

OGG1 et NEIL1, de nouvelles cibles pour la lutte contre le cancer : recherche et caractérisation fonctionnelle et structurale d'inhibiteurs

Julien Cros

▶ **To cite this version:**

Julien Cros. OGG1 et NEIL1, de nouvelles cibles pour la lutte contre le cancer : recherche et caractérisation fonctionnelle et structurale d'inhibiteurs. Cancer. Université d'Orléans, 2021. Français. ⟨NNT : 2021ORLE3145⟩. ⟨tel-03794054⟩

HAL Id: tel-03794054

https://theses.hal.science/tel-03794054v1

Submitted on 3 Oct 2022

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

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



HAL Authorization

UNIVERSITÉ D'ORLÉANS

ÉCOLE DOCTORALE SANTE, SCIENCES BIOLOGIQUES ET CHIMIE DU VIVANT

CENTRE DE BIOPHYSIQUE MOLECULAIRE, CNRS Orléans

THÈSE présentée par :

Julien CROS

soutenue le : 8 décembre 2021

pour obtenir le grade de : **Docteur de l'Université d'Orléans**

Discipline/ Spécialité : Biochimie

OGG1 et NEIL1, de nouvelles cibles pour la lutte contre le cancer : recherche et caractérisation fonctionnelle et structurale d'inhibiteurs

THÈSE dirigée par :

Bertrand CASTAING	Directeur de recherche, CBM, CNRS Orléans

Encadrée par :

Franck COSTE	Ingénieur de recherche, CBM, CNRS Orléans

RAPPORTEURS :

Anna CAMPALANS	Directrice de recherche, CEA, Fontenay-aux-Roses
Murat SAPARBAEV	Directeur de recherche, IGR, Villejuif

JURY :

Anna CAMPALANS	Directrice de recherche, CEA, Fontenay-aux-Roses
Bertrand CASTAING	Directeur de recherche, CBM, CNRS Orléans
Franck COSTE	Ingénieur de recherche, CBM, CNRS Orléans
Isabelle COUILLIN	Directrice de recherche, INEM, CNRS-Université d'Orléans
Murat SAPARBAEV	Directeur de recherche, IGR, Villejuif
Joanna TIMMINS	Directrice de recherche, IBS, CEA-CNRS, Grenoble

Remerciements

Ces travaux ont été réalisés au Centre de Biophysique Moléculaire d'Orléans (CBM, CNRS-UPR 4301) au sein du groupe thématique « Réparation de l'ADN : structure, fonction et dynamique ». Ils ont été financés par une bourse de la Ligue contre le cancer. Je souhaite remercier la Ligue pour sa confiance et ce financement qui m'a permis d'effectuer cette thèse dans de bonnes conditions. Je remercie également le Dr Eva Jakab Toth de m'avoir accueilli au sein du laboratoire.

Je remercie profondément le Dr Anna Campalans et le Dr Murat Saparbaev pour avoir respectivement accepté d'être rapportrice et rapporteur de mon manuscrit de thèse. Je remercie tout autant le Dr Isabelle Couillin et le Dr Joanna Timmins d'avoir également accepté de participer à l'examen de mes travaux en tant que membre du jury.

Je tiens à remercier Bertrand Castaing, mon directeur de thèse, pour m'avoir permis de réaliser ce travail de thèse et pour sa confiance. Je te remercie pour tes conseils, ta bonne humeur et tes encouragements lors de cette thèse. Un merci tout particulier pour ton implication dans la rédaction de ce manuscrit de thèse et pour ton soutien moral. Sur le plan professionnel, j'ai pu rencontrer un chercheur passionné et plein d'idées, et sur le plan humain, une personne bienveillante avec laquelle j'ai apprécié passer ces 3 années de thèse. Tu m'as montré ce qu'était la recherche académique et tu m'as transmis beaucoup de connaissances sur la réparation de l'ADN et la biochimie en général. Ton approche scientifique, ta gentillesse, tes idées et ton éthique resteront pour moi un exemple à suivre pour la suite de ma carrière. Je suis également fier d'avoir pu éviter (et donc survivre) la plupart de tes smashs au ping-pong. Il me semble même avoir pu en renvoyer certains.

Je tiens aussi à remercier Franck Coste, pour son encadrement et ses enseignements tout au long de cette thèse. Merci pour m'avoir initié à la cristallographie, à la préparation de protéines et pour toutes les connaissances techniques que j'ai pu acquérir grâce à toi durant ces 3 ans. J'ai pu beaucoup apprendre grâce à ton expérience et sans toi, cette thèse n'aurait pas pu atteindre les objectifs fixés. Bien que tu ne sois pas aussi bavard que Bertrand (sinon les pauses café ne finiraient jamais), tu resteras tout comme lui un exemple pour l'avenir scientifique que j'envisage aujourd'hui, et j'espère avoir acquis ta rigueur et ton savoir-faire. Je ne peux finir ce paragraphe sans évoquer le fameux « Alors, on en est où ? », qui m'a toujours surpris dans le labo, ou l'intégration au groupe des « Tic et Tac » ou même des « cassos » (dont j'ai l'honneur de partager la présidence avec Alan) quand tu es de bonne humeur. Merci pour ta confiance et ta patience. J'espère cependant qu'un jour tu sauras apprécier la saveur d'une tarte au citron meringuée verte et rose !

Virginie Gaudon, un merci tout particulier à toi. Grâce à toi j'ai pu découvrir les joies d'un laboratoire de biochimie et apprendre qu'on pouvait ajuster un volume sans fioles jaugées, mais avec une éprouvette ! Tu m'as mis à l'aise dès mon arrivée et tu as toujours été disponible pour m'écouter et me conseiller, sur le plan professionnel ou personnel. Merci également pour tes conseils à la paillasse, de m'avoir accepté dans le « labo des filles » et de ne pas trop avoir jugé ma paillasse « parfaitement bien rangée ». En revanche, je ne te remercie pas pour le surnom « chaton » qui a été utilisé par bien trop de personnes à mon goût…

Je souhaite également remercier Stéphane Goffinont, pour m'avoir initié aux tests d'activités et pour avoir partagé brièvement mon bureau. Tu as toujours volontiers gouté à mes expériences culinaires plus ou moins réussies et tu as toujours été là pour m'aider quand j'en avais besoin.

Je tiens à remercier les autres membres de l'équipe, Norbert Garnier et Martine Guerin. Merci pour avoir animé les pauses café et pour les anecdotes amusantes/inquiétantes et les ragots sur la vie à l'université.

Je veux également chaudement remercier Karine Loth, pour son aide et son soutien lors de cette thèse. Sans toi je n'aurais probablement pas trouvé cette thèse et ne serais pas venu sur Orléans. Merci de m'avoir hébergé pendant quelques mois et d'avoir partagé mon régime « salade free » sans rechigner ! Merci aussi pour nos discussions sur des sujets aussi capitaux que les mangas, les fims, les jeux vidéos ! Je reste cependant circonspect sur certaines de tes séries coréennes !

Je tiens aussi à exprimer un vif remerciement à Guillaume Gabant. Merci pour ton travail sur les analyses de masse, que tu as toujours effectué avec volonté et bonne humeur, sans jamais râler sur les échantillons et les manips exotiques que je te donnais. Merci pour ton soutien, tes conseils sur ces techniques auxquelles je n'étais qu'initié. J'aurais aimé pouvoir faire plus de parties de foot sur gazon ou en salle avec toi.

Je remercie Séverine Morisset-Lopez pour m'avoir fourni gracieusement certaines des banques moléculaires qui ont permis à cette thèse d'obtenir des résultats encourageants.

Je ne veux oublier personne alors je tiens à remercier l'ensemble des membres de l'équipe AMV et du CBM. J'ai pris beaucoup de plaisir à venir travailler et à interagir tous les jours avec vous.

Si le travail a eu une part importante dans ma vie, cette thèse m'a également permis de rencontrer de nombreux doctorants et des personnes formidables avec qui j'ai eu beaucoup de plaisir à partager des 3 années.

Je tiens tout d'abord à remercier les membres de l'association ADOC, Chayma, Caroline, Elodie, Justine, Saïda, Vinni, Cédric, Pinpin et Skander. Merci pour tous les moments passés ensemble, à organiser dans la joie, la bonne humeur et l'urgence absolue, des congrès, barbecue et des raclettes pour des permanents affamés. Merci aussi aux autres doctorants, ou non, que j'ai côtoyé lors de diverses soirées,Alan , Florian, Anton, Isabelle, Hugo, Maxence, Lauren, Daouda, Nadira, Charlotte, Medhi, Geoffrey, Edinson, Kevin et tant d'autres.

Justine, meilleure amie, confidente, et « partner in crime », tu as été mon plus grand soutien pendant cette thèse. Merci d'avoir soigneusement tagué mes paillasses avec tes dessins, pour les batailles d'eau sur le campus, pour avoir « improviser » des plats tous plus délicieux les uns que les autres alors que tu n'avais que des « restes », pour toutes tes blagues et ton sourire, qui m'ont permis de toujours garder le moral pendant cette thèse. Merci aussi pour ton sens de la ponctualité, pour avoir activement participé à mes blagues, pour avoir sonné avec moi chez Skander à 3 h du matin avant de partir en courant, à avoir toujours répondu présente pour un « sale coup ». Sans toi, cette thèse n'aurait pas eu autant de souvenirs joyeux et de fous rires, et tu as fait de ces 3 années un bonheur de chaque instant, au labo comme à la maison. Avec toi à mes côtés, je serais prêt à purifier hOGG1 et hNEIL1 tout le reste de ma vie ! Cela dit, j'accepte volontiers de changer ce programme si tu es d'accord, et j'ai hâte de partager l'avenir avec toi.

Je veux remercier tout particulièrement Skander, mon « bro de thèse ». Merci pour ton amitié infaillible, merci pour toutes ces soirées où nous avons bu, chanté, regardé du foot, et des fois, les trois en même temps. Je n'oublie pas que nous sommes invaincus au Beer Pong, sur deux ans d'affilée. Tu es un véritable ami, et j'ai eu l'honneur d'être ton témoin (même si cela a créé un cluster à mon nom !). Merci aussi pour l'introduction au tunisien, avec mes petits favoris « Zebi fi termtek » et « Mala Schlaka ». Si je ne suis pas sûr de pouvoir te rejoindre tout de suite pour une coloc pendant ton post doc (tu sais bien que toi et moi

allons devoir négocier dur avant que ce ne soit envisageable), je suis persuadé d'avoir trouvé un ami pour le reste de ma vie.

Alan, merci pour tout ces moments au laboratoire ou dans le bureau, passer deux années de cette thèse ensemble a été une vraie source de motivation et de joie ! (au grand dam de Franck et Bertrand, j'en ai bien peur…) Ce fut un plaisir de former avec toi les « Cassos » (avec Justine !), Tic et Tac, ou encore Perceval et Karadoc ! Merci aussi pour ces midis de déconnade et les différentes soirées géniales. Je sais que tu es un ami sur lequel je pourrai toujours compter, et pour cela je te remercie. J'ai hâte de venir te voir un peu sur Tours, et tu seras le bienvenu chez moi où que j'aille en post-doc !

Merci également à Florian, pour avoir été le principal organisateur des « week ends » (ou plans foireux selon certaines mauvais langues) entre doctorants. Ta réputation d'agence de tourisme n'est pas usurpée ! Cardiff, Lahaie, Gand, Tallin, Helsinki, Cologne et Maastricht resteront parmi les souvenirs les plus chers de cette thèse. Merci pour tous ces midis pendant plus de 3 ans à frôler la mort avec Skander et moi à la cantine. Tu es et tu resteras un ami cher, et je suis sûr que de nombreux week-ends nous attendent encore. On se rejoindra surement chez Skander, ou chez moi si je trouve une destination exotique !

Merci aussi à Saida, pour m'avoir appris des phrases en tunisien utilisables dans la vraie vie sans finir en prison ! Tout le monde sait que la véritable présidente/tyran de l'ADOC, ce n'était pas Skander, mais bien toi. J'espère que tu m'as pardonné de t'avoir offert une boite de faux chocolats en carton. Malgré le temps que ça m'a pris et la douzaine de plaquettes de chocolat que j'ai dû t'offrir pour me faire pardonner, ça en valait la peine ! Je regrette juste de ne pas avoir pu assister à ton mariage en Tunisie à cause de cette pandémie, mais la fête que tu as organisée en France a été parfaite.

Mohammed, le boucher algérien, à toi je dis ceci : Merci d'avoir épargné mes genoux lors de nos parties de foot en salle, l'usage de mes jambes est un privilège que j'apprécie toujours. Plus sérieusement merci pour ta bonne humeur et ton amitié. Ton seul défaut est d'être un peu fan de biologie, mais personne n'est parfait.

Cédric, je te dois aussi des remerciements. Merci pour nous avoir introduits, Skander et moi, à la douce joie de la présidence d'une association de doctorants. Merci pour tes histoires, ces fous rires lors des week-ends ou à la cantine.

Isabelle, je tiens aussi à te remercier. Tu as été un soutien moral lors des dernières années de cette thèse et ton amitié et ta bonne humeur me sont chères. Je te remercie également pour avoir fourni l'instrument le plus utile de mon appartement, les NERFS. Avant ta soutenance, nous organiserons une bataille au laboratoire, ne t'en fais pas. Merci aussi pour toutes les soirées jeux et pour avoir égayé la monotonie capillaire avec tes couleurs flamboyantes. Je remercie aussi Barnabé pour les dits-jeux, même s'il gagne trop souvent à mon goût.

Anton, agent du KGB ou pas, tu as été une des belles rencontres de cette thèse. Je suis heureux de te compter parmi mes amis, et je suis sûr que nous resterons en contact. Je suis fier d'avoir pu, avec Skander, te battre. Je n'ai malheureusement pas appris le russe aussi vite que tu as appris le français, mais pour honorer ta poésie et la tendresse russe je finirai par « Cyka blyat ! » Merci pour les kilogrammes de confiseries russes que tu m'as offerts, qui ont agrémentés mes pauses au labo ! J'ai en revanche un net doute sur l'intérêt gustatif de ceux fourrés avec 5 mL de vodka pure.

Je n'oublie pas mes amis lyonnais, Marie, Guillaume, Laura, Quentin, Eugénie, Sylvain ! Après les années de césure, les doctorats, heureusement que les post-docs existent comme excuse pour tous se retrouver à l'étranger ! Vous avez fait de ma scolarité, des classes prépas à l'école d'ingénieur un plaisir, et à

mon grand désarroi notre amitié ne semble pas vouloir décliner. En cela il faut bien avouer que vous êtes les « connaissances » les plus fabuleuses que je connaisse !

Je tiens aussi à remercier mes amis de toujours. A Steeven, mon frère de boules, avec qui j'ai eu beaucoup trop de « discussions philosophiques » et de plans foireux sur l'avenir, merci d'avoir toujours été là pour moi, tu sais à quel point ton amitié m'est chère. A Olivier, qui de tous a bien trop subi mon sens de la ponctualité, merci pour toutes les conneries, les bons nos moments passés ensemble, et les fous rires sur tous nos jeux en coop'. Un jour, lorsque tu auras enfin écouté mes conseils d'économie, nous ferons de Marseille Poubelle la plus belle ville du monde. A T.I.B.O pour avoir été le seul à affronter le froid suédois avec moi, merci pour ta patience envers nos blagues idiotes, et ton jus de pomme à venir. Si un jour lointain tu veux faire un JDR de qualité, je serai là pour le tester ! A Jonathan, avec qui un jour j'écrirai un livre. Vous êtes mes meilleurs amis, et je vous considère tous comme des membres de ma famille.

Je tiens également à remercier Arnaud, le toutologue, Emilie et leurs enfants. Nous ne partageons peut-être pas des liens de sang, mais c'est tout comme. J'ai cependant du mal à choisir entre « petit frère agaçant », et « parrain cool » pour Arnaud…

Pour finir, un grand merci à ma famille ! Vous avez toujours été là pour moi. Laurine, maman, papa, mamie grand-père, grand-père, mamie et papi café, Gérald, vous avez littéralement fait de moi ce que je suis aujourd'hui, bien plus que vous ne l'imaginez. Vous avez été pour moi des soutiens, des professeurs, des confidents, des mentors, des exemples à suivre, et encore aujourd'hui je n'arrive pas à imaginer une meilleure enfance et une vie d'adulte que celle que vous avez pu m'offrir. Je vous dois toutes mes réussites, aussi minces soient-elles, et je ne pourrai jamais assez vous remercier, mais j'essaie de vous rendre honneur dans tout ce que j'entreprends.

Sommaire

Liste des abréviations 5

Liste des figures 9

Introduction générale 13

I. Dommages et mécanismes de réparation de l'ADN 15

I.1. Source et formation des lésions de l'ADN 15

I.2. Mécanismes de réparation de l'ADN 21

II. Réparation par excision de base (BER) 29

II.1. Mécanisme général du BER 29

II.2. Les ADN glycosylases 32

II.3. Elimination du site abasique, nettoyage des extrémités et resynthèse 35

II.4. Ligation finale 38

II.5. Ciblage thérapeutique du système BER 38

III. hOGG1 50

III.1. Les protéines 8-oxoG-ADN glycosylases 50

III.2. Considérations structurales et fonctionnelles de hOGG1 53

III.3. Fonctions biologiques principales et impact thérapeutique de l'inhibition de hOGG1 59

IV. hNEIL1 62

IV.1. La superfamille des ADN glycosylases Fpg/Nei 62

IV.2. Considérations structurales et excision des bases endommagées 65

IV.3. Ciblage de hNEIL1 et impact thérapeutique 69

Stratégie et objectifs de la thèse 73

I. Généralités sur le criblage 75

I.1. Origine et évolution du HTS 75

I.2. Etapes principales d'un criblage à visée thérapeutique 76

II. Objectifs 79

Matériel et méthodes 81

I. Préparation des protéines hNEIL1 et hOGG1 83

I.1. Transformation bactérienne 83

I.2. Production 83

I.3. Purification 84

II. Préparation des oligonucléotides 89

II.1. Hybridation des oligonucléotides fluorescents 89

II.2. Hybridation pour études cristallographiques 89

II.3. Marquage des oligonucléotides au ^{32}P 90
III. Criblage 91
III.1. Chimiothèques utilisées 91
III.2. Conditions criblage primaire 92
III.3. Test d'activité orthogonal 95
III.4. Test d'intercalation à l'ADN 98
IV. Caractérisation des inhibiteurs 100
IV.1. Cinétique 100
IV.2. Spectrométrie de masse pour le suivi de la formation d'adduit sur hNEIL1 103
IV.3. Etude de la thermostabilité des protéines par Thermal Shift Assay (TSA) 105
IV.4. Evaluation de la constante K_D de pabAGOG par un test fluorescent 107
V. Etude structurale et cristallogenèse 108
V.1. Principe de la cristallogenèse 108
V.2. Méthodes utilisés pour la formation de cristal 109
V.3. Collecte des données de diffraction des rayons X 110
Résultats et Discussions 111
Partie 1 : Identification et caractérisation d'inhibiteurs de l'enzyme hNEIL1 113
I. Mise en place de la procédure de criblage sur la 8-oxoG-ADN glycosylase bactérienne Fpg de *Lactoccus lactis* 115
I.1. Criblage primaire 115
I.2. Confirmation des hits sur LlFpg par un test orthogonal 119
I.3. Evaluation de l'interaction à l'ADN des composés identifiés 121
II. Recherche d'inhibiteur de la protéine hNEIL1 122
II.1. Criblage primaire 122
II.2. Confirmation des hits 124
II.3. Tests d'intercalation et première sélection des composés 126
II.4. Dose réponse et sélection finale 127
II.5. Expansion des hits et optimisation des inhibiteurs 128
III. Analyses biochimiques et fonctionnelles des inhibiteurs de hNEIL1 132
III.1. Evaluation de la spécificité des inhibiteurs identifiés 132
III.2. Adduits covalents induits par les inhibiteurs de hNEIL1 contenant une fonction nitrile : analyse par spectrométrie de masse (MS) 134
III.3. Effet des inhibiteurs sur la stabilité thermique de hNEIL1 144
III.4. Cinétique d'inhibition de hNEIL1 par la girgensonine 145
IV. Etude structurale de la protéine hNEIL1 en interaction avec les inhibiteurs 146

IV.1. Cristallographie de la protéine seule 147
V. Bilan et conclusion 151
Partie 2 : Identification et caractérisation d'inhibiteurs de hOGG1 153
I. Criblage moléculaire sur hOGG1 155
I.1. Criblage primaire 155
I.2. Confirmation des hits sur hOGG1 157
I.3. Elimination de composés intercalant 157
I.4. Dose-réponse et sélection finale 159
I.5. Expansion des touches et optimisation des inhibiteurs 161
II. Etude du mode d'action des inhibiteurs identifiés 165
II.1. Spécificité des inhibiteurs identifiés 165
II.2. Constantes de dissociation K_D des complexes PabAGOG/inhibiteur 169
II.3. Thermostabilité de hOGG1 en présence d'inhibiteurs 171
II.4. Cinétiques enzymatiques de hOGG1 en présence d'inhibiteurs 171
III. Etude structurale de la protéine hOGG1 et de ses inhibiteurs 173
III.1. Cristallogenèse de la protéine seule 173
IV. Etude structurale d'un analogue fonctionnel de hOGG1, la 8-oxoG-ADN glycosylase d'archée AGOG 174
IV.1. Tentative de cristallisation du complexe PabAGOG/C-02G10 174
IV.2. Etude structurale par cristallographie du complexe PabAGOG/ICOA2-G5-P7 174
V. Bilan et conclusion 181
Conclusion et perspectives 183
Références bibliographiques 191
Annexes 205

Liste des abréviations

1-3D _____ en 1 à 3 dimensions

1-meA _____ 1-méthyladénine

$^1O_2$ _____ oxygène singulet

2TX _____ 2-thioxanthine

^{32}P _____ phosphate radioactif

3-meA _____ 3-méthylguanine

3-meC _____ 3-méthylcytosine

5-fU _____ 5-formyluracile

5-hmU _____ 5-hydroxyméthyluracile

5-meC _____ 5-méthylcytosine

5-OHC _____ 5-hydroxycytosine

5-OHU _____ 5-hydroxyuracile

7-meG _____ 7-méthylguanine

8-oxoA _____ 7,8-dihydro-8-oxo-adénine

8oxodG _____ 8-oxo-2'-désoxyguanosine

8-oxoG _____ 7,8-dihydro-8-oxo-guanine

A – Ala – alanine

A _____ adénine

AAG _____ 3-alkyladénine-ADN glycosylase

ADN _____ acide désoxyribonucléique

ADP _____ adénosine diphosphate

AGOG _____ 8-oxoG-ADN glycosylase d'archée

Ambinter _____ banque totale de GreenPharma

AMP _____ adénosine monophosphate

AP _____ abasique

APE1 _____ apurinique/apyrimidique endonucléase 1

ARN _____ acide ribonucléique

ATM _____ « ataxia telangiectasia mutated »

ATP _____ adénosine triphosphate

ATR _____ « ataxie télangiectasie »

BER _____ réparation par excision de base

BIR _____ « Break Induced Replication »

BRCA2 _____ « breast cancer 2 »

BSA _____ « Bovine serum albumine » ou albumine de sérum bovin

C – Cys – cystéine

C _____ cytosine

Ca^{2+} _____ ion calcium

CETN2 _____ « centrin 2 »

CID _____ dissociation induite par collision

Cm _____ Chloramphénicol

CPD _____ dimère cyclobutylique de pyrimidine

CV _____ volume de colonne

D – Asp – aspartate ou acide aspartique

DDR _____ réponse aux dommages à l'ADN

DHU _____ dihydrouracile

Dna2 _____ « ADN replication helicase 2 »

DNA-PKcs _____ protéine kinase dépendante de l'ADN

dRP _____ groupement phosphate-désoxyribose

DSB _____ cassure double brin

E – Glu – glutamate ou acide glutamique

E. coli _____ *Escherichia coli*

ERCC1 _____ « excision repair cross-complementation group 1 »

ESI _____ « ElectroSpray Ionization »

ETD _____ dissociation par transfert d'électrons

Exo1 _____ exonucléase 1

εA _____ éthénoadénine

εC _____ éthénocytosine

εdA _____ éthénodéoxyadénosine

εdC _____ éthénodéoxycytidine

F – Phe – phénylalanine

F+/F- _____ faux positif /faux négatif

FAH _____ fumarylacétoacétate hydrolase

FapyA _____ 4,6-diamino-5-formamidopyrimidine

FapyG ___ 2,6-diamino-4-hydroxy-5-formamidopyrimidine

FEN-1 _____ « flap endonuclease 1 »

$FeSO_4$ _____ sulfate de fer (II)

FPG_____ FapyG-DNA glycosylase, formamidopyrimidine-DNA *N*-glycosylase ou MutM

G – Gly – glycine

G _____ guanine

GDP _____ guanosine diphosphate

GG-NER _____ réparation par excision de nucléotides du génome global

Gh _____ guanidinohydantoine

GP _____ GPNCL, banque de l'entreprise GreenPharma

GTP _____ guanosine triphosphate

H – His – histidine

H _____ hydrogène

H_2O_2 _____ peroxyde d'hydrogène

H2TH _____ motif « Helix-2 turns-Helix » ou hélice-2 tours-hélice

HAP1 _____ APE1 humaine

HhH _____ motif « Helix-hairpin-Helix » ou à deux hélices α structurées en épingle à cheveux

HhH-GPD _____ superfamille « Helix-hairpin-Helix »

hit _____ molécule « touche »

HLR _____ « HEAT-like repeat » ou répétition de motifs similaires au motif HEAT

HMBG1 _____ « high mobility group box 1 »

hMLH _____ « MutL homolog »

HNE _____ 4-hydroxynonénal

HR _____ recombinaison homologue

HSA _____ « human serum albumine » ou albumine de sérum humain

HTS _____ « Hight Throughtput Screening » ou criblage à haut débit

Hx _____ hypoxanthine

I – Ile – isoleucine

IC _____ complexe d'interrogation

IC_{50} _____ concentration inhibitrice médiane

ICOA1 _____ ICOA-DIVER, banque de diversité de l'Institut de Chimie Organique et Analytique d'Orléans

ICOA2 _____ banque totale de l'Institut de Chimie Organique et Analytique d'Orléans

K – Lys – lysine

Kana _____ kanamycine

K_D _____ constante de dissociation entre un ligand et une protéine

K_i _____ constante d'inhibition

K_M _____ constante de Michaelis

KO _____ « knock-out »

L – Leu – leucine

LB _____ « lysogeni broth »

LCL _____ « lesion caping loop »

Lig I _____ ADN ligase I

Lig III _____ ADN ligase III

LP-BER _____ « Long Patch » BER ou BER à brèche longue

LRC _____ complexe de la reconnaissance de la lésion

LS _____ létalité synthétique

LTSS _____ étude de stabilité à long terme

M – Met – méthionine

MDA _____ malondialdéhyde

Mg^{2+} _____ ion magnésium

MGMT _____ méthylguanine-ADN-méthyltransférase

MMEJ _____ « microhomology-mediated end-joining »

MMR _____ réparation des mésappariements

MMS _____ méthylméthane-sulfonate

MPG _____ N-méthylpurine-ADN glycosylase humaine

Mre11 _____ « meiotic recombination 11 »

MRN _____ complexe Mre11/Rad50/NBS-1

MS _____ spectrométrie de masse

MSH _____ « MutS homolog »

N – Asn – asparagine

NAD^+ _____ nicotinamide adénine dinucleotide

NBS-1 _____ « Nijmegen breakage syndrome-1 »

NEIL1-3 _____ « Nei-like protein » 1 à 3

NER _____ réparation par excision de nucléotides

NHEJ _____ jonction d'extrémités non homologues

NIR _____ réparation par incision de nucléotides

NO_2 _____ dioxyde d'azote

O_2 _____ dioxygène

$O_2^{-\bullet}$ _____ ion superoxyde

O^6-Cl-eG _____ O^6-chloroéthylguanine

O^6-meG _____ O^6-méthylguanine

OGG1 _____ 8-oxoguanine DNA glycosylase

OH _____ groupement hydroxyle

$OH^\bullet$ _____ radical hydroxyle

ONO_2^- _____ peroxynitrite

P – Pro – proline

pADPr _____ poly(ADP-ribose)

PARP1/2 _____ poly(ADN-ribose) polymérase 1/2

PCNA _____ « proliferating cell nuclear agent »

PG _____ PicoGreen

PIKKS _____ phosphatidylinositol 3-kinase

PMS2 _____ « postmeiotic segregation increased 2 »

PNKP _____ polynucléotide kinase phosphatase

Pol β/δ/ε _____ ADN polymérase β/δ ou ε

Prest _____ banque Prestwick Chemical Library de Prestwick

Pt _____ platine

Q – Gln – glutamine

R – Arg – arginine

RC _____ complexe de rencontre

rédox _____ oxydoréduction

RFC _____ facteur de réplication C

ROS _____ espèce réactive de l'oxygène

RPA _____ protéine de réplication A

S – Ser – sérine

SAM _____ adénosylméthionine

SAR _____ relation structure/activité

SDSA _____ « Synthesis-Dependent Strand Annealing »

Sn1/2 _____ Substitution nucléophile de type 1/2

Sp _____ spiroimidohydantoine

SP-BER _____ « Short-Patch » BER ou BER à brèche courte

SSB _____ cassure simple brin

SSBR _____ réparation des cassures simple brin

T – Thr – thréonine

T _____ thymine

T4 Endo V _____ endonucléase V du bactériophage T4

TCEP _____ tris(2-carboxyéthyl)phosphine

TC-NER _____ réparation par excision de nucléotides couplé à la transcription

TDG _____ thymine DNA glycosylase

TDP1 _____ « tyrosyl DNA phosphodiesterase 1 »

TFIIH _____ « transcription factor II human »

Tg _____ thymine glycol

T_m _____ « Melting Temperature » ou température de fusion

TMZ _____ témozolomide

TS _____ thymidilate synthase

TSA _____ Thermal Shift Assay

U _____ uracile

UDG _____ uracile-ADN glycosylase

uHTS _____ ultra criblage à haut débit

UV _____ ultraviolet

V – Val – valine

W – Trp – tryptophane

X _____ xanthine

XRCC1 _____ « X-ray cross complementing gene 1 »

Y – Tyr – tyrosine

Z' _____ facteur équivalent au ratio signal/bruit

Zn^{2+} _____ ion zinc

ZnF _____ « Zinc finger »

ZnLF _____ « Zincless finger »

Liste des figures

Figure 1 : Représentation schématique de la composition chimique d'un nucléotide
Figure 2 : Bases nucléiques de l'ADN
Figure 3 : Appariement canonique de Watson-Crick et non canonique de Hoogsteen
Figure 4 : Principaux sites de modification de l'ADN
Figure 5 : Quelques exemples de composés provoquant des dommages sur l'ADN
Figure 6 : Propriétés d'appariements de type Hoogsteen de la 8-oxoG avec une adénine
Figure 7 : Lésions principales des bases de l'ADN
Figure 8 : Mécanisme de formation des chaînes pADPR pour l'enzyme PARP1 (tiré et traduit de Alesamova *et al.*, 2019)
Figure 9 : Schéma du mécanisme général de HR et NHEJ
Figure 10 : Schéma du mécanisme du MMR
Figure 11 : Schéma général du NER
Figure 12 : Mécanisme général du BER (adapté de Bjoras *et al.*, 2013)
Figure 13 : Les différentes familles structurales des ADN glycosylases en interaction avec l'ADN
Figure 14 : Différentes ADN glycosylases humaines
Figure 15 : Polymérases impliquées dans le BER
Figure 16 : Principe de la létalité synthétique
Figure 17 : Stratégies principales pour l'inhibition du BER (Inspirée de Visnes *et al.*, 2018)
Figure 18 : Exemple d'inhibiteurs de PARP1
Figure 19 : Inhibiteurs de UNG1/2
Figure 20 : Inhibiteurs de MPG
Figure 21 : Inhibiteurs de MTH1 et de NTHL1
Figure 22 : ADN glycosysales de la superfamille structurale HhH
Figure 23 : Schéma représentant les différents modes de diffusion de hOGG1 (inspiré de D'Augustin *et al.*, 2020)
Figure 24 : Résidus de hOGG1 interagissant avec un ADN contenant la 8-oxoG (inspiré de Koval *et al.*, 2014)
Figure 25 : Interaction de hOGG1 avec la cytosine opposée au dommage
Figure 26 : Structure cristallographique d'un complexe entre hOGG1 K249Q et un ADN contenant 8-oxoG (PDB 1EBM)
Figure 27 : Mécanisme suggéré de l'activité catalytique de hOGG1 sur la 8-oxoG par Fromme *et al.*, 2003
Figure 28 : Mécanisme monofonctionnel suggéré de l'activité catalytique de hOGG1 sur la 8-oxoG par Boiteux *et al.*, 2017
Figure 29 : Inhibiteurs identifiés de hOGG1
Figure 30 : ADN glycosylase de la super famille Fpg/Nei
Figure 31 : Triplet d'intercalation de hNEIL1 (P2G) sur un ADN contenant une Tg (PDB 5ITY)
Figure 32 : Motif ZnF/ZnLF des ADN glycosylases Fpg/Nei
Figure 33 : Mécanisme de l'activité bifonctionnelle *N*-glycosylase/AP Lyase de Fpg proposé par Boiteux *et al.*, 2017
Figure 34 : Mécanisme de l'excision de la Tg par hNEIL1 proposé par Zhu *et al.*, 2016
Figure 35 : Comparaison de la boucle LCL de hNEIL1 en conformation *224-in*, *242-in* (conformations fermées) ou de la protéine non complexée à l'ADN (conformation ouverte)
Figure 36 : Inhibiteurs identifiés de hNEIL1
Figure 37 : Matériel de micropipetage fabriqué en 1950 (adapté de Takátsy, 1950)
Figure 38 : Schéma des étapes principales d'un criblage classique pour le développement d'un médicament
Figure 39 : Composition des tampons utilisés pour la purification des protéines hNEIL1dc56 et dc95
Figure 40 : Composition des tampons utilisés pour la purification de pabAGOG
Figure 41 : Composition des tampons utilisés pour la purification des protéines d2hOGG1
Figure 42 : Composition des tampons utilisés pour la purification des protéines dhOGG1
Figure 43 : Composition des tampons utilisés pour la purification des protéines mOGG1

Figure 44 : Composition des chimiothèques utilisées et leurs plaques filles
Figure 45 : Schéma du test d'activité à la fluorescence pour hNEIL1 ou hOGG1
Figure 46 : Schéma du criblage primaire manuel ou semi-automatique
Figure 47 : Test de confirmation à la radioactivité du criblage primaire
Figure 48 : Différentes conditions des tests d'activité selon le type d'activité ou la protéine utilisée
Figure 49 : Schéma représentant le principe du test d'interaction à l'ADN par nos composés en présence de PicoGreen
Figure 50 : Ajouts successifs d'inhibiteurs pour le test d'intercalation à l'ADN
Figure 51 : Représentation graphique de l'activité d'une enzyme sur un substrat et les constantes associées
Figure 52 : Courbes théoriques pour l'évaluation de la cinétique d'une réaction enzymatique
Figure 53 : Représentation théorique d'une courbe suivant l'équation de Lineweaver-Burk
Figure 54 : Tracé des courbes de Lineweaver-Burk pour une protéine en présence d'inhibiteurs de différents mécanismes
Figure 55 : Courbe de TSA théorique et évaluation de la température de fusion d'une protéine
Figure 56 : Test des conditions optimales pour déterminer une valeur précise de T_m avec une valeur de ΔF suffisamment élevée
Figure 57 : Test des conditions optimales pour déterminer une valeur précise de T_m avec une valeur de ΔF suffisamment élevée
Figure 58 : Expression de la valeur de la constante de dissociation K_D
Figure 59 : Position des tryptophanes de pabAGOG
Figure 60 : Diagramme de phases à deux dimensions d'une macromolécule (adapté de Sauter et Giegé, 2010)
Figure 61 : Schéma de la diffusion en phase vapeur par (a) goutte suspendue ou (b) goutte assise
Figure 62 : Exemples de résultats obtenus lors des premiers essais de criblage automatique dans des puits de 30 µL
Figure 63 : Exemples de résultats obtenus pour le deuxième essai de criblage automatique pour un volume final de 60 µL
Figure 64 : Criblage automatique d'une plaque de la banque de diversité de Greenpharma sur LlFpg
Figure 65 : Criblage automatique d'une plaque d'eau ultrapure sur LlFpg
Figure 66 : Criblage manuel d'une plaque de contrôle sur LlFpg
Figure 67 : Criblage « manuel » d'une plaque de GP sur LlFpg
Figure 68 : Résultats du criblage de la banque GP sur LlFpg en pourcentage d'inhibition
Figure 69 : Confirmation des hits du criblage primaire sur LlFpg par un test d'activité orthogonal
Figure 70 : Structure de la purpurine (GP-G10-P4) et de la juglone (GP-F5-P4), deux inhibiteurs confirmés de LlFpg
Figure 71 : Test d'intercalation à l'ADN de la Juglone et de la Purpurine
Figure 72 : Détermination du volume final pour le criblage primaire semi-automatique
Figure 73 : Evaluation du ratio signal/bruit pour le criblage de hNEIL1
Figure 74 : Distribution des facteurs Z' par plaque des cribles primaires sur hNEIL1
Figure 75 : Distribution du pourcentage d'inhibition des composés du criblage primaire sur hNEIL1
Figure 76 : Confirmation des 50 hits obtenus lors du criblage primaire de hNEIL1
Figure 77 : Evaluation du pouvoir intercalant de l'ADN des hits confirmés par le test au PicoGreen
Figure 78 : Composés non-intercalant ayant été confirmés et validés lors des différentes étapes du criblage comme inhibiteurs de hNEIL1
Figure 79 : Evaluation de l'IC_{50} des inhibiteurs de hNEIL1 sortis du criblage
Figure 80 : Analogues de la Girgensonine présents dans la banque d'Ambinter
Figure 81 : Test des analogues des hits de la banque ICOA2 sur hNEIL1

Figure 82 : Réaction d'un composé aminonitrile avec le groupement thiol d'une cystéine
Figure 83 : Structure de l'inhibiteur LA-32
Figure 84 : Evaluation de l'inhibition de la girgensonine sur les enzymes hNEIL1, LlFpg et mvNei
Figure 85 : Courbe d'IC_{50} du composé LA-32 sur les ADN glycosylases LlFpg, mvNei et hNEIl1
Figure 86 : Spectre déconvolué de la protéine hNEIL1 incubée avec la molécule girgensonine
Figure 87 : Structure de la girgensonine et fragmentations potentielles
Figure 88 : Fragmentation de la girgensonine en spectrométrie de masse ESI en faisant varier la tension de l'appareil
Figure 89 : Cinétique de la formation d'adduits de la girgensonine sur hNEIL1dc95 suivie par spectrométrie de masse ESI
Figure 90 : Spectres déconvolués de l'adduit formé par la girgensonine en fonction de la température
Figure 91 : Formation d'adduits avec d'autres inhibiteurs de hNEIL1 suivie par spectrométrie de masse
Figure 92 : Fragmentation des composés LA-32 (A-1), ICOA1-D5-P2 (B-2) et ICOA1-B9-P2 (C-3) en spectrométrie de masse
Figure 93 : Couverture de fragmentation de hNEIL1 en présence de girgensonine par CID ou ETD
Figure 94 : Couverture de séquence pour la protéolyse de hNEIL1dc95 avec deux enzymes différentes
Figure 95 : Position des cystéines C18, C136 et C262 sur hNEIL1
Figure 96 : Cystéines C18, C136 et C262 et leur impact sur la conformation de hNEIL1
Figure 97 : Variation de la température de fusion (T_m) de hNEIL1 (ΔT_m) en présence de différents inhibiteurs issus du criblage ou des analogues
Figure 98 : Cinétiques obtenus avec hNEIL1 et le lecteur ClarioStar
Figure 99 : Diffraction et statistiques d'affinement de hNEIL1 lors du « soaking »
Figure 100 : Structure de hNEIL1 obtenue à 1,43 Å de résolution
Figure 101 : Diffraction de hNEIL1 en co-cristallisation avec les inhibiteurs
Figure 102 : Evaluation du ratio signal/bruit pour le criblage de hOGG1
Figure 103 : Distribution des résultats du criblage et des facteurs Z' du criblage sur hOGG1
Figure 104 : hits obtenus lors des criblages primaires
Figure 105 : Confirmation des 70 hits obtenus lors du criblage primaire de hOGG1
Figure 106 : Résultats du test d'intercalation PicoGreen sur hOGG1
Figure 107 : Structure des composés non intercalant ayant été validés lors des différentes étapes du criblage comme inhibiteur de hOGG1
Figure 108 : IC_{50} des inhibiteurs les plus efficaces du criblage sur hOGG1
Figure 109 : Structure cristallographiques de hOGG1 seule, complexée avec un duplex d'ADN contenant une 8-oxoG ou avec TH5487
Figure 110 : Influence de la concentration de la BSA sur l'inhibition de hOGG1
Figure 111 : Première expansion des hits sur hOGG1
Figure 112 : Deuxième expansion des hits sur hOGG1
Figure 113 : Test des différents analogues des ICOA2-G5-P7
Figure 114 : Récapitulatif des inhibiteurs de hOGG1
Figure 115 : Evaluation de l'inhibition des composés SU268, TH5487, C-02G10 et ICOA26H11-P29 sur hOGG1 entière, dhOGG1 et mOGG1 entière
Figure 116 : Comparaison structurale entre hOGG1, PabAGOG et mOGG1
Figure 117 : Test d'inhibition des principaux inhibiteurs de hOGG1 sur PabAGOG
Figure 118 : Inhibition de PabAGOG par C-02G10, ICOA2-G5-P7 et la 8-oxodG
Figure 119 : Changement conformationnel suite à la complexation de PaeAGOG avec la 8-oxodG
Figure 120 : Détermination de la valeur du K_D pour le composé C-02G10 et 8-oxodG

Figure 121 : Variation de la température de fusion (T_m) de hOGG1 (ΔT_m) en présence de différents inhibiteurs issus de notre criblage ou d'analogues
Figure 122 : Tracé des courbes de Michaelis-Menten et Lineweaver-Burk pour dhOGG1 en absence d'inhibiteur et détermination du V_{max} et du K_M selon les deux techniques
Figure 123 : Courbes de cinétique obtenues pour différentes concentrations d'inhibiteur ICOA2-G5-P7
Figure 124 : Optimisation manuelle des conditions de cristallisation de PabAGOG seule et photos de cristaux obtenus de PabAGOG en présence d'ICOA2-G5-P7
Figure 125 : Diffraction et statistiques d'affinement de PabAGOG lors du trempage d'ICOA2-G5-P7
Figure 126 : Structure et séquence de PabAGOG et PaeAGOG
Figure 127 : Hétérogénéité structurale de l'inhibiteur ICOA2-G5-P7
Figure 128 : Superposition des structures cristallographiques de PabAGOG/ICOA2-G5-P7 et de hOGG1/8-oxodG ou hOGG1/TH5487

Introduction générale

I. Dommages et mécanismes de réparation de l'ADN

I.1. Source et formation des lésions de l'ADN

I.1.1. Introduction biologique

L'ADN, ou acide désoxyribonucléique, est une macromolécule qui est le support de l'information génétique dans chaque cellule vivante. Elle est responsable de l'ensemble des fonctions internes d'un organisme, en codant pour les protéines et ARNs (acide ribonucléique). De fait, sa stabilité est **essentielle pour la pérennité des espèces**. Cependant, l'ADN n'est pas un composé inerte et sa structure chimique peut être modifiée par de **nombreux agents chimiques ou physiques d'origine endogène ou environnementale** pouvant conduire à des changements de sa structure native. Si ces dommages persistent, ils peuvent conduire à la mort cellulaire ou à des mutations. Ces mutations peuvent ensuite contribuer à l'évolution des espèces mais aussi être à l'origine de l'apparition de maladies ou de processus de létalité.

L'ADN est l'assemblage de deux polymères (« simples brins »), généralement structuré en double-hélice décalée droite (« double brin ») de type B, sa forme hélicoïdale la plus courante. Nous pouvons noter qu'il existe d'autres formes moins courantes de l'ADN comme l'ADN en double hélice A ou Z, ou encore des triples hélices et des quadruplexes, mais ces structures ne seront pas détaillées dans cette thèse. Chaque simple brin d'ADN est un polymère formé par une séquence de quatre nucléotides successifs et répétés de manière non aléatoire, donnant ainsi une **séquence nucléotidique**. Chaque nucléotide est composé de trois structures liées covalemment entre elles (Figure 1) :

- Une des quatre **bases nucléiques**, deux bases puriques : Guanine (G), Adénine (A) et deux bases pyrimidiques : Cytosine (C) ou Thymine (T) (Figure 2) ;

- un sucre, le **désoxyribose**, en vert sur la figure (de son nom complet 2'-désoxy-β-D-furanose) ;

- et un **groupement phosphate**, chargé négativement, en rouge sur la figure.

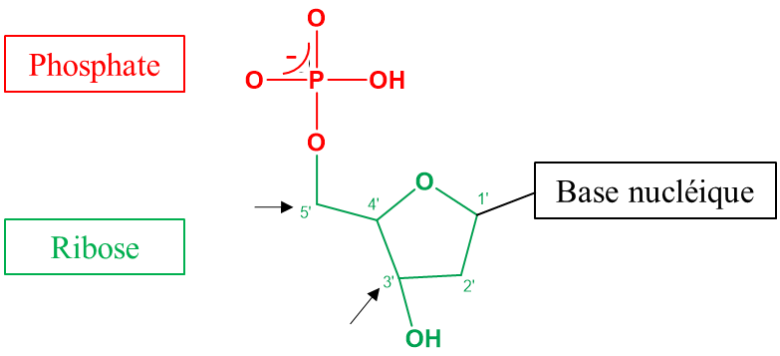


Figure 1 : Représentation schématique de la composition chimique d'un nucléotide

Dans l'ADN, ces nucléotides sont liés entre eux avec la liaison covalente entre le groupement phosphate d'un nucléotide et le désoxyribose d'un autre, formant une liaison dite **3'-5' phosphodiester** (indiqué par les flèches noires sur la Figure 1). La numérotation 3' et 5' découle de la numérotation du ribose, le « ' » indiquant que le désoxyribose est dans l'ADN.

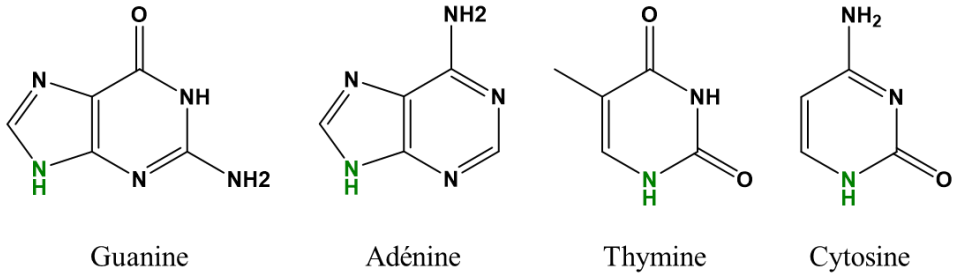


Figure 2 : Bases nucléiques de l'ADN - En vert est indiqué l'atome d'azote avec lequel s'établit la liaison avec le désoxyribose dite « *N*-glycosidique ».

Comme nous l'avons déjà évoqué, l'ADN est une molécule bicaténaire hélicoïdale résultant de l'assemblage antiparallèle 5'→3' et 3'→5' de 2 chaînes de polynucléotides (de deux simples brins). Chaque nucléotide se retrouve en face d'un autre nucléotide dit « complémentaire » avec lequel il interagit via des liaisons hydrogène (liaisons H) qui s'établissent entre les bases nucléiques de chaque brin. L'adénine est complémentaire de la thymine, tandis que la guanine est complémentaire de la cytosine. Ces bases peuvent cependant adopter des orientations différentes, en *anti* ou en *syn*, selon leur rotation autour de la liaison *N*-glycosidique, ce qui va changer la nature des appariements. Dans l'ADN de type B, l'orientation préférentielle des bases azotées est l'orientation *anti/anti*, et donne les appariements de type **Watson et Crick**. Cependant, il est possible d'avoir deux bases appariées aux orientations différentes comme *syn/anti*, ce qui donne des appariements de type **Hoogsteen** caractéristiques de la double-hélice d'ADN de type Z (Figure 3).

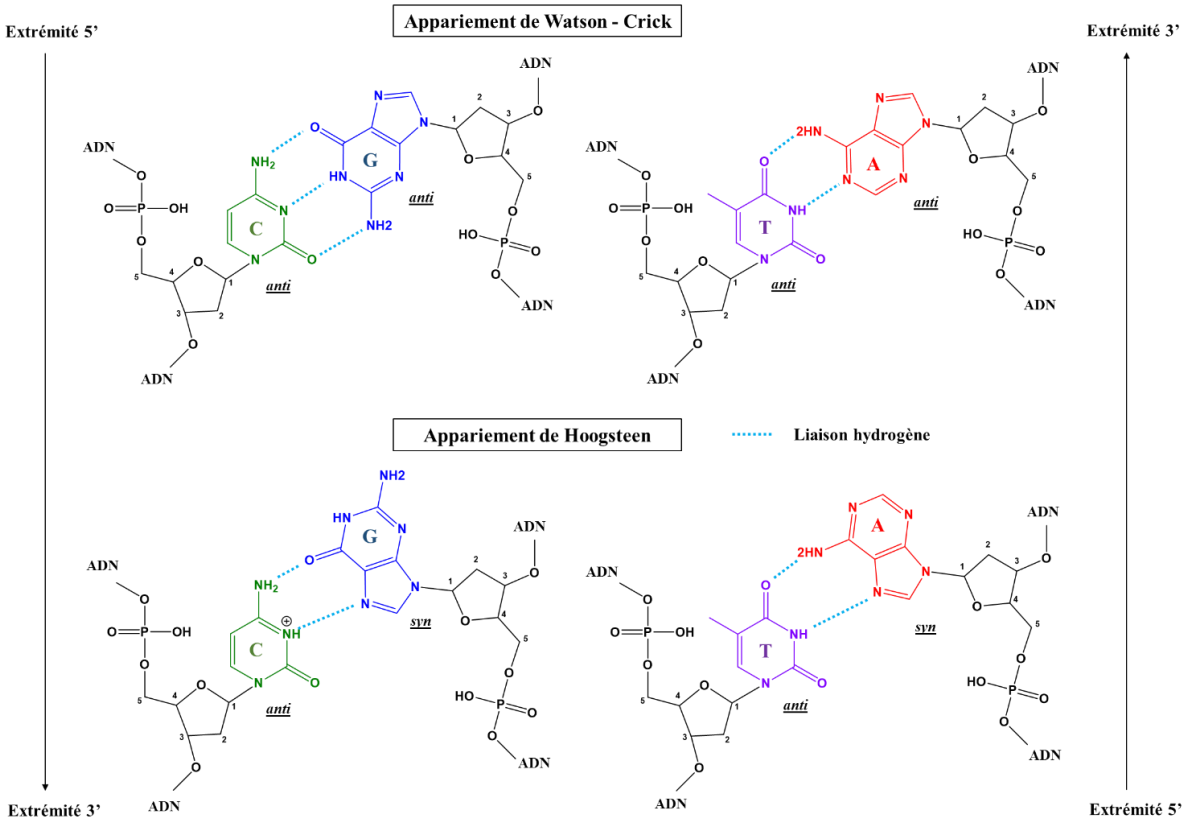


Figure 3 : Appariements canoniques Watson-Crick et non canoniques Hoogsteen - Les liaisons hydrogène sont indiquées en bleu clair et les bases nucléiques sont indiquées : la cytosine (en vert), la guanine (bleu foncé), la thymine (violet) et l'adénine (rouge).

Il est important de noter également que l'on peut retrouver de l'ADN dans deux emplacements cellulaires, le génome nucléaire et le génome mitochondrial. Le génome nucléaire représente l'ADN contenu

dans le noyau des cellules eukaryotes, sous forme de chromosomes, et est hérité pour moitié du père et pour moitié de la mère. Le génome mitochondriale quant à lui représente l'ADN contenue dans les mitochondries, qui se transmet principalement par la mère. La structure générale de l'ADN génomique est très différente de celle de l'ADN nucléaire, et est habituellement représentée sous une forme sphérique[1].

Comme nous l'avons déjà évoqué, l'ADN peut subir différents types de lésions pouvant endommager sa structure et conduire à la mort cellulaire ou à des mutations. C'est ce que nous allons décrire rapidement dans les paragraphes qui suivent.

I.1.2. Sources et types des dommages de l'ADN

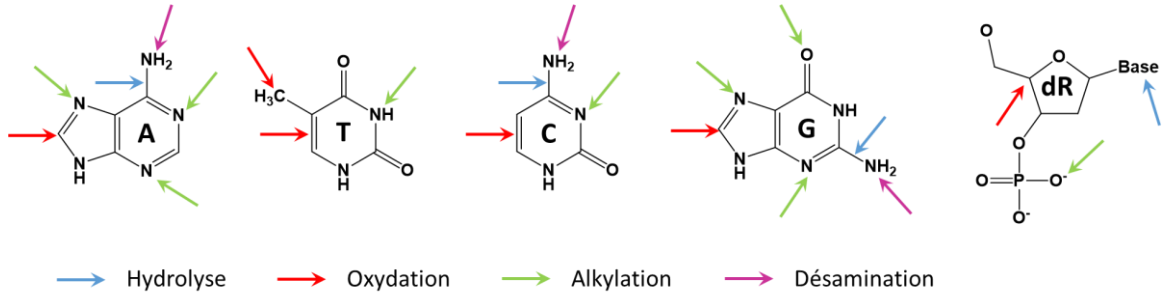


Figure 4 : Principaux sites de modification de l'ADN

Il existe de très nombreux agents (physiques ou chimiques) qui peuvent endommager notre ADN, que ce soit des processus extrinsèques ou intrinsèques à la cellule. Dans ce paragraphe, nous allons décrire plus en détails les différents dommages que peuvent subir les bases nucléiques ou le squelette phosphate-sucre. Il existe trois principales modifications de l'ADN, sur lesquels nous nous concentrerons, et plus particulièrement l'oxydation, qui concerne directement les substrats des protéines d'intérêt hNEIL1 et hOGG1 relatives à mon projet de thèse. En Figure 4 sont décrits les principales attaques possibles sur les bases ou le squelette phosphodiester de l'ADN.

- **Hydrolyse**

L'une des altérations les plus fréquentes est **l'hydrolyse spontanée ou induite du lien *N*-glycosidique** conduisant à la perte de base dans l'ADN (formation de sites abasiques, AP). S'il n'est pas éliminé de l'ADN, le site AP est très mutagène et capable de bloquer la réplication de l'ADN [2,3]. Les hydrolyses spontanées sont possibles dans les conditions physiologiques, mais largement favorisées en condition acide ou avec une élévation de la température et sont favorisées sur les bases puriques G et A [4]. Si les sites AP ne sont pas éliminés, ils peuvent également conduire à la formation de cassures simple brin par β-élimination sur la forme aldéhydique du désoxyribose constituant le site AP [5].

Des mécanismes d'hydrolyse peuvent également conduire à l'élimination de la fonction amine en C4 de la cytosine et de son homologue la 5-méthylcytosine pour former l'uracile (U) ou la thymine, respectivement. La désamination de la 5-méthylcytosine (5-meC) est environ deux à quatre fois plus rapide que celle de la cytosine [6]. Ces dommages peuvent donc mener à des mutations G:C vers A:T (en passant par des mésappariements transitoires G:T et G:U). La thymine dans le mésappariement G:T est réparée très lentement par les systèmes de réparation, et représente de ce fait un tiers des modifications de base qui cause des maladies héréditaires chez l'être humain [7]. Quoique plus rare, l'adénine et la guanine peuvent également subir une désamination hydrolytique, formant respectivement les lésions Hypoxanthine (Hx) et Xanthine (X). Hx est mutagène car elle peut s'apparier avec une cytosine (transition A:T vers G:C)

contrairement à X qui reste apparié à C. La présence du dioxyde d'azote NO_2 dans le cas d'une inflammation par exemple peut également favoriser la désamination des bases de l'ADN [8].

- **Alkylation**

Les **agents alkylants** réagissent pour la plupart avec les azotes cycliques et les oxygènes extracycliques des bases de l'ADN pour former une grande variété d'adduits. Les agents alkylants peuvent réagir sur 12 positions différentes des bases nucléiques [8]. Certains des agents induisant des dommages sur l'ADN sont illustrés en Figure 5. Les lésions générées par un agent alkylant dépendent de plusieurs paramètres : le nombre de sites réactifs de l'agent alkylant (mono ou bi fonctionnel, soit un ou deux sites pouvant alkyler), le type du groupe alkyle ajouté (méthyle, chloroéthyle, etc.), l'ADN cible (simple ou double brin) ou sa réactivité chimique (substitution nucléophile de type Sn1 ou Sn2) [9]. De manière générale, les agents de type Sn2 vont majoritairement agir sur les azotes des bases, tandis que les agents de type Sn1 peuvent modifier à la fois les oxygènes et les azotes. La grande majorité des agents utilisés en chimiothérapie sont de type Sn1.

L'azote N7 de la guanine étant celui disposant de la plus grande réactivité nucléophile, c'est la **cible de la majorité des agents alkylants monofonctionnels**, ce qui conduit à la formation de la 7-méthylguanine (7-meG) (qui compte pour environ 70 % du total des lésions d'alkylation de l'ADN). Cette lésion n'est ni mutagène ni toxique, cependant, elle est instable et susceptible à une dépurination spontanée, qui va créer un site AP, qui lui est mutagène et toxique ou peut former la FapyG pouvant bloquer la réplication de l'ADN [7,10]. La **N3-méthyladénine** (3-meA) représente environ 15 % des lésions d'alkylation et est extrêmement cytotoxique car elle capable de bloquer les ADN polymérases, et donc la réplication de l'ADN [11]. On peut aussi citer la 1-méthyladénine (1-meA) et la 3-méthylcytosine (3-meC), qui sont des composés bloquant également la réplication de l'ADN. Cependant, ces lésions sont beaucoup moins courantes car elles ne peuvent se former que sur un ADN simple brin, les positions N1 et N3 étant protégées dans l'ADN double brin grâce à l'appariement des bases [12].

Pour les oxygènes, c'est la **position O6 de la guanine** qui représente le site majoritaire de méthylation, formant l'O^6-méthylguanine (O^6-meG) ou l'O^6-chloroéthylguanine (O^6-Cl-eG). Bien que minoritaire face à l'alkylation des azotes, ces lésions restent cytotoxiques et mutagènes, pouvant respectivement mener à des erreurs d'appariements avec une thymine (transitions G vers A) ou à des pontages inter-brins de l'ADN qui bloquent la réplication [13,14]. Pour les pyrimidines, c'est la **position en O4** qui est également attaquable et forme les lésions O^4-méthylthymine et O^4-éthylthymine, qui sont des sources de mutation [15].

La principale différence entre les agents alkylants monofonctionnels (comme le témozolomide (TMZ)) et bifonctionnels est leur cytotoxicité. En effet, les agents bifonctionnels, comme les moutardes azotées (dérivés du gaz moutarde) ou les aziridines (mitomycine C, altrétamine, etc.), peuvent non seulement former les mêmes adduits d'alkylation, mais peuvent se lier à deux sites distincts, formant des pontages covalentes inter- ou intra-brins, ou même protéine-ADN, ce qui est particulièrement toxique pour la cellule [9,16]. Les sels de platine (dérivés du cisplatine, un des composés historiquement les plus utilisés dans de nombreuses chimiothérapies anticancéreuses) aussi dits « pseudo agents alkylants » sont des agents nucléophiles dont le platine se lie à l'azote N7 des purines et forme des liens intra- et inter-brins entre 2 purines [17]. Enfin des **agents polluants** tels que les composés nitrosamines ou benzo[a]pyrène (provenant de la fumée de cigarettes) ou des ROS/carbonyles (cigarettes électroniques) peuvent également provoquer la modification de bases de l'ADN en formant des adduits et en augmentant la formation de métabolites méthylés [18,19].

Il existe également des agents alkylants endogènes tels que la S-adénosylméthionine (SAM) [12]. Des adduits sur les bases peuvent également se former suite à la **peroxydation lipidique** qui résulte de l'oxydation des lipides présents dans les membranes des cellules. Cette réaction forme des lipides hydroperoxydes qui peuvent ensuite former différents aldéhydes ou époxydes dont les principaux sont la 4-hydroxynonénal (HNE) et le malondialdéhyde (MDA), respectivement très toxique et mutagène [20]. Ces différents produits peuvent également réagir avec les bases de l'ADN et produire des adduits de type éthéno ou propano, pour la plupart également très mutagènes, comme l'éthénoadénine (εA) ou l'éthénocytosine (εC) [21].

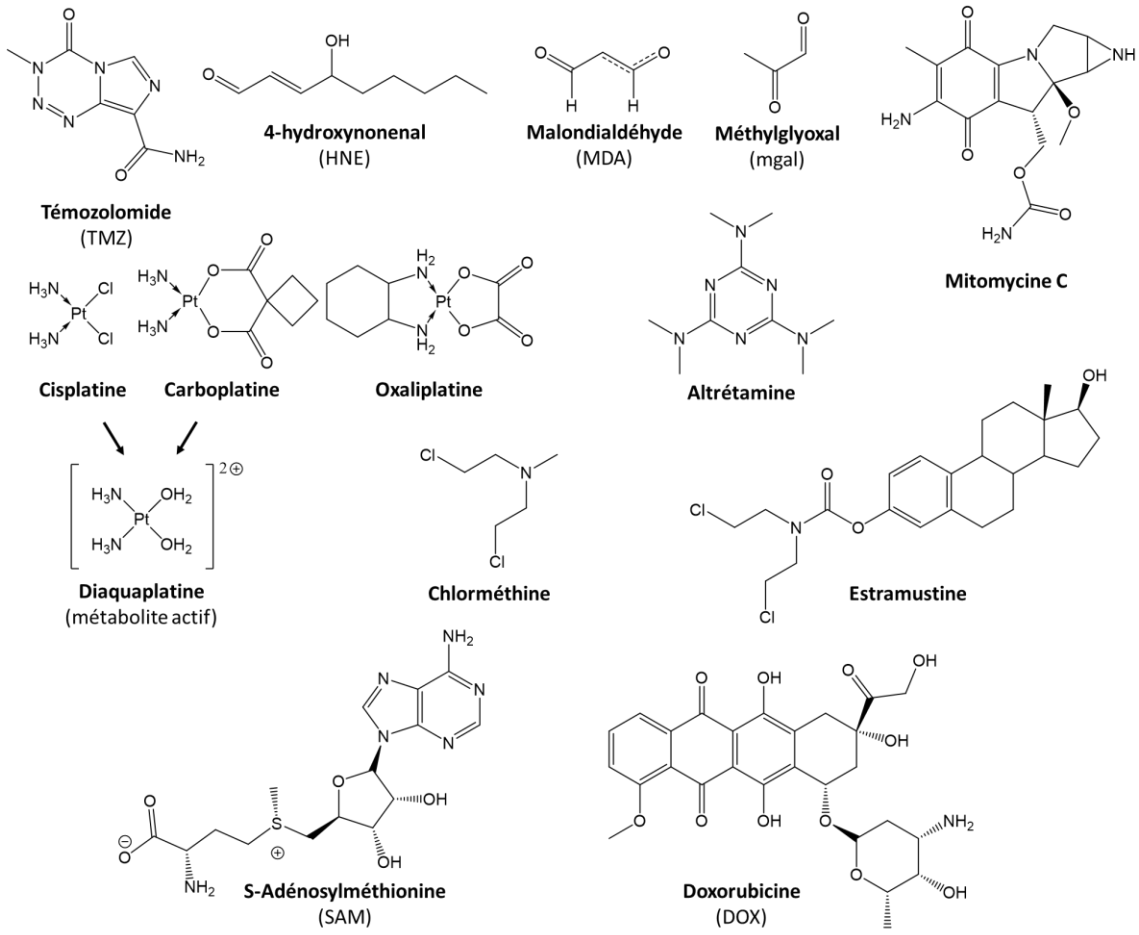


Figure 5 : Quelques exemples de composés provoquant des dommages sur l'ADN

- **Oxydation**

Si pour la plupart des organismes vivants, l'oxygène (ou sa forme dioxygène O_2) est un gaz inerte **indispensable** à leur survie car nécessaire à la production d'énergie (sous la forme d'ATP lors de la respiration mitochondriale) [22], sa présence dans les cellules peut cependant mener à la formation **d'espèces réactives de l'oxygène** (« Reactive Oxygen Species » ou **ROS** en anglais) [23]. Les ROSs sont parmi les principales sources de modification de l'ADN. Ainsi la présence de **radicaux superoxydes** $O_2^{-\bullet}$, de **peroxyde d'hydrogène** H_2O_2, de **radicaux hydroxyles** $OH^\bullet$ ou de **l'oxygène singulet** $^1O_2$ est toxique pour la cellule, contrairement à l'oxygène moléculaire [24]. Ces ROSs peuvent provenir des facteurs exogènes et endogènes divers. Par exemple la respiration mitochondriale, qui génère un anion radical $O_2^{-\bullet}$, transformé en peroxyde d'hydrogène puis en radical hydroxyle en suivant la réaction de Fenton [25] :

$$Fe^{2+} + H_2O_2 \rightarrow HO^\bullet + OH^- + Fe^{3+}$$

Une **réponse inflammatoire** peut également produire du peroxyde d'hydrogène et des anions radicaux superoxyde, qui peuvent réagir selon la réaction de Haber-Weiss [26] :

$$O_2^{-\bullet} + H_2O_2 \rightarrow O_2 + HO^{\bullet} + OH^-$$

D'autres facteurs peuvent générer des ROSs, comme le **monoxyde d'azote** induisant la formation de peroxynitreux et de radicaux hydroxyles $OH^{\bullet}$ ou les métabolites d'œstrogènes pouvant former des ions radicaux superoxydes [24,27]. Finalement, les **rayons ionisants et UV** peuvent également former des ROSs, principalement par radiolyse les molécules d'eau [28,29].

Ces ROSs, bien qu'attaquant l'ADN, sont paradoxalement essentielles pour une cellule, à travers leurs rôles dans la régulation cellulaire, dans l'apoptose et la prolifération des cellules musculaires lisses [30]. Cet aspect oblige donc les cellules à disposer d'un système efficace de régulation de l'oxygène et des ROSs afin de maintenir une homéostasie d'oxydoréduction (rédox) physiologique. Lorsque cette homéostasie est rompue, c'est-à-dire avec une augmentation des ROSs dans la cellule (ou une diminution des capacités anti-oxydantes), on parle de stress oxydant. Les dommages formés par l'oxydation des bases nucléiques peuvent mener à une grande variété de lésions, allant des cassures simple brin (SSB) ou double brin (DSB), de la formation de sites abasiques, la modification des bases ou encore à des pontages protéine-ADN [31]. Parmi toutes ces lésions, et de par son potentiel rédox le plus bas, la **guanine est la base nucléique la plus susceptible d'être oxydée**. Son oxydation mène à la formation de nombreuses lésions, dont les principales et celles qui reviendront le plus dans ce manuscrit, sont la **7,8-dihydro-8-oxo-guanine (8-oxoG)** et la **2,6-diamino-4-hydroxy-5-formamidopyrimidine (FapyG)** [32].

Dans sa conformation majoritaire *anti*, la 8-oxoG forme une paire de bases classique avec la cytosine. En revanche dans sa conformation *syn*, elle est extrêmement **mutagène** car elle peut imiter une thymine et former une paire de bases non canonique mais stable de type Hoogsteen avec une adénine (Figure 6) [33]. Si la 8-oxoG n'est pas éliminée de l'ADN, elle est associée à une haute fréquence de **transversion G:C vers T:A** pendant la réplication de l'ADN. Cette mutation est si commune qu'elle représente la deuxième signature mutationnelle dans les cellules de mélanomes humains [34,35]. Cette lésion est formée plus rapidement encore dans les cellules cancéreuses, indiquant que la 8-oxoG est un facteur important dans le développement des cancers [36]. Le FapyG peut lui aussi mener à des mutations G:C vers T:A, mais de manière moins fréquente, et peut également bloquer la réplication s'il est alkylé en N7 [10]. Réparer ces lésions est donc **particulièrement important pour la stabilité génétique** d'un organisme vivant.

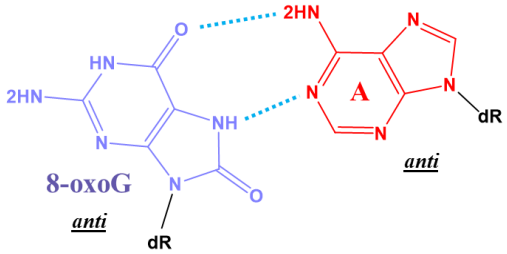


Figure 6 : Propriétés d'appariements de type Hoogsteen de la 8-oxoG avec une adénine - Les liaisons hydrogène sont indiquées en bleu clair.

Les autres bases nucléiques peuvent également être oxydées, comme l'adénine qui forme majoritairement la 4,6-diamino-5-formamidopyrimidine (FapyA) et la 7,8-dihydro-8-oxo-adénine (8-oxoA), mais qui sont peu mutagènes [37]. La thymine ou la cytosine lorsqu'elles sont oxydées peuvent de la même manière former les lésions thymine glycol (Tg), dihydrouracile (DHU), 5-hydroxyuracile (5-OHU), 5-hydroxycytosine (5-OHC), 5-hydroxyméthyluracile (5-hmU) et 5-formyluracile (5-fU) [38] (Figure 7). De plus, la

8-oxoG peut encore subir une oxydation et former la spiroimidohydantoine (Sp) et la guanidinohydantoine (Gh) [39].

Il est à noter que certains composés anticancéreux utilisent également l'intercalation des molécules dans l'ADN pour créer des dommages comme la doxorubicine qui peut former un adduit suite à son intercalation [40] ou les composés bléomycines qui s'intercalent dans l'ADN et oxydent les bases alentours [41].

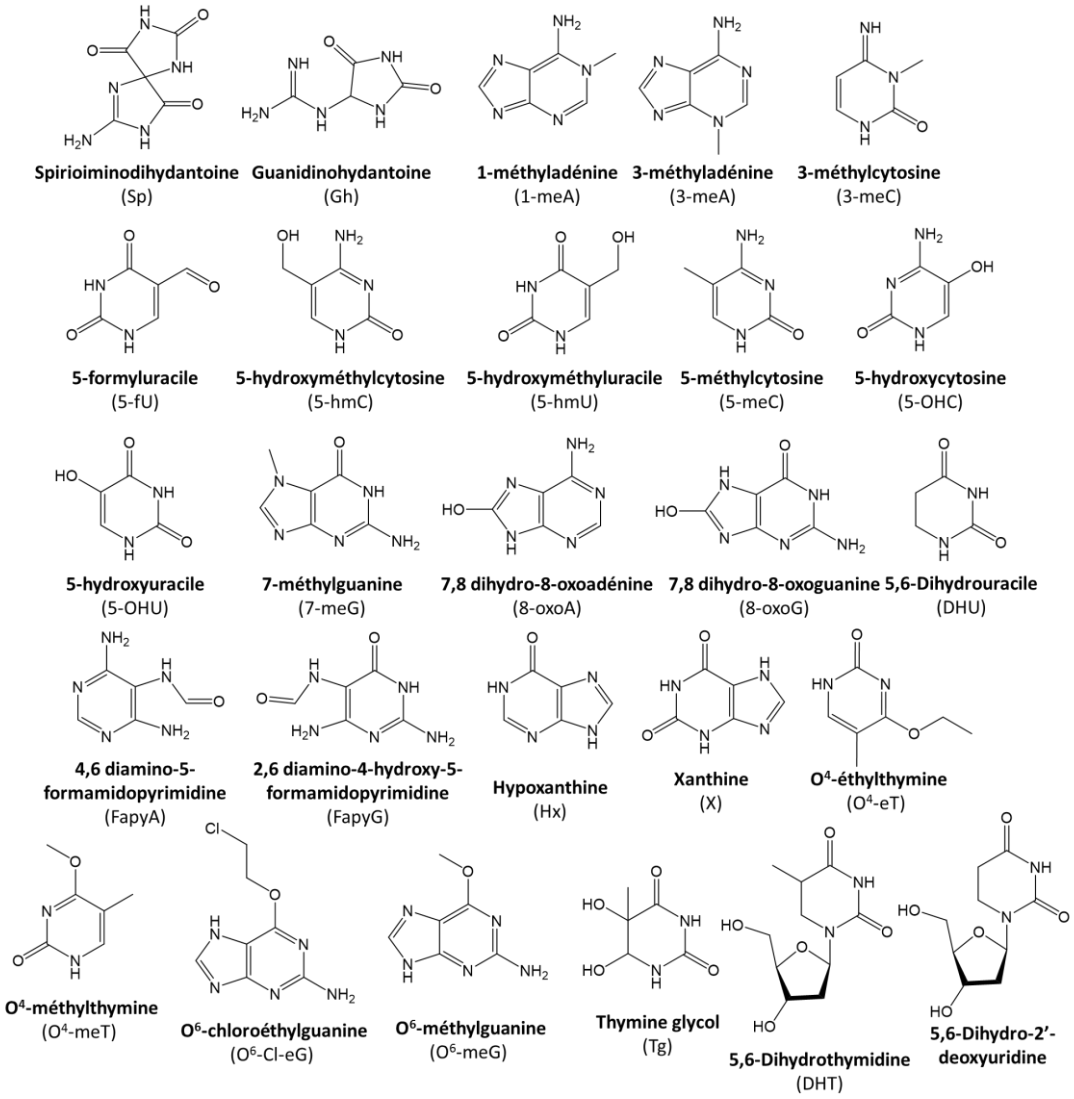


Figure 7 : Lésions principales des bases de l'ADN

I.2. Mécanismes de réparation de l'ADN

Comme nous l'avons vu, de très nombreux agents physiques ou chimiques peuvent induire la formation de lésions dans l'ADN qui peuvent à leur tour avoir des conséquences néfastes pour la cellule (mutations ou mort cellulaire). Afin d'éviter les effets délétères de ces dommages, les organismes ont développé des **systèmes de réparation de leur ADN endommagé**. Nous allons ici décrire brièvement les différents systèmes de réparation de l'ADN. Ma thèse étant principalement reliée à l'un d'entre eux, le **système BER**, un chapitre lui sera entièrement dédié.

I.2.1. La réponse aux dommages de l'ADN

Le premier système à être activé lors de l'apparition de dommages, pour les mammifères, est le **DDR** (« DNA damage response » en anglais). Il est initié par la reconnaissance spécifique des dommages à l'ADN par l'une des trois kinases apparentées à la phosphatidylinositol 3-kinase (**PIKKS**) : la kinase **ATM** (pour Ataxia-telangiectasia), la kinase **ATR** (pour ataxie télangiectasie et protéine associée à Rad3) ou la kinase **DNA-PKcs** (pour protéine kinase dépendante de l'ADN) [42]. La réponse cellulaire dépendra du dommage de l'ADN et de la kinase activée. L'activation de ces kinases déclenche une réaction en chaîne qui inclut le recrutement de la plupart des protéines responsables de la réparation de l'ADN au site lésé. Les cassures double brin activent les trois kinases à la fois [43]. De nombreuses autres enzymes comme le complexe MRN, la protéine RPA, le complexe 9-1-1 ou l'hétérodimère Ku sont impliquées dans cette première reconnaissance de dommages, et la plupart d'entre elles ont déjà été la cible de composés thérapeutiques dans la lutte contre le cancer. Suite à cette réponse initiale, il existe deux grandes stratégies de réparation de l'ADN, la **réparation par réversion directe** ou **par excision/resynthèse**, que nous allons voir en détails.

I.2.2. Réparation par réversion directe

C'est la réparation la plus simple et directe possible. Elle consiste en une **seule étape enzymatique qui permet de restituer l'information initiale sur l'ADN**, en catalysant la réaction chimique inverse à celle qui a conduit à la lésion et ne nécessite aucun remaniement structural important de l'ADN. Les enzymes principalement impliquées dans ce type de réparation sont les méthylguanine-ADN-méthyltransférases, ou MGMT, qui peuvent retirer le groupe alkyle d'une O^6-meG (par son transfert sur une de ces propres cystéine, ce qui en fait une enzyme suicide), et même le groupe alkyle en position 4 s'il est présent [44]. D'autres enzymes eucaryotes comme la ABH2 et la ABH3, ou leur homologue bactérien AlkB peuvent également directement réparer par déméthylation les lésions 1-meA et 3-meC [45]. D'autres enzymes peuvent également réparer directement les lésions en retirant le dommage, comme la photolyase, qui peut réparer les dommages induits par les rayonnements UV tels que les dimères cyclobutilique et photoproduits 6,4, mais cette enzyme n'a pas été identifiée chez l'Homme [46].

I.2.3. Réparation des cassures simple brin

Le système de **réparation des cassures simple brin (SSBR)** repose sur le recrutement à la cassure simple brin de la **protéine Poly(ADN-Ribose) Polymérase 1 (PARP1)** (PARP2 pouvant également jouer ce rôle à moindre mesure) [47]. Cette enzyme fait partie d'une famille de 17 protéines nucléaires isoformes, PARP1 étant l'enzyme prédominante et la plus étudiée. Elle est également impliquée dans d'autres mécanismes de réparation tels que les réparations des cassures double brin (HR et NHEJ) et l'excision/resynthèse (NER, MMR, BER) mais ne participe qu'à la signalisation des dommages à l'ADN, et non à la réparation de l'ADN en elle-même [48]. En cela son action se rapproche de celles des protéines kinases ATM/ATR et DNA-PKcs. Concrètement cette enzyme possède deux fonctions principales ; tout d'abord elle va reconnaitre et se lier aux cassures simple brin, puis va synthétiser des chaînes de poly (ADP-ribose) (pADPr). La formation de pADPr utilise la forme oxydée du nicotinamide adénine dinucléotide (NAD^+) et peut être synthétisée sur différentes protéines cibles ou sur PARP1 elle-même (Figure 8) [49]. Ces chaînes peuvent permettre de recruter différentes enzymes de réparation par des liaisons covalentes ou non covalentes [50], mais également d'interagir avec des histones afin de permettre la relaxation de la chromatine, et de faciliter l'accès des dommages aux protéines de réparation [51]. Enfin, son automodification peut lui permettre de se dissocier du dommage, laissant la place aux enzymes de réparation. Lorsque le dommage est réparé, les pADPrs sont éliminées par les enzymes glycohydrolases [52].

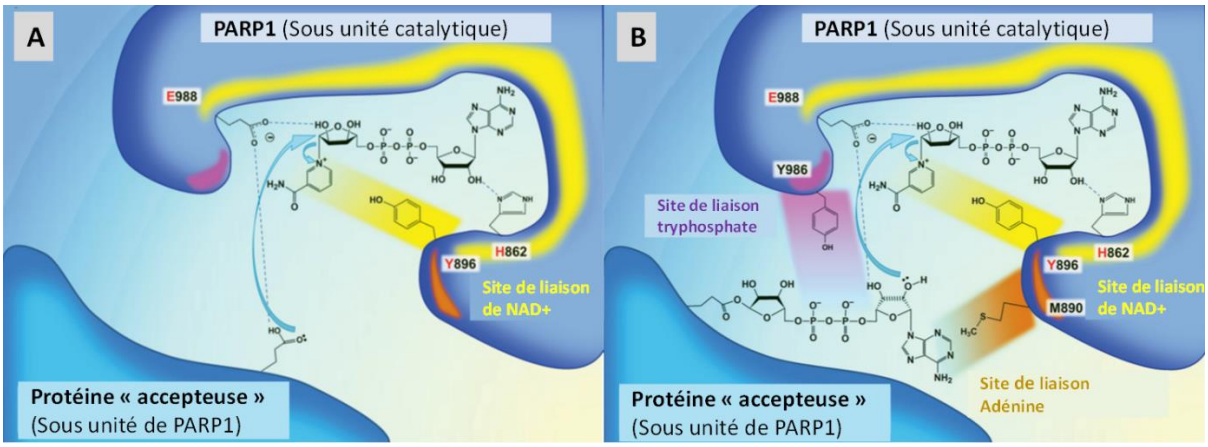


Figure 8 : Mécanisme de formation des chaînes pADPr par l'enzyme PARP1 (tiré et traduit de Alemasova *et al.*, 2019 [49])

Suite à la reconnaissance de la SSB par la PARP1, les principales protéines recrutées pour réparer la cassure simple brin sont l'ADN ligase III (Lig III), l'ADN polymérase β (Pol β) et la protéine « X-ray cross complementing gene 1 » (XRCC1), qui vont venir suturer le squelette phosphodiester. Cette réparation est très efficace, car peu encline aux erreurs de réplication [53].

I.2.4. Réparation des cassures double brin

- **Recombinaison homologue (HR)**

L'un des deux principaux systèmes de réparation des cassures double brin est la **réparation par recombinaison homologue (HR)**. Ce système, essentiel à la survie cellulaire, utilise la redondance de l'information génétique sous la forme de chromosomes homologues ou de chromatides sœurs afin de réparer les DSBs. De ce fait, il est uniquement possible en phase S et G2 [54]. Il est également utilisé pour réparer les fourches de réplication bloquées ou endommagées.

Ce système est tout d'abord initié par la formation du complexe MRN (composé de la protéine Mre11 pour « Meiotic Recombination 11 », de Rad50 et de la protéine NBS-1 pour « Nijmegen breakage Syndrome-1 ») sur le double brin, de chaque côté de la lésion, ainsi que par l'activation d'ATM. Par la suite, le complexe MRN va recruter l'enzyme humaine CtIP (Sae2 chez les levures), afin d'effectuer successivement une activité endonucléase et exonucléase sur l'un des deux brins afin de former une cassure simple brin se terminant par un hydroxyle en 3'. D'autres nucléases telles que l'exonucléase 1 (Exo1) et l'« ADN replication helicase 2 » (Dna2) vont également participer au processus *via* une activité exonucléase de 5' vers 3'. La structure du simple brin résultant est ensuite protégée par son recouvrement par la protéine RPA, qui va ensuite être remplacée par la protéine Rad51 grâce l'enzyme BRCA2 (pour « Breast cancer 2 »). Il en résulte la formation d'un filament de nucléoprotéines RPA/ADN simple brin [55]. Ce filament va ensuite pouvoir s'insérer dans la double hélice d'une chromatides sœurs, ou d'un chromosome homologue et forme une boucle « D-loop ». Suite à la formation de cette boucle, les ADN polymérases vont donc synthétiser le brin complémentaire en respectant les règles d'appariement Watson-Crick. Cela est possible si les séquences sont suffisamment proches pour former des doubles hélices stables [56]. En Figure 9 est indiqué un schéma résumant les différentes étapes de ce mécanisme. Il est à noter qu'en réalité la phase finale peut être différenciée en trois sous mécanismes différents selon les interactions des deux brins d'ADN ou la formation de la « D-Loop » : la formation d'une double jonction de Holliday, la SDSA (« Synthesis-Dependent Strand Annealing ») ou le BIR (« Break Induced Replication ») que nous ne détaillerons pas ici [57].

- **Réparation par jonction d'extrémités non homologues (NHEJ)**

Le second système de réparation des cassures double brin est le **NHEJ** (pour « non-homologous end-joining » en anglais), qui est promu majoritairement par la kinase ATM en phase G1 de la division cellulaire et de manière exclusive par la kinase DNA-PKcs couplée aux protéines Ku [54,58]. Les protéines Ku, formant un complexe Ku en anneau et glissant le long de l'ADN, vont recruter DNA-PKcs sur le site du dommage. La protéine Artemis va par la suite venir se lier à DNA-PKcs, puis après une phosphorylation, va cliver s'ils existent les brins d'ADN « dépassant » de la DSB. Une fois la cassure double brin dénuée d'extrémités sortantes formée, un complexe composé de la ligase IV, de XRCC4 et de XLF va venir refermer la cassure en liant les deux extrémités de la coupure. D'autres enzymes peuvent être impliquées dans le NHEJ, comme l'enzyme PNKP (« polynucleotide phosphatase/kinase ») ou l'enzyme TDP1 (« Tyrosyl DNA phosphodiesterase 1 ») qui prennent en charge certaines extrémités de l'ADN non substrat d'Artemis [59]. Un schéma résumant les différentes étapes de ce second mécanisme et également indiqué en Figure 9.

En réalité, il existe également une forme non-canonique de ce système de réparation, dite **d'alt-NHEJ ou MMEJ** (pour « microhomology-mediated end-joining »). Ce système utilise tout de même de courtes régions homologues pour guider la réparation, souvent présentes sur les bords des cassures double brin en « décalé », et implique notamment l'ADN polymérase θ. Si ce mécanisme est relativement fiable, des erreurs peuvent apparaitre lorsque ces microhomologies ne sont pas parfaitement compatibles [60].

De manière générale, il est ainsi admis que le NHEJ dans son ensemble est moins fidèle que le système HR, car de l'information génétique peut potentiellement être perdue lors de la ligation des deux jonctions, suite à l'élimination des régions simple brin jouxtant la DSB. Cependant HR peut lui aussi être la source de mutations.

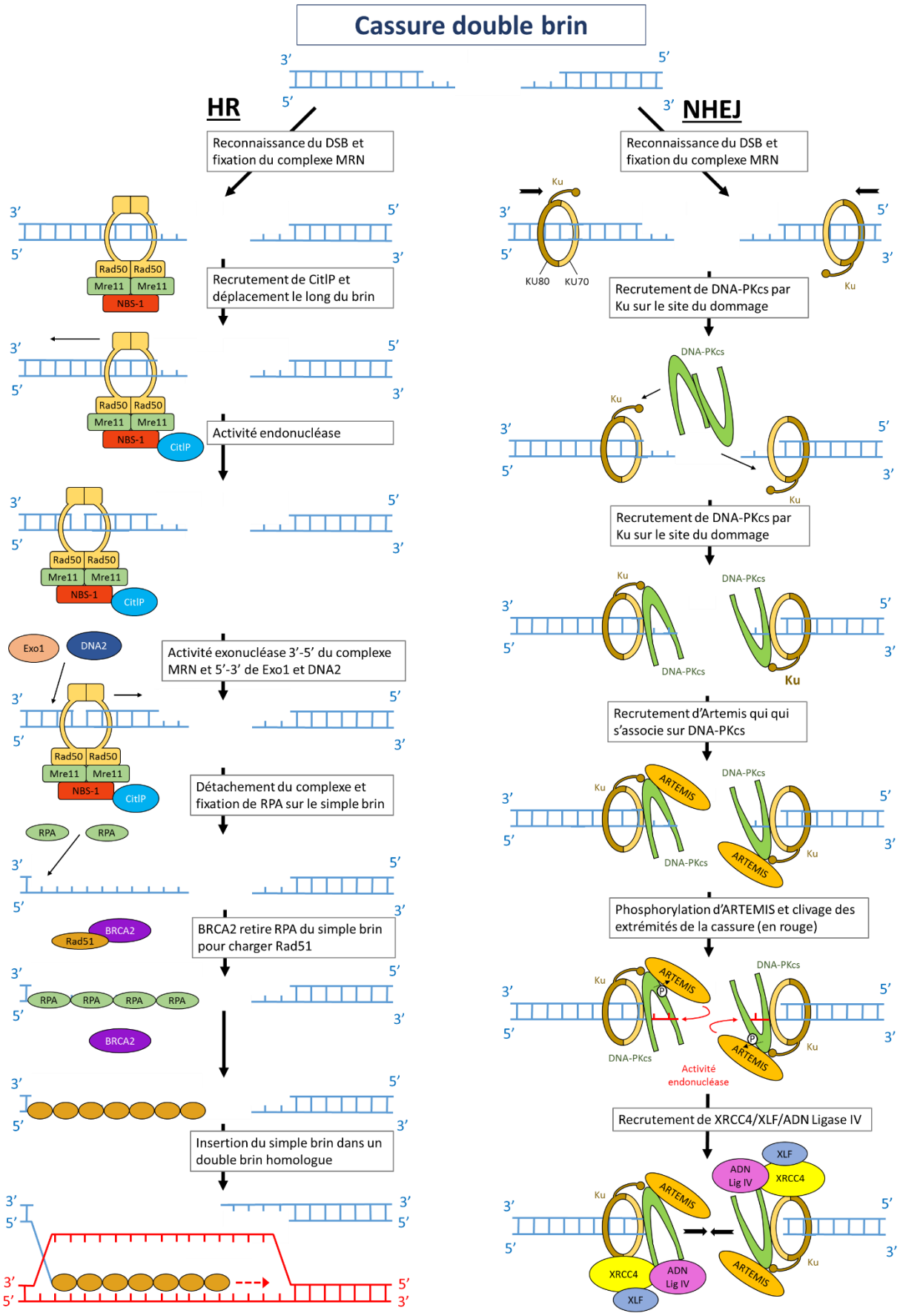


Figure 9 : Schéma du mécanisme général de HR et NHEJ

I.2.5. Réparation par excision et resynthèse

L'un des types de mécanisme les plus importants est le système de réparation par **excision et resynthèse** qui, comme son nom l'indique, consiste à retirer le nucléotide endommagé puis à compléter et réparer le brin d'ADN ainsi excisé par complémentation du brin non endommagé. Différents sous-systèmes existent.

- **Réparation des mésappariements (MMR)**

Le système de **réparation des mésappariements (MMR)** permet de réparer les bases qui ne respectent pas les appariements classiques de type Watson-Crick. Il est en particulier nécessaire à la réparation des erreurs de réplication. Le principe général du MMR consiste en la détection du mésappariement, l'élimination du brin d'ADN contenant la base mésappariée à l'aide d'une exonucléase, puis une resynthèse en utilisant le brin opposé comme matrice. Le MMR a été étudié en profondeur chez *Escherichia coli* (*E. coli*). Il est par ailleurs utilisé comme modèle pour comprendre le système équivalent chez les eucaryotes même si celui-ci diverge fortement, en particulier dans l'étape de reconnaissance du brin contenant la base mésappariée. Il est initié chez *E. coli* par les protéines MutS, MutL et MutH, puis différentes exonucléases (ExoI, ExoVII, ExoX et RecJ) viennent inciser le brin d'ADN contenant la base mal insérée lors de la réplication [61]. L'ADN polymérase III (impliquée aussi dans la réplication) et l'ADN ligase bactérienne viennent ensuite suturer la cassure créée comme dans la réparation des cassures simple brin (SSBR).

Chez l'Homme, des analogues de ces enzymes ont été identifiés : MSH (« MutS Homolog ») et hMLH (« MutL homolog »), ainsi que des analogues des exonucléases comme hExo1. D'autres enzymes entrent aussi en jeu comme PMS2 (« postmeiotic segregation increased 2 »), PCNA (« Proliferating cell nuclear agent ») ou RFC (« facteur de réplication C »). Nous parlerons également de ces enzymes dans d'autres systèmes de réparation. Concrètement chez l'Homme, le mécanisme est similaire au système bactérien MutHLS car seules certaines enzymes varient. Tout d'abord, la base mal appariée est reconnue par le complexe MSH2-MSH6 (ou MSH2-MSH3) qui va se lier à l'ADN au niveau du dommage. Il va ensuite recruter d'autres enzymes comme le complexe MLH1-PMS2, puis RFC et PCNA. La formation de ce complexe va activer l'endonucléase de PMS2, qui va donc créer une SSB au niveau de la base mal appariée. La création de cette SSB va permettre à hExo1 d'effectuer son activité exonucléase et créer une bréche monocaténaire très longue pour éliminer les mésappariements. L'ADN polymérase δ (ou l'ADN polymérase α) va ensuite rajouter les nucléotides manquant en utilisant le brin opposé comme matrice, puis l'ADN ligase I (Lig I) va finir de suturer le brin contenant initialement la base mésappariée (Figure 10) [62,63].

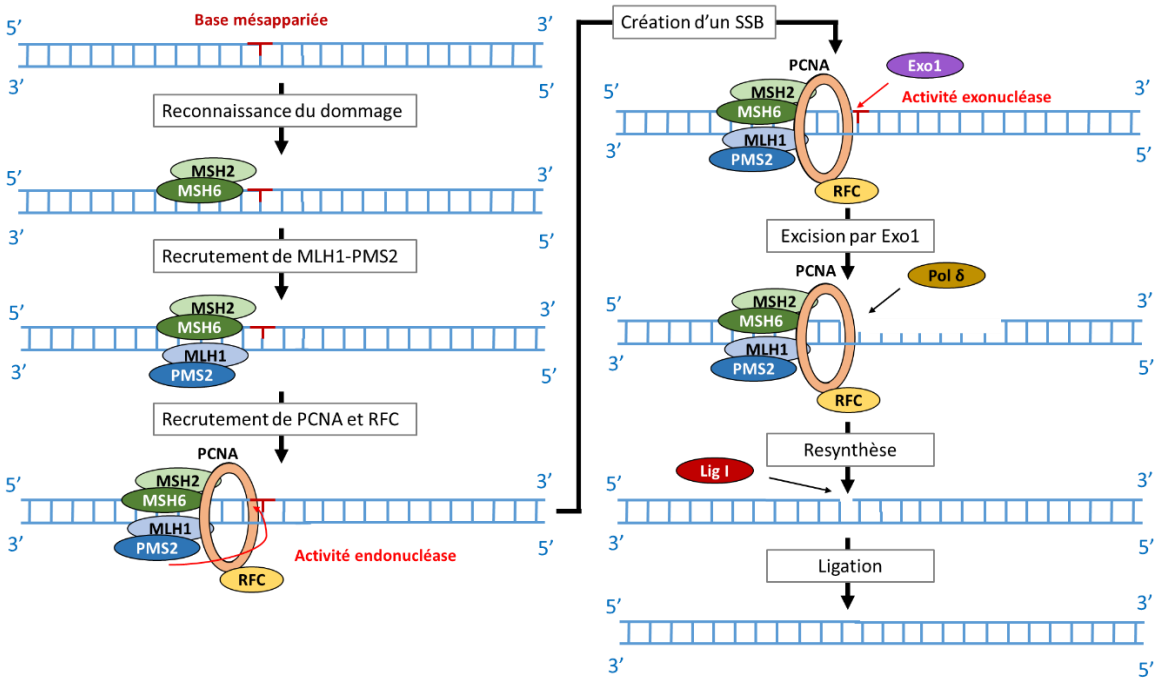


Figure 10 : Schéma du mécanisme du MMR

- **Réparation par excision de nucléotides (NER)**

Le système de **réparation par excision de nucléotides** (**NER** pour « nucleotide excision repair ») est le système principal chez les mammifères pour réparer les dommages et adduits « volumineux » de l'ADN associés à une forte déformation de l'ADN et souvent au blocage de la réplication. Il peut s'agir par exemples de dimères de pyrimidines induits par les UV, de pontages inter- et intra-brin ou de pontages protéine-ADN induits par exemple par le cisplatine [64]. C'est un mécanisme très versatile et capable d'éliminer de nombreux dommages sans que cela n'implique une reconnaissance spécifique de la lésion. C'est l'effet de ces adduits sur la structure de l'ADN qui est déterminant pour que le NER soit mobilisé [65].

Comme les autres systèmes d'excision/resynthèse, le NER se décompose en quatre étapes principales : (i) reconnaissance de la déformation de la double-hélice de l'ADN due au dommage et du brin endommagé ; (ii) excision du dommage sous la forme d'un oligonucléotide contenant le dommage générant ainsi une lacune de plusieurs nucléotides dans l'ADN ; (iii) remplissage de la lacune par resynthèse en utilisant l'ADN complémentaire comme matrice ; puis (iv) ligation du brin d'ADN néosynthétisé. Ce système est initié par deux sous-mécanismes, le NER du génome global (« global genomic NER », **GG-NER**), qui peut être utilisé n'importe où dans le génome, et le NER couplé à la transcription (« transcription coupled NER », **TC-NER**) qui est responsable de la réparation du brin transcrit dans les gènes actifs (ou transcrits). La TC-NER n'élimine les lésions que dans le brin d'ADN transcrit et non dans le brin codant. Le premier est initié par la reconnaissance de la structure endommagée de l'ADN par la protéine XPC couplée avec RAD23B et la protéine CETN2 (« centrin 2 »). Par la suite, ce complexe va faire fondre localement l'ADN (rupture des liaisons H de la paire de bases ciblée) provoquant le relâchement de sa structure, et attire le facteur de transcription TFIIH (« Transcription factor II Human »), qui est composés de 10 protéines différentes. Le second en revanche est initié par l'arrêt de l'activité de l'ARN polymérase II en rencontrant la lésion au cours de la réplication. Différents facteurs y sont utilisés pour retirer l'ADN polymérase du dommage et recruter TFIIH, que nous ne développerons pas ici [66]. A ce stade, les deux mécanismes convergent. Les protéines XPD et XPB, faisant parties du complexe TFIIH, vont créer une « bulle » d'environ 30 nucléotides dans l'ADN, puis différentes protéines comme XPA, RPA et XPD vont venir coordonner les étapes d'excision/resynthèse. Le complexe

TFIIH se couple ensuite avec XPF-ERCC1 (« excision repair cross-complementation group 1 ») et avec XPG. Ces protéines vont exciser un brin de plusieurs nucléotides contenant la lésion. Ensuite des ADN polymérases (δ, ε, κ) vont resynthétiser les nucléotides en utilisant le brin opposé comme matrice, et les ADN ligases I ou III vont venir suturer l'ADN (Figure 11) [64–66]. Chez des eucaryotes inférieurs, des protéines «Alkyltransferase-like » (catalytiquement non-actives) peuvent également, en se liant aux liaisons alkyles volumineuses, favoriser leur réparation par le système NER, mais cet effet n'a pas été observé chez l'Homme [67].

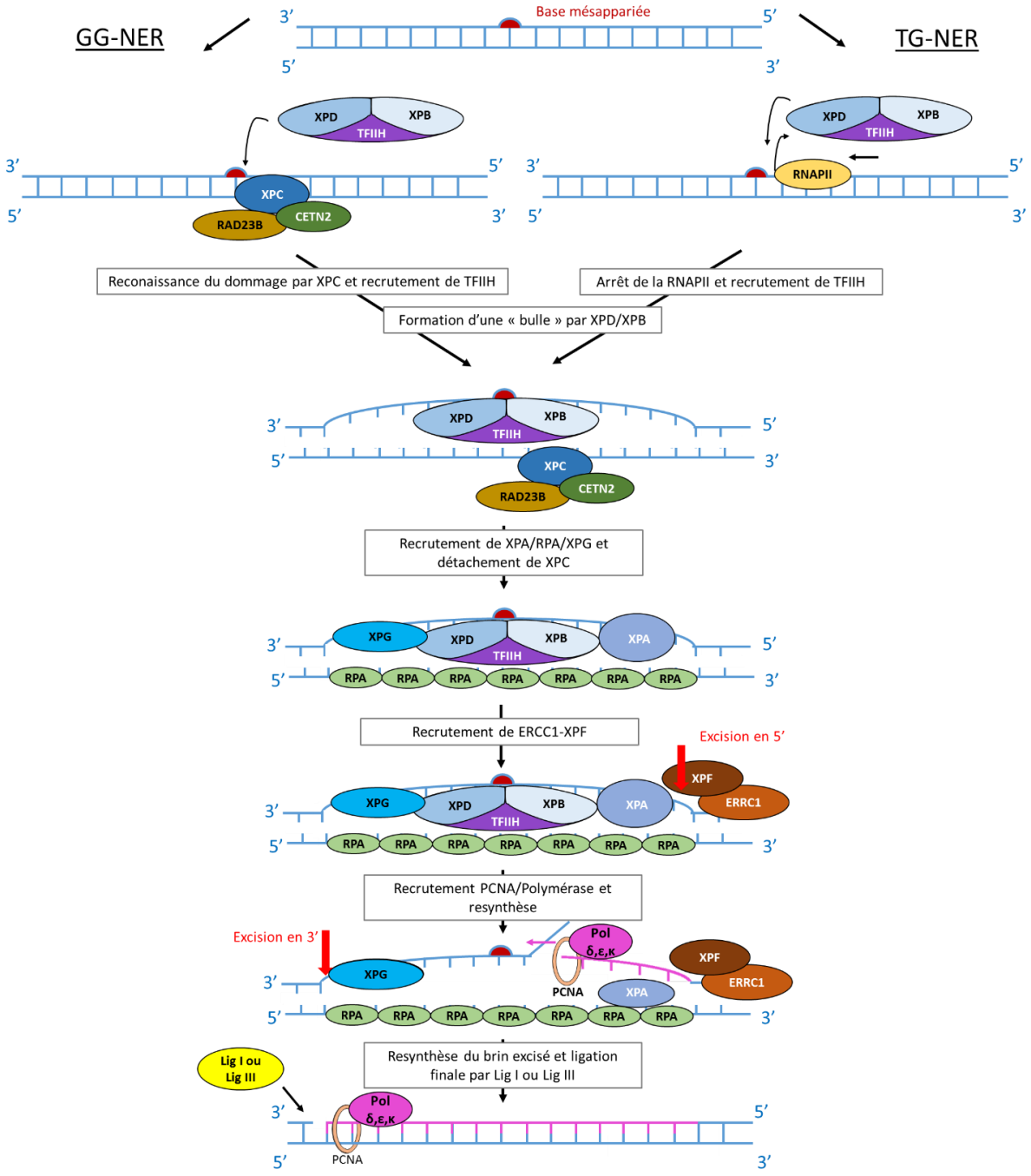


Figure 11 : Schéma général du NER

- **Réparation par incision de nucléotides (NIR)**

Il existe également un **système de réparation NIR** (« **Nucleotide Incision Repair** »), qui a été identifié chez *E. coli* et est initié par l'AP endonucléase Endo IV [68]. Cette enzyme, qui possède comme son nom l'indique une activité AP endonucléase, est capable d'inciser l'ADN en 5' de certaines bases oxydées comme

la 5,6-dihydro-2'-déoxyuridine, la 5,6-dihydrothymidine ou Tg, libérant une extrémité 3'-hydroxyle. Chez l'Homme, c'est APE1 qui possède la même activité [69]. Il est à noter que l'activité AP endonucléase de APE1 ne semble pas reliée à sa capacité à inciser certaines bases oxydées [70]. L'incision ainsi créée par APE1 résulte donc en un groupement hydroxyle en 3' substrat pour une ADN polymérase capable de faire un déplacement de brin, couplé à l'action de l'endonucléase FEN-1 pour éliminer le brin déplacé. Par la suite, une ADN ligase, notamment celles utilisées dans le BER qui sera présenté dans la prochaine partie, va finaliser la réparation.

II. Réparation par excision de base (BER)

La voie majeure de réparation des bases endommagées est le système de **réparation par excision de bases (ou BER)** découvert en 1974 par Tomas Lindahl, lauréat en 2015 du prix Nobel de chimie pour ses travaux sur de décryptage des mécanismes de réparation. Il est aussi quantitativement le plus important pour les cellules de mammifère et joue un rôle important dans les mécanismes de vieillissement, cancérogenèse et neurodégénérescence. On peut également considérer que les SSBs sont des dommages réparés par le BER. Le SSBR ne diffère du BER que par les enzymes qui initient la réparation car elles reconnaissent la base endommagée ou la DSB. Ce système, essentiel pour l'ensemble des organismes vivants, est très bien conservé depuis la bactérie jusqu'à l'Homme. Bien que les enzymes y participant soient différentes, toutes les espèces disposent donc d'une voie BER plus ou moins complexe, avec un mécanisme globalement similaire. Nous allons donc tout d'abord présenter le fonctionnement général de ce mécanisme de réparation, puis nous détaillerons chaque étape ainsi que le rôle des différentes protéines impliquées dans chacune d'entre elles.

II.1. Mécanisme général du BER

Comme dans les cas des systèmes MMR, NER et NIR, le BER peut être subdivisé en 2 phases : une phase d'excision du dommage et une phase de resynthèse réparatrice. Le BER est **initié par la reconnaissance spécifique de la base endommagée par l'une des 11 ADN glycosylases** identifiées chez l'Homme. On distingue deux types d'ADN glycosylases en fonction de leur mode d'action : les ADN glycosylases **monofonctionnelles** et les ADN glycosylases **bifonctionnelles associées à une activité AP lyase**.

Les ADN *N*-glycosylases monofonctionnelles catalysent **l'hydrolyse du lien *N*-glycosidique** entre la base endommagée et le carbone C1' du désoxyribose qui lui est associé (activité ADN glycosylase). La formation d'un **site abasique, ou site AP** (site apurinique/apyrimidique) résulte de l'action de ces enzymes. Le site AP est ensuite clivé au niveau du lien phosphodiester en 5' par une AP endonucléase (APE1 chez l'Homme) qui génère une cassure simple brin présentant une extrémité 3'-hydroxyle (3'-OH) et une extrémité 5'-phosphate-désoxyribose (5'-dRP). Afin de générer une cassure simple brin substrat pour la resynthèse par une ADN polymérase (*i.e.* présentant des extrémités 3'-OH et 5'-P), le dRP est ensuite éliminé par un clivage en 3' selon un mécanisme de β-élimination catalysé par l'ADN polymérase β (Pol β) [71].

Outre leur activité ADN glycosylase, les ADN glycosylases bifonctionnelles possèdent aussi une activité nucléasique dite « **AP lyase** » capable de cliver le lien phosphodiester en 3' du site AP par un mécanisme de **β-élimination** (similaire à l'activité AP lyase associée à Pol β). On parle parfois d'activité « β-lyase ». Il en résulte la formation d'une cassure simple brin présentant une extrémité 3'-phosphate liée à une forme aldéhydique α,β insaturée du désoxyribose et une extrémité 5'-phosphate (5'-P) (cassure simple brin non substrat pour une ADN polymérase, pas plus que pour une ADN ligase). Parfois certaines ADN glycosylases bifonctionnelles sont aussi associées à une activité AP lyase capable de cliver de façon concertée les liens phosphodiester en 3' puis en 5' du site AP par un mécanisme de **β,δ-élimination** résultant en la formation d'une lacune de 1 nucléoside délimitée par des extrémités 3'-phosphate et 5'-phosphate (non

substrat pour une ADN polymérase). On appelle aussi ces enzymes des « β,δ-lyases ». L'aldéhyde insaturé en 3' résultant d'une β-lyase sera éliminé par un clivage hydrolytique en 5' par **l'AP endonucléase majeure, APE1**, tandis que le groupement phosphate en 3' résultant de l'action d'une β,δ-lyase sera hydrolysé par la **polynucléotide kinase phosphatase** (PNKP). Il faut noter que APE1 possède aussi une faible activité 3'-phosphatase et peut donc aussi participer à la déphosphorylation de l'extrémité 3'-P grâce à son activité 3'-phosphodiestérase [72].

Au final la phase d'excision du dommage conduit à une SSB délimitée par des extrémités 3'-OH et 5'-P qui constitue un substrat pour une **ADN polymérase** (Figure 12). Dans la phase dite de « resynthèse » du BER qui suit, c'est **Pol β** qui est majoritairement mobilisée dans le BER (environ 80 %), mais d'autres ADN polymérases peuvent être alternativement recrutées pour la resynthèse comme nous le verrons plus loin. Pol β est une ADN polymérase non-processive spécialisée dans la resynthèse de courts fragments d'ADN à partir d'une extrémité 3'-OH libre (activité « gap-filling »). Finalement, une ADN ligase (Lig I ou Lig III) couplée avec XRCC1 restitue la continuité de la molécule d'ADN. Ce mécanisme est appelé le **BER à brèche courte** (**SP-BER**, « Short-Patch » BER) [73]. Cependant, lorsque Pol β n'arrive pas à exciser le groupement dRP pour diverses raisons (oxydation ou réduction du dRP par exemple), de nouvelles enzymes entrent en jeu, notamment les ADN polymérases réplicatives telles que **Pol δ et Pol ε**. Ces ADN polymérases sont capables d'ajouter à l'extrémité 3'-OH libre de la cassure entre 2 et 8 nucléotides en déplaçant le brin contenant le dRP modifié (aussi appelé le brin pendant ou « flap »). Au fur et à mesure du déplacement, **FEN-1** (« flap endonuclease 1 ») en combinaison avec **PCNA** (facteur de processivité de Pol δ), va exciser le brin pendant d'ADN généré [74,75]. Finalement, la continuité de l'ADN est rétablie par l'ADN ligase I. Cette alternative au SP-BER est appelée le **BER à brèche longue** (**LP-BER** pour « Long-Patch BER ») (Figure 12). Le LP-BER a principalement lieu dans les cellules en prolifération, car il utilise des enzymes également impliquées dans la réplication (Pol δ/ε, PCNA, FEN-1, Lig I), ce qui n'est pas le cas du SP-BER. En définitive, le choix de l'un ou l'autre de ces mécanismes de resynthèse est dépendant du type de dommage sur l'ADN, de l'ADN glycosylase initiatrice, mais également du type cellulaire [76].

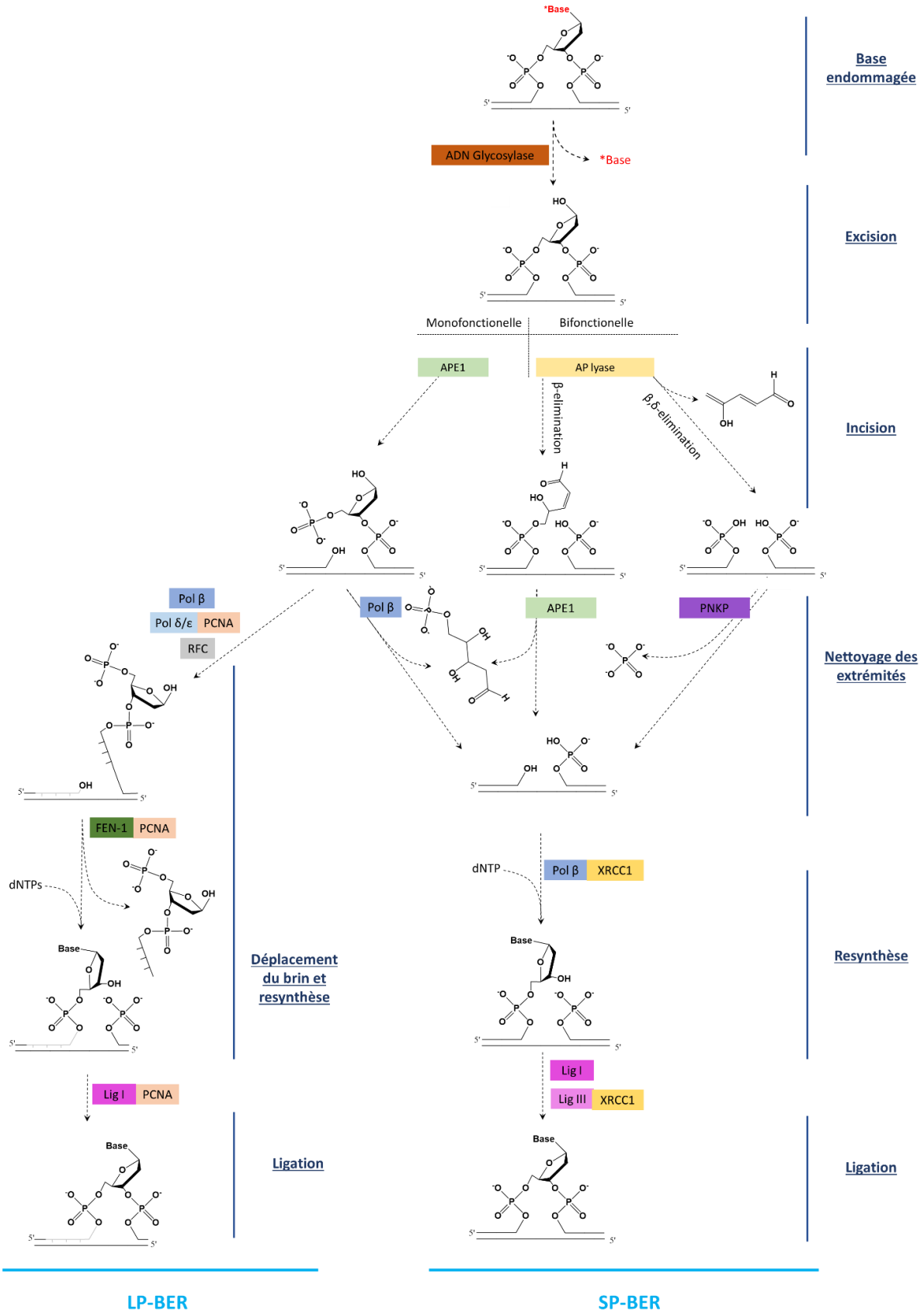


Figure 12 : **Mécanisme général du BER (adapté de Bjoras *et al.*, 2013 [71])**

II.2. Les ADN glycosylases

Il existe 11 ADN glycosylases chez l'Homme [74,77]. Ces enzymes sont généralement de **petite taille** et possèdent un **caractère basique marqué**. Suite à leur fixation à l'ADN, les ADN glycosylases **diffusent le long de ce dernier** à la recherche de dommages aux bases. Les ADN glycosylases localisent généralement le dommage dans un double brin d'ADN (OGG1, MUTYH, TDG, MBD4). Certaines d'entre elles sont capables cependant de reconnaitre la base endommagée dans un ADN simple brin uniquement ou dans l'ADN simple et double brin en même temps et même lorsque celle-ci est mésappariée (UNG, NEIL1,2,3 et SMUG1) [78]. Suite à la reconnaissance de la base endommagée dans l'ADN double brin, les ADN glycosylases induisent une **forte courbure de l'ADN, centrée sur la lésion,** ce qui conduit à son extrusion de la double hélice et à sa stabilisation en position extrahélicale dans la poche du site actif. En effet il a été prouvé que l'ouverture d'une paire de bases était facilitée en cas de courbure de l'ADN [79]. Les lésions présentes dans l'ADN ont également tendance à conférer à l'ADN une plus grande flexibilité, ce qui diminue la barrière énergétique et aide donc le processus d'extrusion. Ce mécanisme dit de **« base flipping » permet d'exposer idéalement le C1' du nucléotide endommagé à une attaque nucléophile dans le site actif de l'enzyme** qui conduira au clivage du lien *N*-glycosidique (activité ADN glycosylase) [80]. Ce mécanisme de **« base flipping »** est partagé par d'autres enzymes de réparation comme la 0^6-guanine alkyl-transférase par exemple [81]. Une fois la base extraite de la double hélice, les ADN glycosylases la maintiennent en position extra-hélicale grâce à **l'intercalation par le petit sillon de résidus de l'enzyme qui comblent le « trou » résultant de l'expulsion du nucléotide endommagé**.

Pour rappel, on peut séparer les ADN glycosylases en trois familles distinctes : les enzymes **monofonctionnelles** (UNG, SMUG1, MBD4, TDG, MPG et MUTYH), les **bifonctionnelles** associées à un mécanisme de **β-élimination** (OGG1, NTHL1 et NEIL3) et celles associées à un mécanisme de **β,δ-élimination** (NEIL1-2) [74,77]. Les enzymes monofonctionnelles clivent la liaison *N*-glycosidique en utilisant une molécule d'eau comme nucléophile, tandis que les bifonctionnelles utilisent un groupement amino activé de la protéine comme nucléophile impliquant la formation transitoire d'un intermédiaire covalent enzyme-ADN de type base de Schiff. Toutes les ADN glycosylases réparant l'uracile sont monofonctionnelles, tandis que les enzymes réparant les bases oxydées sont pour la plupart bifonctionnelles, sauf MUTYH.

Il est également possible de classifier les ADN glycosylases en fonction de leur structure tridimensionnelle. Ainsi, on distingue 6 superfamilles structurales d'ADN glycosylases [82] (Figures 13 et 14).

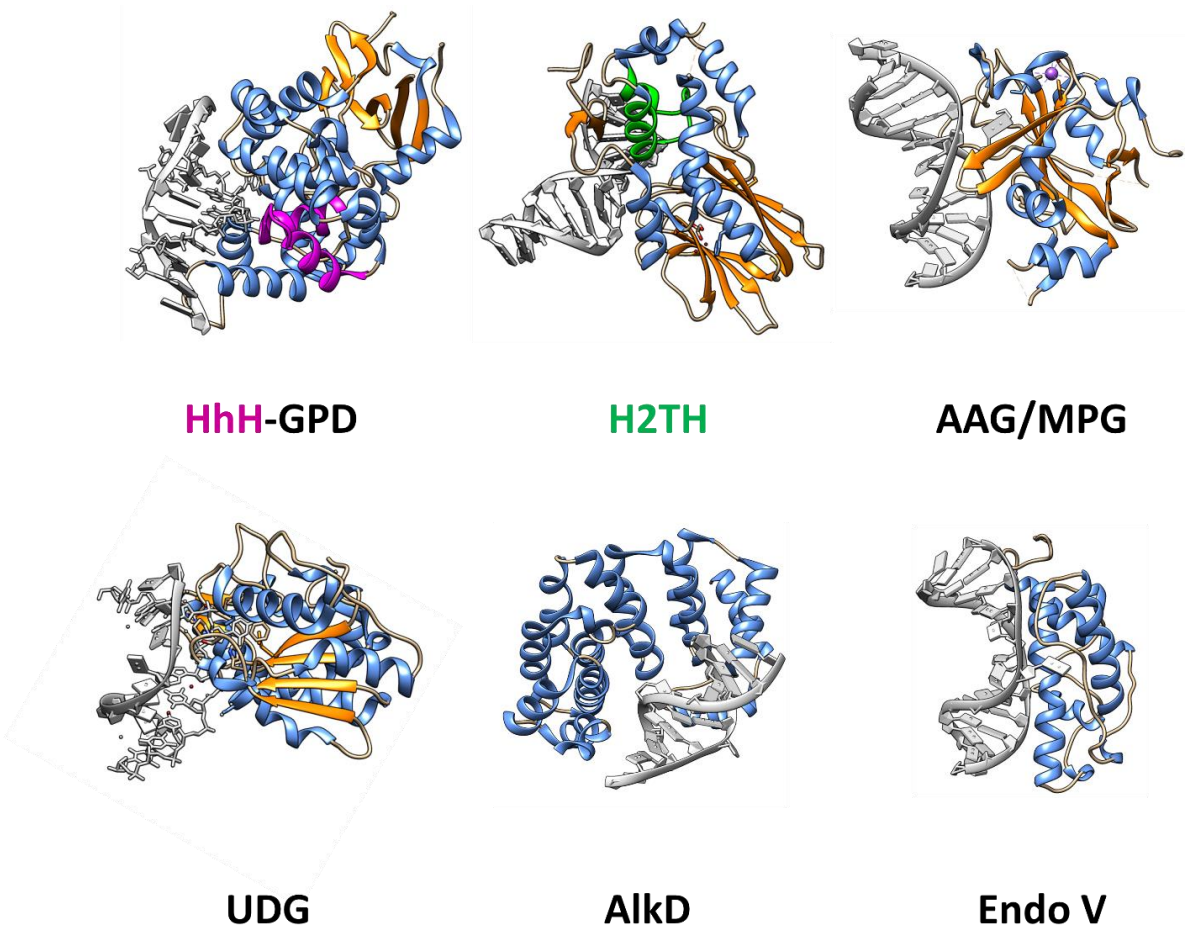


Figure 13 : Les différentes familles structurales des ADN glycosylases en interaction avec l'ADN - Les hélices α sont colorées en bleu, les brins β en orange, le domaine HhH en magenta et H2TH en vert. Les modèles sont : HhH-GPD, OGG1 humaine (1YQK) ; H2TH, NEIL1 humaine (5ITT) ; AAG/MPG, AAG humaine (1EWN) ; UDG, UDG humaine (1EMH) ; ALK, AlkD de *Bacillus cereus* (3JXZ) ; Endo V, T4 pyrimidine dimer Endo V (1VAS).

- La superfamille « **Helix-hairpin-Helix** » (**HhH-GPD**)

C'est la famille la plus diversifiée des ADN glycosylases. Elle est subdivisée en 6 sous familles distinctes d'après des études phylogénétiques et structurales (Ogg2 // AlkA // Mpg II // MutY/Mig // Ogg1 // Nth) [83]. Elle compte de nombreuses protéines telles que les enzymes procaryotes AGOG, EndoIII (ou Nth), AlkA et chez l'Homme les enzymes hOGG1/2, NTHL1, MUTYH et MBD4. Comme son nom l'indique, les enzymes de cette superfamille présentent toutes un motif structural à **deux hélices α en épingle à cheveux** (**HhH**) qui contribue à la structure de la poche du site actif. Il s'agit aussi d'un domaine d'interaction avec l'ADN contribuant à la reconnaissance du squelette phosphate-sucre au voisinage du dommage. Elles possèdent également une **boucle riche en glycines** (Gly - G) et prolines (Pro - P) et contenant un aspartate (Asp - D) catalytique (GPD). Si ce domaine est bien conservé entre les espèces, le reste de la structure de ces enzymes est plus divergent et peut contenir différents motifs structuraux fixant des ions zinc (TAG), un cluster fer-soufre (MpgII, Nth) ou contenant des lysines carbamylées (MagIII) [80]. Ce sont ces différences qui expliquent leurs spécificités propres de substrat, avec leur structure globale expliquant un mode général d'excision similaire [84].

- La superfamille « **Helix-2 turns-Helix** » (hélice-2 tours-hélice, **H2TH** ou Fpg/Nei)

Les enzymes bactériennes Fpg et Nei et eucaryotes NEIL1-3 (NEIL pour « Nei-like ») appartiennent à cette famille. Elles sont généralement constituées de 2 domaines globulaires : **un domaine N-terminal riche**

en brins β et un domaine C-terminal riche en hélices α. Le site actif de ces enzymes se localise à l'interface entre les 2 domaines. Outre le motif **H2TH** (dans le domaine C-terminal) impliqué dans l'interaction de ces protéines avec l'ADN et jouant un rôle similaire au domaine HhH de la famille précédemment citée, elles présentent une séquence N-terminale structurée en 2 hélices α et relativement bien conservée contenant la proline catalytique (Pro1) positionnée dans le site actif. Le domaine C-terminal contient un feuillet β à 2 brins antiparallèles organisé dans la plupart d'entre elles en **doigt à Zinc** à 3 ou 4 cystéines et caractéristique de cette famille d'enzymes [80,85].

- La famille des **alkyladénine-ADN glycosylases** (MPG/AAG)

Le premier représentant des ADN glycosylases de cette famille et qui lui a donné son nom est la N-méthylpurine-ADN glycosylase humaine (MPG) ou 3-alkyladénine-ADN glycosylase (AAG). Ces enzymes initialement identifiées chez les eucaryotes et plus récemment chez les procaryotes ont été très bien caractérisées [86]. Elles sont constituées d'un **seul domaine globulaire très compact de type mixte α/β**, composé de brins β entouré d'hélices α, avec un motif en épingle de brins β qui peut s'insérer dans le petit sillon de l'ADN, ce qui le rend assez différent de son homologue chez *E. coli*, qui dispose du motif HhH [87].

- La famille des **uracile-ADN glycosylases** (UDG)

Cette superfamille a été nommée d'après l'enzyme d'*E. coli* UDG (Uracile-ADN glycosylase), découverte en 1974 par Thomas Lindahl [88]. Depuis, son homologue humain UNG a également été identifié [89]. Malgré une faible homologie de séquence entres elles, les enzymes de cette famille contiennent toutes un motif **en feuillet β à 4 brins parallèles, contenant le site actif, encadré d'hélices α**. On retrouve ce repliement α/β chez les enzymes humaines UNG1/2, SMUG1 et TDG. Cette famille est présente chez les archées, les virus et les eucaryotes et est subdivisée en six sous-catégories en fonction de leurs substrats. La famille I, qui regroupe les enzymes UNG de l'Homme ou de *E. coli*, sont spécifiques de l'uracile et de la 5-fU. La famille II, contenant l'enzyme MUG II d'*E. coli* ou l'enzyme humaine TDG (Thymine DNA-glycosylase), sont des enzymes spécifiques de certains mésappariements impliquant un U, un T ou des dérivés de la 5-meC. Les enzymes de la famille III, telles que SMUG1 ne sont retrouvées que chez les eucaryotes supérieurs et excisent la 5-hmU mais aussi l'uracile dans certains contextes. Les enzymes des trois dernières familles sont uniquement retrouvées chez les archées ou les bactéries thermophiles, et contiennent un domaine à quatre cystéines servant de ligand au cluster métallique fer-soufre (4Fe-4S) [90].

- La famille « **HEAT-like repea**t » (HLR)

Elle a été récemment identifiée, avec AlkC et AlkD chez *Bacillus cereus,* et est très spécifique de l'excision de 3-meA et 7-meG [91]. Cette enzyme est composée principalement de **paires de petites hélices α formant un solénoïde de « superhélice α »**. Si cette famille est présente chez les procaryotes, aucune enzyme n'a été identifiée chez des organismes pluricellulaires [92]. AlkD dispose aussi d'un mécanisme de capture de la lésion différent des autres ADN glycosylases en établissant des interactions avec la base opposée tandis que la base modifiée extrahélicale n'est pas spécifiquement contactée dans le site actif de l'enzyme [93]. Le motif HEAT est souvent utilisé par des enzymes formant des interactions protéine-protéine, notamment les DNA-PKcs, ATM ou ATR dont nous avons déjà parlé, mais seules les ADN glycosylases de cette superfamille structurale contiennent le motif HEAT-like directement impliqué dans l'interaction de l'enzyme avec l'ADN.

- Il existe également une sixième superfamille identifiée uniquement chez les bactériophages, notamment l'endonucléase V du bactériophage T4 (T4 Endo V), qui hydrolyse le lien *N*-glycosidique en 5' des

dimères cyclobutyliques de pyrimidine (CPDs). N'étant pas présente chez l'Homme, et ayant donc moins d'utilité thérapeutique, nous ne présenterons pas ici cette enzyme [71].

Type de lésion	Nom complet	Abréviation	Activité (M, B ou BD)	Substrats	Inactivation du gène chez les souris	ADN cible (sb/db)	Superfamille structurale
Uracile dans sbADN	Uracile-N-Glycosylase	UNG1/2	M	U, 5-Fu, U:A et U/G	Viable	sb et db	UDG (repliement αβ)
	Single-strand-specific monofunctional uracil DNA glycosylase 1	SMUG1	M	U pref U:G>U:A>ssU, 5-hmU, 5-Fu , εC	Légère augmentation de la mutation C → T	sb et db	UDG (repliement αβ)
Dérivés de pyrimidine mésappariés	Methyl-binding domain glycosylase 4	MBD4	M	T:G, U:G, 5-hmU, 5-FU, εC	Viable, augmentation élevée de C→T	db	HhH
	Thymine DNA glycosylase	TDG	M	U:G>T:G 5-FU, εC, 5-hmU, 5-fC, 5-caC	Embryonaire létale	db	UDG (repliement αβ)
Dommages oxydatifs sur les bases	8-oxoG DNA Glycosylase 1	OGG1	B	8-oxoG:C, FaPy:C	Viable, mutation G→ T élevée	db	HhH
	MutY homolog DNA glycosylase	MYH	M	A en opposition à 8oxoG, G ou C et 2-ha opposé à G	Viable, mutation G→ T élevée	db	HhH
Purines alkylées	Methylurine glycosylase	MPG	M	3-meA, 7meG, 3meG, hypoxanthine, εA	Viable, niveau élevé de éthénoA et hypoxanthine	sb et db	MPG/Aag
Fragments aromatiques oxydés ou pyrimidines saturées	Endonuclease III-like 1	NTHL1	B	Tg, FaPyG, 5-hc, 5-hu	viable	db	HhH
	Endonuclease VIII-like glycosylase 1	NEIL1	BD	Tg, FaPyG, FaPy A, 8-oxoG, 5-hU, 5-hC, DHU, Sp et Gh	viable	sb et db	H2TH
	Endonuclease VIII-like glycosylase 2	NEIL2	BD	Tg, FaPyG, FaPy A, 8-oxoG, 5-hU, 5-hC, DHU, Sp et Gh	viable		H2TH
	Endonuclease VIII-like glycosylase 3	NEIL3	BD	FaPyG, FaPyA, Sp, Gh, 8-oxoG	normal	sb (maj)	H2TH

Figure 14 : Différentes ADN glycosylases humaines - Les substrats principaux et secondaires sont indiqués respectivement en rouge et en vert. M = Monofonctionnel, B = Bifonctionnel à β-élimination, BD = Bifonctionel à β,δ-élimination. sb = simple brin, db = double brin.

II.3. Elimination du site abasique, nettoyage des extrémités et resynthèse

II.3.1. Via le BER à brèche courte (SP-BER)

Comme nous l'avons évoqué précédemment, suite à l'ablation de la base endommagée par les ADN glycosylases, le brin d'ADN va être incisé, afin d'éliminer le désoxyribose (site abasique, AP) initialement associé à la base endommagée. Si certaines ADN glycosylases possèdent une activité de clivage des sites AP (AP lyase), dans la plupart des cas ce sont des AP endonucléases qui seront mobilisées pour éliminer le site AP. Dans le cas des cellules de mammifères, c'est **l'apurinique/apyrimidique endonucléase 1** (**APE1**, ou HAP1 pour « human APE1 ») qui va majoritairement venir reconnaitre le site AP formé et inciser par hydrolyse le squelette phosphodiester en 5' de celui-ci. Cette enzyme a été très largement caractérisée, et ses nombreuses fonctions ont été étudiées. Dans le cadre de la réparation des dommages de l'ADN, grâce à l'attaque nucléophile d'une molécule d'eau en 5' du désoxyribose, la protéine va créer une cassure simple brin dans l'ADN double brin, résultant en un **groupement hydroxyle en 3' et un dRP en 5'**. APE1 est une nucléase particulièrement multifonctionnelle, possédant des activités endonucléase, exonucléase (3' → 5'), 3' phosphodiestérase et pouvant également cliver l'ARN [94]. De ce fait, elle peut également retirer le groupement dRP en 3' formé suite à l'action d'une ADN glycosylase/β-lyase [71]. Elle est présente dans le noyau ainsi que dans les mitochondries des cellules (sous une forme tronquée) et possède également des fonctions rédox régulant la transcription [74,95]. Si les deux facettes de sa fonction biologique sont indépendantes, APE1 reste une enzyme indispensable à la survie cellulaire, ce qui en fait donc aussi une cible thérapeutique intéressante.

Suite à l'action d'APE1, les extrémités de la SSB formées sont prises en charge par différentes **ADN polymérases**. Les ADN polymérases incorporent par complémentation en lieu et place du nucléotide endommagé le nucléotide canonique en respectant les règles d'appariement Watson-Crick. Elles sont impliquées dans la réparation de l'ADN mais aussi dans la réplication de l'ADN. On compte 15 ADN polymérases chez l'Homme, dont 7 sont impliquées à des degrés divers dans le BER (Figure 15) [96]. Dans la

réparation de l'ADN, leur rôle ne se limite pas à l'ajout du nucléotide, certaines d'entre-elles pouvant également inciser certaines extrémités afin d'insérer le nucléotide manquant. L'ADN polymérase majoritaire du SP-BER est l'ADN **polymérase β** (Pol β), qui a été identifiée en 1971 chez les mammifères et que l'on retrouve dans toutes les espèces vertébrées [97]. C'est une enzyme de petite taille (39 kDa) par rapport aux ADN polymérases réplicatives, non-processive et qui ne possède pas d'activité exonucléase, mais une activité dRP lyase [98]. C'est notamment Pol β qui est impliquée dans la réparation de la 8-oxoG [99]. Grâce à son activité lyase, elle **permet d'inciser le groupement dRP** résultant de l'activité d'APE1, formant un groupement phosphate en 5' sur le brin d'ADN lors du SP-BER. Pol β incorpore ensuite le nucléotide manquant et contribue au recrutement du complexe XRCC1/Lig III ou Lig I afin de restaurer la continuité du brin d'ADN initialement lésé (Figure 12).

D'autres ADN polymérases sont également impliquées dans le SP-BER, par exemple, Pol ι, qui possède une activité dRP lyase en 5' ainsi que la capacité d'intégrer le nucléotide manquant dans le brin d'ADN. Cette ADN polymérase est principalement nucléaire, mais son activité est environ 30 fois moins efficace que celle de Pol β [100]. Dans les mitochondries, c'est **l'ADN polymérase ϒ** qui est impliquée majoritairement dans l'étape de resynthèse du BER. Si on a longtemps pensé que Pol ϒ était la seule ADN polymérase présente dans les mitochondries permettant de remplir les fonctions essentielles du BER, une étude très récente semble montrer que la Pol β est cependant plus efficace dans ce contexte [101]. Pol λ, une autre ADN polymérase de la famille X et très proche de Pol β, possède également une activité AP lyase et 5'-dRP lyase et a été observée comme contribuant au BER chez les mammifères [102]. Enfin, Pol θ présente une activité dRP et AP lyase et a été caractérisée comme une enzyme pouvant soutenir Pol β, et participer aussi au LP-BER [103]. De manière générale, la plupart de ces enzymes ne peuvent remplacer complètement Pol β. En effet, des souris déficientes en Pol β ne sont pas viables [104]. Ceci souligne l'importance de Pol β dans les 2 voies : SP et LP-BER.

Une variante du SP-BER existe lorsque le BER est initié par une ADN glycosylase bifonctionnelle, comme hNEIL1 par exemple, dont l'action (ADN glycosylase et β,δ-lyase) résulte en l'excision du nucléoside endommagé, et la formation d'une lacune (ou « gap ») délimitée par une **extrémité 3'-P et une extrémité 5'-P**. Pour être substrat de Pol β, la lacune de 1 nucléoside doit être convertie en lacune de 1 nucléotide par déphosphorylation de l'extrémité 3'-P par la **polynucléotide kinase phosphatase** (PNKP) [75]. APE1 peut aussi assurer cette fonction, mais de façon bien moins efficace que la PNKP [105]. Comme la plupart des enzymes impliquées dans la réparation de l'ADN, la PNKP est impliquée dans d'autres systèmes que le BER, comme le NHEJ par exemple ou le SSBR.

Nom	Famille structurale	Mécanisme BER	Divers	Activité utilisé dans le BER
Pol β	X	SP et LP	Enzyme principale	dRP Lyase - AP Lyase - Polymérase
Pol ϒ	A	SP	Uniquement mitochondriale	dRP Lyase - AP Lyase - Polymérase
Pol δ	B	LP		Polymérase
Pol ε	B	LP		Polymérase
Pol λ	X	SP et LP		dRP Lyase - AP Lyase - Polymérase
Pol θ	A	SP et LP		dRP Lyase - AP Lyase - Polymérase
Pol ι	Y	SP		dRP Lyase - AP Lyase - Polymérase

Figure 15 : Polymérases impliquées dans le BER

II.3.2. Via le BER à brèche longue (LP-BER)

Lorsque Pol β ne peut pas prendre en charge le produit de l'activité des ADN glycosylases monofonctionnelles, suite par exemple à une modification chimique du groupement en 5' de la cassure formée suite à l'action d'APE1, d'autres ADN polymérases sont mobilisées au niveau de la cassure simple brin non canonique. C'est la voie dite du LP-BER. C'est principalement le cas des **ADN polymérases réplicatives Pol δ et Pol ε** [106]. Les ADN polymérases λ et θ semblent également pouvoir jouer un rôle dans le LP-BER [96]. Ces enzymes vont prendre la suite de l'insertion du premier nucléotide par Pol β et insérer une dizaine de nucléotides, ce qui résulte en un déplacement du brin contenant l'extrémité 5'-bloquée. Ce brin chevauchant, **dit pendant ou « Flap »**, est pris en charge par FEN-1, qui est une endonucléase impliquée entre autre dans la réplication mais également dans différents systèmes de réparation, notamment les cassures double brin et le BER [107]. FEN-1 permet de cliver les fragments pendants au fur et à mesure de son déplacement par l'ADN polymérase et est essentielle au LP-BER pour créer une cassure simple-brin substrat pour l'ADN ligase I. Il existe un couplage fonctionnel entre FEN-1 et Pol β permettant de stimuler le déplacement du brin (« strand displacement ») par Pol β, ce qui va également permettre la libération du substrat de FEN-1 plus rapide [108].

Si les ADN polymérases et FEN-1 sont les principales responsables de l'action catalytique de l'excision des brins de l'ADN, d'autres enzymes sont également impliquées dans le LP-BER et permettent de stimuler ce mécanisme. C'est notamment le cas de l'enzyme **PCNA**, une enzyme fortement impliquée dans la réplication et capable d'interagir avec FEN-1 et l'ADN ligase I [109]. PCNA fait partie de la famille des « clamps à ADN » dont le rôle essentiel est **d'augmenter fortement la processivité des ADN polymérases** réplicatives telles que Pol δ et Pol ε lors de la réplication mais aussi dans les systèmes de réparation LP-BER, MMR et NER [110]. Même si certaines études *in vitro* suggèrent que le LP-BER peut être opérationnel en absence de PCNA, il semble néanmoins que dans les cellules dépourvues de PCNA, seul le SP-BER est fonctionnel [111].

II.3.3. Co-facteurs du BER

Une autre enzyme intervient fortement dans le BER : la protéine **XRCC1** qui peut se lier avec différentes lésions de l'ADN, comme les SSBs et DSBs. Cette propriété lui permet d'être **une protéine d'échafaudage, ou « scaffold protein »**, en formant des complexes avec de nombreuses autres protéines, permettant de localiser et réguler les mécanismes de réparation par exemple [112]. Elle est de ce fait essentielle au bon maintien de la stabilité génétique. Elle est ainsi impliquée dans la SSBR et le BER, ces systèmes partageant de nombreuses enzymes, le BER formant en effet une cassure simple brin transitoire pendant son mécanisme. Elle interagit par exemple avec les PARP1/2 [113]. Selon la phase du cycle cellulaire de la cellule, XRCC1 peut aussi être impliquée dans la recombinaison ou la réplication de l'ADN. XRCC1 interagit avec de très nombreuses enzymes du BER, notamment certaines ADN glycosylases comme hOGG1, hNEIL2, MPG, TDG ou hNTH1 [114]. Bien que majoritairement impliquée dans le SP-BER, elle est aussi engagée dans le LP-BER. Ainsi, elle forme des complexes fonctionnels avec Pol β, PNKP, APE1 et Lig III lors de l'étape finale de resynthèse/ligation, tout en pouvant également interagir avec la polymérase δ et l'enzyme PCNA [76,112,115,116].

Enfin, d'autres enzymes sont également impliquées dans le BER, comme le facteur de réplication C (RFC) ou la protéine « High mobility group box 1 » (HMGB1). RFC est comme son nom l'indique, impliquée dans le processus de réplication, en permettant le recrutement de PCNA et de Pol δ [117]. De fait, elle permet de mettre en place le LP-BER, mais également d'autres systèmes de réparation comme le NER [118]. RFC peut également interagir avec l'ADN ligase I, que nous développerons dans la prochaine sous partie [119]. HMGB1 est quant à elle une protéine présente dans la chromatine, qui se fixe sur l'ADN sur les régions

internucléosomiques (rôle antagoniste à l'histone H1) et permet de réguler sa structure et l'expression des gènes. Elle est impliquée dans différents systèmes de réparation et notamment le BER en stimulant l'activité de FEN-1 et APE1, et donc en poussant la réparation vers le LP-BER [120]. Il semblerait également qu'elle puisse stimuler la Pol β [121].

II.4. Ligation finale

Quel que soit le mécanisme du BER employé, SP ou LP, l'ADN qui en résulte présente un nucléotide néosynthétisé en 5' faisant face à un groupement hydroxyle en 3'. Les ADN ligases sont des enzymes capables de suturer cette SSB et ont des implications dans la réplication, la réparation par HR, et dans le BER [122].

Il existe trois ADN ligases différentes chez l'Homme et les autres mammifères, les ADN ligases I, III et IV [71]. L'ADN ligase IV est uniquement présente dans le noyau, et est impliquée dans la réparation des DSBs et le système NHEJ, mais n'a pas montré d'implication dans le BER. L'ADN ligase I est elle aussi uniquement présente dans le noyau, mais interagit avec la protéine PCNA, RFC, et est donc requise pour le LP-BER dont elle est la seule ADN ligase [123]. Enfin, l'ADN ligase III est présente à la fois dans le noyau et dans les mitochondries des cellules. Pour le SP-BER, il semble également que Lig I soit l'enzyme majoritaire dans le noyau, tandis que Lig III, est la ligase nécessaire pour le SP-BER mitochondrial [124].

II.5. Ciblage thérapeutique du système BER

II.5.1. Réparation de l'ADN et cancer

Dans ce paragraphe, nous allons examiner en quoi les nombreux systèmes de réparation dont dispose la plupart des organismes vivants peuvent être des cibles thérapeutiques d'intérêt, et particulièrement pour la lutte contre le cancer. Ces considérations sont donc directement reliées à mon projet de thèse. Cependant, on se heurte ici à un paradoxe apparent qui est celui d'inhiber un mécanisme de protection de l'intégrité de l'ADN, afin de lutter contre un cancer, bien souvent lui-même résultant d'une mutation de l'ADN. Pour contourner ce paradoxe, il est important de prendre en compte plusieurs éléments.

Tout d'abord, les cellules cancéreuses ont une capacité de prolifération supérieure aux cellules saines, ce qui encourage donc à viser les différentes étapes du cycle cellulaire. Or, si certains traitements utilisent d'autres approches, comme impacter la division cellulaire par exemple [125], l'une des techniques les plus communes est d'exploiter les effets d'un traitement visant à endommager l'ADN, en utilisant par exemple des agents alkylants [9], des anti-métabolites [126] ou des radiations ionisantes [127]. C'est la principale méthode de la chimiothérapie ou de la radiothérapie en cancérothérapie. En induisant des dommages sur l'ADN, on espère induire la mort cellulaire de manière préférentielle des cellules cancéreuses. En revanche, l'efficacité de ces traitements peut donc être impacté négativement par les systèmes de réparation. Cibler ceux-ci peut donc permettre d'augmenter l'effet thérapeutique de ces traitements.

Certaines cellules cancéreuses ont aussi un système de réparation plus efficace que celui des cellules saines, et peuvent ainsi devenir résistantes aux traitements. Dans ce cas, l'ajout d'un inhibiteur dirigé sélectivement contre ce système de réparation permettrait de rendre ces cellules à nouveau sensibles aux traitements [128]. Les exemples de « re-senbilisation » à un traitement anti-cancéreux sont nombreux *in vitro* et *in vivo*, et nous en développerons certains dans les prochaines parties.

Récemment, des situations de létalité synthétique (LS) permettraient de mettre en place des traitements efficaces contre le cancer, en ayant un effet minime sur les cellules saines. Ce principe de létalité synthétique repose sur une interaction entre deux systèmes biologiques, gènes ou produits de gènes, participant tous deux à un processus essentiel dans la survie cellulaire [129]. L'absence de l'un ou l'autre

n'impacte pas la survie, tandis que l'ablation des deux systèmes va déclencher la mort cellulaire. Comme les cellules cancéreuses ont nécessairement subi une ou plusieurs mutations affectant potentiellement une fonction cellulaire, cette perte de fonction peut ainsi constituer son « talon d'Achille ». Dans le cadre de la mise en place d'une stratégie thérapeutique, l'exploitation du principe de létalité synthétique consiste donc à identifier un deuxième système ayant une relation de létalité synthétique avec ces cellules cancéreuses mais non-essentiel pour les cellules saines. En ciblant ce système, on induira la mort cellulaire préférentiellement des cellules cancéreuses tandis que les cellules saines seront peu impactées (Figure 16).

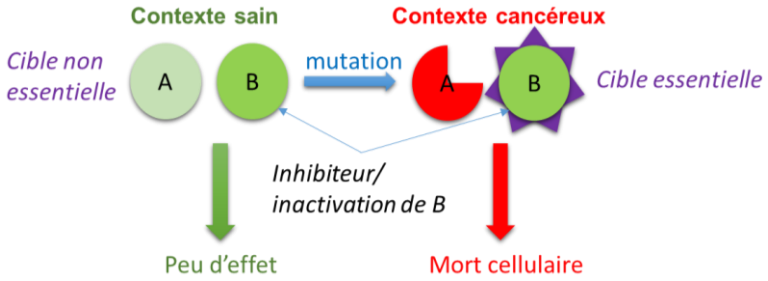


Figure 16 : Principe de la létalité synthétique - A et B étant des gènes ou les produits d'un gène.

Il y a donc de multiples raisons justifiant le ciblage des systèmes de réparation de l'ADN, et l'inhibition de la plupart a donc été fortement envisagée. Ainsi, de nombreuses études portent sur l'inhibition des PIKKS ou de leurs enzymes associées, des réparations de cassures simple/double brin, ou encore des systèmes de réparation par excision/resynthèse comme le NER ou le MMR. Faire un compte rendu de l'immensité des composés existants et de leurs applications potentielles n'étant pas l'objectif de cette thèse, nous invitons les lecteurs intéressés à se référer aux différentes revues suivantes pour plus d'informations [130–133].

II.5.2. Pourquoi inhiber le BER ?

Nous allons ici nous concentrer sur les applications potentielles que présente l'inhibition du système BER. Comme nous l'avons vu, le BER fait appel à de nombreuses protéines, qui sont aussi pour la plupart fortement impliquées dans la réplication de l'ADN ou dans d'autres systèmes de réparation. La mutation de la plupart des enzymes du BER, à l'exception notable des ADN glycosylases, est la cause de formation de tumeurs [134]. Ainsi, l'objectif premier est de trouver un contexte dans lequel ces enzymes sont des cibles pertinentes, ce qui ne sera pas forcément le cas dans n'importe quel type cellulaire. Par exemple, les dommages induits par certains agents génotoxiques utilisés en cancérothérapie sont de bons substrats pour le BER. Ainsi, le ciblage des enzymes du BER peut permettre d'augmenter la sensibilité des cellules cancéreuses à ces agents [135]. De plus, les cellules cancéreuses agressives contiennent plus de lésions (comme l'uracile ou la 8-oxoG) que les cellules saines [136] et dépendent donc plus de la réparation de ces damages que les autres cellules. Si le mécanisme sous-jacent n'est pas encore bien compris, il est admis que certains oncogènes augmentent la présence des ROS, et donc l'oxydation de l'ADN [137]. Enfin, une relation de létalité synthétique peut également être utilisée. Concrètement, l'inhibition sélective du BER peut-être pertinente en cancérothérapie selon les 3 modalités illustrées en Figure 17 [137].

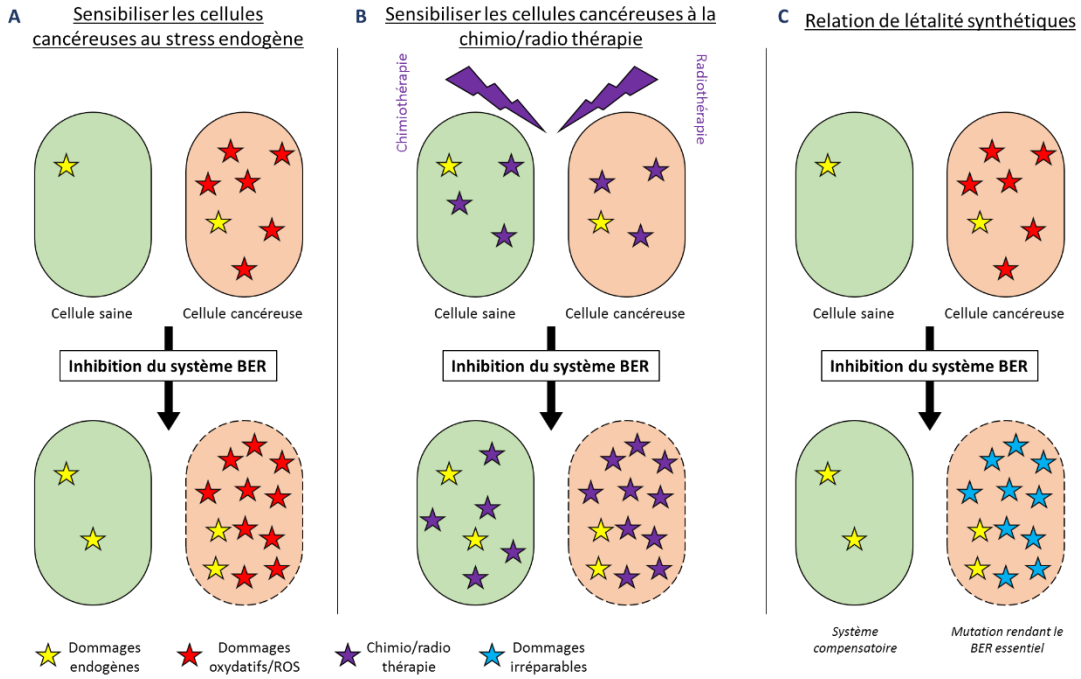


Figure 17 : **Stratégies principales pour l'inhibition du BER (Inspirée de Visnes *et al.*, 2018 [137])** - A) Dans la situation où les cellules cancéreuses produisent plus de ROS et donc de lésions oxydatives de l'ADN, il est possible que l'inhibition du BER permette d'induire la mort cellulaire en introduisant plus de dommages qu'une cellule puisse en supporter. En revanche, pour les cellules saines, subissant moins de dommages endogènes, cette augmentation serait trop faible pour avoir un effet. B) Comme le système BER peut protéger des dommages provoqués en chimio/radiothérapie, l'inhibition de ce système permettrait d'augmenter l'effet du traitement. Cependant, bien que ces traitements soient les plus ciblés possible, cela implique de sensibiliser les cellules saines. C) En exploitant une relation de létalité synthétique entre une enzyme du BER et une ou plusieurs fonctions perdues par mutagenèse et caractéristiques des cellules cancéreuses ciblées, seules les cellules cancéreuses seront touchées par le traitement.

Il nous semble important de présenter ici l'un des exemple les plus significatifs d'une relation de létalité synthétique exploitée avec succès, qui est l'inhibition des enzymes Poly(ADP-ribose) polymérases (PARPs) chez les femmes présentant des mutations pathogènes sur les gènes BRCA1/2. Il s'agit d'un des facteurs de risque les plus puissants de cancer sein et des ovaires puisque plus de 50% des femmes concernées développe la maladie [138]. L'enzyme PARP, bien qu'elle ne soit pas directement engagée dans la réparation du BER, est impliquée dans le recrutement de différentes enzymes du BER (voir paragraphe I.2.3). L'intérêt de cibler les enzymes PARPs a émergé suite à des études ayant montré que les cellules BRCA1/2 ne disposant pas de système HR opérationnel étaient très sensibles aux inhibiteurs dirigés contre PARP1 [139]. En revanche, les cellules disposant d'un HR compétent survivaient à l'action des inhibiteurs de PARP. Ce résultat était en accord avec l'observation que des souris KO (« knock-out ») pour PARP1 sont viables même si elles semblent subir un processus de vieillissement plus rapide et un risque plus élevé de cancérogenèse [140] lié à une sensibilité aux agents alkylants, aux radiations gamma ou à d'autres agents créant des lésions de l'ADN [141]. Or, dans certains cancers héréditaires du sein et des ovaires une déficience du système HR est bien souvent reliée à une mutation dans les gènes BRCA1 et BRCA2 [142]. En effet dans ces cellules cancéreuses, le système de réparation HR n'est plus capable de recruter RAD51 sur les DSBs [143]. Ces cellules BRCA deviennent donc dépendantes du système NHEJ ou alt-NHEJ, qui requiert l'utilisation de PARP1. Il existe donc une relation synthétique létale entre les gènes BRCA1/2 et PARP1.

De nombreux inhibiteurs de PARP1 ont donc été synthétisés et testés sur différentes lignées cellulaires (Figure 18). La recherche d'inhibiteurs de la PARP1 est un domaine de recherche très actif, qui voit l'émergence de nouvelles stratégies anticancéreuses prometteuses. Si les premiers inhibiteurs (PARPi) étaient des analogues de la nicotinamide avec pour objectif de sensibiliser les cellules à des traitements anti-cancéreux [144], de nouveaux composés ont rapidement été synthétisés par différentes entreprises

phramaceutiques : Olaparib (KuDOS/Astrazeneca), Rucaparib (Pfizer/Clovis), Veliparib (Abbvie), Niparib (Merck/Tesaro) ou par des groupes de recherches comme le composé NMS-P118 [145]. La plupart de ces inhibiteurs sont des composés entrant en compétition avec le NAD^+ pour son site de fixation à la PARP1 et induisent un blocage de la formation des pADPrs. En condition d'inhibition de la PARP1, l'enzyme est piégée sur la cassure simple brin (SSB) et empêche le recrutement des enzymes qui terminent la réparation. Il en résulte une accumulation des SSBs qui peuvent se transformer en DSBs après réplication [146]. Pour la prise en charge de ces dommages en présence de PARPi, l'implication du système HR est essentielle, ce qui n'est pas possible pour les cellules mutées sur les gènes BRCA1/2. Les PARPi sont utilisés à la fois en monothérapie en exploitant la relation de LS, mais également en combinaison avec des agents alkylants afin d'augmenter leurs effets cytotoxiques [147]. De très nombreuses études cliniques et précliniques ont montré des résultats enthousiasmants [138,148], et les composés Olaparib, Niraparib et Rucaparib ont déjà été validés par la FDA pour le traitement de différents cancers. Certains mécanismes de ces inhibiteurs ne sont cependant pas encore complètement élucidés, ce qui peut conduire à de nouvelles stratégies prometteuses [147,149–151].

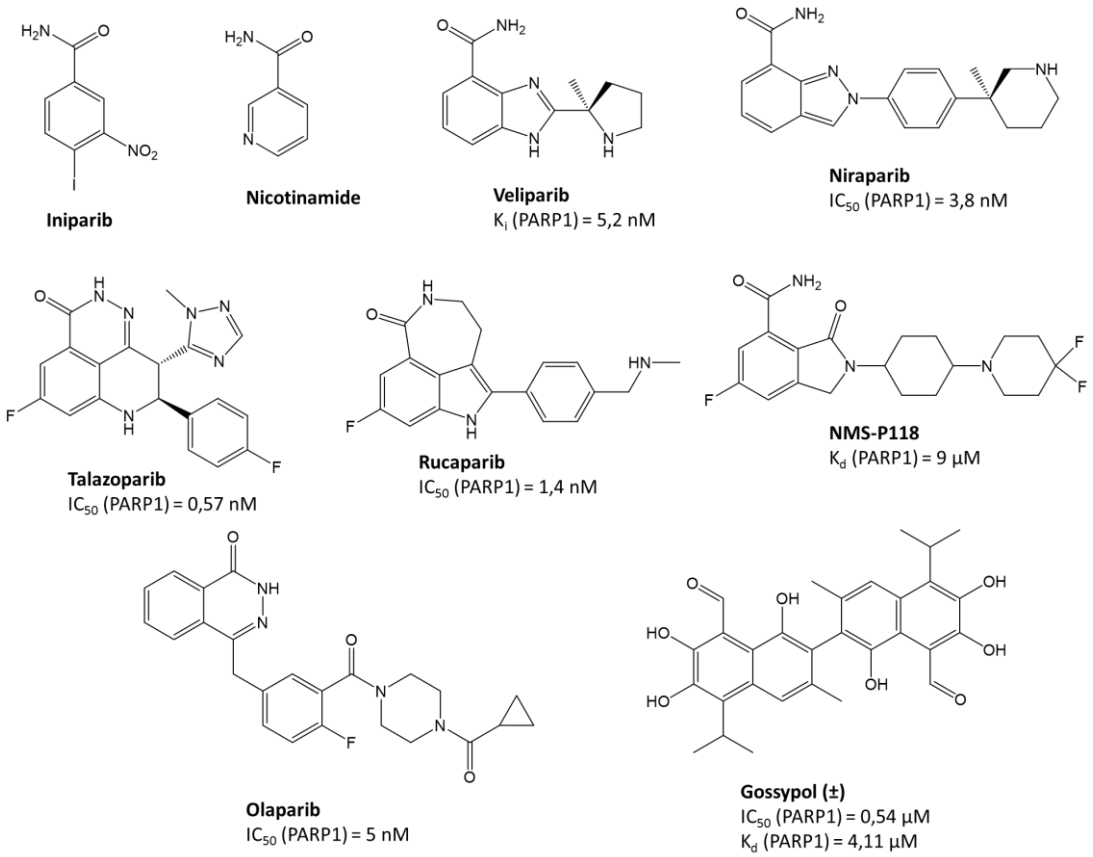


Figure 18 : Exemple d'inhibiteurs de PARP1

Il est important de noter, que les petites molécules bloquant la formation des pADPrs ne sont pas les seuls inhibiteurs prometteurs des PARPs. Ainsi, le gossypol, un polyphénol naturel, a été identifié comme étant un inhibiteur de l'interaction protéine-protéine de PARP/PARP et a montré de bonnes activités à la fois *in vitro* et *in vivo* [152]. L'utilisation d'un ADN double brin DT01 (DBait), composé de 32 paires de bases et imitant les lésions de l'ADN est aussi capable d'« appâter » PARP1, ou les PIKKS et donc d'inhiber le recrutement des enzymes de réparation sur le site du dommage [153]. Ainsi, ces ADN DBaits permettent d'améliorer les effets de différentes thérapies anti-cancéreuses et montrent des résultats très prometteurs [154,155].

L'engouement suscité par les résultats prometteurs des essais de PARP1 a mis en lumière l'intérêt des systèmes de réparation de l'ADN, et l'utilisation des relations de létalité synthétique. Dans le paragraphe suivant, nous allons examiner les différentes enzymes du BER et leur potentiel intérêt thérapeutique.

II.5.3. Inhibition des enzymes du BER suite à la formation d'un site abasique

- **APE1**

Comme nous l'avons expliqué, APE1 est l'une des enzymes principalement requises pour prendre en charge les produits des ADN glycosylases monofonctionnelles ou bifonctionnelles avec une β-élimination. Cette enzyme a également une fonction rédox et peut activer un certain nombre de facteurs de transcription [156]. De fait, son inhibition est problématique car elle impacte de nombreux mécanismes, et son inactivation empêche la survie cellulaire chez la souris [35]. Cependant, de par son rôle central dans les systèmes de réparation et sa grande multifonctionnalité, cette enzyme est considérée comme une cible prometteuse pour la lutte contre le cancer et est l'une des enzymes du BER les plus documentées en ce qui concerne son ciblage [137]. De plus, sa surexpression permet à certains cancers d'accroitre leur résistance à différents agents chimiothérapeutiques [157]. De fait, de nombreux inhibiteurs ont été développés [158], et ont pour la plupart montré des effets de sensibilisation des cellules cancéreuses (cancers du sein, du système nerveux ou encore des gliomes) aux agents génotoxiques (alkylants et radiations ionisantes) [130,159]. L'un des composés ayant été les plus utilisé en clinique sur des cancers du cerveau ou du poumon et en combinaison avec la radiothérapie est le Lucanthone, à l'origine un inhibiteur des topoisomérases II, mais son effet thérapeutique n'a pas vraiment été concluant [160]. Malgré tout, de nouvelles études utilisant d'autres composés montrent un effet sur des cellules de cancer colorectal, et un effet sur les métastases de cancer cervicaux [161]. De même, de nouveaux essais cliniques ont validé la phase I d'inhibiteurs d'APE1 sur des tumeurs solides [162,163]. De très nombreuses applications thérapeutiques sont donc envisageables pour l'inhibition d'APE1, et les lecteurs intéressés sont invités à lire les revues indiquées pour aller plus en profondeur sur ce sujet [158,164–166].

- **Polymérase β**

Il a été démontré que des mutants déficients en Pol β sont hypersensibles à certains agents alkylants, et que sa disruption est associée à une létalité embryonnaire chez la souris [167]. De plus, des mutations sur Pol β ont été observées dans 30 % des cancers solides [168], ce qui en fait une cible intéressante. Ainsi de très nombreuses petites molécules ont été criblées et/ou synthétisées et identifiées comme des inhibiteurs de Pol β [169], mais également des peptides [170] ou des composés naturels [171]. Cependant, la plupart de ces inhibiteurs manquent de spécificité et/ou d'efficacité, même si certains ont sensibilisé des cellules de cancer du côlon à l'agent TMZ [172]. En revanche, plus récemment, l'utilisation d'un inhibiteur irréversible imitant la structure de l'ADN a augmenté l'efficacité de l'agent méthyl-méthanesulfonate dans des cellules HeLa [173].

- **ADN ligases**

L'inhibition des ADN ligases impliquées dans le BER n'a rencontré jusqu'ici qu'un succès mitigé. Si des inhibiteurs de l'ADN ligase A, l'homologue bactérien de Lig I, ont été identifiés [174], peu de composés montrent une spécificité contre les deux ADN ligases humaines I et III impliquées dans le BER. En effet, la plupart des composés inhibant les ADN ligases sont des composés intercalant ayant d'autres interactions biologiques (ions métalliques, pyridoxal phosphate, acides aurintricarboxyliques) [175]. Des analogues de nucléosides ont également été testés [176]. Plus récemment un analogue de benzocumarine-stilbène, a cependant montré des effets *in vitro* et *in vivo* dans des cellules de souris pour la réduction de croissance d'un cancer du sein [177] et est, à notre connaissance, l'inhibiteur le plus abouti des ADN ligases du BER. La

ligase I a également montré une relation de létalité synthétique avec des cellules déficientes en XRCC1, ce qui pourrait conduire à de nouvelles études plus approfondies [178].

- **FEN-1**

Des études récentes ont montré que FEN-1 permettrait une progression tumorale plus rapide, et conférerait une résistance au cisplatine dans les cellules tumorales [179]. Son inhibition augmente également la sensibilité de certaines cellules de cancer cervical à la radiothérapie [180]. De plus, une relation de létalité synthétique a été établie entre FEN-1 et des cellules ayant un système HR défectueux notamment les cellules de cancer ovarien ayant des gènes BRCA1/BRCA2 mutés [181,182]. Ces études récentes montrent donc l'intérêt de l'inhibition de FEN-1 et d'autres études identifieront probablement des inhibiteurs de cette enzyme à l'avenir.

- **PNKP**

L'inhibition de PNKP n'a été étudiée que tardivement, et a déjà montré des résultats enthousiasmants. Ainsi, si des inhibiteurs de PNKP étaient déjà connus (notamment l'héparine, un anti coagulants) peu d'études montrant un effet thérapeutique potentiel avaient été effectuées [183]. L'imidopipéridine, un inhibiteur de PNKP, a permis de sensibiliser des cellules leucémiques ou de cancer de la prostate à différents traitements de radiothérapie [184,185]. De plus, une relation de létalité synthétique a été identifiée entre PNKP et le gène PTEN, un gène de suppression des tumeurs, souvent muté dans des cancers humains, ce qui accentue l'importance de PNKP en tant que cible thérapeutique [186]. Plus récemment encore, un composé imidopipéridine de deuxième génération plus efficace a permis d'établir *in vivo* et *in vitro* la sensibilisation aux radiations de cellules de cancer colorectal mutées sur le gène PTEN [187]. Ce composé peu soluble a dû être au préalable encapsulé par des nano-particules afin d'être délivré dans les cellules [188]. Ceci souligne l'importance du ciblage de la PNKP dans la lutte contre le cancer, mais également l'intérêt des nano-médecines, une solution potentielle pour les composés peu solubles.

- **Co-facteurs**

L'inhibition d'autres facteurs moléculaires du BER tels que HMGB1, XRCC1 ou RFC est en cours d'évaluation dans des contextes différents du BER. Ces enzymes sont impliquées dans de nombreux systèmes biologiques que nous ne développerons donc pas ici [189,190].

II.5.4. Ciblages des ADN glycosylases humaines

Même si de nombreuses enzymes du BER semblent être des cibles pertinentes, leur inhibition a rencontré plus ou moins de succès car leur inactivation (APE1, Pol, XRCC1, Lig III, FEN-1) montre des effets embryonnaires létaux chez la souris. Il reste donc les ADN glycosylases, dont tous les KO sont viables et fertiles, à l'exception de TDG qui est associée à un phénotype embryonnaire létal (probablement à cause de ses fonctions épigénétiques dans la régulation des méthylations de l'ADN [191]). De plus, contrairement aux autres enzymes du BER, les ADN glycosylases initient spécifiquement le BER, et sont donc moins impliquées dans d'autres systèmes de réparation. Le ciblage pharmacologique des ADN glycosylases devrait être mieux toléré par les cellules saines de l'organisme et, particulièrement exploitable dans des situations de relations synthétiques létales. Cet effet limité de l'inactivation des ADN glycosylases s'explique en partie par la redondance de substrats qu'il existe entre les ADN glycosylases. En effet, lors de doubles KO, les phénotypes observés sont beaucoup plus graves. Par exemple, le KO d'OGG1 et de MUTYH induit l'apparition de tumeurs, car ces deux enzymes peuvent se complémenter [192]. De même, NTHL1 et NEIL1 partagent un certain nombre de substrats, et un double KO provoque le développement de cancers du foie ou des reins [193]. Cela renforce donc le rationnel d'une stratégie exploitant l'existence de LS.

Contrairement à l'inhibition de PARP1, qui va impacter l'ensemble du BER, l'inhibition des ADN glycosylases, qui réparent spécifiquement un certain nombre de lésions, est plus limitée et ambivalente. Cela peut s'avérer positif en combinaison avec d'autres molécules créant précisément les dommages reconnus spécifiquement par l'ADN glycosylase cible [194]. En revanche, cela peut avoir un effet négatif, car la réparation du dommage par les ADN glycosylases résulte en un site abasique, généralement plus toxique et peut donc « potentialiser » le dommage génotoxique [195]. En empêchant la formation d'un site abasique par les ADN glycosylases, ces inhibiteurs peuvent donc paradoxalement « diminuer » l'effet de certains traitements, dans des cancers ou leurs effets antibiotiques par exemple [195,196]. Dans le prochain paragraphe, nous détaillerons les différents inhibiteurs des ADN glycosylases humaines, en nous concentrant sur les petites molécules.

- **ADN glycosylases réparant principalement l'uracile**

L'uracile peut apparaitre dans l'ADN (i) suite à l'incorporation de dUTP à la place de dTTP lors de la réplication, (ii) suite à la désamination naturelle d'une cytosine dans l'ADN ce qui conduit à des mutations de type transition G:C vers A:T , ou (iii) suite à l'action d'un agent exogène (des métabolites par exemple, souvent utilisés pour le traitement de certains cancers) [197]. Quatre ADN glycosylases humaines sont capables d'exciser l'uracile : UNG, SMUG1, TDG et MBD4. Dans la plupart des cas, c'est UNG qui est principalement mobilisée [198].

La réparation de l'uracile peut être la cause de résistances aux traitements d'anti-métabolites, ou augmenter l'effet d'autres agents en provoquant une augmentation des SSBs [199,200]. Etant l'enzyme majoritairement mobilisée, l'inhibition d'UNG1/2 semble être l'option la plus prometteuse. L'inactivation de cette enzyme chez des cellules de prostate humaine augmente leur sensibilité à des agents thérapeutiques (doxorubicine et H_2O_2) [201]. On retrouve également un fort taux d'expression de UNG dans des cellules cancéreuses (cancer gastrique ou des poumons), réduisant l'effet de certains agents thérapeutiques génotoxiques [202,203]. Il a été montré en particulier que UNG peut également moduler la résistance au cisplatine [204]. Enfin, une relation de létalité synthétique a été mise en avant entre UNG2 (la forme mitochondriale) et le gène APOBEC3B, un gène mutant présent dans certaines cellules cancéreuses et codant pour une cytidine désaminase. De fait, l'inactivation de UNG2 permet d'induire la mort cellulaire spécifique des cellules mutantes [205]. La plupart des inhibiteurs de UNG sont en phase préclinique et on peut distinguer différentes classes d'inhibiteurs (Figure 19) :

- des inhibiteurs protéiques naturels produits par les bactériophages PBS et PBS2 capables de former un complexe abortif PBS2/UNG. Leur impact thérapeutique n'a pas été étudié mais ils sont utilisés dans la technologie CRISP/Cas9 [206] ;

- des inhibiteurs de type court duplexe d'ADN contenant des aza-nucléotides imitant les états de transition de l'activité du complexe UNG/uracile-ADN [207] ;

- des inhibiteurs synthétiques de type analogue d'uracile [208]. Par la suite, de nouveaux inhibiteurs composés à la fois d'un groupement uracile et d'un groupement phénolique (séparés par une chaîne carbonée) ont été synthétisés et optimisés [209–211]. L'uracile va se placer dans le site actif de l'enzyme, tandis que le phénol se lie dans le domaine d'interaction à l'ADN. Enfin suite à un criblage, des inhibiteurs composés de 3 groupements fonctionnels (contenant un fragment aromatique et des analogues d'uracile) ont également été testés [212].

Pour les autres enzymes cependant, peu d'inhibiteurs sont décrits. Pourtant l'ablation de SMUG1 s'accompagne d'une radio-sensibilisation de certaines lignées cellulaires de souris [213,214], tandis que la suppression de MBD4 sensibilise des fibroblastes d'embryons de souris à certains agents cytotoxiques [215].

Cependant, étant donné la proximité structurale entres les enzymes de cette famille, il est fort probable qu'un inhibiteur de petite taille puisse toutes les inhiber, ce qui diminue l'importance de l'ADN glycosylase ciblée [195]. A notre connaissance aucun inhibiteur spécifique de ces deux enzymes n'a été identifié. TDG présente, quant à elle, le seul KO non viable des ADN glycosylases humaines, et semble donc essentielle au développement de l'embryon [191]. De manière générale, son inhibition ne semble pas provoquer une sensibilisation particulière à des agents génotoxiques [216]. Il a été observé cependant que son inactivation dans les mélanomes induisait spécifiquement la mort cellulaire de ceux-ci [217]. Suite à cette observation, de petites molécules inhibitrices ont été identifiées grâce à un criblage : la Juglone et la Closanthel. Malheureusement, comme nous allons le voir dans ce manuscrit, la Juglone est un composé intercalant, apparaissant souvent positifs dans des criblages. Cette classe d'inhibiteurs n'est certainement pas spécifique. Des inhibiteurs ont aussi été développés en utilisant des complexes d'ADN contenant des cytosines modifiées [218].

Si ce manuscrit se concentre sur l'inhibition des enzymes humaines, il est intéressant de noter que l'inhibition des UNG chez certains virus peut également être utilisée pour lutter contre différentes infections [219]. Par exemple, l'inhibition des UNG des poxvirus réduit leurs réplications [220].

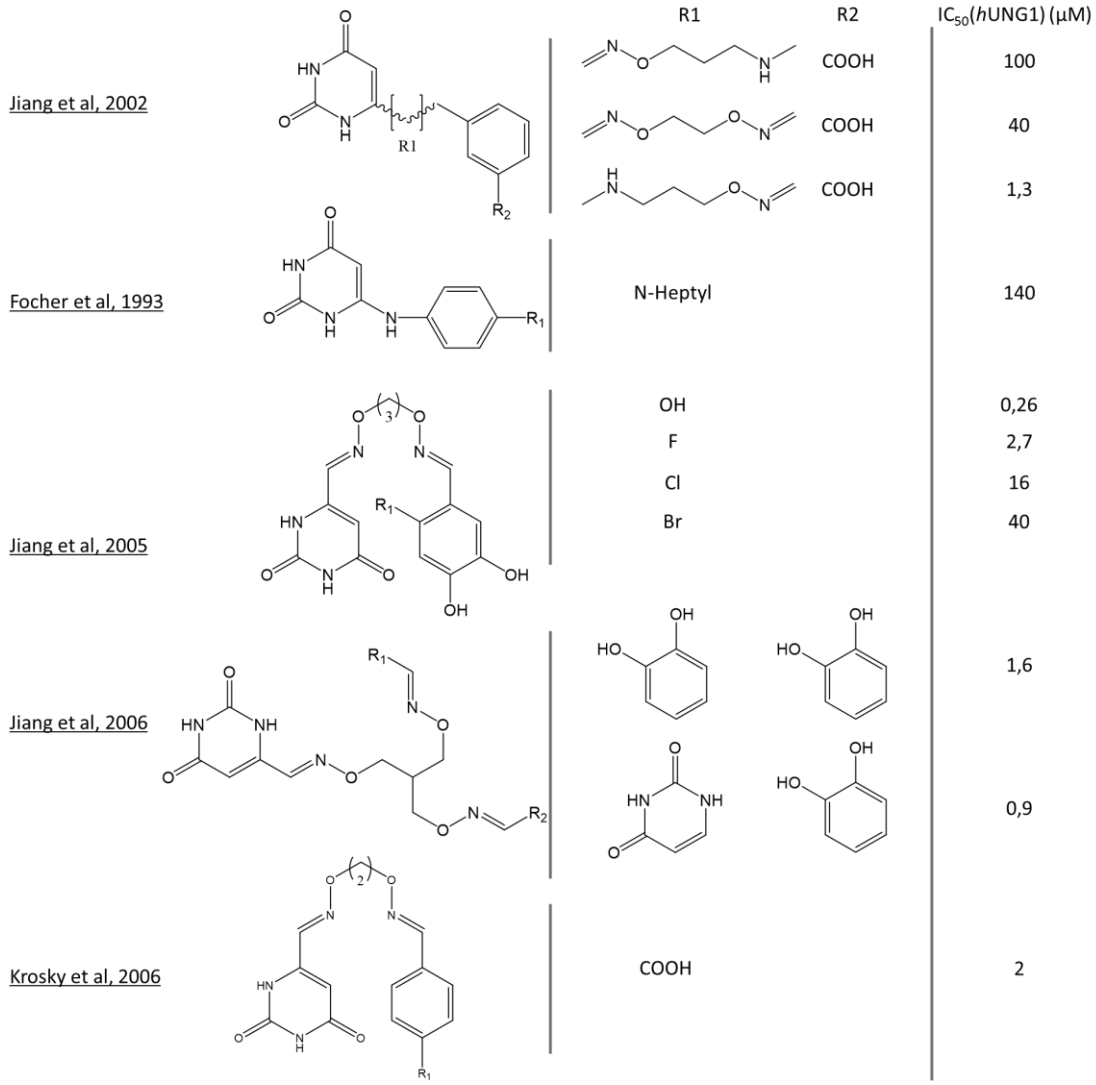


Figure 19 : Exemples d'Inhibiteurs de UNG1/2

- **ADN glycosylases réparant principalement les dommages alkylants**

Pour les bases alkylées de l'ADN, l'Homme dispose de la N-méthylpurine ADN glycosylase (MPG), qui possède une gamme assez large de substrats comme la 7-meG, 3-meG ou 3-meA [221] ainsi qu'un certain nombre de lésions souvent endogènes comme l'éthénoadénine (εA). La plupart des chimiothérapies utilisant un grand nombre d'agents alkylants, l'inhibition de MPG semble donc être une stratégie intéressante. Cependant, nous retrouvons encore une fois le côté ambivalent de l'inhibition des ADN glycosylases.

L'activité ou la surexpression de MPG dans certains types de cancers (colorectaux ou du sein) entraine une sensibilisation des cellules cancéreuses à certains agents alkylants [222,223]. Cela s'explique toujours de la même manière ; la formation de sites abasiques par MPG peut être plus toxique que la lésion en elle-même (pour rappel MPG est une ADN glycosylase monofonctionnelle). Cet effet toxique de MPG est exploité par de nombreuses chimiothérapies [224]. Par exemple, le témozolomide (TMZ), un agent alkylant très utilisé en cancérothérapie, peut être combiné avec la méthoxyamine, qui va bloquer le BER en se liant à la fonction aldéhyde des sites abasiques [225]. Dans ce contexte, la surexpression de MPG va donc augmenter la cytotoxicité du traitement par saturation des acteurs moléculaires du LP-BER. De même, certains agents alkylants dépendant de l'activité de PARP1 voient leur toxicité augmenter en cas de surexpression de MPG [226].

En revanche l'inactivation de MPG peut aussi sensibiliser des cellules de cancers ovariens, des cellules HeLa ou de souris, à des agents alkylants [227]. L'activité de MPG peut ainsi permettre une résistance à TMZ dans des glioblastomes, résistance qui peut être contournée en inhibant également PARP1 [228,229]. Enfin, une surexpression de MPG peut être associée à un cancer des poumons, tandis que le KO de MPG chez la souris induit moins de nécrose tissulaire en cas d'agent attaquant l'ADN [230,231].

Des inhibiteurs de MPG ont été développés, mais peu se sont révélés efficaces et aucune application n'a été effectuée à ce jour (Figure 20). Les premiers inhibiteurs identifiés sont des analogues de substrats, mais servent principalement comme contrôle dans des études fonctionnelles *in vitro* [232]. MPG peut également être inhibée par certains ions métalliques comme le magnésium Mg^{2+} par exemple, mais ce type de composé est évidemment fort peu spécifique et ne permet pas d'application *in vivo* [233]. Des criblages virtuels ont également été effectués, et ont permis d'identifier un composé de la famille des favonols, la Morine, dont l'effet a été confirmé [234]. Ce composé semble également capable de pénétrer les cellules avec une concentration *in cellulo* assez élevée. Son activité cependant assez faible et sa sélectivité réduite ont stimulé la recherche d'autres inhibiteurs plus actifs [235]. Une équipe a également inhibé MPG en utilisant un ADN de 13 paires de bases contenant l'éthénodéoxycytidine (εdC) ou éthénodéoxyadénosine (εdA), avec un IC_{50} de l'ordre du nanomolaire [236]. Ces travaux ont par la suite inspiré l'utilisation de εdC/εdA comme inhibiteur libre de MPG [237]. La structure cristalline de ces inhibiteurs sur MPG a permis de synthétiser une nouvelle génération d'inhibiteurs de la forme aza-nucléosides (qui ont déjà été utilisés pour inhiber d'autres ADN glycosylases comme nous l'avons vu pour UNG). Cependant ces composés présentent une activité bien plus faible que le substrat ADN inhibiteur, avec un IC_{50} de l'ordre du micromolaire. Ils pourront servir de « fragments moléculaires » pour la conception et la synthèse d'inhibiteurs plus efficaces.

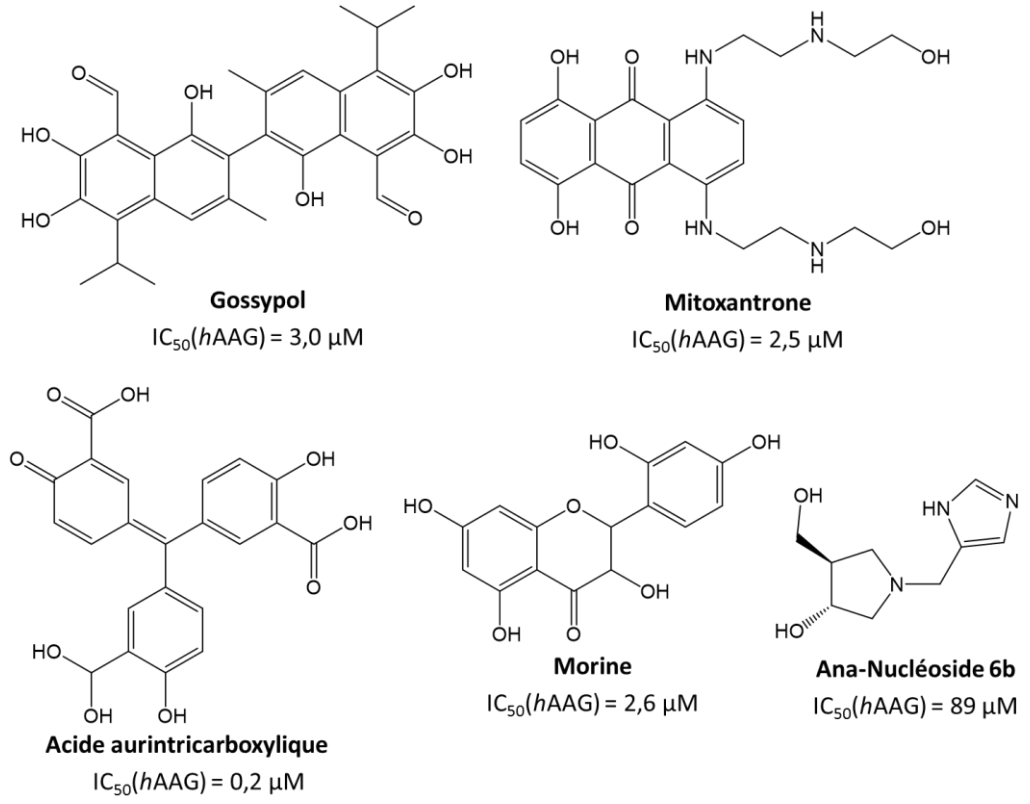


Figure 20 : Inhibiteurs de MPG

- **ADN glycosylases réparant principalement les dommages oxydatifs**

Six ADN glycosylases humaines sont principalement impliquées dans la réparation des bases oxydées : NEIL1/2/3, OGG1, MUTYH et NTHL1. Les enzymes hTDG et UDG sont également capables d'exciser certaines bases oxydées comme la thymine glycol ou la 8-oxoAdénine et la 5-hydroxyuracile respectivement [238,239]. Comme nous avons déjà développé les inhibiteurs de ces enzymes précédemment, nous ne les évoquerons pas dans cette partie. Comme NEIL1 et OGG1 constituent les objets d'étude de cette thèse, nous développerons les aspects de leurs ciblages dans les paragraphes suivants qui leurs seront entièrement consacrés.

Avant d'étudier les ADN glycoylases humaines, il nous semble important de rappeler rapidement un autre système de réparation. La guanine, qui possède le potentiel rédox le plus bas des bases nucléiques, est la plus susceptible d'être oxydée par les ROS. Son produit majeur d'oxydation, la 8-oxoguanine (8-oxoG), est très mutagène [32]. Chez la bactérie, notamment *E. coli*, la 8-oxoG est prise en charge par le « système GO », qui consiste en l'utilisation de trois enzymes Fpg (MutM), MutY et MutT. Tout d'abord Fpg excise la 8-oxoG en face d'une cytosine et conduit à la formation d'une lacune de 1 nucléoside prise en charge par les enzymes agissant en aval dans le SP-BER [23]. Cependant, si la 8-oxoG n'est pas éliminée, ces propriétés particulières d'appariement à l'adénine peuvent conduire après 2 cycles de réplication à des transversions G:C vers T:A. En effet, la 8-oxoG est capable de former une paire de bases non canonique de type Hoogsteen avec l'adénine. Dans la situation où Fpg n'a pas eu le temps d'éliminer la 8-oxoG et qu'un 1er cycle de réplication conduit au mésappariement 8-oxoG:A, l'ADN glycosylase MutY est capable d'exciser l'adénine opposée à la 8-oxoG et reconduire à une paire 8-oxoG :C substrat de Fpg [240]. En ce sens, MutY fait office d'enzyme de « secours ». Enfin, MutT peut hydrolyser le 8-oxodGTP, afin de directement prévenir l'incorporation de la 8-oxoG dans l'ADN en face d'une adénine [241]. Bien que ce système soit bien caractérisé, son implication dans des bactéries pathogènes reste peu étudiée. Les inhibiteurs de ces systèmes bactériens peuvent aussi être à

double tranchant. En effet, ces enzymes de réparation protègent les bactéries et contribuent à leur pouvoir infectieux, ce qui stimule la recherche d'inhibiteurs agissant comme des antibiotiques [242,243]. Cependant, une inactivation du système GO pourrait également mener à une mutation plus élevée des bactéries, et donc à de nouvelles souches résistantes aux différents traitements [244]. Peu d'inhibiteurs du système GO ont cependant été identifiés. Fpg est principalement inhibé par des analogues de bases endommagées [245]. La 2-thioxanthine et certains de ces dérivés ont été identifiés dans notre équipe et sont capables d'attaquer le doigt à zinc de l'enzyme [246,247]. Quelques inhibiteurs de MutY ont été reportés, notamment la suramine, un médicament antiprotozoaire, mais aucune application n'a été envisagée à ce jour [248].

Pour MUTYH (homologue fonctionnel de MutY) et NTHL1, la recherche d'inhibiteurs a connu peu de succès. Cela s'explique tout d'abord par l'intérêt thérapeutique limité du ciblage de ces enzymes. En effet, avec un fort recouvrement de substrats entre ces enzymes, une inhibition sans contexte particulier ou létalité synthétique a peu de chance de montrer de forts effets. Par exemple pour MUTYH, bien qu'une mutation de cette enzyme soit très fortement impliquée dans l'apparition de nombreux type de cancers [249,250] ou dans le développement de polyposes [251], cette enzyme n'a pas encore montré de relation de létalité synthétique, ni d'effet négatif d'une surexpression. Une mutation de NTHL1 est également impliquée dans le développement de polyposes [252]. Cependant là encore, son inhibition semble avoir un intérêt limité, car son inactivation ne montre pas de sensibilisation aux agents oxydants [253]. Cependant si l'inhibition directe de NTHL1 n'a pas encore montré d'intérêt, des équipes ont pu identifier une application en combinaison avec le cisplatine. En effet, NTHL1 interagit avec le facteur de transcription YB1, et ce complexe semble induire une résistance au cisplatine [254]. Ainsi, l'utilisation d'un composé inhibant, non pas l'activité catalytique de NTHL1, mais son interaction avec YB1, a permis de partiellement restaurer l'efficacité du cisplatine (Figure 21) [255].

Enfin, il existe une autre enzyme impliquée dans la réparation des dommages oxydatifs chez l'Homme, l'enzyme homologue de MutT (MTH1). Bien que ce ne soit pas une ADN glycosylase, elle fait partie intégrante du système GO chez l'Homme, et montre un très bon intérêt thérapeutique. En effet, une surexpression de MTH1 dans certaines cellules est considérée comme pro-oncogène en promouvant des métastases et est donc associée à un diagnostic négatif pour certains types de cancers des poumons [256,257]. De plus, MTH1 est requise pour la prolifération de tumeurs thyroïdales [258]. Son inactivation est également toxique sur certaines cellules de tumeurs, à cause de l'accumulation de dommages oxydatifs [195]. De ce fait, cette enzyme a été particulièrement étudiée, et de nombreux inhibiteurs ont été identifiés et caractérisés (Figure 21). Tout d'abord, le composé SCH5134 a été identifié comme un inhibiteur de MTH1, puis la (S)-crizonitib a également été découverte [259]. Cette molécule, dont l'énantiomère (R) est un inhibiteur de tyrosine kinase, a pu démontrer un effet apoptotique sur des cellules de cancer du poumon [260], un effet anti-tumoral sur des cellules de cancer colorectal [261], et semble également réduire la croissance de cellules de cancer gastrique [262]. D'autres inhibiteurs ont été également découverts, notamment les composés TH287, TH588 et TH1579 (Karutonib) [263]. Le Karutonib, le composé le plus soluble, semble posséder une activité anti-tumorale contre les mélanomes et les cellules de cancer de la prostate [264–266] en réduisant également la croissance tumorale *in vivo* et *in vitro* [264]. Deux études cliniques de phase I sont en cours sur ce composé, et les résultats devraient être publiés sous peu [267,268]. Bien que moins soluble, le composé TH287 semble tout de même arrêter le développement de cancer gastrique [269] et son problème de solubilité a été contourné en l'encapsulant dans des nanoparticules afin de le délivrer directement dans des cellules d'ostéosarcome, induisant leur apoptose [270]. D'autres inhibiteurs ont également été identifiés, comme le composé MPD15095 [271]. Le SU0383 est particulièrement original car il est composé de l'assemblement de deux fragments inhibiteurs respectivement de hOGG1 et de MTH1, lui permettant donc d'inhiber ces deux enzymes [272]. Très récemment encore, un nouveau composé a été identifié suite à des criblages à haut débit

(HTS pour « Hight Throughtput Screening ») de petites molécules fragments, la molécule M10639, dont l'efficacité reste à optimiser [273]. De même un HTS de produits naturels a permis d'extraire la molécule émodine, qui semble également diminuer la croissance de certaines cellules de cancer du poumon [274]. En revanche, ce composé semble intercalant, et des études sur sa spécificité devront être conduites pour aller plus loin. Toutes ces études montrent que l'inhibition de MTH1 semble particulièrement prometteuse.

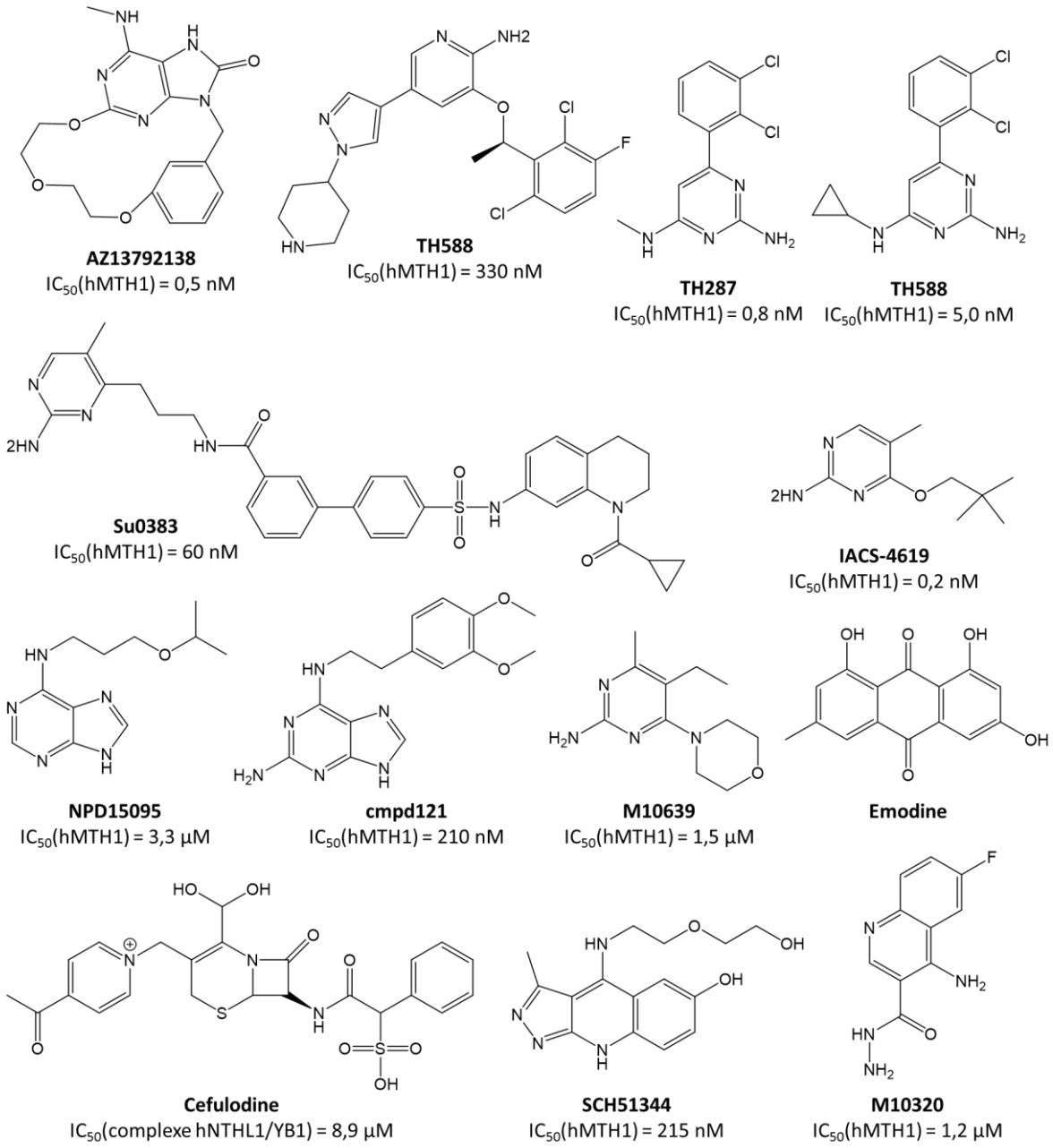


Figure 21 : Inhibiteurs de MTH1 et NTHL1

Cependant, l'enzyme principale pour exciser les purines oxydées, et plus particulièrement la 8-oxoG, est hOGG1, tandis que les pyrimidines oxydées sont plutôt prises en charges par hNEIL1/2/3 (et NTHL1 que l'on a déjà évoqué), qui sont précisément les objets d'étude de cette thèse. Il est à noter cependant que NEIL1 semble pouvoir exciser la 8-oxoG à proximité de l'extrémité 3' d'une cassure simple brin d'ADN, ce qui n'est pas le cas de hOGG1 [275]. Dans les prochaines parties nous allons donc nous concentrer sur ces enzymes, en décrivant leurs structures, leurs mécanismes, et l'intérêt thérapeutique de leurs ciblages.

III. hOGG1

III.1. Les protéines 8-oxoG-ADN glycosylases

III.1.1. Identification de Ogg1 et de son homologue humain

La première ADN glycosylase identifiée prenant en charge les lésions 8-oxoG, mais également FapyG et FapyA, est la protéine Fpg (pour « Formamidopyrimidine-DNA N-glycosylase » ou FapyG-DNA glycosylase) mise en évidence chez *E. coli* [276]. Cette enzyme fait partie du système GO que nous avons déjà évoqué et qui prend majoritairement en charge la lésion 8-oxoG. Chez l'Homme, c'est la protéine hOGG1 (pour « OxoGuanine-DNA Glycosylase 1 ») qui est l'homologue fonctionnel de Fpg. Cette enzyme a tout d'abord été identifiée chez la levure *S. cerevisiae* (yOgg1) et son gène a été cloné par complémentation fonctionnelle d'une souche bactérienne *E. coli fpg⁻/mutY⁻* avec une banque d'ADN génomique de levure [277,278]. La surexpression de yOGG1 dans des souches d'*E. coli* déficientes en Fpg et MutY annihile le caractère mutateur G:C vers T:A de ces souches bactériennes [279].

Le gène OGG1 chez *S. cerevisiae* est situé sur le chromosome XIII et code pour une protéine basique de 43 kDa constituée de 376 acides aminés. Cette enzyme est capable d'exciser 8-oxoG et Me-FapyG d'un double brin d'ADN, et catalyse le clivage en 3' d'un site AP via une réaction de β-élimination. En revanche, bien que ce soit des homologues fonctionnels, les protéines Fpg de *E. coli* et *y*OGG1 ne présentent aucune homologie structurale (voir paragraphe II.2). En effet, si Fpg fait partie de la superfamille structurale H2TH, yOGG1 est une ADN glycosylase appartenant à la famille HhH-GPD, qui partage plutôt un ancêtre commun avec Nth/Endo III de *E. coli* [277]. Il existe également des différences de substrat entre ces deux enzymes, notamment les FapyA qui peuvent être excisées par Fpg mais pas par *y*OGG1.

Le clonage chez la levure a été une première étape essentielle dans la découverte de ses homologues chez d'autres mammifères. Ainsi, un an seulement après son identification chez *S. cerevisiae*, l'homologue humain a été identifié et cloné, avec 60 % d'identité de séquence entre ces deux protéines [280]. Des homologues d'autres mammifères comme la souris par exemple ont également été identifiés [281]. De plus, ce gène a par la suite été identifié chez certaines plantes, insectes et bactéries [282–284]. Comme nous l'avons évoqué, on retrouve chez toutes ces enzymes un domaine HhH, et des domaines de fixation à l'ADN très conservés (Figure 22 B).

III.1.2. Régulation de l'expression de hOGG1

Il faut noter que chez l'Homme, l'analyse de la séquence génomique a révélé un épissage alternatif qui est à l'origine de plusieurs isoformes de hOGG1 notées 1a/1b/1c/2a/2b/2c/2d/2e [285]. Parmi ces isoformes, la forme 1a notée α-OGG1 est une protéine de 345 acides aminés (39 kDa) et la seule se situant dans le noyau, ce qui en fait donc l'unique protéine nucléaire hOGG1 fonctionnelle [286]. Du fait d'un domaine C-terminal non conservé, il semble que la forme 2a, ou β-OGG1 (une protéine de 424 acides aminés soit 47 kDa et l'enzyme mitochondriale majoritaire), déjà exprimée et purifiée, ne présente pas d'activité ADN glycosylase. En revanche la forme 1b semble disposer des mêmes activités que la forme α-OGG1, et semble donc être la responsable de la réparation de la 8-oxoG dans les mitochondries, bien que la forme 1a puisse également tenir ce rôle [285,287]. Les cinq autres isoformes n'ont pas encore été isolées et leurs fonctions restent donc inconnues.

L'expression du gène OGG1 semble globalement ubiquitaire, dépendant d'un promoteur riche en îlots CpG (et dépourvu de boîte TATA ou CAAT) [288]. En revanche, d'autres paramètres peuvent augmenter son expression comme des stress oxydatifs [23], ou la présence d'un agent alkylant, le méthylméthane-

sulfonate (MMS) [289] tandis que le cadmium peut la réduire [290]. Lors du cycle cellulaire, l'expression peut également varier, avec un pic durant les phases S et G2 [291].

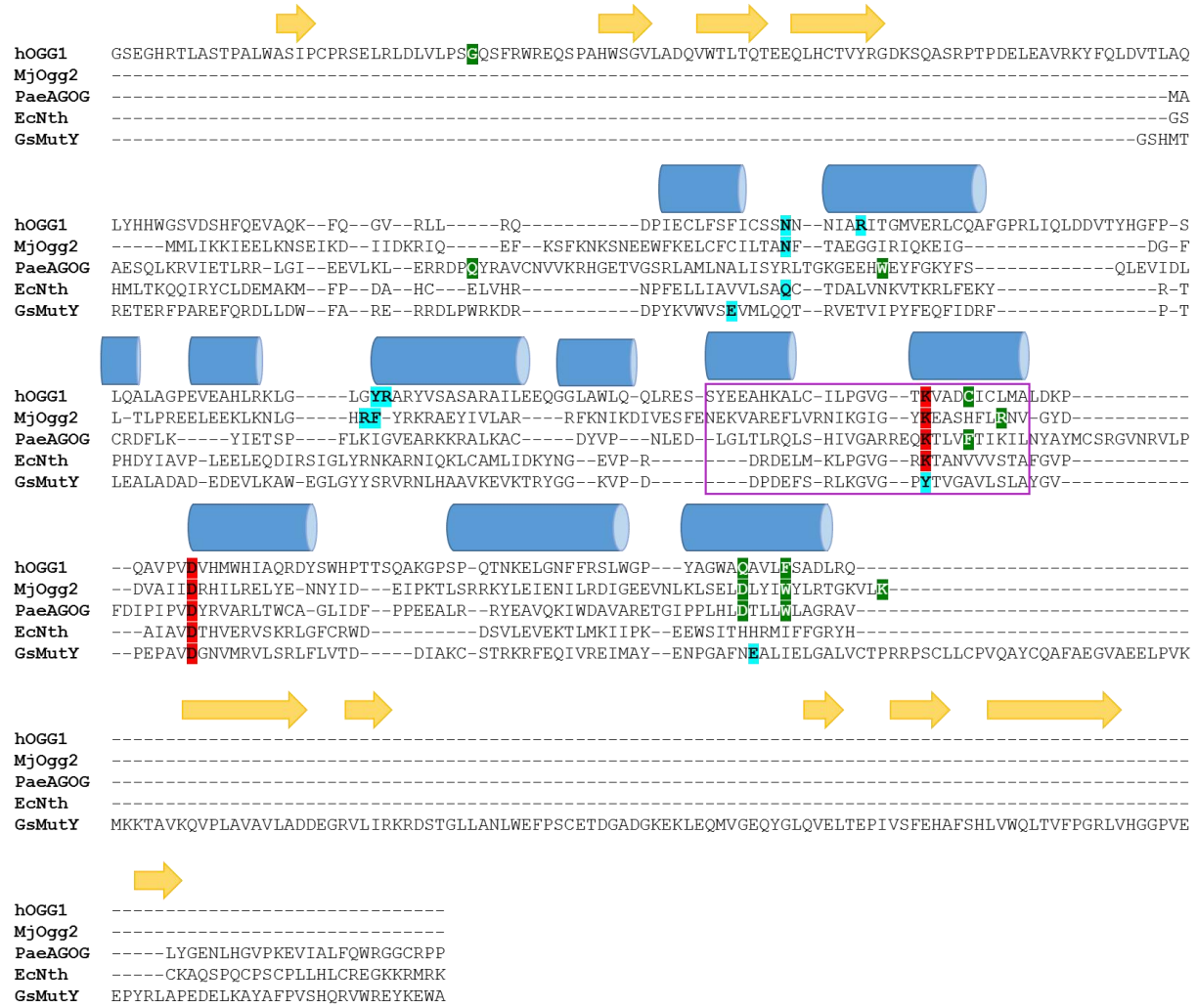
III.1.3. OGG1 chez les mammifères et autres sous-familles

Comme nous l'avons évoqué, OGG1 est présente dans différentes espèces, comme la souris, la drosophile ou des bactéries. Toutes ces enzymes possèdent globalement les mêmes substrats, Me-FapyG/C, 8-oxoG/C, 8-oxoA/C et FapyG/C, avec pourtant une homologie de séquence de seulement 30 à 40 % par rapport à yOGG1 [292]. De manière générale, ces enzymes peuvent exciser de façon préférentielle les purines oxydées, mais pas les pyrimidines oxydées. De plus α-OGG1 ne peut pas exciser les sous-produits de l'oxydation de 8-oxoG que sont Sp et Gh à l'inverse de Fpg et NEIL1 [293]. Il faut également rajouter que l'activité *N*-glycosylase de ces enzymes est fortement dépendante de la présence d'une cytosine en face des bases endommagées. hOGG1 peut même exciser le FapyA lorsque celui-ci est opposé à une cytosine. L'enzyme humaine est bifonctionelle, mais son activité AP lyase de type β-élimination est découplée et plutôt peu efficace [23]. Il existe deux autres sous familles de protéines s'apparentant à OGG1, les protéines OGG2 et AGOG, qui correspondent globalement aux procaryotes et archées, respectivement.

Le terme OGG2 a tout d'abord été attribué à une protéine isolée provenant de *S. cerevisiae*, et dont le gène correspond en réalité à celui de NTG1 (homologue de NTHL1), codant donc pour cette enzyme de la superfamille HhH incapable d'exciser la 8-oxoG [294]. Ce terme a également été utilisé pour les protéines avec des activités semblables à OGG1, mais présentant une identité de séquence très faible. Par exemple, elle a été identifiée chez des bactéries comme *Methanocaldococcus janischii* ou des archées comme *Thermogata maritima* [33]. Si on exclue le corps catalytique, les membres de cette famille possèdent une structure relativement différente de OGG1 et sont dépourvues du domaine N-terminal de OGG1. Pour MjOgg2, le domaine C-terminal se terminant toujours par une lysine joue un rôle dans la reconnaissance de la 8-oxoG [295]. En revanche, bien que les résidus soient différents, les interactions avec la base en position extra-hélicale semblent similaires pour les deux sous-familles, à la différence près que OGG2 interagit moins avec la base opposée. Ceci peut expliquer pourquoi cette enzyme n'a pas autant de spécificité pour la base opposée au dommage [296].

La dernière sous-famille de protéines correspond aux AGOG, identifiées tout d'abord chez l'archée hyperthermophile *Pyrobaculum aerophylum* [297]. Depuis d'autres protéines de cette famille ont été identifiées chez d'autres archées, comme *Thermococcus kodakarensis* [298] ou *Thermococcus gammatolerans* [299], et également *Pyrococcus abyssi*. Ces enzymes ne disposent de presque aucune homologie de séquence avec les enzymes OGG1, et peuvent exciser la 8-oxoG dans l'ADN simple ou double brin quelle que soit la base opposée. Bien que les séquences soient différentes des deux autres sous-familles, un alignement structural montre que, comme OGG2, les AGOG ne disposent pas du domaine N-terminal additionnel de OGG1 mais possèdent une structure 3D proche de celle de OGG2 et du corps catalytique de OGG1 (Figure 22 A). Elles disposent également du domaine HhH-GPD caractéristique, superposable avec les autres ADN glycosylases de cette superfamille [300]. Malgré ces différences, il est supposé que les mécanismes catalytiques de ces trois sous-familles soient très similaires.

A



B

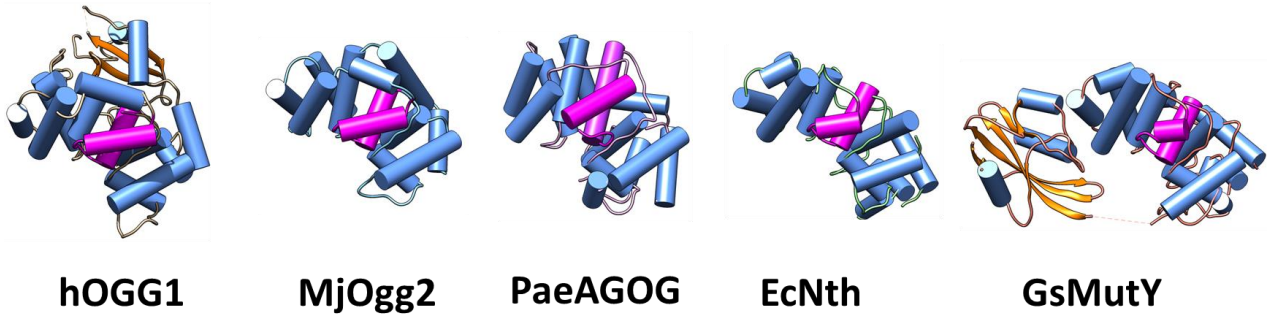


Figure 22 : ADN glycosysales de la superfamille structurale HhH – A) Alignement de séquence basé sur la structure des protéines. Les résidus catalytiques sont indiqués sur fond rouge, les résidus d'intercalation en cyan, et ceux interagissant avec la base opposée en vert. B) Les hélices α sont colorées en bleu, les brins β en orange et le domaine HhH en magenta. Enzyme humaine hOGG1 (PDB : 1EBM), *Methanocaldococcus janischii* MjOgg2 (PDB : 3KNT), *Pyrobaculum aerophylum* PaeAGOG (PDB : 1XQP), *Escherichia coli* EcNTH (PDB : 1P59) et *Geobacillus stearothermophilus* GsMutY (PDB : 1RRQ).

III.2. Considérations structurales et fonctionnelles de hOGG1

Comme nous l'avons déjà évoqué plusieurs fois, hOGG1 possède un motif de fixation non spécifique à la séquence de l'ADN (domaine HhH) suivi du motif GPD. hOGG1 possède également deux domaines en hélice α partagés par les autres membres de la famille HhH ainsi qu'un feuillet β présent uniquement chez AlkA. Des structures 3D de cette enzyme sont disponibles avec ou sans substrat ADN, avec des intermédiaires de base de Schiff ou encore avec des produits de réaction, et permettent donc de donner quelques éléments de relation structure-fonction afin de comprendre les différents mécanismes moléculaires qui permettent à hOGG1 d'exciser la 8-oxoG dans l'ADN.

III.2.1. Déplacement sur l'ADN

Tout d'abord, afin de pouvoir détecter et reconnaitre la lésion 8-oxoG parmi des millions de bases normales, la protéine hOGG1, comme les autres ADN glycosylases, diffuse sur l'ADN. Il existe deux options à l'enzyme pour trouver la 8-oxoG dans l'ADN [301] :

- Le « Sliding » qui consiste à diffuser le long de l'ADN en gardant un contact constant (diffusion 1D) ;

- Le « Hoping » qui consiste à interagir avec l'ADN de manière discontinue, en « sautant » de séquence en séquence sur l'ADN (diffusion 3D).

Il faut noter que la diffusion ne consomme pas d'énergie biochimique, mais est plutôt dirigée thermiquement [301]. Les premières expériences effectuées sur hOGG1 semblent indiquer que hOGG1 est capable de glisser très rapidement le long d'un ADN linéaire, pouvant scanner 1 kb en 0,1 secondes dans les conditions physiologiques [302]. Ce mécanisme de glissement dirigé est également accompagné d'une rotation constante permettant de placer le domaine d'interaction à l'ADN en contact avec ce dernier de manière continue [303]. En revanche, il est peu probable que hOGG1 puisse diffuser de manière continue sur l'ADN *in vivo*, des enzymes déjà liées à l'ADN pouvant certainement constituer des obstacles physiques sur sa route. De plus, il a été observé que lors de sa diffusion le long de l'ADN, sa durée de résidence sur l'ADN est relativement courte et est associée à des sauts microscopiques inter ou intra-brin sur l'ADN. Il est donc probable que la recherche du dommage par hOGG1 sur l'ADN soit une combinaison des mécanismes de « sliding » entrecoupé de « hoping » lui permettant de contourner les obstacles sur l'ADN (Figure 23) [304]. La diffusion de la protéine sur l'ADN dépend donc de la concentration en enzyme, de l'environnement moléculaire et de la compaction de l'ADN. Ce profil de diffusion est cohérent avec les observations obtenues sur différentes ADN glycosylases [305].

Comme nous pouvons nous y attendre, la diffusion des enzymes de réparation sur l'ADN chromatinien est beaucoup plus contrainte que celle sur un ADN linéaire, relaxé et nu [306], rendant la recherche du dommage aléatoire et ce d'autant plus que la 8-oxoG n'induit aucune distorsion locale dans l'ADN. Ainsi l'efficacité de hOGG1 est réduite sur un nucléosome reconstitué par rapport à un ADN linéaire, l'excision de 8-oxoG étant également altérée en présence de l'histone H1 [307]. La présence d'un complexe de remodelage de la chromatine permet de retirer cette réduction d'activité, et le relâchement de l'ADN autour des histones facilite l'activité de hOGG1 [308]. Par exemple, il a été montré que la réparation des dommages oxydés était favorisée dans les régions ouvertes de la chromatine [309]. Ainsi, l'activité de hOGG1 pourrait nécessiter la présence de protéines ADN chaperonnes, mais plus d'études seront nécessaires pour appréhender dans son ensemble ces aspects.

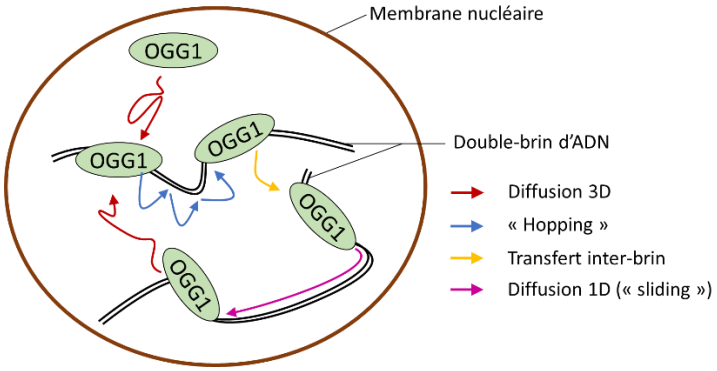


Figure 23 : Schéma représentant les différents modes de diffusion de hOGG1 (inspiré de D'Augustin *et al.*, 2020 [301])

III.2.2. Reconnaissance de la 8-oxoG

Suite à sa diffusion sur l'ADN, hOGG1 va finir par rencontrer la 8-oxoG et établir un complexe enzyme/substrat spécifique mobilisant de nombreux résidus d'acides aminés de l'enzyme et conduisant à son extrusion de la double-hélice et à sa stabilisation dans la poche du site actif de l'enzyme.

De nombreux travaux ont étudié finement la structure cristallographique des complexes hOGG1/ADN et contribué à mieux comprendre les bases moléculaires de la reconnaissance de la 8-oxoG [310–314].Ces études structure/fonction ont permis d'établir que l'enzyme extrude la base endommagée en induisant une torsion de l'ADN au niveau du dommage. La première hypothèse, pour expliquer la spécificité de reconnaissance de la 8-oxoG par hOGG1 suggérait une instabilité supérieure de la paire 8-oxoG:C par rapport à la paire G:C [315]. Cependant, la torsion de l'ADN par hOGG1 permet également de placer une guanine en position extra-hélicale en dehors du site actif (exo site) [316,317]. La torsion simple de l'ADN ne peut donc pas expliquer à elle seule la reconnaissance de 8-oxoG.

Des études de dynamique moléculaire ont en revanche suggéré que la reconnaissance de la base oxydée se faisait en plusieurs étapes transitoires : i) La formation d'un « complexe d'interrogation » (IC) correspondant à la liaison entre la protéine et l'ADN sans dommage ; ii) La formation d'un « complexe de rencontre » (RC) avec la reconnaissance de la 8-oxoG dans l'ADN; iii) un « complexe de reconnaissance » de la lésion (LRC) avec l'extrusion de la base endommagée [318,319]. Il est donc proposé que la sélectivité de hOGG1 pour la 8-oxoG par rapport à la G s'effectue après l'extrusion de la base ciblée. Les bases structurales de la reconnaissance de la 8-oxoG ont été décryptées grâce à l'obtention de la structure d'un complexe abortif entre le mutant K249Q catalytiquement inactif et un fragment d'ADN contenant une 8-oxoG unique [310,311].

- **Interaction avec le squelette phosphate-sucre de l'ADN**

La structure cristallographique indique que la surface d'interaction entre hOGG1 et l'ADN endommagé représente une surface assez faible seulement 2,268 Å² [311]. Alors que la plupart des protéines interagissant avec l'ADN possèdent une région de contact composée majoritairement de lysines et d'arginines, hOGG1 présente un nombre limité de résidus interagissant directement avec les phosphates du squelette de l'ADN. Le résultat de ce complexe est particulier en cela que hOGG1 contient plusieurs hélices α avec un domaine N-terminal dirigé vers l'ADN, ce qui promeut des interactions de type dipôle électrostatique plutôt que des ponts salins entre l'enzyme et son substrat ADN [23,320]. On peut également noter que hOGG1 établit exclusivement des contacts avec le squelette du brin endommagé. Seule la cytosine opposée à la 8-oxoG est contactée dans le brin complémentaire.

Le domaine HhH interagit avec l'ADN en 3' de la lésion et semble impliqué dans le positionnement de la lésion dans la poche du site actif (Figure 24) [23]. En effet, l'une des deux hélices du motif contient les résidus Val250 et Gln249 en contact direct avec le groupement phosphate p_{-2}, et le résidu Gly245 qui forme une liaison hydrogène avec le groupement phosphate p_{-3} [320]. La protéine va également interagir avec les autres phosphates p_{-1}, p_0 et p_1 autour de la 8-oxoG. C'est le cas avec le triplet bien conservé NNN (Asn149-Asn150-Asn151) dont les résidus 150 et 151 vont stabiliser la structure du complexe en formant des liaisons hydrogène avec le phosphate p_{-1} [311]. Le résidu His270 forme également une liaison H avec le phosphate p_0 en 5' de la lésion. Enfin, il faut noter qu'un ion calcium Ca^{2+}, présent dans le milieu de cristallisation, permet également de stabiliser le complexe, en interagissant directement ou *via* des molécules d'eau, avec les phosphates p_1 et p_{-1}. Dans les conditions physiologiques, l'ion Ca^{2+} pourrait être remplacé par un ion magnésium Mg^{2+}.

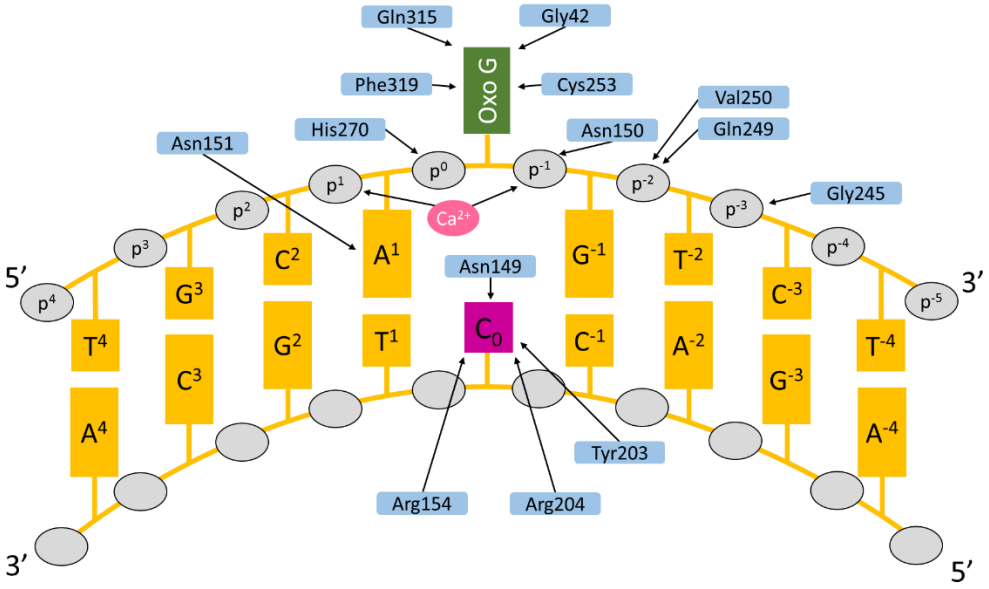


Figure 24 : Résidus de hOGG1 interagissant avec un ADN contenant la 8-oxoG (inspiré de Koval *et al.*, 2014 [320])

- **Interaction de hOGG1 avec la base opposée au dommage**

La fixation à l'ADN endommagé de hOGG1 dans le complexe LRC induit une forte courbure de l'axe de la double-hélice centrée sur la 8-oxoG (Figure 25). Suite, ou de façon concomitante, à l'extrusion de la 8-oxoG, la Tyr203 s'intercale par le petit sillon en 5' de la cytosine C_0 opposée à la 8-oxoG tandis que les résidus Arg154 et Arg204 forment des liaisons hydrogène avec les N1, N3 et O2 de la cytosine C_0, en s'intercalant par le petit sillon de l'ADN. Enfin, l'Asn149, incluse dans la séquence NNN strictement conservée, comble le vide laissé par la 8-oxoG extrudée, et forme une liaison hydrogène entre le carbonyle de son amide et le NH_2 exo-cyclique de la cytosine C_0. L'ensemble de ces interactions avec C_0 permet à l'enzyme de spécifier la pyrimidine opposée au dommage (expliquant la spécificité de hOGG1 pour la cytosine) et permet de la maintenir intra-hélicale empêchant ainsi l'hélice de s'écraser sur elle-même localement [311]. Ces observations structurales et leurs interprétations ont été validées fonctionnellement. Ainsi, le mutant R154H supprime en partie la sélectivité de hOGG1 pour la base opposée au dommage [321].

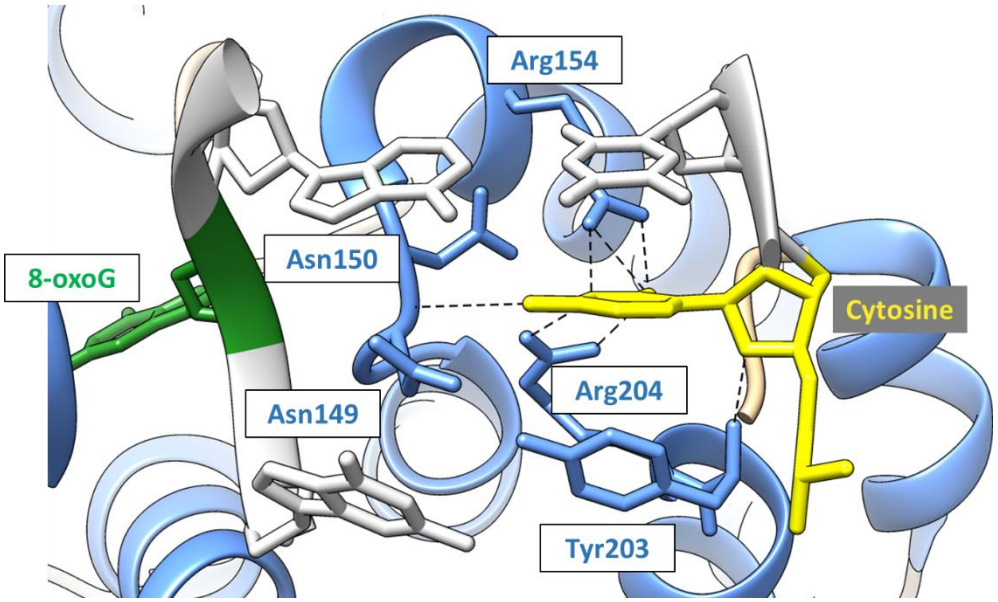


Figure 25 : Interaction de hOGG1 avec la cytosine opposée au dommage (PDB 1EBM) - Les hélices α sont colorées en bleu, les brins β en orange, le domaine HhH en magenta, l'ADN en gris clair, la 8-oxoG en vert et la cytosine en jaune. Les liaisons hydrogène sont indiquées en noir.

- **Reconnaissance de la 8-oxoG**

La principale interaction entre la 8-oxoG extra-hélicale et hOGG1 dans la poche du site actif est la liaison hydrogène formée entre le proton en N7 de la 8-oxoG et le résidu Gln42, le seul résidu du feuillet β conservé chez AlkA et contribuant à l'interface hOGG1/ADN [322]. Comme le N7 d'une G est déprotoné, cela suggère que le mécanisme de discrimination entre la 8-oxoG et la G réside dans la possibilité de faire cette interaction avec la base lésée [323]. Cette discrimination peut aussi s'opérer avant l'extrusion de la 8-oxoG dans le complexe RC par la reconnaissance dans le grand sillon du N7 protoné de la 8-oxoG.

D'autres résidus sont également impliqués dans la reconnaissance du dommage (Figure 26). Les acides aminés Phe319 et Cys253 se lient de chaque côté de la 8-oxoG (par « π-stacking » et liaison hydrogène, respectivement), bloquant celle-ci dans le site actif. Le résidu Gln315 forme également plusieurs liaisons avec la 8-oxoG, à travers des liaisons hydrogène avec son groupement carbonyle, ou en utilisant son groupement amide *via* des molécules d'eau pour interagir avec le O6 de la 8-oxoG [320].

L'étude comparative des structures cristallines de hOGG1 libre et hOGG1/ADN permet d'identifier d'autres résidus importants [310]. Pour le complexe hOGG1/ADN, la formation d'une liaison hydrogène entre les résidus Pro266 et Gln315 permet la formation d'une poche active adéquate pour l'insertion de la 8-oxoG. Le résidu Histidine 270, forme un « π-stacking » avec les résidus Phe319 et Gln315 et interagit avec Asp322 dans l'enzyme libre, empêchant la potentielle entrée d'une base endommagée dans le site actif. Or His270 peut former une liaison hydrogène avec la 8-oxoG et « libérer » Phe319, ce qui permet ainsi l'entrée de la base endommagée dans le site actif. On peut donc considérer ces quatre résidus comme formant une entité permettant d'obtenir une conformation ouverte ou fermée du site actif [310]. Le résidu catalytique Lys249 (impliqué dans l'attaque nucléophile du C1' de la 8-oxoG) interagit également avec la Cys253 dans le complexe hOGG1/ADN, et semble essentiel pour le bon alignement final de la 8-oxoG [324,325]. Dans la structure de hOGG1 libre, Lys249 forme des liaisons H avec Asp268 et Asn149 qui disparaissent dans le complexe hOGG1/ADN.

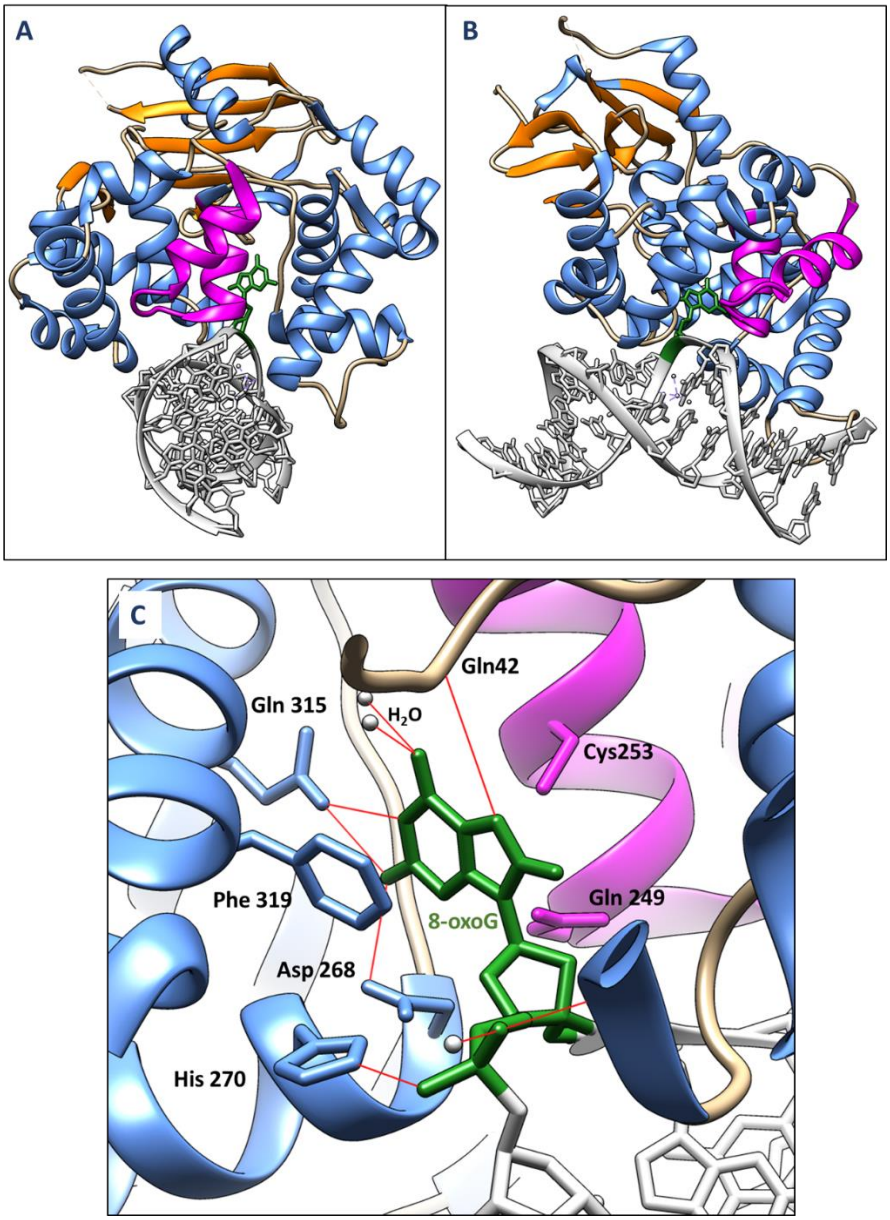


Figure 26 : Structure cristallographique d'un complexe entre hOGG1 K249Q et un ADN contenant 8-oxoG (PDB 1EBM) - A/B) Structure cristallographique de hOGG1 K249K complexée à un ADN avec une 8-oxoG extra-hélicale. Les hélices α sont colorées en bleu, les brins β en orange, le domaine HhH en magenta, l'ADN en gris clair et la 8-oxoG en vert. Les liaisons hydrogène sont indiquées en rouge, et les molécules d'eau sont des sphères blanches. C) Interaction de la 8-oxoG avec les résidus de la poche du site actif de hOGG1.

III.2.3. Mécanismes catalytiques de hOGG1

hOGG1 possède à la fois une activité ADN glycosylase et β-lyase, hydrolysant le lien *N*-glycosidique de la base endommagée, suivi par un clivage par β-élimination en 3' du lien phosphodiester du site abasique résultant. Le premier mécanisme suggéré a tout d'abord été l'attaque du C1' du ribose de la base endommagée par le groupe ε-amino déprotoné du résidu Lys249, proposition soutenue par le fait que la mutation K249Q de hOGG1 abolit complètement les activités catalytiques de l'enzyme [310] (Figure 27). Cette attaque résulte en la formation d'un anion 8-oxoG$^-$ (Figure 27 **2**), qui va en retour de nouveau déprotoner le ε-NH_2 de Lys249, formant un intermédiaire aminal (Figure 27 **4**). Cet intermédiaire va par la suite former une base de Schiff, qui va transférer son proton à 8-oxoG$^-$ et promouvoir l'élimination du groupement phosphate en 3' par β-élimination. Un nouvel intermédiaire est formé, avec une base de Schiff α,β-insaturée et chargée positivement. Après hydrolyse de cette base de Schiff, l'ADN est libéré avec un groupement 4-hydroxy-2,3-

penténal-1 en 3'. Dans ce scénario, le résidu Asp268 du motif GPD est requis pour activer Lys249 en assurant la déprotonation de son groupe ε-amino [311].

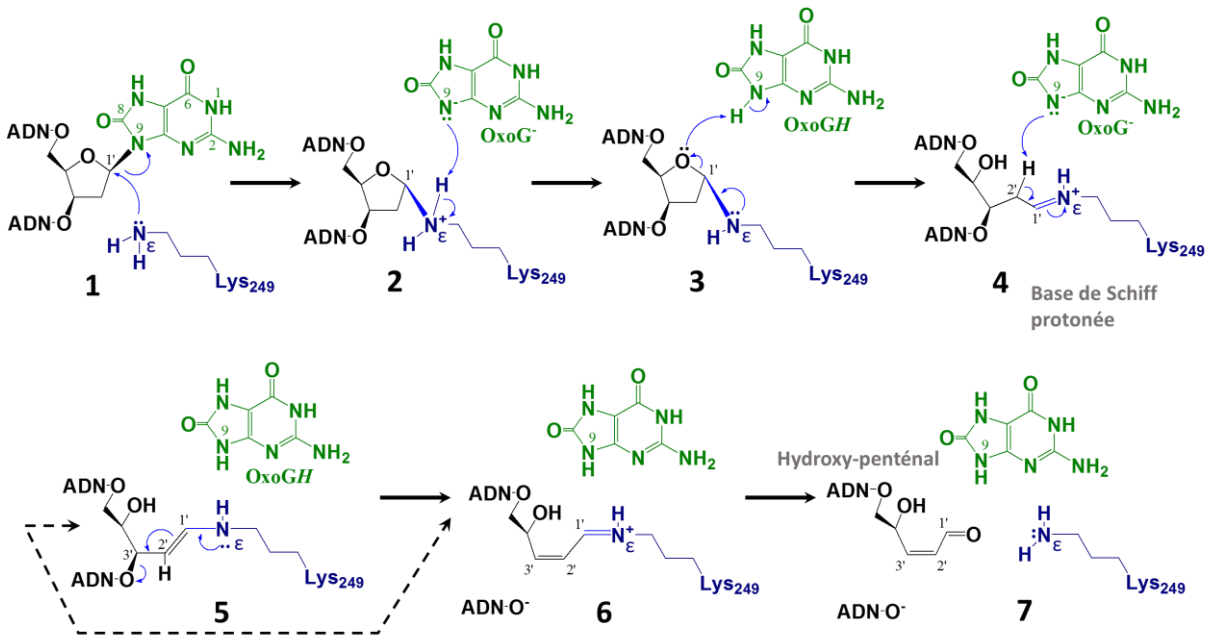


Figure 27 : Mécanisme suggéré de l'activité catalytique de hOGG1 sur la 8-oxoG par Fromme *et al.*, 2003 [312]

Bien que le résidu Asp268 soit essentiel pour l'activité catalytique (un mutant de hOGG1 D268A ne montre aucune activité, mais conserve la reconnaissance), l'hypothèse de la déprotonation de Lys249 par Asp268 est contredite par la basicité trop faible de l'acide aspartique, et par une distance trop élevée entre les deux résidus [320]. Un contact est observé entre Asp268 et la Lys249 uniquement pour hOGG1 libre. Le groupement amino de la Lys249 forme en effet une liaison H avec la Cys253, elle-même interagissant avec la 8-oxoG par une interaction de type Van Der Waals, la mettant dans une conformation non propice à sa déprotonation par Asp268 [23]. Une double mutation K249C/C253K permet toujours d'exciser la 8-oxoG mais supprime toute activité AP lyase, tandis qu'une triple mutation K249C/C253K/D268A supprime toute excision de la 8-oxoG [324]. Finalement, il semblerait que l'impact d'Asp268 soit plutôt sur la charge du O4' du désoxyribose de 8-oxoG [313].

Les activités ADN glycosylase et AP lyase de hOGG1 sont en réalité très fortement découplées, avec une activité AP lyase très faible [326]. En effet, des études semblent indiquer que OGG1 exercerait ces deux activités à des vitesses différentes [327]. Ce découplage pourrait donc suggérer que hOGG1 se dissocie du site abasique avant de s'y réassocier et de le cliver en 3' par un mécanisme de β-élimination. Ainsi, le mécanisme de hOGG1 se rapprocherait de celui d'une ADN glycosylase monofonctionelle, avec APE1 venant cliver par la suite le site AP en 5' [328]. Ainsi, il a même été proposé que le mécanisme catalytique du clivage du lien *N*-glycosidique se fasse, comme pour les autres ADN glycosylases, via une molécule d'eau [324]. Ce mécanisme monofonctionnel est soutenu par des calculs de mécanique quantique qui suggèrent que la Lys249 est également responsable de l'activation nucléophile d'une molécule d'eau, permettant à Asp268 de stabiliser l'intermédiaire oxocarbénium durant toute la réaction [329] comme présenté sur la Figure 28.

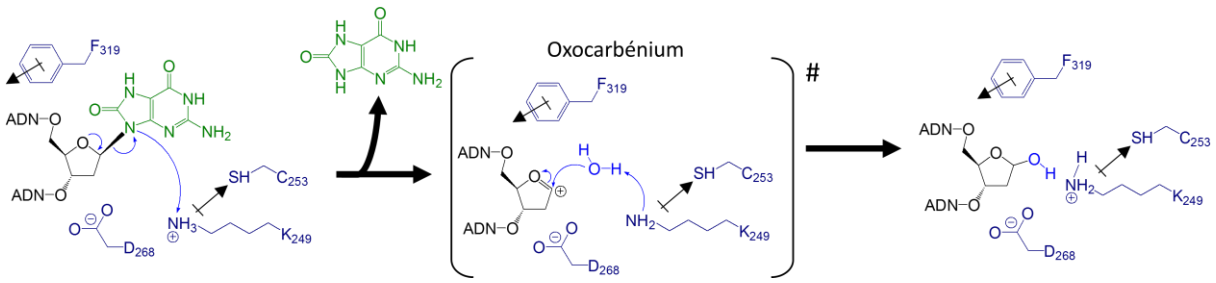


Figure 28 : **Mécanisme monofonctionnel suggéré de l'activité catalytique de hOGG1 sur la 8-oxoG d'après Boiteux *et al.*, 2017 [23]**

III.3. Fonctions biologiques principales et impact thérapeutique de l'inhibition de hOGG1

III.3.1. hOGG1, une cible thérapeutique pertinente

Comme expliqué précédemment, hOGG1 prend en charge certains dommages oxydatifs tels que le produit majeur d'oxydation des purines, la 8-oxoG. Elle permet donc d'éviter les effets délétères ou mutagènes de ces dommages. Elle pourrait également posséder d'autres fonctions biologiques comme le maintien de la taille des télomères grâce à la réparation des guanines oxydées au niveau des séquences des télomères, ce qui a été observé pour yOGG1 [330]. Une autre fonction de hOGG1 pourrait être de jouer un rôle dans la régulation du métabolisme énergétique des cellules. On observe par exemple que des souris soumises à un régime très riche en calories surexpriment mOGG1 en réponse à un stress oxydant [331]. Enfin hOGG1 semble avoir un effet protecteur contre le vieillissement de certaines cellules, ce qui a été observé notamment pour les cellules des globes oculaires [332], et même contre certains accidents cardiovasculaires [333]. Cependant, comme toutes les enzymes de réparation de l'ADN que nous avons présenté, il existe des contextes dans lesquels l'inhibition de hOGG1 peut avoir un intérêt thérapeutique.

Comme tout système réparant l'ADN, hOGG1 a de forts liens avec la cancérogénèse. Il est ainsi montré que l'inactivation des gènes de OGG1/MUTYH/MTH1 chez la souris, responsables de la prise en charge de la 8-oxoG, provoque des mutations spontanées G:C vers T:A et T:A vers G:C [23]. Comme nous l'avons déjà évoqué, l'inactivation double de OGG1 et de MTH1 induit également l'apparition de tumeurs, tandis que l'inactivation d'un seul de ces gènes ne présente pas de phénotype particulièrement sévère [192]. Etrangement, pour certaines études, c'est l'inactivation seule de OGG1 qui induit la formation de tumeurs, tandis que la double inactivation OGG1/MTH1 supprime cet effet [334]. Il a également été montré que certains polymorphismes de hOGG1 étaient à l'origine de l'apparition de cancers gastriques, des reins, ou des poumons et potentiellement des seins [335–340].

Dans le contexte où les traitements anticancéreux font appel à des agents génotoxiques, les systèmes de réparation de l'ADN peuvent induire des résistances aux traitements [132]. Ainsi, la surexpression de OGG1 peut avoir un effet protecteur contre les agents alkylants [196,341–343]. De plus l'inactivation d'OGG1 permet également de sensibiliser des cellules à la bléomycine [344]. Cependant les résultats contradictoires de l'influence de hOGG1 sur l'apparition des cancers, ainsi que le phénotype viable du KO chez les souris, ont initialement modéré l'intérêt de hOGG1 en tant que cible thérapeutique, et peu d'études sur son ciblage thérapeutique ont été envisagées. En revanche, depuis l'apparition de la notion de létalité synthétique et son utilisation avec les inhibiteurs de PARP, différentes études montrent l'intérêt potentiel du ciblage pharmacologique de hOGG1 pour le développement de stratégies anti-cancéreuses innovantes. De plus, la forte présence de mutations dans certains cancers associés en particulier avec le polymorphisme S326C de hOGG1 [345,346], pourrait être une opportunité de découvrir des relations de létalité synthétique avec hOGG1.

Par exemple, une relation de létalité synthétique entre les ADN glycosylases hOGG1 et hNEIL1 et la voie de la thymidilate synthase (TS) a été proposée. Ainsi, l'inhibition de hOGG1 ou hNEIL1 pourrait augmenter l'effet des inhibiteurs de TS déjà existants [347]. Une autre étude montre que l'utilisation du BER par des cellules cancéreuses avec une protéine RAS mutée permet d'éviter les dommages oxydatifs empêchant la prolifération [348]. Dans ce contexte cancéreux, inhiber hOGG1 pourrait donc être une stratégie intéressante. Une stratégie similaire a déjà été validée en ciblant MTH1 [349]. Des études ont également montré que l'inhibition de hOGG1 sur des cellules de neuroblastomes pouvait augmenter les dégâts oxydatifs des radiations ionisantes dans celles-ci [350]. Cependant l'effet reste limité et les cellules survivent au traitement. L'inhibition de hOGG1 et d'un facteur d'activation de celle-ci, la protéine CUX1, permet la sensibilisation de cellules de cancer colorectal aux radiations ionisantes [351].

L'inhibition de hOGG1 peut également trouver un impact thérapeutique dans d'autres domaines que la lutte contre le cancer. En effet, son inactivation pourrait être utile dans le traitement de maladies neurodégénératives ou d'attaques cardiaques [352]. Par exemple pour la maladie d'Huntington, OGG1 semble initier le cycle itératif d'oxydation/excision qui contribue à l'expansion et la répétition du tri-nucléotide CAG dans le gène Huntingtin [353]. Cette expansion va perturber la synthèse de différentes protéines, et est directement reliée à la maladie. Cette maladie, principalement d'origine héréditaire, évolue au fil du temps et de l'expansion du triplet. L'inhibition sélective de hOGG1, et potentiellement de hNEIL1, pourrait donc ralentir, voire empêcher la progression de la maladie [354,355].

hOGG1 joue aussi un rôle dans la régulation de l'inflammation. Ainsi, l'inactivation de OGG1 dans des souris infectées par *H. Pylori* a diminué la sévérité de l'inflammation des muqueuses gastriques [356] ou l'inflammation des poumons des souris en cas de réaction allergique [357]. De même, des souris déficientes en OGG1, résistent mieux aux endotoxines par rapport au groupe contrôle [358]. Afin d'expliquer le rôle de hOGG1 dans l'inflammation, il semblerait que suite à l'excision de la 8-oxoG par hOGG1, la molécule 8-oxoG libre puisse ensuite former un complexe non catalytique avec hOGG1. Ce complexe pourrait ensuite activer les protéines RAS, catalysant la conversion de la guanosine diphosphate (GDP) en guanosine triphosphate (GTP) [359]. L'inactivation d'OGG1 abolit ainsi l'activation des protéines RAS chez la souris [360]. D'autres GTPases comme RAC1 ou RHO semblent aussi pouvoir être activées par le complexe hOGG1/8-oxoG [361,362]. La régulation de hOGG1 pourrait également être une stratégie intéressante pour lutter contre la progression du cancer du poumon, en limitant la réponse inflammatoire [363]. Il existe un autre modèle proposant que la 8-oxoG puisse être une marque épigénétique du stress oxydatif, en particulier dans les régions promotrices de certains gènes et que l'excision et/ou la reconnaissance non-productive de la 8-oxoG par OGG1 dans ces régions particulières (souvent organisées en quadruplex) moduleraient l'expression de ces gènes. Dans cette hypothèse, OGG1 devient un régulateur transcriptionnel des gènes impliqués entre dans la réponse aux stress oxydatifs et inflammatoire [364,365].

III.3.2. Inhibiteurs identifiés de hOGG1

Les premiers inhibiteurs identifiés pour hOGG1 sont des analogues de nucléotides contenus dans des oligonucléotides de synthèse et ont principalement été développés pour les études structurales par cristallographie de complexes stables hOGG1/ADN [82,366,367]. Ces analogues de substrat contenus dans de courts fragments d'ADN sont reconnus avec une haute affinité par la protéine, mais sont non métabolisables par celle-ci. Ils peuvent être classés en deux catégories principales : des ADNs contenant (i) des analogues de sites AP ou (ii) des analogues de nucléotides avec un lien sucre-nucléobase (ou *N*-glycosidique) stable et non clivable par les ADN glycosylases. Il a également été observé que dans le cas de la maladie tyrosinémie héréditaire de type 1 (TH1), une maladie autosomale récessive rare causée par une déficience de l'enzyme

fumarylacétoacétate hydrolase (FAH), l'accumulation de fumarylacétoacétate (Figure 29) pouvait inhiber certaines ADN glycosylases, notamment hOGG1, hNEIL1, hNEIL2 et hNTH1 [368].

Suite au regain d'intérêt pour le ciblage de hOGG1, de nouveaux inhibiteurs ont été identifiés (Figure 29). Tout d'abord un criblage HTS a été mis en place et a conduit à la découverte de nouvelles petites molécules inhibitrices prometteuses contenant une hydrazine, comme le composé 08, avec un IC_{50} submicromolaire [369]. Ce premier criblage a ouvert la voie à d'autres études, qui ont également identifiées et synthétisées de nouveaux inhibiteurs efficaces de hOGG1. Tahara *et al.*, ont par exemple identifié un premier composé, Su01, possédant un IC_{50} de 2 µM. Après une optimisation chimique impressionnante, ils ont réussi à obtenir l'inhibiteur le plus efficace à ce jour : le composé Su0268 [370]. Par la suite ils ont également fusionné ce composé avec un autre fragment inhibiteur de MTH1, obtenant ainsi un composé Su0383 pouvant inhiber spécifiquement OGG1 et MTH1. La présence de 16 µM de Su0383 dans des cellules MCF7 de cancer du sein en présence de H_2O_2 a réduit de 20 % la viabilité des cellules, indiquant la sensibilisation de ces cellules au stress oxydatif [272]. Cependant cette concentration reste assez élevée et risque d'être peu applicable *in vivo*.

De même, un autre criblage a été effectué par Visnes *et al.*, qui ont tout d'abord identifié un composé possédant un IC_{50} de 8,6 µM sur l'activité catalytique, et après une expansion des hits, un composé TH5487 a été identifié (Figure 29). Ce composé, bien accepté par la cellule, a déjà montré des propriétés anti-inflammatoires dans des cellules de poumon de souris [371]. Une structure cristallographique de TH5487 avec hOGG1 montre que ce composé se place dans le site actif de la protéine et bloque l'excision de la 8-oxoG. Bien que très prometteur, ce composé a toutefois montré ses limites *in vivo* car il est absorbé à 99 % par l'albumine sérique ce qui a pour conséquence de limiter drastiquement son effet [372]. Ces différentes études montrent l'intérêt naissant de l'inhibition de hOGG1 et la nécessité de développer de nouveaux composés afin d'étoffer la gamme de molécules disponibles pour de potentielles applications *in cellulo*/*in vivo*.

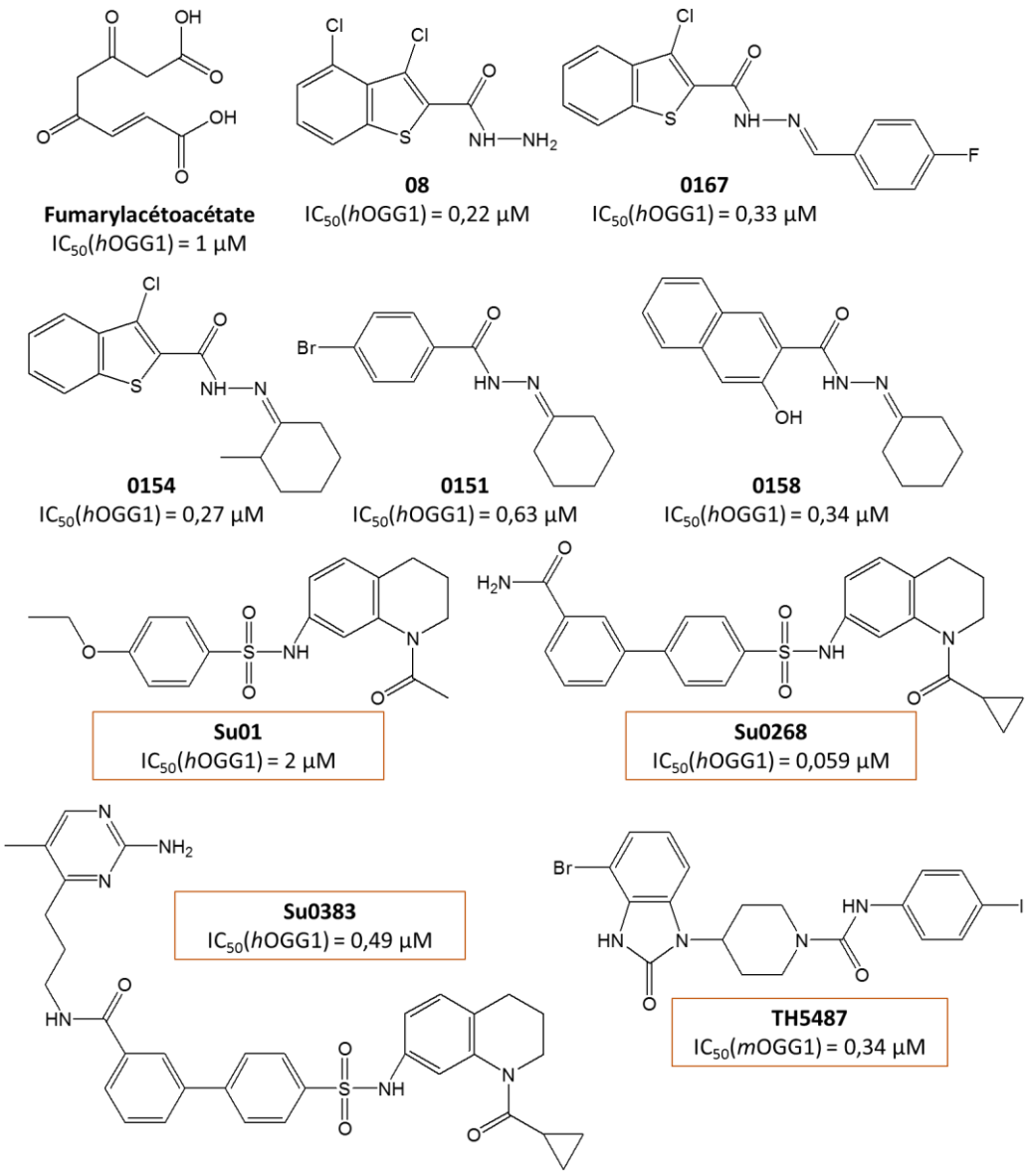


Figure 29 : Inhibiteurs identifiés de hOGG1 - Afin d'illustrer l'état de l'art au début de cette thèse, les inhibiteurs identifiés au cours de celle-ci ont été encadrés en rouge.

IV.hNEIL1

IV.1. La superfamille des ADN glycosylases Fpg/Nei

IV.1.1.Identification de cette famille structurale

Comme nous l'avons évoqué, hNEIL1 fait partie de la superfamille structurale Fpg/Nei. Le premier représentant identifié de cette famille est la protéine Fpg (ou MutM), que nous avons déjà évoqué. Fpg est une ADN glycosylase bifonctionnelle très efficace qui associe une activité ADN *N*-glycosylase, une activité AP lyase qui procède selon un mécanisme de β,δ-élimination et une activité 5'dRP-lyase [276,373]. Elle peut également exciser la 8-oxoG lorsque celle-ci est opposée à une cytosine ou une thymine. Peu après, une autre ADN glycosylase a été identifiée chez *E. coli*, l'endonucléase VIII ou Nei, proche structuralement de Fpg, et pouvant également catalyser une réaction AP lyase par un mécanisme de β,δ-élimination, mais plutôt spécialisée dans l'excision les pyrimidines oxydées [374]. Seules les protéobactéries comme *E. coli* possèdent en même temps Fpg et Nei. Après un séquençage du génome humain, trois homologues structuraux ont été

identifiés et nommés hNEIL1, hNEIL2 et hNEIL3 pour « Nei-like protein » 1 à 3 [375–377]. Ces trois protéines présentent des substrats similaires à ceux de Nei et, contrairement à Fpg, sont incapables d'exciser la 8-oxoG. Plus récemment, une autre enzyme de cette famille a été identifiée chez le virus géant *Mimivirus* MvNei1, une ADN glycosylase très similaire à hNEIL1 [378].

IV.1.2.Caractéristique de cette superfamille

La superfamille Fpg/Nei est définie par la présence de différents motifs structuraux communs et des résidus catalytiques conservés. Tout d'abord, chaque enzyme de cette famille possède un domaine H2TH (pour « helix-two-turn-helix ») (Figure 30) [23]. Elles possèdent également un motif en épingle à cheveux constituée de 2 brins β antiparallèles, contenant ou non un ion zinc (on parle alors de « Zinc finger » ZnF ou de « Zincless finger » ZnLF), dont le rôle est exactement le même pour l'interaction à l'ADN. Par exemple Fpg, Nei, NEIL2 et NEIL3 possèdent un ZnF (consistant en un ion zinc Zn^{2+} coordinné par 4 cystéines ou 3 cystéines et une histidine), tandis que hNEIL1 ou mvNei1 n'en possède pas [379]. Le repliement 3D global du domaine des enzymes de cette famille est également assez similaire, avec deux domaines globulaires reliés entre eux par un domaine charnière flexible. Le domaine C-terminal est riche en hélices α et contient le H2TH et le doigt ZnF/ZnLF. Le domaine N-terminal commence presque toujours par une proline Pro1 (ou Val1 pour NEIL3) et un acide glutamique Glu2 catalytiques, et est riche en structures β (Figure 30) [23]. D'un point de vue mécanistique, l'ensemble des ADN glycosylases de la famille semblent être bifonctionnelles et capables de cliver en 3' et 5' du site AP par une β,δ-élimination qui résulte en la formation d'une lacune de 1 nucléoside.

Malgré des structures similaires, comme nous l'avons évoqué, il existe des différences entre les protéines de cette famille. La plus conséquente d'entre-elles consiste en la différence de substrat entre Fpg et les protéines Nei/NEIL. Nei, capable d'exciser les pyrimidines oxydées, FapyA, FapyG, Sp ou Gh, ne peut exciser la 8-oxoG de manière efficace (elle est 10 à 100 fois moins efficace que Fpg) [380]. De même, NEIL1 peut exciser 5-OHU, 5-OHC, DHU, Sp et Gh, mais elle est très faiblement efficace sur la 8-oxoG. NEIL2 va quant à elle exciser 5-OHU, Sp et Gh quel que soit la base opposée, mais pas 8-oxoG [78,381]. Enfin NEIL3 peut cliver les pyrimidines oxydées ainsi que FapyA, FapyG, Sp et Gh mais pas la 8-oxoG [382]. Comme nous l'avons vu l'homologue fonctionnel de Fpg pour la prise en charge de la 8-oxoG chez l'Homme est hOGG1.

Il est important de noter également que les adduits volumineux de l'ADN peuvent être des substrats de NEIL1. Ainsi, le produit aflotoxin-AFB1 déoxyguanosine à cycle ouvert imidazole, hautement mutagène, est excisé par NEIL1. L'aflotoxin peut induire la formation de carcinomes hépatocellulaires, et les souris ne possédant pas NEIL1 sont plus sensibles à ces tumeurs que les souris contrôles [383]. NEIL1 peut également exciser le mono-adduit du composé psoralen sur la thymine [384].

La différence de substrat entre les enzymes de cette superfamille vient de la présence d'un motif structural non conservé chez Nei/NEIL mais présent chez Fpg ; la boucle LCL (« lesion caping loop ») [385,386]. Cette boucle se caractérise par un grand nombre d'interactions avec les bases altérées extrudée, par exemple le Tyr220 qui forme une liaison hydrogène avec le N7 de la 8-oxoG. Cette boucle est très mal conservée dans les autres ADN glycosylases de la famille, et est essentielle au clivage de la 8-oxoG par Fpg. En effet la mutation de certains résidus de cette boucle supprime la capacité de Fpg à cliver la 8-oxoG [387].

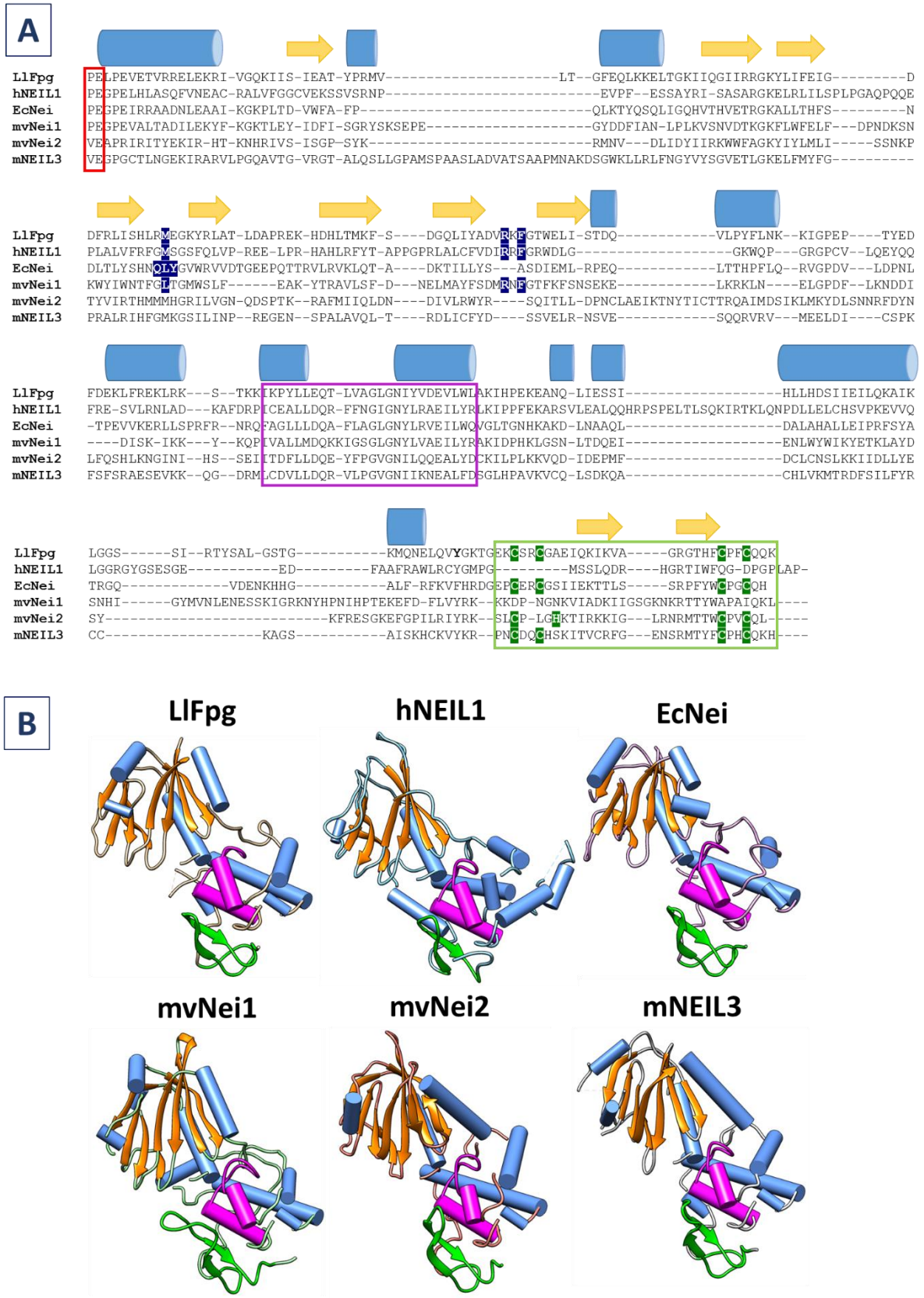


Figure 30 : ADN glycosylase de la superfamille Fpg/Nei - A) Alignement de séquence basé sur la structure des protéines. Les résidus catalytiques sont encadrés en rouge, les résidus d'intercalation sont surlignés en bleu et les cystéines du doigt de zinc sont surlignées en vert. B) Les hélices α sont colorées en bleu, les brins β en orange, le domaine H2TH en magenta et les ZnF ou ZnLF en vert clair. Enzyme humaine hNEIL1 (PDB : 1TDH), *Lactococcus lactis* LlFpg (PDB : 4CIS), *Escherichia coli* EcNei (PDB : 1K3X), *Mimivirus* mvNEI1/2 (PDB : 3A46/4MB7) et murine mNEIL3 (PDB : 3W0F).

IV.2. Considérations structurales et excision des bases endommagées

Les bases moléculaires et mécanistiques de la reconnaissance et de l'excision des bases endommagées ont été élucidées en partie grâce à l'étude par cristallographie des complexes abortifs entre ces protéines et des petits fragments d'ADN contenant des analogues de substrat non clivables, comme des carba-nucléotides [388,389]. Il est à noter que tout comme hOGG1, les ADN glycosylases de la famille Fpg/Nei suivent le même processus de reconnaissance et d'extrusion de la lésion, avec la formation de trois complexes successifs : le complexe d'interrogation IC, le complexe de rencontre RC, et le complexe de la reconnaissance de la lésion LRC [23].

IV.2.1.Reconnaissance du dommage

Concrètement le mécanisme d'action des enzymes Fpg/Nei est similaire, et consiste en une forte torsion de l'ADN centrée sur le nucléoside endommagé qui va permettre de l'extraire de la double-hélice et de le stabiliser en position extra-hélicale dans la poche du site actif [23]. Cette extrusion de la base est soutenue par l'intercalation d'une triade de résidus très bien conservés dans le petit sillon, qui vont combler le vide laissé par la lésion et former des liaisons avec la base opposée. Pour hNEIL1, cette triade comporte les résidus Met81, Arg118 et Phe120, des résidus situés dans le domaine N-terminal [379] (Figure 31).

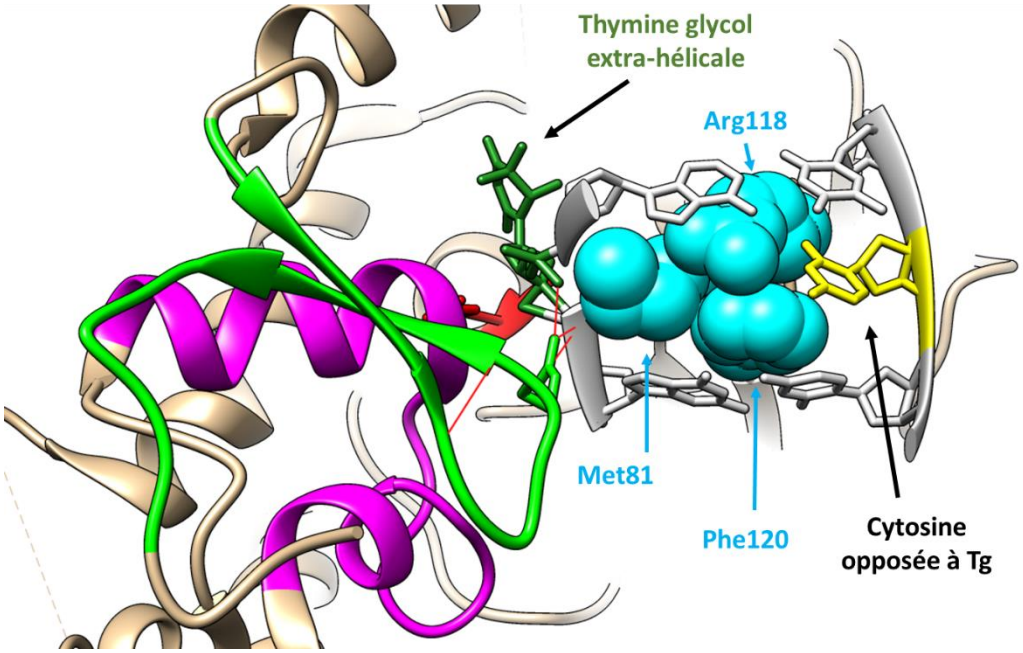


Figure 31 : Triplet d'intercalation de hNEIL1 (P2G) sur un ADN contenant une Thymine glycol (PDB 5ITY) - Les résidus Met81, Arg118 et Phe120 sont affichés en sphères de couleur cyan, le domaine H2TH en magenta, le ZnLF en vert clair, l'ADN en gris clair. La thymine glycol (Tg) est en vert foncé tandis que la cytosine opposée est en jaune.

Dans les enzymes de la famille Fpg/Nei, le motif ZnF ou ZnLF est capital pour l'interaction avec l'ADN (Figure 32). Il a été ainsi montré que les protéines avec un ZnF/ZnLF déstructuré n'étaient plus capables de se lier à l'ADN [390,391]. hNEIL1 contient donc un ZnLF, qui est composé des résidus 267 à 285, avec l'Arg276 qui est extrêmement bien conservée pour l'ensemble des enzymes de la famille Fpg/Nei [379]. En effet, Arg276 est directement en contact avec l'ADN et permet de stabiliser les groupements phosphates de part et d'autre de la lésion, tout en participant à la torsion de l'ADN, qui va faciliter l'extrusion de la base endommagée (Figure 32).

Enfin le motif H2TH, comprenant les résidus 161 à 185 pour hNEIL1, le principal motif structural de cette superfamille, est également fortement impliqué dans la fixation à l'ADN, principalement sur les

groupements phosphate en 5' [388]. hNEIL1 possède également une boucle non structurée en C-terminal, qui est impliquée dans l'interaction avec différents facteurs du BER comme RPA, FEN-1, PCNA, RFC ou encore Pol β, mais ne semble pas indispensable à l'activité catalytique [392–395].

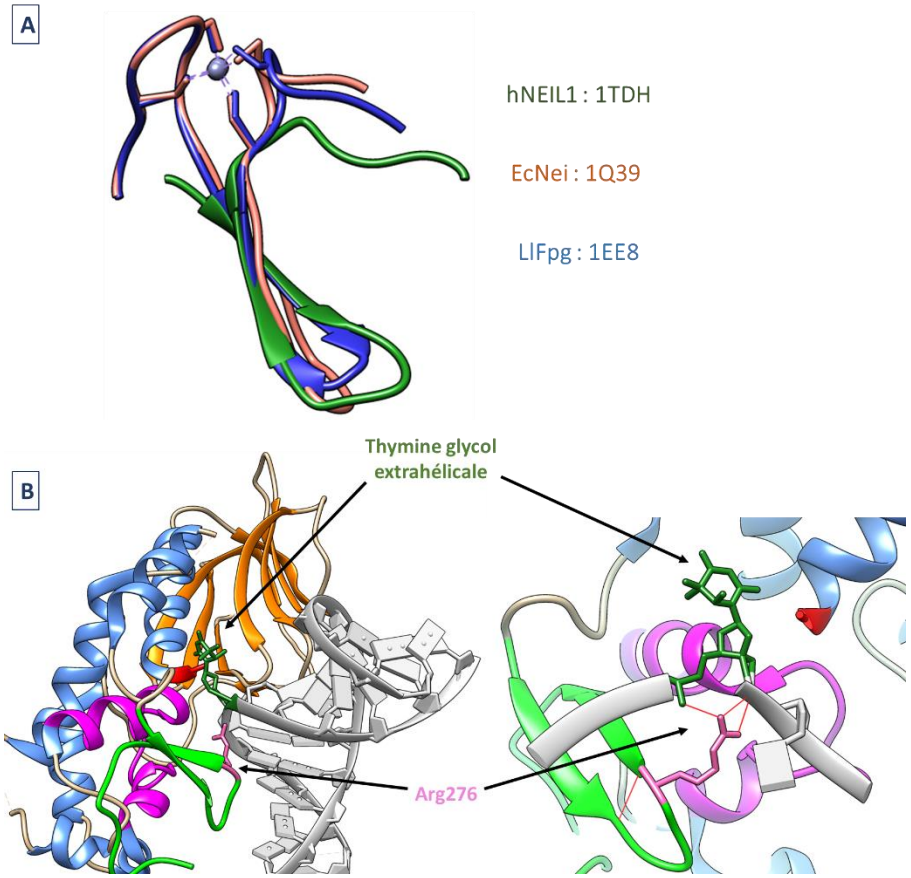


Figure 32 : Motif ZnF/ZnLF des ADN glycosylases Fpg/Nei - A) Superposition des motifs ZnF de TtFpg (résidus 231-266), EcNei (résidus 230-262), et ZnLF de hNEIL1 (résidus 263-290). Les cystéines sont indiquées pour LlFpg et EcNei. (adapté de Doublié *et al*, 2004 [379]). B) Interaction entre le ZnLF et l'ADN lors de la reconnaissance de la thymine glycol par hNEIL1 (P2G) (PDB : 5ITY). Les hélices α sont colorées en bleu, les brins β en orange, le domaine H2TH en magenta, le ZnLF est affiché en vert clair, la thymine glycol en vert foncé, le résidu Arg276 en rose clair, et les résidus catalytiques en rouge. Les liaisons hydrogène sont indiquées en rouge, avec l'ADN (sauf la Tg) en gris clair.

IV.2.2.Mécanisme catalytique

L'activité catalytique repose principalement sur les deux résidus catalytiques de Fpg/Nei, Pro1 et Glu2, situés sur l'hélice α N-terminale très bien conservée (Figure 30A). Le mécanisme de Fpg est très bien documenté, nous allons donc nous en servir pour présenter le mécanisme d'action général des enzymes Fpg/Nei, en prenant pour exemple l'excision de la 8-oxoG (Figure 33). Une fois dans le site actif, le C1' du nucléotide endommagé est exposé à l'attaque nucléophile du résidu Pro1, avec le support du deuxième résidu catalytique Glu2 [396]. Cela va entrainer une activité bifonctionnelle *N*-glycosylase/AP lyase, non découplée, contrairement à hOGG1. Ce mécanisme résulte en un clivage de la base endommagée, et donc la formation d'un site abasique. De façon concomitante, à l'excision de la 8-oxoG, un complexe covalent transitoire protéine-ADN de type base de Schiff se forme entre le Nα de Pro1 et le C1' du nucléotide endommagé. Cet intermédiaire base de Schiff peut être isolé par sa stabilisation irréversible au moyen d'un réducteur fort comme le borohydrure de sodium. Il a été proposé que la déglycosylation nécessitait l'ouverture du 2'-désoxyribose, par une protonation par exemple, avant le clivage du lien *N*-glycosidique [23,389]. La protonation du 2'-désoxyribose est soutenue par des calculs de chimie quantique [397,398]. Il en résulte un intermédiaire transitoire *N,N'*-bisubstitué hémiaminal, qui se réarrange en

base de Schiff suite à l'effet acido-basique de Glu2. Ensuite, une réaction de β,δ-élimination et une hydrolyse finale permettent de former une lacune ou « gap » de 1 nucléoside arborant des extrémités 3' et 5'-phosphate.

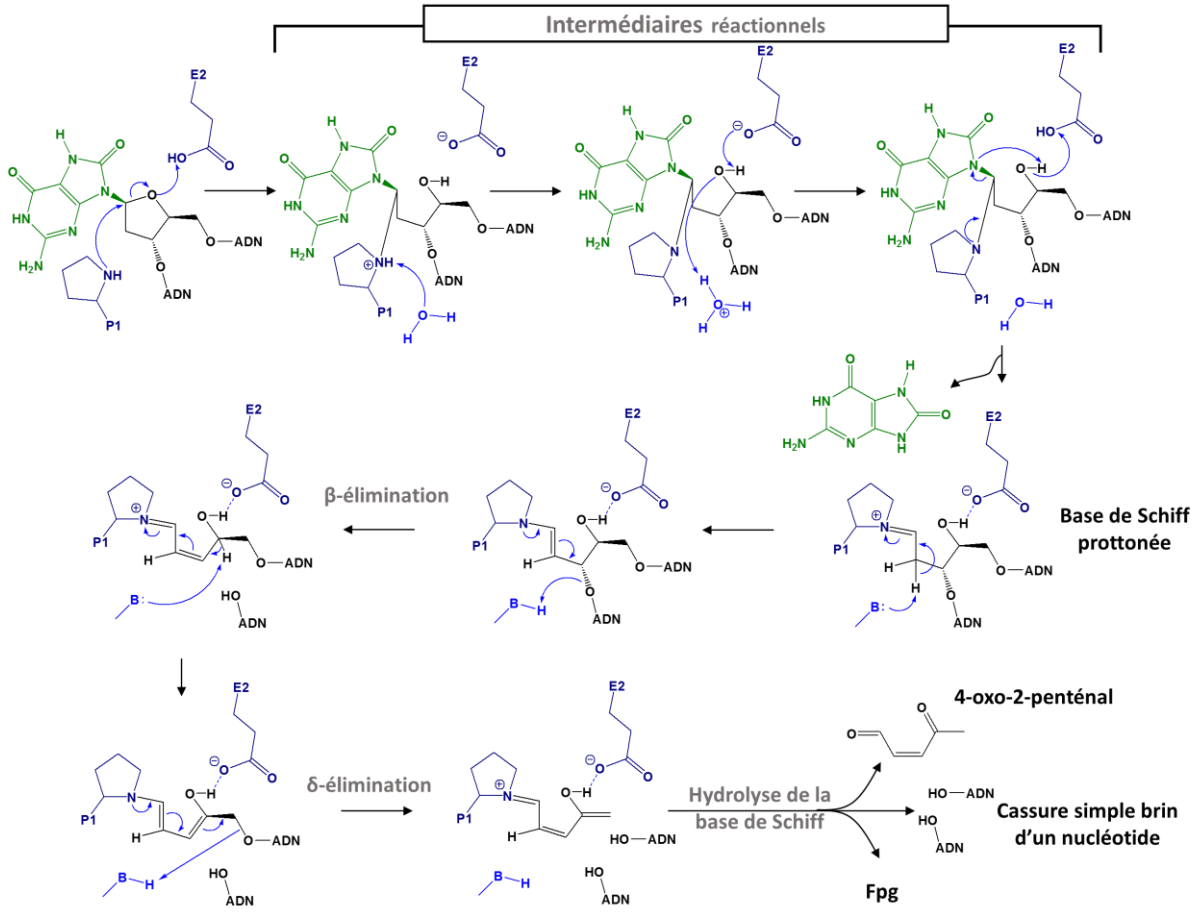


Figure 33 : Mécanisme de l'activité bifonctionnelle *N*-glycosylase/AP lyase de Fpg proposé par Boiteux *et al.*, 2017 [23]

En revanche, hNEIL1 ne peut pas exciser la 8-oxoG comme peut le faire Fpg (à part pour une 8-oxoG située proche de l'extrémité 3' d'une cassure simple brin [275]), ce qui pourrait suggérer un mécanisme différent. Il est à noter cependant que hNEIL1 peut exciser la 8-oxoG à proximité de l'extrémité 3' d'une cassure simple brin d'ADN. Cependant, des études cristallographiques ont montré que l'excision de la Thymine glycol (Tg) par hNEIL1 passait également par cette protonation du 2'-désoxyribose [399]. Zhu *et al.* ont en effet obtenu les premières structures de hNEIL1 complexé avec un double brin d'ADN contenant THF ou Tg, en utilisant une version inactive de l'enzyme, tronquée en C-terminal et mutée P2G. Ils ont également observé que la déprotonation de la Tg passe par deux mécanismes différents selon la tautomérisation de Tg (Figure 34).

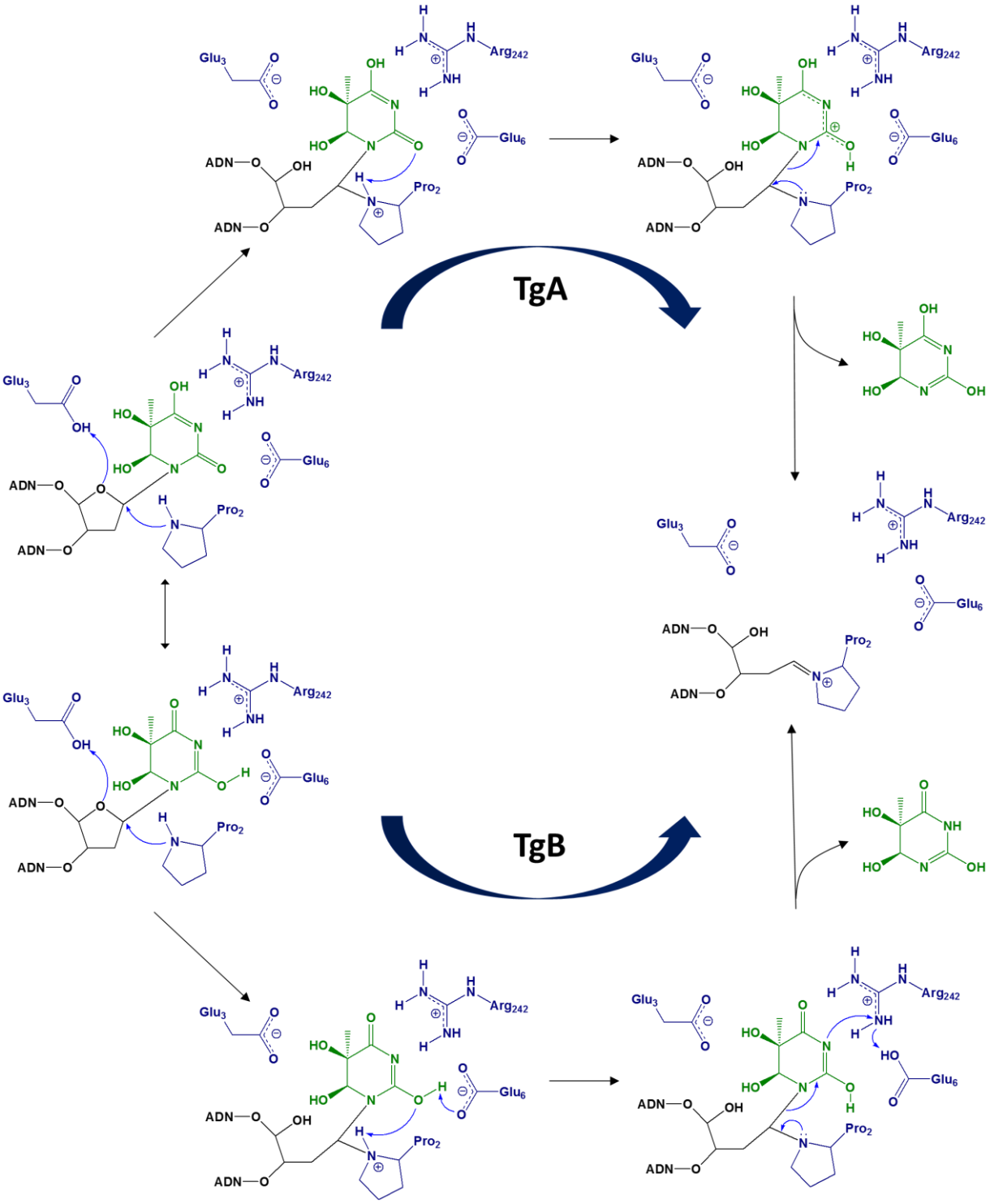


Figure 34 : Mécanisme de l'excision de la Tg par hNEIL1 proposé par Zhu *et al.*, 2016 [399] - D'après des calculs de thermodynamique, la voie TgB serait légèrement favorisée.

Une autre observation découlant de cette étude cristallographique a également révélé la présence d'une simili boucle LCL. En effet, bien que la LCL de Fpg lui confère la capacité d'exciser la 8-oxoG, hNEIL1 possède également une LCL, mais bien plus courte, impliquant ses résidus 240 à 252. Ici aussi la boucle peut être en conformation ouverte en l'absence d'ADN, et fermée en présence de substrat (Figure 35). En présence de la Tg, on voit que l'Arg242 est dirigée vers la lésion, tandis qu'elle est dans une direction opposée pour l'enzyme non complexée [399]. Si elle ne permet pas de reconnaitre la lésion 8-oxoG, cette boucle

« atrophiée » permet bien de reconnaitre les autres substrats de hNEIL1 et de contribuer à bien les positionner dans le site actif pour la catalyse.

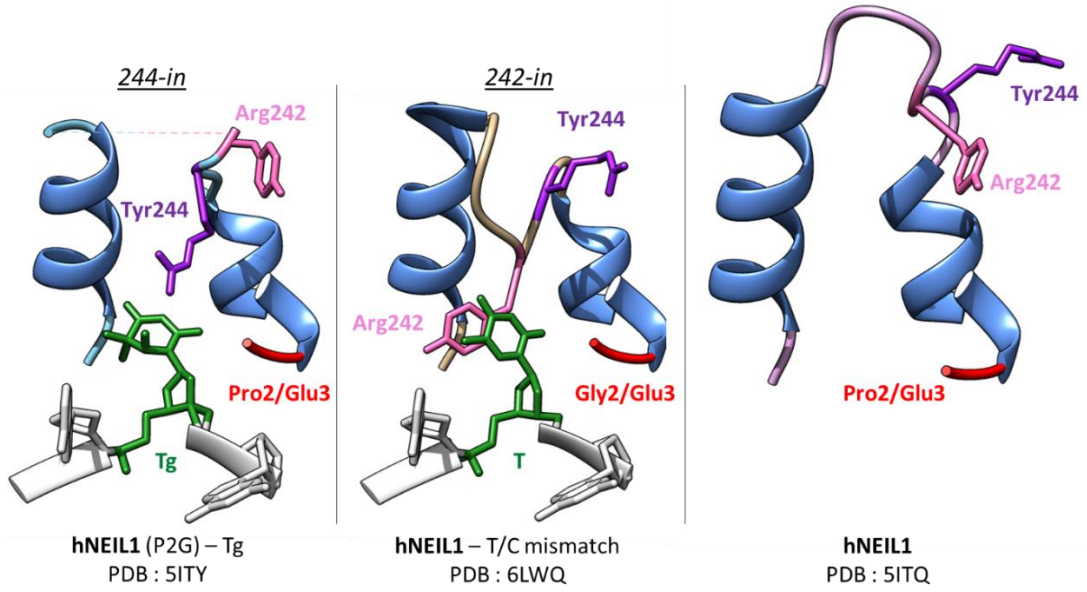


Figure 35 : Comparaison de la boucle LCL de hNEIL1 en conformation *224-in*/*242-in* (conformations fermées) ou de la protéine non complexée à l'ADN (conformation ouverte)

Cependant, malgré la mutation P2G, la lésion n'a pas pu être observée intacte dans le site actif par cristallographie, probablement à cause d'une activité catalytique résiduelle. Très récemment, la même équipe a publié de nouvelles structures de hNEIL1 en complexe avec l'ADN, avec une forme tronquée mais toujours active de hNEIL1 (notée dc95), et une double mutation P2G E3Q [400]. Ces nouvelles structures ont permis d'identifier certains résidus importants pour la différenciation des différents substrats de hNEIL1 par rapport aux bases classiques de l'ADN. Ainsi en présence de thymine, le résidu Arg242 prend la place de la Tyr244, qui en retour va s'empiler sur la base via des interactions de type π-π (Figure 35). La boucle de reconnaissance va également pivoter, en utilisant Gly240 et Gly249 comme pivot afin de libérer plus de place pour la base. En revanche, avec cette conformation hNEIL1 est inactive et protège la base normale de tout clivage. Ainsi, l'alternance entre ces deux conformations « 242-in » et « 244-in » semble permettre à hNEIL1 de discriminer une base endommagée substrat d'une base non-substrat. Il existe une version génomique de type sauvage de hNEIL1 K242, qui présente différents substrats et une activité aberrante, par exemple le résidu thymine dans le mismatch T:C. En réalité l'expression de la protéine hNEIL1-R242, beaucoup moins aberrante et sujet aux erreurs, résulte de l'action de l'enzyme ADAR qui désamine l'adénine dans l'ARNmNEIL1 [400,401].

IV.3. Ciblage de hNEIL1 et impact thérapeutique

IV.3.1.NEIL1 une cible thérapeutique intéressante mais peu étudiée

De par son implication dans le BER et sa capacité à réparer différentes bases endommagées dans l'ADN, NEIL1 est une enzyme participant à l'intégrité de notre patrimoine génétique, réparant également les dommages venant de la réplication [402,403]. Son absence ou son dysfonctionnement sont associés à l'apparition de nombreux cancers [404,405]. Son importance dans la réparation est également illustrée par le fait que son inhibition semble promouvoir la perte de fonction motrice dans la maladie de Parkinson, et son inhibition simultanée avec NTH1 augmente la cancérogénèse des cellules [193,406].

Cependant, comme pour les autres enzymes du BER que nous avons exposées, dans certains contextes une inhibition de cette enzyme peut être positive. Malheureusement, moins de données disponibles mettent en lumière l'intérêt de l'inhibition de NEIL1 par rapport à d'autres ADN glycosylases comme hOGG1 par exemple. Son inhibition a donc de fait été bien moins étudiée.

Même si les données sont moins nombreuses, il existe un certain nombre de contextes favorables à la suppression de NEIL1. Tout d'abord, comme les autres ADN glycosylases, son KO chez la souris est viable, malgré une plus forte propension à l'hypertension et à l'obésité [407]. En revanche, son inactivation permet de sensibiliser les cellules à des radiations UVB ou gamma [408,409]. De plus, la sous-expression de NEIL1 dans des cellules de cancer colorectal réduit la prolifération cellulaire, tandis que sa surexpression protège les cellules cancéreuses de l'apoptose [410]. Une relation de létalité synthétique a été potentiellement identifiée entre les contextes génétiques de l'anémie de Fanconi (maladie génétique rare associée à une aplasie médullaire et un risque de leucémie et cancers), sur lesquels l'inhibition de NEIL1 est toxique [411]. Une autre relation de létalité synthétique a également été identifiée en présence d'inhibiteurs de Thymidilate synthétase comme pour OGG1 [347]. Comme hOGG1, hNEIL1 semble participer à la promotion de l'expansion du triplet CAG, responsable de la maladie d'Huntington [412,413]. Enfin, très récemment, une étude a montré que la présence de hNEIL1 pouvait contrebalancer l'effet du composé TH5487 (un inhibiteur de OGG1). Une utilisation d'un inhibiteur de NEIL1 conjointement à celle de TH5487 pourrait donc également être justifiée [414].

IV.3.2.Inhibiteurs identifiés de hNEIL1

Certains ions ont été identifiés comme inhibiteurs, notamment le sulfate de fer (II) $FeSO_4$, mais ici encore, la spécificité de ces inhibiteurs est très faible, et leur utilisation concrète *quasi* impossible [415]. Comme pour hOGG1, le Fumarylacétoacétate est également capable d'inhiber hNEIL1.

Les principaux inhibiteurs viennent de deux études en particulier. Tout d'abord, suite à un criblage HTS, Jacobs *et al.* ont identifié différents composés, des analogues de purines (noté Px, Figure 36), avec un IC_{50} de l'ordre du micromolaire [416]. A notre connaissance, aucune structure 3D présentant hNEIL1 et ces inhibiteurs n'a été publiée, et le mécanisme d'inhibition reste inconnu. Il est cependant probable que celui-ci passe par l'occupation du site actif par les composés Px.

Notre équipe a également identifié des inhibiteurs de hNEIL1, basés autour d'analogues de la 2-thioxanthine (2TX). A l'origine ces composés ont été identifiés comme inhibiteurs de Fpg de *Lactococcus lactis*, mais pas de son homologue humain [247]. L'inhibition des premiers composés identifiés sur Fpg, passe par l'attaque des cystéines du doigt de Zinc, libérant l'ion Zn^{2+}, et coupant l'activité catalytique. En revanche, du fait que hNEIL1 ne possède pas de doigt zinc, ni de cystéines sur le ZnLF, ces premiers composés montrent donc une inhibition anecdotique de hNEIL1. Par la suite cependant, une deuxième génération de composés, principalement des composés dithiols, a été synthétisée et identifiée comme des inhibiteurs de Fpg mais également de hNEIL1 [246]. Ces composés dithiols sont capables de former des trimères cycliques, et des modélisations de dynamique moléculaire montrent que ces composés peuvent interagir avec plusieurs sites sur l'enzyme, notamment le site actif. En revanche, aucune structure 3D expérimentale n'a été obtenue avec cette nouvelle classe d'inhibiteurs sur hNEIL1.

En conclusion, le nombre limité d'inhibiteurs de hNEIL1 et leurs efficacités relatives, induit la nécessité de recherches approfondies et l'identification de nouveaux composés plus efficaces et sélectifs.

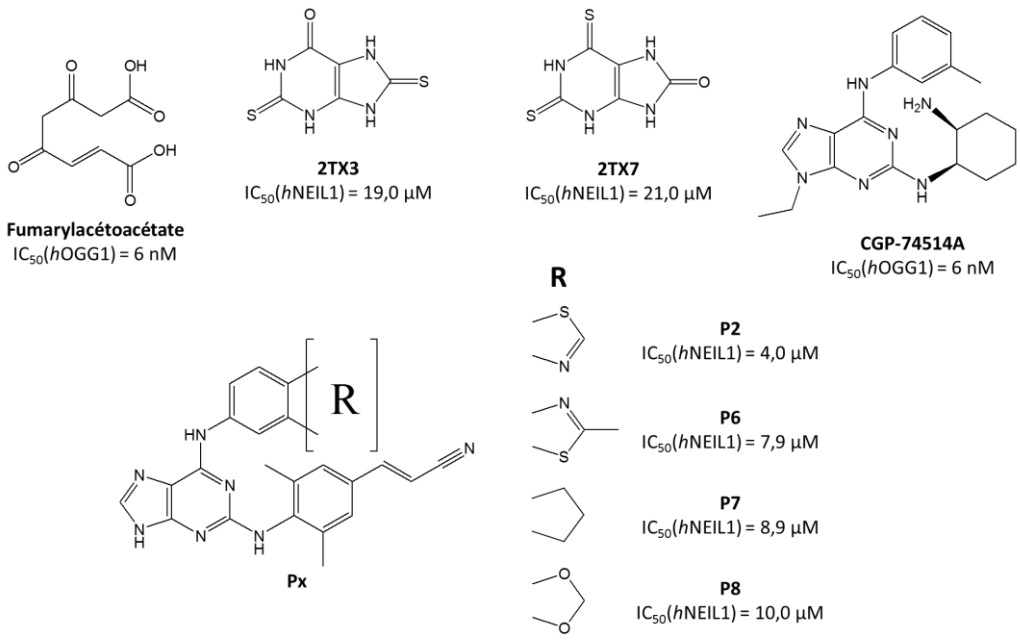


Figure 36 : Inhibiteurs identifiés de hNEIL1

Stratégie et objectifs de la thèse

I. Généralités sur le criblage

De nos jours, pour répondre à la nécessité du développement de nouveaux médicaments, la recherche pharmaceutique se base principalement sur des criblages à haut débit. En effet, ceux-ci permettent de tester un grand nombre de composés synthétiques à travers une miniaturisation d'essais *in vitro* identifiant les modulateurs d'une cible biologique d'intérêt. Le volume de composés testés peut aller jusqu'à des millions de molécules, ce qui serait irréalisable avec des essais biochimiques classiques. Cette nécessité de tester un grand nombre de composés vient du fait que le développement d'un médicament est un procédé particulièrement long et couteux, pouvant varier entre 10 et 15 ans pour en moyenne 1.5 milliard d'euros [417,418]. De plus, il dépend de plusieurs étapes critiques et peut donc être stoppé à tout moment, et l'on considère généralement que 92 % des molécules atteignant les phases cliniques I, ne sont finalement pas retenues [419]. Disposer de plusieurs composés de structures chimiques différentes est donc essentiel pour aboutir à un potentiel candidat médicament. Cela justifie également la recherche d'inhibiteurs différents sur des cibles biologiques pour lesquelles des inhibiteurs ont déjà été identifiés, comme c'est le cas pour cette thèse. Nous allons ici brièvement développer le principe d'un criblage moléculaire classique, afin d'établir les étapes charnières du criblage que nous allons effectuer, et situer ce travail de thèse dans le procédé de l'identification d'un composé thérapeutique.

I.1. Origine et évolution du HTS

Comme la plupart des technologies avancées, le criblage moléculaire tient son origine d'une succession d'innovations techniques modestes. Dans les années 50, une sévère pandémie grippale causée par le virus Influenza toucha fortement la Hongrie. Il en résulta une pénurie de petits matériels de laboratoire et il devint donc compliqué pour les laboratoires hongrois de l'époque de continuer les études sur le virus et les titrages des différents sérums. Pour résoudre ce problème, le Dr Takátsy développa un dispositif pipeteur à base de boucles circulaires calibrées réutilisables permettant de faire des micro-dilutions avec un volume minimal de 25 µL [420] sans utiliser de pipette (Figure 37). De plus, dans le but d'utiliser ce dispositif sans consommer de tubes et réduire ainsi davantage les consommables, il inventa aussi la première plaque microtitre (dont les dispositions des puits 6*12 ou 8*12 viennent de la possibilité de prendre 6 ou 8 boucles avec une seule main [421,422]).

Suite à ce premier développement qui se popularisa rapidement dans le monde, de nouvelles améliorations furent apportées. En 1953, John Liner créa ainsi une version moulée d'une plaque microtitre de 96 puits et commença la production industrielle de celle-ci en fondant American Linbro [423]. Cette production industrielle de microplaques motiva le développement de dispositifs de dilution plus efficaces, particulièrement dans l'industrie biochimique, comme en 1967 où Tom Astle et l'entreprise Astec (renommée TomTec depuis) produisirent l'Autotier, une machine de dilution complètement automatisée.

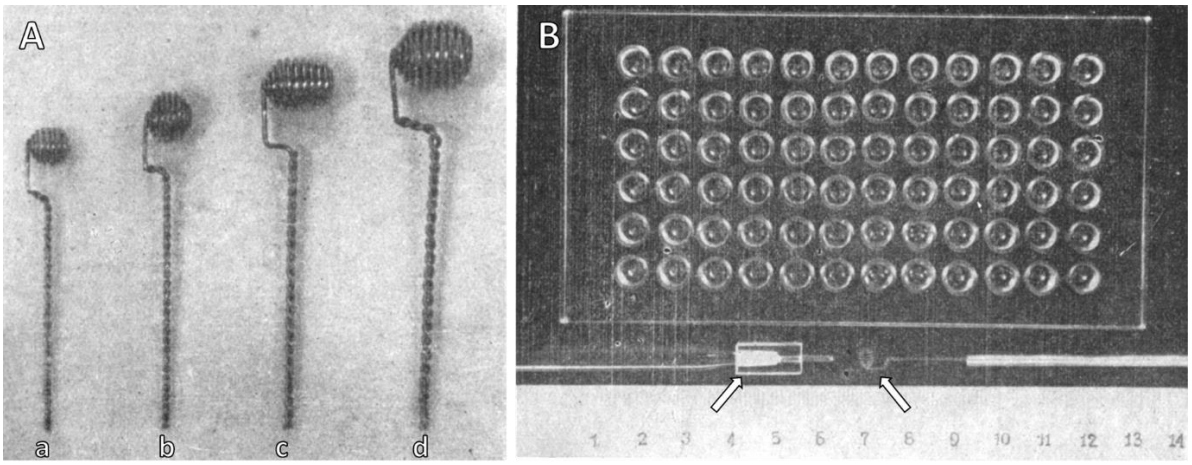


Figure 37 : Matériel de micropipetage fabriqué en 1950 (adapté de Takátsy, 1950[420]) - A) Boucle en spirale créée par Gyula Takátsy pouvant contenir 0,025, 0,05, 0,1 et 0,2 mL pour (a), (b), (c,) et (d) respectivement. B) Plaque en polyméthacrylate de méthyle (ou Lucite) de 6*12 puits. En bas à gauche est indiqué par une flèche un instillateur couplé à une pipette, tandis qu'à droite la boucle (a) est indiquée.

De nombreuses évolutions diverses suivirent, et ces développements techniques facilitèrent l'usage des microplaques en biochimie, et de ce fait, le développement des lecteurs de microplaques [424]. C'est dans les années 1990, que l'industrie des microplaques se développa drastiquement, et avec elle le début des criblages de molécules [425]. Différentes plaques furent développées et la Société pour le Criblage Biomoléculaire (aujourd'hui la « Society for Laboratory Automation and Screening » SLAS en anglais) établit les premiers standards de format de plaque, avec l'accord des principaux fabricants (SBS ANSI 2004) [426]. L'augmentation de la taille des plaques (96/384/1536 puits) et différents développements permirent de mettre en place les premiers criblages de molécule à haut débit, avec entre 10^4 et 10^5 molécules testées par jour. De nos jours, on parle même d'ultra criblage à haut débit (uHTS) avec plus de 10^5 composés testés par jour [427].

I.2. Etapes principales d'un criblage à visée thérapeutique

Comme nous l'avons précisé précédemment, développer un médicament est un procédé coûteux fait de multiples étapes (Figure 38) et demandant une forte interdisciplinarité de la biologie à la chimie médicinale. Nous allons ici décrire les étapes principales d'un criblage moléculaire, qu'il soit moyen ou haut débit, pour produire un candidat médicament capable d'être proposé dans un test clinique.

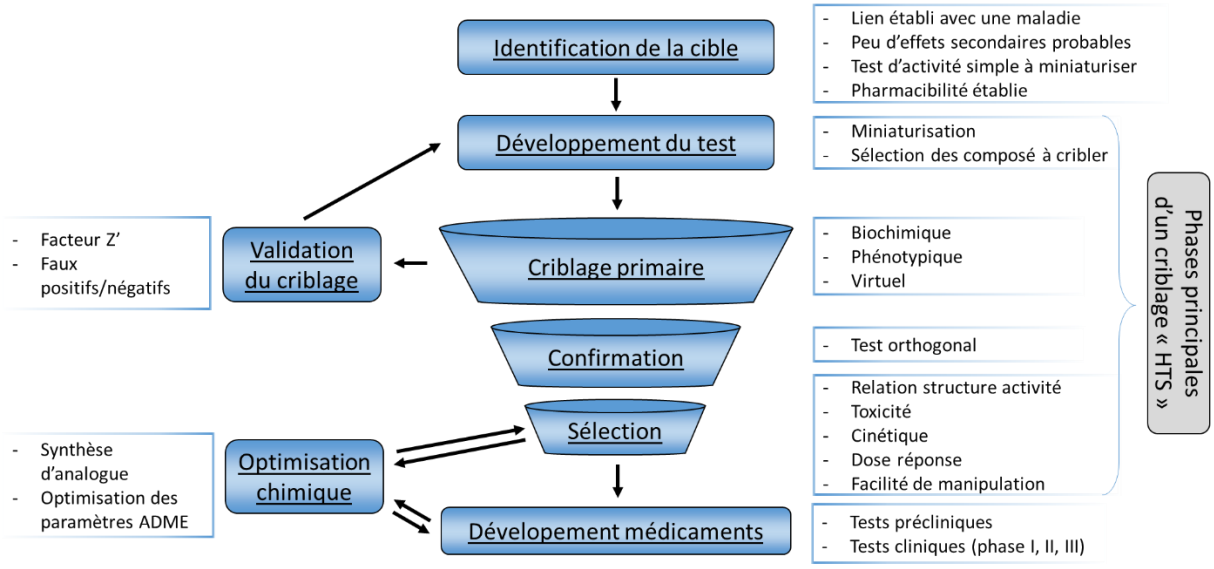


Figure 38 : Schéma des étapes principales d'un criblage classique pour le développement d'un médicament

I.2.1. Validation de la cible biologique d'intérêt

La première étape est donc d'identifier une cible biologique d'intérêt. Communément, cela passe par une protéine fortement impliquée dans une maladie [428]. Récemment, une approche de « ciblage multiple », avec une molécule visant plusieurs mécanismes biologiques différents, est aussi étudiée, mais il reste généralement considéré que cela peut souvent mener à des complications [429]. Un composé hautement sélectif est donc l'objectif majeur de la majorité des criblages.

Afin de déterminer si la cible est intéressante, différents critères sont évalués [430] :

- La **validité** de la cible : Evaluer à quel degré la cible est impliquée dans une maladie donnée (pathologie) et si sa modulation impacte peu les conditions physiologiques normales ;
- La **facilité** avec laquelle un essai est envisageable pour évaluer son activité, ce qui est nécessaire pour mettre en place un criblage efficace et robuste ;
- La **pharmacibilité** (« **druggability** » en anglais), qui évalue la capacité de la cible à interagir avec des petites molécules de manière sélective et efficace. Ce critère théorique utilise les structures 3D de la cible, ou d'un homologue [431] ;
- Une **prédiction favorable d'effet secondaire** suite à des études phénotypiques.

Malheureusement, il faut noter que la validité d'une cible est souvent inversement reliée à son originalité. Les cibles ayant un lien bien documenté avec une maladie ont en effet probablement déjà eu de nombreux criblages les ciblant. En revanche, une cible moins étudiée demandera une mise en place plus importante, sans réussite assurée [432].

I.2.2. Criblage primaire

La deuxième étape est de mettre en place un criblage primaire (« primary screen » en anglais) afin d'évaluer la modulation de la cible biologique en présence des molécules testées et d'identifier des molécules « touches », ou « hits ». A ce stade, il y a différents types de criblages.

Tout d'abord le **criblage biochimique**, le plus couramment utilisé, de par sa relative simplicité de mise en place. A l'origine assez consommateur en temps et en matériel, il est aujourd'hui généralement basé sur une technique simple et homogène « mélanger et lire » (« mix and read » en anglais) [433]. Cette technique

emploie donc souvent des méthodes optiques telles que la fluorescence, la luminescence ou l'absorbance afin d'évaluer l'activité de la cible. Cela permet de lire directement les plaques de criblage sans utiliser de technique de séparation complexe ce qui économise du matériel et permet de cribler des plaques de plus grand format [434]. Parmi ces techniques de détection, la plus utilisée est la fluorescence, par sa sensibilité très élevée [435].

Ensuite le **criblage phénotypique**, plus récent, qui permet d'effectuer un crible directement dans un contexte cellulaire [436]. Ceci permet de détecter toute molécule ayant un effet sur un mécanisme particulier de la cellule, alors qu'elle aurait pu être mise de côté dans un contexte biochimique [437], mais aussi d'évaluer de potentiels effets toxiques [438]. Cela consiste souvent à tester directement les composés sur des lignées cellulaires d'intérêt, et évaluer la viabilité des cellules. Cette stratégie a permis de développer de nombreux traitements anti-cancéreux. En revanche, aucune information sur le mode d'action de l'inhibiteur ne peut être obtenue sans des tests de relation structure/activité (SAR) ou de biochimie plus approfondis [439,440].

Enfin, le **criblage virtuel**, qui consiste à modéliser la cible biologique par informatique et à chercher dans des banques de données toutes les molécules susceptibles d'interagir avec celle-ci [441]. Cela passe souvent par de la dynamique moléculaire (qui prend en compte l'évolution du système au cours du temps) et/ou alors de l'amarrage moléculaire (« docking ») qui va classer les composés en fonction de leurs possibles interactions avec la cible fixe [442]. Cette technique, rapide et peu chère (si l'installation informatique est déjà mis en place) vient très souvent en amont des criblages précédents, et nécessite dans tous les cas une validation *in vivo*/*in vitro*.

Evidemment, d'autres techniques de criblage existent, couplées à d'autres méthodes de détection, comme la **RMN** [443,444], la **cristallographie** [445,446] ou encore la **spectrométrie de masse** [447,448], mais ne font pas encore partie des criblages les plus communs, principalement par leur difficulté de mise en place.

I.2.3. Validation et confirmation du criblage

Afin de valider le premier criblage primaire, il est primordial de mettre en place une validation formelle des hits obtenues et de la robustesse du criblage. Il faut tout d'abord mettre en place des **contrôles inhérents** à l'expérience choisie pour le criblage primaire, qui vont pouvoir attester de sa **robustesse**.

Une importante valeur statistique est le **facteur Z'**, dont la valeur est calculée avec l'Equation 1. Ce facteur remplace le ratio signal sur bruit qui était à l'origine utilisé, et permet de prendre en compte aussi bien la moyenne des résultats que leurs écarts relatifs, donnant une valeur statistique bien plus robuste [449]. En théorie, le meilleur moyen d'avoir la valeur de **Z'** la plus robuste possible serait de répartir les puits contrôle aléatoirement sur la plaque. Cependant, en pratique, la génération de plaques aléatoires prendrait beaucoup trop de temps, que ce soit des plaques préparées manuellement ou avec un robot. En conséquence, les contrôles sont souvent placés sur une même colonne, en évitant les bords de plaques. **Z'** est une valeur numérique sans unité comprise entre 0 et 1. Plus la valeur est proche de 1, plus le test sera considéré comme robuste. Si l'on considère qu'un **Z'** supérieur à 0,5 est suffisant [450], il est maintenant conseillé d'obtenir un **Z'** supérieur à 0,7 pour attester de la très bonne robustesse du test.

$$Z' = 1 - \frac{3\sigma_{(+)} + 3\sigma_{(-)}}{\left|\mu_{(+)} - \mu_{(-)}\right|}$$

Equation 1 : Facteur Z' - $\mu_{(+)}$ et $\sigma_{(+)}$ sont la moyenne et l'écart relatif du signal maximum du test tandis que $\mu_{(-)}$ et $\sigma_{(-)}$ sont la moyenne et l'écart relatif du bruit de fond, qui est donc le signal le plus faible obtenu (souvent en se plaçant dans un contexte sans cible biologique).

Lorsque le criblage primaire est validé, il est capital de confirmer les hits issus du crible. Pour ce faire, un criblage secondaire doit être mis en place afin d'éliminer les faux positifs, et confirmer l'effet modulateur des hits sur la cible (il est à noter que les faux négatifs sont perdus). Pour être efficace, ce second test doit être orthogonal au premier, le plus souvent en utilisant une technique de détection différente [451]. Cela permet d'éliminer les molécules ayant un effet sur la première technique de détection, et non sur la cible en elle-même (par exemple une molécule interférant avec le signal fluorescent). Enfin d'autres tests de caractérisation (études structurales, cinétiques ou encore toxicité) peuvent aussi être utilisés pour sélectionner les têtes de série (« lead » en anglais) et améliorer au maximum leurs propriétés pharmacologiques, en synthétisant des composés analogues par exemple.

I.2.4. Tests précliniques et cliniques

Lorsque les têtes de série ont été sélectionnées et suffisamment caractérisées, on peut commencer les différents tests cliniques. Ils consistent à évaluer le comportement pharmacologique et la cinétique des têtes de série, ainsi que leur toxicité, de la cellule au corps humain entier. Les tests dits « **précliniques** » reposent principalement sur une expérimentation chez l'animal. Ces tests permettent généralement d'évaluer les doses optimales en prévision des futurs tests dits « cliniques », ainsi qu'optimiser la composition du médicament. Lors de cette étape, les caractéristiques pharmaceutiques du composé (métabolisme, toxicité, distribution, absorption, profil ADMET, activité) sont vérifiées, et optimisées par des modifications synthétiques ou méthodologiques si nécessaire. Cette étape peut durer jusqu'à 15 mois, bien que des temps plus courts soient souvent envisageables. Enfin, des études de stabilité à long terme (LTSS) seront menées sur les composés retenus durant 6 mois afin d'évaluer la dégradation des molécules.

Suite à ces études sur l'animal, des études cliniques sont mises en place. Celles-ci doivent permettre *in fine* d'obtenir une autorisation de mise sur le marché. Elles sont séparées en 3 phases :

- **Phase I**, consiste à définir les profils pharmacocinétiques du médicament sur un nombre restreint de candidats volontaires sains. Cela permet d'évaluer la sécurité d'emploi du composé, sa demi-vie et l'excrétion du composé.

- Si les résultats sont positifs, on passe à la **phase II**, qui va établir l'efficacité des composés sur une gamme de patients volontaires. La tolérance et la dose à prescrire seront également déterminées.

- Enfin, la **phase III** va venir conclure ces études en portant sur la sécurité et l'efficacité des conditions préalablement déterminées sur des milliers, voire des dizaines de milliers de patients. Des études en double aveugle seront mises en place pour étudier l'effet placebo et l'effet du composé face au traitement de référence. Si une efficacité réelle est prouvée, la mise sur le marché est acceptée. La fabrication du médicament continue d'être optimisée, ainsi que la phase de pharmacovigilance.

II. Objectifs

L'objectif de ma thèse est de contribuer à l'identification de nouveaux inhibiteurs des ADN glycosylases hOGG1 et hNEIL1. En effet, comme nous l'avons évoqué, ces protéines représentent des cibles intéressantes, notamment dans la lutte contre le cancer. Il est à noter qu'au début de cette thèse, très peu d'inhibiteurs étaient identifiés sur ces deux enzymes. Depuis, de nouveaux inhibiteurs efficaces et prometteurs ont été identifiés, tout particulièrement sur hOGG1. Loin d'amoindrir l'intérêt du développement de nouveaux inhibiteurs, ces nouvelles molécules montrent plutôt l'intérêt croissant pour ces cibles prometteuses et leur validation biologique. Cela encourage donc la découverte de nouvelles molécules. En effet, comme nous l'avons évoqué, pour hOGG1, mais également pour l'ensemble des

composés identifiés sur les systèmes de réparation, la diversité structurale des inhibiteurs est la clé pour optimiser les chances d'obtenir un traitement efficace. Pour ce faire, nous avons identifié de nouveaux inhibiteurs, et nous les avons caractérisés en utilisant des approches biochimiques et structurales.

Dans un premier temps, mon travail de thèse s'est concentré sur la mise en place et la validation d'un criblage primaire robuste et fonctionnel, en utilisant l'enzyme bactérienne LlFpg comme modèle. Par la suite, nous avons sélectionné différentes chimiothèques, et effectué des criblages sur les ADN glycosylases humaines hOGG1 et hNEIL1. Il est important de noter, que nos criblages s'inscrivent dans une démarche académique, ce qui signifie que le volume des criblages effectués est moindre comparé à ceux effectués par l'industrie pharmaceutique. Cependant les différentes étapes du criblage ont été développées afin de sélectionner les composés les plus sélectifs et efficaces, en optimisant au maximum l'originalité des structures et le matériel utilisé. Ainsi, une fois des hits identifiés, nous avons confirmé leur effet inhibiteur à travers différentes étapes de validation, puis nous avons évalué leur efficacité.

Dans un second temps, nous avons mis en place une expansion des hits obtenues lors du criblage primaire, afin d'optimiser nos composés.

En parallèle, nous avons étoffé la caractérisation de nos inhibiteurs sur hOGG1/hNEIL1 par différentes études biochimiques, afin de comprendre au mieux les mécanismes d'inhibition de ces composés, d'évaluer leur spécificité et influer potentiellement la conception de nouveaux analogues. Ces différentes études ont eu pour but le développement d'une seconde génération de composés plus efficaces. Enfin nous avons mis en place de nombreuses études structurales (principalement cristallographiques mais également de la spectrométrie de masse). Ces étapes, conditionnées par notre capacité à cristalliser nos protéines en présence des inhibiteurs, devaient permettre la compréhension des relations structure/activité entre nos inhibiteurs et les ADN glycosylases humaines ciblées.

Matériel et méthodes

I. Préparation des protéines hNEIL1 et hOGG1

I.1. Transformation bactérienne

50 µL de bactéries calcium-compétentes, BL21-DE3 ou Rosetta 2-DE3, sont incubées avec environ 50-100 ng de plasmides (pET28a+ pour PabAGOG, mOGG1, d2hOGG1 pFlag NdeI/XmaI pour LlFpg, pET30a pour hOGG1, hNEIL1, hNEIL1dc95, hNEIL1dc56 et dhOGG1, pet22b+ pour mvNei) pendant 30 min dans la glace. Un choc thermique à 42 °C est ensuite appliqué pendant 45 secondes. Les bactéries sont ensuite rapidement remises dans la glace. Après environ 2 minutes, 500 µL de milieu de culture LB (pour « lysogeni broth » - 20 g/L : un milieu nutritif standard composé de 10 $g.L^{-1}$ Tryptone, 5 $g.L^{-1}$ extrait de levures, 5 $g.L^{-1}$ NaCl) sont ajoutés et la culture est incubée pendant une heure sous agitation à 37 °C. On étale ensuite 10 ou 100 µL de cette culture non diluée sur une boite de Petri contenant 20 mL de LB-Agar (1,5 g Agar/100 mL LB) supplémenté de Kanamycine (Kana - 50 µg/mL), et selon les bactéries utilisées de Chloramphénicol (Cm - 34 µg/mL). Les boites sont placées dans des étuves à 37 °C sur la nuit. Les colonies isolées seront ensuite utilisées pour ensemencer des cultures plus importantes.

I.2. Production

I.2.1. Autoinduction

On ensemence du milieu non-inducteur MDG [452] (2 $mmol.L^{-1}$ $MgSO_4$; 0,25 % aspartate ; 0,5 % glucose ; 25 $mmol.L^{-1}$ NA_2HPO_4 ; 25 $mmol.L^{-1}$ KH_2PO_4 ; 50 $mmol.L^{-1}$ NH_4Cl ; 5 $mmol.L^{-1}$ Na_2SO_4 ; 0,2X du stock 1000X de « trace elements » contenant [50 $mmol.L^{-1}$ $FeCl_3$; 20 $mmol.L^{-1}$ $CaCl_2$; 10 $mmol.L^{-1}$ $MnCl_2$; 10 $mmol.L^{-1}$ $ZnSO_4$; 2 $mmol.L^{-1}$ $CoCl_2$; 2 $mmol.L^{-1}$ $CuCl_2$; 2 $mmol.L^{-1}$ $NiCl_2$; 2 $mmol.L^{-1}$ Na_2MoO_4 ; 2 $mmol.L^{-1}$ Na_2SeO_3 ; 2 $mmol.L^{-1}$ H_3BO_3]), supplémenté de Kanamycine (Kana - 50 µg/mL) et Chloramphénicol (Cm - 34 µg/mL) avec une colonie isolée Rosetta 2-DE3 possédant le plasmide désiré (hNEIL1dc95/dhOGG1/pabAGOG). Les bactéries sont ensuite amplifiées à 37 °C sur la nuit sous agitation.

0,5 mL de cette préculture est ensuite utilisé pour ensemencer 500 mL de milieu ZYM-5052 (10 $g.L^{-1}$ tryptones ; 5 $g.L^{-1}$ extraits de levures ; 25 $mmol.L^{-1}$ Na_2HPO_4 ; 25 $mmol.L^{-1}$ KH_2PO_4 ; 50 $mmol.L^{-1}$ NH_4Cl ; 5 $mmol.L^{-1}$ Na_2SO_4 ; 0,5 % glycérol ; 0,05 % glucose ; 0,2 % lactose ; 2 $mmol.L^{-1}$ $MgSO_4$; 0,2X du stock 1000X de « trace elements ») auto-inducteur sélectif, supplémenté de Kana et Cm, dans un erlenmeyer avec déflecteurs. Les cultures sont ensuite incubées à 37 °C sous agitation pendant environ 4 h. Elles sont alors placées à 18 °C sous agitation pendant une durée de 16 à 20 h. Les bactéries sont enfin collectées par centrifugation (6100 g à 4 °C pendant 45 min sur Super-lite® SLA-3000 rotor, Sorvall®) avant d'être reprises dans un tampon de lavage (10 mM tris/HCl pH 8, 150 mM NaCl) et reculottées en falcon de 50 mL (3000 g à 4 °C pendant 30 min sur centrifugeuse 5804R Eppendorf®, rotor A-4-44). Les culots obtenus sont finalement pesés afin d'établir la masse humide obtenue, et conservés à -20 °C.

Pour les protéines que nous souhaitions utiliser en cristallographie, notamment d2hOGG1, hNEIL1dc56 et mOGG1, nous avons également utilisé une induction à l'IPTG « classique » ou avec une température plus faible.

I.2.2. Production IPTG

On ensemence du milieu LB (200 mL) supplémenté de Kanamycine (50 µg/mL) avec une colonie isolée de bactéries BL21-DE3, possédant le plasmide désiré (hNEIL1dc56). Les bactéries sont ensuite amplifiées à 37 °C sur la nuit sous agitation.

Cette préculture est ensuite culottée (500 rpm à 4 °C pendant 20 min), reprise dans du LB-Kana frais, et utilisée pour ensemencer les cultures de travail (4*0,5L de LB-Kana) telles que la turbidité de départ de ces cultures de travail soit comprise entre 0,1 et 0,15 UA_{600nm}. Elles sont ensuite incubées à 37 °C sous agitation pendant environ 2 h. Lorsque la turbidité a atteint une DO d'environ 0,8 UA_{600nm}, l'expression des protéines est induite par l'ajout de 0,5 mM d'IPTG. Après environ 2 à 3 h d'incubation (160 rpm à 37 °C), les bactéries sont collectées par centrifugation (6000 rpm à 4 °C pendant 45 min) avant d'être reprises dans un tampon de lavage (10 mM tris/HCl pH 8, 150 mM NaCl) et reculottées en falcon de 50 mL (3000 g à 4 °C pendant 30 min sur centrifugeuse 5804R Eppendorf®, rotor A-4-44). Les culots obtenus sont finalement pesés afin d'établir la masse humide obtenue, et conservés à -20 °C.

I.2.3. Production IPTG à faible température

On ensemence du milieu LB (200 mL) supplémenté de Kanamycine (50 µg/mL) avec une colonie isolée de bactéries BL21-DE3, possédant le plasmide désiré (d2hOGG1). Les bactéries sont ensuite amplifiées à 37 °C sur la nuit sous agitation.

Cette préculture est ensuite culottée (500 rpm à 4 °C pendant 20 min), reprise dans du LB-Kana frais, et utilisée pour ensemencer les cultures de travail LB-Kana enrichi en Sorbitol (0,5 M) et en Bétaine (25 mM), telle que la turbidité de départ de ces cultures de travail soit comprise entre 0,1 et 0,15 UA_{600nm}. Elles sont ensuite incubées à 37 °C sous agitation pendant environ 2 h. Lors que la turbidité a atteint une DO d'environ 0,8 UA_{600nm}, la culture est alors placée à 18 °C. Après une heure d'équilibration de la température, l'expression des protéines est induite par l'ajout de 0,5 mM d'IPTG, et placée sous agitation pendant une durée de 16 à 18 h. Les bactéries sont enfin collectées par centrifugation (6100 g à 4 °C pendant 45 min sur Super-lite® SLA-3000 rotor, Sorvall®) avant d'être reprises dans un tampon de lavage (10 mM tris/HCl pH 8, 150 mM NaCl) et reculottées en falcon de 50 mL (3000 g à 4 °C pendant 30 min sur centrifugeuse 5804R Eppendorf®, rotor A-4-44). Les culots obtenus sont finalement pesés afin d'établir la masse humide obtenue, et conservés à -20 °C.

Pour mOGG1, le protocole est analogue à celui de d2hOGG1, à l'exception prêt que nous avons utilisé des cultures de travail qui sont des milieux LB-Kana enrichis en glycérol (0,4 %) et en glucose (0,8 %), en suivant les descriptions indiquées par Visnes *et al.*, 2018 [371].

I.3. Purification

Pour toutes les purifications, l'ensemble des étapes a été effectué à 4 °C, notamment les étapes de chromatographies d'affinités et échangeuses d'ion qui ont été effectuées en chambre froide. L'unique exception est la purification de pabAGOG dont les étapes ont été effectuées à température ambiante lorsque c'était possible. Les étapes de purification impliquant différentes colonnes de chromatographie ont requis l'utilisation des Akta « Purifier », « Pure » et « Start » (GE Healthcare). Afin d'évaluer la purification ou la production des protéines, les différentes étapes de purification ont été suivies en analysant les protéines présentes par migration électrophorétique sur gel en conditions dénaturantes (SDS PAGE 12 %). Après une électrophorèse, les protéines sont colorées avec le kit InstantBlue Coomassie Protein Staine (Expedeon®). L'homogénéité finale des protéines purifiées est également évaluée par spectrométrie de masse en conditions dénaturantes.

I.3.1. Lyse et extraction

Les culots humides des protéines à purifier sont repris à froid avec le Tp 1 (Figures 39 à 43) (à raison de 10 $mL.g^{-1}$ de culot). On ajoute ensuite le lysozyme (entre 0,5 et 1 $mg.mL^{-1}$) et on laisse incuber 30 minutes dans un bain marie à 30 °C. Trois cycles de congélation (azote liquide) et décongélation (bain marie 30 °C) et

5 cycles de sonication (sondes S&M-220, 40 % amplitude avec 5 secondes de pulse et 5 secondes de pause) sont pratiqués pour finir la lyse. Le lysat, est ensuite centrifugé (90 000 g ultracentrifugeuse Beckman-Coulter rotor 45Ti) à 4 °C pendant 60 min. Le surnageant est filtré à 0,45 µm et un échantillon « extrait brut » est récupéré pour les analyses de la purification finale.

I.3.2. Purification des protéines hNEIL1dc95 et dc56 et mvNei

Le surnageant récupéré lors de la lyse est chargé sur une colonne de chromatographie d'affinité de 5 mL contenant une résine métallique au cobalt ou au nickel (HisPur™ Cobalt resin, ThermoScientific/HisTrap™ HP, GE Healthcare). La colonne est lavée avec 10 volumes de colonne (CV) du tampon 1, 10 CV du tampon 2 et rééquilibrée avec 10 CV du tampon 1. Les protéines taguées (His-6) sont ensuite éluées avec 10 CV du tampon 3. Les fragments contenant la protéine sont regroupés et diluées avec le tampon 4 de façon à obtenir une concentration en NaCl de 100 mM. Avant cette étape pour hNEIL1dc56, la protéine étant beaucoup plus basique que la forme dc95, une dialyse est également effectuée sur la nuit afin de changer le tampon en tampon 3bis et d'éliminer l'imidazole. Après cette étape, l'échantillon contenant la protéine désirée est chargé sur une colonne échangeuse de cations HS20 (Poros™ HS20 8 mL, ThermoScientific), préalablement équilibrée avec 10 % de tampon 5 et 90 % de tampon 4. Les protéines sont ensuite éluées avec un gradient de concentration de 10 à 90 % de tampon 5 en 40 CV. Les fractions d'élutions analysées par SDS-PAGE contenant la protéine, sont regroupées, puis concentrées par centrifugation successive de 10 à 15 min à 4000 g à 4 °C dans des tubes Amicon®Ultra 10K (Millipore), jusqu'à environ 2 mL (selon la quantité de protéines obtenues). Cet échantillon est ensuite chargé sur une colonne d'exclusion S75 (Hiload, 16/600, 120 mL, Superdex75pg, GE Healthcare Life Sciences), prééquilibrée avec le tampon 5. Suite à cette colonne, les fractions contenant la protéine sont rassemblées et concentrées à environ 5 mg/mL (la concentration est évaluée par lecture de la densité optique à 280 nm) et filtrées avec un filtre 0,22 µm. Les protéines sont ensuite aliquotée et stockées à -80 °C pour les tests biochimiques. Pour les essais de cristallographies en revanche, les protéines sont utilisées immédiatement.

Tampons hNEIL1 dc95

Tp 1	Tp 2	Tp 3
50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 150 mM NaCl 5 mM Imidazole 10 % glycerol 2 mM β–mercapto-éthanol	50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 300 mM NaCl 10 mM Imidazole 10 % glycerol 2 mM β–mercapto-éthanol	50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 150 mM NaCl 250 mM Imidazole 10 % glycerol 2 mM β–mercapto-éthanol
Tp 4	**Tp 5**	**Tp 6**
50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 10 % glycerol 1 mM DTT	50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 1M NaCl 1mM DTT	20 mM Hepes, pH7,5 300 mM NaCl 10 % glycerol 1mM TCEP 0,1 mDM EDTA

Tampons hNEIL1 dc56

Tp 1	Tp 2	Tp 3	Tp 3bis
50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 500 mM NaCl 5 mM Imidazole 10 % glycerol 2 mM β–mercapto-éthanol	50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 1 M NaCl 25 mM Imidazole 10 % glycerol 2 mM β–mercapto-éthanol	50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 500 M NaCl 250 mM Imidazole 10 % glycerol 2 mM β–mercapto-éthanol	50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 500 mM NaCl 10 % glycerol 2 mM β–mercapto-éthanol
Tp 4	**Tp 5**	**Tp 6**	
50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 10 % glycerol 1 mM DTT	50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 1M NaCl 10 % glycerol 5mM DTT	20 mM Hepes, pH7,5 150mM NaCl 10 % glycerol 1 mM TCEP	

Tp hNEIL1 mvNei

Tp 1	Tp 2	Tp 3
50 mM KP pH 7,5 500 mM NaCl 10 mM Imidazole 10% glycerol PMSF 0,1mM/DTT 1mM/EDTA 1mM	50 mM KP pH 7,5 500 mM NaCl 20 mM Imidazole 10% glycerol PMSF 0,1mM/DTT 1mM/EDTA 1mM	50 mM KP pH 7,5 500 mM NaCl 250 mM Imidazole 10% glycerol PMSF 0,1mM/DTT 1mM/EDTA 1mM
Tp 4	**Tp 5**	**Tp 6**
20 mM Hepes pH 7,6 10% glycerol PMSF 0,1mM/DTT 1mM/EDTA 1mM	20 mM Hepes pH 7,6 1 M NaCl 10% glycerol PMSF 0,1mM/DTT 1mM/EDTA 1mM	20 mM Hepes pH 7,6 500 mM NaCl 10% glycerol 1mM DTT

Figure 39 : Composition des tampons utilisés pour la purification des protéines hNEIL1dc56 et dc95

I.3.3. Purification de la protéine pabAGOG

Pour la purification de PabAGOG, nous avons suivi les mêmes étapes générales que les purifications décrites ci-dessus. Une étape de lavage supplémentaire pour la résine cobalt a cependant été effectuée avec 10 CV de tampon 2bis. De plus, le tag Histidine a été clivé par la trypsine (0,2 mg/mL trypsine pour 1,5 mg/mL dans Tp 5 sur la nuit) avant la charge sur la S75. Suite à cette colonne, les fractions contenant la protéine sont rassemblées et concentrées à environ 30 mg/mL dans le tampon 6.

Tp 1	Tp 2	Tp 2bis	Tp 3
20 mM Hepes pH 7,6 500 mM NaCl 5 mM Imidazole Protease Inhibitor Cocktail™ (Merck)	20 mM Hepes pH 7,6 500 mM NaCl 20 mM Imidazole	20 mM Hepes pH 7,6 2 M NaCl	20 mM Hepes pH 7,6 500 mM NaCl 200 mM Imidazole
Tp 4	**Tp 5**	**Tp 6**	
20 mM Hepes pH 7,6	20 mM Hepes pH 7,6 1 M NaCl	20 mM Hepes pH 7,6 250 mM NaCl 1 mM TCEP	

Figure 40 : Composition des tampons utilisés pour la purification de pabAGOG

I.3.4. Purification de la protéine d2hOGG1

La lyse des culots de la protéine d2hOGG1 ainsi que la charge et l'élution sur une résine métallique au cobalt ou nickel a été effectuée en suivant le protocole décrit pour hNEIL1dc95, avec les tampons décrit en Figure 41. Ensuite, les fragments contenant la protéine sont regroupés et le tampon est changé par dialyse en tampon 3bis. L'échantillon est concentré jusqu'à une DO_{280nm} = 2. Pour dhOGG1, on rajoute environ 1 mL de Thrombine (Millipore®) diluée au 1/100ème pour 10 mL de protéine, et le tube est placée sous agitation à 4 °C sur la nuit. Le lendemain, l'échantillon est filtré avec un filtre 0,22 µm, puis chargé sur une colonne échangeuse de cation SPHP (HiTrap^{TM} SPHP, GE Healthcare), préalablement équilibrée avec 10 % de tampon 5 et 90 % de tampon 4. Les protéines sont ensuite éluées avec un gradient de concentration de 10 à 80 % de tampon 5 en 50 CV. Les fractions d'élutions analysées par SDS-PAGE contenant la protéine, sont regroupées, puis concentrées par centrifugation successive de 10 à 15 min à 4000 g à 4 °C dans des tubes Amicon®Ultra 10K (Millipore), jusqu'à environ 2 mL (selon la quantité de protéines obtenues). Cet échantillon est ensuite chargé sur une colonne d'exclusion S75 (Hiload, 16/600, 120 mL, GE Healthcare Life Sciences), prééquilibrée avec le tampon 5. Suite à cette colonne, les fractions contenant la protéine sont rassemblées, concentrées à environ 20 mg/mL et filtrées avec un filtre 0,22 µm. La protéine est ensuite aliquotée et stockée à -80 °C pour les tests biochimiques. Pour les essais de cristallographie, les protéines sont utilisées immédiatement. Afin de faire varier la concentration en sel et en glycérol des échantillons pour les essais cristallographiques, les tampons 6bis et 6bis2 ont été utilisés pour d2hOGG1. Pour certains tests cristallographiques, l'étape de clivage du tag histidine n'a pas été effectuée.

Tp 1 50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 300 mM NaCl 5 mM Imidazole 10 % glycerol 2 mM β–mercapto-éthanol	**Tp 2** 50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 500 mM NaCl 15 mM Imidazole 10 % glycerol 2 mM β–mercapto-éthanol	**Tp 3** 50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 300 mM NaCl 200 mM Imidazole 10 % glycerol 2 mM β–mercapto-éthanol
Tp3 bis 50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 100 mM NaCl 10 % glycerol 10 mM β–mercapto-éthanol	**Tp 4** 20 mM MES-HCl ph 6,0 10 % glycerol 10 mM β–mercapto-éthanol	**Tp 5** 20 mM MES-HCl ph 6,0 1 M NaCl 10 % glycerol 10 mM β–mercapto-éthanol
Tp 6 20 mM MES-HCl ph 6,0 500 mM NaCl 10 % glycerol 10 mM β–mercapto-éthanol 0,5 mM EDTA	**Tp 6 bis** 20 mM MES-HCl ph 6,0 10 % glycerol 10 mM β–mercapto-éthanol 0,5 mM EDTA	**Tp 6 bis2** 20 mM MES-HCl ph 6,0 500 mM NaCl 10 mM β–mercapto-éthanol 0,5 mM EDTA

Figure 41 : Composition des tampons utilisés pour la purification des protéines d2hOGG1

I.3.5. Purification de la protéine dhOGG1

La purification de dhOGG1 suit le même protocole que pour l'autre forme tronquée d2hOGG1, avec les tampons indiqués en Figure 42. En revanche le Tag est clivé en utilisant 30 ng/µL d'entérokinase à 4 °C sur la nuit en sortie de la colonne d'affinité. Le protocole est identique pour la colonne échangeuse d'ions et d'exclusion. En sortie de colonne S75, le tampon de la protéine est changé par concentration/dilution successives pour le tampon 7. La protéine est finalement concentrée à environ 5 mg/mL, filtrée avec un filtre 0,22 µm, aliquotée et conservées à -80 °C pour les tests biochimiques.

Tp 1 50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 300 mM NaCl 5mM Imidazole 10 mM β–mercapto-éthanol	**Tp 2** 50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 500 mM NaCl 20 mM Imidazole 10 mM β–mercapto-éthanol	**Tp 3** 50 mM Na_2HPO_4/NaH_2PO_4, pH8,0 300 mM NaCl 200 mM Imidazole 10 mM β–mercapto-éthanol
Tp3 bis 20 mM Tris-HCl pH 8,0, 300mM NaCl,	**Tp 4** 10 mM MES-HCl ph 6,0 10 mM β–mercapto-éthanol	**Tp 5** 20 mM MES-HCl ph 6,0 1 M NaCl 10 mM β–mercapto-éthanol
Tp 6 10 mM MES-HCl ph 6,0 500 mM NaCl 10 mM β–mercapto-éthanol	**Tp 7** 10 mM MES-HCl ph 6,0 500 mM NaCl 10 mM β–mercapto-éthanol 40 % glycerol	

Figure 42 : Composition des tampons utilisés pour la purification des protéines dhOGG1

I.3.6. Purification de la protéine mOGG1

Nous avons utilisé le protocole décrit par Visnes *et al.*, 2018 [371]. Il est à noter que la lyse des bactéries a été testée avec le BugBuster Protein Extraction Reagent®, la benzonase nucléase et le cocktail « complete protéase inhibitor (EDTA-free) » comme indiqué sur cette publication, mais les résultats n'ont pas été concluants. Nous avons donc utilisé la technique de lyse précédemment décrite.

Les étapes de lyse et les colonnes de chromatographie d'affinité (Nickel), échangeuse d'ions (SPHP) et d'exclusion (S75) sont similaires à celles utilisées pour d2hOGG1. Cependant il n'y a pas d'étape de clivage du Tag (6-His). La protéine est finalement concentrée jusqu'à environ 20 mg/mL (la concentration est évaluée par lecture de la densité optique à 280 nm) et filtrée avec un filtre 0,22 µm. Les protéines sont ensuite aliquotées et stockées à -80 °C pour les tests biochimiques. Pour les essais de cristallographies en revanche, les protéines sont utilisées immédiatement.

Tp 1 25 mM K_2HPO_4/KH_2PO_4, pH7,6 250 mM NaCl 10 mM Imidazole 5 % glycerol	**Tp 2** 25 mM K_2HPO_4/KH_2PO_4, pH7,6 250 mM NaCl 200 mM Imidazole 5 % glycerol	**Tp 3** 25 mM K_2HPO_4/KH_2PO_4, pH7,4 50 mM NaCl 10 % glycerol
Tp 4 25 mM K_2HPO_4/KH_2PO_4, pH7,4 1 M NaCl 10 % glycerol	**Tp 5** 25 mM K_2HPO_4/KH_2PO_4, pH7,4 250 mM NaCl 10 % glycerol	

Figure 43 : Composition des tampons utilisés pour la purification des protéines mOGG1

II. Préparation des oligonucléotides

Pour les différentes étapes du criblage, nous avons utilisé des ADN double brin que nous avons hybridé à partir d'oligonucléotides simple brin commandés chez Eurogentec. La séquence de ces oligonucléotides est rappelée en Annexe 1. Ces simples brins sont repris dans de l'eau ultrapure pour une concentration de 1 mM.

De manière générale, on hybride le brin contenant la lésion avec un excès de brin complémentaire, afin de s'assurer que la totalité du brin contenant le dommage est hybridé. En effet, si un simple brin contenant la lésion n'est pas hybridé, celui-ci pourra interférer avec les activités des ADN glycosylases, en étant lui-même un substrat.

II.1. Hybridation des oligonucléotides fluorescents

Pour la préparation des ADN substrats fluorescents, qui serviront lors du criblage, mais également lors des études de cinétique ou pour l'évaluation du K_D de pabAGOG, nous avons utilisé les simples brins 17mer_SUP_5'TAMRA_NEIL et 17mer_SUP_5'TAMRA_Ogg en brin contenant la lésion et le fluorophore, et le simple brin 17mer_INF_3'_BHQ2 en brin complémentaire. La séquence de ces oligonucléotides est rappelée en Annexe 1. Nous avons hybridé ces brins d'ADN avec un ratio de 1/1,2 en fluorophore/complémentaire, dans de l'eau ultrapure. Le mélange d'ADN est séché pendant environ une heure et demie à vide et à -50 °C dans un Speed-Vac. Ensuite, l'extrait sec est repris avec 50 µL du tampon ADN (10 mM Tris, pH 8,5, 1 mM EDTA, 100 mM NaCl), afin d'obtenir une concentration finale de 100 µM. Le tube est placé à -4 °C pendant la nuit pour faciliter l'hybridation, puis conservé à -20 °C.

II.2. Hybridation pour études cristallographiques

La cristallogenèse dépend de nombreux paramètres. Afin d'obtenir des structures de haute résolution nous avons donc du tester différentes conditions. Ainsi une bonne affinité entre l'ADN et la protéine peut permettre de favoriser la stabilité et améliorer l'empilement de ce complexe pour la formation de cristaux. Pour la cristallogenèse des complexes ADN glycosylase/ADN nous avons préparé des ADN double brin à partir des oligonucléotides décrits en Annexe 1. Afin de stabiliser un potentiel empilement des doubles

brins d'ADN, nous avons choisi d'utiliser des duplexes à « bords cohésifs », c'est-à-dire avec un décalage d'une base sur chaque bord. Le brin complémentaire est toujours mis en excès avec un ratio de 1,2 face au brin contenant la lésion. Les deux simples brins complémentaires sont dissouts dans de l'eau ultrapure, et le mélange est séché à l'aide d'un Speed-Vac pendant 1,5 h environ. Le temps de séchage peut varier selon la quantité de volume utilisé. Le culot obtenu est repris dans du tampon d'hybridation (10 mM Tris, pH 8,5, 1 mM EDTA, 100 mM NaCl) afin d'avoir une concentration de 2 mM finale. Ensuite, ce mélange est placé dans un bain marie chauffé à 90 °C, dont on fait lentement descendre la température jusqu'à 4 °C. L'ADN est ensuite conservé à -20 °C.

La complexation de ces ADN avec les protéines a été effectuée en respectant un ratio 1,2/1, 1,5/1 ou 2/1 où la proportion en duplex d'ADN est majoritaire. Les échantillons (protéines et ADN) sont incubés 30 min dans la glace avant d'être utilisés en cristallographie.

II.3. Marquage des oligonucléotides au ^{32}P

Ce type de marquage est réalisé pour nos tests d'activité non-fluorescents. Seul le brin contenant la lésion est marqué. Cela va permettre par la suite de quantifier les brins clivés par les ADN glycosylases des brins non clivés, et donc d'évaluer l'activité de l'enzyme. Selon le type d'activité et la protéine désirée pour ce test, différents duplexes peuvent être utilisés. Pour évaluer l'activité glycosylase-lyase de hNEIL1 par exemple, nous avons utilisé le duplexe Act_Neil1, contenant la lésion DHU:G, tandis que pour évaluer les activités glycosylase et glycosylase-lyase de hOGG1, pabAGOG ou LlFpg nous avons utilisé le duplexe Act_OGG1, contenant la lésion 8-oxoguanine:C (Annexe 1). Pour évaluer l'activité AP lyase de pabAGOG et de hOGG1 nous avons utilisé le duplexe Act_ApLyase que nous avons traité par la suite avec de l'uracile DNA glycosylase (UDG) afin d'obtenir un site abasique. Pour chaque type de duplexe désiré, 100 picomoles de simples brins (soit 24mer_OxoG/24mer/DHU) sont mélangés avec 20 U de PolyNucléotide Kinase (PNK) (M02015 NE Biolabs) et 20 µCi de [γ-32P]- Adénosine triphosphate (ATP, Perkin Elmer) dans un tampon 1X Tp PNK (B02015 NEBiolabs). La source est manipulée dans un environnement adéquat de radioactivité et en portant les différents dosimètres nécessaires (torse + main). On incube ensuite pendant 30 minutes à 37 °C, puis le mélange réactionnel est passé sur une minicolonne (Spin Column Wash tubes, BioRad), préalablement lavée 3 fois avec 100 µl d'eau ultrapure. On centrifuge la minicolonne pendant 5 min à 1000 g, et on récupère le mélange en sortie de colonne. On rajoute ensuite 160 picomoles du brin complémentaire 24mer_comp (on se place en excès de brin complémentaire pour les mêmes raisons que pour les autres hybridations, afin d'assurer l'hybridation complète du brin contenant la lésion), et on complémente le tampon avec 1X de tampon tris-EDTA, afin d'obtenir un mélange d'une concentration de 1 µM de duplexe. Ce mélange est incubé 5 minutes à 75 °C, puis laissé à 4 °C pendant la nuit, et stocké à -20 °C.

La procédure pour préparer le duplexe présentant un site abasique est identique sur les premières étapes, soit l'ajout du brin 24mer_U avec la PNK, le [γ-32P]-ATP et les différents tampons. Après 30 minutes d'incubation à 37 °C, et la centrifugation du mélange sur minicolonne, on rajoute 2 U de UDG (EN0361, ThermoScientific) dans un tampon de réaction (Tris-HCl 200 mM, pH 8,2, EDTA 10 mM, NaCl 100 mM). Ce nouveau mélange est de nouveau incubé 30 minutes à 37 °C. Ensuite, l'ajout du brin complémentaire et l'hybridation finale est identique aux autres ADN double brin.

III. Criblage

III.1. Chimiothèques utilisées

III.1.1. Choix des banques de molécules

Nous avons choisi 4 banques de molécules différentes à cribler. Tout d'abord il est capital d'introduire le fait qu'aucun criblage n'avait été effectué dans l'équipe où j'ai effectué cette thèse. Ainsi, nous ne disposions pas des appareils usuels pour un criblage à haut débit par une industrie ou un laboratoire spécialisé (robot pipeteur adapté, préparation et stockage des microplaques automatiques, logiciels conçus pour le suivi des résultats, etc.). Nous avons donc dû faire des choix rationnels afin d'effectuer ce criblage. Tout d'abord nous avons dû privilégier l'originalité des composés à leur quantité. En effet, la plupart des criblages HTS peuvent cribler des centaines de milliers de composés facilement, ce qui n'était pas notre cas. Nous avons donc choisi d'utiliser deux banques de diversité, ICOA-DIVER une banque de diversité de l'Institut de Chimie Organique et Analytique d'Orléans (que nous appellerons **ICOA1** pour la suite) et GPNCL, une banque de l'entreprise GreenPharma, que nous nommerons **GP**. Chacune de ces banques est composée de 480 molécules. L'avantage de ces banques dites de « diversité » réside dans le fait que chaque composé est supposé représenter une famille structurale de molécules dans une banque de composés plus importante (2 millions de molécules pour Ambinter, la chimiothèque totale de GreenPharma). Ainsi on diminue le nombre de composés testés, tout en conservant une diversité structurale importante. Le second avantage du choix de cribler ces banques reposent sur la proximité des chimistes ayant participé à l'élaboration de celles-ci. L'ICOA et GreenPharma sont installés tous les deux sur le campus d'Orléans, ce qui a permis de justifier un financement régional soutenant le projet, mais cela permet aussi de faciliter le développement d'analogues de molécules identifiées. En ce sens, nous avons aussi utilisé la banque totale de l'ICOA, composés de 2480 molécules (que nous appellerons **ICOA2** pour la suite). Afin de compléter l'étendue des molécules testées, nous avons choisi d'utiliser la banque Prestwick Chemical Library de Prestwick (**Prest**). Cette chimiothèque est très souvent utilisée dans les criblages classiques, car elle est composée de 1180 médicaments hors brevet (« off patent » en anglais) présentant une grande diversité chimique et pharmacologique et dont la biodisponibilité et l'innocuité sont bien connues chez l'Homme (composés ayant franchis la phase 1 des études cliniques). Ces composés ont donc vu leur toxicité être déjà évaluée et cela peut faciliter la suite des études. Le plan général de ces plaques ainsi que la composition des chimiothèques est indiqué en Annexe 2.

III.1.2. Préparation des plaques de criblage

La composition des banques utilisées est indiquée en Figure 44. Ces 4 banques sont réparties dans des plaques de 96 puits, à raison de 80 composés par plaque, dilués dans 100 % de DMSO. Afin de ne pas contaminer les plaques mères lors du criblage, des plaques filles ont été préparées par le robot pipetteur Freedom EvoTecan dans des plaques de 96 puits en polypropylène à fond conique (Thermo Scientific Nunc 249944). Ce robot est principalement utilisé pour la préparation des plaques de criblage cristallographique, et son volume est adapté pour la préparation des plaques filles. Pour les deux banques fournies en concentration massique (GP et ICOA1) aucune dilution n'a été effectuée pour la préparation de ces plaques filles. En revanche pour les chimiothèques fournies en concentration molaire, Prestwick et ICOA2, des dilutions dans 100 % DMSO ont été effectuées. En effet, s'il est compliqué de comparer ces différentes banques entres elles au vu de leurs concentrations différentes, nous avons estimé, au vu de la masse moléculaire moyenne des composés de GP et de la faible concentration de ICOA1, que les deux autres chimiothèques étaient beaucoup plus concentrées, et risquaient ainsi de sortir trop de composés peu efficaces. Les différents contrôles (positifs ou négatifs) ont aussi été rajoutés dans les plaques filles selon le

plan indiqué en Annexe 2. En tout, 4560 composés ont donc été préparés pour le criblage sur hNEIL1 et hOGG1.

Chimiothèques				
Provenance	GreenPharma	Pretwick	ICOA	
Nom	GPNCL	Prestwick Chemical Library	ICOA-DIVER	ICOA-TT
Type	Banque de diversité	Banque de diversité	Banque de diversité	Banque totale
Abréviation	GP	Prest	ICOA1	ICOA2
Nombres de molécules	480	1120	480	2480
Concentration plaque mère	2 $mg.mL^{-1}$	10 mM	0,3 $mg.mL^{-1}$	20 mM
Concentration plaque filles	2 $mg.mL^{-1}$	1,25 mM	0,3 $mg.mL^{-1}$	1,25mM
Concentration crible	160 $\mu g.mL^{-1}$	50 µM (x 2)	24 $\mu g.mL^{-1}$	50 µM (x 2)

Figure 44 : Composition des chimiothèques utilisées et leurs plaques filles

III.2. Conditions criblage primaire

III.2.1. Principe

Afin d'identifier des inhibiteurs originaux de hNEIL1 et hOGG1, nous avons opté pour une stratégie de criblage biochimique. Pour ce faire, nous nous sommes largement inspirés de l'essai mis en place par Jacobs *et al.,* 2013 [416], qui exploitait l'activité glycosylase/AP-lyase de la protéine hNEIL1. Le principe est schématisé sur la Figure 45. Un oligonucléotide de 17 bases contenant un substrat reconnu par hOGG1 ou hNEIL1 est hybridé avec un brin complémentaire. Sur le brin contenant la lésion, un fluorophore (TAMRA) est présent à l'extrémité 5', tandis que sur le brin complémentaire, un extincteur de fluorescence (« quencher » en anglais) (BHQ2) est opposé au fluorophore à son extrémité 3'. Le TAMRA possède en effet un pic d'émission de fluorescence à 583 nm tandis que BHQ2 a un spectre d'absorption dont le maxima est à 579 nm. Le spectre d'émission du TAMRA et d'absorption de BHQ2, ainsi que leurs structures sont rappelés en Annexe 3. Lors d'une hybridation entre ces deux brins, BHQ2 va donc quencher le signal produit par le fluorophore. En revanche, en présence d'une activité glycosylase-lyase associée aux protéines hNEIL1 ou hOGG1, l'oligonucléotide clivé par l'enzyme sera libéré durant l'incubation à 37 °C et une émission significative de fluorescence devrait être mesurable. En présence d'une molécule inhibitrice de l'enzyme, on doit s'attendre dans les mêmes conditions à empêcher l'apparition de l'émission de fluorescence.

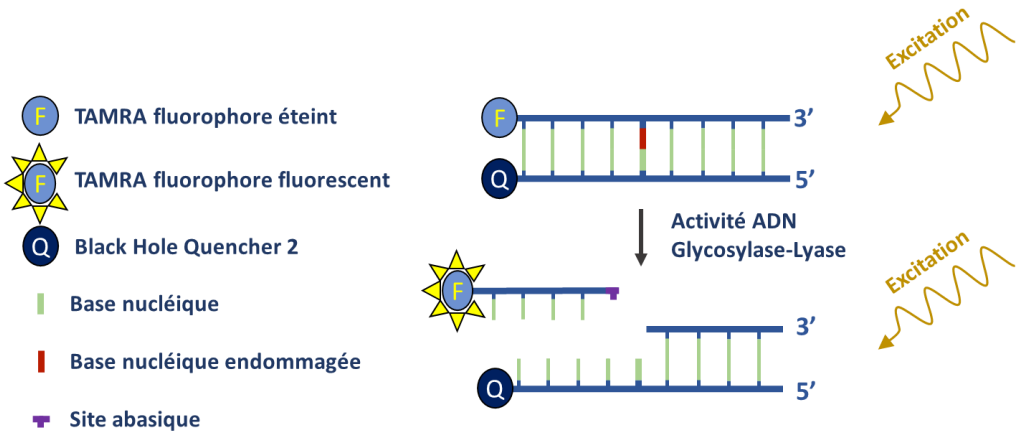


Figure 45 : Schéma du test d'activité à la fluorescence pour hNEIL1 ou hOGG1

Le principe du criblage repose donc sur l'incubation de l'ADN glycosylase en présence des molécules à cribler, puis, après quelques minutes d'incubation l'ADN substrat est rajouté au mélange réactionnel. Dans ces conditions, le composé testé n'est pas en réelle compétition avec le substrat. Comme indiqué précédemment, chaque plaque d'inhibiteurs est composée de 80 molécules réparties dans une plaque de 96 puits. Ces composés seront transférés dans une plaque de 384 puits en plusieurs réplicas, auquel on ajoutera la protéine hNEIL1 ou hOGG1 ainsi que divers tampons. A ce stade, il semble important de rappeler que notre criblage ne prétend pas atteindre le débit nécessaire à un HTS. En effet, si l'ajout des inhibiteurs et du substrat est effectué par un robot de pipetage Mosquito, après beaucoup d'optimisation manuelle que nous détaillerons dans la partie résultat, les plaques de 384 puits contenant la protéine à cribler ont toutes été faites manuellement, ainsi que l'ajout et le mélange des différents tampons. De même l'addition des contrôles positifs et négatifs ont été fait manuellement avec une micro-pipette. Les différents plans de plaque utilisés lors de cette étape sont décrits en Annexe 2. Pour augmenter la fiabilité du crible, nous disposons sur notre plaque de plusieurs contrôles :

- **Contrôle catalytique** : il contient un excès de protéine et l'ADN substrat à la concentration utilisée dans l'essai ce qui fournit le signal maximum de fluorescence observable ;

- **Contrôle bruit de fond** : il ne contient pas de protéine mais contient l'ADN substrat, ce qui donne donc le signal de bruit de fond du fluorophore éteint par le « quencher » ;

- **Contrôle positif** : il contient la protéine et l'ADN substrat dans les mêmes conditions que le reste du criblage en présence ou absence de DMSO, le solvant dans lequel les inhibiteurs ont été préparés ;

- **Contrôle négatif** : il contient la protéine et l'ADN substrat dans les mêmes conditions que le reste du criblage en présence d'un inhibiteur déjà connu.

Les contrôles catalytiques et du bruit de fond permettent de calculer le facteur **Z'**, tandis que les contrôles positifs et négatifs permettent d'identifier les hits du criblage. Ainsi, le pourcentage d'inhibition est calculé selon l'Equation 2 :

$$\mathrm{I(X)} = 100 * \left[1 - \frac{x - \mathrm{M_}}{\mathrm{M_+} - \mathrm{M_}}\right]$$

Equation 2 : Pourcentage d'inhibition I calculé pour chaque composé X - M- et M+ sont les moyennes des contrôles positifs et négatifs respectivement ; *x* correspond au signal de fluorescence obtenu.

Lorsque les plaques contenant la protéine et les molécules à tester sont prêtes, elles sont centrifugées à froid afin de s'assurer que des bulles ne se forment pas dans le puits et affectent la lecture. Ensuite, l'ADN substrat est rajouté et les plaques sont incubées à 37 °C dans le lecteur de fluorescence. La fluorescence est donc lue sur l'ensemble de la plaque après un certain temps, dépendant de l'enzyme criblée. Le pourcentage d'inhibition des composés est établi en fonction de leurs niveaux de fluorescence.

III.2.2. Types de plaques utilisées

Le criblage primaire a été optimisé pour des plaques de 384 puits. Au cours de cette thèse, différents types de plaques ont été utilisés. Pour le criblage manuel de la banque GP sur LlFpg lors de la preuve de concept, des plaques de 384 puits rond en polystyrène noir à fond plat (Cornin 3820), avec un volume maximal de 50 µL ont tout d'abord été utilisés pour les premières étapes d'optimisation. Lors des tests d'optimisation, pour l'augmentation du volume du test (60 µL), des plaques de 384 puits carrés en polystyrène noir à fond plat (Cornin 3715), avec un volume maximal de 112 µL ont été utilisés. Enfin, pour le criblage semi-automatique de hOGG1 et hNEIL1, avec un volume final de 15 µL, de nouvelles plaques ont été utilisées, des plaques de 384 puits carrés à fond rond (Corning 4514).

III.2.3. Criblage manuel

Afin de mettre en place la preuve de concept, le criblage a été effectué manuellement. Cette preuve de concept a été mise en place sur l'enzyme de *Lactoccocus lactis* LlFpg, homologue structural de hNEIL1 et fonctionnel de hOGG1. Dans chaque puits, la protéine est incubée à 5 nM avec les composés des plaques filles de 96 puits à hauteur de 8 % DMSO final dans un tampon de réaction (10 mM Hepes, pH 7,6, 120 mM NaCl, 5 % glycérol, 0,2 mM EDTA, 0,1 mg/mL BSA). Le signal catalytique est obtenu pour 250 nM LlFpg dans les conditions précédentes. Après 5 min d'incubation à température ambiante, l'ADN **TMR_OGG1** est ajouté à 50 nM et le mélange est incubé à 37 °C pendant 15 minutes. La fluorescence est ensuite lue avec le lecteur de plaque Victor™ 3V (Perkin Elmer).

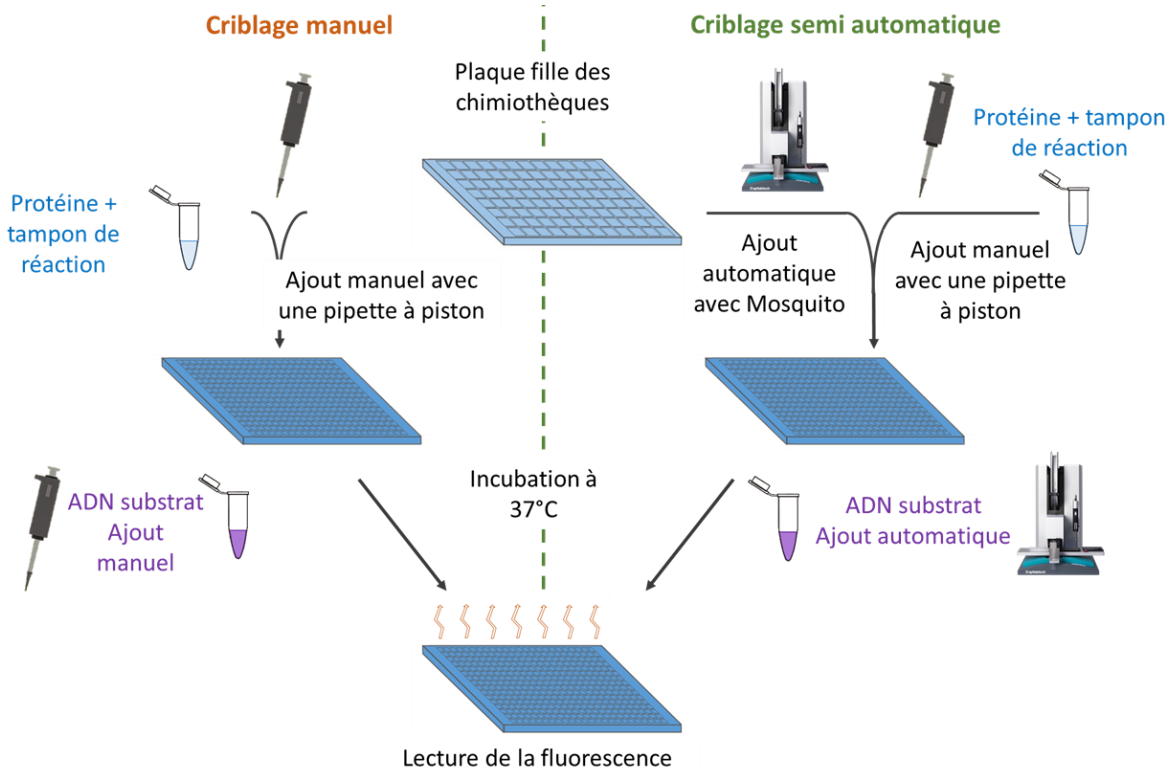


Figure 46 : Schéma du criblage primaire manuel ou semi-automatique

III.2.4. Criblage semi-automatique

Par la suite, le procédé de criblage a été optimisé pour hNEIL1 et hOGG1, et est donc un criblage semi-automatique. Le procédé est identique, mais certaines étapes sont plus rapides et automatisées (Figure 46).

Pour le criblage de hNEIL1, nous avons utilisé la forme entière avec le tag Histidine (Annexe 4). La protéine est ajoutée manuellement dans chaque puits d'une plaque de 384 puits, avec le tampon de réaction à une concentration de 100 nM. Chaque plaque est ensuite centrifugée à 4 °C pendant 1 min à 2000 g (Eppendorf Centrifuge 5804R), afin d'éviter la formation de bulles dans les puits. Grâce à un robot pipeteur Mosquito (sptlabtech, Mosquito® Xtal3), les composés des plaques filles de 96 puits sont rajoutés à raison de 8 % DMSO final. Les plaques sont de nouveau centrifugées à 4 °C pendant 1 min à 2000 g. Après une incubation pendant 5 min à température ambiante, 100 nM de l'ADN **TMR_NEIL1** est ajouté avec le Mosquito pour lancer la réaction. La fluorescence est lue grâce au lecteur Victor™ 3V (Perkin Elmer), après une incubation de 20 minutes à 37 °C. Les conditions finales de ce criblage sont donc les suivantes :

100 nM de **hNEIL1** / 100 nM de **TMR_NEIL1** avec un tampon final de 20 mM Hepes pH 7,5, 150 mM NaCl, 5 % glycérol, 0,1 mM TCEP, 0,2 mM EDTA, 0,1 mg/mL BSA et 8 % DMSO, avec une concentration d'inhibiteurs variable.

Pour le second criblage, nous avons utilisé la protéine tronquée dhOGG1 (Annexe 4). Les différentes étapes sont les mêmes que pour le criblage semi-automatique de hNEIL1, hormis une incubation à 37 °C de 30 minutes après l'ajout de l'ADN. La concentration de protéine utilisée est également différente, pour des conditions finales suivantes :

100 nM de **dhOGG1** / 50 nM de **TMR_OGG1** avec un tampon final de 17,6 mM Hepes pH 7,5, 118 mM NaCl, 7,4 % glycérol, 1 mM β-mercaptoéthanol, 0,2 mM EDTA, 8 % DMSO avec une concentration d'inhibiteurs variable.

Pour ces deux criblages, les contrôles du signal catalytique (signal maximum obtenu) ainsi que ceux du bruit du fond ont été effectués à la pipette, hormis l'ajout d'ADN qui a été effectué par le Mosquito. Le signal catalytique a été obtenu avec 2 µM de hNEIL1 ou de hOGG1 en présence de 100 nM et 50 nM d'ADN substrat respectivement.

III.3. Test d'activité orthogonal

III.3.1. Principe

Pour confirmer les résultats obtenus lors du premier criblage, un deuxième test d'activité doit donc être mis en place. Le principe de ce test est schématisé sur la Figure 47. De la même manière que pour le criblage, un ADN contenant une lésion reconnue par hNEIL1 ou hOGG1 est utilisé (dont les séquences sont indiquées en Annexe 1). En revanche à la place du fluorophore, cet ADN a été marqué en 5' avec un phosphate radioactif ^{32}P. Ainsi après coupure de l'ADN, on peut donc « suivre » la radioactivité du fragment marqué ou du brin complet marqué.

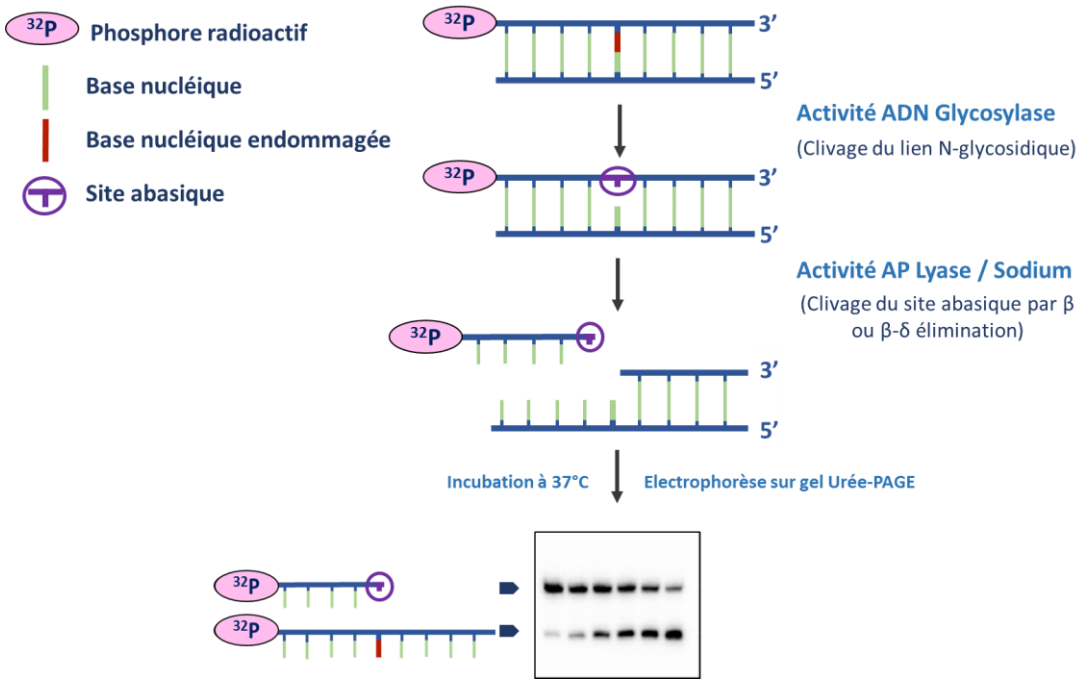


Figure 47 : Test de confirmation à la radioactivité du criblage primaire

Pour le test d'activité orthogonal, nous avons utilisé les duplexes marqués au ^{32}P en 5' décrits précédemment. Les conditions des tests d'activité dépendent de l'activité souhaitée et de la protéine utilisée, mais le principe reste similaire. Les différentes conditions sont indiquées dans un tableau en Figure 48.

Dans ce test, l'ADN substrat radioactif (**Act_NEIL1**, **Act_OGG1** ou **Act_ApLyase** traité par l'UDG) est préincubé avec le composé à tester, pour 8 % DMSO, à 37 °C pendant 5 min, dans le tampon de réaction (10 mM Tris-HCl pH 7,5, 150 mM NaCl, 1 mM EDTA, 0,1 mg/mL BSA). Contrairement au criblage, ce test s'effectue donc dans des conditions de compétition entre le substrat et l'inhibiteur.

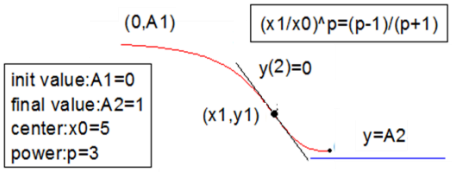
Au moment de lancer la réaction, la protéine est dans chaque tube. Immédiatement après l'ajout de la protéine, chaque tube est centrifugé quelques secondes sur une centrifugeuse de paillasse, et incubé 15 minutes à 37 °C sur un chauffe-tube Thermomix Comfort. Pour arrêter la réaction, les mélanges réactionnels sont incubés 3 minutes à 76 °C en présence de 40 % formamide final et de 0,1 % de bleu de bromophénol, un marqueur de migration électrophorétique (formant le mélange « BF »), et agité grâce à un vortex. Pour les activités glycosylases, la réaction est tout d'abord arrêtée avec l'ajout de NaOH à 0,2 µM à 50 °C pendant 5 minutes, puis ensuite par l'ajout du BF. L'ajout de la soude va ainsi permettre de cliver le squelette phosphodiester du simple brin contenant la lésion retirée par l'activité glycosylase de la protéine, se substituant ainsi à l'activité lyase.

Dans ces conditions, la protéine et l'ADN substrat sont dénaturés. Ce mélange est ensuite analysé par électrophorèse en gel de polyacrylamide (19/1) en présence d'urée (20 %) « Urea-PAGE », ce qui permet de maintenir des conditions dénaturantes durant la migration électrophorétique (500 V et 20 mA / gel pendant 20 min). Les molécules d'ADN sous leur forme simple brin sont séparées en fonction de leur taille comme indiqué dans la Figure 47. En effet, le brin endommagé non-clivé par l'enzyme (correspondant au substrat intact) migre moins loin que le brin clivé par l'enzyme (correspondant au produit de réaction). Après 1 h d'exposition sous un écran d'enregistrement luminescent au phosphore, et lecture de l'écran exposé grâce à un scanner laser Typhoon™ FLA 9500 (GE Healthcare), on peut quantifier la quantité d'ADN clivé et

celle non clivé, et donc évaluer l'activité en pourcentage des ADN glycosylases étudiées, et leur potentielle inhibition.

Ce test a également été utilisé pour établir les valeurs d'IC_{50} des inhibiteurs, en faisant varier les concentrations des composés, avec 3 réplicas par molécule. La courbe de dose-réponse qui en résulte a ensuite été « fittée » grâce au logiciel Origin, en suivant l'Equation 3, donnant ainsi accès à des valeurs d'IC_{50}. Pour rappel, une valeur d'IC_{50} est la concentration nécessaire pour atteindre 50 % d'inhibition de la protéine.

$$y = \frac{A_1 - A_2}{1 + (x/x_0)^p} + A_2$$



Equation 3 : Equation du « fitting » des courbes d'inhibition des inhibiteurs et calcul d'IC50 - A_1 étant la valeur en ordonnée initiale (0 % d'activité dans notre cas), A_2 la valeur finale attendue (100 % d'activité). La valeur d'IC_{50} est donc la valeur en ordonnée pour x= (A_1+A_2)/2 (50 % dans le cas d'un IC_{50}, mais le principe est le même pour les IC_{20} ou IC_{80}). p représente le coefficient de Hill.

Test d'activité radio pour 50 % de coupure de l'ADN substrat							
Protéine	**LlFpg**	**hNEIL1**	**hOGG1**		**pabAGOG**		
Activité	Glyoxylase-Lyase	Glyoxylase-Lyase	Glyoxylase-Lyase	Glycosylase	Glyoxylase-Lyase	Glycosylase	Ap Lyase
Concentration protéine	2 nM	20 nM	30 nM	2 nM	5 nM	2 nM	1,5 nM
ADN substrat	Act_OGG1	ActNeil1	Act_OGG1	Act_OGG1	Act_OGG1	Act_OGG1	Act_ApLyase
Concentration ADN substrat	20 nM	20 nM	20 nM	20 nM	20 nM	20 nM	20 nM
Temps d'incubation	15 min	15 min	15 min	15 min	15 min	15 min	15 min
Température	37 °C	37 °C	37 °C	37 °C	37 °C	37 °C	37 °C
Ajout de soude	X	X	X	+ 1,5µL NaOH 2M 50°C / 5 min	X	+ 1,5µL NaOH 2M 50°C / 5min	X
Arrêt de la réaction	+ 20µL BF 76°C / 3 min	+ 20µL BF 76°C / 3 min	+ 20µL BF 76°C / 3 min	+ 20µL BF 76°C / 3 min	+ 20µL BF 76°C / 3 min	+ 20µL BF 76°C / 3 min	+ 20µL BF 76°C / 3 min

Figure 48 : Différentes conditions des tests d'activité selon le type d'activité ou la protéine utilisée

III.3.2. Avantages et limitations de ce test d'activité

Ce test permet d'évaluer plusieurs types d'activité, notamment l'activité glycosylase-lyase, lyase mais aussi l'activité glycosylase des enzymes dont l'activité est fortement découplée comme hOGG1, ce qui n'était pas possible (du moins simplement) dans le test de criblage. C'est également un test compétitif, le substrat et les inhibiteurs étant mis au contact avec la protéine au même instant. Certaines études ont prouvé que l'addition du substrat en premier permettait de réduire l'effet des inhibiteurs non spécifiques en comparaison avec un protocole incubant la protéine en premier (comme c'est le cas pour le criblage que nous avons mis en place) [453]. C'est finalement l'intérêt de notre test orthogonal qui doit confirmer le statut des composés sélectionnés comme des inhibiteurs. Enfin ce test est beaucoup plus robuste et fiable que le criblage primaire, et permet donc d'avoir des doses-réponses si nécessaire. En revanche, ce test est beaucoup plus consommateur en matériel et en temps, c'est donc pour cela que nous l'avons utilisé comme test de confirmation, et non comme technique de criblage.

III.4. Test d'intercalation à l'ADN

III.4.1. L'intercalation sur l'ADN, un paramètre important à prendre en compte

La plupart des criblages nécessite un seul test de confirmation suite au criblage primaire. Cependant, notre criblage repose sur 3 éléments, la protéine d'intérêt hNEIL1 ou hOGG1, les composés des chimiothèques, mais aussi le substrat ADN. Ainsi, s'il est important de confirmer que l'effet d'inhibition est bien dû aux composés en eux-mêmes et non à leur potentiel interaction avec la technique d'analyse effectuée, la fluorescence dans notre cas, nous devons aussi nous assurer que l'effet d'inhibition que nous observons est dû à l'interaction des composés avec les ADN glycosylases, et non à l'interaction avec l'autre élément du criblage, l'ADN substrat.

Comme nous l'avons évoqué précédemment dans l'introduction, de nombreux médicaments viennent s'intercaler dans l'ADN afin d'avoir un effet thérapeutique (doxorubicine et irinotecan par exemple). Une molécule intercalante n'est donc pas à rejeter à tout prix et dans tous les cas. La plupart des molécules ayant un pouvoir d'intercalation sont toxiques, ce qui peut être utile dans certains contextes, cependant le principe thérapeutique derrière l'inhibition des ADN glycosylases repose sur celui de la **létalité synthétique**. Ainsi une molécule intercalante a fort peu de chances d'être sans effet sur une cellule saine. De plus, l'effet du composé provenant de son intercalation à l'ADN, celui-ci risque fort d'être non sélectif, ce qui est un autre souci pour notre étude. Pour toutes ces raisons, nous avons donc effectué un 3ème test afin d'évaluer le pouvoir intercalant de l'ADN de nos composés. Ce test servira à classer les composés inhibiteurs selon leur « force » d'intercalation, et de sélectionner les molécules finales du criblage.

III.4.2. Principe du test

Pour effectuer ce test, nous avons choisi d'utiliser la solution **PicoGreen « PG »** (Quant-iT^{TM} Picogreen^{TM} ds DNA reagent, Invitrogen). Cette solution, dont la structure a été publiée [454], est un composé fluorescent très sensible, utilisé principalement pour la quantification d'ADN double brin [455], mais aussi pour de très nombreuses applications, notamment l'interaction de composés médicamenteux sur l'ADN [456,457]. Le PG, appartenant à la famille des cyanines, est un composé qui s'intercale dans l'ADN et qui va produire une forte fluorescence après son intercalation (un peu comme le bromure d'éthidium). Le principe de ce test repose donc sur la fluorescence apparaissant lors de l'intercalation du PG sur un ADN double brin (dont la séquence est rappelée en Annexe 1) (Figure 49).

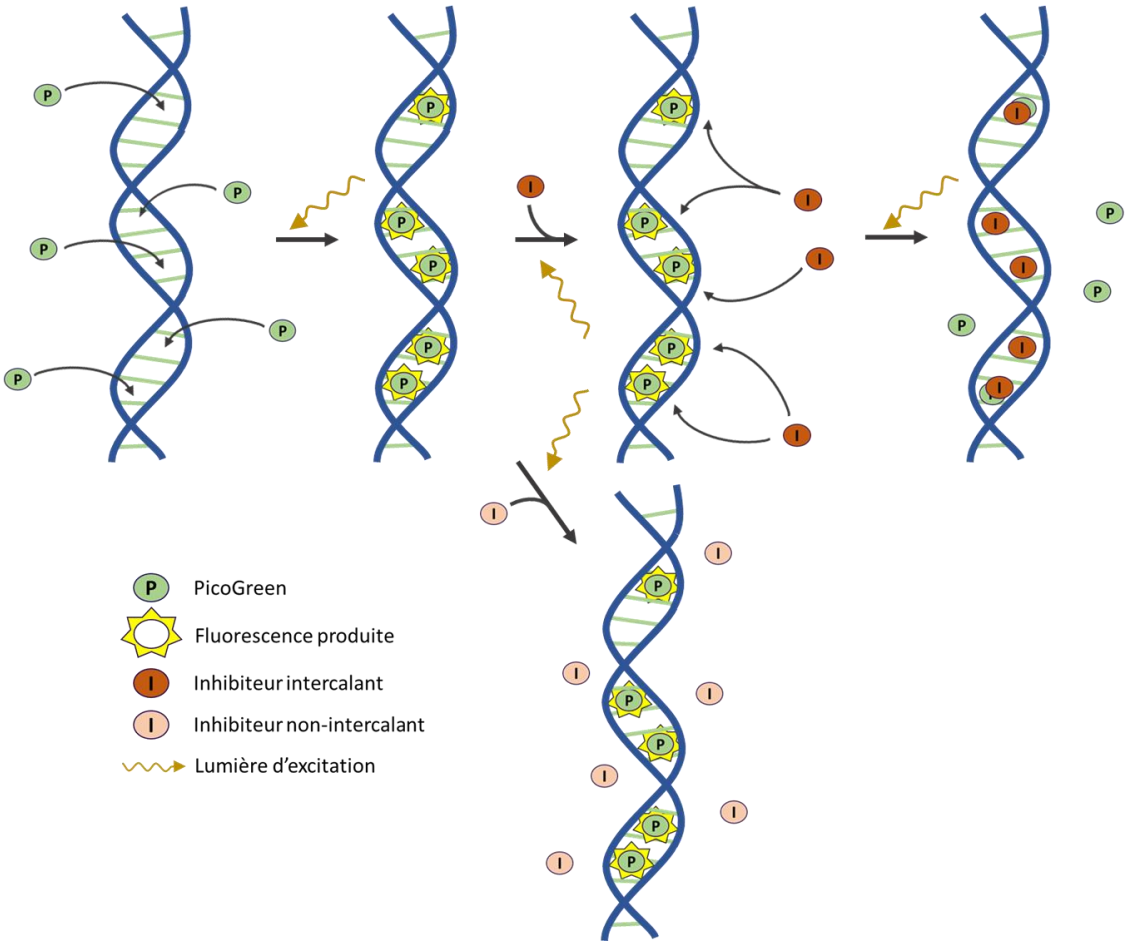


Figure 49 : Schéma représentant le principe du test d'interaction à l'ADN par nos composés en présence de PicoGreen

Concrètement, une solution de travail de PG diluée au 1/100ème est incubée avec un ADN double brin à une concentration de 80 nM dans un tampon de réaction 10 mM TrisHCl, 1 mM EDTA pH 7,5. La solution PicoGreen se dégradant à la lumière, un film d'aluminium est placé autour de tous les tubes. Le mélange est mis à température ambiante pendant 5 minutes, et la fluorescence est lue sur un NanoDrop™ 1000 v3.8 (Thermo Scientific). Par la suite, on rajoute un volume croissant des inhibiteurs. Entre chaque ajout, le mélange est agité doucement, centrifugé, et laissé 5 min à température ambiante. Les ajouts successifs sont indiqués dans la Figure 50 ci-dessous. Si le composé s'intercale fortement avec l'ADN, il remplacera le PG, et la fluorescence diminuera en conséquence. Un composé intercalant peut aussi « éteindre » (« quencher ») la fluorescence de PicoGreen en interagissant avec lui. On admettra que la baisse de la fluorescence induit une interaction de nos composés sur l'ADN, et ces composés ne seront donc pas retenus.

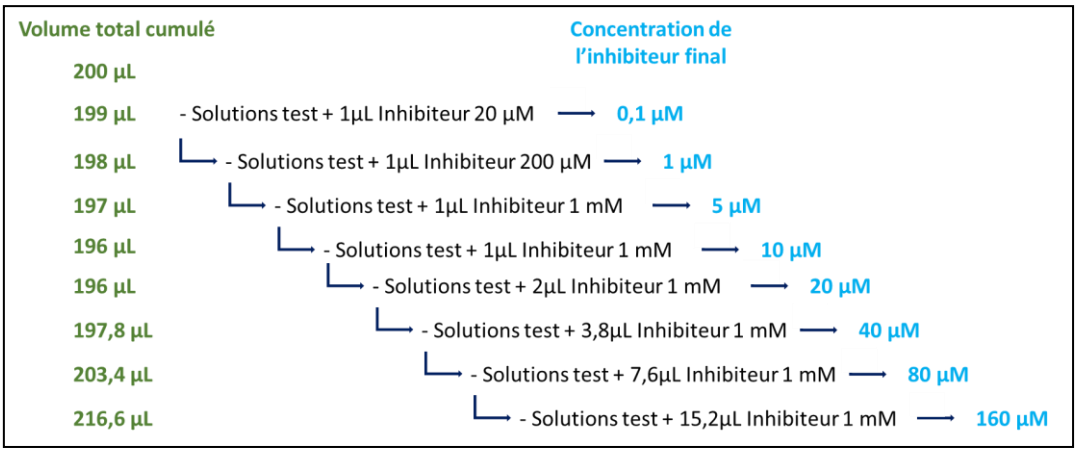


Figure 50 : Ajouts successifs d'inhibiteurs pour le test d'intercalation à l'ADN

IV. Caractérisation des inhibiteurs

IV.1. Cinétique

IV.1.1. Principe général et équations principales

Une des caractérisations usuelles de l'activité enzymatique d'une protéine est d'évaluer la cinétique de son activité. Pour ce faire nous allons utiliser l'équation de Michaelis-Menten. Celle-ci requiert qu'une enzyme E forme avec son substrat S un complexe ES. Ce complexe peut ensuite se re-dissocier en E et S, ou former le produit P, comme décrit en Figure 51. Dans ce postulat, nous supposons que la formation de ce produit est irréversible et les valeurs k_1, k_{-1} et k_2 seront les constantes de vitesse de cette réaction.

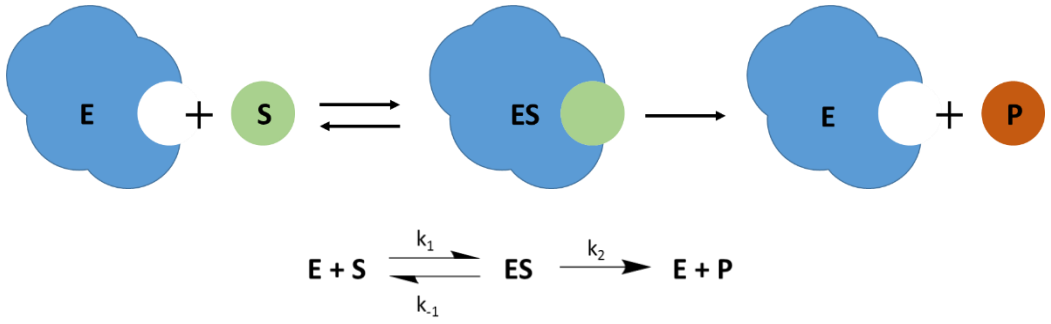


Figure 51 : Représentation graphique de l'activité d'une enzyme sur un substrat et les constantes associées

La vitesse d'une réaction à un instant *t* correspond donc à la tangente de la pente de la courbe d'apparition du produit (Figure 52). En début de réaction cette vitesse est quasiment constante (dans la phase dite stationnaire) et correspond à la vitesse initiale, notée v_0. Pour les enzymes qui suivent cette règle, il est donc possible de tracer une courbe représentant la vitesse initiale de la réaction en fonction de la concentration de l'enzyme. Ainsi, à basse concentration, la vitesse augmente rapidement, jusqu'à atteindre un plateau à plus haute concentration, qui correspond à la vitesse maximale de l'enzyme, indépendant de la concentration de l'enzyme, voir Figure 52.

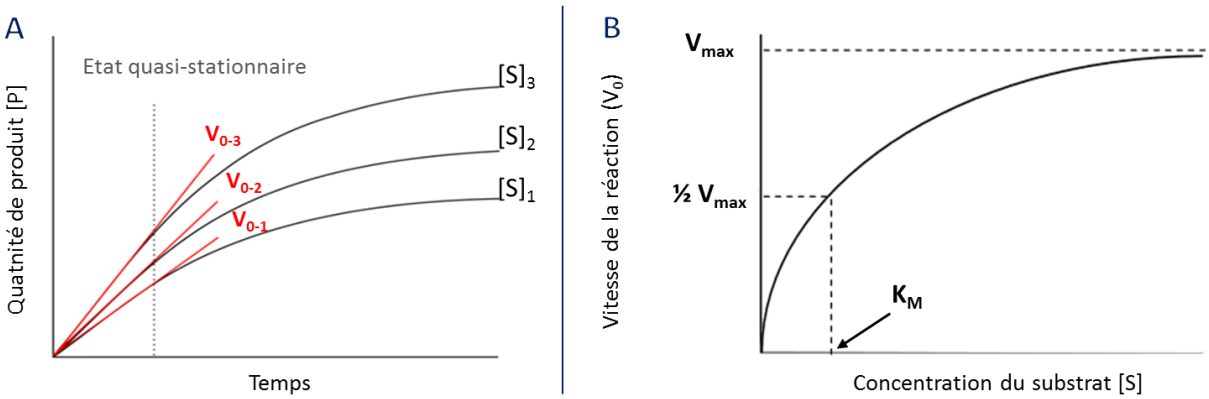


Figure 52 : **Courbes théoriques pour l'évaluation de la cinétique d'une réaction enzymatique -** A) Tracé de l'apparition du produit en fonction du temps, permettant avec la tangente à l'origine de déterminer les vitesses initiales. Chaque courbe correspond à une quantité de substrat différente et une quantité de protéine identique, suivant l'ordre $[S]_3>[S]_2<[S]_1$. B) Tracé des valeurs des v_0 en fonction de la concentration du substrat associé. Avec l'équation de Michaelis-Menten, nous pouvons déterminer les valeurs de V_{max} et de K_M.

L'équation de Michaelis-Menten permet donc de relier la vitesse de réaction d'une enzyme à sa concentration.K_M est la constante de Michaelis et correspond à la concentration de substrat donnant une vitesse de réaction égale à la moitié de la vitesse maximale. Cela permet donc de donner une concentration requise pour obtenir un fort effet catalytique. K_M est donc reliée aux différentes constantes de vitesse selon l'Equation 4 qui suit :

$$v_0 = V_{max}\frac{[S]}{[S]+K_M}$$

Equation 4 : Michaelis-Menten

Il peut être également utile de rappeler que les constantes K_M et K_D, que nous avons déjà évoqué, bien que reliées entre elles, ne sont pas identiques. En effet, on considère que K_D et K_M sont équivalentes pour une valeur de k_{-1} bien supérieure à k_2, ce qui n'est pas tout le temps le cas. De plus, K_D est un paramètre thermodynamique tandis que K_M est un paramètre cinétique, c'est-à-dire qu'il ne prend pas seulement en compte l'affinité du substrat, mais également la vitesse à laquelle le complexe protéine-ligand va créer un produit.

Cependant, l'un des désavantages de l'équation de Michaelis-Menten est qu'il y a bien souvent assez peu de points expérimentaux, et que la détermination graphique est nécessairement imprécise, ce qui peut donner des valeurs de K_M et V_{max} moins précises. Une façon d'obtenir des valeurs plus précises est d'utiliser la réciproque des vitesses et de la concentration, pour donner l'équation de Lineweaver-Burk (Equation 5) et dont la représentation graphique est la suivante (Figure 53) :

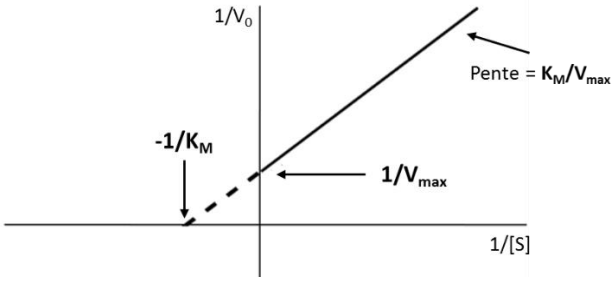


Figure 53 : Représentation théorique d'une courbe suivant l'équation de Lineweaver-Burk

$$\frac{1}{v_0} = \frac{1}{V_{max}} + \frac{K_M}{V_{max} * [S]}$$

Equation 5 : Lineweaver-Burk

Avec cette équation de type $\frac{1}{v_0} = a\frac{1}{[S]} + b$, la pente de la droite est donc $a = K_M/V_{max}$ et l'ordonnée à l'origine $b = 1/V_{max}$. Ainsi on peut facilement déterminer graphiquement K_M, qui correspond à l'intersection de la droite avec l'axe des y, et V_{max}, qui dépend de la pente de la courbe. De plus, l'un des autres avantages de cette représentation est de pouvoir simplement et graphiquement déterminer le mécanisme d'inhibition d'un ligand, c'est-à-dire, selon l'allure des courbes obtenues, que l'inhibiteur sera compétitif, incompétitif ou non compétitif. En effet un inhibiteur compétitif n'aura aucun impact sur V_{max} mais va augmenter la valeur de K_M, ce qui va résulter en une courbe possédant une pente différente, mais qui intersectera l'axe des ordonnées au même point que la courbe sans inhibiteur. Un inhibiteur incompétitif va en revanche diminuer de manière équivalente V_{max} et K_M, résultant en une courbe possédant la même pente (le ratio étant identique) mais une intersection sur les courbes des abscisses et des ordonnées différentes. Enfin un inhibiteur non-compétitif diminuera V_{max} sans changer la valeur de K_M. Ainsi, la courbe aura une pente et une intersection de l'axe des ordonnées différentes, mais coupera l'axe des abscisses au même point que la courbe sans inhibiteur. Les différentes représentations graphiques sont décrites en Figure 54.

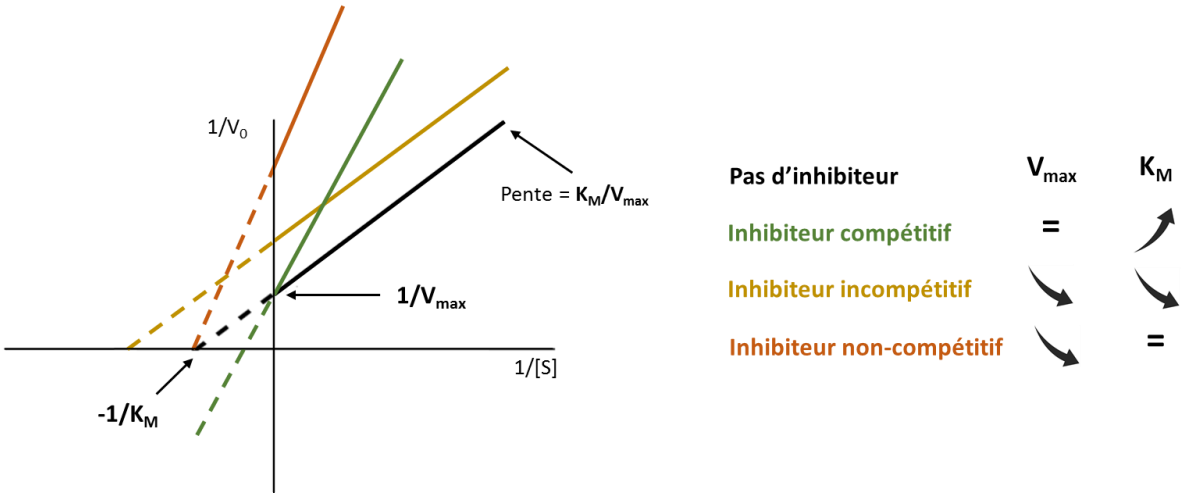


Figure 54 : Tracé des courbes de Lineweaver-Burk pour une protéine en présence d'inhibiteurs de différents mécanismes

IV.1.2.Cinétique test radioactivité pour dhOGG1

Les études cinétiques de dhOGG1 ont été effectuées en suivant le même principe du test d'activité glycosylase radioactif précédemment décrit. Dans ce test, l'ADN substrat radioactif (Act_OGG1) à concentrations variables est préincubé 5 min à 37 °C avec l'inhibiteur, dilué dans 100 % DMSO, dans le tampon de réaction (10 mM Tris-HCl pH 7,5, 150 mM NaCl, 1 mM EDTA, 0,1 mg/mL BSA +8 % DMSO). Pour évaluer les différentes vitesses à l'origine, nous avons utilisé 7 concentrations finales d'ADN substrat : 5/10/20/40/80/120/160 nM.

Au moment de lancer la réaction, la protéine diluée dans le tampon de réaction est ajoutée dans le mélange, à la concentration finale de 2 nM. Immédiatement après l'ajout de la protéine, le mélange est centrifugé quelques secondes sur une centrifugeuse de paillasse, et incubé à 37 °C sur un chauffe-tube Thermomix Comfort. Pour chaque point de cinétique, un aliquot est extrait du mélange réactionnel et incubé

en présence de 0,2 µM de NaOH à 50°C pendant 5 minutes. Les mélanges sont ensuite fortement agités au vortex et incubés en présence d'un volume équivalent de BF pendant 3 minutes à 76 °C et, afin de stopper la réaction. Pour chaque concentration d'ADN, des points de cinétique ont été effectués à 0, 1, 2, 4, 8, 16, 32 et 64 minutes. Chaque expérience a été effectuée en 3 réplicas avec des dilutions indépendantes. Finalement, les échantillons sont analysés par électrophorèse en gel de polyacrylamide (19/1) en présence d'urée (20 %) « Urea-PAGE », ce qui permet de maintenir des conditions dénaturantes durant la migration électrophorétique (500 V et 20 mA / gel pendant 20 min). Après 1 h d'exposition des gels obtenus sous un écran d'enregistrement luminescent au phosphore, celui-ci est lu grâce à un scanner laser Typhoon™ FLA 9500 (GE Healthcare). On peut ainsi quantifier la quantité d'ADN clivé et celle non clivé.

IV.1.3.Cinétique du test fluorescent pour hNEIL1

Pour hNEIL1 en revanche, la cinétique d'activité est trop rapide pour être évaluée par le test radioactif. Nous avons donc utilisé un test d'activité fluorescent analogue à celui utilisé lors du criblage primaire, en évaluant l'activité glycosylase-lyase de hNEIL1. L'ADN fluorescent TMR_NEIL1 est disposé dans une plaque noire polystyrène de 384 puits (Cornin 3715) avec un tampon de réaction (10 mM Tris-HCl pH 7,5, 150 mM NaCl, 1 mM EDTA, 0,1 mg/mL BSA) et préincubé 5 min à 37 °C dans le ClarioStar^{plus} (BMG Labtech). Un injecteur permet ensuite d'injecter la protéine dans chaque tube, et de suivre l'évolution de la fluorescence. Un point de cinétique est effectuée toutes les 0,25 sec pendant 60 minutes, et chaque puits est lu séparément (excitation 515/555 nm, émission 555/615 nm, 5 flashs par lecture). Les données sont ensuite extraites avec le logiciel Mars Data Analysis (BMG Labtech).

IV.2. Spectrométrie de masse pour le suivi de la formation d'adduit sur hNEIL1

Outre les analyses en conditions dénaturantes (MALDI-TOF) en fin de purification, la formation d'adduits sur hNEIL1 a également été suivie. Les analyses de spectrométrie de masse ont été réalisées sur la Plateforme de Spectrométrie de Masse et Protéomique du CBM par Guillaume Gabant. Après incubation de hNEIL1 (2 µM) avec un inhibiteur (800 µM) à 4 °C pendant au minium 20 minutes, ou en absence d'inhibiteur, pour un tampon final de 20 mM Hepes pH 7,5, 150 mM NaCl, 10 % glycérol, 2 % DMSO et 1 mM TCEP, deux méthodes différentes ont été utilisés pour fragmenter la protéine. Tout d'abord, l'approche top-down formz des fragmentations peptidiques au moyen de mécanismes physiques, comme la « dissociation induite par collision » (CID) ou la « dissociation par transfert d'électrons » (ETD). En revanche l'approche « bottom-up », se base sur plusieurs protéases pour fragmenter la protéine en peptides les plus représentatifs possibles de l'ensemble de la protéine. La combinaison de ces différentes approches permet d'obtenir une couverture de séquence plus importante.

IV.2.1.Analyse du complexe en LC (dessalage)-HRMS

La formation des adduits sur hNEIL1 a été mesurée par LC (dessalage)-HRMS. Les analyses ont été réalisées sur un système UltiMate™ 3000 NanoRSLC (Dionex, Sunnyvale, CA, USA) connecté à un spectromètre de masse 4-GHz MaXis ultra haute résolution quadripole-TOF (Bruker Daltonics) équippé d'une source d'ions électrospray. Le système LC-MS a été contrôlé par le logiciel HyStar™ version 3.2 (Bruker). Les protéines ont été déssalées en ligne sur une cartouche de phase inverse MassPREP™, 2.1 x 10 mm, phenyl 1000 Å (Waters) et éluées à un débit de 500 µL/min en utilisant un gradient de 5 à 90% d'acétonitrile dans 0.1% d'acide formique.

Les spectres ESI-QTOF ont été acquis en mode positif sur une gamme de *m/z* 150–3000 avec une pression de gaz nébuliseur de 2 bars. Le débit de gaz séchant était de 10 L/min, la température était de 200 °C, and le paramètre isCID (in source collision induced dissociation) était 30eV. La fréquence d'acquisition

était de 1 Hz correspondant à une somme de 5428 spectres. Une calibration externe a été réalisée avec l'ESI-L Low Concentration Tuning Mix (Agilent Technologies). Les spectres de masses ont été retraités et déconvolués en utilisant le logiciel DataAnalysis 3.1 (Bruker Daltonics) and l'algorithme MaxEnt.

IV.2.2.Localisation des résidus de hNEIL1 modifiés

- **MALDI-TOF MS**

Les peptides ont été dilués à une concentration finale de 0.05 µM dans une solution d'acide 4-hydroxy-α-cyano-cinnamique (4-HCCA) saturée dans un mélange acétonitrile:eau:acide trifluoroacétique 33.3:66.6:0.1 (TWA). Le mélange échantillon-matrice a ensuite été déposé sur une cible en utilisant la méthode en couche ultramince[458]. Les dépôts ont ensuite été lavés par du TFA 0.1% avant l'acquisition. Les analyses ont été réalisées sur un MALDI-TOF/TOF LIFT ultraflexXtreme (Bruker Daltonics). Les spectres ont été acquis en mode réflectron positif dans la gamme de *m/z* 500-5000. Une calibration externe a été réalisée en utilisant le Pepmix I (Bruker Daltonics). L'acquisition des spectres a été réalisée *via* le logiciel Flexcontrol 3.3 (Bruker Daltonics) et l'exploitation des données a été réalisée *via* le logiciel Flexanalysis 3.3 (Bruker Daltonics).

- **nanoLC-MS/MS**

Les peptides ont été analysés sur un système UltiMateTM 3000 nanoUltraHPLC (Dionex, Sunnyvale, CA) connecté à une trappe ionique HCT Ultra PTM Discovery (Bruker) équipée d'une source électrospray. La configuration LC-MS était pilotée par le logiciel HystarTM version 3.2 (Bruker). Les peptides ont été pré-concentrés en ligne sur une précolonne en phase inverse Thermo Acclaim PepMap100 C18 (diamètre interne 100 µm, longueur 2 cm, taille des particules 5 µm, taille des pores 100 Å), puis séparés sur une colonne analytique Thermo Acclaim Pepmap100 C18 (diamètre interne 75 µm, longueur 25 cm, taille des particules 2 µm, taille des pores 100 Å) à l'aide d'un gradient de 2 à 35% d'acétonitrile dans 0.1% d'acide formique réalisé à un débit de 450 nL/min. La colonne était thermostatée à 40°C. La sortie de la nano-colonne a été reliée directement à la source nanoESI en ligne en utilisant un NanoFlow ESI Sprayer (nanosprayer) (Bruker).

Au niveau de la trappe ionique, une tension de - 4000 V a été appliquée au niveau de l'aiguille de nébulisation du nanosprayer. Le gaz de séchage (N2) était à une température de 300°C avec un débit de 5 L/min. Les spectres ont été enregistrés en mode positif entre 300 et 1500 de *m/z* en MS et entre 100 et 2800 de *m/z* en MS/MS. Les acquisitions ont été réalisées avec une résolution de balayage améliorée (8100 *m/z*.s-1). La valeur cible d'accumulation des ions (ICC pour Ion Charge Control) a été fixée à 90 000. La fragmentation CID a été réalisée avec une énergie de collision de 1.20 V et de l'hélium (gaz de collision), et d'une manière préférentielle sur les ions dichargés. La fragmentation ETD a été réalisée en accumulant du fluoranthène jusqu'à atteindre une ICC de 400 000. Le temps de la réaction de transfert d'électrons a été fixé à 100 ms. Les spectres MS/MS ont été acquis en utilisant une liste d'inclusion des m/z des peptides modifiés observés en MALDI. Les fragmentations CID et ETD ont été effectuées dans des analyses nanoLC-ITMS/MS séparées. Une calibration externe a été réalisée en utilisant le calibrant ESI-T Tuning Mix (Agilent Technologies).

IV.2.3.Analyse des inhibiteurs seuls par infusion directe

Les inhibiteurs en solution dans le DMSO ont été dilués à une concentration finale située entre 1 et 20 µM dans une solution d'acétonitrile:eau:acide formique 50/49.8/0.2 (WAFA, pour Water Acetonitrile Formic Acid). Les échantillons ont été infusés à un débit de 180 µL/h à l'aide d'un pousse seringue et d'une aiguille de nébulisation ESI (Agilent Technologies). Les paramètres d'ionisation, de désolvatation et d'analyse MS et MS/MS sont identiques à ceux utilisés pour le couplage nanoRSLC. L'acquisition des spectres de masse

été réalisé *via* le logiciel Esquire Control 6.2, et le traitement des données *via* Data Analysis 4.0 (Bruker Daltonics).

IV.3. Etude de la thermostabilité des protéines par Thermal Shift Assay (TSA)

L'évaluation de la thermostabilité des protéines par fluorescence ou TSA est une des techniques largement utilisé pour évaluer la stabilité des protéines suite à l'interaction avec un ligand. Cette méthode repose sur l'évaluation de la dénaturation d'une protéine en fonction de la température. Pour ce faire, nous utilisons la solution « Sypro™ Orange protein gel stain » (Invitrogen), qui est un ligand non spécifique de la plupart des protéines, et dont la fluorescence augmente dans un environnement hydrophobe. Ainsi lors de la chauffe de la protéine, celle-ci va se dénaturer et se déplier, exposant ainsi son cœur hydrophobe dans lequel la solution Sypro va venir se lier, et donc produire de la fluorescence. Le signal maximal est obtenu pour une dénaturation complète de la protéine, puis il diminue suite à la dissociation du ligand sur la protéine. La première partie de la courbe de fluorescence produite en fonction de la température est donc une sigmoïde qui correspond à la fraction de protéine dénaturée. Ainsi, le point d'inflexion correspond à la température de fusion de la protéine (« Melting Temperature » ou T_m) de la protéine, température à laquelle 50 % de la protéine est dénaturée. Il est nécessaire pour chaque protéine de tester les conditions optimales pour évaluer une valeur de T_m. En effet une concentration trop faible en enzyme ou en Sypro Orange ne donnera pas de courbe exploitable ou donnera une courbe dont la dérivée ne fournira pas un puits unique, rendant la détermination du T_m moins précise (Figure 55). Au contraire un signal trop important va saturer le détecteur et rendra aussi la lecture du T_m moins précise.

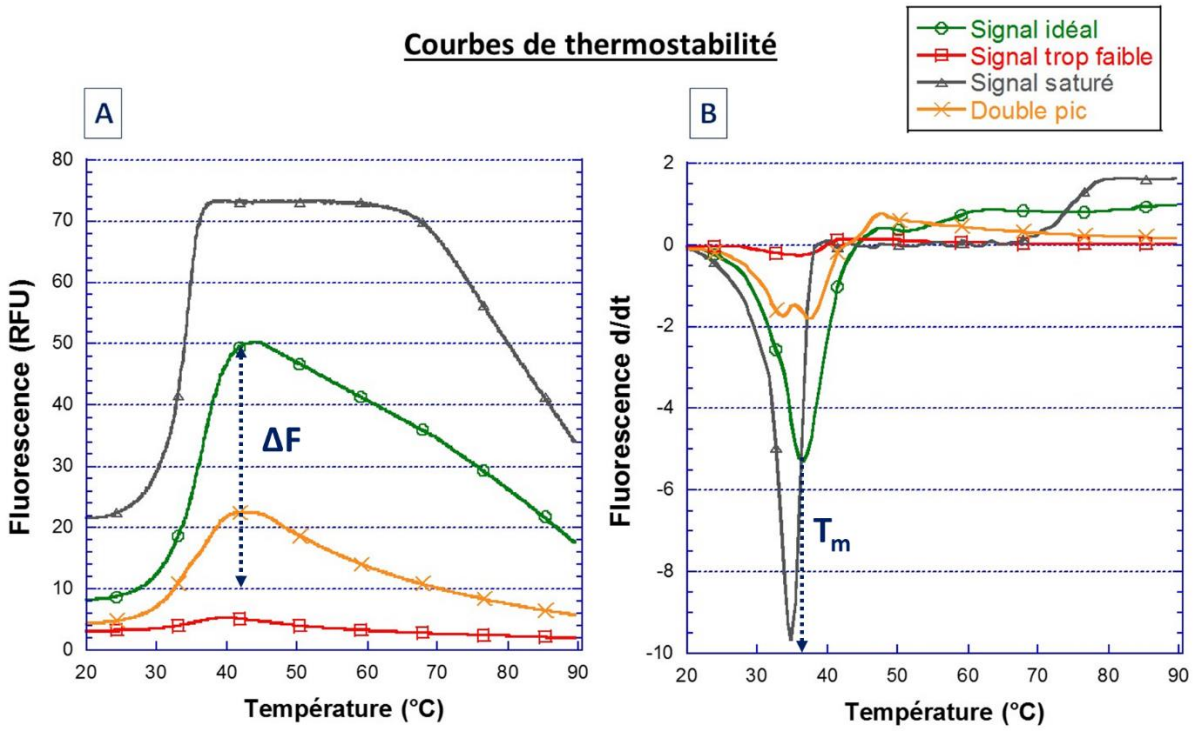


Figure 55 : Courbe de TSA théorique et évaluation de la température de fusion d'une protéine - A) La fluorescence produite par le Sypro Orange est lue en fonction de la température de l'échantillon. La différence entre le signal minimal et le *maximum* de fluorescence est noté ΔF. 4 courbes sont indiquées : une courbe avec un ΔF trop faible pour une lecture correcte en rouge, une courbe dont la fluorescence est trop élevée et fait saturer le lecteur en noir, et une courbe en orange dont le sommet est mal formé et possède deux *maxima* différents. Enfin en vert est représentée une courbe idéale. B) Sur le graphique est indiqué la dérivée de chacune des 4 courbes précédentes. Le *minimum* obtenu permet en abscisse d'obtenir la valeur du T_m.

Nous avons donc mesuré l'évolution du T_m de la protéine en présence des différents inhibiteurs dont nous disposions, que ce soit pour hOGG1 ou hNEIL1. Si le T_m de la protéine est plus faible, le ligand déstabilise la protéine, tandis qu'à l'inverse, si le T_m augmente, le ligand stabilisera la protéine, la rendant moins sensible à la dénaturation.

Concrètement, la protéine est incubée avec la solution Sypro™ Orange dans des microplaques blanches de 96 puits (FrameStar 480/96, 4titude), puis la plaque est centrifugée 1000 g à 4 °C pendant 2 minutes. La plaque est ensuite chauffée de 20 °C à 90 °C en 20 minutes, et la fluorescence est lue en direct par le thermocycler LightCycler™ 480 II (Roche) et les valeurs de T_m sont obtenues après analyse sur le logiciel intégré (LightCycler® 480 Software 1.5.0).

IV.3.1.Mise en place sur hNEIL1

La condition optimale pour ce test a tout d'abord été évaluée en préparant une plaque de 96 puits contenant différentes concentrations de la protéine à tester et de Sypro Orange. Chaque condition a été testée en quatre réplicas. Pour les conditions donnant une valeur de T_m exploitable, l'intérêt est de déterminer la valeur ΔF maximale afin de disposer du signal le plus clair possible. Ainsi pour la détermination des conditions de test optimales, lorsque les signaux étaient exploitables, nous avons sélectionné la condition donnant le ΔF le plus élevé pour la consommation la plus faible en matériel, et plus particulièrement la concentration en protéine. Les résultats obtenus sont indiqués en Figure 56. Les conditions retenues pour hNEIL1 sont donc **2,5X Sypro/2 µM hNEIL1**.

A

x4	1	2	3	4	5	6
A	2,5 X Sypro 0,25 µM Protéine	2,5 X Sypro 0,50 µM Protéine	2,5 X Sypro 1,0 µM Protéine	2,5 X Sypro 2,0 µM Protéine	2,5 X Sypro 4,0 µM Protéine	2,5 X Sypro 8,0 µM Protéine
B	5 X Sypro 0,25 µM Protéine	5 X Sypro 0,50 µM Protéine	5 X Sypro 1,0 µM Protéine	5 X Sypro 2,0 µM Protéine	5 X Sypro 4,0 µM Protéine	5 X Sypro 8,0 µM Protéine
C	10 X Sypro 0,25 µM Protéine	10 X Sypro 0,50 µM Protéine	10 X Sypro 1,0 µM Protéine	10 X Sypro 2,0 µM Protéine	10 X Sypro 4,0 µM Protéine	10 X Sypro 8,0 µM Protéine
D	20 X Sypro 0,25 µM Protéine	20 X Sypro 0,50 µM Protéine	20 X Sypro 1,0 µM Protéine	20 X Sypro 2,0 µM Protéine	20 X Sypro 4,0 µM Protéine	20 X Sypro 8,0 µM Protéine

B

Signal inexploitable
Double pic
Signal saturé
Signal exploitable

Valeur du Tm (°C)	1	2	3	4	5	6
A	36,95	36,50	36,79	38,11	36,28	34,65
B	x	x	34,76	36,96	36,58	34,81
C	x	x	31,62	31,90	33,60	34,63
D	x	x	x	31,33	31,93	34,53

Valeur du ΔF (°C)	1	2	3	4	5	6
A	2,31	4,14	12,96	27,87	50,61	46,99
B	x	x	8,42	18,63	53,04	50,04
C	x	x	0,77	1,66	6,86	14,46
D	x	x	x	1,15	3,77	13,15

Figure 56 : Test des conditions optimales pour déterminer une valeur précise de T_m avec une valeur de ΔF suffisamment élevée

IV.3.2.Mise en place sur hOGG1

Comme pour hNEIL1, afin d'évaluer le changement de température de fusion de la protéine, il a donc fallu déterminer la concentration optimale pour ce test. Ainsi, nous avons préparé une plaque de 96 puits contenant différentes concentrations de la protéine à tester et de Sypro Orange. Chaque condition a été testée en quatre réplicas. Les résultats obtenus sont indiqués en Figure 57. Les conditions retenues pour hOGG1 sont donc **5,0X Sypro/2 µM hOGG1**.

A

x4	1	2	3	4	5	6
A	2,5 X Sypro 0,25 µM Protéine	2,5 X Sypro 0,50 µM Protéine	2,5 X Sypro 1,0 µM Protéine	2,5 X Sypro 2,0 µM Protéine	2,5 X Sypro 4,0 µM Protéine	2,5 X Sypro 8,0 µM Protéine
B	5 X Sypro 0,25 µM Protéine	5 X Sypro 0,50 µM Protéine	5 X Sypro 1,0 µM Protéine	5 X Sypro 2,0 µM Protéine	5 X Sypro 4,0 µM Protéine	5 X Sypro 8,0 µM Protéine
C	10 X Sypro 0,25 µM Protéine	10 X Sypro 0,50 µM Protéine	10 X Sypro 1,0 µM Protéine	10 X Sypro 2,0 µM Protéine	10 X Sypro 4,0 µM Protéine	10 X Sypro 8,0 µM Protéine
D	20 X Sypro 0,25 µM Protéine	20 X Sypro 0,50 µM Protéine	20 X Sypro 1,0 µM Protéine	20 X Sypro 2,0 µM Protéine	20 X Sypro 4,0 µM Protéine	20 X Sypro 8,0 µM Protéine

B

Signal inexploitable
Double pic
Signal saturé
Signal exploitable

Valeur du Tm (°C)	1	2	3	4	5	6
A	45,57	45,63	45,29	45,13	44,57	44,39
B	42,95	44,43	44,92	**44,83**	44,57	44,09
C	x	x	x	42,97	44,44	44,50
D	x	x	x	43,43	44,24	44,28

Valeur du ΔF (°C)	1	2	3	4	5	6
A	3,67	5,07	14,02	16,78	26,16	25,20
B	2,11	5,61	17,63	**29,77**	47,99	56,93
C	x	x	x	11,55	66,77	81,71
D	x	x	x	20,08	55,83	103,26

Figure 57 : Test des conditions optimales pour déterminer une valeur précise de T_m avec une valeur de ΔF suffisamment élevée - A) Différentes conditions testées B) résultats obtenus : valeurs de T_m et ΔF.

IV.4. Evaluation de la constante K_D de pabAGOG par un test fluorescent

Suite à l'identification de composés inhibiteurs sur pabAGOG, le K_D de ceux-ci sur cette protéine ont été évalué. Pour rappel, la valeur de K_D représente la constante de dissociation entre un ligand et une protéine, quantifiant ainsi l'équilibre entre le ligand en solution et celui lié à la protéine, comme décrit en Figure 58.

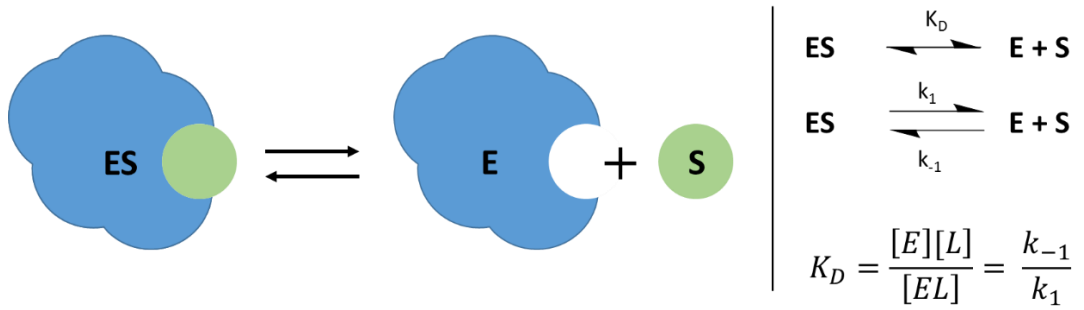


Figure 58 : Expression de la valeur de la constante de dissociation K_D

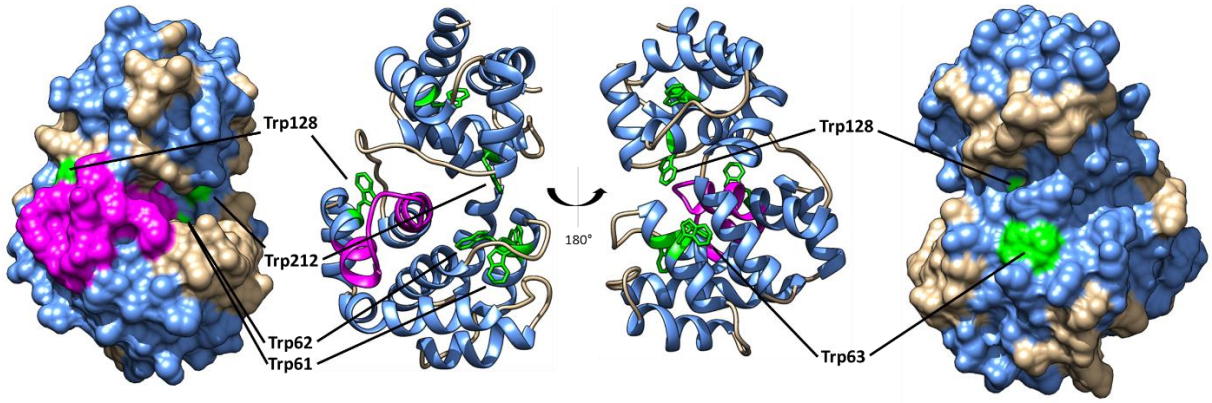


Figure 59 : Position des tryptophanes de pabAGOG

Pour cela nous avons utilisé le même protocole que Lingaraju *et al.* [459], en utilisant la fluorescence intrinsèque des différents tryptophanes de PabAGOG (dont les positions sont indiquées en Figure 59). Le tryptophane 212 est le résidu tout particulièrement intéressant, car il se situe dans le site actif. En effet, l'interaction de différents nucléosides avec les tryptophanes de PabAGOG peuvent diminuer la fluorescence produite. Ainsi, cela permet, en mesurant le maximum observé entre 290 et 450 nm pour une excitation à 280 nm, d'évaluer une valeur de constante de dissociation K_D selon l'Equation 6 de « fitting » des courbes obtenues.

$$F = f_1 * \mathrm{L_0} + \frac{1}{2}(f_2 - f_1) * \left[(\mathrm{P_0} + \mathrm{L_0} + \mathrm{K_d}) - \sqrt{(\mathrm{P_0} + \mathrm{L_0} + \mathrm{K_d})^2 - 4 * \mathrm{P_0} * \mathrm{L_0}}\right]$$

Equation 6 : **Equation de l'extinction de la fluorescence pour la détermination du K_D.** f_1 est le signal de bruit de fond de fluorescence, f_2 représente le signal de fluorescence à la concentration de protéine maximale. $\mathbf{P_0}$ est la concentration totale de pabAGOG et $\mathbf{L_0}$ la concentration totale du ligand.

La protéine pabAGOG, diluée dans un tampon de réaction (20 mM Hepes pH 7,6, 100 mM NaCl) pour une concentration finale de 0,1 µM est incubée avec les composés dilués dans le DMSO, de façon à obtenir une concentration de 8 % DMSO. Ce mélange est incubé à 4 °C pendant 30 minutes, puis la fluorescence est évaluée avec le lecteur Cary Eclipse Fluorescence (Agilent Technologies) (excitation à 280 nm, largeur faisceau 2,5 nm / scan du signal émission de 290 à 450 nm avec 30 nm/min, un intervalle de 0,5 nm, 10 nm de largeur faisceau et un voltage faible). Le même lot de protéines est utilisé pour chaque concentration de composé à tester.

V. Etude structurale et cristallogenèse

V.1. Principe de la cristallogenèse

La cristallographie biologique est une technique basée sur la diffraction des rayons X par un cristal de macromolécules, permettant d'obtenir son organisation tridimensionnelle.

La cristallogenèse correspond au mécanisme de formation d'un cristal, forme solide et ordonnée, à partir d'une macromolécule biologique à l'état soluble. Il est nécessaire de travailler dans des conditions de pureté maximale afin de faciliter l'interaction entre les protéines. Des agrégats non spécifiques peuvent mener à la précipitation de la protéine, mais dans des conditions favorables, ces interactions peuvent former des agrégats plus ordonnés, qui peuvent évoluer vers le « germe cristallin ». Chaque protéine possédant des propriétés spécifiques et uniques, la difficulté est donc d'identifier les conditions favorables pour la formation de ce cristal. Pour cela il est nécessaire de faire varier les différents paramètres qui régissent la solubilité d'une protéine. L'utilisation d'un agent cristallisant et la variation de la concentration des protéines permettent d'agir sur la solubilité des protéines en solution. Ainsi, on peut décrire l'évolution de la solubilité de la macromolécule avec un diagramme de phases, qui permet de mettre en évidence les différents états de la protéine (Figure 60). Dans la zone métastable, au-dessus de la courbe de solubilité, la protéine est sursaturée mais reste limpide et ne précipite pas. Dans la zone de précipitation au contraire, comme son nom l'indique, la protéine précipite. Il existe une zone intermédiaire dite « zone de nucléation ». C'est uniquement dans un scénario ou la sursaturation atteint cette zone que le cristal va pouvoir se former. Par la suite la formation du cristal va diminuer la concentration en protéine, jusqu'à atteindre la zone de solubilité, et la croissance cristalline s'arrête.

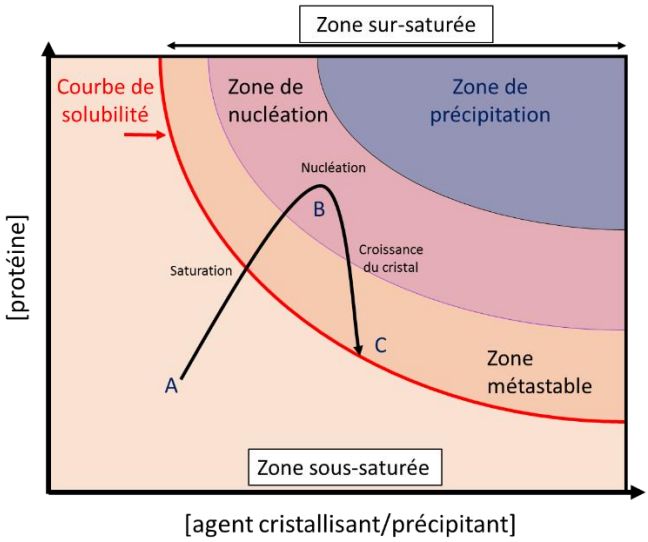


Figure 60 : Diagramme de phases à deux dimensions d'une macromolécule (adapté de Sauter et Giegé, 2010) [460]
- A) Macromolécule soluble, B) Nucléation du cristal et C) Cristal final stable.

V.2. Méthodes utilisés pour la formation de cristal

De nombreuses méthodes existent pour la cristallisation de macromolécules biologiques, mais lors de cette thèse nous avons utilisé la méthode de diffusion en phase vapeur. Cette technique consiste à exploiter un équilibre en phase vapeur dans un système clos (Figure 61). Ainsi, un grand volume de solution de cristallisation, dit « réservoir », est mis en présence d'une goutte de solution de protéine mélangée avec la solution de cristallisation. Ainsi, la concentration de la protéine et celle de la solution de cristallisation est plus faible dans la goutte, que dans la solution « réservoir ». Comme le réservoir est plus concentré en agent cristallisant, de la vapeur d'eau s'échappe de la goutte vers le réservoir, concentrant ainsi la protéine, jusqu'à ce que la goutte atteigne la même concentration en agent cristallisant que le réservoir. La protéine peut donc entrer en sursaturation, et cristalliser dans les bonnes conditions. L'avantage de cette approche est sa facilité de mise en place, et le faible volume nécessaire pour l'appliquer. Il est donc possible de l'effectuer avec des robots de cristallisation comme le robot pipetteur Mosquito® (sptlabtech) par exemple. Lors de cette thèse, nous avons appliqué cette méthode à travers deux techniques, un système de « goutte assise », et un système de « goutte suspendue ». Le système de goutte assise, a été utilisé pour les criblages cristallographiques, avec un réservoir de 80 µL, et une goutte formée de 100 nL de solution protéine et 100 nL de solution de cristallisation, dans une plaque transparente de 96 puits. Le système de goutte suspendue a été utilisé pour l'optimisation manuelle des conditions de cristallisation, dans des plaques de 24 puits (HR3-306 VDXm, Hampton Research), avec un réservoir de 750 µL et un volume de goutte allant de 0,8 à 3,2 µL (selon les ratios utilisés entre la solution protéine et la solution de cristallisation).

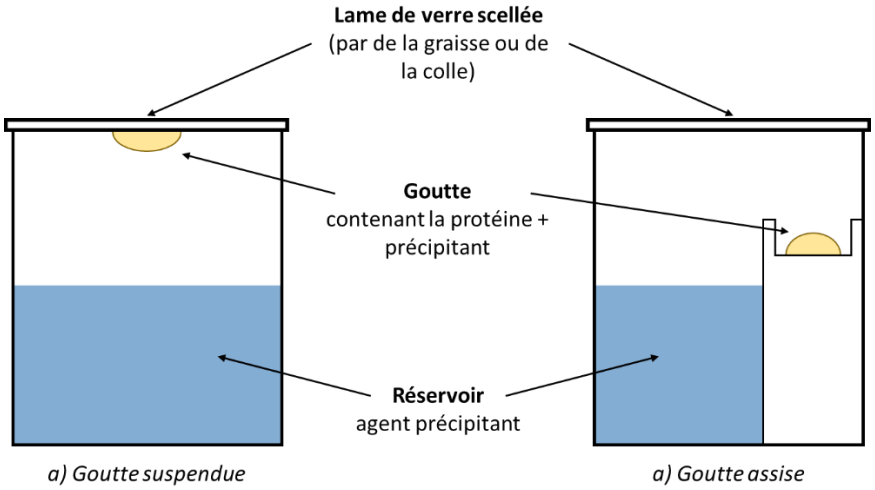


Figure 61 : Schéma de la diffusion en phase vapeur par (a) goutte suspendue ou (b) goutte assise

V.3. Collecte des données de diffraction des rayons X

Quand nécessaire, les cristaux obtenus ont été plongés dans une solution contenant un cryoprotectant avant d'être trempés dans de l'azote liquide. Dans le cas de hNEIL1dc95 et hNEIL1dc56 1µL d'éthylène glycol à 100% a été ajouté à 3µL de la solution du réservoir.

La diffraction aux rayons X de ces cristaux a été effectuée sur les lignes PROXIMA-1 et PROXIMA-2 du synchrotron Soleil (CNRS, CEA, Paris-Saclay) et les images de diffraction ont été collectées à 100 K. Les données obtenues ont été intégrées et mises à l'échelle grâce aux logiciels XDS/AIMLESS[461] ou autoPROC[462]. La structure 3D des molécules cristallisées a été déterminée par remplacement moléculaire avec PHASER de la suite PHENIX[463] (modèle PDB 5ITQ pour hNEIL1 et 1XG4 pour pabAGOG). Les modèles atomiques ont été affinés avec la fonction phenix.refine de la suite PHENIX et améliorés manuellement avec le logiciel COOT[464]. Le tableau des données de collection et des statistiques est présenté dans les paragraphes concernant les résultats de la cristallographie de hNEIL1 et pabAGOG. Enfin, la validation stéréochimique des modèles finaux a été effectuée avec Molprobity et les images ont été produites avec le logiciel UCSF Chimera[465].

Résultats et Discussions

Partie 1 : Identification et caractérisation d'inhibiteurs de l'enzyme hNEIL1

I. Mise en place de la procédure de criblage sur la 8-oxoG-ADN glycosylase bactérienne Fpg de *Lactoccus lactis*

L'objectif de cette thèse est l'identification d'inhibiteurs des ADN glycosylases humaines NEIL1 et OGG1. Afin de valider la procédure de notre criblage ainsi que ses différentes étapes, celui-ci a été effectué en amont avec une protéine bien connue de l'équipe, la Formamidopyrimidine-ADN glycosylase de *Lactococcus lactis* (Fpg ou MutM). Celle-ci est un homologue fonctionnel de hOGG1 et structural de hNEIL1, ce qui en fait un candidat idéal comme preuve de concept. De plus, des inhibiteurs ont déjà été identifiés sur cette protéine, des composés thiols ou dithiols, attaquant le doigt de zinc de la protéine [247]. Une publication dont je suis co-auteur et sur laquelle j'ai participé aux différents tests d'activités, a notamment été publiée en 2020 sur des dérivés de la 2-thioxanthine (2TX, Figure 36) comme inhibiteurs de hNEIL1 [246]. Cependant, cette thèse ayant principalement pour objectif d'identifier des composés novateurs, nous ne nous attarderons pas sur ces molécules, que nous avons princiaplement utilisés comme contrôle positif pour la mise en place du criblage.

I.1. Criblage primaire

I.1.1. Mise en place et optimisation

Pour la mise en place du criblage primaire, nous avons utilisé des plaques de 384 puits, avec un volume d'essai de **30 µL** final, préparés au moyen d'un robot pipeteur EVO TECAN dans un premier temps (Annexe 2). Dans chaque essai était mélangé un composé des chimiothèques, le tampon de réaction, la protéine à cribler et le substrat fluorescent, selon un plan de plaque indiqué en Annexe 2. En Figure 62 sont montrés quelques exemples des différents résultats obtenus lors de ces tests initiaux. On remarque déjà que les valeurs de Z' sont aberrantes, certaines étant négatives, d'autres étant très largement supérieures à 1. De plus, on voit rapidement que les contrôles (positifs et négatifs) ne donnent pas de résultats reproductibles car ils montrent une importante différence de fluorescence observée d'un test à l'autre, et parfois même entre plusieurs réplicas sur une même plaque. L'explication la plus probable est que le faible volume des pipetages ne permettait pas au robot d'être précis, bien que nous soyons au-dessus de la limite indiquée par le fournisseur. Nous avons donc décidé d'augmenter les volumes utilisés par deux, pour un volume final dans l'essai de **60 µL**.

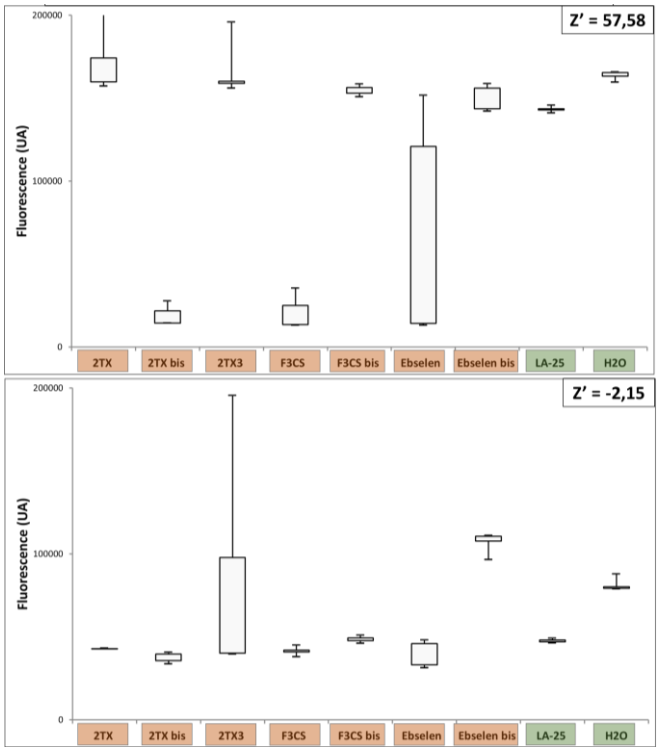


Figure 62 : Exemples de résultats obtenus lors des premiers essais de criblage automatique dans des puits de 30 µL - Les contrôles positifs sont indiqués en rouge tandis que les contrôles négatifs sont indiqués en vert.

Nous avons donc de nouveau criblé une petite plaque de contrôles connus afin d'évaluer si les résultats étaient plus satisfaisants (Figure 63). Les nouvelles valeurs de Z' obtenues (supérieures à 0,5) sont bien plus satisfaisantes et les essais contrôles ont le comportement attendu et semblent relativement reproductibles, malgré quelques écarts relatifs importants (c'est le cas de l'eau et de LA1).

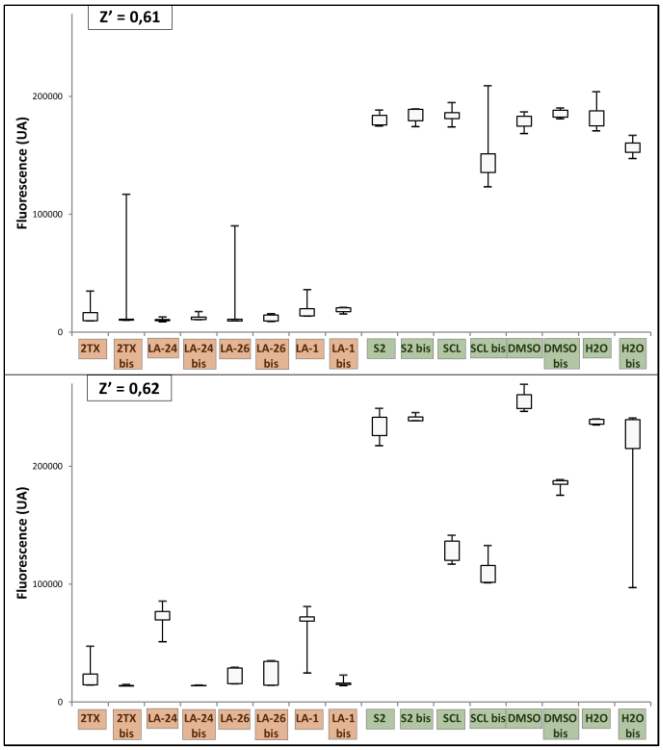


Figure 63 : Exemples de résultats obtenus pour le deuxième essai de criblage automatique pour un volume final de 60 µL - Les contrôles positifs sont indiqués en rouge tandis que les contrôles négatifs sont indiqués en vert. Cette notation sera conservée pour les figures de la mise en place du criblage.

Nous avons donc décidé de conserver ces conditions pour cribler la chimiothèque de Greenpharma (Figure 64). Si l'augmentation du nombre de puits à préparer sur la plaque ne montre pas d'influence négative sur la valeur du facteur Z', certains contrôles en revanche, ne sont pas reproductibles et affichent à nouveau des résultats absurdes (l'eau ultra pure inhibitrice, par exemple). A ce stade de nos investigations, nous avons remarqué qu'à cause du plus grand volume aspiré dans la seringue par le robot pipeteur, et lié au plus grand nombre de puits à préparer, celui-ci perdait fortement en précision, provoquant des résultats « aléatoires ». Cet effet est d'autant plus visible lors du criblage d'une plaque de 96 puits remplis d'eau ultrapure contre LlFpg (Figure 65). Nous voyons bien ici que les résultats du criblage, quelle que soit la valeur de Z' obtenue, ne peuvent être fiables. Au vu des résultats précédents, il n'était donc envisageable ni de diminuer, ni d'augmenter le volume final afin d'augmenter les performances du robot. Nous aurions pu envisager de programmer le robot différemment, afin qu'un volume moins important soit aspiré par la seringue, tout en conservant un volume final important, cependant cela entrainait un temps de préparation beaucoup plus long à température ambiante, et donc une détérioration potentielle de la protéine. Nous avons donc décidé de préparer les plaques du criblage manuellement.

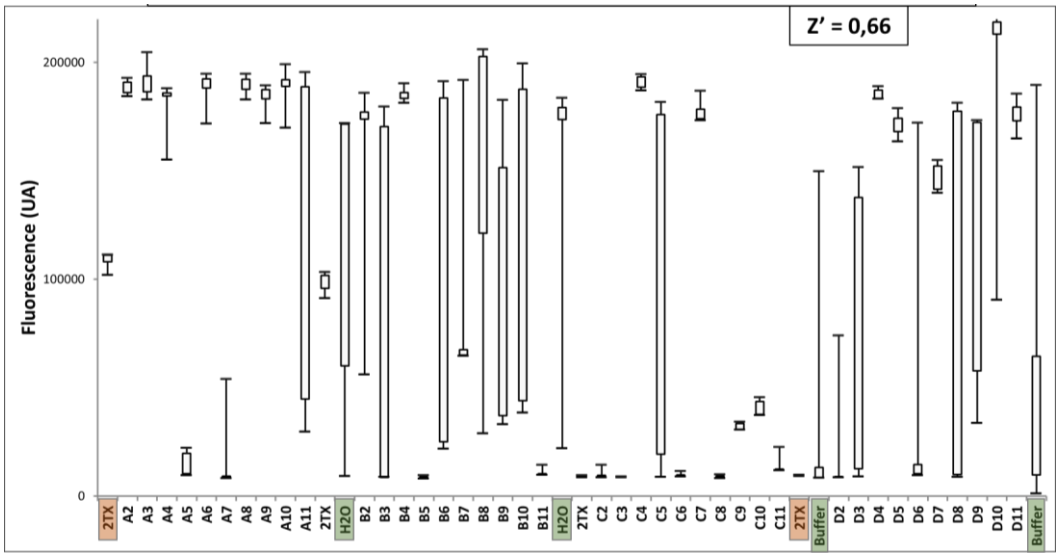


Figure 64 : Criblage automatique d'une plaque de la banque de diversité de Greenpharma sur LlFpg

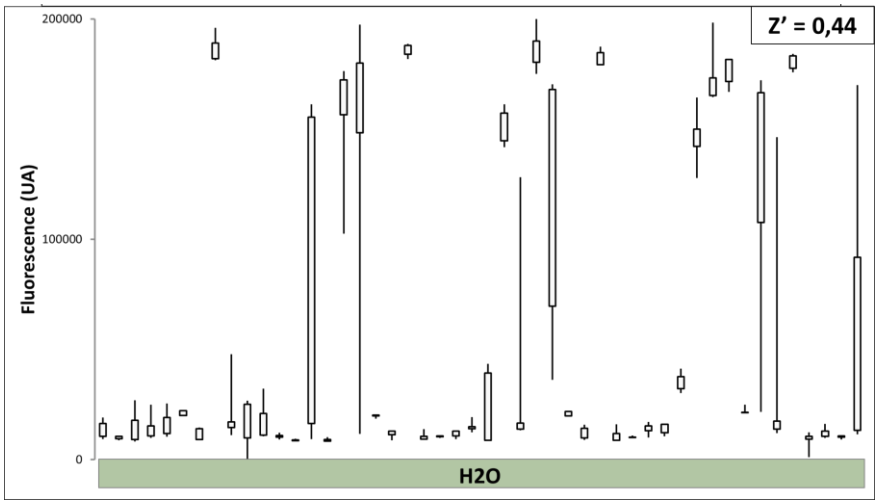


Figure 65 : Criblage automatique d'une plaque d'eau ultrapure sur LlFpg

Une plaque de 384 puits a donc été préparée à la pipette, en suivant le plan de plaque indiqué en Annexe 2, ainsi que le rajout de l'ADN substrat, puis la fluorescence a été évaluée. La valeur de Z' obtenue est excellente, tandis que les contrôles sont reproductibles. Ces résultats sont donc particulièrement en faveur du criblage manuel (Figure 66).

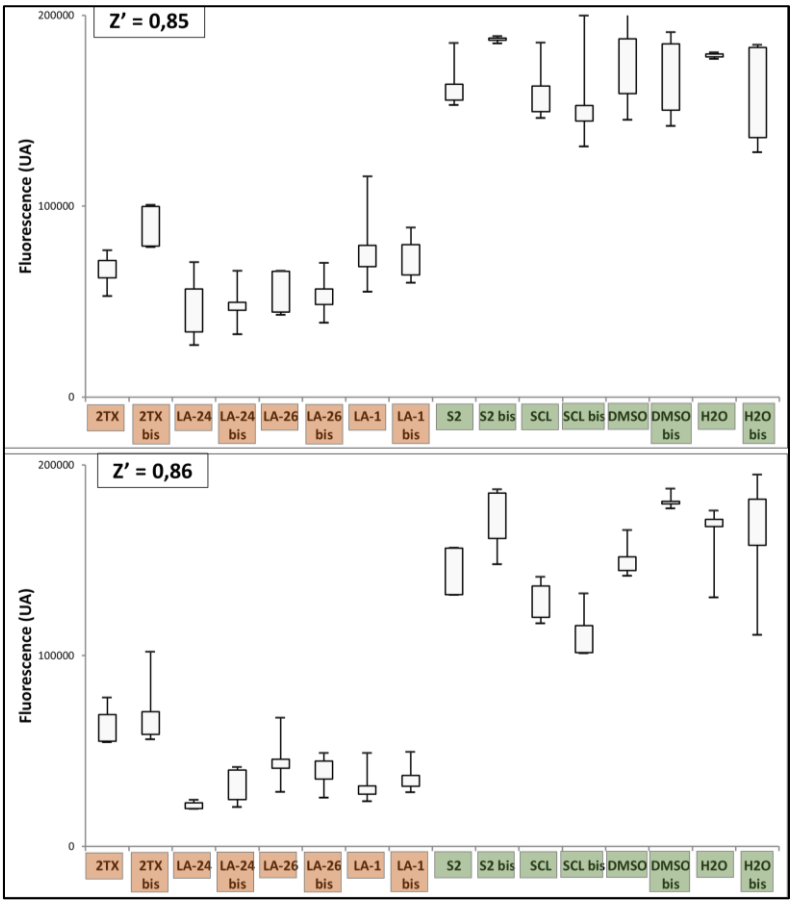


Figure 66 : Criblage manuel d'une plaque de contrôle sur LlFpg

I.1.2. Criblage manuel de la chimiothèque GreenPharma sur LlFpg

Malgré une consommation en temps évidemment plus élevée qu'un criblage automatique, nous avons donc décidé d'effectuer le crible manuel des 480 composés de la banque GreenPharma (GP). Le plan de plaque utilisé pour ce criblage « manuel » est indiqué en Annexe 2. Sur chaque plaque de 384 puits, nous pouvons tester 40 composés en 5 réplicas. Il nous a donc fallu 12 plaques de 384 puits pour tester les 480 composés de cette chimiothèque. Un exemple de résultats du criblage de 40 composés sur LlFpg est indiqué en Figure 67. Les contrôles positifs et négatifs sont bien reproductibles, ce qui valide donc bien la fiabilité de ce criblage manuel pour un plus grand volume de composés testés.

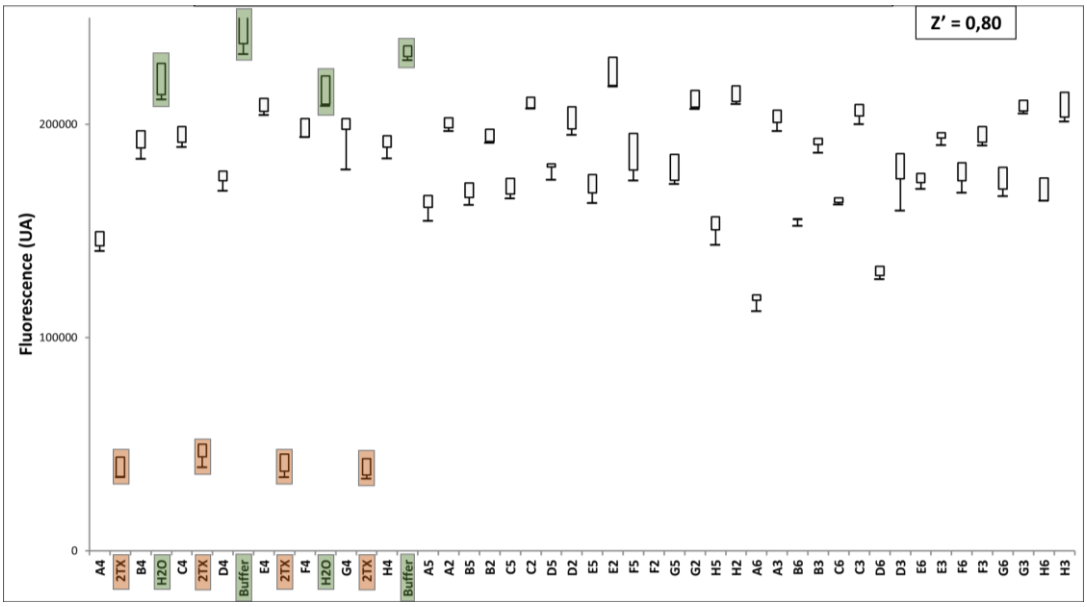


Figure 67 : Criblage « manuel » d'une plaque de GP sur LlFpg

Les résultats globaux de ce criblage manuel sont indiqués en Figure 68. 15 composés y ont été identifiés comme hits, avec un seuil de sélection supérieur à 75 % d'inhibition. La valeur moyenne des Z' obtenus pour les 12 plaques testées est de 0,83 ± 0,11, ce qui est une valeur très correcte et qui nous a donc encouragés à utiliser cette procédure pour identifier des inhibiteurs pour les ADN glycosylases humaines.

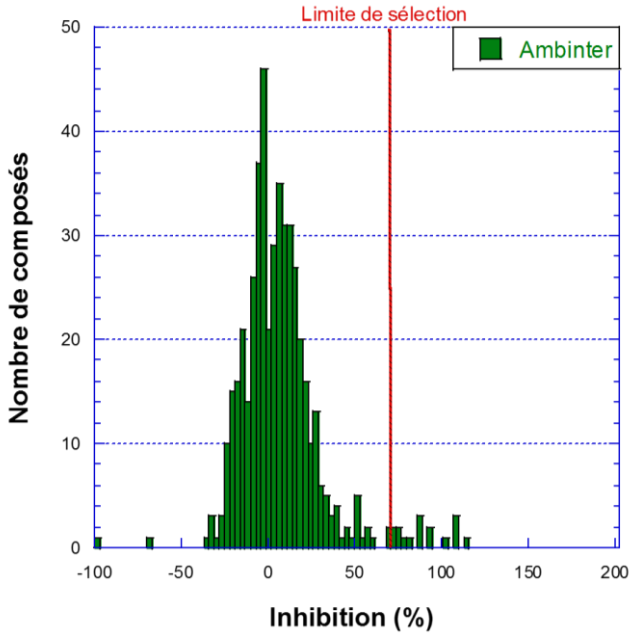


Figure 68 : Résultats du criblage de la banque GP sur LlFpg en pourcentage d'inhibition

I.2. Confirmation des hits sur LlFpg par un test orthogonal

Afin de confirmer ces hits, nous avons donc mis en place le test orthogonal utilisant un ADN substrat radioactif. Après incubation, les mélanges réactionnels sont analysés par Urée-PAGE (voir Chapitre Mat & Met). Nous avons donc testé les différents hits obtenus lors du précédent criblage, à une concentration fixe d'inhibiteur de 100 µM, pour 20 nM d'ADN substrat et 2 nM de LlFpg. L'objectif est de se placer dans des

conditions donnant environ 50 % d'activité de l'enzyme (en condition d'enzyme limitante), afin de pouvoir bien évaluer les variations d'activité en présence de nos composés à confirmer (Figure 69).

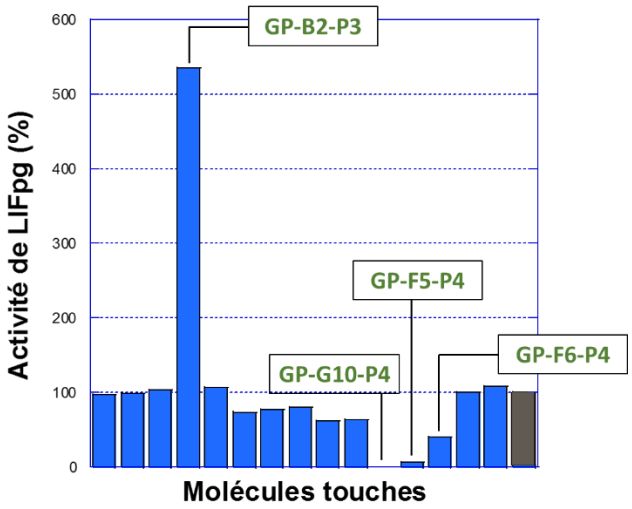


Figure 69 : Confirmation des hits du criblage primaire sur LlFpg par un test d'activité orthogonal - Le contrôle avec 8 % DMSO est indiqué en gris foncé sur l'histogramme.

Tout d'abord, on remarque rapidement que l'un des composés, la GP-B2-P3, ou Harmine (Figures 69 et 70), semble stimuler l'activité au lieu de l'inhiber. Ce composé a probablement été perçu comme inhibiteur lors du criblage primaire en produisant de la fluorescence dans les mêmes longueurs d'ondes que le fluorphore utilisé. Des essais de dose-réponse ont été effectués avec cette molécule sur LlFpg en faisant varier la concentration, mais les résultats n'ont pas été suffisamment reproductibles pour être véritablement exploitables. Hormis ce résultat inattendu, sur les 15 hits identifiés lors du crible, 3 composés ont montré une inhibition supérieure à 50 % pour une concentration de 100 µM. Parmi ces composés, nous avons sélectionné les deux molécules les plus inhibitrices, **GP-F5-P4** (Juglone) et **GP-G10-P4** (Purpurine) (Figures 69 et 70) induisant une activité de LlFpg inférieure à 10 % dans les conditions du test. Nous avons mis de côté le composé **GP-F6-P4**, ou Korseveriline (Figure 70) car son inhibition était de 40 % pour 100 µM, ce que nous avons considéré comme trop faible. Nous aurions pu évaluer plus de composés inhibiteurs confirmés, mais nous sommes ici dans une preuve de concept. Nous n'avons donc pas besoin de faire une étude exhaustive sur tous les composés identifiés, seuls les plus efficaces (plus efficaces que nos contrôles positifs tels que 2TX) nous intéressent pour prouver qu'identifier une molécule inhibitrice est possible avec ce criblage. Nous avons donc évalué la valeur d'IC_{50} de ces deux composés, et déterminé des valeurs de 16,37 ± 1,72 µM pour **GP-F5-P4** et de 11,34 ± 2,32 µM pour **GP-G10-P4**. Ces molécules sont donc plus efficaces sur LlFpg que la 2-thoxanthine (2TX) identifiée précédemment par l'équipe comme un inhibiteur des ADN glycosylases de la superfamille structurale Fpg/Nei, qui possède un IC_{50} de 48,40 µM sur LlFpg. D'autres inhibiteurs restent plus efficaces que ces composés, comme la TX13 qui présente un IC_{50} de 7,5 µM sur LlFpg [246,247].

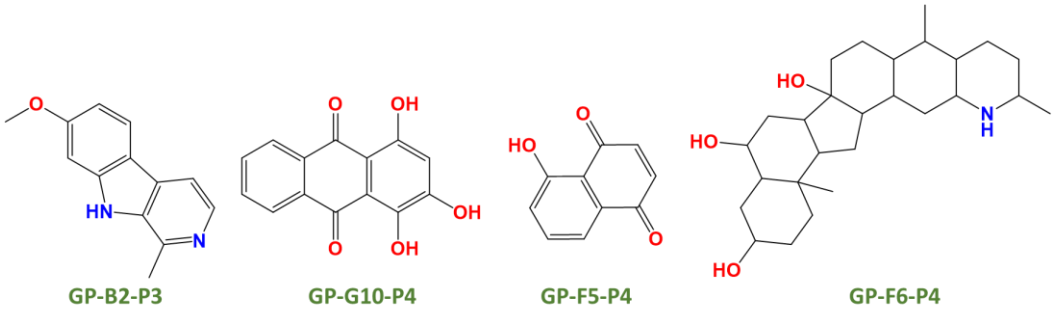


Figure 70 : Structure de la purpurine (GP-G10-P4) et de la juglone (GP-F5-P4), deux inhibiteurs confirmés de LlFpg

I.3. Evaluation de l'interaction à l'ADN des composés identifiés

Avec l'identification de deux nouveaux inhibiteurs de LlFpg, l'objectif visant à mettre en place un test de criblage fiable et utilisable pour d'autres ADN glycosylases semblait donc atteint. Cependant, de par leur fort potentiel d'intercalation dans l'ADN, les composés de type quinones tels que la purpurine et la juglone peuvent interférer avec nos essais enzymatiques en séquestrant le substrat nucléique. Afin d'évaluer ce phénomène, nous avons donc mis en place un test nous permettant d'évaluer le pouvoir intercalant de nos composés (le test Picogreen, voir chapitre Matériel et méthodes). Tout d'abord, nous avons évalué la fluorescence d'un ADN double brin contenant le substrat en présence du Picogreen, avec une concentration croissante de DMSO. Ensuite, nous avons évalué la fluorescence de ce même ADN en présence du Picogreen et en présence de concentrations croissantes de nos deux inhibiteurs. Le bleu de méthylène, un intercalant fort de l'ADN, a été utilisé dans cet essai comme contrôle positif de l'intercalation. La fluorescence associée au Picogreen chute fortement en présence de **GP-F5-P4** ou de **GP-G10-P4**, de manière assez similaire au bleu de méthylène (Figure 71). Ces résultats confirment donc bien que ces deux composés sont fortement intercalants.

Finalement, même si ces molécules ne sont probablement pas des inhibiteurs de LlFpg et exercent leur action plutôt sur son substrat, elles valident l'intérêt de disposer de plusieurs tests pour confirmer le caractère inhibiteur d'un composé sur une ADN glycosylase dont le substrat est un acide nucléique. Nous avons donc validé cette procédure de criblage et de validation pour identifier des inhibiteurs des ADN glycosylases humaines hNEIL1 et hOGG1.

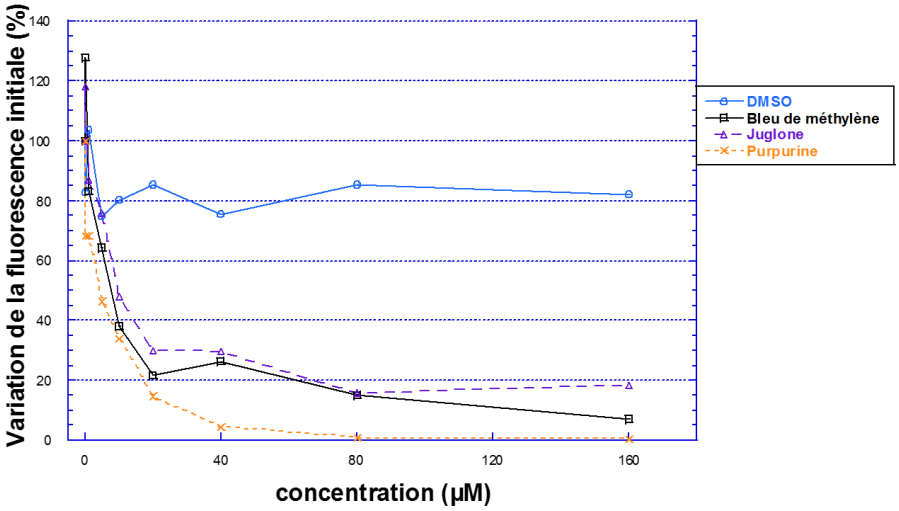


Figure 71 : Test d'intercalation à l'ADN de la Juglone et de la Purpurine

II. Recherche d'inhibiteur de la protéine hNEIL1

II.1. Criblage primaire

II.1.1. Criblage semi-automatique et optimisation

Le principe de criblage ayant été validé lors du précédent test, le criblage primaire a pu être mis en place pour l'ADN glycosylase hNEIL1. En revanche, le criblage manuel de l'intégralité des 4560 composés, soit 114 plaques, aurait pris trop de temps pour cette thèse (environ 1 à deux plaques par jour). Afin d'accélérer et de faciliter le processus, nous avons dû effectuer quelques changements. Tout d'abord, 3 réplicas au lieu de 5 ont été réalisés, ce qui a permis de tester 80 composés par plaque de 384 puits, réduisant par deux le nombre de plaques à préparer (Annexe 2). De plus, nous avons opté pour un criblage combiné, pour les banques de molécules les plus conséquentes que nous avons utilisées (ICOA2 et Prestwick), consistant à tester deux composés par puits. Pour les banques moins volumineuses (ICOA1 et GP) en revanche, un seul composé par puits a été testé. De plus, l'ajout de l'ADN substrat et des inhibiteurs a été effectué par un robot plus rapide et adapté pour le pipetage de petits volumes, le Mosquito (robot classiquement utilisé pour les essais de cristallisation et capable de délivrer des volumes de 50 nL avec précision.). Cela a permis d'effectuer ce criblage de manière « semi-automatique », avec un débit moyen. Comme ce robot ne peut pas manipuler de gros volumes de liquide, nous avons dû réduire le volume final des tests. Cela a permis par la même occasion de diminuer la quantité de protéine nécessaire au criblage. Le ratio signal/bruit obtenu avec le Mosquito, noté R1 et R2 respectivement pour 0 et 15 minutes d'incubation à 37 °C a été évalué en fonction du volume utilisé (pour une concentration d'ADN substrat de 50 nM et 50 nM de hNEIL1). Si le plus haut ratio signal/bruit est obtenu pour les volumes de 20 et 30 µL, nous pouvons aussi remarquer que la différence de ratio signal/bruit change peu entre 0 et 15 minutes (Figure 72). Cela indique que l'activité de l'enzyme est peu détectée. Pour évaluer ce paramètre nous avons calculé le rapport R1/R2, afin d'évaluer la condition montrant la plus grande différence d'activité de l'enzyme. Ainsi, la plus grande valeur de R1/R2 après 15 minutes correspond au volume de **15 µL**. Nous avons donc considéré que 15 µl serait le volume de travail le plus optimal pour la suite du criblage. Ces différentes améliorations nous ont permis de cribler les 4560 molécules sur hNEIL1 sans faire appel à une plateforme extérieure.

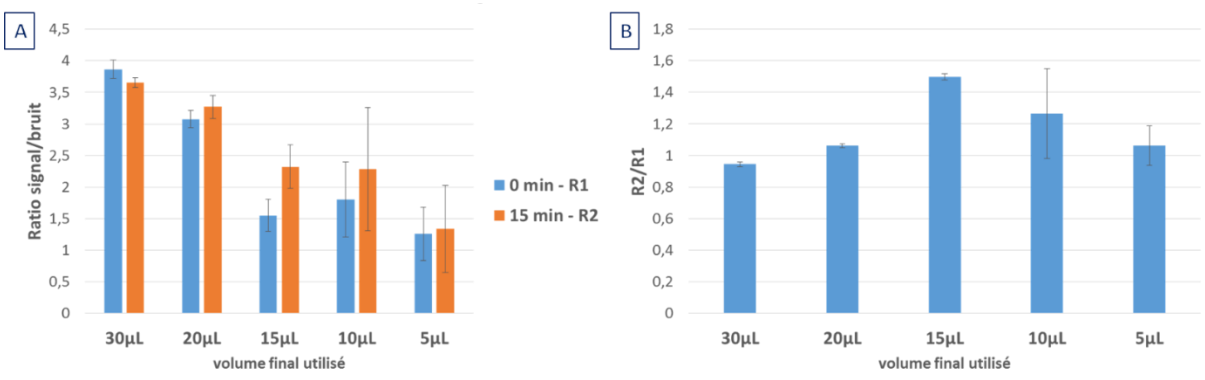


Figure 72 : Détermination du volume final pour le criblage primaire semi-automatique - A) Evaluation de R1 et R2, les ratios signal/bruit pour 0 et 15 minutes d'incubation respectivement. B) Calcul de R1/R2 pour les différents volumes utilisés.

II.1.2. Choix des conditions du criblage

Le principe de ce criblage s'inspire largement du test effectué par Jacobs *et al.,* pour hNEIL1 [416]. Cependant les tampons de conservation de la protéine diffèrent, ainsi que ceux du crible. Nous avons gardé un temps d'incubation de **15 minutes** à 37 °C similaire à celui de la publication, ainsi qu'un volume final dans

le puits de **15 µL**. De plus, nous avons testé 3 concentrations finales (50-75-100 nM) de hNEIL1 et 3 concentrations finales d'ADN substrat (50-75-100 nM) afin d'évaluer le ratio signal/bruit sur 3 réplicas (Figure 73). La condition retenue à l'issue du test est celle donnant le ratio signal/bruit le plus élevé, soit **100 nM d'ADN substrat et 100 nM hNEIL1** (quadrillé en rouge sur le graphique).

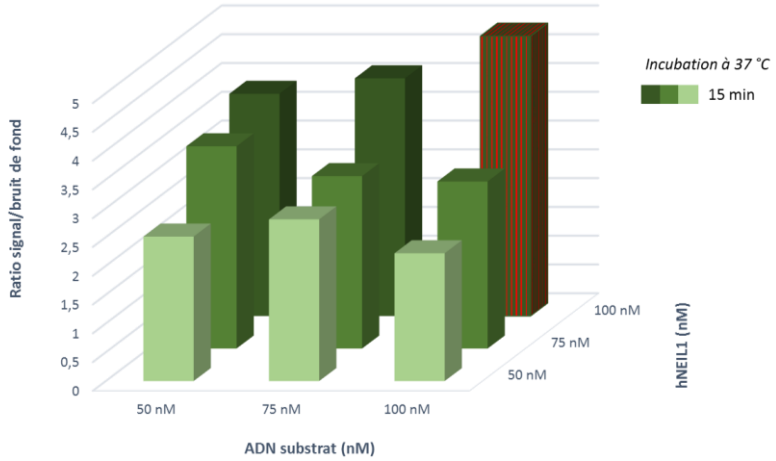


Figure 73 : Evaluation du ratio signal/bruit pour le criblage de hNEIL1

II.1.3. Molécules touches sélectionnées

En utilisant les conditions optimisées précédemment décrites, 4560 molécules ont été testées sur hNEIL1 lors du criblage primaire. Comme indiqué plus haut, chaque composé a été testé 3 fois sur la même plaque. Un **facteur Z'** a été calculé pour chacune des plaques. Ce criblage s'est révélé robuste avec une valeur de Z' moyenne de 0,84 ± 0,04 (Figure 74). Ces valeurs de Z' sont excellentes, d'autant plus pour un criblage « semi-automatique » plus long et fastidieux qu'un HTS classique, et permettent de valider le criblage sur hNEIL1.

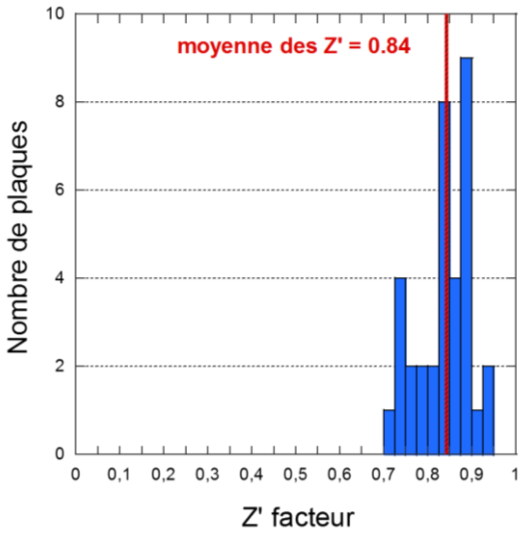


Figure 74 : Distribution des facteurs Z' par plaque des cribles primaires sur hNEIL1

Pour retenir un hit, un seuil de sélection (ou « cut-off ») a dû être déterminé par rapport à des inhibiteurs déjà connus au laboratoire, la Juglone et la Purpurine (voir plus haut). Ces deux molécules montrent un effet d'inhibition d'environ 65 % et 60 % respectivement sur hNEIL1, pour une concentration de 50 µM. Par analogie avec l'IC_{50} de ces composés sur LlFpg, nous avons considéré dans un premier temps que

seules les molécules montrant une inhibition plus forte seraient retenues. Un « cut-off » de 70 % d'inhibition a donc été choisi. En effet, il est généralement admis qu'un criblage idéal doit produire ente 0,5 et 2 % de hits. Un nombre de hits trop important peut être la manifestation d'une concentration d'inhibiteur trop importante ou d'une protéine instable dans les conditions du test. Au contraire, un nombre de hits trop faible, élimine probablement des composés intéressants mais non détectables dans les conditions du test.

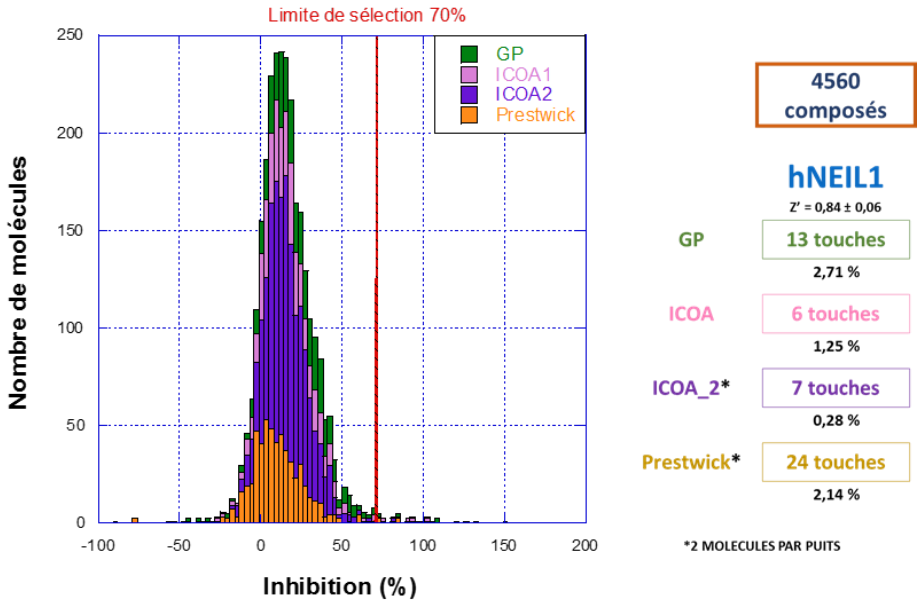


Figure 75 : Distribution du pourcentage d'inhibition des composés du criblage primaire sur hNEIL1

Les composés criblés sont classés en fonction de leur capacité d'inhibition de hNEIL1 (Figure 75). Sur la base des critères que nous avons définis plus haut, 50 molécules ont été retenues, soit 1,09 % de hits ce qui correspond au résultat attendu d'un criblage idéal.

II.2. Confirmation des hits

Afin de confirmer les hits par l'essai enzymatique orthogonal (chapitre Matériel et Méthodes dédié), nous nous sommes placés à une concentration fixe de 50 µM de composé. En effet, toujours dans la logique d'une comparaison avec la Juglone et la Purpurine, une molécule ne montrant pas d'effet supérieur à 50 % d'inhibition à 50 µM ne sera pas retenue. De plus, contrairement au criblage primaire où les concentrations de certains composés pouvaient légèrement varier selon la composition des plaques d'inhibiteurs, les concentrations molaires sont ici identiques pour tous les composés testés. Comme dans le test de criblage, nous nous sommes placés à 8 % final de DMSO dans l'essai afin de ne pas détériorer l'activité de hNEIL1 et de s'assurer une bonne solubilité des composés.

Sur les 50 hits obtenus lors du criblage de hNEIL1, seulement **18 ont été confirmés** comme inhibiteurs de l'activité glycosylase-lyase (Figure 76). La structure des hits confirmés est aussi rappelée en Annexe 5. Au final, 36 % des hits ont été confirmés ce qui correspond à 0,39 % des composés testés dans le criblage primaire. Nous avons donc 64 % de composés « faux positifs ». Il peut s'agir de composés possédant une fluorescence à une longueur d'excitation proche de celle du fluorophore utilisé, faussant donc les résultats. Des effets de plaque peuvent également influer sur les résultats apparents d'un criblage. L'un des composés semble « stimuler » l'activité de hNEIL1 à 150%, mais cette stimulation est faible et ce résultat est probablement dû à des variations expérimentales. Enfin, il faut noter que pour les banques ICOA2 et

Prestwick, deux composés ont été testés dans un même puits. Les composés spectateurs parmi ces deux molécules ont donc été éliminés lors de cette étape.

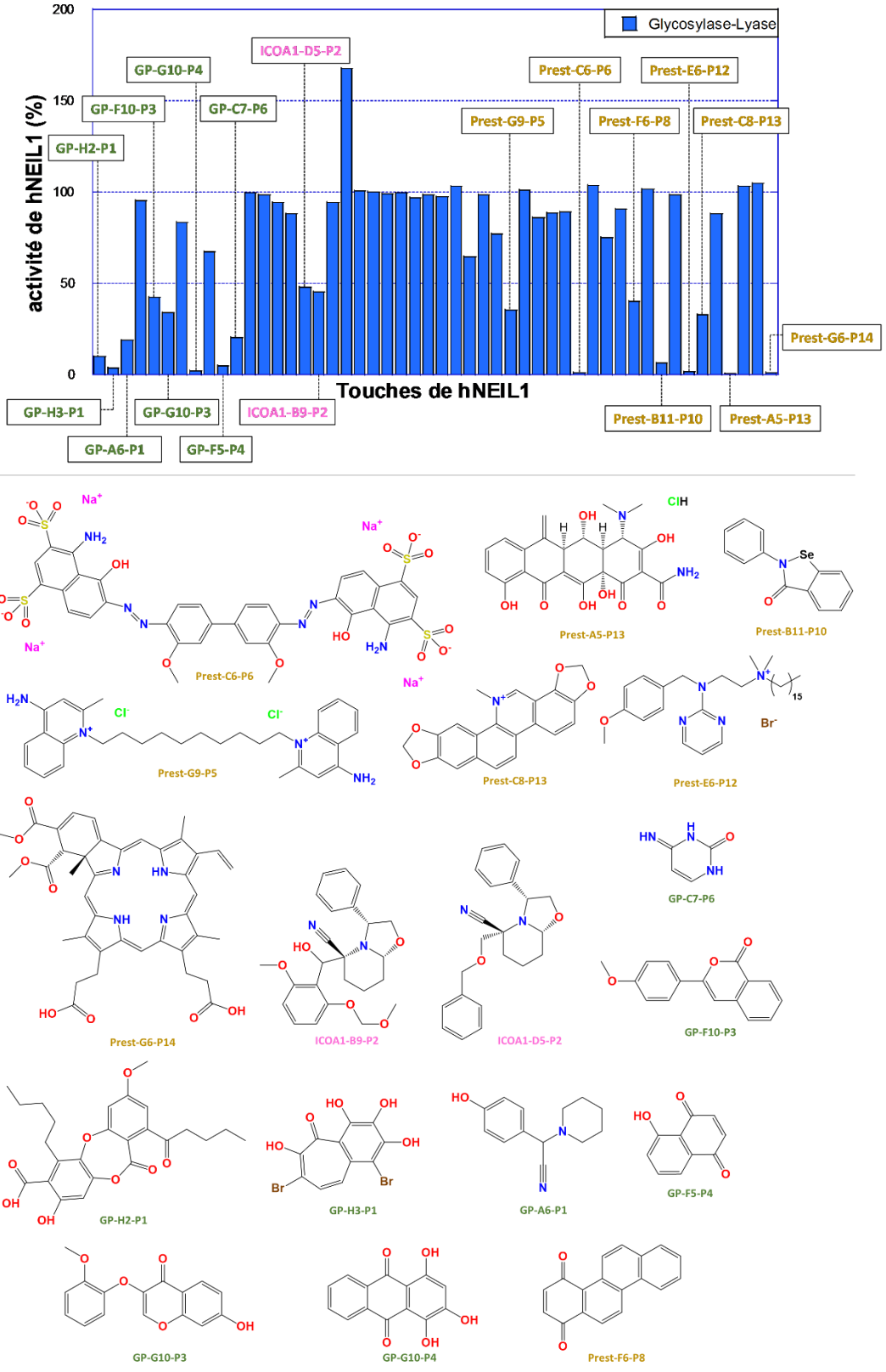


Figure 76 : Confirmation des 50 hits obtenus lors du criblage primaire de hNEIL1

II.3. Tests d'intercalation et première sélection des composés

Les composés inhibiteurs confirmés précédemment ont ensuite été testés en présence de PicoGreen et d'ADN double brin (Figure 77). Trois catégories de composés peuvent être distingués : les composés « non intercalant », qui ne font pas diminuer la fluorescence lue de plus de 20 % ; les composés « faiblement intercalant » qui se situent entre 40 et 80 % de la fluorescence initiale ; et enfin les composés « fortement intercalant », dont la fluorescence est inférieure à 40 %.

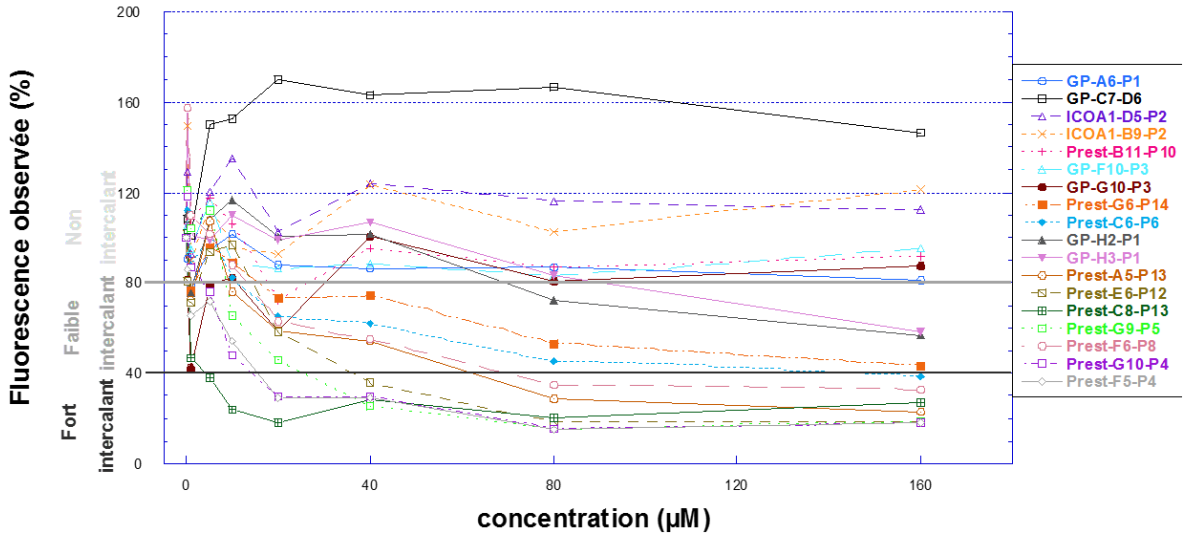


Figure 77 : Evaluation du pouvoir intercalant de l'ADN des hits confirmés par le test au PicoGreen

Concernant la séparation entre les catégories « faiblement » et « fortement intercalant », comme nous l'avons précisé, la baisse de la fluorescence est due à l'intercalation de la molécule dans l'ADN et du déplacement du PicoGreen (PG). Ainsi, l'impact sur le niveau de fluorescence dépendrait donc de l'affinité des composés avec l'ADN, expliquant ces variations de profil. Cependant des études indiquent que la mesure de la fluorescence de PG et son interaction avec des composés pharmaceutiques n'est pas quantitative, et peut donc varier fortement sans lien avec l'affinité du composé sur l'ADN [466]. Il est donc compliqué d'expliquer cette différence de profil de manière quantitative. En revanche, notre objectif n'étant pas de classifier l'affinité de nos composés pour l'ADN mais simplement de répondre qualitativement sur leur pouvoir intercalant. Ainsi, seuls les composés qui se montrent non-intercalant sont sélectionnés. Il est aussi à noter que la Juglone et la Purpurine, présents dans la banque de diversité **GP**, sont logiquement sortis du criblage et ont été confirmés par le test orthogonal. Comme nous nous y attendions, ces composés se sont révélés très intercalant, et n'ont donc pas été sélectionné. **7 composés inhibiteurs de hNEIL1** ont donc été retenus (Figure 78).

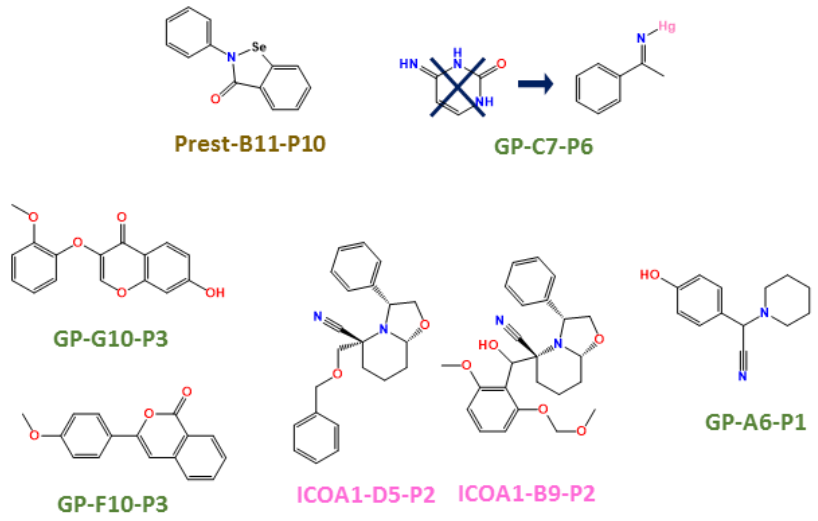


Figure 78 : Composés non-intercalant ayant été confirmés et validés lors des différentes étapes du criblage comme inhibiteurs de hNEIL1

II.4. Dose réponse et sélection finale

L'IC_{50} des composés sélectionnés a été déterminé par des essais de dose-réponse avec le test orthogonal utilisant un substrat radioactif. Bien que plus consommateur en temps et en matériels, cette méthode est cependant bien plus robuste et fiable que le test d'activité à la fluorescence utilisé dans le criblage primaire.

Afin de sélectionner les inhibiteurs les plus prometteurs parmi ces 7 composés, nous avons mis de côté les composés **GP-F10-P3** et **GP-G10-P3**, dont les IC_{50} étaient supérieurs à 50 µM. Nous avons ensuite dû faire au cas par cas pour étudier les composés à conserver. La molécule **Prest-B11-P10** est l'Ebselen, un composé bien connu pour ses capacités oxydantes et qui est utilisé dans de nombreux tests précliniques. Cependant, nous avons voulu vérifier si son mécanisme d'inhibition passait uniquement par son pouvoir oxydant. En absence de tris(2-carboxyéthyl)phosphine (TCEP), l'Ebselen inhibe bien hNEIL1. Par contre en présence de TCEP (0,1 mM), l'Ebselen n'a plus aucun effet sur hNEIL1. Nous avons donc rejeté cette molécule.

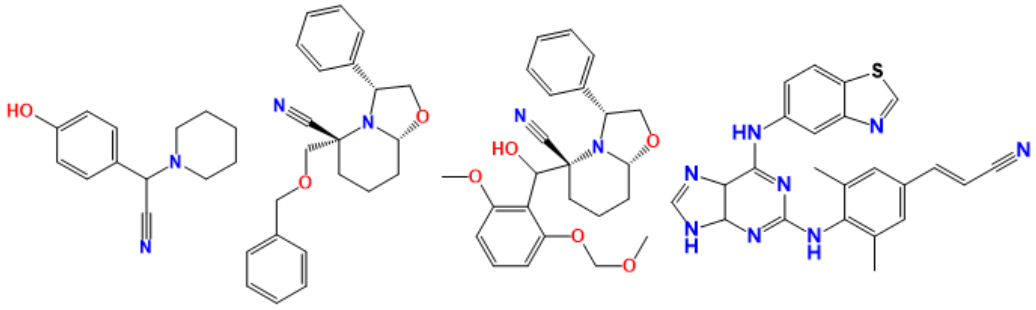
Un autre des 7 composés semblait être intéressant, le composé **GP-C7-P6** de la banque de GreenPharma, une entreprise spécialisée dans les petites molécules naturelles. Cependant certaines des molécules de la banque sont des molécules naturelles synthétisées, en proportion racémique. Le composé **GP-C7-P6** était supposé être la Cytosine ce qui était assez surprenant. Suite à ce résultat, nous avons donc tenté de tester les bases de l'ADN (C, T, G, A commerciales) pour un rapide test d'inhibition sur hNEIL1, et aucune inhibition n'a été observée, que ce soit pour les autres bases ou pour la cytosine contrôle. La cytosine utilisée dans ce second test provenait d'une poudre commerciale de Sigma Aldrich, tandis que celle du criblage provenait directement de la plaque commercialisée par GreenPharma. Nous avons donc effectué une analyse par spectrométrie de masse et renvoyé le composé à analyser à GreenPharma. Après quelques analyses, ce composé s'est vite révélé être une molécule différente de la cytosine annoncée. Par analyse RMN, GreenPharma nous a ensuite proposé une structure probable pour **GP-C7-P6**, qui contient un mercure Hg (Figure 78). Le mercure étant un composé souvent toxique, et la structure de la molécule n'étant qu'un modèle « probable » nous avons donc décidé de la mettre de côté.

Finalement, il reste **3 inhibiteurs finaux** de hNEIL1 (Figure 79). Sans ambiguïté, la molécule **GP-A6-P1** est l'inhibiteur le plus efficace de nos touches et bien meilleur que ceux de la littérature tels que P2. Ainsi, P2 affiche un IC_{50} de 4 ou 28 µM selon la méthode utilisée par Jacob *et al,.* (test radioactif ou par

fluorescence) [416]. La valeur la plus élevée ayant été obtenue avec un test similaire à notre test radioactif, les IC_{50} peuvent donc être comparés. Il reste important de rappeler que l'IC_{50} n'est pas une valeur absolue, et peut donc varier selon les conditions du test appliqué. Plus la concentration de l'enzyme sera importante, plus la quantité d'inhibiteurs, aussi affins soient-ils, devra logiquement être importante pour montrer le même effet d'inhibition. Les conditions de Jacobs *et al.*, utilisant une concentration quatre fois moins importante de hNEIL1 que celle que nous utilisons dans nos propres expériences, nous pouvons donc raisonnablement considérer que la molécule **GP-A6-P1**, ou sous son nom usuel la **Girgensonine**, est l'inhibiteur de hNEIL1 le plus efficace identifié à ce jour.

Les deux autres hits, **ICOA1-D5-P2** et **ICOA1-B9-P2** sont deux molécules structuralement très proches, mais avec des IC_{50} sensiblement plus élevés (Figure 79). Cette observation est particulièrement intéressante car elle peut permettre d'identifier un noyau structural inhibiteur, et donc guider une optimisation de nos composés par la synthèse d'analogues (pharmacomodulation).

L'élément structural commun à ces trois inhibiteurs est la fonction **α-aminonitrile** et pourrait être directement responsable de l'inhibition de hNEIL1 par ces composés. Le composé P2 possède lui aussi une fonction nitrile, mais conjuguée avec une double liaison, ce qui en fait un accepteur de Michael, ce qui n'est pas le cas de nos inhibiteurs. Malgré ces similarités structurales, nos composés peuvent donc avoir un mécanisme différent des inhibiteurs de la littérature.



	IC_{50} sur l'activité Glycosylase-Lyase (µM)	
Composé	*Valeur publiée*	*Test "maison"*
GP-A6-P1	--	4,23 ±0,62
ICOA1-D5-P2	--	17,59 ±3,21
ICOA1-B9-P2	--	46,03 ±6,74
P2	28,00 // 4,00	

Figure 79 : Evaluation de l'IC_{50} des inhibiteurs de hNEIL1 sortis du criblage - Le détail des différentes courbes permettant le calcul des IC_{50} est indiqué en Annexe 6.

II.5. Expansion des hits et optimisation des inhibiteurs

Suite à l'identification finale des hits du criblage précédent, notre objectif a été d'optimiser ces molécules afin d'améliorer leurs IC_{50}. En effet, bien que l'IC_{50} de la Girgensonine soit le plus efficace à ce jour, il reste de l'ordre du micromolaire, ce qui risque de ne pas être assez efficace pour des conditions *in cellulo*. De plus, l'identification d'un groupement commun chez nos inhibiteurs encourage également le développement d'analogues possédant, ou non, cette fonction afin d'évaluer son influence sur l'inhibition. Nous avons donc mis en place une stratégie d'expansion des molécules touches (« Hits expansion » en anglais), qui consiste généralement à tester des analogues structuraux des molécules dans leurs banques respectives. Des données de biologie structurale comme une structure cristallographique des inhibiteurs en présence de l'enzyme peuvent également permettre de mieux diriger l'optimisation chimique, en identifiant

les fonctions chimiques importantes dans le mécanisme d'inhibition. Cependant, comme nous le verrons dans les prochaines parties, les étapes de biologie structurale furent compliquées, voire impossible à finaliser, nous avons donc dû effectuer une expansion de nos molécules « à l'aveugle ».

II.5.1. Analogues de la girgensonine

La **Girgensonine** étant l'inhibiteur le plus efficace que nous ayons obtenu, nous avons donc cherché à tester des molécules similaires. Pour ce faire, nous sommes allés chercher dans la banque totale de GreenPharma (Ambinter), des molécules structuralement proches. Nous en avons identifié 6, avec l'aide de GreenPharma, dont les structures sont indiquées en Figure 80 A.

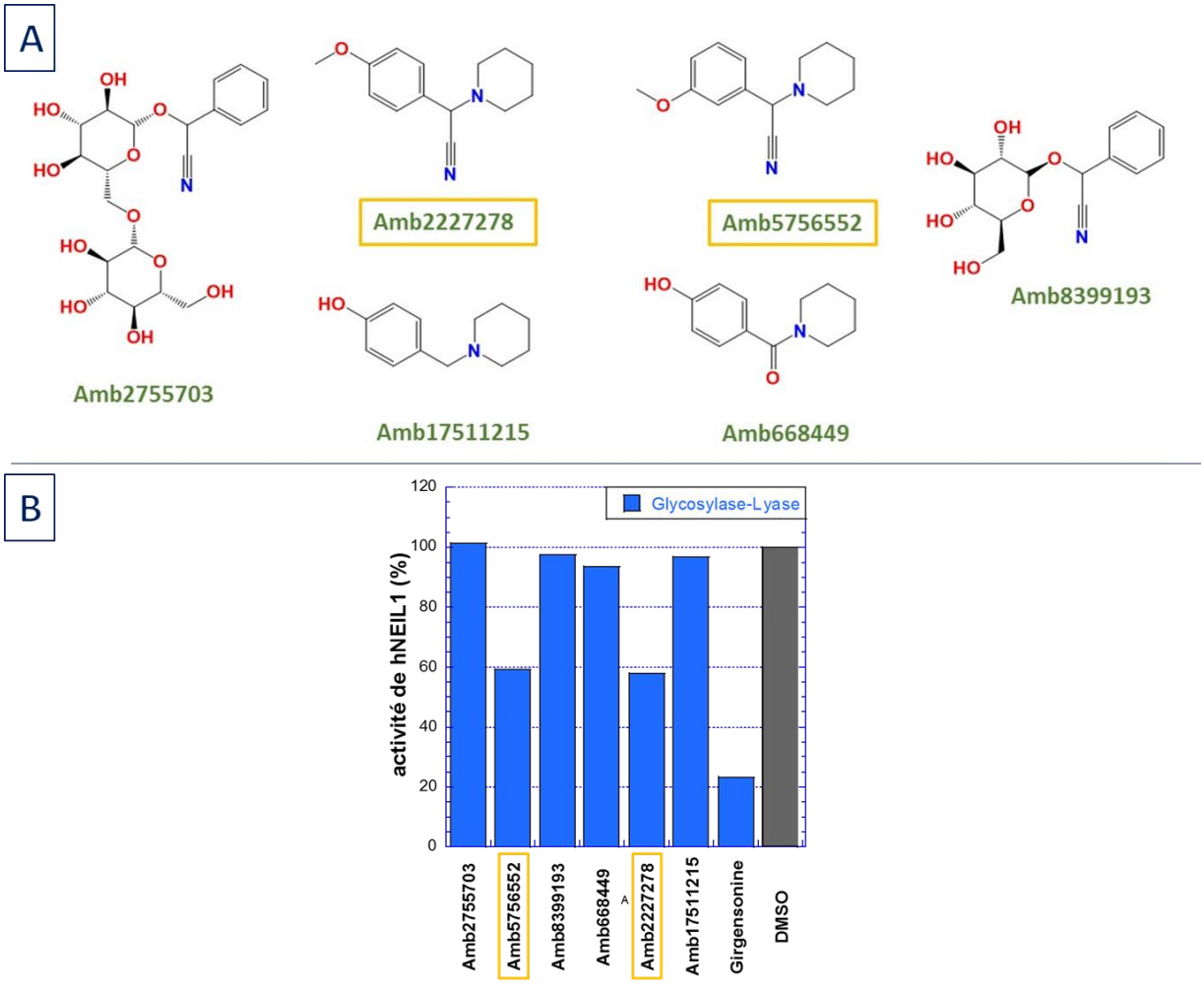


Figure 80 : Analogues de la Girgensonine présents dans la banque d'Ambinter - A) Structure des analogues testés. B) test d'activité de hNEIL1 en présence de 50 µM des analogues. Le contrôle DMSO est indiqué en gris foncé.

Malheureusement, aucune de ces molécules analogues ne s'est révélée être plus efficace que la **Girgensonine** (Figure 80). On peut remarquer d'ailleurs que seul deux d'entre elles sont effectivement inhibitrices de hNEIL1, du moins à la concentration de 50 µM, les molécules **Amb2227278** et **Amb5756552**.

Plusieurs informations peuvent être déduites de ces résultats. Tout d'abord, entre la Girgensonine et les molécules **Amb2227278** et **Amb5756552**, on voit un « shift » au niveau des liaisons H. Tandis que le phénol est un donneur de liaison H, sa forme éther-oxydée, l'anisol est un accepteur. On peut donc suggérer que l'inhibition de hNEIL1 est facilitée par cette liaison hydrogène, qui pourrait interagir avec un groupement carbonyle C=O du squelette de la protéine, un glutamate ou un aspartate, deux acides aminés étant généralement toujours déprotonés dans les protéines à pH physiologique, et donc, bons accepteurs de liaison

H. A ce propos, nous avons vu précédemment (voir Introduction) que le Glu2 est un résidu catalytique essentiel à l'activité de hNEIL1. Il est donc possible que la girgensonine puisse interagir avec son groupement C=O et se placer proche du site actif. Ensuite, on voit que les deux molécules ne disposant plus de la fonction nitrile (Amb17511215 et Amb668469, Figure 80) ne montrent plus d'effet d'inhibition, ce qui semble confirmer l'importance de ce groupement dans le mécanisme d'action des inhibiteurs. Les deux derniers composés quant à eux, possèdent bien une fonction nitrile, mais ne sont pas inhibiteurs pour autant. On peut donc suggérer que l'encombrement stérique autour de cette fonction nitrile est défavorable pour l'inhibition de hNEIL1, l'empêchant peut-être de se lier à hNEIL1 et de bloquer le clivage de la base oxydée. Cela suggère aussi une spécificité de la girgensonine pour hNEIL1. Si aucune amélioration n'a été obtenue lors de cette expansion des hits, des précisions sur l'importance de certaines fonctions chimiques ont pu être établies.

II.5.2. Analogues de B9-2 et D5-2

De la même manière, nous avons sélectionné, avec l'aide des chimistes de l'ICOA, les molécules analogues des composés **ICOA1-B9-P2** et **ICOA1-D5-P2** dans leur banque totale. Six composés plus ou moins proches structuralement ont donc été testés à 50 µM sur l'activité glycosylase-lyase de hNEIL1 (Figure 81).

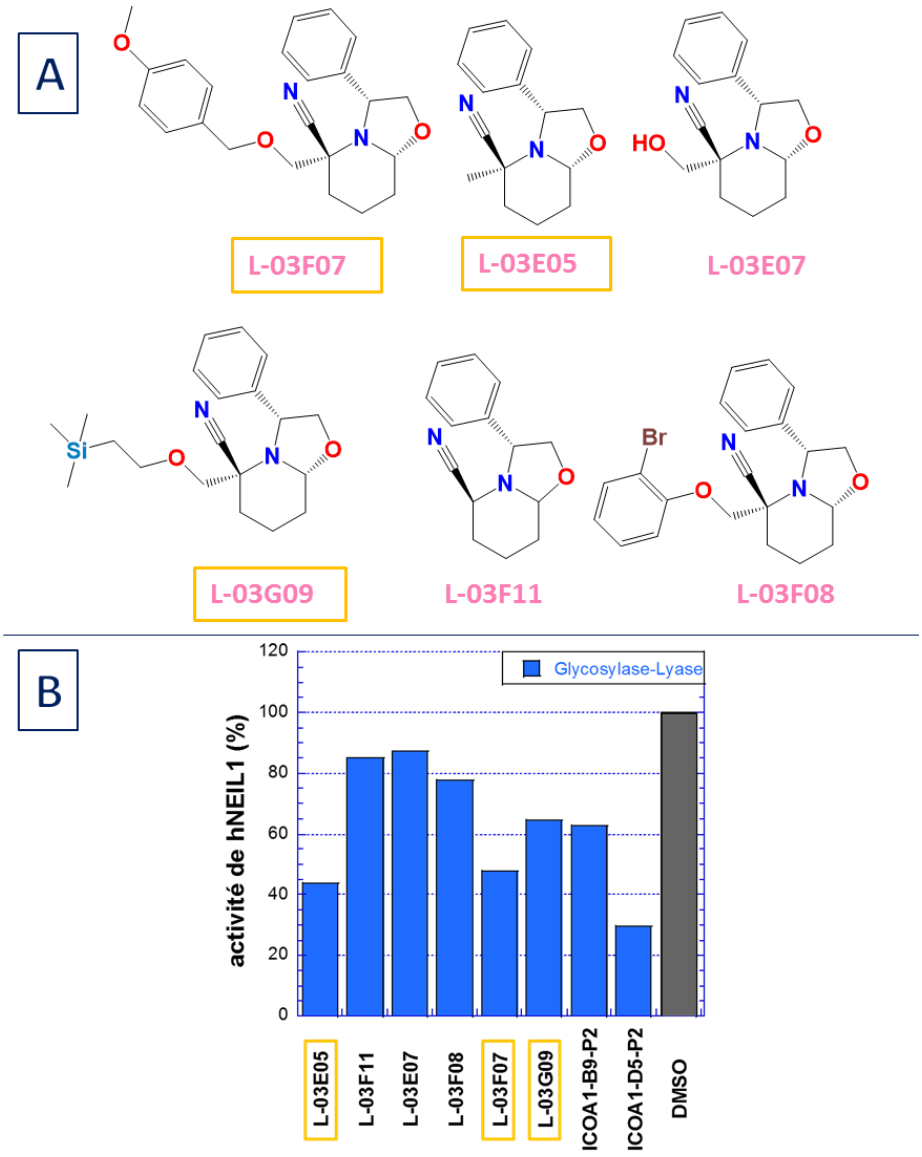


Figure 81 : Test des analogues des hits de la banque ICOA2 sur hNEIL1 - A) Structure des composés. B) Test d'activité glycosylase-lyase de hNEIL1 en présence de 50 µM. Le contrôle DMSO est indiqué en gris foncé.

De la même manière que pour les analogues précédents, aucune des molécules n'a montrée d'amélioration significative de l'inhibition de hNEIL1. Deux composés se sont révélés légèrement supérieurs à **ICOA1-B9-P2**, les molécules **L-03E05** et **L-03F07**, avec entre 60 et 50 % d'inhibition pour une concentration de 50 µM. En comparant les inhibiteurs du criblage, nous pensions que **ICOA1-D5-P2** pourrait être plus efficace grâce à un encombrement stérique plus faible autour de la fonction nitrile. Cela ne semble donc pas être aussi direct que cela, car **L-03E05** n'est pas plus efficace. Un autre composé, **L-03G09**, possède une efficacité comparable à **ICOA2-B9-P2**. Si l'on compare **L-03G09** avec l'**ICOA1-D5-P2**, nous pouvons supposer que le benzyle, plus apolaire que le triméthylsilyle, permet de se positionner plus facilement dans une poche hydrophobe, rendant le composé plus efficace. De même, pour l'**ICOA2-B9-P2**, l'éthylène glycol sur l'aromatique le rend également plus hydrophile que le groupement benzyle, expliquant potentiellement sa différence d'efficacité avec **ICOA2-D5-P2**. Pour le composé **L-03F08**, bien moins efficace, il est possible que l'atome de brome forme une gêne stérique trop importante pour que le benzyle puisse se retrouver dans une potentielle poche hydrophobe.

II.5.3. Molécules diverses

Compte tenu des différents résultats du criblage et du « hit expansion », il est devenu assez évident que s'il était complexe d'optimiser nos inhibiteurs, la fonction nitrile semblait ressortir comme la fonction chimique la plus essentielle de ces inhibiteurs. Il reste important de noter que la fonction nitrile seule ne signifie pas un effet d'inhibition, comme nous l'avons vu lors de l'expansion, mais aussi en regardant attentivement l'intégralité des composés testés lors du criblage, dont certains possédaient bien une fonction nitrile, sans être des hits. Ceci dit, nous avons cherché avec nos collaborateurs de l'ICOA d'autres molécules à tester possédant cette fonction nitrile. Nous avons donc pu tester différents composés possédant un nitrile, notamment la molécule **LA-32**, un triazine inhibiteur de cathepsines, dont la fonction nitrile attaque spécifiquement la cystéine catalytique de ces protéases à cystéine [467]. En effet, cette molécule peut former des liaisons réversibles avec les cystéines en passant par la formation d'un thioimidate résultant de l'attaque nucléophile du thiol sur le carbone à la base du nitrile (Figure 82). Par la suite ce thiomidate pourrait également évoluer en thiazoline de manière irréversible [468].

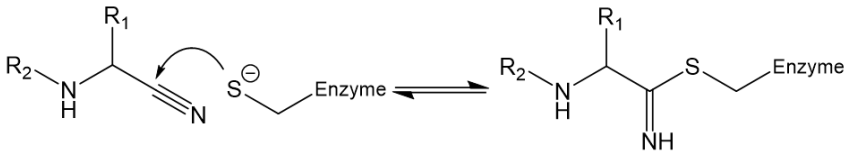


Figure 82 : Réaction d'un composé aminonitrile avec le groupement thiol d'une cystéine

LA-32 s'est révélée être un inhibiteur de hNEIL1, avec un IC_{50} = 21,4 µM (Figure 83). Les courbes dose-réponse ayant permis le calcul de l'IC_{50} de ce composé sont indiqués en Annexe 6. Nous pouvons donc imaginer que nos autres composés inhibiteurs, présentant également une fonction aminonitrile, puissent former le même type de liaison que ce composé avec les cystéines de hNEIL1. Le lien entre cet effet et l'inhibition de hNEIL1 sera étudié plus en détails dans les prochaines parties.

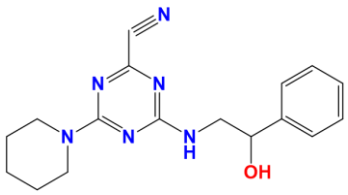


Figure 83 : Structure de l'inhibiteur LA-32

III. Analyses biochimiques et fonctionnelles des inhibiteurs de hNEIL1

Dans cette partie nous allons détailler les différentes caractérisations biochimiques de nos inhibiteurs et de leurs effets sur hNEIL1. En effet, l'objectif est d'étoffer notre compréhension de l'impact de nos inhibiteurs sur cette protéine et pour cela nous avons tenté différentes approches. Tout d'abord nous avons cherché à évaluer la **spécificité de nos inhibiteurs** sur des protéines proches. En effet, il est important de tester ces inhibiteurs sur des protéines plus proches structuralement et partageant des substrats. Nous avons également testé la stabilité de la protéine en présence des inhibiteurs et effectué des études enzymatiques de hNEIL1, afin d'évaluer potentiellement des **valeurs de K_D** de la protéine sur son substrat en présence de nos inhibiteurs, et déterminer le **mécanisme d'inhibition** de nos molécules « Leads », particulièrement la girgensonine.

III.1. Evaluation de la spécificité des inhibiteurs identifiés

Comme nous l'avons déjà évoqué, la spécificité d'un inhibiteur est essentielle pour envisager une utilisation thérapeutique de celui-ci. Bien que ce projet de thèse soit éloigné d'une application thérapeutique proche, il reste important de s'assurer que nos composés possèdent une certaine spécificité.

hNEIL1 fait partie de la superfamille structurale des ADN glycosylases Fpg/Nei. De ce fait, nous avons donc évalué l'activité du meilleur inhibiteur identifié, la girgensonine, sur d'autres enzymes de cette même famille, comme LlFpg. Pour rappel, LlFpg possède un doigt de zinc ZnF, ce qui n'est pas le cas de hNEIL1. Nous avons aussi décidé de tester la girgensonine sur mvNei1, qui comme hNEIL1 possède un ZnLF, et qui en plus ne contient aucune cystéine, afin de disposer d'enzymes présentant une certaine diversité structurale dans cette même famille.

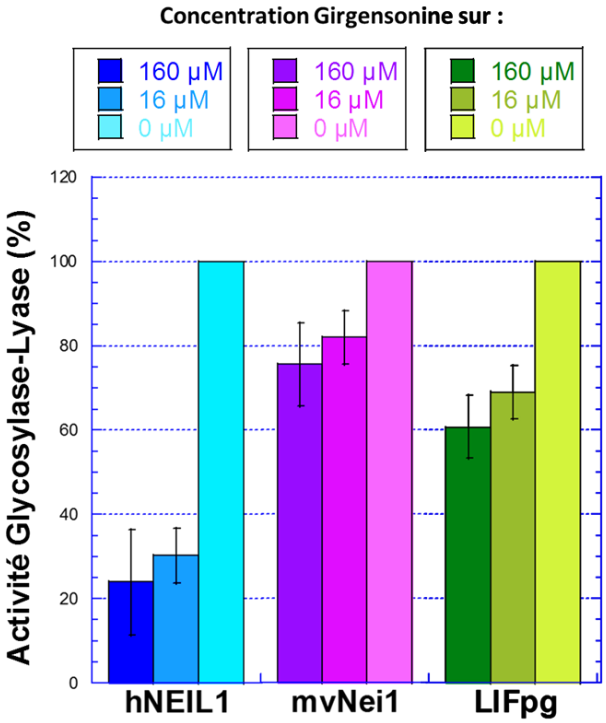
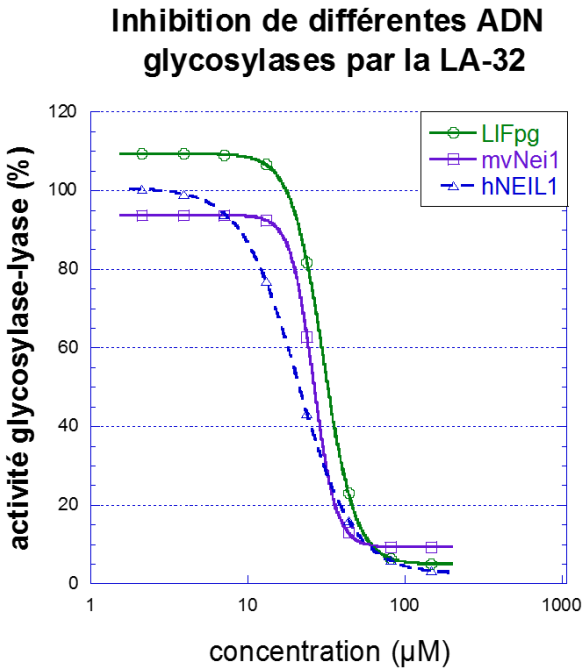


Figure 84 : Evaluation de l'inhibition de la girgensonine sur les enzymes hNEIL1, LlFpg et mvNei1 - Il est à noter que l'ADN substrat utilisé pour LlFpg possède une lésion 8-oxoG:C, tandis que pour hNEIL1 et mvNei1, l'ADN présente une lésion DHU:G. Les conditions des tests utilisés pour ces différentes enzymes sont indiquées dans le paragraphe Matériel et Méthodes dédié.

Nous avons testé l'effet de la girgensonine sur l'activité glycoylase/lyase de ces trois enzymes (Figure 84). Même si les trois enzymes sont plus ou moins bien inhibées par la girgensonine et ce de façon concentration dépendante, il semble que hNEIL1 soit la plus sensible. A ce stade, il est important de rappeler les premières hypothèses évoquées suite au criblage. En effet, nous avions évoqué la possibilité que la fonction nitrile des inhibiteurs puisse avoir une action sur une ou plusieurs cystéines de la protéine, comme c'est le cas pour LA-32 sur les cathepsines par exemple (Figure 82). Dans le cas de mvNei1, nous pouvons écarter cette hypothèse puisque cette enzyme ne possède pas de cystéine. La faible inhibition de mvNeil1 observée ne passe donc pas par un pontage covalent entre cystéine et girgensonine via son groupe nitrile. La fonction nitrile étant également capable de former un pontage avec des sérines, on peut suggérer que cette inhibition pourrait provenir d'une liaison avec une sérine. mvNei1 possède de nombreuses sérines, dont la plupart sont situées en N-terminal et sont peu impliquées dans la reconnaissance ou la prise en charge de l'ADN contenant une lésion. En revanche, la Sérine 272 située sur le ZnlF de mvNei1, stabilise la protéine sur l'ADN et un pontage sur ce résidu pourrait expliquer cette faible inhibition [378]. Pour LlFpg, qui possède 4 cystéines toutes impliquées dans la coordination de l'atome de zinc dans le domaine ZnF (et donc non libres), il n'est pas exclu que l'une ou plusieurs de ces cystéines puissent être modifiées par la girgensonine. Dans le cas de hNEIL1, la situation est différente car elle ne contient pas de domaine ZnF à cystéines. En revanche, elle possède 8 cystéines libres, qui sont donc de potentielles cibles pour la girgensonine. A ce stade de notre étude, il est donc difficile d'affirmer qu'une ou plusieurs cystéines de hNEIL1 ou LlFpg sont des cibles pour la girgensonine selon le mécanisme proposé (Figure 82) puisque mvNei1, dépourvue de cystéine, est tout de même faiblement inhibée par la girgensonine. La seule chose que nous pouvons avancer est qu'il existe une corrélation (peut-être complètement indirecte) entre la présence de cystéines dans les protéines Fpg/Nei et la capacité accrue de la girgensonine à les inhiber. La fixation non-covalente de la girgensonine cumulée à la possibilité de ponter des cystéines libres pourraient expliquer que, parmi les ADN glycosylases Fpg/Nei testées ici, hNEIL1 est l'enzyme la plus sensible à la girgensonine.

En parallèle, nous avons comparé l'inhibition de LA-32 sur ces trois mêmes protéines (Figure 85). LA-32 inhibe les ADN glycosylases Fpg/Nei de façon très similaire, ce qui suggère que son mécanisme d'inhibition ne passe pas uniquement par sa capacité à former des pontages covalents avec des cystéines. On peut cependant relever que la girgensonine est un inhibiteur bien plus sélectif et efficace pour hNEIL1 que LA32.



Protéine cible	IC_{50} (µM)
LlFpg	30,41 ± 1,32
mvNei1	26,21 ± 1,36
hNEIL1	20,84 ± 1,54

Figure 85 : Courbe d'IC_{50} du composé LA-32 sur les ADN glycosylases LlFpg, mvNei et hNEIl1

En conclusion, si le mécanisme d'inhibition reste flou, nous avons pu cependant identifier deux profils différents, avec d'une part un inhibiteur relativement spécifique, la girgensonine, et d'autre part LA-32, un inhibiteur semblant avoir un effet similaire sur les différentes enzymes de la superfamille Fpg/Nei. D'autres investigations peuvent être envisagées pour aller plus loin. Tester nos inhibiteurs sur les autres enzymes humaines hNEIL2 (7 cystéines dont 3 engagées dans le ZnF) et hNEIL3 (27 cystéines dont probablement plus de la moitié engagées dans plusieurs ZnF), pourrait également être considéré, afin de préciser la sélectivité de ces inhibiteurs sur les différentes enzymes humaines de la famille Fpg/Nei, et d'évaluer l'impact d'un potentiel pontage inhibiteur/cystéine.

III.2. Adduits covalents induits par les inhibiteurs de hNEIL1 contenant une fonction nitrile : analyse par spectrométrie de masse (MS)

Afin de caractériser l'effet de nos inhibiteurs sur hNEIL1, nous avons également évalué l'apparition d'un adduit de nos composés sur l'enzyme. Cette approche est motivée par la similarité entre nos composés, et notamment le groupement aminonitrile. Comme nous l'avons déjà évoqué, le nitrile peut former un lien covalent avec une cystéine (Figure 82). L'objectif est donc d'étudier si un tel adduit peut se former avec hNEIL1, et si un tel évènement pourrait en partie expliquer l'inhibition de l'enzyme par nos composés.

III.2.1. Mise en évidence d'un adduit covalent entre la girgensonine et hNEIL1

- **Première observation**

La Girgensonine à 800µM (8% DMSO) a été incubée avec 2µM de la protéine hNEIL1 entière à 4°C pendant la nuit. Le mélange d'incubation a ensuite été analysé par spectrométrie de masse ESI-QTOF (« ElectroSpray Ionization Quadrupole Time Of Flight» en anglais). Le spectre déconvolué qui en résulte est présenté en Figure 86.

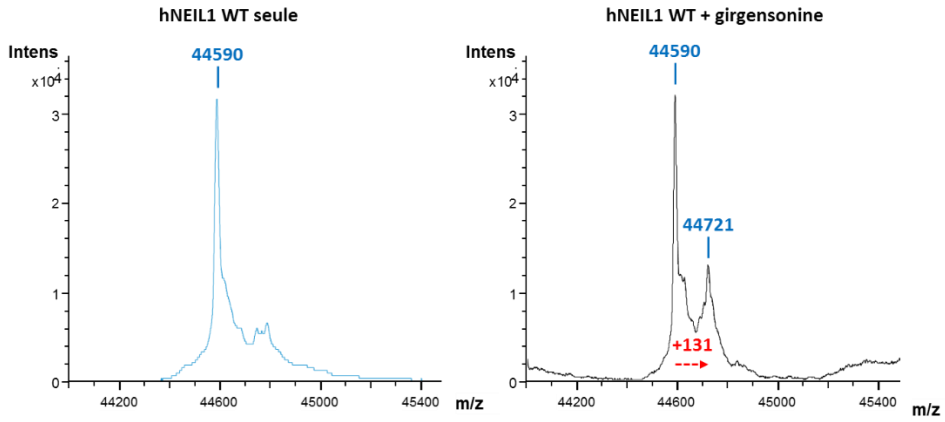


Figure 86 : Spectre déconvolué de la protéine hNEIL1 incubée avec la molécule girgensonine - La masse théorique de hNEIL1 entière non modifiée est égale à 44590,9 Da.

Nous pouvons voir sur ce spectre qu'une partie de la protéine a été modifiée par un adduit de 131 Da. Cependant, cet écart de masse ne correspond pas à celui attendu par une réaction covalente avec la girgensonine (215 Da), soit une différence de masse de -84 Da. Cette perte de -84 Da pourrait venir d'une fragmentation de la girgensonine au cours de l'analyse par spectrométrie de masse (Figure 87). Afin de tester cette hypothèse, nous avons décidé d'étudier la stabilité de la girgensonine dans le spectromètre de masse .

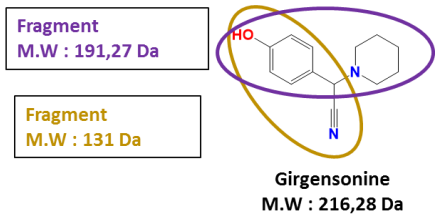


Figure 87 : Structure de la girgensonine et fragmentations potentielles

- **Etude de la stabilité de la girgensonine au cours de l'analyse par spectrométrie de masse ESI-IT (ElectroSpray Ion Trap)**

Nous avons analysé le comportement de la girgensonine seule par ESI-IT en augmentant l'énergie transmise aux ions lors de l'analyse (Figure 88). Cette énergie transmise est reflétée par le paramètre « compound stability ». Comme nous pouvons le voir, une augmentation de l'énergie induit une diminution du pic *m/z* 217 et une augmentation du *m/z* 132. On remarque aussi un fragment de *m/z* 190.1, ce qui pourrait correspondre à la perte du nitrile (entouré en violet, Figure 87). On peut conclure de cette expérience que la girgensonine est sensible à l'analyse MS. Concernant l'analyse MS du complexe hNEIL1/girgensonine, un seul adduit covalent est observé (+131 Da). On peut donc émettre l'hypothèse que la girgensonine est fragmentée dans le spectromètre de masse au cours de l'analyse, et que cet adduit réagit ensuite sur la protéine.

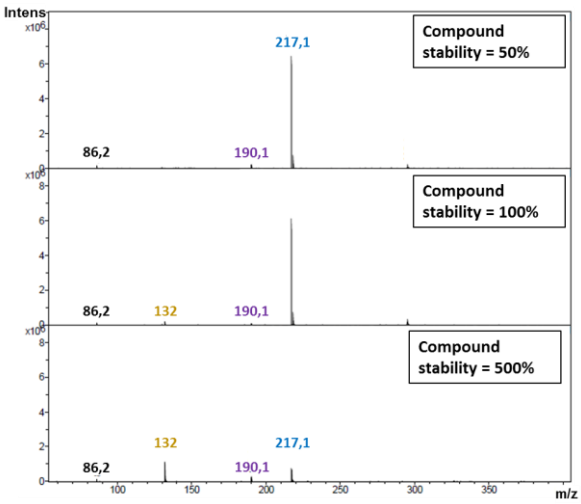


Figure 88 : **Evaluation de la stabilité de la girgensonine en ESI-IT MS en faisant varier la tension de l'appareil. La valeur du « compound stability » est liée à la tension au niveau de la transmission des ions vers l'analyseur. Cette valeur représente plus généralement l'énergie transmise aux ions. Une valeur de 100% correspond à la tension standard de lecture du spectromètre de masse IT.**

- **Cinétique de la formation de l'adduit suivi par spectrométrie de masse**

Afin de vérifier si cet adduit est uniquement un artefact de l'analyse par spectrométrie de masse, nous avons décidé d'effectuer quelques expériences de cinétique afin de vérifier si des profils différents pouvaient être obtenus. Pour ce faire, l'évolution des mélanges réactionnels contenant 2 µM de la protéine hNEIL1dc95 incubés avec 800 µM ou 20 mM de girgensonine ont été suivis au cours du temps par spectrométrie de masse (Figure 89). Comme attendu, on observe que la quantité de molécules de hNEIL1 modifiées par l'adduit de 131 Da augmente en fonction du temps, et également en fonction de la concentration de girgensonine. De fait, cet adduit n'est donc pas uniquement un artefact généré par l'expérience de spectrométrie de masse, mais découle bien d'une réaction qui se fait en solution. Il faut noter que nous utilisons ici une forme tronquée de hNEIL1 (hNEIL1dc95) ne possédant pas les 95 derniers acides aminés suggérant donc que le site de pontage de la girgensonine n'est pas contenu dans ses 95 derniers résidus d'acides aminés. On peut émettre l'hypothèse qu'après interaction de la girgensonine avec hNEIL1, celle-ci est capable de se lier covalemment à la protéine hNEIL1, et que dans les conditions de spectrométrie de masse utilisées, l'adduit initial attendu de 215 Da (girgensonine intacte) est immédiatement fragmenté en l'adduit de 131 Da (Figure 87).

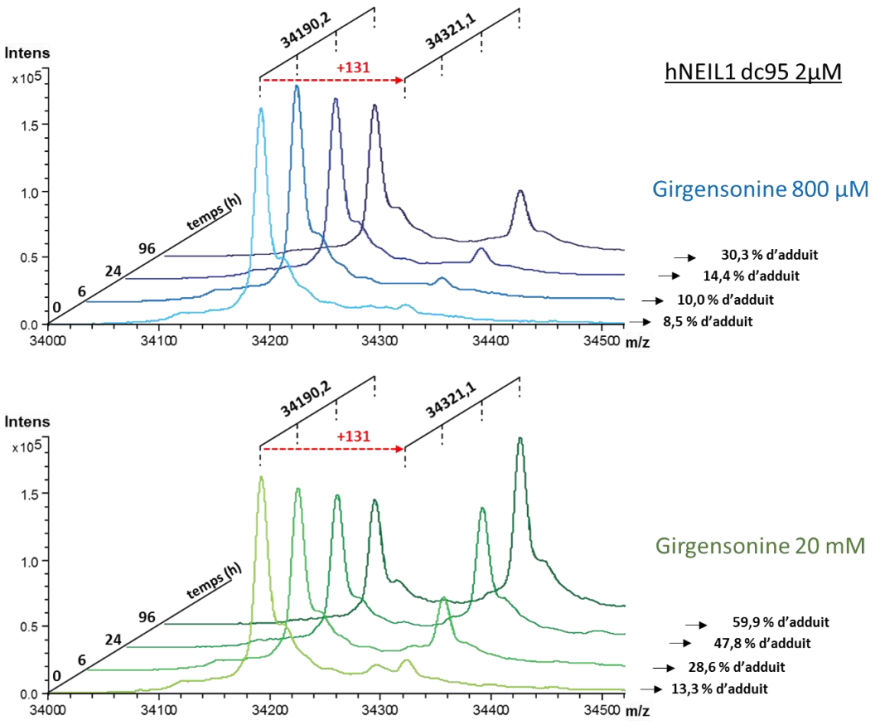


Figure 89 : Cinétique de la formation d'adduits de la girgensonine sur hNEIL1dc95 suivie par spectrométrie de masse ESI - La masse moyenne théorique de la forme tronquée hNEIL1dc95 est égale à 34189,6 Da.

En revanche, malgré plus de 96h d'incubation, le pic de protéine non modifiée (34190.2) est toujours observée. Il est donc peu probable d'obtenir 100 % de complexe entre la protéine et le ligand, la protéine se dégradant après quelques jours. Or, obtenir un échantillon uniquement constitué du complexe covalent hNEIL1-girgensoninepourrait être intéressant pour identifier le site protéique modifié par l'inhibiteur en déterminant la structure tridimensionnelle du complexe covalent par diffraction des rayons X. Dans ce but, nous avons évalué l'effet de différents paramètres afin de comprendre la nature de cette liaison ou d'augmenter le pourcentage de complexe obtenu. Ainsi, l'impact de la température a été évalué en incubant la protéine hNEIL1 entière ou dc95, en présence de girgensonine pour deux températures différentes, 4 °C et 20 °C (Figure 90).

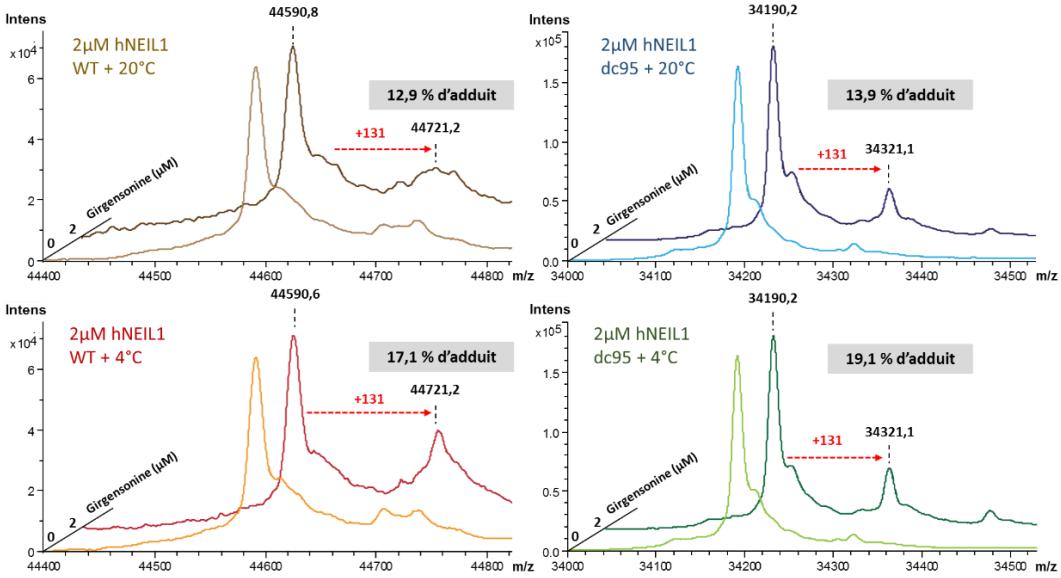


Figure 90 : Spectres déconvolués de l'adduit formé par la girgensonine en fonction de la température

Si l'on aurait pu s'attendre à une réaction plus rapide à température ambiante, nous pouvons observer que ce n'est pas le cas, et qu'au contraire, une température plus élevée semble même être défavorable. De plus, la protéine se conserve très mal à température ambiante, et a tendance à précipiter rapidement. Pour la suite des expériences, nous garderons donc nos échantillons à 4 °C. On ne remarque pas de différence significative pour la forme entière ou tronquée de hNEIL1, qui forment le même adduit comme précédemment (Figure 90). Cependant, il semble que la proportion de cet adduit soit plus élevée pour la forme tronquée hNEIL1dc95, qui se conserve également mieux et est plus stable. Pour la suite des expériences, nous utiliserons hNEIL1dc95 uniquement. En revanche, aucune franche amélioration n'a été observée, et la formation de 100 % d'adduit n'a pas pu être atteinte.

Il est possible que l'inhibition de hNEIL1 par la girgensonine passe par la fonction nitrile de celle-ci, qui pourrait elle-même être impliquée dans le pontage covalent avec un ou plusieurs thiols de la protéine. En revanche, avec une telle concentration d'inhibiteur et une cinétique de pontage aussi lente, il est peu probable que l'adduit observé soit à lui seul responsable de l'inhibition de la protéine. L'inhibition de hNEIL1 par la girgensonine observée dans des temps courts d'incubation est probablement liée à une fixation spécifique et non-covalente, ou du moins réversible, de l'inhibiteur sur un ou plusieurs sites discrets de la protéine pouvant ensuite évoluer vers un pontage covalent. Cette hypothèse est aussi supportée par l'observation que mvNei1 (dépourvu de cystéine) est aussi inhibée par la girgensonine.

III.2.2. Mise en évidence de la formation d'adduits covalents avec d'autres inhibiteurs de hNEIL1 contenant une fonction nitrile

Suite à ces différents résultats avec la girgensonine, nous avons voulu vérifier si les autres inhibiteurs que nous avons identifiés sur hNEIL1, permettait également la formation d'un adduit covalent. Nous avons donc incubé les composés ICOA1-D5-P2, LA-32 et ICOA1-B9-2 à 800 µM avec 2 µM de hNEIL1dc95 sur la nuit, à 4 °C. Nous avons également testé un analogue de la girgensonine, dont nous avons déjà évoqué l'activité, le composé Amb2227278. Les mélanges d'incubation ont ensuite été analysés par spectrométrie de masse dans les mêmes conditions qu'avec la girgensonine (Figure 91). Nous observons des adduits pour les composés Amb2227278 et ICOA1-D5-P2, mais aucun pour LA-32 malgré une fonction nitrile sur cet inhibiteur. Pour le composé ICOA1-B9-P2, un adduit est très faiblement visible, avec une protéine ayant fortement précipité au cours de l'incubation. Ces adduits sont tous différents en masse, mais différentes informations peuvent en être extraites.

Tout d'abord pour le composé Amb2227278, l'adduit observé correspond à la perte du même groupe chimique que celui observé pour la girgensonine. Cela semble donc confirmer que l'adduit observé est bien dû à la perte du groupement pipéridine. Pour les composés ICOA1-B9-P2 et ICOA1-D5-P2, on observe également la même perte de masse pour ces deux composés, soit 119 Da, bien qu'il soit compliqué d'isoler un fragment réaliste de masse identique sur ces deux composés (Figure 92).

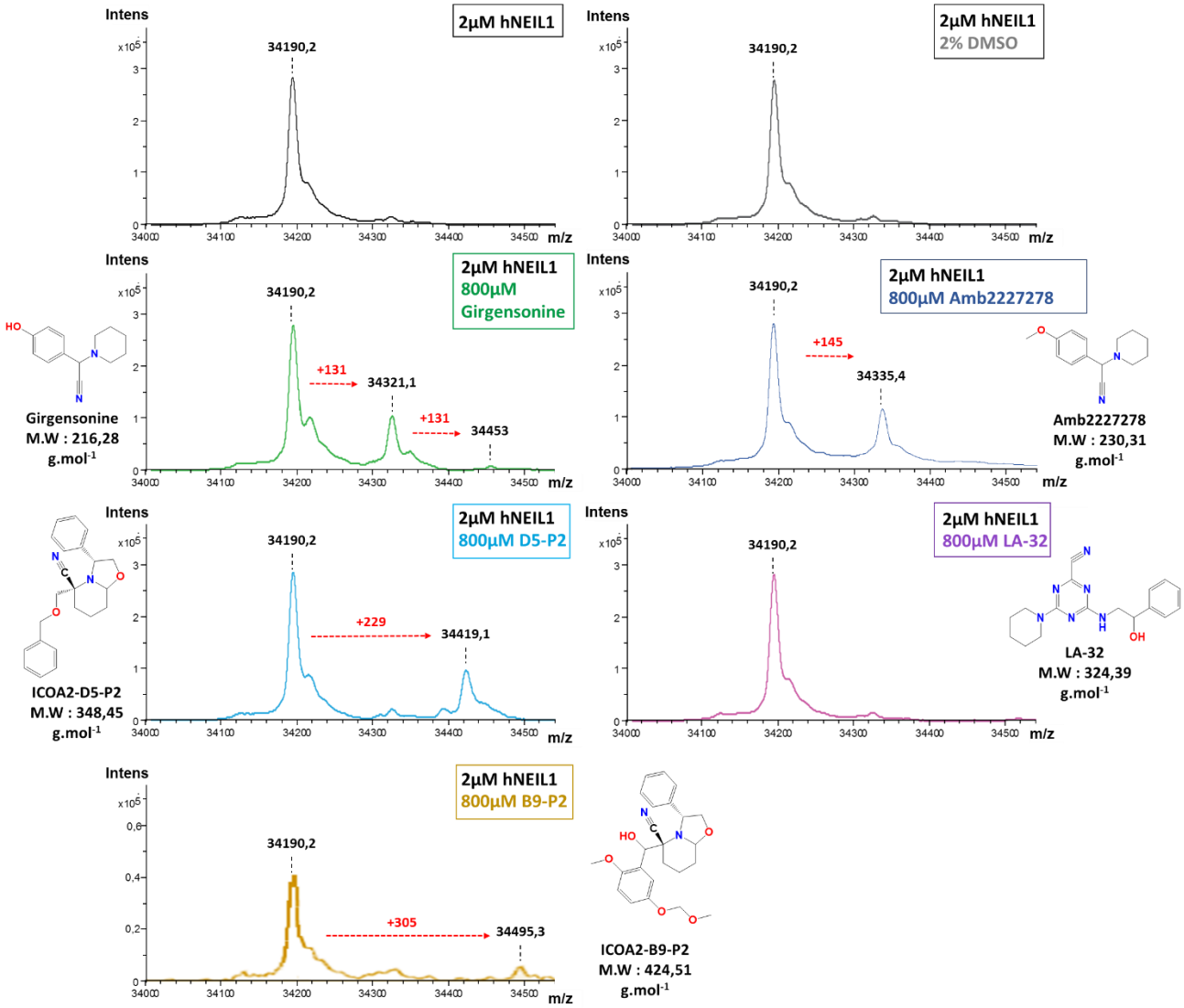


Figure 91 : Formation d'adduits avec d'autres inhibiteurs de hNEIL1 suivie par spectrométrie de masse - Masse moyenne théorique de la protéine tronquée hNEIL1dc95 = 34189,6 Da et des inhibiteurs ; girgensonine = 216,28 Da, LA-32 = 324,39 Da, ICOA1-D5-P2 = 348,45 Da, ICOA1-B9-P2 = 424,51 Da, Amb2227278 = 230.31 Da.

De même que pour la girgensonine, nous avons évalué si les adduits observés résultaient d'une fragmentation de l'adduit initial induite par l'expérience de spectrométrie de masse (Figure 92). Pour les trois inhibiteurs capables de former des adduits sur hNEIL1, l'analyse par spectrométrie de masse est associée à une perte d'un fragment de 25 Da favorisée par une augmentation de la tension dans le spectromètre et correspondant à une perte du nitrile. Pour le composé ICOA1-D5-P2, la perte du nitrile forme un fragment de 323 Da. On peut également voir l'apparition d'un fragment de 230 Da, correspondant à l'adduit observé en masse. En revanche pour ICOA1-B9-P2, aucun fragment ne correspond à l'adduit observé précédemment. Enfin le composé Amb2227278 se comporte comme la girgensonine, avec des fragmentations correspondant à la perte progressive du groupement pipéridine et du nitrile. De ces résultats, nous pouvons conclure que c'est encore l'expérience de masse qui induit une fragmentation des inhibiteurs libres en solution ou pontés sur hNEIL1. Ces composés semblent partager des groupements partants similaires en fonction de leurs structures, comme la perte du nitrile ou de la pipéridine.

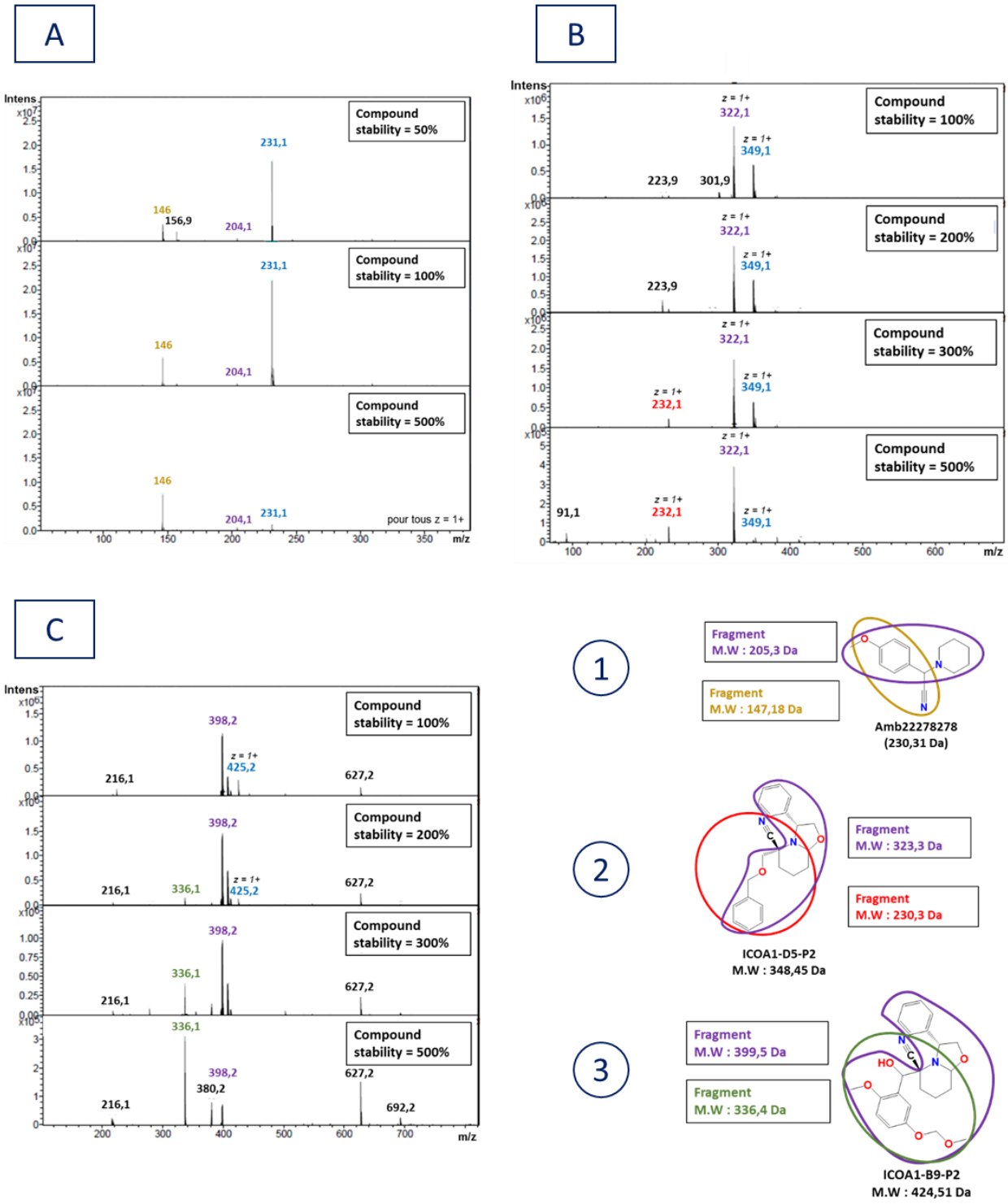


Figure 92 : Fragmentation des composés LA-32 (A-1), ICOA1-D5-P2 (B-2) et ICOA1-B9-P2 (C-3) en spectrométrie de masse

III.2.3. Localisation des adduits de la girgensonine sur hNEIL1 par fragmentation de la protéine

Nous avons ensuite tenté de déterminer où les adduits de la girgensonine pouvait se lier sur notre protéine d'intérêt. Pour ce faire, nous avons utilisé différentes techniques de fragmentation d'une protéine : une approche dite « bottom-up » et une autre dite « top-down » qui diffèrent par la méthode utilisée pour fragmenter les protéines (le principe de ces stratégies est indiqué en Matériel et Méthodes). L'idée générale

de ces deux méthodes est la même et consiste à segmenter la protéine en plusieurs fragments, et parmi ceux-ci à identifier le(s) fragment(s) porteurs d'un ou plusieurs adduits. L'approche « top-down » va consister à fragmenter la protéine directement dans le spectromètre de masse, tandis que pour l'approche «bottom-up », la fragmentation se fait en amont avec l'aide d'une protéolyse enzymatique.

- **Approche « top-down »**

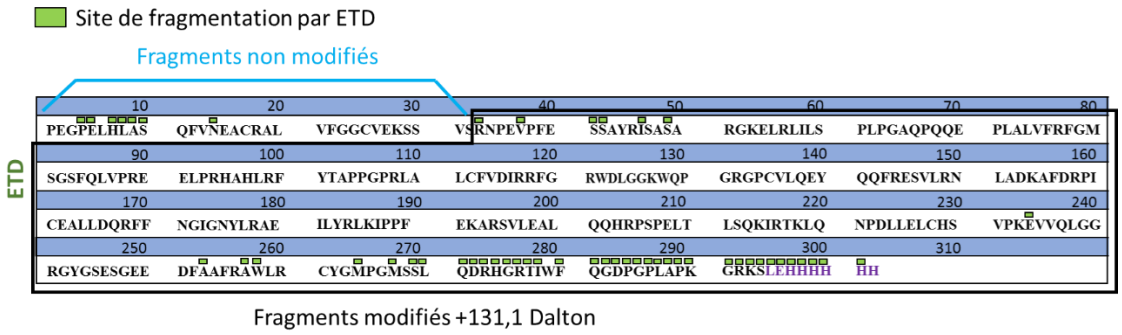


Figure 93 : Couverture de fragmentation de hNEIL1 en présence de girgensonine par ETD

Nous avons représenté en Figure 93 la couverture de fragmentation obtenue grâce à la fragmentation ETD de hNEIL1 incubée avec la girgensonine, c'est-à-dire les différents acides aminés où les clivages peptidiques ont été observés. Comme nous pouvons le voir, la couverture n'est pas optimale, et nous fragmentons principalement les extrémités de la séquence. Les résultats de l'analyse sont détaillés en Annexe 7. Lors de la fragmentation par ETD, on remarque la présence d'adduits de 131 Da dans les fragments en N-terminal, à partir de l'arginine 33. Cela pourrait donc correspondre à certaines cystéines de la protéine, et d'autres expériences ont été mises en place pour tenter d'isoler plus précisément le résidu modifié.

- **Approche « bottom-up »**

Suite à ces résultats incomplets en terme de couverture de séquence, nous avons utilisé l'approche « bottom-up » avec la trypsine comme protéase et le MALDI-TOF MS pour identifier les peptides (Figure 94). Les expériences de protéolyse ont été réalisées sur la protéine en présence de la girgensonine ou de ICOA1-D5-P2. Les résultats sont décrits en Annexe 7. Nous avons pu observer des adduits de +25 Da sur les peptides [1-18] et [133-144] pour la girgensonine et sur le peptide [1-18] pour ICOA1-D5-P2. On peut aussi remarquer un faible adduit à 131 Da pour la girgensonine. L'adduit de 25 Da pourrait correspondre au nitrile C≡N, qui résulterait d'une sous fragmentation de l'adduit initial.

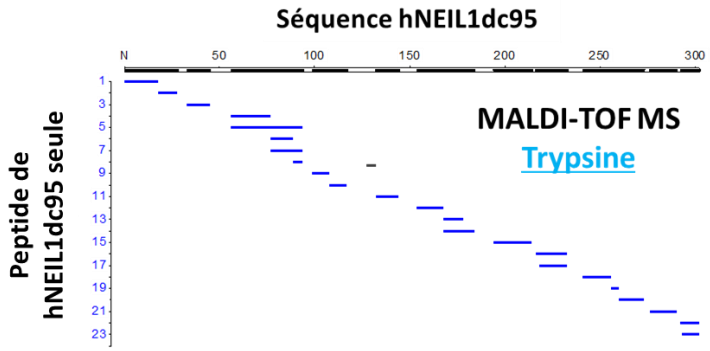


Figure 94 : Couverture de séquence pour la protéolyse de hNEIL1dc95 par la trypsine

En définitif, si l'on recoupe les résultats des différentes approches « bottom up » et « top down », il semble bien qu'un adduit covalent entre ce type d'inhibiteur et hNEIL1 se réalise de façon préférentielle avec un ou des résidus de l'enzyme. Ces résidus sont inclus dans un peptide interne de hNEIL1 s'étendant du résidu 133 au résidu 144 (situé à l'interface entre les 2 domaines globulaires de la protéine et assurant leur connexion physique), ou dans les 18 premiers acides aminés en N-terminal de l'enzyme.

Afin d'identifier plus précisément les acides aminés potentiellement modifiés, nous avons fragmenté les peptides trypsiques par nanoLC-MS/MS CID. Les résultats sont décrits en Annexe 7b. Cette analyse montre que les cystéines C18, C136 et C262 sont modifiées, que ce soit avec la girgensonine ou avec ICOA1-D5-P2. Ces cystéines peu exposées au solvant ne semblent pas être impliquées directement dans l'activité catalytique de hNEIL1 mais pourraient jouer un rôle dans la structure tertiaire de hNEIL1 en assurant le bon l'empilement de structures secondaires localisées dans des sites clés de l'enzyme (Figure 95).

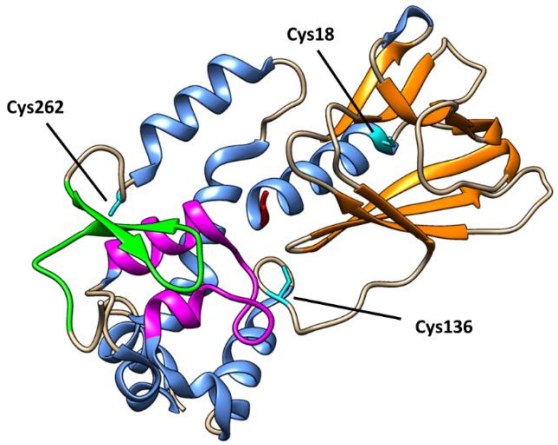


Figure 95 : Position des cystéines C18, C136 et C262 sur hNEIL1 - Les hélices α sont colorées en bleu, les brins β en orange, le domaine H2TH en magenta et les ZnF ou ZnLF en vert clair, les résidus catalytiques en rouge, et les cystéines C18/136/262 en cyan (PDB : 4NRV).

Tout d'abord, C18 est localisée en C-terminal de l'hélice α portant la Pro1 catalytique (Figure 96 A). Elle forme ainsi avec la Leu21 et Val14 des liaisons H ce qui permet un bon empilement de l'hélice α sur le feuillet β du domaine globulaire N-terminal et contribue ainsi à un positionnement optimal de la Pro1 au niveau du site actif de l'enzyme. Sa modification par l'inhibiteur pourrait être associée, à un changement conformationnel local et perturber l'activité de hNEIL1. C262 est à la base de la structure en épingle (hairpin) bêta du ZnLF impliqué dans l'interaction avec l'ADN (Figure 96 B). Elle forme des liaisons H avec la Glu181 et Met265 et interagit également avec le domaine H2TH. Ici aussi un adduit covalent sur la cystéine pourrait limiter la capacité de hNEIL1 à interagir avec l'ADN. Enfin, C136 se situe en bord du peptide « linker » inter-domaine. Sa modification pourrait avoir une influence sur l'orientation relative des domaines globulaires N- et C-terminaux de hNEIL1 et impacter ainsi le bon fonctionnement de l'enzyme.

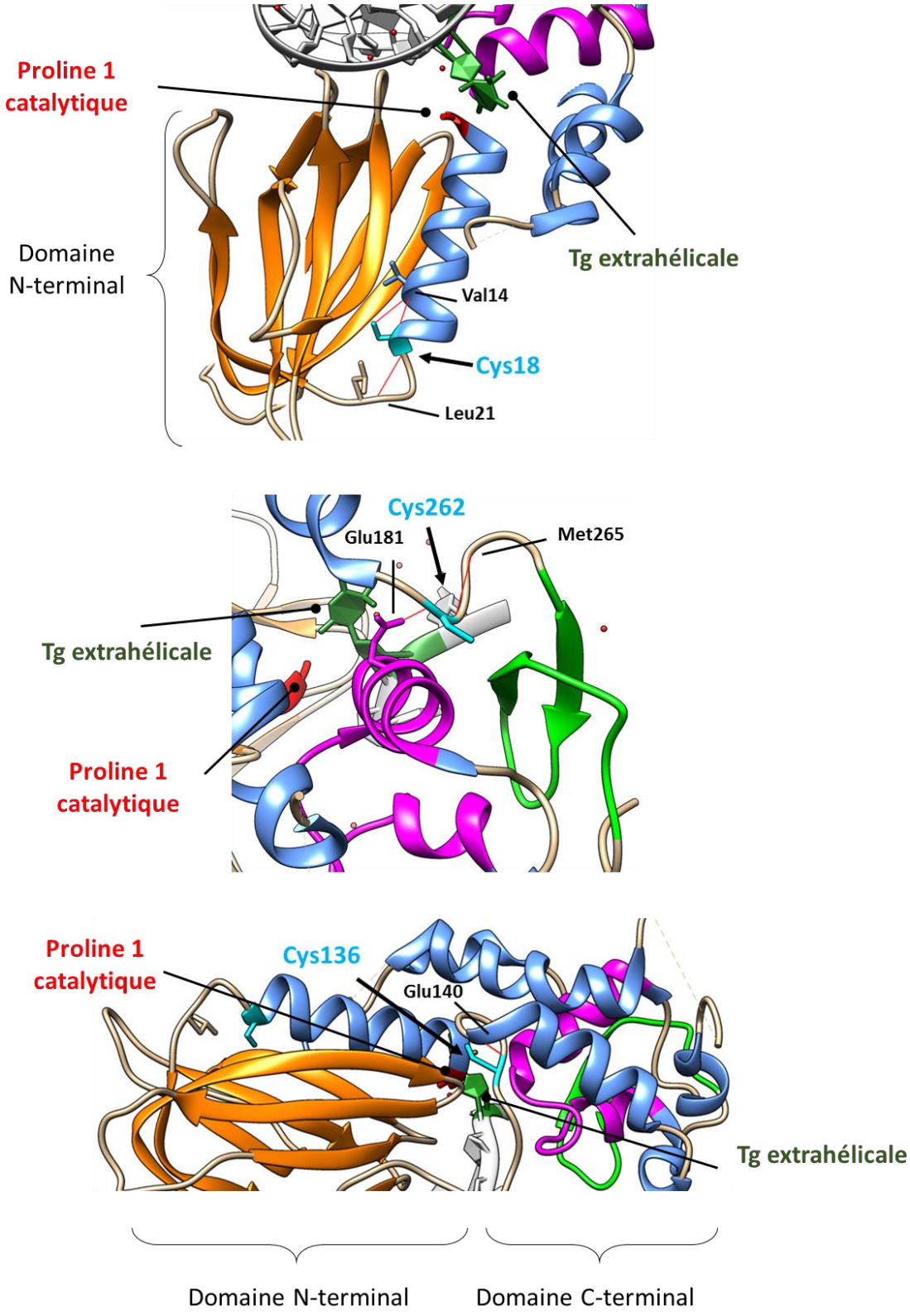


Figure 96 : Cystéines C18, C136 et C262 et leur impact sur la conformation de hNEIL1 - Les hélices α sont colorées en bleu, les brins β en orange, le domaine H2TH en magenta et les ZnF ou ZnLF en vert clair, les résidus catalytiques en rouge, la thymine glycol en vert foncé et les cystéines C18/136/262 en cyan. Les liaisons H sont indiquées en rouge (PDB : 5ITX).

Plus généralement, comme nous l'avons vu, il nous a été impossible d'atteindre 100 % de formation de l'adduit sur la protéine, et ce, quels que soit les inhibiteurs ou le temps d'incubation. Or, avec un tel excès d'inhibiteur par rapport à la protéine, en condition de tests d'activité, aucune activité ne serait détectée. Le

composé a donc « attaqué » l'ensemble des protéines présentes en solution. Cependant, dans ces expériences de spectrométrie de masse, ce n'est pas le cas. De plus le temps d'incubation (de quelques heures jusqu'à plusieurs jours) est également bien plus grand que pour les tests d'activité (15 minutes). Il en résulte, que la formation de cet adduit covalent n'est, sans aucun doute, pas le premier responsable de l'inhibition de hNEIL1 par ces inhibiteurs. Néanmoins, il n'est pas impossible que les cystéines pontées se localisent dans les sites d'interaction non-covalents des inhibiteurs associés à l'inhibition. Cette conclusion est renforcée par la faible formation d'un adduit avec B9-P2 comparé à D5-P2, qui n'en forme aucun, ce qui est inversement proportionnel à leur capacité d'inhibition. Pour ces 2 inhibiteurs, on peut imaginer que le groupement thiol de la cystéine n'est pas positionné idéalement par rapport à la fonction nitrile de l'inhibiteur, ou que leur site d'interaction se situe loin des cystéines évoquées. Dans ce cas, la formation d'un adduit avec les cystéines pourrait plutôt limiter l'effet inhibiteur.

Pour conclure, il est fort probable que cet adduit se forme de façon sélective entre la fonction nitrile de ces différents inhibiteurs et la fonction thiol de cystéines particulières de la protéine. Il n'est cependant pas le premier responsable de l'inhibition de l'enzyme car certains de ces inhibiteurs sont capables d'inhiber mvNEI1, une enzyme structuralement très proche de hNEIL1 mais dépourvue de cystéines. Les sites contenant les cystéines modifiées sont d'ailleurs structuralement proches, malgré des acides aminés différents.

III.3. Effet des inhibiteurs sur la stabilité thermique de hNEIL1

Nous avons évalué l'effet de nos inhibiteurs sur la température de fusion (T_m) de hNEIL1 par la méthode TSA (voir le paragraphe Matériel & Méthode dédié) (Figure 97). Ce type d'expériences peut donner des indices sur le mode d'action des inhibiteurs, ou du moins peut permettre de séparer nos composés en différentes « classes » d'inhibiteurs : ceux stabilisant la protéine, ceux qui au contraire la déstabilise et ceux qui n'ont pas d'impact significatif. Les composés « non déterminés » sont ceux dont la température n'a pu être évaluée à cause d'une courbe de fusion incohérente, potentiellement par une trop forte déstabilisation de hNEIL1. Nous avons testé tous les inhibiteurs confirmés de hNEIL1, y compris les inhibiteurs que nous n'avons pas retenus suite au test d'interaction à l'ADN, et nous avons calculé le **ΔT_m**, soit la différence des **T_m observée entre le contrôle sans inhibiteur et celui avec**. La concentration des inhibiteurs est de 50 µM pour 2 µM de hNEIL1. Nous pouvons tout d'abord remarquer que très peu d'inhibiteurs stabilisent réellement l'enzyme. Cependant, les inhibiteurs les plus efficaces, que ce soit la girgensonine ou ces analogues, ICOA-D5-P2, ICOA1-B9-P2 ou LA-32, ne semblent pas déstabiliser fortement la structure de la protéine. Si une stabilisation de la protéine aurait pu être un signe encourageant pour une étude du complexe hNEIL1/inhibiteur par cristallographie, un effet neutre a au moins le mérite de ne pas être négatif pour la stabilité de la protéine.

ΔTm (°C)	1	2	3	4	5	6	7	8	9	10	11	12
A	Prest-C6-P6 x	Prest-F6-P8 -4,13 ± 0,79	Prest-G9-P5 -0,56 ± 0,29	Prest-B11-P10 -0,21 ± 0,12	Prest-E6-P2 x	Prest-C8-P13 -0,42 ± 0,08	Prest-A5-P13 -0,30 ± 0,01	Prest-G6-P14 x	ICOA1-D5-P2 -0,625 ± 0,29	ICOA1-B9-P2 -0,30 ± 0,01	LA-32 -0,33 ± 0,13	
B	GP-H2-P1 x	GP-H3-P1 0,08 ± 0,20	GP-C7-P6 -2,18 ± 0,08	Juglone -0,27 ± 0,04	Purpurine -0,39 ± 0,13	Amb22227278 -0,30 ± 0,08	Amb5756552 -0,66 ± 0,25					
C	L-03E05 -0,53 ± 0,16	L-03F11 -0,30 ± 0,01	L-03E07 -0,48 ± 0,09	L-03F08 -0,42 ± 0,17	L-03F07 -0,52 ± 0,14	L-03G09 -0,39 ± 0,21	Girgensonine 0,41 ± 0,16	Contrôle hNEIL1				

Effet observé	
	Effet stabilisant (ΔTm > 1 °C)
	Peu d'effet (-1 °C > ΔTm > 1 °C)
	Effet déstabilisant (ΔTm < -1 °C)
	Non déterminé

Figure 97 : Variation de la température de fusion (T_m) de hNEIL1 (ΔT_m) en présence de différents inhibiteurs issus du criblage ou des analogues

III.4. Cinétique d'inhibition de hNEIL1 par la girgensonine

Afin de mieux comprendre le mécanisme d'inhibition de nos composés, nous avons mis en place des études de cinétiques enzymatiques avec hNEIL1. Malheureusement en condition de « multi turnover », hNEIL1 atteint une vitesse constante très rapidement, ce qui complique les études cinétiques. Même en utilisant un lecteur ClarioStar, permettant d'évaluer la fluorescence produite par l'activité de hNEIL1 sur un ADN substrat fluorescent, avec un point toutes les 0,25 secondes, nous n'avons pas pu obtenir de résultats interprétables (Figure 98). Les vitesses initiales ne semblent pas suivre une courbe décrite par l'équation de Michaelis-Menten (Equation 4). Il est intéressant de noter que les courbes indiquées utilisent une unité arbitraire de fluorescence et non une quantité de produit comme c'est usuellement le cas. En effet, pour convertir le signal de fluorescence obtenu en quantité d'ADN clivé, donc en quantité de produit, nous avons effectué une courbe de dosage de l'ADN simple brin contenant le fluorophore indiqué en Figure 98 C. Cependant, si la courbe de dosage est bien linéaire comme attendue, les valeurs obtenues ne correspondent pas à celles obtenues en cinétique, et ne permettent donc pas de calculer la quantité de produit. Ceci peut s'expliquer par une différence de fluorescence entre le simple brin seul et le double brin clivé, avec un potentiel « quenching » résiduel du brin contenant le « blackhole quencher », en solution. Cependant, malgré ce souci d'unité, l'allure des courbes ne changerait pas. Nous n'avons donc pas pu déterminer une valeur cohérente précise de V_{max} ni de K_M.

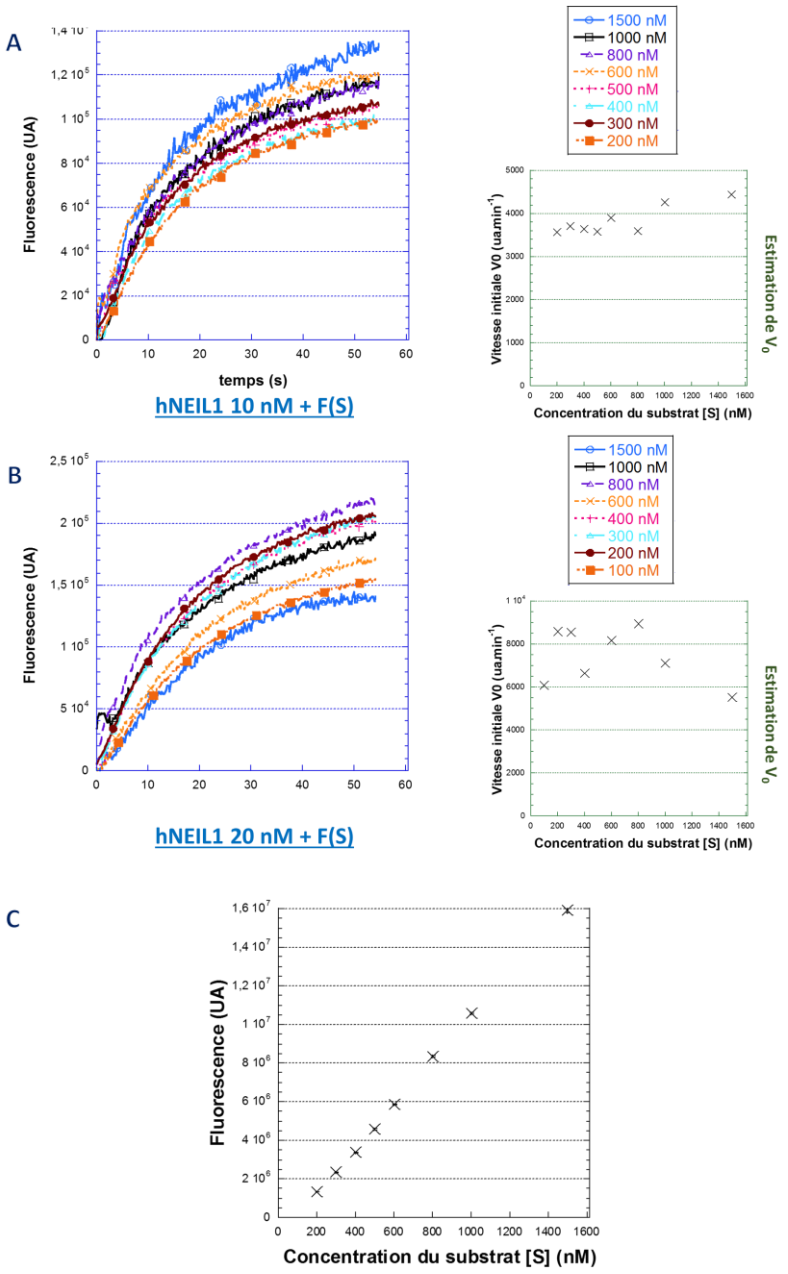


Figure 98 : Cinétiques obtenus avec hNEIL1 et le lecteur ClarioStar - A) et B) représentent le tracé de la formation du produit en fonction du temps pour deux concentrations de hNEIL1 respectivement 10 et 20 nM. C) Analyse du signal fluorescent en fonction de concentrations croissantes de l'ADN simple brin substrat (non-quenché).

Il semble donc que hNEIL1 ne soit tout simplement pas une enzyme Michaelienne. Le fait que ce soit une enzyme bifonctionnelle pourrait expliquer cet aspect singulier de hNEIL1, avec une activité glycosylase et AP lyase efficace, mais suffisamment découplée pour perturber les cinétiques. De plus dans l'ensemble de la littérature, aucune valeur de constante cinétique n'a été proposée pour cette enzyme, ce qui témoigne probablement de la difficulté d'interpréter les cinétiques enzymatiques de hNEIL1. Tous ces éléments ne nous permettent donc pas d'étudier le mode d'inhibition des inhibiteurs de hNEIL1 que nous avons identifiés.

IV. Etude structurale de la protéine hNEIL1 en interaction avec les inhibiteurs

Suite à l'échec des études de cinétique, nous n'avons pu conclure sur le mode d'action de nos inhibiteurs (compétitif, non compétitif, incompétitif). Nous avons donc mis en place des études

cristallographiques pour obtenir la structure tridimensionnelle de hNEIL1 en présence d'inhibiteur afin de comprendre le site et le mode d'interaction de nos composés avec l'enzyme.

IV.1. Cristallographie de la protéine seule

IV.1.1.« Soaking » des inhibiteurs avec les cristaux de hNEIL1

- **Cristallogenèse**

Tout d'abord, nous avons cherché à obtenir des cristaux de la protéine seule. En effet, l'objectif est d'observer le ligand dans la structure tridimensionnelle de la protéine. Pour cela, avoir obtenu le cristal de la protéine seule peut permettre, par un procédé dit de trempage (ou « soaking » en anglais), de faire diffuser ce ligand dans le cristal déjà formé et donc d'obtenir une structure de la protéine en interaction avec le ligand. Cette technique est la méthode la plus simple et la plus directe pour observer un ligand dans une structure cristallographique d'une protéine. Nous avons décidé d'utiliser une forme tronquée de hNEIL1, en retirant 95 acides aminés du domaine C-terminal. Ces 95 résidus forment une chaîne déstructurée et sont donc supposés déstabiliser la formation de complexe cristallin. La séquence de cette forme tronquée, notée hNEIL1dc95, est indiquée en Annexe 4. Nous avons ensuite produit et purifié hNEIL1dc95 en suivant le protocole établi dans le chapitre Matériel et Méthodes. La protéine purifiée a ensuite été concentrée jusqu'à 5 mg/mL dans un tampon final (20 mM Hepes pH 7,6, 100 mM NaCl, 1 mM TCEP), et des essais de cristallisation ont ensuite été réalisés avec les kits JCSG+, Morpheus de Molecular Dimensions, Wizard 1+2 de Rigaku Reagents et HR1+2 de Hampton Research (Annexe 8), soit **384 conditions au total**. Les gouttes ont été préparées par le robot Mosquito en goutte assise et en mélangeant 100 nL de la protéine purifiée et 100 nL de solutions du kit. Paramètre capital au vu de la « fragilité » de hNEIL1, la température des plaques a été maintenue à 4 °C. Suite à ce premier criblage, et à de nombreuses étapes d'optimisation, nous avons obtenu des cristaux avec la condition de solution de cristallisation **200 mM citrate de tri-potassium, 20 % PEG 3350,** avec un ratio de goutte de **2 : 1 (protéine : solution du kit).**

Nous avons donc effectué le « soaking » sur ces cristaux en déposant, sur chaque goutte contenant un cristal, nos différents inhibiteurs sous forme de poudre. Le principe étant qu'une partie de la poudre se dissolvant dans la goutte, les inhibiteurs en solution peuvent rentrer dans le cristal par des canaux remplis de solvant et se lier à la protéine. Nous avons effectué ces étapes de « soaking » avec l'inhibiteur de la littérature P2 ainsi qu'avec deux inhibiteurs identifiés lors du criblage, LA-32 et la girgensonine. Après une semaine d'incubation, un cristal unique est récupéré, transféré dans une solution contenant le tampon de cristallisation et un cryoprotectant (3 µL de la solution du puits + 1 µL d'éthylène glycol), puis les cristaux sont plongés dans de l'azote liquide. Les cristaux ainsi figés sont ensuite envoyés au synchrotron pour la collection des données de diffraction X.

Il est intéressant de noter que les conditions de cristallisation obtenues sont différentes de celles indiquées dans la littérature (35 % PEG 3350, 0,25 M Sulfate de lithium pH 8,5 [404] ou 0,1 M acide cacodylique, 0,1 M NaCl, 0,05 M $MgCl_2$, 24 % PEG 8000 [399]). Ces différences peuvent venir de la légère différence entre les séquences de hNEIL1 utilisées. Nous avons tronqué 95 résidus en C-terminal, tandis que les structures obtenues dans les conditions précédentes sont tronquées de 100 résidus.

- **Données de diffraction**

Les données de diffraction pour la protéine seule ont été collectées jusqu'à 1,46 Å de résolution au synchrotron SOLEIL sur la ligne PROXIMA-2. Ces données nous ont permis de résoudre la structure de hNEIL1 par remplacement moléculaire en utilisant une structure de la protéine disponible dans la PDB (code 5ITQ). Le modèle a été affiné jusqu'à des facteurs R de 16,3 % et R_{free} de 18,2 % dans la gamme de résolution 32,92-

1,46 Å (Figure 99). Si les conditions de cristallisation sont différentes, nous obtenons le même groupe d'espace H3 que les structures présentes dans la PDB.

Statistiques de la collecte de données			
	hNeil1	hNeil1 + girgensonine (soaking)	hNeil1 + P2 (soaking)
Source de rayonnement	SOLEIL PROXIMA-2	SOLEIL PROXIMA-2	SOLEIL PROXIMA-2
Longueur d'onde (Å)	0.98010	0.98011	0.98011
Groupe d'espace	*H*3	*H*3	*H*3
Dimension de la maille: *a, b, c* (Å)	132.00, 132.00, 50.81	132.79, 132.79, 50.65	132.72, 132.72, 51.16
α, β, γ (°)	90.00, 90.00, 120.00	90.00, 90.00, 120.00	90.00, 90.00, 120.00
Gamme de résolution (Å)	32.92-1.46 (1.54-1.46)	66.40-1.72 (1.81-1.72)	66.36-1.49 (1.57-1.49)
Observation totales	578106 (87434)	354648 (54506)	542843 (80424)
Réflexions uniques	57240 (8369)	35442 (5200)	54641 (8014)
Complétude (%)	100.0 (100.0)	100.0 (99.9)	99.6 (100.0)
Multiplicité	10.1 (10.4)	10.0 (10.5)	9.9 (10.0)
R_{merge} [a] (%)	4.6 (122.5)	8.1 (65.3)	6.4 (55.5)
I/σ moyen *(I)*	18.9 (1.5)	14.6 (2.6)	16.2 (2.7)
$CC_{1/2}$ (%)	100.0 (75.0)	99.9 (91.0)	99.9 (91.7)
Affinement et statistiques du modèle			
Gamme de résolution (Å)	32.92-1.46	-	-
Nombre de réflections utilisées	57235	-	-
R_{work} [b] / R_{free} [c] (%)	16.3/18.2	-	-
Facteur B moyen (Å^2)			
Tous les atomes	34.54		
Atomes de la protéine	33.58		
Atome de citrate	63.65		
Atomes du éthane-1,2-diol	48.41		
Atomes de potassium	35.54		
Atome H_2O	43.00		
RMSD par rapport à l'idéal			
Longueur des liaisons (Å)	0.005		
Angles des liaisons (°)	0.830		
Analyse ramachandran (% de résidus)			
Régions favorables / autorisées / interdites	97.9/2.1/0.0		
Nombre d'atomes			
Protéine	2258		
Citrate	13		
Ethane-1,2-diol	20		
Potassium	1		
Eau	257		

[a] $R_{\text{merge}} = \sum_h \sum_i |I_{h,i} - \langle I \rangle_h| / \sum_h \sum_i I_{h,i}$ où $\langle I \rangle_h$ représente l'intensité moyenne des réflexions symétriques équivalentes.
[b] $R_{\text{work}} = \sum_h ||F_o| - |F_c|| / \sum_h |F_o|$, où F_o et F_c sont les amplitudes des facteurs de structures observés et calculés respectivement pour les réflexions h.
[c] R_{free} est la valeur R extradite de 5% des données de réflexion, qui ne sont pas inclues dans le processus d'affinement.

Figure 99 : Diffraction et statistiques d'affinement de hNEIL1 lors du « soaking »

Concernant les structures après trempage avec des inhibiteurs, la girgensonine n'est pas visible dans la structure obtenue qui est superposable aux structures de la protéine libre présentes dans la PDB (voir le RMSD sur Figure 99). L'absence de la girgensonine dans le site actif de hNEIL1 pourrait être liée aux conditions de cristallisation. En effet, une molécule de citrate est observée précisément dans le site actif, formant des liaisons hydrogène avec les résidus Arg133, Arg78, Gly80 et la proline catalytique Pro2 (Figure 100). La concentration du citrate dans la solution de cristallisation étant de 200 mM, il est potentiellement en large excès face à la concentration de girgensonine en solution dans l'essai et donc en compétition déloyale avec la girgensonine. Suite à ces résultats, nous avons évalué l'activité de hNEIL1 en présence de citrate, mais

aucun effet inhibiteur n'a été observé. Le trempage des cristaux de hNEIL1 en présence des inhibiteurs P2, B9-P2 et D5-P2 n'ont pas non plus permis d'observer les complexes hNEIL1/inhibiteur correspondant.

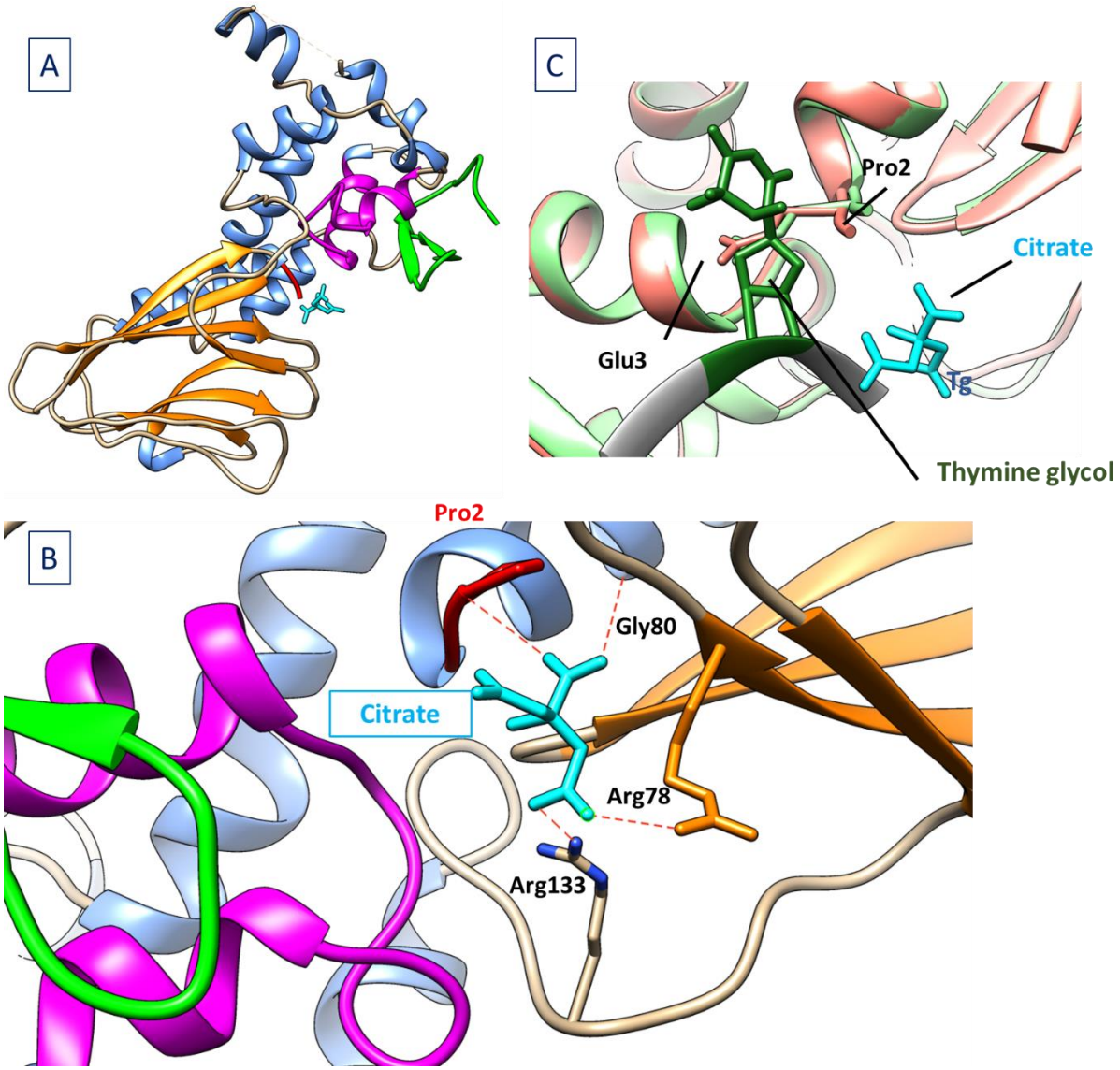


Figure 100 : Structure de hNEIL1 obtenue à 1,43 Å de résolution - A) Structure de hNEIL1 obtenue. Les hélices α sont colorées en bleu, les brins β en orange, le domaine H2TH en magenta et le ZnLF en vert clair. L'anion citrate présent dans le site actif est indiqué en cyan, tandis que les liaisons hydrogène sont indiquées en rouge B) Site actif contenant le citrate C) Superposition de la structure obtenue lors de cette thèse (en saumon) présentant un citrate dans son site actif (en cyan) et hNEIL1 complexée à un ADN (en vert pale, PDB 5ITX) possédant une Tg (en vert foncé).

IV.1.2.Cristallogenèse de la protéine seule en présence de DMSO

- **Cristallogenèse**

Par la suite, nous avons cherché à évaluer si des cristaux de la protéine pouvaient être obtenus en présence d'inhibiteurs, en « co-cristallisation ». En effet, lors du trempage, les interactions entre protéines peuvent masquer les zones d'interaction protéine/inhibiteur et l'absence de DMSO limite la concentration en solution de l'inhibiteur. Nous avons d'abord testé manuellement les conditions précédentes (5 mg/mL de hNEIL1dc95 et 200 mM citrate de tri-potassium), en faisant varier différents paramètres comme l'agent précipitant (16 %, 18 %, 20 % PEG 3350), le volume et le ratio des gouttes utilisées, et le pourcentage de DMSO dans la goutte (1 %, 2 %, 4 %, 8 %). Au total, **52 conditions** ont été testées, toujours à 4 °C, dont 12 permettent la formation de cristaux. Les résultats obtenus ont permis d'établir différentes conditions :

i) la formation des cristaux est favorisée avec une concentration d'agent précipitant plus faible (16 %). Ce résultat semble logique, car l'ajout de DMSO peut aussi agir en tant qu'agent précipitant.

ii) Une concentration trop élevée de DMSO (4 %, 8 %) empêche la cristallisation et induit la précipitation de la protéine.

iii) Un ratio de 1 µL de la protéine + 1 µL de la solution-réservoir est le ratio donnant les cristaux les mieux formés.

La condition retenue est donc **200 mM citrate de tri-potassium, 16 % PEG 3350, avec une goute 1+1 et 2 % de DMSO**, condition très proche de celle obtenue pour la protéine apo. Cela peut indiquer qui indique que le DMSO, à cette concentration, ne perturbe pas la formation du même complexe cristallin que celui obtenu précédemment. Cependant cela ne permet pas de s'affranchir du citrate. Nous avons ensuite incubé la protéine avec 400 µM d'inhibiteurs pour 2 % de DMSO (LA-32, girgensonine, B9-2, D5-2 et P2). Des cristaux se sont formés après une semaine à 4 °C. Ces cristaux ont été plongés dans une solution cryoprotectante contenant de l'éthylène glycol, puis congelés dans l'azote liquide et envoyés au synchrotron.

- **Données de diffraction**

Les données de diffraction ont été collectées au synchrotron SOLEIL sur la ligne PROXIMA-1 et analysées par Franck COSTE. J'ai suivi ces analyses sous sa supervision. Ces données nous ont permis de résoudre la structure de hNEIL1 par remplacement moléculaire en utilisant la structure obtenue ci-dessus. Malheureusement, les structures obtenues étaient identiques à celles du « soaking », et aucun inhibiteur n'a pu être observé. L'affinement n'a donc pas été terminé pour la co-cristallisation (Figure 101). Ici encore, il est possible que le citrate entre en compétition avec la girgensonine, empêchant l'inhibiteur d'interagir avec hNEIL1 dans le cristal.

Statistiques de la collecte de données		
	hNeil1 + girgensonine (co-cristallisation)	hNeil1 + P2 (co-cristallisation)
Source de rayonnement	SOLEIL PROXIMA-1	SOLEIL PROXIMA-1
Longueur d'onde (Å)	0.97856	0.97856
Groupe d'espace	*H*3	*H*3
Dimension de la maille: *a, b, c* (Å)	133.00, 133.00, 50.88	133.03, 133.03, 50.97
α, β, γ (°)	90.00, 90.00, 120.00	90.00, 90.00, 120.00
Gamme de résolution (Å)	46.55-1.74 (1.84-1.74)	38.40-1.76 (1.86-1.76)
Observation totales	309324 (47929)	297918 (49981)
Réflexions uniques	30412 (5004)	30741 (4839)
Complétude (%)	88.5 (100.0)	92.8 (100.0)
Multiplicité	10.2 (9.6)	9.7 (10.3)
R_{merge}[a] (%)	8.7 (51.1)	10.3 (58.9)
I/σ moyen *(I)*	13.6 (2.7)	11.7 (2.4)
$CC_{1/2}$ (%)	99.8 (93.5)	99.7 (93.5)

[a] $R_{merge} = \sum_h \sum_i |I_{h,i} - \langle I \rangle_h| / \sum_h \sum_i I_{h,i}$ où $\langle I \rangle_h$ représente l'intensité moyenne des réflexions symétriques équivalentes.

Figure 101 : Diffraction de hNEIL1 en co-cristallisation avec les inhibiteurs

Afin de s'affranchir du citrate, nous avons lancé de nouveaux criblages cristallographiques de co-cristallisation, dans l'espoir de trouver des conditions ne comportant pas cette molécule. Pour ce faire, nous avons utilisé deux stocks de hNEIL1dc95 à 5 mg/mL que nous avons incubé avec 100 µM de P2 et de girgensonine respectivement, pendant 20 minutes à 4 °C, avec les kits HR1+2 de Hampton Research, Wizard 1+2 de Rigaku Reagent et Morpheus de Molecular Dimension (Annexe 8), soit **288 conditions au total pour chaque inhibiteur**. Nous avons pu observer des microcristaux dans 1 % des conditions pour 100 µM de

girgensonine et dans 0,7 % pour 100 µM de P2, mais les essais d'optimisation réalisés ensuite n'ont pas permis d'obtenir des cristaux ayant un pouvoir diffractant suffisant pour une étude de diffraction des rayons X complète.

IV.1.3.Essai avec une autre forme tronquée de hNEIL1

- **Cristallogenèse**

Nous avons également utilisé une autre forme tronquée de hNEIL1 utilisée par Bandaru *et al.*, hNEIL1dc56, dont la séquence est tronquée de 56 acides aminés en C-terminal [379,469]. En effet, tout comme pour dc95, les alignements de séquence effectués avec différents membres de la superfamille Fpg/Nei tel que AthFpg (*Arabidopsis thaliana*) suggèrent que la délétion de ces 56 acides aminés impacte peu l'activité glycosylase. De plus, une structure a été obtenue par cette même équipe et est disponible dans la littérature. hNEIL1dc56 y est décrite comme beaucoup plus soluble, ce qui est un avantage conséquent pour les cribles cristallographiques. En effet si hNEIL1dc95 conserve effectivement bien l'activité glycosylase, elle n'est pas très soluble. De très nombreuses purifications ont dû être effectuées afin de disposer de la quantité nécessaire de protéine pour un criblage cristallographique. La protéine a ainsi tendance à précipiter rapidement et ne peut être conservée sur plusieurs jours sans être congelée à -80 °C. De plus, les rendements de purification étaient assez faibles, avec de fortes pertes au cours des différentes étapes. **Concrètement, une nouvelle purification a été nécessaire pour chaque criblage cristallographique (environ 2-4 plaques de 96 conditions)**, ainsi que pour les optimisations manuelles.

Nous avons donc purifié cette forme tronquée de hNEIL1, en suivant le protocole décrit dans le chapitre dédié du chapitre Matériel et méthodes et nous l'avons incubé à 5 mg/mL avec **2 % DMSO/20 mM girgensonine** final pendant 20 minutes à 4 °C. Ce lot a été testé avec les kits Q_Classic, JCSG+ et Morpheus de Molecular Dimension (Annexe 8), soit **288 conditions au total**. Après optimisation des résultats obtenus (certains cristaux sont indiqués en Annexe 9), pour 100 nouvelles conditions, les meilleurs cristaux ont été obtenus pour **0,2 M malonate de tri-potassium pH 7, 20 % PEG 3350**. Il est à noter que nous n'avons pas réussi à reproduire les conditions de cristallisation obtenue par Bandaru *et al.* [469].

- **Données de diffraction**

Les données de diffraction ont été collectées au synchrotron SOLEIL sur la ligne PROXIMA-1. Les cristaux obtenus avec hNEIL1dc56, en présence de 20 mM de girgensonine ont suffisamment diffractés pour fournir un modèle structural à 3,4 Å de résolution. Malheureusement, aucune densité électronique pouvant correspondre à la girgensonine n'était visible. Comme seule la protéine était visible dans ce modèle cristallographique, identique à celui proposé dans la PDB, nous n'avons pas cherché à affiner cette structure.

V. Bilan et conclusion

L'objectif de cette partie de mon travail de thèse était d'identifier des inhibiteurs efficaces et sélectifs de l'enzyme humaine hNEIL1. Pour ce faire, nous avons décidé de mettre en place un criblage moyen débit sur différentes chimiothèques, choisies pour leur originalité, leur diversité structurale et leur complémentarité, pour un total de **4560 composés**. Avant d'effectuer ce criblage sur les enzymes humaines, il nous a fallu mettre en place une procédure pilote de criblage sur l'enzyme bactérienne LlFpg. En effet, cette enzyme est un homologue structural de hNEIL1, ce qui en fait un modèle pertinent. De plus, nous disposions également d'inhibiteurs contrôles sur cette enzyme. Ainsi, après de nombreuses optimisations et contraintes méthodologiques, nous avons mis en place un criblage moyen débit robuste et pertinent sur les ADN glycosylases humaines. Cette preuve de concept nous a permis d'identifier des molécules intercalantes et

inhibitrices de LlFpg, hNEIL1 et hOGG1 : la **Juglone** et la **Purpurine**, qui ont été utilisées comme contrôle lors des criblages sur les ADN glycosylases humaines.

Le criblage moléculaire de hNEIL1 a tout d'abord permis d'identifier **50 hits**. Après des étapes de confirmation de l'effet inhibiteur, des étapes de validation et de sélection, 3 composés non intercalant ont finalement été identifiés : la **girgensonine** et les composés **ICOA1-B9-P2** et **ICOA1-D5-P2**. Sur la base des données de la littérature scientifique et sauf omission de notre part, la girgensonine est à ce jour l'inhibiteur connu le plus efficace de hNEIL1. Suite à l'identification de ces composés, une « hit expansion » a été mise en place, en testant des analogues de ces composés. Si aucun composé plus efficace n'a été identifié, nous avons pu identifier quelques fonctions chimiques essentielles à l'inhibition, notamment la **fonction aminonitrile**. Cette observation a permis l'identification d'un quatrième composé, **LA-32**. Lors des études de l'activité des analogues de la girgensonine, nous avons pu également constater que la perte d'un groupement donneur de proton diminue l'efficacité des inhibiteurs. Il est possible que cette perte puisse impacter une potentielle interaction avec le résidu catalytique Glu2. Un encombrement stérique trop important autour du nitrile annule également l'inhibition, ce qui pourrait suggérer que l'inhibiteur se place dans une poche restreinte.

La spécificité de ces inhibiteurs a ensuite été évaluée, avec LA-32 inhibant aussi bien LlFpg, hNEIL1 et mvNei1, tandis que la girgensonine inhibe de manière préférentielle l'enzyme humaine. Le choix des enzymes mvNei1 et LlFpg a été dirigé par le fait qu'elles sont très proches structuralement et par la présence ou non de cystéines libres dans leur séquence. Nous avons également tenté d'établir le mode d'action de ces inhibiteurs sur la cinétique enzymatique de hNEIL1. Malgré de nombreuses tentatives, nous avons conclu que **hNEIL1 n'était pas une enzyme michaélienne** et donc nous ne pouvons pas garantir que ces inhibiteurs agissent directement dans le site actif (inhibition compétitive) ou indirectement en dehors du site actif (inhibition non-compétitive ou incompétitive).

En parallèle, nous avons mis en évidence par spectrométrie de masse la capacité de certains inhibiteurs à former des adduits covalents avec hNEIL1. La formation de **cet adduit** semble directement reliée à la présence d'une fonction nitrile dans certains inhibiteurs, et semble viser certaines cystéines de hNEIL1. En revanche il est peu probable que cet adduit soit responsable de l'inhibition de hNEIL1 par la girgensonine. En réalité, l'hypothèse que nous proposons, est que la formation d'adduits entre la girgensonine et hNEIL1 est une réaction secondaire. Il est cependant possible que cet adduit provienne d'une fixation primaire et sélective de l'inhibiteur au voisinage de ces cystéines, principalement responsable de l'inhibition.

Enfin, nous avons tenté d'**identifier le site d'interaction** de certains inhibiteurs et de mieux comprendre leurs mécanismes d'action en mettant en place des études de cristallographie aux rayons X. Pour cela nous avons testé de nombreuses conditions, que ce soit avec des formes tronquées de hNEIL1, en présence de diverses concentrations d'inhibiteurs. Nous avons ainsi obtenu les **structures de hNEIL1 seule**. Malheureusement, la girgensonine ou les autres inhibiteurs, n'ont jamais été observés dans ces structures. Ce résultat est d'autant plus décevant que la girgensonine ne semble pas déstabiliser la protéine. L'une des explications potentielles est la présence de citrate dans le site actif, empêchant d'observer la girgensonine. Une autre explication serait la faible solubilité de nos composés en absence de DMSO.

Finalement, même si aucune structure n'a pu être obtenue malgré nos efforts, l'identification du composé inhibiteur de hNEIL1 le plus efficace reste un résultat satisfaisant et encourageant. D'autres expériences pourront être mises en place, notamment en évaluant l'effet de nos inhibiteurs sur les autres ADN glycosylases humaines hNEIL2 et hNEIL3. Des études *in cellulo* seront également capitales afin d'évaluer ou non si ces composés sont des molécules intéressantes à développer.

Partie 2 : Identification et caractérisation d'inhibiteurs de hOGG1

I. Criblage moléculaire sur hOGG1

I.1. Criblage primaire

I.1.1. Choix des conditions

Lors de criblage nous avons utilisé la forme tronquée, dhOGG1, qui a 10 acides aminés en moins en N-terminal, contenant la séquence d'adressage de hOGG1 à la mitochondrie, et 20 acides aminés en moins en C-terminal, contenant la séquence d'adressage au noyau. Ces résidus sont prédits non structurés et ne participent pas à l'activité catalytique de l'enzyme. La troncation de ces séquences n'a donc aucun effet sur l'activité de hOGG1. Les concentrations du criblage ont dû être déterminées, principalement en évaluant les ratios signal/bruit. Pour des raisons pratiques de préparation de plaques et pour les raisons évoquées précédemment, le même volume final de **15 µL** a été conservé. En revanche, le temps d'incubation à 37 °C a été étudié, ainsi que la concentration en ADN substrat et en protéine hOGG1 (Figure 102). Les meilleurs ratios signal/bruit ont été obtenus pour un temps d'incubation long, avec un ratio de protéine/ADN supérieur à celui du criblage de hNEIL1. Cela n'est pas surprenant, car hOGG1 possède une activité glycosylase-lyase bien plus médiocre que hNEIL1. La condition finale retenue (quadrillée en rouge sur la Figure 102) a été de **100 nM hOGG1 pour 50 nM d'ADN substrat**, avec un temps d'incubation de seulement **30 minutes**. En effet, nous avons sélectionné le ratio signal/bruit supérieur à 3 avec le temps d'incubation le plus bas. Un long temps d'incubation peut endommager la protéine, et surestimer l'inhibition potentielle des composés testés. Cette condition permet également de conserver une concentration homogène de protéine avec le criblage de hNEIL1, afin de pouvoir comparer les résultats d'inhibition plus facilement.

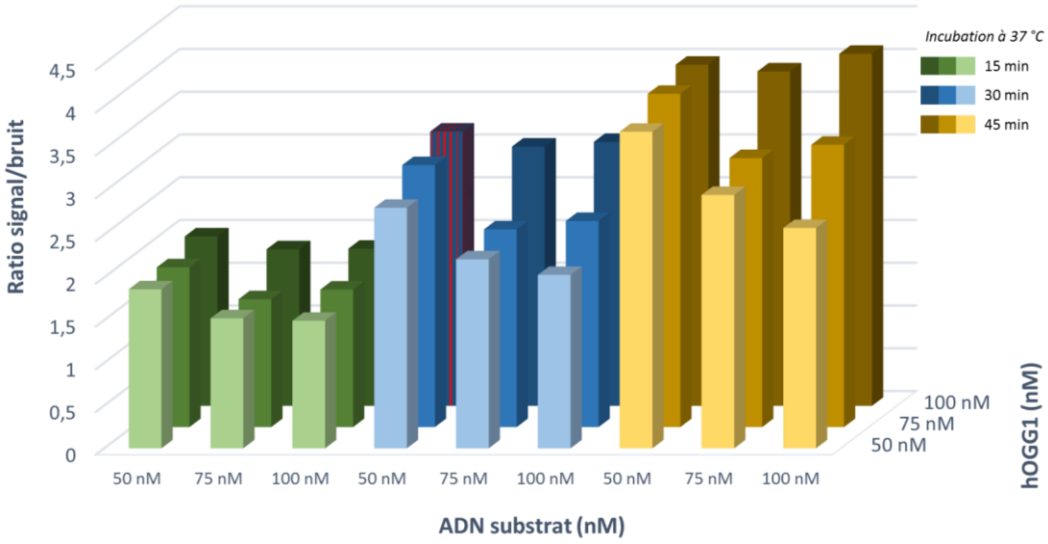


Figure 102 : Evaluation du ratio signal/bruit pour le criblage de hOGG1

I.1.2. Criblage primaire sur hOGG1

Sur l'ensemble du crible, un facteur moyen Z' de 0,80 ± 0,07 a été obtenu, attestant de la fiabilité des résultats. En utilisant le même « cutoff » de sélection, soit 70 % d'inhibition, nous avons identifié 70 molécules touches, soit 1,53 % des 4560 composés testés. La distribution des résultats est indiquée sur la Figure 103.

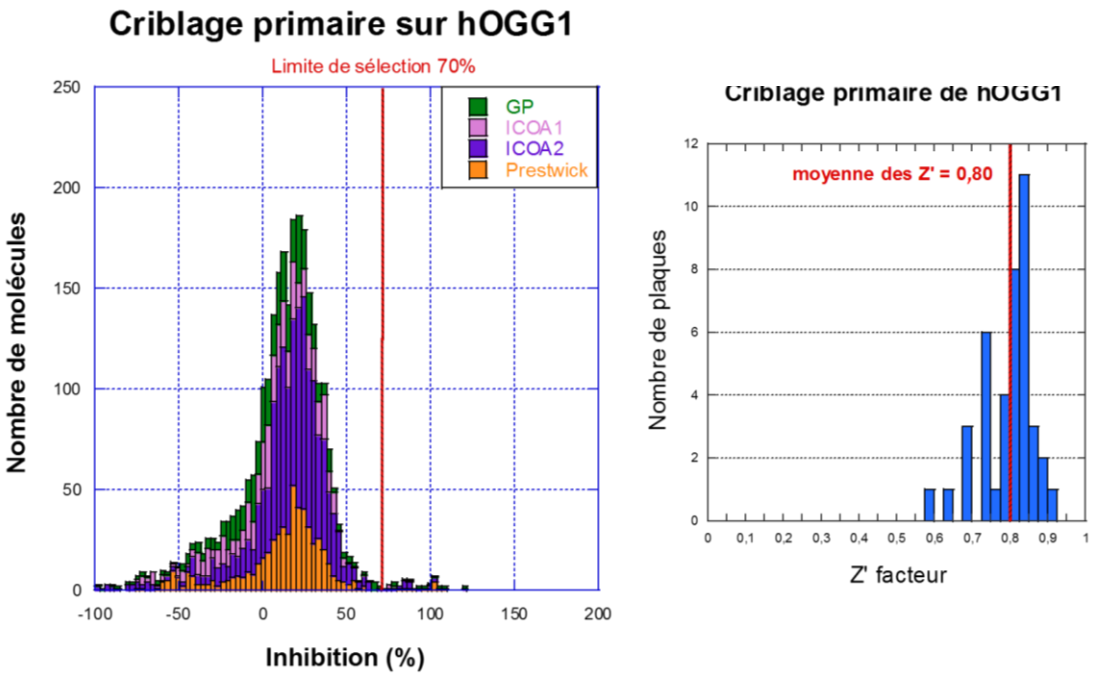


Figure 103 : Distribution des résultats du criblage et des facteurs Z' du criblage sur hOGG1

Il est intéressant de noter qu'un certain nombre de hits se recouvrent entre les deux criblages : 17 molécules sont des hits déjà obtenus avec hNEIL1 (Figure 104). Ces molécules peuvent avoir un spectre d'absorption proche du BHQ2 et donc agir comme des extincteurs de fluorescence d'où leur sélection lors du crible primaire (faux positifs). Ces molécules peuvent aussi interagir avec l'ADN et empêcher n'importe quelle protéine de se fixer à celui-ci, n'inhibant donc pas la protéine directement, mais séquestrant son substrat. Dans tous les cas, le test d'activité orthogonal va permettre de se séparer de la plupart de ces composés.

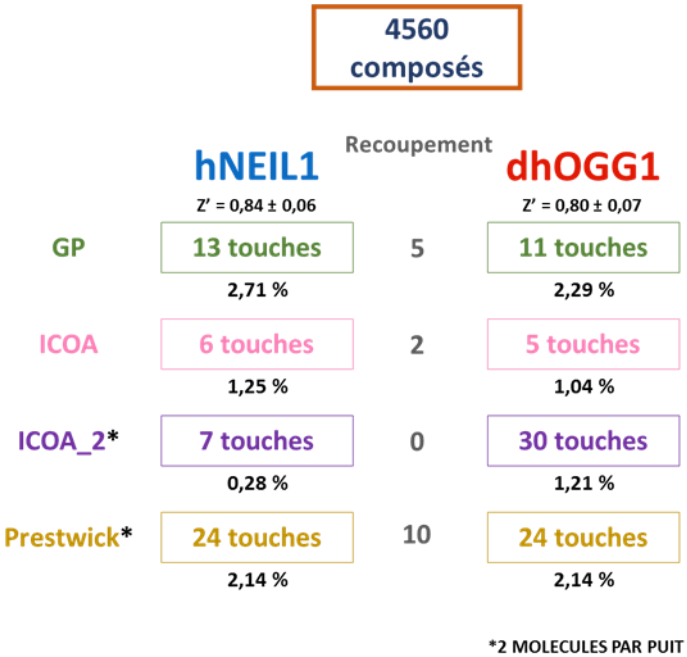


Figure 104 : hits obtenus lors des criblages primaires

I.2. Confirmation des hits sur hOGG1

Les 70 molécules identifiées comme hits sur hOGG1 ont donc été testées par le second test d'activité avec une concentration de 50 µM. L'activité glycosylase et glycosylase-lyase a été évaluée (Figure 105). On remarque que la plupart des inhibiteurs confirmés sont plus efficaces sur l'activité glycosylase-lyase que sur l'activité glycosylase. hOGG1 possédant une activité 8-oxoG glycosylase-lyase médiocre, on peut considérer que celle-ci sera plus « facilement » inhibée, bien que cela ne soit pas systématique. Cela peut s'expliquer par le fait qu'avec un substrat 8-oxoG comme c'est le cas dans ce test, un composé capable d'affecter l'activité 8-oxoG-ADN glycosylase affectera également l'AP lyase couplée (i.e., l'activité glycosylase/lyase). Cependant, les composés qui nous intéressent le plus sont les molécules inhibant en particulier l'activité 8-oxoG ADN glycosylase, l'activité spécifique de hOGG1. Un composé inhibant l'activité AP lyase pourrait être moins sélectif. Enfin, il est également important de noter que parmi les molécules testées en combinaison avec un autre composé (banque ICOA2 et Prestwick), seule une molécule par puits a été identifiée, comme on pouvait s'y attendre. Au total **24 composés** ont été confirmés comme inhibiteurs de hOGG1 par le test orthogonal s'affranchissant de l'utilisation de la fluorescence. La structure des composés confirmés est indiquée en Annexe 5.

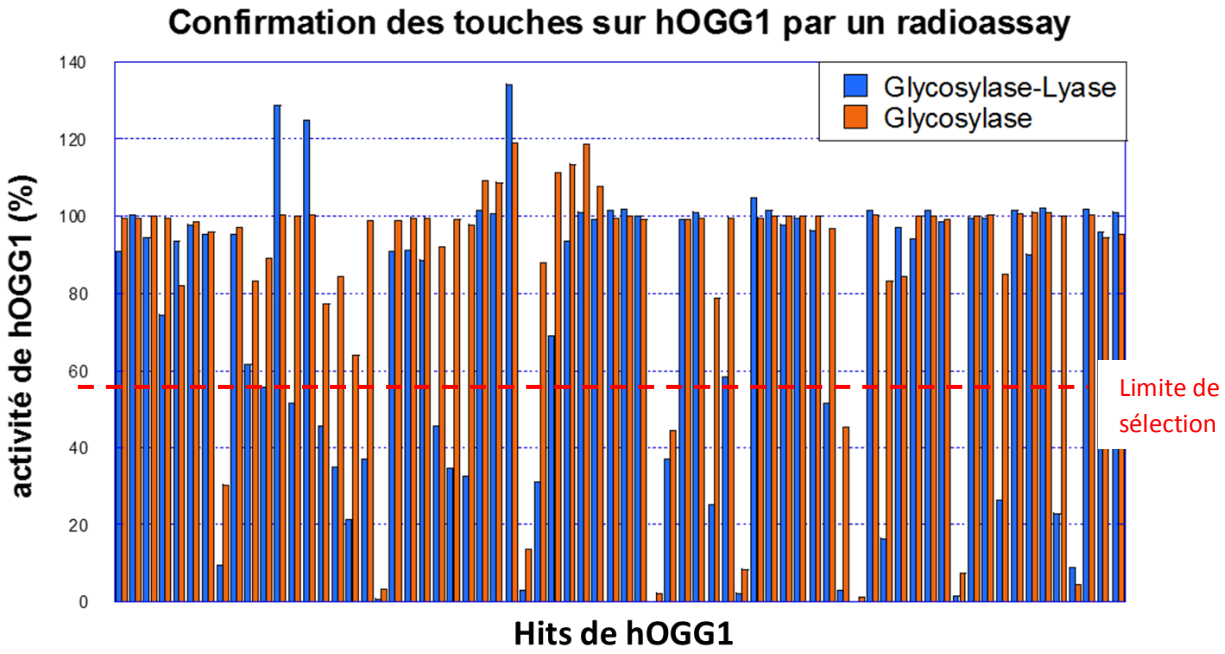


Figure 105 : Confirmation des 70 hits obtenus lors du criblage primaire de hOGG1 - L'activité glycosylase-lyase est indiquée en bleu clair, tandis que l'activité glycosylase est indiquée en orange.

Comme nous pouvions nous y attendre, le nombre de composés communs entre hOGG1 et hNEIL1 a également diminué lors de cette étape, avec seulement 7 molécules confirmées communes aux deux criblages parmi les 24 confirmées au premier test.

I.3. Elimination de composés intercalant

L'intercalation à l'ADN des 24 composés confirmés a été évaluée (Figure 106). Nous avons de nouveau séparé les composés « fortement intercalant » (avec une baisse de fluorescence supérieure à 60 %), « faiblement intercalant » (avec une baisse de fluorescence de 20 à 60 %) et les non intercalant. Au final, 15 inhibiteurs de hOGG1 se sont révélés non intercalant (Figure 107). Il est intéressant de noter que la plupart des composés inhibant à la fois hOGG1 et hNEIL1 sont des intercalants de l'ADN, à l'exception du composé

Prest-B11-P10 (l'Ebselen) et GP-C7-P6. Hors, comme nous l'avons établi pour le criblage de hNEIL1, ces deux composés ne sont pas intéressants pour la suite de cette thèse. Au final, c'est donc **13 inhibiteurs spécifiques de hOGG1** que nous avons identifiés. Parmi eux, nous n'avons pu obtenir la structure du composé ICOA1-H11-P6, pour des raisons de confidentialité de la part de l'ICOA.

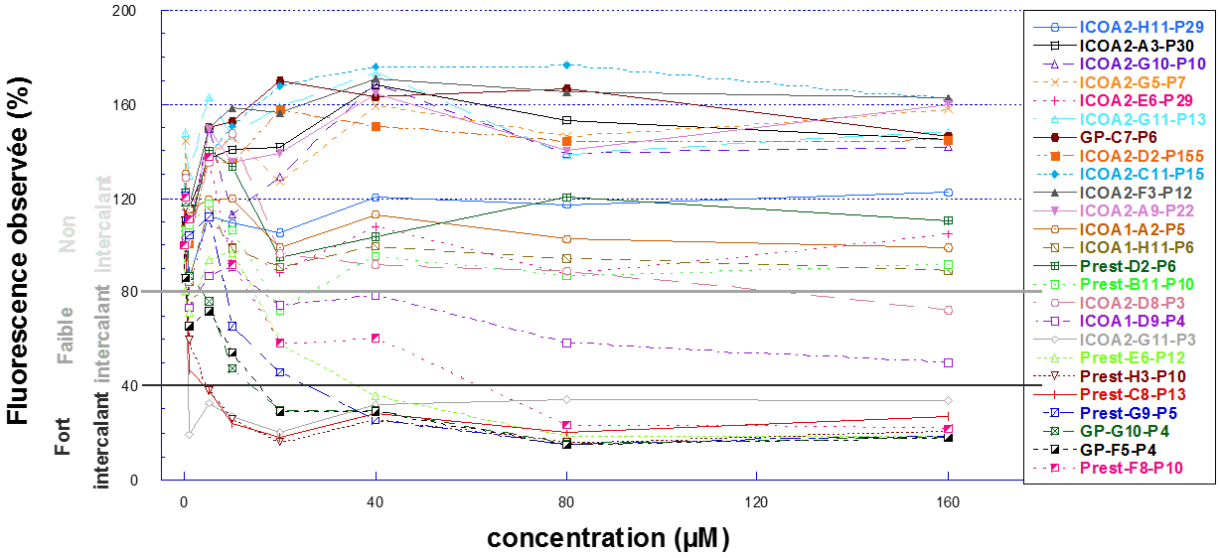


Figure 106 : Résultats du test d'intercalation PicoGreen sur hOGG1

Enfin, un des points remarquables de ces résultats est la présence de composés intercalant, mais sélectifs, n'inhibant qu'une seule des deux ADN glycosylases. En effet, il était attendu de voir les composés intercalant inhiber indifféremment toutes les ADN glycosylases. Or certains composés intercalant n'ont été confirmés que sur hOGG1, ou que sur hNEIL1. Il est possible que ces composés ne soient pas réellement « sélectifs ». S'ils n'ont pas été identifiés durant le crible, ou la confirmation, comme inhibiteur des deux enzymes, cela ne signifie pas pour autant qu'ils ne soient pas du tout inhibiteurs de cette enzyme. En effet, le crible ayant été fait à une concentration fixe, un effet d'inhibition pourrait être observé à plus forte concentration. Ensuite, malgré cette relative spécificité, comme nous l'avons déjà précisé, la plupart des composés intercalant sont toxiques pour les cellules, ce qui n'est pas en accord avec l'objectif de cette thèse. Malgré leurs sélectivités apparentes pour hOGG1, nous ne sommes donc pas allés plus loin dans la caractérisation de ces composés intercalant. Enfin nous pourrions malgré tout aussi imaginer que leur effet inhibiteur provienne d'une interaction à l'interface entre l'ADN et la protéine. Cette hypothèse est difficilement vérifiable sans structure associée, et comme nous le verrons dans les chapitres suivants, obtenir une structure avec un inhibiteur peut s'avérer long, ce qui induit donc de sélectionner un nombre de molécules restreint. Malgré l'intérêt fondamental que ces composés peuvent représenter, nous n'avons donc pas sélectionné ces composés pour des études plus approfondies, le temps de cette thèse ne nous le permettant pas.

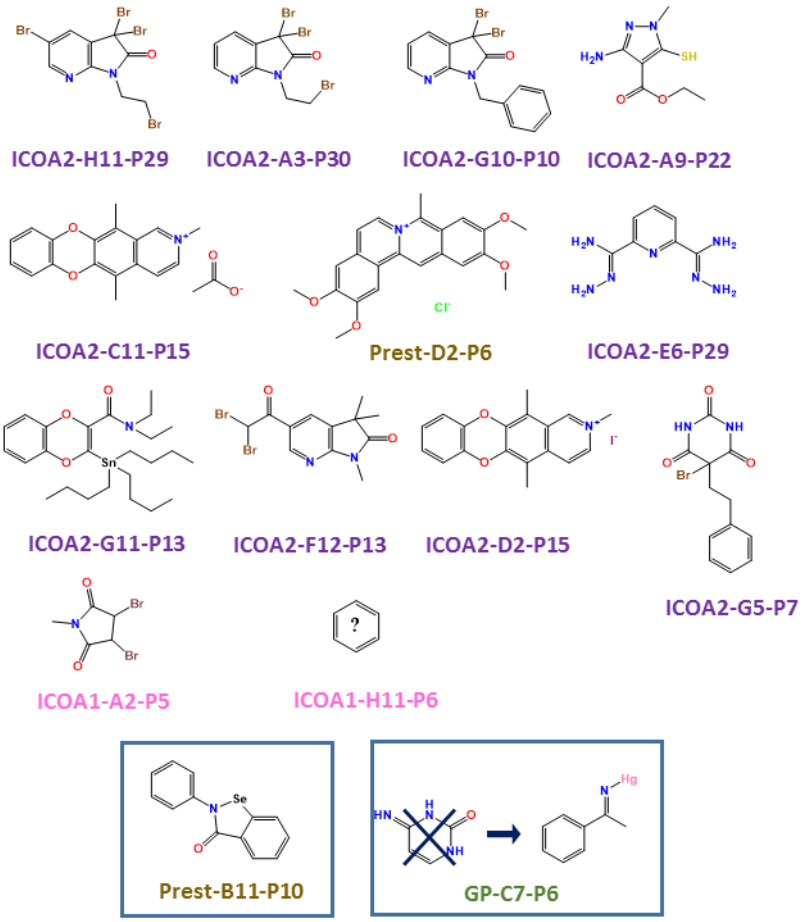


Figure 107 : Structure des composés non intercalant ayant été validés lors des différentes étapes du criblage comme inhibiteur de hOGG1 - Les composés ayant également une activité sur hNEIL1 ont été encadrés en bleu, et n'ont pas été retenus pour la suite de cette thèse.

I.4. Dose-réponse et sélection finale

L'objectif de cette étape a consisté à réduire le nombre d'inhibiteurs à étudier et à classifier les composés confirmés et sélectionnés restants en fonction de leur pouvoir inhibiteur. Tout d'abord nous avons éliminé les composés qui n'affichaient pas un IC_{50} estimé à moins de 50 µM. Ainsi les composés ICOA2-A9-22 / C11-P15 / D2-P15 / F12-P13, ICOA1-H11-6 / A2-P5 et Prest-D2-P6 n'ont pas été retenus pour les étapes suivantes.

	IC_{50} sur l'activité Glycosylase (µM)	
Composé	*Valeur publiée*	*Test "maison"*
ICOA2-A3-P30		0,455 ± 0,012
ICOA2-H11-P29		0,408 ± 0,008
ICOA2-G11-P13		14,101 ± 1,154
ICOA2-G5-P7		0,734 ± 0,037
ICOA2-E6-P29		2,152 ± 0,308
ICOA2-G10-P10		0,634 ± 0,012
TH5487	0,342	2,268 ± 0,660
O8	0,220 ± 0,080	0,674 ± 0,156
Su268	0,059	0,492 ± 0,105

Figure 108 : IC_{50} des inhibiteurs les plus efficaces du criblage sur hOGG1

Nous avons évalué uniquement l'IC_{50} des 6 composés restants (Figure 108). Les 3 principaux inhibiteurs publiés sont aussi affichés en comparaison. Sur les 6 composés, 3 sont très proches structuralement : **ICOA2-H11-P29**, **ICOA2-A3-P30** et **ICOA2-G10-P10**. Ces trois composés affichent aussi les IC_{50} les plus faibles du criblage entre 0,4 et 0,6 µM. Le composé **ICOA2-G5-P7** est aussi assez proche et possède un IC_{50} similaire. Comme pour les résultats sur hNEIL1, avoir une famille d'inhibiteurs proche structuralement est rassurant et très intéressant pour essayer de déterminer quels sont les éléments de structure communs de ces composés qui pourraient être associés à l'inhibition. De plus, on peut identifier dans ces inhibiteurs des éléments structuraux analogues de la purine, ce qui pourrait indiquer qu'ils peuvent occuper le site actif de l'enzyme, un peu comme le ferait la base endommagée. Dans ce cas, ce serait des inhibiteurs compétitifs. Cependant, une ressemblance structurale avec une purine n'induit pas obligatoirement que ces composés vont dans le site actif. En effet parmi les inhibiteurs de la littérature, on peut noter que pour la molécule **TH5487,** la partie de l'inhibiteur contenant l'analogue de 8-oxoG ne se retrouve pas à l'emplacement de la base endommagée extrahélicale comme cela aurait pu être envisagé [371] (Figure 109).

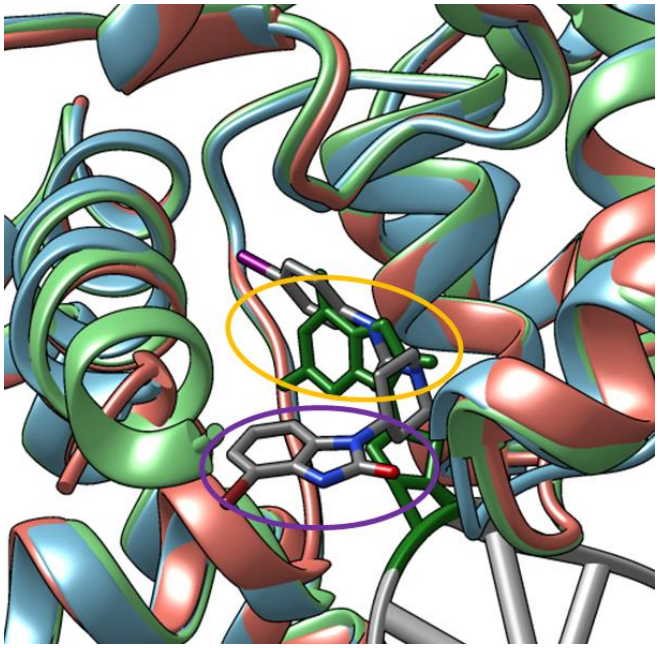


Figure 109 : Structure cristallographiques de hOGG1 seule (turquoise)**, complexée avec un duplex d'ADN contenant une 8-oxoG** (saumon) **ou avec TH5487** (vert pale) - L'ADN est en gris clair, la 8-oxoG en vert foncé, et TH5487 en gris foncé. La partie analogue de 8-oxoG de TH5487 est entourée d'un cercle violet, tandis que la 8-oxoG extrahélicale est entourée d'un cercle jaune. Les codes PDB sont 1KO9 (hOGG1 seule), 1EBM (hOGG1 complexée à l'ADN) et 6RLW (hOGG1 avec TH5487).

Enfin, les deux autres composés **ICOA2-E6-P29** et **ICOA2-G11-P13** sont moins efficaces et présentent un IC_{50} de 2,11 µM et 14,14 µM, respectivement. Leurs structures chimiques sont très différentes des autres inhibiteurs sélectionnés. Au vu de l'efficacité des composés précédents, nous nous sommes donc concentrés sur cette famille de molécules.

On peut noter aussi que l'IC_{50} que nous obtenons, bien qu'étant submicromolaire, reste plus élevé que les meilleurs inhibiteurs de la littérature. Comme l'IC_{50} est une valeur relative, nous avons donc déterminé l'IC_{50} des molécules de la littérature dans nos propres conditions afin de pouvoir véritablement comparer les valeurs (le TH5487 et O8 sont commerciaux et le SU268 a été resynthétisé par nos collègues Luigi Agrofoglio et Vincent Roy de l'ICOA à l'Université d'Orléans). L'IC_{50} que nous obtenons pour les inhibiteurs de la littérature est bien supérieur à celui publié. Cela fait de nos inhibiteurs, avec **Su268**, les

molécules plus efficaces **identifiées à ce jour** pour hOGG1. Ce résultat était assez étonnant, mais en réalité, il n'est pas rare que les IC_{50} varient selon les expérimentateurs. Par exemple, l'équipe de Mas Claret *et al.*, testant des inhibiteurs de MPG, a voulu utiliser la morine, un inhibiteur déjà publié comme « contrôle », et l'IC_{50} qu'ils ont évalué était également bien plus important que prévu (88 µM contre 2,6 µM publié) [237]. Une valeur d'IC_{50} ne peut être que comparée à une valeur ayant été mesurée dans des conditions rigoureusement identiques.

De plus, dans une publication plus récente de Visnes *et al.*, les IC_{50} sont mesurés en absence d'albumine de sérum bovin (« Bovine serum albumine » ou BSA en anglais) et sont donc fortement sous-estimés [372]. Selon cette même publication, le TH5487 ne semble pas avoir d'effet antitumoral *in vivo* chez la souris. Dans les conditions physiologiques, l'équivalent humain de la BSA est bien présent dans le sang (« human serum albumine », HSA) et est bien connue pour adsorber les petits molécules anioniques et hydrophobes [453]. C'est le cas de nombreuses drogues, et le HSA peut donc s'opposer à leur action thérapeutique en les détournant de leurs cibles. Pour illustrer *in vitro* l'influence de la BSA sur l'inhibition d'une enzyme, nous avons testé son effet sur le pouvoir inhibiteur de nos molécules et de celles de la littérature sur hOGG1 (Figure 110). Sans BSA, l'efficacité de tous les inhibiteurs est bien meilleure et ils inhibent complétement l'enzyme à une concentration finale de **0,5 µM**. Outre l'effet évoqué plus haut, la BSA permet aussi de stabiliser les protéines lorsque celles-ci sont très diluées comme c'est le cas dans les essais permettant de mesurer l'IC_{50} d'un inhibiteur. Dans le cas de hOGG1, l'enzyme est deux fois moins active sans BSA qu'avec 0,1 mg/mL de BSA final dans l'essai. Sans BSA, il est donc plus facile d'inhiber hOGG1. *In vitro*, la mesure des IC_{50} en présence de BSA nous parait donc bien plus pertinente pour évaluer le pouvoir inhibiteur d'un composé pour se rapprocher le plus des études de l'effet de ces inhibiteurs dans les cellules et plus tard *in vivo*. En conclusion, les inhibiteurs de hOGG1 que nous avons identifiés semblent bien être parmi les composés les plus efficaces identifiés à ce jour pour inhiber hOGG1.

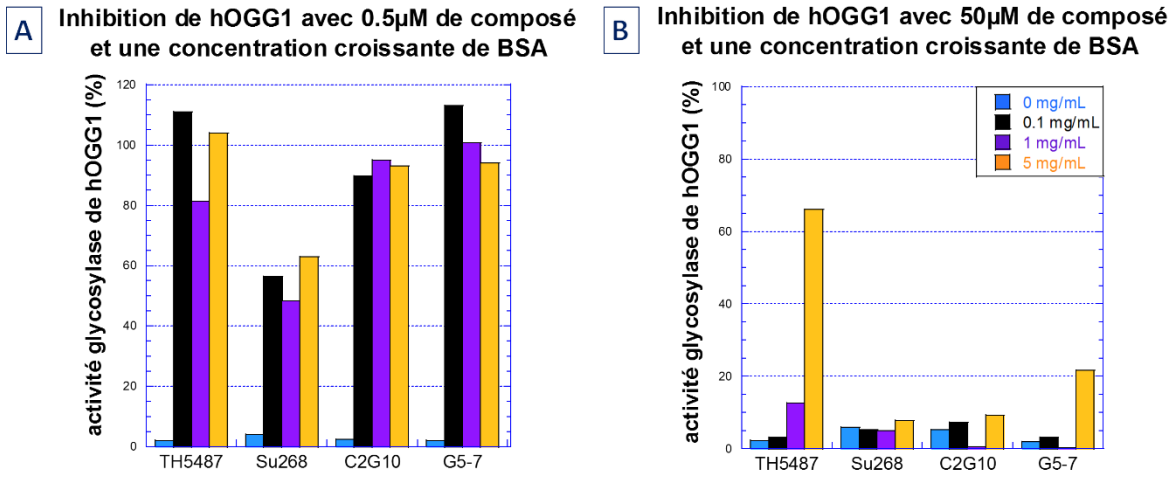


Figure 110 : Influence de la concentration de la BSA sur l'inhibition de hOGG1 - A) Activité glycosylase de hOGG1 en présence de 0,5 µM de quatre inhibiteurs confirmés (indiqués en abscisse), en fonction d'une gamme croissante de BSA (0 mg/mL en bleu, 0,1 mg/mL en noir, 1 mg/mL en violet et 5 mg/mL en orange). B) Activité glycosylase de hOGG1 en présence de 50 µM des quatre même inhibiteurs confirmés, en fonction d'une gamme identique de BSA.

I.5. Expansion des touches et optimisation des inhibiteurs

I.5.1. Expansions des touches du crible

Suite aux résultats du criblage, nous avons donc demandé à nos collaborateurs de l'ICOA de l'Université d'Orléans de nous fournir des analogues des molécules inhibitrices de hOGG1. Tout d'abord, nous

avons testé 46 molécules de structures assez diverses, supposées représenter des analogues des différents hits du crible (Figure 111).

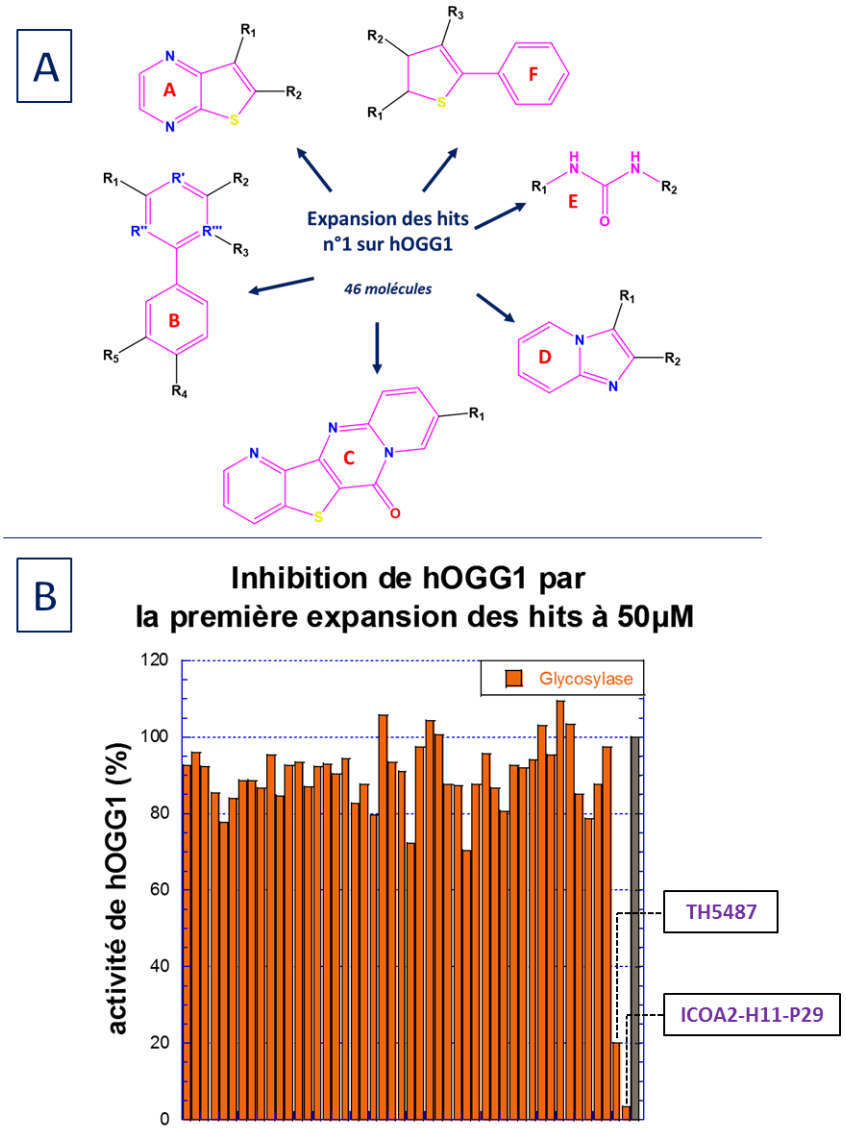


Figure 111 : Première expansion des hits sur hOGG1 - A) Structure générale des composés. B) Evaluation de l'activité glycosylase de hOGG1 en présence de 50 µM des composés. Les molécules TH5487 et ICOA2-H11-P29 ont été utilisées en contrôles positifs, et le contrôle négatif avec du DMSO est indiqué en gris foncé.

Cependant, aucune de ces molécules n'a montré un effet d'inhibition à 50 µM. Nous avons donc décidé de sélectionner uniquement des analogues des trois meilleurs inhibiteurs que nous avons identifiés : **ICOA2-A3-P30**, **ICOA2-H11-P29** et **ICOA2-G10-P10**, très proches structuralement (Figure 112). Trois de ces nouveaux analogues se sont révélés de bons inhibiteurs de hOGG1 dans la gamme de concentration considérée, et l'IC_{50} de ces nouvelles molécules a donc été évalué.

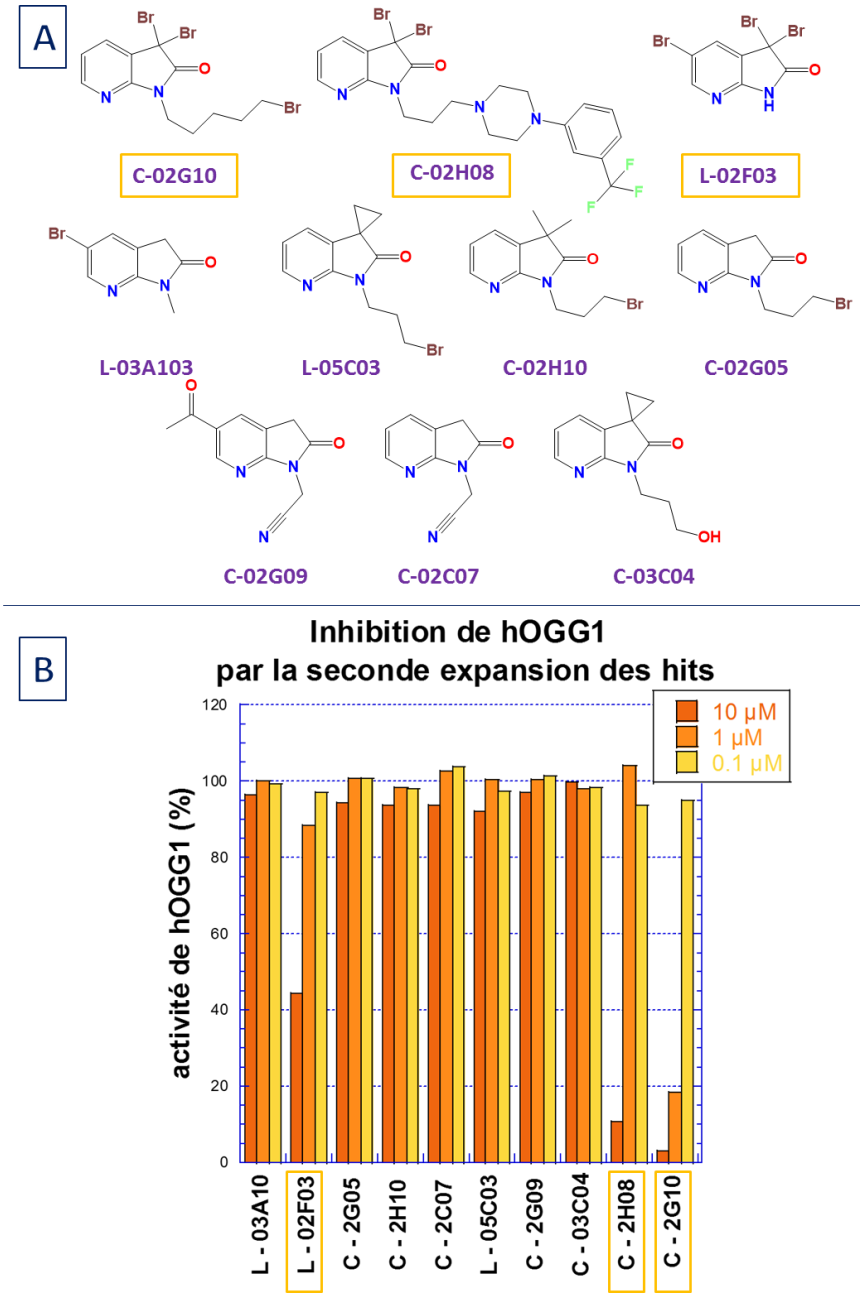


Figure 112 : Deuxième expansion des hits sur hOGG1 - A) Structure générale des composés. B) Evaluation de l'activité glycosylase de hOGG1 en présence de 3 concentrations de composés.

Parmi ces composés, C-02G10 était l'inhibiteur le plus efficace identifié à ce jour avec un IC_{50} de 0,291 µM (Figure 112 ou Annexe 5). Les courbes de dose-réponse à l'origine de la valeur d'IC_{50} sont indiquées en Annexe 6. A partir de cette étude comparative, certaines hypothèses sur les fonctions chimiques essentielles pour l'inhibition ont pu être établies. Tout d'abord, ces nouveaux inhibiteurs restent structuralement proches d'une purine oxydée, suggérant donc un effet sur le site actif. Ensuite les deux atomes de brome Br en position 7 semblent essentiels. Le remplacement ou l'absence de ces liaisons enlèvent l'effet inhibiteur des composés (comme pour la molécule **L-03A103** par exemple). En revanche, un atome de brome en position 5 sur la fonction pyridine ne semble pas avoir de réel impact sur l'effet inhibiteur. Enfin, la chaîne sur l'azote du cycle pyrrolidine semble aussi jouer un rôle important. Comme nous le voyons avec le composé **L-02F03**, une fonction amine secondaire ne semble pas être favorable, en comparaison avec **ICOA2-H11-P29** par exemple, où une chaîne aliphatique bromée sur l'amine rend l'inhibiteur plus efficace. En revanche, une chaîne trop encombrée ne semble pas non plus très favorable. **C-02H08** par exemple a un

IC_{50} plus grand que **C-02G10** (Figure 112 A), qui possède une chaîne plus encombrée avec un cycle azoté et un groupe trifluorométhyle -CF_3, qui est pourtant un bio-isostère du brome.

I.5.1. Analogues du noyau barbiturique

Nous avons ensuite voulu tester des analogues de la molécule **ICOA2-G5-P7**, un inhibiteur dont la structure était légèrement différente des 3 molécules précédentes, mais dont l'IC_{50} était proche. Pour ce faire, nous avons fait synthétiser par l'ICOA des analogues de ce composé en conservant le noyau barbiturique. La synthèse de ces composés a été plus difficile que prévue et nous avons donc dû nous orienter vers des composés commerciaux, dont la structure est indiquée en Figure 113. Deux composés ont montré un effet inhibiteur dans une gamme de concentration considérée. Malheureusement, avec un IC_{50} estimé entre 3 et 4 µM, ces inhibiteurs restent moins efficaces que **l'ICOA2-G5-P7**, mais soulignent néanmoins l'importance de la chaîne de carbone séparant le noyau barbiturique du benzène.

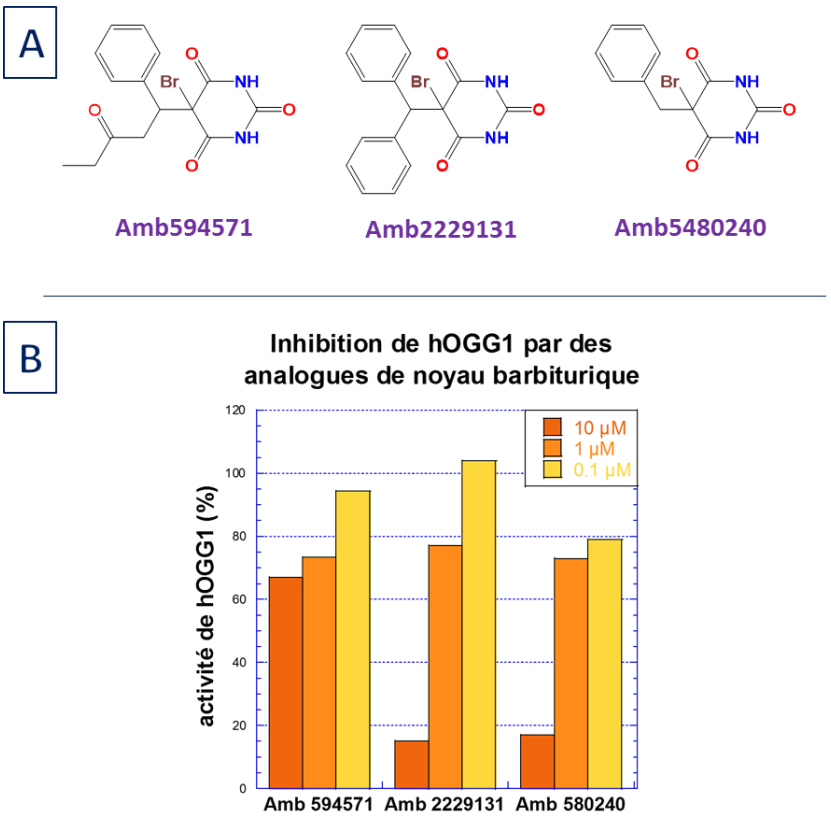
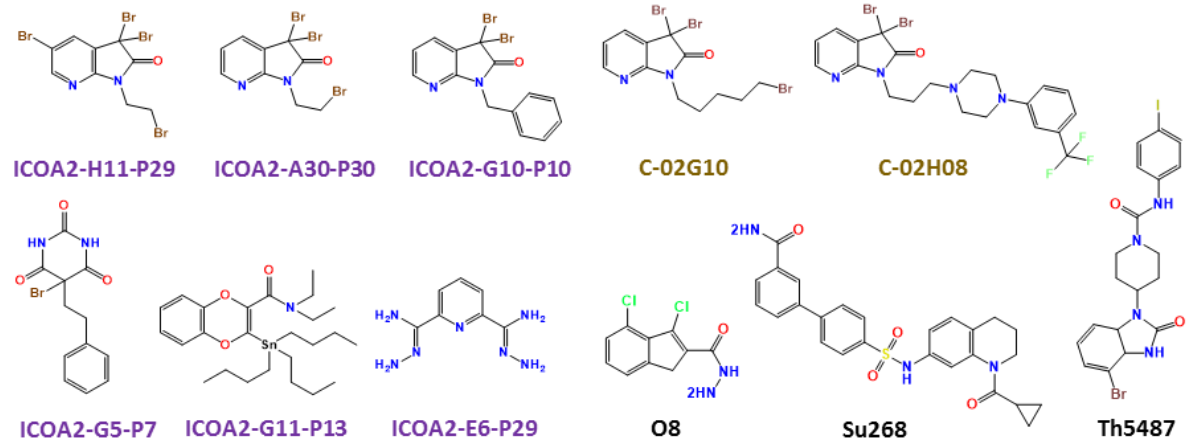
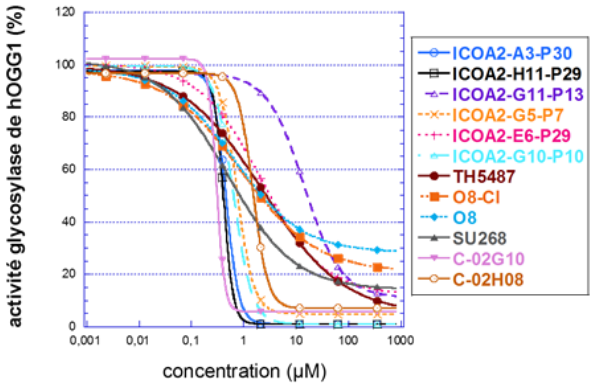


Figure 113 : Test des différents analogues des ICOA2-G5-P7 - A) Structure des composés. B) Evaluation de l'activité glycosylase de hOGG1 en présence de 3 concentrations de composés.

I.5.2. Bilan de l'expansion des touches

Finalement, l'expansion des inhibiteurs de hOGG1 nous a permis d'obtenir le composé **C-02G10**, avec l'IC_{50} le plus faible de tous les inhibiteurs identifiés dans notre équipe. Il semble que ce composé soit aussi efficace que le SU268 identifié récemment [370]. Nous avons donc identifié une nouvelle famille structurale d'inhibiteurs de hOGG1. En Figure 114 sont indiquées les valeurs d'IC_{50} finales des composés les plus efficaces. Les courbes d'IC_{50} sont présentées en Annexe 6.





	IC_{50} sur l'activité Glycosylase (µM)	
Composé	*Valeur publiée*	*Test "maison"*
ICOA2-A3-P30		0,455 ± 0,012
ICOA2-H11-P29		0,408 ± 0,008
ICOA2-G11-P13		14,101 ± 1,154
ICOA2-G5-P7		0,734 ± 0,037
ICOA2-E6-P29		2,152 ± 0,308
ICOA2-G10-P10		0,634 ± 0,012
TH5487	0,342	2,268 ± 0,660
O8	0,220 ± 0,080	0,674 ± 0,156
Su268	0,059	0,492 ± 0,105
C-02G10		0,300 ± 0,022
C-02H08		1,483 ± 0,319

Figure 114 : Récapitulatif des inhibiteurs de hOGG1

II. Etude du mode d'action des inhibiteurs identifiés

II.1. Spécificité des inhibiteurs identifiés

II.1.1. Analyse comparative de l'inhibition de hOGG1 entière et de la forme murine

Lors du crible nous avons utilisé une forme tronquée de OGG1 humaine, que nous avons noté dhOGG1. Dans cette partie, nous allons donc évaluer l'inhibition de la forme entière de la protéine humaine hOGG1. Dans ces expériences, nous avons aussi utilisé la forme murine de OGG1, notée mOGG1, une enzyme dont la fonction est très proche de hOGG1, possédant également une activité glycosylase sur la 8-oxoguanine et une activité AP lyase très faible. Structuralement ces deux enzymes sont également très similaires.

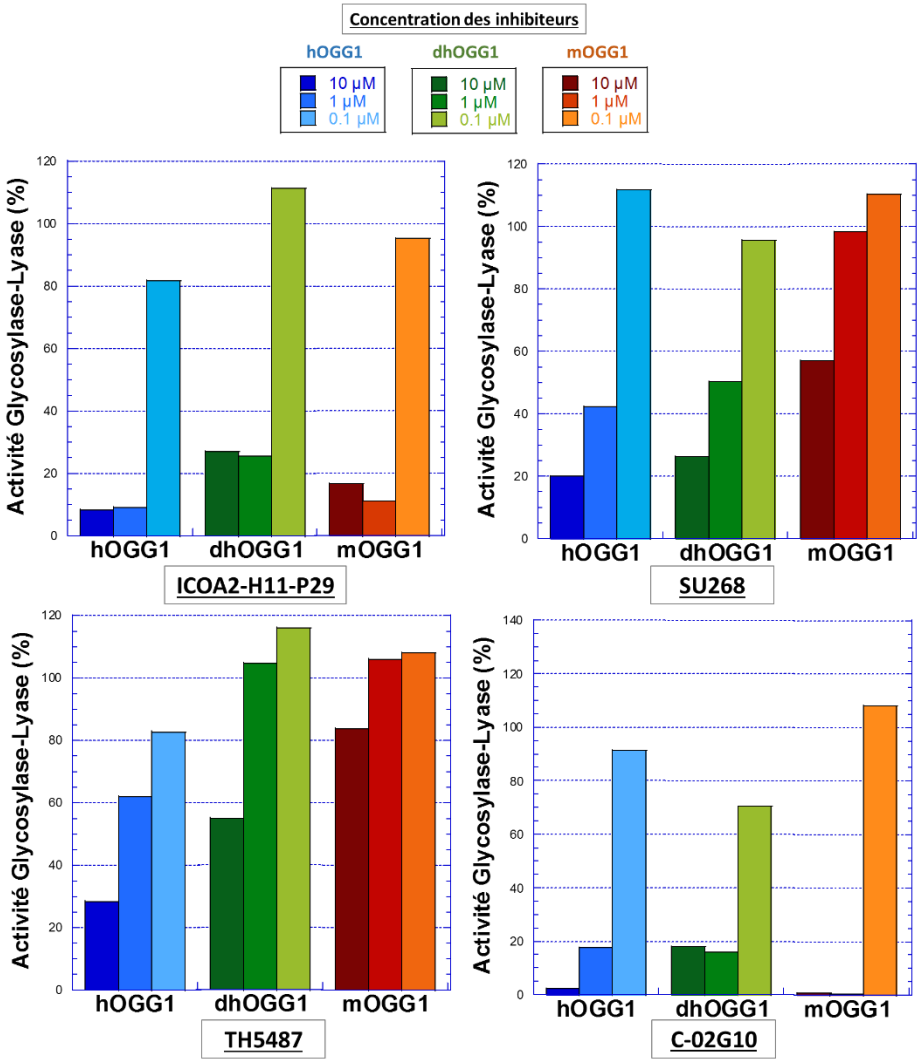


Figure 115 : Evaluation de l'inhibition des composés SU268, TH5487, C-02G10 et ICOA26H11-P29 sur hOGG1 entière, dhOGG1 et mOGG1 entière

Afin de disposer d'un contrôle, nous avons utilisé les deux meilleurs inhibiteurs de la littérature contre hOGG1 : SU268 et TH5487. Pour nos propres composés, nous avons testé les deux meilleurs inhibiteurs issus du criblage, C-02G10 et ICOA2-H11-P29. Trois concentrations de ces inhibiteurs (0,1, 1 et 10 µM) encadrant la valeur de leurs IC_{50} ont été testées (Figure 115). Il est assez clair que peu de différences sont observées entre la forme entière et la forme tronquée de hOGG1. En revanche, pour les deux inhibiteurs de la littérature, l'effet inhibiteur semble parfois moins prononcé en particulier pour la forme murine. Pour les deux inhibiteurs issus du criblage, l'inhibition reste constante. Si cela rend ces composés moins « sélectifs » entre l'enzyme humaine et murine, cela reste positif. En effet, cela justifie l'utilisation potentielle de la forme murine pour des études structurales par cristallographie et pour des études précliniques chez la souris.

II.1.2. Effet des inhibiteurs de hOGG1 sur son homologue AGOG chez les archées

Notre équipe travaille également en ce moment sur la 8-oxoG-ADN glycosylase de l'archée *Pyrrococus abyssii* (PabAGOG), qui est un homologue fonctionnel et structural de hOGG1, mais dont la taille et la séquence primaire sont très différentes de celles de hOGG1 (Figure 116). C'est l'une des représentantes des trois sous familles structurales des ADN glycosylases ayant la 8-oxoguanine pour substrat, les deux autres étant OGG1 et OGG2. Ainsi, il a été intéressant de tester nos inhibiteurs sur cette classe différente d'enzymes

afin d'évaluer leur spécificité, mais aussi de pouvoir potentiellement identifier des mécanismes d'action différents selon nos inhibiteurs.

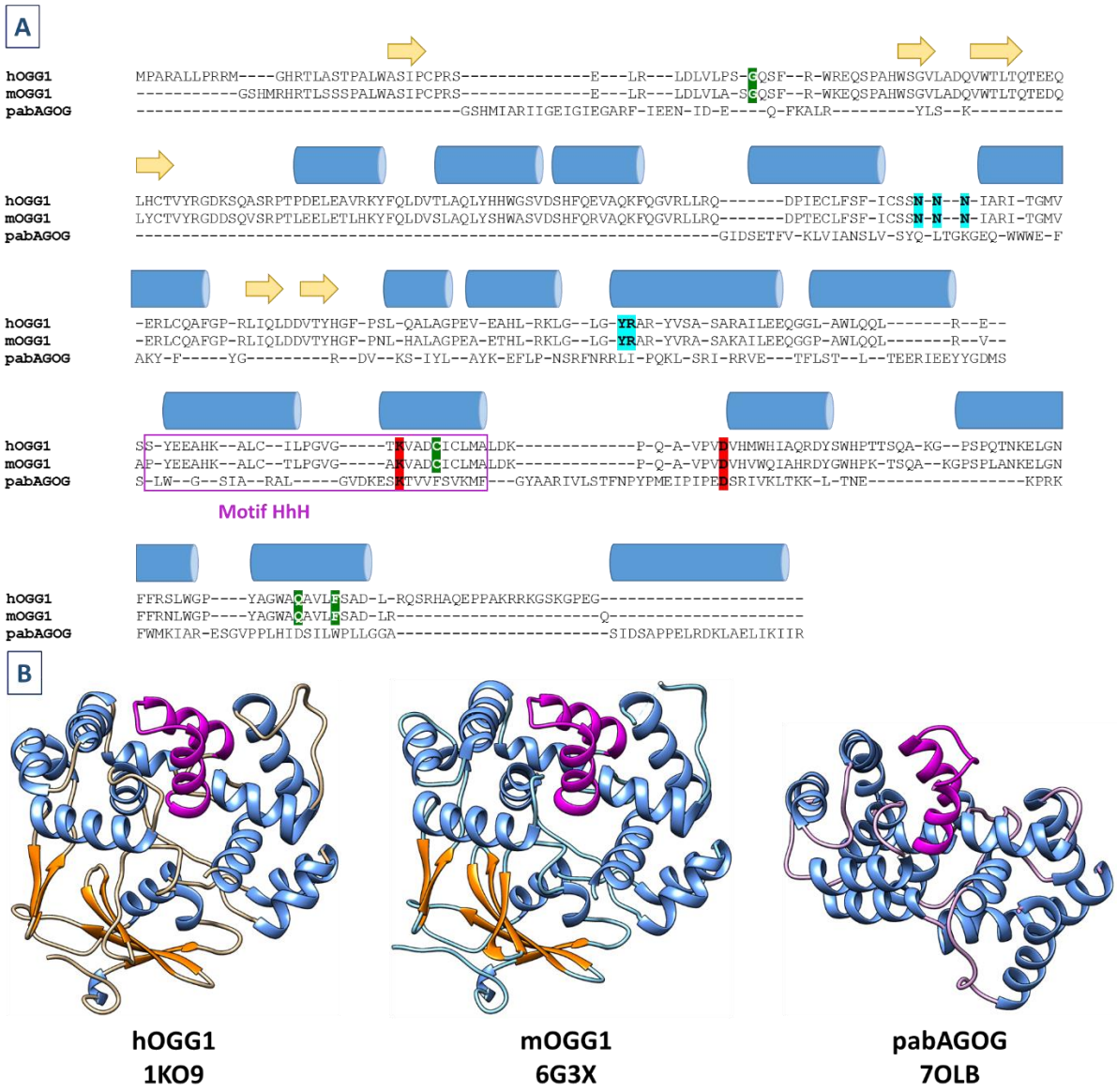


Figure 116 : Comparaison structurale entre hOGG1, PabAGOG et mOGG1 - A) Alignement de séquence de PabAGOG, mOGG1 et hOGG1 basé sur la structure de ces protéines. Les hélices α conservées sont indiquées en bleu foncé. Les brins β conservés chez mOGG1 et hOGG1 sont représentés par des flèches orange. Les résidus catalytiques sont indiqués sur fond rouge, les résidus d'intercalation en cyan, et ceux interagissant avec la base opposée en vert. B) Structure cristallographique de ces enzymes. Les hélices α sont colorées en bleu, les brins β en orange et le domaine HhH en magenta. Les PDB utilisées pour hOGG1, mOGG1 et PabAGOG sont indiquées sur la figure.

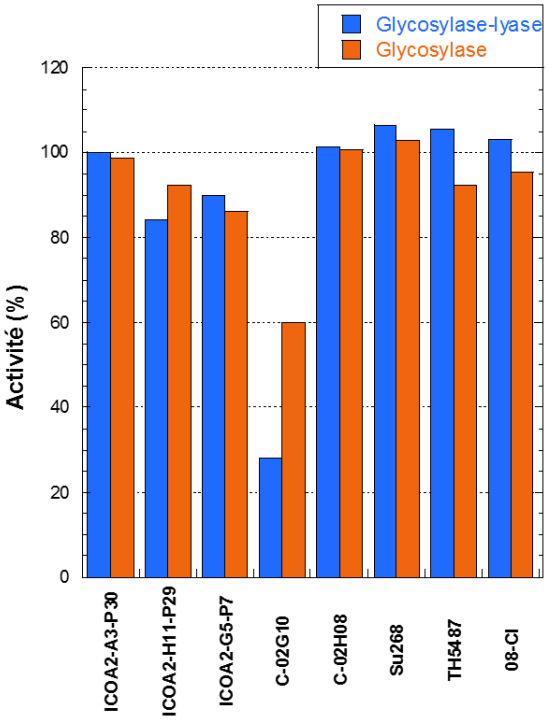


Figure 117 : Test d'inhibition des principaux inhibiteurs de hOGG1 sur PabAGOG

Parmi les inhibiteurs de hOGG1 testés, trois molécules ont montré des effets inhibiteurs : les composés ICOA2-H11-P29, ICOA2-G5-P7 et particulièrement C-02G10.

Suite à ce premier résultat, nous avons sélectionné les inhibiteurs **ICOA2-G5-P7** et **C-02G10**, et déterminé leurs IC_{50} sur les différentes activités de PabAGOG. En effet, bien que l'inhibition par le composé **ICOA2-G5-P7** soit très faible, ce composé présente une structure différente de **C-02G10**, contrairement à **H11-P29**. De fait, tester deux composés avec des structures distinctes peut permettre d'identifier des paramètres structuraux important d'inhibition. En parallèle, nous avons également tenté d'évaluer l'IC_{50} de la 8-oxo-2'-désoxyguanosine (8oxodG), un nucléoside ayant la capacité de se loger dans le site actif des AGOGs et qui a été utilisé comme modèle pour la reconnaissance de la 8-oxoG par AGOG dans les études structurales par cristallographie [300]. Nous n'avons pas pu mettre en évidence un effet inhibiteur significatif de la 8-oxodG sur hOGG1, pas plus que sur AGOG dans la gamme de concentration choisie, et ce bien qu'il est été publié que ce nucléoside soit capable de se fixer dans la poche du site actif de la AGOG de *Pyrobaculum aerophilum* (PaeAGOG) avec une constante de dissociation K_D de 345 nM [300] (Figure 118). On peut rapprocher ce résultat de l'observation du citrate dans le site actif de hNEIL1, sans observer d'inhibition de l'enzyme. Cela illustre que la compétition entre un ADN et une petite molécule, dans un test d'activité par exemple, est bien souvent complètement déloyale.

		IC_{50} (µM)		
	Substrat	**8oxoG-ADN Glycosylase AP Lyase**	***8oxoG ADN glycosylase***	***AP Lyase***
C-02G10	*dbGO/C*	43,87 ± 2,67	65,37 ± 10,25	NA
	sbGO/C	NA	39,46 ± 3,50	NA
	dbAP/C	NA	NA	73,7 ± 4,40
G5-P7	*dbGO/C*	208,8 ± 28,7	231,5 ± 53,2	NA
	sbGO/C	NA	138,1 ± 17,4	NA
	dbAP/C	NA	NA	--
8oxodG	*dbGO/C*	--	> 500	NA
	sbGO/C	NA	--	NA
	dbAP/C	NA	NA	> 2000

Figure 118 : Inhibition de PabAGOG par C-02G10, ICOA2-G5-P7 et la 8-oxodG Les substrats ADN sont indiqués comme suit : db/sb pour double/simple brin, contenant les lésions GO = 8-oxoG et AP = site abasique.

Bien que moins efficace que sur hOGG1, le composé C-02G10 est l'inhibiteur le plus efficace sur AGOG avec des IC_{50} de l'ordre de 40 à 70 µM pour l'ensemble des différentes activités de PabAGOG. Le composé ICOA2-G5-P7 présente un pouvoir inhibiteur 3 à 4 fois moins fort, mais reste cependant largement plus efficace que la 8-oxodG. Cette différence d'inhibition entre les composés est bien plus marquée sur PabAGOG que sur hOGG1. Cela peut s'expliquer par une spécificité plus importante de ICOA2-G5-P7 pour hOGG1.

II.2. Constantes de dissociation K_D des complexes PabAGOG/inhibiteur

Pour évaluer une constante de dissociation pour les complexes PabAGOG avec C-02-G10 et la 8-oxodG, nous avons suivi le même protocole que Lingaraju *et al.* [459] qui utilisent la fluorescence intrinsèque des différents tryptophanes localisés dans le site actif de PaeAGOG et dont la chaîne latérale en particulier celle du Trp222 subit un changement conformationnel très important lors de la fixation de la 8-oxoG dans le site actif (Figure 119). Tout changement conformationnel de ce tryptophane devrait être associé à une modification du spectre d'émission de fluorescence si l'inhibiteur se fixe au niveau du site actif comme le fait la 8-oxodG (Figure 119).

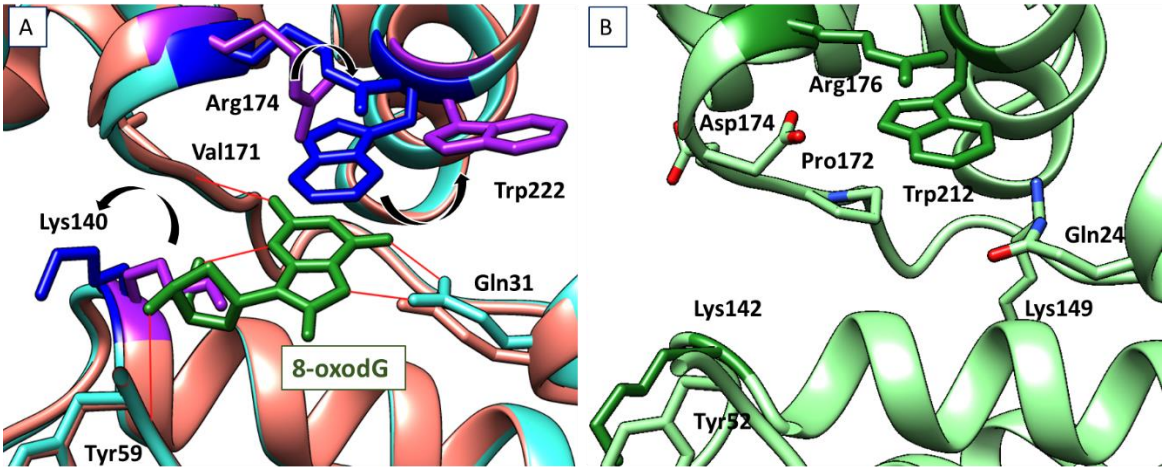
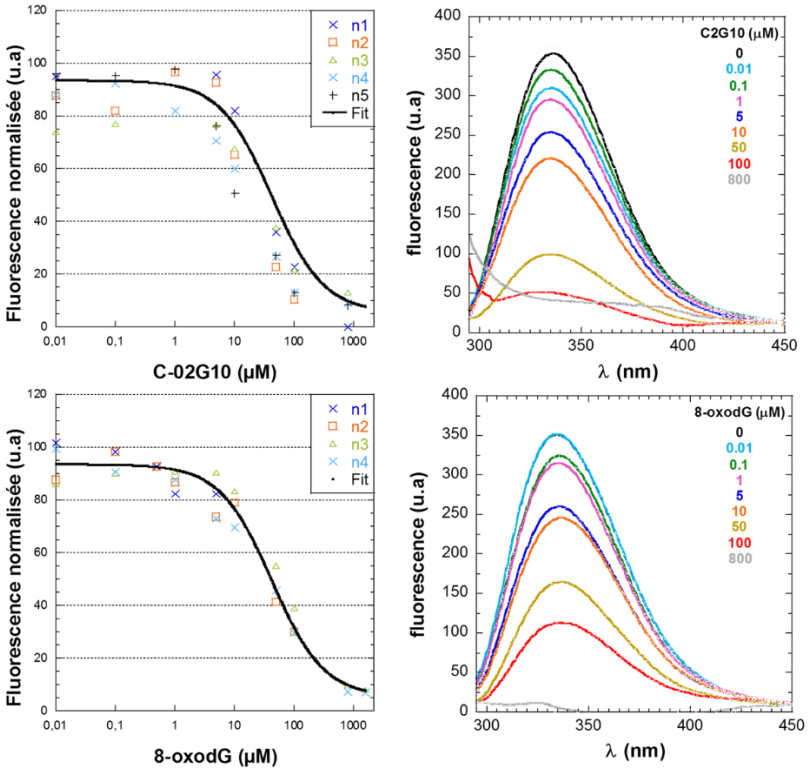


Figure 119 : Changement conformationnel suite à la complexation de PaeAGOG avec la 8-oxodG - A) Changement conformationnel entre la poche du site actif de PaeAGOG seul (en saumon, PDB 1XQO) et PaeAGOG complexée (en bleu clair, PDB 1XQP) à la 8-oxodG (en vert foncé). Les résidus dont la position change sont indiqués en bleu foncé (enzyme complexée) et en violet (enzyme libre), et la direction du changement est indiquée par des flèches noires. B) Site actif de l'enzyme PabAGOG, avec les acides aminés très bien conservés par rapport à PaeAGOG.

En incubant PabAGOG en présence de 8-oxodG ou de la **C-02G10**, la fluorescence intrinsèque des tryptophanes peut varier ce qui peut permettre d'estimer une constante de dissociation apparante (K_{Dapp}) en exploitant les spectres d'émission de fluorescence obtenus à différentes concentrations en inhibiteur (Figure 120). Le K_D estimé par cette méthode avec C-02G10 est environ fois deux plus faible que celui avec la 8-oxodG. Cela signifie que l'affinité de PabAGOG pour l'inhibiteur est plus forte. Ces K_D restent similaire, tandis que les inhibitions de ces deux molécules sont très différentes. Le K_D étant le rapport des constantes de vitesse d'équilibre ($K_D = k_{on}/k_{off}$), une valeur de K_D peut correspondre à différents k_{on} et k_{off}. Pour la C-02G10, nous pouvons imaginer que l'effet inhibiteur sera favorisé par un temps de résidence important dans le site actif, avec un k_{on} important et un faible k_{off}. Ainsi la C-02G10 pourrait rester potentiellement plus longtemps dans le site actif que la 8-oxodG par exemple, et serait moins facilement déplacée par le substrat ADN. Il est à noter que nous avons évalué une constante de dissociation, K_D, et non une constante d'inhibition, K_i, notre test n'évaluant pas l'activité de PabAGOG.

Il faut aussi noter que la méthode utilisée a ses limites. Ainsi, la fluorescence intrinsèque de la protéine peut varier indépendamment de la formation d'un complexe avec son ligand, par exemple avec le ligand formant un filtre et absorbant la lumière d'excitation [470]. Afin de valider plus en profondeur les valeurs obtenues et d'éliminer les potentiels doutes d'un « quenching » de la fluorescence par d'autres mécanismes, il pourrait être intéressant de vérifier l'effet de la C-02G10 sur un tryptophane libre, de faire varier le paramètre température ou de calculer le paramètre Stern-Volmer, qui permet d'évaluer l'influence du « collisional quenching » (quand un un fluorophore excité est désactivé par la le contact d'autres moléucles en solution).



Ligand	K_D (µM)
C-02G10	22,5 ± 4,5
8-oxodG	42,2 ± 6,4

Figure 120 : Détermination de la valeur du K_D pour le composé C-02G10 et 8-oxodG

II.3. Thermostabilité de hOGG1 en présence d'inhibiteurs

Nous avons testé par TSA l'ensemble des inhibiteurs confirmés sur hOGG1, y compris les inhibiteurs que nous n'avons pas retenus suite au test d'interaction à l'ADN, et nous avons évalué le **ΔT_m** de la protéine en leur présence. Les résultats obtenus sont indiqués en Figure 122.

ΔTm (°C)	1	2	3	4	5	6	7	8	9	10	11	12
A	ICOA2-G11-P3 x	ICOA2-D8-P3 -3,16 ± 0,22	ICOA2-G5-P7 -1,51 ± 0,4	ICOA2-G10-P10 -3,27 ± 0,31	ICOA2-F3-P12 -3,15 ± 0,18	ICOA2-G11-P13 x	ICOA2-D2-P15 -1,68 ± 0,51	ICOA2-A9-P22 -0,23 ± 0,07	ICOA2-E6-P29 -3,14 ± 0,19	ICOA2-H11-P29 -3,77 ± 0,19	ICOA2-A3-P30 -3,21 ± 0,12	ICOA2-C11-P15 -3,90 ± 0,39
B	Prest-H9-P4 x	Prest-G9-P5 x	Prest-D2-P6 0,09 ± 0,10	Prest-H3-P10 -2,31 ± 0,13	Prest-F8-P10 -0,92 ± 0,16	Prest-B11-P10 1,05 ± 0,25	Prest-E6-P12 x	Prest-C8-1P3 -0,57 ± 0,03	Juglone x	Purpurine x	GP-H3-P1 x	GP-C7-P6 -3,22 ± 0,26
C	ICOA1-H11-P6 -3,29 ± 0,21	ICOA1-A2-5 0,39 ± 0,21	C-02C07 -3,75 ± 0,18	C-03C04 3,76 ± 0,12	C-02G10 x	C-05C03 -2,93 ± 0,21	C-02G05 -2,54 ± 0,17	C-02H10 -2,04 ± 0,32	C-02H08 -0,69 ± 0,21	C-02F03 -4,03 ± 0,14	C-05A10 -6,82 ± 0,31	C-02G09 -2,08 ± 0,19
D	LA-32 -1,96 ± 0,23	Prest-G6-P14 x	Prest-F6-P8 x									
E	Contrôle hOGG1											

Effet observé

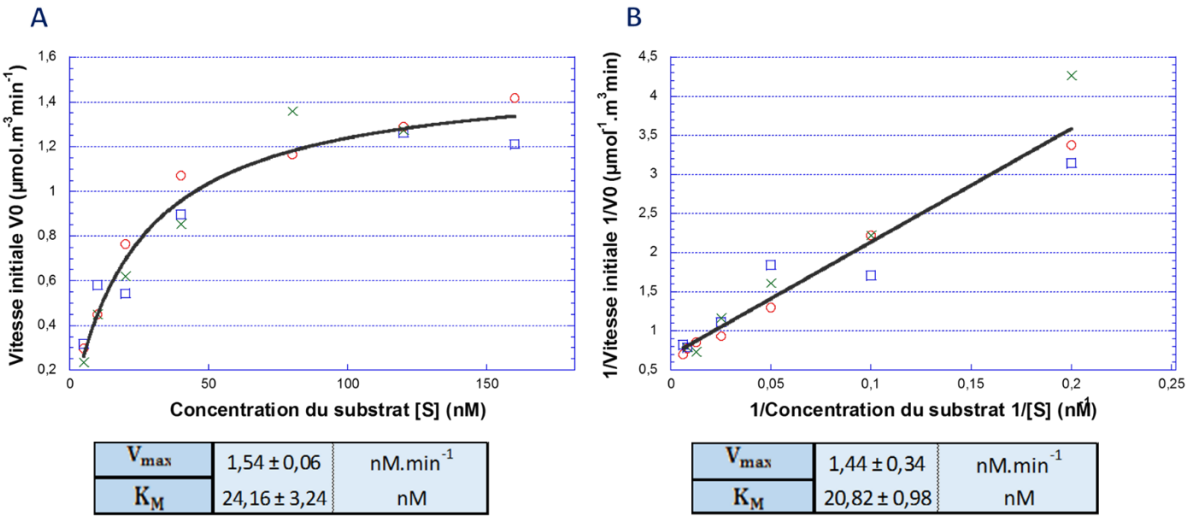
	Effet stabilisant (ΔTm > 1 °C)
	Peu d'effet (-1 °C > ΔTm > 1 °C)
	Effet déstabilisant (ΔTm < -1 °C)
	Non applicable

Figure 121 : Variation de la température de fusion (T_m) de hOGG1 (ΔT_m) en présence de différents inhibiteurs issus de notre criblage ou d'analogues

Très peu d'inhibiteurs stabilisent réellement notre enzyme d'intérêt. Seul le composé Prest-B11-P10 (ou Ebselen) augmente le T_m de hOGG1 d'à peine 1 °C. A l'exception de C-02H08, la plupart des composés testés de la même famille structurale, comme ICOA2-A3-P30 et ICOA2-H11-P29 déstabilisent légèrement l'enzyme. Pour C-02G10, aucune courbe n'a pu être obtenue, potentiellement à cause d'une déstabilisation trop élevée de la protéine. Le composé ICOA2-G5-P7 a quant à lui moins d'effet sur la stabilité thermique de hOGG1. Cet aspect peut en faire un meilleur candidat que les autres meilleurs composés pour des études structurales par cristallographie de hOGG1/inhibiteur. Ces différences peuvent également indiquer un mécanisme d'inhibition différent entre ces composés pourtant similaires. A titre de comparaison, l'inhibiteur TH5487 stabilisent plutôt la protéine, avec une augmentation du T_m publiée de 3-4°C avec des concentrations équivalentes, impliquant surement un mécanisme d'inhibition différent [371].

II.4. Cinétiques enzymatiques de hOGG1 en présence d'inhibiteurs

Afin d'évaluer le mécanisme d'inhibition des composés identifiés sur hOGG1, nous avons tenté de faire des études cinétiques sur hOGG1 en présence d'inhibiteur. Pour ce faire nous avons utilisé la forme tronquée dhOGG1, que nous avions déjà utilisée pour l'estimation des IC_{50} et le criblage. Le test d'activité utilisé est le même que celui utilisé lors de la confirmation des hits du criblage primaire avec comme substrat un court duplexe d'ADN contenant une 8-oxoG opposée à un C et radiomarqué en 5' par du ^{32}P sur le brin endommagé. Dans les cinétiques présentées ici, seule l'activité 8-oxoG-ADN glycosylase a été considérée (voir Matériel & Méthodes). Le protocole est sensiblement identique à celui utilisé lors des tests d'activité « classiques » si ce n'est qu'à chaque point de cinétique, une aliquote est prélevée dans le mélange réactionnel et stoppé indépendamment. Le principe de la cinétique étant d'observer le changement de vitesse de réaction en fonction de la quantité de substrat, il nous fallait donc déterminer une concentration fixe de protéine avec laquelle effectuer ces expériences. Comme nous souhaitions comparer différentes cinétiques en présence d'inhibiteurs, nous nous sommes placés à la même concentration d'enzyme que celles utilisées lors de la détermination des IC_{50}, soit 2 nM, pour une température d'incubation de 37 °C. Les vitesses initiales ont été extraites d'au moins 3 cinétiques enzymatiques suivant la formation du produit en fonction du temps d'incubation à une concentration en substrat donnée (Figure 122).



V_{max}	1,54 ± 0,06	nM.min^{-1}
K_M	24,16 ± 3,24	nM

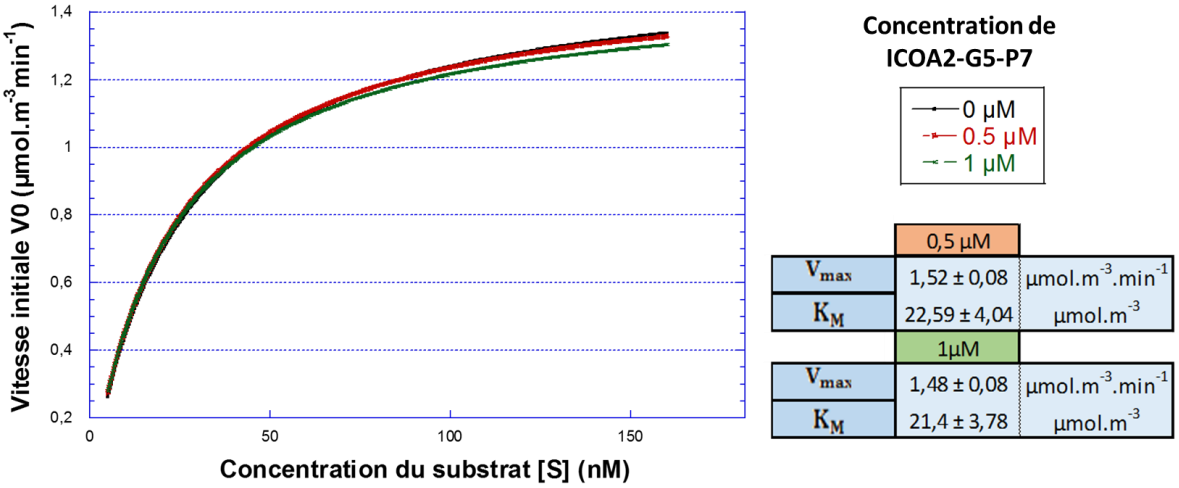
V_{max}	1,44 ± 0,34	nM.min^{-1}
K_M	20,82 ± 0,98	nM

Figure 122 : Tracé des courbes de Michaelis-Menten et Lineweaver-Burk pour dhOGG1 en absence d'inhibiteur et détermination du V_{max} et du K_M selon les deux techniques

Les valeurs de V_{max} et du K_M nous permettent de calculer une constante de vitesse de réaction $k_{cat} = V_{Max}/[hOGG1]_0 = 777$ min^{-1} et une efficacité catalytique (turnover) $k_{cat}/K_M = 0{,}032$ min^{-1}.nM^{-1}. Ces valeurs sont relativement du même ordre de grandeur que celles indiquées dans la littérature ($k_{cat} = 1080$ min^{-1} et $k_{cat}/K_M = 0{,}011$ min^{-1}.nM^{-1}) [471]. On peut noter que notre enzyme semble avoir une plus grande efficacité catalytique. Cet effet peut être dû à un meilleur titre de hOGG1 dans nos échantillons. Par exemple, nous avons effectué ces expériences sans congélation de l'enzyme, ce qui peut expliquer une efficacité légèrement supérieure. De manière générale, l'efficacité catalytique de hOGG1 apparait relativement faible. Cependant, son efficacité *in vivo* pour éliminer la 8-oxoG est illustrée par son effet antimutateur incontestable de la transversion G:C vers T:A. Il existe donc nécessairement dans le noyau cellulaire des partenaires moléculaires capables de stimuler son activité, comme cela a été montré in vitro avec APE1 par exemple [471]. Cette absence de turnover performant *in vitro* reflète probablement une affinité relativement forte de l'enzyme pour ses produits de réaction (rétro-inhibition) ce qui ralentit fortement son recyclage, propriété partagée par de nombreuses ADN glycosylases.

Nous avons ensuite étudié l'effet des inhibiteurs sur les valeurs de K_M en reproduisant les cinétiques en présence d'inhibiteurs afin de pouvoir proposer un mode d'inhibition et éventuellement une constante d'inhibition K_i. Nous avons sélectionné les inhibiteurs **ICOA2-G5-P7** et **ICOA2-H11-P29**. Cependant, malgré de nombreuses tentatives, aucune donnée cinétique n'a pu être exploitable. Tout d'abord, les données obtenues avec **ICOA2-H11-P29** sont difficilement exploitables. En effet, à haute concentration, aucune activité n'est présente, tandis qu'à basse concentration, aucune différence n'est observée avec la protéine contrôle. De plus, de nombreux réplicas ont été testés et de grandes différences ont été observées, ce qui rend l'analyse sans fondement. En premier lieu, on pourrait expliquer le profil de ces résultats en comparant avec la courbe d'IC_{50}, dont la pente est très élevée, ce qui pourrait donc expliquer ce phénomène d'inhibition très forte, puis d'une faible inhibition sans intermédiaire entre deux concentrations relativement proches. Cette tendance peut souvent s'expliquer par des inhibiteurs ayant pour effet d'agglomérer les protéines et de provoquer leur précipitation. Au vu des résultats de thermostabilité précédents, cette hypothèse peut sembler cohérente. Avec une instabilité intrinsèque de la protéine et possiblement des inhibiteurs, il devient difficile d'obtenir des données fiables et reproductibles. Nous avons donc renoncé à aller plus loin avec l'inhibiteur **ICOA2-H11-P29.**

Pour **ICO2-G5-P7**, les données obtenues sont également difficilement exploitables. Ainsi, les profils observés entre deux points de concentration ne différent pas de la courbe sans inhibiteur (Figure 123). De même que précédemment, si l'on compare les conditions de cinétique identiques aux conditions de tests d'activité effectués pour déterminer la valeur d'IC_{50}, on ne remarque aucun effet de cet inhibiteur. L'hypothèse la plus probable est donc que l'inhibiteur ait été dégradé dans le DMSO lors des cinétiques. En effet nous avions peu de quantité de ces produits de synthèse, et il nous a fallu utiliser des stocks déjà congelés plusieurs fois. Une autre hypothèse pourrait être liée à un composé non pur, que nous évoquerons dans la partie cristallographique de ce manuscrit. A ce stade, il est important de noter que l'intégralité de ces expériences a été faite en toute fin de la partie expérimentale de cette thèse. Nous n'avons donc pas pu reproduire celles-ci avec des stocks d'inhibiteurs frais pour vérifier nos hypothèses. Il est également possible que la protéine ait été partiellement dégradée.



	0,5 µM	
V_{max}	1,52 ± 0,08	µmol.m^{-3}.min^{-1}
K_M	22,59 ± 4,04	µmol.m^{-3}
	1µM	
V_{max}	1,48 ± 0,08	µmol.m^{-3}.min^{-1}
K_M	21,4 ± 3,78	µmol.m^{-3}

Figure 123 : Courbes de cinétique obtenues pour différentes concentrations d'inhibiteur ICOA2-G5-P7 - en rouge 0,5 µM, en vert 1 µM et en noir, la protéine seule.

Nous n'avons donc pas été capable d'établir des valeurs de K_M en présence de deux de nos inhibiteurs : **C-02G10** et **G5-P7**. Afin d'obtenir tout de même des valeurs probantes de K_M et d'établir un mécanisme d'inhibition pour nos composés, différentes solutions peuvent être envisagées. Tout d'abord, utiliser des composés fraichement synthétisés permettra d'être plus confiant sur les résultats obtenus, sans crainte de dégradation des molécules. De plus, comme nous le verrons plus tard dans ce manuscrit, des problèmes de pureté de la G5-P7 notamment, peuvent aussi venir perturber les résultats. En parallèle, il est également prévu d'utiliser des oligonucléotides spécialement conçus pour évaluer directement l'activité glycosylase de hOGG1, sans ajout de soude. Malheureusement, ces composés, synthétisés par notre collègue Didier Gasparutto (CEA, Grenoble) n'ont pas eu le temps d'arriver dans le temps imparti de la thèse. Cela permettra d'avoir une lecture plus directe et rapide de la cinétique de la réaction, et facilitera grandement ces expériences. Enfin, des expériences de cinétique avec d'autres inhibiteurs, tels que **C-02H08** par exemple, avec un effet différent sur la protéine, pourrait également être envisagées.

III. Etude structurale de la protéine hOGG1 et de ses inhibiteurs

III.1. Cristallogenèse de la protéine seule

Suite aux différents résultats obtenus avec les inhibiteurs de hOGG1, nous avons cherché à obtenir des cristaux de cette protéine seule en présence de nos composés. Cette démarche était motivée par la présence dans la littérature de conditions de cristallisation de hOGG1. Pour ces expériences de cristallisation, nous avons utilisé une deuxième forme tronquée de hOGG1, notée d2hOGG1. Elle est similaire à dhOGG1,

mais les extrémités de sa séquence ont été légèrement plus raccourcies. Ainsi nous avons retiré 22 acides aminés en C-terminal, et 13 acides aminés en N-terminal. Tous ces résidus forment des chaînes prédites déstructurées, sans aucun effet sur l'activité *in vitro* de hOGG1, et pouvant nuire à la cristallisation de la protéine. Nous avons tout d'abord cherché à répéter les conditions de cristallisation établies par Zander *et al.* [472] ainsi que par Verdine *et al*. [371]. Malheureusement, les essais réalisés ont conduit à la précipitation de la protéine (plus de 60% des conditions) et nous n'avons pas obtenu de cristaux. En parallèle, nous avons mis en place des criblages cristallographiques sur la protéine seule en utilisant les kits JCSG+, MDL_Structure de Molecular Dimension, Q_Compas, Q_Pact de Qiagen et Wizard 1+2 de Rigaku Reagents (dont les compositions sont indiquées en Annexe 8), pour un total de 768 conditions, mais aucun cristal n'a été obtenu.

Suite à l'échec de la cristallisation de la forme humaine de d2hOGG1, nous avons décidé de tester la cristallisation de la forme murine mOGG1. En effet Visnes *et al*. [371] l'ont utilisée avec succès pour obtenir une structure de leur inhibiteur Th5487 avec cette protéine. De plus, les composés C-02G10 et ICOA2-H11-P29 ont également un effet inhibiteur de cette protéine (Figure 115), ce qui en fait un bon modèle. Nous avons donc reproduit les conditions de cristallisation de mOGG1 décrites, et mis en place un criblage de 384 conditions de cristallisation (kits Q_Classics de Qiagen, MDL_Structure, Morpheus de Molecular Dimension et HR1+2 de Hampton Research). Malheureusement, ici encore, aucun cristal n'a été obtenu.

IV. Etude structurale d'un analogue fonctionnel de hOGG1, la 8-oxoG-ADN glycosylase d'archée AGOG

IV.1. Tentative de cristallisation du complexe PabAGOG/C-02G10

Devant notre incapacité à cristalliser mOGG1 ou d2hOGG1, malgré la présence de leurs structures dans la littérature, nous avons décidé d'étudier un autre analogue de hOGG1, la 8-oxoG-ADN glycosylase d'archée (AGOG) de *Pyrococcus abyssi* (PabAGOG). En effet, les inhibiteurs de hOGG1 **C-02G10** et **ICOA2-G5-P7** sont aussi capables d'inhiber PabAGOG (Figures 117 et 118). De plus, le structure 3D de cette protéine a été résolue récemment dans notre équipe (données non publiées). La protéine PabAGOG a été purifiée en suivant le protocole établi dans l'équipe et concentrée à 30 mg/mL dans son tampon final (250 mM NaCl, 20 mM Hepes pH 7,6, voir paragraphe Matériel et Méthodes dédié). Nous avons d'abord tenté de cristalliser la protéine en présence de **20 mM final de C-02G10 et 1 % DMSO final,** avec un criblage de 192 conditions cristallographiques (kits JCSG+ de Molecular Dimension et Q_Classics de Qiagen). Malheureusement ces différents essais de cristallisation n'ont pas abouti à l'obtention de cristaux, avec un fort taux de précipitation. De plus, la C2-02G10 pure se présente sous la forme d'une « huile », ce qui ne permet pas de faire un procédé de « soaking ». Cette huile est également très peu soluble dans autre chose que le DMSO, ce qui ne permet donc pas de tester d'autres conditions.

Nous avons donc décidé de changer d'inhibiteur et de prendre le composé **ICOA2-G5-P7** qui, bien que moins efficace sur PabAGOG, semble de manière générale être plus soluble et moins déstabilisant pour les protéines, comme nous avons pu le voir lors des études de thermostabilité (Figure 121).

IV.2. Etude structurale par cristallographie du complexe PabAGOG/ICOA2-G5-P7

IV.2.1. Trempage des cristaux en présence d'inhibiteur (soaking)

PabAGOG a été cristallisée dans les conditions définies au laboratoire (Solution **1.26 du kit JCSG+ soit 1 M Chlorure de Lithium pH 4, 0 1M Citrate de sodium, 20 % PEG 6000)**. Cependant, avec une solution faite

au laboratoire, les cristaux étaient de moins bonne qualité que ceux obtenus avec le kit. Il nous a donc fallu optimiser les conditions de cristallisation en faisant varier :

- La concentration de l'agent précipitant PEG 6000 (19 %, 20 %, 21 %, 22 %, 23 % et 24 %)

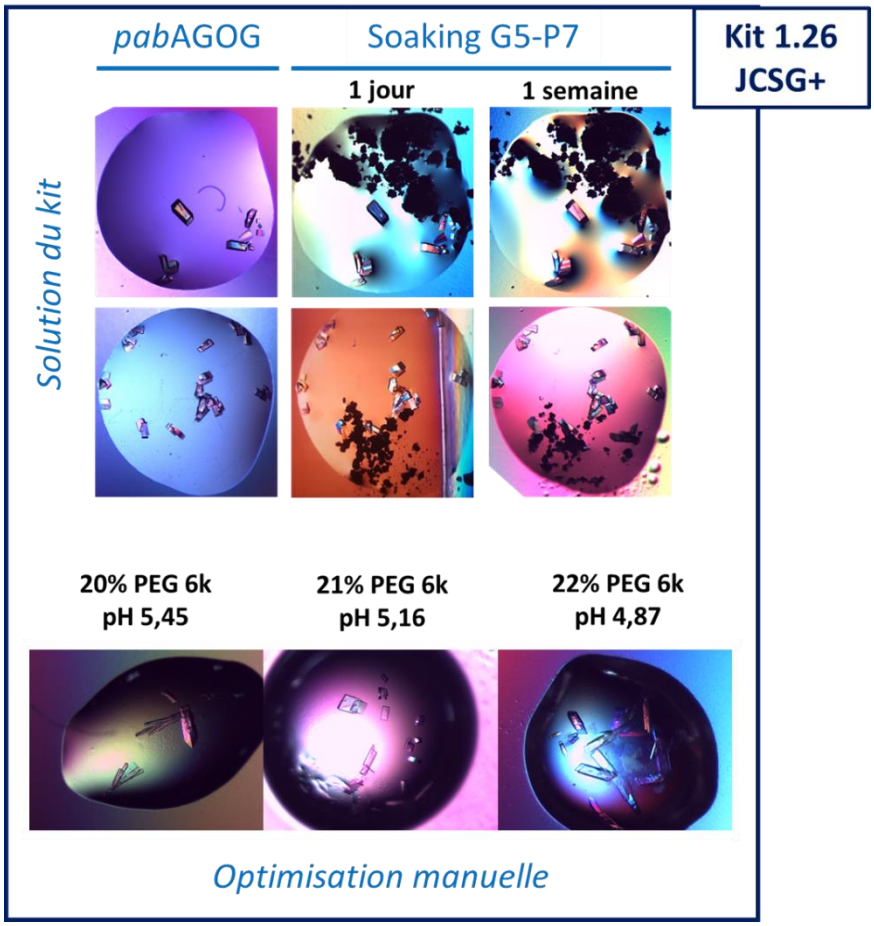
- Le pH de la solution (4, 4,29, 4,58, 4,87, 5,16, 5,45, 5,74 et 6)

- La concentration du LiCl (0,75 M, 1 M et 1,25 M)

En tout, **96 conditions** ont été préparées, et nous avons obtenu des cristaux pour 21 d'entre elles. De ces résultats nous avons pu conclure que :

- Le LiCl a peu d'impact sur la formation des cristaux, sauf à la concentration 1,25 M, solubilisant probablement trop la protéine.

- Une concentration plus élevée de PEG 6000 permet l'apparition de cristaux mieux formés, à condition que le pH diminue, avec cependant un fort enchevêtrement des cristaux.



PEG 6000	LiCl	*pH*
20,0%	0,75-1 M	5,45 -> 6
21,0%	0,75-1 M	4,87 -> 6
22,0%	0,75-1 M	4,29 -> 5,45

Figure 124 : Optimisation manuelle des conditions de cristallisation de PabAGOG seule et photos de cristaux obtenus de PabAGOG en présence d'ICOA2-G5-P7

Sur les gouttes contenant les cristaux de PabAGOG, nous avons directement déposé l'inhibiteur ICOA2-G5-P7 en poudre. Après une semaine d'incubation, les cristaux sont figés dans l'azote liquide sans ajout de cryoprotectant et envoyés au synchrotron Soleil pour la collecte des données de diffraction des rayons X.

IV.2.2.Structure cristallographique du complexe PabAGOG/ICOA2-G5-P7

La qualité des cristaux obtenus suite à leur trempage en présence de l'inhibiteur ICOA2-G5-P7 a permis de collecter des jeux de données de diffraction exploitables pour une analyse cristallographique complète. Les données de diffraction ont été collectées au synchrotron SOLEIL sur la ligne PROXIMA-1. Ces données nous ont permis de résoudre la structure de pabAGOG à 1,25 Å par remplacement moléculaire en utilisant la structure de la protéine seule obtenue dans l'équipe (code 7OLB) (Figure 125).

Statistiques de la collecte de données	
	PabAGOG + G5-P7 (soaking)
Source de rayonnement	SOLEIL PROXIMA-1
Longueur d'onde (Å)	0.97857
Groupe d'espace	*P*1
Dimension de la maille: *a, b, c* (Å)	35.71, 35.74, 46.57
α, β, γ (°)	92.00, 95.25, 108.61
Gamme de résolution (Å)	33.80-1.25 (1.27-1.25)
Observation totales	383475 (18589)
Réflexions uniques	58877 (2887)
Complétude (%)	98.1 (96.0)
Multiplicité	6.4
R_{merge}[a] (%)	5.0 (30.6)
I/□ moyen (*I*)	18.8 (5.0)
$CC_{1/2}$ (%)	99.7 (96.0)
Affinement et statistiques du modèle	
Gamme de résolution (Å)	33.80-1.25
Nombre de réflections utilisées	58865
R_{work}[b] / R_{free}[c] (%)	18.0/19.1
Facteur B moyen (Å^2)	
Tous les atomes	19.83
Atomes de la protéine	18.61
Atomes de G5-P7	19.29
Atomes du PEG	31.70
Atomes du éthane-1,2-diol	26.27
Atomes de sodium	20.06
Atome de chlorure	19.50
Atome H_2O	29.78
RMSD par rapport à l'idéal	
Longueur des liaisons (Å)	0.006
Angles des liaisons (°)	0.786
Analyse ramachandran (% de résidus)	
Régions favorables / autorisées / interdites	97.9/2.1/0.0
Nombre d'atomes	
Protéine	1973
G5-P7	18
PEG	7
Ethane-1,2-diol	8
Sodium	1
Chlorure	1
Eau	228

[a] $R_{merge} = \sum_h \sum_i |I_{h,i} - \langle I \rangle_h| / \sum_h \sum_i I_{h,i}$ où $\langle I \rangle_h$ représente l'intensité moyenne des réflexions symétriques équivalentes.
[b] $R_{work} = \sum_h ||F_o| - |F_c|| / \sum_h |F_o|$, où F_o et F_c sont les amplitudes des facteurs de structures observés et calculés respectivement pour les réflexions h.
[c] R_{free} est la valeur R extradite de 5% des données de réflexion, qui ne sont pas inclues dans le processus d'affinement.

Figure 125 : Diffraction et statistiques d'affinement de PabAGOG lors du trempage d'ICOA2-G5-P7

Un examen fin de la carte de densité électronique du modèle indique sans ambiguïté une densité supplémentaire correspondant à l'inhibiteur complexé à la protéine (Figure 126). Cependant la structure de la molécule observée ne correspond pas exactement à la structure attendue de ICOA2-G5-P7. Ce composé a son atome de brome situé en **para** sur le cycle benzène, au lieu d'être sur le noyau barbiturique.

Evidemment, ce résultat était une surprise. Nous avons donc envoyé l'échantillon ICOA2-G5-P7 pour une analyse RMN et nous avons pu constater que le composé était très hétérogène (Annexe 10). Plusieurs composés apparentés à ICOA2-G5-P7 cohabitaient, avec le brome respectivement sur les positions para, méta, ortho, sans brome, ou avec le brome sur le noyau barbiturique (Figure 126). Sur la carte de densité électronique, il semblerait également que le composé avec le brome en **méta** puisse également être présent dans le site actif en faible proportion (Figure 126).

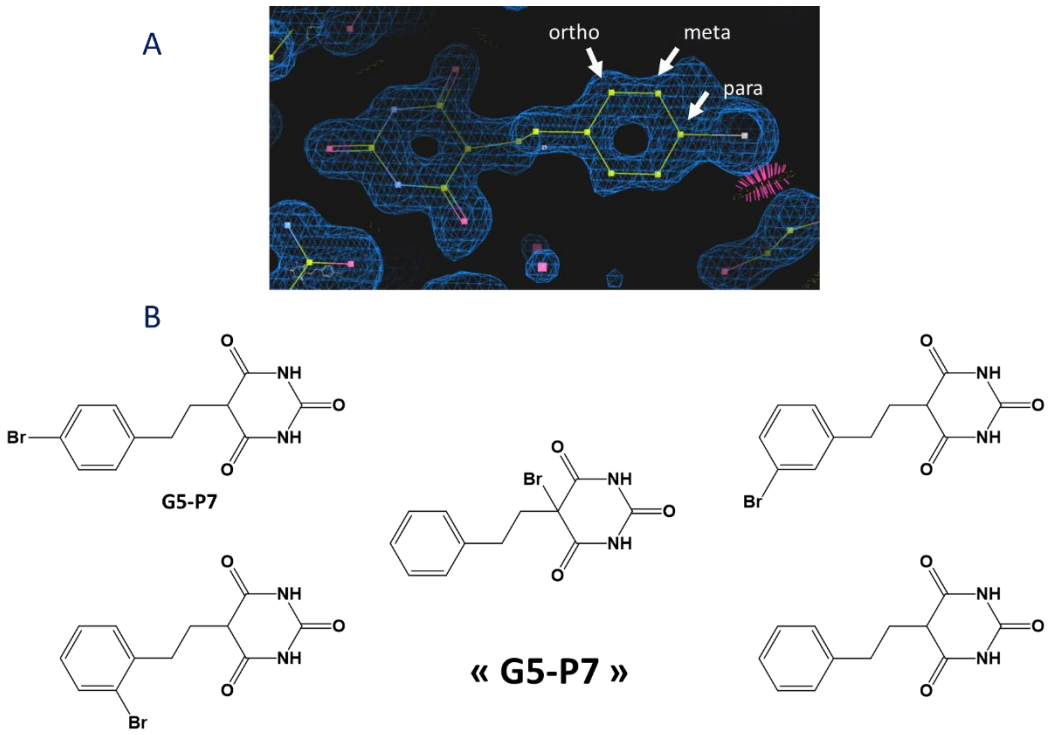


Figure 126 : Hétérogénéité structurale de l'inhibiteur ICOA2-G5-P7 - A) Carte de densité électronique 2fo-fc autour de la « G5-P7 » observé en cristallographie. B) Ces différentes structures ont été proposées par une analyse par chromatographie couplée à une détection par spectrométrie masse et une analyse par RMN du mélange.

En utilisant la structure publiée du complexe PaeAGOG/8-oxodG comme modèle [300], nous pouvons observer que ce composé est présent dans le site actif de l'enzyme PabAGOG (Figure 127). En effet, ces deux enzymes sont très proches structuralement avec un RMSD de 2,4 Å sur les Cα, et PaeAGOG est donc un bon modèle de comparaison. Cette molécule forme des liaisons hydrogène avec les résidus Asp208 et Gln24, des acides aminés également impliqués dans la reconnaissance de la 8-oxoG. Le noyau barbiturique de ce composé est aussi stabilisé par l'empilement des chaînes latérales aromatiques de la Phe146, du domaine HhH de AGOG, et le Trp212. Cette position de l'inhibiteur est très similaire à celle de la 8-oxodG dont le cycle purique dans le complexe PaeAGOG/8-oxodG est également pris en sandwich par la Phe146 et le Trp212 (Figure 127).

A

```
PabAGOG      -----GSHMIARIIGEIGIEGARFIEENIDEQFKALRYLSKGIDSETFVK
Pae_AGOG     MAAESQLKRVIETLRRLGIEEVLKLERR-DPQYRAVCNVVKRHGETVGSR

PabAGOG      LVIANSLVSYQLTGKGEQWWWEFAKYFYGRDVKSIYLAYKEFLPNSRFNR
Pae_AGOG     LAMLNALISYRLTGKGEEHWEYFGKYFSQLEVIDLCRDFLKYIETSPFLK

PabAGOG      RLIPQKLSRIRRVETFLSTLTEERIEEYYGDMSSLWGSIARALGVDKESK
Pae_AGOG     IGVEARKKRALKACDYV--P---NL-E---DLGLTLRQLSHIVGARREQK

PabAGOG      TVVFSVKMFGYAARIVLSTFNPYPMEIPIPEDSRIVKLTKK-LTNE----
Pae_AGOG     TLVFTIKILNYAYMCSRGVNRVLPFDIPIPVDYRVARLTWCAGLI-DFPP

PabAGOG      --------KPRKFWMKIARESGVPPLHIDSILWPLLGG-----ASIDSAP
Pae_AGOG     EEALRRYEAVQKIWDAVARETGIPPLHLDTLLWLAGRAVLYGENLH----

PabAGOG      PELRDK--LAELIKIIR----------
Pae_AGOG     ------GVPKEVIALFQWRGGCRPPSE
```

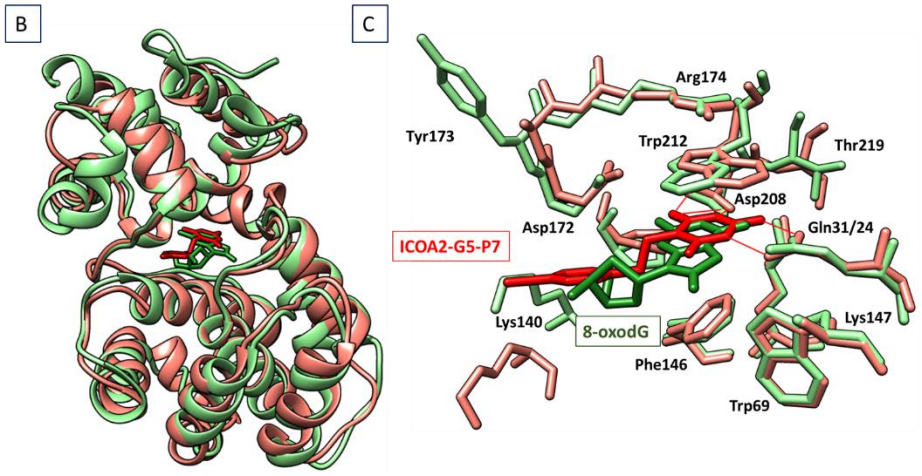


Figure 127 : Structure et séquence de PabAGOG et PaeAGOG - A) Alignement de séquence basé sur la structure de PabAGOG et PaeAGOG (1XQP). B) Superposition de PabAGOG (saumon) et PaeAGOG (vert clair). C) Superposition du site actif de PaeAGOG contenant la 8-oxodG (vert foncé) et du site actif de PabAGOG contenant le composé inhibiteur (rouge foncé). Les liaisons hydrogènes de la G5-P7 avec les résidus Asp208 et Gln24 sont en rouge.

Bien que ce résultat soit inattendu, c'est un résultat prometteur. En effet, suite à ce résultat nous avons demandé à nos collègues de l'ICOA de resynthétiser les différents composés potentiellement présents dans l'échantillon ICOA2-G5-P7 de la banque de diversité de l'ICOA, et tout particulièrement la forme observée dans le site actif de la structure cristallographique. Il est prévu d'évaluer la force inhibitrice de ces différents composés, et de relancer des études cristallographiques avec ceux-ci. En effet, les valeurs d'IC_{50} obtenues avec le composé G5-P7 sont en réalité celles d'un mélange. Nous pouvons donc raisonnablement faire l'hypothèse que nous sous-évaluons l'activité inhibitrice d'un de ces composés, potentiellement celui observé dans la structure cristallographique.

En superposant les sites actifs de hOGG1 complexée à une 8-oxodG et la PabAGOG complexée à ICOA2-G5-P7, nous pouvons proposer un modèle réaliste pour le complexe hOGG1/ICOA2-G5-P7 (Figure 128). En effet, les sites actifs sont très proches structuralement, les acides aminés interagissant avec ICOA2-G5-P7 chez PabAGOG ont des équivalent structuraux chez hOGG1. Les liaisons H formées entre l'inhibiteur et les résidus Asp208 et Gln24 de PabAGOG peuvent également se former avec les résidus équivalents Gln215 et Gly42 du site actif de hOGG1. De même la stabilisation de l'ICOA-G5-P7 dans le site actif par la Trp212 et Phe146 de PabAGOG semble également possible avec les résidus correspondant chez hOGG1,

respectivement Phe319 et Cys253 (Figure 128.C.D). Ces différents résidus sont directement reliés à la reconnaissance de la 8-oxoG, qui se superpose avec le noyau barbiturique de l'ICOA2-G5-P7 (Figure 128.A.B). En ce sens, nous pouvons émettre l'hypothèse que l'ICOA2-G5-P7 est également capable de s'insérer dans le site actif de hOGG1, et d'inhiber de façon compétitive l'activité 8-oxoG-ADN glycosylase de cette enzyme. Afin d'expliquer la différence d'efficacité de ICOA2-G5-P7 sur hOGG1 et PabAGOG, nous pouvons imaginer que le groupement bromobenzyle de l'inhibiteur puisse pivoter autour de la chaîne carbonée centrale pour venir interagir avec le résidu His270, permettant de stabiliser un peu plus l'inhibiteur dans le site actif, comme c'est le cas avec la TH5487 par exemple, ou la 8-oxodG (Figure 128.B.F.H). Cette hypothèse est renforcée par le fait que le groupement bromobenzyle est gêné par les résidus Gln249 et Asp268, et ne pourrait probablement pas être orienté de cette manière dans hOGG1. Etant en contact avec le solvant, sa position a donc de grande chance d'être mobile, et pourrait donc entrer en contact avec His270. Chez PabAGOG en revanche, l'Arg176 (situé à l'emplacement de His270) pointe dans une autre direction et n'interagit pas avec le bromobenzyle de ICOA2-G5-P7. Cependant il faut rappeler que ces intérprétations proviennent uniquement de la superposition des différentes structures cristallographiques. Afin de proposer un véritable modèle pour l'action de ICOA2-G5-P7 sur hOGG1, des études de docking ou de dynamique moléculaire pourrait être envisagées.

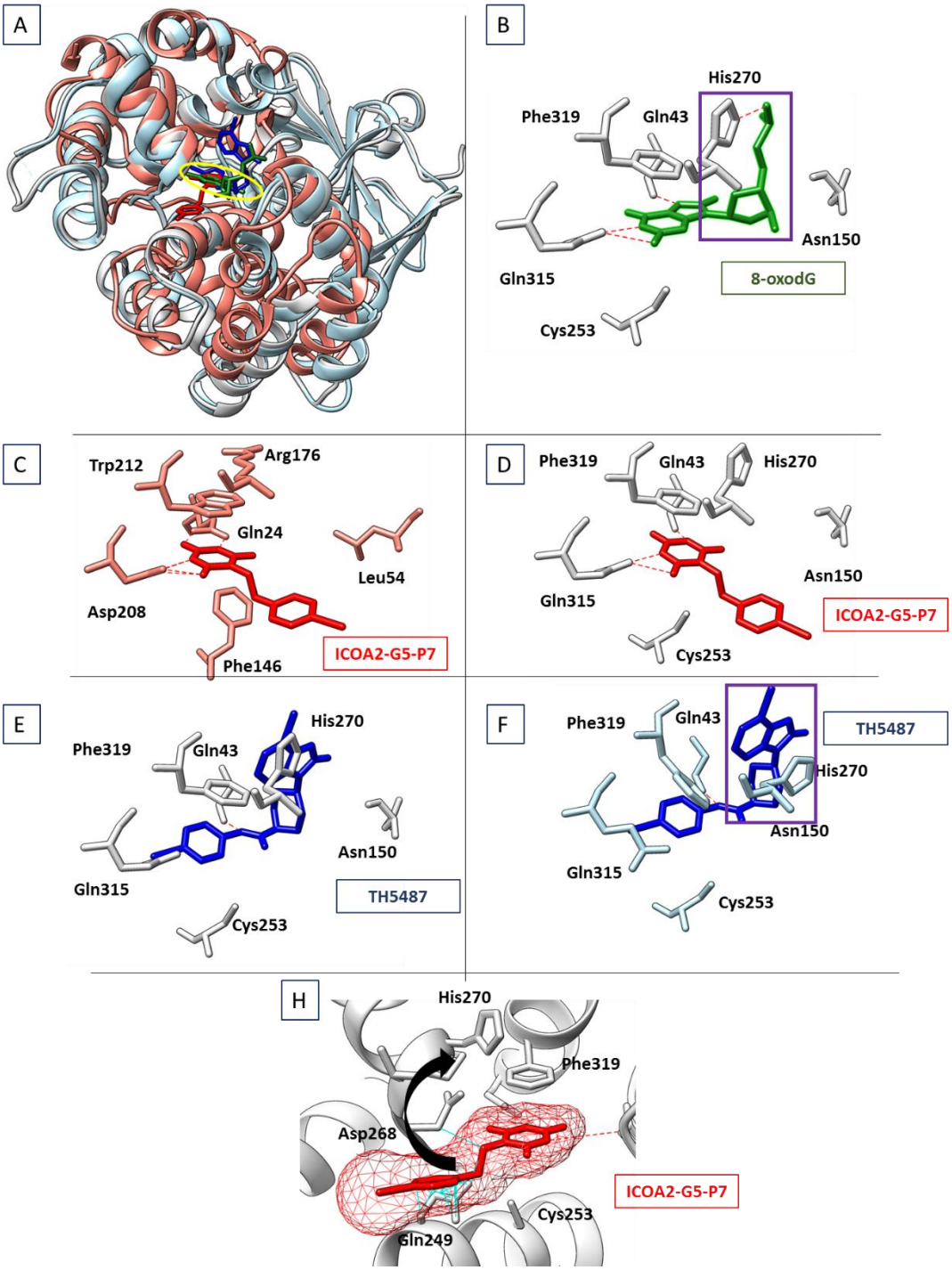


Figure 128 : Superposition des structures cristallographiques de PabAGOG/ICOA2-G5-P7 et de hOGG1/8-oxodG ou hOGG1/TH5487 - A) PabAGOG est indiquée en saumon, hOGG1/8-oxodG en gris (code PDB 1EBM) et hOGG1/TH5487 en bleu clair (code PDB 6RLW). ICOA2-G5-P7 est en rouge vif, TH5487 en bleu foncé et la 8-oxodG en vert foncé. Un cercle jaune indique l'emplacement de la partie purine dans le site actif. B) Intéraction de la 8-oxodG dans le site actif de hOGG1. Les liaisons H sont indiquées par des pointillés rouges et un carré violet indique la zone d'intéraction avec His270 C) Intéraction du composé ICOA2-G5-P7 avec le site actif de PabAGOG D) Superposition de ICOA2-G5-P7 dans le site actif de hOGG1 (grâce à la superposition) E Superposition de TH5487 dans la structure de hOGG1 seule F) Intéraction de TH5487 avec mOGG1. H) Surface de ICOA2-G5-P7 dans le site actif de hOGG1. Les clashs stériques sont indiqués en cyan, et une flèche noire indique une rotation potentielle d el'inhibiteur afin d'intéragir avec His270 et éviter les clashs.

V. Bilan et conclusion

L'objectif de cette partie de la thèse était d'identifier des inhibiteurs efficaces de hOGG1. Nous avons utilisé la même méthodologie de criblage que pour hNEIL1. Le criblage primaire utilisant un test d'activité fluorescent a permis d'identifier 70 molécules touches. Ces hits ont ensuite subi un test de confirmation, consistant en un test d'activité orthogonal de l'activité glycosylase-lyase et glycosylase de hOGG1, qui nous a permis de confirmer 24 de ces molécules. Enfin, un test d'intercalation a permis d'éliminer les composés intercalant. Les composés les moins efficaces ont également été mis de côté pour obtenir finalement 6 inhibiteurs efficaces de hOGG1. Parmi ces composés, on retrouve une famille de 3 molécules analogues, pouvant s'apparenter à des simili-purines. Ces composés sont les inhibiteurs les plus efficaces identifiés lors de ce crible. D'autres structures ont également été identifiées, notamment le composé ICOA2-G5-P7 possédant un noyau barbiturique.

Contrairement à hNEIL1, l'expansion des hits « à l'aveugle » de hOGG1 a permis d'optimiser et d'améliorer l'IC_{50} de nos composés, avec la molécule C-02G10, possédant l'IC_{50} submicromolaire le plus faible, soit 300 nM. Les inhibiteurs identifiés sont donc beaucoup plus efficaces que les inhibiteurs de hNEIL1. Cependant certains de ces composés, tout particulièrement la C-02G10, semblent déstabiliser fortement la protéine, ce qui a pu être un facteur particulièrement négatif pour les tentatives de cristallisation du complexe hOGG1/inhibiteur. Malheureusement, l'étude cinétique d'inhibition de hOGG1 n'a pas été concluante. De par une faible disponibilité des produits et la nécessité de congeler/décongeler plusieurs fois un même échantillon, il est fort probable que les molécules aient été dégradées au cours de l'étude. Il est donc envisageable de reprendre ces expériences et d'espérer des résultats plus positifs.

Les inhibiteurs identifiés sont aussi malheureusement peu solubles en solution aqueuse, et ce paramètre a été une contrainte durant l'ensemble de la thèse, particulièrement pour la cristallogenèse de hOGG1. Ainsi la majorité des essais de cristallisation en présence d'inhibiteurs comme C-02G10 ou ICOA2-H11-P29, ont précipité.

Ce résultat très frustrant nous a cependant indiqué la nécessité de trouver un modèle adapté pour contourner cette difficulté. Nous avons tout d'abord essayé l'enzyme murine mOGG1, dont les conditions de cristallisation étaient, elles aussi, publiées, mais nous n'avons obtenu aucun cristal. Finalement, nous avons opté pour l'homologue fonctionnel de hOGG1 chez l'archée *Pyrococcus abyssi* PabAGOG, une enzyme d'archée excisant la 8-oxoG, avec une séquence très différente de hOGG1, mais une structure générale du corps catalytique très proche. Seuls 2 inhibiteurs de la hOGG1 ont montré des effets sur PabAGOG : C-02G10 et ICOA2-G5-P7. Comme pour hOGG1, les essais de cristallogenèse avec C-02G10 n'ont pas abouti. En revanche, des cristaux ont été obtenus en présence de ICOA2-G5-P7 et ont montré une qualité de diffraction des rayons X permettant une étude cristallographique complète. La structure tridimensionnelle du complexe PabAGOG/ ICOA2G5-P7 montre que l'inhibiteur occupe le site actif de l'enzyme. De manière surprenante, le composé ICOA2-G5-P7 n'était pas un produit de synthèse pur, mais un mélange de différentes isoformes où la position du brome changeait. Le composé observé avec un brome en para du cycle du benzène, majoritairement retrouvé dans le site actif, est donc probablement un produit secondaire de réaction. Cela pourrait donc indiquer que la valeur d'IC_{50} de ce composé soit sous-estimée par rapport à celui du « mélange » ICOA2-G5-P7.

Finalement, nous avons identifié des composés parmi les plus efficaces à ce jour sur hOGG1, et mis en évidence une famille structurale d'inhibiteurs. La structure de PabAGOG /ICOA2-G5-$P7_{para}$ est un résultat particulièrement encourageant, et qui permet pour la première fois d'observer le site d'interaction d'un de nos inhibiteurs sur une ADN glycosylase. Cette structure permet de proposer un modèle pour l'interaction entre hOGG1 et ICOA2-G5-$P7_{para}$. L'évaluation de nos inhibiteurs sur d'autres ADN glycosylases humaines permettrait également d'étoffer la caractérisation de ces composés.

Conclusion et perspectives

Les ADN glycosylases sont des enzymes impliquées dans l'élimination des bases nucléiques endommagées et contribuent ainsi à la stabilité génétique en évitant l'apparition de mutations ou de processus de létalité. Elles sont présentes dans l'immense majorité des espèces vivantes connues, des mammifères aux procaryotes en passant même par certains virus/virus géants et bactériophages. Pourtant, ce sont deux de ces enzymes humaines que nous proposons de cibler pharmacologiquement par la recherche d'inhibiteurs spécifiques: hNEIL1, spécialisée dans la réparation des pyrimidines oxydées, et hOGG1, spécialisée dans la réparation des purines oxydées et plus particulièrement de la 8-oxoguanine, le produit majoritaire d'oxydation des purines. Cet apparent paradoxe découle en réalité du concept de létalité synthétique, et de la complexité des mécanismes biologiques dans lesquels ces enzymes sont impliquées. En effet, bien que bénéfiques dans la plupart des cas, ces enzymes peuvent cependant être impliquées dans le déclenchement de certaines maladies neurodégénératives comme la maladie d'Huntington par exemple[353]. De plus, la lutte contre le cancer repose majoritairement sur trois bras armés : la chirurgie, la chimiothérapie et la radiothérapie. Or, deux de ces stratégies, la radio et la chimiothérapie, ont souvent pour but de créer des lésions dans l'ADN des cellules malades afin d'induire leur mort cellulaire. C'est dans ce contexte que les ADN glycosylases peuvent fournir une protection non désirée aux cellules cancéreuses que l'on cherche à éliminer. Les cellules cancéreuses découlant inévitablement d'une ou plusieurs mutations dans un ou plusieurs gènes, ces mutations peuvent être leur « talon d'Achille » et les rendre dépendantes d'une ADN glycosylase en particulier. Ainsi, ces protéines deviennent des cibles particulièrement intéressantes, car leur ciblage devrait avoir peu d'impact sur les cellules saines, celles-ci possédant d'autres systèmes biologiques pour survivre. Identifier des inhibiteurs sélectifs et efficaces, ainsi que comprendre leur mode d'interaction avec ces différentes enzymes afin de les optimiser par la suite, permettrait donc de disposer d'outils potentiels supplémentaires pour la lutte contre le cancer. Ces inhibiteurs constituent aussi des outils très intéressant pour la recherche fondamentale. C'est dans cette recherche combinant criblage moléculaire, caractérisation biochimique et études structurales, que s'inscrit mon travail de thèse.

Afin d'identifier des inhibiteurs des enzymes hNEIL1 et hOGG1, nous avons tout d'abord mis en place une procédure de criblage semi-automatique à moyen débit que nous avons validée sur l'enzyme bactérienne LlFpg, un homologue fonctionnel de hOGG1 et structural de hNEIL1. Suite à la mise en place de cette procédure pilote, 4560 molécules provenant de banques chimiques diverses et choisies pour leur diversité et leur disponibilité (produits naturels, de synthèses, médicaments hors brevet, etc.) ont été criblées pour leur capacité à moduler l'activité de hOGG1 ou hNEIL1. Les criblages à haut débit sont principalement l'apanage des industries pharmaceutiques ou des laboratoires disposant de plusieurs dizaines de personnes exclusivement dédiées à cette tâche. Ici, il est important de noter que ce criblage est le premier à avoir été mis en place dans notre équipe. Ainsi, l'un des principaux défis que nous avons eu à relever lors de cette thèse, était d'effectuer un criblage robuste et d'identifier des composés originaux, avec des moyens limités et dans un temps limité tout en conservant les différentes étapes et les points critiques de tout criblage. La procédure de criblage est donc perfectible et a été optimisée tout au long de cette thèse.

Malgré toutes les difficultés rencontrées, nous avons pu identifier des inhibiteurs efficaces sur nos deux protéines d'intérêt. Pour hNEIL1, quatre inhibiteurs efficaces ont été identifiés, dont l'un semble être l'inhibiteur le plus efficace identifié à ce jour : la **girgensonine**. Pour hOGG1, une dizaine de molécules inhibitrices ont été sélectionnées suite à la procédure de criblage, dont certaines présentent des motifs structuraux similaires. Ces composés ont pu être optimisés avec des étapes dites de « hits expansion », et ont permis d'identifier l'un des inhibiteurs de hOGG1 les plus efficaces : le composé **C-02G10**. D'autres composés de structures chimiques différentes ont également été identifiés, notamment le composé ICOA2-G5-P7.

La suite du criblage a consisté à tenter d'élucider au mieux le mode d'action de nos inhibiteurs sur hOGG1 et hNEIL1. Nous nous sommes heurtés à des des comportements enzymatiques atypiques sortant des modèles connus, mais aussi à des propriétés physico-chimiques parfois incompatibles avec l'analyse biochimique et enzymologique. Nous avons par exemple pu identifier le groupement nitrile de certains inhibiteurs comme essentiel à l'inhibition de hNEIL1 et potentiellement capable d'induire une inhibition irréversible de l'enzyme suite à sa fixation covalente sur une ou plusieurs cystéine de l'enzyme. Malheureusement, les études de cinétiques enzymatiques n'ont pas pu aboutir car hNEIL1 ne présente pas un comportement micahélien. Pour hOGG1, nos molécules semblent aussi former une classe bien différente d'inhibiteurs en comparaison des composés identifiés récement dans la littérature, Th5487 et Su268. Alors que Th5487 et Su268 semblent stabiliser hOGG1, nos inhibiteurs semblent plutôt la déstabiliser. Enfin, des études préliminaires conduites dans l'équipe et étudiant l'effet des inhibiteurs sur l'interaction de hOGG1 avec l'ADN indique que nos inhibiteurs ne déplacent pas l'enzyme d'une molécule d'ADN contenant une 8-oxoG contrairement à Th5487 et Su268. De façon intéressante, nos inhibiteurs et Th5487 et Su268 sont incapables de déplacer hOGG1 d'une molécule d'ADN contenant un site abasique. Cette propriété pourrait être intéressante à exploiter dans l'avenir. Tout cela indique que le mécansime d'inhibition de hOGG1 par nos inhibiteurs est très différent de celui de Th5487 et Su268. Nos inhibiteurs permettent donc d'étoffer la palette de composés thérapeutiques potentiels, en présentant une efficacité similaire, voire supérieure, et un mécanisme d'inhibition très différent.

Outre l'identification d'inhibiteurs originaux, l'un des objectifs principaux de cette thèse, était d'étudier les bases structurales de l'inhibition de hOGG1 et hNEIL1 en conduisant des études cristallographiques systématiques des complexes enzyme/inhibiteur. Pour hNEIL1, nous avons obtenu une structure de la protéine seule avec une résolution de 1,45 Å (résolution meilleure que les modèles cristallographiques connus dans la PDB), qui correspond aux structures déjà existantes dans la PDB. En revanche, nous ne sommes pas parvenus à observer nos inhibiteurs dans ces cristaux. Nous avons cherché en parallèle à collecter un maximum d'informations structurales avec notamment une étude de la formation d'un adduit covalent avec la girgensonine suivi par spectrométrie de masse. Si la formation de cet adduit sur l'enzyme n'est raisonnablement pas l'unique origine de l'inhibition observée, il donne cependant des informations sur le(s) site(s) potentiel(s) d'interaction de nos inhibiteurs à la surface de hNEIL1. Il pourrait être intéressant d'évaluer la possibilité de formation d'adduits de la girgensonine ou de ICOA1-D5-P2 sur d'autres enzymes de la même famille structurale ne possédant pas de cystéine (ou de cystéine libre), comme mvNei et LlFpg pour écarter toute autre pontage covalent avec un autre acide aminé que la cystéine. De plus les conditions de cristallisation de ces 2 enzymes sont maitrisées dans l'équipe. Pour hOGG1 et malgré nos effords pour y parvenir, aucun cristal n'a été obtenu, et ce malgré l'utilisation d'une forme tronquée et de l'enzyme murine, dont les structures sont pourtant déjà publiées. Même si ce résultat est aussi étonnant que frustrant, il n'est pas rare en cristallographie que des résultats ne soient pas répétables d'un laboratoire à un autre ou que certains détails expérimentaux soient omis ne permettant pas de reproduire l'expérience. Par exemple nos conditions de cristallisation de hNEIL1 nous ayant permis de résoudre sa structure cristallographique sont significativement différentes de celles ayant permis de résoudre les structures de hNEIL1 déjà publiées. Cependant, pour contourner cette difficulté, nous avons effectué l'étude cristallographique d'une enzyme homologue structurale et fonctionnelle de hOGG1, la 8-oxoG-ADN glycosylase archéenne PabAGOG. La structure cristallographique de PabAGOG/ICOA2-G5-P7 montre dans le site actif de l'enzyme un (ou plusieurs) isomère(s) de position pour un substituant brome sur le cycle benzénique de ICOA2-G5-P7. Le site actif de hOGG1 et PabAGOG étant assez proche, la structure de PabAGOG/ICOA2-G5-P7 nous a permis de proposer un modèle du mode d'interaction de cet inhibiteur dans le site actif de hOGG1. Ce premier modèle peut nous guider dans la synthèse de nouveaux composés plus

affins de l'enzyme. Il est également prévu d'approfondir les études sur ce composé en synthétisant les isomères purs et en réévaluant leur activité sur PabAGOG et hOGG1.

Finalement l'ensemble de ces résultats a permis l'identification et la caractérisation de molécules novatrices comme inhibiteurs de hNEIL1 et hOGG1, dont les structures chimiques sont différentes de celles des inhibiteurs déjà connus, élargissant ainsi les possibilités thérapeutiques, tout en apportant des pistes sur le mécanisme d'inhibition de ceux-ci.

La thèse, comme tout travail de recherche, est un projet s'inscrivant sur un temps long, et permet souvent d'ouvrir de nouvelles pistes pour des projets futurs. Ainsi plusieurs pistes mériteraient d'être explorées suite à ce travail.

4 560 petites molécules synthètiques ou naturelles ont été criblées lors de ma thèse. Cela peut paraitre beaucoup à l'échelle de l'équipe de recherche mais, en définitif et malgré une certaine diversité structurale des composés criblés, cette étude reste modeste au regard de ce qui se fait en industrie ou dans des laboratoires spécialisés. Par exemple, la découverte du Su268 résulte d'un criblage primaire de 26 000 composés de synthèse tandis que le composé Th5487 provient d'un criblage primaire de 17 490 composés soit 4 à 6 fois plus de molécules que le notre [370,371]. Il faut rappeler que ces études ont bénéficié également de l'accès à des plateformes de criblage automatisées ce qui permet d'aller beaucoup plus vite (à noter qu'au départ de cette thèse, ces molécules n'étaient pas encore identifiées). Un autre élément non négligeable à prendre en compte est le coût de ces banques de molécules. Pour notre part, nous avons eu accès gratuitement, à la banque de diversité de nos collègues de l'Institut de Chimie Organique et Analytique de l'Université d'Orléans. Une collègue du laboratoire, nous a également permis d'avoir accès à la banque commerciale Prestwick. A titre indicatif, la seule banque de GreenPharma a couté à mon équipe 20 000 € soit près de 45 €/molécules dont le stock est prévu pour seulement 2 à 3 criblages. Loin d'être triviales, ces considérations sont des facteurs capitaux à prendre en compte dans un projet de recherche académique actuel, et m'ont permis de me rendre compte de l'importance des soutiens financier publics. Pour être à la hauteur de la compétition internationale, il faudrait avoir accès par exemple à la Chimiothèque Nationale (banque de diversité de plus de 70 000 composés pour un coût d'environ 0.3€ /molécule livrée sur plaques de 96 ou 384 puits). Néanmoins, pour un si grand nombre de molécules, un criblage semi-automatique ne sera pas suffisant pour une telle quantité de composés. Pour un tel criblage, il faut un criblage primaire robuste, ce qui est notre cas, mais également une plateforme automatisée et un buget dédié, ce dont mon équipe ne dispose pas. Malgré tout, il faut noter que la diversité des différentes banques utilisées par nous et les autres équipes, est suffisamment importante pour que trois criblages sur une même protéine produisent 3 inhibiteurs absolument pas reliés structuralement. En cela, notre stratégie s'est révélée payante.

Lors de notre criblage, nous avons éliminé les composants intercalants, préférant privilégier les composés n'interagissant pas avec l'ADN, et donc moins susceptibles d'être toxiques. De plus, les composés intercalants étaient pour la plupart non spécifiques entre nos deux enzymes d'intérêt. En revanche, certains d'entre eux, montraient une spécificité à l'une ou l'autre des ADN glycosylases, tout en étant fortement intercalant. Ainsi, étudier plus en profondeur ces composés pourrait permettre d'identifier des composés intéressants. Nous pourrions par exemple envisager que l'effet de ces composés se fasse à l'interface entre l'ADN et la protéine, comme c'est le cas par exemple de certains inhibiteurs de topoisomérases [473,474].

Une autre piste qui nous semble intéressante à explorer est d'étudier l'effet de la stéréochimie sur le pouvoir inhibiteur de certains composés que nous avons sélectionnés. En effet, il faut garder à l'esprit par exemple que les composés de la banque de GreenPharma sont fournis sous forme de mélanges racémiques.

Les effets d'un composé peuvent varier selon son énantiomérie. Par exemple, le crizonitib est un inhibiteur de MTH1 sous sa configuration (S), mais un inhibiteur de tyrosine kinase sous sa configuration (R) [259]. La girgensonine possédant un centre asymétrique, il est donc possible qu'un seul des deux énantiomères possède une activité sur hNEIL1, ce qui pourrait indiquer que nous avons sous-estimé son IC_{50} à partir du mélange racémique. Une stucture cristallographique du complexe hNEIL1/girgensonine aurait pu être très informative et indiquer un énantiomère préférentiel.

Ne pas obtenir de structure cristallographique avec nos inhibiteurs sur hNEIL1 a également été un des objectifs manqués de cette thèse. Cependant, tout comme l'utilisation de PabAGOG a pu fournir un modèle intéressant pour un de nos inhibiteurs, nous pouvons envisager l'utilisation d'autres modèles pour les inhibiteurs de hNEIL1. Par exemple, le composé LA-32, qui possède une activité similaire sur l'enzyme bactérienne LlFpg, pourrait suggérer la formation d'un complexe LA-32/LlFpg/ADN-THF. Cette perspective est renforcée par des résultats récents obtenus dans l'équipe montrant que ce composé ne déplace pas le complexe hNEIL1/ADN, et pourrait au contraire le renforcer.

Afin d'identifier de nouveaux composés ayant des structures originales et inconnues, nous pourrions considérer le développement de nouvelles méthodes de criblage de mélanges complexes de composés naturels. A ce propos, la Chimiotèque Nationale propose 15 000 extraits naturels originaux. Ainsi, nous pourrions envisager de fixer notre protéine d'intérêt sur une colonne de chromatographie classique, puis de faire passer différents extraits naturels sur cette colonne. Les composés présentant une forte affinité pour l'enzyme immobilisée pourraient donc être sélectionés par cette colonne d'affinité. Après différentes étapes de lavage, l'enzyme immobilisée de façon non irréversible sur la colonne et compléxée à ces composés de haute affinité nous permettrait de récupérer des composés naturels originaux et potentiellement des inhibiteurs. Cependant la séparation de ces potentiels composés nécessiterait une séparation chromatographique et une analyse fine par spectrométrie de masse et RMN pour clarifier leurs structures. Automatiser un tel dispositif semble complexe, mais pourrait être envisagé.

Dans le cadre d'une approche thérapeutique, un inhibiteur performant sélectionné sur une cible purifiée est encore bien loin de sa validation dans un contexte plus complexe telle que la cellule. En effet, de nombreuses molécules très actives sur la cible isolée sont finalement abandonnées à cause d'effets dit indésirables sur d'autres composés cellulaires (effets « off target »). Les exemples sont nombreux mais c'est notamment le cas du composé SCR7, un inhibiteur de l'ADN Ligase IV, utilisé comme inhibiteur spécifique du système NHEJ. Malheureusement, des effets « off target » de cette molécule induise une toxicité cellulaire élevée et rendent ce composé obsolète pour des applications thérapeutiques [475]. De fait, de nouvelles études doivent être mise en place afin d'optimiser l'inhibition de l'ADN Ligase IV. De même, le composé RI-1, un inhibiteur très efficace de Rad51, formant des liaisons covalentes avec le groupement thiol d'une cystéine de cette protéine, a nécessité le développement d'un nouveau composé RI-2 afin d'éviter les nombreux effets « off target » potentiels [476]. Une façon d'éviter cet écueil est de disposer d'un test de criblage spécifique permettant de mesurer l'activité d'une ADN glycosylase directement dans des cellules. Dans le cas de hOGG1, la mise en place d'une telle procédure est à l'étude. L'intérêt ici est de sélectionner des molécules inhibitrices qui passent sans problème les membranes cellulaires. L'un des problèmes majeurs rencontrés lors de cette thèse est la faible solubilité des inhibiteurs identifiés, principalement pour les inhibiteurs de hOGG1. Cette faible solubilité a posé des soucis lors de nombreuses expériences, et peut expliquer certains résultats non concluants. En revanche, la faible solubilité n'est pas nécessairement un frein incontournable pour des applications thérapeutiques. Bien que mon équipe ne puisse mettre en place cette méthode, il faut rappeler que l'utilisation de nano particules pour encapsuler des composés et les délivrer à

la cible biologique désirée est une stratégie prometteuse, comme le montrent des études récentes sur certains inhibiteurs des systèmes de réparation [188,270].

D'autres pistes existent pour s'affranchir des petites molécules inhibitrices dont la faible spécificité et sélectivité est souvent un obstacle pour le développement d'un nouveau médicament. L'une de ces approches consiste à concevoir des leurres moléculaires très proches du substrat naturel de l'enzyme. Dans le cas des ADN glycosylases, les études structurales par cristallographie ont souvent nécessité la conception et la synthèse de petits fragments d'ADN contenant des analogues de bases lésées reconnues spécifiquement par les enzymes sans pour autant être métabolisables, comme nous l'avons évoqué en introduction de cette thèse. Mon équipe d'accueil a participé activement au développement de ce type d'approches pour décrypter par exemple les déterminants structuraux et fonctionnels impliqués dans la reconnaissance et dans l'élimination des purines oxydées par la 8-oxoG-ADN glycosylase bactérienne Fpg [385,389,477–479]. Outre leur spécificité/sélectivité et leur très forte affinité (présentant souvent un K_D de l'ordre du subnanomolaire) avec l'ADN glycosylase, ils se comportent comme des inhibiteurs très performants de ces enzymes et sont de vrais compétiteurs du substrat ADN. L'exploitation de ces connaissances initialement dédiées à la recherche fondamentale est à l'étude dans mon équipe pour le développement d'une nouvelle classe d'inhibiteurs basée sur des duplex d'ADN agissant comme des leurres moléculaires pour les ADN glycosylases. On parle souvent d'ADN interférents (ADNi). Ces composés ont également l'avantage de ne pas présenter de problèmes de solubilité. La difficulté avec les ADNi réside essentiellement dans le passage au travers des membranes biologiques. Il existe de nombreuses stratégies pour contourner cette difficulté. Mon équipe d'accueil propose de conjuguer les ADNi à un partenaire (peptide/cholestérol) lui permettant de franchir aisément les membranes biologiques et si possible de le délivrer dans le noyau cellulaire et/ou dans la mitochondrie, les 2 compartiments cellulaires dans lesquels hOGG1 et hNEIL1 sont présentes. A titre d'exemple de ce type de stratégie, nous pouvons citer le DBait, une molécule oligonucléotidique en phase de développement cliniques et agissant sur les PARPs et sur les systèmes de réparation des cassures double-brin, qui montre des résultats très prometteurs dans les thérapies anticancéreuses [153–155]. Une autre piste que souhaite développer mon équipe est le ciblage de hOGG1 et hNEIL1 par des intrabodies (nanobodies vectorisés à l'intérieur des cellules). Ces nanobodies sont des fragments d'anticorps à chaîne unique de camélidés contenant uniquement la partie variable de l'anticorps. Cette stratégie est un domaine en plein développement particulièrement en cancérologie mais aucun travaux publié à ce jour ne propose de cibler les ADN glycosylases [480,481] Dans le domaine de la réparation de l'ADN, cette stratégie a été surtout développée pour suivre en temps réel des protéines comme l'histone γ-H2AX et la PARP1 dans les cellules vivantes par exemple[482,483]. Utiliser des nanobodies neutralisant l'activité de hOGG1 ou hNEIL1 couplés à de courts peptides pour assurer le passage dans les cellules sera prochainement développé dans l'équipe. Cela pourrait permettre d'obtenir des inhibiteurs très sélectifs et spécifiques sans effet « off target ».

Mon travail de thèse s'inscrit donc dans un projet plus global et a permis d'identifier de nouveaux inhibiteurs sur hOGG1 et hNEIL1. Cela m'a également permis de mettre en perspective la recherche thérapeutique et d'apprécier les limites et les difficultés de travailler avec des petites molécules. Ce travail a été très formateur sur un sujet où de nombreux aspects restent à découvrir et à développer.

Références bibliographiques

[1] P. A. Gammage, C. Frezza, *BMC Biology* **2019**, *17*, 53.
[2] S. Boiteux, M. Guillet, *DNA Repair* **2004**, *3*, 1.
[3] S. Boiteux, J. Laval, *Biochemistry* **1982**, *21*, 6746.
[4] N. H. Williams, in *Artificial Nucleases*, (Ed: M. A. Zenkova), Springer, Berlin, Heidelberg, **2004**, pp. 3–17.
[5] T. Lindahl, *Nature* **1993**, *362*, 709.
[6] C. A. Lewis, J. Crayle, S. Zhou, R. Swanstrom, R. Wolfenden, *PNAS* **2016**, *113*, 8194.
[7] R. De Bont, N. van Larebeke, *Mutagenesis* **2004**, *19*, 169.
[8] J. Krwawicz, K. D. Arczewska, E. Speina, A. Maciejewska, E. Grzesiuk, *Acta Biochim Pol* **2007**, *54*, 413.
[9] D. Fu, J. A. Calvo, L. D. Samson, *Nat Rev Cancer* **2012**, *12*, 104.
[10] M. Dizdaroglu, G. Kirkali, P. Jaruga, *Free Radic Biol Med* **2008**, *45*, 1610.
[11] F. Drabløs, E. Feyzi, P. A. Aas, C. B. Vaagbø, B. Kavli, M. S. Bratlie, J. Peña-Diaz, M. Otterlei, G. Slupphaug, H. E. Krokan, *DNA Repair (Amst)* **2004**, *3*, 1389.
[12] N. Shrivastav, D. Li, J. M. Essigmann, *Carcinogenesis* **2010**, *31*, 59.
[13] L. L. Gerchman, D. B. Ludlum, *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis* **1973**, *308*, 310.
[14] T. Nikolova, W. P. Roos, O. H. Krämer, H. M. Strik, B. Kaina, *Biochim Biophys Acta Rev Cancer* **2017**, *1868*, 29.
[15] B. Singer, S. J. Spengler, H. Fraenkel-Conrat, J. T. Kuśmierek, *Proc Natl Acad Sci U S A* **1986**, *83*, 28.
[16] G. S. More, A. B. Thomas, S. S. Chitlange, R. K. Nanda, R. L. Gajbhiye, *Anticancer Agents Med Chem* **2019**, *19*, 1080.
[17] S. Dasari, P. B. Tchounwou, *Eur J Pharmacol* **2014**, *740*, 364.
[18] F. Jin, J. Thaiparambil, S. R. Donepudi, V. Vantaku, D. W. B. Piyarathna, S. Maity, R. Krishnapuram, V. Putluri, F. Gu, P. Purwaha, S. K. Bhowmik, C. R. Ambati, F.-C. von Rundstedt, F. Roghmann, S. Berg, J. Noldus, K. Rajapakshe, D. Gödde, S. Roth, S. Störkel, S. Degener, G. Michailidis, B. A. Kaipparettu, B. Karanam, M. K. Terris, S. M. Kavuri, S. P. Lerner, F. Kheradmand, C. Coarfa, A. Sreekumar, Y. Lotan, R. El-Zein, N. Putluri, *Cancer Prev Res (Phila)* **2017**, *10*, 588.
[19] F. Javed, S. V. Kellesarian, I. K. Sundar, G. E. Romanos, I. Rahman, *Oral Dis* **2017**, *23*, 1052.
[20] H. Esterbauer, P. Eckl, A. Ortner, *Mutat Res* **1990**, *238*, 223.
[21] B. Tudek, D. Zdżalik-Bielecka, A. Tudek, K. Kosicki, A. Fabisiewicz, E. Speina, *Free Radic Biol Med* **2017**, *107*, 77.
[22] C. Migdal, M. Serres, *Med Sci (Paris)* **2011**, *27*, 405.
[23] S. Boiteux, F. Coste, B. Castaing, *Free Radical Biology and Medicine* **2017**, *107*, 179.
[24] R. Radi, *Proc Natl Acad Sci U S A* **2018**, *115*, 5839.
[25] J. A. Imlay, S. M. Chin, S. Linn, *Science* **1988**, *240*, 640.
[26] A. L. Jackson, L. A. Loeb, *Mutat Res* **2001**, *477*, 7.
[27] C. C. Faria, M. S. Peixoto, D. P. Carvalho, R. S. Fortunato, *Oxid Med Cell Longev* **2019**, *2019*, 2514312.
[28] W. L. Santivasi, F. Xia, *Antioxid Redox Signal* **2014**, *21*, 251.
[29] J. P. Pouget, T. Douki, M. J. Richard, J. Cadet, *Chem Res Toxicol* **2000**, *13*, 541.
[30] M. Gardès-Albert, *Annales Pharmaceutiques Françaises* **2006**, *64*, 365.
[31] M. Dizdaroglu, *Cancer Letters* **2012**, *327*, 26.
[32] J. Cadet, S. Loft, R. Olinski, M. D. Evans, K. Bialkowski, J. Richard Wagner, P. C. Dedon, P. Møller, M. M. Greenberg, M. S. Cooke, *Free Radical Research* **2012**, *46*, 367.
[33] F. Faucher, S. Doublié, Z. Jia, *Int J Mol Sci* **2012**, *13*, 6711.
[34] E. D. Pleasance, R. K. Cheetham, P. J. Stephens, D. J. McBride, S. J. Humphray, C. D. Greenman, I. Varela, M.-L. Lin, G. R. Ordóñez, G. R. Bignell, K. Ye, J. Alipaz, M. J. Bauer, D. Beare, A. Butler, R. J. Carter, L. Chen, A. J. Cox, S. Edkins, P. I. Kokko-Gonzales, N. A. Gormley, R. J. Grocock, C. D. Haudenschild, M. M. Hims, T. James, M. Jia, Z. Kingsbury, C. Leroy, J. Marshall, A. Menzies, L. J. Mudie, Z. Ning, T. Royce, O. B. Schulz-Trieglaff, A. Spiridou, L. A. Stebbings, L. Szajkowski, J. Teague, D. Williamson, L. Chin, M. T. Ross, P. J. Campbell, D. R. Bentley, P. A. Futreal, M. R. Stratton, *Nature* **2010**, *463*, 191.
[35] E. C. Friedberg, L. B. Meira, *DNA Repair (Amst)* **2006**, *5*, 189.

[36] T. Lindahl, D. E. Barnes, *Cold Spring Harb Symp Quant Biol* **2000**, *65*, 127.
[37] M. A. Kalam, K. Haraguchi, S. Chandani, E. L. Loechler, M. Moriya, M. M. Greenberg, A. K. Basu, *Nucleic Acids Res* **2006**, *34*, 2305.
[38] J. Cadet, K. J. A. Davies, *Free Radic Biol Med* **2017**, *107*, 2.
[39] S. Bjelland, E. Seeberg, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **2003**, *531*, 37.
[40] R. A. Forrest, L. P. Swift, A. Rephaeli, A. Nudelman, K.-I. Kimura, D. R. Phillips, S. M. Cutts, *Biochem Pharmacol* **2012**, *83*, 1602.
[41] J. Hay, S. Shahzeidi, G. Laurent, *Arch Toxicol* **1991**, *65*, 81.
[42] A. Ciccia, S. J. Elledge, *Molecular Cell* **2010**, *40*, 179.
[43] A. N. Blackford, S. P. Jackson, *Molecular Cell* **2017**, *66*, 801.
[44] B. Verbeek, T. D. Southgate, D. E. Gilham, G. P. Margison, *Br Med Bull* **2008**, *85*, 17.
[45] B. Sedgwick, *Nat Rev Mol Cell Biol* **2004**, *5*, 148.
[46] M. Zhang, L. Wang, D. Zhong, *Photochem Photobiol* **2017**, *93*, 78.
[47] J. Yélamos, V. Schreiber, F. Dantzer, *Trends Mol Med* **2008**, *14*, 169.
[48] C. Flohr, A. Bürkle, J. P. Radicella, B. Epe, *Nucleic Acids Res* **2003**, *31*, 5332.
[49] E. E. Alemasova, O. I. Lavrik, *Nucleic Acids Res* **2019**, *47*, 3811.
[50] J.-P. Gagné, M. Isabelle, K. S. Lo, S. Bourassa, M. J. Hendzel, V. L. Dawson, T. M. Dawson, G. G. Poirier, *Nucleic Acids Res* **2008**, *36*, 6959.
[51] M. Javle, N. J. Curtin, *Br J Cancer* **2011**, *105*, 1114.
[52] M. Isabelle, X. Moreel, J.-P. Gagné, M. Rouleau, C. Ethier, P. Gagné, M. J. Hendzel, G. G. Poirier, *Proteome Sci* **2010**, *8*, 22.
[53] A. P. Shah, C. N. Patel, D. K. Sureja, K. P. Sanghavi, *Folia Med (Plovdiv)* **2018**, *60*, 39.
[54] J. R. Chapman, M. R. G. Taylor, S. J. Boulton, *Mol. Cell* **2012**, *47*, 497.
[55] E. Fanning, V. Klimovich, A. R. Nager, *Nucleic Acids Res.* **2006**, *34*, 4126.
[56] W. D. Wright, S. S. Shah, W.-D. Heyer, *J Biol Chem* **2018**, *293*, 10524.
[57] R. Scully, A. Panday, R. Elango, N. A. Willis, *Nat Rev Mol Cell Biol* **2019**, *20*, 698.
[58] A. J. Davis, B. P. C. Chen, D. J. Chen, *DNA Repair* **2014**, *17*, 21.
[59] H. H. Y. Chang, N. R. Pannunzio, N. Adachi, M. R. Lieber, *Nat Rev Mol Cell Biol* **2017**, *18*, 495.
[60] S. Espejel, M. Martín, P. Klatt, J. Martín-Caballero, J. M. Flores, M. A. Blasco, *EMBO Rep* **2004**, *5*, 503.
[61] G.-M. Li, *Cell Res* **2008**, *18*, 85.
[62] Z. Li, A. H. Pearlman, P. Hsieh, *DNA Repair (Amst)* **2016**, *38*, 94.
[63] N. Pećina-Šlaus, A. Kafka, I. Salamon, A. Bukovac, *Frontiers in Molecular Biosciences* **2020**, *7*, 122.
[64] O. D. Schärer, *Cold Spring Harb Perspect Biol* **2013**, *5*, a012609.
[65] G. Spivak, *DNA Repair (Amst)* **2015**, *36*, 13.
[66] J. A. Marteijn, H. Lans, W. Vermeulen, J. H. J. Hoeijmakers, *Nat Rev Mol Cell Biol* **2014**, *15*, 465.
[67] V. F. Latypov, J. L. Tubbs, A. J. Watson, A. S. Marriott, G. McGown, M. Thorncroft, O. J. Wilkinson, P. Senthong, A. Butt, A. S. Arvai, C. L. Millington, A. C. Povey, D. M. Williams, M. F. Santibanez-Koref, J. A. Tainer, G. P. Margison, *Mol Cell* **2012**, *47*, 50.
[68] A. A. Ischenko, M. K. Saparbaev, *Nature* **2002**, *415*, 183.
[69] L. Gros, A. A. Ishchenko, H. Ide, R. H. Elder, M. K. Saparbaev, *Nucleic Acids Res* **2004**, *32*, 73.
[70] A. A. Ishchenko, E. Deprez, A. Maksimenko, J.-C. Brochon, P. Tauc, M. K. Saparbaev, *Proc Natl Acad Sci U S A* **2006**, *103*, 2564.
[71] H. E. Krokan, M. Bjørås, *Cold Spring Harb Perspect Biol* , DOI:10.1101/cshperspect.a012583.
[72] W. A. Beard, S. H. Wilson, *Chem Rev* **2006**, *106*, 361.
[73] W. A. Beard, J. K. Horton, R. Prasad, S. H. Wilson, *Annual Review of Biochemistry* **2019**, *88*, 137.
[74] D. Svilar, E. M. Goellner, K. H. Almeida, R. W. Sobol, *Antioxid Redox Signal* **2011**, *14*, 2491.
[75] M. L. Hegde, T. K. Hazra, S. Mitra, *Cell Res* **2008**, *18*, 27.
[76] A. Hanssen-Bauer, K. Solvang-Garten, O. Sundheim, J. Peña-Diaz, S. Andersen, G. Slupphaug, H. E. Krokan, D. M. Wilson, M. Akbari, M. Otterlei, *Environ Mol Mutagen* **2011**, *52*, 623.
[77] A. L. Jacobs, P. Schär, *Chromosoma* **2012**, *121*, 1.
[78] H. Dou, S. Mitra, T. K. Hazra, *Journal of Biological Chemistry* **2003**, *278*, 49679.
[79] J. Ramstein, R. Lavery, *PNAS* **1988**, *85*, 7231.
[80] S. C. Brooks, S. Adhikary, E. H. Rubinson, B. F. Eichman, *Biochim. Biophys. Acta* **2013**, *1834*, 247.
[81] J. L. Tubbs, A. E. Pegg, J. A. Tainer, *DNA Repair (Amst)* **2007**, *6*, 1100.

[82] B. Dalhus, J. K. Laerdahl, P. H. Backe, M. Bjørås, *FEMS Microbiol. Rev.* **2009**, *33*, 1044.
[83] D. R. Denver, S. L. Swenson, M. Lynch, *Mol Biol Evol* **2003**, *20*, 1603.
[84] N. A. Kuznetsov, O. S. Fedorova, *Adv. Exp. Med. Biol.* **2020**, *1241*, 1.
[85] M. Sugahara, T. Mikawa, T. Kumasaka, M. Yamamoto, R. Kato, K. Fukuyama, Y. Inoue, S. Kuramitsu, *EMBO J* **2000**, *19*, 3857.
[86] R. M. Aamodt, P. Ø. Falnes, R. F. Johansen, E. Seeberg, M. Bjørås, *J. Biol. Chem.* **2004**, *279*, 13601.
[87] M. D. Wyatt, J. M. Allan, A. Y. Lau, T. E. Ellenberger, L. D. Samson, *BioEssays* **1999**, *21*, 668.
[88] T. Lindahl, *Proc Natl Acad Sci U S A* **1974**, *71*, 3649.
[89] C. D. Mol, A. S. Arvai, G. Slupphaug, B. Kavli, I. Alseth, H. E. Krokan, J. A. Tainer, *Cell* **1995**, *80*, 869.
[90] N. Schormann, R. Ricciardi, D. Chattopadhyay, *Protein Sci* **2014**, *23*, 1667.
[91] I. Alseth, T. Rognes, T. Lindbäck, I. Solberg, K. Robertsen, K. I. Kristiansen, D. Mainieri, L. Lillehagen, A.-B. Kolstø, M. Bjørås, *Mol Microbiol* **2006**, *59*, 1602.
[92] B. Dalhus, I. H. Helle, P. H. Backe, I. Alseth, T. Rognes, M. Bjørås, J. K. Laerdahl, *Nucleic Acids Res* **2007**, *35*, 2451.
[93] E. H. Rubinson, A. S. P. Gowda, T. E. Spratt, B. Gold, B. F. Eichman, *Nature* **2010**, *468*, 406.
[94] M. Li, D. M. Wilson, *Antioxid Redox Signal* **2014**, *20*, 678.
[95] A. M. Whitaker, B. D. Freudenthal, *DNA Repair (Amst)* **2018**, *71*, 93.
[96] U. Hübscher, S. Spadari, G. Villani, G. Maga, *DNA Polymerases: Discovery, Characterization and Functions in Cellular DNA Transactions*, WORLD SCIENTIFIC, **2010**.
[97] L. M. Chang, F. J. Bollum, *J Biol Chem* **1971**, *246*, 5835.
[98] W. A. Beard, S. H. Wilson, *Mutat Res* **2000**, *460*, 231.
[99] E. K. Braithwaite, P. S. Kedar, D. J. Stumpo, B. Bertocci, J. H. Freedman, L. D. Samson, S. H. Wilson, *PLoS One* **2010**, *5*, e12229.
[100] J. McIntyre, *DNA Repair* **2020**, *86*, 102753.
[101] B. A. Baptiste, S. L. Baringer, T. Kulikowicz, J. A. Sommers, D. L. Croteau, R. M. Brosh, V. A. Bohr, *DNA Repair (Amst)* **2021**, *99*, 103050.
[102] N. A. Lebedeva, N. I. Rechkunova, S. V. Dezhurov, S. N. Khodyreva, A. Favre, L. Blanco, O. I. Lavrik, *Biochim Biophys Acta* **2005**, *1751*, 150.
[103] K. Asagoshi, W. Lehmann, E. K. Braithwaite, L. Santana-Santos, R. Prasad, J. H. Freedman, B. Van Houten, S. H. Wilson, *Nucleic Acids Res* **2012**, *40*, 670.
[104] *The EMBO Journal* **2000**, *19*, 1397.
[105] M. Meijer, F. Karimi-Busheri, T. Y. Huang, M. Weinfeld, D. Young, *J Biol Chem* **2002**, *277*, 4050.
[106] M. Stucki, B. Pascucci, E. Parlanti, P. Fortini, S. H. Wilson, U. Hübscher, E. Dogliotti, *Oncogene* **1998**, *17*, 835.
[107] G. L. Dianov, U. Hübscher, *Nucleic Acids Res* **2013**, *41*, 3483.
[108] R. Prasad, G. L. Dianov, V. A. Bohr, S. H. Wilson, *J Biol Chem* **2000**, *275*, 4460.
[109] D. Slade, *Genes (Basel)* **2018**, *9*, E416.
[110] A. González-Magaña, F. J. Blanco, *Biomolecules* **2020**, *10*, E570.
[111] G. Frosina, P. Fortini, O. Rossi, F. Carrozzino, G. Raspaglio, L. S. Cox, D. P. Lane, A. Abbondandolo, E. Dogliotti, *J Biol Chem* **1996**, *271*, 9573.
[112] J. Fan, M. Otterlei, H.-K. Wong, A. E. Tomkinson, D. M. Wilson, *Nucleic Acids Res* **2004**, *32*, 2193.
[113] L.-Y. Hu, C.-C. Chang, Y.-S. Huang, W.-C. Chou, Y.-M. Lin, C.-C. Ho, W.-T. Chen, H.-M. Shih, C.-N. Hsiung, P.-E. Wu, C.-Y. Shen, *Hum Mol Genet* **2018**, *27*, 2306.
[114] A. Campalans, S. Marsin, Y. Nakabeppu, T. R. O'connor, S. Boiteux, J. P. Radicella, *DNA Repair (Amst)* **2005**, *4*, 826.
[115] S. Marsin, A. E. Vidal, M. Sossou, J. Ménissier-de Murcia, F. Le Page, S. Boiteux, G. de Murcia, J. P. Radicella, *J Biol Chem* **2003**, *278*, 44068.
[116] C. J. Whitehouse, R. M. Taylor, A. Thistlethwaite, H. Zhang, F. Karimi-Busheri, D. D. Lasko, M. Weinfeld, K. W. Caldecott, *Cell* **2001**, *104*, 107.
[117] T. Tsurimoto, B. Stillman, *Proc Natl Acad Sci U S A* **1990**, *87*, 1023.
[118] R. M. Overmeer, A. M. Gourdin, A. Giglia-Mari, H. Kool, A. B. Houtsmuller, G. Siegal, M. I. Fousteri, L. H. F. Mullenders, W. Vermeulen, *Mol Cell Biol* **2010**, *30*, 4828.
[119] S. Vijayakumar, B. Dziegielewska, D. S. Levin, W. Song, J. Yin, A. Yang, Y. Matsumoto, V. P. Bermudez, J. Hurwitz, A. E. Tomkinson, *Mol Cell Biol* **2009**, *29*, 2042.
[120] Y. Liu, R. Prasad, S. H. Wilson, *Biochim Biophys Acta* **2010**, *1799*, 119.

[121] A. Balliano, F. Hao, C. Njeri, L. Balakrishnan, J. J. Hayes, *Biochemistry* **2017**, *56*, 647.
[122] S. Shuman, *Journal of Biological Chemistry* **2009**, *284*, 17365.
[123] T. R. L. Howes, A. E. Tomkinson, *Subcell Biochem* **2012**, *62*, 327.
[124] D. Simsek, A. Furda, Y. Gao, J. Artus, E. Brunet, A.-K. Hadjantonakis, B. Van Houten, S. Shuman, P. J. McKinnon, M. Jasin, *Nature* **2011**, *471*, 245.
[125] M. A. Dickson, G. K. Schwartz, *Curr Oncol* **2009**, *16*, 36.
[126] M. Tiwari, *J Cancer Res Ther* **2012**, *8*, 510.
[127] G. Delaney, S. Jacob, C. Featherstone, M. Barton, *Cancer* **2005**, *104*, 1129.
[128] L. Li, Y. Guan, X. Chen, J. Yang, Y. Cheng, *Frontiers in Pharmacology* **2021**, *11*, 2520.
[129] J. C. Lucchesi, *Genetics* **1968**, *59*, 37.
[130] N. S. Gavande, P. S. VanderVere-Carozza, H. D. Hinshaw, S. I. Jalal, C. R. Sears, K. S. Pawelczak, J. J. Turchi, *Pharmacology & Therapeutics* **2016**, *160*, 65.
[131] D. Woods, J. J. Turchi, *Cancer biology & therapy* **2013**, *14*, 379.
[132] T. Helleday, E. Petermann, C. Lundin, B. Hodgson, R. A. Sharma, *Nat Rev Cancer* **2008**, *8*, 193.
[133] A. Klinakis, D. Karagiannis, T. Rampias, *Cell. Mol. Life Sci.* **2020**, *77*, 677.
[134] J. B. Sweasy, T. Lang, D. DiMaio, *Cell Cycle* **2006**, *5*, 250.
[135] S. Adhikari, S. Choudhury, P. S. Mitra, J. J. Dubash, S. P. Sajankila, R. Roy, *Anticancer Agents Med Chem* **2008**, *8*, 351.
[136] E. Markkanen, *DNA Repair* **2017**, *59*, 82.
[137] T. Visnes, M. Grube, B. M. F. Hanna, C. Benitez-Buelga, A. Cázares-Körner, T. Helleday, *DNA Repair* **2018**, *71*, 118.
[138] C. J. Lord, A. Ashworth, *Science* **2017**, *355*, 1152.
[139] H. Farmer, N. McCabe, C. J. Lord, A. N. J. Tutt, D. A. Johnson, T. B. Richardson, M. Santarosa, K. J. Dillon, I. Hickson, C. Knights, N. M. B. Martin, S. P. Jackson, G. C. M. Smith, A. Ashworth, *Nature* **2005**, *434*, 917.
[140] T. S. Piskunova, M. N. Yurova, A. I. Ovsyannikov, A. V. Semenchenko, M. A. Zabezhinski, I. G. Popovich, Z.-Q. Wang, V. N. Anisimov, *Curr Gerontol Geriatr Res* **2008**, *2008*, 754190.
[141] J. Horton, S. Wilson, *Frontiers in Oncology* **2013**, *3*, 257.
[142] N. M. Tung, J. E. Garber, *Br J Cancer* **2018**, *119*, 141.
[143] B. Pellegrino, A. Musolino, A. Llop-Guevara, V. Serra, P. De Silva, Z. Hlavata, D. Sangiolo, K. Willard-Gallo, C. Solinas, *Transl Oncol* **2020**, *13*, 410.
[144] M. R. Purnell, W. J. Whish, *Biochem J* **1980**, *185*, 775.
[145] G. Papeo, H. Posteri, D. Borghi, A. A. Busel, F. Caprera, E. Casale, M. Ciomei, A. Cirla, E. Corti, M. D'Anello, M. Fasolini, B. Forte, A. Galvani, A. Isacchi, A. Khvat, M. Y. Krasavin, R. Lupi, P. Orsini, R. Perego, E. Pesenti, D. Pezzetta, S. Rainoldi, F. Riccardi-Sirtori, A. Scolaro, F. Sola, F. Zuccotto, E. R. Felder, D. Donati, A. Montagnoli, *J Med Chem* **2015**, *58*, 6875.
[146] K. Bixel, J. L. Hays, *PGPM* **2015**, *8*, 127.
[147] C. Franchet, J.-S. Hoffmann, F. Dalenc, *Cancers (Basel)* **2021**, *13*, 4132.
[148] K. A. Reiss, R. Mick, M. H. O'Hara, U. Teitelbaum, T. B. Karasic, C. Schneider, S. Cowden, T. Southwell, J. Romeo, N. Izgur, Z. M. Hannan, R. Tondon, K. Nathanson, R. H. Vonderheide, M. M. Wattenberg, G. Beatty, S. M. Domchek, *JCO* **2021**, *39*, 2497.
[149] S. Sinha, S. Molla, C. N. Kundu, *Med Oncol* **2021**, *38*, 118.
[150] M. Macchini, F. Centonze, U. Peretti, G. Orsi, A. M. Militello, M. M. Valente, S. Cascinu, M. Reni, *Cancer Treat Rev* **2021**, *100*, 102262.
[151] A. Alhusaini, A. Cannon, S. G. Maher, J. V. Reynolds, N. Lynam-Lennon, *Biomedicines* **2021**, *9*, 1024.
[152] Z. Na, B. Peng, S. Ng, S. Pan, J.-S. Lee, H.-M. Shen, S. Q. Yao, *Angewandte Chemie International Edition* **2015**, *54*, 2515.
[153] C. Le Tourneau, B. Dreno, Y. Kirova, J. J. Grob, T. Jouary, C. Dutriaux, L. Thomas, C. Lebbé, L. Mortier, P. Saiag, M. F. Avril, E. Maubec, P. Joly, P. Bey, J. M. Cosset, J. S. Sun, B. Asselain, F. Devun, M. E. Marty, M. Dutreix, *Br J Cancer* **2016**, *114*, 1199.
[154] S. Thierry, W. Jdey, S. Alculumbre, V. Soumelis, P. Noguiez-Hellin, M. Dutreix, *Mol Cancer Ther* **2017**, *16*, 2817.
[155] J. Biau, F. Devun, P. Verrelle, M. Dutreix, *Bull Cancer* **2016**, *103*, 227.
[156] N. C. Bauer, A. H. Corbett, P. W. Doetsch, *Nucleic Acids Res* **2015**, *43*, 10083.
[157] G. Kaur, R. P. Cholia, A. K. Mantha, R. Kumar, *J Med Chem* **2014**, *57*, 10241.
[158] R. I. Al-Safi, S. Odde, Y. Shabaik, N. Neamati, *Curr Mol Pharmacol* **2012**, *5*, 14.

[159] M. Poletto, M. C. Malfatti, D. Dorjsuren, P. L. Scognamiglio, D. Marasco, C. Vascotto, A. Jadhav, D. J. Maloney, D. M. Wilson, A. Simeonov, G. Tell, *Mol Carcinog* **2016**, *55*, 688.
[160] Spectrum Pharmaceuticals, Inc, *An International, Multi-Center, Randomized, Double Blind Placebo Controlled Phase II Study to Evaluate the Safety and Efficacy of Lucanthone Administered as an Adjunct to Radiation and Temozolomide for Primary Therapy of Glioblastoma Multiforme*, clinicaltrials.gov, **2021**.
[161] Q. Li, Z.-W. Zhou, W. Duan, C.-Y. Qian, S.-N. Wang, M.-S. Deng, D. Zi, J.-M. Wang, C.-Y. Mao, G. Song, D. Wang, K. D. Westover, C.-X. Xu, *J Exp Clin Cancer Res* **2021**, *40*, 220.
[162] L. Chu, A. K. L. Anderson, M. A. Landers, Y. Wang, M. R. Kelley, R. A. Messmann, *JCO* **2019**, *37*, e14531.
[163] Apexian Pharmaceuticals, Inc., *A Phase 1 Study of APX3330 in Patients With Advanced Solid Tumors*, clinicaltrials.gov, **2020**.
[164] S. Choi, H. K. Joo, B. H. Jeon, *Chonnam Med J* **2016**, *52*, 75.
[165] G. D. Hartman, N. A. Lambert-Cheatham, M. R. Kelley, T. W. Corson, *Int J Mol Sci* **2021**, *22*, 10279.
[166] R. A. Caston, S. Gampala, L. Armstrong, R. A. Messmann, M. L. Fishel, M. R. Kelley, *Drug Discov Today* **2021**, *26*, 218.
[167] H. Gu, J. D. Marth, P. C. Orban, H. Mossmann, K. Rajewsky, *Science* **1994**, *265*, 103.
[168] D. Starcevic, S. Dalal, J. B. Sweasy, *Cell Cycle* **2004**, *3*, 998.
[169] K. H. Barakat, M. M. Gajewski, J. A. Tuszynski, *Drug Discov Today* **2012**, *17*, 913.
[170] I. Kuriyama, A. Miyazaki, Y. Tsuda, H. Yoshida, Y. Mizushina, *Bioorganic & Medicinal Chemistry* **2013**, *21*, 403.
[171] Y. Mizushina, H. Suzuki-Fukudome, T. Takeuchi, K. Takemoto, I. Kuriyama, H. Yoshida, S. Kamisuki, F. Sugawara, *Bioorganic & Medicinal Chemistry* **2014**, *22*, 1070.
[172] A. S. Jaiswal, S. Banerjee, R. Aneja, F. H. Sarkar, D. A. Ostrov, S. Narayan, *PLOS ONE* **2011**, *6*, e16691.
[173] R. Paul, S. Banerjee, M. M. Greenberg, *ACS Chem. Biol.* **2017**, *12*, 1576.
[174] A. Alomari, R. Gowland, C. Southwood, J. Barrow, Z. Bentley, J. Calvin-Nelson, A. Kaminski, M. LeFevre, A. J. Callaghan, H. A. Vincent, D. M. Gowers, *Molecules* **2021**, *26*, 2508.
[175] D. K. Singh, S. Krishna, S. Chandra, M. Shameem, A. L. Deshmukh, D. Banerjee, *Medicinal Research Reviews* **2014**, *34*, 567.
[176] S. W. Yang, P. Huang, W. Plunkett, F. F. Becker, J. Y. Chan, *J Biol Chem* **1992**, *267*, 2345.
[177] M. K. Hussain, D. K. Singh, A. Singh, M. Asad, M. I. Ansari, M. Shameem, S. Krishna, G. R. Valicherla, V. Makadia, S. Meena, A. L. Deshmukh, J. R. Gayen, M. Imran Siddiqi, D. Datta, K. Hajela, D. Banerjee, *Sci Rep* **2017**, *7*, 10715.
[178] R. Ali, M. Alabdullah, M. Algethami, A. Alblihy, I. Miligy, A. Shoqafi, K. A. Mesquita, T. Abdel-Fatah, S. Y. Chan, P. W. Chiang, N. P. Mongan, E. A. Rakha, A. E. Tomkinson, S. Madhusudan, *Theranostics* **2021**, *11*, 8350.
[179] L. He, L. Luo, H. Zhu, H. Yang, Y. Zhang, H. Wu, H. Sun, F. Jiang, C. S. Kathera, L. Liu, Z. Zhuang, H. Chen, F. Pan, Z. Hu, J. Zhang, Z. Guo, *Mol Oncol* **2017**, *11*, 640.
[180] J.-L. Li, J.-P. Wang, H. Chang, S.-M. Deng, J.-H. Du, X.-X. Wang, H.-J. Hu, D.-Y. Li, X.-B. Xu, W.-Q. Guo, Y.-H. Song, Z. Guo, M.-X. Sun, Y.-W. Wu, S.-B. Liu, *Cancer Med* **2019**, *8*, 7774.
[181] P. S. Patel, A. Algouneh, R. Hakem, *Oncogene* **2021**, *40*, 3001.
[182] K. A. Mesquita, R. Ali, R. Doherty, M. S. Toss, I. Miligy, A. Alblihy, D. Dorjsuren, A. Simeonov, A. Jadhav, D. M. Wilson, I. Hickson, N. J. Tatum, E. A. Rakha, S. Madhusudan, *Cancers (Basel)* **2021**, *13*, 1866.
[183] C. Ma, S. Jin, H. Liu, K. Xia, J. Tang, K. Wang, J. Wang, *Mol Cell Probes* **2015**, *29*, 500.
[184] A. Zereshkian, J. V. Leyton, Z. Cai, D. Bergstrom, M. Weinfeld, R. M. Reilly, *Nucl Med Biol* **2014**, *41*, 377.
[185] P. Srivastava, A. Sarma, C. M. Chaturvedi, *PLoS One* **2018**, *13*, e0190516.
[186] T. R. Mereniuk, M. A. M. El Gendy, A. M. Mendes-Pereira, C. J. Lord, S. Ghosh, E. Foley, A. Ashworth, M. Weinfeld, *Mol Cancer Ther* **2013**, *12*, 2135.
[187] Z. Shire, M. R. Vakili, T. D. R. Morgan, D. G. Hall, A. Lavasanifar, M. Weinfeld, *Mol Pharm* **2018**, *15*, 2316.
[188] S. M. A. Sadat, I. M. Paiva, Z. Shire, F. Sanaee, T. D. R. Morgan, M. Paladino, F. Karimi-Busheri, R. S. Mani, G. R. Martin, F. R. Jirik, D. G. Hall, M. Weinfeld, A. Lavasanifar, *J Control Release* **2021**, *334*, 335.

[189] S. Giovannini, M.-C. Weller, H. Hanzlíková, T. Shiota, S. Takeda, J. Jiricny, *Nucleic Acids Res* **2020**, *48*, 4928.
[190] L. You, H. Cui, F. Zhao, H. Sun, H. Zhong, G. Zhou, X. Chen, *Bioengineered* **2021**, *12*, 4995.
[191] S. Cortellino, J. Xu, M. Sannai, R. Moore, E. Caretti, A. Cigliano, M. Le Coz, K. Devarajan, A. Wessels, D. Soprano, L. K. Abramowitz, M. S. Bartolomei, F. Rambow, M. R. Bassi, T. Bruno, M. Fanciulli, C. Renner, A. J. Klein-Szanto, Y. Matsumoto, D. Kobi, I. Davidson, C. Alberti, L. Larue, A. Bellacosa, *Cell* **2011**, *146*, 67.
[192] Y. Xie, H. Yang, C. Cunanan, K. Okamoto, D. Shibata, J. Pan, D. E. Barnes, T. Lindahl, M. McIlhatton, R. Fishel, J. H. Miller, *Cancer Res.* **2004**, *64*, 3096.
[193] M. K. Chan, M. T. Ocampo-Hafalla, V. Vartanian, P. Jaruga, G. Kirkali, K. L. Koenig, S. Brown, R. S. Lloyd, M. Dizdaroglu, G. W. Teebor, *DNA Repair (Amst.)* **2009**, *8*, 786.
[194] P. D. Lawley, D. H. Phillips, *Mutat. Res.* **1996**, *355*, 13.
[195] G. V. Mechetin, A. V. Endutkin, E. A. Diatlova, D. O. Zharkov, *International Journal of Molecular Sciences* **2020**, *21*, 3118.
[196] Y.-H. He, Y. Xu, M. Kobune, M. Wu, M. R. Kelley, W. J. Martin, *American Journal of Physiology-Lung Cellular and Molecular Physiology* **2002**, *282*, L50.
[197] C. F. Errol, C. W. Graham, S. Wolfram, D. W. Richard, A. S. Roger, E. Tom, *DNA Repair and Mutagenesis, Second Edition*, American Society of Microbiology, **2006**.
[198] B. Doseth, T. Visnes, A. Wallenius, I. Ericsson, A. Sarno, H. S. Pettersen, A. Flatberg, T. Catterall, G. Slupphaug, H. E. Krokan, B. Kavli, *J. Biol. Chem.* **2011**, *286*, 16669.
[199] A. D. Bulgar, L. D. Weeks, Y. Miao, S. Yang, Y. Xu, C. Guo, S. Markowitz, N. Oleinick, S. L. Gerson, L. Liu, *Cell Death Dis* **2012**, *3*, e252.
[200] Y. Yan, Y. Qing, J. J. Pink, S. L. Gerson, *Mol Cancer Res* **2018**, *16*, 212.
[201] S. M. K. Pulukuri, J. A. Knost, N. Estes, J. S. Rao, *Mol Cancer Res* **2009**, *7*, 1285.
[202] L. D. Weeks, G. E. Zentner, P. C. Scacheri, S. L. Gerson, *Cell Death Dis* **2014**, *5*, e1045.
[203] F. Ni, H. Tang, C. Wang, Z. Wang, F. Yu, B. Chen, L. Sun, *Cancer Manag Res* **2019**, *11*, 8391.
[204] A. Sawant, A. M. Floyd, M. Dangeti, W. Lei, R. W. Sobol, S. M. Patrick, *DNA Repair* **2017**, *51*, 46.
[205] A. A. Serebrenik, G. J. Starrett, S. Leenen, M. C. Jarvis, N. M. Shaban, D. J. Salamango, H. Nilsen, W. L. Brown, R. S. Harris, *Proc Natl Acad Sci U S A* **2019**, *116*, 22158.
[206] K. Nishida, T. Arazoe, N. Yachie, S. Banno, M. Kakimoto, M. Tabata, M. Mochizuki, A. Miyabe, M. Araki, K. Y. Hara, Z. Shimatani, A. Kondo, *Science* , DOI:10.1126/science.aaf8729.
[207] Y. L. Jiang, A. C. Drohat, Y. Ichikawa, J. T. Stivers, *J. Biol. Chem.* **2002**, *277*, 15385.
[208] F. Focher, A. Verri, S. Spadari, R. Manservigi, J. Gambino, G. E. Wright, *Biochem J* **1993**, *292*, 883.
[209] Y. L. Jiang, D. J. Krosky, L. Seiple, J. T. Stivers, *J Am Chem Soc* **2005**, *127*, 17412.
[210] S. Chung, J. B. Parker, M. Bianchet, L. M. Amzel, J. T. Stivers, *Nat Chem Biol* **2009**, *5*, 407.
[211] D. J. Krosky, M. A. Bianchet, L. Seiple, S. Chung, L. M. Amzel, J. T. Stivers, *Nucleic Acids Res* **2006**, *34*, 5872.
[212] Y. L. Jiang, S. Chung, D. J. Krosky, J. T. Stivers, *Bioorganic & Medicinal Chemistry* **2006**, *14*, 5666.
[213] Q. An, P. Robins, T. Lindahl, D. E. Barnes, *Cancer Res* **2007**, *67*, 940.
[214] M.-J. An, G.-S. Shin, H.-M. Lee, J.-W. Kim, *Genes (Basel)* **2021**, *12*, 201.
[215] D. P. Turner, S. Cortellino, J. E. Schupp, E. Caretti, T. Loh, T. J. Kinsella, A. Bellacosa, *Cancer Res.* **2006**, *66*, 7686.
[216] H. S. Pettersen, T. Visnes, C. B. Vågbø, Eva. K. Svaasand, B. Doseth, G. Slupphaug, B. Kavli, H. E. Krokan, *Nucleic Acids Res* **2011**, *39*, 8430.
[217] P. Mancuso, R. Tricarico, V. Bhattacharjee, L. Cosentino, Y. Kadariya, J. Jelinek, E. Nicolas, M. Einarson, N. Beeharry, K. Devarajan, R. A. Katz, D. G. Dorjsuren, H. Sun, A. Simeonov, A. Giordano, J. R. Testa, G. Davidson, I. Davidson, L. Larue, R. W. Sobol, T. J. Yen, A. Bellacosa, *Oncogene* **2019**, *38*, 3710.
[218] Q. Dai, X. Lu, L. Zhang, C. He, *Tetrahedron* **2012**, *68*, 5145.
[219] T. Suksangpleng, U. Leartsakulpanich, S. Moonsom, S. Siribal, U. Boonyuen, G. E. Wright, P. Chavalitshewinkoon-Petmitr, *Malar J* **2014**, *13*, 149.
[220] N. Schormann, C. I. Sommers, M. N. Prichard, K. A. Keith, J. W. Noah, M. Nuth, R. P. Ricciardi, D. Chattopadhyay, *Antimicrob Agents Chemother* **2011**, *55*, 5054.
[221] C.-Y. I. Lee, J. C. Delaney, M. Kartalou, G. M. Lingaraju, A. Maor-Shoshani, J. M. Essigmann, L. D. Samson, *Biochemistry* **2009**, *48*, 1850.
[222] M. L. Fishel, Y. R. Seo, M. L. Smith, M. R. Kelley, *Cancer Res* **2003**, *63*, 608.

[223] N. M. Leguisamo, H. C. Gloria, A. N. Kalil, T. V. Martins, D. B. Azambuja, L. B. Meira, J. Saffi, *Oncotarget* **2017**, *8*, 54199.
[224] R. W. Sobol, M. Kartalou, K. H. Almeida, D. F. Joyce, B. P. Engelward, J. K. Horton, R. Prasad, L. D. Samson, S. H. Wilson, *J. Biol. Chem.* **2003**, *278*, 39951.
[225] J. Tang, D. Svilar, R. N. Trivedi, X. Wang, E. M. Goellner, B. Moore, R. L. Hamilton, L. A. Banze, A. R. Brown, R. W. Sobol, *Neuro Oncol* **2011**, *13*, 471.
[226] J. A. Calvo, C. A. Moroski-Erkul, A. Lake, L. W. Eichinger, D. Shah, I. Jhun, P. Limsirichai, R. T. Bronson, D. C. Christiani, L. B. Meira, L. D. Samson, *PLoS Genet* , DOI:10.1371/journal.pgen.1003413.
[227] J. Paik, T. Duncan, T. Lindahl, B. Sedgwick, *Cancer Res.* **2005**, *65*, 10472.
[228] I. C. Sorribes, S. K. Handelman, H. V. Jain, *Journal of The Royal Society Interface* **2020**, *17*, 20190722.
[229] J. R. Silber, M. S. Bobola, D. D. Kolstoe, A. Blank, M. C. Chamberlain, *Front. Oncol.* , DOI:10.3389/fonc.2012.00176.
[230] M. R. Ebrahimkhani, A. Daneshmand, A. Mazumder, M. Allocca, J. A. Calvo, N. Abolhassani, I. Jhun, S. Muthupalani, C. Ayata, L. D. Samson, *PNAS* **2014**, *111*, E4878.
[231] P. A. J. Crosbie, A. J. Watson, R. Agius, P. V. Barber, G. P. Margison, A. C. Povey, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **2012**, *732*, 43.
[232] B. Tudek, A. A. VanZeeland, J. T. Kusmierek, J. Laval, *Mutation Research/DNA Repair* **1998**, *407*, 169.
[233] S. Adhikari, J. A. Toretsky, L. Yuan, R. Roy, *J Biol Chem* **2006**, *281*, 29525.
[234] M. Dixon, J. Woodrick, S. Gupta, S. K. Karmahapatra, S. Devito, S. Vasudevan, S. Dakshanamurthy, S. Adhikari, V. M. Yenugonda, R. Roy, *Bioorg Med Chem* **2015**, *23*, 1102.
[235] P. Mladenka, L. Zatloukalová, T. Filipský, R. Hrdina, *Free Radic Biol Med* **2010**, *49*, 963.
[236] G. M. Lingaraju, C. A. Davis, J. W. Setser, L. D. Samson, C. L. Drennan, *J Biol Chem* **2011**, *286*, 13205.
[237] E. Mas Claret, B. Al Yahyaei, S. Chu, R. M. Elliott, M. Imperato, A. Lopez, L. B. Meira, B. J. Howlin, D. K. Whelligan, *Bioorganic & Medicinal Chemistry* **2020**, *28*, 115507.
[238] J.-H. Yoon, S. Iwai, T. R. O'Connor, G. P. Pfeifer, *Nucleic Acids Res* **2003**, *31*, 5399.
[239] Q. An, P. Robins, T. Lindahl, D. E. Barnes, *EMBO J* **2005**, *24*, 2205.
[240] M. L. Michaels, C. Cruz, A. P. Grollman, J. H. Miller, *Proc Natl Acad Sci U S A* **1992**, *89*, 7022.
[241] F. Taddei, H. Hayakawa, M. Bouton, A. Cirinesi, I. Matic, M. Sekiguchi, M. Radman, *Science* **1997**, *278*, 128.
[242] R. Eutsey, G. Wang, R. J. Maier, *DNA Repair (Amst)* **2007**, *6*, 19.
[243] T. Davidsen, E. K. Amundsen, E. A. Rødland, T. Tønjum, *FEMS Immunol Med Microbiol* **2007**, *49*, 243.
[244] R. Gutiérrez, Y. Ram, J. Berman, K. Carstens Marques de Sousa, Y. Nachum-Biala, M. Britzi, D. Elad, G. Glaser, S. Covo, S. Harrus, *Mol Biol Evol* **2021**, msab196.
[245] E. Speina, J. M. Cieśla, M.-A. Graziewicz, J. Laval, Z. Kazimierczuk, B. Tudek, *Acta Biochim Pol* **2005**, *52*, 167.
[246] C. Rieux, S. Goffinont, F. Coste, Z. Tber, J. Cros, V. Roy, M. Guérin, V. Gaudon, S. Bourg, A. Biela, V. Aucagne, L. Agrofoglio, N. Garnier, B. Castaing, *Int J Mol Sci* **2020**, *21*, E2058.
[247] A. Biela, F. Coste, F. Culard, M. Guerin, S. Goffinont, K. Gasteiger, J. Cieśla, A. Winczura, Z. Kazimierczuk, D. Gasparutto, T. Carell, B. Tudek, B. Castaing, *Nucleic Acids Res* **2014**, *42*, 10748.
[248] R. J. Eberle, M. A. Coronado, R. S. Peinado, F. R. de Moraes, D. Olivier, T. Dreyer, D. de Oliveira Lopes, B. S. R. da Luz, V. Azevedo, R. K. Arni, *Int J Biol Macromol* **2019**, *125*, 459.
[249] M. L. Hutchcraft, H. H. Gallion, J. M. Kolesar, *Diagnostics (Basel)* **2021**, *11*, 84.
[250] A. Vincent-Salomon, G. Bataillon, L. Djerroudi, *Ann Pathol* **2020**, *40*, 78.
[251] R. D. Weren, M. J. Ligtenberg, A. Geurts van Kessel, R. M. De Voer, N. Hoogerbrugge, R. P. Kuiper, *J Pathol* **2018**, *244*, 135.
[252] J. E. Grolleman, R. M. de Voer, F. A. Elsayed, M. Nielsen, R. D. A. Weren, C. Palles, M. J. L. Ligtenberg, J. R. Vos, S. W. Ten Broeke, N. F. C. C. de Miranda, R. A. Kuiper, E. J. Kamping, E. A. M. Jansen, M. E. Vink-Börger, I. Popp, A. Lang, I. Spier, R. Hüneburg, P. A. James, N. Li, M. Staninova, H. Lindsay, D. Cockburn, O. Spasic-Boskovic, M. Clendenning, K. Sweet, G. Capellá, W. Sjursen, H. Høberg-Vetti, M. C. Jongmans, K. Neveling, A. Geurts van Kessel, H. Morreau, F. J. Hes, R. H. Sijmons, H. K. Schackert, C. Ruiz-Ponte, D. Dymerska, J. Lubinski, B. Rivera, W. D. Foulkes, I.

P. Tomlinson, L. Valle, D. D. Buchanan, S. Kenwrick, J. Adlard, A. J. Dimovski, I. G. Campbell, S. Aretz, D. Schindler, T. van Wezel, N. Hoogerbrugge, R. P. Kuiper, *Cancer Cell* **2019**, *35*, 256.
[253] N. Yang, M. A. Chaudhry, S. S. Wallace, *DNA Repair (Amst)* **2006**, *5*, 43.
[254] D. Guay, C. Garand, S. Reddy, C. Schmutte, M. Lebel, *Cancer Sci* **2008**, *99*, 762.
[255] M. Senarisoy, C. Barette, F. Lacroix, S. De Bonis, M. Stelter, F. Hans, J.-P. Kleman, M.-O. Fauvarque, J. Timmins, *ACS Chem Biol* **2020**, *15*, 990.
[256] D.-N. Li, C.-C. Yang, J. Li, Q.-G. Ou Yang, L.-T. Zeng, G.-Q. Fan, T.-H. Liu, X.-Y. Tian, J.-J. Wang, H. Zhang, D.-P. Dai, J. Cui, J.-P. Cai, *Biochim Biophys Acta Mol Cell Res* **2021**, *1868*, 118895.
[257] Y. Kumagae, M. Hirahashi, K. Takizawa, H. Yamamoto, M. Gushima, M. Esaki, T. Matsumoto, M. Nakamura, T. Kitazono, Y. Oda, *Oncol Lett* **2018**, *16*, 1765.
[258] K. D. Arczewska, W. Krasuska, A. Stachurska, K. Karpińska, J. Sikorska, M. Kiedrowski, D. Lange, T. Stępień, B. Czarnocka, *DNA Repair (Amst)* **2020**, *95*, 102954.
[259] K. V. M. Huber, E. Salah, B. Radic, M. Gridling, J. M. Elkins, A. Stukalov, A.-S. Jemth, C. Göktürk, K. Sanjiv, K. Strömberg, T. Pham, U. W. Berglund, J. Colinge, K. L. Bennett, J. I. Loizou, T. Helleday, S. Knapp, G. Superti-Furga, *Nature* **2014**, *508*, 222.
[260] X. Dai, G. Guo, P. Zou, R. Cui, W. Chen, X. Chen, C. Yin, W. He, R. Vinothkumar, F. Yang, X. Zhang, G. Liang, *J Exp Clin Cancer Res* **2017**, *36*, 120.
[261] L. M. van der Waals, J. Laoukili, J. M. J. Jongen, D. A. Raats, I. H. M. Borel Rinkes, O. Kranenburg, *Sci Rep* **2019**, *9*, 819.
[262] J. Ji, W. Chen, W. Lian, R. Chen, J. Yang, Q. Zhang, Q. Weng, Z. Khan, J. Hu, X. Chen, P. Zou, X. Chen, G. Liang, *Cell Death Dis* **2018**, *9*, 1.
[263] U. Warpman Berglund, K. Sanjiv, H. Gad, C. Kalderén, T. Koolmeister, T. Pham, C. Gokturk, R. Jafari, G. Maddalo, B. Seashore-Ludlow, A. Chernobrovkin, A. Manoilov, I. S. Pateras, A. Rasti, A.-S. Jemth, I. Almlöf, O. Loseva, T. Visnes, B. O. Einarsdottir, F. Z. Gaugaz, A. Saleh, B. Platzack, O. A. Wallner, K. S. A. Vallin, M. Henriksson, P. Wakchaure, S. Borhade, P. Herr, Y. Kallberg, P. Baranczewski, E. J. Homan, E. Wiita, V. Nagpal, T. Meijer, N. Schipper, S. G. Rudd, L. Bräutigam, A. Lindqvist, A. Filppula, T.-C. Lee, P. Artursson, J. A. Nilsson, V. G. Gorgoulis, J. Lehtiö, R. A. Zubarev, M. Scobie, T. Helleday, *Ann Oncol* **2016**, *27*, 2275.
[264] X. Hua, K. Sanjiv, H. Gad, T. Pham, C. Gokturk, A. Rasti, Z. Zhao, K. He, M. Feng, Y. Zang, J. Zhang, Q. Xia, T. Helleday, U. Warpman Berglund, *Ther Adv Med Oncol* , DOI:10.1177/1758835919866960.
[265] I. Das, H. Gad, L. Pudelko, L. Bräutigam, R. Tuominen, T. Helleday, J. Hansson, S. E. Brage, U. W. Berglund, , DOI:10.1016/J.JID.2018.03.1195.
[266] M. Hu, J. Ning, L. Mao, Y. Yu, Y. Wu, *Oncol Lett* **2021**, *21*, 62.
[267] Beijing Konruns Pharmaceutical Co., Ltd., *Evaluation Kanitinib Treatment of Recurrent or Metastatic Malignant Solid Tumors Safety, Tolerability and Pharmacokinetic Open, Dose Escalation Phase I Clinical Study*, clinicaltrials.gov, **2020**.
[268] Beijing Konruns Pharmaceutical Co., Ltd., *A Phase 1, Multicenter, Open-label, Dose Escalation/Expansion Study Evaluating the Safety, Pharmacokinetics and Preliminary Efficacy of CX1003 (Kanitinib) in the Patients With Relapsed Advanced or Metastatic Solid Tumors*, clinicaltrials.gov, **2020**.
[269] D. Zhan, X. Zhang, J. Li, X. Ding, Y. Cui, J. Jia, *Cancer Biother Radiopharm* **2020**, *35*, 223.
[270] X. Huang, J. Chen, W. Wu, W. Yang, B. Zhong, X. Qing, Z. Shao, *Acta Biomater* **2020**, *109*, 229.
[271] A. Kumar, T. Kawamura, M. Kawatani, H. Osada, K. Y. J. Zhang, *Chem Biol Drug Des* **2017**, *89*, 862.
[272] Y. Tahara, A. M. Kietrys, M. Hebenbrock, Y. Lee, D. L. Wilson, E. T. Kool, *ACS Chem Biol* **2019**, *14*, 2606.
[273] C. Peng, Y.-H. Li, C.-W. Yu, Z.-H. Cheng, J.-R. Liu, J.-L. Hsu, L.-W. Hsin, C.-T. Huang, H.-F. Juan, J.-W. Chern, Y.-S. Cheng, *Bioorg Chem* **2021**, *110*, 104813.
[274] D. Wahi, D. Soni, A. Grover, *J Cancer* **2021**, *12*, 652.
[275] J. L. Parsons, D. O. Zharkov, G. L. Dianov, *Nucleic Acids Research* **2005**, *33*, 4849.
[276] S. Boiteux, T. R. O'Connor, F. Lederer, A. Gouyette, J. Laval, *J Biol Chem* **1990**, *265*, 3916.
[277] H. M. Nash, S. D. Bruner, O. D. Schärer, T. Kawate, T. A. Addona, E. Spooner, W. S. Lane, G. L. Verdine, *Current Biology* **1996**, *6*, 968.
[278] R. de Oliveira, P. A. van der Kemp, D. Thomas, A. Geiger, P. Nehls, S. Boiteux, *Nucleic Acids Res* **1994**, *22*, 3760.
[279] P. A. van der Kemp, D. Thomas, R. Barbey, R. de Oliveira, S. Boiteux, *Proc Natl Acad Sci U S A* **1996**, *93*, 5197.

[280] K. Arai, K. Morishita, K. Shinmura, T. Kohno, S.-R. Kim, T. Nohmi, M. Taniwaki, S. Ohwada, J. Yokota, *Oncogene* **1997**, *14*, 2857.
[281] R. Lu, H. M. Nash, G. L. Verdine, *Curr Biol* **1997**, *7*, 397.
[282] A. L. Dany, A. Tissier, *Mol Genet Genomics* **2001**, *265*, 293.
[283] C. Dherin, J. P. Radicella, M. Dizdaroglu, S. Boiteux, *Nucleic Acids Res* **1999**, *27*, 4001.
[284] S. M. Robey-Bond, R. Barrantes-Reynolds, J. P. Bond, S. S. Wallace, V. Bandaru, *Biochemistry* **2008**, *47*, 7626.
[285] C. Furihata, *Genes Environ* **2015**, *37*, 21.
[286] J. P. Radicella, C. Dherin, C. Desmaze, M. S. Fox, S. Boiteux, *Proc Natl Acad Sci U S A* **1997**, *94*, 8010.
[287] D. Lia, A. Reyes, J. T. A. de Melo Campos, T. Piolot, J. Baijer, J. P. Radicella, A. Campalans, *J Cell Sci* **2018**, *131*, jcs213538.
[288] A. Dhénaut, S. Boiteux, J. P. Radicella, *Mutat Res* **2000**, *461*, 109.
[289] M.-R. Lee, S.-H. Kim, H.-J. Cho, K.-Y. Lee, A. R. Moon, H. G. Jeong, J.-S. Lee, J.-W. Hyun, M.-H. Chung, H. J. You, *J Biol Chem* **2004**, *279*, 9857.
[290] C.-K. Youn, S.-H. Kim, D. Y. Lee, S. H. Song, I.-Y. Chang, J.-W. Hyun, M.-H. Chung, H. J. You, *Journal of Biological Chemistry* **2005**, *280*, 25185.
[291] L. Luna, V. Rolseth, G. A. Hildrestrand, M. Otterlei, F. Dantzer, M. Bjørås, E. Seeberg, *Nucleic Acids Res* **2005**, *33*, 1813.
[292] M.-V. García-Ortiz, R. R. Ariza, T. Roldán-Arjona, *Plant Mol Biol* **2001**, *47*, 795.
[293] D. Gasparutto, E. Muller, S. Boiteux, J. Cadet, *Biochim Biophys Acta* **2009**, *1790*, 16.
[294] S. D. Bruner, H. M. Nash, W. S. Lane, G. L. Verdine, *Curr Biol* **1998**, *8*, 393.
[295] F. Faucher, S. S. Wallace, S. Doublié, *J Mol Biol* **2010**, *397*, 46.
[296] A. Gogos, N. D. Clarke, *Journal of Biological Chemistry* **1999**, *274*, 30447.
[297] A. A. Sartori, G. M. Lingaraju, P. Hunziker, F. K. Winkler, J. Jiricny, *Nucleic Acids Res* **2004**, *32*, 6531.
[298] A. M. Gehring, K. M. Zatopek, B. W. Burkhart, V. Potapov, T. J. Santangelo, A. F. Gardner, *DNA Repair* **2020**, *86*, 102767.
[299] L. Zhang, Y. Li, H. Shi, D. Zhang, Z. Yang, P. Oger, J. Zheng, *Appl. Microbiol. Biotechnol.* **2019**, *103*, 8021.
[300] G. M. Lingaraju, A. A. Sartori, D. Kostrewa, A. E. Prota, J. Jiricny, F. K. Winkler, *Structure* **2005**, *13*, 87.
[301] O. D'Augustin, S. Huet, A. Campalans, J. P. Radicella, *Int J Mol Sci* **2020**, *21*, E8360.
[302] P. C. Blainey, G. Luo, S. C. Kou, W. F. Mangel, G. L. Verdine, B. Bagchi, X. S. Xie, *Nat Struct Mol Biol* **2009**, *16*, 1224.
[303] A. Ahmadi, I. Rosnes, P. Blicher, R. Diekmann, M. Schüttpelz, K. Glette, J. Tørresen, M. Bjørås, B. Dalhus, A. D. Rowe, *Nat Commun* **2018**, *9*, 5381.
[304] M. M. Rowland, J. D. Schonhoft, P. L. McKibbin, S. S. David, J. T. Stivers, *Nucleic Acids Research* **2014**, *42*, 9295.
[305] A. J. Lee, S. S. Wallace, *Free Radical Biology and Medicine* **2017**, *107*, 170.
[306] H. Nilsen, T. Lindahl, A. Verreault, *EMBO J* **2002**, *21*, 5943.
[307] H. Menoni, J. H. J. Hoeijmakers, W. Vermeulen, *Journal of Cell Biology* **2012**, *199*, 1037.
[308] K. Bilotti, E. E. Kennedy, C. Li, S. Delaney, *DNA Repair* **2017**, *59*, 1.
[309] R. Amouroux, A. Campalans, B. Epe, J. P. Radicella, *Nucleic Acids Research* **2010**, *38*, 2878.
[310] M. Bjørås, E. Seeberg, L. Luna, L. H. Pearl, T. E. Barrett, *Journal of Molecular Biology* **2002**, *317*, 171.
[311] S. D. Bruner, D. P. G. Norman, G. L. Verdine, *Nature* **2000**, *403*, 859.
[312] J. C. Fromme, S. D. Bruner, W. Yang, M. Karplus, G. L. Verdine, *Nat Struct Mol Biol* **2003**, *10*, 204.
[313] D. P. G. Norman, S. J. Chung, G. L. Verdine, *Biochemistry* **2003**, *42*, 1564.
[314] C. T. Radom, A. Banerjee, G. L. Verdine, *Journal of Biological Chemistry* **2007**, *282*, 9182.
[315] X. Cheng, C. Kelso, V. Hornak, C. de los Santos, A. P. Grollman, C. Simmerling, *J Am Chem Soc* **2005**, *127*, 13906.
[316] P. C. Blainey, A. M. van Oijen, A. Banerjee, G. L. Verdine, X. S. Xie, *PNAS* **2006**, *103*, 5752.
[317] C. M. Crenshaw, K. Nam, K. Oo, P. S. Kutchukian, B. R. Bowman, M. Karplus, G. L. Verdine, *J Biol Chem* **2012**, *287*, 24916.
[318] H. Li, A. V. Endutkin, C. Bergonzo, L. Fu, A. Grollman, D. O. Zharkov, C. Simmerling, *J. Am. Chem. Soc.* **2017**, *139*, 2682.

[319] N. A. Kuznetsov, V. V. Koval, G. A. Nevinsky, K. T. Douglas, D. O. Zharkov, O. S. Fedorova, *Journal of Biological Chemistry* **2007**, *282*, 1029.
[320] V. V. Koval, D. G. Knorre, O. S. Fedorova, *Acta Naturae* **2014**, *6*, 52.
[321] M. Audebert, J. P. Radicella, M. Dizdaroglu, *Nucleic Acids Res* **2000**, *28*, 2672.
[322] M. M. Thayer, H. Ahern, D. Xing, R. P. Cunningham, J. A. Tainer, *EMBO J* **1995**, *14*, 4108.
[323] A. Banerjee, W. Yang, M. Karplus, G. L. Verdine, *Nature* **2005**, *434*, 612.
[324] B. Dalhus, M. Forsbring, I. H. Helle, E. S. Vik, R. J. Forstrøm, P. H. Backe, I. Alseth, M. Bjørås, *Structure* **2011**, *19*, 117.
[325] T. E. Tyugashev, Y. N. Vorobjev, A. A. Kuznetsova, M. V. Lukina, N. A. Kuznetsov, O. S. Fedorova, *J. Phys. Chem. B* **2019**, *123*, 4878.
[326] I. Morland, L. Luna, E. Gustad, E. Seeberg, M. Bjørås, *DNA Repair (Amst)* **2005**, *4*, 381.
[327] D. O. Zharkov, T. A. Rosenquist, S. E. Gerchman, A. P. Grollman, *J Biol Chem* **2000**, *275*, 28607.
[328] A. E. Vidal, I. D. Hickson, S. Boiteux, J. P. Radicella, *Nucleic Acids Res* **2001**, *29*, 1285.
[329] J. L. Kellie, K. A. Wilson, S. D. Wetmore, *J. Phys. Chem. B* **2015**, *119*, 8013.
[330] J. Lu, Y. Liu, *EMBO J* **2010**, *29*, 398.
[331] H. Sampath, R. S. Lloyd, *DNA Repair (Amst)* **2019**, *81*, 102667.
[332] X.-C. Liu, X.-H. Guo, B. Chen, Z.-H. Li, X.-F. Liu, *Int Ophthalmol* **2018**, *38*, 1451.
[333] P. Li, R. A. Stetler, R. K. Leak, Y. Shi, Y. Li, W. Yu, M. V. L. Bennett, J. Chen, *Neuropharmacology* **2018**, *134*, 208.
[334] K. Sakumi, Y. Tominaga, M. Furuichi, P. Xu, T. Tsuzuki, M. Sekiguchi, Y. Nakabeppu, *Cancer Res* **2003**, *63*, 902.
[335] M. Audebert, S. Chevillard, C. Levalois, G. Gyapay, A. Vieillefond, J. Klijanienko, P. Vielh, A. K. E. Naggar, S. Oudard, S. Boiteux, J. P. Radicella, *Cancer Res* **2000**, *60*, 4740.
[336] H. Blons, J. P. Radicella, O. Laccourreye, D. Brasnu, P. Beaune, S. Boiteux, P. Laurent-Puig, *Mol Carcinog* **1999**, *26*, 254.
[337] S. Chevillard, J. P. Radicella, C. Levalois, J. Lebeau, M.-F. Poupon, S. Oudard, B. Dutrillaux, S. Boiteux, *Oncogene* **1998**, *16*, 3083.
[338] K. Shinmura, T. Kohno, H. Kasai, K. Koda, H. Sugimura, J. Yokota, *Japanese Journal of Cancer Research* **1998**, *89*, 825.
[339] M. Audebert, J. B. Charbonnier, S. Boiteux, J. P. Radicella, *DNA Repair* **2002**, *1*, 497.
[340] A. Sanjari Moghaddam, M. Nazarzadeh, R. Noroozi, H. Darvish, A. Mosavi Jarrahi, *Iran J Cancer Prev* **2016**, *9*, e3467.
[341] Y. Xu, W. K. Hansen, T. A. Rosenquist, D. A. Williams, M. Limp-Foster, M. R. Kelley, *J Pharmacol Exp Ther* **2001**, *296*, 825.
[342] M. Kobune, Y. Xu, C. Baum, M. R. Kelley, D. A. Williams, *Cancer Res* **2001**, *61*, 5116.
[343] T. J. Preston, J. T. Henderson, G. P. McCallum, P. G. Wells, *Mol Cancer Ther* **2009**, *8*, 2015.
[344] M. Wu, Z. Zhang, W. Che, *Toxicology and Applied Pharmacology* **2008**, *228*, 395.
[345] R. D. Montero-Montoya, R. López-Vargas, A. Méndez-Serrano, I. Galicia-Alonso, G. García-Vargas, L. Serrano-García, R. Beltrán-Portugal, S. Rosado-Zaidi, A. Albores-Medina, L. Oropeza-Hernández, L. Hernández-Cadena, F. Mercado-Calderón, E. Alvarado-Toledo, S. Herrera-Morales, O. Arellano-Aguilar, *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **2020**, *853*, 503170.
[346] V. Simonelli, S. Camerini, F. Mazzei, B. Van Loon, A. Allione, M. D'Errico, F. Barone, A. Minoprio, F. Ricceri, S. Guarrera, A. Russo, B. Dalhus, M. Crescenzi, U. Hübscher, M. Bjørås, G. Matullo, E. Dogliotti, *Free Radical Biology and Medicine* **2013**, *63*, 401.
[347] L. Taricani, F. Shanahan, R. H. Pierce, T. J. Guzi, D. Parry, *Cell Cycle* **2010**, *9*, 4876.
[348] Z. M. Ramdzan, C. Vadnais, R. Pal, G. Vandal, C. Cadieux, L. Leduy, S. Davoudi, L. Hulea, L. Yao, A. N. Karnezis, M. Paquet, D. Dankort, A. Nepveu, *PLOS Biology* **2014**, *12*, e1001807.
[349] K. V. M. Huber, E. Salah, B. Radic, M. Gridling, J. M. Elkins, A. Stukalov, A.-S. Jemth, C. Göktürk, K. Sanjiv, K. Strömberg, T. Pham, U. W. Berglund, J. Colinge, K. L. Bennett, J. I. Loizou, T. Helleday, S. Knapp, G. Superti-Furga, *Nature* **2014**, *508*, 222.
[350] X. Wang, C. Liu, Q. Ma, W. Feng, L. Yang, Y. Lu, Z. Zhou, Z. Yu, W. Li, L. Zhang, *Cell Physiol Biochem* **2015**, *37*, 1075.
[351] Z. M. Ramdzan, V. Ginjala, J. B. Pinder, D. Chung, C. M. Donovan, S. Kaur, L. Leduy, G. Dellaire, S. Ganesan, A. Nepveu, *Oncotarget* **2017**, *8*, 19021.
[352] M. Bosshard, E. Markkanen, B. Van Loon, *International Journal of Molecular Sciences* **2012**, *13*, 16172.

[353] I. V. Kovtun, Y. Liu, M. Bjoras, A. Klungland, S. H. Wilson, C. T. McMurray, *Nature* **2007**, *447*, 447.
[354] S. Ayala-Peña, *Free Radic Biol Med* **2013**, *62*, 102.
[355] H. Budworth, F. R. Harris, P. Williams, D. Y. Lee, A. Holt, J. Pahnke, B. Szczesny, K. Acevedo-Torres, S. Ayala-Peña, C. T. McMurray, *PLOS Genetics* **2015**, *11*, e1005267.
[356] E. Touati, V. Michel, J.-M. Thiberge, P. Avé, M. Huerre, F. Bourgade, A. Klungland, A. Labigne, *Helicobacter* **2006**, *11*, 494.
[357] A. Bacsi, L. Aguilera-Aguirre, B. Szczesny, Z. Radak, T. K. Hazra, S. Sur, X. Ba, I. Boldogh, *DNA Repair (Amst)* **2013**, *12*, 18.
[358] J. G. Mabley, P. Pacher, A. Deb, R. Wallace, R. H. Elder, C. Szabó, *FASEB J* **2005**, *19*, 290.
[359] X. Ba, L. Aguilera-Aguirre, Q. T. A. N. Rashid, A. Bacsi, Z. Radak, S. Sur, K. Hosoki, M. L. Hegde, I. Boldogh, *International Journal of Molecular Sciences* **2014**, *15*, 16975.
[360] L. Aguilera-Aguirre, K. Hosoki, A. Bacsi, Z. Radák, S. Sur, M. L. Hegde, B. Tian, A. Saavedra-Molina, A. R. Brasier, X. Ba, I. Boldogh, *Free Radical Biology and Medicine* **2015**, *89*, 20.
[361] G. Hajas, A. Bacsi, L. Aguilera-Aguirre, M. L. Hegde, K. H. Tapas, S. Sur, Z. Radak, X. Ba, I. Boldogh, *Free Radical Biology and Medicine* **2013**, *61*, 384.
[362] J. Luo, K. Hosoki, A. Bacsi, Z. Radak, M. L. Hegde, S. Sur, T. K. Hazra, A. R. Brasier, X. Ba, I. Boldogh, *Free Radical Biology and Medicine* **2014**, *73*, 430.
[363] S. Vlahopoulos, M. Adamaki, N. Khoury, V. Zoumpourlis, I. Boldogh, *Pharmacol Ther* **2019**, *194*, 59.
[364] N. Müller, A. Khobta, *Redox Biology* **2021**, *43*, 101997.
[365] X. Ba, I. Boldogh, *Redox Biology* **2018**, *14*, 669.
[366] O. D. Schärer, J. Jiricny, *Bioessays* **2001**, *23*, 270.
[367] A. M. Chu, J. C. Fettinger, S. S. David, *Bioorganic & Medicinal Chemistry Letters* **2011**, *21*, 4969.
[368] Y. T. Bliksrud, A. Ellingsen, M. Bjørås, *Journal of Inherited Metabolic Disease* **2013**, *36*, 773.
[369] N. Donley, P. Jaruga, E. Coskun, M. Dizdaroglu, A. K. McCullough, R. S. Lloyd, *ACS Chem. Biol.* **2015**, *10*, 2334.
[370] Y. Tahara, D. Auld, D. Ji, A. A. Beharry, A. M. Kietrys, D. L. Wilson, M. Jimenez, D. King, Z. Nguyen, E. T. Kool, *J. Am. Chem. Soc.* **2018**, *140*, 2105.
[371] T. Visnes, A. Cázares-Körner, W. Hao, O. Wallner, G. Masuyer, O. Loseva, O. Mortusewicz, E. Wiita, A. Sarno, A. Manoilov, J. Astorga-Wells, A.-S. Jemth, L. Pan, K. Sanjiv, S. Karsten, C. Gokturk, M. Grube, E. J. Homan, B. M. F. Hanna, C. B. J. Paulin, T. Pham, A. Rasti, U. W. Berglund, C. von Nicolai, C. Benitez-Buelga, T. Koolmeister, D. Ivanic, P. Iliev, M. Scobie, H. E. Krokan, P. Baranczewski, P. Artursson, M. Altun, A. J. Jensen, C. Kalderén, X. Ba, R. A. Zubarev, P. Stenmark, I. Boldogh, T. Helleday, *Science* **2018**, *362*, 834.
[372] T. Visnes, C. Benítez-Buelga, A. Cázares-Körner, K. Sanjiv, B. M. F. Hanna, O. Mortusewicz, V. Rajagopal, J. J. Albers, D. W. Hagey, T. Bekkhus, S. Eshtad, J. M. Baquero, G. Masuyer, O. Wallner, S. Müller, T. Pham, C. Göktürk, A. Rasti, S. Suman, R. Torres-Ruiz, A. Sarno, E. Wiita, E. J. Homan, S. Karsten, K. Marimuthu, M. Michel, T. Koolmeister, M. Scobie, O. Loseva, I. Almlöf, J. E. Unterlass, A. Pettke, J. Boström, M. Pandey, H. Gad, P. Herr, A.-S. Jemth, S. El Andaloussi, C. Kalderén, S. Rodriguez-Perales, J. Benítez, H. E. Krokan, M. Altun, P. Stenmark, U. W. Berglund, T. Helleday, *Nucleic Acids Res* **2020**, *48*, 12234.
[373] S. Boiteux, *Journal of Photochemistry and Photobiology B: Biology* **1993**, *19*, 87.
[374] D. Jiang, Z. Hatahet, J. O. Blaisdell, R. J. Melamede, S. S. Wallace, *J Bacteriol* **1997**, *179*, 3773.
[375] T. K. Hazra, T. Izumi, I. Boldogh, B. Imhoff, Y. W. Kow, P. Jaruga, M. Dizdaroglu, S. Mitra, *PNAS* **2002**, *99*, 3523.
[376] I. Morland, V. Rolseth, L. Luna, T. Rognes, M. Bjørås, E. Seeberg, *Nucleic Acids Res* **2002**, *30*, 4926.
[377] M. Liu, S. Doublié, S. S. Wallace, *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **2013**, *743–744*, 4.
[378] K. Imamura, A. Averill, S. S. Wallace, S. Doublié, *J Biol Chem* **2012**, *287*, 4288.
[379] S. Doublié, V. Bandaru, J. P. Bond, S. S. Wallace, *Proc Natl Acad Sci U S A* **2004**, *101*, 10284.
[380] T. K. Hazra, T. Izumi, R. Venkataraman, Y. W. Kow, M. Dizdaroglu, S. Mitra, *Journal of Biological Chemistry* **2000**, *275*, 27762.
[381] M. K. Hailer, P. G. Slade, B. D. Martin, T. A. Rosenquist, K. D. Sugden, *DNA Repair* **2005**, *4*, 41.
[382] M. Liu, V. Bandaru, J. P. Bond, P. Jaruga, X. Zhao, P. P. Christov, C. J. Burrows, C. J. Rizzo, M. Dizdaroglu, S. S. Wallace, *PNAS* **2010**, *107*, 4925.

[383] V. Vartanian, I. G. Minko, S. Chawanthayatham, P. A. Egner, Y.-C. Lin, L. F. Earley, R. Makar, J. R. Eng, M. T. Camp, L. Li, M. P. Stone, M. R. Lasarev, J. D. Groopman, R. G. Croy, J. M. Essigmann, A. K. McCullough, R. S. Lloyd, *Proc Natl Acad Sci U S A* **2017**, *114*, 4207.
[384] S. Couvé-Privat, G. Macé, F. Rosselli, M. K. Saparbaev, *Nucleic Acids Res* **2007**, *35*, 5672.
[385] F. Coste, M. Ober, Y.-V. Le Bihan, M. A. Izquierdo, N. Hervouet, H. Mueller, T. Carell, B. Castaing, *Chem Biol* **2008**, *15*, 706.
[386] J. C. Fromme, G. L. Verdine, *Journal of Biological Chemistry* **2003**, *278*, 51543.
[387] S. Duclos, P. Aller, P. Jaruga, M. Dizdaroglu, S. S. Wallace, S. Doublié, *DNA Repair (Amst)* **2012**, *11*, 714.
[388] D. O. Zharkov, G. Golan, R. Gilboa, A. S. Fernandes, S. E. Gerchman, J. H. Kycia, R. A. Rieger, A. P. Grollman, G. Shoham, *EMBO J* **2002**, *21*, 789.
[389] K. Pereira de Jésus, L. Serre, C. Zelwer, B. Castaing, *Nucleic Acids Res* **2005**, *33*, 5936.
[390] B. Castaing, A. Geiger, H. Seliger, P. Nehls, J. Laval, C. Zelwer, S. Boiteux, *Nucleic Acids Research* **1993**, *21*, 2899.
[391] A. Prakash, S. Doublié, S. S. Wallace, *Prog Mol Biol Transl Sci* **2012**, *110*, 71.
[392] P. M. Hegde, A. Dutta, S. Sengupta, J. Mitra, S. Adhikari, A. E. Tomkinson, G.-M. Li, I. Boldogh, T. K. Hazra, S. Mitra, M. L. Hegde, *J Biol Chem* **2015**, *290*, 20919.
[393] S. S. Wallace, *Environmental and Molecular Mutagenesis* **2013**, *54*, 691.
[394] M. L. Hegde, C. A. Theriot, A. Das, P. M. Hegde, Z. Guo, R. K. Gary, T. K. Hazra, B. Shen, S. Mitra, *Journal of Biological Chemistry* **2008**, *283*, 27028.
[395] H. Dou, C. A. Theriot, A. Das, M. L. Hegde, Y. Matsumoto, I. Boldogh, T. K. Hazra, K. K. Bhakat, S. Mitra, *Journal of Biological Chemistry* **2008**, *283*, 3130.
[396] D. O. Zharkov, R. A. Rieger, C. R. Iden, A. P. Grollman, *J Biol Chem* **1997**, *272*, 5335.
[397] K. Sadeghian, D. Flaig, I. D. Blank, S. Schneider, R. Strasser, D. Stathis, M. Winnacker, T. Carell, C. Ochsenfeld, *Angewandte Chemie International Edition* **2014**, *53*, 10044.
[398] S. Sowlati-Hashjin, S. D. Wetmore, *J. Phys. Chem. B* **2014**, *118*, 14566.
[399] C. Zhu, L. Lu, J. Zhang, Z. Yue, J. Song, S. Zong, M. Liu, O. Stovicek, Y. Q. Gao, C. Yi, *PNAS* **2016**, *113*, 7792.
[400] M. Liu, J. Zhang, C. Zhu, X. Zhang, W. Xiao, Y. Yan, L. Liu, H. Zeng, Y. Q. Gao, C. Yi, *Nat Commun* **2021**, *12*, 4108.
[401] P. J. Teoh, O. An, T.-H. Chung, J. Y. Chooi, S. H. M. Toh, S. Fan, W. Wang, B. T. H. Koh, M. J. Fullwood, M. G. Ooi, S. de Mel, C. Y. Soekojo, L. Chen, S. B. Ng, H. Yang, W. J. Chng, *Blood* **2018**, *132*, 1304.
[402] M. S. Albelazi, P. R. Martin, S. Mohammed, L. Mutti, J. L. Parsons, R. H. Elder, *Genes (Basel)* **2019**, *10*, 315.
[403] P. R. Martin, S. Couvé, C. Zutterling, M. S. Albelazi, R. Groisman, B. T. Matkarimov, J. L. Parsons, R. H. Elder, M. K. Saparbaev, *Sci Rep* **2017**, *7*, 17438.
[404] A. Prakash, B. L. Carroll, J. B. Sweasy, S. S. Wallace, S. Doublié, *DNA Repair (Amst)* **2014**, *14*, 17.
[405] H. Mori, R. Ouchida, A. Hijikata, H. Kitamura, O. Ohara, Y. Li, X. Gao, A. Yasui, R. S. Lloyd, J.-Y. Wang, *DNA Repair (Amst)* **2009**, *8*, 1328.
[406] B. Xue, W. Xiao, H. Tian, *Biochem Biophys Res Commun* **2020**, *521*, 245.
[407] V. Vartanian, B. Lowell, I. G. Minko, T. G. Wood, J. D. Ceci, S. George, S. W. Ballinger, C. L. Corless, A. K. McCullough, R. S. Lloyd, *Proc Natl Acad Sci U S A* **2006**, *103*, 1864.
[408] T. A. Rosenquist, E. Zaika, A. S. Fernandes, D. O. Zharkov, H. Miller, A. P. Grollman, *DNA Repair (Amst)* **2003**, *2*, 581.
[409] M. J. Calkins, V. Vartanian, N. Owen, G. Kirkali, P. Jaruga, M. Dizdaroglu, A. K. McCullough, R. S. Lloyd, *DNA Repair (Amst)* **2016**, *48*, 43.
[410] W. Xue, Y. Liu, N. Xin, J. Miao, J. Du, Y. Wang, H. Shi, Y. Wei, H. Zhang, Y. Chen, Y. Gao, D. Li, Y. Feng, J. Yan, J. Zhang, N. Hou, C. Huang, J. Han, *Biomed Res Int* **2020**, *2020*, 5053975.
[411] R. D. Kennedy, C. C. Chen, P. Stuckert, E. M. Archila, M. A. De la Vega, L. A. Moreau, A. Shimamura, A. D. D'Andrea, *J Clin Invest* **2007**, *117*, 1440.
[412] A.-V. Goula, K. Merienne, *Genes* **2013**, *4*, 375.
[413] L. Møllersen, A. D. Rowe, J. L. Illuzzi, G. A. Hildrestrand, K. J. Gerhold, L. Tveterås, A. Bjølgerud, D. M. Wilson, M. Bjørås, A. Klungland, *Hum Mol Genet* **2012**, *21*, 4939.
[414] B. M. F. Hanna, M. Michel, T. Helleday, O. Mortusewicz, *Int J Mol Sci* **2021**, *22*, 4542.

[415] M. L. Hegde, P. M. Hegde, L. M. F. Holthauzen, T. K. Hazra, K. S. J. Rao, S. Mitra, *J Biol Chem* **2010**, *285*, 28812.
[416] A. C. Jacobs, M. J. Calkins, A. Jadhav, D. Dorjsuren, D. Maloney, A. Simeonov, P. Jaruga, M. Dizdaroglu, A. K. McCullough, R. S. Lloyd, *PLOS ONE* **2013**, *8*, e81667.
[417] S. M. Paul, D. S. Mytelka, C. T. Dunwiddie, C. C. Persinger, B. H. Munos, S. R. Lindborg, A. L. Schacht, *Nat Rev Drug Discov* **2010**, *9*, 203.
[418] M. Dickson, J. P. Gagnon, *Nat Rev Drug Discov* **2004**, *3*, 417.
[419] M. Hay, D. W. Thomas, J. L. Craighead, C. Economides, J. Rosenthal, *Nat Biotechnol* **2014**, *32*, 40.
[420] G. Takátsy, *Kiserl. Orvostud* **1950**, *5*, 393.
[421] G. Takátsy, *Acta Microbiol Immunol Hung* **2003**, *50*, 369.
[422] E. Farkas, 1.
[423] J. Noah, *IJHTS* **2010**, 141.
[424] D. Barh, K. Blum, M. A. Madigan, *OMICS: Biomedical Perspectives and Applications*, CRC Press, **2016**.
[425] D. A. Pereira, J. A. Williams, *Br J Pharmacol* **2007**, *152*, 53.
[426] T. Ferragamo, M. J. Wildey, *J Biomol Screen* **1999**, *4*, 175.
[427] L. M. Mayr, P. Fuerst, *J Biomol Screen* **2008**, *13*, 443.
[428] A. S. Pina, A. Hussain, A. C. A. Roque, in *Ligand-Macromolecular Interactions in Drug Discovery: Methods and Protocols*, (Ed: A. C. A. Roque), Humana Press, Totowa, NJ, **2010**, pp. 3–12.
[429] A. L. Hopkins, *Nat Chem Biol* **2008**, *4*, 682.
[430] I. Gashaw, P. Ellinghaus, A. Sommer, K. Asadullah, *Drug Discov Today* **2012**, *17 Suppl*, S24.
[431] E. B. Fauman, B. K. Rai, E. S. Huang, *Curr Opin Chem Biol* **2011**, *15*, 463.
[432] P. G. Wyatt, I. H. Gilbert, K. D. Read, A. H. Fairlamb, *Curr Top Med Chem* **2011**, *11*, 1275.
[433] J. H. Toney, A. Ogawa, M. Blair, Y.-W. Park, *ASSAY and Drug Development Technologies* **2003**, *1*, 521.
[434] E. England, P. Newton, F. Neal, L. Kitching, C. Colley, C. J. Rossant, *J Biomol Screen* **2015**, *20*, 536.
[435] C. Eggeling, L. Brand, D. Ullmann, S. Jäger, *Drug Discov Today* **2003**, *8*, 632.
[436] K. Gorshkov, M. Pradhan, M. Xu, S. Yang, E. M. Lee, C. Z. Chen, M. Shen, W. Zheng, *J Pharmacol Exp Ther* **2020**, *374*, 500.
[437] J. Oeljeklaus, F. Kaschani, M. Kaiser, *Angew Chem Int Ed Engl* **2013**, *52*, 1368.
[438] B. Wang, J. Deng, Y. Gao, L. Zhu, R. He, Y. Xu, *Fitoterapia* **2011**, *82*, 1141.
[439] J. G. Moffat, F. Vincent, J. A. Lee, J. Eder, M. Prunotto, *Nat Rev Drug Discov* **2017**, *16*, 531.
[440] W. Zheng, N. Thorne, J. C. McKew, *Drug Discov Today* **2013**, *18*, 1067.
[441] D. Stumpfe, P. Ripphausen, J. Bajorath, *Future Med Chem* **2012**, *4*, 593.
[442] A. Lavecchia, C. Di Giovanni, *Curr Med Chem* **2013**, *20*, 2839.
[443] C. Dalvit, E. Ardini, M. Flocco, G. P. Fogliatto, N. Mongelli, M. Veronesi, *J Am Chem Soc* **2003**, *125*, 14620.
[444] T. Tengel, T. Fex, H. Emtenäs, F. Almqvist, I. Sethson, J. Kihlberg, *Org. Biomol. Chem.* **2004**, *2*, 725.
[445] N. S. Troelsen, M. H. Clausen, *Chemistry* **2020**, *26*, 11391.
[446] M. O'Reilly, A. Cleasby, T. G. Davies, R. J. Hall, R. F. Ludlow, C. W. Murray, D. Tisi, H. Jhoti, *Drug Discov Today* **2019**, *24*, 1081.
[447] M. Rohman, J. Wingfield, *Methods Mol Biol* **2016**, *1439*, 47.
[448] J. Gabriel, G. Höfner, K. T. Wanner, *Frontiers in Chemistry* **2019**, *7*, 665.
[449] null Zhang, null Chung, null Oldenburg, *J Biomol Screen* **1999**, *4*, 67.
[450] R. Copeland, *Evaluation of Enzyme Inhibitors in Drug Discovery*, Wiley, 2nd edn., **2013**.
[451] A. Simeonov, M. I. Davis, in *Assay Guidance Manual*, (Eds: S. Markossian, G. S. Sittampalam, A. Grossman, K. Brimacombe, M. Arkin, D. Auld, C. P. Austin, J. Baell, J. M. M. Caaveiro, T. D. Y. Chung, N. P. Coussens, J. L. Dahlin, V. Devanaryan, T. L. Foley, M. Glicksman, M. D. Hall, J. V. Haas, S. R. J. Hoare, J. Inglese, P. W. Iversen, S. D. Kahl, S. C. Kales, S. Kirshner, M. Lal-Nag, Z. Li, J. McGee, O. McManus, T. Riss, P. Saradjian, O. J. Trask, J. R. Weidner, M. J. Wildey, M. Xia, X. Xu), Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda (MD), **2004**.
[452] F. W. Studier, *Protein Expr Purif* **2005**, *41*, 207.
[453] Y. Liu, M. H. Beresini, A. Johnson, R. Mintzer, K. Shah, K. Clark, S. Schmidt, C. Lewis, M. Liimatta, L. O. Elliott, A. Gustafson, C. E. Heise, *J Biomol Screen* **2012**, *17*, 225.
[454] H. Zipper, H. Brunner, J. Bernhagen, F. Vitzthum, *Nucleic Acids Research* **2004**, *32*, e103.

[455] V. L. Singer, L. J. Jones, S. T. Yue, R. P. Haugland, *Anal Biochem* **1997**, *249*, 228.
[456] N. Ashley, J. Poulton, *Biochem Biophys Res Commun* **2009**, *378*, 450.
[457] K. Fox, Ed., *Drug'DNA Interaction Protocols*, Humana Press, **1997**.
[458] G. Gabant, M. Cadene, *Methods* **2008**, *46*, 54.
[459] G. M. Lingaraju, A. E. Prota, F. K. Winkler, *DNA Repair* **2009**, *8*, 857.
[460] R. Giegé, C. Sauter, *HFSP J* **2010**, *4*, 109.
[461] W. Kabsch, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 125.
[462] C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack, G. Bricogne, *Acta Crystallogr D Biol Crystallogr* **2011**, *67*, 293.
[463] P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr D Biol Crystallogr* **2010**, *66*, 213.
[464] P. Emsley, K. Cowtan, *Acta Crystallogr D Biol Crystallogr* **2004**, *60*, 2126.
[465] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, *J Comput Chem* **2004**, *25*, 1605.
[466] M. Koba, A. Szostek, J. Konopa, *Acta Biochim Pol* **2007**, *54*, 883.
[467] E. Plebanek, F. Chevrier, V. Roy, T. Garenne, F. Lecaille, D. Warszycki, A. J. Bojarski, G. Lalmanach, L. A. Agrofoglio, *European Journal of Medicinal Chemistry* **2016**, *121*, 12.
[468] A. Berteotti, F. Vacondio, A. Lodola, M. Bassi, C. Silva, M. Mor, A. Cavalli, *ACS Med. Chem. Lett.* **2014**, *5*, 501.
[469] V. Bandaru, W. Cooper, S. S. Wallace, S. Doublié, *Acta Crystallogr D Biol Crystallogr* **2004**, *60*, 1142.
[470] M. van de Weert, L. Stella, *Journal of Molecular Structure* **2011**, *998*, 144.
[471] J. W. Hill, T. K. Hazra, T. Izumi, S. Mitra, *Nucleic Acids Res* **2001**, *29*, 430.
[472] U. Zander, G. Hoffmann, I. Cornaciu, J.-P. Marquette, G. Papp, C. Landret, G. Seroul, J. Sinoir, M. Röwer, F. Felisaz, S. Rodriguez-Puente, V. Mariaule, P. Murphy, M. Mathieu, F. Cipriani, J. A. Márquez, *Acta Crystallogr D Struct Biol* **2016**, *72*, 454.
[473] Y. Pommier, *ACS Chem Biol* **2013**, *8*, 82.
[474] Y. Pommier, C. Marchand, *Nat Rev Drug Discov* **2011**, *11*, 25.
[475] G. E. Greco, Y. Matsumoto, R. C. Brooks, Z. Lu, M. R. Lieber, A. E. Tomkinson, *DNA Repair (Amst)* **2016**, *43*, 18.
[476] B. Budke, J. H. Kalin, M. Pawlowski, A. S. Zelivianskaia, M. Wu, A. P. Kozikowski, P. P. Connell, *J Med Chem* **2013**, *56*, 254.
[477] B. Castaing, S. Boiteux, C. Zelwer, *Nucleic Acids Research* **1992**, *20*, 389.
[478] Y.-V. Le Bihan, M. Angeles Izquierdo, F. Coste, P. Aller, F. Culard, T. H. Gehrke, K. Essalhi, T. Carell, B. Castaing, *Nucleic Acids Res* **2011**, *39*, 6277.
[479] F. Coste, M. Ober, T. Carell, S. Boiteux, C. Zelwer, B. Castaing, *J Biol Chem* **2004**, *279*, 44074.
[480] A. Messer, D. C. Butler, *Neurobiol Dis* **2020**, *134*, 104619.
[481] M. Casula, E. Olmastroni, M. Gazzotti, F. Galimberti, A. Zambon, A. L. Catapano, *Pharmacol Res* **2020**, *160*, 105060.
[482] A. Buchfellner, L. Yurlova, S. Nüske, A. M. Scholz, J. Bogner, B. Ruf, K. Zolghadr, S. E. Drexler, G. A. Drexler, S. Girst, C. Greubel, J. Reindl, C. Siebenwirth, T. Romer, A. A. Friedl, U. Rothbauer, *PLOS ONE* **2016**, *11*, e0151041.
[483] M. Rajan, O. Mortusewicz, U. Rothbauer, F. D. Hastert, K. Schmidthals, A. Rapp, H. Leonhardt, M. C. Cardoso, *FEBS Open Bio* **2015**, *5*, 779.

Annexes

Annexe 1 : Séquence des oligonucléotides utilisés

Oligonucléotides utilisés pour la cristallographie

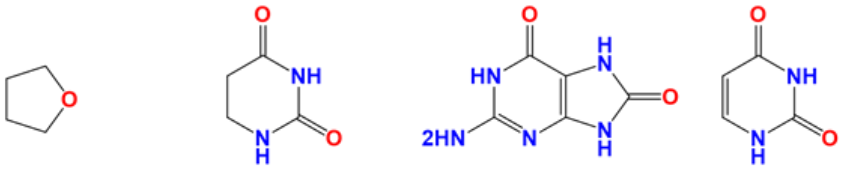
13mer_THF	CGT-CCA-XGT-CTA-C	Tm = 38,0 °C 3761,5 g.mol^{-1}
13mer_comp	TAG-ACC-TGG-ACG-G	Tm = 42,0 °C 3999,7 g.mol^{-1}
16mer_THF	AGC-GTC-CAX-GTC-TAC-C	Tm = 48,0 °C 4693,1 g.mol^{-1}
16mer_comp	TGG-TAG-ACC-TGG-ACG-C	Tm = 52,0 °C 4922,3 g.mol^{-1}

Oligonucléotides utilisés pour le criblage fluorescent

17mer_SUP_5'TAMRA_Neil	TCA-CCY-TCG-TAC-GAC-TC	Tm = 50,0 °C 5660,9 g.mol^{-1}
17mer_SUP_5'TAMRA_Ogg	TCA-CCZ-TCG-TAC-GAC-TC	Tm = 50,0 °C 5714,0 g.mol^{-1}
17mer_INF_3'_BHQ2	GAG-TCG-TAC-GAG-GGT-GA	Tm = 54,0 °C 5872,0 g.mol^{-1}

Oligonucléotides utilisés pour le test d'activité radioactif

24mer_DHU	CTG-ATC-GAT-GAY-GCC-TGA-CAT-GAT	Tm = 50,6 °C 7355,9 g.mol^{-1}
24mer_Oxog	CTG-ATC-GAT-GAC-ZCC-TGA-CAT-GAT	Tm = 50,6 °C 7369,8 g.mol^{-1}
24mer_U	CTG-ATC-GAT-GAC-UCC-TGA-CAT-GAT	Tm = 50,6 °C 7313,8 g.mol^{-1}
24mer_comp	ATC-ATG-TCA-GGC-GTC-ATC-GAT-CAG	Tm = 52,3 °C 7352,8 g.mol^{-1}



X : THF (Tétrahydrofurane) — Y : DHU (5,6 Dihydrouracile) — Z : 8-oxoG (7,8-Dihydro-8-oxoguanine) — U : Uracile

Hybridation de simples brins

13mer_THF + 13mer_comp	→	**13mer_NEIL1**
16mer_THF + 16mer_comp	→	**16mer_OGG1**
17mer_SUP_5'TAMRA_Neil + 17mer_INF_3'_BHQ2	→	**TMR_NEIL1**
17mer_SUP_5'TAMRA_Ogg + 17mer_INF_3'_BHQ2	→	**TMR_OGG1**
24mer_DHU + 24mer_comp	→	**Act_NEIL1**
24mer_Oxog + 24mer_comp	→	**Act_OGG1**
24mer_U + 24mer_comp	→	**Act_ApLyase**

Annexe 2 : Plan des plaques de criblage et composition des chimiothèques

Plan de plaque pour les plaques stocks d'inhibiteurs

	1	2	3	4	5	6	7	8	9	10	11	12
A		Mol 1	Mol 2	Mol 3	Mol 4	Mol 5	Mol 6	Mol 7	Mol 8	Mol 9	Mol 10	
B		Mol 11	Mol 12	Mol 13	Mol 14	Mol 15	Mol 16	Mol 17	Mol 18	Mol 19	Mol 20	
C		Mol 21	Mol 22	Mol 23	Mol 24	Mol 25	Mol 26	Mol 27	Mol 28	Mol 29	Mol 30	
D		Mol 31	Mol 32	Mol 33	Mol 34	Mol 35	Mol 36	Mol 37	Mol 38	Mol 39	Mol 40	
E		Mol 41	Mol 42	Mol 43	Mol 44	Mol 45	Mol 46	Mol 47	Mol 48	Mol 49	Mol 50	
F		Mol 51	Mol 52	Mol 53	Mol 54	Mol 55	Mol 56	Mol 57	Mol 58	Mol 59	Mol 60	
G		Mol 61	Mol 62	Mol 63	Mol 64	Mol 65	Mol 66	Mol 67	Mol 68	Mol 69	Mol 70	
H		Mol 71	Mol 72	Mol 73	Mol 74	Mol 75	Mol 76	Mol 77	Mol 78	Mol 79	Mol 80	

Plan de plaque pour les plaques filles d'inhibiteurs

	1	2	3	4	5	6	7	8	9	10	11	12
A	Purpurine	Mol 1	Mol 2	Mol 3	Mol 4	Mol 5	Mol 6	Mol 7	Mol 8	Mol 9	Mol 10	
B	H2O	Mol 11	Mol 12	Mol 13	Mol 14	Mol 15	Mol 16	Mol 17	Mol 18	Mol 19	Mol 20	
C	Purpurine	Mol 21	Mol 22	Mol 23	Mol 24	Mol 25	Mol 26	Mol 27	Mol 28	Mol 29	Mol 30	
D	DMSO	Mol 31	Mol 32	Mol 33	Mol 34	Mol 35	Mol 36	Mol 37	Mol 38	Mol 39	Mol 40	
E	Juglone	Mol 41	Mol 42	Mol 43	Mol 44	Mol 45	Mol 46	Mol 47	Mol 48	Mol 49	Mol 50	
F	H2O	Mol 51	Mol 52	Mol 53	Mol 54	Mol 55	Mol 56	Mol 57	Mol 58	Mol 59	Mol 60	
G	Juglone	Mol 61	Mol 62	Mol 63	Mol 64	Mol 65	Mol 66	Mol 67	Mol 68	Mol 69	Mol 70	
H	DMSO	Mol 71	Mol 72	Mol 73	Mol 74	Mol 75	Mol 76	Mol 77	Mol 78	Mol 79	Mol 80	

Plan de plaque pour la mise en place du criblage primaire

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A																								
B																								
C			LA-1					DMSO					LA-24					S2					Contrôle catalytique	
D			DMSO					DMSO					H2O					SCL						
E			LA-26					H2O					H2O					S2						
F			LA-24					2TX					S2					SCL						
G			SCL					2TX					LA-1					LA-1						
H			2TX					LA-26					LA-24					LA-26						
I																							Contrôle bruit de fond	
J																								
K																								
L																								
M																								
N																								
O																								
P																								

	Contrôle sans inhibiteur
	Contrôle avec inhibiteurs
	Composés à tester
	Signal cataltytiqe (protéine en excès)
	Bruit de fond

Plan de plaque pour le criblage manuel

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A				Inhib 1					Inhib 2					Inhib 3										
B				2TX					Inhib 4					Inhib 5					Contrôle catalytique					
C				Inhib 6					Inhib 7					Inhib 8										
D				H2O					Inhib 9					Inhib 10										
E				Inhib 11					Inhib 12					Inhib 13										
F				2TX					Inhib 14					Inhib 15										
G				Inhib 16					Inhib 17					Inhib 18										
H				Buffer					Inhib 19					Inhib 20										
I				Inhib 21					Inhib 22					Inhib 23					Contrôle bruit de fond					
J				2TX					Inhib 24					Inhib 25										
K				Inhib 26					Inhib 27					Inhib 28										
L				H2O					Inhib 29					Inhib 30										
M				Inhib 31					Inhib 32					Inhib 33										
N				2TX					Inhib 34					Inhib 35										
O				Inhib 36					Inhib 37					Inhib 38										
P				Buffer					Inhib 39					Inhib 40										

Plan de plaque pour le criblage semi-automatique

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A		Neg Cont			Inhib 1				Inhib 2			Inhib 3			Inhib 4			Inhib 5		Contrôle catalytique				
B		Inhib 6			Inhib 7				Inhib 8			Inhib 9			Inhib 10									
C		Pos Cont			Inhib 11				Inhib 12			Inhib 13			Inhib 14			Inhib 15						
D		Inhib 16			Inhib 17				Inhib 18			Inhib 19			Inhib 20									
E		Neg Cont			Inhib 21				Inhib 22			Inhib 23			Inhib 24			Inhib 25						
F		Inhib 26			Inhib 27				Inhib 28			Inhib 29			Inhib 30									
G		Pos Cont			Inhib 31				Inhib 32			Inhib 33			Inhib 34			Inhib 35						
H		Inhib 36			Inhib 37				Inhib 38			Inhib 39			Inhib 40									
I		Neg Cont			Inhib 41				Inhib 42			Inhib 43			Inhib 44			Inhib 45		Contrôle bruit de fond				
J		Inhib 46			Inhib 47				Inhib 48			Inhib 49			Inhib 50									
K		Pos Cont			Inhib 51				Inhib 52			Inhib 53			Inhib 54			Inhib 55						
L		Inhib 56			Inhib 57				Inhib 58			Inhib 59			Inhib 60									
M		Neg Cont			Inhib 61				Inhib 62			Inhib 63			Inhib 64			Inhib 65						
N		Inhib 66			Inhib 67				Inhib 68			Inhib 69			Inhib 70									
O		Pos Cont			Inhib 71				Inhib 72			Inhib 73			Inhib 74			Inhib 75						
P		Inhib 76			Inhib 77				Inhib 78			Inhib 79			Inhib 80									

	Contrôle sans inhibiteur
	Contrôle avec inhibiteurs
	Composés à tester
	Signal cataltytiqe (protéine en excès)
	Bruit de fond

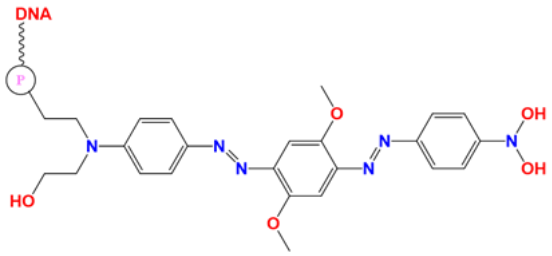
Chimiothèques				
Provenance	GreenPharma	Pretwick	ICOA	
Nom	GPNCL	Prestwick Chemical Library	ICOA-DIVER	ICOA-TT
Type	Banque de diversité	Banque de diversité	Banque de diversité	Banque totale
Abréviation	GP	Prest	ICOA1	ICOA2
Nombres de molécules	480	1120	480	2480
Concentration plaque mère	2 $mg.mL^{-1}$	10 mM	0,3 $mg.mL^{-1}$	20 mM
Concentration plaque filles	2 $mg.mL^{-1}$	1,25 mM	0,3 $mg.mL^{-1}$	1,25mM
Concentration crible	160 $\mu g.mL^{-1}$	50 µM (x 2)	24 $\mu g.mL^{-1}$	50 µM (x 2)

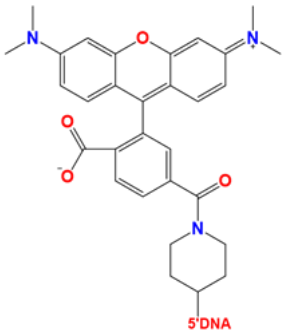
Annexe 3 : Spectre d'émission du TAMRA et d'absorption de BHQ2



BHQ 2 3' – Black Hole Quencher 2
MW = 556,47 g.mol^{-1}
Absorbance max : 579 nm
Quenching range : 550-650 nm

TAMRA 5' – Carboxytetramethylrhodamine
MW = 527,53 g.mol-1
Emission max : 580 nm
Absorbance max : 565 nm





Annexe 4 : Séquences des protéines utilisées

hNEIL1

MPEGPELHLASQFVNEACRALVFGGCVEKSSVSRNPEVPFESSAYRISASARGKELRLILSPLPGAQPQQEPLAL
VFRFGMSGSFQLVPREELPRHAHLRFYTAPPGPRLALCFVDIRRFGRWDLGGKWQPGRGPCVLQEYQQFRENVLR
NLADKAFDRPICEALLDQRFFNGIGNYLRAEILYRLKIPPFEKARSVLEALQQHRPSPELTLSQKIRTKLQNPDL
LELCHSVPKEVVQLGGKGYGSESGEEDFAAFRAWLRCYGMPGMSSLQDRHGRTIWFQGDPGPLAPKGRKSRKKKS
KATQLSPEDRVEDALPPSKAPSRTRRAKRDLPKRTATQRPEGTSLQQDPEAPTVPKKGRRKGRQAASGHCRPRKV
KADIPSLEPEGTSASLEHHHHHH

Tag 6-Histidine

Séquence tronquée pour hNEIL1dc56

Séquence tronquée pour hNEIL1dc95

	hNEIL1 dc56	hNEIL1 dc95	hNEIL1 WT + TAG
Masse moléculaire	38802,51 g.mol^{-1}	34320,33 g.mol^{-1}	44590,9 g.mol^{-1}
Acides aminés	342	303	398
pI théorique	9,80	9,21	9,92

mvNei1

MPEGPEVALTADILEKYFKGKTLEYIDFISGRYSKSEPEGYDDFIANLPLKVSNVDTKGKFLWFELFDPNDKSNK
WYIWNTFGLTGMWSLFEAKYTRAVLSFDNELMAYFSDMRNFGTFKFSNSEKELKRKLNELGPDFLKNDDIDISKI
KKYKQPIVALLMDQKKIGSGLGNYLVAEILYRAKIDPHKLGSNLTDQEIENLWYWIKYETKLAYDSNHIGYMVNL
ENESSKIGRKNYHPNIHPTEKEFDFLVYRKKKDPNGNKVIADKIIGSGKNKRTTYWAPAIQKLEHH HHHH

LlFpg

MPELPEVETVRRELEKRIVGQKIISIEATYPRMVLTGFEQLKKELTGKTIQGISRRGKYLIFEIGDDFRLISHLR
MEGKYRLATLDAPREKHDHLTMKFADGQLIYADVRKFGTWELISTDQVLPYFLKKKIGPEPTYEDFDEKLFREKL
RKSTKKIKPYLLEQTLVAGLGNIYVDEVLWLAKIHPEKETNQLIESSIHLLHDSIIEILQKAIKLGGSSIRTYSA
LGSTGKMQNELQVYGKTGEKCSRCGAEIQKIKVAGRGTHFCPVCQQK

	LlFpg WT
Masse moléculaire	31064,1 g.mol^{-1}
Acides aminés	272
pI théorique	xx

hOGG1

MHHHHHHSSGLVPRGSGMKETAAAKFERQHMDSPDLGTDDDDKAMADIGSEMPARALLPRRMGHRTLASTPALWA
SIPCPRSELRLDLVLPSGQSFRWREQSPAHWSGVLADQVWTLTQTEEQLHCTVYRGDKSQASRPTPDELEAVRKY
FQLDVTLAQLYHHWGSVDSHFQEVAQKFQGVRLLRQDPIECLFSFICSSNNNIARITGMVERLCQAFGPRLIQLD
DVTYHGFPSLQALAGPEVEAHLRKLGLGYRARYVSASARAILEEQGGLAWLQQLRESSYEEAHKALCILPGVGTK
VADCICLMALDKPQAVPVDVHMWHIAQRDYSWHPTTSQAKGPSPQTNKELGNFFRSLWGPYAGWAQAVLFSADLR
QSRHAQEPPAKRRKGSKGPEG

Tag 6-Histidine
Site coupure de l'entérokinase

Séquence tronquée pour dhOGG1

Séquence tronquée pour d2hOGG1

d2hOGG1

MGSSHHHHHHSSGLVPRGSHMRTLASTPALWASIPCPRSELRLDLVLPSGQSFRWREQSPAHWSGVLADQVWTLT
QTEEQLHCTVYRGDKSQASRPTPDELEAVRKYFQLDVTLAQLYHHWGSVDSHFQEVAQKFQGVRLLRQDPIECLF
SFICSSNNNIARITGMVERLCQAFGPRLIQLDDVTYHGFPSLQALAGPEVEAHLRKLGLGYRARYVSASARAILE
EQGGLAWLQQLRESSYEEAHKALCILPGVGTKVADCICLMALDKPQAVPVDVHMWHIAQRDYSWHPTTSQAKGPS
PQTNKELGNFFRSLWGPYAGWAQAVLFSADL

Tag 6-Histidine
Site coupure de la thrombine

dhOGG1

MHHHHHHSSGLVPRGSGMKETAAAKFERQHMDSPDLGTDDDDKAMADIGSEGHRTLASTPALWASIPCPRSELRL
DLVLPSGQSFRWREQSPAHWSGVLADQVWTLTQTEEQLHCTVYRGDKSQASRPTPDELEAVRKYFQLDVTLAQLY
HHWGSVDSHFQEVAQKFQGVRLLRQDPIECLFSFICSSNNNIARITGMVERLCQAFGPRLIQLDDVTYHGFPSLQ
ALAGPEVEAHLRKLGLGYRARYVSASARAILEEQGGLAWLQQLRESSYEEAHKALCILPGVGTKVADCICLMALD
KPQAVPVDVHMWHIAQRDYSWHPTTSQAKGPSPQTNKELGNFFRSLWGPYAGWAQAVLFSADLRQ

	d2hOGG1	dhOGG1	hOGG1 WT
Masse moléculaire	35267,07 $g.mol^{-1}$	36107,97 $g.mol^{-1}$	39557,03 $g.mol^{-1}$
Acides aminés	314	322	353
pI théorique	6,67	6,33	8,61

mOGG1

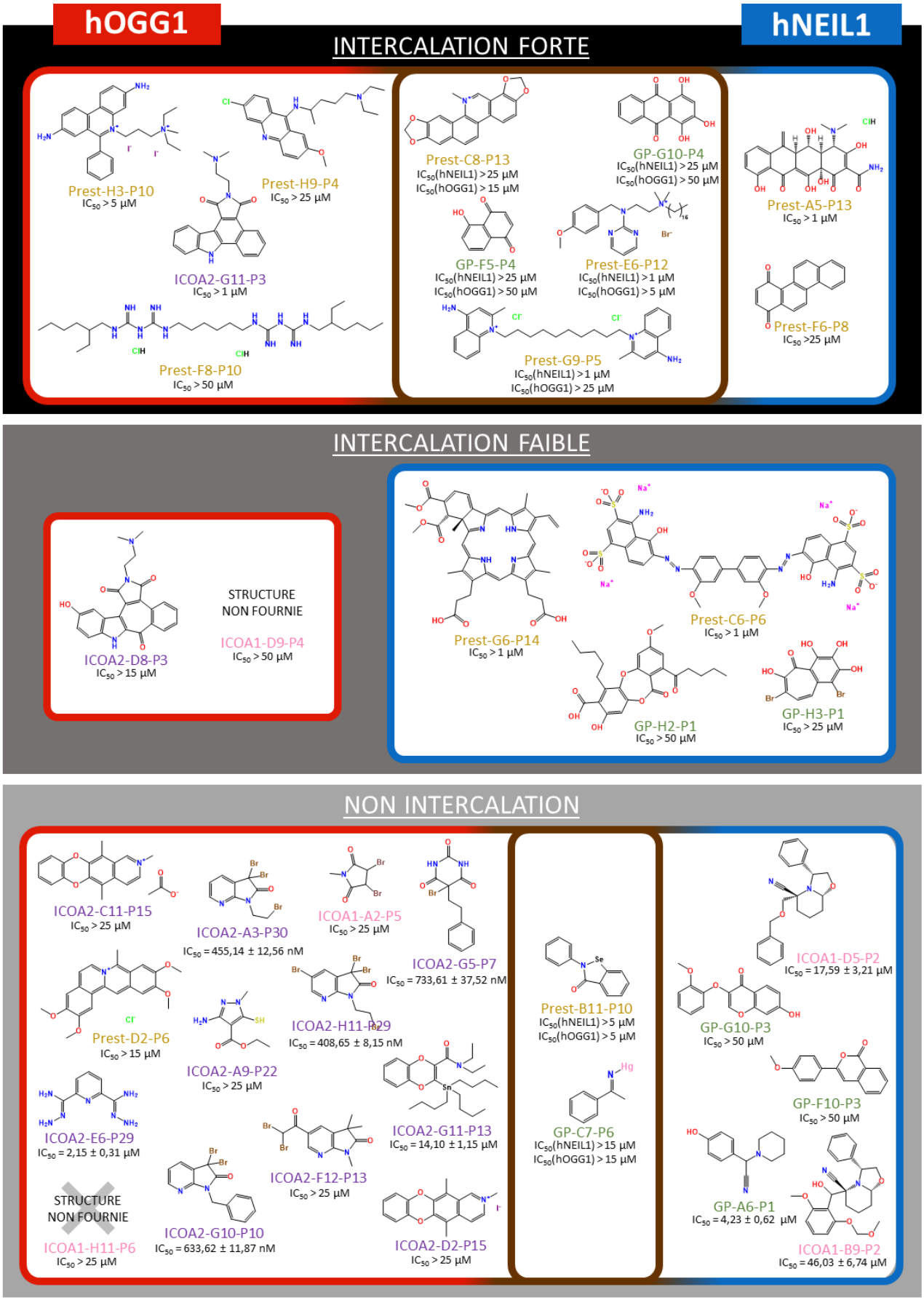
MLFRSWLPSSMRHRTLSSSPALWASIPCPRSELRLDLVLASGQSFRWKEQSPAHWSGVLADQVWTLTQTEDQLYC
TVYRGDDSQVSRPTLEELETLHKYFQLDVSLAQLYSHWASVDSHFQRVAQKFQGVRLLRQDPTECLFSFICSSNN
NIARITGMVERLCQAFGPRLIQLDDVTYHGFPNLHALAGPEAETHLRKLGLGYRARYVRASAKAILEEQGGPAWL
QQLRVAPYEEAHKALCTLPGVGAKVADCICLMALDKPQAVPVDVHVWQIAHRDYGWHPKTSQAKGPSPLANKELG
NFFRNLWGPYAGWAQAVLFSADLRQPSLSREPPAKRKKGSKRPEG

PabAGOG

GSHMIARIIGEIGIEGARFIEENIDEQFKALRYLSKGIDSETFVKLVIANSLVSYQLTGKGEQWWWEFAKYFYGR
DVKSIYLAYKEFLPNSRFNRRLIPQKLSRIRRVETFLSTLTEERIEEYYGDMSSLWGSIARALGVDKESKTVVFS
VKMFGYAARIVLSTFNPYPMEIPIPEDSRIVKLTKKLTNEKPRKFWMKIARESGVPPLHIDSILWPLLGGASIDS
APPELRDKLAELIKIIR

	mOGG1 WT	PabAGOG WT
Masse moléculaire	38883 $g.mol^{-1}$	27943 $g.mol^{-1}$
Acides aminés	345	xx
pI théorique	xx	9,5

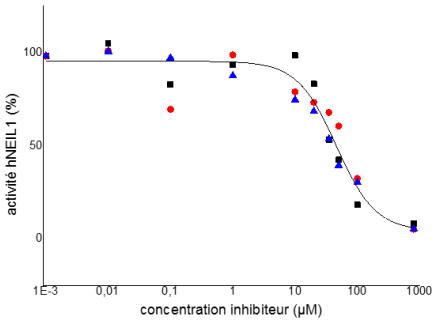
Annexe 5 : Structure des hits confirmés et IC_{50} estimés

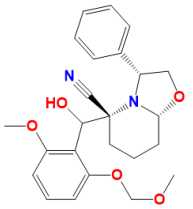


Annexe 6 : Courbes permettant le calcul des IC_{50} des meilleurs inhibiteurs

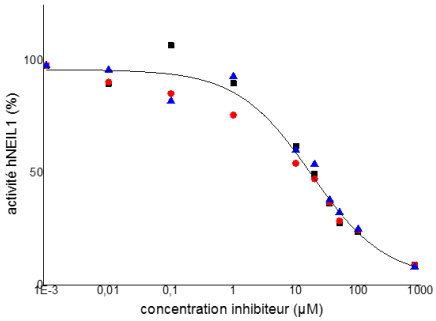
hNEIL1

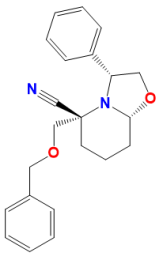
ICOA1-B9-P2
IC_{50} = 46,03 ± 6,74 µM
R^2 ajusté = 0,92747



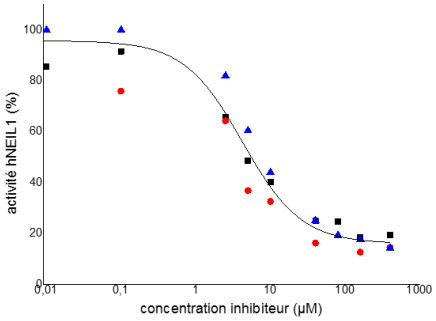


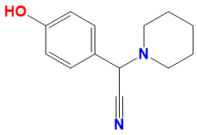
ICOA1-D5-P2
IC_{50} = 17,59 ± 3,21 µM
R^2 ajusté = 0,97437





GP-A6-P1 // Girgensonine
IC_{50} = 4,23 ± 0,63 µM
R^2 ajusté = 0,95857

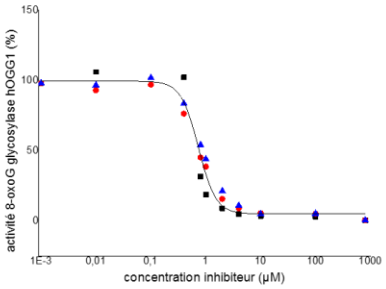




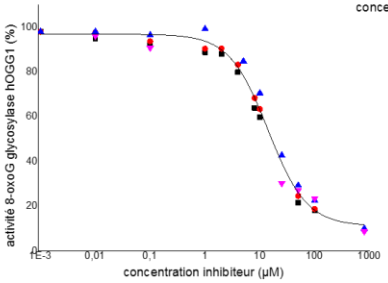
hOGG1

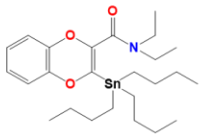


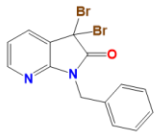
« ICOA2-G5-P7 »
IC_{50} = 733,61 ± 37,52 nM
R^2 ajusté = 0,97541



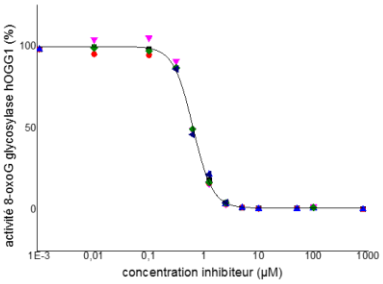
ICOA2-G11-P13
IC_{50} = 14,10 ± 1,15 µM
R^2 ajusté = 0,98683



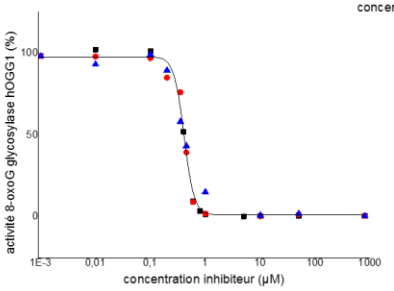


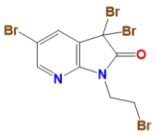


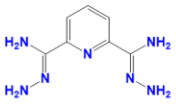
ICOA2-G10-P10
IC_{50} = 633,62 ± 11,87 nM
R^2 ajusté = 0,99744



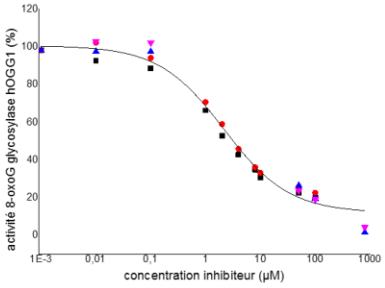
ICOA2-H11-P29
IC_{50} = 408,65 ± 8,15 nM
R^2 ajusté = 0,9903



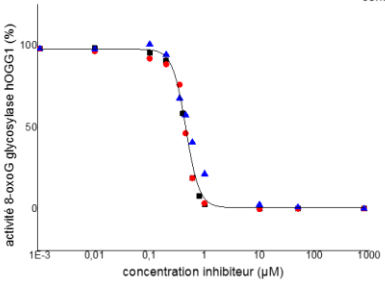


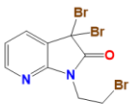


ICOA2-E6-P29
IC_{50} = 2,15 ± 0,31 µM
R^2 ajusté = 0,98414



ICOA2-A3-P30
IC_{50} = 455,14 ± 12,56 nM
R^2 ajusté = 0,98659

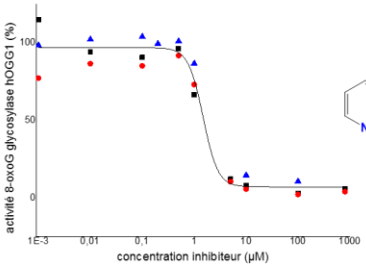


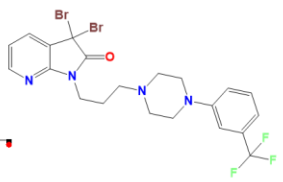


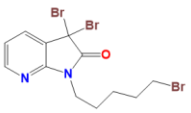
C-02H08

IC_{50} = 1,48 ± 0,32 µM

R^2 ajusté = 0,96608



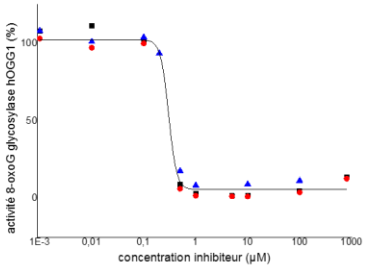




C-02G10

IC_{50} = 300,65 ± 21,88 nM

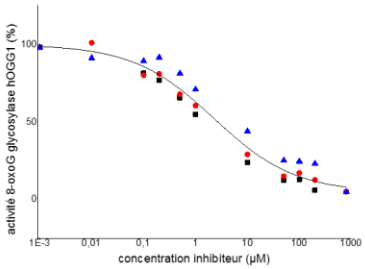
R^2 ajusté = 0,99165

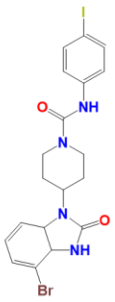


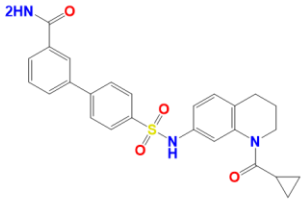
TH5487

IC_{50} = 2,27 ± 0,66 µM

R^2 ajusté = 0,97146



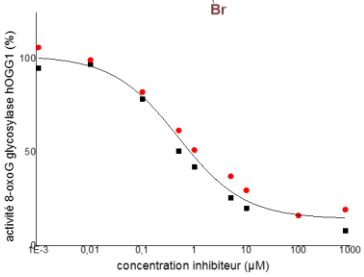




SU268

IC_{50} = 491,81 ± 105,4 nM

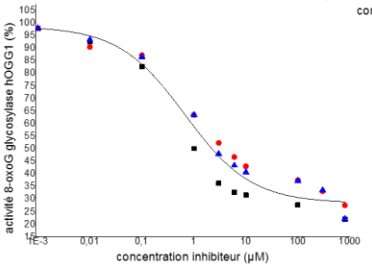
R^2 ajusté = 0,9781



O8

IC_{50} = 673,89 ± 155,7 nM

R^2 ajusté = 0,968

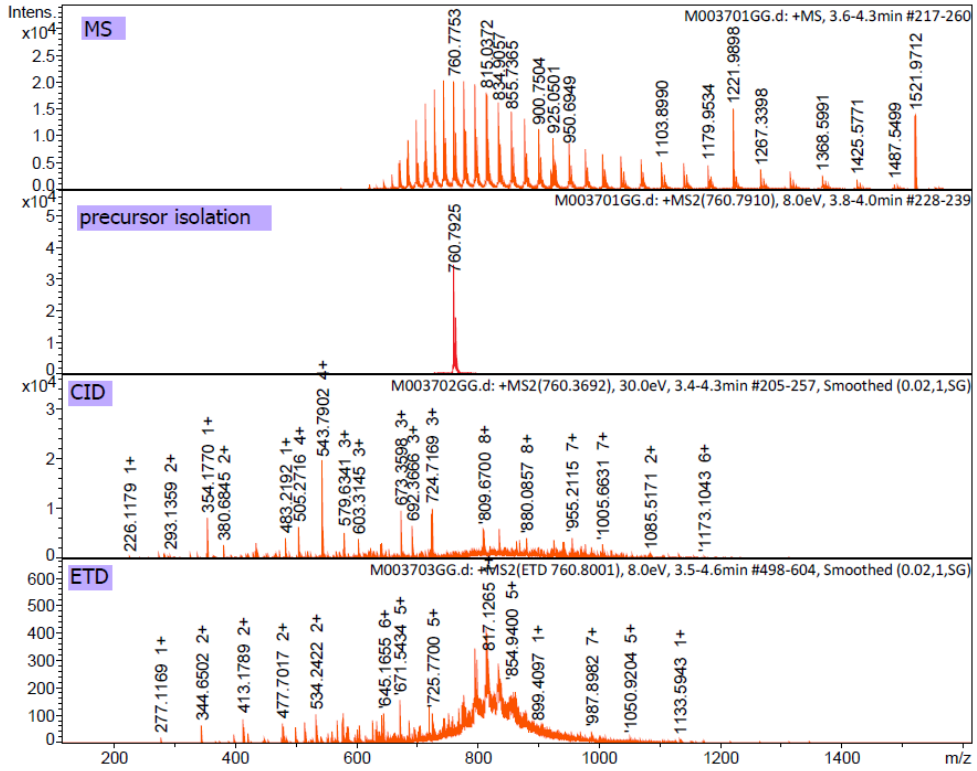




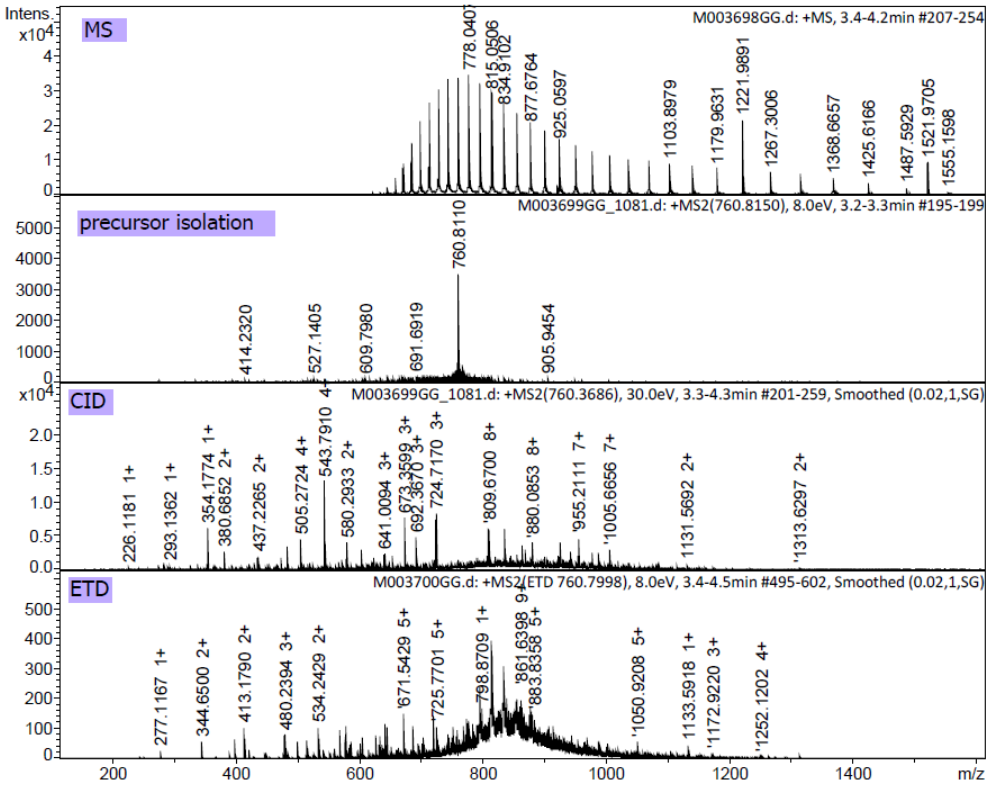
Annexe 7 : Les résultats de l'analyse de fragmentation de hNEIL1

Approche top-down

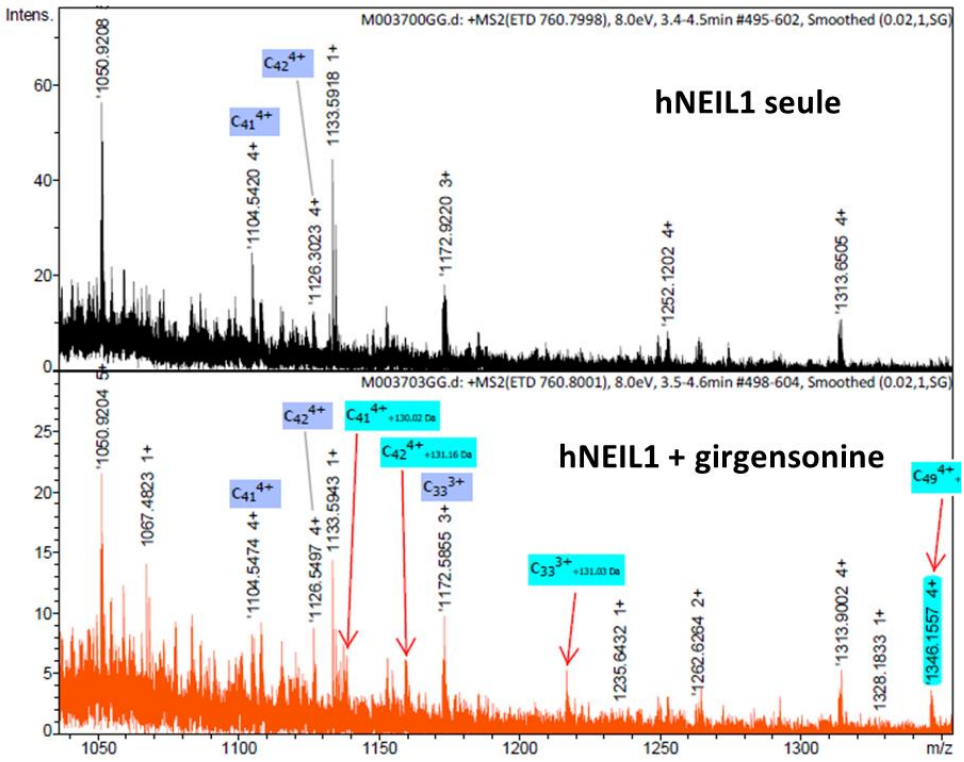
hNEIL1



hNEIL1+ Girgensonine



ETD hNEIL1dc95



MALDI-TOF MS

[Trypsine]

Peptides non modifiés | **Peptides modifiés**

Contrôle

#	Mass	Matching sequence (± error)
1	2244.299	E[90-108]R (-0.13)
2	1619.824	H[95-108]R (+0.02)
3	2192.222	F[100-118]R (-0.05)
4	3643.567	C[261-293]K (+0.21)
5	2218.300	H[274-293]K (-0.11)

Fragment 1-18 et 133-144 observés modifiés (+25Daltons // Ajout du nitrile +CN)

D5-2

#	Mass	Matching sequence (± error)
1	1996.963	P[1-18]R (-0.02)
2	1619.836	H[95-108]R (+0.01)
3	3643.613	C[261-293]K (+0.16)

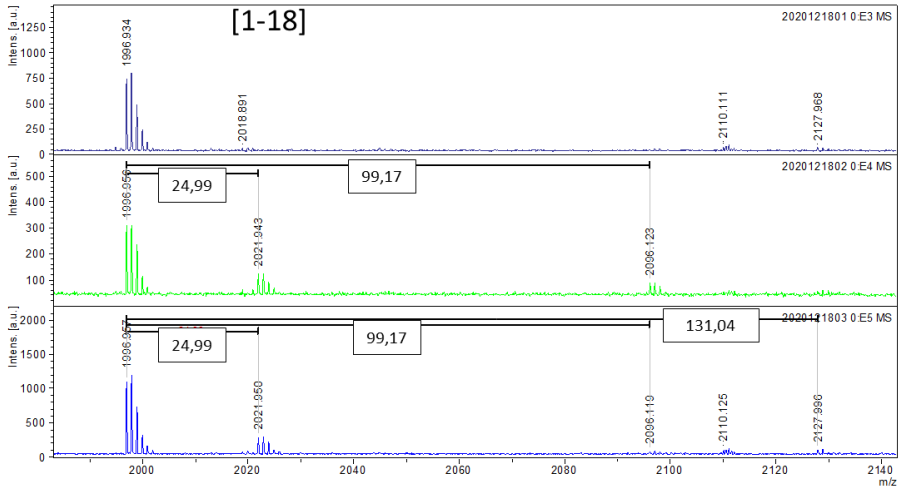
Girgensonine

#	Mass	Matching sequence (± error)
1	1996.970	P[1-18]R (-0.02)
2	1619.841	H[95-108]R (+0.01)
3	2192.259	F[100-118]R (-0.09)
4	1467.723	G[133-144]R (-0.02)
5	3643.600	C[261-293]K (+0.18)

MALDI-TOF MS

Trypsine

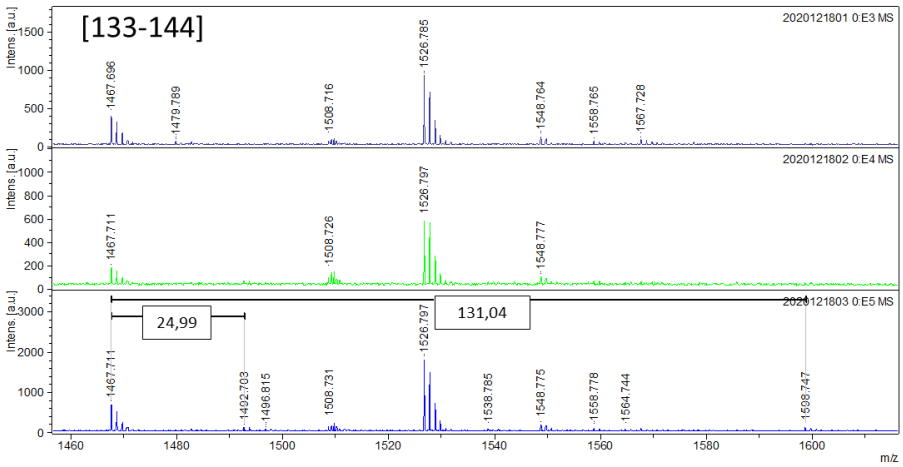
Peptides modifiés + CN (+25 Da) et autres



Contrôle

D5-2

Girgensonine



Contrôle

D5-2

Girgensonine

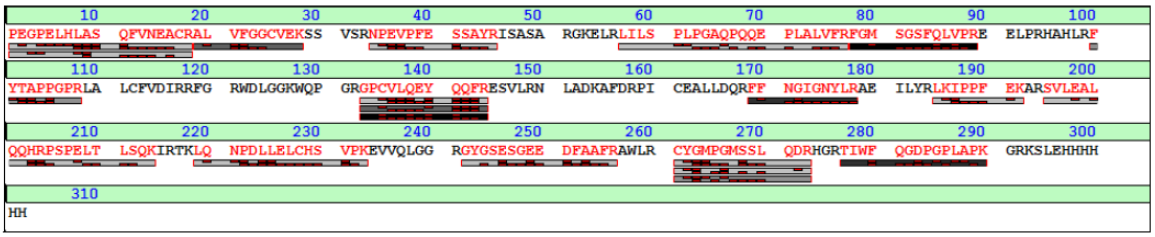
Annexe 7b : Les résultats de l'analyse MS-MS des peptides de hNEIL1

Protein Report
RecProt_Tr_IT_CID_2021-10-15 11:17:15

Results

Protein 1: >sp|hNEIL|hNEILdc95

Accession:	sp|hNEIL|hNEILdc95	**Score:**	1154.1
Database:	RecProt	**Seq. Coverage [%]:**	62.6
MW [kDa] / pI:	34.2 / 9.2	**No. of Peptides:**	19
Modification(s):	Girg-Frag2, Girg-Frag1, Oxidation		

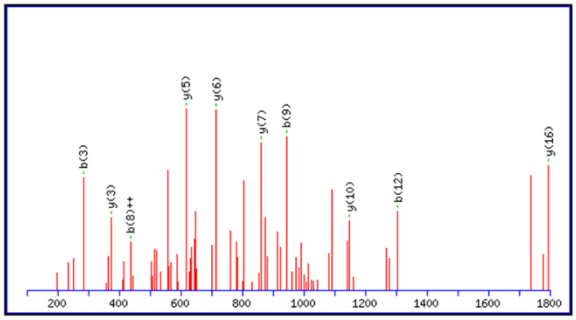


MS/MS Peptide Matches

Cmpd.	m/z meas.	z	Δ m/z [ppm]	Δ m/z [Da]	Rt [min]	Score	No. of Cmpds.	Site [%]	P	Sequence	Modification	Range	Type
41	666.8742	3	827.18	0.5512	10.79	117.33	2	0.0	0	-.PEGPELHLASQFVNEACR.A		1-18	CID
42	674.9656	3	460.71	0.3108	10.82	37.49	1	0.0	0	-.PEGPELHLASQFVNEACR.A	Girg-Frag2: 17	1-18	CID
15	511.7565	2	-27.56	-0.0141	9.89	31.14	1	0.0	0	R.ALVFGGCVEK.S		19-28	CID
14	698.2988	2	-45.54	-0.0318	9.88	64.45	3	0.0	0	R.NPEVPFESSAYR.I		34-45	CID
82	1144.0857	2	-66.92	-0.0766	12.47	95.18	4	0.0	0	R.LILSPLPGAQPQQEPLALVFR.F		57-77	CID
47	663.2908	2	-69.91	-0.0464	10.95	72.26	1	0.0	0	R.FGMSGSFQLVPR.E		78-89	CID
4	503.2719	2	21.14	0.0106	9.36	28.51	1	0.0	0	R.FYTAPPGPR.L		100-108	CID
32	734.3347	2	-29.13	-0.0214	10.51	71.01	1	0.0	0	R.GPCVLQEYQQFR.E		133-144	CID
30	747.2932	2	588.45	0.4395	10.43	56.55	1	0.0	0	R.GPCVLQEYQQFR.E	Girg-Frag2: 3	133-144	CID
50	799.7995	2	-93.95	-0.0752	11.05	79.8	1	0.0	0	R.GPCVLQEYQQFR.E	Girg-Frag1: 3	133-144	CID
49	600.7954	2	-27.04	-0.0162	11.00	45.89	3	0.0	0	R.FFNGIGNYLR.A		169-178	CID
5	486.3233	2	48.22	0.0234	9.43	40.53	4	0.0	1	R.LKIPPFEK.A		185-192	CID
17	1131.0438	2	-63.87	-0.0722	9.96	24.74	1	0.0	0	R.SVLEALQQHRPSPELTLSQK.I		195-214	CID
36	853.4206	2	-35.29	-0.0301	10.66	85.92	1	0.0	0	K.LQNPDLLELCHSVPK.E		219-233	CID
26	811.2907	2	-63.10	-0.0512	10.26	110.17	4	0.0	0	R.GYGSESGEEDFAAFR.A		242-256	CID
27	735.2824	2	-26.86	-0.0197	10.27	44.47	2	0.0	0	R.CYGMPGMSSLQDR.H	Girg-Frag2: 1	261-273	CID
10	743.2326	2	-90.14	-0.0670	9.64	28.16	5	0.0	0	R.CYGMPGMSSLQDR.H	Girg-Frag2: 1; Oxidation: 7	261-273	CID
8	743.2531	2	-62.56	-0.0465	9.56	44.15	4	0.997:	0	R.CYGMPGMSSLQDR.H	Girg-Frag2: 1; Oxidation: 4	261-273	CID
52	763.8784	2	-33.16	-0.0253	11.12	76.4	8	0.0	0	R.TIWFQGDPGPLAPK.G		277-290	CID

hNEIL1/girgensonine Trypsine CID
Fragment [1-18] / Adduit 25 Da

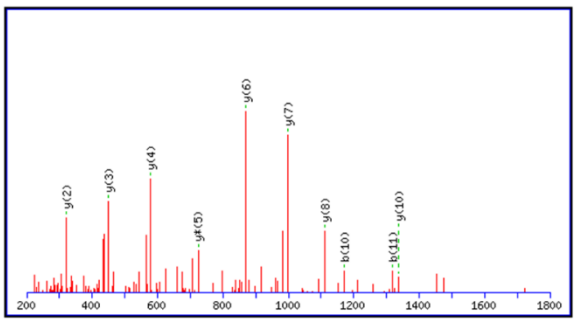
[1 – 18] PEGPELHLASQFVNEACR (Ions score 37)



#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	98.0600	49.5337					P							18
2	227.1026	114.0550			209.0921	105.0497	E	1924.8970	962.9521	1907.8705	954.4389	1906.8865	953.9469	17
3	*284.1241*	142.5657			266.1135	133.5604	G	*1795.8544*	898.4309	1778.8279	889.9176	1777.8439	889.4256	16
4	381.1769	191.0921			363.1663	182.0868	P	1738.8330	869.9201	1721.8064	861.4068	1720.8224	860.9148	15
5	510.2195	255.6134			492.2089	246.6081	E	1641.7802	821.3937	1624.7537	812.8805	1623.7696	812.3885	14
6	623.3035	312.1554			605.2930	303.1501	L	1512.7376	756.8724	1495.7111	748.3592	1494.7270	747.8672	13
7	760.3624	380.6849			742.3519	371.6796	H	1399.6535	700.3304	1382.6270	691.8171	1381.6430	691.3251	12
8	873.4465	437.2269			855.4359	428.2216	L	1262.5946	631.8010	1245.5681	623.2877	1244.5841	622.7957	11
9	*944.4836*	472.7454			926.4730	463.7402	A	*1149.5106*	575.2589	1132.4840	566.7456	1131.5000	566.2536	10
10	1031.5156	516.2615			1013.5051	507.2562	S	1078.4735	539.7404	1061.4469	531.2271	1060.4629	530.7351	9
11	1159.5742	580.2907	1142.5477	571.7775	1141.5636	571.2855	Q	991.4414	496.2244	974.4149	487.7111	973.4309	487.2191	8
12	*1306.6426*	653.8250	1289.6161	645.3117	1288.6321	644.8197	F	*863.3828*	432.1951	846.3563	423.6818	845.3723	423.1898	7
13	1405.7110	703.3592	1388.6845	694.8459	1387.7005	694.3539	V	*716.3144*	358.6609	699.2879	350.1476	698.3039	349.6556	6
14	1519.7540	760.3806	1502.7274	751.8673	1501.7434	751.3753	N	*617.2460*	309.1266	600.2195	300.6134	599.2355	300.1214	5
15	1648.7966	824.9019	1631.7700	816.3886	1630.7860	815.8966	E	503.2031	252.1052	486.1765	243.5919	485.1925	243.0999	4
16	1719.8337	860.4205	1702.8071	851.9072	1701.8231	851.4152	A	*374.1605*	187.5839	357.1340	179.0706			3
17	1847.8381	924.4227	1830.8116	915.9094	1829.8275	915.4174	C	303.1234	152.0653	286.0968	143.5521			2
18							R	175.1190	88.0631	158.0924	79.5498			1

hNEIL1/girgensonine Trypsine CID
Fragment [133-144] / Adduit 25 Da

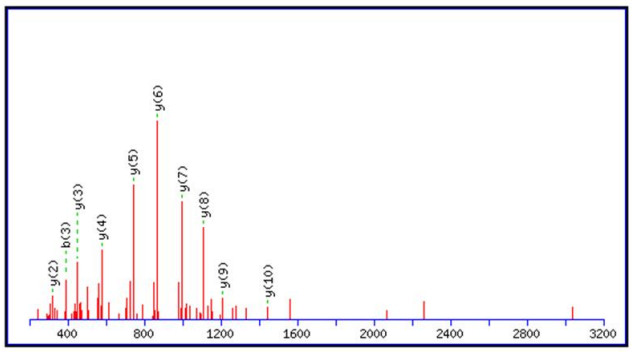
[133 – 144] GPCVLQEYQQFR (Ions score 57)



#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	58.0287	29.5180					G							12
2	155.0815	78.0444					P	1435.6787	718.3430	1418.6521	709.8297	1417.6681	709.3377	11
3	283.0859	142.0466					C	*1338.6259*	669.8166	1321.5994	661.3033	1320.6154	660.8113	10
4	382.1544	191.5808					V	1210.6215	605.8144	1193.5950	597.3011	1192.6109	596.8091	9
5	495.2384	248.1228					L	*1111.5531*	556.2802	1094.5265	547.7669	1093.5425	547.2749	8
6	623.2970	312.1521	606.2704	303.6389			Q	*998.4690*	499.7381	981.4425	491.2249	980.4585	490.7329	7
7	752.3396	376.6734	735.3130	368.1602	734.3290	367.6681	E	*870.4104*	435.7089	853.3839	427.1956	852.3999	426.7036	6
8	915.4029	458.2051	898.3764	449.6918	897.3924	449.1998	Y	741.3679	371.1876	724.3413	362.6743			5
9	1043.4615	522.2344	1026.4349	513.7211	1025.4509	513.2291	Q	*578.3045*	289.6559	561.2780	281.1426			4
10	1171.5201	586.2637	1154.4935	577.7504	1153.5095	577.2584	Q	*450.2459*	225.6266	433.2194	217.1133			3
11	1318.5885	659.7979	1301.5619	651.2846	1300.5779	650.7926	F	*322.1874*	161.5973	305.1608	153.0840			2
12							R	175.1190	88.0631	158.0924	79.5498			1

hNEIL1/girgensonine Trypsine CID
Fragment [133-144] / Adduit 131 Da

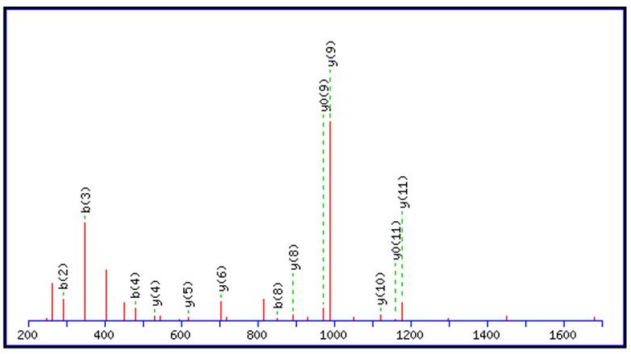
[133 – 144] GPCVLQEYQQFR (Ions score 80)



#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	58.0287	29.5180					G							12
2	155.0815	78.0444					P	1541.7206	771.3639	1524.6940	762.8506	1523.7100	762.3586	11
3	389.1278	195.0675					C	*1444.6678*	722.8375	1427.6413	714.3243	1426.6572	713.8323	10
4	488.1962	244.6017					V	*1210.6215*	605.8144	1193.5950	597.3011	1192.6109	596.8091	9
5	601.2803	301.1438					L	*1111.5531*	556.2802	1094.5265	547.7669	1093.5425	547.2749	8
6	729.3389	365.1731	712.3123	356.6598			Q	*998.4690*	499.7381	981.4425	491.2249	980.4585	490.7329	7
7	858.3815	429.6944	841.3549	421.1811	840.3709	420.6891	E	*870.4104*	435.7089	853.3839	427.1956	852.3999	426.7036	6
8	1021.4448	511.2260	1004.4182	502.7128	1003.4342	502.2207	Y	*741.3679*	371.1876	724.3413	362.6743			5
9	1149.5034	575.2553	1132.4768	566.7420	1131.4928	566.2500	Q	*578.3045*	289.6559	561.2780	281.1426			4
10	1277.5619	639.2846	1260.5354	630.7713	1259.5514	630.2793	Q	*450.2459*	225.6266	433.2194	217.1133			3
11	1424.6304	712.8188	1407.6038	704.3055	1406.6198	703.8135	F	*322.1874*	161.5973	305.1608	153.0840			2
12							R	175.1190	88.0631	158.0924	79.5498			1

hNEIL1/girgensonine Trypsine CID
Fragment [261-273] / Adduit 25 Da

[261– 273] CYGMPGMSSLQDR (Ions score 44)

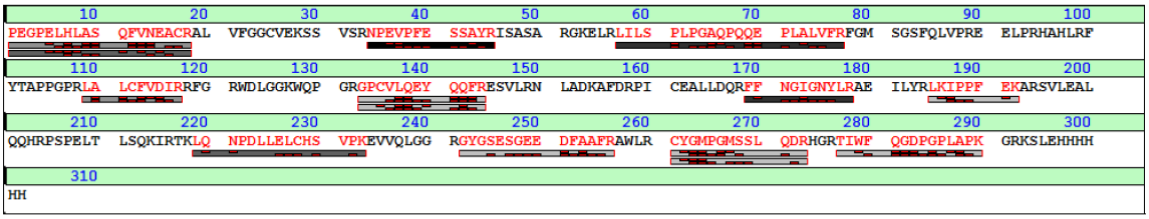


#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	129.0117	65.0095					C							13
2	**292.0750**	146.5412					Y	1341.5926	671.2999	1324.5660	662.7867	1323.5820	662.2946	12
3	**349.0965**	175.0519					G	*1178.5293*	589.7683	1161.5027	581.2550	1160.5187	580.7630	11
4	**480.1370**	240.5721					M	*1121.5078*	561.2575	1104.4812	552.7443	1103.4972	552.2523	10
5	577.1898	289.0985					P	*990.4673*	495.7373	973.4408	487.2240	972.4567	486.7320	9
6	634.2112	317.6092					G	*893.4145*	447.2109	876.3880	438.6976	875.4040	438.2056	8
7	765.2517	383.1295					M	836.3931	418.7002	819.3665	410.1869	818.3825	409.6949	7
8	**852.2837**	426.6455			834.2732	417.6402	S	*705.3526*	353.1799	688.3260	344.6667	687.3420	344.1747	6
9	939.3158	470.1615			921.3052	461.1562	S	*618.3206*	309.6639	601.2940	301.1506	600.3100	300.6586	5
10	1052.3998	526.7035			1034.3893	517.6983	L	*531.2885*	266.1479	514.2620	257.6346	513.2780	257.1426	4
11	1180.4584	590.7328	1163.4318	582.2196	1162.4478	581.7276	Q	418.2045	209.6059	401.1779	201.0926	400.1939	200.6006	3
12	1295.4853	648.2463	1278.4588	639.7330	1277.4748	639.2410	D	290.1459	145.5766	273.1193	137.0633	272.1353	136.5713	2
13							R	175.1190	88.0631	158.0924	79.5498			1

Results

Protein 1: >sp|hNEIL|hNEILdc95

Accession:	sp|hNEIL|hNEILdc95	**Score:**	777.6
Database:	RecProt	**Seq. Coverage [%]:**	48.7
MW [kDa] / pI:	34.2 / 9.2	**No. of Peptides:**	14
Modification(s):	D5-2 Frag2, Oxidation		

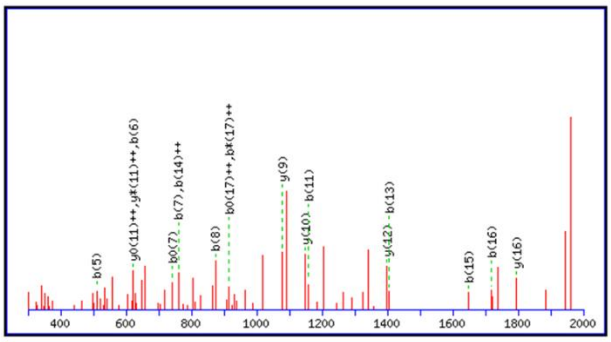


MS/MS Peptide Matches

Cmpd.	m/z meas.	z	Δ m/z [ppm]	Δ m/z [Da]	Rt [min]	Score	No. of Cmpds.	Site [%]	P	Sequence	Modification	Range	Type
23	999.4043	2	423.83	0.4234	10.76	90.67	1	0.0	0	-.PEGPELHLASQFVNEACR.A		1-18	CID
24	1011.3712	2	-106.11	-0.1073	10.81	24.86	1	0.0	0	-.PEGPELHLASQFVNEACR.A	D5-2 Frag2: 17	1-18	CID
10	698.3135	2	-24.49	-0.0171	10.37	56.09	3	0.0	0	R.NPEVPFESSAYR.I		34-45	CID
71	1144.0824	2	-69.81	-0.0799	12.44	107.71	2	0.0	0	R.LILSPLPGAQPQQEPLALVFR.F		57-77	CID
44	525.2864	2	-14.93	-0.0078	11.46	54.43	1	0.0	0	R.LALCFVDIR.R		109-117	CID
19	734.3356	2	-27.91	-0.0205	10.66	60.54	2	0.0	0	R.GPCVLQEYQQFR.E		133-144	CID
18	746.8172	2	-48.90	-0.0365	10.66	33.95	1	0.0	0	R.GPCVLQEYQQFR.E	D5-2 Frag2: 3	133-144	CID
33	600.7895	2	-36.86	-0.0221	11.06	46.5	2	0.0	0	R.FFNGIGNYLR.A		169-178	CID
2	486.3112	2	23.33	0.0113	10.04	35.42	2	0.0	1	R.LKIPPFEK.A		185-192	CID
25	569.7181	3	729.30	0.4152	10.82	38.78	1	0.0	0	K.LQNPDLLELCHSVPK.E		219-233	CID
15	811.277	2	-79.98	-0.0649	10.55	89.85	3	0.0	0	R.GYGSESGEEDFAAFR.A		242-256	CID
16	735.2565	2	-62.08	-0.0456	10.56	24.54	1	0.0	0	R.CYGMPGMSSLQDR.H	D5-2 Frag2: 1	261-273	CID
6	743.2785	2	-28.39	-0.0211	10.19	35.46	2	0.727	0	R.CYGMPGMSSLQDR.H	D5-2 Frag2: 1; Oxidation: 7	261-273	CID
34	763.8562	2	-62.23	-0.0475	11.13	78.83	12	0.0	0	R.TIWFQGDPGPLAPK.G		277-290	CID

hNEIL1/ICOA2-D5-P2 Trypsine CID
Fragment [1-18] / Adduit 25 Da

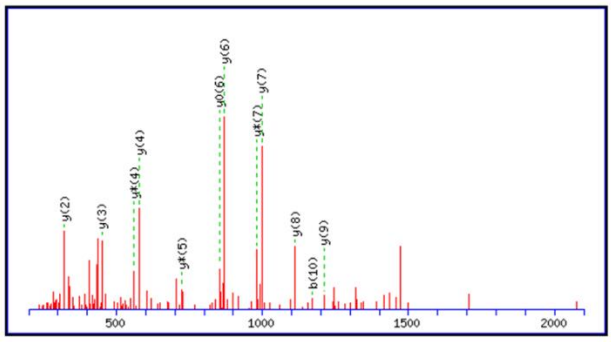
[1 – 18] PEGPELHLASQFVNEACR (Ions score 24)



#	b	b^{++}	b*	b*$^{++}$	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y*$^{++}$	y^{0}	y^{0++}	#
1	98.0600	49.5337					P							18
2	227.1026	114.0550			209.0921	105.0497	E	1924.8970	962.9521	1907.8705	954.4389	1906.8865	953.9469	17
3	284.1241	142.5657			266.1135	133.5604	G	1795.8544	898.4309	1778.8279	889.9176	1777.8439	889.4256	16
4	381.1769	191.0921			363.1663	182.0868	P	1738.8330	869.9201	1721.8064	861.4068	1720.8224	860.9148	15
5	*510.2195*	255.6134			492.2089	246.6081	E	1641.7802	821.3937	1624.7537	812.8805	1623.7696	812.3885	14
6	*623.3035*	312.1554			605.2930	303.1501	L	1512.7376	756.8724	1495.7111	748.3592	1494.7270	747.8672	13
7	*760.3624*	380.6849			742.3519	371.6796	H	1399.6535	700.3304	1382.6270	691.8171	1381.6430	691.3251	12
8	*873.4465*	437.2269			855.4359	428.2216	L	1262.5946	631.8010	1245.5681	623.2877	1244.5841	622.7957	11
9	944.4836	472.7454			926.4730	463.7402	A	1149.5106	575.2589	1132.4840	566.7456	1131.5000	566.2536	10
10	1031.5156	516.2615			1013.5051	507.2562	S	1078.4735	539.7404	1061.4469	531.2271	1060.4629	530.7351	9
11	*1159.5742*	580.2907	1142.5477	571.7775	1141.5636	571.2855	Q	991.4414	496.2244	974.4149	487.7111	973.4309	487.2191	8
12	1306.6426	653.8250	1289.6161	645.3117	1288.6321	644.8197	F	863.3828	432.1951	846.3563	423.6818	845.3723	423.1898	7
13	*1405.7110*	703.3592	1388.6845	694.8459	1387.7005	694.3539	V	716.3144	358.6609	699.2879	350.1476	698.3039	349.6556	6
14	1519.7540	760.3806	1502.7274	751.8673	1501.7434	751.3753	N	617.2460	309.1266	600.2195	300.6134	599.2355	300.1214	5
15	*1648.7966*	824.9019	1631.7700	816.3886	1630.7860	815.8966	E	503.2031	252.1052	486.1765	243.5919	485.1925	243.0999	4
16	*1719.8337*	860.4205	1702.8071	851.9072	1701.8231	851.4152	A	374.1605	187.5839	357.1340	179.0706			3
17	1847.8381	924.4227	1830.8116	915.9094	1829.8275	915.4174	C	303.1234	152.0653	286.0968	143.5521			2
18							R	175.1190	88.0631	158.0924	79.5498			1

hNEIL1/ICOA2-D5-P2 Trypsine CID
Fragment [1-18] / Adduit 25 Da

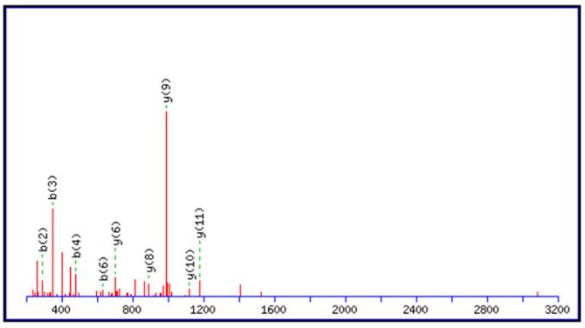
[133 – 144] GPCVLQEYQQFR (Ions score 34)



#	b	b^{++}	b*	b*$^{++}$	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y*$^{++}$	y^{0}	y^{0++}	#
1	58.0287	29.5180					G							12
2	155.0815	78.0444					P	1435.6787	718.3430	1418.6521	709.8297	1417.6681	709.3377	11
3	283.0859	142.0466					C	1338.6259	669.8166	1321.5994	661.3033	1320.6154	660.8113	10
4	382.1544	191.5808					V	*1210.6215*	605.8144	1193.5950	597.3011	1192.6109	596.8091	9
5	495.2384	248.1228					L	*1111.5531*	556.2802	1094.5265	547.7669	1093.5425	547.2749	8
6	623.2970	312.1521	606.2704	303.6389			Q	*998.4690*	499.7381	981.4425	491.2249	980.4585	490.7329	7
7	752.3396	376.6734	735.3130	368.1602	734.3290	367.6681	E	*870.4104*	435.7089	853.3839	427.1956	852.3999	426.7036	6
8	915.4029	458.2051	898.3764	449.6918	897.3924	449.1998	Y	741.3679	371.1876	724.3413	362.6743			5
9	1043.4615	522.2344	1026.4349	513.7211	1025.4509	513.2291	Q	*578.3045*	289.6559	561.2780	281.1426			4
10	1171.5201	586.2637	1154.4935	577.7504	1153.5095	577.2584	Q	*450.2459*	225.6266	433.2194	217.1133			3
11	1318.5885	659.7979	1301.5619	651.2846	1300.5779	650.7926	F	*322.1874*	161.5973	305.1608	153.0840			2
12							R	175.1190	88.0631	158.0924	79.5498			1

hNEIL1/ICOA2-D5-P2 Trypsine CID
Fragment [1-18] / Adduit 25 Da

[261– 273] CYGMPGMSSLQDR (Ions score 25)



#	b	b^{++}	b*	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y^{*++}	y^{0}	y^{0++}	#
1	129.0117	65.0095					C							13
2	*292.0750*	146.5412					Y	1341.5926	671.2999	1324.5660	662.7867	1323.5820	662.2946	12
3	*349.0965*	175.0519					G	*1178.5293*	589.7683	1161.5027	581.2550	1160.5187	580.7630	11
4	*480.1370*	240.5721					M	*1121.5078*	561.2575	1104.4812	552.7443	1103.4972	552.2523	10
5	577.1898	289.0985					P	*990.4673*	495.7373	973.4408	487.2240	972.4567	486.7320	9
6	*634.2112*	317.6092					G	*893.4145*	447.2109	876.3880	438.6976	875.4040	438.2056	8
7	765.2517	383.1295					M	836.3931	418.7002	819.3665	410.1869	818.3825	409.6949	7
8	852.2837	426.6455			834.2732	417.6402	S	*705.3526*	353.1799	688.3260	344.6667	687.3420	344.1747	6
9	939.3158	470.1615			921.3052	461.1562	S	618.3206	309.6639	601.2940	301.1506	600.3100	300.6586	5
10	1052.3998	526.7035			1034.3893	517.6983	L	531.2885	266.1479	514.2620	257.6346	513.2780	257.1426	4
11	1180.4584	590.7328	1163.4318	582.2196	1162.4478	581.7276	Q	418.2045	209.6059	401.1779	201.0926	400.1939	200.6006	3
12	1295.4853	648.2463	1278.4588	639.7330	1277.4748	639.2410	D	290.1459	145.5766	273.1193	137.0633	272.1353	136.5713	2
13							R	175.1190	88.0631	158.0924	79.5498			1

Annexe 8 : Composition des kits de cristallogenèse utilisés

JCSG+ // Molecular Dimensions

Well	Kit #	Salt	Buffer	pH	Precipitant
A1	1.1	0.2M Lithium sulfate	0.1M Sodium acetate	4.5	50% w/v PEG 400
B1	1.2	none	0.1M Sodium citrate	5.5	20% w/v PEG 3000
C1	1.3	0.2M Ammonium citrate dibasic	none		20% w/v PEG 3350
D1	1.4	0.02M Calcium chloride dihydrate	0.1M Sodium acetate	4.6	30% v/v MPD
E1	1.5	0.2M Magnesium formate dihydrate	none		20% w/v PEG 3350
F1	1.6	0.2M Lithium sulfate	0.1M Phosphate/citrate	4.2	20% w/v PEG 1000
G1	1.7	none	0.1M CHES	9.5	20% w/v PEG 8000
H1	1.8	0.2M Ammonium formate	none		20% w/v PEG 3350
A2	1.9	0.2M Ammonium chloride	none		20% w/v PEG 3350
B2	1.10	0.2M Potassium formate	none		20% w/v PEG 3350
C2	1.11	0.2M Ammonium phosphate monobasic	0.1M Tris	8.5	50% v/v MPD
D2	1.12	0.2M Potassium nitrate	none		20% w/v PEG 3350
E2	1.13	0.8M Ammonium sulfate	0.1M Citrate	4.0	none
F2	1.14	0.2M Sodium thiocyanate	none		20% w/v PEG 3350
G2	1.15	none	0.1M BICINE	9.0	20% w/v PEG 6000
H2	1.16	none	0.1M HEPES	7.5	10% w/v PEG 8000 / 8% v/v Ethylene glycol
A3	1.17	none	0.1M Sodium cacodylate	6.5	40% v/v MPD / 5% w/v PEG 8000
B3	1.18	none	0.1M Phosphate/citrate	4.2	40% v/v Ethanol / 5% w/v PEG 1000
C3	1.19	none	0.1M Sodium acetate	4.6	8% w/v PEG 4000
D3	1.20	0.2M Magnesium chloride hexahydrate	0.1M Tris	7.0	10% w/v PEG 8000
E3	1.21	none	0.1M Citrate	5.0	20% w/v PEG 6000
F3	1.22	0.2M Magnesium chloride hexahydrate	0.1M Sodium cacodylate	6.5	50% v/v PEG 200
G3	1.23	1.6M Sodium citrate tribasic dihydrate	none	6.5	none
H3	1.24	0.2M Potassium citrate tribasic monohydrate	none		20% w/v PEG 3350
A4	1.25	0.2M Sodium chloride	0.1M Phosphate/citrate	4.2	20% w/v PEG 8000
B4	1.26	1M Lithium chloride	0.1M Citrate	4.0	20% w/v PEG 6000
C4	1.27	0.2M Ammonium nitrate	none		20% w/v PEG 3350
D4	1.28	none	0.1M HEPES	7.0	10% w/v PEG 6000
E4	1.29	0.8M Sodium phosphate monobasic monohydrate 0.8M Potassium phosphate monobasic	0.1M Sodium HEPES	7.5	none
F4	1.30	none	0.1M Phosphate/citrate	4.2	40% v/v PEG 300
G4	1.31	0.2M Zinc acetate dihydrate	0.1M Sodium acetate	4.5	10% w/v PEG 3000
H4	1.32	none	0.1M Tris	8.5	20% v/v Ethanol
A5	1.33	none	0.1M Sodium/potassium phosphate	6.2	25% v/v 1,2-Propanediol / 10%v/v Glycerol
B5	1.34	none	0.1M BICINE	9.0	10% w/v PEG 20.000 / 2% v/v 1,4-Dioxane
C5	1.35	2M Ammonium sulfate	0.1M Sodium acetate	4.6	none
D5	1.36	none	none		10% w/v PEG 1000 / 10% w/v PEG 8000
E5	1.37	none	none		24% w/v PEG 1500 / 20% v/v Glycerol
F5	1.38	0.2M Magnesium chloride hexahydrate	0.1M Sodium HEPES	7.5	30% v/v PEG 400
G5	1.39	0.2M Sodium chloride	0.1M Sodium/potassium phosphate	6.2	50% v/v PEG 200
H5	1.40	0.2M Lithium sulfate	0.1M Sodium acetate	4.5	30% w/v PEG 8000
A6	1.41	none	0.1M HEPES	7.5	70% v/v MPD
B6	1.42	0.2M Magnesium chloride hexahydrate	0.1M Tris	8.5	20% w/v PEG 8000
C6	1.43	0.2M Lithium sulfate	0.1M Tris	8.5	40% v/v PEG 400
D6	1.44	none	0.1M Tris	8.0	40% v/v MPD
E6	1.45	0.17M Ammonium sulfate	none		25.5% w/v PEG 4000 / 15% v/v Glycerol
F6	1.46	0.2M Calcium acetate hydrate	0.1M Sodium cacodylate	6.5	40% v/v PEG 300
G6	1.47	0.14M Calcium chloride dihydrate	0.07M Sodium acetate	4.6	14% v/v 2-Propanol / 30% v/v Glycerol
H6	1.48	0.04M Potassium phosphate monobasic	none		16% w/v PEG 8000 / 20% v/v Glycerol
A7	2.1	1M Sodium citrate tribasic dihydrate	0.1M Sodium cacodylate	6.5	none
B7	2.2	2M Ammonium sulfate 0.2M Sodium chloride	0.1M Sodium cacodylate	6.5	none
C7	2.3	0.2M Sodium chloride	0.1M HEPES	7.5	10% v/v 2-Propanol
D7	2.4	1.26M Ammonium sulfate 0.2M Lithium sulfate	0.1M Tris	8.5	none
E7	2.5	none	0.1M CAPS	10.5	40% v/v MPD
F7	2.6	0.2M Zinc acetate dihydrate	0.1M Imidazole	8.0	20% w/v PEG 3000
G7	2.7	0.2M Zinc acetate dihydrate	0.1M Sodium cacodylate	6.5	10% v/v 2-Propanol
H7	2.8	1M Ammonium phosphate dibasic	0.1M Sodium acetate	4.5	none
A8	2.9	1.6M Magnesium sulfate heptahydrate	0.1M MES	6.5	none
B8	2.10	none	0.1M BICINE	9.0	10% w/v PEG 6000
C8	2.11	0.16M Calcium acetate hydrate	0.08M Sodium cacodylate	6.5	14.4% w/v PEG 8000 / 20% v/v Glycerol
D8	2.12	none	0.1M Imidazole	8.0	10% w/v PEG 8000
E8	2.13	0.05M Cesium chloride	0.1M MES	6.5	30% v/v Jeffamine® M-600
F8	2.14	3.2M Ammonium sulfate	0.1M Citrate	5.0	none
G8	2.15	none	0.1M Tris	8.0	20% v/v MPD
H8	2.16	none	0.1M HEPES	7.5	20% v/v Jeffamine® M-600
A9	2.17	0.2M Magnesium chloride hexahydrate	0.1M Tris	8.5	50% v/v Ethylene glycol
B9	2.18	none	0.1M BICINE	9.0	10% v/v MPD
C9	2.19	0.8M Succinic acid	none	7.0	none
D9	2.20	2.1M DL-Malic acid	none	7.0	none
E9	2.21	2.4M Sodium malonate dibasic monohydrate	none	7.0	none
F9	2.22	1.1M Sodium malonate dibasic monohydrate	0.1M HEPES	7.0	0.5% v/v Jeffamine® ED-2003
G9	2.23	1M Succinic acid	0.1M HEPES	7.0	1% w/v PEG 2000 MME
H9	2.24	none	0.1M HEPES	7.0	30% v/v Jeffamine® M-600
A10	2.25	none	0.1M HEPES	7.0	30% v/v Jeffamine® ED-2003
B10	2.26	0.02M Magnesium chloride hexahydrate	0.1M HEPES	7.5	22% w/v Poly(acrylic acid sodium salt) 5100
C10	2.27	0.01M Cobalt(II) chloride hexahydrate	0.1M Tris	8.5	20% w/v Polyvinylpyrolidone
D10	2.28	0.2M TMAO	0.1M Tris	8.5	20% w/v PEG 2000 MME
E10	2.29	0.005M Cobalt(II) chloride hexahydrate 0.005M Cadmium chloride hemi(pentahydrate) 0.005M Magnesium chloride hexahydrate 0.005M Nickel(II) chloride hexahydrate	0.1M HEPES	7.5	12% w/v PEG 3350
F10	2.30	0.2M Sodium malonate dibasic monohydrate	none		20% w/v PEG 3350
G10	2.31	0.1M Succinic acid	none		15% w/v PEG 3350
H10	2.32	0.15M DL-Malic acid	none		20% w/v PEG 3350
A11	2.33	0.1M Potassium thiocyanate	none		30% w/v PEG 2000 MME
B11	2.34	0.15M Potassium bromide	none		30% w/v PEG 2000 MME
C11	2.35	2M Ammonium sulfate	0.1M Bis-Tris	5.5	none
D11	2.36	3M Sodium chloride	0.1M Bis-Tris	5.5	none
E11	2.37	0.3M Magnesium formate dihydrate	0.1M Bis-Tris	5.5	none
F11	2.38	1M Ammonium sulfate	0.1M Bis-Tris	5.5	1% w/v PEG 3350
G11	2.39	none	0.1M Bis-Tris	5.5	25% w/v PEG 3350
H11	2.40	0.2M Calcium chloride dihydrate	0.1M Bis-Tris	5.5	45% v/v MPD
A12	2.41	0.2M Ammonium acetate	0.1M Bis-Tris	5.5	45% v/v MPD
B12	2.42	0.1M Ammonium acetate	0.1M Bis-Tris	5.5	17% w/v PEG 10.000
C12	2.43	0.2M Ammonium sulfate	0.1M Bis-Tris	5.5	25% w/v PEG 3350
D12	2.44	0.2M Sodium chloride	0.1M Bis-Tris	5.5	25% w/v PEG 3350
E12	2.45	0.2M Lithium sulfate	0.1M Bis-Tris	5.5	25% w/v PEG 3350
F12	2.46	0.2M Ammonium acetate	0.1M Bis-Tris	5.5	25% w/v PEG 3350
G12	2.47	0.2M Magnesium chloride hexahydrate	0.1M Bis-Tris	5.5	25% w/v PEG 3350
H12	2.48	0.2M Ammonium acetate	0.1M HEPES	7.5	45% v/v MPD

Wizard 1+2 // Molecular Dimensions

Well	Kit #	Salt	Buffer	pH	Precipitant
A1	1.1	none	0.1M CHES	9.5	20% (w/v) PEG-8000
B1	1.2	0.1M NaCl	0.1M HEPES	7.5	10% (v/v) 2-propanol
C1	1.3	none	0.1M CHES	9.5	15% (v/v) ethanol
D1	1.4	0.1M MgCl2	0.1M imidazole	8.0	35% (v/v) 2-methyl-2.4-pentanediol
E1	1.5	none	0.1M CAPS	10.5	30% (v/v) PEG-400
F1	1.6	none	0.1M citrate	5.5	20% (w/v) PEG-3000
G1	1.7	0.1M Zn(OAc)2	0.1M MES	6.0	10% (w/v) PEG-8000
H1	1.8	none	0.1M citrate	5.5	2.0 M $(NH_4)_2SO_4$
A2	1.9	none	0.1M acetate	4.5	1.0 M $(NH_4)_2HPO_4$
B2	1.10	none	0.1M Tris	7.0	20% (w/v) PEG-2000 MME
C2	1.11	0.1M Li2SO4	0.1M MES	6.0	20% (v/v) 1.4-butanediol
D2	1.12	0.1M Ca(OAc)2	0.1M imidazole	8.0	20% (w/v) PEG-1000
E2	1.13	none	0.1M cacodylate	6.5	1.26 M $(NH_4)_2SO_4$
F2	1.14	none	0.1M cacodylate	6.5	1.0 M sodium citrate
G2	1.15	0.1M Li2SO4	0.1M imidazole	8.0	10% (w/v) PEG-3000
H2	1.16	none	0.1M Na/K phosphate	6.2	2.5 M NaCl
A3	1.17	0.1M Li2SO4	0.1M acetate	4.5	30% (w/v) PEG-8000
B3	1.18	0.1M NaCl	0.1M imidazole	8.0	1.0 M K/Na tartrate
C3	1.19	none	0.1M Tris	7.0	20% (w/v) PEG-1000
D3	1.20	0.1M NaCl	0.1M imidazole	8.0	0.4 M NaH_2PO_4/1.6 M K_2HPO_4
E3	1.21	none	0.1M HEPES	7.5	20% (w/v) PEG-8000
F3	1.22	none	0.1M Tris	8.5	10% (v/v) 2-propanol
G3	1.23	0.1M MgCl2	0.1M imidazole	8.0	15% (v/v) ethanol
H3	1.24	0.1M NaCl	0.1M Tris	7.0	35% (v/v) 2-methyl-2.4-pentanediol
A4	1.25	0.1M MgCl2	0.1M Tris	8.5	30% (v/v) PEG-400
B4	1.26	none	0.1M CHES	9.5	10% (w/v) PEG-3000
C4	1.27	0.1M Li2SO4	0.1M CAPS	10.5	1.2 M NaH_2PO_4/0.8 M K_2HPO_4
D4	1.28	0.1M NaCl	0.1M HEPES	7.5	20% (w/v) PEG-3000
E4	1.29	0.1M NaCl	0.1M CHES	9.5	10% (w/v) PEG-8000
F4	1.30	0.1M NaCl	0.1M acetate	4.5	1.26 M $(NH_4)_2SO_4$
G4	1.31	0.1M NaCl	0.1M phosphate-citrate	4.2	20% (w/v) PEG-8000
H4	1.32	none	0.1M Na/K phosphate	6.2	10% (w/v) PEG-3000
A5	1.33	0.1M Li2SO4	0.1M CAPS	10.5	2.0 M $(NH_4)_2SO_4$
B5	1.34	none	0.1M imidazole	8.0	1.0 M $(NH_4)_2HPO_4$
C5	1.35	none	0.1M acetate	4.5	20% (v/v) 1.4-butanediol
D5	1.36	none	0.1M imidazole	8.0	1.0 M sodium citrate
E5	1.37	none	0.1M imidazole	8.0	2.5 M NaCl
F5	1.38	0.1M Li2SO4	0.1M CHES	9.5	1.0 M K/Na tartrate
G5	1.39	0.1M Li2SO4	0.1M phosphate-citrate	4.2	20% (w/v) PEG-1000
H5	1.40	0.1M Ca(OAc)2	0.1M MES	6.0	10% (v/v) 2-propanol
A6	1.41	none	0.1M CHES	9.5	30% (w/v) PEG-3000
B6	1.42	none	0.1M Tris	7.0	15% (v/v) ethanol
C6	1.43	none	0.1M Na/K phosphate	6.2	35% (v/v) 2-methyl-2.4-pentanediol
D6	1.44	0.1M Ca(OAc)2	0.1M acetate	4.5	30% (v/v) PEG-400
E6	1.45	none	0.1M acetate	4.5	20% (w/v) PEG-3000
F6	1.46	0.1M Ca(OAc)2	0.1M imidazole	8.0	10% (w/v) PEG-8000
G6	1.47	0.1M Li2SO4	0.1M Tris	8.5	1.26 M $(NH_4)_2SO_4$
H6	1.48	0.1M Zn(OAc)2	0.1M acetate	4.5	20% (w/v) PEG-1000
A7	2.1	0.2M Zn(OAc)2	0.1M acetate	4.5	10% (w/v) PEG-3000
B7	2.2	0.2M Li2SO4	0.1M MES	6.0	35% (v/v) 2-methyl-2,4-pentanediol
C7	2.3	0.2M MgCl2	0.1M Tris	8.5	20% (w/v) PEG-8000
D7	2.4	0.2M NaCl	0.1M cacodylate	6.5	2.0 M (NH4)2SO4
E7	2.5	0.2M NaCl	0.1M HEPES	7.5	20% (v/v) 1,4-butanediol
F7	2.6	0.2M Li2SO4	0.1M phosphate-citrate	4.2	10% (v/v) 2-propanol
G7	2.7	0.2M NaCl	0.1M Tris	7.0	30% (w/v) PEG-3000
H7	2.8	0.2M NaCl	0.1M Na/K phosphate	6.2	10% (w/v) PEG-8000
A8	2.9	none	0.1M phosphate-citrate	4.2	2.0 M (NH4)2SO4
B8	2.10	none	0.1M Tris	8.5	1.0 M (NH4)2HPO4
C8	2.11	0.2M Zn(OAc)2	0.1M cacodylate	6.5	10% (v/v) 2-propanol
D8	2.12	0.2M Li2SO4	0.1M cacodylate	6.5	30% (v/v) PEG-400
E8	2.13	0.2M Li2SO4	0.1M citrate	5.5	15% (v/v) ethanol
F8	2.14	0.2M NaCl	0.1M Na/K phosphate	6.2	20% (w/v) PEG-1000
G8	2.15	none	0.1M HEPES	7.5	1.26 M (NH4)2SO4
H8	2.16	none	0.1M CHES	9.5	1.0 M sodium citrate
A9	2.17	0.2M MgCl2	0.1M Tris	7.0	2.5 M NaCl
B9	2.18	0.2M Ca(OAc)2	0.1M Tris	7.0	20% (w/v) PEG-3000
C9	2.19	none	0.1M phosphate-citrate	4.2	1.6 M NaH2PO4/0.4 M K2HPO4
D9	2.20	0.2M Zn(OAc)2	0.1M MES	6.0	15% (v/v) ethanol
E9	2.21	none	0.1M acetate	4.5	35% (v/v) 2-methyl-2,4-pentanediol
F9	2.22	none	0.1M imidazole	8.0	10% (v/v) 2-propanol
G9	2.23	0.2M MgCl2	0.1M HEPES	7.5	15% (v/v) ethanol
H9	2.24	0.2M NaCl	0.1M imidazole	8.0	30% (w/v) PEG-8000
A10	2.25	0.2M NaCl	0.1M HEPES	7.5	35% (v/v) 2-methyl-2,4-pentanediol
B10	2.26	none	0.1M CHES	9.5	30% (v/v) PEG-400
C10	2.27	0.2M MgCl2	0.1M cacodylate	6.5	10% (w/v) PEG-3000
D10	2.28	0.2M Ca(OAc)2	0.1M MES	6.0	20% (w/v) PEG-8000
E10	2.29	0.2M NaCl	0.1M CHES	9.5	1.26 M (NH4)2SO4
F10	2.30	0.2M Zn(OAc)2	0.1M imidazole	8.0	20% (v/v) 1,4-butanediol
G10	2.31	0.2M NaCl	0.1M Tris	7.0	1.0 M sodium citrate
H10	2.32	none	0.1M Tris	8.5	20% (w/v) PEG-1000
A11	2.33	0.2M NaCl	0.1M citrate	5.5	1.0 M (NH4)2HPO4
B11	2.34	none	0.1M imidazole	8.0	10% (w/v) PEG-8000
C11	2.35	none	0.1M acetate	4.5	0.8 M NaH2PO4/1.2 M K2HPO4
D11	2.36	0.2M NaCl	0.1M phosphate-citrate	4.2	10% (w/v) PEG-3000

E11	2.37	0.2M Li2SO4	0.1M Tris	7.0	1.0 M K/Na tartrate
F11	2.38	0.2M Li2SO4	0.1M acetate	4.5	2.5 M NaCl
G11	2.39	0.2M NaCl	0.1M CAPS	10.5	20% (w/v) PEG-8000
H11	2.40	0.2M Zn(OAc)2	0.1M imidazole	8.0	20% (w/v) PEG-3000
A12	2.41	0.2M Li2SO4	0.1M Tris	7.0	2.0 M (NH4)2SO4
B12	2.42	0.2M NaCl	0.1M HEPES	7.5	30% (v/v) PEG-400
C12	2.43	0.2M MgCl2	0.1M Tris	7.0	10% (w/v) PEG-8000
D12	2.44	0.2M MgCl2	0.1M cacodylate	6.5	20% (w/v) PEG-1000
E12	2.45	none	0.1M MES	6.0	1.26 M (NH4)2SO4
F12	2.46	0.2M NaCl	0.1M imidazole	8.0	1.0 M (NH4)2HPO4
G12	2.47	0.2M Zn(OAc)2	0.1M imidazole	8.0	2.5 M NaCl
H12	2.48	none	0.1M MES	6.0	1.0 M K/Na tartrate

Q_Classic // Molecular Dimensions

Well	Kit #	Salt	Buffer	pH	Precipitant
A1	1.1	0.01 M Cobalt chloride	0.1 M Sodium acetate	4.6	1.0 M 1,6-Hexanediol
B1	1.2	none	0.1 M tri-Sodium citrate	5.6	2.5 M 1,6-Hexanediol
C1	1.3	0.2 M Magnesium chloride	0.1 M TRIS	8.5	3.4 M 1,6-Hexanediol
D1	1.4	2.0 M Ammonium sulfate	none		5 %(v/v) Isopropanol
E1	1.5	none	0.1 M HEPES sodium salt	7.5	10 %(v/v) Isopropanol / 20 %(w/v) PEG 4000
F1	1.6	0.2 M Calcium chloride	0.1 M Sodium acetate	4.6	20 %(v/v) Isopropanol
G1	1.7	none	0.1 M tri-Sodium citrate	5.6	20 %(v/v) Isopropanol / 20 %(w/v) PEG 4000
H1	1.8	0.2 M tri-Sodium citrate	0.1 M HEPES sodium salt	7.5	20 %(v/v) Isopropanol
A2	1.9	0.2 M tri-Sodium citrate	0.1 M Sodium cacodylate	6.5	30 %(v/v) Isopropanol
B2	1.10	0.2 M Magnesium chloride	0.1 M HEPES sodium salt	7.5	30 %(v/v) Isopropanol
C2	1.11	0.2 M Ammonium acetate	0.1 M TRIS.HCl	8.5	30 %(v/v) Isopropanol
D2	1.12	1.5 M Sodium chloride	none		10 %(v/v) Ethanol
E2	1.13	none	0.1 M TRIS	8.5	20 %(v/v) Ethanol
F2	1.14	none	none		25 %(v/v) Ethylene glycol
G2	1.15	0.02 M Calcium chloride	0.1 M Sodium acetate	4.6	30 %(v/v) MPD
H2	1.16	0.2 M Sodium chloride	0.1 M Sodium acetate	4.6	30 %(v/v) MPD
A3	1.17	0.2 M Ammonium acetate	0.1 M tri-Sodium citrate	5.6	30 %(v/v) MPD
B3	1.18	0.2 M Magnesium acetate	0.1 M Sodium cacodylate	6.5	30 %(v/v) MPD
C3	1.19	0.2 M tri-Sodium citrate	0.1 M HEPES sodium salt	7.5	30 %(v/v) MPD
D3	1.20	0.5 M Ammonium sulfate	0.1 M HEPES	7.5	30 %(v/v) MPD
E3	1.21	0.2 M Ammonium phosphate	0.1 M TRIS	8.5	50 %(v/v) MPD
F3	1.22	none	0.1 M HEPES	7.5	70 %(v/v) MPD
G3	1.23	none	0.1 M TRIS	8.5	25 %(v/v) tert-Butanol
H3	1.24	none	0.1 M tri-Sodium citrate	5.6	35 %(v/v) tert-Butanol
A4	1.25	none	none		0.4 M Ammonium phosphate
B4	1.26	none	0.1 M tri-Sodium citrate	5.6	1.0 M Ammonium phosphate
C4	1.27	none	0.1 M TRIS.HCl	8.5	2.0 M Ammonium phosphate
D4	1.28	none	0.1 M HEPES	7.5	2.0 M Ammonium formate
E4	1.29	none	0.1 M Sodium acetate	4.6	2.0 M Ammonium sulfate
F4	1.30	none	0.1 M TRIS.HCl	8.5	2.0 M Ammonium sulfate
G4	1.31	none	none		2.0 M Ammonium sulfate
H4	1.32	0.1 M Sodium chloride	0.1 M HEPES	7.5	1.6 M Ammonium sulfate
A5	1.33	0.01 M Cobalt chloride	0.1 M MES	6.5	1.8 M Ammonium sulfate
B5	1.34	0.2 M K/Na tartrate	0.1 M tri-Sodium citrate	5.6	2.0 M Ammonium sulfate
C5	1.35	none	none		1.0 M Imidazole pH 7.0
D5	1.36	none	none		0.4 M K/Na tartrate
E5	1.37	none	0.1 M HEPES sodium salt	7.5	0.8 M K/Na tartrate
F5	1.38	none	0.1 M Imidazole	6.5	1.0 M Sodium acetate
G5	1.39	0.05 M Cadmium sulfate	0.1 M HEPES	7.5	1.0 M Sodium acetate
H5	1.40	none	0.1 M Sodium cacodylate	6.5	1.4 M Sodium acetate
A6	1.41	none	0.1 M Sodium acetate	4.6	2.0 M Sodium chloride
B6	1.42	0.1 M Sodium phosphate 0.1 M Potassium phosphate	0.1 M MES	6.5	2.0 M Sodium chloride
C6	1.43	none	0.1 M HEPES	7.5	4.3 M Sodium chloride
D6	1.44	none	0.1 M HEPES sodium salt	7.5	1.4 M tri-Sodium citrate
E6	1.45	none	none		1.6 M tri-Sodium citrate pH 6.5
F6	1.46	none	0.1 M HEPES sodium salt	7.5	0.8 M Sodium phosphate / 0.8 M Potassium phosphate
G6	1.47	none	0.1 M Sodium acetate	4.6	2.0 M Sodium formate
H6	1.48	none	none		4.0 M Sodium formate
A7	2.1	none	0.1 M BICINE	9.0	2 %(v/v) Dioxane / 10 %(w/v) PEG 20000
B7	2.2	1.6 M Ammonium sulfate	0.1 M MES	6.5	10 %(v/v) Dioxane
C7	2.3	none	none		35 %(v/v) Dioxane
D7	2.4	0.5 M Sodium chloride	0.1 M tri-Sodium citrate	5.6	2 %(v/v) Ethylene imine polymer
E7	2.5	1.5 M Ammonium sulfate	0.1 M TRIS	8.5	12 %(v/v) Glycerol
F7	2.6	0.5 M Sodium chloride 0.01 M Magnesium chloride	none		0.01 M CTAB
G7	2.7	0.01 M Ferric chloride	0.1 M tri-Sodium citrate	5.6	10 %(v/v) Jeffamine M-600
H7	2.8	none	0.1 M HEPES	7.5	20 %(v/v) Jeffamine M-600
A8	2.9	0.5 M Ammonium sulfate	0.1 M tri-Sodium citrate	5.6	1.0 M Lithium sulfate
B8	2.10	0.01 M Nickel chloride	0.1 M TRIS	8.5	1.0 M Lithium sulfate
C8	2.11	none	0.1 M HEPES sodium salt	7.5	1.5 M Lithium sulfate
D8	2.12	none	0.1 M BICINE	9.0	2.0 M Magnesium chloride
E8	2.13	none	none		0.2 M Magnesium formate
F8	2.14	none	0.1 M MES	6.5	1.6 M Magnesium sulfate
G8	2.15	none	0.1 M TRIS.HCl	8.5	8 %(w/v) PEG 8000
H8	2.16	none	0.1 M HEPES	7.5	10 %(w/v) PEG 8000
A9	2.17	0.5 M Lithium sulfate	none		15 %(w/v) PEG 8000
B9	2.18	0.2 M Zinc acetate	0.1 M Sodium cacodylate	6.5	18 %(w/v) PEG 8000
C9	2.19	0.2 M Calcium acetate	0.1 M Sodium cacodylate	6.5	18 %(w/v) PEG 8000
D9	2.20	0.2 M Magnesium acetate	0.1 M Sodium cacodylate	6.5	20 %(w/v) PEG 8000
E9	2.21	0.05 M Potassium phosphate	none		20 %(w/v) PEG 8000
F9	2.22	0.2 M Ammonium sulfate	0.1 M Sodium cacodylate	6.5	30 %(w/v) PEG 8000
G9	2.23	0.2 M Sodium acetate	0.1 M Sodium cacodylate	6.5	30 %(w/v) PEG 8000
H9	2.24	0.2 M Ammonium sulfate	none		30 %(w/v) PEG 8000
A10	2.25	2.0 M Ammonium sulfate	0.1 M HEPES sodium salt	7.5	2 %(v/v) PEG 400

B10	2.26	0.2 M Calcium chloride	0.1 M HEPES sodium salt	7.5	28 %(v/v) PEG 400
C10	2.27	0.1 M Cadmium chloride	0.1 M Sodium acetate	4.6	30 %(v/v) PEG 400
D10	2.28	0.2 M Magnesium chloride	0.1 M HEPES sodium salt	7.5	30 %(v/v) PEG 400
E10	2.29	0.2 M tri-Sodium citrate	0.1 M TRIS.HCl	8.5	30 %(v/v) PEG 400
F10	2.30	0.1 M Sodium chloride	0.1 M BICINE	9.0	20 %(w/v) PEG 550 MME
G10	2.31	0.01 M Zinc sulfate	0.1 M MES	6.5	25 %(w/v) PEG 550 MME
H10	2.32	none	none		10 %(w/v) PEG 1000 / 10 %(w/v) PEG 8000
A11	2.33	none	none		30 %(w/v) PEG 1500
B11	2.34	0.01 M Nickel chloride	0.1 M TRIS	8.5	20 %(w/v) PEG 2000 MME
C11	2.35	0.2 M Ammonium sulfate	0.1 M Sodium acetate	4.6	30 %(w/v) PEG 2000 MME
D11	2.36	none	0.1 M Sodium acetate	4.6	8 %(w/v) PEG 4000
E11	2.37	0.2 M Ammonium sulfate	0.1 M Sodium acetate	4.6	25 %(w/v) PEG 4000
F11	2.38	0.2 M Ammonium acetate	0.1 M Sodium acetate	4.6	30 %(w/v) PEG 4000
G11	2.39	0.2 M Ammonium acetate	0.1 M tri-Sodium citrate	5.6	30 %(w/v) PEG 4000
H11	2.40	0.2 M Magnesium chloride	0.1 M TRIS.HCl	8.5	30 %(w/v) PEG 4000
A12	2.41	0.2 M Lithium sulfate	0.1 M TRIS.HCl	8.5	30 %(w/v) PEG 4000
B12	2.42	0.2 M Sodium acetate	0.1 M TRIS.HCl	8.5	30 %(w/v) PEG 4000
C12	2.43	0.2 M Ammonium sulfate	none		30 %(w/v) PEG 4000
D12	2.44	0.2 M Ammonium sulfate	0.1 M MES	6.5	30 %(w/v) PEG 5000 MME
E12	2.45	none	0.1 M HEPES	7.5	10 %(w/v) PEG 6000 / 5 %(v/v) MPD
F12	2.46	2.0 M Sodium chloride	none		10 %(w/v) PEG 6000
G12	2.47	none	0.1 M HEPES	7.5	20 %(w/v) PEG 10000 / 8 %(v/v) Ethylene glycol
H12	2.48	none	0.1 M MES	6.5	12 %(w/v) PEG 20000

HR 1+2 // Hampton Research

Well	Kit #	Salt	Buffer	pH	Precipitant
A1	1.1	0.02M Calcium chloride dihydrate	0.1M Sodium acetate trihydrate	4.6	30% v/v (+/-)-2-Methyl-2,4-pentanediol
B1	1.2	none	none		0.4M Potassium sodium tartrate tetrahydrate
C1	1.3	none	none		0.4M Ammonium phosphate monobasic
D1	1.4	none	0.1M TRIS hydrochloride	8.5	2.0M Ammonium sulfate
E1	1.5	0.2M Sodium citrate tribasic dihydrate	0.1M HEPES sodium	7.5	30% v/v (+/-)-2-Methyl-2,4-pentanediol
F1	1.6	0.2M Magnesium chloride hexahydrate	0.1M TRIS hydrochloride	8.5	30% w/v Polyethylene glycol 4,000
G1	1.7	none	0.1M Sodium cacodylate trihydrate	6.5	1.4M Sodium acetate trihydrate
H1	1.8	0.2M Sodium citrate tribasic dihydrate	0.1M Sodium cacodylate trihydrate	6.5	30% v/v 2-Propanol
A2	1.9	0.2M Ammonium acetate	0.1M Sodium citrate tribasic dihydrate	5.6	30% w/v Polyethylene glycol 4,000
B2	1.10	0.2M Ammonium acetate	0.1M Sodium acetate trihydrate	4.6	30% w/v Polyethylene glycol 4,000
C2	1.11	none	0.1M Sodium citrate tribasic dihydrate	5.6	1.0M Ammonium phosphate monobasic
D2	1.12	0.2M Magnesium chloride hexahydrate	0.1M HEPES sodium	7.5	30% v/v 2-Propanol
E2	1.13	0.2M Sodium citrate tribasic dihydrate	0.1M TRIS hydrochloride	8.5	30% v/v Polyethylene glycol 400
F2	1.14	0.2M Calcium chloride dihydrate	0.1M HEPES sodium	7.5	30% v/v Polyethylene glycol 400
G2	1.15	0.2M Ammonium sulfate	0.1M Sodium cacodylate trihydrate	6.5	30% w/v Polyethylene glycol 8,000
H2	1.16	none	0.1M HEPES sodium	7.5	1.5M Lithium sulfate monohydrate
A3	1.17	0.2M Lithium sulfate monohydrate	0.1M TRIS hydrochloride	8.5	30% w/v Polyethylene glycol 4,000
B3	1.18	0.2M Magnesium acetate tetrahydrate	0.1M Sodium cacodylate trihydrate	6.5	20% w/v Polyethylene glycol 8,000
C3	1.19	0.2M Ammonium acetate	0.1M TRIS hydrochloride	8.5	30% v/v 2-Propanol
D3	1.20	0.2M Ammonium sulfate	0.1M Sodium acetate trihydrate	4.6	25% w/v Polyethylene glycol 4,000
E3	1.21	0.2M Magnesium acetate tetrahydrate	0.1M Sodium cacodylate trihydrate	6.5	30% v/v (+/-)-2-Methyl-2,4-pentanediol
F3	1.22	0.2M Sodium acetate trihydrate	0.1M TRIS hydrochloride	8.5	30% w/v Polyethylene glycol 4,000
G3	1.23	0.2M Magnesium chloride hexahydrate	0.1M HEPES sodium	7.5	30% v/v Polyethylene glycol 400
H3	1.24	0.2M Calcium chloride dihydrate	0.1M Sodium acetate trihydrate	4.6	20% v/v 2-Propanol
A4	1.25	none	0.1M imidazole	6.5	1.0M Sodium acetate trihydrate
B4	1.26	0.2M Ammonium acetate	0.1M Sodium citrate tribasic dihydrate	5.6	30% v/v (+/-)-2-Methyl-2,4-pentanediol
C4	1.27	0.2M Sodium citrate tribasic dihydrate	0.1M HEPES sodium	7.5	20% v/v 2-Propanol
D4	1.28	0.2M Sodium acetate trihydrate	0.1M Sodium cacodylate trihydrate	6.5	30% w/v Polyethylene glycol 8,000
E4	1.29	none	0.1M HEPES sodium	7.5	0.8M Potassium sodium tartrate tetrahydrate
F4	1.30	0.2M Ammonium sulfate	none		30% w/v Polyethylene glycol 8,000
G4	1.31	0.2M Ammonium sulfate	none		30% w/v Polyethylene glycol 4,000
H4	1.32	none	none		2.0M Ammonium sulfate
A5	1.33	none	none		4.0M Sodium formate
B5	1.34	none	0.1M Sodium acetate trihydrate	4.6	2.0M Sodium formate
C5	1.35	none	0.1M HEPES sodium	7.5	0.8M Sodium phosphate monobasic monohydrate 0.8M Potassium phosphate monobasic
D5	1.36	none	0.1M TRIS hydrochloride	8.5	8% w/v Polyethylene glycol 8,000
E5	1.37	none	0.1M Sodium acetate trihydrate	4.6	8% w/v Polyethylene glycol 4,000
F5	1.38	none	0.1M HEPES sodium	7.5	1.4M Sodium citrate tribasic dihydrate
G5	1.39	none	0.1M HEPES sodium	7.5	2% v/v Polyethylene glycol 400 2.0M Ammonium sulfate
H5	1.40	none	0.1M Sodium citrate tribasic dihydrate	5.6	20% v/v 2-Propanol 20% w/v Polyethylene glycol 4,000
A6	1.41	none	0.1M HEPES sodium	7.5	10% v/v 2-Propanol 20% w/v Polyethylene glycol 4,000
B6	1.42	0.05M Potassium phosphate monobasic	none		20% w/v Polyethylene glycol 8,000
C6	1.43	none	none		30% w/v Polyethylene glycol 1,500
D6	1.44	none	none		0.2M Magnesium formate dihydrate
E6	1.45	0.2M Zinc acetate dihydrate	0.1M Sodium cacodylate trihydrate	6.5	18% w/v Polyethylene glycol 8,000
F6	1.46	0.2M Calcium acetate hydrate	0.1M Sodium cacodylate trihydrate	6.5	18% w/v Polyethylene glycol 8,000
G6	1.47	none	0.1M Sodium acetate trihydrate	4.6	2.0M Ammonium sulfate
H6	1.48	none	0.1M TRIS hydrochloride	8.5	2.0M Ammonium phosphate monobasic
A7	2.1	2.0M Sodium chloride	none		10% w/v Polyethylene glycol 6,000
B7	2.2	0.5M Sodium chloride 0.01M Magnesium chloride hexahydrate 0.01M Hexadecyltrimethylammonium bromide	none		none
C7	2.3	none	none		25% v/v Ethylene glycol
D7	2.4	none	none		35% v/v 1,4-Dioxane
E7	2.5	2.0M Ammonium sulfate	none		5% v/v 2-Propanol
F7	2.6	none	none		1.0M Imidazole pH 7.0
G7	2.7	none	none		10% w/v Polyethylene glycol 1,000 10% w/v Polyethylene glycol 8,000
H7	2.8	1.5M Sodium chloride	none		10% v/v Ethanol
A8	2.9	none	0.1M Sodium acetate trihydrate	4.6	2.0M Sodium chloride

B8	2.10	0.2M Sodium chloride	0.1M Sodium acetate trihydrate	4.6	30% v/v (+/-)-2-Methyl-2,4-pentanediol
C8	2.11	0.01M Cobalt(II) chloride hexahydrate	0.1M Sodium acetate trihydrate	4.6	1.0M 1,6-Hexanediol
D8	2.12	0.1M Cadmium chloride hydrate	0.1M Sodium acetate trihydrate	4.6	30% v/v Polyethylene glycol 400
E8	2.13	0.2M Ammonium sulfate	0.1M Sodium acetate trihydrate	4.6	30% w/v Polyethylene glycol monomethyl ether 2,000
F8	2.14	0.2M Potassium sodium tartrate tetrahydrate	0.1M Sodium citrate tribasic dihydrate	5.6	2.0M Ammonium sulfate
G8	2.15	0.5M Ammonium sulfate	0.1M Sodium citrate tribasic dihydrate	5.6	1.0M Lithium sulfate monohydrate
H8	2.16	0.5M Sodium chloride	0.1M Sodium citrate tribasic dihydrate	5.6	2% v/v Ethylene imine polymer
A9	2.17	none	0.1M Sodium citrate tribasic dihydrate	5.6	35% v/v tert-Butanol
B9	2.18	0.01M Iron(III) chloride hexahydrate	0.1M Sodium citrate tribasic dihydrate	5.6	10% v/v Jeffamine M-600
C9	2.19	none	0.1M Sodium citrate tribasic dihydrate	5.6	2.5M 1,6-Hexanediol
D9	2.20	none	0.1M MES monohydrate	6.5	1.6M Magnesium sulfate heptahydrate
E9	2.21	0.1M Sodium phosphate monobasic monohydrate 0.1M Potassium phosphate monobasic	0.1M MES monohydrate	6.5	2.0M Sodium chloride
F9	2.22	none	0.1M MES monohydrate	6.5	12% w/v Polyethylene glycol 20,000
G9	2.23	1.6M Ammonium sulfate	0.1M MES monohydrate	6.5	10% v/v 1,4-Dioxane
H9	2.24	0.05M Cesium chloride	0.1M MES monohydrate	6.5	30% v/v Jeffamine M-600
A10	2.25	0.01M Cobalt(II) chloride hexahydrate	0.1M MES monohydrate	6.5	1.8M Ammonium sulfate
B10	2.26	0.2M Ammonium sulfate	0.1M MES monohydrate	6.5	30% w/v Polyethylene glycol monomethyl ether 5,000
C10	2.27	0.01M Zinc sulfate heptahydrate	0.1M MES monohydrate	6.5	25% v/v Polyethylene glycol monomethyl ether 550
D10	2.28	none	1.6M Sodium citrate tribasic dihydrate	6.5	none
E10	2.29	0.5M Ammonium sulfate	0.1M HEPES sodium	7.5	30% v/v (+/-)-2-Methyl-2,4-pentanediol
F10	2.30	none	0.1M HEPES sodium	7.5	10% w/v Polyethylene glycol 6,000 5% v/v (+/-)-2-Methyl-2,4-pentanediol
G10	2.31	none	0.1M HEPES sodium	7.5	20% v/v Jeffamine M-600
H10	2.32	0.1M Sodium chloride	0.1M HEPES sodium	7.5	1.6M Ammonium sulfate
A11	2.33	none	0.1M HEPES sodium	7.5	2.0M Ammonium formate
B11	2.34	0.05M Cadmium sulfate hydrate	0.1M HEPES sodium	7.5	1.0M Sodium acetate trihydrate
C11	2.35	none	0.1M HEPES sodium	7.5	70% v/v (+/-)-2-Methyl-2,4-pentanediol
D11	2.36	none	0.1M HEPES sodium	7.5	4.3M Sodium chloride
E11	2.37	none	0.1M HEPES sodium	7.5	10% w/v Polyethylene glycol 8,000 8% v/v Ethylene glycol
F11	2.38	none	0.1M HEPES sodium	7.5	20% w/v Polyethylene glycol 10,000
G11	2.39	0.2M Magnesium chloride hexahydrate	0.1M TRIS hydrochloride	8.5	3.4M 1,6-Hexanediol
H11	2.40	none	0.1M TRIS hydrochloride	8.5	25% v/v tert-Butanol
A12	2.41	0.01M Nickel(II) chloride hexahydrate	0.1M TRIS hydrochloride	8.5	1.0M Lithium sulfate monohydrate
B12	2.42	1.5M Ammonium sulfate	0.1M TRIS hydrochloride	8.5	12% v/v Glycerol
C12	2.43	0.2M Ammonium phosphate monobasic	0.1M TRIS hydrochloride	8.5	50% v/v (+/-)-2-Methyl-2,4-pentanediol
D12	2.44	none	0.1M TRIS hydrochloride	8.5	20% v/v Ethanol
E12	2.45	0.01M Nickel(II) chloride hexahydrate	0.1M TRIS hydrochloride	8.5	20% w/v Polyethylene glycol monomethyl ether 2,000
F12	2.46	0.1M Sodium chloride	0.1M BICINE	9.0	20% v/v Polyethylene glycol monomethyl ether 550
G12	2.47	none	0.1M BICINE	9.0	2.0M Magnesium chloride hexahydrate
H12	2.48	none	0.1M BICINE	9.0	2% v/v 1,4-Dioxane 10% w/v Polyethylene glycol 20,000

Morpheus // Molecular Dimension

Well	Kit #	Salt	Buffer	pH	Precipitant
A1	1.1	0.06M Divalents	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 1
B1	1.2	0.06M Divalents	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 2
C1	1.3	0.06M Divalents	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 3
D1	1.4	0.06M Divalents	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 4
E1	1.5	0.06M Divalents	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 1
F1	1.6	0.06M Divalents	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 2
G1	1.7	0.06M Divalents	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 3
H1	1.8	0.06M Divalents	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 4
A2	1.9	0.06M Divalents	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 1
B2	1.10	0.06M Divalents	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 2
C2	1.11	0.06M Divalents	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 3
D2	1.12	0.06M Divalents	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 4
E2	1.13	0.09M Halogens	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 1
F2	1.14	0.09M Halogens	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 2
G2	1.15	0.09M Halogens	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 3
H2	1.16	0.09M Halogens	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 4
A3	1.17	0.09M Halogens	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 1
B3	1.18	0.09M Halogens	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 2
C3	1.19	0.09M Halogens	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 3
D3	1.20	0.09M Halogens	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 4
E3	1.21	0.09M Halogens	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 1
F3	1.22	0.09M Halogens	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 2
G3	1.23	0.09M Halogens	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 3
H3	1.24	0.09M Halogens	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 4
A4	1.25	0.09M NPS	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 1
B4	1.26	0.09M NPS	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 2
C4	1.27	0.09M NPS	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 3
D4	1.28	0.09M NPS	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 4
E4	1.29	0.09M NPS	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 1
F4	1.30	0.09M NPS	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 2
G4	1.31	0.09M NPS	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 3
H4	1.32	0.09M NPS	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 4

A5	1.33	0.09M NPS	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 1
B5	1.34	0.09M NPS	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 2
C5	1.35	0.09M NPS	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 3
D5	1.36	0.09M NPS	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 4
E5	1.37	0.12M Alcohols	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 1
F5	1.38	0.12M Alcohols	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 2
G5	1.39	0.12M Alcohols	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 3
H5	1.40	0.12M Alcohols	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 4
A6	1.41	0.12M Alcohols	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 1
B6	1.42	0.12M Alcohols	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 2
C6	1.43	0.12M Alcohols	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 3
D6	1.44	0.12M Alcohols	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 4
E6	1.45	0.12M Alcohols	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 1
F6	1.46	0.12M Alcohols	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 2
G6	1.47	0.12M Alcohols	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 3
H6	1.48	0.12M Alcohols	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 4
A7	2.1	0.12M Ethylene glycols	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 1
B7	2.2	0.12M Ethylene glycols	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 2
C7	2.3	0.12M Ethylene glycols	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 3
D7	2.4	0.12M Ethylene glycols	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 4
E7	2.5	0.12M Ethylene glycols	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 1
F7	2.6	0.12M Ethylene glycols	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 2
G7	2.7	0.12M Ethylene glycols	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 3
H7	2.8	0.12M Ethylene glycols	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 4
A8	2.9	0.12M Ethylene glycols	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 1
B8	2.10	0.12M Ethylene glycols	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 2
C8	2.11	0.12M Ethylene glycols	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 3
D8	2.12	0.12M Ethylene glycols	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 4
E8	2.13	0.12M Monosaccharides	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 1
F8	2.14	0.12M Monosaccharides	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 2
G8	2.15	0.12M Monosaccharides	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 3
H8	2.16	0.12M Monosaccharides	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 4
A9	2.17	0.12M Monosaccharides	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 1
B9	2.18	0.12M Monosaccharides	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 2
C9	2.19	0.12M Monosaccharides	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 3
D9	2.20	0.12M Monosaccharides	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 4
E9	2.21	0.12M Monosaccharides	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 1
F9	2.22	0.12M Monosaccharides	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 2
G9	2.23	0.12M Monosaccharides	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 3
H9	2.24	0.12M Monosaccharides	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 4
A10	2.25	0.1M Carboxylic acids	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 1
B10	2.26	0.1M Carboxylic acids	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 2
C10	2.27	0.1M Carboxylic acids	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 3
D10	2.28	0.1M Carboxylic acids	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 4
E10	2.29	0.1M Carboxylic acids	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 1
F10	2.30	0.1M Carboxylic acids	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 2
G10	2.31	0.1M Carboxylic acids	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 3
H10	2.32	0.1M Carboxylic acids	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 4
A11	2.33	0.1M Carboxylic acids	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 1
B11	2.34	0.1M Carboxylic acids	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 2
C11	2.35	0.1M Carboxylic acids	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 3
D11	2.36	0.1M Carboxylic acids	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 4
E11	2.37	0.1M Amino acids	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 1
F11	2.38	0.1M Amino acids	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 2
G11	2.39	0.1M Amino acids	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 3
H11	2.40	0.1M Amino acids	0.1M Buffer System 1	6.5	50% v/v Precipitant Mix 4
A12	2.41	0.1M Amino acids	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 1
B12	2.42	0.1M Amino acids	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 2
C12	2.43	0.1M Amino acids	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 3
D12	2.44	0.1M Amino acids	0.1M Buffer System 2	7.5	50% v/v Precipitant Mix 4
E12	2.45	0.1M Amino acids	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 1
F12	2.46	0.1M Amino acids	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 2
G12	2.47	0.1M Amino acids	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 3
H12	2.48	0.1M Amino acids	0.1M Buffer System 3	8.5	50% v/v Precipitant Mix 4

Mix name	Composition	Catalogue Number (100 & 250mL)
Divalents	0.3M Magnesium chloride hexahydrate; 0.3M Calcium chloride dihydrate	MD2-100(250)-70
Halogens	0.3M Sodium fluoride; 0.3M Sodium bromide; 0.3M Sodium iodide	MD2-100(250)-71
NPS†	0.3M Sodium nitrate, 0.3 Sodium phosphate dibasic, 0.3M Ammonium sulfate	MD2-100(250)-72
Alcohols	0.2M 1,6-Hexanediol; 0.2M 1-Butanol 0.2M 1,2-Propanediol; 0.2M 2-Propanol; 0.2M 1,4-Butanediol; 0.2M 1,3-Propanediol	MD2-100(250)-73
Ethylene glycols	0.3M Diethylene glycol; 0.3M Triethylene glycol; 0.3M Tetraethylene glycol; 0.3M Pentaethylene glycol	MD2-100(250)-74
Monosaccharides	0.2M D-Glucose; 0.2M D-Mannose; 0.2M D-Galactose; 0.2M L-Fucose; 0.2M D-Xylose; 0.2M N-Acetyl-D-Glucosamine	MD2-100(250)-75
Carboxylic acids	0.2M Sodium formate; 0.2M Ammonium acetate; 0.2M Sodium citrate tribasic dihydrate; 0.2M Sodium potassium tartrate tetrahydrate; 0.2M Sodium oxamate	MD2-100(250)-76
Amino acids	0.2M DL-Glutamatic acid monohydrate; 0.2M DL-Alanine; 0.2M Glycine; 0.2M DL-Lysine monohydrochloride; 0.2M DL-Serine	MD2-100(250)-77

†**NPS;** Nitrate Phosphate Sulfate

Table 3: Buffer systems used in Morpheus®

Mix name	Conc.	pH @ 20°C	Composition	Catalogue Number (100 & 250mL)
Buffer System 1	1.0M	6.5	Imidazole; MES monohydrate (acid)	MD2-100(250)-100
Buffer System 2	1.0M	7.5	Sodium HEPES; MOPS (acid)	MD2-100(250)-101
Buffer System 3	1.0M	8.5	Tris (base); BICINE	MD2-100(250)-102

Table 4: Mixes of Precipitants[1] used in Morpheus®

Mix name	Old Mix Name	Composition	Catalogue Number (100 & 250mL)
Precipitant Mix 1	P500MME_P20K	40% v/v PEG 500* MME; 20 % w/v PEG 20000	MD2-100(250)-81
Precipitant Mix 2	EDO_P8K	40% v/v Ethylene glycol; 20 % w/v PEG 8000	MD2-100(250)-82
Precipitant Mix 3	GOL_P4K	40% v/v Glycerol; 20% w/v PEG 4000	MD2-100(250)-83
Precipitant Mix 4	MPD_P1K_P3350	25% v/v MPD; 25% PEG 1000; 25% w/v PEG 3350	MD2-100(250)-84

*N.B. The PEG 550 MME that was originally used in this screen has been discontinued and replaced with PEG 500 MME.

Helix // Molecular Dimension

Well	Kit #	Salt	Additive	Buffer	pH	Precipitant
A1	1.1	0.05M Potassium chloride 0.1M Lithium chloride	0.012M Spermine tetrahydrochloride	0.05M MES	6.5	25% v/v PEG 400
B1	1.2	none	0.001M Spermine tetrahydrochloride	0.05M MES	6.5	25% v/v PEG 400
C1	1.3	0.1M Sodium chloride 0.05M Lithium chloride	0.01M Magnesium chloride hexahydrate	0.05M MES	6.5	25% v/v PEG 400
D1	1.4	0.2M Potassium chloride	0.005M Hexammine cobalt(III) chloride	0.05M MES	6.5	25% v/v PEG 400
E1	1.5	0.1M Lithium chloride	0.01M Manganese(II) chloride tetrahydrate	0.05M MES	6.5	25% v/v PEG 400
F1	1.6	0.1M Sodium chloride	0.005M Strontium chloride hexahydrate	0.05M MES	6.5	25% v/v PEG 400
G1	1.7	0.05M Potassium chloride 0.1M Lithium chloride	0.012M Spermine tetrahydrochloride	0.05M MES	6.5	35% v/v MPD
H1	1.8	none	0.001M Smprmine tetrahydrochloride	0.05M MES	6.5	35% v/v MPD
A2	1.9	0.1M Sodium chloride 0.1M Lithium chloride	0.01M Magnesium chloride hexahydrate	0.05M MES	6.5	35% v/v MPD
B2	1.10	0.2M Potassium chloride	0.005M Hexammine cobalt(III) chloride	0.05M MES	6.5	35% v/v MPD
C2	1.11	0.1M Lithium chloride	0.01M Manganese(II) chloride tetrahydrate	0.05M MES	6.5	35% v/v MPD
D2	1.12	0.1M Sodium chloride	0.005M Strontium chloride hexahydrate	0.05M MES	6.5	35% v/v MPD
E2	1.13	0.05M Potassium chloride 0.1M Lithium chloride	0.012M Spermine tetrahydrochloride	0.05M MES	6.5	17% v/v PEG 4000
F2	1.14	none	0.001M Spermine tetrahydrochloride	0.05M MES	6.5	17% v/v PEG 4000
G2	1.15	0.1M Sodium chloride 0.05M Lithium chloride	0.01M Magnesium chloride hexahydrate	0.05M MES	6.5	17% v/v PEG 4000
H2	1.16	0.2M Potassium chloride	0.005M Hexammine cobalt(III) chloride	0.05M MES	6.5	17% v/v PEG 4000
A3	1.17	0.1M Lithium chloride	0.01M Manganese(II) chloride tetrahydrate	0.05M MES	6.5	17% v/v PEG 4000
B3	1.18	0.1M Sodium chloride	0.005M Strontium chloride hexahydrate	0.05M MES	6.5	17% v/v PEG 4000
C3	1.19	0.05M Potassium chloride 0.1M Lithium chloride	0.012M Spermine tetrahydrochloride	0.05M MES	6.5	2.6M Sodium malonate dibasic monohydrate

D3	1.20	none	0.001M Ssormine tetrahydrochloride	0.05M MES	6.5	2.6M Sodium malonate dibasic monohydrate
E3	1.21	0.1M Sodium chloride 0.05M Lithium chloride	0.01M Magnesium chloride hexahydrate	0.05M MES	6.5	2.6M Sodium malonate dibasic monohydrate
F3	1.22	0.2M Potassium chloride	0.005M Hexammine cobalt(III) chloride	0.05M MES	6.5	2.6M Sodium malonate dibasic monohydrate
G3	1.23	0.1M Lithium chloride	0.01M Manganese(II) chloride tetrahydrate	0.05M MES	6.5	2.6M Sodium malonate dibasic monohydrate
H3	1.24	0.1M Sodium chloride	0.002M Strontium chloride hexahydrate	0.05M MES	6.5	2.6M Sodium malonate dibasic monohydrate
A4	1.25	0.1M Ammonium sulfate	none	0.05M Bis-Tris	7.0	44% v/v PEG 200
B4	1.26	0.2M Potassium chloride	0.01M Calcium chloride dihydrate	0.05M Bis-Tris	7.0	40% v/v PEG 400
C4	1.27	0.1M Sodium chloride	0.002M Spermine tetrahydrochloride	0.05M Bis-Tris	7.0	37% v/v PEG 1000
D4	1.28	0.1M Sodium chloride 0.2M Potassium chloride	0.02M Magnesium chloride hexahydrate	0.05M Bis-Tris	7.0	35% v/v PEG 2000
E4	1.29	0.05M Potassium chloride	0.005M Spermine tetrahydrochloride	0.05M Bis-Tris	7.0	33% v/v PEG 3350
F4	1.30	0.1M Ammonium sulfate 0.1M Sodium chloride	1.0M Sodium formate	0.05M Bis-Tris	7.0	20% v/v PEG 8000
G4	1.31	none	1.0M Sodium formate	0.05M Bis-Tris	7.0	20% v/v PEG 20,000
H4	1.32	0.05M Lithium sulfate	none	0.05M Bis-Tris	7.0	40% v/v PEG 500 MME
A5	1.33	0.1M Sodium citrate tribasic dihydrate	none	0.05M Bis-Tris	7.0	40% v/v PEG 500 MME
B5	1.34	0.1M Ammonium sulfate	0.002M Spermine tetrahydrochloride	0.05M Bis-Tris	7.0	40% v/v MPD
C5	1.35	none	none	0.05M Bis-Tris	7.0	35% v/v MPD
D5	1.36	0.1M Sodium chloride	0.02M Magnesium chloride hexahydrate	0.05M Bis-Tris	7.0	30% v/v MPD
E5	1.37	0.1M Potassium chloride	none	0.05M Bis-Tris	7.0	27% v/v MPD
F5	1.38	none	5% v/v PEG 400	0.05M Bis-Tris	7.0	24% v/v MPD
G5	1.39	0.2M Lithium sulfate	none	0.05M Bis-Tris	7.0	20% v/v MPD
H5	1.40	none	20% v/v Glycerol 5% w/v PEG 3350	0.05M Bis-Tris	7.0	1.0M Ammonium sulfate
A6	1.41	none	10% v/v Glycerol 5% v/v PEG 500 MME	0.05M Bis-Tris	7.0	1.2M Ammonium sulfate
B6	1.42	none	25% v/v Glycerol	0.05M Bis-Tris	7.0	1.4M Ammonium sulfate
C6	1.43	0.15M Potassium chloride	25% v/v Glycerol	0.05M Bis-Tris	7.0	1.6M Ammonium sulfate
D6	1.44	0.1M Sodium chloride	25% v/v Glycerol	0.05M Bis-Tris	7.0	1.8M Ammonium sulfate
E6	1.45	none	5% v/v PEG 400	0.05M Bis-Tris	7.0	2.0M Ammonium sulfate
F6	1.46	0.1M Potassium chloride	none	0.05M Bis-Tris	7.0	22% w/v PEG 2000 MME
G6	1.47	0.2M Lithium chloride	none	0.05M Bis-Tris	7.0	22% w/v PEG 2000 MME
H6	1.48	0.1M Sodium chloride	none	0.05M Bis-Tris	7.0	22% w/v PEG 2000 MME
A7	2.1	0.1M Potassium chloride	none	0.05M HEPES	6.5	20% v/v MPD
B7	2.2	none	0.005M Spermine tetrahydrochloride	0.05M HEPES	6.5	10% v/v MPD
C7	2.3	0.1M Sodium chloride 0.2M Potassium chloride	none	0.05M HEPES	6.5	10% v/v MPD
D7	2.4	0.2M Sodium chloride	none	0.05M HEPES	6.5	25% v/v MPD
E7	2.5	0.05M Lithium sulfate	none	0.05M HEPES	6.5	15% v/v MPD
F7	2.6	0.3M Sodium chloride	none	0.05M HEPES	6.5	30% v/v MPD
G7	2.7	0.05M Sodium chloride	0.01M Magnesium chloride hexahydrate	0.05M HEPES	6.5	25% v/v PEG 400
H7	2.8	0.2M Potassium chloride	none	0.05M Bis-Tris	7.0	10% v/v PEG 400
A8	2.9	0.05M Lithium sulfate	none	0.05M HEPES	6.5	18% v/v PEG 400
B8	2.10	none	none	0.05M Bis-Tris	7.0	1.5M Lithium sulfate
C8	2.11	0.1M Sodium chloride 0.1M Potassium chloride	none	0.05M HEPES	6.5	1.9M Lithium sulfate
D8	2.12	none	0.02M Magnesium chloride hexahydrate	0.05M Bis-Tris	7.0	15% w/v PEG 2000 MME
E8	2.13	0.05M Lithium sulfate	none	0.05M HEPES	6.5	10% w/v PEG 2000 MME
F8	2.14	none	none	0.05M Bis-Tris	7.0	14% w/v PEG 2000 MME
G8	2.15	0.1M Lithium chloride	none	0.05M Bis-Tris	7.0	10% w/v PEG 500 MME
H8	2.16	0.05M Potassium chloride	none	0.05M Bis-Tris	7.0	1.1M Ammonium sulfate
A9	2.17	0.05M Lithium sulfate	none	0.05M HEPES	6.5	1.4M Ammonium sulfate
B9	2.18	0.1M Sodium chloride	0.01M Lithium chloride	0.05M HEPES	6.5	1.8M Ammonium sulfate
C9	2.19	0.3M Lithium sulfate	none	0.05M Bis-Tris	7.0	18% w/v PEG 1000
D9	2.20	0.3M Sodium chloride	none	0.05M HEPES	6.5	15% w/v PEG 1000
E9	2.21	0.25M Potassium chloride	none	0.05M HEPES	6.5	12% w/v PEG 1000
F9	2.22	0.1M Lithium sulfate	none	0.05M Bis-Tris	7.0	25% v/v PEG 350 MME
G9	2.23	none	none	0.05M HEPES	6.5	20% v/v PEG 350 MME
H9	2.24	0.05M Lithium sulfate	none	0.05M HEPES	6.5	15% v/v PEG 350 MME
A10	2.25	none	0.005M Spermine tetrahydrochloride	0.05M Sodium acetate	4.5	18% v/v MPD
B10	2.26	0.1M Sodium chloride	none	0.05M Sodium acetate	4.5	24% v/v MPD
C10	2.27	0.15M Potassium chloride	none	0.05M Sodium acetate	4.5	32% v/v MPD
D10	2.28	none	0.01M Magnesium chloride hexahydrate	0.05M Sodium acetate	5.0	21% v/v MPD
E10	2.29	0.1M Sodium chloride 0.1M Potassium chloride	0.005M Spermine tetrahydrochloride	0.05M Sodium acetate	5.0	14% v/v MPD
F10	2.30	none	0.05M Lithium chloride	0.05M Sodium acetate	5.0	15% v/v PEG 400
G10	2.31	none	none	0.05M MES	5.5	21% v/v MPD
H10	2.32	none	0.02M Magnesium chloride hexahydrate	0.05M MES	5.5	21% v/v MPD
A11	2.33	none	none	0.05M MES	5.5	12% w/v PEG 2000
B11	2.34	none	0.02M Magnesium sulfate heptahydrate	0.05M MES	5.5	21% v/v MPD
C11	2.35	none	0.005M Spermine tetrahydrochloride	0.05M MES	5.5	15% v/v PEG 400
D11	2.36	none	10% v/v Glycerol	0.05M MES	5.5	1.5M Ammonium sulfate
E11	2.37	0.05M Sodium chloride 0.1M Potassium chloride	0.01M Calcium chloride dihydrate	0.05M Bis-Tris	7.0	25% w/v PEG 2000
F11	2.38	0.25M Sodium chloride	0.005M Spermine tetrahydrochloride	0.05M Bis-Tris	7.0	20% w/v PEG 3350
G11	2.39	0.1M Potassium chloride	0.005M Barium chloride dihydrate	0.05M Bis-Tris	7.0	12% w/v PEG 8000
H11	2.40	0.3M Sodium chloride	none	0.05M Bis-Tris	7.0	8% w/v PEG 20,000
A12	2.41	0.05M Sodium chloride 0.2M Lithium chloride	0.01M Calcium chloride dihydrate	0.05M Bis-Tris	7.0	20% w/v PEG 2000
B12	2.42	0.1M Potassium chloride	0.002M Spermine tetrahydrochloride	0.05M Bis-Tris	7.0	15% w/v PEG 2000 MME
C12	2.43	0.05M Potassium chloride 0.1M Lithium sulfate	0.02M Magnesium sulfate heptahydrate	0.05M Bis-Tris	7.0	15% w/v PEG 2000 MME
D12	2.44	0.1M Sodium chloride	0.002M Barium chloride dihydrate	0.05M Bis-Tris	7.5	6% w/v PEG 20,000
E12	2.45	0.1M Potassium chloride	0.002M Spermine tetrahydrochloride	0.05M Bis-Tris	7.5	7% w/v PEG 8000
F12	2.46	0.05M Lithium sulfate	0.03M Magnesium sulfate heptahydrate	0.05M Bis-Tris	8.0	15% w/v PEG 3350
G12	2.47	0.2M Sodium chloride	none	0.05M Bis-Tris	8.5	9% w/v PEG 3350
H12	2.48	none	0.052M Magnesium chloride hexahydrate	0.05M Bis-Tris	8.5	18% w/v PEG 3350

Q_Compass // Molecular Dimension

Well	Kit #	Salt	Buffer	pH	Precipitant
A1	1.1	0.1 M Potassium chloride	none		12 %(w/v) PEG 8000 / 5 %(w/v) Glycerol
B1	1.2	0.5 M Potassium chloride	none		12 %(w/v) PEG 8000 / 10 %(w/v) Glycerol
C1	1.3	0.2 M Ammonium sulfate	none		15 %(w/v) PEG 8000
D1	1.4	0.5 M Lithium sulfate	none		15 %(w/v) PEG 8000
E1	1.5	0.2 M Sodium acetate	0.1 M MES	6.5	15 %(w/v) PEG 8000
F1	1.6	0.05 M Ammonium sulfate	0.1 M Sodium citrate		15 %(w/v) PEG 8000
G1	1.7	0.2 M Calcium acetate	0.1 M HEPES	7.5	18 %(w/v) PEG 8000
H1	1.8	0.1 M Sodium acetate	0.1 M HEPES	7.5	18 %(w/v) PEG 8000 / 2 %(w/v) Isopropanol
A2	1.9	0.2 M Lithium sulfate	0.1 M TRIS	8.5	18 %(w/v) PEG 8000
B2	1.10	none	0.1 M HEPES	7.5	20 %(w/v) PEG 8000
C2	1.11	0.2 M Magnesium acetate	0.1 M MES	6.5	20 %(w/v) PEG 8000
D2	1.12	none	0.1 M CHES	9.5	20 %(w/v) PEG 8000
E2	1.13	0.2 M Ammonium sulfate	0.1 M MES	6.5	22 %(w/v) PEG 8000
F2	1.14	0.2 M Lithium chloride	none		25 %(w/v) PEG 8000
G2	1.15	0.2 M Ammonium sulfate	none		30 %(w/v) PEG 8000
H2	1.16	none	0.1 M Sodium acetate	4.6	8 %(w/v) PEG 10000
A3	1.17	none	0.1 M Imidazole	8.0	14 %(w/v) PEG 10000
B3	1.18	none	0.1 M TRIS	8.5	16 %(w/v) PEG 10000
C3	1.19	0.1 M Sodium chloride	0.1 M TRIS	8.5	18 %(w/v) PEG 10000 / 20 %(w/v) Glycerol
D3	1.20	none	0.1 M HEPES	7.5	20 %(w/v) PEG 10000
E3	1.21	none	0.1 M TRIS	8.5	30 %(w/v) PEG 10000
F3	1.22	none	0.1 M MES	6.5	10 %(w/v) PEG 20000
G3	1.23	0.1 M Magnesium chloride	0.1 M TRIS	8.5	17 %(w/v) PEG 20000
H3	1.24	none	none		20 %(w/v) PEG 20000
A4	1.25	0.01 M Sodium acetate	none		50 %(w/v) MPD / 15 %(w/v) Ethanol
B4	1.26	0.05M Sodium chloride	0.05 M Sodium acetate		50 %(w/v) MPD / 20 %(w/v) Isopropanol
C4	1.27	0.1 M Ammonium phosphate	0.1 M TRIS	8.5	50 %(w/v) MPD
D4	1.28	none	none		55 %(w/v) MPD
E4	1.29	0.01 M Calcium chloride	0.1 M Sodium acetate	4.6	60 %(w/v) MPD
F4	1.30	0.02 M Sodium acetate	none		60 %(w/v) MPD
G4	1.31	none	0.1 M MES	6.5	70 %(w/v) MPD
H4	1.32	none	0.1 M TRIS	8.5	70 %(w/v) MPD
A5	1.33	0.01 M Calcium chloride	0.1 M TRIS	8.5	20 %(w/v) Methanol
B5	1.34	none	0.1 M TRIS	8.5	2 %(w/v) Ethanol
C5	1.35	none	0.1 M HEPES	7.5	5 %(w/v) Ethanol / 5 %(w/v) MPD
D5	1.36	0.2 M Sodium chloride	0.1 M TRIS	8.5	5 %(w/v) Ethanol / 5 %(w/v) MPD
E5	1.37	none	0.1 M TRIS	8.5	10 %(w/v) Ethanol
F5	1.38	none	0.1 M Sodium acetate	4.6	12 %(w/v) Ethanol / 4 %(w/v) PEG 400
G5	1.39	none	0.1 M TRIS	8.5	14 %(w/v) Ethanol / 5 %(w/v) Glycerol
H5	1.40	none	0.1 M TRIS	8.5	18 %(w/v) Ethanol
A6	1.41	none	none		20 %(w/v) Ethanol
B6	1.42	none	none		20 %(w/v) Ethanol / 10 %(w/v) Glycerol
C6	1.43	0.1 M Sodium acetate	none		30 %(w/v) Ethanol / 10 %(w/v) PEG 6000
D6	1.44	none	none		45 %(w/v) Ethanol
E6	1.45	0.01 M Sodium acetate	none		50 %(w/v) Ethanol
F6	1.46	0.05 M Sodium acetate	none		60 %(w/v) Ethanol / 1.5 %(w/v) PEG 6000
G6	1.47	0.1 M Sodium chloride	none		60 %(w/v) Ethanol
H6	1.48	0.01 M Magnesium sulfate	0.1 M TRIS	8.5	2 %(w/v) Isopropanol
A7	2.1	none	0.1 M HEPES	7.5	5 %(w/v) Isopropanol
B7	2.2	0.2 M Calcium chloride	0.1 M Sodium acetate	4.6	10 %(w/v) Isopropanol
C7	2.3	0.2 M Sodium citrate	0.1 M HEPES	7.5	10 %(w/v) Isopropanol
D7	2.4	0.01 M Magnesium chloride	0.1 M TRIS	8.5	10 %(w/v) Isopropanol
E7	2.5	0.05 M Sodium chloride	0.1 M TRIS	8.5	12 %(w/v) Isopropanol
F7	2.6	0.2 M Sodium citrate	0.1 M MES	6.5	15 %(w/v) Isopropanol
G7	2.7	0.2 M Sodium citrate	0.1 M HEPES	7.5	15 %(w/v) Isopropanol
H7	2.8	0.2 M Magnesium chloride	0.1 M HEPES	7.5	15 %(w/v) Isopropanol
A8	2.9	0.2 M Ammonium acetate	0.1 M TRIS	8.5	15 %(w/v) Isopropanol
B8	2.10	0.2 M Calcium chloride	0.1 M Sodium acetate	4.6	20 %(w/v) Isopropanol
C8	2.11	0.2 M Sodium citrate	0.1 M HEPES	7.5	20 %(w/v) Isopropanol
D8	2.12	0.1 M Magnesium chloride	0.1 M HEPES	7.5	25 %(w/v) Isopropanol
E8	2.13	0.2 M Sodium citrate	0.1 M MES	6.5	30 %(w/v) Isopropanol
F8	2.14	0.2 M Magnesium chloride	0.1 M HEPES	7.5	30 %(w/v) Isopropanol
G8	2.15	0.2 M Ammonium acetate	0.1 M TRIS	8.5	30 %(w/v) Isopropanol
H8	2.16	0.1 M Calcium chloride	0.1 M TRIS	8.5	25 %(w/v) tert-Butanol
A9	2.17	none	0.1 M Sodium citrate	5.6	35 %(w/v) tert-Butanol
B9	2.18	none	none		0.2 M Ammonium dihydrogen phosphate
C9	2.19	none	none		0.2 M Potassium/Sodium tartrate
D9	2.20	none	none		0.2 M Magnesium acetate
E9	2.21	none	none		0.4 M Ammonium dihydrogen phosphate
F9	2.22	none	none		0.4 M Potassium/Sodium tartrate
G9	2.23	none	0.1 M TRIS	8.5	0.4 M Potassium/Sodium tartrate
H9	2.24	0.2 M Sodium citrate	none		0.5 M Ammonium dihydrogen phosphate
A10	2.25	none	0.1 M Imidazole	8.0	0.5 M Sodium acetate
B10	2.26	none	0.1 M HEPES	7.5	0.7 M Sodium citrate
C10	2.27	none	0.1 M TRIS	8.5	0.7 M Lithium sulfate
D10	2.28	none	0.1 M HEPES	7.5	0.8 M Potassium/Sodium tartrate
E10	2.29	none	0.1 M Sodium citrate	5.6	1.0 M Ammonium dihydrogen phosphate
F10	2.30	none	0.1 M TRIS	8.5	1.0 M Ammonium dihydrogen phosphate
G10	2.31	0.01 M Nickel chloride	0.1 M TRIS	8.5	1.0 M Lithium sulfate
H10	2.32	none	0.1 M Imidazole	8.0	1.0 M Sodium acetate
A11	2.33	none	0.1 M Sodium acetate	4.6	1.0 M Sodium formate
B11	2.34	none	0.1 M MES	6.5	1.4 M Sodium acetate
C11	2.35	none	0.1 M HEPES	7.5	1.4 M Sodium citrate
D11	2.36	none	0.1 M TRIS	8.5	1.5 M Lithium sulfate
E11	2.37	none	1 M Sodium citrate	6.5	none
F11	2.38	none	0.1 M MES	6.5	1.6 M Magnesium sulfate
G11	2.39	none	0.1 M MES	6.5	1.6 M Potassium/Sodium tartrate
H11	2.40	none	0.1 M MES	6.5	2.0 M Ammonium formate
A12	2.41	none	0.1 M TRIS	8.5	2.0 M Ammonium dihydrogen phosphate
B12	2.42	none	none		2.0 M Sodium formate
C12	2.43	none	0.1 M TRIS	8.5	2.0 M Magnesium chloride
D12	2.44	0.2 M Sodium acetate	0.1 M MES	6.5	2.0 M Sodium chloride
E12	2.45	none	0.1 M Sodium acetate	4.6	2.0 M Sodium formate
F12	2.46	none	0.1 M TRIS	8.5	1.0 M Ammonium dihydrogen phosphate / 30 %(w/v) Glycerol
G12	2.47	none	0.1 M HEPES	7.5	4.0 M Sodium chloride
H12	2.48	none	none		3.0 M Sodium formate

Q_Pact // Molecular Dimension

Well	Kit #	Salt	Buffer	pH	Precipitant
A1	1.1	none	0.1M SPG buffer	4.0	25% w/v PEG 1500
B1	1.2	none	0.1M SPG buffer	5.0	25% w/v PEG 1500
C1	1.3	none	0.1M SPG buffer	6.0	25% w/v PEG 1500
D1	1.4	none	0.1M SPG buffer	7.0	25% w/v PEG 1500
E1	1.5	none	0.1M SPG buffer	8.0	25% w/v PEG 1500
F1	1.6	none	0.1M SPG buffer	9.0	25% w/v PEG 1500
G1	1.7	0.2M Sodium chloride	0.1M Sodium acetate	5.0	20% w/v PEG 6000
H1	1.8	0.2M Ammonium chloride	0.1M Sodium acetate	5.0	20% w/v PEG 6000
A2	1.9	0.2M Lithium chloride	0.1M Sodium acetate	5.0	20% w/v PEG 6000
B2	1.10	0.2M Magnesium chloride	0.1M Sodium acetate	5.0	20% w/v PEG 6000
C2	1.11	0.2M Calcium chloride	0.1M Sodium acetate	5.0	20% w/v PEG 6000
D2	1.12	0.01M Zinc chloride	0.1M Sodium acetate	5.0	20% w/v PEG 6000
E2	1.13	none	0.1M MIB buffer	4.0	25% w/v PEG 1500
F2	1.14	none	0.1M MIB buffer	5.0	25% w/v PEG 1500
G2	1.15	none	0.1M MIB buffer	6.0	25% w/v PEG 1500
H2	1.16	none	0.1M MIB buffer	7.0	25% w/v PEG 1500
A3	1.17	none	0.1M MIB buffer	8.0	25% w/v PEG 1500
B3	1.18	none	0.1M MIB buffer	9.0	25% w/v PEG 1500
C3	1.19	0.2M Sodium chloride	0.1M MES	6.0	20% w/v PEG 6000
D3	1.20	0.2M Ammonium chloride	0.1M MES	6.0	20% w/v PEG 6000
E3	1.21	0.2M Lithium chloride	0.1M MES	6.0	20% w/v PEG 6000
F3	1.22	0.2M Magnesium chloride	0.1M MES	6.0	20% w/v PEG 6000
G3	1.23	0.2M Calcium chloride	0.1M MES	6.0	20% w/v PEG 6000
H3	1.24	0.01M Zinc chloride	0.1M MES	6.0	20% w/v PEG 6000
A4	1.25	none	0.1M PCB buffer	4.0	25% w/v PEG 1500
B4	1.26	none	0.1M PCB buffer	5.0	25% w/v PEG 1500
C4	1.27	none	0.1M PCB buffer	6.0	25% w/v PEG 1500
D4	1.28	none	0.1M PCB buffer	7.0	25% w/v PEG 1500
E4	1.29	none	0.1M PCB buffer	8.0	25% w/v PEG 1500
F4	1.30	none	0.1M PCB buffer	9.0	25% w/v PEG 1500
G4	1.31	0.2M Sodium chloride	0.1M Hepes	7.0	20% w/v PEG 6000
H4	1.32	0.2M Ammonium chloride	0.1M Hepes	7.0	20% w/v PEG 6000
A5	1.33	0.2M Lithium chloride	0.1M Hepes	7.0	20% w/v PEG 6000
B5	1.34	0.2M Magnesium chloride	0.1M Hepes	7.0	20% w/v PEG 6000
C5	1.35	0.2M Calcium chloride	0.1M Hepes	7.0	20% w/v PEG 6000
D5	1.36	0.01M Zinc chloride	0.1M Hepes	7.0	20% w/v PEG 6000
E5	1.37	none	0.1M MMT buffer	4.0	25% w/v PEG 1500
F5	1.38	none	0.1M MMT buffer	5.0	25% w/v PEG 1500
G5	1.39	none	0.1M MMT buffer	6.0	25% w/v PEG 1500
H5	1.40	none	0.1M MMT buffer	7.0	25% w/v PEG 1500
A6	1.41	none	0.1M MMT buffer	8.0	25% w/v PEG 1500
B6	1.42	none	0.1M MMT buffer	9.0	25% w/v PEG 1500
C6	1.43	0.2M Sodium chloride	0.1M Tris	8.0	20% w/v PEG 6000
D6	1.44	0.2M Ammonium chloride	0.1M Tris	8.0	20% w/v PEG 6000
E6	1.45	0.2M Lithium chloride	0.1M Tris	8.0	20% w/v PEG 6000
F6	1.46	0.2M Magnesium chloride	0.1M Tris	8.0	20% w/v PEG 6000
G6	1.47	0.2M Calcium chloride	0.1M Tris	8.0	20% w/v PEG 6000
H6	1.48	none	0.1M Tris	8.0	20% w/v PEG 6000
A7	2.1	0.2M Sodium fluoride	none		20% w/v PEG 3350
B7	2.2	0.2M Sodium bromide	none		20% w/v PEG 3350
C7	2.3	0.2M Sodium iodide	none		20% w/v PEG 3350
D7	2.4	0.2M Potassium thiocyanate	none		20% w/v PEG 3350
E7	2.5	0.2M Sodium nitrate	none		20% w/v PEG 3350
F7	2.6	0.2M Sodium formate	none		20% w/v PEG 3350
G7	2.7	0.2M Sodium acetate	none		20% w/v PEG 3350
H7	2.8	0.2M Sodium sulphate	none		20% w/v PEG 3350
A8	2.9	0.2M Potassium/sodium tartrate	none		20% w/v PEG 3350
B8	2.10	0.2M Sodium/potassium phosphate	none		20% w/v PEG 3350
C8	2.11	0.2M Sodium citrate	none		20% w/v PEG 3350
D8	2.12	0.2M Sodium malonate	none		20% w/v PEG 3350
E8	2.13	0.2M Sodium fluoride	0.1M Bis Tris propane	6.5	20% w/v PEG 3350
F8	2.14	0.2M Sodium bromide	0.1M Bis Tris propane	6.5	20% w/v PEG 3350
G8	2.15	0.2M Sodium iodide	0.1M Bis Tris propane	6.5	20% w/v PEG 3350
H8	2.16	0.2M Potassium thiocyanate	0.1M Bis Tris propane	6.5	20% w/v PEG 3350
A9	2.17	0.2M Sodium nitrate	0.1M Bis Tris propane	6.5	20% w/v PEG 3350
B9	2.18	0.2M Sodium formate	0.1M Bis Tris propane	6.5	20% w/v PEG 3350
C9	2.19	0.2M Sodium acetate	0.1M Bis Tris propane	6.5	20% w/v PEG 3350
D9	2.20	0.2M Sodium sulphate	0.1M Bis Tris propane	6.5	20% w/v PEG 3350
E9	2.21	0.2M Potassium/sodium tartrate	0.1M Bis Tris propane	6.5	20% w/v PEG 3350
F9	2.22	0.2M Sodium/potassium phosphate	0.1M Bis Tris propane	6.5	20% w/v PEG 3350
G9	2.23	0.2M Sodium citrate	0.1M Bis Tris propane	6.5	20% w/v PEG 3350
H9	2.24	0.2M Sodium malonate	0.1M Bis Tris propane	6.5	20% w/v PEG 3350
A10	2.25	0.2M Sodium fluoride	0.1M Bis Tris propane	7.5	20% w/v PEG 3350
B10	2.26	0.2M Sodium bromide	0.1M Bis Tris propane	7.5	20% w/v PEG 3350
C10	2.27	0.2M Sodium iodide	0.1M Bis Tris propane	7.5	20% w/v PEG 3350
D10	2.28	0.2M Potassium thiocyanate	0.1M Bis Tris propane	7.5	20% w/v PEG 3350
E10	2.29	0.2M Sodium nitrate	0.1M Bis Tris propane	7.5	20% w/v PEG 3350
F10	2.30	0.2M Sodium formate	0.1M Bis Tris propane	7.5	20% w/v PEG 3350
G10	2.31	0.2M Sodium acetate	0.1M Bis Tris propane	7.5	20% w/v PEG 3350
H10	2.32	0.2M Sodium sulphate	0.1M Bis Tris propane	7.5	20% w/v PEG 3350
A11	2.33	0.2M Potassium/sodium tartarte	0.1M Bis Tris propane	7.5	20% w/v PEG 3350
B11	2.34	0.2M Sodium/potassium phosphate	0.1M Bis Tris propane	7.5	20% w/v PEG 3350
C11	2.35	0.2M Sodium citrate	0.1M Bis Tris propane	7.5	20% w/v PEG 3350
D11	2.36	0.2M Sodium malonate	0.1M Bis Tris propane	7.5	20% w/v PEG 3350

E11	2.37	0.2M Sodium fluoride	0.1M Bis Tris propane	8.5	20% w/v PEG 3350
F11	2.38	0.2M Sodium bromide	0.1M Bis Tris propane	8.5	20% w/v PEG 3350
G11	2.39	0.2M Sodium iodide	0.1M Bis Tris propane	8.5	20% w/v PEG 3350
H11	2.40	0.2M Potassium thiocyanate	0.1M Bis Tris propane	8.5	20% w/v PEG 3350
A12	2.41	0.2M Sodium nitrate	0.1M Bis Tris propane	8.5	20% w/v PEG 3350
B12	2.42	0.2M Sodium formate	0.1M Bis Tris propane	8.5	20% w/v PEG 3350
C12	2.43	0.2M Sodium acetate	0.1M Bis Tris propane	8.5	20% w/v PEG 3350
D12	2.44	0.2M Sodium sulphate	0.1M Bis Tris propane	8.5	20% w/v PEG 3350
E12	2.45	0.2M Potassium/sodium tartrate	0.1M Bis Tris propane	8.5	20% w/v PEG 3350
F12	2.46	0.2M Sodium/potassium phosphate	0.1M Bis Tris propane	8.5	20% w/v PEG 3350
G12	2.47	0.2M Sodium citrate	0.1M Bis Tris propane	8.5	20% w/v PEG 3350
H12	2.48	0.2M Sodium malonate	0.1M Bis Tris propane	8.5	20% w/v PEG 3350

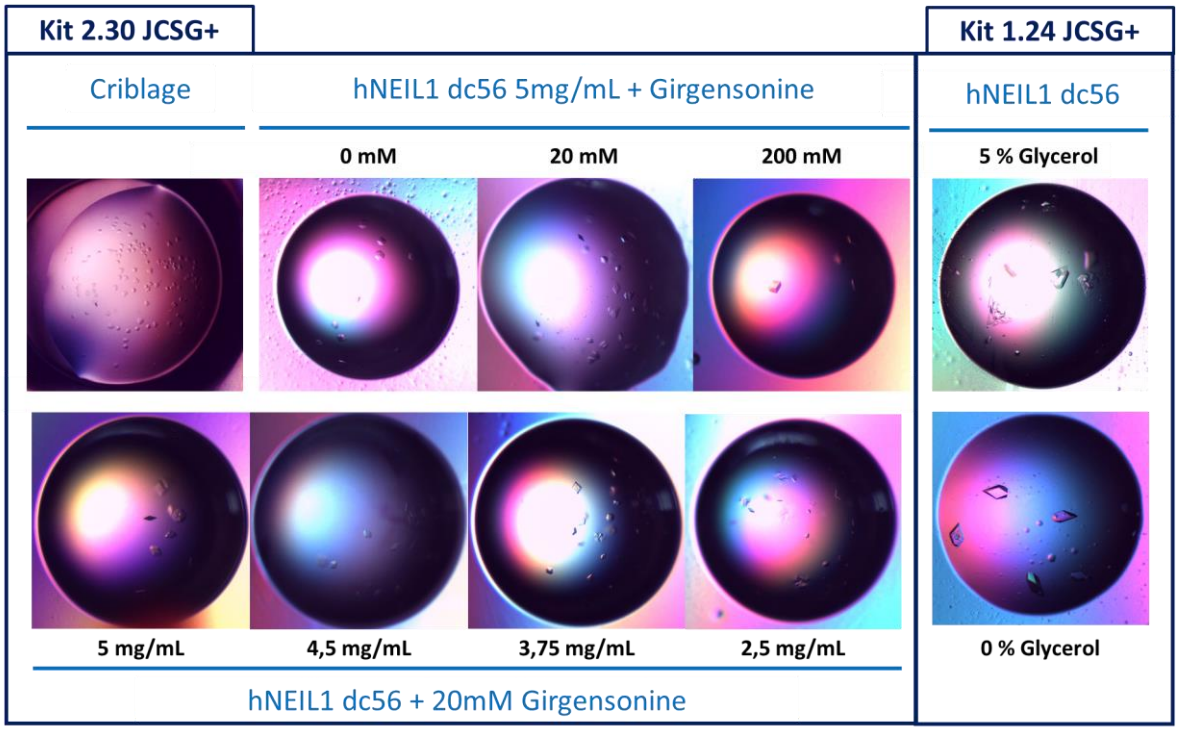
Q_Pact // Molecular Dimension

Well	Kit #	Salt	Buffer	pH	Precipitant
A1	1.1	0.02M Calcium chloride dihydrate	0.1M Sodium acetate	4.6	30% v/v MPD
B1	1.2	0.2M Ammonium acetate	0.1M Sodium acetate	4.6	30% w/v PEG 4000
C1	1.3	0.2M Ammonium sulfate	0.1M Sodium acetate	4.6	25% w/v PEG 4000
D1	1.4	2.0M Sodium formate	0.1M Sodium acetate	4.6	none
E1	1.5	2.0M Ammonium sulfate	0.1M Sodium acetate	4.6	none
F1	1.6	none	0.1M Sodium acetate	4.6	8% w/v PEG 4000
G1	1.7	0.2M Ammonium acetate	0.1M Sodium citrate	5.6	30% w/v PEG 4000
H1	1.8	0.2M Ammonium acetate	0.1M Sodium citrate	5.6	30% v/v MPD
A2	1.9	none	0.1M Sodium citrate	5.6	20% w/v PEG 4000 20% v/v 2-Propanol
B2	1.10	1.0M Ammonium phosphate monobasic	0.1M Sodium citrate	5.6	none
C2	1.11	0.2M Calcium chloride dihydrate	0.1M Sodium acetate	4.6	20% v/v 2-Propanol
D2	1.12	1.4M Sodium acetate trihydrate	0.1M Sodium cacodylate	6.5	none
E2	1.13	0.2M Sodium citrate tribasic dihydrate	0.1M Sodium cacodylate	6.5	30% v/v 2-Propanol
F2	1.14	0.2M Ammonium sulfate	0.1M Sodium cacodylate	6.5	30% w/v PEG 8000
G2	1.15	0.2M Magnesium acetate tetrahydrate	0.1M Sodium cacodylate	6.5	20% w/v PEG 8000
H2	1.16	0.2M Magnesium acetate tetrahydrate	0.1M Sodium cacodylate	6.5	30% v/v MPD
A3	1.17	1.0M Sodium acetate trihydrate	0.1M Imidazole	6.5	none
B3	1.18	0.2M Sodium acetate trihydrate	0.1M Sodium cacodylate	6.5	30% w/v PEG 8000
C3	1.19	0.2M Zinc acetate dihydrate	0.1M Sodium cacodylate	6.5	18% w/v PEG 8000
D3	1.20	0.2M Calcium acetate hydrate	0.1M Sodium cacodylate	6.5	18% w/v PEG 8000
E3	1.21	0.2M Sodium citrate tribasic dihydrate	0.1M Sodium HEPES	7.5	30% v/v MPD
F3	1.22	0.2M Magnesium chloride hexahydrate	0.1M Sodium HEPES	7.5	30% v/v 2-Propanol
G3	1.23	0.2M Calcium chloride dihydrate	0.1M Sodium HEPES	7.5	28% v/v PEG 400
H3	1.24	0.2M Magnesium chloride hexahydrate	0.1M Sodium HEPES	7.5	30% v/v PEG 400
A4	1.25	0.2M Sodium citrate tribasic dihydrate	0.1M Sodium HEPES	7.5	20% v/v 2-Propanol
B4	1.26	0.8M Potasium sodium tartrate tetrahydrate	0.1M Sodium HEPES	7.5	none
C4	1.27	1.5M Lithium sulfate	0.1M Sodium HEPES	7.5	none
D4	1.28	0.8M Sodium phosphate monobasic monohydrate 0.8M Potasium phosphate monobasic	0.1M Sodium HEPES	7.5	none
E4	1.29	1.4M Sodium citrate tribasic dihydrate	0.1M Sodium HEPES	7.5	none
F4	1.30	2.0M Ammonium sulfate	0.1M Sodium HEPES	7.5	2% v/v PEG 400
G4	1.31	none	0.1M Sodium HEPES	7.5	20% w/v PEG 4000 10% v/v 2-Propanol
H4	1.32	2.0M Ammonium sulfate	0.1M Tris	8.5	none
A5	1.33	0.2M Magnesium chloride hexahydrate	0.1M Tris	8.5	30% w/v PEG 4000
B5	1.34	0.2M Sodium citrate tribasic dihydrate	0.1M Tris	8.5	30% v/v PEG 400
C5	1.35	0.2M Lithium sulfate	0.1M Tris	8.5	30% w/v PEG 4000
D5	1.36	0.2M Ammonium acetate	0.1M Tris	8.5	30% v/v 2-Propanol
E5	1.37	0.2M Sodium acetate trihydrate	0.1M Tris	8.5	30% w/v PEG 4000
F5	1.38	none	0.1M Tris	8.5	8% w/v PEG 8000
G5	1.39	2.0M Ammonium phosphate monobasic	0.1M Tris	8.5	none
H5	1.40	0.4M Potasium sodium tartrate tetrahydrate	none		none
A6	1.41	0.4M Ammonium phosphate monobasic	none		none
B6	1.42	0.2M Ammonium sulfate	none		30% w/v PEG 8000
C6	1.43	0.2M Ammonium sulfate	none		30% w/v PEG 4000
D6	1.44	2.0M Ammonium sulfate	none		none
E6	1.45	4.0M Sodium formate	none		none
F6	1.46	0.05M Potassium phosphate monobasic	none		20% w/v PEG 8000
G6	1.47	none	none		30% w/v PEG 1500
H6	1.48	0.2M Magnesium formate dihydrate	none		none
A7	2.1	0.1M Sodium chloride	0.1M BICINE	9.0	30% v/v PEG 500 MME
B7	2.2	2.0M Magnesium chloride hexahydrate	0.1M BICINE	9.0	none
C7	2.3	none	0.1M BICINE	9.0	10% w/v PEG 20,000 2% v/v 1,4-Dioxane
D7	2.4	0.2M Magnesium chloride hexahydrate	0.1M Tris	8.5	3.4M 1,6-Hexanediol
E7	2.5	none	0.1M Tris	8.5	25% v/v tert-Butanol
F7	2.6	1.0M Lithium sulfate 0.01M Nickel(II) chloride hexahydrate	0.1M Tris	8.5	none
G7	2.7	1.5M Ammonium sulfate	0.1M Tris	8.5	12% v/v Glycerol
H7	2.8	0.2M Ammonium phosphate monobasic	0.1M Tris	8.5	50% v/v MPD
A8	2.9	none	0.1M Tris	8.5	20% v/v Ethanol
B8	2.10	0.01M Nickel(II) chloride hexahydrate	0.1M Tris	8.5	20% w/v PEG 2000 MME
C8	2.11	0.5M Ammonium sulfate	0.1M Sodium HEPES	7.5	30% v/v MPD
D8	2.12	none	0.1M Sodium HEPES	7.5	10% w/v PEG 6000 5% v/v MPD
E8	2.13	none	0.1M Sodium HEPES	7.5	20% v/v Jeffamine® M-600
F8	2.14	1.6M Ammonium sulfate 0.1M Sodium chloride	0.1M Sodium HEPES	7.5	none
G8	2.15	2.0M Ammonium formate	0.1M Sodium HEPES	7.5	none
H8	2.16	1.0M Sodium acetate trihydrate 0.05M Cadmium sulfate 8/3-hydrate	0.1M Sodium HEPES	7.5	none
A9	2.17	none	0.1M Sodium HEPES	7.5	70% v/v MPD
B9	2.18	4.3M Sodium chloride	0.1M Sodium HEPES	7.5	none

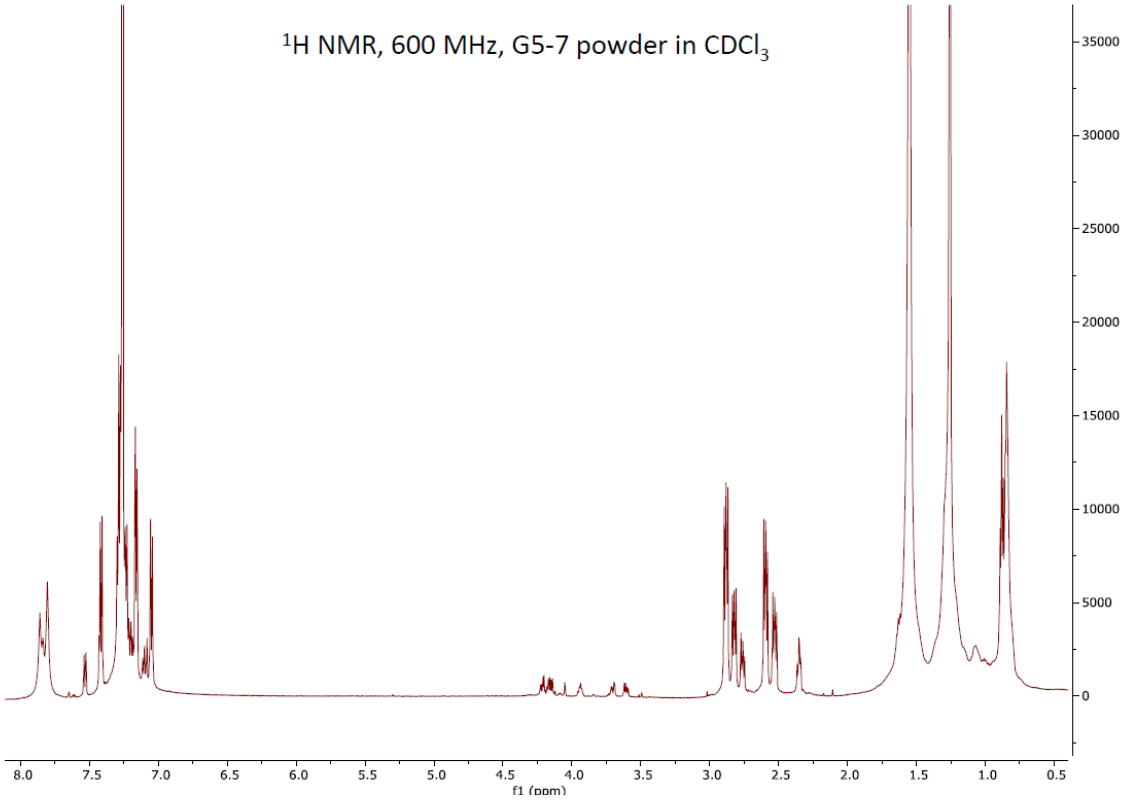
C9	2.19	none	0.1M Sodium HEPES	7.5	10% w/v PEG 8000 8% v/v Ethylene glycol
D9	2.20	1.6M Magnesium sulfate heptahydrate	0.1M MES	6.5	none
E9	2.21	2.0M Sodium chloride 0.1M Potassium phosphate monobasic 0.1M Sodium phosphate monobasic monohydrate	0.1M MES	6.5	none
F9	2.22	none	0.1M MES	6.5	12% w/v PEG 20,000
G9	2.23	1.6M Ammonium sulfate	0.1M MES	6.5	10% v/v 1,4-Dioxane
H9	2.24	0.05M Cesium chloride	0.1M MES	6.5	30% v/v Jeffamine® M-600
A10	2.25	0.01M Cobalt(II) chloride hexahydrate 1.8M Ammonium sulfate	0.1M MES	6.5	none
B10	2.26	0.2M Ammonium sulfate	0.1M MES	6.5	30% w/v PEG 5000 MME
C10	2.27	0.01M Zinc sulfate heptahydrate	0.1M MES	6.5	25% v/v PEG 500 MME
D10	2.28	none	0.1M Sodium HEPES	7.5	20% w/v PEG 10,000
E10	2.29	2.0M Ammonium sulfate 0.2M Potassium sodium tartrate tetrahydrate	0.1M Sodium citrate	5.6	none
F10	2.30	1.0M Lithium sulfate 0.5M Ammonium sulfate	0.1M Sodium citrate	5.6	none
G10	2.31	0.5M Sodium chloride	0.1M Sodium citrate	5.6	4% v/v Polyethyleneimine
H10	2.32	none	0.1M Sodium citrate	5.6	35% v/v tert-Butanol
A11	2.33	0.01M Iron(III) chloride hexahydrate	0.1M Sodium citrate	5.6	10% v/v Jeffamine® M-600
B11	2.34	0.01M Manganese(II) chloride tetrahydrate	0.1M Sodium citrate	5.6	2.5M 1,6-Hexanediol
C11	2.35	2.0M Sodium chloride	0.1M Sodium acetate	4.6	none
D11	2.36	0.2M Sodium chloride	0.1M Sodium acetate	4.6	30% v/v MPD
E11	2.37	0.01M Cobalt(II) chloride hexahydrate	0.1M Sodium acetate	4.6	1.0M 1,6-Hexanediol
F11	2.38	0.1M Cadmium chloride hemi(pentahydrate)	0.1M Sodium acetate	4.6	30% v/v PEG 400
G11	2.39	0.2M Ammonium sulfate	0.1M Sodium acetate	4.6	30% w/v PEG 2000 MME
H11	2.40	2.0M Sodium chloride	none		30% w/v PEG 6000
A12	2.41	0.5M Sodium chloride 0.1M Magnesium chloride hexahydrate 0.01M CTAB	none		none
B12	2.42	none	none		25% v/v Ethylene glycol
C12	2.43	none	none		35% v/v 1,4-Dioxane
D12	2.44	2.0M Ammonium sulfate	none		5% v/v 2-Propanol
E12	2.45	none	1.0M Imidazole	7.0	none
F12	2.46	none	none		10% w/v PEG 1000 10% w/v PEG 8000
G12	2.47	1.5M Sodium chloride	none		10% v/v Ethanol
H12	2.48	none	1.6M Sodium citrate	6.5	none

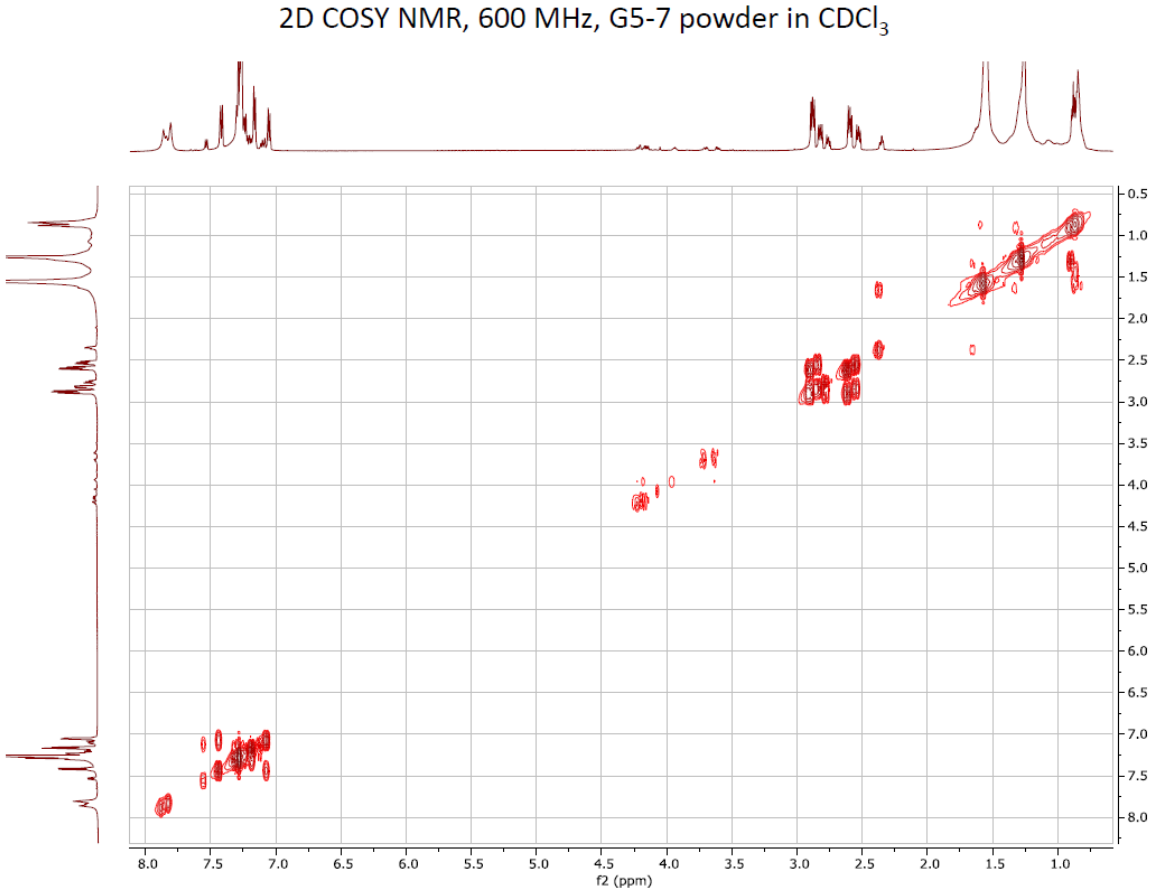
Annexe 9 : Cristaux obtenus lors de l'optimisation du complexe et de hNEIL1dc56

hNEIL1 dc56



Annexe 10 : Analyse RMN de ICOA2-G5-P7





Annexe 11 : Articles et productions scientifiques

Articles dans une revue

1. Tber Z., Biteau N.G., Agrofoglio L., **Cros J.,** Goffinont S., Castaing B., Cyril Nicolas, and Vincent Roy (2019). **Microwave-Assisted Suzuki–Miyaura and Sonogashira Coupling of 4-Chloro-2-(trifluoromethyl)pyrido[1,2-e]purine Derivatives.** *European journal of Organic Chemistry*. **33**, 5756-5767.
2. Rieux C., Goffinont S., Coste F., Tber Z., **Cros J.,** Roy V., Guérin M., Gaudon V., Bourg S., Biela A., Aucagne V., Agrofoglio., Garnier N., Castaing B (2020). **Thiopurine derivative-induced Fpg/Nei DNA 2 glycosylase inhibition: structural, dynamic and 3 functional insights.** *International Journal of Molecular Sciences.* **21**(6) 2058.
3. **Cros J.,** Coste F., Goffinont S., Mimosa Silva de Campo J F., Gaudon V., Berteina-Raboin S., Castaing B. **New small molecule inhibitors of the human 8-oxoguanine-DNA glycosylase hOGG1 : search and preliminary functional and structural characterizations.** *(en préparation)*
4. **Cros J.,** Coste F., Goffinont S., Gabant G., Gaudon V., Roy V., Bernard P., Agrofoglio L., Cadin M., Castaing B. **Nitrile-carrying synthetic or natural compounds are high effective hNEIL1 inhibitors.** *(en préparation)*

Communication orale

Cros J., Coste F., Tber Z., Goffinont S., Gaudon V., Roy V., Agrofoglio L., Castaing B. Small molecule inhibitors of the human DNA glycosylases hOGG1 and NEIL1, pertinent therapeutic targets in cancer and neurodegenerative disease therapies. **25th Young Research Fellows Meeting.** 03/2018, Orléans, France.

Cros J., Coste F., Goffinont S., Gaudon V., Roy V., Garnier N., Guerin M., Berteina-Rabouin S., Routier S., Bernard P., Agrofoglio L., Castaing B. Discovery and characterization of inhibitors of the DNA repair enzymes hNEIL1 and hOGG1, pertinent therapeutic targets for cancer therapy**. International Associated Laboratory.** 06/2019, Military Institue of Medicine, Varsovie, Pologne.

Poster

International or national event

Cros J., Coste F., Goffinont S., Gaudon V., Roy V., Garnier N., Guérin M., Berteina-Rabouin S., Routier S., Agrofoglio L., Castaing B. Screening, characterisation and structural studies of small molecules inhibitors of the DNA glycosylase hNEIL1 and hOGG1, potential therapeutic targets in cancer and neurodegenerative disease. **1s French congress on Integrative Structural Biology (BSI).** 10/2019, Toulouse, France.

Cros J., Coste F., Goffinont S., Gaudon V., Roy V., Garnier N., Guérin M., Berteina-Rabouin S., Routier S., Bernard P., Agrofoglio L., Castaing B. Search for small molecule inhibitors of the DNA repair enzymes hNEIL1 and hOGG1, potential therapeutic targets in cancer. **13th French 3R Meeting.** 05/2019, Presqu'île de Giens, France.

Cros J., Coste F., Goffinont S., Gaudon V., Roy V., Agrofoglio L., Castaing B. Recherche d'inhibiteurs des enzymes de reparation de l'AND hNEIL1 et hOGG1, de nouvelles cibles pertinentes pour le développement de therapies anti-cancéreuses. **31eme Colloque Biotechnocentre.** 10/2018, Seillac, France.

Cros J., Coste F., Tber Z., Goffinont S., Nadan V., Roy V., Agrofoglio L., Castaing B. Search for small molecule inhibitors of DNA glycosylases, pertinent therapeutic targets in cancer and neurodegenerative disease therapies. **5th day of the International Associated Laboratory**. 05/2018, Orléans, France.

Local event

Cros J., Coste F., Goffinont S., Gaudon V., Roy V., Garnier N., Guérin M., Berteina-Rabouin S., Routier S., Bernard P., Agrofoglio L., Castaing B. Search for small molecule inhibitors of the DNA repair enzymes hNEIL1 and hOGG1, potential therapeutic targets in cancer. **17th ADOC Symposium**. 07/2019, Orléans, France.

Cros J., Coste F., Tber Z., Goffinont S., Nadan V., Roy V., Agrofoglio L., Castaing B. Search for small molecule inhibitors of DNA glycosylases, pertinent therapeutic targets in cancer and neurodegenerative disease therapies. **16th ADOC symposium.** 04/2018, Orléans, France.

Cros J., Coste F., Goffinont S., Gaudon V., Roy V., Garnier N., Guerin M., Berteina-Rabouin S., Routier S., Bernard P., Agrofoglio L., Castaing B. Search for small molecule inhibitors of the DNA repair enzymes hNEIL1 and hOGG1, potential therapeutic targets in cancer. **Journée fédération FR2708.** 04/2019, Orléans, France.

Julien CROS

OGG1 et NEIL1, de nouvelles cibles pour la lutte contre le cancer : recherche et caractérisation fonctionnelle et structurale d'inhibiteurs

L'ADN de tout organisme est continuellement endommagé par des agents physiques ou chimiques d'origines endogène ou exogène. Les dommages de l'ADN qui en résultent peuvent être à l'origine de l'apparition de mutations ou de mort cellulaire. Pour pallier à ces effets délétères, les organismes ont développé des systèmes de réparation de l'ADN. Le système de réparation par excision de base (BER) est la voie majeure de réparation des bases endommagées et est initié par des ADN glycosylases telles que hOGG1 et hNEIL1. Ces enzymes reconnaissent spécifiquement les bases lésées et les éliminent. Paradoxalement, les réparations initiées par ces protéines peuvent diminuer l'effet thérapeutique de certains traitements, notamment anti-cancéreux. En exploitant le principe de létalité synthétique, le ciblage thérapeutique de hOGG1 et hNEIL1 pourrait être pertinent pour lutter contre certains cancers, mais aussi contre des maladies neurodégénératives (Huntington) ou des processus inflammatoires pathologiques.

Au travers du criblage « moyen débit » de banques de petites molécules naturelles ou synthétiques, mon travail de thèse a consisté en l'identification de nouveaux inhibiteurs sélectifs de hNEIL1 et hOGG1 et en la caractérisation de leur mode d'action par des études biochimiques et structurales. Si nous avons pu mettre en lumière des fonctions chimiques essentielles et quelques déterminants structuraux et fonctionnels relatifs à leurs modes d'action, de nombreuses zones d'ombre demeurent et mériterons d'être explorés dans le futur. En revanche, l'utilisation d'un homologue archéen de hOGG1, l'enzyme PabAGOG, a permis de proposer un modèle tridimensionnel cohérent d'un complexe hOGG1/inhibiteur pour l'un des meilleurs inhibiteurs que nous avons identifiés. Finalement, ces nouveaux composés comptent parmi les meilleurs inhibiteurs de hOGG1 et hNEIL1 identifiés à ce jour et certains d'entre eux devraient bénéficier d'une évaluation *in cellulo.*

Mots clé : Criblage, enzymes de réparation, cristallographie, BER, hOGG1, hNEIL1, inhibition

hNEIL1 and hOGG1, potential therapeutic targets in cancer : search, functional and structural characterization studies on inhibitors.

DNA is continuously damaged by physical or chemical agents from endogenous or exogenous sources. These damages can induce mutation or cell death. To counteract these deleterious effects, organisms have developed DNA repair systems. The Base Excision Repair System (BER) is the major pathway to repair damaged bases. It is initiated by DNA Glycosylases such as hNEIL1 and hOGG1 which specifically recognize and remove oxidized bases. However, in some situation like conventional cancer treatment, the repairs initiated by these enzymes can lead to therapeutic resistance. Therefore, based on the synthetic lethality concept, selective inhibition of hNEIL1 and hOGG1 could be relevant in some pathologic contexts like cancer, but also against neurodegenerative diseases (Huntington) or pathological inflammatory processes.

This thesis work aimed to identify new selective inhibitors of hNEIL1 and hOGG1 with a "medium throughput" screening of natural or synthetic small molecule libraries, and to characterize their mode of action through biochemical and structural studies. Although we identified some essentials chemicals functions and some leads to understand the action mode of our inhibitors, many areas remains unveiled and deserve to be explored in the future. However, the use of an archaeal homologue of hOGG1, the enzyme PabAGOG, has allowed us to propose a coherent three-dimensional model of an hOGG1/inhibitor complex for one of the best inhibitors we have identified. Finally, these new inhibitors are among the best inhibitors of hOGG1 and hNEIL1 identified to date and some of them will undergo *in cellulo* evaluation.

Keywords : Screening, reparation enzyme, crystallography, BER, hOGG1, hNEIL1, inhibition

CBM, CNRS, Orléans