimées avec SmPRMT3 dans ce système d'expression. Nous avons montré que le domaine de fixation à l'ARN de rpS2 est nécessaire et suffisant à sa fixation à PRMT3 et sommes récemment parvenus à purifier plusieurs constructions solubles de ce complexe. Nous tentons maintenant de caractériser ce complexe afin de procéder à son étude structurale par cryo-microscopie électronique.

En 2016, la protéine humaine PDCD2L (39 kDa) a été identifiée dans un complexe ternaire avec PRMT3 et rpS2. Ce complexe de plus haut poids moléculaire représente un intérêt non négligeable pour la compréhension de la reconnaissance PRMT–substrat et dans le processus de biogenèse du ribosome et nous allons caractériser ce complexe en vue d'études structurales.

B - Étude structurale et fonctionnelle de PRMT2

PRMT2 est la moins caractérisée des PRMT de type I. Elle est connue pour son rôle de coactivateur des récepteurs nucléaires et son implication dans la régulation de l'épissage, mais jusqu'en 2017 sa structure atomique n'était pas connue et aucun substrat n'avait été identifié. Au sein de l'équipe, au cours de cette thèse, nous avons résolu les structures cristallographiques à haute résolution des PRMT2 de souris et de poisson zèbre. Ce travail a donné lieu à une publication (Cura et al., 2017).

Les structures montrent que PRMT2 forme un dimère caractéristique des PRMT et possède un site actif conservé caractéristique des PRMT de type I. Bien que PRMT2 de souris ait été cristallisée entière et que la protéine soit entière dans les cristaux, l'extrémité amino-terminal comportant un domaine SH3 n'est pas visible dans les cartes de densité électronique, ce qui suggère qu'il est flexible. La fonction exacte de ce domaine n'est pas bien caractérisée, mais nous avons démontré que sa présence stimule l'activité enzymatique de PRMT2, suggérant un rôle dans la fixation du substrat. Durant cette étude, nous avons identifié un premier inhibiteur de PRMT2 et déterminé la structure tridimensionnelle de PRMT2 avec plusieurs autres inhibiteurs.

Nous avons également montré que le facteur d'épissage rsf1 est partenaire de PRMT2 et que le complexe formé est stable et homogène en solution. Les analyses par spectrométrie de masse ont montré que PRMT2–rsf1 forme un complexe ternaire constitué d'un dimère de PRMT2 fixant un monomère de rsf1. Les essais de co-cristallisation de PRMT2 et rsf1 n'ont pas encore permis d'obtenir de cristaux de ce complexe.

Dans les derniers mois de la thèse, nous avons abordé l'étude structurale de ce complexe par cryo-microscopie électronique. Une première structure à une résolution de 7 Å a été obtenue. Le modèle montre un dimère de PRMT2 semblable au dimère cristallographique arborant sur un seul des monomères une densité supplémentaire correspondant probablement à la protéine rsf1 lié au domaine amino-terminal de PRMT2. Ce modèle contient donc le complexe dans son entier, confirme la stœchiométrie mesurée en spectrométrie de masse et représente, bien qu'à basse résolution, le premier modèle d'une PRMT fixé à un substrat complet. Ce modèle apporte une évidence supplémentaire que le domaine SH3 de PRMT2 est nécessaire à la fixation du substrat et montre une certaine hétérogénéité conformationnelle dans le positionnement du domaine SH3 et du substrat qui lui associé.

Nous nous efforçons aujourd'hui d'améliorer la résolution de ce modèle. La préparation des grilles de microscopie électronique a été considérablement améliorée depuis les premiers travaux et le traitement des données d'une récente acquisition indique que nous devrions pouvoir pousser la résolution du modèle au-delà de 7 Å. Pour atteindre une haute résolution, un nombre plus important de particules sera nécessaire et la composition du tampon pourrait encore être améliorée pour en diminuer le bruit de fond. De plus, l'hétérogénéité conformationnelle du complexe pourra être réduite si nous souhaitons extraire des informations à plus haute résolution sur les contacts intermoléculaires régissant la fixation de rsf1 au domaine SH3 de PRMT2. Pour cela, une approche par réticulation pourrait être envisagée.

C - Étude structurale des complexes CARM1-peptides

En continuité des travaux développés dans l'équipe depuis de nombreuses années (Ajebar, 2012 ; Mailliot, 2013), nous avons mis au point et utilisé des outils moléculaires mimes de l'état de transition de la réaction de méthylation pour stabiliser les complexes PRMT avec leurs substrats.

Des composés constitués d'un substrat/produit polypeptidique lié à un analogue du cofacteur SAM ont été conçus et synthétisés. Le pontage covalent d'un substrat à un cofacteur s'est avéré dans d'autres cas comme une stratégie efficace pour capturer et caractériser des complexes enzyme-substrat peu stables. La stratégie induite par cette approche est que ces analogues ou mines de l'état de transition pourraient servir de base à la conception de nouveaux inhibiteurs spécifiques ou sélectifs d'une PRMT donnée.

Avant cette thèse, la nature chimique des analogues SAM-peptides utilisés n'avaient pas permis de visualiser les peptides dans les cartes de densité électronique. Dans la continuité de cette approche, nous avons mis au point durant cette thèse de nouveaux types d'analogues SAM-peptides qui ont permis de capturer et de stabiliser des complexes CARM1-peptides.

Quatre structures cristallographiques à haute résolution du complexe CARM1-peptide PABP1 ont été obtenues, démontrant que cette approche par ancrage permet la stabilisation de complexes CARM1-peptides. En comparant nos structures avec les structures cristallographiques de CARM1 humaine en complexe avec des peptides PABP1 non modifié obtenues fin 2015 par Copeland et ses collaborateurs, nous avons montré que le peptide modifié adopte la même configuration que le peptide naturel, ce qui indique que l'arginine modifiée ne perturbe pas sa fixation. De plus, l'ancrage du peptide a permis de mieux structurer une plus grande partie du peptide à la surface du module catalytique de CARM1. Nous avons aussi montré que l'utilisation de dérivés peptidiques n'introduit pas de distorsion dans le site catalytique de CARM1, contrairement aux structures résolues par Copeland et ses collaborateurs. Nos composés figent donc le site actif de

la PRMT dans un instantané de la réaction de méthylation. Ce travail a donné lieu à une publication (van Haren et al., 2017).

Suite à ce premier succès, nous avons commencé à généraliser cette approche à d'autres substrats de CARM1 et à des substrats d'autres PRMT, avec plusieurs objectifs :

- établir un modèle commun de reconnaissance PRMT-peptide ou au contraire démontrer que les différentes PRMT n'adoptent pas les mêmes stratégies de reconnaissance de leurs substrats,
- expliquer comment les PRMT peuvent méthyler sélectivement plusieurs sites différents sur un même substrat,
- comprendre les bases structurales du dialogue entre les modifications épigénétiques des histones et la méthylation des arginines par les PRMT.

Des peptides dérivés de H3, H4 et rsf1 respectivement spécifiques de CARM1, PRMT1 / PRMT6 / PRMT7 et PRMT2 ont été synthétisés.

Des peptides dérivés de H3 (13-31) et H3 (1-41), plus longs que ceux utilisés précédemment, ont été synthétisés afin de comprendre les mécanismes de reconnaissance à courte et longue distance identifiés avec CARM1. Deux versions de ces peptides, avec ou sans acétylation sur la lysine 18 (proximale du site de méthylation R17), ont aussi été synthétisés pour déterminer les bases structurales du dialogue entre ces deux modifications post-traductionnelles. Des cristaux de complexes mCARM1-dérivés de H3 ont été obtenus. 48 jeux de données de diffraction ont été collectés et les trois meilleurs d'entre eux ont été totalement affinés.

Les structures cristallographiques montrent que l'utilisation de peptides longs (1-41) ne permet pas d'augmenter le nombre de résidus visibles dans la structure comparés aux peptides courts (13-31). L'acétylation de K18 ajoute quelques interactions supplémentaires qui ne suffisent pas à elles seules à expliquer l'antériorité de l'acétylation de K18 sur la méthylation de R17 *in vivo*. En plus de ces interactions supplémentaires, la neutralisation de la charge de la lysine due à son acétylation facilite probablement la désolvatation du segment de substrat, ce qui pourrait faciliter sa fixation à CARM1. La modification de H3K18 pourrait également agir par l'intermédiaire d'une protéine reconnaissant l'acétylation puis recrutant CARM1 pour la méthylation subséquente de R17.

Plusieurs directions vont maintenant être prises pour donner suite à cette approche. Des peptides H3 centrés sur le second site de méthylation par CARM1, H3R26, avec et sans acétylation de K27 ont été récemment synthétisés puis co-cristallisés avec CARM1. Ces peptides nous serviront à explorer les mécanismes de sélectivité de cette enzyme. Nous envisageons également la synthèse de dérivés peptidiques dont l'arginine modifiée incorpore un groupement méthyl supplémentaire, afin d'avoir des mimes de la seconde étape de la réaction de diméthylation.

Finalement, nous espérons pouvoir étendre la stratégie des dérivés peptidiques à l'ancrage de substrats complets. Cela pourrait être réalisé par des approches de ligation chimique en condition native, adaptée à la ligation de segments protéiques non structurés, ou par incorporation *in vivo* de résidus non naturels, employée pour des segments non structurés comme pour des sections globulaires.

- Par ligation chimique en conditions natives : cette méthode consiste à faire réagir un peptide arborant une fonction chimique électrophile sur sa chaîne principale avec un second peptide arborant une fonction chimique nucléophile, par exemple la chaîne latérale d'une cystéine. La réaction engendre la formation d'une liaison covalente. Nous souhaiterions utiliser cette technologie pour relier un peptide incorporant l'arginine modifiée au reste du substrat correspondant, et nous réaliserons des premiers essais avec le complexe PRMT2-rsf1 et CARM1-H3.
- Par incorporation *in vivo* de résidus non naturels lors de la synthèse des protéines : cette méthode élégante nécessite l'ingénierie génétique d'un organisme pour lui permettre d'incorporer un acide aminé non naturel ajouté au milieu de culture, dans notre cas une arginine arborant un groupement adénosine, lors de la synthèse de la protéine. Le signal d'incorporation étant un codon stop détourné. Cette méthode nécessite cependant la conception d'un couple aminoacyl-ARNt-synthétases/ARNt adapté.

D - Analyse structurale et fonctionnelle du mutant P409A

Les résidus strictement concernés sont très peu nombreux chez les PRMT, et la plupart sont directement impliqués dans le mécanisme catalytique et appartiennent aux motifs caractérisant cette famille d'enzyme. Un de ces résidus strictement conservés dans les structures des PRMT de type I, II et III, est une proline (P409 chez CARM1 de souris) situé directement en amont du motif IV. Les alignements de séquences et structures de PRMT montre que sa présence est corrélée à un état productif du module catalytique. Ce résidu se localise sur les surfaces externes du dimère catalytique, et est systématiquement placé à la verticale d'un résidu aromatique placé en N-1 d'un second résidu aromatique interagissant lui-même avec le tryptophane strictement conservé du motif IV dans le site actif.

Ce résidu n'avait jusqu'à aujourd'hui été le sujet d'aucune étude, mais sa fréquence et le contexte structural dans lequel il se trouve nous a poussés à nous y intéresser. Des analyses biophysiques, enzymatiques et structurales ont été réalisées sur le mutant CARM1 P409A afin d'étudier l'importance de ce résidu.

Nous avons montré que la mutation P409A n'affecte pas la stabilité thermique comparée à l'enzyme non mutante. La structure cristallographique de CARM1 P409A a aussi montré que la mutation ne modifie pas la structure du module catalytique, ni globalement, ni localement autour de

la mutation, la perte de la chaîne latérale de la proline étant compensée par une molécule d'eau. Ces données suggèrent que cette proline n'est pas à elle seule nécessaire à l'organisation structurale du motif IV, mais que l'ensemble du contexte structural environnant y participe probablement.

Des mesures de l'activité enzymatique ont montré que la mutation ne perturbe pas l'activité de l'enzyme et nous avons mesuré par thermophorèse que les enzymes WT et mutante possèdent une affinité comparable pour le substrat H3. Ce dernier point reste néanmoins à confirmer.

Ces données ne permettent pas aujourd'hui d'expliquer la conservation stricte de cette proline dans tous les modules catalytiques productifs des PRMT. Situé dans un tournant de type II nécessaire au positionnement du motif IV, il est possible que la proline soit le résidu le plus adapté du point de vue évolutif pour cette structure locale. Néanmoins, cela n'explique pas sa non-conservation dans les modules C-terminal inactifs de PRMT7 et PRMT9.

Des analyses enzymologiques supplémentaires devront être réalisées afin de déterminer si ce résidu joue un rôle dans la cinétique de la réaction de méthylation. Aussi, cette étude devrait être menée sur d'autres PRMT de type I, II et III afin de proposer une explication globale sur la fonction de ce résidu. Pour finir, il est possible que ce résidu participe à un mécanisme commun de régulation de l'activité des PRMT, et l'identification des éléments régulateurs sera alors nécessaire pour comprendre l'importance de cette proline.

E - Conception d'inhibiteurs sélectifs des PRMT

Depuis 2008, l'équipe utilise des composés analogues du SAM pour servir de base à la conception de nouveaux inhibiteurs spécifiques et sélectifs d'une PRMT donnée. Lors de cette thèse, nous avons travaillé sur quatre familles de composés analogues du SAM, dans la continuité des approches développées dans l'équipe (Ajebbar, 2012 ; Mailliot, 2013). Ces familles de composés se différencient par la nature chimique des groupements liés à la base adénosine du SAM, permettant une exploration plus diversifiée des différentes poches du site catalytique des PRMT. Les quatre types de composés étudiés sont succinctement présentés ci-dessous :

- les composés U (U1 à U3) ont été développés par l'équipe de Nathaniel Martin (université d'Utrecht),
- les composés EML ont été développés par l'équipe de Gianluca Sbardella (université de Salerne) et sont les premiers inhibiteurs qui ciblent à la fois PRMT1 et PRMT7,
- les composés LH ont été développés par l'équipe de Paola Arimondo (à l'époque à ETaC Toulouse). Ces composés avaient initialement été conçus comme des inhibiteurs des DNMT,
- les composés SATT ont été mis au point au sein de l'équipe durant cette thèse. Ces inhibiteurs interagissent avec une poche nouvellement identifiée sur le module catalytique des PRMT.

1 - [Les composés LH](#)

Les composés LH, analogues de l'état de transition de DNMT, ont été initialement conçus pour inhiber les ADN et ARN méthyltransférases. N'ayant montré aucune inhibition pour cette classe d'enzymes, nous avons montré que ces composés inhibaient sélectivement les PRMT. Nous avons co-cristallisé ces inhibiteurs avec CARM1 et PRMT6 et avons résolu les structures cristallographiques à haute résolution de quatre complexes CARM1–LH et deux complexes PRMT6–LH. Ces structures montrent que l'inhibition des PRMT par ces composés est corrélée avec le déplacement de résidus essentiels du site catalytique. Ce travail a donné lieu à une publication (Halby et al., 2018).

2 - [Les composés U](#)

Dans le cadre de notre collaboration avec l'équipe de Nathaniel Martin, nous avons co-cristallisé les composés U avec PRMT2, CARM1 et PRMT6. Nous avons résolu les structures cristallographiques à haute résolution de PRMT2 et CARM1 en complexe avec trois composés U, et PRMT6 en complexe avec deux composés U. L'analyse des interactions dans la poche de méthylation a mis en évidence les différences subtiles qui établissent la sélectivité de ces inhibiteurs pour PRMT2, CARM1 ou PRMT6. Une partie du travail a été publiée avec la structure de PRMT2 (Cura et al., 2017), le reste fait partie d'un article en préparation.

3 - [Les composés EML](#)

Notre collaborateur Gianluca Sbardella a montré que les composés EML sont les premiers inhibiteurs sélectifs de PRMT1 et PRMT7, et dans une moindre mesure de CARM1. Aucune structure cristallographique de PRMT7 et PRMT1 n'a pu être obtenue avec ces composés. Néanmoins, nous avons pu résoudre la structure à haute résolution de PRMT6 en complexe avec deux composés EML et de CARM1 en complexe avec trois composés EML. Pour CARM1, les structures révèlent le mode de fixation des composés EML et mettent en évidence les différences qui établissent la sélectivité de ces inhibiteurs. Les structures de PRMT6 en complexe avec des EML dévoilent que ceux-ci adoptent une configuration radicalement différente qui pourrait expliquer la très mauvaise inhibition mesurée. Une publication sur ce travail est en préparation.

4 - [Développement des composés SATT](#)

L'identification d'une région fixant un additif de cristallisation, le BTP, à la surface de mCARM1 nous a conduit à caractériser un potentiel site pouvant permettre de développer une nouvelle famille d'inhibiteurs. Ce site se localise à la position normalement occupée par l'hélice α -W de CARM1, à proximité directe du groupement adénine du SAM fixé dans le site catalytique. En

collaboration avec l'équipe du docteur Alain Wagner, nous avons conçus et synthétisé deux composés incorporant des groupements phénylpropanediol ou boc-sérinol liés à la position 7 de la structure du composé mimant l'adénine. Nous avons montré que ces deux composés, SATT3001 et SATT3002, inhibent sélectivement PRMT1 et CARM1 à des concentrations sub-millimolaires.

Nous avons co-cristallisé CARM1 en présence de chacun de ces 2 inhibiteurs. Les structures cristallographiques confirment que les groupements phénylpropanediol et boc-sérinol occupent le site « SATT » tel qu'imaginé lors de la conception de ces composés. De plus, les densités électroniques autour des groupements phénylpropanediol et boc-sérinol suggèrent que ceux-ci sont légèrement flexibles. Cette instabilité est due au peu de liaisons établies avec CARM1, ce qui indique qu'en choisissant des groupements chimiques permettant de mieux interagir avec l'enzyme nous pourrions significativement améliorer l'affinité et la sélectivité de ces composés. Une publication est en préparation sur ce travail.

Les prochaines étapes seront de déterminer les groupements chimiques permettant d'interagir étroitement avec le site SATT. Ensuite, nous devons remplacer les groupements adénosine et guanidinium par des structures chimiques plus adaptées à l'utilisation comme composés thérapeutiques, c'est-à-dire par des fonctions chimiques plus hydrophobes, comprenant moins de donneurs et accepteurs de liaisons hydrogène, moins flexibles et dont la structure serait moins métabolisable par l'organisme. Ces composés seront testés sur des lignées cellulaires cancéreuses et saines pour mesurer leur effet anti-prolifératif.

Index des figures

Figure 1 - Evolution du nombre de publications sur l'épigénétique depuis 1980.....	9
Figure 2 - Représentation schématique de la compaction de l'ADN.....	10
Figure 3 - Schéma des familles de complexes impliqués dans la régulation épigénétique.....	11
Figure 4 - Représentation schématique des modifications des histones.....	13
Figure 5 - Méthylation des arginines par les PRMT.....	17
Figure 6 - Représentation schématique des neuf PRMT caractérisées chez les mammifères.....	18
Figure 7 - Activation de la transcription par PRMT1 en synergie avec p300.....	22
Figure 8 - Coactivation de la transcription par les récepteurs nucléaires par CARM1 et p300/CBP.....	29
Figure 9 - Module catalytique des PRMT.....	41
Figure 10 - Poches de fixation des substrats des PRMT.....	42
Figure 11 - Motifs caractéristiques du module catalytique des PRMT.....	43
Figure 12 - Structure des motifs conservés dans les PRMT de types I, II et III.....	45
Figure 13 - Comparaison de l'organisation spatiale des PRMT.....	46
Figure 14 - Modèle de PRMT2.....	47
Figure 15 - Modèle de PRMT3.....	48
Figure 16 - Particularités de CARM1.....	49
Figure 17 - Structure cristallographique du dimère de PRMT5.....	50
Figure 18 - Structure du complexe PRMT5-MEP50.....	50
Figure 19 - Structure et spécificités de la PRMT7 de souris.....	51
Figure 20 - Schéma de la mono-méthylation d'une arginine par CARM1.....	53
Figure 21 - Déterminants structuraux de la diméthylation asymétrique.....	54
Figure 22 - Déterminants structuraux de la diméthylation symétrique.....	55
Figure 23 - Reconnaissance des diméthylarginines par les domaines Tudor.....	58
Figure 24 - Mode de fixation du domaine jmjC de JMJD5 à ses substrats.....	60
Figure 25 - Inhibiteurs compétitifs naturels des PRMT.....	61
Figure 26 - Sites de fixation des inhibiteurs des PRMT.....	62
Figure 27 - Inhibiteurs analogues du SAM.....	65
Figure 28 - Exemples de composés SA dans le site catalytique de CARM1 de souris.....	66
Figure 29 - Inhibiteurs non analogues du SAM : les AMI.....	67
Figure 30 - Inhibiteurs non analogues du SAM : allantodapsone et ses dérivés.....	67
Figure 31 - Inhibiteurs non analogues du SAM : les amidines.....	68
Figure 32 - Inhibiteurs non analogues du SAM : les cyanines.....	69
Figure 33 - Inhibiteurs non analogues du SAM : les dérivés de la curcumine.....	70
Figure 34 - Inhibiteurs non analogues du SAM : les dérivés du pyrazole et de l'imidazole.....	71
Figure 35 - Modes de fixation des composés CMPD-1 et CMPD-2 à CARM1.....	72
Figure 36 - inhibiteurs non analogues du SAM : les allostériques.....	73
Figure 37 - Mode de fixation de SGCZ07 à la PRMT3 humaine.....	74
Figure 38 - Inhibiteurs non analogues du SAM : autres cas.....	75
Figure 39 - Modes de fixation des composés EPZ015866 et EZM2302 à PRMT5 et CARM1.....	76
Figure 40 - Carte de répartition de la schistosomiase.....	79
Figure 41 - Modèle par homologie de SmPRMT3.....	81
Figure 42 - Trypsinolyse de SmPRMT3-C1.....	83
Figure 43 - Modèle par homologie de SmPRMT3 avec indication des constructions utilisées.....	84
Figure 44 - Purification de 6his-SmPRMT3-fl.....	86
Figure 45 - Purification de SmPRMT3-C1.....	87
Figure 46 - Purification des constructions SmPRMT3-C2/-C3 et 6his-Trx-TEV-SmPRMT3-C3.....	88
Figure 47 - Optimisation du tampon de purification pour la construction SmPRMT3-C3.....	89
Figure 48 - Purification de 6his-C3 et -C4.....	90
Figure 49 - Méthylation de rsf1 de <i>S. frugiperda</i> par SmPRMT3-fl.....	91
Figure 50 - Caractérisation de SmPRMT3-fl et -C3 : diffusion dynamique de la lumière.....	92
Figure 51 - Caractérisation de SmPRMT3-fl et -C3 : diffusion des rayons-X aux petits angles.....	94
Figure 52 - Caractérisation de SmPRMT3-fl / C3 : spectrométrie de masse.....	95

Figure 53 - Caractérisation de SmPRMT3-C3 : stabilité thermique.....	96
Figure 54 - Cristallisation de SmPRMT3-fl et -C3.....	98
Figure 55 - Observation de 6his-SmPRMT3-fl en cryo-EM en transmission.....	100
Figure 56 - Traitement des données collectées sur le microscope Titan Krios.....	102
Figure 57 - La protéine eucaryote rpS2 dans son contexte ribosomique.....	106
Figure 58 - Expression et solubilité de SmrpS2 dans <i>E. coli</i>	107
Figure 59 - Interaction des complexes SmPRMT3 – SmrpS2.....	108
Figure 60 - Purification des complexes PRMT3-rpS2.....	110
Figure 61 - Reconstruction du complexe PRMT2-rsf1.....	127
Figure 62 - Analyse par spectrométrie de masse du complexe PRMT2 – rsf1.....	128
Figure 63 - Criblage des conditions de préparation des grilles pour le complexe PRMT2-rsf1.....	130
Figure 64 - Structure à basse résolution du complexe PRMT2-rsf1.....	131
Figure 65 - Schéma d'organisation d'une grille de cryo-EM perforée de type R22.....	132
Figure 66 - Second échantillon du complexe PRMT2 – rsf1.....	133
Figure 67 - Stratégies d'ancrage de peptides aux PRMT.....	140
Figure 68 - Analogues « SAM-peptide ».....	141
Figure 69 - Schéma de la reconnaissance CARM1 peptide.....	144
Figure 70 - Sites de méthylation par CARM1 et d'acétylation par p300 sur l'histone H3.....	145
Figure 71 - Sites de fixation des molécules de PEG sur CARM1.....	148
Figure 72 - Structures cristallographiques de mCARM1 en complexe avec les dérivés de H3.....	150
Figure 73 - Modes de fixation des peptides H3 (13-31) K18 et K18 ^{Ac} à mCARM1.....	153
Figure 74 - Sites de méthylation sur H4 et rsf1.....	155
Figure 75 - Conservation de la structure primaire du module catalytique des PRMT.....	159
Figure 76 - Stabilité thermique de mCARM1 P409A vs WT.....	162
Figure 77 - Structure cristallographique du mutant mCARM1 P409A.....	163
Figure 78 - Activité enzymatique de CARM1.....	164
Figure 79 - Affinité de mCARM1 pour PABP1 (437-488).....	165
Figure 80 - Représentation schématique des composés EML, U et SATT.....	175
Figure 81 - Structures cristallographiques de CARM1 en complexe avec les composés U.....	178
Figure 82 - Structures cristallographiques de PRMT2 en complexe avec les composés U.....	181
Figure 83 - Structures cristallographiques de mPRMT6 en complexe avec les inhibiteurs U.....	182
Figure 84 - Structures cristallographiques de mCARM1 en complexe avec les inhibiteurs EML..	186
Figure 85 - Structure cristallographique de mPRMT6 en complexe avec les inhibiteurs EML.....	189
Figure 86 - Identification d'une nouvelle poche à la surface de CARM1.....	191
Figure 87 - Poches de fixation de l'arginine substrat, du SAM et du site SATT.....	192
Figure 88 - Résultats de docking des composés WS2 et WS3 dans la structure de mCARM1.....	193
Figure 89 - Structure de mCARM1 en complexe avec les inhibiteurs SATT300X.....	195

Index des alignements

Alignement 1 - Alignement de séquences des modules catalytiques des PRMT.....	19
Alignement 2 - Alignement de séquences des PRMT de type I.....	82
Alignement 3 - Alignement de séquences de rpS2 eucaryotes et rpS5 bactérien.....	105
Alignement 4 - Alignement de séquences de quelques PRMT de type I, II et III.....	160

Index des tableaux

Tableau 1 - Identités de séquence entre les modules catalytiques des PRMT.....	20
Tableau 2 - Répartition des PRMT chez les eucaryotes.....	20
Tableau 3 - Fonctions, localisations cellulaires et substrats des PRMT.....	21
Tableau 4 - Inhibiteurs de PRMT publiés et cristallisés avec leurs cibles spécifiques.....	63
Tableau 5 - Conditions d'expression et comportements en purification des constructions utilisées..	86
Tableau 6 - Caractérisation de SmPRMT3-fl et -C3 : diffusion dynamique de la lumière.....	93
Tableau 7 - Caractérisation de SmPRMT3-fl et -C3 : Diffusion des rayons-X aux petits angles.....	93
Tableau 8 - Caractérisation de SmPRMT3-fl / C3 : spectrométrie de masse.....	95
Tableau 9 - Résumé des conditions de cristallisation testés pour SmPRMT3.....	97
Tableau 10 - Résumé des cristaux de complexes mCARM1-dérivés de H3 collectés.....	146
Tableau 11 - Résumé des structures obtenues à partir des conditions en malonate.....	149
Tableau 12 - Inhibition (IC ₅₀) des PRMT par les composés U.....	176
Tableau 13 - Stabilisation thermique des PRMT par les composés U.....	177
Tableau 14 - Structures cristallographiques et interactions de mCARM1 avec les composés U.....	179
Tableau 15 - Structures cristallographiques et interactions de mPRMT2 avec les composés U.....	181
Tableau 16 - Structures cristallographiques et interactions de mPRMT6 avec les composés U.....	183
Tableau 17 - Inhibition des PRMT par les composés EML.....	184
Tableau 18 - Stabilisation thermique des PRMT par les composés EML.....	185
Tableau 19 - Structures cristallographiques et interactions des complexes mCARM1-EML.....	187
Tableau 20 - Structures cristallographiques et interactions des complexes mPRMT6-EML.....	189
Tableau 21 - Inhibition des PRMT par les composés SATT3001 et SATT3002.....	194

Index des annexes

Annexe 1 - Séquences des constructions de SmPRMT3 fl à C2.....	216
Annexe 2 - Séquences des constructions de SmPRMT3 C3 à C5.....	217
Annexe 3 - Données cristallographiques pour SmPRMT3.....	218
Annexe 4 - Données cristallographiques pour les structures de mCARM1 WT et P409A.....	219
Annexe 5 - Données cristallographiques pour les structures de mPRMT2 et zPRMT2.....	220
Annexe 6 - Données cristallographiques pour les complexes mCARM1-dérivés peptidiques (1)..	221
Annexe 7 - Données cristallographiques pour les complexes mCARM1-dérivés peptidiques (2)..	222
Annexe 8 - Données cristallographiques pour les complexes mCARM1-dérivés peptidiques (3)..	223
Annexe 9 - Données cristallographiques pour les complexes mCARM1-LH (1).....	224
Annexe 10 - Données cristallographiques pour les complexes mCARM1-LH (2).....	225
Annexe 11 - Données cristallographiques pour les complexes mPRMT6-LH.....	226
Annexe 12 - Données cristallographiques pour les complexes mCARM1-U.....	227
Annexe 13 - Données cristallographiques pour les complexes mPRMT2-U.....	228
Annexe 14 - Données cristallographiques pour les complexes mPRMT6-U.....	229
Annexe 15 - Données cristallographiques pour les complexes mCARM1-EML.....	230
Annexe 16 - Données cristallographiques pour les complexes mPRMT6-EML.....	231
Annexe 17 - Données cristallographiques pour les complexes mCARM1-SATT.....	232

Annexes

VIII - Annexes

SmPRMT3 (1-553 / fl) MW 63382,4 Da pI 5,13 Eps(red) 69790 L.mol⁻¹.cm⁻¹

MTDSNPMEYD CDSDDWQEVV LPDLSLSEKY TCLMCENHFN TAPLFFQHLT LQHDWVNFLD
GEGKSIFNDQ YDWIRFVNFV RKTTPVDLRQ AFYDMVWKT DLDLLYPVLP DDEVLSIDIE
SYLLDNQRSN QSPGIVVHEK STNQQPVSQE MSRIMAENKS LRTQLNNCKR LLSTMSAQSV
RRKSPCDQSS NYDLDMEPND SSYFGSYGHF EIHGEMINDR VRTESYVNF I LSNAEKYFKH
KIILDVSGSGS GILSIIAAQA GASHVYGVEA ADEIYAASHE TLRVNLLER VTFIHGOAES
VELPVKKVDV IISEWMGYFL FFESMLDSVL KMASKYLSRD GHIFPRHYTL NLLGVQCSEQ
LRKRRLEHWN NVYGYNMPAL RRAALSEAHV LNLTNEHVTP PISPITILTQ SFELVALDLD
DMHRNRIYNL SNHCSLLCEQ KFHLTIQPTT DINNNSSSSS YELDAIVGYF DVRFDDADC
KVEFSTSPTT PLTHWKQTLT FLDKPIRVKP GDKISGIITI RRATTDNRGL EINLLIGETE
NSLEIKQTFD LIG

SmPRMT3 (141-553 / C1) MW 46877,04 Da pI 5,85 Eps(red) 38850 L.mol⁻¹.cm⁻¹

MTDSNPMEYD CDSDDWQEVV LPDLSLSEKY TCLMCENHFN TAPLFFQHLT LQHDWVNFLD
GEGKSIFNDQ YDWIRFVNFV RKTTPVDLRQ AFYDMVWKT DLDLLYPVLP DDEVLSIDIE
SYLLDNQRSN QSPGIVVHEK STNQQPVSQE MSRIMAENKS LRTQLNNCKR LLSTMSAQSV
RRKSPCDQSS NYDLDMEPND SSYFGSYGHF EIHGEMINDR VRTESYVNF I LSNAEKYFKH
KIILDVSGSGS GILSIIAAQA GASHVYGVEA ADEIYAASHE TLRVNLLER VTFIHGOAES
VELPVKKVDV IISEWMGYFL FFESMLDSVL KMASKYLSRD GHIFPRHYTL NLLGVQCSEQ
LRKRRLEHWN NVYGYNMPAL RRAALSEAHV LNLTNEHVTP PISPITILTQ SFELVALDLD
DMHRNRIYNL SNHCSLLCEQ KFHLTIQPTT DINNNSSSSS YELDAIVGYF DVRFDDADC
KVEFSTSPTT PLTHWKQTLT FLDKPIRVKP GDKISGIITI RRATTDNRGL EINLLIGETE
NSLEIKQTFD LIG

SmPRMT3 (171-553 / C2) MW 43402,10 Da pI 5,59 Eps(red) 38850 L.mol⁻¹.cm⁻¹

MTDSNPMEYD CDSDDWQEVV LPDLSLSEKY TCLMCENHFN TAPLFFQHLT LQHDWVNFLD
GEGKSIFNDQ YDWIRFVNFV RKTTPVDLRQ AFYDMVWKT DLDLLYPVLP DDEVLSIDIE
SYLLDNQRSN QSPGIVVHEK STNQQPVSQE MSRIMAENKS LRTQLNNCKR LLSTMSAQSV
RRKSPCDQSS NYDLDMEPND SSYFGSYGHF EIHGEMINDR VRTESYVNF I LSNAEKYFKH
KIILDVSGSGS GILSIIAAQA GASHVYGVEA ADEIYAASHE TLRVNLLER VTFIHGOAES
VELPVKKVDV IISEWMGYFL FFESMLDSVL KMASKYLSRD GHIFPRHYTL NLLGVQCSEQ
LRKRRLEHWN NVYGYNMPAL RRAALSEAHV LNLTNEHVTP PISPITILTQ SFELVALDLD
DMHRNRIYNL SNHCSLLCEQ KFHLTIQPTT DINNNSSSSS YELDAIVGYF DVRFDDADC
KVEFSTSPTT PLTHWKQTLT FLDKPIRVKP GDKISGIITI RRATTDNRGL EINLLIGETE
NSLEIKQTFD LIG

Annexe 1 - Séquences des constructions de SmPRMT3 fl à C2.

La séquence correspondant aux constructions est représentée en noir ou en couleur. Les domaines amino-terminal, de Rossmann, le tonneau β et le bras de dimérisation sont colorés respectivement en vert, orange, bleu clair et violet. Les motifs I, II III et IV sont colorés respectivement en bleu foncé, rose, rouge et vert. MW : poids moléculaire. pI : point isoélectrique. Eps(red) : coefficient d'extinction moléculaire en considérant que les cystéines sont réduites.

SmPRMT3 (184-553 / C3) MW 41943,36 Da pI 5,36 Eps(red) 38850 L.mol⁻¹.cm⁻¹

MTDSNPMEYD CDSDDWQEV T LPDLSLSEKY TCLMCENHFN TAPLFFQHLT LQHDWVNFLD
GEGKSIFNDQ YDWIRFVNFV RKTTPVDLRQ AFYDMVWKT D SLDLLYPVLP DDEVLSIDIE
SYLLDNQRSN QSPGIVVHEK STNQOPVSQE MSRIMAENKS LRTQLNNCKR LLSTMSAQSV
RRKSPCDQSS NYDLDMEPND SSYFGSYGHF EIHGEMINDR VRTESYVNF I LSNAEKYFKH
KIILDVGS GS GILSIIAAQA GASHVYGVEA ADEIYAASHE TLRVNNLLER VTFIHGQAES
VELPVKKVDV IISEWMGYFL FFESMLDSVL KMASKYLSRD GHIFPRHYTL NLLGVQCSEQ
LRKRRLEHWN NVYGYNMPAL RRAALSEAHV LNLTNEHVTP PISPITILTQ SFELVALDLD
DMHRNRIYNL SNHCSSLCEQ KFHLTIQPTT DINNNSSSSS YELDAIVGYF DVRFDDADC
KVEFSTSPTT PLTHWKQTLL FLDKPIRVKP GDKISGIITI RRATTDNRGL EINLLIGETE
NSLEIKQTFD LIG

SmPRMT3 (199-553 / C4) MW 40260,61 Da pI 5,66 Eps(red) 37360 L.mol⁻¹.cm⁻¹

MTDSNPMEYD CDSDDWQEV T LPDLSLSEKY TCLMCENHFN TAPLFFQHLT LQHDWVNFLD
GEGKSIFNDQ YDWIRFVNFV RKTTPVDLRQ AFYDMVWKT D SLDLLYPVLP DDEVLSIDIE
SYLLDNQRSN QSPGIVVHEK STNQOPVSQE MSRIMAENKS LRTQLNNCKR LLSTMSAQSV
RRKSPCDQSS NYDLDMEPND SSYFGSYGHF EIHGEMINDR VRTESYVNF I LSNAEKYFKH
KIILDVGS GS GILSIIAAQA GASHVYGVEA ADEIYAASHE TLRVNNLLER VTFIHGQAES
VELPVKKVDV IISEWMGYFL FFESMLDSVL KMASKYLSRD GHIFPRHYTL NLLGVQCSEQ
LRKRRLEHWN NVYGYNMPAL RRAALSEAHV LNLTNEHVTP PISPITILTQ SFELVALDLD
DMHRNRIYNL SNHCSSLCEQ KFHLTIQPTT DINNNSSSSS YELDAIVGYF DVRFDDADC
KVEFSTSPTT PLTHWKQTLL FLDKPIRVKP GDKISGIITI RRATTDNRGL EINLLIGETE
NSLEIKQTFD LIG

SmPRMT3 (210-553 / C5) MW 39045,41 Da pI 5,69 Eps(red) 34380 L.mol⁻¹.cm⁻¹

MTDSNPMEYD CDSDDWQEV T LPDLSLSEKY TCLMCENHFN TAPLFFQHLT LQHDWVNFLD
GEGKSIFNDQ YDWIRFVNFV RKTTPVDLRQ AFYDMVWKT D SLDLLYPVLP DDEVLSIDIE
SYLLDNQRSN QSPGIVVHEK STNQOPVSQE MSRIMAENKS LRTQLNNCKR LLSTMSAQSV
RRKSPCDQSS NYDLDMEPND SSYFGSYGHF EIHGEMINDR VRTESYVNF I LSNAEKYFKH
KIILDVGS GS GILSIIAAQA GASHVYGVEA ADEIYAASHE TLRVNNLLER VTFIHGQAES
VELPVKKVDV IISEWMGYFL FFESMLDSVL KMASKYLSRD GHIFPRHYTL NLLGVQCSEQ
LRKRRLEHWN NVYGYNMPAL RRAALSEAHV LNLTNEHVTP PISPITILTQ SFELVALDLD
DMHRNRIYNL SNHCSSLCEQ KFHLTIQPTT DINNNSSSSS YELDAIVGYF DVRFDDADC
KVEFSTSPTT PLTHWKQTLL FLDKPIRVKP GDKISGIITI RRATTDNRGL EINLLIGETE
NSLEIKQTFD LIG

Annexe 2 - Séquences des constructions de SmPRMT3 C3 à C5.

La séquence correspondant aux constructions est représentée en noir ou en couleur. Les domaines amino-terminal, de Rossmann, le tonneau β et le bras de dimérisation sont colorés respectivement en vert, orange, bleu clair et violet. Les motifs I, II III et IV sont colorés respectivement en bleu foncé, rose, rouge et vert. MW : poids moléculaire. pI : point isoélectrique. Eps(red) : coefficient d'extinction moléculaire en considérant que les cystéines sont réduites.

PDB ID / Cristal		SP045C4n	
échantillon	Protéine	SmPRMT3 ₁₈₄₋₅₅₃	
	Cofacteur/Inhibiteur	SFG 2 mM	
	Substrat	-	
Cristallisation	Réservoir	100 mM BTP pH 9 10 % PEG 5000 MME 100 mM NaCl	
	Température (°C)	20	
Collecte	Source	ESRF ID30b	
	Détecteur	EIGER	
	Longueur d'onde (Å)	0.9677	
	Distance (mm)	308.7	
	#Images	3600	
	Oscillations (°)	0.1	
	Temps d'exposition (s)	0.2	
Traitement	Groupe d'espace	H32	
	Paramètres de maille	$a - b - c$ (Å) $\alpha - \beta - \gamma$ (°)	159.1 159.1 617.1 90 90 120
	Résolution (Å)	49.33 - 7.29 (8.15 - 7.29)	
	R_{merge} / R_{sym}^*	0.142 (1.366)	
	I/σ_I	15.1 (2.7)	
	$CC_{1/2}$	0.989 (0.307)	
	Complétude (%)	99.6 (99.9)	
	Redondance	20.1 (20.4)	
	Affinement	Resolution (Å)	-
		#Réflexions	-
R_{work} (%)		-	
R_{free} (%)		-	
#Atomes		Protéine Ligand Solvant	- - -
Wilson B (Å ²)		-	
Validation		RMSD	Bond lengths (Å) Bond angles (°)
	Ramachandran favored (%)	-	
	Ramachandran outliers (%)	-	
	Rotamer Outliers (%)	-	
	Cbeta Deviations	-	
	Average B (Å ²)	-	

Annexe 3 - Données cristallographiques pour SmPRMT3.

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées sur la couche de plus haute résolution.

PDB ID / Cristal		5IH3	Non déposée	
échantillon	<i>Protéine</i>	mCARM1 ₁₃₀₋₄₈₇	mCARM1 ₁₃₀₋₄₈₇ P409A	
	<i>Cofacteur/Inhibiteur</i>	SAH 2 mM	SAH 2 mM	
	<i>Substrat</i>	Guanidine Me1 2 mM	-	
Cristallisation	<i>Réservoir</i>	100 mM Tris-HCl pH 8 19% PEG 2000 MME 167 mM NaCl	100 mM Tris-HCl pH 8.5 22% PEG 3350 200 mM A.S.	
	<i>Température (°C)</i>	24	20	
Collecte	<i>Source</i>	ESRF ID14-4	ESRF ID30b	
	<i>Détecteur</i>	CCD ADSC Q315r	PILATUS 6M	
	<i>Longueur d'onde (Å)</i>	0.9304	0.9919	
	<i>Distance (mm)</i>	240.5	384.5	
	<i>#Images</i>	900	2400	
	<i>Oscillations (°)</i>	0.2	0.05	
	<i>Temps d'exposition (s)</i>	0.1	0.2	
	Traitement	<i>Groupe d'espace</i>	P2 ₁ -2 ₁ -2	P2 ₁ -2 ₁ -2
<i>Paramètres de maille</i>		<i>a - b - c (Å)</i>	74.6 - 97.9 - 205.0	74.8 98.3 207.2
		<i>α - β - γ (°)</i>	90 - 90 - 90	90 90 90
<i>Résolution (Å)</i>		47.62 - 1.77 (1.80 - 1.77)	47.81 - 2.04 (2.11 - 2.04)	
<i>R_{merge} / R_{sym} *</i>		0.129 (1.40)	0.213 (1.09)	
<i>I/σ₁</i>		11.2 (1.5)	5.41 (1.07)	
<i>CC_{1/2}</i>		0.998 (0.407)		
<i>Complétude (%)</i>		99.9 (98.6)	97.40 (98.61)	
<i>Redondance</i>		7.3 (6.8)	4.5 (4.4)	
Affinement		<i>Resolution (Å)</i>	47.62 - 1.77 (1.79 - 1.77)	47.81 - 2.04 (2.11 - 2.04)
	<i>#Réflexions</i>	146548	96055	
	<i>R_{work} (%)</i>	16.40 (34.82)	0.20 (0.31)	
	<i>R_{free} (%)</i>	19.05 (38.65)	0.23 (0.33)	
	<i>#Atomes</i>	<i>Protéine</i>	10989	11053
		<i>Ligand</i>	249	226
		<i>Solvant</i>	1051	1264
	<i>Wilson B (Å²)</i>	20.79	20.05	
Validation	<i>RMSD</i>	<i>Bond lengths (Å)</i>	0.009	0.002
		<i>Bond angles (°)</i>	1.158	0.52
	<i>Ramachandran favored (%)</i>	97.08	96.05	
	<i>Ramachandran outliers (%)</i>	0.00	0.00	
	<i>Rotamer Outliers (%)</i>	0.66	0.83	
	<i>Cbeta Deviations</i>	0	0	
	<i>Average B (Å²)</i>	31.00	27.36	

Annexe 4 - Données cristallographiques pour les structures de mCARM1 WT et P409A.

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées pour la couche de plus haute résolution. La structure 5IH3 n'a pas été résolue lors de cette thèse, mais est présentée en tant que structure de référence.

PDB ID / Cristal		5FUB	5FUL	
échantillon	<i>Protéine</i>	zPRMT2 ₇₃₋₄₀₇	mPRMT2 ₁₋₄₄₅	
	<i>Cofacteur/Inhibiteur</i>	SAH 1 mM	SAH 0.5 mM	
	<i>Substrat</i>	-	-	
Cristallisation	<i>Réservoir</i>	100 mM MES pH 6.0 9% PEG 20000 300 mM NaCl	100 mM HEPES pH 7.0 16% PEG 6000 100 mM CaCl2	
	<i>Température (°C)</i>	20	20	
Collecte	<i>Source</i>	SOLEIL Proxima 2	ESRF ID23-2	
	<i>Détecteur</i>	ADSC Q315	PILATUS 2M	
	<i>Longueur d'onde (Å)</i>	0.9801	0.8726	
	<i>Distance (mm)</i>	291.6	241.5	
	<i>#Images</i>	180	2500	
	<i>Oscillations (°)</i>	0.5	0.1	
	<i>Temps d'exposition (s)</i>	-	0.1	
	Traitement	<i>Groupe d'espace</i>	H 32	C 2 2 21
<i>Paramètres de maille</i>		<i>a - b - c (Å)</i>	147.1 147.1 127.4	65.5 115.4 133.0
		<i>α - β - γ (°)</i>	90 90 120	90 90 90
<i>Résolution (Å)</i>		56.97 - 2.00 (2.05 - 2.00)	44.32 - 1.89 (1.93 - 1.98)	
<i>R_{merge} / R_{sym} *</i>		0.102 (1.710)	0.091 (1.901)	
<i>I/σ_i</i>		9.2 (1.0)	15.4 (1.2)	
<i>CC_{1/2}</i>		0.998 (0.308)	0.999 (0.448)	
<i>Complétude (%)</i>		99.9 (100)	99.8 (97.3)	
<i>Redondance</i>		5.3 (5.3)	9.2 (8.5)	
Affinement	<i>Resolution (Å)</i>	38.41 - 2.00 (2.05 - 2.00)	43.26 - 1.89 (1.96 - 1.89)	
	<i>#Réflexions</i>	35917	40528	
	<i>R_{work} (%)</i>	18.17 (31.36)	16.57 (29.55)	
	<i>R_{free} (%)</i>	21.75 (31.44)	19.24 (29.64)	
	<i>#Atomes</i>	<i>Protéine</i>	2696	2722
		<i>Ligand</i>	57	50
		<i>Solvant</i>	181	309
<i>Wilson B (Å²)</i>	35.4	30.1		
Validation	<i>RMSD</i>	<i>Bond lengths (Å)</i>	0.004	0.019
		<i>Bond angles (°)</i>	0.603	1.289
	<i>Ramachandran favored (%)</i>	97.61	97.03	
	<i>Ramachandran outliers (%)</i>	0.00	0.00	
	<i>Rotamer Outliers (%)</i>	0.00	0.33	
	<i>Cbeta Deviations</i>	0	0	
	<i>Average B (Å²)</i>	46.00	37.00	

Annexe 5 - Données cristallographiques pour les structures de mPRMT2 et zPRMT2.

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées pour la couche de plus haute résolution.

PDB ID / Cristal		5LGP	5LGQ	5LGR	
échantillon	<i>Protéine</i>	mCARM1 ₁₃₀₋₄₈₇	mCARM1 ₁₃₀₋₄₈₇	mCARM1 ₁₃₀₋₄₈₇	
	<i>Cofacteur/Inhibiteur</i>	-	-	-	
	<i>Substrat</i>	P1C3s 2 mM	P2C3s 2 mM	P1C3u 2 mM	
Cristallisation	<i>Réservoir</i>	100 mM Tris-HCl pH 8.5 16% PEG 3350 200 mM A.S.	100 mM Tris-HCl pH 8.5 14% PEG 3350 200 mM A.S.	100 mM Tris-HCl pH 8.5 14% PEG 3350 200 mM A.S.	
	<i>Température (°C)</i>	20	20	20	
Collecte	<i>Source</i>	SOLEIL Proxima1	ESRF ID 29	SOLEIL Proxima1	
	<i>Détecteur</i>	PILATUS 6M	CCD ADSC Q315r	PILATUS 6M	
	<i>Longueur d'onde (Å)</i>	0.97857	1.073	0.97857	
	<i>Distance (mm)</i>	369.4	403.2	393.4	
	<i>#Images</i>	900	1800	900	
	<i>Oscillations (°)</i>	0.2	0.1	0.2	
	<i>Temps d'exposition (s)</i>	0.2	0.02	0.2	
Traitement	<i>Groupe d'espace</i>	P2 ₁ -2 ₁ -2	P2 ₁ -2 ₁ -2	P2 ₁ -2 ₁ -2	
	<i>Paramètres de maille</i>	<i>a - b - c (Å)</i> 74.8 98.7 207.3	74.7 99.0 207.3	74.8 98.5 208.3	
		<i>α - β - γ (°)</i> 90 90 90	90 90 90	90 90 90	
	<i>Résolution (Å)</i>	48.01 - 2.04 (2.07 - 2.04)	48.15 - 2.11 (2.15 - 2.11)	47.95 - 2.00 (2.03 - 2.00)	
	<i>R_{merge} / R_{sym} *</i>	0.15 (2.33)	0.15 (2.02)	0.17 (3.23)	
	<i>I/σ_i</i>	9.7 (0.8)	9.0 (0.7)	8.0 (0.6)	
	<i>CC_{1/2}</i>	0.998 (0.313)	0.998 (0.325)	0.998 (0.308)	
	<i>Complétude (%)</i>	100 (100)	99.9 (99.6)	100 (99.9)	
	<i>Redondance</i>	6.6 (6.4)	6.5 (4.7)	9.7 (6.6)	
Affinement	<i>Resolution (Å)</i>	48.03 - 2.04 (2.06 - 2.04)	48.15 - 2.11 (2.13 - 2.11)	46.05 - 2.00 (2.02 - 2.00)	
	<i>#Réflexions</i>	98404	89162	104667	
	<i>R_{work} (%)</i>	18.73 (33.18)	18.66 (35.59)	19.20 (36.23)	
	<i>R_{free} (%)</i>	21.93 (33.82)	23.60 (34.02)	23.37 (38.86)	
		<i>Protéine</i>	11443	11346	11457
		<i>#Atomes Ligand</i>	26	85	142
		<i>Solvant</i>	755	839	765
	<i>Wilson B (Å²)</i>	31.63	32.74	29.67	
Validation	<i>RMSD Bond lengths (Å)</i>	0.006	0.002	0.002	
		<i>Bond angles (°)</i>	0.691	0.530	0.540
		<i>Ramachandran favored (%)</i>	97.08	96.86	96.87
		<i>Ramachandran outliers (%)</i>	0.00	0.07	0.00
		<i>Rotamer Outliers (%)</i>	0.25	0.33	0.33
		<i>Cbeta Deviations</i>	0	0	0
		<i>Average B (Å²)</i>	37.59	41.67	38.65

Annexe 6 - Données cristallographiques pour les complexes mCARM1-dérivés peptidiques (1).

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées pour la couche de plus haute résolution.

PDB ID / Cristal		5LGS	X121B2v-3	X126B5v	
échantillon	<i>Protéine</i>	mCARM1 ₁₃₀₋₄₈₇	mCARM1 ₁₃₀₋₅₀₇	mCARM1 ₁₃₀₋₄₉₇	
	<i>Cofacteur/Inhibiteur</i>	-	-	-	
	<i>Substrat</i>	P2C3u 2 mM	H3 (13-31) K18Ac 0.5 mM	H3 (13-31) 2 mM	
Cristallisation	<i>Réservoir</i>	100 mM Tris-HCl pH 8.5 20% PEG 3350 200 mM A.S.	10 mM TrisHCL pH 8.5 12 % PEG 3350 200 mM A.S.	100 mM Tris pH 8.5 16 % PEG3350 300 mM A.S.	
	<i>Température (°C)</i>	20	20	20	
Collecte	<i>Source</i>	SOLEIL Proxima1	SOLEIL Proxima1	source IGBMC	
	<i>Détecteur</i>	PILATUS 6M	PILATUS 6M	EIGER	
	<i>Longueur d'onde (Å)</i>	0.97857	0.9786	1.5418	
	<i>Distance (mm)</i>	369.4	369.5	117	
	<i>#Images</i>	900	1200	3860	
	<i>Oscillations (°)</i>	0.2	0.1	0.13	
	<i>Temps d'exposition (s)</i>	0.2	0.1	30	
	Traitement	<i>Groupe d'espace</i>	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2
<i>Paramètres de maille</i>		<i>a - b - c (Å)</i>	75.0 98.6 208.2	75.1 98.5 207.7	75.1 98.3 208.3
		<i>α - β - γ (°)</i>	90 90 90	90 90 90	90 90 90.0
<i>Résolution (Å)</i>		47.97 - 2.10 (2.14 - 2.10)	47.93 - 2.01 (2.04 - 2.01)	47.84 - 2.12 (2.15 - 2.12)	
<i>R_{merge} / R_{sym} *</i>		0.13 (2.75)	0.13 (6.29)	0.137 (0.986)	
<i>I/σ_i</i>		10.4 (0.7)	5.4 (0.1)	7.1 (0.9)	
<i>CC_{1/2}</i>		0.999 (0.320)	0.998 (0.096)	0.996 (0.456)	
<i>Complétude (%)</i>		100 (99.9)	98.9 (93.8)	99.3 (91.0)	
<i>Redondance</i>		6.7 (6.9)	3.6 (3.5)	7.2 (3.5)	
Affinement		<i>Resolution (Å)</i>	47.97 - 2.10 (2.12 - 2.10)	47.93 - 2.38 (2.42 - 2.38)	47.84 - 2.12 (2.14 - 2.12)
	<i>#Réflexions</i>	90760	61925	87901	
	<i>R_{work} (%)</i>	18.63 (33.86)	19.91 (34.14)	21.80 (32.02)	
	<i>R_{free} (%)</i>	21.47 (37.70)	24.27 (35.33)	26.01 (34.73)	
	<i>#Atomes</i>	<i>Protéine</i>	11266	11321	11305
		<i>Ligand</i>	135	441	429
		<i>Solvant</i>	584	402	715
	<i>Wilson B (Å²)</i>	37.00	47.38	31.89	
Validation	<i>RMSD</i>	<i>Bond lengths (Å)</i>	0.002	0.003	0.585
		<i>Bond angles (°)</i>	0.516	0.530	0,05
	<i>Ramachandran favored (%)</i>	96.85	96.13	95,43	
	<i>Ramachandran outliers (%)</i>	0.00	0.00	0.070	
	<i>Rotamer Outliers (%)</i>	0.49	0.80	1.43	
	<i>Cbeta Deviations</i>	0	0	0	
	<i>Average B (Å²)</i>	51.00	57.00	46.20	

Annexe 7 - Données cristallographiques pour les complexes mCARM1-dérivés peptidiques (2).

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées pour la couche de plus haute résolution. Le cristal X121B2v-3 (non déposé) a été affiné à 2,38 Å bien que la résolution limite soit de 2,01 Å après traitement, car une chute des valeurs de R_{free} et R_{work} a été observée au-delà de 2,38 Å.

PDB ID / Cristal		X130C1n	X131A3b	X131D2r
échantillon	<i>Protéine</i>	mCARM1 ₁₃₀₋₄₉₇	mCARM1 ₁₃₀₋₄₉₇	mCARM1 ₁₃₀₋₄₉₇
	<i>Cofacteur/Inhibiteur</i>	-	-	-
	<i>Substrat</i>	H3 (13-31) K18Ac 2 mM	H3 (13-31) 2 mM	H3 (1-41) K18Ac 1 mM
Cristallisation	<i>Réservoir</i>	100 mM MES pH 7 1M Na malonate 200 mM NaCl	100 mM MES pH 5.5 1.2 M Na malonate 200 mM NaCl	100 mM Tris pH 7.5 1.1 M Na malonate 200 mM NaCl
	<i>Température (°C)</i>	20	20	20
Collecte	<i>Source</i>	source IGBMC	ESRF ID30b	ESRF ID30b
	<i>Détecteur</i>	EIGER	EIGER	EIGER
	<i>Longueur d'onde (Å)</i>	1.54178	0.9677	0.9677
	<i>Distance (mm)</i>	117.0	135.9	161.7
	<i>#Images</i>	2298	3600	3600
	<i>Oscillations (°)</i>	0.1	0.05	0.05
	<i>Temps d'exposition (s)</i>	15	0.2	0.2
	<i>Groupe d'espace</i>	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2
Traitement	<i>Paramètres de maille</i>	<i>a - b - c (Å)</i> 74.9 99.6 208.1	74.4 98.6 206.6	75.1 100.4 208.0
		<i>α - β - γ (°)</i> 90 90 90	90 90 90	90 90 90
	<i>Résolution (Å)</i>	46.11 - 2.0 (2.04 - 2.0)	45.76 - 2.54 (2.62 - 2.54)	46.17 - 2.40 (2.46 - 2.40)
	<i>R_{merge}/R_{sym} *</i>	0.102 (1.21)	0.129 (2.284)	0.166 (2.148)
	<i>I/σ_i</i>	7.4 (0.7)	8.6 (0.6)	5.9 (0.6)
	<i>CC_{1/2}</i>	0.998 (0.355)	0.999 (0.475)	0.995 (0.350)
	<i>Complétude (%)</i>	95.0 (73.6)	97.9 (87.4)	99.5 (97.2)
	<i>Redondance</i>	4.9 (3.3)	6.3 (5.5)	5.8 (5.4)
Affinement	<i>Resolution (Å)</i>	45.32 - 2.20 (2.23 - 2.20)	45.78 - 2.54 (2.59 - 2.54)	46.17 - 2.40 (2.44 - 2.40)
	<i>#Réflexions</i>	76938	49687	61909
	<i>R_{work} (%)</i>	20.93 (33.30)	21.51 (38.96)	21.76 (34.33)
	<i>R_{free} (%)</i>	25.68 (38.20)	25.80 (40.41)	25.84 (40.88)
	<i>Protéine</i>	11366	11281	11279
	<i>#Atomes</i>	<i>Ligand</i> 442	434	352
		<i>Solvant</i> 636	197	316
<i>Wilson B (Å²)</i>	34.36	54.98	47.82	
Validation	<i>RMSD</i>	<i>Bond lengths (Å)</i> 0,587	0.003	0.002
		<i>Bond angles (°)</i> 0,045	0.519	0.570
	<i>Ramachandran favored (%)</i>	95,85	95.57	95.58
	<i>Ramachandran outliers (%)</i>	0.00	0.07	0.07
	<i>Rotamer Outliers (%)</i>	0,95	1.43	0.81
	<i>Cbeta Deviations</i>	0	0	0
	<i>Average B (Å²)</i>	48.00	67.60	58.30

Annexe 8 - Données cristallographiques pour les complexes mCARM1-dérivés peptidiques (3).

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées pour la couche de plus haute résolution. La structure correspondant au cristal X130C1n (non déposé) a été affinée à 2,20 Å bien que la résolution limite soit de 2,00 Å après traitement, car une chute des valeurs de R_{free} et R_{work} a été observée au-delà de 2,20 Å.

PDB ID / Cristal		5TBH	5LV2	5TBI
échantillon	<i>Protéine</i>	mCARM1 ₁₃₀₋₄₈₇	mCARM1 ₁₃₀₋₄₈₇	mCARM1 ₁₃₀₋₄₈₇
	<i>Cofacteur/Inhibiteur</i>	LH1236 2 mM	LH1246 2 mM	LH1427 2 mM
	<i>Substrat</i>	-	-	-
Cristallisation	<i>Réservoir</i>	100 mM Tris-HCl pH 8 19% PEG 2000 MME 100 mM NaCl	100 mM Tris-HCl pH 8 17% PEG 2000 MME 100 mM NaCl	100 mM Tris-HCl pH 8 21% PEG 2000 MME 100 mM NaCl
	<i>Température (°C)</i>	20	20	20
Collecte	<i>Source</i>	SOLEIL Proxima1	SOLEIL Proxima1	SOLEIL Proxima1
	<i>Détecteur</i>	PILATUS 6M	PILATUS 6M	PILATUS 6M
	<i>Longueur d'onde (Å)</i>	0.97857	0.97857	0.97857
	<i>Distance (mm)</i>	440.5	440.5	463.8
	<i>#Images</i>	900	900	600
	<i>Oscillations (°)</i>	0.2	0.2	0.2
	<i>Temps d'exposition (s)</i>	0.2	0.2	0.2
Traitement	<i>Groupe d'espace</i>	P2 ₁ -2 ₁ 2	P2 ₁	P2 ₁ -2 ₁ 2
	<i>Paramètres de maille</i> <i>a - b - c (Å)</i> <i>α - β - γ (°)</i>	74.6 98.5 206.4 90 90 90	98.7 74.8 206.6 90 90.53 90	74.5 98.0 205.7 90 90 90
	<i>Résolution (Å)</i>	49.26 - 2.34 (2.40 - 2.34)	49.34 - 2.29 (2.33 - 2.29)	49.01 - 2.29 (2.35 - 2.29)
	<i>R_{merge} / R_{sym} *</i>	0.28 (1.84)	0.16 (1.16)	0.16 (1.15)
	<i>I/σ_i</i>	7.4 (1.8)	5.5 (1.0)	6.9 (1.2)
	<i>CC_{1/2}</i>	0.979 (0.328)	0.985 (0.412)	0.991 (0.495)
	<i>Complétude (%)</i>	99.6 (95.3)	99.1 (90.9)	99.3 (92.3)
	<i>Redondance</i>	6.3 (5.9)	3.4 (3.2)	4.4 (4.1)
Affinement	<i>Resolution (Å)</i>	49.26 - 2.34 (2.37 - 2.34)	49.34 - 2.29 (2.32 - 2.29)	45.54 - 2.29 (2.32 - 2.29)
	<i>#Réflexions</i>	64692	134021	67860
	<i>R_{work} (%)</i>	19.36 (31.06)	20.51 (34.26)	19.77 (33.57)
	<i>R_{free} (%)</i>	22.98 (34.29)	23.80 (34.33)	22.84 (34.26)
	<i>#Atomes</i> <i>Protéine</i>	10989	21990	10989
	<i>Ligand</i>	218	365	209
	<i>Solvant</i>	406	1166	600
<i>Wilson B (Å²)</i>	30.84	31.60	31.25	
Validation	<i>RMSD</i> <i>Bond lengths (Å)</i>	0.006	0.005	0.004
	<i>Bond angles (°)</i>	1.024	0.912	0.855
	<i>Ramachandran favored (%)</i>	96.99	96.96	96.92
	<i>Ramachandran outliers (%)</i>	0.00	0.04	0.00
	<i>Rotamer Outliers (%)</i>	0.42	0.50	0.17
	<i>Cbeta Deviations</i>	0	0	0
	<i>Average B (Å²)</i>	38.00	42.00	40.00

Annexe 9 - Données cristallographiques pour les complexes mCARM1-LH (1).

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées pour la couche de plus haute résolution.

PDB ID / Cristal		5TBJ	5LV3	
échantillon	<i>Protéine</i>	mCARM1 ₁₃₀₋₄₈₇	mCARM1 ₁₃₀₋₄₈₇	
	<i>Cofacteur/Inhibiteur</i>	LH1452 2 mM	LH1561Br 2 mM	
	<i>Substrat</i>	-	-	
Cristallisation	<i>Réservoir</i>	100 mM Tris-HCl pH 8 19% PEG 2000 MME 100 mM NaCl	100 mM Tris-HCl pH 8 15% PEG 2000 MME 100mM NaCl	
	<i>Température (°C)</i>	20	20	
Collecte	<i>Source</i>	SOLEIL Proxima1	ESRF ID23-2	
	<i>Détecteur</i>	PILATUS 6M	PILATUS 2M-F	
	<i>Longueur d'onde (Å)</i>	0.97857	0.8726	
	<i>Distance (mm)</i>	440.5	211.9	
	<i>#Images</i>	600	4250	
	<i>Oscillations (°)</i>	0.2	0.05	
	<i>Temps d'exposition (s)</i>	0.2	0.04	
Traitement	<i>Groupe d'espace</i>	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	
	<i>Paramètres de maille</i>	<i>a - b - c (Å)</i>	74.8 99.2 207.3	75.1 98.5 208.1
		<i>α - β - γ (°)</i>	90 90 90	90 90 90
	<i>Résolution (Å)</i>	19.97 - 2.32 (2.38 - 2.32)	30.0 - 1.80 (1.83 - 1.80)	
	<i>R_{merge} / R_{sym} *</i>	0.23 (1.80)	0.16 (1.87)	
	<i>I/σ_i</i>	4.6 (0.9)	8.7 (1.1)	
	<i>CC_{1/2}</i>	0.972 (0.317)	0.997 (0.302)	
	<i>Complétude (%)</i>	98.6 (87.5)	100 (100)	
<i>Redondance</i>	4.4 (4.2)	7.9 (7.9)		
Affinement	<i>Resolution (Å)</i>	19.97 - 2.32 (2.35 - 2.32)	29.95 - 1.80 (1.82 - 1.80)	
	<i>#Réflexions</i>	66234	143435	
	<i>R_{work} (%)</i>	20.22 (35.39)	17.35 (31.53)	
	<i>R_{free} (%)</i>	23.73 (35.62)	20.30 (34.06)	
	<i>#Atomes</i>	<i>Protéine</i>	11000	12241
		<i>Ligand</i>	224	178
		<i>Solvant</i>	754	1126
<i>Wilson B (Å²)</i>	30.26	21.80		
Validation	<i>RMSD</i>	<i>Bond lengths (Å)</i>	0.004	0.005
		<i>Bond angles (°)</i>	0.904	0.814
	<i>Ramachandran favored (%)</i>	97.36	96.6	
	<i>Ramachandran outliers (%)</i>	0.15	0.00	
	<i>Rotamer Outliers (%)</i>	0.42	0.84	
	<i>Cbeta Deviations</i>	0	0	
	<i>Average B (Å²)</i>	39.00	30.85	

Annexe 10 - Données cristallographiques pour les complexes mCARM1-LH (2).

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées pour la couche de plus haute résolution.

PDB ID / Cristal		5LV4	5LV5
échantillon	<i>Protéine</i>	mPRMT6 ₁₋₃₇₈ C53S	mPRMT6 ₁₋₃₇₈
	<i>Cofacteur/Inhibiteur</i>	LH1236 1 mM	LH1458 1.125 mM
	<i>Substrat</i>	-	-
Cristallisation	<i>Réservoir</i>	100 mM Hepes pH 8.0 200 mM MgCl ₂ 10% PEG 6000	100 mM Tris pH 8.0 200 mM MgCl ₂ 10% PEG 6000
	<i>Température (°C)</i>	20	20
Collecte	<i>Source</i>	SOLEIL PROXIMA1	SOLEIL PROXIMA2
	<i>Détecteur</i>	PILATUS 6M	Q315
	<i>Longueur d'onde (Å)</i>	0.97779	0.9801
	<i>Distance (mm)</i>	321.8	237.88
	<i>#Images</i>	1000	500
	<i>Oscillations (°)</i>	0.2	0.25
	<i>Temps d'exposition (s)</i>	0.1	0.3 s
	<i>Groupe d'espace</i>	I4 ₁	I4 ₁
Traitement	<i>Paramètres de maille</i>	<i>a - b - c (Å)</i> 79.9 79.9 119.7 <i>α - β - γ (°)</i> 90 90 90	77.7 77.7 117.7 90 90 90
	<i>Résolution (Å)</i>	41.08 - 1.66 (1.69 - 1.65)	40.16 - 1.80 (1.84 - 1.80)
	<i>R_{merge} / R_{sym} *</i>	0.142 (1.508)	0.105 (1.261)
	<i>I/σ_i</i>	7.7 (1.0)	10.7 (1.2)
	<i>CC_{1/2}</i>	0.996 (0.474)	0.998 (0.421)
	<i>Complétude (%)</i>	99.7 (94.5)	99.8 (97.3)
	<i>Redondance</i>	7.2 (5.5)	5.2 (4.7)
	<i>Resolution (Å)</i>	41.1 - 1.65 (1.69 - 1.65)	40.2 - 1.8 (1.84 - 1.80)
Affinement	<i>#Réflexions</i>	44207	32113
	<i>R_{work} (%)</i>	15.90 (18.80)	22.00 (24.60)
	<i>R_{free} (%)</i>	29.30 (31.00)	35.00 (40.60)
	<i>#Atomes</i>		
	<i>Protéine</i>	2504	2533
	<i>Ligand</i>	38	39
	<i>Solvant</i>	331	221
<i>Wilson B (Å²)</i>	21.70	22.9	
Validation	<i>RMSD</i>		
	<i>Bond lengths (Å)</i>	0.009	0.010
	<i>Bond angles (°)</i>	1.079	1.060
	<i>Ramachandran favored (%)</i>	99.05	98.73
	<i>Ramachandran outliers (%)</i>	0.00	0.232
	<i>Rotamer Outliers (%)</i>	0.37	0.00
	<i>Cbeta Deviations</i>	0	0
	<i>Average B (Å²)</i>	30.85	30.63

Annexe 11 - Données cristallographiques pour les complexes mPRMT6-LH.

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées pour la couche de plus haute résolution.

PDB ID / Cristal		5K8V	5K8W	5K8X
échantillon	<i>Protéine</i>	mCARM1 ₁₃₀₋₄₈₇	mCARM1 ₁₃₀₋₄₈₇	mCARM1 ₁₃₀₋₄₈₇
	<i>Cofacteur/Inhibiteur</i>	U1 1 mM	U2 1 mM	U3 1 mM
	<i>Substrat</i>	-	-	-
Cristallisation	<i>Réservoir</i>	100 mM Tris-HCl pH 8 15% PEG 2000 MME 100 mM NaCl	100 mM Tris-HCl pH 8 15% PEG 2000 MME 100 mM NaCl	100 mM Tris-HCl pH 8 15% PEG 2000 MME 100 mM NaCl
	<i>Température (°C)</i>	20	20	20
Collecte	<i>Source</i>	SOLEIL Proxima2	SOLEIL Proxima2	ESRF ID 23-1
	<i>Détecteur</i>	CCD ADSC Q315r	CCD ADSC Q315r	PILATUS 6M
	<i>Longueur d'onde (Å)</i>	0.9801	0.9801	0.9795
	<i>Distance (mm)</i>	329.0	291.7	392.8
	<i>#Images</i>	820	618	1800
	<i>Oscillations (°)</i>	0.2	0.2	0.1
	<i>Temps d'exposition (s)</i>	0.3	0.3	0.1
Traitement	<i>Groupe d'espace</i>	P2 ₁ -2 ₁ -2	P2 ₁ -2 ₁ -2	P2 ₁ -2 ₁ -2
	<i>Paramètres de maille</i>	<i>a - b - c (Å)</i> 75.2 98.1 206.3	74.9 98.8 207.5	74.8 98.5 205.5
	<i>α - β - γ (°)</i>	90 90 90	90 90 90	90 90 90
	<i>Résolution (Å)</i>	47.72 - 2.25 (2.30 - 2.25)	48.06 - 2.10 (2.14 - 2.10)	49.24 - 1.99 (2.02 - 1.99)
	<i>R_{merge} / R_{sym} *</i>	0.28 (2.13)	0.31 (1.98)	0.16 (2.04)
	<i>I/σ_i</i>	5.7 (1.1)	5.8 (1.6)	7.3 (0.8)
	<i>CC_{1/2}</i>	0.986 (0.349)	0.983 (0.333)	0.996 (0.325)
	<i>Complétude (%)</i>	99.8 (97.6)	99.8 (99.9)	99.2 (84.8)
<i>Redondance</i>	6.5 (6.6)	4.9 (5.1)	6.6 (5.7)	
Affinement	<i>Resolution (Å)</i>	47.72 - 2.25 (2.28 - 2.25)	45.93 - 2.10 (2.12 - 2.10)	45.55 - 1.99 (2.01 - 1.99)
	<i>#Réflexions</i>	72983	90444	103907
	<i>R_{work} (%)</i>	19.24 (30.02)	18.26 (27.65)	16.99 (35.47)
	<i>R_{free} (%)</i>	23.48 (32.20)	24.02 (33.34)	21.15 (37.17)
	<i>Protéine</i>	10989	11019	10993
	<i>#Atomes Ligand</i>	170	195	270
	<i>Solvant</i>	596	1198	834
<i>Wilson B (Å²)</i>	29.61	22.95	29.19	
Validation	<i>RMSD Bond lengths (Å)</i>	0.002	0.004	0.008
	<i>Bond angles (°)</i>	0.477	0.586	1.002
	<i>Ramachandran favored (%)</i>	96.55	96.88	97.17
	<i>Ramachandran outliers (%)</i>	0.00	0.00	0.00
	<i>Rotamer Outliers (%)</i>	0.50	0.83	0.41
	<i>Cbeta Deviations</i>	0	0	0
	<i>Average B (Å²)</i>	41.00	30.00	39.00

Annexe 12 - Données cristallographiques pour les complexes mCARM1-U.

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées pour la couche de plus haute résolution.

PDB ID / Cristal		5FWD	5JMQ	
échantillon	Protéine	mPRMT2 ₁₋₄₄₅	mPRMT2 ₁₋₄₄₅	
	Cofacteur/Inhibiteur	U2 1 mM	U3 1 mM	
	Substrat	-	-	
Cristallisation	Réservoir	100 mM HEPES pH 7.0 20% PEG 6000 100 mM CaCl2	100 mM HEPES pH 7.0 22% PEG 6000 100 mM CaCl2	
	Température (°C)	20	20	
Collecte	Source	ESRF ID23-1	ESRF ID23-1	
	Détecteur	PILATUS 6M	PILATUS 6M	
	Longueur d'onde (Å)	0.9795	0.9795	
	Distance (mm)	383.5	360.2	
	#Images	1100	1800	
	Oscillations (°)	0.15	0.1	
	Temps d'exposition (s)	0.1	0.1	
	Traitement	Groupe d'espace	C 2 2 21	C 2 2 21
Paramètres de maille		<i>a</i> - <i>b</i> - <i>c</i> (Å)	66.0 114.5 132.1	65.7 114.7 132.2
		α - β - γ (°)	90 90 90	90 90 90
Résolution (Å)		44.03 - 2.00 (2.05 - 2.00)	44.05 - 1.79 (1.83 - 1.79)	
R_{merge} / R_{sym}^*		0.161 (1.444)	0.113 (1.957)	
I/σ_I		6.9 (1.2)	9.8 (0.7)	
$CC_{1/2}$		0.995 (0.619)	0.988 (0.217)	
Complétude (%)		99.9 (100)	98.7 (84.9)	
Redondance		5.9 (6.1)	6.5 (5.1)	
Affinement		Resolution (Å)	43.26 - 2.00 (2.06 - 2.00)	43.31 - 1.80 (1.83 - 1.79)
	#Réflexions	34075	46333	
	R_{work} (%)	18.82 (28.40)	17.92 (36.16)	
	R_{free} (%)	20.94 (28.86)	20.14 (39.16)	
	#Atomes	Protéine	2723	2723
		Ligand	64	55
		Solvant	268	321
Wilson B (Å ²)	30.8	28.7		
Validation	RMSD	Bond lengths (Å)	0.007	0.006
		Bond angles (°)	0.699	0.878
	Ramachandran favored (%)	97.92	98.22	
	Ramachandran outliers (%)	0.00	0.00	
	Rotamer Outliers (%)	0.66	0.33	
	Cbeta Deviations	0	0	
	Average B (Å ²)	35.00	33.00	

Annexe 13 - Données cristallographiques pour les complexes mPRMT2-U.

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées pour la couche de plus haute résolution.

PDB ID / Cristal		L46D2	L46A3	
échantillon	<i>Protéine</i>	mPRMT6 ₃₄₋₃₇₈	mPRMT6 ₃₄₋₃₇₈	
	<i>Cofacteur/Inhibiteur</i>	U1 2 mM	U2 2 mM	
	<i>Substrat</i>	-	-	
Cristallisation	<i>Réservoir</i>	100 mM Tris pH 7.5 20% PEG 3,350 200 mM NaI	100 mM Tris pH 7.5 200 mM NaI 20% PEG 3,350	
	<i>Température (°C)</i>	20	20	
Collecte	<i>Source</i>	source IGBMC	source IGBMC	
	<i>Détecteur</i>	PILATUS 3K	PILATUS 3K	
	<i>Longueur d'onde (Å)</i>	1.54178	1.54178	
	<i>Distance (mm)</i>	65	65	
	<i>#Images</i>	2606	5320	
	<i>Oscillations (°)</i>	0.1	0.1	
	<i>Temps d'exposition (s)</i>	5	5	
	Traitement	<i>Groupe d'espace</i>	P 2 ₁	P 2 ₁
<i>Paramètres de maille</i>		<i>a - b - c (Å)</i>	41.5 118.0 72.0	41.9 118.0 72.1
		<i>α - β - γ (°)</i>	90 103.3 90	90 103.5 90
<i>Résolution (Å)</i>		45.11 - 2.15 (2.22 - 2.15)	45.17 - 1.70 (1.73 - 1.70)	
<i>R_{merge}/R_{sym} *</i>		0.088 (0.871)	0.053 (0.934)	
<i>I/σ₁</i>		10.4 (1.0)	16.5 (1.2)	
<i>CC_{1/2}</i>		0.996 (0.294)	0.999 (0.454)	
<i>Complétude (%)</i>		99.8 (99.5)	99.8 (98.7)	
<i>Redondance</i>		4.0 (2.2)	4.8 (3.0)	
Affinement		<i>Resolution (Å)</i>	29.17 - 2.15 (2.23 - 2.15)	35.06 - 1.70 (1.76 - 1.70)
	<i>#Réflexions</i>	36561	74687	
	<i>R_{work} (%)</i>	25.98 (30.50)	18.13 (21.09)	
	<i>R_{free} (%)</i>	34.54 (38.59)	28.06 (31.63)	
	<i>#Atomes</i>	<i>Protéine</i>	5039	5208
		<i>Ligand</i>	46	48
		<i>Solvant</i>	10	446
	<i>Wilson B (Å²)</i>	34.41	20.42	
Validation	<i>RMSD</i>	<i>Bond lengths (Å)</i>	0.002	0.004
		<i>Bond angles (°)</i>	0.497	0.688
	<i>Ramachandran favored (%)</i>	98.42	98.17	
	<i>Ramachandran outliers (%)</i>	0.16	0.15	
	<i>Rotamer Outliers (%)</i>	0.93	0.54	
	<i>Cbeta Deviations</i>	0	0	
	<i>Average B (Å²)</i>	45.90	27.85	

Annexe 14 - Données cristallographiques pour les complexes mPRMT6-U.

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées pour la couche de plus haute résolution.

PDB ID / Cristal		X123D2r2	X126D6-2	X123D1v
échantillon	Protéine	mCARM1 ₁₃₀₋₄₈₇	mCARM1 ₁₃₀₋₄₉₇	mCARM1 ₁₃₀₋₄₈₇
	Cofacteur/Inhibiteur	EML709 1 mM	EML734 2 mM	EML736 0.5 mM
	Substrat	-	-	-
Cristallisation	Réservoir	100 mM Tris-HCl pH 8 12% PEG 2000 MME 100mM NaCl	100 mM Tris-HCl pH 8 18% PEG 2000 MME 100mM NaCl	100 mM Tris-HCl pH 8 10% PEG 2000 MME 100mM NaCl
	Température (°C)	20	20	20
Collecte	Source	SOLEIL Proxima1	ESRF ID30b	SOLEIL Proxima1
	Détecteur	PILATUS 6M	EIGER	PILATUS 6M
	Longueur d'onde (Å)	0.978570	0.9677	0.978570
	Distance (mm)	346.6	186.9	346.6
	#Images	1200	1800	1200
	Oscillations (°)	0.1	0.1	0.1
	Temps d'exposition (s)	0.1	0.1	0.1
Traitement	Groupe d'espace	P2 ₁ -2 ₁ -2	P2 ₁ -2 ₁ -2	P2 ₁ -2 ₁ -2
	Paramètres de maille	a - b - c (Å) 74.7 98.8 207.0	75.0 99.5 208.5	74.9 98.7 207.3
		α - β - γ (°) 90 90 90	90 90 90	90 90 90
	Résolution (Å)	48.06 - 2.30 (2.35 - 2.30)	48.44 - 2.25 (2.30 - 2.25)	45.89 - 1.85 (1.88 - 1.85)
	R _{merge} / R _{sym} *	0.13 (2.12)	0.16 (5.85)	0.12 (5.26)
	I/σ _i	7.3 (0.5)	9.3 (0.4)	6.7 (0.2)
	CC _{1/2}	0.998 (0.279)	0.999 (0.278)	0.998 (0.086)
	Complétude (%)	99.8 (99.4)	99.6 (95.9)	99.7 (98.6)
Affinement	Redondance	4.3 (3.4)	6.9 (7.0)	4.4 (3.8)
	Resolution (Å)	42.53 - 2.30 (2.38 - 2.30)	48.40 - 2.25 (2.28 - 2.25)	45.90 - 2.02 (2.04 - 2.02)
	#Réflexions	68660	74659	101175
	R _{work} (%)	21.17 (37.86)	21.18 (40.07)	20.38 (34.97)
	R _{free} (%)	24.64 (40.22)	24.74 (44.47)	23.26 (37.24)
	#Atomes	Protéine 10978	11309	11001
		Ligand 224	292	212
		Solvant 272	302	605
Wilson B (Å ²)	48.67	43.27	38.52	
Validation	RMSD	Bond lengths (Å) 0.004	0.002	0.002
		Bond angles (°) 0.618	0.569	0.54
	Ramachandran favored (%)	96.03	96.79	96.77
	Ramachandran outliers (%)	0.00	0.07	0.00
	Rotamer Outliers (%)	1.09	0.73	0.33
	Cbeta Deviations	0	0	0
	Average B (Å ²)	48.67	57.93	43.63

Annexe 15 - Données cristallographiques pour les complexes mCARM1-EML.

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées pour la couche de plus haute résolution.

PDB ID / Cristal		L45B2n	L48A4	
échantillon	<i>Protéine</i>	mPRMT6 ₃₄₋₃₇₈	mPRMT6 ₃₄₋₃₇₈	
	<i>Cofacteur/Inhibiteur</i>	EML734 2 mM	EML736 2 mM	
	<i>Substrat</i>	-	-	
Cristallisation	<i>Réservoir</i>	100 mM Tris pH 7.5 200 mM NaNO3 20% PEG 3,350	100 mM Tris pH 8.0 200 mM HCOONa 19% PEG 3,350	
	<i>Température (°C)</i>	20	20	
Collecte	<i>Source</i>	source IGBMC	source IGBMC	
	<i>Détecteur</i>	PILATUS 3K	EIGER 4M	
	<i>Longueur d'onde (Å)</i>	1.54178	1.54178	
	<i>Distance (mm)</i>	40	75	
	<i>#Images</i>	11843	2218	
	<i>Oscillations (°)</i>	0.1	0.1	
	<i>Temps d'exposition (s)</i>	5	5	
	Traitement	<i>Groupe d'espace</i>	P 2 ₁	P 2 ₁
<i>Paramètres de maille</i>		<i>a - b - c (Å)</i>	41.8 118.1 72.0	41.8 118.5 71.9
		<i>α - β - γ (°)</i>	90 104.3 90	90 103.1 90
<i>Résolution (Å)</i>		45.08 - 1.70 (1.73 - 1.70)	45.26 - 2.00 (2.05 - 2.00)	
<i>R_{merge}/R_{sym} *</i>		0.111 (1.559)	0.097 (0.629)	
<i>I/σ₁</i>		17.1 (1.7)	14.4 (2.0)	
<i>CC_{1/2}</i>		0.999 (0.798)	0.996 (0.679)	
<i>Complétude (%)</i>		100 (99.6)	98.7 (91.4)	
<i>Redondance</i>	19.4 (12.3)	3.8 (3.3)		
Affinement	<i>Resolution (Å)</i>	39.5 - 1.70 (1.76 - 1.70)	38.53 - 2.00 (2.07 - 2.00)	
	<i>#Réflexions</i>	74256	45652	
	<i>R_{work} (%)</i>	16.64 (19.86)	26.36 (29.39)	
	<i>R_{free} (%)</i>	24.20 (29.78)	34.54 (34.52)	
	<i>#Atomes</i>	<i>Protéine</i>	5166	5126
		<i>Ligand</i>	88	92
		<i>Solvant</i>	403	0
<i>Wilson B (Å²)</i>	21.59	27.9		
Validation	<i>RMSD</i>	<i>Bond lengths (Å)</i>	0.007	0.004
		<i>Bond angles (°)</i>	0.794	0.709
	<i>Ramachandran favored (%)</i>	98.77	98.61	
	<i>Ramachandran outliers (%)</i>	0.00	0.00	
	<i>Rotamer Outliers (%)</i>	0.36	0.55	
	<i>Cbeta Deviations</i>	0	0	
	<i>Average B (Å²)</i>	32.35	18.53	

Annexe 16 - Données cristallographiques pour les complexes mPRMT6-EML.

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées pour la couche de plus haute résolution.

PDB ID / Cristal		X123C3n2	X123C3b2	
échantillon	<i>Protéine</i>	mCARM1 ₁₃₀₋₄₈₇	mCARM1 ₁₃₀₋₄₈₇	
	<i>Cofacteur/Inhibiteur</i>	SATT3001 1 mM	SATT3002 1 mM	
	<i>Substrat</i>	-	-	
Cristallisation	<i>Réservoir</i>	100 mM Tris-HCl pH 8 14% PEG 2000 MME 100mM NaCl	100 mM Tris-HCl pH 8 14% PEG 2000 MME 100mM NaCl	
	<i>Température (°C)</i>	20	20	
Collecte	<i>Source</i>	SOLEIL Proxima1	SOLEIL Proxima1	
	<i>Détecteur</i>	PILATUS 6M	PILATUS 6M	
	<i>Longueur d'onde (Å)</i>	0.978570	0.978570	
	<i>Distance (mm)</i>	345.4	369.5	
	<i>#Images</i>	1200	1200	
	<i>Oscillations (°)</i>	0.1	0.1	
	<i>Temps d'exposition (s)</i>	0.1	0.1	
	Traitement	<i>Groupe d'espace</i>	P2 ₁ ² ₁ ²	P2 ₁ ² ₁ ²
<i>Paramètres de maille</i>		<i>a - b - c (Å)</i>	74.8 98.3 207.0	74.6 98.4 206.1
		<i>α - β - γ (°)</i>	90 90 90	90 90 90
<i>Résolution (Å)</i>		47.83 - 2.00 (2.03 - 2.00)	49.20 - 1.95 (1.98 - 1.95)	
<i>R_{merge} / R_{sym} *</i>		0.08 (1.29)	0.07 (1.51)	
<i>I/σ₁</i>		8.9 (1.0)	10.7 (0.9)	
<i>CC_{1/2}</i>		0.998 (0.352)	0.999 (0.371)	
<i>Complétude (%)</i>		99.9 (100.0)	99.9 (99.9)	
<i>Redondance</i>		4.4 (4.5)	4.4 (4.6)	
Affinement	<i>Resolution (Å)</i>	47.83 - 2.0 (2.07 - 20.0)	45.64 - 1.95 (2.02 - 1.95)	
	<i>#Réflexions</i>	103709	111075	
	<i>R_{work} (%)</i>	20.59 (34.31)	20.24 (34.19)	
	<i>R_{free} (%)</i>	23.63 (33.48)	21.99 (35.74)	
	<i>#Atomes</i>	<i>Protéine</i>	10753	10930
		<i>Ligand</i>	172	150
		<i>Solvant</i>	555	597
<i>Wilson B (Å²)</i>	39.80	40.18		
Validation	<i>RMSD</i>	<i>Bond lengths (Å)</i>	0.003	0.003
		<i>Bond angles (°)</i>	0.62	0.64
	<i>Ramachandran favored (%)</i>	96.70	96.16	
	<i>Ramachandran outliers (%)</i>	0.00	0.00	
	<i>Rotamer Outliers (%)</i>	0.43	0.59	
	<i>Cbeta Deviations</i>	0	0	
	<i>Average B (Å²)</i>	46.00	45.06	

Annexe 17 - Données cristallographiques pour les complexes mCARM1-SATT.

Se référer à (Karplus et al., 2015) pour les définitions des facteurs cristallographiques. Les valeurs entre parenthèses correspondent aux statistiques calculées pour la couche de plus haute résolution.

Articles joints et communications scientifiques

Articles

Hijacking DNA methyltransferase transition state analogues to produce chemical scaffolds for PRMT inhibitors. Halby, L., **Marechal, N.**, Pechalrieu, D., Cura, V., Franchini, D.-M., Faux, C., Alby, F., Troffer-Charlier, N., Kudithipudi, S., Jeltsch, A., et al. (2018). *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 373.

Transition state mimics are valuable mechanistic probes for structural studies with the arginine methyltransferase CARM1. van Haren, M.J., **Marechal*, N.**, Troffer-Charlier, N., Cianciulli, A., Sbardella, G., Cavarelli, J., and Martin, N.I. (2017). *Proc. Natl. Acad. Sci.* 114, 3625–3630.

Structural studies of protein arginine methyltransferase 2 reveal its interactions with potential substrates and inhibitors. Cura, V., **Marechal, N.**, Troffer-Charlier, N., Strub, J.-M., van Haren, M.J., Martin, N.I., Cianféroni, S., Bonnefond, L., and Cavarelli, J. (2017). *FEBS J.* 284, 77–96.

*Co-premier auteur

Présentations orales et posters.

Transition state mimics are valuable mechanistic probes for structural studies with the arginine methyltransferase CARM1. **Marechal*, N.**, van Haren, M.J., Troffer-Charlier, N., Cianciulli, A., Sbardella, G., Cavarelli, J., and Martin, N.I. (2017).

- Séminaire interne IGBMC, 22 février 2018. **Exposé oral.**
- Retraite des départements de biologie structurale IGBMC/IBMC. 13 Décembre 2017. **Exposé oral.**
- 31st Rhine-Knee Regiomeeting, 28 Septembre 2017, Munster, France. **Exposé oral.**
- IGBMC. 29 Juin 2017. **Poster.**
- French-American Doctoral Exchange Seminar 6 Mars 2017, Institut Gustave Roussy, Villejuif, France (séminaire sur sélection et invitation). **Exposé oral.**
- Colloque SFB-GTBio, « structural biology meets biophysics », 15 Décembre 2016, Obernai, France. **Exposé oral.**

Structure - function analysis of PRMT2. Cura, V., **Marechal, N.**, Troffer-Charlier, N., Strub, J.-M., van Haren, M.J., Martin, N.I., Cianféroni, S., Bonnefond, L., and Cavarelli, J. (2017). *FEBS J.* 284, 77–96.

- Retraite des départements de biologie structurale IGBMC/IBMC. 13 Décembre 2017. **Poster.**
- IGBMC. 23 Juin 2016. **Poster.**

Protein-Arginine-N-methyltransferase : From 3D structures to the rational design of compounds with pharmaceutical functions. **Marechal N**, Mailliot J, Cura V, Troffer-Charlier N, Wurtz JM, Wagner A, Cianféroni S, Bonnefond L, Cavarelli J.

- Retraite des doctorants et post-doctorants de l'IGBMC, 10 Avril 2015. **Poster.**
- IGBMC. 18 Juin 2015. **Poster.**

Bibliographie

- Abdulrahman, W., Uhring, M., Kolb-Cheynel, I., Garnier, J.-M., Moras, D., Rochel, N., Busso, D., and Poterszman, A. (2009). A set of baculovirus transfer vectors for screening of affinity tags and parallel expression strategies. *Anal. Biochem.* *385*, 383–385.
- Ahmad, K., and Henikoff, S. (2002). The Histone Variant H3.3 Marks Active Chromatin by Replication-Independent Nucleosome Assembly. *Mol. Cell* *9*, 1191–1200.
- Ajebbar, S. (2012). Synthèse de ligands à la protéine CARM1 pour l'étude de son activité enzymatique et la synthèse d'inhibiteurs sélectifs (Université de Strasbourg, France: Thèse de doctorat).
- Al-Dhaheri, M., Wu, J., Skliris, G.P., Li, J., Higashimoto, K., Wang, Y., White, K.P., Lambert, P., Zhu, Y., Murphy, L., et al. (2011). CARM1 Is an Important Determinant of ER -Dependent Breast Cancer Cell Differentiation and Proliferation in Breast Cancer Cells. *Cancer Res.* *71*, 2118–2128.
- Allan, M., Manku, S., Therrien, E., Nguyen, N., Styhler, S., Robert, M.-F., Goulet, A.-C., Petschner, A.J., Rahil, G., Robert MacLeod, A., et al. (2009). N-Benzyl-1-heteroaryl-3-(trifluoromethyl)-1H-pyrazole-5-carboxamides as inhibitors of co-activator associated arginine methyltransferase 1 (CARM1). *Bioorg. Med. Chem. Lett.* *19*, 1218–1223.
- An, W., Kim, J., and Roeder, R.G. (2004). Ordered Cooperative Functions of PRMT1, p300, and CARM1 in Transcriptional Activation by p53. *Cell* *117*, 735–748.
- Anlar, B., and Gunel-Ozcan, A. (2012). Tenascin-R: Role in the central nervous system. *Int. J. Biochem. Cell Biol.* *44*, 1385–1389.
- Antonyamy, S., Bonday, Z., Campbell, R.M., Doyle, B., Druzina, Z., Gheyi, T., Han, B., Jungheim, L.N., Qian, Y., Rauch, C., et al. (2012). Crystal structure of the human PRMT5:MEP50 complex. *Proc. Natl. Acad. Sci.* *109*, 17960–17965.
- Aryal, R.P., Kwak, P.B., Tamayo, A.G., Gebert, M., Chiu, P.-L., Walz, T., and Weitz, C.J. (2017). Macromolecular Assemblies of the Mammalian Circadian Clock. *Mol. Cell* *67*, 770-782.e6.
- Atan, A., and Güzel, Ö. (2013). How should prostate specific antigen be interpreted? *Turk. J. Urol.* *39*, 188–193.
- Avery, O.T., MacLeod, C.M., and McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from *Pneumococcus* type III. 1944. *J. Exp. Med.* *79*, 137–158.
- Bachand, F., and Silver, P.A. (2004). PRMT3 is a ribosomal protein methyltransferase that affects the cellular levels of ribosomal subunits. *EMBO J.* *23*, 2641–2650.
- Baeuerle, P.A., and Baltimore, D. (1988). 1KB: A Specific Inhibitor of the NE-KB Transcription Factor. *Science* *242*, 8.
- Baker, N.A., Sept, D., Joseph, S., Holst, M.J., and McCammon, J.A. (2001). Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci.* *98*, 10037–10041.

- Balasubramanyam, K., Varier, R.A., Altaf, M., Swaminathan, V., Siddappa, N.B., Ranga, U., and Kundu, T.K. (2004). Curcumin, a Novel p300/CREB-binding Protein-specific Inhibitor of Acetyltransferase, Represses the Acetylation of Histone/Nonhistone Proteins and Histone Acetyltransferase-dependent Chromatin Transcription. *J. Biol. Chem.* 279, 51163–51171.
- Ball, D.J., Gross, D.S., and Garrard, W.T. (1983). 5-methylcytosine is localized in nucleosomes that contain histone H1. *Proc. Natl. Acad. Sci.* 80, 5490–5494.
- Bandyopadhyay, S., Harris, D.P., Adams, G.N., Lause, G.E., McHugh, A., Tillmaand, E.G., Money, A., Willard, B., Fox, P.L., and DiCorleto, P.E. (2012). HOXA9 Methylation by PRMT5 Is Essential for Endothelial Cell Expression of Leukocyte Adhesion Molecules. *Mol. Cell. Biol.* 32, 1202–1213.
- Banfalvi, G. (2008). Chromatin Fiber Structure and Plectonemic Model of Chromosome Condensation in *Drosophila* Cells. *DNA Cell Biol.* 27, 65–70.
- Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395.
- Barrero, M.J., and Malik, S. (2006). Two functional modes of a nuclear receptor-recruited arginine methyltransferase in transcriptional activation. *Mol. Cell* 24, 233–243.
- Basta, J., and Rauchman, M. (2015). The Nucleosome Remodeling and Deacetylase (NuRD) Complex in Development and Disease. *Transl. Res. J. Lab. Clin. Med.* 165, 36–47.
- Battle, E., Sancho, E., Francí, C., Domínguez, D., Monfar, M., Baulida, J., and García de Herreros, A. (2000). The transcription factor Snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nat. Cell Biol.* 2, 84–89.
- Bauer, U.-M. (2002). Methylation at arginine 17 of histone H3 is linked to gene activation. *EMBO Rep.* 3, 39–44.
- Baylin, S.B., and Schuebel, K.E. (2007). The epigenomic era opens. *Nature* 448, 548–549.
- Beard, W.A., Prasad, R., and Wilson, S.H. (2006). Activities and Mechanism of DNA Polymerase β . In *Methods in Enzymology*, (Elsevier), pp. 91–107.
- Beitz, E. (2000). TEXshade: shading and labeling of multiple sequence alignments using LATEX2 epsilon. *Bioinforma. Oxf. Engl.* 16, 135–139.
- Berry, D.R., and Abbott, B.J. (1978). Incorporation of ¹⁴C-labeled compounds into sinefungin (A9145), a nucleoside antifungal antibiotic. *J. Antibiot. (Tokyo)* 31, 185–191.
- Bertoli, C., Skotheim, J.M., and de Bruin, R.A.M. (2013). Control of cell cycle transcription during G1 and S phases. *Nat. Rev. Mol. Cell Biol.* 14, 518–528.
- Bielefeld-Sevigny, M. (2009). AlphaLISA immunoassay platform- the “no-wash” high-throughput alternative to ELISA. *Assay Drug Dev. Technol.* 7, 90–92.
- Bissinger, E.-M., Heinke, R., Spannhoff, A., Eberlin, A., Metzger, E., Cura, V., Hassenboehler, P., Cavarelli, J., Schüle, R., Bedford, M.T., et al. (2011). Acyl derivatives of p-aminosulfonamides and dapsone as new inhibitors of the arginine methyltransferase hPRMT1. *Bioorg. Med. Chem.* 19, 3717–3731.
- Blanc, R.S., Vogel, G., Chen, T., Crist, C., and Richard, S. (2016). PRMT7 Preserves Satellite Cell Regenerative Capacity. *Cell Rep.* 14, 1528–1539.

- Blythe, S.A., Cha, S.-W., Tadjuidje, E., Heasman, J., and Klein, P.S. (2010). β -Catenin Primes Organizer Gene Expression by Recruiting a Histone H3 Arginine 8 Methyltransferase, Prmt2. *Dev. Cell* 19, 220–231.
- Boisvert, F.-M., Déry, U., Masson, J.-Y., and Richard, S. (2005a). Arginine methylation of MRE11 by PRMT1 is required for DNA damage checkpoint control. *Genes Dev.* 19, 671–676.
- Boisvert, F.-M., Rhie, A., Richard, S., and Doherty, A.J. (2005b). The GAR motif of 53BP1 is arginine methylated by PRMT1 and is necessary for 53BP1 DNA binding activity. *Cell Cycle Georget. Tex* 4, 1834–1841.
- Bonnefond, L., Stojko, J., Mailliot, J., Troffer-Charlier, N., Cura, V., Wurtz, J.-M., Cianférani, S., and Cavarelli, J. (2015). Functional insights from high resolution structures of mouse protein arginine methyltransferase 6. *J. Struct. Biol.* 191, 175–183.
- Boriack-Sjodin, P.A., Jin, L., Jacques, S.L., Drew, A., Sneeringer, C., Scott, M.P., Moyer, M.P., Ribich, S., Moradei, O., and Copeland, R.A. (2016). Structural Insights into Ternary Complex Formation of Human CARM1 with Various Substrates. *ACS Chem. Biol.* 11, 763–771.
- Brahms, H., Meheus, L., De Brabandere, V., Fischer, U., and LüHrmann, R. (2001). Symmetrical dimethylation of arginine residues in spliceosomal Sm protein B/B' and the Sm-like protein LSm4, and their interaction with the SMN protein. *RNA* 7, 1531–1542.
- Branscombe, T.L., Frankel, A., Lee, J.-H., Cook, J.R., Yang, Z., Pestka, S., and Clarke, S. (2001). PRMT5 (Janus Kinase-binding Protein 1) Catalyzes the Formation of Symmetric Dimethylarginine Residues in Proteins. *J. Biol. Chem.* 276, 32971–32976.
- Braun, D.A., Fribourg, M., and Sealfon, S.C. (2013). Cytokine Response Is Determined by Duration of Receptor and Signal Transducers and Activators of Transcription 3 (STAT3) Activation. *J. Biol. Chem.* 288, 2986–2993.
- Buhr, N., Carapito, C., Schaeffer, C., Kieffer, E., Van Dorselaer, A., and Viville, S. (2008). Nuclear proteome analysis of undifferentiated mouse embryonic stem and germ cells. *ELECTROPHORESIS* 29, 2381–2390.
- Burg, J.M., Link, J.E., Morgan, B.S., Heller, F.J., Hargrove, A.E., and McCafferty, D.G. (2015). KDM1 Class Flavin-Dependent Protein Lysine Demethylases. *Biopolymers* 104, 213–246.
- Butler, J.S., Zurita-Lopez, C.I., Clarke, S.G., Bedford, M.T., and Dent, S.Y.R. (2011). Protein-arginine Methyltransferase 1 (PRMT1) Methylates Ash2L, a Shared Component of Mammalian Histone H3K4 Methyltransferase Complexes. *J. Biol. Chem.* 286, 12234–12244.
- Cabezas-Cruz, A., Lancelot, J., Caby, S., Oliveira, G., and Pierce, R.J. (2014). Epigenetic control of gene function in schistosomes: a source of therapeutic targets? *Front. Genet.* 5.
- Cadigan, K.M. (2012). TCFs and Wnt/ β -catenin signaling: more than one way to throw the switch. *Curr. Top. Dev. Biol.* 98, 1–34.
- Cha, B., and Jho, E.-H. (2012). Protein arginine methyltransferases (PRMTs) as therapeutic targets. *Expert Opin. Ther. Targets* 16, 651–664.
- Chang, B., Chen, Y., Zhao, Y., and Bruick, R.K. (2007). JMJD6 Is a Histone Arginine Demethylase. *Science* 318, 444–447.

- Chang, Y.-I., Hua, W.-K., Yao, C.-L., Hwang, S.-M., Hung, Y.-C., Kuan, C.-J., Leou, J.-S., and Lin, W.-J. (2010). Protein-arginine Methyltransferase 1 Suppresses Megakaryocytic Differentiation via Modulation of the p38 MAPK Pathway in K562 Cells. *J. Biol. Chem.* *285*, 20595–20606.
- Chan-Penebre, E., Kuplast, K.G., Majer, C.R., Boriack-Sjodin, P.A., Wigle, T.J., Johnston, L.D., Rioux, N., Munchhof, M.J., Jin, L., Jacques, S.L., et al. (2015). A selective inhibitor of PRMT5 with in vivo and in vitro potency in MCL models. *Nat. Chem. Biol.* *11*, 432–437.
- Chari, A., Golas, M.M., Klingenhäger, M., Neuenkirchen, N., Sander, B., Englbrecht, C., Sickmann, A., Stark, H., and Fischer, U. (2008). An assembly chaperone collaborates with the SMN complex to generate spliceosomal SnRNPs. *Cell* *135*, 497–509.
- Chen, D. (1999). Regulation of Transcription by a Protein Methyltransferase. *Science* *284*, 2174–2177.
- Chen, C., Jin, J., James, D.A., Adams-Cioaba, M.A., Park, J.G., Guo, Y., Tenaglia, E., Xu, C., Gish, G., Min, J., et al. (2009). Mouse Piwi interactome identifies binding mechanism of Tdrkh Tudor domain to arginine methylated Miwi. *Proc. Natl. Acad. Sci.* *106*, 20336–20341.
- Chen, D., Huang, S.-M., and Stallcup, M.R. (2000). Synergistic, p160 Coactivator-dependent Enhancement of Estrogen Receptor Function by CARM1 and p300. *J. Biol. Chem.* *275*, 40810–40816.
- Chen, S.L., Loffler, K.A., Chen, D., Stallcup, M.R., and Muscat, G.E.O. (2002). The Coactivator-associated Arginine Methyltransferase Is Necessary for Muscle Differentiation: CARM1 coactivates myocyte enhancer factor-2. *J. Biol. Chem.* *277*, 4324–4333.
- Cheng, D., Yadav, N., King, R.W., Swanson, M.S., Weinstein, E.J., and Bedford, M.T. (2004). Small Molecule Regulators of Protein Arginine Methyltransferases. *J. Biol. Chem.* *279*, 23892–23899.
- Cheng, D., Côté, J., Shaaban, S., and Bedford, M.T. (2007). The Arginine Methyltransferase CARM1 Regulates the Coupling of Transcription and mRNA Processing. *Mol. Cell* *25*, 71–83.
- Cheng, D., Valente, S., Castellano, S., Sbardella, G., Di Santo, R., Costi, R., Bedford, M.T., and Mai, A. (2011a). Novel 3,5-Bis(bromohydroxybenzylidene)piperidin-4-ones as Coactivator-Associated Arginine Methyltransferase 1 Inhibitors: Enzyme Selectivity and Cellular Activity. *J. Med. Chem.* *54*, 4928–4932.
- Cheng, Y., Frazier, M., Lu, F., Cao, X., and Redinbo, M.R. (2011b). Crystal Structure of the Plant Epigenetic Protein Arginine Methyltransferase 10. *J. Mol. Biol.* *414*, 106–122.
- Chiao, P.J., Shin, D.M., Sacks, P.G., Hong, W.K., and Tainsky, M.A. (1992). Elevated expression of the ribosomal protein S2 gene in human tumors. *Mol. Carcinog.* *5*, 219–231.
- Chikayama, E., Kigawa, T., Saito, K., Koshiba, S., Inoue, M., and Yokoyama, S. (2005). Solution structure of the SH3 domain of the Protein arginine N-methyltransferase 2. *BE Publ.*
- Chiou, N.-T., and Lynch, K.W. (2014). Mechanisms of spliceosomal assembly. *Methods Mol. Biol. Clifton NJ* *1126*, 35–43.
- Choi, S., Jung, C.-R., Kim, J.-Y., and Im, D.-S. (2008). PRMT3 inhibits ubiquitination of ribosomal protein S2 and together forms an active enzyme complex. *Biochim. Biophys. Acta BBA - Gen. Subj.* *1780*, 1062–1069.
- Chow, J.C., Yen, Z., Ziesche, S.M., and Brown, C.J. (2005). Silencing of the mammalian X chromosome. *Annu. Rev. Genomics Hum. Genet.* *6*, 69–92.

- Chu, Y., Li, G., and Guo, H. (2013). QM/MM MD and free energy simulations of the methylation reactions catalyzed by protein arginine methyltransferase PRMT3. *Can. J. Chem.* *91*, 605–612.
- Clevers, H., and van de Wetering, M. (1997). TCF/LEF factor earn their wings. *Trends Genet. TIG* *13*, 485–489.
- Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* *50*, 760–763.
- Compere, S.J., and Palmiter, R.D. (1981). DNA methylation controls the inducibility of the mouse metallothionein-I gene lymphoid cells. *Cell* *25*, 233–240.
- Copeland, R.A. (2018). Protein methyltransferase inhibitors as precision cancer therapeutics: a decade of discovery. *Philos. Trans. R. Soc. B Biol. Sci.* *373*.
- Côté, J., and Richard, S. (2005). Tudor Domains Bind Symmetrical Dimethylated Arginines. *J. Biol. Chem.* *280*, 28476–28483.
- Couture, J.-F., Collazo, E., and Trievel, R.C. (2006). Molecular recognition of histone H3 by the WD40 protein WDR5. *Nat. Struct. Mol. Biol.* *13*, 698–703.
- Covic, M., Hassa, P.O., Saccani, S., Buerki, C., Meier, N.I., Lombardi, C., Imhof, R., Bedford, M.T., Natoli, G., and Hottiger, M.O. (2005). Arginine methyltransferase CARM1 is a promoter-specific regulator of NF- κ B-dependent gene expression. *EMBO J.* *24*, 85–96.
- Coward, J., Borek, E., Zappia, V., Williams-Ashman, H., and Schlenk, F. (1977). Chemical mechanisms of methyl transfer reactions: comparison of methylases with nonenzymic ‘model reactions.’ *The Biochemistry of Adenosylmethionine* 127–144.
- Cura, V., Troffer-Charlier, N., Wurtz, J.-M., Bonnefond, L., and Cavarelli, J. (2014). Structural insight into arginine methylation by the mouse protein arginine methyltransferase 7: a zinc finger freezes the mimic of the dimeric state into a single active site. *Acta Crystallogr. D Biol. Crystallogr.* *70*, 2401–2412.
- Cura, V., Marechal, N., Troffer-Charlier, N., Strub, J.-M., van Haren, M.J., Martin, N.I., Cianférani, S., Bonnefond, L., and Cavarelli, J. (2017). Structural studies of protein arginine methyltransferase 2 reveal its interactions with potential substrates and inhibitors. *FEBS J.* *284*, 77–96.
- Cuthbert, G.L., Daujat, S., Snowden, A.W., Erdjument-Bromage, H., Hagiwara, T., Yamada, M., Schneider, R., Gregory, P.D., Tempst, P., Bannister, A.J., et al. (2004). Histone Deimination Antagonizes Arginine Methylation. *Cell* *118*, 545–553.
- Daley, J.M., and Sung, P. (2014). 53BP1, BRCA1, and the Choice between Recombination and End Joining at DNA Double-Strand Breaks. *Mol. Cell. Biol.* *34*, 1380–1388.
- Daujat, S., Bauer, U.-M., Shah, V., Turner, B., Berger, S., and Kouzarides, T. (2002). Crosstalk between CARM1 Methylation and CBP Acetylation on Histone H3. *Curr. Biol.* *12*, 2090–2097.
- Deans, C., and Maggert, K.A. (2015). What Do You Mean, “Epigenetic”? *Genetics* *199*, 887–896.
- Debler, E.W., Jain, K., Warmack, R.A., Feng, Y., Clarke, S.G., Blobel, G., and Stavropoulos, P. (2016). A glutamate/aspartate switch controls product specificity in a protein arginine methyltransferase. *Proc. Natl. Acad. Sci.* *113*, 2068–2073.

- Diebold, M.-L., Fribourg, S., Koch, M., Metzger, T., and Romier, C. (2011). Deciphering correct strategies for multiprotein complex assembly by co-expression: Application to complexes as large as the histone octamer. *J. Struct. Biol.* *175*, 178–188.
- van Dijk, T.B., Gillemans, N., Stein, C., Fanis, P., Demmers, J., van de Corput, M., Essers, J., Grosveld, F., Bauer, U.-M., and Philipson, S. (2010). Friend of Prmt1, a Novel Chromatin Target of Protein Arginine Methyltransferases. *Mol. Cell. Biol.* *30*, 260–272.
- Dowden, J., Hong, W., Parry, R.V., Pike, R.A., and Ward, S.G. (2010). Toward the development of potent and selective bisubstrate inhibitors of protein arginine methyltransferases. *Bioorg. Med. Chem. Lett.* *20*, 2103–2105.
- Dowden, J., Pike, R.A., Parry, R.V., Hong, W., Muhsen, U.A., and Ward, S.G. (2011). Small molecule inhibitors that discriminate between protein arginine N-methyltransferases PRMT1 and CARM1. *Org. Biomol. Chem.* *9*, 7814–7821.
- Drew, A.E., Moradei, O., Jacques, S.L., Rioux, N., Boriack-Sjodin, A.P., Allain, C., Scott, M.P., Jin, L., Raimondi, A., Handler, J.L., et al. (2017). Identification of a CARM1 Inhibitor with Potent In Vitro and In Vivo Activity in Preclinical Models of Multiple Myeloma. *Sci. Rep.* *7*.
- Drulyte, I., Johnson, R.M., Hesketh, E.L., Hurdiss, D.L., Scarff, C.A., Porav, S.A., Ranson, N.A., Muench, S.P., and Thompson, R.F. (2018). Approaches to altering particle distributions in cryo-electron microscopy sample preparation. *Acta Crystallogr. Sect. Struct. Biol.* *74*, 560–571.
- Du, J., Johnson, L.M., Jacobsen, S.E., and Patel, D.J. (2015). DNA methylation pathways and their crosstalk with histone methylation. *Nat. Rev. Mol. Cell Biol.* *16*, 519–532.
- Duncan, K.W., Rioux, N., Boriack-Sjodin, P.A., Munchhof, M.J., Reiter, L.A., Majer, C.R., Jin, L., Johnston, L.D., Chan-Penebre, E., Kuplast, K.G., et al. (2015). Structure and Property Guided Design in the Identification of PRMT5 Tool Compound EPZ015666. *ACS Med. Chem. Lett.* *7*, 162–166.
- Eberharter, A., and Becker, P.B. (2002). Histone acetylation: a switch between repressive and permissive chromatin. *EMBO Rep.* *3*, 224–229.
- Eden, S., and Cedar, H. (1994). Role of DNA methylation in the regulation of transcription. *Curr. Opin. Genet. Dev.* *4*, 255–259.
- Ehrlich, M. (2009). DNA hypomethylation in cancer cells. *Epigenomics* *1*, 239–259.
- El Messaoudi, S., Fabrizio, E., Rodriguez, C., Chuchana, P., Fauquier, L., Cheng, D., Theillet, C., Vandel, L., Bedford, M.T., and Sardet, C. (2006). Coactivator-associated arginine methyltransferase 1 (CARM1) is a positive regulator of the Cyclin E1 gene. *Proc. Natl. Acad. Sci.* *103*, 13351–13356.
- El-Andaloussi, N., Valovka, T., Touelle, M., Steinacher, R., Focke, F., Gehrig, P., Covic, M., Hassa, P.O., Schär, P., Hübscher, U., et al. (2006). Arginine Methylation Regulates DNA Polymerase β . *Mol. Cell* *22*, 51–62.
- Engilberge, S., Riobé, F., Di Pietro, S., Lassalle, L., Coquelle, N., Arnaud, C.-A., Pitrat, D., Mulatier, J.-C., Madern, D., Breyton, C., et al. (2017). Crystallophore: a versatile lanthanide complex for protein crystallography combining nucleating effects, phasing properties, and luminescence. *Chem. Sci.* *8*, 5909–5917.

- Fabbrizio, E., El Messaoudi, S., Polanowska, J., Paul, C., Cook, J.R., Lee, J.-H., Nègre, V., Rousset, M., Pestka, S., Le Cam, A., et al. (2002). Negative regulation of transcription by the type II arginine methyltransferase PRMT5. *EMBO Rep.* 3, 641–645.
- Felsenfeld, G. (2014). A Brief History of Epigenetics. *Cold Spring Harb. Perspect. Biol.* 6, a018200–a018200.
- Feng, Q., He, B., Jung, S.-Y., Song, Y., Qin, J., Tsai, S.Y., Tsai, M.-J., and O'Malley, B.W. (2009). Biochemical Control of CARM1 Enzymatic Activity by Phosphorylation. *J. Biol. Chem.* 284, 36167–36174.
- Feng, Y., Li, M., Wang, B., and Zheng, Y.G. (2010). Discovery and Mechanistic Study of a Class of Protein Arginine Methylation Inhibitors. *J. Med. Chem.* 53, 6028–6039.
- Feng, Y., Xie, N., Jin, M., Stahley, M.R., Stivers, J.T., and Zheng, Y.G. (2011). A Transient Kinetic Analysis of PRMT1 Catalysis. *Biochemistry (Mosc.)* 50, 7033–7044.
- Feng, Y., Maity, R., Whitelegge, J.P., Hadjikyriacou, A., Li, Z., Zurita-Lopez, C., Al-Hadid, Q., Clark, A.T., Bedford, M.T., Masson, J.-Y., et al. (2013). Mammalian Protein Arginine Methyltransferase 7 (PRMT7) Specifically Targets R X R Sites in Lysine- and Arginine-rich Regions. *J. Biol. Chem.* 288, 37010–37025.
- Ferreira de Freitas, R., Eram, M.S., Smil, D., Szewczyk, M.M., Kennedy, S., Brown, P.J., Santhakumar, V., Barsyte-Lovejoy, D., Arrowsmith, C.H., Vedadi, M., et al. (2016). Discovery of a Potent and Selective Coactivator Associated Arginine Methyltransferase 1 (CARM1) Inhibitor by Virtual Screening. *J. Med. Chem.* 59, 6838–6847.
- Filipescu, D., Szenker, E., and Almouzni, G. (2013). Developmental roles of histone H3 variants and their chaperones. *Trends Genet.* 29, 630–640.
- Fischer, U., Liu, Q., and Dreyfuss, G. (1997). The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell* 90, 1023–1029.
- Franke, D., and Svergun, D.I. (2009). *DAMMIF*, a program for rapid *ab-initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* 42, 342–346.
- Frankel, A., Yadav, N., Lee, J., Branscombe, T.L., Clarke, S., and Bedford, M.T. (2002). The Novel Human Protein Arginine N -Methyltransferase PRMT6 Is a Nuclear Enzyme Displaying Unique Substrate Specificity. *J. Biol. Chem.* 277, 3537–3543.
- Friesen, W.J., Paushkin, S., Wyce, A., Massenet, S., Pesiridis, G.S., Van Duyne, G., Rappsilber, J., Mann, M., and Dreyfuss, G. (2001a). The Methylosome, a 20S Complex Containing JBP1 and pICln, Produces Dimethylarginine-Modified Sm Proteins. *Mol. Cell. Biol.* 21, 8289–8300.
- Friesen, W.J., Massenet, S., Paushkin, S., Wyce, A., and Dreyfuss, G. (2001b). SMN, the Product of the Spinal Muscular Atrophy Gene, Binds Preferentially to Dimethylarginine-Containing Protein Targets. *Mol. Cell* 7, 1111–1117.
- Friesen, W.J., Wyce, A., Paushkin, S., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). A Novel WD Repeat Protein Component of the Methylosome Binds Sm Proteins. *J. Biol. Chem.* 277, 8243–8247.
- Fuhrmann, J., and Thompson, P.R. (2016). Protein Arginine Methylation and Citrullination in Epigenetic Regulation. *ACS Chem. Biol.* 11, 654–668.

- Fujiwara, T., Mori, Y., Chu, D.L., Koyama, Y., Miyata, S., Tanaka, H., Yachi, K., Kubo, T., Yoshikawa, H., and Tohyama, M. (2006). CARM1 Regulates Proliferation of PC12 Cells by Methylating HuD. *Mol. Cell. Biol.* 26, 2273–2285.
- Fulcher, A.J., Sivakumaran, H., Jin, H., Rawle, D.J., Harrich, D., and Jans, D.A. (2016). The protein arginine methyltransferase PRMT6 inhibits HIV-1 Tat nucleolar retention. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* 1863, 254–262.
- Gana, R., Rao, S., Huang, H., Wu, C., and Vasudevan, S. (2013). Structural and functional studies of S-adenosyl-L-methionine binding proteins: a ligand-centric approach. *BMC Struct. Biol.* 13, 6.
- Gao, G., Dhar, S., and Bedford, M.T. (2017). PRMT5 regulates IRES-dependent translation via methylation of hnRNP A1. *Nucleic Acids Res.* gkw1367.
- Gao, L., Sun, J., Liu, F., Zhang, H., and Ma, Y. (2016). Higher expression levels of the HOXA9 gene, closely associated with MLL-PTD and EZH2 mutations, predict inferior outcome in acute myeloid leukemia. *OncoTargets Ther.* 711.
- Gary, J.D., and Clarke, S. (1998). RNA and protein interactions modulated by protein arginine methylation. *Prog. Nucleic Acid Res. Mol. Biol.* 61, 65–131.
- Gershey, E.L., Vidali, G., and Allfrey, V.G. (1968). Chemical studies of histone acetylation. The occurrence of epsilon-N-acetyllysine in the f2a1 histone. *J. Biol. Chem.* 243, 5018–5022.
- Gherardini, P.F., Ausiello, G., Russell, R.B., and Helmer-Citterich, M. (2010). Modular architecture of nucleotide-binding pockets. *Nucleic Acids Res.* 38, 3809–3816.
- Gibbs, A.C., Bjorndahl, T.C., Hodges, R.S., and Wishart, D.S. (2002). Probing the Structural Determinants of Type II' β -Turn Formation in Peptides and Proteins. *J. Am. Chem. Soc.* 124, 1203–1213.
- Goldman, A.D., Beatty, J.T., and Landweber, L.F. (2016). The TIM Barrel Architecture Facilitated the Early Evolution of Protein-Mediated Metabolism. *J. Mol. Evol.* 82, 17–26.
- Gonsalvez, G.B., Tian, L., Ospina, J.K., Boisvert, F.-M., Lamond, A.I., and Matera, A.G. (2007). Two distinct arginine methyltransferases are required for biogenesis of Sm-class ribonucleoproteins. *J. Cell Biol.* 178, 733–740.
- Granger, D.N., and Senchenkova, E. (2010). *Inflammation and the Microcirculation* (San Rafael (CA): Morgan & Claypool Life Sciences).
- Gros, L., Delaporte, C., Frey, S., Decesse, J., de Saint-Vincent, B.R., Cavarec, L., Dubart, A., Gudkov, A.V., and Jacquemin-Sablon, A. (2003). Identification of new drug sensitivity genes using genetic suppressor elements: protein arginine N-methyltransferase mediates cell sensitivity to DNA-damaging agents. *Cancer Res.* 63, 164–171.
- Guccione, E., Bassi, C., Casadio, F., Martinato, F., Cesaroni, M., Schuchlantz, H., Lüscher, B., and Amati, B. (2007). Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. *Nature* 449, 933–937.
- Gui, S., Wooderchak, W.L., Daly, M.P., Porter, P.J., Johnson, S.J., and Hevel, J.M. (2011). Investigation of the Molecular Origins of Protein-arginine Methyltransferase I (PRMT1) Product Specificity Reveals a Role for Two Conserved Methionine Residues. *J. Biol. Chem.* 286, 29118–29126.

- Gui, S., Gathiaka, S., Li, J., Qu, J., Acevedo, O., and Hevel, J.M. (2014). A Remodeled Protein Arginine Methyltransferase 1 (PRMT1) Generates Symmetric Dimethylarginine. *J. Biol. Chem.* 289, 9320–9327.
- Halby, L., Marechal, N., Pechalrieu, D., Cura, V., Franchini, D.-M., Faux, C., Alby, F., Troffer-Charlier, N., Kudithipudi, S., Jeltsch, A., et al. (2018). Hijacking DNA methyltransferase transition state analogues to produce chemical scaffolds for PRMT inhibitors. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 373.
- Han, H.-S., Jung, C.-Y., Yoon, Y.-S., Choi, S., Choi, D., Kang, G., Park, K.-G., Kim, S.-T., and Koo, S.-H. (2014). Arginine Methylation of CRT2 Is Critical in the Transcriptional Control of Hepatic Glucose Metabolism. *Sci. Signal.* 7, ra19–ra19.
- Handrkova, H., Petrak, J., Halada, P., Pospisilova, D., and Cmejla, R. (2011). Tyrosine 87 is vital for the activity of human protein arginine methyltransferase 3 (PRMT3). *Biochim. Biophys. Acta BBA - Proteins Proteomics* 1814, 277–282.
- Hang, R., Liu, C., Ahmad, A., Zhang, Y., Lu, F., and Cao, X. (2014). *Arabidopsis* protein arginine methyltransferase 3 is required for ribosome biogenesis by affecting precursor ribosomal RNA processing. *Proc. Natl. Acad. Sci.* 111, 16190–16195.
- van Haren, M., van Ufford, L.Q., Moret, E.E., and Martin, N.I. (2015). Synthesis and evaluation of protein arginine N-methyltransferase inhibitors designed to simultaneously occupy both substrate binding sites. *Org. Biomol. Chem.* 13, 549–560.
- van Haren, M.J., Marechal, N., Troffer-Charlier, N., Cianciulli, A., Sbardella, G., Cavarelli, J., and Martin, N.I. (2017). Transition state mimics are valuable mechanistic probes for structural studies with the arginine methyltransferase CARM1. *Proc. Natl. Acad. Sci.* 114, 3625–3630.
- Hasegawa, M., Toma-Fukai, S., Kim, J.-D., Fukamizu, A., and Shimizu, T. (2014a). Protein arginine methyltransferase 7 has a novel homodimer-like structure formed by tandem repeats. *FEBS Lett.* 588, 1942–1948.
- Hasegawa, M., Toma-Fukai, S., Kim, J.-D., Fukamizu, A., and Shimizu, T. (2014b). Protein arginine methyltransferase 7 has a novel homodimer-like structure formed by tandem repeats. *FEBS Lett.* 588, 1942–1948.
- Hassenboehler, P. (2010). Etude structurale et fonctionnelle de la protéine arginine méthyltransférase CARM1 (Université de Strasbourg, France: Thèse de doctorat).
- Hebbes, T.R., Thorne, A.W., and Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J.* 7, 1395–1402.
- Heinke, R., Spannhoff, A., Meier, R., Trojer, P., Bauer, I., Jung, M., and Sippl, W. (2009). Virtual Screening and Biological Characterization of Novel Histone Arginine Methyltransferase PRMT1 Inhibitors. *ChemMedChem* 4, 69–77.
- Heinrichs, V., Bach, M., Winkelmann, G., and Lührmann, R. (1990). U1-specific protein C needed for efficient complex formation of U1 snRNP with a 5' splice site. *Science* 247, 69–72.
- Heitz, E. (1928). Das Heterochromatin der Moose. *Jahrb Wiss Botanik* 69, 762–818.
- Hernandez, S.J., Dolivo, D.M., and Dominko, T. (2017). PRMT8 demonstrates variant-specific expression in cancer cells and correlates with patient survival in breast, ovarian and gastric cancer. *Oncol. Lett.* 13, 1983–1989.

- Herrmann, F., Lee, J., Bedford, M.T., and Fackelmayer, F.O. (2005). Dynamics of Human Protein Arginine Methyltransferase 1 (PRMT1) in Vivo. *J. Biol. Chem.* *280*, 38005–38010.
- Ho, M.-C., Wilczek, C., Bonanno, J.B., Xing, L., Seznec, J., Matsui, T., Carter, L.G., Onikubo, T., Kumar, P.R., Chan, M.K., et al. (2013). Structure of the Arginine Methyltransferase PRMT5-MEP50 Reveals a Mechanism for Substrate Specificity. *PLoS ONE* *8*, e57008.
- Holliday, R., and Pugh, J.E. (1975). DNA modification mechanisms and gene activity during development. *Science* *187*, 226–232.
- Hong, W., Li, J., Laughton, C.A., Yap, L.F., Paterson, I.C., and Wang, H. (2014). Investigating the binding preferences of small molecule inhibitors of human protein arginine methyltransferase 1 using molecular modelling. *J. Mol. Graph. Model.* *51*, 193–202.
- Howe, F.S., Boubriak, I., Sale, M.J., Nair, A., Clynes, D., Grijzenhout, A., Murray, S.C., Woloszczuk, R., and Mellor, J. (2014). Lysine Acetylation Controls Local Protein Conformation by Influencing Proline Isomerization. *Mol. Cell* *55*, 733–744.
- Hsu, S.-C., Galceran, J., and Grosschedl, R. (1998). Modulation of Transcriptional Regulation by LEF-1 in Response to Wnt-1 Signaling and Association with β -Catenin. *Mol. Cell. Biol.* *18*, 4807–4818.
- Hu, H., Owens, E.A., Su, H., Yan, L., Levitz, A., Zhao, X., Henary, M., and Zheng, Y.G. (2015). Exploration of Cyanine Compounds as Selective Inhibitors of Protein Arginine Methyltransferases: Synthesis and Biological Evaluation. *J. Med. Chem.* *58*, 1228–1243.
- Huang, J., Vogel, G., Yu, Z., Almazan, G., and Richard, S. (2011). Type II Arginine Methyltransferase PRMT5 Regulates Gene Expression of Inhibitors of Differentiation/DNA Binding Id2 and Id4 during Glial Cell Differentiation. *J. Biol. Chem.* *286*, 44424–44432.
- Human Genome Sequencing Consortium, I. (2004). Finishing the euchromatic sequence of the human genome. *Nature* *431*, 931–945.
- Hyllus, D., Stein, C., Schnabel, K., Schiltz, E., Imhof, A., Dou, Y., Hsieh, J., and Bauer, U.-M. (2007). PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation. *Genes Dev.* *21*, 3369–3380.
- Iberg, A.N., Espejo, A., Cheng, D., Kim, D., Michaud-Levesque, J., Richard, S., and Bedford, M.T. (2008). Arginine Methylation of the Histone H3 Tail Impedes Effector Binding. *J. Biol. Chem.* *283*, 3006–3010.
- Imataka, H., Gradi, A., and Sonenberg, N. (1998). A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J.* *17*, 7480–7489.
- Iwasaki, H., Kovacic, J.C., Olive, M., Beers, J.K., Yoshimoto, T., Crook, M.F., Tonelli, L.H., and Nabel, E.G. (2010). Disruption of Protein Arginine N-Methyltransferase 2 Regulates Leptin Signaling and Produces Leanness In Vivo Through Loss of STAT3 Methylation. *Circ. Res.* *107*, 992–1001.
- Jeanes, A., Gottardi, C., and Yap, A. (2008). Cadherins and cancer: how does cadherin dysfunction promote tumor progression? *Oncogene* *27*, 6920–6929.

- Jelinic, P., Stehle, J.-C., and Shaw, P. (2006). The Testis-Specific Factor CTCFL Cooperates with the Protein Methyltransferase PRMT7 in H19 Imprinting Control Region Methylation. *PLoS Biol.* 4, e355.
- Jiang, H., Zhou, Z., Jin, S., Xu, K., Zhang, H., Xu, J., Sun, Q., Wang, J., and Xu, J. (2018). PRMT9 promotes hepatocellular carcinoma invasion and metastasis via activating PI3K/Akt/GSK-3 β /Snail signaling. *Cancer Sci.* 109, 1414–1427.
- Jiang, W., Roemer, M.E., and Newsham, I.F. (2005). The tumor suppressor DAL-1/4.1B modulates protein arginine N-methyltransferase 5 activity in a substrate-specific manner. *Biochem. Biophys. Res. Commun.* 329, 522–530.
- Jobert, L., Argentini, M., and Tora, L. (2009). PRMT1 mediated methylation of TAF15 is required for its positive gene regulatory function. *Exp. Cell Res.* 315, 1273–1286.
- Johnson, E.C.B., and Kent, S.B.H. (2006). Insights into the mechanism and catalysis of the native chemical ligation reaction. *J. Am. Chem. Soc.* 128, 6640–6646.
- Josling, G.A., Selvarajah, S.A., Petter, M., and Duffy, M.F. (2012). The Role of Bromodomain Proteins in Regulating Gene Expression. *Genes* 3, 320–343.
- Jung, G.A., Shin, B.S., Jang, Y.S., Sohn, J.B., Woo, S.R., Kim, J.E., Choi, G., Lee, K.-M., Min, B.H., Lee, K.-H., et al. (2011). Methylation of eukaryotic elongation factor 2 induced by basic fibroblast growth factor *via* mitogen-activated protein kinase. *Exp. Mol. Med.* 43, 550.
- Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D Biol. Crystallogr.* 66, 133–144.
- Kaniskan, H.Ü., Szewczyk, M.M., Yu, Z., Eram, M.S., Yang, X., Schmidt, K., Luo, X., Dai, M., He, F., Zang, I., et al. (2015). A Potent, Selective and Cell-Active Allosteric Inhibitor of Protein Arginine Methyltransferase 3 (PRMT3). *Angew. Chem. Int. Ed.* 54, 5166–5170.
- Karkhanis, V., Wang, L., Tae, S., Hu, Y.-J., Imbalzano, A.N., and Sif, S. (2012). Protein Arginine Methyltransferase 7 Regulates Cellular Response to DNA Damage by Methylating Promoter Histones H2A and H4 of the Polymerase δ Catalytic Subunit Gene, *POLD1*. *J. Biol. Chem.* 287, 29801–29814.
- Karplus, P.A., and Diederichs, K. (2015). Assessing and maximizing data quality in macromolecular crystallography. *Curr. Opin. Struct. Biol.* 34, 60–68.
- Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780.
- Katsanis, N., Yaspo, M.-L., and Fisher, E.M.C. (1997). Identification and mapping of a novel human gene, HRMT1L1, homologous to the rat protein arginine N-methyltransferase 1 (PRMT1) gene. *Mamm. Genome* 8, 526–529.
- Khorasanizadeh, S. (2004). The nucleosome: from genomic organization to genomic regulation. *Cell* 116, 259–272.
- Kim, D., Lee, J., Cheng, D., Li, J., Carter, C., Richie, E., and Bedford, M.T. (2010a). Enzymatic Activity Is Required for the *in Vivo* Functions of CARM1. *J. Biol. Chem.* 285, 1147–1152.
- Kim, J.-D., Park, K.-E., Ishida, J., Kako, K., Hamada, J., Kani, S., Takeuchi, M., Namiki, K., Fukui, H., Fukuhara, S., et al. (2015). PRMT8 as a phospholipase regulates Purkinje cell dendritic arborization and motor coordination. *Sci. Adv.* 1, e1500615–e1500615.

- Kim, Y.-R., Lee, B.K., Park, R.-Y., Nguyen, N.T.X., Bae, J.A., Kwon, D.D., and Jung, C. (2010b). RDesiefafrccehraertinctletial CARM1 expression in prostate and colorectal cancers. 13.
- Kimanius, D., Forsberg, B.O., Scheres, S.H., and Lindahl, E. (2016). Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *ELife* 5, e18722.
- Klingl, S., Scherer, M., Stamminger, T., and Muller, Y.A. (2015). Controlled crystal dehydration triggers a space-group switch and shapes the tertiary structure of cytomegalovirus immediate-early 1 (IE1) protein. *Acta Crystallogr. D Biol. Crystallogr.* 71, 1493–1504.
- Koh, S.S., Li, H., Lee, Y.-H., Widelitz, R.B., Chuong, C.-M., and Stallcup, M.R. (2002). Synergistic Coactivator Function by Coactivator-associated Arginine Methyltransferase (CARM) 1 and β -Catenin with Two Different Classes of DNA-binding Transcriptional Activators. *J. Biol. Chem.* 277, 26031–26035.
- Kölbel, K., Ihling, C., Bellmann-Sickert, K., Neundorff, I., Beck-Sickinger, A.G., Sinz, A., Kühn, U., and Wahle, E. (2009). Type I Arginine Methyltransferases PRMT1 and PRMT-3 Act Distributively. *J. Biol. Chem.* 284, 8274–8282.
- Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003). *PRIMUS*: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* 36, 1277–1282.
- Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* 184, 868–871.
- Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. *Science* 184, 865–868.
- Kousaka, A., Mori, Y., Koyama, Y., Taneda, T., Miyata, S., and Tohyama, M. (2009). The distribution and characterization of endogenous protein arginine N-methyltransferase 8 in mouse CNS. *Neuroscience* 163, 1146–1157.
- Kozbial, P.Z., and Mushegian, A.R. (2005). Natural history of S-adenosylmethionine-binding proteins. *BMC Struct. Biol.* 26.
- Krauss, I.R., Sica, F., Mattia, C.A., and Merlino, A. (2012). Increasing the X-ray Diffraction Power of Protein Crystals by Dehydration: The Case of Bovine Serum Albumin and a Survey of Literature Data. *Int. J. Mol. Sci.* 13, 3782–3800.
- Kremer, J.R., Mastrorade, D.N., and McIntosh, J.R. (1996). Computer Visualization of Three-Dimensional Image Data Using IMOD. *J. Struct. Biol.* 116, 71–76.
- Kuhn, P., Chumanov, R., Wang, Y., Ge, Y., Burgess, R.R., and Xu, W. (2011). Automethylation of CARM1 allows coupling of transcription and mRNA splicing. *Nucleic Acids Res.* 39, 2717–2726.
- Kurochkina, N., and Guha, U. (2012). SH3 domains: modules of protein–protein interactions. *Biophys. Rev.* 5, 29–39.
- Kwak, Y.T., Guo, J., Prajapati, S., Park, K.-J., Surabhi, R.M., Miller, B., Gehrig, P., and Gaynor, R.B. (2003). Methylation of SPT5 Regulates Its Interaction with RNA Polymerase II and Transcriptional Elongation Properties. *Mol. Cell* 11, 1055–1066.
- Lacroix, M., Messaoudi, S.E., Rodier, G., Le Cam, A., Sardet, C., and Fabrizio, E. (2008). The histone-binding protein COPR5 is required for nuclear functions of the protein arginine methyltransferase PRMT5. *EMBO Rep.* 9, 452–458.

- Lakowski, T.M., and Frankel, A. (2008). A Kinetic Study of Human Protein Arginine N - Methyltransferase 6 Reveals a Distributive Mechanism. *J. Biol. Chem.* *283*, 10015–10025.
- Lakowski, T.M., and Frankel, A. (2009). Kinetic analysis of human protein arginine N-methyltransferase 2: formation of monomethyl- and asymmetric dimethyl-arginine residues on histone H4. *Biochem. J.* *421*, 253–261.
- Landry-Voyer, A.-M., Bilodeau, S., Bergeron, D., Dionne, K.L., Port, S.A., Rouleau, C., Boisvert, F.-M., Kehlenbach, R.H., and Bachand, F. (2016). Human PDCD2L Is an Export Substrate of CRM1 That Associates with 40S Ribosomal Subunit Precursors. *Mol. Cell. Biol.* *36*, 3019–3032.
- Langley, G.W., Brinkø, A., Münzel, M., Walport, L.J., Schofield, C.J., and Hopkinson, R.J. (2016). Analysis of JmjC Demethylase-Catalyzed Demethylation Using Geometrically-Constrained Lysine Analogues. *ACS Chem. Biol.* *11*, 755–762.
- Lawrence, T. (2009). The Nuclear Factor NF- κ B Pathway in Inflammation. *Cold Spring Harb. Perspect. Biol.* *1*.
- Le Romancer, M., Treilleux, I., Leconte, N., Robin-Lespinasse, Y., Sentis, S., Bouchekioua-Bouzaghrou, K., Goddard, S., Gobert-Gosse, S., and Corbo, L. (2008). Regulation of estrogen rapid signaling through arginine methylation by PRMT1. *Mol. Cell* *31*, 212–221.
- Lee, J. (2002). PABP1 identified as an arginine methyltransferase substrate using high-density protein arrays. *EMBO Rep.* *3*, 268–273.
- Lee, H.W., Kim, S., and Paik, W.K. (1977). S-adenosylmethionine: protein-arginine methyltransferase. Purification and mechanism of the enzyme. *Biochemistry (Mosc.)* *16*, 78–85.
- Lee, J., Sayegh, J., Daniel, J., Clarke, S., and Bedford, M.T. (2005a). PRMT8, a New Membrane-bound Tissue-specific Member of the Protein Arginine Methyltransferase Family. *J. Biol. Chem.* *280*, 32890–32896.
- Lee, J.-H., Cook, J.R., Yang, Z.-H., Mirochnitchenko, O., Gunderson, S.I., Felix, A.M., Herth, N., Hoffmann, R., and Pestka, S. (2005b). PRMT7, a New Protein Arginine Methyltransferase That Synthesizes Symmetric Dimethylarginine. *J. Biol. Chem.* *280*, 3656–3664.
- Lee, P.K.M., Goh, W.W.B., and Sng, J.C.G. (2017). Network-based characterization of the synaptic proteome reveals that removal of epigenetic regulator *Prmt8* restricts proteins associated with synaptic maturation. *J. Neurochem.* *140*, 613–628.
- Lee, W.-C., Lin, W.-L., Matsui, T., Chen, E.S.-W., Wei, T.-Y.W., Lin, W.-H., Hu, H., Zheng, Y.G., Tsai, M.-D., and Ho, M.-C. (2015). Protein Arginine Methyltransferase 8: Tetrameric Structure and Protein Substrate Specificity. *Biochemistry (Mosc.)* *54*, 7514–7523.
- Lei, N., Zhang, X., Chen, H., Wang, Y., Zhan, Y., Zheng, Z., Shen, Y., and Wu, Q. (2009). A feedback regulatory loop between methyltransferase PRMT1 and orphan receptor TR3. *Nucleic Acids Res.* *37*, 832–848.
- Li, G., and Zhu, P. (2015). Structure and organization of chromatin fiber in the nucleus. *FEBS Lett.* *589*, 2893–2904.
- Li, Y., Li, Z., and Zhu, W.-G. (2017). Molecular mechanisms of epigenetic regulators as activatable targets in cancer theranostics. *Curr. Med. Chem.*

- Li, Z., Lee, I., Moradi, E., Hung, N.-J., Johnson, A.W., and Marcotte, E.M. (2009). Rational Extension of the Ribosome Biogenesis Pathway Using Network-Guided Genetics. *PLoS Biol.* 7, e1000213.
- Lin, W.-J., Gary, J.D., Yang, M.C., Clarke, S., and Herschman, H.R. (1996). The Mammalian Immediate-early TIS21 Protein and the Leukemia-associated BTG1 Protein Interact with a Protein-arginine N -Methyltransferase. *J. Biol. Chem.* 271, 15034–15044.
- Lipson, R.S., Webb, K.J., and Clarke, S.G. (2010). Rmt1 catalyzes zinc-finger independent arginine methylation of ribosomal protein Rps2 in *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.* 391, 1658–1662.
- Liu, F., Li, F., Ma, A., Dobrovetsky, E., Dong, A., Gao, C., Korboukh, I., Liu, J., Smil, D., Brown, P.J., et al. (2013). Exploiting an Allosteric Binding Site of PRMT3 Yields Potent and Selective Inhibitors. *J. Med. Chem.* 56, 2110–2124.
- Liu, H., Wang, C., Lee, S., Deng, Y., Wither, M., Oh, S., Ning, F., Dege, C., Zhang, Q., Liu, X., et al. (2017). Clipping of arginine-methylated histone tails by JMJD5 and JMJD7. *Proc. Natl. Acad. Sci.* 114, E7717–E7726.
- Liu, H., Wang, C., Lee, S., Ning, F., Wang, Y., Zhang, Q., Chen, Z., Zang, J., Nix, J., Dai, S., et al. (2018). Specific Recognition of Arginine Methylated Histone Tails by JMJD5 and JMJD7. *Sci. Rep.* 8.
- Liu, K., Chen, C., Guo, Y., Lam, R., Bian, C., Xu, C., Zhao, D.Y., Jin, J., MacKenzie, F., Pawson, T., et al. (2010). Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain. *Proc. Natl. Acad. Sci.* 107, 18398–18403.
- Liu, Q., Fischer, U., Wang, F., and Dreyfuss, G. (1997). The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell* 90, 1013–1021.
- Lorenzen, S., Peters, B., Goede, A., Preissner, R., and Frömmel, C. (2004). Conservation of cis prolyl bonds in proteins during evolution. *Proteins Struct. Funct. Bioinforma.* 58, 589–595.
- Luger, K., der, A.W.M., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *389*, 12.
- Mai, A., Cheng, D., Bedford, M.T., Valente, S., Nebbioso, A., Perrone, A., Brosch, G., Sbardella, G., De Bellis, F., Miceli, M., et al. (2008). Epigenetic Multiple Ligands: Mixed Histone/Protein Methyltransferase, Acetyltransferase, and Class III Deacetylase (Sirtuin) Inhibitors. *J. Med. Chem.* 51, 2279–2290.
- Mailliot, J. (2013). Étude structurale de l’histoneméthyltransférase « CARM1 » et de ses complexes biologiquement significatifs : des structures 3D vers la conception rationnelle de composés à action pharmacologique (Université de Strasbourg, France: Thèse de doctorat).
- Malygin, E.G., and Hattman, S. (2012). DNA methyltransferases: mechanistic models derived from kinetic analysis. *Crit. Rev. Biochem. Mol. Biol.* 47, 97–193.
- Mao, R., Shao, J., Zhu, K., Zhang, Y., Ding, H., Zhang, C., Shi, Z., Jiang, H., Sun, D., Duan, W., et al. (2017). Potent, Selective, and Cell Active Protein Arginine Methyltransferase 5 (PRMT5) Inhibitor Developed by Structure-Based Virtual Screening and Hit Optimization. *J. Med. Chem.* 60, 6289–6304.

- Marek, M., Kannan, S., Hauser, A.-T., Moraes Mourão, M., Caby, S., Cura, V., Stolfa, D.A., Schmidtkunz, K., Lancelot, J., Andrade, L., et al. (2013). Structural Basis for the Inhibition of Histone Deacetylase 8 (HDAC8), a Key Epigenetic Player in the Blood Fluke *Schistosoma mansoni*. *PLoS Pathog.* 9, e1003645.
- Marfella, C.G.A., and Imbalzano, A.N. (2007). The Chd Family of Chromatin Remodelers. *Mutat. Res.* 618, 30–40.
- Marmorstein, R., and Roth, S.Y. (2001). Histone acetyltransferases: function, structure, and catalysis. *Curr. Opin. Genet. Dev.* 11, 155–161.
- Marsden, M.P., and Laemmli, U.K. (1979). Metaphase chromosome structure: evidence for a radial loop model. *Cell* 17, 849–858.
- Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* 6, 838–849.
- Marzluff, W.F., and Duronio, R.J. (2002). Histone mRNA expression: multiple levels of cell cycle regulation and important developmental consequences. *Curr. Opin. Cell Biol.* 14, 692–699.
- Mastrorade, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* 152, 36–51.
- Matsuoka, T., Tahara, M., Yokoi, T., Masumoto, N., Takeda, T., Yamaguchi, M., Tasaka, K., Kurachi, H., and Murata, Y. (1999). Tyrosine Phosphorylation of STAT3 by Leptin through Leptin Receptor in Mouse Metaphase 2 Stage Oocyte. *Biochem. Biophys. Res. Commun.* 256, 480–484.
- Maxam, A.M., and Gilbert, W. (1977). A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U. S. A.* 74, 560–564.
- McGraw, W.T., Potempa, J., Farley, D., and Travis, J. (1999). Purification, Characterization, and Sequence Analysis of a Potential Virulence Factor from *Porphyromonas gingivalis*, Peptidylarginine Deiminase. *Infect Immun* 67, 9.
- McPherson, A. (2001). A comparison of salts for the crystallization of macromolecules. *Protein Sci. Publ. Protein Soc.* 10, 418–422.
- Meister, G. (2002). Assisted RNP assembly: SMN and PRMT5 complexes cooperate in the formation of spliceosomal UsnRNPs. *EMBO J.* 21, 5853–5863.
- Meyer, R., Wolf, S.S., and Obendorf, M. (2007). PRMT2, a member of the protein arginine methyltransferase family, is a coactivator of the androgen receptor. *J. Steroid Biochem. Mol. Biol.* 107, 1–14.
- Miao, F., Li, S., Chavez, V., Lanting, L., and Natarajan, R. (2006). Coactivator-Associated Arginine Methyltransferase-1 Enhances Nuclear Factor- κ B-Mediated Gene Transcription through Methylation of Histone H3 at Arginine 17. *Mol. Endocrinol.* 20, 1562–1573.
- Michaud-Levesque, J., and Richard, S. (2009). Thrombospondin-1 Is a Transcriptional Repression Target of PRMT6. *J. Biol. Chem.* 284, 21338–21346.
- Migliori, V., Müller, J., Phalke, S., Low, D., Bezzi, M., Mok, W.C., Sahu, S.K., Gunaratne, J., Capasso, P., Bassi, C., et al. (2012). Symmetric dimethylation of H3R2 is a newly identified histone mark that supports euchromatin maintenance. *Nat. Struct. Mol. Biol.* 19, 136–144.

- Minakhina, S., Naryshkina, T., Changela, N., Tan, W., and Steward, R. (2016). Zfp8/PDCD2 Interacts with RpS2 Connecting Ribosome Maturation and Gene-Specific Translation. *PLoS ONE* 11.
- Miranda, T.B., and Jones, P.A. (2007). DNA methylation: The nuts and bolts of repression. *J. Cell. Physiol.* 213, 384–390.
- Miranda, T.B., Miranda, M., Frankel, A., and Clarke, S. (2004). PRMT7 Is a Member of the Protein Arginine Methyltransferase Family with a Distinct Substrate Specificity. *J. Biol. Chem.* 279, 22902–22907.
- Mitchell, L.H., Drew, A.E., Ribich, S.A., Rioux, N., Swinger, K.K., Jacques, S.L., Lingaraj, T., Boriack-Sjodin, P.A., Waters, N.J., Wigle, T.J., et al. (2015). Aryl Pyrazoles as Potent Inhibitors of Arginine Methyltransferases: Identification of the First PRMT6 Tool Compound. *ACS Med. Chem. Lett.* 6, 655–659.
- Miyamoto, K., Tomizawa, T., Koshiya, S., Inoue, M., Kigawa, T., and Yokoyama, S. (2004). Solution structure of the C2H2 zinc finger domain of the protein arginine N-methyltransferase 3 from *Mus musculus*. *BE Publ.*
- Miyata, S., Mori, Y., and Tohyama, M. (2010). PRMT3 is essential for dendritic spine maturation in rat hippocampal neurons. *Brain Res.* 1352, 11–20.
- Morettn, A., Baldwin, R.M., and Cote, J. (2015). Arginine methyltransferases as novel therapeutic targets for breast cancer. *Mutagenesis* 30, 177–189.
- Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., and Olson, A.J. (2009). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* 30, 2785–2791.
- Morrison, A.J., and Shen, X. (2005). DNA Repair in the Context of Chromatin. *Cell Cycle* 4, 513–512.
- Mowen, K.A., Schurter, B.T., Fathman, J.W., David, M., and Glimcher, L.H. (2004). Arginine methylation of NIP45 modulates cytokine gene expression in effector T lymphocytes. *Mol. Cell* 15, 559–571.
- Murray, K. (1964). The occurrence of epsilon-N-methyl lysine in histones. *Biochemistry (Mosc.)* 3, 10–15.
- Na, J., Lee, K., Kim, H.-G., Shin, J.-Y., Na, W., Jeong, H., Lee, J.-W., Cho, S., Kim, W.-S., and Ju, B.-G. (2012). Role of Type II Protein Arginine Methyltransferase 5 in the Regulation of Circadian Per1 Gene. *PLoS ONE* 7, e48152.
- Natchiar, S.K., Myasnikov, A.G., Kratzat, H., Hazemann, I., and Klaholz, B.P. (2017). Visualization of chemical modifications in the human 80S ribosome structure. *Nature* 551, 472–477.
- Nelson, C.J., Santos-Rosa, H., and Kouzarides, T. (2006). Proline isomerization of histone H3 regulates lysine methylation and gene expression. *Cell* 126, 905–916.
- Niewmierzycka, A., and Clarke, S. (1999). S-Adenosylmethionine-dependent Methylation in *Saccharomyces cerevisiae*: identification of a novel protein arginine methyltransferase. *J. Biol. Chem.* 274, 814–824.
- Obianyo, O., Osborne, T.C., and Thompson, P.R. (2008). Kinetic Mechanism of Protein Arginine Methyltransferase 1[†]. *Biochemistry (Mosc.)* 47, 10420–10427.

- Ohkura, N., Takahashi, M., Yaguchi, H., Nagamura, Y., and Tsukada, T. (2005). Coactivator-associated Arginine Methyltransferase 1, CARM1, Affects Pre-mRNA Splicing in an Isoform-specific Manner. *J. Biol. Chem.* *280*, 28927–28935.
- Ohtsubo, M., Theodoras, A.M., Schumacher, J., Roberts, J.M., and Pagano, M. (1995). Human cyclin E, a nuclear protein essential for the G1-to-S phase transition. *Mol. Cell. Biol.* *15*, 2612–2624.
- Ord, M.G., and Stocken, L.A. (1967). Phosphate and thiol groups in histone f3 from rat liver and thymus nuclei. *Biochem. J.* *102*, 631–636.
- Ou, C.-Y., LaBonte, M.J., Manegold, P.C., So, A.Y.-L., Ianculescu, I., Gerke, D.S., Yamamoto, K.R., Ladner, R.D., Kahn, M., Kim, J.H., et al. (2011). A Coactivator Role of CARM1 in the Dysregulation of β -Catenin Activity in Colorectal Cancer Cell Growth and Gene Expression. *Mol. Cancer Res.* *9*, 660–670.
- Pahlich, S., Zakaryan, R.P., and Gehring, H. (2008). Identification of proteins interacting with protein arginine methyltransferase 8: The Ewing sarcoma (EWS) protein binds independent of its methylation state. *Proteins Struct. Funct. Bioinforma.* *72*, 1125–1137.
- Pal, S., Vishwanath, S.N., Erdjument-Bromage, H., Tempst, P., and Sif, S. (2004). Human SWI/SNF-Associated PRMT5 Methylates Histone H3 Arginine 8 and Negatively Regulates Expression of ST7 and NM23 Tumor Suppressor Genes. *Mol. Cell. Biol.* *24*, 9630–9645.
- Palmer, D.K., O’Day, K., Trong, H.L., Charbonneau, H., and Margolis, R.L. (1991). Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. *Proc. Natl. Acad. Sci.* *88*, 3734–3738.
- Pande, V. (2016). Understanding the Complexity of Epigenetic Target Space. *J. Med. Chem.* *59*, 1299–1307.
- Paronetto, M.P., Achsel, T., Massiello, A., Chalfant, C.E., and Sette, C. (2007). The RNA-binding protein Sam68 modulates the alternative splicing of Bcl-x. *J. Cell Biol.* *176*, 929–939.
- Penney, J., Seo, J., Kritskiy, O., Elmsaouri, S., Gao, F., Pao, P.-C., Su, S.C., and Tsai, L.-H. (2017). Loss of Protein Arginine Methyltransferase 8 Alters Synapse Composition and Function, Resulting in Behavioral Defects. *J. Neurosci.* *37*, 8655–8666.
- Perreault, A., Bellemer, C., and Bachand, F. (2008). Nuclear export competence of pre-40S subunits in fission yeast requires the ribosomal protein Rps2. *Nucleic Acids Res.* *36*, 6132–6142.
- Perreault, A., Gascon, S., D’Amours, A., Aletta, J.M., and Bachand, F. (2009). A Methyltransferase-independent Function for Rmt3 in Ribosomal Subunit Homeostasis. *J. Biol. Chem.* *284*, 15026–15037.
- Petsko, G.A. (2006). Transformation. *Genome Biol.* *7*, 117.
- Pilka, E.S., James, T., and Lisztwan, J.H. (2015). Structural definitions of Jumonji family demethylase selectivity. *Drug Discov. Today* *20*, 743–749.
- Pollack, B.P., Kotenko, S.V., He, W., Izotova, L.S., Barnoski, B.L., and Pestka, S. (1999a). The Human Homologue of the Yeast Proteins Skb1 and Hsl7p Interacts with Jak Kinases and Contains Protein Methyltransferase Activity. *J. Biol. Chem.* *274*, 31531–31542.

- Pollack, B.P., Kotenko, S.V., He, W., Izotova, L.S., Barnoski, B.L., and Pestka, S. (1999b). The Human Homologue of the Yeast Proteins Skb1 and Hsl7p Interacts with Jak Kinases and Contains Protein Methyltransferase Activity. *J. Biol. Chem.* *274*, 31531–31542.
- Portela, A., and Esteller, M. (2010). Epigenetic modifications and human disease. *Nat. Biotechnol.* *28*, 1057–1068.
- Pu, W.T., Krapivinsky, G.B., Krapivinsky, L., and Clapham, D.E. (1999). pICln Inhibits snRNP Biogenesis by Binding Core Spliceosomal Proteins. *Mol. Cell. Biol.* *19*, 4113–4120.
- Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* *14*, 290–296.
- Purandare, A.V., Chen, Z., Huynh, T., Pang, S., Geng, J., Vaccaro, W., Poss, M.A., Oconnell, J., Nowak, K., and Jayaraman, L. (2008). Pyrazole inhibitors of coactivator associated arginine methyltransferase 1 (CARM1). *Bioorg. Med. Chem. Lett.* *18*, 4438–4441.
- Qi, C., Chang, J., Zhu, Y., Yeldandi, A.V., Rao, S.M., and Zhu, Y.-J. (2002). Identification of Protein Arginine Methyltransferase 2 as a Coactivator for Estrogen Receptor α . *J. Biol. Chem.* *277*, 28624–28630.
- Raijmakers, R., Zendman, A.J.W., Egberts, W.V., Vossenaar, E.R., Raats, J., Soede-Huijbregts, C., Rutjes, F.P.J.T., van Veelen, P.A., Drijfhout, J.W., and Pruijn, G.J.M. (2007). Methylation of Arginine Residues Interferes with Citrullination by Peptidylarginine Deiminases in vitro. *J. Mol. Biol.* *367*, 1118–1129.
- Rawlings, J.S. (2004). The JAK/STAT signaling pathway. *J. Cell Sci.* *117*, 1281–1283.
- Razin, A., and Cedar, H. (1977). Distribution of 5-methylcytosine in chromatin. *Proc. Natl. Acad. Sci.* *74*, 2725–2728.
- Ren, J., Wang, Y., Liang, Y., Zhang, Y., Bao, S., and Xu, Z. (2010). Methylation of Ribosomal Protein S10 by Protein-arginine Methyltransferase 5 Regulates Ribosome Biogenesis. *J. Biol. Chem.* *285*, 12695–12705.
- Rho, J., Choi, S., Seong, Y.R., Cho, W.-K., Kim, S.H., and Im, D.-S. (2001). PRMT5, Which Forms Distinct Homo-oligomers, Is a Member of the Protein-arginine Methyltransferase Family. *J. Biol. Chem.* *276*, 11393–11401.
- Rice, N.R., and Ernst, M.K. (1993). In vivo control of NF-kappa B activation by I kappa B alpha. *EMBO J.* *12*, 4685–4695.
- Rioux, N., Duncan, K.W., Lantz, R.J., Miao, X., Chan-Penebre, E., Moyer, M.P., Munchhof, M.J., Copeland, R.A., Chesworth, R., and Waters, N.J. (2016). Species differences in metabolism of EPZ015666, an oxetane-containing protein arginine methyltransferase-5 (PRMT5) inhibitor. *Xenobiotica Fate Foreign Compd. Biol. Syst.* *46*, 268–277.
- Robinson, P.J.J., Fairall, L., Huynh, V.A.T., and Rhodes, D. (2006). EM measurements define the dimensions of the “30-nm” chromatin fiber: Evidence for a compact, interdigitated structure. *Proc. Natl. Acad. Sci.* *103*, 6506–6511.
- Rossetto, D., Avvakumov, N., and Côté, J. (2012). Histone phosphorylation: A chromatin modification involved in diverse nuclear events. *Epigenetics* *7*, 1098–1108.

- Rossi, F., Forné, T., Antoine, E., Tazi, J., Brunel, C., and Cathala, G. (1996). Involvement of U1 small nuclear ribonucleoproteins (snRNP) in 5' splice site-U1 snRNP interaction. *J. Biol. Chem.* *271*, 23985–23991.
- Rossmann, M.G., Moras, D., and Olsen, K.W. (1974). Chemical and biological evolution of nucleotide-binding protein. *Nature* *250*, 194–199.
- Sack, J.S., Thieffine, S., Bandiera, T., Fasolini, M., Duke, G.J., Jayaraman, L., Kish, K.F., Klei, H.E., Purandare, A.V., Rosettani, P., et al. (2011). Structural basis for CARM1 inhibition by indole and pyrazole inhibitors. *Biochem. J.* *436*, 331–339.
- Saha, P., Eichbaum, Q., Silberman, E.D., Mayer, B.J., and Dutta, A. (1997). p21CIP1 and Cdc25A: competition between an inhibitor and an activator of cyclin-dependent kinases. *Mol. Cell. Biol.* *17*, 4338–4345.
- Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* *234*, 779–815.
- Sanchez, S.E., Petrillo, E., Beckwith, E.J., Zhang, X., Rugnone, M.L., Hernando, C.E., Cuevas, J.C., Godoy Herz, M.A., Depetris-Chauvin, A., Simpson, C.G., et al. (2010). A methyl transferase links the circadian clock to the regulation of alternative splicing. *Nature* *468*, 112–116.
- Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* *74*, 5463–5467.
- Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl. Acad. Sci.* *103*, 1412–1417.
- Schurter, B.T., Koh, S.S., Chen, D., Bunick, G.J., Harp, J.M., Hanson, B.L., Henschen-Edman, A., Mackay, D.R., Stallcup, M.R., and Aswad, D.W. (2001). Methylation of Histone H3 by Coactivator-Associated Arginine Methyltransferase 1[†]. *Biochemistry (Mosc.)* *40*, 5747–5756.
- Seto, E., and Yoshida, M. (2014). Erasers of Histone Acetylation: The Histone Deacetylase Enzymes. *Cold Spring Harb. Perspect. Biol.* *6*, a018713.
- Shia, W.-J., Okumura, A.J., Yan, M., Sarkeshik, A., Lo, M.-C., Matsuura, S., Komeno, Y., Zhao, X., Nimer, S.D., Yates, J.R., et al. (2012). PRMT1 interacts with AML1-ETO to promote its transcriptional activation and progenitor cell proliferative potential. *Blood* *119*, 4953–4962.
- Shin, D.M., Chiao, P.J., Sacks, P.G., Shin, H.J., Hong, W.K., Hittelman, W.N., and Tainsky, M.A. (1993). Activation of ribosomal protein S2 gene expression in a hamster model of chemically induced oral carcinogenesis. *Carcinogenesis* *14*, 163–166.
- Shishkova, E., Zeng, H., Liu, F., Kwiecien, N.W., Hebert, A.S., Coon, J.J., and Xu, W. (2017). Global mapping of CARM1 substrates defines enzyme specificity and substrate recognition. *Nat. Commun.* *8*, 15571.
- Siarheyeva, A., Senisterra, G., Allali-Hassani, A., Dong, A., Dobrovetsky, E., Wasney, G.A., Chau, I., Marcellus, R., Hajian, T., Liu, F., et al. (2012). An Allosteric Inhibitor of Protein Arginine Methyltransferase 3. *Structure* *20*, 1425–1435.
- Sievers, F., and Higgins, D.G. (2014). Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol. Biol.* Clifton NJ *1079*, 105–116.

- Sikorsky, T., Hobor, F., Krizanova, E., Pasulka, J., Kubicek, K., and Stefl, R. (2012). Recognition of asymmetrically dimethylated arginine by TDRD3. *Nucleic Acids Res.* *40*, 11748–11755.
- Sims, R.J., Nishioka, K., and Reinberg, D. (2003). Histone lysine methylation: a signature for chromatin function. *Trends Genet.* *19*, 629–639.
- Simunek, M., Hoßfeld, U., and Wissemann, V. (2011). ‘Rediscovery’ revised - the cooperation of Erich and Armin von Tschermak-Seysenegg in the context of the ‘rediscovery’ of Mendel’s laws in 1899-1901: The ‘rediscovery’ of Mendel’s laws, 1899-1901. *Plant Biol.* *13*, 835–841.
- Singh, V., Miranda, T.B., Jiang, W., Frankel, A., Roemer, M.E., Robb, V.A., Gutmann, D.H., Herschman, H.R., Clarke, S., and Newsham, I.F. (2004). DAL-1/4.1B tumor suppressor interacts with protein arginine N-methyltransferase 3 (PRMT3) and inhibits its ability to methylate substrates in vitro and in vivo. *Oncogene* *23*, 7761–7771.
- Singhroy, D.N., Mesplède, T., Sabbah, A., Quashie, P.K., Falgoutyret, J.-P., and Wainberg, M.A. (2013). Automethylation of protein arginine methyltransferase 6 (PRMT6) regulates its stability and its anti-HIV-1 activity. *Retrovirology* *10*, 73.
- Sinha, S.H., Owens, E.A., Feng, Y., Yang, Y., Xie, Y., Tu, Y., Henary, M., and Zheng, Y.G. (2012). Synthesis and evaluation of carbocyanine dyes as PRMT inhibitors and imaging agents. *Eur. J. Med. Chem.* *54*, 647–659.
- Sivolob, A., and Prunell, A. (2003). Linker Histone-dependent Organization and Dynamics of Nucleosome Entry/Exit DNAs. *J. Mol. Biol.* *331*, 1025–1040.
- Solari, C., Echegaray, C.V., Luzzani, C., Cosentino, M.S., Waisman, A., Petrone, M.V., Francia, M., Sassone, A., Canizo, J., Sevlever, G., et al. (2016). Protein arginine Methyltransferase 8 gene is expressed in pluripotent stem cells and its expression is modulated by the transcription factor Sox2. *Biochem. Biophys. Res. Commun.* *473*, 194–199.
- Song, G., Ouyang, G., and Bao, S. (2005). The activation of Akt/PKB signaling pathway and cell survival. *J. Cell. Mol. Med.* *9*, 59–71.
- Spannhoff, A., Heinke, R., Bauer, I., Trojer, P., Metzger, E., Gust, R., Schüle, R., Brosch, G., Sippl, W., and Jung, M. (2007). Target-Based Approach to Inhibitors of Histone Arginine Methyltransferases. *J. Med. Chem.* *50*, 2319–2325.
- Sparks, A.B., Rider, J.E., Hoffman, N.G., Fowlkes, D.M., Quillam, L.A., and Kay, B.K. (1996). Distinct ligand preferences of Src homology 3 domains from Src, Yes, Abl, Cortactin, p53bp2, PLCgamma, Crk, and Grb2. *Proc. Natl. Acad. Sci. U. S. A.* *93*, 1540–1544.
- Stein, C., Nötzold, R.R., Riedl, S., Bouchard, C., and Bauer, U.-M. (2016). The Arginine Methyltransferase PRMT6 Cooperates with Polycomb Proteins in Regulating HOXA Gene Expression. *PLOS ONE* *11*, e0148892.
- Stopa, N., Krebs, J.E., and Shechter, D. (2015). The PRMT5 arginine methyltransferase: many roles in development, cancer and beyond. *Cell. Mol. Life Sci.* *72*, 2041–2059.
- Strahl, B.D., Briggs, S.D., Brame, C.J., Caldwell, J.A., Koh, S.S., Ma, H., Cook, R.G., Shabanowitz, J., Hunt, D.F., Stallcup, M.R., et al. (2001). Methylation of histone H4 at arginine 3 occurs in vivo and is mediated by the nuclear receptor coactivator PRMT1. *Curr. Biol.* *11*, 996–1000.

- Sun, L., Wang, M., Lv, Z., Yang, N., Liu, Y., Bao, S., Gong, W., and Xu, R.-M. (2011). Structural insights into protein arginine symmetric dimethylation by PRMT5. *Proc. Natl. Acad. Sci.* *108*, 20538–20543.
- Svergun, D.I. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* *25*, 495–503.
- Svergun, D.I. (1999). Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing. *Biophys. J.* *76*, 2879–2886.
- Swiercz, R., Person, M.D., and Bedford, M.T. (2005). Ribosomal protein S2 is a substrate for mammalian PRMT3 (protein arginine methyltransferase 3). *Biochem. J.* *386*, 85–91.
- Swiercz, R., Cheng, D., Kim, D., and Bedford, M.T. (2007). Ribosomal Protein rpS2 Is Hypomethylated in PRMT3-deficient Mice. *J. Biol. Chem.* *282*, 16917–16923.
- Takai, H., Masuda, K., Sato, T., Sakaguchi, Y., Suzuki, T., Suzuki, T., Koyama-Nasu, R., Nasu-Nishimura, Y., Katou, Y., Ogawa, H., et al. (2014). 5-Hydroxymethylcytosine Plays a Critical Role in Glioblastomagenesis by Recruiting the CHTOP-Methylosome Complex. *Cell Rep.* *9*, 48–60.
- Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: An extensible image processing suite for electron microscopy. *J. Struct. Biol.* *157*, 38–46.
- Tang, J., Gary, J.D., Clarke, S., and Herschman, H.R. (1998). PRMT 3, a Type I Protein Arginine *N*-Methyltransferase That Differs from PRMT1 in Its Oligomerization, Subcellular Localization, Substrate Specificity, and Regulation. *J. Biol. Chem.* *273*, 16935–16945.
- Tang, J., Frankel, A., Cook, R.J., Kim, S., Paik, W.K., Williams, K.R., Clarke, S., and Herschman, H.R. (2000). PRMT1 Is the Predominant Type I Protein Arginine Methyltransferase in Mammalian Cells. *J. Biol. Chem.* *275*, 7723–7730.
- Tang, L., Nogales, E., and Ciferri, C. (2010). Structure and Function of SWI/SNF Chromatin Remodeling Complexes and Mechanistic Implications for Transcription. *Prog. Biophys. Mol. Biol.* *102*, 122–128.
- Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* *14*, 1025–1040.
- Tchuem Tchuente, L.-A., Rollinson, D., Stothard, J.R., and Molyneux, D. (2017). Moving from control to elimination of schistosomiasis in sub-Saharan Africa: time to change and adapt strategies. *Infect. Dis. Poverty* *6*.
- Teske, K.A., and Hadden, M.K. (2017). Methyllysine binding domains: Structural insight and small molecule probe development. *Eur. J. Med. Chem.* *136*, 14–35.
- Teyssier, C., Chen, D., and Stallcup, M.R. (2002). Requirement for Multiple Domains of the Protein Arginine Methyltransferase CARM1 in Its Transcriptional Coactivator Function. *J. Biol. Chem.* *277*, 46066–46072.
- Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013). Defining the RGG/RG Motif. *Mol. Cell* *50*, 613–623.
- Therrien, E., Larouche, G., Manku, S., Allan, M., Nguyen, N., Styhler, S., Robert, M.-F., Goulet, A.-C., Besterman, J.M., Nguyen, H., et al. (2009). 1,2-Diamines as inhibitors of co-activator associated arginine methyltransferase 1 (CARM1). *Bioorg. Med. Chem. Lett.* *19*, 6725–6732.

- Thoma, F. (1979). Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J. Cell Biol.* *83*, 403–427.
- Thomson, T., and Lin, H. (2009). The Biogenesis and Function PIWI Proteins and piRNAs: Progress and Prospect. *Annu. Rev. Cell Dev. Biol.* *25*, 355–376.
- Timm, D.E., Bowman, V., Madsen, R., and Rauch, C. (2018). Cryo-electron microscopy structure of a human PRMT5:MEP50 complex. *PLOS ONE* *13*, e0193205.
- Torres-Padilla, M.-E., Parfitt, D.-E., Kouzarides, T., and Zernicka-Goetz, M. (2007). Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* *445*, 214–218.
- Trewick, S.C., McLaughlin, P.J., and Allshire, R.C. (2005). Methylation: lost in hydroxylation? *EMBO Rep.* *6*, 315–320.
- Tripsianes, K., Madl, T., Machyna, M., Fessas, D., Englbrecht, C., Fischer, U., Neugebauer, K.M., and Sattler, M. (2011). Structural basis for dimethylarginine recognition by the Tudor domains of human SMN and SPF30 proteins. *Nat. Struct. Mol. Biol.* *18*, 1414–1420.
- Troffer-Charlier, N., Cura, V., Hassenboehler, P., Moras, D., and Cavarelli, J. (2007). Functional insights from structures of coactivator-associated arginine methyltransferase 1 domains. *EMBO J.* *26*, 4391–4401.
- Uhlen, M., Fagerberg, L., Hallstrom, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, A., Kampf, C., Sjostedt, E., Asplund, A., et al. (2015). Tissue-based map of the human proteome. *Science* *347*, 1260419–1260419.
- Urlaub, H., Raker, V.A., Kostka, S., and Lührmann, R. (2001). Sm protein–Sm site RNA interactions within the inner ring of the spliceosomal snRNP core structure. *EMBO J.* *20*, 187–196.
- Vagin, V.V., Wohlschlegel, J., Qu, J., Jonsson, Z., Huang, X., Chuma, S., Girard, A., Sachidanandam, R., Hannon, G.J., and Aravin, A.A. (2009). Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members. *Genes Dev.* *23*, 1749–1762.
- Vale, N., Gouveia, M.J., Rinaldi, G., Brindley, P.J., Gärtner, F., and Correia da Costa, J.M. (2017). Praziquantel for Schistosomiasis: Single-Drug Metabolism Revisited, Mode of Action, and Resistance. *Antimicrob. Agents Chemother.* *61*, e02582-16.
- Veland, N., Hardikar, S., Zhong, Y., Gayatri, S., Dan, J., Strahl, B.D., Rothbart, S.B., Bedford, M.T., and Chen, T. (2017). The Arginine Methyltransferase PRMT6 Regulates DNA Methylation and Contributes to Global DNA Hypomethylation in Cancer. *Cell Rep.* *21*, 3390–3397.
- Verma, M., Charles, R.C.M., Chakrapani, B., Coumar, M.S., Govindaraju, G., Rajavelu, A., Chavali, S., and Dhayalan, A. (2017). PRMT7 Interacts with ASS1 and Citrullinemia Mutations Disrupt the Interaction. *J. Mol. Biol.* *429*, 2278–2289.
- Vhuiyan, M.I., Pak, M.L., Park, M.A., Thomas, D., Lakowski, T.M., Chalfant, C.E., and Frankel, A. (2017). PRMT2 interacts with splicing factors and regulates the alternative splicing of *BCL-X*: *J. Biochem. (Tokyo)* mvw102.
- Vignali, M., Hassan, A.H., Neely, K.E., and Workman, J.L. (2000). ATP-Dependent Chromatin-Remodeling Complexes. *Mol. Cell. Biol.* *20*, 1899–1910.
- Waddington, C.H. (2012). The Epigenotype. *Int. J. Epidemiol.* *41*, 10–13.

- Waldmann, T., Izzo, A., Kamieniarz, K., Richter, F., Vogler, C., Sarg, B., Lindner, H., Young, N.L., Mittler, G., Garcia, B.A., et al. (2011). Methylation of H2AR29 is a novel repressive PRMT6 target. *Epigenetics Chromatin* 4, 11.
- Walport, L.J., Hopkinson, R.J., Chowdhury, R., Schiller, R., Ge, W., Kawamura, A., and Schofield, C.J. (2016). Arginine demethylation is catalysed by a subset of JmjC histone lysine demethylases. *Nat. Commun.* 7, 11974.
- Wals, K., and Ovaa, H. (2014). Unnatural amino acid incorporation in *E. coli*: current and future applications in the design of therapeutic proteins. *Front. Chem.* 2.
- Walter, T.S., Meier, C., Assenberg, R., Au, K.-F., Ren, J., Verma, A., Nettleship, J.E., Owens, R.J., Stuart, D.I., and Grimes, J.M. (2006). Lysine Methylation as a Routine Rescue Strategy for Protein Crystallization. *Structure* 14, 1617–1622.
- Wan, H., Huynh, T., Pang, S., Geng, J., Vaccaro, W., Poss, M.A., Trainor, G.L., Lorenzi, M.V., Gottardis, M., Jayaraman, L., et al. (2009). Benzo[d]imidazole inhibitors of Coactivator Associated Arginine Methyltransferase 1 (CARM1)—Hit to Lead studies. *Bioorg. Med. Chem. Lett.* 19, 5063–5066.
- Wang, Y. (2004). Human PAD4 Regulates Histone Arginine Methylation Levels via Demethylination. *Science* 306, 279–283.
- Wang, Y.-C., and Li, C. (2012). Evolutionarily conserved protein arginine methyltransferases in non-mammalian animal systems: Evolutionarily conserved PRMTs. *FEBS J.* 279, 932–945.
- Wang, C., Zhu, Y., Caceres, T.B., Liu, L., Peng, J., Wang, J., Chen, J., Chen, X., Zhang, Z., Zuo, X., et al. (2014). Structural Determinants for the Strict Monomethylation Activity by *Trypanosoma brucei* Protein Arginine Methyltransferase 7. *Structure* 22, 756–768.
- Wang, H., Huang, Z.Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B.D., Briggs, S.D., Allis, C.D., Wong, J., Tempst, P., et al. (2001). Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science* 293, 853–857.
- Wang, H., Straubinger, R.M., Aletta, J.M., Cao, J., Duan, X., Yu, H., and Qu, J. (2009a). Accurate localization and relative quantification of arginine methylation using nanoflow liquid chromatography coupled to electron transfer dissociation and Orbitrap mass spectrometry. *J. Am. Soc. Mass Spectrom.* 20, 507–519.
- Wang, L., Gural, A., Sun, X.-J., Zhao, X., Perna, F., Huang, G., Hatlen, M.A., Vu, L., Liu, F., Xu, H., et al. (2011). The Leukemogenicity of AML1-ETO Is Dependent on Site-Specific Lysine Acetylation. *Science* 333, 765–769.
- Wang, M., Hu, Y., and Stearns, M.E. (2009b). RPS2: a novel therapeutic target in prostate cancer. *J. Exp. Clin. Cancer Res.* 28, 6.
- Wärnmark, A., Treuter, E., Wright, A.P.H., and Gustafsson, J.-A. (2003). Activation functions 1 and 2 of nuclear receptors: molecular strategies for transcriptional activation. *Mol. Endocrinol. Baltim. Md* 17, 1901–1909.
- Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinforma. Oxf. Engl.* 25, 1189–1191.

- Wedemeyer, W.J., Welker, E., and Scheraga, H.A. (2002). Proline Cis–Trans Isomerization and Protein Folding. *Biochemistry (Mosc.)* 41, 14637–14644.
- Weerakoon, K.G.A.D., Gobert, G.N., Cai, P., and McManus, D.P. (2015). Advances in the Diagnosis of Human Schistosomiasis. *Clin. Microbiol. Rev.* 28, 939–967.
- Wilczek, C., Chitta, R., Woo, E., Shabanowitz, J., Chait, B.T., Hunt, D.F., and Shechter, D. (2011). Protein Arginine Methyltransferase Prmt5-Mep50 Methylates Histones H2A and H4 and the Histone Chaperone Nucleoplasmin in *Xenopus laevis* Eggs. *J. Biol. Chem.* 286, 42221–42231.
- Wilkins, S.E., Islam, M.S., Gannon, J.M., Markolovic, S., Hopkinson, R.J., Ge, W., Schofield, C.J., and Chowdhury, R. (2018). JMJD5 is a human arginyl C-3 hydroxylase. *Nat. Commun.* 9.
- Will, C.L., and Luhrmann, R. (2011). Spliceosome Structure and Function. *Cold Spring Harb. Perspect. Biol.* 3, a003707–a003707.
- Wu, X., and Zhang, Y. (2017). TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat. Rev. Genet.* 18, 517–534.
- Wu, J., Cui, N., Wang, R., Li, J., and Wong, J. (2012). A Role for CARM1-Mediated Histone H3 Arginine Methylation in Protecting Histone Acetylation by Releasing Corepressors from Chromatin. *PLoS ONE* 7, e34692.
- Wu, Q., Bruce, A.W., Jedrusik, A., Ellis, P.D., Andrews, R.M., Langford, C.F., Glover, D.M., and Zernicka-Goetz, M. (2009). CARM1 is Required in Embryonic Stem Cells to Maintain Pluripotency and Resist Differentiation. *Stem Cells Dayt. Ohio* 27, 2637–2645.
- Xie, B., Invernizzi, C.F., Richard, S., and Wainberg, M.A. (2007). Arginine Methylation of the Human Immunodeficiency Virus Type 1 Tat Protein by PRMT6 Negatively Affects Tat Interactions with both Cyclin T1 and the Tat Transactivation Region. *J. Virol.* 81, 4226–4234.
- Xie, Y., Zhou, R., Lian, F., Liu, Y., Chen, L., Shi, Z., Zhang, N., Zheng, M., Shen, B., Jiang, H., et al. (2014). Virtual screening and biological evaluation of novel small molecular inhibitors against protein arginine methyltransferase 1 (PRMT1). *Org Biomol Chem* 12, 9665–9673.
- Xu, W. (2001). A Transcriptional Switch Mediated by Cofactor Methylation. *Science* 294, 2507–2511.
- Xu, J., and Li, Q. (2003). Review of the *in Vivo* Functions of the p160 Steroid Receptor Coactivator Family. *Mol. Endocrinol.* 17, 1681–1692.
- Yadav, N., Lee, J., Kim, J., Shen, J., Hu, M.C.-T., Aldaz, C.M., and Bedford, M.T. (2003). Specific protein methylation defects and gene expression perturbations in coactivator-associated arginine methyltransferase 1-deficient mice. *Proc. Natl. Acad. Sci.* 100, 6464–6468.
- Yadav, N., Cheng, D., Richard, S., Morel, M., Iyer, V.R., Aldaz, C.M., and Bedford, M.T. (2008). CARM1 promotes adipocyte differentiation by coactivating PPAR γ . *EMBO Rep.* 9, 193–198.
- Yan, L., Yan, C., Qian, K., Su, H., Kofsky-Wofford, S.A., Lee, W.-C., Zhao, X., Ho, M.-C., Ivanov, I., and Zheng, Y.G. (2014). Diamidine Compounds for Selective Inhibition of Protein Arginine Methyltransferase 1. *J. Med. Chem.* 57, 2611–2622.
- Yang, N.J., and Hinner, M.J. (2015a). Getting Across the Cell Membrane: An Overview for Small Molecules, Peptides, and Proteins. In *Site-Specific Protein Labeling*, A. Gautier, and M.J. Hinner, eds. (New York, NY: Springer New York), pp. 29–53.

- Yang, Y., and Bedford, M.T. (2013). Protein arginine methyltransferases and cancer. *Nat. Rev. Cancer* 13, 37–50.
- Yang, F., Li, X., Sharma, M., Sasaki, C.Y., Longo, D.L., Lim, B., and Sun, Z. (2002). Linking β -Catenin to Androgen-signaling Pathway. *J. Biol. Chem.* 277, 11336–11344.
- Yang, H., Ouyang, Y., Ma, H., Cong, H., Zhuang, C., Lok, W.-T., Wang, Z., Zhu, X., Sun, Y., Hong, W., et al. (2017). Design and synthesis of novel PRMT1 inhibitors and investigation of their binding preferences using molecular modelling. *Bioorg. Med. Chem. Lett.* 27, 4635–4642.
- Yang, Y., Lu, Y., Espejo, A., Wu, J., Xu, W., Liang, S., and Bedford, M.T. (2010). TDRD3 Is an Effector Molecule for Arginine-Methylated Histone Marks. *Mol. Cell* 40, 1016–1023.
- Yang, Y., Hadjikyriacou, A., Xia, Z., Gayatri, S., Kim, D., Zurita-Lopez, C., Kelly, R., Guo, A., Li, W., Clarke, S.G., et al. (2015b). PRMT9 is a Type II methyltransferase that methylates the splicing factor SAP145. *Nat. Commun.* 6.
- Yao, R., Jiang, H., Ma, Y., Wang, L., Wang, L., Du, J., Hou, P., Gao, Y., Zhao, L., Wang, G., et al. (2014). PRMT7 Induces Epithelial-to-Mesenchymal Transition and Promotes Metastasis in Breast Cancer. *Cancer Res.* 74, 5656–5667.
- Yi, P., Wang, Z., Feng, Q., Chou, C.-K., Pintilie, G.D., Shen, H., Foulds, C.E., Fan, G., Serysheva, I., Ludtke, S.J., et al. (2017). Structural and Functional Impacts of ER Coactivator Sequential Recruitment. *Mol. Cell* 67, 733-743.e4.
- Yoshimoto, T., Boehm, M., Olive, M., Crook, M.F., San, H., Langenickel, T., and Nabel, E.G. (2006). The arginine methyltransferase PRMT2 binds RB and regulates E2F function. *Exp. Cell Res.* 312, 2040–2053.
- Yu, X.-R., Tang, Y., Wang, W.-J., Ji, S., Ma, S., Zhong, L., Zhang, C.-H., Yang, J., Wu, X.-A., Fu, Z.-Y., et al. (2015). Discovery and structure–activity analysis of 4-((5-nitropyrimidin-4-yl)amino)benzimidamide derivatives as novel protein arginine methyltransferase 1 (PRMT1) inhibitors. *Bioorg. Med. Chem. Lett.* 25, 5449–5453.
- Yue, W.W., Hassler, M., Roe, S.M., Thompson-Vale, V., and Pearl, L.H. (2007). Insights into histone code syntax from structural and biochemical studies of CARM1 methyltransferase. *EMBO J.* 26, 4402–4412.
- Yue, Y., Chu, Y., and Guo, H. (2015). Computational Study of Symmetric Methylation on Histone Arginine Catalyzed by Protein Arginine Methyltransferase PRMT5 through QM/MM MD and Free Energy Simulations. *Molecules* 20, 10032–10046.
- Yun, M., Wu, J., Workman, J.L., and Li, B. (2011). Readers of histone modifications. *Cell Res.* 21, 564–578.
- Zhang, K. (2016). Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* 193, 1–12.
- Zhang, J., and Zheng, Y.G. (2016). SAM/SAH Analogs as Versatile Tools for SAM-Dependent Methyltransferases. *ACS Chem. Biol.* 11, 583–597.
- Zhang, X., and Cheng, X. (2003). Structure of the Predominant Protein Arginine Methyltransferase PRMT1 and Analysis of Its Binding to Substrate Peptides. *Structure* 11, 509–520.
- Zhang, B., Dong, S., Zhu, R., Hu, C., Hou, J., Li, Y., Zhao, Q., Shao, X., Bu, Q., Li, H., et al. (2015). Targeting protein arginine methyltransferase 5 inhibits colorectal cancer growth by decreasing arginine methylation of eIF4E and FGFR3. *Oncotarget* 6, 22799–22811.

- Zhang, H., Guo, X., Feng, X., Wang, T., Hu, Z., Que, X., Tian, Q., Zhu, T., Guo, G., Huang, W., et al. (2017a). MiRNA-543 promotes osteosarcoma cell proliferation and glycolysis by partially suppressing PRMT9 and stabilizing HIF-1 α protein. *Oncotarget* 8.
- Zhang, R., Li, X., Liang, Z., Zhu, K., Lu, J., Kong, X., Ouyang, S., Li, L., Zheng, Y.G., and Luo, C. (2013). Theoretical Insights into Catalytic Mechanism of Protein Arginine Methyltransferase 1. *PLoS ONE* 8.
- Zhang, X., Zhou, L., and Cheng, X. (2000). Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *EMBO J.* 19, 3509–3519.
- Zhang, X., Ouyang, S., Kong, X., Liang, Z., Lu, J., Zhu, K., Zhao, D., Zheng, M., Jiang, H., Liu, X., et al. (2014). The Catalytic Mechanism of Histone Acetyltransferase p300: From the Proton Transfer to Acetylation Reaction. *J. Phys. Chem. B* 118, 2009–2019.
- Zhang, Z., Nikolai, B.C., Gates, L.A., Jung, S.Y., Siwak, E.B., He, B., Rice, A.P., O'Malley, B.W., and Feng, Q. (2017b). Crosstalk between histone modifications indicates that inhibition of arginine methyltransferase CARM1 activity reverses HIV latency. *Nucleic Acids Res.* 45, 9348–9360.
- Zhao, Q., Rank, G., Tan, Y.T., Li, H., Moritz, R.L., Simpson, R.J., Cerruti, L., Curtis, D.J., Patel, D.J., Allis, C.D., et al. (2009). PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. *Nat. Struct. Mol. Biol.* 16, 304–311.
- Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* 14, 331–332.
- Zhong, J., Cao, R.-X., Hong, T., Yang, J., Zu, X.-Y., Xiao, X.-H., Liu, J.-H., and Wen, G.-B. (2011). Identification and expression analysis of a novel transcript of the human PRMT2 gene resulted from alternative polyadenylation in breast cancer. *Gene* 487, 1–9.
- Zhou, B.-R., Jiang, J., Feng, H., Ghirlando, R., Xiao, T.S., and Bai, Y. (2015). Structural Mechanisms of Nucleosome Recognition by Linker Histones. *Mol. Cell* 59, 628–638.
- Zhou, Z., Sun, X., Zou, Z., Sun, L., Zhang, T., Guo, S., Wen, Y., Liu, L., Wang, Y., Qin, J., et al. (2010). PRMT5 regulates Golgi apparatus structure through methylation of the golgin GM130. *Cell Res.* 20, 1023–1033.
- Zondlo, N.J. (2013). Aromatic-Proline Interactions: Electronically Tunable CH/ π Interactions. *Acc. Chem. Res.* 46, 1039–1049.
- Zurita-Lopez, C.I., Sandberg, T., Kelly, R., and Clarke, S.G. (2012). Human Protein Arginine Methyltransferase 7 (PRMT7) Is a Type III Enzyme Forming ω -N^G-Monomethylated Arginine Residues. *J. Biol. Chem.* 287, 7859–7870.
- Zweygarth, E., Schillinger, D., Kaufmann, W., and Röttcher, D. (1986). Evaluation of sinefungin for the treatment of Trypanosoma (Nannomonas) congolense infections in goats. *Trop. Med. Parasitol. Off. Organ Dtsch. Tropenmedizinische Ges. Dtsch. Ges. Tech. Zusammenarbeit GTZ* 37, 255–257.

***Structural studies of protein arginine methyltransferase 2 reveal
its interactions with potential substrates and inhibitors***

Vincent Cura, Nils Marechal, Nathalie Troffer-Charlier, Jean-Marc Strub, Matthijs J. vanHaren,
Nathaniel I. Martin, Sarah Cianféroni, Luc Bonnefond and Jean Cavarelli

Structural studies of protein arginine methyltransferase 2 reveal its interactions with potential substrates and inhibitors

Vincent Cura¹, Nils Marechal¹, Nathalie Troffer-Charlier¹, Jean-Marc Strub², Matthijs J. van Haren³, Nathaniel I. Martin³, Sarah Cianférani², Luc Bonnefond¹ and Jean Cavarelli¹

¹ Department of Integrated Structural Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS UMR7104, INSERM U596, Université de Strasbourg, Illkirch, France

² BioOrganic Mass Spectrometry Laboratory (LSMBO), IPHC, UMR7178, Université de Strasbourg, France

³ Department of Chemical Biology and Drug Discovery, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, The Netherlands

Keywords

CARM1; epigenetics; PRMT2; PRMT4; RSF1

Correspondence

J. Cavarelli, Department of Integrated Structural Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), CNRS UMR7104, INSERM U596, Université de Strasbourg, 1 rue Laurent Fries, Illkirch F-67404, France
Fax: +33 (0)3 88 65 32 01
Tel: +33 (0)3 69 48 52 74
E-mail: cava@igbmc.fr

(Received 4 August 2016, revised 27 September 2016, accepted 2 November 2016)

doi:10.1111/febs.13953

PRMT2 is the less-characterized member of the protein arginine methyltransferase family in terms of structure, activity, and cellular functions. PRMT2 is a modular protein containing a catalytic Ado-Met-binding domain and unique Src homology 3 domain that binds proteins with proline-rich motifs. PRMT2 is involved in a variety of cellular processes and has diverse roles in transcriptional regulation through different mechanisms depending on its binding partners. PRMT2 has been demonstrated to have weak methyltransferase activity on a histone H4 substrate, but its optimal substrates have not yet been identified. To obtain insights into the function and activity of PRMT2, we solve several crystal structures of PRMT2 from two homologs (zebrafish and mouse) in complex with either the methylation product S-adenosyl-L-homocysteine or other compounds including the first synthetic PRMT2 inhibitor (Cp1) studied so far. We reveal that the N-terminal-containing SH3 module is disordered in the full-length crystal structures, and highlights idiosyncratic features of the PRMT2 active site. We identify a new nonhistone protein substrate belonging to the serine/arginine-rich protein family which interacts with PRMT2 and we characterize six methylation sites by mass spectrometry. To better understand structural basis for Cp1 binding, we also solve the structure of the complex PRMT4:Cp1. We compare the inhibitor–protein interactions occurring in the PRMT2 and PRMT4 complex crystal structures and show that this compound inhibits efficiently PRMT2. These results are a first step toward a better understanding of PRMT2 substrate recognition and may accelerate the development of structure-based drug design of PRMT2 inhibitors.

Database

All coordinates and structure factors have been deposited in the Protein Data Bank: zPRMT2₁₋₄₀₈-SFG = [5g02](#); zPRMT2₇₃₋₄₀₈-SAH = [5fub](#); mPRMT2₁₋₄₄₅-SAH = [5ful](#); mPRMT2₁₋₄₄₅-Cp1 = [5fwa](#), mCARM1₁₃₀₋₄₈₇-Cp1 = [5k8v](#).

Abbreviations

mPRMT2, mouse (*Mus musculus*) PRMT2; PDB, Protein Data Bank; PRMT, protein arginine methyltransferase; RMSD, root-square-mean deviation; SAH, S-adenosyl-L-homocysteine; SAM, S-adenosyl-L-methionine; SAXS, small-angle X-ray scattering; SFG, sinefungin; zPRMT2, zebrafish (*Danio rerio*) PRMT2.

Introduction

Protein arginine N-methyltransferases (PRMTs) are eukaryotic enzymes that catalyze the transfer of methyl groups from the methyl donor S-adenosyl-L-methionine (SAM) to guanidine nitrogens in arginine residues of target proteins. The PRMT family has been the focus of many biological researches and at least nine members of the PRMT family (PRMT1 to PRMT9) have been identified in mammalian cells [1–3]. PRMTs are divided into four major classes (type I to type IV) depending on the chemistry of methyl-arginine they generate. Type I and type II PRMT, the two most populated classes have been extensively studied. Type I PRMTs (PRMT1, PRMT2, PRMT3, PRMT4/CARM1, PRMT6, and PRMT8) produce asymmetric dimethylarginine, whereas type II (PRMT5, PRMT9) form symmetric dimethylarginine. PRMT7 produces monomethylarginine residues only, which defines it as the first type III PRMT. A fourth type which results in the monomethylation of the internal δ guanidino nitrogen has only been described in yeast [4]. Thanks to extensive research by many teams, several dozen crystal structures from seven different PRMTs have been determined either free or in complex with cofactor analogs [5–19]. Only PRMT2 and PRMT9 remain to be characterized at the structural level.

Human PRMT2 (or HRMT1L1) was first identified by homology to the yeast methyltransferase *HMT1* gene in 1998 [20]. PRMT2 contains a highly conserved catalytic SAM-binding domain and is characterized by an N-terminal extension partially folded into a Src homology 3 (SH3) domain. PRMT2 harbors a weak methyltransferase activity on histone H4 tail *in vitro*, but no optimal substrate has been found yet [21].

It has been shown that PRMT2 interacts with nuclear hormone receptors, including estrogen receptor alpha, retinoic acid receptors, and androgen receptors, and enhances their transcriptional activity in a ligand-dependent manner [22,23]. In addition to its role as a transcriptional coactivator, PRMT2 has also been found to be involved in diverse cellular processes, through its interaction with partners such as the adenoviral early region 1B-associated protein (E1B-AP5) [24], the retinoblastoma protein [25], β -catenin [26], or the transcription factor STAT3 [27].

Implication of PRMT2 in breast carcinogenesis in correlation with nuclear receptors expression has been reported [28,29]. Loss of PRMT2 nuclear expression in breast cancer cells is linked to increased cyclin D1 expression, thus promoting breast tumor cell proliferation. In addition to the original full-length PRMT2, four alternatively spliced PRMT2 isoforms

(PRMT2L2, PRMT2 α , β , and γ) have been identified. Together, all these isoforms showed increased expression in breast tumor samples [30,31].

So far no inhibitor of PRMT2 has been studied. Because of the lack of structural data and the weak enzyme activity, very few information is available to understand how this protein recognizes its substrates.

In the present study, we report the crystal structures of the PRMT2 catalytic module from zebrafish (*Danio rerio*) and from mouse (*Mus musculus*) in complex with either the methylation product S-adenosyl-L-homocysteine (SAH) or other compounds including a potent CARM1 inhibitor. We identify a new nonhistone protein substrate belonging to the serine-/arginine-rich protein family which interacts with PRMT2. We compare the inhibitor–protein interactions occurring in the PRMT2 and CARM1 complex structures and show that this compound inhibits efficiently PRMT2.

Results

PRMT2 from zebrafish

PRMT2 from zebrafish (*D. rerio*) was expressed in insect cells using the baculovirus system, purified and crystallized. The first crystals were obtained from a full-length construct (zPRMT2₁₋₄₀₈) in the presence of sinefungin, a natural inhibitor. The structure was solved and refined at 2.45 Å resolution (Table 1). The electron density map revealed the absence of residues 1–71 corresponding to the N-terminal part containing the SH3 domain. A SDS/PAGE analysis performed on dissolved crystals showed that the crystallized product was shorter than the purified protein, indicating that proteolysis occurred in the crystallization drop (Fig. 1A). Mass spectrometry LC-MS/MS analysis was performed on a trypsin digest of the gel-extracted protein band to localize the cleavage site. Sequence resulting from tryptic peptide identification revealed that the crystallized molecule started at residue 68, four residues before the catalytic module visible in the zPRMT2 crystal structure. zPRMT2 N-terminal degradation was potentially caused by a bacterial contamination of the sinefungin solution used for crystallization. Subsequent attempts to crystallize the full-length zPRMT2 with a fresh sinefungin solution were successful but the resulting crystals diffracted only at low resolution (5 Å). To improve the initial crystals obtained from a degraded protein, a new construct zPRMT2₇₃₋₄₀₈ restricted to the catalytic domain was cloned. The protein was expressed, purified and crystallized in the presence of SAH. The structure was

Table 1. X-ray data collection and refinement statistics. Statistics for the highest resolution shell are shown in parentheses.

	zPRMT2 ₁₋₄₀₈	zPRMT2 ₇₃₋₄₀₈	mPRMT2 ₁₋₄₄₅	mPRMT2 ₁₋₄₄₅	mCARM1 ₁₃₀₋₄₈₇
Ligand	Sinefungin	SAH	SAH	Compound 1	Compound 1
PDB ID	5g02	5fub	5ful	5fwa	5k8v
Data processing					
Wavelength (Å)	0.979	0.980	0.873	0.979	0.980
Resolution range (Å)	45.3–2.45 (2.55–2.45)	57.0–2.0 (2.05–2.00)	43.3–1.9 (1.93–1.89)	44.2–1.8 (1.84–1.8)	47.7–2.25 (2.30–2.25)
Space group	H32	H32	C222 ₁	C222 ₁	P2 ₁ 2 ₁ 2
Unit cell (Å, °)	148.0 148.0 127.7 90 90 120	147.1 147.1 127.4 90 90 120	65.5 115.4 133.0 90 90 90	66.2 114.7 132.7 90 90 90	75.2 98.1 206.3 90 90 90
Total reflections	128 263 (13 045)	191 435 (14 027)	371 882 (21 162)	308 208 (18 911)	474 150 (28 897)
Unique reflections	19 847 (2193)	35 929 (2639)	40 533 (3933)	47 105 (2780)	73 068 (4372)
Multiplicity	6.5 (5.9)	5.3 (5.3)	9.2 (8.5)	6.5 (6.8)	6.5 (6.6)
Completeness (%)	99.8 (98.7)	99.9 (100.0)	99.8 (97.3)	100 (100)	99.8 (97.6)
//σ(<i>l</i>)	11.8 (1.1)	9.2 (1.0)	15.4 (1.2)	11.7 (1.0)	5.7 (1.1)
Resolution limit for //σ(<i>l</i>) > 2	2.63	2.14	1.99	1.90	2.52
Wilson B-factor (Å ²)	56.9	35.4	30.1	29.9	29.6
<i>R</i> _{meas} ^a	0.087 (1.315)	0.114 (1.904)	0.096 (2.025)	0.097 (1.967)	0.307 (2.314)
CC _{1/2} ^b	0.999 (0.698)	0.998 (0.308)	0.999 (0.448)	0.999 (0.470)	0.986 (0.349)
Refinement					
Resolution range (Å)	45.3–2.45 (2.54–2.45)	38.4–2.0 (2.07–2.0)	35.1–1.9 (1.96–1.89)	35.0–1.8 (1.86–1.80)	47.7–2.2 (2.28–2.25)
% <i>R</i> -work	19.96 (32.73)	18.17 (30.59)	16.50 (27.80)	17.17 (30.52)	19.24 (30.02)
% <i>R</i> -free	24.20 (39.38)	21.75 (31.29)	19.04 (28.43)	19.07 (34.97)	23.48 (32.30)
Number of non-H atoms	2778	2934	3088	3078	11 755
Protein	2728	2722	2749	2723	10 989
Ligands	37	31	24	51	170
Water	13	181	315	304	596
Validation					
RMS (bonds)	0.002	0.004	0.018	0.004	0.002
RMS (angles)	0.68	0.61	1.50	0.86	0.477
Ramachandran favored (%)	96	98	97	98	96.6
Ramachandran outliers (%)	0	0	0	0.3	0
Average B-factor (Å ²)	76.92	45.60	36.28	35.55	41.0
Rotamer outliers (%)	1.0	0	0.33	0	0.5

^a*R*_{meas} is the redundancy-independent merging.

^bSee [73] for definitions. The resolution limits for //σ(*l*) > 2 are reported.

refined to 2.0 Å resolution and was almost identical to the one obtained after proteolysis (RMSD = 0.27 Å). A small difference occurs at the C-terminal extremity which is rotated by 180° after M406, changing the orientation of the two last residues.

PRMT2 from mouse

Full-length PRMT2 from mouse (*M. musculus*), mPRMT2₁₋₄₄₅, was expressed in insect cells using the baculovirus system, purified and crystallized. The structure was solved and refined at 2.0 Å resolution (Table 1). The visible part of the structure is limited to the catalytic module (residues 107–445) as for

zPRMT2. The molecular mass of the purified mPRMT2 measured by mass spectrometry under denaturing conditions is 50 728 Da for a calculated mass of 50 726 Da for the complete protein. Furthermore, a SDS/PAGE performed on a dissolved mPRMT2 crystal revealed no detectable size difference between the dissolved crystal and the purified protein (Fig. 1B). Therefore, unlike zPRMT2, the absence of residues 1–106 in the electron density map was not the result of proteolysis before or during the crystallization process, but could be explained by a high flexibility of the N-terminal region within the crystals.

Matthews coefficient and solvent content calculated for one full-length monomer per asymmetric unit

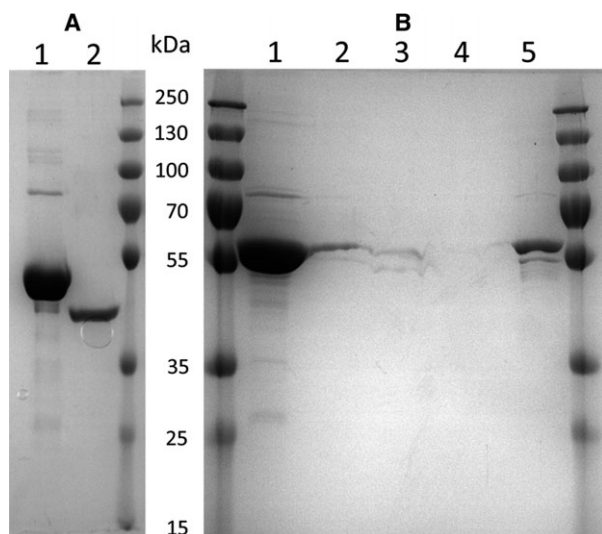


Fig. 1. PAGE analysis of purified PRMT2 and crystals. Proteins were visualized by Coomassie blue staining. For crystals analysis, each crystal was washed three times in 5 μ L of reservoir solution and dissolved in SDS/PAGE loading buffer. (A) Lane 1: Purified zPRMT2 (calculated MW 46 651 Da), lane 2: dissolved zPRMT2 crystal. (B) Lane 1: Purified mPRMT2 (calculated MW 50 726 Da); lanes 2, 3, and 4: crystal wash solutions. Lane 5: dissolved mPRMT2 crystal.

($V_m = 2.48 \text{ \AA}^3 \cdot \text{Da}^{-1}$, $\sim 50\%$ solvent content) indicates that the unit cell could easily accommodate SH3 domain. In the crystal packing, no steric hindrance could prevent the positioning of SH3 domain in the vicinity of residue D107.

Overall structure

Phylogenetic analysis [32] and sequence comparisons have established that PRMT2 methylation module is closely related to all type I PRMTs (35% to 39% sequence identity between PRMT8, PRMT1, PRMT3, PRMT6, and CARM1 from mouse). As expected, the monomeric structure of PRMT2 catalytic module is very similar to that of all the PRMT known structures especially type I PRMTs. It consists in a SAM-binding domain (residues 107–254, mPRMT2 numbering) adopting a Rossmann fold, a helical dimerization arm (residues 266–298), interrupting the β -barrel domain (residues 255–265 and residues 299–445) (Fig. 2). The two domains are connected by the strictly conserved cis-proline 254.

zPRMT2 and mPRMT2 share 55% sequence identity and up to 59% in their catalytic modules. The structure of the catalytic modules is highly similar and can be well superimposed (RMSD = 0.68 \AA) (Fig. 3). The major difference is the orientation of the loop

β 14– β 15 (426–437, 392–400 in zPRMT2). This loop is involved in the crystal packing in mPRMT2 (hydrogen bond between main-chain O L429 and N E205 from a neighbor monomer), whereas in zPRMT2, it is turned toward solvent and opens slightly the β barrel.

SAH-binding pocket

The active site pocket is strictly conserved between the mouse and zebrafish enzymes. The SAH and the sinefungin molecules are well defined in the electronic density. Residues involved in the SAH recognition are conserved among all PRMT2s (Fig. 2). Bifurcated hydrogen bonds involving residue E209 carboxylate and S237 O γ maintain the adenine amino group. N1 interacts with the V208 main-chain carbonyl and the E180 carboxylate forms two hydrogen bonds with the ribose hydroxyl oxygens (Fig. 4). For the homocystein moiety, the carboxylate group binds to R133, the amino group interacts with the C158 carbonyl and M127 makes van der Waals contacts with the S atom. Helix α X harboring the conserved YFxxY motif closes the SAM-binding pocket as observed in the structures of PRMT3 [6], CARM1 [8], and PRMT6 [18]. Helix α X interacts via Y118 and Y114 with catalytic E232 allowing the proper formation of the substrate arginine pocket required for catalysis (Fig. 3).

Sinefungin binds slightly differently from SAH because of a different orientation of the homocystein moiety. The homocystein amino group is too far to form a hydrogen bond but the sinefungin-specific NH2 ϵ is involved in hydrogen bonds with E189 carboxylate O ϵ 1 and carbonyl oxygens (Fig. 4A).

Substrate-binding pocket

The highly conserved active site glutamate residues E223 and E232 of the so-called double-E loop [7] form a pair of salt bridges with a calcium ion in mPRMT2 (Fig. 4C). Calcium was added to the crystallization drop in order to obtain diffracting crystals. In mPRMT2:SAH structure, this ion occupies the site where the guanidino moiety of the arginine substrate would bind. In zPRMT2, calcium is replaced by a water molecule. E198 of the double-E loop interacts with the ammonium of a MES molecule which fills the arginine substrate channel (Fig. 4A,B).

Dimerization

Homodimerization is a feature conserved in the PRMTs and essential for the catalytic activity (for a discussion and review see [16,18,33]). Exception is the

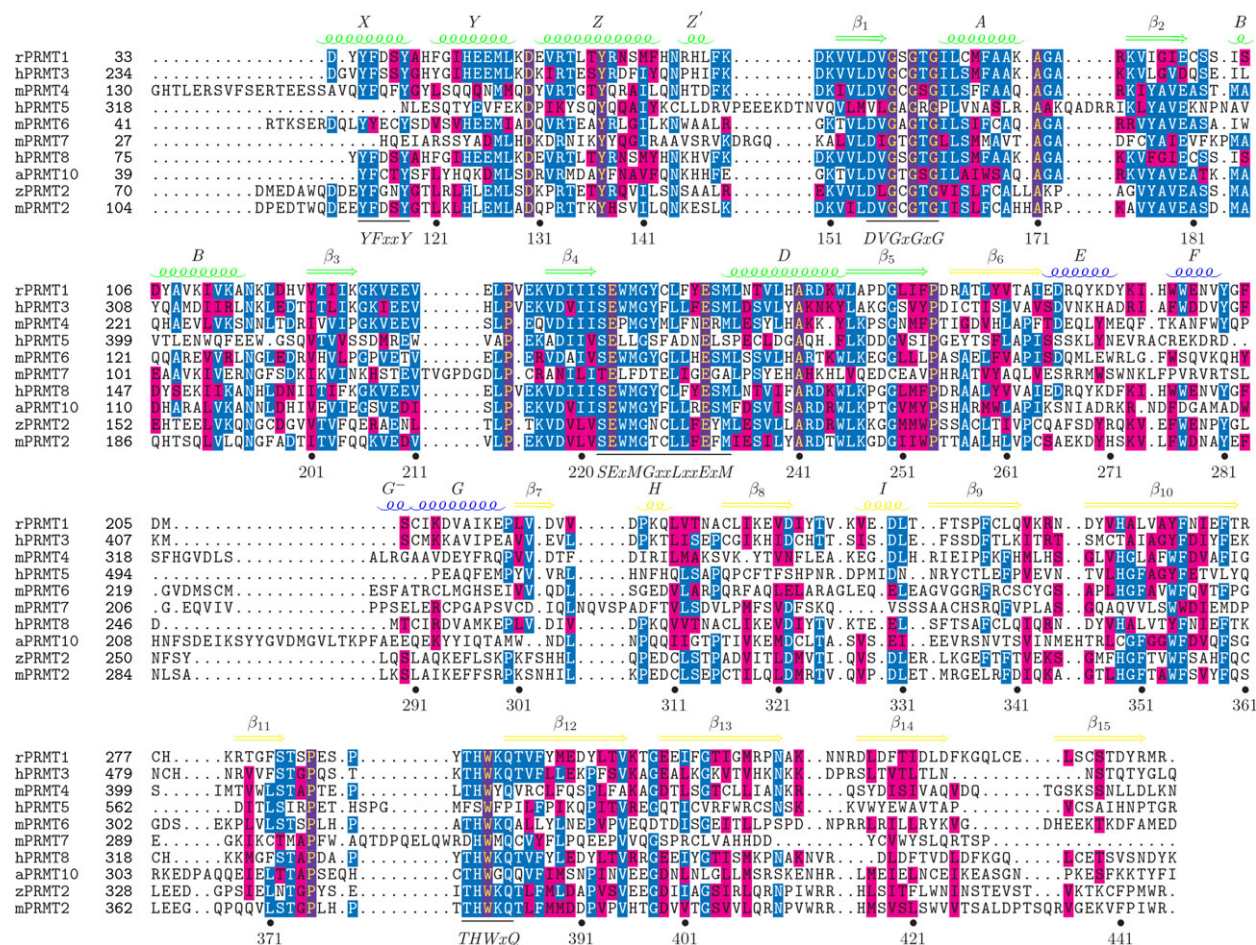


Fig. 2. Structure-sequence alignment of selected PRMTs. Ten PRMTs sequences have been aligned based on their crystal structures. The alignment is restricted to the catalytic core. The secondary structure of mPRMT2 is drawn above the alignment with color. The SAM-binding domain, the β -barrel domains, and the dimerization arm are colored green, yellow, and blue, respectively. The mPRMT2 residue numbering is shown below the sequences. The four signature sequences are localized and their consensus written below. Amino acids are shaded according to similarity to the consensus sequence. Amino acids highlighted are either invariant (violet) or similar (blue) as defined by the following grouping: F, Y, and W; I, L, M, and V; R and K; D and E; G and A; S, T, N, and Q. Abbreviations are as follows: m/*Mus musculus*, h/*Homo sapiens*, r/*Rattus norvegicus*, a/*Arabidopsis thaliana*, and z/zebrafish (*D. rerio*). This figure was produced with the program TEXSHADE [70].

monomeric PRMT7 which consist in two PRMT modules in tandem forming a pseudo dimer [18,34,35]. Higher order oligomers of some PRMTs, resulting of association of dimers, have been also found in solution both *in vivo* and *in vitro* [5,7,8,11,19,36–40]. zPRMT2 quaternary structure was studied using small-angle X-ray scattering. To ensure that zPRMT2₁₋₄₀₈ was not degraded, SDS/PAGE analysis was performed before SAXS experiments. Furthermore, the molecular mass of zPRMT2 monomer was determined from the scattering intensity values $I(0)$ using bovine serum albumin (BSA) as a reference. The resulting value of 47 630 Da corresponds to the size of zPRMT2₁₋₄₀₈ (theoretical mass of 46 845 Da), confirming the presence of the

SH3 domain in the protein sample. The scattering curves show that the full-length protein behaves as a dimer in solution according to the radius of gyration ($R_g = 3.3$ nm) (Table 2). The dimeric zPRMT2₇₃₋₄₀₈ exhibits a R_g of 2.9 nm identical to the R_g calculated from the X-ray structure (Fig. 5). However, the comparison of *ab initio* low-resolution models generated from zPRMT2 full-length and zPRMT2₇₃₋₄₀₈ scattering data did not reveal significant differences (data not shown). Therefore, it was not possible to localize SH3 domain on the whole structure using SAXS, probably because of the lack of rigidity of this N-terminal region. In both mPRMT2 and zPRMT2 crystal structures, the two monomers are related by a

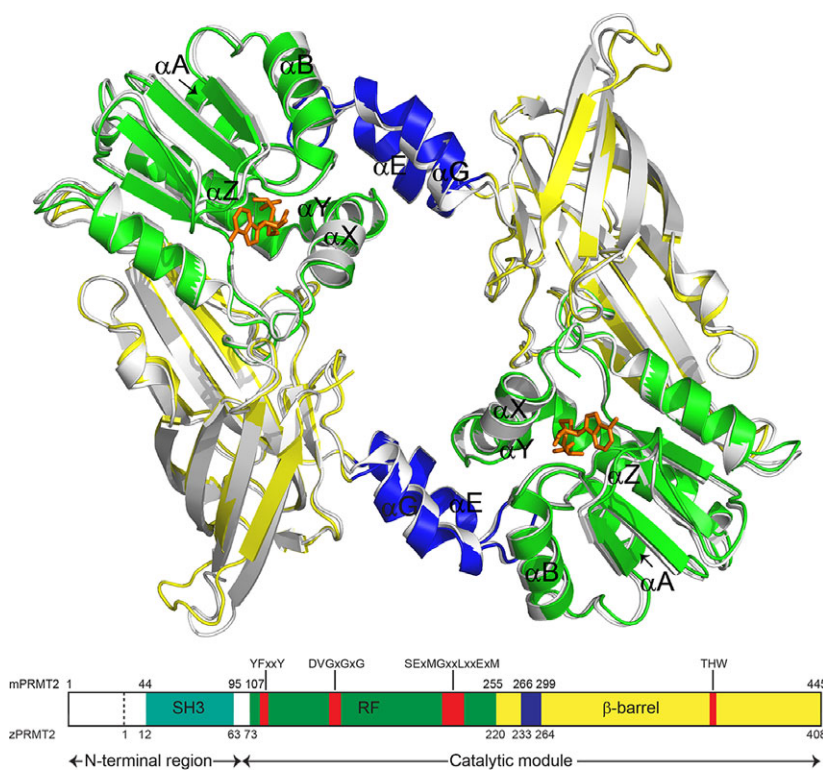


Fig. 3. Superimposed methylation modules of PRMT2 dimers from mouse (in gray) and from zebrafish (in color) and schematic representation illustrating the modular organization of full-length PRMT2. Rossmann fold is shown in green, the beta barrel in yellow, and the dimerization arm in blue. SAH is displayed in orange, as a stick model in the structure. On the scheme, SH3 domain is indicated in cyan and the four motifs in red.

crystallographic twofold axis and hence are structurally identical. The dimeric interface is stabilized by hydrophobic and hydrogen bonds and buries a surface area of 1602 and 1574 Å² for zPRMT2 and mPRMT2, respectively (calculated using PISA) [41]. The dimer formation involves the dimerization arm from one monomer, namely helices αE and αG and two short 3₁₀-helices F and G', and helices αY, αZ, αA, and αB from the other monomer and delineates a central hole of ~ 22 × 12 Å². The orientation of the monomer in the functional dimer observed in zPRMT2 is perfectly conserved in mPRMT2 (Fig. 3). Furthermore, sequence conservation in the dimerization arm suggests that this feature could be common to all PRMT2 orthologs (Fig. 2).

Activity

The *in vitro* methyltransferase activity of recombinant PRMT2 on histones is described as being weak compared to PRMT1, CARM1, and PRMT6, suggesting that the optimal substrate for PRMT2 is probably unknown. A small, but significant, activity has been previously described using the histone H4 as the substrate [21]. *In vivo* evidence has shown that PRMT2 associated with β-catenin deposits an asymmetric dimethyl mark on H3R8 of target gene promoters [23].

Synthetic peptides corresponding to the N-terminal tails of histone H3 (residues 1–34) and H4 (residues 2–22) were used as substrates to test the activity of the purified recombinant mPRMT2. The resulting gels showed incorporation of radioactivity in four different species: weak signal was detected for either H4_{2–22} or H3_{1–34} as expected; a much stronger signal was attributed to PRMT2 likely as a result of automethylation, whereas an additional strong signal was revealed for a minor contaminant with an estimated mass of 16 kDa (Fig. 6). The radioactive signals were all linearly dependent to the concentration in enzyme, confirming that the activity observed is due to PRMT2.

The 16 kDa contaminant copurifying with mPRMT2 was identified by LC-MS/MS mass spectrometry as the repressor splicing factor 1 (RSF1) from the insect expression host *Spodoptera frugiperda*. This splicing repressor antagonizes serine- and arginine-rich (SR) protein function [42] or coregulates alternative splicing with the other SR proteins in drosophila [43]. It has a modular organization with an N-terminal RRM-SRSF3-like RNA-binding domain and a long arginine-/glycine-rich C-terminal part. RSF1 is related to the serine-/arginine-rich (SR) family of splicing regulators, especially with the serine-/arginine-rich splicing factor 7 (SRSF7) involved in pre-mRNA splicing and mRNA export (51% sequence identity in the RRM domain). RSF1 from *S. frugiperda* was expressed

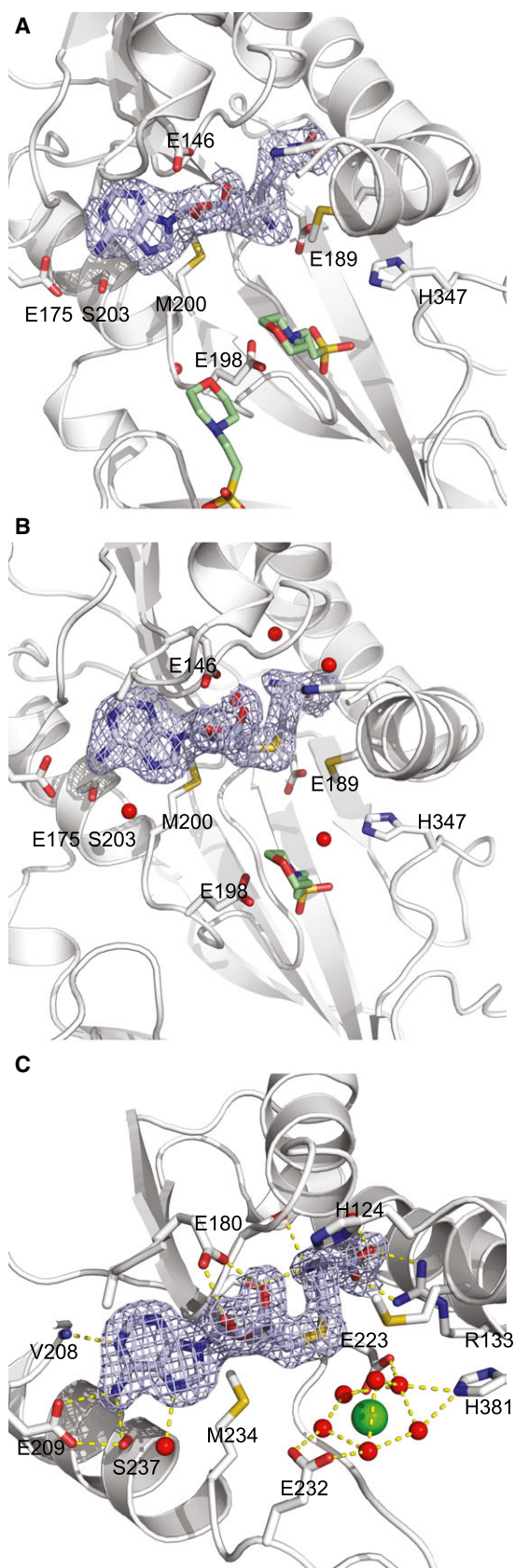


Fig. 4. A detailed view of the amino acid interactions in the active site of zPRMT2 in complex with sinefungin (A), with SAH (B), and mPRMT2 with SAH (C). For clarity, helix X has been removed. SAH or sinefungin are drawn as ball and stick models in light blue. Electron density ($2F_{obs}-F_{calc}$)-weighted maps around sinefungin and SAH are represented as a mesh, contouring level was contoured at 1σ . MES molecules and calcium ion are shown in green. Water molecules are drawn as red spheres. Hydrogen bonds are shown as yellow dotted lines.

in *Escherichia coli* and purified. To localize the methylation sites, recombinant RSF1 was incubated with SAM and mPRMT2 and analyzed by mass spectrometry (Fig. 7). Six methylation sites were identified on RSF1, four arginine residues are monomethylated (position 24, 100, 111, and 139), and two are dimethylated (position 7 and 120). Interestingly, the two dimethylated arginines are both located in short motifs displaying a unique consensus sequence SxGGxRxY. R7 and R24 belong to the RRM domain of RSF1 (Fig. 8). According to a model based on the structure of SRSF7 (PDB ID [2hvx](#)), these residues would be located in an α helix and in a β strand, respectively, with their side chains exposed at the surface of the domain and therefore accessible to the methyltransferase. The four other methylation sites would be located in the unfolded arginine-rich C-terminal part of RSF1.

Effect of calcium on the activity

A calcium ion was found bound in the bottom of active site pocket of mPRMT2:SAH structure, at a position corresponding to the binding site of the charged guanidino group of the target arginine residue. The hydrated calcium atom is surrounded by the two highly conserved glutamate residues of the ‘double-E loop’ and other conserved residues of ‘THW’ loop. This binding site is therefore conserved in all arginine methyltransferases. A calcium atom has also been observed at the same position in PRMT6 [18] and could explain the reduced activity of PRMT7 in the presence of magnesium or calcium [44]. To investigate the effect of calcium on PRMT2, the enzyme was incubated with increasing concentrations of calcium chloride and its activity was quantified (Fig. 9A). These experiments revealed that rather high calcium concentrations are necessary to inhibit the enzyme, as 25% of the initial activity remains in the presence of 10 mM CaCl_2 . At higher concentrations, the solubility of the enzyme was affected, causing consequently an artifactual drop of activity. These results suggest that calcium is a relatively weak inhibitor of PRMT2. The metal chelator EDTA increased slightly enzyme activity with respect to the positive control, indicating a possible contamination of purified mPRMT2 by divalent cations.

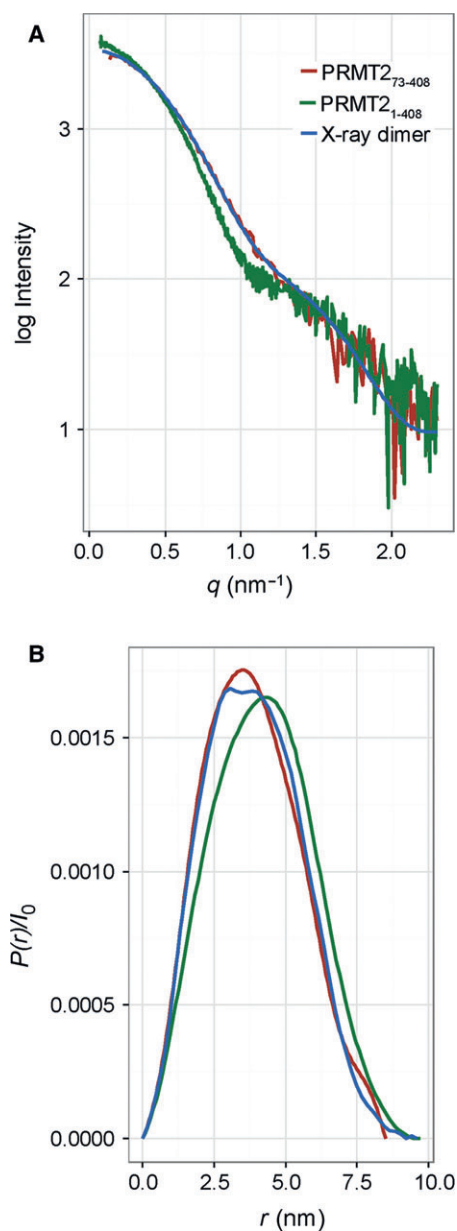


Fig. 5. SAXS data. (A) Comparison of scattering curves and (B) distance distributions $P(r)/I_0$ for zPRMT2₁₋₄₀₈, zPRMT2₇₃₋₄₀₈, and crystallographic structure of zPRMT2₇₃₋₄₀₈ (X-ray dimer) in the presence of SAH.

SH3 domain truncation

PRMT2 is characterized by its N-terminal extension partially folded into a SH3 domain. To check whether this module is involved in PRMT2 methyltransferase function, enzymatic activities of zPRMT2₇₃₋₄₀₈ and zPRMT2₁₋₄₀₈ were compared using RSF1 as substrate (Fig. 9B). A sevenfold decrease in activity was detected for the truncated construct with respect to the full-

Table 2. Small-angle X-ray scattering data. Radius of gyration (R_g) and maximum diameter of the particle (D_{max}) values are indicated for zPRMT2 full-length and catalytic module in the presence of SAH.

	zPRMT2 ₁₋₄₀₈ + SAH	zPRMT2 ₇₃₋₄₀₈ + SAH
Protein concentration	2.2 mg·mL ⁻¹	2.4 mg·mL ⁻¹
R_g (nm) from $P(r)$	3.22 ± 0.005	2.97 ± 0.02
R_g (nm) from Guinier	3.25 ± 0.01	2.90 ± 0.04
D_{max} (nm)	9.00 ± 0.05	8.50 ± 0.05

length enzyme. These results reveal that PRMT2 lacking SH3 domain exhibits activity *in vitro* but less than the full-length enzyme, showing the importance of the N-terminal region for the enzymatic function.

Inhibitor binding

A set of SAM-based compounds have been recently described as potent inhibitors of PRMTs [45]. These molecules were designed to simultaneously occupy both the adenosine and the peptide substrate-binding sites. Among them, compound 1 (Cp1) has been reported as an inhibitor for PRMT1, PRMT6, and PRMT4/CARM1 with the highest selectivity and potency for CARM1 vs. PRMT1 or PRMT6. It was therefore interesting to evaluate the inhibitor potency of Cp1 on PRMT2 and to compare it to the effects observed on related PRMTs.

Cp1 binding on mPRMT2 was first compared to SAH and sinefungin by thermal shift assay experiments (Fig. 10A). Binding of SAH or sinefungin increased the melting temperature by 5.3 °C ± 1.2 °C and 5.0 °C ± 1.4 °C with respect to the apo protein. Surprisingly, when Cp1 was added to mPRMT2, the melting temperature difference reached 15 °C ± 1.9 °C, suggesting a better affinity of PRMT2 for Cp1 than for SAH or sinefungin.

To further investigate the affinity of Cp1 toward PRMT2, we determined IC₅₀ values using RSF1 as a substrate (Fig. 10B). We obtained similar IC₅₀ for Cp1 (16.3 ± 3.8 μM) and for SAH (18.3 ± 2.0 μM). These values are much higher than the ones that we obtained for mouse CARM1 (mCARM1) (460 ± 134 nM for Cp1 and 820 ± 35 nM for SAH) which are in good agreement with the published data [45]. The IC₅₀ values obtained for PRMT2 with Cp1 are in the order of magnitude of those published for PRMT6 (20.23 ± 8.67 μM) and PRMT1 (11.09 ± 2.77 μM). However, the high potency and the selectivity of Cp1 toward CARM1 were not observed for PRMT2. To better understand structural basis for Cp1 binding, we crystallized and solved the structure of the complexes

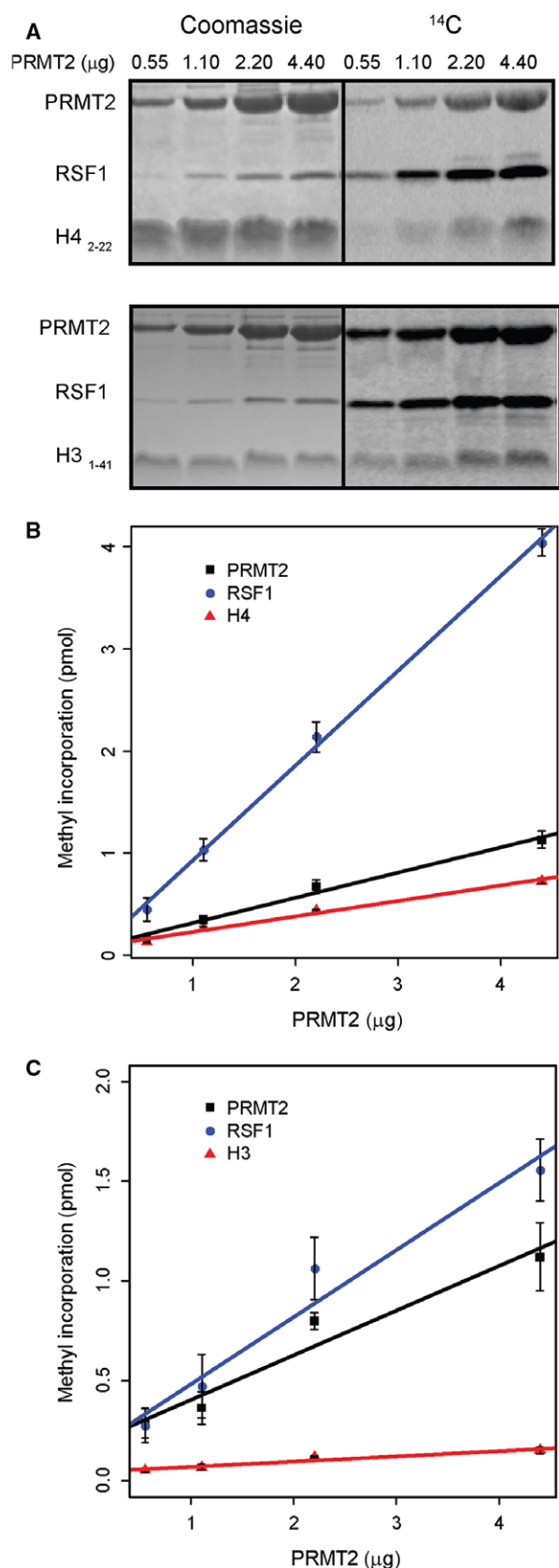


Fig. 6. Methylated product analyses by gel-based assays: (A) Increasing concentrations of mPRMT2 were incubated with H4₂₋₂₂ peptide (upper panel) and H3₁₋₄₁ peptide (lower panel). RSF1 band corresponds to the 16 kDa mPRMT2 contaminant. (B) Curves showing the methylation of the different species with respect to PRMT2 concentration in the presence of H4₂₋₂₂ or (C) H3₁₋₄₁ peptide. The error bars represent SD from experiments run in triplicate.

mPRMT2:Cp1 at 1.8 Å resolution and also mCARM1:Cp1 at 2.25 Å resolution (Table 1).

Structure of mPRMT2:Cp1

Cp1 lacks the methionine moiety typical of all the SAM analogs but harbors a guanidine connected with a spacer to the adenosine moiety (Fig. 11). Structure comparison revealed that the hydrogen bond network observed between the key residues of PRMT2 and the adenosine part of SAH is also well conserved for Cp1 despite a slight conformational change in the ribose cycle. The guanidinium group of Cp1 enters into the arginine pocket where it is maintained by H-bonds and salt bridges between the carboxylates of the motif III residues E223 and E232. The ligand-free methionine pocket is filled by several water molecules. Additional H-bonds occur with the M125 main-chain oxygen and with a water molecule mediating an indirect interaction with the H381 side chain. Thermal shift assays revealed that Cp1 binds more strongly to mPRMT2 than SAH. This result can be explained by the different sets of interactions occurring in the two distinct pockets of the active site, i.e., the methionine pocket for SAH or the arginine pocket for Cp1.

CARM1-PRMT2 Cp1 comparison

Comparing mPRMT2 and mCARM1 structures solved with Cp1 reveals a highly similar positioning of the inhibitor in both the active site pockets. Almost all the key residues involved in Cp1 binding are conserved and their side chains are interacting identically with the inhibitor in both structures (Fig. 11C). However, two differences appear. First, one additional helix W, specific to CARM1 (residues 136–141), packs against the adenine group of Cp1 shielding it from the solvent. This helix, which does not exist in PRMT2, enables an indirect interaction between the guanidino group of residue R141 from CARM1 and N7 from Cp1 via one water molecule. Additionally, F138 forms a hydrophobic interaction with F151, closing the pocket on the adenine ring. Second, in mPRMT2, the histidine residue H124 can form a hydrogen bond with the O3'

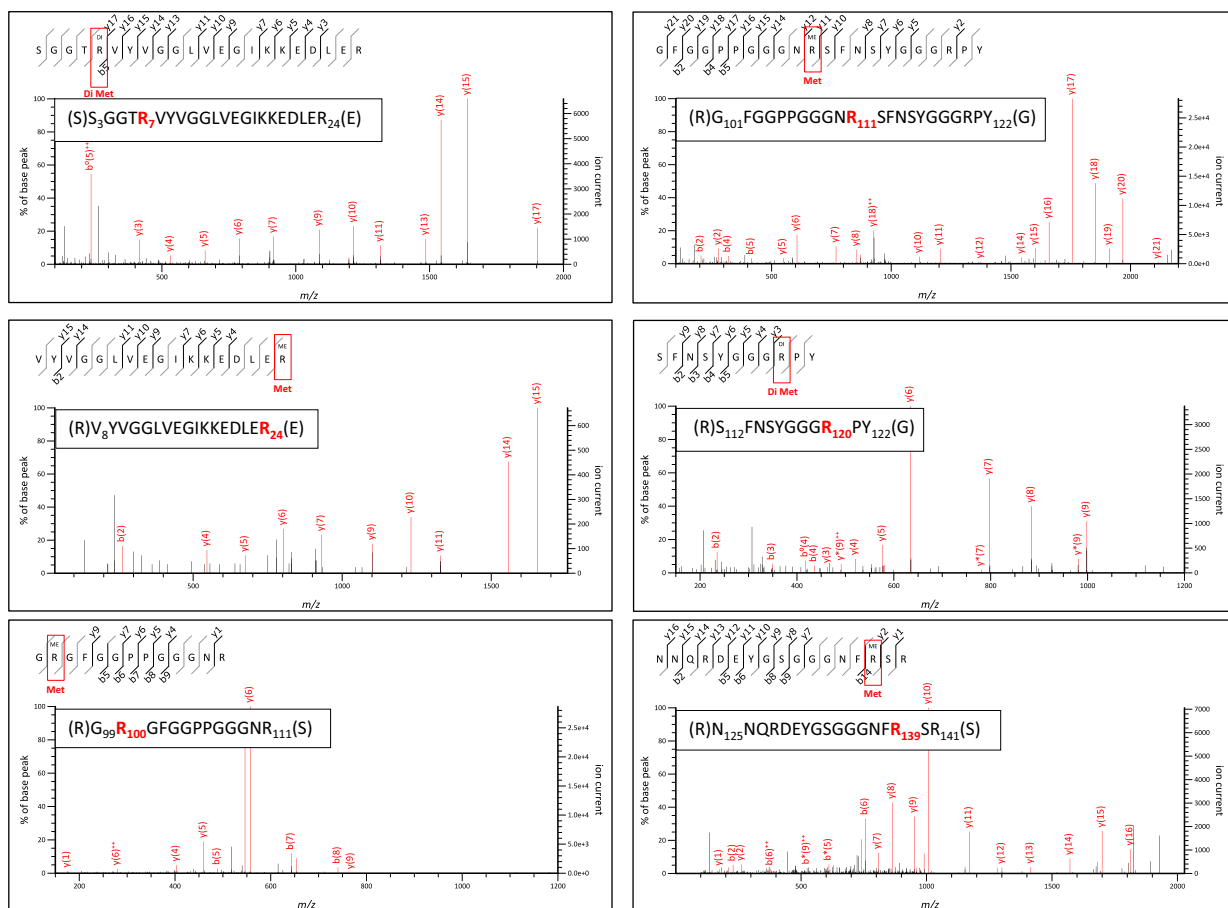


Fig. 7. Analysis of RSF1 methylation by mass spectrometry. See main text for details.

ribose hydroxyl group of Cp1. This residue is replaced by a glutamine (E160) in mCARM1 and the distance to Cp1 is thus too far to be compatible with a hydrogen bond. It is interesting to note that this histidine residue is conserved among almost all PRMT2 sequences known with a few exceptions in which cases it is changed into a glutamine. This histidine residue is also conserved in other PRMTs including PRMT1 and PRMT6. It is probable that these discrepancies impact the effect of Cp1 on the different PRMTs. The high potency and selectivity of Cp1 for CARM1 could be the result of its specific helix W which reinforces the interactions with the inhibitor.

Discussion

The crystal structures of PRMT2 from zebrafish and from mouse revealed a highly conserved dimeric architecture formed by two typical type I PRMT catalytic modules. The global shape of the functional dimer is a consequence of variations in relative orientations of the two monomers. It varies significantly in the

structures of different PRMTs and is linked to active site accessibility and therefore substrate recognition [10,16]. The N-terminal module of mPRMT2 (residues 1–106) partially folded in a SH3 domain is missing in the electron density although present in the crystal. A similar case has already been mentioned in mCARM1 [8], where the PH domain assumes different positions in the crystal structure, behaving as a wobbly domain. This behavior is the consequence of the lack of direct interaction between the N-terminal domain and the catalytic module as well as of the flexibility of the connection between functional modules. SH3 motifs are known to interact to proline-rich sequences in their binding partner, mediating the assembly of large multi-protein complexes. The wobbly SH3 domain would probably be stabilized in its functional position in the presence of partners or PRMT2 substrates.

It is well established that in PRMTs, the methyltransferase function is exclusively borne by the catalytic domain, containing both SAM- and arginine-binding sites. However, the presence of additional domains can strongly affect enzymatic activity, as

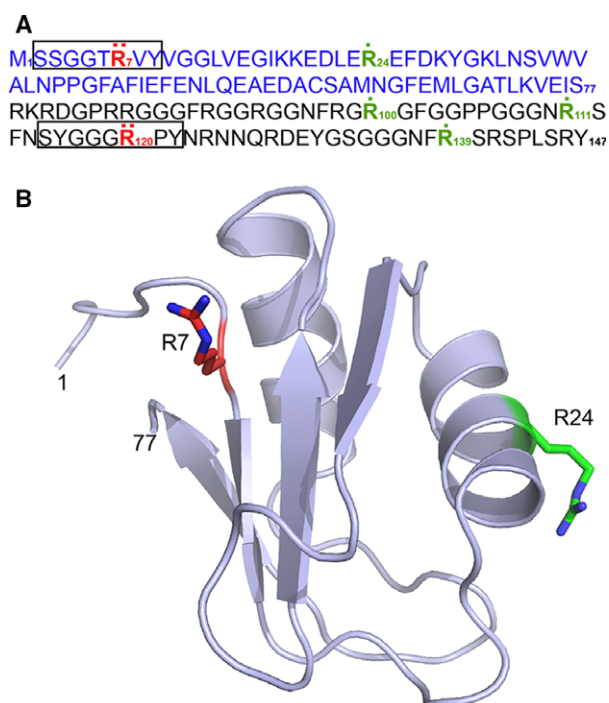


Fig. 8. Methylation sites on RSF1 sequence and model of the RRM domain. (A) Peptidic sequence of *S. frugiperda* RSF1. The folded part corresponding to the RRM domain is shown in blue. Mono- (') and dimethylated (') arginines are in bold. Consensus sequence SxGGxRxY is boxed. (B) 3D-Model of RSF1 RRM domain generated by Swiss-Model [71]. The side chains of the methylated arginine residues are shown in stick.

shown for PRMT5 [46] or PRMT8 [47]. We showed that it is also the case for PRMT2 as N-terminal-truncated enzyme displays reduced methyltransferase activity. Our structural results indicate that this truncation does not impair enzyme stability or dimerization. Therefore, the SH3 domain and/or adjacent sequences belonging to the N-terminal extension of the catalytic module may be also involved in substrate binding or recognition through long distance interactions.

We could identify RSF1, a host-cell contaminant coeluting with PRMT2 and furthermore methylated by the enzyme. Although RSF1 is not present in mouse and PRMT2 has no ortholog in insects, the interaction observed between an arginine methyl transferase and a protein involved in RNA splicing is not surprising. Indeed, arginine methylation is a post-translational modification that occurs in RNA-binding proteins including the spliceosome and SR proteins [48]. It has been shown that PRMT5 catalyzes sDMA formation on the Sm proteins [49] and CARM1 methylates three splicing factors: SmB, U1-C, and SAP49 [50]. Furthermore, methylation of SR proteins by PRMTs has been observed: SF2/ASF and its close paralog SRp30c both

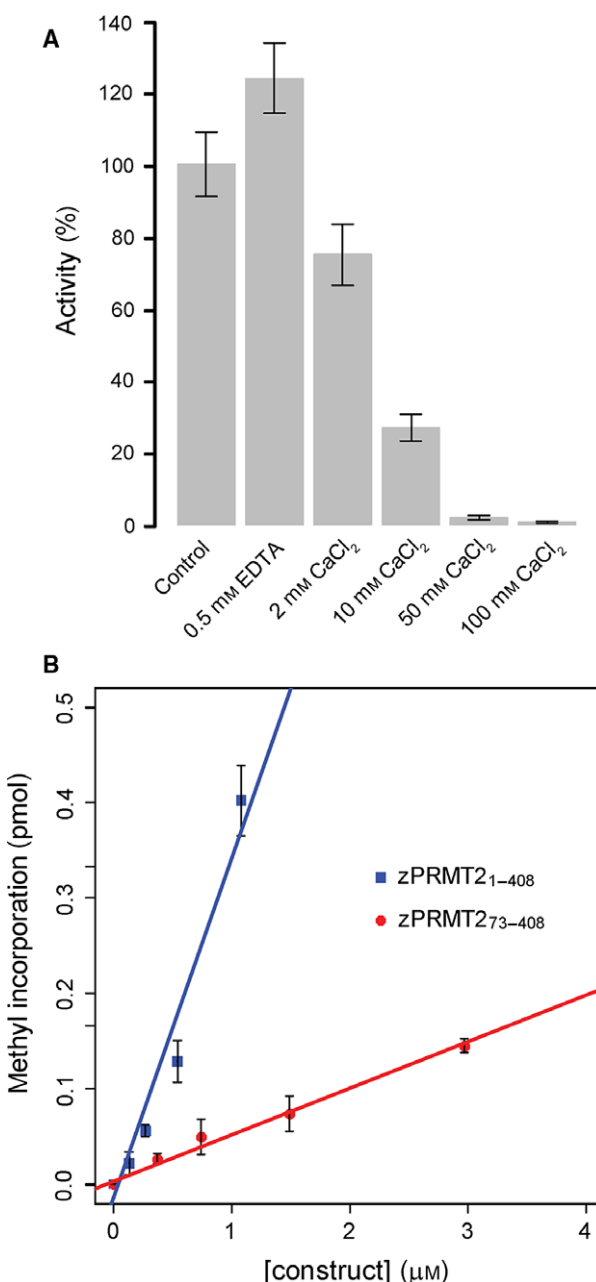


Fig. 9. Methyltransferase activity quantified by gel-based assays: (A) Effect of calcium on mPRMT2 activity. Relative activity was quantified using recombinant RSF1 as a substrate and 1 μM mPRMT2₁₋₄₄₅. The positive control was performed with no additive and used as a reference. (B) Effect of SH3 domain deletion on zPRMT2 activity. Curves comparing the methyltransferase activity of zPRMT2₁₋₄₀₈ (blue) and zPRMT2₇₃₋₄₀₈ (red). The error bars represent the SD of three sets of experiments.

belonging to the SR family are methylated by PRMT1 [51,52]. More recently, it has been shown that human PRMT9 symmetrically dimethylates arginine residues on splicing factor SF3B2 [35]. Taken together, these

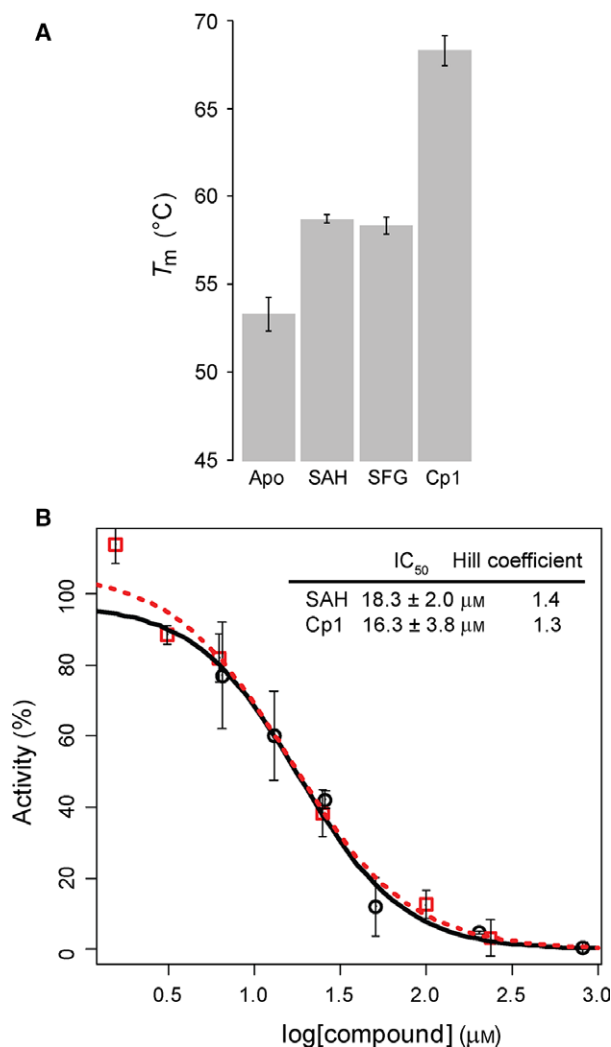


Fig. 10. Effect of inhibitor binding. (A) Melting temperatures (T_m) determined by thermal shift assay for mPRMT2 alone (apo) or incubated with 1 mM SAH, sinefungin (SFG), and Cp1. (B) Dose–response curves for SAH (circles, plain line) and Cp1 (square, red dotted line) at 1.2 μM of mPRMT2. The error bars represent the SD of three sets of experiments.

results and our initial findings suggest that substrates for PRMT2 could be found in the SR family. We tried to methylate *in vitro* mouse SRSF7 expressed and purified from insect cell culture (data not shown). This preliminary attempt was unsuccessful. Nevertheless, it is not excluded that other members of SR protein

family are PRMT2 methylation targets. So far, PRMT2 has not been mentioned to interact with splicing regulators or RNA-binding proteins but this possibility needs to be further investigated.

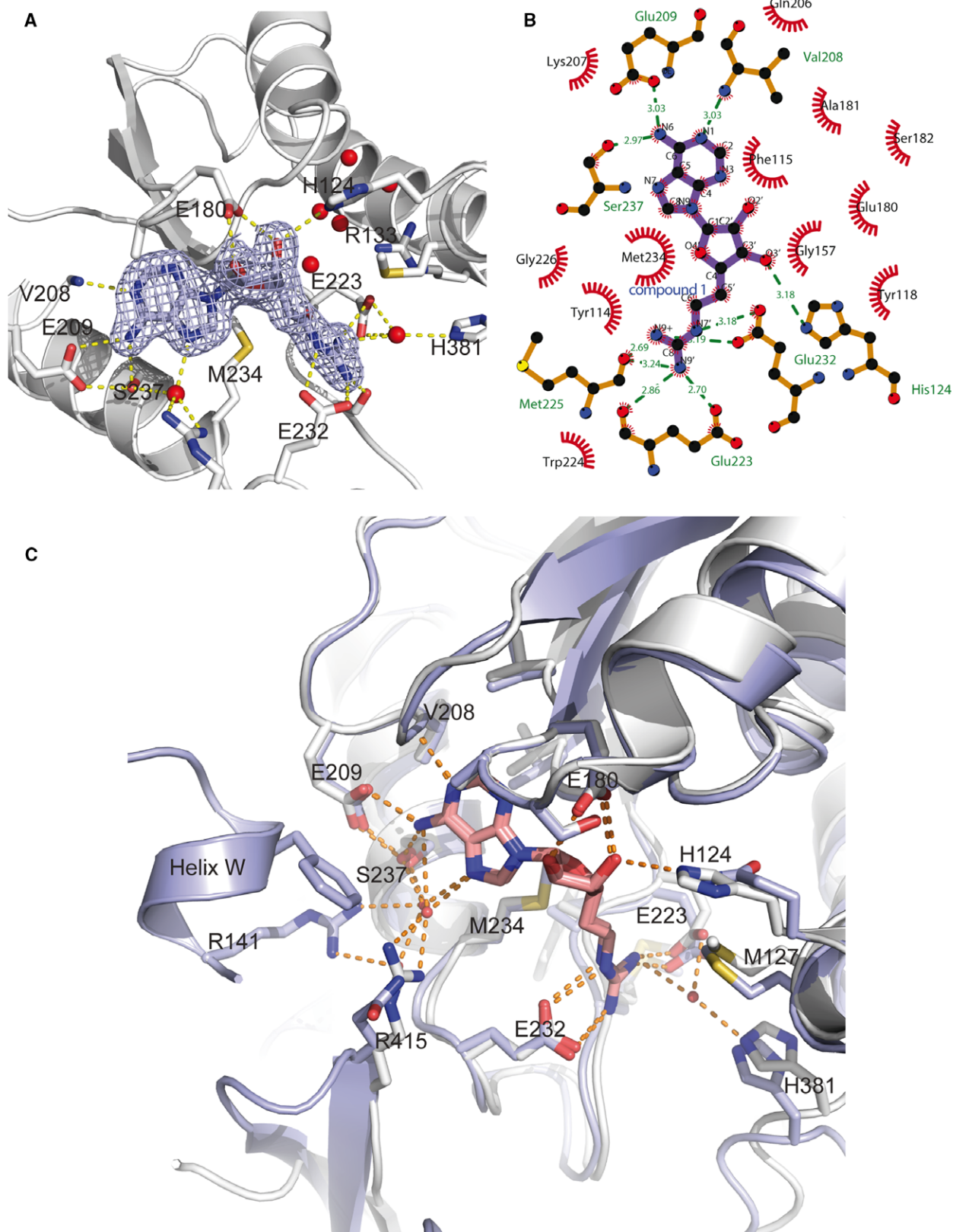
The crystal structure of mPRMT2 in complex with Cp1, a specific CARM1 inhibitor, has been solved and analyzed. The IC_{50} values obtained for Cp1 with mPRMT2 were higher than with CARM1, revealing a weaker affinity for mPRMT2. Nevertheless, Cp1 inhibits mPRMT2 slightly more efficiently than SAH and can be then considered as the first synthetic PRMT2 inhibitor studied so far. Altogether, these results are a first step toward a better understanding of PRMT2 substrate recognition and may accelerate the development of structure-based drug design of PRMT2 inhibitors.

Material and methods

PRMT2 cloning, expression, and purification

The *D. rerio* *prmt2* gene sequence (Genbank [BC129171.1](#)) corresponding to the full-length protein and to the PRMT core (residues 73–408) were amplified by PCR from a cDNA clone (Source BioScience, Nottingham, UK; IRAK293-C16). The *M. musculus* *prmt2* full-length gene sequence (Genbank [BC122563.1](#)) was amplified by PCR from a cDNA clone (Source BioScience IRCL32A04). The sequences were cloned in the pDONR207TM (Fisher Scientific, Illkirch, France) vector using a BP reaction (Gateway[®] Cloning; Thermo Fisher Scientific, Waltham, MA, USA). The positive clones were confirmed by sequencing (GATC BIOTECH, Konstanz, Germany). The sequences were subcloned in a pDEST20TM vector (for *D. rerio* *prmt2*) or a pBacGGWA vector [53] (for *M. musculus* *prmt2*) using a LR reaction. DH10Bac competent cells containing the baculovirus genome were transformed with the pDEST20TM-PRMT2 plasmids and plated onto LB agar media containing 15 mg·mL⁻¹ tetracycline, 7 mg·mL⁻¹ gentamicin, 50 mg·mL⁻¹ kanamycin, 25 mg·mL⁻¹ X-Gal, and 40 mg·mL⁻¹ isopropyl thio- β -d-galactoside (IPTG). Bacmid DNA purified from recombination-positive white colonies was transfected into Sf21 cells using the Lipofectin reagent (Fisher Scientific). Viruses were harvested 10 days after transfections. The resulting recombinant proteins harbor an amino-terminal glutathione S-transferase (GST) tag followed by a Tobacco etch virus (TEV) protease cleavage site.

Fig. 11. Overview of Cp1-binding site on mPRMT2 and mCARM1. (A) View of Cp1 interactions in mPRMT2 SAM- and arginine-binding pockets. Electron density (2Fobs-Fcalc)-weighted maps around Cp1, represented as a mesh, was contoured at 1 σ . (B) Flat representation done with LigPlot [72] showing contacts between Cp1 (compound 1) and active site residues in mPRMT2. The ligand bonds are colored in purple. Hydrogen bonds are shown as dotted lines. The spoked arcs represent protein residues making nonbonded contacts with Cp1. (C) Superimposition of the mPRMT2 (gray) and mCARM1 (purple) active sites. Cp1 is drawn in pink in both structures. Numbered residues belong to mPRMT2 except R141 belonging to mCARM1. Helix W specific to mCARM1 is also shown.



mPRMT2 also harbors a R445W point mutation due to a mutation in the PCR product. Sf21 cells were grown at 27 °C in suspension culture in SF900II (Thermo Fisher Scientific) using Erlenmeyer flasks. Two liter of Sf21 cell culture (at 0.8×10^6 cells·mL⁻¹) was infected with recombinant GST-zPRMT2 viruses with an infection multiplicity of 1. Cells were harvested 48 h postinfection. Cells lysis was performed by sonication in 100 mL buffer *A* [50 mM Tris/HCl pH 8.0, 250 mM NaCl, 5% glycerol, 10 mM DTT, 0.01% NP40, and anti-proteases (Roche, Basel, Switzerland, Complete™, EDTA-free)] and cellular debris were removed after centrifugation of the lysate at 40 000 g for 30 min. The supernatant was incubated at 4 °C for 4 h with 3 mL glutathione Sepharose 4B resin (GE Healthcare, Velizy, France). After a short centrifugation, the supernatants were discarded and the beads were poured in an Econo-column (Bio-Rad, Hercules, CA, USA). After two wash steps with 10 mL buffer *A*, 3 mL buffer *A* supplemented with in-house produced TEV protease were applied to the columns and digestion was performed overnight at 22 °C with gentle mixing. The digests were concentrated with an Amicon Ultra 10K (Merck Millipore, Darmstadt, Germany), loaded on a gel-filtration column (HiLoad 16/60 Superdex S200, GE Healthcare) and eluted at 1 mL·min⁻¹ with buffer *B* [20 mM Tris/HCl pH 8.0, 75 mM NaCl, 5 mM TCEP and 1 mM EDTA] for zPRMT2, using an ÄKTA Explorer device (GE Healthcare). Fractions containing zPRMT2 were pooled and concentrated to 17 mg·mL⁻¹. mPRMT2 and zPRMT2₇₃₋₄₀₈ were purified from baculovirus-infected cells using the same protocol with the following modifications: the crude extract was incubated with glutathione sepharose overnight at 4 °C, TEV digestion was performed during 4 h at 30 °C and buffer *C* was used for gel-filtration (20 mM Tris/HCl pH 8.0, 100 mM NaCl, and 5 mM TCEP). Fractions containing mPRMT2 were pooled and concentrated to 3.5 mg·mL⁻¹. Fractions containing zPRMT2₇₃₋₄₀₈ were pooled and concentrated to 14 mg·mL⁻¹.

PRMT2 crystallization

Crystallization conditions were screened using commercially or home-made available kits by the sitting-drop vapor-diffusion method in 96-well MRC2 plates (Swissci, Zug, Switzerland), and employing a Mosquito robot (TTP Labtech, Melbourn, UK) dispensing droplets made of 100 nL protein solution and 100 nL reservoir solution. The plates were placed in a Rock Imager (Formulatrix, Bedford, MA, USA) at 20 °C and monitored periodically.

The zPRMT2 crystal used for X-ray diffraction was obtained in a 96-well MRC2 plates in sitting drops made with a zPRMT2 sample at 17 mg·mL⁻¹, mixed with 1 mM sinefungin, and a reservoir solution containing 100 mM MES pH 6.0, 9% (w/v) PEG 20 000, and 100 mM LiCl. Single diffracting crystals grew to 120 × 100 × 20 μm within 4 weeks.

The zPRMT2₇₃₋₄₀₈ crystal used for X-ray diffraction was obtained in a 96-well MRC2 plates in sitting drops made with a zPRMT2 sample at 14 mg·mL⁻¹, mixed with 1 mM SAH, and a reservoir solution containing 100 mM MES pH 6.0, 9% (w/v) PEG 20 000, and 300 mM NaCl. Single diffracting crystals grew to 50 × 40 × 20 μm within 3 days.

First hits for mPRMT2 were obtained in several conditions of the PACT (Molecular Dimensions, Newmarket, UK) and The Peg Suite (Qiagen, Hilden, Germany) crystallization screens. Crystallization conditions were further optimized in 24-well VDX plates (Hampton Research, Aliso Viejo, CA, USA). Crystals used for X-ray diffraction were obtained within days at 20 °C in a hanging drop made with 1 μL mPRMT2 at 3.5 mg·mL⁻¹, mixed with 1 mM inhibitor, and 1 μL reservoir solution containing 16% (w/v) PEG 6000, 100 mM Hepes-NaOH pH 7.0 and 100 mM CaCl₂.

Data collection and structure solution

zPRMT2 crystals were soaked in the reservoir solution supplemented with 15–20% (v/v) PEG 400 or 20% (v/v) ethylene glycol and flash-cooled in liquid nitrogen.

The diffraction datasets were collected on the SOLEIL PROXIMA1 and PROXIMA2 beamlines, using a PILATUS 6M (DECTRIS, Baden-Dättwil, Switzerland) or a Quantum 315 (ADSC, Poway, CA, USA) detector and processed with XDS [54], Imosflm [55], and HKL-2000 [56]. The crystals belonged to the *H32* space group with one zPRMT2 molecule in the asymmetric unit. The dataset was submitted to the BALBES server for automated molecular replacement at York Structural Biology Laboratory [57] after unsuccessful manual attempts with known PRMT structures or models. An initial solution with a BALBES template based on CARM1 from *Homo sapiens* (PDB ID 2y1w) and the sequence of zPRMT2 was obtained with one monomer in the asymmetric unit.

mPRMT2 crystals were cryoprotected by a short soak in 20% (w/v) PEG 6000, 100 mM Hepes pH 7.0, 100 mM CaCl₂, 15% (v/v) PEG 400, and flash-cooled in liquid nitrogen for data collection. Diffraction datasets were collected on the ID23-2 beamline at the ESRF using a PILATUS 2M detector (DECTRIS) and processed with XDS. The crystal belonged to the *C222*₁ space group with one mPRMT2 molecule in the asymmetric unit. The structure was solved by molecular replacement using PHASER and zPRMT2 structure as a starting model.

Iterative cycles of model building and refinement were carried out using COOT [58], PHENIX [59], and BUSTER [60]. TLS refinement with four groups per polypeptide chain was used. All other crystallographic calculations were carried out with the ccp4 package [61]. A summary of the data statistics for the refined structure is provided in Table 1. The atomic coordinates and experimental data (PDB ID 5g02, 5fub, 5ful, 5fwa) have been deposited in the Protein Data Bank.

CARM1 purification, crystallization, and structure determination

A construct of mouse CARM1 from residues 130 to 487 (mCARM1₁₃₀₋₄₈₇) was expressed as a GST fusion protein using the baculovirus system, purified as previously described [62] and crystallized in 15% (w/v) PEG 2000MME, 100 mM Tris/HCl pH 8.0, 100 mM NaCl in the presence of 1 mM Cp1 at 20 °C. mCARM1 crystals were cryoprotected by a short soak in 22% (w/v) PEG 2000MME, 100 mM Tris/HCl pH 8.0, 100 mM NaCl, 15% (v/v) PEG 400, and flash-cooled in liquid nitrogen. Diffraction datasets were collected on SOLEIL PROXIMA2 beamline, using a Quantum 315 (ADSC) detector and processed with XDS. The structure was solved and refined using PHENIX. Model building and refinement were carried out using COOT [58], PHENIX [59], and BUSTER [60]. TLS refinement with three groups per polypeptide chain was used. All other crystallographic calculations were carried out with the CCP4 package [61]. A summary of the data statistics for the refined structure is provided in Table 1. The atomic coordinates and experimental data (PDB ID 5k8v) have been deposited in the Protein Data Bank.

RSF1 expression and purification

A synthetic gene encoding RSF1 from *S. frugiperda* and codon-optimized for the expression in *E. coli* (Eurofins Genomics, Ebersberg, Germany) was inserted into the *NdeI* and *BamHI* sites of a pET-15b vector (Novagen, Madison, WI, USA). RSF1 was overexpressed in *E. coli* BL21 (DE3) as a N-terminal His-tagged fusion protein in 1 L LB/ampicillin medium. Protein expression was induced for 3 h at 37 °C by the addition of 0.4 mM IPTG when A600 nm reached 0.6. Harvested cells were suspended in 20 mM Tris/HCl pH 8.0, 250 mM NaCl and lysed by sonication. The lysate was centrifuged for 30 min at 40 000 g and the crude extract was applied to a Ni-NTA Superflow Agarose (Qiagen) column equilibrated in the same buffer. His6-RSF1 was eluted with 250 mM imidazole and concentrated at 0.9 mg·mL⁻¹ in 20 mM Tris/HCl pH 8.0, 100 mM NaCl, and 0.5 mM DTT.

Gel-based activity assays

Methylation assays were performed at 23 °C in 20 mM Tris/HCl pH 8, 100 mM NaCl, 0.5 mM DTT. PRMT2 at varying concentrations was incubated for 3–5 h with a substrate (500 μM histone H3 or H4 tail peptides, or 5 μM RSF1) and 0.02 mM S-adenosyl-L-[Me-¹⁴C]methionine (¹⁴C-SAM) at 50 mCi·mmol⁻¹ (PerkinElmer, Richmond, VA, USA). The reaction was stopped by addition of SDS loading buffer 4× and analyzed on 16% Tricine/SDS/PAGE gel. After electrophoresis, the gel was stained with Coomassie brilliant blue, dried on a Whatman 3MM paper sheet and exposed for 48 h against a BAS-MS 2025

imaging plate (GE Healthcare). The amount of radioactivity incorporated by the substrate was quantified by phosphorimager analysis (Typhoon; GE Healthcare). The amount of arginine methylation was deduced from the radioactive signal using spots of ¹⁴C-SAM as a reference.

IC₅₀ determination

All the SAM analogs were dissolved in water and their concentration was determined by UV absorption ($\epsilon_{260} = 15\,400\text{ M}^{-1}\cdot\text{cm}^{-1}$). Inhibition assays were adapted from the gel-based activity assay described above: Solutions containing either 10 μM RSF1 and 1.2 μM mPRMT2 in 20 mM Tris/HCl pH 8.0, 100 mM NaCl, 0.5 mM DTT or 0.8 μM mCARM1, 10 μM H3 peptide in 50 mM Tris/HCl pH 8, 200 mM NaCl, 0.5 mM DTT were incubated at 23 °C with increasing concentrations of SAM analogs (0.1–500 μM). The reaction was initiated by adding 20 μM ¹⁴C-labeled SAM and stopped after 4 h for PRMT2 or 30 min in the case of CARM1 assays. IC₅₀ values were determined by fitting the data for sigmoidal dose–response curves using the R version 2.13.0 software package [63].

Thermal shift assay

Thermal shift assay experiments were performed on a MiniOpticon Real-Time PCR System MJ Mini (Bio-Rad). The protein sample (4 μM) was incubated in 20 mM Tris/HCl pH 8.0, 100 mM NaCl with 1 mM of potential ligands in a total volume of 23 μL for 30 min at room temperature. The samples were heated from 20 °C to 94 °C at 0.1 °C·s⁻¹ after addition of 7 μL Sypro Orange 25× (Thermo Fisher Scientific). The fluorescence intensity was plotted as a function of temperature using the MJ OPTICON MONITOR ANALYSIS software. Melting temperature (T_m) was given by the inflection point of the fluorescence curve.

SAXS experiments and data processing

zPRMT2₇₃₋₄₀₈ SAXS data were collected on a BioSAXS-1000 camera (Rigaku, Ettlingen, Germany) mounted on a MicroMax 007 HF Rigaku rotating anode at a sample-detector distance of 500 mm, covering the range of momentum transfer $0.008 < q < 0.65\text{ \AA}^{-1}$ ($q = 4\pi \sin(\theta)/\lambda$, where 2θ is the scattering angle and $\lambda = 0.154\text{ nm}$ is the X-ray wavelength) in four frames (30 min each) to check for possible radiation damage. The q scale was calibrated by silver behenate powder diffraction and all data were collected up to a maximum q of 0.35 \AA^{-1} . Scattering measurements were carried out at 20 °C in a thermostated quartz capillary with a diameter of 1.5 mm. zPRMT2 full-length SAXS data were collected at the SAXS beamline P12 at the PETRA III storage ring (Deutsches Elektronen-Synchrotron, Hamburg, Germany). The exposure time was 0.045 s.

All zPRMT2 samples were solubilized in 20 mM Tris/HCl pH 8.0, 100 mM NaCl, 5 mM TCEP, 1 mM SAH and measured at several concentrations. The SAXS data were processed by standard procedures using the ATSAS software package (version 2.7.2). The scattering intensity $I(0)$ and the radius of gyration R_g were evaluated using the Guinier approximation with PRIMUS [64]. These parameters were also determined from the entire scattering pattern using GNOM which computes the distance distribution function $P(r)$ and the maximum particle dimension D_{\max} [65]. The program CRY SOL was used to fit the theoretical scattering pattern to the experimental data [66]. For *ab initio* model construction, multiple runs were performed using the program DAMMIF [67] to verify the stability of the solution. Ten models were averaged using DAMAVER [68] to produce a starting model for a second modeling step using DAMMIN. The final model was superimposed to the X-ray structure using SUPCOMB [69].

Mass spectrometry analysis

Protein preparation for liquid digestion

The Rapid Gest was reconstituted with 1 mL of 25 mM NH_4HCO_3 . Eighty-five microliter of the Rapid Gest solution was added to 15 μL of sample. The reduction was done with 5 mM DTT (final concentration) in 25 mM NH_4HCO_3 (1 h at 57 °C) and alkylation was performed with 15 mM iodoacetamide (final concentration) in 25 mM NH_4HCO_3 (30 min in the dark). The trypsin (Promega, Madison, WI, USA; V511A) was suspended in 100 μL of 25 mM NH_4HCO_3 . Five microliter was added to the sample and incubated overnight at 37 °C. The digestion was stopped by the addition of 1 μL of TFA.

A second digestion was performed with the pepsin. The pepsin was suspended at 1 mg·mL⁻¹ in water at pH 1. Five microliter was added to the sample and incubated 3 h at 37 °C. The digestion was stopped by heating the sample at 90 °C during 5 min. The sample was centrifuged 5 min at 10 000 g and the supernatant analyses by nanoLC-MS/MS.

Chromatography conditions on NanoAcquity

The analysis was performed on a nanoACQUITY Ultra-Performance-LC (UPLC; Waters, Milford, MA, USA). The samples were trapped on a 20 × 0.18 mm, 5 μm Symmetry C18 precolumn (Waters Corp.), and the peptides were separated on a ACQUITY UPLC® BEH130 C18 column (Waters Corp.), 75 μm × 250 mm, 1.7 μm particle size. The solvent system consisted of 0.1% formic acid in water (solvent A) and 0.1% formic acid in acetonitrile (solvent B). Trapping was performed during 3 min at 5 $\mu\text{L}\cdot\text{min}^{-1}$ with 99% of solvent A and 1% of solvent B. Elution was performed at a flow rate of 400 nL·min⁻¹, using 1–40% gradient (solvent B) over 35 min at 45 °C followed by 65% (solvent B) over 5 min.

MS and MS/MS conditions on SYNAPT mass spectrometer

The MS and MS/MS analyses were performed on the SYNAPT™, a hybrid quadrupole orthogonal acceleration time-of-flight tandem mass spectrometer (Waters) equipped with a Z-spray ion source and a lock-mass system. The capillary voltage was set at 3.5 kV and the cone voltage at 35 V. Mass calibration of the TOF was achieved using Glu-fibrino-peptide B fragments obtained at 27 V on trap cell on the (50–2000) m/z range in positive mode. Online correction of this calibration was performed with Glu-fibrino-peptide B as the lock-mass. The ion $(M + 2H)^{2+}$ at m/z 785.8426 is used to calibrate MS data and the fragment ion $(M + H)^+$ at m/z 684.3469 is used to calibrate MS/MS data during the analysis.

For tandem MS experiments, the system was operated with automatic switching between MS and MS/MS modes (MS 0.5 s/scan on m/z range [250–1500] and MS/MS 0.7 s/scan on m/z range [50–2000]). The three most abundant peptides (intensity threshold 60 counts·s⁻¹), preferably doubly and triply charged ions, were selected on each MS spectrum for further isolation and CID fragmentation with two energies set using collision energy profile. Fragmentation was performed using argon as the collision gas. The complete system was fully controlled by MassLynx 4.1 (SCN 566; Waters). Raw data collected during nanoLC-MS/MS analysis were processed and converted with ProteinLynx Browser 2.5 (Waters) into. pkl peak list format. Normal background subtraction type was used for both MS and MS/MS with 5% threshold and polynomial correction of order 5, and deisotoping was performed.

MS and MS/MS conditions on TripleTOF 5600+ mass spectrometer

The MS and MS/MS analyses were performed on the TripleTOF 5600, a hybrid quadrupole orthogonal acceleration time-of-flight tandem mass spectrometer (AB Sciex, Villebon, France). The mass spectrometer was operated in positive mode, with the following settings: ion spray voltage floating (ISVF) 2300 V, curtain gas (CUR) 25 psi, interface heater temperature (IHT) 75 °C, ion source gas 1 (GS1) 2 psi, declustering potential (DP) 100 V. Information-dependent acquisition (IDA) mode was used with top-five MS/MS scans. The MS scan had an accumulation time of 250 ms on m/z (400–1250) range and the MS/MS scans had 100 ms m/z (150–1600) range in high sensitivity mode. Switching criteria were set to ions with charge state of 2–4 and an abundance threshold of more than 150 counts, exclusion time was set at 12 s. IDA rolling collision energy script was used for automatically adapting the CE. Mass calibration of the analyzer was achieved using peptides from digested BSA. The complete system was fully controlled by AnalystTF 1.6 (AB Sciex).

MS/MS data interpretation

The peak list has been searched against a Swiss Prot-derived combined target-decoy database using MASCOT (version 2.4.1; Matrix Science, London, UK). The database contained sequences of human proteins including common contaminants (human keratins) and porcine trypsin and was created using an in-house database generation toolbox (<http://msda.u-strasbg.fr>). During database search, up to one missed cleavage by trypsin and four variable modifications [oxidation of methionine (+16 Da), carbamidomethylation of cysteine (+57 Da), and methylation or dimethylation on arginine] were considered. The search window was set to 15 ppm for precursor ions and 0.05 Da for fragment ions. MASCOT result files (.dat) were imported into SCAFFOLD 3 software (version 3.6.5; Proteome Software Inc., Portland, OR, USA) and filtering criteria based on probability-based scoring of the identified peptides were taken into account: peptides having an (Mascot ion score minus identity score) above 0 and a Mascot Ion score above 30 were validated. The false discovery rate (FDR) was calculated to be < 1% based on the number of decoy hits.

Acknowledgements

This work was supported by grants from CNRS, Université de Strasbourg, INSERM, Instruct, part of the European Strategy Forum on Research Infrastructures (ESFRI) supported by national member subscriptions as well as the French Infrastructure for Integrated Structural Biology (FRISBI) (ANR-10-INSB-05-01, grant ANR-10-LABX-0030-INRT), a French State fund managed by the Agence Nationale de la Recherche under the frame program Investissements d'Avenir labeled ANR-10-IDEX-0002-02 to JC and IGBMC, grants from the French Proteomic Infrastructure (ANR-10-INBS-08; ProFI project) to SC and grants from Association pour la Recherche contre le Cancer (ARC) (No. A09/4/5005, No. SFI20121205902). We thank the members of the IGBMC common services and the members of the structural biology platform of IGBMC for technical assistance. We thank Alastair McEwen for help during data collection at synchrotron sites. We thank Johann Stojko for help during MS analysis. We thank Natacha Rochel for help during SAXS data collection and PETRA III P12 beamline staff for technical assistance during data collection at the Deutsches Elektronen-Synchrotron. We thank members of SOLEIL Proxima1/Proxima2 beamlines and the European Synchrotron Radiation Facility–European Molecular Biology Laboratory joint Structural Biology groups for the use of beamlines facilities and for help during X-ray data collection.

Author contributions

VC, NM, NT-C performed the biochemical and biophysical experiments. VC, NM, LB, and JC performed the structure determination by X-ray crystallography. JMS and SC performed and analyzed the mass spectrometry experiments. MJH and NIM provided the cp1 compounds. VC, LB, SC, and JC wrote the article. JC directed the study.

References

- 1 Bedford M & Clarke S (2009) Protein arginine methylation in mammals: who, what, and why. *Mol Cell* **33**, 1–13.
- 2 Lee Y-H & Stallcup MR (2009) Minireview: protein arginine methylation of nonhistone proteins in transcriptional regulation. *Mol Endocrinol* **23**, 425–433.
- 3 Yang Y & Bedford MT (2013) Protein arginine methyltransferases and cancer. *Nat Rev Cancer* **13**, 37–50.
- 4 Niewmierzycka A & Clarke S (1999) S-Adenosylmethionine-dependent methylation in *Saccharomyces cerevisiae*. Identification of a novel protein arginine methyltransferase. *J Biol Chem* **274**, 814–824.
- 5 Weiss VH, McBride AE, Soriano MA, Filman DJ, Silver PA & Hogle JM (2000) The structure and oligomerization of the yeast arginine methyltransferase, Hmt1. *Nat Struct Biol* **7**, 1165–1171.
- 6 Zhang X, Zhou L & Cheng X (2000) Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *EMBO J* **19**, 3509–3519.
- 7 Zhang X & Cheng X (2003) Structure of the predominant protein arginine methyltransferase PRMT1 and analysis of its binding to substrate peptides. *Structure* **11**, 509–520.
- 8 Troffer-Charlier N, Cura V, Hassenboehler P, Moras D & Cavarelli J (2007) Functional insights from structures of coactivator-associated arginine methyltransferase 1 domains. *EMBO J* **26**, 4391–4401.
- 9 Yue WW, Hassler M, Roe SM, Thompson-Vale V & Pearl LH (2007) Insights into histone code syntax from structural and biochemical studies of CARM1 methyltransferase. *EMBO J* **26**, 4402–4412.
- 10 Cheng Y, Frazier M, Lu F, Cao X & Redinbo MR (2011) Crystal structure of the plant epigenetic protein arginine methyltransferase 10. *J Mol Biol* **414**, 106–122.
- 11 Antonyamy 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 USA* **109**, 17960–17965.
- 12 Ho MC, Wilczek C, Bonanno JB, Xing L, Seznec J, Matsui T, Carter LG, Onikubo T, Kumar PR, Chan MK *et al.* (2013) Structure of the arginine

- methyltransferase PRMT5-MEP50 reveals a mechanism for substrate specificity. *PLoS One* **8**, e57008.
- 13 Wang C, Zhu Y, Chen J, Li X, Peng J, Zou Y, Zhang Z, Jin H, Yang P, Wu J *et al.* (2014) Crystal structure of arginine methyltransferase 6 from *Trypanosoma brucei*. *PLoS One* **9**, e87267.
 - 14 Wang C, Zhu Y, Caceres Tamar B, Liu L, Peng J, Wang J, Chen J, Chen X, Zhang Z, Zuo X *et al.* (2014) Structural determinants for the strict monomethylation activity by *Trypanosoma brucei* protein arginine methyltransferase 7. *Structure* **22**, 756–768.
 - 15 Hasegawa M, Toma-Fukai S, Kim J-D, Fukamizu A & Shimizu T (2014) Protein arginine methyltransferase 7 has a novel homodimer-like structure formed by tandem repeats. *FEBS Lett* **588**, 1942–1948.
 - 16 Cura V, Troffer-Charlier N, Wurtz JM, Bonnefond L & Cavarelli J (2014) Structural insight into arginine methylation by the mouse protein arginine methyltransferase 7: a zinc finger freezes the mimic of the dimeric state into a single active site. *Acta Crystallogr D Biol Crystallogr* **70**, 2401–2412.
 - 17 Lee WC, Lin WL, Matsui T, Chen ES, Wei TY, Lin WH, Hu H, Zheng YG, Tsai MD & Ho MC (2015) Protein arginine methyltransferase 8: tetrameric structure and protein substrate specificity. *Biochemistry* **54**, 7514–7523.
 - 18 Bonnefond L, Stojko J, Mailliot J, Troffer-Charlier N, Cura V, Wurtz JM, Cianferani S & Cavarelli J (2015) Functional insights from high resolution structures of mouse protein arginine methyltransferase 6. *J Struct Biol* **191**, 175–183.
 - 19 Toma-Fukai S, Kim JD, Park KE, Kuwabara N, Shimizu N, Krayukhina E, Uchiyama S, Fukamizu A & Shimizu T (2016) Novel helical assembly in arginine methyltransferase 8. *J Mol Biol* **428**, 1197–1208.
 - 20 Scott HS, Antonarakis SE, Lalioti MD, Rossier C, Silver PA & Henry MF (1998) Identification and characterization of two putative human arginine methyltransferases (HRMT1L1 and HRMT1L2). *Genomics* **48**, 330–340.
 - 21 Lakowski TM & Frankel A (2009) Kinetic analysis of human protein arginine N-methyltransferase 2: formation of monomethyl- and asymmetric dimethyl-arginine residues on histone H4. *Biochem J* **421**, 253–261.
 - 22 Qi C, Chang J, Zhu Y, Yeldandi AV, Rao SM & Zhu Y-J (2002) Identification of protein arginine methyltransferase 2 as a coactivator for estrogen receptor alpha. *J Biol Chem* **277**, 28624–28630.
 - 23 Meyer R, Wolf SS & Obendorf M (2007) PRMT2, a member of the protein arginine methyltransferase family, is a coactivator of the androgen receptor. *J Steroid Biochem Mol Biol* **107**, 1–14.
 - 24 Kzhyshkowska J, Kremmer E, Hofmann M, Wolf H & Dobner T (2004) Protein arginine methylation during lytic adenovirus infection. *Biochem J* **383**, 259–265.
 - 25 Yoshimoto T, Boehm M, Olive M, Crook MF, San H, Langenickel T & Nabel EG (2006) The arginine methyltransferase PRMT2 binds RB and regulates E2F function. *Exp Cell Res* **312**, 2040–2053.
 - 26 Blythe SA, Cha SW, Tadjuidje E, Heasman J & Klein PS (2010) beta-Catenin primes organizer gene expression by recruiting a histone H3 arginine 8 methyltransferase, Prmt2. *Dev Cell* **19**, 220–231.
 - 27 Iwasaki H, Kovacic JC, Olive M, Beers JK, Yoshimoto T, Crook MF, Tonelli LH & Nabel EG (2010) Disruption of protein arginine N-methyltransferase 2 regulates leptin signaling and produces leanness *in vivo* through loss of STAT3 methylation. *Circ Res* **107**, 992–1001.
 - 28 Zhong J, Cao RX, Liu JH, Liu YB, Wang J, Liu LP, Chen YJ, Yang J, Zhang QH, Wu Y *et al.* (2014) Nuclear loss of protein arginine N-methyltransferase 2 in breast carcinoma is associated with tumor grade and overexpression of cyclin D1 protein. *Oncogene* **33**, 5546–5558.
 - 29 Oh TG, Bailey P, Dray E, Smith AG, Goode J, Eriksson N, Funder JW, Fuller PJ, Simpson ER, Tilley WD *et al.* (2014) PRMT2 and RORgamma expression are associated with breast cancer survival outcomes. *Mol Endocrinol* **28**, 1166–1185.
 - 30 Zhong J, Cao RX, Hong T, Yang J, Zu XY, Xiao XH, Liu JH & Wen GB (2011) Identification and expression analysis of a novel transcript of the human PRMT2 gene resulted from alternative polyadenylation in breast cancer. *Gene* **487**, 1–9.
 - 31 Zhong J, Cao RX, Zu XY, Hong T, Yang J, Liu L, Xiao XH, Ding WJ, Zhao Q, Liu JH *et al.* (2012) Identification and characterization of novel spliced variants of PRMT2 in breast carcinoma. *FEBS J* **279**, 316–335.
 - 32 Wang YC, Wang JD, Chen CH, Chen YW & Li C (2015) A novel BLAST-based relative distance (BBRD) method can effectively group members of protein arginine methyltransferases and suggest their evolutionary relationship. *Mol Phylogenet Evol* **84**, 101–111.
 - 33 Morales Y, Caceres T, May K & Hevel JM (2016) Biochemistry and regulation of the protein arginine methyltransferases (PRMTs). *Arch Biochem Biophys* **590**, 138–152.
 - 34 Miranda TB, Miranda M, Frankel A & Clarke S (2004) PRMT7 is a member of the protein arginine methyltransferase family with a distinct substrate specificity. *J Biol Chem* **279**, 22902–22907.
 - 35 Hadjikyriacou A, Yang Y, Espejo A, Bedford MT & Clarke SG (2015) Unique features of human protein arginine methyltransferase 9 (PRMT9) and its substrate RNA splicing factor SF3B2. *J Biol Chem* **290**, 16723–16743.
 - 36 Lin WJ, Gary JD, Yang MC, Clarke S & Herschman HR (1996) The mammalian immediate-early TIS21

- protein and the leukemia-associated BTG1 protein interact with a protein-arginine N-methyltransferase. *J Biol Chem* **271**, 15034–15044.
- 37 Rho J, Choi S, Seong YR, Cho WK, Kim SH & Im DS (2001) Prmt5, which forms distinct homo-oligomers, is a member of the protein-arginine methyltransferase family. *J Biol Chem* **276**, 11393–11401.
- 38 Lim Y, Kwon YH, Won NH, Min BH, Park IS, Paik WK & Kim S (2005) Multimerization of expressed protein-arginine methyltransferases during the growth and differentiation of rat liver. *Biochim Biophys Acta* **1723**, 240–247.
- 39 Herrmann F, Pably P, Eckerich C, Bedford MT & Fackelmayer FO (2009) Human protein arginine methyltransferases *in vivo*—distinct properties of eight canonical members of the PRMT family. *J Cell Sci* **122**, 667–677.
- 40 Feng Y, Xie N, Jin M, Stahley MR, Stivers JT & Zheng YG (2011) A transient kinetic analysis of PRMT1 catalysis. *Biochemistry* **50**, 7033–7044.
- 41 Krissinel EB, Winn MD, Ballard CC, Ashton AW, Patel P, Potterton EA, McNicholas SJ, Cowtan KD & Emsley P (2004) The new CCP4 Coordinate Library as a toolkit for the design of coordinate-related applications in protein crystallography. *Acta Crystallogr D Biol Crystallogr* **60**, 2250–2255.
- 42 Labourier E, Bourbon H-M, Gallouzi I-E, Fostier M, Allemand E & Tazi J (1999) Antagonism between RSF1 and SR proteins for both splice-site recognition *in vitro* and *Drosophila* development. *Genes Dev* **13**, 740–753.
- 43 Bradley T, Cook ME & Blanchette M (2015) SR proteins control a complex network of RNA-processing events. *RNA* **21**, 75–92.
- 44 Feng Y, Hadjikyriacou A & Clarke SG (2014) Substrate specificity of human protein arginine methyltransferase 7 (PRMT7): the importance of acidic residues in the double E loop. *J Biol Chem* **289**, 32604–32616.
- 45 van Haren M, van Ufford LQ, Moret EE & Martin NI (2015) Synthesis and evaluation of protein arginine N-methyltransferase inhibitors designed to simultaneously occupy both substrate binding sites. *Org Biomol Chem* **13**, 549–560.
- 46 Pesiridis GS, Diamond E & Van Duyne GD (2009) Role of pICln in methylation of Sm proteins by PRMT5. *J Biol Chem* **284**, 21347–21359.
- 47 Sayegh J, Webb K, Cheng D, Bedford M & Clarke S (2007) Regulation of protein arginine methyltransferase 8 (PRMT8) activity by its N-terminal domain. *J Biol Chem* **282**, 36444–36453.
- 48 Yu MC (2011) The role of protein arginine methylation in mRNP dynamics. *Mol Biol Int* **2011**, 163827.
- 49 Friesen WJ, Paushkin S, Wyce A, Massenet S, Pesiridis GS, Duyne GV, Rappsilber J, Mann M & Dreyfuss G (2001) The methylosome, a 20S complex containing JBP1 and pICln, produces dimethylarginine-modified Sm proteins. *Mol Cell Biol* **21**, 8289–8300.
- 50 Cheng D, Cote J, Shaaban S & Bedford MT (2007) The arginine methyltransferase CARM1 regulates the coupling of transcription and mRNA processing. *Mol Cell* **25**, 71–83.
- 51 Sinha R, Allemand E, Zhang Z, Karni R, Myers MP & Krainer AR (2010) Arginine methylation controls the subcellular localization and functions of the oncoprotein splicing factor SF2/ASF. *Mol Cell Biol* **30**, 2762–2774.
- 52 Bressan GC, Moraes EC, Manfioli AO, Kuniyoshi TM, Passos DO, Gomes MD & Kobarg J (2009) Arginine methylation analysis of the splicing-associated SR protein SFRS9/SRP30C. *Cell Mol Biol Lett* **14**, 657–669.
- 53 Abdulrahman W, Uhring M, Kolb-Cheynel I, Garnier JM, Moras D, Rochel N, Busso D & Poterszman A (2009) A set of baculovirus transfer vectors for screening of affinity tags and parallel expression strategies. *Anal Biochem* **385**, 383–385.
- 54 Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132.
- 55 Battye TGG, Kontogiannis L, Johnson O, Powell HR & Leslie AGW (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* **67**, 271–281.
- 56 Otwinowski Z & Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. In *Macromolecular Crystallography* (Carter CW & Sweet RM, eds), pp. 307–326. Academic Press, New York.
- 57 Long F, Vagin AA, Young P & Murshudov GN (2008) BALBES: a molecular-replacement pipeline. *Acta Crystallogr D Biol Crystallogr* **64**, 125–132.
- 58 Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126–2132.
- 59 Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213–221.
- 60 Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, Roversi P, Sharff A, Smart OS, Vornrhein C *et al.* (2011) BUSTER version 2.10.0. Global Phasing Ltd., Cambridge, UK.
- 61 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* **67**, 235–242.
- 62 Troffer-Charlier N, Cura V, Hassenboehler P, Moras D & Cavarelli J (2007) Expression, purification,

- crystallization and preliminary crystallographic study of isolated modules of the mouse coactivator-associated arginine methyltransferase 1. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **63**, 330–333.
- 63 R Development Core Team (2005) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- 64 Konarev PV, Volkov VV, Sokolova AV, Koch MHJ & Svergun DI (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr* **36**, 1277–1282.
- 65 Svergun D (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J Appl Crystallogr* **25**, 495–503.
- 66 Svergun D, Barberato C & Koch MHJ (1995) CRY SOL – a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J Appl Crystallogr* **28**, 768–773.
- 67 Franke D & Svergun DI (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr* **42**, 342–346.
- 68 Volkov VV, Lapuk VA, Kayushina RL, Shtykova EV, Varlamova EY, Malfois M & Svergun DI (2003) Low-resolution structure of immunoglobulins IgG1, IgM and rheumatoid factor IgM-RF from solution X-ray scattering data. *J Appl Crystallogr* **36**, 503–508.
- 69 Kozin MB & Svergun DI (2001) Automated matching of high- and low-resolution structural models. *J Appl Crystallogr* **34**, 33–41.
- 70 Beitz E (2000) TEXshade: shading and labeling of multiple sequence alignments using LATEX2 epsilon. *Bioinformatics* **16**, 135–139.
- 71 Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Cassarino TG, Bertoni M, Bordoli L *et al.* (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* **42**, W252–W258.
- 72 Laskowski RA & Swindells MB (2011) LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Inf Model* **51**, 2778–2786.
- 73 Karplus PA & Diederichs K (2015) Assessing and maximizing data quality in macromolecular crystallography. *Curr Opin Struct Biol* **34**, 60–68.

Transition state mimics are valuable mechanistic probes for structural studies with the arginine methyltransferase CARM1

Matthijs J. van Haren, Nils Marechal, Nathalie Troffer-Charlier, Agostino Cianciulli, Gianluca Sbardella, Jean Cavarelli, and Nathaniel I. Martin



Transition state mimics are valuable mechanistic probes for structural studies with the arginine methyltransferase CARM1

Matthijs J. van Haren^{a,1}, Nils Marechal^{b,1}, Nathalie Troffer-Charlier^b, Agostino Cianciulli^c, Gianluca Sbardella^c, Jean Cavarelli^{b,2}, and Nathaniel I. Martin^{a,2}

^aDepartment of Chemical Biology & Drug Discovery, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CG Utrecht, The Netherlands;

^bDepartment of Integrated Structural Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire, Université de Strasbourg, CNRS UMR 7104, INSERM U 964, Illkirch, F-67404, France; and ^cEpigenetic Med Chem Lab, Dipartimento di Farmacia, Università degli Studi di Salerno, I-84084 Fisciano (SA), Italy

Edited by Jocelyn Côté, University of Ottawa, Ottawa, Canada, and accepted by Editorial Board Member Gregory A. Petsko February 28, 2017 (received for review November 6, 2016)

Coactivator associated arginine methyltransferase 1 (CARM1) is a member of the protein arginine methyltransferase (PRMT) family and methylates a range of proteins in eukaryotic cells. Overexpression of CARM1 is implicated in a number of cancers, and it is therefore seen as a potential therapeutic target. Peptide sequences derived from the well-defined CARM1 substrate poly(A)-binding protein 1 (PABP1) were covalently linked to an adenosine moiety as in the AdoMet cofactor to generate transition state mimics. These constructs were found to be potent CARM1 inhibitors and also formed stable complexes with the enzyme. High-resolution crystal structures of CARM1 in complex with these compounds confirm a mode of binding that is indeed reflective of the transition state at the CARM1 active site. Given the transient nature of PRMT–substrate complexes, such transition state mimics represent valuable chemical tools for structural studies aimed at deciphering the regulation of arginine methylation mediated by the family of arginine methyltransferases.

protein arginine *N*-methyltransferase | PRMT | CARM1 | transition state mimics | cocrystal structures

The methylation of arginine residues by protein arginine *N*-methyltransferases (PRMTs) plays a vital role in a variety of key cellular functions including gene regulation, signal transduction, RNA processing, and DNA repair (1, 2). Conversely, dysregulated PRMT activity is known to be linked to a variety of cancers and other diseases (3, 4). Coactivator associated arginine methyltransferase 1 (CARM1, also known as PRMT4) is directly involved in numerous cellular processes via methylation of the histone H3 tail peptide, splicing factors, RNA binding proteins, and coactivation of nuclear receptors (1, 5). Growing evidence indicates associations with the up-regulation of CARM1 in breast (6, 7), colon (8, 9), prostate (9, 10), and liver (11) cancers, making it an appealing potential therapeutic target.

Like all lysine and arginine methyltransferases, CARM1 uses the cofactor *S*-adenosine-*L*-methionine (AdoMet) as the source of the methyl group that is transferred to the guanidine group of the arginine side chain in a target protein or peptide with concomitant generation of *S*-adenosine-*L*-homocysteine (AdoHcy) (1). As for all type 1 PRMTs, the first methylation step catalyzed by CARM1 gives monomethyl-arginine (MMA), followed by a second methylation to form asymmetric dimethyl-arginine (ADMA). Although the different PRMTs vary significantly in overall size and sequence, they share a common active site architecture defined by specific amino acids (belonging to the so-called motifs I to IV) known to be key for catalysis (Fig. 1A) (1–4, 12). Most notable in this regard are residues involved in hydrogen bonding to the adenosine moiety of AdoMet and two conserved glutamate residues that form the so-called “double-E loop” critical for chelating and orienting the guanidine group of the target arginine residue. In doing so, the target guanidine

moiety is precisely presented to the electrophilic methylsulfonium group of AdoMet, facilitating an “S_N2-like” substitution reaction.

Recently, we reported a series of small molecule PRMT inhibitors wherein a guanidine moiety was covalently linked to an adenosine group meant to mimic that of AdoMet (13). Interestingly, some of these inhibitors displayed potent and specific inhibition of CARM1 with IC₅₀ values in the nanomolar range (13). As a means of enhancing CARM1 selectivity, we report here an approach wherein known CARM1 substrate peptides are covalently tethered to an adenosine unit to create transition state mimics of the CARM1 methylation reaction (Fig. 1B). Such transition state mimics were also envisioned to serve as mechanistic probes and chemical tools for structural studies. Previous reports have highlighted the challenges associated with cocrystallizing PRMTs with their various peptide substrates (14–18). Because of these challenges, a complete structural understanding of how PRMTs achieve specificity in arginine methylation at different target sites remains elusive. Furthermore, although structures of isolated PRMTs have been published, few studies exist that provide an atomic-level characterization of the

Significance

The posttranslational methylation of arginine is a widespread epigenetic modification catalyzed by the family of protein arginine methyltransferases (PRMTs). Dysregulation of PRMT expression is implicated in the pathogenesis of many diseases including human cancers. An atomic-scale understanding of the PRMT catalytic mechanism is crucial for both fundamental biological and pharmacological applications. Despite intense efforts, crystal structures of PRMT complexes with long peptides and full-length substrates have not been solved because of their inherent instability. To address this issue, we describe peptide-based transition state mimics that form stable complexes with the PRMT enzyme coactivator associated arginine methyltransferase 1 resulting in high-resolution cocrystal structures. Our findings provide an exciting approach to understanding PRMT substrate recognition and the regulation of arginine methylation.

Author contributions: G.S., J.C., and N.I.M. designed research; M.J.v.H., N.M., N.T.-C., and A.C. performed research; M.J.v.H., N.M., N.T.-C., A.C., J.C., and N.I.M. analyzed data; and M.J.v.H., G.S., J.C., and N.I.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. J.C. is a guest editor invited by the Editorial Board.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID codes 5LGP–5LGS).

¹M.J.v.H. and N.M. contributed equally to this work.

²To whom correspondence may be addressed. Email: n.i.martin@uu.nl or cava@igbmc.fr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1618401114/-DCSupplemental.

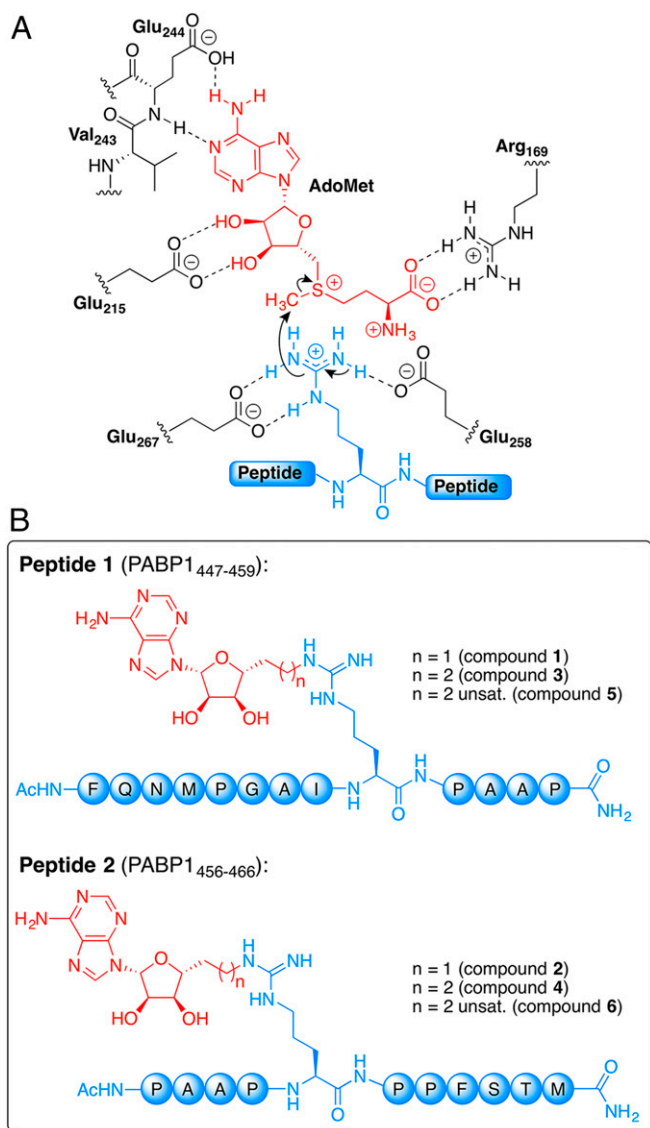


Fig. 1. The CARM1 active site and design of transition state mimics. (A) CARM1 employs an “S_N2-like” bisubstrate mechanism. Key active site residues are indicated including the conserved Glu₂₆₇ and Glu₂₅₈ residues comprising the guanidine-binding “double-E loop” (*Mus musculus* CARM1 numbering scheme). The AdoMet cofactor is highlighted in red and the target arginine-containing peptide in blue. (B) Structures of PABP1-derived constructs prepared in this study as CARM1 transition state mimics.

functional macromolecular complexes formed between PRMTs and their targets. Importantly, the transition-state mimicking compounds we here describe were found to cocrystallize with CARM1, forming stable complexes suitable for structural studies.

Results

Design of Transition State Mimics and Inhibition Studies. In selecting the CARM1-specific peptide substrates to be used in designing the transition state mimics, we focused on the nonhistonic poly (A)-binding protein 1 (PABP1), one of the most highly methylated CARM1 targets. Previous studies using peptide fragments derived from PABP1 identified the strongest CARM1 methylation sites as R455 and R460 (19). We therefore selected two PABP1 fragments, comprised of residues 447–459 and 456–466 (each containing a single arginine) to serve as the peptide segment of the transition state mimics (Fig. 1B). Building on methodology

recently developed in our group (20–22), we here report a convenient on-resin approach for preparing peptides in which an adenosine group can be covalently linked to the guanidine moiety of a specific arginine residue (compounds 1–6; Fig. 1B, see *SI Appendix, Materials and Methods* for complete synthetic details). Spacer lengths of two and three carbon atoms were explored because previous investigations indicated these closest mimic the PRMT transition state (13). In addition, unsaturated analogs of the three-carbon spacer constructs were also prepared to examine the effect of reduced linker flexibility. Compounds 1–6 were specifically designed so as to mimic the transition state of the first methylation reaction catalyzed by CARM1. Given that the first catalytic step is common to all PRMTs, this strategy represents a general approach that should be applicable for studying substrate binding by any PRMT enzyme (type I, II, or III) and its corresponding substrate(s). Also prepared were compounds 7 and 8, in which the target arginine of the PABP1 fragments was incorporated as the asymmetrically dimethylated species. Compounds 7 and 8 thus represent product inhibitors of CARM1 and were included as a means of evaluating the relative inhibitory potency of transition state mimics 1–6.

The inhibitory activity of compounds 1–8 toward CARM1 was investigated and compared with the ability to inhibit PRMT1, the most abundant arginine methyltransferase. An initial inhibition screen performed at a fixed inhibitor concentration of 50 μ M revealed that compounds 1–6 were all effective CARM1 inhibitors, whereas the dimethylated PABP1-derived peptides 7 and 8 showed no effect under these conditions. Based on these preliminary findings, complete IC₅₀ curves were generated for compounds 1–6 against both CARM1 and PRMT1 (Table 1). Analogs containing three-atom saturated linkers (compounds 3 and 4) and unsaturated linkers (compounds 5 and 6) display the most potent inhibition of CARM1 with IC₅₀ values <100 nM. Shortening the spacer from three to two carbon atoms results in a significant decrease in activity, suggesting that a three-atom linker between the guanidine moiety and the adenosine group is optimal for mimicking the transition state geometry. When tested against PRMT1, compounds 1–6 displayed significantly reduced inhibitory activity. In light of these findings, we next examined the application of the more potent CARM1 inhibiting compounds 3–6 as transition state mimics for use in cocrystallization studies.

Crystallization and Structure Determination of CARM1 Complexes with Transition State Mimicking Compounds 3–6. Cocrystallization studies were performed with the PABP1-derived transition state mimics and an isolated catalytic domain of mmCARM1 (*Mus musculus* CARM1, residues 130–487) (15, 23). All structures

Table 1. IC₅₀ values measured for compounds 1–8 against CARM1 and PRMT1

Compound	IC ₅₀ values*	
	CARM1	PRMT1
AdoHcy	0.276 ± 0.052	7.65 ± 2.98
1 (Peptide 1, n = 1)	1.93 ± 0.55	16.28 ± 4.49
2 (Peptide 2, n = 1)	5.90 ± 0.87	16.29 ± 5.68
3 (Peptide 1, n = 2)	0.0920 ± 0.0132	12.34 ± 3.65
4 (Peptide 2, n = 2)	0.0901 ± 0.0107	25.54 ± 7.40
5 (Peptide 1, n = 2, unsat.)	0.0817 ± 0.0116	4.29 ± 1.04
6 (Peptide 2, n = 2, unsat.)	0.0876 ± 0.0092	16.62 ± 5.96
7 (Peptide 1, aDMA) [†]	>50	>50
8 (Peptide 2, aDMA) [†]	>50	>50

*IC₅₀ values reported in micromolars. IC₅₀ values from triplicate data obtained at a range of 7–10 concentrations ± SD (see *SI Appendix, IC₅₀ curves*).
[†]In compounds 7 and 8, the central arginine residue is present in asymmetrically dimethylated form.

were solved and refined to 2-Å resolution in the space group $P2_12_12$ as seen for mmCARM1 catalytic domain (15) and contain one copy of the CARM1 tetramer in the asymmetric unit (*SI Appendix, Table S1*). In all cases, each CARM1 monomer binds one molecule of the transition state mimic investigated (Fig. 2*A*). As indicated in Fig. 2*B*, the structures solved include near-complete assignment of all residues in the peptide fragments of compounds 3–6. Compounds 3 and 5 are comprised of a 13-aa sequence from PABP1 (residues 447–459) centered around R455, flanked by eight residues on the N-terminal side and four residues on the C-terminal side. The electron density maps obtained in the cocrystallization studies with both 3 and 5 reveal the conformation of 12 of 13 residues in two monomers and for the full-length peptide (13 residues) in the two other CARM1 monomers (Fig. 3*A*). Compounds 4 and 6 are comprised of an 11-aa sequence from PABP1 (residues 456–466) centered around R460 flanked by four residues on the N-terminal side and six residues on the C-terminal side. The electron density maps obtained in the cocrystallization studies with both 4 and 6 reveal the conformation of 8 of 11 residues in three of the monomers and 9 of 11 residues in the remaining CARM1 monomer (Fig. 3*B*). The last two residues on the C-terminal side of 4 and 6 are not seen in the electron density map. As expected for the structures obtained with compounds 3–6, the adenosine moiety occupies the AdoMet binding pocket. In the other CARM1 structures reported to date, this same pocket is occupied by either AdoHcy or sinefungin (15–18).

The three carbon-atom linker appears to provide the ideal spacer to position and orient both the adenosine group and the

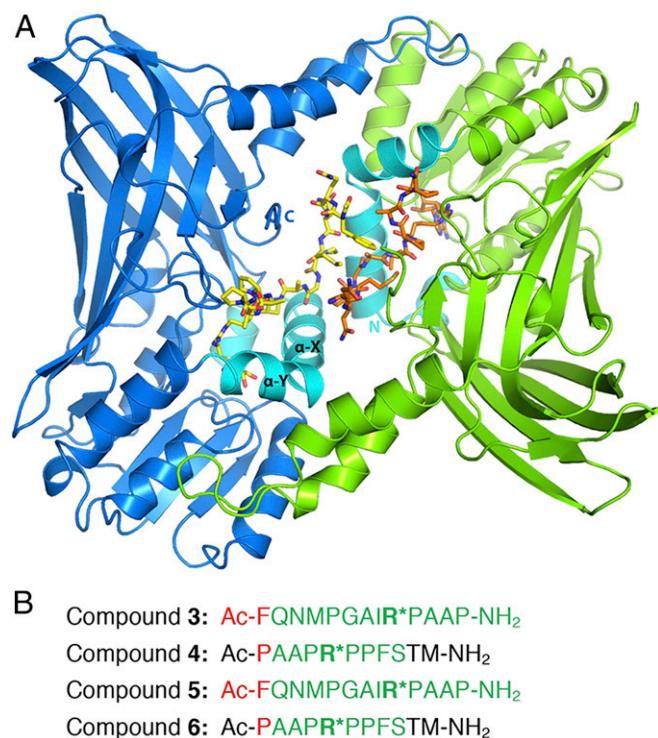


Fig. 2. Crystallization and structure determination of CARM1-transition state mimic complexes. (*A*) Dimer of mmCARM1 with compound 3 bound. CARM1 monomers within the dimer (green, blue) are represented as a cartoon. Compound 3 is represented as sticks. N-terminal helices of each monomer (not shown in other figures) are here highlighted in light blue. (*B*) Peptide sequences used and results from crystallization studies with compounds 3–6. Residues visible in all four monomers of the asymmetric unit indicated in green and those seen in at least one monomer are shown in red. Compounds 3 and 4 contain a saturated three-atom spacer, whereas compounds 5 and 6 contain an unsaturated three-atom spacer.

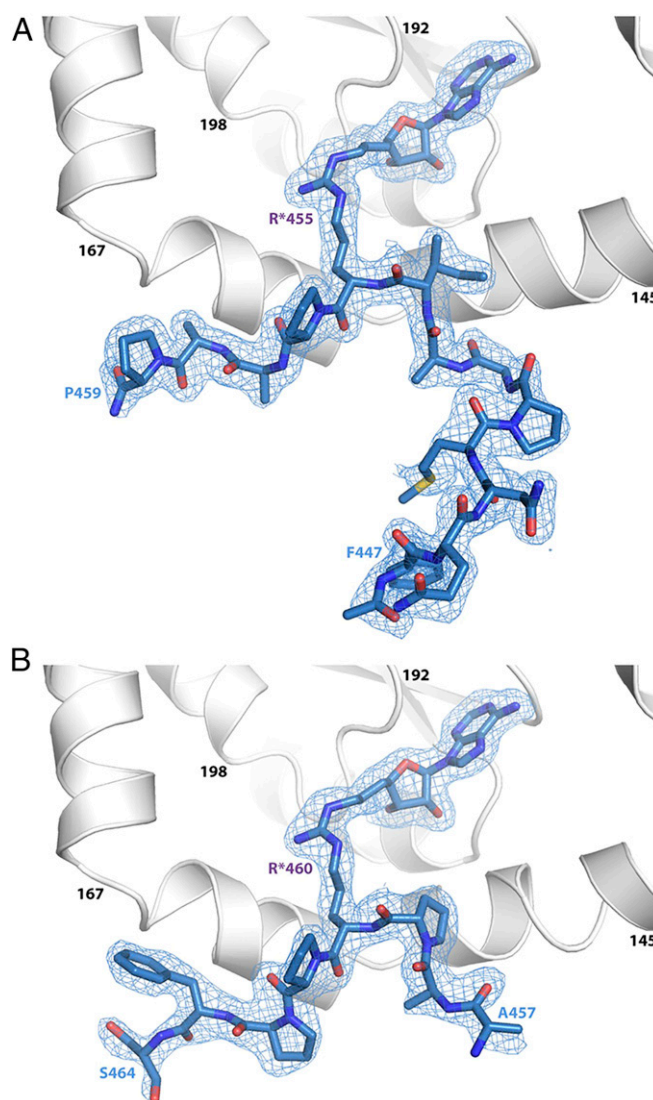


Fig. 3. Electron density ($2F_{obs} - F_{calc}$) weighted maps. Compound 5 (*A*) and compound 6 (*B*) bound to subunit B of mmCARM1. CARM1 is represented as cartoon, and peptide-based transition state mimics 5 and 6 are represented as sticks. Maps are represented as a mesh, contouring level set to 1 σ .

modified arginine residue present in the peptide-based transition state mimics (Fig. 4). Specifically, the four PRMT signature motifs surround the binding pocket occupied by the covalently linked adenosine and arginine components of compounds 3–6. The catalytic site is locked on one side by the PRMT motif I ($Y_{150}F_{151}xxY_{154}$ mmCARM1 numbering) and the guanidine moiety of the arginine side chain is hydrogen-bonded to E258 and E267 belonging to the double-E loop and H415 of the THW loop. Of particular interest is the comparison of the binding modes for the transition state mimics that contain a fully saturated spacer (3 and 4) with those containing an unsaturated spacer (5 and 6), revealing both to have similar conformations (Fig. 4 and *SI Appendix, Figs. S3 and S4*). These findings indicate that the unsaturated spacers in 5 and 6 lead to a preformed conformation suitable for binding in the CARM1 active site. The additional hydrogen atoms present in the saturated linkers of compounds 3 and 4 result in a modification of the conformation of M269 and a slightly different network of water molecules (Fig. 4 and *SI Appendix, Fig. S3*).

Another striking result is the finding that the conformations observed for the peptide components of the CARM1-bound

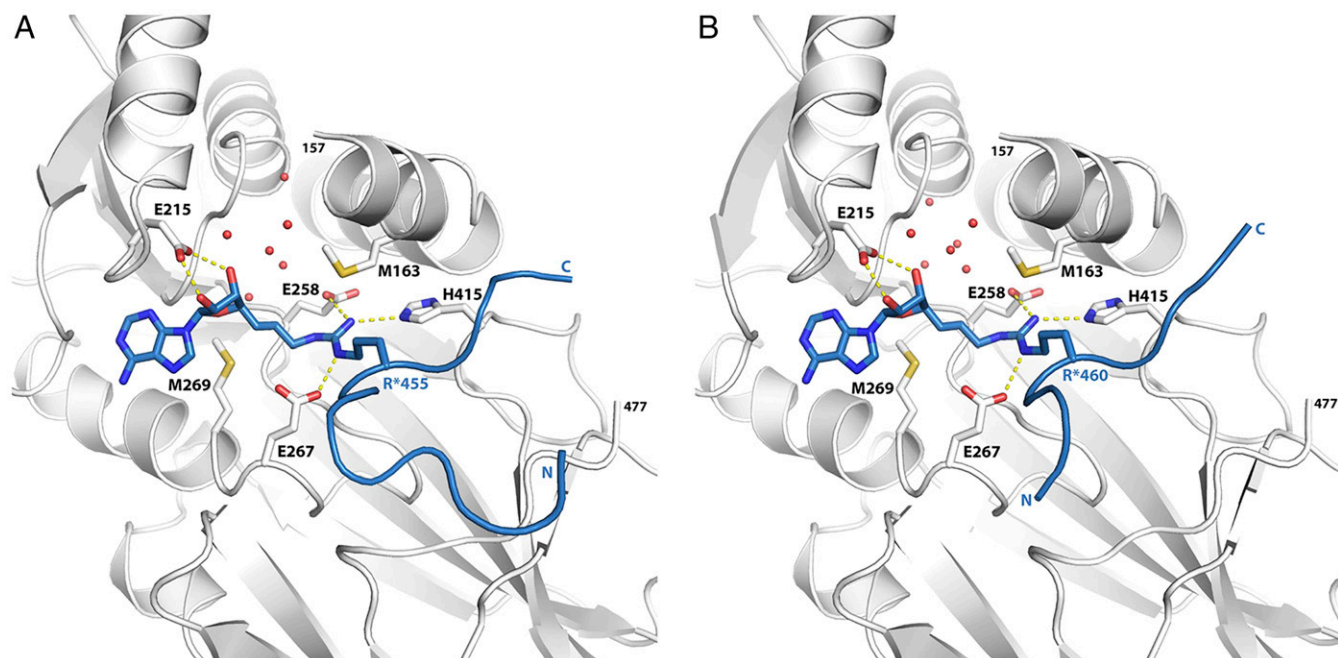


Fig. 4. Cocystal structures of CARM1 with transition state mimics **5** and **6** revealing interacting active site residues and water molecules. Cartoon and stick representation of the transition state mimics bound to mmCARM1. Compound **5** is derived from PABP1_{447–459} (A) and compound **6** is derived from PABP1_{456–466} (B). Both compounds contain an unsaturated linker between the guanidine moiety and the adenosine unit. For clarity, N-terminal helices of CARM1 are not shown.

transition state mimics are identical to those observed with unmodified PABP1 fragments bound to CARM1 in a recently reported cocrystallization study (18). Superimposition of the structures of the binary complexes formed between mmCARM1 and transition state mimics **3** or **5** with the structures recently reported for the ternary complexes of hsCARM1, sinefungin, and a PABP1-derived peptide containing R455 (18) reveal an identical conformation for the visible PABP1 substrate peptide in both complexes (Fig. 5A and *SI Appendix*, Figs. S5 and S6). A similar result is obtained when superimposing the structures of binary complexes of mmCARM1 and transition state mimics **4** and **6** with the ternary complex of hsCARM1, sinefungin, and another PABP1-derived peptide containing R460 (Fig. 5B and *SI Appendix*, Fig. S6). Moreover, the electron density maps of the complexes of mmCARM1 with **3** and **5** reveal the conformation of four more residues at the N-terminal side of the peptide compared with the ternary complexes of hsCARM1, sinefungin, and the PABP1-derived peptide containing R455 (Fig. 5A and *SI Appendix*, Fig. S5). In the case of the complexes formed between mmCARM1 and transition state mimics **4** and **6**, two additional N-terminal residues are also visible compared with the ternary complex of hsCARM1, sinefungin, and a PABP1-derived peptide containing R460 (Fig. 5B and *SI Appendix*, Fig. S6). Furthermore, it is interesting to note that the recently reported hsCARM1-PABP1 peptide complexes could only be solved in the presence of the AdoMet-competitive methyltransferase inhibitor sinefungin (18). The inclusion of sinefungin in the AdoMet binding site has an impact on the conformation of R169 that normally interacts with the methionine moiety of AdoMet. This distortion in turn leads to a small alteration in the conformation of E258, resulting in a change in the arrangement of bound water molecules. By comparison, transition state mimics **3–6** do not introduce such disturbances in the active site. The three carbon-atom linker used in these conjugates serves to properly orient and position the substrate peptide while simultaneously docking the adenosine moiety in the AdoMet binding site. In doing so, a “frozen” image of the transition state of the methylation reaction is effectively achieved.

Discussion

Given the challenges associated with obtaining stable complexes of PRMTs and their corresponding protein substrates, the majority of PRMT crystal structures reported to date are of free enzyme, often bound to an analog of the AdoMet cofactor. These structures reveal a strong conservation of catalytic domain architecture and specifically the AdoMet binding pocket. In contrast, no structures have been solved for PRMTs in complex with full-length target proteins, and only a few structures have been reported for PRMTs in complex with small, well-resolved, peptide substrate fragments (18, 24, 25). The covalent tethering of substrates to enzymes has proven to be an effective strategy for capturing difficult to characterize enzyme–substrate complexes, leading to their structural characterization (26, 27). In a related approach, we describe here the application of transition state mimics for enabling structural studies with the PRMTs. Specifically, we designed transition state mimics **1–6** based on CARM1 peptide substrates wherein the arginine that is normally methylated is instead covalently linked to an adenosine moiety mimicking that present in the AdoMet cofactor. Compounds **3–6** were found to be particularly potent inhibitors of CARM1, and the subsequently solved cocystal structures reveal that they are effective transition-state mimics capable of revealing biologically relevant peptide conformations. This assertion is supported by the near-identical conformations seen when superimposing the binary complexes described here with the recently reported ternary structures of hsCARM1, sinefungin, and a PABP1-derived peptide (18). Furthermore, it is interesting to note that ternary hsCARM1–PABP1 peptide complexes could only be solved in the presence of the AdoMet-competitive methyltransferase inhibitor sinefungin. The inclusion of sinefungin in the AdoMet binding site has an impact on the conformation of several residues. By comparison, the transition state mimics here described do not introduce any unnatural perturbations of the AdoMet binding site.

It has been shown that long distance interactions are crucial for substrate recognition and binding by PRMTs (15–18, 28, 29). However, despite extensive efforts, crystal structures of PRMT

$$Y = \frac{100}{\left(1 + 10^{\left(\frac{\text{Log}(IC_{50}-X)}{\text{HillSlope}}\right)}\right)}, \quad [1]$$

where Y = percent activity, X = the logarithmic concentration of the compound, and Hill Slope = slope factor or Hill coefficient. The IC_{50} value was determined by the concentration resulting in a half-maximal percent activity. The IC_{50} values measured for AdoHcy, which served as a reference compound, are similar to those reported (13). Full IC_{50} curves are presented in *SI Appendix, IC₅₀ curves*.

CARM1 Cloning, Expression, and Purification. The *Mus musculus* CARM1 gene sequence corresponding to the PRMT core (residues 130–487, mmCARM1₁₃₀₋₄₈₇) was amplified by PCR from the original GST-CARM1 construct (15, 31) (see *SI Appendix, Materials and Methods* for full experimental details).

X-Ray Data Collection and Structure Determination.

Crystallization. Transition state mimics 3–6 were solubilized in water before addition to the protein solution (2 mg·mL⁻¹) at the final concentration of 2 mM. The protein–peptide solution was incubated 20 min at room temperature before use. Vapor diffusion method using hanging drop trays with a 0.5-mL reservoir was used for crystallization. Typically, 2 μL of protein-ligand solution were added to 1 μL of well solution consisting of 14–16% (v/v) PEG 3350, 100 mM Tris-HCl pH 8.5, and 200 mM (NH₄)₂SO₄. Crystals grew in a few days at 293 K.

Data collection and structure solution. Crystals were flash-frozen in liquid nitrogen after a brief transfer to 5-μL reservoir solution containing 15% (vol/vol) PEG 400 as a cryoprotectant and were stored in liquid nitrogen. The diffraction datasets were collected on the SOLEIL PROXIMA1 and ESRF ID-29 beamlines, using a PILATUS 6M (Dectris) or a Quantum 315 (ADSC) detector and processed with XDS (32) and HKL-2000 (33). The crystals belonged to the P2₁2₁2 space group with four monomers of CARM1 in the asymmetric unit. The structures were solved by molecular replacement using CARM1 structure as a probe (15).

Model building and refinement were carried out by using Coot (34) and PHENIX (35). TLS refinement with three groups per polypeptide chain was used. All other crystallographic calculations were carried out with the CCP4 package (36). Structure figures were generated with PyMol.

Data availability. All structures and structure factors have been deposited in the Protein Data Bank (CARM1-Compound 3 = 5LGP; CARM1-Compound 4 = 5LQG; CARM1-Compound 5 = 5LGR; CARM1-Compound 6 = 5LGS).

Additional information. Supplementary information, chemical compound information, and source data are available in *SI Appendix, Materials and Methods*.

ACKNOWLEDGMENTS. We thank Javier Sastre Toraño for high-resolution mass spectrometry analysis, the members of the Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC) common services, and the members of the structural biology platform of IGBMC for technical assistance; Vincent Cura and Luc Bonnefond for help in the structural studies, and members of SOLEIL Proxima1 beamlines and the European Synchrotron Radiation Facility–European Molecular Biology Laboratory joint Structural Biology groups for the use of beamline facilities and for assistance during X-ray data collection. Financial support provided by Utrecht University and the Netherlands Organization for Scientific Research is acknowledged (M.J.v.H. and N.I.M.). J.C., N.M., and N.T.-C. are supported by grants from CNRS; Université de Strasbourg; INSERM; Instruct, part of the European Strategy Forum on Research Infrastructures supported by national member subscriptions and the French Infrastructure for Integrated Structural Biology (ANR-10-INSB-05-01); Grant ANR-10-LABX-0030-INRT, a French State fund managed by the Agence Nationale de la Recherche under the frame program Investissements d’Avenir labeled ANR-10-IDEX-0002-02 and Association pour la Recherche contre le Cancer Grant SFI20121205902. G.S. is supported by grants from the Italian Ministero dell’ Istruzione, dell’ Università e della Ricerca (MIUR), Progetti di Ricerca di Interesse Nazionale (PRIN) Grants 2012ZH9YH and 20152TE5PK, and from the Università di Salerno (Italy). G.S. and N.I.M. are supported by European Cooperation in Science and Technology (COST, Action CM1406). A.C. was supported by a Short-Term Scientific Mission Grant from the COST Action CM1406.

- Fuhrmann J, Clancy KW, Thompson PR (2015) Chemical biology of protein arginine modifications in epigenetic regulation. *Chem Rev* 115:5413–5461.
- Bedford MT, Clarke SG (2009) Protein arginine methylation in mammals: Who, what, and why. *Mol Cell* 33:1–13.
- Yang Y, Bedford MT (2013) Protein arginine methyltransferases and cancer. *Nat Rev Cancer* 13:37–50.
- Cha B, Jho EH (2012) Protein arginine methyltransferases (PRMTs) as therapeutic targets. *Expert Opin Ther Targets* 16:651–664.
- Wei H, Mundade R, Lange KC, Lu T (2014) Protein arginine methylation of non-histone proteins and its role in diseases. *Cell Cycle* 13:32–41.
- Morettn A, Baldwin RM, Côté J (2015) Arginine methyltransferases as novel therapeutic targets for breast cancer. *Mutagenesis* 30:177–189.
- Cheng H, et al. (2013) Overexpression of CARM1 in breast cancer is correlated with poorly characterized clinicopathologic parameters and molecular subtypes. *Diagn Pathol* 8:129.
- Ou CY, et al. (2011) A coactivator role of CARM1 in the dysregulation of β-catenin activity in colorectal cancer cell growth and gene expression. *Mol Cancer Res* 9:660–670.
- Kim YR, et al. (2010) Differential CARM1 expression in prostate and colorectal cancers. *BMC Cancer* 10:197.
- Hong H, et al. (2004) Aberrant expression of CARM1, a transcriptional coactivator of androgen receptor, in the development of prostate carcinoma and androgen-independent status. *Cancer* 101:83–89.
- Osada S, et al. (2013) Elevated expression of coactivator-associated arginine methyltransferase 1 is associated with early hepatocarcinogenesis. *Oncol Rep* 30:1669–1674.
- Cura V, Troffer-Charlier N, Wurtz JM, Bonnefond L, Cavarelli J (2014) Structural insight into arginine methylation by the mouse protein arginine methyltransferase 7: A zinc finger freezes the mimic of the dimeric state into a single active site. *Acta Crystallogr D Biol Crystallogr* 70:2401–2412.
- van Haren M, van Ufford LQ, Moret EE, Martin NI (2015) Synthesis and evaluation of protein arginine N-methyltransferase inhibitors designed to simultaneously occupy both substrate binding sites. *Org Biomol Chem* 13:549–560.
- Osborne T, Roska RL, Rajski SR, Thompson PR (2008) In situ generation of a bisubstrate analogue for protein arginine methyltransferase 1. *J Am Chem Soc* 130:4574–4575.
- Troffer-Charlier N, Cura V, Hassenboehler P, Moras D, Cavarelli J (2007) Functional insights from structures of coactivator-associated arginine methyltransferase 1 domains. *EMBO J* 26:4391–4401.
- Yue WW, Hassler M, Roe SM, Thompson-Vale V, Pearl LH (2007) Insights into histone code syntax from structural and biochemical studies of CARM1 methyltransferase. *EMBO J* 26:4402–4412.
- Sack JS, et al. (2011) Structural basis for CARM1 inhibition by indole and pyrazole inhibitors. *Biochem J* 436:331–339.
- Boriack-Sjodin PA, et al. (2016) Structural insights into ternary complex formation of human CARM1 with various substrates. *ACS Chem Biol* 11:763–771.
- Lee J, Bedford MT (2002) PABP1 identified as an arginine methyltransferase substrate using high-density protein arrays. *EMBO Rep* 3:268–273.
- Martin NI, Liskamp RM (2008) Preparation of N(G)-substituted L-arginine analogues suitable for solid phase peptide synthesis. *J Org Chem* 73:7849–7851.
- Lakowski TM, t Hart P, Ahern CA, Martin NI, Frankel A (2010) Nη-substituted arginyl peptide inhibitors of protein arginine N-methyltransferases. *ACS Chem Biol* 5:1053–1063.
- t Hart P, Lakowski TM, Thomas D, Frankel A, Martin NI (2011) Peptidic partial bi-substrates as inhibitors of the protein arginine N-methyltransferases. *Chembiochem* 12:1427–1432.
- Troffer-Charlier N, Cura V, Hassenboehler P, Moras D, Cavarelli J (2007) Expression, purification, crystallization and preliminary crystallographic study of isolated modules of the mouse coactivator-associated arginine methyltransferase 1. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 6:330–333.
- Antonyamy S, et al. (2012) Crystal structure of the human PRMT5:MEP50 complex. *Proc Natl Acad Sci USA* 109:17960–17965.
- Wang C, et al. (2014) Structural determinants for the strict monomethylacylation activity by trypanosoma brucei protein arginine methyltransferase 7. *Structure* 22:756–768.
- Banerjee A, Yang W, Karplus M, Verdine GL (2005) Structure of a repair enzyme interrogating undamaged DNA elucidates recognition of damaged DNA. *Nature* 434:612–618.
- Yang M, et al. (2007) Structural basis of histone demethylation by LSD1 revealed by suicide inactivation. *Nat Struct Mol Biol* 14:535–539.
- Mailliot J (2013) Étude structurale de l’histoneméthyltransférase “CARM1” et de ses complexes biologiquement significatifs: Des structures 3D vers la conception rationnelle de composés à action pharmacologique. PhD thesis (University of Strasbourg, Strasbourg, France).
- Bonnefond L, et al. (2015) Functional insights from high resolution structures of mouse protein arginine methyltransferase 6. *J Struct Biol* 191:175–183.
- Vedadi M, et al. (2011) A chemical probe selectively inhibits G9a and GLP methyltransferase activity in cells. *Nat Chem Biol* 7:566–574.
- Chen D, et al. (1999) Regulation of transcription by a protein methyltransferase. *Science* 284:2174–2177.
- Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* 66:125–132.
- Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
- Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132.
- Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221.
- Winn MD, et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67:235–242.

Supplementary Information for:

Transition state mimics are valuable mechanistic probes for structural studies with the arginine methyltransferase CARM1

Matthijs J. van Haren^{a,1}, Nils Marechal^{b,1}, Nathalie Troffer-Charlier^b, Agostino Cianciulli^c, Gianluca Sbardella^c, Jean Cavarell^{b,2}, and Nathaniel I. Martin^{a,2}

^aDepartment of Chemical Biology & Drug Discovery, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Universiteitsweg 99, 3584 CG Utrecht, The Netherlands.

^bDepartment of Integrated Structural Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Université de Strasbourg, CNRS UMR7104, INSERM U 964, 1 rue Laurent Fries, Illkirch, F-67404, France.

^cEpigenetic Med Chem Lab, Dipartimento di Farmacia, Università degli Studi di Salerno, Via Giovanni Paolo II 132, I-84084 Fisciano (SA), Italy

¹These authors contributed equally to this work.

²Correspondence and requests for materials should be addressed to J.C. (cava@igbmc.fr) and N.I.M. (n.i.martin@uu.nl)

Table of Contents

Page

S2-11	Experimental Procedures
S12-13	IC ₅₀ Curves for AdoHcy and compounds 1-6
S14-S15	CARM1 cloning, expression, and purification
S15-S17	Supplemental table and figures for structural studies
S18-21	¹ H and ¹³ C NMR spectra for compounds 10-12
S22-25	HPLC traces for compounds 1-8
S26	References

Experimental Procedures

Reagents and general methods

All reagents employed were of American Chemical Society (ACS) grade or finer and were used without further purification unless otherwise stated. The amine precursors **9-11** were synthesized according to previously described procedures (1). All known compounds prepared had NMR spectra and HRMS data consistent with the assigned structures. All reactions and fractions from column chromatography were monitored by thin layer chromatography (TLC) using plates with a UV fluorescent indicator (normal SiO₂, Merck 60 F²⁵⁴). One or more of the following methods were used for visualization: UV absorption by fluorescence quenching; phosphomolybdic acid: ceric sulfate: sulfuric acid: H₂O (10 g:1.25 g: 12 mL:238 mL) staining; KMnO₄ staining; PPh₃ staining; ninhydrin staining.

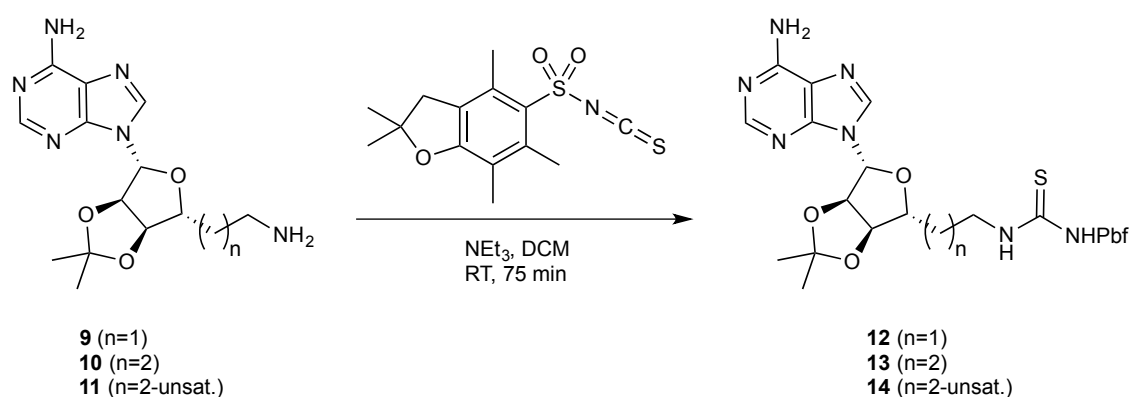
Flash chromatography was performed using Merck type 60, 230–400 mesh silica gel. The final compounds **1-8** were purified by preparative scale RP-HPLC using a Reprosil-Pur C18 column (10 μm, 250 × 22 mm) eluted with a water–acetonitrile gradient moving from 0% to 50% acetonitrile (0.1% TFA) over 60 minutes at a flow-rate of 12.0 mL·min⁻¹ with UV detection at 214 nm. Purity was confirmed to be ≥95% by analytical RP-HPLC using a Phenomenex Kinetex C18 column (5 μm, 250 × 4.6 mm) eluted with a water–acetonitrile gradient moving from 0% to 50% CH₃CN (0.1% TFA) over 30 minutes at a flow rate of 1.0 mL·min⁻¹ with UV detection at 214 nm and 254 nm.

Instrumentation for compound characterization

¹H NMR spectra were recorded at 400 MHz or 500 MHz with chemical shifts reported in parts per million (ppm) downfield relative to tetramethylsilane (TMS). ¹H NMR data are reported in the following order: multiplicity (s, singlet; d, doublet; t, triplet; q, quartet and m, multiplet), coupling constant (*J*) in hertz (Hz) and the number of protons. When appropriate, the multiplicity is preceded by br, indicating that the signal was broad. ¹³C NMR spectra were

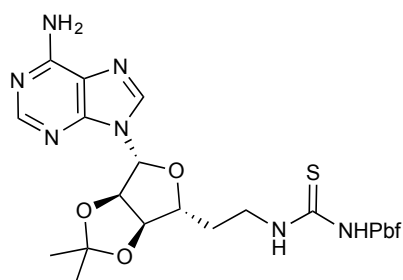
recorded at 101 MHz with chemical shifts reported relative to CDCl₃ (δ 77.16). High-resolution mass spectrometry (HRMS) analysis was performed using a Bruker micrOTOF instrument.

Synthetic procedures and compound characterization



Supplemental Scheme S1. Synthesis of Pbf-protected adenosine-thioureas

N-((2-((3*aR*,4*R*,6*R*,6*aR*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)ethyl)carbamothioyl)-2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-sulfonamide (**12**)

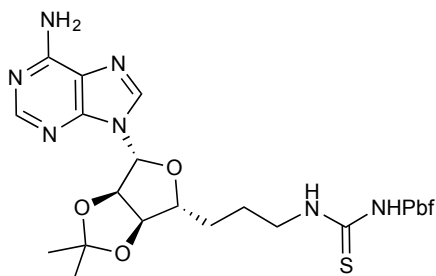


Adenosine ethylamine **9** (300 mg, 0.78 mmol, 1.0 eq) was dissolved in DCM (7.8 mL) at 0 °C, triethylamine was added (0.43 mL, 3.1 mmol, 4.0 eq), followed by dropwise addition of a 0.1 M solution of Pbf-NCS² (11.8 mL, 1.18 mmol, 1.5 eq) in DCM. After 75 min, the mixture was diluted with DCM, washed with water (3x) and brine (1x), dried with Na₂SO₄ and concentrated. The crude product was purified by column chromatography (2-4% MeOH

gradient in DCM) yielding **12** (287 mg, 0.45 mmol, 58%) as a white solid.

Analytical data: ^1H NMR (500 MHz, CDCl_3) δ 8.36 (s, 1H), 8.10 (br s, 1H), 7.96 (s, 1H), 6.03 (s, 1H), 5.87 (br s, 2H), 5.49 (s, 1H), 4.95 (s, 1H), 4.20 (t, $J = 4.1$ Hz, 1H), 3.76-3.62 (m, 1H), 3.62-3.44 (m, 1H), 2.91 (s, 2H), 2.45 (s, 3H), 2.35 (s, 3H), 2.06 (s, 3H), 2.13-1.90 (m, 2H), 1.60 (s, 3H), 1.45 (s, 6H), 1.38 (s, 3H); ^{13}C NMR (126 MHz, CDCl_3 , extracted from HSQC and HMBC) δ 160.8, 155.3, 153.0, 149.0, 140.6, 139.7, 133.8, 126.8, 125.4, 118.2, 114.3, 90.0, 87.2, 85.8, 83.9, 83.6, 43.0, 42.9, 31.3, 28.5, 27.1, 25.2, 18.9, 17.3, 12.2; HRMS (m/z): $[\text{M}]^+$ calcd. for $\text{C}_{28}\text{H}_{38}\text{N}_7\text{O}_6\text{S}_2$, 632.2319, found 632.2342.

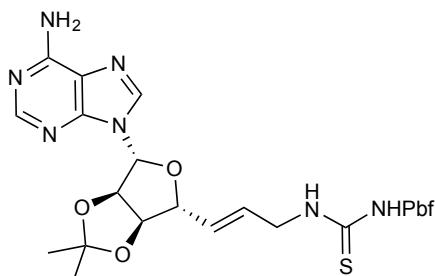
***N*-((3-((3*aR*,4*R*,6*R*,6*aR*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)propyl)carbamothioyl)-2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-sulfonamide (13)**



Adenosine propylamine **10** (334 mg, 1.0 mmol, 1.0 eq) was dissolved in DCM (10 mL) at 0°C, triethylamine was added (0.277 mL, 2.0 mmol, 2.0 eq), followed by dropwise addition of a 0.1 M solution of Pbf-NCS (10.0 mL, 1.0 mmol, 1.0 eq) in DCM. After 90 min, the mixture was diluted with DCM, washed with water (3x) and brine (1x), dried with Na_2SO_4 and concentrated. The crude product was purified by column chromatography (2-4% MeOH gradient in DCM) yielding **13** (380 mg, 0.59 mmol, 59%) as a white solid.

Analytical data: ^1H NMR (400 MHz, CDCl_3) δ 8.38 (s, 1H), 8.04 (s, 1H), 7.92 (t, $J = 5.3$ Hz, 1H), 6.39 (br s, 2H), 6.03 (d, $J = 2.3$ Hz, 1H), 5.42 (dd, $J = 6.5, 2.2$ Hz, 1H), 4.79 (dd, $J = 6.4, 4.1$ Hz, 1H), 4.13-4.05 (m, 1H), 3.60-3.41 (m, 2H), 2.91 (s, 2H), 2.52 (s, 3H), 2.43 (s, 3H), 2.05 (s, 3H), 1.71-1.50 (m, 4H), 1.61 (s, 3H), 1.43 (s, 3H), 1.37 (s, 3H); ^{13}C NMR (101 MHz, CDCl_3) δ 178.9, 160.8, 155.7, 153.5, 149.4, 140.4, 140.2, 134.4, 127.1, 125.8, 120.0, 118.8, 114.9, 90.2, 87.5, 86.3, 84.2, 84.1, 45.3, 42.9, 30.5, 28.6, 27.3, 25.6, 24.8, 19.4, 17.7, HRMS (m/z): $[\text{M}]^+$ calcd. for $\text{C}_{29}\text{H}_{40}\text{N}_7\text{O}_6\text{S}_2$, 646.2476, found 646.2446

***N*-(((*E*)-3-((3*aR*,4*R*,6*R*,6*aR*)-6-(6-amino-9*H*-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-*d*][1,3]dioxol-4-yl)allyl)carbamoithiyl)-2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-sulfonamide (14)**

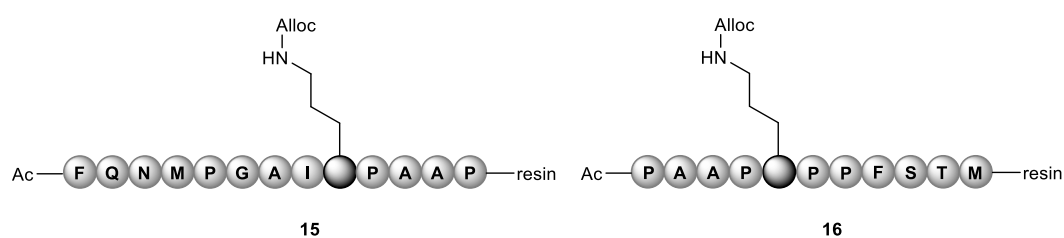


Adenosine propenylamine **11** (165 mg, 0.5 mmol, 1.0 eq) was dissolved in DCM (5 mL) at 0 °C, triethylamine was added (0.14 mL, 1.0 mmol, 2.0 eq), followed by dropwise addition of a 0.1 M solution of Pbf-NCS (5.0 mL, 0.5 mmol, 1.0 eq) in DCM. After 90 min, TLC showed full conversion and the mixture was concentrated. The crude product was purified by column chromatography (2-4% MeOH gradient in DCM) yielding **14** (220 mg, 0.34 mmol, 68%) as a clear oil.

Analytical data: ^1H NMR (400 MHz, CDCl_3) δ 8.40 (s, 1H), 8.05 (s, 1H), 7.90 (t, $J = 5.4$ Hz, 1H), 6.56 (s, 2H), 6.11 (d, $J = 1.6$ Hz, 1H), 5.72 (dd, $J = 15.6, 6.9$ Hz, 1H), 5.57 (dt, $J = 15.5, 5.7$ Hz, 1H), 5.45 (dd, $J = 6.3, 1.6$ Hz, 1H), 4.95 (dd, $J = 6.2, 3.4$ Hz, 1H), 4.62 (dd, $J = 6.7,$

3.1 Hz, 1H), 4.17-3.98 (m, 2H), 2.90 (s, 2H), 2.51 (s, 3H), 2.41 (s, 3H), 2.04 (s, 3H), 1.60 (s, 3H), 1.46 – 1.40 (m, 6H), 1.37 (s, 3H); ^{13}C NMR (101 MHz, CDCl_3) δ 179.2, 160.8, 155.8, 153.5, 149.3, 140.4, 140.3, 134.4, 130.4, 128.0, 127.0, 125.8, 119.8, 118.7, 114.6, 90.5, 87.5, 84.6, 84.5, 46.6, 42.9, 28.6, 27.2, 25.5, 19.4, 17.6, 12.6; HRMS (m/z): $[\text{M}]^+$ calcd. for $\text{C}_{29}\text{H}_{38}\text{N}_7\text{O}_6\text{S}_2$, 644.2319, found 644.2284

Synthesis of Alloc-protected ornithine precursor peptides

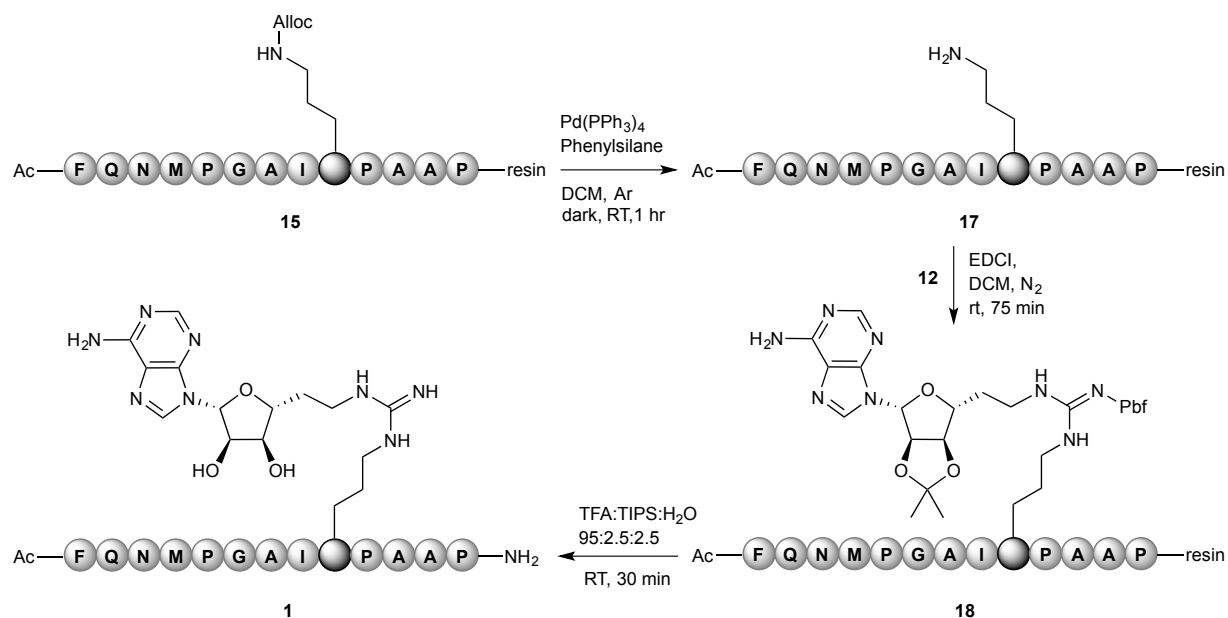


Supplemental Figure S1. Structures of the Alloc-protected ornithine containing peptides

The Alloc-protected ornithine-containing peptides **15** and **16** were synthesized using a Symphony peptide synthesizer on Rinkamide tentagel S resin (455 mg with a resin loading of 0.22 mmol/g) using standard Fmoc SPPS protocols. Each peptide was synthesized in three batches of 0.10 mmol and peptide couplings were performed by using protected Fmoc amino acid (4.0 eq), BOP reagent (4.0 eq), and DIPEA (8.0 eq) in DMF (total volume 5 mL) at ambient temperature for 1 hour.

Upon completion of SPPS, the N-terminal was acetylated on resin using acetic anhydride (100 eq, 0.94 mL) and DiPEA (100 eq, 1.74 mL) in DMF (10 mL) for 1 hour at room temperature under nitrogen atmosphere. The resin was then flushed with DMF (4 x 10 mL) and DCM (2 x 10 mL). The peptides were not cleaved from the resin, but used directly in the next step.

On-resin synthesis of adenosine-peptide conjugates



Supplemental Scheme S2. Representative synthetic scheme for the final adenosine-peptide conjugates, presented for compound **1**.

Alloc-protected peptide **15** was deprotected on the resin using tetrakis(triphenylphosphine)-palladium(0) and phenylsilane in DCM following a literature procedure³. Subsequently, compound **12** (0.15 mmol, 1.5 eq) was coupled to the ornithine side-chain of the peptide (0.1 mmol, 1.0 eq) using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDCI) (0.15 mmol, 1.5 eq) in DCM (10 mL). The mixture was swirled under nitrogen for 1 hour. The mixture was drained and the resin was washed with DCM (5 × 10 mL), DMF (5 × 10 mL) and DCM (2 × 10 mL). Protected peptide-adenosine conjugate **18** was deprotected and cleaved from the resin using TFA/TIPS/H₂O (95:2.5:2.5). Precipitation in MTBE/hexanes (1:1) yielded the crude peptide, which was purified by preparative RP-HPLC. The purity and identity were confirmed via analytical HPLC and high-resolution mass spectrometry.

Compounds **2-6** were synthesized in a similar fashion starting from the appropriate thiourea and resin-bound peptide.

Analytical data for peptides 1-6:

1: HRMS (m/z): $[M]^+$ calcd. for $C_{74}H_{113}N_{24}O_{19}S$, 1673.8329, found 1673.8369

2: HRMS (m/z): $[M]^+$ calcd. for $C_{66}H_{99}N_{20}O_{17}S$, 1475.7212, found 1475.7251

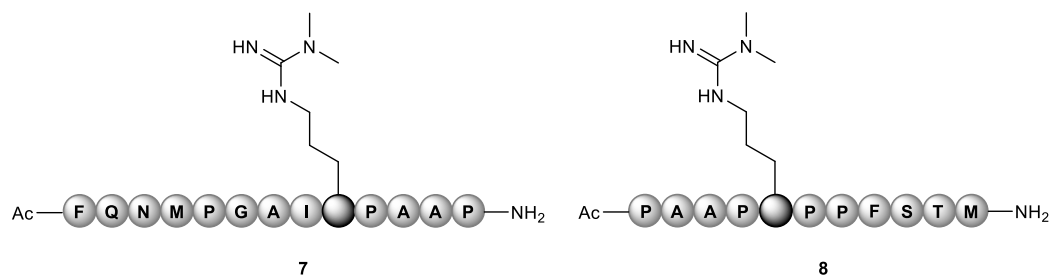
3: HRMS (m/z): $[M]^+$ calcd. for $C_{75}H_{115}N_{24}O_{19}S$, 1687.8486, found 1687.8467

4: HRMS (m/z): $[M]^+$ calcd. for $C_{67}H_{101}N_{20}O_{17}S$, 1489.7369, found 1489.7322

5: HRMS (m/z): $[M]^+$ calcd. for $C_{75}H_{113}N_{24}O_{19}S$, 1685.8329, found 1685.8317

6: HRMS (m/z): $[M]^{2+}$ calcd. for $C_{67}H_{99}N_{20}O_{17}S$, 744.3649, found 744.3656

Synthesis of aDMA-peptides 7 and 8



Supplemental Figure S2. Structure of the asymmetrically dimethylated arginine containing peptides

Asymmetrically dimethylated arginine (aDMA)-containing peptides **7** and **8** were prepared on an automated peptide synthesizer following standard Fmoc SPPS protocols. The required asymmetrically dimethylated arginine building block was prepared as previously described (2).

Peptide couplings were performed by using protected Fmoc amino acids (4.0 eq), BOP reagent (4.0 eq), and DIPEA (8.0 eq) in DMF (total volume 5 mL) at ambient temperature for 1 hour.

Upon completion of SPPS, the N-terminal was acetylated on resin using acetic anhydride (100 eq, 0.94 mL) and DiPEA (100 eq, 1.74 mL) in DMF (10 mL) for 1 hour at room temperature under nitrogen atmosphere. The resin was then flushed with DMF (4 x 10 mL) and DCM (2 x 10 mL) and the peptides were cleaved from the resin and deprotected by using TFA/TIS/H₂O (95:2.5:2.5). Precipitation in MTBE/hexanes (1:1) yielded the crude peptides. Peptides **7** and **8** were purified by preparative RP-HPLC and their purity and identity were confirmed via analytical HPLC and high-resolution mass spectrometry.

Analytical data for peptides 7 and 8:

7: HRMS (*m/z*): [M]⁺ calcd. for C₆₅H₁₀₄N₁₉O₁₆S, 1438.7624, found 1438.7668;

8: HRMS (*m/z*): [M]⁺ calcd. for C₅₇H₉₀N₁₅O₁₄S, 1240.6507, found 1240.6523

Methyltransferase Inhibition Assay

Methyltransferase inhibition assays were performed as previously described (3) using commercially available chemiluminescent assay kits for PRMT1 and CARM1 (purchased from BPS Bioscience, San Diego, CA, USA). The enzymatic reactions were conducted in triplicate at room temperature for 1 hour in substrate-coated well plates at a final reaction volume of 50 μ L containing the manufacturer's proprietary assay buffer, AdoMet (at a concentration of 5-times the respective K_m value for each enzyme), the methyltransferase enzyme: PRMT1 (80 ng per reaction) and CARM1 (200 ng per reaction) and inhibitors **1-6** and AdoHcy (in the range of concentrations: 0.001–250 μ M in water). Before addition of AdoMet, the enzyme was first incubated with the inhibitor for 15 minutes at 37 °C. Positive controls were performed in the absence of inhibitors using water to keep the final volume consistent. Blanks and substrate controls were performed in the absence of the enzyme and AdoMet respectively. Following the enzymatic reactions, 100 μ L of primary antibody (recognizing the respective immobilized asymmetrically dimethylated arginine

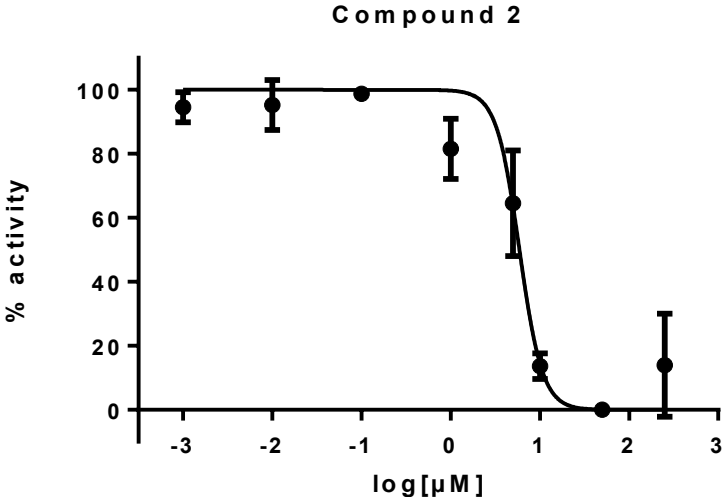
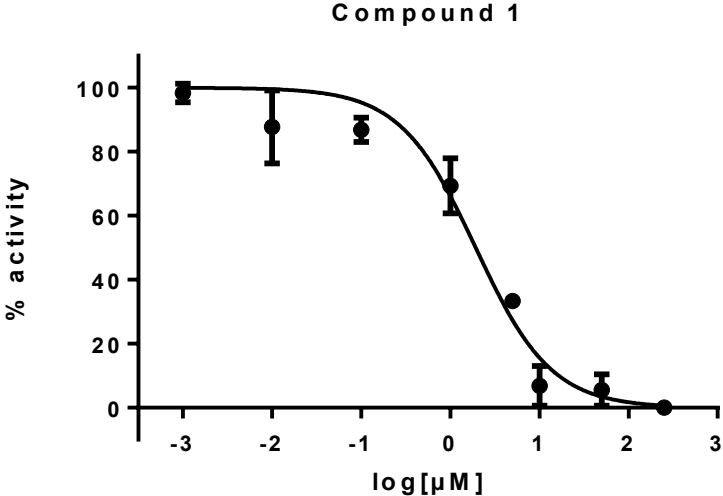
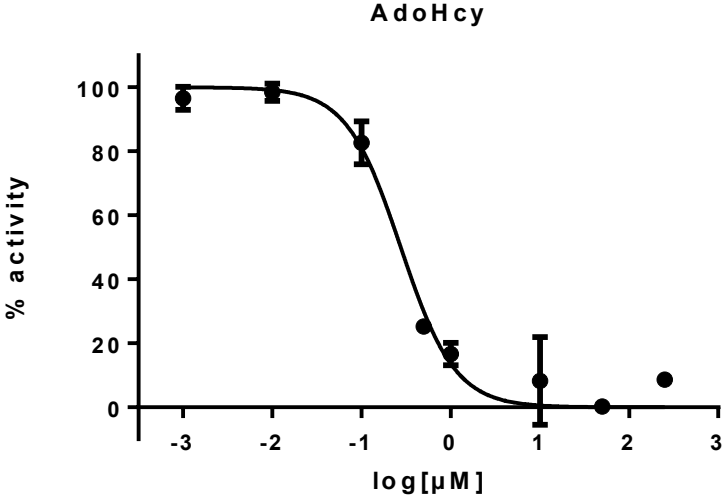
product) was added to each well and the plate was incubated at room temperature for an additional 1 hour. Then 100 μ L of secondary horseradish peroxidase (HRP)-conjugated antibody was added to each well and the plate was incubated at room temperature for additional 30 minutes. Finally, 100 μ L of an HRP substrate mixture was added to the wells and the luminescence was measured directly using a standard microplate reader. In all cases enzyme activity measurements were performed in triplicate at each of the inhibitor concentrations evaluated. The luminescence data were analysed using GraphPad Prism (version 6.02).

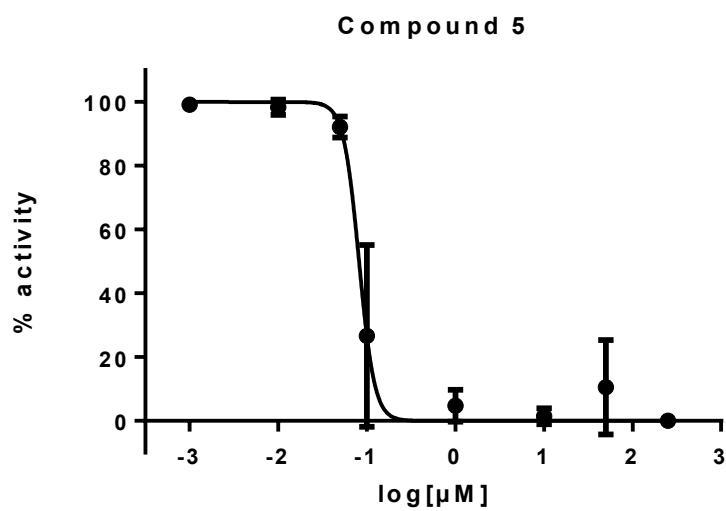
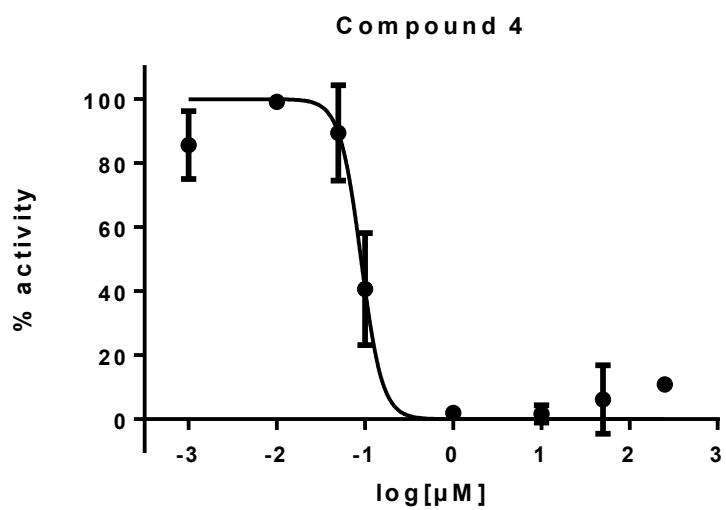
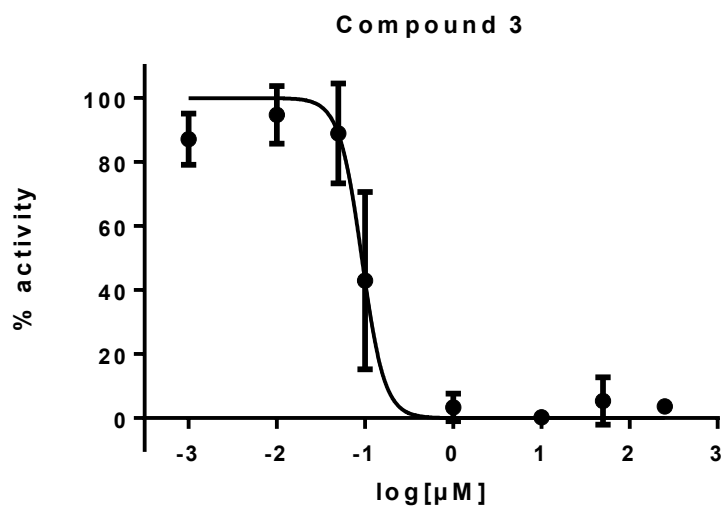
The luminescence data was normalized with the highest value obtained in a concentration range defined as 100% activity and the lowest value defined as 0%. Blank and positive controls were performed to analyse the validity of the assay method. Blank values obtained were generally higher than the lowest value obtained in a concentration range. Positive controls were generally within \pm 15% of the highest values obtained in a concentration range, but occasionally the highest values were up to 33% higher than the positive control. The percent activity values were plotted as a function of inhibitor concentrations and fitted using non-linear regression analysis of the sigmoidal dose–response curve generated using normalized data and a variable slope following equation 1,

$$Y = \frac{100}{(1 + 10^{((\text{Log}IC_{50} - X) * \text{HillSlope}))}} \quad (\text{equation 1})$$

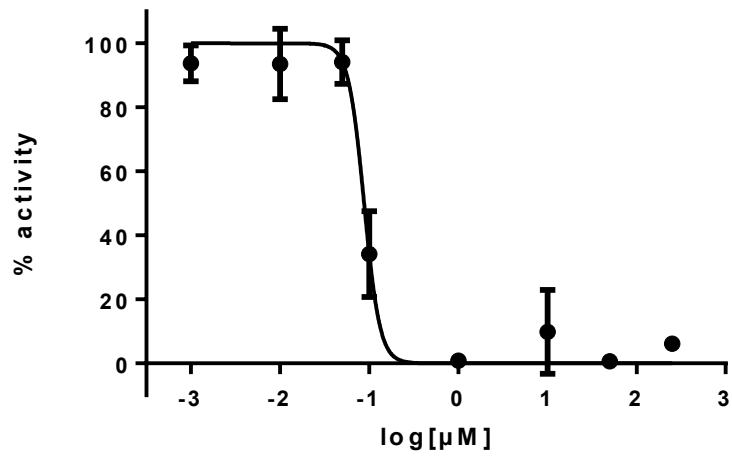
where Y = percent activity, X = the logarithmic concentration of the compound and Hill Slope = slope factor or Hill coefficient. The IC_{50} value was determined by the concentration resulting in a half-maximal percent activity. The standard errors were reported using the symmetrical CI function. The IC_{50} values measured for AdoHcy, which served as a reference compound, are similar to those previously reported (1).

IC₅₀ curves for AdoHcy and compounds 1-6





Compound 6



CARM1 cloning, expression, and purification

The *Mus musculus* CARM1 gene sequence corresponding to the PRMT core (residues 130 to 487, mmCARM1₁₃₀₋₄₈₇) were amplified by PCR from the original GST-CARM1 construct (4). The sequences were cloned in the pDONR207TM (Invitrogen) vector using a BP reaction (Gateway® Cloning, Life Technologies). The positive clones were confirmed by sequencing (GATC). The sequences were subcloned in a pDEST20TM vector using a LR reaction. The resulting recombinant protein harbor an amino-terminal glutathione S-transferase (GST) tag followed by a Tobacco etch virus (TEV) protease cleavage site. DH10Bac competent cells containing the baculovirus genome were transformed with the pDEST20TM-CARM1 plasmids and plated onto LB agar media containing 15 mg·mL⁻¹ tetracycline, 7 mg·mL⁻¹ gentamicin, 50 mg·mL⁻¹ kanamycin, 25 mg·mL⁻¹ X-Gal and 40 mg·mL⁻¹ IPTG. Bacmid DNA purified from recombination-positive white colonies was transfected into Sf9 cells using the Lipofectin reagent (Invitrogen). Viruses were harvested 10 days after transfection. Sf9 cells were grown at 300K in suspension culture in Grace medium (Gibco) using Bellco spinner flasks. 1L of Sf9 cell culture (at 0.8 x 10⁶ cells·mL⁻¹) was infected with recombinant GST-mmCARM1 virus with an infection multiplicity of 1. Cells were harvested 48 h post-infection. Cells lysis was performed by sonication in 50 mL buffer A [50 mM Tris-HCl pH 8.0, 250 mM NaCl, 5% glycerol, 5 mM TCEP, 0.01% NP40 and anti-proteases (Roche, CompleteTM, EDTA-free)] and cellular debris were sedimented by centrifugation of the lysate at 40,000g for 30 min. The supernatant was incubated overnight at 277K with 2 mL glutathione Sepharose resin (GE Healthcare). After a short centrifugation, the supernatants were discarded and the beads were poured in an Econo-column (Bio-Rad). After two wash steps with 10 mL buffer A, 2 mL buffer A supplemented with in-house produced TEV protease were applied to the columns and digestion was performed 4 hours at 303K with gentle mixing. The digest was concentrated with an Amicon Ultra 10K (Milipore), loaded on a gel-filtration column (HiLoad 16/60 Superdex S200, GE

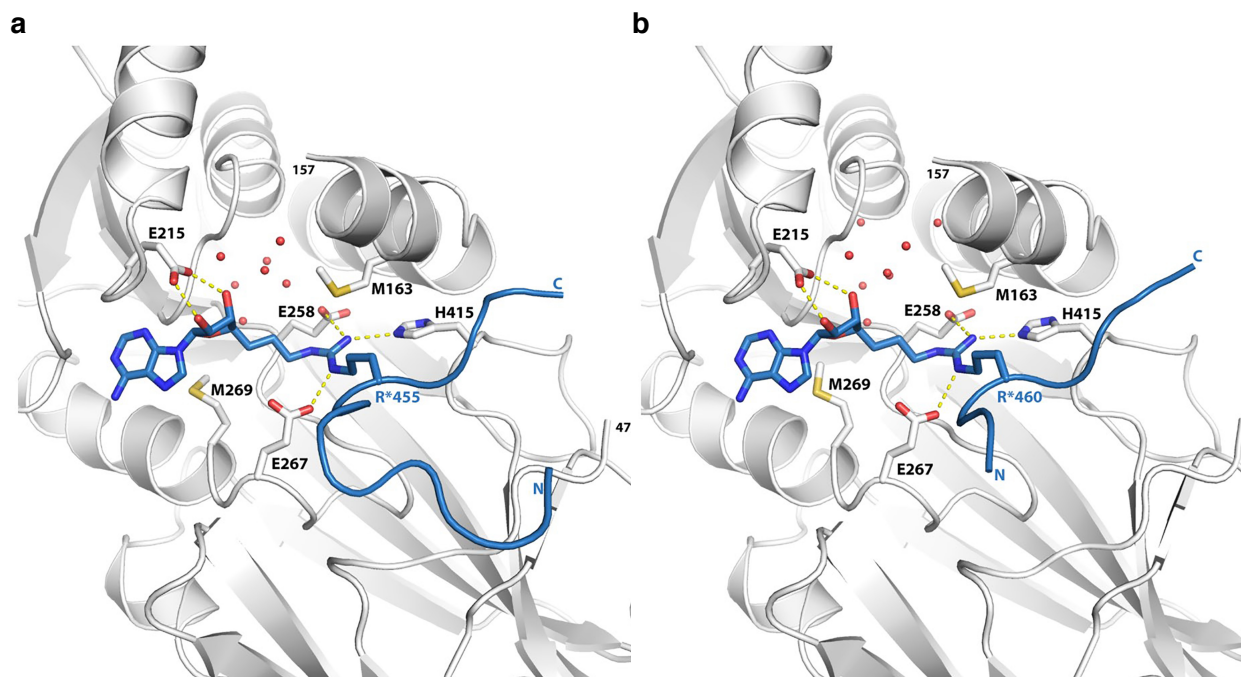
Healthcare) and eluted at 1 mL·min⁻¹ with buffer B [20mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM TCEP] using an ÄKTA Purifier device (GE Healthcare). Fractions containing mmCARM1₁₃₀₋₄₈₇ were pooled and concentrated to 5 mg·mL⁻¹.

Supplemental table and figures for structural studies

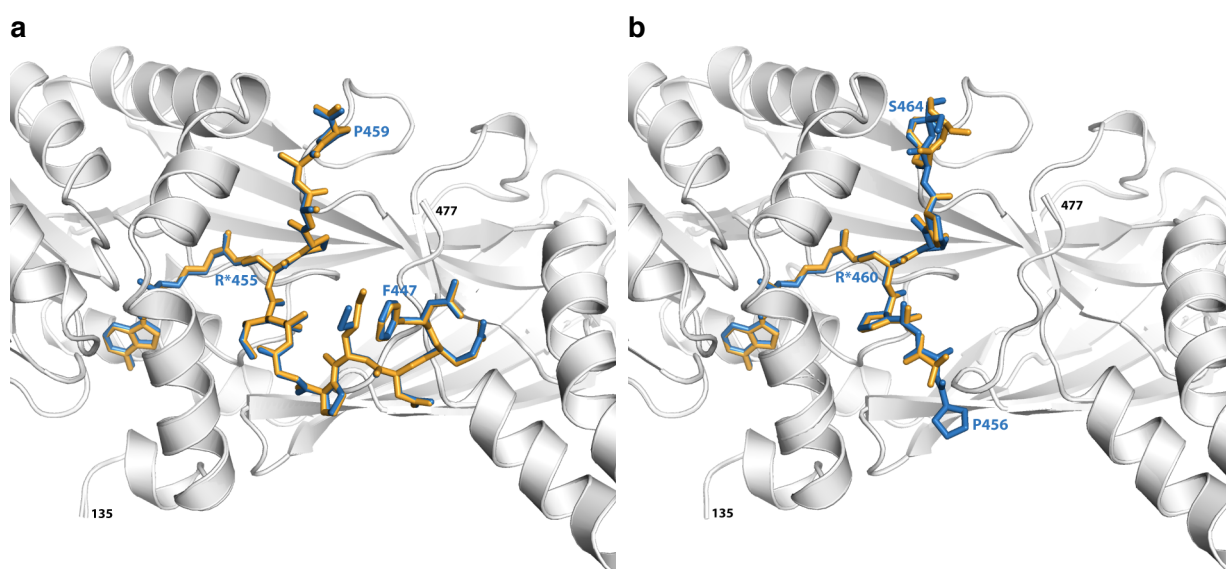
Supplementary Table S1: X-Ray data collection and refinement statistics for mmCARM1 complexes with compounds 3-6

PDB ID	5LGP	5LGR	5LGQ	5LGS
Ligand	Compound 3	Compound 5	Compound 4	Compound 6
Data processing				
Resolution (Å)	48.01 - 2.04 (2.07 - 2.04)	46.05 - 2.00 (2.03 - 2.00)	48.15 - 2.11 (2.15 - 2.11)	47.97 - 2.10 (2.14 - 2.10)
Space group	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2
Cell parameters				
a - b - c (Å)	74.8 98.7 207.3	74.8 98.5 208.3	74.7 99.0 207.3	75.0 98.6 208.2
α - β - γ (°)	90 90 90	90 90 90	90 90 90	90 90 90
Total reflections	654505 (31049)	700081 (34013)	581767 (21071)	608599 (30695)
Unique reflections	98498 (4845)	104813 (5156)	89266 (4491)	90877 (4456)
R _{merge} or R _{sym}	0.148 (2.331)	0.174 (3.226)	0.150 (2.021)	0.132 (2.753)
R _{meas} *	0.161 (2.538)	0.189 (3.500)	0.163 (2.281)	0.143 (2.978)
R _{pim}	0.062 (0.997)	0.073 (1.345)	0.063 (1.030)	0.055 (1.130)
I/σ ₁	9.7 (0.8)	8.0 (0.6)	9.0 (0.7)	10.4 (0.7)
CC _{1/2}	0.998 (0.313)	0.998 (0.308)	0.998 (0.325)	0.999 (0.320)
Completeness (%)	100 (100)	100 (99.9)	99.9 (99.6)	100 (99.9)
Multiplicity	6.6 (6.4)	9.7 (6.6)	6.5 (4.7)	6.7 (6.9)
Wilson B (Å ²)	31.7	29.7	32.7	37.0
Resolution limit for I/σ(I) > 2.0	2.31	2.38	2.40	2.38
Refinement				
Resolution (Å)	48.01 - 2.04 (2.06 - 2.04)	46.05 - 2.00 (2.02 - 2.00)	48.15 - 2.11 (2.13 - 2.11)	47.97 - 2.10 (2.12 - 2.10)
R _{work} (%)	18.73 (33.18)	19.20 (36.23)	18.66 (35.59)	18.63 (33.86)
R _{free} (%)	21.93 (33.82)	23.37 (38.86)	23.60 (34.02)	21.47 (37.70)
Number of non H atoms				
protein	11001	11063	11128	11068
compound	448	423	323	323
ligand	20	60	85	55
water	755	765	839	584
Validation				
RMSD				
bond lengths (Å)	0.006	0.002	0.002	0.002
bond angles (°)	0.691	0.540	0.530	0.516
Ramachandran favored (%)	97.08	96.87	96.86	96.85
Ramachandran outliers (%)	0.0	0	0.07	0
Rotamer outliers (%)	0.25	0.33	0.33	0.49
Average B (Å ²) (all atoms)	41.0	42.0	46.0	51.0

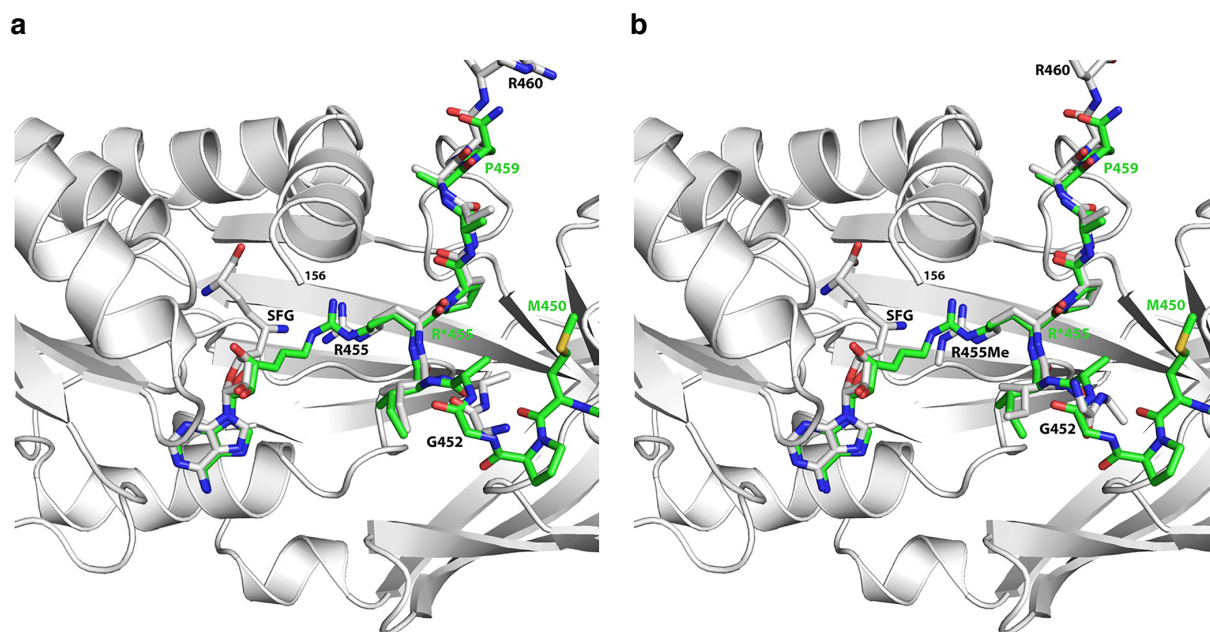
Statistics for the highest-resolution shell are shown in parentheses. *See reference (5) for crystallographic definitions. The resolution limits for I/σ(I) > 2 are reported.



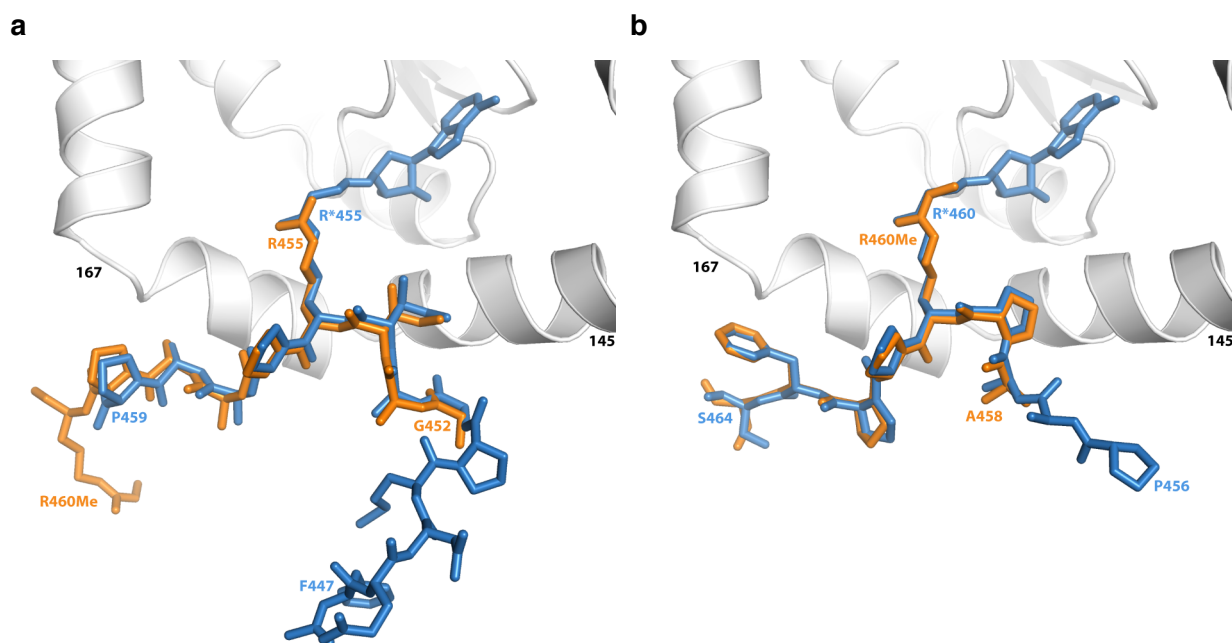
Supplementary Figure S3. Co-crystal structures of CARM1 with transition state mimics **3** and **4** revealing interacting active site residues and water molecules. Cartoon and stick representation of the transition state mimics bound to mmCARM1. (a) Compound **3** is derived from PABP1₄₄₇₋₄₅₉ and (b) Compound **4** is derived from PABP1₄₅₆₋₄₆₆. Both compounds contain a saturated linker between the guanidine moiety and the adenosine unit. For clarity, N-terminal helices of CARM1 are not shown.



Supplementary Figure S4. Superimposition of saturated and unsaturated peptides bound to subunit B of mmCARM1. (a) Superimposition of compound **5** (blue sticks) and compound **3** (orange sticks) bound to mmCARM1. (b) Superimposition of compound **6** (blue sticks) and compound **4** (orange sticks) bound to mmCARM1



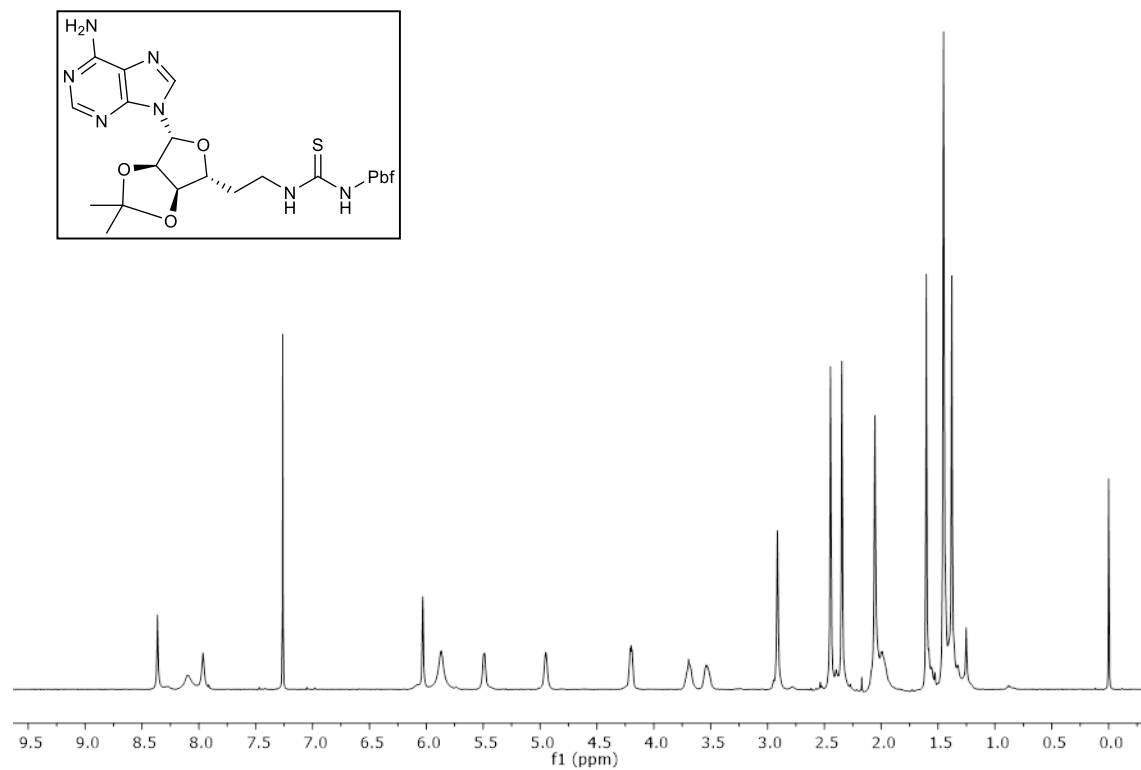
Supplementary Figure S5. Structures of transition state mimics overlaid with isolated peptides in presence of SFG. (a) Superimposition of compound 5 (green sticks) bound to mmCARM1 and SFG-PABP1(R455) (white sticks) bound to hsCARM1 (PDB code 5DX1). (b) Superimposition of compound 5 (green sticks) bound to mmCARM1 and SFG-PABP1(R455-MMA) (white sticks) bound to hsCARM1 (PDB code 5DX8). SFG molecule is shown for hsCARM1. A similar conformation is obtained using compound 3, compound 4 and compound 6 (views not shown).



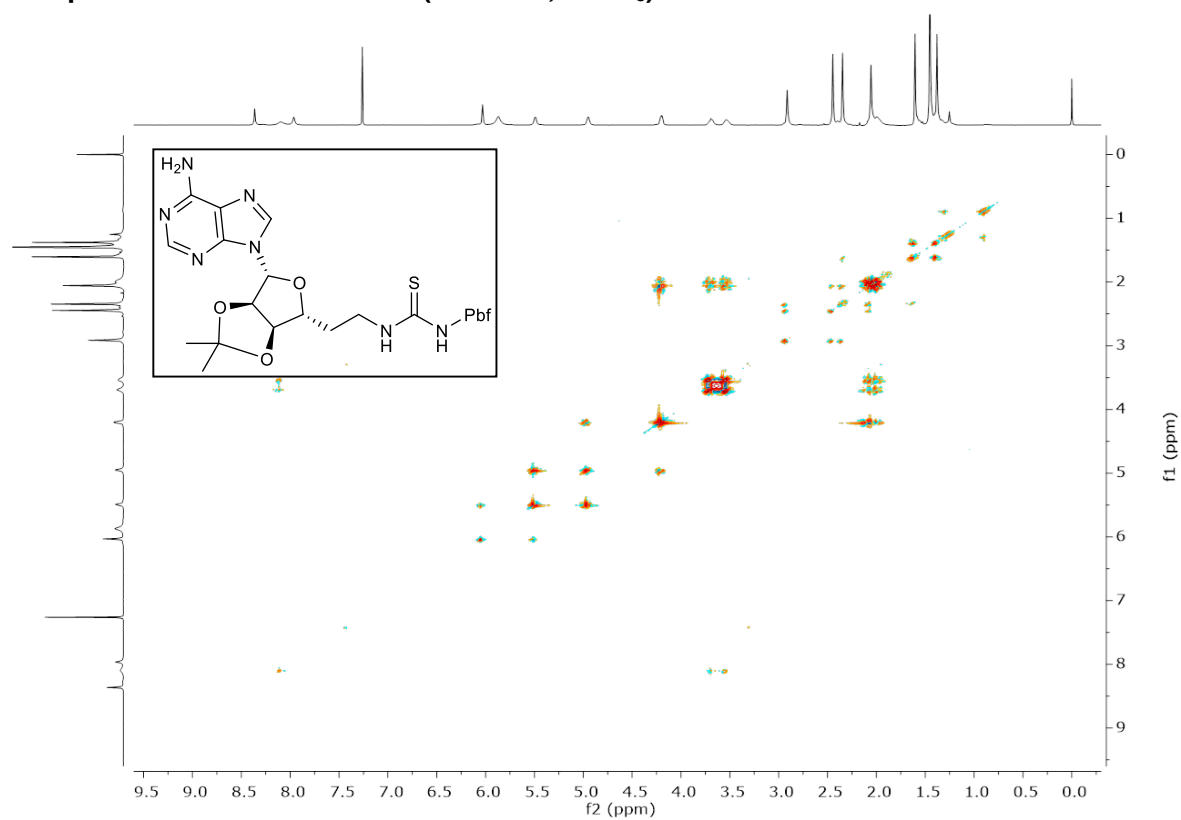
Supplementary Figure S6. Superimposition of transition state analogues 3 and 4 and free peptides on CARM1. (a) Superimposition of compound 3 (blue sticks) bound to mmCARM1 to SFG-PABP1(R455) (orange sticks) bound to hsCARM1 (PDB code 5DX1). SFG molecule is not shown for hsCARM1. (b) Superimposition of compound 4 (blue sticks) bound to mmCARM1 to SFG-PABP1(R460-MMA) (orange sticks) bound to hsCARM1 (PDB code 5DXA). SFG molecule is not shown for hsCARM1.

NMR spectra

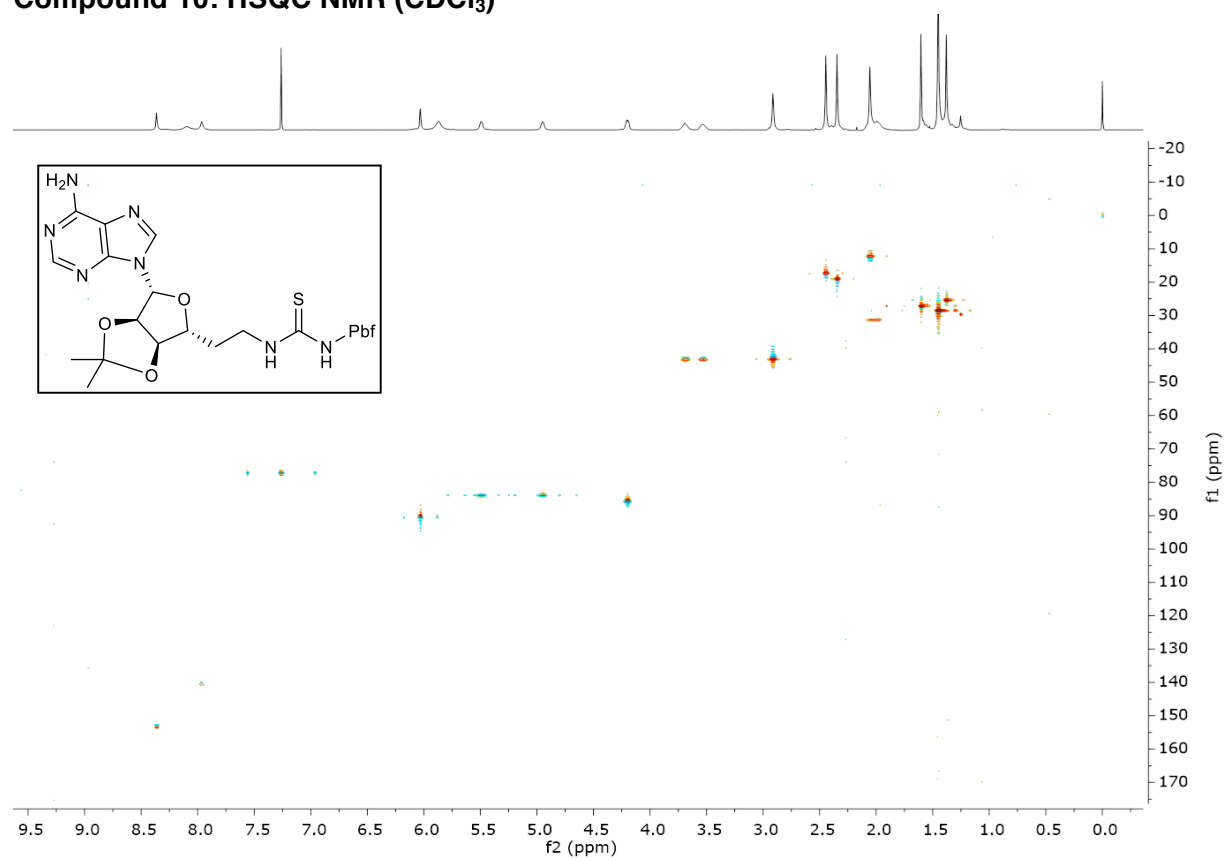
Compound 10: ^1H NMR (500 MHz, CDCl_3)



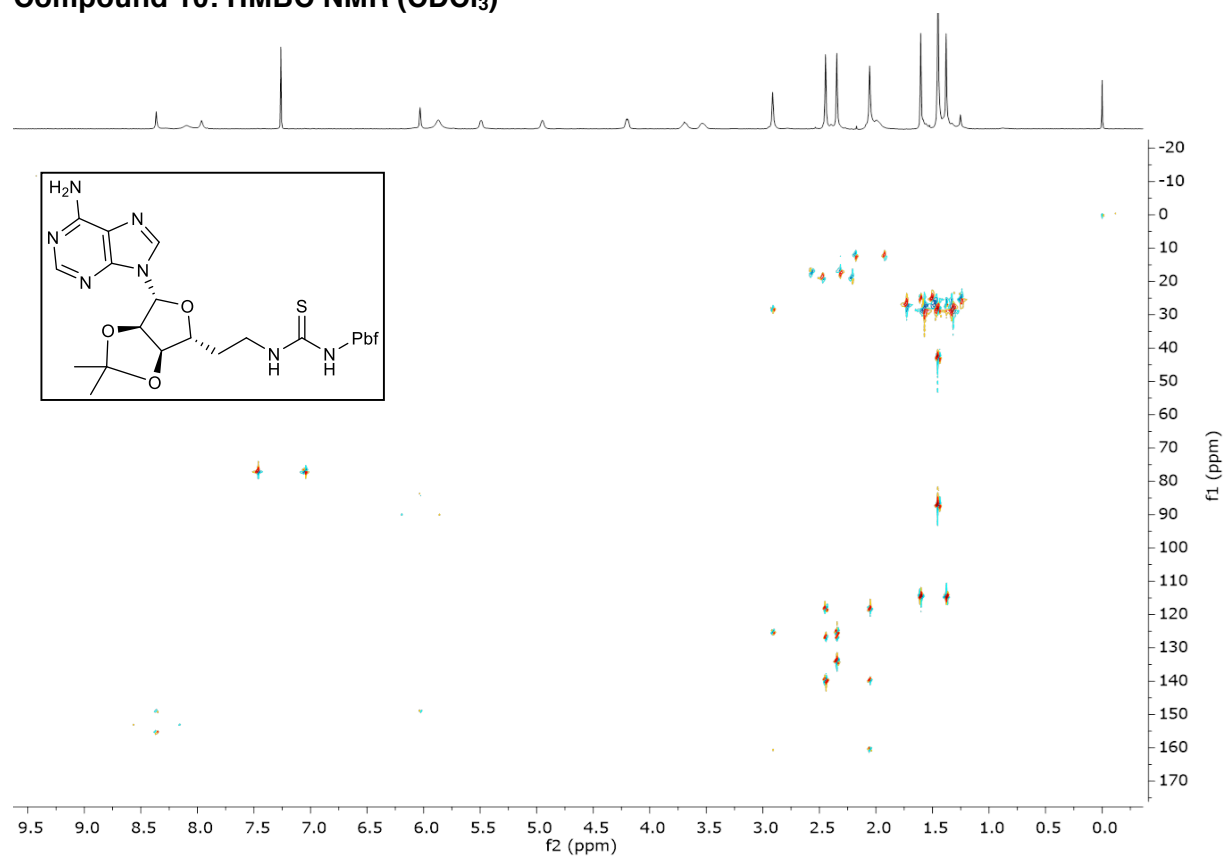
Compound 10: ^1H COSY NMR (500 MHz, CDCl_3)



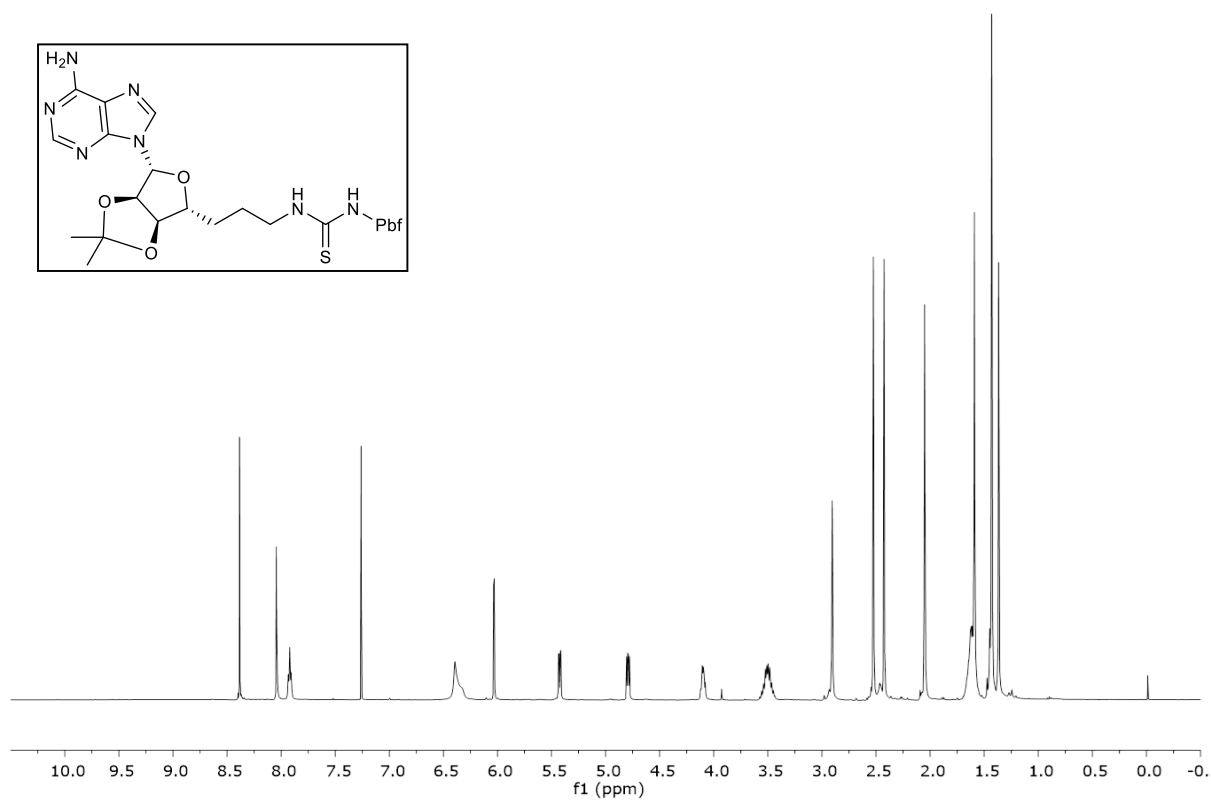
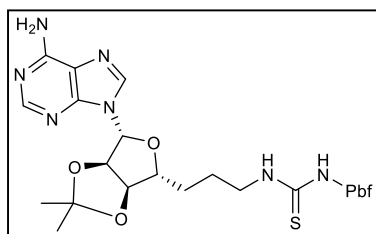
Compound 10: HSQC NMR (CDCl₃)



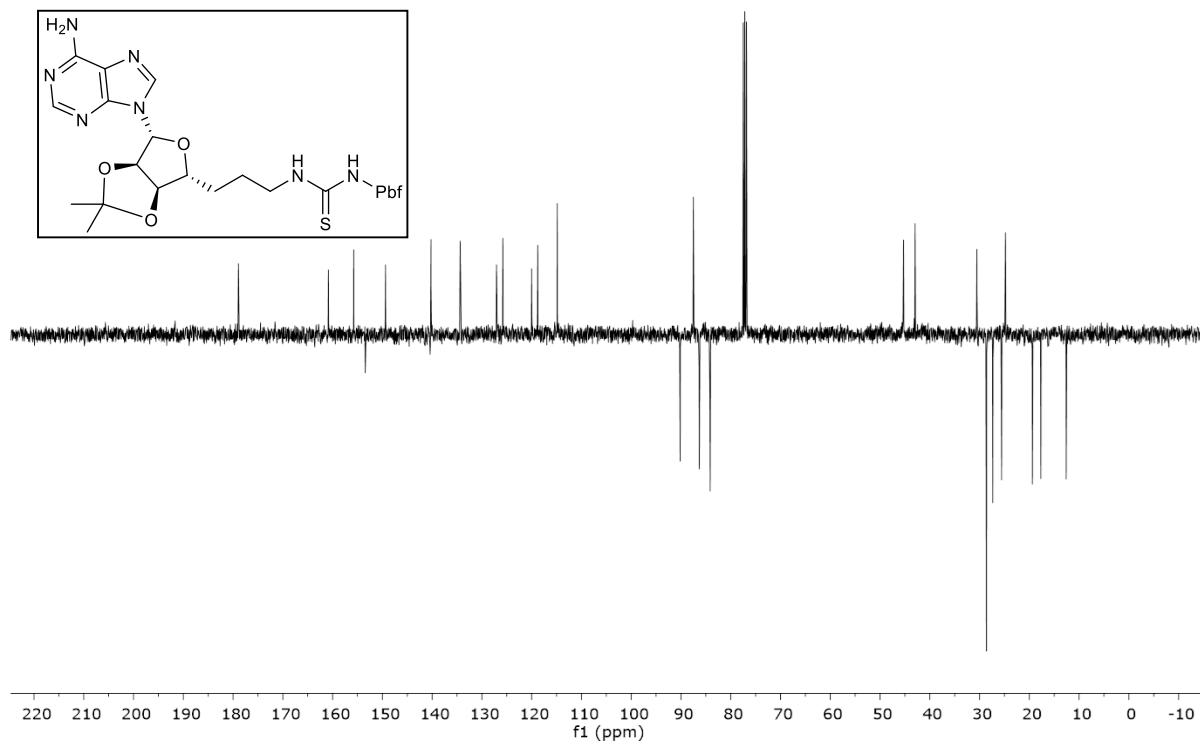
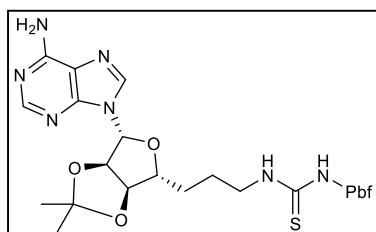
Compound 10: HMBC NMR (CDCl₃)



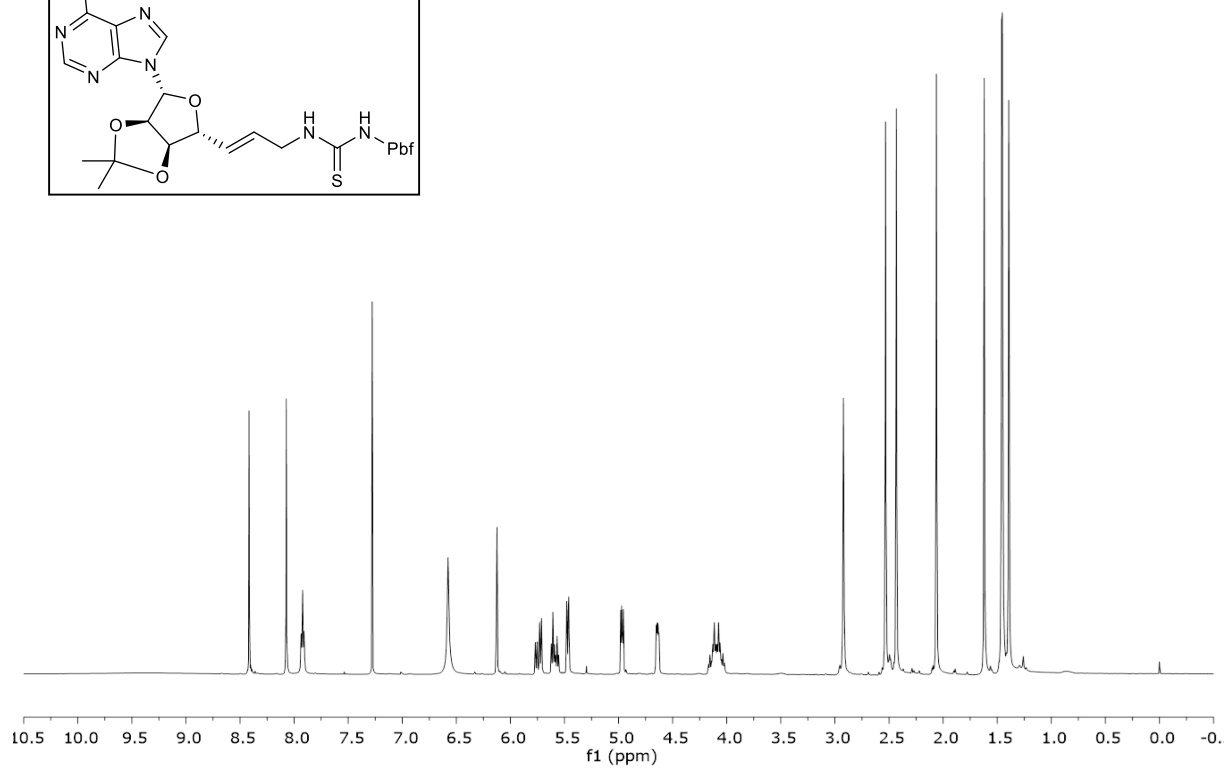
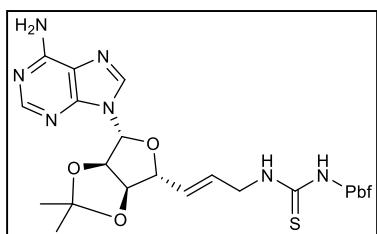
Compound 11: ^1H NMR (400 MHz, CDCl_3)



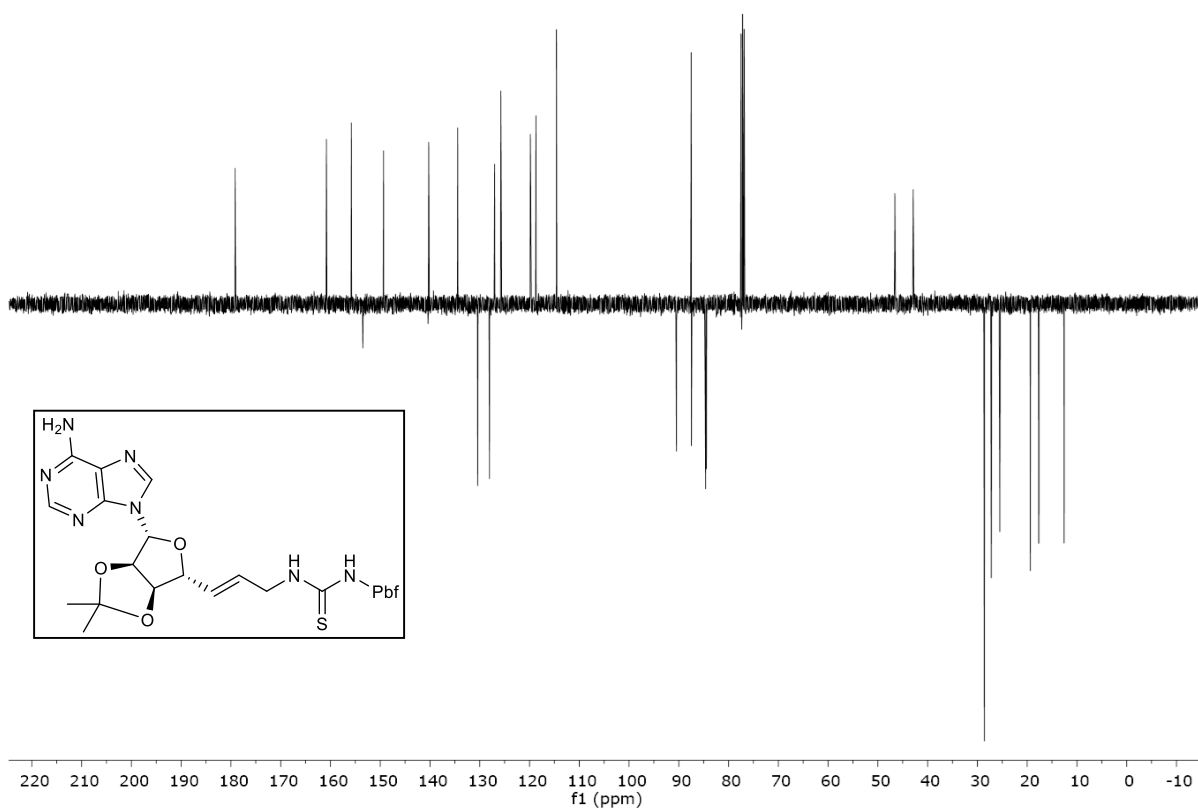
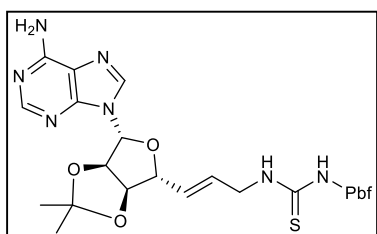
Compound 11: ^{13}C NMR (101 MHz, CDCl_3)



Compound 12: ^1H NMR (400 MHz, CDCl_3)

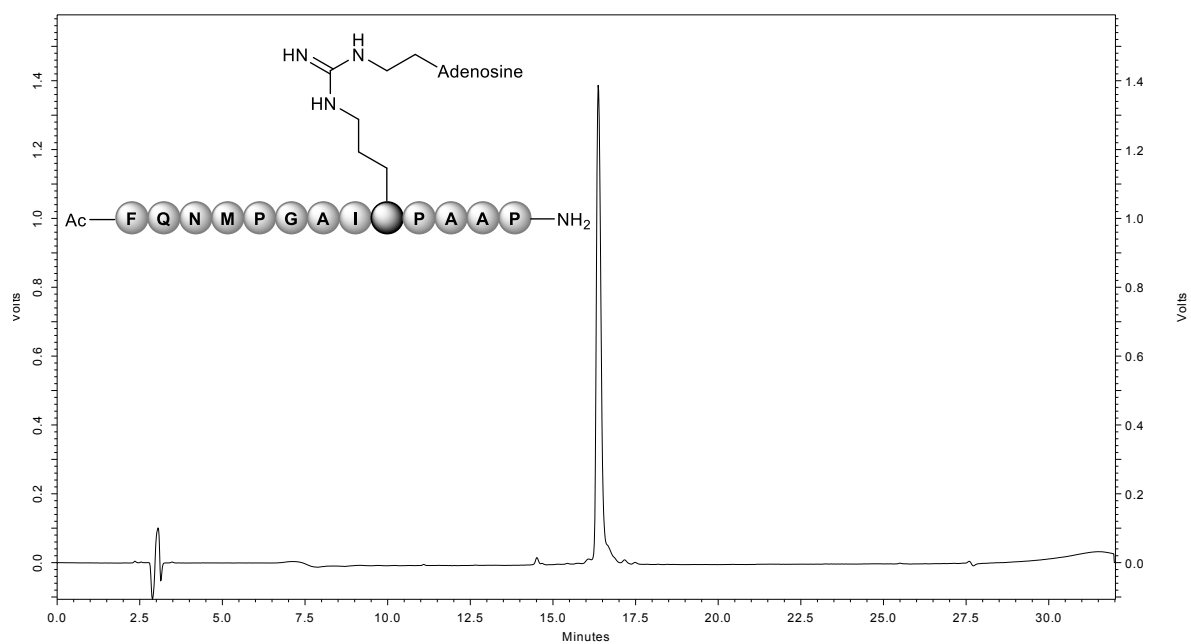


Compound 12: ^{13}C NMR (101 MHz, CDCl_3)

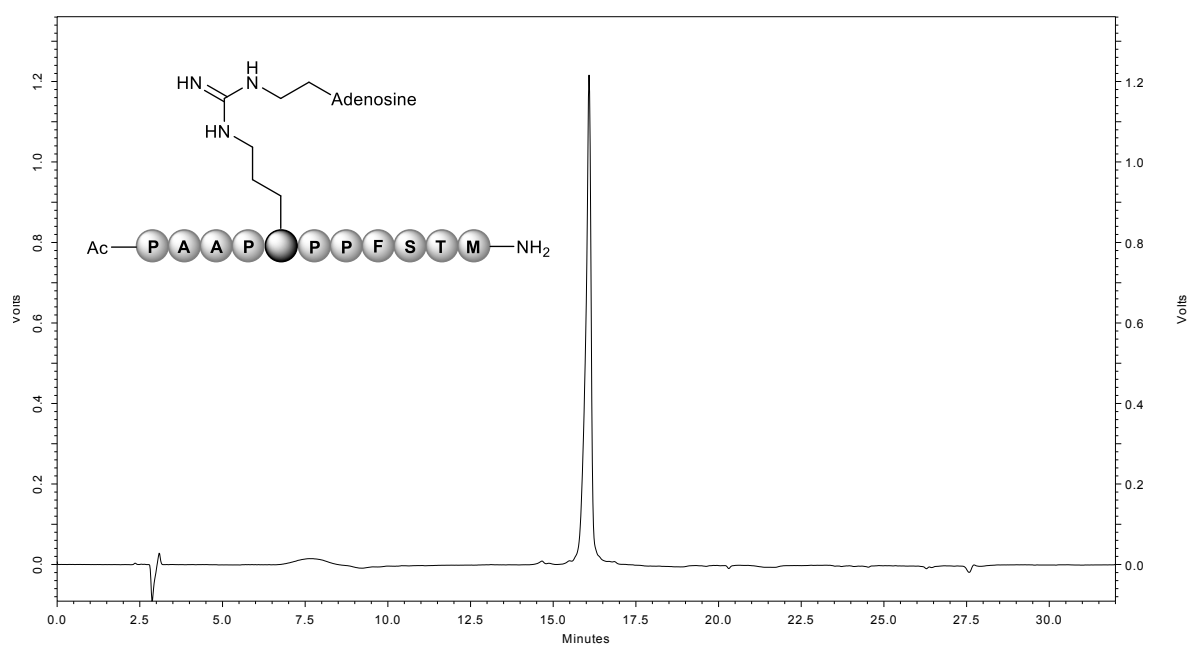


HPLC traces of final compounds

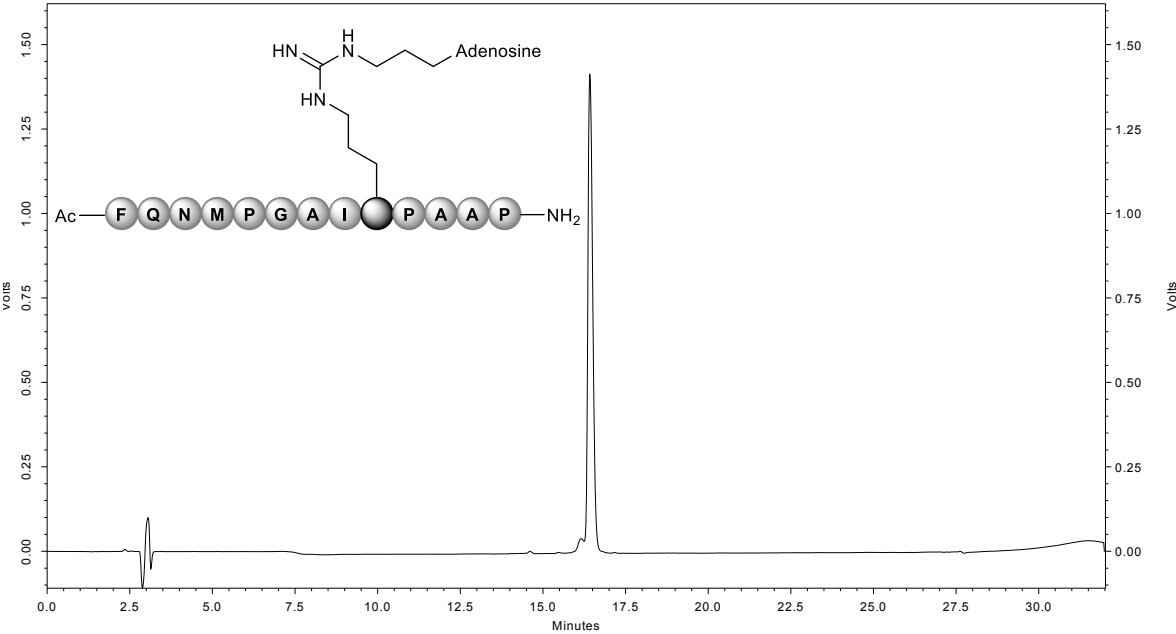
Compound 1



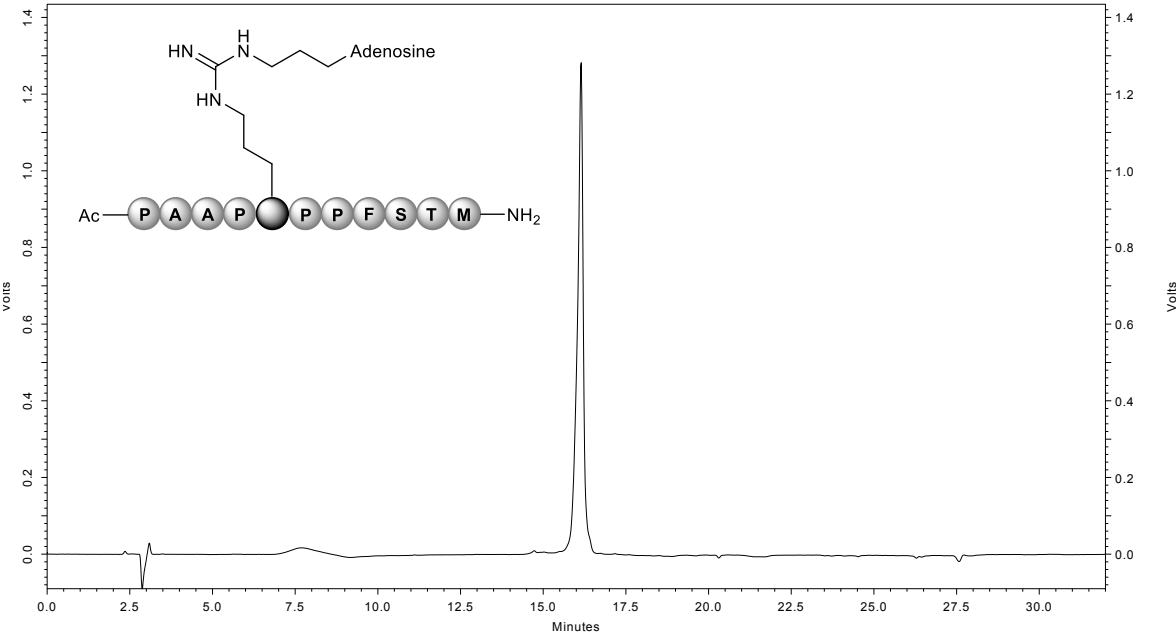
Compound 2



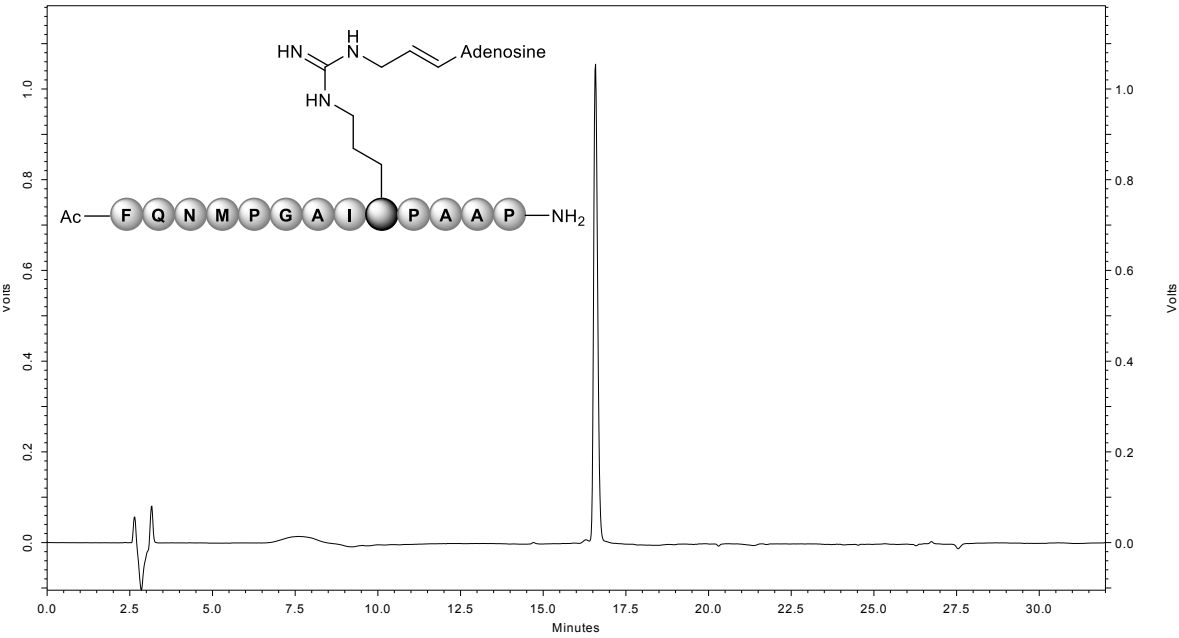
Compound 3



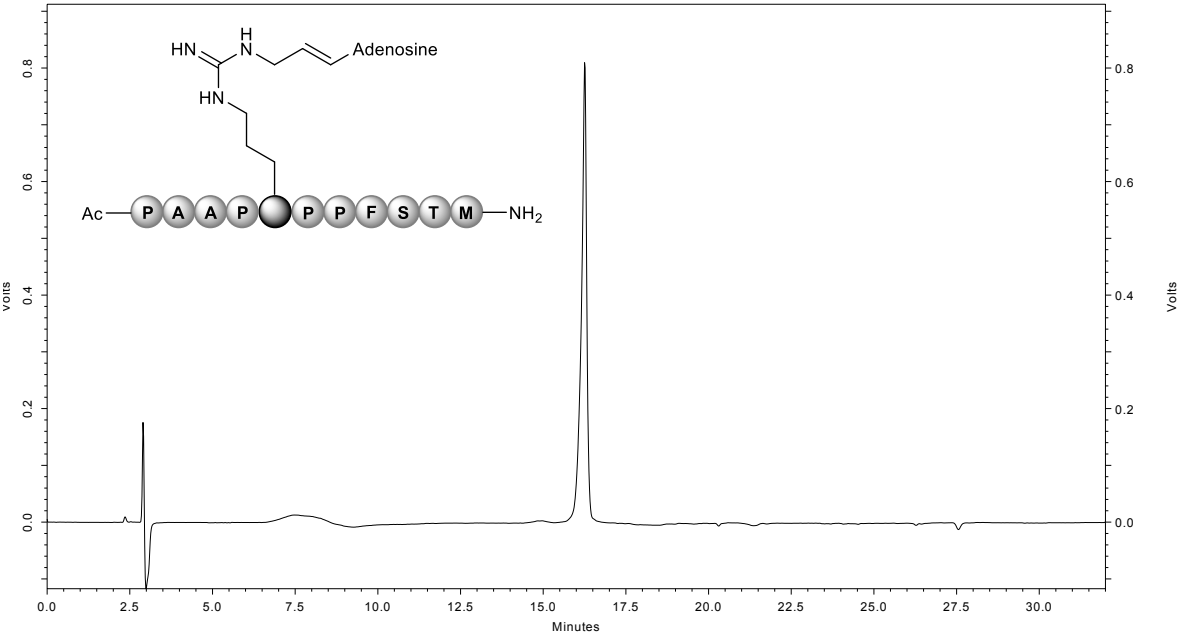
Compound 4



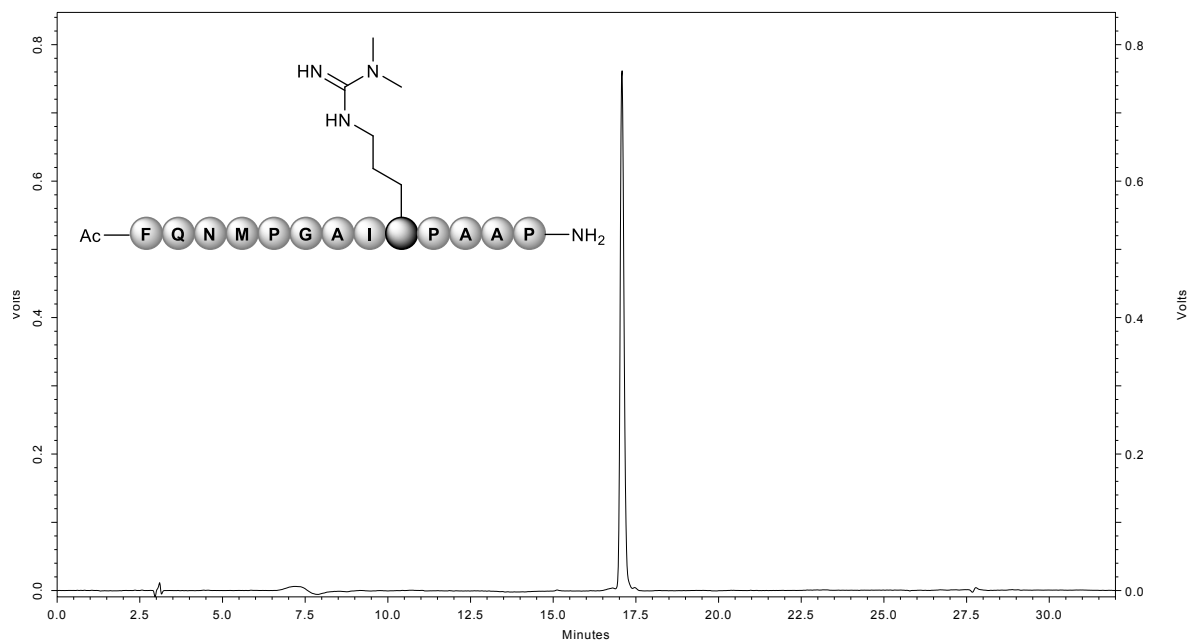
Compound 5



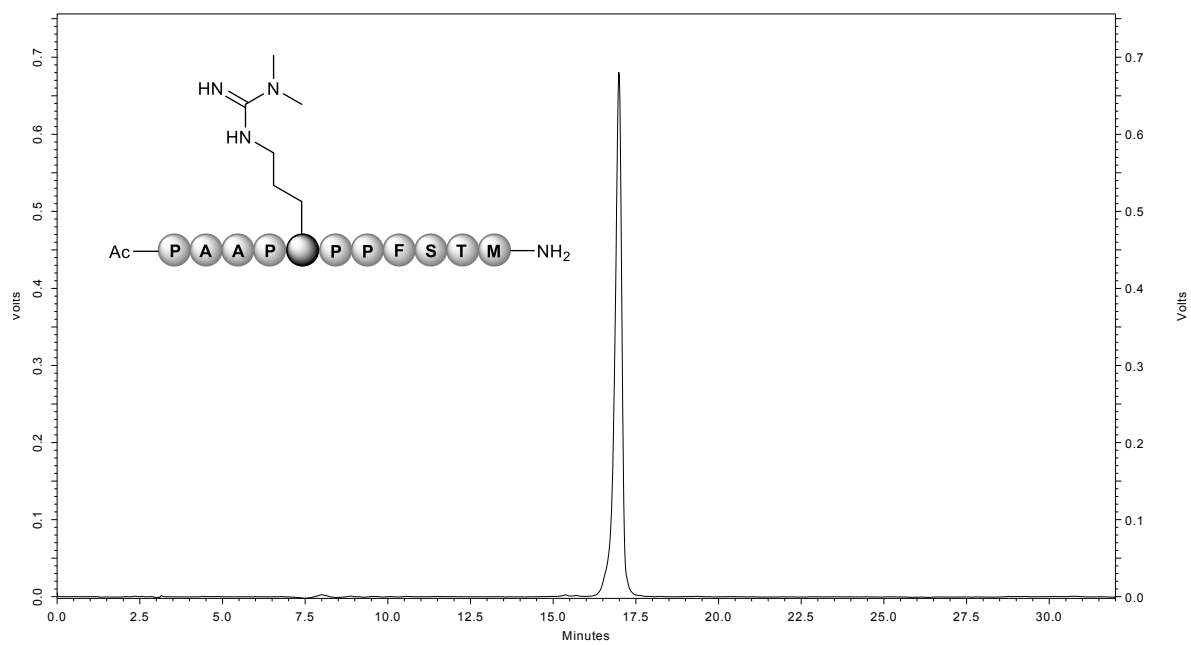
Compound 6



Compound 7



Compound 8



References

1. van Haren M, van Ufford LQ, Moret EE, & Martin NI (2015) Synthesis and evaluation of protein arginine N-methyltransferase inhibitors designed to simultaneously occupy both substrate binding sites. *Org Biomol Chem* 13(2):549-560.
2. Martin NI & Liskamp RL (2008) Preparation of N^G-substituted L-arginine analogues suitable for solid phase peptide synthesis. *J Org Chem* 73(19):7849-7851.
3. Vedadi M, *et al.* (2011) A chemical probe selectively inhibits G9a and GLP methyltransferase activity in cells. *Nat Chem Biol* 7(8):566-574.
4. Chen D, *et al.* (1999) Regulation of transcription by a protein methyltransferase. *Science (New York, N.Y.)* 284(5423):2174-2177.
5. Karplus PA & Diederichs K (2015) Assessing and maximizing data quality in macromolecular crystallography. *Curr Opin Struct Biol* 34:60-68.

***Hijacking DNA methyltransferase transition state analogues to
produce chemical scaffolds for PRMT inhibitors***

Ludovic Halby, Nils Marechal, Dany Pechalrieu, Vincent Cura, Don-Marc Franchini, Céline Faux,
Frédéric Alby, Nathalie Troffer-Charlier, Srikanth Kudithipudi, Albert Jeltsch, Wahiba Aouadi,
Etienne Decroly, Jean-Claude Guillemot, Patrick Page, Clotilde Ferroud, Luc Bonnefond,
Dominique Guianvarc'h, Jean Cavarelli, and Paola B. Arimondo

Research



Cite this article: Halby L *et al.* 2018 Hijacking DNA methyltransferase transition state analogues to produce chemical scaffolds for PRMT inhibitors. *Phil. Trans. R. Soc. B* **373**: 20170072.
<http://dx.doi.org/10.1098/rstb.2017.0072>

Accepted: 11 September 2017

One contribution of 18 to a discussion meeting issue 'Frontiers in epigenetic chemical biology'.

Subject Areas:
biochemistry

Keywords:
epigenetics, DNA methylation, histone methylation, transition state analogues, PRMT inhibitor, chemical probes

Authors for correspondence:

Jean Cavarelli
e-mail: cava@igbmc.fr
Paola B. Arimondo
e-mail: paola.arimondo@etac.cnrs.fr

[†]Present address: Centre de Recherches en Cancérologie de Toulouse, INSERM UMR1037, Toulouse, France.

[‡]Shared last co-authorship.

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.4025908>.

Hijacking DNA methyltransferase transition state analogues to produce chemical scaffolds for PRMT inhibitors

Ludovic Halby^{1,2}, Nils Marechal^{3,4,5,6}, Dany Pechalrieu¹, Vincent Cura^{3,4,5,6}, Don-Marc Franchini^{1,†}, Céline Faux¹, Frédéric Alby⁷, Nathalie Troffer-Charlier^{3,4,5,6}, Srikanth Kudithipudi⁸, Albert Jeltsch⁸, Wahiba Aouadi⁹, Etienne Decroly⁹, Jean-Claude Guillemot⁹, Patrick Page¹⁰, Clotilde Ferroud¹¹, Luc Bonnefond^{3,4,5,6}, Dominique Guianvarc'h^{12,13}, Jean Cavarelli^{3,4,5,6,‡} and Paola B. Arimondo^{1,14,‡}

¹CNRS FRE3600 EtAc, bât. IBCG, 31062 Toulouse, France

²Maison Française d'Oxford, CNRS, MEAE, 2-10 Norham Road, Oxford, UK

³Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch, France

⁴Centre National de la Recherche Scientifique, UMR7104, Illkirch, France

⁵Institut National de la Santé et de la Recherche Médicale, U 1258, Illkirch, France

⁶Université de Strasbourg, Illkirch, France

⁷Laboratoire Pierre Fabre, 3 avenue H. Curien, 31100 Toulouse, France

⁸Institute of Biochemistry, Faculty of Chemistry, University Stuttgart, Pfaffenwaldring 55, 70569 Stuttgart, Germany

⁹Laboratoire Architecture Fonction des Macromolécules Biologiques (AFMB-UMR 7257), Aix-Marseille Université & CNRS, 163 avenue de Luminy, 13288 Marseille cedex 09, France

¹⁰Epirem SAS, 1 Rue des Pénitents Blancs, 31000 Toulouse, France

¹¹Laboratoire de chimie moléculaire, CMGPCE, EA7341, Conservatoire National des Arts et Métiers, 2 rue Conté, 75003 Paris, France

¹²Sorbonne Universités, UPMC Université Paris 06, Ecole Normale Supérieure, CNRS, Laboratoire des Biomolécules (LBM), 4 place Jussieu, 75005 Paris, France

¹³Département de Chimie, Ecole Normale Supérieure, PSL Research University, UPMC Université Paris 06, CNRS, Laboratoire des Biomolécules (LBM), 75005 Paris, France

¹⁴Churchill College, CB3 0DS Cambridge, UK

PBA, 0000-0001-5175-4396

DNA, RNA and histone methylation is implicated in various human diseases such as cancer or viral infections, playing a major role in cell process regulation, especially in modulation of gene expression. Here we developed a convergent synthetic pathway starting from a protected bromomethylcytosine derivative to synthesize transition state analogues of the DNA methyltransferases. This approach led to seven 5-methylcytosine-adenosine compounds that were, surprisingly, inactive against hDNMT1, hDNMT3Acat, TRDMT1 and other RNA human and viral methyltransferases. Interestingly, compound 4 and its derivative 2 showed an inhibitory activity against PRMT4 in the micromolar range. Crystal structures showed that compound 4 binds to the PRMT4 active site, displacing strongly the S-adenosyl-L-methionine cofactor, occupying its binding site, and interacting with the arginine substrate site through the cytosine moiety that probes the space filled by a substrate peptide methylation intermediate. Furthermore, the binding of the compounds induces important structural switches. These findings open new routes for the conception of new potent PRMT4 inhibitors based on the 5-methylcytosine-adenosine scaffold.

This article is part of a discussion meeting issue 'Frontiers in epigenetic chemical biology'.

1. Background

RNA, DNA and histone methylation is implicated in various human diseases such as cancer or viral infections, playing a major role in cell processes regulation,

especially in modulation of gene expression [1,2]. Methylation of deoxycytidines (dC) in DNA is the most stable epigenetic mark and occurs at CpG sites, which are grouped in islands and essentially located in promoters, repeated sequences and CpG island shores [3]. Hypermethylation of promoters induces gene silencing while hypomethylation is associated with gene expression [4]. The enzymes responsible for DNA methylation are the DNA methyltransferases (DNMTs). DNMTs add a methyl group on the carbon-5 (C5) position of the deoxycytidine at the CpG site in the DNA by using *S*-adenosyl-L-methionine (AdoMet/SAM) as methyl donor, like most of the methyltransferases (MTases) [5]. Alteration of DNA methylation patterns is associated with various diseases, such as cancer, and DNA methylation represents an attractive therapeutic strategy [6].

Formerly DNMT2, now TRDMT1, structurally close to DNMTs, weakly methylates DNA at cytosine but rather methylates tRNA, especially cytosine 38 in the anticodon loop of aspartic acid tRNA [7]. tRNA methylation is implicated in metabolic activity of cancer cells and its inhibition has anti-proliferative effects [8]. More globally RNA cytosine methylation functions are still unknown and newly studied. While only the C5 position of cytosine is methylated in DNA, the four bases of RNA (pyrimidines U, C and purines G, A) are methylated at a large number of sites [9]. The study of the role of RNA methylation in diseases is emerging and the search for new inhibitors could result in chemical probes to study the biology of this modification [10].

Histone methyltransferases (HMTs) catalyse the transfer of one, two or three methyl groups to lysine and arginine residues of histone proteins. Two major types of HMT exist, lysine-specific (HKMTs) and arginine-specific (PRMTs). In eukaryotic cells, histone methylation at certain sites is implicated in the condensation and decondensation of chromatin. We focused on the proteins arginine methyltransferases (PRMTs), a family containing nine members (PRMT1 to PRMT9) identified in mammalian cells [11,12]. Expressed in nearly all mammalian cell types and tissues, PRMTs are involved in a wide variety of functional processes, like RNA processing, DNA damage and repair, cell signalling and most importantly transcription [13–15]. Moreover, arginine methylation of histones increases the complexity of the histone code by modulating the interaction of nuclear factors with other nearby histone marks.

Abnormal expression or activity of methylation-regulating enzymes has been noted in several types of human cancers, suggesting associations between histone methylation and malignant transformation of cells or formation of tumours [12,16]. Therefore, there is an emerging interest in the discovery of new inhibitors of DNMTs, RNA MTases and PRMTs. In the case of transferases, a potent chemical strategy to design inhibitors is to use transition state analogues. Interestingly, DNMTs, RNA MTases and PRMTs all share the same methyl donor, the AdoMet/SAM, and, concerning the substrate that is methylated, cytosine, guanine and arginine share some common pharmacological features. We thus developed a convergent synthetic pathway to synthesize transition state analogues of the DNMT methylation reaction based on the coupling of cytosine analogues to adenosine. Such compounds can be useful chemical tools for structural studies. We explored their potential and selectivity toward several DNMTs, but also RNA MTases and PRMTs. Inhibition data revealed that such compounds unexpectedly failed to inhibit DNMTs and RNA MTases but rather inhibited PRMTs. Crystal structures of PRMT4 and PRMT6 in the presence of the compounds revealed

that the compounds compete with the *S*-adenosyl-L-homocysteine (SAH) and occupy the two substrate pockets of the PRMTs, the SAM and arginine pocket. Thus the 5-methylcytosine-adenosine scaffold resulted in a potent starting point for the conception of new inhibitors of PRMT4.

2. Methods

(a) Chemicals

Commercial chemicals were from Sigma-Aldrich and Alfa-Aesar; oligonucleotides were bought from Eurogentec. The chemical synthesis and analysis of the conjugates are detailed in the electronic supplementary material. All compounds were prepared as 10^{-2} M aliquots in 100% DMSO, stored at -20°C and used as dilutions freshly prepared.

(b) Chemical synthesis

(i) *N*1-(Methoxymethyl)-*N*⁴-benzoyl-5-methylcytosine (**9**)

To a solution of *N*⁴-benzoyl-5-methylcytosine (**8**) synthesized according to [17] (500 mg, 2.18 mmol) in 70 ml of dichloromethane (DCM) was added *N,O*-bis(trimethylsilyl)acetamide (1.3 ml; 5.4 mmol), and the mixture was stirred at RT for 1 h. Methoxymethyl chloride (2.87 mmol) was added and the mixture was stirred for 18 h at RT. The solvent was removed and the residue was purified by silica gel chromatography using the eluent cyclohexane/ethyl acetate (4/6). Compound **9** was obtained as a white powder (554 mg; 2.03 mmol; 93%). ¹H NMR (500 MHz; CDCl₃) δ 13.21 (s, 1H), 8.31 (m, 2H), 7.52 (m, 1H), 7.44 (m, 2H), 7.25 (q, 1H, *J* = 1.2 Hz), 5.11 (s, 2H), 3.40 (s, 3H), 2.11 (d, 3H, *J* = 1.2 Hz). ¹³C NMR (125 MHz, CDCl₃) δ 179.8, 160.0, 148.9, 139.8, 137.1, 132.6, 130.0, 128.2, 112.6, 78.2, 57.1, 13.5. HRMS-ESI (*m/z*) calculated for C₇H₁₁N₃Na₁O₂ [M + Na]⁺: 296.1006; found: 296.0991.

(ii) *N*1-(Methoxymethyl)-*N*⁴-benzoyl-5-bromomethylcytosine (**10**)

To a solution of **9** (550 μmol) in dry benzene (3 ml) under argon atmosphere were added *N*-bromosuccinimide (NBS) (125 mg; 720 μmol) and 2,2'-azobis(2-methylpropionitrile) (AIBN) (17 mg; 0.1 mmol) and the mixture was heated at 70°C for 1 to 2 h. The reaction was followed by thin layer chromatography to prevent side product formation (cyclohexane/ethylacetate 4:6). After cooling, the solvent was removed by vacuum and the residue was co-evaporated with DCM twice. The crude product was used immediately without further purification.

(iii) 2',3'-*O*-Isopropylidene-5'-deoxy-5'-(2-nitrobenzenesulfonamido)adenosine (**13**) and 2',3'-*O*-isopropylidene-5'-deoxy-5'-(2-nitrobenzenesulfonamidomethyl)adenosine (**14**)

To a 0.2 M solution of **11** or **12** in pyridine was added 2-nitrophenylsulfonyl chloride (2.5 eq.) portion-wise and the mixture was stirred at RT overnight. The solvent was removed by vacuum and the residue was purified by silica gel flash chromatography using a linear gradient of 1 N NH₃ in MeOH (0 → 10% MeOH/NH₃) in DCM. 2',3'-*O*-Isopropylidene-5'-deoxy-5'-(2-nitrobenzenesulfonamido)adenosine (**13**) was obtained from 49 μmol of **11** as a white amorphous solid (200 mg; 41 μmol; 83%) and 2',3'-*O*-isopropylidene-5'-deoxy-5'-(2-nitrobenzenesulfonamidomethyl)adenosine (**14**) was obtained from 3.13 mmol of **12** as a white amorphous solid (1.21 g; 2.40 mmol; 77%). 2',3'-*O*-Isopropylidene-5'-deoxy-5'-(2-nitrobenzenesulfonamido)adenosine (**13**): ¹H NMR (500 MHz; DMSO-*d*₆) δ 8.64 (brs, 1H), 8.29 (s, 1H), 8.12 (s, 1H), 7.92 (dd, *J* = 1.2, 8.0 Hz, 1H), 7.84 (dd, *J* = 1.2, 8.0 Hz, 1H), 7.79 (ddd, *J* = 1.2, 6.4, 7.8 Hz, 1H), 7.71 (ddd, *J* = 1.2, 6.4, 7.7 Hz, 1H), 7.41 (brs, 2H), 6.10 (d, *J* = 2.9 Hz, 1H), 5.36 (dd, *J* =

3.0, 6.3 Hz, 1H), 4.94 (dd, $J = 2.9, 6.3$ Hz, 1H), 4.22 (ddd, $J = 2.9, 5.7, 11.5$ Hz, 1H), 3.29 (dd, $J = 5.8, 13.6$ Hz, 1H), 3.21 (dd, $J = 5.8, 13.6$ Hz, 1H), 1.51 (s, 3H), 1.27 (s, 3H). ^{13}C NMR (125 MHz; DMSO- d_6) δ 156.7, 153.0, 148.8, 148.0, 140.6, 132.4, 129.8, 124.8, 119.7, 113.9, 89.9, 85.7, 83.2, 82.0, 45.1, 27.4, 25.5. HRMS-ESI (m/z) calculated for $\text{C}_{33}\text{H}_{46}\text{N}_7\text{O}_8$ $[\text{M} + \text{H}]^+$: 492.1296; found: 492.1290. *2',3'-O*-Isopropylidene-5'-deoxy-5'-(2-nitrobenzenesulfonamidomethyl)adenosine (**14**): ^1H NMR (500 MHz; CDCl_3) δ 8.29 (s, 1H), 8.08–8.05 (m, 1H), 7.88 (s, 1H), 7.82–7.77 (m, 1H), 7.72–7.69 (m, 2H), 6.01 (brt, $J = 5.4$ Hz, 1H), 5.98 (d, $J = 2.4$ Hz, 1H), 5.70 (brs, 1H), 5.46 (dd, $J = 2.6, 6.4$ Hz, 1H), 5.00 (dd, $J = 4.5, 6.4$ Hz, 1H), 4.24 (dt, $J = 4.5, 7.9$ Hz, 1H), 3.31–3.18 (m, 1H), 2.10–1.98 (m, 1H), 1.61 (s, 3H), 1.38 (s, 3H). ^{13}C NMR (125 MHz; CDCl_3) δ 155.5, 153.2, 149.2, 147.8, 140.2, 133.8, 133.4, 132.6, 130.8, 125.3, 120.4, 114.9, 90.2, 84.8, 83.7, 83.5, 40.7, 45.1, 27.3, 25.4. HRMS-ESI (m/z) calculated for $\text{C}_{20}\text{H}_{24}\text{N}_7\text{O}_7\text{S}$ $[\text{M} + \text{H}]^+$: 506.1452; found: 506.1480.

(iv) *tert*-Butyl 2-(Boc-amino)-4-((2',3'-*O*-isopropylidene-5'-deoxyadenosin-5'-yl)amino)butanoate (**15**) and *tert*-butyl 2-(Boc-amino)-4-(((2',3'-*O*-isopropylidene-5'-deoxyadenosin-5'-yl)methyl)amino)butanoate (**16**)

To a solution of derivative **13** or **14** (305 μmol) in *N,N*-dimethylformamide (DMF) (5 ml) was added K_2CO_3 (126 mg; 915 μmol) and freshly prepared (*S*)-2-*tert*-butoxycarbonylamino-4-bromobutyric acid *tert*-butylester (125 mg; 369 μmol) and the mixture was stirred at RT for 6 h. Thiophenol (50 μl ; 500 μmol) was added and the mixture was stirred at RT overnight. The mixture was filtered, the solvent removed by vacuum and the residue purified by silica gel flash chromatography using a linear gradient 0 \rightarrow 10% MeOH/ NH_3 1 N in DCM to give **15** (143 mg; 253 μmol ; 83%) or **16** (155 mg; 266 μmol ; 88%) as pale yellow foams. *tert*-Butyl 2-(Boc-amino)-4-((2',3'-*O*-isopropylidene-5'-deoxyadenosin-5'-yl)amino)butanoate (**15**): ^1H NMR (500 MHz; CDCl_3) δ 8.83 (s, 1H), 8.20 (s, 1H), 6.01 (d, $J = 3.3$ Hz, 1H), 5.52 (m, 1H), 5.11 (m, 1H), 4.42 (m, 1H), 4.29 (m, 1H), 3.23 (s, 1H), 3.01 (m, 1H), 2.90 (m, 1H), 2.80 (m, 1H), 2.71 (m, 1H), 2.01 (m, 1H), 1.84 (m, 1H), 1.65 (m, 3H), 1.51–1.37 (m, 21H). ^{13}C NMR (125 MHz; CDCl_3) δ 171.6, 155.7, 152.7, 151.3, 149.7, 142.3, 123.7, 114.9, 91.0, 85.4, 83.3, 82.0, 81.8, 79.6, 52.6, 51.1, 46.2, 32.2, 28.3, 28.04, 27.7, 27.2, 25.5. HRMS-ESI (m/z) calculated for $\text{C}_{26}\text{H}_{42}\text{N}_7\text{O}_7$ $[\text{M} + \text{H}]^+$: 564.2140; found: 564.2153. *tert*-Butyl 2-(Boc-amino)-4-(((2',3'-*O*-isopropylidene-5'-deoxyadenosin-5'-yl)methyl)amino)butanoate (**16**): ^1H NMR (500 MHz; CDCl_3) δ 8.32 (s, 1H), 8.17 (s, 1H), 7.32 (brs, 2H), 7.32 (bd, $J = 7.6$ Hz, 2H), 6.09 (d, $J = 2.7$ Hz, 1H), 5.47 (dd, $J = 2.7, 6.4$ Hz, 1H), 4.88 (dd, $J = 3.7, 6.4$ Hz, 1H), 4.19–4.15 (m, 1H), 3.90–3.84 (m, 1H), 2.49–2.42 (m, 4H), 1.77–1.57 (m, 4H), 1.58 (m, 3H), 1.40–1.34 (m, 18H), 1.32 (m, 3H). ^{13}C NMR (125 MHz; CDCl_3) δ 172.4, 156.7, 155.9, 153.2, 149.4, 140.3, 119.5, 113.9, 88.7, 84.3, 84.1, 83.5, 80.5, 78.4, 53.2, 46.2, 45.8, 33.8, 31.2, 28.6, 28.1, 27.5, 25.7. HRMS-ESI (m/z) calculated for $\text{C}_{27}\text{H}_{44}\text{N}_7\text{O}_7$ $[\text{M} + \text{H}]^+$: 578.3297; found: 578.3301.

(v) *tert*-Butyl 2-(Boc-amino)-4-(((1-(methoxymethyl)cytosin-5-yl)methyl)(2',3'-*O*-isopropylidene-5'-deoxyadenosin-5'-yl)amino)butanoate (**17**) and *tert*-butyl 2-(Boc-amino)-4-(((1-(methoxymethyl)-cytosin-5-yl)methyl)((2',3'-*O*-isopropylidene-5'-deoxyadenosin-5'-yl)methyl)amino)butanoate (**18**)

Crude **10** was added to a solution of **15** or **16** (0.1 mmol) in DMF (3 ml) and *N,N*-diisopropylethylamine (DIPEA) (100 μl). After 1 h at RT, the mixture was diluted with ethyl acetate, the organic phase was washed with 1 M NaHCO_3 and brine, dried on MgSO_4 and the solvent was removed by vacuum. The residue was dissolved in 7 N NH_3 in MeOH and stirred at RT for 3 h.

The solvent was removed and the residue was purified by silica gel flash chromatography using a linear gradient of 1 N NH_3 in MeOH (0 \rightarrow 10% MeOH/ NH_3) in DCM, then by reversed phase high pressure liquid chromatography (HPLC) using a linear CH_3CN gradient with 0.2% TEA (0 \rightarrow 80% CH_3CN) to afford **17** (47 mg; 64 μmol ; 64%) or **18** (34 mg; 44 μmol ; 44%) as white powders. *tert*-Butyl 2-(Boc-amino)-4-(((1-(methoxymethyl)cytosin-5-yl)methyl)(2',3'-*O*-isopropylidene-5'-deoxyadenosin-5'-yl)amino)butanoate (**17**): ^1H NMR (500 MHz; DMSO- d_6) δ 8.29 (s, 1H), 8.12 (s, 1H), 7.51 (s, 1H), 7.36 (m, 3H), 7.10 (d, $J = 8.0$ Hz, 1H), 6.73 (brs, 1H), 6.18 (d, $J = 2.2$ Hz, 1H), 5.46 (dd, $J = 2.2, 6.3$ Hz, 1H), 5.0 (d, $J = 9.6$ Hz, 1H), 4.98 (dd, $J = 3.5, 6.3$ Hz, 1H), 4.89 (d, $J = 9.6$ Hz, 1H), 4.25 (m, 1H), 3.77 (m, 1H), 3.31 (m, 2H), 3.21 (s, 3H), 3.15 (m, 1H), 2.84 (m, 1H), 2.55 (m, 2H), 2.32 (m, 1H), 1.71 (m, 2H), 1.53 (s, 3H), 1.36 (s, 9H), 1.34 (s, 9H), 1.32 (s, 3H). ^{13}C NMR (125 MHz; DMSO- d_6) δ 172.2, 165.9, 156.6, 156.0, 155.9, 153.1, 149.0, 144.8, 140.6, 119.6, 114.0, 102.4, 89.0, 84.2, 83.7, 83.3, 80.7, 78.8, 78.5, 56.3, 55.6, 53.4, 53.0, 50.1, 28.6, 28.0, 27.7, 27.4, 25.8. HRMS-ESI (m/z) calculated for $\text{C}_{33}\text{H}_{51}\text{N}_{10}\text{O}_9$ $[\text{M} + \text{H}]^+$: 731.3835; found: 731.3829. *tert*-Butyl 2-(Boc-amino)-4-(((1-(methoxymethyl)-cytosin-5-yl)methyl)((2',3'-*O*-isopropylidene-5'-deoxyadenosin-5'-yl)methyl)amino)butanoate (**18**): ^1H NMR (500 MHz; DMSO- d_6) δ 8.30 (s, 1H), 8.16 (s, 1H), 7.52 (s, 1H), 7.32 (brs, 3H), 7.32 (bd, $J = 7.6$ Hz, 1H), 6.86–6.77 (m, 1H), 6.08 (d, $J = 2.6$ Hz, 1H), 5.46 (dd, $J = 2.6, 6.3$ Hz, 1H), 4.97 (s, 2H), 4.84 (dd, $J = 3.7, 6.3$ Hz, 1H), 4.11–4.01 (m, 1H), 3.83–3.75 (m, 1H), 3.27–3.14 (m, 4H), 2.48–2.32 (m, 4H), 1.87–1.58 (m, 4H), 1.52 (m, 3H), 1.37–1.32 (m, 18H), 1.31 (m, 3H). ^{13}C NMR (125 MHz; DMSO- d_6) δ 172.1, 166.1, 156.6, 156.0, 155.9, 153.2, 149.3, 144.3, 140.4, 119.6, 114.9, 102.7, 88.8, 84.5, 84.0, 83.5, 80.7, 78.8, 78.5, 56.2, 53.0, 52.9, 49.5, 48.9 (Cb1), 29.8, 28.6, 28.0, 27.9, 27.5, 25.7. HRMS-ESI (m/z) calculated for $\text{C}_{34}\text{H}_{53}\text{N}_{10}\text{O}_9$ $[\text{M} + \text{H}]^+$: 745.3991; Found: 745.3984.

(vi) 2-Amino-4-(((1-(methoxymethyl)cytosin-5-yl)methyl)(5'-deoxyadenosin-5'-yl)amino)butanoic acid (**1**) and 2-amino-4-(((1-(methoxymethyl)-cytosin-5-yl)methyl)((5'-deoxyadenosin-5'-yl)methyl)amino)butanoic acid (**2**)

Water (100 μl) was added to a solution of **17** (40 mg; 55 μmol) or **18** (30 mg; 40 μmol) in trifluoroacetic acid (TFA) (1 ml) and stirred at RT for 1 h. The solvent was removed and the residue was purified by reversed phase HPLC using a linear 0 \rightarrow 60% CH_3CN gradient to obtain **1** (19 mg; 36 μmol ; 65%) or **2** (19 mg; 35 μmol ; 87%) as a colourless foam. 2-Amino-4-(((1-(methoxymethyl)cytosin-5-yl)methyl)(5'-deoxyadenosin-5'-yl)amino)butanoic acid (**1**): ^1H NMR (500 MHz; D_2O) δ 8.10 (s, 1H), 8.08 (s, 1H), 7.51 (s, 1H), 5.90 (d, $J = 2.1$ Hz, 1H), 4.84 (s, 1H), 4.79 (s, 1H), 4.67 (m, 1H), 4.25–4.17 (m, 2H), 3.66–3.55 (m, 2H), 3.47 (dd, $J = 2.6, 13.6$ Hz, 1H), 3.15 (s, 3H), 2.93 (m, 2H), 2.78 (m, 2H), 2.00 (m, 1H), 1.92 (m, 1H). ^{13}C NMR (125 MHz; D_2O) δ 173.8, 164.1, 154.7, 156.0, 151.6, 148.6, 146.3, 140.7, 118.9, 115.1, 117.4, 88.6, 80.9, 72.8, 71.9, 79.6, 56.1, 54.7, 53.4, 52.4, 50.7, 26.7. HRMS-ESI (m/z) calculated for $\text{C}_{21}\text{H}_{31}\text{N}_{10}\text{O}_7$ $[\text{M} + \text{H}]^+$: 535.2372; Found: 534.2376. 2-Amino-4-(((1-(methoxymethyl)-cytosin-5-yl)methyl)((5'-deoxyadenosin-5'-yl)methyl)amino)butanoic acid (**2**): ^1H NMR (500 MHz; DMSO- d_6) δ 9.05 (brs, 1H, HNH), 8.42 (s, 1H), 8.28 (s, 1H), 8.02 (s, 1H), 7.98 (brs, 2H), 8.89 (d, $J = 5.3$ Hz, 1H), 5.06 (s, 2H), 4.67 (t, $J = 5.3$ Hz, 1H), 4.10 (t, $J = 4.8$ Hz, 1H), 3.98–3.88 (m, 2H), 3.83–3.65 (m, 2H), 3.28 (m, 4H), 2.98–2.78 (m, 4H), 2.11–1.89 (m, 4H). ^{13}C NMR (125 MHz; DMSO- d_6) δ 171.1, 161.5, 154.7, 150.8, 149.4, 141.3, 119.6, 88.8, 82.4, 79.3, 73.8, 73.4, 56.9, 50.7, 52.9, 49.5, 48.9, 29.8, 25.6. HRMS-ESI (m/z) calculated for $\text{C}_{22}\text{H}_{33}\text{N}_{10}\text{O}_7$ $[\text{M} + \text{H}]^+$: 546.2528; found: 546.2531.

(vii) 1-(Methoxymethyl)-5-((2',3'-O-isopropylidene-5'-deoxyadenosin-5'-yl)aminomethyl)cytosine (**19**) and 1-(methoxymethyl)-5-((2',3'-O-isopropylidene-5'-deoxyadenosin-5'-yl)aminomethyl)cytosine (**20**)

To a solution of derivative **13** or **14** (102 μmol) in DMF (2 ml) was added K_2CO_3 (41 mg; 306 μmol) and freshly prepared **10** (350 μmol). The mixture was stirred at RT for 1 h. The mixture was diluted with ethyl acetate and washed with water. The organic phase was washed with brine and dried with MgSO_4 . Solvent was removed and the residue was dissolved in 7N NH_3 in MeOH. The mixture was stirred at RT for 6 h. The solvent was removed under reduced pressure. The residue was dissolved in a suspension of K_2CO_3 (41 mg; 306 μmol) in DMF (1 ml). Thiophenol (25 μl ; 250 μmol) was added and the mixture was stirred at RT overnight, filtered and the solvent was removed. The residue was purified by reversed phase HPLC using a linear 0 \rightarrow 60% CH_3CN gradient to obtain **19** (10 mg; 21 μmol ; 21%) or **20** (17 mg; 36 μmol ; 35%) as white foams. 1-(Methoxymethyl)-5-((2',3'-O-isopropylidene-5'-deoxyadenosin-5'-yl)aminomethyl)cytosine (**19**): ^1H NMR (500 MHz; $\text{DMSO}-d_6$) δ 8.33 (s, 1H), 8.08 (s, 1H), 7.49 (s, 1H), 6.06 (d, $J = 3.3$ Hz, 1H), 5.44 (dd, $J = 3.3, 6.3$ Hz, 1H), 4.97 (m, 3H), 4.23 (m, 1H), 3.40 (d, $J = 6.1$ Hz, 2H), 3.20 (s, 3H), 2.69 (m, 2H), 1.54 (s, 3H), 1.33 (s, 3H). ^{13}C NMR (125 MHz; $\text{DMSO}-d_6$) δ 166.0, 156.4, 155.5, 153.1, 149.2, 143.9, 140.6, 119.4, 113.9, 104.1, 89.5, 83.0, 84.5, 82.6, 78.8, 56.3, 50.2, 46.9, 27.4, 25.7. HRMS-ESI (m/z) calculated for $\text{C}_{20}\text{H}_{28}\text{N}_9\text{O}_5$ [$\text{M} + \text{H}$] $^+$: 474.2208; found: 474.2202. 1-(Methoxymethyl)-5-((2',3'-O-isopropylidene-5'-deoxyadenosin-5'-yl)aminomethyl)cytosine (**20**): ^1H NMR (500 MHz; $\text{DMSO}-d_6$) δ 8.32 (s, 1H), 8.17 (s, 1H), 7.47 (s, 1H), 7.33 (brs, 2H), 7.13 (brs, 1H), 6.09 (d, $J = 2.7$ Hz, 1H), 5.37 (dd, $J = 2.7, 6.4$ Hz, 1H), 4.97 (s, 2H), 4.89 (dd, $J = 3.8, 6.4$ Hz, 1H), 4.18–4.15 (m, 1H), 3.34 (s, 2H), 3.21 (s, 3H), 2.48–2.47 (m, 2H), 1.80–1.71 (m, 1H), 1.53 (s, 3H), 1.33 (s, 3H). ^{13}C NMR (125 MHz; $\text{DMSO}-d_6$) δ 166.2, 156.6, 156.1, 153.2, 149.4, 143.4, 140.4, 119.5, 114.0, 104.1, 88.8, 84.5, 84.1, 83.5, 78.8, 56.2, 47.2, 45.1, 33.5, 27.5, 25.7. HRMS-ESI (m/z) calculated for $\text{C}_{21}\text{H}_{30}\text{N}_9\text{O}_5$ [$\text{M} + \text{H}$] $^+$: 488.2364; found: 488.2362.

(viii) 1-(Methoxymethyl)-5-(*N*-(5'-deoxyadenosin-5'-yl)aminomethyl)cytosine (**3**) and 1-(methoxymethyl)-5-(*N*-(5'-deoxyadenosin-5'-yl)aminomethyl)cytosine (**4**)

Water (50 μl) was added to a solution of **19** (10 mg; 21 μmol) or of **20** (15 mg; 31 μmol) in TFA (0.5 ml) and stirred at RT for 1 h. The solvent was removed and the residue was co-evaporated three times with 1N NH_3 in MeOH under reduced pressure to obtain **3** (8.1 mg; 19 μmol ; 89%) or **4** (9.2 mg; 19 μmol ; 89%) as a colourless foam. 1-(Methoxymethyl)-5-(*N*-(5'-deoxyadenosin-5'-yl)aminomethyl)cytosine (**3**): ^1H NMR (500 MHz; D_2O) δ 8.26 (s, 1H), 8.12 (s, 1H), 7.83 (s, 1H), 6.10 (d, $J = 4.6$ Hz, 1H), 5.07 and 5.01 (d, $J = 10.6$ Hz, 2H), 4.83 (t, $J = 4.5$ Hz, 1H), 4.53 (t, $J = 5.2$ Hz, 1H), 4.49 (q, $J = 5.8$ Hz, 1H), 4.17 and 4.12 (d, $J = 14.6$ Hz, 2H), 3.60 (d, $J = 6.0$ Hz, 2H), 3.30 (s, 3H). ^{13}C NMR (125 MHz; D_2O) δ 164.8, 157.1, 155.5, 152.3, 149.2, 148.2, 140.7, 119.0, 98.6, 89.3, 79.8, 79.3, 72.8, 71.3, 56.1, 48.0, 44.0. HRMS-ESI (m/z) calculated for $\text{C}_{17}\text{H}_{23}\text{N}_9\text{NaO}_5$ [$\text{M} + \text{Na}$] $^+$: 456.1714; found: 456.1699. 1-(Methoxymethyl)-5-(*N*-(5'-deoxyadenosin-5'-yl)aminomethyl)cytosine (**4**): ^1H NMR (500 MHz; $\text{DMSO}-d_6$) δ 8.57 (brs, 2H), 8.40 (s, 1H), 8.24 (s, 1H), 7.98 (s, 1H), 7.72 (brs, 1H), 5.90 (d, $J = 4.23$ Hz, 1H), 5.06 (s, 2H), 4.62 (t, $J = 4.7$ Hz, 1H), 4.20 (t, $J = 5.4$ Hz, 1H), 3.98–3.91 (m, 2H), 3.29 (s, 3H), 3.09–3.01 (m, 2H), 2.08–1.97 (m, 1H). ^{13}C NMR (125 MHz; $\text{DMSO}-d_6$) δ 162.6, 155.4, 151.7, 150.6, 149.4, 140.4, 119.6, 97.3, 88.8, 81.0, 79.3, 73.6, 73.5, 56.9, 44.2, 43.2, 29.9. HRMS-ESI (m/z) calculated for $\text{C}_{18}\text{H}_{20}\text{N}_9\text{O}_5$ [$\text{M} + \text{H}$] $^+$: 488.2051; found: 488.2052.

(ix) 1-(Methoxymethyl)uracil (**22**)

To a solution of uracil (**21**) (0.88 g; 7.93 mmol) in 250 ml of DCM was added *N,O*-bis(trimethylsilyl)acetamide (BSA) (4.8 ml; 19.4 mmol) and the mixture was stirred for 1 h at RT. Methoxymethyl chloride (MOMCl) (784 μl ; 10.32 mmol) was added and stirred at RT for 18 h. The solvent was removed and the residue was purified by silica gel chromatography using the eluent cyclohexane/ethyl acetate (3:7) to give **22** (1.17 g; 7.53 mmol; 95%) as a white powder. ^1H NMR (500 MHz; $\text{DMSO}-d_6$) δ 11.32 (s, 1H), 7.70 (d, $J = 8.0$ Hz, 1H), 5.61 (d, $J = 8.0$ Hz, 1H), 5.02 (s, 2H), 3.27 (s, 3H). ^{13}C (125 MHz, $\text{DMSO}-d_6$) δ 164.0, 151.5, 145.4, 101.9, 78.0, 56.4. HRMS-ESI (m/z) calculated for $\text{C}_6\text{H}_8\text{N}_2\text{NaO}_3$ [$\text{M} + \text{Na}$] $^+$: 179.0427; found: 179.0416.

(x) 2',3'-O-Isopropylidene-5'-deoxy-5'-(((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)methyl)adenosine (**23**) and 2',3'-O-isopropylidene-5'-deoxy-5'-(((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)ethyl)adenosine (**24**)

To a solution of 1,2,4-triazole (138 mg; 2.0 mmol) and POCl_3 (60 μl ; 0.65 mmol) in 1.5 ml of acetonitrile at 0°C was added 280 μl of triethylamine (TEA) dropwise and the mixture was stirred at 0°C for 40 min and then for 30 min at RT. Compound **22** (380 mg; 0.24 mmol) was added and vigorously stirred at RT overnight. The solvent was removed and 5 ml of **11** (65 mg; 0.20 mmol) or **12** (89 mg; 0.20 mmol) in MeCN was added. The reaction mixture was stirred for 6 h at RT. The mixture was diluted with ethyl acetate. The organic phase was washed with water and brine and dried over Na_2SO_4 . The solvent was removed and the residue was purified by silica gel flash chromatography using a linear gradient of 1N ammonia in methanol (0 \rightarrow 10% MeOH/ NH_3) in DCM to afford **23** (42 mg; 94.6 μmol ; 47%) or **24** (34 mg; 74.2 μmol ; 37%) as a white powder. 2',3'-O-isopropylidene-5'-deoxy-5'-(((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)methyl)adenosine (**23**): ^1H NMR (500 MHz; $\text{DMSO}-d_6$) δ 8.35 (s, 1H), 8.21 (s, 1H), 7.95 (t, $J = 6.1$ Hz, 1H), 7.56 (d, $J = 7.2$ Hz, 1H), 7.36 (brs, 2H), 6.18 (d, $J = 2.7$ Hz, 1H), 5.73 (d, $J = 7.2$ Hz, 1H), 5.47 (dd, $J = 2.8, 6.3$ Hz, 1H), 5.02–4.95 (m, 3H), 4.34–4.29 (m, 1H), 3.68–3.26 (m, 2H), 3.24 (s, 3H), 1.54 (s, 3H), 1.33 (s, 3H). ^{13}C NMR (125 MHz; $\text{DMSO}-d_6$) δ 164.5, 156.6, 156.0, 153.2, 149.2, 145.1, 140.5, 119.7, 114.0, 95.0, 89.5, 84.3, 83.5, 82.3, 78.9, 56.3, 42.4, 27.5, 25.7. HRMS-ESI (m/z) calculated for $\text{C}_{19}\text{H}_{25}\text{N}_8\text{O}_5$ [$\text{M} + \text{H}$] $^+$: 445.1942; found: 445.1946. 2',3'-O-isopropylidene-5'-deoxy-5'-(((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)ethyl)adenosine (**24**): ^1H NMR (500 MHz; $\text{DMSO}-d_6$) δ 8.34 (s, 1H), 8.17 (s, 1H), 7.77 (t, $J = 5.4$ Hz, 1H), 7.52 (d, $J = 7.3$ Hz, 1H), 7.33 (brs, 2H), 6.13 (d, $J = 2.6$ Hz, 1H), 5.67 (d, $J = 7.3$ Hz, 1H), 5.49 (dd, $J = 2.7, 6.3$ Hz, 1H), 4.98 (s, 2H), 4.94 (dd, $J = 3.5, 6.3$ Hz, 1H), 4.19–4.14 (m, 1H), 3.27–3.20 (m, 5H), 1.98–1.76 (m, 2H), 1.54 (s, 3H), 1.33 (s, 3H). ^{13}C NMR (125 MHz; $\text{DMSO}-d_6$) δ 164.3, 156.6, 156.1, 153.2, 149.1, 144.7, 140.4, 119.5, 114.0, 95.0, 88.8, 83.9, 83.8, 83.5, 78.9, 56.3, 37.0, 32.8, 27.5, 25.7. HRMS-ESI (m/z) calculated for $\text{C}_{20}\text{H}_{27}\text{N}_8\text{O}_5$ [$\text{M} + \text{H}$] $^+$: 459.2099; found: 459.2096.

(xi) 5'-Deoxy-5'-(((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)adenosine (**5**) and 5'-deoxy-5'-(((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)methyl)adenosine (**6**)

Water (50 μl) was added to a solution of adenosine-cytosine derivative **23** (27 mg; 60.8 μmol) or **24** (28 mg; 61.1 μmol) in TFA (0.5 ml) and stirred at RT for 2 h. The solvent was removed and the residues were co-evaporated three times with 1N NH_3 in MeOH under reduce pressure. The residues were purified by

reversed phase HPLC using a linear acetonitrile gradient with 0.2% TEA (0 → 80% CH₃CN). Compound **5** (21 mg; 52.0 μmol; yield 85%) or **6** (19 mg; 45.3 μmol; yield 74%) was obtained as a colourless foam. 5'-Deoxy-5'-((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)adenosine (**5**): ¹H NMR (500 MHz; DMSO-d₆) δ 8.36 (s, 1H), 8.20 (s, 1H), 8.09 (t, *J* = 5.8 Hz, 1H), 7.58 (d, *J* = 7.2 Hz, 1H), 7.34 (brs, 2H), 5.88 (d, *J* = 6.2 Hz, 1H), 5.78 (d, *J* = 7.3 Hz, 1H), 5.47 (d, *J* = 6.3 Hz, 1H), 5.33 (d, *J* = 4.8 Hz, 1H), 5.00 (d, *J* = 4.2 Hz, 2H), 4.69 (ddd, *J* = 5.4, 6.1, 11.5 Hz, 1H), 4.15–4.07 (m, 2H), 3.65 (t, *J* = 5.7 Hz, 2H), 3.24 (s, 3H). ¹³C NMR (125 MHz; DMSO-d₆) δ 164.5, 156.6, 156.1, 153.0, 149.1, 145.1, 140.7, 119.9, 95.0, 88.2, 83.4, 78.9, 73.1, 71.8, 56.3, 42.6. HRMS-ESI (*m/z*) calculated for C₁₆H₂₁N₈O₅ [M + H]⁺: 405.1629; found: 405.1641. 5'-Deoxy-5'-(((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)methyl)adenosine (**6**): ¹H NMR (500 MHz; DMSO-d₆) δ 8.35 (s, 1H), 8.16 (s, 1H), 7.82 (t, *J* = 5.4 Hz, 1H), 7.53 (d, *J* = 7.4 Hz, 1H), 7.28 (brs, 2H), 5.88 (d, *J* = 5.4 Hz, 1H), 5.69 (d, *J* = 7.1 Hz, 1H), 5.51 (brs, 1H), 5.29 (brs, 1H), 4.99 (s, 2H), 4.69–4.63 (m, 1H), 4.14–4.06 (m, 1H), 3.97–3.91 (m, 1H), 3.43–3.35 (m, 1H), 3.30–3.20 (m, 4H), 2.04–1.81 (m, 2H). ¹³C NMR (125 MHz; DMSO-d₆) δ 164.3, 156.5, 156.2, 153.1, 149.8, 144.6, 140.4, 119.6, 95.1, 88.0, 81.9, 78.9, 73.8, 73.4, 56.3, 37.1, 33.0. HRMS-ESI (*m/z*) calculated for C₁₇H₂₃N₈O₅ [M + H]⁺: 419.1786; found: 419.1793.

(xii) 2',3'-*O*-Isopropylidene-5'-deoxy-5'-(*N*-((2-(Boc-amino)ethyl)-2-nitrophenylsulfonamido)methyl)adenosine (**25**)

To a solution of adenosine derivative **13** (158 mg; 0.52 mmol) in 3 ml of DMF was added K₂CO₃ (215 mg; 1.56 mmol) and 2-(Boc-amino)ethyl bromide (140 mg; 0.63 mmol) and the mixture was stirred for 12 h at RT. The solvent was removed and the residue was purified by silica gel chromatography using a linear gradient 0 → 10% MeOH/NH₃ in DCM. Compound **25** (207 mg; 0.33 mmol; 63%) was obtained as a pale yellow foam. ¹H NMR (500 MHz; CDCl₃) δ 8.34 (s, 1H), 7.87–7.83 (m, 2H), 7.59–7.51 (m, 2H), 7.50–7.43 (m, 1H), 6.01 (d, *J* = 2.0 Hz, 1H), 5.82 (brs, 2H), 5.39 (dd, *J* = 1.4, 6.4 Hz, 1H), 5.10 (dd, *J* = 4.0, 6.4 Hz, 1H), 4.91 (m, 1H), 4.33 (dt, *J* = 4.0, 8.3 Hz, 1H), 3.86–3.58 (m, 2H), 3.57–3.47 (m, 1H), 3.34–3.11 (m, 3H), 1.60 (s, 3H), 1.42 (s, 12H), 1.38 (s, 3H). ¹³C NMR (125 MHz; CDCl₃) δ 155.8, 155.5, 153.2, 149.0, 147.8, 140.1, 133.4, 132.9, 131.5, 131.1, 124.1, 120.3, 114.8, 90.4, 85.1, 84.0, 82.4, 79.4, 49.8, 48.8, 38.4, 28.3, 27.1, 25.3. HRMS-ESI (*m/z*) calculated for C₂₆H₃₅N₈O₉S [M + H]⁺: 635.2242; found: 635.2249.

(xiii) 2',3'-*O*-Isopropylidene-5'-deoxy-5'-(*N*-((2-aminoethyl)-2-nitrophenylsulfonamido)methyl)adenosine (**26**)

A solution of **25** (200 mg; 0.32 mmol) in dry TFA (0.5 ml) was stirred at RT for 1 h. The solvent was removed and the residue was purified by silica gel chromatography using a linear gradient 0 → 10% MeOH/NH₃ in DCM. Compound **26** (154 mg; 0.29 mmol; 90%) was obtained as a pale yellow foam. ¹H NMR (500 MHz; CDCl₃) δ 8.30 (s, 1H), 7.98–7.91 (m, 2H), 7.69–7.57 (m, 3H), 6.08 (d, *J* = 1.9 Hz, 1H), 6.00 (brs, 1H), 5.39 (dd, *J* = 1.9, 6.5 Hz, 1H), 5.08 (dd, *J* = 4.1, 6.2 Hz, 1H), 4.48 (dt, *J* = 3.9, 9.3 Hz, 1H), 3.96–3.88 (m, 1H), 3.66–3.54 (m, 2H), 3.33–3.22 (m, 1H), 2.94–2.81 (m, 2H), 1.60 (s, 3H), 1.39 (s, 3H). ¹³C NMR (125 MHz; CDCl₃) δ 155.6, 153.1, 148.8, 148.0, 140.4, 133.9, 131.9, 131.2, 124.3, 120.2, 115.0, 90.4, 85.9, 83.9, 82.3, 53.4, 48.8, 38.8, 27.0, 25.4. HRMS-ESI (*m/z*) calculated for C₂₁H₂₇N₈O₇S [M + H]⁺: 535.1718; found: 535.1719.

(xiv) 2',3'-*O*-Isopropylidene-5'-deoxy-5'-(((*N*-2-((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)ethyl)-2-nitrophenyl)sulfonamido)methyl)adenosine (**27**)

To a solution of triazole (276 mg; 4.0 mmol) and POCl₃ (120 μl; 1.29 mmol) in 3 ml of acetonitrile at 0°C was added TEA

(560 μl drop-wise) and the mixture was stirred for 40 min at 0°C then at RT for 30 min. Compound (**22**) (75 mg; 0.48 mmol) was added and the mixture was vigorously stirred at RT overnight. The solvent was removed and 5 ml of a solution of previously prepared **26** (134 mg; 250 μmol) was added and the mixture was stirred at RT for 6 h. The mixture was diluted with ethyl acetate. The organic phase was washed with water and brine and dried over Na₂SO₄. The solvent was removed and the residue was purified by silica gel flash chromatography using a linear gradient 0 → 10% MeOH/NH₃ in DCM to afford **27** (81 mg; 0.12 mmol; 48%) as a white powder. ¹H NMR (500 MHz; DMSO-d₆) δ 8.34 (s, 1H), 8.20 (s, 1H), 7.99–7.81 (m, 3H), 7.76 (ddd, *J* = 1.2, 7.7, 15.4 Hz, 1H), 7.64 (ddd, *J* = 1.2, 7.7, 15.4 Hz, 1H), 7.61–7.50 (m, 2H), 6.19 (d, *J* = 2.2 Hz, 1H), 5.39 (dd, *J* = 2.2, 6.4 Hz, 1H), 5.06 (dd, *J* = 3.7, 6.4 Hz, 1H), 5.01 (s, 2H), 4.33 (m, 1H), 3.82–60 (m, 2H), 3.47–3.41 (m, 2H), 3.33–3.26 (m, 1H), 3.24 (s, 3H), 3.14–3.04 (m, 2H), 1.52 (s, 3H), 1.31 (s, 3H). ¹³C NMR (125 MHz; DMSO-d₆) δ 163.9, 156.0, 155.4, 152.5, 148.9, 147.8, 145.1, 140.8, 132.7, 131.9, 130.4, 124.7, 129.52, 114.1, 95.1, 89.2, 84.5, 83.7, 82.3, 79.9, 56.3, 50.0, 46.9, 38.7, 27.4, 25.6. HRMS-ESI (*m/z*) calculated for C₂₇H₃₂N₁₀O₉S [M + H]⁺: 673.2147; found: 673.2152.

(xv) 2',3'-*O*-Isopropylidene-5'-deoxy-5'-((*N*-2-((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)ethyl)amino)methyl)adenosine (**28**)

To a solution of **27** (70 mg; 104 μmol) and K₂CO₃ (43 mg; 31.3 μmol) in acetonitrile (0.5 ml) was added thiophenol (10 μl; 100 μmol) and the mixture was stirred at RT overnight. The mixture was filtered and the solvent was removed by vacuum. The residue was purified by silica gel chromatography using a linear gradient 0 → 10% MeOH/NH₃ in DCM. Compound **28** (33 mg; 67.8 μmol; 65%) was obtained as a white powder. ¹H NMR (500 MHz; DMSO-d₆) δ 8.38 (s, 1H), 8.18 (s, 1H), 8.01 (brs, 1H), 7.60 (d, *J* = 7.2 Hz, 1H), 6.25 (d, *J* = 2.5 Hz, 1H), 5.73 (d, *J* = 7.3 Hz, 1H), 5.41 (dd, *J* = 2.5, 6.2 Hz, 1H), 5.09 (dd, *J* = 3.4, 6.2 Hz, 1H), 5.01 (s, 2H), 4.51–4.42 (m, 1H), 4.15–4.04 (m, 2H), 3.50–3.41 (m, 2H), 3.33–3.22 (m, 2H), 3.24 (s, 3H), 3.03–2.93 (m, 2H), 1.57 (s, 3H), 1.34 (s, 3H). ¹³C NMR (125 MHz; DMSO-d₆) δ 164.7, 156.6, 155.8, 153.2, 149.1, 145.2, 140.6, 119.7, 114.1, 95.0, 89.6, 83.6, 82.7, 82.2, 78.9, 56.3, 49.5, 47.7, 37.7, 27.4, 25.6. HRMS-ESI (*m/z*) calculated for C₂₁H₃₀N₉O₅ [M + H]⁺: 488.2364; found: 488.2364.

(xvi) 5'-Deoxy-5'-((*N*-2-((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)ethyl)amino)methyl)adenosine (**7**)

Water (50 μl) was added to a solution of adenosine-cytosine derivative **28** (30 mg; 62 μmol) in TFA (0.5 ml) and the mixture was stirred at RT for 2 h. The solvent was removed and the residues were co-evaporated three times with 1 N ammonia in methanol under reduced pressure. The residues were purified by reversed phase HPLC using a linear acetonitrile gradient with 0.2% TEA (0 → 80% CH₃CN). Compound **7** was obtained (23 mg; 51.3 μmol; yield 83%) as a colourless foam. ¹H NMR (500 MHz; DMSO-d₆) δ 8.97 (brs, 2H), 8.49 (s, 1H), 8.32 (s, 1H), 8.04 (brs, 1H), 7.65 (d, *J* = 7.4 Hz, 1H), 5.99 (d, *J* = 5.5 Hz, 1H), 5.77 (d, *J* = 7.4 Hz, 1H), 5.01 (s, 2H), 4.71 (t, *J* = 5.1 Hz, 1H), 4.32–4.21 (m, 2H), 3.67–3.36 (m, 4H), 3.24 (s, 3H), 3.20–2.09 (m, 2H). ¹³C NMR (125 MHz; DMSO-d₆) δ 164.3, 155.2, 154.8, 150.7, 149.2, 145.6, 141.7, 119.8, 95.1, 89.1, 80.2, 79.0, 73.2, 72.0, 56.4, 49.3, 47.7, 37.3. HRMS-ESI (*m/z*) calculated for C₁₈H₂₆N₉O₅ [M + H]⁺: 448.2051; found: 448.2051.

(c) Methyltransferase activity

DNMT3A and *DNMT1* assays were previously described [18,19] and are detailed in the electronic supplementary material.

The percentage of inhibition was calculated according to equation (2.1), where X is the signal determined in the absence of the inhibitor and Y is the signal obtained in the presence of the inhibitor:

$$\% \text{ inh} = \left[\frac{X - Y}{X} \right] \times 100. \quad (2.1)$$

TRDMT1 assays were conducted as described in [20] using inhibitor concentrations of 1 mM and 100 μM .

The *RNA MTase filter binding assay* (FBA) using various viral and cellular MTases was previously described [21–23]. Briefly, FBA is based on the incubation of the RNA MTase at the corresponding concentration in the presence of the RNA substrate and the radiolabelled triitated AdoMet (^3H SAM) in Tris-HCl pH 8 40 mM, DTT 1 mM, SAM 10 μM . Reaction mixtures were incubated for 20 min at 30°C and stopped by adding a 10-fold dilution of the reaction mixture in 20 μM ice-cold AdoHcy (New England Biolabs). Samples were then transferred to diethylaminoethylcellulose (DEAE) filters (Perkin Elmer) by using a Filtermat Harvester apparatus (Packard Instruments). Unincorporated ^3H SAM was removed from the DEAE filters by several washes with 0.01 M ammonium formate, pH 8.0, H_2O and absolute ethanol. After drying, filters were incubated with BetaplateScint (Wallac) scintillation fluid before quantification of the ^3H methylation of the RNA substrates using a Wallac 1450 MicroBetaTriLux liquid scintillation counter. Details are found in the electronic supplementary material.

PRMT assays were carried out by Eurofins Scientific (France), based on the incorporation of radioactive methyl groups from [methyl- ^3H]AdoMet on the corresponding peptide substrates in the presence of the human recombinant purified proteins [24–27] (details in the electronic supplementary material).

Permeability studies were carried out by Eurofins Scientific (France) and are detailed in the electronic supplementary material.

PRMT crystallography studies: The *Mus musculus* PRMT4 gene sequence corresponding to the PRMT core (residues 130 to 487, MmPRMT4) [28] was obtained and *M. musculus* PRMT6 cloning, expression and purification were carried out as described previously [29,30], and are detailed in the electronic supplementary material.

(d) PRMT4 and PRMT6 crystallization and structure determination with the compounds

Compounds were solubilized in water before addition to the protein solution (2 mg ml⁻¹ for PRMT4 and 6 mg ml⁻¹ for PRMT6) at the final concentration of 2 mM for PRMT4 and 1 mM for PRMT6. The vapour diffusion method on hanging drop trays (VDX plate, Hampton Research) with a 1 ml reservoir was used for crystallization. Typically, 1 μl of protein–ligand solution was added to 1 μl of well solution consisting of 15–25% (v/v) PEG2000MME, 100 mM Tris-HCl pH 8.0 and 100 mM NaCl for PRMT4, and of 100 mM Tris-HCl or HEPES-NaOH pH 8.0, 200 mM NaCl and 10% (v/v) PEG 6000 for PRMT6. Crystals were grown at 20°C and flash-frozen in liquid nitrogen after a brief transfer to 5 μl reservoir solution containing 15% (v/v) PEG400 as a cryoprotectant and were stored in liquid nitrogen.

Datasets were collected at the SOLEIL PROXIMA1 and ESRF ID23-2 beamlines, using a PILATUS 6M detector, PILATUS 2M-F (Dectris) or Q315 detector (ADSC), and processed with XDS [31] and HKL-2000 [32]. Five structures of MmPRMT4 were solved and refined. The crystals belong to the $P2_12_12_1$ space group with four MmPRMT4 molecules in the asymmetric unit (AU) or the $P2_1$ space group with eight MmPRMT4 molecules in the AU. Two structures of MmPRMT6 were solved and refined. The crystals belong to the $I4_1$ space group with one MmPRMT6 molecule in the AU. The structures were solved by molecular replacement using previously solved MmPRMT4 and MmPRMT6 structures.

Iterative cycles of model building and refinement were carried out using Coot [33], PHENIX [34] and BUSTER [35]. A summary of the data statistics for the five solved and refined structures is provided in table 1. The atomic coordinates and experimental data have been deposited in the Protein Data Bank (PDB IDs: 5TBH, 5TBI, 5TBJ, 5LV2, 5LV3, 5LV4 and 5LV5). All structure figures were prepared with Cuemol (<http://www.cuemol.org>).

3. Results and discussion

(a) Rational design

Recently, the transition state of DNMT1 has been characterized by means of experimental kinetic isotopic effects showing that the transfer of the methyl group is the rate-limiting step [37]. Thus, in order to mimic the transition state of the methylation reaction catalysed by the DNMTs (figure 1a), we coupled the adenosine moiety of the AdoMet cofactor to a cytosine analogue methylated at position 5. Position N1 of the cytosine ring was substituted by a methoxymethyl group (MOM) to mimic the sugar. We kept the amino acid substituent of the AdoMet, the methionine, replaced the sulfur atom by nitrogen to increase chemical stability and varied the length of the *N*-alkyl chain, affording compounds 1 and 2 (figure 1b).

Next we omitted the amino acid substituent (3 and 4, figure 1c) and explored the attachment to the amine on position C4 of the cytosine and the length of the *N*-alkyl chain between the cytosine and the adenosine (5, 6 and 7, figure 1c).

(b) Chemistry

The synthesis of these analogues required the optimization of a specific pathway because of the uncommon chemical reactivity of the cytosine base compared with thymine. The key step was the development of a convergent synthetic pathway starting from the protected bromomethylcytosine derivative 10 (scheme 1). Owing to its high reactivity, position N1 of the cytosine derivative 8 was substituted by a methoxymethyl group (MOM) in order to avoid polymerization. This MOM group was chosen because it can also mimic a fragment of the sugar moiety. Amine derivatives 1 to 4 were obtained by *N*-alkylation by cytosine bromide derivative 10 (scheme 1) of adenosine derivatives. The key synthon 9 was freshly prepared under Vorbrüggen conditions from the protected 5-methylcytosine derivative 8 which was obtained following the procedure described by Buchini *et al.* [17]. An *N1*-alkylation with a MOM group was performed to obtain derivative 9. Then radical bromination in the benzylic position of 9 was achieved with *N*-bromosuccinimide (NBS) and 2,2'-azo-bis-isobutyronitrile (AIBN) leading to the desired unstable bromide 10, which was used immediately without any purification.

Transition state analogues 1 and 2 were synthesized starting from amines 11 and 12 described previously [38,39], which were protected by addition of 2-nitrophenylsulfonyl chloride, affording the nosylated derivatives 13 and 14 (scheme 2). They were alkylated by (*S*)-2-*tert*-butoxycarbonylamino-4-bromobutyric acid *tert*-butylester [40] in the presence of TEA, then treated with thiophenol to remove the nosyl protective group to give 15 and 16. The secondary amines 15 and 16 were alkylated by the protected bromomethylcytosine 10 in the presence of K_2CO_3 , followed by treatment with ammonia in methanol to remove the benzoyl group of the cytosine moiety, leading to 17 and 18, respectively. The compounds 1 and 2 were obtained from compounds 17 and 18 after cleavage

Table 1. X-ray data collection and refinement statistics. Statistics for the highest-resolution shell are shown in parentheses. The resolution limits for $l/\sigma(l) > 2$ are reported.

	MmPRMT4	MmPRMT4	MmPRMT4	MmPRMT4	MmPRMT6	MmPRMT6
compound	1	3	7	4	1	2
PDB ID	5TBH	5LV2	5TBI	5TBJ	5LV4	5LV5
<i>data processing</i>						
wavelength (Å)	0.97857	0.97857	0.97857	0.97857	0.97779	0.9801
resolution range (Å)	49.26–2.34 (2.40–2.34)	49.34–2.29 (2.33–2.29)	45.54–2.29 (2.35–2.29)	19.97–2.32 (2.38–2.32)	41.10–1.65 (1.69–1.65)	40.16–1.80 (1.84–1.80)
space group	P2 ₁ 2 ₁ 2	P2 ₁	P2 ₁ 2 ₁ 2	P2 ₁ 2 ₁ 2	I4 ₁	I4 ₁
unit cell (Å, °)	74.6 98.5 206.4 90 90 90	98.7 74.8 206.6 90 90.53 90	74.5 98.0 205.7 90 90 90	74.8 99.2 207.3 90 90 90	79.9 79.9 119.7 90 90 90	77.7 77.7 117.7 90 90 90
total reflections	407 736 (25 394)	456 198 (19 462)	299 557 (17 440)	291 218 (16 419)	318 563 (11 302)	168 294 (8906)
unique reflections	64 779 (4286)	134 100 (6011)	67 916 (4232)	66 284 (3886)	44 207 (2053)	32 113 (1875)
multiplicity	6.3 (5.9)	3.4 (3.2)	4.4 (4.1)	4.4 (4.2)	7.2 (5.5)	5.2 (4.7)
completeness (%)	99.6 (95.3)	99.1 (90.9)	99.3 (92.3)	98.6 (87.5)	99.7 (94.5)	99.8 (97.3)
$l/\sigma(l)$	7.4 (1.8)	5.5 (1.0)	6.9 (1.2)	4.6 (0.9)	7.7 (1.0)	10.7 (1.2)
resolution limit for $l/\sigma(l) > 2$	2.40	2.55	2.49	2.63	1.79	1.89
Wilson <i>B</i> -factor (Å ²)	31.59	32.27	31.86	30.97	21.69	23.11
R_{meas}^a	0.301 (2.016)	0.188 (1.388)	0.185 (1.314)	0.263 (2.062)	0.153 (1.656)	0.117 (1.417)
$CC_{1/2}^b$	0.979 (0.328)	0.985 (0.412)	0.991 (0.495)	0.972 (0.317)	0.996 (0.474)	0.998 (0.421)
<i>refinement</i>						
resolution range (Å)	49.26–2.34 (2.37–2.34)	49.34–2.29 (2.32–2.29)	45.54–2.29 (2.32–2.29)	19.97–2.32 (2.35–2.32)	41.10–1.65 (1.69–1.65)	40.16–1.80 (1.86–1.80)
% <i>R</i> -work	19.36 (31.06)	20.51 (34.26)	19.77 (33.57)	20.22 (35.39)	15.90 (29.30)	15.40 (27.59)
% <i>R</i> -free	22.98 (34.29)	23.80 (34.33)	22.84 (34.26)	23.73 (35.62)	18.80 (31.00)	19.52 (34.04)
number of non-H atoms						
protein	11 041	22 150	11 041	11 009	2504	2533
ligands	166	261	157	224	38	39
water	406	1166	600	754	331	221
<i>validation</i>						
RMS (bonds)	0.006	0.005	0.004	0.004	0.009	0.010
RMS (angles)	1.024	0.912	0.855	0.904	1.079	1.060
Ramachandran favoured (%)	96.99	96.96	96.52	97.36	99.05	98.73
Ramachandran outliers (%)	0.0	0.04	0.0	0.15	0.0	0.32
rotamer outliers (%)	0.42	0.50	0.17	0.42	0.37	0.0
average <i>B</i> -factor (Å ²)	34.89	38.11	36.50	36.24	30.85	30.63
protein	34.79	38.11	36.43	36.06	29.91	39.90
ligands	42.32	41.76	42.27	41.70	20.79	35.26
water	36.60	37.26	36.17	37.32	39.14	38.20

^a R_{meas} is the redundancy-independent merging.^bSee [36] for definitions.

of the protective groups upon TFA/water treatment. Since the deprotected group compounds were difficult to purify, we chose to first isolate compounds **17** and **18** at high purity and to cleave all the acid-labile protective groups only in the final step. Most importantly, in the synthetic pathway, the choice

of the nosyl group was crucial because of its ability to play a double role of activating and protecting group [41]. A direct *N*-alkylation of the primary amines **11** and **12** was tried first but provided the unprotected secondary amine with a very poor yield, even upon use of caesium hydroxide as base,

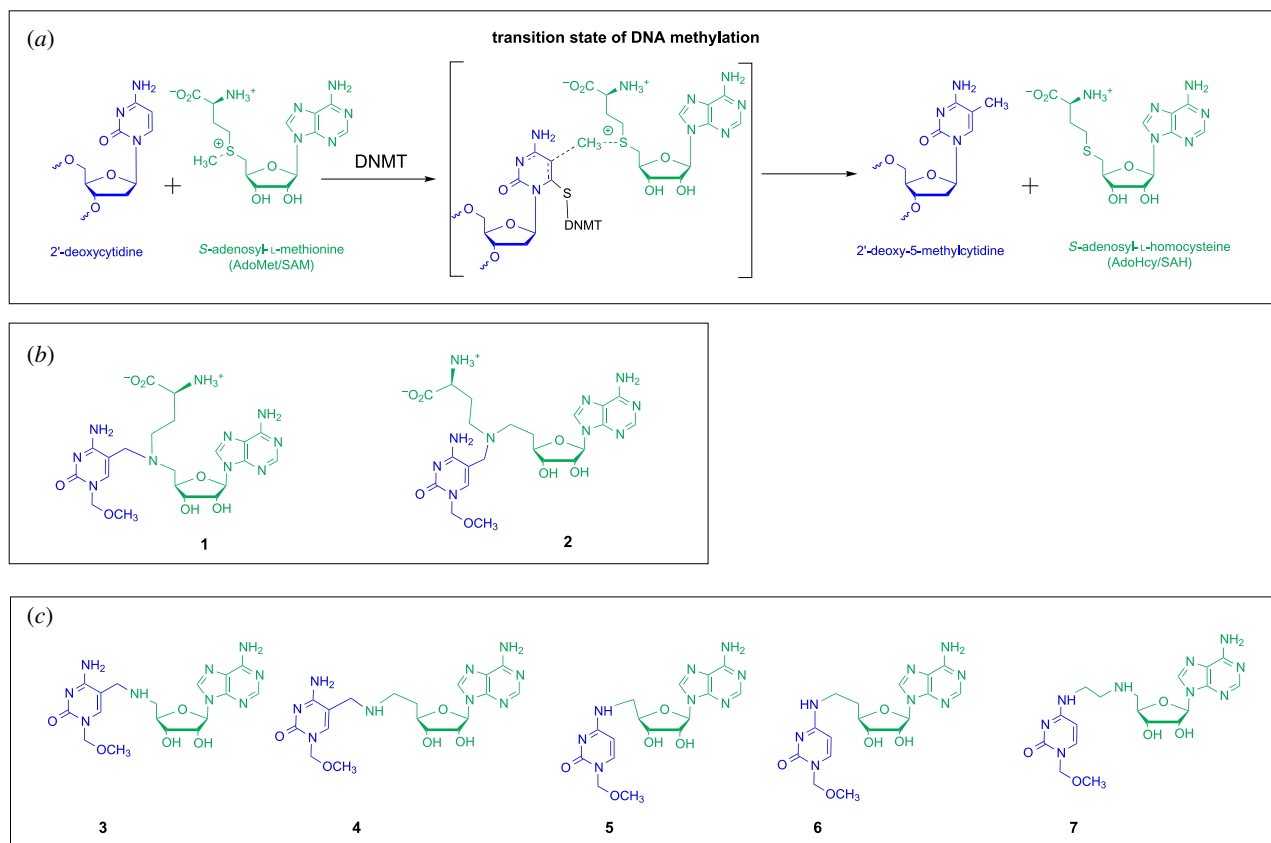
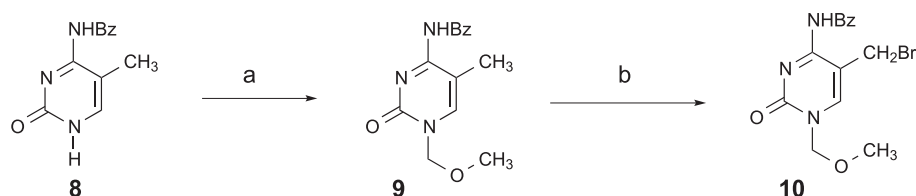


Figure 1. (a) Schematic representation of the methylation reaction into the catalytic pocket of DNMT1 and DNMT3A. The methyl-donor SAM and its analogues are shown in green, and the cytidine in the DNA and its cytosine analogues in blue. (b) Synthesized transition state analogues. (c) Exploring the length of the linkers and the attachment point on the cytosine.



Scheme 1. Synthesis of cytosine analogue **10**. (a) MOMCl, BSA, DCM, RT, 18 h. (b) NBS, AIBN, benzene, 70°C, 2 h.

which is known to induce a chemoselectivity in secondary amine formation [42]. In contrast, the use of the nosyl group afforded more than 64% yield over the three steps of protection, coupling and deprotection.

Compounds **3** and **4**, lacking the amino acid moiety, were obtained using the same synthetic pathway (scheme 2) by *N*-alkylation of the nosylated derivatives **13** and **14** with the protected bromomethylcytosine **10** in the presence of K_2CO_3 , followed by ammonia treatment in methanol to remove the benzoyl group of the cytosine moiety and then thiophenol treatment, affording compounds **19** and **20**. Treatment with TFA in water afforded the desired bisubstrate analogues **3** and **4**. Of note, the order of the deprotection steps is crucial: the benzoyl protective group had to be removed first because we observed that the presence of the secondary amine made it difficult to remove. Indeed, the presence of an NH group on the C5 methyl group of the cytosine moiety seems to stabilize the benzoyl protective group on the amine at the C4 position. Neither NH_3 nor KOH in methanol or water at RT or under reflux removed the benzoyl group.

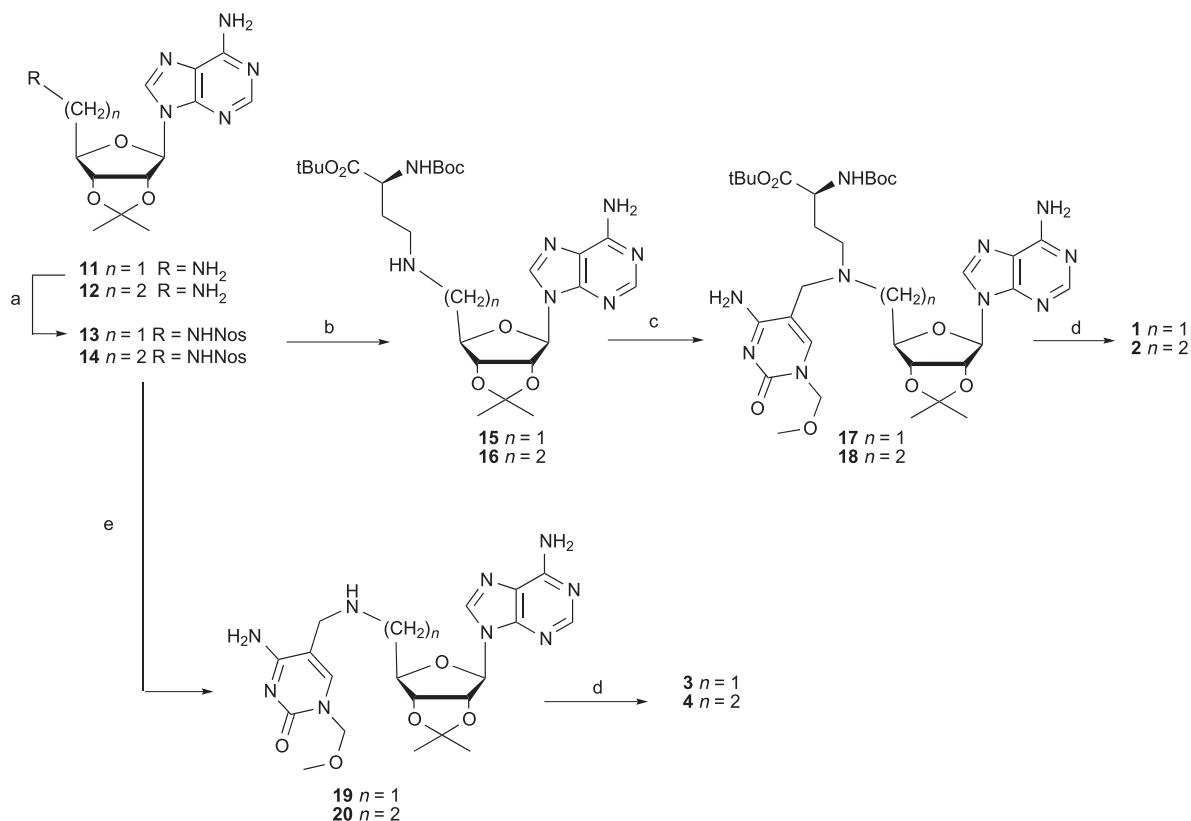
Analogues **5** and **6**, linked to the adenosine moiety by the N4 position of cytosine, were obtained starting from amines

11 and **12** (scheme 3). *N*1-MOM-protected uracil derivative **22** was prepared from uracil (**21**) using the same conditions as for **9** (scheme 1). Next, **22** was activated with 1,2,4-triazole and $POCl_3$ in the presence of TEA and reacted with the amine derivatives **11** and **12** to afford compounds **5** and **6**, respectively, after removal of the isopropylidene groups of intermediates **23** and **24** upon acidic treatment.

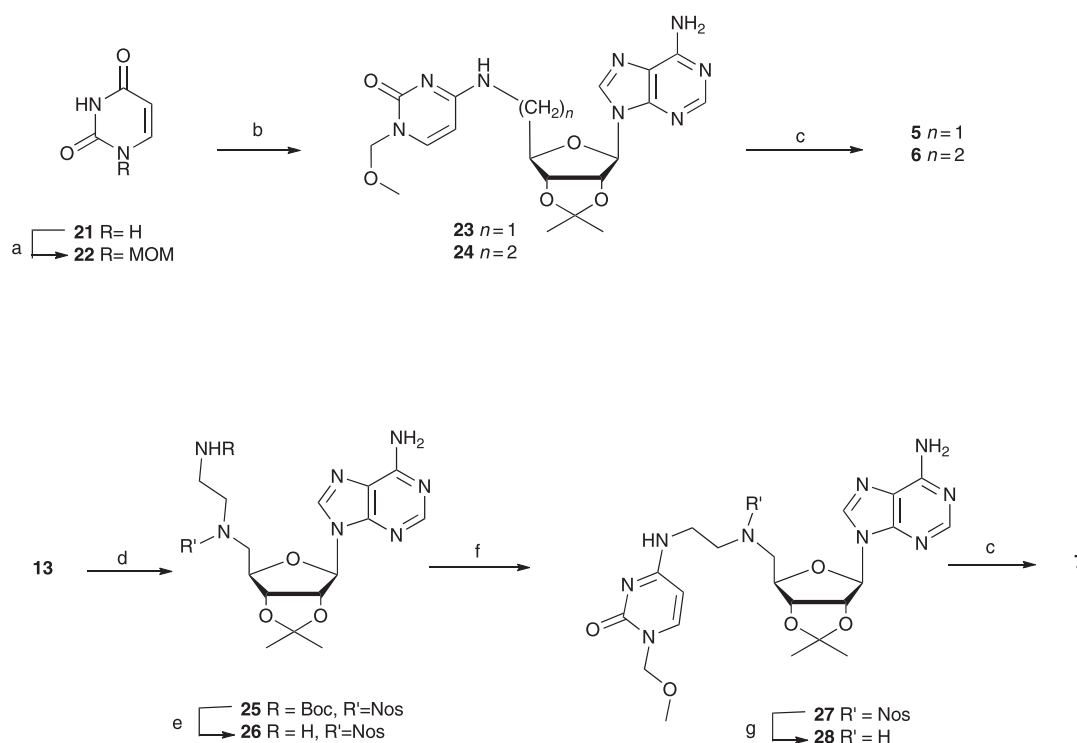
Finally, analogue **7**, containing an ethyl spacer between the adenosine and the cytosine moieties, was obtained from nosylated **13**, which was *N*-alkylated by 2-(Boc-amino)-ethyl bromide to give **25** (scheme 3). After Boc-deprotecting by TFA, the obtained amine **26** was reacted with activated uracil derivative **22** to give **27**. Desired compound **7** was obtained under acidic conditions from **28** obtained after deprotection of **27** by thiophenol.

(c) Biological activity

Compounds **1** to **7** were tested for their ability to inhibit the methylation of DNA duplexes by human catalytic DNMT3A2 and full length DNMT1 using the assays previously developed [19]. None of the compounds, even compounds **1** and **2**, which



Scheme 2. Synthesis of transition state analogues **1** and **2**. (a) NosCl, pyridine, RT, 6 h, 83% for **13** and 77% for **14**. (b) (i) (*S*)-2-*tert*-butoxycarbonylamino-4-bromobutyric acid *tert*-butylester, DMF, K_2CO_3 , RT, 2 h. (ii) PhSH, K_2CO_3 , DMF, RT, 18 h, 83% for **15** and 88% for **16**. (c) (i) **10**, DMF, DiPEA, RT, 1 h. (ii) 7 N NH_3 in MeOH, 64% for **17** and 44% for **18**. (d) TFA, water, RT, 1 h, 65% for **1**, 87% for **2**, 89% for **3** and 89% for **4**. (e) (i) **10**, DMF, K_2CO_3 , RT, 1 h. (ii) 7 N NH_3 in MeOH. (iii) PhSH, K_2CO_3 , DMF, RT, 18 h, 21% for **19** and 35% for **20**.



Scheme 3. (a) MOMCl, BSA, DCM, RT, 18 h, 95%. (b) (i) **22**, POCl_3 , 1,2,4-triazole, TEA, MeCN, 12 h, 0°C to RT; (ii) **11** or **12**, MeCN, RT, 6 h, 47% for **23** and 37% for **24**. (c) TFA, water, RT, 1 h, 85% for **5**, 74% for **6** and 83% for **7**. (d) 2-(Boc-amino)-ethyl bromide, TEA, DMF, 65°C, 2 h, 63%. (e) TFA, RT, 1 h, 90%. (f) (i) **22**, POCl_3 , 1,2,4-triazole, TEA, MeCN, 12 h, 0°C to RT; (ii) **11**, MeCN, RT, 6 h, 48%. (g) PhSH, K_2CO_3 , DMF, RT, 18 h, 65%.

mimic more closely the major transition state of DNMTs [37], was found to be effective in inhibiting hDNMT3Acet and hDNMT1 (table 2). This is surprising, since Wahhab and colleagues observed a strong inhibition of DNMT1 and DNMT3b2 ($\text{IC}_{50} < 10 \mu\text{M}$) with amino AdoHcy derivatives, in

which the 5' amine group of the adenosine was functionalized with a benzyl substituent such as 4-nitrobenzyl group [43–45].

Next, we tested the compounds for their ability to inhibit TRDM1, since this cytosine RNA MTase is structurally related to the DNMTs, being discovered by sequence homology to the

Table 2. Screening for inhibitory potential of transition state analogues against DNA and RNA MTases. The percentage of inhibition of the enzymatic activity is reported for experiments run in duplicates. —, not determined.

Cpd	DNMT		RNA MTase							
	DNMT1	DNMT3A	TRDMT1	WNV-MTase	MERS 2'O-MTase	MERS N7-MTase	SARS 2'O-MTase	SARS N7-MTase	hRNMT	DV-MTase
	32 μ M	20 μ M	1 mM	50 μ M	50 μ M	50 μ M	50 μ M	50 μ M	50 μ M	50 μ M
1	0	0	46 \pm 9	35	0	0	0	40	90	10
2	0	0	0	10	0	0	0	60	60	0
3	0	0	0	0	0	0	0	0	0	0
4	0	0	0	10	0	0	0	0	0	0
5	0	0	—	10	0	0	0	0	0	0
6	0	0	0	25	0	0	0	10	45	5
7	0	0	0	0	0	0	0	0	0	0

Table 3. Screening for PRMT inhibition. Percentage of inhibition at the indicated concentration. —, not determined.

Cpd	PRMT1		PRMT4			PRMT6		PRMT7	
	100 μ M	10 μ M	100 μ M	10 μ M	IC ₅₀ (μ M)	100 μ M	10 μ M	100 μ M	10 μ M
1	24 \pm 2	0	59 \pm 0.6	21 \pm 7	—	61 \pm 2	3 \pm 9	81 \pm 2	30 \pm 0.5
2	42 \pm 6	0	88 \pm 2	54 \pm 2	14 \pm 1	47 \pm 1	8 \pm 12	83 \pm 7	46 \pm 3
3	—	—	21 \pm 13	−7 \pm 0.3	—	—	—	12	0
4	13 \pm 5	0	99 \pm 0.2	81 \pm 1	1.5 \pm 0.5	20 \pm 2	−3 \pm 2	89 \pm 1	33 \pm 1
5	—	—	95 \pm 0.5	18 \pm 0.8	—	—	—	88 \pm 6	19 \pm 6
6	—	—	50 \pm 0.8	5 \pm 0.6	—	—	—	36 \pm 9	9 \pm 9
7	—	—	66 \pm 2	25 \pm 1	—	—	—	63 \pm 4	4 \pm 1

DNMTs and once called DNMT2 till it was discovered that it methylated RNA [7,8]. Only compound 1, the transition state amino analogue bearing the amino acid group, showed some inhibition at a high concentration of 1 mM (table 2) with an inhibition activity of 20% at 100 μ M. This incited us to screen the compounds against several viral and human RNA MTases involved in the methylation of the RNA cap structure. These MTases have been previously shown to methylate cap structure either at the N7 position of guanine (WNV-MTase, MERS-N7-MTase, SARS N7-MTase, hRNMT) or/and at the 2'OH position of the first nucleotide ribose (WNV-MTase, MERS 2'O-MTase, SARS 2'O-MTase, DV MTase). Again only the amino acid derivatives showed some inhibitory potency at 50 μ M, in particular compound 1, against hRNMT.

Next the compounds were screened against selected human PRMTs (table 3), identified to play a role in cancer development and maintenance, and co-crystallization studies were performed with compounds 1, 2, 3, 4 and 7 with an isolated catalytic domain of MmPRMT4 (residues 130 to 487) [46,47] or with PRMT6 [30]. Interestingly, compound 2 and its analogue 4 showed an activity against PRMT4, also called coactivator-associated arginine methyltransferase 1 (CARM1), and some activity against PRMT7 (table 3). IC₅₀ values against PRMT4 of 1.9 and 14 μ M were measured for 4 and 2, respectively. The

structural data brought insights on how the 5-methylcytosine-adenosine transition state analogues inhibit the PRMTs.

(d) Structural insights on the inhibition of PRMTs

Concerning PRMT4, except for compound 2, high-resolution crystal structures were obtained and the structures were solved and refined in two closely related space groups P2₁2₁2 and P2₁ as seen for the MmPRMT4 catalytic domain. The orthorhombic space group contains one copy of the PRMT4 tetramer in the asymmetric unit, while the monoclinic space group has two PRMT4 tetramers in the asymmetric unit (table 1).

These co-crystal structures revealed that the compounds are able to displace the SAH. The purified MmPRMT4 construct naturally contains SAH molecules bound in the active site. By comparing the PRMT4–compound structures with PRMT4–SAH structures (PDB ID: 5IH3), two classes of compounds can be defined: (i) that displacing all SAH molecules natively bound to the PRMT4 (compound 4) and (ii) those partially displacing the SAH molecules (compounds 1, 3 and 7). The compounds bind to the PRMT active site and occupy the two substrate binding sites (figure 2): the SAM cofactor and the arginine site (to be methylated)—the cytosine moieties probing the space filled by a substrate peptide

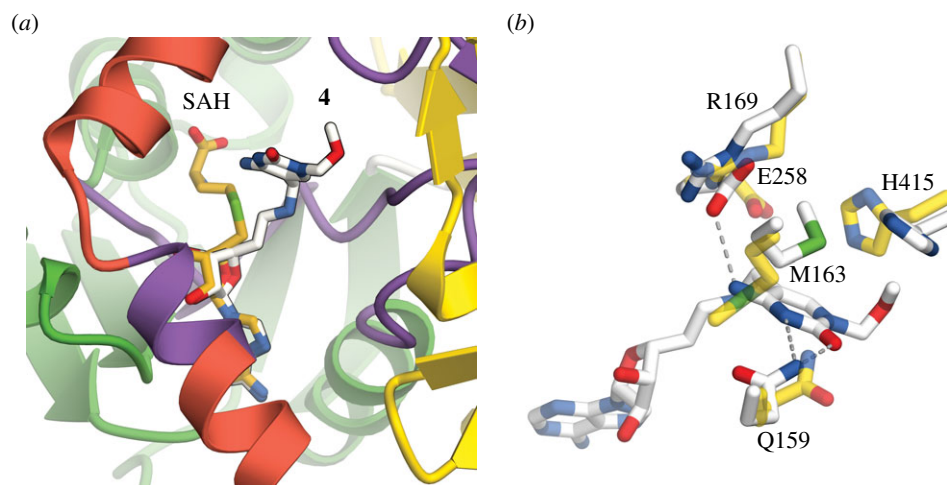


Figure 2. Binding mode and interactions of compound **4** on PRMT4. Conformation of the inhibitor **4** in the catalytic site as compared with SAH (*a*). The SAH molecule is coloured in yellow. The colour code for PRMT4 is the one defined in electronic supplementary material, figure S1. Concerted side chain movement upon **4** binding as compared with SAH (*b*). Residues are drawn as stick models and labelled (PDB: 5TBJ). Hydrogen bonds are drawn as grey dashed lines. Residues from the SAH structure (PDB: 5IH3) are coloured in yellow and slightly transparent. Water molecules are not shown.

methylation intermediate. This is clearly represented by comparison with previously solved structures [29,48] (electronic supplementary material, figure S1). For all inhibitors, the interactions with the base and sugar moieties are strictly conserved and identical to what is observed for SAH [46].

(i) Strong inhibition of PRMT4 by analogue **4**

Compound **4** shows the strongest potency against PRMT4 in the inhibition assays. It lacks the amino acid moiety and it is distant from the 5' position of the adenosine by an additional carbon. In the co-crystal structure, the cytosine moiety occupies the place of the target arginine (figure 2*a*; electronic supplementary material, figure S1*c,d*). Binding of **4** induces the concerted movement of mainly five residue side chains, Gln159, Met163, Arg169, Glu258 and His415 (figure 2*b* and electronic supplementary material, figure S2). All PRMTs share a common active site architecture defined by specific amino acids (belonging to the so-called motifs I to IV) known to be key for catalysis [11–13,49,50]. Glu258 belongs to the double-E loop (motif 3) and His415 to the THW loop (motif 4), conserved residues crucial for catalysis by all PRMTs. Arg169 is also conserved in all PRMTs as it interacts with the carboxylate atoms of the SAM/SAH, and Met169 is conserved in all type-I PRMTs. Gln159 is a crucial residue, only present in PRMT4 sequences, and here interacts by two hydrogen bonds with the cytosine moiety (figure 2*b* and electronic supplementary material, figure S1*c,d*). This latter interaction seems to be the driving force for the observed inhibition selectivity of the compounds on PRMT4.

(ii) The other compounds

Of note, compound **2** differs from **4** by the presence of the amino acid moiety, which results in a decrease of inhibition (1.9 to 14 μM for IC_{50}). However, no co-crystal structure could be obtained with PRMT4.

Compounds **1**, **3** and **7** are weak inhibitors of PRMT4 (table 3) and displaced only 50% of the naturally bound SAH molecules on PRMT4 (only one SAH is replaced by the compound per dimer, figure 3). Compounds **1** and **3** are derivatives of **2** and **4**, respectively, but with only one carbon atom in

the linker between the sugar and the nitrogen atom, while **7** is branched through position N4 of the cytosine. In the case of all three compounds, the co-crystal structures show that the cytosine moiety goes less deep into the arginine substrate binding site (overlay of the bound compounds in the co-crystal shown in electronic supplementary material, figure S2), which can also partly explain the reduced inhibition. In addition, the co-crystals with **1** (figure 3*a,b*) and **7** (figure 3*c,d*) reveal that binding of each compound produces no conformational changes except a small reorientation of Gln159 and a different water molecule network (electronic supplementary material, figure S3). In contrast, the weakest inhibitor compound **3** induces conformational changes of three residues, His415, Arg169 and Arg169 side chains, involved in the interaction with the cytosine moiety as shown by compound **4** (figure 3*e,f* and electronic supplementary material, figure S3).

(iii) Complementary insights from PRMT6 co-crystals

Only compounds **1** and **2** bearing the amino acid moiety were able to co-crystallize with PRMT6. The two compounds differ by the length of the link between the sugar and the nitrogen. The cytosine moiety has very similar conformations for both compounds and their binding is without any conformational changes compared with SAH binding on PRMT6 (figure 4 and electronic supplementary material, figure S4), explaining the inhibitory data (table 3).

(e) Cancer cell experiments

Because of the role of PRMT4 in breast, colon and prostate cancer, we tested compounds **2** and **4** in prostate DU145, colon HCT116 and breast MCF7 cancer cell lines. These cell lines were treated with **2** and **4** at 1 and 10 μM and analysed after 3, 7 or 11 days of treatment. The compounds did not alter the proliferation of the cells (electronic supplementary material, figure S5) and compound **4** had no effect on the methylation of H3R17, a substrate of PRMT4, even after 11 days of treatment (electronic supplementary material, figure S6). In view of the elevated polar surface area of compound **4**, permeability studies were further carried out in Caco-2 cells. The apparent permeability coefficient (P_{app}) was

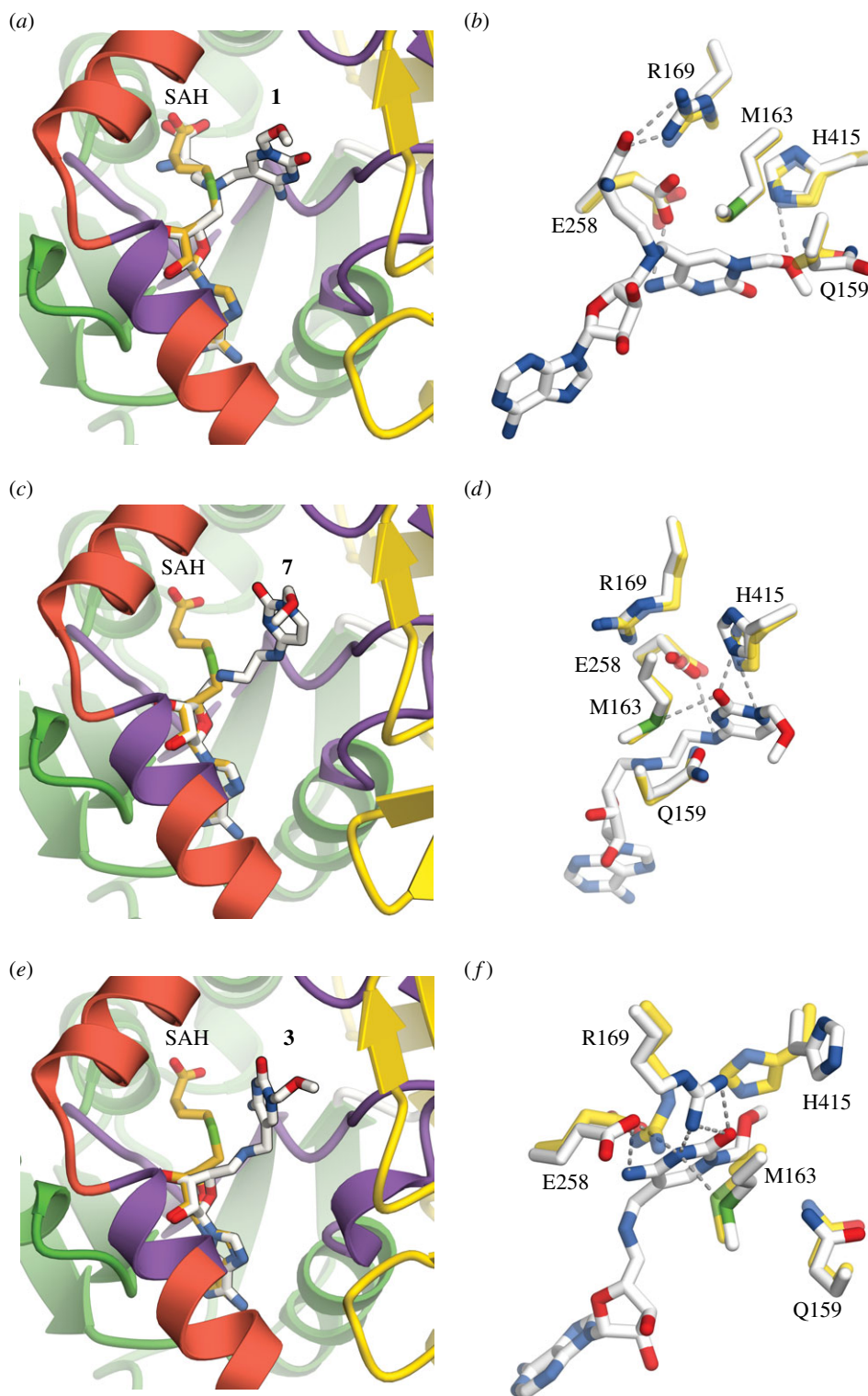


Figure 3. Binding mode and interactions of **1**, **7** and **3** on PRMT4. The colour code for PRMT4 is the one defined in figure 2. Hydrogen bonds are drawn as grey dashed lines. Residues from the SAH structure are coloured in yellow and slightly transparent. Water molecules are not shown. Conformation of **1** in the catalytic site as compared with SAH (a) and concerted side chain movement upon **1** binding as compared with SAH (b). Residues are drawn as stick models and labelled (PDB: 5TBH). Conformation of the inhibitor **7** in the catalytic site as compared with SAH (c) and concerted side chain movement upon **7** binding as compared with SAH (d). Residues are drawn as stick models and labelled (PDB: 5TBI). Conformation of **3** in the catalytic site as compared with SAH (e). Concerted side chain movement upon **3** binding as compared with SAH (f). Residues are drawn as stick models and labelled (PDB: 5LV2).

determined in the apical-to-basolateral (A–B; pH6.5/7.4)) and/or the B–A directions across cell monolayers cultured on 96-well polycarbonate membrane filters. Compound **4** showed good solubility and was tested at 10 μM with a final DMSO concentration of 1%. Compound **4** showed a $P_{\text{app}}(\text{A–B})$ of $1.3 \times 10^{-6} \text{ cm s}^{-1}$ and $P_{\text{app}}(\text{B–A})$ of $0.3 \times 10^{-6} \text{ cm s}^{-1}$,

demonstrating a lack of efflux through the P-gp transporter. Both P_{app} values are in agreement with a very poor ability of compound **4** to cross the cell membrane. As a result, the lack of antiproliferative activity of compound **4** is most likely due to poor target engagement due to low concentration levels of the chemical entity in the intracellular medium.

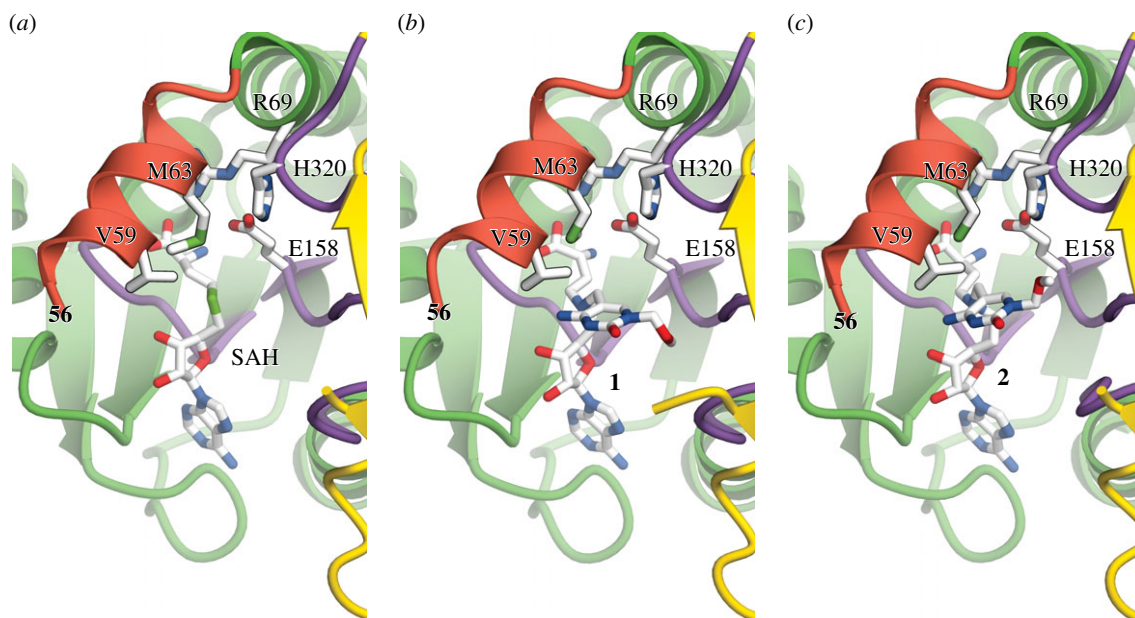


Figure 4. Binding mode and interactions of SAH, **1** and **2** on PRMT6. Binding mode of SAH to PRMT6 (PDB: 5FQN) (a). Conformation of critical residues upon **1** binding on PRMT6 (b) and **2** binding on PRMT6 (c). PRMT6 monomer is drawn as a cartoon model with the same colour code used in previous figures. Residue side chains and the ligands are drawn as stick models. The number of the first visible N-terminal residue is indicated. Water molecules are not shown.

4. Conclusion

In conclusion, we developed a convergent synthetic pathway starting from a protected bromomethylcytosine derivative to obtain transition state analogues of the DNMTs bearing a 5-methylcytosine scaffold linked to adenosine via an *N*-alkyl linker. This approach led to seven cytosine-adenosine compounds that were, surprisingly, inactive against hDNMT1 and hDNMT3Acat. Three reasons could explain this: either (1) the inability of the analogues to enter the pocket and displace the cofactor, or (2) their geometry—for example the work of Du *et al.* [37] suggests that an sp^3 configuration of the C5 carbon of the cytosine moiety could be necessary instead of the sp^2 configuration of **1** to **4**, or (3) the distances of the bonds do not align with the transition state complex. Further analogues will be needed to test these hypotheses.

The compounds were also tested against several RNA MTases, such as TRDMT1, structurally closely related to the DNMTs, and viral RNA MTases. Weak activity was observed for a transition state analogue containing the amino acid side chain.

We next screened the compounds against several PRMTs, selected for their potential role in cancer. Interestingly, compound **4** and its derivative **2** showed inhibitory activity against PRMT4 at the micromolar level.

Crystal structures of PRMT4 and PRMT6 in the presence of the compounds revealed that some inhibitors displace totally (compound **4**) or partially (compounds **3** and **7**) the SAM cofactor from the PRMTs (figures 2–4). The four co-crystal structures of PRMT4 (PRMT4–4, PRMT4–3, PRMT4–7, PRMT4–1) and two co-crystal structures of PRMT6 (PRMT6–2, PRMT6–1) revealed the following. First, the compounds bind to the PRMT active site and occupy the two substrate binding sites: the SAM and the arginine site (to be methylated) (electronic supplementary material, figure S1). Most interestingly, the cytosine moiety fills in the space left by the arginine, explaining the inhibition data and confirming our working hypothesis (electronic supplementary material,

figures S3 and S4). Second, five structural switches are involved in the strongest *in vitro* inhibition. These conformation changes involve residues strongly conserved in type-I PRMTs and amino acids specific to PRMT4.

Interestingly, increasing evidence shows the specific involvement of PRMT4 in promoting tumorigenesis in solid and haematological cancers [9,51–56]. Thus the discovery of small molecule inhibitors of PRMTs has been the focus of many studies during the last years and several compounds have been described [24,25,29,57–61]. High-throughput screening efforts led to the identification of pyrazole and benzimidazole derivatives as potent PRMT4 inhibitors, the most potent showing IC_{50} values of 27 and 30 nM [61]. These inhibitors bind to the substrate arginine-binding cavity of PRMT4 but require the presence of bound SAH. Among the mimics of the SAM scaffold and more particularly the bisubstrate analogues designed to simultaneously occupy substrate and cofactor binding sites ([62] for review), the most potent and specific inhibition shows nanomolar range IC_{50} values [60].

Here we identified a new scaffold derived from the design of transition state analogues of the DNMTs to inhibit PRMTs and in particular PRMT4. Compound **4** is able to displace the four bound SAH molecules from the catalytic site and the cytosine moiety to occupy the space filled by the substrate peptide methylation intermediate, inducing crucial structural changes. The scaffold has not been optimized and further pharmacophoric modulations are needed to increase the potency and selectivity. The co-crystal structures highlighted the chemical modifications that can be done and will guide the optimization. In addition the cellular penetration properties need improvement to obtain an effect in cancer cells. These findings open new routes for the conception of new PRMT4 inhibitors based on the 5-methylcytosine scaffold.

Data accessibility. All structures and structure factors have been deposited in the Protein Data Bank (PRMT4–1: 5TBH; PRMT4–3: 5LV2; PRMT4–7: 5TBI; PRMT4–4: 5TBJ, PRMT6–1: 5LV4, PRMT6–2: 5LV5).

Authors' contributions. L.H. participated in the design of the compounds, carried out the chemical laboratory work, participated in data analysis and helped draft the manuscript; D.P. and C.Fa. carried out the DNMT assays; D.-M.F. and F.A. carried out the cellular assays; W.A., E.D. and J.-C.G. carried out the RNA methylases assays; A.J. and S.K. carried out the TRDM1 assays. D.G. and C.Fe. participated in the conception of the study, and D.G. proofread carefully the manuscript. N.M., V.C., N.T.-C. and L.B. performed the crystallography studies and analysis. P.P. was in charge of the stability and permeability data. J.C. conceived and directed the structural studies and participated in drafting the manuscript. P.B.A. conceived, designed and coordinated the study, and drafted the manuscript. All authors gave their approval for publication.

Competing interests. We have no competing interests.

Funding. This work was supported by funding to P.B.A. from Fondation InNaBioSanté, PlanCancer2014 (no. EPIG201401), CNRS and the National Research Agency (ANR: Agence Nationale de la Recherche) 'Investissement d'avenir' (ANR-11-PHUC-001, CAPTOR research programme). N.M., V.C., L.B., N.T.-C. and J.C. are supported by grants from CNRS, Université de Strasbourg, INSERM,

Instruct, part of the European Strategy Forum on Research Infrastructures (ESFRI) supported by national member subscriptions as well as the French Infrastructure for Integrated Structural Biology (FRISBI) (ANR-10-INSB-05-01), grant ANR-10-LABX-0030-INRT, a French State fund managed by the Agence Nationale de la Recherche under the frame programme Investissements d'Avenir (ANR-10-IDEX-0002-02 to J.C.) and IGBMC and grants from Association pour la Recherche contre le Cancer (ARC) (nos A09/4/5005, no. SFI20121205902 and PJA 20161204817)). N.M. is an IGBMC International PhD Programme fellow supported by LabEx INRT funds.

Acknowledgements. We thank Dr Alain Wagner (CNRS, University of Strasbourg) and Jean-Marie Wurtz (IGBMC, Illkirch) for helpful discussions. We thank the members of the IGBMC common services and the members of the structural biology platform of IGBMC for technical assistance. We thank members of SOLEIL Proxima1 beamlines, the SOLEIL staff for smoothly running the facility, the staff of the European Synchrotron Radiation Facility–European Molecular Biology Laboratory joint Structural Biology groups for the use of beamline facilities, support in using beamline(s) and for assistance during X-ray data collection.

References

- Baylin SB, Jones PA. 2011 A decade of exploring the cancer epigenome—biological and translational implications. *Nat. Rev. Cancer* **11**, 726–734. (doi:10.1038/nrc3130)
- Zhao BS, Roundtree IA, He C. 2017 Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.* **18**, 31–42. (doi:10.1038/nrm.2016.132)
- Gros C, Fahy J, Halby L, Dufau I, Erdmann A, Gregoire JM, Ausseil F, Vispe S, Arimondo PB. 2012 DNA methylation inhibitors in cancer: recent and future approaches. *Biochimie* **94**, 2280–2296. (doi:10.1016/j.biochi.2012.07.025)
- Jones PA. 2012 Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484–492. (doi:10.1038/nrg3230)
- Jurkowska RZ, Jurkowski TP, Jeltsch A. 2011 Structure and function of mammalian DNA methyltransferases. *Chembiochem* **12**, 206–222. (doi:10.1002/cbic.201000195)
- Lopez M, Halby L, Arimondo PB. 2016 DNA methyltransferase inhibitors: development and applications. *Adv. Exp. Med. Biol.* **945**, 431–473. (doi:10.1007/978-3-319-43624-1_16)
- Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh CL, Zhang X, Golik KG, Jacobsen SE, Bestor TH. 2006 Methylation of tRNA^{Asp} by the DNA methyltransferase homolog Dnmt2. *Science* **311**, 395–398. (doi:10.1126/science.1120976)
- Schaefer M, Lyko F. 2010 Solving the Dnmt2 enigma. *Chromosoma* **119**, 35–40. (doi:10.1007/s00412-009-0240-6)
- Cantara WA, Crain PF, Rozenski J, McCloskey JA, Harris KA, Zhang X, Vendeix FA, Fabris D, Agris PF. 2011 The RNA modification database, RNAMDB: 2011 update. *Nucleic Acids Res.* **39**, D195–D201. (doi:10.1093/nar/gkq1028)
- Traube FR, Carell T. 2017 The chemistries and consequences of DNA and RNA methylation and demethylation. *RNA Biol.* **14**, 1099–1107. (doi:10.1080/15476286.2017.1318241)
- Bedford MT, Clarke SG. 2009 Protein arginine methylation in mammals: who, what, and why. *Mol. Cell* **33**, 1–13. (doi:10.1016/j.molcel.2008.12.013)
- Yang Y, Bedford MT. 2013 Protein arginine methyltransferases and cancer. *Nat. Rev. Cancer* **13**, 37–50. (doi:10.1038/nrc3409)
- Fuhrmann J, Clancy KW, Thompson PR. 2015 Chemical biology of protein arginine modifications in epigenetic regulation. *Chem. Rev.* **115**, 5413–5461. (doi:10.1021/acs.chemrev.5b00003)
- Blanc RS, Richard S. 2017 Arginine methylation: the coming of age. *Mol. Cell* **65**, 8–24. (doi:10.1016/j.molcel.2016.11.003)
- Morales Y, Caceres T, May K, Hevel JM. 2016 Biochemistry and regulation of the protein arginine methyltransferases (PRMTs). *Arch. Biochem. Biophys.* **590**, 138–152. (doi:10.1016/j.abb.2015.11.030)
- Hu H, Qian K, Ho MC, Zheng YG. 2016 Small molecule inhibitors of protein arginine methyltransferases. *Expert Opin. Investig. Drugs* **25**, 335–358. (doi:10.1517/13543784.2016.1144747)
- Buchini S, Leumann CJ. 2006 2'-O-Aminoethyl oligoribonucleotides containing novel base analogues: synthesis and triple-helix formation at pyrimidine/purine inversion sites. *Eur. J. Org. Chem.* **14**, 3152–3168. (doi:10.1002/ejoc.200600182)
- Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X. 2007 Structure of Dnmt3a bound to Dnmt3L suggests a model for *de novo* DNA methylation. *Nature* **449**, 248–251. (doi:10.1038/nature06146)
- Gros C, Chauvigne L, Poulet A, Menon Y, Ausseil F, Dufau I, Arimondo PB. 2013 Development of a universal radioactive DNA methyltransferase inhibition test for high-throughput screening and mechanistic studies. *Nucleic Acids Res.* **41**, e185. (doi:10.1093/nar/gkt753)
- Elhardt W, Shanmugam R, Jurkowski TP, Jeltsch A. 2015 Somatic cancer mutations in the DNMT2 tRNA methyltransferase alter its catalytic properties. *Biochimie* **112**, 66–72. (doi:10.1016/j.biochi.2015.02.022)
- Bouvet M, Debarnot C, Imbert I, Selisko B, Snijder EJ, Canard B, Decroly E. 2010 *In vitro* reconstitution of SARS-coronavirus mRNA cap methylation. *PLoS Pathog.* **6**, e1000863. (doi:10.1371/journal.ppat.1000863)
- Barral K *et al.* 2013 Development of specific dengue virus 2'-O- and N7-methyltransferase assays for antiviral drug screening. *Antiviral Res.* **99**, 292–300. (doi:10.1016/j.antiviral.2013.06.001)
- Aouadi W, Blanjoie A, Vasseur JJ, Debart F, Canard B, Decroly E. 2017 Binding of the methyl donor S-adenosyl-L-methionine to Middle East respiratory syndrome coronavirus 2'-O-methyltransferase nsp16 promotes recruitment of the allosteric activator nsp10. *J. Virol.* **91**, e02217-16. (doi:10.1128/JVI.02217-16)
- Selvi BR *et al.* 2010 Identification of a novel inhibitor of coactivator-associated arginine methyltransferase 1 (CARM1)-mediated methylation of histone H3 Arg-17. *J. Biol. Chem.* **285**, 7143–7152. (doi:10.1074/jbc.M109.063933)
- Cheng D, Yadav N, King RW, Swanson MS, Weinstein EJ, Bedford MT. 2004 Small molecule regulators of protein arginine methyltransferases. *J. Biol. Chem.* **279**, 23 892–23 899. (doi:10.1074/jbc.M401853200)
- Iberg AN, Espejo A, Cheng D, Kim D, Michaud-Levesque J, Richard S, Bedford MT. 2008 Arginine methylation of the histone H3 tail impedes effector binding. *J. Biol. Chem.* **283**, 3006–3010. (doi:10.1074/jbc.C700192200)
- Zurita-Lopez CI, Sandberg T, Kelly R, Clarke SG. 2012 Human protein arginine methyltransferase 7 (PRMT7) is a type III enzyme forming ω -*N*⁶-monomethylated arginine residues. *J. Biol. Chem.* **287**, 7859–7870. (doi:10.1074/jbc.M111.336271)
- Chen D, Ma H, Hong H, Koh SS, Huang SM, Schurter BT, Aswad DW, Stallcup MR. 1999 Regulation of transcription by a protein methyltransferase. *Science* **284**, 2174–2177. (doi:10.1126/science.284.5423.2174)

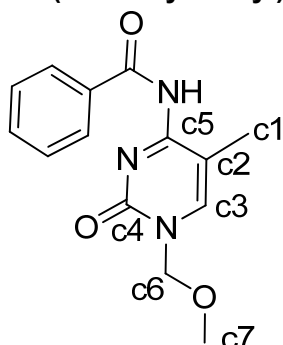
29. van Haren MJ, Marechal N, Troffer-Charlier N, Cianciulli A, Sbardella G, Cavarelli J, Martin NI. 2017 Transition state mimics are valuable mechanistic probes for structural studies with the arginine methyltransferase CARM1. *Proc. Natl Acad. Sci. USA* **114**, 3625–3630. (doi:10.1073/pnas.1618401114)
30. Bonnefond L, Stojko J, Mailliot J, Troffer-Charlier N, Cura V, Wurtz JM, Cianferani S, Cavarelli J. 2015 Functional insights from high resolution structures of mouse protein arginine methyltransferase 6. *J. Struct. Biol.* **191**, 175–183. (doi:10.1016/j.jsb.2015.06.017)
31. Kabsch W. 2010 XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132. (doi:10.1107/S0907444909047337)
32. Otwinowski Z, Minor W. 1997 Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326. (doi:10.1016/S0076-6879(97)76066-X)
33. Emsley P, Cowtan K. 2004 Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132. (doi:10.1107/S0907444904019158)
34. Adams PD *et al.* 2010 PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221. (doi:10.1107/S0907444909052925)
35. Bricogne G *et al.* 2011 *BUSTER version 2.10.0*. Cambridge, UK: Global Phasing.
36. Karplus PA, Diederichs K. 2015 Assessing and maximizing data quality in macromolecular crystallography. *Curr. Opin. Struct. Biol.* **34**, 60–68. (doi:10.1016/j.sbi.2015.07.003)
37. Du Q, Wang Z, Schramm VL. 2016 Human DNMT1 transition state structure. *Proc. Natl Acad. Sci. USA* **113**, 2916–2921. (doi:10.1073/pnas.1522491113)
38. Mai V, Comstock LR. 2011 Synthesis of an azide-bearing *N*-mustard analogue of *S*-adenosyl-*L*-methionine. *J. Org. Chem.* **76**, 10 319–10 324. (doi:10.1021/jo2019637)
39. Lerner C, Masjost B, Ruf A, Gramlich V, Jakob-Roetne R, Zurcher G, Borroni E, Diederich F. 2003 Bisubstrate inhibitors for the enzyme catechol-*O*-methyltransferase (COMT): influence of inhibitor preorganisation and linker length between the two substrate moieties on binding affinity. *Org. Biomol. Chem.* **1**, 42–49. (doi:10.1039/B208690P)
40. Day FA, Launay DFM, Charlton MH, Moffat DFC. 2010 Pyrrolo [2, 3-*d*] pyrimidines as inhibitors of hsp90. Patent WO 2010043867 A1.
41. Diouy F, Guéné E, Di Scala-Roulleau A, Ferroud C, Guy A, Port M. 2005 Synthesis of a new tricyclic tetraazatriacetic acid as ligand for gadolinium(III). *Tetrahedron Lett.* **46**, 611–613. (doi:10.1016/j.tetlet.2004.11.133)
42. Salvatore RN, Nagle AS, Jung KW. 2002 Cesium effect: high chemoselectivity in direct *N*-alkylation of amines. *J. Org. Chem.* **67**, 674–683. (doi:10.1021/jo010643c)
43. Wahhab A. 2006 Inhibitors of DNA methyltransferase. Patent WO2006078752.
44. Saavedra OM *et al.* 2009 SAR around (*L*)-*S*-adenosyl-*L*-homocysteine, an inhibitor of human DNA methyltransferase (DNMT) enzymes. *Bioorg. Med. Chem. Lett.* **19**, 2747–2751. (doi:10.1016/j.bmcl.2009.03.113)
45. Isakovic L *et al.* 2009 Constrained (*L*)-*S*-adenosyl-*L*-homocysteine (SAH) analogues as DNA methyltransferase inhibitors. *Bioorg. Med. Chem. Lett.* **19**, 2742–2746. (doi:10.1016/j.bmcl.2009.03.132)
46. Troffer-Charlier N, Cura V, Hassenboehler P, Moras D, Cavarelli J. 2007 Functional insights from structures of coactivator-associated arginine methyltransferase 1 domains. *EMBO J.* **26**, 4391–4401. (doi:10.1038/sj.emboj.7601855)
47. Troffer-Charlier N, Cura V, Hassenboehler P, Moras D, Cavarelli J. 2007 Expression, purification, crystallization and preliminary crystallographic study of isolated modules of the mouse coactivator-associated arginine methyltransferase 1. *Acta Crystallogr. F* **63**, 330–333. (doi:10.1107/S1744309107011785)
48. Boriack-Sjodin PA *et al.* 2016 Structural insights into ternary complex formation of human CARM1 with various substrates. *ACS Chem. Biol.* **11**, 763–771. (doi:10.1021/acscchembio.5b00773)
49. Cha B, Jho EH. 2012 Protein arginine methyltransferases (PRMTs) as therapeutic targets. *Expert Opin. Ther. Targets* **16**, 651–664. (doi:10.1517/14728222.2012.688030)
50. Cura V, Troffer-Charlier N, Wurtz JM, Bonnefond L, Cavarelli J. 2014 Structural insight into arginine methylation by the mouse protein arginine methyltransferase 7: a zinc finger freezes the mimic of the dimeric state into a single active site. *Acta Crystallogr. D Biol. Crystallogr.* **70**, 2401–2412. (doi:10.1107/S1399004714014278)
51. Wang L *et al.* 2016 CARM1 methylates chromatin remodeling factor BAF155 to enhance tumor progression and metastasis. *Cancer Cell* **30**, 179–180. (doi:10.1016/j.ccell.2016.06.013)
52. Wang L, Zeng H, Wang Q, Zhao Z, Boyer TG, Bian X, Xu W. 2015 MED12 methylation by CARM1 sensitizes human breast cancer cells to chemotherapy drugs. *Sci. Adv.* **1**, e1500463. (doi:10.1126/sciadv.1500463)
53. El Messaoudi S *et al.* 2006 Coactivator-associated arginine methyltransferase 1 (CARM1) is a positive regulator of the *Cyclin E1* gene. *Proc. Natl Acad. Sci. USA* **103**, 13 351–13 356. (doi:10.1073/pnas.0605692103)
54. Hong H *et al.* 2004 Aberrant expression of CARM1, a transcriptional coactivator of androgen receptor, in the development of prostate carcinoma and androgen-independent status. *Cancer* **101**, 83–89. (doi:10.1002/ncr.20327)
55. Morettin A, Baldwin RM, Cote J. 2015 Arginine methyltransferases as novel therapeutic targets for breast cancer. *Mutagenesis* **30**, 177–189. (doi:10.1093/mutage/geu039)
56. Al-Dhaheiri M *et al.* 2011 CARM1 is an important determinant of ER α -dependent breast cancer cell differentiation and proliferation in breast cancer cells. *Cancer Res.* **71**, 2118–2128. (doi:10.1158/0008-5472.CAN-10-2426)
57. Ferreira de Freitas R *et al.* 2016 Discovery of a potent and selective coactivator associated arginine methyltransferase 1 (CARM1) inhibitor by virtual screening. *J. Med. Chem.* **59**, 6838–6847. (doi:10.1021/acs.jmedchem.6b00668)
58. Huynh T *et al.* 2009 Optimization of pyrazole inhibitors of coactivator associated arginine methyltransferase 1 (CARM1). *Bioorg. Med. Chem. Lett.* **19**, 2924–2927. (doi:10.1016/j.bmcl.2009.04.075)
59. Kaniskan HU *et al.* 2016 Design and synthesis of selective, small molecule inhibitors of coactivator-associated arginine methyltransferase 1 (CARM1). *MedChemComm* **7**, 1793–1796. (doi:10.1039/C6MD00342G)
60. van Haren M, van Ufford L, Moret EE, Martin NI. 2015 Synthesis and evaluation of protein arginine *N*-methyltransferase inhibitors designed to simultaneously occupy both substrate binding sites. *Org. Biomol. Chem.* **13**, 549–560. (doi:10.1039/C4OB01734J)
61. Sack JS *et al.* 2011 Structural basis for CARM1 inhibition by indole and pyrazole inhibitors. *Biochem. J.* **436**, 331–339. (doi:10.1042/BJ20102161)
62. Zhang J, Zheng YG. 2016 SAM/SAH analogs as versatile tools for SAM-dependent methyltransferases. *ACS Chem. Biol.* **11**, 583–597. (doi:10.1021/acscchembio.5b00812)

Supplementary Information

- NMR attribution page 1
- Complementary Material and Methods page 12
- Complementary Biological results page 17
- References page 23

NMR attribution

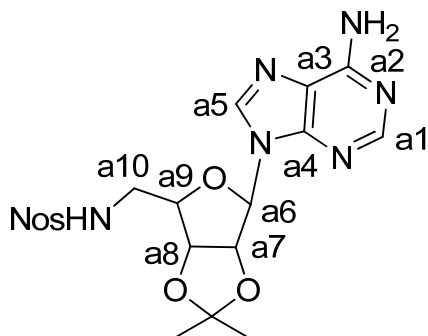
N1-(methoxymethyl)-N⁴-benzoyl-5-methyl-cytosine (9).



¹H NMR (500MHz ; CDCl₃) δ 13.21 (s, 1H, HNH) , 8.31 (m, 2H, HBz), 7.52 (m, 1H, HBz), 7.44 (m, 2H, HBz), 7.25 (q, 1H, *J*= 1.2 Hz, Hc3), 5.11 (s, 2H, Hc6), 3.40 (s, 3H, Hc7), 2.11 (d, 3H, *J*= 1.2 Hz, Hc1).

¹³C (125MHz, CDCl₃) δ 179.8 (CBz), 160.0 (Cc5), 148.9 (Cc4), 139.8 (Cc3), 137.1 (CBz), 132.6 (CBz), 130.0 (CBz), 128.2 (CBz), 112.6 (Cc2), 78.2 (Cc6), 57.1 (Cc7), 13.5 (Cc1).

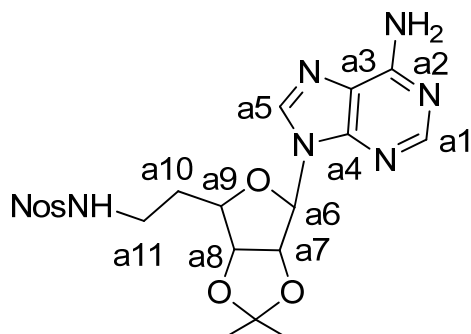
2',3'-O-isopropylidene- 5'-deoxy-5'-(2-nitrobenzenesulfonamido)adenosine (13)



¹H NMR (500MHz; DMSO-*d*6) δ 8.64 (brs, 1H, HNH), 8.29 (s, 1H, Ha1), 8.12 (s, 1H, Ha5), 7.92 (dd, *J*=1.2, 8.0Hz, 1H, HNos), 7.84 (dd, *J*=1.2, 8.0Hz, 1H, HNos), 7.79 (ddd, *J*=1.2, 6.4, 7.8Hz, 1H, HNos), 7.71 (ddd, *J*=1.2, 6.4, 7.7Hz, 1H, HNos), 7.41 (brs, 2H, HNH₂), 6.10 (d, *J*=2.9Hz, 1H, Ha6), 5.36 (dd, *J*=3.0, 6.3Hz, 1H, Ha7), 4.94 (dd, *J*=2.9, 6.3Hz, 1H, Ha8), 4.22 (ddd, *J*=2.9, 5.7, 11.5Hz, 1H, Ha9), 3.29 (dd, *J*=5.8, 13.6 Hz, 1H, Ha10A), 3.21 (dd, *J*=5.8, 13.6 Hz, 1H, Ha10B), 1.51 (s, 3H, HMe), 1.27 (s, 3H, HMe).

¹³C NMR (125MHz; DMSO-*d*6) δ 156.7 (Ca2), 153.0 (Ca1), 148.8 (Ca4), 148.0 (CNos), 140.6 (Ca5), 132.4 (CNos), 129.8 (CNos), 124.8 (CNos), 119.7 (Ca3), 113.9 (CiPr), 89.9 (Ca6), 85.7 (Ca9), 83.2 (Ca7), 82.0 (Ca8), 45.1 (Ca10), 27.4 (CMe), 25.5 (CMe).

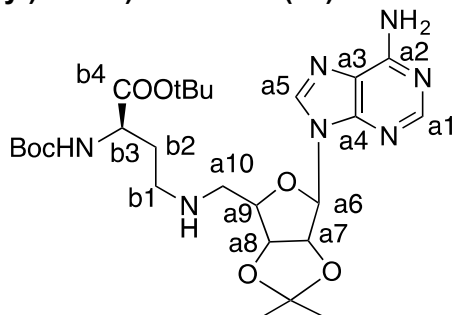
2',3'-O-isopropylidene- 5'-deoxy-5'-(2-nitrobenzenesulfonamidomethyl)adenosine (14)



¹H NMR (500MHz; CDCl₃) δ 8.29 (s, 1H, Ha1), 8.08-8.05 (m, 1H, HNos), 7.88 (s, 1H, Ha5), 7.82-7.77 (m, 1H, HNos), 7.72-6.9 (m, 2H, HNos), 6.01 (brt, *J*=5.4Hz, 1H, HNH2), 5.98 (d, *J*=2.4Hz, 1H, Ha6), 5.70 (brs, 1H, HNH), 5.46 (dd, *J*=2.6, 6.4Hz, 1H, Ha7), 5.00 (dd, *J*=4.5, 6.4Hz, 1H, Ha8), 4.24 (dt, *J*=4.5, 7.9Hz, 1H, Ha9), 3.31-3.18 (m, 1H, Ha11), 2.10-1.98 (m, 1H, Ha10), 1.61 (s, 3H, HMe), 1.38 (s, 3H, HMe).

¹³C NMR (125MHz; CDCl₃) δ 155.5 (Ca2), 153.2 (Ca1), 149.2 (Ca4), 147.8 (CNos), 140.2 (Ca5), 133.8 (CNos), 133.4 (CNos), 132.6 (CNos), 130.8 (CNos), 125.3 (CNos), 120.4 (Ca3), 114.9 (CiPr), 90.2 (Ca6), 84.8 (Ca9), 83.7 (Ca7), 83.5 (Ca8), 40.7 (Ca11), 45.1 (Ca10), 27.3 (CMe), 25.4 (CMe).

***tert*-butyl 2-((Boc)amino)-4-(((2',3'-*O*-isopropylidene-5'-deoxyadenosiny-5'-yl)amino)butanoate (15)**

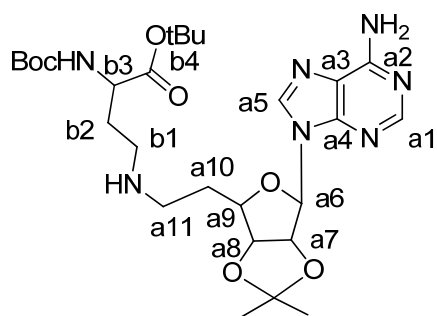


¹H NMR (500MHz; CDCl₃) δ 8.83 (s, 1H, Ha5), 8.20 (s, 1H, Ha1), 6.01 (d, *J*=3.3Hz, 1H, Ha6), 5.52 (m, 1H, Ha7), 5.11 (m, 1H, Ha8), 4.42 (m, 1H, Ha9), 4.29 (m, 1H, Ha9), 3.23 (s, 1H, Hb3), 3.01 (m, 1H, Ha10), 2.90 (m, 1H, Ha10), 2.80 (m, 1H, Hb1), 2.71 (m, 1H, Hb1), 2.01 (m, 1H, Hb2), 1.84 (m, 1H, Hb2), 1.65 (m, 3H, HiPr), 1.51-1.37 (m, 21H, HiPr and Ht-Bu and HBoc).

¹³C NMR (125MHz; CCl₃) δ 171.6 (Cb4), 155.7 (CBoc), 152.7 (Ca1), 151.3 (Ca2), 149.7 (Ca4), 142.3 (Ca5), 123.7 (Ca3), 114.9 (CiPr), 91.0 (Ca6), 85.4 (Ca9), 83.3 (Ca7), 82.0 (Ca8), 81.8 (CtBu), 79.6 (CBoc), 52.6 (Cb3), 51.1 (Ca10), 46.2 (Cb1), 32.2 (Cb2), 28.3 (CtBu), 28.04 (CBoc), 27.7 (Cb2), 27.2 (CiPr), 25.5 (CMe).

HRMS-ESI(m/z) calculated for C₂₆H₄₂N₇O₇ [M+H]⁺ : 564.2140; Found: 564.2153.

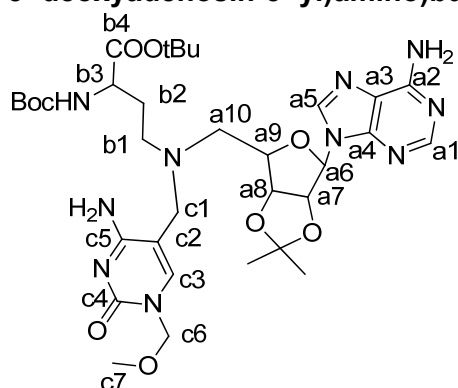
***tert*-butyl 2-((Boc)amino)-4-(((2',3'-*O*-isopropylidene-*N*⁴-benzoyl-5'-deoxyadenosiny-5'-yl) methylamino)butanoate (16)**



¹H NMR (500MHz; DMSO-*d*₆) δ 8.32 (s, 1H, Ha5), 8.17 (s, 1H, Ha1), 7.32 (brs, 2H, HNH2), 7.32 (bd, $J=7.6$ Hz, 2H, HNH), 6.09 (d, $J=2.7$ Hz, 1H, Ha6), 5.47 (dd, $J=2.7, 6.4$ Hz 1H, Ha7), 4.88 (dd, $J=3.7, 6.4$ Hz, 1H, Ha8), 4.19-4.15 (m, 1H, Ha9), 3.90-3.84 (m, 1H, Hb3), 2.49-2.42 (m, 4H, Hb1 and Ha11), 1.77-1.57 (m, 4H, Ha10 and Hb2), 1.58 (m, 3H, HiPr), 1.40-1.34 (m, 18H, Ht-Bu and HBoc), 1.32 (m, 3H, HiPr).

¹³C NMR (125MHz; CCl₃) δ 172.4 (Cb4), 156.7 (Ca2), 155.9 (CBoc), 153.2 (Ca1), 149.4 (Ca4), 140.3 (Ca5), 119.5 (Ca3), 113.9 (CiPr), 88.7 (Ca6), 84.3 (Ca9), 84.1 (Ca7), 83.5 (Ca8), 80.5 (CtBu), 78.4 (CBoc), 53.2 (Cb3), 46.2 (Ca11), 45.8 (Cb1), 33.8 (Ca10), 31.2 (Cb2), 28.6 (CBoc), 28.1 (CtBu), 27.5 (CiPr), 25.7 (CiPr).

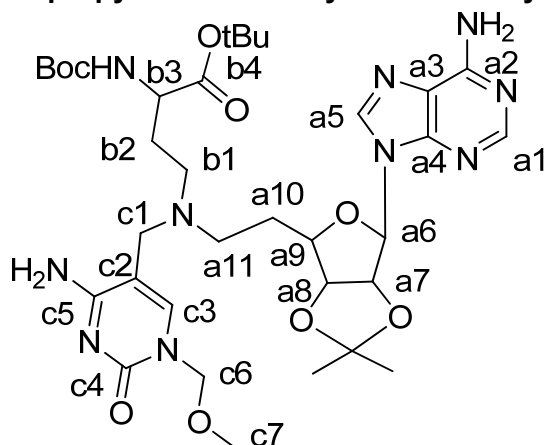
***tert*-butyl 2-((Boc)amino)-4-((1-(methoxymethyl)cytosin-5-yl)methyl)(2',3'-isopropylidene-5'-deoxyadenosin-5'-yl)amino)butanoate (17)**



¹H NMR (500MHz; DMSO-*d*₆) δ 8.29 (s, 1H, Ha5), 8.12 (s, 1H, Ha1), 7.51 (s, 1H, Hc3), 7.36 (m, 3H, HNH and HNH2), 7.10 (d, $J=8.0$ Hz, 1H, HNH), 6.73 (brs, 1H, HNH), 6.18 (d, $J=2.2$ Hz, 1H, Ha6), 5.46 (dd, $J=2.2, 6.3$ Hz, 1H, Ha7), 5.0 (d, $J=9.6$ Hz, 1H, Hc6), 4.98 (dd, $J=3.5, 6.3$ Hz, 1H, Ha8), 4.89 (d, $J=9.6$ Hz, 1H, Hc6), 4.25 (m, 1H, Ha9), 3.77 (m, 1H, Hb3), 3.31 (m, 2H, Hc1), 3.21 (s, 3H, Hc7), 3.15 (m, 1H, Hc1), 2.84 (m, 1H, Hb1), 2.55 (m, 2H, Hb1 and Ha10), 2.32 (m, 1H, Ha10), 1.71 (m, 2H, Hb2), 1.53 (s, 3H, HMe), 1.36 (s, 9H, Ht-Bu), 1.34 (s, 9H, Boc), 1.32 (s, 3H, HMe).

¹³C NMR (125MHz; DMSO-*d*₆) δ 172.2 (Cb4), 165.9 (Cc5), 156.6 (Ca2), 156.0 (Cc4), 155.9 (CBoc), 153.1 (Ca1), 149.0 (Ca4), 144.8 (Cc3), 140.6 (Ca5), 119.6 (Ca3), 114.0 (CiPr), 102.4 (Cc2), 89.0 (Ca6), 84.2 (Ca9), 83.7 (Ca7), 83.3 (Ca8), 80.7 (Ct-Bu), 78.8 (Cc6), 78.5 (CBoc), 56.3 (Cc7), 55.6 (Cb1), 53.4 (Cc1), 53.0 (Cb3), 50.1 (Ca10), 28.6 (Ct-Bu), 28.0 (CBoc), 27.7 (Cb2), 27.4 (CMe), 25.8 (CMe).

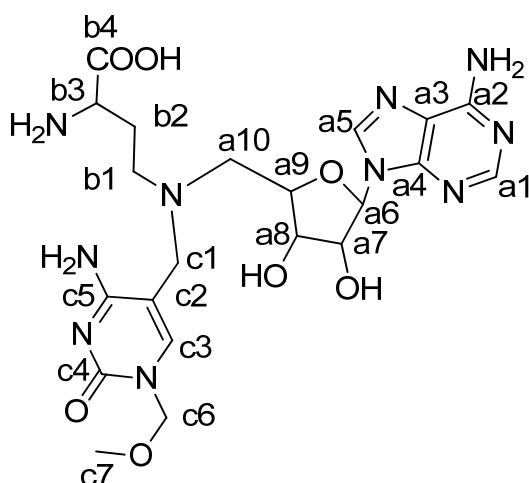
tert-butyl 2-((Boc)amino)-4-(((1-(methoxymethyl)-cytosin-5-yl)methyl)(5'-(2',3'-isopropylidene-5'-deoxy-adenosin-5'-yl)methyl)amino)butanoate (18).



¹H NMR (500MHz; DMSO-*d*6) δ 8.30 (s, 1H, Ha5), 8.16 (s, 1H, Ha1), 7.52 (s, 1H, Hc3), 7.32 (brs, 3H, HNH), 7.32 (bd, *J*=7.6Hz, 1H, HNH), 6.86-6.77 (m, 1H, HNH), 6.08 (d, *J*=2.6Hz, 1H, Ha6), 5.46 (dd, *J*=2.6, 6.3Hz 1H, Ha7), 4.97 (s, 2H, Hc6), 4.84 (dd, *J*=3.7, 6.3Hz, 1H, Ha8), 4.11-4.01 (m, 1H, Ha9), 3.83-3.75 (m, 1H, Hb3), 3.27-3.14 (m, 4H, Hc7 and Hc1), 2.48-2.32 (m, 4H, Hb1 and Ha11), 1.87-1.58 (m, 4H, Ha10 and Hb2), 1.52 (m, 3H, HiPr), 1.37-1.32 (m, 18H, Ht-Bu and HBoc), 1.31 (m, 3H, HiPr).

¹³C NMR (125MHz; CCl₃) δ 172.1 (Cb4), 166.1 (Cc5), 156.6 (Ca2), 156.0 (CBoc), 155.9 (Cc4), 153.2 (Ca1), 149.3 (Ca4), 144.3 (Cc3), 140.4 (Ca5), 119.6 (Ca3), 114.9 (CiPr), 102.7 (Cc2), 88.8 (Ca6), 84.5 (Ca9), 84.0 (Ca7), 83.5 (Ca8), 80.7 (CtBu), 78.8 (Cc6), 78.5 (CBoc), 56.2 (Cc7), 53.0 (Cb3), 52.9 (Cc1), 49.5 (Ca11), 48.9 (Cb1), 29.8 (Cb2), 28.6 (CBoc), 28.0 (CtBu), 27.9 (Ca10), 27.5 (CiPr), 25.7 (CiPr).

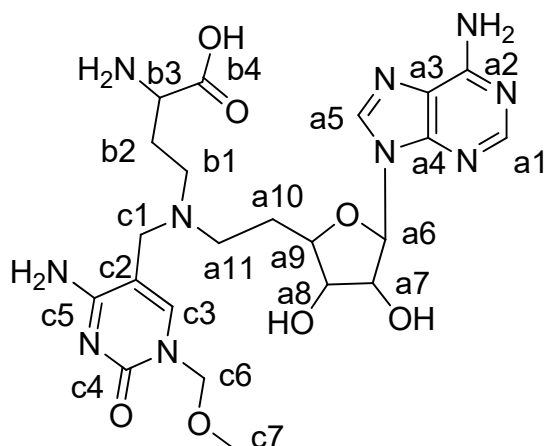
2-amino-4-(((1-(methoxymethyl)cytosin-5-yl)methyl)((5'-deoxy-5'-(adenosin-5'-yl)amino)butanoic acid (1)



¹H NMR (500MHz; D₂O) δ 8.10 (s, 1H, Ha5), 8.08 (s, 1H, Ha1), 7.51 (s, 1H, Hc3), 5.90 (d, *J*=2.1Hz, 1H, Ha6), 4.84 (s, 1H, Hc6), 4.79 (s, 1H, Hc6), 4.67 (m, 1H, Ha7), 4.25-4.17 (m, 2H, Ha8 and Ha9), 3.66-3.55 (m, 2H, Hb3 and Ha10), 3.47 (dd, *J*=13.6Hz, 1H, Ha10), 3.15 (s, 3H, Hc7), 2.93 (m, 2H, Hb1), 2.78 (m, 2H, Hc1), 2.00 (m, 1H, Hb2), 1.92 (m, 1H, Hb2).

¹³C NMR (125MHz; D₂O) δ 173.8 (Cb4), 164.1 (Cc5), 154.7 (Ca2), 156.0 (Cc4), 151.6 (Ca1), 148.6 (Ca4), 146.3 (Cc3), 140.7 (Ca5), 118.9 (Ca3), 115.1 (CiPr), 117.4 (Cc2), 88.6 (Ca6), 80.9 (Ca9), 72.8 (Ca7), 71.9 (Ca8), 79.6 (Cc6), 56.1 (Cc7), 54.7 (Cb1), 53.4 (Cb3), 52.4 (Ca10), 50.7 (Cc1), 26.7 (Cb2).

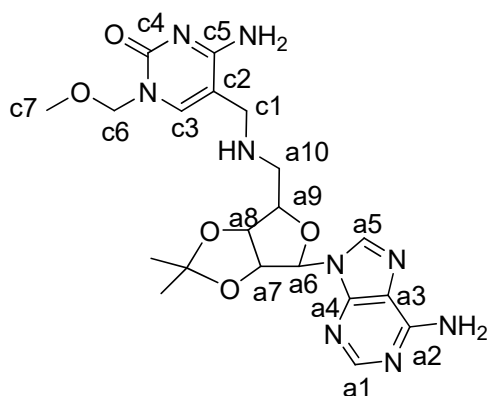
2-amino- 4-(((1-(methoxymethyl)-cytosin-5-yl)methyl)(5'-(2',3'-isopropylidene-5'-deoxyadenosin-5'-yl)methyl)amino)butanoic acid (2).



¹H NMR (500MHz; DMSO-*d*6) δ 9.05 (brs, 1H, HNH), 8.42 (s, 1H, Ha5), 8.28 (s, 1H, Ha1), 8.02 (s, 1H, Hc3), 7.98 (brs, 2H, HNH), 8.89 (d, *J*=5.3Hz, 1H, Ha6), 5.06 (s, 2H, Hc6), 4.67 (t, *J*=5.3Hz 1H, Ha7), 4.10 (t, *J*=4.8Hz, 1H, Ha8), 3.98-3.88 (m, 2H, Ha9 and Hb3), 3.83-3.65 (m, 2H, Hc1), 3.28 (m, 4H, Hc7), 2.98-2.78 (m, 4H, Hb1 and Ha11), 2.11-1.89 (m, 4H, Ha10 and Hb2).

¹³C NMR (125MHz; CCl₃) δ 171.1 (Cb4), 161.5 (Cc5), 154.7 (Ca2), 150.8 (Ca1), 149.4 (Ca4), 141.3 (Ca5), 119.6 (Ca3), 88.8 (Ca6), 82.4 (Ca9), 79.3 (Cc6), 73.8 (Ca8), 73.4 (Ca7), 56.9 (Cc7), 50.7 (Cb3), 52.9 (Cc1), 49.5 (Ca11), 48.9 (Cb1), 29.8 (Cb2), 25.6 (Ca10).

1-(methoxymethyl)-5-((2',3'-*O*-isopropylidene-5'-deoxyadenosin-5'-yl)aminomethyl)cytosine (19)

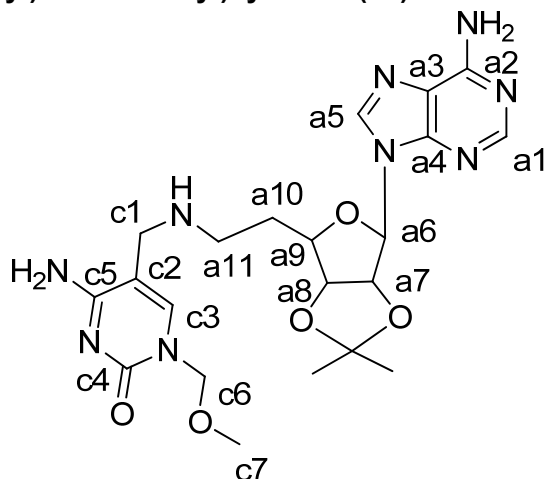


¹H NMR (500MHz; DMSO-*d*6+D₂O) δ 8.33 (s, 1H, Ha5), 8.08 (s, 1H, Ha1), 7.49 (s, 1H, Hc3), 6.06 (d, *J*=3.3Hz, 1H, Ha6), 5.44 (dd, *J*=3.3, 6.3Hz, 1H, Ha9), 4.97 (m, 3H, Ha7 and Hc7), 4.23 (m, 1H, Ha8), 3.40 (d, *J*=6.1Hz, 2H, Hc1), 3.20 (s, 3H, Hc7), 2.69 (m, 2H, Ha10), 1.54 (s, 3H, HiPr), 1.33 (s, 3H, HiPr).

¹³C NMR (125MHz; DMSO-*d*6+ D₂O) δ 166.0 (Cc5), 156.4 (Ca2), 155.5 (Cc4), 153.1 (Ca1), 149.2 (Ca4), 143.9 (Cc3), 140.6 (Ca5), 119.4 (Ca3), 113.9 (CiPr), 104.1 (Cc2), 89.5 (Ca6), 83.0

(Ca9), 84.5 (Ca8), 82.6 (Ca7), 78.8 (Hc6), 56.3 (Cc7), 50.2 (Ca10), 46.9 (Cc1), 27.4 (CiPr), 25.7 (CiPr).

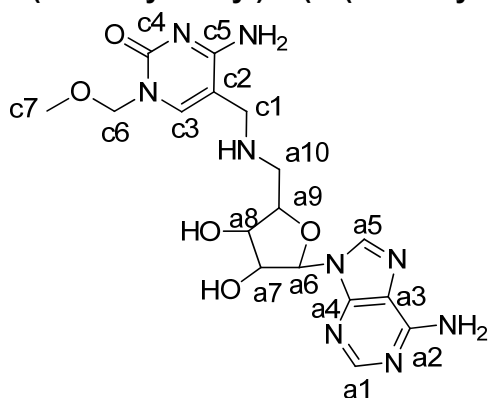
1-(methoxymethyl)-5-((2',3'-O-isopropylidene-5'-deoxyadenosin-5'-yl)aminomethyl)cytosine (20).



¹H NMR (500MHz; DMSO-*d*6) δ 8.32 (s, 1H, Ha5), 8.17 (s, 1H, Ha1), 7.47 (s, 1H, Hc3), 7.33 (brs, 2H, HNH2a), 7.13 (brs, 1H, HNH), 6.09 (d, *J*=2.7Hz, 1H, Ha6), 5.37 (dd, *J*=2.7, 6.4Hz, 1H, Ha7), 4.97 (s, 2H, Hc6), 4.89(dd, *J*=3.8, 6.4Hz, 1H, Ha8), 4.18-4.15 (m, 1H, Ha9), 3.34 (s, 2H, Hc1), 3.21 (s, 3H, Hc7), 2.48-2.47 (m, 2H, Ha11), 1.80-1.71 (m, 1H, Ha10), 1.53 (s, 3H, Hp), 1.33 (s, 3H, HiPr).

¹³C NMR (125MHz; DMSO-*d*6) δ 166.2 (Cc5), 156.6 (Ca2), 156.1 (Cc4), 153.2 (Ca1), 149.4 (Ca4), 143.4 (Cc3), 140.4 (Ca5), 119.5 (Ca3), 114.0 (CiPr), 104.1 (Cc2), 88.8 (Ca6), 84.5 (Ca8), 84.1 (Ca7), 83.5 (Ca9), 78.8 (Cc6), 56.2 (Cc7), 47.2 (Cc1), 45.1 (Ca11), 33.5 (Ca10), 27.5 (CiPr), 25.7 (CiPr).

1-(methoxymethyl)-5-(*N*-(5'-deoxyadenosin-5'-yl)aminomethyl)cytosine (3)

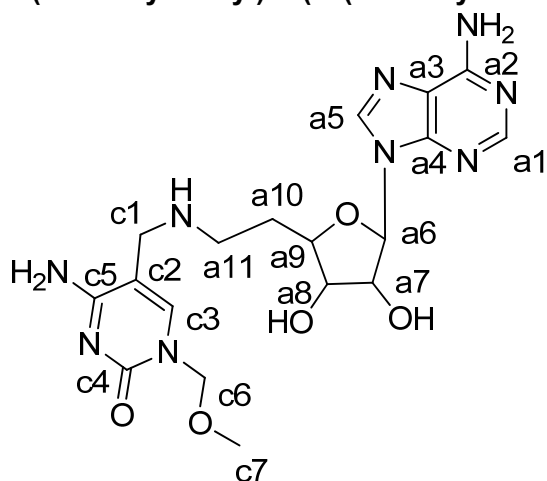


¹H NMR (500MHz; D₂O) δ 8.26 (s, 1H, Ha1), 8.12 (s, 1H, Ha5), 7.83 (s, 1H, Hc3), 6.10 (d, *J*=4.6Hz, 1H, Ha6), 5.07 and 5.01 (d, *J*=10.6Hz, 2H, Hc6), 4.83 (t, *J*=4.5Hz, 1H, Ha7), 4.53 (t, *J*=5.2Hz, 1H, Ha8), 4.49 (q, *J*=5.8Hz, 1H, Ha9), 4.17 and 4.12 (d, *J*=14.6Hz, 2H, Ha10), 3.60 (d, *J*=6.0Hz, 2H, Hc1), 3.30 (s, 3H, Hc7).

¹³C NMR (125MHz; D₂O) δ 164.8 (Cc5), 157.1 (Cc4), 155.5 (Ca2), 152.3 (Ca1), 149.2 (Cc3), 148.2 (Ca4), 140.7 (Ca5), 119.0 (Ca3), 98.6 (Cc2), 89.3 (Ca6), 79.8 (Cc6), 79.3 (Ca9), 72.8 (Ca7), 71.3 (Ca8), 56.1 (Cc7), 48.0 (Cc1), 44.0 (Ca10).

HRMS-ESI(*m/z*) calculated for C₁₇H₂₃N₉NaO₅ [M+Na]⁺: 456.1714 ; Found: 456.1699.

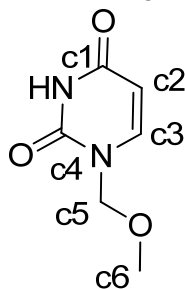
1-(methoxymethyl)-5-(*N*-(5'-deoxyadenosin-5'-yl)aminomethyl)cytosine (4)



¹H NMR (500MHz; DMSO-*d*6) δ 8.57 (brs, 2H, HNH₂), 8.40 (s, 1H, Ha₅), 8.24 (s, 1H, Ha₁), 7.98 (s, 1H, Hc₃), 7.72 (brs, 1H, HNH), 5.90 (d, *J*=4.2Hz, 1H, Ha₆), 5.06 (s, 2H, Hc₆), 4.62 (t, *J*=4.7Hz, 1H, Ha₇), 4.20 (t, *J*=5.4Hz, 1H, Ha₈), 3.98-3.91 (m, 2H, Ha₉ and Hc₁), 3.29 (s, 3H, Hc₇), 3.09-3.01 (m, 2H, Ha₁₁), 2.08-1.97 (m, 1H, Ha₁₀).

¹³C NMR (125MHz; DMSO-*d*6) δ 162.6 (Cc₅), 155.4 (Ca₂), 151.7 (Ca₁), 150.6 (Cc₃), 149.4 (Ca₄), 140.4 (Ca₅), 119.6 (Ca₃), 97.3 (Cc₂), 88.8 (Ca₆), 81.0 (Ca₉), 79.3 (Cc₆), 73.6 (Ca₇), 73.5 (Ca₈), 56.9 (Cc₇), 44.2 (Ca₁₁), 43.2 (Cc₁), 29.9 (Ca₁₀).

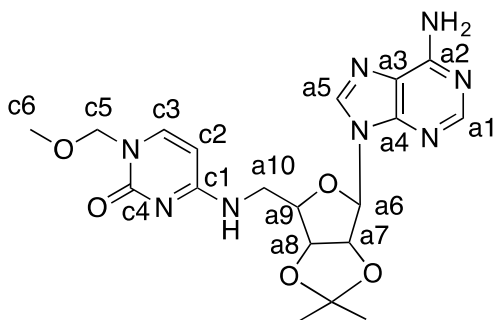
1-(methoxymethyl)uracil (22).



¹H NMR (500MHz ; DMSO-*d*6) δ 11.32 (s, 1H, HNH) , 7.70 (d, *J*=8.0Hz, 1H, Hc₃), 5.61 (d, *J*=8.0Hz, 1H, Hc₂), 5.02 (s, 2H, Hc₅), 3.27 (s, 3H, Hc₆).

¹³C (125MHz, DMSO-*d*6) δ 164.0 (Cc₁), 151.5 (Cc₄), 145.4 (Cc₃), 101.9 (Cc₂), 78.0 (Cc₅), 56.4 (C₆).

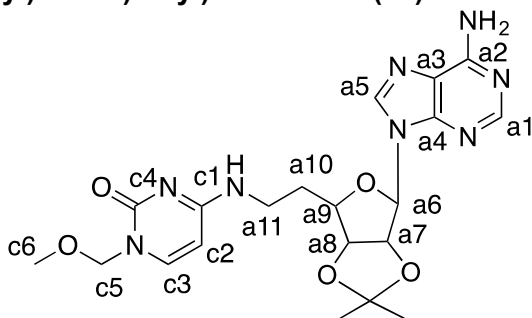
2',3'-*O*-isopropylidene-5'-deoxy-5'-(((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)methyl)adenosine. (23)



¹H NMR (500MHz; DMSO-*d*₆) δ 8.35 (s, 1H, Ha1), 8.21 (s, 1H, Ha5), 7.95 (t, $J=6.1$ Hz, 1H, HNH), 7.56 (d, $J=7.2$ Hz, 1H, Hc3), 7.36 (brs, 2H, HNH), 6.18 (d, $J=2.7$ Hz, 1H, Ha6), 5.73 (d, $J=7.2$ Hz, 1H, Hc4), 5.47 (dd, $J=2.8, 6.3$ Hz, 1H, Ha7), 5.02-4.95 (m, 3H, Ha8 and Hc5), 4.34-4.29 (m, 1H, Ha9), 3.68-3.26 (m, 2H, Ha10), 3.24 (s, 3H, Hc8), 1.54 (s, 3H, HiPr), 1.33 (s, 3H, HiPr).

¹³C NMR (125MHz; DMSO-*d*₆) δ 164.5 (Cc1), 156.6 (Cc4), 156.0 (Ca2), 153.2 (Ca1), 149.2 (Ca4), 145.1 (Cc3), 140.5 (Ca5), 119.7 (Ca3), 114.0 (CiPr), 95.0 (Cc2), 89.5 (Ca6), 84.3 (Ca7), 83.5 (Ca8), 82.3 (Ca9), 78.9 (Cc5), 56.3 (Cc6), 42.4 (Ca10), 27.5 (CiPr), 25.7 (CiPr).

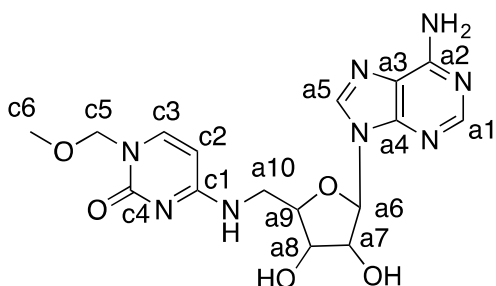
2',3'-O-isopropylidene-5'-deoxy-5'-((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)ethanol (24)



¹H NMR (500MHz; DMSO-*d*₆) δ 8.34 (s, 1H, Ha1), 8.17 (s, 1H, Ha5), 7.77 (t, $J=5.4$ Hz, 1H, HNH), 7.52 (d, $J=7.3$ Hz, 1H, Hc3), 7.33 (brs, 2H, HNH), 6.13 (d, $J=2.6$ Hz, 1H, Ha6), 5.67 (d, $J=7.3$ Hz, 1H, Hc2), 5.49 (dd, $J=2.7, 6.3$ Hz, 1H, Ha7), 4.98 (s, 2H, Hc5), 4.94 (dd, $J=3.5, 6.3$ Hz, 1H, Ha8), 4.19-4.14 (m, 1H, Ha9), 3.27-3.20 (m, 5H, Ha11 and Hc6), 1.98-1.76 (m, 2H, Hc10), 1.54 (s, 3H, HiPr), 1.33 (s, 3H, HiPr).

¹³C NMR (125MHz; DMSO-*d*₆) δ 164.3 (Cc1), 156.6 (Cc4), 156.1 (Ca2), 153.2 (Ca1), 149.1 (Ca4), 144.7 (Cc3), 140.4 (Ca5), 119.5 (Ca3), 114.0 (CiPr), 95.0 (Cc2), 88.8 (Ca6), 83.9 (Ca7), 83.8 (Ca8), 83.5 (Ca9), 78.9 (Cc5), 56.3 (Cc6), 37.0 (Ca11), 32.8 (Ca10), 27.5 (CiPr), 25.7 (CiPr).

5'-deoxy-5'-((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)adenosine (5)

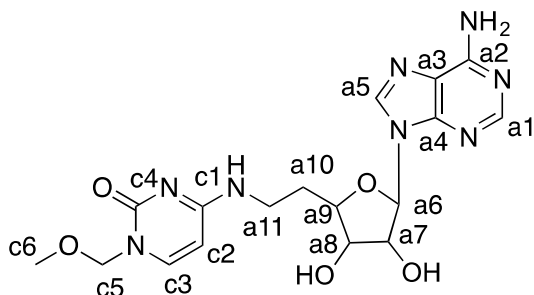


¹H NMR (500MHz; DMSO-*d*₆) δ 8.36 (s, 1H, Ha1), 8.20 (s, 1H, Ha5), 8.09 (t, $J=5.8$ Hz, 1H, HNH), 7.58 (d, $J=7.2$ Hz, 1H, Hc3), 7.34 (brs, 2H, HNH), 5.88 (d, $J=6.2$ Hz, 1H, Ha6), 5.78 (d,

$J=7.3\text{Hz}$, 1H, Hc2), 5.47 (d, $J=6.3\text{Hz}$, 1H, HOH), 5.33 (d, $J=4.8\text{Hz}$, 1H, HOH), 5.00 (d, $J=4.2\text{Hz}$, 2H, Hc5), 4.69 (ddd, $J=5.4$, 6.1, 11.5Hz, 1H, Ha7), 4.15-4.07 (m, 2H, Ha8 and Ha9), 3.65 (t, $J=5.7\text{Hz}$, 2H, Ha10), 3.24 (s, 3H, Hc6).

^{13}C NMR (125MHz; DMSO-*d*6) δ 164.5 (Cc1), 156.6 (Cc4), 156.1 (Ca2), 153.0 (Ca1), 149.1 (Ca4), 145.1 (Cc3), 140.7 (Ca5), 119.9 (Ca3), 95.0 (Cc2), 88.2 (Ca6), 83.4 (Ca9), 78.9 (Cc5), 73.1 (Ca7), 71.8 (Ca8), 56.3 (Cc6), 42.6 (Ca10).

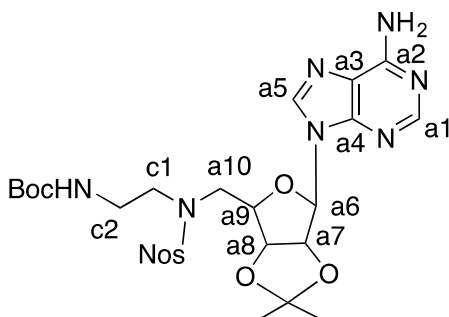
5'-deoxy-5'-((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)methyl)adenosine. (6).



^1H NMR (500MHz; DMSO-*d*6) δ 8.35 (s, 1H, Ha1), 8.16 (s, 1H, Ha5), 7.82 (t, $J=5.4\text{Hz}$, 1H, HNH), 7.53 (d, $J=7.4\text{Hz}$, 1H, Hc3), 7.28 (brs, 2H, HNH), 5.88 (d, $J=5.4\text{Hz}$, 1H, Ha6), 5.69 (d, $J=7.1\text{Hz}$, 1H, Hc2), 5.51 (brs, 1H, HOH), 5.29 (brs, 1H, HOH), 4.99 (s, 2H, Hc5), 4.69-4.63 (m, 1H, Ha7), 4.14-4.06 (m, 1H, Ha8), 3.97-3.91 (m, 1H, Ha9), 3.43-3.35 (m, 1H, Ha11), 3.30-3.20 (m, 4H, Ha11 and Hc6), 2.04-1.81 (m, 2H, Ha10).

^{13}C NMR (125MHz; DMSO-*d*6) δ 164.3 (Cc1), 156.5 (Cc4), 156.2 (Ca2), 153.1 (Ca1), 149.8 (Ca4), 144.6 (Cc3), 140.4 (Ca5), 119.6 (Ca3), 95.1 (Cc2), 88.0 (Ca6), 81.9 (Ca7), 78.9 (Cc5), 73.8 (Ca8), 73.4 (Ca9), 56.3 (Cc6), 37.1 (Ca11), 33.0 (Ca10).

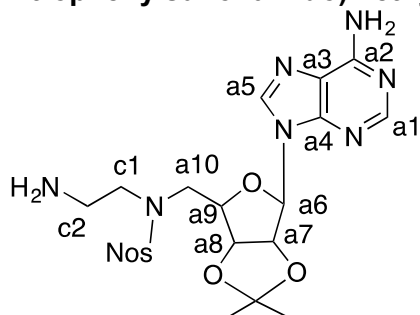
2',3'-O-isopropylidene-5'-deoxy-5'-(N-((2-(Boc-amino)ethyl)-2-nitrophenylsulfonamido)methyl)adenosine (25)



^1H NMR (500MHz; CDCl_3) δ 8.34 (s, 1H, Ha1), 7.87-7.83 (m, 2H, HNos and Ha5), 7.59-7.51 (m, 2H, HNos), 7.50-7.43 (m, 1H, HNos), 6.01 (d, $J=2.0\text{Hz}$, 1H, Ha6), 5.82 (brs, 2H, HNH2), 5.39 (dd, $J=1.4$, 6.4Hz, 1H, Ha7), 5.10 (dd, $J=4.0$, 6.4Hz, 1H, Ha8), 4.91 (m, 1H, HNH), 4.33 (dt, $J=4.0$, 8.3Hz, 1H, Ha9), 3.86-3.79 and 3.66-3.58 (m, 2H, Hc1), 3.57-3.47 (m, 1H, Ha10), 3.34-3.11 (m, 3H, Ha10 and Hc2), 1.60 (s, 3H, HiPr), 1.42 (HBoc), 1.38 (s, 3H, HiPr).

^{13}C NMR (125MHz; CDCl_3) δ 155.8 (CBoc), 155.5 (Ca2), 153.2 (Ca1), 149.0 (Ca4), 147.8 (CNos), 140.1 (Ca5), 133.4 (CNos), 132.9 (CNos), 131.5 (CNos), 131.1 (CNos), 124.1 (CNos), 120.3 (Ca3), 114.8 (CiPr), 90.4 (Ca6), 85.1 (Ca7), 84.0 (Ca8), 82.4 (Ca9), 79.4 (CBoc), 49.8 (Cc1), 48.8 (Ca10), 38.4 (Cc2), 28.3 (CBoc), 27.1 (CiPr), 25.3 (CiPr)

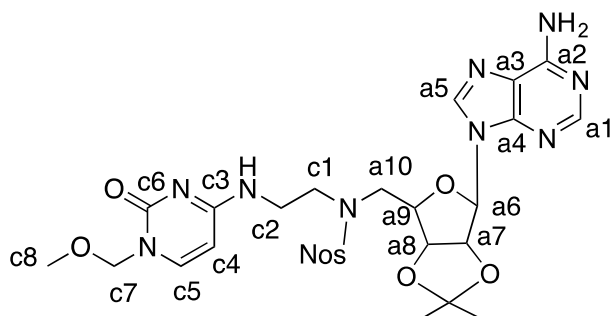
2',3'-O-isopropylidene-5'-deoxy-5'-(N-((2-aminoethyl)-2-nitrophenylsulfonamido)methyl)adenosine (26)



¹H NMR (500MHz; CDCl₃) δ 8.30 (s, 1H, Ha1), 7.98-7.91 (m, 2H, HNos and Ha5), 7.69-7.57 (m, 3H, HNos), 6.08 (d, *J*=1.9Hz, 1H, Ha6), 6.00 (brs, 1H, HNH), 5.39 (dd, *J*=1.9, 6.5Hz, 1H, Ha7), 5.08 (dd, *J*=4.1, 6.2Hz, 1H, Ha8), 4.48 (dt, *J*=3.9, 9.3Hz, 1H, Ha9), 3.96-3.88 (m, 1H, Hc1), 3.66-3.54 (m, 2H, Ha10 and Hc1), 3.33-3.22 (m, 1H, Ha10), 2.94-2.81 (m, 2H, Hc2), 1.60 (s, 3H, HiPr), 1.39 (s, 3H, HiPr).

¹³C NMR (125MHz; CDCl₃) δ 155.6 (Ca2), 153.1 (Ca1), 148.8 (Ca4), 148.0 (CNos), 140.4 (Ca5), 133.9 (CNos), 131.9 (CNos), 131.2 (CNos), 124.3 (CNos), 120.2 (Ca3), 115.0 (CiPr), 90.4 (Ca6), 85.9 (Ca7), 83.9 (Ca8), 82.3 (Ca9), 53.4 (Cc1), 48.8 (Ca10), 38.8 (Cc2), 27.0 (CiPr), 25.4 (CiPr)

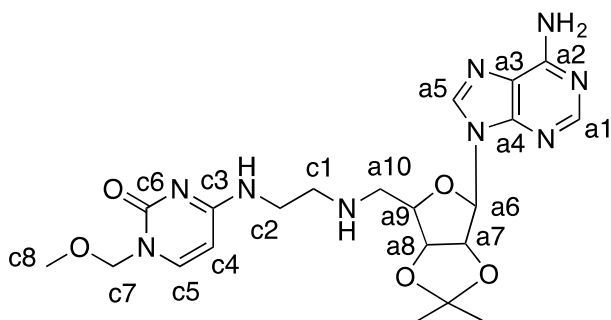
2',3'-O-isopropylidene-5'-deoxy-5'-(((N-(2-((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)ethyl)-2-nitrophenyl)sulfonamido)methyl)adenosine (27).



¹H NMR (500MHz; DMSO-*d*₆) δ 8.34 (s, 1H, Ha1), 8.20 (s, 1H, Ha5), 7.99-7.81 (m, 3H, HNos and HNH), 7.76 (ddd, *J*=1.2, 7.7, 15.4Hz, 1H, HNos), 7.64 (ddd, *J*=1.2, 7.7, 15.4Hz, 1H, HNos), 7.61-7.50 (m, 2H, Hc5 and HNH), 6.19 (d, *J*=2.2Hz, 1H, Ha6), 5.39 (dd, *J*=2.2, 6.4Hz, 1H, Ha7), 5.06 (dd, *J*=3.7, 6.4Hz, 1H, Ha8), 5.01 (s, 2H, Hc7), 4.33 (m, 1H, Ha9), 3.82-60 (m, 2H, Ha10), 3.47-3.41 (m, 2H, Hc1), 3.33-3.26 (m, 1H, Ha10), 3.24 (s, 3H, Hc8), 3.14-3.04 (m, 2H, Hc2), 1.52 (s, 3H, HiPr), 1.31 (s, 3H, HiPr).

¹³C NMR (125MHz; DMSO-*d*₆) δ 163.9 (Cc3), 156.0 (Cc6), 155.4 (Ca2), 152.5 (Ca1), 148.9 (Ca4), 147.8 (CNos), 145.1 (Cc5), 140.8 (Ca5), 132.7 (CNos), 131.9 (CNos), 130.4 (CNos), 124.7 (CNos), 129.52 (Ca3), 114.1 (CiPr), 95.1 (Cc4), 89.2 (Ca6), 84.5 (Ca7), 83.7 (Ca8), 82.3 (Ca9), 79.9 (Cc7), 56.3 (Cc8), 50.0 (Ca10), 46.9 (Cc1), 38.7 (Cc2), 27.4 (CiPr), 25.6 (CiPr).

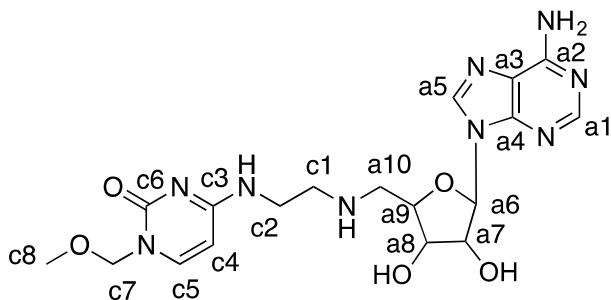
2',3'-O-isopropylidene-5'-deoxy-5'-(((N-(2-((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)ethyl)-amino)methyl)adenosine (28).



¹H NMR (500MHz; DMSO-*d*6) δ 8.38 (s, 1H, Ha1), 8.18 (s, 1H, Ha5), 8.01 (brs, 1H, HNH), 7.60 (d, *J*=7.2Hz, 1H, Hc5), 6.25 (d, *J*=2.5Hz, 1H, Ha6), 5.73 (d, *J*=7.3Hz, 1H, Hc4), 5.41 (dd, *J*=2.5, 6.2Hz, 1H, Ha7), 5.09 (dd, *J*=3.4, 6.2Hz, 1H, Ha8), 5.01 (s, 2H, Hc7), 4.51-4.42 (m, 1H, Ha9), 4.15-4.04 (m, 2H, HNH), 3.50-3.41 (m, 2H, Hc2), 3.33-3.22 (m, 2H, Ha10), 3.24 (s, 3H, Hc8), 3.03-2.93 (m, 2H, Hc1), 1.57 (s, 3H, HiPr), 1.34 (s, 3H, HiPr).

¹³C NMR (125MHz; DMSO-*d*6) δ 164.7 (Cc3), 156.6 (Cc6), 155.8 (Ca2), 153.2 (Ca1), 149.1 (Ca4), 145.2 (Cc5), 140.6 (Ca5), 119.7 (Ca3), 114.1 (CiPr), 95.0 (Cc4), 89.6 (Ca6), 83.6 (Ca7), 82.7 (Ca8), 82.2 (Ca9), 78.9 (Cc7), 56.3 (Cc8), 49.5 (Ca10), 47.7 (Cc1), 37.7 (Cc2), 27.4 (CiPr), 25.6 (CiPr).

5'-deoxy-5'-((*N*-(2-((1-(methoxymethyl)-2-oxo-1,2-dihydropyrimidin-4-yl)amino)ethyl)amino)methyl)adenosine (7)



¹H NMR (500MHz; DMSO-*d*6) δ 8.97 (brs, 2H, HNH), 8.49 (s, 1H, Ha1), 8.32 (s, 1H, Ha5), 8.04 (brs, 1H, HNH), 7.65 (d, *J*=7.4Hz, 1H, Hc5), 5.99 (d, *J*=5.5Hz, 1H, Ha6), 5.77 (d, *J*=7.4Hz, 1H, Hc4), , 5.01 (s, 2H, Hc7), 4.71 (t, *J*=5.1Hz, 1H, Ha7), 4.32-4.21 (m, 2H, Ha8 and Ha9), 3.67-3.36 (m, 4H, Hc2 and Ha10), 3.24 (s, 3H, Hc8), 3.20-2.09 (m, 2H, Hc1).

¹³C NMR (125MHz; DMSO-*d*6) δ 164.3 (Cc3), 155.2 (Cc6), 154.8 (Ca2), 150.7 (Ca1), 149.2 (Ca4), 145.6 (Cc5), 141.7 (Ca5), 119.8 (Ca3), 95.1 (Cc4), 89.1 (Ca6), 80.2 (Ca9), 79.0 (Cc7), 73.2 (Ca7), 72.0 (Ca8), 56.4 (Cc8), 49.3 (Ca10), 47.7 (Cc1), 37.27 (Cc2).

Supplementary Materials and Methods

DNMT inhibition assays

DNMT3A enzyme inhibition was adapted from the restriction-based fluorescence assay protocol described by [1]. Briefly, a 5'-labelled biotin oligonucleotide hybridized to its complementary strand labelled with 6-carboxyfluorescein at the 3'-end was transferred into a 384-well microplate (black Optiplates; PerkinElmer) pre-coated with avidin. The duplex contains a unique CpG site overlapping with a restriction site of a methylation-sensitive restriction enzyme. The human C-terminal DNMT3A (a.a. 623-908) was added to each well (200 ng/well) and mixed with the chemical compounds at the desired concentration and freshly prepared AdoMet (20 μ M final concentration) to start the reaction in a total volume of 50 μ L. After incubation at 37°C (1h), each well was washed and specific fluorescence signals were detected with the methylation-sensitive restriction enzyme HpyCH4IV (New England Biolabs, Ipswich, MA, USA) on a PerkinElmer Envision detector.

His-DNMT1 (182 kDa, human) inhibition assays[2] were performed in a total reaction volume of 10 μ L in low-volume nonbinding surface (NBS) 384-well microplates (Corning Inc.), containing test compound (up to 1% **DMSO-d6**), 1 μ M of a S-adenosyl-L-methionine (AdoMet)/[methyl-³H]AdoMet (3 TBqmmol⁻¹, PerkinElmer) mix in a ratio of 3:1 (isotopic dilution 1*:3), 0.3 μ M of biotinylated hemimethylated DNA duplex and 90 nM of DNMT1 in methylation buffer (20 mM HEPES (pH 7.2), 1 mM EDTA, 50 mM KCl, 25 μ g/mL of bovine serum albumin). The reaction was incubated at 37°C for 2 h, then an aliquot (8 μ L) was transferred into a streptavidin 96- well scintillant-coated FlashPlate (PerkinElmer) containing 20 μ M S-adenosyl-L-homocysteine (AdoHcy; 190 μ L) in 50 mM Tris-HCl (pH 7.4). The FlashPlate was agitated at RT for 1 h, washed three times with 200 μ L of 0.05% Tween-20 in 50 mM Tris-HCl (pH 7.4), and read in 200 μ L of 50mM Tris-HCl (pH 7.4) on TopCount NXT (PerkinElmer).

RNA MTase inhibition assays

The **RNA MTase** filter binding assay (FBA) is based on the incubation of the RNA MTase at the concentration indicated in the table below in the presence of the RNA substrate and the radiolabeled tritiated SAM ([³H]SAM) in Tris-HCl pH 8 40mM, DTT 1 mM, SAM 10 μ M. Reaction mixtures were incubated for 20 min. at 30°C and stopped by adding a 10-fold dilution of the reaction mixture in 20 μ M ice-cold SAH (AdoHcy, New England Biolabs). Samples were then transferred to diethylaminoethylcellulose (DEAE) filters (Perkin Elmer) by using a Filtermat Harvester apparatus (Packard Instruments). Unincorporated [³H]SAM was removed from the DEAE filters by several washes with 0.01 M ammonium formate, pH 8.0, H₂O, and absolute ethanol. After drying, filters were incubated with BetaplateScint (Wallac) scintillation fluid before quantification of the ³H methylation of the RNA substrates using a Wallac 1450 MicroBetaTriLux liquid scintillation counter

<u>MTase</u>	<u>MgCl₂</u>	<u>RNA Substrate</u>	<u>Enzyme</u>
<u>WNV-MTase</u>	=	<u>GpppACCCC</u>	<u>1 μM</u>
<u>MERS 2'O-MTase</u>	<u>1mM</u>	<u>^mGpppACCCC</u>	<u>1 μM</u>
<u>MERS N7-MTase</u>	=	<u>GpppACCCC</u>	<u>0.25μM</u>
<u>SARS 2'O -MTase</u>	<u>1mM</u>	<u>^mGpppACCCC</u>	<u>1 μM</u>
<u>SARS N7-MTase</u>	=	<u>GpppACCCC</u>	<u>0.1 μM</u>
<u>hRNMT</u>	=	<u>GpppACCCC</u>	<u>0.25μM</u>
<u>DV-MTase</u>	=	<u>GpppACCCC</u>	<u>1 μM</u>

WNV-MTase: MTase domain from the West Nile virus (nsp5 protein); MERS 2'O-MTase: MTase from the Middle East Respiratory Syndrome Coronavirus (MERS-CoV) (nsp10/nsp16 complex); MERS-N7-MTase: MTase from the MERS-CoV (NSP14 protein); SARS 2'O-MTase: MTase from the Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) (nsp10/nsp16 complex); SARS N7-MTase: MTase from the SARS-CoV (nsp14) ; hRNMT: human RNA N7-methyltransferase; DV MTase: MTase domain of the Dengue virus NS5 protein.

PRMT inhibition assays

Enzyme		Experimental conditions		Readout	
hPRMT1	human recombinant (E. coli)	Histone H4 full length _{SEP} (25 nM) _{SEP} [³ H]SAM (60 nM)	90 min RT	Methylated Histone H4 full length	Scintillation counting
hPRMT4	human recombinant (Sf9 cells)	histone H3 full length _{SEP} (25 nM) _{SEP} [³ H]SAM (60 nM)	60 min RT	Methylated histone H3 full length	Scintillation counting
hPRMT6	human recombinant (E. coli)	Histone H3 full length _{SEP} (25 nM) _{SEP}	120 min RT	Methylated histone H3 full length	Scintillation counting
hPRMT7	human recombinant (Sf9 cells)	Histone H2B (21-41),biotin labeled (4 nM) [³ H]SAM (250 nM)	90 min RT	Methylated Histone H2B (21-41)	Scintillation counting

Crystallography studies

PRMT4 cloning, expression and purification

The *mus musculus Prmt4* gene sequence corresponding to the PRMT core (residues 130 to 487) were amplified by PCR from the original GST-PRMT4 construct [3]. The sequences were cloned in the pDONR207™ (Invitrogen) vector using a BP reaction (Gateway® Cloning, Life Technologies). The positive clones were confirmed by sequencing (GATC). The sequences

were subcloned in a pDEST20™ vector using a LR reaction. The resulting recombinant protein harbor an amino-terminal glutathione S-transferase (GST) tag followed by a Tobacco etch virus (TEV) protease cleavage site. DH10Bac competent cells containing the baculovirus genome were transformed with the pDEST20™-PRMT4 plasmids and plated onto LB agar media containing 15 mg.mL⁻¹tetracycline, 7 mg.mL⁻¹gentamicin, 50 mg. mL⁻¹ kanamycin, 25 mg. mL⁻¹ X-Gal and 40 mg. mL⁻¹ IPTG. Bacmid DNA purified from recombination-positive white colonies was transfected into Sf9 cells using the Lipofectin reagent (Invitrogen). Viruses were harvested 10 days after transfection. Sf9 cells were grown at 300 K in suspension culture in Grace medium (Gibco) using Bellco spinner flasks. 2 l of sf9 cell culture (at 0.8 x 10⁶ cells.mL⁻¹) were infected with recombinant GST-mPRMT4 virus with an infection multiplicity of 1. Cells were harvested 48 h post-infection. Cells lysis was performed by sonication in 100 ml buffer A [50 mM Tris-HCl pH 8.0, 250 mM NaCl, 5% glycerol, 5 mM DTT, 0.01% NP40 and anti-proteases (Roche, Complete™, EDTA-free)] and cellular debris were sedimented by centrifugation of the lysate at 40000 g for 30 min. The supernatant was incubated overnight at 277 K with 3 ml glutathione Sepharose resin (GE Healthcare). After a short centrifugation, the supernatants were discarded and the beads were poured in an Econo-column (Bio-Rad). After two wash steps with 10 ml buffer A, 3 ml buffer A supplemented with in-house produced TEV protease were applied to the columns and digestion was performed 4 hours at 303K with gentle mixing. The digest was concentrated with an Amicon Ultra 10K (Milipore), loaded on a gel-filtration column (HiLoad 16/60 Superdex S200, GE Healthcare) and eluted at 1 mL.min⁻¹ with buffer B [20mM Tris-HCl pH 8.0, 100mM NaCl, 1mM TCEP] using an ÄKTA Explorer device (GE Healthcare). Fractions containing mPRMT4 were pooled and concentrated to 5 mg.mL⁻¹.

PRMT4 crystallization and structure determination

Compounds were solubilized in water before addition to the protein solution (2 mg mL⁻¹) at the final concentration of 2 mM. The protein–ligand solution was incubated 20 min at RT before use. Vapor diffusion method utilizing hanging drop trays (VDX plate, Hampton Research) with a 1 mL reservoir was used for crystallization. Typically, 1 µL of protein-ligand solution was added to 1 µL of well solution consisting of 15-25% (v/v) PEG2000MME, 100 mM Tris-HCl pH 8.0 and 100 mM NaCl. Crystals grew in a few days at 20°C. Crystals were flash-frozen in liquid nitrogen after a brief transfer to 5 µL reservoir solution containing 15% (v/v) PEG400 as a cryoprotectant and were stored in liquid nitrogen.

Datasets were collected either at the SOLEIL PROXIMA1 and ESRF ID23-2 beamlines, using a PILATUS 6M detector or PILATUS 2M-F (Dectris), and processed with XDS [4] and HKL-2000 [5]. Five structures of MmCARM1 have been solved and refined. The crystals belong to the P 21 21 2 space group with four MmPRMT4 molecules in the asymmetric unit (AU) or the P 21 space group with eight MmPRMT4 molecules in the AU. The structures were solved by molecular replacement using previously solved Mm PRMT4 structures as the search model [6] Iterative cycles of model building and refinement were carried out using Coot [7], PHENIX [8] and Buster [9]. A summary of the data statistics for the five solved and refined structures is provided in Table 1. The atomic coordinates and experimental data (PDBID 5TBH, 5TBI, 5TBJ, 5LV2 and 5LV3) have been deposited in the Protein Data Bank. All structure figures were prepared with Cuemol (<http://www.cuemol.org>).

PRMT6 cloning, expression and purification

The *Mus musculus* PRMT6 gene sequence (UP Q6NZB1) corresponding to the full length protein was cloned in the pNEA-vHis vector [10] between the *Nde*I and *Xho*I restriction sites and used to transform *E. coli* BL21 (DE3) pRARE2 cells. The resulting recombinant protein harbors an amino-terminal hexahistidine tag followed by a Tobacco etch virus (TEV) protease cleavage site. The protein also contains a F315L point mutation, reported for the entry as a natural variant. The protein expression was induced for 4h at 310 K, with 0.1 mM IPTG in LB medium. The cells were harvested by centrifugation and resuspended in 10 ml per pellet g of buffer A [50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 7 mM β-ME, 20 mM imidazole) supplemented with 1 mM PMSF]. After sonication, the lysate was centrifuged at 40,000 x g for 25 min at 4°C. The supernatant was mixed with 0.5 ml of NiNTA superflow resin (Qiagen) per liter of culture, incubated for 1h at

4°C, and then loaded on an Econo column (Bio-Rad). After washing the column with 10 column volumes (c.v.) of buffer A, the bound protein was eluted with 5 c.v. of buffer A supplemented with 500 mM imidazole. The fractions containing the recombinant protein were detected by a Bradford assay, pooled and supplemented with TEV protease. The sample was then dialyzed in a 10,000 MWCO dialysis membrane (Spectrum) against 1 liter of buffer B [50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 7 mM β-ME and 1 mM EDTA] for 16 h at 4°C. The dialyzed sample was loaded onto a HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare), equilibrated in buffer C [20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM DTT]. The protein was finally concentrated on an Amicon Ultra-4 10,000 MWCO filter (Millipore) to a 6 mg ml⁻¹ final concentration. Samples were either directly used for crystallization or flash-frozen as 25 μl aliquots in liquid nitrogen and stored at -80°C.

PRMT6 crystallization, data collection and structure determination

The compounds were solubilized in water before addition to the protein solution (6 mg ml⁻¹) at the final concentration of 1 mM. Vapor diffusion method utilizing hanging drop trays (VDX plate, Hampton Research) with a 1 ml reservoir was used for crystallization. Typically, 1 μl of protein-ligand solution was added to 1 μl of well solution consisting of 100 mM Tris-HCl or HEPES-NaOH pH 8.0, 200 mM NaCl and 10% (v/v) PEG 6,000. Crystals grew in a 4 to 5 weeks at 20°C. Crystals were flash-frozen in liquid nitrogen after a brief transfer to 5 μl reservoir solution containing 15% (v/v) PEG400 as a cryoprotectant and were stored in liquid nitrogen.

Datasets were collected either at the SOLEIL PROXIMA1 and PROXIMA2 beamlines, using a PILATUS 6M detector (Dectris) or Q315 detector (ADSC), and processed with XDS [4] and HKL-2000 [5]. Two structures of MmPRMT6 have been solved and refined. The crystals belong to the I41 space group with one MmPRMT6 molecule in the AU. The structures were solved by molecular replacement using previously solved MmPRMT6 structures as the search model. Iterative cycles of model building and refinement were carried out using Coot [7], PHENIX [8] and Buster [9]. A summary of the data statistics for the five solved and refined structures is provided in Table 1. The atomic coordinates and experimental data (PDBID 5LV4 and 5LV5) have been deposited in the Protein Data Bank. All structure figures were prepared with Cuemol (<http://www.cuemol.org>).

Permeability studies

Fluorescein assessment for permeability assays.

Fluorescein was used as the cell monolayer integrity marker. Fluorescein permeability assessment (in the A-B direction at pH 7.4 on both sides) was performed after the permeability assay for the test compound. The cell monolayer that had a fluorescein permeability of less than 1.5 x 10⁻⁶ cm/s for Caco-2 and MDR1-MDCKII cells and 2.5 x 10⁻⁶ cm/s for MDCKII cells was considered intact, and the permeability result of the test compound from intact cell monolayer is reported.

Permeability.

The apparent permeability coefficient (P_{app}) of the test compound was calculated as follows:

$$P_{app}(cm/s) = \frac{V_R \cdot C_{R,end}}{\Delta t} \cdot \frac{1}{A \cdot (C_{D,mid} - C_{R,mid})}$$

where V_R is the volume of the receiver chamber. C_{R,end} is the concentration of the test compound in the receiver chamber at the end time point, Δt is the incubation time and A is the surface area of the cell monolayer. C_{D,mid} is the calculated mid-point concentration of the test compound in the donor side, which is the mean value of the donor concentration at time 0 minute and the donor concentration at the end time point. C_{R,mid} is the mid-point concentration of the test compound in the receiver side, which is one half of the receiver concentration at the end time point. Concentrations of the test compound were expressed as peak areas of the test compound.

Recovery of the test compound from the permeability assay.

The recovery of the test compound was calculated as follows:

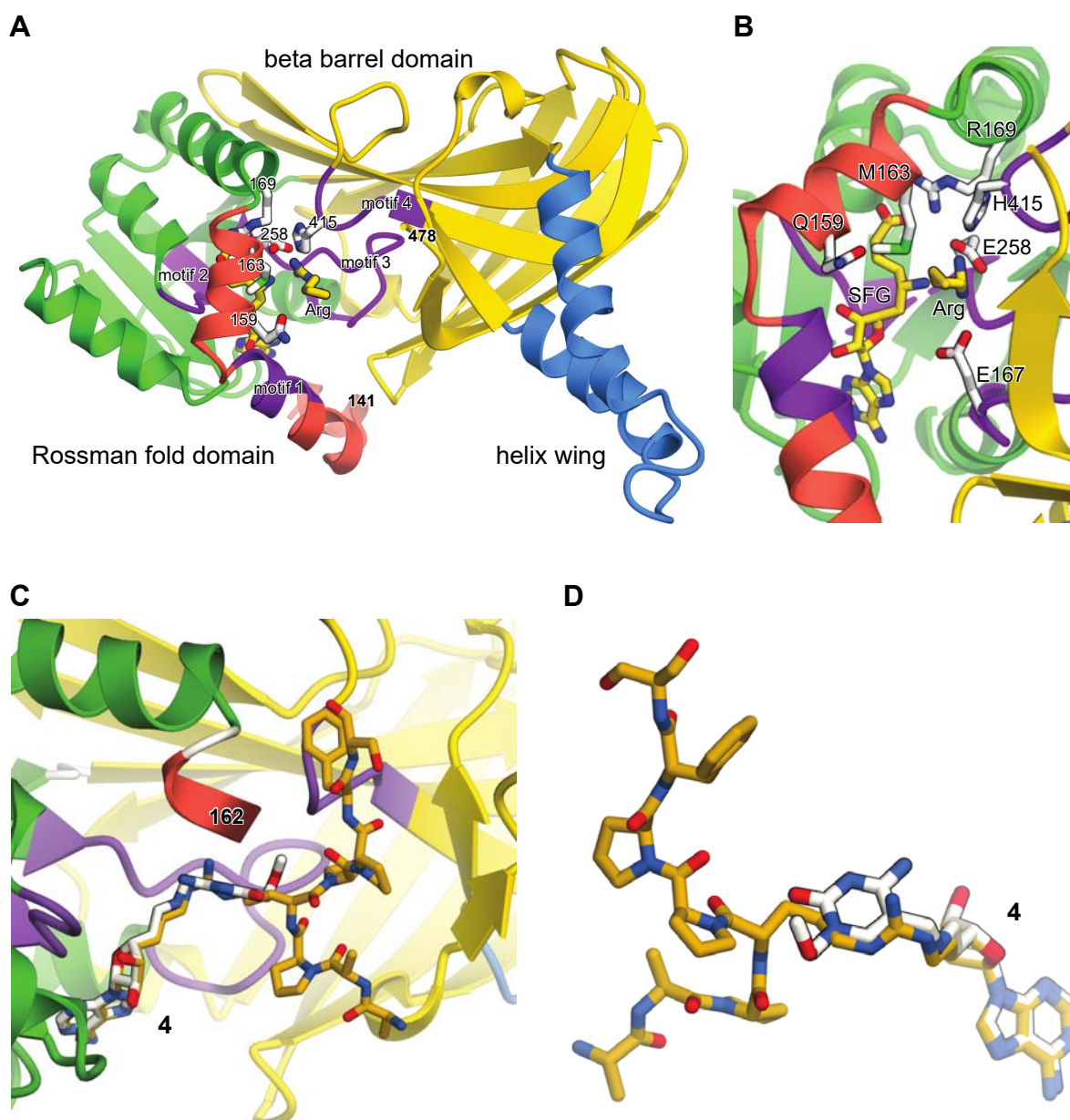
$$\text{Recovery(\%)} = \frac{V_D \cdot C_{D,\text{end}} + V_R \cdot C_{R,\text{end}}}{V_D \cdot C_{D0}} \cdot 100$$

where V_D and V_R are the volumes of the donor and receiver chambers, respectively. $C_{D,\text{end}}$ is the concentration of the test compound in the donor sample at the end time point. $C_{R,\text{end}}$ is the concentration of the test compound in the receiver sample at the end time point. C_{D0} is the concentration of the test compound in the donor sample at time zero. Concentrations of the test compound are expressed as peak areas of the test compound.

Supplementary Biological results

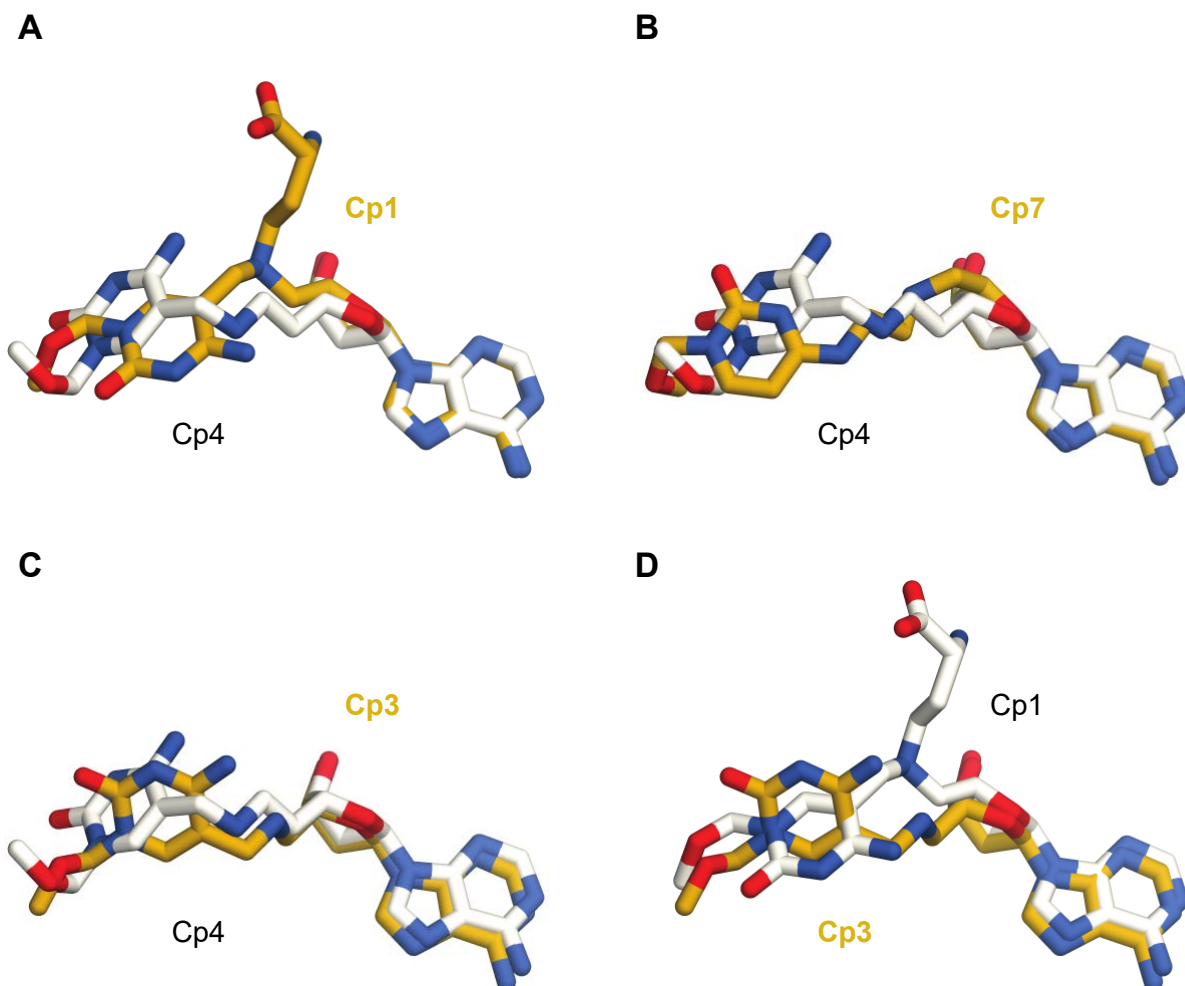
Supplementary Figure 1.

PRMT4 monomer organization and important residues for inhibitors recognition. Overall view of a PRMT4 monomer (**A**) and close-up in the catalytic pocket (**B**). The monomer are drawn as cartoon models with colours corresponding to the different domains and motifs: green for the Rossmann-fold domain, yellow for the beta-barrel domain, blue for the helix wing, red for the first alpha helices and purple for the four signature motifs. Side chains of residues involved in the inhibitors recognition are drawn as white sticks models. The sinefungine (SFG) inhibitor and the target Arginine residue are drawn as yellow sticks models. Water molecules are not shown. Figure drawn using PRMT4-SFG structure PDB ID 5IS6. **Compound 4 occupies the two substrates binding sites: the SAH co-factor and the arginine site.** Binding mode and interactions of Compound 4 on PRMT4 (**C**). For clarity, N-terminal helices of CARM1 are not shown. Conformation of the **4** in the catalytic site as compared to SAM-peptide mimics in yellow (**D**) [11]. The SAM-peptide mimics (PDB: 5LGS) is coloured in yellow. Orientation has been changed between panel C and D.



Supplementary Figure 2.

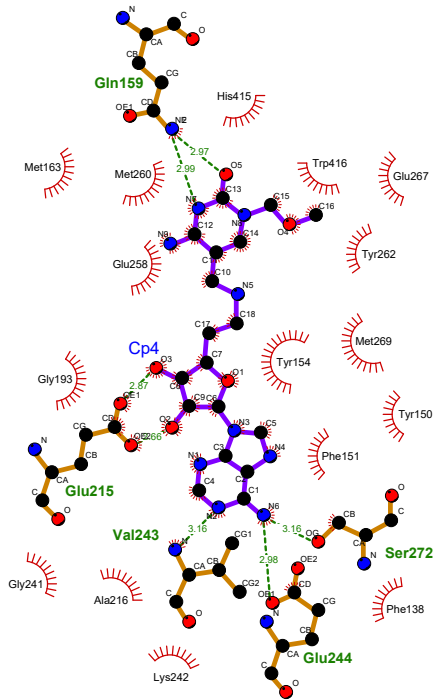
Superposition of bound compounds in PRMT4. (A): compounds **1** and **4**; **(B)** **4** and **7**; **(C)** **4** and **3**; and **(D)** **1** and **3**. Compounds are drawn as stick models and labelled.



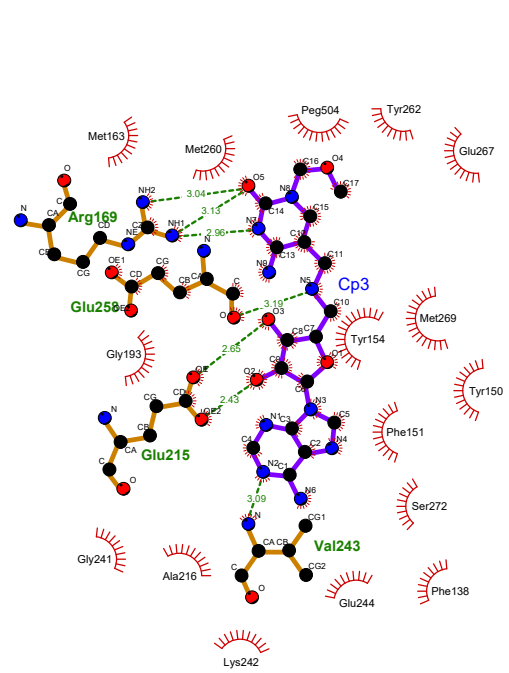
Supplementary Figure 3.

Overview of compounds binding sites in mPRMT4. Flat representation done with LigPlot [12] showing contacts between compounds and active site residues in PRMT4. The ligand bonds are coloured in purple. Hydrogen bonds are shown as dotted lines. The spoked arcs represent protein residues making non-bonded contacts with compound 4. (A): Cp4 (compound 4); (B) Cp3 (compound 3); (C) Cp1 (compound 1); (D) Cp 7 (compound 7).

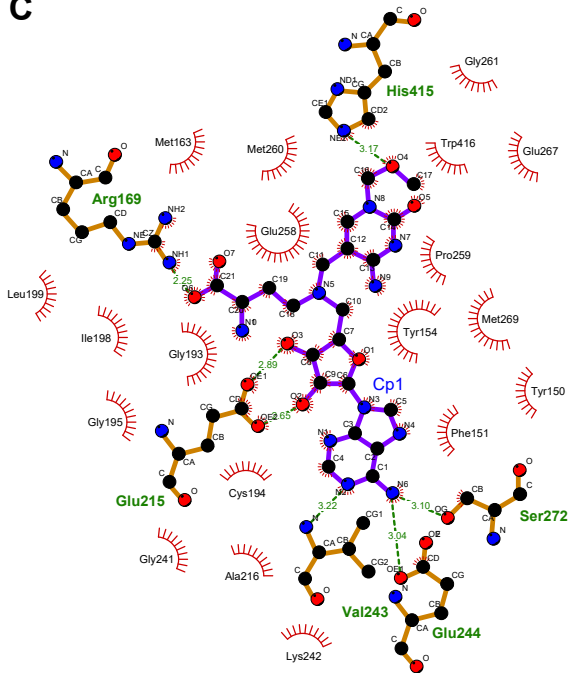
A



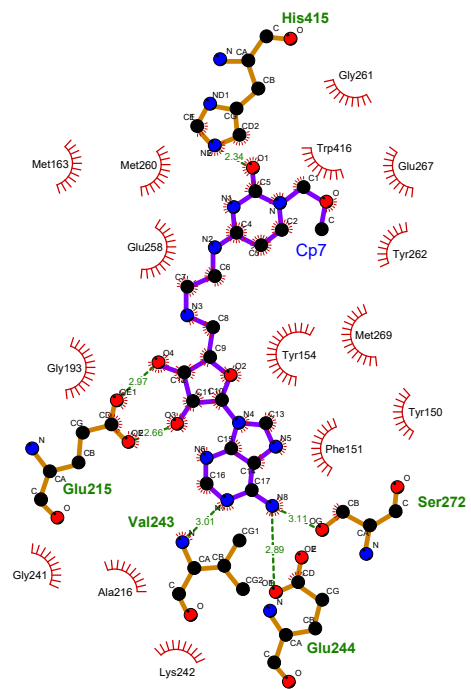
B



C

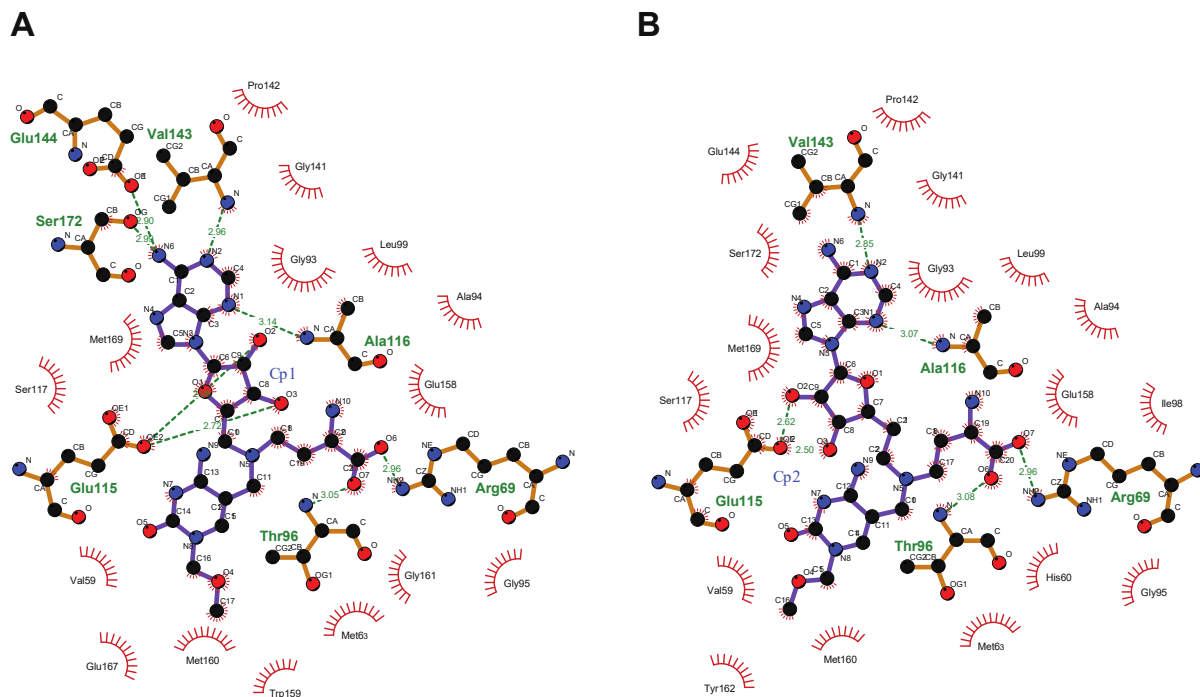


D



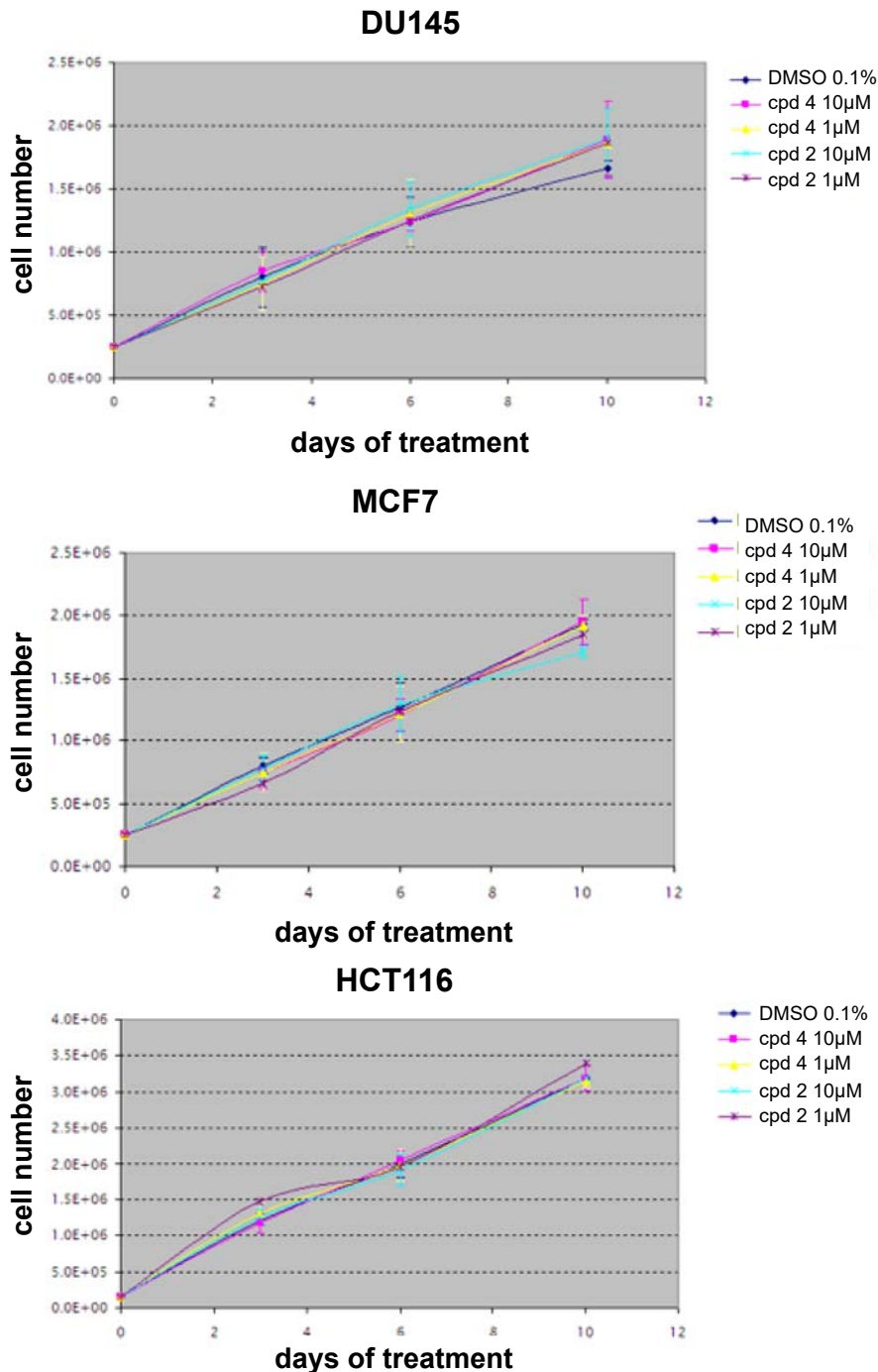
Supplementary Figure 4.

Overview of compounds binding sites in PRMT6. Flat representation done with LigPlot [12] showing contacts between compounds and active site residues in PRMT6. The ligand bonds are coloured in purple. Hydrogen bonds are shown as dotted lines. The spoked arcs represent protein residues making non-bonded contacts with Cp4. (A): Cp1 (compound 1); (B) Cp2 (compound 2)



Supplementary Figure 5. Cell proliferation assays

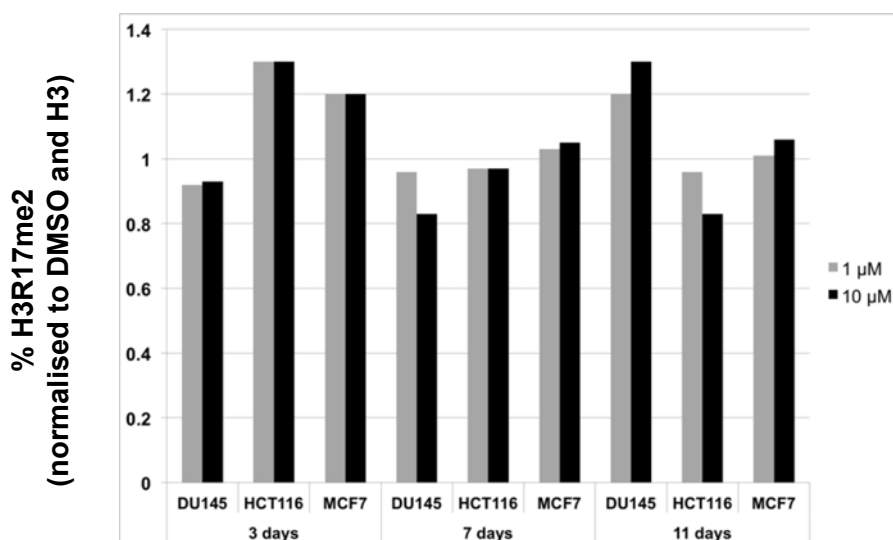
For antiproliferative activity, cells were seeded ($3 \cdot 10^4$ cells/mL for DU145, $2 \cdot 10^4$ cells/mL for MCF7 and $1.5 \cdot 10^4$ cells/mL for HCT116) at day 0 in a 96-well plate. Cells were treated with test compound solutions at a dose range from $3.2 \cdot 10^{-9}$ M to $1 \cdot 10^{-5}$ M. Treatment was repeated every 72h and the antiproliferative activity of compounds was measured using the ATP quantification method "ATPlite one step Luminescence Assay System" (Perkin Elmer). The values presented are the mean of at least two independent experiments run in duplicates.



Supplementary Figure 6.

H3R17 dimethylation analysis in cancer cells.

Commercial DU145, HCT116 and MCF7 cells cultivated in DMEM with 10% SVF at 37°C and 5% CO₂ were treated with compound **4** and harvested after 3, 7 or 11 days. The compound was added at each cell passage in fresh media. Cells were fixed with freshly made 2% PFA (in PBS) for 20 min at 4°C and permeabilised with 0.1% Triton X-100 (0.05% final) (in cold 1x PBS) for 10 min on ice. After centrifugation at 300 g for 5 min, the pellets were washed twice with cold 1x PBS and cells were saturated with 1% BSA in 1x PBS containing 0.01% Tween-20 for 2 h at 4°C (or O/N). After centrifugation at 300 g for 5 min, cells were immunostained with a primary antibody for H3R17me2 (Abcam, #ab8284) and histone H3 (Abcam, #ab1791) for 1 h at RT. Samples were washed and incubated with the secondary antibody (Life Technologies, Alexa Fluor 647-conjugated goat anti-mouse, #A21236, or Alexa Fluor 488-conjugated goat anti-rabbit, #A11034, Ig antibodies) for 30 min at RT. After a final wash, cells were stained with propidium iodide (1:200 dilution in 1x PBS // Sigma #P4864 - 1mg/ml) for 10 min at RT and labeling was analyzed by flow cytometry on a LSRII cytometer (BD Biosciences).



References

- [1] Ceccaldi, A., Rajavelu, A., Champion, C., Rampon, C., Jurkowska, R., Jankevicius, G., Senamaud-Beaufort, C., Ponger, L., Gagey, N., Ali, H. D., et al. 2011 C5-DNA methyltransferase inhibitors: from screening to effects on zebrafish embryo development. *Chembiochem*. **12**, 1337-1345.
- [2] Gros, C., Chauvigne, L., Poulet, A., Menon, Y., Ausseil, F., Dufau, I. & Arimondo, P. B. 2013 Development of a universal radioactive DNA methyltransferase inhibition test for high-throughput screening and mechanistic studies. *Nucleic Acids Res.* (DOI:gkt753 [pii] 10.1093/nar/gkt753 [doi]).
- [3] Chen, D., Ma, H., Hong, H., Koh, S. S., Huang, S. M., Schurter, B. T., Aswad, D. W. & Stallcup, M. R. 1999 Regulation of transcription by a protein methyltransferase. *Science (New York, N.Y.)* **284**, 2174-2177.
- [4] Kabsch, W. 2010 XDS. *Acta Crystallogr Sect D Biol Crystallogr*. **66**, 125-132.
- [5] Otwinowski, Z. & Minor, W. 1997 Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol*. **276**, 307-326.
- [6] Troffer-Charlier, N., Cura, V., Hassenboehler, P., Moras, D. & Cavarelli, J. 2007 Functional insights from structures of coactivator-associated arginine methyltransferase 1 domains. *EMBO J* **26**, 4391-4401. (DOI:10.1038/sj.emboj.7601855).
- [7] Emsley, P. & Cowtan, K. 2004 Coot: Model-Building Tools for Molecular Graphics. *Acta Cryst.* **D60**, 2126-2132.
- [8] Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., et al. 2010 PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D Biological Crystallography* **66**, 213-221. (DOI:10.1107/S0907444909052925 S0907444909052925 [pii]).
- [9] Bricogne G., B. E., Brandl M., Flensburg C., Keller P., Paciorek W., Roversi P, Sharff A., Smart O.S., Vonrhein C., Womack T.O. 2011 BUSTER version 2.10.0. (Cambridge, United Kingdom, Global Phasing Ltd.
- [10] Romier, C., Ben Jelloul, M., Albeck, S., Buchwald, G., Busso, D., Celie, P. H., Christodoulou, E., De Marco, V., van Gerwen, S., Knipscheer, P., et al. 2006 Co-expression of protein complexes in prokaryotic and eukaryotic hosts: experimental procedures, database tracking and case studies. *Acta Crystallogr D Biol Crystallogr* **62**, 1232-1242. (DOI:10.1107/S0907444906031003).
- [11] van Haren, M. J., Marechal, N., Troffer-Charlier, N., Cianciulli, A., Sbardella, G., Cavarelli, J. & Martin, N. I. 2017 Transition state mimics are valuable mechanistic probes for structural studies with the arginine methyltransferase CARM1. *Proc Natl Acad Sci U S A* **114**, 3625-3630. (DOI:10.1073/pnas.1618401114).
- [12] Laskowski, R. A. & Swindells, M. B. 2011 LigPlot+: Multiple Ligand-Protein Interaction Diagrams for Drug Discovery. *J Chem Inf Model* **51**, 2778-2786. (DOI:10.1021/ci200227u).

Étude structurale des Protéine Arginine Méthyltransférases

Reconnaissance des substrats et conception rationnelle de modulateurs

Résumé

Les protéine arginine méthyltransférases (PRMT) sont impliquées dans de nombreux processus cellulaires, incluant la régulation de l'expression des gènes, le contrôle de l'épissage, le maintien de l'intégrité du génome et la transduction du signal. De nombreuses études montrent que la dérégulation de l'activité des PRMT est associée au développement de pathologies, et en particulier de cancers. Les PRMT constituent ainsi une des nouvelles cibles potentielles en chimiothérapie. Les travaux présentés dans ce manuscrit portent sur trois cibles : PRMT2, PRMT3 et PRMT4/CARM1. Combinant des approches biochimiques, biophysiques et structurales (cristallographie et cryo-microscopie électronique), ces travaux comportent deux aspects : (I) comprendre au niveau atomique la régulation de la réaction de méthylation des protéines (reconnaissance protéines-protéines et interactions entre modifications post-traductionnelles) ; (II) découvrir des inhibiteurs spécifiques et puissants de plusieurs PRMT cibles.

Mots clés : PRMT ; epigenetique ; methylation des arginines, analogues « SAM-peptides » ; inhibiteurs ; biologie structurale.

Summary

Protein arginine methyltransferases (PRMT) are involved in many cellular processes, including gene expression regulation, splicing control, maintenance of genome integrity and signal transduction. Since deregulation of those biological processes appears to be implicated in the pathogenesis of different diseases, PRMTs have emerged as potential new targets for the development of novel therapeutic modulators. Despite the large amount of biological and structural data on PRMTs, two challenges remain to be solved by structural biology ; (I) understanding how PRMTs recognize and bind their full-length substrates ; (II) revealing how PRMTs achieve specific arginine methylation on different target sites. The works presented here focused on 3 targets: PRMT2, PRMT3 and PRMT4/CARM1. We used biochemical, biophysical and structural methods (bio-crystallography and cryo-electron microscopy) to decipher structural clues that drive PRMT-substrate recognition. We developed new chemical probes that can be used in early drug discovery for the conception of PRMT inhibitors.

Keywords : PRMTs ; epigenetic ; arginine methylation ; SAM-peptides analogs ; inhibitors ; structural biology.