



HAL
open science

Caractérisation structurale et fonctionnelle de mutations oncogéniques de l'histone méthyltransférase SETD2 et de l'histone acétyltransférase MOZ

Christina Michail

► **To cite this version:**

Christina Michail. Caractérisation structurale et fonctionnelle de mutations oncogéniques de l'histone méthyltransférase SETD2 et de l'histone acétyltransférase MOZ. Biologie moléculaire. Université Paris Cité, 2022. Français. NNT : 2022UNIP5198 . tel-04660097

HAL Id: tel-04660097

<https://theses.hal.science/tel-04660097v1>

Submitted on 23 Jul 2024

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.

Thèse de doctorat

Université Paris Cité

Ecole doctorale Bio Sorbonne Paris Cité (BioSPC) 562

Unité de Biologie Fonctionnelle et Adaptative (CNRS UMR 8251)

Équipe Réponses Moléculaires et Cellulaires aux Xénobiotiques (RMCX)

Caractérisation structurale et fonctionnelle de mutations oncogéniques de l'histone méthyltransférase SETD2 et de l'histone acétyltransférase MOZ

Par Christina MICHAIL

Thèse de doctorat de Biologie Cellulaire et Moléculaire

Dirigée par le Pr. Fernando RODRIGUES-LIMA

Présentée et soutenue publiquement le 15 décembre 2022

Devant un jury composé de :

Laurent DELVA, DR, Université de Bourgogne
Eve DE ROSNY, MCU HDR, Université Grenoble Alpes
Souhila MEDJKANE, MCU HDR, Université Paris Cité
Fabien GUIDEZ, CR HDR, Université de Bourgogne
Valérie SERRE, PU, Université Paris Cité
Ahmed HAOUZ, PhD HDR, Institut Pasteur
Fernando RODRIGUES-LIMA, PU, Université Paris Cité

Rapporteur
Rapporteuse
Examinatrice
Examineur
Présidente
Membre Invité
Directeur de Thèse

Caractérisation structurale et fonctionnelle de mutations oncogéniques de l'histone méthyltransférase SETD2 et de l'histone acétyltransférase MOZ

Chez l'Homme, SET Domain containing 2 (SETD2) est la principale histone méthyltransférase responsable de la triméthylation de la lysine 36 de l'histone H3 (H3K36) pour donner la marque épigénétique H3K36me3. Cette enzyme fait partie de la famille des histone méthyltransférases à domaine SET et catalyse le transfert d'un groupe méthyle sur la lysine cible à partir du cofacteur S- adénylméthionine (SAM). La marque H3K36me3 est impliquée dans plusieurs processus cellulaires dont la régulation de la transcription et la réparation de l'ADN. SETD2 est décrite comme un suppresseur de tumeur parmi les plus fréquemment mutés dans les cancers (carcinomes rénaux à cellules claires et leucémies). Malgré l'importance de cet enzyme, les données concernant la caractérisation de ces mutations restent parcellaires. Ce travail de thèse a consisté à caractériser au niveau structural et fonctionnel la mutation L1609P de SETD2 par des approches de biochimie moléculaire et cellulaire et de cristallographie. Cette mutation a été identifiée chez des patients pédiatriques atteints de leucémie. Nous avons pu obtenir la structure cristallographique du mutant SETD2 L1609P à 2.2 Å. Celle-ci révèle que la mutation entraîne des changements conformationnels au niveau du site actif l'enzyme. Cette observation est en accord avec des expériences de dosage d'activité de SETD2 montrant que le mutant est peu actif sur l'histone H3 purifiée ou présente au sein de nucléosomes (25% d'activité résiduelle). L'utilisation de cellules HEK293T modifiées par approche CRISPR/Cas9 pour exprimer le mutant SETD2 L1609P confirme l'impact de la mutation sur l'activité de l'enzyme qui entraîne une baisse significative de la marque H3K36me3 dans les cellules. Nous avons par ailleurs montré que le mutant L1609P est peu exprimé du fait de la formation d'agrégats (puncta) nucléaires "ubiquitinylés" et dégradés par le protéasome. Notre travail montre également que la mutation L1609P de SETD2 confère une résistance à l'apoptose suite à un traitement des cellules à l'étoposide. Ce travail a permis de poser les bases moléculaires et structurales de l'impact d'une mutation oncogénique de SETD2.

En parallèle à ces travaux, nous nous sommes également intéressés à l'étude structurale et fonctionnelle d'une mutation oncogénique de l'histone acétyltransférase MOZ (MOncocytic leukemia Zing finger). Cette enzyme fait partie de la famille des lysine acétyltransférases à domaine MYST et a pour substrat l'histone H3 (notamment H3K23) et des protéines non-histone telles que p53. Des altérations génétiques de MOZ (mutations ou translocations génomiques) sont impliquées dans des hémopathies malignes et d'autres types de cancers. L'impact moléculaire et fonctionnel de ces altérations reste toutefois peu documenté. Notre étude a porté sur la mutation T605P identifiée dans des cancers colorectaux. Cette mutation se situe dans le domaine catalytique de l'enzyme en juxtaposition avec la lysine 604 (K604) dont l'auto-acétylation est décrite comme d'importance pour l'activité enzymatique. Par des approches de biochimie des protéines et d'enzymologie, nous avons montré que le mutant MOZ T605P était incapable d'acétyler l'histone H3 et la protéine p53 mais gardait sa capacité à s'auto-acétyler. L'analyse de cellules HEK293T modifiées par CRISPR/Cas9 exprimant le mutant MOZ T605P confirme la perte d'activité de l'enzyme et la diminution concomitante de la marque épigénétique H3K23ac. Par ailleurs, nous avons résolu la structure cristallographique de MOZ T605P à 2.9 Å et montrons que la mutation entraîne un réarrangement du site actif conduisant à un encombrement stérique pour la fixation de la lysine cible. La structure du mutant MOZ T605P confirme également sa capacité à s'auto-acétyler sur la lysine 604. Ces travaux permettent à la fois d'appréhender l'impact moléculaire et fonctionnel de mutations oncogéniques de MOZ et de mieux comprendre la machinerie catalytique de cette enzyme.

Mots clefs : Épigenétique, Mutations oncogéniques, SETD2, H3K36me3, MOZ

Structural and functional characterization of oncogenic mutations in histone methyltransferase SETD2 and histone acetyltransferase MOZ

SET Domain containing 2 (SETD2) is the major human histone methyltransferase responsible for the trimethylation of lysine 36 of histone H3 (H3K36) and the deposition of the epigenetic mark H3K36me3. This enzyme is a member of the SET domain histone methyltransferases family and catalyzes the transfer of a methyl group onto the target lysine using the cofactor S-adenosylmethionine (SAM). The H3K36me3 epigenetic mark is involved in several cellular processes including transcription regulation and DNA repair. SETD2 is described as one of the most frequently mutated tumor suppressors in cancer (clear cell renal cell carcinomas and leukemias). Despite the importance of this enzyme, the characterization of these mutations remains poorly documented. The aim of this thesis was to characterize structurally and functionally the SETD2 L1609P mutation by molecular and cellular biochemistry and crystallography approaches. This mutation was identified in paediatric leukaemia patients. We were able to obtain the crystallographic structure of the SETD2 L1609P mutant at 2.2 Å. Our structure reveals that the mutation leads to conformational changes at the active site of the enzyme. This observation is in agreement with our activity assays data showing that the mutant is only slightly active on purified histone H3 or on nucleosomes (25% residual activity). The use of HEK293T cells modified by CRISPR/Cas9 approach to express the SETD2 L1609P mutant confirms the impact of the mutation on the activity of the enzyme which results in a significant decrease of the H3K36me3 mark in cells. We have also shown that the L1609P mutant is poorly expressed due to the formation of ubiquitinated nuclear aggregates (puncta) degraded by the proteasome. Our work also shows that the L1609P mutation of SETD2 confers resistance to apoptosis following etoposide treatment of cells. This study provides the molecular and structural basis for the understanding of the impact of an oncogenic SETD2 mutation associated with leukemia. In parallel to this work, we also studied the structural and functional effect of an oncogenic mutation of the histone acetyltransferase MOZ (MONocytic leukemia Zing finger). This enzyme is a member of the MYST domain family of lysine acetyltransferases and its substrates are the H3 histone (notably H3K23) and non-histone proteins such as p53. Genetic alterations of MOZ (mutations or genomic translocations) are implicated in hematological malignancies and other types of cancers. However, the molecular and functional impact of these alterations remains poorly documented. Our study focused on the T605P mutation identified in colorectal cancers. This mutation is located in the catalytic domain of the enzyme in juxtaposition with the lysine 604 (K604) which auto-acetylation is known to be important for enzymatic activity. Using protein biochemistry and enzymology approaches, we showed that the MOZ T605P mutant was unable to acetylate histone H3 and p53 protein but retained its ability to self-acetylate. Analysis of CRISPR/Cas9-modified HEK293T cells expressing the MOZ T605P mutant confirmed the loss of activity of the enzyme and the concomitant decrease in the epigenetic mark H3K23ac. Furthermore, we solved the crystallographic structure of MOZ T605P at 2.9 Å and showed that the mutation causes a conformational change of the active site leading to a steric clash that may alter binding of the target lysine. The structure of the MOZ T605P mutant also confirms its ability to self-acetylate on lysine 604. This work provides a description at the molecular and functional impact of an oncogenic mutation of MOZ and a better understanding of the catalytic machinery of this enzyme.

Keywords: Epigenetics, Oncogenic mutations, SETD2, H3K36me3, MOZ

Table des matières

REMERCIEMENTS	7
LISTE DES PUBLICATIONS.....	19
ABBREVIATIONS	21
FIGURES ET TABLEAUX	25
INTRODUCTION.....	29
I. LA MARQUE EPIGENETIQUE H3K36ME3	31
A. Méthylation des histones	31
1) Méthylation du résidu H3K4.....	33
2) Méthylation du résidu H3K9.....	34
3) Méthylation du résidu H3K27	36
4) Méthylation du résidu H3K36.....	37
5) Méthylation du résidu H3K79.....	39
6) Méthylation du résidu H4K20.....	40
B. La marque H3K36me3.....	41
1) Régulation de la marque H3K36me3.....	42
2) Reconnaissance de la marque H3K36me3.....	44
3) Implication de la marque H3K36me3.....	49
4) Mutations oncogéniques d'H3K36.....	55
II. SET DOMAIN CONTAINING II, SETD2.....	57
A. Les domaines de SETD2 et leur fonction.....	60
1) Les domaines structuraux AWS et Post-SET	61
2) Le domaine catalytique SET.....	62
3) Les domaines d'interactions protéiques WW et PRR	64
4) Les domaines d'interaction avec les RNA pol II (SRI) et hnRNP-L (SHI).....	65
B. Activité lysine méthyltransférase de SETD2.....	67
1) Le complexe ternaire SETD2 – SAM – H3K36/H3M36	67
2) Autorégulation de SETD2	71
3) La stabilité de SETD2.....	73
4) Les substrats non-histone de SETD2	74
C. SETD2 et pathologies	79
1) SETD2 et ccRCC.....	80
2) SETD2 et hémopathies	81
3) SETD2, gliomes et mutations d'H3G34.....	83

4) Le Syndrome de Sotos.....	84
III. MONOCYTIC LEUKEMIA ZINC FINGER, MOZ	85
A. <i>Les Histone AcétylTransférases (HAT)</i>	85
1) Les GNAT	86
2) CBP/p300	87
3) Les MYST	89
B. <i>MOZ (MONocytic leukemia Zinc finger)</i>	93
1) Les domaines de MOZ.....	93
2) La structure du domaine MYST de MOZ.....	95
3) La dyade catalytique de MOZ	96
4) L'auto-acétylation de MOZ.....	98
5) Les rôles de MOZ	100
C. <i>MOZ et pathologies</i>	101
1) Le Syndrome KAT6A.....	102
2) Les mutations de MOZ dans les cancers.....	103
3) Expression de MOZ dans les cancers.....	105
4) Les translocations chromosomiques de <i>Moz</i>	107
OBJECTIFS GÉNÉRAUX	109
CARACTERISATION STRUCTURALE ET FONCTIONNELLE D'UN MUTANT ONCOGENIQUE DE SETD2.....	113
I. INTRODUCTION	115
II. RESULTATS	117
III. DISCUSSION	134
IV. MATERIELS ET METHODES.....	140
IV. ANNEXES.....	159
CARACTERISATION STRUCTURALE ET FONCTIONNELLE D'UN MUTANT ONCOGENIQUE DE MOZ.....	163
I. INTRODUCTION	165
II. RESULTATS	166
III. DISCUSSION	172
IV. MATERIELS ET METHODES.....	176
V. ANNEXES	188
CONCLUSION GÉNÉRALE	193
RÉFÉRENCES	197
ANNEXES (PUBLICATIONS)	217

REMERCIEMENTS

Qui aurait cru que les remerciements seraient la partie de ce manuscrit où j'aurais la plus grande difficulté à rédiger ! Non seulement je ne sais pas par où commencer, mais surtout comme le savent les personnes qui feront partie de ces paragraphes, je ne sais absolument pas où m'arrêter !

Dans un premier temps, je souhaiterais remercier mon jury de thèse. Un grand merci aux rapporteurs Drs. Laurent Delva et Eve De Rosny qui ont accepté d'évaluer mon travail mais également aux Drs. Souhila Medjkane et Fabien Guidez d'être présents le jour de ma soutenance pour échanger sur mon projet. Merci, également, à la Pr. Valérie Serre, d'avoir accepté de présider mon jury de thèse, et d'avoir fait partie de mes comités de suivi individuel. Un grand merci également au Dr. Ahmed Haouz d'être présent comme membre invité en ce jour spécial pour moi.

Je voudrais également remercier le Pr. Bertrand Friguet, qui a suivi l'avancement de mon projet en tant que deuxième membre de mon jury de comité de suivi individuel. C'était un grand plaisir de pouvoir échanger avec vous et le Pr. Valérie Serre, aussi bien par visioconférences depuis Perpignan, qu'en présentiel au laboratoire.

Ce travail de thèse n'aurait pas pu être réalisé sans l'accord et le bon accueil à l'Unité et au laboratoire par le Pr. Jean-Marie Dupret. Merci d'être simple et humain avec nous (qui peut oublier le petit apéro avec des restes d'un congrès, entre le directeur de l'Unité et deux petits thésards), avec de belles discussions cinématographiques, et d'ailleurs merci de nous avoir autorisés à organiser la première raclette du labo !

Pendant mes quatre années de thèse, j'ai eu la chance de construire des liens avec des personnes avec qui je garderai contact longtemps après mon départ du laboratoire. Premier, et spécial dans cette liste, mon directeur de thèse, le Pr. Fernando Rodrigues-Lima, mon 'papa spirituel'. Fernando, je te remercie de la confiance et le support que tu m'as montré, dans tous les niveaux professionnels et personnels ! Tes conseils sont et seront importants pour moi, aussi bien pour ma carrière que pour mon développement personnel. Merci d'être à l'écoute à chaque fois que j'avais besoin et merci pour tous les conseils scientifiques et de vie que tu m'as donnés. Je dois une grande partie de la personne et de la

jeune chercheuse que je suis aujourd'hui à toi et je te remercie pour cela. Je garde énormément de souvenirs de nos discussions enrichissantes et intéressantes autour de la science mais pas que, nos petites séances d'organisation et de nettoyage du labo (mention spéciale de la centri), et les petits moments gourmands avec le saucisson portugais flambé à l'éthanol absolue du labo! Je choisirai comme souvenir spéciale le guide touristique à Saint-Malo et j'attends mon tour pour être le guide locale en Grèce !

À mon arrivée au laboratoire, j'ai eu l'honneur d'être prise sous les ailes du thésard senior de l'époque, Jérémy. Jérémy, tu as été pour moi un grand frère que je n'ai jamais eu. Tu es la personne qui m'a conseillé, qui m'a poussé et qui m'a soutenu tout au long de cette aventure. En vrai, tu as eu plusieurs casquettes, commençant par celle du mentor, en finissant par celle d'un très bon ami. C'est à toi aussi que je dois une grande partie de la 'jeune scientifique' que je suis aujourd'hui, car c'est toi qui m'a appris comment faire ma toute première purification, mes tout premiers cristaux et en général je te dois une grande partie de mes compétences. Merci de m'avoir passé la torche de SETD2, j'espère que j'ai été à la hauteur ! Nous avons passé plein de moments ensemble au labo et en dehors. Nous parlerons toujours du marathon Avengers, des 'escape game', des petites journées concert au labo et même de tous les repas et les verres ensemble. Nous avons été un bon duo pendant ces années. Je suis sûre que nous allons continuer à se voir et à rigoler. Merci d'avoir été là et de m'avoir écouté et conseillé pendant toutes mes crises.

La petite famille que j'ai eue la chance d'avoir au laboratoire, ne peut pas exclure Linh, la personne qui était mon support au bureau du jour 0 jusqu'à aujourd'hui. Nous avons commencé par t'appeler, malgré tes « 18 ans », notre maman du laboratoire, mais au final tu es beaucoup plus que ça... Merci de m'avoir aidé dans tous les niveaux ! Tu m'as énormément appris en tant que scientifique sur l'HPLC, la qPCR, les cellules... Je pense que ce qui va te décrire pendant tes prochains re-18 ans, sera ta passion pour les billes magnétiques et les CHIP à Saint Louis. Il faudra que tu penses à ça, chaque fois que ta chaîne HPLC a des problèmes... Un énorme merci pour toute ton aide les mercredis (et pas que) de cette année. Sans toi je n'aurai jamais pu avancer sur ma thèse et mes manip en parallèle ! Merci pour toutes nos discussions, les échanges de bonheur et de problèmes que nous avons eu...

Je voudrais effectivement remercier Florent et Émile. Florent, merci pour ton aide avec les PCR et la procédure de ma candidature pour le monitorat. Ça m'a fait également plaisir d'écouter tes petites histoires diverses et passionnantes. Émile, tu seras toujours la mascotte de ce laboratoire, mais je pense que le point que je vais garder en mémoire de toi, est toute ton aide que tu m'as donné lors de ma chute et mon accident de travail.

Une grande partie de l'histoire de ce labo est son petit 'gang asiatique' de Chine, avec qui nous avons eu le plaisir de partager l'expérience de thèse. Qing, Wenchao, monsieur Ron, Guiling, vous avez tous marqué avec votre gentillesse et votre petit caractère mes quatre années au laboratoire ! Qing, 'my lady', j'espère que nous aurons un jour l'opportunité d'assister ensemble à un concert de Maroon 5. Je suis très contente pour ton poste, ton mari et ton 'baby Qing'. Je me rappelle que c'est tout ce dont tu as rêvé. Wenchao, mon petit voisin du bureau, nos séances de cristallographie pleines d'émotions, nos échanges culturels et gastronomiques resteront pour longtemps dans ma mémoire. Je suis persuadée que tu feras des brillantes publications qui te permettront d'avoir le poste de tes rêves en Chine. Rony, ta présence au labo nous a tous marqué. Ton caractère « de l'espace » sera impossible d'être remplacé. Tu nous as fait rigoler un nombre de fois incalculable et tu es sans aucun doute la personne avec le plus de surnoms donnés par son prénom et sa personnalité. Vous me manquez tous les trois et j'espère vous revoir bientôt ! Guiling, je suis très contente que tu sois parmi nous au laboratoire, toujours souriante, et que nous pouvons échanger sur les manips, les séances de sport et les petites balades à Paris.

Je voudrais aussi remercier Justine et Louise pour leur bonne humeur. Justine, la femme star de ce laboratoire, je suis sûre qu'un jour nous allons assister à un grand spectacle de 'batchatcha' dans lequel tu vas participer (qui sait, même organiser). En attendant, j'attends de ton retour par rapport à l'organisation du vidéo clip du labo. Louise, je sais que la période qui arrive va être chargée et je te souhaite bon courage.

Au cours de ma dernière année de thèse, nous avons eu le plaisir d'avoir des nouvelles recrues. Mireille, merci pour votre sourire et de vos conseils autour de mes manips. Ça me fait un très grand plaisir d'échanger avec vous. Votre relation 'Statler et Waldorf de Muppet Show' avec Fernando est juste adorable. Frédérique (Fred), nous sommes vite devenues 'partenaires du matin', et c'est ce qui me rend heureuse de savoir que je vais arriver au labo

tôt et que je vais te retrouver avec tes bonnes ondes, pour discuter de tout et de rien. Merci pour toute ton aide, surtout pour ma manip de FACS pour le cycle cellulaire. Tessa, toute jeune mais pas toute innocente, merci pour ton petit côté fufou qui pimente nos journées. Je t'assure que nous irons boire de bières (et du vin pour moi) ensemble pour t'aider à tenir à ton tour le long de ta thèse.

Depuis la première année de ma thèse, j'ai eu l'opportunité de travailler et de beaucoup apprendre par Valentina. Valentina, merci pour toutes les belles IF et pour m'avoir initié dans le monde de sénescence. Notre moment phare était l'encadrement des apprentis – chercheurs, et c'est ça qui nous a permis de mieux se connaître et de se rapprocher. Merci pour les échanges que nous avons eu dans tous les niveaux, personnels et professionnels. D'ailleurs, je ne peux pas parler de Valentina sans parler de Pascal. Pascal, merci pour toutes les discussions que nous avons eu, toutes les blagues que nous avons fait et du dessert que tu m'avais préparé pour mon dernier anniversaire, c'était un délice.

Un grand merci aux autres personnes de l'équipe, Armelle, Stéphanie, Sonja, Dorian et Alice. Merci également à Chloé qui m'a dépanné un certain nombre de fois avec son kit d'extraction d'ARN, et à qui je voudrais également souhaiter bon courage pour la suite de sa thèse. Merci aussi à Françoise pour toute son aide sur les problèmes administratifs, ainsi qu'à Olja, Jean-Luc, Roseline et Tina.

Pendant ces quatre ans, il y a eu également beaucoup de départs du laboratoire. Maylan, Chloé, Dani, Oliver et Claire, votre présence ici a été marquée par votre bienveillance et votre bonne humeur. En même temps, nous avons eu le plaisir d'accueillir des stagiaires où chacun a laissé sa propre note. Nabila, Violette, Margaux, Émilie, merci de m'avoir permis d'évoluer en tant que jeune encadrante. Etienne et Nhutien, merci pour toutes les petites blagues et les moments conviviaux que nous avons eu ensemble.

Passons de l'autre côté du couloir, à l'équipe de Processus Dégénératifs, Stress et Vieillesse. Je voudrais remercier Manon, Janany et Maria pour leurs grand sourire, les petites pauses café et les 'cake party' que nous avons fait ensemble ! Manon, mon 'petit chou', nous sommes liées maintenant par notre séjour au forum de l'école doctorale. Je suis

contente que cela m'ait permis de trouver une amie à ton visage et je te souhaite la 'vie en rose' comme cela te décrit. Ma petite Janany, je te souhaite de trouver ton chemin vers là où tu seras heureuse, mais dans tous les cas je sais que t'auras un future brillant dans l'art de la coiffure. Maria, merci pour toutes les petites discussions de couloir, de tes bonnes ondes et de ta bonne humeur qui font du bien.

Je voudrais également remercier Marie, Guillaume et Angélique du centre Epigénétique et Destin Cellulaire. Marie ou autrement 'bella', je suis honorée d'avoir partagé ce bout de chemin de la vie avec toi. Je ne pourrai jamais oublier nos fous rires du vendredi midi autour d'un bon burger. J'admire énormément ta détermination à suivre ce que tu veux. Mon plus grand regret, restera Singapore, mais qui sait, peut-être qu'au futur nous aurons une opportunité pareil... Guillaume, ou 'Jean-Gui', merci pour toute ton aide avec les manips de radioactivité. C'est un énorme plaisir et une source de motivation de voir des gens passionnés par la recherche comme toi ! Ma petite Angélique, malgré ton arrivée récente, nous nous sommes bien rapprochées grâce à ta bienveillance. J'ai hâte pour la suite de notre manip avec les inhibiteurs et surtout de notre excursion Disney.

Le trajet de ma thèse ne s'arrête pas là. Pendant une grande période, j'ai travaillé avec la plate-forme de Cristallographie de l'Institut Pasteur. Un grand merci à Ahmed, pour tous ces conseils, sa bienveillance, et la patience qu'il a montré à moi et mon impatience. Ahmed, merci pour tout le travail et le temps que tu m'as consacré. Merci de faire également partie de mon jury. Je voulais absolument que tu sois présent, car c'est toujours enrichissant de discuter avec toi. Bien évidemment, merci à Ariel pour tout son temps pour la résolution de mes structures cristallographiques, à Fred pour son énorme gentillesse et intérêt, ainsi qu'à Cédric et Patrick, les deux mascottes de la plateforme, qui ont rendu agréables et mémorables mes visites. Je suis impatiente de vous voir l'année prochaine pour des nouvelles aventures.

Petit retour à l'Unité de Biologie Fonctionnelle et Adaptative pour remercier Leslie. Leslie, grâce à ton travail nous avons pu analyser au mieux les structures cristallographiques que nous avons résolues. Notre travail ne s'arrête pas là, ce qui me rend particulièrement

heureuse car ta bonne humeur et ton sens de l'humour te rendent agréable. Merci également à Delphine pour les analyses de dynamique moléculaire et sa bienveillance.

Une grande partie de cette thèse a été faite suite au travail de Dr. Haopeng Yang qui a permis de modifier des cellules HEK293T par la technologie CRISPR/Cas9. Même si nous ne nous sommes jamais rencontrés personnellement, je voudrais le remercier. Je voudrais également remercier le Dr. Michael R. Green qui a permis la réalisation de cette partie.

Pendant ma dernière année de thèse, j'ai eu aussi l'opportunité et le plaisir de faire des expériences à l'Institut de Recherche Saint-Louis. Je voudrais remercier le Dr. Nicolas Duplhy pour m'avoir accueilli dans son laboratoire, et Valeria pour toute son aide, son temps et sa bonne humeur. Merci également à Lin-Pierre et Emilie pour leur accueil chaleureux et les belles discussions que nous avons eu ensemble. Fabien, merci pour toute ton aide avec mes expériences de ChIP. Celine, ma petite voisine libanaise, ça me fait très plaisir le fait que nous avons commencé par être deux inconnues qui manipulaient dans la même pièce et qu'aujourd'hui nous sommes amies.

Je voudrais aussi remercier Magalie du Collège de France pour son aide avec le FACS, son professionnalisme et sa grande volonté, ainsi que le Dr. Véronique Arluison de CEA pour tout son intérêt et son aide avec mon expérience de dichroïsme circulaire.

Ma thèse a été aussi l'opportunité d'enseigner des travaux pratiques de biologie moléculaire et de génétique. Je voudrais remercier le Dr. India Leclercq de m'avoir donné cette occasion. Un grand merci à Nathalie pour toute l'organisation exemplaire des TP, et également à Samia et Sandrine, le petit duo de folie avec qui j'ai eu plaisir de rigoler, et échanger.

Une grande leçon de ma vie pendant cette période de quatre années a été qu'il faut trouver un bon équilibre entre la vie professionnelle et personnelle. J'avais pour longtemps la tendance de regarder que mes études et mon côté carriériste, jusqu'au moment où toi, mon chéri, mon Loïc tu m'as appris qu'il faut un ensemble pour une vie heureuse. Merci d'être

mon compagnon dans la vie de tous les jours et d'être là pour tous les moments où j'ai besoin de toi. Tu as été mon plus grand admirateur et soutien. Grâce à toi et toute ton aide pour les petits et les grands détails, j'ai pu avancer sur ma thèse de manière sereine. Merci d'être à l'écoute quand je me plain tout le temps des problèmes de mes manips, et pas que. Merci d'être là pour fêter tous les événements de bonheur de ma courte carrière.

À mes côtés durant cette aventure, mes amis, qui étaient là pour me changer les idées et pour passer de bons moments ensemble. Médina, tu as été pour moi la personne vers qui je pouvais compter pour quoi que ce soit. Je ne vais jamais oublier ni les petites excursions chez toi à Strasbourg qui étaient une vraie détente, ni ta mission pompier à Chypre. Laëtitia, nos petits verres et nos moments de bien-être étaient une oasis pendant cette période. Nos discussions de filles étaient ce dont j'avais vraiment besoin pour déconnecter. Mes petits chatons, Léa et Quentin, merci pour toutes vos ondes positives et vos grands sourires chaque fois que l'on se voyait. Sarah, t'as une capacité unique de me transmettre ton énergie de folie, dont j'avais besoin. Merci pour tous les amis !

Je voudrais également remercier ma famille qui a cru en moi et qui m'a toujours rempli d'amour. Mention spéciale à mes deux grands-mères, Χριστίνα et Πόπη, mon grand-père Σάββα, et à ma tante Βάσω, que j'espère avoir rendu fières. Merci pour nos appels qui m'ont permis de reprendre des forces et continuer. Un grand merci aussi à ma belle-famille, surtout à Daniel et Françoise qui ont été des grands supporteurs de mon travail. Daniel, merci pour votre gentillesse et Françoise, merci pour ton humanité et ton aide avec la relecture de mon français.

En dernier lieu, je voudrais remercier ma famille proche, mes parents Κατερίνα et Ιάσωνα et ma petite sœur Σαββιάννα. 'Σας αγαπώ και σας ευχαριστώ για όλα' ! Maman, papa, merci pour tous les sacrifices que vous avez fait pour moi depuis ma naissance. Je pense que je ne pourrai jamais vous retourner tout ça, mais j'espère que ce travail vous rendra fiers de votre fille. Merci de m'avoir supporté durant toutes les épreuves de ma vie et de de votre amour inconditionnel. Σαββιάννα, merci d'enseigner ma vie avec ton petit caractère et ton grand sourire. Je suis la plus fière de toi et je suis sûre qu'un jour tu feras du bien dans ce monde ! Merci à tous les trois d'être mes piliers de la vie.

Un dernier merci, le plus difficile pour moi, à mon petit frère Γιώργο qui malheureusement n'est plus là. La pensée envers toi me donne du courage chaque fois que j'en ai besoin. Tu es la raison pour laquelle je fais de la science aujourd'hui car grâce à toi j'ai appris que quand nous n'avons plus de solutions, nous avons toujours l'espoir par la recherche. Merci de m'avoir mis dans cette voie mais surtout, merci de veiller à moi.

Cette thèse est un hommage pour toi mon frère...

LISTE DES PUBLICATIONS

Berthelet J, Michail C, Bui LC, Le Coadou L, Sirri V, Wang L, Dulphy N, Dupret JM, Chomienne C, Guidez F, Rodrigues-Lima F. The Benzene Hematotoxic and Reactive Metabolite 1,4-Benzoquinone Impairs the Activity of the Histone Methyltransferase SET Domain Containing 2 (SETD2) and Causes Aberrant Histone H3 Lysine 36 Trimethylation (H3K36me3). **Mol Pharmacol.** 2021 Sep;100(3):283-294. doi: 10.1124/molpharm.121.000303

Zhang W, Berthelet J, Michail C, Bui LC, Gou P, Liu R, Duval R, Renault J, Dupret JM, Guidez F, Chomienne C, Rodrigues Lima F. Human CREBBP acetyltransferase is impaired by etoposide quinone, an oxidative and leukemogenic metabolite of the anticancer drug etoposide through modification of redox-sensitive zinc-finger cysteine residues. **Free Radic Biol Med.** 2021 Jan;162:27-37. doi: 10.1016/j.freeradbiomed.2020.11.027

Xu X, Zhang W, Berthelet J, Liu R, Michail C, Chaffotte AF, Dupret JM, Rodrigues-Lima F. From transglutaminases (TGs) to arylamine N-acetyltransferases (NATs): Insight into the role of a spatially conserved aromatic amino acid position in the active site of these two families of enzymes. **Biochem Biophys Res Commun.** 2020 Apr 30;525(2):308-312. doi: 10.1016/j.bbrc.2020.02.082

ABBREVIATIONS

ac: groupement acétyl

AcCoA: AcetylCoA

ALL: Acute Lymphoblastic Leukemia

AML: Acute Myeloid Leukemia

AWS: Associated With Set

BRPF1: BRomodomain and PHD Finger containing 1

CBP: CREB Binding Protein

CCND1: G1/S-specific cyclin-D1

ccRCC: clear cell Renal Cell Carcinoma

ChIP: Chromatin ImmunoPrecipitation

COSMIC: Catalogue Of Somatic Mutations In Cancer

CPK: Corey Pauling Koltun

DMEM: Dulbecco's Modified Eagle Medium

DMSO: DiMethyl SulfOxide

DTT: DiThioThreitol

EZH2: Enhancer - of - zest homolog 2

GFP: Green Fluorescent Protein

H3K14: residu lysine 14 de l'histone H3

H3K23: residu lysine 23 de l'histone H3

H3K36: residu lysine 36 de l'histone H3

HAT: Histone AcetylTransferase

HCl: acide chlorhydrique

HClO₄: acide perchlorique

hMOF: human Males absent On the First

HMT: Histone MethylTransferase

HPLC: High Pressure Liquid Chromatography

IgG: ImmunoGlobulin G

IPTG: IsoPropyl β -D-1-ThioGalactopyranoside

KAT: Lysin AcetylTransferase

kcat: constante catalytique

K_M: constante de Michaelis

KMT: Lysine MethylTransferase

LB: Luria Bertani

LC-MS: Liquid Chromatography – Mass Spectrometry

L_{IN} : boucle auto-inhibitrice

me: groupement methyl

MMR: DNA MisMatch Repair

MOZ: MOnocytic leukemia Zinc finger

MSH6: MutS Homolog 6

MYST: MOZ, YBF2, SAS2, Tip60

Ni-NTA: Nickel-NiTrilotriacetic Acid-agarose chromatography

PAR: 4-2-Pyridylazoresorcinol

PBS: Phosphate Buffered Saline

PBST: Phosphate Buffered Saline Tween

PDB: Protein Data Bank

PRC2: Polycomb Repressive Complex 2

PWWP: Pro - Trp - Trp - Pro

q-PCR: quantitative Polymerase Chain Reaction

RFP: Red Fluorescent Protein

RMSF: Root Mean Square Fluctuation

RMSD: Root Mean Square Deviation

RNA pol II: RNA polymerase II

rpm: rotations per minute

RPMI: Roswell Park Memorial Institute

SAH: S-adénosylhomocysteine

SAM: S-adenosylmethionine

SDS: Sodium Dodecyl Sulfate

SDS-PAGE: Sodium Dodecyl Sulfate-PolyAcrylamide Gel Electrophoresis

SET: Su(var) (Suppressor of Variegation), E(z) (Enhancer of Zeste) et Trx (Trithorax)

SETD2: SET Domain containing 2

SRI: Set2 Rpb1 Interacting

SVF: Sérum de Veau Fœtal

TCA: TriChloroacetic Acid

T_M: Melting Temperature

TSA: Thermal Shift Assay

Ub: Ubiquitin

WB: Western Blot

Figures et Tableaux

INTRODUCTION	29
--------------------	----

I. LA MARQUE EPIGENETIQUE H3K36ME3	31
--	----

Figure I – 1 : Méthylation des histones par les HMT..... 32

Figure I – 2 : Disposition des stades de méthylation d'H3K4 au sein d'un gène à transcription active..... 33

Tableau I – 1 : HMT de H3K4, leurs rôles principaux et des maladies associées..... 33

Tableau I – 2 : Les HMT d'H3K9 et des maladies associées à leurs fonctions..... 34

Figure I – 3 : Méthylation d'H3K9 dans l'hétérochromatine et l'euchromatine..... 35

Figure I – 4 : Disposition des trois stades de méthylation d'H3K27 au sein des gènes silencieux et actifs 36

Figure I – 5 : Processivité de méthylation d'H3K36 par les huit HMT des mammifères 38

Tableau I – 3 : HMT majoritaires d'H3K36 et des maladies associées à leurs fonctions 38

Figure I – 6 : DOT1L, un cas particulier d'HMT 39

Figure I – 7 : Mécanisme séquentiel de méthylation d'H3K9 et d'H4K20 par l'intermédiaire d'HP-1..... 40

Figure I – 8 : Enzymes responsables de la transition entre la marque H3K36me2 et la marque H3K36me3 41

Figure I – 9 : Mécanisme schématique de déméthylation du résidu H3K36 triméthylé par JMJD2 43

Figure I – 10 : Isomérisation cis-trans de la Proline 38 de l'histone 3 43

Tableau I – 4 : Domaines des protéines 'readers' et lysines méthylées reconnues..... 45

Figure I – 11 : Reconnaissance d'H3K36me3 par le domaine PWWP de BRPF1 46

Figure I – 12 : Reconnaissance d'H3K36me3 par le domaine Tudor de PRF1..... 48

Figure I – 13 : Représentation simplifiée de la régulation de la transcription par la marque H3K36me3.. 50

Figure I – 14 : Mécanisme de réparation DSB par la signalisation d'H3K36me3..... 53

Figure I – 15 : Communication (« cross-talk ») de la marque H3K36me3 avec la méthylation/ acétylation des histones et la méthylation d'ADN 54

Figure I – 16 : Représentation simplifiée de l'effet d'H3K36M sur les marques d'H3K36me3 et H3K27me3 56

II. SET DOMAIN CONTAINING II, SETD2	57
---	----

Tableau II – 1 : Familles des HMT à domaine SET et lysine de substrat histone..... 58

Figure II – 1 : Représentation schématique des domaines catalytiques des 7 familles des HMT à domaine SET 59

Figure II – 2 : Les domaines de SETD2..... 60

Figure II – 3 : Les domaines à doigts de zinc structuraux de SETD2, AWS et Post-SET	62
Figure II – 4 : L'architecture du domaine catalytique SET	63
Figure II – 5 : Autorégulation de l'interaction de SETD2 avec ses partenaires, par son domaine WW	65
Figure II – 6 : Le domaine SRI de SETD2 et les résidus importants pour l'interaction avec la RNA pol II	66
Figure II – 7 : Fixation de ligands SAH et substrat peptidique modifié (H3K36M) au domaine SET	68
Tableau II – 2 : Liaisons Hydrogène directes et interactions hydrophobes entre les résidus de SETD2 et du peptide H3K36M.....	70
Figure II – 8 : Le complexe entre SETD2, nucléosome et le cofacteur SAH	71
Figure II – 9 : Conformation 'ouverte' et 'fermée' de la boucle L _{IN} de SETD2	72
Figure II – 10 : Les substrats non-histone de SETD2 et les fonctions associées à leur méthylation	74
Figure II – 11 : Représentation simplifiée de la réponse immunitaire médiée par la signalisation INF α – STAT1, par l'intermédiaire de SETD2.....	78
Figure II – 12 : Setd2, un gène moteur de cancérogenèse	79
Figure II – 13 : Mutations de Setd2 et leurs effets cellulaires lors de l'avancement des stades de ccRCC... 81	
Figure II – 14 : Mutations de SETD2 dans les hémopathies.....	82
Figure II – 15 : Modèle du processus de résistance des cellules aux traitements chimiothérapeutiques....	83
III. MONOCYTIC LEUKEMIA ZINC FINGER, MOZ	85
Figure III – 1 : Mécanisme simplifié d'acétylation des histones par des HAT avec ouverture de la chromatine.....	85
Figure III – 2 : Structures du domaine catalytique HAT des hGCN5 et CPAF.....	87
Figure III – 3 : Mécanisme d'activité des CBP/p300.....	88
Figure III – 4 : Mécanisme catalytique de type ping-pong des MYST	89
Tableau III – 1 : Les MYST humaines en tant que sous-unités catalytiques des complexes multi-protéiques	91
Tableau III – 2 : Les membres de la famille MYST retrouvés chez l'Homme et leurs fonctions.....	91
Figure III – 5 : Alignement des séquences des cinq MYST humaines.....	92
Figure III – 6 : Les cinq domaines de MOZ.....	93
Tableau III – 3 : Domaines de MOZ et leurs fonctions	94
Figure III – 7 : Structure du domaine DPF, caractéristique des MYST	95
Figure III – 8 : Architecture du domaine MYST de MOZ	96
Figure III – 9 : Architecture catalytique de MOZ.....	97
Figure III – 10 : Auto-acétylation de MOZ	99
Figure III – 11 : Représentation simplifiée des mécanismes d'activation/inhibition de senescence par MOZ	101
Figure III – 12 : Représentation simplifiée des mutations non sens et faux sens liées au syndrome KAT6A	102
Figure III – 13 : Principales mutations de chaque domaine, enregistrées dans la banque de COSMIC....	103

Tableau III – 4 : Distribution des mutations du gène *MOZ*, enregistrées dans la banque des données des mutations somatiques aux cancers (COSMIC) 104

Figure III – 14 : Dérégulation de l'expression de *MOZ* dans les cancers..... 106

Tableau III – 5 : Les principales translocations chromosomiques de *MOZ*..... 108

CARACTERISATION STRUCTURALE ET FONCTIONNELLE D'UN MUTANT ONCOGENIQUE DE SETD2..... 113

Figure S-1 : Mutations de *SETD2* identifiées chez des patients atteints d'ALL..... 116

Figure S-2 : Alignement des séquences des majeurs HMT chez l'Homme, expression de *SETD2* WT et *SETD2* L1609P recombinantes et détermination de leur activité méthyltransférase et de leur thermostabilité 119

Figure S-3 : Analyse de la marque H3K36me3 et détermination de la stabilité de GFP-*SETD2* WT et GFP-*SETD2* L1609P, transfectée dans des cellules HEK293T KO *SETD2*..... 123

Figure S-4 : Analyse de la marque H3K36me3, détermination de la stabilité de *SETD2* endogène et résistance au traitement à l'étoposide des cellules HEK293T CRISPR *SETD2* WT et *SETD2* L1609P 126

Figure S-5 : Comparaison globale des structures cristallographiques du domaine catalytique de *SETD2* WT et de *SETD2* L1609P et études de dynamique moléculaire 129

Tableau S – 1 : Statistiques de collecte et d'affinement des données de la structure cristallographique du domaine catalytique de *SETD2* L1609P 130

Figure S-6 : Effets de la mutation L1609P de *SETD2* sur les résidus environnants..... 132

Figure S-7 : Effets fonctionnels et structuraux de la mutation L1609P sur *SETD2* 139

Tableau S – 2 : Liste des différents substrats utilisés pour le dosage de l'activité HMT de *SETD2* recombinante 144

Tableau S-3 : Anticorps utilisés pour la détection par Western Blot..... 145

Figure S-1 Supplémentaire : Effet de la mutation *SETD2* L1609P sur la fixation du cofacteur SAM et du peptide substrat H3M36, et effet de la température sur les doigts de zinc de l'enzyme mutante 159

Figure S-2 Supplémentaire : Expression de l'ARNm *SETD2* WT et L1609P après transfection dans les cellules HEK293T *SETD2* KO 160

Figure S-3 Supplémentaire : Effet du traitement à l'étoposide sur le cycle cellulaire des cellules HEK293T CRISPR *SETD2* WT et HEK293T CRISPR *SETD2* L1609P..... 161

CARACTERISATION STRUCTURALE ET FONCTIONNELLE D'UN MUTANT ONCOGENIQUE DE MOZ 163

Tableau M – 1 : Mutations faux sens les plus représentées dans le domaine MYST de *MOZ* 165

Figure M-1 : Alignement des séquences des MYST, expression de MOZ WT et MOZ T605P recombinants, détermination de leur capacité de liaison à l'AcétylCoA et de leur activité lysine acétyltransférase.....	167
Tableau M – 2 : Statistiques de collecte et d'affinement des données de la structure cristallographique du domaine catalytique de MOZ T605P.....	169
Figure M-2 : Alignement structural des MOZ WT et MOZ T605P, effet de la mutation sur le site actif et l'accessibilité du substrat	170
Figure M-3 : Proposition d'un modèle simplifié de la "perte de fonction" du mutant MOZ T605P.....	173
Tableau M – 3 : Les différents substrats utilisés pour le dosage de l'activité HAT de MOZ recombinante	181
Tableau M – 4 : Conditions d'HPLC selon le peptide substrat utilisé	181
Tableau M – 5 : Anticorps utilisés pour la détection par Western Blot.....	182
Figure M-1 Supplémentaire : Purification par Ni-NTA et clivage de l'étiquette 6-histidine par thrombine, suivi par purification par chromatographie d'exclusion pour des essais de cristallogénèse.....	188
Figure M-2 Supplémentaire : Effet de la mutation MOZ T605P sur la stabilisation de la lysine auto-acétylée et sur le tunnel de l'accessibilité du substrat au niveau du site actif.....	189
Figure M-3 Supplémentaire : Modélisation de la fixation du substrat H4K16CoA dans les structures MOZ WT et MOZ T605P.....	190
Figure M-4 Supplémentaire : Effet du traitement à l'étoposide sur l'induction de sénescence	191

INTRODUCTION

I. La marque épigénétique H3K36me3

A. Méthylation des histones

La méthylation des histones est un processus complexe, permettant selon l'emplacement de cette modification dans le génome, la régulation de la transcription dans les locus euchromatiniens ou l'inhibition de la transcription dans les régions de l'hétérochromatine. Les cibles de méthylation des histones sont les résidus basiques : la lysine (K) qui peut avoir trois états de méthylation (mono-, di-, tri-, Fig.I-1A) et l'arginine (R) qui, elle, peut être mono- ou di- méthylée (symétrique ou asymétrique, Fig.1B). Cette réaction s'effectue grâce à des enzymes appelées Histone MéthylTransférases (HMT) ou Lysine MéthylTransférases (KMT), qui utilisent comme cofacteur le donneur universel de méthyl S-AdénosylMéthionine (SAM).

La particularité de la méthylation par rapport aux autres modifications covalentes, comme l'acétylation et la phosphorylation, est le fait qu'elle ne change pas la charge de sa cible. En conséquence, les résidus basiques restent chargés positivement et les interactions avec les groupements phosphates de l'ADN ne sont pas déstabilisées. Cela suggère que la forme compacte de la chromatine ne soit pas déstabilisée par des interactions électrostatiques, mais probablement par des interactions hydrophobes. Néanmoins, à l'image de l'acétylation et de la phosphorylation, la méthylation permet le recrutement et la liaison des protéines avec des fonctions en aval. En effet, les modifications épigénétiques servent de signaux et sont perçues par des acteurs cellulaires, ayant comme résultat la 'communication' entre les différentes marques de la chromatine et les processus physiologiques. Chaque état de méthylation est considéré comme un signal différent ¹⁻³.

À ce jour, la méthylation des lysines est mieux documentée que celle des arginines. Les lysines méthylées les mieux connues se situent sur l'histone 3 (H3) et l'histone 4 (H4). Les lysines méthylables le mieux décrites dans l'histone H3 sont les résidus K4, K9, K27, K36, K79

et dans l'histone H4 le résidu K20. Les arginines méthylables sont les acides aminés R2, R17, R26 d'H3 et la R3 d'H4. Selon la combinaison et l'état de méthylation de chacun de ces résidus, la cellule adopte un niveau différent de régulation transcriptionnelle. En général, les méthylations de résidus H3K4, H3K36 et H3K79 sont considérées comme des marques activatrices de la transcription, contrairement aux méthylations des résidus H3K9, H3K27 et H4K20 qui sont des marques répressives ⁴.

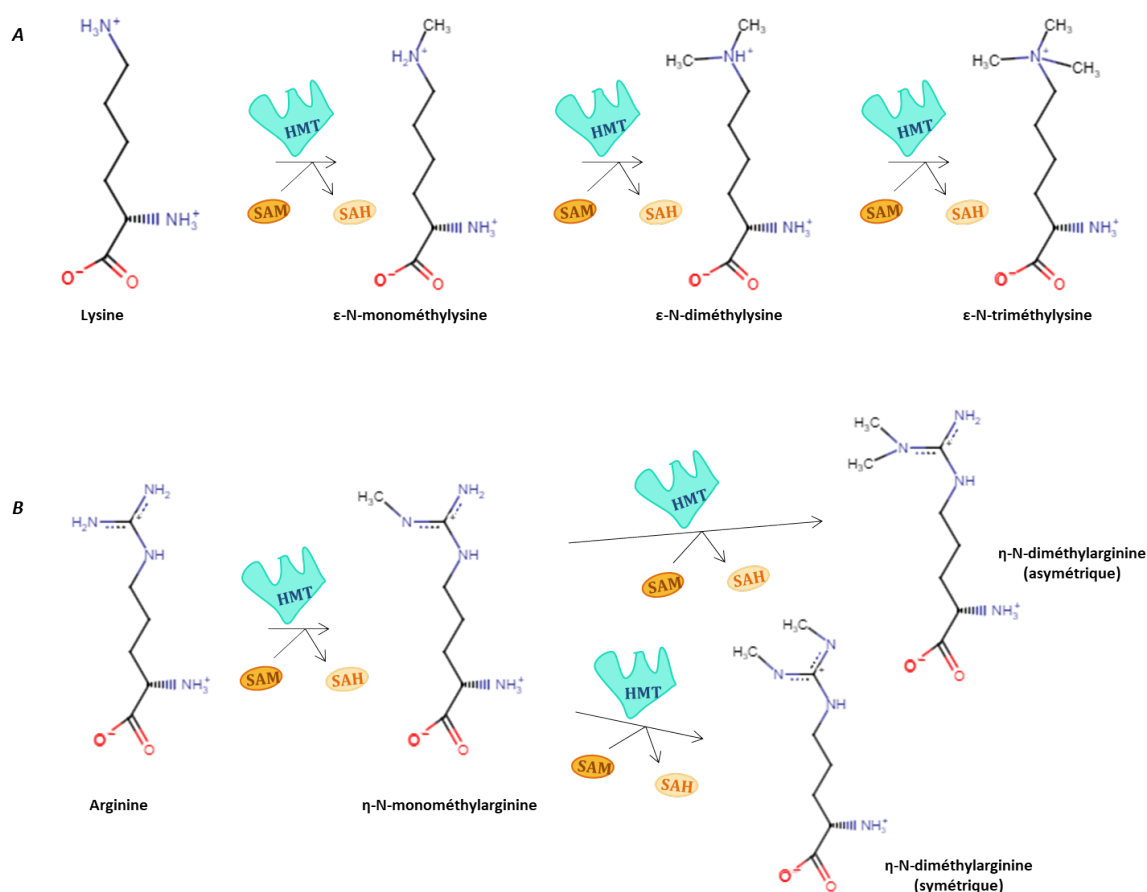


Figure I – 1 : Méthylation des histones par les HMT

A. Mécanisme séquentiel de méthylation des lysines par les HMT

B. Méthylation des arginines par les HMT

Les résidus basiques (K ou R) des histones peuvent être méthylés grâce aux HMT. Ces enzymes utilisent comme cofacteur la SAM (S-adénosylméthionine), qui donne son groupement méthyle sans changer la charge du substrat. Les produits libérés sont le métabolite de la SAM, la SAH (S-adénosylhomocystéine) ainsi que le substrat histone méthylé. La lysine peut être mono-, di- ou tri- méthylée sur le même atome d'azote ζ . L'arginine peut être diméthylée en portant un groupement méthyle à chacun de ces deux azotes η , générant un produit symétrique, ou en portant les deux groupements méthyle sur le même azote η , formant ainsi un produit asymétrique.

1) Méthylation du résidu H3K4

La méthylation de la lysine 4 d'H3 est une marque des histones liée à l'activation de la transcription. Les stades de méthylation, H3K4me1, H3K4me2, H3K4me3 marquent de manière différente les gènes à transcription active. La marque H3K4me1 est enrichie au niveau des enhanceurs, la marque H3K4me2 est majoritaire au niveau du 5' des gènes actifs, alors que la marque H3K4me3 se situe au niveau du promoteur de ces gènes (Fig.1-2) ⁵. En outre, cette méthylation est liée à l'acétylation de l'histone H3 (H3ac), mais aussi à des régions d'ADN non méthylées au niveau des promoteurs (îlots CpG). Plus particulièrement, la triméthylation du résidu H3K4 permet de créer un environnement favorable pour la communication avec la machinerie transcriptionnelle ⁶. Le Tableau I-1 récapitule les HMT responsables de la marque H3K4 méthylée, ainsi que leurs principales fonctions et les types des pathologies dans lesquels elles sont retrouvées ⁴.

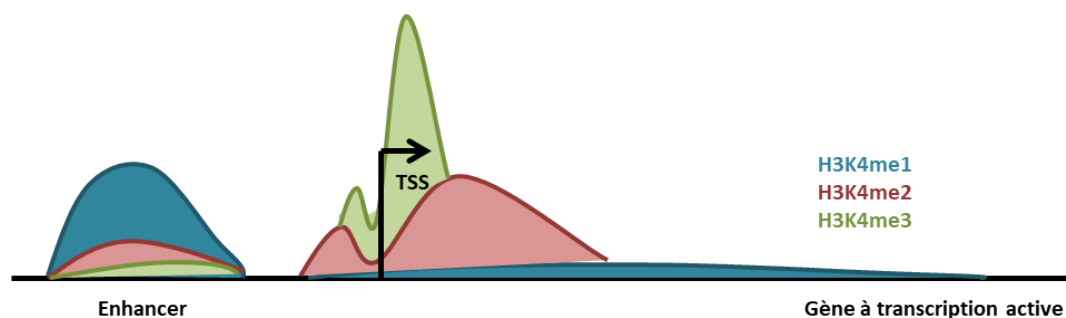


Figure I – 2 : Disposition des stades de méthylation d'H3K4 au sein d'un gène à transcription active ⁶

TSS : Transcription Start Site

HMT/ KMT	Rôle	Maladie Associée
hSET1A (SET domain containing 1A)/ KMT2F	Remodelage de la chromatine	Hépatocarcinogenèse / Epilepsie
hSET1B/ KMT2G	Remodelage de la chromatine	Développement Intellectuel
MLL1 (Mixed Lineage Leukemia 1)/ KMT2A	Sous unité du facteur de transcription II (TFII)	Leucémie / Syndrome Wiedemann-Steiner
MLL2/ KMT2B	Dialogue K4me2/3 et H3-Ac	Dystonie
MLL3/ KMT2C	Arret du cycle cellulaire (G/M)	Syndrome Kleeftstra
MLL4/ KMT2D	Liaison aux îlots CpG	Syndrome Kabuki/ Lymphomagenèse

Tableau I – 1 : HMT de H3K4, leurs rôles principaux et des maladies associées ⁴

2) Méthylation du résidu H3K9

La méthylation du résidu H3K9 est une caractéristique de l'hétérochromatine et de la transcription silencieuse dans l'euchromatine. La marque H3K9me3 est typique de l'hétérochromatine péricentromérique, permettant la ségrégation adéquate des chromosomes lors de la mitose. Par ailleurs, les marques H3K9me1/2 sont trouvées dans les régions géniques non transcrites de l'euchromatine, permettant la méthylation de l'ADN au niveau des îlots CpG, une autre marque épigénétique de répression de la transcription. La méthylation d'H3K9 dans l'euchromatine permet la création d'un environnement local d'hétérochromatine ³. Selon la localisation de la lysine cible, plusieurs enzymes peuvent intervenir.

Dans l'hétérochromatine, les trois stades de méthylation sont contrôlés par plusieurs HMT (Fig.1-3). SETDB1 (SET Domain Bifurcated 1 ou KMT1E) est responsable de la monométhylation d'H3K9, permettant ainsi de créer la marque H3K9 monométhyl qui sert comme substrat aux SUV39H1/2 (Suppressor of Variegation 3-9 Homolog 1/2 ou KMT1A/B) pour la formation des marques H3K9 di- et tri- méthyl ^{7,8}. Dans l'euchromatine, l'hétérodimère G9a (KMT1C) – GLP (G9a Like Protein ou KMT1D) conduit à la formation de la marque H3K9 mono- et di- méthyl et à la répression de la transcription active. La dérégulation de ces enzymes est associée à de nombreuses maladies neurodégénératives, dont la maladie d'Alzheimer et d'Huntington, et des cancers où G9a est souvent surexprimée, induisant l'hyperméthylation des gènes suppresseurs des tumeurs et l'inhibition de leur expression (Tableau I-2) ⁴.

HMT/ KMT	Maladie Associée
SETDB1/ KMT1E	Maladie d'Huntington
SUV39H1/2 / KMT1A/B	Instabilité Chromosomique
G9a / KMT1C	Cancer du sein
GLP / KMT1D	Syndrome Kleeftstra

Tableau I – 2 : Les HMT d'H3K9 et des maladies associées à leurs fonctions ^{4,9}

Tableau non exhaustif des maladies, dont des cancers, dans lesquels les HMT d'H3K9 sont retrouvées modifiées (mutations ou niveau d'expression altéré).

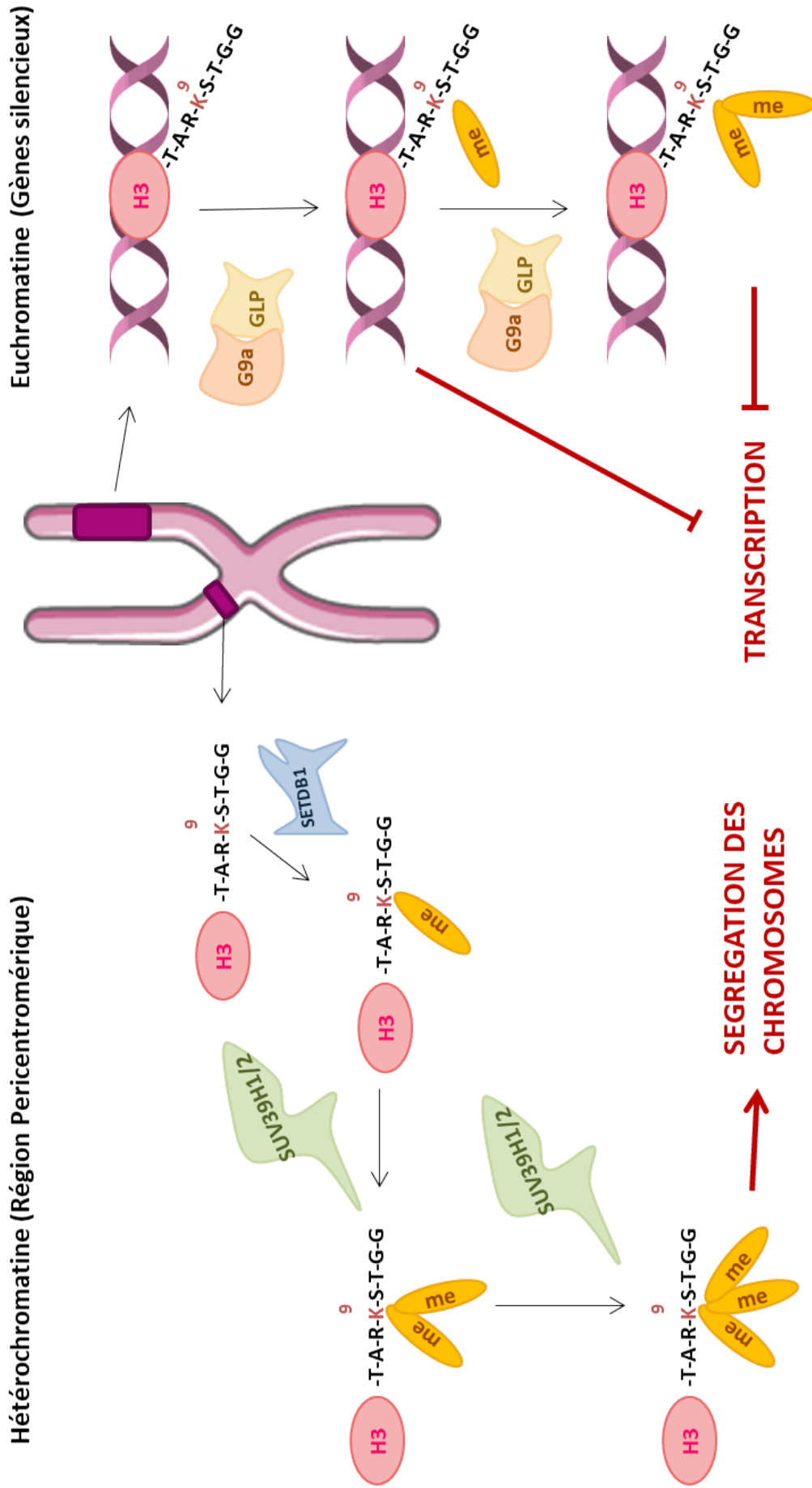


Figure 1 – 3 : Méthylation d'H3K9 dans l'hétérochromatine et l'euchromatine

Dans le cas de l'hétérochromatine à la région pericentromérique, SETDB1 et SUV39H1/2 permettent la formation de la marque H3K9me3 qui lui est spécifique. En outre, dans la chromatine au niveau des gènes silencieux, l'hétérodimère G9a/GLP forme les H3K9me1/2, responsables de l'inhibition de la transcription.

3) Méthylation du résidu H3K27

Les trois étapes de méthylation du résidu H3K27 (mono, di, tri) sont issues de l'activité d'EZH2 (Enhancer - of - zest homolog 2/ KMT6A ou, en moindre partie de son homologue EZH1), la sous unité catalytique du complexe PRC2 (Polycomb Repressible Complex 2) ¹⁰. EZH2, incapable de se lier aux nucléosomes, présente une faible activité de méthyltransférase quand elle ne fait pas partie de PRC2 ¹¹. Ce complexe est composé de trois autres protéines qui sont les SUZ12 (SUppressor of Zeste 12 protein homolog), EED (Embryonic Ectoderm Development protein) et RBBP4/7 (RetinoBlastoma – Binding – Protein 4/7). Les SUZ12 et RBBP4/7 permettent au complexe de se lier aux nucléosomes, alors que EED contribue non seulement à cette liaison, mais augmente également l'activité d'EZH2 ¹¹.

La marque d'H3K27 méthylée a un rôle différent selon son niveau de méthylation, mais en globalité elle a une fonction de répression de la transcription liée à la marque H3K27me3. La triméthylation d'H3K27, responsable de l'inhibition transcriptionnelle, ne peut pas coexister avec des marques activatrices de la transcription comme la marque H3K4me3, alors qu'elle peut être présente en parallèle de la marque H3K9 méthylée. Au contraire, la marque H3K27me1 est présente dans les gènes en cours de transcription et elle coexiste avec la triméthylation du résidu H3K36, quant à la marque H3K27me2, elle est très répandue dans la globalité d'H3 (environ 70%), suggérant un rôle préventif de l'activation anormale de certains enhancers (Fig.1-4) ¹². De ce fait, la mauvaise régulation de la méthylation du résidu H3K27 est associée avec la cancérogenèse et les métastases, notamment par hyper expression ou hyper activation d'EZH2.

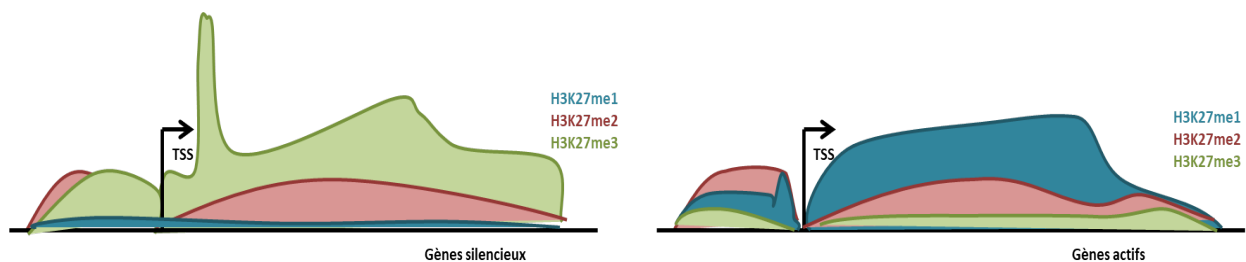


Figure 1 – 4 : Disposition des trois stades de méthylation d'H3K27 au sein des gènes silencieux et actifs ¹³

4) Méthylation du résidu H3K36

À ce jour, chez les mammifères huit HMT sont connues comme étant responsables de la méthylation de la lysine 36 de l'histone H3 (Fig.1-5). NSD1 (Nuclear receptor SET Domain – containing 1 ou KMT3B), NSD2 (ou KMT3G), NSD3 (KMT3F), ASH1L (Absent Small and Homeotic doscs 1 - Like ou KMTH2), SMYD2 (SET and MYND domain - containing 2 ou KMT3C), SETD3 (SET Domain – containing 3) et SETMAR (SET domain and MARiner transposase fusion gene – containing) sont décrites comme les méthyltransférases capables de mono- et di- méthyler le résidu H3K36. Parmi elles, les mieux décrites sont les NSD1/2/3, qui malgré le fait qu'elles partagent le même site de méthylation, elles possèdent des gènes cibles différents. La seule enzyme apte à triméthyler la lysine 36 est SETD2 (SET Domain – containing 2 ou KMT3A) ^{4,14}.

En général, la méthylation de l'acide aminé H3K36 est associée à la transcription active de l'euchromatine, permettant l'élongation de la transcription, avec un pic de di- et tri-méthylation en 3' des régions en cours de transcription. Cela suggère que cette modification possède un rôle à la fin de la transcription ou potentiellement aux stades précoces des modifications de l'ARNm ¹⁵. La présence de la marque H3K36 méthylée a été également identifiée dans les régions promotrices de plusieurs gènes avec un rôle antagoniste de l'activité inhibitrice du complexe PRC2, permettant ainsi de supporter la notion de communication ('cross-talk') entre les différentes modifications épigénétiques ¹⁶.

Par ailleurs, cette marque épigénétique est pourvue d'un rôle dans le contrôle de la réplication d'ADN, mais la machinerie exacte de cet effet reste inconnue. En outre, la méthylation du résidu H3K36 est liée aux réponses aux dommages de l'ADN. Dans le cas des cassures double brins (DSB – 'Double Strand Breaks), la marque H3K36me2 facilite la jonction des extrémités non homologues (NHEJ – 'Non Homologous end – joining') ¹⁷, mais la marque H3K36me3 induit dans les régions de mésappariement de l'ADN, la réparation par la machinerie appropriée (MMR – DNA MisMatch Repair) également ¹⁸.

Des altérations de la méthylation du résidu H3K36 sont associées à plusieurs maladies. Dans le cas des patients atteints de cancers, des mutations faux-sens sont retrouvées sur les résidus de l’histone H3, dont la lysine 36, mutée le plus souvent en méthionine (H3K36M) ou en isoleucine (H3K36I). Ces mutations inhibent la formation de la marque H3K36 méthylée et de la dérégulation cellulaire dépendante de ces marques^{19,20}. De même, des mutations des enzymes responsables de la méthylation du résidu H3K36 ont été identifiées dans différentes pathologies, dont les principales maladies associées à ces mutations, sont listés au Tableau I-3.

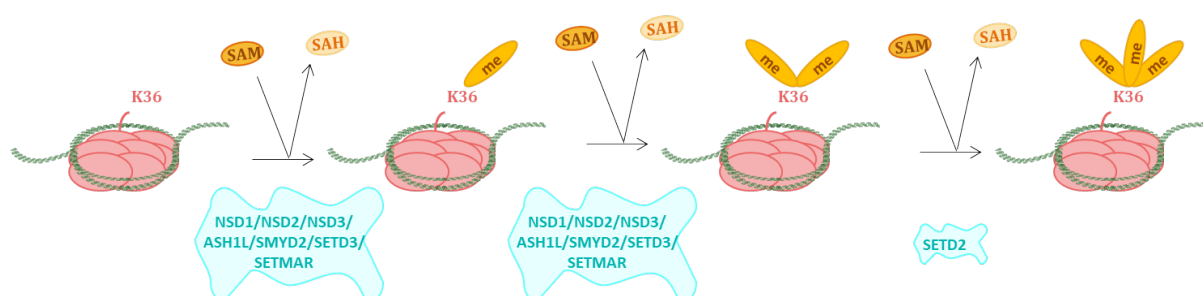


Figure I – 5 : Processivité de méthylation d’H3K36 par les huit HMT des mammifères¹⁴

HMT/ KMT	Maladie Associée
NSD1	Syndrome Sotos/ Leucémie Myéloïde Aiguë (LAM)
NSD2	Syndrome Wolf-Hirschhorn
NSD3	Cancer du sein/ Leucémie Myéloïde Aigue (AML)
SETD2	Carcinome des cellules rénales (ccRCC) / Leucémie Myéloïde Aigue (AML)/ Lymphomes T

Tableau I – 3 : HMT majoritaires d’H3K36 et des maladies associées à leurs fonctions²¹⁻²⁷

Tableau non exhaustif des maladies dans lesquelles les HMT d’H3K36 sont altérées (mutations ou niveau d’expression). NSD1/3, sont notamment liées avec l’AML par des translocations chromosomiques avec la protéine nucléoporine NUP98.

5) Méthylation du résidu H3K79

La méthylation du résidu H3K79 est une marque particulière en ce qui concerne sa localisation, mais également sa formation. Plus précisément, la lysine 79 se situe dans la partie globulaire de l'H3, en contraste des autres lysines méthylables de cette histone (Fig. I-6A). De plus, les trois étapes de méthylation du résidu H3K79 sont effectuées par DOT1L (Disrupter Of Telomere silencing 1 Like protein ou KMT4). DOT1L est la seule HMT qui ne possède pas le domaine catalytique SET (décrit au chapitre suivant). En revanche, DOT1L possède une structure plus proche d'une Arginine Méthyltransférase (PRMT), malgré leur différence sur la spécificité de substrat (Fig-6B) ²⁸. L'acide aminé H3K79 méthylé est lié à la transcription active, notamment à l'élongation de la transcription, étant souvent associée à des facteurs de transcription ²⁹. De ce fait, dans le cas des leucémies, DOT1L est souvent retrouvée délocalisée suite aux translocations des protéines MLL, ses partenaires de fusion étant des facteurs de transcription qui interagissent avec DOT1L ³⁰.

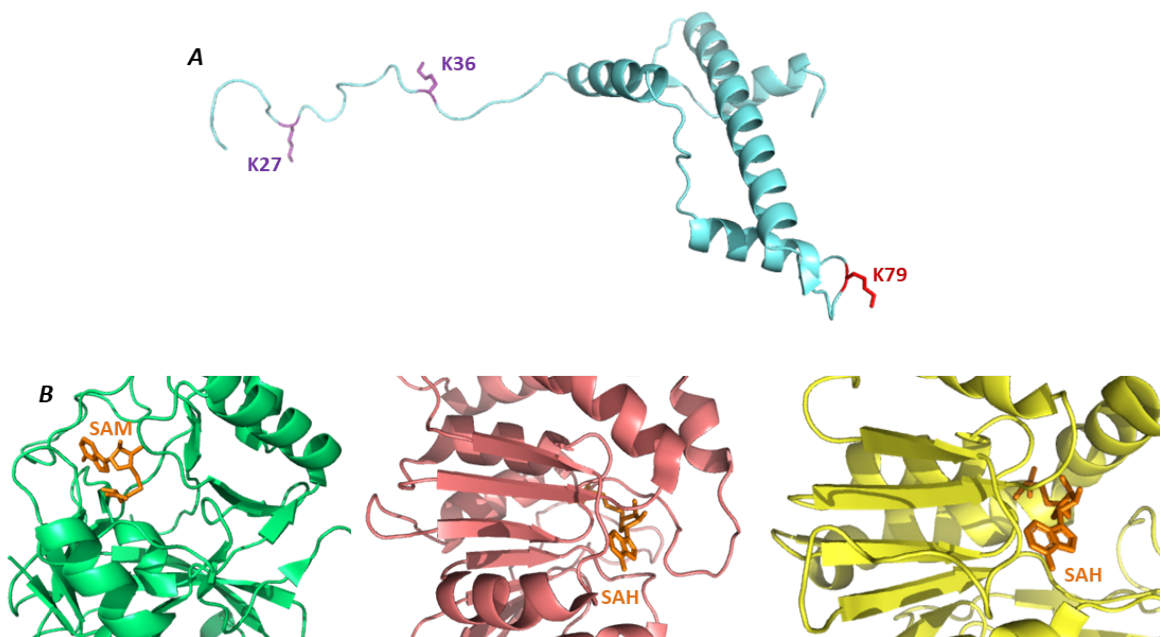


Figure I – 6 : DOT1L, un cas particulier d'HMT

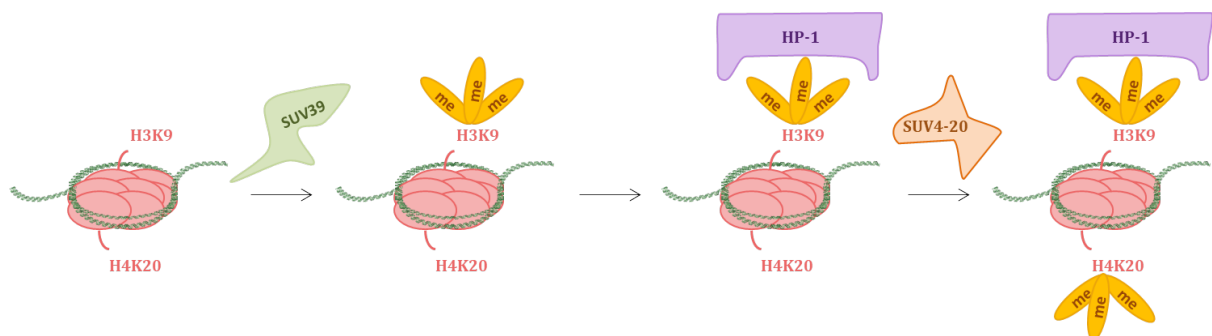
A. Position des lysines 79 (rouge), 27 et 36 (violet) au sein de l'histone 3 (PDB : 1A0I) 28

B. Comparaison des sites de liaison de substrat de SETD2 (vert, PDB : 5JJY), DOT1L (rose, PDB : 1NW3) et PRMT1 (jaune, PDB : 1ORI)

6) Méthylation du résidu H4K20

Les étapes de méthylation progressive (mono-, di-, tri- méthylation) de la lysine 20 de l'histone 4 (H4K20), une autre marque d'hétérochromatine, sont catalysées par des enzymes différentes. La monométhylation est le résultat de l'activité spécifique de SET8 (SET domain containing 8 ou KMT5A), qui est le substrat des SUV4-20H1/2 (SUppressor of Variegation 4 – 20 Homolog 1/2 ou KMT5B/C) pour former à la suite les marques H4K20me2 et H4K20me3^{31,32}. De manière intéressante, la méthylation du résidu H4K20 est dépendante et régulée par la méthylation du résidu H3K9. En effet, SUV4-20 est recrutée au niveau de sa cible, grâce à la protéine de l'hétérochromatine 1 (HP1) qui reconnaît la marque H3K9 méthylée, suggérant ainsi un mécanisme séquentiel de méthylation du résidu H3K9, puis du résidu H4K20 (Fig.I-7)³³.

L'H4K20 méthylée a un rôle important pour l'intégrité génomique et la régulation du cycle cellulaire, avec l'observation de niveaux différents d'H4K20me1 au cours des phases du cycle³⁴. Elle est également importante pour la réparation de l'ADN via le recrutement du facteur de réparation 53BP1 (p53 Binding Protein 1) par interaction directe avec la marque H4K20me2³⁵. La méthylation aberrante du résidu H4K20 est liée au processus tumoral mais, à ce jour, les bases moléculaires de cette corrélation nécessitent d'être élucidées.



*Figure I – 7 : Mécanisme séquentiel de méthylation d'H3K9 et d'H4K20 par l'intermédiaire d'HP-1*³³

B. La marque H3K36me3

Au début des années 2000, plusieurs hypothèses ont été posées par rapport à la stabilité des marques de méthylation d'histones. Avec une demi-vie des histones égale à celle d'une méthyle-lysine (~132 jours), la méthylation était considérée comme une marque permanente. L'implication de cette dernière dans la régulation des processus essentiels à la cellule, dont la transcription, a fait émerger l'hypothèse d'un processus dynamique et réversible ³⁶. Ça n'a été qu'en 2004 que ce dogme a été ébranlé suite à l'identification de la première histone lysine déméthylase, LSD1 (Lysine Specific Demethylase 1), capable de reverser la mono- et di- méthylation de la lysine 4 de l'histone H3 ³⁷.

Aujourd'hui il est établi que la présence d'une marque épigénétique sur les histones est la résultante du rôle de deux acteurs modificateurs, les méthyltransférases ('writers') et les déméthylases ('erasers'). Dans le cas de la marque H3K36me3, SETD2 est la triméthyltransférase et JMJD2 (JuMonJi Domain containing 2) la déméthylase (Figure I-8). En outre, la marque H3K36me3 endosse le rôle de 'messenger', en étant reconnue par des protéines contenant des domaines capables d'interagir et de 'lire' ('readers') ce signal et en déclenchant de nombreux processus cellulaires majeurs, comme la réparation des dommages de l'ADN et l'élongation de la transcription.

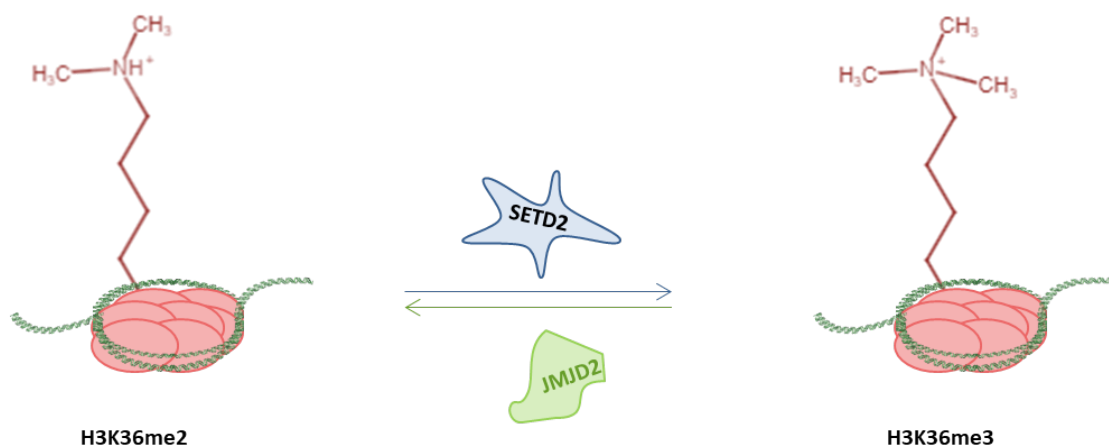


Figure I - 8 : Enzymes responsables de la transition entre la marque H3K36me2 et la marque H3K36me3

1) Régulation de la marque H3K36me3

La triméthylation est un processus particulier par rapport aux autres stades de méthylation du résidu H3K36, puisqu'elle est effectuée par une seule enzyme, SETD2. De manière intéressante, cette protéine est capable aussi de mono- et di- méthyliser le résidu H3K36 *in vitro*, mais sa délétion *in vivo* n'a pas d'incidence sur ces deux premières étapes de méthylation. Cela suggère que dans un contexte cellulaire, SETD2 est uniquement responsable de la dernière étape de ce processus. Cependant, malgré le fait que SETD2 possède une activité méthyltransférase intrinsèque, elle nécessite vraisemblablement la présence d'hnRNP-L (Heterogeneous Nuclear RiboNucleoProtein L) pour exercer cette fonction. hnRNP-L est une protéine capable d'interagir avec l'ARN, permettant ainsi de supposer une relation entre la chromatine et la modification d'ARNm³⁸.

La marque H3K36me3 est considérée aujourd'hui comme une marque réversible, grâce à l'histone déméthylase JMJD2 (ou JHMD3 - JmjC domain - containing histone demethylation protein, ou KDM4 – Lysine specific demethylase 4) qui peut aussi déméthyliser le résidu H3K36 diméthylé, mais aussi le résidu H3K9 di- et tri- méthylé³⁹. Il existe quatre gènes de *Jmjd2* chez l'Homme : *Jmjd2A*, *Jmjd2B*, *Jmjd2C* et *Jmjd2D*. Les JMJD2A/B/C peuvent déméthyliser l'H3K36me3 et l'H3K9me3, alors que le JMJD2D est spécifique d'H3K9me3. L'activité de ces enzymes est due aux deux domaines catalytiques JmjC et JmjN, et leur mécanisme de catalyse nécessite un atome de fer divalent (Fig.I-9)⁴⁰. Malgré tout, l'effet de cette déméthylation reste peu documenté et l'efficacité de la déméthylation d'H3K36me3 est plus faible par rapport à celle d'H3K9me3⁴¹.

Curieusement, la triméthylation du résidu H3K36 est aussi régulée par d'autres facteurs, qui ne sont pas dépendants de SETD2 ou JMJD2. En 2006, Nelson *et al* ont démontré que dans le cas de la levure, l'isomérisation de la proline 38 de l'histone 3 (H3P38) par l'isomérase Fpr4 (FK506 binding protein 4) inhibe la méthylation du résidu H3K36 par SET2 (l'homologue de SETD2)⁴². En effet, l'H3P38, située à la queue N-terminale de l'histone H3, est accessible à l'isomérisation cis-trans par Fpr4 (Fig.I-10). Cette isomérisation peut influencer le processus de méthylation, favorisée par la forme trans, mais aussi la déméthylation qui est

avantageée par la conformation cis. Cette régulation a été confirmée aussi dans le cas de SETD2 humaine, son activité étant incompatible avec la conformation cis⁴³.

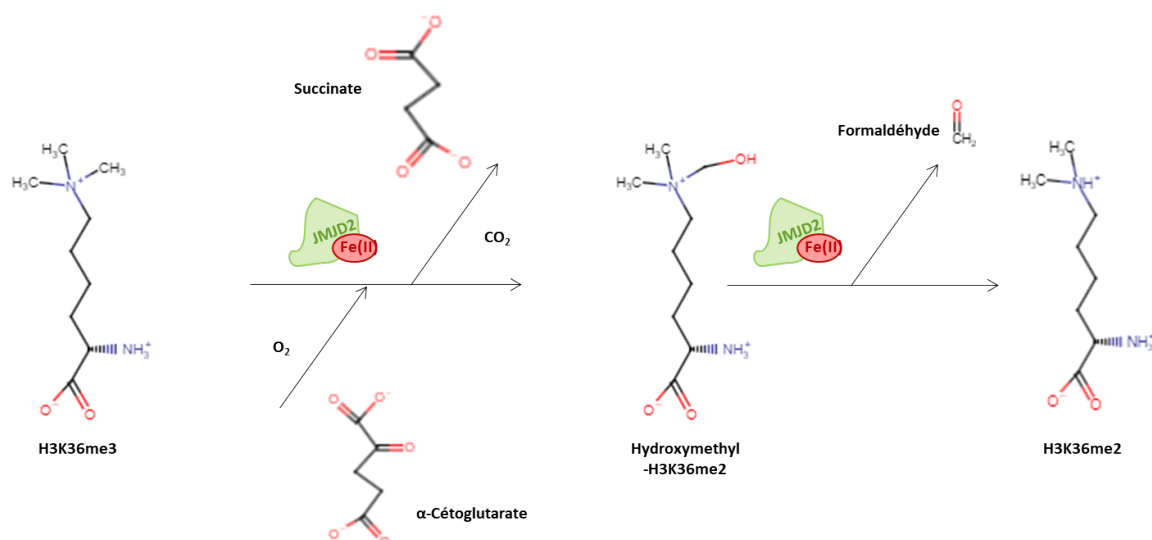


Figure I – 9 : Mécanisme schématisé de déméthylation du résidu H3K36 triméthylé par JMJD2

JMJD2 catalyse une réaction de dioxygénation dépendante de la chélation d'un atome de fer Fe²⁺ par le domaine JmjC qui permet de coordonner une molécule d'oxygène et d'α-cétoglutarate, de manière à former un intermédiaire d'hydroxyméthyl-lysine. En parallèle, l'α-cétoglutarate qui sert comme donneur d'électron est clivé en succinate et dioxyde de carbone. À la fin de la réaction, une molécule de formaldéhyde est libérée (issue de l'hydroxyméthyl-lysine).

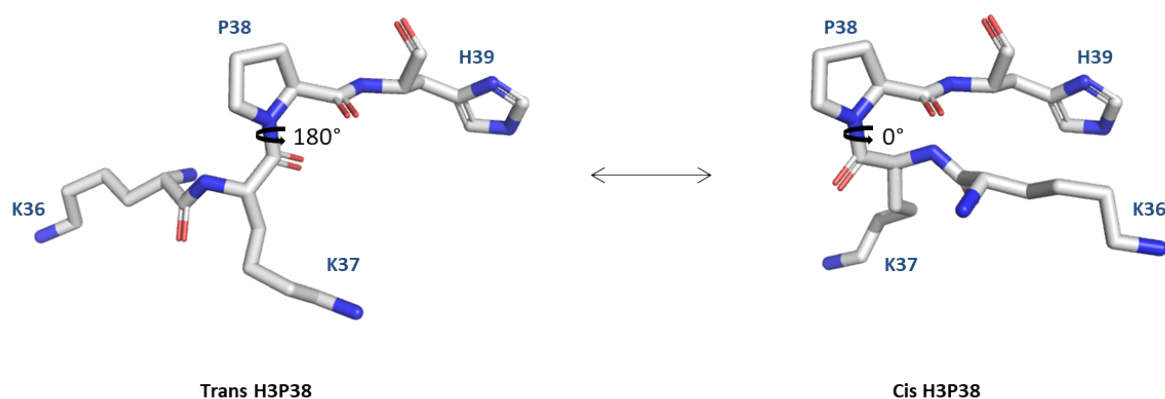


Figure I – 10 : Isomérisation cis-trans de la Proline 38 de l'histone 3⁴²

Cette isomérisation est spontanée dans le cas des peptides et des régions déstructurées comme la queue des histones. Les isomérases, comme la Frp4, permettent d'accélérer cette réaction où l'angle dièdre (ω) de la liaison peptidique entre la P38 et le résidu d'avant (lysine 37) est modifié de 180°

2) Reconnaissance de la marque H3K36me3

Les modifications épigénétiques peuvent affecter la compaction et la structure de la chromatine par altération des liaisons entre les histones et l'ADN, mais aussi par le recrutement de protéines qui peuvent 'lire' les modifications post - traductionnelles (PTM). À la suite de cette reconnaissance, des facteurs du réseau de signalisation nucléaire sont recrutés au niveau de la chromatine de manière à déclencher des processus fondamentaux pour la cellule, dont la transcription, la réplication, la réparation d'ADN ou la modification d'autres résidus de la chromatine. Dans le cas de la méthylation des lysines des histones, comme la triméthylation du résidu H3K36, la compaction de la chromatine n'est pas affectée. En conséquence, les protéines 'readers' sont essentielles pour la reconnaissance de la méthylation des histones et pour l'initiation des processus en aval de ces PTM ⁴⁴.

Aujourd'hui, les protéines 'readers' des lysines méthylées sont les plus étudiées et leur fonction est due à différents domaines responsables de la reconnaissance de la lysine méthylée (Tableau I-4). La liaison à la cible est généralement effectuée par une cage aromatique formée de deux à quatre résidus hydrophobes qui forment des empilements-pi ('pi-stacking') et des interactions hydrophobes et Van-der-Waals avec le groupement méthylamonium. La taille et la composition de cette cage aromatique sont responsables pour le choix de la lysine et de l'état de méthylation que la protéine 'reader' reconnaît. Ceci suggère que la mono- et la di- méthylation sont insérées dans une poche de taille plus restreinte qui peut contenir des résidus acides (Asp ou Glu), alors que la triméthylation nécessite une cavité plus large, composée uniquement des résidus hydrophobes ⁴⁵. La marque H3K36me3 est reconnue par les domaines Chromo-barrel, PWWP et Tudor qui possèdent une architecture caractéristique en tonneau β .

Le domaine PWWP, avant d'être identifié comme un domaine reconnaissant les histones modifiées, dont le résidu H3K36me3 (Fig.I-11A) était connu comme capable de se lier à l'ADN. Il est composé de 100-130 résidus, formant un tonneau β avec cinq brins antiparallèles, suivis de 1-5 hélices α (Fig.11-B). La lysine cible se fixe à la fin de ce tonneau dans une poche composée de trois résidus hydrophobes dirigés de manière perpendiculaire entre eux (Fig.I-11C). Les résidus de surface de la protéine 'reader' forment aussi des liaisons

hydrogènes avec les résidus Gly33 et Gly34 de l’histone 3 qui se situent en amont du résidu Lys36 et qui donnent également une flexibilité à cette partie de l’histone. Cette reconnaissance de la marque H3K36me3 par les ‘readers’ à domaine PWWP est unique grâce à la présence des résidus H3G33/34, mais aussi du résidu Val35 d’H3 qui s’insère dans une poche hydrophobe entre le brin β 2 et l’hélice α 1 (Fig.I-11D) ^{46,47}.

Domaine du ‘reader’	Marque de méthylation associée
ADD (ATRX – DNMT3 –DNMT3L)	H3K9me3
Ankyrin	H3K9me2, H3K9me1
BAH (‘Bromo – Adjacent Homology)	H4K20me2
Chromo-barrel	H3K36me3 , H3K36me2, H4K20me1, H3K4me1
Chromodomain	H3K9me3, H3K9me2, H3K27me3, H3K27me2
DCD (‘Double Chromo Domain)	H3K4me3, H3K4me2, H3K4me1
MBT (Malignant Brain Tumor)	H3Kme1, H3Kme2, H4Kme1, H4Kme2
PHD (Plant HomeoDomain)	H3K4me3, H3K4me2, H3K9me3
PWWP (Pro-Trp-Trp-Pro)	H3K36me3 , H4K20me1, H4K20me3, H3K79me3
TTD (Tandem Tudor Domain)	H3K4me3, H3K9me3, H4K20me2
Tudor	H3K36me3
WD40	H3K27me3, H3K9me3
zf-CW (zinc-finger CW)	H3K4me3

Tableau I – 4 : Domaines des protéines ‘readers’ et lysines méthylées reconnues ⁴⁵

Les domaines Chromo-barrel, PWWP et Tudor, représentés en rouge, sont responsables de la reconnaissance d’H3K36me3. Ces trois domaines, ainsi que les Chromodomaine, MBT et TTD, constituent la famille ‘Royale’, avec une topologie caractéristique d’un tonneau β (β -barrel) ⁴⁸.

(ATRX : ATP – dépendant helicase ou X - linked helicase II, DNMT3 : DNA cytosine - 5 - methyltransferase 3, DNMT3L : DNA cytosine - 5 - methyltransferase 3 - like, WD40 : dipeptide Trp Asp conservé – longueur de 40 résidus, zf-CW : domaine de deux Cys et deux à quatre Trp conservés)

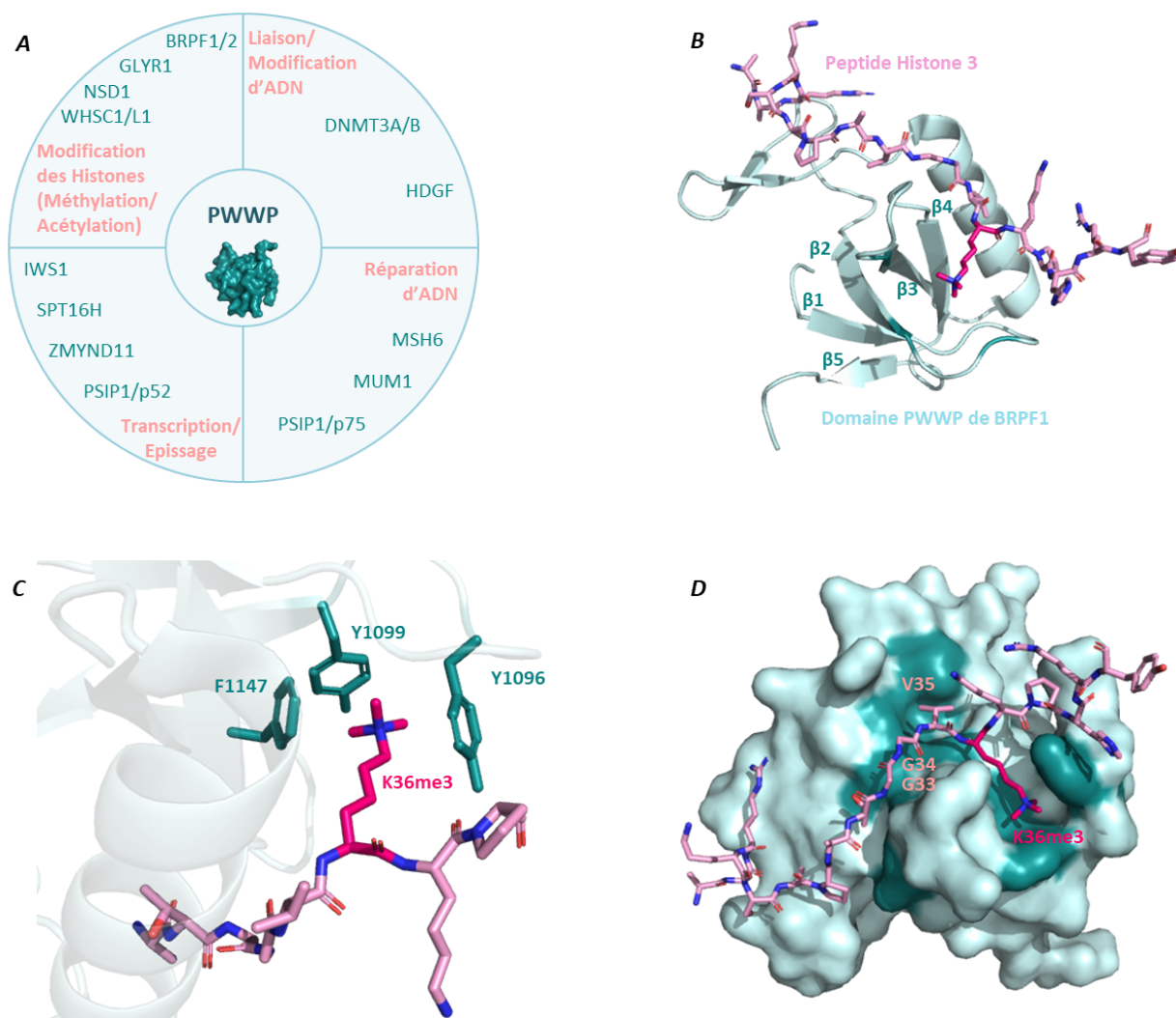


Figure I – 11 : Reconnaissance d'H3K36me3 par le domaine PWWP de BRPF1

A. 'Readers' d'H3K36me3 à domaine PWWP et leurs fonctions associées ⁴⁹

B. Domaine PWWP de BRPF1 en interaction avec le peptide histone 3 ⁴⁶

C. Poche hydrophobe de PWWP de BRPF1 et la reconnaissance d'H3K36me3

D. Surface de PWWP – BRPF1 et les interactions avec le peptide H3

BRPF1/2 : BRomodomain and PHD Finger containing 1/2, GLYR1 : GLYoxylate Reductase 1, WHSC1L1 : Wolf – Hirschhorn Syndrome Candidate 1 - Like 1, HDGF : Hepatoma – Derived Growth Factor, MSH6 : MutS Homolog 6, MUM1 : Melanoma Associated Antigen1, PSIP1 : PC4 and SFRS1 interacting protein (p52 – petit isoforme, p76 grand isoforme), IWS1 : Interacts With SUPT6H, Suppressor of Ty 16 Homolog, ZMYND11 : Zinc Finger MYND – Type Containing 11

Le domaine PWWP de BRPF1 (PDB : 2X4Y) est représenté en cyan clair et le peptide histone 3 (résidus 24-41) en rose. Les résidus en cyan foncé sont les résidus responsables de la reconnaissance d'H3K36me3 (Y1096, Y1099 et F1147). De même, à la surface de PWWP, en cyan foncé sont les résidus qui interagissent avec les G33 et G34 d'H3 (R1152 et W1154), ainsi qu'avec la V35 (V1127, L1130 et W1156).

Le domaine Chromo (Chromodomain) présente une forte ressemblance en termes de structure de son tonneau β avec le domaine PWWP, composé de cinq brins β suivis par une partie hélicoïdale, malgré leur faible similarité en séquence. Le domaine Chromo est assez petit, d'une taille d'environ 30-50 résidus. La reconnaissance de la lysine méthylée se fait à nouveau grâce à la poche hydrophobe composée de 3-4 résidus hydrophobes située à la partie C-terminal du tonneau β . MORF4L1 (MORtality Factor 4 Like 1) impliquée dans l'épissage alternatif, et MSL3 (Male Spécifique Lethal 3), importante pour l'acétylation de la lysine 16 d'histone 4, sont les deux exemples les plus connus de protéines 'readers' de la marque H3K36me3, à domaine Chromo^{50,51}.

Le domaine Tudor, présente la même topologie (tonneau de cinq brins β mais pas de partie en hélice, Fig.I-12-A) que les domaines PWWP et Chromo, mais avec une plus forte affinité et spécificité de liaison à sa cible. L'interaction de la cage hydrophobe (composée de quatre résidus) avec la lysine triméthylée est renforcée par des liaisons hydrogène, formées par un résidu acide et la partie triméthylamonium (Fig.I-12-B). De manière intéressante, le domaine Tudor sous sa forma apo est structuré en tonneau β , mais la poche de liaison de la lysine cible n'est formée qu'après la liaison de l'histone H3 (Fig.12-C). De plus, les résidus T32-R40 d'H3 forment des liaisons avec la surface du domaine Tudor, favorisant ainsi cette interaction (Fig.12-D). Néanmoins, le domaine Tudor peut exister en forme simple ou en double répétition en tandem, contenant deux tonneaux β (TTD) et seule la forme simple peut reconnaître la marque H3K36me3. Les deux "readers' à domaine Tudor connus sont les PHF1 (PHD Finger protein 1) et PHF19 qui ont un impact sur la méthylation d'H3K27 par régulation de l'activité du complexe PRC2^{45,52,53}.

Globalement, la marque H3K36me3 peut être 'lue' par différentes protéines 'readers'. Selon la protéine 'reader', un 'signal' différent est émis dans la cascade de signalisation nucléaire, ayant comme résultat une réponse différente au niveau cellulaire. Les fonctions les plus connues suite à la reconnaissance de la marque H3K36me3 sont le contrôle de la transcription, l'épissage de l'ARNm, la réparation d'ADN et le contrôle d'autres marques épigénétiques.

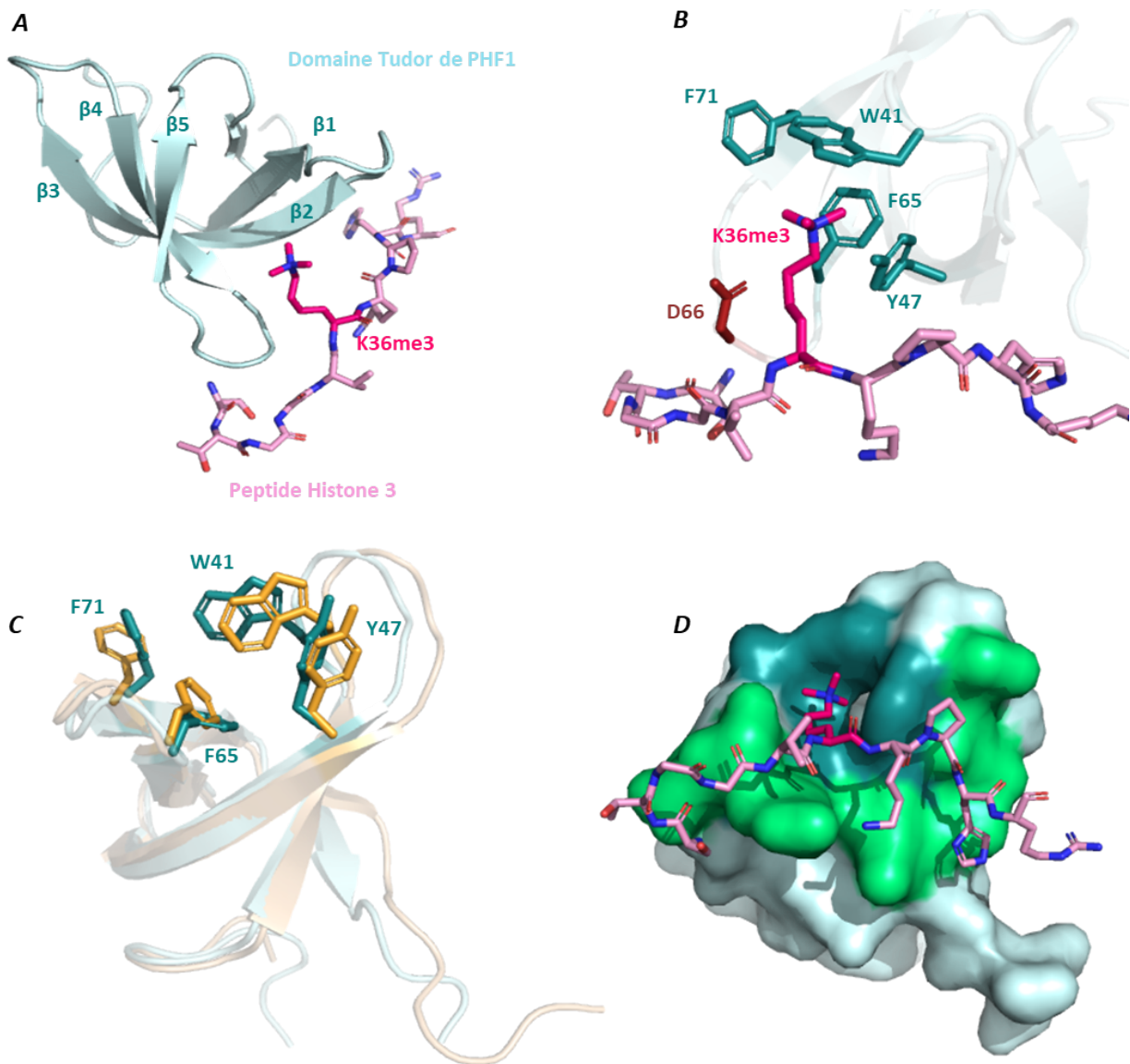


Figure I – 12 : Reconnaissance d'H3K36me3 par le domaine Tudor de PRF1

A. Tonneau β du domaine Tudor en interaction avec le peptide histone 3 ⁵²

B. Poche hydrophobe de Tudor de PRF1 et les résidus interagissant avec l'H3K36me3

C. Structuration de la poche hydrophobe Tudor, après la liaison avec l'H3

D. Surface de Tudor – PRF1 et les interactions avec l'H3K36me3 et le peptide H3

La partie haute du panel, à gauche, le domaine Tudor de PRF1 (PDB : 4HCZ) est représenté en cyan clair et le peptide histone 3 (résidus 31-40) en rose. À droite, les résidus en cyan foncé sont les résidus responsables de la reconnaissance d'H3K36me3 (W41, Y47, F65 et F71) et le résidu en rouge (D66) permet la formation d'une liaison hydrogène avec la partie triméthylamonium d'H3K36me3.

Dans la partie basse du panel, à gauche, sont représentés en orange les résidus de la poche hydrophobe de la forme Tudor Apo (PDB : 2E5P) et en cyan foncé les mêmes résidus suite à la liaison d'H3 et la formation correcte de la poche. À droite, les résidus en cyan foncé sont ceux qui accueillent l'H3K36me3 et ceux en vert sont ceux qui interagissent avec le reste d'H3 de manière à stabiliser, à renforcer et à rendre la reconnaissance d'H3K36me3 spécifique.

3) Implication de la marque H3K36me3

a. L'élongation de la transcription et la transcription cryptique

La fonction prépondérante de la marque H3K36me3 est la participation dans les processus de transcription, notamment l'élongation et la prévention de la transcription cryptique. Cette marque est retrouvée dans plusieurs stades de la biosynthèse de l'ARNm. Malgré le lien global de la marque H3K36me3 avec l'euchromatine, les fonctions de cette marque épigénétique ne sont pas toujours associées avec la transcription active. Son effet est plutôt défini par la protéine 'reader' qui est recrutée à ce niveau.

Lors de la transcription, SETD2 est recrutée au niveau de l'ARN polymérase II (RNAPII) de manière à triméthyliser le résidu H3K36. Ce 'dépôt' induit le recrutement du complexe FACT (FACilitates Chromatin Transcription), par la sous-unité SPT16H de ce complexe. Lors de l'association de ce complexe chaperon d'histones à la chromatine, il y a un désassemblage des nucléosomes, avec la libération du dimère des histones H2A-H2B, sans influencer les histones H3-H4. Cela induit l'accessibilité de l'ADN et la transcription peut ainsi progresser. Une fois la RNAPII passée, le complexe FACT intervient à nouveau pour remplacer le dimère H2A-H2B et réassembler le nucléosome, en refermant la chromatine de manière à inhiber la transcription cryptique ⁵⁴.

Par ailleurs, la marque H3K36me3 permet, après le passage de la RNAPII le recrutement de protéines comme ZMYND11 et PHF19, afin de réprimer la transcription des gènes qui viennent d'être transcrits et de permettre à la RNAPII d'avancer. En effet, ZMYND19 permet d'inhiber l'activité de la polymérase par le recrutement des déméthylases et des désacétylases, induisant ainsi le passage de la chromatine à son état compact, non accessible à la transcription ^{55,56}. En outre, PHF19 a un effet concomitant à celui de ZMYND11 en induisant la liaison du complexe PRC2 et de la déméthylase JMJD2 pour enlever la marque d'euchromatine active H3K36me3 et générer la marque d'inhibition de la transcription H3K27me3 (Fig.I-13) ⁵³.

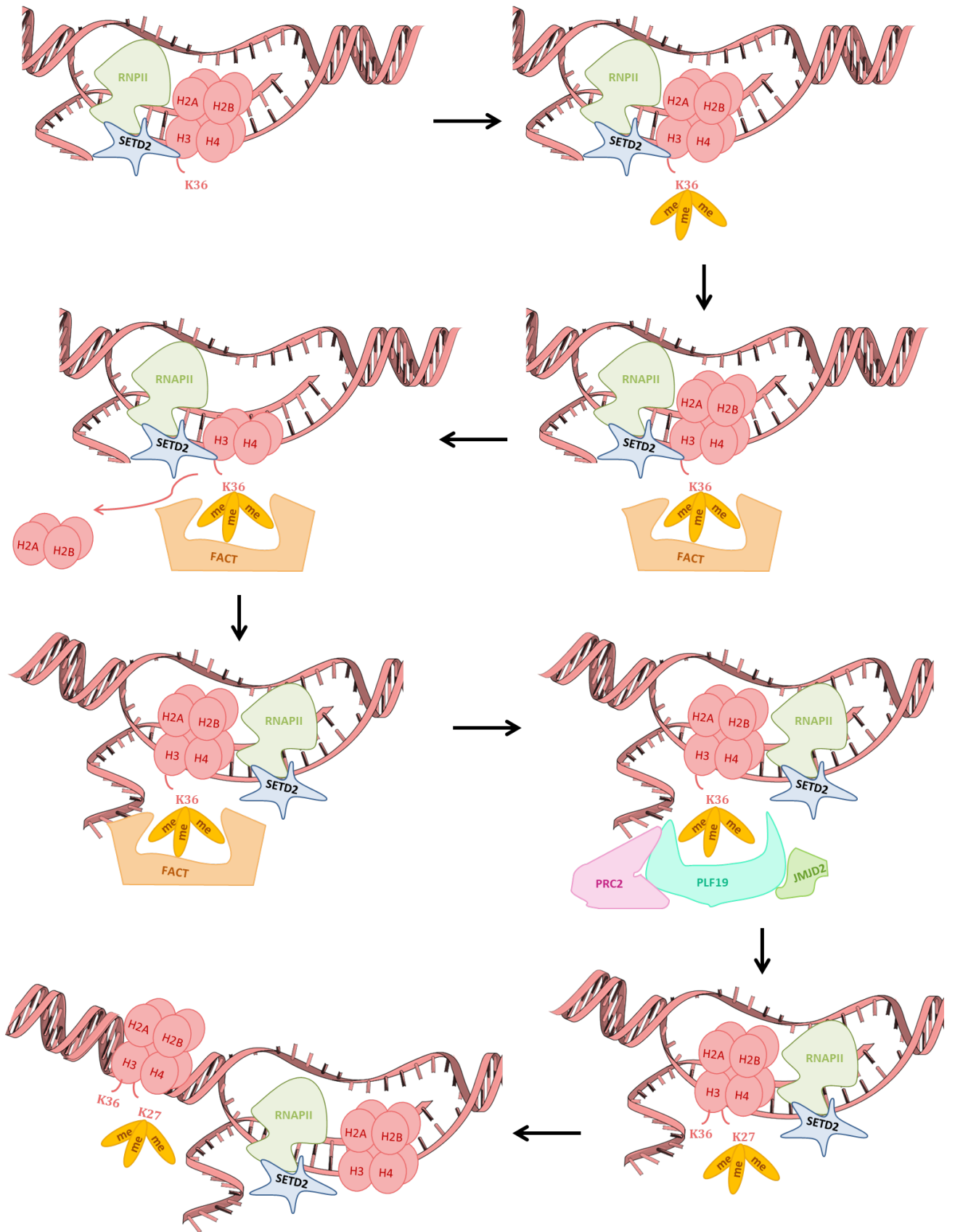


Figure I - 13 : Représentation simplifiée de la régulation de la transcription par la marque H3K36me3

b. Les modifications post-transcriptionnelles des ARNm

Aujourd'hui, le lien dynamique entre l'épissage et les modifications de la chromatine est bien documenté. L'idée que les nucléosomes sont présents de manière prépondérante dans les régions des exons est supportée par le fait que la taille moyenne d'un exon est d'environ 147pb, mais aussi par le fait qu'il y a une communication entre les modifications épigénétiques et l'épissage¹⁴. Dans le cas d'H3K36me3, il a été documenté que sa présence est diminuée dans le cas de gènes sans introns⁵⁷. En parallèle, au sein d'un gène contenant des introns et des exons, la triméthylation du résidu H3K36 marque plutôt les régions des exons que celles des introns.

D'autre part, la marque H3K36me3 est également moins présente dans le cas des exons qui vont être exclus lors des processus d'épissage alternatif⁵⁸. Plus particulièrement, elle peut être reconnue par MORF4L1 qui se lie au régulateur d'épissage PTBP1 (Polypyrimidine Tract Binding Protein 1), permettant ainsi d'exclure des exons des ARNm, en générant des transcrits de taille différente⁵⁹. En outre, PSIP1/p52, qui lie aussi la marque H3K36me3, induit le recrutement des facteurs d'épissage, dont Srsf1 (Serine/ Arginine Rich Splicing Factor 1) qui est délocalisé dans la chromatine en absence de PSIP1/p52⁶⁰.

La triméthylation du résidu H3K36 est liée à d'autres modifications d'ARNm, dont l'ajout de N6-méthyladénosine (m6A) qui permet entre autre l'export d'ARNm du noyau. Pour cela, la marque H3K36me3 est reconnue par METTL14 (N6-adenosine-methyltransferase non-catalytic subunit), un composant du complexe de m⁶-méthyltransférase (MTC). Ce recrutement permet l'interaction directe de MTC avec la RNAPII, aboutissant à la méthylation en cours de transcription de la m⁶A cible⁶¹. Le fait que la marque H3K36me3 a un rôle dans l'export d'ARNm du noyau est supporté par la liaison d'IWS1 (Interacts With SUPT6H 1). L'IWS1 a été montré comme étant l'intermédiaire de liaison d'ARNm à exporter, avec le facteur d'export Aly/REF (Ally of AML-1 and LEF-1)⁶².

c. La réparation de l'ADN

La cellule se trouve constamment en présence d'agents ou de processus qui provoquent des lésions à l'ADN. Cela est la raison de l'existence de plusieurs mécanismes de défense et de réparation. Selon le type de lésion, il existe deux grands types de mécanismes de réparation de l'ADN, les ruptures de double brin (DSB – Double Strand Breaks) et les mésappariements d'ADN (MMR – MisMatch Repair). Ces processus reposent sur des machineries nucléaires qui peuvent se lier à l'ADN grâce aux marques issues des modifications de la chromatine, dont la marque H3K36me3, qui permettent de préserver l'intégrité génomique. La marque H3K36me3 a été décrite comme étant impliquée dans les deux processus, selon les protéines 'readers' qui la reconnaissent.

Dans le cas de DSB, PSIP1/p75 se lie à la marque H3K36me3 et permet le recrutement de CtIP (C-terminal binding protein Interacting Protein). Cette dernière est responsable de l'ablation du site de cassure pour que RAD15 (RAD51 recombinaison) puisse à la suite générer les deux nouveaux brins d'ADN (Fig.I-14) ^{63,64}. En parallèle, la marque H3K36me3 est impliquée dans la machinerie de MMR qui permet de corriger les erreurs de réplication. Ces mésappariements sont reconnus par le complexe MutS α/β (Mutator S – α/β). MutS α est le partenaire majoritaire de ce complexe avec un ratio de 10 :1, et il est composé des MSH2/6 (MutS Homologue protein 2/6). Grâce à son domaine PWWP, MSH6 se lie aux régions des mésappariements, marquées par le résidu H3K36me3. L'abondance de ces facteurs dans le noyau dépend de la phase du cycle cellulaire avec un pic en début de la phase S. Une fois la région de mésappariement reconnue par le complexe MutS α/β , ce dernier se lie directement à l'ADN de manière à initier le processus de réparation ¹⁸.

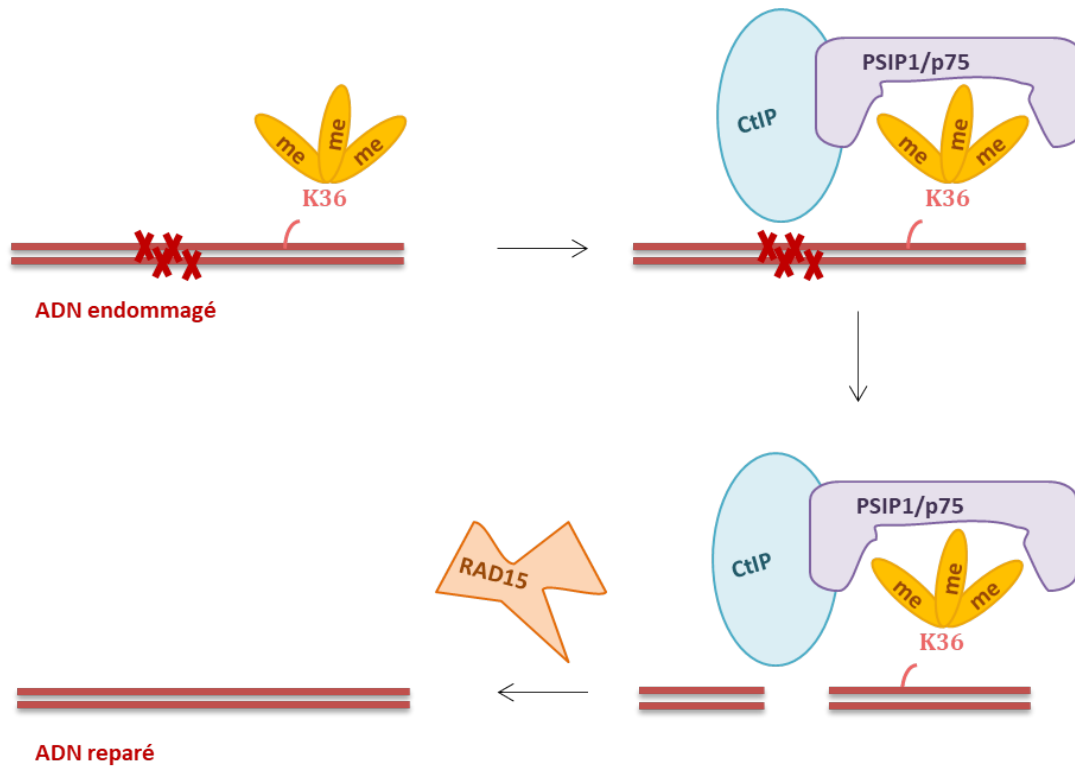


Figure I – 14 : Mécanisme de réparation DSB par la signalisation d'H3K36me3

d. Contrôle d'autres marques épigénétiques

La notion de communication ('cross-talk') des différentes marques épigénétiques et de la présence d'un code génétique est aujourd'hui supportée notamment par l'existence des protéines 'readers', responsables de la régulation d'autres PTM. Ainsi, le niveau de la marque H3K36me3 peut être régulé par le domaine PWWP des NSD1/2/3. Ces méthyltransférases vont à la suite générer des marques H3K36 mono- et di- méthyl. Cette dernière permet à SETD2 de former la marque H3K36me3⁶⁵. De plus, comme il a été décrit précédemment, grâce aux PHF1/19, l'H3K36me3 peut réguler par activation ou inhibition la méthylation du résidu H3K27 via les complexes PRC2^{52,53}. Néanmoins, la marque H3K36me3 peut communiquer avec d'autres marques épigénétiques comme l'acétylation des histones et la méthylation de l'ADN (Fig.I-15).

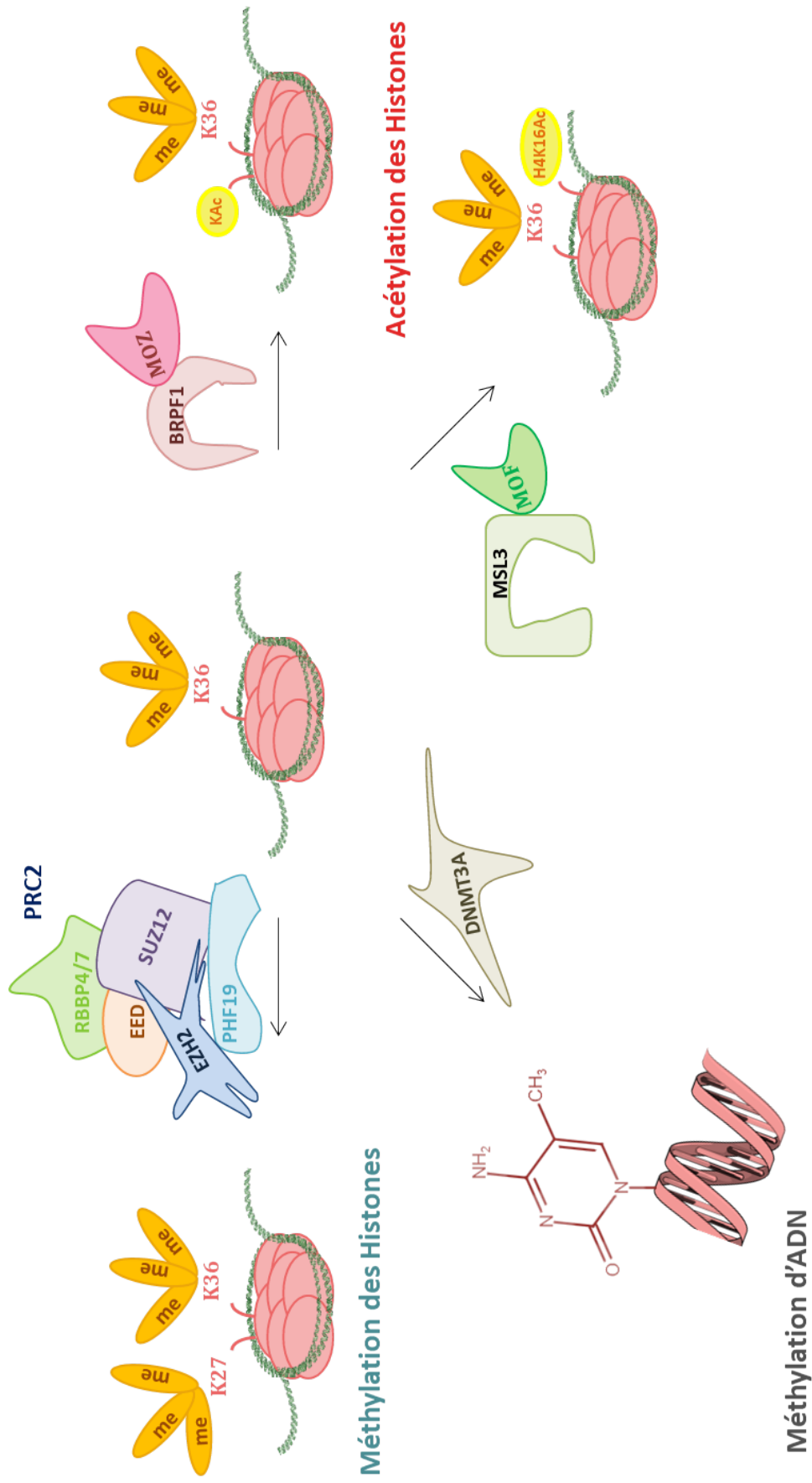


Figure 1 – 15 : Communication (« cross-talk ») de la marque H3K36me3 avec la méthylation/ acétylation des histones et la méthylation d'ADN

Cela est le cas de la communication entre la marque épigénétique H3K36me3 et l'acétylation des histones par MOZ (MOnocytic leukemia Zing Finger) et MORF (MOz Related Factor), deux acétyltransférases d'histones homologues qui constituent les sous-unités catalytiques des complexes multi-protéiques. Un des acteurs de ces complexes est BRPF1, qui se lie à la chromatine et permet le recrutement des MOZ/MORF. Cette liaison est permise par les domaines Bromo et PWWP qui lient l'acétylation et la méthylation des histones. Le domaine PWWP de BRPF1 peut reconnaître la marque H3K36me3, induisant ainsi l'acétylation par MOZ et MORF au niveau des lysines proches du résidu H3K36 triméthylé, comme les acides aminés K14 et K23 de l'histone H3⁴⁶. La triméthylation du résidu H3K36 peut aussi influencer l'acétylation du résidu H4K16 par l'histone acétyltransférase MOF (Males absent On the First) qui est recrutée au niveau de la marque H3K36me3 par le 'reader' MSL3⁶⁶. En outre, il a été démontré que ces deux marques, l'H3K36me3 et l'H4K16Ac, ont un rôle synergique dans la réparation des lésions d'ADN par l'intermédiaire de PSPI1/75⁶⁷.

Par ailleurs, la méthylation de l'ADN s'inscrit aussi dans le cas des modifications épigénétiques qui peuvent 'communiquer' avec la méthylation des histones. Notamment, il a été prouvé que les DNMT3A/B reconnaissent la marque H3K36me3 et permettent la méthylation *de novo* des îlots CpG⁶⁸. La méthylation des îlots CpG est fortement associée à l'inhibition de la transcription, et est souvent retrouvée dans le corps des gènes activement transcrits. L'emplacement commun entre les îlots CpG méthylés et la marque H3K36me3, ainsi que leur rôle commun d'inhibition de la transcription cryptique, permet de supporter la notion de 'communication' entre les deux⁶⁹.

4) Mutations oncogéniques d'H3K36

La triméthylation du résidu H3K36 par SETD2 pour former la marque H3K36me3 est impactée par des mutations oncogéniques d'H3K36 sur le variant H3.3. Ces altérations ont été identifiées dans 90% des patients atteints de chondroblastomes et également dans le

cas de sarcomes. La mutation oncogénique majoritaire est l'H3K36M (mutation de la lysine 36 en méthionine) et des mutations H3K36I (lysine en isoleucine) ont été aussi documentées dans le cas de sarcomes. La diminution de la triméthylation du résidu H3K36 est spécifique de ces mutations car il n'y a pas d'effet similaire sur l'expression des autres marques de méthylation d'histones ^{19,20}.

En 2016, Fang *et al* et Lu *et al* ont démontré que ces mutations altèrent l'expression des gènes impliqués dans des processus tumoraux. En effet, les mutations H3K36M et H3K36I diminuent la triméthylation du résidu H3K36 et inhibent ainsi la différenciation des cellules progénitrices et leur prolifération aberrante. La perte de di- et de tri- méthylation du résidu H3K36 n'est pas due à un effet sur l'expression des méthyltransférases responsables, NSD2 et SETD2, respectivement. Les méthyltransférases SETD2 et NSD2 sont retrouvées séquestrées au niveau de la mutation H3K36M car ces enzymes ont une plus forte affinité pour le résidu méthionine que le résidu lysine ^{19,20}.

Les histones avec des mutations H3K36M ou H3K36I jouent donc le rôle d'inhibiteur de SETD2 ou NSD2 (Fig.I-16) ¹⁹. La perte concomitante de la marque H3K36me3 résulte d'un dépôt non régulé et d'une augmentation du niveau de triméthylation du résidu H3K27 par le complexe de PRC2. Il a été démontré effectivement que PHF1, qui fait le lien entre la marque H3K36me3 et l'inhibition de PRC2, ne peut plus se lier au mutant oncogénique H3K36M. En outre, la liaison de BRPF1 est inhibée dans le cas de H3K36M, ce qui suggère qu'il y a une altération du niveau d'acétylation avec implication du 'reader' BRPF1 ⁷⁰.

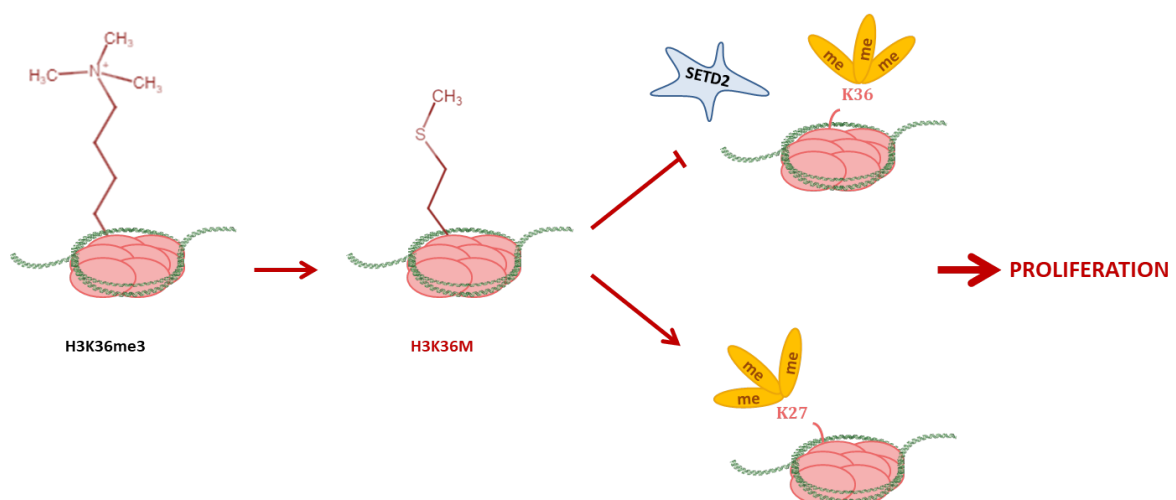


Figure I – 16 : Représentation simplifiée de l'effet d'H3K36M sur les marques d'H3K36me3 et H3K27me3

II. SET Domain Containing II, SETD2

La méthylation des histones est catalysée par la famille des Histone MéthylTransférases (HMT) qui transfèrent un groupement méthyle sur leur substrat en utilisant comme cofacteur la SAM. Chez l'Homme, sont documentées environ 50 HMT qui peuvent méthyler les lysines de la queue N-terminale des histones (DOT1L méthyliant la partie globulaire d'H3). À l'exception de DOT1L, le point commun des HMT est la présence du domaine catalytique SET, appelé d'après les gènes *Su(var)* (Suppressor of Variegation), *E(z)* (Enhancer of Zeste) et *Trx* (Trithorax), qui codent pour des protéines impliquées dans des processus épigénétiques chez la Drosophile ⁷¹.

Le domaine SET, composé d'environ 130 résidus, est conservé dans toutes les HMT et ne présente pas d'homologie en termes de séquence et de structure avec d'autres méthyltransférases, suggérant ainsi que les protéines à domaine SET constituent une nouvelle superfamille de méthyltransférases. Par ailleurs, une caractéristique importante des HMT est la présence de régions de fixation du substrat et du cofacteur distinctes mais proches dans la structure protéique. Les domaines adjacents au domaine SET, riches en Cys, ont été identifiés comme essentiels pour l'activité enzymatique, probablement en ayant un rôle dans la stabilisation de la structure du domaine SET. De manière intéressante, malgré l'importance de ces régions, elles ne sont pas strictement conservées au sein des HMT ⁷².

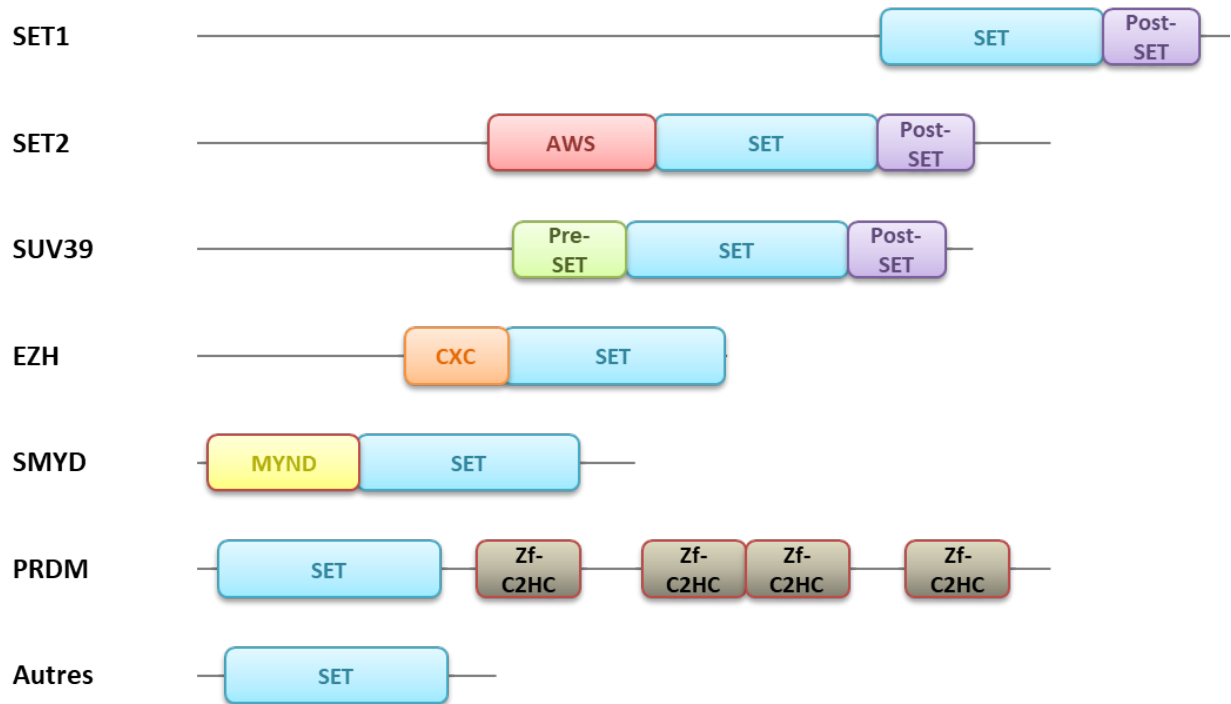
Selon la conservation de la séquence des régions aux extrémités du domaine SET, les méthyltransférases d'histones (à domaine SET) peuvent être classées en 7 familles, avec une spécificité de substrat pour chacune d'elles. Le Tableau II-1 représente ces familles alors que la Figure II-1 illustre leurs différences sur les régions autour du domaine SET. SET Domain containing 2 (SETD2) fait partie de la famille SET2, responsable de la méthylation d'H3K36. La caractéristique de cette famille est la présence des domaines AWS (Associated With Set) et post-SET, respectivement sur les parties N- et C-terminal flanquant le domaine SET. Ces deux régions sont des doigts de zinc structuraux, dont la présence est indispensable à l'activité enzymatique.

FAMILLE	HMT	Lysine d’Histone méthylée
SET1	MLL1/2/3/4	H3K4
	hSET1A/B	
SET2	NSD1/2/3	H3K36
	ASH1L	
	SETD2	
SUV39	SUV39H1/2	H3K9
	G9a	
	GLP	
	SETDB1	
	SETMAR	H3K36
EZH	EZH1/2	H3K27
SMYD	SMYD1/2/3	H3K4
PRDM	PRDM2	H3K9
	PRDM9	H3K4
Autres	SET8	H4K20
	SUV4-20H1/2	
	SET7/9	H3K4

*Tableau II – 1 : Familles des HMT à domaine SET et lysine de substrat histone*⁷³

Le tableau représente les membres principaux et mieux documentés de chaque famille.

PRDM : PR domain containing



*Figure II – 1 : Représentation schématique des domaines catalytiques des 7 familles des HMT à domaine SET*⁷³

AWS : Associated With Set, CXC : Tesmin/TSO1-like, domaine riche en Cys, Zf-C2HC : doigt - de - zinc de type 3 Cys et 1 His

Bien que SETD2 soit considérée comme une protéine ubiquitaire, elle a été identifiée dans un premier temps dans les cellules souches hématopoïétiques, CD34+^{74,75}. Avant d'être caractérisée en 2005 comme une histone lysine méthyltransférase à domaine SET et un co-activateur transcriptionnel, SETD2 a été initialement découverte en 1998 comme une protéine partenaire de la protéine Huntingtin^{76,77}. En effet, l'Huntingtin, retrouvée altérée dans le cas de la maladie d'Huntington, interagit par sa région riche en prolines avec le domaine WW (Trp-Trp) de SETD2. À ce titre, SETD2 était nommée à l'époque HYPB pour Huntingtin Yeast interaction Binding Protein B. Cependant, en 2008 Edmunds *et al* ont renommé HYPB en SETD2 et l'ont caractérisée comme la seule HMT, capable de trimétyler le résidu H3K36 chez les mammifères⁷⁸.

A. Les domaines de SETD2 et leur fonction

SETD2 est une protéine de grande taille, composée de 2564 résidus (Fig.II-2A). L'activité catalytique de l'enzyme est due au domaine SET qui nécessite toutefois la présence des domaines AWS et Post-Set. Le triplet AWS – SET – Post-SET constitue ainsi le 'cœur catalytique' de SETD2 (Fig.II-2B). À ce jour, quatre autres domaines sont bien documentés. Les domaines SHI (SETD2 hnRNP Interaction), WW (Trp-Trp) et SRI (Set2 Rpb1 Interacting), situés en C-ter (C-terminal), possèdent comme fonction commune l'interaction avec d'autres protéines. Le domaine PRR (Proline Rich Region) permet l'auto-inhibition de SETD2. Par ailleurs, la région en N-ter (N-terminal) de SETD2 est la moins caractérisée. Ce n'est qu'en 2020 que *Bhattacharya and Workman* ont démontré que cette partie de l'enzyme est instable et permet de réguler la demi-vie de SETD2 et par conséquent, sa fonction et son activité ⁷⁹.

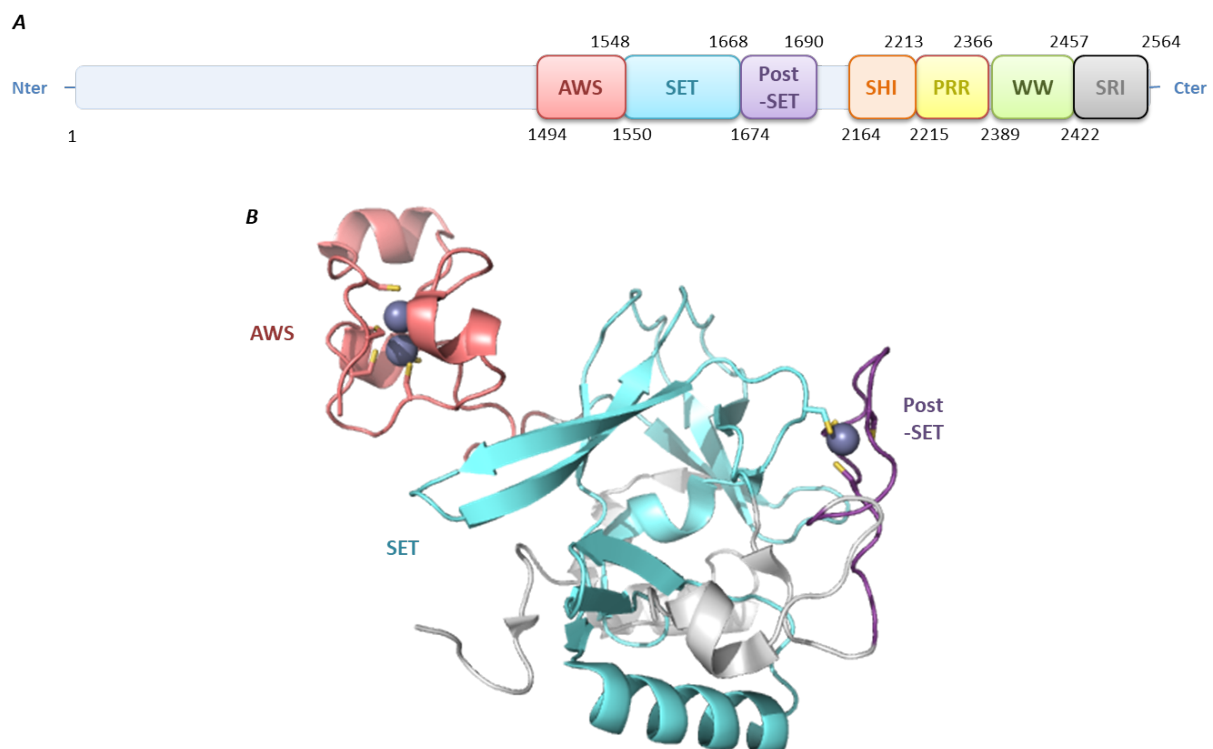


Figure II – 2 : Les domaines de SETD2

A. Représentation schématique des domaines de SETD2

B. Domaines AWS – SET – Post-SET

Les atomes de zinc des domaines AWS et Post-SET sont représentés en sphère (gris). Les chaînes latérales des Cystéines impliquées dans les doigts de zinc sont en sticks.

1) Les domaines structuraux AWS et Post-SET

a. Le domaine AWS

Le domaine AWS est caractéristique de la famille SET2, suggérant ainsi une spécificité de celui-ci pour la méthylation du résidu H3K36. Il est composé d'environ 50 résidus, dont 7 Cys, qui coordonnent 2 atomes de Zinc (Zn^{2+}). Ces éléments forment une structure particulière de Zn_2C_7 (Fig.II-3A). Le domaine AWS a été découvert lors d'analyses computationnelles pour l'identification des nouveaux domaines protéiques. Sa fonction précise reste peu documentée, malgré son importance pour l'activité enzymatique⁸⁰. Il est toutefois supposé que sa fonction est similaire à celle du domaine Pre-SET, qui permet de stabiliser la structure du domaine SET, en interagissant avec des résidus éloignés du site catalytique⁸¹.

b. Le domaine Post-SET

Le domaine Post-SET adopte la forme d'une boucle flexible. Il est composé également d'un doigt de zinc de type C4, formé par 4 résidus cystéines, dont les 3 se trouvent au niveau du domaine Post-SET et la quatrième dans la partie C-terminale du domaine SET (C1631, Fig.3B). De manière intéressante, le positionnement du résidu Histidine 1629 (domaine SET) permet de supposer la présence potentielle d'un cinquième résidu coordinateur de ce doigt de zinc. Plusieurs familles des HMT nécessitent la présence de ce domaine, suggérant ainsi un rôle important de ce dernier dans le fonctionnement enzymatique⁸².

En effet, la boucle du domaine Post-SET (appelée aussi boucle auto-inhibitrice ou L_{IN}) peut adopter plusieurs conformations qui permettent d'auto-réguler l'activité enzymatique. De plus, les résidus du domaine Post-SET participent à la formation de la poche de fixation du substrat. De manière intéressante, le domaine Post-SET se trouve déstructuré en amont de la liaison du substrat de l'enzyme. Ce n'est qu'après sa fixation que ce domaine se stabilise, avec la coordination de l'atome de Zn^{2+} et la formation d'une boucle autour du substrat, de manière à stabiliser son interaction avec l'enzyme^{43,83}.

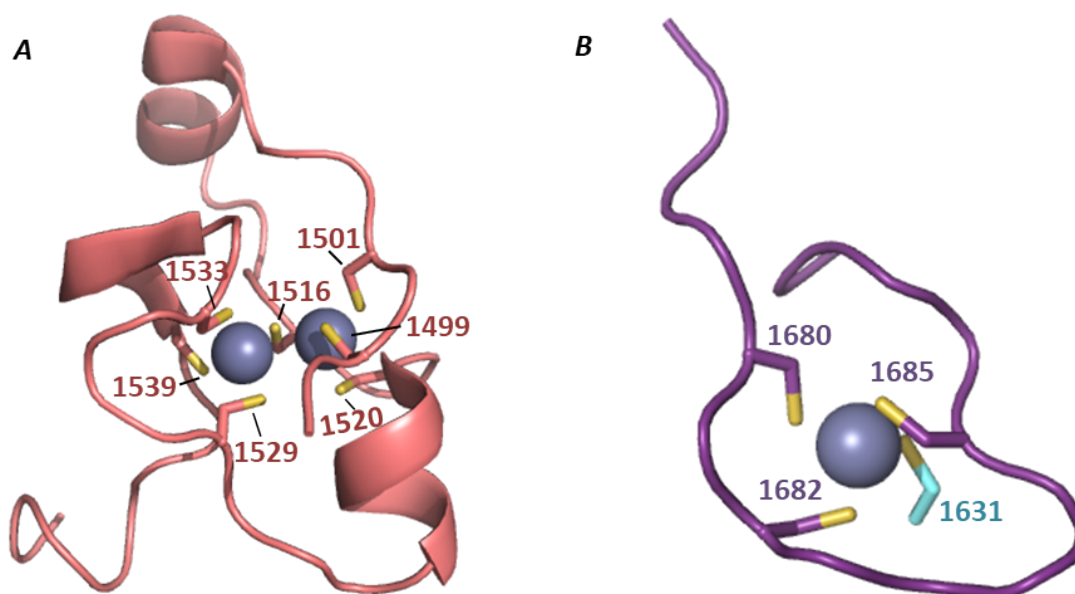


Figure II – 3 : Les domaines à doigts de zinc structuraux de SETD2, AWS et Post-SET

A. Domaine AWS

B. Domaine Post-SET

Les Cystéines 1499, 1501, 1516, 1520, 1529, 1533, 1539 coordonnent les deux atomes de zinc pour la formation du double doigt – de – zinc du domaine AWS. Les Cystéines 1678, 1680 et 1685 du domaine Post-SET, ainsi que la 1631 du domaine SET forment le C4 du domaine Post-SET. La chaîne latérale de la C1631 en cyan CPK (Corey Pauling Koltun).

2) Le domaine catalytique SET

Le domaine SET, conservé au cours l'évolution des HMT, est responsable de la catalyse enzymatique par la formation d'un complexe ternaire avec le cofacteur et le substrat (Fig.II-4A). La fixation du cofacteur et du substrat dans deux sites distincts contribue à la processivité de l'enzyme ⁷². Le domaine SET s'organise sous une forme globulaire et est sous-divisé en 3 régions (Fig.II-4B,C), SET-N (N-ter), SET-I (Insert) et SET-C (C-ter). Les régions SET-N et SET-C forment chacune 3-4 petits feuillets β et une hélice qui sont reliés entre eux par des boucles. Ces feuillets forment une structure triangulaire, caractéristique de ce domaine.

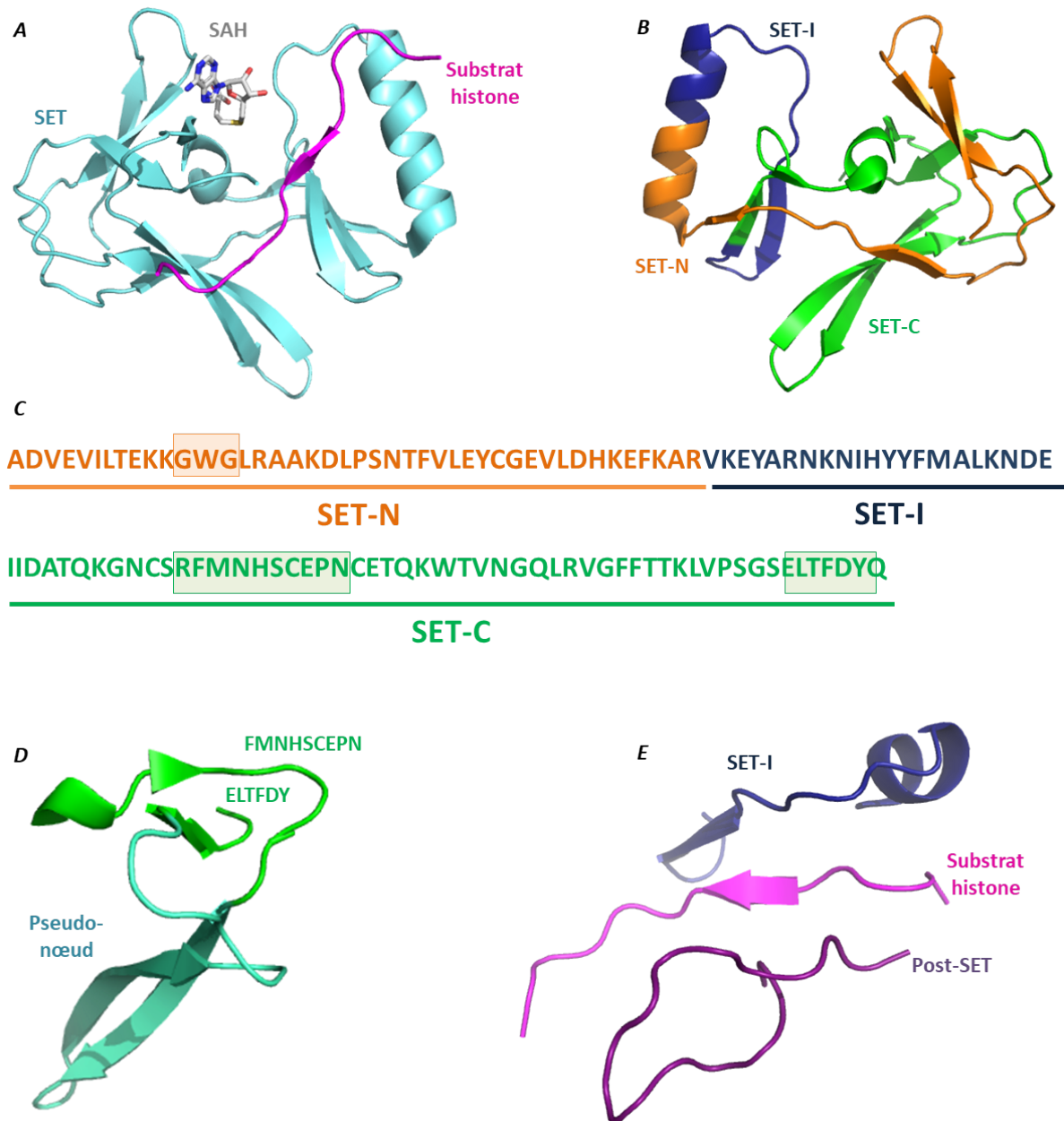


Figure II – 4 : L'architecture du domaine catalytique SET

A. Représentation du domaine SET (cyan) en complexe avec le substrat histone (magenta) et le cofacteur SAH (gris, CPK)

B. Les 3 régions du domaine SET : SET-N (A1550-R1592, orange), SET-I (V1593-E1613, bleu), SET-C (I1614-Q1667, vert)

C. La séquence des 3 régions du domaine SET et la position dans la séquence des motifs $G_{1561}XG_{1563}$, $R_{1625}FINHXCXPN_{1634}/E_{1661}LXFDY_{1666}$

D. Pseudo-nœud (cyan) qui permet le rapprochement des motifs formant le site actif (vert)

E. Sillon de liaison du substrat (magenta), formé par les SET-I (bleu) et Post-SET (violet)

La région SET-C s'organise sous une forme de pseudo-nœud (par le passage d'un feuillet au milieu d'une boucle), qui permet de rapprocher les deux motifs les plus conservés du domaine SET, les motifs R₁₆₂₅FINHXCXPN₁₆₃₄/E₁₆₆₁LXFDY₁₆₆₆ (Fig.II-4D). Ces derniers forment le site actif de l'enzyme, avec des résidus clés pour la catalyse (comme les R₁₆₂₅⁸⁴ et Y₁₆₆₆⁸³) et se situent entre les poches de liaison du substrat et du cofacteur. Il est important de noter que la Cys du R₁₆₂₅FINHXCXPN₁₆₃₄ est le dernier résidu coordinateur du doigt de zinc du domaine Post-SET, suggérant ainsi l'importance de ce motif. De plus, le R₁₆₂₅FINHXCXPN₁₆₃₄, ainsi que le motif G₁₅₆₁XG₁₅₆₃ de la région SET-N, permettent l'interaction avec le cofacteur SAM⁸⁵.

SET-I est une région statique de taille variable, mais de structure conservée entre les différentes HMT à domaine SET (e.g. 21 résidus pour SETD2 et 363 pour SETDB1)⁷². Étant donné la faible similarité de séquence de la région SET-I des HMT, il est supposé que cette dernière participe de manière non exclusive à la spécificité de reconnaissance du substrat⁸⁶. En effet, en coordination, la région SET-I et le domaine Post-SET forment le sillon de taille variable qui permet la liaison du substrat (Fig.II-4E) et contribuent, à un moindre degré, à la formation de la poche de liaison du cofacteur⁸⁷.

3) Les domaines d'interactions protéiques WW et PRR

Le domaine WW est un motif protéique conservé, impliqué dans des interactions protéine-protéine. Il est composé de 35-40 résidus et nommé d'après les deux résidus Tryptophane (Trp) strictement conservés et espacés de 20-22 amino-acides (W₂₃₉₅ et W₂₄₁₇ pour SETD2). Après repliement, il adopte une structure globulaire, stable, composée d'un feuillet β de 3 brins antiparallèles qui entourent le cœur hydrophobe⁸⁸. Sa structure lui permet de reconnaître des motifs riches en Prolines (Pro), dont PPxY ou PPLP⁸⁹.

Ce domaine a permis de découvrir SETD2 (HYPB) comme protéine interagissant avec la région riche en Pro de Huntingtin (Htt) ⁷⁷. De manière intéressante, SETD2 possède une région poly-Pro, appelée domaine PRR, située en amont du domaine WW. Ce domaine constitue une forme d'autorégulation de l'interaction de SETD2 avec ses partenaires, notamment l'Htt. En effet, les domaines WW et PRR de SETD2 peuvent interagir au sein de la même molécule, induisant une conformation fermée qui inhibe l'interaction du domaine WW de SETD2 avec d'autres PRR (Fig.II-5) ⁹⁰.

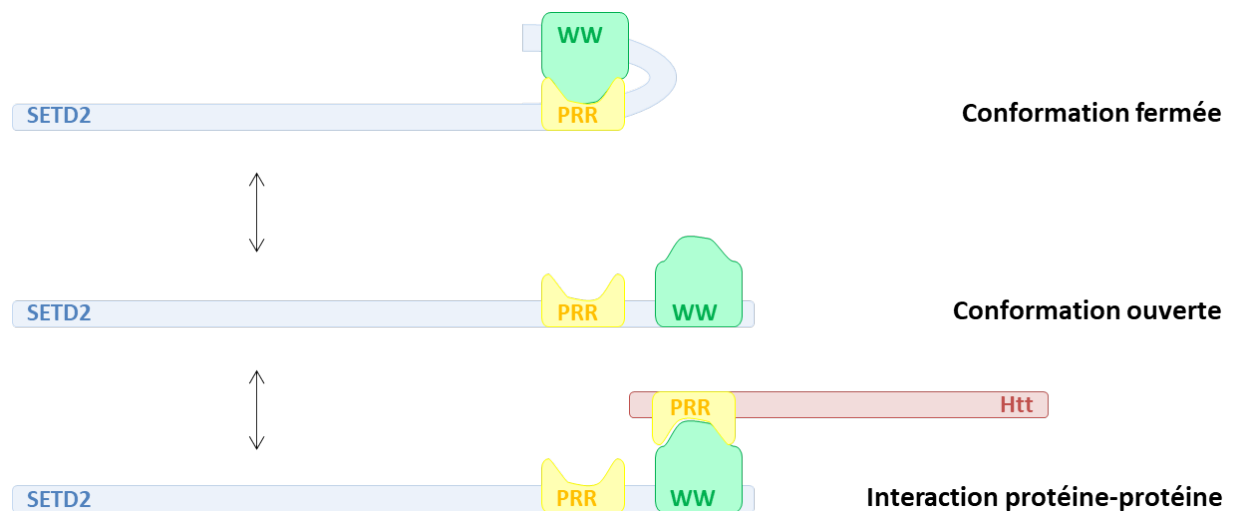
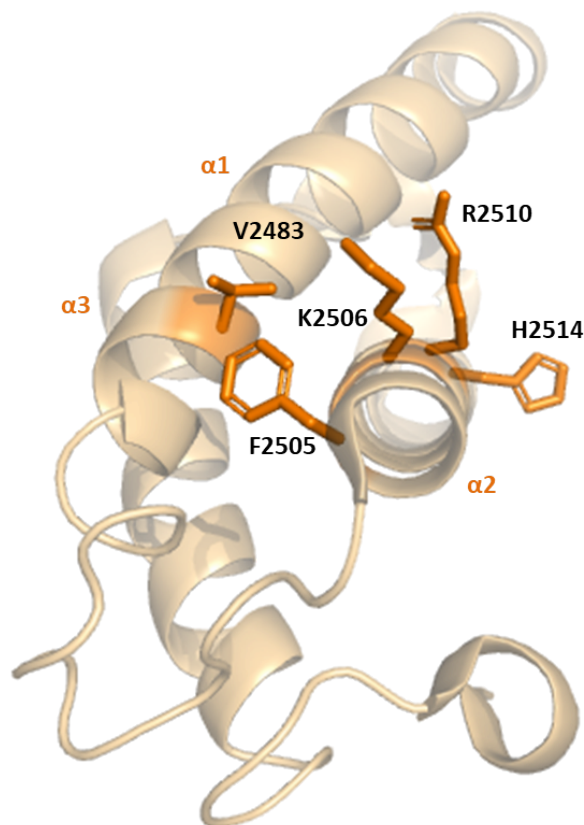


Figure II – 5 : Autorégulation de l'interaction de SETD2 avec ses partenaires, par son domaine WW

4) Les domaines d'interaction avec les RNA pol II (SRI) et hnRNP-L (SHI)

La plus grande sous-unité de la RNA polymérase II (RNA pol II) chez les eucaryotes est la CTD (C-Terminal repeat Domain), composée de multiples répétitions de la séquence $Y_1S_2P_3T_4S_5P_6S_7$. Lors de l'initiation de la transcription, ce domaine est non phosphorylé, et pour permettre le redémarrage de transcription, la CTD de la RNA pol II doit être phosphorylée au niveau des Sérines 2 et 5. La forme de CTD hyperphosphorylée, et

uniquement celle-là, est reconnue par le domaine SRI de SETD2, permettant ainsi de coupler l'élongation de la transcription avec la triméthylation du résidu H3K36 par SETD2. En effet, la fonction prépondérante liée à la marque H3K36me3 est l'élongation de la transcription. SETD2 est recrutée au niveau de la RNA pol II grâce à son domaine SRI et permet l'avancement du processus de transcription, notamment via la formation de la marque H3K36me3. De plus, des mutations de SETD2 ou des défauts de dépôt d'H3K36me3, entraînent une délocalisation de la RNA pol II ^{91,92}.



La résolution de la structure du domaine SRI de SETD2 a permis de démontrer que celui-ci est organisé en 3 hélices α , permettant ainsi de le différencier par les autres domaines d'interaction avec la CTD qui sont décrits à ce jour. L'interaction de SRI avec la RNA pol II implique les hélices $\alpha1$ et $\alpha2$ et plus particulièrement les résidus V2483, F2505, K2506, R2510, H2514 (Fig. II-6) ⁹³.

Figure II – 6 : Le domaine SRI de SETD2 et les résidus importants pour l'interaction avec la RNA pol II

En effet, en plus de son domaine SRI, SETD2 possède un domaine de liaison à hnRNP-L, appelé SHI (SETD2 hnRNP Interaction). hnRNP-L se lie aux régions riches en cytosines et adénines des pre-ARNm et permet de réguler l'épissage. Par ailleurs, il est admis que les processus d'épissage et de transcription sont liés. De manière intéressante, SETD2 permet de faire le lien entre ces deux processus majeurs de la cellule. Cela distingue SETD2 comme un candidat idéal pour faciliter le processus d'épissage, *via* sa capacité à recruter hnRNP-L au niveau du site de transcription par la RNA pol-II ^{94,95}.

B. Activité lysine méthyltransférase de SETD2

En 2012, Zheng *et al* ont résolu la structure tridimensionnelle du complexe binaire du ‘cœur catalytique’ AWS – SET – Post-SET de SETD2 avec son cofacteur SAM ⁹⁶. Quelques années plus tard, Yang *et al* et Zhang *et al* ont obtenu les deux premières structures cristallographiques du cœur catalytique en complexe ternaire avec la SAM et un peptide substrat contenant la séquence du résidu H3K36 ^{43,83}. Ces études ont permis de poser les bases moléculaires et structurales de la reconnaissance du substrat et du cofacteur, l’auto-régulation et la stabilité de l’enzyme ⁸⁴.

1) Le complexe ternaire SETD2 – SAM – H3K36/H3M36

La formation du complexe ternaire SETD2 – SAM – H3K36 est le résultat d’un mécanisme séquentiel non-ordonné qui permet la fixation du cofacteur SAM et du substrat à leurs sites distincts sur l’enzyme. Le cofacteur SAM se fixe dans une poche formée par les boucles L_{β1β2}, L_{α7β7} et L_{POST} (Fig. II-7A), permettant ainsi la stabilisation du domaine SET. Pour documenter l’interaction de SETD2 avec son substrat histone H3, Yang *et al* et Zhang *et al* ont utilisé un peptide contenant les résidus A29 à R42 de H3, où le résidu K36 est remplacé par une méthionine (peptide H3K36M). L’utilisation d’un peptide modifié repose sur le fait que l’enzyme a plus d’affinité pour le résidu H3K36M que H3K36. Ce peptide se positionne dans le site actif de l’enzyme grâce à la formation d’un pseudo-feuillet β, issu des interactions avec les boucles L_{α6β5} et L_{IN} (boucle d’auto-inhibition, Fig. II-7B).

L’interaction de SETD2 avec le résidu M36 du peptide est stabilisée par quatre résidus qui forment une cage hydrophobe autour de la lysine cible (Y1579, M1607, F1664, Y1666, Fig. II-7-C). Par ailleurs, les autres résidus du peptide renforcent cette reconnaissance par des

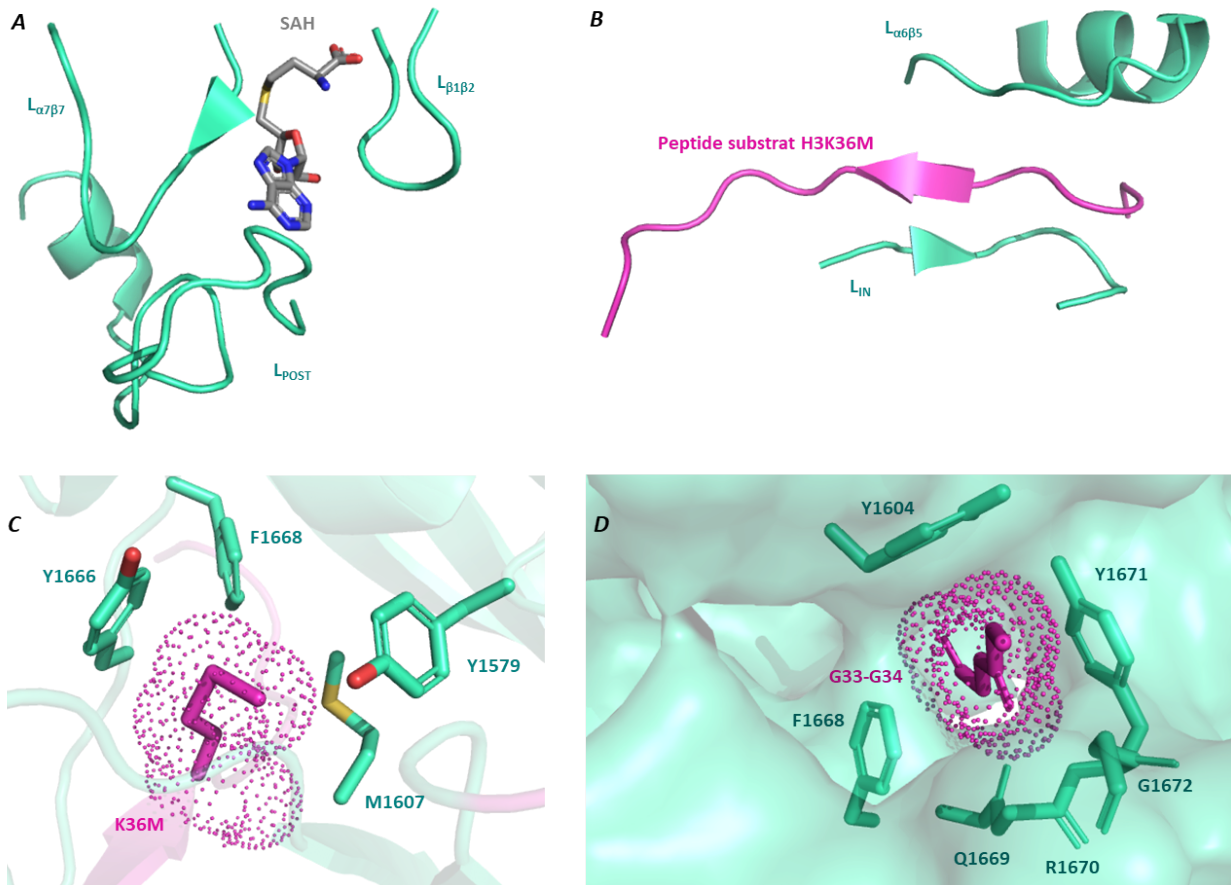


Figure II – 7 : Fixation de ligands SAH et substrat peptidique modifié (H3K36M) au domaine SET

A. Poche de fixation du cofacteur SAH (gris, CPK) formée par les boucles $L_{\beta 1\beta 2}$ (L1556-W1562), $L_{\alpha 7\beta 7}$ (M1627-C1635) et L_{POST} (Q1676-K1703)

B. Pseudo-feuillet β entre le peptide substrat H3K36M (magenta), la $L_{\alpha 6\beta 5}$ (R1598-F1606) et la L_{IN} (Q1667-A1675)

C. Cœur hydrophobe qui permet la fixation d'H3K36M

D. Tunnel étroit permettant la fixation parfaite des H3G33 et H3G34

Pour ces illustrations, SETD2 est représentée en vert-cyan (PDB : 5JJY) ⁸³

liaisons d'hydrogène et des interactions hydrophobes (14 et 15 respectivement, Tableau II-2). D'une part, les résidus A29-V35 étant de petits résidus non-chargés, s'insèrent dans une poche étroite électropositive (formée par les résidus F1589, Y1604, F1606, F1668). D'autre part, les résidus K37-R42 sont positionnés dans une région acide qui permet la liaison avec ces résidus basiques. Notamment, les résidus H39, R40 et Y41 du peptide forment des liaisons hydrogène avec les résidus E1636 et T1637 de SETD2, grâce à la configuration *trans* du résidu H3P38.

Néanmoins, une particularité de la reconnaissance du résidu H3K36 par SETD2, et en général par les méthyltransférases de la famille SET2, est la présence des deux résidus glycine (Gly) en proximité d'H3K36 (G33-G34). La structure de SETD2 démontre que ces deux Gly, qui ont la plus petite chaîne latérale (atome d'hydrogène), s'insèrent parfaitement dans un tunnel étroit, formé par les noyaux aromatiques des résidus Y1604, F1668, Y1671 et les chaînes principales des résidus Q1669-G1672. Tout autre résidu à cette position est incompatible avec ce tunnel et ne permet pas la fixation de l'enzyme à son substrat (Fig. II-7D).

En complément de ces études cristallographiques, Liu *et al* ont résolu en 2021 la première structure de SETD2 liée au nucléosome et contenant la mutation oncogénique H3K36M (Fig. II-8, PDB : 7EA8)⁹⁷. Leur travail a permis de confirmer les données cristallographiques précédemment décrites. La structure tridimensionnelle (3D) obtenue montre également la fixation d'H3K36M dans une poche hydrophobe. Les résidus V35, M36 de K37 d'H3 forment avec les résidus F1606, M1607 et A1608 de SETD2 une structure intermoléculaire en feuillets β , stabilisée par des interactions entre leurs chaînes principales. De plus, les résidus Y41, R49 et R52 d'H3 interagissent avec des résidus du domaine AWS de SETD2, ce qui permet de suggérer l'importance de ce domaine dans l'interaction avec l'histone au sein du nucléosome. De manière intéressante, la présence dans les nucléosomes de l'Histone 1 (H1) induit la compaction et la stabilisation des nucléosomes conduisant à l'inhibition de la liaison de SETD2 à l'histone H3 et à la diminution de la marque H3K36me3.

Liaisons Hydrogène		Interactions Hydrophobes	
Résidu SETD2	Résidu peptide H3K36M	Résidu SETD2	Résidu peptide H3K36M
F1606	M36	Y1579	M36
A1608	M36	1589	V35
E1636	R40, Y41	V1593	P30
T1637	H39	N1601	T32
Y1666	K37	Y1604	T32, G33
Q1669	V35, K37	M1607	M36, P38
Y1671	G33	Q1638	R40
G1672	G33	T1653	Y41
K1673	P30	F1664	M36
E1674	S31	Q1667	K37, H39
		F1668	G34, V35
		R1670	G33
		I1697	P30
		R1694	A29
		A1700	V35

Tableau II – 2 : Liaisons Hydrogène directes et interactions hydrophobes entre les résidus de SETD2 et du peptide H3K36M

Le tableau est non exhaustif. Il est basé sur l'analyse 'Ligplot' de la structure 5JJY⁸³

Les résidus T1637-H39 et E1636-R40 forment des doubles liaisons hydrogène.

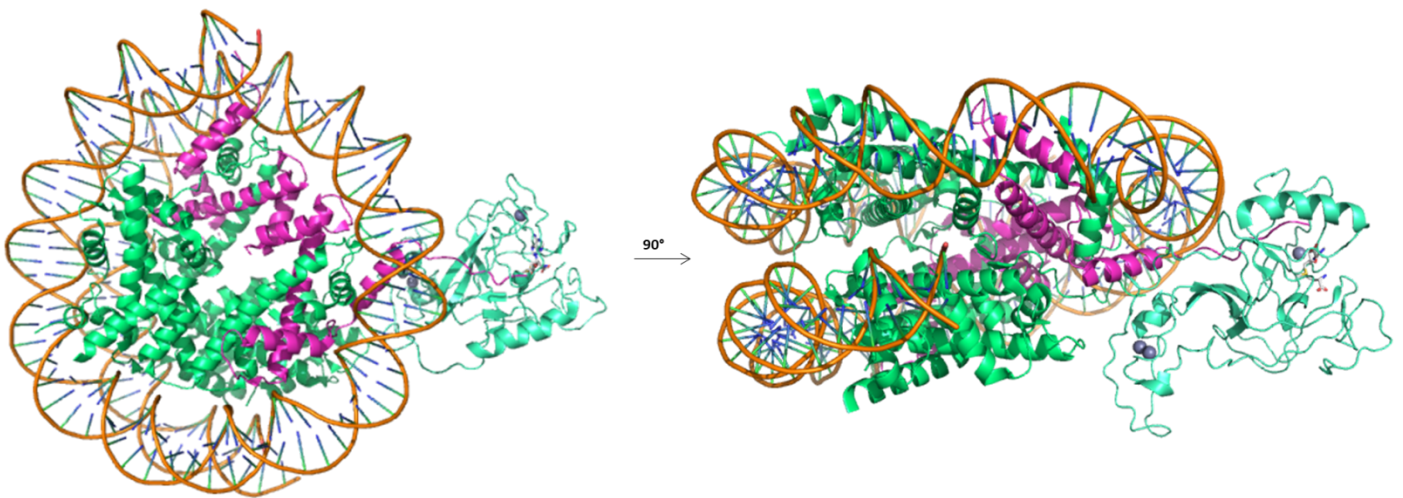


Figure II – 8: Le complexe entre SETD2, nucléosome et le cofacteur SAH (PDB : 7AE8)

SETD2 est représentée en cyan, l'histone H3 en magenta, les histones H2A, H2B et H4 et vert et l'ADN en orange. Le cofacteur SAH est en gris et en CPK. Les structures de gauche et de droite sont identiques.

2) Autorégulation de SETD2

La comparaison des structures de SETD2 en complexe binaire et ternaire permet de mettre en évidence un mécanisme d'autorégulation enzymatique grâce au domaine Post-SET et, plus particulièrement grâce à la boucle L_{IN} , (Q1667-A1675) qui relie les domaines SET et Post-SET (boucle auto-inhibitrice). En effet, ce dernier est complètement déstructuré en absence de liaison du substrat. Dans ce cas, la chaîne latérale du résidu R1670 est dirigée vers le site actif de l'enzyme, où elle occupe la position de la lysine 36 (ou méthionine 36 dans le cas du mutant) du substrat histone. Cette topologie est appelée 'conformation fermée'.

Pour passer à la 'conformation ouverte', qui est la conformation active de l'enzyme, le résidu R1670 s'éloigne de $\sim 15\text{\AA}$ du site de fixation du substrat par une rotation de 90° de la boucle auto-inhibitrice L_{IN} (Fig. II-9). L'enzyme reste sous sa forme ouverte par la formation d'une nouvelle cage hydrophobe par les résidus Y1604, Y1666, F1668, L1689 et la chaîne

latérale de R1670. Des mutations de ce dernier conduisent à une enzyme pas ou peu fonctionnelle. Ceci suggère qu'à la fois la charge et l'hydrophobie du résidu R1670 sont importantes pour la catalyse enzymatique. Par ailleurs, suite à la fixation du substrat, les résidus en C-ter du domaine Post-SET (E1692-K1703) se dirigent vers le site de liaison du substrat et forment une hélice α ($\alpha 8$). Cette dernière permet de stabiliser la liaison du substrat avec l'enzyme par des interactions avec l'hélice $\alpha 6$ et la formation d'un cluster hydrophobe en proximité du résidu V35 de H3^{83,96}. La position des résidus formant le tunnel étroit permettant la fixation des résidus G33-G34 de l'histone H3 (chaîne latérale des résidus Y1604, F1668, Y1671 et chaîne principale des résidus Q1669-G1672) reste invariable dans la conformation fermée et ouverte.

SETD2 a été également décrite comme méthyltransférase capable de s'autométhylter. La méthylation de SETD2 n'est pas dépendante de la liaison du substrat histone. Néanmoins, il est supposé qu'elle a un effet sur l'activité HMT de l'enzyme. Malgré tout, il n'y a pas d'étude qui permette de détailler ce processus et son rôle dans l'activité de SETD2⁷⁶.

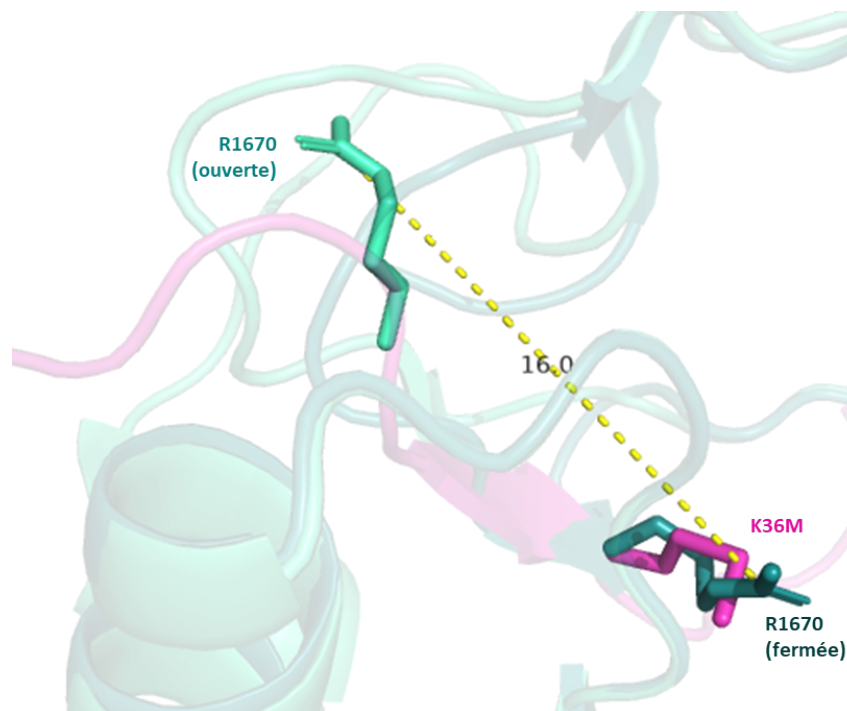


Figure II – 9: Conformation 'ouverte' (vert-cyan) et 'fermée' (bleu) de la boucle LIN de SETD2
Conformation fermée, PDB : 4H12

3) La stabilité de SETD2

De manière intéressante, malgré le fait que la délétion de SETD2 soit délétère pour la cellule, son niveau d'expression ne doit pas être en excès. Lors d'une surexpression de SETD2, cette dernière triméthyle le résidu H3K36 de manière RNA pol II indépendante⁷⁹. En conséquence, les fonctions liées à la marque H3K36me3, dont la transcription et le 'cross-talk' avec d'autres marques épigénétiques sont altérées. De plus, la surexpression de SETD2 a comme résultat son agrégation. Le fait que SETD2 soit une protéine sujette à l'agrégation, suggère que son niveau d'expression peut être régulé par un mécanisme permettant le fonctionnement adéquat de son activité.

En effet, SETD2 est une protéine instable, avec une faible demi-vie, qui est régulée par le système Ubiquitine – Protéasome (UPS – Ubiquitin Proteasome System). L'inhibition de l'UPS par l'inhibiteur du protéasome MG132, permet l'accumulation de SETD2. L'autophagie ne semble pas impliquée à la régulation de SETD2 car l'inhibition des lysosomes par la chloroquine n'influence pas le niveau protéique de SETD2. L'instabilité de l'enzyme est principalement due à sa région N-ter⁷⁹.

Cette région de SETD2 ne possède pas de domaine protéique défini et n'a pas de similarité de séquence avec d'autres protéines connues. Par ailleurs, la région N-ter de SETD2 est constituée d'acides aminés chargés avec une faible hydrophobie. Ce type de région, appelées Régions Intrinsèquement Désordonnées (IDR – Intrinsically Disordered Regions), est très flexible et n'a pas de structure définie. Cette absence de structuration définie rend ces régions plus sensibles à la protéolyse notamment par le protéasome. Au total, SETD2 possède 17 IDRs dont 11 se situent dans la partie N-ter et régulent le niveau d'expression de l'enzyme. *A contrario*, le 'cœur catalytique' AWS – SET – Post-SET de SETD2 apparaît beaucoup plus stable⁹⁸.

4) Les substrats non-histone de SETD2

Les HMT sont capables de générer des modifications post-transcriptionnelles (PMT – Post-Translational Modifications) sur des protéines non-histones. Grâce à cette nouvelle fonction des HMT, elles ont été renommées ‘KMT’ (Lysine MéthylTransférases). Comme indiqué précédemment SETD2 (KMT3 dans la nomenclature KMT) est connue comme l’unique triméthyltransférase du résidu H3K36. Ce résidu est la cible unique de SETD2 au niveau des histones. Depuis 2016, des substrats non-histone de SETD2 ont été découverts (Fig. II-10). Ces nouveaux substrats, notamment la tubuline⁹⁹ et l’actine¹⁰⁰, permettent de faire un lien entre SETD2 et le cytosquelette. La méthylation de STAT1 par SETD2¹⁰¹ suggère l’implication de SETD2 dans la signalisation cellulaire. La méthylation d’EZH2 (méthyltransférase du complexe PRC2) par SETD2¹⁰² souligne les propriétés (suppresseur de tumeur) de SETD2.

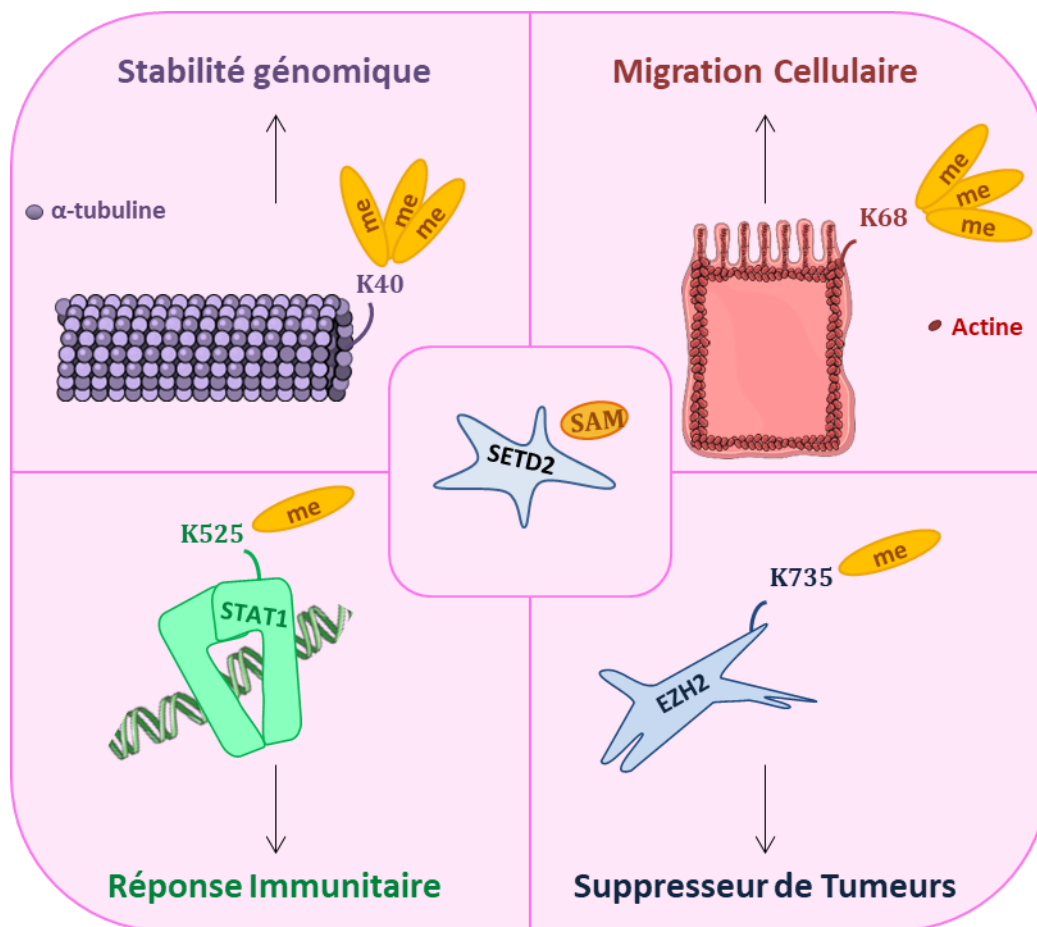


Figure II – 10: Les substrats non-histone de SETD2 et les fonctions associées à leur méthylation

a. Tubuline

Les microtubules, acteurs majeurs du cytosquelette, sont composés d'hétérodimères de tubulines α et β . Ces structures cellulaires sont nécessaires pour des fonctions diverses et majeures de la cellule, notamment la mitose. En 2016, Park *et al* ont été les premiers à démontrer un lien potentiel entre l'épigénome et le cytosquelette par l'identification d'un rôle de SETD2 sur le remodelage de ce dernier ⁹⁹. En effet, SETD2 est capable de méthyliser spécifiquement l' α -tubuline sur la lysine 40 (K40). De manière intéressante en proximité de ce résidu, se trouvent deux résidus glycine et un résidu de proline qui sont nécessaires pour la reconnaissance du substrat par SETD2. De plus, le résidu K40 est une cible connue des PTM car il est retrouvé également acétylé.

La méthylation (TubK40me3) et l'acétylation (TubK40Ac) du résidu K40 ont un rôle opposé et sont mutuellement exclusifs sur la fonction des microtubules. Cela est en accord avec les fonctions opposées des marques d'acétylation et de méthylation des lysines sur les histones. Alors que la modification TubK40Ac permet la stabilisation des microtubules, la marque TubK40me3 régule la dépolymérisation des microtubules pendant la mitose. En particulière, la marque TubK40me3 est retrouvée spécifiquement au fuseau mitotique et au 'midbody' pendant la cytokinèse. La marque TubK40me3 est essentielle également pour la polarisation et la migration des neurones car elle favorise la nucléation et la formation des microtubules. Par ailleurs, la déplétion de SETD2 induit la perte du fuseau mitotique et la présence des nombreux défauts mitotiques et des anomalies de cytokinèse ^{99,103}.

Les déterminants moléculaires de la reconnaissance de la tubuline par SETD2 ont été récemment déterminés. La méthylation de la tubuline est favorable lorsqu'elle se trouve sous sa forme dimérique et non sous forme de microtubules polymérisées. En outre, SETD2 reconnaît l' α -tubuline grâce à son domaine SRI. Des mutations de ce dernier abolissent la méthylation de la tubuline, suggérant ainsi un rôle de ce domaine pour la liaison avec l' α -tubuline. La particularité de cette interaction est que les résidus impliqués sont différents de ceux responsables pour l'interaction avec la RNA pol II. La reconnaissance de l' α -tubuline se fait grâce aux résidus chargés du domaine SRI qui interagissent avec la région chargée négativement du C-ter de la tubuline, permettant ainsi la formation d'interactions

électrostatiques. En revanche, SETD2 interagit avec la RNA pol II par des résidus aromatiques et le SRI n'est pas indispensable pour la méthylation d'H3K36¹⁰⁴.

De la même manière que pour les marques épigénétiques, les PTM du cytosquelette sont reconnues par des protéines 'readers'. Pour la première fois, la coordination fonctionnelle au niveau du cytosquelette entre un modificateur et un 'reader' épigénétique a été décrite. La marque TubK40me3 de SETD2 est reconnue par PBRM1 (PolyBRoMo – 1) grâce à son domaine BAH. Ce domaine est connu comme 'reader' de méthylation des histones, mais il n'est pas associé à la marque H3K36me3 de SETD2 (Tableau I-4). Le recrutement de PBRM1 au niveau de la tubuline méthylée permet la stabilité génomique et le bon déroulement de la mitose¹⁰⁵.

b. Actine

Le double rôle de SETD2, comme 'remodeler' du réseau de la chromatine et du cytosquelette, est à nouveau suggéré par son activité de méthyltransférase de l'actine. En effet, SETD2 triméthyle la lysine 68 de l'actine (ActK68me3), dans les régions présentant une réorganisation du cytosquelette et dans les structures particulières du réseau d'actine, dont les lamellipodes. La lysine 68 se trouve en proximité d'un résidu Pro (P70) avec un agencement similaire aux résidus P38 et K36 dans l'histone H3¹⁰⁰.

In vitro, SETD2 méthyle à la fois les monomères et les polymères d'actine alors que dans les cellules, la méthylation de la forme polymérisée (F-actine) est 100 à 1000 fois supérieure que celle de la forme globulaire non polymérisée (G-actine). L'interaction de SETD2 avec l'actine s'effectue par l'intermédiaire d'HTT, ce qui permettrait d'expliquer la présence des défauts du cytosquelette d'actine et de la migration cellulaire dans le cas de la maladie d'Huntington. Par ailleurs, la déplétion de SETD2 conduit à la perte des contacts cellulaires et à des défauts de migration. Cela suggère que SETD2 régule la migration cellulaire et l'organisation du réseau d'actine¹⁰⁰.

c. STAT1 (Signal Transducer and Activator of Transcription 1)

SETD2 pourrait être impliquée également dans la réponse immunitaire et l'inflammation, par son implication dans la réponse anti-virale, *via* l'interféron α (IFN α) et la signalisation JAK-STAT (JAnus Kinase - Signal Transducer and Activator of Transcription). SETD2 interviendrait dans ce processus, suite à l'infection par le virus de l'Hépatite B (HBV – Hepatitis B Virus). En effet, SETD2 catalyserait la monométhylation de la lysine 525 de STAT1 (K525me1) et induirait également la triméthylation d'H3K36 au niveau des promoteurs des gènes stimulés par l'interféron (ISGs – Interferon Stimulated Genes), de manière sélective à ceux qui sont dépendants de STAT1, comme *Isg15* (Fig. II-11). La méthylation du résidu K525 est essentielle pour la phosphorylation de STAT1 et sa translocation nucléaire, sans avoir un effet sur sa dimérisation ¹⁰¹.

d. EZH2

La communication ('cross-talk') entre les différentes marques épigénétiques est aujourd'hui bien décrite. Dans le cas de la marque H3K36me3 médiée par SETD2, elle est inversement corrélée avec la marque H3K27me3 issue de l'activité d'EZH2 (sous-unité catalytique du complexe PRC2). De manière intéressante, ces deux HMT ont une relation antagoniste. En effet, SETD2 monométhyle EZH2 sur la lysine 735 (K735me1). Cette méthylation induit l'ubiquitinylation et la dégradation d'EZH2 par le protéasome. SETD2 interagit avec EZH2 par les résidus 1520-1530 du domaine AWS SETD2. Plus précisément, le résidu R1523 de SETD2 forme des liaisons d'hydrogène avec les résidus R78, G79 et T80 d'EZH2. Des mutants non méthylables de K735 dans EZH2, ont une demi-vie plus élevée et un niveau de poly-ubiquitinylation diminué. Cette fonction de SETD2 permettrait l'inhibition des métastases du cancer de la prostate, suggérant ainsi une fonction de suppresseur de tumeur de SETD2 dépendante de la méthylation d'EZH2 par SETD2 ¹⁰².

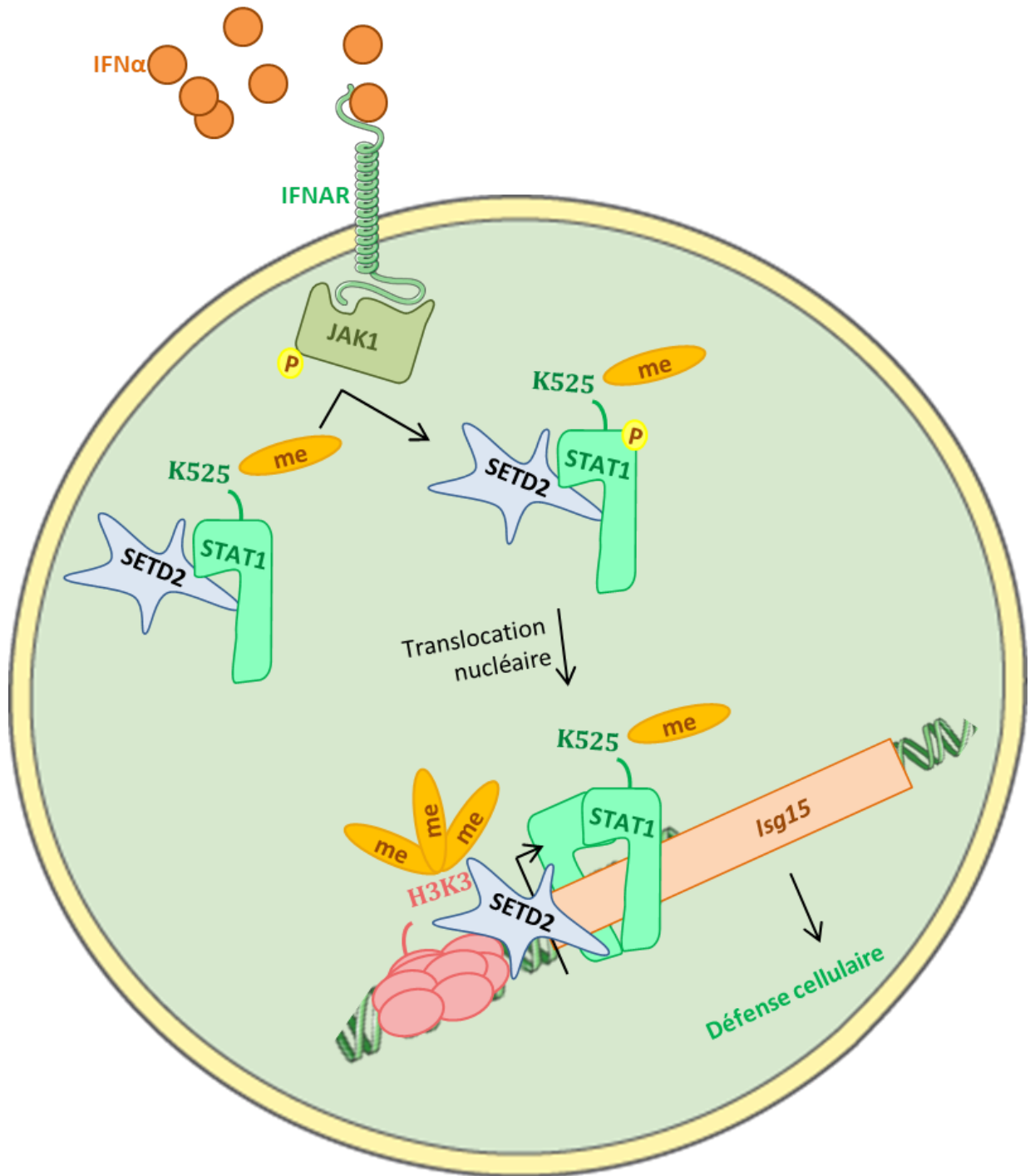


Figure II - 11: Représentation simplifiée de la réponse immunitaire médiée par la signalisation IFN α - STAT1, par l'intermédiaire de SETD2

IFNAR : Interferon α Receptor

C. SETD2 et pathologies

Setd2 est décrite comme un suppresseur de tumeurs et gène ‘moteur’ de la cancérogénèse (‘cancer driver gene’) ¹⁰⁶. Ces gènes sont définis par une différence significative entre le nombre de mutations observées et le nombre des mutations dues au hasard. Des mutations des ‘cancer driver genes’ confèrent aux cellules des avantages sélectifs par rapport aux cellules normales et contribuent à la progression tumorale. Dans le cas de SETD2, des mutations sont retrouvées dans le gène entier, mais elles sont surtout localisées au niveau des domaines AWS – SET – Post-SET (‘cœur catalytique’ de l’enzyme, Fig. II-12A). Selon les données issues de la banque de données ‘COSMIC’ (Catalogue Of Somatic Mutations In Cancer), 46% des mutations du gène et 74% des mutations du ‘cœur catalytique’ (AWS – SET – Post-SET) sont des mutations faux-sens (Fig. II-12B). Ces mutations sont observées majoritairement dans le cas des carcinomes rénaux à cellulaires claires (ccRCC – Clear Cell Renal Cell Carcinoma), d’hémopathies malignes et des gliomes.

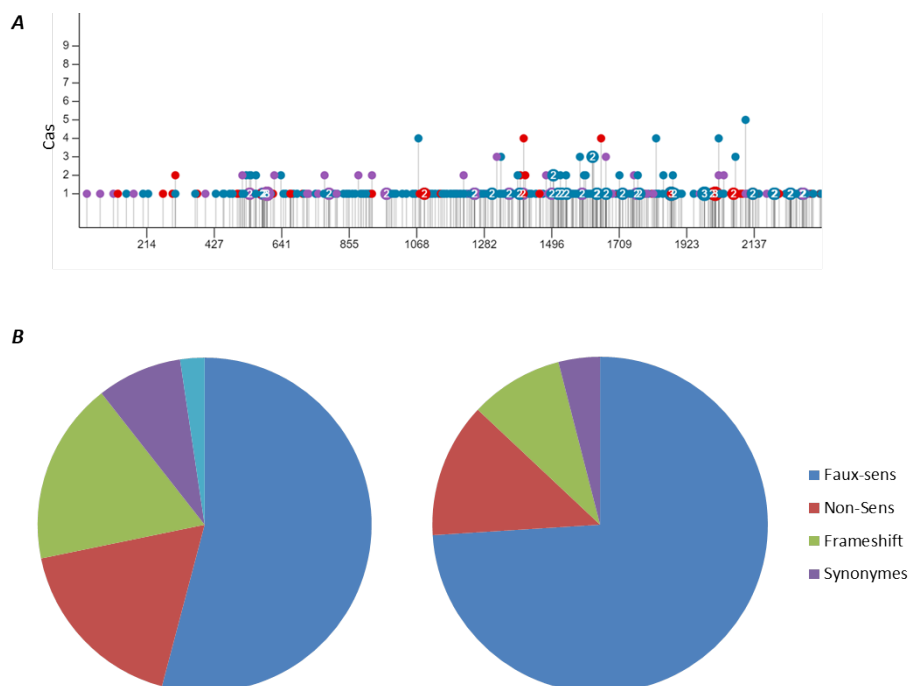


Figure II – 12: Setd2, un gène moteur de cancérogenèse

A. Mutations de *Setd2* et leurs positions selon les données de TCGA

B. Distribution des mutations somatiques au sein du gène total de SETD2 (gauche) et du triplet catalytique (droite)

TCGA : The Cancer Genome Atlas

1) SETD2 et ccRCC

Le ccRCC est le sous-type de cancer du rein prépondérant à l'âge adulte, avec 4-8% des patients présentant des mutations de SETD2. Les premières mutations dans le gène de *Setd2* ont été découvertes en 2010¹⁰⁷. L'analyse génomique démontre dans un premier temps une diminution d'expression de SETD2, suite à la perte d'un des deux bras du chromosome 3 (3p où *Setd2* se situe), sans avoir un effet sur la triméthylation d'H3K36me3. Néanmoins, cette mutation mono-allélique de SETD2 induit la perte d'activité vis-à-vis à l' α -tubuline et donc une instabilité génomique¹⁰⁸.

Dans le cas des stades cancéreux plus avancés et des métastases, il y a également présence des mutations caractérisées comme 'perte de fonction' dans l'allèle restant qui induisent un niveau d'H3K36me3 plus faible (Fig. II-13)^{109,110}. La banque de données COSMIC cite *Setd2* comme l'un des cinq gènes les plus mutés pour les cas de ccRCC, démontrant ainsi la spécificité et l'importance de SETD2 pour le développement et l'avancement de ce type de cancer.

Setd2 est un gène connu pour la régulation de la transcription et le maintien de la forme de l'hétérochromatine de manière à éviter la transcription cryptique. De manière intéressante, dans le cas des patients atteints de ccRCC, les mutations fonctionnelles de SETD2 sont associées à l'hypométhylation d'ADN, ce qui induit l'altération de l'expression génique et du processus de transcription¹¹⁰. Par ailleurs, la perte d'activité de SETD2 induit des défauts de recrutement du complexe FACT (Facilitates Chromatin Transcription), l'augmentation du stress cellulaire suite à l'endommagement de l'ADN qui ne peut pas être réparé, ainsi que la dérégulation du processus d'épissage^{109,111}.

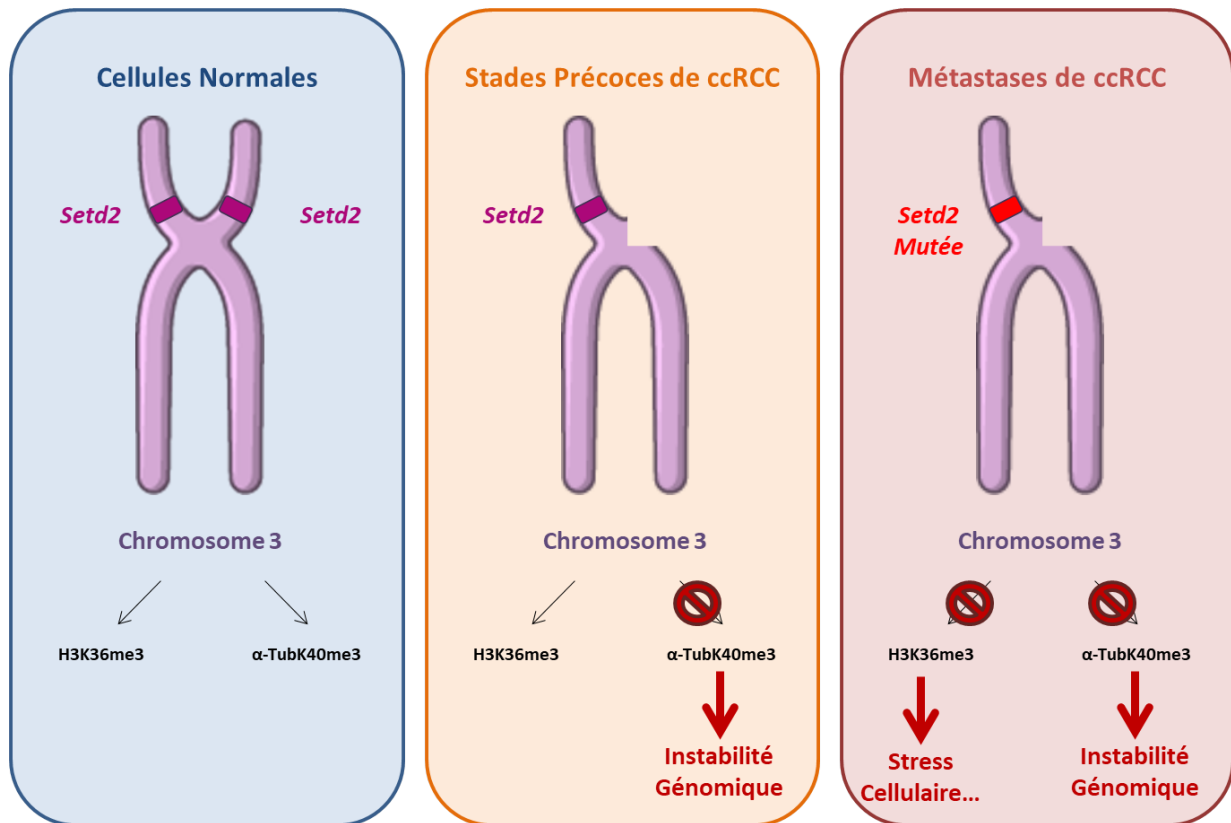


Figure II – 13: Mutations de Setd2 et leurs effets cellulaires lors de l'avancement des stades de ccRCC

2) SETD2 et hémopathies

Des mutations de SETD2 sont également documentées dans des hémopathies (Fig. II-14). SETD2 est retrouvée mutée dans 6% des leucémies myéloïdes aiguës (AML - Acute Myeloid Leukemia). Ce type de leucémie est caractérisé par des réarrangements chromosomiques (majoritairement des réarrangements de *Mll* dans plus de 50% des cas) et nécessite des perturbations moléculaires pour arriver au stade de malignité. De manière récurrente, des mutations de *Setd2* sont documentées dans le cas des patients présentant des réarrangements chromosomiques de *Mll*. Par ailleurs, les mutations de SETD2 sont décrites comme 'perte de fonction' car elles entraînent une diminution de la marque H3K36me3. La déficience de SETD2 induit l'activation des voies de signalisation mTOR (mammalian Target Of Rapamycin) et JAK/STAT qui sont connues comme étant impliquées dans la

leucemogénèse. SETD2 est décrite comme une protéine suppresseur de tumeurs avec un rôle dans l'initiation et la progression de la leucemogénèse ²⁷.

La perte de *Setd2* dans les cellules myélodysplasiques accélère la formation d'AML. Ceci est due à l'augmentation de la capacité de prolifération des cellules souches et progénitrices hématopoïétiques (HSPC – Hematopoietic Stem and Progenitor Cells), à l'altération du cycle cellulaire, et à la diminution de la capacité de différenciation et de mort cellulaire. En général, le faible niveau d'expression de SETD2 est lié à un mauvais pronostic ¹¹². De manière intéressante, une autre étude démontre que la perte bi-allélique de *Setd2* arrête la prolifération et favorise la différenciation des HSPC. Néanmoins, elle est également associée à une augmentation du signal d'endommagement de l'ADN et à un niveau plus élevé du marqueur de stress génomique, γ H2A.X. Par conséquent, ces cellules ont une accumulation des mutations, ce qui contribue à la dérégulation hématopoïétique, en parallèle de la diminution du nombre des HSPC ¹¹³.

SETD2 est également impliquée dans 10% des leucémies lymphoblastiques aiguës (ALL – Acute Lymphoblastic Leukemia). L'étude de Mar *et al* a permis d'identifier SETD2 comme un des modificateurs épigénétiques impliqués dans la rechute d'ALL. En effet, *Setd2* présente des mutations *de novo* dans le cas de patients pédiatriques en rechute d'ALL, qui est documentée dans 15-20% de ces patients. Ces patients sont résistants aux traitements chimiothérapeutiques et ont un faible niveau de survie ¹¹⁴. De plus, le même groupe a décrit que des mutations de *Setd2* sont résistantes aux agents chimiothérapeutiques qui endommagent l'ADN, mais n'ont pas d'effet sur les agents non endommageant de l'ADN (Fig. II-15). Cette résistance peut être restaurée par l'inhibition de la déméthylase d'H3K36, KDM4A, suggérant ainsi une potentielle piste de traitement de ces patients ¹¹⁵.

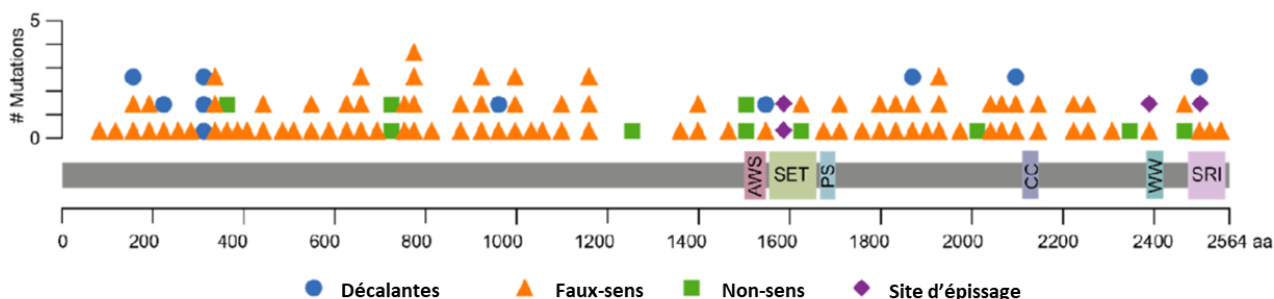


Figure II – 14: Mutations de SETD2 dans les hémopathies

Schéma modifié par Skucha *et al* ¹¹⁶

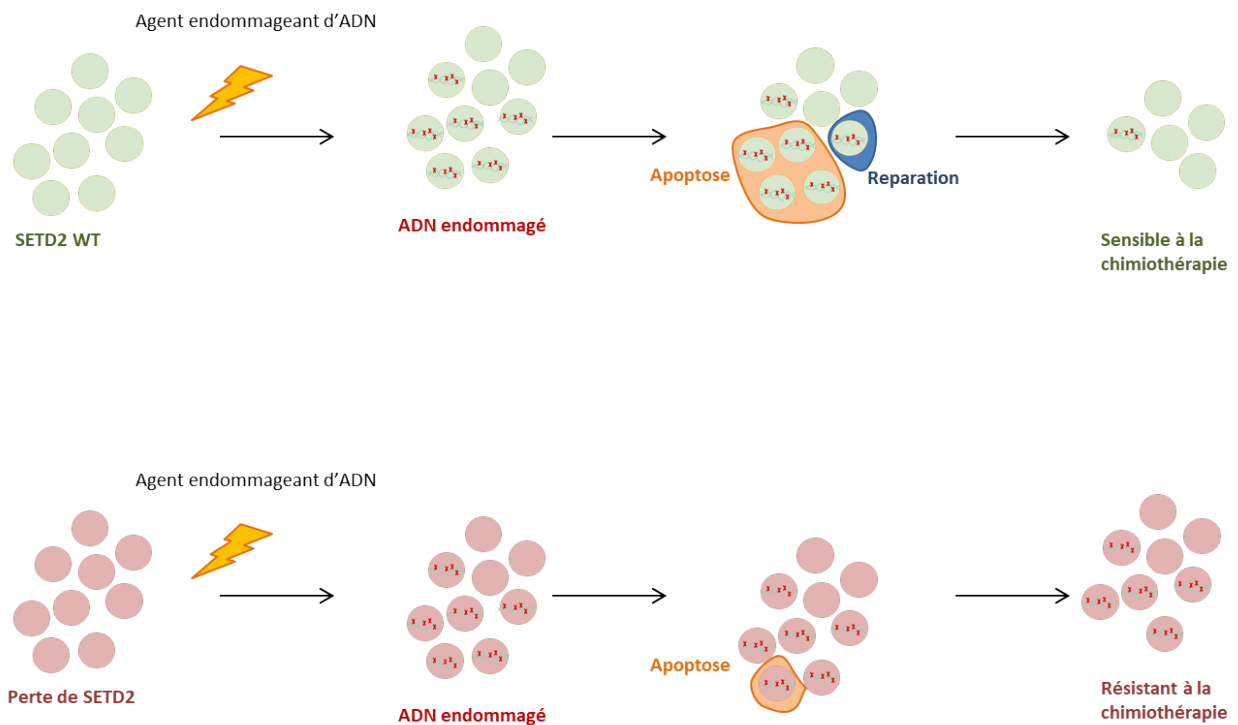


Figure II – 15: *Modèle du processus de résistance des cellules aux traitements chimiothérapeutiques*¹¹⁵

3) SETD2, gliomes et mutations d'H3G34

De manière similaire à la mutation oncogénique H3K36M, la Glycine 34 (G34) de l'histone H3 est également retrouvée mutée dans 50% des gliomes pédiatriques. En effet, le résidu G34 est muté entre autres en Valine (V), Arginine (R), Aspartate (D). Ces mutations impactent la méthylation du résidu H3K36 par SETD2, entraînant ainsi une diminution de la marque H3K36me3. Plus précisément, la G34V/R/D bloque la reconnaissance d'H3K36 par SETD2. SETD2 dispose d'un tunnel étroit, important pour l'interaction avec son substrat. Dû à la taille de ce tunnel, l'accès est restreint à des résidus Glycine, provoquant ainsi un encombrement stérique quand un autre résidu plus volumineux se trouve à cette position. L'absence de la marque H3K36me3 conduit au non recrutement des protéines 'readers' comme la machinerie de réparation des mésappariements d'ADN (MutS α)¹¹⁷.

4) Le Syndrome de Sotos

Des mutations de SETD2 sont également décrites dans le cas des patients atteints de Syndrome de Sotos (SoS). Ce syndrome pédiatrique, décrit pour la première fois en 1964, est caractérisé par une surcroissance pendant l'enfance. Les enfants présentant le SoS ont une très grande taille et/ou avec des tissus présentant de surcroissance. Par ailleurs, ils ont des défauts d'apprentissage et des troubles sociaux. *Nsd1* a été le premier gène lié au SoS, avec 60-90% des patients représentant des mutations de celui-ci. Récemment, des séquençages génomiques des enfants avec ce syndrome, ont démontré la présence des mutations des *Setd2* et *Dnmt3a*, dans le cas d'absence des mutations de *Nsd1*.

Le fait que SETD2 et NSD1 soient toutes les deux des méthyltransférases d'H3K36, et que DNMT3A est un 'reader' d'H3K36 méthylé, permettent de suggérer qu'il y a un lien entre ce syndrome et la méthylation du résidu H3K36^{118,119}. De plus, des souris SETD2 knock-out conditionnels présentent un phénotype 'SoS-like' avec des anomalies similaires que les patients atteints de syndrome de Sotos. Ces souris ont des défauts dans l'interaction sociale, l'apprentissage moteur et la mémoire spatiale. De plus leur analyse a démontré que SETD2 a un rôle dans la connexion neuronale¹²⁰.

III. MOncytic leukemia Zinc finger, MOZ

A. Les Histone AcétylTransférases (HAT)

Les Histone AcétylTransférases (HAT), appelées aussi Lysine (K) AcétylTransférases (KAT) ont été initialement découvertes comme responsables de l'ajout d'un groupement acétyl (issu du cofacteur-donneur universel d'acétyl, l'AcétylCoA) au niveau des lysines situées en N-terminal des histones. Cet ajout d'une charge négative permet de neutraliser la charge positive des histones, induisant ainsi la déstabilisation du complexe avec l'ADN (chargé négativement) et l'ouverture de la chromatine (Fig. III-1) ¹²¹. Il est important de noter que les histones acétylées interagissent toujours avec l'ADN. Cette interaction est plus faible que leur forme non acétylée et permet la formation des mononucléosomes ^{122,123}.

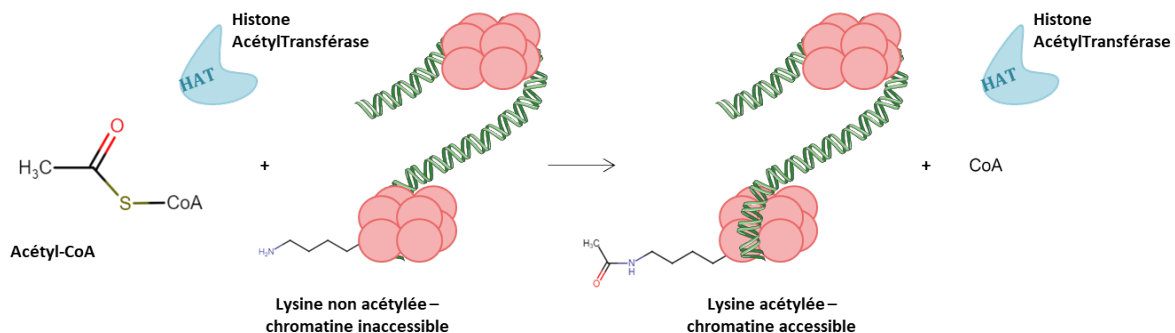


Figure III - 1 : Mécanisme simplifié d'acétylation des histones par des HAT avec ouverture de la chromatine

La forme ouverte de la chromatine suite à l'acétylation des histones est associée principalement à l'activation de la transcription, mais aussi à la réplication. Comme il a été initialement décrit dans le cas du facteur de transcription TFIIIA, il est nécessaire que les histones soient acétylées de manière à stabiliser l'interaction avec des facteurs de transcription ¹²⁴. Cette interaction est inhibée dans le cas des histones non acétylées en raison de la formation du tétramère 2 x (H3-H4), également déstabilisé par acétylation des

lysines. En parallèle, lors du processus de réplication, les histones H3 et H4 néosynthétisées sont recrutées respectivement sous leur forme mono ou di-acétylée, ce qui facilite leur positionnement dans les nucléosomes et l'assemblage de la chromatine ¹²⁵.

L'acétylation des lysines s'effectue principalement sur l'histone H3 au niveau des lysines 9, 14, 18, 23 (H3K9, H3K14, H3K18, H3K23), mais aussi au niveau des lysines 5, 8, 12 et 16 de l'histone H4 (H4K5, H4K8, H4K12, H4K16). En 2000, Strahl *et* Allis ont remarqué de manière intéressante que toutes les lysines qui représentent des cibles d'acétylation potentielles au niveau des histones se situent dans un espace régulier et périodique entre elles, similaire à celui nécessaire pour la formation d'une hélice α (3,6 Å) ².

À ce jour, on distingue trois familles de HAT chez l'Homme, appelées GNAT, p300/CBP et MYST ^{44,126}.

1) Les GNAT

La famille des GNAT (General Control Non-repressible 5-related histone N-AcetylTransferases) doit leur nom au membre 'historique' yGCN5 (yeast General Control of amino acid synthesis protein 5). yGCN5 est en effet la première acétyltransférase identifiée, localisée dans le noyau et liée à la transcription ¹²⁷. hGCN5 et PCAF (p300/CBP Associated Factor), appelées aujourd'hui respectivement KAT2A et KAT2B, sont les deux seules GNAT présentes chez les mammifères et sont capables d'acétyler les lysines 9, 14, 18 de H3 (H3K9, H3K14, H3K18) ainsi que l'histone 2B (H2B) ¹²⁸.

De manière intéressante, la famille des GNAT regroupe des HAT mais également d'autres enzymes capables d'acétyler des substrats autres que des protéines, dont des antibiotiques, des vitamines et des catécholamines. La caractéristique commune des membres de cette famille est leur structure tertiaire, composée de six hélices α et neuf feuillets β , ainsi que leur capacité à former des multimères ¹²⁹.

Au milieu des années 2000, les structures des γ GCN5 et PCAF ont été résolues (Fig. III-2) et ont permis de comprendre le mécanisme catalytique qui nécessite la formation d'un complexe ternaire, suggérant ainsi un mode de catalyse séquentiel¹³⁰⁻¹³². Ces structures ont révélé des résidus importants pour la spécificité de la liaison entre l'histone et ces enzymes. Notamment, en ce qui concerne la catalyse de l'acétylation du résidu H3K14, le motif nécessaire pour la reconnaissance du site de liaison enzymatique est G-K14-X-P, évoquant l'importance des résidus G13 et P16¹³³.

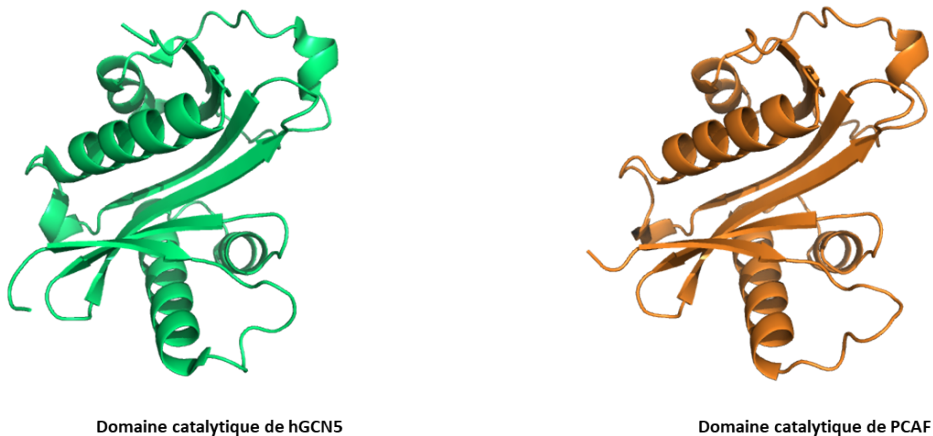


Figure III - 2 : Structures du domaine catalytique HAT des hGCN5 (PDB : 1Z4R) et CPAF (PDB : 1CMO).

2) CBP/p300

Creb Binding Protein (CBP ou KAT3A) et p300 (KAT3B, protéine d'environ 300kDa) sont deux enzymes paralogues qui forment la famille de CBP/p300, avec une faible identité de séquence avec les autres HAT. Elles ont été identifiées respectivement en 1985 et 1993 en tant que partenaires de 'Early region 1A (E1A)' - un facteur de transcription oncogénique d'adénovirus - et de 'cAMP response element binding protein (CREB)'^{134,135}. En 1996, deux groupes, Ogryzko *et al* et Bannister *et Kouzarides* ont prouvé que CBP et p300 possèdent une activité intrinsèque de HAT et sont capables d'acétyler les core histones au sein des mononucléosomes^{136,137}.

En 2008, Liu *et al* ont résolu la structure cristallographique de p300 (PDB : 3BIY) et ont démontré que cette famille utilise un mécanisme catalytique différent des autres HAT (mécanisme séquentiel Theorell – Chance, Fig. III-3A) ¹³⁸. Dans le cas de p300, le cofacteur AcétylCoA se fixe dans un premier temps avec l'enzyme et ensuite, le substrat histone interagit très faiblement avec p300, permettant le transfert rapide du groupement acétyle à la lysine cible. Ce type de mécanisme sans formation de complexe ternaire est en accord avec la structure résolue, qui dévoile un site de fixation de substrat acide. Ce mécanisme a été étendu également dans le cas de CBP grâce à la résolution de la structure de son domaine catalytique qui n'est arrivée qu'en 2017 (Fig. III-3B) ¹³⁹.

CBP et p300 présentent une analogie de régulation interne avec les kinases qui sont capables de s'auto-activer par phosphorylation. Leur domaine catalytique HAT contient une boucle (boucle AiL) composée de treize lysines capables d'être acétylées. Cette région basique a une forte affinité pour le sillon acide qui correspond au site de fixation du substrat, induisant de cette manière l'inactivation de l'enzyme. Par contre, lorsque cette boucle est acétylée, elle ne se situe plus dans le site actif de l'enzyme, permettant l'activation de ce dernier. La famille des KAT3 est donc capable de s'autoréguler grâce à sa boucle auto-inhibitrice et sa capacité d'auto-acétylation ^{140,141}.

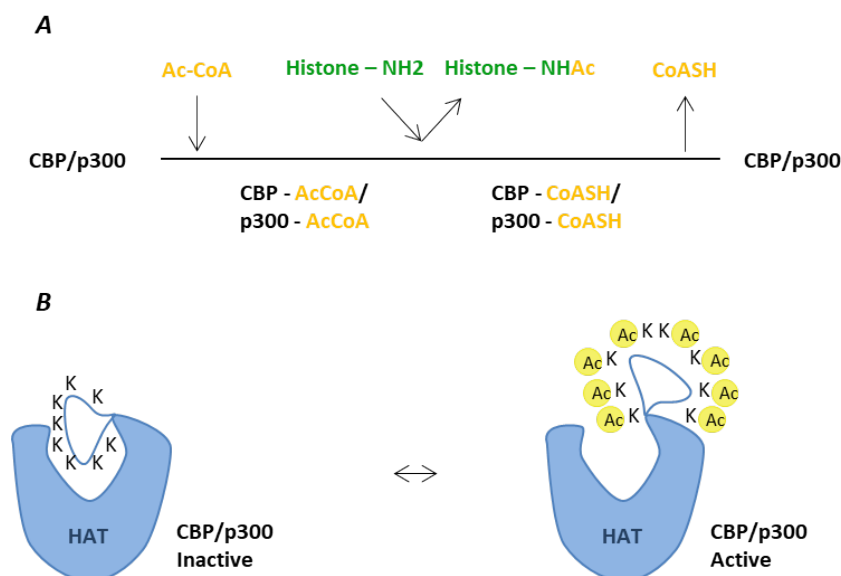


Figure III – 3 : Mécanisme d'activité des CBP/p300

A. Mécanisme catalytique des CBP/p300 selon le modèle Theorell –Chance

B. Mécanisme d'autorégulation des CBP/p300

3) Les MYST

La famille des MYST doit son nom aux membres fondateurs MOZ (MOncyclic leukemia Zinc finger)¹⁴², YBF2/SAS3 (Identical with SAS2/ Something About Silencing 3)^{143,144}, SAS2 (Something About Silencing 2)¹⁴⁴ et TIP60 (Tat Interactive Protein 60kDA)¹⁴⁵. Ces enzymes ont été identifiées dans les années 1996-1997 au cours d'études menées sur certains mécanismes cellulaires comme l'activation de la transcription mais aussi des anomalies génétiques comme les translocations chromosomiques. Leur seul point commun est le domaine catalytique MYST qui est composé d'un site de fixation du cofacteur AcétylCoA et d'un doigt de zinc de type C2HC (composé de trois Cystéines et une Histidine).

Les MYST possèdent un mécanisme catalytique différent des autres HAT, celui-ci étant de type ping-pong. Le site catalytique est composé d'une dyade cystéine (Cys-C) - glutamate (Glu-E), deux résidus qui sont proches dans la séquence et la structure et qui sont conservés dans toutes les MYST. Lors de la catalyse, il y a formation d'une forme intermédiaire de cystéine acétylée (Cys-Acétylée) grâce à la déprotonation du résidu Cys par le résidu Glu et l'attaque nucléophile de la Cys déprotonée sur l'AcétylCoA. À la suite, le Glu est responsable de la déprotonation de la lysine cible qui se lie dans le site actif et qui, à la fin de la réaction, se trouve acétylée par attaque nucléophile de la Cys-Acétylée (Fig.III-4)¹⁴⁶.

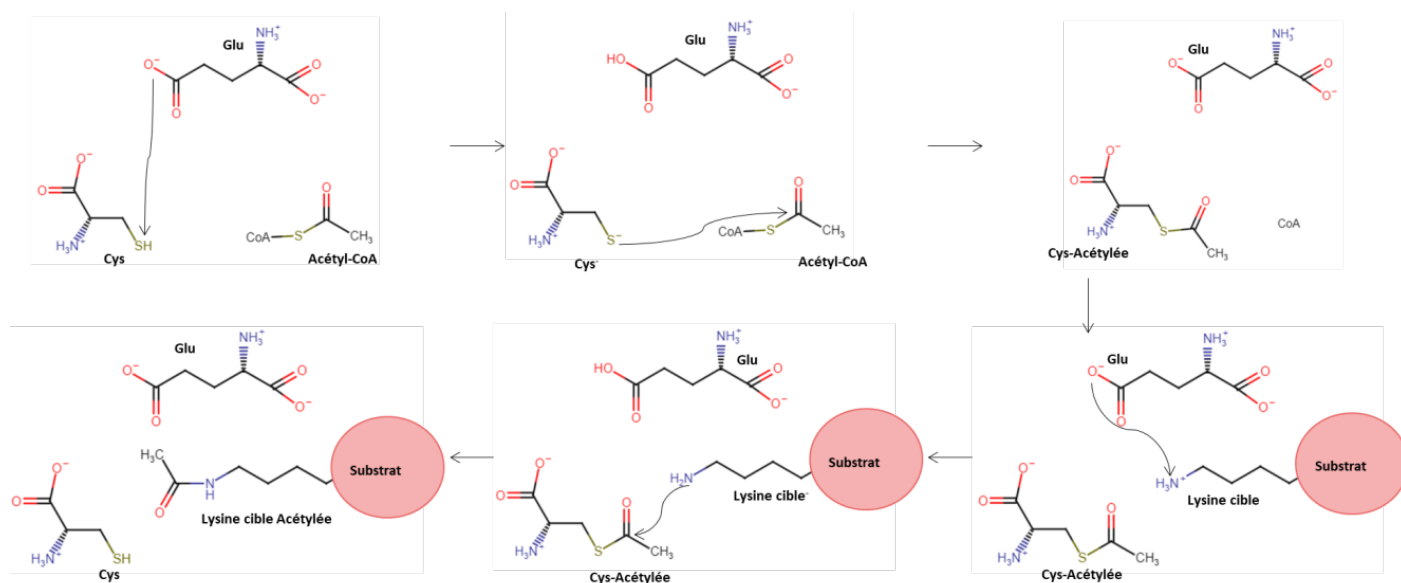


Figure III - 4 : Mécanisme catalytique de type ping-pong des MYST

De même que les CBP et p300, les MYST ont besoin d'être auto-acétylées pour être actives. La résolution de la structure d'une MYST chez la levure (ESA1) et d'une MYST chez l'Homme (hMOF) par Yuan *et al* en 2012 démontre la présence d'un groupement acétyle lié à une lysine présente dans le site actif et conservée au sein des MYST. Cette acétylation qui se produit en *cis* (au sein de la molécule) entraîne l'altération de l'espace occupé par le site actif. En effet, cette lysine se trouve sur une boucle flexible et lorsque l'acétylation se produit, la boucle se stabilise et entraîne l'accessibilité du site actif au substrat. Dans le cas contraire, cette lysine non acétylée reste dans la région du site actif et perturbe l'accession du substrat ¹⁴⁷.

Le dernier élément qui s'avère important pour le domaine catalytique des MYST est le domaine en doigt de zinc C2HC qui est 'insolite' et spécifique de ces HAT (sauf pour ESA1). Ce domaine est nécessaire pour l'activité enzymatique en facilitant la liaison avec la chromatine et l'interaction avec les substrats histones ^{148,149}. Par contre, dans certains cas comme celui d'YBF2/SAS3, l'interaction du domaine catalytique avec le substrat est présente uniquement dans le cas des histones libres et non dans le cas des histones liées à l'ADN ^{148,150}. Cette préférence de substrat peut être liée au fait que les MYST sont considérées comme les sous unités catalytiques de complexes transcriptionnels qui permettent leur recrutement au niveau chromatinien.

Chez l'Homme, on distingue cinq MYST - les homologues MOZ et MORF, et les TIP60, hMOF et HBO1. Ces HAT font partie des complexes multi-protéiques, leur formation étant nécessaire pour que les enzymes exercent leur activité. Le Tableau III-1 récapitule les principaux membres des complexes protéiques des MYST humaines. Il est important de souligner que ces protéines sont aussi capables d'acétyler des protéines non histones comme p53, qui a été le premier exemple de protéine non-histone acétylée par des HAT ¹⁵¹. Leurs cibles histones ainsi que leurs principales fonctions cellulaires sont présentées dans le Tableau III-2. De même, l'alignement de leurs séquences avec représentation des résidus importants peut être trouvé dans la figure III-5 (Fig.III-5).

MOZ	MORF	HBO1	hMOF	TIP60
BRPF1	BRPF1	JADE1/2/3	hMSL1	EPC1
ING5	ING5	ING4/5	hMSL2	ING3
hEAF6	hEAF6	hEAF6	hMSL3L1	hEAF6

Tableau III – 1 : Les MYST humaines en tant que sous-unités catalytiques des complexes multi-protéiques

Dans le cas de TIP60 sont notées uniquement les sous-unités équivalentes des autres complexes (MOZ/MORF, HBO1,hMOF). Cela est fait pour des raisons de simplicité, dues au fait que TIP60 fait partie de très gros complexes.

MYST	Cible histone	Cible non histone	Fonction
MOZ (<i>MYST3/KAT6A</i>)	H3K9, H3K14, H3K23, H4K16	p53K120, p53K382	Activation de la transcription Renouvellement des cellules souches hématopoïétiques
MORF (<i>MYST4/KAT6B</i>)	H3K9, H3K14, H3K23, H4K16	p53K120, p53K382	Activation de la transcription Neurogenèse et skeletogenèse
HBO1 (<i>MYST2/KAT7</i>)	H4K5, H4K8, H4K12, H3K14	JADE1, MCM	Réplication de l'ADN Régulation de la transcription
hMOF (<i>MYST1/KAT8</i>)	H4K16	p53K120, MSL1/3	Activation de la transcription Compensation du chromosome X
TIP60 (<i>KAT5</i>)	H4K5, H4K8, H4K12, H4K16, H2A	ATM, FOXP3	Co-Activateur des protéines clés de la cellule (<i>e.g.</i> NF-κB, c-MYC, E2F1) Facteur de Transcription

Tableau III – 2 : Les membres de la famille MYST retrouvés chez l'Homme et leurs fonctions

Ce tableau est non exhaustif. Pour des raisons de synthèse et de simplicité, il représente uniquement les modifications et les fonctions les plus importantes et les plus communes. Dans le contexte cellulaire, les cibles ainsi que les fonctions de chacune de ces protéines varient selon le groupe protéique dans lequel elles se trouvent.

Tip60	TRMKNIECIELGHRHLKPWYFSPYPQELTTLPLVYL CE F CL KYGRSLKCLQR HL TKCDLR	60
hMOF	TKVKYVDKIHIGNYEIDAWYFSPFPEDYQKPKLWL CE Y CL KYMKYEKSYR FL G QC QWR	60
HBO1	EGSNMIKTIAFGRYELDTWYHSYPPEYARLGR LY M CE F CL KYMKSQ T ILRR HMA K CV WK	60
MOZ	PQVRCPSVIEFGKYEIHTWYSSYPYQEYSRLPK LY L CE F CL KYMK S RTIL QQ H M K CGWF	60
MORF	DCGRYPSVIEFGKYEIQTWYSSYPYQEYARLPK LY L CE F CL KYMK S KNILLR H SK K CGWF	60
	. . * :*. . . . ** *:*:*:	* :*:*:**** : . * :*
Tip60	HPPGNEIYRKGTISFFEIDGRKNKSYSQNLCLLAKCFLDH K TLYYDTPFLFYVMTEYDC	120
hMOF	QPPGKEIYRKSNI SVYEVVDGDKHKIYCQNLCLLAKLFLDH K TLYFDVEPFVFIYLTEVDR	120
HBO1	HPPGDEIYRKGSISVFEVDGKKNKIYCQNLCLLAKLFLDH K TLYYDVEPFLFYVMTEADN	120
MOZ	HPPANEIYRKNNISVFEVDGNVSTIYCQNLCLLAKLFLDH K TLYYDVEPFLFYVLTQNDV	120
MORF	HPPANEIYRRKDL SVFEVDGNMSKIYCQNLCLLAKLFLDH K TLYYDVEPFLFYVLTKNDE	120
	:**..****: :*:*:*:* . *.***** ***** * ***:*.:**:*:*:*: *	
Tip60	KGFHIVGYFSKEKESTEDYNV CI ILTLPPYQRRGYG K LLIEFSYELSKVEGKTG TP E K PL	180
hMOF	QGAHIVGYFSKEKESPDGNNV CI ILTLPPYQRRGYG K FLIAFSYELSKLESTVGS P E K PL	180
HBO1	TGCHLIGYFSKEKNSFLNYNV CI ILTMPQYMRQGYG K MLIDFSYLLSKVEEKVGS P E R PL	180
MOZ	KGCHLVGYFSKEKHCQQKYNV CI MILPQYQRKG Y GRFLIDFSYLLSKREGQAGS P E K PL	180
MORF	KGCHLVGYFSKEKLCQQKYNV CI MIMPQHQRQ G FRFLIDFSYLLSRREGQAGS P E K PL	180
	* *:*:***** . **:*:*:*:*:*:*:*:*:*:* ** *** **:* * .*:**:*:	
Tip60	SDLGLLSYRSYWSQTILEILMGLKSESGERPQITINEISEITSIKKEDVI ST LQYLN LI N	240
hMOF	SDLGKLSYRSYWSVWVLEILRDFR----G-TLSIKDLSQMTSITQNDI IST LQSL N MV K	234
HBO1	SDLGLISYRSYWKEVLLR YL H N FQ----GKEISIKEISQETAVNPVDIV ST LQALQ M L K	235
MOZ	SDLGRLSYMAWKS V ILECLY H QN----DKQISIKKLSKLTGIC P QDIT ST LHHL R ML D	235
MORF	SDLGRLSYLA W KSVILEYLY H HH----ERHISIKAIS R ATGMCP H DIAT T LQHL H MI D	235
	**** :* :*. .:* . * . :*: :*. *.: * :*: * .:..	
Tip60	YYKGQYILTLSEDIVDGERAML--K R LLR I DSK L H F TP	278
hMOF	YWK G QH V IC V TP K LVE H L K SA Q Y K - K PP I TV D SV C L K W A P	274
HBO1	YWK G K H L V L K R Q DL I DE W IA K E A K R S N S N K T MD P S C L K W T P	276
MOZ	FRSD Q F V I R RE K L I Q D H M A K L Q L N -LR P VD V D P E C L R W T P	275
MORF	KRD G R F V I IR R E K L I L S H M E K L K T C -SR A N E L D P D S L R W T P	275
 :. . :* .*:*:*:	

Figure III – 5 : Alignement des séquences des cinq MYST humaines.

Le code couleur vert représente la région du doigt de zinc de type C2HC et les trois cystéines (C) avec l’histidine (H) qui le composent sont colorées en violet. La région en couleur saumon correspond au site de fixation du cofacteur AcétylCoA. La lysine (K) en rouge est la lysine auto-acétylée. Cette dernière, avec les deux résidus cystéine (C) et glutamate (E) colorés en bleu, constituent la dyade catalytique des MYST.

B. MOZ (MOnocytic leukemia Zinc finger)

MOZ est un des membres de la famille des HAT à domaine MYST identifié en 1996 par Borrow *et al.* Cette découverte provient de l'étude de la translocation chromosomique récurrente t(8 ;16)(p11 ;p13) qui est une marque caractéristique de leucémie myéloïde aigue. L'un des deux partenaires de fusion, situé dans le chromosome 8 est *MOZ*, dont la présence et la fonction étaient à l'époque inconnues ¹⁴². Ce n'est qu'en 2001 qu'un groupe canadien a démontré que la protéine MOZ possédait une activité HAT intrinsèque ¹⁵².

1) Les domaines de MOZ

MOZ est une grande protéine de haut poids moléculaire (~250kDa) composée de 2004 résidus. En plus du domaine catalytique, quatre autres domaines ont été identifiés, deux en N-terminal (N-ter) et deux autres en C-terminal (C-ter) par rapport au domaine catalytique MYST. Du N-ter au C-ter sont retrouvés les domaines NEMM (N-terminal part of Enok, MOZ or MORF), DPF (PHD1/ PHD2 – Plant HomeoDomain-linked), ED (un long étirement acide Glutamate -E-, Aspartate -D-) et SM (région riche en Serine -S- et Méthionine -M-). La Figure III-6 représente de manière schématique MOZ et ses domaines et le Tableau III-3 explique leurs fonctions.

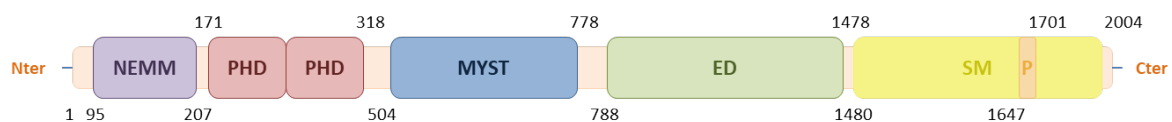


Figure III – 6 : Les cinq domaines de MOZ

NEMM

À ce jour, la fonction du domaine NEMM n'est pas claire. La partie C-ter de ce domaine présente une similarité de séquence avec les histones H1 et H5, c'est pour cela que le C-ter du NEMM est appelé H15. En revanche, le sous-domaine H15 est mieux documenté. Il pourrait être impliqué dans la localisation de l'enzyme au niveau du noyau ¹⁵³.

DPF

Le double domaine PHD en tandem est une caractéristique rare de certaines MYST, dont MOZ. Il s'agit de deux doigts - de - zinc structuraux de type C4HC3. Le PHD2 est assez similaire à d'autres PHD, alors que le PHD1 est plutôt atypique. Ce domaine intervient dans la reconnaissance de l'histone H3, induisant la conformation de cette dernière en hélice α pour faciliter la liaison de MOZ avec l'histone H3 et l'acétylation concomitante des lysines d'H3 (Fig.III-7)^{154,155}.

MYST

Le domaine MYST est le seul domaine de MOZ en commun avec les autres MYST. L'homologie de séquence de ce domaine avec celui des autres MYST a permis initialement de suggérer que MOZ pouvait être une HAT jusqu'en 2001 où cette fonction lui a été bien attribuée ¹⁵². Dans le cas de MOZ, ce domaine est aussi composé d'un doigt de zinc C2HC et d'un site de fixation du cofacteur AcétylCoA.

ED

Cette région acide n'a pas de fonction connue à ce jour. Elle est néanmoins bien documentée comme la région la plus sujette à des mutations non sens et des anomalies génétiques telles que des translocations, produisant ainsi des protéines tronquées ou des chimères. Ces altérations sont retrouvées dans les cas de leucémies et du syndrome KAT6A ¹⁵⁶.

SM

La région riche en sérines et méthionines a été un peu plus caractérisée que la région ED. En 1999, Champagne *et al* ont démontré que le domaine SM de MORF, l'homologue de MOZ, est capable d'activer la transcription. Deux ans plus tard, le même groupe a confirmé cette fonction pour MOZ ^{152,157}.

Tableau III – 3 : Domaines de MOZ et leurs fonctions

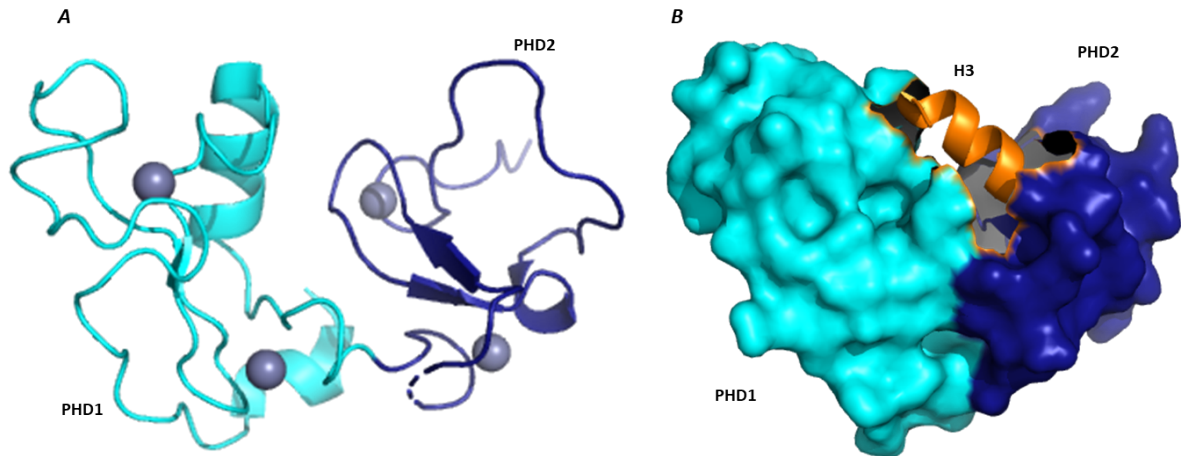


Figure III – 7 : Structure du domaine DPF, caractéristique des MYST

A. Les deux doigts-de-zinc C4HC3 du domaine DPF (PDB : 4LJN)

B. Reconnaissance et la conformation d'H3 en hélice α induite par DPF (PDB : 4LK9)

Dans les deux figures le domaine PHD1 est représenté en cyan et le domaine PHD2 en bleu foncé. L'hélice (orange) dans la figure III-7B correspond à l'hélice de la queue de l'histone H3 formée par le DPF. La cavité des deux PHD est observée dans le cas de la figure III-7B de manière à mieux illustrer la reconnaissance d'H3.

2) La structure du domaine MYST de MOZ

Les deux premières structures du domaine catalytique de MOZ, très similaires entre elles, ont été résolues en 2007 (Fig.III-8A). La comparaison de ces structures avec celles d'Esa1 et hMOF a démontré que la région centrale de MOZ est conservée dans toutes les MYST, alors que les régions N et C terminales sont divergentes. Dans le cas de MOZ, en N-ter il y a le motif doigt de zinc (C2HC) permettant la liaison à l'ADN et en C-ter un faisceau de trois hélices, contenant un motif hélice tour hélice qui est le motif typique de liaison à l'ADN (Fig.III-8B) ¹⁵⁰. Ces deux caractéristiques structurales permettent à MOZ de se lier à la fois aux nucléosomes et aux histones libres, à la différence d'autres HAT comme SAS3.

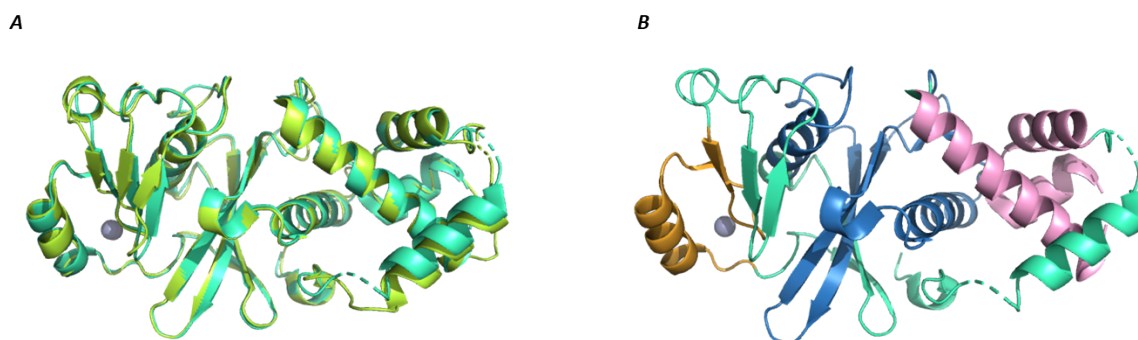


Figure III – 8 : Architecture du domaine MYST de MOZ

A. Les deux structures du domaine MYST de MOZ (vert-cyan, PDB : 2OZU et lime, PDB : 2RC4)

B. Les sous-domaines du domaine MYST de MOZ

Les deux sous-domaines qui permettent la liaison à l'ADN sont représentés en orange et en rose et correspondent respectivement au doigt-de_zinc C2HC et au faisceau de trois hélices. En bleu, il y a la région centrale de MYST qui est conservée dans toutes les protéines de cette famille et qui contient les résidus importants pour la catalyse enzymatique.

3) La dyade catalytique de MOZ

Le mécanisme catalytique de MOZ est déduit par homologie avec les autres HAT de la famille des MYST, dont ESA1 pour laquelle ce mécanisme a été détaillé ¹⁴⁶. Il s'agit d'un mécanisme de type ping-pong avec deux résidus catalytiques qui, dans le cas de MOZ, sont la cystéine 646 (C646) et le glutamate 680 (E680). Cette dyade est située proche du site de fixation de l'AcétylCoA (Fig.III-9A), où ce cofacteur s'insère dans une poche formée à la surface protéique (Fig.III-9B). En ce qui concerne la lysine du substrat, elle se fixe parfaitement dans l'espace créé entre ceux deux résidus catalytiques et de manière parallèle (Fig.III-9C). Cette architecture de l'enzyme permet d'avoir tous les acteurs de la catalyse à proximité de manière à faciliter la réaction enzymatique (Fig.III-9D).

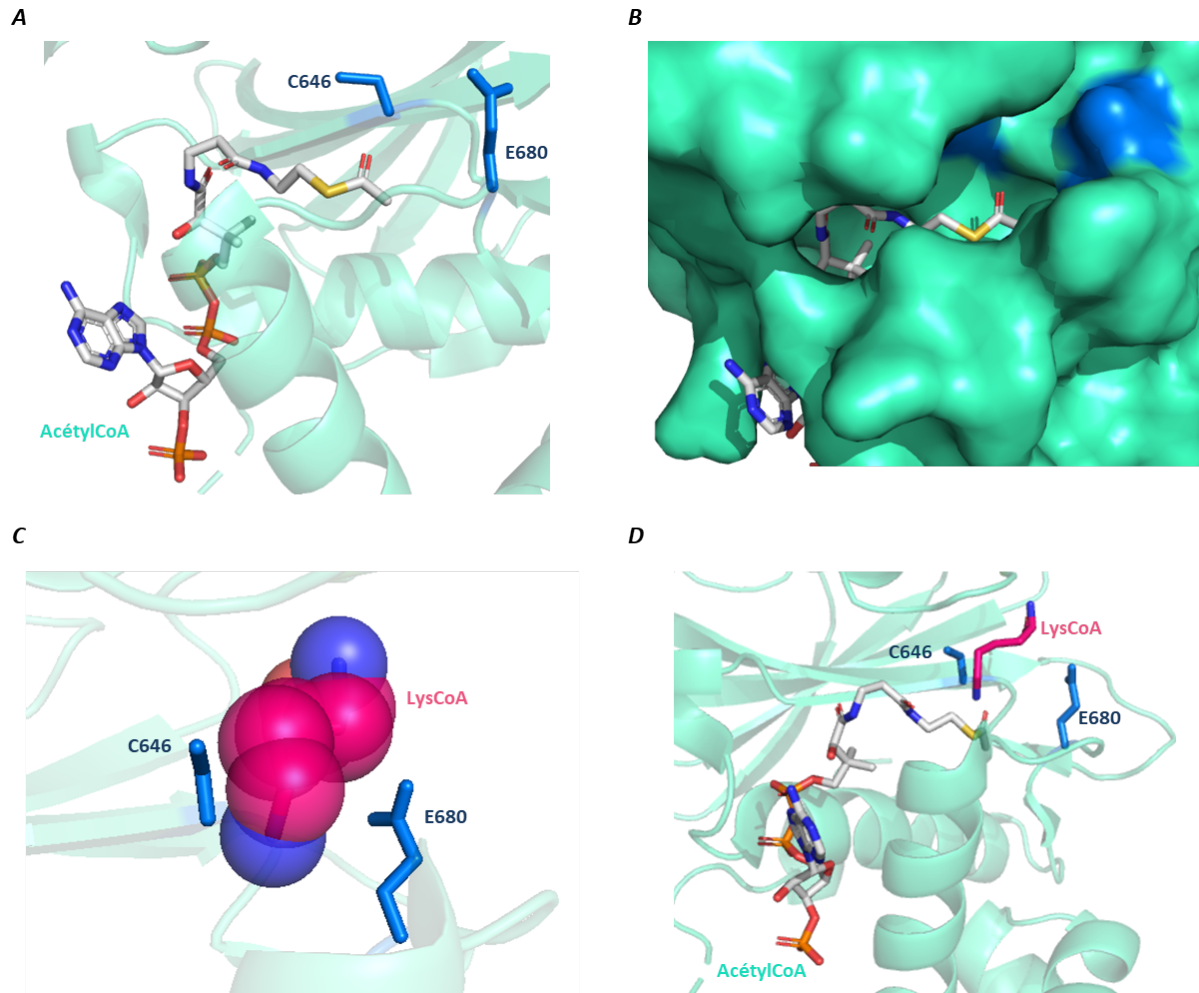


Figure III - 9 : Architecture catalytique de MOZ

A. La dyade catalytique de MOZ (bleu, PDB : 2OZU) et le site de fixation de l'AcétylCoA

B. La poche de fixation de l'AcétylCoA à la surface protéique (vert-cyan)

C. La liaison du substrat lysine (LysCoA, magenta) entre les deux résidus catalytiques

D. L'architecture catalytique de MOZ

Le LysCoA représenté dans ces figures est issu de la structure d'ESA1 co-cristallisée avec l' H4K16-CoA (PDB : 3TO6). Les deux structures (2OZU et 3TO6) ont été alignées pour cette modélisation.

4) L'auto-acétylation de MOZ

Comme il a été décrit précédemment, le domaine MYST doit dans un premier temps s'acétyler lui-même de manière à être capable d'acétyler à la suite ses substrats. L'auto-acétylation de MOZ a été décrite en 2007, avec la résolution de la structure du domaine catalytique de MOZ (PDB : 2OZU) par Wu *et al* (non publié). Cette acétylation s'effectue en cis, ce qui signifie que chaque molécule est capable de s'auto-acétyler intrinsèquement. Dans le cas de MOZ, la lysine correspondante à la lysine d'auto-acétylation est la 604 (K604). Elle se situe dans le site actif, en face de la dyade catalytique C646-E680.

L'auto-acétylation est aujourd'hui mieux décrite dans le cas de hMOF, qui est un homologue de MOZ et permet de comprendre par analogie le mécanisme d'auto-acétylation de cette dernière (Fig.III-10A). Yuan *et al* ont résolu en 2012 la structure de hMOF qui démontre que la boucle sur laquelle se situe la lysine auto-acétylée (K274) adopte deux conformations (Fig.III-10B). Dans un cas, la lysine K274 est acétylée et se dirige en dehors du site actif, alors que dans l'autre cas elle reste non-acétylée et retournée de 90° vers le site actif, formant une liaison hydrogène avec le glutamate catalytique (E350) et bloquant ainsi l'accès au site actif de la lysine substrat (Fig.III-10C).

De manière à stabiliser la conformation ouverte et l'accessibilité du site actif, la lysine auto-acétylée forme des interactions hydrophobes avec les résidus de son environnement. Les phénylalanines 283 et 285 (F283, F285) forment des liaisons de Van der Waals avec le groupement méthyle de la lysine K274 acétylée, alors que la tyrosine 301 (Y301) et la sérine 303 (S303) interagissent par des liaisons hydrogène avec la partie carbonyle de la lysine (Fig.III-10D). Ces résidus sont strictement conservés dans toutes les MYST, suggérant un mécanisme similaire¹⁴⁷.

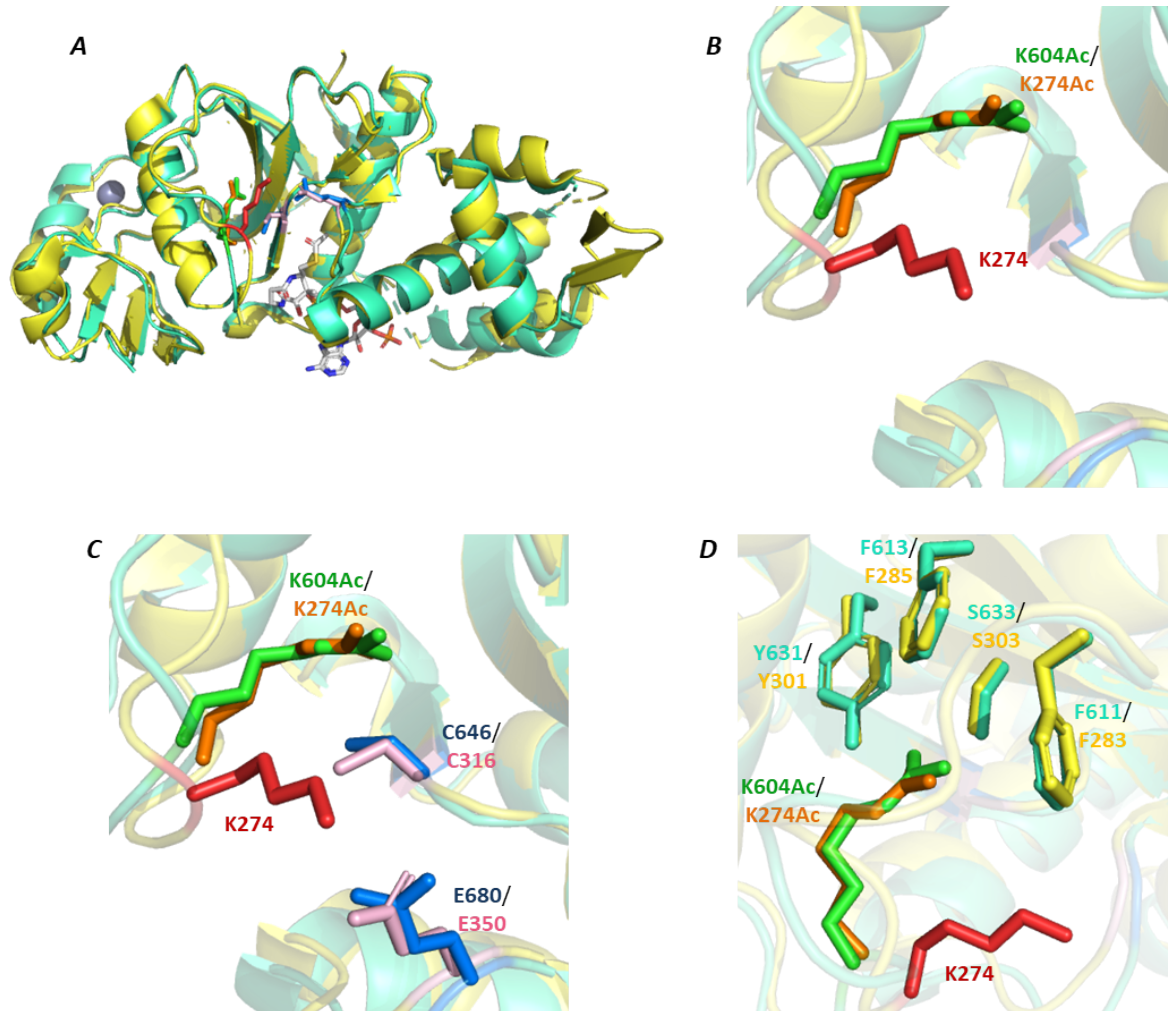


Figure III - 10 : Auto-acétylation de MOZ

A. Le domaine MYST des MOZ (vert-cyan, PDB : 2OZU) et hMOF (jaune, PDB : 3TOA)

B. Les lysines auto-acétylées des MOZ (vert) et hMOF (orange) et la lysine non acétylée de hMOF (rouge)

C. Les sites actifs des MOZ (bleu) et hMOF (rose) avec les lysines 'catalytiques'

D. L'environnement interagissant avec la lysine catalytique (vert-cyan, MOZ et jaune hMOF)

5) Les rôles de MOZ

MOZ est un co-activateur transcriptionnel lié à la prolifération et à la différenciation des cellules souches hématopoïétiques (CSH). Elle interagit principalement par son domaine SM avec des facteurs de transcription comme AML1 (Acute Myeloid Leukemia protein 1), RUNX2 (Runt domain transcription factor 2) et PU.1 qui sont responsables de la génération des CSH et de la différenciation des lymphocytes T, B en macrophages^{153,158,159}. Cette interaction participe à la régulation de l'expression de plusieurs gènes hématopoïétiques d'intérêt, notamment MIP-1 α (Microphage Inflammatory Protein 1 α), un inhibiteur de prolifération des CSH présentes dans les cellules hématopoïétiques¹⁶⁰.

L'expression de *MOZ* n'est pas spécifique à un tissu. Selon l'organe où elle s'exprime elle participe à différents processus liés au développement. L'implication précise de MOZ dans ces mécanismes reste inconnue. Néanmoins, elle a été identifiée comme acteur dans la neurogenèse, la skeletogenèse et le développement craniofacial¹⁶¹. De ce fait, la double délétion chromosomique du gène *Moz* est létale pour les embryons de souris contrairement à une délétion sur l'un des deux chromosomes.

MOZ est également impliquée dans la régulation du cycle cellulaire, et plus particulièrement dans l'arrêt de ce dernier et dans l'induction ou l'inhibition de la sénescence selon le mécanisme activé. Lors d'un stress cellulaire MOZ, acétyle p53 aux niveaux des lysines 120 et 382 induisant l'expression du gène *p21* et l'arrêt du cycle cellulaire en phase G1^{162,163}. En l'absence de dommages à l'ADN, MOZ inhibe la sénescence cellulaire par activation de la voie INK4/ARF et l'expression des régulateurs du cycle cellulaire p16 et p19, codés par *CDKN2A* (Cyclin Dependant Kinase inhibitor 2A)^{164,165}. Les résultats contradictoires de ces études suggèrent que selon le contexte cellulaire, MOZ peut agir sur des voies différentes, conduisant à l'activation de différents processus cellulaires (Fig.III-11).

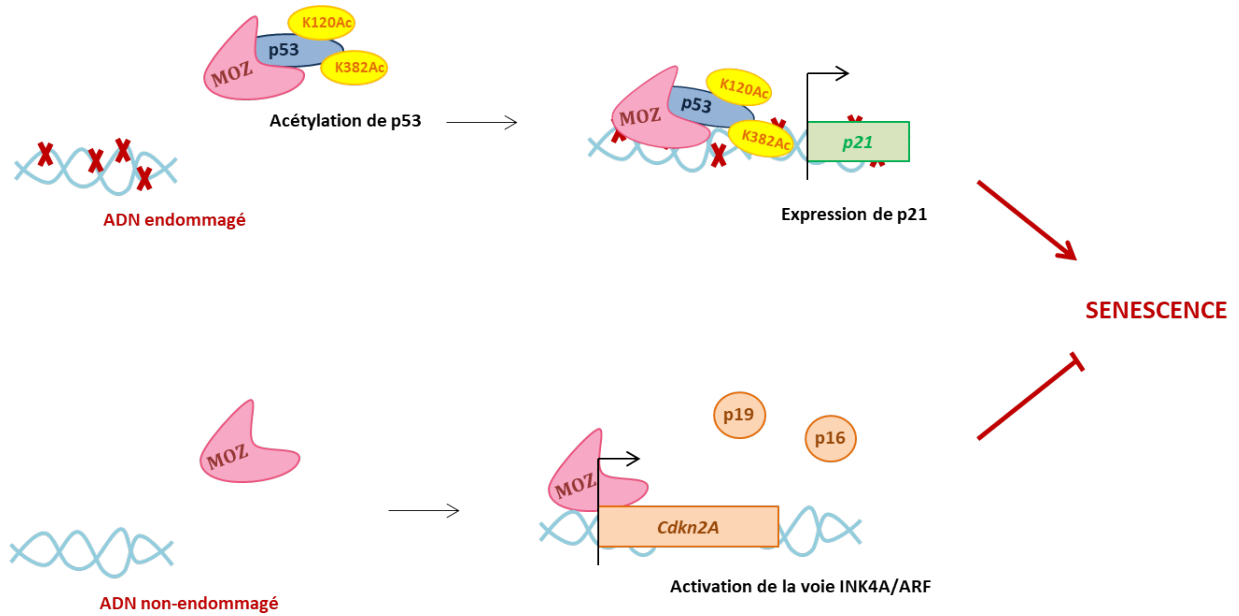


Figure III – 11 : Représentation simplifiée des mécanismes d’activation/inhibition de senescence par MOZ

C. MOZ et pathologies

Des altérations du gène *MOZ* sont associées à des altérations développementales. Des mutations de ce gène sont décrites aujourd’hui dans un syndrome rare, appelé ‘Syndrome KAT6A’ mais aussi dans différents cancers et plus particulièrement le cancer du sein et les glioblastomes. En 1996, *MOZ* a été identifiée comme étant le deuxième partenaire du gène de fusion *MOZ-CBP* issu de la translocation chromosomique t(8 ;16)(p11 ;p13) qui fusionne les deux histones acétyltransférases *CBP* et *MOZ*. Depuis, plusieurs translocations chromosomiques impliquant le gène *MOZ* ont été retrouvées dans le génome de patients atteints de cancers, notamment des leucémies ¹⁶¹.

1) Le Syndrome KAT6A

Le syndrome KAT6A est une maladie neurodéveloppementale rare qui a été identifiée pour la première fois en 2015 suite au séquençage du génome des enfants présentant un retard développemental du cerveau et du système nerveux. Ce syndrome est lié à des mutations hétérozygotes qui sont majoritairement *de novo*, dominantes et non sens, amenant à la synthèse des protéines tronquées au niveau du domaine acide. De même, des mutations faux sens ont été également documentées dans le cas de résidus conservés dans le domaine MYST ou SM (Fig.III-12). Par des analyses génomiques, ces variants ont pu être identifiés pour la première fois dans le cas des enfants KAT6A. Des variants de KAT6B sont également identifiés. Le syndrome est appelé aujourd’hui ‘Syndrome KAT6’ et à ce jour 334 cas sont documentés ^{166–168}.

Tous les enfants qui ont le syndrome KAT6A présentent une invalidité neurologique et intellectuelle, avec un retard du développement. Dans la grande majorité il a été constaté un retard de langage, des difficultés d'alimentation, de l'hypotonie et des défauts de vision. Plus récemment, il a été décrit des cas présentant des crises épileptiques, du strabisme et des anomalies cardiaques. Depuis la découverte de ce syndrome, les cas enregistrés augmentent graduellement ^{169–171}. En 2017, la fondation KAT6A a été créée par des parents d'enfants malades, de manière à supporter la recherche sur ce syndrome qui a pour but de développer des traitements qui permettront un meilleur niveau de vie pour ces enfants.

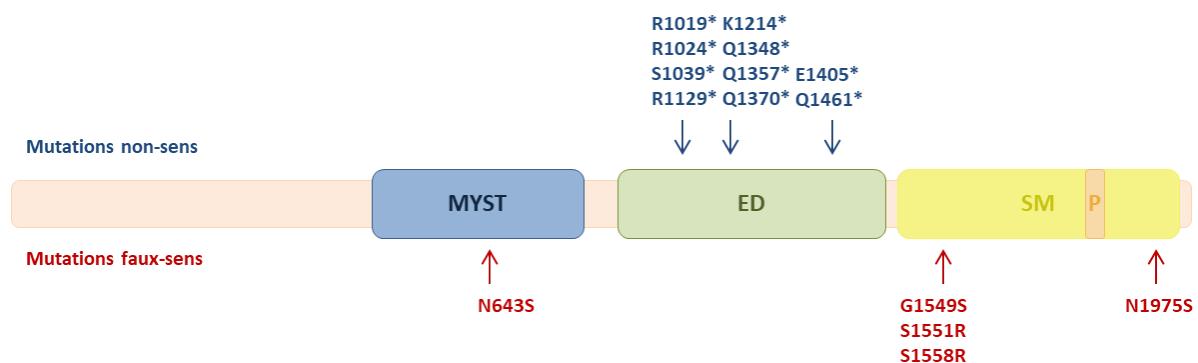


Figure III – 12 : Représentation simplifiée des mutations non sens (bleu) et faux sens (rouge) liées au syndrome KAT6A

2) Les mutations de MOZ dans les cancers

D'après les données enregistrées dans les banques de génomique des cancers (cBioPortal et COSMIC), MOZ a une faible fréquence de mutations somatiques qui est égale à 0,9%. Étant une protéine qui n'a pas de spécificité d'expression tissulaire, la distribution des mutations enregistrées est variable et inclut plusieurs tissus (Tableau III-4). L'œil est l'organe avec le pourcentage de mutations le plus élevé, suivi principalement par des organes de l'appareil digestif et du tractus gastro-intestinal, comme l'estomac, le foie et le gros intestin. Parmi ces mutations, les plus représentées sont les mutations faux sens (45%). D'autres mutations sont décrites comme les mutations synonymes (17%) et les mutations non sens (3%).

Au total, 1448 mutations sont enregistrées dans le site de COSMIC (Catalogue Of Somatic Mutations In Cancer). Parmi eux, 498 se trouvent dans le domaine acide (ED) qui est décrit comme étant la région la plus susceptible à être altérée. Toujours dans la partie C-ter de MOZ, la région SM est souvent mutée avec 231 cas. 109 mutations sont identifiées dans le domaine catalytique MYST, conservé au sens de toutes les MYST porte 109 mutations. La partie N-ter est la moins altérée avec 48 mutations dans le domaine DPF et 24 mutations dans le domaine NEMM. La figure III-13 représente les principales mutations identifiées dans chacun de ces domaines (Fig.III-13).

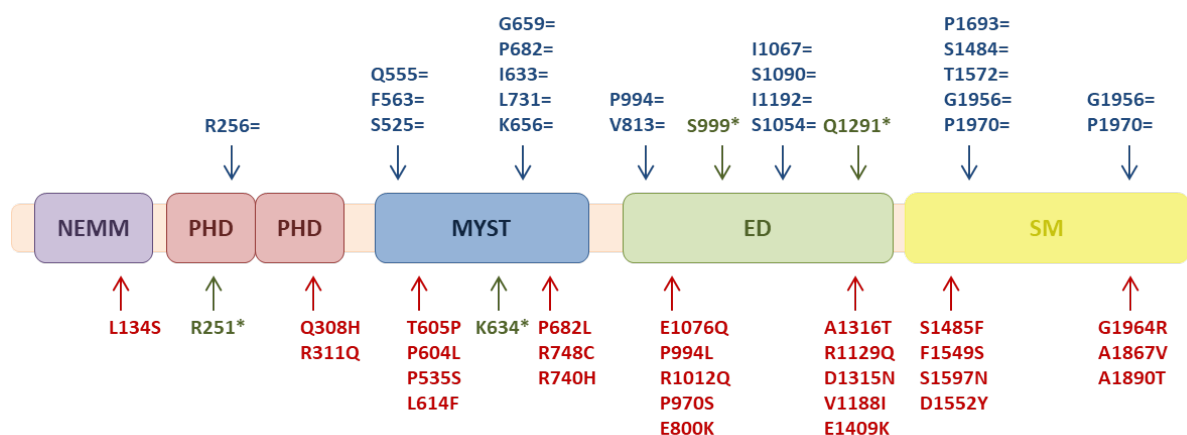


Figure III – 13 : Principales mutations de chaque domaine, enregistrées dans la banque de COSMIC

En rouge figurent les mutations faux sens, en bleu les mutations synonymes et en vert les mutations non sens.

Tissu	Testés	Mutés	% Mutés
Œil	123	65	52.8
Tractus gastro-intestinal	39	4	10.3
Peau	1877	164	8.7
Méninges	168	12	7.1
Estomac	1455	101	6.9
Foie	2303	149	6.5
Placenta	31	2	6.5
Gros intestin	3154	191	6.1
Voie biliaire	886	50	5.6
Appareil urinaire	810	40	4.9
Tissu mou	636	31	4.9
Sein	2947	132	4.5
Poumon	2988	123	4.1
Prostate	1990	80	4.0
Ovaire	1023	41	4.0
Col de l'utérus	334	13	3.9
Glande salivaire	110	4	3.6
Œsophage	1597	50	3.1
Pancréas	2073	59	2.8
Voies aérodigestives supérieures	1411	34	2.4
Rein	2564	43	1.7
Système nerveux central	2749	36	1.3
Hématopoïétique et lymphoïde	5207	62	1.2
Intestin grêle	251	3	1.2
Hypophyse	83	1	1.2
Testicule	229	2	0.9
Thyroïde	1669	11	0.7
Plèvre	269	1	0.4
Os	573	2	0.3
Gland surrénale	624	1	0.2

*Tableau III – 4 :
Distribution des
mutations du
gène MOZ,
enregistrées
dans la banque
des données des
mutations
somatiques aux
cancers
(COSMIC)*

Les données
sont
représentées
dans l'ordre
décroissant du
pourcentage
des Mutés.

Malgré de nombreuses mutations présentes sur le gène *MOZ*, à ce jour il n’y a pas d’étude sur la caractérisation de ces mutations, ainsi que sur leur effet enzymatique. Les mutations détectées sont issues directement des analyses génomiques des patients et nécessitent une évaluation future afin de mieux comprendre la conséquence de ces mutations. Il est important de noter que les mutations de *MOZ* présentes dans le cas des patients atteints de cancer sont différentes de celles identifiées dans le cas des patients avec le syndrome KAT6A, supposant ainsi un mécanisme et un effet différent entre les deux types des mutations.

3) Expression de MOZ dans les cancers

L’analyse des données enregistrées dans les banques de données du génome des patients de cancer, TCGA (The Cancer Genome Atlas project) et d’expression du génome selon le tissu, GTEx (The Genotype-Tissue Expression project), a démontré qu’il y a une variation de l’expression de *MOZ* dans les tissus tumoraux par rapport aux tissus normaux. Dans la majorité des cancers, l’expression de *MOZ* augmente, comme cela a été décrit pour la leucémie myéloïde aiguë (LAML), le glioblastome (GBM) et le carcinome d’œsophage (ESCA). Dans d’autres types de cancers, comme l’adénocarcinome des poumons (LUAD) et le carcinosacrome (UCS) son expression diminue (Fig.III-14A) ¹⁶¹.

Par ailleurs, l’expression de *MOZ* dans certains types de cancers, comme les cancers du sein à récepteur œstrogènes positifs (ER α +), corrèle fortement avec un mauvais pronostic. Plus précisément, il a été démontré que la diminution de l’expression de *MOZ* dans les cellules de cancer du sein, inhibe la croissance tumorale, suggérant ainsi que *MOZ* peut être impliquée dans la régulation de la prolifération cellulaire. Cela est dû à l’activation de l’expression d’ER α , un des marqueurs les plus importants du cancer du sein, identifiant ainsi *MOZ* comme la première acétyltransférase capable de réguler l’expression d’ER α (Fig.III-14B) ¹⁷².

De même, dans le cas des glioblastomes, l'expression de MOZ est particulièrement élevée, évoquant la corrélation du niveau d'expression de MOZ avec le grade du gliome, mais aussi avec la diminution du temps de survie après le diagnostic des patients. MOZ acétyle la lysine 23 de l'histone H3 (H3K23), générant une marque épigénétique non canonique (H3K23Ac), qui induit ainsi la liaison du facteur intermédiaire de transcription TRIM24 (TRIM24 Motif containing 24), permettant l'activation de la voie PI3K/AKT et de la tumorigenèse avec, pour conséquence, la prolifération accrue des cellules tumorales. L'inhibition de l'activité de MOZ permet d'arrêter la croissance de ces tumeurs avec un pronostic défavorable (Fig.III-14C) ¹⁷³.

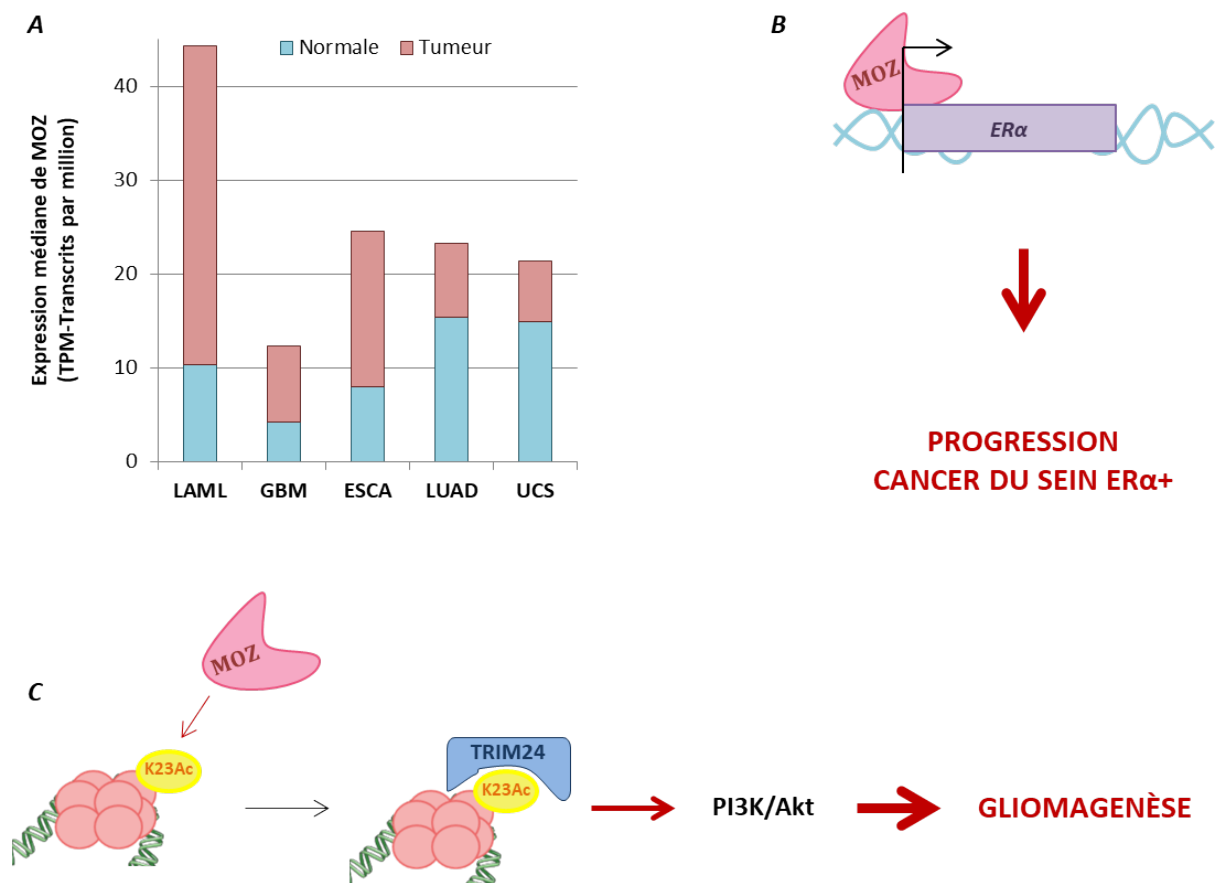


Figure III – 14 : Dérégulation de l'expression de MOZ dans les cancers

A. Expression de MOZ dans différents types de cancer par rapport aux tissus normaux

B. L'implication de MOZ au développement des cancers du sein ERα+

C. L'effet de surexpression de MOZ sur la progression des gliomes

4) Les translocations chromosomiques de *MOZ*

L'étude de la translocation chromosomique t(8;16)(p11;p13) a permis l'identification du gène *MOZ* par sa fusion avec *Cbp* et la formation de la protéine chimère MOZ-CBP¹⁴². Cette translocation récurrente est une des caractéristiques du sous type M4/M5 de leucémie myéloïde aiguë (AML) avec des pronostics défavorables et est associée avec un blocage de la différenciation monocyttaire et l'érythrophagocytose. La chimère MOZ-CBP résulte de l'association des 1547 premiers résidus de MOZ qui sont alliés avec CBP à partir du résidu 266. De manière intéressante, lors de la fusion de ces deux histones acétyltransférases, il y a conservation des deux domaines catalytiques (MYST de MOZ et HAT de CBP) entraînant une modification aberrante de la chromatine¹⁴². D'autres études ont décrit des points de rupture différents, produisant ainsi des chimères de taille et de composition variables^{174,175}.

La formation de cette chimère est responsable de la dérégulation de certains processus cellulaires. Comme il a été décrit précédemment, MOZ interagit par son domaine SM avec le facteur de transcription AML1 et conduit à la différenciation des M1 (cellules sanguines) en monocytes et macrophages. En revanche, la chimère MOZ-CBP qui ne possède pas de domaine SM inhibe la transcription dépendante d'AML1. Cela suggère que MOZ-CBP inhibe la différenciation des M1 qui continuent de proliférer de manière exponentielle¹⁵³. La protéine chimère MOZ-CBP dérégule d'autres mécanismes cellulaires comme la transcription par le facteur de transcription nucléaire NF-κB, dont l'effet de la chimère MOZ-CBP est opposé à celui de MOZ-CBP sur l'AML1. En effet, la capacité activatrice de NF-κB par MOZ est fortement augmentée lorsque MOZ est sous forme de chimère¹⁷⁶.

Le gène *MOZ* se situe dans une région du chromosome 8 susceptible d'être impliquée dans des translocations chromosomiques et qui est liée à la leucémogénèse. Ainsi, de nouveaux partenaires de fusion de MOZ ont été caractérisés. Certains de ces nouveaux partenaires de fusion sont le paralogue de CBP, p300, et le 'Transcriptional Intermediaire Factor 2' (TIF2). La chimère MOZ-p300 est issue de la translocation t(8;16)(p11;q13) avec des effets similaires à ceux de MOZ-CBP¹⁷⁷. MOZ-TIF2 est issue de l'inversion inv(8)(p11;q13) et est probablement responsable de la délocalisation de TIF2 et ses partenaires d'une part et d'autre part de l'augmentation de l'activité de MOZ due au recrutement de CBP par

TIF2^{178,179}. A ce jour, il y a de plus en plus de partenaires de MOZ décrits dans des cas de cancers et, plus particulièrement, dans les leucémies. Le Tableau III-5 récapitule les partenaires décrits dans la littérature, ainsi que la jonction responsable de la formation de la protéine chimère.

Partenaire	Protéine Chimère	Jonction	Type de Cancer
CBP	MOZ-CBP	t(8 ;16)(p11 ;p13)	AML M4/M5
p300	MOZ-p300	t(8 ;16)(p11 ;q13)	AML M4/M5
TIF2	MOZ-TIF2	inv(8)(p11 ;q13)	AML M4/M5
NCOA3	MOZ-NCOA3	t(8 ;20)(p11 ;q13)	AML M5
LEUTX	LEUTX-MOZ	t(8 ;19)(p11 ;q13)	AML M4/M5
ASXL2	ASXL2-MOZ	t(2 ;8)(p23 ;q11)	MDS (Syndrome myélodysplasique)

Tableau III - 5 : Les principales translocations chromosomiques de MOZ^{142,174,177,179-181}

OBJECTIFS GÉNÉRAUX

Ce travail de thèse a eu comme objectif de poser les bases moléculaires et structurales de l'impact de mutations identifiées dans des cancers sur des enzymes d'importance épigénétique, l'histone méthyltransférase SETD2 et l'histone acétyltransférase MOZ.

SETD2 est une protéine suppresseur de tumeur, décrite comme la principale histone méthyltransférase responsable de la triméthylation de la lysine 36 de l'histone H3 (H3K36). L'absence d'activité de SETD2 est liée avec l'initiation et la progression des processus cancéreux. Des mutations du gène de *Setd2* sont identifiées de manière récurrente dans une variété de cancers. Cependant la caractérisation de ces mutations reste peu documentée.

MOZ est une histone acétyltransférase de la famille des histones acétyltransférases à domaine MYST. Sa fonction d'acétylation de l'histone H3 et de p53 est impliquée dans des processus développementaux et l'hématopoïèse. Par ailleurs, des altérations génétiques de MOZ telles que des translocations génomiques et/ou des mutations sont observées dans différents cancers. À ce jour, l'impact moléculaire et fonctionnel de ces altérations reste peu documenté.

Mon travail de thèse se divise en deux axes. L'axe majeur porte sur la caractérisation d'une mutation oncogénique de SETD2. En parallèle, je me suis intéressée à la caractérisation d'une mutation oncogénique de MOZ.

- 1) Nous nous sommes intéressés à l'étude de la mutation SETD2 L1609P, identifiée chez des patients pédiatriques atteints de leucémie. La position de cette mutation dans le domaine SET-I de l'enzyme, décrit comme impliqué dans la fixation du substrat protéique de l'enzyme, nous a permis de poser l'hypothèse d'une dérégulation de l'activité enzymatique. Nous avons utilisé des approches d'enzymologie, de biochimie moléculaire et biologie cellulaire, ainsi que de cristallographie, afin de caractériser au niveau fonctionnel et structural ce mutant de SETD2.
- 2) En parallèle, nous avons étudié la mutation MOZ T605P, située dans le site catalytique de l'enzyme, en juxtaposition avec la lysine d'auto-acétylation (K604) critique pour l'activité de l'enzyme. Nous avons procédé à une analyse biochimique et structurale de cette mutation, afin de mieux comprendre son

impacte. Ces travaux permettent également de mieux décrire la machinerie catalytique de MOZ.

**CARACTERISATION STRUCTURALE
ET FONCTIONNELLE D'UN
MUTANT ONCOGENIQUE DE SETD2**

I. Introduction

SETD2 est décrite comme un suppresseur de tumeur parmi les plus fréquemment mutés dans les cancers. Dalglish *et al* ont décrit pour la première fois en 2010 des mutations inactivatrices de SETD2 identifiées chez des patients atteints de carcinomes rénaux à cellules claires (ccRCC) ¹⁰⁷. Dans le cas des ccRCC, *Setd2* est classé parmi les 5 gènes le plus mutés et est impliqué dans l'initiation et la progression tumorale via des mutations mono- ou bi- alléliques ^{108,109}. Depuis les travaux de Dalglish *et al*, des études de séquençage d'exomes entiers ont associé des mutations somatiques de *Setd2* avec une pléthore de cancers, dont des hémopathies et des glioblastomes (données issues de la base de données COSMIC). Dans le cas des hémopathies malignes, SETD2 est également impliquée dans des leucémies de type AML et ALL ^{27,114}. De plus, plusieurs études mettent aujourd'hui en évidence des liens entre des mutations de SETD2 et un phénotype chimiorésistant ^{115,182,183}.

L'importance de SETD2 est fortement liée à sa fonction majeure d'histone méthyltransférase du résidu H3K36, résultant dans la formation de la marque épigénétique H3K36me3 ¹⁴. Cette marque est impliquée dans plusieurs processus cellulaires dont l'élongation de la transcription et la réparation de l'ADN. Par ailleurs, de manière intéressante, la majorité des mutations identifiées de SETD2 sont centrées autour des régions AWS – SET – Post-SET, qui constituent le domaine catalytique de l'enzyme. Malgré le nombre de mutations documentées, les données concernant la caractérisation structurale et fonctionnelle de celles-ci restent parcellaires. À ce jour, seuls les travaux de Hacker *et al* en 2016 ont permis d'étudier une mutation oncogénique récurrente, d'un résidu d'importance majeure ⁸⁴. Leur travail biochimique sur la caractérisation de la mutation R1625C a permis de mettre en évidence le caractère de 'perte d'activité' de cette mutation.

Dans un contexte similaire, nous nous sommes intéressés à l'étude d'une autre mutation située dans le domaine SET de SETD2, la mutation L1609P. Cette mutation a été identifiée chez des patients pédiatriques atteints de leucémie lymphoblastique aiguë, grâce aux travaux de Mar *et al* (Fig. S-1) ¹¹⁴. Dans leur étude, les auteurs ont identifié un

enrichissement des mutations dans le cas des patients pédiatriques en rechute d'ALL, et 12 % des patients de la cohorte totale étudiée ont présenté des mutations dans le gène de *Setd2*.

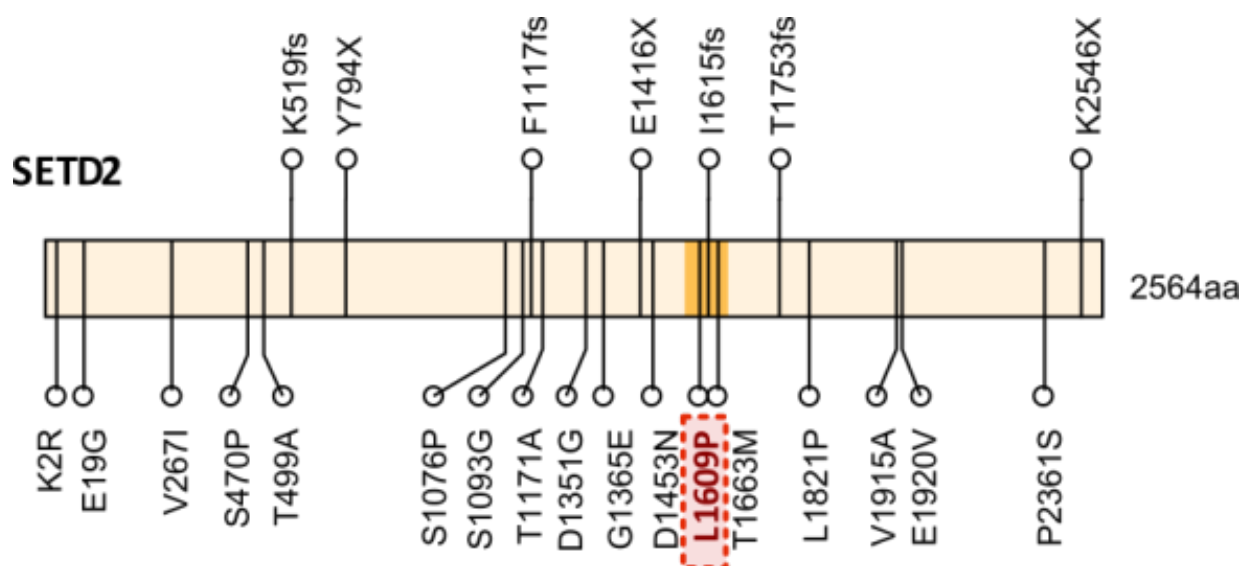


Figure S-1 : Mutations de SETD2 identifiées chez des patients atteints d'ALL

Schéma modifié, à partir de *Mar et al*¹¹⁴

Le domaine catalytique est représenté en jaune. 2 mutations faux sens sont identifiées dans ce domaine (L1609P et T1663M). Nous nous sommes intéressés à la mutation L1609P proche du site actif de l'enzyme et des résidus importants pour la catalyse (résidus 1604-1608).

Notre travail a consisté à caractériser au niveau structural et fonctionnel la mutation L1609P. Le résidu L1609 est présent dans le domaine catalytique de l'enzyme (sous-domaine SET) proche d'un motif impliqué dans la catalyse (résidus 1604-1608). Nous avons utilisé des méthodes de biochimie cellulaire et moléculaire afin de comprendre l'effet engendré par la mutation sur l'activité et la stabilité du mutant SETD2 L1609P. Grâce à nos études structurales utilisant des approches de cristallographie, nous avons également obtenu la structure 3D du mutant SETD2 L1609P. Ce résultat nous a permis de mieux comprendre l'effet de la mutation sur l'activité et la stabilité de l'enzyme.

II. Résultats

1) Activité et stabilité du mutant SETD2 L1609P *in vitro*

L'alignement multiple des séquences du domaine SET des principales histones méthyltransférases humaines montre que le résidu L1609 est fortement conservé (Fig. S-2A) et par ailleurs l'alignement suggère la nécessité d'un acide aminé hydrophobe et large à cette position, qui se trouve en position C-ter d'une structure secondaire en brin β (brin β 5 du domaine SET). Cette structure secondaire proche du site de fixation du substrat protéique de SETD2 (notamment H3), comme démontré par la structure cristallographique de SETD2 obtenue avec un substrat histone (PDB 5JJY).

Nous avons exprimé et purifié le domaine catalytique (comportant les régions AWS – SET – Post-SET) de SETD2 dans sa forme sauvage (WT) et mutante (L1609P). En bactérie *E.Coli* le mutant L1609P s'exprime aussi bien que le WT, et les deux enzymes obtenues sont pures et solubles. De manière intéressante, nous observons un décalage de migration du mutant SETD2 L1609P vers les hauts poids moléculaires (Fig. S-2B). Or, dans le cas d'une mutation leucine en proline, le résidu proline possède un plus faible poids moléculaire (115.13 Da) que le résidu leucine (131.17 Da), suggérant ainsi une différence conformationnelle visible en SDS-PAGE entre les deux formes de l'enzyme.

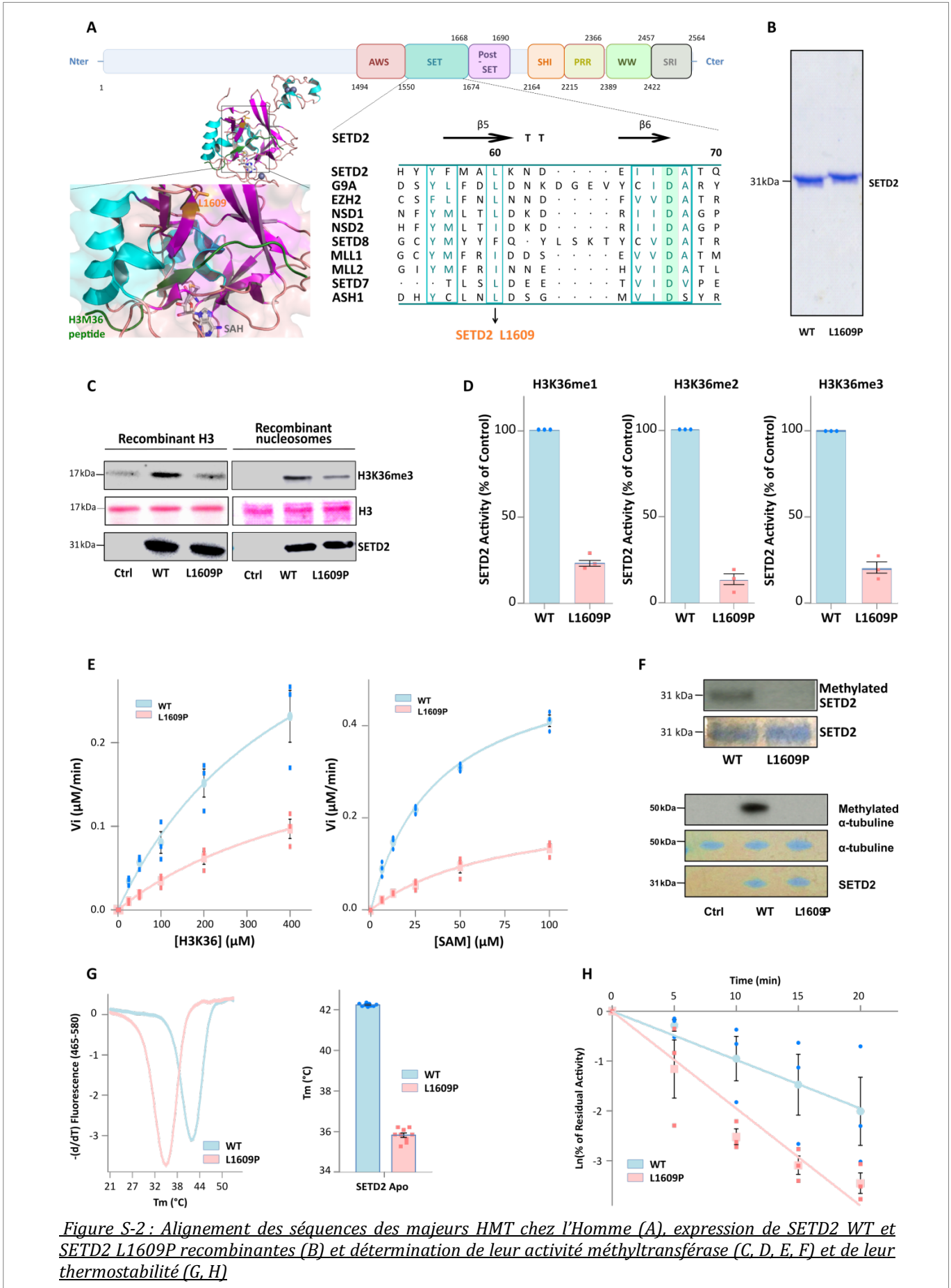
Afin de comprendre l'effet de la mutation L1609P sur l'activité histone méthyltransférase de l'enzyme, nous avons procédé à des essais de dosages enzymatiques. Nous avons incubé SETD2 en présence de son cofacteur SAM et de l'histone H3 (purifiée ou au sein de nucléosomes, Fig. S-2C). Nous montrons qu'en effet la mutation SETD2 L1609P altère l'activité histone méthyltransférase de l'enzyme, avec une baisse significative de la marque H3K36me3 qui est décrite comme spécifique de SETD2.

Par ailleurs, Sun *et al* ont montré que SETD2 est capable de catalyser les trois étapes de méthylation du résidu H3K36 (mono-, di- et tri- méthylation) *in vitro* ⁷⁶. Afin de mieux

comprendre l'altération de l'activité de SETD2 issue de la mutation L1609P, nous avons voulu déterminer si la mutation abolit l'activité de méthyltransférase de SETD2 sur toutes les étapes de méthylation ou s'il s'agit uniquement d'un défaut de son activité de triméthylation. Pour ce faire, nous avons procédé à des dosages d'activité de SETD2 WT et SETD2 L1609P sur des substrats peptidiques couplés à la fluoresceine (H3K36, H3K36me1, H3K36me2)¹⁸⁴. Les mélanges enzymatiques sont ensuite séparés et quantifiés par HPLC (Fig. S-2D). Les résultats montrent que le mutant SETD2 L1609P est un mutant 'perte d'activité' pour les trois étapes de méthylation du résidu H3K36, avec une activité résiduelle moyenne de ~ 25 %. En outre, l'efficacité catalytique (kcat/Km) du mutant est significativement plus faible que celle de la forme WT, avec une efficacité catalytique de $2 \cdot 10^{-4} \mu\text{M}^{-1} \cdot \text{min}^{-1}$ pour SETD2 WT et de $8 \cdot 10^{-5} \mu\text{M}^{-1} \cdot \text{min}^{-1}$ pour SETD2 L1609P (Fig. S-2E) pour le substrat peptidique H3K36 et $2,5 \cdot 10^{-4} \mu\text{M}^{-1} \cdot \text{min}^{-1}$ pour SETD2 WT et de $5 \cdot 10^{-4} \mu\text{M}^{-1} \cdot \text{min}^{-1}$ pour SETD2 L1609P (Fig. S-2E) pour le cofacteur SAM.

Par ailleurs, SETD2 est plus largement décrite comme une lysine méthyltransférase, capable de méthyler des substrats non histone, comme l' α -tubuline⁹⁹, mais également capable de s'auto-méthyler⁷⁶. Par des expériences de dosages radioactifs (cofacteur SAM tritiée), nous avons comparé l'activité lysine méthyltransférase de SETD2 WT et SETD2 L1609P vis-à-vis de l' α -tubuline, ainsi que son activité d'auto-méthylation. Comme observé précédemment pour son activité histone méthyltransférase, l'activité lysine méthyltransférase du mutant SETD2 L1609P sur l' α -tubuline est fortement impactée. De même, l'auto-méthylation de SETD2 L1609P est fortement perturbée (Fig. S-2F).

L'ensemble de ces résultats indiquent que la mutation SETD2 L1609P entraîne la perte significative de l'activité de l'enzyme. SETD2 étant décrite comme une protéine intrinsèquement instable⁷⁹, nous avons également procédé à la caractérisation de la stabilité (notamment thermostabilité) du mutant L1609P. Par des expériences de décallage thermique (TSA assay), nous avons identifié la température de fusion (T_M) des deux formes de l'enzyme (Fig. S-2G). De manière intéressante, le mutant SETD2 L1609P est beaucoup moins stable que la forme sauvage (T_M SETD2 WT = 42,3 °C, T_M SETD2 L1609P = 35,8 °C). Nous avons également déterminé ce paramètre en présence du cofacteur SAM et d'un substrat peptidique H3M36 dérivé de la séquence de l'histone H3 (peptide H3K36M). Ces



A : Alignement des séquences du domaine SET des membres principaux des HMT à domaine SET chez l'Homme. L'alignement a été fait par le serveur Clustal Ω , suivi d'une analyse par le logiciel ESPript 3¹⁸⁵.

B : Séparation des protéines SETD2 sauvage (WT) et SETD2 L1609P recombinantes, purifiées par chromatographie d'affinité (Ni-NTA). 1 μ g de protéine a été séparé par SDS-PAGE, suivi de coloration au bleu de Coomassie R-250.

C : Dosage de l'activité d'histone méthyltransférase de SETD2. 5 μ M de SETD2 ont été incubés en présence de 1 μ g de substrat H3 (recombinant ou nucléosomes) et 100 μ M de SAM pendant 16 heures à température ambiante. Les échantillons ont été séparés par SDS-PAGE et transférés sur une membrane de nitrocellulose. La marque H3K36me3 a été détectée par l'utilisation d'un anticorps spécifique.

D : Dosage de l'activité de SETD2 pour les 3 étapes de méthylation du résidu H3K36. 5 μ M de SETD2 (WT ou L1609P) ont été incubés avec 100 μ M SAM et 75 μ M de peptide substrat H3K36, H3K36me1 ou H3K36me2. Les peptides substrat-produits ont été séparés et détectés par HPLC.

E : Cinétiques apparentes K_m , k_{cat} et de l'efficacité enzymatique k_{cat}/K_m pour le substrat peptidique H3K36 et le cofacteur SAM.

F : Dosage de l'activité méthyltransférase de SETD2 pour l' α -tubuline et de l'activité d'autométhylation. 5 μ M de SETD2 ont été incubés avec 100 μ M de SAM tritiée en présence ou en absence de 1 μ g d' α -tubuline purifiée. Les échantillons ont été séparés par SDS-PAGE. Le gel a été séché pendant 90 minutes et exposé à un film d'autoradiographie.

G : La stabilité thermique de l'enzyme a été mesurée grâce à la technique de décallage thermique (TSA). 10 μ M de SETD2 ont été mises en présence de réactif SYPRO. Les échantillons ont été soumis à une augmentation graduelle de la température et la dénaturation thermique de la protéine a été suivie par mesure de la fluorescence au cours du temps.

H : L'activité résiduelle de SETD2 WT et L1609P, après incubation à des temps variables à 37 °C a été mesurée par HPLC. SETD2 a été préincubée à 37 °C (pendant 0, 5, 10, 15, 20 minutes), puis incubée pendant une nuit avec 100 μ M de SAM et 75 μ M de substrat peptidique H3K36. Les peptides substrat et produit ont été séparés par HPLC et l'aire des pics a été intégrée de manière à pouvoir calculer l'activité résiduelle de l'enzyme.

ligands entraînent aussi bien la stabilisation de la forme WT que de la forme mutante (suggérant qu'ils se fixent sur les deux formes enzymatiques, Fig. S-1A Supp.). Toutefois, l'écart ΔT_M entre les deux formes de l'enzyme en présence de ces ligands reste invariable ($\Delta T_M = 7$ °C).

En parallèle de ces approches de décallage thermique, des expériences de dénaturation thermique suivies de dosage de l'activité résiduelle ont été effectuées et confirment que le mutant SETD2 L1609P est inactivé plus rapidement que la forme WT par préincubation à 37 °C (Fig. S-2H).

L'enzyme SETD2 nécessite la présence de doigt de zinc structuraux (présents dans les domaines AWS et Post-SET) afin d'être correctement repliée et fonctionnelle. Nous avons voulu déterminer si la moindre stabilité thermique du mutant L1609P impliquait également ces doigts de zinc. Pour ce faire, nous avons chauffé l'enzyme à 37 °C et avons suivi la libération des ions Zn^{2+} dans la solution (Fig. S-1B Supp.). Nous observons un rélargage des ions Zn^{2+} plus rapide pour la forme SETD2 L1609P par rapport à celle de SETD2 WT, confirmant que la forme mutante est plus instable que la forme sauvage de l'enzyme.

2) Activité et stabilité de SETD2 L1609P *in cellulo*

a. Surexpression de SETD2 (domaine catalytique) dans des cellules HEK293T (KO SETD2)

De manière à mieux comprendre les effets de la mutation L1609P nous avons exprimé le domaine catalytique de SETD2 sous forme de protéine couplée à la GFP (Green Fluorescent Protein) par des approches de transfection. Pour ce faire, nous avons exprimé la forme GFP-SETD2 WT et GFP-SETD2 L1609P dans des cellules HEK293T n'exprimant pas SETD2 (SETD2 KO). Comme attendu, ces cellules KO SETD2 n'ont pas de marque épigénétique H3K36me3⁸⁴. L'expression dans ces cellules de la forme GFP-SETD2 WT permet de restaurer le phénotype de triméthylation du résidu H3K36, alors que l'expression du mutant SETD2 L1609P ne permet pas de restaurer la marque H3K36me3 (Fig. S-3A).

Par ailleurs, lors de la vérification de l'expression des formes GFP-SETD2 WT et GFP-SETD2 L1609P dans les cellules, nous avons constaté que le mutant s'exprime peu par rapport à la forme sauvage (Fig. S-3A). Ce défaut d'expression protéique n'est pas due à une différence de transfection, ni à une différence d'expression des ARNm de GFP-SETD2 WT et GFP-SETD2 L1609P. En effet, des expériences de qPCR montrent que les niveaux d'expression d'ARNm de SETD2 WT et L1609P dans les cellules transfectées sont équivalents (Fig. S-2 Supp.).

Suite aux résultats obtenus précédemment avec les formes recombinantes purifiées du domaine catalytique de SETD2 montrant que le mutant L1609P est plus instable que la forme sauvage, nous nous sommes demandés si la différence d'expression du mutant L1609P dans les cellules est due à sa dégradation. En effet, Bhattacharya *et al* ont montré que SETD2 est une protéine intrinsèquement instable dont l'expression est régulée par dégradation par le protéasome⁷⁹. Ces auteurs ont montré par ailleurs que la partie la plus stable de SETD2 est le domaine catalytique. Des études de transfection des cellules KO avec les constructions GFP-SETD2 suivies d'analyses par microscopie de fluorescence montrent que le mutant GFP-SETD2 L1609P s'exprime moins que la forme sauvage et qu'il forme des agrégats nucléaires (décrits précédemment comme des puncta)⁷⁹. Comme observé précédemment, ces expériences indiquent que le mutant SETD2 L1609P est plus instable

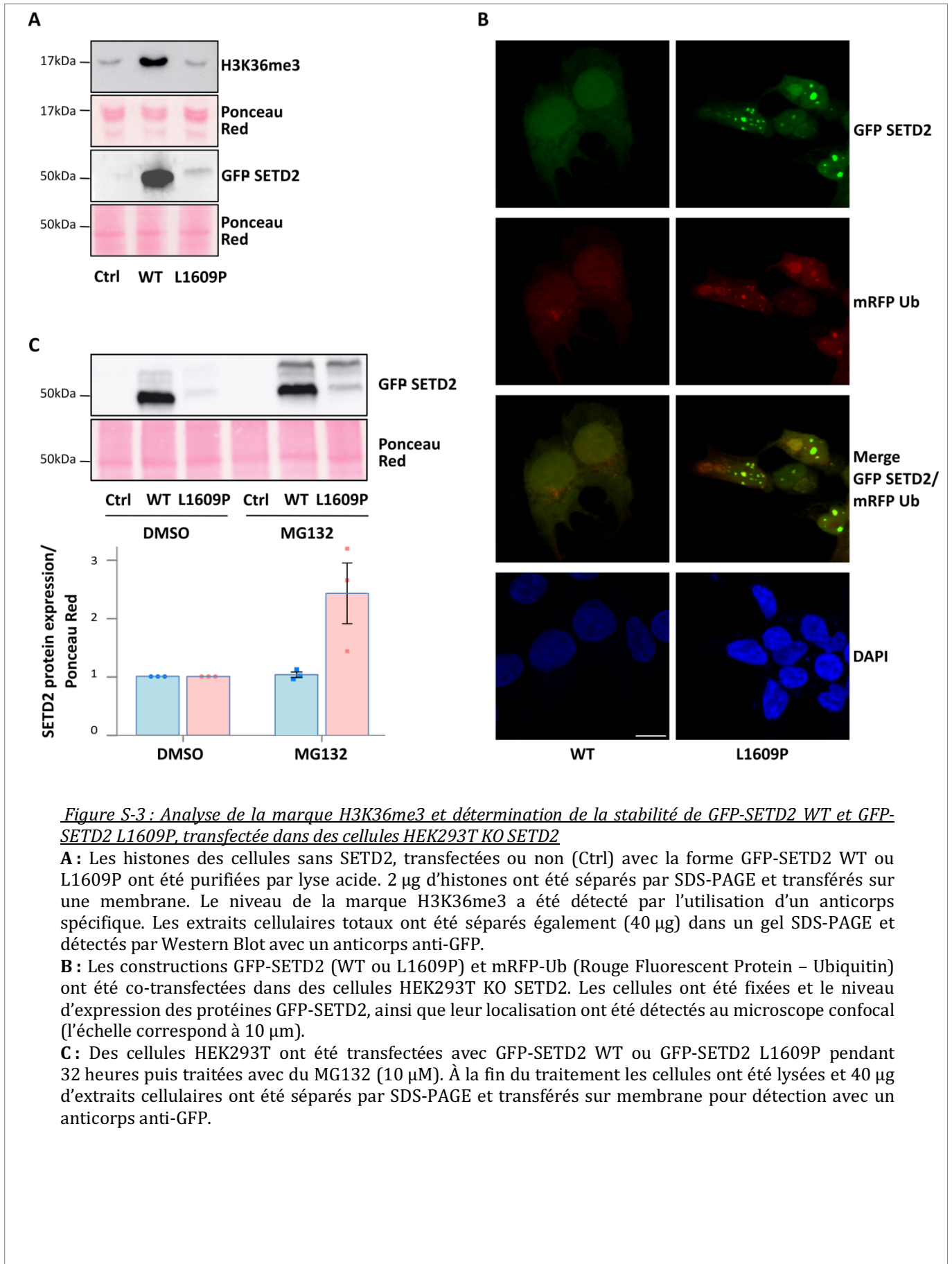
que la forme sauvage dans un contexte cellulaire (Fig. S-3B) et sont en accord avec les travaux de Bhattacharya *et al*⁷⁹ qui montrent que les propriétés de stabilité intrinsèque de SETD2 peuvent conduire à la formation d'agrégats et à la dégradation de l'enzyme par le protéasome. Des expériences de co-transfection des constructions GFP-SETD2 avec de l'ubiquitine couplée à la RFP (Rouge Fluorescent Protein) montrent une colocalisation des agrégats de GFP-SETD2 L1609P avec l'ubiquitine. Par ailleurs, nous montrons que l'inhibition du protéasome dans les cellules permet d'augmenter l'expression du mutant SETD2 L1609P dans les cellules transfectées (Fig. S-3C). Ces données indiquent que la forme mutante SETD2 L1609P s'aggrège et est dégradée dans la cellule conduisant à une faible expression du mutant en comparaison de la forme sauvage.

b. Expression de SETD2 L1609P dans des cellules HEK293T modifiées par approche CRISPR/Cas9

Les données obtenues précédemment confirment le phénotype de 'perte d'activité, de stabilité et d'expression' du mutant SETD2 L1609P. Pour confirmer les résultats précédents, nous avons modifié par approche CRISPR/Cas9 des cellules HEK293T, de manière à ce que ces cellules expriment de manière constitutive le mutant SETD2 L1609P ou la forme sauvage de l'enzyme.

En accord avec les résultats précédents, les cellules exprimant le mutant SETD2 L1609P présentent un niveau très faible de la marque H3K36me3 comme montré par des expériences de Western Blot des histones cellulaires et par immunofluorescence (Fig. S-4A). Par ailleurs, l'expression de la construction GFP-SETD2 WT permet de restaurer le phénotype de triméthylation du résidu H3K36 et la marque H3K36me3 dans les histones cellulaires (Fig. S-4B). *A contrario*, comme observé précédemment, l'expression de la forme GFP-SETD2 L1609P dans ces cellules ne permet pas de restaurer le niveau de la marque H3K36me3.

Afin de mieux comprendre l'effet de la mutation SETD2 L1609P au niveau génomique, nous avons procédé à des expériences de CHIP-qPCR avec un anticorps H3K36me3 sur 3 exons du



gène *myc* connues pour porter des marques H3K36me3 sur leurs nucléosomes⁸⁴ (Fig. S-4C). En accord avec les résultats précédents, nous observons que dans les cellules exprimant le mutant SETD2 L1609P, la marque H3K36me3 au niveau de ces trois exons du gène *myc*, présente une diminution significative par rapport aux cellules exprimant la forme sauvage.

Comme observé précédemment dans les expériences de transfection de la construction GFP-SETD2 L1609P dans des cellules HEK293T KO SETD2, l'expression du mutant SETD2 L1609P dans les cellules HEK293T modifiées par technologie CRISPR/Cas9 est très fortement diminuée par rapport aux cellules exprimant la forme sauvage de l'enzyme (Fig. S-4D). Par des expériences de détection par Western Blot, nous avons observé à nouveau un plus faible niveau d'expression de SETD2 dans le cas du mutant L1609P, et avons également confirmé la dégradation accrue de la forme mutante par le protéasome (Fig. S-4E).

Après confirmation du phénotype de 'perte d'activité, de stabilité et d'expression' du mutant SETD2 L1609P dans des cellules HEK293T modifiées, nous avons voulu comprendre si la mutation engendre d'autres modifications phénotypiques de ces cellules. Nous avons démontré que dans les cellules HEK293T CRISPR SETD2 L1609P, le niveau de la marque H3K36me3 est diminué par rapport aux cellules exprimant la forme sauvage de SETD2. Or, la diminution de la marque H3K36me3 a été associée à l'altération des processus de réparation de l'ADN et à des phénotypes de chimiorésistance à des anti-cancéreux génotoxiques tels que l'étoposide¹¹⁵. Nous avons ainsi étudié si le phénotype des cellules HEK293T CRISPR exprimant SETD2 L1609P ou la forme sauvage de l'enzyme présentait une altération de la réparation de l'ADN (via la détection du niveau de phosphorylation de l'histone H2A.X), de l'apoptose, et du cycle cellulaire suite à un traitement à l'étoposide.

Pour ce faire, nous avons traités pendant 72 heures les cellules HEK293T CRISPR SETD2 WT et HEK293T CRISPR SETD2 L1609P. À la fin du traitement, nous avons marqué les cellules apoptotiques avec l'Annexine V et nous avons trié au cytomètre les cellules selon leur niveau de marquage (Fig. S-4F). En effet, suite au traitement à l'étoposide, moins de cellules mortes sont présentes dans le cas des cellules HEK293T CRISPR SETD2 L1609P par rapport aux cellules HEK293T CRISPR SETD2 WT, suggérant une résistance accrue à des cellules SETD2 L1609P à l'apoptose induite par l'étoposide. Ces résultats sont en accord avec des

travaux montrant une résistance de cellules hématopoïétiques SETD2 KO à l'apoptose suite à un traitement à l'étoposide ¹¹⁵.

En parallèle, nous avons vérifié les niveaux de la marque γ H2A.X (histone H2A.X phosphorylé sur la sérine 139) dans les cellules suite à ce traitement. γ H2A.X est une marque de cassures et de stress de l'ADN associée à la présence de dommages à l'ADN ⁶³. Il a été montré précédemment que cette marque était perturbée dans des cellules avec des défauts d'activité de SETD2 (SETD2 KO) ⁸⁴. Nous avons traité les cellules HEK293T CRISPR SETD2 WT et HEK293T CRISPR SETD2 L1609P avec l'étoposide pendant 72 heures et nous les avons laissé récupérer pendant 0, 3 ou 5 jours après le traitement (Fig. S-3A Supp.). Nous avons constaté que le signal γ H2A.X est plus important dans le cas des cellules exprimant le mutant SETD2 L1609P. Ceci est en accord avec des travaux précédents montrant la perturbation de la réparation de l'ADN dans des cellules HEK293T SETD2 KO traitées avec des drogues génotoxiques ^{84,182}.

Par ailleurs, une caractéristique souvent associée à la résistance aux drogues chimiothérapeutiques (notamment étoposide) est la dérégulation du cycle cellulaire. Par des expériences de marquage des cellules à l'iodure de propidium et de leur séparation par cytométrie en flux (Fig. S-3B Supp.) nous avons observé une altération des différentes phases du cycle cellulaire avec une accumulation de cellules en phase G2/M. Ces résultats sont accord avec des données obtenus récemment avec des cellules SETD2 KO traitées par ara-C, une drogue génotoxique ¹⁸².

L'ensemble de ces résultats indiquent que les cellules HEK293T exprimant la forme SETD2 L1609P ont un phénotype de réponse à l'étoposide différent des cellules exprimant la forme sauvage. Un phénotype similaire (résistance à l'apoptose, altération de la marque γ H2A.X, blocage du cycle cellulaire en G2/M) a été observé dans des travaux portant sur des cellules déficientes en SETD2 (et en marque épigénétique H3K36me3) et exposées à des drogues génotoxiques.

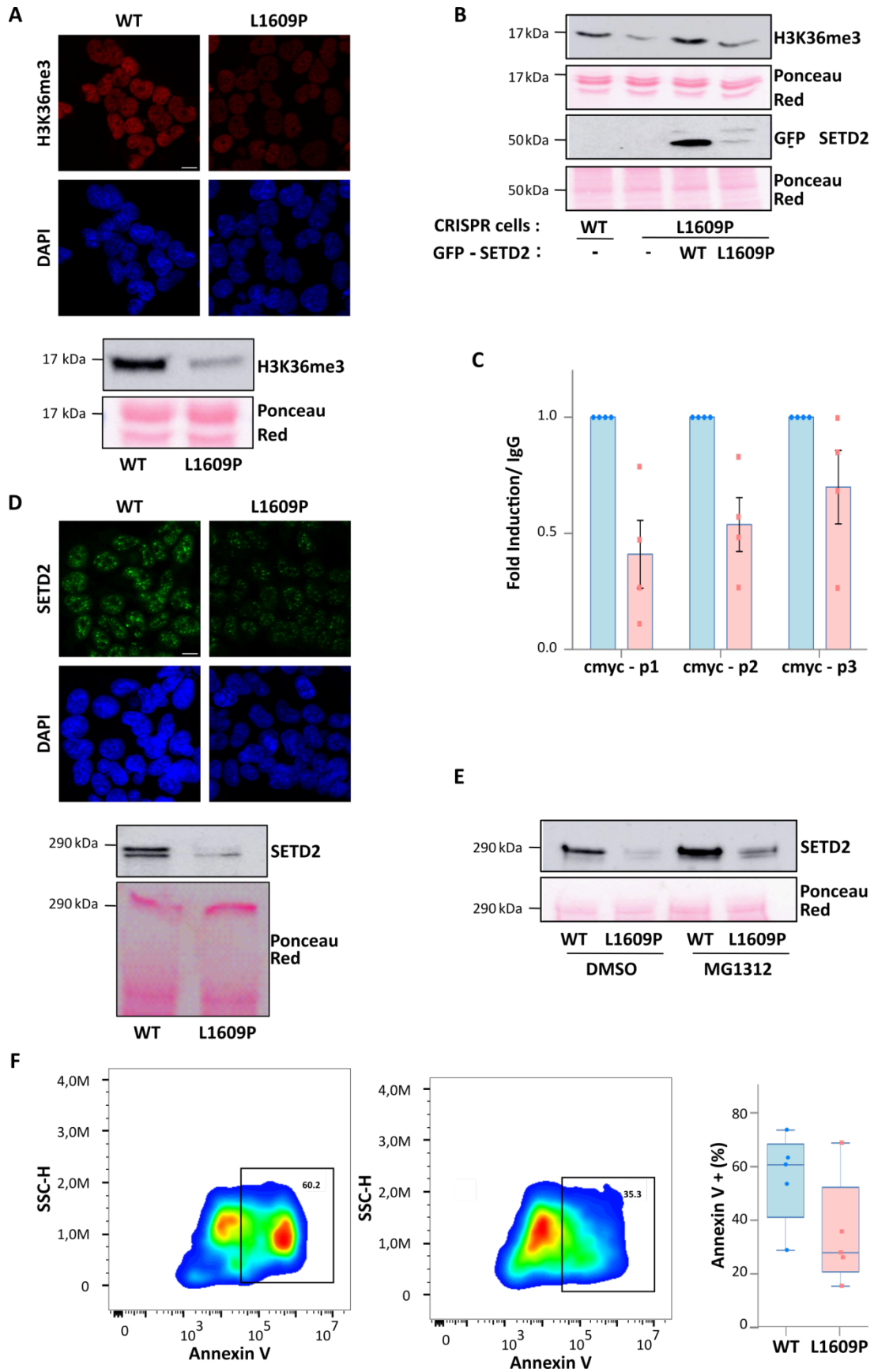


Figure S-4 : Analyse de la marque H3K36me3, détermination de la stabilité de SETD2 endogène et résistance au traitement à l'étoposide des cellules HEK293T CRISPR SETD2 WT et SETD2 L1609P

A : Niveau de triméthylation du résidu H3K36 dans les cellules HEK293T CRISPR SETD2 WT et HEK293T CRISPR SETD2 L1609P. Les cellules ont été fixées et la marque H3K36me3 a été détectée par immunofluorescence au microscope confocale (échelle à 10 µm). En parallèle, les histones endogènes ont été purifiées par extraction acide et 2 µg ont été séparés par SDS-PAGE et transférés sur une membrane pour une détection avec un anticorps contre H3K36me3.

B : Surexpression de GFP-SETD2 WT et GFP-SETD2 L1609P dans les cellules HEK293T CRISPR SETD2 L1609P. Les cellules ont été récupérées et lysées dans une solution PBS Triton 1 %, puis soniquées et centrifugées pendant 15 minutes à 15000 g à 4 °C. Le surnageant a été récupéré et 40 µg ont été séparés par SDS-PAGE, puis transférés sur membrane pour détection avec un anticorps anti-GFP. La fraction non soluble a été utilisée pour faire une extraction acide des histones et 2 µg ont été séparés par SDS-PAGE. Les protéines ont été transférées sur membrane et détectées avec un anti-H3K36me3.

C : Immunoprécipitation de la chromatine par des anticorps contre la marque H3K36me3 et l'histone H3 des cellules HEK293T CRISPR SETD2 WT et L1609P. Le niveau d'ADN co-immunoprécipité a été détecté par qPCR en utilisant des amorces contre 3 exons différents du gène *myc*.

D : Détection du niveau d'expression protéique de SETD2 dans les cellules HEK293T CRISPR SETD2 WT et L1609P. Les extraits totaux (40 µg) ont été séparés par SDS-PAGE et SETD2 a été détectée par Western Blot avec un anticorps anti-SETD2. En parallèle les cellules ont été fixées et marquées avec le même anticorps puis analysées par immunofluorescence au microscope confocal (échelle à 10 µm).

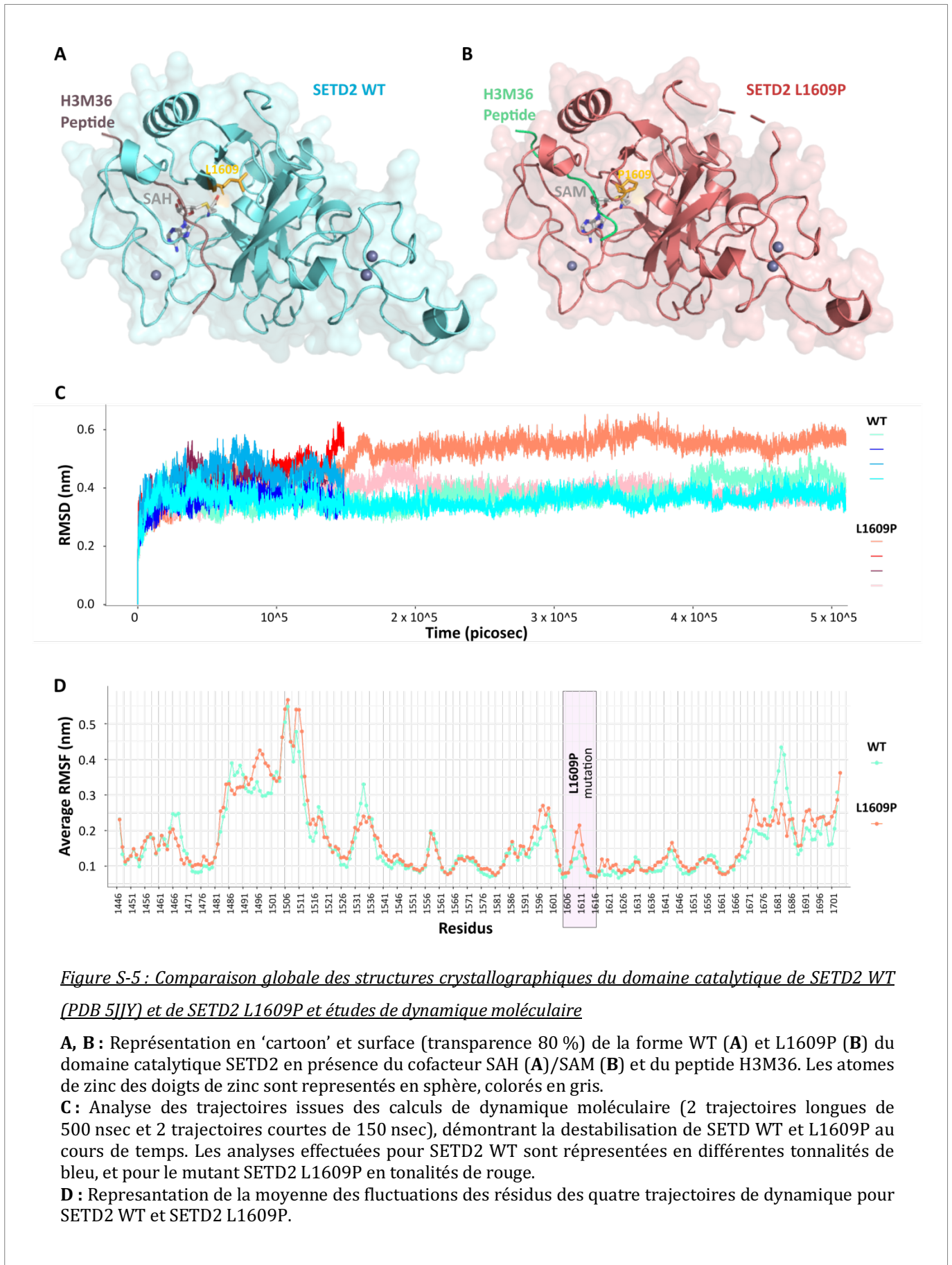
E : Les cellules ont été traitées pendant 16 heures avec du MG132 (10 µM). Le niveau de SETD2 des extraits cellulaires a été détecté par l'utilisation d'un anticorps anti-SETD2.

F : Détection par cytométrie en flux du pourcentage des cellules positives en Annexine V suite au traitement de 72 heures avec 400 nM d'étoposide. À la fin du traitement, les cellules ont été récupérées, marquées à l'Annexine V et analysées par cytométrie en flux. Le logiciel Flowjo a été utilisé pour l'analyse des résultats.

3) Effet de la mutation L1609P sur la structure de SETD2

Par nos travaux en biochimie moléculaire et cellulaire, nous avons démontré que le mutant SETD2 L1609P présente un phénotype de 'perte d'activité et de stabilité'. Cependant, de manière à comprendre à quoi est due cette perte d'activité, nous avons procédé à des expériences de cristallisation du domaine catalytique de SETD2 L1609P. Malgré l'instabilité relative de ce mutant de SETD2, nous avons réussi à obtenir la structure cristallographique du domaine catalytique de ce mutant de SETD2 à 2,2 Angström (Å) de résolution. Nous avons comparé dans un premier temps cette structure L1609P avec la structure SETD2 WT utilisée comme référence (code PDB 5JJY). SETD2 WT issue de la structure 5JJY est cocrystallisée avec le peptide H3M36 (centré sur la lysine 36 qui est mutée en méthionine) et le cofacteur SAH (Fig. S-5A). La structure de SETD2 L1609P a été obtenue en présence du cofacteur SAM et du peptide H3M36 (Fig. S-5B). La comparaison des deux structures ne présente pas de changements structuraux globaux majeurs (RMSD = 0,275 nm). Toutefois, l'analyse de ces structures montre des altérations conformationnelles locales dues à la mutation L1609P, comme le montrent les représentations des structures secondaires et la surface de deux formes enzymatiques.

Nous avons précédemment décrit que le mutant SETD2 L1609P est plus instable que la forme sauvage. Afin de mieux comprendre l'impact de la mutation sur la structure du domaine catalytique de SETD2, nous avons effectué des expériences de dynamique moléculaire. Deux simulations 'longues' de 500 nsec et 'courtes' de 150 nsec ont été conduites et les variations des positions des résidus (C α) ont été suivies par mesure des RMSD (Root Mean Square Deviation) et RMSF (Root Mean Square Fluctuations) (Fig. S-5C). Ces simulations montrent qu'en globalité, le mutant SETD2 L1609P fluctue plus que la forme SETD2 WT. Ensuite, nous avons moyenné les quatre simulations pour chacune des structures et avons observé les fluctuations au niveau de chacun des résidus de la structure. Plusieurs régions des deux protéines présentent des différences de fluctuations (Fig. S-5D), notamment au niveau des résidus 1609-1612 proches du résidu muté (L1609), et qui apparaissent comme plus flexibles dans le cas du mutant L1609P.



Wavelength	
Resolution range (Å)	40.42 - 2.19 (2.268 - 2.19)
Space group	P 21 21 21
Unit cell (Å, °)	59.718 77.408 77.884, 90 90 90
Total reflections	256568 (25912)
Unique reflections	19159 (1868)
Multiplicity	13.4 (13.9)
Completeness (%)	99.83 (99.68)
Mean I/sigma(I)	20.66 (2.69)
Wilson B-factor (Å ²)	49.14
R-merge	0.07346 (0.9469)
R-meas	0.07643 (0.983)
R-pim	0.02088 (0.2625)
CC1/2	0.999 (0.92)
CC*	1 (0.979)
Reflections used in refinement	19131 (1863)
Reflections used for R-free	964 (74)
R-work	0.2154 (0.3117)
R-free	0.2561 (0.3810)
CC(work)	0.940 (0.898)
CC(free)	0.934 (0.835)
Number of non-hydrogen atoms	2170
macromolecules	2084
Ligands	30
Solvent	56
Protein residues	261
RMS(bonds, Å)	0.008
RMS(angles, °)	1.02
Ramachandran favored (%)	96.86
Ramachandran allowed (%)	2.75
Ramachandran outliers (%)	0.39
Rotamer outliers (%)	1.77
Clashscore	7.23
Average B-factor (Å)	60.12
Macromolecules (Å)	60.47
Ligands (Å)	47.51
Solvent (Å)	53.99

Tableau S – 1 : Statistiques de collecte et d'affinement des données de la structure cristallographique du domaine catalytique de SETD2 L1609P

Les statistiques pour l'enveloppe de plus haute résolution sont indiquées entre parenthèses.

Nous avons par la suite focalisé notre analyse sur la région de la mutation L1609P qui est proche du site de fixation du peptide H3M36 (site de fixation des substrats protéiques, Fig. S-6A). De manière intéressante, la mutation L1609P située à la fin d'un brin β , entraîne une déformation de cette structure secondaire et son 'remplacement' par une structure en boucle. Ceci est en accord avec les données issues des analyses computationnelles de dynamique moléculaire qui montrent que la structure du mutant est plus flexible, notamment au niveau de la mutation et des résidus environnants. En parallèle à ce changement structural, nous avons observé un réarrangement important des résidus K1610-E1613 qui constituent la structure en épingle à cheveux entre les brins β 5 et β 6 (Fig. S-6B). Par ailleurs, une des distances caractéristiques d'une épingle à cheveux (entre le C α du résidu K1610 et le C α du résidu E1613) est significativement modifiée, avec un doublement de sa taille (4,8 Å pour la structure SETD2 WT et 8,4 Å pour la structure SETD2 L1609P).

Un des résidus de la structure en épingle à cheveux le plus impacté est le résidu K1610. Les angles de torsion Phi et Psi de cette lysine sont impactés suite à la mutation L1609P et ceci est observé dans toutes les trajectoires de dynamique moléculaire (Fig. S-6C). Ce changement d'angles de la K1610 (basculement de 150 °) engendre un encombrement stérique et un changement conformationnel d'autres résidus de la structure SETD2 L1609P (Fig. S-6D). Plus précisément, contrairement à la structure de SETD2 WT la chaîne latérale du résidu K1610 du mutant est dirigée vers la chaîne latérale du résidu K37 du peptide H3M36 qui en raison de cet encombrement, adopte une autre conformation significativement plus éloignée du résidu K1610. De plus, la chaîne latérale du résidu K1610 se dirige vers la chaîne latérale du résidu K1639. L'encombrement stérique entre les résidus K1610 et K1639 entraîne également un changement conformationnel de la K1639, dont l'angle chi1 est modifié dans toutes les trajectoires de dynamique moléculaire. Par ailleurs, les changements conformationnels des résidus K1610 et K1639 induisent le rapprochement de ces résidus vers le site de fixation du substrat peptidique, qui change d'orientation à son extrémité C-ter à partir du résidu P39 du peptide H3M36 (Fig. S-6D). La poche de fixation du peptide est également d'une dimension plus restreinte dans le cas du mutant SETD2 L1609P, entraînant ainsi un changement conformationnel du peptide au niveau du résidu P39 (Fig. S-6E). De manière curieuse, alors que nous avons utilisé un peptide identique que

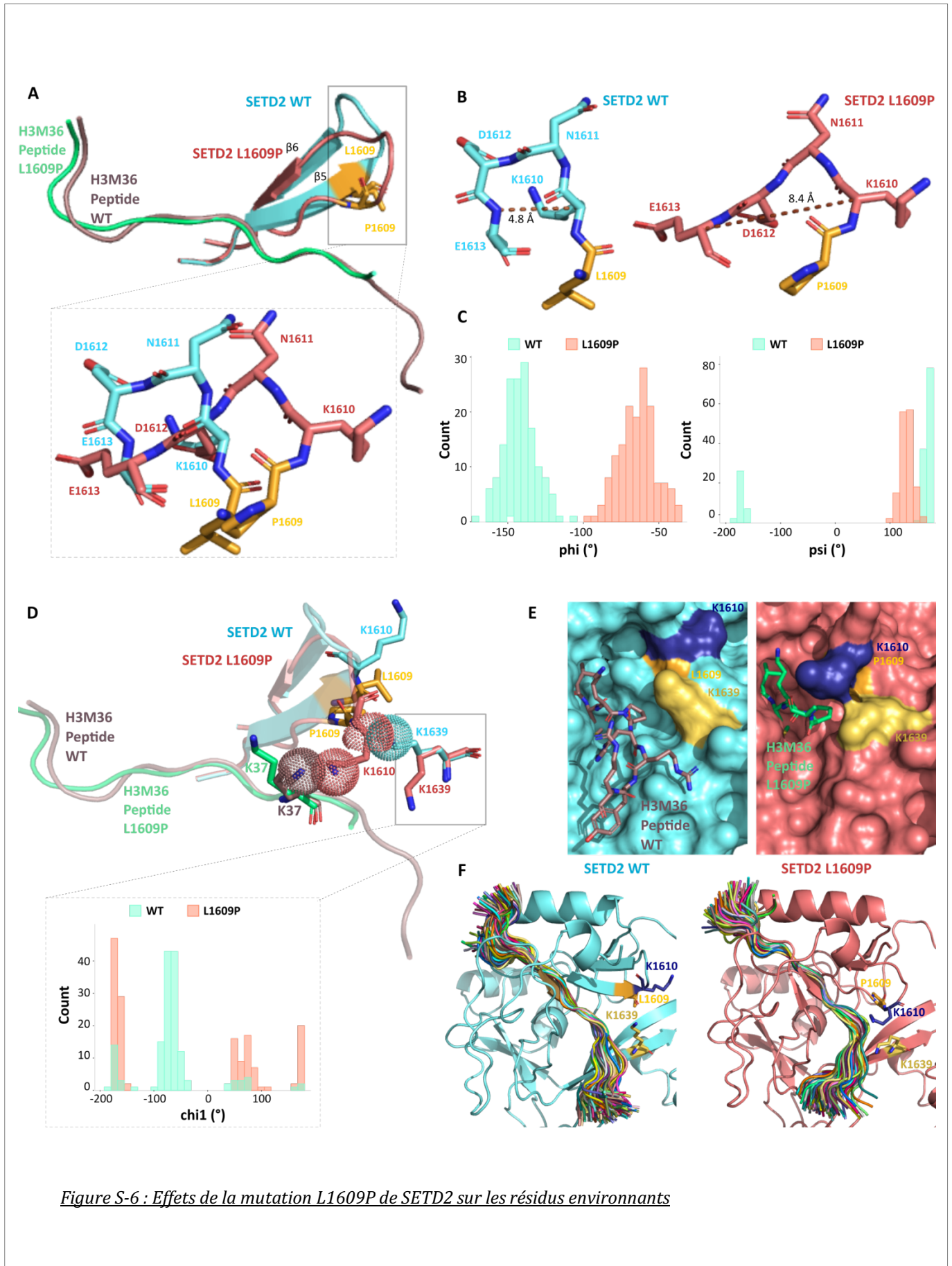


Figure S-6 : Effets de la mutation L1609P de SETD2 sur les résidus environnants

A : Alignement structural du 'β hairpin' formé par les brins β5-β6 en présence du peptide substrat H3M36. Représentation en 'cartoon'. SETD2 WT est colorée en aquamarine et SETD2 L1609P en deepsalmon. La mutation L1609P est colorée en brightorange (stick). Le peptide H3M36 dans la structure SETD2 WT est coloré en violet et le peptide H3M36 dans la structure SETD2 L1609P est vert.

B : Représentation en stick (CPK) des résidus du 'β hairpin' entre des brins β5-β6 (L1609-E1613).

C : Analyse issue des calculs de dynamique moléculaire, avec représentation des angles phi et psi du résidu K1610, observés au cours des trajectoires.

D : Alignement structural du 'hairpin' formé par les brins β5-β6 des structures SETD2 WT (PDB : 5JJY, aquamarine) et SETD2 L1609P (deepsalmon) en présence du peptide substrat H3M36. Les résidus 1609 (orange), 1610, 1639 de SETD2 et K37 du peptide H3M36 sont représentés en stick CPK. Les différents angles χ₁ du résidu K1639 au cours des trajectoires de dynamique moléculaire sont représentés en barplot.

E : Représentation en surface de la région de fixation du peptide H3M36 de SETD2 WT et SETD2 L1609P. SETD2 WT est colorée en aquamarine et SETD2 L1609P en deepsalmon. La surface du résidu 1609 est représentée en brightorange, celle du résidu 1610 en bleu et celle du résidu K1639 en jaune. Le peptide H3M36 est en stick (CPK), coloré en violet dans la structure SETD2 WT et en vert dans la structure SETD2 L1609P.

F : Les différentes conformations du peptide H3M36 au cours des trajectoires de dynamique moléculaire sont représentées en 'cartoon'. La structure SETD2 WT est colorée en aquamarine et la structure SETD2 L1609P en deepsalmon.

celui utilisé dans le cas de la structure 5JJY, la densité électronique en C-ter du peptide de notre structure est faible et les derniers résidus du peptide ne peuvent pas être correctement résolus. Nous avons émis l'hypothèse que ceci peut être en lien avec le changement conformationnel du peptide au niveau du résidu P39, et que cette nouvelle conformation induit une flexibilité plus importante du peptide à ce niveau. Pour vérifier notre hypothèse, nous avons effectué des expériences de dynamique moléculaire des structures SETD2 WT et SETD2 L1609P, en utilisant dans les deux cas le peptide H3M36 issue de la structure 5JJY (Fig. S-6F). Nous observons que l'extrémité C-ter du peptide est beaucoup plus flexible dans le cas de la dynamique avec la structure SETD2 L1609P, suggérant ainsi un impact de la mutation L1609P sur la fixation du substrat.

III. Discussion

SETD2 est une histone méthyltransférase décrite comme un suppresseur de tumeur associé à l'initiation et au développement tumoral. Une pléthore de mutations affectant l'enzyme est documentée dans plusieurs types de cancers, notamment des carcinomes rénaux à cellules claires, des leucémies et des lymphomes. Ces mutations, majoritairement faux sens, sont situées principalement dans le domaine catalytique de SETD2 (constitué des sous-domaines AWS – SET – Post-SET). Malgré le nombre important et croissant de mutations oncogéniques de SETD2 décrites chez les patients atteints de cancer, les effets de ces altérations moléculaires sur la structure et la fonction de SETD2 restent mal décrits. À ce jour, l'étude de Hacker *et al* est la seule qui porte sur la caractérisation d'une mutation oncogénique située dans le domaine catalytique de SETD2⁸⁴. Les auteurs ont procédé à l'analyse de l'effet de la mutation SETD2 R1625C sur l'activité et la stabilité de l'enzyme. La mutation R1625C altérerait la fixation du substrat protéique (histone H3) sur l'enzyme et entraînerait la diminution de l'activité enzymatique vis-à-vis de l'histone H3.

Notre travail a porté sur la caractérisation fine de l'effet d'une mutation oncogénique de SETD2 au niveau structural et fonctionnel. Nous nous sommes intéressés à l'étude de la mutation SETD2 L1609P, identifiée chez des patients pédiatriques atteints de leucémie lymphoblastique aiguë¹¹⁴. D'une manière générale, le résidu leucine en position 1609 est très conservé dans la plupart des histones méthyltransférases, notamment G9a, EZH2 et MLL1. L'alignement de séquences indique qu'à cette position est majoritairement présent un résidu leucine, isoleucine ou phénylalanine, suggérant la nécessité d'un résidu hydrophobe et volumineux à cette position. Le résidu L1609 se situe dans le motif SET-I du domaine SET de SETD2. Ce motif SET-I (faisant partie du domaine catalytique AWS – SET – Post-SET) est de taille variable selon les méthyltransférases, mais de structure conservée entre les différentes HMT à domaine SET⁷². Étant donné la faible similarité de séquence de la région SET-I des HMT, il est supposé que cette dernière participe à la spécificité de reconnaissance des substrats protéiques⁸⁶. Ceci permet de supposer que la mutation d'un

résidu dans cette région, pourrait altérer les interactions avec les substrats et l'activité de l'enzyme.

Afin de mieux comprendre l'impact de cette mutation d'un point de vue fonctionnel, nous avons utilisé des approches de biochimie moléculaire et cellulaire, basées sur l'utilisation de protéines recombinantes et de modèles cellulaires. En effet, dans un premier temps, nous avons exprimé en bactérie et purifié le domaine catalytique (AWS – SET – Post-SET) de SETD2 WT et SETD2 L1609P. Par des dosages d'activité lysine méthyltransférase, nous avons montré que la mutation L1609P entraîne une perte d'activité importante de l'enzyme sur ses substrats protéiques (histone H3 et α -tubuline), ainsi qu'un défaut de l'auto-méthylation de l'enzyme (Sun *et al*)⁷⁶. Par ailleurs, la forme mutante L1609P apparaît comme intrinsèquement moins stable que la forme sauvage. Cette instabilité structurale, due à la mutation L1609P affecte l'activité intrinsèque de l'enzyme et conduit à son agrégation et sa dégradation dans les cellules⁷⁹.

Comme indiqué ci-dessus des analyses cinétiques montrent que l'efficacité catalytique de l'enzyme vis-à-vis de peptides H3 est perturbée par la mutation (k_{cat}/K_M de SETD2 WT = $2.10^{-4} \mu M^{-1}.min^{-1}$ et de SETD2 L1609P = $8.10^{-5} \mu M^{-1}.min^{-1}$). Bien que la mutation se situe dans une région décrite comme impliquée dans l'interaction avec les substrats protéiques (H3 notamment), les analyses cinétiques ne montrent pas d'altérations significatives de la liaison du peptide H3. Toutefois les cinétiques ont été faites avec des peptides qui ne reflètent pas la complexité de l'interaction entre l'enzyme et l'histone H3 dans les nucléosomes. Aussi, nous avons voulu tester l'effet de la mutation sur la liaison à l'histone H3 dans les nucléosomes par une expérience de retard sur gel en conditions natives. Ce type d'expérience a déjà été conduit par Liu *et al* afin de déterminer l'importance potentielle du domaine SRI sur la liaison aux nucléosomes⁹⁷. Pour ce faire, les auteurs ont utilisé deux constructions contenant les résidus 1382-1915 avec ou sans le domaine SRI. Ce dernier ne semble pas essentiel pour la fixation aux nucléosomes. Du fait du pI (point isoélectrique) du domaine catalytique de SETD2, ces expériences n'ont pas été conclusives. De ce fait, des expériences de mesure de la constante de dissociation (K_D) de l'enzyme par décalage thermique (TSA) sont en cours de réalisation. Cette mesure va nous permettre d'avoir des informations sur la capacité de l'enzyme à fixer son substrat, et de

voir si celle-ci est perturbée par la mutation L1609P. Des approches par microcalorimétries sont également envisagées.

À ce jour, aucune structure cristallographique d'un mutant de SETD2 n'a été publiée ou disponible dans la banque PDB. Hacker *et al* ont toutefois procédé à des études d'analyse structurale *in silico* et démontré un impact éventuel de la mutation SETD2 R1625C sur la fixation du cofacteur SAM⁸⁴. Notre travail a permis de résoudre la structure cristallographique du mutant SETD2 L1609P en présence du cofacteur SAM et d'un substrat peptide H3M36. Cette structure nous a permis de montrer un impact potentiel de la mutation sur la région de fixation des substrats protéiques. En effet, la mutation L1609P induit des changements conformationnels significatifs autour de la mutation, avec une forte perturbation au niveau d'une structure en épingle à cheveux, formée par les résidus K1610-E1613. En conséquence, la position de la chaîne latérale du résidu K1610 est altérée et dirigée vers le site de fixation du substrat, entraînant un encombrement stérique ('clash' stérique) avec le résidu K1639 dont la chaîne latérale s'oriente également vers le site de fixation du substrat. Ces changements conformationnels entraînent vraisemblablement la diminution de la taille de la poche de fixation du substrat.

En parallèle, nous avons procédé à des analyses de dynamique moléculaire, afin de mieux comprendre l'effet structural engendré par la mutation L1609P sur SETD2. Cette analyse a été effectuée sur une forme semi-Holo, en présence uniquement du peptide substrat et en absence du cofacteur SAM. Ceci est dû à des difficultés de paramétrage du cofacteur SAM dans les outils de dynamique moléculaire. Les trajectoires de dynamique moléculaire ont permis toutefois de comprendre l'impact de la mutation sur la stabilité intrinsèque de l'enzyme, en montrant une augmentation significative de la mobilité de la région proche de la mutation au cours des trajectoires de dynamique.

Pour mieux comprendre l'effet de la mutation dans un contexte cellulaire, nous avons utilisé des approches de transfection de cellules HEK293T avec des constructions du domaine catalytique de SETD2 couplé à la GFP. Le domaine catalytique de SETD2 est décrit par Bhattacharya *et Workman* comme la partie la plus stable de l'enzyme. En effet, SETD2 est considérée comme protéine globalement instable (présence de plusieurs régions non structurées à faible complexité) dont l'expression est finement régulée par dégradation par

le protéasome⁷⁹. Les auteurs ont ainsi démontré que la partie N-terminale de SETD2 forme des agrégats ubiquitinylés qui ne sont pas observés avec la partie de l'enzyme contenant le domaine catalytique. Par ailleurs, Hacker *et al* ont montré que la mutation SETD2 R1625C conduit à une forme enzymatique moins stable que l'enzyme sauvage, le mutant étant plus facilement dégradé par le protéasome⁸⁴. Les données obtenues nous ont permis de montrer une faible expression du mutant SETD2 L1609P, résultant vraisemblablement d'une plus importante dégradation par le protéasome⁷⁹. En complément de l'étude Bhattacharya *et Workman*, nous avons observé l'agrégation du domaine catalytique du mutant SETD2 L1609P, qui n'est pas observée avec l'enzyme sauvage. Parallèlement, nous observons que contrairement à la forme SETD2 WT, la transfection de la construction GFP-SETD2 L1609P ne permet pas de restaurer significativement le niveau de la marque H3K36me3 dans les cellules. La perte significative de l'activité de méthylation de l'histone H3 par SETD2 L1609P dans les cellules est vraisemblablement due à l'instabilité (formation d'agrégats), à la faible expression de cette forme mutante et à sa faible activité intrinsèque.

En parallèle à ces approches de transfection de cellules KO SETD2, nous avons généré des cellules HEK293T modifiées par approche CRISPR/Cas9 pour exprimer le mutant SETD2 L1609P. Ce modèle est le plus pertinent biologiquement et nous permet de mieux décrire l'effet de la mutation L1609P sur l'enzyme SETD2 dans la cellule. Cependant, notre modèle n'est toutefois pas optimal car la mutation SETD2 L1609P est identifiée dans des cellules leucémiques (plus difficiles à modifier par approche CRISPR/Cas9), les cellules HEK293T étant des cellules épithéliales rénales. Ce modèle HEK293T exprimant le mutant L1609P toutefois un modèle biologique très important et utile pour comprendre les effets moléculaires et cellulaires de la mutation L1609P.

La comparaison des cellules HEK293T CRISPR SETD2 WT et HEK293T CRISPR SETD2 L1609P a permis de confirmer les données obtenues avec les protéines recombinantes et les cellules SETD2 KO en montrant le phénotype 'perte d'activité, de stabilité et d'expression' dû à la mutation L1609P. Ce modèle cellulaire a clairement permis de montrer que la mutation SETD2 L1609P conduit à une diminution de la marque épigénétique H3K36me3 dans les cellules. Des expériences de CHIP-qPCR sur différents exons du gène *myc* confirment cette diminution de la marque H3K36me3. De manière intéressante, le niveau de la marque

H3K36me3 des nucléosomes présents dans les régions exoniques de ce gène a été décrit comme significativement diminué dans des cellules exprimant le mutant inactif SETD2 R1625C⁸⁴. Des expériences de ChIP-Seq sur les cellules HEK293T CRISPR SETD2 WT et L1609P sont en cours d'analyse. Toutefois les données préliminaires montrent, comme attendu, une baisse de la marque H3K36me3.

En outre, nous avons montré que la mutation L1609P confère des propriétés phénotypiques de résistance à l'étoposide, agent chimiothérapeutique génotoxique. De manière intéressante, plusieurs études ont permis de mettre en évidence ce type de phénotype dans des cellules SETD2 KO (notamment des cellules hématopoïétiques) pour lesquelles la marque épigénétique H3K36me3 était drastiquement diminuée^{115,182,183}. Nous avons observé une diminution de la mort par apoptose dans le cas de cellules HEK293T CRISPR SETD2 L1609P vraisemblablement liée à une dérégulation du cycle cellulaire. Cette différence à la réponse aux traitements génotoxiques a déjà été documentée par Mar *et al* et Dong *et al* dans le cas d'utilisation de cellules SETD2 KO^{115,182}. Par ailleurs, l'étude du signal γ H2A.X après traitement à l'étoposide confirme une altération du processus de réparation de l'ADN dans les cellules HEK293T CRISPR SETD2 L1609P. Ceci est en accord avec le rôle de la marque H3K36me3 dans ce processus. Par ailleurs, Hacker *et al* ont également observé une altération de la réparation des cassures de l'ADN et une persistance du signal γ H2A.X dues à la mutation inactivatrice SETD2 R1625C⁸⁴.

Le modèle cellulaire HEK293T CRISPR SETD2 L1609P développé au cours de nos travaux est un outil très important pour la compréhension des fonctions de SETD2 et des processus biologiques altérés par l'expression de mutants pathogéniques. À titre d'exemple, des expériences de métabolomique ainsi que d'études du cytosquelette et de la dynamique cellulaire sont d'ores et déjà lancées et devraient apporter de nouveaux éléments sur les fonctions de SETD2.

Globalement, ce travail représente la première étude complète de l'impact moléculaire, cellulaire et fonctionnel d'une mutation oncogénique affectant le domaine catalytique de SETD2 (Fig. S-7).

IV. Matériels et méthodes

1) Alignement des séquences du domaine SET

Les séquences des principaux membres de la famille des histones méthyltransférases à domaine SET chez l'Homme (SETD2, G9A, EZH2, NSD1, NSD2, SETD8, MLL1, MLL2, SETD7, ASH1) ont été obtenues à partir de la base de données Uniprot/Swiss-prot (Q9BYW2, Q96KQ7, Q15910, Q96L73, O96028, Q9NQR1, Q03164, O14686, Q8WTS6, Q9NR48, respectivement). L'alignement multiple des séquences du domaine SET de ces enzymes a été obtenu grâce au serveur Clustal Ω (EMBL-EBI) et analysé par le logiciel ESPrict 3¹⁸⁵.

2) Mutagenèse dirigée, construction du mutant L1609P

Vecteur d'expression bactérienne : Le vecteur d'expression bactérienne de type pet28a-MHL, contenant l'insert d'ADN qui code le domaine catalytique (constitué des sous domaines AWS – SET – Post-SET) de la protéine humaine SETD2 (résidus 1433-1711) a été obtenu auprès de la société Addgene (#25348, plasmide pet28a-SETD2 WT). La mutation SETD2 L1609P (Leu en Pro) a été introduite dans ce plasmide par l'utilisation du kit Agilent QuickChange II XL selon les instructions du fabricant en utilisant les amorces complémentaires suivantes :

SETD2_F_L1609P : 5' - CACTATTACTTTATGGCTC **C** GAAAAATGATGAAATTATCGACGCG- 3'

SETD2_R_L1609P : 5' - CGCGTCGATAATTTTCATCATTTTTTC **G** GAGCCATAAAGTAATAGTG- 3'

Le plasmide portant la mutation SETD2 L1609P a été utilisé pour transformer des bactéries compétentes *DH5α*, qui ont été sélectionnées sur des boîtes LB (Luria Bertani) agar contenant 30 µg/µl de kanamycine. La présence de la mutation a été vérifiée par séquençage (Eurofins).

De manière à exprimer SETD2, les plasmides codant la forme SETD2 WT et SETD2 L1609P ont été insérés par choc thermique à 42°C dans des bactéries compétentes *E.Coli BL21 (DE3)*. Les bactéries avec le plasmide d'expression pet28a-SETD2 WT ou pet28a-SETD2 L1609P ont été sélectionnées sur des boîtes LB agar contenant 30 µg/µl de kanamycine.

Vecteur d'expression eucaryote : Pour les études de SETD2 au niveau cellulaire, nous avons obtenu le vecteur d'expression eucaryote pEGFP-SETD2 contenant l'insert d'ADN complémentaire humain codant pour SETD2 (résidus 500-2564), auprès du laboratoire du Dr. Sergio de Almeida (Portugal). Nous avons également sous cloné dans ce même vecteur la partie de *Setd2* codant les domaines AWS – SET – Post-SET (1433-1711). Pour ce faire, nous avons utilisé les sites de restriction du plasmide BglII et PstI. Nous avons nommé cette construction 'GFP-SETD2 domaine catalytique' et nous avons utilisé à nouveau le kit Agilent QuickChange II XL pour introduire la mutation L1609P. Nous avons utilisé comme amorces complémentaires les amorces suivantes :

SETD2 F L1609P : 5' - CATCCATTACTATTTTCATGGCCCGGAAGAATGATGAGATAATAG- 3'

SETD2 R L1609P : 5' - CTATTATCTCATCATTCTTCGGGCCATGAAATAGTAATGGATG- 3'

Le plasmide portant la mutation SETD2 L1609P a été utilisé pour transformer des bactéries compétentes *DH5α*, qui ont été sélectionnées sur des boîtes LB (Luria Bertani) agar contenant 30 µg/µl de kanamycine. La présence de la mutation a été vérifiée par séquençage (Eurofins).

3) Expression et purification de SETD2 recombinante

Pré-culture bactérienne : Les bactéries *E.Coli BL21 (DE3)/ pet28a -SETD2 WT* et *E.Coli BL21 (DE3)/ pet28a-SETD2 L1609P* ont étéensemencées dans du milieu LB (Sigma) en présence de 30 µg/µL kanamycine (Euromedex) sur la nuit sous agitation (150 rpm) à 37 °C.

Culture bactérienne et induction de l'expression protéique : Du milieu LB contenant de la kanamycine (30 µg/µl) etensemencé avec la pré-culture bactérienne a été placé à 37 °C sous agitation (150 rpm). Lorsque l'absorbance du milieu à 595 nm a atteint 0,6-0,7 unités d'absorption, l'expression protéique a été induite par l'ajout de 500 µM IPTG (IsoPropyl β-D-1-ThioGalactopyranoside, Euromedex). À la suite, les cultures bactériennes ont été placées pour une nuit à 16 °C sous agitation (150 rpm). Les bactéries ont été récoltées par centrifugation à 4400 rpm pendant 20 minutes. Les bactéries ont été lavées avec du PBS et centrifugées à nouveau à 4400 rpm pendant 15 minutes. Les bactéries ont été stockées à -80 °C ou utilisées directement.

Purification de la protéine recombinante : La lyse bactérienne a été optimisée par choc thermique (2 répétitions de 'flash freeze' dans l'azote liquide, suivies par une courte incubation à 37 °C). Les bactéries ont été resuspendues dans du tampon de lyse composé de PBS pH 8, 300 mM NaCl, 1 % triton X-100, 1 mg/mL lysozyme, cocktail d'inhibiteurs de protéases (Roche) et incubées à 4 °C pendant 30 minutes sous agitation. Le lysat bactérien a été soniqué dans la glace pendant 10 minutes effectives (Branson Digital Sonicator model 450), par des impulsions de 10 secondes à 20 % de la puissance maximale du sonicateur, séparées par 20 secondes de pause. L'extrait a été ensuite centrifugé à 17000 g à 4 °C pendant 30 minutes. Le surnageant a été récupéré et incubé pendant 2 heures à 4 °C sous agitation, en présence de 10 mM imidazole et de billes de nickel-agarose (Sigma) préalablement lavées 3 fois avec du PBS pH 8, 300 mM NaCl. À la fin de l'incubation, les billes ont été placées dans une colonne de purification (équilibrée avec du PBS pH 8, 300 mM NaCl). Les billes ont été lavées avec du PBS pH 8, 300 mM NaCl, 0,2 % triton X-100, puis avec du PBS pH 8, 300 mM NaCl. L'élution des protéines recombinantes a été faite avec du PBS pH 8, 300 mM NaCl, 300mM imidazole. Les protéines éluées ont été réduites pendant 20 minutes avec du 10 mM DTT dans la glace, puis transférées dans une colonne de

dessalage PD-10 (Cytiva) et éluées sur le tampon final. Pour les expériences *in vitro*, le tampon final est composé de Tris 50 mM, NaCl 50 mM, pH 8 (tampon méthyltransférase), alors que pour la cristallogenèse le tampon final est composé de PIPES 20 mM, NaCl 250 mM, pH 6,5. Avant chaque utilisation, les protéines ont été centrifugées à 15000 g à 4 °C pendant 15 minutes et la concentration protéique a été déterminée par le Nanodrop.

Purification protéique par chromatographie d'exclusion : Les protéines ont été centrifugées à 15000 g à 4 °C pendant 15 minutes. L'échantillon a été ensuite injecté dans une colonne de chromatographie d'exclusion Superdex 200 16/60 (Cytiva), préalablement équilibrée avec 20 mM PIPES, 250 mM NaCl, pH 6,5. Les fractions qui correspondent à la protéine d'intérêt, une fois identifiées selon le chromatogramme, ont été groupées et concentrées grâce à des concentrateurs Vivaspin 10 MWO (Sartorius), par centrifugation à 3700 g à 4 °C. Les échantillons ont été réduits avec 10 mM de DTT et conservés dans la glace jusqu'à l'utilisation pour le criblage. La pureté de la protéine a été déterminée par SDS-PAGE et coloration au bleu de Coomassie R250.

4) Dosage de l'activité histone méthyltransférase

Dosage par Western Blot (WB) : Le dosage de l'activité histone méthyltransférase de SETD2 a été réalisé dans du tampon Tris 50 mM, NaCl 50 mM, pH 8. Plus précisément, 5 µM de SETD2 recombinante ont été incubés avec 0,5 µg de substrat (Tableau S-2), 150 µM SAM et 1 mM DTT pendant une nuit à température ambiante. La réaction a été arrêtée par l'ajout de tampon Laemmli contenant 400 mM de β-mercaptoéthanol (Sigma) et les échantillons ont été analysés par Western Blot comme décrit ci-dessous.

Dosage par chromatographie liquide en haute pression (HPLC) : L'activité méthyltransférase de SETD2 pour les trois étapes de méthylation du résidu H3K36 a été déterminée par HPLC à l'aide de substrats peptidiques fluorescents, en utilisant la méthode développée et publiée par le laboratoire¹⁸⁴. L'enzyme (5 µM) a été incubée en présence de

75 μ M peptide substrat (peptide de 7 à 9 acides aminés, centré sur le résidu H3K36, Tableau S-2), 100 μ M SAM et 1 mM DTT, dans du tampon méthyltransférase (Tris 50 mM, NaCl 50 mM, pH 8). Pour la caractérisation des paramètres enzymatiques de SETD2 pour son substrat H3K36, nous avons varié la concentration en substrat (0, 25, 50, 100, 200, 400 μ M) et la concentration de SAM était constante (100 μ M). Pour la caractérisation des paramètres enzymatiques du cofacteur SAM, nous avons varié la concentration de ce dernier (0 ; 6,25 ; 12,5 ; 25 ; 50 ; 100 μ M SAM) et la concentration du substrat H3K36 était constante (75 μ M). La réaction enzymatique a été effectuée à température ambiante pendant une nuit et arrêtée le lendemain par l'ajout d'acide perchlorique HClO₄ 15 % v/v (Sigma).

Dix μ L de l'échantillon ont été injectés sur la colonne HPLC (Shimadzu) en condition isocratique (80 % phase A composée d'eau avec 0,1 % HClO₄ et 20 % phase B composée d'acétonitrile avec 0,1 % acide trifluoroacétique), avec un flux de 1 mL/min à 40 °C pendant 27 minutes. Les peptides substrat et produit (peptide substrat méthylé) ont été séparés par une colonne Kromasil (100-5-C18, 4,6 x 250 mm, taille de particule 5 μ m, Phenomenex). À la sortie de la colonne, les peptides ont été détectés par fluorescence (λ_{ex} = 485 nm, λ_{em} = 530 nm). La quantification et l'analyse des données ont été effectuées avec le logiciel LabSolutions (Shimadzu). L'aire des pics correspondants aux peptides substrat et produit a été intégrée et quantifiée, et a permis de déterminer l'activité de l'enzyme.

Substrat	Fournisseur	Dosage
H3.1 recombinant	Actif Motif (31294)	WB
Mono-nucléosomes	Epcypher (16-0009)	WB
tubuline purifiée	Dr. Nicolas Joly (Institut Jacques Monod)	WB
Peptide H3K36	Proteogenix (FAM-TGGVKPHR-NH ₂)	HPLC
Peptide H3K36me1	Proteogenix (FAM-TGGVKmePHR-NH ₂)	HPLC
Peptide H3K36me2	Proteogenix (FAM-TGGVKme ₂ PHR-NH ₂)	HPLC
Peptide H3M36	Proteogenix (TGGVMPHR-NH ₂)	TSA

Tableau S - 2 : Liste des différents substrats utilisés pour le dosage de l'activité HMT de SETD2 recombinante

5) SDS-PAGE et Western Blot

Les protéines ont été séparées par électrophorèse SDS-PAGE (gels tris-glycine, 14 % acrylamide pour SETD2 recombinante seule, 18 % acrylamide pour les dosages H3 recombinant et nucléosomes) en conditions réductrices (25 mM Tris, 190 mM glycine, 0,5 % SDS) à 180 V pendant 60 - 90 minutes. À la fin de la migration, selon l'expérience, le gel a été coloré au bleu de Coomassie ou transféré sur une membrane de nitrocellulose (0,2 µm) à 200 mA pendant 1 heure, en utilisant un tampon de transfert composé de 25 mM Tris, 190 mM glycine, 20 % éthanol. La membrane a été ensuite rincée et colorée au rouge ponceau (Sigma), puis saturée pendant 30 minutes dans du PBST (PBS Tween 0,02 %) – lait 5 %. La membrane a été rincée avec du PBST puis incubée pendant une nuit à 4 °C sous agitation, avec l'anticorps primaire (Tableau S-3) dilué dans du PBST – 1 % lait. La membrane a été lavée 3 fois pendant 5 minutes avec du PBST et incubée pendant 1 heure à température ambiante avec l'anticorps secondaire (Tableau S-3) dilué dans du PBST - 1 % lait. Après 3 nouveaux lavages de 5 minutes avec du PBST, le signal issu de la détection des protéines a été révélé par chimioluminescence, en utilisant le réactif ECL Prime (Cytiva) et l'appareil de détection Amersham Imager 600 (GE Healthcare).

Anticorps	Fournisseur	Référence	Dilution
H3K36me3	Abcam	Ab9050	1/5000
Histag	Sigma	H1029	1/10000
GFP	Santa Cruz	sc-9996	1/10000
SETD2	Thermofisher	PA5-43071	1/5000
γH2A.X (S139)	Thermofisher	CR55T33	1/3000
Igg souris	Sigma	A9044	1/100000
Igg lapin	Santa Cruz	sc-2357	1/100000

Tableau S-3 : Anticorps utilisés pour la détection par Western Blot

6) Auto-méthylation de SETD2 et activité sur tubuline

Afin de mesurer l'activité méthyltransférase de SETD2 sur son substrat α -tubuline, ainsi que sa capacité à s'auto-méthyler, 3 μ M d'enzyme WT ou de mutant L1609P ont été incubés en présence de SAM radioactive ^3H (16,7 kBq, PerkinElmer), 1 mM DTT et 1 μ g α -tubuline (Tableau S-2) dans du tampon de méthyltransférase (Tris 50 mM, 50 mM NaCl, pH 8) à température ambiante pendant une nuit. La réaction a été arrêtée par l'ajout du tampon Laemmli (0,02 % bleu de bromophenol, 4 % SDS, 20 % Glycérol, 120 mM Tris, pH 6,8) contenant 100 mM de β -mercaptoéthanol et les échantillons ont été migrés sur un gel SDS-PAGE 4-12 %, pendant 1 heure à 180 V (Invitrogen) en conditions réductrices. Le gel radioactif a été bloqué dans une solution contenant 50 % éthanol, 5 % acide acétique et 5 % PEG 300 pendant 1 heure sous agitation, puis coloré à l'Instant Blue (Merck) et incubé pendant 1 heure supplémentaire dans une solution permettant de rehausser le signal radioactif (Enlightning™ Rapid Autoradiography Enhancer, PerkinElmer). Le gel a été séché sous vide sur un papier Whatman pendant 80 minutes à 80 °C et exposé sur un film d'autoradiographie (Cytiva) à -80 °C.

7) Décalage thermique (TSA – Thermal Shift Assay)

La stabilité des protéines SETD2 WT et SETD2 L1609P a été mesurée en absence ou en présence du cofacteur SAM et du peptide H3M36 (Apo, SAM, SAM + H3M36), dans une plaque PCR de 384 puits opaques (Roche) et dans du tampon de méthyltransférase (Tris 50 mM, NaCl 50 mM, pH 8) à température ambiante. Dix μ M de SETD2 ont été incubés en présence du colorant SYPRO (5X, Sigma) et, selon la condition, 400 μ M SAM et 1 mM de peptide H3M36. La plaque a été placée dans le LightCycler 480 (Roche). L'émission de fluorescence a été détectée selon les propriétés du SYPRO ($\lambda_{\text{ex}} = 465 \text{ nm}$, $\lambda_{\text{em}} = 580 \text{ nm}$). Le thermocycler a été programmé pour une première incubation à 20 °C pendant 15 secondes,

suivie d'une augmentation graduelle de la température jusqu'à 95 °C (montée de 0,05 °C/ sec et 12 lectures/degré) et d'une incubation finale à 20 °C pendant 15 secondes. La quantification et l'analyse des données ont été effectuées avec le logiciel LightCycler 480 SW 1.5 (Roche). La T_M (température de dénaturation) a été déterminée par la valeur minimale de la dérivée première de l'émission de fluorescence en fonction de la température (dFluo/dT).

8) Activité résiduelle après thermodénaturation à 37 °C

Vingt-cinq μM de SETD2 ont été incubés pendant 0, 5, 10, 15, 20 minutes à 37 °C. À l'issue de chaque incubation, 5 μM de l'enzyme ont été récupérés et incubés avec 75 μM substrat H3K36, 100 μM SAM, 1 mM DTT dans du Tris 50 mM, NaCl 50 mM, pH 8, pendant une nuit à température ambiante. Le lendemain, la réaction a été arrêtée avec de l' HClO_4 et l'activité résiduelle a été mesurée par HPLC comme précédemment décrit.

9) Rélargage de zinc

Dix μM de SETD2 et 100 μM de PAR (4-2-Pyridylazoresorcinol) ont été déposés dans une plaque 96 puits et incubés à 37 °C avec une mesure directe de l'absorbance à 490 nm. La réaction a été faite dans du Tris 50 mM, NaCl 50 mM, pH 8.

10) Culture Cellulaire

Lignées cellulaires utilisées : Pour l'étude *in cellulo*, nous avons utilisé deux lignées cellulaires différentes. Les cellules HEK293T dépourvues du gène *Setd2* (modifiées par la technologie TALEN), ont été offertes par le laboratoire du Dr. W. Kimryn Rathmell (USA). Ces cellules (SETD2 KO) ont été utilisées principalement pour les études de surexpression de SETD2. En collaboration avec le laboratoire du Dr. Michael R. Green (Houston, USA), nous avons pu mettre au point des lignées HEK293T CRISPR SETD2 WT et HEK293T CRISPR SETD2 L1609P. Ces cellules ont été modifiées par la méthode CRISPR/Cas9 de manière à exprimer la forme WT ou L1609P de SETD2.

Génération des lignées HEK293T CRISPR SETD2 WT et HEK293T CRISPR SETD2 L1609P : Les cellules HEK293T (ATCC CRL-3216) ont été électroporées avec la ribonucléoprotéine Cas9-ccRNA-tracrRNA (RNP). Plus précisément, le ccRNA (Integrated DNA Technology) ciblant le gène *Setd2* a été hybridé avec du tracrARN marqué à l'Atto550 (Integrated DNA Technology). Cet hybride a ensuite formé une RNP avec l'Alt-R® S.p. Cas9 Nuclease V3 (Integrated DNA Technology) avant l'électroporation. Cinquante mille cellules ont été mélangées avec 1,5 µM de RNP et 4 µM de matrice donneur oligonucléotide simple brin, suivi d'une électroporation avec le système de transfection Neon™ (ThermoFisher) selon les instructions du fabricant. Les cellules uniques positives à la phycoerythrine (PE) ont été triées 2 jours après transfection. Les colonies ont été multipliées et validées par séquençage Sanger de la région ciblée amplifiée. Toutes les cellules ont été maintenues dans un milieu DMEM (Corning) contenant 10% de sérum de veau fœtal (SVF) et pénicilline/streptomycine (Corning).

ccRNA :

SETD2 L1609P crRNA1: 5'-CATGGCCCTGAAGAATGATGAGG-3'

SETD2 L1609P crRNA2: 5'-GCTTACCTCATCATTCTCAGGG-3'

Matrice Donneur Oligonucléotide Simple Brin :

5' GTCCTAGAATATTGTGGAGAGGTACTIONCGATCATAAAGAGTTTAAAGCTCGAGTGAAGGAGTATG
CACGAAACAAAACATACATTACTATTTTCATGGCACCGAAGAATGATGAAGTAAGCAGCTTGGATAT
GTTTGCCTTGAAACAGATCTGTTAGGTTTGAAGGAGCCATTGATACTGATTTGATTTTTTTTTT 3'

Amorces de Séquencage Sanger :

SETD2 F L1609P : 5' - CCCTACTGCTCCTAGGTAGGAAGT - 3'

SETD2 R L1609P : 5' - AACGTGGTGAAACCCCGT - 3'

Entretien des cellules : Les cellules ont été cultivées en flasque T75 en milieu RPMI complet composé de RPMI 1640 (Gibco) complété avec 10 % de SVF (ThermoFisher), 10 % de Glutamax (Gibco) et placées en incubateur à 37 °C sous 5 % CO₂.

Transfection des cellules HEK293T : Nous avons procédé à la transfection par lipofection (métafectène, Biontex) des cellules HEK293T SETD2 KO avec le plasmide eucaryote pEGFP-SETD2 catalytique WT ou pEGFP-SETD2 catalytique L1609P. Trois cent mille cellules/cm² ont étéensemencées et transfectées avec une solution de transfection, composé d'un ratio de 1 µg plasmide : 2 µL métafectène dans du milieu DMEM. Pour les expériences d'expression protéique nous avons transfecté 2 µg de plasmide et pour les expériences d'expression d'ARNm nous avons transfecté 20 ng de plasmide. Nous avons également procédé à des co-transfections de pEGFP-SETD2 avec le plasmide EGFP-C1 codant mRFP-Ub (Addgene #11935). Pour la co-transfection nous avons transfecté 1 µg de plasmide totale (ratio 1 : 1). Les cellules transfectées sont ensuite placées en culture pendant 48 heures.

Traitement des cellules au MG132 : Les cellules HEK293T SETD2 KO transfectées avec pEGFP-SETD2 ou des cellules HEK293T CRISPR SETD2 WT et HEK293T CRISPR SETD2 L1609P ont été traitées avec 10 µM MG132, 16 heures avant leur récupération. Des cellules traitées uniquement au DMSO ont été utilisées en contrôle.

Analyse de l'expression des protéines SETD2 endogène et GFP-SETD2, et des marques H3K36me3 et γ H2A.X

Lyse cellulaire : Les cellules ont été récupérées et centrifugées à 200 g pendant 5 minutes. Les culots cellulaires ont été lavés avec du PBS et centrifugés à nouveau à 11000 g pendant 1 minute. Les cellules ont été ensuite lysées avec du tampon de lyse cellulaire composé de PBS, 1 % Triton, cocktail d'inhibiteurs protéases et incubées à 4 °C pendant 30 minutes sous rotation. Les lysats ont été soniqués deux fois pendant 2 secondes à 10 % de la puissance du sonicateur et centrifugés à 15000 g pendant 15 minutes à 4 °C. Les surnageants (extraits totaux cellulaires) ont été transférés dans des nouveaux tubes et utilisés pour la détection de SETD2-GFP, alors que les culots ont été utilisés pour l'extraction acide des histones. Les extraits (40 μ g) ont été migrés en conditions dénaturantes dans un gel gradient 4-15 % et transférés dans une membrane de nitrocellulose pour détection avec un anticorps spécifique comme précédemment décrit (Tableau S-3).

Extraction acide des histones : Deux volumes d'HCl 0,2 N ont été ajoutés aux culots obtenus à la fin de la lyse cellulaire. Les échantillons ont été ensuite soniqués 2 fois pendant 2 secondes à 10 % de la puissance du sonicateur et incubés à 4 °C sous rotation pendant une nuit. Le lendemain, les histones ont été centrifugées à 15600 g pendant 20 minutes à 4 °C. Les surnageants ont été récupérés et dosés à l'aide du réactif de Bradford. Pour comparer le niveau de méthylation d'histones entre les cellules, 2 μ g d'histones ont été mélangés avec du Laemmli pH 11, contenant 200 mM de β -mercaptoéthanol, migrés dans un gel SDS-PAGE et transférés sur une membrane de nitrocellulose pour détection par Western Blot comme décrit précédemment.

Fixation de cellules et détection par immunofluorescence : Les cellules HEK293T SETD2 KO co-transfectées avec GFP-SETD2 et mRFP-Ub, ainsi que des cellules HEK293T CRISPR SETD2 WT et HEK293T CRISPR SETD2 L1609P ont été cultivées sur des lames ou lamelles. Les cellules HEK293T SETD2 KO co-transfectées avec GFP-SETD2 et mRFP-Ub ont été fixées avec du paraformaldehyde 4 % pendant 10 minutes à température ambiante et ensuite perméabilisées avec une solution PBS, 0,1 % Triton. Les cellules HEK293T CRISPR SETD2 WT et HEK293T CRISPR SETD2 L1609P ont été fixées avec du méthanol à -20 °C pendant 20 minutes. Les lames ont été ensuite séchées pendant 15 minutes. Les lavages et les

incubations avec les différents anticorps ont été faites avec du PBS. L'incubation des anticorps primaires (anti-SETD2 à une dilution 1/1000 ou anti-H3K36me3 à une dilution 1/3000 (Tableau S-3)) a été faite pendant une nuit à 4 °C. Ces échantillons ont été lavés trois fois le lendemain avec du PBS et les lames ont été incubées avec l'anticorps secondaire anti-lapin (GaR 488 pour SETD2 ou GaR 594 pour H3K36me3) pendant 45 minutes à température ambiante. Toutes les lames ont été montées avec un milieu de montage contenant du DAPI (Fluorshield, Sigma). Les cellules ont été imagées au microscope confocal (Leica SP5, objectif 40, 4 coupes de 0,5 µm), les images ont été enregistrées avec le logiciel (Leica Application Suite Advanced Fluorescence Software), et traitées ensuite avec le logiciel ImageJ (Fiji).

Analyse des ARNm GFP-SETD2 catalytique

Extraction des ARNm : Les cellules ont été lavées 2 fois avec du PBS puis récupérées avec du tri-reagent/trizol et conservées à -80 °C jusqu'au moment d'utilisation. Avant l'extraction, les échantillons ont été incubés pendant 5 minutes à température ambiante. 1/5 volume de trizole a été rajouté aux échantillons qui ont été vortexés pendant 15 secondes et incubés pendant 7 minutes à température ambiante. Les lysats ont été centrifugés à 12000 g pendant 15 minutes à 4 °C et la phase aqueuse des ARN a été récupérée avant d'être complétée avec 1/2 volume d'isopropanol. Les échantillons ont été mélangés, incubés pendant 7 minutes à température ambiante et centrifugés à 12000 g pendant 10 minutes à 4 °C. Les culots des ARNm ont été lavés avec 1/2 volume d'éthanol 75 %, centrifugés à 7500 g pendant 5 minutes à 4 °C et séchés avant d'être resuspendus dans 1/20 volume d'eau pure. Dix µg d'ARNm ont été ré-purifiés avec le kit d'extraction d'ARN (Macherey-Nagel).

Reverse Transcription : Deux µg d'ARNm ont été mélangés avec 1 µg des oligo-dT, chauffés pendant 5 minutes à 70 °C puis incubés dans la glace pendant 5 minutes. Un échantillon composé du tampon spécifique de l'enzyme M-MLV ; 1,25 mM des dNTPs et 20U de M-MLV RT ont été rajoutés aux échantillons des ARNm. Les échantillons ont été incubés dans un

thermocycler (Biorad). Pour la réaction nous avons utilisé un programme de 60 minutes à 42 °C, 5 minutes à 95 °C et une dernière étape de 30 minutes à 20 °C.

Analyse de l'expression des ARNm par qPCR : Vingt ng d'ADNc de chaque échantillon ont été mélangés avec une solution composée de Sybergreen (1X, Roche) et 0,3 µM de chacun des amorces de qPCR, dans une plaque PCR 384 puits opaques (Roche). Une gamme de différentes concentrations d'ADNc provenant d'un mélange de tous les échantillons (40 ; 20 ; 10 ; 5 ; 2,5 ; 1,25 ; 0 ng) a été également effectuée afin de vérifier l'efficacité des amorces utilisés. La plaque a été ensuite analysée grâce à l'appareil LightCycler 480 (Roche). Les échantillons ont été pré-incubés à 95 °C pendant 5 minutes, suivis d'une amplification de 45 cycles (95 °C pendant 20 secondes, 58 °C pendant 20 secondes et 72 °C pendant 20 secondes). Ensuite, la formation des dimères des amorces a été vérifiée par une courbe de dissociation (95 °C pendant 30 secondes, 55 °C pendant 30 secondes et 95 °C d'enregistrement de fluorescence en continu jusqu'au moment d'atteindre les 95 °C). À la fin, les échantillons ont été refroidis pendant 30 secondes à 40 °C. L'analyse des données par quantification relative vis-à-vis d'un gène de ménage (RPL13A) a été effectuée grâce au logiciel LightCycler 480 SW 15.

hSETD2cata qPCR 2F : 5' - GGACAACTGAGGGTTGGGTTT - 3'

hSETD2cata qPCR 2R : 5' - TCTGGGCTTCTTTCCATATCTCT - 3'

RPL13A F : 5' - TGTTTGACGGCATCCCAC - 3'

RPL13A R : 5' - CTGTCAGCTGCCTGGTACTTC - 3'

Analyse par ChIP-qPCR des exons *myc* co-immunoprecipités avec le résidu H3K36me3

Cross-Link des cellules : Huit millions de cellules HEK293T CRISPR SETD2 WT et HEK293T CRISPR SETD2 L1609P, ont étéensemencées et lavées 2 fois avec du PBS. Les cellules ont été ensuite incubées pendant 10 minutes sous agitation à 37 °C avec du PBS complété avec

16 % de formaldéhyde. Le milieu a été ôté et les cellules ont été incubées pendant 5 minutes dans la glace avec une solution PBS Glycine 125 mM (froid). Le milieu a été à nouveau jeté et les cellules ont été récupérées dans du PBS froid complété avec des inhibiteurs de protéases, puis centrifugées à 2600 rpm pendant 15 minutes à 4 °C. Le culot cellulaire a été resuspendu dans du PBS contenant des inhibiteurs de protéases et centrifugé à nouveau pendant 1 minute à 11000 g à 4 °C. Le surnageant a été ôté et les cellules ont été ‘flash freezed’ dans l’azote liquide et conservées à -80 °C jusqu’à utilisation.

Immunoprécipitation de la chromatine (ChIP) par des anticorps anti-H3K36me3 : Pour cette expérience nous avons utilisé le kit True Microchip (Diagenode) selon les conditions d’utilisation du fabricant. L’ADN a été immunoprécipité avec les anticorps α -H3K36me3 et α -H3 (control positif). Un control négatif d’immunoprécipitation avec un anticorps α -lapin (control IgG) a été également effectué. Les immunocomplexes ont été isolés avec des billes de protéine A magnétiques, et l’ADN a été isolé avec les colonnes MicroChIP DiaPure (Diagenode). Nous avons analysé 3 régions exoniques du gène *cmyc*⁸⁴. L’enrichissement des régions génomiques en H3K36me3 (fold enrichment) a été calculé par la comparaison de ‘ChIP H3K36me3’ avec la ‘ChIP control IgG’ et normalisation par la ‘ChIP H3’.

cmyc_exon_1_F : 5’ - GCCGCATCCACGAACTTT - 3’

cmyc_exon_1_R : 5’ - TCCTTGCTCGGGTGTGTAAG- 3’

cmyc_exon_2_F : 5’ - TGCCCCTCAACGTTAGCTTC- 3’

cmyc_exon_2_R : 5’ - GGCTGCACCGAGTCGTAGTC- 3’

cmyc_exon_3_F : 5’ - CCTGAGCAATCACCTATGAACTTG- 3’

cmyc_exon_3_R : 5’ - CAAGGTTGTGAGGTTGCATTTG- 3’

Analyse du phénotype chimiorésistant

Traitement des cellules à l'Étoposide: Les cellules HEK293T CRISPR SETD2 WT et HEK293T CRISPR SETD2 L1609P ont étéensemencées (20000 cellules/cm²) et traitées avec 400 nM étoposide ou DMSO et mises en culture pendant 72 heures.

Analyse des cellules apoptotiques par cytométrie en flux : À la fin du traitement, les cellules lavées et récoltées (digestion par trypsine) ont été centrifugées pendant 5 minutes à 250 g. Le surnageant a été ôté et dans chaque échantillon un mélange réactionnel composé d'APC-Annexine V (550474, BD Biosciences), de PerCP 7AAD (559925, BD Biosciences) et du tampon Annexine V 1X a été ajouté. Les échantillons ont été incubés pendant 15 minutes à l'obscurité à température ambiante avant d'être analysés à l'aide du cytomètre (Aurora, CYTEK). Les données obtenues ont été analysées par le logiciel Flowjo (BD Biosciences).

Analyse du cycle cellulaire par cytométrie en flux : Les cellules ont été traitées comme indiqué précédemment. À la fin du traitement, les cellules ont été fixées ou mises à nouveau en culture dans du milieu RPMI 1640 complété avec 10 % SVF et 10 % Glutamax, pendant 3 ou 5 jours. Pour la fixation, les cellules ont été lavées dans du PBS froid et comptées, lavées à nouveau, puis resuspendues dans du PBS. La fixation a été faite par l'ajout aux échantillons de l'éthanol absolu (froid). Les cellules fixées ont été conservées à -20 °C jusqu'à l'utilisation. Pour procéder au marquage, nous avons lavé les cellules avec du PBS Triton 0,1 % avant de les resuspendre dans un volume adapté pour une concentration finale de $2,5 \cdot 10^6$ cellules/mL. Nous avons ensuite rajouté de la RNase A (0,5 mg/mL, Sigma) et de l'iodure de propidium (50 µg/µL, Invitrogen). Les échantillons ont été incubés pendant 30 minutes à température ambiante à l'obscurité avant d'être analysés à l'aide du cytomètre (BD Accuri C6 Plus). Les données cellulaires ont été analysées par le logiciel Flowjo (BD Biosciences).

Analyse de la marque γ H2A.X (WB): Les cellules ont été traitées comme indiqué. À la fin du traitement, les cellules ont été récupérées ou mises à nouveau en culture dans du milieu RPMI 1640 complété avec 10 % SVF et 10 % Glutamax pendant 3 ou 5 jours. À chaque temps défini, nous avons récupéré les cellules. Nous avons procédé à une analyse protéique de γ H2A.X comme indiqué précédemment.

11) Cristallisation du mutant SETD2 L1609P

Cristallogénèse de SETD2 L1609P : La protéine SETD2 L1609P recombinante a été produite et purifiée comme décrit précédemment (purification d'affinité par Ni-NTA et purification par chromatographie d'exclusion). La mesure d'activité enzymatique par HPLC et la pureté de l'échantillon ont été également vérifiées. Un échantillon contenant 10 mg/mL de SETD2 L1609P ; 3,2 mM SAM ; 1,6 mM peptide H3M36 (ratio molaire entre protéine, peptide, cofacteur à 1 : 5 : 10 respectivement) a été préparé dans un volume final de 250 µL du tampon de cristallographie (PIPES 20 mM, NaCl 250 mM, pH 6,5) et incubé à 4 °C sur la nuit.

Un criblage initial de 7 plaques multipuits (Greiner Bio-One) contenant au total 672 conditions commerciales de cristallogénèse a permis d'identifier les meilleures conditions pour l'obtention de cristaux. Le criblage a été effectué à l'aide du Mosquito™ nanoliter (TPP Labtech), selon la méthode de diffusion de vapeur en goutte assise. En effet, 200 nL d'échantillon ont été mélangés avec un volume équivalent de la solution du réservoir provenant des plaques multipuits. Les plaques ont été stockées à 18 °C dans le système RockImager (Formulatrix), qui permet de suivre la croissance des cristaux par la prise régulière des images des gouttes générées. Les cristaux de SETD2 L1609P ont été obtenus quatre jours plus tard et reproduits à la main, dans une plaque de 24 puits à 18 °C, utilisant la méthode de goutte suspendue. La condition qui a donné les meilleurs cristaux est composée de 32 % w/v PEG 4K, 100 mM Tris, 800 mM chlorure de lithium, pH 8,5 (plaque JBS 1 & 4, puits F12). Les cristaux ont été cryoprotégés à l'aide d'huile et congelés dans l'azote liquide jusqu'au moment de leur diffraction.

Diffraction des cristaux : Les jeux de données de diffraction des Rayons X ont été collectés sur la ligne PROXIMA-1 au synchrotron SOLEIL (St Aubin, France). Les données de diffraction ont été analysées avec le programme XDS et la suite Collaborative Computational Project, Number 4 (CCP4i) ^{186,187}.

Résolution et analyse de la structure : La structure du domaine catalytique de SETD2 L1609P a été résolue par remplacement moléculaire en utilisant comme modèle la structure de SETD2 WT complexée avec le cofacteur SAH et le peptide substrat H3M36 (PDB : 5JJY)

grâce au logiciel Phaser¹⁸⁸. La structure finale à 2,2 Å a été obtenue après ajustement et raffinement par les logiciels respectifs Coot et REFMAC^{189,190}. Les analyses de la structure ont été faites avec le logiciel PyMol (*The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC*) en comparaison avec la structure 5JJY.

Analyse par dynamique moléculaire : Nous avons effectué des simulations de dynamique moléculaire en utilisant les structures SETD2 WT (PDB : 5JJY) et SETD2 L1609P. À partir de ces deux structures, nous avons fait plusieurs simulations en utilisant différents systèmes : SETD2 WT et SETD2 L1609P en forme 'Apo', et SETD2 WT et SETD2 L1609P en complexe avec le peptide substrat H3M36.

Préparation des systèmes

La structure SETD2 WT (PDB : 5JJY) est en complexe avec un peptide H3M36 composé de 14 acides aminés (peptide H3M36 WT, APSTGGVMKPHRYR). La structure du mutant SETD2 L1609P est liée à un peptide plus petit, composé de 11 acides aminés (peptide H3M36 L1609P, APSTGGVMKPH). Pour exclure l'impact éventuel de la différence de taille des peptides pendant les analyses des simulations de dynamique moléculaire, nous avons remplacé le peptide court dans la structure mutante L1609P (peptide H3M36 L1609P) par le peptide plus long (peptide H3M36 WT). Pour ce faire, le peptide court a été retiré de la structure mutante L1609P. Ensuite, la structure mutante L1609P a été superposée de manière optimale à la structure SETD2 WT en utilisant la fonction align du logiciel PyMOL. Le peptide H3M36 WT de la structure SETD2 WT a été copié dans la structure du mutant SETD2 L1609P.

Les structures SETD2 WT et L1609P comportent une région non observable dans les cristaux de 12 et 9 acides aminés respectivement. Nous avons d'abord modélisé ces régions à l'aide du logiciel Modeller 10.1¹⁹¹. Pour ce faire, les molécules d'eau et le cofacteur (SAH ou SAM respectivement) ont été retirés du fichier PDB et les trois cations Zn²⁺ ont été conservés. Les régions manquantes ont été modélisées dans les structures de SETD2 WT et L1609P en l'absence ou en présence du peptide long. Un total de 100 modèles de chaque condition a

été généré. Pour chaque système, le modèle présentant la meilleure fonction objective a été choisi.

Simulations de dynamique moléculaire

Les analyses ont été réalisées avec les structures SETD2 WT et L1609P constituées du domaine catalytique sous leur forme 'Apo' ou en complexe avec le peptide H3M36 WT. Les simulations de dynamique moléculaire ont été exécutées avec le logiciel Gromacs (version 2018.3) ¹⁹² et le champ de force AMBER-99SB ¹⁹³. Le système a été solvatoé avec une boîte d'eau cubique (modèle d'eau TIP3P), avec une distance minimale entre tout atome et les parois de la boîte de 10 Å. Pour neutraliser électriquement le système, des contre ions Cl⁻ et Na⁺ ont été ajoutés dans chaque système. Toutes les interactions non liantes ont été calculées explicitement jusqu'à une distance de 10 Å. La méthode Ewald à maille de particules (PME) a été utilisée pour les interactions électrostatiques à longue distance. Toutes les liaisons ont été contraintes en utilisant l'algorithme LINCS, permettant un pas de temps de 2 fsec entre deux configurations.

Initialement, les systèmes ont été minimisés énergétiquement, puis équilibrés pendant 2 nsec à 1 bar et 300 K. Enfin, une étape de production de 150 ou 500 nsec a été exécutée. Pour les structures SETD2 WT et SETD2 L1609P 'Apo', deux simulations de 150 nsec et deux simulations de 500 nsec ont été calculées, alors que pour leur forme complexée avec le peptide H3M36 WT, 1 simulation de 150 nsec et 2 simulations de 500 nsec ont été calculées.

La comparaison des différentes simulations des deux systèmes (forme sauvage et forme mutante) a été réalisée à l'aide de deux critères : le RMSD (Root Mean Square Deviation) et le RMSF (Root Mean Square Fluctuation).

Le RMSD est une mesure de distance structurale entre deux conformations différentes après superposition optimale des deux conformations. Le RMSD est calculé sur les carbones C α des résidus selon la formule suivante :

$$RMSD(t, 0) = \sqrt{\frac{1}{n} \sum_{i=1}^n \|r_i(t) - r_i(0)\|^2}$$

Où n est le nombre de carbones $C\alpha$ dans la protéine, $r_i(t)$ le vecteur position de l'atome i au temps t et $r_i(0)$ le vecteur position de l'atome i au temps 0 .

Nous avons calculé le RMSD prise toutes les picosecondes le long de la trajectoire et la conformation de référence prise au temps au temps 0 . Les RMSD ont été calculés pour les différentes trajectoires obtenues à partir de la structure sauvage et mutantes.

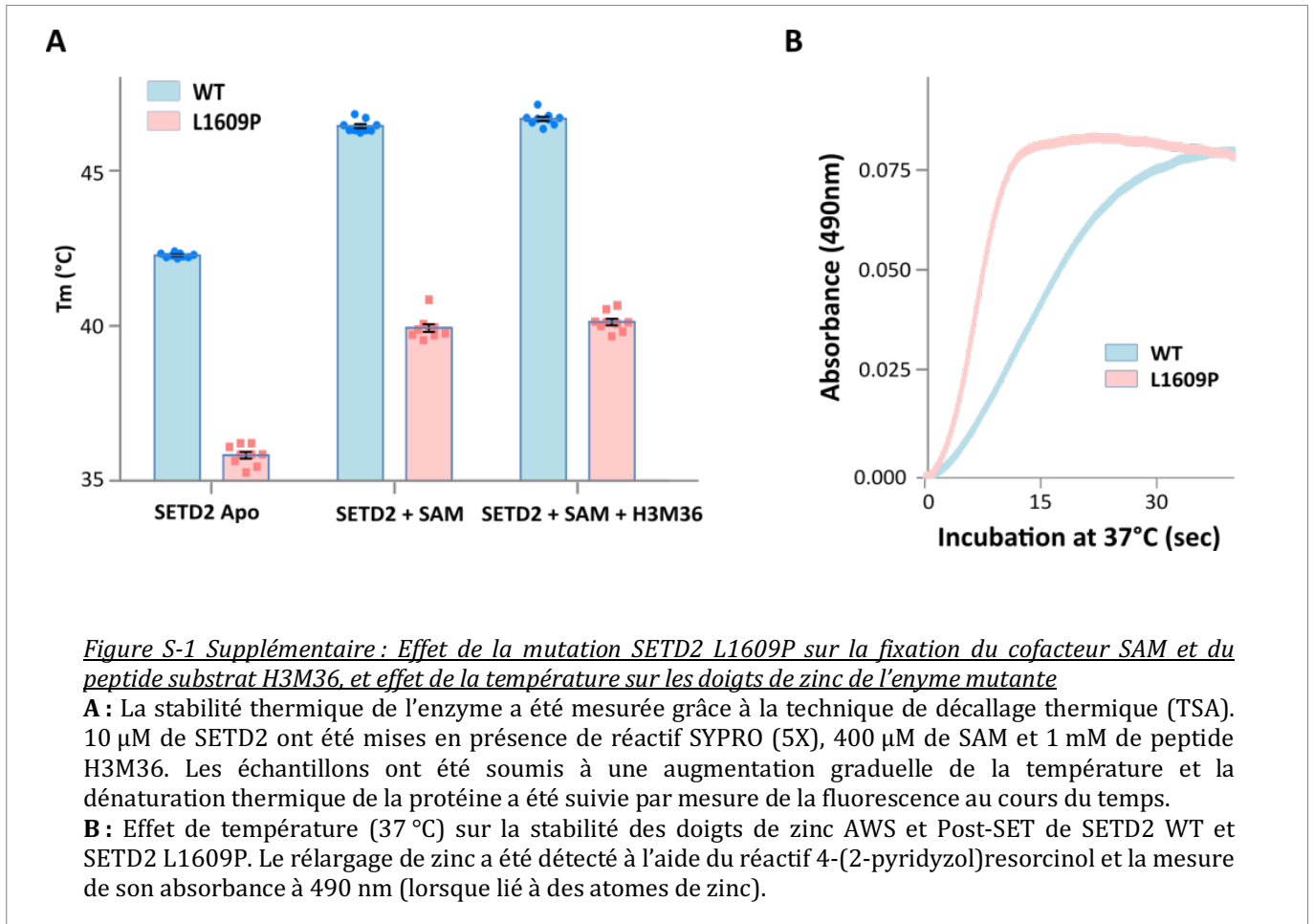
Le RMSF est une mesure des fluctuations atomiques de la protéine au cours du temps. Le RMSF des résidus au cours du temps a été calculée sur les carbones $C\alpha$ de la protéine selon la formule suivante :

$$RMSF_i = \sqrt{\frac{1}{nsteps} \sum_{i=1}^{nstep} \|r_i(t) - \langle r_i \rangle\|^2}$$

avec $nsteps$ le nombre de conformations au cours du temps, $r_i(t)$ le vecteur position de l'atome i au temps t et $\langle r_i \rangle$ le vecteur position de la position atomique moyenne calculée le long de la trajectoire.

Les RMSF par résidus ont été calculés pour chacune des trajectoires obtenues à partir de la structure sauvage et mutantes 'apo' en extrayant une conformation toutes les picosecondes. Pour chaque système (structure sauvage et structure mutante), la moyenne des RMSF par résidus a été calculée à partir des trois trajectoires correspondantes.

IV. Annexes



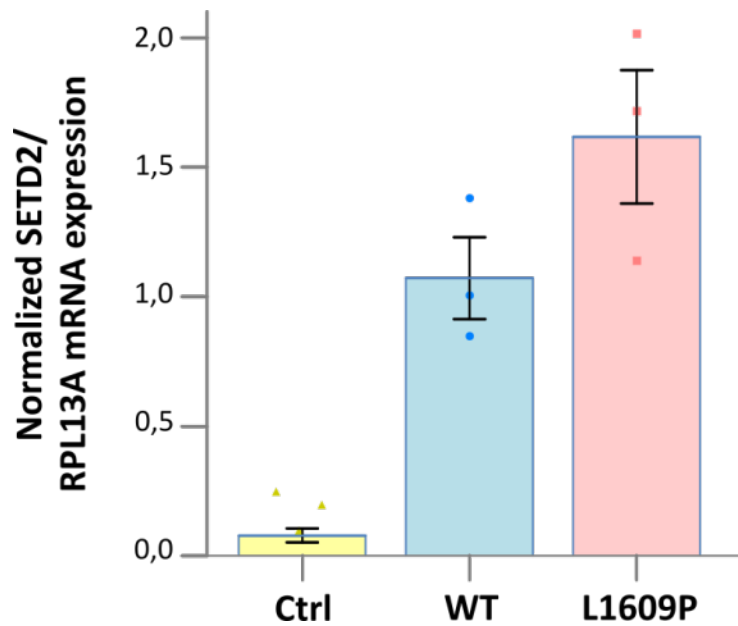


Figure S-2 Supplémentaire : Expression de l'ARNm SETD2 WT et L1609P après transfection dans les cellules HEK293T SETD2 KO

Les cellules HEK293T dépourvues de SETD2 (SETD2 KO) ont été transfectées avec 20 ng de SETD2 WT ou SETD2 L1609P. Des cellules non transfectées ont été utilisées comme contrôle. L'ARNm a été purifié et puis transcrit en ADNc. Un contrôle sans reverse transcription a été fait. Le niveau d'expression de l'ARNm SETD2 a été déterminé par qPCR.

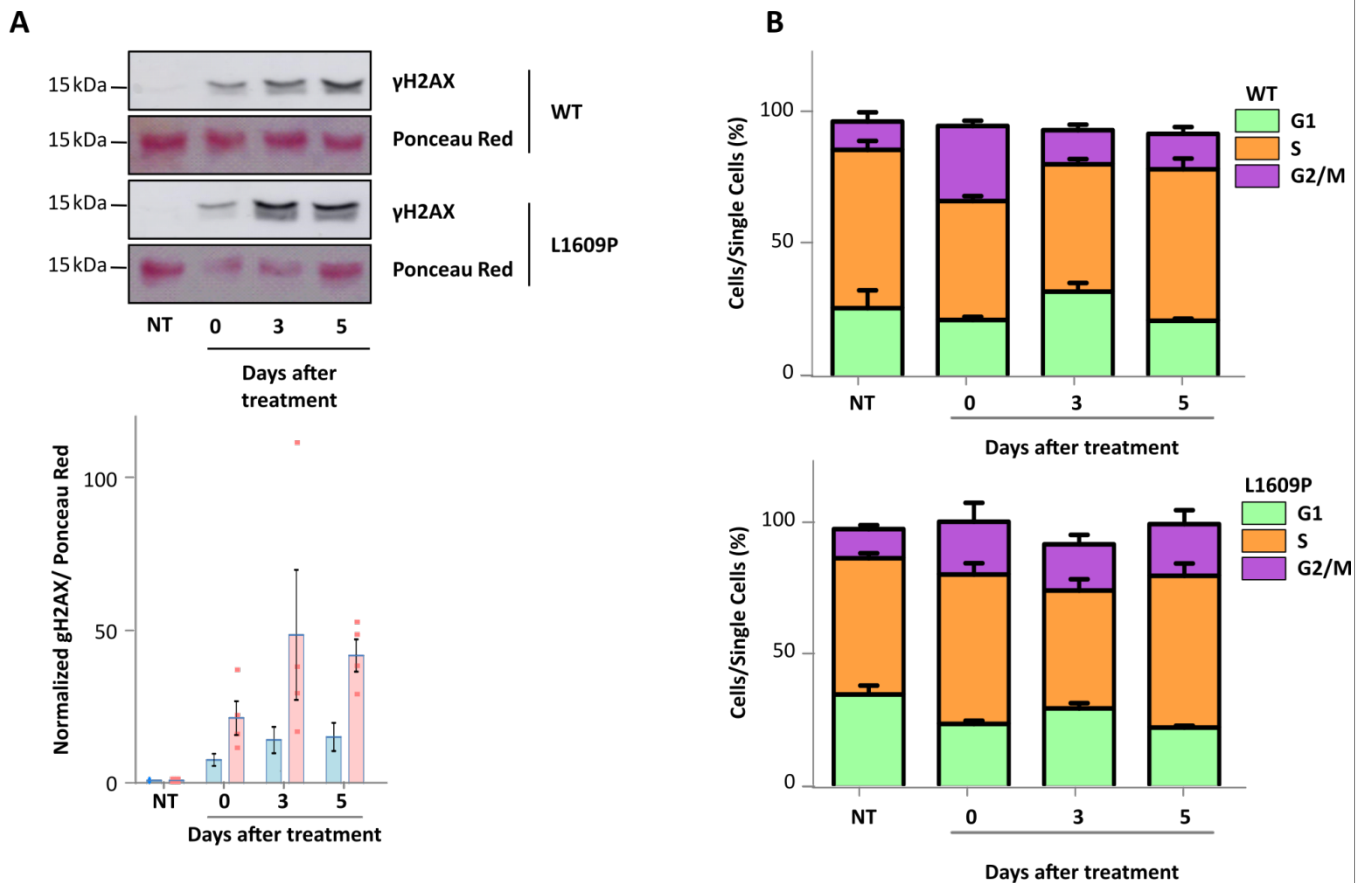


Figure S-3 Supplémentaire : Effet du traitement à l'étoposide sur le cycle cellulaire des cellules HEK293T CRISPR SETD2 WT et HEK293T CRISPR SETD2 L1609P

Les cellules ont été traitées pendant 72 heures avec 400 nM d'étoposide et puis remises en culture dans du milieu frais.

A : Les extraits protéiques ont été analysés et le niveau de la marque γ H2AX des deux lignées a été détecté par Western Blot.

B : Les cellules ont été fixées à l'éthanol et marquées à l'iodure de propidium. Les différentes phases du cycle cellulaire ont été analysées par cytométrie en flux.

**CARACTERISATION STRUCTURALE
ET FONCTIONNELLE D'UN
MUTANT ONCOGENIQUE DE MOZ**

I. Introduction

MOZ (MYST3/KAT6A) est une histone acétyltransférase, membre de la famille des HAT à domaine MYST, capable d'acétyler des résidus lysine de l'histone H3 (H3K14, H3K23) et de p53 (p53K120, p53K382). MOZ est impliquée dans des maladies neurodégénératives (Syndrome KAT6A) et des cancers. Selon les données extraites de la base COSMIC (mutations observées dans les cancers), 61 % des mutations identifiées dans le domaine catalytique (MYST) de MOZ sont des mutations faux-sens. Le Tableau M-1 représente les mutations faux sens de MOZ le plus souvent observées. Parmi elles, la mutation T605P est particulièrement pertinente pour étudier l'activité et le mécanisme catalytique de l'enzyme. En effet, cette dernière se situe dans une boucle importante pour le fonctionnement de l'enzyme, à côté du résidu catalytique K604. Cette lysine est décrite comme capable d'être auto-acétylée et d'avoir ainsi un rôle sur l'activité enzymatique ¹⁴⁷.

Position	Mutation	Nombre de mutations
664	D664N	5
504	P504L	2
535	P535S	2
605	T605P	2
614	L614F	2
682	P682L	2
740	R748H	2
748	R748C	2

Tableau M - 1 : Mutations faux sens les plus représentées dans le domaine MYST de MOZ
Les données sont issues de la base de données COSMIC.

T605P est une mutation somatique de MOZ, identifiée dans le cas de patients atteints de carcinome colorectal ^{194,195}. Notre travail a consisté à caractériser cette mutation d'un point de vue biochimique et structural. Dans un premier temps, nous avons focalisé l'étude de l'effet de la mutation T605P sur l'activité de MOZ vis-à-vis de substrats connus. Dans un second temps, nous avons procédé à la caractérisation structurale de cette mutation qui nous a permis de mieux comprendre l'effet de la mutation sur l'activité enzymatique.

II. Résultats

1) Effet de la mutation T605P sur l'activité de MOZ

La mutation T605P se situe sur la boucle $\alpha 2\beta 7$ (L599-P612) du domaine catalytique MYST. Cette région est connue pour son implication dans l'auto-acétylation des MYST et pour son importance dans la régulation de l'activité enzymatique¹⁴⁷. Chez l'Homme, les résidus MOZ L599-P612 sont majoritairement conservés entre les membres de la famille MYST. Parmi eux, le résidu T605 de MOZ est juxtaposé à la lysine décrite comme auto-acétylée (K604). Cette lysine est strictement conservée dans la famille des MYST (Fig. M-1A).

Le domaine catalytique MYST du mutant MOZ T605P s'exprime aussi bien sous sa forme recombinante (Fig. M-1B) que le domaine catalytique de l'enzyme sauvage (WT). Des études de thermostabilité par approche TSA suggèrent que les formes sauvage et mutante présentent deux populations en solution. En présence du cofacteur AcétylCoA, une population est observée (Fig. M-1B, M-1C). Suite à la fixation du cofacteur AcétylCoA, le mutant MOZ T605P peut également s'auto-acétyler (Fig. M-1D) avec une moindre efficacité que celle du WT. Cependant, l'activité lysine acétyltransférase du mutant T605P de MOZ sur H3 et p53 est fortement perturbée. Le mutant MOZ T605P, *in vitro*, n'est pas actif vis-à-vis de peptides substrats histone (peptides centrés sur les lysines cibles H3K14 et H3K23) (Fig. M-1E, M-1F), de l'histone H3 et de core histones (Fig. M-1G, M-1H), ainsi que son substrat non histone p53 (Fig. M-1I).

L'effet sur l'activité d'histone acétyltransférase du mutant MOZ T605P est également confirmé *in cellulo* grâce à l'utilisation de cellules HEK293T exprimant de manière constitutive la forme sauvage ou mutante T605P de MOZ (modifiées par approche CRISPR/Cas9). Comme attendu, le niveau d'acétylation des marques H3K14ac et H3K23ac, ainsi que le niveau d'acétylation global des histones cellulaires, sont diminués dans le cas des cellules exprimant la forme mutante T605P (Fig. M-1J).

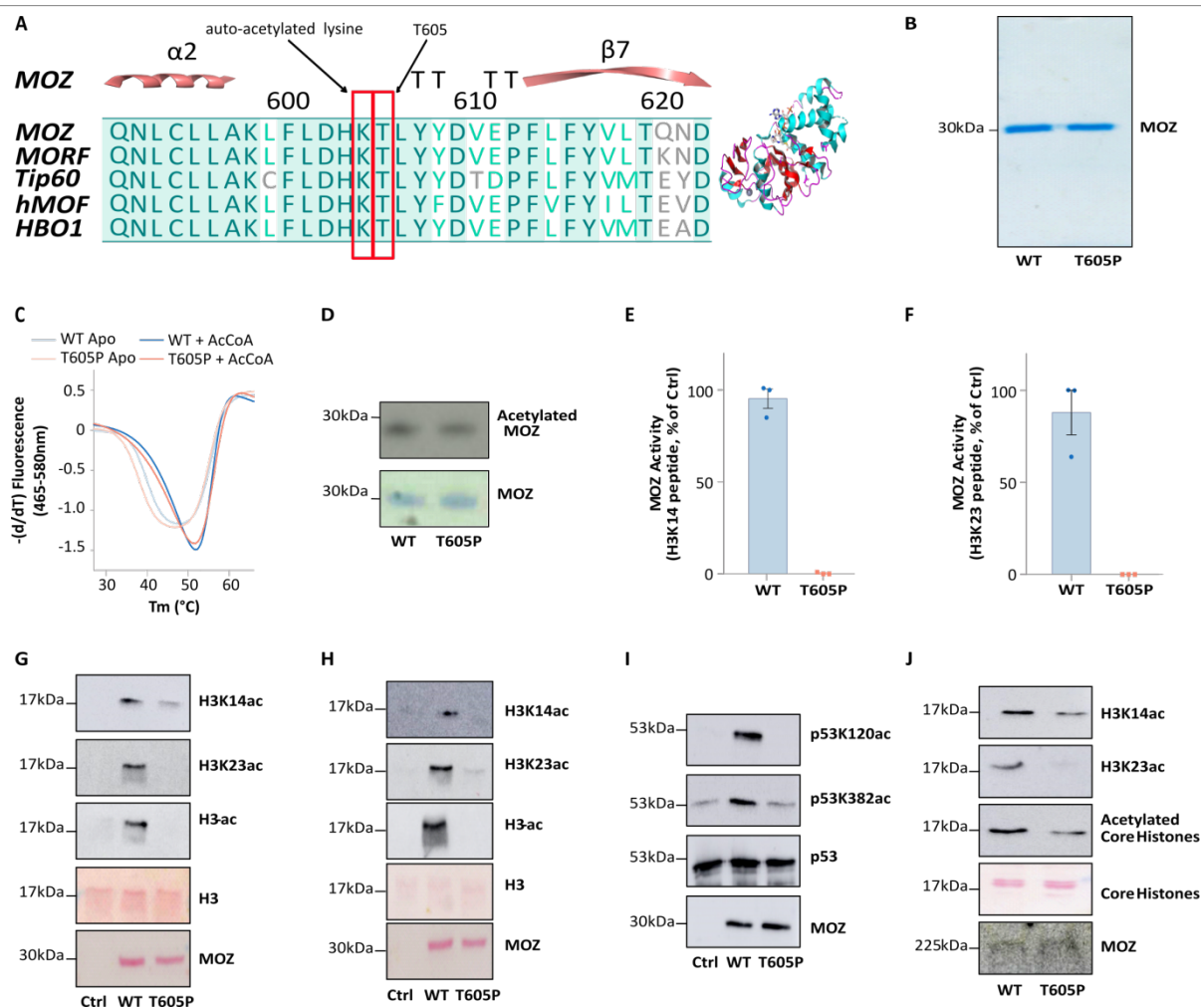


Figure M-1 : Alignement des séquences des MYST, expression de MOZ WT et MOZ T605P recombinants, détermination de leur capacité de liaison à l'AcétylCoA et de leur activité lysine acétyltransférase

A : Alignement des séquences du domaine MYST des membres de la famille MYST chez l'Homme (MOZ, MORF, Tip60, hMOF et HBO1) par le serveur Clustal Ω, suivi d'une analyse par le logiciel ESPript 3¹⁸⁵.

B : Séparation des protéines MOZ sauvage (WT) et MOZ T605P recombinantes, purifiées par chromatographie d'affinité (Ni-NTA) suivie par chromatographie d'exclusion. 1 µg de protéines a été séparé par SDS-PAGE, suivi d'une coloration au bleu de Coomassie R-250.

C : La capacité de l'enzyme à se lier à son cofacteur a été mesurée grâce à la technique de décalage thermique (TSA). 10 µM de MOZ ont été mis en présence de réactif SYPRO avec ou sans 400 µM d'AcétylCoA. Les échantillons ont été soumis à une augmentation graduelle de la température. La dénaturation thermique de la protéine a été suivie par mesure de la fluorescence au cours du temps.

D : Dosage de l'activité d'auto-acétylation de MOZ en présence d'AcétylCoA tritié. 2 µM de MOZ ont été incubés en présence de 3,7 kBq AcétylCoA radioactif pendant 4 heures à 30 °C. Les échantillons ont été séparés par SDS-PAGE. Le gel a été séché pendant 90 minutes et exposé à un film radiosensible à -80 °C.

E, F : Dosage de l'activité histone acétyltransférase de MOZ. 2 µM de MOZ (WT ou T605P) ont été incubés pendant 4 heures à 30 °C en présence de 100 µM AcétylCoA, 1 mM DTT, 10 mM sodium butyrate et 100 µM de peptide substrat H3K14 (**E**), ou H3K23 (**F**).

G, H, I : Dosage de l'activité acétyltransférase de MOZ pour ses substrats H3 recombinant (**G**), core histones purifiées (**H**) et p53 recombinant (**I**). 2 µM de MOZ ont été incubés pendant 4 heures à 30 °C en présence de 100 µM AcétylCoA, 1 mM DTT, 10 mM sodium butyrate et 1 µg de substrat (histone H3, core histones, p53).

J : Détection du niveau d'acétylation des marques H3K14ac et H3K23ac, et du niveau d'acétylation global des histones issues des cellules HEK293T modifiées par CRISPR/Cas9 afin d'exprimer la forme WT ou T605P de MOZ. 2 µg de core histones purifiées de ces cellules ont été séparées par SDS-PAGE et le niveau d'acétylation a été détecté par Western Blot à l'aide d'anticorps spécifiques.

2) Effet de la mutation T605P sur la structure de MOZ

Le domaine MYST de MOZ T605P (504-778) a été exprimé en bactéries et purifié pour des expériences de cristallisation. Nous avons résolu la structure cristallographique du domaine MYST de MOZ T605P (2.9 Å) en présence du cofacteur CoA (Tableau M-2, Fig. M-1 Supp.). La comparaison de cette structure avec la structure MOZ WT de référence (PDB : 2OZU) a permis de mettre en évidence des différences entre les régions qui sont différentes, notamment au voisinage du site de mutation (Fig. M-2A). Les deux structures sont néanmoins très similaires et portent les caractéristiques communes du domaine MYST, dont le doigt de zinc et le site de fixation du cofacteur (Fig. M-2B, M-2C), qui se situent dans les deux cas dans une poche à la surface protéique (Fig. M-2D, M-2E).

Les résidus catalytiques, ainsi que la lysine auto-acétylée, se trouvent positionnés de manière identique sur la structure WT (Fig. M-1F, panel de droite) et la structure mutante T605P (Fig. M-1F, panel du centre), à l'exception de la cystéine catalytique (C646). L'orientation différente de la chaîne latérale de cette cystéine n'est pas due à la mutation. En effet, la comparaison de ces deux structures avec une troisième structure de MOZ WT (PDB : 6BA4) nous a permis d'observer la présence de deux conformères de la cystéine C646. Dans la structure MOZ T605P, cette cystéine est positionnée de la même manière que celle du conformère, issue de la structure WT 6BA4 (Fig. M-1F, panel de droite).

La structure de MOZ T605P nous a également permis de confirmer la capacité du mutant MOZ T605P à s'auto-acétyler sur la lysine 604 positionnée dans le site actif de l'enzyme. Sous sa forme acétylée, le résidu K604 forme des interactions de Van der Waals et des liaisons hydrogène avec des résidus en proximité (F613, F615 et Y631, S633 respectivement), permettant la stabilisation de la boucle $\alpha 2\beta 7$ où se situe le résidu K604 (Fig. M-2A Supp.). La force de ces liaisons est très probablement perturbée dans le cas de MOZ T605P car les résidus impliqués dans ces interactions (F613, F615, Y631, S633) sont plus éloignés du résidu K604ac (Fig. M-2B Supp.).

L'étude cristallographique du mutant MOZ T605P confirme également que le résidu muté (605) se situe dans le site actif de l'enzyme, en proximité directe des résidus catalytiques

Wavelength	
Resolution range (Å)	44.18 - 3.0 (3.107 - 3.0)
Space group	I 4 2 2
Unit cell (Å, °)	110.1 110.1 148.11, 90 90 90
Total reflections	250892 (22995)
Unique reflections	9451 (912)
Multiplicity	26.5 (25.2)
Completeness (%)	99.89 (100.00)
Mean I/sigma(I)	26.90 (1.16)
Wilson B-factor (Å ²)	136.15
R-merge	0.07684 (2.349)
R-meas	0.07842 (2.397)
R-pim	0.01534 (0.4716)
CC1/2	0.999 (0.565)
CC*	1 (0.85)
Reflections used in refinement	9448 (912)
Reflections used for R-free	471 (46)
R-work	0.2297 (0.4057)
R-free	0.2521 (0.4391)
CC(work)	0.963 (0.555)
CC(free)	0.972 (0.462)
Number of non-hydrogen atoms	2326
Macromolecules	2276
Ligands	50
Solvent	0
Protein residues	273
RMS(bonds, Å)	0.010
RMS(angles, °)	1.35
Ramachandran favored (%)	91.04
Ramachandran allowed (%)	8.58
Ramachandran outliers (%)	0.37
Rotamer outliers (%)	1.17
Clashscore	19.40
Average B-factor (Å)	148.21
Macromolecules (Å)	148.87
Ligands (Å)	118.48
Number of TLS groups	8

Tableau M – 2 : Statistiques de collecte et d'affinement des données de la structure cristallographique du domaine catalytique de MOZ T605P

Les statistiques pour l'enveloppe de plus haute résolution sont indiquées entre parenthèses.

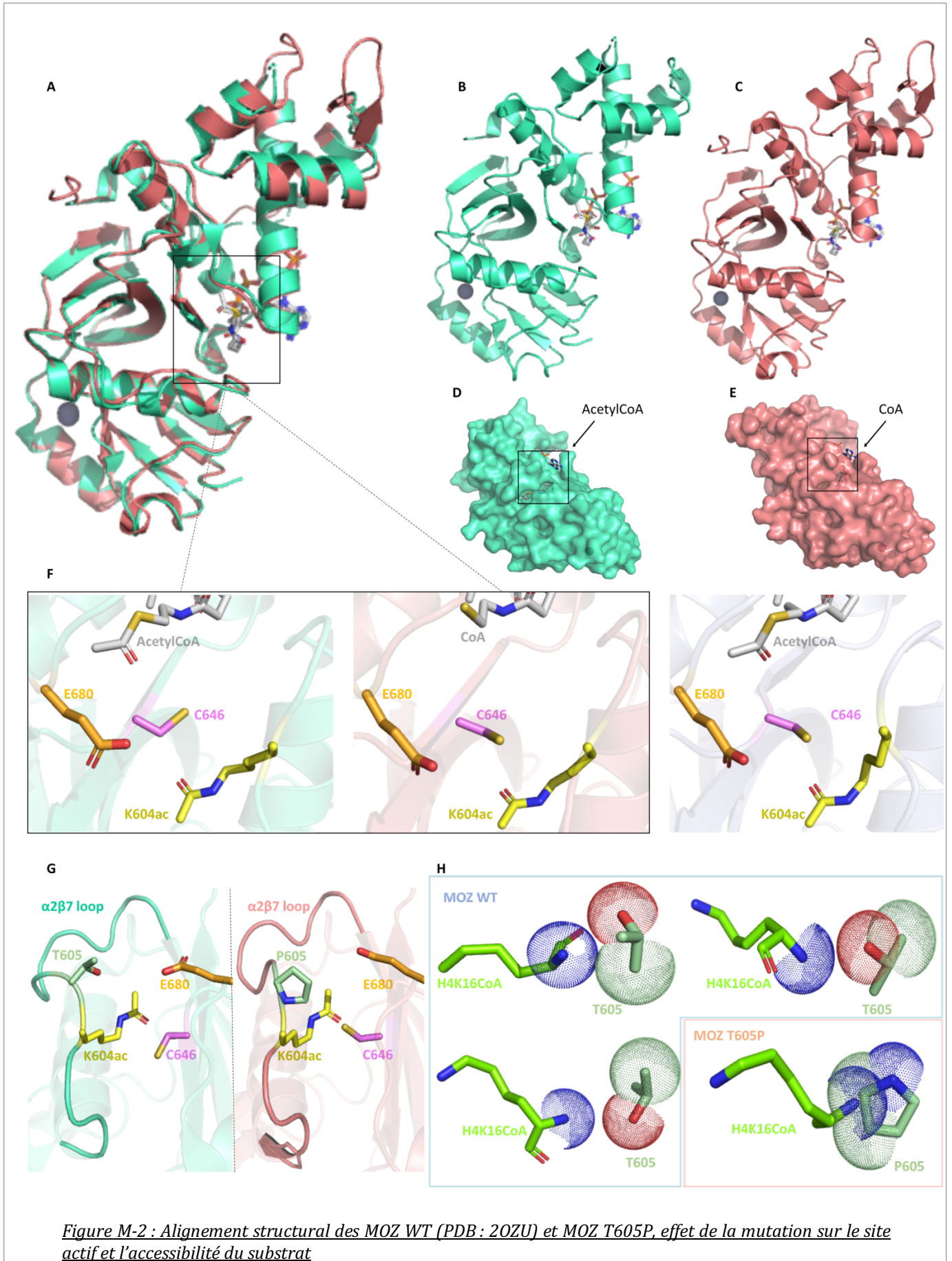


Figure M-2 : Alignement structural des MOZ WT (PDB : 2OZU) et MOZ T605P, effet de la mutation sur le site actif et l'accessibilité du substrat

A : Alignement des structures du domaine MYST de MOZ WT (PDB : 2OZU, 'cartoon', coloré en greencyan) et MOZ T605P ('cartoon', coloré en deepsalmon). Les deux enzymes sont cocrystallisées avec le cofacteur AcétylCoA (WT) ou CoA (T605P), représenté en gris CPK. L'analyse structurale a été effectuée grâce au logiciel PyMol.

B, C : Représentation schématique de la forme WT (**B**) et T605P (**C**) de MOZ en présence du cofacteur AcétylCoA. L'atome de zinc du doigt de zinc est représenté en sphère, coloré en gris.

D, E : Représentation en surface de la forme WT (**D**) et T605P (**E**) de MOZ, et de la poche de fixation du cofacteur dans chacune des structures.

F : Comparaison des résidus du site actif de MOZ WT (PDB 2OZU à gauche et PDB 6BA4 à droite) et MOZ T605P (au centre). Tous les résidus sont représentés en stick, CPK. La lysine auto-acétylée (K604ac) est colorée en jaune, la cystéine catalytique (C646) en violet et le glutamate catalytique (E680) en brightorange. Le cofacteur AcétylCoA ou CoA est également représenté en gris CPK.

G : Représentation schématique de la boucle $\alpha 2\beta 7$ du domaine MYST de MOZ (WT PDB 2OZU à gauche en greencyan, T605P à gauche en deepsalmon) où se situent la lysine auto-acétylée (stick, jaune CPK) et le résidu T605/ P605 (stick, palegreen CPK), proches de la dyade catalytique C646 (stick, violet CPK) et E680 (stick, brightorange CPK).

H : Représentation des interactions Van der Waals entre le résidu 605 (stick, palegreen, CPK) et le substrat H4K16CoA (stick, vert CPK, issu de la structure de ESA1 WT 3TO6). Le panel en haut à gauche montre le résidu T605 issu de la structure de MOZ WT (PDB 2OZU), le panel en haut à droite le résidu T605 issu de la structure MOZ WT (PDB 2RC4), le panel en bas à gauche le résidu T605 de la structure MOZ WT (PDB 6BA4) et le panel en bas à droite le résidu P605 de la structure MOZ T605P.

C646 et E680. La mutation entraîne un changement de la conformation de la boucle $\alpha 2\beta 7$ qui, au niveau de la mutation, se trouve plus en proximité du site actif en comparaison avec la structure 2OZU (Fig. M-2G). L'alignement des structures MOZ WT (PDB : 2OZU, 2RC4, 6BA4) et MOZ T605P avec ESA1 (analogue de MOZ chez la levure), cocrystallisées avec H4K16CoA (PDB : 3TO6), a permis d'identifier le site de fixation de la lysine substrat de MOZ, qui se trouve entre les deux résidus catalytiques (C646, E680) (Fig. M-3 Supp.). Dans le cas des structures 2OZU (Fig. M-2H, panel haut gauche), 2RC4 (Fig. M-2H, panel haut droite) et 6BA4 (Fig. M-2H, panel bas gauche), la chaîne latérale de la T605 n'a pas d'incidence sur le positionnement du substrat H4K16CoA. En revanche, dans le cas du mutant MOZ T605P, la chaîne latérale de la proline 605 entraîne un encombrement stérique avec la chaîne principale du substrat lysine (Fig. M-2H, panel bas droite). De plus, le tunnel d'accessibilité au solvant dans le site actif est modifié au niveau de la mutation (Fig. M-2C Supp.). Dans le cas de la mutation T605P, la chaîne latérale de la proline 605 occupe une partie du tunnel (Fig. M-2D Supp., panel du centre), entraînant ainsi une déformation de ce dernier, de manière à s'adapter à la nouvelle conformation du site actif (Fig. M-2D Supp., panel de gauche).

III. Discussion

MOZ est une protéine complexe, souvent impliquée dans les translocations génomiques avec d'autres histone acétyltransférases, notamment CBP, ¹⁴² mais aussi dans des syndromes neurodégénératifs, dont le syndrome KAT6A ¹⁶⁹. Ces anomalies génétiques sont principalement dues à des mutations ou des points de rupture situés dans les domaines MYST, ED et SM de MOZ. Parmi ces domaines, le domaine MYST confère à MOZ son activité de lysine acétyltransférase ¹⁵². Or, l'effet fonctionnel des mutations dans ce domaine n'est pas décrit. Ce travail a permis de caractériser une mutation oncogénique de MOZ, située dans le domaine MYST, au site actif de l'enzyme. Nous avons résolu pour la première fois une structure cristallographique d'un mutant pathologique de MOZ à 2.9 Å et nous avons démontré que le changement structural a un effet sur l'activité de l'enzyme, due à un remodelage structural du site actif.

Par des expériences de biochimie, nous avons démontré que le mutant MOZ T605P est un mutant 'perte de fonction', étant toutefois capable de lier son cofacteur AcétylCoA et de s'auto-acétyler. Ce résultat suggère que le mécanisme d'auto-acétylation est différent de celui permettant l'acétylation des substrats. Nous avons également confirmé l'effet de la 'perte de fonction' de la mutation MOZ T605P par l'utilisation de cellules HEK293T obtenues avec la technologie de CRISPR/Cas9 de manière à exprimer constitutivement la forme mutante de MOZ. Ces cellules présentent un défaut d'acétylation globale des histones, et notamment un défaut d'acétylation des marques H3K14ac et H3K23ac, connues pour être des marques dépendantes de MOZ.

La structure de MOZ T605P montre que la mutation T605P entraîne un réarrangement du site actif (Fig. M-3). Plus précisément, la boucle $\alpha 2\beta 7$ sur laquelle sont positionnés les résidus 604 (lysine auto-acétylée) et 605 (mutation) s'oriente vers le site de liaison du substrat et la dyade catalytique cystéine 646 – glutamate 680 (identifiés par homologie structurale avec l'orthologue de MOZ chez la levure, ESA1 ¹⁴⁶). Ce changement conformationnel de la boucle $\alpha 2\beta 7$ perturbe vraisemblablement la fixation de la lysine cible

des substrats protéiques. La chaîne latérale de la proline 605 empêcherait par encombrement stérique la fixation de la lysine cible.

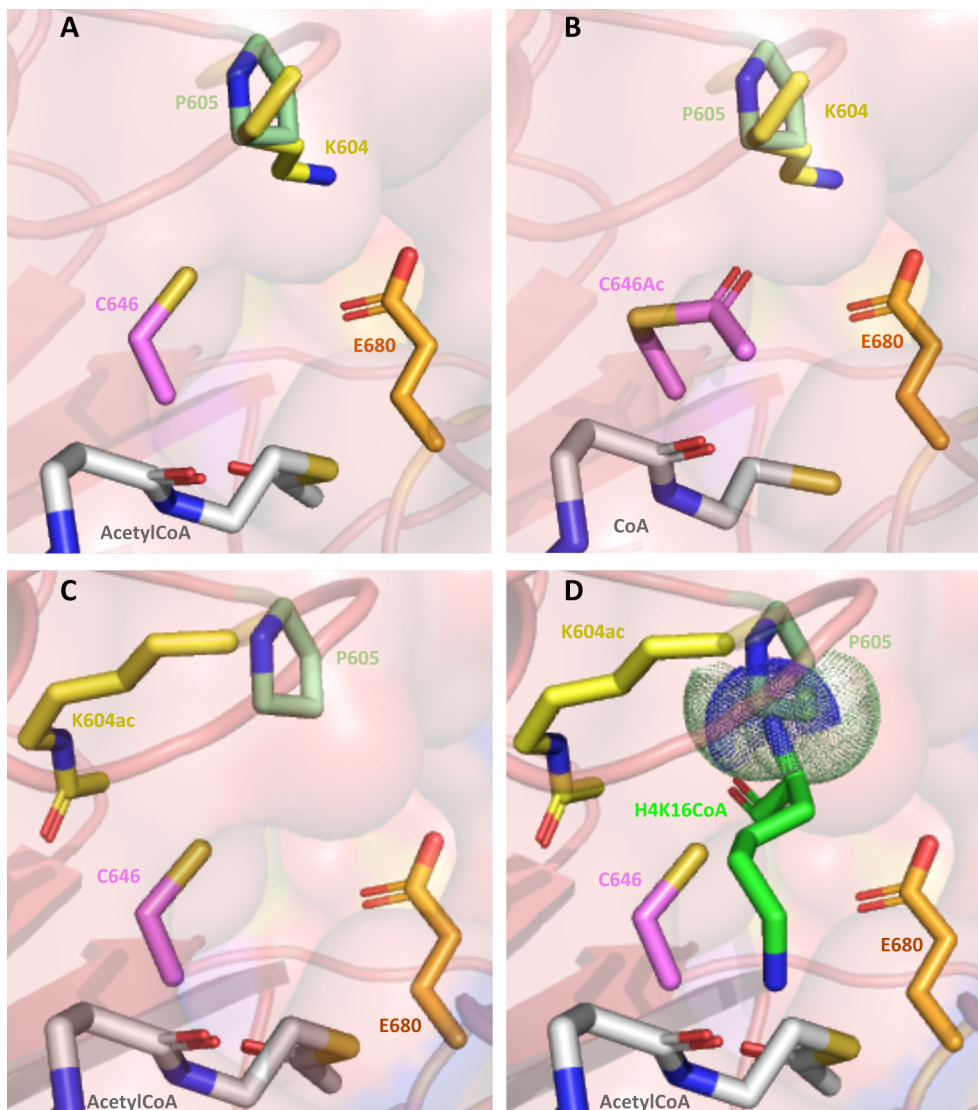


Figure M-3 : Proposition d'un modèle simplifié de la "perte de fonction" du mutant MOZ T605P

(Les structures PDB 6BA4, 3TOA, 3TO6, 1MJA et MOZ T605P sont utilisées pour la proposition de ce modèle)

A : Le site actif de l'enzyme est constitué des résidus K604, C646 et E680 (représentés en stick, colorés respectivement en jaune, violet et brightorange). La lysine 604 est non acétylée et se dirige vers la dyade catalytique composée des résidus C464, E680. Le cofacteur AcétylCoA est également présent.

B : Le résidu E680 active le résidu C646 qui, par attaque nucléophile sur le groupement acétyl du cofacteur de l'AcétylCoa, devient acétylé et libère une molécule de CoA.

C : Le résidu K604, activé par le résidu E680, s'acétyle à son tour par attaque nucléophile sur le résidu C646Ac. La chaîne latérale s'éloigne de la dyade catalytique des résidus C646-E680 et cette fois c'est le site de la mutation P605 qui est proche de ces deux résidus (stick, CPK, palegreen). Une molécule d'AcétylCoA est à nouveau fixée proche du site actif.

D : L'accès du substrat H4K16CoA est bloqué par encombrement stérique avec le résidu P605.

Néanmoins, la possibilité que le substrat protéique puisse être fixé avec une très faible affinité ne peut pas être exclue. La proline 605 se situe sur une boucle décrite comme mobile ($\alpha 2\beta 7$) et importante pour la fixation du substrat. Plus précisément, selon le niveau d'acétylation de la lysine 604, cette boucle adopte une conformation différente, qui permet de réguler la fixation du substrat et l'activité de l'enzyme ¹⁴⁷. La flexibilité de la boucle $\alpha 2\beta 7$ du mutant MOZ T605P nécessite d'être élucidée par de futures études de dynamique moléculaire, ce qui permettra de conclure sur l'espace accessible au solvant et sur la capacité de MOZ T605P à lier la lysine substrat. Des expériences de spectrométrie de masse native vont également être conduites afin d'étudier la capacité du mutant MOZ T605P à former un complexe avec l'histone H3.

De plus, des essais de cristallisation de MOZ T605P en complexe avec l'analogue de substrat LysCoA devraient permettre de mieux comprendre l'impact du changement conformationnel de la boucle $\alpha 2\beta 7$ sur la fixation du substrat. En effet, l'analyse structurale a été effectuée à l'aide de la structure d'ESA1 (orthologue de MOZ) chez la levure, cocrystallisée avec un peptide H4K16CoA ¹⁴⁷. L'alignement des deux structures a permis d'identifier le site de fixation du substrat. Malgré cette homologie, des particularités de fixation du substrat protéique sur MOZ, et notamment par le mutant MOZ T605P, ne peuvent pas être mises en évidence avec une telle analyse. Par ailleurs, nous avons essayé de cristalliser le mutant T605P de MOZ avec le cofacteur AcétylCoA et le peptide substrat p53K382 qui n'est pas présent dans la structure cristallographique. Cela est probablement dû à la faible affinité de MOZ pour ce substrat peptidique. De manière intéressante, à ce jour, aucune structure de MYST avec un peptide substrat ou une protéine substrat n'a été résolue.

Le changement conformationnel, engendré au site actif de l'enzyme suite à la mutation et à l'encombrement stérique avec la lysine cible, expliquerait l'effet délétère de la mutation sur l'activité de l'enzyme. On observe une faible activité résiduelle de l'enzyme mutant *in vitro*, en particulier sur le résidu H3K14, suggérant que, peut-être, ce motif est plus facilement acétylé que le motif H3K23. Ce résultat coïncide avec le fait que la boucle qui entraîne l'inhibition de la fixation du substrat pourrait être flexible et qu'elle pourrait adopter des

conformations qui permettent de manière transitoire la fixation du substrat protéique. Les analyses de dynamique moléculaire permettront également de mieux comprendre ce point.

L'utilisation des cellules HEK293T, modifiées par technologie CRISPR/Cas9 et exprimant soit la forme WT ou la forme mutante T605P de MOZ, a permis d'ores et déjà de confirmer le phénotype de 'perte d'activité' acétyltransférase de ce mutant. Le faible niveau d'acétylation du résidu H3K14 dans le cas des cellules MOZ T605P est probablement dû au faible niveau d'acétylation par le mutant T605P de MOZ, comme observé *in vitro*. Cela peut être également dû à d'autres enzymes car l'activité de MOZ est redondante avec celle d'autres HAT, dont l'homologue de MOZ, MORF¹⁹⁶. Ces cellules vont être utilisées pour de futures expériences de CHIP-Seq et RNA-Seq qui permettront de définir l'impact de la mutation T605P de MOZ sur les régions génomiques affectées par la perte des marques H3K23ac et H3K14ac et sur le transcriptome des cellules mutées.

Ces travaux ont permis également de mettre en évidence des effets de la mutation MOZ T605P sur la physiologie des cellules. MOZ est décrite comme une enzyme impliquée dans la régulation du cycle cellulaire et, plus précisément, dans l'arrêt de ce dernier. Selon le contexte, MOZ participe à l'inhibition ou à l'induction de la sénescence des cellules^{162,164}. Des résultats préliminaires montrent que les cellules portant la mutation MOZ T605P adoptent un état senescent après traitement avec l'étoposide, un médicament génotoxique (Fig. M-4 Supp.), contrairement aux cellules MOZ WT. Ce résultat nécessite d'être complété par des expériences de cytométrie en flux de manière à étudier le cycle cellulaire. Cela permettra de vérifier que le cycle de ces cellules est bien arrêté mais aussi d'identifier la phase du cycle dans laquelle les cellules HEK293T MOZ T605P sont bloquées. Des marqueurs de sénescence et de la phase du cycle où les cellules sont bloquées, seront également étudiés.

IV. Matériels et méthodes

1) Alignement des séquences des MYST humaines

Les séquences des membres de la famille des MYST chez l'Homme (MOZ, MORF, Tip60, hMOF, HBO1) ont été obtenues à partir de la base de données Uniprot/Swiss-prot (respectivement Q92794, Q8WYB, Q92993, Q9H7Z6, Q95251). L'alignement multiple a été fait par Clustal Ω (EMBL-EBI) et analysé par ESPrpt 3¹⁸⁵.

2) Mutagenèse dirigée, construction du mutant MOZ T605P

Le vecteur d'expression bactérienne de type pet28a-LIC, contenant l'insert d'ADN codant le domaine MYST de la protéine humaine MOZ/MYST3/KAT6A (résidus 497-780), a été obtenu auprès de la société Addgene (#25181, plasmide pet28a-MOZ WT). La mutation MOZ T605P (Thr en Pro) a été introduite par l'utilisation du kit Agilent QuickChange II XL. L'insertion de la mutation a été effectuée par PCR, en utilisant les amorces complémentaires suivantes :

MOZ F T605P : 5' - GTTGTTTCTTGACCACAAA **C**CCCTCTATTACGATGTGG – 3'

MOZ R T605P : 5' - CCACATCGTAATAGAGGG **G**TTTGTGGTCAAGAAACAAC - 3'

Le plasmide pet28a-LIC-MOZ T605P a ensuite été sélectionné par digestion de l'ADN matrice par l'enzyme de restriction *Dpn1* (inclus dans le kit Agilent). Le plasmide non digéré a été utilisé pour transformer des bactéries compétentes *DH5α* qui ont été sélectionnées sur des boîtes LB (Luria Bertani) agar contenant 30 µg/µl de kanamycine. La présence de la mutation a été vérifiée par séquençage (Eurofins).

De manière à exprimer la forme sauvage et mutée de MOZ, les plasmides pet28a-LIC-WT et pet28a-LIC-T605P ont été transformés par choc thermique à 42°C dans des bactéries compétentes *E.Coli BL21 (DE3)*. Les bactéries avec le plasmide d'expression pet28a-LIC-WT ou pet28a-LIC-T605P ont été sélectionnées sur des boîtes LB agar contenant 30 µg/µl de kanamycine.

3) Expression et purification de MOZ recombinante

Pré-culture bactérienne : Les bactéries *E.Coli BL21 (DE3)/pet28a-LIC-MOZ WT (497-780)* et *E.Coli BL21 (DE3)/pet28a-LIC-MOZ T605P (497-780)* ont étéensemencées dans du milieu LB (Sigma) en présence de 30 µg/µL kanamycine (Euromedex) sur la nuit sous agitation (150 rpm) à 37 °C.

Culture bactérienne et induction de l'expression protéique: Du milieu LB contenant de la kanamycine (30 µg/µl) a étéensemencé avec la pré-culture bactérienne et placé à 37 °C sous agitation (150 rpm). Lorsque l'absorbance du milieu à 595 nm atteint 0,6-0,7 unités d'absorbance, l'expression protéique a été induite par l'ajout de 250 µM d'IPTG (Euromedex). À la suite, les bactéries ont été placées pour une nuit à 16 °C sous agitation (150 rpm) et récoltées le lendemain par centrifugation à 4400 rpm pendant 20 minutes. Les bactéries ont été lavées avec du PBS et centrifugées à nouveau à 4400 rpm pendant 15 minutes. Les bactéries ont été stockées à -80 °C ou utilisées directement.

Purification de la protéine recombinante : De manière à optimiser la lyse protéique, les bactéries ont été fragilisées par choc thermique (2 répétitions de 'flash freeze' dans l'azote liquide, suivies par une courte incubation à 37 °C). Les bactéries ont été resuspendues dans du tampon de lyse composé de PBS pH 8, 300 mM NaCl, 1 % triton X-100, 1 mg/mL lysozyme, et d'un cocktail d'inhibiteurs des protéases (Roche) et incubées à 4 °C pendant 30 minutes sous agitation. Le lysat bactérien a été soniqué dans la glace pendant 10 minutes effectives (Branson Digital Sonicator model 450) par des impulsions de 10 secondes à 20 % de la puissance maximale du sonicateur, séparées par 20 secondes de pause. L'extrait a été ensuite centrifugé à 17000 g à 4 °C pendant 30 minutes. Le surnageant a été récupéré et incubé pendant 2 heures à 4 °C sous agitation, en présence de 10 mM d'imidazole et des billes de nickel-agarose (Sigma) préalablement lavées 3 fois avec du PBS pH 8, 300 mM NaCl. À la fin de l'incubation, les billes ont été placées dans une colonne de purification (lavée avec du PBS pH 8, 300 mM NaCl). Après deux lavages successifs des billes avec du PBS pH 8, 300 mM NaCl, 0.2 % triton X-100 dans un premier temps et du PBS pH 8, 300 mM NaCl, 20mM imidazole dans un second temps, les protéines recombinantes ont été éluées avec du PBS pH 8, 300 mM NaCl, 300mM imidazole. Les protéines éluées ont été réduites pendant 20 minutes avec 10 mM DTT dans la glace, transférées dans une colonne de dessalage PD-10 (Cytiva) et éluées dans un tampon final selon l'utilisation de la protéine. Pour les expériences *in vitro*, le tampon final est composé de Tris 40 mM, 100 mM NaCl, 1 mM DTT, pH 8 (tampon acétyltransférase), alors que, pour la cristallogénèse le tampon final est composé de 20 mM sodium citrate (27883, VWR), 100 mM NaCl, pH 5,5. La concentration des échantillons a été mesurée par le test de Bradford.

Clivage de l'étiquette 6-histidine (uniquement pour la cristallogénèse): La protéine obtenue après l'étape de purification par nickel-agarose a été incubée avec la thrombine (T6884, Sigma) pendant une nuit à 15 °C sous agitation (10U de thrombine/ mg de protéine). Le lendemain, l'échantillon a été incubé avec des billes de benzamidine-agarose (A7155, Sigma) pendant 30 minutes à 4 °C sous agitation (1 mL de résine/ 8mg de thrombine utilisée) et centrifugé à 15000 g, à 4 °C pendant 15 minutes. Le surnageant a été récupéré et conservé dans la glace jusqu'au moment de l'utilisation. Le clivage de l'étiquette 6-histidine a été vérifié par Western Blot, avec un anticorps anti-Histag (Tableau M-5).

Purification protéique par chromatographie d'exclusion : La protéine a été centrifugée à 15000 g à 4 °C pendant 15 minutes. Le surnageant a été récupéré et l'échantillon a été ensuite injecté dans une colonne de chromatographie d'exclusion Superdex 200 16/60 (Cytiva), préalablement équilibrée avec le tampon de la protéine (tampon acétyltransférase ou tampon de cristallographie). Les fractions qui correspondent à la protéine d'intérêt, une fois identifiées par migration SDS-PAGE et coloration au bleu de Coomassie R, ont été groupées et concentrées grâce à des concentrateurs Vivaspin 10 MWO (Sartorius) par centrifugation à 4400 rpm à 4 °C. Pour l'utilisation en cristallographie, les échantillons ont été réduits avec 10 mM de DTT et conservés dans la glace jusqu'à l'utilisation pour criblage. Pour l'utilisation en biochimie, les protéines ont été aliquotées et conservées à -80 °C. Avant chaque utilisation, elles ont été centrifugées à 15000 g à 4 °C pendant 15 minutes et la concentration des échantillons a été mesurée par le test de Bradford.

4) Auto-acétylation de MOZ – Autoradiographie

Afin de mesurer l'activité d'auto-acétylation de MOZ, 2 µM d'enzyme WT ou de mutant T605P ont été incubés en présence d'AcétylCoA radioactif ³H (3,7 kBq, PerkinElmer) et 10 mM de sodium butyrate dans du tampon acétyltransférase (Tris 40 mM, 100 mM NaCl, 1 mM DTT, pH 8) à 30 °C pendant 4 heures. La réaction a été arrêtée par l'ajout de tampon Laemmli (0,02 % bleu de bromophenol, 4 % SDS, 20 % Glycérol, 120 mM Tris, pH 6,8) contenant 100 mM β-mercaptoéthanol et les échantillons ont été migrés sur un gel SDS-PAGE 4-12 %, pendant 1 heure à 180 V (Invitrogen) en conditions réductrices. Le gel radioactif a été bloqué dans une solution contenant 50 % éthanol, 5 % acide acétique et 5 % PEG 300 pendant 1 heure sous agitation, puis coloré par Instant Blue (Merck) et incubé pendant 1 heure supplémentaire dans une solution permettant de rehausser la signal radioactif (Enlightning™ Rapid Autoradiography Enhancer, PerkinElmer). Le gel a été séché sous vide sur un papier Whatman pendant 80 minutes à 80 °C et exposé sur un film d'autoradiographie (Cytiva) à -80 °C.

5) Dosage de l'activité acétyltransférase de MOZ

Dosage par Western Blot (WB) : Le dosage d'activité lysine acétyltransférase de MOZ a été réalisé dans du tampon Tris 40 mM, 100 mM NaCl, 1 mM DTT, pH 8. Plus précisément, 2 μ M de MOZ recombinante ont été incubés avec 1 μ g de substrat (selon le Tableau M-3), 100 μ M d'AcétylCoA et 10 mM de sodium butyrate pendant 4 heures à 30 °C. La réaction a été arrêtée par l'ajout de tampon Laemmli contenant 400 mM de β -mercaptoéthanol (Sigma) et les échantillons ont été analysés par Western Blot comme décrit ci-dessous.

Dosage par chromatographie liquide en haute pression (HPLC) : L'activité d'histone acétyltransférase de MOZ a été déterminée par HPLC à l'aide de substrats peptidiques fluorescents ¹⁸⁴. L'enzyme (2 μ M) a été incubée en présence de 100 μ M peptide substrat (peptide de 7-9 acides aminés, dont la séquence correspond à celle du substrat histone de MOZ, Tableau M-3), 100 μ M AcétylCoA et 10 mM sodium butyrate dans le tampon d'acétyltransférase, Tris 40 mM, 100 mM NaCl, 1 mM DTT, pH 8. La réaction enzymatique a été effectuée à 30 °C pendant 4 heures et a été ensuite arrêtée par l'ajout d'acide perchlorique HClO₄ 15 % v/v (Sigma). Dix μ L de l'échantillon ont été injectés en HPLC (Shimadzu) en condition isocratique (Tableau M-4). Les peptides substrat et produit (peptide substrat acétylé) ont été séparés par la colonne Kromasil (100-5-C18, 4,6 x 250 mm, taille de particule 5 μ m, Phenomenex). À la sortie de la colonne, les peptides ont été détectés par fluorescence (λ_{ex} = 485 nm, λ_{em} = 530 nm). La quantification et l'analyse des données ont été effectuées avec le logiciel LabSolutions (Shimadzu). L'aire des pics correspondants aux peptides substrat et produit a été intégrée et quantifiée, et a permis de déterminer l'activité de l'enzyme.

Substrat	Fournisseur	Dosage
H3.1 recombinant	Actif Motif (31294)	WB
Core Histones	Purifiées au laboratoire (voir la suite)	WB
p53 recombinant	Purifié au laboratoire	WB
Peptide H3K14	Proteogenix (FAM-STGGKAPRR-NH ₂)	HPLC
Peptide H3K23	Proteogenix (FAM-QLATKAARR-NH ₂)	HPLC

Tableau M – 3 : Les différents substrats utilisés pour le dosage de l'activité HAT de MOZ recombinante

Peptide	Phase Mobile	Temps de run (minutes)	Temperature (°C)	Flux (mL/min)
H3K14	83 % Phase A, 17 % Phase B	25	40	1
H3K23	25 % Phase A, 75 % Phase B	30	40	1

Tableau M – 4 : Conditions d'HPLC selon le peptide substrat utilisé

Phase A = eau LC-MS avec 0,1% HClO₄, Phase B = acétonitrile LC-MS avec 0,1 % acide trifluoroacétique TFA (Biosolve)

6) SDS-PAGE et Western Blot

Les protéines ont été séparées par électrophorèse SDS-PAGE (gels tris-glycine, 14 % acrylamide pour MOZ et p53 recombinantes, 18 % acrylamide pour H3 recombinant et core histones) en conditions réductrices (25 mM Tris, 190 mM glycine, 0,5 % SDS) à 180 V pendant 60 - 90 minutes. À la fin de la migration, selon l'expérience, le gel a été coloré au bleu de Coomassie ou transféré sur une membrane de nitrocellulose (0,2 µm) à 200 mA

pendant 1 heure, en utilisant un tampon de transfert composé de 25 mM Tris, 190 mM glycine, 20 % éthanol. La membrane a ensuite été rincée et colorée au rouge Ponceau (Sigma), puis saturée pendant 30 minutes dans du PBST (PBS Tween 0,02 %) - lait 5 %. La membrane a été rincée avec du PBST et incubée pendant une nuit à 4 °C sous agitation, avec l'anticorps primaire (Tableau M-5) dilué dans du PBST - 1 % lait. La membrane a été lavée avec du PBST, puis incubée pendant 1 heure à température ambiante avec l'anticorps secondaire (Tableau M-5) dilué dans du PBST - 1 % lait. Finalement, la membrane a été lavée avec du PBST et le signal, issu de la détection des protéines a été révélé par chimioluminescence, en utilisant le réactif ECL Prime (Cytiva) et l'appareil de détection Amersham Imager 600 (GE Healthcare).

Anticorps	Fournisseur	Référence	Dilution
H3K14Ac	Abcam	ab52946	1/2500
H3K23Ac	Diagenode	C15410344	1/2500
panKAc	Cell Signaling	9441S	1/5000
p53K120Ac	Millipore	ABE286	1/5000
p53K382Ac	Cell Signaling	2525S	1/2500
p53	Sigma	P6749	1/10000
MOZ	Santa Cruz	sc-293282	1/2000
Histag	Sigma	H1029	1/10000
Igg souris	Sigma	A9044	1/100000
Igg lapin	Santa Cruz	sc-2357	1/100000

Tableau M – 5 : Anticorps utilisés pour la détection par Western Blot

7) Décalage thermique (TSA – Thermal Shift Assay)

La stabilité des protéines MOZ WT et MOZ T605P a été mesurée en absence ou en présence du cofacteur AcétylCoA (Apo ou AcCoA). Dans une plaque PCR de 384 puits opaques (Roche), 10 μ M de MOZ ont été incubés en présence de 10X du colorant SYPRO (Sigma) et, selon la condition, 400 μ M AcétylCoA dans du tampon d'acétyltransférase (40 mM Tris, 100 mM NaCl, 1 mM DTT, pH 8) à température ambiante. La plaque a été centrifugée à 800 rpm pendant 10 secondes et placée dans le LightCycler 480 (Roche). L'émission de fluorescence a été détectée selon les propriétés du SYPRO (λ_{ex} = 465 nm, λ_{em} 580 nm). Le thermocycler a été programmé pour une première incubation à 20 °C pendant 15 secondes, suivie d'une augmentation graduelle de température jusqu'à 95 °C (montée de 0,05 °C/ sec et 12 lectures/degré) et d'une incubation finale à 20 °C pendant 15 secondes. La quantification et l'analyse des données ont été effectuées avec le logiciel LightCycler 480 SW 1.5 (Roche). La T_M a été déterminée par la valeur minimale de la dérivée première de l'émission de fluorescence en fonction de la température (dFluo/dT).

8) Culture Cellulaire

Génération des lignées HEK293T MOZ WT et HEK293T MOZ T605P : Les cellules HEK293T (ATCC CRL-3216) ont été électroporées avec la ribonucléoprotéine Cas9-ccRNA-tracrRNA (RNP). Plus précisément, le ccRNA (Integrated DNA Technology) ciblant le gène de *KAT6A*, a été hybridé avec du tracrARN marqué à l'Atto550 (Integrated DNA Technology). Cet hybride a ensuite formé une RNP avec l'Alt-R® S.p. Cas9 Nuclease V3 (Integrated DNA Technology) avant l'électroporation. Cinquante mille cellules ont été mélangées avec 1,5 μ M de RNP et 4 μ M de matrice donneur oligonucléotide simple brin, suivi d'une électroporation avec le système de transfection Neon™ (ThermoFisher) selon les instructions du fabricant. Les cellules uniques positives à la phycoérythrine (PE) ont été triées 2 jours après la

transfection. Les colonies ont été multipliées et validées par séquençage Sanger de la région ciblée amplifiée. Toutes les cellules ont été maintenues dans un milieu DMEM (Corning) avec 10% de sérum de veau fœtal (SVF) et pénicilline/streptomycine (Corning).

ccRNA :

KAT6A T605P crRNA1: 5'-CATCGTAATAGAGGGTTTTGTGG-3'

KAT6A T605P crRNA2: 5'-TGGCTCCACATCGTAATAGAGGG-3'

Matrice Donneur Oligonucléotide Simple Brin :

5' - GACTGTTTTGTTCTGTTTATTTCTCTAGGTTGATGGGAATGTGAGTACCATTTATTGTCAAACCT
GTGTCTTTTGGCAAAGTTGTTTCTTGATCACAACCTCTCTATTACGATGTAGAGCCATTTCTTTTTTA
TGTAACACAGAATGATGTCAAGGGCTGCCACCTTGTGGCTACTTTTCTAAGGTAAGTAT – 3'

Amorces de Séquençage Sanger :

MOZ F T605P : 5' - GTCTGGACTGTGGATATCTAGAGATGCT – 3'

MOZ R T605P : 5' - CTAGCACCTGGTAAGTTTCAAATGT – 3'

Entretien des cellules : Les cellules ont été cultivées en flasque T75 en milieu RPMI complet composé de RPMI 1640 (Gibco) complémenté avec 10 % de SVF (ThermoFisher), 10 % de Glutamax (Gibco) et placées en incubateur à 37 °C sous 5 % CO₂.

Lyse cellulaire: Six millions de cellules ont été récupérées et centrifugées à 200 g pendant 5 minutes. Les culots cellulaires ont été lavés avec du PBS et centrifugés à nouveau à 11000 g pendant 1 minute. Les cellules ont été ensuite lysées avec du tampon de lyse cellulaire composé de PBS, 1 % Triton, 10 mM sodium butyrate, cocktail inhibiteurs protéases et incubées à 4 °C pendant 30 minutes sous rotation. Les lysats ont été soniqués deux fois pendant 2 secondes à 10 % de la puissance du sonicateur et centrifugés

à 15000 g pendant 15 minutes à 4 °C. Les surnageants (extraits totaux cellulaires) ont été transférés dans des nouveaux tubes et traités à la suite par l'acide trichloracétique et les culots ont été utilisés pour l'extraction acide des histones.

Précipitation à l'acide trichloracétique (TCA) : La fraction soluble, issue de la lyse cellulaire, a été incubée avec de l'acide trichloracétique 100 % w/v (Sigma) à 4 °C pendant 10 minutes (4 volumes de lysat cellulaire : 1 volume de TCA). L'échantillon a été centrifugé à 14000 g à 4 °C pendant 5 minutes. Après avoir ôté le surnageant, le culot a été lavé avec de l'acétone froide, suivi d'une centrifugation de 14000 g à 4 °C pendant 5 minutes. Le culot a été lavé une deuxième fois avec de l'acétone (CarloEbra) froide et centrifugé comme la première fois. À la fin du deuxième lavage, le culot a été séché par chauffage à 95 °C pendant 10 minutes et dissous dans du Laemmli pH 11 contenant du β -mercaptoéthanol avant migration SDS-PAGE et détection de MOZ par Western Blot. Dans ce cas, un gel tris-glycine gradient 4-15 % a été utilisé pour la migration des échantillons en conditions réductrices, comme décrit précédemment. Après migration, les protéines ont été transférées sur membrane de nitrocellulose (0,2 μ m) à 250 mA à 4 °C pendant 2,5 heures, en utilisant un tampon de transfert, composé de 25 mM Tris, 190 mM glycine, 10 % éthanol, 0,09 % SDS. La coloration au rougeau Ponceau et la saturation de la membrane avec du lait, ainsi que la détection ont été effectuées, comme décrit précédemment.

Extraction acide des histones : Deux volumes d'HCl 0,2 N ont été ajoutés aux culots obtenus à la fin de la lyse cellulaire. Les échantillons ont été ensuite soniqués 2 fois pendant 2 secondes à 10 % de la puissance du sonicateur et incubés à 4 °C sous rotation pendant une nuit. Le lendemain, les histones ont été centrifugées à 15600 g pendant 20 minutes à 4 °C. Les surnageants ont été récupérés et dosés à l'aide du réactif de Bradford. Pour comparer le niveau d'acétylation d'histones entre les deux lignées cellulaires, 2 μ g d'histones ont été mélangées avec du Laemmli pH 11, contenant 200 mM de β -mercaptoéthanol, migrées dans un gel SDS-PAGE et transférées sur une membrane de nitrocellulose pour détection par Western Blot comme décrit précédemment.

Traitement des cellules à l'Étoposide: Six millions de cellules ont été traitées avec 25 μ M d'Étoposide et placées pendant 2 heures en culture. Le milieu a été ôté, les cellules ont été lavées deux fois avec du PBS et ensuite,ensemencées dans une plaque de culture de 6 puits

(TPP) et remises en culture dans du milieu RPMI complet pendant 1, 2, 3, 6 et 8 jours. Un contrôle des cellules traitées avec du DMSO (Euromedex) a été également effectué.

Fixation des cellules et coloration β -galactosidase : À l'issue du traitement à l'étoposide, les cellules ont été fixées et colorées pendant une nuit à l'aide du kit de coloration de sénescence β -galactosidase (9860, Cell Signaling). Des images des cellules ont été obtenues au microscope optique (Leica LED Fluo) et traitées avec le logiciel ImageJ (Fiji).

9) Cristallisation du mutant T605P MOZ

Cristallogénèse de MOZ T605P : La protéine MOZ T605P recombinante a été produite et purifiée comme décrit précédemment (purification, digestion par thrombine de l'étiquette 6-histidines, purification par chromatographie d'exclusion). La mesure de l'activité enzymatique par HPLC, la pureté et l'absence de l'étiquette 6-histidines ont été également vérifiées. Un échantillon contenant 5,5 mg/mL de MOZ T605P ; 1,83 mM AcétylCoA ; 0,92 mM peptide p53K382 (ratio molaire entre protéine, peptide, cofacteur à 1 : 5 : 10 respectivement) a été préparé dans du tampon de cristallographie (20 mM sodium citrate, 100 mM NaCl pH 5,5) et incubé à 4 °C pour une nuit.

Un criblage initial de 7 plaques multipuits (Greiner Bio-One) contenant au total 672 conditions commerciales de cristallogénèse a permis d'identifier les meilleures conditions pour l'obtention de cristaux. Le criblage a été effectué à l'aide du Mosquito™ nanoliter (TPP Labtech), selon la méthode de diffusion de vapeur en goutte assise. Deux cents nL de l'échantillon ont été mélangés avec un volume équivalent de la solution du réservoir provenant des plaques multipuits. Les plaques ont été stockées à 18 °C dans le système RockImager (Formulatrix), qui permet de suivre la croissance des cristaux par la prise régulière des images des gouttes générées. Les cristaux de MOZ T605P ont été obtenus dans la journée et reproduits à la main, dans une plaque de 24 puits à 18 °C, utilisant la méthode de goutte suspendue. La condition qui a donné les meilleurs cristaux est la 0,1 M MES

pH 6,5, 12 % w/v polyéthylène glycol 20k (plaque MDL 1 & 2, puits F10). Les cristaux ont été cryoprotégés à l'aide de glycérol et de la solution de cristallisation, et conservés dans l'azote liquide jusqu'au moment de leur diffraction.

Diffraction des cristaux : Les cristaux ont été diffractés sur la ligne PROXIMA-1 au synchrotron SOLEIL (St Aubin, France). Les données de diffraction ont été analysées avec le programme XDS et la suite Collaborative Computational Project, Number 4 (CCP4i) ^{186,187}.

Résolution et analyse de la structure : La structure de MOZ T605P a été résolue par remplacement moléculaire en utilisant comme modèle les structures de MOZ WT complexée avec le cofacteur AcétylCoA (PDB : 2OZU et 2RC4) grâce au logiciel Phaser ¹⁸⁸. La structure finale à 2,9 Å a été obtenue après ajustement et raffinement par les logiciels respectifs Coot et REFMAC ^{189,190}. Les analyses de la structure ont été faites avec le logiciel PyMol et l'analyse des cavités grâce au logiciel Caver^{WEB} ¹⁹⁷. Les structures 2OZU, 2RC4, 6BA4, 3TO6, 3TOA, 1MJA ont été utilisées pour l'analyse de la structure MOZ T605P par PyMol (*The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC*).

V. Annexes

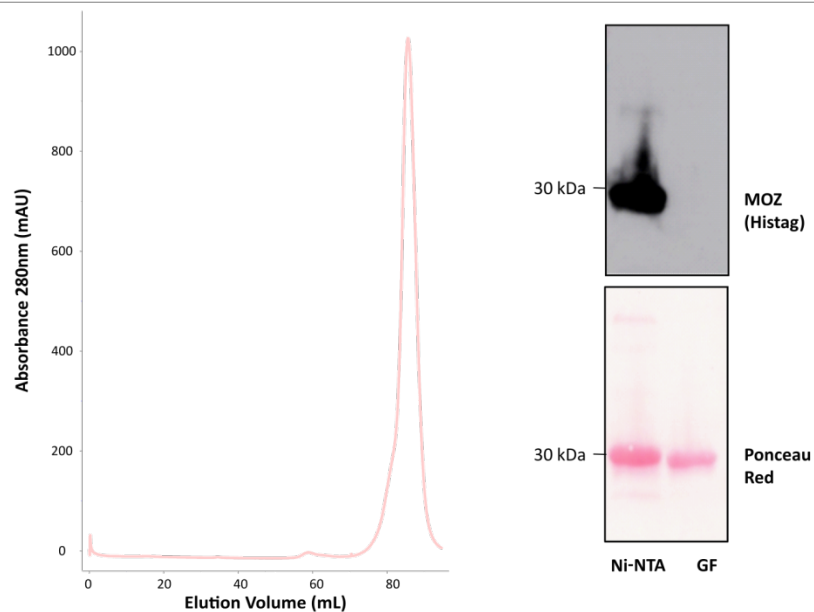


Figure M-1 Supplémentaire : Purification par Ni-NTA et clivage de l'étiquette 6-histidine par thrombine, suivie par purification par chromatographie d'exclusion (GF) pour des essais de cristallogénèse

Le domaine MYST du MOZ T605P recombinante, couplé à une étiquette 6-histidine (N-terminale), a été purifié par affinité avec des billes de nickel-agarose. L'étiquette 6-histidine a été clivée par incubation de la protéine avec de la thrombine (10 U thrombine/ mg de protéine) pendant une nuit à 15 °C sous agitation. L'échantillon a été ensuite incubé avec des billes de benzamidine-agarose pendant 30 minutes à 4 °C et purifié à nouveau par chromatographie d'exclusion (GF). La pureté de la protéine, ainsi que l'absence de l'étiquette 6-histidine suite au clivage par la thrombine, ont été vérifiées par Western Blot. 10 µL de l'échantillon issu de chaque étape d'intérêt (sortie des colonnes de Ni-NTA et GF) ont été séparés par migration en conditions réductrices (SDS-PAGE). Les protéines du gel ont été transférées sur une membrane de nitrocellulose 0,2 µm et détectées par un anticorps contre l'étiquette histidine.

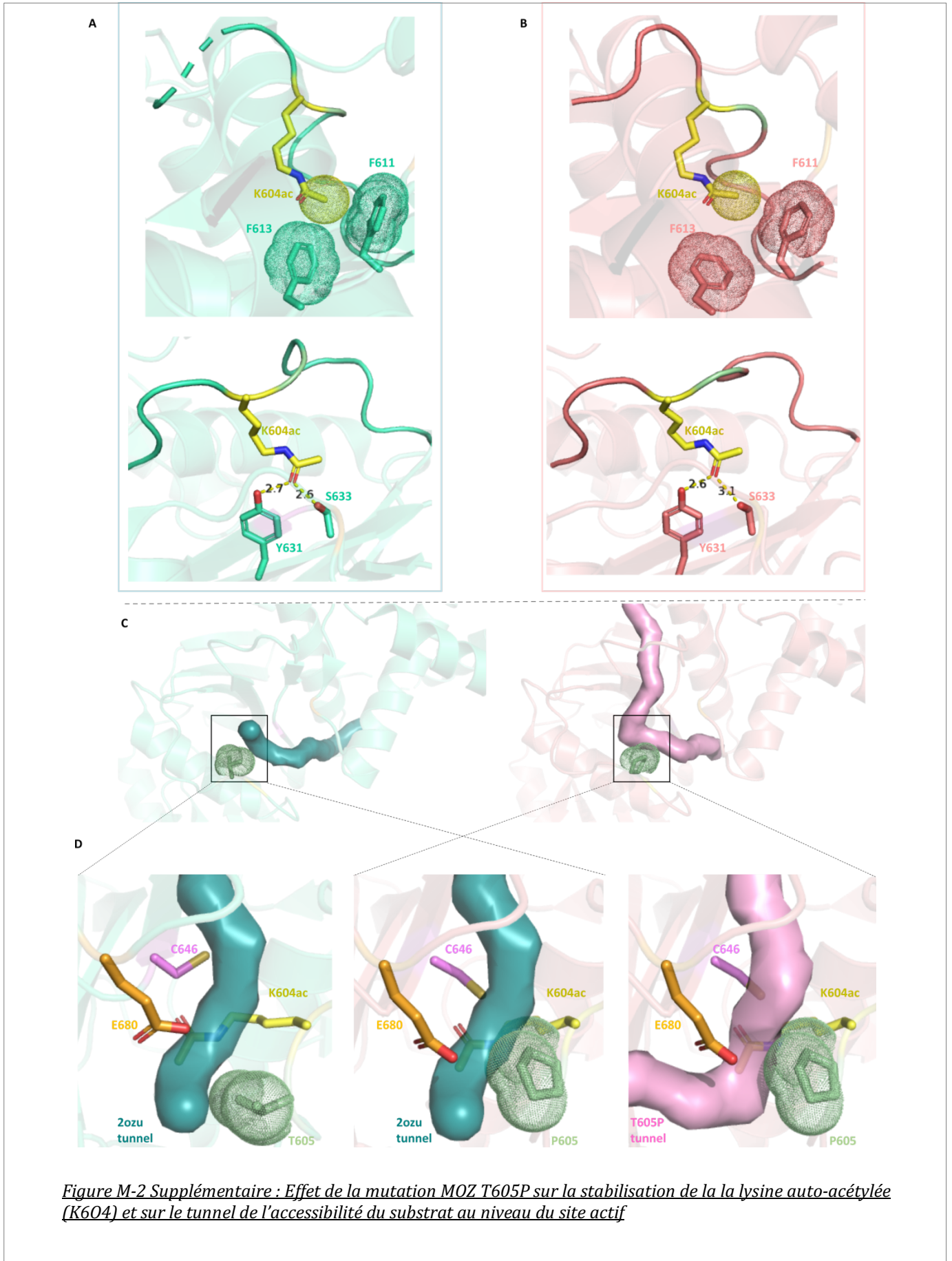


Figure M-2 Supplémentaire : Effet de la mutation MOZ T605P sur la stabilisation de la lysine auto-acétylée (K604) et sur le tunnel de l'accessibilité du substrat au niveau du site actif

La position de la lysine auto-acétylée (K604ac), ainsi que la boucle $\alpha 2\beta 7$, sont stabilisées par des interactions du résidu K604ac avec les résidus de son environnement. Le groupement méthyle de la K604ac fait des interactions de Van der Waals avec les noyaux aromatiques des phénylalanines F611 et F613. Le groupement carbonyle de la K604ac forme des liaisons hydrogène avec les Y631 et S633.

A : Les résidus permettant la stabilisation de la K604ac (F611 - F613 en haut, Y631 – S633 en bas) dans la structure MOZ WT (PDB : 2OZU).

B : Les résidus permettant la stabilisation de la K604ac (F611 - F613 en haut, Y631 – S633 en bas) dans la structure MOZ T605P.

C : Les tunnels d'accessibilité du solvant au niveau du site actif de l'enzyme sont calculés avec le logiciel CAVER^{WEB 197}. Les tunnels obtenus par cette analyse (tunnel issu de la structure MOZ WT 2OZU en bleu, tunnel issu de la structure MOZ T605P en rose) sont alignés avec les structures de MOZ WT (greencyan) et MOZ T605P (deepsalmon).

D : Les chaînes latérales des résidus du site actif (K604ac colorée en jaune CPK, C646 colorée en violet CPK, E680 brightorange CPK) sont représentées en stick. La chaîne latérale du résidu 605 est représentée en palegreen. Le panel de gauche représente le tunnel WT 2OZU, aligné avec la structure MOZ WT 2OZU. Le panel central représente le tunnel WT 2OZU aligné avec la structure MOZ T605P. Le panel de droite

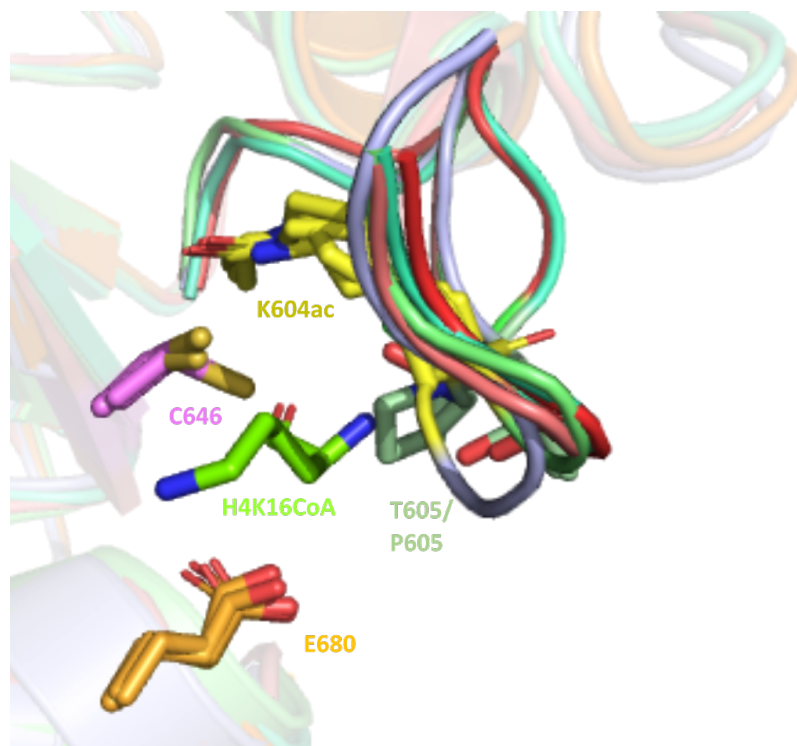


Figure M-3 Supplémentaire : Modélisation de la fixation du substrat H4K16CoA (PDB : 3T06) dans les structures MOZ WT (PDB : 2OZU, 2RC4, 6BA4) et MOZ T605P

Les structures MOZ WT (PDB 2OZU colorée en greencyan, 2RC4 colorée en lime, 6BA4 colorée en lightblue), MOZ T605P (colorée en deepsalmon) et ESA1 WT (PDB 3T06 colorée en rouge) ont été alignées. Les chaînes latérales des résidus catalytiques, issus de chaque structure (K604 colorée en jaune CPK, C646 colorée en violet CPK, E680 en brightorange CPK), ainsi que le résidu 605 (T ou P, coloré en palegreen CPK), sont représentées en stick. Le substrat H4K16CoA cocrystallisé avec l'enzyme ESA1 WT, est représenté en vert CPK.

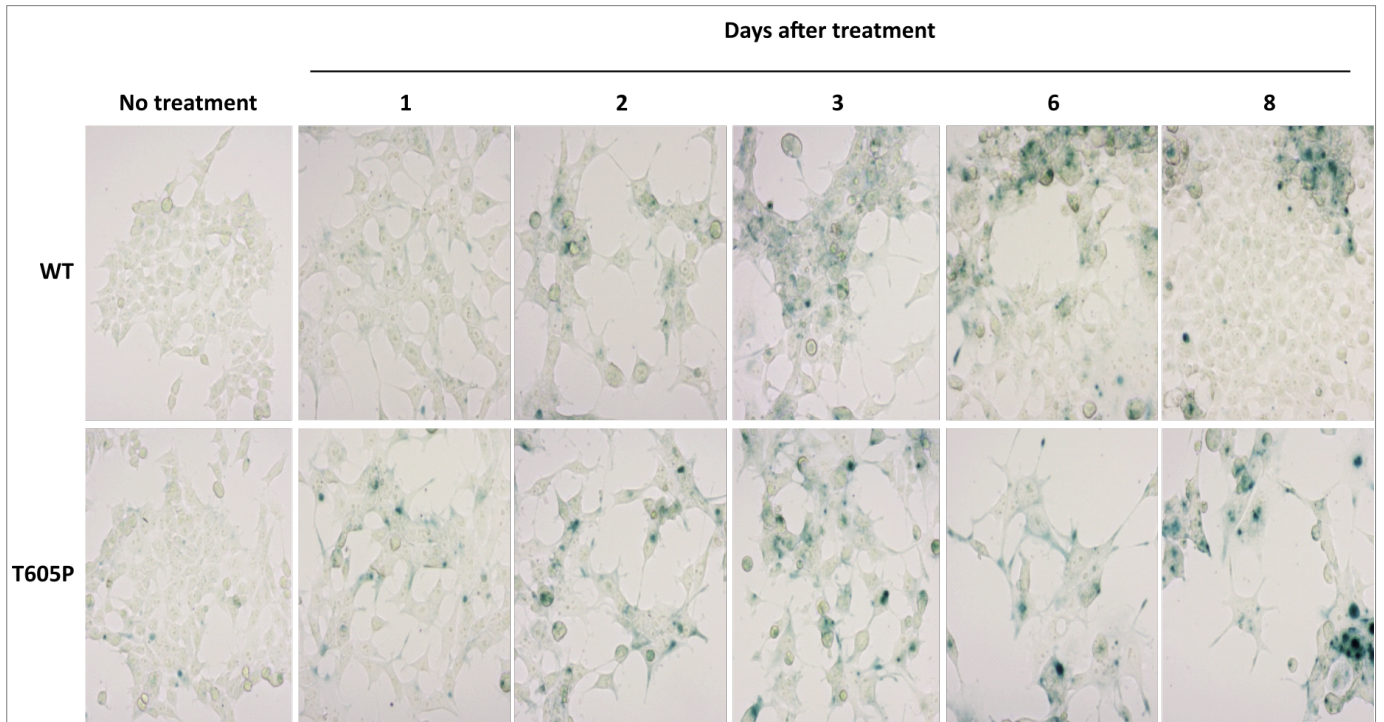


Figure M-4 Supplémentaire : Effet du traitement à l'étoposide sur l'induction de sénescence

Les cellules HEK, exprimant de manière constitutive la forme sauvage WT ou mutante T605P de MOZ, ont été traitées pendant 2 heures avec 25 μ M d'étoposide. À la fin du traitement, les cellules ont été mises en culture pendant 1, 2, 3, 6 ou 8 jours. À chaque temps défini, les cellules ont été fixées et colorées avec le kit de coloration β -galactosidase pour marquer les cellules sénescents.

CONCLUSION GÉNÉRALE

Ce travail de thèse a permis de poser les bases sur la caractérisation structurale et fonctionnelle de mutations de l'histone méthyltransférase SETD2 et de l'histone acétyltransférase MOZ identifiées dans certains cancers. Les deux mutations étudiées, SETD2 L1609P et MOZ T605P, entraînent des 'pertes d'activité' enzymatique par des altérations au niveau du site actif.

La dérégulation de l'activité de SETD2 par la mutation L1609P est également fortement liée à une perte de stabilité de l'enzyme, qui est par ailleurs généralement décrite comme un exemple de protéine instable. Nous avons également utilisé pour la première fois des approches de dynamique moléculaire afin de mieux caractériser les effets structuraux de cette mutation. Ces analyses sont en accord avec d'autres expériences qui démontrent la plus grande instabilité du mutant SETD2 L1609P comparée à la forme sauvage. Ce phénotype est souvent associé avec la diminution de l'efficacité de réparation de cassures de l'ADN et la capacité de résistance aux agents chimiothérapeutiques. Des cellules modifiées par approche CRISPR/Cas9 exprimant le mutant L1609P confirment ces effets.

Les cellules exprimant le mutant SETD2 L1609P présentent également des différences en termes d'adhérence et de morphologie cellulaire. L'analyse de cet effet est pertinente, car SETD2 est décrite aujourd'hui comme capable de méthyliser l' α -tubuline et l'actine, qui sont des protéines majeures du cytosquelette. L'effet de ces méthylations pourrait expliquer en partie la différence observée. L'analyse de ce phénotype est en cours et permettra de mieux comprendre l'implication de la mutation SETD2 L1609P sur le cytosquelette et sa dynamique.

En ce qui concerne l'actétyltransférase MOZ, la mutation T605P induit également la dérégulation de son activité et l'altération subséquente de l'acétylation de l'histone H3 et de p53. Une partie importante de notre étude est la résolution de la structure cristallographique de ce mutant de MOZ, qui nous a permis de mettre en évidence une altération du site actif de l'enzyme. Pour mieux décrire les conséquences fonctionnelles de la mutation MOZ T605P, nous sommes actuellement en cours d'analyse de l'expression génique par des approches de génomique fonctionnelle (ChIP-Seq, RNA-Seq). De plus, nous sommes en train d'analyser plus en détail la capacité du mutant T605P à induire un phénotype sénescence, souvent associé à des défauts d'activité de MOZ. Des cellules

modifiées par approche CRISPR/Cas9 et exprimant le mutant MOZ T605P suggèrent d'ores et déjà des différences des propriétés de sénescence de ces cellules.

RÉFÉRENCES

1. Jenuwein T, Allis CD. Translating the histone code. *Science*. 2001;293(5532):1074–1080.
2. Strahl BD, Allis CD. The language of covalent histone modifications. *Nature*. 2000;403(6765):41–45.
3. Martin C, Zhang Y. The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* 2005;6(11):838–849.
4. Hyun K, Jeon J, Park K, Kim J. Writing, erasing and reading histone lysine methylations. *Exp. Mol. Med.* 2017;49(4):e324.
5. Barski A, Cuddapah S, Cui K, et al. High-resolution profiling of histone methylations in the human genome. *Cell*. 2007;129(4):823–837.
6. Hughes AL, Kelley JR, Klose RJ. Understanding the interplay between CpG island-associated gene promoters and H3K4 methylation. *Biochim. Biophys. acta. Gene Regul. Mech.* 2020;1863(8):194567.
7. Loyola A, Tagami H, Bonaldi T, et al. The HP1alpha-CAF1-SetDB1-containing complex provides H3K9me1 for Suv39-mediated K9me3 in pericentric heterochromatin. *EMBO Rep.* 2009;10(7):769–775.
8. Rice JC, Briggs SD, Ueberheide B, et al. Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol. Cell*. 2003;12(6):1591–1598.
9. Ryu H, Lee J, Hagerty SW, et al. ESET/SETDB1 gene expression and histone H3 (K9) trimethylation in Huntington's disease. *Proc. Natl. Acad. Sci. U. S. A.* 2006;103(50):19176–19181.

10. Cao R, Wang L, Wang H, et al. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science*. 2002;298(5595):1039–1043.
11. Nekrasov M, Wild B, Müller J. Nucleosome binding and histone methyltransferase activity of Drosophila PRC2. *EMBO Rep*. 2005;6(4):348–353.
12. Ferrari KJ, Scelfo A, Jammula S, et al. Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity. *Mol. Cell*. 2014;53(1):49–62.
13. Laugesen A, Højfeldt JW, Helin K. Molecular Mechanisms Directing PRC2 Recruitment and H3K27 Methylation. *Mol. Cell*. 2019;74(1):8–18.
14. Wagner EJ, Carpenter PB. Understanding the language of Lys36 methylation at histone H3. *Nat. Rev. Mol. Cell Biol*. 2012;13(2):115–126.
15. Bannister AJ, Schneider R, Myers FA, et al. Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *J. Biol. Chem*. 2005;280(18):17732–17736.
16. Yuan W, Xu M, Huang C, et al. H3K36 methylation antagonizes PRC2-mediated H3K27 methylation. *J. Biol. Chem*. 2011;286(10):7983–7989.
17. Fnu S, Williamson EA, De Haro LP, et al. Methylation of histone H3 lysine 36 enhances DNA repair by nonhomologous end-joining. *Proc. Natl. Acad. Sci. U. S. A*. 2011;108(2):540–545.
18. Li F, Mao G, Tong D, et al. The histone mark H3K36me3 regulates human DNA mismatch repair through its interaction with MutS α . *Cell*. 2013;153(3):590–600.
19. Lu C, Jain SU, Hoelper D, et al. Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. *Science*. 2016;352(6287):844–849.
20. Fang D, Gan H, Lee J-H, et al. The histone H3.3K36M mutation reprograms the epigenome of chondroblastomas. *Science*. 2016;352(6291):1344–1348.
21. Kurotaki N, Imaizumi K, Harada N, et al. Haploinsufficiency of NSD1 causes Sotos syndrome. *Nat. Genet*. 2002;30(4):365–366.
22. Jaju RJ, Fidler C, Haas OA, et al. A novel gene, NSD1, is fused to NUP98 in the

- t(5;11)(q35;p15.5) in de novo childhood acute myeloid leukemia. *Blood*. 2001;98(4):1264–1267.
23. Nimura K, Ura K, Shiratori H, et al. A histone H3 lysine 36 trimethyltransferase links Nkx2-5 to Wolf-Hirschhorn syndrome. *Nature*. 2009;460(7252):287–291.
 24. Angrand PO, Apiou F, Stewart AF, et al. NSD3, a new SET domain-containing gene, maps to 8p12 and is amplified in human breast cancer cell lines. *Genomics*. 2001;74(1):79–88.
 25. Rosati R, La Starza R, Veronese A, et al. NUP98 is fused to the NSD3 gene in acute myeloid leukemia associated with t(8;11)(p11.2;p15). *Blood*. 2002;99(10):3857–3860.
 26. Duns G, van den Berg E, van Duivenbode I, et al. Histone methyltransferase gene SETD2 is a novel tumor suppressor gene in clear cell renal cell carcinoma. *Cancer Res*. 2010;70(11):4287–4291.
 27. Zhu X, He F, Zeng H, et al. Identification of functional cooperative mutations of SETD2 in human acute leukemia. *Nat. Genet*. 2014;46(3):287–293.
 28. Nguyen AT, Zhang Y. The diverse functions of Dot1 and H3K79 methylation. *Genes Dev*. 2011;25(13):1345–1358.
 29. Steger DJ, Lefterova MI, Ying L, et al. DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. *Mol. Cell. Biol*. 2008;28(8):2825–2839.
 30. Mueller D, Bach C, Zeisig D, et al. A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. *Blood*. 2007;110(13):4445–4454.
 31. Fang J, Feng Q, Ketel CS, et al. Purification and functional characterization of SET8, a nucleosomal histone H4-lysine 20-specific methyltransferase. *Curr. Biol*. 2002;12(13):1086–1099.
 32. Yang H, Pesavento JJ, Starnes TW, et al. Preferential dimethylation of histone H4 lysine 20 by Suv4-20. *J. Biol. Chem*. 2008;283(18):12085–12092.

33. Schotta G, Lachner M, Sarma K, et al. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev.* 2004;18(11):1251–1262.
34. Jørgensen S, Schotta G, Sørensen CS. Histone H4 lysine 20 methylation: key player in epigenetic regulation of genomic integrity. *Nucleic Acids Res.* 2013;41(5):2797–2806.
35. Botuyan MV, Lee J, Ward IM, et al. Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. *Cell.* 2006;127(7):1361–1373.
36. Bannister AJ, Schneider R, Kouzarides T. Histone methylation: dynamic or static? *Cell.* 2002;109(7):801–806.
37. Shi Y, Lan F, Matson C, et al. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell.* 2004;119(7):941–953.
38. Yuan W, Xie J, Long C, et al. Heterogeneous nuclear ribonucleoprotein L is a subunit of human KMT3a/Set2 complex required for H3 Lys-36 trimethylation activity in vivo. *J. Biol. Chem.* 2009;284(23):15701–15707.
39. Klose RJ, Yamane K, Bae Y, et al. The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature.* 2006;442(7100):312–316.
40. Young NL, Dere R. Mechanistic insights into KDM4A driven genomic instability. *Biochem. Soc. Trans.* 2021;49(1):93–105.
41. Hillringhaus L, Yue WW, Rose NR, et al. Structural and evolutionary basis for the dual substrate selectivity of human KDM4 histone demethylase family. *J. Biol. Chem.* 2011;286(48):41616–41625.
42. Nelson CJ, Santos-Rosa H, Kouzarides T. Proline isomerization of histone H3 regulates lysine methylation and gene expression. *Cell.* 2006;126(5):905–916.
43. Zhang Y, Shan C-M, Wang J, et al. Molecular basis for the role of oncogenic histone mutations in modulating H3K36 methylation. *Sci. Rep.* 2017;7:43906.

44. Kouzarides T. Chromatin modifications and their function. *Cell*. 2007;128(4):693–705.
45. Musselman CA, Lalonde M-E, Côté J, Kutateladze TG. Perceiving the epigenetic landscape through histone readers. *Nat. Struct. Mol. Biol.* 2012;19(12):1218–1227.
46. Vezzoli A, Bonadies N, Allen MD, et al. Molecular basis of histone H3K36me3 recognition by the PWWP domain of Brpf1. *Nat. Struct. Mol. Biol.* 2010;17(5):617–619.
47. Wu H, Zeng H, Lam R, et al. Structural and histone binding ability characterizations of human PWWP domains. *PLoS One*. 2011;6(6):e18919.
48. Maurer-Stroh S, Dickens NJ, Hughes-Davies L, et al. The Tudor domain “Royal Family”: Tudor, plant Agenet, Chromo, PWWP and MBT domains. *Trends Biochem. Sci.* 2003;28(2):69–74.
49. Li J, Duns G, Westers H, et al. SETD2: an epigenetic modifier with tumor suppressor functionality. *Oncotarget*. 2016;7(31):50719–50734.
50. Zhang P, Du J, Sun B, et al. Structure of human MRG15 chromo domain and its binding to Lys36-methylated histone H3. *Nucleic Acids Res.* 2006;34(22):6621–6628.
51. Kim D, Blus BJ, Chandra V, et al. Corecognition of DNA and a methylated histone tail by the MSL3 chromodomain. *Nat. Struct. Mol. Biol.* 2010;17(8):1027–1029.
52. Musselman CA, Avvakumov N, Watanabe R, et al. Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. *Nat. Struct. Mol. Biol.* 2012;19(12):1266–1272.
53. Ballaré C, Lange M, Lapinaite A, et al. Phf19 links methylated Lys36 of histone H3 to regulation of Polycomb activity. *Nat. Struct. Mol. Biol.* 2012;19(12):1257–1265.
54. Carvalho S, Raposo AC, Martins FB, et al. Histone methyltransferase SETD2 coordinates FACT recruitment with nucleosome dynamics during transcription. *Nucleic Acids Res.* 2013;41(5):2881–2893.
55. Wen H, Li Y, Xi Y, et al. ZMYND11 links histone H3.3K36me3 to transcription

- elongation and tumour suppression. *Nature*. 2014;508(7495):263–268.
56. Wen H, Li Y, Li H, Shi X. ZMYND11: an H3.3-specific reader of H3K36me3. *Cell Cycle*. 2014;13(14):2153–2154.
 57. de Almeida SF, Grosso AR, Koch F, et al. Splicing enhances recruitment of methyltransferase HYPB/Setd2 and methylation of histone H3 Lys36. *Nat. Struct. Mol. Biol.* 2011;18(9):977–983.
 58. Kolasinska-Zwierz P, Down T, Latorre I, et al. Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat. Genet.* 2009;41(3):376–381.
 59. Luco RF, Pan Q, Tominaga K, et al. Regulation of alternative splicing by histone modifications. *Science*. 2010;327(5968):996–1000.
 60. Pradeepa MM, Sutherland HG, Ule J, Grimes GR, Bickmore WA. Psip1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLoS Genet.* 2012;8(5):e1002717.
 61. Huang H, Weng H, Zhou K, et al. Histone H3 trimethylation at lysine 36 guides m(6)A RNA modification co-transcriptionally. *Nature*. 2019;567(7748):414–419.
 62. Yoh SM, Lucas JS, Jones KA. The Iws1:Spt6:CTD complex controls cotranscriptional mRNA biosynthesis and HYPB/Setd2-mediated histone H3K36 methylation. *Genes Dev.* 2008;22(24):3422–3434.
 63. Pfister SX, Ahrabi S, Zalmas L-P, et al. SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Rep.* 2014;7(6):2006–2018.
 64. Aymard F, Bugler B, Schmidt CK, et al. Transcriptionally active chromatin recruits homologous recombination at DNA double-strand breaks. *Nat. Struct. Mol. Biol.* 2014;21(4):366–374.
 65. Zhang M, Yang Y, Zhou M, et al. Histone and DNA binding ability studies of the NSD subfamily of PWWP domains. *Biochem. Biophys. Res. Commun.* 2021;569:199–206.

66. Li X, Wu L, Corsa CAS, Kunkel S, Dou Y. Two mammalian MOF complexes regulate transcription activation by distinct mechanisms. *Mol. Cell.* 2009;36(2):290–301.
67. Li L, Wang Y. Cross-talk between the H3K36me3 and H4K16ac histone epigenetic marks in DNA double-strand break repair. *J. Biol. Chem.* 2017;292(28):11951–11959.
68. Dhayalan A, Rajavelu A, Rathert P, et al. The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J. Biol. Chem.* 2010;285(34):26114–26120.
69. Hahn MA, Wu X, Li AX, Hahn T, Pfeifer GP. Relationship between gene body DNA methylation and intragenic H3K9me3 and H3K36me3 chromatin marks. *PLoS One.* 2011;6(4):e18844.
70. Klein BJ, Krajewski K, Restrepo S, et al. Recognition of cancer mutations in histone H3K36 by epigenetic writers and readers. *Epigenetics.* 2018;13(7):683–692.
71. Jenuwein T, Laible G, Dorn R, Reuter G. SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cell. Mol. Life Sci.* 1998;54(1):80–93.
72. Marmorstein R. Structure of SET domain proteins: a new twist on histone methylation. *Trends Biochem. Sci.* 2003;28(2):59–62.
73. Völkel P, Angrand P-O. The control of histone lysine methylation in epigenetic regulation. *Biochimie.* 2007;89(1):1–20.
74. Mao M, Fu G, Wu JS, et al. Identification of genes expressed in human CD34(+) hematopoietic stem/progenitor cells by expressed sequence tags and efficient full-length cDNA cloning. *Proc. Natl. Acad. Sci. U. S. A.* 1998;95(14):8175–8180.
75. Zhang QH, Ye M, Wu XY, et al. Cloning and functional analysis of cDNAs with open reading frames for 300 previously undefined genes expressed in CD34+ hematopoietic stem/progenitor cells. *Genome Res.* 2000;10(10):1546–1560.
76. Sun X-J, Wei J, Wu X-Y, et al. Identification and characterization of a novel human histone H3 lysine 36-specific methyltransferase. *J. Biol. Chem.* 2005;280(42):35261–35271.

77. Faber PW, Barnes GT, Srinidhi J, et al. Huntingtin interacts with a family of WW domain proteins. *Hum. Mol. Genet.* 1998;7(9):1463–1474.
78. Edmunds JW, Mahadevan LC, Clayton AL. Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J.* 2008;27(2):406–420.
79. Bhattacharya S, Workman JL. Regulation of SETD2 stability is important for the fidelity of H3K36me3 deposition. *Epigenetics Chromatin.* 2020;13(1):40.
80. Doerks T, Copley RR, Schultz J, Ponting CP, Bork P. Systematic identification of novel protein domain families associated with nuclear functions. *Genome Res.* 2002;12(1):47–56.
81. Wilson JR, Jing C, Walker PA, et al. Crystal structure and functional analysis of the histone methyltransferase SET7/9. *Cell.* 2002;111(1):105–115.
82. Zhang X, Yang Z, Khan SI, et al. Structural basis for the product specificity of histone lysine methyltransferases. *Mol. Cell.* 2003;12(1):177–185.
83. Yang S, Zheng X, Lu C, et al. Molecular basis for oncohistone H3 recognition by SETD2 methyltransferase. *Genes Dev.* 2016;30(14):1611–1616.
84. Hacker KE, Fahey CC, Shinsky SA, et al. Structure/Function Analysis of Recurrent Mutations in SETD2 Protein Reveals a Critical and Conserved Role for a SET Domain Residue in Maintaining Protein Stability and Histone H3 Lys-36 Trimethylation. *J. Biol. Chem.* 2016;291(40):21283–21295.
85. Dillon SC, Zhang X, Trievel RC, Cheng X. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol.* 2005;6(8):227.
86. Qian C, Zhou MM. SET domain protein lysine methyltransferases: Structure, specificity and catalysis. *Cell. Mol. Life Sci.* 2006;63(23):2755–2763.
87. Schapira M. Structural Chemistry of Human SET Domain Protein Methyltransferases. *Curr. Chem. Genomics.* 2011;5(Suppl 1):85–94.

88. Chen HI, Sudol M. The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. *Proc. Natl. Acad. Sci. U. S. A.* 1995;92(17):7819–7823.
89. Kay BK, Williamson MP, Sudol M. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 2000;14(2):231–241.
90. Gao Y-G, Yang H, Zhao J, Jiang Y-J, Hu H-Y. Autoinhibitory structure of the WW domain of HYPB/SETD2 regulates its interaction with the proline-rich region of huntingtin. *Structure.* 2014;22(3):378–386.
91. Kizer KO, Phatnani HP, Shibata Y, et al. A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Mol. Cell. Biol.* 2005;25(8):3305–3316.
92. Rebehmed J, Revy P, Faure G, de Villartay J-P, Callebaut I. Expanding the SRI domain family: a common scaffold for binding the phosphorylated C-terminal domain of RNA polymerase II. *FEBS Lett.* 2014;588(23):4431–4437.
93. Li M, Phatnani HP, Guan Z, et al. Solution structure of the Set2-Rpb1 interacting domain of human Set2 and its interaction with the hyperphosphorylated C-terminal domain of Rpb1. *Proc. Natl. Acad. Sci. U. S. A.* 2005;102(49):17636–17641.
94. Bhattacharya S, Levy MJ, Zhang N, et al. The methyltransferase SETD2 couples transcription and splicing by engaging mRNA processing factors through its SHI domain. *Nat. Commun.* 2021;12(1):1443.
95. Bhattacharya S, Wang S, Reddy D, et al. Structural basis of the interaction between SETD2 methyltransferase and hnRNP L paralogs for governing co-transcriptional splicing. *Nat. Commun.* 2021;12(1):6452.
96. Zheng W, Ibáñez G, Wu H, et al. Sinefungin derivatives as inhibitors and structure probes of protein lysine methyltransferase SETD2. *J. Am. Chem. Soc.* 2012;134(43):18004–18014.

97. Liu Y, Zhang Y, Xue H, et al. Cryo-EM structure of SETD2/Set2 methyltransferase bound to a nucleosome containing oncohistone mutations. *Cell Discov.* 2021;7(1):32.
98. Bhattacharya S, Lange JJ, Levy M, et al. The disordered regions of the methyltransferase SETD2 govern its function by regulating its proteolysis and phase separation. *J. Biol. Chem.* 2021;297(3):101075.
99. Park IY, Powell RT, Tripathi DN, et al. Dual Chromatin and Cytoskeletal Remodeling by SETD2. *Cell.* 2016;166(4):950–962.
100. Seervai RNH, Jangid RK, Karki M, et al. The Huntingtin-interacting protein SETD2/HYPB is an actin lysine methyltransferase. *Sci. Adv.* 2020;6(40):.
101. Chen K, Liu J, Liu S, et al. Methyltransferase SETD2-Mediated Methylation of STAT1 Is Critical for Interferon Antiviral Activity. *Cell.* 2017;170(3):492-506.e14.
102. Yuan H, Han Y, Wang X, et al. SETD2 Restricts Prostate Cancer Metastasis by Integrating EZH2 and AMPK Signaling Pathways. *Cancer Cell.* 2020;38(3):350-365.e7.
103. Xie X, Wang S, Li M, et al. α -TubK40me3 is required for neuronal polarization and migration by promoting microtubule formation. *Nat. Commun.* 2021;12(1):4113.
104. Kearns S, Mason FM, Rathmell WK, et al. Molecular determinants for α -tubulin methylation by SETD2. *J. Biol. Chem.* 2021;297(1):100898.
105. Karki M, Jangid RK, Anish R, et al. A cytoskeletal function for PBRM1 reading methylated microtubules. *Sci. Adv.* 2021;7(14):.
106. Bailey MH, Tokheim C, Porta-Pardo E, et al. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell.* 2018;173(2):371-385.e18.
107. Dalgliesh GL, Furge K, Greenman C, et al. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature.* 2010;463(7279):360–363.
108. Chiang Y-C, Park I-Y, Terzo EA, et al. SETD2 Haploinsufficiency for Microtubule Methylation Is an Early Driver of Genomic Instability in Renal Cell Carcinoma. *Cancer Res.* 2018;78(12):3135–3146.

109. Ho TH, Park IY, Zhao H, et al. High-resolution profiling of histone h3 lysine 36 trimethylation in metastatic renal cell carcinoma. *Oncogene*. 2016;35(12):1565–1574.
110. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature*. 2013;499(7456):43–49.
111. Kanu N, Grönroos E, Martinez P, et al. SETD2 loss-of-function promotes renal cancer branched evolution through replication stress and impaired DNA repair. *Oncogene*. 2015;34(46):5699–5708.
112. Chen B-Y, Song J, Hu C-L, et al. SETD2 deficiency accelerates MDS-associated leukemogenesis via S100a9 in NHD13 mice and predicts poor prognosis in MDS. *Blood*. 2020;135(25):2271–2285.
113. Skucha A, Ebner J, Schmöllerl J, et al. MLL-fusion-driven leukemia requires SETD2 to safeguard genomic integrity. *Nat. Commun*. 2018;9(1):1983.
114. Mar BG, Bullinger LB, McLean KM, et al. Mutations in epigenetic regulators including SETD2 are gained during relapse in paediatric acute lymphoblastic leukaemia. *Nat. Commun*. 2014;5:3469.
115. Mar BG, Chu SH, Kahn JD, et al. SETD2 alterations impair DNA damage recognition and lead to resistance to chemotherapy in leukemia. *Blood*. 2017;130(24):2631–2641.
116. Skucha A, Ebner J, Grebien F. Roles of SETD2 in Leukemia-Transcription, DNA-Damage, and Beyond. *Int. J. Mol. Sci*. 2019;20(5):.
117. Fang J, Huang Y, Mao G, et al. Cancer-driving H3G34V/R/D mutations block H3K36 methylation and H3K36me3-MutS α interaction. *Proc. Natl. Acad. Sci. U. S. A*. 2018;115(38):9598–9603.
118. Luscan A, Laurendeau I, Malan V, et al. Mutations in SETD2 cause a novel overgrowth condition. *J. Med. Genet*. 2014;51(8):512–517.
119. Tlemsani C, Luscan A, Leulliot N, et al. SETD2 and DNMT3A screen in the Sotos-like

- syndrome French cohort. *J. Med. Genet.* 2016;53(11):743–751.
120. Xu L, Zheng Y, Li X, et al. Abnormal neocortex arealization and Sotos-like syndrome-associated behavior in *Setd2* mutant mice. *Sci. Adv.* 2021;7(1):.
 121. Garcia-Ramirez M, Rocchini C, Ausio J. Modulation of chromatin folding by histone acetylation. *J. Biol. Chem.* 1995;270(30):17923–17928.
 122. Hong L, Schroth GP, Matthews HR, Yau P, Bradbury EM. Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 “tail” to DNA. *J. Biol. Chem.* 1993;268(1):305–314.
 123. Mutskov V, Gerber D, Angelov D, et al. Persistent interactions of core histone tails with nucleosomal DNA following acetylation and transcription factor binding. *Mol. Cell. Biol.* 1998;18(11):6293–6304.
 124. Lee DY, Hayes JJ, Pruss D, Wolffe AP. A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell.* 1993;72(1):73–84.
 125. Allis CD, Chicoine LG, Richman R, Schulman IG. Deposition-related histone acetylation in micronuclei of conjugating *Tetrahymena*. *Proc. Natl. Acad. Sci. U. S. A.* 1985;82(23):8048–8052.
 126. Marmorstein R, Zhou M-M. Writers and readers of histone acetylation: structure, mechanism, and inhibition. *Cold Spring Harb. Perspect. Biol.* 2014;6(7):a018762.
 127. Brownell JE, Zhou J, Ranalli T, et al. *Tetrahymena* histone acetyltransferase A: a homolog to yeast *Gcn5p* linking histone acetylation to gene activation. *Cell.* 1996;84(6):843–851.
 128. Allis CD, Berger SL, Cote J, et al. New nomenclature for chromatin-modifying enzymes. *Cell.* 2007;131(4):633–636.
 129. Shirmast P, Ghafoori SM, Irwin RM, et al. Structural characterization of a GNAT family acetyltransferase from *Elizabethkingia anophelis* bound to acetyl-CoA reveals a new dimeric interface. *Sci. Rep.* 2021;11(1):1274.

130. Tanner KG, Trievel RC, Kuo MH, et al. Catalytic mechanism and function of invariant glutamic acid 173 from the histone acetyltransferase GCN5 transcriptional coactivator. *J. Biol. Chem.* 1999;274(26):18157–18160.
131. Schuetz A, Bernstein G, Dong A, et al. Crystal structure of a binary complex between human GCN5 histone acetyltransferase domain and acetyl coenzyme A. *Proteins.* 2007;68(1):403–407.
132. Clements A, Rojas JR, Trievel RC, et al. Crystal structure of the histone acetyltransferase domain of the human PCAF transcriptional regulator bound to coenzyme A. *EMBO J.* 1999;18(13):3521–3532.
133. Rojas JR, Trievel RC, Zhou J, et al. Structure of Tetrahymena GCN5 bound to coenzyme A and a histone H3 peptide. *Nature.* 1999;401(6748):93–98.
134. Yee SP, Branton PE. Detection of cellular proteins associated with human adenovirus type 5 early region 1A polypeptides. *Virology.* 1985;147(1):142–153.
135. Chrivia JC, Kwok RP, Lamb N, et al. Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature.* 1993;365(6449):855–859.
136. Ogryzko V V, Schiltz RL, Russanova V, Howard BH, Nakatani Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell.* 1996;87(5):953–959.
137. Bannister AJ, Kouzarides T. The CBP co-activator is a histone acetyltransferase. *Nature.* 1996;384(6610):641–643.
138. Liu X, Wang L, Zhao K, et al. The structural basis of protein acetylation by the p300/CBP transcriptional coactivator. *Nature.* 2008;451(7180):846–850.
139. Park S, Stanfield RL, Martinez-Yamout MA, et al. Role of the CBP catalytic core in intramolecular SUMOylation and control of histone H3 acetylation. *Proc. Natl. Acad. Sci. U. S. A.* 2017;114(27):E5335–E5342.
140. Thompson PR, Wang D, Wang L, et al. Regulation of the p300 HAT domain via a novel activation loop. *Nat. Struct. Mol. Biol.* 2004;11(4):308–315.

141. Delvecchio M, Gaucher J, Aguilar-Gurrieri C, Ortega E, Panne D. Structure of the p300 catalytic core and implications for chromatin targeting and HAT regulation. *Nat. Struct. Mol. Biol.* 2013;20(9):1040–1046.
142. Borrow J, Stanton VPJ, Andresen JM, et al. The translocation t(8;16)(p11;p13) of acute myeloid leukaemia fuses a putative acetyltransferase to the CREB-binding protein. *Nat. Genet.* 1996;14(1):33–41.
143. Reifsnyder C, Lowell J, Clarke A, Pillus L. Yeast SAS silencing genes and human genes associated with AML and HIV-1 Tat interactions are homologous with acetyltransferases. *Nat. Genet.* 1996;14(1):42–49.
144. Ehrenhofer-Murray AE, Rivier DH, Rine J. The role of Sas2, an acetyltransferase homologue of *Saccharomyces cerevisiae*, in silencing and ORC function. *Genetics.* 1997;145(4):923–934.
145. Kamine J, Elangovan B, Subramanian T, Coleman D, Chinnadurai G. Identification of a cellular protein that specifically interacts with the essential cysteine region of the HIV-1 Tat transactivator. *Virology.* 1996;216(2):357–366.
146. Yan Y, Harper S, Speicher DW, Marmorstein R. The catalytic mechanism of the ESA1 histone acetyltransferase involves a self-acetylated intermediate. *Nat. Struct. Biol.* 2002;9(11):862–869.
147. Yuan H, Rossetto D, Mellert H, et al. MYST protein acetyltransferase activity requires active site lysine autoacetylation. *EMBO J.* 2012;31(1):58–70.
148. Takechi S, Nakayama T. Sas3 is a histone acetyltransferase and requires a zinc finger motif. *Biochem. Biophys. Res. Commun.* 1999;266(2):405–410.
149. Akhtar A, Becker PB. The histone H4 acetyltransferase MOF uses a C2HC zinc finger for substrate recognition. *EMBO Rep.* 2001;2(2):113–118.
150. Holbert MA, Sikorski T, Carten J, et al. The human monocytic leukemia zinc finger histone acetyltransferase domain contains DNA-binding activity implicated in chromatin targeting. *J. Biol. Chem.* 2007;282(50):36603–36613.

151. Gu W, Roeder RG. Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell*. 1997;90(4):595–606.
152. Champagne N, Pelletier N, Yang XJ. The monocytic leukemia zinc finger protein MOZ is a histone acetyltransferase. *Oncogene*. 2001;20(3):404–409.
153. Kitabayashi I, Aikawa Y, Nguyen LA, Yokoyama A, Ohki M. Activation of AML1-mediated transcription by MOZ and inhibition by the MOZ-CBP fusion protein. *EMBO J*. 2001;20(24):7184–7196.
154. Ali M, Yan K, Lalonde M-E, et al. Tandem PHD fingers of MORF/MOZ acetyltransferases display selectivity for acetylated histone H3 and are required for the association with chromatin. *J. Mol. Biol*. 2012;424(5):328–338.
155. Dreveny I, Deeves SE, Fulton J, et al. The double PHD finger domain of MOZ/MYST3 induces α -helical structure of the histone H3 tail to facilitate acetylation and methylation sampling and modification. *Nucleic Acids Res*. 2014;42(2):822–835.
156. Yang X-J. MOZ and MORF acetyltransferases: Molecular interaction, animal development and human disease. *Biochim. Biophys. Acta*. 2015;1853(8):1818–1826.
157. Champagne N, Bertos NR, Pelletier N, et al. Identification of a human histone acetyltransferase related to monocytic leukemia zinc finger protein. *J. Biol. Chem*. 1999;274(40):28528–28536.
158. Pelletier N, Champagne N, Stifani S, Yang X-J. MOZ and MORF histone acetyltransferases interact with the Runt-domain transcription factor Runx2. *Oncogene*. 2002;21(17):2729–2740.
159. Katsumoto T, Aikawa Y, Iwama A, et al. MOZ is essential for maintenance of hematopoietic stem cells. *Genes Dev*. 2006;20(10):1321–1330.
160. Bristow CAP, Shore P. Transcriptional regulation of the human MIP-1alpha promoter by RUNX1 and MOZ. *Nucleic Acids Res*. 2003;31(11):2735–2744.
161. Wiesel-Motiuk N, Assaraf YG. The key roles of the lysine acetyltransferases KAT6A and KAT6B in physiology and pathology. *Drug Resist. Updat. Rev. Comment*.

- Antimicrob. Anticancer Chemother.* 2020;53:100729.
162. Rokudai S, Aikawa Y, Tagata Y, et al. Monocytic leukemia zinc finger (MOZ) interacts with p53 to induce p21 expression and cell-cycle arrest. *J. Biol. Chem.* 2009;284(1):237–244.
 163. Rokudai S, Laptenko O, Arnal SM, et al. MOZ increases p53 acetylation and premature senescence through its complex formation with PML. *Proc. Natl. Acad. Sci. U. S. A.* 2013;110(10):3895–3900.
 164. Baell JB, Leaver DJ, Hermans SJ, et al. Inhibitors of histone acetyltransferases KAT6A/B induce senescence and arrest tumour growth. *Nature.* 2018;560(7717):253–257.
 165. Sheikh BN, Phipson B, El-Saafin F, et al. MOZ (MYST3, KAT6A) inhibits senescence via the INK4A-ARF pathway. *Oncogene.* 2015;34(47):5807–5820.
 166. Arboleda VA, Lee H, Dorrani N, et al. De novo nonsense mutations in KAT6A, a lysine acetyl-transferase gene, cause a syndrome including microcephaly and global developmental delay. *Am. J. Hum. Genet.* 2015;96(3):498–506.
 167. Tham E, Lindstrand A, Santani A, et al. Dominant mutations in KAT6A cause intellectual disability with recognizable syndromic features. *Am. J. Hum. Genet.* 2015;96(3):507–513.
 168. Millan F, Cho MT, Retterer K, et al. Whole exome sequencing reveals de novo pathogenic variants in KAT6A as a cause of a neurodevelopmental disorder. *Am. J. Med. Genet. A.* 2016;170(7):1791–1798.
 169. Kennedy J, Goudie D, Blair E, et al. KAT6A Syndrome: genotype-phenotype correlation in 76 patients with pathogenic KAT6A variants. *Genet. Med.* 2019;21(4):850–860.
 170. Urreizti R, Lopez-Martin E, Martinez-Monseny A, et al. Five new cases of syndromic intellectual disability due to KAT6A mutations: widening the molecular and clinical spectrum. *Orphanet J. Rare Dis.* 2020;15(1):44.
 171. Troisi S, Maitz S, Severino M, et al. Epilepsy in KAT6A syndrome: Description of two individuals and revision of the literature. *Eur. J. Med. Genet.* 2022;65(1):104380.

172. Yu L, Liang Y, Cao X, et al. Identification of MYST3 as a novel epigenetic activator of ER α frequently amplified in breast cancer. *Oncogene*. 2017;36(20):2910–2918.
173. Lv D, Jia F, Hou Y, et al. Histone Acetyltransferase KAT6A Upregulates PI3K/AKT Signaling through TRIM24 Binding. *Cancer Res*. 2017;77(22):6190–6201.
174. Chaffanet M, Mozziconacci MJ, Fernandez F, et al. A case of inv(8)(p11q24) associated with acute myeloid leukemia involves the MOZ and CBP genes in a masked t(8;16). *Genes. Chromosomes Cancer*. 1999;26(2):161–165.
175. Panagopoulos I, Isaksson M, Lindvall C, et al. RT-PCR analysis of the MOZ-CBP and CBP-MOZ chimeric transcripts in acute myeloid leukemias with t(8;16)(p11;p13). *Genes. Chromosomes Cancer*. 2000;28(4):415–424.
176. Chan EM, Chan RJ, Comer EM, et al. MOZ and MOZ-CBP cooperate with NF-kappaB to activate transcription from NF-kappaB-dependent promoters. *Exp. Hematol*. 2007;35(12):1782–1792.
177. Chaffanet M, Gressin L, Preudhomme C, et al. MOZ is fused to p300 in an acute monocytic leukemia with t(8;22). *Genes. Chromosomes Cancer*. 2000;28(2):138–144.
178. Carapeti M, Aguiar RC, Goldman JM, Cross NC. A novel fusion between MOZ and the nuclear receptor coactivator TIF2 in acute myeloid leukemia. *Blood*. 1998;91(9):3127–3133.
179. Liang J, Prouty L, Williams BJ, Dayton MA, Blanchard KL. Acute mixed lineage leukemia with an inv(8)(p11q13) resulting in fusion of the genes for MOZ and TIF2. *Blood*. 1998;92(6):2118–2122.
180. Esteyries S, Perot C, Adelaide J, et al. NCOA3, a new fusion partner for MOZ/MYST3 in M5 acute myeloid leukemia. *Leukemia*. 2008;22(3):663–665.
181. Chinen Y, Taki T, Tsutsumi Y, et al. The leucine twenty homeobox (LEUTX) gene, which lacks a histone acetyltransferase domain, is fused to KAT6A in therapy-related acute myeloid leukemia with t(8;19)(p11;q13). *Genes. Chromosomes Cancer*. 2014;53(4):299–308.

182. Dong Y, Zhao X, Feng X, et al. SETD2 mutations confer chemoresistance in acute myeloid leukemia partly through altered cell cycle checkpoints. *Leukemia*. 2019;33(11):2585–2598.
183. Kim I-K, McCutcheon JN, Rao G, et al. Acquired SETD2 mutation and impaired CREB1 activation confer cisplatin resistance in metastatic non-small cell lung cancer. *Oncogene*. 2019;38(2):180–193.
184. Duval R, Fritsch L, Bui L-C, et al. An acetyltransferase assay for CREB-binding protein based on reverse phase-ultra-fast liquid chromatography of fluorescent histone H3 peptides. *Anal. Biochem*. 2015;486:35–37.
185. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. 2014;42(Web Server issue):W320-4.
186. Kabsch W. XDS. *Acta Crystallogr. D. Biol. Crystallogr*. 2010;66(Pt 2):125–132.
187. Winn MD, Ballard CC, Cowtan KD, et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr*. 2011;67(Pt 4):235–242.
188. McCoy AJ, Grosse-Kunstleve RW, Adams PD, et al. Phaser crystallographic software. *J. Appl. Crystallogr*. 2007;40(Pt 4):658–674.
189. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr*. 2010;66(Pt 4):486–501.
190. Murshudov GN, Skubák P, Lebedev AA, et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr*. 2011;67(Pt 4):355–367.
191. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol*. 1993;234(3):779–815.
192. Berendsen HJC, van der Spoel D, van Drunen R. GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun*. 1995;91(1):43–56.

193. Tian C, Kasavajhala K, Belfon KAA, et al. ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* 2020;16(1):528–552.
194. Giannakis M, Mu XJ, Shukla SA, et al. Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma. *Cell Rep.* 2016;15(4):857–865.
195. Giannakis M, Hodis E, Jasmine Mu X, et al. RNF43 is frequently mutated in colorectal and endometrial cancers. *Nat. Genet.* 2014;46(12):1264–1266.
196. Klein BJ, Jang SM, Lachance C, et al. Histone H3K23-specific acetylation by MORF is coupled to H3K14 acylation. *Nat. Commun.* 2019;10(1):4724.
197. Stourac J, Vavra O, Kokkonen P, et al. Caver Web 1.0: identification of tunnels and channels in proteins and analysis of ligand transport. *Nucleic Acids Res.* 2019;47(W1):W414–W422.

ANNEXES (PUBLICATIONS)



From transglutaminases (TGs) to arylamine *N*-acetyltransferases (NATs): Insight into the role of a spatially conserved aromatic amino acid position in the active site of these two families of enzymes

Ximing Xu ^{a, b, 1}, Wenchao Zhang ^{a, 1}, Jérémy Berthelet ^a, Rongxing Liu ^a, Christina Michail ^a, Alain F. Chaffotte ^c, Jean-Marie Dupret ^a, Fernando Rodrigues-Lima ^{a, *}

^a Université de Paris, BFA, UMR 8251, CNRS, 75013, Paris, France

^b Key Laboratory of Marine Drugs, Chinese Ministry of Education, School of Medicine and Pharmacy, Ocean University of China, Qingdao, 266003, China

^c Institut Pasteur, Unité de Résonance Magnétique Nucléaire des Biomolécules, 75015, Paris, France

ARTICLE INFO

Article history:

Received 10 February 2020

Accepted 12 February 2020

Available online 20 February 2020

Keywords:

Arylamine *N*-Acetyltransferases

Transglutaminases

Active site

Catalytic mechanism

Transition state

ABSTRACT

Transglutaminases (TG) and arylamine *N*-acetyltransferases (NAT) are important family of enzymes. Although they catalyze different reactions and have distinct structures, these two families of enzymes share a spatially conserved catalytic triad (Cys, His, Asp residues). In active TGs, a conserved Trp residue located close to the triad cysteine is crucial for catalysis through stabilization of transition states.

Here, we show that in addition to sharing a similar catalytic triad with TGs, functional NAT enzymes also possess in their active site an aromatic residue (Phe, Tyr or Trp) occupying a structural position similar to the Trp residue of active TGs. More importantly, as observed in active TGs, our data indicates that in functional NAT enzymes this conserved aromatic residue is also involved in stabilization of transition states. These results thus indicate that in addition to the three triad residues, these two families of enzymes also share a spatially conserved aromatic amino acid position important for catalysis. Identification of residues involved in the stabilization of transition states is important to develop potent inhibitors. Interestingly, NAT enzymes have been shown as potential targets of clinical interest.

© 2020 Elsevier Inc. All rights reserved.

1. Introduction

Transglutaminases (TGs) (EC 2.3.2.13) are a family of calcium-dependent enzymes that catalyze the formation of an isopeptide bond between glutamine side chains and the ϵ -amino groups of lysine residue side chains [1–3]. This process results in cross-linking of proteins involved in important biological processes such as coagulation. Crystal structures of TGs have shown that the catalytic center of these enzymes shares common structural properties with the papain-like family of cysteine protease, notably the presence of a catalytic triad of Cys-His-Asp [2,4].

Arylamine *N*-acetyltransferases (NATs) (EC 2.3.1.5) are an important family of xenobiotic-metabolizing enzymes that catalyze the acetyl-coenzyme A (AcCoA)-dependent acetylation of arylamines including drugs and carcinogens [5,6]. NAT enzymes are

found in a range of eukaryotic and prokaryotic species where they may play diverse functions [6–9]. Several eukaryotic and prokaryotic NAT enzymes have been described at both functional and structural levels [9–16]. All these structural studies identified a common fold that comprises three domains and a Cys-His-Asp/Glu catalytic triad in the active site of NAT enzymes [9,14,16,17]. Although TG and NAT enzymes have otherwise distinct 3D structures, catalyze different reactions and assume different functions, the structural studies on TGs and NATs indicate that these two families of enzymes both share a catalytic center and notably a structurally conserved catalytic triad [2,18]. In addition, the catalytic reaction of TG and NAT enzymes relies on ping-pong mechanisms where the cysteine of the triad acts as the catalytic nucleophile residue [2,12,14,18,19]. Further studies on TGs identified a strictly conserved aromatic residue (corresponding to Trp 241 in human TG2) in the active site that is essential for activity of these enzymes [2,4]. This Trp residue is conserved in all enzymatically active TGs and has been shown to serve mainly on the stabilization of transition-states through interaction involving the side chain of

* Corresponding author.

E-mail address: fernando.rodrigues-lima@u-paris.fr (F. Rodrigues-Lima).

¹ These authors contributed equally and should be considered as co-first authors.

the Trp residue [4]. By using different structural and enzymatic approaches, we show that in addition to sharing a similar catalytic triad with TGs, NAT enzymes also possess in their active site an aromatic residue (Phe, Tyr or Trp) occupying a structural position similar to the Trp residue of functional TGs. The identification in active site of NATs of a conserved position involved in stabilization of transition states is of importance to better understand the catalytic mechanisms of this family of enzymes and to design potent inhibitors [20].

2. Materials and methods

2.1. Materials

Acetyl coenzyme A (AcCoA), p-nitrophenylacetate (PNPA), p-dimethylaminobenzaldehyde (DMAB), p-aminosalicylate (PAS), 2-aminofluorene (2-AF) were from Sigma-Aldrich.

2.2. Expression and purification of wild-type *Mesorhizobium loti* NAT1 and mutant enzymes

The pET28a-*M. loti* NAT1 vector encodes the enzyme as a (6xHis) tagged fusion protein and was used as a template to make Phe42-Tyr, Phe42Trp, Phe42His, Phe42Ile and Phe42Ala mutants as described previously [21]. Expression of the NAT proteins in *E. coli* BL21(DE3) and purification was carried out as reported previously [21]. Protein concentrations were determined using the Bradford assay (Bio-rad). Proteins were also analyzed by SDS-PAGE and stained using Coomassie stain.

2.3. NATs activity assays

PNPA assays were carried out in 96-well plates as described previously [8]. Briefly, the mixtures contained NAT enzyme (5 µg/ml final) and p-AS or 2-AF (500 µM final) in a volume of 100 µl of 25 mM Tris-HCl pH 7.5. The reaction was initiated by addition of PNPA (2 mM final) dissolved in DMSO. The rates were determined by monitoring the increase in absorbance at 405 nm using a thermostated (37 °C) plate reader (BioTek, France). The data were corrected by subtracting the non specific hydrolysis of PNPA in the absence of NAT enzyme. All assays were done in triplicates and are shown as mean ± SD.

DMAB assays were also carried out in 96-well plates as described previously [8]. Briefly, 50 µL reaction mixtures, containing NAT enzyme (10 µg/mL), p-AS (0–300 µM) and AcCoA (0–400 µM) in 25 mM Tris-HCl pH 7.5. The reaction was quenched by adding 40 µL of cold trichloroacetic acid (40%). Finally, 100 µL DMAB (2%) in 9:1 acetonitrile and water were added to the mixture. Reaction rate was monitored according to the decrease of absorbance at 450 nm. All assays were done in triplicates and are shown as mean ± SD.

NATs catalyze acetyl transfer using a ping pong bi-bi mechanism. Accordingly, kinetic constants for *M. loti* NAT1 (wild type) and mutants were determined by nonlinear fitting to equation (1) or equation (2) (in case of substrate inhibition behavior at high concentration of substrates) [22]:

$$v = \frac{V_{\max AB}}{K_{AB} + K_{BA} + AB} \quad (1)$$

$$v = \frac{V_{\max AB}}{K_{AB} \left(1 + \frac{B}{K_{iB}} \right) + K_{BA} + AB} \quad (2)$$

where V_{\max} is the maximum velocity; A and B are the concentrations of the two substrates; K_A and K_B are Michaelis-Menten constants for substrate A and B, respectively. K_{iB} is the inhibition constant for p-AS. The overall rate constant k_{cat} was calculated by using $k_{\text{cat}} = V_{\max}/[E_0]$, where $[E_0]$ is total enzyme concentration.

Thermodynamic parameters ($\Delta\Delta G_B$ and $\Delta\Delta G_{TS}$) were calculated as previously described [4,23] (see equations (3) and (4)) and represent, respectively, the difference between WT and mutant enzymes in free energy required to form the enzyme-substrate complex (ES) from E + S, and the difference in free energy between WT and mutants to stabilize the transition-state complex.

$$\Delta\Delta G_B = \Delta G_{Bmut} - \Delta G_{Bwt} = RT \ln \left(\frac{K_{mmut}}{K_{mwt}} \right) \quad (3)$$

$$\Delta\Delta G_{TS} = \Delta G_{TSmut} - \Delta G_{TSwt} = RT \ln \left[\frac{(k_{\text{cat}}/K_m)_{wt}}{(k_{\text{cat}}/K_m)_{mut}} \right] \quad (4)$$

2.4. Circular dichroism (CD) spectroscopy

Circular dichroism analyzes were carried out as described previously [24]. *M. loti* NAT1 and mutants (0.5 mg/ml or 2 mg/ml) were dialyzed against 20 mM potassium phosphate buffer (pH 7.5).

2.5. Structural analyzes

Coordinates of the 3D structures of human tissue transglutaminase 2 (PDB: 1KV3), human tissue transglutaminase 3 (PDB: 1L9M), human factor XIII (PDB: 1QRK), human NAT2 (PDB: 2PFR), *Mycobacterium abscessus* NAT (PDB: 4GUZ) and *Mesorhizobium loti* NAT1 (PDB: 2BSZ) were retrieved from the Protein Data Bank. Structure were aligned, analyzed and visualized using Chimera software [25].

2.6. Gaussian Network Model (GNM) calculation

The GNM of *M. loti* NAT1 structure (PDB: 2BSZ) and human TG2 (PDB: 1KV3) were analyzed by Prody program [26,27]. Kirchhoff matrix was built firstly, with cut-off distance of 7 Å, and gamma value of 1.0. The slowest mode was selected for square fluctuation analysis.

3. Results and discussion

Fig. 1A shows the conservation of the Cys-His-Asp catalytic triad residues of three different active transglutaminase enzymes (including human TG2) that have been structurally aligned. This alignment also clearly indicates the structural conservation of the Trp residue (Trp 241 for human TG2) in the active site of TGs in a position close to the catalytic cysteine side chain (a distance of 5.9 Å between the thiol group of the catalytic cysteine and the Trp side chain). We also aligned the 3D structures of three well characterized and enzymatically active NATs (human NAT2, *Mycobacterium abscessus* NAT and *Mesorhizobium loti* NAT1). As shown in Fig. 1B, a similar structural arrangement of the Cys-His-Asp catalytic triad and of a Phe residue (Phe42 for *M. loti* NAT1) in the active site of these enzymes was found. Moreover, in the aligned NAT structures, this Phe residue (or Trp residue for *M. abscessus* NAT) is also positioned very close to the catalytic cysteine (distance of 5.5 Å between the thiol group of Cys73 and the Phe42 side chain in *M. loti* NAT1) as observed for the Trp residue in TGs structures (Fig. 1A and B). Further structural alignment of the catalytic center of TGs and NAT enzymes (for clarity only human TG2 and *M. loti* NAT1 are

shown in Fig. 1C) supports the observation that the Phe and Trp residues occupy a very similar spatial position in the catalytic site of these two different enzymes (Fig. 1C). In NATs, the Phe residue is part of a highly conserved motif (known as the “PFEN” motif) which participates to the active site pocket of NAT enzymes and contributes to the stability of their structures [28]. Multiple sequence alignment of the amino acid sequences of 41 enzymatically active eukaryotic and prokaryotic NAT enzymes clearly shows the high degree of conservation of this motif (Supplementary Figure 1). More importantly, the alignment indicates that in active NATs, only a Phe, Tyr or Trp residue is present in the motif thus strongly suggesting that an aromatic residue is needed at this position for NAT activity (Fig. 1D and supplementary Figure 1). In active TG enzymes, so far, only a Trp residue has been found at this equivalent position [2,4]. However, it has been shown that Tyr could partially

functionally compensate for Trp in catalysis [4]. As indicated in Fig. 1E, protein flexibility analysis using Gaussian Network Model (GNM) approaches further show that the catalytic triad amino acids and the Phe/Tyr/Trp residue (referred thereafter in the manuscript as “F/Y/W” residue) of the “PFEN” motif of *M. loti* NAT1 display little flexibility (with square fluctuations below 0.1) as observed in TG enzymes (Fig. 1E; for clarity only human TG2 and/or *M. loti* NAT1 square fluctuations per residue number are shown). Reduced flexibility of these catalytic residues is consistent with the need to maintain their spatial arrangement for catalysis (notably for transition state stabilization) [4,12]. As shown previously, in active TG enzymes the conserved Trp residue has been shown to be crucial for stabilization of transition state intermediates [4]. Therefore, our results reported above strongly suggest that the “F/Y/W” residue of the “PFEN” motif of NAT enzymes fulfils a role in catalysis similar to that of the active site Trp residue of functional TGs.

Biochemical and enzymatic analyses were carried out using *M. loti* NAT1 as a model of active NAT enzymes. This NAT enzyme is well characterized both at the structural and functional level and is easily amenable to mutagenesis, heterologous expression in *E. coli*, purification to homogeneity and enzymatic assays [21,24]. In *M. loti* NAT1, the aromatic residue present in the “PFEN” motif is a Phe residue (supplementary Figure 1). Therefore, WT *M. loti* NAT1 and several mutants where the Phe42 residue was mutated to Tyr, Trp, His, Ala, or Ile were expressed in *E. Coli* as done previously with TG2 [4]. No differences in the growth and amount of *E. coli* cells were observed suggesting that no mutant was deleterious for the bacterium. All cultures (1 L of Luria-Berthani medium) yielded ~30 mg of pure and soluble proteins excepted for the F42A mutant (~5 mg protein) which was prone to aggregation at high concentrations (>0.5 mg/ml) (Fig. 2A). In order to evaluate possible structural changes due to the mutations of Phe42, far and near-UV circular dichroism (CD) spectroscopy analysis was carried out with the purified proteins. As shown in Fig. 2B, the mutations had no impact on the secondary structure of the enzyme. Similarly, near-UV CD spectroscopy showed no major differences between WT *M. loti* NAT1 and the mutants, excepted for the F42W for which a small difference in molar ellipticity was observed between 290 and 310 nm due to the introduction of the Trp residue was observed (Fig. 2C). Therefore, these experiments suggest that F42Y, F42W, F42H, and F42I mutant have an overall fold similar to the WT enzyme. This is in agreement with the crystal structure of the F42W *M. loti* NAT1 mutant which has been shown to fold exactly as the WT *M. loti* NAT1 enzyme [24]. As expected from the expression and purification experiments, analysis of the F42A mutant by CD gave inconsistent results due to its tendency to aggregate at high concentrations (>0.5 mg/ml). This further suggests that the presence at position 42 of a residue with a large side chain (larger than Ala) is needed for protein integrity. This is also consistent with data indicating that amino acids of the “PFEN” motif are involved in stabilizing the structure of NAT enzymes [28].

We next tested whether the different mutants were functional using two prototypic NAT arylamine substrates, 2-aminofluorene (2-AF) and *p*-aminosalicylate (*p*-AS). As shown in Fig. 2D, NAT activity was readily measured for WT enzyme (with a Phe at position 42), F42Y and F42W mutants. Conversely, F42H, F42I and F42A mutants displayed very poor activity. Moreover, while F42W mutant had activity similar to those of the WT enzyme, the F42Y mutant displayed activity higher than WT. These data clearly indicate that the presence of an aromatic residue at position 42 is crucial for NAT enzyme activity. This is in agreement with the strict conservation of an aromatic residue at this position in functional NAT enzymes characterized to date (Fig. 1D and supplementary Figure 1).

To better evaluate the role of the F/Y/W residue of the “PFEN”

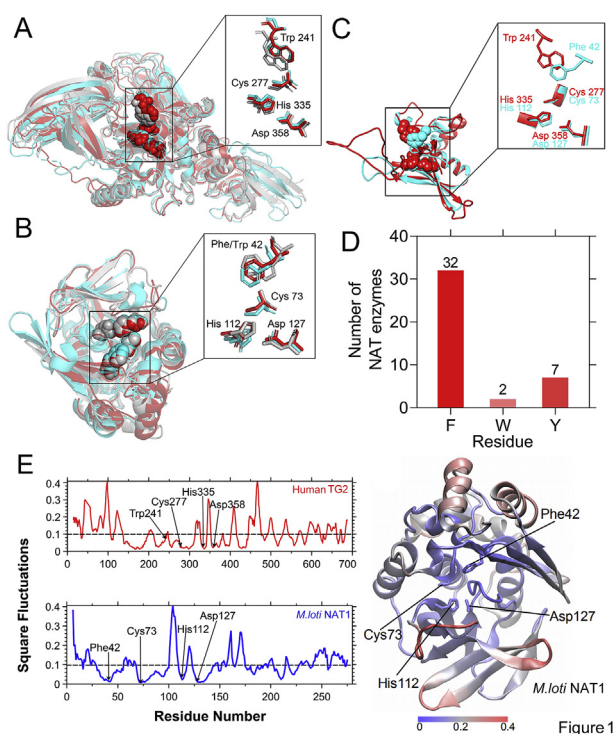


Fig. 1. Structural comparison of enzymatically active transglutaminases and arylamine N-acetyltransferases.

(A) Structural alignment of human TG2 (red color), human TG3 (grey color) and human factor XIII (cyan color). Catalytic triad residues (Cys-His-Asp) and the conserved Trp residue are shown. Numbering corresponds to amino acids of human TG2.

(B) Structural alignment of human *M. loti* NAT1 (red color), *M. abscessus* NAT (grey color) and human NAT2 (cyan color). Catalytic triad residues (Cys-His-Asp) and the conserved Phe42/Trp42 (*M. loti* NAT1 and *M. abscessus* NAT, respectively) are shown.

(C) Structural alignment of the catalytic center of human TG2 (amino acids 231–372, red color) and *M. loti* NAT1 (amino acids 38–138, cyan color). The catalytic triad residues and the Trp241 and Phe42 residues of human TG2 and *M. loti* NAT1 are shown in red and cyan colors, respectively.

(D) Residue type in the “PFEN” motif of functional NAT enzymes. The numbers were calculated from the multiple alignment of the amino acid sequence of the 41 active NAT enzymes (see supplementary Fig. 1). Occurrence of Phe (F), Trp (W) and Tyr (Y) residue at the second position of the “PFEN” motif of NAT enzymes is shown above each bar.

(E) Residue mobility of *M. loti* NAT1 and human TG2 were calculated using Gaussian Network Model (GNM) as described in Materials and Methods. Square fluctuations of human TG2 (PDB 1KV3) and *M. loti* NAT1 (PDB 2BSZ) were generated from the slowed mode of GNM and are shown for each residue. Mean value of all the residues for each protein is drawn in dashed line. The right panel shows a cartoon diagram of *M. loti* NAT1 with color scale of residue mobility (square fluctuations ranging from dark blue to red). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

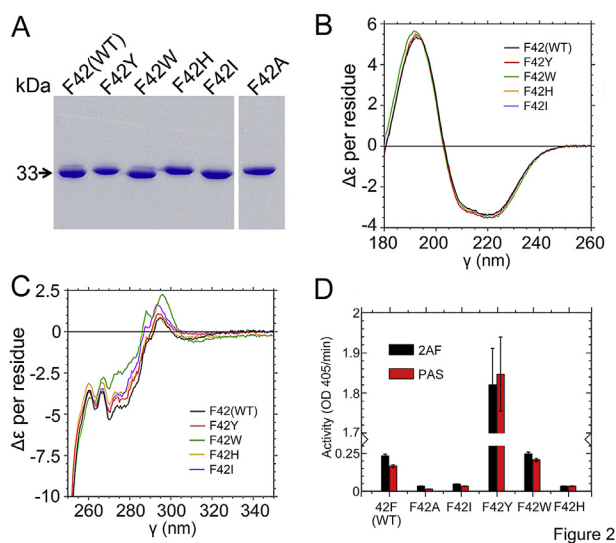


Fig. 2. Heterologous expression, purification and determination of NAT activity of WT *M. loti* NAT1 and F42Y, F42W, F42H, F42I and F42A mutants.

(A) Coomassie staining of 3 µg of WT *M. loti* NAT1, F42Y, F42W, F42H, F42I and F42A mutants.

(B) Far-UV CD spectra (180–260 nm) of WT *M. loti* NAT1, F42Y, F42W, F42H and F42I mutants.

(C) Near-UV CD spectra (250–350 nm) of WT *M. loti* NAT1, F42Y, F42W, F42H and F42I mutants.

(D) Enzymatic activity of WT *M. loti* NAT1, F42Y, F42W, F42H, F42I and F42A mutants towards 2-AF and p-AS substrates. Data are represented as means ± SD of triplicate experiments.

motif of NAT enzymes, notably in transition-state stabilization, kinetic parameters (K_m , k_{cat} , k_{cat}/K_m) were determined for the active *M. loti* NAT1 proteins (WT, F42W and F42Y) using AcCoA and PAS as model substrates. Activity of NAT enzymes occurs via a bi-bi ping pong mechanism, the initial rates were fitted to equation (1), which reflects this (see Materials and Methods and supplementary Figure 2). As substrate inhibition was observed for the F42Y mutant, the equation to which the initial rates for this mutant were fitted includes a substrate inhibition parameter (Equation (2), see Materials and Methods and supplementary Figure 2). Kinetic constants for the three *M. loti* NAT1 proteins are shown in Table 1. All enzymes were found to be highly active with k_{cat} values $> 1 \text{ s}^{-1}$. Notably, the k_{cat} of the F42W and F42Y were 5 and 60 times higher than those of the WT enzyme. The K_m values for AcCoA and PAS were also moderately increased by the mutations (up to 13 times for K_m AcCoA of the F42Y compared to WT). Accordingly, $\Delta\Delta G_B$ values for AcCoA and PAS were also increased (values ranging from 0.57 to 1.86 kcal/mol). This is in agreement with different structural studies on *M. loti* NAT1, *P. aeruginosa* NAT, *M. Abscessus* NAT, *M. Marinum* NAT, *M. smegmatis* NAT and human NAT2 showing that the F/Y/W residue interacts with the cofactor CoA and/or aromatic substrates [10,24,28–30]. Our results are also in agreement with the data obtained in TG enzymes, where substitution of the Trp residue have a greater impact on the catalytic rate constant k_{cat} than on the Michaelis constant K_m parameter [4]. Taking into account the data reported above, the catalytic efficiencies (k_{cat}/K_{mAcCoA} and k_{cat}/K_{mPAS}) of F42W and F42Y were found to be increased compared to WT, notably for the F42Y mutant (k_{cat}/K_{mAcCoA} and k_{cat}/K_{mPAS} 5 and 24 times higher than WT, respectively (Table 1). These increases in catalytic efficiency are largely attributable to increases in k_{cat} rather than decreased in K_m and are associated with reduced $\Delta\Delta G_{TS}$ (–0.91 kcal/mol for the enzyme acetylation step and –1.53 kcal/mol for the PAS acetylation step by the F42Y mutant). Introduction of a hydroxyl and indole group in

Table 1

Kinetic and thermodynamic parameters for PAS acetylation by (RHILO)NAT1 WT and F42W and F42Y mutants.

	WT	F42W	F42Y
k_{cat} (s^{-1})	1.36 ± 0.61	6.86 ± 1.22	82.12 ± 18.14
$K_m(\text{AcCoA})$ (μM)	94.11 ± 43.5	404.8 ± 125.2	1228.23 ± 252.4
$K_m(\text{PAS})$ (μM)	164.03 ± 49.05	665.72 ± 231.5	427.12 ± 150.3
$k_{cat}/K_m(\text{AcCoA})$ ($\text{s}^{-1}\mu\text{M}^{-1}$)	0.014	0.017	0.067
$k_{cat}/K_m(\text{PAS})$ ($\text{s}^{-1}\mu\text{M}^{-1}$)	0.008	0.010	0.192
$\Delta\Delta G_B(\text{AcCoA})$ (kcal/mol)		0.86	1.52
$\Delta\Delta G_B(\text{PAS})$ (kcal/mol)		0.83	0.57
$\Delta\Delta G_{TS}(\text{AcCoA})$ (kcal/mol)		–0.09	–0.91
$\Delta\Delta G_{TS}(\text{PAS})$ (kcal/mol)		–0.13	–1.53

Kinetic constants and $\Delta\Delta G$ values were determined as described in Materials and Methods. Values are mean ± SD of triplicate determinations.

$\Delta\Delta G_B$ = differences between WT and mutant enzymes in free energy required to form the Michaelis complex.

$\Delta\Delta G_{TS}$ = differences between WT and mutant enzymes in free energy required to stabilize the transition-state complexes.

F42Y and F42W mutants may further contribute to stabilization of the transition state intermediate through hydrogen bonds as shown for TG enzymes [4]. Overall, our data indicate that the F/Y/W residue of the ‘PFEN’ motif of NAT enzymes is likely to be involved in transition state stabilization as shown for its counterpart in TG enzymes (Trp241 in human TG2) [4].

Although more modest than its impact on catalysis, the conserved Trp residue of TG enzymes has also been shown to contribute to formation of the Michaelis complex [4]. Accordingly, we also confirmed that the F/Y/W residue of the ‘PFEN’ motif of NATs also contributes to substrate binding (Table 1) [10,24,29–31]. Identification of active site residues involved in the stabilization of transition state residues is important to understand catalytic mechanisms and to develop potent inhibitors [20]. As the inhibition of bacterial NAT enzymes has been shown to be of clinical interest, our data may thus be helpful in the design of NAT inhibitors [9,31,32].

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgements

This work was supported by University Paris Diderot and CNRS. We thank the China Scholarship Council (CSC) for supporting the PhD of WZ, XX and RL. CM is supported by a PhD fellowship from Paris Diderot University (Ecole Doctorale BioSPC). All authors declare they have no conflict of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.bbrc.2020.02.082>.

References

- C.S. Lee, H.H. Park, Structural aspects of transglutaminase 2: functional, structural, and regulatory diversity, *Apoptosis: Int. J. Program. Cell Death* 22 (2017) 1057–1068, <https://doi.org/10.1007/s10495-017-1396-9>.
- S.N. Murthy, S. Iismaa, G. Begg, D.M. Freymann, R.M. Graham, L. Lorand, Conserved tryptophan in the core domain of transglutaminase is essential for catalytic activity, *Proc. Natl. Acad. Sci. U.S.A.* 99 (2002) 2738–2742, <https://doi.org/10.1073/pnas.052715799>.
- H. Sun, M.T. Kaartinen, Transglutaminases in monocytes and macrophages, *Med. Sci.* 6 (2018), <https://doi.org/10.3390/medsci6040115>.
- S.E. Iismaa, S. Holman, M.A. Wouters, L. Lorand, R.M. Graham, A. Husain, Evolutionary specialization of a tryptophan indole group for transition-state stabilization by eukaryotic transglutaminases, *Proc. Natl. Acad. Sci. U.S.A.*

- 100 (2003) 12636–12641, <https://doi.org/10.1073/pnas.1635052100>.
- [5] D.W. Hein, Molecular genetics and function of NAT1 and NAT2: role in aromatic amine metabolism and carcinogenesis, *Mutat. Res.* 506–507 (2002) 65–77, [https://doi.org/10.1016/S0027-5107\(02\)00153-7](https://doi.org/10.1016/S0027-5107(02)00153-7).
- [6] E. Sim, N. Lack, C.J. Wang, H. Long, I. Westwood, E. Fullam, A. Kawamura, Arylamine N-acetyltransferases: structural and functional implications of polymorphisms, *Toxicology* 254 (2008) 170–183, <https://doi.org/10.1016/j.tox.2008.08.022>.
- [7] A.E. Glenn, E.P. Karagianni, A. Ulndreaj, S. Boukouvala, Comparative genomic and phylogenetic investigation of the xenobiotic metabolizing arylamine N-acetyltransferase enzyme family, *FEBS Lett.* 584 (2010) 3158–3164, <https://doi.org/10.1016/j.febslet.2010.05.063>.
- [8] N. Laurieri, J. Dairou, J.E. Egleton, L.A. Stanley, A.J. Russell, J.M. Dupret, E. Sim, F. Rodrigues-Lima, From arylamine N-acetyltransferase to folate-dependent acetyl CoA hydrolase: impact of folic acid on the activity of (HUMAN)NAT1 and its homologue (MOUSE)NAT2, *PLoS One* 9 (2014), e96370, <https://doi.org/10.1371/journal.pone.0096370>.
- [9] E. Sim, G. Fakis, N. Laurieri, S. Boukouvala, Arylamine N-acetyltransferases—from drug metabolism and pharmacogenetics to identification of novel targets for pharmacological intervention, *Adv. Pharmacol.* 63 (2012) 169–205, <https://doi.org/10.1016/B978-0-12-398339-8.00005-7>.
- [10] A.M. Abuhammad, E.D. Lowe, E. Fullam, M. Noble, E.F. Garman, E. Sim, Probing the architecture of the Mycobacterium marinum arylamine N-acetyltransferase active site, *Protein Cell* 1 (2010) 384–392, <https://doi.org/10.1007/s13238-010-0037-7>.
- [11] S.J. Holton, J. Dairou, J. Sandy, F. Rodrigues-Lima, J.M. Dupret, M.E. Noble, E. Sim, Structure of Mesorhizobium loti arylamine N-acetyltransferase 1, *Acta Crystallogr. Sect. F, Struct. Biol. Cryst. Commun.* 61 (2005) 14–16, <https://doi.org/10.1107/S1744309104030659>.
- [12] X. Kubiak, I. Li de la Sierra-Gallay, A.F. Chaffotte, B. Pluvinage, P. Weber, A. Haouz, J.M. Dupret, F. Rodrigues-Lima, Structural and biochemical characterization of an active arylamine N-acetyltransferase possessing a non-canonical Cys-His-Glu catalytic triad, *J. Biol. Chem.* 288 (2013) 22493–22505, <https://doi.org/10.1074/jbc.M113.468595>.
- [13] M. Martins, B. Pluvinage, I. Li de la Sierra-Gallay, F. Barbault, J. Dairou, J.M. Dupret, F. Rodrigues-Lima, Functional and structural characterization of the arylamine N-acetyltransferase from the opportunistic pathogen Nocardia farcinica, *J. Mol. Biol.* 383 (2008) 549–560, <https://doi.org/10.1016/j.jmb.2008.08.035>.
- [14] J.C. Sinclair, J. Sandy, R. Delgoda, E. Sim, M.E. Noble, Structure of arylamine N-acetyltransferase reveals a catalytic triad, *Nat. Struct. Biol.* 7 (2000) 560–564, <https://doi.org/10.1038/76783>.
- [15] I.M. Westwood, S.J. Holton, F. Rodrigues-Lima, J.M. Dupret, S. Bhakta, M.E. Noble, E. Sim, Expression, purification, characterization and structure of Pseudomonas aeruginosa arylamine N-acetyltransferase, *Biochem. J.* 385 (2005) 605–612, <https://doi.org/10.1042/BJ20041330>.
- [16] H. Wu, L. Dombrovsky, W. Tempel, F. Martin, P. Loppnau, G.H. Goodfellow, D.M. Grant, A.N. Plotnikov, Structural basis of substrate-binding specificity of human arylamine N-acetyltransferases, *J. Biol. Chem.* 282 (2007) 30189–30197, <https://doi.org/10.1074/jbc.M704138200>.
- [17] X. Xu, X. Kubiak, J.M. Dupret, F. Rodrigues-Lima, Arylamine N-acetyltransferases: a structural perspective. Comments regarding the BJP paper by Zhou et al. *Br. J. Pharmacol.* 171 (2013) 279–280, <https://doi.org/10.1111/bph.12273>, 2014.
- [18] F. Rodrigues-Lima, C. Delomenie, G.H. Goodfellow, D.M. Grant, J.M. Dupret, Homology modelling and structural analysis of human arylamine N-acetyltransferase NAT1: evidence for the conservation of a cysteine protease catalytic domain and an active-site loop, *Biochem. J.* 356 (2001) 327–334, <https://doi.org/10.1042/0264-6021:3560327>.
- [19] B. Riddle, W.P. Jencks, Acetyl-coenzyme A: arylamine N-acetyltransferase. Role of the acetyl-enzyme intermediate and the effects of substituents on the rate, *J. Biol. Chem.* 246 (1971) 3250–3258.
- [20] R.A. Copeland, Evaluation of enzyme inhibitors in drug discovery. A guide for medicinal chemists and pharmacologists, *Methods Biochem. Anal.* 46 (2005), 1–265.
- [21] N. Atmane, J. Dairou, D. Flatters, M. Martins, B. Pluvinage, P. Derreumaux, J.M. Dupret, F. Rodrigues-Lima, The conserved glycine/alanine residue of the active-site loop containing the putative acetylCoA-binding motif is essential for the overall structural integrity of Mesorhizobium loti arylamine N-acetyltransferase 1, *Biochem. Biophys. Res. Commun.* 361 (2007) 256–262, <https://doi.org/10.1016/j.bbrc.2007.07.034>.
- [22] A. C.-B, *Fundamentals of Enzyme Kinetics*, second ed. ed., Portland Press, London, 1995.
- [23] A.J. Wilkinson, A.R. Fersht, D.M. Blow, G. Winter, Site-directed mutagenesis as a probe of enzyme structure and catalysis: tyrosyl-tRNA synthetase cysteine-35 to glycine-35 mutation, *Biochemistry* 22 (1983) 3581–3586, <https://doi.org/10.1021/bi00284a007>.
- [24] X. Xu, I. Li de la Sierra-Gallay, X. Kubiak, R. Duval, A.F. Chaffotte, J.M. Dupret, A. Haouz, F. Rodrigues-Lima, Insight into cofactor recognition in arylamine N-acetyltransferase enzymes: structure of Mesorhizobium loti arylamine N-acetyltransferase in complex with coenzyme A, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 71 (2015) 266–273, <https://doi.org/10.1107/S139900471402522X>.
- [25] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera—a visualization system for exploratory research and analysis, *J. Comput. Chem.* 25 (2004) 1605–1612, <https://doi.org/10.1002/jcc.20084>.
- [26] A. Bakan, A. Dutta, W. Mao, Y. Liu, C. Chennubhotla, T.R. Lezon, I. Bahar, Evol and ProDy for bridging protein sequence evolution and structural dynamics, *Bioinformatics* 30 (2014) 2681–2683, <https://doi.org/10.1093/bioinformatics/btu336>.
- [27] A. Bakan, L.M. Meireles, I. Bahar, ProDy: protein dynamics inferred from theory and experiments, *Bioinformatics* 27 (2011) 1575–1577, <https://doi.org/10.1093/bioinformatics/btr168>.
- [28] J. Sandy, A. Mushtaq, S.J. Holton, P. Schartau, M.E. Noble, E. Sim, Investigation of the catalytic triad of arylamine N-acetyltransferases: essential residues required for acetyl transfer to arylamines, *Biochem. J.* 390 (2005) 115–123, <https://doi.org/10.1042/BJ20050277>.
- [29] A. Coccagn, X. Kubiak, X. Xu, G. Garnier, I. Li de la Sierra-Gallay, L. Chi-Bui, J. Dairou, F. Busi, A. Abuhammad, A. Haouz, J.M. Dupret, J.L. Herrmann, F. Rodrigues-Lima, Structural and functional characterization of an arylamine N-acetyltransferase from the pathogen Mycobacterium abscessus: differences from other mycobacterial isoforms and implications for selective inhibition, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 70 (2014) 3066–3079, <https://doi.org/10.1107/S1399004714021282>.
- [30] E. Fullam, I.M. Westwood, M.C. Anderton, E.D. Lowe, E. Sim, M.E. Noble, Divergence of cofactor recognition across evolution: coenzyme A binding in a prokaryotic arylamine N-acetyltransferase, *J. Mol. Biol.* 375 (2008) 178–191, <https://doi.org/10.1016/j.jmb.2007.10.019>.
- [31] J. Sandy, S. Holton, E. Fullam, E. Sim, M. Noble, Binding of the anti-tubercular drug isoniazid to the arylamine N-acetyltransferase protein from Mycobacterium smegmatis, *Protein Sci. : Publ. Protein Soc.* 14 (2005) 775–782, <https://doi.org/10.1110/ps.041163505>.
- [32] B. Pluvinage, J. Dairou, O. Possot, M. Martins, A. Fouet, J.M. Dupret, F. Rodrigues-Lima, Cloning and molecular characterization of three arylamine N-acetyltransferase genes from Bacillus anthracis: identification of unusual enzymatic properties and their contribution to sulfamethoxazole resistance, *Biochemistry* 46 (23) (2007) 7069–7078.

Supplementary file

From transglutaminases (TGs) to arylamine *N*-acetyltransferases (NATs): Insight into the role of a spatially conserved aromatic amino acid position in the active site of these two families of enzymes

Ximing Xu^{1,2,#}, Wenchao Zhang^{1,#}, Jérémy Berthelet¹, Rongxing Liu¹, Christina Michail¹, Alain F. Chaffotte³, Jean-Marie Dupret¹ and Fernando Rodrigues-Lima^{1,*}

¹ Université de Paris, BFA, UMR 8251, CNRS, F-75013, Paris, France

² Key Laboratory of Marine Drugs, Chinese Ministry of Education, School of Medicine and Pharmacy, Ocean University of China, Qingdao, 266003, China

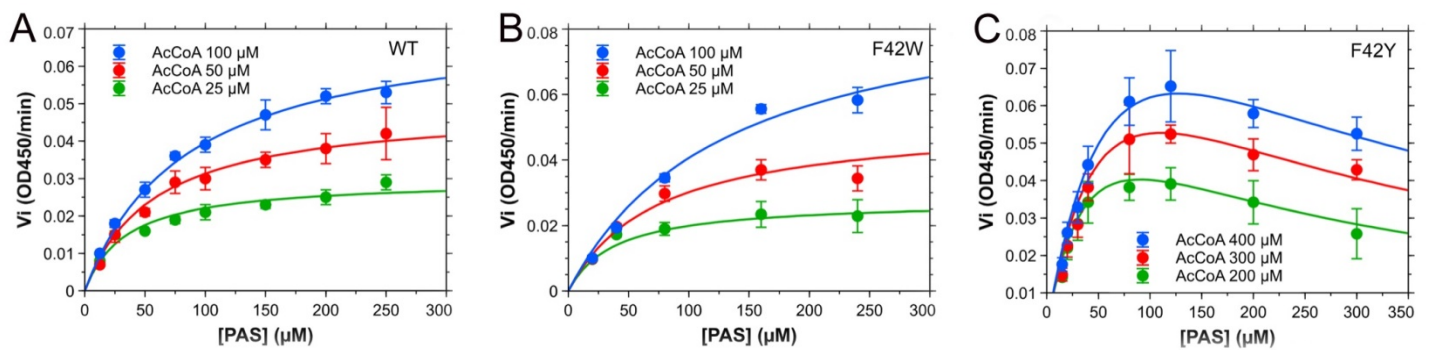
³ Institut Pasteur, Unité de Résonance Magnétique Nucléaire des Biomolécules, Paris, France

Supplementary Figure 1: Multiple alignment showing the conservation of the “PFEN” motif from 41 functional NAT enzymes (from eukaryotes and prokaryotes organisms). *M. loti* NAT1 (RHILONAT1) corresponds to the first sequence. The conserved F/Y/W position at position 42 in *M. loti* NAT1 is boxed. Full descriptions of abbreviated names are indicated below.

	42
RHILONAT1	AHPQAI PFEN IDPFL-G--RPVR-LDLA
MYCMMNAT	AHTGAI PFEN LDPLM-G--VPVDDLSAE
SALTYNAT	AHNCAI PFEN LDVLL-P--REIQ-LDET
MYCS2NAT	AHNRSI PFEN LDPLL-G--IPVADLSAE
PSEAEENAT	AHLRRVA PFEN LDVLLDR----PIEIDAD
ECOLXNAT	KHNCTI PFEN LDVLL-P--REIQ-LDNQ
MYCBONAT	VHSRTI PFEN LDPLL-G--VPVDDLSPQ
MYCTUNAT	VHSRTI PFEN LDPLL-G--VPVDDLSPQ
KLEP7NAT	QHNSAI PFEN LDVLL-P--REIH-LDDG
MYCA1NAT	AHTGSI PFEN LDPLM-G--VPVDDLSPA
BACANNAT	TFAHTI PFEN LDVIA-S---NTNTITME
RHILONAT2	AQMRAI PFEN IDVLL-G---DIPNLTEN
BACCENAT1	SFAHTI PFEN LDVIA-R---NTNQISLE
BACCENAT2	AMAQTV PFEN LNILE-K---NFKEISKE
BACCENAT3	KMGMI LPYEN LDIMA-G---TIKNISKD
BACANNAT1	AMAQTV PFEN LVLE-K---NFKEISQE
BACANNAT2	TFAHTI PFEN LDVIA-S---NTNTITME
BACANNAT3	NMGMI LPYEN LDIMA-G---TIKNISKN
NOCFANAT	RHTTAI PFEN LEAVL-G--RPVR-LDLA
LEGPANAT	AHIKVP YSNFEL RRIANQHPVQRQSL
MYCABSAT	GHVLNI KWEN LDVAVL-H--KHVA-LDIP
TSUPDNAT1	AHATTL PWEN IDA-----V-----
HYPJQNAT1	HHLARV PFES IGLHYSS--HRTLSIDPD
HYPVGNAT1	HHLARV PFES IGLHYSP--HRILSLDPD
HYPVGNAT2	AQITTF PFYEN LSVHYSS--THLVDIQPD
PODANNAT2	LHMARV PFES LSLHYSK--YRTLNLNPE
PODANNAT1	AQITSF PFYEN LSIHYNP--FHRILLDPP
HUMANNAT1	HQIRAV PFEN LNIIHC-G---DAMDGLLE
HUMANNAT2	HQIRAV PFEN LNMHHC-G---QAMELGLE
MOUSENAT1	HQMRAV PFEN LNMHHC-G---EAMHDLQ
MOUSENAT2	HQIRAI PFEN LNIIHC-G---ESMELSLE
MOUSENAT3	HQIRAI PFEN LNIIHC-G---KTMELSLE
RATNAT1	HQMRAV PFEN LSMHHC-G---EAMCLGLE
RATNAT2	HQIRAI PFEN LNIIHC-G---ESMELNLE
RABITNAT1	HQIRTV PFYEN LSIHC-G---ESMELDLE
RABITNAT2	HQIRTV PFYEN LSIHC-G---ESMELDLE
CHICKNAT1	HHIRAV PFEN LSIHC-G---EKITLELE
CHICKNAT2	HHIQAV PFEN LSIHC-G---ETIELDLA
CHICKNAT3	HHIQAI PFEN LSMHHC-G---ETIDLQ
MESAUNAT1	HQMRTI PFEN LNMHHC-G---EAMDGLLE
MACMUNAT1	HQIRAV PFEN LNIIHC-G---EAMDGLLE
MACMUNAT2	HQIRAV PFEN LNMHHC-G---EAMELGLE
	. :...:
Prim. cons.	HHIRAI PFEN LDIHCSGNQREVMELDLE

MYCMMNAT, *Mycobacterium marinum* NAT; SALTYNAT, *Salmonella typhimurium* NAT; MYCS2NAT, *Mycobacterium smegmatis* NAT; PSEAENAT, NAT *Pseudomonas aeruginosa*; ECOLXNAT, *Escherichia coli* NAT; MYCBONAT, *Mycobacterium bovis* NAT; MYCTUNAT, *Mycobacterium tuberculosis* NAT; KLEP7NAT, *Klebsiella pneumoniae subsp pneumoniae* NAT; MYCA1NAT, *Mycobacterium avium* NAT; RHILONAT1, *Mesorhizobium loti* NAT1; RHILONAT2, *Mesorhizobium loti* NAT2; BACCENAT1, *Bacillus cereus* NAT1; BACCENAT2, *Bacillus cereus* NAT2; BACCENAT3, *Bacillus cereus* NAT3; BACANNAT1, *Bacillus anthracis* NAT1; BACANNAT2, *Bacillus anthracis* NAT2; BACANNAT3, *Bacillus anthracis* NAT3; NOCFANAT, *Nocardia farcinica* NAT; LEGPANAT, *Legionella pneumophila str.Paris* NAT; HYPJQNAT1, *Trichoderma reesei* NAT1; HYPVGNAT2, *Trichoderma virens* NAT2; HYPVGNAT1, *Trichoderma virens* NAT1; PODANNAT1, *Podospora anserina* NAT1; PODANNAT2, *Podospora* NAT2; MYCABSNTAT, *Mycobacterium abscessu* NAT; TSUPDNNAT1, *Sukamurella paurometabola* NAT1; HUMANNAT1, *Homo sapiens* NAT1; HUMANNAT2, *Homo sapiens* NAT2; HUMANNAT3, *Homo sapiens* NAT3; MOUSENAT1, *Musculus* NAT1; MOUSENAT2, *Musculus castaneus* NAT2; MOUSENAT3, *Musculus* NAT3; RATNAT1, *Rattus norvegicus* NAT1; RATNAT2, *Rattus norvegicus* NAT2; RABITNAT1, *Oryctolagus cuniculus* NAT1; RABITNAT2, *Oryctolagus cuniculus* NAT2; CHICKNAT1, *Gallus gallus* NAT1; CHICKNAT2, *Gallus gallus* NAT2; CHICKNAT3, *Gallus gallus* NAT3; MESAUNAT1, *Mesocricetus auratus* NAT1; MACMUNAT1, *Macaca mulatta* NAT1; MACMUNAT2, *Macaca mulatta* NAT2.

Supplementary Figure 2: Kinetics analysis of WT *M. loti* NAT1, F42W and F42Y mutants determined with AcCoA and p-AS. Initial velocities were fitted to equations 1 or 2 (For F42Y) and the kinetic parameters k_{cat} , $K_{m(AcCoA)}$ and $K_{m(PAS)}$ determined. Assays were carried in triplicates and data are shown as mean \pm SD.



The Benzene Hematotoxic and Reactive Metabolite 1,4-Benzoquinone Impairs the Activity of the Histone Methyltransferase SET Domain Containing 2 (SETD2) and Causes Aberrant Histone H3 Lysine 36 Trimethylation (H3K36me3)[§]

Jérémy Berthelet¹, Christina Michail, Linh-Chi Bui, Louise Le Coadou, Valentina Sirri, Li Wang, Nicolas Dulphy, Jean-Marie Dupret, Christine Chomienne, Fabien Guidez, and Fernando Rodrigues-Lima

Université de Paris, Unité de Biologie Fonctionnelle et Adaptative (BFA), UMR 8251, CNRS, Paris, France (J.B., C.M., L.-C.B., L.L.C., V.S., J.-M.D., F.R.-L.); The First Affiliated Hospital of Chongqing Medical University, Department of Hematology, Chongqing, China (L.W.); Université de Paris, Institut de Recherche Saint-Louis (IRSL), UMRS 1160 (N.D.), UMRS 1131 (C.C., F.G.), INSERM, Paris, France; and Service de Biologie Cellulaire, Assistance Publique des Hôpitaux de Paris (AP-HP), Hôpital Saint Louis, Paris, France (C.C.)

Received April 15, 2021; accepted June 28, 2021

ABSTRACT

Human SETD2 is the unique histone methyltransferase that generates H3K36 trimethylation (H3K36me3), an epigenetic mark that plays a key role in normal hematopoiesis. Interestingly, recurrent inactivating mutations of SETD2 and aberrant H3K36me3 are increasingly reported to be involved in hematopoietic malignancies. Benzene (BZ) is a ubiquitous environmental pollutant and carcinogen that causes leukemia. The leukemogenic properties of BZ depend on its biotransformation in the bone marrow into oxidative metabolites, in particular 1,4-benzoquinone (BQ). This hematotoxic metabolite can form DNA and protein adducts that result in the damage and the alteration of cellular processes. Recent studies suggest that BZ-dependent leukemogenesis could depend on epigenetic perturbations, notably aberrant histone methylation. We investigated whether H3K36 trimethylation by SETD2 could be impacted by BZ and its hematotoxic metabolites. Herein, we show that BQ, the major leukemogenic metabolite of BZ, inhibits irreversibly the human histone methyltransferase SETD2, resulting in decreased H3K36me3. Our mechanistic studies further indicate that the BQ-dependent inactivation of SETD2 is due to covalent binding of BQ to reactive Zn-finger cysteines within the catalytic domain

of the enzyme. The formation of these quinoprotein adducts results in loss of enzyme activity and protein crosslinks/oligomers. Experiments conducted in hematopoietic cells confirm that exposure to BQ results in the formation of SETD2 crosslinks/oligomers and concomitant loss of H3K36me3 in cells. Taken together, our data indicate that BQ, a major hematotoxic metabolite of BZ, could contribute to BZ-dependent leukemogenesis by perturbing the functions of SETD2, a histone lysine methyltransferase of hematopoietic relevance.

SIGNIFICANCE STATEMENT

Benzoquinone is a major leukemogenic metabolite of benzene. Dysregulation of histone methyltransferase is involved in hematopoietic malignancies. This study found that benzoquinone irreversibly impairs SET domain containing 2, a histone H3K36 methyltransferase that plays a key role in hematopoiesis. Benzoquinone forms covalent adducts on Zn-finger cysteines within the catalytic site, leading to loss of activity, protein crosslinks/oligomers, and concomitant decrease of H3K36me3 histone mark. These data provide evidence that a leukemogenic metabolite of benzene can impair a key epigenetic enzyme.

Introduction

Benzene (BZ) is an aromatic compound of industrial importance that is predominantly used as a solvent or as a primary material for chemical synthesis. BZ is also a carcinogen considered to be a ubiquitous environmental pollutant (Smith, 2010). Exposure to BZ is indeed a well established cause of hematopoietic malignancies (Smith et al., 2011; McHale et al., 2012; Snyder, 2012; Eastmond et al., 2014). The leukemogenic properties of BZ are known to rely on its metabolism in bone marrow cells into oxidative species, notably 1,4-benzoquinone (BQ) (Smith, 2010). This oxidative phenolic species is thought to be one of the major hematotoxic metabolites of BZ (Frantz et al., 1996; Whysner et al., 2004; Smith, 2010; Holmes and

This work was supported by running grants from Université de Paris and Centre National de la Recherche Scientifique, and by grants from ANSES (Agence Nationale de Sécurité Sanitaire de l'Alimentation, de l'Environnement et du Travail) and Instituts Thématiques MultiOrganismes Cancer (plan Cancer-Environnement). J.B. was supported by a Ph.D. fellowship from Cancerpole IDF (Région Ile de France), C.M. and L.L.C. are supported by Ph.D. fellowships from Université de Paris (Ecole Doctorale BioSPC).

The authors declare no conflict of interest.

¹Current affiliation: Université de Paris, Centre Epigénétique et Destin Cellulaire (CEDC), UMR 7216, CNRS, Paris, France.
<https://dx.doi.org/10.1124/molpharm.121.000303>.

[§] This article has supplemental material available at molpharm.aspetjournals.org.

Winn, 2019; North et al., 2020). BQ is known to be particularly reactive and capable of binding proteins through Michael addition on certain reactive cysteine residues (Wang et al., 2006; Bolton and Dunlap, 2017). Although the mechanisms by which BZ induces hematologic malignancies remain poorly understood, different studies indicate that BZ metabolites act through multiple modes of action (McHale et al., 2012; Sauer et al., 2018; North et al., 2020). In particular, inhibition of topoisomerases II by BZ metabolites and subsequent chromosomal damage, alteration of cell signaling pathways, and immune-mediated bone marrow dysfunctions have been reported (Frantz et al., 1996; Lindsey et al., 2004; Sauer et al., 2018; Duval et al., 2019; Lu et al., 2020; North et al., 2020). Increasing evidence indicates that BZ induces epigenetic changes, leading to aberrant DNA and histone methylation, and that epigenetic alterations contribute to BZ-dependent leukemogenesis (Smith, 2010; Chappell et al., 2016; Fenga et al., 2016; Yu et al., 2019; Chung and Herceg, 2020). Interestingly, normal and malignant hematopoiesis is known to rely on epigenetic processes such as histone modifications (Butler and Dent, 2013; Chopra and Bohlander, 2015). As such, epigenetic enzymes, notably histone methyltransferases, are recurrently mutated and dysregulated in hematologic malignancies (Butler and Dent, 2013; Husmann and Gozani, 2019).

The histone mark H3K36me3 is a key epigenetic modification that is linked to transcription, alternative splicing, and DNA repair (Wagner and Carpenter, 2012; Li et al., 2016; Fahey and Davis, 2017; Husmann and Gozani, 2019). The nonredundant histone methyltransferase SETD2 is the sole methyltransferase that mediates trimethylation of H3K36 to generate H3K36me3 (Edmunds et al., 2008; Husmann and Gozani, 2019). Recurrent SETD2-inactivating mutations and altered H3K36me3 levels are found in cancer at a high frequency, and several studies suggest that SETD2 acts as a tumor suppressor (Kudithipudi and Jeltsch, 2014; Fahey and Davis, 2017; Husmann and Gozani, 2019). Importantly, SETD2 and the H3K36me3 mark play a crucial role in hematopoiesis (Wang et al., 2018; Zhang et al., 2018; Zhou et al., 2018; Chen et al., 2020). Accordingly, loss-of-function mutations in SETD2 and subsequent aberrant H3K36 trimethylation are among the most common alterations found in blood malignancies, supporting that disruption of the SETD2-dependent H3K36me3 mark contributes to malignant hematopoiesis (Huether et al., 2014; Mar et al., 2014, 2017; Zhu et al., 2014; Dong et al., 2019). Using molecular and cellular approaches, we provide here mechanistic evidence that human SETD2 is inactivated by BQ, leading to reduced levels of H3K36me3. More specifically, we found that loss of SETD2 methyltransferase activity is due to covalent binding of BQ to cysteine residues within the catalytic domain of SETD2, in particular Zn²⁺-bound cysteines located in the AWS zinc finger. We also observed that generation of BQ-SETD2 quinoproteins also resulted in the formation of SETD2 crosslinks/oligomers, in agreement with the aggregation-prone properties of SETD2 and previous studies with

other enzymes (Shu et al., 2019, 2020a). Taken together, our data suggest that the hematotoxic metabolite of BZ, 1,4-benzoquinone, may contribute to BZ-dependent leukemogenesis by altering the activity of the H3K36 methyltransferase SETD2. More broadly, our work provides molecular and cellular evidence that certain hematotoxic metabolites of benzene could directly impact epigenetic processes of hematopoietic relevance.

Materials and Methods

Chemicals and Reagents. Antibodies used in this paper are as follows: H3K36me3 (ab9050; Abcam), SETD2 (PA5-34934; ThermoFisher), 6× histidine-tag (H1029; Sigma-Aldrich, France), GFP-tag (B2; Santa Cruz), plumbagin [kind gift of Dr. Hiroyuki Tanaka (Kyushu University)], H3 (3638S; Cell Signaling, France), γ -H2AX (CR55T33; eBioscience), secondary anti-rabbit (A1949; Sigma-Aldrich), and secondary anti-mouse (31430; ThermoFisher). 1,4-Benzoquinone, hydroquinone, phenol, benzene, plumbagin, nitro blue tetrazolium, iodoacetamide-fluorescein, S-adenosylmethionine, 4-(2-pyridylazo)resorcinol, *N,N,N',N'*-tetrakis(2-pyridinylmethyl)-1,2-ethanediamine, *N*-ethylmaleimide, dithiothreitol, dimethylsulfoxide, β -mercaptoethanol, His-select nickel resin, protease inhibitors, lysozyme, isopropyl β -D-1-thiogalactopyranoside, formaldehyde, Triton X-100, and horseradish peroxidase were from Sigma-Aldrich (France). pET28-MHL SETD2 (1435–1711) plasmid was from Addgene (40741), and GFP-SETD2 (500-2564) plasmid was a kind gift of Dr. Sérgio De Almeida (Faculdade de Medicina da Universidade de Lisboa). PD-10 desalting column was from GE Healthcare (52-1308-00), protein A-agarose and G-agarose were from Santa Cruz (sc2001 and sc-2002), Metafectene was from Biontex, and Bradford reagent was from Bio-Rad. H3K36 peptide (5-FAM-TGGVKRPHR-NH₂) and its methylated form, H3K36me (5-FAM-TGGVK_{me}RPHR-NH₂), were from Proteogenix. 1,4-Benzoquinone, hydroquinone, phenol, benzene, and plumbagin were diluted in DMSO at a stock concentration of 150 mM.

Recombinant Protein Expression and Purification. cDNA sequence from human SETD2 encoding for the AWS domain (residues 1494–1550) was cloned into pET28a expression vector. pET28-MHL SETD2 and pET28a AWS domain were transformed into HI-Control BL21 (DE3) *Escherichia coli*, and bacteria were grown in Luria-Broth medium at 37°C under agitation until reaching an optical density of 0.6. Protein expression was then induced by adding 750 μ M isopropyl β -D-1-thiogalactopyranoside and lowering the temperature to 16°C overnight under agitation. Cells were pelleted and resuspended in lysis buffer (PBS, pH 8, 300 mM NaCl, 1% Triton X-100, 1 mg/ml lysozyme and protease inhibitors) for 30 minutes at 4°C. The resulting lysate was sonicated (10 minutes, 10 seconds ON, 20 seconds OFF, 20% power) and clarified by centrifugation (15000g, 30 minutes, 4°C). The supernatant was then incubated for 2 hours on ice in the presence of 10 mM imidazole and 1 ml His-select nickel resin under agitation. Beads were loaded on a column equilibrated with PBS, pH 8, and 300 mM NaCl and washed with washing buffer (PBS, pH 8, 300 mM NaCl, 0.1% Triton X-100) followed by a second wash with washing buffer without Triton X-100. Protein was eluted using elution buffer (PBS, pH 8, 300 mM NaCl, and 300 mM imidazole), and highly concentrated fractions were pooled and incubated on ice for 30 minutes in the presence of 10 mM dithiothreitol (DTT). Purified proteins were buffer-exchanged using PD-10 desalting columns to methyltransferase buffer (50 mM Tris-HCl, 50 mM NaCl, pH 8), and its purity was determined by SDS-PAGE and Coomassie

ABBREVIATIONS: AWS, associated with SET; BQ, 1,4-benzoquinone; BZ, benzene; CDK2, Cyclin-Dependent Kinase 2; ChIP, chromatin immunoprecipitation; Ctrl, control; DAPI, 4',6-diamidino-2-phenylindole; DTT, dithiothreitol; FA, formic acid; GFP, green fluorescent protein; H3K36, lysine 36 of histone H3; H3K36me3, trimethylated histone H3 lysine 36; IAF, iodoacetamide-fluorescein; MYC, Myelocytomatosis; NBT, nitroblue tetrazolium; NEM, *N*-ethylmaleimide; PAR, 4-(2-pyridylazo)resorcinol; PBG, plumbagin; PH, phenol; PO, peroxidase; qPCR, quantitative polymerase chain reaction; SAM, S-adenosylmethionine; SET, Su(var)3-9, Enhancer-of-zeste and Trithorax; SETD2, SET domain containing 2; UFLC, ultra-fast liquid chromatography system.

staining. Proteins were kept at -80°C . Before use, recombinant SETD2 or AWS domain were incubated in the presence of 2 mM DTT for 15 minutes on ice. Proteins were then buffer-exchanged using PD-10 desalting columns to methyltransferase buffer, and protein concentration was assessed by Bradford assay.

UFLC-Mediated SETD2 Methyltransferase Activity Assay. The H3K36 methyltransferase activity of SETD2 was determined as previously described (Duval et al., 2015). Recombinant SETD2 was incubated with 75 μM H3K36 peptide and 75 μM *S*-adenosylmethionine (SAM) in 50 μl of methyltransferase buffer overnight at room temperature. Reaction was then stopped with the addition of 50 μl perchloric acid (HClO_4) (15% v/v in water), and 10- μl aliquots of the mixture were injected in a reverse phase ultra-fast liquid chromatography system (UFLC) (Shimadzu, France). The mobile phase used for the separation consisted of two solvents: A was water with 0.1% HClO_4 , and B was acetonitrile with 0.1% trifluoroacetic acid. Separation was performed by an isocratic flow (80% A/20% B) rate of 1 ml/min on a Kromasil column 100-5-C18 (4.6 \times 250 mm at 40°C). H3K36 peptide (substrate) and its methylated form (H3K36me, product) were monitored by fluorescence emission ($\lambda = 530$ nm) after excitation at $\lambda = 485$ nm and quantified by integration of the peak fluorescence area.

Effects of Benzene and Benzene Metabolites on SETD2 Activity. Recombinant SETD2 (3 μM) was incubated with increasing concentrations of benzene (BZ), phenol (PH), hydroquinone (HQ) (25, 50, 100 μM), BQ (3, 6, 12 μM), or vehicle (Ctrl) for 10 minutes at room temperature. In total, 10 μl of the mixture was then diluted with 40 μl methyltransferase buffer containing 75 μM H3K36 peptide and 75 μM SAM (final concentrations). SETD2 H3K36 methyltransferase activity was assayed overnight at room temperature and was assessed by UFLC as described above.

Determination of SETD2 H3K36 Methyltransferase Activity on Core Histone Substrates. Recombinant SETD2 (3 μM) was incubated with 30 μM BZ, PH, HQ, BQ, or vehicle (Ctrl) for 10 minutes at room temperature in methyltransferase buffer. In total, 10 μl of the mixture was then diluted with 10 μl methyltransferase buffer containing 1 μg purified human core histones [from SETD2-KO HEK293 cells (Hacker et al., 2016)] and 100 μM SAM (final concentrations) and incubated overnight at room temperature. Samples were then separated by SDS-PAGE (18% gel) and transferred onto a nitrocellulose membrane. SETD2-dependent methylation of H3K36 was detected using an anti H3K36me3 antibody. Ponceau staining was used to ensure equal protein loading.

Effects of Buffer Exchange and DTT Reducing Agent on BQ-Inactivated SETD2. To test the effect of buffer exchange on BQ-inactivated SETD2, recombinant SETD2 (3 μM) was incubated with 30 μM BQ or vehicle (Ctrl) for 10 minutes at room temperature in methyltransferase buffer. Samples were then buffer-exchanged using a PD-10 column to methyltransferase buffer. SETD2 H3K36 methyltransferase activity was assessed by UFLC as described above.

To test the effect of dithiothreitol (DTT) on BQ-inactivated SETD2, recombinant SETD2 (3 μM) was incubated with 30 μM BQ or vehicle (Ctrl) for 10 minutes at room temperature, followed by a second incubation in the presence of 10 mM DTT for 10 minutes at room temperature in methyltransferase buffer. SETD2 H3K36 methyltransferase activity was assessed by UFLC as described above.

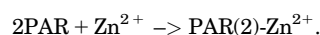
Detection of Quinone-Protein Adducts by Nitroblue Tetrazolium. Detection of quinone-protein adducts were carried out using the nitroblue tetrazolium (NBT) method as reported previously (Shu et al., 2019). Briefly, recombinant SETD2 (5 μg) or recombinant AWS domain (5 μg) were incubated with increasing concentration of BQ (7, 15, 30 μM), plumbagin (PBG; 30 μM), or vehicle (Ctrl) for 10 minutes (BQ) or 15 minutes (PBG) at room temperature in methyltransferase buffer. Reactions were stopped with the addition of Laemmli sample buffer containing 400 mM β -mercaptoethanol, and samples were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. Membrane was then incubated with a solution of 2 M potassium glycinate, 0.6 mg/ml NBT (pH 10.0) at room temperature for 45 minutes in the dark. Upon purple precipitate

apparition, membrane was further blocked with 5% nonfat milk in phosphate buffer containing 1% Tween-20 (PBST) for 1 hour at room temperature and later used for immunodetection of SETD2 protein or ponceau staining.

Labeling of Free Cysteine Residues Using Iodoacetamide-Fluorescein. Labeling of free cysteine residues with iodoacetamide-fluorescein (IAF) was carried out as reported previously (Duval et al., 2019). Recombinant SETD2 (5 μg) or recombinant AWS domain (5 μg) were incubated in the increasing concentration of BQ (0, 7, 15, 30 μM), PBG (30 μM), or vehicle (Ctrl) for 10 minutes (BQ) or 15 minutes (PBG) at room temperature in methyltransferase buffer. In total, 20 μM IAF was added to the mixture, and samples were further incubated for 15 minutes at 37°C . Reactions were stopped by the addition of Laemmli sample buffer containing 400 mM β -mercaptoethanol, and samples were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. IAF fluorescence on labeled cysteine residues was detected using a Fujifilm LAS 4000 detection system ($\lambda_{\text{exc}} = 485$ nm, $\lambda_{\text{em}} = 530$ nm). A ponceau staining of the membrane was done to ensure equal protein loadings.

Kinetics of SETD2 Inactivation by 1,4-Benzoquinone. Recombinant SETD2 (5 μM) was incubated with increasing concentrations of BQ (0, 20, 30, 40 μM) on ice in methyltransferase buffer. Every 30 seconds, an aliquot of the mixture was taken and diluted 10 times, and SETD2 residual activity was measured by UFLC as described above. Data were analyzed as described previously for irreversible inhibitors under pseudo first-order conditions using Origin-Pro software (version 2021; OriginLab Corporation, Northampton, MA) (Copeland, 2005). The equation rate of inhibition of SETD2 by BQ can be represented as follows: $\text{Ln}(\% \text{ residual SETD2 activity}) = -k_{\text{obs}} \times t$, where k_{obs} is apparent first-order inhibition rate constant, and t is time. For each BQ concentration, the k_{obs} values were extracted from the slopes of the natural logarithm of the percentage of SETD2 activity as a function of time. The second-order rate constant (k_{inact}) was obtained from the slope of k_{obs} values plotted as a function of BQ concentration ($k_{\text{inact}} = k_{\text{obs}}/[\text{BQ}]$).

PAR-Mediated Zn^{2+} Release Analysis. Recombinant SETD2 (5 μg) or recombinant AWS domain (5 μg) were incubated with 30 μM BZ, PH, HQ, BQ, PBG, 1 mM *N*-ethylmaleimide (NEM) or vehicle (Ctrl) in the presence of 100 μM 4-(2-pyridylazo)resorcinol (PAR) for 20 minutes at room temperature in methyltransferase buffer. PAR reacts with free Zn^{2+} in solution according to the following equation:



The orange-colored complex $\text{PAR}(2)\text{-Zn}^{2+}$ has an absorption peak at 490 nm. The reaction was started by the addition of benzene, benzene metabolites, PBG, or NEM and was monitored by reading the absorbance of the solution every minute at 490 nm using a microplate reader (Biotek Instruments, France).

Mass Spectrometry Analysis of Plumbagin Adducts. Recombinant SETD2 catalytic domain or AWS Zn-finger domain (3 μM) were incubated with 60 μM PBG for 30 minutes at room temperature in assay buffer. The reaction was stopped with 10 mM DTT and diluted 10 times in methyltransferase buffer, and unmodified cysteines were blocked by addition of 10 mM *N*-ethylmaleimide for 10 minutes. Samples were then incubated overnight at 37°C with trypsin (Promega, France) at 12.5 ng/ μl in 25 mM ammonium bicarbonate (pH 8.0). The supernatant containing peptides was acidified with formic acid (FA), desalted on C18 tips (Pierce C18 tips, Thermo Scientific, France), and eluted in 10 μl 70% ACN, 0.1% FA. Desalted samples were evaporated using a SpeedVac and then taken up in 10 μl of buffer A (water, 0.1% FA), and 5 μl was injected on a nanoLC High performance liquid chromatography (HPLC) system (Thermo Scientific, France) coupled to a hybrid quadrupole-Orbitrap mass spectrometer (Thermo Scientific, France). Peptides were loaded on a reverse phase C18 μ -precolumn (C18 PepMap100, 5 μm , 100 A, 300 μm i.d. \times 5 mm) and separated on a C18 column (EASY-spray C18 column, 75 μm i.d. \times 50 cm) at a constant flow rate of 300 nl/min

with a 120-minute gradient of 2% to 40% buffer B (buffer B: 20% water, 80% ACN, 0.1% FA). Mass spectrometry (MS) analyses were performed by the Orbitrap cell with a resolution of 120,000 [at m/z (mass number/charge number) 200]. Tandem mass spectrometry (MS/MS) fragments were obtained by Higher-energy C-trap dissociation (HCD) activation (collisional energy of 28%) and acquired in the ion trap in top-speed mode with a total cycle of 3 seconds. The database search was performed against the Swissprot database and the *Homo sapiens* taxonomy with Mascot version 2.5.1 software with the following parameters: tryptic peptides only with up to two missed cleavages, variable modifications: cysteine plumbagin and methionine oxidation. MS and MS/MS error tolerances were set respectively to 7 ppm and 0.5 Da. Peptide identifications were validated using a 1% false discovery rate threshold obtained by Proteome Discoverer (version 2.2; Thermo Scientific, France) and the percolator algorithm. The candidate sequences modified by plumbagin were manually inspected for de novo sequencing.

Bioactivation of HQ into BQ Using Peroxidase Hydroquinone-Bioactivation System. Bioactivation of HQ into BQ by peroxidase (PO) was carried out as previously described (Eastmond et al., 2005). Horseradish PO (0.03 mg/ml) was incubated in the presence of hydroquinone (300 μ M) and H_2O_2 (50 μ M) for 30 minutes at room temperature in methyltransferase buffer. Heat-inactivated PO (100°C, 30 minutes) was used as a negative bioactivation control. Conversion of HQ to BQ was assessed by spectrophotometry (UV-1650PC; Shimadzu, France) by reading the absorbance spectrum of the solution between 220 and 340 nm. Recombinant SETD2 (3 μ M) was then incubated with a 1:10 dilution of the previous bioactivation mixture for 10 minutes at room temperature in methyltransferase buffer. Residual SETD2 H3K36 methyltransferase activity was assessed by UFLC as described above.

Western Blotting. Samples were separated by SDS-PAGE followed by a transfer onto a nitrocellulose membrane (0.22 μ m, GE Healthcare) for 1 hour at 4°C. Membranes were blocked in 5% nonfat milk in PBS 0.1% Tween-20 (PBST) for 1 hour and incubated overnight with primary antibody in 1% nonfat milk PBST at 4°C. The next day, membranes were washed three times with PBST prior to incubation with secondary antibody for 1 hour at room temperature. Membranes were then washed three more times, and the signal was detected by chemiluminescence using Enhanced Chemiluminescence (ECL) Prime reagent (GE Healthcare) on an Amersham Imager 600 detection system (GE Healthcare, France).

Cell Culture, Transfection of HEK293T Cells, and Cell Treatment. HEK293T, K562, and HeLa cells were grown in RPMI 1640 medium with 10% heat-inactivated FBS and 1 mM L-glutamine at 37°C under 5% CO_2 . Human CD34+ cells (from a healthy human donor) were maintained in RPMI 1640 medium with 20% heat-inactivated FBS, penicillin (100 U/ml), streptomycin (100 μ g/ml), and 1 mM L-glutamine at 37°C under 5% CO_2 . Cells were routinely tested for mycoplasma using DAPI staining. For transfection, HEK293T cells were seeded at 30,000 cells/cm² in a 100-cm² Petri dish and directly transfected using a solution containing 9 μ g of GFP-SETD2 plasmid and 18 μ l metafectene. Cells were then put back in the incubator at 37°C and 5% CO_2 for 48 hours.

For 1,4-benzoquinone treatment, K562, CD34+, and HeLa cells (10×10^6) were exposed to BQ or vehicle (Ctrl) for 30 minutes at 37°C, 5% CO_2 , in RPMI 1640 medium and then put back in the incubator for 3 more hours (6 hours for HeLa) at 37°C, 5% CO_2 , in RPMI 1640 medium. BQ concentrations used were 0, 1, 2, 5, and 10 μ M for K562 cells; 10 μ M for CD34+ cells; and 20 μ M for HeLa cells.

Immunofluorescence Detection of SETD2 in BQ-Treated HeLa Cells. HeLa cells (10×10^6) were fixed on slides using methanol at -20°C for 20 minutes. Methanol was then evaporated, and the slides were dried for 10 minutes. Cells were permeabilized with a PBS solution containing 0.1% Triton X-100 for 10 minutes at room temperature and washed three times with PBS before being incubated overnight with SETD2 (1:1000) and γ -H2AX (1:2000) antibodies in PBST at 4°C. After three washes in PBS, the slides were

incubated for 45 minutes at room temperature in PBST containing a suitable secondary antibody α -rabbit (Alexa Fluor 488, 1:500) or α -mouse (Alexa Fluor 594, 1:500). The slides were then washed with PBS, mounted, and marked with DAPI (Fluorshield).

Acid Extraction of Endogenous Histones. Treated K562 cells were lysed using cell lysis buffer (PBS 150 mM NaCl, pH 7.5, 1% Triton X-100, protease inhibitors) for 30 minutes at 4°C under agitation. The lysate was sonicated for 2 seconds (10% power) and then centrifuged for 15 minutes at 15,000g (4°C). The soluble fraction corresponds to total protein cell extract, and the pellet corresponds to insoluble chromatin and membranes. Total protein cell extracts were kept to be further used in Western blot analysis, and pellets were further incubated overnight in the presence of 0.2 N HCl at 4°C under agitation. The next day, the acid extracted histones were centrifuged for 15 minutes at 15,000g (4°C), and the supernatant containing the extracted histones was kept. Histone concentrations were determined using Bradford assay. In total, 2.5 μ g of extracted histones were separated by SDS-PAGE (18% gel) and transferred onto a nitrocellulose membrane. SETD2-dependent methylation of H3K36 was detected using an anti-H3K36me3 antibody. A ponceau staining of the membrane was done to ensure equal protein loading.

Treatment of Cell Extracts and Immunoprecipitation of GFP-SETD2. GFP-SETD2-transfected HEK293T cells were lysed using cell lysis buffer (PBS 150 mM NaCl, pH 7.5, 1% Triton X-100, protease inhibitors) for 30 minutes at 4°C under agitation. The lysate was sonicated for 2 seconds (10% power) and then centrifuged for 15 minutes at 15,000g (4°C). In all, 1 mg of total protein extract (soluble fraction) was incubated with increasing concentrations of BQ (0, 25, 50 μ M) or vehicle (Ctrl) for 30 minutes on ice and further incubated overnight at 4°C with 1.5 μ g anti-GFP antibody and protease inhibitors under agitation. The following day, samples were rocked for 2 more hours at 4°C with 30 μ l protein G-agarose. Beads were then washed three times with cell lysis buffer and two times with methyltransferase buffer and split in two: one-fifth of the beads was mixed with 20 μ l Laemmli sample buffer containing 400 mM β -mercaptoethanol to be further used in Western blot analysis, and the rest was incubated with 30 μ l methyltransferase buffer containing 75 μ M H3K36 peptide, 75 μ M SAM, 1 mM DTT, and protease inhibitors overnight at room temperature. The reaction was then stopped with the addition of 50 μ l perchloric acid ($HClO_4$) (15% v/v in water), and SETD2 H3K36 methyltransferase activity was assessed by UFLC as described above.

Chromatin Immunoprecipitation of BQ-Treated K562 Cells. K562 cells (10×10^6) were fixed with 1% formaldehyde for 10 minutes at 37°C. The reaction was stopped with the addition of 0.125 M glycine for 5 minutes at 4°C. Cells were then proceeded for chromatin immunoprecipitation (ChIP) using True Microchip kit (Diagenode SA) according to the manufacturer's instructions. DNA was immunoprecipitated using H3K36me3 and H3 (positive control) antibodies. Rabbit IgG was also used as a negative control. Immunocomplexes were isolated with protein A-agarose beads. DNA was isolated using the MicroChIP DiaPure columns (Diagenode SA). The primers used for qPCR were from Hacker et al. (2016) and correspond to *MYC* exon 2 and *CDK2* exon 6.

Human MYC_F: TGCCCCCTCAACGTTAGCTTC
 Human MYC_R: GGCTGCACCGAGTCGTAGTC
 Human CDK2_F: CCCTATTCCTGGAGATTCTG
 Human CDK2_R: CTCCGTCCATCTTCATCCAG

Fold enrichments of ChIP experiments were calculated by comparing the H3K36me3 ChIP with the IgG control and then normalized to the H3 antibody internal control of ChIP quality.

Statistical Analysis. The experiments carried out in this manuscript have not been designed to address a specific quantifiable statistical null hypothesis and thus have an exploratory character. The computed *P* values are therefore only descriptive. Values are presented as means \pm S.D. of three independent experiments and analyzed by either *t* test or one-way ANOVA using Rstudio [R Core Team (2017); R Foundation for Statistical Computing, Vienna, Austria] or Prism GraphPad 5 (GraphPad Prism version 5.0.0 for Windows; GraphPad

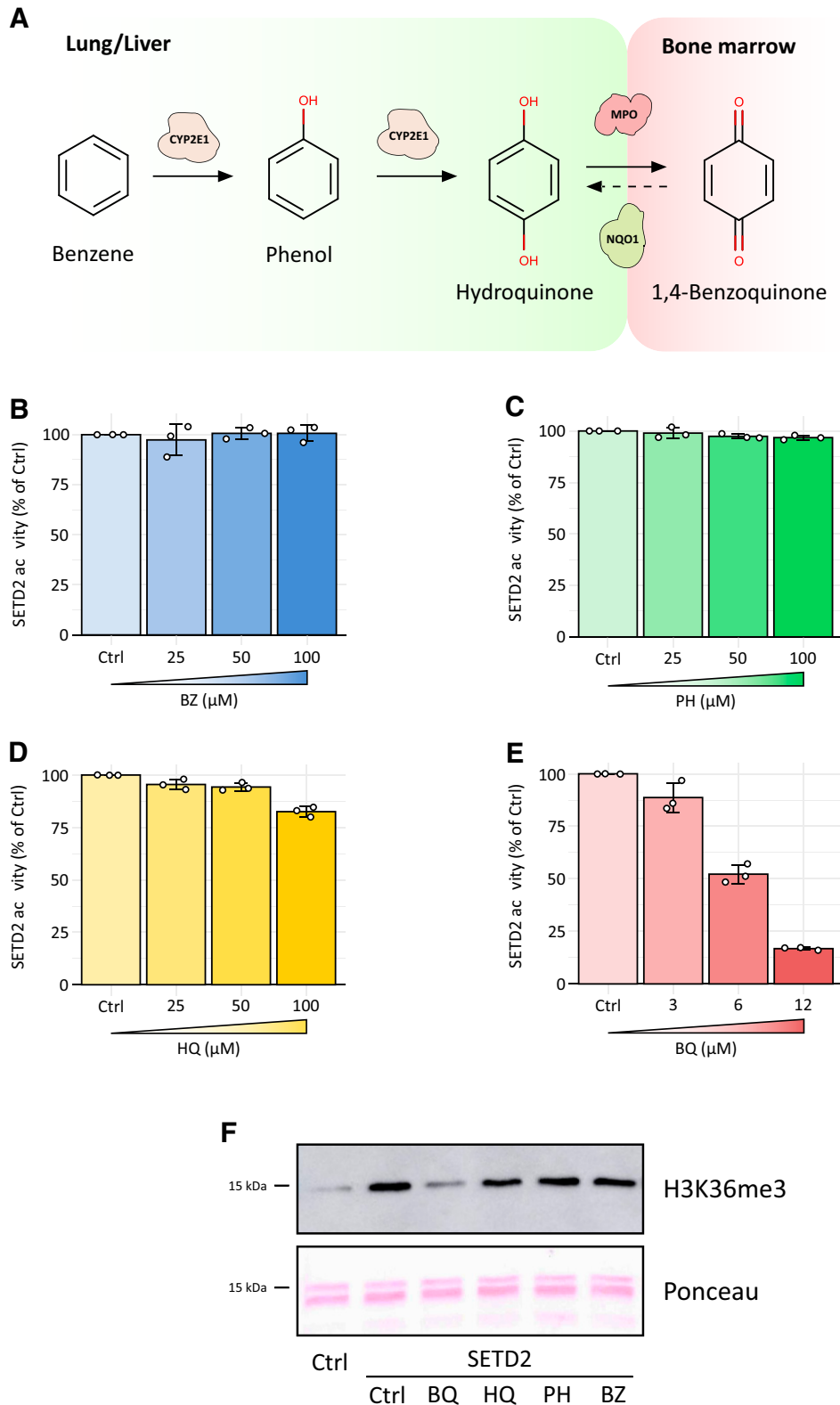


Fig. 1. Inhibition of the histone methyltransferase activity of human SETD2 by BZ, PH, HQ, and BQ. (A) Schematic representation of BZ bioactivation into BQ in bone marrow. MPO: myeloperoxidase; NQO1: NAD(P)H Quinone Dehydrogenase 1. (B–E) Recombinant SETD2 (3 μM) was incubated with increasing concentrations of BZ (B, 0–100 μM), PH (C, 0–100 μM), HQ (D, 0–100 μM), or BQ (E, 0–12 μM) for 10 minutes at room temperature. SETD2 residual activity was measured by UFLC using a H3K36 fluorescent peptide. Bar plots and error bars represent the means and the S.D. of three independent experiments. (F) SETD2 was incubated in the presence of 30 μM BZ, PH, HQ, or BQ for 10 minutes at room temperature. Residual SETD2 methyltransferase activity was assessed by incubating the enzyme with purified human core histones and detection of the H3K36me3 histone mark on nitrocellulose membranes using a specific antibody. Ponceau staining was carried out to ensure equivalent protein loading.

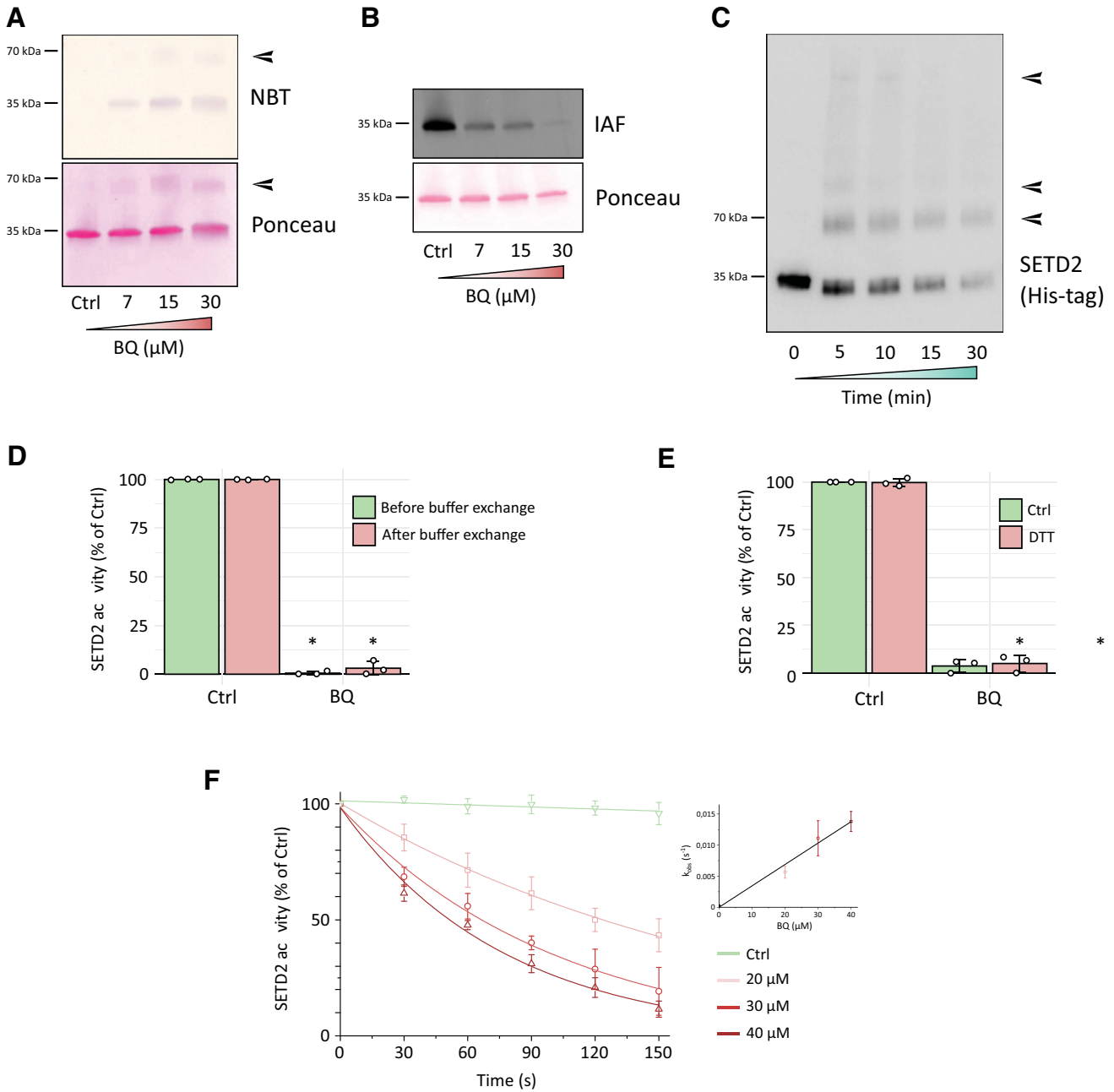


Fig. 2. BQ forms covalent adducts with SETD2 cysteine residues and protein crosslinks/oligomers. (A) The Benzene Hematotoxic and Reactive Metabolite 1,4-Benzoquinone Impairs the Activity of the Histone Methyltransferase SET Domain Containing 2 (SETD2) and Causes Aberrant Histone H3 Lysine 36 Trimethylation (H3K36me3). (B) Recombinant SETD2 was incubated with increasing concentrations of BQ (0–30 μM) for 10 minutes at room temperature, followed by a second incubation with 20 μM IAF for 10 minutes at 37°C. Samples were then separated by SDS-PAGE and transferred onto a nitrocellulose membrane. IAF-labeled cysteine residues were detected by fluorescence emission ($\lambda_{\text{exc}} = 485 \text{ nm}$, $\lambda_{\text{em}} = 530 \text{ nm}$). Ponceau staining was carried out to ensure equal protein loading. (C) Recombinant SETD2 was incubated with 30 μM BQ at room temperature. At different time points, an aliquot of the mixture was taken and diluted into Laemmli buffer containing 400 mM β -mercaptoethanol. Samples were then separated by SDS-PAGE and transferred onto a nitrocellulose membrane. SETD2 was detected using an antibody against 6 \times Histidine-tag. Black arrows represent high-molecular-weight SETD2 protein crosslinks/oligomers. (D) Recombinant SETD2 was incubated with 30 μM BQ for 10 minutes at room temperature and buffer-exchanged using PD-10 columns into methyltransferase buffer (see *Materials and Methods*). Residual SETD2 activity was measured by UFLC using a H3K36 fluorescent peptide. Bar plots and error bars represent the means and the S.D. of three independent experiments, respectively. * $P < 0.05$ compared with Ctrl. (E) Recombinant SETD2 was preincubated with 30 μM BQ for 10 minutes at room temperature, followed by an incubation with 10 mM DTT for 10 minutes at room temperature. Residual SETD2 was measured by UFLC using a H3K36 fluorescent peptide. Bar plots and error bars represent the means and the S.D. of three independent experiments, respectively. * $P < 0.05$ compared with Ctrl. (F) Recombinant SETD2 was incubated with increasing concentrations of BQ (0–40 μM) on ice. Every 30 seconds, an aliquot of the mixture was taken and diluted 10 times, and residual SETD2 activity was measured by UFLC using a H3K36 fluorescent peptide. For each BQ concentration, the percentage (%) of SETD2 residual activity was plotted as a function of time. The apparent first-order inhibition constants (k_{obs}) were calculated from the slope of the natural logarithm–based linear regression of the previous representation (see *Material and Methods*). The k_{obs} values were then plotted against BQ concentrations, and the second-order inhibition constant (k_{inact}) was extracted from the slope (inset). Dot plots and error bars represent the means and the S.D. of three independent experiments, respectively.

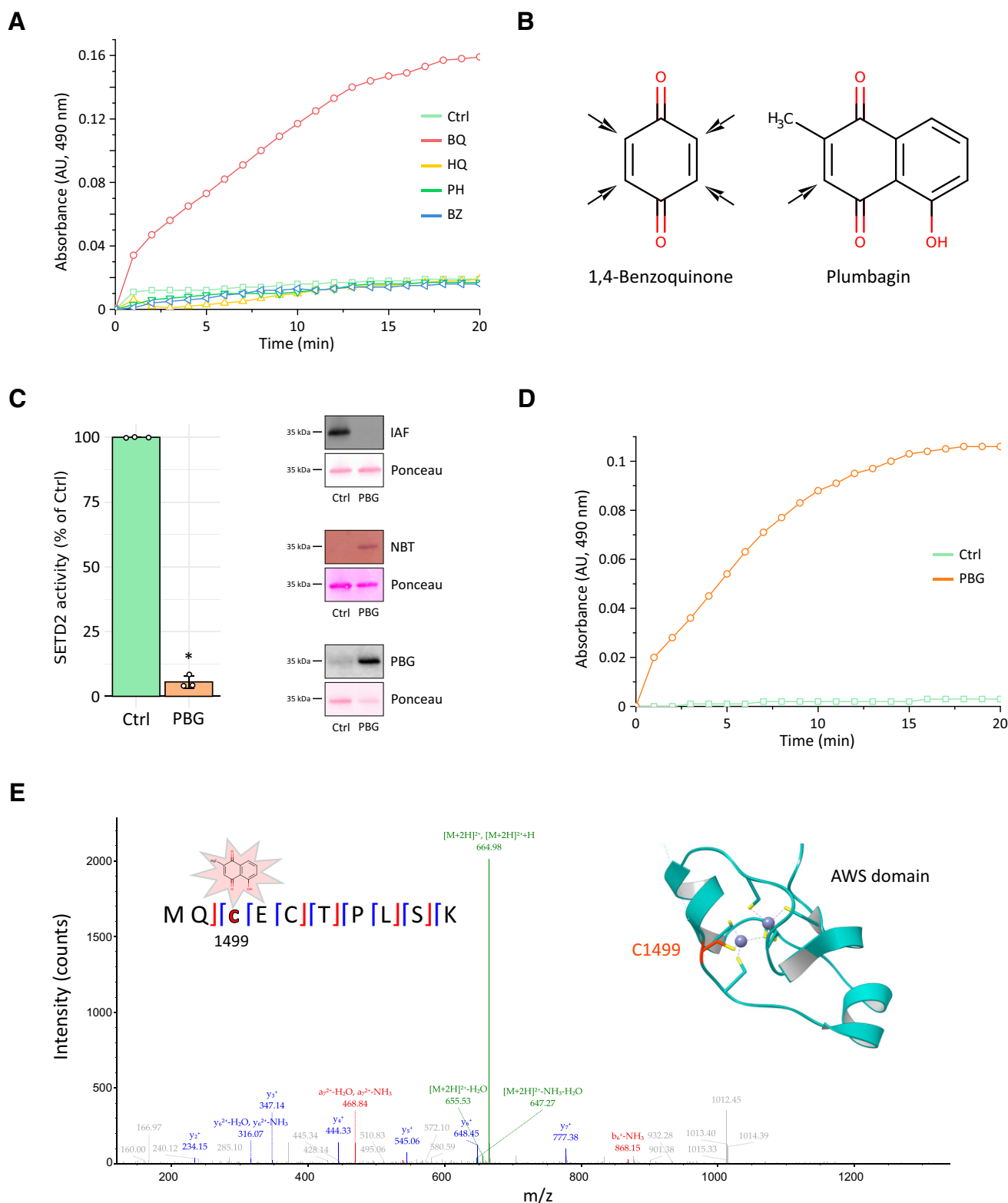


Fig. 3. BQ reacts with zinc-finger cysteine residues within the AWS domain of SETD2, leading to zinc release. (A) Recombinant SETD2 was incubated with 30 μM BZ, PH, HQ, or BQ and 100 μM PAR for 20 minutes at room temperature. The absorbance of PAR-Zn²⁺ complexes (490 nm) was followed every minute using a plate reader. (B) Comparison of the chemical structures of BQ and PBG. Black arrows represent potential Michael addition sites. (C) Left panel: recombinant SETD2 was incubated in the presence of 30 μM PBG for 15 minutes at room temperature. Residual SETD2 activity was then measured by UFLC using a H3K36 fluorescent peptide. Bar plots and error bars represent the means and the S.D. of three independent experiments, respectively. * $P < 0.05$ compared with Ctrl. Right panel: recombinant SETD2 was incubated with 30 μM PBG for 15 minutes at room temperature. For the NBT staining of quinoprotein adducts, samples were then separated by SDS-PAGE and transferred onto a nitrocellulose membrane. Plumbagin adducts on SETD2 were then detected using NBT staining as described in *Materials and Methods*. For the IAF cysteine labeling, samples were incubated with 20 μM IAF for 10 minutes at 37°C, separated by SDS-PAGE, and transferred onto a nitrocellulose membrane. IAF-labeled cysteine residues were detected by fluorescence emission ($\lambda_{\text{exc}} = 485 \text{ nm}$, $\lambda_{\text{em}} = 530 \text{ nm}$). For the

Software, San Diego, CA) followed by two-tailed Dunnett's test. $P < 0.05$ was used to consider differences as statistically significant.

Results

BQ Inhibits SETD2 H3K36 Methyltransferase Activity. BZ is known to be readily biotransformed into oxidative metabolites in bone marrow cells (Fig. 1A) (Smith, 2010). We tested the effects of BZ and its major oxidative metabolites on SETD2 catalytic domain (residues 1435–1711). As shown in Fig. 1, B–E, the H3K36 methyltransferase activity of SETD2 was significantly inactivated by low micromolar concentrations of BQ ($IC_{50} \approx 6 \mu\text{M}$). Conversely, no or very modest effects of phenol or hydroquinone were observed at concentrations up to $100 \mu\text{M}$. Moreover, BQ generated from hydroquinone by an *in vitro* peroxidase/ H_2O_2 system that mimics the formation of BQ in bone marrow was able to completely inactivate SETD2, thus further underlining the reactivity of this BZ metabolite toward SETD2 (Frantz et al., 1996) (Supplemental Fig. 1). The levels of BQ in bone marrow tissues upon BZ exposure have not been reported. However, different studies indicate that metabolites of BZ (including PH and HQ) could reach high micromolar concentrations ($\approx 180 \mu\text{M/ppm}$ of BZ) (Kim et al., 2006b; Rappaport et al., 2010). Levels of HQ (the precursor of BQ) close to $50 \mu\text{M}$ have been measured in urine of workers exposed to BZ (Kim et al., 2006a). Histone methyltransferase assays conducted with purified core histones confirmed that BQ was the sole BZ metabolite able to impair the trimethylation of H3K36 by SETD2 (Fig. 1F).

BQ Forms Michael Adducts on Active Site Zinc-Finger Cysteines and Impairs Irreversibly SETD2 Histone Methyltransferase Activity. BQ is known to be capable of reacting with certain redox-sensitive cysteines forming covalent quinone-thiol Michael adducts (Shu et al., 2019). As shown in Fig. 2, A and B, NBT redox staining and free cysteine labeling with IAF supported the formation of BQ-protein adducts on cysteines within the catalytic domain of SETD2. We also found that BQ led to the formation of non-reducible SETD2 protein crosslinks/oligomers as previously observed for other enzymes inhibited by BQ (Fig. 2, A and C) (Shu et al., 2020a, 2020b). Interestingly, SETD2 has been shown to be inherently prone to protein crosslink/aggregation, which can be detected by SDS-PAGE/Western blot or immunofluorescence (Bhattacharya and Workman, 2020). We found that the activity of SETD2 could not be restored by buffer exchange or reduction by DTT, which is consistent with the formation of irreversible cysteine-BQ adducts and protein crosslinks/oligomers (Fig. 2, D and E). Moreover, the inactivation of SETD2 by BQ was concentration- and time-dependent, with a second-order inhibition rate constant (k_i) equal to $4 \times 10^2 \text{ M}^{-1}\cdot\text{second}^{-1}$, thus further supporting the irreversible nature of the inactivation (Fig. 2F).

Cysteine residues are chemically and kinetically favored sites for BQ adduction, and as such, certain enzymes can be impaired by formation of BQ adducts on reactive cysteines (Bender et al., 2007; Mbiya et al., 2013; Duval et al., 2019; Shu et al., 2020b). SETD2 catalytic domain possesses two Zn-finger regions (AWS and SET/post-SET) containing three Zn atoms chelated by cysteines that contribute to the folding of the domain and ensure enzyme function (Yang et al., 2016; Zhang et al., 2017) (Supplemental Fig. 2A). Moreover, electrophilic agents such as BQ can react with certain reactive Zn-bound cysteines in Zn-fingers causing Zn ejection and protein unfolding (Lee et al., 2013). In agreement with this, we observed that BQ (but not BZ or the other BZ metabolites) was able to cause Zn release from SETD2 catalytic domain (Fig. 3A). Mass spectrometry experiments were carried out to identify the cysteines adducted by BQ. However, in initial attempts, no individual amino acid residues were identified. This is likely due to the ability of BQ to crosslink multiple residues in SETD2, as observed previously for topoisomerase II α (Bender et al., 2007). To circumvent this technical difficulty, plumbagin (a quinone that has one single reactive site for adduction) was used as a surrogate of BQ as previously described by Bender et al. (2007) (Fig. 3B). As observed with BQ, we confirmed that PBG formed covalent adducts with cysteine residues within SETD2 catalytic domain with concomitant loss of enzyme activity and Zn release (Fig. 3, C and D). Mass spectrometry analysis confirmed the presence of a PBG adduct on Cys1499 of the AWS Zn-finger (Fig. 3E). Accordingly, NBT and IAF staining confirmed that PBG could form quinoprotein adducts on AWS Zn-finger cysteines and induce Zn release (Supplemental Fig. 2, B–D). Interestingly, the 1499 cysteine residue chelates a Zn atom present in the AWS Zn-finger (Yang et al., 2016). However, as the peptide coverage was rather low ($\sim 30\%$), we cannot rule out that other cysteines could be adducted. Further experiments carried out with the purified AWS Zn-finger domain confirmed that BQ could form covalent adducts with AWS Zn-finger cysteines and induce Zn release and protein crosslinks/oligomers (Supplemental Fig. 2, D and E). Consistent with our results, we found that *N,N,N,N*-tetrakis-(2-pyridylmethyl)ethylenediamine (TPEN), a well known Zn chelator, could inactivate SETD2 catalytic domain by depleting Zn atoms (Supplemental Fig. 3). Altogether, our results suggest that BQ binds covalently to AWS domain Zn-finger cysteines within SETD2 catalytic domain, resulting in Zn release, loss of histone methyltransferase activity, and protein crosslinks/oligomers formation. This is in agreement with the key structural and functional role of the AWS Zn-finger domain of SETD2 (Yang et al., 2016; Zhang et al., 2017).

Impairment of SETD2 Histone Methyltransferase Activity and Decreased H3K36me3 Levels in Cells Exposed to BQ. We further tested the potential relevance of BQ-dependent inactivation of SETD2 by conducting

immunodetection of plumbagin adducts, plumbagin-treated SETD2 samples were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. Plumbagin adducts were detected using an antibody against plumbagin. Ponceau staining of each nitrocellulose membrane is carried out to ensure equal protein loading. (D) Recombinant SETD2 was incubated with $30 \mu\text{M}$ PBG and $100 \mu\text{M}$ PAR for 20 minutes at room temperature. The absorbance of PAR-Zn²⁺ complexes (490 nm) was followed every minute using a plate reader. (E) Recombinant AWS domain was incubated with $30 \mu\text{M}$ PBG for 30 minutes at room temperature. The sample was then reduced with 10 mM DTT and alkylated with NEM prior to trypsin digestion and liquid chromatography–MS/MS analysis. The mass spectrum shows the presence of a plumbagin adduct on the Zn-finger cysteine residue 1499. The localization of the adduct on cysteine 1499 in AWS domain is displayed (inset).

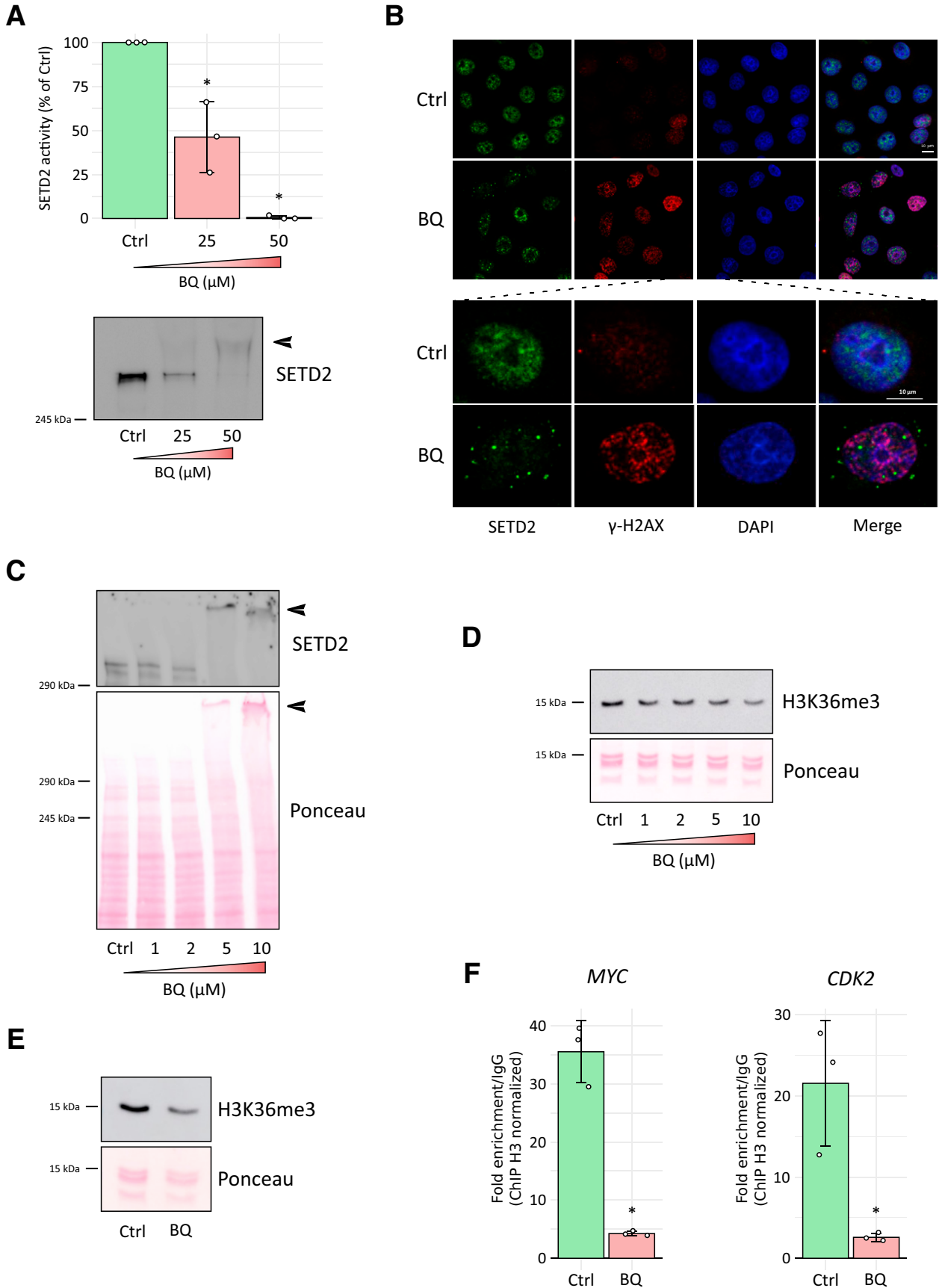


Fig. 4. BQ inhibits SETD2 activity and decreases H3K36me3 epigenetic mark in cells. (A) HEK293T cells transfected with full-length GFP-SETD2 plasmid were lysed, and 1 mg of cellular protein extracts was incubated with increasing concentrations of BQ (0–50 μM) for 30 minutes on ice. GFP-SETD2 was then immunoprecipitated, and residual SETD2 activity was measured by UFLC using a H3K36 fluorescent peptide. Bar

experiments in cells. To this end, the effects of BQ were first analyzed in HEK293T cells expressing GFP-tagged SETD2. This cellular approach has notably been used to analyze the formation of SETD2 aggregates/oligomers by Western blot (Bhattacharya and Workman, 2020). As shown in Fig. 4A, exposure of transfected cells to BQ led to the formation of inactive SETD2 crosslinks/oligomers as previously observed for other enzymes inhibited by BQ (Shu et al., 2020a, 2020b). SETD2 is inherently prone to protein crosslink/aggregation, and SETD2 aggregates/oligomers have been observed by immunofluorescence as puncta in cells (Bhattacharya and Workman, 2020). As shown in Fig. 4B, immunofluorescence studies showed the presence of puncta of SETD2 upon exposure of cells to BQ, thus supporting the formation of SETD2 crosslinks/oligomers as previously observed by Bhattacharya and Workman (2020). Further experiments were conducted in human hematopoietic K562 cells, which express functional SETD2. Consistent with the results described above, we found that exposure of K562 cells to BQ led to the formation of endogenous SETD2 crosslinks/oligomers (Fig. 4C). Concomitantly, and in agreement with the data reported above, we observed significantly decreased levels of the H3K36me3 mark on histones extracted from BQ-treated K562 cells (Fig. 4D). Congruent with the results obtained with the K562 cell line, exposure of human primary hematopoietic CD34+ stem cells to BQ led also to decrease of the H3K36me3 mark, which is consistent with inactivation of SETD2 by BQ (Fig. 4E). H3K36me3 levels have been shown to be present along the gene body with preference for exons, notably in *MYC* and *CDK2* genes (Eastmond et al., 2005). Using ChIP-qPCR, we analyzed the H3K36me3 levels at exons of *MYC* and *CDK2* in K562 cells exposed to BQ, as described previously by Hacker et al. (2016). Congruent with the above results, significantly decreased levels of H3K36me3 were observed for the exons studied, further supporting that BQ alters SETD2-dependent H3K36me3 epigenetic marks (Fig. 4F).

Discussion

BQ is considered as the major hematotoxic metabolite of BZ (Smith, 2010; North et al., 2020). The leukemogenic properties of BQ are thought to rely, at least in part, on the formation of covalent Michael adducts with cysteine residues that may affect the structure and the function of proteins involved in the regulation of hematopoietic cells (Smith et al.,

2011; Bolton and Dunlap, 2017; North et al., 2020; Shu et al., 2020b). So far, inhibition of topoisomerase II enzymes by BQ is considered as one of the key mechanisms contributing to BZ hematotoxicity (Bender et al., 2007; Eastmond et al., 2014; Holmes and Winn, 2019). However, recent studies suggest that epigenetic mechanisms, notably DNA and histone methylation, could be altered in BZ-induced leukemogenesis (Yu et al., 2019; Chung and Herceg, 2020). We show here that the histone H3K36-specific trimethyltransferase SETD2 is inactivated by BQ. SETD2 is the sole lysine methyltransferase that catalyzes the formation of H3K36me3, a key epigenetic mark involved in transcriptional activation and DNA repair (Husmann and Gozani, 2019). SETD2 is considered as a tumor suppressor, and recurrent inactivating mutations of SETD2 and aberrant H3K36me3 levels have been reported in hematopoietic malignancies (Mar et al., 2014; Zhu et al., 2014; Husmann and Gozani, 2019). We found that BQ-dependent irreversible inactivation of SETD2 activity occurs through the formation of covalent BQ adducts on Zn-fingers cysteines present in the catalytic domain of the enzyme, notably within the AWS subdomain (Fig. 3; Supplemental Fig. 3). The formation of these quinoprotein adducts on SETD2 result in Zn release from the enzyme. It has been shown that Zn-bound cysteines in Zn-fingers are redox-sensitive and react readily with electrophilic compounds to form covalent adducts that are accompanied by Zn ejection and alteration of protein structure and function (Lee et al., 2013). Interestingly, the histone methyltransferases G9a and GLP are inhibited by electrophilic compounds able to eject Zn from Zn-fingers from their catalytic domain (Lenstra et al., 2018). Moreover, quinone-containing compounds have been reported to inhibit the histone demethylase KDM4A and the CREBBP/p300 acetyltransferase at least in part through Zn ejection from their catalytic Zn-finger domains (Jayatunga et al., 2015; Zhang et al., 2021). We also found that generation of BQ adducts on cysteine Zn-fingers within SETD2 catalytic domain led to the formation of protein crosslinks/oligomers that may contribute to enzyme inactivation (Figs. 2 and 4). This is consistent with observations indicating that disruption of Zn-bound cysteines in Zn-fingers, notably through oxidation or reaction with electrophiles, can lead to Zn release and subsequent protein unfolding (Quintal et al., 2011; Lee et al., 2013; Kluska et al., 2018). Interestingly, SETD2 has been shown to be a rather unstable protein and prone to aggregation (Bhattacharya and Workman, 2020).

plots and error bars represent the means and the S.D. of three independent experiments, respectively. * $P < 0.05$ compared with Ctrl. In parallel, half of the immunoprecipitated samples was separated by SDS-PAGE and transferred onto a nitrocellulose membrane. GFP-SETD2 protein was detected using an anti-GFP antibody against GFP. Black arrows represent high-molecular-weight SETD2 protein crosslinks/oligomers. (B) HeLa cells were exposed to 20 μM BQ for 30 minutes at 37°C, washed, and further cultured for 6 hours at 37°C in fresh medium. Cells were fixed with methanol for 15 minutes, and endogenous SETD2 was detected by immunofluorescence using a specific anti-SETD2 antibody. An antibody against $\gamma\text{-H2AX}$ protein was also used to evaluate BQ-dependent DNA breaks. DAPI staining is featured for nuclear localization. Optical sections are shown with scale bars of 10 μm . (C) K562 cells were exposed to increasing concentrations of BQ (0–10 μM) for 30 minutes at 37°C and then washed and further cultured for 3 hours at 37°C in fresh medium. Cells were then lysed, and protein extracts were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. Endogenous SETD2 protein was detected using an anti-SETD2 antibody. Ponceau staining is shown. Black arrows represent high-molecular-weight SETD2 protein crosslinks/oligomers. (D) Endogenous histones were extracted from K562 cell lysates, and 2.5 μg of extracted histones was separated by SDS-PAGE and transferred onto a nitrocellulose membrane. Levels of the H3K36me3 mark were detected using an anti-H3K36me3 antibody. Ponceau staining was carried out to ensure equal protein loading. (E) Human hematopoietic CD34+ cells were exposed to 10 μM BQ for 30 minutes at 37°C, washed, and further incubated for 3 hours in fresh medium. Cells were then lysed into Laemmli Buffer containing 400 mM $\beta\text{-mercaptoethanol}$. Samples were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. Levels of the H3K36me3 mark were detected using an anti-H3K36me3 antibody. Ponceau staining was carried out to ensure equal protein loading. (F) K562 cells were exposed to 10 μM BQ for 30 minutes at 37°C, washed, and further cultured for 3 hours at 37°C in fresh medium. ChIP was carried out using antibodies against H3K36me3 histone mark and H3 as described in *Materials and Methods*. RT-qPCR experiments were subsequently carried out on human *MYC* and *CDK2* genes using primers specific for exons 3 and 6, respectively, located in the gene bodies.

More broadly, protein crosslinking upon covalent binding of BQ to cysteines has been shown to occur with enzymes such as glyceraldehyde-3-phosphate dehydrogenase, creatine kinase or thioredoxin 1, and is thought to be responsible, at least in part, for their inhibition (Shu et al., 2020a, 2020b). Moreover, the findings obtained with the purified SETD2 catalytic domain were further supported by experiments carried out in cells. Indeed, we found that exposure to BQ of cells expressing ectopically GFP-SETD2 or endogenous SETD2 resulted in protein crosslinks/oligomers and loss of SETD2 activity as observed with the purified SETD2 catalytic domain (Figs. 2 and 4). Consistent with this, we found that exposure to BQ caused a marked decrease of the H3K36me3 mark on histones from human hematopoietic K562 or primary bone marrow CD34+ cells. SETD2 activity and trimethylation of H3K36 are involved in normal hematopoiesis, and alteration of SETD2 and H3K36me3 levels is recurrently observed in hematopoietic malignancies (Mar et al., 2017; Chen et al., 2020). Of note, dysregulation of epigenetic processes such as DNA and histone methylation by BZ has been reported over the last years (Smith, 2010; Fenga et al., 2016; Chung and Herceg, 2020). Our data support these observations and provide mechanistic evidence that BQ, an oxidative and hematotoxic metabolite of BZ, is able to react with and perturb the activity of a key epigenetic enzyme.

Acknowledgments

We thank the technical platform “BioProfiler” (BFA, CNRS UMR 8251 Unit, Université de Paris) for provision of UFLC facilities.

Authorship Contributions

Participated in research design: Berthelet, Guidez, Rodrigues-Lima.

Conducted experiments: Berthelet, Michail, Bui, Le Coadou, Sirri, Guidez.

Performed data analysis: Berthelet, Michail, Bui, Le Coadou, Sirri, Wang, Dulphy, Dupret, Chomienne, Guidez, Rodrigues-Lima.

Wrote or contributed to the writing of the manuscript: Berthelet, Rodrigues-Lima.

References

Bender RP, Ham A-JJL, and Osheroff N (2007) Quinone-induced enhancement of DNA cleavage by human topoisomerase II α : adduction of cysteine residues 392 and 405. *Biochemistry* **46**:2856–2864.

Bhattacharya S and Workman JL (2020) Regulation of SETD2 stability is important for the fidelity of H3K36me3 deposition. *Epigenetics Chromatin* **13**:40.

Bolton JL and Dunlap T (2017) Formation and biological targets of quinones: Cytotoxic versus cytoprotective effects. *Chem Res Toxicol* **30**:13–37.

Butler JS and Dent SYR (2013) The role of chromatin modifiers in normal and malignant hematopoiesis. *Blood* **121**:3076–3084.

Chappell G, Pogribny IP, Guyton KZ, and Rusyn I (2016) Epigenetic alterations induced by genotoxic occupational and environmental human chemical carcinogens: a systematic literature review. *Mutat Res Rev Mutat Res* **768**:27–45.

Chen BY, Song J, Hu CL, Chen SB, Zhang Q, Xu CH, Wu JC, Hou D, Sun M, Zhang YL, et al. (2020) SETD2 deficiency accelerates MDS-associated leukemogenesis via S100a9 in NHD13 mice and predicts poor prognosis in MDS. *Blood* **135**:2271–2285.

Chopra M and Bohlander SK (2015) Disturbing the histone code in leukemia: translocations and mutations affecting histone methyl transferases. *Cancer Genet* **208**:192–205.

Chung FFL and Herceg Z (2020) The promises and challenges of toxico-epigenomics: Environmental chemicals and their impacts on the epigenome. *Environ Health Perspect* **128**:15001.

Copeland RA (2005) Evaluation of enzyme inhibitors in drug discovery. A guide for medicinal chemists and pharmacologists. *Methods Biochem Anal* **46**:1–265.

Dong Y, Zhao X, Feng X, Zhou Y, Yan X, Zhang Y, Bu J, Zhan D, Hayashi Y, Zhang Y, et al. (2019) SETD2 mutations confer chemoresistance in acute myeloid leukemia partly through altered cell cycle checkpoints. *Leukemia* **33**:2585–2598.

Duval R, Bui LC, Mathieu C, Nian Q, Berthelet J, Xu X, Haddad I, Vinh J, Dupret JM, Busi F, et al. (2019) Benzoquinone, a leukemogenic metabolite of benzene, catalytically inhibits the protein tyrosine phosphatase PTPN2 and alters STAT1 signaling. *J Biol Chem* **294**:12483–12494.

Duval R, Fritsch L, Bui L-C, Berthelet J, Guidez F, Mathieu C, Dupret J-M, Chomienne C, Ait-Si-Ali S, and Rodrigues-Lima F (2015) An acetyltransferase assay for fluorescent histone H3 peptides. *Anal Biochem* **486**:35–37.

Eastmond DA, Keshava N, and Sonawane B (2014) Lymphohematopoietic cancers induced by chemicals and other agents and their implications for risk evaluation: An overview. *Mutat Res* **761**:40–64.

Eastmond DA, Mondrala ST, and Hasegawa L (2005) Topoisomerase II inhibition by myeloperoxidase-activated hydroquinone: a potential mechanism underlying the genotoxic and carcinogenic effects of benzene. *Chem Biol Interact* **153-154**:207–216.

Edmunds JW, Mahadevan LC, and Clayton AL (2008) Dynamic histone H3 methylation during gene induction: HYPB/Setd2 mediates all H3K36 trimethylation. *EMBO J* **27**:406–420.

Fahey CC and Davis IJ (2017) SETting the stage for cancer development: SETD2 and the consequences of lost methylation. *Cold Spring Harb Perspect Med* **7**:a026468.

Fenga C, Gangemi S, and Costa C (2016) Benzene exposure is associated with epigenetic changes (review). *Review Mol Med Rep* **13**:3401–3405.

Frantz CE, Chen H, and Eastmond DA (1996) Inhibition of human topoisomerase II in vitro by bioactive benzene metabolites. *Environ Health Perspect* **104** (Suppl 6):1319–1323.

Hacker KE, Fahey CC, Shinsky SA, Chiang YJ, DiFiore JV, Jha DK, Vo AH, Shavit JA, Davis LJ, Strahl BD, et al. (2016) Structure/function analysis of recurrent mutations in SETD2 protein reveals a critical and conserved role for a SET domain residue in maintaining protein stability and histone H3 Lys-36 trimethylation. *J Biol Chem* **291**:21283–21295.

Holmes TH and Winn LM (2019) DNA damage and perturbed topoisomerase II α as a target of 1,4-benzoquinone toxicity in murine fetal liver cells. *Toxicol Sci* **171**:339–346.

Huehner R, Dong L, Chen X, Wu G, Parker M, Wei L, Ma J, Edmonson MN, Hedlund EK, Rusch MC, et al. (2014) The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. *Nat Commun* **5**:3630.

Husmann D and Gozani O (2019) Histone lysine methyltransferases in biology and disease. *Nat Struct Mol Biol* **26**:880–889.

Jayatunga MKP, Thompson S, McKee TC, Chan MC, Reece KM, Hardy AP, Sekirnik R, Seden PT, Cook KM, McMahon JB, et al. (2015) Inhibition of the HIF1 α -p300 interaction by quinone- and indandione-mediated ejection of structural Zn(II). *Eur J Med Chem* **94**:509–516.

Kim S, Vermeulen R, Waidyanatha S, Johnson BA, Lan Q, Rothman N, Smith MT, Zhang L, Li G, Shen M, et al. (2006a) Using urinary biomarkers to elucidate dose-related patterns of human benzene metabolism. *Carcinogenesis* **27**:772–781.

Kim S, Vermeulen R, Waidyanatha S, Johnson BA, Lan Q, Smith MT, Zhang L, Li G, Shen M, Yin S, et al. (2006b) Modeling human metabolism of benzene following occupational and environmental exposures. *Cancer Epidemiol Biomarkers Prev* **15**:2246–2252.

Kluska K, Adamczyk J, and Krężel A (2018) Metal binding properties, stability and reactivity of zinc fingers. *Coord Chem Rev* **367**:18–64.

Kudithipudi S and Jeltsch A (2014) Role of somatic cancer mutations in human protein lysine methyltransferases. *Biochim Biophys Acta* **1846**:366–379.

Lee YM, Wang YT, Duh Y, Yuan HS, and Lim C (2013) Identification of labile Zn sites in drug-target proteins. *J Am Chem Soc* **135**:14028–14031.

Lenstra DC, Al Temimi AHK, and Mecnović J (2018) Inhibition of histone lysine methyltransferases G9a and GLP by ejection of structural Zn(II). *Bioorg Med Chem Lett* **28**:1234–1238.

Li J, Duns G, Westers H, Sijmons R, van den Berg A, and Kok K (2016) SETD2: an epigenetic modifier with tumor suppressor functionality. *Oncotarget* **7**:50719–50734.

Lu PCW, Shahbaz S, and Winn LM (2020) Benzene and its effects on cell signaling pathways related to hematopoiesis and leukemia. *J Appl Toxicol* **40**:1018–1032.

Mar BG, Bullinger LB, McLean KM, Grauman PV, Harris MH, Stevenson K, Neuberger DS, Sinha AU, Sallan SE, Silverman LB, et al. (2014) Mutations in epigenetic regulators including SETD2 are gained during relapse in paediatric acute lymphoblastic leukaemia. *Nat Commun* **5**:3469.

Mar BG, Chu SH, Kahn JD, Krivtsov AV, Koche R, Castellano CA, Kotler JL, Zon RL, McConkey ME, Chabon J, et al. (2017) SETD2 alterations impair DNA damage recognition and lead to resistance to chemotherapy in leukemia. *Blood* **130**:2631–2641.

Mbiya W, Chipinda I, Siegel PD, Mhike M, and Simoyi RH (2013) Substituent effects on the reactivity of benzoquinone derivatives with thiols. *Chem Res Toxicol* **26**:112–123.

McHale CM, Zhang L, and Smith MT (2012) Current understanding of the mechanism of benzene-induced leukemia in humans: implications for risk assessment. *Carcinogenesis* **33**:240–252.

North CM, Roseboom M, Kocabas NA, Schnatter AR, Faulhammer F, and Williams SD (2020) Modes of action considerations in threshold expectations for health effects of benzene. *Toxicol Lett* **334**:78–86.

Quintal SM, dePaula QA, and Farrell NP (2011) Zinc finger proteins as templates for metal ion exchange and ligand reactivity. Chemical and biological consequences. *Metallomics* **3**:121–139.

Lindsey RH Jr, Bromberg KD, Felix CA, and Osheroff N (2004) 1,4-Benzoquinone is a topoisomerase II poison. *Biochemistry* **43**:7563–7574.

Rappaport SM, Kim S, Lan Q, Li G, Vermeulen R, Waidyanatha S, Zhang L, Yin S, Smith MT, and Rothman N (2010) Human benzene metabolism following occupational and environmental exposures. *Chem Biol Interact* **184**:189–195.

Sauer E, Gauer B, Nascimento S, Nardi J, Göthel G, Costa B, Correia D, Matte U, Charão M, Arbo M, et al. (2018) The role of B7 costimulation in benzene immunotoxicity and its potential association with cancer risk. *Environ Res* **166**:91–99.

Shu N, Cheng Q, Arnér ESJ, and Davies MJ (2020a) Inhibition and crosslinking of the selenoprotein thioredoxin reductase-1 by p-benzoquinone. *Redox Biol* **28**:101335.

Shu N, Häggglund P, Cai H, Hawkins CL, and Davies MJ (2020b) Modification of Cys residues in human thioredoxin-1 by p-benzoquinone causes inhibition of its catalytic activity and activation of the ASK1/p38-MAPK signalling pathway. *Redox Biol* **29**:101400.

- Shu N, Lorentzen LG, and Davies MJ (2019) Kinetics and biological consequences of quinone-induced protein adduction. *Free Radic Biol Med* **120**:S55.
- Smith MT (2010) Advances in understanding benzene health effects and susceptibility. *Annu Rev Public Health* **31**:133–148, 2, 148.
- Smith MT, Zhang L, McHale CM, Skibola CF, and Rappaport SM (2011) Benzene, the exposome and future investigations of leukemia etiology. *Chem Biol Interact* **192**:155–159.
- Snyder R (2012) Leukemia and benzene. *Int J Environ Res Public Health* **9**:2875–2893.
- Wagner EJ and Carpenter PB (2012) Understanding the language of Lys36 methylation at histone H3. *Nat Rev Mol Cell Biol* **13**:115–126.
- Wang L, Niu N, Li L, Shao R, Ouyang H, and Zou W (2018) H3K36 trimethylation mediated by SETD2 regulates the fate of bone marrow mesenchymal stem cells. *PLoS Biol* **16**:e2006522.
- Wang X, Thomas B, Sachdeva R, Arterburn L, Frye L, Hatcher PG, Cornwell DG, and Ma J (2006) Mechanism of arylating quinone toxicity involving Michael adduct formation and induction of endoplasmic reticulum stress. *Proc Natl Acad Sci USA* **103**:3604–3609.
- Whysner J, Reddy MV, Ross PM, Mohan M, and Lax EA (2004) Genotoxicity of benzene and its metabolites. *Mutat Res* **566**:99–130.
- Yang S, Zheng X, Lu C, Li G-MM, Allis CD, and Li H (2016) Molecular basis for oncohistone H3 recognition by SETD2 methyltransferase. *Genes Dev* **30**:1611–1616.
- Yu C-HH, Li Y, Zhao X, Yang S-QQ, Li L, Cui N-XX, Rong L, and Yi Z-CC (2019) Benzene metabolite 1,2,4-benzenetriol changes DNA methylation and histone acetylation of erythroid-specific genes in K562 cells. *Arch Toxicol* **93**:137–147.
- Zhang W, Berthelet J, Michail C, Bui LC, Gou P, Liu R, Duval R, Renault J, Dupret JM, Guidez F, et al. (2021) Human CREBBP acetyltransferase is impaired by etoposide quinone, an oxidative and leukemogenic metabolite of the anticancer drug etoposide through modification of redox-sensitive zinc-finger cysteine residues. *Free Radic Biol Med* **162**:27–37.
- Zhang Y-LL, Sun J-WW, Xie Y-YY, Zhou Y, Liu P, Song J-CC, Xu C-HH, Wang L, Liu D, Xu A-NN, et al. (2018) Setd2 deficiency impairs hematopoietic stem cell self-renewal and causes malignant transformation. *Cell Res* **28**:476–490.
- Zhang Y, Shan C-MM, Wang J, Bao K, Tong L, and Jia S (2017) Molecular basis for the role of oncogenic histone mutations in modulating H3K36 methylation. *Sci Rep* **7**:43906.
- Zhou Y, Yan X, Feng X, Bu J, Dong Y, Lin P, Hayashi Y, Huang R, Olsson A, Andreassen PR, et al. (2018) *Setd2* regulates quiescence and differentiation of adult hematopoietic stem cells by restricting RNA polymerase II elongation. *Haematologica* **103**:1110–1123.
- Zhu X, He F, Zeng H, Ling S, Chen A, Wang Y, Yan X, Wei W, Pang Y, Cheng H, et al. (2014) Identification of functional cooperative mutations of SETD2 in human acute leukemia. *Nat Genet* **46**:287–293.

Address correspondence to: Fernando Rodrigues-Lima, Université de Paris, BFA, UMR 8251, CNRS, F-75013, Paris, France. E-mail: fernando.rodrigues-lima@univ-paris-diderot

SUPPLEMENTAL FIGURES

The benzene hematotoxic and reactive metabolite 1,4-benzoquinone impairs the activity of the histone methyltransferase SETD2 and causes aberrant H3K36 trimethylation (H3K36me3)

Jérémy Berthelet^{1,#}, Christina Michail¹, Linh-Chi Bui¹, Louise Le Coadou¹, Valentina Sirri¹, Li Wang², Nicolas Dulphy³, Jean-Marie Dupret¹, Christine Chomienne^{4,5}, Fabien Guidez⁴ and Fernando Rodrigues-Lima¹

¹ Université de Paris, Unité de Biologie Fonctionnelle et Adaptative (BFA), UMR 8251, CNRS, 75013 Paris, France

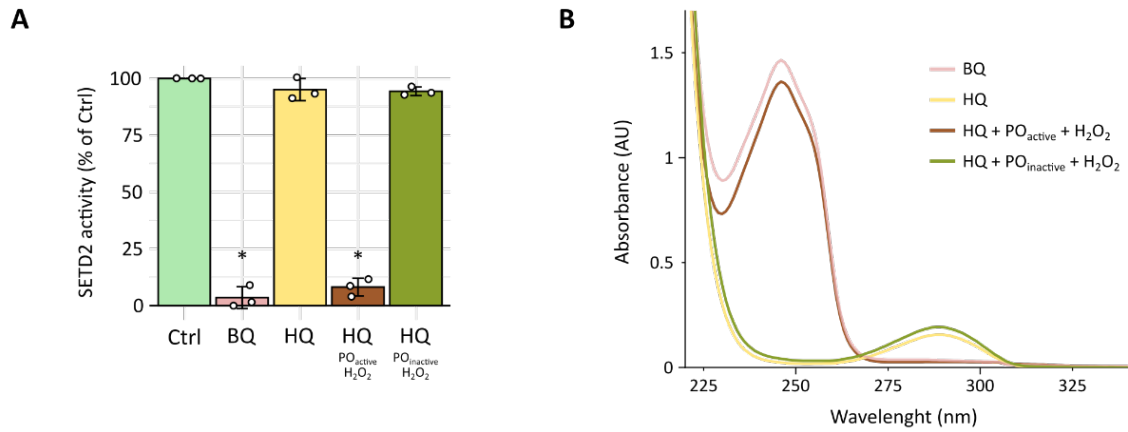
² The First Affiliated Hospital of Chongqing Medical University, Department of Hematology, Chongqing, China

³ Université de Paris, Institut de Recherche Saint-Louis (IRSL), UMRS 1160, INSERM, 75010 Paris, France

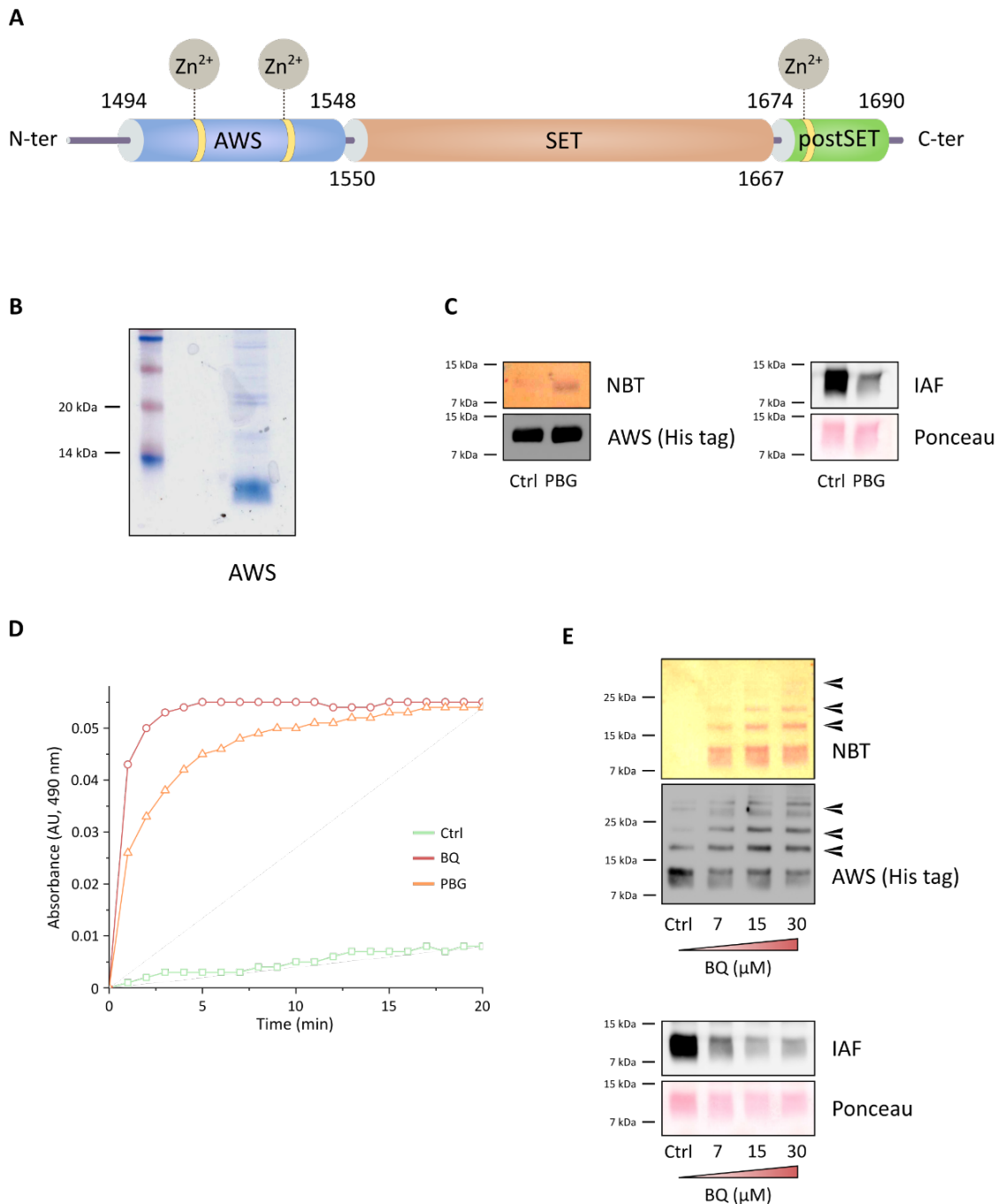
⁴ Université de Paris, Institut de Recherche Saint-Louis (IRSL), UMRS 1131, INSERM, 75010 Paris, France

⁵ Service de Biologie Cellulaire, Assistance Publique des Hôpitaux de Paris (AP-HP), Hôpital Saint Louis, Paris, France

Present address: Université de Paris, Centre Epigénétique et Destin Cellulaire (CEDC), UMR 7216, CNRS, 75013, Paris, France.

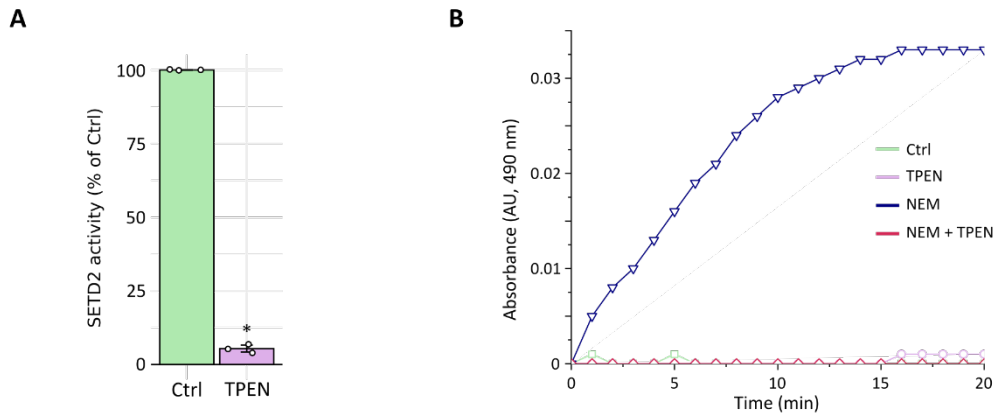


Sup figure 1. (A) Recombinant SETD2 was incubated 30 μ M BQ or HQ or HQ preincubated with active (HQ PO active) or inactive peroxidase (HQ inactive PO) hydroquinone-bioactivation system (see Materials and Methods) for 10 min at room temperature. Residual SETD2 activity was measured by UFLC using a H3K36 fluorescent peptide. Barplots and error bars represent the means and the SD of three independent experiments, respectively. * : *p*.value < 0.05 compared with control (Ctrl). **(B)** Absorbance spectrum (220-340 nm) of BQ, or HQ or HQ in the presence of active (HQ PO active) or inactive peroxidase (HQ PO inactive) bioactivation system (see Materials and Methods). Co-incubation of HQ with active PO hydroquinone-bioactivation system leads to a shift in the HQ absorbance spectrum with the loss of the characteristic HQ peak at 290 nm and the apparition of the characteristic BQ peak at 245 nm.



Sup figure 2. (A) Schematic representation of the catalytic domain of SETD2. Locations of the zinc fingers are represented using grey spheres. The AWS domain is displayed in blue, SET domain in light brown and post-SET domain in green. **(B)** Recombinant AWS domain of SETD2 was separated by SDS-PAGE followed by Coomassie staining. **(C)** Recombinant AWS domain of SETD2 was incubated with 30 μ M plumbagin (PBG) for 15 min at room temperature. Left panel: For the nitro blue tetrazolium (NBT) staining of quinoprotein adducts, samples were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. Plumbagin adducts on the AWS domain were then detected using NBT staining as described in Materials and Methods. Right panel: For the iodoacetamide-fluorescein (IAF) cysteine labelling, samples were incubated with 20 μ M iodoacetamide-fluorescein (IAF) for 10 min at 37°C. Samples were then separated by SDS-PAGE and transferred onto a nitrocellulose membrane. IAF-labelled cysteine residues were detected by fluorescence emission (λ_{exc} = 485 nm, λ_{em} = 530 nm). Ponceau staining is shown. **(D)** Recombinant AWS domain of SETD2 (5 μ g) was incubated in the presence of 30 μ M

1,4-benzoquinone (BQ) or plumbagin (PBG) and 100 μM 4-(2-pyridylazo)resorcinol (PAR) for 20 min at room temperature. The absorbance of PAR-Zn²⁺-complexes (490 nm) was followed every minute using a plate reader. **(E)** Recombinant AWS domain of SETD2 was incubated with increasing concentrations of BQ (0-30 μM) for 10 min at room temperature. Top panel: For the nitro blue tetrazolium (NBT) staining of quinoprotein adducts, samples were separated by SDS-PAGE and transferred onto a nitrocellulose membrane. BQ adducts on the AWS domain were then detected using NBT staining as described in Materials and Methods. The same samples were also analyzed by Western blot using an anti-6xHistidine tag antibody. Black arrows represent high molecular weight SETD2 protein cross-links/oligomers. Bottom panel: For the iodoacetamide-fluorescein (IAF) cysteine labelling, samples were incubated with 20 μM iodoacetamide-fluorescein (IAF) for 10 min at 37°C, separated by SDS-PAGE and transferred onto a nitrocellulose membrane. IAF-labelled cysteine residues were detected by fluorescence emission (λ_{exc} = 485 nm, λ_{em} = 530 nm). Ponceau staining is shown.



Sup figure 3. (A) Recombinant SETD2 (3 μ M) was incubated with 1 mM *N,N,N',N'*-tetrakis(2-pyridinylmethyl)-1,2-ethanediamine (TPEN) for 10 min at room temperature, followed by a buffer exchange using PD-10 columns to methyltransferase buffer (see Materials and Methods). Residual SETD2 activity was measured by UFLC using H3K36 fluorescent peptide. Barplots and error bars represent the means and the SD of three independent experiments, respectively. * : *p.value* < 0.05 compared with control (Ctrl). **(B)** Recombinant SETD2 (3 μ M) was incubated in the presence of 1 mM TPEN for 10 min at room temperature, followed by a buffer exchange using PD-10 columns to SETD2 activity buffer (see Materials and Methods). SETD2 (1 μ M) was then incubated with 1 mM *N*-ethylmaleimide (NEM) and 100 μ M 4-(2-pyridylazo)resorcinol (PAR) for 20 min at room temperature. The absorbance of PAR-Zn²⁺-complexes (490 nm) was followed every minute using a plate reader.



Contents lists available at ScienceDirect

Free Radical Biology and Medicine

journal homepage: www.elsevier.com/locate/freeradbiomed

Human CREBBP acetyltransferase is impaired by etoposide quinone, an oxidative and leukemogenic metabolite of the anticancer drug etoposide through modification of redox-sensitive zinc-finger cysteine residues

Wenchao Zhang^a, Jérémy Berthelet^{a,b}, Christina Michail^a, Linh-Chi Bui^a, Panhong Gou^c, Rongxing Liu^a, Romain Duval^d, Justine Renault^a, Jean-Marie Dupret^a, Fabien Guidez^c, Christine Chomienne^{c,e}, Fernando Rodrigues Lima^{a,*}

^a Université de Paris, BFA, UMR 8251, CNRS, F-75013, Paris, France

^b Université de Paris, CEDC, UMR 7216, CNRS, F-75013, Paris, France

^c Université de Paris, Institut de Recherche Saint-Louis, UMRS 1131, INSERM, F-75010, Paris, France

^d Université de Paris, BIGR, UMRS 1134, INSERM, F-75015, Paris, France

^e Service de Biologie Cellulaire, Assistance Publique des Hôpitaux de Paris (AP-HP), Hôpital Saint Louis, Paris, France

ARTICLE INFO

Keywords:

CREBBP
Quinone
Oxidative metabolite
Inhibition
Zinc-finger cysteines
Histone acetylation

ABSTRACT

Etoposide is an extensively prescribed anticancer drug that, unfortunately, causes therapy-related leukemia. The mechanisms by which etoposide induces secondary hematopoietic malignancies are poorly documented. However, etoposide-related leukemogenesis is known to depend on oxidative metabolites of etoposide, notably etoposide quinone, that can react with protein cysteine residues such as in topoisomerases II. CREBBP is a major histone acetyltransferase that functions mainly as a transcriptional co-activator. This epigenetic enzyme is considered as a tumor suppressor that plays a major role in hematopoiesis. Genetic alterations affecting CREBBP activity are highly common in hematopoietic malignancies. We report here that CREBBP is impaired by etoposide quinone. Molecular and kinetic analyses show that this inhibition occurs through the rapid and covalent ($k_{inhib} = 16.10^2 \text{ M}^{-1} \cdot \text{s}^{-1}$) adduction of etoposide quinone with redox sensitive cysteine residues within the RING and PHD Zn²⁺-fingers of CREBBP catalytic core leading to subsequent release of Zn²⁺. In agreement with these findings, experiments conducted in cells and in mice treated with etoposide showed irreversible inhibition of endogenous CREBBP activity and decreased H3K18 and H3K27 acetylation. As shown for topoisomerases II, our work thus suggests that the leukemogenic metabolite etoposide quinone can impair the epigenetic CREBBP acetyltransferase through reaction with redox sensitive cysteine residues.

1. Introduction

Lysine acetyltransferase (KAT) enzymes acetylate histone and non-histone proteins through the transfer of an acetyl group from acetyl-coenzyme A (AcCoA) to specific lysine residues [1,2]. CREB binding protein (CREBBP, also known as CBP or KAT3A) is a major mammalian KAT enzyme that functions as a transcriptional co-activator notably through acetylation of histones (in particular H3K18 and H3K27) and transcription factors (such as p53) [3–5]. CREBBP is a tumor suppressor that is important in normal development. Germline loss-of-function mutations in *CREBBP* cause Rubinstein-Taybi syndrome, a

developmental disorder that includes predisposition for leukemia [6]. CREBBP is a critical regulator of normal hematopoiesis and alteration of the acetyltransferase activity of the enzyme contributes to lymphomagenesis and leukemogenesis [7–11]. Moreover, loss of function mutations in *CREBBP* are indeed common in lymphoma and leukemia and are associated with aberrant acetylation of histone H3 (and non-histone proteins [9,10,12–14]). Multiple chromosomal translocations associated with hematological malignancies involve the KAT domain of CREBBP resulting in fusion proteins that contribute to leukemogenesis through abnormal histone acetylation [7,15,16]. The crucial role of CREBBP in normal and malignant hematopoiesis is further underlined

Abbreviations: CREBBP, CREB Binding Protein; KAT, lysine acetyltransferase; NBT, nitroblue tetrazolium; IAF, iodoacetamide-fluorescein.

* Corresponding author.

E-mail address: fernando.rodrigues-lima@u-paris.fr (F.R. Lima).

<https://doi.org/10.1016/j.freeradbiomed.2020.11.027>

Received 20 October 2020; Received in revised form 22 November 2020; Accepted 23 November 2020

Available online 2 December 2020

0891-5849/© 2020 Elsevier Inc. All rights reserved.

by different mouse models that show that lack of functional CREBBP leads to serious hematological defects [17–19].

Etoposide is a semi-synthetic podophyllotoxin derivative widely used for the treatment of many types of cancer. This chemotherapeutic drug acts as a topoisomerase II poison that inhibits the enzyme's ability to ligate DNA, resulting in an accumulation of double-strand breaks and subsequent apoptosis [20]. Although etoposide is commonly prescribed as a front-line treatment of a variety of human malignancies, this chemotherapeutic drug is also known to cause secondary leukemia [21–23]. However, the molecular basis underlying etoposide-induced leukemogenesis is not well-defined. In humans, etoposide is readily metabolized by CYP3A4 into etoposide catechol which is further oxidized by cellular oxidases/peroxidases into the reactive metabolite etoposide quinone [24–26]. Several studies have implicated oxidative etoposide metabolites, in particular etoposide quinone, as the primary contributors to the leukemogenic process [27–31]. Notably, etoposide quinone has been shown to react with redox sensitive cysteine residues of topoisomerase II enzymes [28–30]. In addition, efficacy of etoposide quinone against type II topoisomerases was found to be considerably higher than that of etoposide thus emphasizing the key role of this oxidative metabolite in the biological effects of etoposide and putting forward impairment of topoisomerase II functions by etoposide quinone as a major mechanism in etoposide-induced leukemia [28–30]. Nevertheless, reaction of etoposide quinone with additional cellular proteins besides topoisomerase II inhibition have been suggested [32–36]. In particular, recent studies indicate that exposure to etoposide may have epigenetic effects unrelated to topoisomerases II [37–40]. In line with these observations, we report herein molecular and kinetic evidence showing that the histone acetyltransferase CREBBP is irreversibly impaired by etoposide quinone, the major oxidative metabolite of etoposide. We found that this inhibition occurs through the rapid adduction by etoposide quinone of redox-sensitive cysteine residues located in the PHD/RING Zn²⁺-fingers of CREBBP and known to be crucial for its acetyltransferase activity [41–43]. Experiments in hematopoietic cells and in mice further confirm that exposure to etoposide irreversibly impairs CREBBP acetyltransferase activity with concomitant decrease in cellular histone H3 acetylation. Overall, our data indicate that, in addition to the alteration of topoisomerase II functions, etoposide quinone can impair CREBBP acetyltransferase by reacting with key cysteine residues within the catalytic core of the enzyme. Inhibition of a key epigenetic enzyme such as CREBBP by etoposide quinone may be an additional mechanism contributing to etoposide-dependent leukemogenesis.

2. Materials and methods

2.1. Chemicals and reagents

Etoposide and etoposide quinone were from Toronto Research Chemicals (Canada). Recombinant human H3.1 histone was from Actif Motif (Belgium). Antibodies used in this paper are: H3K18Ac (9675S, Cell Signaling), H3K27Ac (ab4729, Abcam, UK), H3 (3638S, Cell Signaling, France), CREBBP (A22, Santa Cruz, Germany), 6xHis (H1029, Sigma-Aldrich, France), secondary anti-rabbit (A1949, Sigma-Aldrich, France) and secondary anti-mouse (31430, ThermoFisher, France). All other chemicals, reagents and cells were from Sigma-Aldrich (France) otherwise noted.

2.2. Protein expression and purification

cdNA sequences from human CREBBP encoding residues 1096–1751 (CREBBP catalytic core), 1205–1279 (CREBBP RING domain) and 1280–1321 (CREBBP PHD domain) were cloned into pET28a expression vector and transformed into competent *E. coli* C41 (DE3). Bacteria were grown in Luria-Broth (LB) medium containing 30 µg/ml kanamycin and 20 µg/ml chloramphenicol. Recombinant protein expressions were

induced by 0.5 mM of IPTG and the cultures were shaken overnight at 16 °C. Bacteria pellets were resuspended in buffer A (20 mM Tris-HCl, 300 mM NaCl, 1% Triton X-100, 0.2 mM β-mercaptoethanol (β-ME), 1 mg/ml lysozyme, protease inhibitors, pH 8.0) and lysed for 30 min at 4 °C under agitation. The lysate was then sonicated for 10 min (10 s ON, 20 s OFF, 20% power) and the soluble fraction was clarified by centrifugation (15000 g, 4 °C, 30 min). The supernatant was incubated with His-Select Nickel Affinity Gel beads for 2 h on ice under agitation in presence of 10 mM imidazole. The proteins were eluted with buffer B (20 mM Tris-HCl, 300 mM NaCl, pH 8.0) containing 300 mM imidazole and buffer exchanged with a desalting column (PD-10) to buffer C (Tris-HCl 20 mM, NaCl 150 mM, 10% glycerol pH 8.0). Protein purity was assessed by SDS PAGE followed by Coomassie staining. Proteins were stored at –80 °C.

2.3. CREBBP acetyltransferase activity assay

CREBBP acetyltransferase activity was assayed as previously described [44]. CREBBP catalytic core was incubated with 100 µM AcCoA, 30 µM H3K18 or H3K27 peptide, 1 mM DTT in 100 µl of assay buffer (25 mM Tris-HCl pH 8.0) for 20 min at room temperature. The reaction was stopped with 100 µl of HClO₄ (15% v/v in water) and 20 µl aliquots of the mixture were injected in a reverse phase ultra-fast liquid chromatography system (RP-UFLC, Shimadzu, France). Peptides were separated under isocratic conditions (80% acetonitrile, 20% water, 0.12% TFA, flow rate of 0.6 ml/min) on a Shim-pack XR-ODS column at 40 °C. H3K18 and H3K27 substrate peptides and their acetylated product were detected by fluorescence (excitation: 485 nm, emission: 530 nm) and quantified by integration of the peak fluorescence areas.

2.4. Effects of etoposide quinone on CREBBP acetyltransferase activity

CREBBP catalytic core (500 nM) was incubated for 10 min at room temperature in the presence of different concentrations (0, 5, 10, 20, 40 µM) of etoposide quinone (EQ) or the parent drug (ET) in 20 µl assay buffer. Then, 80 µl of assay buffer (100 µM AcCoA, 30 µM H3K18 or H3K27 peptide, final concentrations) were added to the treated enzyme mixture and CREBBP acetyltransferase activity was assayed for 20 min at room temperature as described above.

2.5. Effects of reducing agents and buffer exchange on etoposide quinone-inhibited CREBBP

CREBBP catalytic core (500 nM) was first incubated with etoposide quinone (40 µM) or etoposide (ET) for 10 min at room temperature. Samples were either buffer exchanged using a desalting column (PD SpinTrap G-25, GE Healthcare, France) to remove free etoposide quinone in solution or further incubated with 10 mM DTT. Protein concentration was assessed after buffer exchange using Bradford assay (Bio-Rad, France). Residual acetyltransferase activity of CREBBP was assayed as described above.

2.6. Kinetic analysis of CREBBP inhibition by etoposide quinone

CREBBP catalytic core (500 nM) was incubated in the presence of different concentrations of etoposide quinone (0, 1, 2.5, 5, 7.5, 10 µM) in assay buffer. Every 30 s, an aliquot of the solution was taken and diluted 10 times in assay buffer to stop the reaction. Residual activity of CREBBP was then measured by UFLC as described above. The data were analyzed as described previously for irreversible inhibitors under pseudo first order conditions using Prism Graphpad and Excel softwares [45]. The rate of inhibition of CREBBP by etoposide quinone can be described as follows:

$\ln(\% \text{ residual CREBBP activity}) = -k_{\text{obs}} \times t$, where k_{obs} is apparent first-order inhibition rate constant, and t is time). For each etoposide concentration, the k_{obs} values can be extracted from the slopes of the

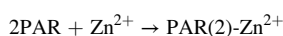
natural logarithm of the percentage of CREBBP activity as a function of time. The second order rate constant (k_{inhib}) was obtained from the slope of k_{obs} values plotted as a function of etoposide concentration ($k_{inhib} = k_{obs}/[\text{etoposide quinone}]$).

2.7. Fluorescein-conjugated iodoacetamide (IAF) labeling and nitro blue tetrazolium (NBT) staining

CREBBP catalytic core (10 μg), CREBBP RING (10 μg) or CREBBP PHD (10 μg) domains were incubated in the presence of increasing concentrations (0, 5, 10, 20, 40 μM) of etoposide quinone for 10 min at room temperature in assay buffer. For IAF labelling of free CREBBP protein cysteines, aliquots were incubated in the presence of 50 μM fluorescein-conjugated iodoacetamide (IAF) for 10 min at 37 °C. The labelling reaction was stopped by the addition of 10 μl sample buffer containing 200 mM β -mercaptoethanol prior to SDS-PAGE and transfer onto a nitrocellulose membrane. IAF fluorescence was detected using a Fujifilm LAS 4000 detection system ($\lambda_{exc} = 485 \text{ nm}$, $\lambda_{em} = 530 \text{ nm}$). For detection of etoposide quinone adducts on CREBBP proteins, NBT staining was carried out as reported by Ref. [46]. Briefly, aliquots of proteins were first separated by SDS-PAGE and transferred onto a nitrocellulose membrane prior to incubation with 2 M potassium glycinate, 0.6 mg/ml NBT, pH 10.0 at room temperature for 30 min in the dark.

2.8. Zn^{2+} release analysis

CREBBP RING (5 μM) or CREBBP PHD (5 μM) domains were incubated in the presence or absence of etoposide quinone (40 μM) in assay buffer containing 100 μM 4-(2-pyridylazo) resorcinol (PAR). PAR reacts with free Zn^{2+} in solution according to the scheme:



The $\text{PAR}(2)\text{-Zn}^{2+}$ complex has an absorption peak around 490 nm. The reaction was initiated by addition of etoposide quinone and monitored by reading the absorbance of the solution at 490 nm using a microplate reader (Biotek Instruments, France).

2.9. Mass spectrometry analysis

CREBBP catalytic core (10 μM) was incubated with 40 μM EQ for 20 min at room temperature in assay buffer. The reaction was stopped with 10 mM DTT and diluted 20 times in assay buffer. Samples were then incubated overnight at 37 °C with trypsin (Promega, France) at 12.5 ng/ μl in 25 mM ammonium bicarbonate pH 8.0. The supernatant containing peptides was acidified with formic acid (FA), desalted on C18 tips (Pierce C18 tips, Thermo Scientific, France), and eluted in 10 μl 70% ACN, 0.1% FA. Desalted samples were evaporated using a SpeedVac then taken up in 10 μl of buffer A (water, 0.1% FA) and 5 μl were injected on a nanoLC HPLC system (Thermo Scientific, France) coupled to a hybrid quadrupole-Orbitrap mass spectrometer (Thermo Scientific, France). Peptides were loaded on a reverse phase C18 μ -precursor (C18 PepMap100, 5 μm , 100A, 300 μm i.d.x5 mm) and separated on a C18 column (EASY-spray C18 column, 75 μm i.d.x50 cm) at a constant flow rate of 300 nl/min, with a 120 min gradient of 2–40% buffer B (buffer B: 20% water, 80% ACN, 0.1% FA). MS analyses were performed by the Orbitrap cell with a resolution of 120,000 (at m/z 200). MS/MS fragments were obtained by HCD activation (collisional energy of 28%) and acquired in the ion trap in top-speed mode with a total cycle of 3 s. The database search was performed against the Swissprot database and the *Homo sapiens* taxonomy with Mascot v2.5.1 software with the following parameters: tryptic peptides only with up to 2 missed cleavages, variable modifications: cysteine etoposide quinone and methionine oxidation. MS and MS/MS error tolerances were set respectively to 7 ppm and 0.5 Da. Peptide identifications were validated using a 1% False Discovery

Rate (FDR) threshold obtained by Proteome Discoverer (version 2.2, Thermo Scientific, France) and the percolator algorithm. The candidate sequences modified by etoposide quinone were manually inspected for *de novo* sequencing.

2.10. Bioactivation of etoposide into etoposide quinone by a myeloperoxidase (MPO)/ H_2O_2 system

Etoposide was bioactivated into etoposide quinone *in vitro* using a myeloperoxidase/ H_2O_2 system as described previously [31]. Briefly, etoposide (100 μM), H_2O_2 (200 μM) and myeloperoxidase (MPO) (5 units/ml) were co-incubated in 1 ml of 100 mM sodium acetate (pH 6.5) for 30 min at 37 °C. Etoposide quinone formation was detected by spectrophotometry (UV-1650PC, Shimadzu, France). Inactive MPO (heated for 30 min at 95 °C) was used as a negative control. 20 μl of bioactivated etoposide were mixed with 30 μl CREBBP (0.5 μM final) in assay buffer and incubated 20 min at room temperature. Then, 50 μl of assay buffer containing 30 μM H3K18 peptide, 100 μM AcCoA and 1 mM DTT (final concentrations) were added to the solution and CREBBP activity was assessed by UFLC as described above.

2.11. Detection of H3 acetylation by dot blot

CREBBP (500 nM) was first incubated with etoposide quinone or etoposide (40 μM) in 20 μl assay buffer for 10 min at room temperature. Then, 80 μl of assay mix (100 μM AcCoA, 300 ng recombinant human H3) were added and further incubated for 20 min at room temperature. Samples (corresponding to 200 and 100 ng H3) were dot-blotted using a Dot-Blot apparatus (Bio-Rad, France) onto nitrocellulose membranes (0.1 μm). Detection of H3K18 and H3K27 acetylation was carried out by western blotting using specific antibodies. The membranes were then stripped and re-probed with anti-H3 antibody.

2.12. Western blot

Samples were separated by SDS-PAGE prior to transfer onto a nitrocellulose membrane (0.45 μm , GE Healthcare, France) for 2 h 30 min at 4 °C. Membranes were blocked in 5% non-fat milk in PBS 0.05% Tween-20 (PBST) for 2 h and incubated overnight with primary antibody in 1% non-fat milk PBST at 4 °C. The next day, membranes were washed 5 times with PBST and incubated with secondary antibody for 2 h at room temperature. Membranes were then washed 5 more times and the signal was detected by chemiluminescence using ECL Prime on an Amersham Imager 600 detection system (GE Healthcare, France).

2.13. Cellular culture and treatment with etoposide

K562 cells were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS) and 1 mM L-glutamine at 37 °C, 5% CO_2 . K562 cells (10×10^6 cells) were washed 3 times with RPMI 1640 medium and then treated with 40 μM etoposide in RPMI 1640 medium containing 1 mM L-glutamine for 24 h at 37 °C, 5% CO_2 . Cells were then washed with PBS before use.

2.14. Acid extraction of histones

K562 cells in 100 mm petri dishes or bone marrow from mice hind limb long bones were lysed for 30 min at 4 °C with 400 μl lysis buffer (PBS pH 7.5, 0.5% Triton X-100, 3.3 μM trichostatin A (TSA) and protease inhibitors) and then centrifuged (15000 g) for 10 min at 4 °C. The pellet (nucleus enriched fraction) was incubated overnight at 4 °C in 400 μl 0.2 M HCl. The next day, the acid extract was centrifuged (10 min at 6500 g, 4 °C) and the supernatant (histone enriched fraction) was taken. Protein concentration was determined using Bradford reagent (Bio-Rad, France). 10 μg of histone extract were separated by SDS-PAGE (18%), transferred onto a nitrocellulose membrane and detected by

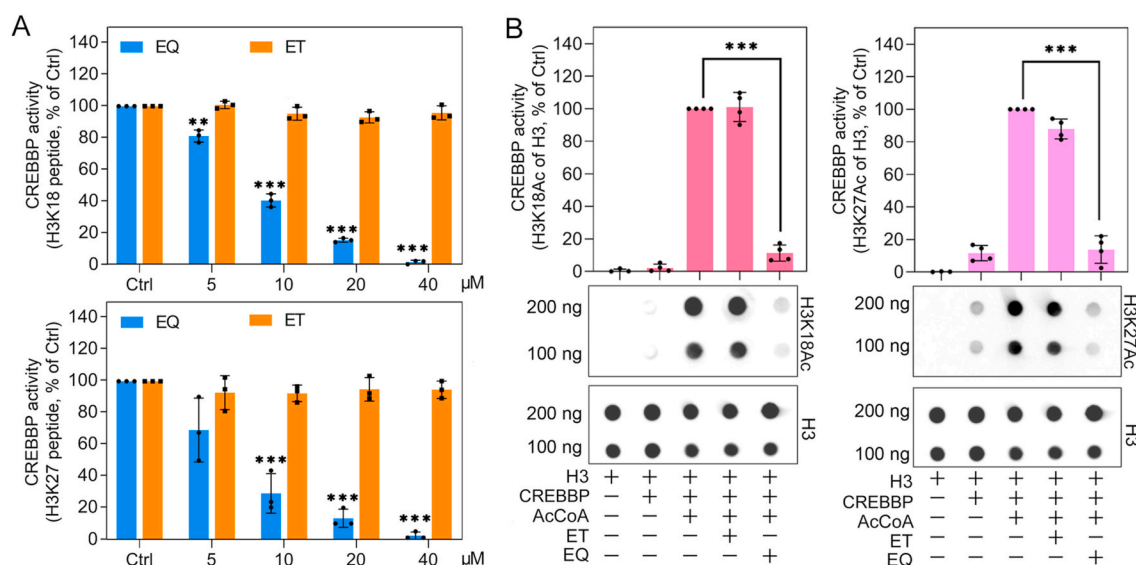


Fig. 1. Etoposide quinone inhibits CREBBP acetyltransferase activity

(A) Purified CREBBP catalytic core (aa 1096–1700) was incubated with increasing concentrations of etoposide quinone (EQ), etoposide (ET) or vehicle (Ctrl) for 20 min at room temperature. Residual CREBBP acetyltransferase activity was then measured by UFLC using H3K18 and H3K27 fluorescent peptides. Results are means of three independent experiments. Error bars indicate S.D. values. ** $p < 0.05$ or *** $p < 0.001$ compared with control (Ctrl). (B) CREBBP catalytic core was first pre-incubated with 40 μ M etoposide quinone (EQ) or etoposide (ET) for 20 min at room temperature prior to further incubation with recombinant histone H3 and acetyl-Coenzyme A (AcCoA) for another 20 min at room temperature. Samples were dot-blotted and H3K18 (left panel) and H3K27 (right panel) acetylation was detected using specific antibodies. Dot-blots were quantified using ImageJ software. Results are means of three independent experiments. Error bars indicate S.D. values. *** $p < 0.001$ compared with control (no EQ or ET).

western-blot using H3K18Ac, H3K27Ac or H3 antibodies. Grayscale quantitative analysis of Western blots was performed using ImageJ software.

2.15. Immunoprecipitation of endogenous CREBBP

K562 cells (10.10^6 cells) or mice bone marrow from hind limb long bones were lysed with lysis buffer (PBS pH 7.5, 1% Triton X-100, 3.3 μ M trichostatin A, 100 μ M DTT and protease inhibitors) for 30 min on ice and then centrifuged for 15 min (15000 g) at 4 $^{\circ}$ C. The supernatant was taken and the concentration of proteins in the extracts was determined by the Bradford reagent (Bio Rad, France). CREBBP was immunoprecipitated from 2.5 mg total protein extracts by an overnight incubation (under rocking conditions) with 1 μ g of anti-CREBBP antibody (Santa Cruz, Germany) at 4 $^{\circ}$ C. The following day, samples were rocked for 2 h at 4 $^{\circ}$ C with 25 μ l protein A agarose beads (Santa Cruz, Germany). The beads were washed 3 times with assay buffer and then incubated at room temperature with 60 μ l of assay buffer containing 30 μ M H3K18 fluorescent peptide, 100 μ M AcCoA and 1 mM DTT. After 1 h of incubation, beads were sedimented by centrifugation and 50 μ l of the supernatant were taken and mixed with 50 μ l of HClO₄ (15% v/v in water) to determine CREBBP activity by UFLC. Beads were then washed with PBS and incubated with sample buffer containing 200 mM β -ME for 5 min at 100 $^{\circ}$ C prior to SDS-PAGE and Western-blot analysis. Grayscale quantitative analysis of Western blots was performed using ImageJ software.

2.16. Treatment of mice with etoposide

C57BL/6 mice (12 weeks old female) were obtained from Janvier (Le Genest, France) and bred in pathogen-free conditions in the animal facility of the Institut de Recherche Saint-Louis IRSL (Hôpital Saint-Louis, Université de Paris). Mice (n = 6) were treated overnight by intraperitoneal injection of 100 mg/kg etoposide in DMSO. Controls (n = 6) were carried out with vehicle only. The following day, the hind limb long bones were taken and washed 3 times with cold PBS. Bone marrow was

flushed from the bones with PBS and the bone marrow cells (5.10^7 cells/mice) were lysed with lysis buffer (PBS pH 7.5, 1% Triton X-100, 3.3 μ M TSA, 1 mM DTT and protease inhibitors) for 30 min on ice under agitation. The supernatant was then used for immunoprecipitation of CREBBP and the pellet for histone extraction as described above. ChIP qPCR experiments were carried out as described below. All the animal procedures performed were approved by the local Committee on the use and care of Animal.

2.17. Chromatin immunoprecipitation (ChIP)

K562 cells (10.10^6 cells) or bone marrow cells (5.10^7 cells/mice) were fixed with 1% formaldehyde for 10 min at 37 $^{\circ}$ C. The reaction was stopped with the addition of 0.125 M glycine for 5 min at room temperature. Cells were then proceeded for ChIP using ChIP-IT® Express Chromatin Immunoprecipitation Kit (Active Motif, USA) according to manufacturer instructions. Anti-H3K18Ac, anti-H3K27Ac, and anti-H3 were used. Rabbit IgG were used as a control. Immunocomplexes were isolated with protein-A-agarose beads (Santa Cruz, Germany). DNA was isolated by phenol/chloroform extraction followed by ethanol precipitation. The primers used for qPCR were:

Human IL6_F: CAGCAGCCAACCTCCTCTAA
 Human IL6_R: CTGTGAGCGGCTGTTGTAGA
 Human MDM2_F: CGGAAAGATGGAGCAAGAAG
 Human MDM2_R: CCCTCAAGACTCCCAGTTT
 Mouse MDM2_F: AGATCTGACGCCCTCTTCTG
 Mouse MDM2_R: GCAGCTTCAGTCCCGATTTA
 Mouse COPS2_F: CATGCGACATGAAATCATCC
 Mouse COPS2_R: CACAAACTGTGCTCAGGAA

2.18. Statistical analysis

Values are presented as mean \pm standard deviation (SD) and analyzed by either t-test or one-way analysis of variance (ANOVA) using

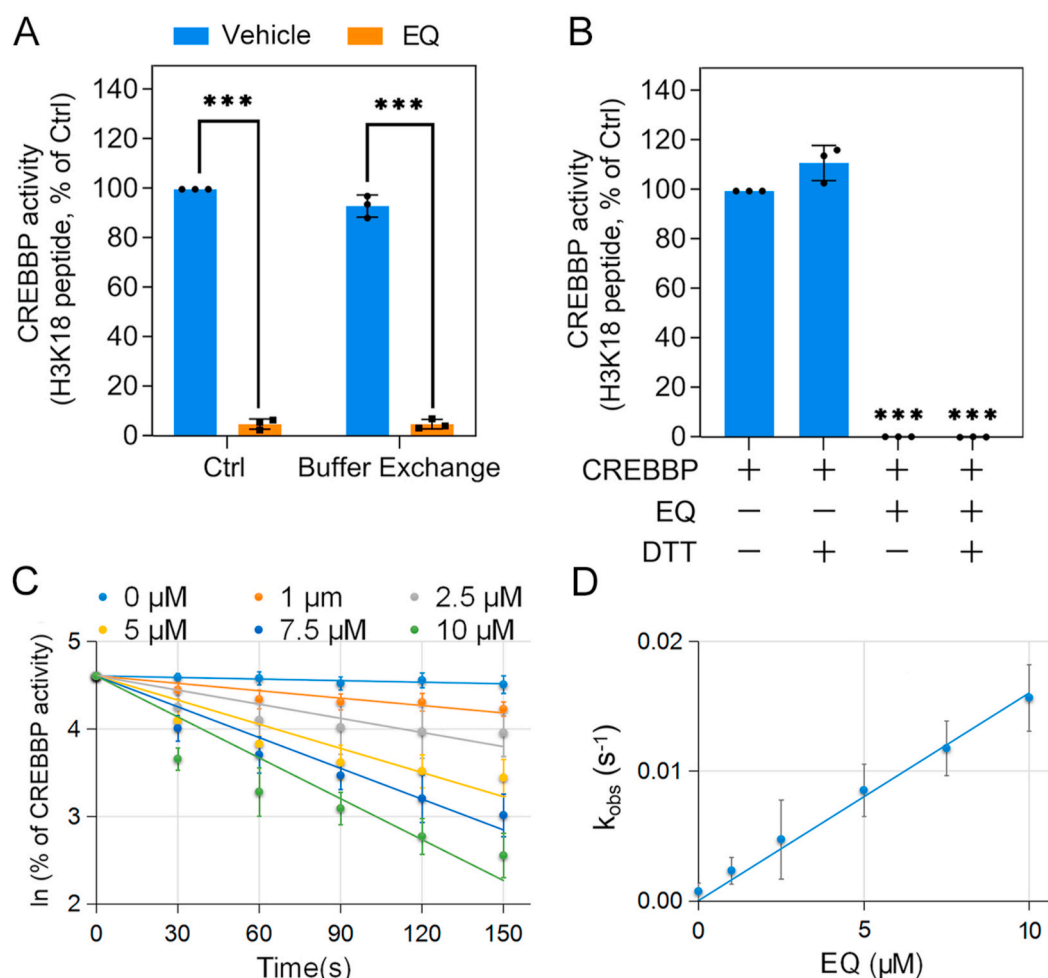


Fig. 2. Inhibition of CREBBP acetyltransferase activity by etoposide quinone is irreversible and fast

(A) CREBBP catalytic core was incubated with 40 μM etoposide quinone (EQ) or vehicle (Ctrl) for 20 min at room temperature. Samples were then buffer exchanged using PD-10 columns and CREBBP residual activity assayed by UFLC using H3K18 fluorescent peptide. Results are means of three independent experiments. Error bars indicate S.D. values. *** $p < 0.001$ compared with control.

(B) CREBBP catalytic core was pre-incubated with 40 μM etoposide quinone (EQ) for 20 min at room temperature followed by incubation with 10 mM DTT for 20 min at room temperature. Samples were then assayed for CREBBP activity by UFLC using H3K18 fluorescent peptide. Results are means of three independent experiments. Error bars indicate S.D. values. *** $p < 0.001$ compared with control.

(C) CREBBP catalytic core was incubated with different concentrations of etoposide quinone (0–10 μM). Aliquots were taken every 30s and residual acetyltransferase activity was measured by UFLC using H3K18 fluorescent peptide. For each etoposide quinone concentration the natural logarithm (ln) of the percentage (%) of CREBBP residual activity was plotted as a function of time. The apparent first order inhibition constants (k_{obs}) were calculated from the linear regressions (slopes). Results are means of three independent experiments. Error bars indicate S.D. values.

(D) The k_{obs} values were plotted against etoposide quinone (EQ) concentrations and the second order inhibition constant (k_{inhib}) was extracted from the slope. Results are means of three independent experiments. Error bars indicate S.D. values.

Prism GraphPad 5 followed by two-tailed Tukey test. A threshold of $p < 0.05$ was used to consider differences as statistically significant.

3. Results

3.1. Etoposide quinone impairs the acetylation of H3K18 and H3K27 by CREBBP *in vitro*

The leukemogenic properties of etoposide resides, at least in part, on its metabolic bioactivation in hematopoietic cells into oxidative quinone metabolites, most notably etoposide quinone which is known to react with cysteine residues and be a strong inhibitor of nuclear topoisomerases II [24,28,29,47]. We asked whether etoposide quinone could affect other nuclear enzymes whose alteration could contribute to leukemogenesis. Interestingly, the histone acetyltransferase CREBBP is known to be a major regulator of hematopoiesis and impairment of its activity is linked to the development of hematological malignancies [7].

In addition, CREBBP activity is known to rely on redox-sensitive cysteine zinc finger within the catalytic core of the enzyme [41–43]. To test the effects of etoposide quinone on human CREBBP HAT activity, the purified recombinant catalytic core domain of the enzyme (residues 1096 to 1700) [48] (Supplementary Fig. 1) was incubated with increasing concentrations of the metabolite or the parent drug prior to measurement of the HAT activity of CREBBP using fluorescent histone H3 peptides as substrates [44]. We found that contrary to etoposide, etoposide quinone was able to inhibit CREBBP HAT activity in a dose-dependent manner with IC_{50} values in the low micromolar ranges (below 10 μM) for both H3K18 and H3K27 peptides (Fig. 1A and Supplementary Fig. 2). Consistently, assays carried out with recombinant histone H3 confirmed that etoposide quinone but not etoposide inhibits CREBBP HAT activity (Fig. 1B). No significant effects of etoposide on CREBBP activity were observed (even at concentrations of etoposide >200 μM (data not shown) further confirming that CREBBP HAT activity is more sensitive to etoposide quinone than to the parent drug etoposide.

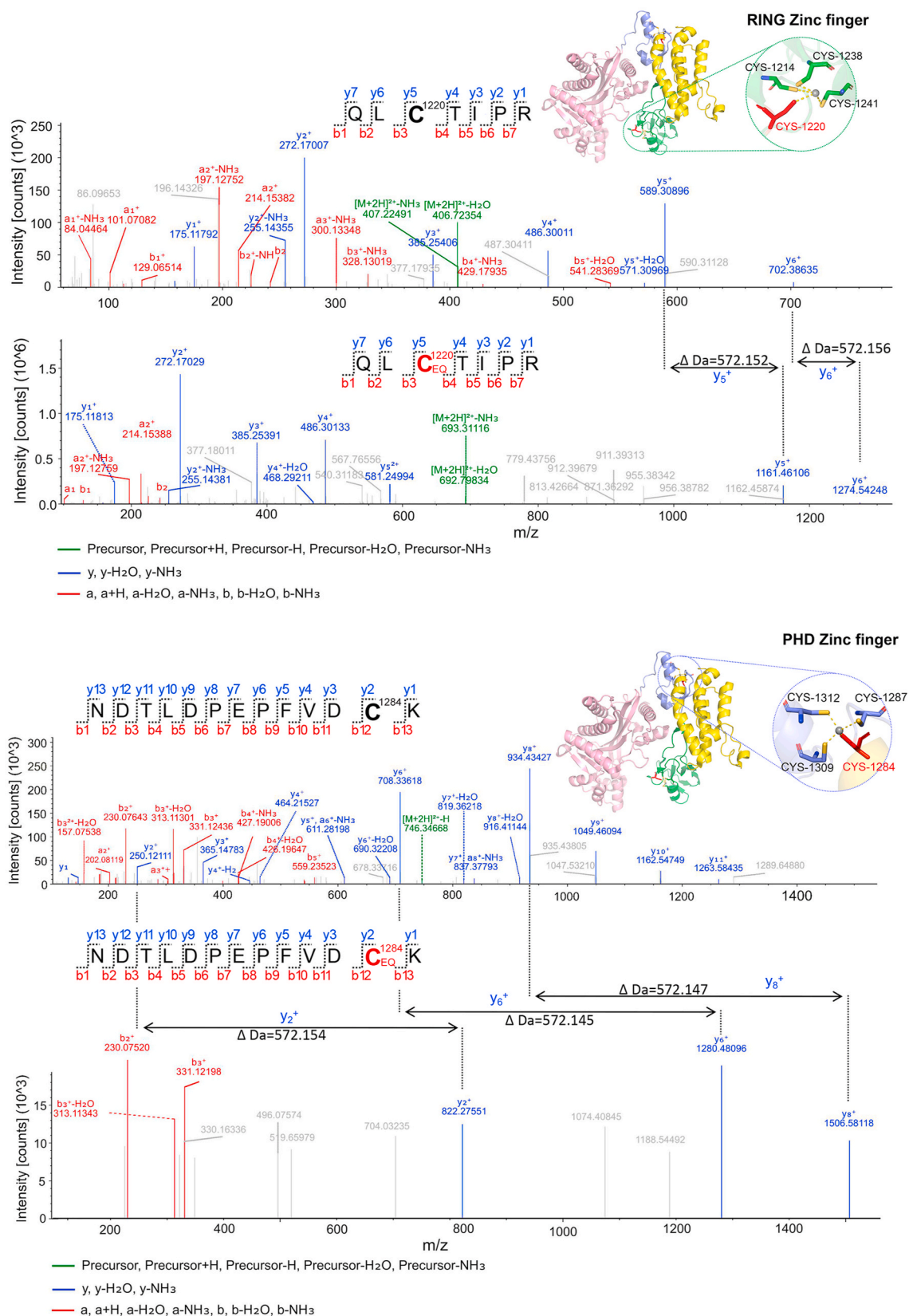


Fig. 3. Mass spectrometry analyses of etoposide quinone adducts on CREBBP catalytic core CREBBP catalytic core was incubated with etoposide quinone (40 μ M) for 20 min at room temperature. The samples were reduced with DTT and alkylated with NEM prior to trypsin digestion and liquid chromatography-MS/MS analysis. A control without etoposide quinone was done in the same conditions. Mass spectra showing presence of an etoposide quinone adduct on cysteine 1220 of the RING Zn²⁺-finger (Upper panels) and on cysteine 1284 of the PHD Zn²⁺-finger (Lower panels) are shown. The sequences of the peptides are displayed on each panel in addition to the etoposide quinone adducts (EQ) on the cysteine residues.

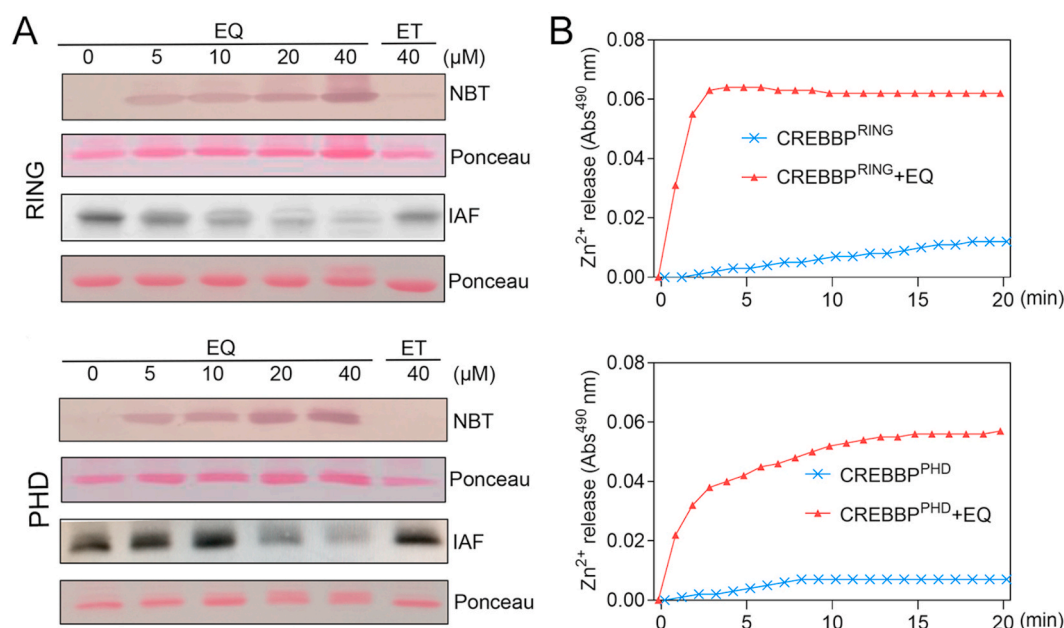


Fig. 4. Modification of cysteine residues of the RING and PHD domains of CREBBP by etoposide quinone and subsequent Zn²⁺ release

(A) Purified RING (aa 1205–1279) (top panel) and PHD (aa 1280–1321) (bottom panel) domains were incubated with increasing concentrations of etoposide quinone (EQ, 0–40 μM) or 40 μM etoposide (ET) for 20 min at room temperature. For the quinone-specific nitro blue tetrazolium (NBT) staining, samples were separated by SDS-PAGE and transferred onto nitrocellulose membranes. Etoposide quinone adducts on the RING and PHD proteins were then detected using NBT as described in Materials and Methods. Ponceau red staining is shown as a loading control. For the iodoacetamide-fluorescein (IAF) labelling, samples were incubated with 10 μM IAF for 10 min at room temperature, separated by SDS-PAGE and transferred onto nitrocellulose membranes. Cysteine residues labelled by IAF were detected by fluorescence ($\lambda_{exc} = 485$ nm, $\lambda_{em} = 530$ nm). Ponceau red staining is shown as a loading control

(B) Purified RING and PHD domains were incubated with or without etoposide quinone (40 μM) in the presence of 100 μM PAR. The absorbance of PAR-Zn²⁺-complexes (490 nm) was followed as a function of time at room temperature using a plate reader. Etoposide quinone had no effect on PAR absorbance (data not shown). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.2. Inhibition of CREBBP by etoposide quinone is fast and irreversible

Etoposide quinone has been shown to be an irreversible and fast inhibitor of human topoisomerases II at micromolar concentrations [29, 30]. To further analyze the inhibition mechanism of CREBBP by etoposide quinone, the enzyme was incubated with etoposide quinone and subsequently subjected to buffer exchange prior to measurement of residual HAT activity. As shown in Fig. 2A, the activity of CREBBP could not be recovered by buffer exchange after inhibition by etoposide quinone. Neither could the activity be restored by addition of the reducing agent DTT (Fig. 2B). To further analyze the inhibition mechanism of CREBBP by etoposide quinone, we performed kinetic experiments under pseudo-first order conditions (Fig. 2C and D) [45]. As shown in Fig. 2C, the semilogarithmic plots of percentage of residual activity against time for various concentrations of etoposide quinone indicated that the enzyme was inhibited in a time- and dose-dependent manner thus confirming the irreversible nature of the inhibition. The second-order rate constant of inhibition (k_{inhib}) was determined by plotting the k_{obs} values obtained from the semilogarithmic plots against etoposide quinone concentration and was found to be equal to 16.10^2 M⁻¹ s⁻¹ (Fig. 2D). Overall, these data show that etoposide quinone inhibits CREBBP rapidly and irreversibly similar to etoposide quinone-dependent inhibition of human topoisomerases II [29,30].

3.3. Etoposide quinone forms covalent adducts with PHD/RING Zn²⁺-finger cysteines of the catalytic core of CREBBP

Etoposide quinone has been previously shown to covalently inhibit human topoisomerases II by reaction with cysteine residues [30]. Moreover, quinone species, such as etoposide quinone, are known to be electrophilic chemicals that can readily react covalently with nucleophilic amino acids in proteins, most notably redox-sensitive cysteine

residues [36,49–51]. The formation of these covalent quinone-protein adducts is considered as a major mechanism for the biological and toxicological effects of quinone species [51–54]. We carried out mass spectrometry analyses to investigate whether amino acids found within the catalytic core of human CREBBP could react with etoposide quinone and form such covalent adducts. LC-MS/MS analysis of tryptic digests of etoposide quinone-inhibited CREBBP allowed the identification of two peptides derived from the RING and PHD Zn²⁺-finger domains with mass increases of 572.15 Da (corresponding to the mass of an etoposide quinone molecule) at cysteine residues 1220 and 1284, respectively (Fig. 3, upper and lower panel for RING and PHD sequences, respectively). To further confirm that etoposide quinone can react covalently with cysteine residues within the RING and PHD domains of the enzyme, we incubated purified recombinant RING and PHD proteins (aa 1205–1279 and 1280–1321, respectively; Supplementary Fig. 1) with etoposide quinone prior to quinone-specific nitro blue tetrazolium (NBT) staining and protein cysteine labeling by fluorescein-iodoacetamide (IAF) [34,46,55]. As shown in Fig. 4A, etoposide-quinone was able to form covalent adducts with both RING and PHD domains as evidenced by NBT staining of RING and PHD proteins treated with etoposide quinone. Covalent adduction of etoposide quinone on the two protein domains was accompanied with the concomitant loss of cysteine labeling by IAF thus supporting that quinone adducts occurred on reactive cysteine residues in these two domains. Zn²⁺-bound cysteines found in Zn²⁺-fingers are known to be redox sensitive and to readily react with electrophilic agents such as quinone species leading to covalent modification of these residues and to subsequent ejection of Zn²⁺ [56,57]. We thus used the metallochromic indicator 4-(2-pyridylazo)resorcinol (PAR) to monitor the release of Zn²⁺ from RING and PHD Zn²⁺-fingers upon exposure to etoposide quinone [58]. As shown in Fig. 4B, exposure to etoposide quinone led to the rapid ejection of Zn²⁺ from both RING and PHD domains.

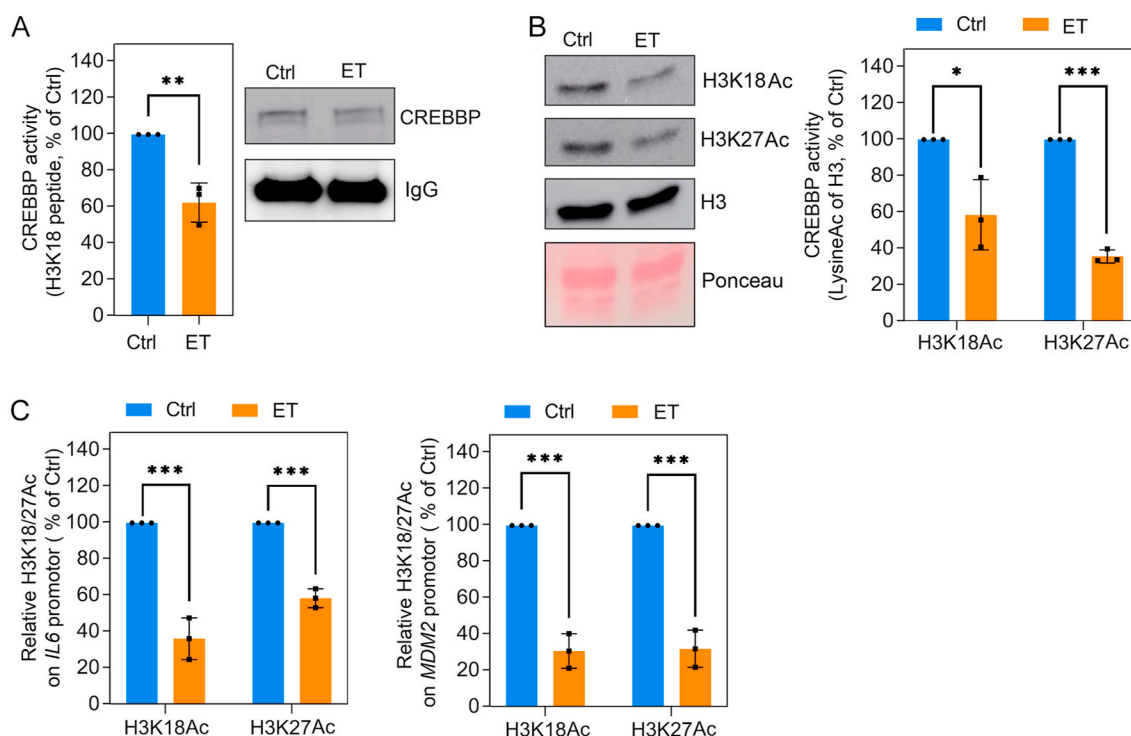


Fig. 5. Exposure of K562 cells to etoposide leads to inhibition of CREBBP and to the loss of H3K18 and H3K27 acetylation

(A) K562 cells were exposed to 40 μ M etoposide (ET) or vehicle (Ctrl) for 24 h, cells were lysed and cellular CREBBP was isolated by immunoprecipitation. Immunoprecipitated CREBBP activity was measured by UFLC. In parallel, CREBBP protein content in immunoprecipitates was determined by Western-blot (Right panel). Results are means of three independent experiments. Western blot is representative of three independent experiments. Error bars indicate S.D. values. ** p < 0.01 compared with control (Ctrl).

(B) Histones were extracted from the cell lysates and H3K18 and H3K27 acetylation was detected by Western blot. Quantification of H3 acetylation in Western blots was carried out using ImageJ software. Results are means of three independent experiments. Western blot is representative of three independent experiments. Error bars indicate S.D. values. * p < 0.05 compared with control (Ctrl), *** p < 0.001.

(C) Chromatin immunoprecipitation (ChIP) was conducted on cells chromatin using anti-H3K18ac and H3K27ac histone marks as described in *Materials and Methods*. RT-qPCR experiments were subsequently carried out on *IL6* and *MDM2* genes using primers specific for their promoters. Results are means of three independent experiments. Error bars indicate S.D. values. *** p < 0.001 compared with control (Ctrl).

Altogether, the data reported above indicate that etoposide quinone covalently reacts with CREBBP catalytic core forming adducts with cysteine residues found within the RING and PHD Zn²⁺-fingers, leading to the subsequent release of Zn²⁺.

3.4. CREBBP activity is irreversibly inhibited in human hematopoietic cells and in mice exposed to etoposide

Upon administration, etoposide is readily metabolized into reactive oxidative compounds in cells, most notably etoposide quinone, that covalently binds to redox sensitive cysteine residues of proteins [25, 59–62]. The biotransformation of etoposide into etoposide quinone is mainly due to cellular oxidative enzymes such as CYP3A4 or peroxidases [24,63]. To confirm that etoposide quinone generated from biotransformation of etoposide can indeed inhibit CREBBP, we took advantage of the previously described *in vitro* peroxidase/H₂O₂ system, that mimics the cellular oxidative bioactivation of etoposide into etoposide quinone [31,47]. As shown in *Supplementary Fig. 3*, etoposide oxidized by a functional peroxidase/H₂O₂ system was able to irreversibly inhibit CREBBP activity and lead to the formation of protein-quinone adducts (NBT staining) as well as loss of cysteine thiol labeling by IAF. In contrast, when a non-functional peroxidase/H₂O₂ system was used to metabolize etoposide, we did not observe inhibition nor formation of quinone adducts on cysteine residues were observed (*Supplementary Fig. 3*). These findings are in agreement with our results described above and confirm that, contrary to etoposide, its quinone metabolite can readily impair CREBBP HAT activity through formation of irreversible

covalent adducts on redox sensitive cysteine residues.

As we mentioned above, metabolization of etoposide into oxidative metabolites, such as etoposide quinone, is prominent in hematopoietic cells treated with this chemotherapeutic drug and has been implicated in the leukemogenic processes [24,59,63]. We thus tested whether the inhibition of CREBBP HAT activity could occur in a cellular context. To this end, K562 cells were exposed to etoposide at concentrations that do not significantly impact cell viability (*Supplementary Fig. 4*) [64]. The endogenous CREBBP enzyme was then immunoprecipitated and assayed for HAT activity. As shown in *Fig. 5A*, endogenous CREBBP HAT activity was irreversibly inhibited in cells exposed to etoposide (~60% residual activity). Remarkably, we also found that the acetylation of histone H3 (H3K18Ac and H3K27Ac) in cells was also reduced upon etoposide treatment (*Fig. 5B*). Consistent with the loss of acetylation mark on histone H3, ChIP-qPCR experiments revealed decreased levels of H3K18Ac and H3K27Ac on the promoters of *IL6* and *MDM2* genes (known to be acetylated and regulated by CREBBP [10,65–67], (*Fig. 5C*). Further investigations carried out on mice (C57BL/6J, n = 6) treated with a clinically-relevant dose of etoposide *via* intraperitoneal injection (100 mg/kg) [68] revealed a decrease in H3K18 and H3K27 acetylation levels in histones isolated from the bone marrow cells of treated mice (*Fig. 6A*). In line with these data, the intrinsic HAT activity of immunoprecipitated CREBBP was significantly reduced in etoposide-treated mice (*Fig. 6A* and *B*). Furthermore, ChIP-qPCR experiments performed on chromatin extracted from mice bone marrow cells were in agreement with the previous results showing that H3K18 and H3K27 acetylation was decreased on the promoters of *COPS2* and

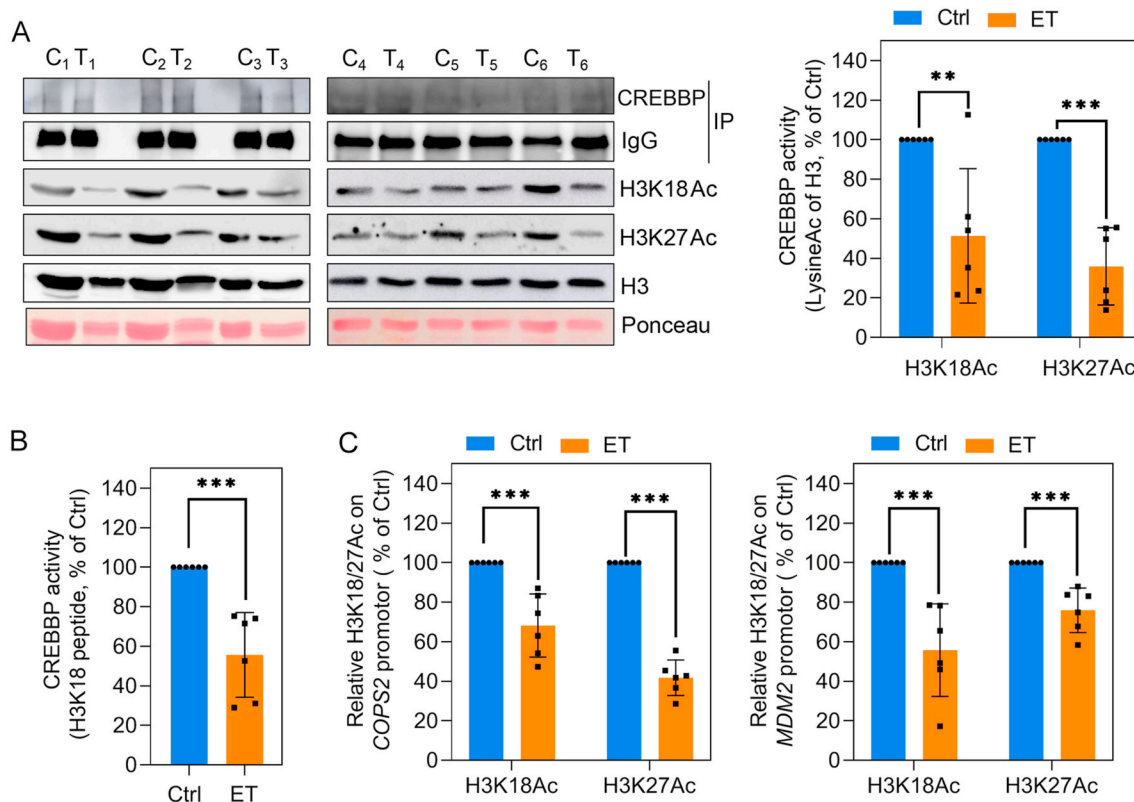


Fig. 6. Exposure of mice to etoposide inhibits CREBBP activity and H3K18 and H3K27 acetylation in bone marrow cells.

(A) C57BL/6J mice ($n = 6$) were treated with 100 mg/kg etoposide (Tn) or vehicle (Cn) for 24 h. Bone marrow was isolated from hind limb long bone. Histones were extracted from cell lysates and H3K18 and H3K27 acetylation was detected by Western blot (Left panels). Quantification of H3 acetylation in Western blots was carried out using ImageJ software (Right panel). Results are means of the six treated (Tn) and non-treated (Cn) mice. Error bars indicate S.D. $**p < 0.01$ compared with control (Ctrl), $***p < 0.001$ compared with Ctrl.

(B) CREBBP was first immunoprecipitated from the bone marrow cell lysates and its acetyltransferase activity of the immunoprecipitates was measured by UFLC. In parallel, CREBBP protein content in the immunoprecipitates was determined by Western-blot (Fig. 6A, Left panel). Results are means of the six treated (Tn) and non-treated (Cn) mice. Error bars indicate S.D. $**p < 0.01$ compared with control (Ctrl)

(C) Chromatin immunoprecipitation (ChIP) was carried on bone marrow cells chromatin using antibodies against the H3K18Ac and H3K27Ac histone marks as described in *Materials and Methods*. RT-qPCR experiments were subsequently carried out on murine *COPS2* and *MDM2* genes using primers specific for their promoters. Results are means of the six treated (Tn) and non-treated (Cn) mice. Error bars indicate S.D. values. $**p < 0.01$ compared with control (Ctrl), $***p < 0.001$ compared with Ctrl.

MDM2 genes of etoposide-treated mice (Fig. 6C). Taken together, these results are consistent with the data obtained with recombinant CREBBP and support the notion that metabolic activation of etoposide into oxidative reactive metabolites, notably etoposide quinone, leads to impairment of CREBBP HAT activity with subsequent aberrant histone acetylation.

4. Discussion

Despite its strong potential as a chemotherapeutic drug, etoposide is well-known to be associated with the development of treatment-related hematopoietic malignancies [22]. The molecular events underlying etoposide-dependent leukemogenesis remain poorly understood, recent studies suggest nonetheless that etoposide quinone, a reactive oxidative metabolite of etoposide, could be a major player in the leukemogenic process, notably through interactions with topoisomerase II enzymes [28–30]. However, as a reactive electrophilic species etoposide quinone likely exerts its toxic effects through alteration of other cellular protein targets [24,32,34,35,50,51]. In addition, recent studies suggest that exposure to etoposide could lead to epigenetic alterations, notably affecting histones, that may contribute to the deleterious effects of the drug [37,38]. In this work, we show that the leukemogenic oxidative metabolite of etoposide, etoposide quinone, is able to impair the activity of the histone acetyltransferase CREBBP. Remarkably, this histone

acetyltransferase is recognized as a tumor suppressor that plays a crucial role in the homeostasis of the hematopoietic system [7,11,17,19,69]. Moreover, mutations and other molecular genetic abnormalities of CREBBP that alter its acetyltransferase activity are commonly observed in leukemia and lymphoma [7,9,10,14]. Interestingly, it has been shown that the histone acetyltransferase p300, the close homologue of CREBBP, is irreversibly inhibited by quinone species such as hydroxynaphthoquinones [70,71]. Our in depth enzymatic analyses indicate that the inhibition of CREBBP by etoposide quinone is irreversible and takes places rapidly with a second-order inhibition rate constant of $16.10^2 \text{ M}^{-1} \cdot \text{s}^{-1}$ consistent with the value found for the irreversible inhibition of topoisomerase II α by etoposide quinone ($28.10^2 \text{ M}^{-1} \cdot \text{s}^{-1}$) [29]. Interestingly, topoisomerases II are known biological targets of etoposide quinone and impairment of these nuclear enzymes by etoposide quinone have been suggested to contribute to etoposide-related leukemogenesis [28–30]. The biological and toxicological effects of numerous quinone species, such as etoposide quinone, are known to depend on their capacity to react rapidly and covalently with protein cysteine residues, to form irreversible adducts that impact protein structure and/or functions [50,51]. Consistently, the irreversible inhibition of topoisomerase II α by etoposide quinone has been suggested to rely on quinone adduction of several reactive cysteine residues of the enzyme [29,72]. In a similar manner, our mass spectrometry and labeling experiments showed that etoposide quinone was indeed able to form covalent adducts with two

cysteine residues within the RING and PHD domains of the CREBBP catalytic core (Figs. 3 and 4). The RING and PHD domains are each constituted of a redox sensitive C4 Zn²⁺-fingers (4 cysteines coordinating the Zn²⁺ atom) that are known to be an integral part of the CREBBP catalytic core and to be essential for its acetyltransferase activity [41,42,48]. Of importance, the two cysteines we identified (Cys1220 and Cys1284) belong, to the RING and PHD Zn²⁺-fingers, respectively where they coordinate the Zn²⁺ atom. Importantly, it was previously shown that mutation of Cys1284 to Ala results in CREBBP inactivation, thus supporting our findings [41]. Moreover, comparatively to our findings, it has been shown that Zn²⁺-bound cysteines in Zn²⁺-fingers readily react with quinone species to form covalent adducts that can lead to Zn²⁺ ejection and subsequent alteration of protein function [56,57]. More generally, release of Zn²⁺ from C4 Zn²⁺-fingers (such as those in the RING and PHD domains of CREBBP) and their subsequent structural alterations are commonly observed when Zn²⁺-finger cysteines are adducted by electrophilic chemicals [56,57]. We indeed observed that reaction of etoposide-quinone with RING and PHD Zn²⁺-finger cysteines of CREBBP led to Zn²⁺ ejection (Fig. 4). Interestingly, it has been reported recently that certain quinones species, were able to react with a Zn²⁺-finger in the catalytic domain of the histone demethylase KDM4A and to induce similar Zn²⁺ ejection [73]. In line with these observations, our results strongly suggest that impairment of CREBBP activity is likely due to the adduction of the bulky etoposide quinone molecule on these critical RING and PHD Zn²⁺-finger cysteine residues, resulting in subsequent release of Zn²⁺ and alteration of the RING/PHD structures.

As shown for other well-known chemical leukemogens such as benzene, the leukemogenic effects of etoposide are increasingly recognized as relying on the metabolic activation of etoposide into reactive oxidative metabolites in the bone marrow, in particular etoposide quinone [24–26,30,31]. Recent studies indicate that etoposide quinone is more potent at impairing topoisomerase II functions than the parent drug etoposide, thus supporting a major role for this quinone metabolite on the topoisomerase II-dependent leukemogenic process of etoposide [28, 30]. In accordance with these observations, we were unable to find any significant effect of etoposide on purified recombinant CREBBP activity *in vitro* in contrast to etoposide quinone. This is in agreement with the electrophilic properties of the quinone metabolite which prompt this species to react covalently with redox sensitive Zn²⁺-finger cysteines [50,51,56]. As stated above, etoposide is readily oxidized in hematopoietic cells into etoposide quinone by cellular cytochromes P450 and peroxidases [24,25,63]. In line with the data obtained with recombinant purified CREBBP, exposure of hematopoietic cells to etoposide led to the inhibition of the endogenous enzyme with concomitant decrease in histone H3K18 and H3K27 acetylation (Fig. 5). Further experiments with mice exposed to etoposide confirmed inhibition of CREBBP acetyltransferase activity and decreased H3K18 and H3K27 acetylation (Fig. 6). Interestingly, a similar decrease in histone H3 acetylation was recently reported in a human neuroblastoma cell line exposed to etoposide [39].

Overall, our study shows that, in addition to the alteration of topoisomerase II functions, etoposide quinone, a leukemogenic oxidative metabolite of etoposide can impair CREBBP acetyltransferase by reacting with key cysteine residues within the catalytic core of the enzyme.

Declaration of competing interest

None declared.

Acknowledgments

This work was supported by running grants from Université de Paris and CNRS, and by grants from ANSES (Agence Nationale de Sécurité Sanitaire de l'Alimentation, de l'Environnement et du Travail) and ITMO Cancer (plan Cancer-Environnement). WZ, PG and RL are

supported by China Scholarship Council PhD fellowships. CM is supported by a PhD fellowship from Université de Paris (ED BioSPC). We thank the technical platform “BioProfiler” (BFA, CNRS UMR 8251 Unit, Université de Paris) for provision of UFLC facilities. We thank Dr. O. Brookes and Dr. C. Mathieu for English language proofreading and comments. We thank Dr. F. Busi for help with PCR primers design and ordering.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.freeradbiomed.2020.11.027>.

References

- [1] P. Chi, C.D. Allis, G.G. Wang, Covalent histone modifications—miswritten, misinterpreted and mis-erased in human cancers, *Nat. Rev. Canc.* 10 (2010) 457–469.
- [2] E. Verdin, M. Ott, 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond, *Nat. Rev. Mol. Cell Biol.* 16 (2015) 258–264.
- [3] W. Gu, R.G. Roeder, Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain, *Cell* 90 (1997) 595–606.
- [4] V.V. Ogryzko, R.L. Schiltz, V. Russanova, B.H. Howard, Y. Nakatani, The transcriptional coactivators p300 and CBP are histone acetyltransferases, *Cell* 87 (1996) 953–959.
- [5] B.M. Dancy, P.A. Cole, Protein lysine acetylation by p300/CBP, *Chem. Rev.* 115 (2015) 2419–2452.
- [6] J.H. Roelfsema, D.J. Peters, Rubinstein-Taybi syndrome: clinical and molecular overview, *Expert Rev. Mol. Med.* 9 (2007) 1–16.
- [7] R. Dutta, B. Tiu, K.M. Sakamoto, CBP/p300 acetyltransferase activity in hematologic malignancies, *Mol. Genet. Metabol.* 119 (2016) 37–43.
- [8] I. Garcia-Ramirez, S. Tadros, I. Gonzalez-Herrero, A. Martin-Lorenzo, G. Rodriguez-Hernandez, D. Moore, L. Ruiz-Roca, O. Blanco, D. Alonso-Lopez, J.L. Rivas, et al., Crebbp loss cooperates with Bcl2 overexpression to promote lymphoma in mice, *Blood* 129 (2017) 2645–2656.
- [9] C.G. Mullighan, J. Zhang, L.H. Kasper, S. Lerach, D. Payne-Turner, L.A. Phillips, S. L. Heatley, L. Holmfeldt, J.R. Collins-Underwood, J. Ma, et al., CREBBP mutations in relapsed acute lymphoblastic leukaemia, *Nature* 471 (2011) 235–239.
- [10] L. Pasqualucci, V. Trifonov, G. Fabbri, J. Ma, D. Rossi, A. Chiarenza, V.A. Wells, A. Grunn, M. Messina, O. Elliott, et al., Analysis of the coding genome of diffuse large B-cell lymphoma, *Nat. Genet.* 43 (2011) 830–837.
- [11] J. Zhang, S. Vlasevska, V.A. Wells, S. Nataraj, A.B. Holmes, R. Duval, S.N. Meyer, T. Mo, K. Basso, P.K. Brindle, et al., The CREBBP acetyltransferase is a haploinsufficient tumor suppressor in B-cell lymphoma, *Canc. Discov.* 7 (2017) 322–337.
- [12] H. Hashwah, C.A. Schmid, S. Kasser, K. Bertram, A. Stelling, M.G. Manz, A. Muller, Inactivation of CREBBP expands the germinal center B cell compartment, down-regulates MHCII expression and promotes DLBCL growth, *Proc. Natl. Acad. Sci. U. S. A* 114 (2017) 9701–9706.
- [13] S.J. Horton, G. Giotopoulos, H. Yun, S. Vohra, O. Sheppard, R. Bashford-Rogers, M. Rashid, A. Clipson, W.I. Chan, D. Sasca, et al., Early loss of Crebbp confers malignant stem cell properties on lymphoid progenitors, *Nat. Cell Biol.* 19 (2017) 1093–1104.
- [14] A. Inthal, P. Zeithofer, M. Zeginigg, M. Morak, R. Grausenburger, E. Fronkova, B. Fahrner, G. Mann, O.A. Haas, R. Panzer-Grumayer, CREBBP HAT domain mutations prevail in relapse cases of high hyperdiploid childhood acute lymphoblastic leukemia, *Leukemia* 26 (2012) 1797–1803.
- [15] M. Qian, H. Zhang, S.K. Kham, S. Liu, C. Jiang, X. Zhao, Y. Lu, C. Goodings, T. N. Lin, R. Zhang, et al., Whole-transcriptome sequencing identifies a distinct subtype of acute lymphoblastic leukemia with predominant genomic abnormalities of EP300 and CREBBP, *Genome Res.* 27 (2017) 185–195.
- [16] X.J. Yang, Lysine acetylation and the bromodomain: a new partnership for signaling, *Bioessays: news and reviews in molecular, cellular and developmental biology* 26 (2004) 1076–1087.
- [17] A.L. Kung, V.I. Rebel, R.T. Bronson, L.E. Ch'ng, C.A. Sieff, D.M. Livingston, T. P. Yao, Gene dose-dependent control of hematopoiesis and hematologic tumor suppression by CBP, *Genes Dev.* 14 (2000) 272–277.
- [18] V.I. Rebel, A.L. Kung, E.A. Tanner, H. Yang, R.T. Bronson, D.M. Livingston, Distinct roles for CREB-binding protein and p300 in hematopoietic stem cell self-renewal, *Proc. Natl. Acad. Sci. U. S. A* 99 (2002) 14789–14794.
- [19] S.N. Zimmer, Q. Zhou, T. Zhou, Z. Cheng, S.L. Abboud-Werner, D. Horn, M. Lecoche, R. White, A.V. Krivtsov, S.A. Armstrong, et al., Crebbp haploinsufficiency in mice alters the bone marrow microenvironment, leading to loss of stem cells and excessive myelopoiesis, *Blood* 118 (2011) 69–79.
- [20] E.L. Baldwin, N. Osheroff, Etoposide, topoisomerase II and cancer. *Current medicinal chemistry, Anti-cancer agents* 5 (2005) 363–372.
- [21] A.R. Mistry, C.A. Felix, R.J. Whitmarsh, A. Mason, A. Reiter, B. Cassinat, A. Parry, C. Walz, J.L. Wiemels, M.R. Segal, et al., DNA topoisomerase II in therapy-related acute promyelocytic leukemia, *N. Engl. J. Med.* 352 (2005) 1529–1538.
- [22] M. Pendleton, R.H. Lindsey Jr., C.A. Felix, D. Grimwade, N. Osheroff, Topoisomerase II and leukemia, *Ann. N. Y. Acad. Sci.* 1310 (2014) 98–110.

- [23] M.A. Smith, L. Rubinstein, J.R. Anderson, D. Arthur, P.J. Catalano, B. Freidlin, R. Heyn, A. Khayat, M. Krailo, V.J. Land, et al., Secondary leukemia or myelodysplastic syndrome after treatment with epipodophyllotoxins, *J. Clin. Oncol.* : official journal of the American Society of Clinical Oncology 17 (1999) 569–577.
- [24] M. Atwal, E.L. Lishman, C.A. Austin, I.G. Cowell, Myeloperoxidase enhances etoposide and mitoxantrone-mediated DNA damage: a target for myeloprotection in cancer chemotherapy, *Mol. Pharmacol.* 91 (2017) 49–57.
- [25] V.E. Kagan, A.I. Kuzmenko, Y.Y. Tyurina, A.A. Shvedova, T. Matsura, J. C. Yalowich, Pro-oxidant and antioxidant mechanisms of etoposide in HL-60 cells: role of myeloperoxidase, *Canc. Res.* 61 (2001) 7777–7784.
- [26] M.V. Relling, H.L. McLeod, L.C. Bowman, V.M. Santana, Etoposide pharmacokinetics and pharmacodynamics after acute and chronic exposure to cisplatin, *Clin. Pharmacol. Ther.* 56 (1994) 503–511.
- [27] C.A. Felix, A.H. Walker, B.J. Lange, T.M. Williams, N.J. Winick, N.K. Cheung, B. D. Lovett, P.C. Nowell, I.A. Blair, T.R. Rebbeck, Association of CYP3A4 genotype with treatment-related leukemia, *Proc. Natl. Acad. Sci. U. S. A* 95 (1998) 13176–13181.
- [28] E.G. Gibson, M.M. King, S.L. Mercer, J.E. Dewese, Two-mechanism model for the interaction of etoposide quinone with topoisomerase II α , *Chem. Res. Toxicol.* 29 (2016) 1541–1548.
- [29] D.A. Jacob, S.L. Mercer, N. Osheroff, J.E. Dewese, Etoposide quinone is a redox-dependent topoisomerase II poison, *Biochemistry* 50 (2011) 5660–5667.
- [30] N.A. Smith, J.A. Byl, S.L. Mercer, J.E. Dewese, N. Osheroff, Etoposide quinone is a covalent poison of human topoisomerase II β , *Biochemistry* 53 (2014) 3229–3236.
- [31] Vlasova II, W.H. Feng, J.P. Goff, A. Giorgianni, D. Do, S.M. Gollin, D.W. Lewis, V. E. Kagan, J.C. Yalowich, Myeloperoxidase-dependent oxidation of etoposide in human myeloid progenitor CD34+ cells, *Mol. Pharmacol.* 79 (2011) 479–487.
- [32] A.E. Alegria, P. Sanchez-Cruz, A. Kumar, C. Garcia, F.A. Gonzalez, A. Orellano, B. Zayas, M. Gordaliza, Thiols oxidation and covalent binding of BSA by cyclolignanic quinones are enhanced by the magnesium cation, *Free Radic. Res.* 42 (2008) 70–81.
- [33] A. Montecucco, F. Zanetta, G. Biamonti, Molecular mechanisms of etoposide, *EXCLI journal* 14 (2015) 95–108.
- [34] Q. Nian, J. Berthelet, W. Zhang, L.C. Bui, R. Liu, X. Xu, R. Duval, S. Ganesan, T. Leger, C. Chomienne, et al., T-cell protein tyrosine phosphatase is irreversibly inhibited by etoposide-quinone, a reactive metabolite of the chemotherapy drug etoposide, *Mol. Pharmacol.* 96 (2019) 297–306.
- [35] E. Rojas, P. Mussali, E. Tovar, M. Valverde, DNA-AP sites generation by etoposide in whole blood cells, *BMC Canc.* 9 (2009) 398.
- [36] B.K. Sinha, C.E. Myers, Irreversible binding of etoposide (VP-16-213) to deoxyribonucleic acid and proteins, *Biochem. Pharmacol.* 33 (1984) 3725–3728.
- [37] E. Chamani, A. Rabbani-Chadegani, Z. Zahraei, Spectroscopic detection of etoposide binding to chromatin components: the role of histone proteins, *Spectrochim. Acta Mol. Biomol. Spectrosc.* 133 (2014) 292–299.
- [38] H.J. Gill Super, A role for epigenetics in the formation of chromosome translocations in acute leukemia, *Cancer genetics* 208 (2015) 230–236.
- [39] T. Groh, J. Hrabeta, M.A. Khalil, H. Doktorova, T. Eckschlager, M. Stiborova, The synergistic effects of DNA-damaging drugs cisplatin and etoposide with a histone deacetylase inhibitor valproate in high-risk neuroblastoma cells, *Int. J. Oncol.* 47 (2015) 343–352.
- [40] N. Bagheri-Sereshtki, B.F. Hales, B. Robaire, The effects of chemotherapeutic agents, bleomycin, etoposide, and cisplatin, on chromatin remodeling in male rat germ cells, *Biol. Reprod.* 94 (2016) 81.
- [41] E. Kalkhoven, H. Teunissen, A. Houweling, C.P. Verrijzer, A. Zantema, The PHD type zinc finger is an integral part of the CBP acetyltransferase domain, *Mol. Cell Biol.* 22 (2002) 1961–1970.
- [42] E. Kalkhoven, J.H. Roelfsema, H. Teunissen, A. den Boer, Y. Ariyurek, A. Zantema, M.H. Breuning, R.C. Hennekam, D.J. Peters, Loss of CBP acetyltransferase activity by PHD finger mutations in Rubinstein-Taybi syndrome, *Hum. Mol. Genet.* 12 (2003) 441–450.
- [43] M. Delvecchio, J. Gaucher, C. Aguilar-Gurrieri, E. Ortega, D. Panne, Structure of the p300 catalytic core and implications for chromatin targeting and HAT regulation, *Nat. Struct. Mol. Biol.* 20 (2013) 1040–1046.
- [44] R. Duval, L. Fritsch, L.C. Bui, J. Berthelet, F. Guidez, C. Mathieu, J.M. Dupret, C. Chomienne, S. Ait-Si-Ali, F. Rodrigues-Lima, An acetyltransferase assay for CREB-binding protein based on reverse phase-ultra-fast liquid chromatography of fluorescent histone H3 peptides, *Anal. Biochem.* 486 (2015) 35–37.
- [45] R.A. Copeland, Evaluation of enzyme inhibitors in drug discovery. A guide for medicinal chemists and pharmacologists, *Methods Biochem. Anal.* 46 (2005) 1–265.
- [46] M.A. Paz, R. Fluckiger, A. Boak, H.M. Kagan, P.M. Gallop, Specific detection of quinoproteins by redox-cycling staining, *J. Biol. Chem.* 266 (1991) 689–692.
- [47] V.E. Kagan, J.C. Yalowich, G.G. Borisenko, Y.Y. Tyurina, V.A. Tyurin, P. Thampatty, J.P. Fabisiak, Mechanism-based chemopreventive strategies against etoposide-induced acute myeloid leukemia: free radical/antioxidant approach, *Mol. Pharmacol.* 56 (1999) 494–506.
- [48] S. Park, R.L. Stanfield, M.A. Martinez-Yamout, H.J. Dyson, I.A. Wilson, P.E. Wright, Role of the CBP catalytic core in intramolecular SUMOylation and control of histone H3 acetylation, *Proc. Natl. Acad. Sci. U. S. A* 114 (2017) E5335–E5342.
- [49] M.S. Elgawish, N. Kishikawa, K. Ohyama, N. Kuroda, Characterization of quinone derived protein adducts and their selective identification using redox cycling based chemiluminescence assay, *Journal of chromatography. A* 1403 (2015) 96–103.
- [50] J.L. Bolton, T. Dunlap, Formation and biological targets of quinones: cytotoxic versus cytoprotective effects, *Chem. Res. Toxicol.* 30 (2017) 13–37.
- [51] N. Shu, L.G. Lorentzen, M.J. Davies, Reaction of quinones with proteins: kinetics of adduct formation, effects on enzymatic activity and protein structure, and potential reversibility of modifications, *Free Radic. Biol. Med.* 137 (2019) 169–180.
- [52] J.L. Bolton, M.A. Trush, T.M. Penning, G. Dryhurst, T.J. Monks, Role of quinones in toxicology, *Chem. Res. Toxicol.* 13 (2000) 135–160.
- [53] X. Wang, B. Thomas, R. Sachdeva, L. Arterburn, L. Frye, P.G. Hatcher, D. G. Cornwell, J. Ma, Mechanism of arylating quinone toxicity involving Michael adduct formation and induction of endoplasmic reticulum stress, *Proc. Natl. Acad. Sci. U. S. A* 103 (2006) 3604–3609.
- [54] I. Klopick, M.S. Dolenc, Chemicals and drugs forming reactive quinone and quinone imine metabolites, *Chem. Res. Toxicol.* 32 (2019) 1–34.
- [55] Y. Wu, K.S. Kwon, S.G. Rhee, Probing cellular protein targets of H₂O₂ with fluorescein-conjugated iodoacetamide and antibodies to fluorescein, *FEBS Lett.* 440 (1998) 111–115.
- [56] S.M. Quintal, Q.A. dePaula, N.P. Farrell, Zinc finger proteins as templates for metal ion exchange and ligand reactivity. Chemical and biological consequences, *Metall : integrated biomaterial science* 3 (2011) 121–139.
- [57] H. Lee, J. Komano, Y. Saitoh, S. Yamaoka, T. Kozaki, T. Misawa, M. Takahama, T. Satoh, O. Takeuchi, N. Yamamoto, et al., Zinc-finger antiviral protein mediates retinoic acid inducible gene I-like receptor-independent antiviral response to murine leukemia virus, *Proc. Natl. Acad. Sci. U. S. A* 110 (2013) 12379–12384.
- [58] K. Miyamoto, K. Togiya, The creation of the artificial RING finger from the cross-brace zinc finger by alpha-helical region substitution, *Biochem. Biophys. Res. Commun.* 394 (2010) 972–975.
- [59] N. Haim, J. Nemeč, J. Roman, B.K. Sinha, Peroxidase-catalyzed metabolism of etoposide (VP-16-213) and covalent binding of reactive intermediates to cellular macromolecules, *Canc. Res.* 47 (1987) 5835–5840.
- [60] N. Zheng, C.A. Felix, S. Pang, R. Boston, P. Moate, J. Scavuzzo, I.A. Blair, Plasma etoposide catechol increases in pediatric patients undergoing multiple-day chemotherapy with etoposide, *Clin. Canc. Res. : an official journal of the American Association for Cancer Research* 10 (2004) 2977–2985.
- [61] Y. Fan, E.M. Schreiber, A. Giorgianni, J.C. Yalowich, B.W. Day, Myeloperoxidase-catalyzed metabolism of etoposide to its quinone and glutathione adduct forms in HL60 cells, *Chem. Res. Toxicol.* 19 (2006) 937–943.
- [62] B.D. Lovett, L. Lo Nigro, E.F. Rappaport, I.A. Blair, N. Osheroff, N. Zheng, M. D. Magonigal, W.R. Williams, P.C. Nowell, C.A. Felix, Near-precise interchromosomal recombination and functional DNA topoisomerase II cleavage sites at MLL and AF-4 genomic breakpoints in treatment-related acute lymphoblastic leukemia with t(4;11) translocation, *Proc. Natl. Acad. Sci. U. S. A* 98 (2001) 9802–9807.
- [63] X. Zhuo, N. Zheng, C.A. Felix, I.A. Blair, Kinetics and regulation of cytochrome P450-mediated etoposide metabolism, *Drug Metabol. Dispos.: the biological fate of chemicals* 32 (2004) 993–1000.
- [64] M.M. Mc Gee, G. Campiani, A. Ramunno, C. Fattorusso, V. Nacci, M. Lawler, D. C. Williams, D.M. Zisterer, Pyrrolo-1,5-benzoxazepines induce apoptosis in chronic myelogenous leukemia (CML) cells by bypassing the apoptotic suppressor bcr-abl, *J. Pharmacol. Exp. Therapeut.* 296 (2001) 31–40.
- [65] M.F. Barber, E. Michishita-Kioi, Y. Xi, L. Tasselli, M. Kioi, Z. Moqtaderi, R. I. Tennen, S. Paredes, N.L. Young, K. Chen, et al., SIRT7 links H3K18 deacetylation to maintenance of oncogenic transformation, *Nature* 487 (2012) 114–118.
- [66] Y. Liu, Z.B. Xing, J.H. Zhang, Y. Fang, Akt kinase targets the association of CBP with histone H3 to regulate the acetylation of lysine K18, *FEBS Lett.* 587 (2013) 847–853.
- [67] A. Sriraman, A. Dickmanns, Z. Najafova, S.A. Johnsen, M. Dobbstein, CDK4 inhibition diminishes p53 activation by MDM2 antagonists, *Cell Death Dis.* 9 (2018) 918.
- [68] F. Marchetti, J.B. Bishop, X. Lowe, W.M. Generoso, J. Hozier, A.J. Wyrobek, Etoposide induces heritable chromosomal aberrations and aneuploidy during male meiosis in the mouse, *Proc. Natl. Acad. Sci. U. S. A* 98 (2001) 3952–3957.
- [69] W.I. Chan, R.L. Hannah, M.A. Dawson, C. Pridans, D. Foster, A. Joshi, B. Gottgens, J.M. Van Deursen, B.J. Huntly, The transcriptional coactivator Cbp regulates self-renewal and differentiation in adult hematopoietic stem cells, *Mol. Cell Biol.* 31 (2011) 5046–5060.
- [70] K.C. Ravindra, B.R. Selvi, M. Arif, B.A. Reddy, G.R. Thanuja, S. Agrawal, S. K. Pradhan, N. Nagashayana, D. Dasgupta, T.K. Kundu, Inhibition of lysine acetyltransferase KAT3B/p300 activity by a naturally occurring hydroxynaphthoquinone, plumbagin, *J. Biol. Chem.* 284 (2009) 24453–24464.
- [71] M.D. Vasudevarao, P. Mizar, S. Kumari, S. Mandal, S. Siddhanta, M.M. Swamy, S. Kaypee, R.C. Kodihalli, A. Banerjee, C. Naryana, et al., Naphthoquinone-mediated inhibition of lysine acetyltransferase KAT3B/p300, basis for non-toxic inhibitor synthesis, *J. Biol. Chem.* 289 (2014) 7702–7717.
- [72] R.P. Bender, A.J. Ham, N. Osheroff, Quinone-induced enhancement of DNA cleavage by human topoisomerase II α : adduction of cysteine residues 392 and 405, *Biochemistry* 46 (2007) 2856–2864.
- [73] M.K. Jayatunga, S. Thompson, T.C. McKee, M.C. Chan, K.M. Reece, A.P. Hardy, R. Sekirnik, P.T. Seden, K.M. Cook, J.B. McMahon, et al., Inhibition of the HIF1 α -p300 interaction by quinone- and indandione-mediated ejection of structural Zn(II), *Eur. J. Med. Chem.* 94 (2015) 509–516.

SUPPLEMENTARY DATA

Human CREBBP acetyltransferase is impaired by etoposide quinone, an oxidative and leukemogenic metabolite of the anticancer drug etoposide through modification of redox-sensitive zinc-finger cysteine residues

Wenchao ZHANG¹, Jérémy BERTHELET^{1,2}, Christina MICHAÏL¹, Linh-Chi BUI¹, Panhong GOU³, Rongxing LIU¹, Romain DUVAL⁴, Justine RENAULT¹, Jean-Marie DUPRET¹, Fabien GUIDEZ³, Christine CHOMIENNE^{3,5} and Fernando RODRIGUES LIMA¹

¹ Université de Paris, BFA, UMR 8251, CNRS, F-75013 Paris, France

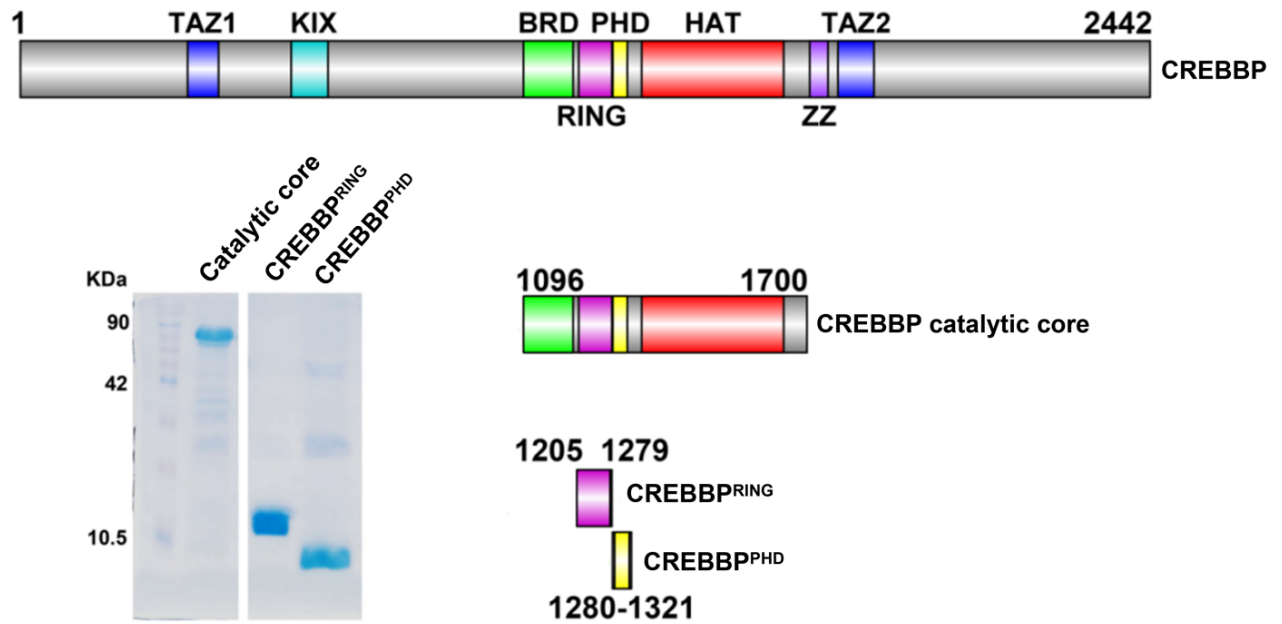
² Université de Paris, CEDC, UMR 7216, CNRS, F-75013 Paris, France

³ Université de Paris, Institut de Recherche Saint-Louis, UMRS 1131, INSERM, F-75010 Paris, France

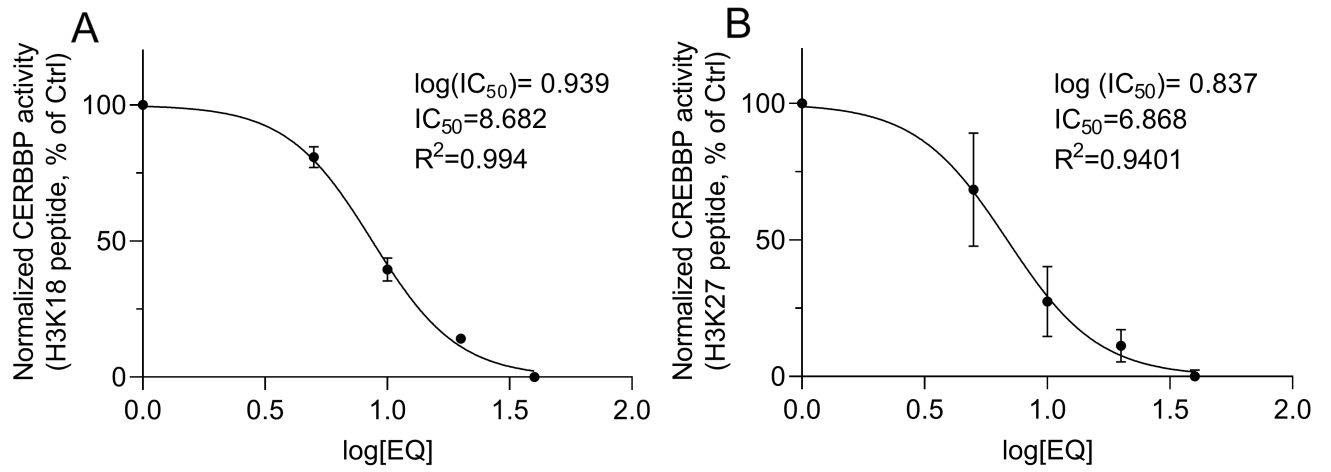
⁴ Université de Paris, BIGR, UMRS 1134, INSERM, F-75015 Paris, France

⁵ Service de Biologie Cellulaire, Assistance Publique des Hôpitaux de Paris (AP-HP), Hôpital Saint Louis, Paris, France.

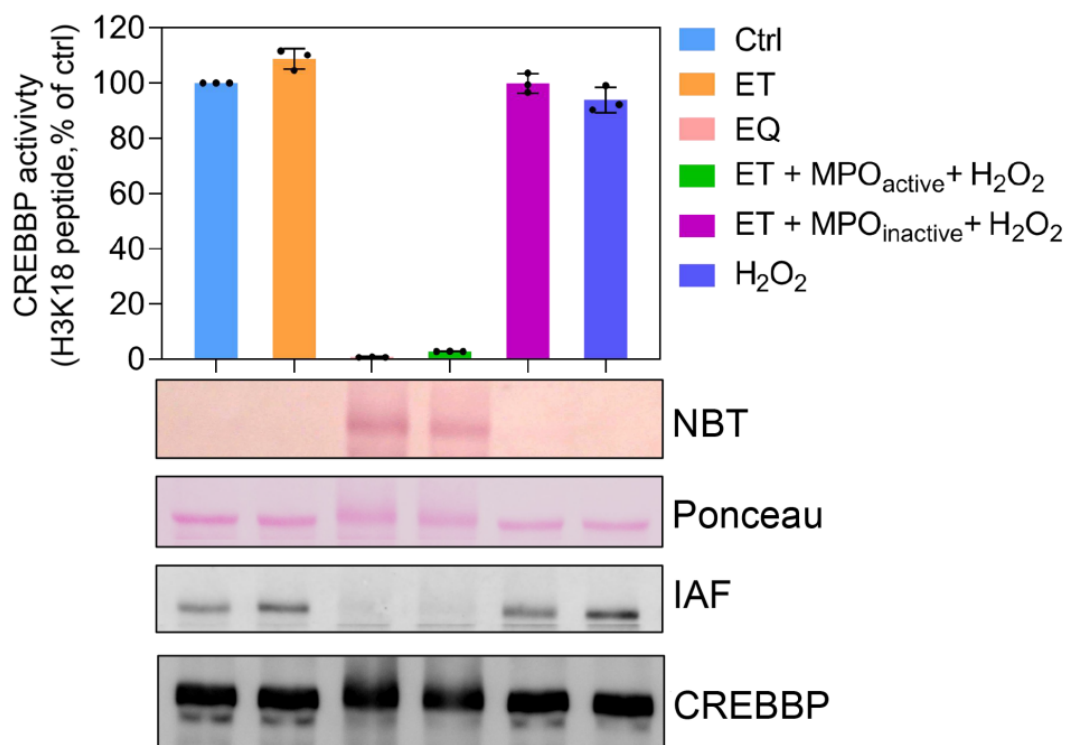
Correspondence and request for materials should be addressed to FRL (fernando.rodrigues-lima@u-paris.fr)



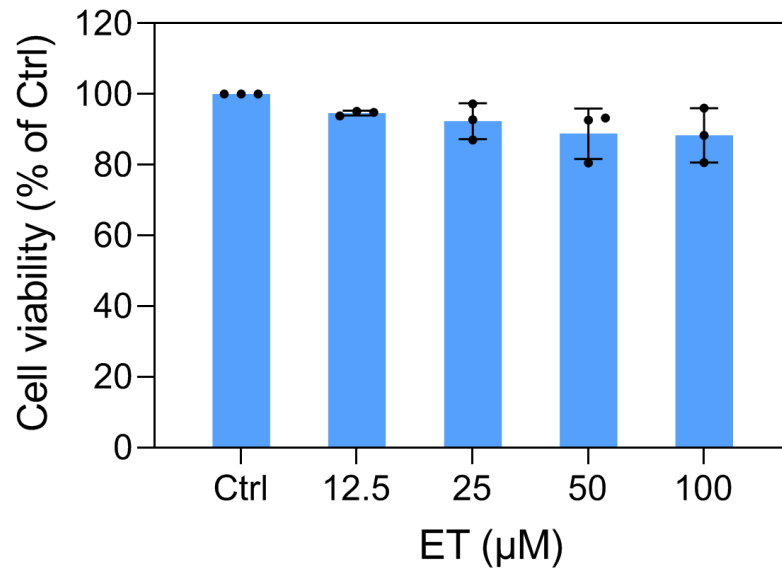
Supplementary Figure 1. Representation of the different domains of human CREBBP. The Commaise staining of the different CREBBP protein used in this study is also shown (CREBBP catalytic core, aa 1096-1700; CREBBP RING domain, aa 1205-1279 and CREBBP PHD domain, aa 1280-1321).



Supplementary Figure 2. Semi-logarithmic plot of the inhibition of CREBBP by etoposide quinone and determination of IC_{50} values. Error bars indicate S.D. values.



Supplementary Figure 3. Bioactivation of etoposide to etoposide quinone by myeloperoxidase inhibits CREBBP activity. Etoposide (100 μ M) was co-incubated with 5 U/ml myeloperoxidase (active or inactivated by boiling) and 200 μ M H₂O₂ in 100 mM sodium acetate buffer pH 6.5 at 37°C for 30 min. The reaction mixture was taken and incubated with CREBBP at room temperature for 20 min. Residual CREBBP activity was measured by UFLC using H3K18 fluorescent peptide. Controls were carried out with etoposide quinone, etoposide and H₂O₂. Results are means of three independent experiments. Error bars indicate S.D. values. *** p <0.001 compared with control (Ctrl). NBT and Ponceau stains, IAF labelling, and anti-6xhistidine-tag western-blot were carried out with the different samples as described in Materials and Methods. Membranes shown are representative of three independent experiments.



Supplementary Figure 4. Effect of etoposide exposure (24h) on K562 cell viability

Cells (96 well plate, 10^4 cells per well) were exposed to etoposide for 24h in RPMI 1640. Cell viability was then assessed by WST-1 assay kit using manufacturer instructions (Roche diagnostic, France). Results are means of three independent experiments. Error bars indicate S.D. values.

#1	a ⁺	a+H ⁺	a+H ²⁺	a ²⁺	b ⁺	b ²⁺	Seq.	y ⁺	y ²⁺	y-H ⁺	y-H ²⁺	#2
1	101.07094	102.07876	51.54302	51.03911	129.06585	65.03657	Q					7
2	214.15500	215.16283	108.08505	107.58114	242.14992	121.57860	L	702.39671	351.70199	701.38888	351.19808	6
3	317.16419	318.17201	159.58964	159.08573	345.15910	173.08319	C	589.31264	295.15996	588.30482	294.65605	5
4	418.21187	419.21969	210.11348	209.60957	446.20678	223.60703	T	486.30346	243.65537	485.29563	243.15146	4
5	531.29593	532.30376	266.65552	266.15160	559.29084	280.14906	I	385.25578	193.13153	384.24795	192.62762	3
6	628.34869	629.35652	315.18190	314.67799	656.34361	328.67544	P	272.17172	136.58950	271.16389	136.08558	2
7							R	175.11895	88.06311	174.11113	87.55920	1

#1	a ⁺	a+H ⁺	a+H ²⁺	a ²⁺	b ⁺	b ²⁺	Seq.	y ⁺	y ²⁺	y-H ⁺	y-H ²⁺	#2
1	101.07094	102.07876	51.54302	51.03911	129.06585	65.03657	Q					7
2	214.15500	215.16283	108.08505	107.58114	242.14992	121.57860	L	1274.54970	637.77849	1273.54187	637.27457	6
3	889.31718	890.32500	445.66614	445.16223	917.31209	459.15969	C-Etoposide-quinone	1161.46563	581.23646	1160.45781	580.73254	5
4	990.36486	991.37268	496.18998	495.68607	1018.35977	509.68352	T	486.30346	243.65537	485.29563	243.15146	4
5	1103.44892	1104.45675	552.73201	552.22810	1131.44384	566.22556	I	385.25578	193.13153	384.24795	192.62762	3
6	1200.50169	1201.50951	601.25839	600.75448	1228.49660	614.75194	P	272.17172	136.58950	271.16389	136.08558	2
7							R	175.11895	88.06311	174.11113	87.55920	1

Supplementary table 1. Mass spectrometry fragment matches value of the RING domain generated by Mascot Software V2.5.1 (Upper panel: non treated protein; lower panel: treated with etoposide quinone). Value type, Theo. Mass [Da] and the data show the ion series.

#1	a ⁺	a+H ⁺	a+H ²⁺	a ²⁺	b ⁺	b ²⁺	Seq.	y ⁺	y ²⁺	y-H ⁺	y-H ²⁺	#2
1	87.05529	88.06311	44.53520	44.03128	115.05020	58.02874	N					13
2	202.08223	203.09006	102.04867	101.54475	230.07715	115.54221	D	1378.61951	689.81339	1377.61168	689.30948	12
3	303.12991	304.13774	152.57251	152.06859	331.12483	166.06605	T	1263.59257	632.29992	1262.58474	631.79601	11
4	416.21397	417.22180	209.11454	208.61063	444.20889	222.60808	L	1162.54489	581.77608	1161.53706	581.27217	10
5	531.24092	532.24874	266.62801	266.12410	559.23583	280.12155	D	1049.46082	525.23405	1048.45300	524.73014	9
6	628.29368	629.30151	315.15439	314.65048	656.28860	328.64794	P	934.43388	467.72058	933.42605	467.21667	8
7	757.33627	758.34410	379.67569	379.17178	785.33119	393.16923	E	837.38112	419.19420	836.37329	418.69028	7
8	854.38904	855.39686	428.20207	427.69816	882.38395	441.69561	P	708.33852	354.67290	707.33070	354.16899	6
9	1001.45745	1002.46528	501.73628	501.23236	1029.45237	515.22982	F	611.28576	306.14652	610.27793	305.64261	5
10	1100.52587	1101.53369	551.27048	550.76657	1128.52078	564.76403	V	464.21735	232.61231	463.20952	232.10840	4
11	1215.55281	1216.56063	608.78396	608.28004	1243.54772	622.27750	D	365.14893	183.07810	364.14111	182.57419	3
12	1318.56199	1319.56982	660.28855	659.78464	1346.55691	673.78209	C	250.12199	125.56463	249.11416	125.06072	2
13							K	147.11280	74.06004	146.10498	73.55613	1

#1	a ⁺	a+H ⁺	a+H ²⁺	a ²⁺	b ⁺	b ²⁺	Seq.	y ⁺	y ²⁺	y-H ⁺	y-H ²⁺	#2
1	87.05529	88.06311	44.53520	44.03128	115.05020	58.02874	N					13
2	202.08223	203.09006	102.04867	101.54475	230.07715	115.54221	D	1950.77250	975.88989	1949.76467	975.38598	12
3	303.12991	304.13774	152.57251	152.06859	331.12483	166.06605	T	1835.74556	918.37642	1834.73773	917.87250	11
4	416.21397	417.22180	209.11454	208.61063	444.20889	222.60808	L	1734.69788	867.85258	1733.69005	867.34866	10
5	531.24092	532.24874	266.62801	266.12410	559.23583	280.12155	D	1621.61381	811.31055	1620.60599	810.80663	9
6	628.29368	629.30151	315.15439	314.65048	656.28860	328.64794	P	1506.58687	753.79707	1505.57905	753.29316	8
7	757.33627	758.34410	379.67569	379.17178	785.33119	393.16923	E	1409.53411	705.27069	1408.52628	704.76678	7
8	854.38904	855.39686	428.20207	427.69816	882.38395	441.69561	P	1280.49151	640.74940	1279.48369	640.24548	6
9	1001.45745	1002.46528	501.73628	501.23236	1029.45237	515.22982	F	1183.43875	592.22301	1182.43093	591.71910	5
10	1100.52587	1101.53369	551.27048	550.76657	1128.52078	564.76403	V	1036.37034	518.68881	1035.36251	518.18489	4
11	1215.55281	1216.56063	608.78396	608.28004	1243.54772	622.27750	D	937.30192	469.15460	936.29410	468.65069	3
12	1890.71498	1891.72281	946.36504	945.86113	1918.70990	959.85859	C-Etoposide-quinone	822.27498	411.64113	821.26715	411.13722	2
13							K	147.11280	74.06004	146.10498	73.55613	1

Supplementary table 2. Mass spectrometry fragment matches value of the PHD domain generated by Mascot Software V2.5.1 (Upper panel: non treated protein; lower panel: treated with etoposide quinone). Value type, Theo. Mass [Da] and the data show the ion series.