



HAL
open science

Synthèse de ligands à la protéine CARM1 pour l'étude de son activité enzymatique et la synthèse d'inhibiteurs sélectifs

Samira Ajebbar

► **To cite this version:**

Samira Ajebbar. Synthèse de ligands à la protéine CARM1 pour l'étude de son activité enzymatique et la synthèse d'inhibiteurs sélectifs. Autre. Université de Strasbourg, 2012. Français. NNT : 2012STRAF002 . tel-00769956

HAL Id: tel-00769956

<https://theses.hal.science/tel-00769956>

Submitted on 4 Jan 2013

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.

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES**Laboratoire Conception et Applications de Molécules Bioactives****THÈSE**

présentée par

Samira AJEBBAR

soutenue le 11 mai 2012

pour obtenir le grade de

Docteur de l'université de Strasbourg

Discipline/ Spécialité : Chimie

Synthèse de ligands à la protéine CARM1 pour l'étude de son activité enzymatique et la conception d'inhibiteurs sélectifs.

THÈSE dirigée par :**Mr. Wagner Alain**

Dr., Université de Strasbourg

RAPPORTEURS :**Mme Arimondo Paola B.**

Dr., CNRS-Pierre Fabre

Mr. Ploux Olivier

Pr., ENSCP Chimie Paris Tech

MEMBRES DU JURY :**Mr. Bourguignon Jean-Jacques**

Dr., Université de Strasbourg

Remerciements

Ces travaux de thèse ont été réalisés au sein du Laboratoire des Systèmes Chimiques Fonctionnels sous la direction du Dr. Alain Wagner. C'est la première personne que je tenais à remercier pour m'avoir fait confiance en me gardant au laboratoire après mon stage et m'avoir ainsi permis de réaliser une thèse. Je tenais également à le remercier pour son encadrement et pour m'avoir transmis une méthode de travail qui m'a aidée à avancer tout au long de ces quatre ans.

Je tiens à témoigner toute ma gratitude aux membres du jury pour m'avoir fait l'honneur de juger mes travaux : Dr. Paola Arimondo, Pr. Olivier Ploux et Dr. Jean-Jacques Bourguignon. Je tiens également à remercier le Pr. Jean Cavarelli, avec qui nous avons collaboré, pour sa contribution à l'avancement du projet.

J'exprime également ma gratitude pour toutes les autres personnes avec qui j'ai collaboré : je vais commencer par le Pr. Jean-Marie Wurtz qui m'a accordée du temps pour m'apprendre à maîtriser les logiciels de modélisation et je le remercie aussi pour ses conseils en informatique. Merci aussi au Dr. Vincent Cura pour son implication dans mon projet, sa disponibilité et pour m'avoir initiée aux tests d'inhibition. Merci au Dr. Pierre Hassenboeler et à Justine Maillot pour avoir testé mes molécules et pour sa relecture de l'introduction.

Un grand merci à Pascal Eberling pour sa gentillesse, sa bonne humeur, sa disponibilité et pour m'avoir dépannée quelquefois pour la purification de quelques uns de mes produits.

Je tenais également à remercier le Dr. Denis Weltin et Céline Schmidt de Phytodia pour leur disponibilité et pour avoir réalisé dans les meilleurs délais les tests MTT.

Je remercie également le Dr. Sarah Sanglier et Stéphanie Piétot pour avoir réalisé les analyses de spectrométrie de masse.

Je remercie chaleureusement tout le service d'analyse de la Faculté de Pharmacie pour leur patience, Cyril Antheaume pour les analyses RMN, Pascale Buisine, Patrick Wehrung, Syndie Berthe et Nicole Rimlinger pour les analyses de masse.

Je voudrais également remercier Nathalie, Marianne et Liliane pour leur gentillesse, leur disponibilité et pour toutes les démarches administratives.

Enfin, je remercie toute l'équipe de m'avoir apportée aide et soutien durant cette thèse en particulier Maria, Coraline, Jérémy, Marie-Charlotte et Jessie pour ces bons conseils et pour m'avoir remontée le morale plus d'une fois.

Je tenais également à remercier les personnes exceptionnelles que j'ai eu la chance de rencontrer à mon arrivée au laboratoire : Laure, Aurélia et Cédric pour leur gentillesse et leur générosité et pour avoir embelli mes journées de thèses parfois difficiles ; merci à Laure pour ton humour et ton soutien jusqu'au bout, merci à Aurélia pour les bon moments et tes bons petits plats et merci à Cédric pour tes bons conseils et pour tes sorties organisées.

Je tenais également à remercier ma meilleure amie, Fatou, pour ses bons conseils et pour m'avoir remonté le moral plus d'une fois tout au long de mes études et surtout pendant la thèse.

Pour finir, je tiens à remercier mes parents et ma famille pour m'avoir soutenue tout au long de mes études, pour leur confiance, leur amour et pour leurs petites attentions qui ont amélioré mon quotidien.

Merci beaucoup.

Sommaire

Abréviations.....	1
Avant-propos.....	3
Chapitre 1 : Introduction générale	11
I. Modifications épigénétiques	13
I.1 Définition de l'épigénétique.....	13
I.2 De l'ADN vers la constitution de la chromatine	13
I.3 Modifications chimiques liées aux modifications épigénétiques	17
<i>I.3.1 La méthylation de l'ADN</i>	<i>18</i>
<i>I.3.2 La Modification des histones</i>	<i>19</i>
II. Les modifications épigénétiques et le cancer	28
II.1 Développement des cellules cancéreuses du à des modifications épigénétiques	29
<i>II.1.1 L'identité d'une cellule</i>	<i>29</i>
II.2 La thérapie épigénétique.....	31
II.3 Les inhibiteurs des cibles épigénétiques	33
<i>II.3.1 Inhibiteurs des DNMTs</i>	<i>33</i>
<i>II.3.2 Les inhibiteurs des HKMTs</i>	<i>37</i>
<i>II.3.3 Les inhibiteurs des PRMTs</i>	<i>38</i>
III. Etude du Coactivator Associated aRginine Methyltransferase 1	45
III.1 Présentation de CARM1	45
<i>III.1.1 Architecture de CARM1</i>	<i>47</i>
<i>III.1.2 Structure tridimensionnelle de CARM1</i>	<i>47</i>
<i>III.1.3 Inhibiteurs de CARM1</i>	<i>51</i>
III.2 Notre stratégie pour étudier le mode de fixation de CARM1 avec ses substrats.....	53

Chapitre 2 : Exploration de la « poche du sulfonium » par co-cristallisation d’analogues de la SAM 65

- I. A stepwise strategy towards selective inhibitors and peptide mimics of coactivator-associated arginine methyltransferase 1.....67**
- II. Essais de co-cristallisation des analogues de la SAM dans mCARM1₁₃₀₋₄₈₇87**
- III. Conclusions et perspectives.....87**

Chapitre 3 : Exploration de la « poche de l’arginine » par co-cristallisation d’analogues de bisubstrats 89

- I. Mécanisme de méthylation proposé dans la littérature pour les PRMTs de type I.....91**
 - I.1 Réactions.....91
 - I.1.1 Postulat du mécanisme de méthylation par une approche cinétique 92*
 - I.1.2 Etude du mécanisme sur les bases des structures cristallographiques de CARM1..... 93*
- II. Etude du mécanisme de transfert de méthyle faisant intervenir la SAM et une guanidine..94**
 - II.1 Etude de l’environnement de la « poche de l’arginine » de CARM1 et PRMT194
 - II.2 Comparaison de la « poche de l’arginine » de CARM1 avec une autre enzyme utilisant la SAM..... 96
 - II.2.1 La guanidine acétate méthyltransférase (GAMT)..... 96*
- III. Modèle en deux dimensions, proposé pour la réaction de méthylation de l’arginine du substrat, dans le site catalytique de CARM198**
 - III.1 Image en deux dimensions du modèle proposé.....98
 - III.2 Image en trois dimensions représentant l’environnement de l’arginine99
- IV. Synthèse des analogues de bisubstrats qui se rapprochent de l’état de transition et qui correspondent à la première étape de méthylation.....101**
 - IV.1 Stratégie pour synthétiser des inhibiteurs des PRMTs 101
 - IV.2 Synthèse d’analogues de bisubstrats..... 104
 - IV.2.1 Synthèse du dérivé 35..... 105*
 - IV.2.2 Synthèse des analogues de bisubstrats 45 à 47 et 51 à 53 106*
- V. Résultats et discussions108**
 - V.1 Tests *in vitro* pour l’inhibition de CARM1 et de PRMT1 sur le tétramère de l’histone (H3H4)₂..... 108

V.2	Evaluation de la cytotoxicité des analogues de bisubstrats 45 à 47 et 51 à 53	114
V.2.1	<i>Principe du test MTT</i>	114
V.2.2	<i>Résultats des tests MTT sur les analogues 45 à 47 et 51 à 53</i>	115
V.3	Etudes structurales.....	116
V.3.1	<i>Structures cristallographiques des analogues 45 et 51</i>	116
V.3.2	<i>Confirmation du modèle</i>	119
V.3.3	<i>Comparaison des structures cristallographiques de CARM1 et PRMT1 pour expliquer la sélectivité</i>	124
VI.	Conclusions et perspectives.....	125
<u>Chapitre 4 : Approche pour l'exploration du « domaine de fixation du peptide » par réaction <i>in-situ</i></u>		129
I.	Stratégie employée pour explorer le « domaine de fixation du peptide » par réaction <i>in-situ</i>	131
II.	Réaction <i>in-situ</i> entre une molécule hameçon azoture et un peptide alcyne par réaction de click.....	133
II.1	Antécédents, dans la littérature, de la réaction de click <i>in-situ</i> employée en chimie combinatoire.....	136
II.1.1	<i>Réaction in-situ au sein de l'acétylcholine estérase (AChE)</i>	136
II.1.2	<i>Réaction in-situ au sein de la protéase du VIH-1 (Pr SF-2-WTQ7K-Pr)</i>	138
II.2	Essais de co-cristallisation du mélange réactionnel dans CARM1.....	140
II.2.1	<i>Synthèse du peptide PJ207</i>	140
II.2.2	<i>Synthèse de la molécule hameçon 54</i>	142
II.2.3	<i>Essais de co-cristallisation du mélange réactionnel</i>	142
III.	Analyse par spectrométrie de spectrométrie de masse du mélange réactionnel.....	144
III.1	Principe de l'analyse des complexes non-covalents par spectrométrie de spectrométrie de masse	144
III.2	Analyse par spectrométrie de masse de la réaction <i>in-situ</i> et du « mélange réactionnel » en conditions dénaturantes ou natives	145
III.2.1	<i>Analyse du mélange réactionnel en conditions dénaturantes</i>	145
III.2.2	<i>Analyse du mélange réactionnel en conditions natives</i>	147
IV.	Réaction <i>in-situ</i> entre une molécule hameçon électrophile et un peptide cystéine par réaction de substitution nucléophile.....	151

IV.1	Molécules hameçons électrophiles pour la capture de peptides cystéine.....	151
IV.1.1	<i>Synthèses des molécules hameçons 63 à 70</i>	153
IV.2	Essais de réactions <i>in-situ</i> entre les molécules hameçons électrophiles et les peptides cystéines puis co-cristallisation dans CARM1	156
IV.2.1	<i>Mesures d'inhibition des molécules hameçons</i>	156
IV.2.2	<i>Essais de co-cristallisation du mélange réactionnel</i>	157
V.	Réaction <i>in-situ</i> entre une molécule hameçon électrophile et un peptide arginine par réaction de condensations	158
V.1.	Molécules hameçons électrophiles pour la capture de peptides arginine.....	159
V.2	Essais de réactions <i>in-situ</i> entre les molécules hameçons électrophiles et les peptides arginines puis co-cristallisation dans CARM1	164
V.2.1	<i>Mesures d'inhibition des molécules hameçons</i>	164
V.2.2	<i>Essais de co-cristallisation du mélange réactionnel</i>	166
VI.	Réaction <i>in-situ</i> entre une molécule hameçon électrophile guanidine et un peptide cystéine par réaction de substitution nucléophile.....	169
VI.1	Molécule hameçon électrophile guanidine pour la capture de peptides cystéine	170
VI.1.1	<i>Synthèse de la molécule hameçon 90</i>	171
VI.2	Essais de réactions <i>in-situ</i> entre la molécules hameçons 90 et les peptides cystéines puis co-cristallisation dans CARM1	172
VI.2.1	<i>Mesures d'inhibition de la molécule hameçon 90</i>	173
VI.2.2	<i>Essais de co-cristallisation du mélange réactionnel</i>	173
VII.	Conclusions et perspectives.....	174
<u>Chapitre 5 : Approche pour l'exploration du « domaine du peptide » par essais de co-cristallisation de molécules ancres dans CARM1</u>.....		179
I.	Stratégie employée pour explorer le « domaine de fixation du peptide »	181
II.	Exploration du « domaine de fixation du peptide » par essais de co-cristallisation d'adduits SAM-peptides obtenus par réaction de cycloaddition de Huisgen	184
II.1	Adduits SAM-peptide obtenus par « Cu (I)-catalyzed azide-alkyne cycloaddition » (ou CuAAC) et « Ru (II)-catalyzed azide-alkyne cycloaddition » (ou RuAAC).....	184
II.2	Synthèse des adduits SAM-peptide 93, 103 à 110	185
II.3	Caractérisations biochimiques des adduits SAM-peptides 93, 103 à 110	190

II.3.1	<i>Mesures d'inhibition des adduits SAM-peptides 93, 103 à 110</i>	190
II.3.2	<i>Essais de co-cristallisation des molécules ancrées 93, 103 à 110 dans mCARM1₁₃₀₋₄₉₀</i>	193
III.	Exploration du « domaine de fixation du peptide » par essais de co-cristallisation d'un adduit SAM-peptide obtenu grâce à une réaction de substitution nucléophile	194
III.1	Adduit SAM-peptide obtenu par substitution nucléophile entre la cystéine modifiée 74 et une molécule hameçon électrophile	194
III.2	Synthèse de l'adduit SAM-peptide 114 et de la molécule hameçon 115	195
III.3	Caractérisations biochimiques de l'adduit SAM-peptide 114 et de la molécule hameçon 115	196
III.3.1	<i>Mesures d'inhibition de l'adduit SAM-peptide 114 et molécule hameçon 115</i>	196
III.3.2	<i>Essais de co-cristallisation de l'adduit SAM-peptide 114 et de la molécule hameçon 115 dans CARM1</i>	198
IV.	Conclusions et perspectives	198
	<u>Conclusions et perspectives</u>	201
	<u>Partie expérimentale</u>	211
I.	General procedure	213
II.	Chemistry	214
III.	Conditions expérimentales pour l'analyse par spectrométrie de masse des assemblages non covalents	274
III.1	Analyse en conditions dénaturantes.....	274
III.2	Analyse en conditions natives	275
III.3	Préparation de l'échantillon	275

Abréviations

Les abréviations sont en français lorsqu'il y a un consensus pour une dénomination française, lorsque le consensus n'est pas connu, elles sont en anglais.

AChE : Acétylcholine estérase	EcoDam : méthyltransférase de l'ADN d' <i>Escherichia coli</i>
ACN : Acétonitrile	EDC : Chlorhydrate du 1-(3-Diméthylaminopropyl)-3-éthylcarbodiimide
ADN : Acide désoxyribonucléique	EGCG : Epigallocatechin-3-gallate
ADP-ribose : Adénosine diphosphate ribose	<i>ESRF</i> : <i>European synchrotron radiation facility</i>
APCI : Ionisation chimique à pression atmosphérique	<i>Ezh2</i> : <i>Enhancer of zeste homolog 2</i>
ApoMB : Apo myoglobine	FAD : Flavine adénine dinucléotide
ARN : Acide ribonucléique	FDA: <i>Food and drug administration</i> (ou en français : « Agence fédérale américaine des produits alimentaires et médicamenteux)
Asp : Acide aspartique	Fmoc : Groupement 9-fluorénylméthyl-oxycarbonyle
5-Aza-CR : 5-azacytidine	FT : Facteur de transcription
5-Aza-CdR : 5-aza-2'-désoxycytidine	
Boc: Groupement <i>tert</i> -butoxycarbonyle	GAA : Guanidinoacétate
BS : Benzène sulfonamide	GAMT : Guanidine acétate méthyltransférase
BSA : Sérum d'albumine bovin	GCN5: <i>General control of amino acid synthesis protein 5-like 2</i>
BTP : Bis-tris propane	Glu : Acide glutamique
	GNATs : <i>Gcn5-N-acetyltransferases</i>
CA : Co-activateurs transcriptionnels	GRIP1: <i>Glucocorticoid receptor interacting protein 1</i>
CARM1 : <i>Coactivator-associated arginine methyltransferase 1</i>	HAT(1) : Acétyltransférase d'histone (1)
CBP/P300: CREB binding protein/ <i>p300</i>	HBTU : 2-(1H-benzotriazol-1-yl)-1,1,3,3-tétraméthyluronium hexafluorophosphate
CHLP -MS: Chromatographie en phase liquide à haute performance couplée à la spectrométrie de masse	HDAC : Déacétylase d'histone
CI ₅₀ : Concentration inhibitrice médiane	hDNMT1: Méthyltransférase 1 humaine de l'ADN
CpG : Cytosine-Guanine	HEPES : Acide N-(2-hydroxyéthyl)-piperazine-N'-(2-éthanesulfonique)
CR : Corépresseurs transcriptionnels	His : Histidine
CuAAC : <i>Cu (I)-catalyzed azide-alkyne cycloaddition</i>	HKDM : Lysine déméthylase d'histone
δ : Déplacement chimique	HKMT : Lysine méthyltransférase d'histone
Da : Dalton	HMT : Méthyltransférase d'histone
DAD : Détecteur à barrette de diode	HOBt: 1-Hydroxybenzotriazole
<i>a</i> DMA : <i>ω-N_G,N_G</i> -asymétrique diméthylarginine	HRE : Élément de réponse aux hormones
<i>s</i> DMA : <i>ω-N_G,N_G</i> -symétrique diméthylarginine	HSQC : <i>Heteronuclear Single Quantum Coherence</i>
DMAP: 4-diméthylaminopyridine	
DMEM : milieu Eagle modifié de Dulbecco	
DMF : N,N-diméthylformamide	
DMSO: Diméthylsulfoxyde	
DNMT : (cytosine-5-)-méthyltransférase de l'ADN	
Dot1: <i>Disruptor of telomeric silencing-1</i>	
DS : Déviation standard	

IC₅₀: *Half maximal inhibitory concentration*
IGBMC : Institut de génétique et de biologie
moléculaire et cellulaire
IR : Infraréd

J : Constante de couplage
JHDM : *JmC domain-containing histone
demethylase*
JMJD2 : *Jumonji domain-containing protein
2*
JmjC : Jumonji C

LSD1 : *Lysine specific demethylase 1
related factor*

MBD : *Methyl-CpG Binding Domain*
MES : Acide 2-morpholino éthanesulfonique
monohydraté
Met : Méthionine
MGO : Méthylglyoxal
MMA : ω -N_G-monométhylarginine
MORF : *Monocyclic leukemia zinc finger
related factor*
MOZ : *Monocytic leukemia zinc finger
protein*

NAD : Adénine nicotinamide dinucléotide
NMR: *Nuclear magnetic resonance*
NS : Naphtyles sulfonates
NU : Naphtyles urées

PAD4/ PADI4 : Peptidyl- arginine déiminase
4
PEG : Polyéthylène glycol
MME : Monométhyl éther
PDB : Banque de données des protéines
PRMT : Méthyltransférase de l'arginine

RMN : Résonance magnétique nucléaire
RN : Récepteurs nucléaires
RS : Rapport de sélectivité
RuAAC : *Ru (II)-catalyzed azide-alkyne
cycloaddition*

SAM: *S*-adénosyle-L-méthionine
SAH: *S*-adénosyle-L-homocystéine
SETs: *Suppressor of variegation, enhancer
of zeste and trithorax*

SFG: Sinéfungine
S_N2 : Substitution nucléophile bimoléculaire
SPPS : Synthèse peptidique sur support solide
Suv39h: *Suppressor of variegation 3-9 homolog
2*

TBBD : 2,3,7,8-Tétrahydroxy(1)benzopyrano
(5,4,3-cde)-(1)benzopyran-5,10-dione
THF : Tétrahydrofurane
Tip 60: *Tat interacting protein 60*
TLC : *Thin layer chromatography*
TRDMT : *tRNA (cytosine(38)-C(5))-
methyltransferase*
Tris : Trishydroxyméthylaminométhane
Tyr : Tyrosine

UV : Ultra-violet

VIH-1 : Virus de l'immunodéficience humaine
de type 1

Avant-propos

Avant-propos

Les modifications épigénétiques jouent un rôle prépondérant dans la régulation de l'expression des gènes en modulant la structure de la chromatine par des modifications chimiques au niveau de l'Acide DésoxyriboNucléique (ADN) et des protéines liées à ce dernier, les histones. Parmi ces modifications chimiques, la méthylation des arginines sur les histones est catalysée par les méthyltransférases d'arginines (PRMTs). Ces enzymes catalysent la réaction de méthylation des arginines sur les histones, *via* le donneur de méthyle naturel, la *S*-Adénosyle-*L*-Méthionine (SAM). Parmi les neuf PRMTs présentes chez les mammifères, PRMT4, plus connue sous le nom de Coactivator-Associated aRginine Methyltransferase 1 (CARM1), est une protéine cruciale impliquée dans de nombreux processus de la vie cellulaire, qui ne seront cependant pas détaillés ici. La dérégulation de ces processus peut induire certains cancers comme celui du sein ou de la prostate. Ainsi, la compréhension à l'échelle moléculaire du mode d'action de CARM1 s'avère très important tant au niveau fondamental que pour des applications pharmacologiques.

Le but de cette thèse est de développer des outils chimiques permettant d'appréhender le mécanisme d'action de l'enzyme CARM1 durant le processus de méthylation afin de comprendre son mode d'interaction avec l'histone. Ce travail a été réalisé en collaboration avec l'équipe des Prs. Jean Cavarelli et Jean-Marie Wurtz de l'IGBMC de Strasbourg.

En 2007, l'équipe de J. Cavarelli a mis en évidence la structure cristallographique du site catalytique de CARM1. Elle a été obtenue seule ou en complexe avec un cofacteur qui était soit la *S*-Adénosyle-*L*-Homocystéine (SAH), produit de la réaction de méthylation, soit la Sinefungine, un inhibiteur naturel des méthyltransférases. L'étude menée a montré que la présence de ces cofacteurs dans l'enzyme a pour effet de stabiliser le site catalytique. Notre

stratégie est donc de créer des molécules hameçons basées sur le motif de la SAM capables de capturer un peptide mimant la séquence de l'histone H3, pour ensuite les co-cristalliser avec CARM1. Ainsi, grâce à la diffraction aux rayons X, les interactions mises en jeu dans le complexe CARM1/ molécules hameçons/peptide pourront être déterminées (**figure 1**).

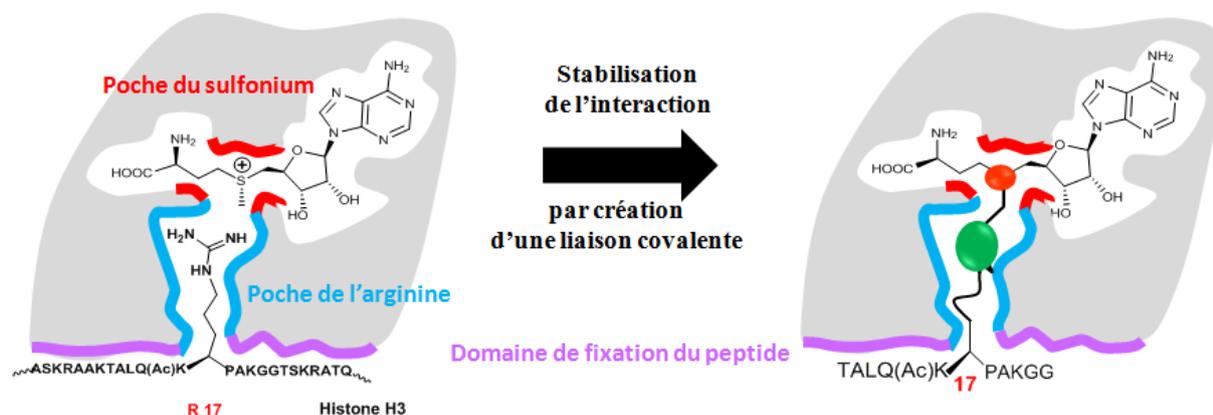


Figure 1 : Stratégie employée pour stabiliser le complexe CARM1/molécules hameçons/peptide : stabilisation de l'interaction en créant une liaison covalente à partir de molécules hameçons ayant pour motif la SAM.

Cette stratégie se déroulera en trois étapes : la première consistera en la synthèse d'analogues de la SAM obtenus grâce à des modifications réalisées autour de l'atome de soufre. Ces composés permettront l'exploration de la « poche du sulfonium ». La seconde étape sera la synthèse d'analogues de bisubstrats nécessaires pour l'exploration de la « poche de l'arginine ». Pour finir, pour étudier le « domaine de fixation du peptide », nous aborderons la synthèse d'adduits SAM-peptide comprenant une ancre SAM, un motif traversant la poche de l'arginine et un peptide.

Ce manuscrit sera composé en cinq parties. La première partie décrira les modifications épigénétiques ainsi que les enzymes impliquées dans la méthylation de l'ADN et des histones. Nous dresserons une liste non exhaustive des inhibiteurs décrits dans la littérature. Pour finir, l'enzyme CARM1 et son rôle dans les processus cellulaires seront présentés.

Dans le second chapitre, nous exposerons nos travaux sur l'exploration de la « poche du sulfonium ». Plusieurs analogues de la SAM seront synthétisés en modifiant le degré d'oxydation du soufre et en effectuant des modifications isostériques. La fonction sulfonium de la SAM sera ainsi remplacée par un atome d'oxygène, un atome de sélénium ou un atome d'azote. Puis, une deuxième série d'analogues azotés sera réalisée en incorporant dans leur structure une chaîne hydrophobe, hydrophile ou encore un hétérocycle. Ces analogues s'avèreront être de bons inhibiteurs sélectifs de CARM1 par rapport à une enzyme appartenant à la même famille, PRMT1. La synthèse de ces analogues sera décrite ainsi que leurs caractérisations biochimiques. Cette deuxième partie sera décrite sous la forme d'une publication. Grâce aux structures cristallographiques obtenues, les analogues nous permettront de commencer à explorer la « poche de l'arginine ».

Dans le troisième chapitre, des analogues se rapprochant de l'état de transition de la réaction de méthylation nous permettront d'explorer la « poche de l'arginine » de manière plus approfondie. Pour cela, des analogues de la SAM incorporant dans leur structure un motif guanidine seront synthétisés. La longueur de la chaîne portant ce motif a été modifiée : elle sera composée de deux ou trois carbones afin de se rapprocher au maximum de l'état de transition. Deux structures cristallographiques nous permettront d'obtenir des informations précieuses sur l'environnement de l'arginine du substrat dans le site catalytique de CARM1. Après validation de cette étape, nous continuerons notre exploration du « domaine de fixation du peptide » suivant deux stratégies. La première consiste à stabiliser un complexe « CARM1/molécules hameçons/peptide » en créant un lien covalent par réaction de click *in situ*. Cette approche sera basée sur la capacité de l'enzyme à rapprocher les deux réactifs jusqu'à ce qu'ils soient connectés. La seconde stratégie consiste à stabiliser un complexe

« CARM1/molécules ancrées » en préformant des adduits SAM-peptides. Les stratégies employées seront décrites dans les deux derniers chapitres.

Dans la quatrième partie, la stratégie employée pour renforcer l'interaction entre une molécule hameçon et le peptide sera détaillée. La première approche choisie reposera sur la création d'un lien covalent entre une molécule hameçon et un peptide mimant la séquence de l'histone H3 dans lequel l'arginine 17, réactive, sera modifiée pour réagir *in-situ* :

- Grâce à une réaction de cycloaddition de Huisgen [3+2] (ou réaction de « click ») entre une molécule hameçon azoture et un peptide alcyne.
- Grâce à une substitution nucléophile entre des molécules hameçons électrophiles capables de capturer un peptide cystéine.
- Par réaction de condensation entre des molécules hameçons électrophiles capables de capturer un peptide arginine.

La synthèse de ces composés ainsi que leurs caractérisations biochimiques seront décrites. Ces complexes CARM1/molécules hameçon/peptide seront soumis à des essais de co-cristallisation.

Malheureusement, tous ces essais se révéleront infructueux car aucune technique n'a mis en évidence la formation d'adduits. Par conséquent, nous présenterons, dans le dernier chapitre, une alternative à cette stratégie et une série d'adduits sera donc préformée pour ancrer un peptide dans l'enzyme CARM1. Ces molécules ancrées seront obtenues par réaction:

- De click entre des molécules hameçons azotures et des peptides alcynes. Ces cycloadditions auront lieu soit par réaction de « Cu(I)-catalyzed azide-alkyne cycloaddition » (CuAAC) ou par réaction de « Ru(II)-catalyzed azide-alkyne cycloaddition » (RuAAC).
- De substitution nucléophile entre une molécule hameçon électrophile et une cystéine modifiée.

Ces molécules ancrées seront ensuite testées en co-cristallisation dans CARM1.

Pour conclure, les résultats obtenus pendant cette thèse et les perspectives envisagées seront discutés.

Chapitre 1

Introduction générale

Chapitre 1

Introduction générale

I. Modifications épigénétiques

I.1 Définition de l'épigénétique

L'épigénétique est l'étude des changements héréditaires dans la fonction des gènes qui ont lieu sans altération de la séquence Acide DésoxyriboNucléique (ADN). Elle détermine comment ces modifications sont susceptibles de désactiver ou d'activer les gènes⁽¹⁾. En 1942, l'invention du terme « épigénétique » est attribuée à Conrad Waddington qui définit l'épigénétique comme « la branche de la biologie qui étudie les relations de cause à effet entre les gènes et leurs produits, faisant apparaître le phénotype ».

Trois processus biochimiques importants se distinguent, définissent l'épigénétique et interagissent ensemble : la méthylation de l'ADN, les modifications des histones et le remodelage de la chromatine.

I.2 De l'ADN vers la constitution de la chromatine

L'ADN est constitué d'une succession de nucléotides qui sont composés d'un sucre (désoxyribose), d'un acide phosphorique et d'une base azotée. Il existe quatre bases azotées divisées en deux grandes familles : les bases puriques (adénine et guanine) et les bases pyrimidiques (thymine et cytosine) (figure **1.1**).

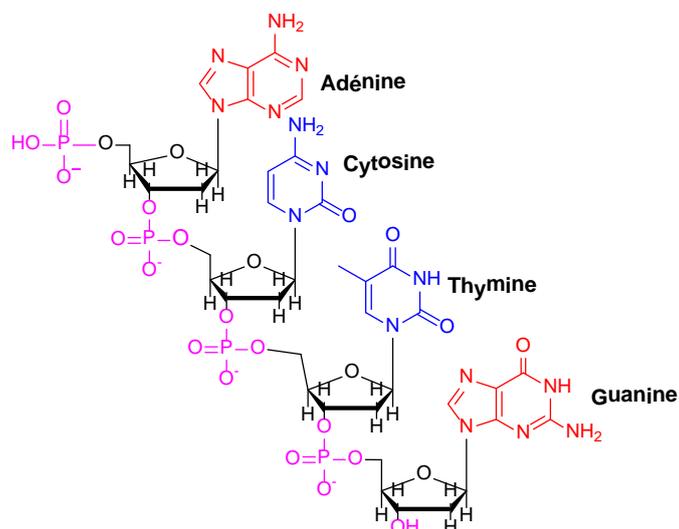


Figure 1.1 : Représentation d'une séquence de nucléotides possédant les bases puriques (en rouge) et pyrimidiques (en bleu). Les nucléosides sont liés entre eux par une jonction phosphate (en violet) et le désoxyribose est dessiné en noir.

Ces brins d'ADN interagissent ensemble par liaisons hydrogènes pour former la double hélice de l'ADN (figure 1.2).

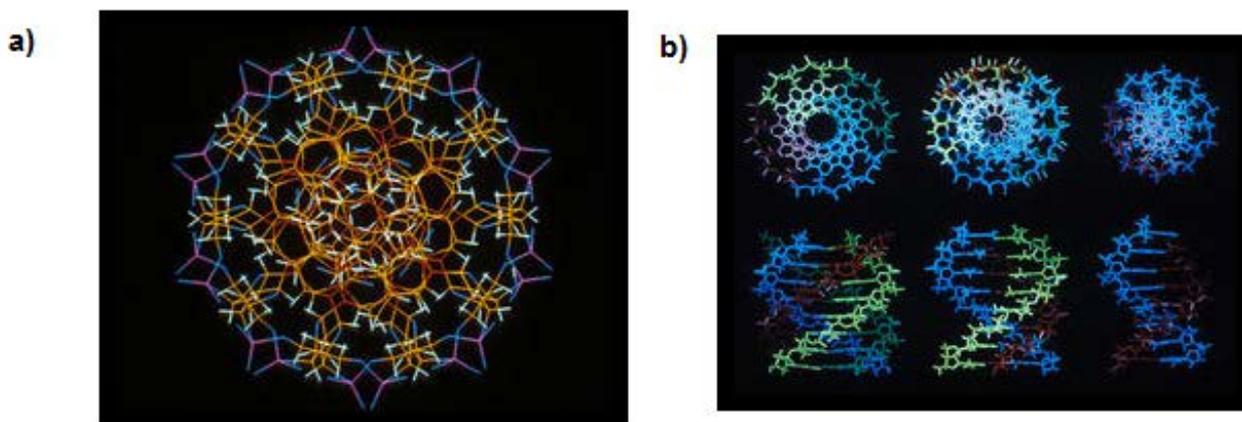


Figure 1.2 : Modélisation de l'ADN. **a)** Vue transversale de l'ADN où les atomes de carbone sont représentés en orange, les oxygènes en bleu, les azotes en rouge, les hydrogènes en blanc et les phosphores en violet⁽²⁾. **b)** La rangée du dessus correspond à une modélisation d'un double, d'un triple et d'un quadruple brin d'ADN. Dans la rangée du dessous, ces mêmes modélisations sont représentées selon l'axe de l'ADN⁽²⁾.

Plusieurs protéines liées à l'ADN peuvent être distinguées, notamment les histones et les protéines chromosomiques non histone, comme la topoisomérase d'ADN II par exemple. Nous nous intéresserons uniquement aux histones.

Les histones sont des protéines basiques étroitement associées à l'ADN formant des structures appelées nucléosomes⁽³⁾. Chaque nucléosome est constitué de 146 paires de bases d'ADN et comprend un cœur protéique (ou octamère d'histones) formé de deux exemplaires de chacune des histones (H3, H4, H2A et H2B). L'ADN s'enroule sur 1,7 tour autour de cet octamère pour former une superhélice⁽⁴⁾. Le nucléosome constitue le premier niveau de compaction de l'ADN dans le noyau. Les histones de la classe H1 sont les histones dites « de liaison » : une histone de cette classe lie l'ADN à l'endroit où celui-ci rentre et sort de la particule cœur, et « scelle » ainsi le nucléosome. La structure cristallographique du nucléosome, représentée dans la figure 1.3, été déposée dans la banque de données de protéines (ou PDB) : elle a pour référence 2NZD et a été obtenue avec une résolution de 2,65 Å⁽⁵⁾.

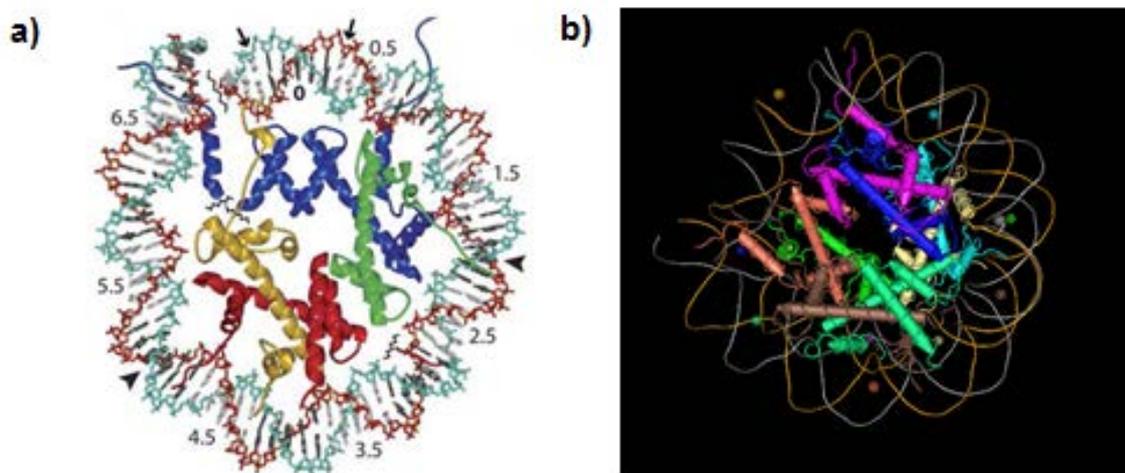


Figure 1.3 : Structure du nucléosome constitué de 146 paires d'ADN et de l'octamère d'histones. **a)** Vue de dessous de la superhélice de l'ADN montrant les bases d'ADN en bleu et rouge. Les protéines des histones sont en bleu pour H3, en vert pour H4, en jaune pour H2A et en rouge pour H2B⁽⁴⁾. **b)** Structure RX du nucléosome qui a été déposée dans la PDB sous la référence 2NZD (résolution de 2,65 Å)⁽⁵⁾.

Les nucléosomes sont ensuite répétés de façon régulière et organisée pour former une structure plus complexe, avec un niveau d'organisation plus compact : la chromatine

(figure 1.4). Cet assemblage atteint un niveau d'organisation maximal pour former le chromosome⁽⁶⁾.

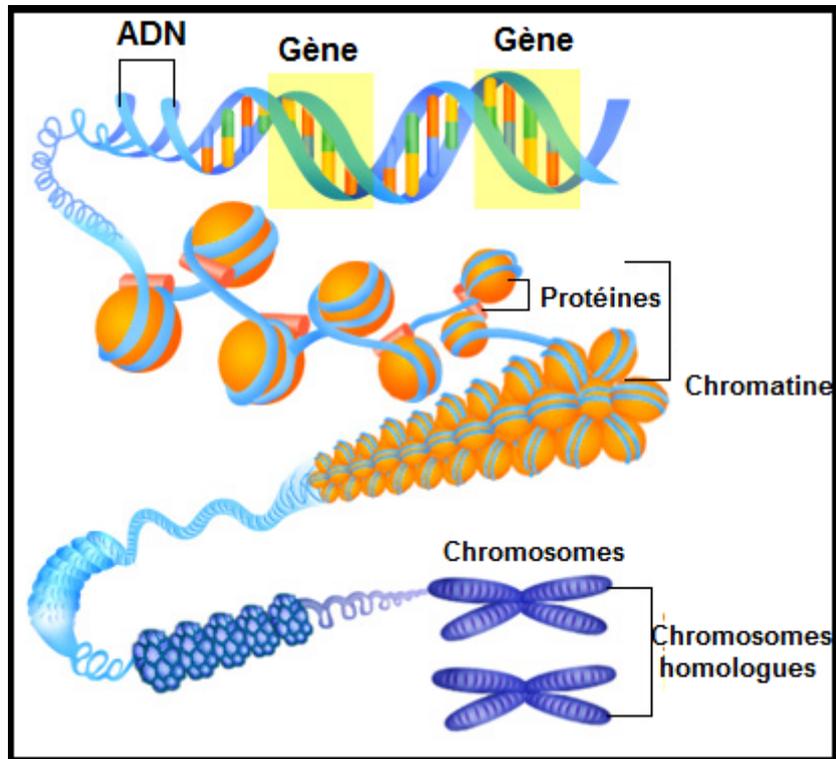


Figure 1.4 : Constitution du chromosome : les gènes sont constitués de séquences nucléotidiques (multicolores) qui forment des couples le long de l'armature de l'ADN (en bleu) ; ce dernier s'enroule ensuite autour des protéines histones (en orange) pour former la chromatine puis les chromosomes (en bleu foncé) à l'intérieur du noyau⁽⁷⁾.

La chromatine est condensée à différents degrés dans le noyau : en effet, les étapes de condensation et de décondensation peuvent être régulées durant le cycle cellulaire. Par exemple, lors de la mitose, la chromatine est fortement compactée afin de permettre la ségrégation adéquate des chromosomes ; en revanche, pendant les autres phases cellulaires, elle est plutôt diffuse et partiellement décondensée⁽⁸⁾. A l'intérieur du noyau cellulaire, deux positions de la chromatine sont observées:

- L'état « ouvert » de la chromatine, ou euchromatine, qui caractérise l'activité transcriptionnelle de l'ADN.

- L'état « fermé » de la chromatine, ou hétérochromatine, qui est associé à la non-transcription de l'ADN.

Les différentes modifications post-traductionnelles observées sur les histones, comme la méthylation ou l'acétylation vont influencer l'état de compaction de la chromatine et permettre ou non l'expression génique ⁽⁹⁾. Cette expression est condition de la reconnaissance d'une zone régulatrice située en amont du gène reconnue par des protéines spécifiques qui vont assurer le recrutement d'autres protéines de la machinerie transcriptionnelle.

Ces différentes modifications chimiques vont être décrites dans le paragraphe suivant.

1.3 Modifications chimiques liées aux modifications épigénétiques

Les modifications covalentes de la chromatine ont lieu soit sur les histones ou soit directement sur l'ADN. Chez la majorité des eucaryotes, la méthylation de l'ADN est située au niveau des promoteurs des gènes, particulièrement ceux associés à des transposons ou à des gènes soumis à empreinte. Elle a lieu sur le carbone C5 des cytosines, habituellement au sein de dinucléotides CpG (ou îlots CpG, Cytosine-Guanine) ou à l'intérieur de trinucleotides CNG (ici, N représente un nucléotide quelconque). Elle provoque généralement la répression de la transcription. La méthylation des îlots CpG situés au sein des promoteurs est un processus bien connu de répression de l'expression des gènes cibles ⁽¹⁰⁾.

I.3.1

La méthylation de l'ADN

Les enzymes (cytosine-5-)-méthyltransférases de l'ADN (DNMTs) sont responsables de la méthylation de l'ADN⁽¹¹⁾. Durant ce processus, elles catalysent le transfert d'un groupe méthyle, *via* le donneur de méthyle naturel, la SAM. La méthylation a lieu sur la position cinq de la cytosine, présente sur un résidu précédant une guanine (figure 1.5).

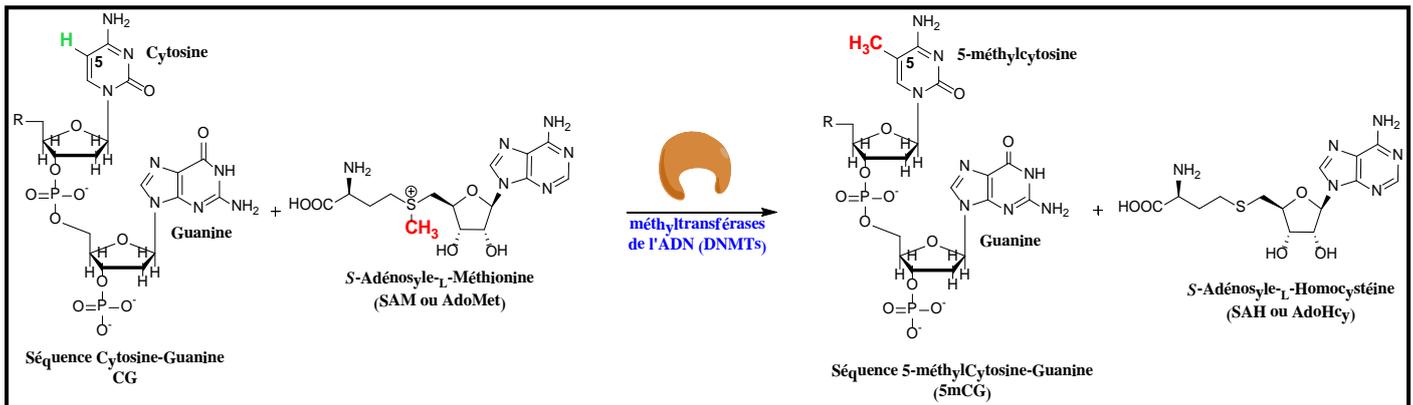


Figure 1.5 : Méthylation de l'ADN sur la position 5 de la cytosine appartenant à la séquence cytosine-guanine (CpG) ; cette étape a lieu *via* le donneur de méthyle naturel, la S-Adénosyle-L-Méthionine (SAM) et est catalysée par une enzyme, méthyltransférase de l'ADN (DNMT). Les produits formés sont la 5-méthylcytosine sur cette séquence et la S-Adénosyle-L-Homocystéine.

Chez les mammifères, quatre méthyltransférases de l'ADN existent : il s'agit de la DNMT1, TRDMT (pour *tRNA* (cytosine(38)-C(5))-methyltransferase), DNMT3A et DNMT3B. La première est impliquée dans le maintien de l'information épigénétique et prend comme modèle le brin d'ADN hémi-méthylé lors de la réplication. En revanche, les deux autres méthyltransférases DNMT3A et 3B peuvent agir sur un brin non méthylé et modifient donc l'information épigénétique (figure 1.6). L'activité de TRDMT n'est pas encore définie : son site catalytique possède la même signature que pour les autres DNMTs mais son activité est faiblement comparable par rapport à ces dernières. Récemment, il a été montré que TRDMT est capable de méthyler l'ARN transfert acide aspartique, ce qui suggère que les enzymes méthyltransférases sont spécifiques de certains substrats⁽¹²⁾.

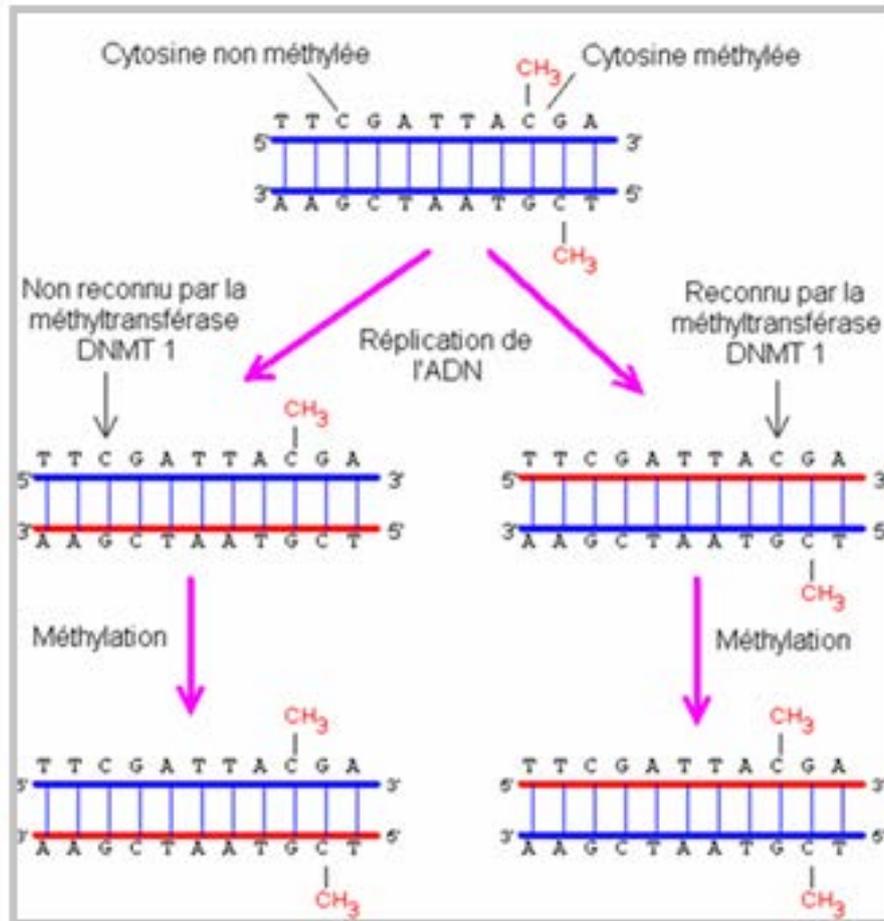


Figure 1.6 : DNMT1 est une méthylase de maintenance ; elle est impliquée dans la transmission du profil de méthylation de la cellule mère aux cellules filles au cours de la réplication de l'ADN. Les deux autres méthyltransférases DNMT3A et 3B établissent les profils de méthylation au cours du développement (méthylases *de novo*)⁽¹³⁾.

I.3.2 La Modification des histones

Plusieurs types de modifications sur les histones existent : l'addition de petits groupements tels qu'un méthyle, un acétyle ou un phosphate est possible. Le transfert de plus gros groupements tels qu'un poly (ADP- ribose) ou des petites protéines ubiquitines peut aussi être observé⁽¹⁴⁾.

Chez les mammifères, certains résidus d'histones subissent plusieurs modifications (figure 1.7) telles que :

- L'acétylation sur certains résidus lysines (K) des histones H2A, H2B, H3 et H4.

- La phosphorylation sur certains résidus sérines (S) de ces quatre mêmes histones.
- La méthylation sur certains résidus lysines et arginines (R) des histones H2A, H3 et H4.
- L'ubiquitination sur les résidus lysines K129 de l'histone H2A et K120 de l'histone H2B.

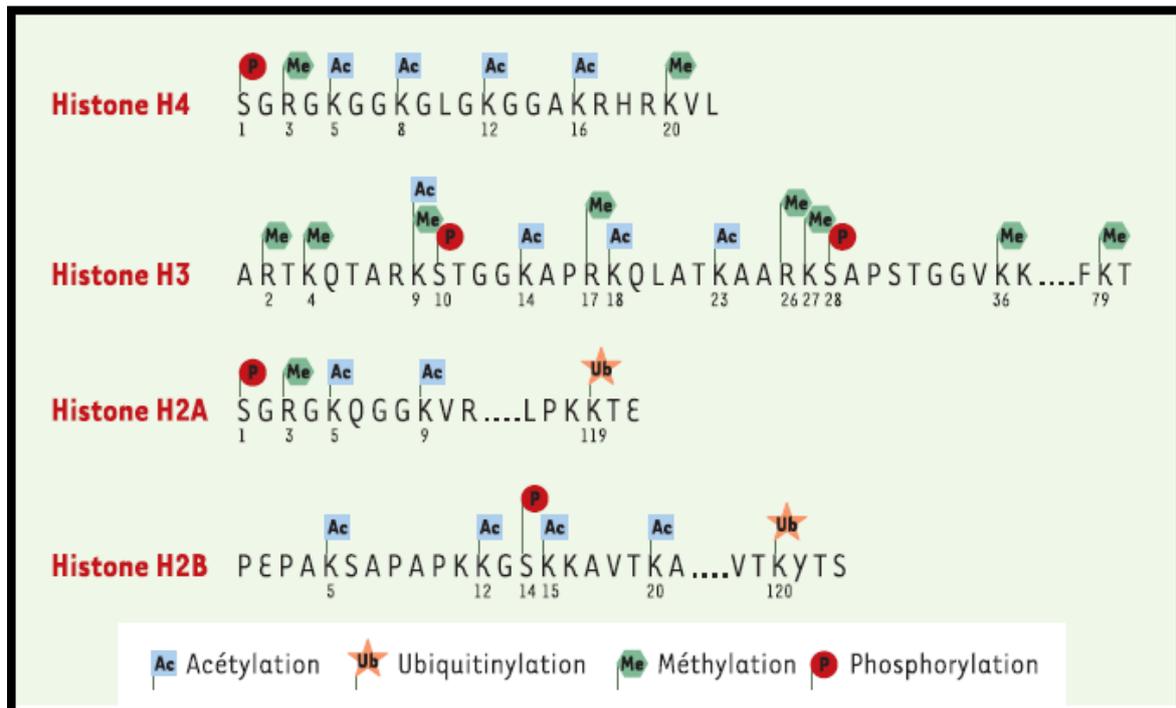


Figure 1.7 : Résidus d'acides aminés des histones sujets à des modifications post traductionnelles chez les mammifères. Par exemple, la lysine 9 de l'histone H3 peut être soit acétylée, soit méthylée, ces modifications étant mutuellement exclusives⁽¹⁶⁾.

Ces modifications peuvent être effectuées selon différentes combinaisons, sur des sites spécifiques qui sont connus sous le nom de « code histone »⁽¹⁵⁾. Elles permettent de comprendre pourquoi il est possible d'obtenir une grande variété d'expression de gènes et de différents phénotypes pour un seul code génétique⁽¹⁶⁾.

La méthylation et l'acétylation sont les modifications chimiques sur les histones qui sont les plus étudiées. Par conséquent, les enzymes impliquées dans ces processus seront

décrites de façon non exhaustive et seules celles présentes chez les mammifères seront présentées.

1.3.2.1 L'acétylation

L'acétylation est la première modification chimique observée où l'amine d'un résidu lysine est acétylée *via* l'acétyl-coenzyme A. Cette transformation conduit à la formation de la lysine acétylée correspondante et au coenzyme A. Cette réaction est catalysée par une enzyme appelée AcétylTransférase d'Histone (HAT), identifiée en 1996 sous la forme d'un co-activateur transcriptionnel (figure 1.8).

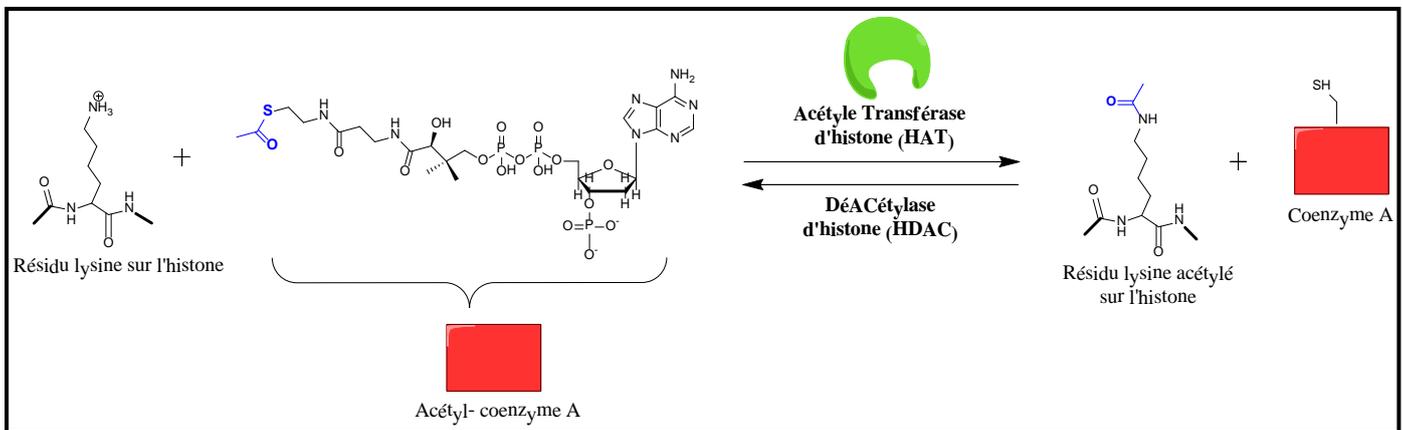


Figure 1.8 : Réaction d'acétylation sur un résidu lysine d'une histone, catalysée par l'enzyme HAT, *via* l'acétyl-coenzyme A. Le processus est réversible et le résidu acétylé peut être déplacé par les enzymes HDAC.

Les HATs sont réparties en plusieurs familles, en fonction des motifs de séquence conservés. Les deux principales familles sont :

- GNATs (apparentées à la Gcn5-N-AcetylTransferases) dont GCN5 (pour *General control of amino acid synthesis protein 5-like 2*)^(17a) et HAT1 (pour Acétyltransférase d'histone 1)^(17b-d) font parties.
- MYST dont MOZ (pour *monocyclic leukemia zinc finger protein*), MORF (pour *monocyclic leukemia zinc finger related factor*) et TIP60 (pour *Tat interacting protein 60*)^(17e-h) font parties.

D'autres qui portent également une activité lysine acétyltransférase, comme CBP/P300 (pour *CREB binding protein/P300*)⁽¹⁷ⁱ⁾, certains co-activateurs de récepteurs nucléaires ou facteurs de transcription.

L'acétylation des histones régule la transcription, la réparation et la réplication de l'ADN, mais est aussi responsable de la condensation de la chromatine. Elle joue un rôle dans la progression du cycle cellulaire, la recombinaison et la réparation de l'ADN, ainsi que dans l'apoptose⁽¹⁷⁾. Concernant le rôle des HATs dans les maladies cancéreuses, rien ne peut être généralisé : en effet, ces enzymes fonctionnent soit comme suppresseurs ou soit comme activateurs de tumeurs en fonction du type et de l'état d'avancement de ces dernières.

Parfois, le processus est réversible et ces résidus modifiés peuvent être désacétylés par d'autres enzymes, appelées histones désacétylases (HDAC). Chez les eucaryotes, les 18 HDACs sont regroupées en quatre classes⁽¹⁸⁾ :

- La classe I : HDACs 1, 2, 3, et 8
- La classe II: HDACs 4, 5, 6, 7, 9 et 10
- La classe III : Sirtuines humaines (SIRT) 1-7
- La classe IV : HDAC 11

Les classes I, II et IV sont des enzymes zinc-dépendantes tandis que la classe III représente l'adénine nicotinamide dinucléotide (NAD⁺).

Les protéines d'histones déacétylases jouent un rôle crucial dans le remodelage de la chromatine et sont impliquées dans une multitude de processus biologiques tels que la régulation de la transcription, la progression du cycle cellulaire, de la survie de la cellule ainsi que de sa différenciation⁽¹⁹⁾.

Passons maintenant à la méthylation des histones, qui est la deuxième modification la plus fréquente de ces unités protéiques.

1.3.2.2 La méthylation

La méthylation des histones peut avoir lieu sur des résidus lysines ou arginines et est catalysée respectivement par les lysines méthyltransférases (HKMTs) ou les arginines méthyltransférases (PRMTs) ^(15b). Comme pour la méthylation de l'ADN, le donneur de méthyle utilisé est la SAM.

1.3.2.2.1 La méthylation des lysines

La méthylation des lysines sur les histones est catalysée par les enzymes lysines méthyltransférases (HKMTs) (figure 1.9).

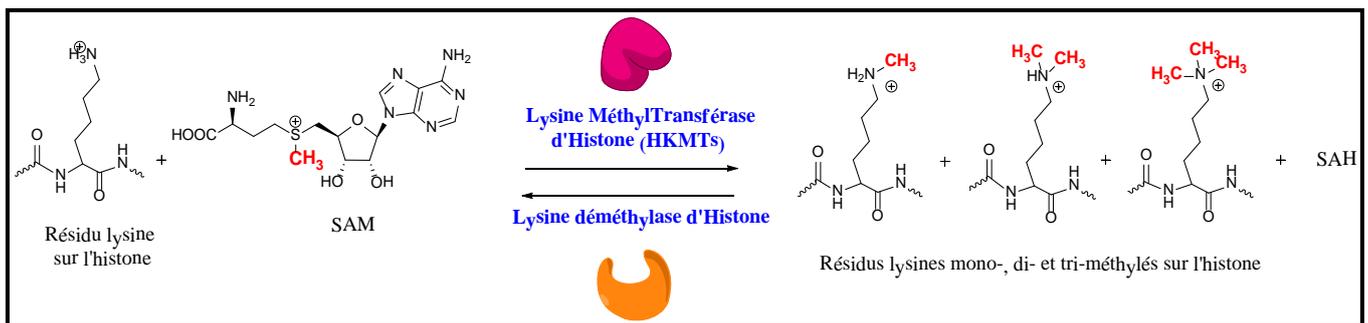


Figure 1.9 : Réaction de méthylation sur le résidu lysine des histones *via* la SAM : la réaction est catalysée par les HKMTs. Cette réaction est réversible et les lysines peuvent être déméthylées par une enzyme lysine déméthylase d'histone, comme par exemple la LSD1 (pour *Lysine specific demethylase 1*).

Deux types de repliements ont été identifiés pour ces protéines :

- La famille Dot1 (pour *Disruptor Of Telomeric silencing-1*)
- Les HKMTs à domaine SET (pour *Suppressor of variegation, Enhancer of zeste and Trithorax*)

Les HKMTs catalysent la mono-, di- ou tri-méthylation et la différence de méthylation donne lieu à diverses fonctions. Par exemple, la diméthylation de la lysine 4 de

l'histone H3 (H3K4) provoque soit l'activation ou la désactivation d'un gène, tandis que la triméthylation de cette même lysine active exclusivement les gènes⁽²⁰⁾. Cette étape de méthylation régule donc la transcription et la réparation de l'ADN.

Le processus est réversible et les résidus méthylés peuvent être déméthylés par les lysines déméthylases d'histones (HKDMs). Deux familles de lysines déméthylases ont été identifiées, il s'agit de la LSD1 (pour *lysine specific demethylase 1*) et la famille des protéines contenant le domaine JmjC (pour Jumonji C). La première est une amine oxydase FAD (pour flavine adénine dinucléotide) dépendante qui est capable de déméthyle la lysine 4 mono- ou di-méthylée de l'histone H3 (H3K4me1/2) (tableau 1.1). En revanche elle ne cible pas les lysines triméthylées comme H3K4me3. Cette action est associée à la répression de la transcription⁽²¹⁾.

La seconde famille de protéines (JmjC) utilise des cofacteurs de Fe(II) ou d'alpha-cétoglutarate pour rendre la réaction de méthylation des lysines réversibles. Au sein de cette famille, deux sous familles peuvent être distinguées :

- JHDM (pour *JmC domain-containing histone demethylase*)
- JMJD2 (pour *Jumonji domain-containing protein 2*)

La plupart des lysines déméthylases d'histone sont substrats spécifiques comme indiqué dans le tableau 1.1 :

HKDMs	Cibles spécifiques	Références
-------	--------------------	------------

LSD1		H3K4Me1/2	21
JHDM	JHDM1A	H3K36me1/2	22
	JHDM2A	H3K9me1/2	23
JMJD2	JMJD2A	H3K9me3	24-25
		H3K36me3	
	JMJD2B	H3K9me3	26
	JMJD2C	H3K9me3	27
		H3K9me2	
	JMJD2D	H3K9me2/3	28

Tableau 1.1 : Les différentes HKDMs sont classées selon leurs cibles associées. Par exemple, LSD1 déméthyle la lysine 4 mono-ou di-méthylée de l’histone H3.

I.3.2.2.2 La méthylation des arginines

L’azote du groupement guanidine des arginines peut être diméthylé *via* la SAM et la réaction est catalysée par les méthyltransférases d’arginines (PRMTs). La diméthylation a lieu de façon asymétrique ou symétrique en fonction de la classe de PRMTs.

Chez les mammifères, la famille des PRMTs contient neuf PRMTs (PRMT1 à PRMT9) qui sont divisées en deux classes principales ⁽²⁹⁾:

- Le type I: PRMT1 à 3, CARM1 (Coactivator-associated Arginine (R) Methyltransferase 1) ou PRMT4, PRMT6, PRMT8
- Le type II: PRMT5 et PRMT9

Les deux classes transfèrent un groupement méthyle sur l'azote du groupement guanidine de l'arginine, *via* la SAM, pour conduire à l'intermédiaire monométhylarginine (MMA). Dans un second temps, un second méthyle est transféré avec une régiosélectivité qui dépend du type de l'enzyme : en effet, les enzymes de type I conduisent à la ω - G_N, G_N -diméthylarginine asymétrique (*a*DMA) tandis que les enzymes de type II forment la ω - G_N, G_N' -diméthylarginine symétrique (*s*DMA) (figure 1.10).

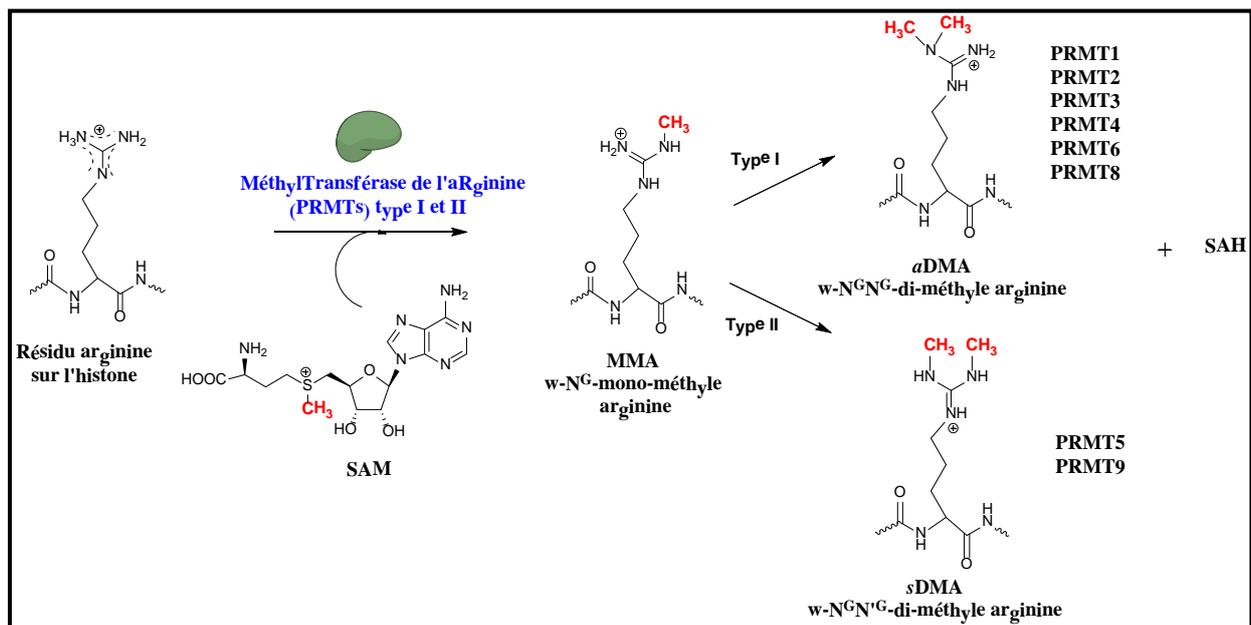


Figure 1.10 : Réaction de méthylation sur le résidu arginine des histones *via* la SAM. La réaction est catalysée par les PRMTs : après la formation du MMA, les PRMTs de type I catalysent la formation de l'*a*DMA tandis que les PRMTs de type II provoquent la formation de la *s*DMA.

Cette méthylation régule la transcription ; de plus, la plupart de ces enzymes sont substrats spécifiques et correspondent à différentes fonctions ^(14b), comme indiqué dans le tableau 1.2 :

PRMTs	Cibles spécifiques	Fonctions
PRMT1		Activation des gènes

H4R3		
CARM1	H3R2, H3R17, H3R26	Activation des gènes
PRMT5	H4R3, H3R8	Répression des gènes

Tableau 1.2 : Les différentes PRMTs sont classées selon leurs cibles associées et leurs fonctions. Par exemple, CARM1 méthyle les arginines 2, 17 et 26 des histones H3 et active les gènes ^(14b).

Les PRMTs sont impliquées dans plusieurs mécanismes cellulaires vitaux, comme la régulation de la transcription de l'ADN, la transduction du signal ou encore la réparation de l'ADN.

La réversion de la méthylation des arginines n'est pas encore très claire, en revanche la déimination des arginines mono-méthylées par une enzyme appelée peptidyl-arginine déiminase 4 (PADI4 ou PAD4) est bien connue ⁽³⁰⁻³¹⁾. Il existe cinq PAD (PAD1-5) et seule PAD4 est localisée dans le noyau ⁽³²⁾. L'équipe de Kouzarides⁽³⁰⁾ a montré que PAD4 convertit des résidus arginines non ou mono-méthylés des arginines 2, 8, 17 et 26 des histones H3 et H4 en citrulline dans la régulation de la transcription d'un gène œstrogène-sensible ⁽³¹⁾ (figure 1.11).

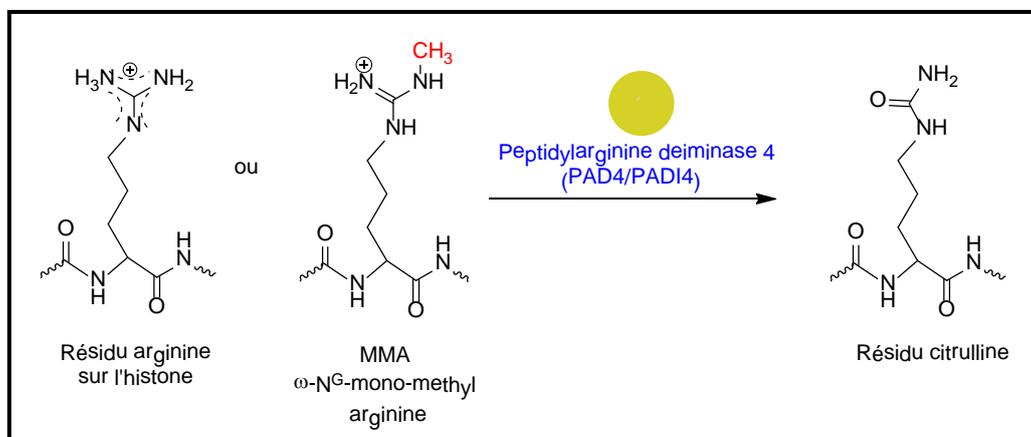


Figure 1.11 : Réaction de déimination sur les résidus arginine non ou mono-méthylée des histones catalysée par PADI4/PAD4 pour former un résidu citrulline.

La déimination des arginines mono-méthylées d'histones en citrulline par PAD4 n'est pas réellement un processus réversible car la citrulline est un acide aminé non naturel qui ne peut pas remplacer l'arginine ⁽³³⁾. Par conséquent, la citrulline serait ensuite recyclée par des aminotransférases inconnues ou l'histone serait remplacée ⁽³⁴⁾.

II. Les modifications épigénétiques et le cancer

La plupart des enzymes responsables des modifications épigénétiques ont été pour la plupart identifiées et caractérisées. Ces modifications ont été largement appréhendées pour comprendre leurs rôles dans l'expression des gènes, la croissance des cellules, la différenciation, l'apoptose ou encore les processus pathophysiologiques. Ces derniers impliquent les cancers et les trois classes de gènes qui favorisent leur développement sont les suivantes :

- **Les gènes oncogènes** : ils sont impliqués dans la division des cellules et ont pour fonction d'accélérer leur prolifération. Si l'un des deux allèles de ces gènes est modifié, alors ces derniers deviennent hyperactifs et augmentent la division cellulaire. Ces gènes jouent un rôle majeur dans la croissance de tumeurs.
- **Les gènes « suppresseurs de tumeurs »** : ils freinent la division cellulaire. S'ils subissent une modification, alors les cellules se multiplieront de manière incontrôlée.
- **Les gènes de « réparation de l'ADN »** ou aussi appelés « **gènes de surveillance** » : ils permettent la détection et la réparation des erreurs qui ont eu lieu lors de la réplication de l'ADN. Si ces gènes sont mutés, alors ils deviennent inactifs : la probabilité d'activer des oncogènes ou d'inactiver des gènes suppresseurs de tumeurs devient alors élevée.

II.1 Développement des cellules cancéreuses du à des modifications épigénétiques

II.1.1 L'identité d'une cellule

L'identité d'une cellule est directement dépendante du programme génétique et épigénétique. Elle dépend donc de son histoire et de son environnement.

II.1.1.1 Cellules cancéreuses

Un certain nombre de gène en général exprimés lors de l'antogénèse et réprimés peuvent être « réactivés » et être à l'origine de cette « cancérisation ». Par ailleurs d'autres gènes capables de bloquer le développement tumoral et de ce fait appelés gènes suppresseur de tumeurs sont susceptibles d'être « réprimés ».

II.1.1.2 Mécanismes épigénétiques dans les cellules cancéreuses

Dans les cellules cancéreuses, deux phénomènes de méthylation peuvent être observés au niveau de l'ADN : l'hyperméthylation locale et l'hypométhylation globale.

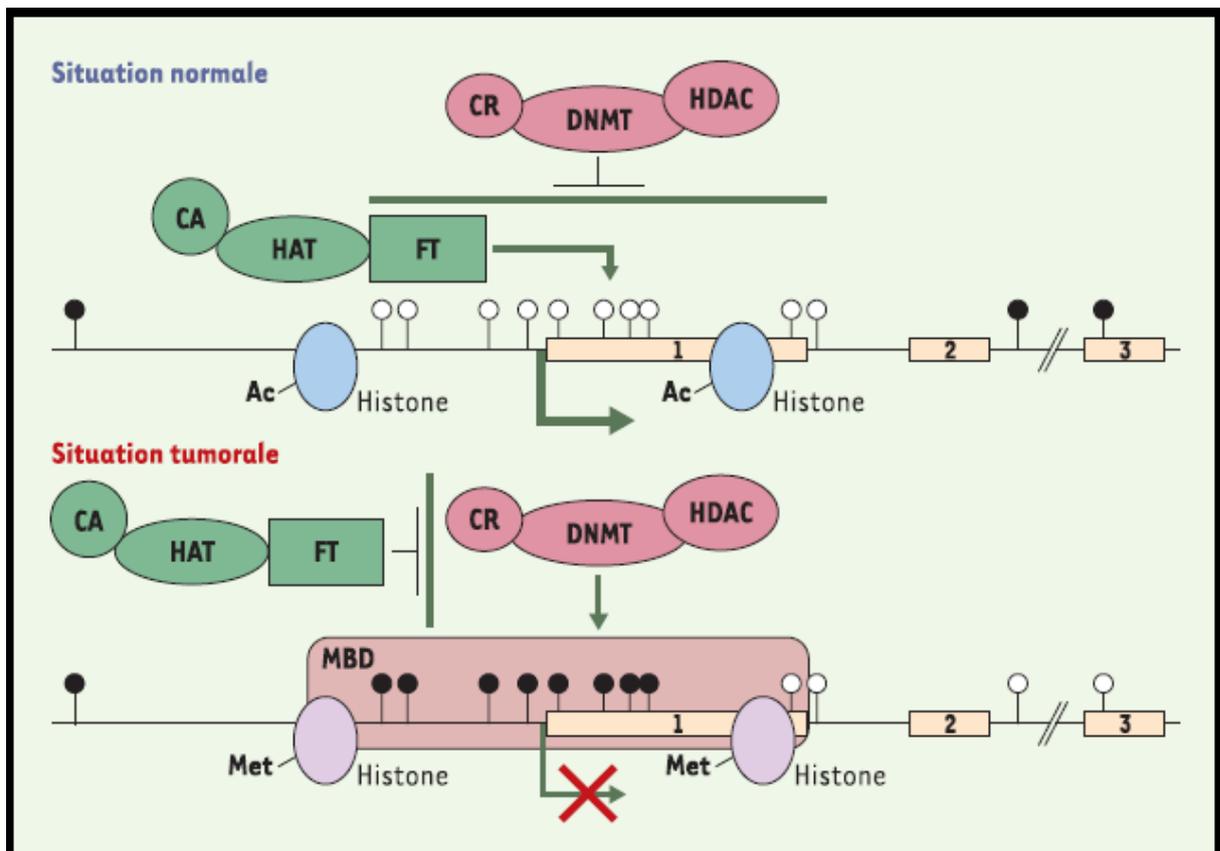
Dans une situation normale, la plupart des CpG qui ne sont pas contenus dans des îlots sont méthylés (cercles noirs), tandis que les CpG, à l'intérieur des îlots, ne sont pas méthylés (cercles blancs) (figure 1.12). Dans cette région promotrice, l'acétylation des histones conduit à un état d'euchromatisation qui ouvre l'accès aux trois complexes de transcription suivants : le facteur de transcription (FT), les protéines acétylantes (HATs) et les co-activateurs transcriptionnels (CA).

Dans une situation tumorale, les profils de méthylation sont inversés : les CpG répartis le long du génome et dans les régions codantes subissent une hypométhylation tandis que les îlots CpG au niveau des promoteurs sont hyperméthylés (figure 1.12). Les protéines contenant un domaine *Methyl-CpG Binding Domain* (MBD) sont capables de se lier aux régions ou aux

îlots CpG méthylés pour permettre le recrutement d'autres protéines capables de modifier les protéines régulatrice de la transcription comme :

- Les enzymes qui provoquent des modifications post-traductionnelles au niveau des histones : les HDACs, les Histones MéthylTransférases HMTs (HKMTs et PRMTs)
- Les protéines responsables de la méthylation de l'ADN, les DNMTs
- Les corépresseurs transcriptionnels (CR)

Dans cette situation tumorale et en considérant l'expression des gènes « supresseurs de tumeurs », cette zone chromatine peut être modifiée chimiquement par ces trois types d'enzymes, provoquant l'hétérochromatisation de la chromatine. Cette zone devient alors



inaccessible aux complexes de transcription, empêchant ainsi l'expression du gènes⁽³⁵⁾.

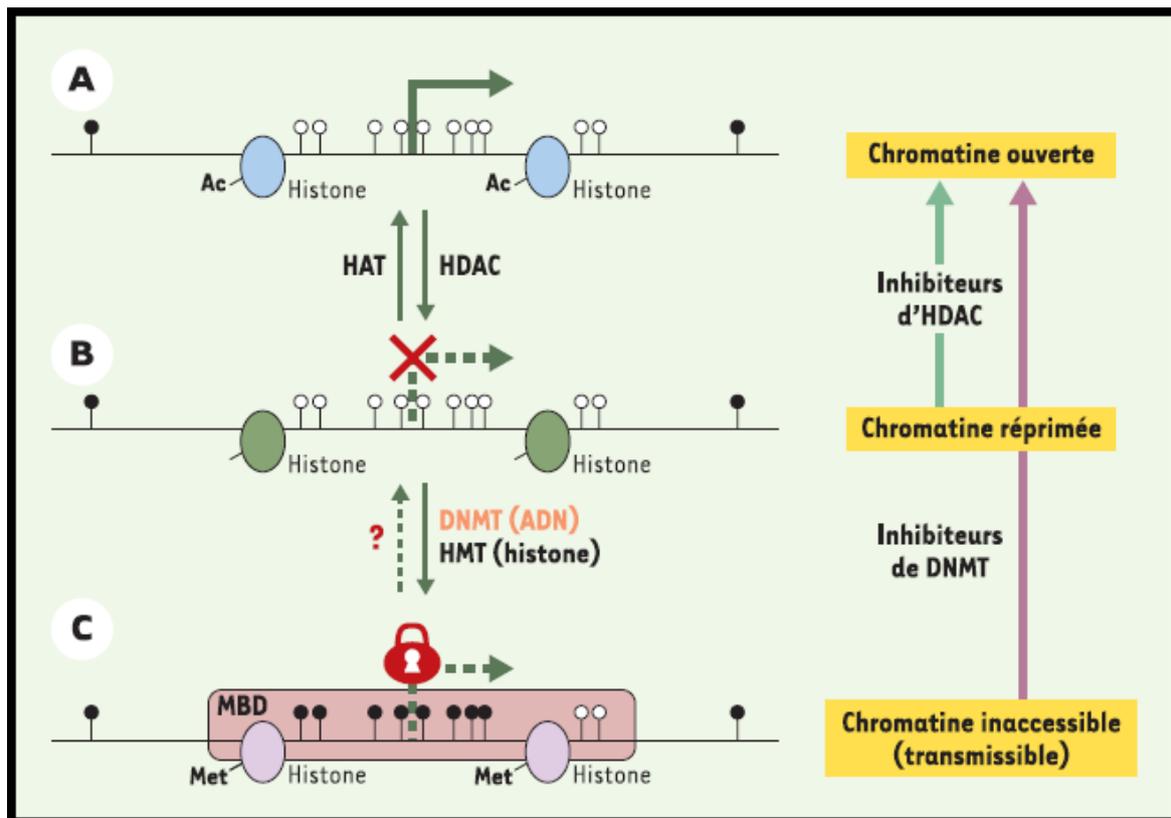
Figure 1.12 : Etats de la chromatine au niveau des « gènes supresseurs de tumeurs » dans une situation normale et tumorale⁽³⁵⁾.

Un des buts spécifiques de la thérapie épigénétique est la restauration du profile de la méthylation de l'ADN ou des histones.

II.2 La thérapie épigénétique

Contrairement aux facteurs génétiques du cancer qui affectent la séquence de l'ADN, les modifications épigénétiques sont réversibles : des perspectives en thérapie ont pu être ainsi envisagées. Dans le cadre de la thérapie épigénétique antitumorale, l'utilisation combinée d'agents modulant l'acétylation des histones et la méthylation de l'ADN réactivent la transcription des gènes silencieux capables de bloquer l'activité tumorale.

A titre d'exemple, les inhibiteurs des HDACs réactivent la transcription de gènes dont les promoteurs ne sont pas méthylés, mais pas celle des gènes dont les promoteurs sont hyperméthylés. En revanche, les inhibiteurs de DNMTs permettent de déverrouiller la transcription en inhibant l'enzyme responsable de la méthylation de l'ADN et en empêchant donc la méthylation sur les îlots CpG de la région promotrice ; ces agents permettent de



conduire ces îlots à un état d'euchromatisation et de réexprimer des gènes. Les inhibiteurs de certaines histones méthyltransférases (HMTs) permettent également d'inhiber les enzymes permettant la réexpression de gènes suppresseur de tumeur (figure 1.13).

Figure 1.13: Thérapie épigénétique⁽³⁵⁾. **a)** L'acétylation des histones permet l'ouverture de la chromatine. **b)** La désacétylation des histones par les HDACs conduit à une chromatine fermée : les inhibiteurs de ces enzymes vont donc rétablir l'état de la chromatine initiale. **c)** La méthylation des histones et de l'ADN par les DNMTs et les HMTs conduit à une chromatine inaccessible : les inhibiteurs de ces enzymes ont pour rôle de rétablir l'état de la chromatine.

Pour certains inhibiteurs, notamment pour les HDACs et les DNMTs, des études et des essais cliniques ont été menés. Dans le paragraphe suivant, nous nous intéresserons uniquement aux inhibiteurs des méthyltransférases de l'ADN et des histones.

II.3 Les inhibiteurs des cibles épigénétiques

Commençons, dans ce paragraphe, par l'étude des inhibiteurs des méthyltransférases de l'ADN.

II.3.1 Inhibiteurs des DNMTs

Les inhibiteurs des DNMTs peuvent être classés en deux groupes : les inhibiteurs nucléosidiques et les inhibiteurs non-nucléosidiques⁽³⁶⁻³⁷⁾. Pour chacune des deux classes, quelques composés seront décrits dans le tableau **1.3**.

II.3.1.1 Inhibiteurs nucléosidiques des DNMTs

Les deux inhibiteurs nucléosidiques les plus utilisés possèdent une cytosine modifiée directement liée à un ribose : il s'agit des composés 5-azacytidine (5-Aza-CR) et 5-aza-2'-désoxycytidine (5-Aza-CdR). D'autres agents peuvent aussi présenter une cytosine mutée qui est, cette fois-ci, liée à un déoxyribose : parmi eux, ce sont le 5-fluoro-2'-désoxycytidine et la zébularine.

Lors d'une modification épigénétique au niveau de l'ADN, la méthylation de la position 5 de la cytosine se fait suivant un processus séquentiel. Dans un premier temps, un groupement carboxylate d'un résidu du site actif d'une DNMT permet d'ancrer les amines greffées sur la cytosine, sauf pour la zébularine. La double liaison en position 5-6 de la cytosine présente ainsi un caractère électrophile ce qui permet à une cystéine présente dans le site actif de l'enzyme de réagir sur cette liaison. Par conséquent, le transfert du groupe méthyle de la SAM vers le carbone en position 5 de la cytosine est alors possible. Finalement,

l'enzyme libre est éjectée du substrat par une réaction de β -élimination et peut ainsi méthyle un second site de CpG (figure 1.14) ⁽³⁶⁾.

Dans la thérapie épigénétique, les composés nucléosidiques inhibent les DNMTs en formant un complexe covalent avec ces dernières, empêchant ainsi l'éjection de l'enzyme hors du substrat qui ne pourra plus méthyle un second site de CpG.

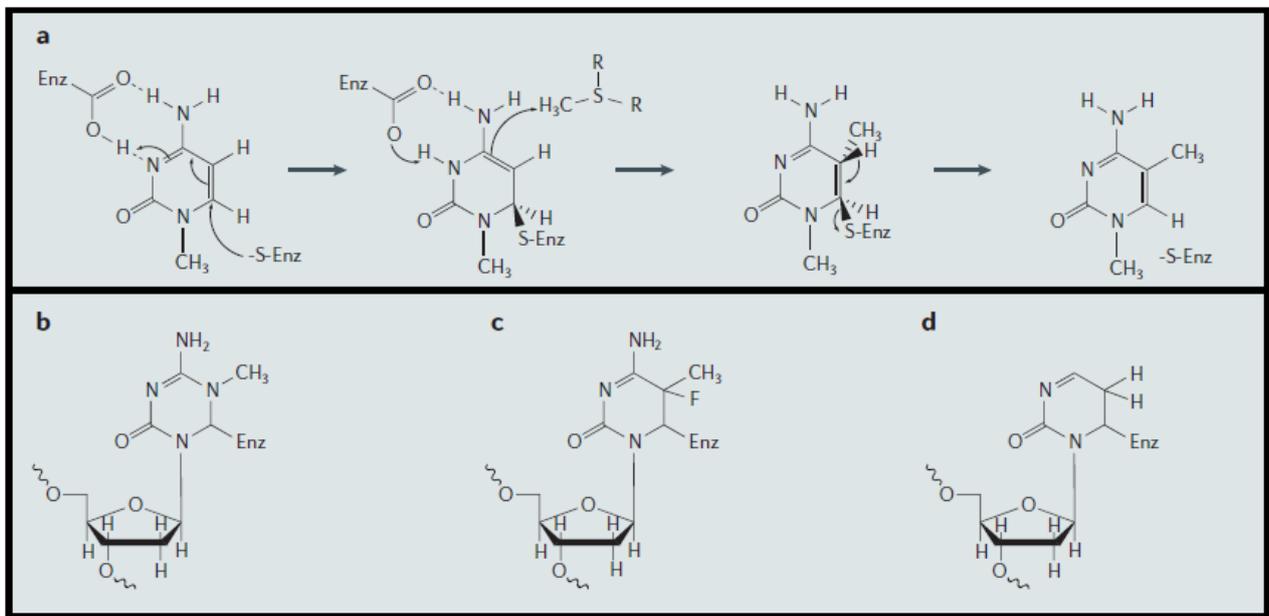


Figure 1.14 : Mécanisme d'inhibition des enzymes DNMTs par les analogues nucléosidiques ⁽³⁶⁾. **a)** Mécanisme de transfert de méthyle *via* la SAM catalysé par les DNMTs. **b)** Complexe covalent formé entre l'enzyme et la 5-Aza-CR ou la 5-Aza-CdR. **c)** Complexe covalent formé entre l'enzyme et la 5-fluoro-2'-desoxycytidine. **d)** Complexe covalent formé entre l'enzyme et la zébularine.

L'analogue 5-azacytidine possède un autre mécanisme d'inhibition : en effet, il est cytotoxique et provoque ainsi le réarrangement chromosomique ⁽³⁸⁻³⁹⁾. Cet inhibiteur a été approuvé par la FDA (pour *food and drug administration*) pour le traitement des syndromes de myélodysplasie. Il est utilisé cliniquement contre le traitement de tumeurs solides ⁽⁴⁰⁾. Ses principaux inconvénients sont sa courte durée de demi-vie et sa faible biodisponibilité ⁽⁴¹⁾.

En revanche, la zébularine présente un temps de demi-vie plus élevé et est moins toxique pour les cellules cancéreuses du sein. Elle est donc plus adaptée aux thérapies épigénétiques.

II.3.1.2 Les inhibiteurs non-nucléosidiques des DNMTs

Les inhibiteurs non-nucléosidiques sont de petites molécules capables d'inhiber la méthylation de l'ADN en se liant au site catalytique des enzymes des DNMTs⁽⁴²⁾.

Le composé _L-RG108 a été synthétisé par le groupe de Lyko et *al.*⁽⁴³⁾ et a montré une inhibition efficace pour les DNMTs *in vitro*. De plus, à de faibles concentrations de l'ordre du micromolaire, il inhibe encore l'enzyme en ne provoquant aucune cytotoxicité dans les cellules cancéreuses du colon (HCT116) (tableau **1.3**).

Décrit par le groupe de Yang et *al.*, l'épigallocatechin-3-gallate (EGCG) appartient à la famille des polyphénols. Cet inhibiteur a deux mécanismes d'action : il inhibe l'enzyme DNMT et inhibe aussi la prolifération de cellules cancéreuses à de faibles concentrations, de l'ordre du micromolaire⁽⁴⁴⁾. Pour finir, le composé EGCG est testé en phase I (tableau **1.3**).

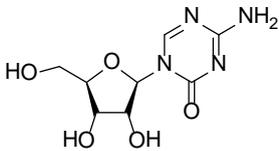
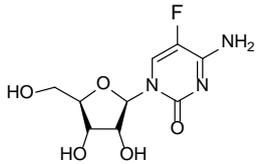
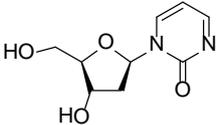
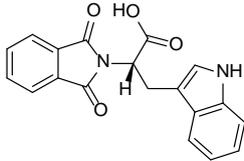
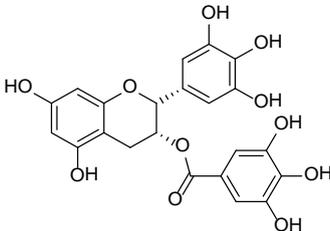
Composés	Structures	Concentrations d'inhibition	Phases d'essais
5-Aza-CR		De l'ordre du μM	FDA
5-fluoro-2'-désoxycytidine		De l'ordre du μM	Clinique
Zébularine		Doses comprises entre le μM et le mM	Pré-clinique
L-RG108		De l'ordre du μM	Pré-clinique
EGCG		De l'ordre du μM	Phase I

Tableau 1.3 : Liste non exhaustive des inhibiteurs nucléosidiques et non-nucléosidiques des DNMTs ⁽⁴⁵⁾.

Récemment, une série d'inhibiteurs des DNMTs a été identifiée par le groupe de Arimondo et *al.*⁽⁴⁶⁾ Ces inhibiteurs appartiennent à la famille des procainamides et ont montré un effet d'inhibition pour les enzymes DNMT3A/3L et DNMT1. Six de ces produits se présentent comme des inhibiteurs potentiels : à titre d'exemple, le composé **12** montre une faible inhibition pour hDNMT1 avec une valeur de CI_{50} égale à environ $5 \mu\text{M}$. De plus, une bonne sélectivité par rapport aux bactéries méthyltransférase d'ADN d'*Escherichia coli*

(EcoDam) et une lysine méthyltransférase (G9a) a été observée avec cet inhibiteur **12**. Cependant, à des concentrations de l'ordre du micromolaire, il provoque des effets cytotoxiques pour une lignée de cellules cancéreuses du colon HCT116 mais aussi pour une lignée de cellules tumorales de la prostate DU145.

II.3.2 *Les inhibiteurs des HKMTs*

Plusieurs enzymes HKMTs sont impliquées dans des cancers⁽⁴⁷⁻⁴⁸⁾ : à titre d'exemple, la protéine Ezh2 (pour *Enhancer of zeste homolog 2*) est surexprimée dans les cancers du sein et de la prostate, ainsi que dans plusieurs myélomes et lymphomes⁽⁴⁹⁻⁵²⁾. De ce fait, plusieurs inhibiteurs ont été développés et quelques composés sont décrits dans le tableau **1.4**. A l'heure actuelle, aucun de ces inhibiteurs sont en phase d'essais cliniques.

En 2005, le groupe de Imhof et *al.*⁽⁵³⁾ a décrit le premier inhibiteur, la chaetocin, qui permet d'inhiber l'activité des HKMTs : il agit comme un compétiteur de la SAM.

Le second produit BIX-01294, décrit en 2007 par Jenuwein et *al.*⁽⁵⁴⁾, inhibe de manière spécifique l'enzyme G9a mais n'est pas compétiteur avec la SAM. Un nouvel inhibiteur, le BIX-01338, a été synthétisé par la même équipe et est capable d'inhiber deux HKMTs, les SuV39h1 (pour *Suppressor of variegation 3-9 homolog 2*) et G9a avec des valeurs respectives de 1,1 μM et de 4,7 μM . Ce dernier composé n'est pas spécifique et pourrait être utilisé comme un pan-inhibiteur des HKMTs.

Finalement, en 2011 l'équipe de Song et *al.*⁽⁵⁵⁾ a découvert un inhibiteur de Dot1L, basé sur le motif de la SAM qui est capable d'inhiber l'activité de cette enzyme à une concentration de l'ordre du nanomolaire ; ce composé se montre aussi sélectif par rapport aux PRMTs (PRMT1 et CARM1).

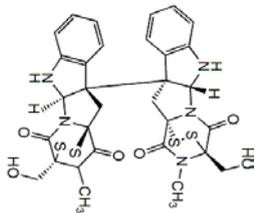
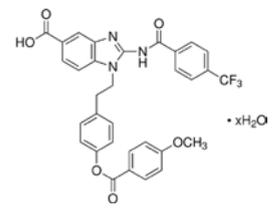
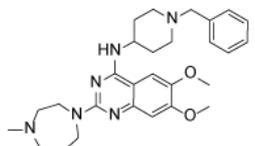
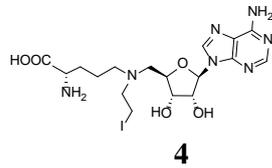
Composés	CI ₅₀ (μM) : Cibles	Références	Composés	CI ₅₀ (μM) : Cibles	Références
 Chaetocin	0,8 : SuV39h1	53	 BIX-01338	4,7 : G9a 1,1 : SuV39h1	54
 BIX-01294	1,7 : G9a	54	 4	0,038 : Dot1	55

Tableau 1.4 : Inhibiteurs des HKMTs

II.3.3 Les inhibiteurs des PRMTs

Les PRMTs représentent des cibles prometteuses pour le développement d'agents anticancéreux, comme cela est illustré par une récente revue écrite par Jin et *al.*⁽⁵⁶⁾ De nombreuses équipes de recherche se sont intéressées à la synthèse d'inhibiteurs des PRMTs. Les premières recherches concernant ces composés se sont tout d'abord focalisées sur PRMT1 et trois stratégies ont été employées. Ces dernières seront toutes abordées dans les paragraphes suivants et nous allons débiter par le criblage d'inhibiteurs par approche *in silico*.

II.3.3.1 Découverte d'inhibiteurs de PRMT1 par approche in silico

En 2004, les premiers travaux publiés ont été réalisés par l'équipe de Bedford et *al.*⁽⁵⁷⁾ : elle a décrit un composé appartenant à la famille des naphtyles sulfonates (NS), capable d'inhiber PRMT1 à des doses de l'ordre du μM .

Bien plus tard, l'équipe de Finn et *al.*⁽⁵⁸⁾ s'est intéressée à la sélectivité de ce composé envers d'autres PRMTs et ont trouvé un facteur de sélectivité égale à 19 avec CARM1 (tableau 1.5).

Concomitamment, d'autres groupes se sont intéressés à la synthèse d'inhibiteurs de PRMT1 comme l'équipe de Jung et *al.*⁽⁵⁹⁾ qui a développé des dérivés du benzène sulfonamide (BS) ; le meilleur de ces composés a présenté une valeur de CI_{50} égale à environ 2 μM (tableau 1.5).

Le groupe de Sbardella et *al.*⁽⁶⁰⁾ a finalement synthétisé des dérivés xanthénones, puis des dérivés naphtyles urées (NU)⁽⁶¹⁾ comme le composé **1c** capable d'inhiber l'activité de PRMT1 avec une concentration de l'ordre du micromolaire. Ce dérivé **1c** est aussi cinq fois plus sélectif par rapport à CARM1.

La structure de ces composés est représentée dans le tableau 1.5. D'autres inhibiteurs des PRMTs ont été aussi ajoutés

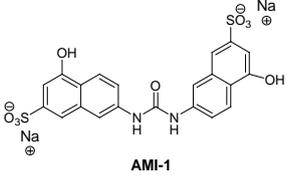
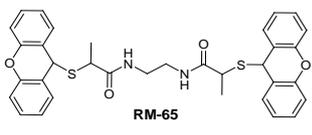
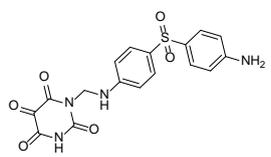
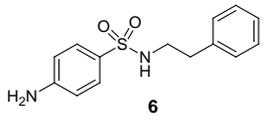
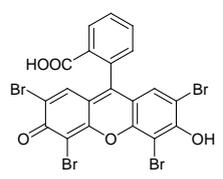
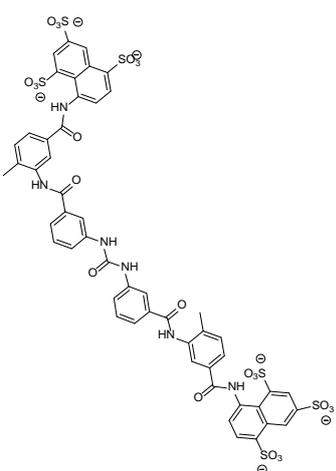
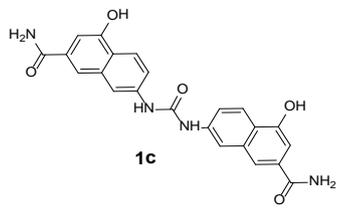
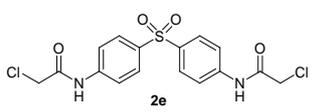
Composés	CI ₅₀ PRMT1 (μM)	Références	Composés	CI ₅₀ PRMT1 (μM)	Références
 <p>AMI-1</p> <p>Naphtyle sulfonate</p>	8,8	57	 <p>RM-65</p> <p>Amide thioglycolique</p>	55,4	62
 <p>Benzène sulfonamide</p>	1,7	59	 <p>6</p> <p>Benzène sulfonamine</p>	12,8	63
 <p>Xanthénone</p>	1,4	60	 <p>Suramine</p>	5,3	64
 <p>1c</p> <p>Naphtyle urée</p>	9,3	61	 <p>2e</p> <p>Benzène sulfonamide</p>	1,5	65

Tableau 1.5: Inhibiteurs de PRMT1 découverts *in silico* et par criblage virtuel.

II.3.3.2 Approche rationnelle pour le développement d'inhibiteurs de PRMT1 compétitifs du cofacteur

La seconde stratégie est basée sur une approche rationnelle où l'inhibiteur est en compétition avec le cofacteur SAM. Deux équipes se sont intéressées à la synthèse d'inhibiteurs qui s'inspirent du motif de la SAM. En 2008, la première équipe était celle de Thompson et *al.*⁽⁶⁶⁾ : elle a décrit des analogues azotés de la SAM (**AAI**) possédant un groupement électrophile capable de réagir *in-situ* avec l'arginine d'un peptide qui mime le substrat (histone H4). Ce dérivé **AAI** est capable d'inhiber PRMT1 avec une valeur de CI_{50} égale à 18,5 μ M.

D'autres analogues azotés de la SAM ont été décrits en 2010 et 2011 par Dowden et *al.*⁽⁶⁷⁻⁶⁸⁾ et possèdent un fort potentiel d'inhibition avec des valeurs de CI_{50} égales à environ 4 μ M. Ces composés se rapprochent de l'état de transition de la réaction entre la SAM et son substrat. Les produits **17** à **19** sont sélectifs par rapport à CARM1 : le composé **17** montre une faible inhibition (supérieure à 100 μ M) tandis que les analogues **18** et **19** ne présentent aucune inhibition.

Dans le tableau **1.6**, pour comparer les valeurs de CI_{50} , les structures de ces inhibiteurs sont indiquées ainsi que celles de la SAM, du SAH et de la SFG⁽⁶⁶⁻⁶⁷⁾; cette dernière est un inhibiteur puissant mais non sélectif des méthyltransférases.

Composés	CI ₅₀ PRMT1 (μM)	Références	Composés	CI ₅₀ PRMT1 (μM)	Références
<p>SAM</p>	15	66	<p>16</p>	2,9	67
<p>SAH</p>	8,3	66	<p>17 : 2,9 18 : 5,6</p>		68
<p>AAI</p>	18,5	66			
<p>SFG</p>	0,7	67	<p>19</p>	3,9	68

Tableau 1.6 : Inhibiteurs de PRMT1, compétitifs de la SAM

II.3.3.3 Approche rationnelle pour le développement d'inhibiteurs de PRMT1 spécifiques du substrat

La troisième stratégie est également basée sur une approche rationnelle, où la structure de l'inhibiteur est inspirée de celle du substrat de l'enzyme. En effet, les équipes de

Thompson et de Martin, puis celle de Frankel et *al.* ont décrit des inhibiteurs de PRMT1, sélectifs par rapport à d'autres PRMTs, notamment CARM1.

En 2010, les équipes de Martin et Frankel⁽⁶⁹⁾ ont synthétisé l'inhibiteur **R1-4** : celui-ci possède une arginine ramifiée et un peptide mimant la séquence du substrat fibrillarine. Ce composé inhibe PRMT1 avec une valeur de CI_{50} égale à 14 μ M et il est douze fois sélectif par rapport à CARM1. Puis, en 2011, ce même groupe de chercheurs a essayé d'optimiser ces résultats, en ajoutant sur l'arginine une chaîne mimant la partie acide aminée de la SAM. Mais cette modification n'a ni conduit à une meilleure inhibition pour PRMT1, ni à une meilleure sélectivité envers CARM1 puisque seul un facteur de sélectivité égal à 2,5 a été déterminé⁽⁷⁰⁾.

Pour finir en 2011, l'équipe de Thompson⁽⁷¹⁾ a récemment publié un inhibiteur, le **C21**, qui possède un peptide mimant la séquence de l'histone H4 ; au sein de ce peptide, le groupement guanidine, initialement présent sur l'arginine en position 3, a été remplacé par un groupement chloroacétimide. Ce composé a été testé et présente une forte inhibition pour PRMT1, ce qui est traduit par une valeur de CI_{50} égale à 1,8 μ M. Ce dérivé est aussi 500 fois sélectif envers CARM1.

Les structures des composés issues de cette troisième stratégie sont représentées dans le tableau **1.7**.

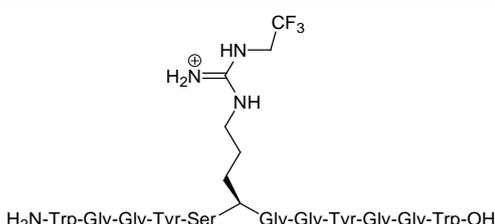
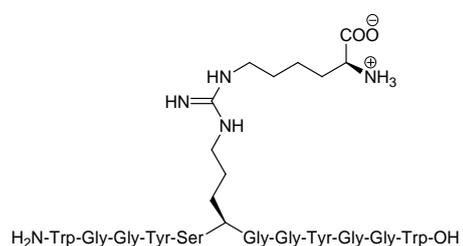
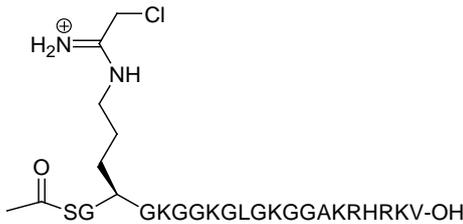
Composés	CI ₅₀ PRMT1 (μM)	Références
 <p>R1-4</p>	13,9	69
 <p>R1-Lys 18</p>	13,9	70
 <p>C21</p>	1,8	71

Tableau 1.7 : Inhibiteurs de PRMT1, substrats spécifiques.

De nombreux composés ont ainsi été décrits pour inhiber l'activité de PRMT1. PRMT4 (ou CARM1) est une autre enzyme responsable de la méthylation des histones et pourrait être une nouvelle cible thérapeutique. La recherche est donc très active pour synthétiser des inhibiteurs sélectifs et spécifiques de CARM1. C'est à cette enzyme que nous nous intéresserons au cours de mes travaux de thèse.

III. Etude du Coactivator Associated aRginine Methyltransferase 1

III.1 Présentation de CARM1

Le Coactivator Associated aRginine Methyltransferase 1 (CARM1) a été identifié par la technique de double-hybride, un procédé biologique qui met en évidence les interactions directes entre deux protéines ; en effet, CARM1 interagit avec la protéine GRIP1 (pour *Glucocorticoid Receptor Interacting Protein 1*)⁽⁷²⁾. Cette dernière appartient à la famille des co-activateurs « p60 » et est impliquée dans le mécanisme d'activation de la transcription dépendante des récepteurs nucléaires (RN). En effet, ces récepteurs sont activés par une hormone spécifique et se lient à un élément de réponse aux hormones (HRE) en amont d'un gène cible. Ce processus permet le recrutement d'un ensemble de coactivateurs, agissant sur le nucléosome (comme par exemple la méthylation de l'histone H3 par la protéine CARM1), et permettant l'assemblage de la machinerie transcriptionnelle au niveau du promoteur du gène⁽⁷³⁻⁷⁶⁾.

En effet, en 1999, CARM1 a été la première PRMT identifiée en tant qu'enzyme méthyltransférase spécifique d'histone⁽⁷²⁾, méthylant *in vitro* les arginines R2, R17 et R26 de l'histone H3, R17 étant le site principal. La méthylation de H3R17 est généralement associée à la fonction coactivatrice de CARM1, en réponse aux récepteurs nucléaires et à d'autres facteurs de transcription. CARM1 est capable de méthyler d'autres protéines qui possèdent d'autres substrats que l'histone H3, comme par exemple l'ARN polymérase II (figure 1.15)⁽⁷⁷⁾.

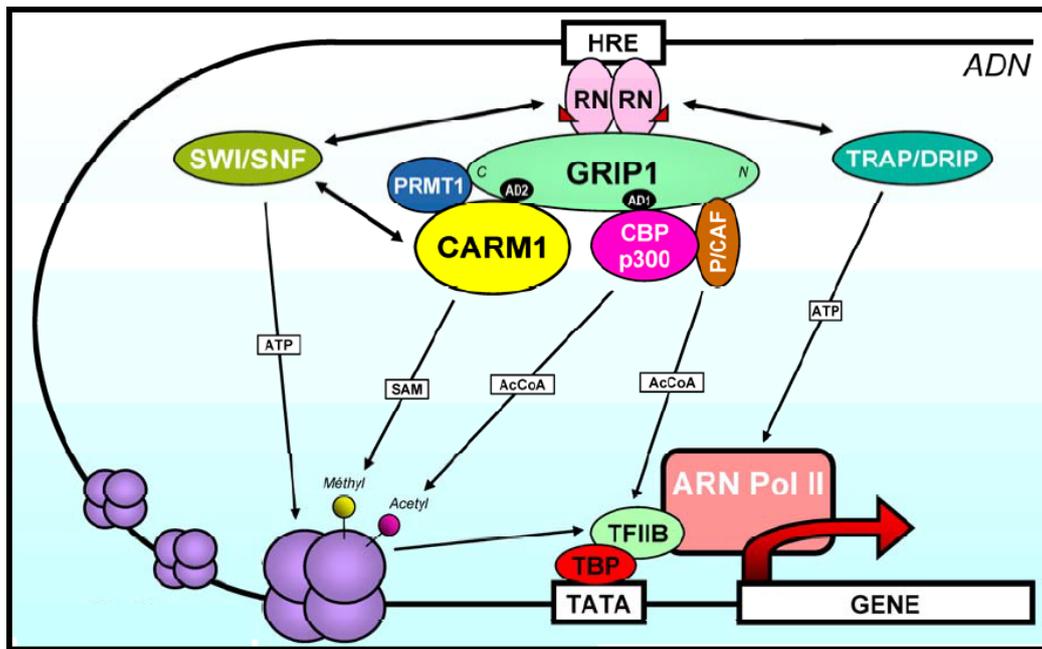


Figure 1.15 : Implication de CARM1 dans le remodelage de la chromatine avec la méthylation de l'histone H3 et dans l'activation de la transcription en tant que co-activateur de la protéine GRIP1 ⁽⁷⁷⁾.

En outre, plusieurs études ont démontré l'implication de CARM1 dans certains cancers. La surexpression conjuguée de cette enzyme est observée dans les tumeurs du cancer du sein⁽⁷⁸⁾. L'inactivation par mutation de CARM1 affecte l'expression des gènes qui est stimulée par les oestrogènes, la progression du cycle cellulaire et la croissance des cellules du cancer du sein^(77, 79). Des concentrations élevées de CARM1 dans les cellules sont observées dans le développement de carcinomes de la prostate⁽⁸⁰⁾. Cette enzyme constitue donc une cible potentielle en chimiothérapie. La découverte d'inhibiteurs spécifiques pourrait ouvrir des voies thérapeutiques nouvelles. Par conséquent, la compréhension à l'échelle atomique du mode d'action de CARM1 s'avère cruciale tant au niveau fondamental que pour des applications pharmacologiques.

III.1.1 Architecture de CARM1

CARM1 contient 608 acides aminés chez l'homme et la souris et son architecture est subdivisée en 3 domaines distincts présentés sur la figure 1.16⁽⁸¹⁾.

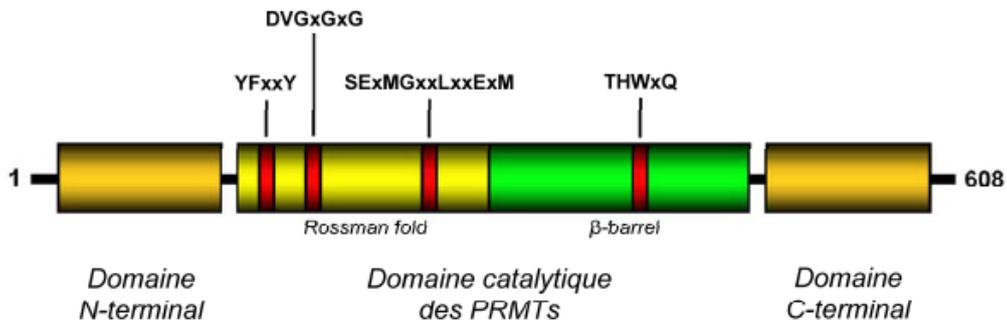


Figure 1.16 : Représentation schématique de l'architecture de CARM1⁽⁸¹⁾ : 608 acides aminés chez l'homme et la souris répartis dans un domaine catalytique (en jaune et vert). En rouge, sont représentés les motifs signatures des PRMTs et aux extrémités, les deux domaines additionnels de CARM1.

CARM1 est construit autour d'un domaine catalytique central qui est une zone contenant les résidus 130 à 507 chez la souris, appelée mCARM1. Cette séquence de résidus est fortement conservée pour toutes les enzymes de la famille des PRMTs. Ce domaine est subdivisé en deux régions structurellement différentes : la première est composée d'un repliement de type Rossman fold (domaine classique de fixation de nucléotides) à l'extrémité N-terminale et la deuxième est constituée d'un tonneau β , du côté C-terminal. Le domaine de type Rossman fold contient les 4 motifs signatures de la famille des PRMTs qui sont impliqués dans la formation du site catalytique.

III.1.2 Structure tridimensionnelle de CARM1

Dans sa forme apo, c'est-à-dire en absence de ligand, les cristaux de CARM1 sont en forme de livre ouvert, ce qui est caractéristique de la formation de cristaux du domaine catalytique de mCARM1₁₄₀₋₄₈₀ (c'est-à-dire les résidus d'acides aminés 140 à 480 de CARM1

chez la souris) (figure 1.17 A). Lorsque CARM1 est en complexe avec l'un de ces ligands, notamment la SAH ou la SAM, le cristal présente la forme indiquée ci-dessous (figures 1.17 B et C). La taille moyenne des cristaux de mCARM1₁₄₀₋₄₈₀ est de 200 x 120 x 40 μm ⁽⁸²⁾.

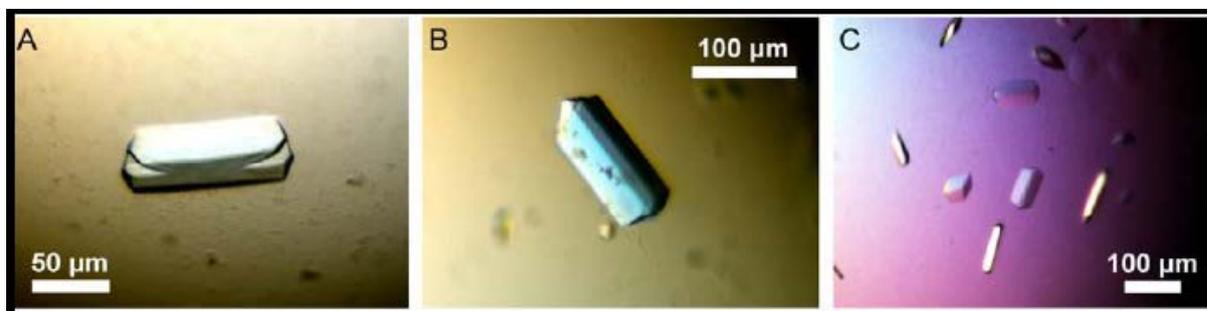


Figure 1.17 : Cristaux de mCARM1₁₄₀₋₄₈₀ ; A : Cristal apo en forme de livre ouvert typique des cristaux du domaine catalytique de CARM1 en PEG ; B : Cristal avec la SAH ; C : Cristaux avec la SAM.

La taille moyenne des cristaux de mCARM1₁₄₀₋₄₈₀ est de 200 x 120 x 40 μm ⁽⁸²⁾.

La structure RX de CARM1 a été obtenue concomitement par les équipes de Cavarelli⁽⁸³⁾ et de Pearl⁽⁸⁴⁾. Les structures obtenues avec les constructions mCARM1₁₃₀₋₅₀₇ et mCARM1₁₃₀₋₄₉₀ ont permis de montrer que CARM1 est sous forme de dimère dans lequel chaque monomère interagit avec l'autre grâce à un bras de dimérisation représenté en bleu sur la figure 1.18.

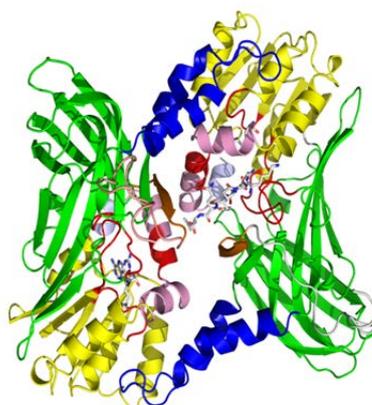


Figure 1.18 : Représentation du dimère du domaine catalytique sur le modèle des résidus 136 à 497 issu des structures cristallographiques de mCARM1₁₃₀₋₅₀₇ et mCARM1₁₃₀₋₄₉₀. Les grandes zones du domaine

catalytique de CARM1 sont les suivantes : le site catalytique avec la SAH, le domaine de type Rossmann fold en jaune, en vert le tonneau β qui contient le bras de dimérisation en bleu ⁽⁸²⁾.

Une seconde observation est le changement de conformation de la partie N-terminale de la protéine en présence de cofacteur.

Sur la figure 1.19, on peut observer un changement de conformation provoqué par la poche du cofacteur. Lorsque CARM1 est sous sa forme apo, l'extrémité N-terminale du domaine catalytique n'est pas repliée (structure visible à partir du résidu 155), alors qu'en présence du cofacteur, l'extrémité N-terminale du domaine catalytique forme une hélice supplémentaire (structure visible à partir du résidu 144) et permet de fermer la poche du site actif.

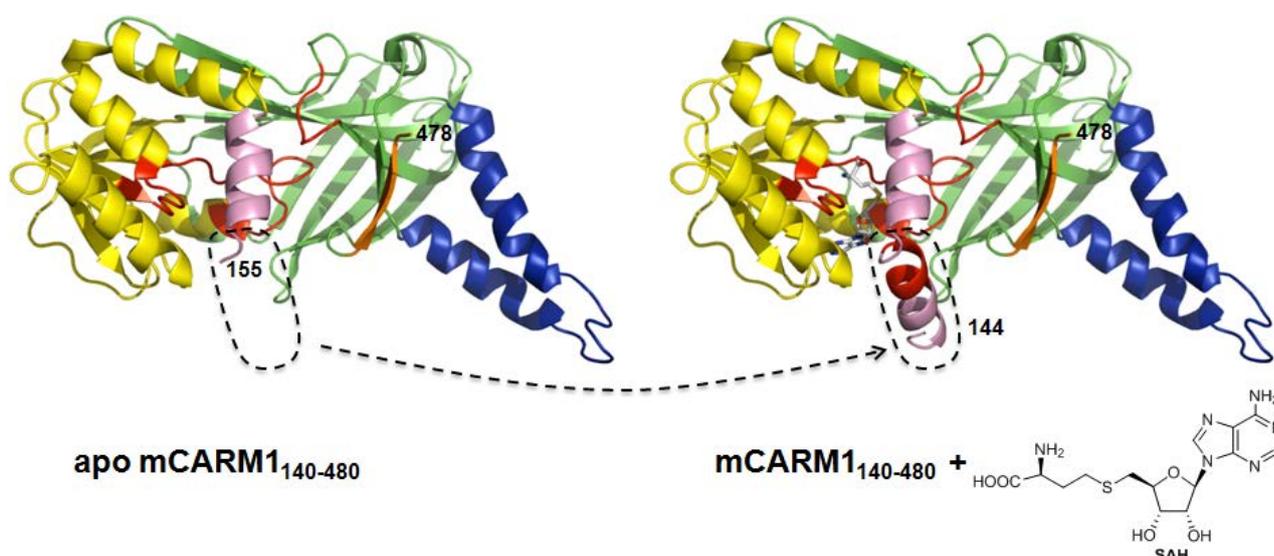


Figure 1.19: Structure cristallographique de la protéine CARM1 dans sa forme apo et en complexe avec un cofacteur, la SAH. **a)** L'enzyme dans sa forme apo-CARM1₂₈₋₅₀₇. **b)** L'enzyme s'organise en présence d'un cofacteur.

Les différentes structures cristallographiques obtenus ont également permis de mettre en évidence une organisation de CARM1 sous forme de gros oligomères, tels que des octamères ou des tétramères (figure 1.20) mais a aussi souligné l'importance de certains résidus du site catalytique de CARM1 (figure 1.21) ⁽⁸²⁾.

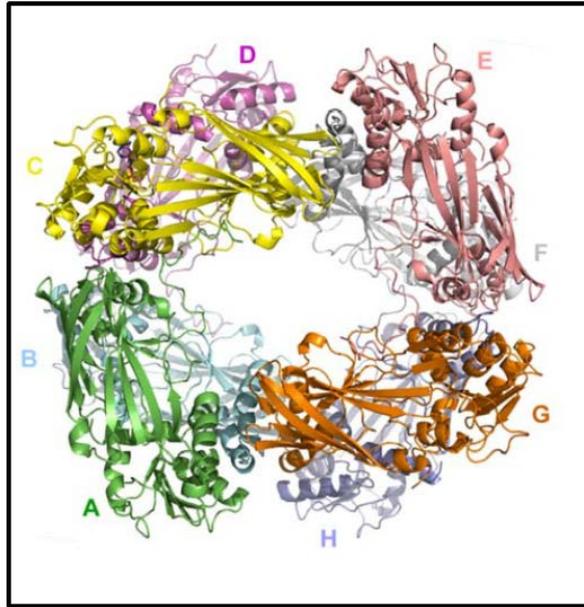


Figure 1.20 : Structure tridimensionnelle de l'enzyme CARM1 : la formation d'un octamère est observée dans les cristaux⁽⁸²⁾.

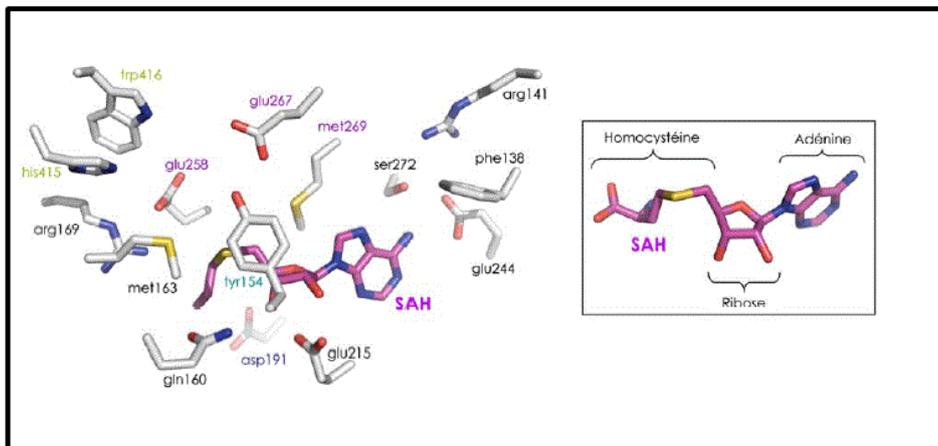


Figure 1.21 : Structure tridimensionnelle de CARM1 : focalisation sur le ligand SAH se liant au site catalytique de CARM1⁽⁸²⁾.

Maintenant que la structure cristallographique est connue, cette dernière aide à la conception d'inhibiteurs sélectifs de CARM1.

III.1.3 *Inhibiteurs de CARM1*

Seuls deux inhibiteurs de CARM1 ont été publiés au début de mes travaux de thèse, en 2008. Le premier a été décrit en 2007 par l'équipe de Conte⁽⁸⁵⁾ et correspond à un composé dérivant de la diphénylpentadione. Il inhibe CARM1 avec une valeur de CI_{50} égale à 28 μM et est 16 fois plus sélectif vis-à-vis de cette enzyme par rapport à PRMT1 (tableau **1.8**).

Le second inhibiteur publié dans la littérature a été synthétisé par le groupe de Purundare et *al.*⁽⁸⁶⁾, en 2008, et correspond au 1H-(pyrazol-1-yl)benzyl)propanamide (composé **7b**) ; ce dernier est capable d'inhiber CARM1 avec une valeur de CI_{50} égale à 0,08 μM mais aucune sélectivité n'a été montrée (tableau **1.8**).

En 2009, grâce à des essais de criblage à haut débit, ce même groupe élabore un meilleur inhibiteur sélectif de CARM1 : il fait partie de la famille des 1H-(pyrazol-1-yl)benzyl) propanamides et inhibe cette enzyme avec une valeur de CI_{50} égale à 0,04 μM ⁽⁸⁷⁾ (tableau **1.8**). Il est sélectif des PRMTs car il présente aussi une faible affinité pour PRMT1 (CI_{50} supérieure à 25 μM). A l'heure actuelle, il reste le meilleur composé. En 2011, une structure RX de cet inhibiteur, en complexe avec la SAH, a récemment été obtenue ; elle a montré qu'il fallait ajouter du SAH pour induire l'organisation dans le site de fixation de CARM1⁽⁸⁸⁾. Ce phénomène d'organisation avait déjà été observé par cristallographie en 2007 par les équipes de Cavarelli⁽⁸³⁾ et Pearl⁽⁸⁴⁾ lorsque que CARM1 était en présence d'un cofacteur, la SAH ou la SFG.

Par la suite, d'autres inhibiteurs ont été publiés, comme le Benzo[d]imidazole **17b**, décrit en 2009 par Wan et *al.*⁽⁸⁹⁾, qui inhibe CARM1 avec une CI_{50} de 0,07 μM . L'acide ellagique (TBBD), identifié et décrit en 2010 par Selvi et *al.*⁽⁹⁰⁾, présente une capacité à bloquer spécifiquement la méthylation de l'arginine 17 de l'histone H3 (tableau **1.8**).

En 2011, le groupe de Cheng et *al.*⁽⁹¹⁾ a décrit un inhibiteur sélectif de CARM1 qui possède un motif benzyl-3,5-dibenzylidènepiperidin-4-one ; ce composé a montré une bonne inhibition avec une valeur de CI_{50} de 8,6 μM pour CARM1, alors qu'il ne présente aucune affinité pour PRMT1 et une lysine méthyltransférase (SET7) ; cette mauvaise inhibition est traduite par des valeurs de CI_{50} supérieures à 667 μM .

Les structures de tous ces inhibiteurs de CARM1 sont représentées dans le tableau 1.8.

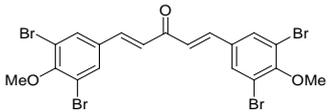
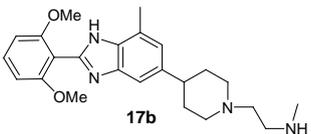
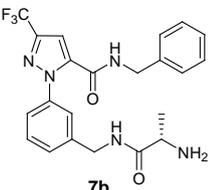
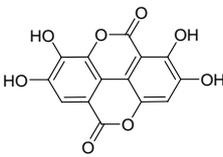
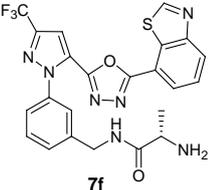
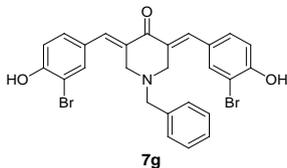
Composés	CI_{50} CARM1 (μM)	Références	Composés	CI_{50} CARM1 (μM)	Références
	27,9	85		0,07	89
Diphenylpentadienone			Benzo[d]imidazole		
	0,08	86		$K_d = 4$ (μM)/H3	90
1H-((pyrazol-1-yl) benzyl)propanamide			TBBD		
	0,04	87		8,6	91
1H-pyrazol-1- yl)benzyl)propanamide			Dibenzylidènepiperidinone		

Tableau 1.8 : Structures des inhibiteurs de CARM1 connus à ce jour dans la littérature.

Au cours de ma thèse, nous nous sommes intéressés au mécanisme d'action de l'enzyme CARM1 durant le processus de méthylation pour comprendre son mode de fixation avec les substrats. Nous allons maintenant aborder ce mécanisme dans le paragraphe suivant.

III.2 Notre stratégie pour étudier le mode de fixation de CARM1 avec ses substrats

A ce jour, le mécanisme d'action de l'enzyme CARM1, ainsi que son mode de reconnaissance avec ses substrats, sont inconnus (figure 1.22).

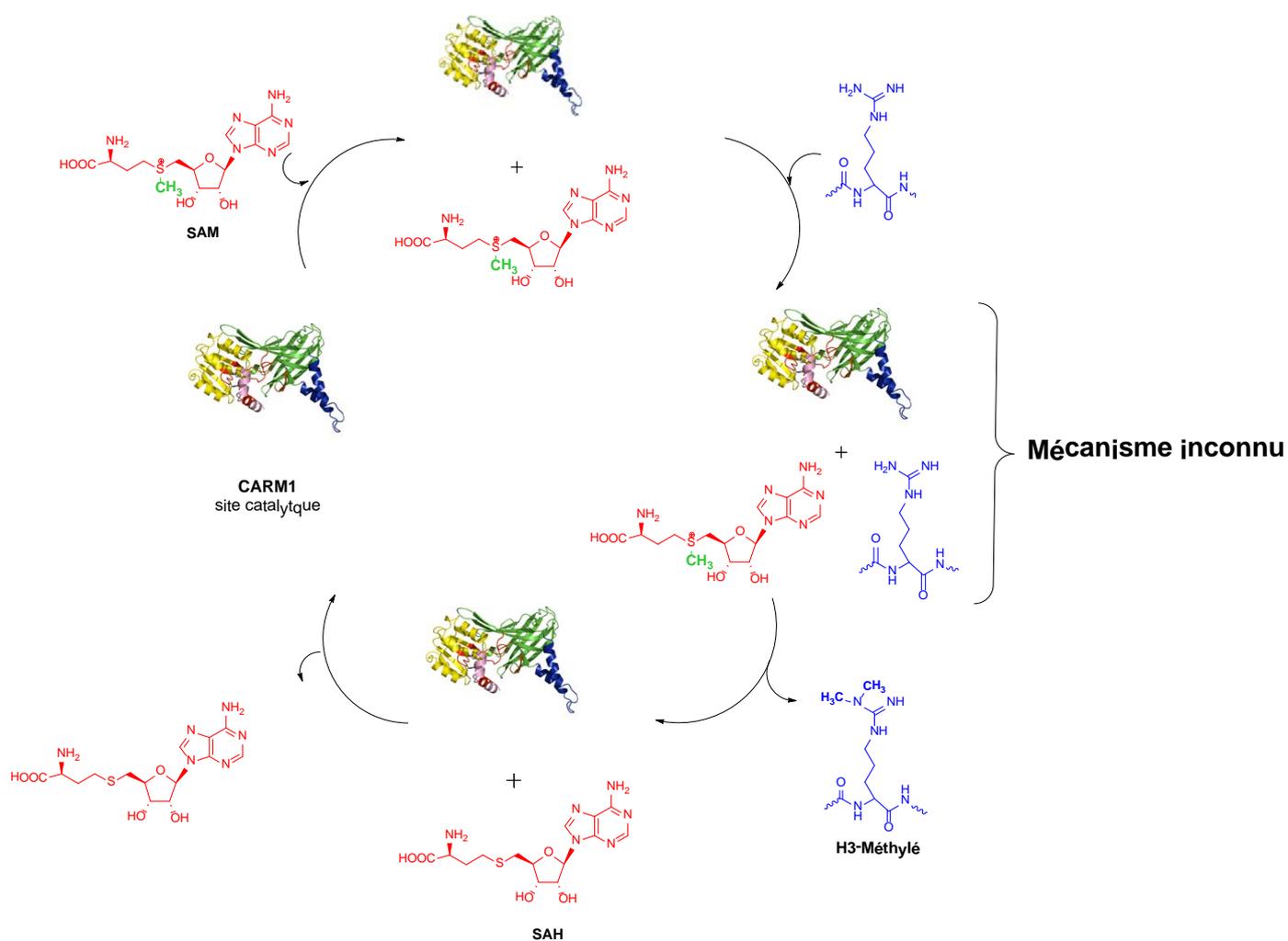


Figure 1.22 : Réaction de méthylation catalysée par l'enzyme CARM1.

Nous aimerions étudier cette réaction en utilisant la diffraction aux rayons X. Le défi est d'obtenir une structure cristallographique du complexe trimérique CARM1/SAM/substrat.

Des travaux menés au préalable ont montré que la présence de cofacteurs dans l'enzyme a pour effet de stabiliser le site catalytique.

Notre stratégie est donc de créer des molécules hameçons ou des molécules ancres basées sur le motif de la SAM capables d'ancrer un peptide mimant la séquence du substrat choisi, l'histone H3, pour ensuite les co-cristalliser avec CARM1. Ainsi, grâce à la diffraction aux rayons X, les interactions mises en jeu dans le complexe CARM1/molécules hameçons/peptide pourront être déterminées. Cette stratégie s'est effectuée en trois étapes (figure 1.23) :

- La première étape a consisté en la synthèse d'analogues de la SAM obtenus grâce à des modifications réalisées autour de l'atome de soufre. Ces composés nous permettront d'explorer la « poche du sulfonium ».
- La seconde étape a été la synthèse d'analogues de bisubstrats nécessaires pour l'exploration de la « poche de l'arginine ».
- Pour finir, dans une dernière étape, nous allons aborder la synthèse d'adduits SAM-peptide pour pouvoir étudier le « domaine de fixation du peptide ».

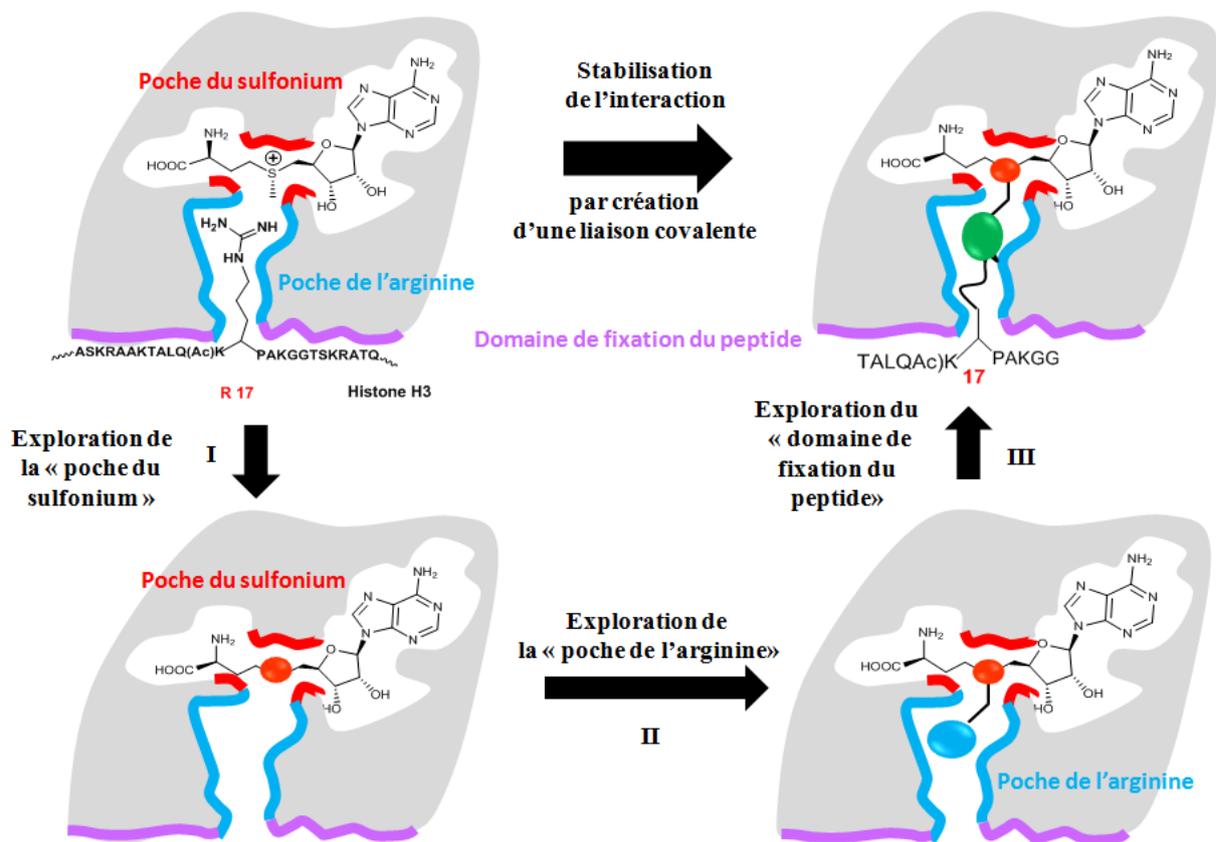


Figure 1.23 : Stratégie en trois étapes employée pour stabiliser le complexe CARM1/molécules hameçons/peptide : stabilisation de l'interaction en créant une liaison covalente à partir de molécules hameçons ou molécules ancres ayant pour motif la SAM. **a)** Etat de transition de la réaction entre la SAM et l'histone H3. **b)** Exploration de la « poche du sulfonium » par co-cristallisation d'analogues de la SAM. **c)** Exploration de la « poche de l'arginine » par co-cristallisation d'analogues de bisubstrats. **d)** Exploration du « domaine de fixation du peptide » par co-cristallisation d'adduits SAM-peptide.

Dans le paragraphe qui suivra, la première étape de notre stratégie va être décrite. Elle nous a permis de découvrir des inhibiteurs sélectifs de CARM1 par rapport à PRMT1. Ce travail à fait l'objet d'une publication qui sera introduite dans le chapitre suivant.

Références

1. A. Bird, *Genes Dev.*, **2002**, *16*, 6–21.
2. Site internet : <http://www.inserm.fr/thematiques/genetique-genomique-et-bioinformatique/dossiers-d-information/epigenetique/>
3. a) K. Luger, A.W. Mader, R.K. Richmond, D.F. Sargent, T.J. Richmond, *Nature*, **1997**, *389*, 251-260. b) K. Luger, *Curr Opin Genet Dev*, **2003**, *13*, 127–135.
4. M.S. Ong, T.J. Richmond, C.A. Davey, *Journal of Molecular Biology*, **2007**, *368*, 1067-1074.
5. Site internet :
<http://www.ncbi.nlm.nih.gov/Structure/mmdb/mmdbsrv.cgi?Dopt=s&uid=45929>
6. P. Ridgway, C. Maison, G. Almouzni, Chromatine : Organisation Fonctionnelle du Génome. Atlas Genet Cytogenet Oncol Haematol. April 2002. A. E. Handel, G.C. Ebers, S.V. Ramagopalan, *Trends Mol. Med*, **2009**. URL : <http://AtlasGeneticsOncology.org/Educ/ChromatinEducFr.html>.
7. Site internet : http://activity.ntsec.gov.tw/lifeworld/english/content/gene_cc5.html
8. a) S.I. Grewal, D. Moazed, *Science*, **2003**, *301*, 798-802. b) S.L. Berger, *Curr Opin Genet Dev*, **2002**, *12*, 142-148. c) W. An, *Subcell Biochem*, **2007**, *41*, 351-369.
9. F. Vandermeers, R. Kettmann, L. Willems, *Biotechnol. Agron. Soc. Environ.*, **2008**, *12*, 211-218.
10. A. Bird, *Science*, **2001**, *294*, 2113–2115.
11. T.H. Bestor, *Hum. Mol. Genet.*, **2000**, *9*, 2395–2402.
12. M. Schaefer, S. Hagemann, K. Hanna, F. Lyko, *Cancer Res.*, **2009**, *69*, 8127–8132.

13. Siteinternet :http://www.facbio.com/content/index.php?option=com_content&task=view&id=52&Itemid=1&limit=1&limitstart=3
14. a) V.G. Allfrey, A.E. Mirsky, *Science*, **1964**, *144*, 559. b) T. Kouzarides, *Cell*, **2007**, *128*, 693–705. c) S.L. Berger, *Science*, **2001**, *292*, 64-65.
15. a) M.S. Cosgrove, J.D. Boeke, C. Wolberger, *Nat Struct Mol Biol.*, **2004**, *11*, 1037–1043. b) T. Kouzarides, *Curr Opin Genet Dev*, **2002**, *12*, 198–209. c) B.D. Strahl, C.D. Allis, *Nature*, **2000**, *403*, 41–45.
16. Lacoste, J. Côté, *médecine sciences*, **2003**, *19*, 955-959.
<http://id.erudit.org/iderudit/007166ar>.
17. a) J.E. Brownell, J. Zhou, T. Ranalli, R. Kobayashi, D.G. Edmondson, S.Y. Roth, C.D. Allis, *Cell.*, **1996**, *84*, 843–851. b) L.Chang, S.S. Loranger, C. Mizzen, S.G. Ernst, C.D. Allis, A.T. Annunziato, *Biochemistry*, **1997**, *36*, 469-480. c) S. Kleff, E.D. Andrulis, C.W. Anderson, R. Sternglanz, *J. Biol. Chem.*, **1995**, *270*, 24674-24677. d) M.R. Parthun, J. Widom, D.E. Gottschling, *Cell*, **1996**, *87*, 85-94. e) N. Champagne, N.R. Bertos, N. Pelletier, A.H. Wang, M. Vezmar, Y. Yang, H.H. Heng, X.J. Yang, *J. Biol. Chem*, **1999**, *274*, 28528-28536. f) A. Hilfiker, D. Hilfiker-Kleiner, A. Pannuti, J.C. Lucchesi, *EMBO J.*, **1997**, *16*, 2054-2060. g) J. Kamine, B. Elangovan, T. Subramanian, D. Coleman, G. Chinnadurai, *Virology*, **1996**, *216*, 357-366. h) E.R. Smith, A. Eisen, W. Gu, M. Sattah, A. Pannuti, J. Zhou, R.G. Cook, J.C. Lucchesi, C.D. Allis, *Proc. Natl. Acad. Sci. USA*, **1998**, *95*, 3561-3565. i) V.V. Ogryzko, R.L. Schiltz, V. Russanova, B.H. Howard, Y. Nakatani, *Cell.*, **1996**, *87*, 953-959.
18. a) S.G. Gray, T.J. Ekstrom, *Exp. Cell Res.*, **2001**, *262*, 75–83. b) A.J. de Ruijter, A.H. van Gennip, H.N. Caron, S. Kemp, A.B. van Kuilenburg, *Biochem. J.*, **2003**, *370*, 737–749. c) I.V. Gregoretti, Y.M. Lee, H.V. Goodson, *J. Mol. Biol.*, **2004**, *338*, 17–31.

19. M.J. Carrozza, R.T. Utley, J.L. Workman, J. Côté, *Trends Genet*, **2003**, *19*, 321-329.
20. H. Santos-Rosa, R. Schneider, A.J. Bannister, J. Sherriff, B.E. Bernstein, N.C. Emre, S.L. Schreiber, J. Mellor, T. Kouzarides, *Nature*, **2002**, *419*, 407-411.
21. Y. Shi, F. Lan, C. Matson, P. Mulligan, J.R. Whetstine, P.A. Cole, R.A. Casero, *Cell* **2004**, *119*, 941-953.
22. Y. Tsukada, J. Fang, H. Erdjument-Bromage, M.E. Warren, C.H. Borchers, P. Tempst, Y. Zhang, *Nature*, **2006**, *439*, 811-816.
23. K. Yamane, C. Toumazou, Y. Tsukada, H. Erdjument-Bromage, P. Tempst, J. Wong, Y. Zhang, *Cell*, **2006**, *125*, 483-495.
24. R.J. Klose, K. Yamane, Y. Bae, D. Zhang, H. Erdjument-Bromage, P. Tempst, J. Wong, Y. Zhang, *Nature* **2006**, *442*, 312-316.
25. J.R. Whetstine, A. Nottke, F. Lan, M. Huarte, S. Smolikov, Z. Chen, E. Spooner, E. Li, G. Zhang, *Cell*, **2006**, *125*, 467-481.
26. B.D. Fodor, S. Kubicek, M. Yonezawa, R.J. O'Sullivan, R. Sengupta, L. Perez-Burgos, S. Opravil, K. Mechtler, G. Schotta, T. Jenuwein, *Genes Dev*, **2006**, *20*, 1557-1562.
27. P.A. Cloos, J. Christensen, K. Agger, A. Maiolica, J. Rappsilber, T. Antal, K.H. Hansen, K. Helin, *Nature*, **2006**, *442*, 307-311.
28. M. Wissmann, N. Yin, J.M. Muller, H. Greschik, B.D. Fodor, T. Jenuwein, C. Vogler, R. Schneider, T. Gunther, R. Buettner, E. Metzger, R. Schule, *Nat Cell Biol*, **2007**, *9*, 347-353.
29. a) A. Di Lorenzo, M. T. Bedford, *FEBS Letters*, **2011**, *585*, 2024-2031. b) C. D. Krause, Z. H. Yang, Y. S. Kim, J.H. Lee, J. R. Cook, S. Pestka *Pharmacol. Ther.*, **2007**, *113*, 50-87. c) Y. H. Lee, M. R. Stallcup, *Mol. Endocrinol.*, **2009**, *23*, 425-433.

30. G.L. Cuthbert, S. Daujat, A.W. Snowden, H. Erdjument-Bromage, T. Hagiwara, M. Yamada, R. Schneider, P.D. Gregory, P. Tempst, A.J. Bannister, T. Kouzarides, *Cell* **2004**, *118*, 545–553.
31. Y. Wang, J. Wysocka, J. Sayegh, Y.H. Lee, J.R. Perlin, L. Leonelli, L.S. Sonbuchner, C.H. McDonald, R.G. Cook, Y. Dou, R.G. Roeder, S. Clarke, M.R. Stallcup, C.D. Allis, S.A. Coonrod, *Science*, **2004**, *306*, 279–283.
32. K. Nakashima, T. Hagiwara, M. Yamada, *J Biol Chem*, **2002**, *277*, 49562–49568.
33. R.B. Denman, *Bioessays*, **2005**, *27*, 242–246.
34. Y.G. Zheng, J. Wu, Z. Chen, M. Goodman, *Med.Res.Rev.*, **2008**, *28*, 645-687.
35. S. Deltour, V. Chopin, D. Leprince, *Medecine et science*, **2005**, *21*, 405-411.
36. C.B. Yoo, P.A. Jones, *Nat Rev Drug Discov*, **2006**, *5*, 37–50.
37. T. Suzuki, N. Miyata, *Curr Med Chem*, **2006**, *13*, 935–958.
38. R. Mabaera, M.R. Greene, C.A. Richardson, S.J. Conine, C.D. Kozul, C.H. Lowrey, *Blood*, **2008**, *111*, 411-420.
39. H.K. Kuo, J.D. Griffith, K.N. Kreuzer, *Cancer Res.*, **2007**, *67*, 8248-8254.
40. F. Chik, M. Szyf, *Carcinogenesis*, **2011**, *32*, 224–232.
41. C.B. Klein, L. Su, D. Bowser, J. Leszczynska, *Environ Health Perspect*, **2002**, *11*, 739–743.
42. K. Schuebel, S. Baylin, *Nat Methods*, **2005**, *2*, 736–738.
43. B. Brueckner, R.G. Boy, P. Siedlecki, T. Musch, H.C. Kliem, P. Zielenkiewicz, S. Suhai, M. Wiessler, F. Lyko, *Cancer Res.*, **2005**, *65*, 6305–6311.

44. M.Z. Fang, Y. Wang, N. Ai, Z. Hou, Y. Sun, H. Lu, W. Welsh, C.S. Yang, *Cancer Res.*, **2003**, *63*, 7563–7570.
45. F-F Cai, C. Kohler, B. Zhang, M-H Wang, W-J Chen, X-Y Zheng, *Int. J. Mol. Sci.*, **2011**, *12*, 4465-4476.
46. L. Halby, C. Champion, C. Sénamaud-Beaufort, S. Ajjan, T. Drujon, A. Rajavelu, A. Ceccaldi, R. Jurkowska, O. Lequin, W. G. Nelson, A. Guy, A. Jeltsch, D. Guianvarc'h, C. Ferroud, P. B. Arimondo, *ChemBioChem*, **2012**, *13*, 157 – 165.
47. S.B. Hake, A. Xiao, C.D. Allis, *Br J Cancer*, **2004**, *90*, 761–769.
48. R. Schneider, A.J. Bannister, T. Kouzarides, *Trends Biochem Sci*, **2002**, *27*, 396–402.
49. S. Varambally, S.M. Dhanasekaran, M. Zhou, T.R. Barrette, C. Kumar-Sinha, M.G. Sanda, D. Ghosh, K.J. Pienta, R.G. Sewalt, A.P. Otte, M.A. Rubin, A.M. Chinnaiyan, *Nature*, **2002**, *419*, 624–629.
50. C.G. Klee, Q. Cao, S. Varambally, R. Shen, I. Ota, S.A. Tomlins, D. Ghosh, R.G. Sewalt, A.P. Otte, D.F. Hayes, M.S. Sabel, D. Livant, S.J. Weiss, M.A. Rubin, A.M. Chinnaiyan, *Proc Natl Acad Sci USA*, **2003**, *100*, 11606–11611.
51. G.J. van Leenders, D. Dukers, D. Hessels, S.W. van den Kieboom, C.A. Hulsbergen, J.A. Witjes, A.P. Otte, C.J. Meijer, *Eur Urol*, **2007**, *52*, 455–463.
52. C. Hwang, V.N. Giri, J.C. Wilkinson, C.W. Wright, A.S. Wilkinson, K.A. Cooney, C.S. Duckett, *Breast Cancer Res Treat*, **2008**, *107*, 235–242.
53. D. Greiner, T. Bonaldi, R. Eskeland, E. Roemer, A. Imhof, *Nat Chem Biol*, **2005**, *1*, 143–145.

54. S. Kubicek, R.J. O'Sullivan, E.M. August, E.R. Hickey, Q. Zhang, M.L. Teodoro, S. Rea, K. Mechtler, J.A. Kowalski, C.A. Homon, T.A. Kelly, T. Jenuwein, *Mol Cell*, **2007**, *25*, 473–481.
55. Y. Yao, P. Chen, J. Diao, G. Cheng, L. Deng, J. L. Anglin, B. V. Venkataram Prasad, Y. Song, *J. Am. Chem. Soc.*, **2011**, *133*, 16746–16749.
56. J. M. Yost, I. Korboukh, F. Liu, C. Gao, J. Jin, *Curr. Chem. Genom.*, **2011**, *5*, 72-84.
57. D. Cheng, N. Yadav, R. W. King, M. S. Swanson, E.J. Weinstein, M. T. Bedford, *J.Biol.Chem.*, **2004**, *279*, 23892-23899.
58. K. Bonham, S. Hemmers, Y. H. Lim, D. M. Hill, M. G. Finn, K. A. Mowen, *FEBS*, **2010**, *277*, 2096-2108.
59. A. Spannhoff, R. Heinke, I. Bauer, P. Trojer, E. Metzger, R. Gust, R. Schuele, G. Brosch, W. Sippl, M. Jung, *J.Med.Chem*, **2007**, *50*, 2319-2325.
60. R. Ragno, S. Simeoni, S. Castellano, C. Vicidomini, A. Mai, A. Caroli, A. Tramontano, C. Bonaccini, P. Trojer, I. Bauer, G. Brosch, G. Sbardella, *J. Med.Chem*, **2007**, *50*, 1241-1253.
61. S. Castellano, C. Milite, R. Ragno, S. Simeoni, A. Mai, V. Limongelli, E. Novellino, I. Bauer, G. Brosch, A. Spannhoff, D. Cheng, M. T. Bedford, G. Sbardella, *ChemMedChem*, **2010**, *5*, 398-414.
62. A. Spannhoff, R. Machmur, R. Heinke, P. Trojer, I. Bauer, G. Brosch, H. Roland, W. Sippl, J. Manfred, *Bioorg.Med.Chem.Lett.*, **2007**, *17*, 4150-4153.
63. R. Heinke, A. Spannhoff, R. Meier, P. Trojer, I. Bauer, M. Jung, W. Sippl, *ChemMedChem*, **2009**, *4*, 69-77.
64. Y. Feng, M. Li, B. Wang, Y.G. Zheng, *J.Med.Chem*, **2010**, *53*, 6028-6039.

65. E-M Bissinger, R. Heinke, A. Spannhoff, A. Eberlin, E. Metzger, V. Cura, P. Hassenboehler, J. Cavarelli, R. Schuele, M.T. Bedford, M. Jung, *Bioorg.Med.Chem.*, **2011**, *19*, 3717-3731.
66. T. Osborne, R. L. Weller Roska, S.R. Rajska, P. R. Thompson, *JACS*, **2008**, *130*, 4574-4575.
67. J. Dowden, W. Hong, R. V. Parry, R. A. Pike, S. G. Ward, *Bioor.Med.Chem.Lett.*, **2010**, *20*, 2103-2105.
68. J. Dowden, R. A. Pike, R. V. Parry, W. Hong, U. A. Muhsen, S.G. Ward, *Org.Biomol.Chem*, **2011**, *9*, 7814-7821.
69. T.M. Lakowski, P. Hart, C.A. Ahern, N.I. Martin, A. Frankel, *Chemical and Biology*, **2010**, *5*, 1053-1063.
70. P. Hart, T.M. Lakowski, D. Thomas, A. Frankel, N. L. Martin, *ChemBioChem*, **2011**, *12*, 1427-1432.
71. K.L. Bicker, O. Obiany, H.L. Rust, P.R. Thompson, *Mol.BioSyst.*, **2011**, *7*, 48-51.
72. D. Chen, H. Ma, H. Hong, S.S. Koh, S.M. Huang, B.T. Schurter, D.W. Aswad, M.R. Stallcup, *Science*, **1999**, *284*, 2174–2177.
73. M. Chevillard-Briet, D. Trouche, L. Vandel, *EMBO J.*, **2002**, *21*, 5457–5466.
74. J. Lee, M.T. Bedford, *EMBO*, **2002**, *3*, 268–273.
75. W. Xu, H. Chen, K. Du, H. Asahara, M. Tini, B.M. Emerson, M. Montminy, R.M. Evans, *Science*, **2001**, *294*, 2507–2511.
76. N. Yadav, J. Lee, J. Kim, J. Shen, M.C.T. Hu, C.M. Aldaz, M.T. Bedford, *Proc Natl Acad Sci U S A*, **2003**, *100*, 6464–6468.

77. S.L. Chen, K.A. Loffler, D. Chen, M.R. Stallcup, G.E.O. Muscat, *J. Biol Chem*, **2002**, 277, 4324-4333.
78. S.E. Messaoudi, E. Fabrizio, C. Rodriguez, P. Chuchana, L. auquier, D.Cheng, C. Theillet, L. Vandiel, M.T. Bedford, C. Sardet, *Proc Natl Acad Sci U S A*, **2006**, 103, 13351–13356.
79. S. Frietze, M. Lupien, P.A. Silver, M. Brown, *Cancer Res*, **2008**, 68, 301–306.
80. S. Majumder, Y. Liu, O.H. Ford, J.L. Mohler, Y.E. Whang, *Prostate*, **2006**, 66, 1292–1301.
81. C. Teyssier, D. Chen, M.R. Stallcup, *J. Biol. Chem.*, **2002**, 277, 46066-46072.
82. P. Hassenboeler, Thèse de l'Université de Strasbourg, **2010**.
83. N. Troffer-Charlier, V. Cura, P. Hassenboehler, D. Moras, J. Cavarelli, *EMBO J*, **2007**, 26, 4391-4401.
84. W.W. Yue, M. Hassler, S. M. Roe1, V. Thompson-Vale, L. H.Pearl, *EMBO J.*, **2007**, 26, 4402-4412.
85. A. Mai, S. Valente, D. Cheng, A. Perrone, R. Ragno, S. Simeoni, G. Sbardella, G. Brosch, A. Nebbioso, M. Conte, *ChemMedChem*, **2007**, 2, 987-991.
86. A. V. Purandare, Z. Chen, T. Huynh, S. Pang, J. Geng, W. Vaccaro, M. A. Poss, J. Oconnell, K. Nowak, L. Jayaraman, *Bioorg. Med. Chem. Lett.*, **2008**, 18, 4438-4441.
87. T. Huynh, Z. Chen, S. Pang, J. Geng, T. Bandiera, S. Bindi, P. Vianello, F. Roletto, S. Thieffine, A. Galvani, W. Vaccaro, M. A. Poss G.L.Trainor, M. V. Lorenzi, M. Gottardis, L. Jayaraman, A.V. Purandare, *Bioorg.Med.Chem.Lett*, **2009**, 19, 2924-2927.

88. J.S. Sack, S. Thieffine, T. Bandiera, M. Fasolini, G.J. Duke, L. Jayaraman, K.F. Kish, E. Herbert A.K. Purandare, P. Rosettani, *Biochemical Journal*, **2011**, *436*, 331-339.
89. H. Wan, T. Huynh, S. Pang, J. Geng, W. Vaccaro, M. A. Poss, G. L. Trainor, M. V. Lorenzi, M. Gottardis, L. Jayaraman, *Bioorg. Med. Chem. Lett.*, **2009**, *19*, 5063-5066.
90. B. R. Selvi, K. Batta, A. H. Kishore, K. Mantelingu, R. A. Varier, K. Balasubramanyam, S. K. Pradhan, D. Dasgupta, S. Sriram, S. Agrawal, *J. Biol. Chem.*, **2010**, *285*, 7143-7152.
91. D. Cheng, S. Valente, S. Castellano, G. Sbardella, R. Di Santo, R. Costi, Mark T. Bedford, A. Mai, *J. Med. Chem.*, **2011**, *54*, 4928-4932.

Chapitre 2

*Exploration de « la poche du sulfonium » par
co-cristallisation d'analogues de la SAM*

Chapitre 2

Exploration de la « poche du sulfonium » par co-cristallisation d'analogues de la SAM

Dans ce chapitre, nous commencerons l'exploration de la « poche du sulfonium ». Pour cela, nous synthétiserons des analogues de la SAM qui seront par la suite testés en co-cristallisation dans l'enzyme CARM1. Plusieurs analogues de la SAM seront obtenus en modifiant le degré d'oxydation du soufre ou en effectuant des modifications isostériques. La fonction sulfonium de la SAM sera ainsi remplacé par un atome d'oxygène, un atome de sélénium ou un atome d'azote. Puis, une deuxième série d'analogues azotés, les « SAM-arginines mimiques », sera réalisée en incorporant dans leur structure une chaîne hydrophobe, hydrophile ou encore un hétérocycle. La synthèse de ces analogues sera décrite ainsi que leurs caractérisations biochimiques dans ce premier chapitre. Cette partie sera décrite sous la forme d'une publication. Puis, nous terminerons par les essais de co-cristallisation de ces analogues dans CARM1 et ces résultats seront abordés à la fin de ce chapitre.

I. A stepwise strategy towards selective inhibitors and peptide mimics of coactivator-associated arginine methyltransferase 1

A stepwise strategy towards selective inhibitors and peptide mimics of coactivator-associated arginine methyltransferase 1.

Samira Ajebar,^{&,[a]} Vincent Cura,^{&,[b]} Pierre Hassenboehler,^[b] Justine Mailliot,^[b] Nathalie Troffer-Charlier,^[b] Jean-Marie Wurtz,^[b] Jean Cavarelli^{*[b]} and Alain Wagner^{*[a]}

Recent studies have shown that, PRMTs are associated with the development of various human diseases, including cancer, inflammation and psychiatric disorders. In particular, abnormal expression of CARM1 (Coactivator-associated Arginine Methyltransferase 1) has been observed in several

Introduction

The PRMT family contains about nine human members (PRMT1 to PRMT9) and is divided into two major classes, type I and type II enzyme.^[1-3] Both classes first transfer the methyl group from *S*-adenosyl-*L*-methionine (SAM, also known as AdoMet) to terminal guanidine nitrogen of arginine to form methylarginine as a transient intermediate and *S*-adenosyl-*L*-homocysteine (SAH, also known as AdoHcy). This transient intermediate then undergoes a second methylation with a regioselectivity that depends on the enzyme's type. Type I enzymes (PRMT1-3, CARM1 or PRMT4, PRMT6, and PRMT8) lead to asymmetric *N*, *N*-dimethylarginine (*a*DMA), whereas type II enzyme (PRMT5, and PRMT9) form symmetric *N,N'*-dimethylarginine side chains (*s*DMA) (Figure 1).

Recent studies have shown that PRMTs are associated with the development of various human diseases, including cancer, inflammation, and psychiatric disorders.^[4-6]

humancancers such as prostate, breast and colorectal cancers. Consequently, CARM1 represents promising target for the development of anticancer agents. We report herein a strategy for the stepwise elaboration of selective inhibitors by sequential exploration of CARM1 sulfonium pocket, arginine pocket and peptide binding surface. From a first series of SAM analogues, a methymamine derivative **21b** was selected as lead compound to investigate the possibility of establishing a specific interaction within the CARM1 catalytic site. Compound **21b** showed an IC₅₀ value equal to 158 μM for CARM1 and a selectivity factor over PRMT1 equal to 12.7 thus underlying the possibility of achieving selectivity with only slight structural changes in the inhibitor. Moreover, best compounds didn't cytotoxic at 250μM on breast tumour cell lines (MCF7). These results will now provide the basis to pursue our strategy and further elaborate "Sam peptide mimic"

In particular, abnormal expression of CARM1 (Coactivator-associated Arginine Methyltransferase 1) has been observed in several human cancers such as, prostate,^[7] breast,^[8] and colorectal cancers.^[9] Consequently, CARM1 represent promising targets for the development of anticancer agents as illustrated by a recent review written by Jin and co-workers.^[10]

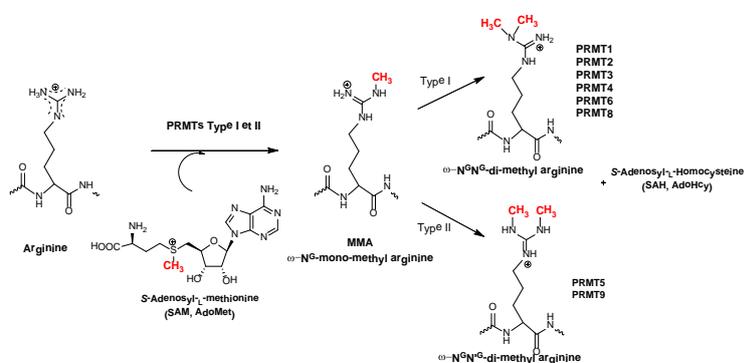


Figure 1. Methylation of arginine residues in proteins by PRMTs. PRMTs are divided into four major classes (type I to type IV) depending on the chemistry of methyl-arginine they generate. Using *S*-adenosyl-*L*-methionine (SAM) as methyl donor, Type I and type II PRMT, the two most populated classes generate monomethylarginine on one of the terminal guanidine nitrogen atoms and then asymmetric or symmetric dimethylarginine.

To synthesize PRMTs inhibitors three main strategies were employed. The first one that relied on high throughput screening (HTS) assays and in-silico approaches allowed to find selective micromolar inhibitor. For example, Bedford and al^[11] first reported naphthyl sulfonate (NS) inhibitors able to

inhibit PRMT1 with an IC_{50} value equal to 8.8 μ M. Later Finn and co-workers^[12] have shown that the best compound of this family is selective for PRMT1 over CARM1 with a selectivity factor equal to 19. Then Sbardella and co-workers^[13] identified through virtual screening approaches related naphthyl urea able to inhibit PRMT1 with an IC_{50} value equal to 9.3 μ M and a 5 fold selectivity over CARM1.

On the other hand HTS strategy enable identifying small potential inhibitors for CARM1.^[14-18] The first heterocyclic inhibitor that was more potent for CARM1 than for PRMT1 was described by Purandare and al.^[19] It belongs to 1H-pyrazol-1-yl)benzyl) propanamide family inhibit CARM1 with an IC_{50} value equal to 0.04 μ M and showed a low PRMT1's affinity ($> 25 \mu$ M). More recently benzyl-3,5-dibenzylidenepiperidin-4-one inhibitor of CARM1 was described by Cheng and al.^[20] They showed an IC_{50} value of 8.6 μ M for CARM1 whereas no inhibition for PRMT1 and a lysine methyltransferase (SET7) with a IC_{50} value superior to 667 μ M was observed.

The second strategy mainly directed toward discovery of PRMT1 inhibitor relied on the rational exploration of the PRMT1 sulfonium and the arginine binding pockets by synthesizing aza-AdoMet analogues able to be competitive with the SAM co-factor. First, Thompson and al^[21] synthesized aziridine AdoMet able to inhibit PRMT1 with an IC_{50} value equal to 18.5 μ M. Then Dowden and al^[22-23] have described guanidine-AdoMet analogues able to interact with the Arginine binding pocket. The best guanidine-AdoMet analogue showed an IC_{50} value of 2.9 μ M for PRMT1 with a good selectivity over CARM1 and over SET7. Interestingly further modification of the amino acid residue did not yield to any selectivity improvement.

The third strategy concerns the discovery of PRMT1's inhibitors by synthesizing peptide inhibitors mimics

of PRMT1 substrate. This strategy relies on the interaction of the inhibitors with the PRMT1 peptide binding surface. Thompson and al,^[24] have identified a peptide inhibitor C21, mimic of the N terminus of histone H4, with an IC_{50} of 1.8 μ M for PRMT1 and a very low affinity for CARM1 ($IC_{50} > 500$). In the same vein Frankel and al^[25] yield peptide inhibitor R1-Lys, mimic of fibrillar substrate, with an IC_{50} value of 14 μ M for PRMT1 and 2-fold selectivity over CARM1 and over PRMT6.

In this paper we report a rational structure-based approach to discover selective CARM1. This strategy, inspired by the work of Thompson^[21] and Dowden^[22-23] combines chemistry and molecular modeling. It relies on the stepwise elaboration of inhibitors by sequential exploration of the CARM1 sulfonium pocket, the arginine pocket and the peptide binding surface (Figure 2).

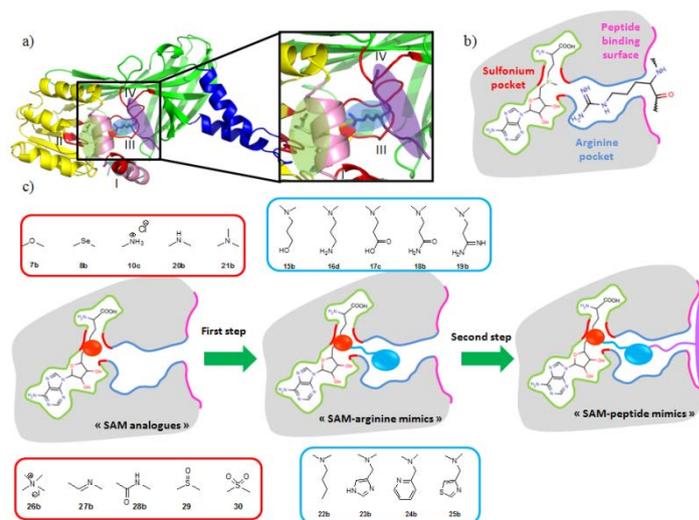


Figure 2. Rational design strategy for CARM1 bisubstrate inhibitors. a) Overview of one catalytic module of CARM1. This module, common to all PRMTs, is folded into two domains; a SAH/SAM binding domain shown in yellow and a α -barrel shown in green. The helix segment shown in blue is involved in the formation of the catalytic dimer. The four motifs characteristic of the PRMT catalytic center (I,II,III,IV) are highlighted in red. Ellipsoids are shown here to define the three main zones: sulfonium pocket in light green, arginine pocket in light blue and the peptide binding surface in light purple. b) Schematic representation of the active site of CARM1 illustrating the definition of pockets used to chemically explore the active site. The peptide substrate is anchored in CARM1 via its arginine residue. c) A three-step strategy to design CARM1 inhibitors. First step consists of designing "SAM analogues" (**7b-10c**; **20b-21b**; **26b-30**) with modifications around the sulfur atom to explore sulfonium pocket. The sulfur atom was first replaced by either oxygen, or selenium or nitrogen atom. Second step consists of designing "SAM arginine mimics" (**15b-19b**; **22b-25b**), to explore arginine pocket. That's why several nitrogen analogues were synthesized with the addition of a tail including either hydrophobic or hydrophilic groups, and heterocyclic derivatives Third step consists of designing "SAM arginine peptide mimics" to explore peptide binding surface.

Several crystal structures of the catalytic domain of CARM1 have shown that cofactor-binding, such as SAM, SAH or sinefungin (SFG), produces large conformational changes in the catalytic domain.^[26-27] These crystal structures clearly illustrate that SAM binding is a prerequisite to build up the productive peptide binding site.^[28]

Our first aim is thus to design inhibitors bearing analogues of SAM motif which should therefore be able to induce the productive conformational change. Two types of analogues have been synthesized and tested, referred to hereafter as "SAM analogues" and "SAM-arginine mimics". (Figure 2) For "SAM analogues", we carried out an isosteric modification at the sulfur atom by varying the nature of the heteroatom, its substitution and its oxidation state. Building on the best sulfonium pocket inhibitor we elaborated "SAM-arginine mimics", by introducing various side chains to establish specific interactions with the amino acid residues within the catalytic site.

Results and discussion

Chemistry

Syntheses of SAM analogues (Scheme 1) were carried out from commercial, boc-aspartic acid **1** and adenosine isopropylidene **4**. Acid **1** was reduced to alcohol **2**^[29] and further oxidized into aldehyde **3** using Dess-Martin periodinane reagent.

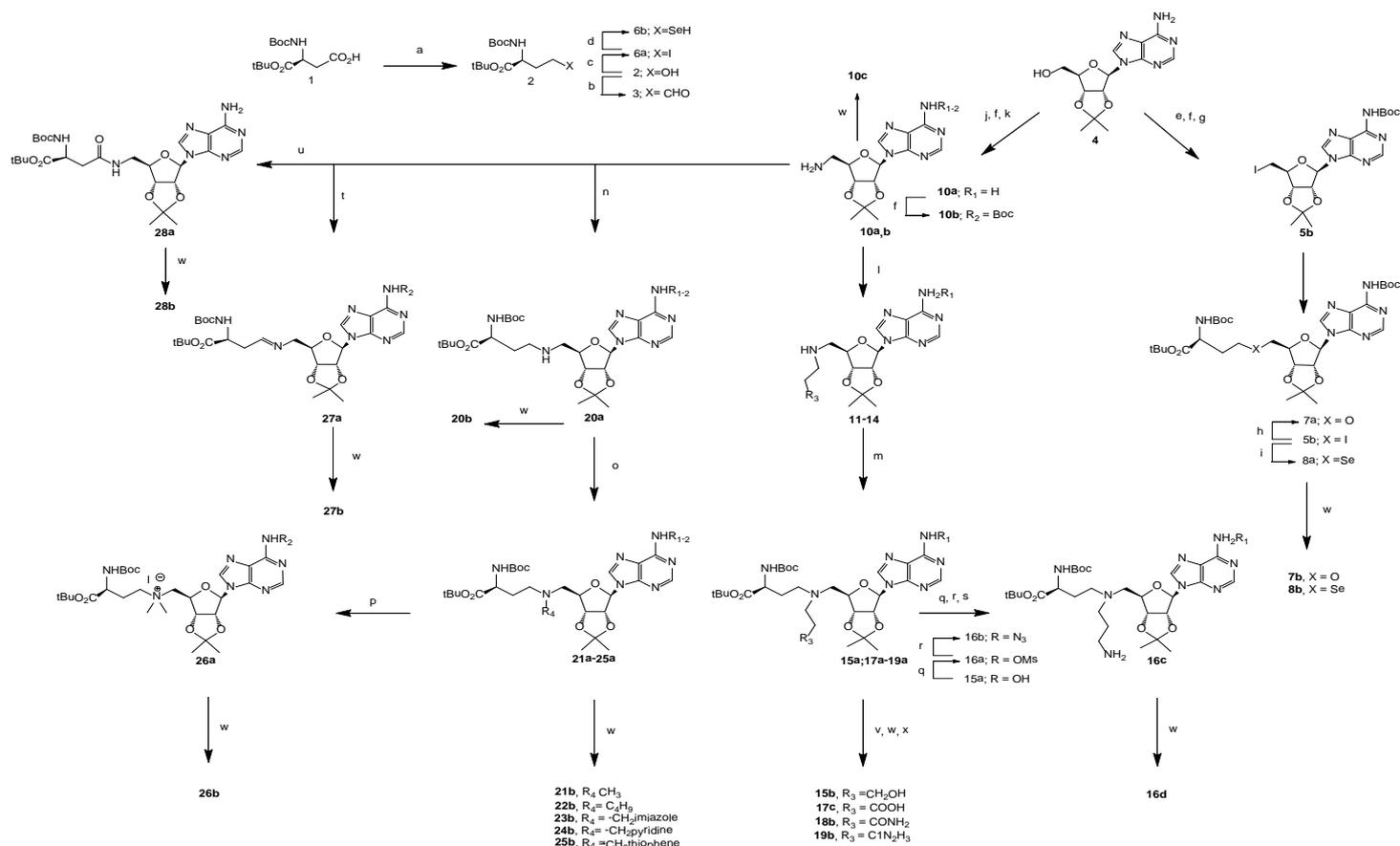
An electrophilic derivative **5b** was synthesized in three steps with an overall yield of 36 %. First the alcohol function of adenosine derivative **4** was activated by a tosylate

group as previously described.^[30] Then, N₆ position of adenine of was protected by Boc₂O to avoid the formation of charged cycloadenosine side product.^[31] Finally, the tosylate derivative **5a** was substituted by sodium iodide to afford derivative **5b**. On the other way a nucleophilic derivative **10a-b** was synthesized using a sequence described^[32-33] in two steps with an overall yield of 63%.

Electrophilic iodo derivative **5b** was reacted with alcohol **2** or selenol **6b**^[34] to afford derivative **7a** and **8a** respectively. They were further deprotected, under acidic conditions to afford desired products **7b** and **8b** with an overall yield of 30% and 50% respectively. Compound **8a** was easily identified by RMN spectroscopy; the shift of the carbon's signal next to the iodo shifted from 5.2 ppm to 36.5 ppm for the seleno derivative **8a**.

Nucleophilic amine derivative **10a** underwent an alkylation^[35] with 3-bromopropanol, 3-bromopropionamide, methyl 3-bromopropionate and 3-bromopropionitrile in THF using DIEA as a base to yield the secondary amines **11-14**. These compounds were submitted to a reductive amination with aldehyde **3** to afford derivatives **15a**, **17a-19a**. Analogue **16c** was synthesized from alcohol **15a** in three steps. The activation of the alcohol by a mesyl group gave derivative **16a**, followed by substitution with sodium azide to obtain compound **16b**.

Subsequent azide reduction under Staudinger's condition afforded amine **16c**, with a 23% yield over the three steps. From compound **17a**, saponification and subsequent deprotection under acidic conditions led to bis-acid derivative **17c**.



Scheme 1: Reagents and conditions: a) **1**, EtOCOCl, DIEA, THF, 45 minutes at -5°C, NaBH₄, H₂O, 12h at room temperature; b) **2**, Dess-Martin reagent, DCM, 6h at room temperature; c) **2**, I₂, PPH₃, imidazole, DCM, 3h at room temperature; d) **Se** (s), DMF, NaOH, N₂H₄, 3h at 60°C, then **6a**, DMF, 18h at 60°C, NaBH₄, EtOH, 2h at 0°C then 12h at room temperature; e) **4**, TsCl, pyridine, 6h at 0°C; f) Boc₂O, NaH, THF/DMF, 12h at 60°C; g) NaI, acetone, 12h at 60°C; h) **2**, NaH, THF, K₂CO₃, 24h at 70°C; i) **6b**, DMF, 12h at 80°C; j) **4**, DPPA, DBU, dioxane, 3h at room temperature, TBAI, NaN₃, 12h at 110°C; k) H₂/Pd/C, MeOH, 12h at room temperature; l) **10 a-b**, R₃X, DIEA, THF, 12h at 60°C; m) **3**, **11-14**, HOAc, NaBH₃CN, MeOH, 12h at room temperature; o) **20a**, R₄CHO, HOAc, NaBH₃CN, MeOH, 12h at room temperature; p) **21a**, MeI, EtOH anhydre, 6 days at room temperature; q) **15a**, MsCl, DIEA, THF, 12h at room temperature; r) **16a**, NaN₃, CH₃CN, 48h at 80°C; s) **16b**, PPH₃, H₂O, THF 12h at 60°C; t) **3**, HOAc, MeOH, 5 days at room temperature; u) **1**, NHS, EDC, DCM, 24h at room temperature, then **10a**, DIEA, THF, 3h at 60°C; v) **17a**, LiOH 1M, THF/MeOH (3/1), 12h at room temperature; w) HCl in dioxane, 1h at room temperature; x) **19a**, HCl in MeOH, 12h at room temperature, NaHCO₃ saturated, NH₄Cl saturated, H₂O, 20h at 37°C.

Finally, ternary amine analogues **15a-18a** were deprotected by treatment with HCl dioxane 4.0 M at room temperature for 2 hours. To remove the protecting groups of **19a** and to form an iminoether intermediary, we used a large excess of HCl methanol 1.25 M at room temperature overnight. Then a basic treatment using NaHCO_{3sat} and NH₄Cl_{sat} at 37°C for 20 hours was performed to afford amidine derivative **19b**. The disappearance of the cyano group band at 2252 cm⁻¹ could be observed by infrared spectroscopy. Moreover, the ¹³C NMR signal corresponding to the cyano group at 118.9 ppm disappeared while a new signal corresponding to the quaternary carbon of amidine appeared at 155.6 ppm. Using this synthetic pathway the final products were obtained with an average overall yield of 25%.

Amine derivatives **20a-25a** were synthesized via a second synthetic route involving two successive reductive aminations of the key intermediate **10a-b** as described by Dowden and al.^[23] Noteworthy, while first reductive amination proceeded without noticeable trouble, the second reductive amination was found to be much slower. For instance, when reacting an aldehyde bearing a heterocyclic substituent, the reaction mixture had to be heated for 3 days at 60°C. Derivative **21a** was alkylated by methyl iodide to obtain the ammonium intermediate **26a**. Finally, **20a-26a** were deprotected as described above to afford the final products **20b-26b**, with overall yield ranging from 25 to 30%. Imine derivative **27a** was obtained from amine **10b** and aldehyde **3** using acetic acid in methanol. No significant hydrolysis could

be observed after deprotection reaction and imine **27b** was obtained with a global yield of 19%. Finally, amine **10a** was subjected to coupling with acid aspartic **1** to afford amide **28a**, which was deprotected to obtain the amide analogue **28b** with an overall yield of 49%. Sulfoxide **29** and sulfone **30** derivatives were synthesized from commercially available SAH.^[36-37]

Biology

The ability of each analogue to compete with the SAM molecule in the catalytic site of either mouse CARM1 or mouse PRMT1 was used to classify the compounds. For this purpose, the methylation activities of CARM1₁₃₀₋₅₀₇ (amino acids from 130 to 507 of mouse CARM1) displaying the same catalytic activity as the full length CARM1 (608 amino acids) (Hassenboehler et al., to be published) and PRMT1 (354 amino acids) were measured using the histone tetramer (H3H4)₂ as a substrate and ¹⁴C SAM as the methyl donor. Incorporation of radioactivity in histone H3 for CARM1 and in histone H4 for PRMT1 was quantified using gel based assays (see experimental procedure).

The residual enzymatic activity was measured using several analogues concentrations from 0.5 μM to 2 mM and was expressed as a percentage of the enzyme activity without an analogue. Relative IC₅₀ values were derived from these experiments by fitting the data to a dose response curve model. Inhibition of the methylation reaction was used to estimate the ability of each compound for the SAM binding pocket. Full experimental data are only shown in Figure 3 for some selected compounds (see discussion below).

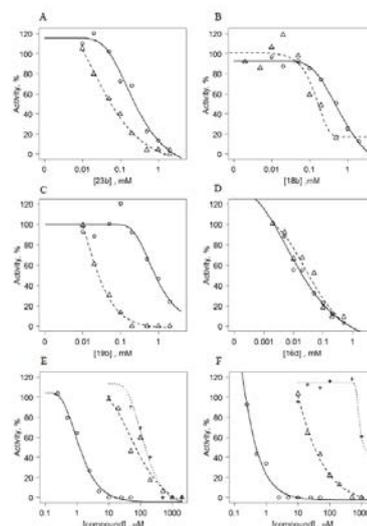


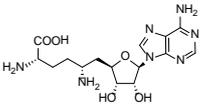
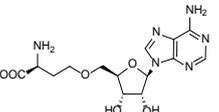
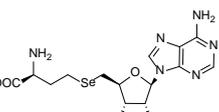
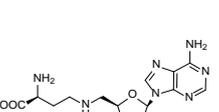
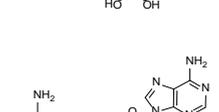
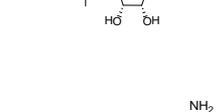
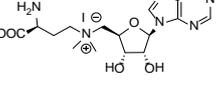
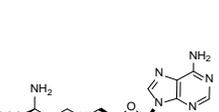
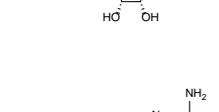
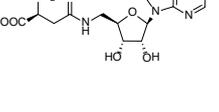
Figure 3. Inhibition of CARM1 and PRMT1 activities. Panels A,B,C,D: The fractional activities of CARM1 (Δ) and of PRMT1 (○) were by plotted with respect to the concentration of 23b (A), 18b(B), 19b(C) and 16d(D). Panels E and F: Effect of increasing concentrations of sinefungin (○), 20b (Δ) and 21b (+) on the activity of CARM1 (E) and PRMT1 (F) respectively.

We decided to only evaluate the CARM1/PRMT1 selectivity for the most active compounds (Table 1-2). For these compounds we define a “selectivity factor” (SF) as the ratio between the IC₅₀ calculated for PRMT1 to the IC₅₀ calculated for CARM1.

We first tested the inhibition capabilities of several analogues containing a modification surrounding the sulfur atom. Sinefungin (SFG), a good but not selective natural inhibitor of biological systems that use SAM as a methyl donor, was used as reference molecule.

These SAM analogues can be divided into two categories. The first group has isosteric oxygen **7b** and selenium **8b** atom or it possesses an amide **28b** function in place of the 5'-sulfur atom of SAH. Finally, sulfur oxidation state has been varied into sulfoxide **29** or a sulfone **30**.

Table 1. Inhibitory activities (IC_{50} values) of SAM analogues having a structural modification around the sulfur atom against CARM1 and over PRMT1 using the histone tetramer ($H3H4$)₂ as substrate.^[b]

Compound	Structure	$IC_{50} \pm SD$ [μ M]		PRMT1 / CARM1
		CARM1	PRMT1	
SFG		0.7 ± 0.3	0.6 ± 0.3	0.9
7b		>1000	ND ^[c]	-
8b		600 ± 50	ND ^[c]	-
20b		37 ± 5	31 ± 5	0.8
21b		158 ± 52	>2mM	-
26b		75 ± 5	>2mM	-
27b		27 ± 9	200 ± 20	7.4
28b		>1000	ND ^[c]	-
29		350 ± 33	ND ^[c]	-
30		250 ± 50	ND ^[c]	-

[a] Histone tetramer ($H3H4$)₂ were used as substrates and ¹⁴C SAM as the methyl donor; [b] Values are means \pm SD derived from at least two separate experiments; [c] Not determined.

Table 2. Inhibitory activities (IC₅₀ values) of SAM analogues having a structural modification around the sulfur atom against CARM1 and over PRMT1 using the histone tetramer (H3H4)₂ as substrate.^[b]

Compound	Structure	IC ₅₀ ± SD [μM] CARM1	IC ₅₀ ± SD [μM] PRMT1	PRMT1 / CARM1
SFG		0.7 ± 0.3	0.6 ± 0.3	1
10c		16 ± 2	380 ± 44	21
15b		76 ± 11	290 ± 50	5
16d		30 ± 10 ^[d]	32 ± 4 ^[d]	1
17c		137 ± 5	260 ± 13	2
18b		115 ± 14	757 ± 38	7
19b		20 ± 12	825 ± 164	41
22b		120 ± 22	ND ^[c]	-
23b		43 ± 24	294 ± 82	7
24b		621 ± 50	ND ^[c]	-
25b		409 ± 40	ND ^[c]	-

[a] Histone tetramer (H3H4)₂ were used as substrates and ¹⁴C SAM as the methyl donor; [b] Values are means ± SD derived from at least two separate experiments; [c] Not determined; [23] Literature value (Reference [23]): 3.9 ± 1.8 μM (PRMT1) and no activity for CARM1, *In vitro* methylation reactions.

It is worth noting that, sulfoxide **29** and sulfone **30** had poor inhibitory properties on CARM1 whereas they are known to be potential inhibitors of SAM dependent DNA methyltransferase^[38] and phosphatidylethanolamine methylase.^[39] This finding is interesting since the three enzymes all use SAM as a co-factor thus underlying the possibility of achieving selectivity with only slight structural changes in the inhibitor. The second group concerns aza analogues of SAM (**20b-21b**; **26b-27b**) which could inhibit CARM1 in the micro molar range (Table 1). Interestingly, compound **20b** a quite potent inhibitor of both CARM1 and PRMT 1 ($IC_{50} = 37$ and $31 \mu M$ respectively) is reported to show a very low inhibition for catechol-o-methyltransferase (COMT) and tRNA methyltransferases (tRNAMTs) with a mM range.^[40]

The methylamine **21b** was selected as a starting point to investigate the possibility of establishing a specific interaction within the PRMT catalytic site. For synthesis reasons **21b** was selected instead of ammonium **26b** because of the possible side reactions when performing amine quaternization on a wide range of functionalized derivatives.

Among these SAM arginine mimics three categories can be distinguished. The first one encompasses inhibitors showing both strong inhibition of CARM1 with a range of 20 to $115 \mu M$ and low inhibition of PRMT1 (**18b-19b**, **23b**, **26b-27b**). The second category comprises inhibitors that show a moderate inhibitory effect for CARM1 with a range of 67- $158 \mu M$, but maintain their selectivity over PRMT1 (**15b**, **17c** and **21b**). The third category of analogues efficiently inhibits CARM1 with a range of 30- $37 \mu M$, but also PRMT1 (**16d** and **20b**). Notably the compounds bearing a heterocyclic substituent (**24b** and **25b**) did not inhibit CARM1, while structurally related analogue **23b** was found to have a good IC_{50} value equal to $43 \mu M$ for CARM1 inhibition.

In this work our aim was to enhance the specificity and selectivity over other PRMTs, such as PRMT1. And above all to validate our strategy enabled us to follow our exploration of the CARM1 peptide binding surface. Interestingly, a related strategy was used by Dowden et al^[22-23] to develop inhibitors of PRMT1 selective over SET7. Dowden synthesized a series of guanidine AdoMet analogues with various chain lengths between the AdoMet nitrogen and the guanidine. He found that an analogue with a tail of four carbon was the most suited linker length to achieve high PRMT1 inhibition ($2.9 \mu M$) and high selectivity over CARM1 and SET7. Noteworthy also **16d** in his hand showed no inhibition for CARM1 while in our experiment we found that **16d** had an IC_{50} value of $30 \mu M$ (Figure 3).

Cytotoxicity tests were measured for compounds **10c**, **15b-23b** and **26b-27b** on breast tumor cell lines (MCF7), and these compounds didn't show cytotoxicity. Indeed, six doses were used (data not shown) and at $250 \mu M$ there are 90% of viability for all compounds.

Ligand binding Affinity for CARM1 and pocket localization.

“Sulfonium pocket”

Ligands synthesized in the present study can be classified into two sets of compounds according to the size of the functional group that replaces the sulfur atom in SAM or the NH_2 moiety of SFG in the middle of these ligands. One set of the ligands can only bind to the "Sulfonium pocket" (see Figure 2) due to the small size of the additional groups, e.g. compounds **7b** or **20b**. The second set of synthetic compounds, with larger moieties (e.g. compounds **16d** to **19b**) binds to the Sulfonium pocket and also partially to the "arginine pocket", which corresponds to the CARM1 arginine pocket.

Regarding the "sulfonium pocket" binding ligands, one observes that most of these ligands are isosteric of SAH or SAM and can bind to CARM1 with an affinity ranging from a few 27 μM to more than 1000 μM . Compounds with low or no affinity are either in steric contact with residues which line the protein pocket or the substituent is not well accommodated in the pocket, for example the oxygen substitution in **7b** (1300 μM) compared to the nitrogen isosteric in **20b** (37 μM). Surprisingly, one compound **10c**, a degradation product, exhibited the best affinity (16 μM) for CARM1. This rather puzzling result will be further explored by comparing the selectivity between CARM1 and PRMT1. It is worth noting that despite the strong affinity of compound **10c** it will be difficult to develop and enhance such a molecule as there are not that many possibilities remaining in terms of synthetic strategies.

"Arginine pocket"

The ligands reaching the "arginine pocket" all contain small functional groups that indicate a rather constrained environment corresponding to the region where the guanidinium moiety is methylated. The environment is relatively polar, negatively charged with the two glutamic acid residues of motif III conserved among PRMT sequences (Glu258, Glu267 in mouse CARM1 numbering) and numerous main chain carbonyls. The characteristics of this environment plausibly explain the differences in affinity of this set of ligands for CARM1. Indeed, the affinity correlates well with the number of hydrogen donors present in the additional/substituted moiety, e.g. consider the series **17c**, **18b**, **19b**. The compound's affinity increases (IC_{50} value decreases) when nitrogen atoms replace the oxygen atoms. Note also the good binding of compound **16d** with an NH_2 group in the substituted chain.

CARM1 and PRMT1 ligand specificity

The "Sulfonium pocket" and the "Arginine pocket" form a continuous region and the distinction between the two areas is helpful in analyzing the specificity observed among these ligands between PRMT1 and CARM1. Indeed, several compounds revealed unexpected results for the two PRMTs tested in the present study: CARM1 and PRMT1. Based on the 3D modeling and docking approaches, a structural explanation can be given for those compounds by exploiting the CARM1 and PRMT1 crystal structures.

Focusing again on the "Sulfonium pocket" binding ligands, the ligands **27b** and **28b** exhibit a relatively rigid SAM skeleton due to either a double bond or an amide bond, respectively. The latter ligand causes steric clashes with the "Sulfonium cavity" explaining the absence of binding. The ligand **27b** introduces a planar geometry in the SAM analogue and is still able to bind to the two PRMTs by adopting a cis geometry. The difference in binding is most likely due to the amino acids sequence; CARM1 has a glutamine (Gln160) residue close to the double bond, while PRMT1 has a histidine (His53) residue. The cis configuration introduces a ligand geometry that extends in the direction of the glutamine or histidine residue in either CARM1 or PRMT1. The former may probably more accommodate these changes whereas the latter may not. In the light of the difference in sequence between the two proteins, one can also understand the difference in affinity of the compound **10c**. Most likely, the NH_2 group at the cleavage point interacts more favorably with the glutamine residue of CARM1 than with the histidine of PRMT1.

For the compound pointing to the "arginine pocket", it is worth to note the difference in specificity of the three compounds, **18b**, **19b** and **23b** between CARM1 and PRMT1. This is rather surprising as the environment around the

different moieties is rather conserved in all PRMTs and therefore lack of specificity may be expected. This is indeed what we observed for compounds (**16d** and **20b**) that exhibit the same affinity for both enzymes. The pocket around the functionalized ligands is mainly delineated by either negatively charged or polar oxygen provided by the two conserved glutamic acid from motif III and two main chain carbonyl oxygen atoms (from Glu258, Met260). The chains of the three ligands point towards the arginine pocket that is rather narrow and has strong steric restraints. Especially, ligand **22b** has

Instead of **A**

As shown in Figure 4, residues around Pro259 in mouse CARM1 and its equivalent residue Trp 153 in mouse PRMT1, differ significantly.

These differences surrounding the arginine pocket are the most likely explanation for the ligand specificity observed for compounds **18b**, **19b** and **23b**. The CARM1 structure seems to be able to accommodate the functional groups of these three ligands, while PRMT1 has a less flexible environment and is unable to form an adequate hydrogen

Conclusion

In this paper, we describe a stepwise approach to discover CARM1 inhibitors, which combines chemistry and molecular modeling. We synthesized and characterized

charged or polar oxygen provided by the two conserved glutamic acid from motif III and two main chain carbonyl oxygen atoms (from Glu258, Met260). The chains of the three ligands point towards the arginine pocket that is rather narrow and has strong steric restraints. Especially, ligand **22b** has

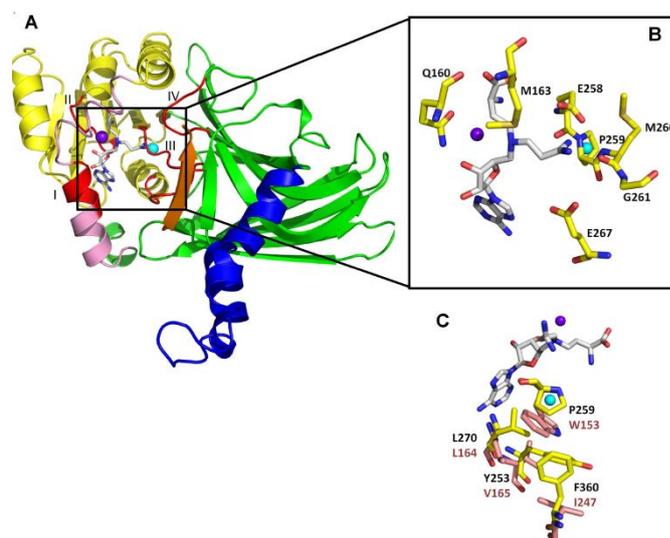


Figure 4. A) ligand **19b** in the pocket of catalytic site of CARM1 with two glutamic acids (Glu 267 and Glu 268) of motif III, and Pro259. Amidine group is surrounded by carbonyl atoms of protein backbone. B) Environment of Pro 259 and Trp 153. C) Residues of catalytic site of CARM1 are represented in white and residues of catalytic site of PRMT1 are represented in purple.

bond network. For example ligand **19b** contains an amidine moiety, which we expect to form two hydrogen bonds with CARM1. In contrast, ligand **18b** can only form one hydrogen bond. The imidazole ligand **23b**, in the protonated form can also form two hydrogen bonds, but it is a more rigid and bulkier compound compared to compound **19b**.

several “SAM analogues” or “SAM-arginine mimics”. We reported that modifying at the sulfur atom on “SAM analogues” lead to poor CARM1 inhibition and no selectivity. In contrast, ternary nitrogen SAM analogues lead to effective

inhibition at values. Inhibition experiments using two different PRMTs (CARM1 and PRMT1) show that such compounds can produce selectivity within the PRMT family. Using available crystal structures and molecular modeling, we provide a rational explanation for the results. This precious

Experimental Section

Chemistry

General. All chemicals and solvents were of reagent grade and were used without further purification. Unless indicated, reactions were carried out under an atmosphere of argon in flame-dried glassware with magnetic stirring. The progress of all reactions and column chromatography were monitored by Thin-layer chromatography (TLC). TLC was performed in using Merck silica gel 60 F₂₅₄ precoated plates and visualized by UV absorption or stained with a solution of ninhydrin in EtOH. Flash chromatography ARMEN was performed using Merck silica gel 45 (230-400 mesh) with an appropriate eluent. Sinefungin (SFG), S-(5'-Adenosyl)-L-homocysteine (SAH) and 2',3'-O-isopropylideneadenosine were obtained from Sigma-Aldrich (Germany), Boc-Aspartic Acid were obtained from Bachem (Germany), bromopropionamide were obtained from TCI Europe (Belgium). Bromopropanol, methyl 3-bromopropionate and bromopropionitrile were obtained from Alfa aesar (Germany). ¹H and ¹³C NMR spectra were recorded on a Bruker Avance DPX 300 Spectrometer (operated at 300 or 75 MHz, respectively) or ultrashield plus 400 Spectrometer (operated at 400 or 100 MHz, respectively) or ultrashield plus 500 Spectrometer (operated at 500 or 125 MHz, respectively). The ¹H NMR chemical shifts are reported in parts per million, using the residual solvent protons (7.24 for CDCl₃, 3.31 for MeOD, 4.78 for D₂O) as reference. The ¹³C NMR chemical shifts are reported in parts per million, using the signal for residual solvent protons as reference (77.23 for CDCl₃, 48.95 and for MeOD, for D₂O: 1 μL of MeOH is used as internal reference). Data were represented as follows, chemical shift, multiplicity (s = singlet, bs = broad singlet, d = doublet, dd = doublet of doublets, t = triplet, q = quadruplet and m = multiplet), integration and coupling (J in Hz). LC-MS data were obtained using a LC/MSD/SL Agilent Technologies; ionising source is APCI/ESI, Agilent technologies 1200 autosampler; pump agilent 1200 binary system; column ThermoScientific (58282): 30 x 1mm, 19 μm, hypersil gold C18, Agilent Technology 1200 DAD detector; Agilent technology MS quadrupole mass spectrometer operating in APCI/ESI mode. InfraRed spectra were obtained neat in the solid phase or from solution in CHCl₃ using a recorded on a Nicolet 380 FT-IR from Thermo Electron Corporation. The HPLC analyses were performed using a Shimadzu LC-20 AD system, equipped with a SPD 20A absorption detector. The HPLC system was as follows: **Method A:** Waters Sunfire C18; 5 μm; 4.6x50 mm column; 1 mL/minutes; eluent B 0 % over 5 minutes, then 0 to 100% over 10 minutes. Eluent A = H₂O, 0.01% TFA and eluent B = CH₃CN. The column was maintained at a temperature of 30°C. **Method B:** Agilent ZORBAX 300SB-C18; 5 μm; 4.6x250 mm column; 1 mL/minutes; eluent B 0 to 100% over 30 minutes. Eluent A = H₂O, 0.01% TFA and eluent B = CH₃CN. The column was maintained at a temperature of 30°C. Purities of final compounds were in all cases greater than 98%, as determined by reverse-phase HPLC analysis.

2-tert-Butoxycarbonylamino-4-oxo-butyrac acid tert-butyl ester (3). Dess-Martin periodinane reagent (6.8 g, 16.0 mmol, 1 eq) was dissolved in anhydrous dichloromethane (2 mL) and stirred at 0°C. Alcohol (2)^[29] (4.3 g, 16.0 mmol, 1 eq) dissolved in anhydrous dichloromethane (8 mL) was dropwise added to the solution. The reaction mixture was warmed to room temperature and stirred for 6 hours. After the reaction was finished, the solution was filtered through celite and extracted with dichloromethane (125 mL). After evaporating the solvent under reduced pressure, the residue was purified by silica flash column chromatography with cyclohexane/acetone (7/3) as eluent to give yellow oil (3.8 g, 86%). ¹H NMR (300MHz, CDCl₃):

information will allow us to pursue our strategy to find more potential inhibitors and further elaborate the "Sam peptide mimics" which, according to results published by, Thompson^[24] and Frankel^[25] and should yield analogues with improved efficacy and selectivity.

δ (ppm) = 1.40 (d, J = 4, 18H); 2.87-3.03 (m, 2H); 4.46 (m, 1H); 5.37 (bs, 1H, NH); 9.70 (s, 1H). ¹³C NMR (75 MHz, CDCl₃): δ (ppm) = 28.2; 46.4; 49.4; 80.1; 82.8; 155.5; 170.1; 199.4. IR ν (cm⁻¹): C-O (1155); C-N (1393); C=O (1719); Csp²-H_{aldehyde} (2933-2978).

Toluene-4-sulfonic acid 6-(6-tert-butoxycarbonylamino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl ester (5a). Tosyl compound^[30] (1.3 g, 2.2 mmol, 1 eq) was dissolved in anhydrous tetrahydrofuran (6 mL) and stirred at 0°C. Then sodium hydride (271 mg, 11.3 mmol, 4 eq) was added in one portion and the solution became dark red. Then di-tert-butyl dicarbonate (1.2 g, 5.6 mmol, 2 eq) dissolved in anhydrous N,N-dimethylformamide (1 mL) was dropwise added at 0°C to the solution. The reaction mixture was stirred at room temperature overnight. Then, the reaction mixture was diluted in ethyl acetate (100 mL) and washed with a solution of NH₄Cl_{sat} (2 x 50 mL) and a solution of HCl 1N (2 x 50 mL). The organic layer was dried over Na₂SO₄, filtered and evaporated. The crude was purified by silica flash column chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a yellow solid (1.3 g, 78%). EF: C₂₅H₃₁N₅O₈S₁. ¹H NMR (300 MHz, CDCl₃): δ (ppm) = 1.28 (s, 3H); 1.40 (s, 9H); 1.51 (s, 3H); 2.32 (s, 3H); 4.15 (bs, 2H); 4.41 (bs, 1H); 4.97 (dd, J₁ = 3.2; J₂ = 8.8, 1H); 5.27 (dd, J₁ = 2.8; J₂ = 8.4, 1H); 6.08 (d, J = 2.8, 1H); 7.16 (d, J = 12.4, 2H); 7.54 (d, J = 12.0, 2H); 8.07 (s, 1H); 8.66 (s, 1H). ¹³C NMR (100MHz, CDCl₃): δ (ppm) = 21.5; 25.3; 27.7; 28.1; 69.3; 82.2; 83.9; 84.1; 84.8; 90.9; 114.5; 120.2; 127.7; 129.5; 132.2; 139.8; 144.6; 148.9; 149.9; 153.3. IR ν (cm⁻¹): Csp²-H (554-843); C-O (1112); C-N (1307); O=S=O (1371); C=N (1599); C=O (1755); NH(2982). LC-MS (ESI-APCI): m/z [M+H]⁺ = 562.

[9-(6-Iodomethyl-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-yl)-9H-purin-6-yl]-carbamic acid tert-butyl ester (5b). Compound (5a) (0.9 g, 1.7 mmol, 1 eq) was dissolved in acetone (15 mL) and stirred at room temperature, then sodium iodide (2.5 g, 16.0 mmol, 10 eq) was added in one portion and the reaction mixture was stirred under reflux overnight. The reaction mixture was diluted in ethyl acetate (100 mL) and washed with a solution of NaOH 1N (3 x 30 mL), then the aqueous layer was extracted with dichloromethane (3 x 30 mL) dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using cyclohexane/ethyl acetate (4/6) as eluent to give a yellow solid (434 mg, 50%). EF: C₁₈H₂₄N₅O₅I. ¹H NMR (400 MHz, CDCl₃): δ (ppm) = 1.38 (s, 3H); 1.43 (s, 9H); 1.60 (s, 3H); 3.25 (dd, J₁ = 7.8; J₂ = 10.4, 1H); 3.42 (dd, J₁ = 7.8; J₂ = 10.4, 1H); 4.40-4.45 (m, 1H); 5.08 (dd, J₁ = 2.1, J₂ = 6.4, 1H); 5.49 (dd, J₁ = 2.1, J₂ = 6.4, 1H); 6.14 (d, J = 2.1, 1H); 8.04 (s, 1H); 8.74 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 5.2; 25.6; 27.2; 28.0; 82.3; 84.1; 84.6; 86.8; 91.3; 115.1; 120.4; 140.1; 144.3; 150.6; 151.2; 152.3. LC-MS (ESI-APCI): m/z [M+H]⁺ = 518.

2-tert-Butoxycarbonylamino-4-[6-(6-tert-butoxycarbonylamino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethoxy]-butyrac acid tert-butyl ester(7a). Alcohol (2)^[29] (11 mg, 0.041 mmol, 1.1 eq) was dissolved in anhydrous tetrahydrofuran (1 mL) and stirred at room temperature, then sodium hydride (2 mg, 0.083 mmol, 2.0 eq) and potassium carbonate (6 mg, 0.041 mmol, 1.0 eq) were successively added in one portion. Compound (5b) (19 mg, 0.037 mmol, 1.0 eq) was dissolved in anhydrous tetrahydrofuran (1 mL) and was added to the alcoholate at room temperature. Then the reaction mixture was stirred under reflux for 2 days. The reaction mixture was diluted in ethyl acetate (20 mL) and washed with a solution of NaOH 1N (3 x 10 mL), the aqueous layer was extracted with

dichloromethane (3 x 10 mL), dried over Na₂SO₄, filtered and evaporated. The crude was purified by silica flash column chromatography using a dichloromethane/methanol gradient (98/2 /1) to give a white solid (12 mg, 48%). EF: C₃₁H₄₈N₆O₁₀. ¹H NMR (300MHz, CDCl₃): δ (ppm) = 1.39 (s, 3H); 1.44 (s, 27H); 1.61 (s, 3H); 2.02-2.22 (m, 2H); 3.23-3.41 (m, 2H); 3.44-3.66 (m, 2H); 4.21-4.34 (m, 1H); 4.41-4.63 (m, 1H); 5.03-5.07 (m, 1H); 5.30-5.33 (m, 1H, NH); 5.45-5.49 (m, 1H); 6.18 (s, 1H); 8.22 (s, 1H); 8.56 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 25.9; 27.6; 28.3; 28.8; 30.2; 37.2; 51.4; 58.8; 80.8; 82.0; 83.6; 84.5; 87.2; 91.7; 115.5; 120.0; 141.1; 144.7; 151.0; 152.8; 155.5; 172.6. IR ν (cm⁻¹): Csp²-H (668-852); C-O (1157); C-N (1277); C=N (1601); C=O (1716); NH (2979). LC-MS (ESI-APCI): m/z [M+H]⁺ = 665.

2-tert-Butoxycarbonylamino-4-[6-(6-tert-butoxycarbonylamino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethylselanyl]-butyric acid tert-butyl ester (8a). Selenol (**6b**)^[34] (22 mg, 0.065 mmol, 1.0 eq) was introduced into the round bottom flask equipped by a condenser with dry ice and cooled at (-78°C), then ammoniac (gas) (10 mL) was added. Sodium (28 mg, 1.217 mmol, 18.7 eq) was added in one portion and the reaction mixture was stirred at (-78°C) for 30 minutes. Then the reaction mixture was slowly warmed to room temperature. Anhydrous ethanol (10 mL) was added and the reaction mixture was filtered on celite and the filtrate was evaporated. Finally, anhydrous N,N'-dimethylformamide (0.5 mL) was added to the crude. Compound (**5b**) (104 mg, 0.20 mmol, 2.9 eq) dissolved in anhydrous N,N'-dimethylformamide (0.5 mL) was dropwise added to the reaction mixture at room temperature. The reaction mixture was stirred at room temperature overnight, then under reflux for 4 hours. The reaction mixture was cooled and evaporated. The crude was purified by silica flash column chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a white solid (10 mg, 21%). EF: C₃₁H₄₈N₆O₉Se₁. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.40 (s, 3H); 1.50 (s, 27H); 1.63 (s, 3H); 1.98-2.39 (m, 2H); 2.84-2.93 (m, 1H); 3.60-3.80 (m, 3H); 4.22-4.34 (m, 1H); 4.51 (bs, 1H); 5.12 (dd, J₁ = 3.2; J₂ = 6.0, 1H); 5.41 (dd, J₁ = 3.2; J₂ = 6.0, 1H); 6.18 (d, J = 2.8, 1H); 8.20 (s, 1H); 8.86 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 25.3; 26.0; 27.1; 28.0; 28.1; 30.5; 36.5; 45.1; 81.8; 82.5; 84.0; 85.3; 86.4; 91.2; 114.9; 122.5; 143.8; 150.4; 152.3; 153.1; 162.4; 171.8. LC-MS (ESI-APCI): m/z [M+H]⁺ = 729.

[9-(6-Azidomethyl-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-yl)-9H-purin-6-yl]-carbamic acid tert-butyl ester (9b). Compound (**9a**)^[32] (4.1 g, 12.0 mmol, 1 eq) was dissolved in anhydrous tetrahydrofuran (40 mL), then NaH (1.0 g, 43.0 mmol, 3.6 eq) was added in one portion and the reaction mixture was stirred for 30 minutes. Di-tert-butyl dicarbonate (1.8 g, 22.0 mmol, 1.8 eq) diluted in anhydrous N,N'-dimethylformamide (10 mL) were added and the reaction mixture was stirred at 60°C overnight. The reaction mixture was diluted in ethyl acetate (100 mL) and washed with a solution of NH₄Cl_{sat} (2 x 100 mL) and with a solution of HCl 1N (2 x 100 mL). The organic layer was separated, dried over Na₂SO₄, filtered and evaporated. The crude was purified by silica flash column chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a white solid (5.1 g, 96%). EF: C₁₈H₂₄N₈O₅. ¹H NMR (300MHz, CDCl₃): δ (ppm) = 1.38 (s, 3H); 1.55 (s, 9H); 1.61 (s, 3H); 3.55 (d, J = 6, 2H); 4.38 (bs, 1H); 5.04 (dd, J₁ = 3.1, J₂ = 9.0, 1H); 5.43 (dd, J₁ = 3.1, J₂ = 9.0, 1H); 6.13 (s, 1H); 7.89 (s, 1H); 8.75 (s, 1H). ¹³C NMR (75 MHz, CDCl₃): δ (ppm) = 25.0; 27.5; 28.2; 52.0; 81.5; 83.5; 84.9; 85.3; 90.4; 114.6; 122.2; 143.9; 149.7; 150.2; 152.2; 156.7. LC-MS (ESI-APCI): m/z [M+H]⁺ = 433.

[9-(6-Aminomethyl-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-yl)-9H-purin-6-yl]-carbamic acid tert-butyl ester (10b). Compound (**9b**) (5.0 g, 12.0 mmol, 1 eq) dissolved in anhydrous methanol (60 mL) was degassed two times under argon. Then Pd/C (0.1 g, 1.1 mmol, 0.1 eq) was added in one portion and the reaction mixture was degassed two times under argon. Finally, hydrogen was added and the reaction mixture was degassed two times. The reaction mixture was stirred at room temperature overnight under pressure (3 bars), under hydrogen. The reaction mixture was filtered on celite and the filtrate was evaporated. The crude gave a yellow solid (3.9 g, 80%), and was used directly in the next step. EF: C₁₈H₂₆N₆O₅. ¹H NMR

(300MHz, CDCl₃): δ (ppm) = 1.28 (s, 3H); 1.43 (s, 9H); 1.54 (s, 3H); 3.31-3.38 (m, 2H); 4.44 (bs, 1H); 4.93 (dd, J₁ = 6.0, J₂ = 15.0, 1H); 5.30 (dd, J₁ = 6.0, J₂ = 15.0, 1H); 5.82 (d, J = 6, 1H); 7.82 (s, 1H); 8.39 (s, 1H). ¹³C NMR (75 MHz, CDCl₃): δ (ppm) = 25.8; 27.9; 29.1; 43.1; 79.7; 81.8; 83.3; 84.6; 92.8; 115.2; 121.2; 140.5; 149.4; 156.7; 157.5. LC-MS (ESI-APCI): m/z [M+H]⁺ = 407.

3-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]-amino]-propan-1-ol (11). Amine (**10a**) (1.05 g, 2.6 mmol, 1 eq) was dissolved in anhydrous tetrahydrofuran (10 mL) and stirred at room temperature, then bromopropanol (226 μL, 2.6 mmol, 1 eq) was dropwise added to the solution. Finally, N,N-Diisopropylethylamine (1.34 mL, 7.8 mmol, 3 eq) was dropwise added and the reaction mixture was stirred under reflux overnight. The reaction mixture was evaporated and the crude was purified by flash chromatography using a dichloromethane/methanol gradient (85/15 to 70/30) to give a yellow solid (390 mg, 40%). EF: C₁₆H₂₄N₆O₄. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.37 (s, 3H); 1.55 (s, 3H); 1.97-2.08 (m, 2H); 3.34-3.35 (m, 2H); 3.47 (t, J = 12, 1H); 3.58-3.68 (m, 2H); 3.72 (t, J = 12, 1H); 4.40 (bs, 1H); 4.87-4.91 (m, 1H); 5.23-5.28 (m, 1H); 5.81-5.88 (dd, J = 4, 1H); 8.07 (s, 1H); 8.72 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 25.3; 27.1; 30.8; 48.1; 52.2; 61.1; 82.1; 83.2; 85.2; 90.5; 114.2; 119.8; 139.8; 148.8; 152.9; 155.9. LC-MS (ESI-APCI): m/z [M+H]⁺ = 365.

3-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]-amino]-propionic acid methyl ester (12). Amine (**10a**) (500 mg, 1.63 mmol, 1 eq) was dissolved in anhydrous tetrahydrofuran (15 mL) and stirred at room temperature. Then methyl 3-bromopropionate (180 μL, 1.63 mmol, 1 eq) was dropwise added to the solution. Finally, N,N-Diisopropylethylamine (829 μL, 4.81 mmol, 3 eq) was dropwise added at room temperature and the reaction mixture was stirred under reflux overnight. The reaction mixture was evaporated and the crude was purified by flash chromatography using a dichloromethane/methanol gradient (85/15 to 70/30) to give a white solid (120 mg, 20 %). EF: C₁₇H₂₄N₆O₅. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.32 (s, 3H); 1.55 (s, 3H); 2.51 (t, J = 13.2, 2H); 2.86-2.96 (m, 4H); 3.58 (s, 3H); 4.38 (bs, 1H); 4.97-4.98 (dd, J₁ = 4.0; J₂ = 6.4, 1H); 5.34-5.36 (dd, J₁ = 3.6, J₂ = 6.4, 1H); 5.95 (d, J = 3.2, 1H); 6.38 (bs, 2H, NH₂); 7.86 (s, 1H); 8.27 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 18.9; 25.5; 27.5; 45.2; 50.7; 82.3; 83.4; 85.5; 91.0; 114.7; 120.6; 140.1; 149.3; 153.2; 155.8; 172.84. LC-MS (ESI-APCI): m/z [M+H]⁺ = 393.

3-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]-amino]-propionitrile (13). Amine (**10a**) (500 mg, 1.63 mmol, 1 eq) was dissolved in anhydrous tetrahydrofuran (30 mL) and stirred at room temperature, then 3-bromopropionitrile (271 μL, 3.26 mmol, 2 eq) was dropwise added to the solution. Finally, N,N-Diisopropylethylamine (844 μL, 4.89 mmol, 3 eq) was dropwise added and the reaction mixture was stirred under reflux overnight. The reaction mixture was evaporated and the crude was purified by flash chromatography using a dichloromethane/methanol gradient (85/15 to 70/30) to give a white solid (100 mg, 17 %). EF: C₁₆H₂₁N₇O₃. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.33 (s, 3H); 1.55 (s, 3H); 2.42 (t, J = 13.2, 2H); 2.78-2.96 (m, 4H); 4.30 (bs, 1H); 5.02 (dd, J₁ = 4; J₂ = 7.2, 1H); 5.41 (dd, J₁ = 3.6, J₂ = 7.6, 1H); 5.95 (d, J = 3.2, 1H); 6.38 (bs, 2H, NH₂); 7.86 (s, 1H); 8.27 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 18.9; 25.5; 27.5; 45.2; 50.7; 82.3; 83.4; 85.5; 91.0; 114.7; 118.8; 120.6; 140.1; 149.4; 153.2; 156.1. IR ν (cm⁻¹): C//N (2252). LC-MS (ESI-APCI): m/z [M+H]⁺ = 360.

3-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]-amino]-propionamide (14). Amine (**10a**) (200 mg, 0.654 mmol, 1 eq) was dissolved in anhydrous tetrahydrofuran (6 mL) and stirred at room temperature, then bromopropionamide (100 mg, 0.654 mmol, 1 eq) was added to the solution. Finally, N,N-Diisopropylethylamine (338 μL, 1.962 mmol, 3 eq) was dropwise added and the reaction mixture was stirred under reflux for 15 hours. The reaction mixture was evaporated

and the crude was purified by flash chromatography using a dichloromethane/methanol gradient (85/15 to 70/30) to give a yellow solid (173 mg, 70 %). EF: C₁₆H₂₃N₇O₄. ¹H NMR (400MHz, MeOD): δ (ppm) = 1.28 (s, 3H); 1.50 (s, 3H); 2.25-2.29 (m, 2H); 2.55-2.69 (m, 2H); 2.75-2.89 (m, 2H); 4.16-4.22 (m, 1H); 4.90-4.94 (m, 1H); 5.35-5.38 (m, 1H); 6.04 (s, 1H); 8.08 (s, 1H); 8.19 (s, 1H). ¹³C NMR (100 MHz, MeOH): δ (ppm) = 25.6; 27.6; 35.8; 46.6; 52.4; 83.3; 85.7; 87.8; 91.9; 115.4; 120.7; 141.3; 150.3; 154.1; 157.5; 177.5. LC-MS (ESI-APCI): m/z [M+H]⁺ = 378.

4-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]- (3-hydroxy-propyl)-amino]-2-tert-butoxycarbonylamino-butyrac acid tert-butyl ester (15a). Compound (11) (150 mg, 0.41 mmol, 1.2 eq) was dissolved in anhydrous methanol (3 mL) and stirred at room temperature. To the solution, the aldehyde (3) (94 mg, 0.34 mmol, 1.0 eq) dissolved in anhydrous methanol (3 mL) was dropwise added at room temperature. Then HOAc (21 μL, 0.34 mmol, 1 eq) was dropwise added, and NaBH₃CN (35 mg, 0.55 mmol, 1.6 eq) was added in one portion. The reaction mixture was stirred at room temperature overnight. Then the reaction mixture was diluted in ethyl acetate (50 mL) and washed with a solution of NaOH 1N (2 x 20 mL), the aqueous layer was extracted with dichloromethane (2 x 20 mL) dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a yellow solid (140 mg, 63%). EF: C₂₉H₄₇N₇O₈. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.32 (s, 3H); 1.38 (s, 18H); 1.54 (s, 3H); 1.54-1.61 (m, 2H); 1.78-1.88 (m, 1H); 2.04 (q, J = 6, 2H); 2.34-2.54 (m, 2H); 2.67-2.80 (m, 2H); 2.98 (bs, 1H); 3.48 (t, J = 6.4, 1H); 3.73 (t, J = 6.0, 1H); 4.17 (bs, 1H); 4.40 (bs, 1H); 5.08 (d, J = 3.6, 1H); 5.37 (d, J = 3.6, 1H); 5.92 (d, J = 3.2, 1H); 7.88 (s, 1H); 8.22 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 25.3; 27.2; 28.2; 28.5; 30.6; 35.3; 46.1; 51.0; 52.8; 55.9; 60.4; 81.5; 82.1; 83.4; 85.0; 85.2; 90.5; 114.2; 120.0; 139.8; 148.9; 149.2; 153.0; 156.0; 171.8. LC-MS (ESI-APCI): m/z [M+H]⁺ = 622.

4-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]- (3-methanesulfonyloxy-propyl)-amino]-2-tert-butoxycarbonylamino-butyrac acid tert-butyl ester (16a). Compound (15a) (100 mg, 0.16 mmol, 1eq) was dissolved in anhydrous tetrahydrofuran (1.4 mL) and stirred at room temperature. To the solution, methanesulfonyl chloride (13 μL, 0.16 mmol, 1 eq) and N,N-Diisopropylethylamine (28 μL, 0.16 mmol, 1 eq) were dropwise added successively. The reaction mixture was stirred at room temperature overnight. Then the reaction mixture was diluted in ethyl acetate (20 mL) and washed with a solution of NaOH 1N (2 x 10mL). The aqueous layer was extracted with dichloromethane (2 x 10mL), dried over Na₂SO₄, filtered and evaporated. The crude (70 mg) was used for the next step. EF: C₃₀H₄₉N₇O₁₀S₁. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.33 (s, 3H); 1.38 (d, J = 4, 18H); 1.59 (s, 3H); 1.63-1.75 (m, 2H); 1.87-1.92 (m, 2H); 2.37-2.41 (m, 2H); 2.48-2.55 (m, 3H); 2.67-2.75 (m, 1H); 2.95 (s, 3H); 4.01-4.21 (m, 2H); 4.23-4.27 (m, 1H); 4.95 (dd, J₁ = 3.6, J₂ = 6.8, 1H); 5.47 (d, J = 6.4, 1H); 5.52 (d, J = 6.4, 1H); 6.00 (s, 1H); 6.24 (bs, NH₂); 7.87 (bs, 1H); 8.26 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 25.6; 27.0; 28.1; 28.5; 29.6; 37.3; 50.6; 50.8; 52.8; 56.2; 68.4; 79.7; 81.9; 83.6; 83.9; 85.8; 91.0; 114.5; 120.4; 140.5; 149.3; 153.2; 155.6; 156.0; 171.8.

4-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]- (3-azido-propyl)-amino]-2-tert-butoxycarbonylamino-butyrac acid tert-butyl ester (16b). Derivative (16a) (70 mg, 0.10 mmol, 1.0 eq) was dissolved in anhydrous acetonitrile (8 mL) and stirred at room temperature. To the solution, sodium azide (9 mg, 0.15 mmol, 1.5 eq) was added in one portion. The reaction mixture was stirred under reflux for 2 days. The reaction mixture was evaporated and purified by flash chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a yellow solid (40 mg, 62%). EF: C₂₉H₄₆N₁₀O₇. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.33 (s, 3H); 1.38 (d, J = 3.2, 18H); 1.55 (s, 3H); 1.70-1.79 (m, 2H); 1.87-1.92 (m, 1H); 2.32-2.41 (m, 2H); 2.45-2.52 (m, 3H); 2.66-2.74 (m, 2H); 3.25 (t, J = 6.4, 2H); 4.04-4.10 (m, 1H); 4.23-4.28 (m, 1H); 4.96 (dd, J₁ = 3.2, J₂ = 6.8, 1H); 5.46 (bs, 1H, NH); 5.58 (d, J = 7.2, 1H); 5.99 (s, 1H); 6.09 (bs, NH₂); 7.85 (bs, 1H); 8.27 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 25.6; 27.3; 28.1; 28.5; 29.4; 29.8; 49.4; 50.8; 51.7; 52.7; 56.3; 79.6; 81.8; 83.6; 83.9; 85.8;

91.1; 114.5; 120.5; 140.3; 149.3; 153.2; 155.6; 155.9; 171.8. IR ν (cm⁻¹): N₃ (2101) LC-MS (ESI-APCI): m/z [M+H]⁺ = 647.

4-{{(3-Amino-propyl)-[6-(6-amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]-amino}-2-tert-butoxycarbonylamino-butyrac acid tert-butyl ester (16c). Compound (16b) (40 mg, 0.062 mmol, 1 eq) was dissolved in a mixture of tetrahydrofuran / water (236 μL/39 μL) and stirred at room temperature. To the solution, triphenylphosphine (32 mg, 0.124 mmol, 2 eq) was added in one portion and the reaction mixture was stirred under reflux overnight. Then the reaction mixture was cooled and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol (50/50) as eluent to give a yellow solid (23 mg, 60%). EF: C₂₉H₄₈N₈O₇. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.36 (s, 3H); 1.40 (s, 9H); 1.43(s, 9H); 1.59 (s, 3H); 1.65-1.76 (m, 1H); 1.81-2.01 (m, 1H); 2.47-2.72 (m, 6H); 2.78-2.82 (m, 2H); 2.96-3.04 (m, 2H); 4.23 (q, J = 3.6, 1H); 4.66 (bs, 1H); 4.99 (dd, J₁ = 3.6, J₂ = 6.4, 1H); 5.43 (dd, J₁ = 3.6, J₂ = 6.4, 1H); 5.92 (bs, NH); 6.00 (d, J = 3.2, 1H); 7.90 (s, 1H); 8.29 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 25.6; 27.5; 28.2; 28.2; 29.9; 38.4; 43.2; 52.3; 56.2; 79.9; 82.0; 83.9; 85.8; 87.6; 90.9; 114.8; 120.5; 143.3; 145.0; 148.1; 151.9; 156.1; 171.3. LC-MS (ESI-APCI): m/z [M+H]⁺ = 621.

4-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]- (2-methoxycarbonyl-ethyl)-amino]-2-tert-butoxycarbonylamino-butyrac acid tert-butyl ester (17a).Compound (12) (120 mg, 0.306 mmol, 1.2 eq) was dissolved in anhydrous methanol (1.3 mL) and stirred at room temperature. To the solution, the aldehyde (3) (70 mg, 0.255 mmol, 1.0 eq) dissolved in anhydrous methanol (1.3 mL) was added to the solution at room temperature. HOAc (15 μL, 0.255 mmol, 1.0 eq) and NaBH₃CN (26 mg, 0.408 mmol, 1.6 eq) were successively added in one portion and the reaction mixture was stirred at room temperature for 21 hours. The reaction mixture was diluted in ethyl acetate (100 mL) and washed with a solution of NaOH 1N (2 x 50 mL), the aqueous layer was extracted with dichloromethane (2 x 50 mL), dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a white solid (110 mg, 67%). EF: C₃₀H₄₇N₇O₉. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.37 (s, 3H); 1.39 (s, 18H); 1.55 (s, 3H); 1.71-1.92 (m, 2H); 2.18-2.38 (m, 2H); 2.55-2.77 (m, 4H); 2.87-2.91 (m, 2H); 3.58 (s, 3H); 4.07-4.22 (m, 1H); 4.30 (bs, 1H); 4.91-5.03 (m, 1H); 5.43-5.46 (m, 1H); 6.01 (bs, 1H); 7.90 (s, 1H); 8.28 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 25.6; 27.3; 28.2; 28.5; 29.5; 32.3; 49.7; 50.7; 51.8; 52.9; 56.0; 79.6; 81.8; 83.4; 83.9; 85.6; 91.1; 114.5; 120.8; 140.4; 149.3; 153.2; 155.6; 155.9; 171.9; 173.0. LC-MS (ESI-APCI): m/z [M+H]⁺ = 650.

4-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]- (2-carboxy-ethyl)-amino]-2-tert-butoxycarbonylamino-butyrac acid tert-butyl ester (17b). Compound (17a) (40 mg, 0.062 mmol, 1 eq) was dissolved in a mixture of methanol/tetrahydrofuran (1.4 mL/2.9 mL) and stirred at room temperature. To the solution, LiOH 1M (15mg in 7 mL of water) was dropwise added and the reaction mixture was stirred overnight at room temperature. Then the reaction mixture was evaporated and the crude was purified by reverse phase flash chromatography using water/acetonitrile (14/86) as eluent to give a white solid (26 mg, 67%). EF: C₂₉H₄₅N₇O₉. ¹H NMR (400MHz, D₂O): δ (ppm) = 1.37 (s, 9H); 1.45 (s, 9H); 1.67 (s, 3H); 1.92 (s, 3H); 2.25 (t, J = 8.4, 3H); 2.53-2.56 (m, 2H); 2.74-2.81 (m, 5H); 3.77-3.81 (bs, 1H); 4.46 (bs, 1H); 5.04 (bs, 1H); 5.51 (d, J = 6.4, 1H); 6.25 (d, J = 2.8, 1H); 8.27 (s, 1H); 8.32 (s, 1H). ¹³C NMR (100 MHz, D₂O): δ (ppm) = 23.6; 24.6; 26.2; 27.9; 34.2; 49.2; 50.7; 55.1; 57.7; 77.2; 78.8; 81.4; 82.9; 84.3; 89.7; 115.7; 118.8; 140.9; 145.1; 153.3; 155.9; 157.7; 171.3; 179.4. LC-MS (ESI-APCI): m/z [M+H]⁺ = 636.

4-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]- (2-carbamoyl-ethyl)-amino]-2-tert-butoxycarbonylamino-butyrac acid tert-butyl ester (18a). Compound (14) (145 mg, 0.385 mmol, 1.2 eq) was dissolved in anhydrous methanol (5 mL) and stirred at room temperature. To the solution, the aldehyde (3) (86 mg,

0.326 mmol, 1.0 eq) dissolved in anhydrous methanol (5 mL) was dropwise added. HOAc (19 μ L, 0.321 mmol, 1 eq) and NaBH₃CN (32 mg, 0.514 mmol, 1.6 eq) were successively added and the reaction mixture was stirred at room temperature overnight. Then the reaction mixture was diluted in ethyl acetate (30 mL) and washed with a solution of NaOH 1N (2 x 10 mL), the aqueous layer was extracted with dichloromethane (2 x 10 mL) dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a yellow solid (100 mg, 48 %). EF: C₂₉H₄₆N₈O₈. ¹H NMR (400MHz, MeOD): δ (ppm) = 1.36 (s, 3H); 1.45 (s, 18H); 1.61 (s, 3H); 2.00-2.08 (m, 2H); 2.73-2.82 (m, 2H); 2.97-3.05 (m, 2H); 3.08-3.21 (m, 2H); 3.63-3.80 (m, 2H); 4.10-4.15 (m, 1H); 4.38 (bs, 1H); 5.05-5.24 (m, 1H); 5.43-5.49 (m, 1H); 6.21 (bs, 1H); 7.84 (s, 1H); 8.26 (s, 1H). ¹³C NMR (100 MHz, MeOD): δ (ppm) = 25.8; 27.6; 28.3; 28.9; 30.8; 33.9; 41.5; 50.6; 51.5; 52.7; 56.2; 80.3; 82.7; 83.4; 84.1; 85.3; 92.2; 115.7; 120.9; 139.9; 150.1; 154.1; 157.5; 158.1; 172.2; 178.4. LC-MS (ESI-APCI): m/z [M+H]⁺ = 635.

4-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]-(2-cyano-ethyl)-amino]-2-tert-butoxycarbonylamino-butyl ester (19a). Compound (13) (1.200 g, 3.34 mmol, 1.2 eq) was dissolved in anhydrous methanol (4 mL) and stirred at room temperature. To the solution the aldehyde (3) (0.760 g, 2.79 mmol, 1.0 eq) dissolved in anhydrous methanol (3 mL) was dropwise added. HOAc (0.167 mL, 2.79 mmol, 1.0 eq) and NaBH₃CN (0.280 mg, 4.46 mmol, 1.6 eq) were successively added and the reaction mixture was stirred at room temperature overnight. Then the reaction mixture was diluted in ethyl acetate (100 mL) and washed with a solution of NaOH 1N (2 x 50 mL), the aqueous layer was extracted with dichloromethane (2 x 50 mL) dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a white solid (820 mg, 40%). EF: C₂₉H₄₄N₈O₇. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.37 (s, 3H); 1.39 (s, 18H); 1.55 (s, 3H); 1.71-1.92 (m, 2H); 2.18-2.38 (m, 2H); 2.55-2.77 (m, 4H); 2.87-2.91 (m, 2H); 4.07-4.22 (m, 1H); 4.30 (bs, 1H); 4.91-5.03 (m, 1H); 5.43-5.46 (m, 1H); 6.01 (bs, 1H); 7.90 (s, 1H); 8.28 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 25.5; 27.3; 28.2; 28.5; 29.8; 42.4; 46.4; 51.5; 53.4; 58.7; 79.5; 81.7; 82.3; 83.7; 85.3; 91.2; 114.5; 118.9; 120.5; 140.0; 149.3; 153.2; 155.6; 156.0; 172.0. LC-MS (ESI-APCI): m/z [M+H]⁺ = 617.

2-tert-Butoxycarbonylamino-4-[[6-(6-tert-butoxycarbonylamino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]-amino]-butyric acid tert-butyl ester (20a). Amine (10b) (1.1 g, 2.6 mmol, 0.9 eq) was dissolved in anhydrous methanol (3 mL) and stirred at room temperature. To the solution, the aldehyde (3) (0.75 g, 2.8 mmol, 1.0 eq) dissolved in anhydrous methanol (3 mL) was dropwise added. HOAc (0.2 mL, 2.8 mmol, 1.0 eq) and NaBH₃CN (0.3 g, 4.6 mmol, 1.7 eq) were successively added and the reaction mixture was stirred at room temperature overnight. Then the reaction mixture was diluted in ethyl acetate (50 mL) and washed with a solution of NaOH 1N (2 x 20 mL), the aqueous layer was extracted with dichloromethane (2 x 20 mL) dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol (95/5) as eluent to give a yellow solid (1.2 g, 71%). EF: C₃₁H₄₉N₇O₉. ¹H NMR (300MHz, CDCl₃): δ (ppm) = 1.35 (s, 3H); 1.44 (s, 9H); 1.46 (s, 18H); 1.61 (s, 3H); 1.70-2.08 (m, 2H); 2.61-2.95 (m, 3H); 3.30-3.75 (m, 1H); 4.21-4.30 (bs, 1H); 4.44-4.47 (m, 1H); 4.84-4.94 (m, 1H); 5.26-5.46 (m, 1H); 5.80-5.88 (m, 1H); 8.35 (s, 1H); 8.70 (s, 1H). ¹³C NMR (75 MHz, CDCl₃): δ (ppm) = 25.0; 27.3; 28.2; 32.3; 46.2; 51.3; 53.5; 81.7; 82.2; 83.9; 84.2; 85.3; 91.0; 114.6; 122.5; 142.3; 149.8; 150.5; 152.2; 156.8; 171.6. LC-MS (ESI-APCI): m/z [M+H]⁺ = 664.

2-tert-Butoxycarbonylamino-4-[[6-(6-tert-butoxycarbonylamino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]-methyl-amino]-butyric acid tert-butyl ester (21a). Amine (20b) (190 mg, 0.29 mmol, 1.1 eq) was dissolved in anhydrous methanol (1.5 mL) and stirred at room temperature. To the solution, formaldehyde (19 μ L, 0.26 mmol, 1.0 eq) was dropwise added. HOAc (16 μ L, 0.26 mmol, 1 eq) and NaBH₃CN (26 mg, 0.42 mmol, 1.6 eq) were successively added and the reaction mixture

was stirred at room temperature for 48 hours. The reaction mixture was diluted in ethyl acetate (50 mL) and washed with a solution of NaOH 1N (2 x 20 mL), the aqueous layer was extracted with dichloromethane (2 x 20 mL) dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a white solid (68 mg, 39%). EF: C₃₂H₅₁N₇O₉. ¹H NMR (300MHz, CDCl₃): δ (ppm) = 1.36 (s, 3H); 1.41 (s, 9H); 1.52 (s, 18H); 1.68 (s, 3H); 1.68-1.88 (m, 2H); 2.19 (s, 3H); 2.39-2.63 (m, 4H); 4.17 (bs, 1H); 4.35 (bs, 1H); 4.91 (bs, 1H); 5.43-5.53 (d, J = 6, 1H); 6.06 (s, 1H); 8.08 (s, 1H); 8.72 (s, 1H). ¹³C NMR (75 MHz, CDCl₃): δ (ppm) = 25.6; 27.4; 28.2; 29.5; 42.8; 53.1; 54.2; 59.6; 81.8; 82.4; 83.3; 84.0; 85.2; 91.2; 114.8; 122.7; 142.0; 149.5; 150.2; 150.4; 152.9; 171.7. LC-MS (ESI-APCI): m/z [M+H]⁺ = 678.

4-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]-butyl-amino]-2-tert-butoxycarbonylamino-butyl ester (22a). Amine (20a) (128 mg, 0.23 mmol, 1.2 eq) was dissolved in anhydrous methanol (3.3 mL) and stirred at room temperature. To the solution, butyraldehyde (17 μ L, 0.19 mmol, 1.0 eq) was dropwise added. HOAc (11 μ L, 0.19 mmol, 1.0 eq) and NaBH₃CN (19 mg, 0.30 mmol, 1.6 eq) were successively added and the reaction mixture was stirred at room temperature for 16 hours. Then the reaction mixture was diluted in ethyl acetate (50 mL) and washed with a solution of NaOH 1N (2 x 20 mL), the aqueous layer was extracted with dichloromethane (2 x 20 mL) dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a white solid (70mg, 60%). EF: C₃₀H₄₉N₇O₇. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 0.82 (t, J = 7.2, 3H); 1.25 (q, J = 7.2, 2H); 1.27-1.33 (m, 2H); 1.36 (s, 3H); 1.41 (s, 9H); 1.52 (s, 9H); 1.68 (s, 3H); 1.58-1.65 (m, 1H); 1.85-1.94 (m, 1H); 2.34-2.57 (m, 5H); 2.72-2.74 (m, 1H); 4.2 (bs, 1H); 4.33 (bs, 1H); 4.95 (bs, 1H); 5.46 (d, J = 3.6, 1H); 6.02 (s, 1H); 7.90 (s, 1H); 8.30 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 14.2; 20.6; 25.6; 27.4; 28.2; 28.6; 29.4; 37.5; 51.0; 53.2; 54.5; 56.2; 79.6; 81.8; 83.6; 84.0; 85.8; 91.2; 114.5; 120.5; 140.4; 149.4; 153.3; 155.7; 155.8; 171.8. LC-MS (ESI-APCI): m/z [M+H]⁺ = 620.

4-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]-(1H-imidazol-2-ylmethyl)-amino]-2-tert-butoxycarbonylamino-butyl ester (23a). Amine (20a) (147 mg, 0.261 mmol, 1.2 eq) was dissolved in anhydrous methanol (2 mL) and stirred at room temperature. Then 4(5) imidazolecarboxaldehyde (21 mg, 0.218 mmol, 1.0 eq) dissolved in anhydrous methanol (2 mL) was dropwise added. HOAc (13 μ L, 0.218 mmol, 1.0 eq) and NaBH₃CN (22 mg, 0.349 mmol, 1.6 eq) were successively added and the reaction mixture was stirred under reflux for 3 days. The reaction mixture was diluted in ethyl acetate (50 mL) and washed with a solution of NaOH 1N (2 x 20 mL), the aqueous layer was extracted with dichloromethane (2 x 20 mL) dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a white solid (45 mg, 32%). EF: C₃₀H₄₅N₉O₇. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.32 (s, 3H); 1.39 (s, 9H); 1.41(s, 9H); 1.54 (s, 3H); 1.69-1.78 (m, 1H); 1.90-2.00 (m, 1H); 2.63-2.84 (m, 4H); 3.59-3.75 (m, 2H); 4.28 (bs, 1H); 4.39 (bs, 1H); 4.96 (bs, 1H); 5.42 (bs, 1H, NH); 5.63 (bs, 1H); 6.07 (bs, 1H); 6.65 (bs, 2H, NH₂); 7.60 (bs, 1H); 8.18 (bs, 2H); 8.47 (bs, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 25.4; 27.2; 28.2; 28.6; 30.9; 49.0; 50.5; 52.0; 56.3; 80.4; 82.6; 83.5; 84.2; 85.4; 90.7; 114.7; 119.9; 128.8; 132.2; 140.8; 148.9; 153.0; 152.9; 155.8; 156.0; 171.8. LC-MS (ESI-APCI): m/z [M+H]⁺ = 644.

4-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]-pyridin-2-ylmethyl-amino]-2-tert-butoxycarbonylamino-butyl ester (24a). Amine (20a) (137 mg, 0.243mmol, 1.2eq) was dissolved in anhydrous methanol (3.4 mL) and stirred at room temperature. Then 2-pyridinecarboxaldehyde (19 μ L, 0.203 mmol, 1.0 eq) was dropwise added to the solution. HOAc (12 μ L, 0.203 mmol, 1.0 eq) and NaBH₃CN (21 mg, 0.325 mmol, 1.6 eq) were successively added and the reaction mixture was stirred under reflux for 3 days. The reaction mixture was diluted in ethyl acetate (50 mL) and washed with a solution of NaOH 1N (2 x 20mL), the aqueous layer was extracted with dichloromethane (2 x 20mL) dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a white solid (266

mg, 58%). EF: C32H46N8O7. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.33 (s, 3H); 1.35 (s, 9H); 1.42 (s, 9H); 1.57 (s, 3H); 1.69-1.84 (m, 1H); 1.98-2.00 (m, 1H); 2.31-2.84 (m, 4H); 3.49-3.79 (m, 2H); 4.05-4.16 (m, 1H); 4.30-4.37 (m, 1H); 4.88-4.94 (m, 1H); 5.35-5.44 (m, 1H); 6.00 (s, 1H); 7.04 (t, *J* = 12.4, 1H); 7.29 (d, *J* = 8.4, 1H); 7.47 (t, *J* = 8.4, 1H); 7.87 (d, *J* = 12.4, 1H); 8.31 (s, 1H); 8.42 (s, 1H). ¹³C RMN (100 MHz, CDCl₃): δ (ppm) = 25.5; 25.6; 28.1; 28.6; 29.5; 51.1; 52.5; 53.1; 56.4; 81.8; 82.2; 83.5; 84.2; 85.6; 90.9; 114.5; 120.5; 122.2; 123.3; 136.5; 140.0; 149.0; 149.3; 153.2; 155.7; 155.9; 159.0; 171.9. LC-MS (ESI-APCI): *m/z* [M+H]⁺ = 655.

4-[[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]-thiazol-2-ylmethyl-amino]-2-tert-butoxycarbonylamino-butyl ester (25a). Amine (20a) (126 mg, 0.224 mmol, 1.2 eq) was dissolved in anhydrous methanol (2 mL) and stirred at room temperature. Then 2-thiazolecarboxaldehyde (21 mg, 0.187 mmol, 1.0 eq) dissolved in anhydrous methanol (1 mL) was dropwise added to the solution. HOAc (11 μL, 0.187 mmol, 1.0 eq) and NaBH₃CN (18 mg, 0.300 mmol, 1.6 eq) were successively added and the reaction mixture was stirred at microwave for 2h15 at 80°C. The reaction mixture was diluted in ethyl acetate (20 mL) and washed with a solution of NaOH 1N (2 x 10 mL), the aqueous layer was extracted with dichloromethane (2 x 10 mL) dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a white solid (100 mg, 68 %). EF: C30H44N8O7S1. ¹H NMR (CDCl₃, 400MHz): δ (ppm) = 1.35 (s, 3H); 1.37 (s, 9H); 1.42 (s, 9H); 1.57 (s, 3H); 1.74-1.83 (m, 1H); 1.98-2.05 (m, 1H); 2.35-2.82 (m, 4H); 3.26-3.31 (d, *J* = 7.6, 2H); 4.05-4.15 (m, 1H); 4.30-4.37 (m, 1H); 4.92-5.02 (m, 1H); 5.35-5.44 (m, 1H); 6.00 (bs, 1H); 7.06 (s, 1H); 7.89 (s, 1H); 8.26 (s, 1H); 8.70 (s, 1H). ¹³C NMR (100 MHz, CDCl₃): δ (ppm) = 25.5; 27.1; 28.1; 28.5; 29.4; 50.7; 53.1; 56.1; 59.1; 79.5; 81.7; 82.7; 84.0; 85.6; 90.8; 114.6; 116.1; 120.2; 140.0; 149.3; 152.5; 153.2; 154.7; 155.9; 158.0; 171.8. IR ν (cm⁻¹): Csp²-H (667-873); C-O (1077); C-N (1295); C=N (1599); C=O (1647); NH (2931); NH₂ (3197). LC-MS (ESI-APCI): *m/z* [M+H]⁺ = 661.

[6-(6-tert-Butoxycarbonylamino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]-3-tert-butoxycarbonyl-3-tert-butoxycarbonylamino-propyl)-dimethyl-ammonium (26a). Compound (21a) (37 mg, 0.06 mmol, 1eq) was dissolved in anhydrous ethanol (1 mL) and stirred at room temperature. To the solution, methyl iodide (34 μL, 0.54 mmol, 10 eq) was dropwise added. The reaction mixture was stirred at room temperature for 10 days. Then the reaction mixture was evaporated and 20 mL of Et₂O was added. We obtained a precipitate and then it was filtered to give a white powder (15 mg, 39%). EF: C33H54N7O9. ¹H NMR (300MHz, CDCl₃): δ (ppm) = 1.32 (s, 3H); 1.41 (s, 18H); 1.52 (s, 9H); 1.59 (s, 3H); 1.82-2.09 (m, 2H); 3.06 (s, 6H); 3.37-3.47 (m, 4H); 3.85-3.90 (m, 1H); 4.69 (bs, 1H); 5.15 (d, *J* = 5.4, 1H); 5.28 (d, *J* = 5.4, 1H); 6.35 (s, 1H); 8.40 (s, 1H); 8.57 (s, 1H). ¹³C NMR (75 MHz, CDCl₃): δ (ppm) = 25.6; 27.4; 28.2; 28.5; 29.9; 52.5; 55.3; 59.9; 65.0; 79.9; 80.5; 82.1; 83.4; 84.5; 90.9; 116.3; 120.6; 144.2; 147.3; 151.2; 152.5; 160.1; 169.7. LC-MS (ESI-APCI): *m/z* [M+H]⁺ = 693.

[3-[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethylimino]-1-methyl-propyl]-carbamic acid tert-butyl ester (27a). Aldehyde (3) (178 mg, 0.65 mmol, 1 eq) was dissolved in anhydrous methanol (0.5 mL) and stirred at room temperature. To the solution, amine (10a) (200 mg, 0.65 mmol, 1 eq) dissolved in anhydrous methanol (0.5 mL) was dropwise added at room temperature. Then HOAc (78 μL, 1.3 mmol, 2 eq) was dropwise added and the reaction mixture was stirred at room temperature for 5 days. The reaction mixture was evaporated and the crude was purified by silica flash column chromatography using a dichloromethane/methanol (94/6) as eluent to give a white solid (87 mg, 24%). EF: C26H39N7O7. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.36 (s, 3H); 1.43 (s, 18H); 1.60 (s, 3H); 2.75-2.81 (dd, *J*₁ = 4.8, *J*₂ = 13.6, 1H); 2.97-3.02 (dd, *J*₁ = 4.8, *J*₂ = 14.0, 1H); 3.55 (d, *J* = 4.2, 2H); 4.35-4.37 (m, 1H); 4.44-4.47 (m, 1H); 4.99-5.02 (dd, *J*₁ = 3.6, *J*₂ = 6.4, 1H); 5.38-5.40 (dd, *J*₁ = 2.4; *J*₂ = 6.4, 1H); 5.55 (m, 1H); 6.08 (d, *J* = 2.4, 1H); 6.95 (bs, 2H, NH₂); 7.95 (s, 1H); 8.24 (s, 1H). ¹³C NMR (300 MHz, CDCl₃): δ (ppm) = 25.5; 27.3; 28.1; 28.6; 37.4; 50.7; 52.5; 80.1; 82.1; 82.4; 84.3; 85.8; 90.9; 115.1; 119.8; 140.0; 149.1; 152.5; 152.2; 155.9; 170.4; 175.8. LC-MS (ESI-APCI): *m/z* [M+H]⁺ = 562.

N-[6-(6-Amino-purin-9-yl)-2,2-dimethyl-tetrahydro-furo[3,4-d][1,3]dioxol-4-ylmethyl]-2-tert-butoxycarbonylamino-succinamic acid tert-butyl ester (28a). Acid Boc-Asp-OtBu (1) (1.02 g, 3.51 mmol, 1.0 eq) was suspended in dichloromethane (10 mL) and stirred at room temperature. Then N-hydroxysuccinimide (0.560 g, 4.87 mmol, 1.4 eq) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (0.780 g, 4.06 mmol, 1.1 eq) were successively added in one portion to the solution. Then the reaction mixture was stirred at room temperature overnight. The reaction mixture was diluted in water (3 x 10 mL) and extracted with dichloromethane (3 x 10 mL), dried over Na₂SO₄, filtered and evaporated. The crude (1.06 g, 2.75 mmol, 1.0 eq) was dissolved in anhydrous tetrahydrofuran (11 mL) and stirred at room temperature. Derivative (10a) (0.857 g, 2.8 mmol, 1.02 eq) and N,N-Diisopropylethylamine (474 μL, 2.75 mmol, 1.0eq) were successively added to the solution and the reaction mixture was stirred under reflux for 3 hours. Then the reaction mixture was diluted in ethyl acetate (50 mL) and washed with water (2 x 20 mL) and a solution of NH₄Cl sat (2 x 20 mL), dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol (98/2) as eluent to give a yellow solid (1.1 g, 72%). EF: C26H39N7O8. ¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.39 (s, 3H); 1.40 (s, 18H); 1.53 (s, 3H); 2.87 (bs, 2H); 3.17 (d, *J* = 12, 1H); 4.08 (t, *J* = 10.4, 1H); 4.41 (bs, 2H); 4.75 (d, *J* = 5.2, 1H); 5.76 (d, *J* = 4.4, 1H); 5.84 (d, *J* = 5.6, 1H); 6.60 (bs, 2H, NH₂); 7.82 (s, 1H); 8.29 (s, 1H); 8.45 (bs, 1H, NH). ¹³C NMR (400 MHz, CDCl₃): δ (ppm) = 25.4; 27.6; 38.4; 41.1; 53.6; 79.7; 81.6; 82.1; 82.2; 83.4; 92.7; 114.7; 120.9; 140.5; 148.8; 153.0; 155.7; 156.3; 170.7; 170.8. LC-MS (ESI-APCI): *m/z* [M+H]⁺ = 578.

General method for compounds (7a-8a, 15b; 16d-28a)

HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound (7a-8a, 15b; 16c-28a) (1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 mL/minutes. Final compounds were quantitatively obtained as a yellowish sticky solid.

2-Amino-4-[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethoxy]-butyric acid (7b). EF: C14N20N6O6. ¹H NMR (500MHz, D₂O): δ (ppm) = 2.08-2.47 (m, 2H); 2.75-2.85 (m, 1H); 3.35-3.70 (m, 3H); 3.75-3.79 (m, 1H); 3.88-4.10 (m, 1H); 4.15-4.22 (m, 1H); 4.40-4.42 (m, 1H); 6.17 (d, *J* = 4.5, 1H); 8.38 (s, 1H); 8.44 (s, 1H). ¹³C NMR (125 MHz, D₂O): δ (ppm) = 41.3; 51.0; 63.2; 67.5; 71.2; 73.1; 78.3; 90.2; 122.7; 143.9; 149.4; 150.9; 155.9; 171.8. IR ν (neat, cm⁻¹): Csp²-H (700-854); C-O (1097); C-N (1251); C=N (1507); C=O (1716); NH (2982); NH₂ (3341). LC-MS (ESI-APCI): *m/z* [M+H]⁺ = 369 (100%). HPLC purity (method A): tr = 1.61, ≥98%.

2-Amino-4-[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethylselenanyl]-butyric acid (8b). EF: C14H20N6O5Se1. ¹H NMR (500MHz, D₂O): δ (ppm) = 2.38-2.54 (m, 2H); 2.99-3.04 (m, 1H); 3.66-3.76 (m, 2H); 3.88-4.15 (m, 3H); 4.28-4.31 (m, 1H); 4.45-4.53 (m, 1H); 6.20 (d, *J* = 6.3, 1H); 8.39 (s, 1H); 8.45 (s, 1H). ¹³C NMR (125 MHz, D₂O): δ (ppm) = 27.4; 27.9; 32.6; 52.3; 71.3; 73.4; 80.8; 89.4; 119.5; 141.1; 148.8; 152.7; 156.0; 171.2. IR ν (neat, cm⁻¹): Csp²-H (710); C-O (1009); C-N (1232); C=N (1410); C=O (1685); NH (2999). LC-MS (ESI-APCI): *m/z* [M+H]⁺ = 433 (100%). HPLC purity (method A): tr = 1.55, ≥98%.

2-Aminomethyl-5-(6-amino-purin-9-yl)-tetrahydro-furan-3,4-diol (10c). EF: C10H14N6O3. ¹H NMR (500MHz, D₂O): δ (ppm) = 3.51 (d, *J* = 8.4, 1H); 4.41 (q, *J* = 4.8, 1H); 4.50 (t, *J* = 3.6, 1H); 4.88 (t, *J* = 5.2, 1H); 6.17 (d,

$J = 5.2, 1\text{H}$); 8.47 (d, $J = 4.4, 2\text{H}$). ^{13}C RMN (100 MHz, D_2O): δ (ppm) = 41.2; 71.2; 73.5; 80.6; 89.7; 119.3; 143.6; 144.4; 145.1; 149.8. LC-MS (ESI-APCI): m/z $[\text{M}+\text{H}]^+ = 267$ (100%). HPLC purity (method A): $\text{tr} = 2.73, \geq 98\%$. HRMS (m/z): $[\text{M}+\text{H}]^+ = 267, 1127$, Found: 267.1208.

2-Amino-4-[[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethyl]-(3-hydroxy-propyl)-amino]-butyric acid (15b). EF: C17H27N7O6. ^1H NMR (500MHz, D_2O): δ (ppm) = 1.85-2.07 (m, 1H); 2.18-2.42 (m, 3H); 3.21 (t, $J = 7.2, 1\text{H}$); 3.33-3.41 (m, 2H); 3.53-3.67 (m, 3H); 4.01-4.05 (m, 1H); 4.11-4.25 (m, 1H); 4.41-4.47 (m, 2H); 4.85-4.88 (m, 2H); 6.14 (d, $J = 6.0, 1\text{H}$); 8.45 (s, 2H). ^{13}C NMR (125 MHz, D_2O): δ (ppm) = 26.4; 27.6; 44.7; 49.2; 51.3; 55.4; 60.1; 71.4; 73.4; 79.6; 89.8; 119.4; 143.6; 144.4; 148.1; 150.4; 171.42. LC-MS (ESI-APCI): m/z $[\text{M}+\text{H}]^+ = 426$ (100%). HPLC purity (method A): $\text{tr} = 1.61, \geq 98\%$. HRMS (m/z): $[\text{M}+\text{H}]^+ = 426, 20986$, Found: 426.21067.

2-Amino-4-[(3-amino-propyl)-[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethyl]-amino]-butyric acid (16d). EF: C17H28N8O5. ^1H NMR (400MHz, D_2O): δ (ppm) = 2.19-2.45 (m, 4H); 3.48-3.67 (m, 4H); 3.73-3.83 (m, 4H); 3.99-4.03 (m, 1H); 4.51 (bs, 2H); 4.83-4.87 (m, 1H); 6.17 (d, $J = 2.6, 1\text{H}$); 8.46 (d, $J = 2.8, 2\text{H}$). ^{13}C NMR (100 MHz, D_2O): δ (ppm) = 24.5; 29.6; 43.3; 44.4; 51.2; 55.5; 71.6; 73.0; 78.2; 90.1; 119.4; 143.6; 144.7; 148.1; 150.0; 171.4. LC-MS (ESI-APCI): m/z $[\text{M}+\text{H}]^+ = 426$ (100%). HPLC purity (method A): $\text{tr} = 1.71, \geq 98\%$. HRMS ($\text{M}+\text{H})^+ = \text{Calculated} : 425, 2255$, Found: 425.2257.

2-Amino-4-[[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethyl]-(2-carboxy-ethyl)-amino]-butyric acid (17c). EF: C17H25N7O7. ^1H NMR (400MHz, D_2O): δ (ppm) = 2.18-2.26 (m, 1H); 2.33-2.41 (m, 1H); 3.50-3.68 (m, 5H); 3.72-3.79 (m, 2H); 3.83-3.87 (m, 2H); 4.49 (t, $J = 5.2, 1\text{H}$); 4.59-4.61 (m, 1H); 4.93 (t, $J = 5.2, 1\text{H}$); 6.14 (d, $J = 4.4, 1\text{H}$); 8.47 (d, $J = 4.4, 2\text{H}$). ^{13}C NMR (100 MHz, D_2O): δ (ppm) = 24.7; 28.4; 34.1; 52.2; 51.8; 55.5; 71.7; 72.9; 78.4; 90.1; 119.4; 143.7; 145.0; 148.2; 150.2; 172.3; 173.6. LC-MS (ESI-APCI): m/z $[\text{M}-\text{H}]^- = 439$ (100%). HPLC purity (method A): $\text{tr} = 1.52, \geq 98\%$. HRMS ($\text{M}-\text{H})^- = \text{Calculated} : 439.18155$, Found: 439.17971.

2-Amino-4-[[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethyl]-(2-carbamoyl-ethyl)-amino]-butyric acid (18b). EF: C17H26N8O6. ^1H NMR (400MHz, D_2O): δ (ppm) = 2.18-2.26 (m, 1H); 2.33-2.41 (m, 1H); 3.50-3.68 (m, 5H); 3.72-3.79 (m, 2H); 3.83-3.87 (m, 2H); 4.45-4.57 (m, 2H); 4.93 (t, $J = 5.2, 1\text{H}$); 6.14 (d, $J = 4.4, 1\text{H}$); 8.46 (d, $J = 4.4, 2\text{H}$). ^{13}C NMR (125 MHz, D_2O): δ (ppm) = 24.7; 28.1; 45.0; 49.1; 51.8; 55.4; 71.7; 72.9; 78.2; 90.0; 119.4; 143.7; 144.5; 148.1; 150.1; 172.1; 174.3. LC-MS (ESI-APCI): m/z $[\text{M}+\text{H}]^+ = 439$ (100%). HPLC purity (method A): $\text{tr} = 2.32, \geq 98\%$. HRMS (m/z): $[\text{M}+\text{H}]^+ = 439, 20481$, Found: 439.20503.

2-Amino-4-[[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethyl]-(2-carbamimidoyl-ethyl)-amino]-butyric acid (19b). Compound (19a) (200 mg, 0.33 mmol, 1 eq) was dissolved in a solution of hydrogen chloride in methanol (2.9 mL, 66.00 mmol, 200 eq) and stirred at 0°C . Then water (119 μL , 6.60 mmol, 20 eq) was dropwise added to the solution and the reaction mixture was gently warmed to room temperature. The reaction mixture was stirred overnight, then the reaction mixture was diluted in water (50 mL) and washed with dichloromethane (2 x 20 mL), the aqueous layer was evaporated. Water (5 mL) was added to the remaining crude powder and it was neutralised with a solution of NaHCO_3 sat (1 mL). After, to the crude, a solution of NH_4Cl sat (12 mg in 474 μL of water) was dropwise added and the mixture was stirred at 37°C for 20 hours. The crude mixture was purified by reverse phase flash chromatography using a water/acetonitrile as eluent to give a white solid (40 mg, 28 %). EF : C17H27N9O5. ^1H NMR (300MHz, D_2O): δ (ppm) = 1.65-1.90 (m, 2H); 2.59-2.71 (m, 4H); 2.84-3.03 (m, 4H);

3.31-3.35 (m, 1H); 4.23-4.29 (m, 1H); 4.35 (t, $J = 5.1, 1\text{H}$); 4.80 (t, $J = 5.7, 1\text{H}$); 6.03 (d, $J = 4.8, 1\text{H}$); 8.25 (s, 1H); 8.30 (s, 1H). ^{13}C NMR (D_2O , 125 MHz): δ (ppm) = 29.9; 40.4; 47.9; 50.6; 53.1; 55.0; 72.2; 78.9; 81.9; 89.8; 119.7; 142.1; 146.3; 148.4; 151.6; 156.3; 171.1. LC-MS (ESI-APCI): m/z $[\text{M}+\text{H}]^+ = 438$ (100%). HPLC purity (method B): $\text{tr} = 2.37, \geq 98\%$. HRMS ($\text{M}+\text{H})^+ = \text{Calculated} : 438, 2208$, Found: 438.2262.

2-Amino-4-[[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethyl]-amino]-butyric acid (20b). EF: C14H21N7O5. ^1H NMR (400MHz, D_2O): δ (ppm) = 2.24-2.32 (m, 1H); 2.36-2.46 (m, 1H); 3.47-3.64 (m, 4H); 3.72-3.75 (m, 1H); 3.80-3.86 (m, 1H); 4.07-4.09 (m, 1H); 4.46-4.52 (m, 1H); 6.14 (d, $J = 3.6, 1\text{H}$); 8.43 (bs, 2H). ^{13}C NMR (100 MHz, D_2O): δ (ppm) = 24.3; 48.9; 50.4; 55.4; 71.5; 73.0; 78.0; 90.1; 119.3; 143.6; 144.4; 148.1; 149.9; 170.5. LC-MS (ESI-APCI): m/z $[\text{M}+\text{H}]^+ = 368$ (100%). HPLC purity (method B): $\text{tr} = 2.98, \geq 98\%$. HRMS ($\text{M}+\text{H})^+ = \text{Calculated} : 368, 1680$, Found: 368.1685.

2-Amino-4-[[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethyl]-methyl-amino]-butyric acid (21b). EF: C15H23N7O5. ^1H NMR (400MHz, D_2O): δ (ppm) = 2.24-2.32 (m, 1H); 2.36-2.46 (m, 1H); 2.92 (s, 3H); 3.47-3.64 (m, 4H); 3.72-3.75 (m, 1H); 3.80-3.86 (m, 1H); 4.07-4.09 (m, 1H); 4.46-4.52 (m, 1H); 6.14 (d, $J = 3.6, 1\text{H}$); 8.43 (bs, 2H). ^{13}C NMR (100 MHz, D_2O): δ (ppm) = 24.6; 38.9; 51.0; 51.5; 55.6; 71.3; 73.1; 78.3; 90.2; 119.5; 143.7; 144.5; 148.2; 150.1; 171.8. LC-MS (ESI-APCI): m/z $[\text{M}+\text{H}]^+ = 382$ (100%). HPLC purity (method A): $\text{tr} = 1.41, \geq 98\%$. HRMS ($\text{M}+\text{H})^+ = \text{Calculated} : 382, 1833$, Found: 382.1847.

2-Amino-4-[[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethyl]-butyl-amino]-butyric acid (22b). EF: C18H28N7O5. ^1H NMR (400MHz, D_2O): δ (ppm) = 0.64 (bs, 3H); 1.08-1.13 (m, 3H); 1.44 (bs, 2H); 1.84-2.00 (m, 1H); 2.07-2.16 (m, 1H); 2.99-3.08 (m, 2H); 3.26-3.32 (m, 2H); 3.48 (bs, 1H); 3.53-3.63 (m, 2H); 4.29 (bs, 2H); 5.96 (d, $J = 3.0, 1\text{H}$); 8.22 (s, 2H). ^{13}C NMR (125 MHz, D_2O): δ (ppm) = 12.7; 19.1; 24.8; 38.5; 48.9; 52.7; 54.9; 62.5; 71.8; 72.8; 78.4; 89.8; 119.3; 143.0; 146.2; 148.3; 151.3; 172.8. LC-MS (ESI-APCI): m/z $[\text{M}+\text{H}]^+ = 424$ (100%). HPLC purity (method A): $\text{tr} = 3.41, \geq 98\%$. HRMS ($\text{M}+\text{H})^+ = \text{Calculated} : 424.23029$, Found: 424.23079.

2-Amino-4-[[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethyl]-(1H-imidazol-2-ylmethyl)-amino]-butyric acid (23b). FB: C18H25N9O5. ^1H NMR (400MHz, D_2O): δ (ppm) = 2.18-2.26 (m, 1H); 2.32-2.43 (m, 1H); 3.42-3.57 (m, 4H); 3.65-3.78 (m, 2H); 3.95-4.10 (m, 2H); 4.39-4.45 (m, 1H); 4.57-4.63 (m, 1H); 6.07 (d, $J = 3.2, 1\text{H}$); 7.66 (s, 1H); 8.34 (d, $J = 2.4, 2\text{H}$); 8.71 (s, 1H). ^{13}C NMR (100 MHz, D_2O): δ (ppm) = 26.6; 49.2; 50.9; 52.9; 57.1; 73.5; 75.1; 79.8; 92.3; 121.3; 124.8; 137.9; 145.5; 146.4; 148.9; 150.0; 152.3; 172.8. LC-MS (ESI-APCI): m/z $[\text{M}+\text{H}]^+ = 448$ (100%). HPLC purity (method A): $\text{tr} = 3.74, \geq 98\%$. HRMS ($\text{M}+\text{H})^+ = \text{Calculated} : 448, 2050$, Found: 448.2048.

2-Amino-4-[[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethyl]-pyridin-2-ylmethyl-amino]-butyric acid (24b). EF: C20H26N8O5. ^1H NMR (400MHz, D_2O): δ (ppm) = 2.20-2.41 (m, 4H); 3.27-3.45 (m, 2H); 3.55-3.63 (m, 2H); 4.12 (t, $J = 4.4, 1\text{H}$); 4.41-4.50 (m, 2H); 4.68 (t, $J = 4.4, 1\text{H}$); 6.06 (d, $J = 3.2, 1\text{H}$); 7.72 (t, $J = 7.6, 1\text{H}$); 7.86 (d, $J = 7.2, 1\text{H}$); 8.25 (t, $J = 8.4, 1\text{H}$); 8.32 (s, 1H); 8.36 (s, 1H); 8.50 (d, $J = 8.4, 1\text{H}$). ^{13}C NMR (100 MHz, D_2O): δ (ppm) = 24.5; 50.7; 51.0; 55.7; 56.3; 71.8; 73.4; 79.8; 90.2; 119.6; 126.6; 127.5; 143.7; 144.5; 145.2; 148.1; 149.3; 150.5; 170.8; 173.6. LC-MS (ESI-APCI): m/z $[\text{M}+\text{H}]^+ = 459$ (100%). HPLC purity (method A): $\text{tr} = 3.92, \geq 98\%$. HRMS ($\text{M}+\text{H})^+ = \text{Calculated} : 459, 2098$, Found: 459.2096.

2-Amino-4-[[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethyl]-thiazol-2-ylmethyl-amino]-butyric acid (25b). EF : C₁₈H₂₄N₈O₅S₁. ¹H NMR (400MHz, D₂O): δ (ppm) = 2.20-2.48 (m, 2H); 3.07-3.11 (m, 2H); 3.42-3.58 (m, 2H); 3.64-3.83 (m, 2H); 4.04-4.15 (m, 1H); 4.44-4.49 (m, 1H); 4.61 (bs, 1H); 4.70 (t, *J* = 4, 1H); 6.09 (d, *J* = 3.6, 1H); 7.78 (s, 1H); 8.36 (d, *J* = 7.2, 2H); 8.86 (s, 1H). ¹³C NMR (100 MHz, D₂O): δ (ppm) = 24.5; 50.4; 52.1; 55.6; 58.4; 71.8; 73.4; 78.2; 90.5; 119.5; 124.8; 143.7; 144.6; 148.2; 150.1; 157.0; 158.5; 170.6. LC-MS (ESI-APCI): *m/z* [M+H]⁺ = 465 (100%). HPLC purity (method A): *tr* = 4.21, ≥98%. HRMS (M+H)⁺ = Calculated : 465,1663, Found : 465,1673.

(3-Amino-3-carboxy-propyl)-[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethyl]-dimethyl-ammoniu iodide (26b). EF : C₁₆H₂₆N₇O₅. ¹H NMR (300MHz, MeOD): δ (ppm) = 2.38-2.45 (m, 2H); 3.10 (s, 6H); 3.44-3.54 (m, 2H); 3.72- 3.76 (m, 2H); 3.83-3.96 (m, 2H); 4.39-4.47 (m, 2H); 6.03 (s, 1H); 8.32 (d, *J* = 6, 2H). ¹³C NMR (100 MHz, D₂O): δ (ppm) = 26.8; 52.0; 55.3; 61.9; 65.9; 73.2; 76.7; 79.8; 91.2; 120.0; 143.5; 148.1; 152.9; 159.5; 169.8. LC-MS (ESI-APCI): *m/z* [M]⁺ = 396 (100%). HPLC purity (method A): *tr* = 1.35, ≥98%. HRMS (M+H)⁺ = Calculated : 396,1989, Found : 396,1995.

2-Amino-4-[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethylimino]-butyric acid (27b). EF: C₁₄H₂₀N₇O₅. ¹H NMR (300MHz, MeOD): δ (ppm) = 3.02 (d, *J* = 5.7, 2H); 3.39-3.54 (m, 2H); 3.66 (d, *J* = 6.0, 1H); 3.91-3.99 (m, 1H); 4.07 (bs, 1H); 4.22 (bs, 1H); 4.29 (bs, 1H); 4.74-4.84 (m, 1H); 6.11 (d, *J* = 3.3, 1H); 8.47 (s, 1H); 8.57 (s, 1H). ¹³C NMR (100 MHz, MeOD): δ (ppm) = 35.1; 50.7; 53.5; 72.5; 75.7; 85.3; 91.1; 120.2; 143.9; 145.2; 146.8; 149.5; 170.5; 173.9. LC-MS (ES+APCI): (M+K)⁺ = 390 (100%). HPLC purity (method B): *tr* = 2.46, ≥98%.

2-Amino-N-[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethyl]-succinamic acid (28b). EF : C₁₄H₁₉N₇O₆. ¹H NMR (400MHz, D₂O): δ (ppm) = 2.96-3.02 (m, 2H); 3.38-3.49 (m, 2H); 3.60 (d, *J* = 5.2, 1H); 3.91-3.99 (m, 1H); 4.07-4.12 (m, 1H); 4.32-4.37 (m, 1H); 6.08 (d, *J* = 4.8, 1H); 8.37 (s, 1H); 8.54 (s, 1H). ¹³C NMR (100 MHz, D₂O): δ (ppm) = 34.1; 40.3; 49.2; 70.5; 73.3; 82.7; 88.5; 119.7; 143.2; 144.2; 148.5; 149.5; 170.5; 170.8. LC-MS (ES+APCI): (M+H)⁺ = 382 (100%). HPLC purity (method A): *tr* = 1.41, 100%.

2-Amino-4-[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethanesulfonyl]-butyric acid (29). Hydrogen peroxid (50% in water) (50 μL, 0.781 mmol, 10 eq) was dropwise added to S-(5-Adenosyl)-L-homocysteine (SAH) (30 mg, 0.078 mmol, 1 eq) and the reaction mixture was stirred at room temperature for 36 hours. Then the reaction mixture was evaporated and purified by reverse phase flash chromatography using a water /acetonitrile as eluent to give yellow solid (20 mg, 66%). EF : C₁₄H₂₀N₆O₆S₁. ¹H NMR (400MHz, D₂O): δ (ppm) = 2.23-2.37 (m, 2H); 2.60-2.82 (m, 2H); 2.98-3.27 (m, 2H); 3.41-3.55 (m, 1H); 3.77-3.91 (m, 1H); 4.44-4.58 (m, 2H); 6.18 (d, *J* = 3.2, 1H); 8.19 (s, 1H); 8.26 (s, 1H). ¹³C NMR (100 MHz, D₂O): δ (ppm) = 23.9; 47.0; 53.2; 54.3; 72.7; 77.7; 80.2; 88.5; 118.8; 140.4; 148.6; 151.0; 154.9; 173.1. IR ν (neat, cm⁻¹): S=O (1020) LC-MS (ES+APCI): (M+H)⁺ = 401 (100%). HPLC purity (method B): *tr* = 1.41, 100%.

2-Amino-4-[5-(6-amino-purin-9-yl)-3,4-dihydroxy-tetrahydro-furan-2-ylmethanesulfonyl]-butyric acid (30). Ammonium molybdate (48 mg, 0.04 mmol, 0.3 eq) was dissolved in water (1.5 mL) and stirred at room temperature. Perchloric acid (0.4 mL) was dropwise added to the solution and the reaction mixture was heated at 100°C for 5 minutes. Then the reaction mixture was cooled at room temperature and SAH (50 mg, 0.130 mmol, 1.0 eq) was added in one portion to the solution. The reaction mixture

was cooled at 0°C and hydrogen peroxid (50% in water) (1 mL) was dropwise added. The reaction mixture became yellow, and then it was gently warmed at room temperature and stirred for 24 hours. Propylamine (3 mL) was added and pH of the reaction mixture changed from 6 to 10. Then, ethanol (10 mL) was added, but no precipitate was formed, consequently the reaction mixture was evaporated and purified by reverse phase flash chromatography using a water /acetonitrile as eluent to give yellow solid (40 mg, 74%). EF : C₁₄H₂₀N₆O₇S₁. ¹H NMR (400MHz, D₂O): δ (ppm) = 1.55-1.67 (m, 1H); 2.23-2.30 (m, 1H); 2.86 (t, *J* = 7.8, 1H); 3.13-3.17 (m, 1H); 3.26-3.43 (m, 2H); 3.69 (dd, *J*₁ = 2.7, *J*₂ = 14.8, 1H); 3.87 (t, *J* = 6.8, 1H); 3.92-3.98 (m, 1H); 4.44 (t, *J* = 4.8, 1H); 6.06 (d, *J* = 4.8, 1H); 8.36 (s, 2H). ¹³C NMR (100 MHz, D₂O): δ (ppm) = 22.3; 49.9; 52.2; 55.3; 72.9; 73.4; 78.2; 89.8; 119.0; 143.2; 148.1; 150.5; 162.7; 171.9. IR ν (neat, cm⁻¹): S=O (1021); O=S=O (1330).LC-MS (ESI-APCI): *m/z* [M+H]⁺ = 417 (100%). HPLC purity (method B): *tr* = 2.29, 100≥98%.

Biochemistry

Cloning

A construct of mouse CARM1 from residues 130 to 507 (CARM1_{130_507}) was cloned into the expression vector pDEST20 (Invitrogen). A TEV protease cleavage site was inserted between the GST Tag and the CARM1 coding sequence. The gene of PRMT1 was amplified by PCR from a mouse cDNA bank and cloned into the pET15b (Novagen) expression vector.

Protein production

CARM1_{130_507} was expressed as a GST fusion protein using the baculovirus system. A 1 L cell culture was harvested by centrifugation, resuspended in 50 mM Tris-HCl pH8, 250 mM NaCl, 5% glycerol (v:v), 2 mM beta mercaptoethanol and lysed by sonication. The lysate clarified by centrifugation was incubated with 1 ml of glutathion sepharose 4B resin (GE Healthcare). The resin was washed twice with the lysis buffer and incubated with 200 μg TEV protease for 3 h at 30 °C to cleave the GST tag. The resin was then washed in the same buffer to elute CARM1. The protein was further purified using a S200 gel-filtration column (GE Healthcare) equilibrated in 20mM Tris-HCl pH8, 100mM NaCl. The purified protein was stored at -20 °C in 50% glycerol. The Escherichia coli BL21(DE3) cells overexpressing a mouse His6-PRMT1 construct from residue 11 to 353 were grown in LB medium containing 50 mg/L kanamycin at 37 °C. When OD 600nm reached 0.6, the expression of His6-PRMT1 was induced with 0.2 mM IPTG. After 15 hours at 20°C, cells were harvested by centrifugation, resuspended in 50 mM Tris-HCl pH8.2, 250 mM NaCl, 5% glycerol (v:v) and lysed by sonication. Cell debris were removed by centrifugation at 25000 g for 30 min. The supernatant was loaded on a Ni -NTA agarose (Qiagen) column and washed with the lysis buffer supplemented with 10 mM imidazole. His6-PRMT1 was eluted with 200 mM imidazole in the same buffer. The fractions containing His-PRMT1 were pooled and dialysed overnight against 50 mM Tris-HCl pH8.2, 50 mM NaCl, 0.5 mM EDTA and 0.5 mM beta mercaptoethanol. His-PRMT1 was further purified using a monoQ column (GE Healthcare) equilibrated in the same buffer. The protein was eluted with a linear gradient from 50 mM to 250 mM NaCl. Fractions containing His-PRMT1 were concentrated using an Amicon Ultra with a cutoff of 30 kDa to a final concentration of 2.8 mg.mL⁻¹ and stored at -20°C in 50% glycerol.

Expression and purification of soluble recombinant (H3H4)₂

Escherichia coli BL21(DE3) cells transformed with an expression plasmid encoding His-tagged human histone H4 and histone H3 (Diebold et al., 2011) were grown in 3 x 1 L of Terrific Broth containing 100 mg/L ampicillin at 37 °C to an OD600nm of 0.6. Co-expression of H3 and H4 was induced with 0.2 mM IPTG for 2 hours at 37°C. The cells were harvested by centrifugation at 4000 g for 20 min. The cells were resuspended in 50 mM Tris-HCl pH8.2, 2 M NaCl, 0.1 % PMSF and lysed by sonication. The lysate was clarified by centrifugation at 25000 g for 30 min and applied to a 1.5 mL Ni-NTA agarose (Qiagen) column. (H3H4)₂ heterotetramers were eluted with 200 mM imidazole in 50 mM Tris-HCl pH8, 2 M NaCl. The fraction containing (H3H4)₂ were concentrated using an Amicon Ultra with a cutoff of 30 kDa and were applied to a Sephacryl 200 16/60 gel filtration column (GE Healthcare) equilibrated in 50mM Tris-HCl pH8.0, 2 M NaCl, 0.5 mM

EDTA, 0.5 mM DTT. Purified (H3H4)₂ was concentrated to 9 mg / mL on a Amicon Ultra 30.

Inhibition assays

All the SAM analogues were dissolved in water and their concentration was determined by UV absorption ($\epsilon_{260}=15400 \text{ M}^{-1}\text{cm}^{-1}$). The substrate and enzyme concentrations were adapted to ensure a linear accumulation of the product with **time during the methylation reaction**. that the amount of methylated product versus time was linear

A solution containing 7.5 μM of (H3H4)₂, 10 $\mu\text{g/mL}$ of bovin serum albumin and 200 nM of CARM1 or PRMT1 in 50 mM Tris HCl pH8, 200 mM NaCl, 0.5 mM EDTA, 0.5 mM beta -mercaptoethanol was incubated with increasing concentrations of SAM analogues (0.5 μM to 2 mM) for 30 min at room temperature. A centrifugation step at 17000 g for 1 min was performed to eliminate any precipitated protein. The reaction was initiated by addition of 10 μM ¹⁴C-labeled SAM at 50 mCi.mmol⁻¹ (Perkin Elmer), for 20 min at 25°C, stopped by mixing to 3 μL of SDS PAGE loading buffer and analysed on a 12% polyacrylamide denaturing gel. After electrophoresis, the gel was stained with Coomassie Brilliant Blue, destained, dried on a Whatman 3MM paper sheet and exposed for 24 h to 5 days against aGP2025 imaging plate (Kodak). The amount of radioactivity incorporated specifically by histone H3 or H4 was quantified by phosphorimager analysis (Typhoon, Molecular Dynamic). IC₅₀ values were determined by fitting the data to the Weibull model for sigmoidal dose response curves using the R version 2.13.0 software package.^[41] Molecular graphics figure were generated using PyMOL.^[42]

Docking of ligands in the CARM1 catalytic center

The compound **19b** shown in Figure 6 (as well as other compounds: **16d**, **17c**, **18b**) has been built up starting from the SAH skeleton as observed in the crystal structure of the SAH-CARM1 complex 9 (PDB code 3B3F). The compound has been built using the Maestro building tool 39. Once a given compound is constructed and aligned according to the SAH template, the chain extending towards the arginine channel has been oriented in order to avoid steric clashes with the protein atoms and favoring a hydrogen bond network with the conserved glutamic acids and carbonyl groups of motif III (Glu258, Glu267 in mouse CARM1 numbering). The environment around the guanidinium moiety is rather restricted and does not allow many possibilities in terms of chain orientation. Unfortunately, the hydrogen bond network cannot be derived exactly as the NH (or OH) hydrogen bond donors can form hydrogen bonds with the negative groups of the two conserved glutamic acids or the carbonyl groups of the main chain lining this region. One suitable hydrogen bond network has been selected and the compounds have been minimized in the pocket by leaving flexible only the amino acid side chains that are close ((Glu258, Glu267 in mouse CARM1 numbering) to the compound's chain (**16d**, **17c**, **18b** and **19b**). The minimization has been done in Maestro by using the Tools/Minimize command.^[39]

Glossary

PRMTs, Proteine aRginine MethylTransferases; RNA, Ribonucleic Acid; DNA, Deoxyribonucleic Acid; CARM1, Co-activator Associated aRginine Methyltransferase1; mCARM1, mouse Co-activator Associated aRginine Methyltransferase 1; PRMT1, Proteine aRginine MethylTransferase 1; hPRMT1, human Proteine aRginine MethylTransferase 1; Boc2O, Di-tert-butylidylcarbonate; THF, Tetrahydrofuran; DIEA, N,N-Diisopropylethylamine; HCl, Hydrogen Chloride; NaHCO₃ sat, Sodium hydrogenocarbonate saturated; NH₄Cl sat, Ammonium chloride saturated; NMR, Nuclear Magnetic Resonance; ppm, parts per million; SAM; AdoMet, S-adenosyl-L-methionine; SAH, S-adenosyl-L-homocysteine; SFG, sinefungin; EDTA, Ethylenediaminetetraacetic acid; TEV, Tobacco Etch Virus GST, Glutathione S-transferase; SDS, Sodium dodecyl sulfate; Tris, Tris(hydroxymethyl)aminomethane; IPTG, Isopropyl β -D-1 thio galactopyranoside; Ni-NTA, Nitrilotriacetic acid complexed to Ni²⁺ - ion

Acknowledgements

This work was supported by French National Agency (ANR, ANR-08-PCVI

0037-01) through the "Physique et chimie du Vivant" Program, funds from Association pour la Recherche sur le Cancer (contracts N°3910 and N°3505), Fondation de la Recherche Médicale (grant to P.H.), the French Research Ministry (PhD grants to J.M. and P.H.), Centre National de la Recherche Scientifique (CNRS), the Institut National de la Santé et de la Recherche Médicale (INSERM), the Université de Strasbourg (UDS). We thank Pr. Michael R. Stallcup (University of Southern California, USA) for the pGEX-CARM1 plasmid. We thank Louise Chisholm for careful reading of the manuscript. We thank Pascale Buisine, Dr. Patrick Wehrung and Dr. Sarah Sanglier for the mass spectra. NMR technical assistance provided by Cyril Antheaun is gratefully acknowledged.

Keywords: ((protein arginine methyltransferase · coactivator-associated arginine methyltransferase-1· PRMT4· PRMT1· inhibitors))

- [1]. A. Di Lorenzo, M. T. Bedford, *FEBS Letters*, **2011**, 585, 2024–2031.
- [2]. C. D. Krause, Z. H. Yang, Y. S. Kim, J.H. Lee, J. R. Cook, S. Pestka *Pharmacol. Ther.*, **2007**, 113, 50-87.
- [3]. Y. H. Lee, M. R. Stallcup, *Mol. Endocrinol.*, **2009**, 23, 425-433.
- [4]. R.A. Copeland, M.E. Solomon, V.M. Richon, *Nat. Rev. Drug. Discovery*, **2009**, 8, 724-732.
- [5]. N. Tsankova, W. Renthall, A. Kumar, E. J. Nesler, *Nat. Rev. Neurosci.*, **2007**, 8, 355-67.
- [6]. P.A. Jones, S.B. Baylin, *Cell*, **2007**, 128, 683-92.
- [7]. H. Hong, C. Kao, M.H. Jeng, J. N. Eble, M. O. Koch, T. A. Gardner, S. Zhang, L. Li, C.X. Pan, Z. Hu, *Cancer*, **2004**, 101, 83-89.
- [8]. S. El Messaoudi, E. Fabbrizio, C. Rodriguez, P. Chuchana, L. Fauquier, D. Cheng, C. Theillet, L. Vandell, M. T. Bedford, C. Sardet, *Proc. Natl. Acad. Sci. U.S.A.*, **2006**, 103, 13351-13356.
- [9]. Y. R. Kim, B. K. Lee, R. Y. Park, N. T. X. Nguyen, J. A. Bae, D. D. Kwon, C. Jung, *BMC Cancer*, **2010**, 10, 197.
- [10]. J. M. Yost, I. Korbouk, F. Liu, C. Gao, J. Jin, *Curr. Chem. Genom.*, **2011**, 5, 72-84.
- [11]. D. Cheng, N. Yadav, R. W. King, M. S. Swanson, E.J. Weinstein, M. T. Bedford, *J. Biol. Chem.*, **2004**, 279, 23892-23899.
- [12]. K. Bonham, S. Hemmers, Y. H. Lim, D. M. Hill, M. G. Finn, K. A. Mowen, *FEBS*, **2010**, 277, 2096-2108.
- [13]. S. Castellano, C. Milite, R. Ragno, S. Simeoni, A. Mai, V. Limongelli, E. Novellino, I. Bauer, G. Brosch, A. Spannhoff, D. Cheng, M. T. Bedford, G. Sbardella, *ChemMedChem*, **2010**, 5, 398-414.
- [14]. A. V. Purandare, Z. Chen, T. Huynh, S. Pang, J. Geng, W. Vaccaro, M. A. Poss, J. Oconnell, K. Nowak, L. Jayaraman, *Bioorg. Med. Chem. Lett.*, **2008**, 18, 4438-4441.
- [15]. T. Huynh, Z. Chen, S. Pang, J. Geng, T. Bandiera, S. Bindi, P. Vianello, F. Roletto, S. Thieffine, A. Galani, *Bioorg. Med. Chem. Lett.*, **2009**, 19, 2924-2927.
- [16]. E. Therrien, G. Larouche, S. MankuAllan, M. Nguyen, N. Styhler, S. Robert, M. F. Goulet, A. C. Besterman, J. M. Nguyen, *Bioorg. Med. Chem. Lett.*, **2009**, 19, 6725-6732.
- [17]. H. Wan, T. Huynh, S. Pang, J. Geng, W. Vaccaro, M. A. Poss, G. L. Trainor, M. V. Lorenzi, M. Gottardis, L. Jayaraman, *Bioorg. Med. Chem. Lett.*, **2009**, 19, 5063-5066.
- [18]. B. R. Selvi, K. Batta, A. H. Kishore, K. Mantelingu, R. A. Varier, K. Balasubramanyam, S. K. Pradhan, D. Dasgupta, S. Sriram, S. Agrawal, *J. Biol. Chem.*, **2010**, 285, 7143-7152.
- [19]. T. Huynh, Z. Chen, S. Pang, J. Geng, T. Bandiera, S. Bindi, P. Vianello, F. Roletto, S. Thieffine, A. Galvani, W. Vaccaro, M. A. Poss, G.L. Trainor, M. V. Lorenzi, M. Gottardis, L. Jayaraman, A.V. Purandare, *Bioorg. Med. Chem. Lett.*, **2009**, 19, 2924-2927.
- [20]. D. Cheng, S. Valente, S. Castellano, G. Sbardella, R. Di Santo, R. Costi, Mark T. Bedford, A. Mai, *J. Med. Chem.*, **2011**, 54, 4928-4932.
- [21]. T. Osborne, R. L. Weller Roska, S.R. Rajski, P. R. Thompson, *JACS*, **2008**, 130, 4574-4575.
- [22]. J. Dowden, W. Hong, R. V. Parry, R. A. Pike, S. G. Ward, *Bioorg. Med. Chem. Lett.*, **2010**, 20, 2103-2105.
- [23]. J. Dowden, R. A. Pike, R. V. Parry, W. Hong, U. A. Muhsen, S.G. Ward, *Org. Biomol. Chem.*, **2011**, 9, 7814-7821.
- [24]. O. Obianyo, C. P. Causey, T. C. Osborne, J. E. Jones, Y.H. Lee, V.M. R. Stallcup, P. R. Thompson, *ChemBioChem*, **2010**, 11, 1219-1223.
- [25]. P. Hart, T.M. Lakowski, D. Thomas, A. Frankel, N. L. Martin, *ChemBioChem*, **2011**, 12, 1427-1432.
- [26]. W. W. Yue, M. Hassler, S. M. Roe, V. Thompson-Vale, L. H. Pearl, *EMBO J*, **2007**, 26, 4402-4412.
- [27]. N. Troffer-Charlier, V. Cura, P. Hassenboehler, D. Moras, J. Cavarelli, *EMBO J*, **2007**, 26, 4391-4401.
- [28]. J. S. Sack, S. Thieffine, T. Bandiera, M. Fasolini, G. J. Duke, L. Jayaraman, K. F. Kish, H. E. Klei, A. V. Purandare, P. Rosettani, *Biochemical Journal*, **2011**, 436, 331-339.
- [29]. K. Ramsamy, R. K. Olsen, T. Emery, *Synthesis*, **1982**, 42-43.
- [30]. H. M. Cohen, A. D. Griffiths, D. S. Tawfik, D. Loakes, *Org. Biomol. Chem.*, **2005**, 3, 152-161.
- [31]. S. A. Sikchi, P. G. Hultin, *J. Org. Chem.*, **2006**, 71, 5888-5891.
- [32]. F. Liu, D. J. Austin, *Tetrahedron Letters*, **2001**, 42, 3153-3154.
- [33]. G. Ceulemans, F. Vandendriessche, J. Rozenski, P. Herdewijn, *Nucleosides and Nucleotides*, **1995**, 14, 117.

- [34]. A. H. G. Siebum, W. S. Woo, J. Raap, J. Lugtenburg, *European Journal of Organic Chemistry* **2004**, 13, 2905-2913.
- [35]. S. G. Petersen, S.R. Rajski, *Journal of Organic Chemistry*, **2005**, 70, 5833- 5839.
- [36]. (a) R. Kuhn, W. Jahn, *Chem. Ber.*, **1965**, 98, 1699. (b) R. T.Borchardt, Y. S. Wu, *J. Med. Chem.*, **1974**, 17, 862.
- [37]. C. Guérard, M. Bréard, F. Courtois, T. Drujon, O. Ploux, *Bioorg. Med. Chem. Lett.* **2004**, 14, 1661-1664.
- [38]. J. Lim, W. C Winkler, S. Nakamura, V. Scott, R. R. Breaker, *Angew. Chem. Int. Ed. Engl.* **2006**, 45, 964-968.
- [39]. P. Fonlupt, C. Rey, H. Pacheco, *From Life Sciences* **1982**, 31, 655-659.
- [40]. C. D. Chang, J. K. Coward, *J. Med. Chem.* **1976**, 19, 684-691.
- [41]. Team, R. D. C. R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria, 2011.
- [42]. DeLano, W. L. The PyMOL User's Manual. DeLano Scientific: San Carlos, CA, USA., 2002.
- [43]. Maestro. version 9.1; Schrödinger, LLC: New York, NY, 2010.

Received: ((will be filled in by the editorial staff))

Published online: ((will be filled in by the editorial staff))

II. Essais de co-cristallisation des analogues de la SAM dans mCARM1₁₃₀₋₄₈₇

Des essais de co-cristallisation ont été entrepris pour les analogues de la SAM dans mCARM1₁₃₀₋₄₈₇. Des cristaux ont été uniquement obtenus pour les analogues **10c**, **16d** et **23b** :

- Une image de l'analogue **10c** a été obtenue avec une résolution de 2,8 Å.
- De même, le produit **16d** a été visualisé dans CARM1 avec une résolution de 2,4 Å. Néanmoins, la structure cristallographique obtenue n'est pas complète car l'amine terminale de la chaîne n'y est pas présente.
- Finalement, une image du composé **23b** dans CARM1 a été obtenue avec une résolution de 1,9 Å. Cette structure cristallographique nous a permis d'acquérir des informations supplémentaires sur « la poche de l'arginine » ainsi que d'expliquer la sélectivité observée pour certains analogues entre CARM1 et PRMT1.

Ces résultats vont nous permettre de continuer notre stratégie en explorant « la poche de l'arginine » par des essais de co-cristallisation avec des analogues de bisubstrats dans CARM1.

III. Conclusions et perspectives

Dans la première partie de ce chapitre, une dizaine d'analogues de la SAM ont été synthétisés pour explorer la « poche du sulfonium ». Les tests *in vitro* utilisés nous ont permis de mesurer l'inhibition de ces composés dans CARM1 ; ils nous ont montré que ce sont les analogues possédant un atome d'azote à la place d'un atome de soufre qui présentent les meilleures valeurs d'inhibition (CI₅₀) comprises entre 37 µM et 158 µM.

Pour commencer à explorer la « poche de l'arginine », des analogues de la « SAM arginine mimiques » ont été synthétisés : trois d'entre eux (**18b**, **19b** et **23b**) ont montré une bonne sélectivité vis-à-vis de PRMT1. Cette sélectivité et cette spécificité ont pu être expliquées de façon rationnelle grâce aux structures cristallographiques obtenues grâce aux composés **10c**, **16d** et **23b** mais aussi grâce à la modélisation moléculaire. Ces structures cristallographiques nous ont également apporté des informations sur certains points d'ancrage du site actif.

Ces résultats encourageants valident donc la première étape de notre stratégie. Nous allons maintenant poursuivre nos travaux dans le prochain chapitre en explorant le « poche de l'arginine » grâce à des analogues bisubstrats.

Chapitre 3

Exploration de la « poche de l'arginine » par co-cristallisation d'analogues de bisubstrats

Chapitre 3

Exploration de la « poche de l'arginine » par co-cristallisation d'analogues de bisubstrats

Dans ce chapitre, nous explorerons de manière approfondie la « poche de l'arginine » de CARM1 afin de synthétiser des molécules hameçons. Pour cela, nous étudierons les différentes interactions qui peuvent avoir lieu entre le site actif de CARM1 et le peptide. Cette analyse nous permettra de proposer un modèle sur le mécanisme de transfert de méthyle de l'arginine par cette enzyme. A partir de ces données, des analogues de bisubstrats seront synthétisés puis caractérisés biochimiquement ; ils se révéleront être de bons inhibiteurs sélectifs des PRMTs.

I. Mécanisme de méthylation proposé dans la littérature pour les PRMTs de type I

Dans ce paragraphe, les postulats mécanistiques qui ont été émis pour la réaction de méthylation des PRMTs de type I seront présentés. Ils ont été établis sur la base d'une structure cristallographique ou à l'aide d'une étude cinétique.

I.1 Réactions

Les PRMTs de type I (PRMT 1, 2, 3, 4, 6 et 8) catalysent le transfert du méthyle de la SAM vers l'azote de la chaîne latérale de l'arginine : la ω - N_G -monométhylarginine (MMA) est tout d'abord obtenue. Cette étape est commune à toutes les PRMTs. Puis s'ensuit la formation de la ω - N_G, N_G -asymétrique diméthylarginine (*a*DMA)⁽¹⁾ (figure 3.1).

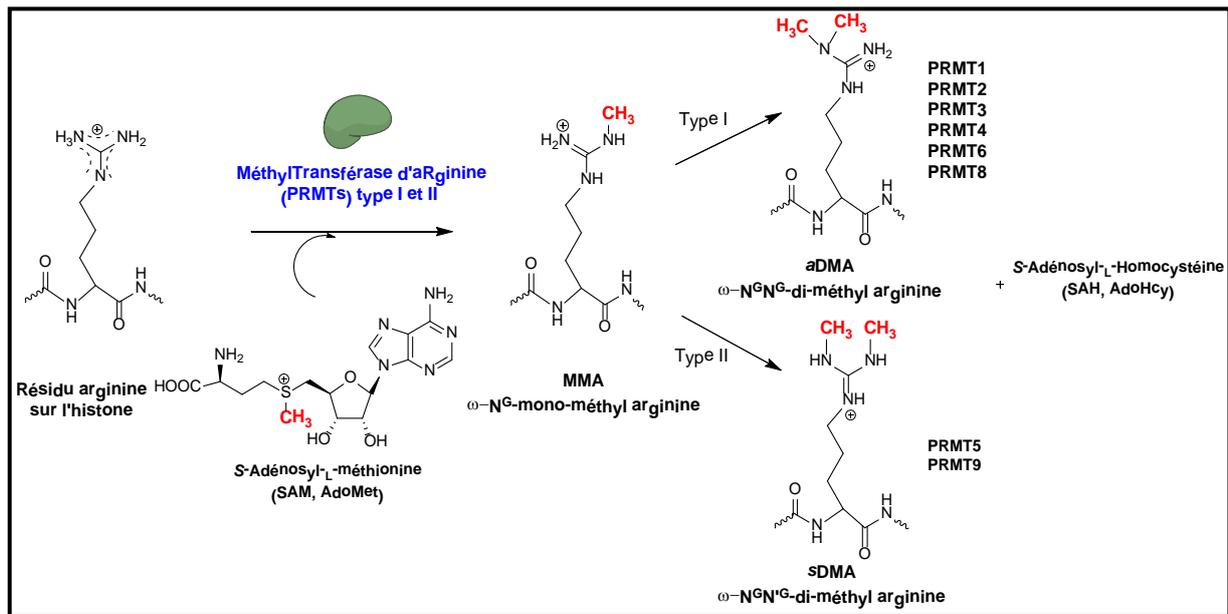


Figure 3.1 : Réaction de méthylation des arginines pour les PRMTs de type I *via* la SAM

Aujourd'hui, le déroulement exact des étapes de méthylation impliquées dans cette réaction n'est pas connu, en particulier l'ordre de liaison du cofacteur et du substrat. L'enchaînement des étapes ne semble particulièrement pas évident pour la seconde méthylation.

Des études ont été menées pour mieux comprendre ces étapes et des postulats mécanistiques ont été proposés. Ces travaux ont été réalisés grâce à une approche cinétique ou une approche structurale ; cette dernière repose sur les structures cristallographiques des PRMTs déjà caractérisées.

I.1.1 *Postulat du mécanisme de méthylation par une approche cinétique*

Deux enzymes appartenant à la famille des PRMTs ont été étudiées : PRMT1 et PRMT6 ; les substrats utilisés pour ces études sont respectivement deux peptides, l'un mimant l'histone H4, l'autre mimant la séquence des substrats fibrillaires.

En 2008, l'équipe de Thompson⁽²⁾ a axé ses recherches sur l'ordre de réaction de diméthylation pour PRMT1. Selon ses travaux, cette enzyme utiliserait un mécanisme de catalyse aléatoire où l'ordre ne serait pas important : la fixation du substrat et de la SAM, la libération de la SAH et du peptide diméthylé hors du site actif se produiraient sans ordre déterminant.

La même année, l'équipe de Frankel et *al.*⁽³⁾ a exploré cette réaction, catalysée par PRMT6, en utilisant des peptides mono- et di-méthylés qui miment les substrats fibrillarins. Cette étude a révélé que ce mécanisme était totalement différent de celui proposé par le groupe de Thompson et *al.* : il s'agirait d'un mécanisme de type séquentiel au cours duquel la SAM se lierait en premier à l'enzyme. Dans un second temps, la réaction de diméthylation pourrait avoir lieu et le peptide diméthylé serait alors dissocié avant la libération de la SAH.

En 2011, l'équipe de Zheng et *al.*⁽⁴⁾ s'est intéressé à PRMT1 et a analysé la cinétique de réaction de méthylation d'un peptide, mimant la séquence de l'histone H4, en présence de cette enzyme.

Leurs travaux sont en accord avec ceux proposés par Frankel et *al.*: PRMT1 serait aussi impliquée dans un mécanisme séquentiel où le peptide, une fois mono-méthylé, serait extrait du site catalytique de l'enzyme avant la seconde méthylation.

Ces conclusions concordent avec celles obtenues grâce à une étude cristallographique des enzymes PRMT1 et CARM1. Ces résultats vont être maintenant abordés.

I.1.2 Etude du mécanisme sur les bases des structures cristallographiques de CARM1

En 2007, l'équipe de J. Cavarelli⁽⁵⁾ et celle de Pearl⁽⁶⁾ ont publié concomitamment deux structures cristallographiques de CARM1. Les structures obtenues par l'équipe de J. Cavarelli

contiennent l'enzyme seule ou l'enzyme co-cristallisée avec l'un de ses cofacteurs, la SAH. Les références respectives de ces structures disponibles dans la banque de donnée de protéine (pdb) sont 3B3G et 3B3F.

Deux observations ont été notées :

- Un changement de conformation a eu lieu en présence du cofacteur.
- CARM1 est sous forme de dimère.

Deux conclusions ont été tirées à partir de ces observations : tout d'abord, le changement de conformation de CARM1 implique que son cofacteur, la SAH, se lie en premier. La deuxième observation suggère que le dimère de l'enzyme est capable d'éjecter le produit de mono-méthylation de manière transitoire, pour permettre le réapprovisionnement en co-facteur, ou de façon définitive : le second site actif de CARM1 serait alors capable de se lier avec le produit de mono-méthylation pour obtenir le peptide diméthylé⁽⁶⁾

II. Etude du mécanisme de transfert de méthyle faisant intervenir la SAM et une guanidine

II.1 Etude de l'environnement de la « poche de l'arginine » de CARM1 et PRMT1

En superposant les séquences d'acides aminés contenus dans les sites catalytiques de CARM1 et de PRMT1, des similarités ont été observées dans la poche de l'arginine, où plusieurs résidus sont identiques : il s'agit de deux glutamates (Glu), d'une tyrosine (Tyr), de deux méthionines (Met), d'une histidine (His) et d'un acide aspartique (Asp). Ces résidus sont représentés dans le tableau **3.1**. Dans la « poche de l'arginine », le groupement guanidinium est entouré de résidus qui possèdent principalement des charges négatives, dues à la présence des deux acides glutamiques. Ces résidus présentent aussi des oxygènes polaires provenant des groupements carbonyles de la chaîne principale de l'acide glutamique et de la méthionine.

La structure cristallographique de PRMT1, de référence PDB 1OR8, a révélé que les deux acides glutamiques du site actif (Glu 144 et Glu 153) stabilisent le groupement guanidinium de l'arginine en formant des liaisons électrostatiques avec cette dernière. Pour vérifier si ces acides aminés sont cruciaux concernant l'activité catalytique de PRMT1, des structures mutantes ont été réalisées où les deux glutamates ont été mutés par des glutamines ; la fonction acide est donc remplacée par une fonction amide. Ces modifications ont provoqué la perte d'activité de méthylation de PRMT1⁽⁷⁻⁸⁾.

L'équipe de Lee et *al* a réalisé des études de mutagenèse, effectuées sur CARM1 : elles ont montré que la mutation de le Glu 267 en glutamine conduisait également à la perte d'activité de méthylation de cette enzyme⁽⁹⁾.

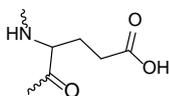
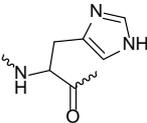
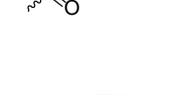
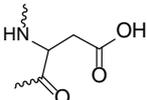
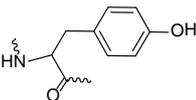
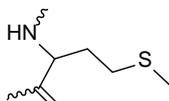
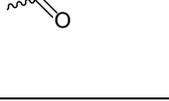
CARM1	PRMT1	Structures	CARM1	PRMT1	Structures
Glu 258	Glu 144		His 415	His 294	
Glu 267	Glu 153		Asp 166	Asp 58	
Tyr 262	Tyr 147				
Met 163	Met 48				
Met 269	Met 155				

Table 3.1 : Résidus conservés dans la poche de l'arginine de CARM1 et de PRMT1

II.2 Comparaison de la « poche de l'arginine » de CARM1 avec une autre enzyme utilisant la SAM

La guanidine acétate méthyltransférase (GAMT) est une enzyme qui utilise la SAM comme donneur de méthyle ; son substrat est une guanidine.

Cette enzyme nous a interpellés, car l'ancrage entre le substrat et les acides aminés de son site actif est similaire à celui de CARM1. Le groupement guanidinium est dans le plan et forme la même triade avec les acides glutamiques et aspartiques que dans notre site actif.

II.2.1 La guanidine acétate méthyltransférase (GAMT)

II.2.1.1 Réaction

La guanidinoacétate (GAA) est méthylée par la SAM pour conduire à la formation de la créatine et de la SAH ; cette réaction est catalysée par la GAMT (figure 3.2) ⁽¹⁰⁻¹¹⁾.

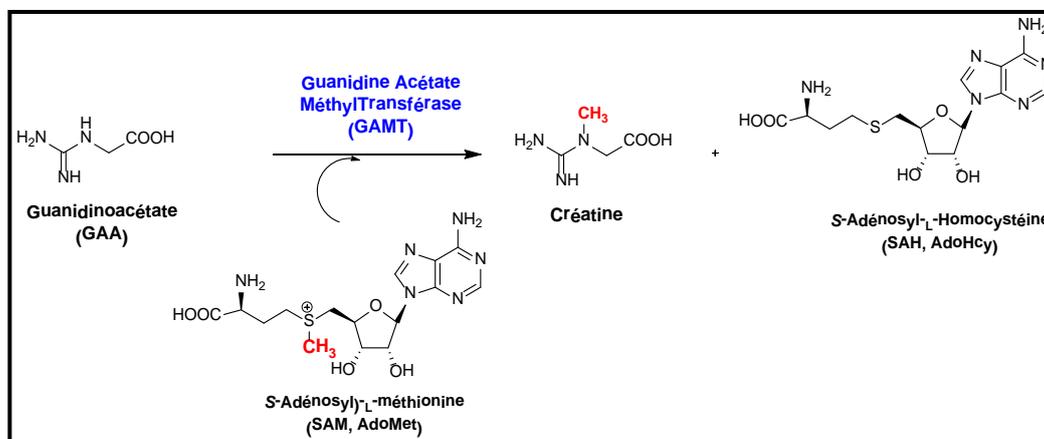


Figure 3.2 : Réaction de méthylation de la guanidinoacétate catalysée par la GAMT via la SAM.

En 2006, l'équipe de Himo et *al.*⁽¹²⁾ a proposé un mécanisme réactionnel basé sur du calcul théorique et des données structurales obtenues sur l'enzyme GAMT, en complexe avec un ligand, la SAM, et un substrat, la GAA (code PDB IXCJ). Ces travaux ont montré que la SAM et la GAA se liaient d'abord au site actif dans lequel le groupement guanidine du GAA est ancré par un résidu d'acide aspartique (Asp 134) et un résidu glutamate (Glu 45) (figure 3.3.a). Puis le proton de l'azote N_E (représenté en bleu sur la figure 3.3) du GAA était déprotonné par une interaction directe avec le résidu Asp 134 présent dans ce site (figure 3.3.b). L'environnement de la protéine permet d'augmenter le pK_a de l'Asp 134⁽¹¹⁾, rendant cette étape possible. Cette étape de déprotonation rend l'atome N_E très nucléophile et donc capable de réagir sur le groupe méthyle de la SAM pour réaliser la réaction de type S_N2 conduisant à la formation de la créatine (figure 3.3.d). Finalement, la méthylation de la GAA rompt la liaison hydrogène établie avec l'Asp 134 et extrait ainsi la créatine du site actif⁽¹²⁾.

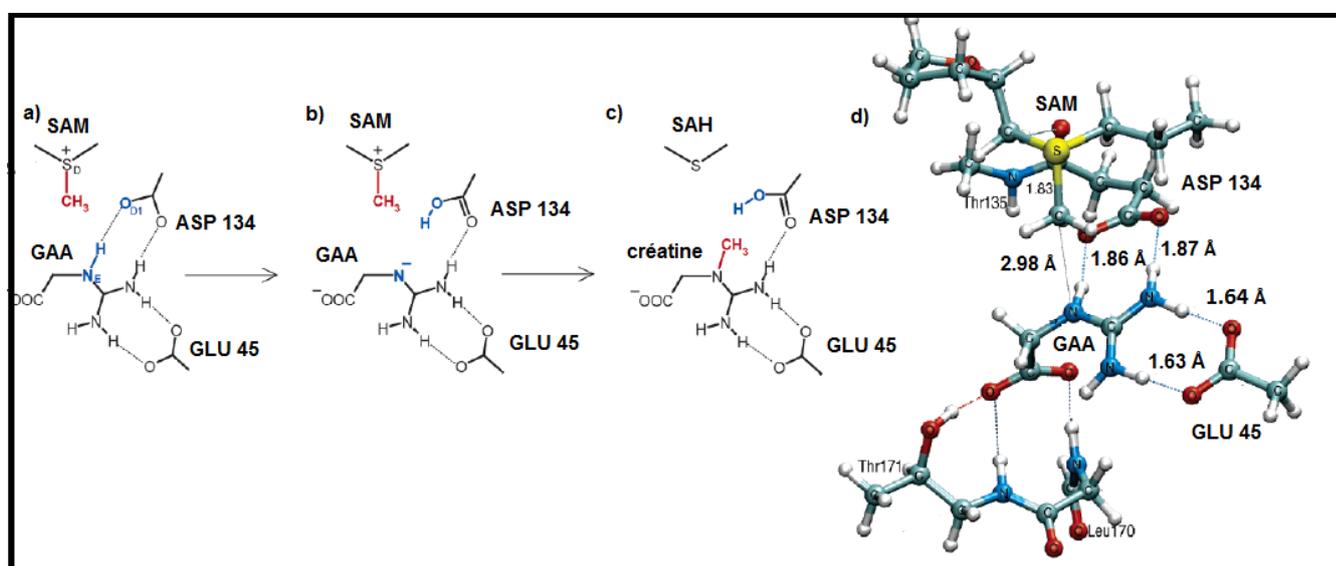


Figure 3.3: Réaction de méthylation de la GAA catalysée par la GAMT *via* la SAM⁽¹²⁾. a) Au départ, ancrage de la GAA au niveau de l' Asp134. b) Transfert de proton et de méthyle. c) Produits de réaction. d) Structure optimisée de l'état de transition. Code de couleurs pour les atomes : l'atome de carbone est en gris, l'atome d'azote en bleu, l'atome d'oxygène en rouge, l'atome d'hydrogène en blanc et l'atome de soufre en jaune.

Dans le cadre de cette thèse, nous nous sommes inspirés du mécanisme postulé par l'équipe de Himo. Nous nous sommes aussi basés sur les travaux déjà publiés sur le site catalytique de CARM1 qui ont été obtenus grâce à sa structure cristallographique⁽⁵⁻⁶⁾. A partir de ces données, un modèle mécanistique a été proposé par le Pr. Jean-Marie Wurtz de l'IGMBC de Strasbourg, avec qui nous avons collaboré.

III. Modèle en deux dimensions, proposé pour la réaction de méthylation de l'arginine du substrat, dans le site catalytique de CARM1

Plusieurs études concernant le mécanisme de méthylation des PRMTs ont été présentées dans le paragraphe précédent ; elles ont montré que cette réaction respectait un ordre séquentiel, comme par exemple le mécanisme de transfert de méthyle sur la guanidinoacétate, catalysée par une méthyltransférase. A partir de ces données, nous avons émis un postulat sur l'état de transition de la réaction de méthylation de l'arginine du substrat. Ce transfert est catalysé par CARM1 et le donneur de méthyle reste la SAM.

III.1 Image en deux dimensions du modèle proposé

Le groupement guanidinium est plan : il peut donc former des liaisons hydrogènes avec les acides glutamiques présents dans le site actif (Glu 258 et Glu 267). Ces deux acides aminés ont un rôle bien précis puisque le Glu 267 permet l'ancrage du substrat tandis que le Glu 258 joue un rôle de pompe à proton (figure **3.4 a**). Une des formes limite du groupement guanidinium est représentée, cette polarisation permet d'effectuer une attaque nucléophile sur le méthyle de la SAM (figure **3.4 b**). En effet, à l'approche du groupement méthyle, le proton N-H de la guanidine interagit avec le Glu 258 et cet atome d'azote adopte alors une géométrie plus pyramidale qui facilite l'attaque nucléophile sur le groupement $-CH_3$. Le proton de la liaison N-H est ensuite transféré complètement vers le Glu 258. A l'état de transition, les

transferts du méthyle et du proton semblent être couplés et ont lieu de manière concertée (figure 3.4 c). Finalement, après le transfert du groupe méthyle sur l'atome d'azote, ce dernier retrouve son état original d'hybridation sp^2 et le Glu 258 est protoné (figure 3.4 d).

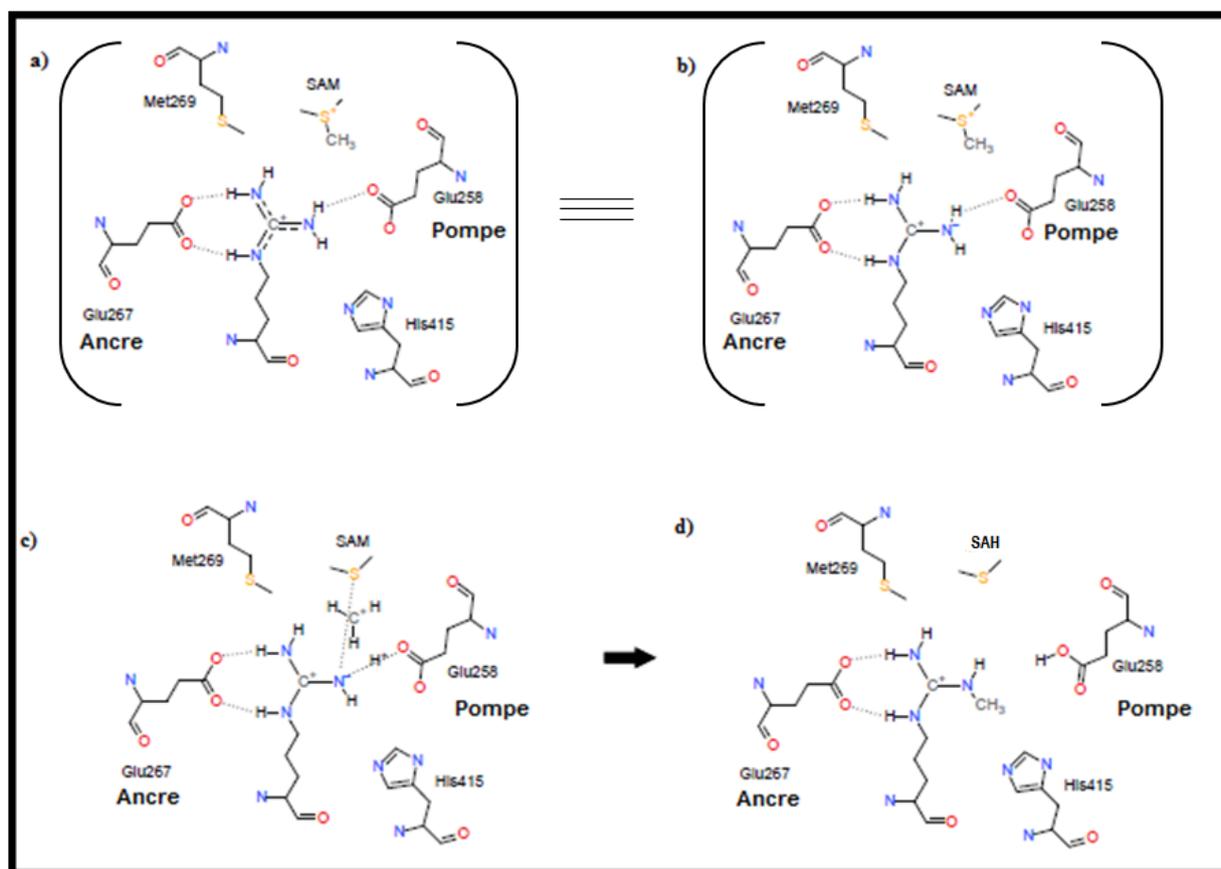


Figure 3.4: Modèle mécanistique proposé pour comprendre la réaction de méthylation de l'arginine du substrat, catalysée par CARM1. a) L'arginine représentée dans lequel le cation guanidinium est représenté sous sa forme moyenne. Le guanidinium est dans le plan et il est ancré dans le site actif de CARM1 grâce au Glu 267 ; en même temps, le Glu 258 joue le rôle de pompe à proton. b) Une des formes limite est représentée pour montrer que le groupement guanidinium est polarisé pour permettre l'attaque nucléophile de l'arginine sur le méthyle. c) Durant l'état de transition, le méthyle est transféré ; en même temps, la déprotonation du guanidinium a lieu. d) L'arginine est méthylée et le Glu 258 est protoné.

III.2 Image en trois dimensions représentant l'environnement de l'arginine

Pour construire cette image, le Pr. Jean-Marie Wurtz s'est basé sur les données bibliographiques et cristallographiques du site actif de mCARM1₁₄₀₋₄₈₀ co-cristallisé avec le cofacteur SAH. La référence de cette structure disponible dans la banque de donnée de

protéine est pdb 3B3F. En s'inspirant des données structurales, une image en trois dimensions (figure 3.5) a été créée grâce au logiciel de modélisation Chimera⁽¹³⁾ et les molécules ont été construites avec le logiciel Avogadro⁽¹⁴⁾. L'arginine a été positionnée dans le site catalytique à une distance de 4,65 Å du méthyle de la SAM. L'arginine est positionnée pour que le groupement guanidine soit plan et qui soit capable d'interagir avec les acides glutamiques. Deux des azotes du groupement guanidine forment une liaison hydrogène avec les oxygènes des carbonyles du Glu 267, à une distance de 2,51 Å et de 2,85 Å. Le troisième azote forme une liaison hydrogène de 2,98 Å avec le Glu 258. L'angle entre le soufre de la SAM et son méthyle est de 131,14°; l'angle des plans des atomes C-S (de la SAM) et des atomes N-C (de l'arginine) mesure 179,93° permettant ainsi le transfert du groupement méthyle.

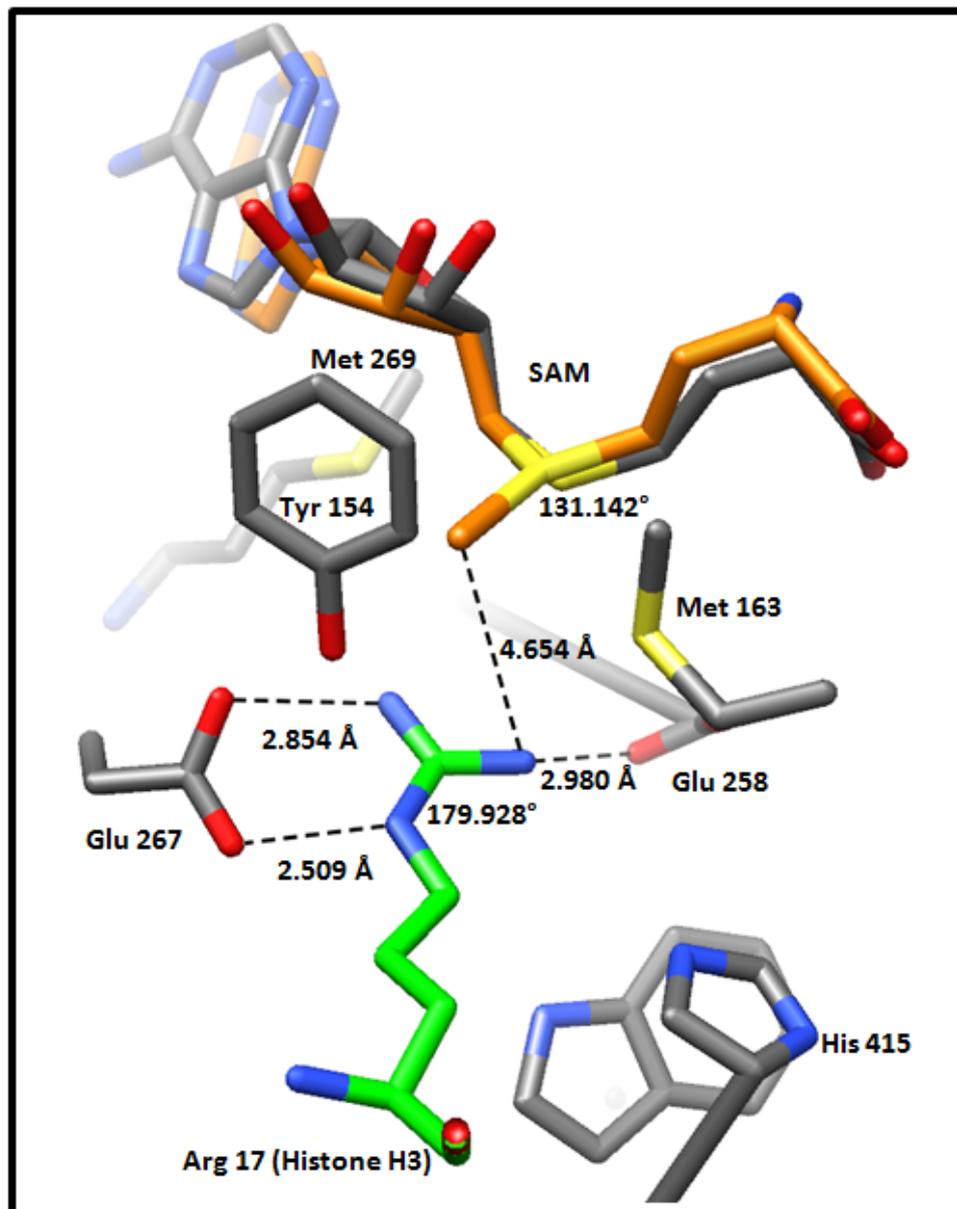


Figure 3.5 : Image 3D représentant l'environnement de l'arginine dans le site actif, avec le cofacteur SAM et les acides aminés. Le code couleur est le suivant : les atomes d'azote sont en bleu, les atomes d'oxygène en rouge et les atomes de soufre en jaune. Les atomes de carbone pour le cofacteur SAH ainsi que les résidus de la structure cristallographique 3B3F sont en gris tandis que le cofacteur SAM modélisé est en orange. L'arginine du substrat peptidique est modélisé en vert.

A partir de ce modèle en trois dimensions, des analogues de bisubstrats ont été proposés ; des produits correspondant à la première étape de méthylation ont été aussi conçus.

IV. Synthèse des analogues de bisubstrats qui se rapprochent de l'état de transition et qui correspondent à la première étape de méthylation

L'obtention de ces analogues nous offrent deux perspectives : explorer « la poche de l'arginine » et synthétiser des inhibiteurs des PRMTs.

IV.1 Stratégie pour synthétiser des inhibiteurs des PRMTs

Pour inhiber les PRMTs, notre stratégie est de synthétiser des analogues azotés de la SAM qui se rapprochent de l'état de transition entre cette dernière et l'arginine. Pour connecter l'azote de la SAM et le groupement guanidinium du substrat, nous avons choisi deux longueurs de chaîne carbonée pour contrôler la flexibilité de ces composés dans le site actif (figure 3.6).

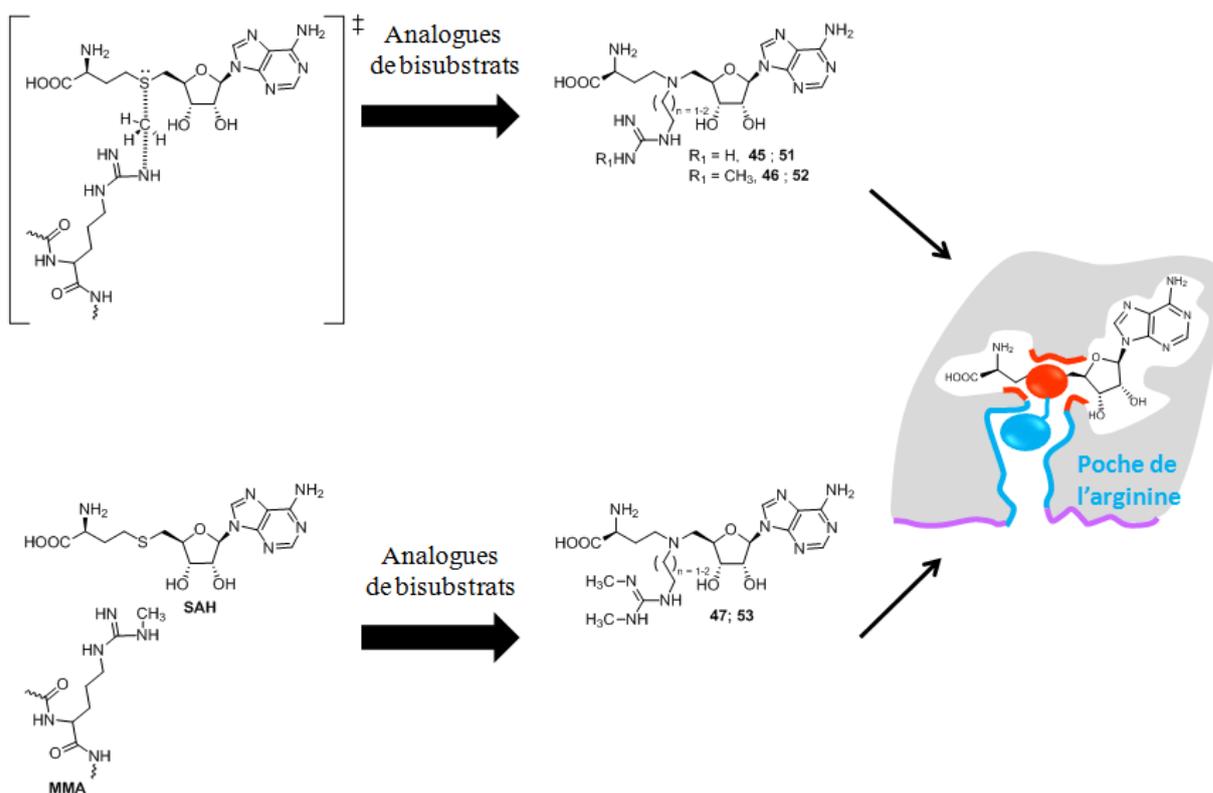


Figure 3.6 : Stratégie représentant les équivalents covalents de synthèse qui se rapprochent de l'état de transition. a) Analogues de bisubstrats (**45** et **46** ; **51** et **52**) qui correspondent à la position de l'arginine et de la SAM avant réaction. b) Analogues de bisubstrats (**47** et **53**) qui correspondent à l'étape de monométhylation de l'arginine.

Sur la figure **3.6 a**, sont représentés les analogues de bisubstrats (**45** et **46** ; **51** et **52**) : ils se rapprochent de l'équivalent covalent de l'état de transition entre le cofacteur et l'arginine, dans le site actif de CARM1, avant réaction. Deux autres composés guanidines (**47** et **53**) ont été synthétisés pour se rapprocher de l'état de transition ; ce dernier correspond à l'étape de mono-méthylation qui est commune aux deux types I et II de PRMTs (figure **3.6 b**).

Au cours de ma thèse, deux publications nous ont permis de confirmer notre idée. L'équipe de Dowden et *al.*⁽¹⁵⁾ a tout d'abord synthétisé un analogue de bisubstrat qui possède trois carbones entre l'azote de la SAM et le groupement guanidinium. Ce produit a présenté une inhibition de $13,3 \pm 8,7 \mu\text{M}$ pour PRMT1. En 2011, cette même équipe a proposé une deuxième série d'analogues de bisubstrats : ces derniers possèdent des chaînes carbonées plus

longues entre l'azote de la SAM et le groupement guanidinium afin d'obtenir une sélectivité au sein des PRMTs⁽¹⁶⁾.

Avant de synthétiser les molécules hameçons, nous avons voulu explorer la « poche de l'arginine » grâce à des analogues de bisubstrats. Dans ce chapitre, six analogues ont donc été synthétisés : ils possèdent soit une chaîne carbonée à deux carbones (composés **45**, **46** et **47**), soit une chaîne carbonée à trois carbones (composés **51**, **52** et **53**) représentés sur la figure **3.7**.

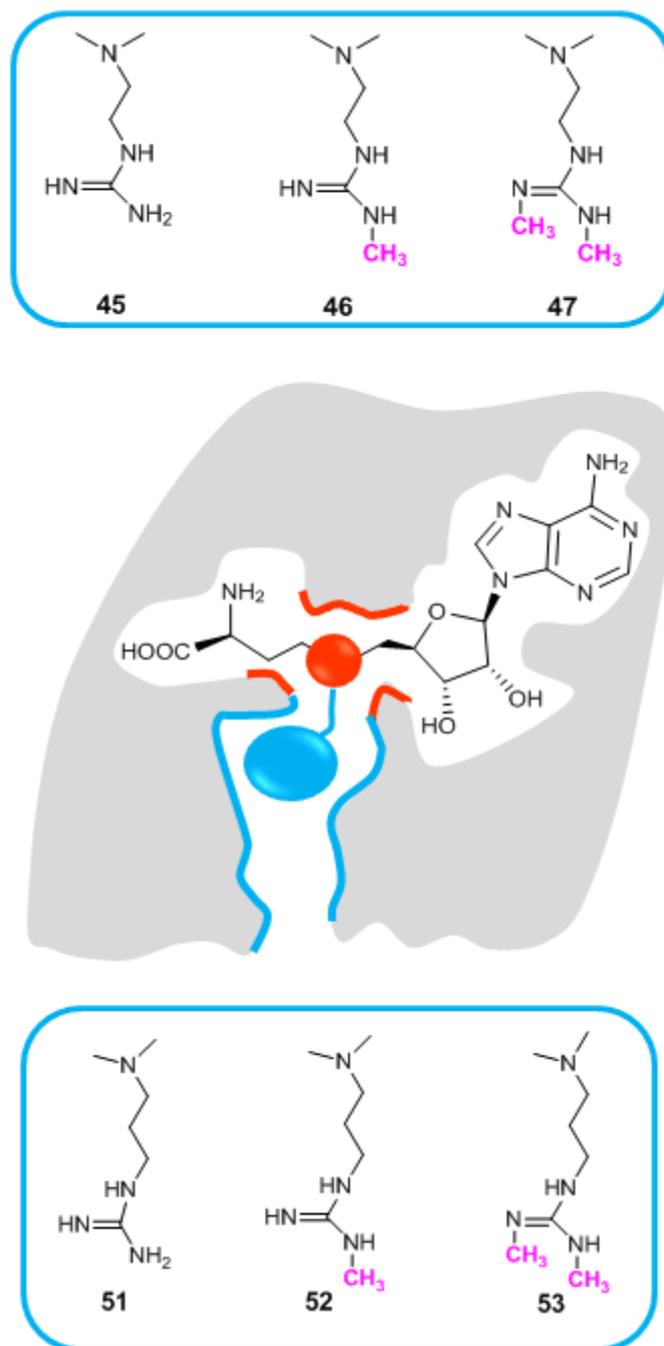


Figure 3.7 : Six analogues de bisubstrats (45 à 47 et 51 à 53) ont été synthétisés pour explorer la « poche de l'arginine ».

IV.2 Synthèse d'analogues de bisubstrats

Les analogues 45 à 47 possèdent un bras espaceur de deux carbones et sont synthétisés en deux étapes à partir de l'amine 35 ; celle-ci subit une réaction de guanylation qui conduit

aux dérivés **42** à **44** ; ces derniers sont déprotégés en milieu acide pour former les analogues souhaités.

Les analogues **51** à **53** présentent le bras espaceur le plus long et sont synthétisés en utilisant l'intermédiaire **16c**, décrit dans le chapitre précédent.

Abordons maintenant la synthèse du précurseur **35** qui est l'intermédiaire pivot pour obtenir les analogues **45** à **47** (figure 3.8).

IV.2.1 Synthèse du précurseur 35

Le dérivé **35** a été synthétisé à partir de l'amine **10a** décrite dans le chapitre précédent (figure 3.8). Cette dernière est alkylée par du bromoéthanol pour conduire à l'amine secondaire **31** avec un rendement de 45%. Dans une seconde étape, l'amine tertiaire **32** a été obtenue avec un rendement de 60% grâce à une réaction d'amination réductrice, entre l'aldéhyde **3** et le dérivé **31**.

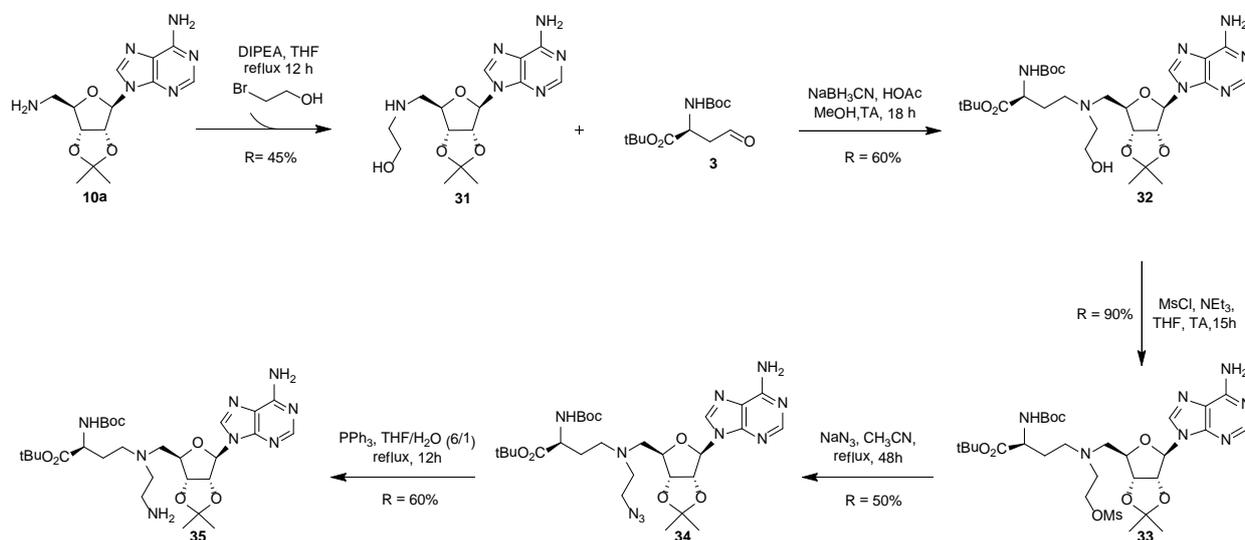


Figure 3.8 : Synthèse du précurseur **35**

La fonction alcool du composé **32** a ensuite été activée par un groupement mésyle pour conduire quantitativement au dérivé **33**.

Ce dernier a été chauffé 48 heures à reflux, en présence d'azoture de sodium : la substitution nucléophile du mésyle par un groupement azoture a été ainsi réalisée et le dérivé **34** a été obtenu avec un rendement de 50%. Cet azoture est ensuite réduit dans les conditions de Staudinger pour obtenir le précurseur recherché **35**, qui a été isolé avec 60% de rendement.

Les précurseurs **16c** et **35** vont maintenant nous permettre de synthétiser les dérivés de bisubstrats souhaités (figure **3.9**).

IV.2.2 *Synthèse des analogues de bisubstrats 45 à 47 et 51 à 53*

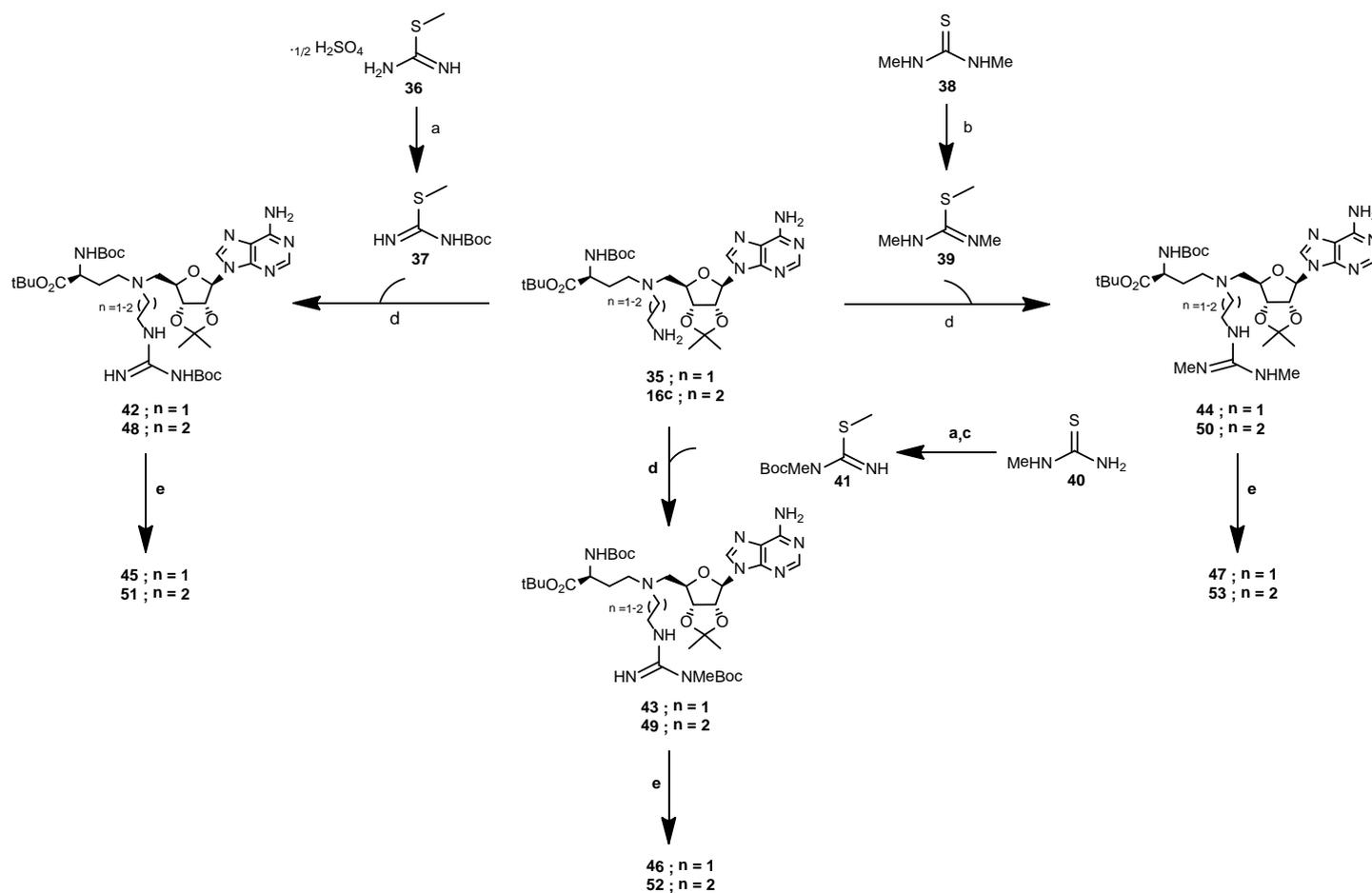
Le principe de cette synthèse est le suivant : les dérivés guanidines **42** à **44** et **48** à **50** sont obtenus respectivement à partir des amines **35** et **16c** et des S-méthyl-iso-thiourées **37**, **39** et **41** (figure **3.9**). Les amines et thiourées sont ensuite chauffées sous reflux d'acétonitrile pendant 48 heures pour former les guanidines correspondantes avec des rendements compris entre 40 et 60%.

La première étape a été l'obtention des S-méthyl-iso-thiourées **37**, **39** et **41** qui ont été synthétisées en utilisant trois procédures différentes. Nous nous sommes inspirés du protocole décrit par Jan Izdebski et *al.*⁽¹⁷⁾ pour synthétiser la N(Boc)-S-méthyl-iso-thiourée **37** ; cette dernière a été formée à partir de la S-méthyl-iso-thiourée commerciale **36** et isolée avec un rendement de 80%. Dans un second temps, la S-1,2,3-triméthyl-iso-thiourée **39** a été obtenue quantitativement en respectant la procédure décrite par LaMattina et *al.*⁽¹⁸⁾ et a été synthétisée grâce à une réaction de méthylation entre l'iodure de méthyle et la 1,3-diméthyl-thiourée **40** commerciale.

Finalement, pour former en deux étapes la N(Boc)-S-1,2-diméthyl-iso-thiourée **41**, nous avons utilisé la 1,3-méthyl-thiourée commerciale **40**. Celle-ci est méthylée à l'aide du diméthylsulfate pour conduire à la thiourée correspondante, puis l'amine de cette dernière est

protégée par un groupement Boc ; la thiourée **41** est ainsi obtenue avec 64% de rendement sur les deux étapes. Ce chemin réactionnel est décrit sur la figure 3.9.

La protection de la S-méthyl-iso-thiourée intermédiaire permet d'avoir deux composés **43** et **49** moins polaires lors de l'étape de guanylation ce qui facilite leurs purifications par chromatographie sur gel de silice.



Conditions réactionnelles : a) dioxane/H₂O (1/1), NaOH 1N, Boc₂O, 12h, TA; b) THF, MeI, K₂CO₃, 2h à 0°C puis 3h à TA; c) THF, (CH₃)₂SO₄, K₂CO₃, 2h à 0°C, puis TA ; d) CH₃CN, 48h, 80°C, 40-60%; e) HCl (1,25M dans le dioxane), 2h, TA, rendement supérieur à 98%.

Figure 3.9 : Synthèse des dérivés de bisubstrats.

Ces trois S-méthyl-iso-thiourées sont mises en présence des amines **35** et **16c**, à reflux, dans l'acétonitrile. Les dérivés guanidines **42** à **44** et **48** à **50** sont ainsi obtenus puis déprotégés en milieu acide grâce à une solution d'acide chlorhydrique à 1,25M dans le

dioxane ; cette réaction a lieu à température ambiante pendant 2 heures, pour obtenir les analogues **45** à **47** et **51** à **53** souhaités avec un rendement supérieur à 98%.

Ces derniers ont ensuite été testés biologiquement ; pour cela, nous avons collaboré avec l'équipe dirigée par les Prs. Jean Cavarelli et Jean-Marie Wurtz de l'IGBMC de Strasbourg.

Ces produits ont subi des tests biologiques qui permettent de déterminer leurs concentrations inhibitrices médianes (CI₅₀) sur l'enzyme CARM1 ; ces mesures ont été effectuées par le Dr. Vincent Cura. Ces analogues ont également été testés en co-cristallisation avec l'enzyme CARM1 par le Dr. Pierre Hassenboeler et Justine Maillot.

V. Résultats et discussions

V.1 Tests in vitro pour l'inhibition de CARM1 et de PRMT1 sur le tétramère de l'histone (H3H4)₂

Les tests d'inhibition ont été réalisés en mettant en compétition le cofacteur SAM, et les différents analogues, dans le site catalytique de CARM1 ou de PRMT1. Pour étudier ces enzymes, deux séquences d'acides aminés appartenant à la souris ont été construites : CARM1 souris (mCARM1₁₃₀₋₅₀₇) comprend les acides aminés 130 à 507 tandis que PRMT1 souris (mPRMT1) en contient 354. La méthylation de ces deux enzymes de souris a été mesurée en utilisant le tétramère de l'histone (H3H4)₂ comme substrat et le ¹⁴C SAM radioactif comme donneur de méthyle.

La radioactivité est quantifiée en utilisant des gels dans l'histone H3 pour CARM1 et dans l'histone H4 pour PRMT1. La compétition entre la SAM et chaque analogue pourra être ainsi mesurée dans le site catalytique de mCARM1₁₃₀₋₅₀₇ ou mPRMT1. Ces tests nous ont permis de classer toutes les molécules.

Pour déterminer la valeur de CI_{50} de chaque analogue, l'activité enzymatique résiduelle a été mesurée pour différents concentrations variant de 0,5 μ M à 2 mM. Les résultats d'inhibition obtenus avec ces analogues et leur sélectivité envers PRMT1 sont représentés dans le tableau **3.2** et **3.3**.

La SFG est un analogue naturel de la SAM qui est produite par *Streptomyces griseolus*. Elle est également un inhibiteur potentiel des méthyltransférases⁽¹⁹⁾. La structure cristallographique de CARM1⁽⁵⁾ a montré que ce composé se fixe au même endroit que la SAM dans son site catalytique. La SFG joue ainsi le rôle d'inhibiteur compétitif de la SAM. Comme ce composé se lie dans la « poche du sulfonium » il est donc utilisé comme molécule de référence pour nos tests d'activité.

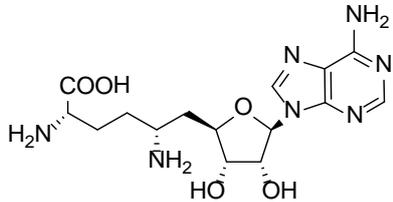
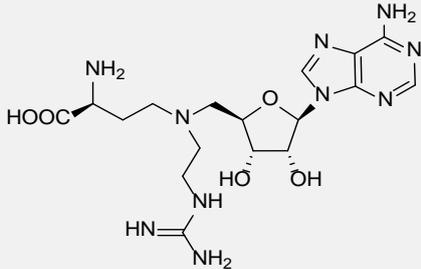
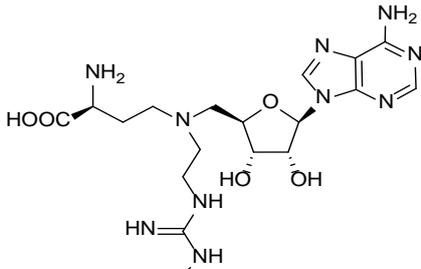
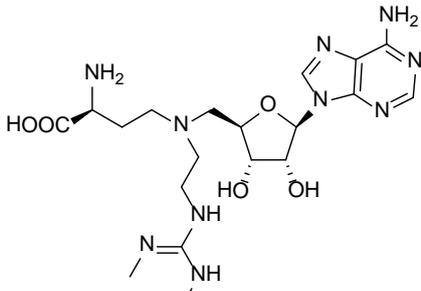
Composés	Structures	CI ₅₀ ± DS [μM]		RS ^(c)
		CARM1 ^(a,b)	PRMT1 ^(a,b)	
SFG		0,7 ± 0,3	0,6 ± 0,3	1
45		2,1 ± 0,6	113,0 ± 43,0	54
46		9,0 ± 2,0	194,0 ± 12,0	22
47		9,1 ± 3,2	125,0 ± 8,0	14

Table 3.2 : Activités d'inhibition (valeurs de CI₅₀) des analogues de bisubstrats contre CARM1 et PRMT1. a) Le substrat utilisé est le tétramère d'histone (H3H4)₂ et le ¹⁴C SAM radioactif est le donneur de méthyle; b) Les valeurs ont été déterminées à partir de deux expériences séparées au minimum; c) RS = rapport de sélectivité : PRMT1/CARM1.

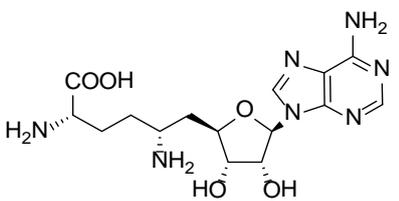
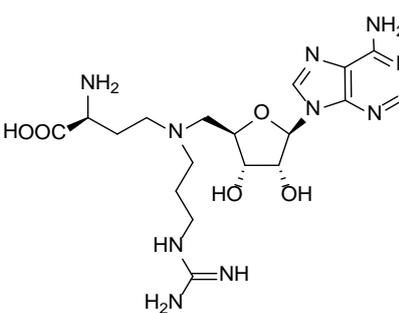
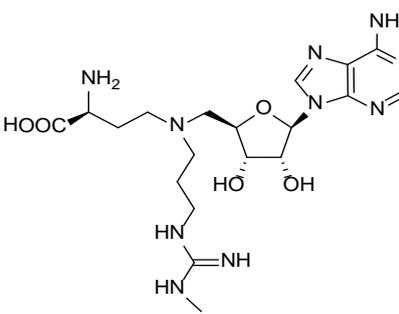
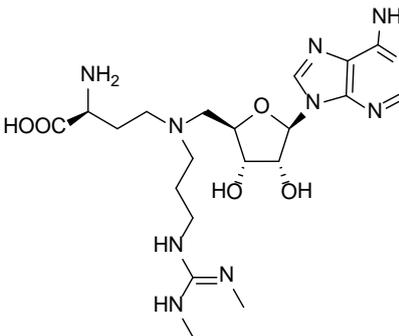
Composés	Structures	CI ₅₀ ± DS [μM]		RS ^(c)
		CARM1 ^(a,b)	PRMT1 ^(a,b)	
SFG		0,7 ± 0,3	0,6 ± 0,3	1
51		8,2 ± 0,3 ^(d)	22,0 ± 2,0 ^(d)	3
52		10,0 ± 1,0	21,0 ± 8,0	2
53		8,3 ± 4,2	34,0 ± 1	4

Table 3.3 : Activités d'inhibition (valeurs de CI₅₀) des analogues de bisubstrats contre CARM1 et PRMT1. a) Le substrat utilisé est le tétramère d'histone (H3H4)₂ et le ¹⁴C SAM radioactif est le donneur de méthyle; b) Les valeurs ont été déterminées à partir de deux expériences séparées au minimum; c) RS = rapport de sélectivité : PRMT1/CARM1. d) Valeurs de la littérature (Référence [16]): 6,2 ± 3,8 μM (PRMT1) et 1,3 ± 8,7 (CARM1), par réaction de méthylation *in vitro*.

Pour cette classe d'analogues, les valeurs de CI_{50} sont en général comprises entre 2,1 μM et 10,0 μM . Nous pouvons constater que la longueur de la chaîne carbonée permet d'induire une sélectivité au sein de la même famille d'enzyme. En effet, la chaîne à deux carbones permet d'avoir un facteur de sélectivité compris entre 14 et 54.

En revanche, les inhibiteurs contenant une chaîne à trois carbones ne sont seulement que deux à quatre fois sélectifs envers CARM1. Ces résultats sont d'autant plus intéressants, puisque d'après la bibliographie, l'équipe de Dowden et *al.*⁽¹⁵⁾ avait montré l'efficacité du composé **51** sur PRMT1. De plus, en 2011, elle avait aussi prouvé qu'une sélectivité au sein des PRMTs pouvait être obtenue en modifiant la longueur de la chaîne de ces analogues⁽¹⁶⁾. Pour une chaîne de trois carbones, nous avons pratiquement la même sélectivité pour CARM1 et PRMT1 puisque les ratios de leurs CI_{50} restent compris entre 2 et 4. En revanche, lorsque la chaîne est raccourcie d'un carbone, une meilleure sélectivité pour CARM1 est observée. En revanche, une meilleure sélectivité est observée pour PRMT1 lorsque la chaîne est rallongée d'un carbone supplémentaire (tableau **3.4**).

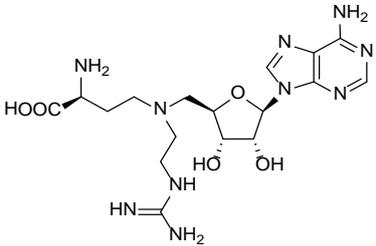
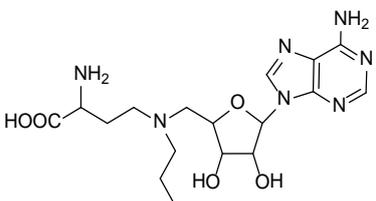
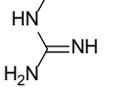
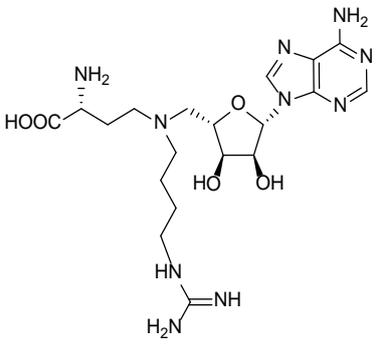
Composés	Structures	CI ₅₀ ± DS [µm]		RS ^(c)
		CARM1	PRMT1	
45		2,1± 0,6 ^(a)	113± 43 ^(a)	54
16		6,2 ± 3,8 ^(b)	13,3± 8,7 ^(b)	2
		8,2± 0,3 ^(a)	22± 2 ^(a)	3
17		2,9 ± 0,8 ^(b)	>100 ^(b)	-

Tableau 3.4: Tests d'activité comparant les analogues de bisubstrats décrits dans la littérature⁽¹⁵⁻¹⁶⁾ et nos composés. (a) Nos composés ; (b) Inhibiteurs décrits par Dowden et al⁽¹⁵⁻¹⁶⁾. c) RS = rapport de sélectivité : PRMT1/CARM1.

Compte tenu des bons résultats obtenus concernant les valeurs d'inhibition, des tests de cytotoxicité ont été effectués.

V.2 *Evaluation de la cytotoxicité des analogues de bisubstrats 45 à 47 et 51 à 53*

La cytotoxicité de ces analogues a été évaluée grâce à des tests MTT (ou bromure de 3-(4,5-diméthylthiazol-2-yl)-2,5-diphényltétrazolium) sur des cellules cancéreuses du sein (MCF-7). Ces cellules ont été choisies car les enzymes PRMTs sont fréquemment associées aux cancers du sein et de la prostate. Les tests ont été réalisés à Phytodia par le Dr. Denis Weltin et Céline Schmidt.

V.2.1 *Principe du test MTT*

Le test MTT permet de déterminer la cytotoxicité des produits ; c'est un test de colorimétrie standard qui mesure la prolifération des cellules par spectrophotométrie ultraviolet (UV).

Les sels de tétrazolium, comme le MTT, peuvent être réduits par les mitochondries des cellules vivantes actives en un composé coloré, le bleu de formazan. Le nombre de cellules vivantes pourra ainsi être déterminé par l'intensité de la coloration.

Six solutions à 1 mM dans l'eau sont préparées, puis les tests ont été réalisés sur des cellules cancéreuses du sein MCF-7 confluentes et placées dans une plaque 96 puits. Chacun des analogues est alors ajouté à huit concentrations différentes aux cellules qui sont ensuite incubées pendant 24 heures à 37°C dans du DMEM (à 0.1% de BSA). Le réactif MTT est ensuite additionné dans chaque milieu cellulaire et les cellules sont de nouveau incubées deux heures à 37°C. Chaque puits cellulaire est ensuite placé dans 100 µL de DMSO pour dissoudre le précipité de formazan et la densité optique de chaque échantillon est mesurée par spectrophotométrie UV à 595 nm pour déterminer le nombre de cellules actives métaboliquement.

Chaque essai a été répété trois fois et deux expériences de contrôle ont été exécutées en parallèle : les cellules seules sans aucun produit et ces mêmes cellules en présence d'un composé, l'Amytal, qui les détruit toutes.

V.2.2 *Résultats des tests MTT sur les analogues 45 à 47 et 51 à 53*

Les taux de viabilité cellulaire pour les deux échantillons sont représentés sur la figure **3.10**.

Pour la molécule de référence, la SFG, aucune cytotoxicité n'est observée puisque les taux de viabilité restent proches de 100%, à l'exception de la forte concentration de 500 μM .

Le même résultat est observé pour l'analogue **45** qui provoque une diminution de 30% des taux de viabilité des cellules à 500 μM . Ce composé est en revanche non toxique pour des concentrations plus faibles.

En ce qui concerne l'analogue **46**, les taux de viabilité des cellules chutent fortement à partir d'une concentration de 125 μM , pour atteindre uniquement 25% de viabilité à 500 μM . Cet analogue induit donc un effet cytotoxique à partir de 125 μM . Ce résultat est surprenant, mais pour l'instant, aucune explication n'a pu être proposée.

Le phénomène inverse se produit pour les analogues **47**, **51** à **53** : une augmentation de l'activité cellulaire est observée entre 4 et 500 μM , prouvant que ces analogues n'induisent aucune cytotoxicité sur les cellules MCF-7 (figure **3.10**).

Ces tests nous prouvent que, mis à part l'analogue **46**, ces composés ne sont pas toxiques entre 4 et 250 μM puisque leurs taux de viabilité restent proches ou supérieurs à 100%.

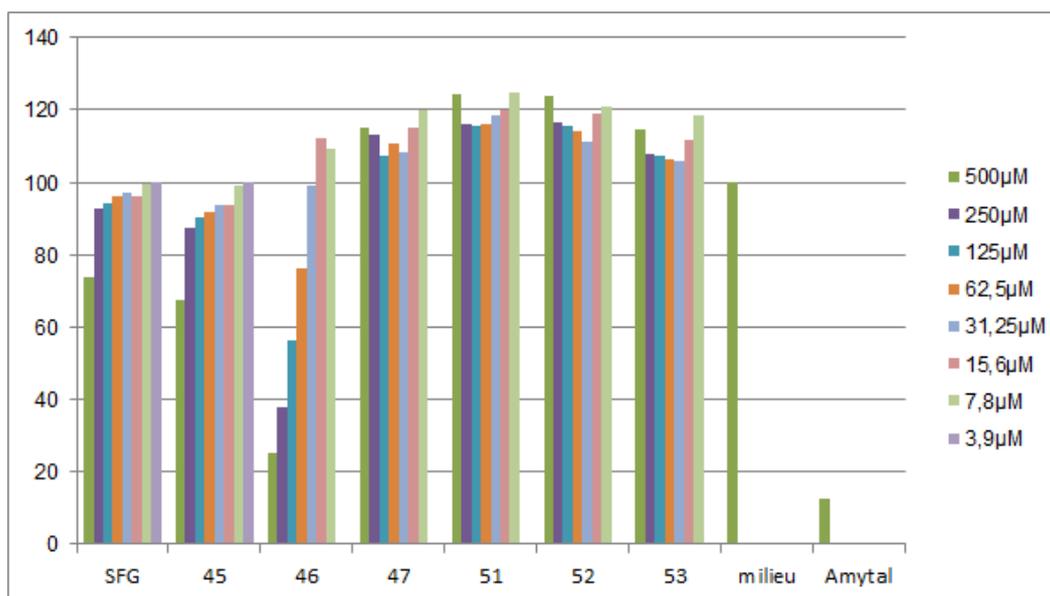


Figure 3.10 : Taux de viabilité cellulaire après 24 heures d'incubation des cellules MCF-7 au contact des analogues **45** à **47** et **51** à **53**.

V.3 Etudes structurales

V.3.1 Structures cristallographiques des analogues **45** et **51**

Les structures cristallines des analogues de bisubstrats **45** et **51**, en co-cristallisation dans CARM1, ont été obtenues par l'équipe du Pr. Jean Cavarelli et du Pr. Jean-Marie Wurtz. La structure de l'analogue **45** a été obtenue au synchrotron à l'ESRF de Grenoble, avec la construction mCARM1₁₃₀₋₅₀₇, et une résolution de 2,6 Å (figure 3.11). Sur cette structure cristalline, nous pouvons constater que le groupement guanidinium forme deux liaisons hydrogènes avec le Glu 267 à une distance de 2,25 Å et de 3,19 Å. Cet analogue est aussi capable de réaliser une liaison hydrogène avec une molécule d'eau à une distance de 1,95 Å. Cette même molécule d'eau forme également une liaison hydrogène avec le Glu 258 à une distance de 2,91 Å.

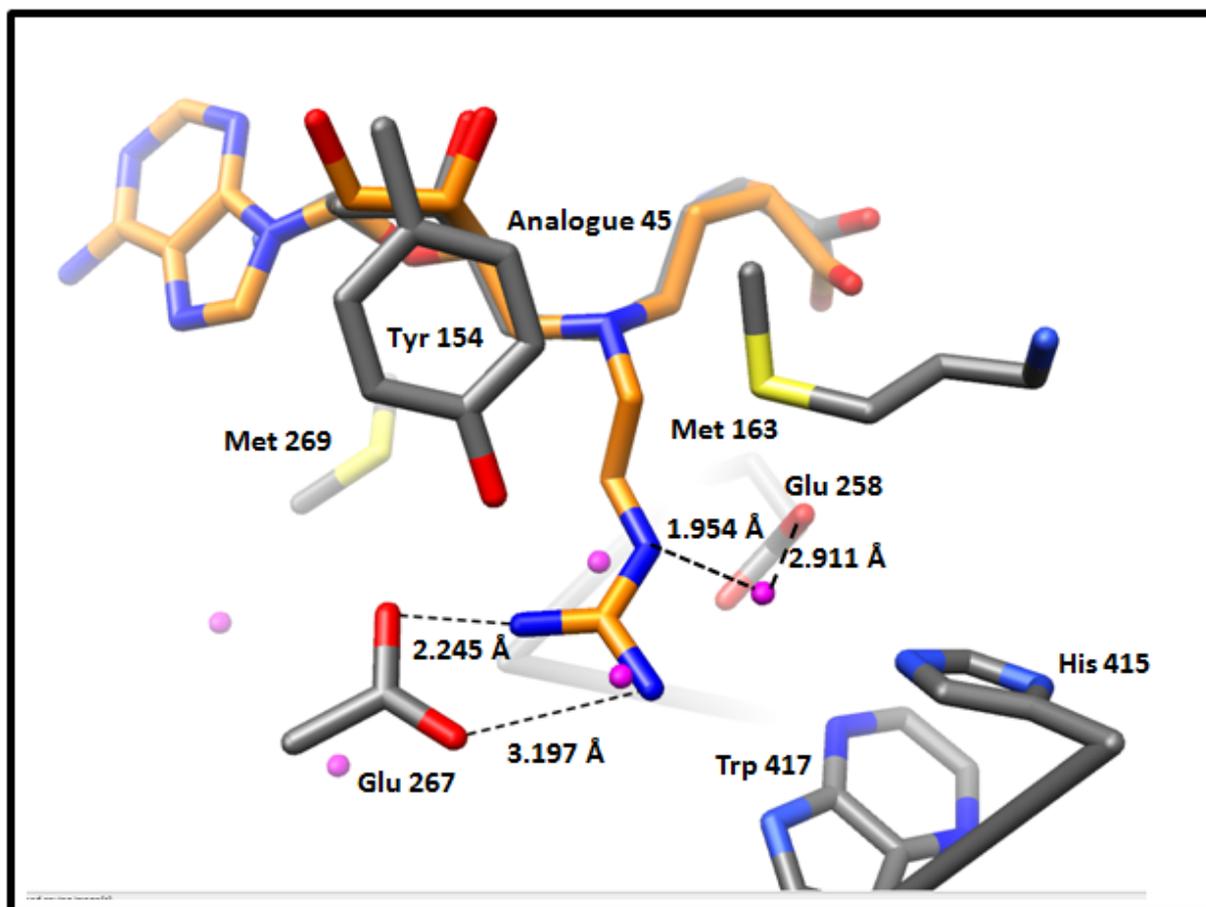


Figure 3.11 : Structure cristallographique de l'analogue **45** dans le site catalytique de mCARM1₁₃₀₋₅₀₇ avec une résolution de 2,6 Å . Le code couleur est le suivant : les atomes d'azote sont en bleu, les atomes d'oxygène en rouge et les atomes de soufre en jaune. Les atomes de carbone pour le cofacteur SAH ainsi que les résidus de la structure cristallographique 3B3F sont en gris tandis que l'analogue **45** est en orange. Les molécules d'eau sont en magenta.

La structure cristalline de l'analogue **51** à été visualisée au synchrotron du soleil à Paris avec la construction mCARM1₁₃₀₋₄₈₇ et une résolution de 2,1 Å (figure **3.12**). Sur cette structure cristalline, le groupement guanidinium forme deux liaisons hydrogènes avec le Glu 267 l'une à une distance de 2,04 Å et l'autre à une distance de 3,39 Å. Cet analogue forme également une liaison hydrogène avec une molécule d'eau à une distance de 2,23 Å. Cette molécule d'eau interagit également avec le Glu 258 en formant une liaison hydrogène à une distance de 2,91 Å.

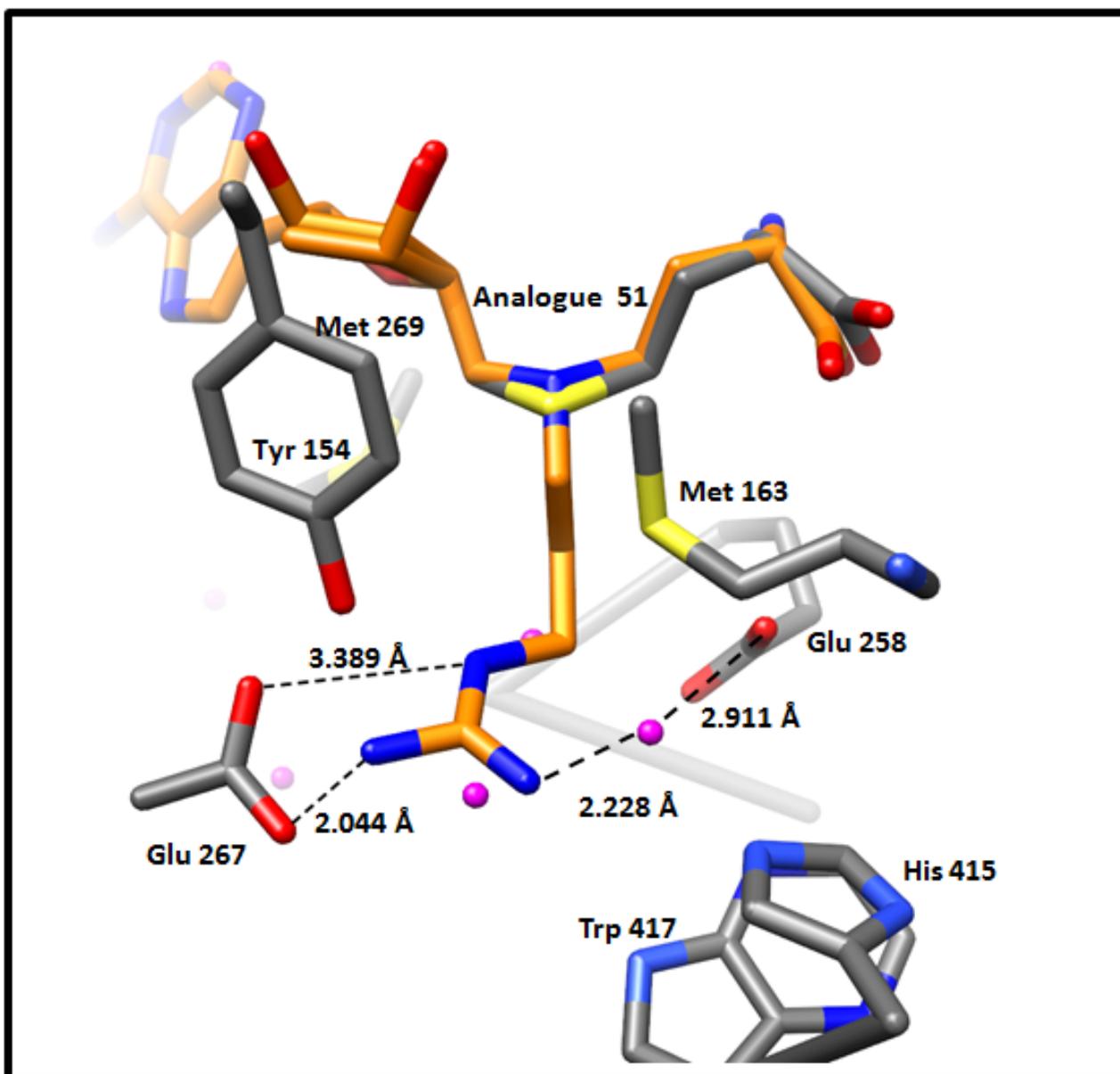


Figure 3.12 : Structure cristallographique de l'analogue **51** dans le site catalytique de mCARM1₁₃₀₋₄₈₇ avec une résolution de 2,1 Å. Le code couleur est le suivant : les atomes d'azote sont en bleu, les atomes d'oxygène en rouge et les atomes de soufre en jaune. Les atomes de carbone pour le cofacteur SAH ainsi que les résidus de la structure cristallographique 3B3F sont en gris tandis que l'analogue **51** est en orange. Les molécules d'eau sont en magenta.

Ces structures cristallographiques permettent, d'une part, de vérifier le modèle proposé par le Pr. Jean-Marie Wurtz et, d'autre part, d'expliquer la sélectivité des analogues entre CARM1 et PRMT1.

V.3.2 *Confirmation du modèle*

Dans le modèle proposé, l'arginine a été positionnée de façon à ce que le groupement guanidine du substrat soit plan: il peut ainsi interagir avec les deux acides glutamiques du site catalytique de CARM1, le Glu 267 et le Glu 258, en formant trois liaisons hydrogènes.

Sur la structure cristalline obtenue avec l'analogue **45**, nous pouvons constater que le groupement guanidine forme deux liaisons hydrogènes avec le Glu 267 à une distance de 2,2 Å et de 3,2 Å.

La longueur de la chaîne carbonée impose une certaine conformation qui, dans ce cas, est proche du modèle proposé. En effet, les distances entre les atomes d'azotes N1, N2 de l'arginine et le Glu 267 sont respectivement de 2,85 Å et 2,91 Å. La distance entre l'atome d'azote N3 de l'arginine et le Glu 258 est de 2,98 Å alors que la distance entre l'atome d'azote du groupement guanidine de l'analogue **45** et le Glu 258 est de 3,55 Å (figure **3.13**).

En ce qui concerne l'angle de liaison entre l'atome d'azote et la chaîne carbonée, il est de 113,8 degrés au lieu de 131,1 degrés dans le modèle ; de plus, l'angle de torsion des atomes N-C-C-N de l'analogue **45** est de 171,1 degrés. Sur la figure **3.13**, l'arginine modélisée dans le site actif de CARM1 se superpose pratiquement à l'analogue de bisubstrat **45**.

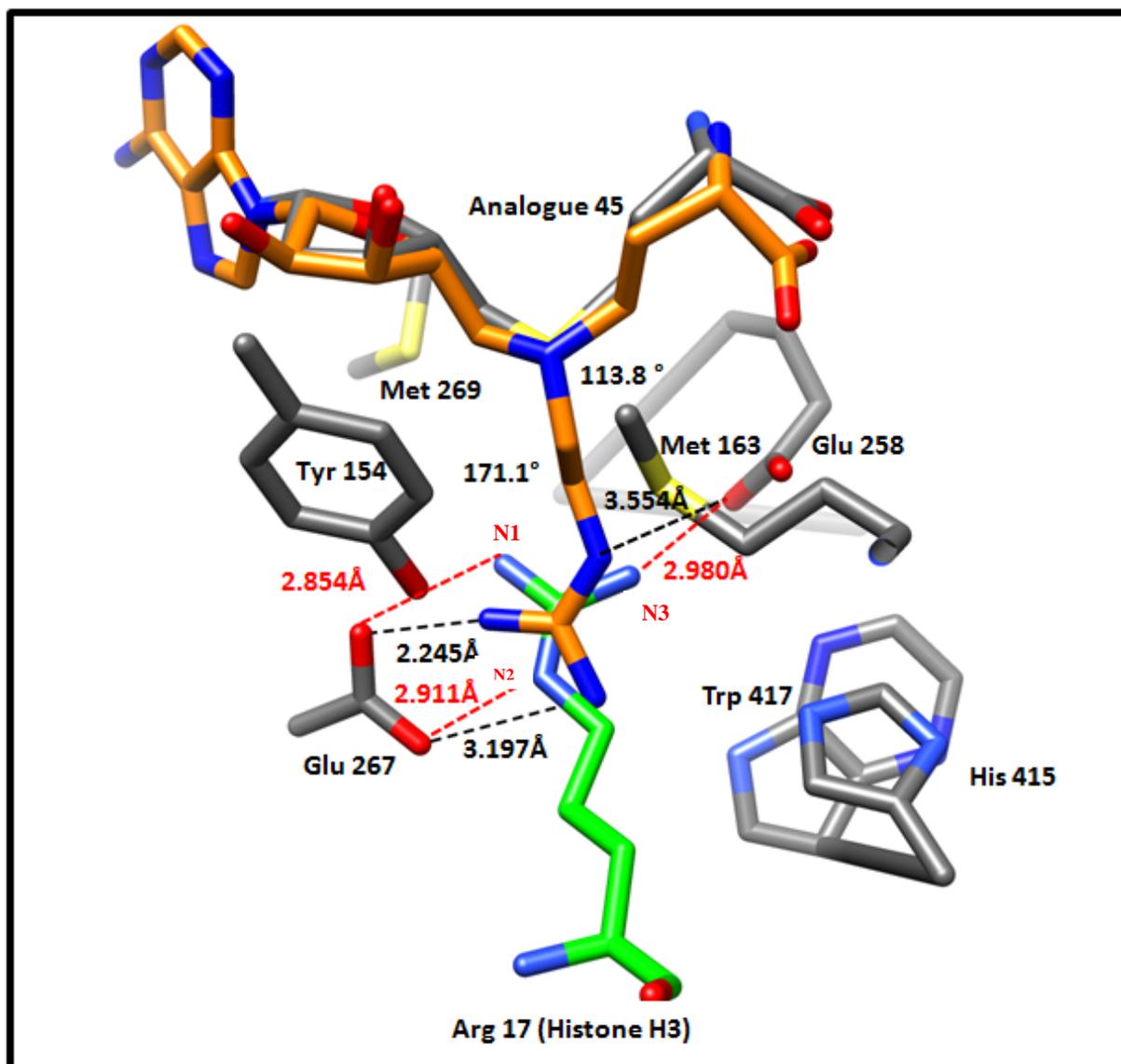


Figure 3.13 : Superposition de l'arginine du modèle et de l'analogue **45** dans le site actif de CARM1. Les distances entre les atomes d'azote du groupement guanidine de l'analogue **45** et les Glu 258 et 267 sont représentées en noir. Les distances entre les atomes d'azote du groupement guanidine de l'arginine et les Glu 258 et 267 sont représentées en rouge.

Pour l'analogue **51**, la chaîne carbonée est plus longue d'un carbone et l'ancrage de cet analogue dans CARM1 est un peu différent. Le groupement guanidinium est légèrement décalé vers la sortie de la poche de l'arginine et il est donc mal positionné au niveau des acides glutamiques.

Le groupement guanidine forme ainsi deux liaisons hydrogènes avec le Glu 267, l'une à une distance plus courte et l'autre à une distance plus longue par rapport au modèle proposé. Ces deux distances correspondent respectivement à 2,0 Å et 3,4 Å. Les distances entre les atomes d'azotes N1 de l'arginine, N2 et le Glu 267 sont respectivement de 2,9 Å et 2,5 Å. La distance entre l'atome d'azote N3 de l'arginine et le Glu 258 est de 2,9 Å alors que la distance entre l'atome d'azote du groupement guanidine de l'analogue **51** et le Glu 258 est de 5,4 Å (figure **3.14**).

La longueur de la chaîne impose une conformation légèrement différente avec un angle de liaison entre l'atome d'azote et la chaîne carbonée de 112° et un angle de torsion des atomes N-C-C-N de l'analogue **51** égal à 179,5°. Sur la figure **3.14**, le groupement guanidinium de cet analogue est légèrement en décalage par rapport à l'arginine modélisée.

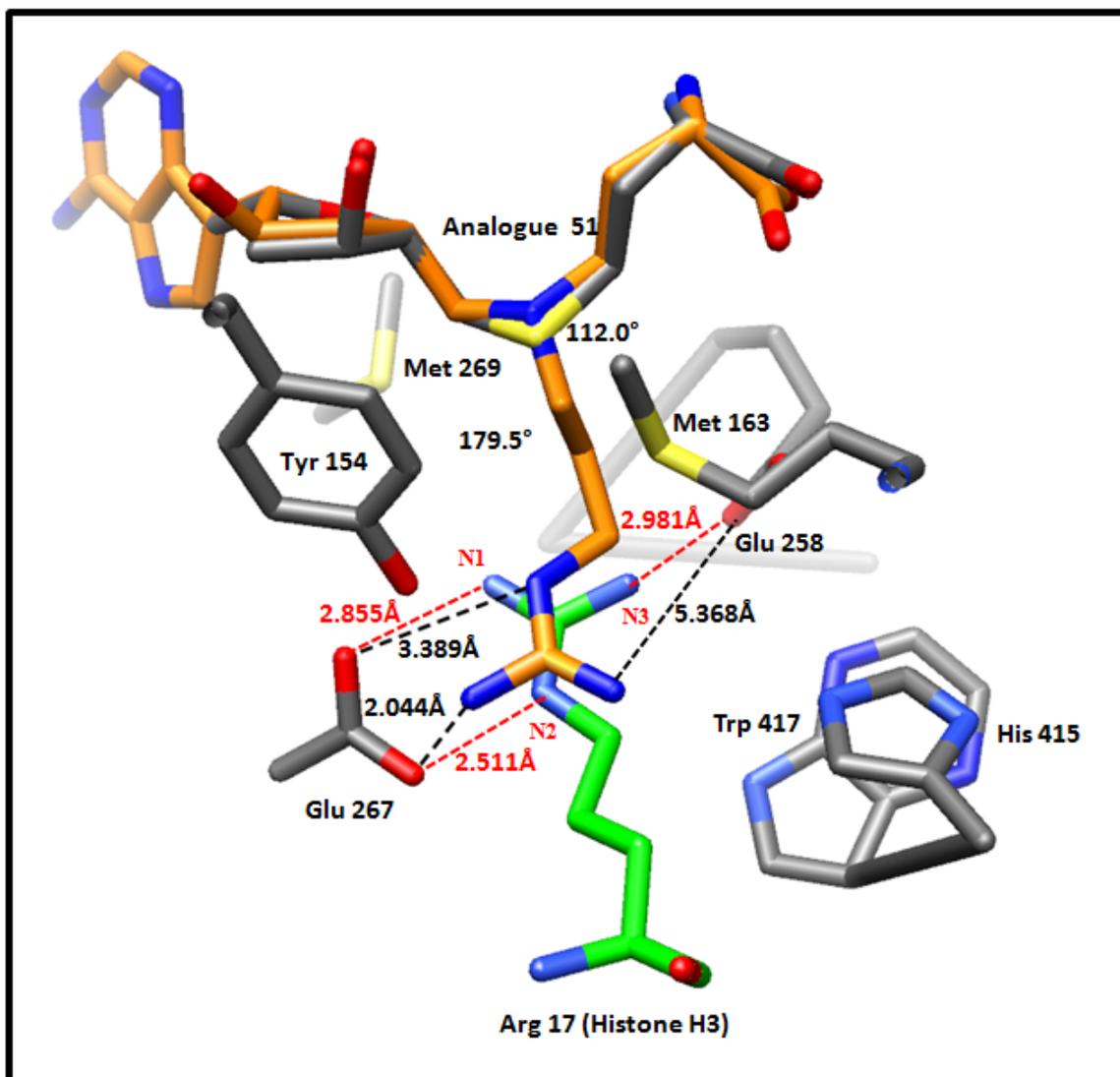


Figure 3.14: Superposition de l'arginine du modèle et de l'analogue **51** dans le site actif de CARM1. Les distances entre les atomes d'azote du groupement guanidine de l'analogue **51** et les Glu 258 et 267 sont représentées en noir. Les distances entre les atomes d'azote du groupement guanidine de l'arginine et les Glu 258 et 267 sont représentées en rouge.

Les distances entre les acides aminés importants du site actif du modèle et les deux analogues **45** et **51** sont représentés dans le tableau **3.4** :

Acide aminés	Distance avec le groupement guanidine de l'arginine dans le modèle (Å)	Distance du groupement guanidine 45 dans la structure RX (Å)	Distance du groupement guanidine 51 dans la structure RX (Å)
Met 163	4,01	3,14	3,78
Met 269	4,61	5,81	5,78
Tyr 154	3,22	3,80	3,40
His 415	4,82	3,77	3,40

Tableau 3.4: Distances en angströms entre l'arginine et des analogues **45** et **51** avec les acides aminés importants du site actif.

Nous pouvons constater que les distances entre les analogues **45** et **51** avec les acides aminés du site actif sont proches du modèle.

Ces structures cristallographiques valident le modèle, la justesse de la position de l'arginine dans le site actif, ainsi que l'environnement de celle-ci. L'ancrage de ces deux analogues est légèrement différent puisque la longueur de leurs chaînes carbonées impose une certaine conformation. Ces structures RX mettent également en évidence l'importance de certains acides aminés qui permettent d'expliquer la sélectivité observée pour ces analogues envers CARM1 et PRMT1.

V.3.3 *Comparaison des structures cristallographiques de CARM1 et PRMT1 pour expliquer la sélectivité*

L'analogue **45** présente les meilleurs effets d'inhibition avec une valeur de CI_{50} égale à 2,1 μ M pour CARM1 ; il possède aussi une sélectivité 54 fois meilleure par rapport à PRMT1.

Deux raisons peuvent expliquer l'efficacité de cet analogue au sein de CARM1 :

- La formation de liaisons hydrogènes est possible avec le Glu 267.
- Les analogues avec les guanidines méthylées ont des valeurs d'inhibition plus faibles que les guanidines simples. Ces observations sont dues à l'encombrement stérique induit par les méthyles du groupement guanidinium avec l'His 415 et la Tyr 157 qui sont présentes dans le site catalytique de CARM1.

Grâce à sa structure cristalline, nous avons observé que l'analogue **45** était capable de former des liaisons hydrogènes avec les Glu 267 et 258. Les analogues **46** et **47** possèdent respectivement un et deux méthyles sur le groupement guanidinium et ont une valeur d'inhibition plus basse que pour l'analogue **45**. Ces résultats peuvent s'expliquer par la difficulté des analogues **46** et **47** à former des liaisons hydrogènes avec le Glu 267 due à la présence des groupements méthyles sur le guanidinium. De plus, ces analogues sont aussi contraints stériquement par la Tyr 154 et l'His 415 du site actif. Pour les analogues **51** à **53** à chaîne carbonée plus longue, il est également plus difficile de former des liaisons hydrogènes avec le Glu 267 et la contrainte stérique avec la tyrosine et l'histidine en est d'autant plus forte.

Pour expliquer la sélectivité entre PRMT1 et CARM1, il faut comparer les sites actifs de ces deux enzymes. En comparant les structures de leurs sites actifs, nous pouvons constater que certains de leurs acides aminés se superposent : les Glu 258 et 144, les Tyr 46 et 154, les

Met 269 et 154 et les His 415 et 293 (respectivement pour CARM1 et PRMT1). Néanmoins, le Glu 267 dans CARM1 ne se superpose pas au Glu 153 de PRMT1 ; en effet, la première pointe vers la poche du sulfonium alors que la seconde se situe à l'extérieur de la poche de l'arginine⁽⁶⁾.

Dans notre modèle, c'est la formation de liaisons hydrogènes et de liaisons électrostatiques entre le guanidinium et le Glu 267 qui explique l'efficacité de nos inhibiteurs. Par conséquent, la sélectivité entre CARM1 et PRMT1 peut s'expliquer par l'impossibilité de cette dernière à former des liaisons hydrogènes et électrostatiques avec le Glu 153 dans son site catalytique. Les positions relatives de la SAM et de la guanidine jouent donc un rôle dans la sélectivité au sein des PRMTs.

VI. Conclusions et perspectives

Pour l'exploration de la « poche de l'arginine », nous nous sommes basés sur la structure cristallographique de CARM1 et sur un mécanisme proposé dans la littérature. Celui-ci met en jeu une méthyltransférase, la GAMT, un donneur de méthyle, la SAM et un substrat, la GAA. Ce substrat et la SAM présentent le même ancrage au niveau du site actif de cette enzyme que dans CARM1 ; c'est pourquoi nous avons pu nous inspirer de cette réaction pour proposer un modèle mécanistique.

Pour cette étude, six analogues de bisubstrats ont été synthétisés et deux structures cristallographiques ont été obtenues pour les analogues **45** et **51** qui ont co-cristallisé dans CARM1. Ces structures RX nous ont permis de valider notre modèle et d'avoir une idée précise des interactions qui peuvent avoir lieu à l'état de transition entre l'arginine et les résidus du site actif de CARM1. Nous avons montré que ces interactions sont gouvernées par un environnement électrostatique et stérique imposé par les différents acides aminés du site catalytique.

Ce modèle permet également de confirmer l'importance du Glu 267 dans le site actif de CARM1 et de répondre à la question posée par Pearl et *al.*⁽⁶⁾ : la conformation de l'arginine dans CARM1 peut être différente que dans les autres PRMTs.

De plus, des tests d'activité ont été effectués sur ces composés qui ont montré de très bonnes inhibitions envers CARM1 ; l'analogue **45** est ainsi cinquante-quatre fois plus sélectif par rapport à PRMT1.

Ces analogues nous ont donc permis d'explorer la « poche de l'arginine » ; dans le prochain chapitre, nous avons décidé de continuer notre étude en sortant de cette poche et en explorant le « domaine de fixation du peptide ». La première stratégie employée pour cette exploration est une approche par réaction *in-situ* entre une molécule hameçon capable de piéger un peptide mimant la séquence du substrat. Pour ces travaux, nous avons choisi comme substrat l'histone H3.

Références

1. J. D., Gary, S. Clarke, *Prog. Nucleic Acid Res. Mol. Biol.*, **1998**, *61*, 65–131.
2. Obianyoy, T.C. Osborne, P.R. Thompson, *Biochemistry*, **2008**, *47*, 10420–10427.
3. T.M. Lakowski, A. Frankel, *J. Biol. Chem.*, **2008**, *283*, 10015–10025.
4. Y. Feng, N. Xie, M. Jin, M. R. Stahley, J. T. Stivers, Y. G. Zheng, *Biochemistry*, **2011**, *50*, 7033–7044.
5. N. Troffer-Charlier, V. Cura, P. Hassenboehler, D. Moras, J. Cavarelli, *EMBO J.*, **2007**, *6*, 4391–4401.
6. W.W. Yue, M. Hassler, S. M. Roe1, V. Thompson-Vale, L. H.Pearl, *EMBO J.*, **2007**, *26*, 4402-4412.
7. X. Zhang, X. Cheng, *Structure*, **2003**, *11*, 509-20.
8. X. Zhang, L. Zhou, X. Cheng, *EMBO J.*, **2000**, *19*, 3509–3519.
9. Y.H. Lee, S.S. Koh, X. Zhang, X. Cheng, M.R. Stallcup, *MolCell Biol*, **2002**, *22*, 3621–3632.
10. a) M. Wyss, R. Kaddurah-Douk, *Physiol Rev*, **2000**, *80*, 1107–1213. b) R. Murphy, G. McConell , D. Cameron-Smith, K. Watt, L. Ackland, B. Walzel, T. Wallimann, R. Snow, *Am J Physiol Cell Physiol*, **2001**, *280*, 415–422.
11. J. Komoto, T. Yamada, Y. Takata, K. Konishi, H. Ogawa, T. Gomi, M. Fujioka, F.Takusagawa, *Biochemistry*, **2004**, *43*, 14385–14394.
12. P. Velichkova, F. Himo, *J. Phys. Chem. B*, **2006**, *110*, 16-19.
13. E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, *J. Com-put. Chem.*, **2004**, *25*, 1605–1612.
14. Avogadro V1.0. (2010). Retrieved from Avogadro: an open-source molecular builder and visualization tool. Version 1.XX. <http://avogadro.openmolecules.net/>

15. J. Dowden, W. Hong, R. V. Parry, R. A. Pike, S. G. Ward, *Bioorg.Med.Chem.Lett.*, **2010**, *20*, 2103-2105.
16. J. Dowden, R. A. Pike, R. V. Parry, W. Hong, U. A. Muhsen, S.G. Ward, *Org.Biomol.Chem*, **2011**, *9*, 7814-7821.
17. T. Gers, D. Kuncce, P. Markowski, J. Izdebski, *Synthesis*, **2004**, *1*, 37-42.
18. J.L. LaMattina, P.A. McCarthy, L.A. Reiter, W.F. Holt, L.A. Yeh, *J. Med. Chem.* , **1990**, *33*, 543-52.
19. C.S. Pugh, R. T. Borchardt, H. O. Stone, *J. Biol. Chem.*, **1978**, *253*, 4075-4077.

Chapitre 4

Approche pour l'exploration du « domaine de fixation du peptide » par réaction in-situ

Chapitre 4

*Approche pour l'exploration du « domaine de fixation du peptide » par réaction *in-situ**

Après avoir exploré la « poche du sulfonium » et la « poche de l'arginine », nous avons envisagé de poursuivre notre stratégie grâce à l'exploration du « domaine de fixation du peptide ». Cette stratégie va nous permettre d'étudier les interactions entre le cofacteur et le substrat. Pour ce faire, des molécules hameçons ont été choisies pour réagir *in-situ* avec des peptides du substrat. Dans le paragraphe suivant, la stratégie employée va être détaillée.

I. Stratégie employée pour explorer le « domaine de fixation du peptide » par réaction *in-situ*

Pour explorer le « domaine de fixation du peptide », nous avons tenté de capturer des peptides à l'aide de molécules hameçons présentant le motif de la SAM. Cette stratégie repose sur la formation d'un lien covalent entre le peptide et la molécule hameçon au sein de la poche catalytique. Elle nécessite que l'enzyme rapproche les deux réactifs pour favoriser la formation d'adduits. Afin de mettre en évidence la formation de la liaison entre la molécule hameçon et le peptide cible au sein du site actif, nous avons choisi, comme méthode d'analyse, la diffraction aux rayons X. Cette dernière a été réalisée par l'équipe du Pr. Jean Cavarelli. Pour former ce lien covalent entre les molécules hameçons et le substrat, trois types de réaction ont été utilisés (figure 4.1) :

- Une réaction de click⁽¹⁾ entre un peptide alcyne et une molécule hameçon possédant une fonction azoture (figure 4.1 a)
- Une réaction de substitution nucléophile entre des molécules hameçons électrophiles capables de piéger un peptide cystéine (figure 4.1 b)

- Une réaction de condensation entre des molécules hameçons électrophiles capables de capturer un peptide arginine (figure 4.1 c).

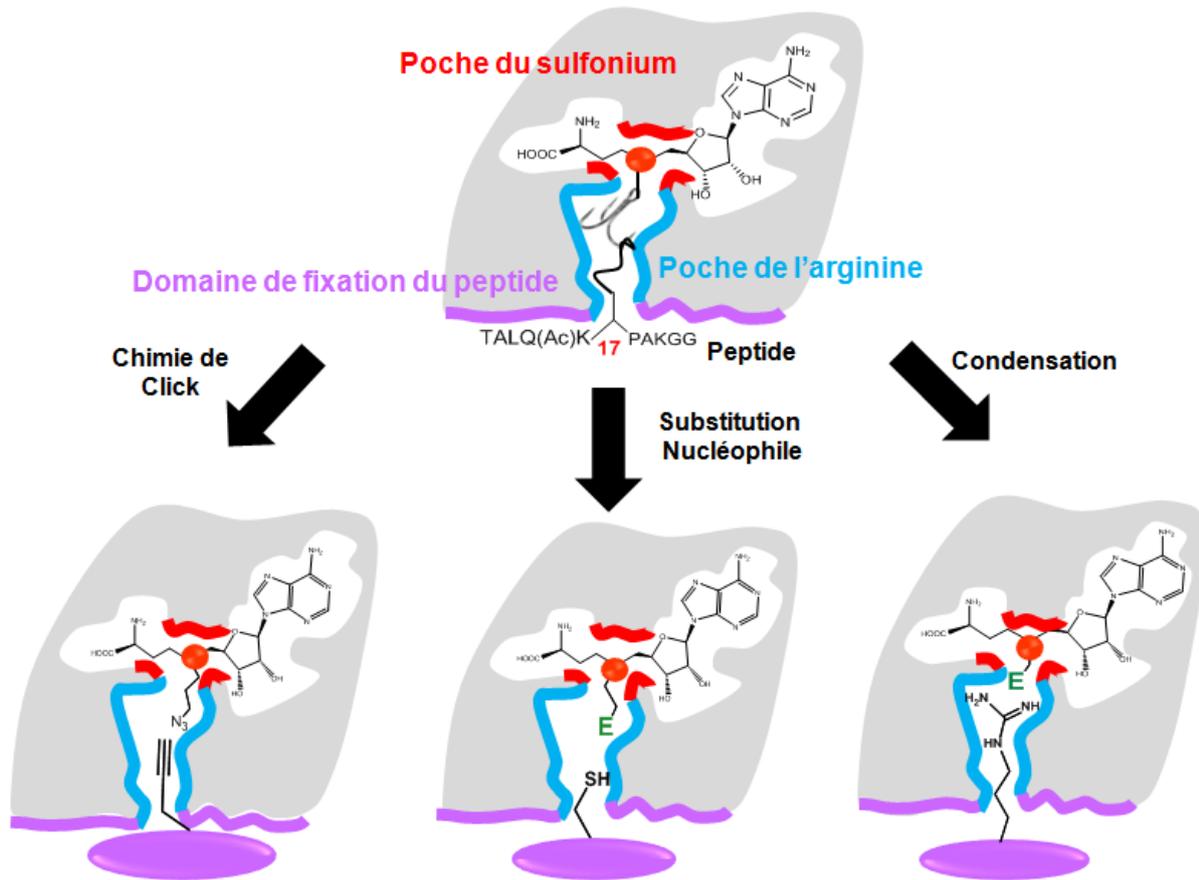


Figure 4.1 : Pour capturer un peptide dans le site catalytique de CARM1, la stratégie qui a été employée est une réaction *in-situ* entre une molécule hameçon et un peptide. a) Réaction de click entre un peptide alcyne et une molécule hameçon avec une fonction azoture. b) Réaction de substitution nucléophile entre une molécule hameçon électrophile et un peptide cystéine. c) Réaction entre une molécule hameçon électrophile et un peptide arginine.

Ces trois approches seront développées en détail dans les paragraphes suivants. Pour la première approche, le principe et les antécédents exposés dans la littérature seront décrits ; nous finirons par la description des résultats que nous avons obtenus. Pour les deux autres approches de la stratégie, seuls le principe et les résultats seront explicités.

II. Réaction *in-situ* entre une molécule hameçon azoture et un peptide alcyne par réaction de click

Pour commencer la stratégie exposée précédemment, nous avons voulu capturer le substrat grâce à des analogues *via* une réaction de cycloaddition. Pour cela, une fonction azoture a été incorporée dans la structure de l'analogue azoté **54** de la SAM. En parallèle, pour les substrats, un peptide **PJ207** a été synthétisé, à l'intérieur duquel une fonction alcyne remplace l'arginine 17 ; ce peptide permet de mimer la séquence 12-22 de l'histone H3.

Dans un premier temps, la molécule hameçon azoture **54** a été co-cristallisée dans CARM1 afin de vérifier si ce composé conservait son affinité pour l'enzyme, malgré la présence de l'azoture. Cette étape ayant été validée, des essais de click *in-situ* dans CARM1 ont été réalisés entre cette molécule et le peptide **PJ207**. Ces essais se déroulent en respectant la procédure suivante : la molécule **54** et le peptide **PJ207** sont utilisés à une concentration de 2 mM et sont mélangés dans de l'eau ; puis l'enzyme CARM1, utilisée à une concentration de 40 μ M, est introduite au mélange réactionnel. Ce dernier est ensuite incubé à 37°C pendant une heure, puis est centrifugé pendant dix minutes pour enlever le précipité éventuellement formé par la protéine. Le surnageant est ensuite mis en co-cristallisation en présence d'additifs, appelés agents de cristallisation, afin d'obtenir des cristaux. Suivant les conditions de cristallisation utilisées, les cristaux peuvent apparaître ou non et leur taille peut également varier. Les additifs utilisés sont :

- Une solution tampon qui permet d'obtenir les meilleures conditions de pH pour que l'enzyme soit active, c'est-à-dire à pH = 8 pour CARM1.
- Un polymère qui agit par exclusion du solvant au voisinage de la protéine ; le polymère le plus fréquemment utilisé est le polyéthylène glycol (PEG).

- Un sel, non dénaturant pour la protéine, qui modifie la force ionique et augmente les interactions hydrophobes entre les molécules de la protéine en excluant le solvant : généralement, ce sont le chlorure de sodium, le sulfate et acétate d'ammonium qui sont employés.

Dans le tableau ci-dessous, sont indiqués les additifs et leur concentration pour les essais de co-cristallisation dans CARM1 (tableau **4.1**).

Les additifs les plus utilisés pour les essais de co-cristallisation sont le tampon Tris HCl (100 mM), un polyéthylène glycol (2000MME 15-25%) et une solution de chlorure de sodium (100-300 mM).

Tampons	Agents précipitants	Sels
Tris HCl, pH = 8, 100 mM	PEG1500 10-30%	NaCl 100-300 mM
HEPES, pH = 8, 100 mM	PEG2000MME 15-25%	-
BTP, pH = 8, 100 mM	PEG3350 10-30%	-

Tableau 4.1 : Additifs utilisés pour les essais de co-cristallisation dans CARM1.

La cristallisation se fait par la méthode de "cristallisation par échange de vapeur en gouttes suspendues" au cours de laquelle des gouttes de cristallisation sont formées à partir de 2 μL de mélange réactionnel et de 2 μL d'additifs. Ces gouttes sont ensuite conservées à une température de 24°C jusqu'à l'obtention de cristaux ayant la bonne taille pour la diffraction. Cette étape peut être longue et peut durer environ un mois.

L'autre méthode utilisée pour identifier la présence d'adduits est la spectrométrie de masse des complexes non covalents (figure 4.2).

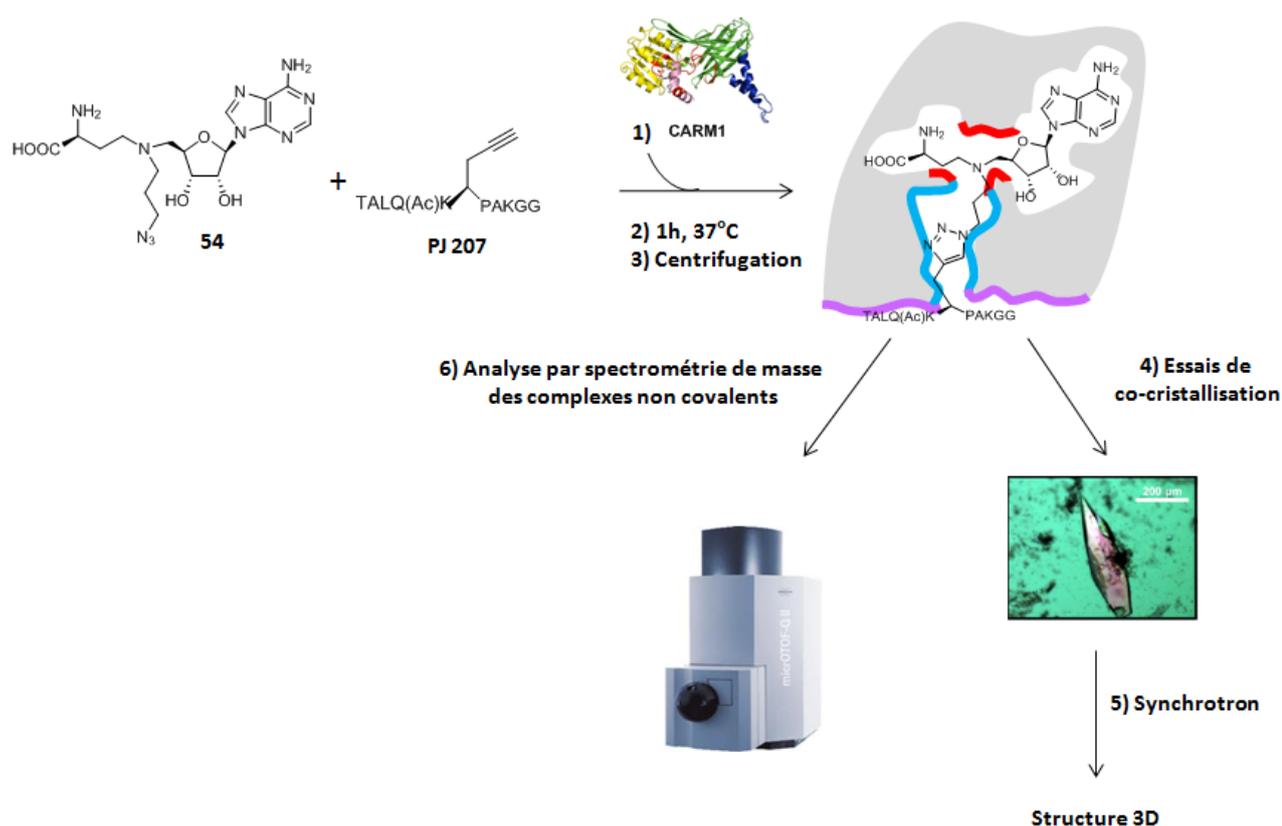


Figure 4.2 : Stratégie pour capturer le peptide par réaction de click *in-situ*. 1) Introduction de l'enzyme CARM1 au mélange molécule hameçon/peptide. 2) Incubation du mélange réactionnel à 37°C pendant une heure. 3) Centrifugation. 4) Essais de co-cristallisation. 5) Analyse des cristaux au synchrotron. 6) Analyse du mélange réactionnel par spectrométrie de masse des complexes non-covalents.

Avant de décrire les résultats obtenus, une étude bibliographique portant sur la chimie de click va être présentée, notamment celle utilisée en chimie combinatoire.

II.1 Antécédents, dans la littérature, de la réaction de click *in-situ* employée en chimie combinatoire

Dans cet exemple, une enzyme permet de catalyser la réaction de click, grâce à l'adoption d'une bonne conformation, et ce, sans utiliser de catalyseur chimique. Cette technique a été initiée par l'équipe de Mock⁽²⁾ puis développée par Sharpless et *al.*⁽³⁾ en 2002. Contrairement aux réactions chimiques catalysées au cuivre ou au ruthénium, ces réactions nécessitent un peu plus de temps, suivant la quantité de protéine utilisée. Le but est d'identifier le produit formé par la protéine puis d'en obtenir une quantité suffisante pour l'analyser par spectrométrie de masse. En général, la technique utilisée est la spectrométrie de masse couplée à la chromatographie liquide à haute performance (CHLP-MS). Ainsi, les deux régioisomères, obtenus après la réaction de click, peuvent être séparés et identifiés (figure 4.3)⁽⁴⁾.

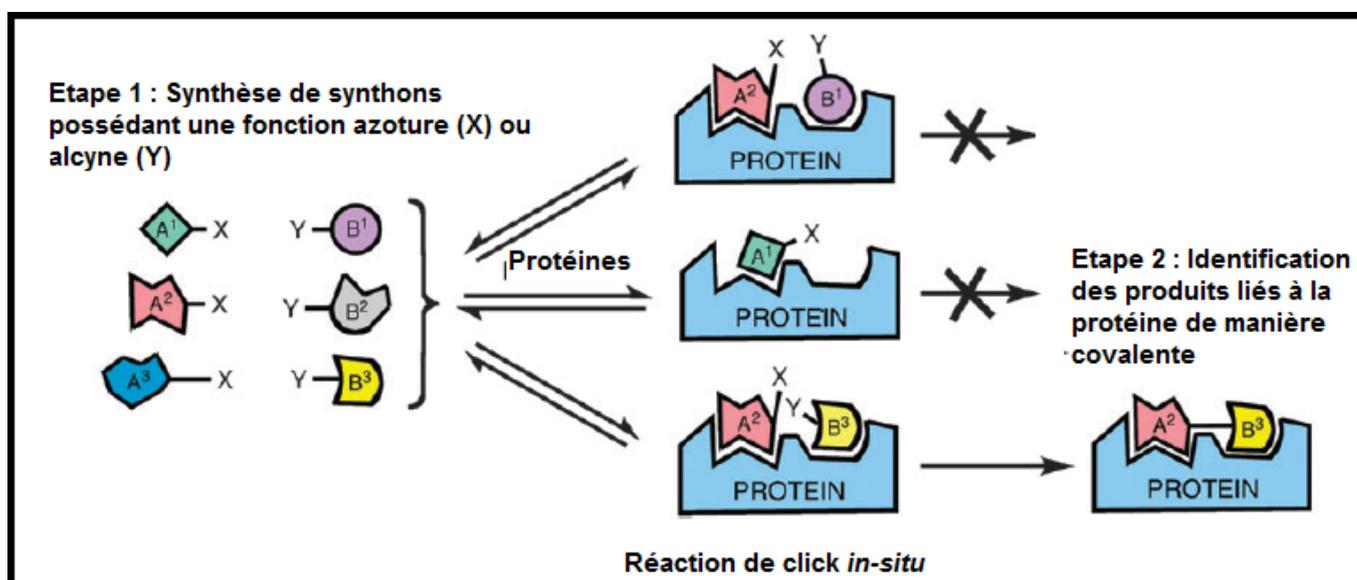


Figure 4.3 : Réaction *in-situ* entre des synthons possédant une fonction azoture ou alcyne, puis identification des isomères issus de la réaction de click par CHLP-MS⁽⁴⁾.

II.1.1 Réaction *in-situ* au sein de l'acétylcholine estérase (AChE)

La première cible étudiée pour la réaction *in-situ* était l'acétylcholine estérase (AChE), une cible importante impliquée dans les fonctions neurobiologiques. Cette enzyme a un site actif divisé en deux domaines. Plusieurs inhibiteurs bivalents ont été synthétisés pour inhiber la cible efficacement⁽⁵⁾. La stratégie employée par Sharpless et *al.*⁽³⁾ a été d'inhiber les deux sites actifs de l'enzyme en incorporant deux ligands pouvant se lier de façon covalente par réaction *in-situ*. Cette réaction a été réalisée à 22°C en utilisant une solution tampon de citrate d'ammonium à une valeur de pH égale à 7 (figure 4.4).

Les ligands synthétisés ont été mis en présence de l'AChE et ont montré des constantes d'affinités (Kd) de l'ordre du nanomolaire ; la réaction de cycloaddition a eu lieu entre ces inhibiteurs et l'enzyme et, après 90 minutes, une conversion de 80 % a été observée. Pour finir, les deux produits de cycloaddition obtenus sont encore plus affins avec l'enzyme, car ils présentent des constantes d'affinité de l'ordre du femtomolaire.

Dans le butanol et sans enzyme ou catalyseurs chimiques, la même conversion est observée après neuf jours. Pour finir, si la réaction devait avoir lieu à 18°C et dans la solution tampon, il faudrait attendre 40 ans pour atteindre le même taux de conversion de 80%.

Ces travaux montrent l'importance de l'acétylcholine estérase pour catalyser les réactions de cycloadditions *in-situ*.

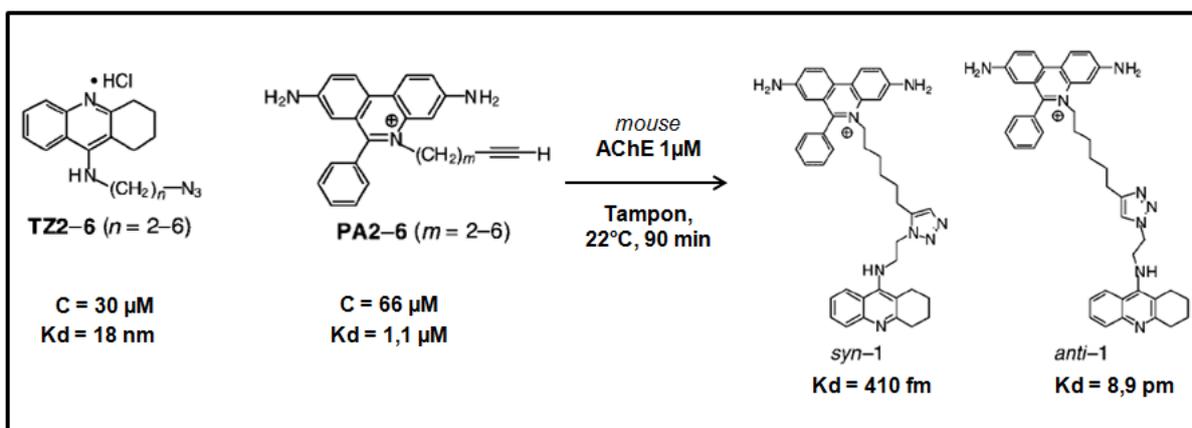


Figure 4.4 : Réaction *in-situ* dans l'acétylcholine estérase (AChE)⁽³⁾.

Compte tenu des résultats intéressants obtenus avec l'acétylcholine estérase, différentes réactions de cycloaddition *in-situ* ont été réalisées grâce à d'autres enzymes. A titre d'exemple, nous allons maintenant exposer un exemple de réaction catalysée par la protéase VIH-1.

II.1.2 Réaction *in-situ* au sein de la protéase du VIH-1 (Pr SF-2-WTQ7K-Pr)

La protéase VIH-1 est une enzyme cruciale impliquée dans le SIDA⁽⁶⁾ et la découverte d'inhibiteurs de cette enzyme est donc primordiale. L'équipe de Fokin et *al.*⁽⁷⁾ a mis au point la réaction *in-situ* entre un azoture et un alcyne possédant des valeurs d'inhibition de l'ordre du μM. Ces deux réactifs ont été introduits dans une solution tampon d'acide 2-morpholinoéthanesulfonique monohydraté (MES), à pH = 5 et mélangés pendant 24 heures à 23°C, en présence de la protéase du VIH-1. Le produit de cycloaddition **10** obtenu s'est révélé être un inhibiteur plus efficace que les réactifs : il présente une valeur d'inhibition de l'ordre du nanomolaire.

Dans les mêmes conditions, avec la protéase du VIH-1, des traces de triazole **10** ont pu être obtenues (figure 4.5). Grâce à ces travaux, nous avons de nouveau souligné l'importance des enzymes pour les réactions *in-situ*.

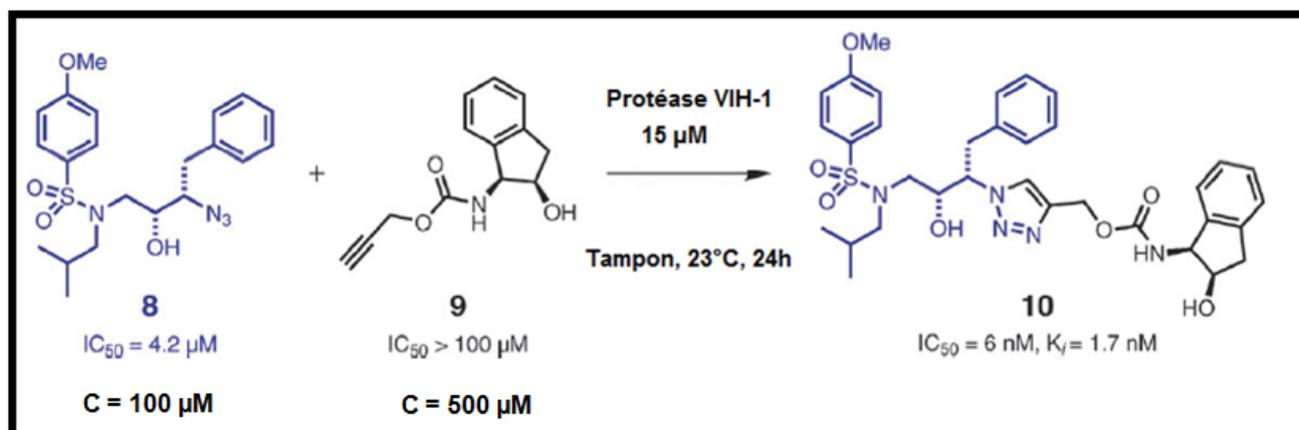


Figure 4.5 : Inhibiteur de la protéase VIH-1 obtenu à partir d'inhibiteurs de l'enzyme possédant une fonction alcyne ou une fonction azoture au sein de l'enzyme ⁽⁷⁾.

Le point important de cette expérience est la formation d'un produit de click à partir de synthons utilisés à différentes concentrations. En effet, la réaction *in-situ* décrite ici est réalisée en présence de synthons possédant une concentration en ligand sept fois plus élevée que celle utilisée pour l'expérience décrite par Sharpless et une concentration en enzyme quinze fois plus élevée également. Ces concentrations en ligand et en enzyme conduisent à une réaction plus lente et moins efficace ; en effet, après 24 heures, seules des traces de produit de cycloaddition (0,3% de conversion) sont observées tandis qu'une conversion de 80% est obtenue en seulement 90 minutes, en présence d'AChE.

Compte tenu de ces résultats, nous avons tout de même envisagé de tester cette réaction *in-situ* dans notre système. En effet, dans notre cas, il faut utiliser les synthons à une concentration de l'ordre du mM pour essayer de saturer un maximum de sites de l'enzyme CARM1 qui en compte huit. Dans nos conditions de co-cristallisation, le mélange réactionnel est chauffé à 37°C pendant 1 heure, puis l'étape de co-cristallisation est très lente et dure en

moyenne un mois. Même si, dans notre cas, la réaction de cycloaddition entre l'azoture **54** et le peptide **PJ207** se révélerait lente, elle aurait largement le temps de se produire pendant la période de co-cristallisation d'un mois.

Pour réaliser la réaction *in-situ*, trois composants seront mis en jeu : l'enzyme CARM1, une molécule hameçon et un peptide. Tout au long de ce chapitre, le terme « mélange réactionnel » sera attribué à ces trois composants.

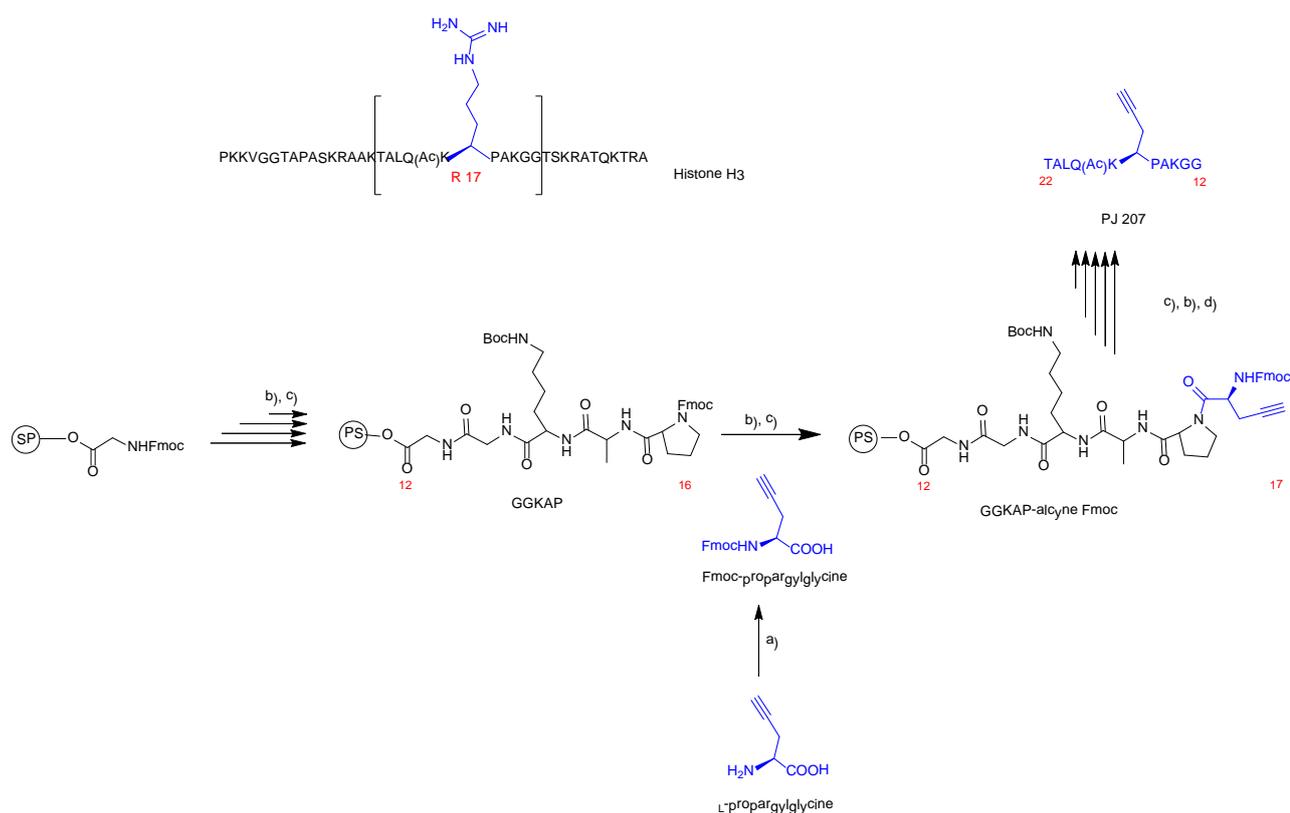
II.2 Essais de co-cristallisation du mélange réactionnel dans CARM1

II.2.1 Synthèse du peptide PJ207

Le peptide **PJ207**, ainsi que ceux présentés dans le chapitre suivant, ont été synthétisés par Pascal Eberling, de l'IGBMC. La particularité de ce peptide de onze acides aminés est la présence d'un alcyne à la place de l'arginine 17 ; de plus, ces acides aminés sont reconnus par CARM1. Ce peptide a été synthétisé selon la méthode en phase solide (SPPS) mise au point par Merrifield⁽⁸⁾, à l'aide d'un synthétiseur automatique de peptides ABI433A (figure **4.6**).

La chimie utilisée est basée sur l'utilisation du groupement 9-fluorénylméthoxy-carbonyle (Fmoc)⁽⁹⁻¹⁰⁾ pour la protection de la fonction amine de l'acide aminé, en combinaison avec des groupements tertio-butyles et trityles pour la protection des fonctions secondaires des acides aminés trifonctionnels. Des résines de type Wang⁽⁹⁾ sont utilisées pour la synthèse du **PJ207**. La première étape consiste à faire une réaction d'estérification entre l'acide aminé souhaité et la fonction hydroxyle de la résine. Aujourd'hui, les résines Wang pré-chargées avec le Fmoc-acide aminé désiré sont disponibles commercialement. Pour la synthèse du **PJ207**, une résine de ce type a été utilisée : la résine Fmoc-glycine-Wang. Lors de la synthèse, l'accroissement de la chaîne peptidique se fait selon une série d'étapes de déprotection (Fmoc) et couplage peptidique suivant le nombre d'acides

aminés de la séquence. Les Fmoc-acides aminés sont couplés par un agent de couplage qui est le *o*-benzotriazole-*N,N,N',N'*-tetraméthyl-uronium-hexafluoro-phosphate (HBTU) associé à l'hydroxybenzotriazole (HOBT) dans la DMF et en présence de *N,N'*-diisopropyléthylamine comme base. La quantité de peptide synthétisée est de l'ordre de 0,1 mM. Lorsque la séquence GGKAP est synthétisée alors la proline 16 est déprotégée en milieu basique pour pouvoir introduire la Fmoc-propargylglycine en position 17 par une réaction de couplage. Puis les étapes sont répétées jusqu'à obtenir la séquence désirée de onze acides aminés pour finalement déprotéger le dernier Fmoc-acide aminé côté N-terminal par traitement basique (pipéridine 20%, DMF). S'en suit une étape de clivage du peptide du support solide (résine) et la déprotection des chaînes latérales des acides aminés par un traitement acide (TFA) (figure 4.6).



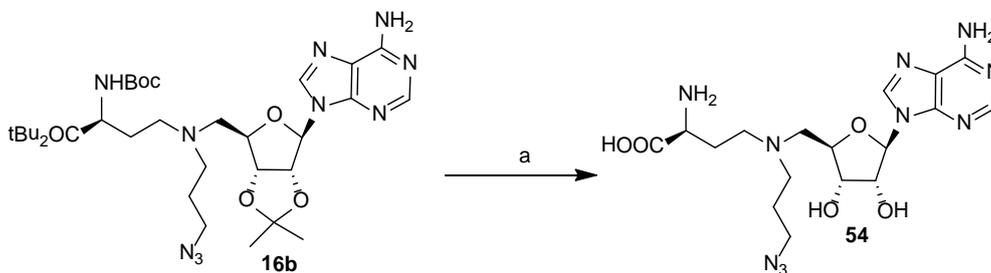
Conditions réactionnelles pour la synthèse de l'alkyne Fmoc à partir de la *L*-propargylglycine : a) Na_2CO_3 (10%), Fmoc-Osu, Dioxane, 0°C à température ambiante, 12h, rendement quantitatif. b) Déprotection du Fmoc : pipéridine (20%), DMF. c) Synthèse peptidique sur support solide : méthode de couplage peptidique : HBTU (10 éq), HOBT (10 éq), DIEA (15 éq), DMF. d) Clivage de la résine : TFA.

Figure 4.6 : Synthèse du peptide **PJ207**

Le peptide brut de synthèse obtenu après coupure est purifié par CHLP semi préparative en phase inverse. La caractérisation du produit fini se fait par CHLP analytique et par spectrométrie de masse.

II.2.2 Synthèse de la molécule hameçon **54**

L'analogue **54** est obtenu par déprotection en milieu acide de l'azoture **16b** décrit dans le chapitre 2 (figure 4.7).



Conditions réactionnelles : a) HCl (1,25M dans le dioxane), 2h à température ambiante, rendement quantitatif.

Figure 4.7: Synthèse de la molécule **54**.

II.2.3 Essais de co-cristallisation du mélange réactionnel

Avant de faire des essais de co-cristallisation du mélange réactionnel, c'est-à-dire entre la molécule hameçon **54**, le peptide **PJ207** et l'enzyme mCARM1₁₃₀₋₄₈₇, la première étape a consisté à co-cristalliser l'analogue **54** dans l'enzyme CARM1.

II.2.3.1 Essai de co-cristallisation de la molécule hameçon **54** dans mCARM1₁₃₀₋₄₈₇

La molécule **54** a été co-cristallisée dans mCARM1₁₃₀₋₄₉₀ et les cristaux obtenus ont permis d'obtenir une structure cristallographique avec une résolution de 2,4 Å. Ce résultat nous a permis de continuer notre stratégie en réalisant la réaction *in-situ* dans CARM1 entre l'analogue **54** et le peptide **PJ207**.

II.2.3.2 Essais de co-cristallisation du mélange réactionnel dans mCARM1₁₃₀₋₄₉₀

Les conditions utilisées dans la littérature pour réaliser la réaction de click *in-situ* sont les suivantes (tableau 4.2) :

Entrées	Alcyne	Azoture	Enzymes	Temps	T°C	Tampons	Taux de conversion
1	66µM	30µM	AChE (1µM)	90 min	22°C	Citrate d'ammonium (2mM), pH = 7-7,3	80%
2	500µM	100µM	Protéase VIH-1 : 15µM	24 h	23°C	0,1M (MES) et NaCl (0,2M)	0,3%

Tableau 4.2 : Conditions réactionnelles testées pour réaliser la réaction *in-situ*^(3,7).

Les conditions que nous avons testées pour réaliser les essais de réaction de click sont résumées dans le tableau 4.3. Nous avons utilisé des concentrations élevées (de l'ordre du millimolaire) en molécule **54** et en peptide **PJ207** car nous voulions occuper les huit sites de fixation de l'enzyme CARM1. En effet, cette enzyme se présente toujours sous la forme d'un dimère et chaque cristal contient quatre unités de ce dimère. En utilisant une concentration égale à 2 mM en cofacteur (SAH ou SFG), seuls quatre sites sont occupés ; ces molécules sont pourtant affines avec CARM1.

La procédure utilisée a été décrite dans le paragraphe II de ce chapitre. Tout d'abord, un essai a été réalisé uniquement avec le mélange réactionnel. Des cristaux ont été obtenus et la structure cristallographique a uniquement révélé la présence de la molécule **54**, c'est pourquoi un nouvel essai a été effectué en ajoutant un catalyseur chimique, CuSO₄.5H₂O. En général, ces conditions sont utilisées pour des réactions de cycloaddition de Huisgen dans les cellules ou dans des bactéries comme *Escherichia coli*⁽¹¹⁾. Dans ces systèmes, la quantité de

cuivre utilisée varie de 500 μM à 1 mM. Nous ne pouvions pas utiliser de telles quantités car, à partir d'une concentration de 12 μM en sulfate de cuivre, nous observons déjà la précipitation de la protéine ce qui nous a empêché d'obtenir des cristaux (tableau 4.3).

Entrées	54	PJ207	$\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$	Ascorbate de sodium	CARM1
1	2 mM	2 mM	0	0	40 μM
2	2 mM	2 mM	0,012 mM	0,010 mM	40 μM

Tableau 4.3 : Essais de co-cristallisation du mélange réactionnel.

D'après les résultats obtenus, aucune conclusion n'a pu être tirée sur la validité de l'approche. Il nous a fallu trouver une autre technique pour mettre en évidence l'apparition d'éventuels adduits, c'est pourquoi nous avons analysé cette réaction *in-situ* par spectrométrie de masse utilisés pour les assemblages non-covalents. L'objectif de ces expériences est de déterminer si la réaction entre la molécule 54 et le peptide PJ207 peut avoir lieu en présence de mCARM1₁₃₀₋₅₀₇.

III. Analyse par spectrométrie de masse du mélange réactionnel

III.1 Principe de l'analyse des complexes non-covalents par spectrométrie de masse

La spectrométrie de masse des complexes non-covalents est une technique qui permet d'analyser, en phase gazeuse, des complexes non-covalents de macromolécules biologiques⁽¹²⁾ ou de macromolécules et de petites molécules. Ces complexes sont maintenus par interactions faibles de type électrostatique, liaisons hydrogènes, interactions de type Van der Waals et interactions hydrophobes lors du processus d'ionisation par électronébulisation⁽¹³⁻¹⁴⁾. Pour mettre en évidence l'existence d'un complexe non-covalent par spectrométrie de masse, deux

analyses doivent être réalisées. La première consiste à analyser le complexe en conditions « dénaturantes », c'est-à-dire en ajoutant un mélange de solvant 50/50 d'eau et d'acétonitrile en présence de 1% d'acide formique. Dans de telles conditions, les interactions non-covalentes sont rompues et la protéine est dénaturée. Les masses moléculaires des espèces individuelles formant le complexe sont alors mesurées. La deuxième analyse se fait dans les conditions « natives » c'est-à-dire en solution tampon, à pH et force ionique contrôlés permettant à la fois le maintien de l'intégrité structurale en solution des complexes biologiques étudiés et l'analyse par spectrométrie de masse utilisant comme source l'électronébulisation. Si l'on observe une différence entre la spectrométrie de masse mesurée en conditions dénaturantes et en conditions natives, nous pourrions conclure en l'existence d'une interaction non-covalente et déduire directement la masse du composé en interaction.

III.2 Analyse par spectrométrie de masse de la réaction in-situ et du « mélange réactionnel » en conditions dénaturantes ou natives

III.2.1 Analyse du mélange réactionnel en conditions dénaturantes

III.2.1.1 Mesure de mCARM1₁₃₀₋₅₀₇

L'analyse en condition dénaturante de CARM1 utilise une concentration en monomère égale à 2 μ M qui est dilué dans un mélange eau/acétonitrile/acide formique (50/50/1). La spectrométrie de masse de la protéine mCARM1₁₃₀₋₅₀₇ observée en condition dénaturante est de $43066,3 \pm 0,6$ Da ce qui est en accord avec la spectrométrie de masse de la protéine attendue (43065,9 Da) (figure **4.8**).

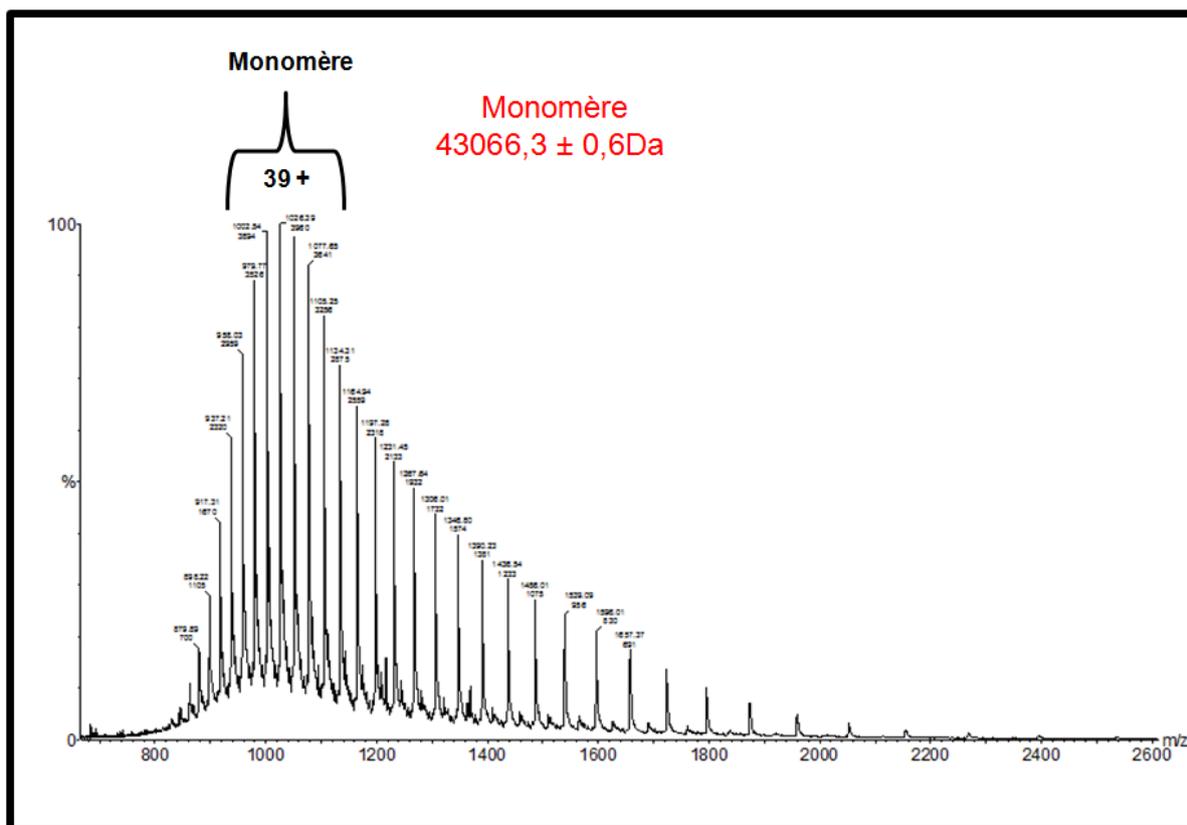


Figure 4.8 : Analyse de mCARM1₁₃₀₋₅₀₇ en conditions dénaturantes.

III.2.1.2 Analyse par spectrométrie de masse du mélange réactionnel en conditions dénaturantes

Lors de cette analyse, nous nous attendions à détecter la masse moléculaire du produit de réaction qui correspond à la somme des masses moléculaires de l'analogue **54** (450 Da) et du peptide **PJ207** (1107 Da) ; l'adduit présenterait donc une masse égale à 1557,8 Da.

Sur le spectre de masse (figure **4.9**), nous observons bien la présence de la protéine CARM1 dans le bruit de fond, c'est à dire dans la région correspondante aux masses comprises entre 800 Da et 1600 Da. La masse prédominante est celle du peptide **PJ207** qui correspond au pic à 1108 Da (figure **4.9.a**) En revanche, nous n'observons aucun pic correspondant à la masse moléculaire du produit attendu qui est égale à 1557,8 Da (figure **4.9.b**).

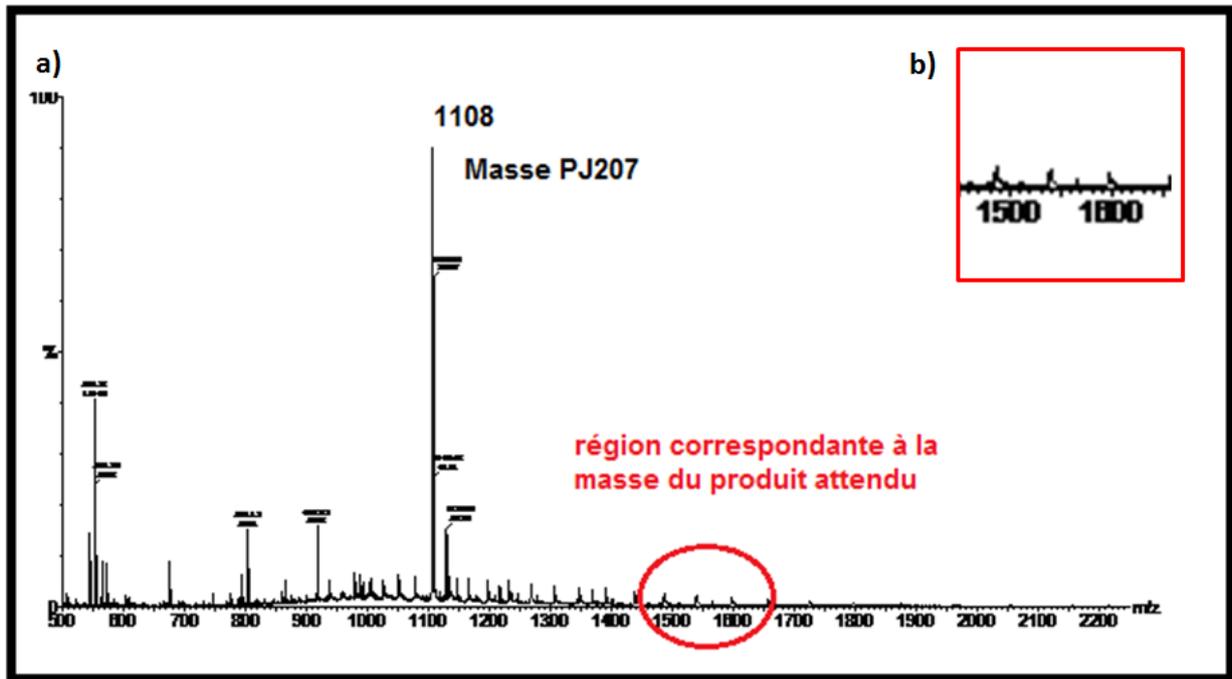


Figure 4.9 : Analyse en conditions dénaturante du mélange réactionnel. a) Spectre correspondant à l'analyse du mélange réactionnel. b) Agrandissement de la région qui correspond à la masse du produit attendu.

Nous avons alors décidé de tenter cette réaction en conditions natives qui nous permettrait d'observer plus facilement le complexe CARM1/adduit.

III.2.2 *Analyse du mélange réactionnel en conditions natives*

III.2.2.1 *Analyse de l'enzyme mCARM1₁₃₀₋₅₀₇ en conditions natives*

L'analyse en conditions natives consiste à utiliser une concentration en monomère égale à 20 μ M dans une solution tampon d'acétate d'ammonium à 50 mM, ajustée à pH = 7. Cette analyse nous a permis de déterminer que la protéine mCARM1₁₃₀₋₅₀₇ forme un tétramère non-covalent de 173259 Da et un octamère de 354564 Da à un pH égal à 7 (figure 4.10).

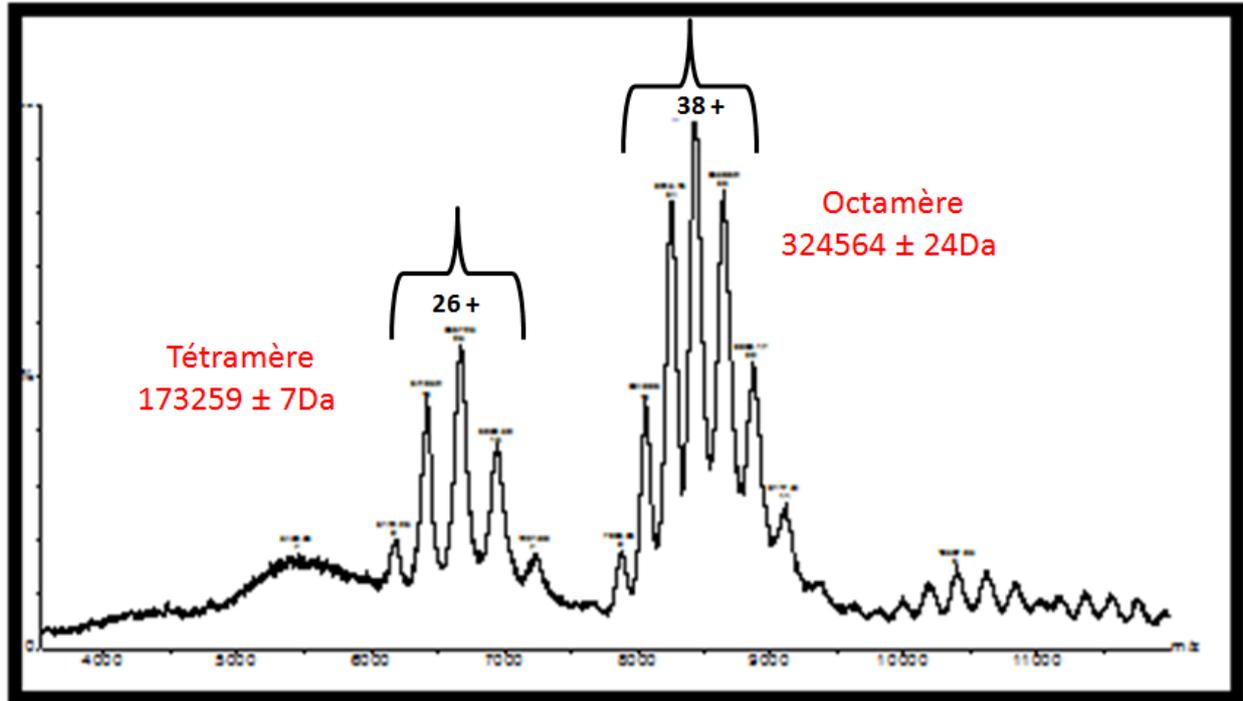


Figure 4.10 : Analyse en conditions natives de mCARM1₁₃₀₋₅₀₇ qui met en évidence la forme tétramérique et octamérique de la protéine à pH = 7.

III.2.2.2 *Analyse par spectrométrie de masse du mélange réactionnel en conditions natives*

L'analyse de la protéine en conditions natives met en évidence la présence d'un pic correspondant au peptide **PJ207**. Néanmoins, la région du spectre qui correspond aux masses moléculaires de la protéine CARM1 ne montre qu'une faible ondulation : aucune trace des deux pics massifs correspondant aux formes tétramériques et octamériques de CARM1 n'a pu être observée. Cette analyse ne nous permet donc pas d'étudier la fixation de notre analogue **54** au sein de la protéine CARM1 car un trop large excès de ce composé et du peptide **PJ207** est nécessaire ; il n'est donc plus possible de visualiser l'enzyme et les adduits qui devraient se former (figure **4.11**).

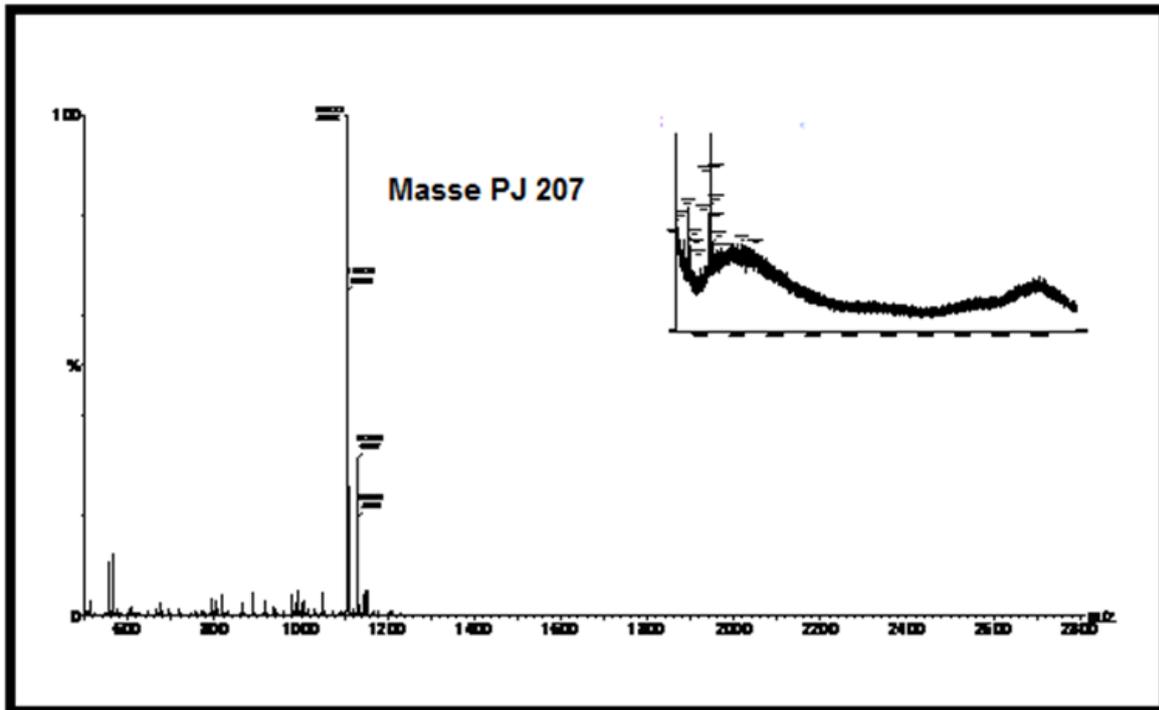


Figure 4.11 : Analyse du mélange réactionnel en conditions natives. Un agrandissement du spectre entre 2000 et 12000 m/z est représenté en haut à droite.

Ces premiers résultats ne nous ont donc pas permis de mettre en évidence la formation des adduits attendus. Cependant, nous ne pouvons pas conclure quant à la formation du produit de cycloaddition car le taux de conversion de cette réaction varie suivant les concentrations utilisées en synthons. Plus ces concentrations sont faibles, plus le taux de conversion est élevé. Lorsque les concentrations en synthons sont sept fois plus élevées que celles utilisées pour l'AChE, le taux de conversion chute de 80 % (tableau 4.4). Dans notre cas, les concentrations utilisées en synthons sont quatorze fois plus élevées que celle utilisées pour l'AChE, le taux de conversion devrait être proche de zéro. Dans ce cas, l'appareil de spectrométrie utilisé ne nous permettrait pas de mettre en évidence cette très faible quantité de produit ; en effet, la concentration seuil de détection de cet appareil est de l'ordre du picomolaire.

Entrées	Alcyne	Azoture	Enzymes	Temps	T°C	Tampons	Taux de conversion
1	66µM	30µM	AChE (1µM)	90 min	22°C	Citrate d'ammonium (2mM) pH = 7,0-7,3	80%
2	500µM	100µM	Protéase VIH-1 : 15µM	24 h	23°C	MES (0,1M) et NaCl (0,2M) pH = 5,3	0,3%
4	1mM	1mM	CARM1 : 83µM	1 h	37°C	Tris HCl, (100 mM) pH = 8,0	-

Tableau 4.4 : Conditions utilisées pour réaliser les essais de click *in-situ* décrites dans la littérature et pour notre mélange réactionnel analysé par spectrométrie de masse ^(3,7).

L'approche par réaction de click *in-situ* n'a pas abouti en utilisant la méthode de co-cristallisation : sans l'utilisation de cuivre, aucun adduit n'a pu être détecté. De plus, en présence du catalyseur $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, même introduit à de très faibles concentrations (de l'ordre de 12 µM), la protéine a précipité.

Afin de mettre en évidence l'apparition d'éventuels adduits, une nouvelle technique à été utilisée, la spectrométrie de masse des complexes non covalents. Cependant, aucune conclusion n'a pu être tirée suite à ces expériences.

L'approche par réaction de click n'est donc pas assez réactive. Il est donc nécessaire de trouver une alternative : après analyse de la littérature, une nouvelle approche a été définie qui consiste à faire réagir spontanément une molécule hameçon électrophile avec un peptide cystéine. Dans cette nouvelle approche, seule la technique par co-cristallisation sera utilisée,

car les expériences de spectrométrie de masse sont difficiles à mettre en œuvre, coûteuses et les conditions restent à être optimisées.

IV. Réaction *in-situ* entre une molécule hameçon électrophile et un peptide cystéine par réaction de substitution nucléophile

Dans cette stratégie, une molécule hameçon électrophile est capable de capturer un peptide mimant la séquence de l'histone H3 ; ce peptide possède une cystéine à la place de l'arginine 17 et est appelé peptide cystéine. Le mélange réactionnel sera ensuite testé en co-cristallisation et analysé par diffraction aux rayons X suivant le même protocole décrit dans le paragraphe II de ce chapitre (figure 4.12).

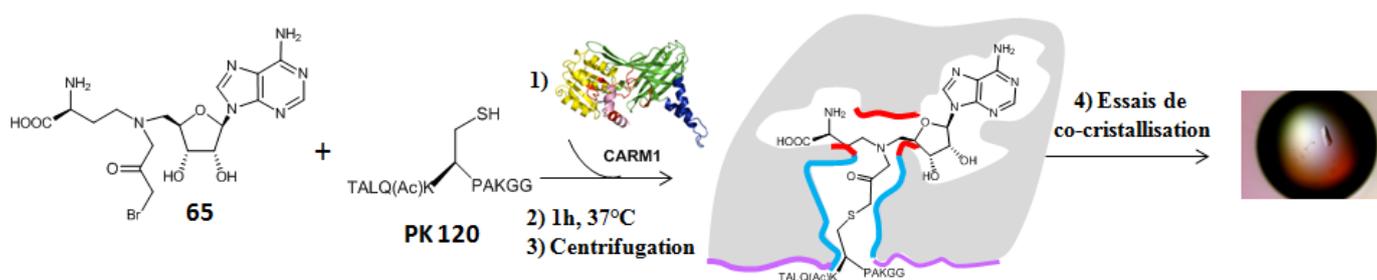


Figure 4.12 : Stratégie pour capturer un peptide cystéine par réaction de substitution nucléophile entre des molécules hameçons électrophiles et un peptide cystéine. 1) Introduction de l'enzyme CARM1 au mélange molécule hameçon/peptide. 2) Incubation du mélange réactionnel à 37°C pendant une heure. 3) Centrifugation. 4) Essais de co-cristallisation et analyse au synchrotron.

IV.1 Molécules hameçons électrophiles pour la capture de peptides cystéine

La littérature nous a servi de base pour concevoir les molécules hameçons électrophiles capables de réagir avec une cystéine modifiée, mais nous nous sommes également inspirés de la réactivité chimique et de l'environnement du site actif pour obtenir des molécules hameçons originales. Elles sont décrites sur la figure 4.13 et elles possèdent comme fonction chimique :

- un groupement époxyde (63)

- un groupement acrylate (**64**)
- un groupement bromocétone (**65**)
- un groupement céto-ester (**66**)
- un groupement iodé (**68**)
- un groupement thiocyanate (**70**)

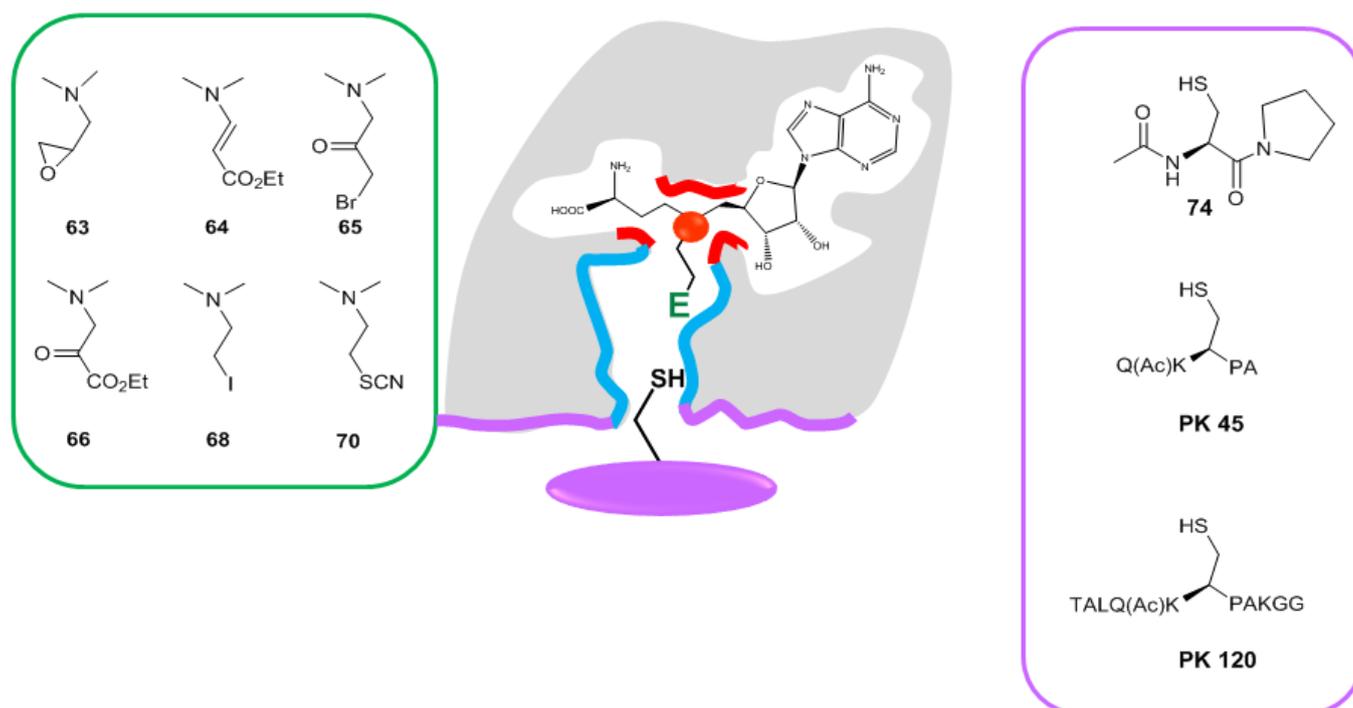


Figure 4.13 : Molécules hameçons électrophiles et peptides cystéines utilisés pour réaliser la réaction *in-situ*.

Les groupements des molécules hameçons électrophiles **64** et **65** ont été choisis car ils sont connus pour être réactifs envers les cystéines modifiées⁽¹⁵⁻¹⁶⁾. Les peptides cystéines utilisés sont le peptide **PK 45**, qui mime la séquence 15-19 de l'histone H3 et **PK 120**, qui mime la séquence 12-22 de l'histone H3. Une cystéine modifiée seule à également été utilisée.

IV.1.1 Synthèses des molécules hameçons 63 à 70

IV.1.1.1 Synthèses des molécules 63 à 66

Les molécules hameçons choisies ont été synthétisées en deux étapes à partir de l'amine **10a**. Les analogues **63** à **66** sont obtenus par une réaction d'alkylation sur l'amine **10a** suivie d'une amination réductrice. Le dérivé est ensuite déprotégé en milieu acide (figure **4.14**).

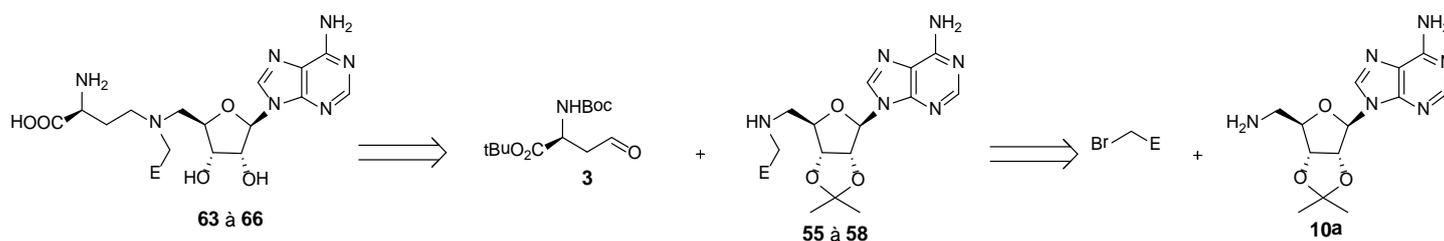


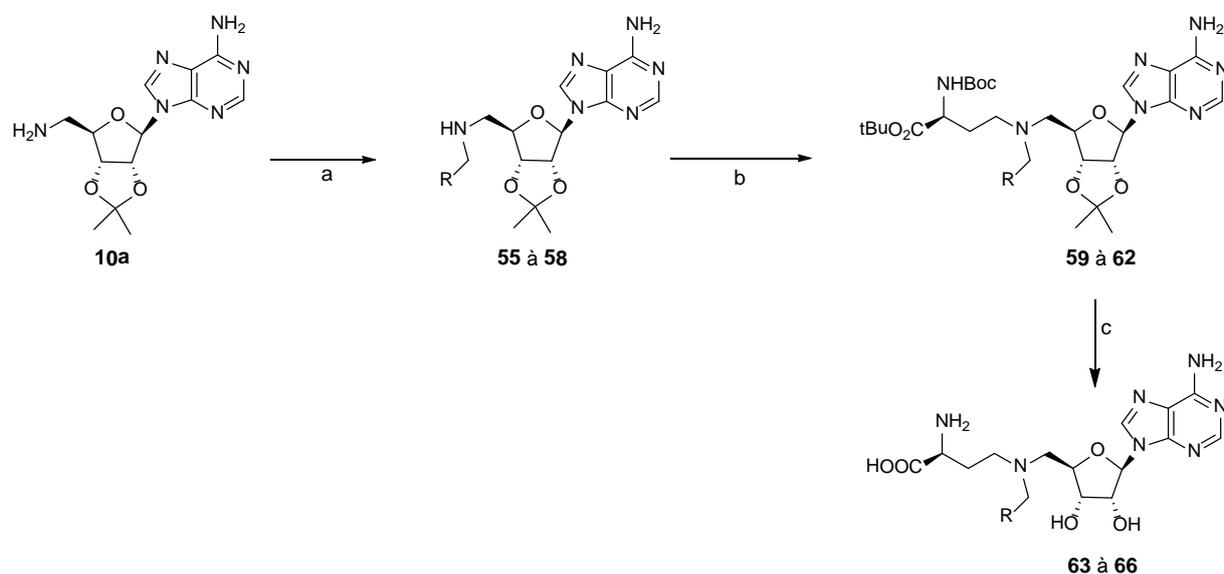
Figure 4.14 : Schéma rétrosynthétique pour l'obtention des molécules hameçon **63** à **66**. Ces analogues sont obtenus à partir d'une alkylation puis d'une amination réductrice suivie d'une déprotection en milieu acide.

Les rendements de ces deux étapes, suivant l'électrophile introduit, est décrit dans le tableau ci-dessous (tableau **4.5**).

Entrées	Electrophiles	Dérivés : Rendements (%)	Dérivés : Rendements (%)
1		55 : 30	59 : 27
2		56 : 50	60 : 61
3		57 : 50	61 : 38
4		58 : 29	62 : 38

Tableau 4.5: Dérivés bromés utilisés pour la synthèse des électrophiles. Les rendements des étapes d'alkylation et d'amination réductrice sont indiqués pour chaque électrophile utilisé.

Finalement, ces dérivés **59** à **62** sont déprotégés en milieu acide pour conduire à une première série de molécules hameçons **63** à **66** (figure **4.15**).

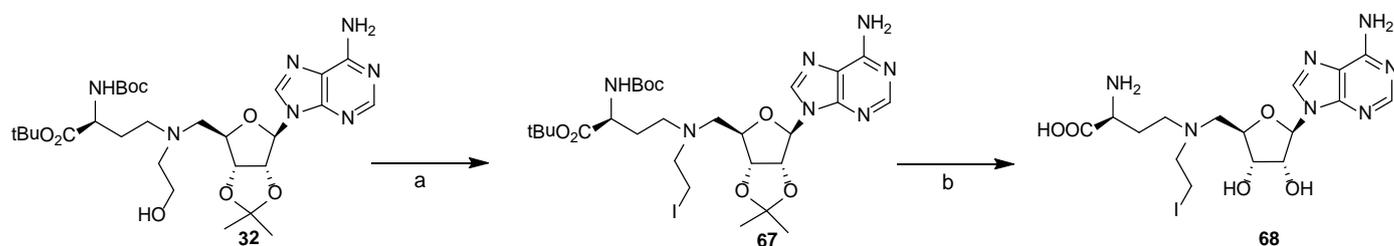


Conditions réactionnelles : a) RBr, DIEA, THF, 18h, 50°C . b) **3**, HOAc, NaBH₃CN, MeOH, 12h, température ambiante. c) HCl (1,25M dans le dioxane), 2h, température ambiante, rendement quantitatif.

Figure 4.15 : Synthèses des molécules hameçons **63** à **66**.

IV.1.1.2 Synthèse de la molécule 68

La molécule **68** a été synthétisée à partir du dérivé **32** dans lequel la fonction alcool est substituée par de l'iode en présence de triphénylphosphine pour former le dérivé **67** avec 34% de rendement. Puis, ce dernier est déprotégé quantitativement en milieu acide pour obtenir une nouvelle molécule hameçon **68** (figure **4.16**).

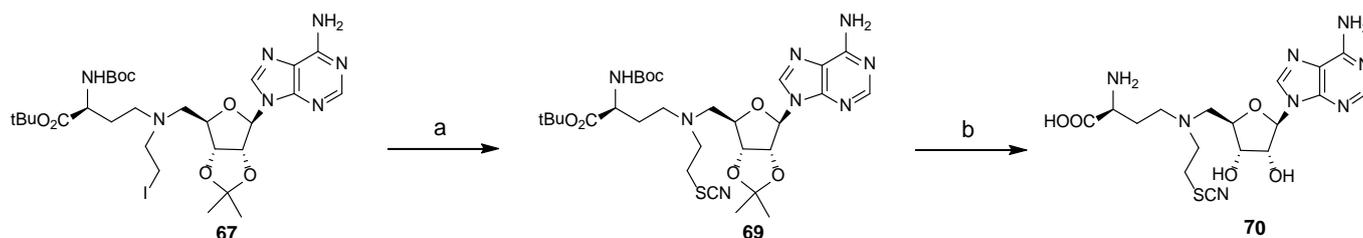


Conditions réactionnelles : a) I₂, PPh₃, imidazole, DCM, 12h, 0°C puis température ambiante, 34%. b) HCl (1,25M dans le dioxane), 2h, température ambiante, rendement quantitatif.

Figure 4.16 : Synthèse de la molécule **68**.

IV.1.1.3 Synthèse de la molécule thiocyanate **70**

Le dérivé **69** est synthétisé par substitution nucléophile de l'iode du dérivé **67** par le groupement thiocyanate, en conditions douces, en utilisant les réactifs TMSNCS et TBAF⁽¹⁶⁾. Après déprotection en milieu acide, l'analogue **70** est obtenu avec un rendement de 28% sur les deux étapes (figure 4.17).



Conditions réactionnelles : a) TMSNCS, TBAF (0,1M dans le THF), 12h, 0°C puis température ambiante, 77% .
b) HCl (1,25M dans le dioxane), 2h, température ambiante, 37%.

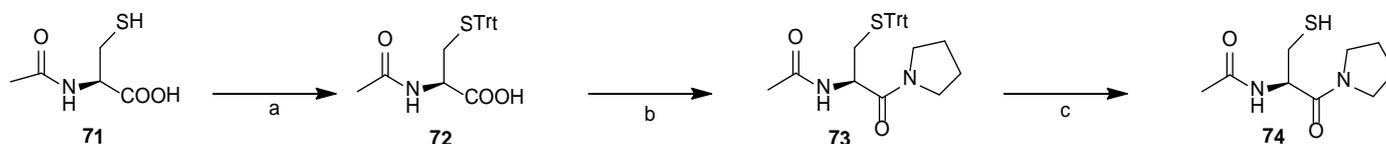
Figure 4.17 : Synthèse de la molécule thiocyanate **70**.

Les molécules hameçons étant prêtes à être testées, les peptides correspondants doivent à présent être synthétisés. Les peptides **PK45** et **PK120** ont été synthétisés suivant le même protocole que le **PJ207**. La synthèse de la cystéine modifiée **74** est décrite ci-dessous (figure 4.18).

IV.1.1.4 Synthèse de la cystéine modifiée **74**

La cystéine modifiée **74** est synthétisée à partir de *N*-acétyl-L-cystéine modifiée commerciale **71** en trois étapes (figure 4.18). La première étape consiste à protéger le thiol par un groupement trityle⁽¹⁷⁾ pour éviter la formation d'un pont disulfure. La cystéine modifiée **72** obtenue subit un couplage peptidique avec la pyrrolidine en présence d'EDC et de DIEA dans le dichlorométhane pendant 12 heures ; le dérivé **73** est obtenu avec un rendement de 42% et sa fonction thiol est finalement déprotégée en présence de triisopropylsilane et de TFA

pendant 12 heures à température ambiante pour conduire à la cystéine modifiée désirée **74** quantitativement.



Conditions réactionnelles : a) HOAc, Tri-OH, BF₃.Et₂O, 18h, température ambiante, 83%. b) EDC.HCl, DIEA, Pyrrolidine, DCM, 16h, température ambiante, 42%. c) Triisopropylsilane, TFA, DCM, 12h, 0°C puis température ambiante, rendement quantitatif.

Figure 4.18 : Synthèse de la cystéine modifiée **74**

IV.2 Essais de réactions *in-situ* entre les molécules hameçons électrophiles et les peptides cystéines puis essais de co-cristallisation dans CARM1

Dans un premier temps, les molécules hameçons ont été caractérisées biochimiquement en mesurant leurs valeurs d'inhibition (CI₅₀). Ces valeurs nous permettent d'avoir une indication la capacité des molécules à se lier à l'enzyme.

IV.2.1 Mesures d'inhibition des molécules hameçons

Les mesures d'inhibition ont été réalisées en respectant le protocole décrit dans le chapitre précédent. La molécule de référence est l'analogue de bisubstrat **45** qui pour le moment est notre meilleur inhibiteur. L'analogue époxyde **63** possède une valeur d'inhibition de 114,8 μM. Lorsque la fonction époxyde est remplacée par un accepteur de Mickael (analogue **64**) ou une fonction bromocétone (analogue **65**), la valeur d'inhibition n'est pas meilleure. La fonction céto-ester (analogue **66**) a une valeur d'inhibition moins bonne avec une CI₅₀ égale à 309,7μM. Finalement, l'analogue iodé (analogue **68**) a un CI₅₀ de 153,0 μM et la substitution de l'iode par la fonction thiocyanate (analogue **70**) conduit à une inhibition de 212,4 μM (tableau **4.6**).

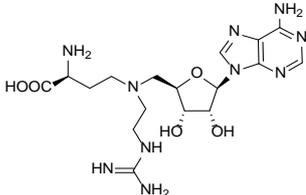
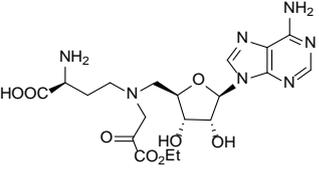
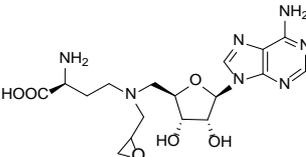
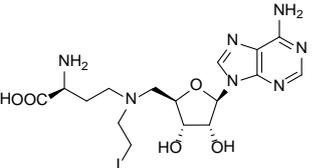
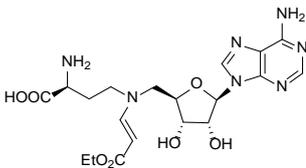
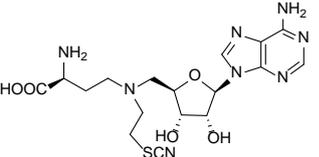
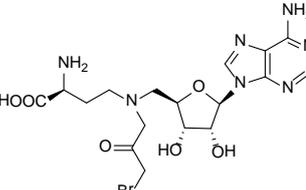
Molécules hameçons	Structures	CI ₅₀ [μM] CARM1	Molécules hameçons	Structures	CI ₅₀ [μM] CARM1
45		2,1	66		309,7
63		114,8	68		153,0
64		192,7	70		212,4
65		131,7			

Tableau 4.6 : Valeurs de CI₅₀ des molécules hameçons contre CARM1. Le substrat utilisé est le tétramère d'histone (H3H4)₂ et le ¹⁴C SAM est le donneur de méthyle radioactif.

Bien que ces molécules hameçons présentent des valeurs de CI₅₀ moyennes qui sont comprises entre 100 μM et 300 μM, chacun de ces analogues va être testé en co-cristallisation dans CARM1.

IV.2.2 Essais de co-cristallisation du mélange réactionnel

Des essais de réaction *in-situ* puis de co-cristallisation ont été entrepris pour les molécules hameçons **63** à **66**, **68** et **70** dans mCARM1₁₃₀₋₄₈₇ avec les cystéines modifiées **74**,

PK45 et PK 120. Des cristaux ont été uniquement obtenus pour les molécules **65** et **70**. Dans ces cristaux, seule la densité électronique des molécules hameçons ont été visualisées et aucune trace de peptide ou d'adduit n'a été observée. La molécule **65** a été visualisée dans CARM1 avec une résolution de 2,6 Å et une image de la molécule **70** a aussi été obtenue avec une résolution de 2,5 Å. Ces structures cristallographiques nous ont apporté deux informations :

- La première nous indique que nous n'avons pas identifié la formation d'un éventuel adduit sur la carte de densité électronique.
- La seconde nous indique que, malgré les valeurs moyennes de CI_{50} de ces deux molécules (131,7 μ M pour la molécule **65** et 212,4 μ M pour la molécule **70**), il est possible d'obtenir une structure cristallographique.

Par conséquent, pour réaliser la réaction *in-situ*, le choix du peptide est primordial. Jusqu'à présent, le peptide **PK110** est le seul à avoir montré une fixation dans l'enzyme par analyse de spectrométrie de masse, c'est pourquoi nous avons décidé de réaliser les réactions *in-situ* avec ce dernier malgré la difficulté à modifier l'arginine. Notre troisième approche consiste donc à piéger l'arginine d'un peptide mimant la séquence de l'histone H3 par des molécules hameçons adéquates.

V. Réaction *in-situ* entre une molécule hameçon électrophile et un peptide arginine par réaction de condensations

Dans cette dernière stratégie, des molécules hameçons électrophiles seront capables de piéger, par réaction de condensation, des peptides qui présentent une arginine 17. Dans un premier temps, les molécules hameçons seront caractérisées biochimiquement, puis des essais de co-cristallisation entre ces analogues et ces peptides seront réalisés en présence de

l'enzyme CARM1. Finalement, les cristaux obtenus seront analysés au synchrotron (figure 4.19).

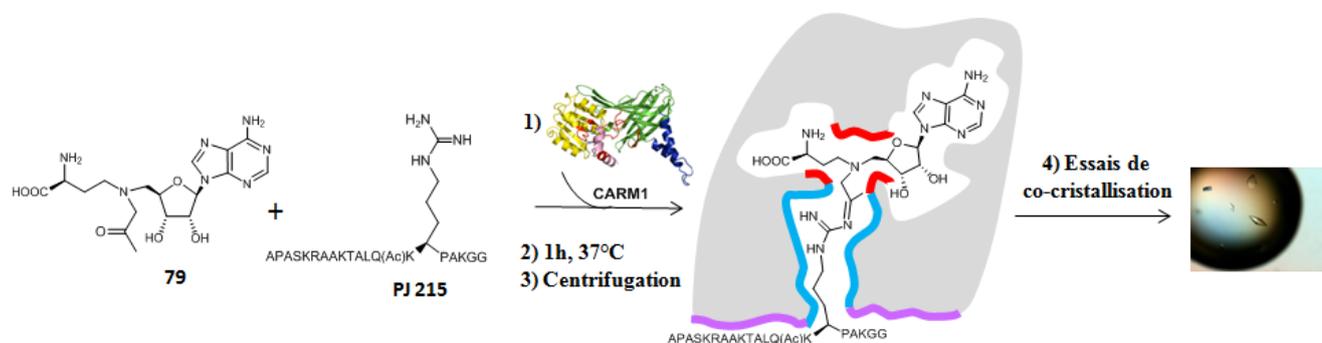


Figure 4.19 : Stratégie pour capturer le peptide par réaction de condensation avec des molécules hameçons électrophiles. 1) Introduction de l'enzyme CARM1 au mélange molécule hameçon/peptide. 2) Incubation du mélange réactionnel à 37°C pendant une heure. 3) Centrifugation. 4) Essais de co-cristallisation puis analyse au synchrotron.

En s'inspirant de la littérature et de la réactivité chimique, des molécules hameçons capables de réagir avec l'arginine de l'histone H3 ont été conçues.

V.1. Molécules hameçons électrophiles pour la capture de peptides arginine

En tenant compte des contraintes de synthèse, du site actif de CARM1 et en s'inspirant de la réactivité chimique, quatre molécules hameçons ont été synthétisées:

- Un analogue cyano (**77**)
- Un analogue cétone (**79**)
- Un analogue dicarbonylé (**84**)
- Un analogue aziridine (**85**)

Les peptides utilisés sont le peptide **PJ 215** qui mime la séquence 12-29 de l'histone H3, le **PI 84** qui mime la séquence 3-25 de l'histone H3 et le peptide **PK110** qui mime la

séquence 1-34 de l'histone H3. Une arginine et une arginine non-méthylée ont également été utilisées (figure 4.20).

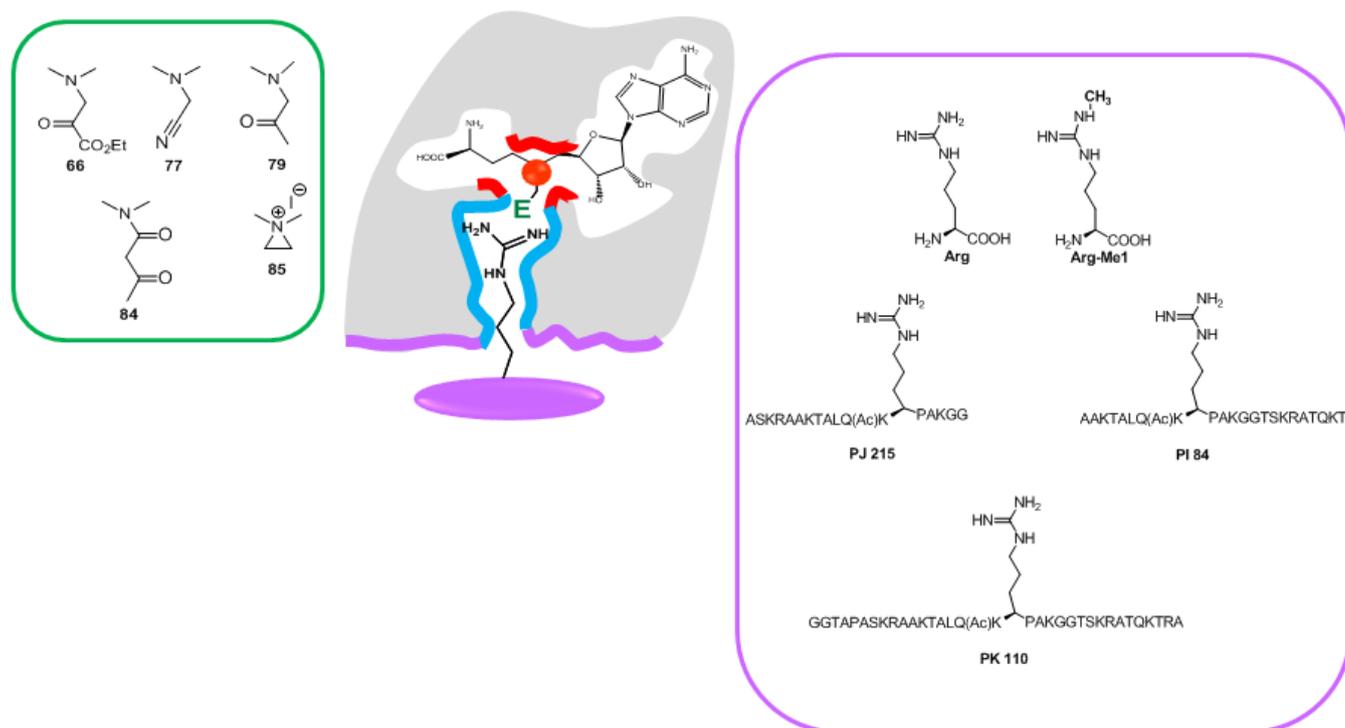


Figure 4.20 : Molécules hameçons électrophiles capables de piéger un peptide arginine.

Les groupements des analogues **66**, **68**, **79**, **84** et **85** ont été choisis car ils sont connus pour être réactifs envers l'arginine^(15, 19-22).

Les molécules hameçons choisies ont été synthétisées en trois étapes à partir de l'amine **10a**, suivant deux voies différentes. La voie A a permis d'obtenir l'analogue **77** par une réaction d'alkylation suivit d'une amination réductrice puis d'une déprotection. Alors que les analogues **79** et **84** ont été obtenus par deux animations réductrices successives en suivant la voie B puis d'une déprotection (figure 4.21).

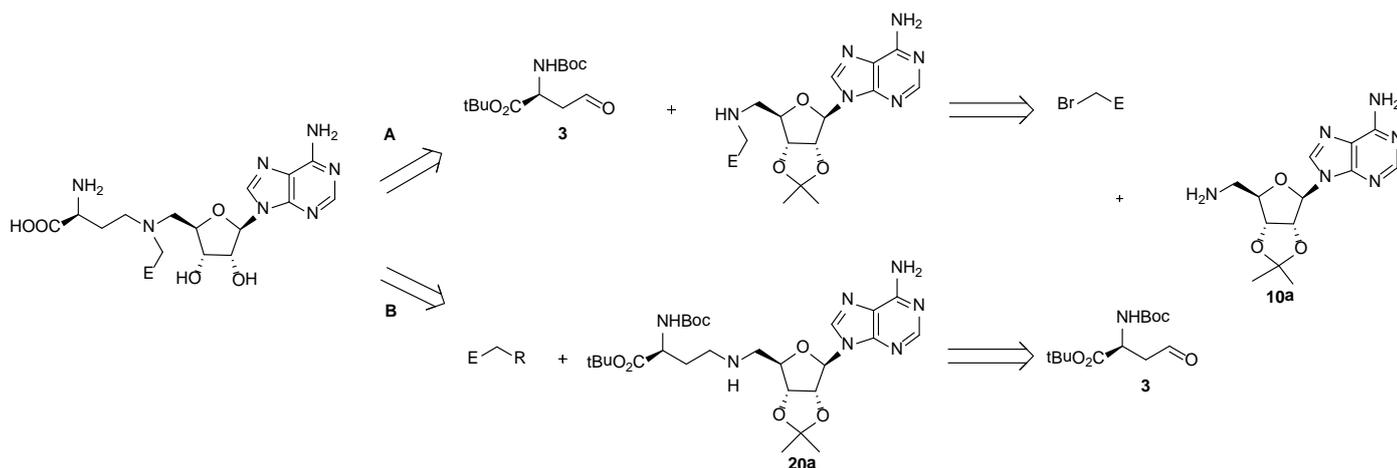
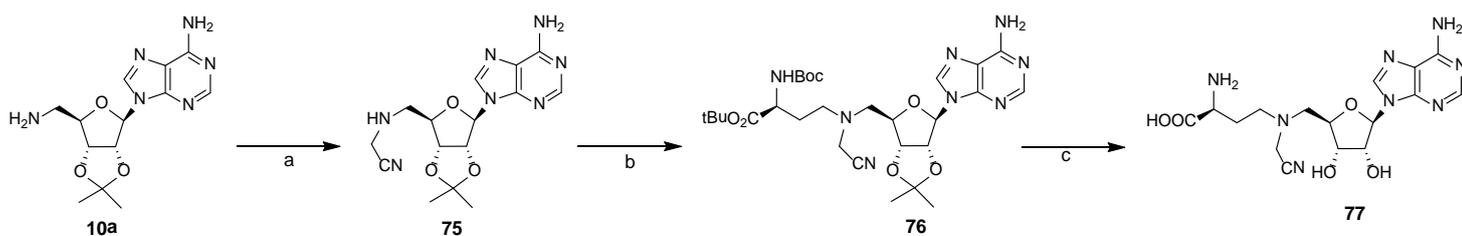


Figure 4.21 : Schéma rétrosynthétique pour l'obtention des molécules hameçons **77**, **79** et **84**.

V.1.1.1 Synthèse de la molécule **77**

La molécule **77** a été synthétisée à par la voie A. Le dérivé **75** est obtenu par alkylation entre l'amine **10a** et le 2-bromoacétonitrile avec un rendement de 66%. Puis, une amination réductrice avec l'aldéhyde **3** et une déprotection en milieu acide conduit à l'analogue désiré **77** (figure 4.22).

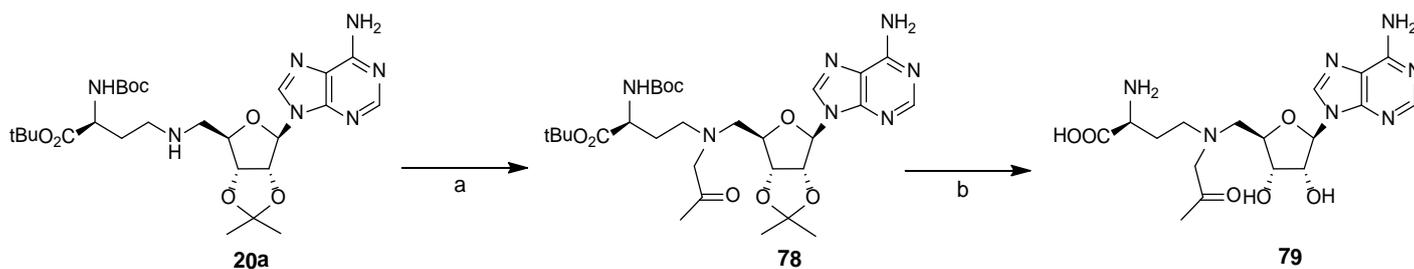


Conditions réactionnelles : a) 2-bromoacétonitrile, DIEA, THF, 12h, 50°C, 66% . b) RCHO, HOAc, NaBH₃CN, MeOH, 12h à température ambiante 39 % . c) HCl /dioxane : 1,25M, 2h à température ambiante, quantitatif.

Figure 4.22 : Synthèse de la molécule **77**.

V.1.1.2 Synthèse de la molécule 79

La molécule **79** a été synthétisée en deux étapes suivant la voie B décrite dans la figure 4.27. Le composé **20a** subit ainsi une première amination réductrice en présence de méthylglyoxal (MGO), puis le composé **78** obtenu est déprotégé en milieu acide pour former le composé **79** avec un rendement de 34% sur ces deux étapes (figure 4.23).



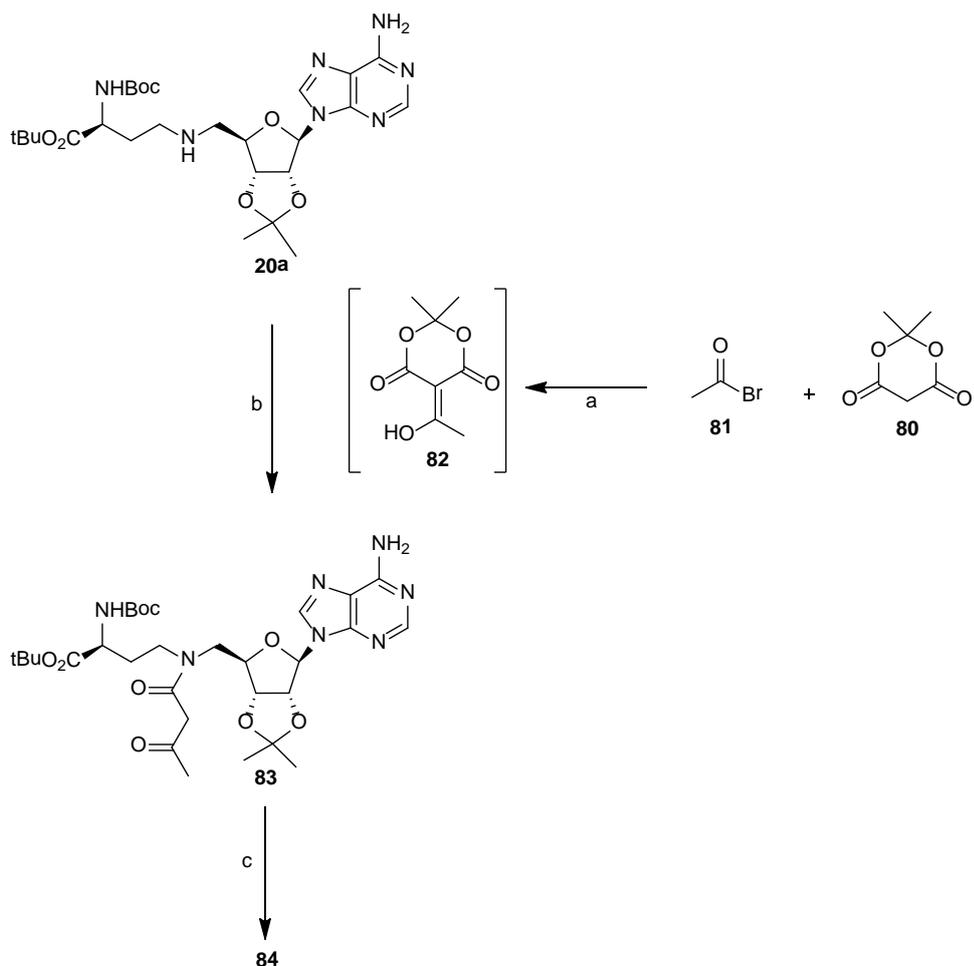
Conditions réactionnelles : a) MGO, HOAc, NaBH₃CN, MeOH, 12h, température ambiante, 48%. b) HCl (1,25M dans le dioxane), 2h, température ambiante, 70%.

Figure 4.23 : Synthèse de la molécule **79**.

V.1.1.3 Synthèse de la molécule 84

L'acide de Meldrum⁽²³⁾ **80** réagit avec le bromure d'acétyle **81** en présence de pyridine et de 4-diméthylaminopyridine pour conduire à l'intermédiaire **82**. Ce dernier est mis en réaction avec l'amine **20a** dans un mélange dichlorométhane / toluène et agité à température ambiante pendant 20 heures pour former le dérivé **83**⁽²⁴⁾.

La dernière étape de déprotection en milieu acide conduit à la formation de la molécule hameçon désirée **84** avec un rendement de 27% sur les trois étapes (figure 4.24).

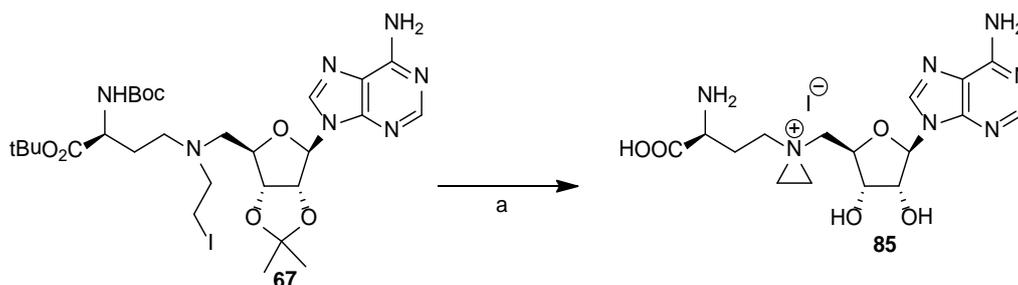


Conditions réactionnelles : a) DCM, pyridine, DMAP, 12h, 0°C puis température ambiante. b) DCM/Toluène (1/4), 20h, température ambiante, 31% sur les trois étapes. c) HCl (1,25M dans le dioxane), 2h, température ambiante, rendement quantitatif.

Figure 4.24: Synthèse de la molécule hameçon **84**.

V.1.1.4 Synthèse de la molécule 85

La molécule **85** est obtenue à partir du dérivé iodé **67** en deux étapes. Tout d'abord, la molécule **67** est déprotégée en milieu acide chlorhydrique concentré puis au bout de deux heures le milieu réactionnel est dilué dans de l'eau et lavé avec du dichlorométhane. La phase aqueuse est évaporée puis une solution de soude à 0,5M⁽²⁵⁾ est ajoutée et le milieu réactionnel est agité à température ambiante pendant 14 heures. La molécule hameçon **85** est ainsi obtenue avec 71 % de rendement (figure 4.25).



Conditions réactionnelles : a) HCl (1,25M dans le dioxane), NaOH 0,5M, 14h, température ambiante, 71%.

Figure 4.25 : Synthèse de l'aziridine **85**.

Les molécules hameçons **77**, **79**, **84** et **85** ont ensuite été testées biologiquement. Dans un premier temps, les valeurs de CI_{50} de ces composés dans l'enzyme CARM1 vont être déterminées puis les analogues **77**, **79**, **84** et **85** ainsi que les analogues **66** et **68**, décrites dans le paragraphe précédent, seront testés en co-cristallisation dans CARM1 avec le peptide natif.

V.2 Essais de réactions *in-situ* entre les molécules hameçons électrophiles et les peptides arginines puis co-cristallisation dans CARM1

V.2.1 Mesures d'inhibition des molécules hameçons

Les tests d'inhibition ont été réalisés suivant la méthode décrite dans le chapitre 3 c'est-à-dire en mettant en compétition le cofacteur, la SAM, et les différentes molécules hameçon dans le site catalytique de CARM1. La référence utilisée pour nos tests d'activité est l'analogue **45**. Les valeurs de CI_{50} pour cette classe de molécules varient fortement puisqu'elles sont comprises entre 13,3 et 863,8 μ M (tableau 4.7). Nous pouvons constater que suivant le groupement électrophile, le potentiel d'inhibition varie fortement dans CARM1. Les valeurs de CI_{50} des molécules carbonylés (**79** et **84**) se situent entre 20,6 μ M et 80,7 μ M. Cette dernière valeur correspond à la cétone **79** tandis que la meilleure valeur correspond au dicarbonylé **84**. Pour les molécules inspirées par la réactivité chimique, les résultats observés sont totalement différents, car la molécule **77** possédant un groupement cyano n'a plus d'effet d'inhibition sur l'enzyme avec une valeur de CI_{50} égale à 863,8 μ M. En revanche, l'aziridine

85 a une valeur d'inhibition de 13,3 μM meilleure que celle obtenue avec la molécule hameçon **84**.

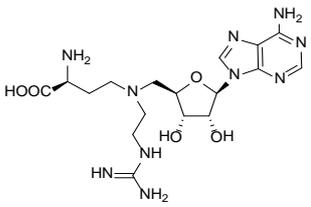
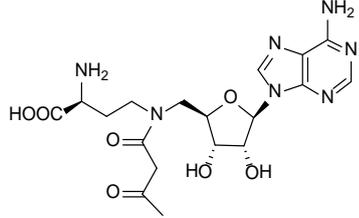
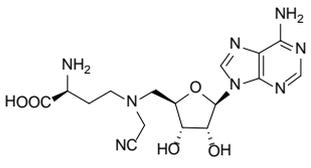
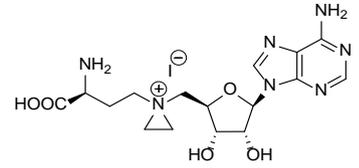
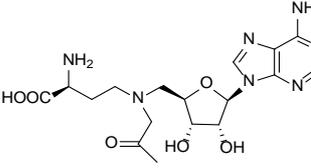
Composés	Structures	CI ₅₀ [μM] CARM1	Composés	Structures	CI ₅₀ [μM] CARM1
45		2,1	84		20,6
77		863,8	85		13,3
79		80,7			

Tableau 4.7 : Valeurs de CI₅₀ des molécules hameçons contre CARM1. Le substrat utilisé est le tétramère d'histone (H3H4)₂ et le ¹⁴C SAM est le donneur de méthyle radioactif.

Le groupement carbonyle semble être un bon choix pour la conception de nos molécules hameçons, mais la position de la fonction cétone sur la molécule doit se faire judicieusement.

Afin d'éviter toutes contraintes stériques et interactions défavorables qui pourraient y avoir entre ces groupements et les acides aminés du site actif de CARM1 (tableau **4.8**).

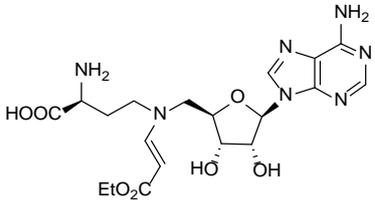
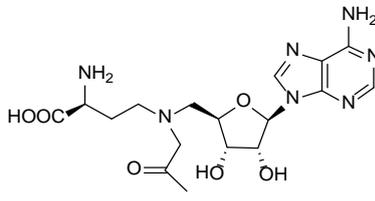
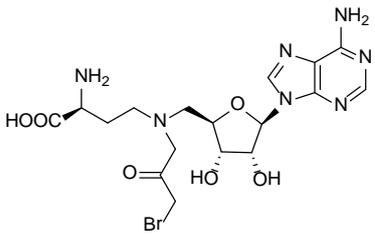
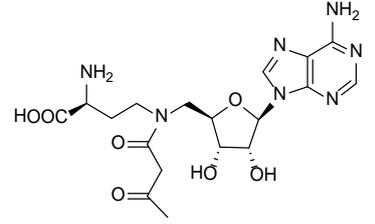
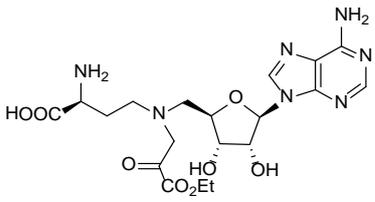
Molécules hameçons	Structures	CI ₅₀ [μM] CARM1	Molécules hameçons	Structures	CI ₅₀ [μM] CARM1
64		192,7	79		80,7
65		131,7	84		20,6
66		309,7			

Tableau 4.8 : Valeurs de CI₅₀ des molécules hameçons contre CARM1. Le substrat utilisé est le tétramère d'histone (H3H4)₂ et le ¹⁴C SAM est le donneur de méthyle radioactif.

Toutes ces molécules hameçons (**66**, **68**, **77**, **79**, **84** et **85**) ont ensuite été engagées dans des essais de co-cristallisation avec CARM1.

V.2.2 Essais de co-cristallisation du mélange réactionnel

Des essais de co-cristallisations ont été entrepris entre les molécules hameçons électrophiles (**66**, **68**, **77**, **79**, **84** et **85**) et les peptides arginine (**PI84**, **PJ215** et **PK110**). Des essais ont également été réalisés en utilisant une arginine seule, pour voir s'il n'y avait pas de gênes stériques créés par le peptide dans le site actif de l'enzyme, et une arginine mono-méthylée pour augmenter la nucléophilie de cette dernière. Les conditions utilisées pour

réaliser les essais de co-cristallisation sont décrites dans le paragraphe II de ce chapitre : elles peuvent être ainsi comparées à celles utilisées dans la littérature. Les réactifs utilisés sont les suivants (tableau **4.9**) :

- La 1,3-butanedione qui est capable de réagir sur une arginine de la transferrine en présence de tampon borate nécessaire pour stabiliser la formation du diol vicinal. La réaction est rapide car, en deux minutes, 76 % du produit est formé⁽²⁰⁾.
- La 1,2-cyclohexadione réagit sur une arginine de la myoglobine (ApoMB) pour former, en deux heures, 20% de diol⁽²¹⁾.
- Le phenylglyoxal réagit avec une arginine de la transferrine pour obtenir 50 % de produit désiré en 5 minutes⁽²²⁾.
- La molécule hameçon **68** réagit avec un peptide arginine de l'histone H4 catalysé par PRMT1 pour former 25 % d'adduit en 30 minutes⁽¹⁹⁾.

Les conditions utilisées en dérivés carbonylés sont de l'ordre du mM et en large excès par rapport à l'enzyme, ces conditions se rapprochent de celles utilisées pour les essais de co-cristallisation.

Entrées	Enzymes	Dérivés dicarboxylés	Substrats	Tampons	pHs	T°C	Temps (min)	Taux de conversion (%)
1	Transferrine (0,52 mM)	1,3-butanedione (20 mM)		Borate (200 mM)	7,8	25	2	76
2	Apo MB (23,5 µM)	1,2-cyclohexadione (930 µM)		NaOH 1N	10	25	120	20
3	Transferrine (65 µM)	Phenylglyoxal (17 mM)		NaHCO ₃ (100 mM)	7,8	25	5	50
4	PRMT1 (387 µM)	Composé 68 (1,67 mM)	Peptide AcH ₄₋₂₁ (50 µM)		8,0	37	30	25
5	CARM1 (40 µM)	Molécules hameçons (2 mM)	Peptide H3 (2 mM)	Tris HCl (100 mM)	8,0	37	60	-

Tableau 4.9 : Conditions réactionnelles pour modifier une arginine au sein d'une protéine ⁽¹⁹⁻²²⁾.

Deux structures cristallographiques ont été obtenues avec les molécules **68** et **85**.

- Dans la structure cristallographique obtenue avec la molécule **68**, le peptide n'a pas été visualisé sur la carte de densité électronique. De plus, la molécule **68** n'est pas visible entièrement car l'atome d'iode n'est pas présent dans la structure.
- Dans la structure cristallographique obtenue avec la molécule **85**, aucun peptide n'a été observé sur la carte de densité électronique. De plus, la résolution obtenue est trop faible (2,6 Å) pour que nous puissions affirmer que

c'est bien la structure de la molécule **85** et non celle de la SAM présente naturellement dans le site actif.

Cette dernière observation montre qu'il faut concevoir une molécule hameçon possédant une grande affinité avec l'enzyme pour :

- Réussir à extraire la molécule de la SAM de CARM1
- Réussir à occuper tous les sites libres de CARM1 avec un ratio 4 : 4 afin de réussir à visualiser nos adduits sur la carte de densité électronique

Jusqu'à présent, les composés les plus efficaces dans CARM1 que nous avons obtenus sont les analogues de bisubstrats. Ainsi, nous avons tenté une dernière approche qui consiste à concevoir une molécule hameçon possédant un groupement guanidine électrophile capable de réagir avec un peptide cystéine.

VI. Réaction *in-situ* entre une molécule hameçon électrophile guanidine et un peptide cystéine par réaction de substitution nucléophile

Dans cette partie, nous avons conçu une molécule hameçon électrophile guanidine dans laquelle est incorporée une chaîne aliphatique bromée (molécule **90**). Ce dernier serait capable de capturer un peptide cystéine. Le mélange réactionnel, composé de la molécule hameçon **90**, d'un peptide cystéine et de l'enzyme CARM1, sera testé en co-cristallisation et analysé par diffraction aux rayons X suivant le même protocole décrit dans le paragraphe II de ce chapitre (figure **4.26**)

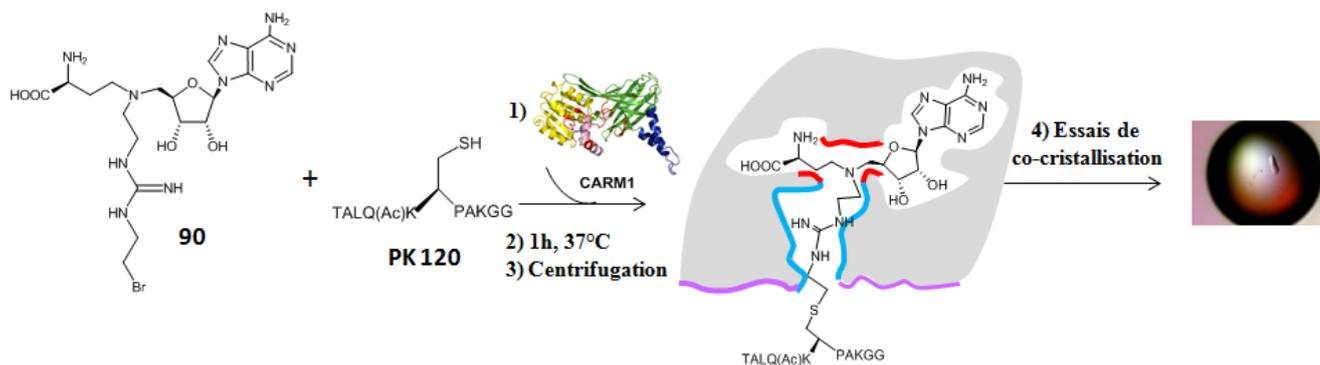


Figure 4.26. : Stratégie pour capturer un peptide cystéine par réaction de substitution nucléophile avec la molécule hameçon électrophile guanidine **90**. 1) Introduction de l'enzyme CARM1 au mélange molécule hameçon/peptide. 2) Incubation du mélange réactionnel à 37°C pendant une heure. 3) Centrifugation. 4) Essais de co-cristallisation et analyse au synchrotron.

VI.1 Molécule hameçon électrophile guanidine pour la capture de peptides cystéine.

La molécule hameçon **90** possède un groupement guanidine dans lequel un groupement électrophile a été introduit pour pouvoir réagir avec les peptides **PK45** et **PK120** ainsi que la cystéine modifiée **74**. La longueur de chaîne a été choisie en se basant sur l'état de transition de la réaction entre la SAM et l'arginine de l'histone (figure **4.27**).

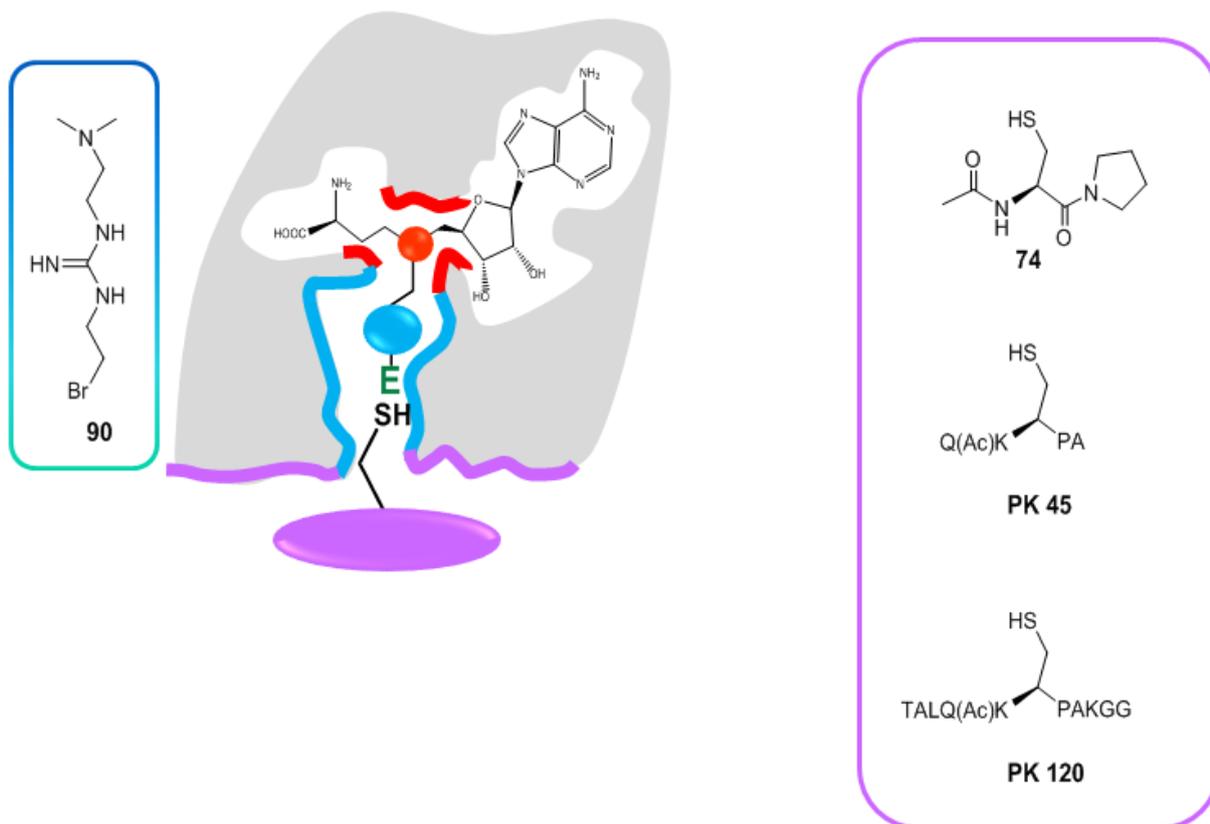


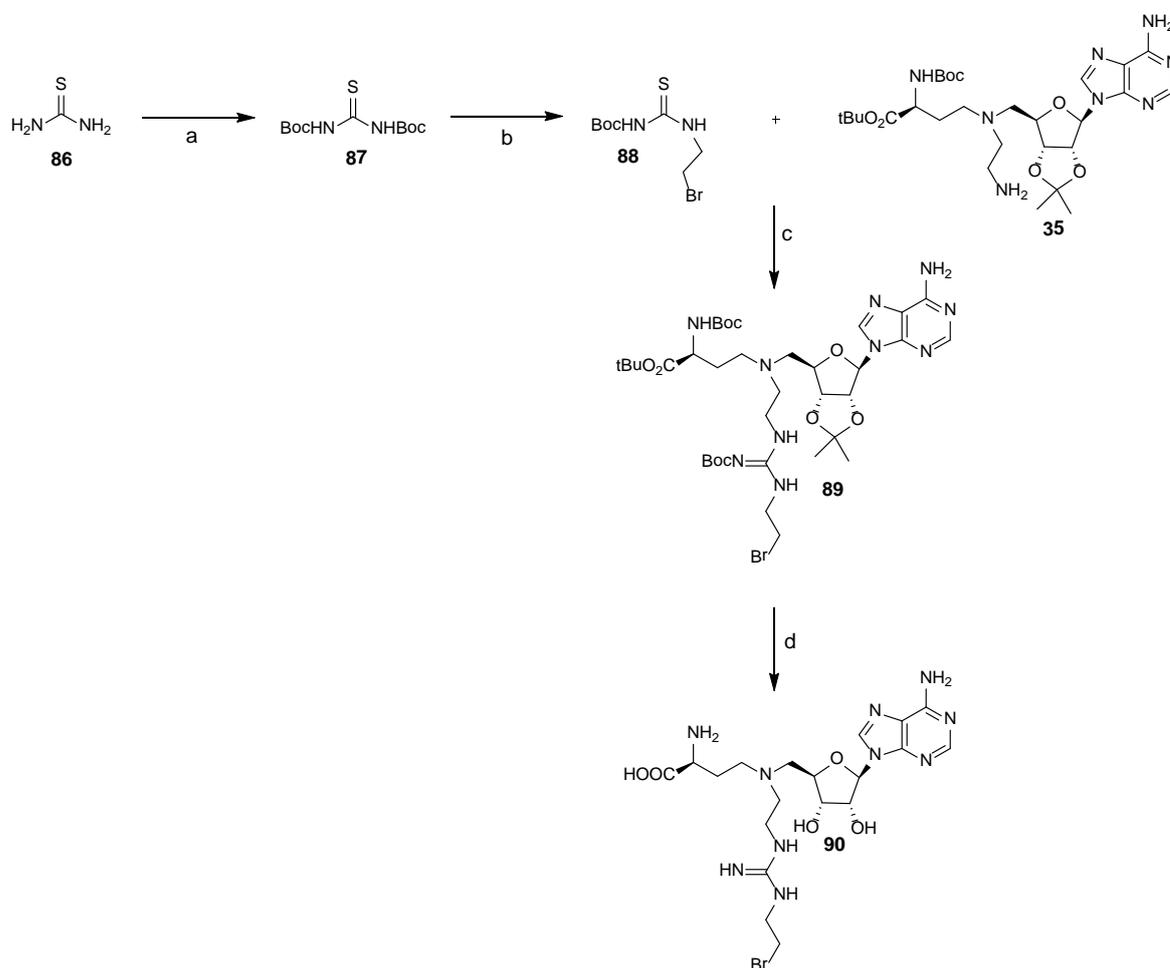
Figure 4.27 : Molécule hameçon **90** et les peptides **PK45**, **PK120** et de la cystéine modifiée **74** utilisés pour réaliser la réaction *in-situ*.

La première étape de cette approche a été de synthétiser la molécule hameçon **90**.

VI.1.1 Synthèse de la molécule hameçon **90**

La première étape a été l'obtention de la thiourée **88** qui a été synthétisée en deux étapes à partir de la thiourée **86** commerciale en s'inspirant du protocole décrit par Yin et al⁽²⁶⁾. La molécule **90** a ensuite été synthétisée avec 63% de rendement à partir de la thiourée **88** et l'amine **35** en présence d'EDC dans la N,N-diméthylformamide en chauffant à 70°C pendant 12 heures⁽²⁷⁾.

Ce chemin réactionnel est décrit sur la figure 4.28.



Conditions réactionnelles : a) NaH, THF, 10 minutes à 0°C puis Boc₂O, 40 minutes à 0°C puis température ambiante, 14 heures, 97%. b) NaH, THF, 1 heure à 0°C, TFAA, 1h à 0°C, bromoéthylamine.HCl, DIEA, à température ambiante, 50h, 17%. c) EDC.HCl, DIEA, DMF, 70°C, 12 heures, 63%. d) HCl (1,25M dans le dioxane) 2h, température ambiante, rendement quantitatif.

Figure 4.28 : Synthèse de la molécule guanidine électrophile 90.

VI.2 Essais de réactions *in-situ* de la molécule hameçon 90 et les peptides cystéines puis co-cristallisation dans CARM1

La molécule hameçon 90 a été testée en inhibition avant d'être testée en co-cristallisation dans le mélange réactionnel.

VI.2.1 Mesures d'inhibitions de la molécule hameçon 90

Les tests d'inhibition ont été réalisés suivant la méthode décrite dans le chapitre 3 c'est-à-dire en mettant en compétition le cofacteur, la SAM et les différents analogues, dans le site catalytique de CARM1. La référence utilisée pour nos tests d'activité est l'analogue **45**. La valeur de CI_{50} obtenue pour la molécule **90** est de 14 μ M (tableau **4.10**).

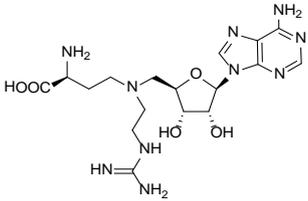
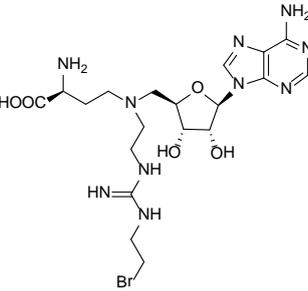
Analogues	Structures	CI_{50} DS [μ M] CARM1
45		2 ± 1
90		14 ± 5

Tableau 4.10 : Valeur de CI_{50} de la molécule **90** contre CARM1; le substrat utilisé est le tétramère d'histone (H3H4)₂ et le ¹⁴C SAM est le donneur de méthyle radioactif.

L'électrophile **90** inhibe CARM1 avec une valeur de CI_{50} de 14 μ M ; il est, pour l'instant, la meilleure molécule hameçon que nous ayons obtenue.

VI.2.2 Essais de co-cristallisation du mélange réactionnel

Des essais de réaction *in-situ* impliquant cette molécule hameçon, les peptides cystéines **PK45** et **PK120** ainsi que la cystéine modifiée **74** ont été réalisés puis des essais de co-cristallisation ont été faits.

La structure cristallographique obtenue ne montre pas la présence d'un peptide, seule la molécule **90** est visualisée mais pas entièrement car l'atome de brome n'est pas visible. De plus, cette molécule occupe uniquement un site sur quatre du tétramère de CARM1. Ce résultat suggère que, malgré l'utilisation d'une molécule hameçon ayant une bonne capacité à se lier à l'enzyme, la réaction *in-situ* n'a pas eu lieu. De plus, nous sommes confrontés à autre problème : la résolution de la structure.

Par conséquent, pour explorer « le domaine de fixation du peptide », nous avons décidé de préformer des adduits chimiquement puis de réaliser des essais de co-cristallisation dans CARM1. Cette approche sera abordée dans le dernier chapitre.

VII. Conclusions et perspectives

Pour explorer le « domaine de fixation du peptide », nous voulions obtenir une structure cristallographique du complexe CARM1/molécule hameçon/peptide. Des essais de click *in-situ* ont été réalisés pour pouvoir capturer le peptide et ainsi obtenir une structure cristallographique de ce complexe. La méthode de choix était la création d'un lien covalent en utilisant trois approches.

Dans une première approche, des essais de click *in-situ* ont été réalisés entre une molécule hameçon possédant une fonction azoture (analogue **54**) et un peptide mimant la séquence 12-22 de l'histone H3 (**PJ207**) dans l'enzyme CARM1. Ce mélange réactionnel a été mis en co-cristallisation et les cristaux obtenus n'ont pas mis en évidence la présence d'adduits ; en revanche, une structure de la molécule **54** a été obtenue avec une résolution de 2,4 Å. Un nouvel essai a été réalisé en ajoutant dans le mélange réactionnel du sulfate de cuivre à une concentration de 12 µM ; malheureusement cette quantité de cuivre a fait précipiter la protéine. Nous nous sommes alors tournés vers une autre technique pour analyser

le mélange réactionnel : la spectrométrie de masse pour les complexes non-covalents. Cette technique ne nous a pas permis d'obtenir plus d'informations sur la formation d'un nouvel adduit. Face à ces difficultés, une seconde approche a été mise en œuvre.

Cette seconde approche a été de réaliser des réactions *in-situ* entre des molécules hameçons électrophiles capable de capturer un peptide cystéine. Ces peptides possèdent une cystéine à la place de l'arginine 17 de l'histone H3 et présentent les différentes séquences d'acides aminés 12-22 et 15-19 et une cystéine modifiée **74**. Les essais de réaction *in-situ* n'ont cependant pas abouti. Les cartes de densités électroniques obtenues ont uniquement mis en évidence la présence de deux molécules **65** et **70** avec une résolution respective de 2,6 Å et 2,5 Å. Aux vues de ces résultats décevants, une troisième approche a été entreprise.

Cette approche a consisté à réaliser des réactions de condensation *in-situ* entre un peptide arginine, comprenant différentes séquences d'acides aminés 3-25, 12-29 et 1-34, et des molécules hameçons électrophiles. Ces essais *in-situ* n'ont pas été concluants et la structure cristallographique obtenue avec la molécule **85** n'a pas présenté une assez bonne résolution pour savoir si c'était cette molécule présente au sein de CARM1 ou bien une molécule de la SAM présente naturellement dans le site actif et qui n'aurait pas été extraite.

Par la suite, une molécule hameçon **90**, possédant un motif guanidine, a été synthétisée pour réagir avec les peptides **PK45** et **PK120** et la cystéine modifiée **74**, malheureusement aucun adduits n'a été observé sur la carte de densité électronique. Seule la structure cristallographique de la molécule **90** à été observée dans un site sur quatre de l'enzyme CARM1.

Ces résultats montrent que, malgré l'efficacité de la molécule hameçon pour CARM1, il est nécessaire que le peptide soit aussi très affiné avec l'enzyme. Des analyses de masse en

cours ont montré que le peptide **PK110** se fixe seulement sur un site sur quatre de l'enzyme. C'est pourquoi des études sont maintenant réalisées par nos collaborateurs pour déterminer quel peptide présente une haute affinité avec CARM1.

Nous pourrions aussi envisager des analyses par marquage impliquant des anticorps de la SAM pour étudier le transfert de molécules hameçons électrophiles vers un peptide.

Aux vues de ces résultats, une nouvelle stratégie a été élaborée. Elle consiste à préformer un adduit chimiquement puis à le co-cristalliser dans l'enzyme. La structure cristallographique obtenue nous permettra certainement d'explorer le « domaine de fixation du peptide ». Cette stratégie va être développée dans le dernier chapitre.

Références

1. H.C. Kolb, M.G. Finn, K.B. Sharpless, *Angew. Chem. Int. Ed.*, **2001**, *40*, 2004-2021.
2. a) W.L. Mock, T.A. Irra, J.P. Wepsiec, T.L. Manimaran, *J. Org. Chem.*, **1983**, *48*, 3619-3620. b) W.L. Mock, T.A. Irra, J.P. Wepsiec, M. Adhya, *J. Org. Chem.*, **1989**, *54*, 5302-5308.
3. W.G. Lewis, L.G. Green, F. Grynszpan, Z. Radic, P.R. Carlier, P. Taylor, M.G. Finn, K.B. Sharpless, *Angew. Chem. Int. Ed.*, **2002**, *41*, 1053-1057.
4. S.K. Mamidyala, M.G. Finn, *Chem. Soc. Rev.*, **2010**, *39*, 1252-1261.
5. P.R. Carlier, D.M. Du, Y.F. Han, J. Liu, E. Perola, I.D. Williams, Y.P. Pang, *Angew. Chem., Int. Ed.*, **2000**, *39*, 1775-1777.
6. L. Tong, *Chem. Rev.*, **2002**, *102*, 4609-4626.
7. M. Whiting, J. Muldoon, Y.C. Lin, M. Silverman, W. Lindstrom, A. Olson, H.C. Kolb, M.G. Finn, K.B. Sharpless, J.H. Elder, V.V. Fokin, *Angew. Chem. Int. Ed.*, **2006**, *45*, 1435-1439.
8. B. Merrifield, *J. Am. Chem. Soc.*, **1963**, *85*, 2149-2154.
9. S.W. Wang, *J. Org. Chem.*, **1976**, *41*, 3258-3261.
10. G. Lu et al., *J. Org. Chem.*, **1981**, *46*, 3433-3436.
11. A.J. Link, M.K.S. Vink, D.A. Tirrell, *J. Am. Chem. Soc.*, **2004**, *126*, 10598-10602.
12. Sarah Sanglier, Alain Van Dorsselaer, Valérie Vivat-Hannah, Jean-Paul Renaud, *spectra analyse*, **2005**, *245*, 26-35.
13. B. Ganem, Y.T Li, J. Henion, *J. Am. Chem. Soc.*, **1991**, *113*, 6294-6296.
14. V. Katta, B.T. Chait, *J. Am. Chem. Soc.*, **1991**, *113*, 8534-8535.
15. A.N. Glazer, *Annu. Rev. Biochem.*, **1970**, *39*, 101-130.

16. J.M. Chalker, G.J.L. Bernandes, Y.A. Lin, B.G. Davis, *Chem. Asian J.*, **2009**, *4*, 630-640.
17. P.Y. Renard, H. Schwebel, P. Vayron, E. Leclerc, S. Dias, C. Mioskowski, *Tetrahedron. Lett.*, **2001**, *42*, 8479-8481.
18. M. Bodanszky, A. Bodanszky, *In The Practice of Peptide Synthesis*, 2nd ed.; Springer-Verlag: Berlin Heidelberg, Germany, **1994**; pp 68–69.
19. Osborne, R. L. Weller Roska, S.R. Rajski, P. R. Thompson, *J. Am. Chem. Soc.*, **2008**, *130*, 4574-4575.
20. A.L. Grossberg, D. Pressman, *Biochemistry*, **1968**, *7*, 272-279.
21. M.Z. Atassi, A.V. Thomas, *Biochemistry*, **1969**, *8*, 3385-3394.
22. K. Takahashi, *J. Biol. Chem.*, **1968**, *243*, 6171-6179.
23. D.L. Smith, J.A. Mc Closkey, J.D. Scribner, *J. Org. Chem.*, **1978**, *43*, 2087-2088.
24. C. Marchi, E. Trepas, M. Moreno-Manas, A. Vallribera, E. Molins, *Tetrahedron*, **2002**, *58*, 5699-5708.
25. Y. Minoura, M. Takebayashi, C.C. Price, *J. Am. Chem. Soc.*, **1959**, *81*, 4689-4692.
26. B-L. Yin, Z-G. Liu, J-C Zhang, Z-R Li, *Synthesis*, **2010**, *6*, 991-999.
27. N.L. Martin, R.M. Liskamp, *J. Org. Chem.*, **2008**, *73*, 7849-7851.

Chapitre 5

Approche pour l'exploration du « domaine de fixation du peptide » par essais de co-cristallisation de molécules ancrées dans CARM1

Chapitre 5

Approche pour l'exploration du « domaine de fixation du peptide » par essais de co-cristallisation de molécules ancres dans CARM1

Dans le chapitre précédent, nous avons tenté d'explorer le « domaine de fixation du peptide » par réaction *in-situ* entre des molécules hameçons et des peptides. Malheureusement, cette stratégie à échoué, c'est pourquoi une seconde alternative a été envisagée : elle consiste à préformer des molécules ancres puis à les co-cristalliser dans CARM1. Ce chapitre sera divisé en trois parties :

- La première décrira la stratégie employée pour préformer ces adduits.
- La synthèse des adduits SAM-peptides sera abordée dans la seconde partie. Ces adduits seront obtenus grâce à une réaction de cycloaddition de Huisgen. Dans ce paragraphe, nous ajouterons aussi les caractérisations biochimiques de ces adduits.
- La troisième et dernière partie présentera enfin la synthèse d'un adduit SAM-peptide obtenu par réaction de substitution nucléophile entre une molécule hameçon électrophile et un acide aminé **74** portant un motif cystéine modifié. Les caractérisations biochimiques de ce composé seront aussi décrites.

I. Stratégie employée pour explorer le « domaine de fixation du peptide »

La synthèse chimique d'adduits SAM-peptides a été l'alternative choisie pour explorer le « domaine de fixation du peptide ». Deux voies de synthèse ont été utilisées pour former ces adduits :

- La première consiste en une réaction de cycloaddition de Huisgen⁽¹⁾ entre une molécule hameçon azoture et un acide aminé modifié portant un motif propargyle (figure **5.1 a**)
- La deuxième est une réaction de substitution nucléophile entre une molécule électrophile et la cystéine modifiée **74**. (figure **5.1 b**)

Après avoir isolé ces adduits, nous les mettrons ensuite en présence de l'enzyme CARM1 pour tenter de les faire réagir ensemble. Pour cela, pour chaque essai de co-cristallisation réalisé avec ces adduits, le protocole suivant sera toujours respecté : le mélange réactionnel est incubé à 37°C pendant une heure, puis est centrifugé. Le surnageant récupéré est ensuite testé en co-cristallisation ; des additifs y sont ajoutés pour obtenir plus facilement des cristaux. Ces derniers sont finalement analysés par diffraction aux rayons X (figure **5.1**).

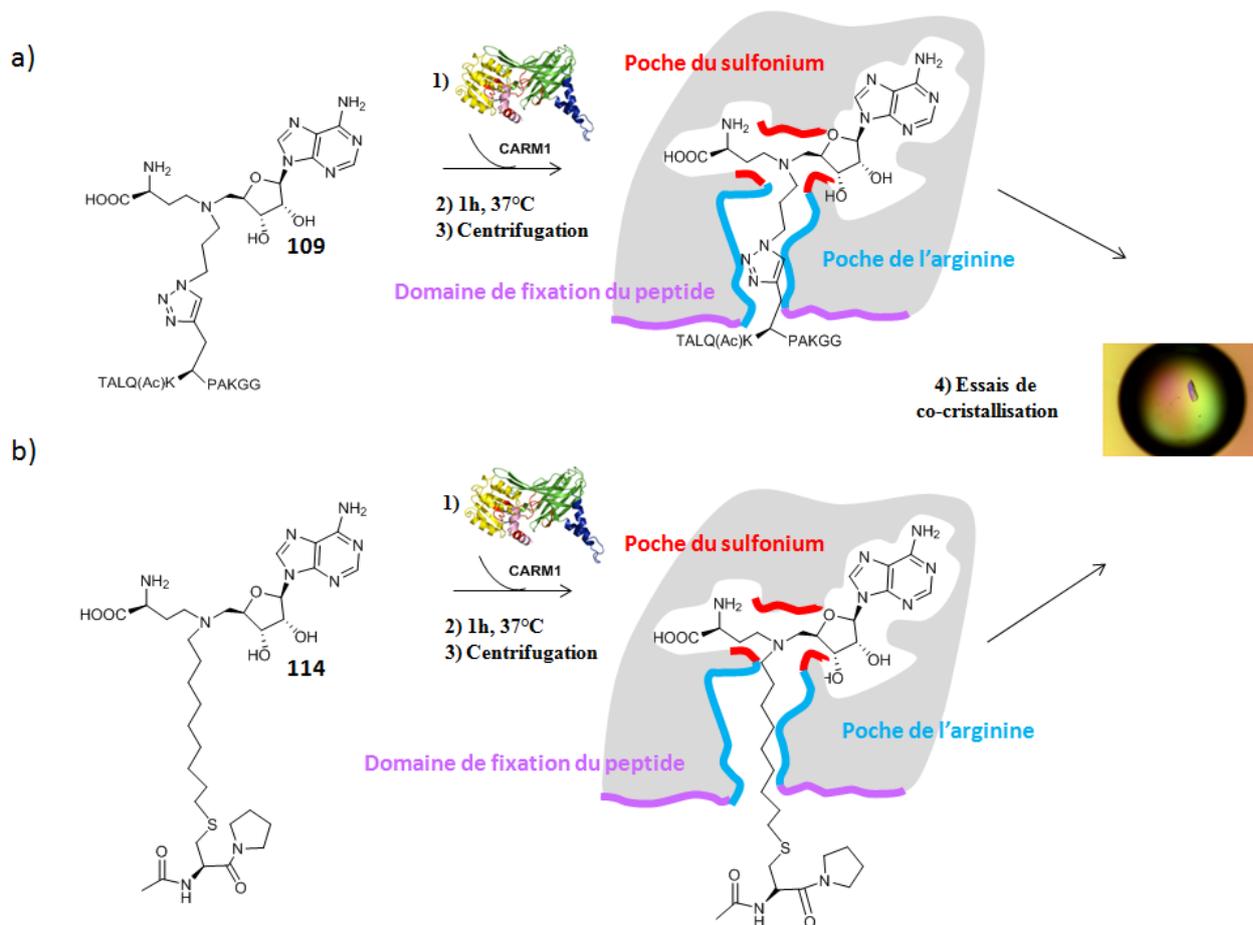


Figure 5.1: Stratégie employée pour explorer le « domaine de fixation du peptide ». a) Adduits SAM-peptide obtenus par réaction de cycloaddition de Huisgen puis testés en co-cristallisation dans CARM1. b) Adduits SAM-peptide obtenus par réaction de substitution nucléophile puis testés en co-cristallisation dans CARM1.

Abordons maintenant la synthèse d'adduits obtenus par cycloaddition de Huisgen et les résultats obtenus avec ceux-ci pour les essais de co-cristallisation.

II. Exploration du « domaine de fixation du peptide » par essais de co-cristallisation d'adduits SAM-peptides obtenus par réaction de cycloaddition de Huisgen

II.1 Adduits SAM-peptide obtenus par « Cu (I)-catalyzed azide-alkyne cycloaddition » (ou CuAAC) et « Ru (II)-catalyzed azide-alkyne cycloaddition » (ou RuAAC)

Pour explorer le « domaine de fixation du peptide », huit adduits SAM-peptides ont été sélectionnés (figure 5.2) :

- Nous avons tout d'abord envisagé de modifier la longueur de la chaîne entre l'atome d'azote de l'analogue et le triazole pour être plus proche de l'état de transition (composés **103** et **104**).
- Nous avons voulu ensuite déterminer si l'isomérisation du triazole pouvait avoir de l'importance : les 1,4- régioisomères **103** et **104** ont été synthétisés grâce à une réaction de « Cu (I)-catalyzed azide-alkyne cycloaddition » (CuAAC)⁽²⁻³⁾ tandis que les 1,5- régioisomères **105** et **106** ont été isolés après une réaction de « Ru (II)-catalyzed azide-alkyne cycloaddition » (RuAAC)⁽⁴⁻⁷⁾.
- Finalement, nous avons modifié la longueur de chaîne du peptide et protégé les amines et acides carboxyliques de ce dernier pour étudier les interactions qui pourraient avoir lieu entre le substrat et CARM1 (adduits **107** à **110**).

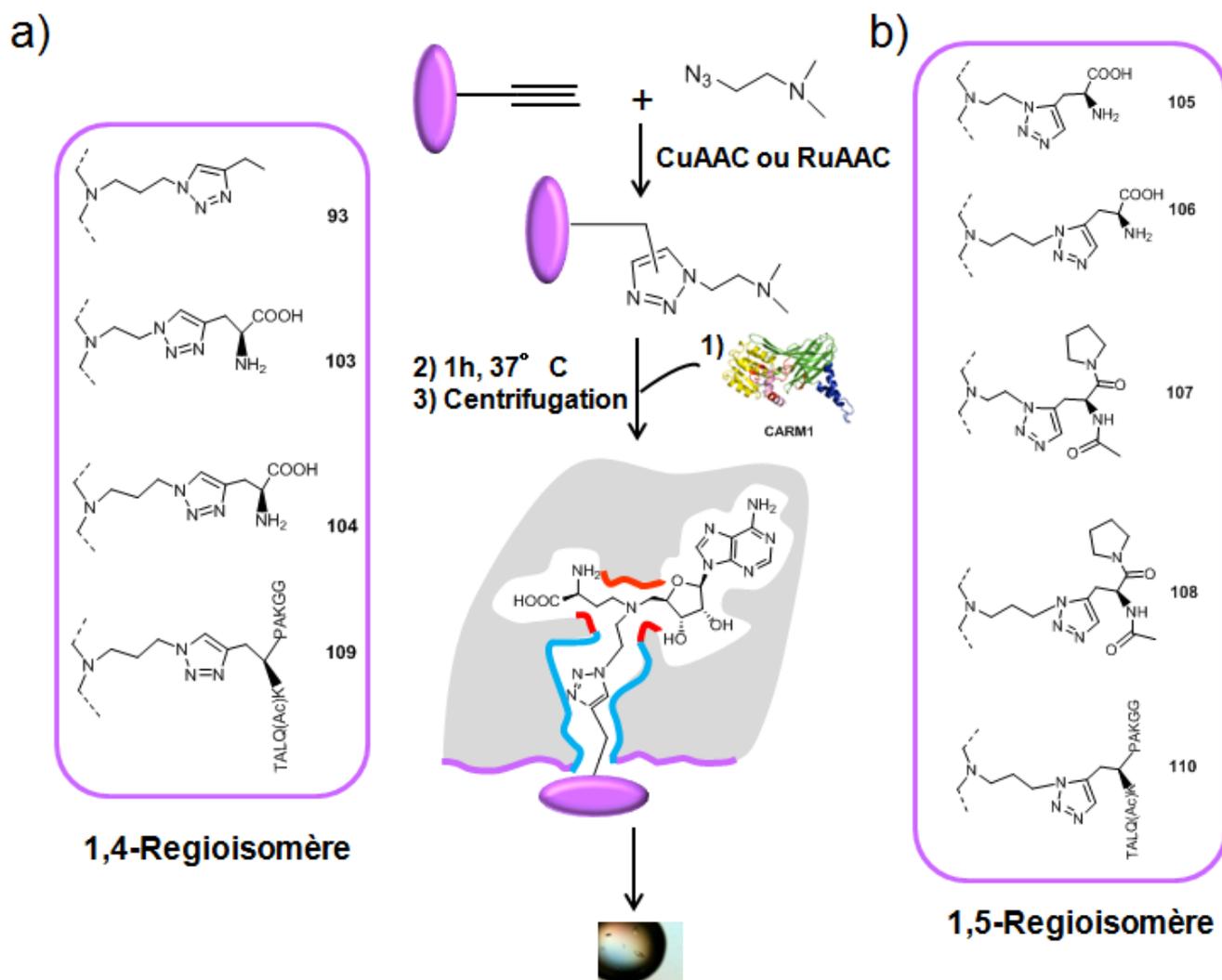


Figure 5.2 : Adduits SAM-peptide (**93** ; **103** à **110**). a) Obtention d'adduits SAM-peptide (**93**, **103**, **104**, **109**) par « Cu (I)-catalyzed azide-alkyne cycloaddition » (CuAAC). b) Obtention d'adduits SAM-peptide (**105** à **108**, **110**) par « Ru (II)-catalyzed azide-alkyne cycloaddition » (RuAAC).

La synthèse de ces adduits SAM-peptide va être détaillée dans le paragraphe suivant.

II.2 Synthèse des adduits SAM-peptide 93, 103 à 110

Les adduits SAM-peptide ont été synthétisés par réaction de cycloaddition de Huisgen et ont été obtenus grâce aux dérivés azotures **16b** et **34**, au butyne, aux alcynes **95**, **96** et **PJ207**.

Les alcynes **95** et **96** ont été synthétisés à partir de la *L*-propargylglycine commerciale **94** (figure 5.3.a). L'alcyne **95** a été isolé avec un rendement de 83% après protection de son amine par un groupement Boc⁽⁸⁾. Quant à l'alcyne **96**, il a été formé en deux étapes : la première a consisté en un couplage peptidique entre la *L*-propargylglycine et la pyrrolidine. Ce couplage a été suivi par une acétylation de l'amine restante, réalisée en présence d'anhydride acétique et de pyridine. Le rendement global de ces deux étapes est de 70%⁽⁹⁾.

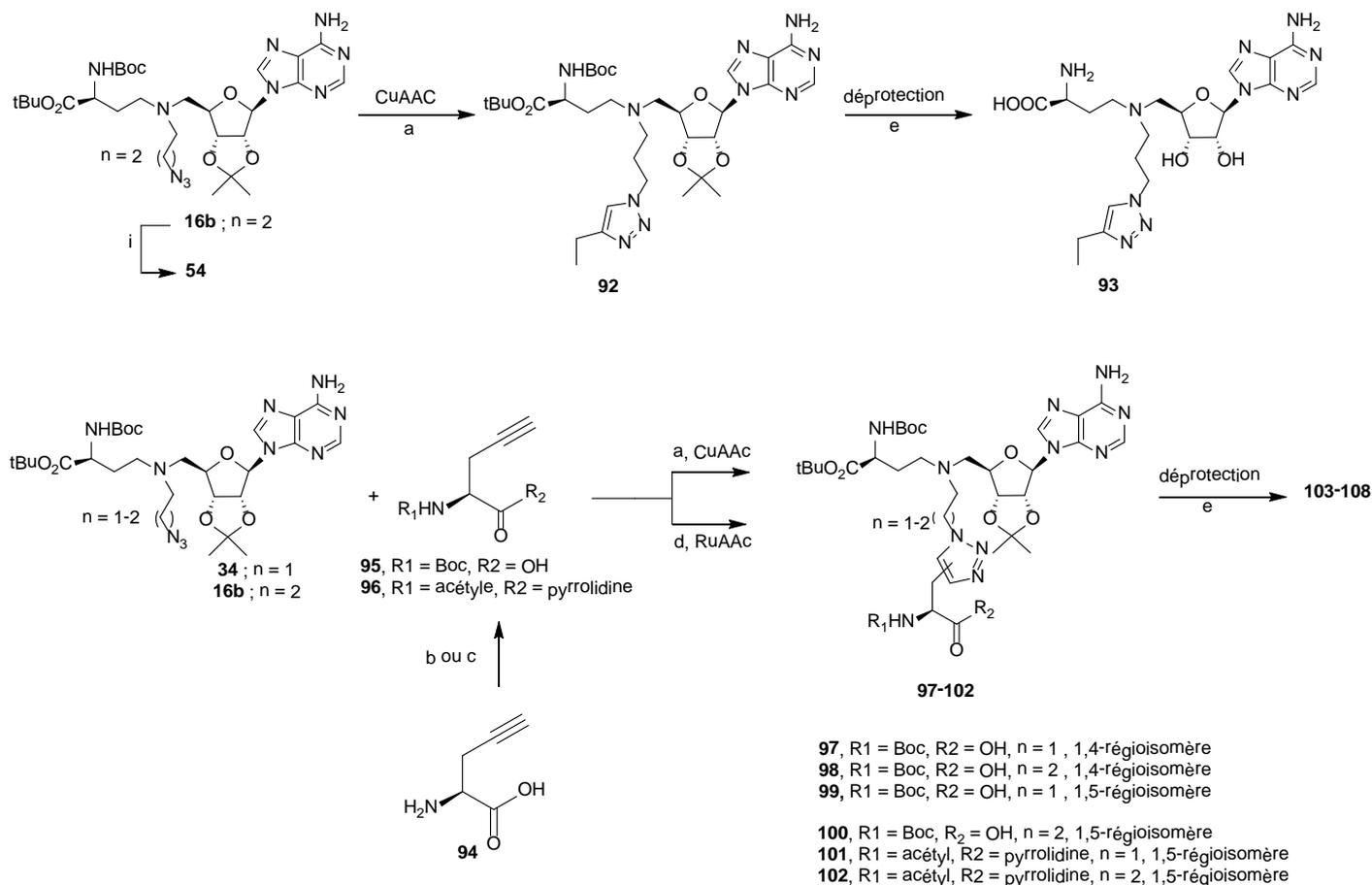
La synthèse du peptide **PJ207** a été décrite dans le chapitre précédent.

Les triazoles 1,4-disubstitués **92**, **97** et **98** ont été synthétisés à partir des dérivés **16b** et **34**, du butyne et de l'alcyne **95**; ils ont été formés grâce à une réaction de CuAAC, réalisée en présence de sulfate de cuivre et d'ascorbate de sodium⁽³⁾.

Pour obtenir respectivement les triazoles **97** et **98**, les azotures **34** et **16b** ont été mis en présence de l'alcyne **95** et ont été chauffés à 60°C dans un mélange de solvant eau/*tert*-Butanol (1v/1v). En utilisant les mêmes conditions opératoires, il a été possible de synthétiser l'analogue **92** à partir du butyne commercial et de l'azoture **16b**. Chacun des adduits **92**, **97** et **98** a été obtenu avec un rendement moyen de 70% (figure 5.3.a).

En parallèle, les triazoles 1,5-disubstitués **99** à **102** ont été obtenus avec un rendement moyen de 50% grâce à une réaction de RuAAC : l'alcyne **95** a été mis en présence de l'azoture **34** ou **16b** pour former respectivement l'analogue **99** ou **100**. De même, à partir du composé **96** et des analogues **34** et **16b**, ont pu être isolés les triazoles **101** et **102** respectivement. Chacune de ces réactions de RuAAC a été réalisée sous reflux de tétrahydrofurane et en présence de 5% molaire de catalyseur Cp*Ru (PPh₃)₂Cl⁽⁸⁾.

Après déprotection des analogues **92**, **97** à **102** en milieu acide chlorhydrique, les composés désirés **93**, **103** à **108** ont été isolés avec un rendement quantitatif (figure 5.3.a).

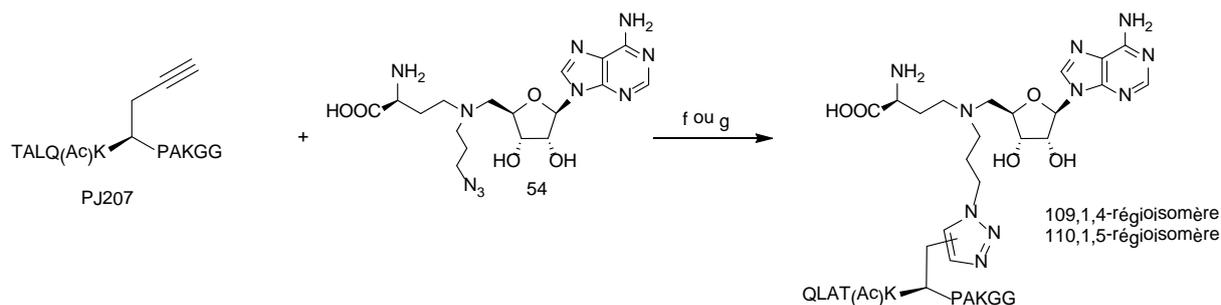


Conditions réactionnelles: a) Butyne, CuSO₄, sodium ascorbate, tBuOH/H₂O (1:1), 60 degrés, 12h, 70%, (CuAAC). b) 1,4-dioxane/H₂O (2:1), 0 degré, solution 2M NaOH, Boc₂O, 2h, TA, 83%. c) EDC.HCl, DIEA, DCM, pyrrolidine, TA, 12h. DCM, pyridine, anhydride acétique, 12h, TA (70% sur les deux étapes). d) Cp*Ru(PPh₃)₂Cl (5 mol%), THF, 60 degrés, 12h, 50-67%, (RuAAC). e) HCl (1,25M dans le dioxane), 2h, TA, quantitatif.

Figure 5.3.a : Synthèse des adduits SAM-peptide **93**; **103-108**.

Dans un dernier temps, les adduits SAM-peptide **109** et **110** ont été envisagés et obtenus à partir de l'azoture **54** et du peptide **PJ207**. Ces deux composés ont été mélangés dans de l'eau et chauffés à 37°C pendant 12 heures pour former le dérivé **109** grâce à une CuAAC. Le rendement de cette cycloaddition est de 67% figure 5.3.b).

Quant à l'adduit **110**, il a été obtenu en mélangeant l'analogue **54** et le peptide **PJ207** dans un mélange de solvant eau/méthanol ; ce milieu réactionnel a été chauffé à 37°C pendant 12 heures, en présence de 5 mol% de catalyseur Cp*Ru (PPh₃)₂Cl, utilisé pour la formation des triazoles de régiosélectivité 1,5-. Cette cycloaddition a présenté un rendement moyen de 30% (figure **5.3.b**).



Conditions réactionnelles : f) **54**, CuSO₄, sodium ascorbate, H₂O, 37 degrés, 12h, 67%. g) **54**, Cp*Ru(PPh₃)₂Cl (5 mol%), MeOH/H₂O (1:1), 37 degrés, 12h, 30%.

Figure 5.3.b: Synthèse des adduits SAM-peptide **109-110**.

La formation du triazole peut être observée par spectroscopie infra rouge, grâce à la disparition de la bande à 2100 cm^{-1} qui est caractéristique de la fonction azoture $\text{N}_3^{(9)}$. Grâce à la résonance magnétique nucléaire HSQC (pour *Heteronuclear Single Quantum Coherence*), nous pouvons aussi observer une différence entre les triazoles 1,4- et 1,5-disubstitués grâce aux déplacements chimiques du proton et du carbone éthylénique (tableau 5.1).

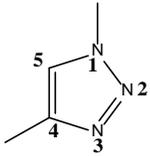
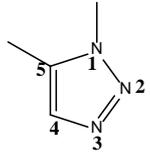
Isomères	Positions	H ¹ (ppm)	C ¹³ (ppm)
<u>1, 4-Régioisomère</u>			
	5	7,9	125
<u>1, 5-Régioisomère</u>			
	4	8,4	132

Tableau 5.1 : Déplacements chimiques des protons en position 4 et 5 du triazole des 1,4- et 1,5-régioisomères.

Grâce à ces différents déplacements et à la spectroscopie infra-rouge, nous avons pu confirmer la formation des ces deux triazoles d'isomérisation différente. Les adduits **93**, **103** à **110** synthétisés doivent maintenant être caractérisés biochimiquement et ces analyses vont être décrites dans le paragraphe suivant.

II.3 Caractérisations biochimiques des adduits 93, 103 à 110

II.3.1 Mesures d'inhibition des adduits SAM-peptides 93, 103 à 110

Les tests d'inhibition ont été effectués suivant le protocole décrit dans le chapitre 3. Les valeurs d'inhibition obtenues avec les adduits **93, 103 à 110** varient suivant la longueur de la chaîne et l'isomérisation des différents composés (tableau **5.2**).

D'après ces tests, le triazole 1,4-disubstitué **93** montre une inhibition de 198 μM . En greffant un acide aminé sur le du triazole (composé **104**), une meilleure inhibition a été obtenue, égale à 70 μM . De plus, en raccourcissant la longueur de la chaîne d'un carbone entre l'atome d'azote de l'analogue et le triazole (adduit **103**), l'inhibition est encore meilleure et s'élève ainsi à 50 μM .

Pour le triazole 1,5-disubstitué, le phénomène inverse est observé : en effet, plus la longueur de la chaîne entre l'atome d'azote de l'analogue et le triazole est longue (adduit **106**) plus la valeur d'inhibition est bonne. Ainsi, l'adduit **106** présente une valeur de CI_{50} égale à 13,3 μM tandis que le même 1,5-régioisomère **105**, à chaîne plus courte, possède une CI_{50} plus forte de 200 μM .

Nous avons voulu vérifier si les acides aminés greffés sur le triazole des adduits pouvaient avoir un effet lors de leur ancrage au sein de CARM1 ; c'est pourquoi, nous avons synthétisé deux composés de 1,5-régioisomérisation **107** et **108** dont les acides carboxyliques ont été protégés par un groupement proline et les amines par un groupement acétyle. Pour les produits **105** et **107**, avec une chaîne courte de deux carbones entre l'azote et le triazole, aucun effet n'a pu être observé puisque les valeurs de CI_{50} sont pratiquement identiques et égales à environ 200 μM . En revanche, pour les adduits présentant une chaîne plus longue, nous observons une moins bonne inhibition, égale à 145 μM , pour l'adduit protégé **107** : les

acides aminés libres du triazole augmentent le potentiel d'inhibition de l'adduit **106** au sein de CARM1 et doivent probablement interagir avec d'autres acides aminés du site actif de l'enzyme.

Enfin, compte tenu des bons résultats d'inhibition obtenus avec l'analogue **106**, deux autres SAM-peptides **109** et **110** ont été synthétisés avec cette isométrie 1,5 ; nous avons voulu synthétiser des adduits dont la structure se rapprocherait au plus près de l'état de transition entre le cofacteur (la SAM) et son substrat (l'histone H3). Dans la structure des composés **109** et **110**, l'acide aminé libre a donc été remplacé par un peptide. En allongeant la longueur du peptide, nous pouvons observer que l'inhibition chute fortement pour le 1,4-régioisomère **109** avec une valeur de CI_{50} égale à 227 μM . En revanche, pour l'autre isomère **110**, nous obtenons une valeur d'inhibition de 56 μM , comprise entre celles des adduits **106** et **108** (tableau 5.2). Pour le composé **110**, l'augmentation de la longueur du peptide a donc permis d'augmenter le potentiel d'inhibition au sein de CARM1. Parmi les composés synthétisés, c'est le composé **106** de régioisométrie 1,5 qui s'est détaché en présentant les meilleurs résultats avec une valeur d'inhibition égale à 13 μM . En le comparant avec le dérivé **105**, nous nous sommes aperçus de l'importance de la longueur de la chaîne carbonée qui doit comprendre trois carbones, entre l'atome d'azote de l'analogue et le triazole, pour augmenter le potentiel d'inhibition des 1,5- régioisomères. Finalement, en greffant sur le triazole des peptides, nous avons constaté que ces derniers permettaient d'inhiber CARM1 avec une bonne valeur de CI_{50} , c'est le cas de l'adduit **110** avec une valeur de CI_{50} égale à 56 μM .

SAM-peptides	Structures	CI ₅₀ [μm] CARM1	SAM-peptides	Structures	CI ₅₀ [μm] CARM1
93		198	107		203
103		55	108		145
104		70	109		227
105		208	110		56

Tableau 5.2 : Activité d'inhibition (valeurs de CI₅₀) des adduits SAM-peptides (**93**, **103** à **110**) contre CARM1. a) le substrat utilisé est le tétramère d'histone (H3H4)₂ et le ¹⁴C SAM est le donneur de méthyle radioactif.

Tous ces adduits SAM-peptides **93**, **103** à **110** sont ensuite testés en co-cristallisation dans CARM1.

II.3.2 *Essais de co-cristallisation des adduits SAM-peptides 93, 103 à 110 dans mCARM1₁₃₀₋₄₉₀*

Les adduits **93**, **103** à **110** ont tous été testés en co-cristallisation suivant le protocole décrit dans le paragraphe I.1 de ce chapitre. Avec ces adduits, il nous serait alors possible de mieux comprendre comment le substrat se fixe au sein de l'enzyme.

Des cristaux ont été obtenus pour chacun des adduits, malheureusement les structures visualisées n'ont pas été complètes. L'adduit **104** est le seul qui a été visualisé presque entièrement : sa structure est visible jusqu'au carbone en α de l'acide aminé du triazole.

Ces résultats nous laissent cependant penser que ces adduits ne nous permettent pas de stabiliser suffisamment CARM1. Ainsi, pour pouvoir explorer le « domaine de fixation du peptide », un autre adduit SAM-peptide va être envisagé puis testé en co-cristallisation. Cette adduit possédera une chaîne aliphatique qui devrait permettre d'éviter d'éventuelles contraintes stériques dans le site actif de l'enzyme. Cet adduit va être décrit dans la dernière partie.

III. Exploration du « domaine de fixation du peptide » par essais de co-cristallisation d'un adduit SAM-peptide obtenu grâce à une réaction de substitution nucléophile

III.1 Adduit SAM-peptide obtenu par substitution nucléophile entre la cystéine modifiée 74 et une molécule hameçon électrophile

Nous avons choisi de synthétiser rapidement un adduit SAM-peptide par substitution nucléophile entre la cystéine modifiée **74** décrite dans le chapitre précédent et une molécule hameçon possédant un brome. La chaîne choisie est longue pour pouvoir sortir du site actif et possède donc neuf carbones (figure **5.4**).

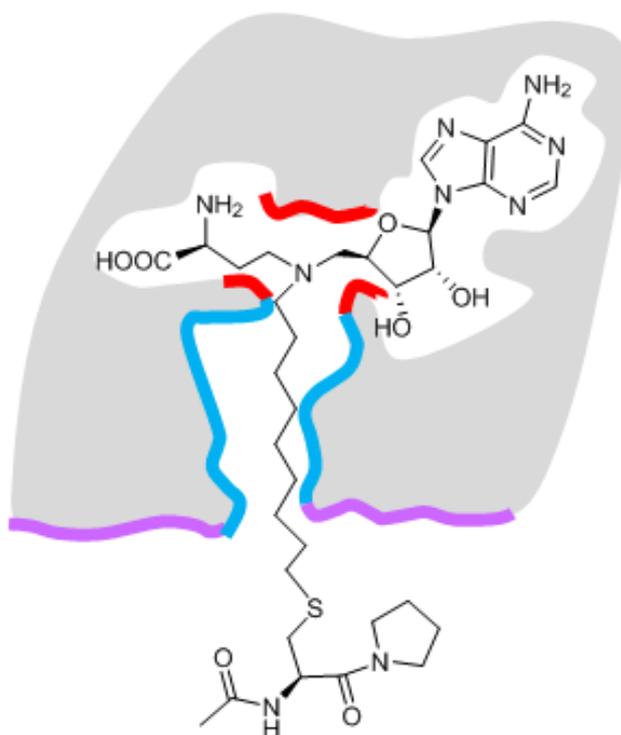
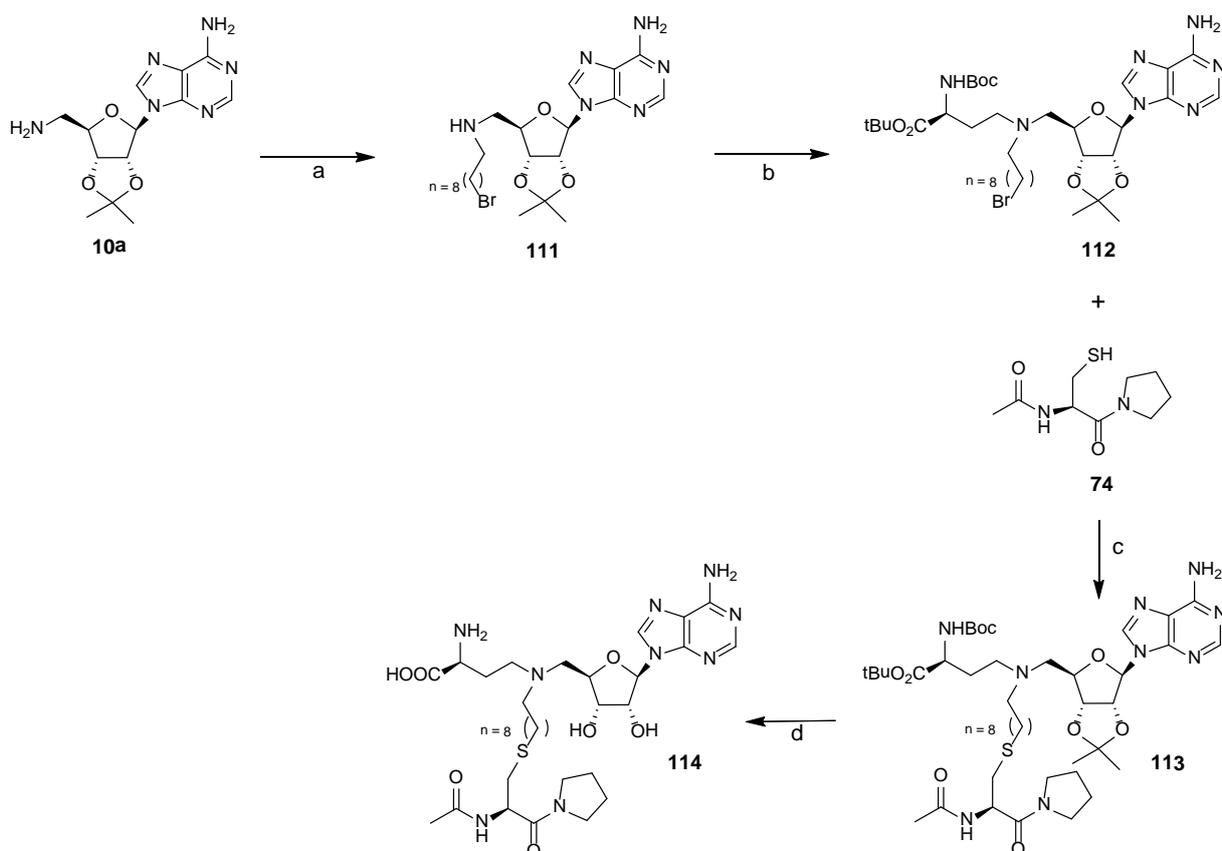


Figure 5.4: Adduit SAM-peptide obtenu par substitution nucléophile entre une molécule hameçon possédant un brome et la cystéine modifiée **74**.

La synthèse de cette molécule hameçon ainsi que celle de cet adduit va être détaillée dans le paragraphe suivant.

III.2 Synthèse de l'adduit SAM-peptide 114 et d'une molécule hameçon 115

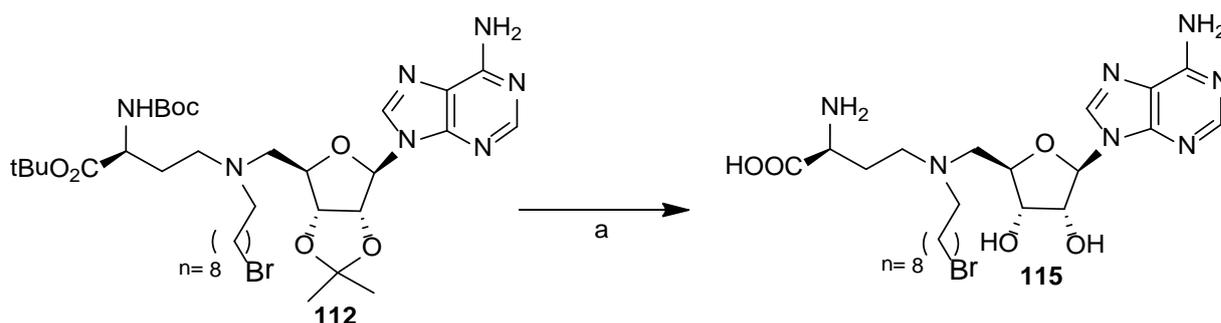
L'adduit **113** est synthétisé grâce à une réaction de substitution nucléophile entre la cystéine modifiée **74** et le dérivé bromé **112**. Ce dernier est obtenu en deux étapes : la première consiste à alkyler l'amine **10a** par le 1,9-dibromononane pour obtenir le dérivé **111** avec un rendement de 75%. Cette alkylation est suivie d'une amination réductrice où le dérivé **111** réagit avec l'aldéhyde **3** pour former le composé **112**. Le rendement de cette étape est de 36%. L'analogue **112** est mis en présence de la cystéine modifiée **74** pour former le produit **113** qui est finalement déprotégé en milieu acide pour obtenir l'adduit **114** avec un rendement quantitatif (figure 5.5).



Conditions réactionnelles a) 1,9-dibromononane, DIEA, THF, 12h, 60°C, 75%. b) **3**, HOAc, NaBH₃CN, MeOH, 12h, température ambiante, 36%. c) K₂CO₃, THF/DMF (1:2), 18h, température ambiante, 10%. d) HCl (1,25M dans le dioxane), 2h, température ambiante, rendement quantitatif.

Figure 5.5 : Synthèse de l'adduit **114**.

En plus de l'adduit **114**, nous avons voulu également tester l'intermédiaire réactionnel **112**. Ce dernier est déprotégé en milieu acide pour obtenir l'analogue **115** (figure 5.6).



Condition réactionnelle : a) HCl (1,25M dans le dioxane), 2h, température ambiante, rendement quantitatif.

Figure 5.6 : Synthèse de l'analogue **115**.

III.3. Caractérisations biochimiques de l'adduit SAM-peptide 114 et de la molécule hameçon 115

III.3.1. Mesures d'inhibition de l'adduit 114 et de la molécule hameçon 115

La référence utilisée pour ces tests est l'analogue de bisubstrat **45** qui présente une valeur d'inhibition de 2 μ M. Ces tests montrent que ces deux composés présentent de bons effets d'inhibition : en effet, l'adduit **114** inhibe l'enzyme avec une valeur de CI₅₀ de 40 μ M tandis que la molécule **115** présente une meilleure inhibition avec une valeur de CI₅₀ plus faible égale à 31 μ M (tableau 5.3).

Molécules	Structures	CI ₅₀ [μM] CARM1
45		2
114		40
115		31

Table 5.3: Activités d'inhibition (valeurs de CI₅₀) de l'adduit SAM-peptide **114** et de la molécule hameçon **115** contre CARM1; le substrat utilisé est le tétramère d'histone (H3H4)₂ et le ¹⁴C SAM est le donneur de méthyle radioactif.

Ces deux composés sont ensuite testés en co-cristallisation.

III.3.2. Essais de co-cristallisation de l'adduit SAM-peptide 114 et de la molécule hameçon 115

Des essais de co-cristallisation dans l'enzyme CARM1 ont été réalisés, en suivant toujours le même protocole. Malheureusement, ces composés ont provoqué la précipitation de l'enzyme CARM1, c'est pourquoi aucun cristal n'a pu être obtenu. Cette précipitation est certainement due à la longueur de la chaîne qui donne un caractère hydrophobe à la protéine. Il faudra donc synthétiser d'autres adduits SAM-peptide possédant une chaîne plus hydrophile comme un polyéthylène glycol ou une spermine.

IV. Conclusions et perspectives

Pour pouvoir explorer le « domaine de fixation du peptide », huit adduits de la SAM-peptide **93**, **103** à **110** ont été synthétisés par réaction de cycloaddition de Huisgen : les 1,4-régioisomères ont été formés en utilisant une catalyse au cuivre tandis que les 1,5-régioisomères ont été obtenus grâce à une cycloaddition catalysée au ruthénium.

Parmi ces analogues, une première série de composés a été obtenue en faisant varier la longueur de chaîne carbonée de l'analogue azoture. En plus de la chaîne carbonée, l'isomérisation du triazole de ces adduits a aussi été modifiée. Enfin, une dernière série de composés a été synthétisée en rallongeant la partie peptidique de l'adduit SAM-peptide.

Le potentiel d'inhibition de chaque composé a été mesuré contre CARM1 et il s'est avéré que le composé **106**, un 1,2,3-triazole 1,5-disubstitué possédant une chaîne de trois carbones, entre l'atome d'azote de la SAM et le triazole, présentait la meilleure inhibition avec une valeur de CI_{50} de 13 μ M.

Les tests de co-cristallisation de ces adduits avec mCARM1₁₃₀₋₄₉₀ ont été infructueux car les structures obtenues n'ont pas montré le composé entièrement dans le site actif de l'enzyme. Le seul adduit qui a été visualisé presque complètement est le dérivé **104** : la structure est visible jusqu'au carbone en α de l'acide aminé. Ce résultat montre donc que la longueur de la chaîne de l'acide aminé du triazole est trop courte : il faudra donc synthétiser de nouveaux adduits possédant une chaîne plus longue.

Un essai de spectrométrie de masse non-covalente a été réalisée avec le composé **109** qui présente une valeur de CI_{50} de 227 μ M : aucune fixation n'a été observée. L'analyse par spectrométrie de masse n'est donc pas adaptée à notre système, peut-être à cause de l'oligomérisation de la protéine ou alors de la faible affinité de nos molécules pour CARM1.

La seconde approche a consisté à réaliser des essais de co-cristallisation dans CARM1 avec l'adduit **115** formé à partir d'une molécule hameçon bromée et une cystéine modifiée. La longue chaîne servait à émerger de la poche de l'arginine afin de sortir vers le « domaine de fixation du peptide ». Malheureusement, cet adduit a provoqué la précipitation de la protéine et aucun cristal n'a été ainsi obtenu. Pour la suite du projet il faudra introduire une chaîne plus hydrophile sur cet adduit afin d'éviter la précipitation.

Références

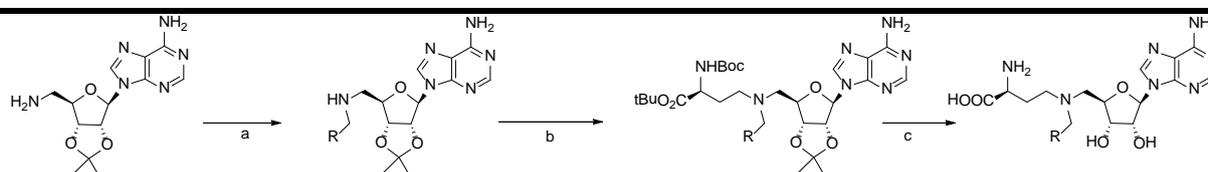
1. H.C. Kolb, M.G. Finn, K.B. Sharpless, *Angew. Chem. Int. Ed.*, **2001**, *40*, 2004.
2. C. W. Tornøe, C. Christensen, M. Meldal, *J. Org. Chem.* **2002**, *67*, 3057.
3. V. V. Rostovtsev, L. G Green, V.V. Fokin, K. B Sharpless, *Angew. Chem., Int. Ed.* **2002**, *41*, 2596.
4. L. Zhang, X.G. Chen, P. Xue, H.H.Y. Sun, I.D. Williams, K.B. Sharpless, V.V. Fokin, *J. Am. Chem. Soc.*, **2005**, *127*, 15998.
5. M.M. Majireck, S.M. Weinreb, *J. Org. Chem.* **2006**, *71*, 8680.
6. L.K. Rasmussen, B.C. Boren, V.V. Fokin, *Org. Lett*, **2007**, *9*, 5337.
7. B.C. Boren, S. Narayan, L.K. Rasmussen, L. Zhang, H. Zhao, Z. Lin, G. Jia, V.V. Fokin, *J. Am. Chem. Soc.*, **2008**, *130*, 8923.
8. N. J. Pedersen; D. Sejer; B.Nielsen ; T. Kvist ; J.M. Mathiesen ; H. Braeuner-Osborne ; D.K. Taylor ; A.D. Abell, *Bioorg.Med.Chem.Lett*, **2010**, *20*, 7512.
9. a) C.J. Nielsen, C.E. Sjøfjogren, *J.Mol.Struct. (THEOCHEM)*, 1987, *150*, 361. b) F.F. Chen, F. Wang, *Molecules*, **2009**, *14*, 2656.

Conclusion générale et perspectives

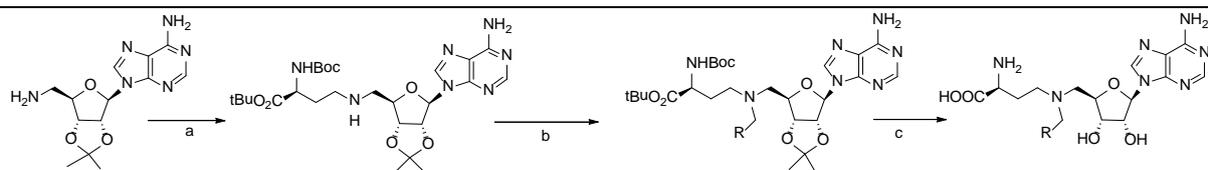
Conclusion générale et perspectives

Le but de ma thèse a été de développer des outils chimiques permettant d'appréhender le mécanisme d'action de l'enzyme CARM1 durant le processus de méthylation afin de comprendre son mode d'interaction avec l'histone. Dans un premier temps, il a fallu synthétiser des analogues de la SAM ainsi plusieurs voies de synthèse ont été envisagées. Les synthèses utilisées sont récapitulées dans le tableau 1 ci-dessous :

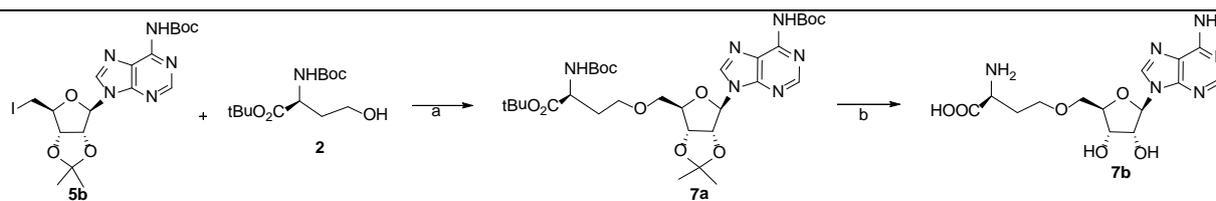
Méthodes utilisées pour la synthèse des analogues de la SAM



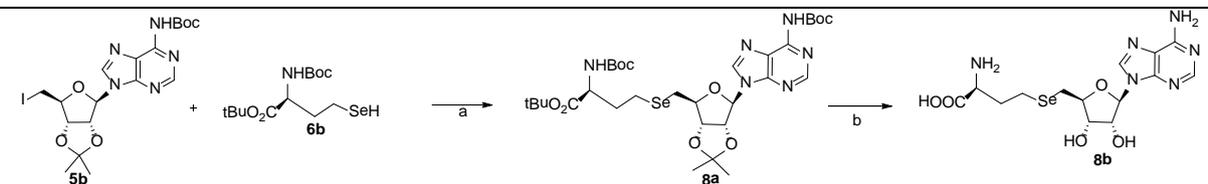
a) RBr, DIEA, THF, 18h, 50°C ; b) **3**, HOAc, NaBH₃CN, MeOH, 12h, température ambiante ; c) HCl (1,25M dans le dioxane), 2h, température ambiante, rendement quantitatif.



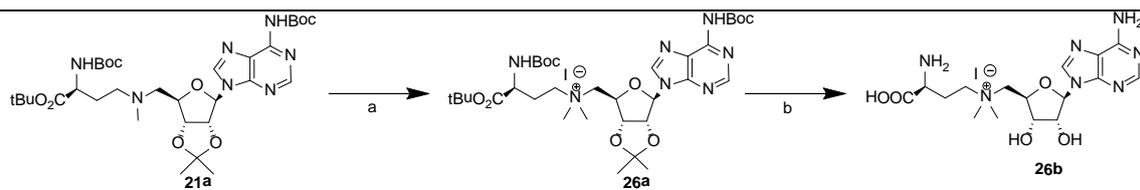
a) RCHO, HOAc, NaBH₃CN, MeOH, 12h, température ambiante, 48% ; b) HCl (1,25M dans le dioxane), 2h, température ambiante, 70%.



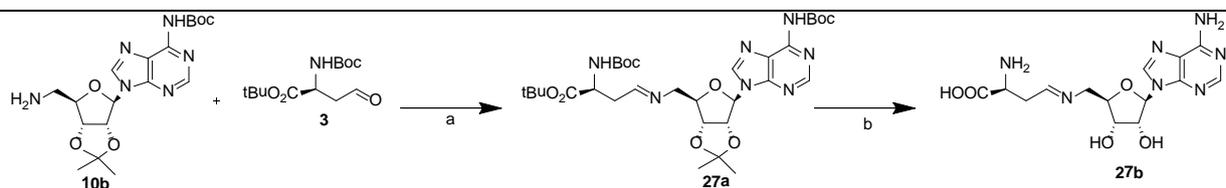
a) NaH, THF, K₂CO₃, 24h à 70°C, 48%, b) HCl (1,25M dans le dioxane), 2h, température ambiante, rendement quantitatif.



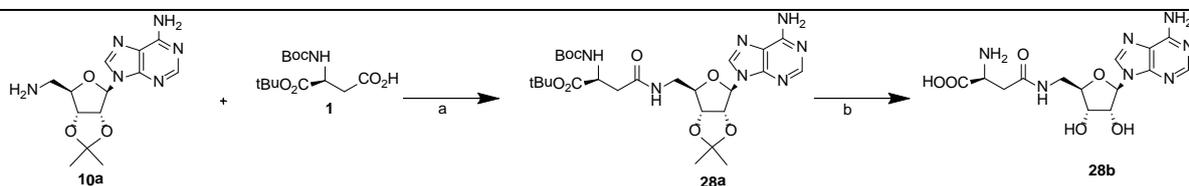
6b, DMF, 12h à 80°C, 21%; b) HCl (1,25M dans le dioxane), 2h, température ambiante, rendement quantitatif.



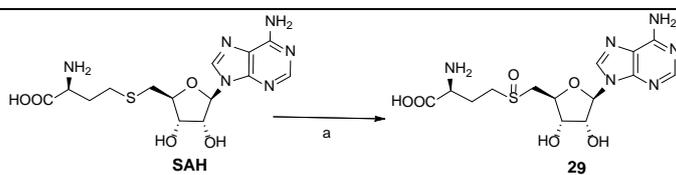
a) MeI, EtOH anhydre, 6 jours à température ambiante b) HCl (1,25M dans le dioxane), 2h, température ambiante, rendement quantitatif.



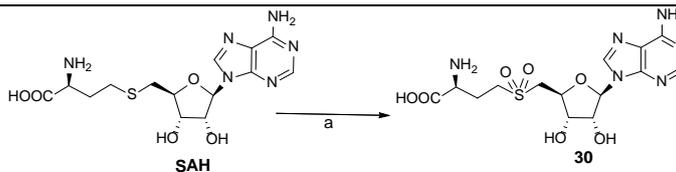
a) HOAc, MeOH, 5 jours à température ambiante, 24% ; b) HCl (1,25M dans le dioxane), 2h, température ambiante, rendement quantitatif.



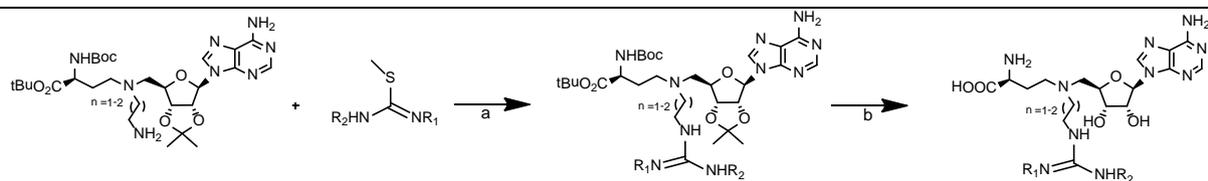
a) **1**, NHS, EDC, DCM, 24h à température ambiante, puis **10a**, DIEA, THF, 3h à 60°C b) HCl (1,25M dans le dioxane), 2h, température ambiante, rendement quantitatif.



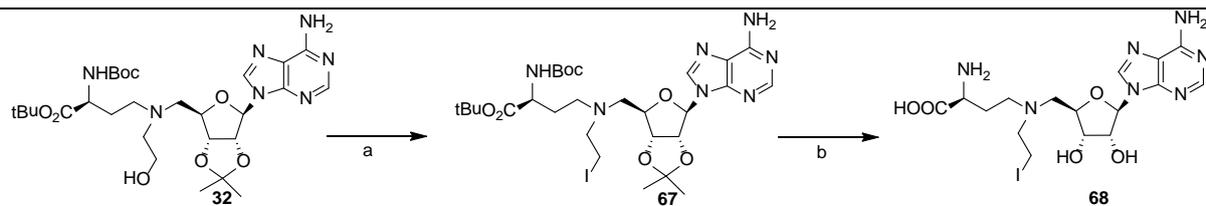
a) H₂O₂ (50% dans l'eau), température ambiante, 24h, 66%.



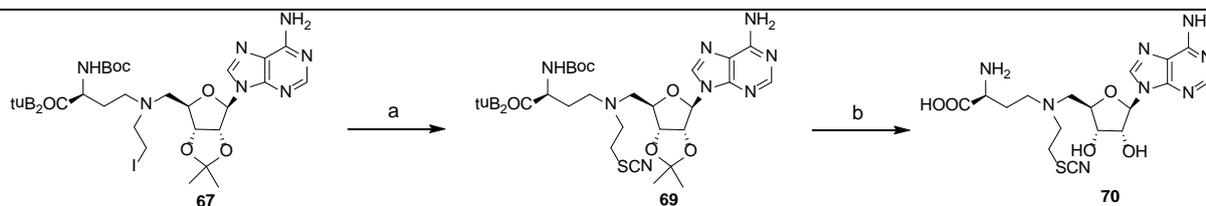
a) (NH₄)₂MoO₄, HClO₄, H₂O₂, H₂O, 0°C-TA, 24h, 74%.



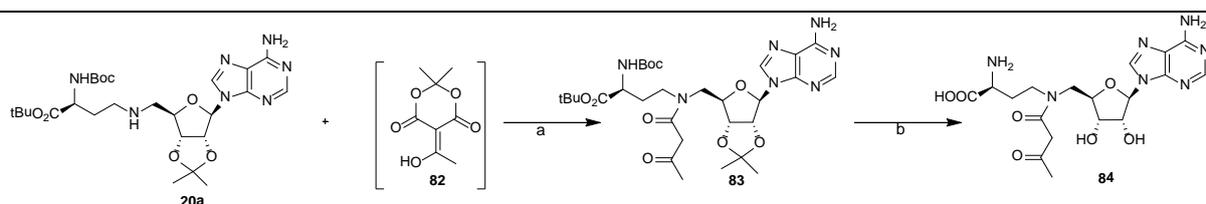
a) CH₃CN, 48h, 80°C, 40-60% ; b) HCl (1,25M dans le dioxane), 2h, TA, rendement supérieur à 98%.



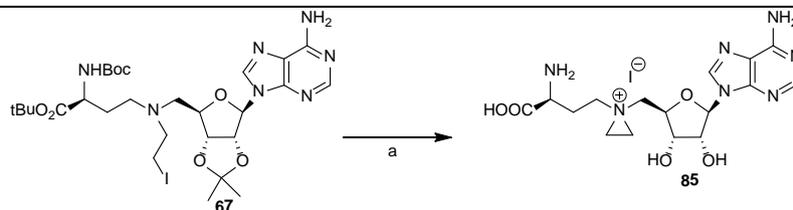
a) I_2 , PPH_3 , imidazole, DCM, 12h, $0^\circ C$ puis température ambiante, 34% ; b) HCl (1,25M dans le dioxane), 2h, température ambiante, rendement quantitatif.



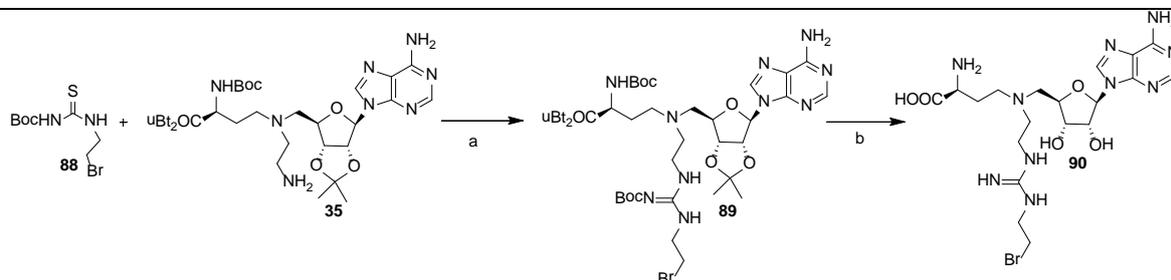
a) TMSNCS, TBAF (0,1M dans THF), 12h, $0^\circ C$ puis température ambiante, 77% ; b) HCl (1,25M dans le dioxane), 2h, température ambiante, 37%.



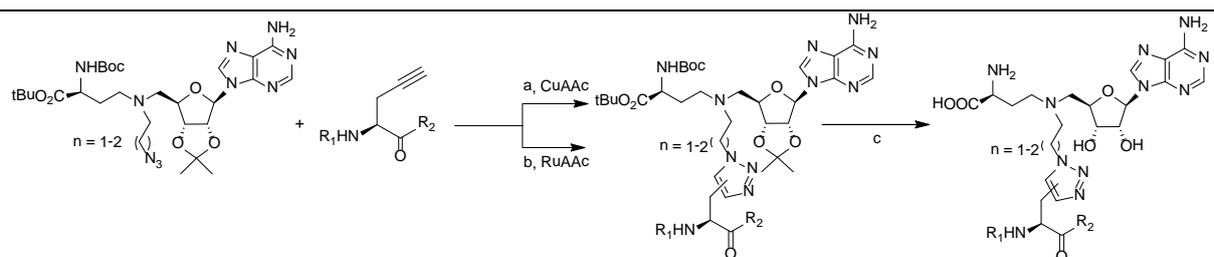
a) DCM, pyridine, DMAP, 12h, $0^\circ C$ puis température ambiante ; DCM/Toluène (1/4), 20h, température ambiante, 31% sur les trois étapes ; b) HCl (1,25M dans le dioxane), 2h, température ambiante, rendement quantitatif.



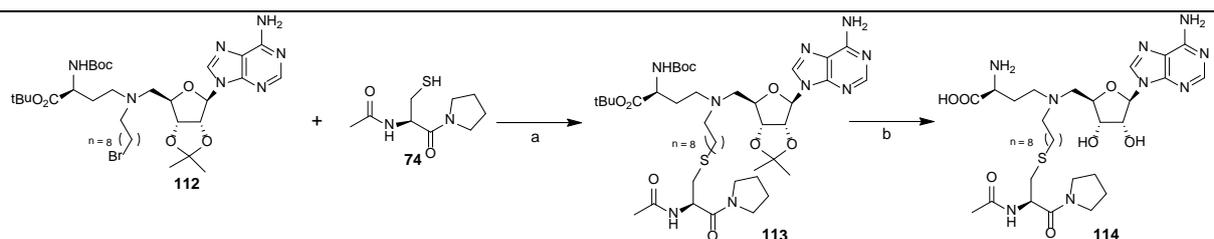
a) HCl (1,25M dans le dioxane), NaOH 0,5M, 14h, température ambiante, 71%.



a) EDC.HCl, DIEA, DMF, $70^\circ C$, 12 heures, 63% ; b) HCl (1,25M dans le dioxane) 2h, température ambiante, rendement quantitatif.



a) Butyne, CuSO_4 , sodium ascorbate, $\text{tBuOH}/\text{H}_2\text{O}$ (1:1), 60 degrés, 12h, 70%, (CuAAC). b) $\text{Cp}^*\text{Ru}(\text{PPh}_3)_2\text{Cl}$ (5 mol%), THF, 60 degrés, 12h, 50-67%, (RuAAC). c) HCl (1,25M dans le dioxane), 2h, TA, quantitatif.



a) K_2CO_3 , THF/DMF (1:2), 18h, température ambiante, 10% ; b) HCl (1,25M dans le dioxane), 2h, température ambiante, rendement quantitatif.

Tableau 1 : Récapitulatif des méthodes utilisées pour la synthèse des analogues de la SAM

Dans une première partie, nous avons exploré la « poche du sulfonium » en synthétisant des analogues de la SAM obtenus grâce à des modifications réalisées autour de l'atome de soufre. En remplaçant la fonction sulfonium par un atome d'azote, nous avons gardé l'affinité au sein de l'enzyme CARM1.

Pour commencer à explorer la « poche de l'arginine », des analogues de la « SAM arginine mimiques » ont été envisagés ; trois d'entre eux (**18b**, **19b** et **23b**) ont montré une bonne sélectivité vis-à-vis de PRMT1. Le produit **18b** inhibe l'enzyme CARM1 avec une valeur de CI_{50} de 115 μM avec un rapport de sélectivité (RS) égale à 7. L'analogue **19b**, qui incorpore dans sa structure une fonction imidate, inhibe CARM1 avec une valeur de CI_{50} de 20 μM et un rapport de sélectivité égale à 40. Finalement, l'analogue **23b** présente une valeur d'inhibition pour CARM1 de 43 μM et un rapport de sélectivité égale à 7. Cette sélectivité a

pu être expliquée de façon rationnelle grâce à la structure cristallographique obtenue avec l'analogue **23b**. Cette structure nous a également apporté des informations supplémentaires sur certains points d'ancrage du site actif. Ces premiers résultats nous ont permis de valider cette étape; nous avons décidé de poursuivre davantage en explorant le « poche de l'arginine » grâce à des analogues de bisubstrats.

Dans une seconde partie, la « poche de l'arginine » a pu être explorée grâce aux structures RX obtenues avec les analogues de bisubstrats synthétisés **45** et **51**. Ces structures nous ont permis d'avoir une idée précise des interactions qui peuvent avoir lieu à l'état de transition entre l'arginine et les résidus du site actif de CARM1. Nous avons montré que ces interactions sont gouvernées par un environnement électrostatique et stérique imposé par les différents acides aminés du site catalytique. Un modèle sur le mécanisme de transfert de méthyle de la réaction catalysée par l'enzyme CARM1 a été proposé et les structures obtenues nous ont permis de valider notre modèle. Finalement, ces analogues de bisubstrats ont été soumis à des tests d'activité qui ont montré de très bonnes inhibitions contre CARM1 ; l'analogue **45** est ainsi 54 fois plus sélectif par rapport à PRMT1. Ces analogues nous ont donc permis d'explorer la « poche de l'arginine » et de valider la seconde étape de notre stratégie. Nous avons donc décidé de continuer notre étude en sortant de cette poche et en explorant le « domaine de fixation du peptide ».

Deux stratégies ont été employées, la première stratégie employée pour cette exploration est une approche par réaction *in-situ* de molécules hameçons capables de piéger un peptide mimant la séquence du substrat. Cette approche s'est avérée infructueuse car les techniques employées n'ont jamais mis en évidence la présence d'éventuels adduits, ni par co-cristallisation, ni par spectrométrie de masse. Nous avons suspecté un manque d'affinité de

nos molécules, c'est pourquoi une molécule hameçon électrophile guanidine **90** a été synthétisée puis testée en co-cristallisation avec des peptides cystéines. Malheureusement, aucun adduit n'a été observé sur la structure RX obtenue, seul une partie de la molécule **90** a été observée. Ces résultats montrent que, malgré la capacité de la molécule hameçon **90** à se lier à CARM1, il est nécessaire que le peptide soit aussi très affiné avec l'enzyme. Des analyses de spectrométrie de masse ont montré que le peptide **PK110** se fixait seulement sur un site du tétramère de CARM1. C'est pourquoi des études sont maintenant réalisées par nos collaborateurs pour déterminer quel peptide présente une haute affinité avec CARM1. Des analyses par marquage impliquant des anticorps de la SAM pourraient être envisagées pour étudier le transfert de molécule hameçon électrophiles sur des peptides.

Aux vues de ces résultats, une nouvelle stratégie a été élaborée : elle a consisté à préformer des adduits chimiquement puis à essayer de les co-cristalliser dans l'enzyme. Dans cette dernière partie, des adduits SAM-peptide ont été synthétisés pour pouvoir explorer le « domaine de fixation du peptide ». Le potentiel d'inhibition de chaque composé a été mesuré contre CARM1 et il s'est avéré que le composé **106** a présenté la meilleure inhibition avec une valeur de CI_{50} de 13 μM ; de même, l'adduit **110** a également une bonne valeur de CI_{50} de 56 μM . Les tests de co-cristallisation de ces adduits avec mCARM1₁₃₀₋₄₉₀ ont été infructueux car les structures obtenues n'ont pas montré le composé entièrement dans le site actif de l'enzyme. Le seul adduit qui a été visualisé presque complètement est le dérivé **104** : la structure est visible jusqu'au carbone en α de l'acide aminé.

La seconde approche a consisté à réaliser des essais de co-cristallisation dans CARM1 avec l'adduit **115** formé à partir d'une molécule hameçon électrophile **112** et la cystéine modifiée **74**. La longue chaîne a servi à émerger de la poche de l'arginine afin de sortir vers le

« domaine de fixation du peptide ». Malheureusement, cet adduit a provoqué la précipitation de la protéine et aucun cristal n'a été ainsi obtenu.

Pour la suite du projet, il faudra continuer dans cette voie en optimisant la synthèse de ces adduits pour ensuite les tester en co-cristallisation. Une étape d'optimisation est nécessaire pour les deux classes d'adduits suivants :

- Pour les adduits obtenus par réaction de click, il faudra utiliser des alcynes plus longs afin d'émerger de la « poche du domaine de fixation » et obtenir une structure cristallographique complète.
- Pour les adduits obtenus par réaction de substitution nucléophile entre une molécule hameçon électrophile et une cystéine modifiée, il faudra modifier la nature de la longueur de la chaîne pour éviter le phénomène de précipitation.

Cette optimisation pourra se faire soit en remplaçant la chaîne carbonée par un PEG soit par une spermine ou encore en faisant réagir la molécule hameçon **90** avec une cystéine modifiée mimant la séquence du peptide **PK110 (P-SH)** (figure 1).

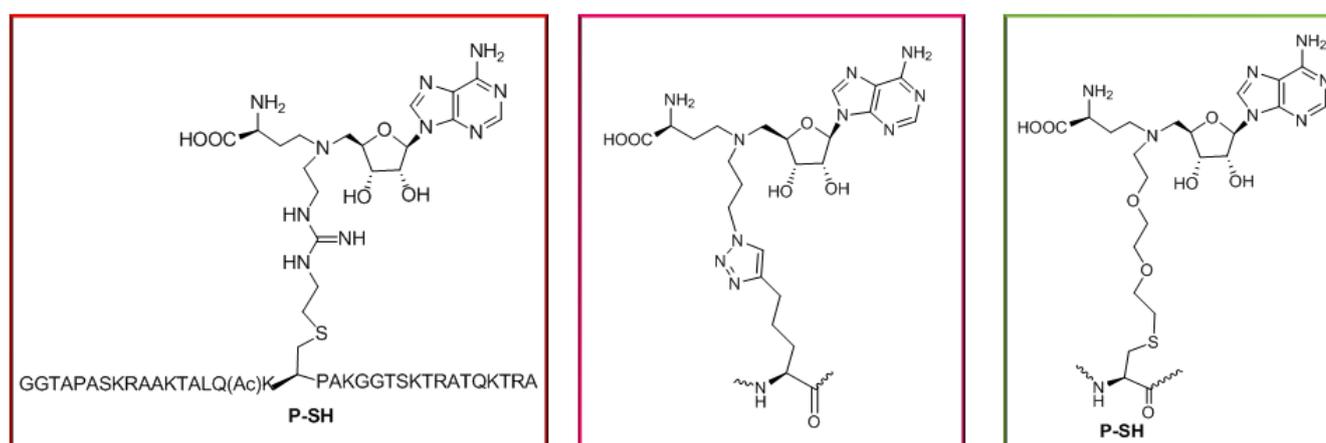


Figure 1 : Adduits SAM-peptide, de seconde génération, envisagés pour les tests de co-cristallisation dans CARM1.

Cette dernière réaction nous permettra d'avoir un adduit SAM-peptide qui se rapprocherait en même temps de l'état de transition et pourrait interagir avec les points d'ancrage du site actif qui sont nécessaires à la stabilisation d'un tel adduit dans CARM1.

Notre stratégie utilisée pour étudier CARM1 pourrait être utilisée pour d'autres PRMTs (PRMT1, PRMT5, PRMT6 et PRMT7).

Nous pourrions également mesurer le potentiel d'inhibition de nos analogues de la SAM au sein d'autres méthyltransférases.

Partie expérimentale

Partie expérimentale

I. General procedure

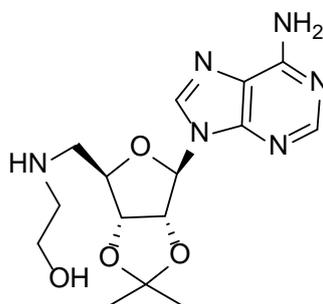
All chemicals and solvents were of reagent grade and were used without further purification. Unless indicated, reactions were carried out under an atmosphere of argon in flame-dried glassware with magnetic stirring. The progress of all reactions and column chromatography were monitored by Thin-layer chromatography (TLC). TLC was performed in using Merck silica gel 60 F₂₅₄ precoated plates and visualized by UV absorption or stained with a solution of ninhydrin in EtOH. Flash chromatography ARMEN was performed using Merck silica gel 45 (230-400 mesh) with an appropriate eluent. ¹H and ¹³C NMR spectra were recorded on a Bruker Avance DPX 300 Spectrometer (operated at 300 or 75 MHz, respectively) or ultrashield plus 400 Spectrometer (operated at 400 or 100 MHz, respectively) or ultrashield plus 500 Spectrometer (operated at 500 or 125 MHz, respectively). The ¹H NMR chemical shifts are reported in parts per million, using the signal for residual solvent protons (7.24 for CDCl₃, 3.31 for MeOD, 4.78 for D₂O) as reference. The ¹³C NMR chemical shifts are reported in parts per million, using the signal for residual solvent protons as reference (77.23 for CDCl₃, 48.95 and for MeOD, for D₂O: 1 μL of MeOH is used as internal reference). Data were represented as follows, chemical shift, multiplicity (s = singlet, bs = broad singlet, d = doublet, dd = doublet of doublets, t = triplet, q = quadruplet, qu = quintuplet, s = sextuplet, sp = septuplet and m = multiplet), integration and coupling (J in Hz). LC-MS data were obtained using a LC/MSD/SL Agilent Technologies; ionising source is APCI/ESI, Agilent technologies 1200 autosampler; pump agilent 1200 binary system; column ThermoScientific (58282): 30 x 1mm, 19 μm, hypersil gold C18, Agilent Technology 1200 DAD detector; Agilent technology MS quadrupole mass spectrometer operating in APCI/ESI mode. InfraRed spectra were obtained neat in the solid phase or from solution in CHCl₃ using a recorded on a Nicolet 380 FT-IR from Thermo Electron Corporation. The HPLC analyses were performed using a Shimadzu LC-20 AD system, equipped with a SPD 20A absorption detector. The HPLC system was as follows:

- **Method** : Waters Sunfire C18; 5 μ ; 4.6x50 mm column; 1 mL/minutes; eluent B 0 % over 5 minutes, then 0 to 100% over 10 minutes. Eluent A = H₂O, 0.01% TFA and eluent B = CH₃CN. The column was maintained at a temperature of 30°C.
- Purities of final compounds were in all cases greater than 98%, as determined by reverse-phase HPLC analysis.

II. Chemistry

Preparation of 2-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)ethanol **31**

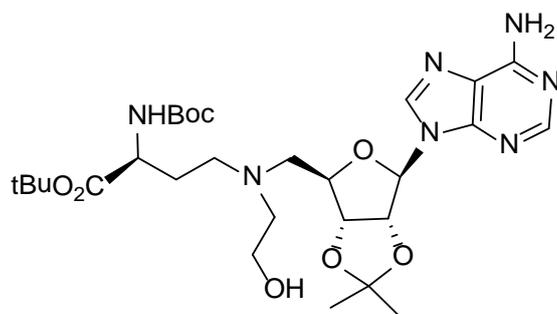
Amine **10a** (3 g, 9 mmol, 1 eq) was dissolved in anhydrous tetrahydrofuran (88 mL) and stirred at room temperature, then bromoethanol (1.3 mL, 18 mmol, 2 eq) was dropwise added to the solution. Finally, N,N-diisopropylethylamine (4.1 mL, 13 mmol, 2 eq) was dropwise added and the reaction mixture was stirred under reflux overnight. The reaction mixture was evaporated and the crude was purified by flash chromatography using a dichloromethane/methanol gradient (85/15 to 70/30) to give a yellow solid (1.4 g, 45%).



¹H NMR (CDCl₃, 400MHz): δ (ppm) = 1.32 (s, 3H); 1.55 (s, 3H); 2.67-2.77 (m, 2H); 2.89-2.91 (m, 2H); 3.61 (t, J = 4.8 Hz, 2H); 4.34 (bs, 1H); 4.97 (dd, J_1 = 3.6 Hz, J_2 = 6.4 Hz, 1H); 5.41 (dd, J_1 = 3.6 Hz, J_2 = 6.4 Hz, 1H); 5.96 (dd, J = 2.8 Hz, 1H); 6.49 (bs, 2H); 7.89 (s, 1H); 8.23 (s, 1H). ¹³C NMR (CDCl₃, 100 MHz): δ (ppm) = 25.5; 27.4; 50.5; 51.1.1; 51.4; 60.9; 82.4; 83.7; 85.7; 91.0; 114.7; 120.2; 140.0; 149.3; 153.3; 155.9. IR ν (cm⁻¹): Csp₂-H (703-868); C-O (1079); C-N (1298); C=N (1477); C=O (1713); NH₂ (2854-2977). LC-MS (ESI-APCI): m/z [M+H]⁺ = 352.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(2-hydroxyethyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **32**

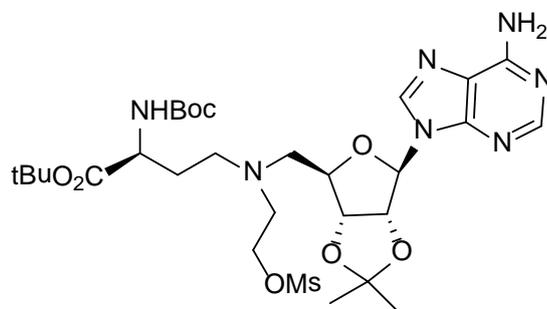
Compound **31** (1.4 g, 3.9 mmol, 1.2 eq) was dissolved in anhydrous methanol (40 mL) and stirred at room temperature. To the solution, aldehyde **3** (900 mg, 3.3 mmol, 1.0 eq) dissolved in anhydrous methanol (16 mL) was dropwise added at room temperature. Then HOAc (200 μ L, 0.71 mmol, 1 eq) was dropwise added, and NaBH₃CN (322 mg, 5.3 mmol, 1.6 eq) was added in one portion. The reaction mixture was stirred at room temperature overnight. Then the reaction mixture was diluted in ethyl acetate (50 mL) and washed with a solution of NaOH 1N (2 x 20 mL), the aqueous layer was extracted with dichloromethane (2 x 20 mL) dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a yellow solid 1.2 g, 60%.



¹H NMR (CDCl₃, 400MHz): δ (ppm) = 1.30 (s, 3H); 1.35 (s, 18H); 1.52 (s, 3H); 1.59-1.68 (m, 1H); 1.85-1.96 (m, 1H); 2.45-2.62 (m, 4H); 2.74-2.79 (m, 2H); 3.26-3.39 (m, 1H); 3.46-3.51 (m, 1H); 3.97-4.09 (m, 1H); 4.24 (bs, 1H); 4.98 (bs, 1H); 5.38 (d, J = 4.4, 1H); 5.98 (s, 1H); 6.51 (bs, NH₂, 2H); 7.86 (s, 1H); 8.23 (s, 1H). ¹³C NMR (CDCl₃, 100 MHz): δ (ppm) = 25.6; 27.2; 28.1; 28.4; 29.8; 50.9; 52.7; 56.3; 57.2; 59.4; 79.7; 81.9; 83.4; 84.0; 85.4; 90.5; 114.6; 120.3; 140.1; 149.1; 153.2; 155.7; 156.1; 172.1. IR ν (cm⁻¹): Csp₂-H (713-909); C-O (1151); C-N (1366); C=N (1598); C=O (1705); NH (2977). LC-MS (ESI-APCI): m/z [M+H]⁺ = 609.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(2-((methylsulfonyl)oxy)ethyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **33**

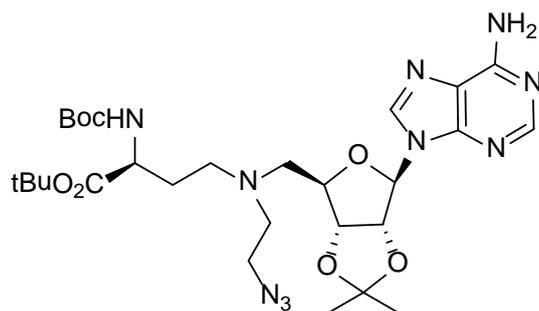
Compound **32** (1 g, 1.6 mmol, 1eq) was dissolved in anhydrous tetrahydrofuran (9 mL) and stirred at room temperature. To the solution, methanesulfonyl chloride (127 μ L, 1.6 mmol, 1 eq) and N,N-diisopropylethylamine (283 μ L, 1.6 mmol, 1 eq) were dropwise added successively. The reaction mixture was stirred at room temperature overnight. Then the reaction mixture was diluted in ethyl acetate (20 mL) and washed with a solution of NaOH 1N (2 x 10 mL). The aqueous layer was extracted with dichloromethane (2 x 10 mL), dried over Na₂SO₄, filtered and evaporated. The crude (1.1 g) was used for the next step.



¹H NMR (CDCl₃, 400MHz): δ (ppm) = 1.35 (s, 3H); 1.46 (s, 18H); 1.55 (s, 3H); 1.66-1.86 (m, 2H); 2.36-2.66 (m, 3H); 2.68 (s, 3H); 2.70-2.76 (m, 2H); 2.88-2.99 (m, 1H); 3.34 (t, J = 4 Hz, 1H); 3.41-3.52 (m, 1H); 4.09 (bs, 1H); 4.14-4.19 (m, 1H); 4.91 (dd, J₁ = 4 Hz, J₂ = 6 Hz, 1H); 5.38 (d, J = 5.6 Hz, 1H); 5.67 (bs, NH, 1H); 5.94 (s, 1H); 6.61 (bs, NH₂, 2H); 7.85 (s, 1H); 8.16 (s, 1H). ¹³C NMR (CDCl₃, 100 MHz): δ (ppm) = 25.3; 27.0; 27.9; 28.3; 29.4; 39.7; 51.0; 52.2; 53.4; 56.3; 60.2; 79.2; 81.5; 83.1; 83.6; 85.5; 90.7; 114.2; 120.1; 140.2; 149.9; 153.0; 155.5; 155.9; 171.9.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(2-azidoethyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **34**

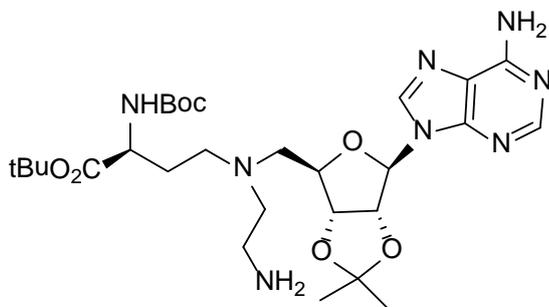
Compound **33** (1.1g, 1.3 mmol, 1.0 eq) was dissolved in anhydrous acetonitrile (30 mL) and stirred at room temperature. To the solution, sodium azide (151 mg, 2.4 mmol, 1.5 eq) was added in one portion. The reaction mixture was stirred under reflux for 2 days. The reaction mixture was evaporated and purified by flash chromatography using a dichloromethane/methanol gradient (98/2 to 95/5) to give a yellow solid (500 mg, 50%).



¹H NMR (CDCl₃, 400MHz): δ (ppm) = 1.36 (s, 3H); 1.40 (s, 18H); 1.58 (s, 3H); 1.63-1.74 (m, 1H); 1.87-1.96 (m, 1H); 2.47-2.85 (m, 6H); 3.16-3.27 (m, 1H); 3.44 (t, *J* = 8, 1H); 4.15 (bs, 1H); 4.27-4.32 (m, 1H); 4.99 (bs, 1H); 5.47 (bs, 1H); 5.92 (bs, 2H); 6.01 (d, *J* = 4Hz, 1H); 7.89 (s, 1H); 8.30 (s, 1H). **¹³C NMR** (CDCl₃, 100 MHz): δ (ppm) = 25.6; 27.3; 28.1; 28.5; 29.8; 49.4; 50.8; 51.7; 52.7; 56.3; 79.6; 81.8; 83.6; 84.2; 85.8; 91.1; 114.6; 120.5; 140.4; 149.4; 153.3; 155.7; 155.9; 171.9. **IR** ν(cm⁻¹): Csp₂-H (725-905); C-O (1153); C-N (1367); C=O (1705); C=N (1642); N₃ (2101); NH(2980); NH₂ (3326). **LC-MS** (ESI-APCI): *m/z* [M+H]⁺ = 634.

Preparation of tert-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(2-aminoethyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate **35**

Compound **34** (450 mg, 0.71mmol, 1 eq) was dissolved in a mixture of tetrahydrofuran / water (2.7 mL/450 μL) and stirred at room temperature. To the solution, triphenylphosphine (373 mg, 1.42 mmol, 2 eq) was added in one portion and the reaction mixture was stirred under reflux overnight. Then the reaction mixture was cooled and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol (50/50) as eluent to give a yellow solid (176 mg, 60%).

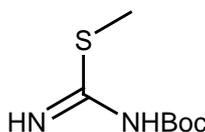


¹H NMR (CDCl₃, 400MHz): δ (ppm) = 1.33 (s, 3H); 1.36 (s, 9H); 1.46 (s, 9H); 1.58 (s, 3H); 1.69-1.77 (m, 1H); 1.90-1.98 (m, 1H); 2.68-2.97 (m, 8H); 4.21 (bs, 1H); 4.35 (bs, 1H); 4.80 (t,

$J = 4.4$ Hz, 1H); 5.20 (d, $J = 4.8$ Hz, 1H); 5.39 (d, $J = 6.4$ Hz, 1H); 6.09 (s, 1H); 8.22 (s, 1H); 8.41 (s, 1H). ^{13}C NMR (CDCl_3 , 100 MHz): δ (ppm) = 25.4; 27.1; 28.1; 28.4; 29.9; 43.2; 51.5; 51.7; 51.8; 56.1; 80.8; 81.9; 83.2; 84.1; 89.9; 90.9; 114.8; 119.9; 143.3; 148.1; 151.9; 155.7; 156.1; 171.3. IR $\nu(\text{cm}^{-1})$: $\text{Csp}_2\text{-H}$ (725-905); C-O (1153); C-N (1367); C=O (1705); C=N (1642); NH (2980); NH_2 (3326). LC-MS (ESI-APCI): m/z $[\text{M}+\text{H}]^+ = 607$.

Preparation of N, (Boc)-S-methyl-iso-thiourea **37**

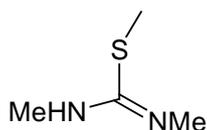
S-methyl-iso-thiourea **36** (1.0 g, 3.6 mmol, 1.0 eq) was dissolved in water (11 mL) and stirred at room temperature. To the solution, di-*tert*-butyl dicarbonate (1.0 g, 4.7 mmol, 1.3 eq) dissolved in anhydrous dioxane (11 mL) was dropwise added to the solution at room temperature. Then a solution of NaOH 1N (7 mL) was dropwise added to the reaction mixture and stirred at room temperature overnight. The reaction mixture was filtered and washed with water (50 mL). The filtrate was extracted with dichloromethane (3 x 30 mL). The organic layer was dried over Na_2SO_4 , filtered and evaporated. The resulting crude was a white oil (543 mg, 80%) and was used directly in the next step.



^1H NMR (CDCl_3 , 300MHz): δ (ppm) = 1.40 (s, 9H); 2.40 (s, 3H). ^{13}C NMR (CDCl_3 , 75 MHz) : δ (ppm) = 14.1; 28.2; 80.1; 155.5; 170.1.

Preparation of S-1,2,3-trimethyl-iso-thiourea **39**

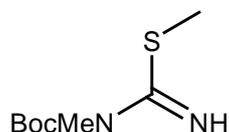
1,3-dimethyl-thiourea **38** (1.0 g, 9.7 mmol, 1.0 eq) was dissolved in anhydrous tetrahydrofuran (100 mL) and stirred at room temperature. To the solution, K_2CO_3 (2.0 g, 14 mmol, 1.5 eq) and methyl iodide (0.7 mL, 12 mmol, 1.2 eq) were successively added to the solution at 0°C . The reaction mixture was stirred at 0°C for 2h and at room temperature for 3h. The reaction mixture was filtered and neutralised with HOAc (3 mL). The aqueous layer was extracted with AcOEt (2 x 100 mL), and the organic layer was dried over Na_2SO_4 , filtered and evaporated. The resulting crude was a white solid (1.1 g, 100%) and was used directly in the next step.



¹H NMR (D₂O, 400MHz): δ (ppm) = 2.40 (s, 3H) ; 2.98 (d, J = 16 Hz, 6H). **¹³C NMR** (D₂O, 400MHz) : δ (ppm) = 13.3; 30.3; 169.2.

Preparation of N(Boc)-S-1, 2-dimethyl-iso-thiourea **41**

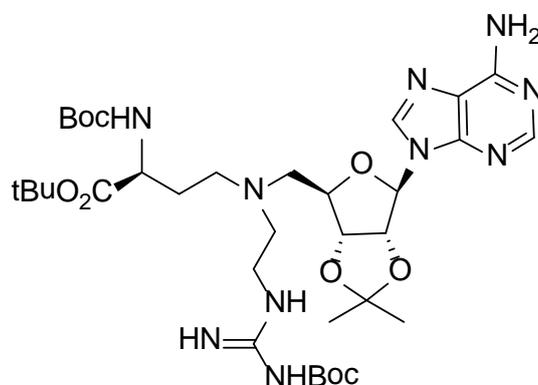
1,3-diméthyl-thiourea **40** (2.0 g, 22 mmol, 1.0 eq) was dissolved in anhydrous tetrahydrofuran (228 mL) and stirred at room temperature. To the solution, K₂CO₃ (4.6 g, 33 mmol, 1.5 eq) and dimethylsulfate (2.5 mL, 26 mmol, 1.2 eq) were successively added to the solution at 0°C. The reaction mixture was stirred at 0°C for 2h and at room temperature overnight. The reaction mixture was filtered and washed with water (50 mL). The organic layer was evaporated. The resulting crude was a white oil (1.6 g, 70%) and was used directly in the next step. The crude (1.6 g, 15.1 mmol, 1.0 eq) was dissolved in water (16 mL) and stirred at room temperature. To the solution, di-*tert*-butyl dicarbonate (4.3 g, 19.6 mmol, 1.3 eq) dissolved in anhydrous dioxane (16 mL) was dropwise added to the solution at room temperature. Then a solution of NaOH 1N (10.5 mL) was dropwise added to the reaction mixture and stirred at room temperature overnight. The reaction mixture was filtered and washed with water (100 mL). The filtrate was extracted with dichloromethane (3 x 30 mL). The organic layer was dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a cyclohexane/ethyl acetate gradient (95/5) to give a white oil (1.3 g, 78%).



¹H NMR (CDCl₃, 300MHz) : δ (ppm) = 1.50 (s, 9H) ; 2.38 (s, 3H) ; 3.14 (s, 3H). **¹³C NMR** (CDCl₃, 75 MHz) : δ (ppm) = 14.7; 28.1; 30.2; 82.5; 153.5; 163.4.

Preparation of *tert*-butyl 10-((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)-13-((*tert*-butoxycarbonyl)amino)-6-imino-2,2-dimethyl-4-oxo-3-oxa-5,7,10-triazatetradecan-14-oate **42**

Compound **35** (20 mg, 0.033 mmol, 1.0 eq) was dissolved in anhydrous acetonitrile (250 μ L) and stirred at room temperature. To the solution, thiourea **37** (44 mg, 0.099 mmol, 3.0 eq) dissolved in anhydrous acetonitrile (250 μ L) was dropwise added to the solution at room temperature. Then the reaction mixture was heated under reflux for 48h. The reaction mixture was cooled at room temperature and evaporated. The crude was purified by reverse phase flash chromatography using water/methanol gradient (10/90) to give a yellow solid (15 mg, 60 %).

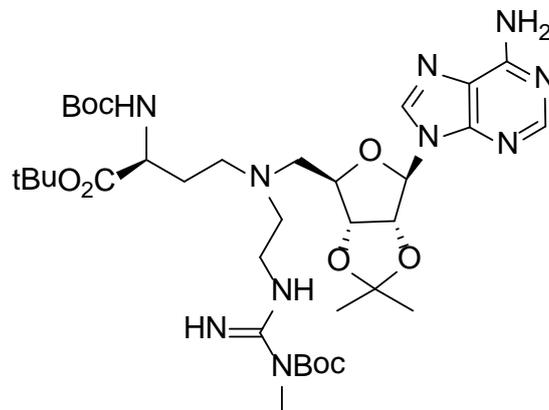


$^1\text{H NMR}$ (MeOD, 300MHz): δ (ppm) = 1.38 (s, 3H); 1.43 (s, 18H); 1.45 (s, 9H); 1.62 (s, 3H); 1.72-1.85 (m, 1H); 2.02-2.16 (m, 1H); 2.83-3.15 (m, 4H); 3.41-3.42 (m, 4H); 4.07 (bs, 1H); 4.53 (q, $J = 4.8$ Hz, 1H); 5.08 (dd, $J_1 = 4.8$ Hz, $J_2 = 6$ Hz, 1H); 5.36 (dd, $J_1 = 4.8$ Hz, $J_2 = 6$ Hz, 1H); 6.31 (s, 1H); 8.42 (d, $J = 3.6$ Hz, 1H). $^{13}\text{C NMR}$ (CDCl_3 , 100 MHz): δ (ppm) = 25.4; 27.1; 28.1; 28.4; 29.9; 36.9; 51.5; 51.7; 51.8; 56.1; 79.7; 80.8; 81.9; 83.2; 84.1; 89.9; 114.8; 119.9; 143.3; 148.1; 151.9; 156.1; 162.6. 171.3. **LC-MS** (ESI-APCI): m/z $[\text{M-Boc}]^+ = 649$.

Preparation of *tert*-butyl 10-((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)-13-((*tert*-butoxycarbonyl)amino)-6-imino-2,2,5-trimethyl-4-oxo-3-oxa-5,7,10-triazatetradecan-14-oate **43**

Compound **35** (25 mg, 0.041 mmol, 1.0 eq) was dissolved in anhydrous acetonitrile (325 μ L) and stirred at room temperature. To the solution, thiourea **41** (37 mg, 0.123 mmol, 3.0 eq) dissolved in anhydrous acetonitrile (300 μ L) was dropwise added to the solution at room temperature. Then the reaction mixture was heated under reflux for 48h. The reaction

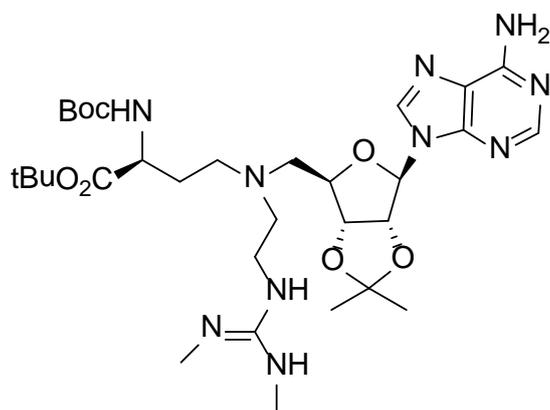
mixture was cooled at room temperature and evaporated. The crude was purified by reverse phase flash chromatography using water/methanol gradient (10/90) to give a yellow solid (12 mg, 40 %).



$^1\text{H NMR}$ (CDCl_3 , 300MHz): δ (ppm) = 1.36 (s, 3H); 1.41 (s, 18H); 1.45 (s, 9H); 1.56 (s, 3H); 2.52-2.72 (m, 6H); 2.77-2.79 (m, 2H); 3.39 (s, 3H); 3.67-3.74 (m, 2H); 4.21 (bs, 1H); 4.29 (bs, 1H); 4.98 (bs, 1H); 5.13 (bs, NH, 1H); 5.51 (d, $J = 5.4$ Hz, 1H); 5.84 (bs, NH_2 , 2H); 6.07 (s, 1H); 8.04 (s, 1H); 8.30 (s, 1H). $^{13}\text{C NMR}$ (CDCl_3 , 100 MHz): δ (ppm) = 25.7; 27.4; 28.0; 28.4; 28.6; 29.5; 30.5; 34.6; 50.9; 52.9; 56.3; 79.8; 81.9; 83.5; 84.1; 85.6; 91.1; 114.7; 120.6; 140.4; 149.5; 153.3; 155.7; 160; 171.9. **LC-MS** (ESI-APCI): m/z $[\text{M}-\text{Boc}]^+ = 663$.

Preparation of (E)-tert-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(2-(2,3-dimethylguanidino)ethyl)amino)-2-((tert-butoxycarbonyl)amino)butanoate **44**

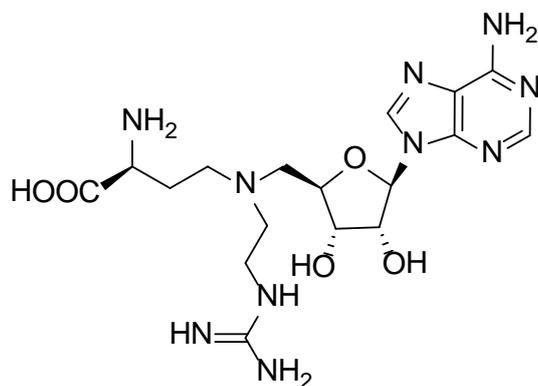
Compound **35** (50 mg, 0.082 mmol, 1.0 eq) was dissolved in anhydrous acetonitrile (1 mL) and stirred at room temperature. To the solution, thiourea **39** (29 mg, 0.241 mmol, 3.0 eq) dissolved in anhydrous acetonitrile (500 μL) was dropwise added to the solution at room temperature. Then the reaction mixture was heated under reflux for 48h. The reaction mixture was cooled at room temperature and evaporated. The crude was purified by reverse phase flash chromatography using water/methanol gradient (10/90) to give a yellow solid (20 mg, 36 %).



¹H NMR (CDCl₃, 400MHz): δ (ppm) = 1.34 (s, 3H); 1.39 (s, 9H); 1.44 (s, 9H); 1.58 (s, 3H); 1.95 (bs, 1H); 2.59-2.66 (m, 4H); 2.67 (d, *J* = 4.8 Hz, 6H); 2.72-3.10 (m, 5H); 4.21 (bs, 1H); 4.33 (bs, 1H); 4.89 (dd, *J*₁ = 3.2 Hz, *J*₂ = 6.8 Hz, 1H); 5.23 (bs, NH, 1H); 5.40 (dd, *J*₁ = 2.8 Hz, *J*₂ = 6.4 Hz, 1H); 6.08 (s, 1H); 7.62 (bs, NH, 3H); 8.16 (s, 1H); 8.25 (s, 1H). **¹³C NMR** (CDCl₃, 100 MHz): δ (ppm) = 25.2; 27.1; 28.0; 28.3; 29.5; 30.1; 39.2; 50.6; 51.9; 54.6; 56.3; 80.2; 82.5; 82.9; 83.9; 84.6; 91.1; 114.8; 120.0; 141.2; 148.5; 150.7; 155.7; 156.3; 162.5; 171.6. **LC-MS** (ESI-APCI): *m/z*[M+H]⁺ = 677.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(2-guanidinoethyl)amino)butanoic acid **45**

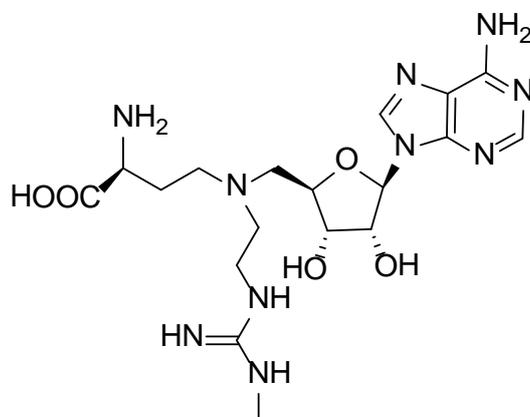
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **42** (10 mg, 0.013 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **45** was obtained as a yellowish sticky solid (6mg, 98%).



¹H NMR (D₂O, 500MHz): δ (ppm) = 2.17-2.21 (m, 1H); 2.31-2.42 (m, 1H); 3.40-3.47 (m, 1H); 3.56-3.62 (m, 6H); 3.76 (d, *J* = 4.2 Hz, 2H); 3.90 (d, *J* = 6 Hz, 1H); 4.50 (bs, 2H); 6.16 (s, 1H); 8.44 (s, 2H). **¹³C NMR** (D₂O, 125 MHz): δ (ppm) = 24.7; 36.2; 53.1; 53.7; 55.6; 71.6; 72.7; 78.8; 89.8; 119.3; 142.7; 148.2 ; 151.5; 156.7; 167.9; 172.9. **HRMS** (m/z) : [M-H]⁻ = 451.21659, Found: 451.21558.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(2-(3-methylguanidino)ethyl)amino)butanoic acid **46** :

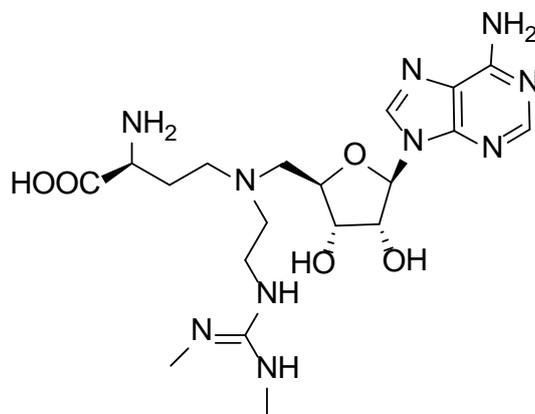
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **43** (10 mg, 0.013 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **46** was obtained as a yellowish sticky solid (6mg, 98%).



¹H NMR (D₂O, 400MHz): δ (ppm) = 2.17-2.21 (m, 1H); 2.31-2.42 (m, 1H); 2.81 (s, 3H); 3.22-3.31 (m, 1H); 3.50-3.52 (m, 1H); 3.52-3.58 (m, 4H); 3.65-3.76 (m, 3H); 3.94-3.96 (m, 1H) 4.40-4.42 (m, 1H); 4.48-4.50 (m, 1H); 6.17 (s, 1H); 8.46 (s, 2H). **¹³C NMR** (MeOD, 100 MHz): δ (ppm) = 23.9; 33.2; 36.0; 52.4; 54.9; 56.6; 82.7; 84.7; 86.8; 91.9; 123.7; 142.5; 151.1 ; 154.4; 157.7; 158.2; 174.2. **HRMS** (m/z) : [M+H]⁺ = 467.2467, Found: 467.2517.

Preparation of (E)-2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(2-(2,3-dimethylguanidino)ethyl)amino)butanoic acid **47**

HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **44** (10 mg, 0.015 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **47** was obtained as a yellowish sticky solid (7 mg, 98%).

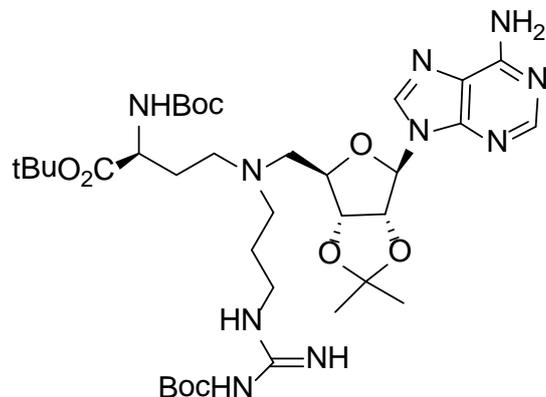


¹H NMR (D₂O, 400MHz): δ (ppm) = 2.19-2.24 (m, 1H); 2.33-2.42 (m, 1H); 2.71 (s, 5H); 2.82 (s, 1H); 3.42-3.49 (m, 2H); 3.56-3.61 (m, 5H); 3.76-3.78 (m, 2H); 3.92-3.96 (m, 1H); 4.45-4.52 (m, 2H); 6.15 (s, 1H); 8.45 (s, 2H). **¹³C NMR** (D₂O 100 MHz): δ (ppm) = 24.9; 27.7; 36.3; 52.6; 53.4; 54.2; 55.8; 71.8; 73.2; 78.9; 90.4; 119.5; 143.7; 148.3 ; 150.3; 156.0; 172.5; 175.9. **LC-MS** (ESI-APCI): m/z [M+H]⁺ = 481 (\geq 98%). **HRMS** (m/z) : [M+H]⁺ = 481.2629, Found: 481.2726.

Preparation of *tert*-butyl 11-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)-14-((*tert*-butoxycarbonyl)amino)-6-imino-2,2-dimethyl-4-oxo-3-oxa-5,7,11-triazapentadecan-15-oate **48**

Compound **16c** (60 mg, 0.097 mmol, 1.0 eq) was dissolved in anhydrous acetonitrile (1 mL) and stirred at room temperature. To the solution, thiourea **37** (84 mg, 0.291 mmol, 3.0 eq) dissolved in anhydrous acetonitrile (800 μ L) was dropwise added to the solution at room temperature. Then the reaction mixture was heated under reflux for 48h. The reaction

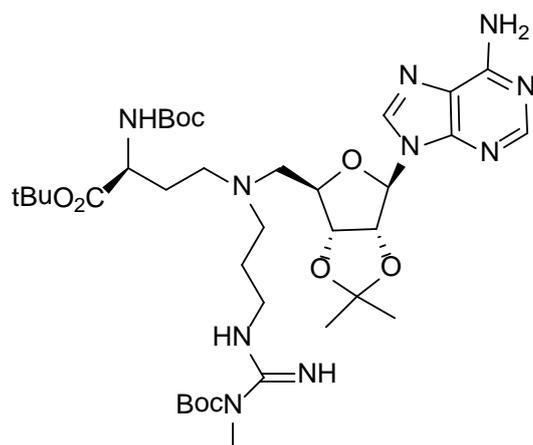
mixture was cooled at room temperature and evaporated. The crude was purified by reverse phase flash chromatographie using water/MeOH gradient (10/90) to give a yellow solid (33 mg, 40 %).



$^1\text{H NMR}$ (MeOD, 400MHz): δ (ppm) = 1.38 (s, 3H); 1.43 (s, 18H); 1.55 (s, 9H); 1.63 (s, 3H); 1.86-2.03 (m, 3H); 2.11-2.20 (m, 2H); 3.18-3.33 (m, 5H); 3.66-3.83 (m, 2H); 3.92-3.95 (m, 1H); 4.64-4.68 (m, 1H); 5.15 (dd, $J_1 = 4.4$ Hz ; $J_2 = 6.4$ Hz, 1H); 5.36 (dd, $J_1 = 4.4$ Hz ; $J_2 = 6.4$ Hz, 1H) ; 6.03(s, 1H); 8.40 (s, 2H). $^{13}\text{C NMR}$ (MeOD, 100 MHz): δ (ppm) = 25.6; 27.0; 27.6; 28.3; 28.4; 28.9; 39.9; 52.5; 53.5; 56.4; 81.9; 83.6; 84.2; 85.6; 86.2; 92.3; 116.5; 121.3; 140.4; 143.8; 149.6; 153.6; 155.4; 160.2; 172.0. **LC-MS** (ESI-APCI): m/z $[\text{M-Boc}]^+ = 663$.

Preparation of *tert*-butyl 11-((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)-14-((*tert*-butoxycarbonyl)amino)-6-imino-2,2,5-trimethyl-4-oxo-3-oxa-5,7,11-triazapentadecan-15-oate **49**

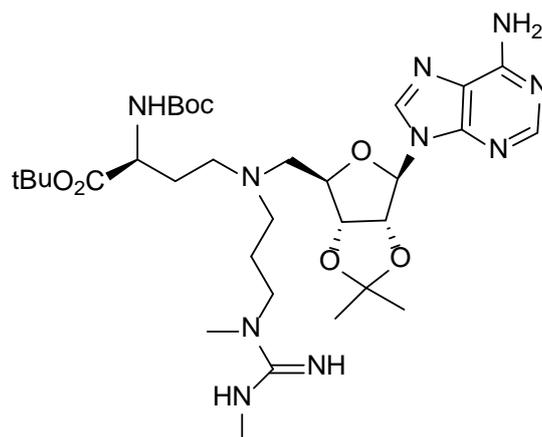
Compound **16c** (66 mg, 0.106 mmol, 1.0 eq) was dissolved in anhydrous acetonitrile (2 mL) and stirred at room temperature. To the solution, thiourea **41** (65mg, 0.318 mmol, 3.0 eq) dissolved in anhydrous acetonitrile (650 μL) was dropwise added to the solution at room temperature. Then the reaction mixture was heated under reflux for 48h. The reaction mixture was cooled at room temperature and evaporated. The crude was purified by reverse phase flash chromatographie using water/MeOH gradient (10/90) to give a yellow solid (39 mg, 50 %).



$^1\text{H NMR}$ (MeOD, 300MHz): δ (ppm) = 1.39 (s, 3H); 1.43 (s, 18H); 1.44 (s, 9H); 1.60 (s, 3H); 1.65-1.72 (m, 2H); 1.85-2.0 (m, 2H); 2.44-2.48 (m, 1H); 2.60-2.63 (m, 4H); 2.81-2.89 (m, 2H); 3.39 (s, 3H); 3.55-3.67 (m, 1H); 4.02 (bs, 1H); 4.32 (bs, 1H); 5.05 (bs, 1H); 5.49 (s, 1H); 6.17 (s, 1H); 8.23 (d, $J = 15$ Hz, 2H). $^{13}\text{C NMR}$ (MeOD, 100 MHz): δ (ppm) = 25.8; 26.0; 27.6; 28.4; 28.8; 29.9; 30.7; 33.1; 34.9; 52.2; 54.3; 57.4; 80.5; 81.1; 82.6; 84.7; 85.1; 86.7; 91.8; 115.4; 120.8; 142.1; 150.2; 154.0; 157.4; 158.0; 173.6. **LC-MS** (ESI-APCI): m/z $[\text{M-Boc}]^+ = 677$.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(3-(1,3-dimethylguanidino)propyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **50**

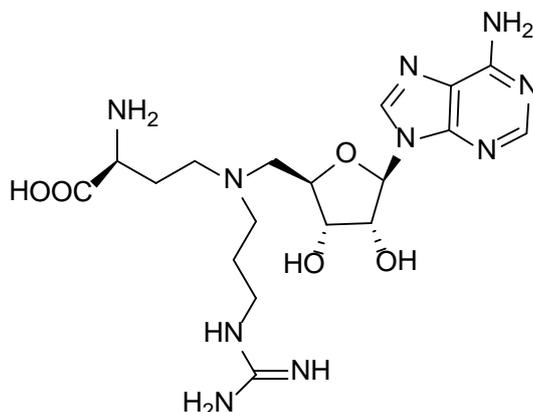
Compound **16c** (50 mg, 0.081 mmol, 1.0 eq) was dissolved in anhydrous acetonitrile (1 mL) and stirred at room temperature. To the solution, thiourea **39** (29 mg, 0.242 mmol, 3.0 eq) dissolved in anhydrous acetonitrile (500 μL) was dropwise added to the solution at room temperature. Then the reaction mixture was heated under reflux for 48h. The reaction mixture was cooled at room temperature and evaporated. The crude was purified by reverse phase flash chromatographic using water/MeOH gradient (10/90) to give a yellow solid (25 mg, 45 %).



¹H NMR (MeOD, 400MHz): δ (ppm) = 1.39 (s, 3H); 1.47 (s, 18H); 1.61 (s, 3H); 1.88-2.20 (m, 4H); 2.45-2.53 (m, 2H); 2.38 (s, 3H); 2.77 (s, 3H); 2.99-3.17 (m, 3H); 3.19-3.23 (q, J = 7.2 Hz, 1H); 3.40-3.65 (m, 2H); 4.08 (bs, 1H); 4.35 (bs, 1H); 4.52-4.58 (m, 1H); 5.05-5.15 (m, 1H); 6.14 (bs, 1H); 8.23-8.27 (m, 2H). **¹³C NMR** (MeOD, 100 MHz): δ (ppm) = 25.8; 26.0; 27.6; 28.4; 28.8; 29.9; 30.7; 33.1; 34.9; 52.2; 54.3; 57.4; 80.5; 82.6; 84.7; 85.1; 86.7; 91.8; 115.4; 120.8; 142.1; 148.4; 150.2; 154.0; 157.4; 158.0; 173.6. **LC-MS** (ESI-APCI): m/z $[M+H]^+$ = 691.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(3-guanidinopropyl)amino)butanoic acid **51**

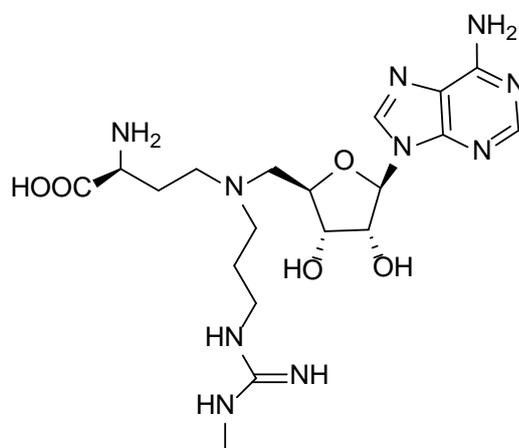
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **48** (20 mg, 0.026 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **51** was obtained as a yellowish sticky solid (10 mg, 98%).



NMR (D₂O, 400MHz): δ (ppm) = 1.97-2.03 (m, 2H); 2.16-2.24 (m, 1H); 2.32-2.42 (m, 1H); 3.18-3.39 (m, 4H); 3.54-3.55 (m, 2H); 3.70-3.96 (m, 5H); 4.51 (bs, 1H); 6.19 (s, 1H); 8.45 (s, 2H). **¹³C NMR** (D₂O, 125 MHz): δ (ppm) = 24.7; 26.8; 36.2; 53.1; 53.7; 55.6; 71.6; 72.7; 78.8; 89.8; 119.3; 142.7; 148.2 ; 151.5; 156.7; 167.9; 172.9. **LC-MS** (ESI-APCI): m/z $[M+H]^+$ = 467 (100%). HPLC purity : t_r = 1.78, \geq 98%.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(3-(3-methylguanidino)propyl)amino)butanoic acid **52**

HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **49** (10 mg, 0.013 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **52** was obtained as a yellowish sticky solid (6 mg, 98%).

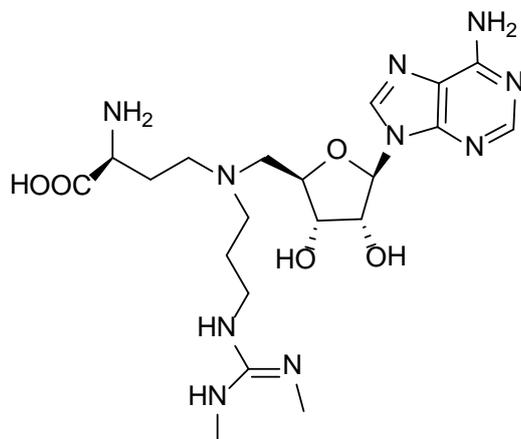


¹H NMR (D₂O, 400MHz): δ (ppm) = 2.05-2.28 (m, 5H); 2.56 (s, 3H); 3.38-3.55 (m, 4H); 3.73-3.79 (m, 6H); 4.44-4.52 (m, 1H); 6.14 (d, *J* = 6 Hz, 1H); 8.38 (d, *J* = 6.8 Hz, 2H). ¹³C NMR (D₂O, 100 MHz): δ (ppm) = 25.2; 27.5; 31.5; 36.9; 48.9; 52.9; 54.4; 55.5; 71.4; 72.5; 78.5; 89.3; 119.9; 140.7; 148.2; 150.2; 153.1; 165.0; 172.9. LC-MS (ESI-APCI): *m/z* [M+H]⁺ = 481 (≥98%). HPLC purity : *tr* = 1.45 minutes, ≥98%.

Preparation of (E)-2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(3-(2,3-dimethylguanidino)propyl)amino)butanoic acid **53**

HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **50** (10 mg, 0.014 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100%

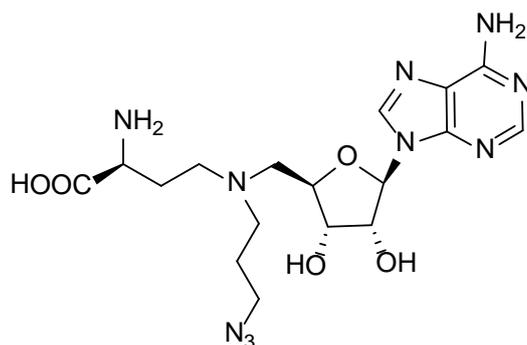
acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **53** was obtained as a yellowish sticky solid (7 mg, 98%).



¹H NMR (D₂O, 400MHz): δ (ppm) = 2.05-2.46 (m, 4H); 2.57 (s, 3H); 2.98 (s, 3H); 3.36-3.65 (m, 8H) ; 3.72-3.97 (m, 2H); 4.50 (bs, 2 H); 6.18 (s, 1H); 8.45 (s, 2H). ¹³C RMN (D₂O, 100 MHz): δ (ppm) = 24.7; 26.9; 29.9; 30.9; 36.2; 49.4; 52.2; 54.0; 71.7; 72.9; 77.8; 90.1; 119.8; 142.9; 144.8; 147.2; 152.5; 160.2; 170.7. LC-MS (ESI-APCI): m/z [M+H]⁺ = 495 (≥98%). HPLC purity : tr = 1.51, ≥98%.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(3-azidopropyl)amino)butanoic acid **54**

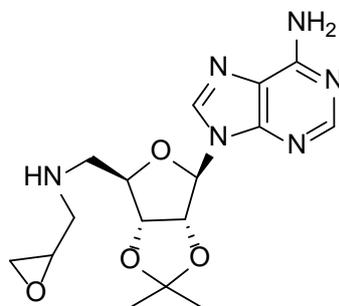
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **16b** (55 mg, 0.085 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **54** was quantitatively obtained as a yellowish sticky solid (35 mg).



¹H NMR (D₂O, 400MHz): δ (ppm) = 1.95-2.17 (m, 1H); 2.26-2.49 (m, 3H); 3.43-3.63 (m, 6H); 3.73-3.89 (m, 3H); 4.12 (m, 1H); 4.52 (bs, 2 H); 6.17 (s, 1H); 8.42 (s, 2H). **¹³C NMR** (D₂O, 100 MHz): δ (ppm) = 25.5; 27.1; 49.1; 51.7; 56.4; 56.6; 72.7; 74.2; 79.4; 91.3; 120.5; 144.8; 145.6; 149.3; 151.1; 171.7. **IR** ν(cm⁻¹): Csp₂-H (703-868); C-O (1079); C-N (1298); C=N (1681); C=O (1732); N₃ (2104); NH₂ (2928). **LC-MS** (ESI-APCI): m/z [M+H]⁺ = 451 (≥98%). **HRMS** (m/z) : [M+H]⁺ = 451.2088, Found: 451.2172.

Preparation of 9-(2,2-dimethyl-6-(((oxiran-2-ylmethyl)amino)methyl)tetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-9H-purin-6-amine **55**

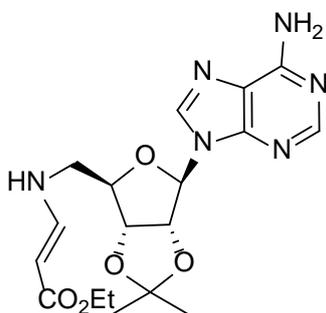
Amine (**10a**) (500 mg, 1.63 mmol, 1 eq) was dissolved in anhydrous tetrahydrofuran (10 mL) and stirred at room temperature. Then 2-(bromomethyl)oxirane (140 μL, 1.63 mmol, 1 eq) was dropwise added to the solution. Finally, N,N-diisopropylethylamine (846 μL, 4.90 mmol, 3 eq) was dropwise added at room temperature and the reaction mixture was stirred under reflux overnight. The reaction mixture was evaporated and the crude was purified by flash chromatography using a dichloromethane/methanol gradient (85/15 to 70/30) to give a yellow solid (200 mg, 30 %).



¹H NMR (CDCl₃, 400MHz): δ (ppm) = 1.34 (s, 3H); 1.57 (s, 3H); 2.19-2.40 (m, 1H); 2.51-2.59 (m, 1H); 2.67-2.73 (m, 1H); 2.87-3.04 (m, 4H); 4.32-4.39 (m, 1H); 4.48 (bs, 1H); 4.99-5.03 (m, 1H); 5.41-5.43 (m, 1H); 6.21 (s, 1H); 7.87 (s, 1H); 8.28 (s, 1H). **¹³C NMR** (CDCl₃, 100 MHz): δ (ppm) = 25.4; 27.4; 29.8; 35.9; 45.4; 51.5; 51.8; 82.4; 84.3; 85.7; 91.0; 114.8; 120.1; 140.0; 149.3; 150.3; 153.3; 155.9.

Preparation of (E)-ethyl 3-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)acrylate **56**

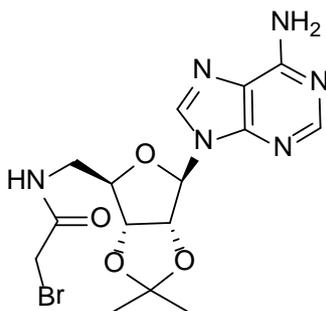
Amine **10a** (344 mg, 1.12 mmol, 1 eq) was dissolved in anhydrous tetrahydrofuran (7 mL) and stirred at room temperature. Then ethyl 3-bromoacrylate (136 μ L, 1.12 mmol, 1 eq) was dropwise added to the solution. Finally, N,N-diisopropylethylamine (582 μ L, 3.37 mmol, 3 eq) was dropwise added at room temperature and the reaction mixture was stirred under reflux overnight. The reaction mixture was evaporated and the crude was purified by flash chromatography using a dichloromethane/methanol gradient (85/15 to 70/30) to give a yellow solid (227 mg, 50 %).



$^1\text{H NMR}$ (CDCl_3 , 400MHz): δ (ppm) = 1.20 (s, $J = 5.6$ Hz, 3H); 1.37 (s, 3H); 1.57 (d, $J = 6$ Hz, 3H); 3.34-3.48 (m, 2H); 4.06 (sp, $J = 6.4$ Hz, 2H); 4.41-4.47 (m, 1H); 4.90-5.02 (m, 1H); 5.25-5.33 (m, 1H); 6.27 (s, 1H); 6.40 (s, 1H); 7.06-7.12 (m, 1H); 7.90 (s, 1H); 8.28 (s, 1H). $^{13}\text{C NMR}$ (CDCl_3 , 100 MHz): δ (ppm) = 14.6; 25.4; 27.5; 50.1; 59.8; 81.5; 83.7; 85.4; 86.4; 90.04; 115.07; 120.01; 125.6; 139.9; 148.9; 152.3; 153.3; 156.1; 170.8.

Preparation of N-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)-2-bromoacetamide **57**

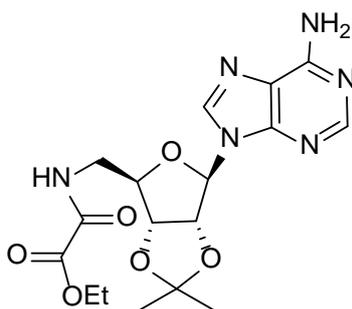
Amine **10a** (200 mg, 0.654 mmol, 1 eq) was dissolved in anhydrous tetrahydrofuran (4mL) and stirred at room temperature. Then 1,3-dibromopropan-2-one (184 mg, 0.850 mmol, 1.3 eq) was dropwise added to the solution. Finally, N,N-diisopropylethylamine (338 μ L, 1.96 mmol, 3 eq) was dropwise added at room temperature and the reaction mixture was stirred under reflux overnight. The reaction mixture was evaporated and the crude was purified by flash chromatography using a dichloromethane/methanol gradient (85/15 to 70/30) to give a yellow solid (144 mg, 50 %).



¹H NMR (CDCl₃, 300MHz): δ (ppm) = 1.33 (s, 3H); 1.61 (s, 3H); 3.26-3.30 (m, 2H); 3.67-3.72 (m, 2H); 4.11-4.20 (m, 1H); 4.50 (bs, 1H); 4.82-4.85 (m, 1H); 5.26 (bs, 1H); 5.76 (bs, NH₂, 2H); 5.83 (s, 1H); 7.82 (s, 1H); 8.29 (s, 1H). **¹³C NMR** (CDCl₃, 100 MHz): δ (ppm) = 25.4; 27.7; 29.9; 39.9; 81.7; 82.8; 83.7; 92.8; 115.0; 120.1; 140.6; 149.8; 153.4; 156.3; 161.9.

Preparation of ethyl 2-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)-2-oxoacetate **58**

Amine **10a** (500 mg, 1.634 mmol, 1 eq) was dissolved in anhydrous tetrahydrofuran (15 mL) and stirred at room temperature. Then ethyl-3-bromo-2-oxopropanoate (205 μL, 1.634 mmol, 1 eq) was dropwise added to the solution. Finally, N,N-Diisopropylethylamine (846 μL, 4.902 mmol, 3 eq) was dropwise added at room temperature and the reaction mixture was stirred under reflux overnight. The reaction mixture was evaporated and the crude was purified by flash chromatography using a dichloromethane/methanol gradient (85/15 to 70/30) to give a yellow solid (206 mg, 30 %).

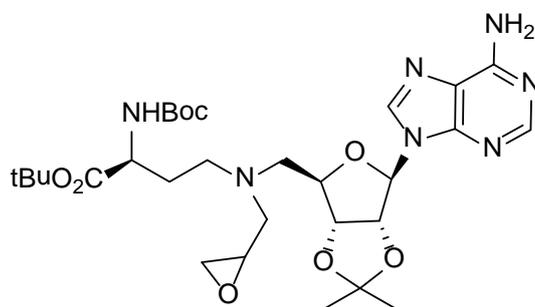


¹H NMR (CDCl₃, 400MHz): δ (ppm) = 1.24 (t, *J* = 5.4 Hz, 3H); 1.38 (s, 3H); 1.55 (s, 3H); 3.06-3.09 (m, 1H); 3.18-3.22 (m, 1H); 4.19 (q, *J* = 5.2 Hz, 2H); 4.38-4.41 (m, 1H); 4.50 (bs, 1H); 5.01 (dd, *J*₁ = 2.8 Hz, *J*₂ = 6.4 Hz, 1H); 5.28 (dd, *J*₁ = 2.4 Hz, *J*₂ = 6.4 Hz, 1H); 6.06 (s, 1H); 6.23 (bs, NH₂, 2H); 7.84 (s, 1H); 8.23 (s, 1H). **¹³C NMR** (CDCl₃, 100 MHz): δ (ppm) =

14.0; 25.4; 27.3; 49.3; 61.7; 81.5; 84.2; 85.3; 90.9; 114.7; 120.0; 139.7; 149.2; 150.3; 153.3; 155.9; 160.9.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(oxiran-2-ylmethyl)amino)-2-((*tert*butoxycarbonyl)amino)butanoate **59**

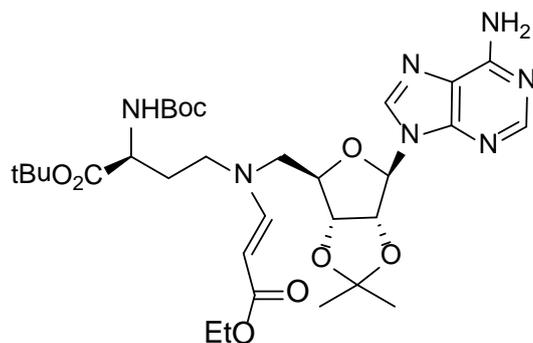
Compound **55** (200 mg, 0.552 mmol, 1.2 eq) was dissolved in anhydrous methanol (800 μ L) and stirred at room temperature. To the solution, aldehyde **3** (136 mg, 0.498 mmol, 1.1 eq) dissolved in anhydrous methanol (200 μ L) was dropwise added. HOAc (28 μ L, 0.460 mmol, 1.0 eq) and NaBH₃CN (46 mg, 0.730 mmol, 1.6 eq) were successively added and the reaction mixture was stirred at room temperature overnight. Then the reaction mixture was diluted in ethyl acetate (50 mL) and washed with a solution of NaOH 1N (2 x 20 mL), the aqueous layer was extracted with dichloromethane (2 x 20 mL), dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol (95/5) as eluent to give a yellow solid (83 mg, 27%).



¹H NMR (CDCl₃, 300MHz): δ (ppm) = 1.37 (s, 3H); 1.44 (s, 18H); 1.58 (s, 3H); 1.69-1.93 (m, 2H); 2.32-3.03 (m, 9H); 4.19 (bs, 1H); 4.40-4.48 (m, 1H); 4.97-5.04 (m, 1H); 5.29-5.31 (m, 1H); 6.08 (s, 1H); 7.84 (s, 1H); 8.32 (s, 1H). ¹³C NMR (CDCl₃, 75 MHz): δ (ppm) = 25.4; 27.3; 27.9; 28.4; 28.7; 30.5; 36.7; 44.5; 50.3; 51.1; 58.4; 80.4; 81.5; 82.4; 84.3; 85.2; 91.2; 114.8; 120.1; 140.3; 149.8; 153.3; 155.8; 156.8; 172.2. LC-MS (ESI-APCI): m/z [M+H]⁺ = 620.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(3-ethoxy-3-oxoprop-1-en-1-yl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **60**

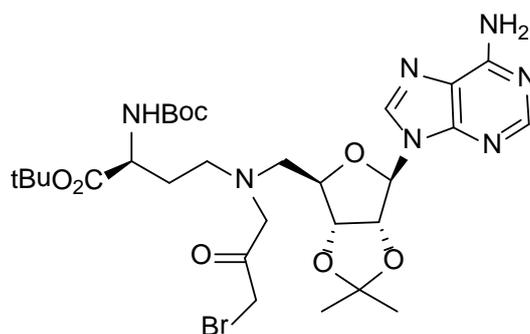
Compound **56** (100 mg, 0.248 mmol, 1.2 eq) was dissolved in anhydrous methanol (400 μ L) and stirred at room temperature. To the solution, aldehyde **3** (136 mg, 0.206 mmol, 1.0 eq) dissolved in anhydrous methanol (200 μ L) was dropwise added. HOAc (12 μ L, 0.206 mmol, 1.0 eq) and NaBH₃CN (21 mg, 0.730 mmol, 1.6 eq) were successively added and the reaction mixture was stirred at room temperature overnight. Then the reaction mixture was diluted in ethyl acetate (50 mL) and washed with a solution of NaOH 1N (2 x 20 mL), the aqueous layer was extracted with dichloromethane (2 x 20 mL) dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol (95/5) as eluent to give a yellow solid (82 mg, 61%).



¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.17 (t, J = 7.2 Hz, 3H); 1.39 (s, 3H); 1.44 (s, 18H); 1.56 (s, 3H); 1.66-1.72 (m, 1H); 1.84-1.94 (m, 1H); 2.62-2.90 (m, 4H); 4.07 (q, J = 7.2 Hz, 2H); 4.25-4.35 (m, 1H); 4.37-4.49 (m, 1H); 4.93-4.98 (m, 1H); 5.38 (dd, J_1 = 3.2 Hz, J_2 = 7.2 Hz, 1H); 5.44-5.52 (m, 1H); 5.98-6.01 (m, 1H); 6.16 (bs, NH₂, 2H); 7.89 (s, 1H); 8.28 (s, 1H). ¹³C NMR (75 MHz, CDCl₃): δ (ppm) = 14.3; 25.6; 27.3; 28.1; 28.5; 29.8; 45.2; 50.7; 53.6; 60.6; 79.6; 81.8; 82.5; 83.9; 85.6; 91.1; 114.7; 120.4; 125.6; 139.9; 140.3; 149.3; 153.2; 155.6; 155.9; 171.9; 172.7. LC-MS (ESI-APCI): m/z [M+H]⁺ = 662.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(3-bromo-2-oxopropyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **61**

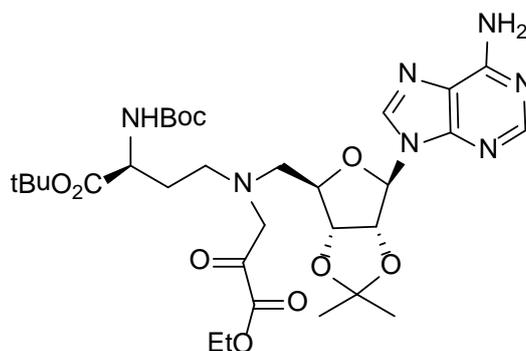
Compound **57** (30 mg, 0.068 mmol, 1.2 eq) was dissolved in anhydrous methanol (200 μ L) and stirred at room temperature. To the solution, aldehyde **3** (15 mg, 0.057 mmol, 1.0 eq) dissolved in anhydrous methanol (100 μ L) was dropwise added. HOAc (3 μ L, 0.057 mmol, 1.0 eq) and NaBH₃CN (6 mg, 0.091 mmol, 1.6 eq) were successively added and the reaction mixture was stirred at room temperature overnight. Then the reaction mixture was diluted in ethyl acetate (50 mL) and washed with a solution of NaOH 1N (2 x 20 mL), the aqueous layer was extracted with dichloromethane (2 x 20 mL) dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol (95/5) as eluent to give a yellow solid (15 mg, 38%).



¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.36 (s, 3H); 1.44 (s, 18H); 1.60 (s, 3H); 1.78-1.98 (m, 2H); 2.13-2.25 (m, 2H); 2.38-2.95 (m, 3H); 3.30-3.46 (m, 1H); 3.59-3.67 (m, 2H); 4.13-4.20 (m, 1H); 4.40-4.50 (m, 1H); 4.83-4.95 (m, 1H); 5.33-5.39 (m, 1H); 5.82 (s, 1H); 6.03 (bs, NH₂, 2H); 7.83 (s, 1H); 8.32 (s, 1H). **¹³C NMR** (CDCl₃, 75 MHz): δ (ppm) = 25.2; 27.0; 27.8; 28.2; 29.5; 35.6; 55.3; 56.1; 58.2; 60.5; 79.3; 81.4; 82.8; 83.6; 85.3; 90.4; 114.3; 119.9; 139.8; 144.7; 148.9; 152.6; 155.6; 165.8; 171.6. **LC-MS** (ESI-APCI): m/z [M-Br]⁻ = 633.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(3-ethoxy-2,3-dioxopropyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **62**

Compound **58** (200 mg, 0.476 mmol, 1.2 eq) was dissolved in anhydrous methanol (800 μ L) and stirred at room temperature. To the solution, aldehyde **3** (118 mg, 0.432 mmol, 1.1 eq) dissolved in anhydrous methanol (200 μ L) was dropwise added. HOAc (24 μ L, 0.397 mmol, 1.0 eq) and NaBH₃CN (40 mg, 0.635 mmol, 1.6 eq) were successively added and the reaction mixture was stirred at room temperature overnight. Then the reaction mixture was diluted in ethyl acetate (50 mL) and washed with a solution of NaOH 1N (2 x 20 mL), the aqueous layer was extracted with dichloromethane (2 x 20 mL) dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol (98/2) as eluent to give a yellow solid (110 mg, 38%).

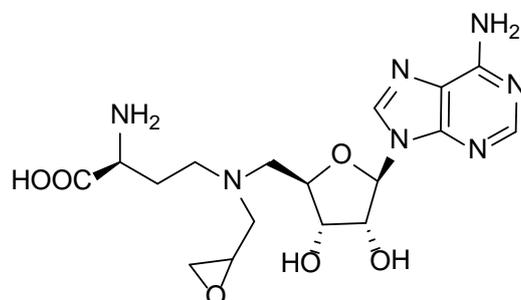


¹H NMR (400MHz, CDCl₃): δ (ppm) = 1.36 (s, 3H); 1.44 (s, 18H); 1.52 (t, J = 8 Hz, 3H); 1.54 (s, 3H); 1.98-2.05 (m, 2H); 2.79-3.02 (m, 2H); 3.05-3.07 (m, 2H); 3.65 (q, J = 8Hz, 2H); 4.28 (bs, 1H); 4.40-4.50 (m, 1H); 4.85-4.95 (m, 1H); 5.28-5.40 (m, 1H); 5.99 (s, 1H); 7.87 (s, 1H); 8.23(s, 1H). ¹³C NMR (CDCl₃, 75 MHz): δ (ppm) = 25.2; 27.0; 27.8; 28.2; 29.5; 35.6; 55.3; 56.1; 58.2; 60.5; 79.3; 81.4; 82.8; 83.6; 85.3; 90.4; 114.3; 119.9; 139.8; 144.7; 148.9; 152.6; 155.6; 155.9; 165.8; 171.6. LC-MS (ESI-APCI): m/z [M-H]⁻ = 677.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(oxiran-2-ylmethyl)amino)butanoic acid **63**

HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **59** (20 mg, 0.032 mmol, 1 eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and

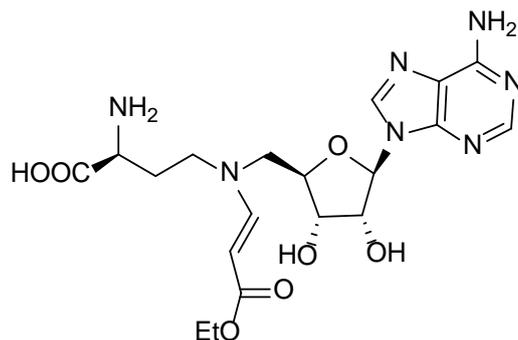
the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **63** was quantitatively obtained as a yellowish sticky solid (14 mg).



$^1\text{H NMR}$ (D_2O , 300MHz): δ (ppm) = 2.16-2.46 (m, 2H); 3.02-3.10 (m, 1H); 3.36-3.71 (m, 5H) ; 3.80 (bs, 2H); 3.91-3.95 (m, 1H); 4.17-4.37 (m, 2H); 4.44-4.58 (m, 2H); 6.17 (s, 1H); 8.45 (s, 2H). $^{13}\text{C NMR}$ (D_2O , 100 MHz): δ (ppm) = 25.6; 47.2; 50.8; 51.5; 55.1; 57.2; 62.6; 71.6; 73.1; 78.1; 90.2; 119.3; 140.9; 143.6; 148.0 ; 149.9; 170.7. HPLC purity : tr = 1.35, $\geq 98\%$.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(3-ethoxy-3-oxoprop-1-en-1-yl)amino)butanoic acid **64**

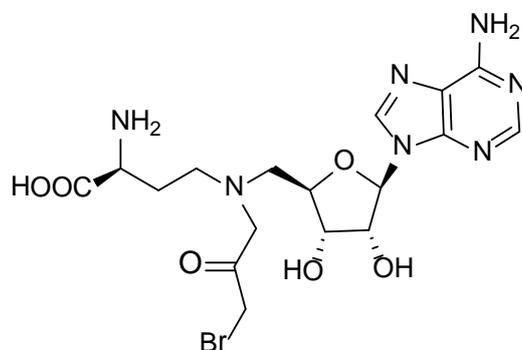
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **60** (20 mg, 0.030 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **64** was quantitatively obtained as a yellowish sticky solid (14 mg).



¹H NMR (D₂O, 300MHz): δ (ppm) = 2.20-2.41 (m, 3H); 2.80 (t, J = 6 Hz, 1H); 2.91 (bs, 2H); 3.56-3.44 (m, 3H); 3.67-3.91 (m, 4H); 4.22 (bs, 1H); 4.52 (bs, 1H); 4.78 (bs, 1H); 4.87 (bs, 1H); 6.16 (s, 1H); 8.43 (s, 2H). **¹³C NMR** (D₂O, 100 MHz): δ (ppm) = 13.9; 27.3; 52.5; 53.4; 54.6; 61.9; 71.9; 73.1; 78.5; 90.4; 102.09; 119.6; 143.8; 148.1 ; 150.1; 153.4; 166.8; 172.2. HPLC purity : tr = 1.31, ≥98%.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(3-bromo-2-oxopropyl)amino)butanoic acid **65**

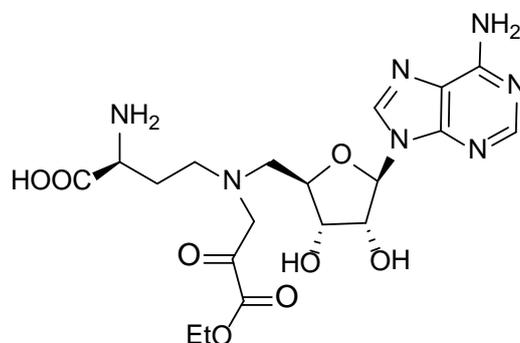
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **61** (15 mg, 0.021 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **65** was quantitatively obtained as a yellowish sticky solid (11 mg).



¹H NMR (D₂O, 300MHz): δ (ppm) = 2.06-2.46 (m, 4H); 2.75-2.82 (m, 1H); 2.91 (bs, 2H); 3.79 (t, J = 8.4 Hz, 2H); 4.10-4.14 (m, 2H); 4.38-4.46 (m, 2H); 4.59 (t, J = 8.8 Hz, 1H); 6.13 (s, 1H); 8.45 (s, 2H). **¹³C NMR** (D₂O, 100 MHz): δ (ppm) = 25.1; 42.4; 53.7; 54.5; 55.2; 62.3; 71.9; 73.1; 78.5; 90.4; 119.4; 143.8; 144.8; 148.1; 150.1; 172.2; 191.9. **LC-MS** (ESI-APCI): m/z [M+H]⁺ = 502. HPLC purity: tr = 1.30, ≥98%.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(3-ethoxy-2,3-dioxopropyl)amino)butanoic acid **66**

HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **62** (10 mg, 0.015 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **66** was quantitatively obtained as a yellowish sticky solid (7 mg).

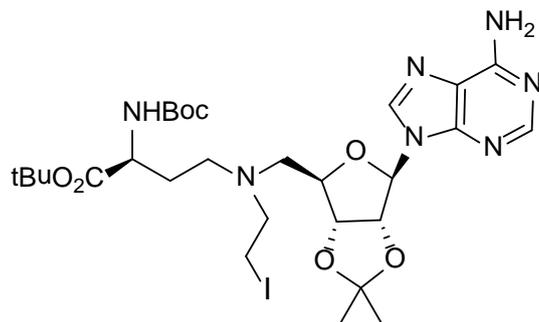


¹H NMR (D₂O, 300MHz): δ (ppm) = 1.33 (d, *J* = 6.6 Hz, 3H); 2.02-2.60 (m, 2H); 3.19-3.23 (m, 1H); 3.50-3.80 (m, 5H); 3.98-4.16 (m, 2H); 4.23-4.68 (m, 4H); 6.23 (s, 1H); 8.48 (s, 2H).
¹³C RMN (D₂O, 100 MHz): δ (ppm) = 13.3; 26.8; 52.4; 54.7; 58.5; 62.8; 71.8; 73.6; 78.7; 90.3; 119.9; 143.4; 148.3; 149.2; 150.6; 166.7; 171.9; 195.2. LC-MS (ESI-APCI): *m/z* [M+H]⁺ = 481. HPLC purity : *tr* = 1.38, ≥98%.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(2-iodoethyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **67**

I₂ (83 mg, 0.329 mmol, 2 equiv) was added to triphenylphosphine (86 mg, 0.329 mmol, 2 equiv) and imidazole (22 mg, 0.329 mmol, 2 equiv) in dichloromethane (1 mL) at 0°C. Compound **32** (100 mg, 0.164 mmol, 1 equiv) in dichloromethane (300 μL) was then added, and the reaction mixture was stirred for 1 h. The reaction was diluted with ice-chilled CH₂Cl₂ and H₂O, and the organic layer was washed twice with H₂O before being placed back under argon and chilled to 0°C. HCl-dioxane (4N, 300 μL, 23 equiv) was added, and the mixture was stirred overnight. The reaction mixture was filtered and evaporated. The

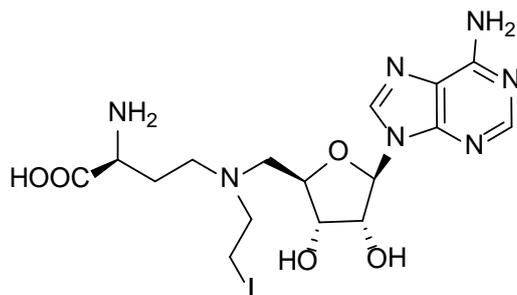
crude was purified by flash chromatography using as eluent dichloromethane/méthanol (98/2) to afford compound **67** as a white solid (40 mg, 34%).



$^1\text{H NMR}$ (300MHz, CDCl_3): δ (ppm) = 1.36 (s, 3H); 1.44 (s, 18H); 1.58 (s, 3H); 1.68-1.70 (m, 1H); 1.86-2.05 (m, 2H); 2.39-2.70 (m, 4H); 2.73-2.82 (m, 2H); 3.03-3.09 (m, 1H); 4.09 (bs, 1H); 4.31 (m, 1H); 4.95-5.00 (m, 1H); 5.30-5.45 (m, 1H); 5.86 (bs, NH_2); 6.01 (s, 1H); 7.89 (s, 1H); 8.31 (s, 1H). $^{13}\text{C NMR}$ (75 MHz, CDCl_3): δ (ppm) = 3.26; 25.6; 27.3; 28.2; 28.6; 29.9; 50.9; 52.9; 56.3; 57.7; 79.7; 81.9; 83.5; 84.1; 85.9; 90.9; 114.6; 120.6; 140.4; 149.3; 153.3; 155.6; 155.9; 171.8.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(2-iodoethyl)amino)butanoic acid **68**

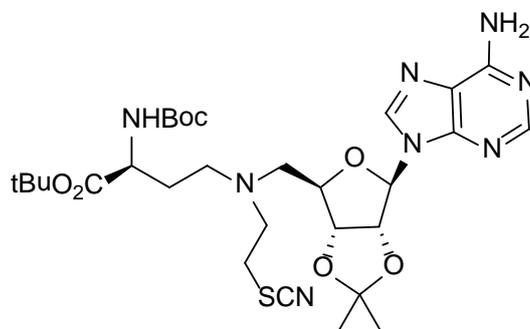
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **67** (50 mg, 0.07 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **68** was quantitatively obtained as a yellowish sticky solid (36 mg).



$^1\text{H NMR}$ (MeOD, 400MHz): δ (ppm) = 2.35-2.58 (m, 2H); 3.53-3.67 (m, 4H); 3.75-3.79 (m, 2H); 3.90-3.96 (m, 1H); 4.17 (t, J = 4.8 Hz, 2H); 4.48-4.56 (m, 2H); 4.74 (t, J = 4.0 Hz, 1H); 6.16 (d, J = 5.2 Hz, 1H); 8.53 (s, 1H); 8.56 (s, 1H). $^{13}\text{C NMR}$ (MeOD, 100 MHz): δ (ppm) = -8.1; 26.0; 51.7; 52.1; 56.8; 57.0; 73.5; 74.9; 80.1; 92.2; 120.6; 144.9; 145.9; 149.9 ; 151.9; 170.7. HPLC purity : tr = 1.87, $\geq 98\%$.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(2-thiocyanatoethyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **69**

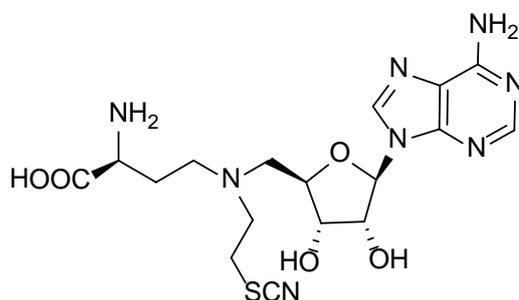
Trimethylsilyl isothiocyanate (12 μL , 0.084 mmol, 1.5 eq) was added at room temperature to compound **67** (40 mg, 0.056 mmol, 1.0 eq) in anhydrous tetrahydrofuran (260 μL). Tetrabutylammonium fluoride (24 μL of a 1.0 M solution in tetrahydrofuran) was then added dropwise. The reaction mixture was stirred overnight, then tetrahydrofuran was evaporated. The crude was submitted to flash chromatography using a dichloromethane/methanol eluent (95/5) yielding compound **69** as a with solid (27 mg, 77%).



$^1\text{H NMR}$ (300MHz, CDCl_3): δ (ppm) = 1.32 (s, 3H); 1.40 (s, 18H); 1.60 (s, 3H); 1.90-2.10 (m, 2H); 2.12-2.30 (m, 2H); 3.00-3.30 (m, 5H); 3.48-3.50 (m, 1H); 4.03 (bs, 1H); 4.75 (bs, 1H); 5.13 (bs, 1H); 5.38-5.50 (m, 1H); 6.19 (s, 1H); 7.82 (bs, 1H); 8.34 (s, 1H). $^{13}\text{C NMR}$ (75 MHz, CDCl_3): δ (ppm) = 25.6; 27.3; 28.1; 28.5; 29.9; 32.1; 50.8; 52.2; 59.4; 80.3; 81.0; 82.8; 83.1; 84.5; 90.1; 112.5; 116.1; 120.0; 133.6; 142.3; 148.3; 153.3; 155.6; 155.8; 170.4. $\text{IR v (cm}^{-1}\text{)} = \text{SCN} : 2154$.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(2-thiocyanatoethyl)amino)butanoic acid **70**

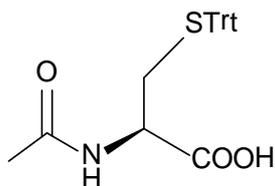
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **69** (27 mg, 0.04 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **69** were obtained as a yellowish sticky solid (7 mg, 37%).



¹H NMR (MeOD, 400MHz): δ (ppm) = 2.41-2.49 (m, 3H); 2.87 (t, J = 6 Hz, 1H); 3.62-3.79 (m, 5H); 3.87-3.98 (m, 1H); 4.19 (bs, 2H); 4.50-4.54 (m, 1H); 4.75 (s, 1H); 6.16 (s, 1H); 8.54 (d, J = 8.8 Hz, 2H). ¹³C NMR (MeOD, 100 MHz): δ (ppm) = 26.1; 42.5; 45.9; 51.6; 52.0; 56.9; 73.5; 74.9; 80.1; 92.3; 112.6; 121.3; 144.6; 145.8; 149.9 ; 151.9; 170.6. HPLC purity t_r = 1.32, \geq 98%.

Preparation of 2-acetamido-3-(tritylthio)propanoic acid **72**

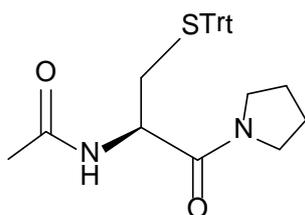
N-acetyl-L-cysteine **71** (2.6 g, 15.93 mmol, 1.0 eq) was dissolved in acetic acid (11.4 mL) followed by the addition of triphenylmethanol (4.23 g, 16.25 mmol, 1.0 eq) and BF₃.etherate (3.03 mL, 23.90 mmol, 1.5 eq) and the solution was stirred for 3 hours at room temperature. The solution was treated with 3 x 15 mL of water and 2 x 12 mL of a saturated sodium acetate solution, chilled on ice, and filtered to yield a white slurry. The slurry was extracted with diethylether, concentrated under reduced pressure, and dried to yield compound **72** as a white solid (5.4 g, 83%).



¹H NMR (CDCl₃, 400MHz): δ (ppm) = 1.88 (s, 3H); 2.46-2.67 (m, 2H); 4.35 (q, *J* = 6.3 Hz, 1H); 7.18 (t, *J* = 12 Hz, 6H); 7.26 (t, *J* = 7.8 Hz, 6H); 7.38 (d, *J* = 7.5 Hz, 3H). **¹³C NMR** (CDCl₃, 100 MHz): δ (ppm) = 23.3; 33.6; 55.3; 67.5; 127.4; 128.5; 129.9; 144.7; 171.8; 174.3.

Preparation of N-(1-oxo-1-(pyrrolidin-1-yl)-3-(tritylthio)propan-2-yl)acetamide **73**

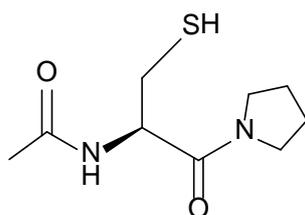
Compound **72** (500 mg, 1.24 mmol, 1.0 eq) was suspended in dichloromethane (12.5 mL) and stirred at room temperature. Then 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (286 mg, 1.49 mmol, 1.2 eq) and *N,N*-diisopropylethylamine (257 μL, 1.49 mmol, 1.2 eq) were successively added in one portion to the solution. Finally, pyrrolidine (123 μL, 1.49 mmol, 1.2 eq) was added to the reaction mixture and stirred at room temperature overnight. The reaction mixture was diluted in ethyl acetate (50 mL) and washed with a hydrogen chloride 1N solution (2 x 25 mL) and a solution of sodium hydroxide 1N (2 x 25 mL), dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/ethyl acetate (70/30) eluent to give a white solid (241 mg, 42%).



¹H NMR (CDCl₃, 300MHz): δ (ppm) = 1.73-1.82 (m, 4H); 1.91 (s, 3H); 2.46-2.62 (m, 2H); 2.99-3.06 (m, 1H); 3.35-3.47 (m, 3H); 4.68 (q, *J* = 6.3 Hz, 1H); 6.97 (d, *J* = 12 Hz, 3H); 7.21 (qu, *J* = 7.8 Hz, 10H); 7.37 (d, *J* = 7.5 Hz, 2H). **¹³C NMR** (CDCl₃, 100 MHz): δ (ppm) = 22.9; 24.0; 25.8; 45.9; 49.9; 66.9; 126.7; 127.9; 129.6; 144.5; 168.7; 169.6.

Preparation of N-(3-mercapto-1-oxo-1-(pyrrolidin-1-yl)propan-2-yl)acetamide 74

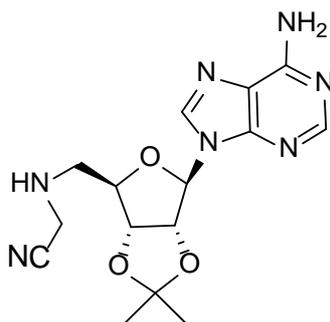
Compound **73** (200 mg, 0.437 mmol, 1.0 eq) was dissolved in dichloromethane (2.7 mL) and stirred at room temperature. Then triisopropylsilane (66 μ L, 0.528 mmol, 1.5 eq) and trifluoroacetic acid (175 μ L, 2.19 mmol, 10.0 eq) were successively added dropwise to the solution. The reaction mixture was stirred at room temperature overnight and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol (90/10) eluent to give quantitatively a white solid (94 mg).



$^1\text{H NMR}$ (MeOD, 400MHz): δ (ppm) = 1.74-1.83 (m, 2H); 1.87 (s, 3H); 1.88-1.95 (m, 2H); 2.56 (dd, $J_1 = 6.8$ Hz, $J_2 = 13.6$ Hz, 1H); 2.72 (dd, $J_1 = 7.6$ Hz, $J_2 = 13.6$ Hz, 1H); 3.31 (t, $J = 6.8$ Hz, 1H); 3.51-3.61 (m, 2H); 4.60 (t, $J = 7.2$ Hz, 1H). $^{13}\text{C NMR}$ (MeOD, 100 MHz): δ (ppm) = 22.4; 25.2; 26.2; 26.9; 47.4; 48.2; 55.3 ; 170.7; 173.1.

Preparation of 2-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)acetonitrile 75

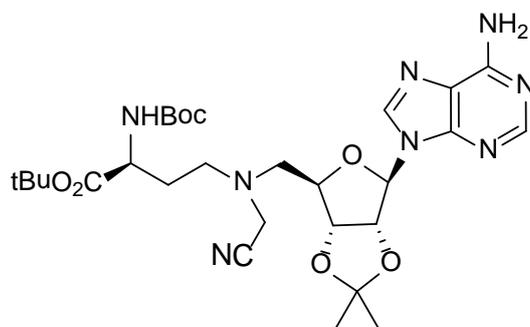
Amine **10a** (100 mg, 0.327 mmol, 1.0 eq) was dissolved in anhydrous tetrahydrofuran (2.8 mL) and stirred at room temperature. Then 2-bromoacetonitrile (23 μ L, 0.327 mmol, 1.0 eq) was dropwise added to the solution. Finally, N,N-diisopropylethylamine (169 μ L, 0.981 mmol, 3 eq) was dropwise added at room temperature and the reaction mixture was stirred under reflux overnight. The reaction mixture was evaporated and the crude was purified by flash chromatography using a dichloromethane/methanol gradient (85/15 to 70/30) to give a white solid (175 mg, 66 %).



$^1\text{H NMR}$ (CDCl_3 , 400MHz): δ (ppm) = 1.34 (s, 3H); 1.57 (s, 3H); 2.94-3.08 (m, 2H) 3.51-3.63 (m, 2H); 4.36 (q, J = 3.6 Hz, 1H); 5.03 (dd, J_1 = 4.0 Hz, J_2 = 6.4 Hz, 1H); 5.41 (dd, J_1 = 3.6 Hz, J_2 = 6.8 Hz, 1H); (d, J = 3.6 Hz, 1H); 7.82 (s, 1H); 8.29 (s, 1H). $^{13}\text{C NMR}$ (CDCl_3 , 100 MHz): δ (ppm) = 25.5; 27.5; 37.6; 50.4; 82.2; 83.2; 84.9; 91.5; 114.9; 117.7; 120.7; 140.2; 149.4; 153.3; 156.1. **IR** ν (cm^{-1}) = CN : 2204.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(cyanomethyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **76**

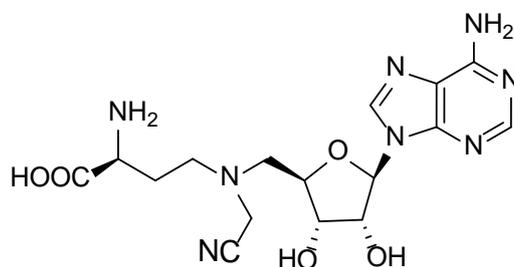
Compound **75** (75 mg, 0.217 mmol, 1.2 eq) was dissolved in anhydrous methanol (300 μL) and stirred at room temperature. To the solution, aldehyde **3** (49 mg, 0.181 mmol, 1.0 eq) dissolved in anhydrous methanol (200 μL) was dropwise added. HOAc (11 μL , 0.181 mmol, 1.0 eq) and NaBH_3CN (18 mg, 0.286 mmol, 1.6 eq) were successively added and the reaction mixture was stirred at room temperature overnight. Then the reaction mixture was diluted in ethyl acetate (50 mL) and washed with a solution of NaOH 1N (2 x 20 mL), the aqueous layer was extracted with dichloromethane (2 x 20 mL) dried over Na_2SO_4 , filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol (98/2) eluent to give a yellow solid (43 mg, 39%).



$^1\text{H NMR}$ (CDCl_3 , 400MHz): δ (ppm) = 1.36 (s, 3H); 1.41 (s, 18H); 1.58 (s, 3H); 1.69-1.74 (m, 1H); 1.77-2.16 (m, 2H); 2.55-2.65 (m, 1H); 2.76-2.90 (m, 2H); 3.50-3.64 (m, 2H); 4.22-4.30 (m, 1H); 4.33-4.37 (m, 1H); 5.02 (dd, J_1 = 4.0 Hz, J_2 = 8.0 Hz, 1H); 5.37-5.45 (m, 1H); 6.01 (s, 1H); 7.89 (s, 1H); 8.32 (s, 1H). $^{13}\text{C RMN}$ (100 MHz, CDCl_3): δ (ppm) = 25.5; 27.5; 28.2; 28.5; 30.4; 43.0; 51.1; 52.5; 55.8; 79.9; 82.3; 82.9; 84.3; 85.9; 90.8; 114.9; 117.7; 120.6; 140.0; 149.4; 153.3; 155.6; 155.9; 171.7.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(cyanomethyl)amino)butanoic acid **77**

HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **76** (43 mg, 0.07 mmol, 1 eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **77** was obtained quantitatively as a yellowish sticky solid (29 mg).

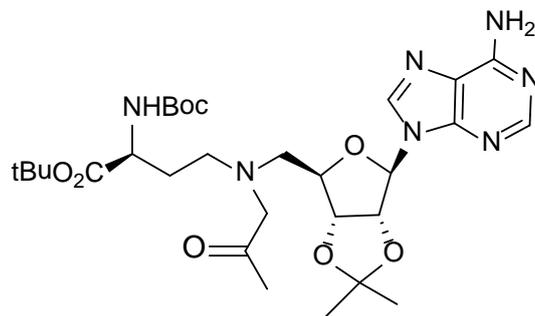


¹H NMR (D₂O, 400MHz): δ (ppm) = 2.03-2.13 (m, 1H); 2.17-2.25 (m, 1H); 2.93-3.02 (m, 2H); 3.08-3.19 (m, 2H); 3.92 (d, *J* = 2.8 Hz, 2H); 4.10 (t, *J* = 6.4 Hz, 1H); 4.31-4.40 (m, 3H); 6.12 (d, *J* = 4Hz, 1H); 8.44 (d, *J* = 7.2 Hz, 2H). **¹³C RMN** (D₂O, 100 MHz): δ (ppm) = 26.3; 42.1; 50.9; 51.9; 56.1; 71.7; 73.2; 81.6; 89.3; 115.7; 119.2; 143.2; 144.5; 148.2 ; 149.9; 171.6. **HRMS** (m/z) : [M+H]⁺ = 407.17859, Found: 407.18037.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(2-oxopropyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **78**

Amine **20a** (200 mg, 0.302 mmol, 1.2 eq) was dissolved in anhydrous methanol (4 mL) and stirred at room temperature. To the solution, 2-oxopropanal (130 μL, 0.252 mmol, 1.0 eq) was dropwise added. HOAc (15 μL, 0.252 mmol, 1.0 eq) and NaBH₃CN (25 mg, 0.403 mmol, 1.6 eq) were successively added and the reaction mixture was stirred at room temperature for 16 hours. Then the reaction mixture was diluted in ethyl acetate (50 mL) and washed with a solution of NaOH 1N (2 x 20 mL), the aqueous layer was extracted with dichloromethane (2 x 20 mL) dried over Na₂SO₄, filtered and evaporated. The crude was

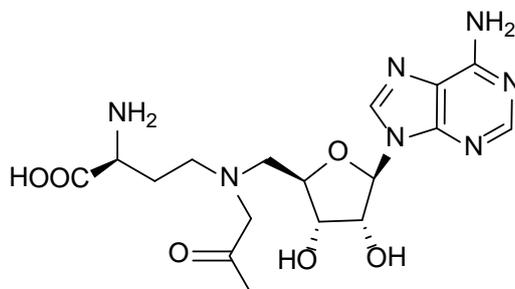
purified by flash chromatography using a dichloromethane/methanol eluent (96/4) to give a white solid (104 mg, 48%).



$^1\text{H NMR}$ (CDCl_3 , 300MHz): δ (ppm) = 1.36 (s, 3H); 1.41 (s, 18H); 1.58 (s, 3H); 2.13 (s, 3H); 2.23-2.90 (m, 6H); 3.23-3.76 (m, 2H); 4.10-4.20 (m, 1H); 4.40-4.50 (m, 1H); 4.92-5.06 (m, 1H); 5.41-5.57 (m, 1H); 5.87 (s, 1H); 7.86 (s, 1H); 8.30 (s, 1H). $^{13}\text{C NMR}$ (CDCl_3 , 100 MHz): δ (ppm) = 20.0; 25.6; 27.3; 28.2; 28.5; 30.1; 51.4; 52.9; 56.5; 63.5; 79.9; 81.5; 83.3; 84.1; 85.5; 90.7; 114.9; 120.6; 140.4; 149.4; 153.3; 155.9; 172.0; 200.3.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(2-oxopropyl)amino)butanoic acid **79**

HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **78** (28 mg, 0.045 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **79** was obtained as a yellowish sticky solid (13 mg, 70%).

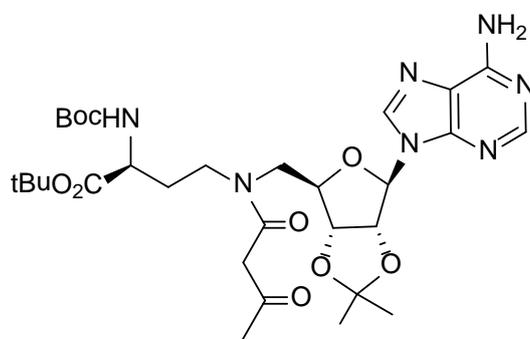


$^1\text{H NMR}$ (D_2O , 400MHz): δ (ppm) = 1.22 (bs, 3H); 2.26-2.42 (m, 2H); 3.15-3.46 (m, 4H); 3.79 (bs, 1H); 3.99-4.09 (m, 2H); 4.22 (bs, 2H); 4.52 (s, 1H); 6.17 (s, 1H); 8.44 (s, 2H).

^{13}C NMR (D_2O , 100 MHz): δ (ppm) = 26.2; 42.9; 50.7; 51.9; 55.3; 65.9; 71.9; 73.2; 79.7; 90.1; 119.9; 143.5; 144.1; 148.1; 150.9; 169.8; 200.3. HRMS (m/z) : $[\text{M}+\text{H}]^+$ = 423.1866, Found: 423.1857.

Preparation of *tert*-butyl 4-(N-((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)-3-oxobutanamido)-2-((*tert*-butoxycarbonyl)amino)butanoate **83**

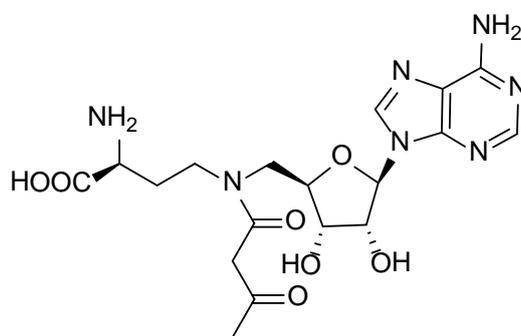
A solution of Meldrum's acid **80** (500mg, 3.47 mmol, 1.0 eq) was dissolved in dichloromethane (13 mL) and cooled at 0°C . Then pyridine (559 μL , 6.94 mmol, 2 eq) was added and the reaction mixture was stirred for 2 hours. To the solution, acetyl chloride (396 μL , 2.51 mmol, 1.5 eq) was dropwise added. The reaction mixture was stirred at room temperature overnight. The reaction mixture was concentrated in vacuo and the crude was used for the next step. To the crude **82** (646 mg, 3.47 mmol, 10 eq) dissolved in anhydrous toluene (20 mL), was added amine **20a** (197 mg, 0.350 mmol, 1.0 eq) dissolved in dichloromethane (5 mL). The reaction mixture was stirred at room temperature for 20 hours. After concentration in vacuo, the residue was purified by flash chromatography using a dichloromethane/methanol eluent (98/2) to afford compound **83** as red solid (70 mg, 31%).



^1H NMR (MeOD, 400MHz): δ (ppm) = 1.39 (s, 3H); 1.42 (s, 18H); 1.60 (s, 3H); 1.75-1.191 (m, 2H); 2.44 (s, 3H); 2.75-2.99 (m, 4H); 3.40-3.77 (m, 2H); 4.05-4.10 (m, 1H); 4.35 (bs, 1H); 5.05 (bs, 1H); 5.49 (bs, 1H); 6.14 (s, 1H); 8.23 (d, $J = 20$ Hz, 2H). ^{13}C NMR (CDCl_3 , 100 MHz): δ (ppm) = 25.8; 26.3; 27.7; 28.4; 28.8; 32.2; 47.2; 52.3; 54.3; 80.7; 82.9; 83.8; 84.9; 86.4; 91.2; 115.9; 121.1; 142.2; 150.4; 154.1; 157.6; 158.1; 170.9; 173.4; 200.5.

Preparation of 2-amino-4-(N-((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)-3-oxobutanamido)butanoic acid **84**

HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **83** (20 mg, 0.031 mmol, 1 eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **84** was obtained quantitatively as a yellowish sticky solid (14 mg).

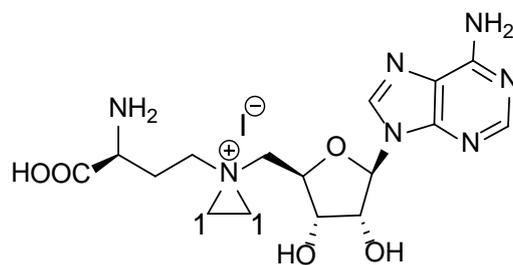


¹H NMR (D₂O, 400MHz): δ (ppm) = 2.14-2.40 (m, 2H); 2.68 (s, 3H); 3.32 (t, *J* = 4.0 Hz, 2H); 3.50-3.64 (m, 2H); 3.82 (t, *J* = 6.0 Hz, 1H); 4.38-4.47 (m, 2H); 4.76 (bs, 2H) 4.89 (t, *J* = 6.0 Hz, 1H); 6.12 (d, *J* = 4.0 Hz, 1H); 8.42 (s, 2H). ¹³C NMR (D₂O, 100 MHz): δ (ppm) = 26.2; 32.6; 51.2; 51.9; 55.3; 66.3; 71.9; 73.2; 76.7; 79.7; 91.1; 120.0; 143.5; 145.1; 148.1; 150.9; 169.8; 173.5; 200.5. HPLC purity : tr = 1.29, ≥98%.

Preparation of 1-(3-amino-3-carboxypropyl)-1-((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)aziridin-1-ium iodide **85**

HCl-dioxane (0.48 M) and water (360 μL) were dropwise added to compound **68** (8 mg, 0.015 mmol, 1 eq) at room temperature successively. Then a solution of sodium hydroxide (0.5 M) was dropwise added. The reaction mixture was monitored by HPLC and after stirring for 14 hours, was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing

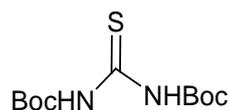
0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **85** was obtained as a white solid (4 mg, 71%).



$^1\text{H NMR}$ (D_2O , 500MHz): δ (ppm) = 1.98-2.11 (m, 2H); 2.98-3.31 (m, 4H); 3.72-3.75 (m, 3H); 3.50-3.64 (m, 3H); 4.29-4.41 (m, 2H); 6.05 (s, 1H); 8.23 (d, $J = 17$ Hz, 2H). $^{13}\text{C NMR}$ (D_2O , 125 MHz): δ (ppm) = 26.0; 40.1 (C1); 51.7; 52.1; 56.8; 57.0; 73.5; 74.9; 80.1; 92.2; 120.6; 144.9; 145.9; 149.9 ; 151.9; 170.7. **LC-MS** (ESI-APCI): m/z $[\text{M}+\text{H}]^+ = 394$. HPLC purity : $t_r = 1.74$, $\geq 98\%$.

Preparation of thiourea **87**

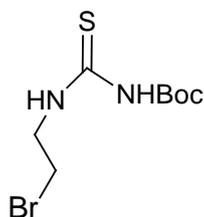
Thiourea **86** (570 mg, 7.5 mmol, 1.0 eq) was dissolved in anhydrous tetrahydrofuran (150 mL) and cooled at 0°C . Then sodium hydride (1.4 mg, 56.25 mmol, 7.5 eq) was added and the reaction mixture was warmed at room temperature and stirred for 10 minutes. Then the reaction mixture was still cooled at 0°C and di-*tert*-butyl dicarbonate (31.6 g, 16.5 mmol, 2.2 eq) was added in one portion to the solution. After stirring at this temperature for 40 minutes, the reaction mixture was warmed to room temperature and stirred overnight. Then the reaction mixture was diluted in water (200 mL) and washed with a solution of sodium hydrogenocarbonate saturated (10 mL). The aqueous layer was extracted with dichloromethane (3 x 50 mL) dried over Na_2SO_4 , filtered and evaporated. The crude was obtained as a yellow solid (2 g) and used for the next step.



$^1\text{H NMR}$ (CDCl_3 , 300MHz): δ (ppm) = 1.40 (s, 18H). $^{13}\text{C NMR}$ (CDCl_3 , 75 MHz): δ (ppm) = 28.2; 28.4; 79.8; 156.9; 159.1.

Preparation of thiourea 88

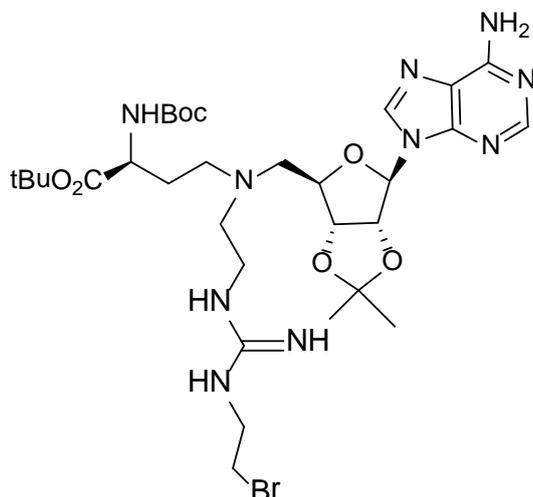
To compound **87** (1 g, 3.6 mmol, 1.0 eq) in anhydrous tetrahydrofuran (72 mL) was added sodium hydride (104 mg, 4.3 mmol, 1.2 eq, 60% in oil) at 0°C. The reaction mixture was stirred at the same temperature for 1 hour, then anhydride trifluoroacetic (551 μ L, 3.9 mmol, 1.1 eq) was added and the stirring was continued for an additional hour. Then, 2-bromoethanamine hydrobromide (683 μ L, 3.9 mmol, 1.1 eq) was added and the resulting reaction was stirred for 50 hours. Water (30 mL) was added to quench the reaction and the mixture was extracted with ethyl acetate (3 x 20 mL). The combined organic layers were dried over Na₂SO₄ and the solvent was removed under reduced pressure. The residue was purified by flash chromatography using a dichloromethane/methanol eluent (98/2) to obtain compound **88** as a yellow solid (200 mg, 17%).



¹H NMR (CDCl₃, 300MHz): δ (ppm) = 1.47 (s, 9H); 3.26 (t, J = 7.8 Hz, 2H); 3.81 (t, J= 7.8 Hz, 2H). ¹³C NMR (CDCl₃, 75 MHz): δ (ppm) = 28.2; 29.9; 41.5; 79.9; 156.9; 159.1.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(2-(3-(2-bromoethyl)guanidino)ethyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate 89

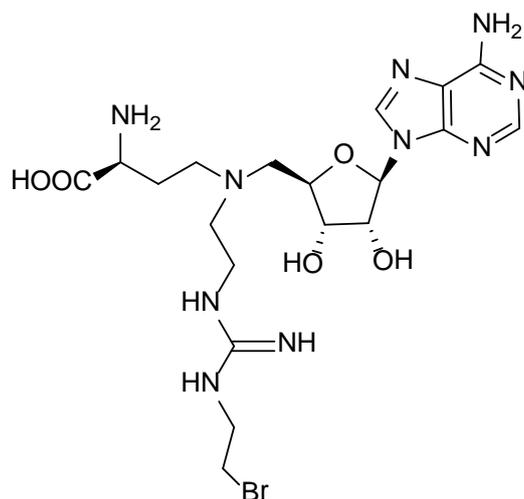
Thiourea **88** (13 mg, 0.036 mmol, 1 eq) was dissolved in anhydrous N,N-dimethylformamide (288 μ L) and the solution was treated with 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (14 mg, 0.071 mmol, 2.0 eq) and N,N-diisopropylethylamine (12 μ L, 0.071 mmol, 2.0 eq) followed by compound **35** (43 mg, 0.071 mmol, 2.0 eq). The reaction mixture was stirred at 70 °C for 12 hours. After stirring, it was directly submitted to reverse phase flash chromatography using eluent water/acetonitril 90/10 to afford compound **89** as white solid (17 mg, 63%).



¹H NMR (MeOD, 300MHz): δ (ppm) = 1.39 (s, 3H); 1.45 (s, 18H); 1.62 (s, 3H); 1.84-2.03 (m, 2H); 2.59-3.12 (m, 9H); 3.55-3.72 (m, 2H); 3.80-4.04 (m, 1H); 4.19 (bs, 1H); 4.33 (bs, 1H) 5.01-5.16 (m, 1H); 5.47 (bs, 1H); 6.20 (s, 1H); 8.28 (s, 2H). **¹³C NMR** (MeOD, 100 MHz): δ (ppm) = 25.7; 27.6; 28.4; 28.9; 29.5; 30.9; 38.9; 40.7; 52.3; 53.5; 55.0; 57.5; 62.1; 80.8; 82.9; 84.6; 85.4; 86.3; 91.3; 115.8; 120.8; 142.4; 150.3; 154.0; 157.5; 158.5; 163.4; 173.8. **LC-MS** (ESI-APCI): m/z $[M+H]^+$ = 755.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(2-(3-(2-bromoethyl)guanidino)ethyl)amino)butanoic acid **90**

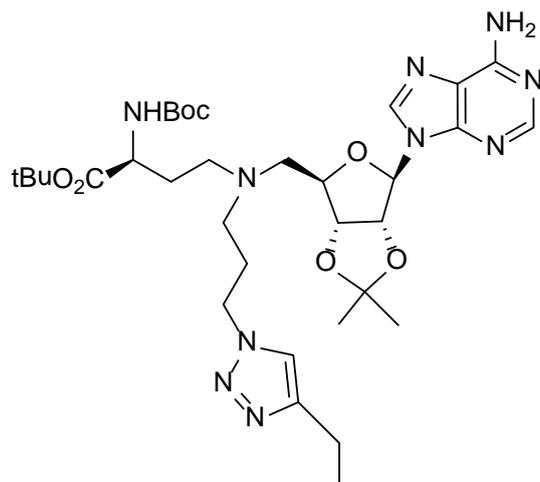
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **85** (10 mg, 0.013 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL) and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **89** was obtained as a yellowish sticky solid (7 mg, 98%).



¹H NMR (D₂O, 300MHz): δ (ppm) = 1.40-1.50 (m, 1H); 1.64-1.70 (m, 1H); 1.95 (bs, 1H); 2.14-2.41 (m, 2H); 2.91 (t, *J* = 8.1 Hz, 1H); 3.17 (s, 1H); 3.40-3.56 (m, 6H); 3.67-3.72 (m, 2H); 3.96 (bs, 1H); 4.70 (s, 2H); 6.10 (s, 1H); 8.40 (s, 2H). **¹³C NMR** (D₂O, 125 MHz): δ (ppm) = 24.5; 30.87; 37.08; 40.0; 46.4; 53.5; 53.8; 56.7; 72.5; 73.9; 82.8; 91.1; 120.6; 144.2; 145.5; 146.4; 148.9 ; 150.7; 170.1. **HRMS** (m/z) : [M+H]⁺ = 560.4117, Found: 560.2303. HPLC purity : tr = 1.47 minutes, ≥98%.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(3-(4-ethyl-1H-1,2,3-triazol-1-yl)propyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **92**

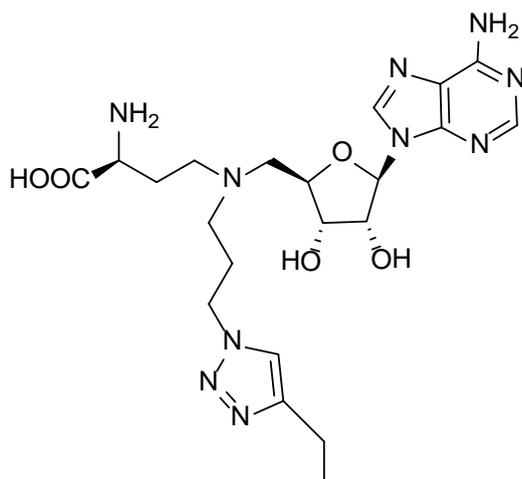
Compound **16b** (100 mg, 0.160 mmol, 1.0 eq) was introduced into the round bottom flask equipped by a condenser with dry ice then it was dissolved in water (640 μL). CuSO_{4(s)} (8 mg, 0.03 mmol, 0.2 eq) was added in one portion followed by water (640 μL) and sodium ascorbate (11mg, 0.06 mmol, 0.4 eq). The reaction mixture was cooled at -78°C and butyne (gas) (20 mL) was added and the reaction mixture was slowly warmed to room temperature for 2 hours. Then the reaction mixture was warmed to 50°C and stirred overnight. After cooling, water (10 mL) was added to the reaction mixture and the aqueous layer was extracted with ethyl acetate (3x 10 mL). The organic layer was dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatographie using as eluent dichloromethane/methanol (85/15) to obtain **92** as a yellow solid (77 mg , 70%).



¹H NMR (CDCl₃, 300MHz): δ (ppm) = 0.84 (d, *J* = 3.9 Hz, 3H); 1.36 (s, 3H); 1.47(s, 18H); 1.60 (s, 3H); 1.71-1.91 (m, 4H); 2.41-2.94 (m, 8H); 3.29-3.66 (m, 2H); 4.15 (bs, 1H); 4.33 (bs, 1H); 4.99 (bs, 1H) 5.40 (bs, 1H); 6.02 (s, 1H); 7.35 (s, 1H); 7.99 (s, 1H); 8.36 (s, 1H). **¹³C NMR** (CDCl₃, 100 MHz,): δ (ppm) = 13.9; 22.7; 25.5; 27.2; 27.9; 28.4; 31.9; 50.9; 52.4; 54.9; 56.4; 63.4; 79.7; 81.9; 83.1; 83.8; 85.6; 90.9; 114.5; 121.2; 129.0; 139.0; 149.5; 153.4; 155.7; 156.1; 172.1. **LC-MS** (ESI-APCI): *m/z* [M+H]⁺ = 701.

Preparation fo 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(3-(4-ethyl-1H-1,2,3-triazol-1-yl)propyl)amino)butanoic acid **93**

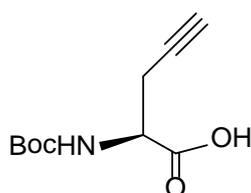
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **92** (25 mg, 0.036 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **93** was obtained quantitatively as a yellowish sticky solid (15 mg).



¹H NMR (D₂O, 300MHz): δ (ppm) = 1.18 (t, J = 6 Hz, 1 H); 1.32 (d, J = 6 Hz, 2H); 2.12-2.26 (m, 3H); 2.47-2.64 (m, 1H); 3.50-4.08 (m, 10H); 4.26 (bs, 1H); 4.54 (s, 2H); 5.43 (s, 1H); 6.18 (s, 1H); 7.96 (1 H) ; 8.45 (2 H). **¹³C NMR** (D₂O, 125 MHz): 12.6; 24.5; 27.2; 27.6; 48.1; 52.2; 54.4; 55.4; 62.6; 71.7; 72.9; 78.5; 90.2; 119.5; 125.2; 140.7; 143.7; 144.6; 150.1; 150.2; 172.1; 175.4. **HRMS** (m/z) : [M+H]⁺ = 505.2557, Found: 505.2632.

Preparation of (2S)-2-[(*tert*-Butoxycarbonyl)amino]pent-4-ynoic acid **95**

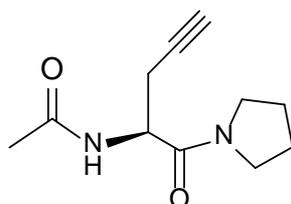
L-Propargyl glycine **94** (300 mg, 2.650 mmol, 1.0 eq) was dissolved in a 2:1 mixture of 1,4-dioxane/H₂O (5.4 mL) and cooled to 0 °C. Then a sodium hydroxide solution (1.3 mL, 2.650 mmol, 1.0 eq) was added dropwise followed by di-*tert*-butyl dicarbonate (635 mg, 2.9 mmol, 1.1 eq). The reaction was stirred for 15 minutes at 0 °C, then allowed to warm to room temperature and stirred for additional 2 hours. The 1, 4-dioxane was removed *in vacuo* and the resulting aqueous residue cooled to 0 °C and washed with ethyl acetate (20 mL). The aqueous solution was acidified to pH \approx 2-3 with an aqueous NaHSO₄ solution and extracted with ethyl acetate (3 \times 40 mL). The combined organic layers were washed with water, dried over MgSO₄, filtered and the solvent removed *in vacuo* to afford the target material as a colorless gum (250 mg, 44%).



$^1\text{H NMR}$ (CDCl_3 , 300 MHz) δ (ppm) = 1.47 (s, 9H), 2.09 (t, J = 2.7 Hz, 1H), 2.77-2.86 (m, 2H), 4.50-4.56 (m, 1H), 5.36 (bs, 1H). $^{13}\text{C NMR}$ (CDCl_3 , 75 MHz) δ (ppm) = 22.5, 28.2, 51.8, 71.2, 78.3, 80.7, 155.4, 175.2.

Preparation of N-(1-oxo-1-(pyrrolidin-1-yl) pent-4-yn-2-yl) acetamide 96

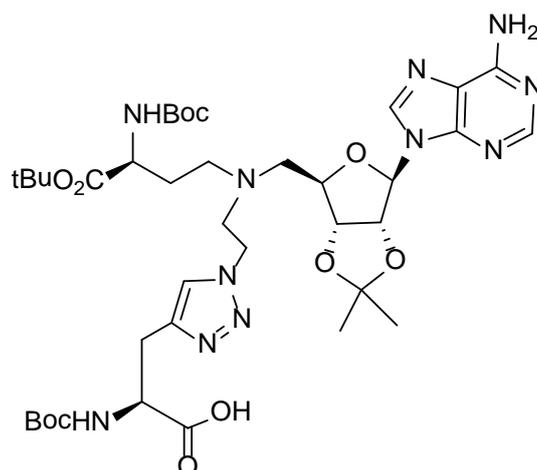
(*L*)-Propargyl glycine **94** (300 mg, 2.65 mmol, 1.0 eq) was dissolved in anhydrous tetrahydrofuran (26.5 mL) and stirred at room temperature. Then 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (612 mg, 3.19 mmol, 1.2 eq) and *N,N*-diisopropylethylamine (550 μL , 3.19 mmol, 1.2 eq) were successively added in one portion to the solution. Finally pyrrolidine (263 μL , 3.19 mmol, 1.2 eq) was added to the reaction mixture and stirred at room temperature overnight. The reaction mixture was diluted in ethyl acetate (50 mL) and washed with a hydrogen chloride 1N solution (2 x 25 mL) and a solution of sodium hydroxide 1N (2 x 25 mL), dried over Na_2SO_4 , filtered and evaporated. The crude was used for the next step. The crude (900 mg, 3.00 mmol, 1.0eq) was dissolved in a mixture 1:1 of anhydrous dichloromethane/pyridine (26 mL) and stirred at 0°C. Then anhydrid acetic was dropwise added and the reaction was warmed to room temperature and stirred for 12 hours. The reaction mixture was reduced *in vacuo* and purified by flash chromatography using a cyclohexane/ethyl acetate (20/80) eluent to give an orange solid (386 mg, 70%).



$^1\text{H NMR}$ (MeOD , 300 MHz) δ (ppm) = 1.88-1.99 (m, 4H), 2.15 (s, 3H), 2.37 (s, 1H), 2.68-2.73 (m, 2H); 3.41 (t, J = 6.4 Hz, 2H), 3.51 (t, J = 6.4 Hz, 2H); 4.54 (t, J = 6.4 Hz, 1H), 4.92 (bs, 1H). $^{13}\text{C NMR}$ (MeOD , 100 MHz) δ (ppm) = 22.4, 25.6; 27.0; 47.0; 52.8; 71.2; 80.3; 171.9; 173.3.

Preparation of 3-(1-(2-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(4-(*tert*-butoxy)-3-((*tert*-butoxycarbonyl)amino)-4-oxobutyl)amino)ethyl)-1H-1,2,3-triazol-4-yl)-2-((*tert*-butoxycarbonyl)amino)propanoic acid **97**

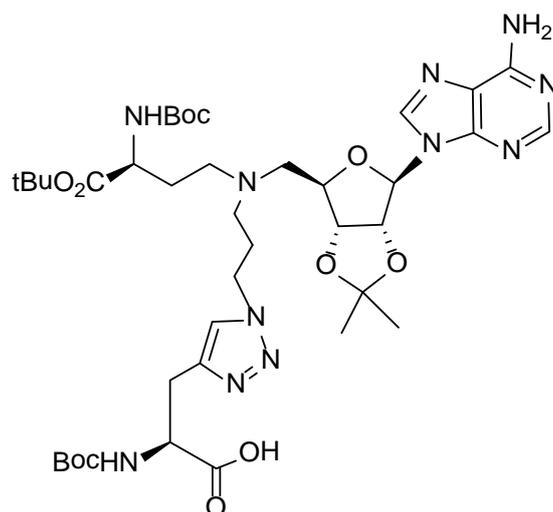
CuSO₄·5H₂O (11.5 mg, 0.046 mmol, 0.2 eq) was dissolved in water (0.9 mL) and added to sodium ascorbate (18 mg, 0.095 mmol, 0.4 eq) at room temperature. Then, the brown solution was added to alkyne **95** (50 mg, 0.23 mmol, 1.0 eq) at room temperature. Finally, azide **34** (148 mg, 0.23 mmol, 1.0 eq) dissolved in *t*-butanol (0.9 mL) was dropwise added to the solution at room temperature and the reaction mixture was warmed to 60°C and stirred for 12 hours. After cooling, water (50 mL) was added to the reaction mixture and the aqueous layer was extracted with ethyl acetate (3 x 20 mL). The organic layer was dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol eluent (85/15) to obtain **97** as a yellow solid (136 mg, 70%).



¹H NMR (CDCl₃, 400MHz): δ (ppm) = 1.38 (s, 3H); 1.44 (s, 18H); 1.46 (s, 9H); 1.56 (s, 3H); 1.68-1.84 (m, 2H); 2.46-2.66 (m, 2H); 2.78-2.93 (m, 1H); 3.15-3.24 (m, 2H); 3.39-3.67 (m, 2H); 3.81-4.11 (m, 3H); 4.31 (bs, 1H); 4.40 (bs, 1H); 4.83-5.07 (m, 1H); 5.11-5.51 (m, 2H); 6.01 (s, 1H); 7.66 (bs, 1H), 7.90 (s, 1H); 8.29 (s, 1H). ¹³C RMN (CDCl₃, 100 MHz,): δ (ppm) = 25.4; 27.1; 28.3; 28.5; 29.8; 46.1; 50.9; 52.0; 54.7; 56.0; 57.2; 59.3; 78.8; 81.8; 82.7; 83.7; 84.2; 90.5; 114.1; 119.8; 122.1; 132.3; 140.1; 148.1; 153.1; 155.5; 155.8; 170.9; 171.5. LC-MS (ESI-APCI): m/z [M+H]⁺ = 847.

Preparation of 3-(1-(3-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(4-(tert-butoxy)-3-((tert-butoxycarbonyl)amino)-4-oxobutyl)amino)propyl)-1H-1,2,3-triazol-4-yl)-2-((tert-butoxycarbonyl)amino)propanoic acid **98**

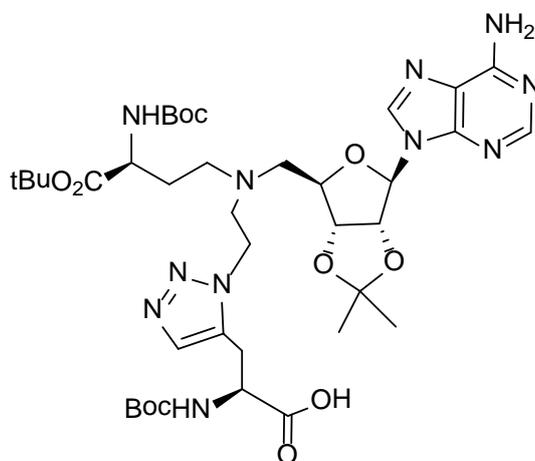
CuSO₄·5H₂O (16 mg, 0.062 mmol, 0.2 eq) was dissolved in water (1.2 mL) and added to sodium ascorbate (23 mg, 0.124 mmol, 0.4 eq) at room temperature. Then, the brown solution was added to alkyne **95** (66 mg, 0.310 mmol, 1.0 eq) at room temperature. Finally, azide **16b** (200 mg, 0.310 mmol, 1.0 eq) dissolved in *t*-butanol (1.2 mL) was dropwise added to the solution at room temperature and the reaction mixture was warmed to 60°C and stirred for 12 hours. After cooling, water (50 mL) was added to the reaction mixture and the aqueous layer was extracted with ethyl acetate (3 x 20 mL). The organic layer was dried over Na₂SO₄, filtered and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol eluent (85/15) to obtain **98** as a yellow solid (186 mg, 70%).



¹H NMR (CDCl₃ 300MHz): δ (ppm) = 1.36 (s, 3H); 1.44 (s, 27H); 1.63 (s, 3H); 1.67-1.93 (m, 2H); 2.14-2.33 (m, 1H); 2.50-2.60 (m, 4H); 3.21-3.26 (m, 3H); 4.08-4.17 (m, 3H); 4.19 (bs, 1H); 4.37 (bs, 1H); 4.60 (bs, 1H); 4.98 (bs, 1H), 5.51 (bs, 2H); 6.05 (bs, 1H); 6.32 (bs, NH₂, 2H); 7.65 (bs, 1H); 7.97 (s, 1H); 8.17 (s, 1H). **¹³C NMR** (100 MHz, δ (ppm) = 25.6; 27.1; 27.8; 28.3; 30.1; 32.6; 34.6; 39.8; 52.9; 55.9; 58.9; 61.3; 59.5; 80.6; 82.7; 83.7; 84.9; 86.5; 91.5; 115.5; 121.4; 132.6; 142.7; 148.5; 152.1; 154.6; 159.3; 171.4; 175.1. **LC-MS** (ESI-APCI): m/z [M+H]⁺ = 860.

Preparation of 3-(1-(2-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(4-(tert-butoxy)-3-((tert-butoxycarbonyl)amino)-4-oxobutyl)amino)ethyl)-1H-1,2,3-triazol-5-yl)-2-((tert-butoxycarbonyl)amino)propanoic acid **99**

Alkyne **95** (100 mg, 0.469 mmol, 1.0 eq) and azide **34** (297 mg, 0.469 mmol, 1.0 eq) were dissolved in anhydrous tetrahydrofuran (4 mL) and the solvent was deoxygenated by bubbling nitrogen through the solution. The catalyst, Cp*Ru(PPh₃)₂Cl (5 mol%), was added in one portion with stirring and the reaction heated under reflux conditions for 12 hours. Volatiles were stripped under vacuum and the residue purified by flash chromatography using as eluent dichloromethane/methanol (85/15) to provide compound **99** as a brown solid (200 mg, 50%).

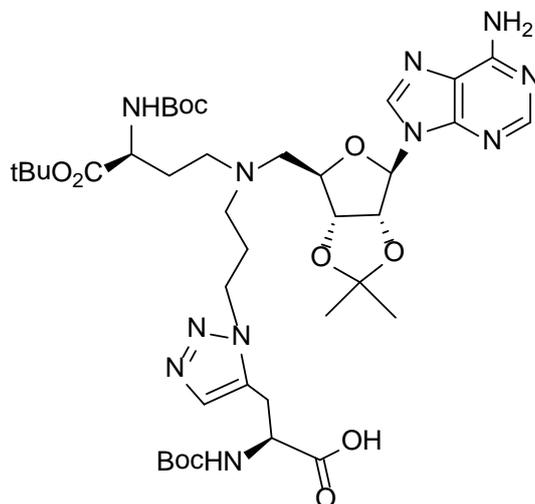


¹H NMR (CDCl₃, 300MHz): δ (ppm) = 1.35 (s, 3H); 1.49 (s, 27H); 1.52 (s, 3H); 1.69-1.80 (m, 2H); 1.93-2.08 (m, 2H); 2.33-2.90 (m, 4H); 3.05-3.55 (m, 2H); 3.56-3.67 (m, 1H); 4.05 (bs, 1H); 4.21 (bs, 1H); 4.33 (bs, 1H); 4.89-5.02 (m, 1H); 5.21-5.35 (m, 1H); 5.44-5.54 (m, 1H); 5.98 (s, 1H), 6.74 (bs, NH₂, 2H); 7.39 (s, 1H); 7.88 (s, 1H); 8.24 (s, 1H).
¹³C NMR (CDCl₃, 100 MHz): δ (ppm) = 25.5; 27.1; 28.0; 28.4; 29.5; 47.2; 49.1; 52.9; 54.2; 57.2; 59.3; 61.5; 79.4; 81.7; 83.2; 83.8; 85.3; 90.5; 114.4; 120.2; 125.2; 133.3; 140.3; 149.1; 153.05; 155.6; 155.9; 171.8; 175.3. LC-MS (ESI-APCI): m/z [M+H]⁺ = 847.

Preparation of 3-(1-(3-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)(4-(tert-butoxy)-3-((tert-butoxycarbonyl)amino)-4-oxobutyl)amino)propyl)-1H-1,2,3-triazol-5-yl)-2-((tert-butoxycarbonyl)amino)propanoic acid **100**

Alkyne **95** (66 mg, 0.309 mmol, 1.0 eq) and azide **16b** (200 mg, 0.309 mmol, 1.0 eq) were dissolved in anhydrous tetrahydrofuran (2.5 mL) and the solvent was deoxygenated by

bubbling nitrogen through the solution. The catalyst, Cp*Ru(PPh₃)₂Cl (5 mol %), was added in one portion with stirring and the reaction heated under reflux conditions for 12 hours. Volatiles were stripped under vacuum and the residue purified by flash chromatography using as eluent dichloromethane/methanol (85/15) to provide compound **100** as brown solid (133 mg, 50%).

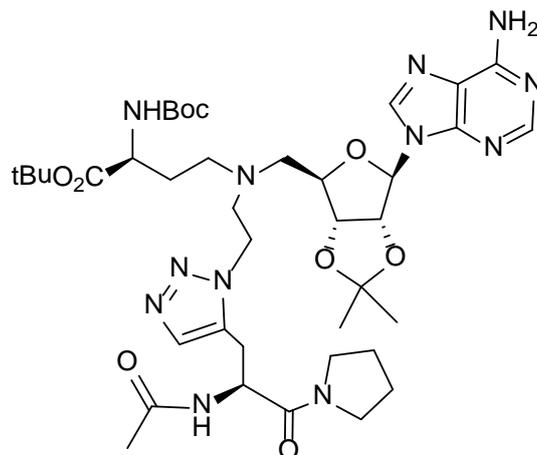


¹H NMR (MeOD, 300MHz): δ (ppm) = 1.39 (s, 3H); 1.45 (s, 27H); 1.60 (s, 3H); 1.84-2.07 (m, 4H); 2.32-2.62 (m, 4H); 2.83-2.90 (m, 2H); 3.56-3.63 (m, 2H); 3.98-4.00 (m, 2H); 4.09 (bs, 1H); 4.30 (bs, 2H); 4.58 (s, 1H); 5.05 (s, 1H); 6.16 (s, 1H); 7.70 (dd, *J*₁ = 9Hz, *J*₂ = 18 Hz, 1H); 8.23 (d, *J* = 9 Hz, 2H). ¹³C NMR (MeOD, 100 MHz): δ (ppm) = 25.5; 26.2; 27.7; 28.5; 28.9; 30.7; 48.5; 50.5; 52.0; 54.5; 57.3; 61.9; 76.7; 80.6; 82.8; 85.0; 86.9; 91.9; 116.2; 120.6; 130.9; 141.9; 150.5; 154.1; 157.5; 158.1; 163.0; 175.6. LC-MS (ESI-APCI): *m/z* [M+H]⁺ = 860.

Preparation of *tert*-butyl 4-((2-(5-(2-acetamido-3-oxo-3-(pyrrolidin-1-yl)propyl)-1H-1,2,3-triazol-1-yl)ethyl)((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **101**

Alkyne **96** (28 mg, 0.136 mmol, 1.0 eq) and azide **34** (86 mg, 0.136 mmol, 1.0 eq) were dissolved in anhydrous tetrahydrofurane (3 mL) and the solvent was deoxygenated by bubbling nitrogen through the solution. The catalyst, Cp*Ru(PPh₃)₂Cl (5 mol%) was added in one portion with stirring and the reaction heated under reflux conditions for 12 hours. Volatiles were stripped under vacuum and the residue was purified by flash chromatography

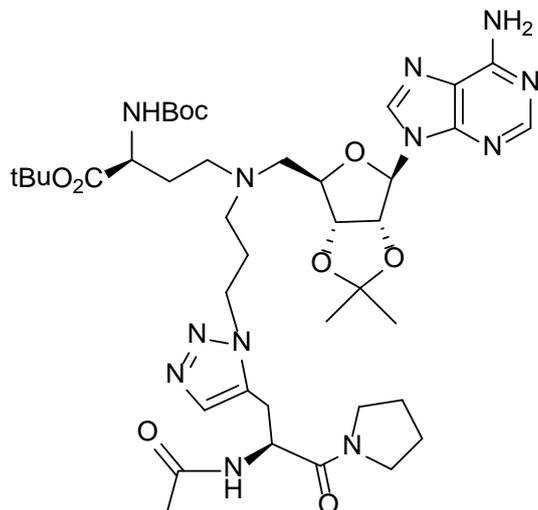
using as eluent dichloromethane/methanol (85/15) to provide compound **101** as brown solid (57 mg, 50%).



¹H NMR (MeOD, 400MHz): δ (ppm) = 1.39 (s, 3H); 1.44 (s, 18H); 1.62 (bs, 6H); 1.96-2.01 (m, 4H); 2.17-2.37 (m, 5H); 2.43-3.14 (m, 9H); 3.48-3.67 (m, 4H); 3.79-3.84 (m, 1H); 4.02-4.17 (m, 2H); 4.35-4.41 (m, 2H); 5.08 (bs, 1H); 6.18 (bs, 1H); 6.98 (bs, NH); 8.10 (d, J = 8 Hz, 1H); 8.25 (d, J = 8 Hz, 2H). **¹³C NMR** (MeOD, 100MHz): δ (ppm) = 22.7, 23.7; 24.9; 25.5; 27.2; 28.4; 28.7; 30.4; 46.8; 51.2; 52.6; 56.1; 57.2; 59.3; 79.7; 81.9; 90.5; 114.7; 120.6; 132.2; 140.4; 149.2; 152.9; 155.5; 171.8; 174.3. **LC-MS** (ESI-APCI): m/z $[M+H]^+$ = 841.

Preparation of *tert*-butyl 4-((3-(5-(2-acetamido-3-oxo-3-(pyrrolidin-1-yl)propyl)-1H-1,2,3-triazol-1-yl)propyl)((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **102**

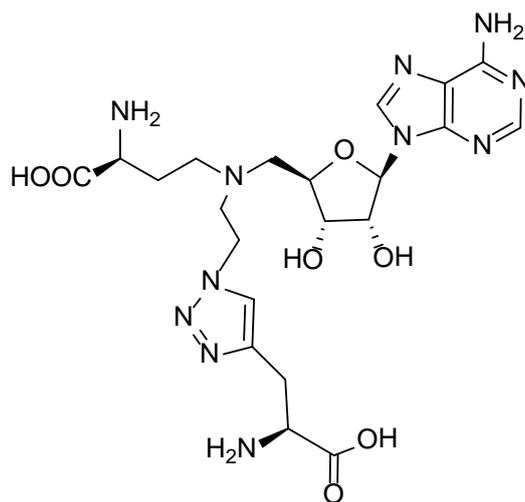
Alkyne **96** (28 mg, 0.136 mmol, 1.0 eq) and azide **16b** (88 mg, 0.136 mmol, 1.0 eq) were dissolved in anhydrous tetrahydrofurane (3 mL) and the solvent was deoxygenated by bubbling nitrogen through the solution. The catalyst, Cp*Ru(PPh₃)₂Cl (5 mol%), was added in one portion with stirring and the reaction heated under reflux conditions for 12 hours. Volatiles were stripped under vacuum and the residue purified by flash chromatography using as eluent dichloromethane/methanol (85/15) to provide compound **102** as brown solid (58 mg, 50%).



¹H NMR (CDCl₃, 500MHz): δ (ppm) = 1.35 (s, 3H); 1.45 (s, 18H); 1.62 (s, 3H); 1.80-1.90 (m, 3H); 2.12 (s, 3H); 2.28-2.78 (m, 9H); 3.44-3.75 (m, 6H); 3.98 (bs, 1H); 4.09 (bs, 3H); 4.27 (bs, 1H) 4.86 (m, 2H); 5.45-5.62 (m, 2H); 6.08 (s, 1H) 7.61 (s, 1H); 7.93 (s, 1H); 8.25 (s, 1H). **¹³C NMR** (CDCl₃, 125 MHz): δ (ppm) = 22.9; 25.05; 25.6; 27.3; 28.1; 28.5; 29.9; 32.1; 45.7; 51.6; 52.4; 56.2; 59.5; 79.8; 80.7; 82.3; 83.0; 84.3; 89.4; 115.9; 119.3; 139.2; 148.1; 152.1; 152.9; 155.4; 171.7; 174.5. **LC-MS** (ESI-APCI): m/z [M+H]⁺ = 855.

Preparation of 2-amino-4-((2-(4-(2-amino-2-carboxyethyl)-1H-1,2,3-triazol-1-yl)ethyl)((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)amino)butanoic acid **103**

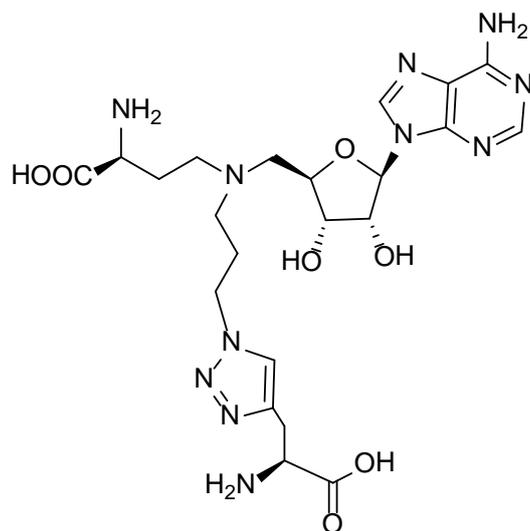
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **97** (20 mg, 0.024 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **103** was obtained quantitatively as a yellowish sticky solid (13 mg).



¹H NMR (D₂O, 400MHz): δ (ppm) = 2.4-2.32 (m, 4H); 3.45-3.62 (m, 4H); 3.70-3.84 (m, 5H); 3.93-4.02 (m, 1H); 4.19 (t, *J* = 4 Hz, 1H); 4.34-4.48 (m, 2H); 6.10 (d, *J* = 4 Hz, 1H); 7.88 (s, 1H); 8.37 (d, *J* = 3.6 Hz, 2H). **¹³C RMN** (D₂O, 100 MHz): δ (ppm) = 24.5; 27.3; 44.7; 50.5; 52.4; 54.5; 55.3; 62.6; 71.7; 73.0; 78.2; 90.0; 119.5; 123.9; 143.3; 144.7; 148.1; 150.00; 171.6. **HRMS** [M+H]⁺ = 550.2409, Found: 550.2482.

Preparation of 2-amino-4-((3-(4-(2-amino-2-carboxyethyl)-1H-1,2,3-triazol-1-yl)propyl)((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)amino)butanoic acid **104**

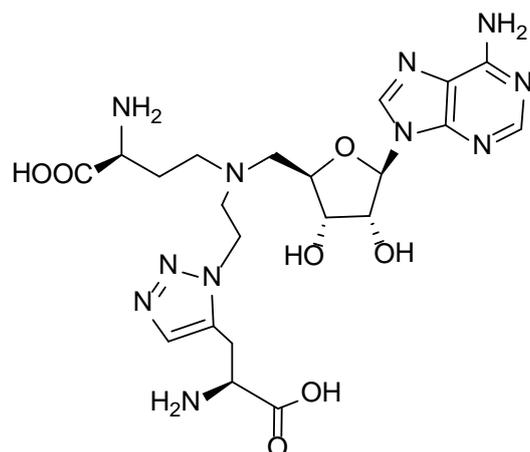
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **98** (60 mg, 0.070 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **104** was obtained quantitatively as a yellowish sticky solid (40 mg).



$^1\text{H NMR}$ (D_2O , 300MHz): δ (ppm) = 2.1-2.35 (m, 5H); 3.48-3.56 (m, 3H); 3.61-3.66 (m, 1H); 3.78-3.86 (m, 2H); 4.06 (bs, 2H); 4.33-4.47 (m, 6H); 6.10 (s, 1H); 7.88 (s, 1H); 8.41 (s, 2H). $^{13}\text{C RMN}$ (D_2O , 100MHz): δ (ppm) = 23.6; 24.4; 25.6; 47.2; 50.6; 51.05; 52.5; 55.1; 62.6; 71.6; 73.0; 78.1; 90.2; 119.4; 125.2; 140.9; 143.6; 144.4; 148.0; 149.9; 170.5; 170.7. **HRMS** $[\text{M}+\text{H}]^+ = 564.2576$, Found: 564.2564.

Preparation of 2-amino-4-((2-(5-(2-amino-2-carboxyethyl)-1H-1,2,3-triazol-1-yl)ethyl)((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)amino)butanoic acid **105**

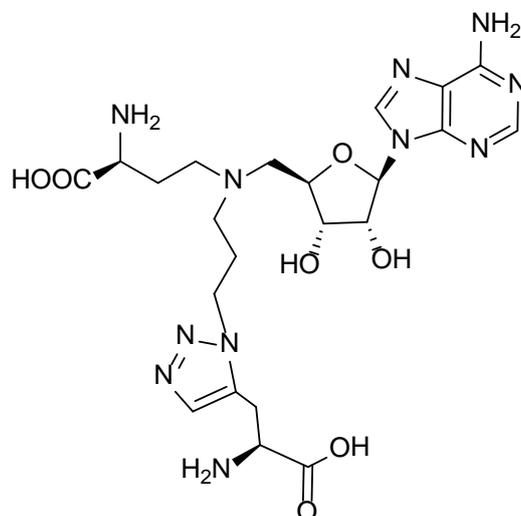
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **99** (30 mg, 0.035 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **105** was obtained quantitatively as a yellowish sticky solid (20 mg).



¹H NMR (D₂O, 300MHz): δ (ppm) = 2.06-2.21 (m, 5H); 3.42-3.59 (m, 5H); 3.59-3.65 (m, 2H); 3.73-3.80 (m, 3H); 4.31-4.37 (m, 2H); 6.06 (d, *J* = 6Hz, 1H); 8.30 (d, *J* = 6 Hz, 2H); 8.37 (s, 1H). **¹³C NMR** (D₂O, 100 MHz): δ (ppm) = 24.5; 25.8; 44.9; 51.6; 53.0; 55.5; 55.8; 57.5; 71.8; 73.2; 78.2; 90.3; 119.7; 125.6; 143.6; 144.7; 148.4; 150.0; 155.4. 170.2; 171.3. **HRMS** [M+H]⁺ = 550.2418, Found: 550.2491.

Preparation of 2-amino-4-((3-(5-(2-amino-2-carboxyethyl)-1H-1,2,3-triazol-1-yl)propyl)((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)amino)butanoic acid **106**

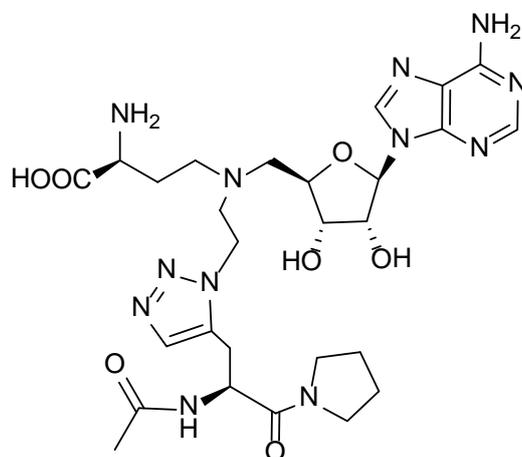
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **100** (20 mg, 0.023 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **106** was obtained quantitatively as a yellowish sticky solid (13 mg).



¹H NMR (D₂O, 300MHz): δ (ppm) = 2.26-2.36 (m, 5H); 2.70-2.75 (m, 1H); 3.02-3.22 (m, 1H); 3.32-3.64 (m, 5H); 3.71-3.83 (m, 2H); 3.96 (bs, 2H); 4.16-4.25 (m, 1H); 4.52 (bs, 2H); 6.19 (bs, 1H); 8.36 (s, 1H); 8.44 (s, 2H). **¹³C NMR** (D₂O, 125 MHz): δ (ppm) = 22.6; 24.5; 25.9; 41.3; 47.9; 51.7; 55.0; 55.4; 71.6; 72.9; 78.3; 90.0; 119.3; 132.09; 143.6; 144.5; 144.7; 148.1; 150.0; 171.8; 176.5. **HRMS** [M+H]⁺ = 564.2576, Found: 564.2647. HPLC purity : tr = 1.34, ≥98%.

Preparation of 4-((2-(5-(2-acetamido-3-oxo-3-(pyrrolidin-1-yl)propyl)-1H-1,2,3-triazol-1-yl)ethyl)((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)amino)-2-aminobutanoic acid **107**

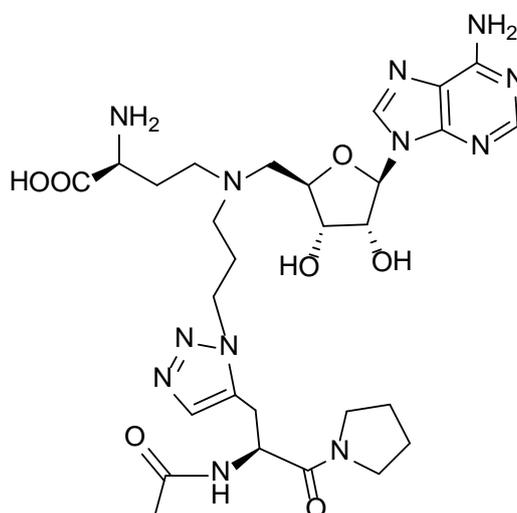
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **101** (20 mg, 0.024 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **107** was obtained quantitatively as a yellowish sticky solid (15 mg).



¹H NMR (D₂O, 500MHz): δ (ppm) = 1.87-1.99 (m, 4H); 2.03-2.06 (m, 2H); 2.09 (s, 3H); 2.65-2.85 (m, 7H); 3.40-3.84 (m, 9H); 4.30-4.35 (m, 2H); 5.08 (t, *J* = 7 Hz, 1H); 6.08 (d, *J* = 4Hz); 8.29 (dd, *J*₁ = 5.5 Hz; *J*₂ = 14Hz, 1H); 8.52 (d, *J* = 5.5 Hz, 2H). **¹³C NMR** (D₂O, 125MHz): δ (ppm) = 23.7; 24.5; 25.4; 26.5; 46.6; 49.1; 50.5; 52.8; 54.8; 55.8; 59.9; 72.0; 72.9; 81.4; 88.2; 119.2; 125.3; 140.2; 152.8; 155.5; 156.0173.3; 188.2. **LC-MS** (ESI-APCD): *m/z* [M+H]⁺ = 645. HPLC purity : *tr* = 1.79, ≥98%.

Preparation of 4-((3-(5-(2-acetamido-3-oxo-3-(pyrrolidin-1-yl)propyl)-1H-1,2,3-triazol-1-yl)propyl)((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)amino)-2-aminobutanoic acid **108**

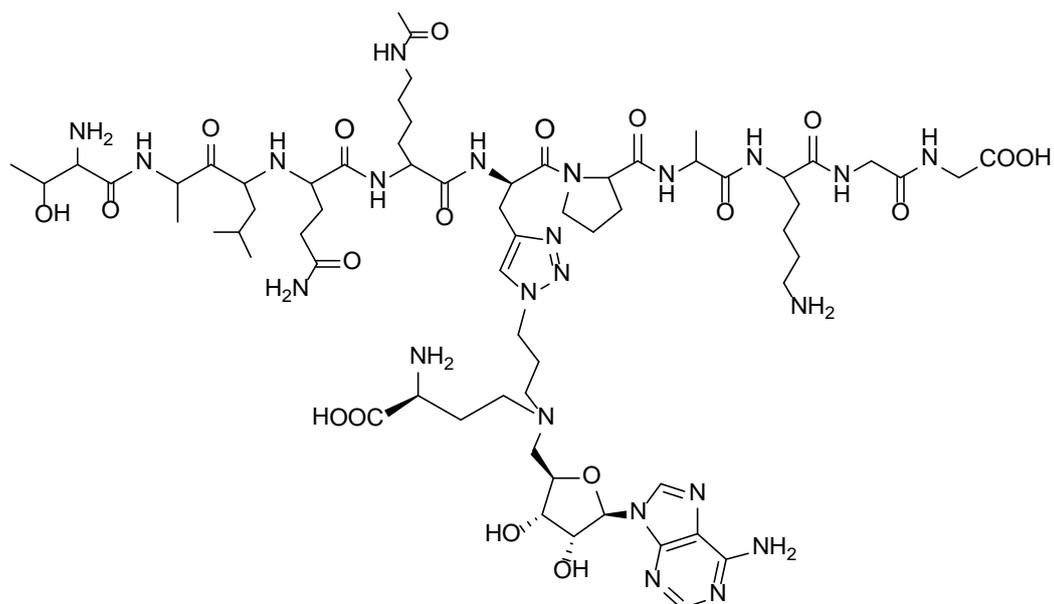
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **102** (10 mg, 0.012 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **108** was obtained quantitatively as a yellowish sticky solid (8 mg).



¹H NMR (D₂O, 500MHz): δ (ppm) = 1.59 (m, 4H); 1.70-1.86 (m, 5H); 2.02 (s, 3H); 2.64-2.73 (m, 2H); 2.77-2.84 (m, 2H); 2.95-3.01 (m, 3H); 3.23 (t, *J* = 5 Hz, 1H); 3.40-3.5 (m, 3H); 3.69-3.75 (m, 2H); 3.77-3.80 (m, 4H); 4.32 (bs, 2H); 6.08 (d, *J* = 4 Hz, 1H); 8.28 (s, 2H); 8.31 (s, 1H). **¹³C NMR** (D₂O, 125MHz): δ (ppm) = 23.7; 24.5; 25.0; 25.4; 26.4; 47.3; 50.5; 51.5; 54.6; 55.7; 60.9; 72.1; 72.8; 81.5; 88.3; 119.3; 126.2; 140.3; 153.1; 155.4; 155.9; 172.9; 188.4. **LC-MS** (ESI-APCI): *m/z* [M+H]⁺ = 659.

Preparation of compound **109**

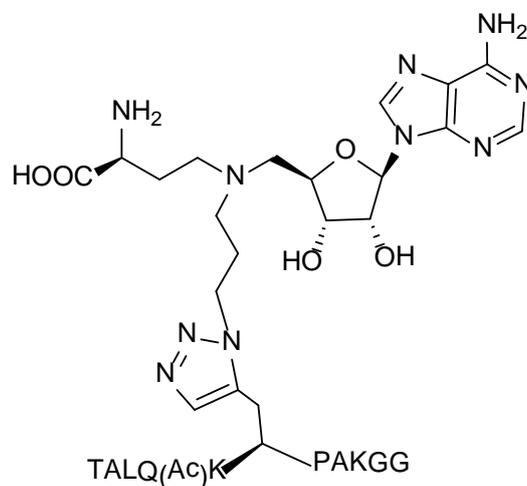
CuSO₄·5 H₂O (0.6 mg, 2.4 × 10⁻³ mmol, 0.3 eq) was dissolved in water (0.2 mL) and added to sodium ascorbate (1.0 mg, 5.3 × 10⁻³ mmol, 0.6 eq) at room temperature. Then, the brown solution was added to alkyne **PJ207** (8.8 mg, 7.9 × 10⁻³ mmol, 1.0 eq) at room temperature. Finally, azide derivative **54** (5.0 mg, 0.011 mmol, 1.4 eq) dissolved in water (0.1 mL) was dropwise added to the solution at room temperature and the reaction mixture was stirred at 37°C under argon. The progress of the reaction was monitored by HPLC. Then, after 12 hours, the reaction mixture was purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. The compound **109** was obtained using a water /acetonitrile (70/30) eluent to give a white solid (12 mg, 67%).



¹H NMR (D₂O, 400MHz): δ (ppm) = 0.83 (bs, 2H); 0.88 (d, *J* = 2.8 Hz, 2H); 0.94 (d, *J* = 4.8 Hz, 2H); 1.16 (d, *J* = 4.4 Hz, 2H); 1.34-1.43 (m, 20H); 1.43-1.66 (m, 9H); 1.92 (s, 3H); 1.97 (s, 3H); 2.05-2.25 (m, 2 H); 2.72-2.75 (m, 3H); 2.91 (t, *J* = 3.6 Hz, 2H); 2.97 (m, 2H); 3.10-3.30 (m, 3H); 3.42 (t, *J* = 8 Hz, 3H); 3.76-3.63 (m, 2H); 3.78 (s, 3H); 3.85-3.88 (m, 2H); 4.01 (d, *J* = 5.2 Hz, 1H); 4.15 (s, 1H); 4.20-4.38 (m, 5H); 4.47-4.42 (m, 2H); 4.61 (t, *J* = 12 Hz, 2H); 5.07-5.1 (m, 1H); 5.44-5.55 (m, 1H); 6.00 (s, 1H); 7.65 (s, 1H) 8.28 (s, 1H); 8.33 (s, 1H). MS (MALDI-TOF): *m/z* [M+H]⁺ = 1557.8.

Preparation of compound 110

Alkyne **PJ207** (10.2 mg, 9.2×10^{-3} mmol, 1.0 eq) and azide **54** (4.2 mg, 9.2×10^{-3} mmol, 1.0 eq) were dissolved in a mixture of water/methanol 1:1 (500 μ L) and the solvent was deoxygenated by bubbling nitrogen through the solution. The catalyst, Cp*Ru(PPh₃)₂Cl (5 mol%), was added in one portion with stirring at 37°C under argon for 12 hours. The progress of the reaction was monitored by HPLC. Then, after stirring, the reaction mixture was purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. The compound **110** was obtained using a water /acetonitrile (70/30) as eluent to give a white solid (4 mg, 30 %).

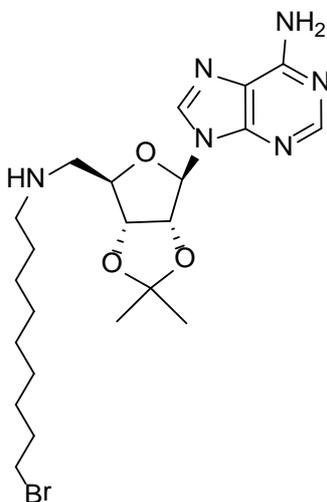


¹H NMR (D₂O, 400MHz): δ (ppm) = 0.88 (d, J = 7.6 Hz, 3H); 0.94 (d, J = Hz, 3H); 1.18 (d, J = 6.4 Hz, 2H); 1.26-1.36 (m, 8H); 1.39-1.40 (m, 6H); 1.45 (d, J = 4.8 Hz, 6H); 1.52-1.65 (m, 6H); 1.92 (s, 3H); 1.98 (s, 3H); 2.05-2.11 (m, 3 H); 2.36 (t, J = 6.4 Hz, 3H); 2.60 (d, J = 6.4 Hz, 2H); 2.75 (d, J = 7.2 Hz, 2H); 2.83-2.85 (m, 4H); 3.16 (t, J = 4 Hz, 2H); 3.38 (d, J = 16.4 Hz, 3H); 3.67 (s, 2H); 3.72 (s, 3H); 3.81-3.85 (m, 2H); 3.98 (d, J = 6 Hz, 1H); 4.15 (d, J = 4.4 Hz, 1H); 4.24-4.38 (m, 5H); 4.45 (q, J = 7.6 Hz, 2H); 4.60 (q, J = 12 Hz, 2H); 6.07 (d, J = 4.4 Hz, 1H); 8.28 (s, 1H); 8.33 (s, 1H); 8.46 (s, 1H). MS (MALDI-TOF): m/z [M+H]⁺ = 1557.8.

Preparation of 9-(6-(((9-bromononyl)amino)methyl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-9H-purin-6-amine 111

Amine **10a** (1.2 g, 3.9 mmol, 1 eq) was dissolved in anhydrous tetrahydrofuran (16 mL) and stirred at room temperature, then 1,9-dibromononane (1.0 mL, 5.1 mmol, 1.3 eq)

was dropwise added to the solution. Finally, N,N-diisopropylethylamine (2.0 mL, 11.7 mmol, 2 eq) was dropwise added and the reaction mixture was stirred under reflux overnight. The reaction mixture was evaporated and the crude was purified by flash chromatography using a dichloromethane/methanol gradient (95/5) to give compound **111** as a red solid 1.5 g, 75%.

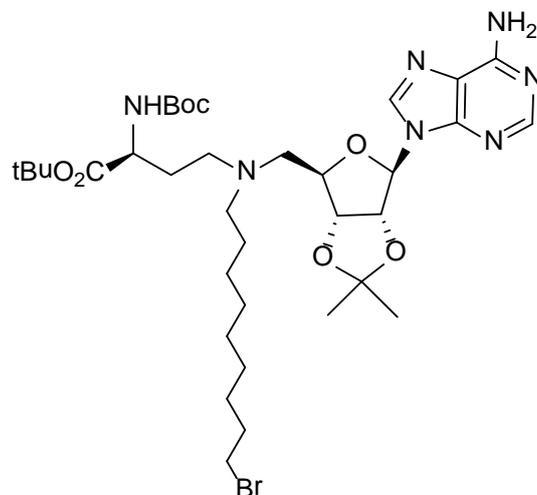


$^1\text{H NMR}$ (CDCl_3 , 300MHz): δ (ppm) = 1.13-1.20 (m, 8H); 1.34 (s, 3H); 1.60 (s, 3H); 1.63-1.72 (m, 5H); 2.89-2.91 (m, 1H); 3.03-3.05 (m, 2H); 3.06-3.11 (m, 2H); 3.57-3.67 (m, 2H); 4.65 (bs, 1H); 5.03 (bs, 1H) 5.22 (bs, 1H); 6.05 (s, 1H); 8.13 (s, 2H). $^{13}\text{C NMR}$ (CDCl_3 , 100 MHz.): δ (ppm) = 19.3; 25.4; 27.4; 29.4; 29.8; 40.6; 43.9; 49.9; 51.3; 81.9; 83.6; 87.6; 90.70; 114.6; 120.4; 139.9; 149.5; 153.09; 155.7. **LC-MS** (ESI-APCI): m/z $[\text{M}+\text{H}]^+ = 511$.

Preparation of *tert*-butyl 4-(((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3] dioxol-4-yl)methyl)(9-bromononyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate
112

Compound **111** (560 mg, 1.1 mmol, 1.2 eq) was dissolved in anhydrous methanol (10 mL) and stirred at room temperature. To the solution, aldehyde **3** (249 mg, 0.9 mmol, 1.0 eq) dissolved in anhydrous methanol (7 mL) was dropwise added at room temperature. Then HOAc (55 μL , 0.9 mmol, 1 eq) was dropwise added, and NaBH_3CN (53 mg, 1.5 mmol, 1.6 eq) was added in one portion. The reaction mixture was stirred at room temperature overnight. Then the reaction mixture was diluted in ethyl acetate (50 mL) and washed with a solution of NaOH 1N (2 x 20 mL), the aqueous layer was extracted with dichloromethane (2 x 20 mL) dried over Na_2SO_4 , filtered and evaporated. The crude was purified by flash

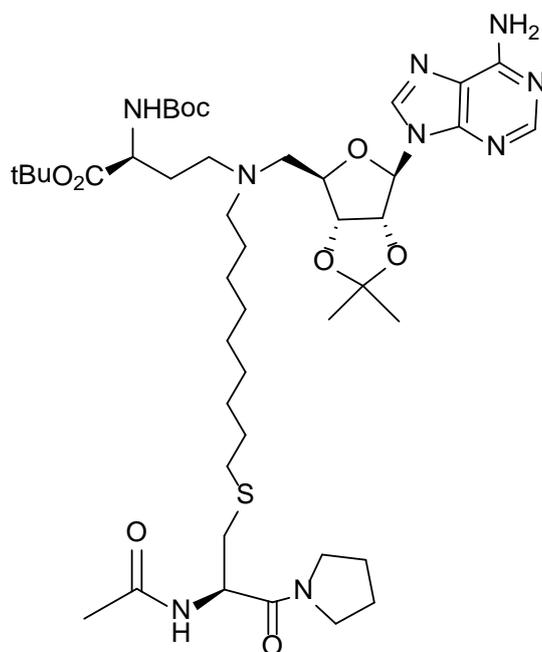
chromatography using a dichloromethane/methanol gradient (98/2) to give compound **112** as an orange solid 250 mg, 36%.



¹H NMR (CDCl₃, 300MHz): δ (ppm) = 1.19-1.31 (m, 9H); 1.38 (s, 3H); 1.45 (s, 18H); 1.62 (s, 3H); 1.73-1.93 (m, 2H); 2.01-2.03 (m, 1H); 2.09-2.11 (m, 1H); 2.24-2.74 (m, 7H); 3.18-3.22 (m, 1H); 3.49-3.71 (m, 2H); 3.93 (bs, 1H); 4.11 (bs, 1H); 4.29 (bs, 1H); 4.93 (dd, $J_1 = 2.8$ Hz, $J_2 = 6.8$ Hz, 1H); 5.43 (bs, 1H); 6.02 (s, 1H); 7.94 (s, 1H); 8.283 (s, 1H).
¹³C NMR (CDCl₃, 100 MHz): δ (ppm) = 25.4; 27.0; 28.2; 28.6; 29.2; 29.4; 30.2; 32.8; 34.1; 52.8; 55.6; 58.3; 59.1; 79.9; 81.3; 82.4; 83.3; 84.1; 90.9; 114.9; 119.9; 140.4; 148.9; 152.7; 155.8; 171.8. LC-MS (ESI-APCI): m/z [M-H]⁻ = 767.

Preparation of *tert*-butyl 4-((9-((2-acetamido-3-oxo-3-(pyrrolidin-1-yl)propyl)thio)nonyl)((6-(6-amino-9H-purin-9-yl)-2,2-dimethyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)methyl)amino)-2-((*tert*-butoxycarbonyl)amino)butanoate **113**

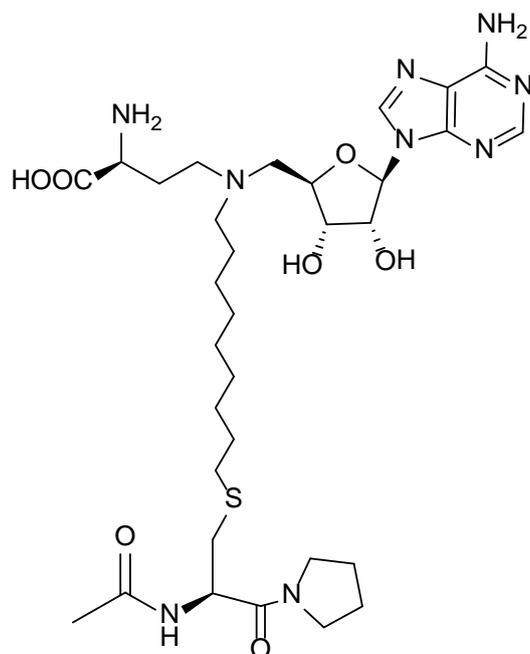
Compound **74** (35 mg, 0.16 mmol, 1.0 eq) was dissolved in anhydrous N,N-dimethylformamide (1.3 mL) and stirred at room temperature. Then K₂CO₃ (54 mg, 0.39 mmol, 2 eq) was added in one portion. To the solution, compound **112** (150 mg, 0.19 mmol, 1.2 eq) dissolved in anhydrous tetrahydrofuran was dropwise added at room temperature. The reaction mixture was stirred at room temperature overnight and evaporated. The crude was purified by flash chromatography using a dichloromethane/methanol gradient (98/2) to give compound **113** as a yellow solid (14 mg, 10%).



¹H NMR (CDCl₃, 300MHz): δ (ppm) = 1.09-1.27 (m, 8H); 1.28-1.31 (m, 4H); 1.37 (s, 3H); 1.41 (s, 18H); 1.62 (s, 3H); 1.73 (bs, 3H); 1.84-1.96 (m, 2H); 2.01-2.28 (m, 3H); 1.84-2.03 (m, 2H); 2.28-2.64 (m, 7H); 2.93-3.12 (m, 1H); 3.26-3.30 (m, 1H); 3.57-3.61 (m, 1H); 3.91-3.95 (m, 1H); 4.08-4.19 (m, 2H); 4.33 (bs, 1H); 4.91 (bs, 1H); 5.40-5.42 (m, 1H); 5.54-5.68 (m, 1H); 6.08 (s, 1H); 7.92 (s, 1H); 8.29 (s, 1H). **¹³C NMR** (CDCl₃, 100 MHz): δ (ppm) = 22.7; 25.2; 27.3; 28.4; 28.7; 28.9; 29.1; 29.3; 28.4; 29.5; 30.1; 32.5; 34.5; 52.7; 53.0; 54.8; 59.9; 79.2; 81.6; 83.8; 85.8; 90.1; 114.1; 120.2; 149.3; 149.4; 152.4; 155.2; 155.6; 170.1; 172.1. **LC-MS** (ESI-APCI): m/z [M-Boc]⁻ = 804.

Preparation of 4-((9-((2-acetamido-3-oxo-3-(pyrrolidin-1-yl)propyl)thio)nonyl)((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)amino)-2-aminobutanoic acid **114**

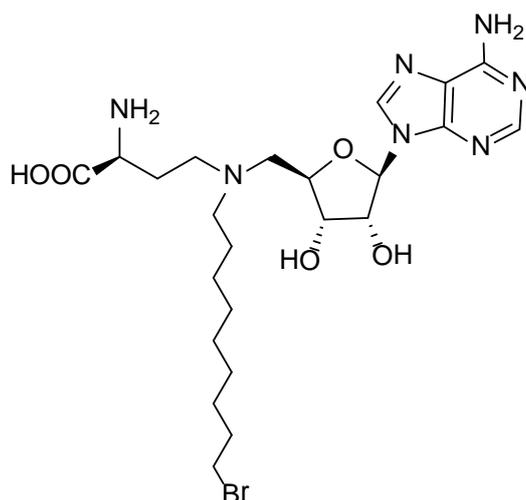
HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **113** (11 mg, 0.012 mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **114** was obtained quantitatively as a yellowish sticky solid (8 mg).



¹H NMR (D₂O, 400MHz): δ (ppm) = 0.78-1.21 (m, 6H); 1.23-1.44 (m, 2H); 1.48-1.67 (m, 2H); 2.22 (s, 3H); 2.28-2.47 (m, 3H); 2.86-2.93 (m, 2H); 3.03-3.15 (m, 5H); 3.48-3.59 (m, 7H); 3.66-3.96 (m, 10H); 4.49 (bs, 2H); 6.13 (bs, 1H); 8.28 (bs, 2H). **LC-MS** (ESI-APCI): m/z [M+H]⁺ = 708.

Preparation of 2-amino-4-(((5-(6-amino-9H-purin-9-yl)-3,4-dihydroxytetrahydrofuran-2-yl)methyl)(9-bromononyl)amino)butanoic acid **115**

HCl-dioxane (2.88 M) and water (0.29 M) were dropwise added to compound **112** (50 mg, 0.065mmol, 1eq) at 0°C successively. The reaction mixture was warmed to room temperature and stirred for 2 hours. The reaction mixture was diluted in water (10 mL), and the aqueous layer was washed with dichloromethane (2 x 10 mL). The aqueous layer was evaporated and purified by reverse phase flash chromatography using a gradient of 0-100% acetonitrile in water containing 0.01% trifluoroacetic acid (TFA) over 30 minutes at a flow rate of 20 ml/min. Compound **115** was obtained quantitatively as a yellowish sticky solid (35 mg).



¹H NMR (D₂O, 300MHz): δ (ppm) = 0.80-1.30 (m, 6H); 1.45-1.60 (m, 2H); 1.82-2.40 (m, 4H); 2.65-2.75 (m, 2H); 3.00-3.22 (m, 2H); 3.30-3.72 (m, 6H); 3.80-4.20 (m, 2H); 4.30-4.50 (m, 4H); 6.08 (s, 1H); 8.38 (s, 2H). **¹³C NMR** (D₂O, 100 MHz): δ (ppm) = 22.8; 24.6; 25.5; 26.8; 28.1; 31.8; 48.9; 51.5; 53.4; 54.4; 55.4; 57.9; 62.1; 71.6; 75.1; 78.4; 90.2; 119.4; 143.6; 144.6; 148.2; 150.0; 171.5. **LC-MS** (ESI-APCI): m/z [M-Br]⁻ = 507.

III. Conditions expérimentales pour l'analyse par spectrométrie de masse des assemblages non covalents

Les échantillons sont dissouts dans une solution tampon d'acétate d'ammonium à 50 mM et ont été analysés en conditions dénaturantes et natives.

III.1 Analyse en conditions dénaturantes

Les analyses des protéines en conditions dénaturantes ont été effectuées à l'aide d'un spectromètre de masse muni d'une source à ionisation électrospray et d'un analyseur à temps de vol (LCT, Waters, Manchester, UK) couplé à un robot d'injection nano-electrospray sur puce microfluidique (Triversa Nanomate Nanomate, Advion Biosciences, Norfolk, UK). Les résultats ont été acquis en mode d'ionisation positif après calibration à l'aide d'une solution de myoglobine de cœur de cheval à 2 μ M dans un solvant eau/acétonitrile/acide formique 50/50/1 ; chaque protéine a été diluée dans un solvant eau/acétonitrile/acide formique 50/50/1 de sorte à obtenir une concentration en monomère de 2 μ M.

III.2 Analyse en conditions natives

Lors des analyses réalisées en conditions natives, l'optimisation des paramètres instrumentaux a été réalisée de sorte à préserver l'intégrité des édifices non-covalents présents en solution lors de leur transfert en phase gazeuse, tout en assurant une désolvatation et une transmission optimale des ions. Cette optimisation concerne en particulier la tension d'accélération appliquée sur le cône d'échantillonnage (V_c) ainsi que la pression dans l'interface (P_i). Chaque protéine a été diluée dans le tampon acétate d'ammonium à 50 mM à pH = 7, pour obtenir une concentration en monomère de 20 μ M.

III.3 Préparation de l'échantillon

Le tampon des échantillons mCARM1₁₃₀₋₅₀₇ a été remplacé par une solution d'acétate d'ammonium à 50 mM dont le pH a été ajusté à la valeur de 7 à l'aide d'une solution d'ammoniaque à 35%. L'échange de tampon a été réalisé par filtration sur gel (colonne Zeba, Pierce) en effectuant 2 cycles de dessalage pour chaque échantillon. La masse moléculaire des trois composants du mélange réactionnel est exprimée en dalton (Da), qui est l'unité de spectrométrie de masse correspondant à un atome d'hydrogène utilisée en biochimie. Les masses moléculaires et les concentrations des réactifs sont indiquées dans le tableau 1 ci-dessous.

	mCARM1₁₃₀₋₅₀₇	PJ207	54
Mwt (Da)	43065,9	1107	450
Concentration (μM)	83	1000	1000

Tableau 1 : Préparation des échantillons pour les analyses. La masse moléculaire ainsi que les concentrations sont indiquées pour les trois composants du mélange réactionnel

