



HAL
open science

**Etude d'interactions protéine-ADN : Interactions
Ku70/Ku80-ADN au sein de la voie NHEJ et
caractérisation d'inhibiteurs d'une interaction
protéine-ADN cible dans un contexte de recherche de
médicaments**

Amandine Gontier

► **To cite this version:**

Amandine Gontier. Etude d'interactions protéine-ADN : Interactions Ku70/Ku80-ADN au sein de la voie NHEJ et caractérisation d'inhibiteurs d'une interaction protéine-ADN cible dans un contexte de recherche de médicaments. Biochimie, Biologie Moléculaire. Université Paris Saclay (COMUE), 2019. Français. NNT : 2019SACLS553 . tel-02884842

HAL Id: tel-02884842

<https://theses.hal.science/tel-02884842>

Submitted on 30 Jun 2020

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.

« Étude d'interactions protéine-ADN :
interactions Ku70/Ku80-ADN au sein
de la voie NHEJ et caractérisation
d'inhibiteurs d'une interaction protéine-ADN
cible dans un contexte de recherche de médicaments »

Thèse de doctorat de l'Université
Paris-Saclay préparée
à l'Université Paris-Sud
CEA Saclay / Sanofi Vitry-Sur-Seine

Ecole doctorale n°569 : Innovation
thérapeutique du fondamental à
l'appliqué (ITFA)

Spécialité de doctorat :
Biochimie et Biologie Structurale

Thèse présentée et soutenue
à Gif-Sur-Yvette,
le 6 Décembre 2019, par

Amandine Gontier

Composition du Jury :

Sylvie Nessler Professeur des Universités, UP Sud	Président
Eric Ennifar Chercheur, IBMC Strasbourg	Rapporteur
Ludovic Sauguet Chercheur, Institut Pasteur Paris	Rapporteur
Mireille Bétermier Directrice de Recherche, I2BC Gif-Sur-Yvette	Examineur
Bertrand Raynal Ingénieur de Recherche, Institut Pasteur Paris	Examineur
Alexey Rak Responsable BSB, Sanofi Vitry-Sur-Seine	Co-directeur de thèse
Jean-Baptiste Charbonnier Chercheur I2BC, CEA Saclay	Directeur de thèse

La persévérance vient à bout de tout.

Proverbe français

Remerciements

L'histoire a débuté le 27 Juillet 2016 lorsque j'ai reçu l'appel téléphonique du Dr Jean-Baptiste Charbonnier m'annonçant que j'avais été sélectionnée pour l'offre de bourse CIFRE. Cela a été le début d'une superbe expérience qui a réellement démarré le 12 Décembre 2016. Ma thèse s'est déroulée au sein de deux laboratoires académique et privé respectivement à l'« Institut de Biologie Intégrative de la Cellule » situé au CEA de Saclay et au sein du laboratoire « Integrated Drug Discovery » situé chez Sanofi à Vitry-Sur-Seine.

Je tiens à remercier tout d'abord mes deux encadrants de thèse le Dr Jean-Baptiste Charbonnier ; mon directeur de thèse et le Dr Alexey Rak ; mon co-directeur de thèse, pour m'avoir choisie pour cette thèse CIFRE et pour leur accueil au sein de leurs laboratoires respectifs. Vous m'avez appris beaucoup tant au niveau professionnel que personnel et je tiens à vous dire merci pour ces trois années très enrichissantes. J'ai passé certes la plupart de mon temps au sein du CEA mais l'ambiance au sein des biophysiciens de Sanofi était tout aussi sympathique.

Jean-Baptiste, je tiens à te remercier pour ces trois années passées dans la « Team JB » ! Tu as su m'écouter, me faire confiance sur les différents projets et me donner certaines responsabilités. Tes précieux conseils m'ont toujours été d'une grande aide. Merci pour ta motivation et ta bonne humeur qui font de la « Team JB » une équipe soudée, motivée et dans laquelle on ressent vraiment cet esprit d'entraide toujours avec la joie de vivre. On n'oubliera pas non plus la journée d'initiation au karaté sur le plateau du Moulon qui nous a permis d'évacuer le stress du vendredi et d'apprendre à compter japonais !

Alexey, je tiens à te remercier pour ta bienveillance et ton soutien depuis le début de thèse. Rigueur, persévérance et management sont les trois mots que je retiens en premier quant à ce que tu m'as appris. Tu as été là à chaque moment pour me soutenir dans mes projets et challenges et je t'en remercie. J'ai beaucoup apprécié ces trois ans de collaboration avec Sanofi et cela a toujours été un plaisir de venir sur le site de Vitry que ce soit pour des manips, pour des conseils que tu as pu me donner. Merci à toi !

Je tiens à remercier le Pr Sylvie Nessler pour avoir accepté de présider le jury de ma soutenance de thèse. Je remercie également le Dr Ludovic Sauguet et le Dr Eric Ennifar d'avoir accepté de juger mon travail et notamment d'avoir pris le temps de lire mon manuscrit. Enfin je remercie

le Dr Mireille Bétermier ainsi que le Dr Bertrand Raynal d'avoir accepté de juger mon travail en tant qu'examineurs.

Je remercie tout particulièrement mes deux équipes de travail au sein des deux laboratoires. J'ai passé de supers moments avec vous tous ; des moments d'ordre professionnel comme à l'extérieur plus festifs !

Virginie, merci beaucoup pour m'avoir accueillie en tant que nouvelle doctorante au CEA et pour m'avoir formée sur les diverses techniques de laboratoire. Tu es une personne au grand cœur, toujours à l'écoute et indispensable. On retiendra « les week-ends synchrotron » et les nuits passées à Soleil avec nos cristaux de Ku qui ont eu du mal à diffracter (cela m'étonne d'ailleurs de ne pas encore m'être transformée un matin en Ku !), les « pourifs » parfois qui nous mettaient en rogne et nous retardaient dans l'avancée des projets mais aussi et surtout les bons moments entre collègues et amies ! Je voudrais remercier mes copines de labo Audrey « Bristol », Jingqi, Ambre, Paloma, Florence, Nico, Audrey, Sayma, Camille, Gaëlle, Elizabeth, Fatima, Fanny, Cyril, Alicia, Nadège avec qui j'aurai passé trois ans, deux ans voire même qu'une seule année à partager paillasses, coups durs, matériel de labo, cafés, chambres (en congrès !), fous rires, confidences et j'en passe. Vous êtes pour moi de supers amies ! Merci du fond du cœur les filles, vous avez fait de ces trois années, un souvenir inoubliable !

Difficile de faire bref lorsque l'on passe trois ans avec des personnes au top ! Chloé, Seydou, Arun, ma petite Rosa, Catherine (comme ... ma seconde maman au labo), Pascal B, Nathalie, Marie V, Bénédicte, Sarah, Fouad, Kévin, Rania, mes colocs de bureau Manon et Chafiaa, Agathe, Maxime, Gwen, Michel, Manu, Marie, Raphaël, Jessica, Pascal D et toutes les autres personnes que je n'aurais pas citées, merci pour votre sympathie, les pauses déj et cafés en ma compagnie (pauses obligatoires pour garder le moral surtout en mode rédaction !).

Je voudrais remercier également l'équipe de Sanofi, notamment Jacques pour sa gentillesse et ses bons conseils lors des expériences biophysiques, Valérie, toujours la joie de vivre et avec qui j'aurai passé de bons moments dans le bureau à papoter de tout et de rien, François et Annick qui m'ont prise sous leurs ailes quelques semaines pour des manips de RMN et avec qui nous formions une bonne équipe ! Merci à vous pour le temps précieux que vous m'avez consacré. Je remercie Laure et Andreas pour leurs disponibilités et leur aide pour certains projets ainsi que toutes les autres personnes avec qui j'ai partagé de super moments que ce soit en manips, au restaurant, ou à l'extérieur, je ne garderai que de bons souvenirs !

Je remercie les équipes de Patrick Calsou avec Philippe Frit, à l'IBPS Toulouse, d'Eli Rothenberg avec Michael Morten au département de biochimie et de pharmacologie moléculaire, NYU école de médecine ainsi que celle de Mauro Modesti, au CRCM Marseille pour les nombreuses interactions que nous avons pu avoir lors de nos collaborations. Merci aussi à l'équipe de Dynamic Biosensors pour m'avoir accueillie une semaine en formation à Munich et notamment merci à Emeline Barbet-Massin pour ses précieux conseils.

Amis physiciens, chimistes, badminton-iens ou parisiens, merci pour ces moments détente en dehors du boulot. On retient notamment nos afters dans notre QG du Shamrock tous les vendredis soirs, les soirées ambiancées à l'appart ou encore dans les bars, les moments à courir le mercredi soir après un volant en papotant notamment avec mes copines Alicia et Carole. Sabrina, Gabrielle, Maryam et Augustin, merci pour ces moments entre amis et pour votre soutien. Ça a été un plaisir de tous vous rencontrer, de belles amitiés se sont créées. Merci pour tout !

Amies lilloises, amis lillois, merci à vous pour vos petits sms de réconfort, vos appels, nos petits week-ends retrouvailles dans ch'Nord ou sur Paris. Le temps fait que nos chemins se sont séparés pour une question professionnelle mais cela prouve une très belle amitié solide et qui perdure malgré la distance qui nous sépare. Merci.

Je voudrais terminer par dire merci aux plus proches : mes parents Nathalie et Bertrand, ma sœur Manon, mon frère Antoine, mes grands-parents. Merci à vous pour votre soutien, vos appels, nos petits moments en famille toujours réconfortants. Vous avez été là à chaque instant pour moi et vos encouragements m'ont fait grandir et tenir. Merci à ma petite Chanel, ma petite canaille qui, chaque soir, depuis Novembre 2018 a su me faire rire avec ses petites bêtises et ses petits moments câlins (Miaouuu). Enfin, je voudrais remercier mon chéri Didier pour m'avoir soutenue, pour m'avoir préparé de bons petits plats et d'être mon rayon de soleil au quotidien. Tu as su m'apporter amour, réconfort et bienveillance à mon égard. Combativité aura été le mot qui m'a boostée jusqu'au bout ! A tous, je ne vous remercierai jamais assez. Grâce à vous, j'ai réussi à garder mon sourire et ma joie de vivre. Cette nouvelle épreuve m'aura fait murir, devenir ce que je suis aujourd'hui et restera gravée à jamais.

Table des matières

Liste des abréviations	8
INTRODUCTION	11
PARTIE I :	12
Inhibition d'une interaction protéine-ADN : un exemple de recherche en <i>drug discovery</i>	12
A) La Science au service du patient	13
1) <i>Les cancers</i>	13
2) <i>Notion de FBDD</i>	15
3) <i>Exemples de drogues par FBDD</i>	20
B) Croisement de méthodes pour la recherche de nouveaux médicaments	21
1) <i>Evaluation de la possibilité de ciblage in vitro et in vivo</i>	22
2) <i>Identification et validation de hits par RX et biophysique</i>	25
3) <i>Vectorisation des médicaments</i>	29
4) <i>L'essor de nouvelles technologies</i>	30
Premier objectif	38
PARTIE II :	39
Etude d'une interaction protéine-ADN, Ku70/80-ADN, au cœur de la voie NHEJ	39
A) Formation des CDB et réparation de ces cassures	40
1) <i>Les origines diverses des CDB</i>	40
2) <i>Formation des CDB lors d'un traitement thérapeutique</i>	41
3) <i>La voie c-NHEJ chez les mammifères</i>	45
B) Ku70/80 : une protéine indispensable à la voie de réparation c-NHEJ	49
1) <i>Structure de l'hétérodimère Ku70/80</i>	49
2) <i>Reconnaissance des extrémités de l'ADN</i>	52
3) <i>Interaction de Ku avec des partenaires du NHEJ</i>	52
Second objectif	60
MATERIEL & METHODES	61
PARTIE I :	62
Etudes structurales et fonctionnelles des interactions de Ku70/80 avec des facteurs du NHEJ et des petites molécules	62
A) A propos de l'hétérodimère Ku70/80	62
B) Préparation des échantillons protéiques	64
1) <i>Clonage de Ku70/80 et XLF dans des vecteurs pFL</i>	65
2) <i>Expression des protéines en cellules d'insectes</i>	66

3) Purification de l'hétérodimère Ku70/80 et de la protéine XLF	67
C) Préparation des ligands	71
1) Peptides pXLF et pAPLF.....	71
2) Petites molécules.....	72
D) Interactions Ku-ADN et ligands	73
1) Mesures d'interactions par microcalorimétrie	73
2) Mesure d'interactions par nanoleviers d'ADN (technologie du switchSENSE).....	76
3) High-Throughput Differential Scanning Fluorimetry (HT-DSF).....	77
4) Mesures d'interactions et de compétitions par RMN	80
E) Cristallographie des protéines	85
PARTIE II :	93
Interactions protéine-ADN en <i>drug discovery</i>	93
A) Mesures d'interactions par la technologie du switchSENSE	93
1) Oligonucléotides, puces et modes de mesure utilisés.....	93
2) Mise en place de la puce.....	94
3) Les logiciels du switchSENSE	95
B) Caractérisation des interactions protéine – ADN	97
1) Informations sur les échantillons protéiques.....	97
2) Informations sur les oligonucléotides utilisés	98
C) Inhibition des interactions protéine – ADN	98
1) Mise en place d'une stratégie de caractérisation	99
2) Titrations.....	99
RESULTATS.....	100
PARTIE I :	101
L'IP6 : une molécule stimulatrice du NHEJ.....	101
A) Expression et purification des facteurs du NHEJ	101
1) Expression des protéines Ku et XLF en cellules d'insecte.....	101
2) Purification de la protéine Ku70/80 ; version tronquée KuCC	102
3) Purification de la protéine XLF	104
4) Analyse par chromatographie d'exclusion de taille des protéines purifiées	106
B) Structure cristallographique du complexe Ku70/Ku80/hDNA en présence de l'inositol phosphate (IP6).....	107
1) Essais de cristallisation Ku-IP6 / Ku-ADN-IP6.....	107
2) Optimisation des conditions de cristallisation et de diffraction de KuCC-hDNA-IP6	109
3) Diffraction et traitement des données collectées pour le complexe KuCC-hDNA-IP6	111
4) Phasage et affinement.....	112

5) Analyse des données cristallographiques du complexe KuCC-hDNA-IP6	115
6) Comparaison des structures Ku-ADN-IP6 et Ku-ADN-pXLF	116
C) Caractérisation biophysique de l'interaction entre Ku et l'IP6	117
D) Interactions simultanées de l'IP6 et du peptide pXLF	128
PARTIE II :	134
Stratégies d'inhibition du NHEJ : un croisement de méthodes	134
A) Un criblage d'une collection de fragments Sanofi sur Ku par DSF	134
B) Etude d'interactions et de compétitions par RMN	137
PARTIE III :	145
Caractérisation d'inhibiteurs d'une interaction protéine-ADN dans un projet de Recherche Sanofi de candidats médicaments	145
A) Caractérisation des interactions protéine-ADN	145
B) Caractérisation d'inhibiteurs par titrations	152
CONCLUSIONS ET PERSPECTIVES	156
PARTIE I :	157
A) Caractérisation des interactions protéine-ADN pour un projet industriel d'oncologie	157
B) Caractérisation d'inhibition des interactions protéine-ADN	157
PARTIE II :	158
A) Caractérisations biochimique et structurale du complexe KuCC-hDNA-IP6	158
B) Etude biophysique de l'interaction Ku-IP6 et Ku-ADN-IP6	160
ANNEXES	163

Liste des abréviations

ADMET: Adsorption Distribution Metabolism Excretion Toxicity studies

ADN : Acide DésoxyriboNucléique

ARNm : Acide RiboNucléique messenger

CAII : Carbonic Anhydrase II

Cter : domaine C-terminal

CDB : Cassure Double Brin

CRISPR : Clustered Regularly Interspaced Short Interspaced Palindromic Repeats

cryo-EM : cryo-Electron Microscopy

CV : Column Volume

DBS : Dynamic BioSensors

DMSO : diméthylsulfoxyde

DPA : Date of Proliferation Arrest

DR : Dynamic Response

DRX² : appareil utilisé durant cette thèse pour les approches switchSENSE®

DSB : Double-Strand Break

DSF : Differential Scanning Fluorimetry

DSS : acide 3-triméthylsilylpropane-1-sulfonique

ϵ_0 : coefficient d'extinction molaire

EDTA : Ethylène Diamine Tétra-Acétique

EMSA : Electrophoretic Mobility Shift Assay

FACS : Fluorescence-Assisted Cell Sorting

FBDD : Fragment-Based Drug Discovery

FDA : Food and Drug Administration

FID : Free Induction Decay

FL : Full-Length

Fluo-AUC : Fluo-Analytical Ultra Centrifugation

FT : Flow-Through

QC : Quality Control

hDNA : hairpin DNA

HT-DSF : High Throughput-Differential Scanning Fluorimetry

IDT : Integrated DNA Technologies

IND : Investigational New Drug
IP6 : Inositol hexakisphosphate
ITC : Isothermal Titration Calorimetry
IR : Ionizing Radiation
KBM : Ku Binding Motif
KO : KnockOut
Ku : protéine hétérodimérique Ku70/80
MOI : Multiplicity Of Infection
MPC : Multi-Purpose bioChip
MST : MicroScale Thermophoresis
NDA : New Drug Application
NHEJ : Non-Homologous End Joining
Ni : Nickel
NMR : Nuclear Magnetic Resonance
NOE : Nuclear Overhauser Effect
NS : Nombre de Scans
Nter : domaine N-terminal
ON : overnight
PAGE : Electrophorèse sur Gel de PolyAcrylamide
PBS : Phosphate Buffer Saline
PDB : Protein Data Bank
ppm : partie par million (10^{-6})
QC : Quality Control
RMN : Résonance Magnétique Nucléaire
ROS : Reactive Oxygen Species
SAXS : Small-Angle X-Ray Scattering
SDS : Sodium DodécylSulfate
SEC : Size Exclusion Chromatography
SPR : Surface Plasmon Resonance
STD : Saturation-Transfer Difference
TCEP : Tris-CarboxylEthylPhosphine
Tm : melting Temperature

TLS : Translation, Libration and Screw rotation

TSA : Thermal Shift Assay

TTD : Therapeutic Target Database

vWA : domaine von Willebrand

WL : Water-Ligand Observation with Gradient Spectroscopy

YFP : Yellow Fluorescent Protein

¹⁹F : sonde fluor

INTRODUCTION

PARTIE I :

Inhibition d'une interaction protéine-ADN : un exemple de recherche en *drug discovery*

Longtemps, les substances actives naturelles ont constitué le cœur des médicaments nouveaux. C'est l'exemple typique de la pénicilline, découverte en 1928 par le médecin britannique Alexander Fleming. Cet antibiotique a montré ses vertus thérapeutiques de manière accidentelle suite à des expériences réalisées sur les propriétés des staphylocoques. Fleming réalisait un grand nombre d'expériences sur des boîtes de culture, souvent de manière désordonnée. Après avoir décidé de mettre un peu de désinfectant pour pallier aux problèmes de contaminations, il se rendit compte lors de la venue d'un de ses visiteurs, que certaines cultures n'ayant pas été totalement immergées dans le désinfectant, possédaient des zones sur lesquelles aucune bactérie n'était présente. Ces zones étaient situées autour d'un champignon nommé *Penicillium notatum*. C'est ce champignon qui secrète une substance naturelle aux propriétés antibiotiques, la pénicilline à l'origine du nom du médicament.

Les années passent, la Science évolue. A côté de la Recherche toujours très active, de nouvelles substances naturelles comme médicaments, des démarches de conception rationnelle ou de cribles haut débit ont vu le jour et se sont développées dans l'industrie pharmaceutique. Des bibliothèques de petites molécules sont alors testées notamment sur des cellules hôtes. Ces petites molécules sont soit des extraits, des substances synthétisées ou encore des produits naturels. Les molécules « gagnantes » sont celles qui ont un effet thérapeutique souhaité dans le domaine de la pharmacologie. Celles-ci sont ainsi sélectionnées et optimisées pour la création de nouveaux médicaments.

C'est ensuite la découverte de la molécule d'ADN dans la seconde moitié du XIX^{ème} siècle par le biochimiste suisse Friedrich Miescher qui fit en quelque sorte rebondir de nouveau la Science. C'est en 1932 que l'on montre que la molécule d'ADN peut se transmettre de cellules en cellules et provoquer certaines maladies. Dans les années 1930, les premières études sur la structure de l'ADN sont effectuées avec la caractérisation de longs filaments composés d'une succession de bases espacées de 0.34 nm et en 1953 la publication de la structure cristallographique de la double hélice d'ADN. Dans les années 2000, le séquençage du génome humain ouvre la voie à l'étude des principaux gènes humains ainsi qu'à leur clonage et à la

purification des protéines correspondantes dont certaines vont constituer de nouvelles cibles thérapeutiques majeures. En parallèle, des bibliothèques de petites molécules de taille plus importante sont construites et testées contre des cibles d'intérêt impliquées dans un grand nombre de maladies. La recherche de nouveaux médicaments se fait alors par des expériences *in vitro* qui sont alors validées *in cellulo* et sur des animaux pour en assurer leur efficacité. Ces cribles peuvent également être réalisés directement *in cellulo*. Ce champ de recherche de nouveaux médicaments est appelé *drug discovery*.

Deux sujets sont traités dans ce manuscrit. Le point commun entre ces deux sujets est l'utilisation des méthodes de biophysique et de biologie structurale pour étudier des interactions protéines-ligands dans une démarche de *drug discovery*.

A) La Science au service du patient

1) Les cancers

A l'heure actuelle, environ 400 000 nouveaux cas de cancers sont détectés en France (données de la Ligue contre le Cancer). Les cas observés chez les hommes sont plus nombreux que ceux observés chez les femmes même si cette différence s'atténue progressivement. Depuis 2005, une baisse des cancers est notée indépendamment du sexe avec notamment chez les femmes le progrès de la médecine contre le cancer du sein.

Commençons par définir quelques termes importants. Le mot cancer est un terme très général pour désigner un grand nombre de maladies. Parfois, ce terme est employé à tort car il ne s'agit pas réellement de ce type de pathologie. Il faut savoir que l'on parle de cancer lorsque des cellules de notre organisme se transforment et deviennent nuisibles en proliférant de manière progressive et excessive. Il faut de plus faire attention à ne pas confondre les termes « génétique » et « héréditaire ». On dit d'un cancer qu'il est génétique car on parle de mutagenèse de gènes mais cela ne veut pas dire que celui-ci est à titre héréditaire. En effet il n'y a pas forcément de lien de génération. On ne recense d'ailleurs aujourd'hui qu'une minorité de cancers héréditaires, de l'ordre de 5 %. Les mutations de certains gènes vont amener des changements dans les voies de signalisation telles que la prolifération cellulaire, la migration cellulaire ou l'homéostasie.

Plusieurs caractéristiques propres au cancer sont énumérées en 2000 (Hanahan et Weinberg 2000) et c'est en 2008 que l'on parle de métabolisme des cellules tumorales (Kroemer et Pouyssegur 2008). Les cellules anormales forment ce qu'on appelle une tumeur maligne. On parle de tumeur à partir d'un regroupement de 100 000 cellules anormales. Elles acquièrent plusieurs caractéristiques importantes : prolifération, résistances et autosuffisance en signaux de croissance. En fonction de l'état d'avancée de la maladie et de sa gravité, les cellules cancéreuses vont migrer plus ou moins rapidement loin de cette première tumeur et coloniser un autre organe. Elles forment alors une nouvelle tumeur que l'on appelle métastase. Celle-ci aura le soutien de cellules avoisinantes pour proliférer et grossir grâce aux vaisseaux lymphatiques ou sanguins qui apportent les besoins en oxygène et nutriments. De plus, une tumeur est en réalité constituée de sous-populations de cellules cancéreuses ayant chacune sa sensibilité à un traitement donné. Pour qu'une thérapie soit complète, il est souvent nécessaire de combiner plusieurs traitements pour minimiser l'indice de récurrence du cancer. Les cellules tumorales sont d'ailleurs plus sensibles à de nombreux stress par rapport aux cellules saines (Luo, Solimini, et Elledge 2009). Lors d'un cancer, le but est de tuer les cellules tumorales et de limiter la mort des cellules saines environnantes. La radiothérapie reste aujourd'hui un des traitements les plus fréquemment utilisés. On compte environ 200 000 personnes traitées par radiothérapie en France par an. Ce traitement crée des cassures de l'ADN dans une région ciblée. Les cellules tumorales sont plus sensibles à ces dommages de l'ADN par rapport aux cellules saines de par leur métabolisme. Différentes recherches visent à identifier des nouvelles molécules susceptibles de potentialiser la radiothérapie. Mais de quoi s'agit-il réellement ?

La radiothérapie utilise les rayonnements X ionisants permettant de créer localement des cassures double brin et simple brin de l'ADN et induire l'apoptose de ces cellules. Chez les mammifères, c'est la voie de réparation Non-Homologue de l'ADN (voie NHEJ) qui est la plus utilisée pour réparer les cassures double brin de l'ADN. De nombreux laboratoires dont celui du CEA de Saclay cherchent à caractériser de nouveaux inhibiteurs du NHEJ pour améliorer l'efficacité de la radiothérapie sur les tumeurs et permettre éventuellement d'utiliser des doses plus faibles de rayons X.

On appelle médicament toute substance naturelle ou chimique contenant un principe actif et ayant un effet curable ou préventif sur l'organisme. Une autre substance appelée excipient peut être ajoutée au principe actif pour permettre au patient une administration plus simple. Cet excipient peut être un conservateur, permettre une forme facilitant l'administration du

médicament ou encore un aliment et ne doit en aucun cas interagir avec le principe actif. De plus, trouver un nouveau médicament est un processus long, coûteux et complexe. Différentes approches permettent leur conception. Parmi elles, une approche occupe une place croissante depuis une quinzaine d'années. Il s'agit de la méthode par criblage de molécules fragments (appelée FBDD pour *Fragment-Based Drug Discovery* en anglais). Tom Blundell a notamment été une des personnes qui a contribué au développement de cette approche par criblage de fragments avec la création en 1999 de la compagnie « Astex Technology » (Figure 1).

Cette compagnie est centrée sur la recherche et le développement de nouvelles drogues pour le domaine de l'oncologie avec de nombreuses molécules issues de l'approche FBDD. Astex réalise à la fois l'analyse structurale de cibles majeures par rayons X mais aussi la sélection de molécules fragments ayant une compatibilité avec la forme du site actif de la cible. Cette entreprise est soutenue et collabore par de grandes industries pharmaceutiques comme Novartis, GSK ou encore AstraZeneca. Leur plateforme « Pyramid™ » intéresse ces industries car elle permet l'identification rapide de nouveaux candidats médicaments pour des pathologies bien ciblées. Peu de temps avant la création de cette entreprise, une seconde compagnie californienne « SuperGen » a été créée et a proposé un traitement contre la leucémie aiguë myéloblastique appelé Dacogen®. En 2011, ces deux entreprises ont fusionné pour donner « Astex Pharmaceuticals » que le groupe « Otsuka » a racheté en 2013 pour 900 millions de dollars (Figure 1). A l'heure actuelle, Astex produit de nombreux candidats médicaments dont certains sont en phase clinique.

2) Notion de FBDD

L'approche fragments est apparue à la fin des années 1990 dans les laboratoires Abbott (Shuker et al. 1996) et dans le laboratoire de Tom Blundell. Les fragments sont des molécules de petite taille (poids moléculaire inférieur à 300 g.mol^{-1}) (Figure 2). Ils correspondent généralement à une fraction des molécules classiquement utilisées pour le criblage haut-débit classique qui sont de taille environ une fois et demie plus grande (criblage *High Throughput Screening* ou HTS avec des molécules dites *lead-like* ou *drug-like*).

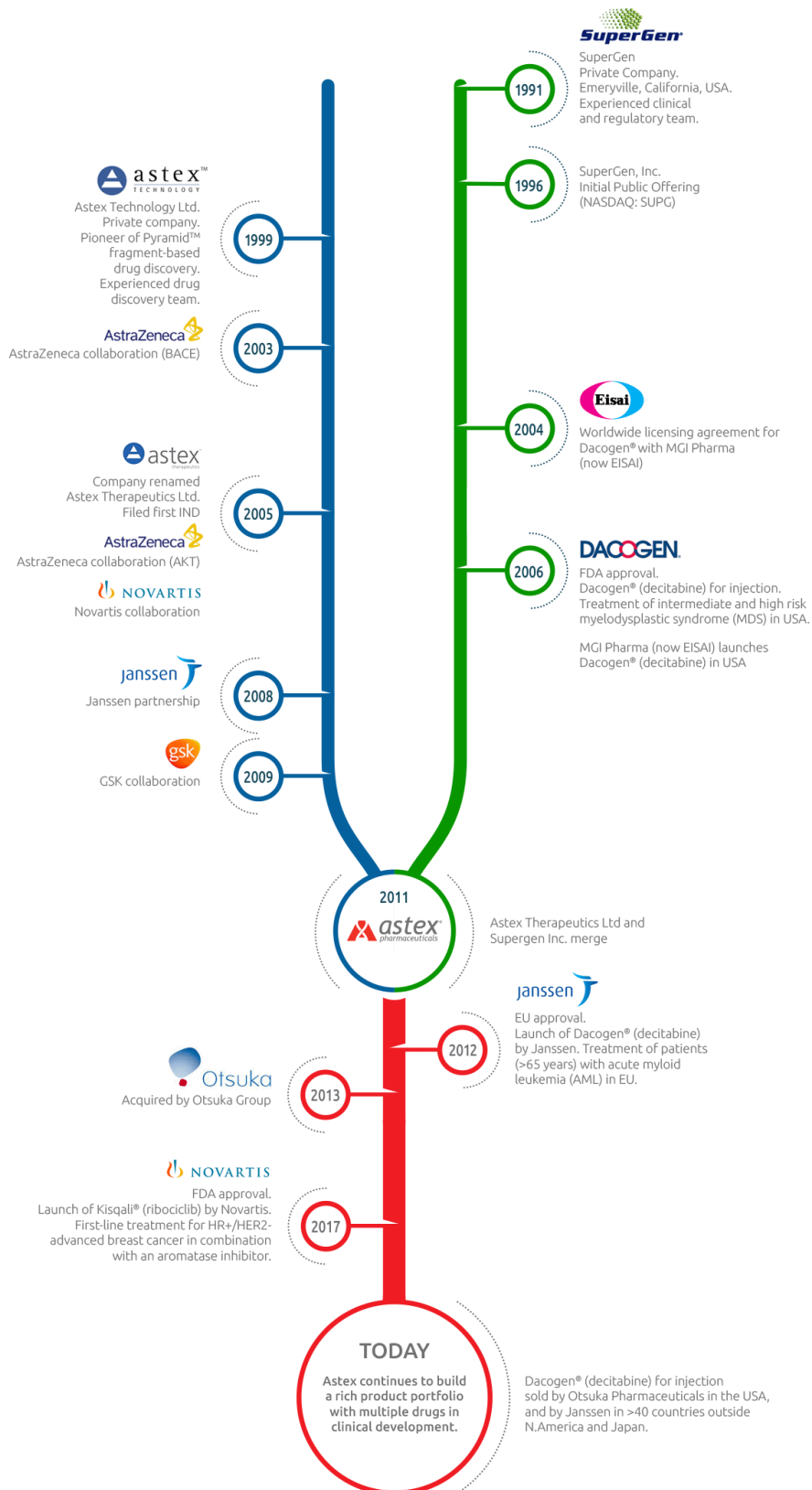


Figure 1 : Historique de l'industrie pharmaceutique « Astex Pharmaceuticals » (<https://astx.com/discover-astex/our-history/>)

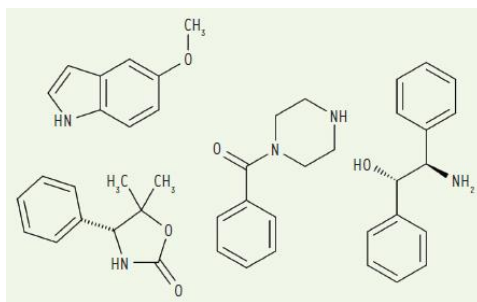


Figure 2 : Exemples de fragments utilisés pour le processus de FBDD (extraits d'une chimiothèque de l'Institut des Sciences Analytiques de l'Université de Lyon 1, librairie FragmentTech utilisée pour du criblage par RMN) (Krimm 2015)

L'approche FBDD a pour objectif de cribler rapidement une banque de taille réduite de composés simples contre la cible thérapeutique étudiée. Les banques de fragments ne dépassent généralement pas les 10 000 composés. Les mesures effectuées lors du premier crible sont des mesures biophysiques *in vitro*. La constante de dissociation mesurée entre les fragments les plus actifs et la cible d'intérêt est fréquemment faible (K_D entre 10 μ M et 10 mM).

L'approche FBDD consiste à sélectionner 2 – 3 fragments de faible affinité pour les faire évoluer vers une forte affinité en les liant ensemble s'ils occupent des poches proches sur la cible (Figure 3). Le FBDD utilise les informations structurales des complexes entre la cible et ces fragments pour guider cette optimisation des fragments dans une seconde étape. Les fragments identifiés lors du premier stade sont appelés « *hits* ». On note qu'en absence de biologie structurale rationnelle, la FBDD ne peut définir le mécanisme de liaison de ces fragments et guider la suite de l'étude par croissance de fragments, de liaison ou de fusion. Plusieurs étapes décrivent l'approche fragments dans le développement et la découverte d'un médicament (Figure 3) :

1. Design d'une librairie de fragments :

Cette librairie peut compter entre une centaine et 10 000 composés. Les fragments sélectionnés lors de la première étape d'un crible ont généralement une faible affinité pour leur cible (K_D entre 10 μ M et 10 mM). Ils sont criblés à haute concentration. Selon la méthode de détection et l'approche utilisée, alors la concentration des fragments utilisée lors du criblage peut varier de 200 μ M à 50 mM. En phase aqueuse, ils sont habituellement concentrés à 1 mM. Ces fragments doivent être très solubles et exempts de toute toxicité particulière. Au départ ils étaient choisis selon trois critères principaux (Någren 2003) :

- Une masse moléculaire inférieure à 300 g.mol⁻¹,
- Un rapport octanol / eau log P inférieur à 3¹ concernant la solubilité des composés chimiques dans les solvants
- Un ratio accepteur / donneur de liaisons hydrogènes inférieur à 3

Aujourd'hui, pour qu'une molécule soit qualifiée de fragment, est ajouté le critère de solubilité devant être supérieur à 500 µM. Cette étape de conception de banque de fragments est une étape clé du processus car ils doivent être appropriés par rapport aux informations disponibles (notamment la structure du site actif, la nature des ligands naturels) à l'étude et capables d'être transformés par la suite.

2. **L'identification et la validation de candidats médicaments par différentes méthodes de criblage :**

Ces étapes font l'objet principal de ma thèse. Je les détaillerai dans une prochaine partie

3. **Elaboration / optimisation du fragment :**

Le but de cette étape est d'obtenir un composé efficace sur la cible à partir de fragments identifiés peu actifs isolément (Scott et al. 2012). Plusieurs procédures sont utilisées : (a) Fusion de plusieurs fragments dont on a pu caractériser par des études structurales l'interaction avec différents sites sur la cible. Ces sites doivent être suffisamment proches pour pouvoir relier les fragments par des liaisons covalentes tout en maintenant leur interaction avec leur site respectif ; (b) Modification chimique du fragment pour ajouter des groupements fonctionnels permettant des interactions supplémentaires avec la cible à partir de données structurales ; (c) Fusion de fragments qui interagissent avec la cible avec des fractions qui se chevauchent. Les modes d'interactions des différents fragments sont utilisés pour dessiner un nouveau fragment comportant la fraction commune et l'ensemble des fonctions des différents fragments identifiés ; (d) Deux composés peuvent se lier simultanément sur la cible grâce à une molécule déjà présente (sur la cible).

Ces différentes stratégies permettent d'augmenter l'affinité et l'activité existant entre la cible et le(s) fragment(s) identifiés.

Approche FBDD

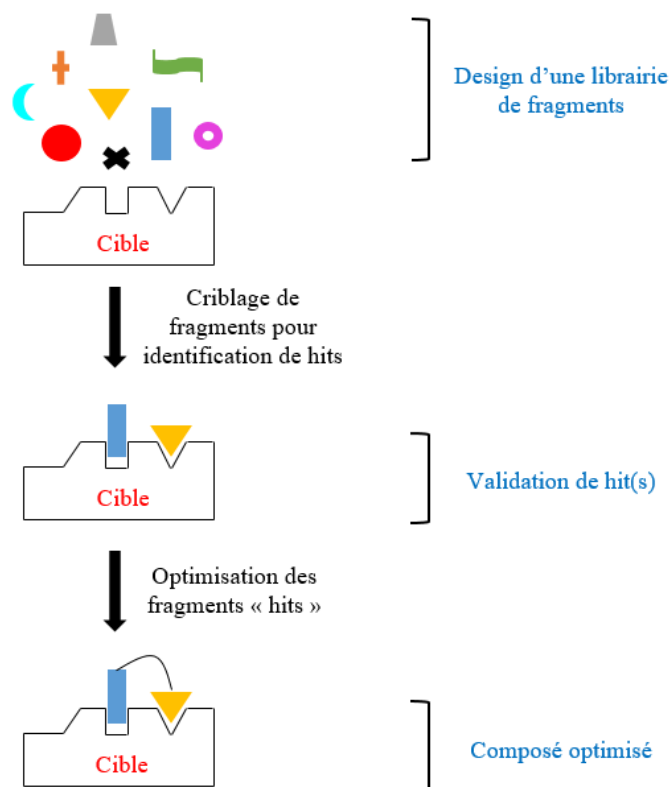


Figure 3 : Schéma décrivant les étapes principales d'une approche fragments en drug discovery

L'approche FBDD présente un certain nombre d'avantages. Elle permet de cribler un très grand ensemble de molécules en terme de composition chimique à partir d'une librairie de départ petite (Scott et al. 2012; Boyd, Turnbull, et Walse 2012). Ceci représente un gain de temps significatif. Les fragments présentent une affinité faible pour la cible d'intérêt en début de processus. Ils ont cependant une capacité de liaison forte. La capacité de liaison est définie comme l'énergie de liaison du complexe divisée par le nombre d'atomes du fragment (exceptés les atomes d'hydrogène). Les fragments ont une capacité de liaison supérieure à 0,3 kcal/mol/atome (Schultes et al. 2010). Cette notion permet d'obtenir l'activité de chaque fragment en fonction de sa taille. Le but est donc lors du processus de FBDD, de lier covalamment entre eux des fragments pour en faire des molécules de plus grande taille plus affines. Cette approche est utilisée par un nombre croissant de compagnies pharmaceutiques et de laboratoires académiques (Figure 4). Cette approche est notamment intéressante dans la recherche de nouveaux inhibiteurs d'interactions protéine-protéine (Mullard 2012; Valkov et al. 2012) notamment pour des cibles complexes à étudier. Les perspectives à venir seraient de

pouvoir utiliser davantage de fragments issus de sources naturelles et d'agrandir l'espace chimique exploité.

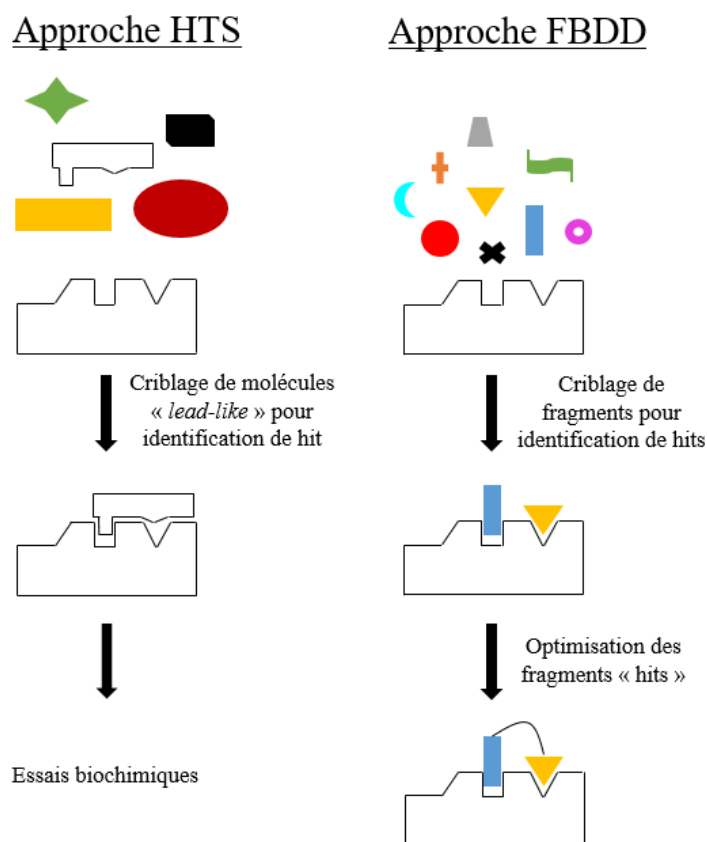


Figure 4 : Schéma comparant les deux approches « criblage haut-débit » (HTS) et « *Fragment Based Drug Discovery* » (FBDD)

3) Exemples de drogues par FBDD

Une vingtaine de médicaments ont été découverts par l'approche FBDD depuis 2012 (Erlanson 2011). Nous citerons ici quatre médicaments à titre d'exemples.

Le premier exemple est le **Vemurafenib** qui a été le premier médicament découvert par FBDD. Cette molécule a été approuvée en 2011 par la *US Food and Drug Administration* (FDA) après des recherches publiées en 2008 (Tsai et al. 2008). Cette molécule cible la sérine/thréonine kinase BRAF et plus particulièrement le mutant V600E de BRAF. Cette mutation est présente chez des patients présentant des mélanomes (60 % des mélanomes ont la mutation V600E). BRAF est une protéine impliquée dans la signalisation cellulaire et la croissance cellulaire. Le Vemurafenib est développé par Plexxikon et Hoffmann-La Roche et utilisé lors de cancers de la peau dits de mélanomes avancés.

Un second succès important du FBDD est la découverte de la molécule **Navitoclax** développée par les laboratoires Abbott. Cette molécule est un inhibiteur oral sélectif des protéines de type leucémie B 2 (que l'on note Bcl-2) qui stoppe ou ralentit le développement de tumeurs. C'est une molécule qui mime le domaine appelé BH3 (pour Bcl-2 *homology 3 domain*) et cible les protéines antiapoptotiques de la famille Bcl-2. Le Navitoclax est actuellement en phase II de recherche clinique pour soigner différents cancers tels que celui de la prostate, des lymphomes ou encore de la leucémie lymphoïde chronique. Son arrivée en phase III risque cependant d'être difficile car ce médicament est faiblement efficace.

Un troisième composé, identifié par l'approche FBDD, semble intéressant pour des patients atteints de la maladie d'Alzheimer. Cette pathologie est caractérisée par des pertes de la mémoire et des troubles des temps et espace dus à une dégénérescence lente des neurones. Des plaques β -amyloïdes sont alors accumulées et produisent une phosphorylation excessive de la protéine Tau et une toxicité des cellules nerveuses. La molécule mise en avant ici est le **Verubecestat** développé par les laboratoires Merck, qui est un inhibiteur de la bêta-sécrétase 1 (appelée BACE), une enzyme nécessaire à la formation des plaques amyloïdes. Elle permet un entretien à long terme pour limiter la production des bêta-amyloïdes. Cette nouvelle molécule permet, à dose maximale (60 mg), de réduire de plus 80 % les taux de protéines bêta-amyloïde. Cette troisième illustration montre de nouveau un aspect prometteur de l'approche FBDD.

B) Croisement de méthodes pour la recherche de nouveaux médicaments

Afin de trouver de nouveaux médicaments, il est primordial de choisir la bonne cible thérapeutique que l'on souhaite viser à travers le traitement. C'est la première étape clé en *drug discovery* qui permet notamment aux scientifiques du monde entier d'avancer sur certaines voies de signalisation et de comprendre certaines interactions à l'échelle moléculaire ou atomique. Il faut comprendre la maladie à traiter. C'est ainsi qu'interviennent les différentes collaborations entre laboratoires académiques et laboratoires industriels. On parle de besoin non satisfait quand une pathologie n'a pas encore de médicament sur le marché ou que celle-ci possède un médicament non-accepté pour cause d'effets secondaires potentiels sur le patient. De plus, lorsqu'une maladie fait son apparition, c'est qu'il y a un dysfonctionnement ou une altération de certaines fonctions de l'organisme. Afin de pallier à la notion de non-satisfait, les chercheurs se doivent d'identifier correctement la cible la plus intéressante.

1) Evaluation de la possibilité de ciblage *in vitro* et *in vivo*

La cible doit être impliquée dans un mécanisme physiopathologique (Gashaw et al. 2011). Plusieurs schémas sont possibles pour aboutir à la validation d'une cible. Le premier schéma part de banques de médicaments connus et ceux-ci sont testés sur plusieurs cibles nouvelles. Le second schéma est *a contrario* de trouver la cible avant de pouvoir cribler des banques de fragments. Dans ce cas, après identification de la cible et de son site actif, des chimiothèques sont alors criblées dans le but de trouver la molécule qui aura un effet sur cette cible (Figure 5). Les deux approches sont complémentaires et peuvent être testées en parallèle pour ainsi permettre un gain de temps. Certes, partir d'un médicament pour voir son effet sur une cible, et dans son contexte biologique semble être une meilleure approche que l'inverse. Seulement, il est plus facile et moins coûteux de démarrer une étude de *drug discovery* en connaissant dès le départ la cible purifiée car elle permet de rationaliser plus rapidement des molécules nouvelles et efficaces. Mon travail de thèse s'est basé sur des études dans lesquelles nous connaissions les cibles. Nous appelons cette approche « *target-based drug discovery* » où la découverte du médicament est fondée sur la cible de départ connue.

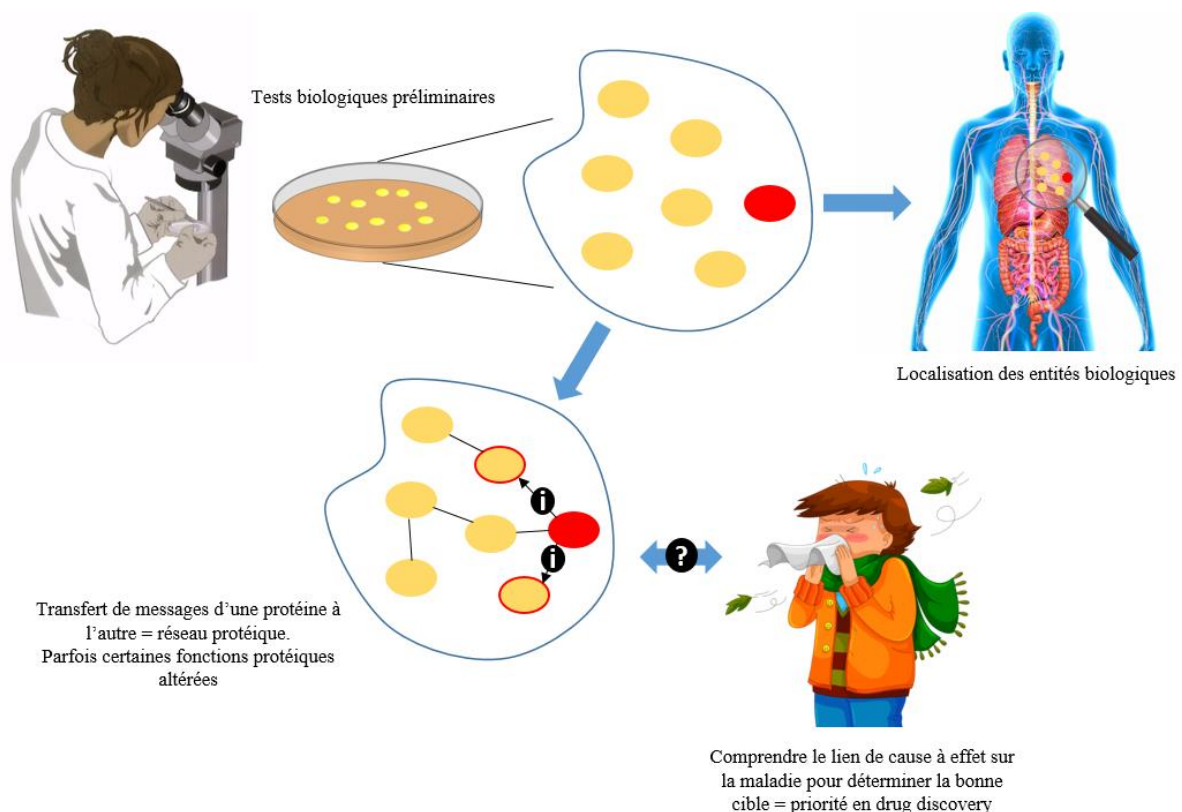


Figure 5 : Schéma désignant la validation d'une cible thérapeutique (adaptation d'une vidéo <https://www.youtube.com/watch?v=3GI0gAcW8rw>)

Mais qu'est-ce qu'une bonne cible thérapeutique ? Une cible peut être de nature diverse. C'est une entité chimique ou biochimique pouvant aller de l'ARN jusqu'aux protéines dont l'activité peut être altérée par un agent thérapeutique. Plusieurs critères de sélection sont proposés pour considérer une cible comme intéressante. Un grand nombre de cibles sont des macromolécules impliquées dans des mécanismes de signalisation que l'on va chercher à modifier soit avec une molécule chimique d'un faible poids moléculaire ou avec une macromolécule (protéine ou ARN) (Figure 6) (Gashaw et al. 2011). La cible doit, si possible, n'être exprimée que dans une ou plusieurs parties du corps que l'on souhaite traiter et non pas dans l'organisme entier. Il est important de pouvoir évaluer les effets possibles de la drogue sur la distribution de la cible et sur son activité. Pour des drogues ayant un effet inhibiteur, il est préférable de connaître les propriétés de mutants ou de souris *knockout* (KO) pour la cible pour connaître les conséquences de son inhibition.

Ces critères déterminent l'intérêt d'une cible d'après la possibilité de moduler *in vivo* ses activités (on parle de *druggability* en anglais). On parle de *ligandability* (en anglais) lorsqu'il est possible de moduler une cible par des molécules synthétiques pouvant interagir spécifiquement avec elle. Ces molécules n'étant pas considérées encore comme des médicaments, on ne peut pas encore parler de *druggability*. Ces deux termes désignent donc le potentiel d'une cible par rapport à la possibilité de développer une petite molécule pour moduler son activité. La *ligandability* est ainsi essentielle mais non suffisante pour trouver de nouveaux médicaments. La *druggability* est plus complexe puisque qu'elle met en jeu plusieurs facteurs de type pharmacodynamiques et pharmacocinétiques qui sont liés à la distribution tissulaire et à la localisation cellulaire. Mon projet de thèse sur une cible liant l'ADN avec Sanofi et sur Ku70/80 au CEA, s'inscrit dans l'évaluation de la capacité à moduler *in vitro* deux cibles et se rattache donc principalement à la notion de *ligandability*.

Un classement au niveau de la puissance de *ligandability* peut être effectué en fonction de l'interaction qui est retrouvée entre une cible et un ligand :

- Faible *ligandability* \Leftrightarrow taux de succès du candidat faible > 1 mM
- *Ligandability* modérée $\Leftrightarrow 0,1 <$ taux de succès du candidat intermédiaire < 1 mM
- Haute *ligandability* \Leftrightarrow taux de succès du candidat $< 0,1$ mM

Il faut noter qu'une faible *ligandability* amène toujours un échec lors d'un criblage haut-débit.

Médicament	Classes de cibles	Mode d'action
Composé chimique de petit poids moléculaire	Enzymes	Inhibiteurs / activateurs
	Récepteurs	Agonistes / antagonistes / modulateurs / activateurs allostériques / sensibilisateurs
	Facteurs de transcription	Inhibiteurs / activateurs
	Canaux ioniques	Inhibiteurs / ouvreurs
	Protéines de transport	Inhibiteurs
	Interfaces protéine-protéine	Inhibiteurs d'interaction protéine-protéine
	Acides nucléiques	Alkylation / réactions de complexation et d'intercalation
	Produits biologiques	Protéines (extracellulaires)
Récepteurs transmembranaires / protéines extracellulaires		Protéines recombinantes
Récepteurs de surface cellulaire		Conjugué anticorps-médicament
Substrats et métabolites		Clivage enzymatique
Acides nucléiques	ARN	ARN interférent

Figure 6 : Classement de trois grandes familles de drogues, des types de facteurs qu'elles ciblent et des effets qu'elles peuvent avoir sur ces facteurs (d'après Gashaw et al. 2011)

Après identification de la cible, sa validation doit être établie. La première est l'aspect reproductibilité de l'expérience réalisée par différents laboratoires. La deuxième étape est l'introduction de variations environnementales. De plus, la modification du type de cellule ne doit pas changer pour autant l'efficacité du ligand. Pour ce faire, différentes méthodes de validation sont utilisées. Les petits ARN interférents (siRNA) sont utilisés (Lindsay 2003; Jain 2004) pour supprimer temporairement la cible d'intérêt via l'inhibition de l'ARN messager (ARNm) du gène codant correspondant. Ils miment en quelque sorte le médicament en modulant l'ARNm. L'approche CRISPR/Cas9 est de plus en plus en plein essor depuis quelques années dans des milliers de laboratoires. Ces « ciseaux moléculaires » sous forme d'ARN guide sont capables de couper une séquence d'ADN double brin ciblée grâce à la collaboration d'une enzyme appelée Cas9 (Choi et Meyerson 2014). Ceci permet notamment de réparer un gène déficient ou de totalement modifier ce gène. Suite à ces approches, si la cible est démontrée comme importante pour le phénotype étudié, alors celle-ci est considérée comme une cible prometteuse.

Aujourd'hui, sur l'ensemble des projets de *drug discovery*, peu sont commercialisés à l'échelle industrielle. Un médicament échoue fréquemment lors des phases cliniques malgré des

propriétés intéressantes dans les études *in vitro* ou *in vivo* en amont. C'est pourquoi il est essentiel d'identifier de bons candidats médicaments.

2) *Identification et validation de hits par RX et biophysique*

Une partie de ma thèse s'inscrit dans la recherche et la validation de candidats fragments potentiellement intéressants contre deux cibles majeures en oncologie : une première cible étudiée par Sanofi et une seconde cible étudiée au CEA : le complexe Ku70/80 de la voie de réparation des cassures double-brin de l'ADN (voir partie suivante). Je décris dans le paragraphe qui suit l'approche que nous avons utilisée pour l'identification et la validation de candidats médicaments sur ces deux projets.

Un nombre élevé de chimiothèques de fragments est disponible. Ces banques sont soit de nature privée lorsque celles-ci appartiennent à des industries pharmaceutiques soit de nature publique lorsque celles-ci sont disponibles librement dans le commerce. Pour mes deux projets en parallèle, j'ai utilisé des collections de fragments de nature privée développées par Sanofi. Lors d'une approche fragments, ceux-ci peuvent être utilisés individuellement ou sous forme de mélange puis par déconvolution des mélanges qui ont donné des pistes (*hits*). Dans ce travail, les fragments sont testés de manière isolée.

L'approche par fragments nécessite une mise au point du crible pour être capable de détecter des affinités faibles au départ. De par leur taille, dans les premières étapes, on attend des affinités entre 10 μ M et 10 mM qui nécessite un essai compatible avec l'utilisation de ligands à des concentrations élevées.

Dans l'étape de validation des fragments intéressants, une seule méthode ne suffit pas à sélectionner les meilleurs candidats. En effet, en raison d'une faible affinité d'interaction et un niveau élevé de faux positifs dans cette gamme de détection d'affinité, l'application d'une combinaison de diverses méthodes biophysiques permet de réduire le signal au bruit et de réduire le niveau des faux positifs. Les laboratoires qui font du FBDD utilisent un ensemble de méthodes biophysiques et des analyses par biologie structurale pour classer les fragments les plus affins et positionner ces fragments sur la structure 3D de la cible. Un certain nombre de méthodes de validation haut-débit des interactions entre les fragments et les cibles sont plus fréquemment utilisées. Nous décrirons brièvement certaines d'entre-elles.

La SPR (pour *Surface Plasmon Resonance* en anglais) est une méthode biophysique qui utilise une interaction physique de type lumière-matière (Englebienne, Hoonacker, et Verhas 2003). La cible est fixée sur une puce recouverte d'une couche métallique. Lorsque les plasmons de surface sont en résonance avec l'onde incidente, alors les électrons libres présents dans le métal oscillent. Ces plasmons sont sensibles à des micro-variations de l'indice de réfraction à proximité de la surface. La SPR est un réfractomètre, un instrument à base d'optique qui mesure la densité de l'échantillon à la surface de la puce. La SPR permet de mesurer des valeurs d'affinité à l'équilibre et d'obtenir les constantes cinétiques d'association et de dissociation de l'interaction. La SPR permet l'interaction de différents types de macromolécules (protéine, ADN, ARN) et dans certains cas de cellules avec des ligands greffés sur la puce.

La microcalorimétrie permet de caractériser les valeurs de thermodynamiques entre cible et ligand. L'affinité de liaison comprend à la fois l'entropie et l'enthalpie. Les données de la littérature indiquent que la mesure de l'enthalpie de l'interaction et son évolution au cours de l'optimisation de candidats est une donnée importante pour guider cette optimisation (Ferenczy et Keserü 2010). Cette enthalpie (notée ΔH) doit être la plus grande possible et négative. La composante enthalpique d'une interaction comprend l'énergie liée aux interactions moléculaires telles que les interactions de Van der Waals, les ponts salins ou encore les ponts hydrogènes et une part de l'énergie liée aux changements d'hydratation du ligand et de la cible. La composante entropique (notée ΔS) comprend l'énergie liée aux interactions hydrophobes et une part de l'énergie liée aux changements d'hydratation du ligand et de sa cible. La microcalorimétrie est une méthode puissante pour déterminer ces valeurs d'enthalpie et d'entropie caractérisant l'affinité de liaison entre un ligand et une cible. L'ITC (pour *Isothermal Titration Calorimetry* en anglais) est une méthode permettant de quantifier la chaleur dégagée ou absorbée lors d'une interaction moléculaire entre un récepteur (dans la cellule de mesure) et un ligand (dans la seringue d'injection) (Krell 2008). De plus cette interaction est caractérisée par l'équation de l'énergie de Gibbs notée :

$$\Delta G^\circ = -RT \ln K_a = RT \ln K_d = \Delta H - T \Delta S$$

avec ΔG° (kcal.M⁻¹) : énergie libre de Gibbs standard,
R (kcal.M⁻¹) : constante des gaz parfaits,
T (K) : température,
K_a (M⁻¹) : constante d'association à l'équilibre
K_d (M) : constante de dissociation à l'équilibre

L'énergie libre est en fait la somme d'un terme enthalpique ΔH (kcal.M⁻¹) et d'un terme entropique $-T\Delta S$ (kcal.M⁻¹) comme indiqué dans la formule. Cette méthode est rarement appliquée dans la méthode de criblage car elle nécessite une quantité assez importante en cible. Cependant, elle est fréquemment utilisée à l'étape d'optimisation. La méthode est décrite dans la partie « Matériel et Méthodes ».

Dans ce projet de thèse, j'ai pu utiliser plusieurs des approches citées ci-dessus. Afin d'appuyer nos données et de valider certains candidats, nous avons décidé de combiner plusieurs méthodes avec l'appui d'une nouvelle technologie : le switchSENSE (Cléry et al. 2017). Chaque technique a ses avantages et la combinaison de ces différentes méthodes permet d'appuyer nos arguments et de confirmer si un fragment est validé comme ligand spécifique. La sensibilité et la rapidité des méthodes de biophysique ont contribué à l'essor de l'approche FBDD ces dernières années. Les laboratoires qui travaillent dans ce domaine effectuent une veille technologique permanente pour suivre l'apparition de nouvelles approches de biophysique susceptibles d'améliorer encore le processus de sélection et de validation des meilleurs fragments. C'est dans ce contexte que l'I2BC et Sanofi ont décidé de collaborer pour évaluer l'apport de la technologie du switchSENSE sur un projet de FBDD de Sanofi.

Les études structurales jouent un rôle clé dans les premières étapes de FBDD pour décrire le site d'interaction des fragments avec la protéine cible. La cristallographie et la RMN sont fréquemment utilisées et plus récemment certaines structures à haute résolution par cryo-EM suggèrent que cette dernière approche pourrait avoir un rôle croissant.

Dans une approche FBDD, la structure 3D de la cible, seule ou en complexe avec un de ses partenaires, est résolue dès le début du projet. Les fragments identifiés dans le crible sont alors mis en contact avec les cristaux par trempage dans la solution de cristallisation ou sont mis en contact dès le départ des essais de cristallisation (co-cristallisation). De nombreuses structures par rayons X peuvent alors être déterminées si le fragment se fixe et si le pouvoir de diffraction des cristaux de la cible n'est pas affecté par le trempage ou la co-cristallisation. Les fortes concentrations utilisées en cristallogénèse (10 mg.mL⁻¹ – 20 mg.mL⁻¹ c'est-à-dire 0,5 mM – 1 mM pour une protéine de 200 acides aminés) sont favorables à la fixation de fragments mis en quantité 10 fois supérieure à la concentration molaire protéique. La cristallographie peut être utilisée comme méthode de criblage directement en trempant un nombre important de fragments dans les cristaux de protéines (Newman et al. 2005). Pour mener ces études de cristallographie

dans le cadre d'une étude FBDD, les robots de cristallisation permettent d'effectuer le trempage à haut-débit. Certaines lignes de synchrotron permettent également de passer en haut débit des cristaux et de déterminer des structures de manière quasi-automatique (Ligne Massif-1 de ESRF, Grenoble). La cristallographie peut être appliquée à des protéines de toute taille et permet d'avoir accès à des informations avec une haute résolution spatiale. C'est la méthode principale utilisée en FBDD pour connaître le mode de liaison d'un ligand sur une cible.

Pour un certain nombre de projets, la RMN (pour Résonance Magnétique Nucléaire) peut s'avérer un outil important dans une étude de FBDD. La cible peut ne pas cristalliser ou les fragments peuvent affecter le pouvoir de diffraction des cristaux. La RMN est alors une technique robuste pour détecter les interactions cible-fragments. Cette technique spectroscopique peut être utilisée de deux façons.

En premier, elle permet de caractériser la structure 3D d'une protéine. Cette approche est rapide pour des protéines de 100 à 200 acides aminés et devient plus délicate à partir de 300 acides aminés. A partir de l'attribution des résonances de chaque atome des résidus de la protéine et de la connaissance de sa structure, on peut enregistrer des spectres en présence de fragments et identifier le site de liaison en suivant le déplacement des pics. La RMN, comme la cristallographie, nécessite une forte concentration en protéine (0,5 mM – 1 mM) et permet de détecter des candidats ayant une faible affinité pour la cible. Cette méthode spectroscopique permet également de mesurer l'affinité de l'interaction (Dalvit 2009; Cala, Guillièrè, et Krimm 2014).

Une seconde approche RMN consiste à marquer les fragments avec du fluor. Cette approche a été utilisée dans le projet et est décrite dans la partie « Matériel et Méthodes ». D'autres méthodes de criblage par RMN sont retrouvées avec les avantages et limites (Wyss et al. 2012) (Figure 7).

Lors de ma thèse j'ai été amenée à utiliser la méthode de détection en suivant le spectre de ligands. J'ai réalisé des expériences de STD (pour *Saturation Transfer Difference* en anglais) (Mayer et Meyer, 2001), de Water-LOGSY (Dalvit 2009) et des compétitions pour la recherche de candidats potentiellement intéressants pour la voie NHEJ (notions expliquées dans la partie « Matériel et Méthodes »). Pour ce faire, la petite molécule d'inositol phosphate (IP6) et l'ADN ont été mis en compétition avec des fragments d'une chimiothèque Sanofi (partie suivante du manuscrit). La détection par suivi en RMN du ligand a été choisie pour quatre raisons : (i) la construction choisie de la protéine Ku a un poids moléculaire de 130 kDa, (ii) les expériences

ne nécessitent pas de marquage isotopique, (iii) la méthode est peu consommatrice en fragments et enfin (iv) cette approche permet d'estimer le pourcentage d'inhibition d'un médicament sur sa cible et obtenir sa valeur d'IC₅₀.

La cristallographie par rayons X et la RMN sont deux méthodes complémentaires et nécessaires pour une étude de *drug discovery*. Pour mon projet avec Sanofi, toutes ces techniques ont été appliquées pour pallier aux limites spécifiques de chaque technique.

Méthode de criblage	Avantages	Limites
Détectées par la cible		
	Détection d'affinités faible et haute	Consommation élevée en protéine marquée
	Basée sur la structure	Limitée aux petites protéines (40 – 60 kDa)
	Informations sur site de liaison	Concentration protéique élevée (25 – 80 µM)
	Détection des sites de liaison spécifiques	Structure 3D requise + attribution des résidus des sites actifs
	Robustesse / fiabilité	
	Étude par SAR pour affinités faibles	
	Étude par SAR pour ligands de haute affinité par compétition	
Détectées par le ligand		
STD / Water-LOGSY	Pas de marquage protéique	Pas de révélation de site de liaison du ligand
	Concentration protéique requise faible (1 – 5 µM)	Signal pouvant être dû à des sites multiples de liaison
	Pas de limitation de taille protéique	
Compétitions STD / Water-LOGSY	Site de liaison & K _i relatif au marqueur avec site de liaison connu et K _D	Requiert un marqueur avec KD connu et/ou le site de liaison
	Détection de ligands d'affinité faible et haute	

Figure 7 : Méthodes de criblage par RMN en FBDD : avantages et limites (adaptée de (Renaud et al. 2016))

3) Vectorisation des médicaments

Après validation de la cible d'intérêt et du ou des candidat(s) médicament(s), la question est de savoir comment envoyer cette nouvelle molécule vers les tissus et cellules ciblés tout en

maintenant l'activité de la molécule. C'est la notion de vectorisation de médicament. Le but ici est de maximiser la distribution du principe actif spécifiquement sur la cible d'intérêt et de limiter les risques de toxicité. Cette vectorisation des médicaments est en plein essor. La molécule est encapsulée dans des vecteurs de taille nanométrique de natures diverses. On trouve des médicaments encapsulés par des lipides. Prenons l'exemple du premier nanomédicament commercialisé appelé Doxil. Celui-ci fut accepté par la FDA en 1995. Il s'agit d'un nanoliposome de la molécule de doxorubicine, un antibiotique antinéoplasique. Depuis 2014, une quarantaine de médicaments vectorisés ont été commercialisés dont une trentaine en France notamment dans le domaine de la cancérologie. Cette innovation biotechnologique est en plein essor. L'encapsulation permet de faire entrer plus facilement les médicaments dans les cellules au travers des barrières naturelles. L'encapsulation permet de protéger le principe actif jusqu'à son arrivée sur la cible.

Il est possible de greffer de petits ligands à l'interface de ces nanomédicaments afin de cibler le principe actif. Des recherches sont en cours pour développer des nanoparticules radio-opaques afin de suivre le traitement par imagerie.

4) *L'essor de nouvelles technologies*

La cristallographie aux rayons X, la RMN, l'ITC ou la SPR sont quatre techniques robustes utilisées depuis de nombreuses années pour déterminer les candidats médicaments intéressants. En parallèle, les laboratoires suivent l'émergence de nouvelles technologies pouvant optimiser la caractérisation et l'optimisation de candidats médicaments.

Parmi les techniques les plus récentes utilisées au cours de ma thèse, nous retrouvons les approches DSF (pour *Differential Scanning Fluorimetry* en anglais) et nano-DSF qui permettent de déterminer la ou les température(s) de dénaturation d'une cible seule ou en complexe en suivant le déploiement d'une protéine en fonction de la température (Lo et al. 2004; Niesen, Berglund, et Vedadi 2007). En DSF, la fluorescence est émise par une sonde fluorescente (le SYPRO Orange dans notre cas d'étude) capable de se lier aux parties hydrophobes exposées de la cible et notamment au cœur hydrophobe qui est exposé lorsque la protéine se dénature. En nano-DSF, on suit la fluorescence intrinsèque des acides aminés aromatiques (marquage non-requis de la cible). Ces deux approches permettent de mettre en évidence l'interaction de ligands avec la protéine. On distingue des molécules stabilisatrices de

la cible lorsque la température de dénaturation (*melting Temperature* en anglais - T_m) est augmentée ou des molécules déstabilisatrices de la cible lorsque le T_m est diminué. Ces deux approches sont robotisées et requièrent de très faibles quantités de matériel (typiquement 10 μ L de solution protéique à 0,5 – 1 mg/mL par mesure).

La MST (pour *MicroScale Thermophoresis* en anglais) mesure un changement de fluorescence induit par un gradient de température créé par un laser infrarouge. Un déplacement du ligand et du récepteur est observé à l'endroit où est appliqué le laser sur une courte durée et une zone limitée du capillaire. Le laser entraîne une augmentation locale de la température de la solution éclairée. Les molécules diffusent vers l'extérieur de cette zone et reviennent lorsque le laser est arrêté. La diffusion du récepteur est liée à plusieurs paramètres dont son hydratation et son rayon hydrodynamique. Dans une première approche, le ligand ou le récepteur sont marqués avec une sonde de fluorescence dont on suit la disparition et le retour avec l'utilisation du laser (Seidel et al. 2013). C'est donc un laser infrarouge qui est utilisé pour détecter le changement de fluorescence (Renaud et al. 2016). Il est possible lorsque la cible possède plusieurs aromatiques tels qu'un tryptophane, d'utiliser la MST dite *label-free*. Dans ce cas, c'est le changement de fluorescence intrinsèque de la cible qui est mesuré (Seidel et al. 2012).

SwitchSENSE est le nom donné à une technologie proposée en 2012 par la compagnie Dynamic Biosensors (DBS) GmbH Martinsried. On trouve actuellement deux types d'appareils électro-optiques : le DRX possédant une seule couleur (diode jaune) et le DRX² possédant deux couleurs (diodes rouge et vert). L'I2BC possède cette deuxième version.

Cette technique permet d'étudier différents types d'interactions moléculaires notamment les interactions protéine-protéine/peptide, protéine-ADN/ARN, protéine-petite molécule (Figure 8). Le switchSENSE est une technologie de microfluidique. Elle permet principalement de mesurer par fluorescence les cinétiques d'association et de dissociation (k_{on} , k_{off} et K_D), la taille et la forme d'une macromolécule ou d'un complexe. On peut mesurer par cette technique des valeurs de K_D allant du micromolaire (10^{-6} M) au picomolaire (10^{-10} M à 10^{-12} M) (Cléry et al. 2017). Les échantillons sont envoyés sur une puce recouverte d'or et composée de quatre canaux. Chacun de ces canaux est constitué de six électrodes. Sur chaque électrode, se trouvent des nanoleviers d'ADN simple brin (ssDNA) liés covalamment à la puce en 5'. En 3' du nanolevier d'ADN, est lié de manière covalente un fluorophore fluorophore de couleur rouge ou verte. Les milliers de molécules d'ADN simple brin déjà greffés sur la puce sont ensuite hybridés par injection d'ADN complémentaires. Cette étape d'hybridation est appelée étape de

fonctionnalisation de la puce, L'ADN complémentaire peut également être hybridé indépendamment avec une protéine covalamment liée en 5'. Nous obtenons un ADN double-brin greffé sur la puce avec ou sans protéine liée qui permet de mesurer ensuite des interactions respectivement de type protéine/analyte ou ADN/analyte.

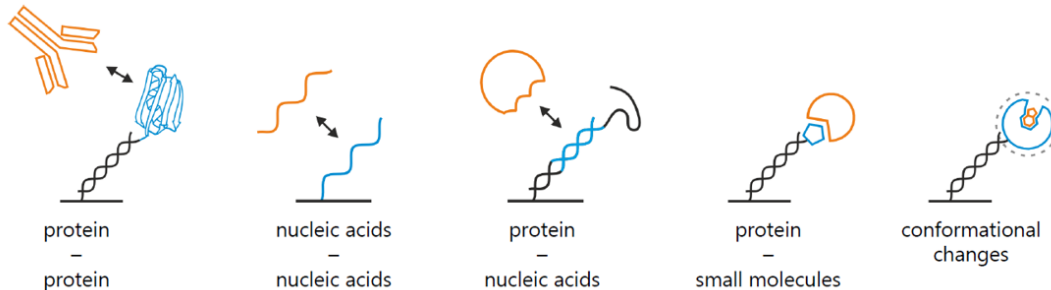


Figure 8 : Représentation schématique des différents types d'interactions pouvant être analysées en switchSENSE (<https://www.dynamic-biosensors.com/>)

Deux modes peuvent être utilisés pour les mesures. Le premier mode est le mode « statique ». La fluorescence est mesurée en fonction de l'environnement chimique présent autour du fluorophore.

Lorsqu'un analyte interagit avec le ligand (sur le nanolevier double brin) à proximité du fluorophore, la fluorescence émise peut être modifiée (diminuée le plus fréquemment, ou augmentée). Dans ce mode, un potentiel négatif est appliqué, L'ADN étant chargé négativement, alors celui-ci se retrouve en position verticale sur la puce, Le deuxième mode est le mode « dynamique ». Il permet, grâce à l'application d'un potentiel électrique alternatif, de mettre en mouvement les oligonucléotides présents sur la puce (Figure 9). Du fait de la charge négative de l'ADN, les nanoleviers vont passer d'une position verticale (lorsque le champ appliqué est de signe négatif ; répulsion) à une position horizontale (lorsque le champ appliqué est de signe positif ; attraction). Lorsque le nanolevier est en position horizontale, alors l'intensité de la fluorescence diminue. Lorsque le nanolevier est en position verticale, alors l'intensité de la fluorescence est augmentée : il y a émission de fluorescence. On mesure avec ce mode dynamique, la « réponse dynamique » correspondant à l'aire sous la courbe de la différence « fluorescence émise lorsque le nanolevier est libre » - « fluorescence émise lorsque l'analyte est lié » (Figure 10).



Figure 9 : Représentation schématique du mode dynamique de switchSENSE (Langer et al. 2013)

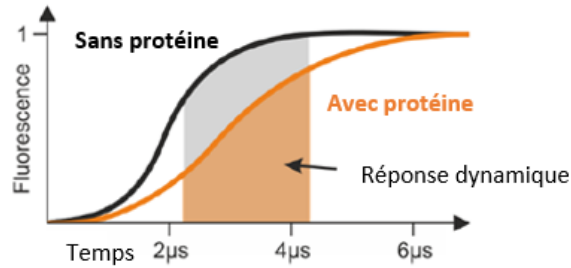


Figure 10 : Graphique représentant la fluorescence émise en fonction du temps et la réponse dynamique en switchSENSE (<https://www.dynamic-biosensors.com/>)

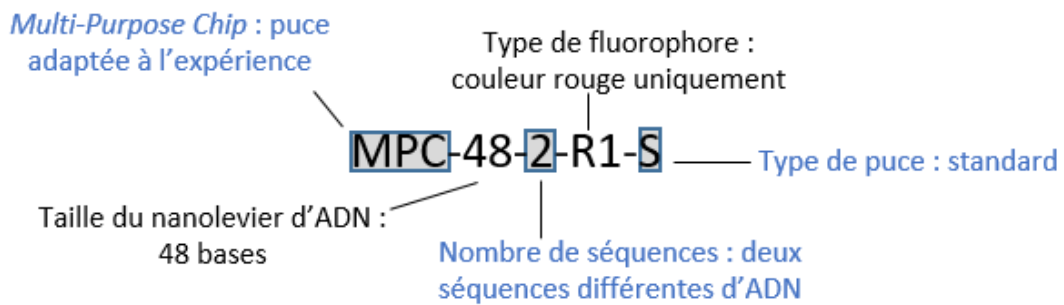


Figure 11 : Exemple de nomenclature de puce de switchSENSE

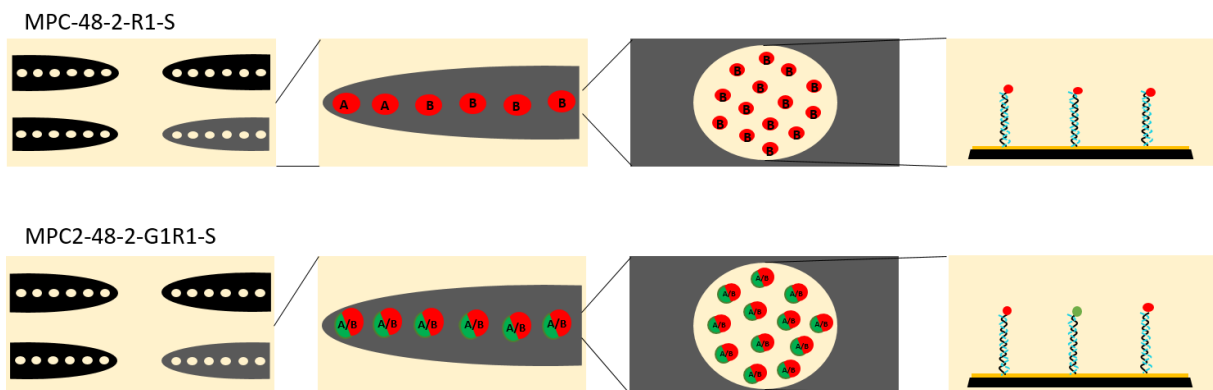


Figure 12 : Représentation schématique de deux exemples de puces de switchSENSE. La première image montre la puce recouverte d'or et ses quatre canaux (1 en bas à gauche, 2 en haut à gauche, 3 en haut à droite, 4 en bas à droite). La seconde image montre la vue d'ensemble d'un canal constitué de six électrodes, Ces électrodes sont composées de nanoleviers constitués des séquences A et/ou B et des fluorophores rouge et/ou vert. La troisième image est un zoom d'une électrode composée d'une multitude de nanoleviers d'ADN. La quatrième image est un zoom décrivant les nanoleviers en position verticale et hybridés par un ADN complémentaire.

Plusieurs types de puces existent. Une nomenclature particulière est utilisée pour les reconnaître (Figure 11). Certaines puces ont un seul type de fluorophore ou deux fluorophores, ou désormais aucun fluorophore - le fluorophore est apporté dans un second temps avec un ADN - (ce qu'on appelle les puces noires ou *black chips* en anglais). Il est de plus possible d'utiliser une puce possédant des électrodes composées de deux séquences d'ADN différentes A et B (dites respectivement de référence et de mesure) ou possédant sur les deux premières électrodes une même séquence d'ADN A (dite de référence) et sur les quatre dernières électrodes une séquence d'ADN B (dite de mesure) (Figure 12). Pour mon projet avec Sanofi, j'ai utilisé une puce de type MPC-48-2-R1-S. Elle possède un fluorophore rouge, les deux premières électrodes servent de référence et les quatre électrodes suivantes servent d'électrodes de mesure.

Le nombre de nanoleviers d'ADN sur une électrode est variable en fonction du type de puce utilisée. La puce peut être de haute densité, d'une densité modérée ou de faible densité. Lorsque l'utilisateur a besoin de séquences oligonucléotidiques très longues, il est préférable d'utiliser des puces de faible densité afin d'éviter le chevauchement des nanoleviers en mode dynamique. Sur une puce standard, la distance moyenne entre deux nanoleviers est de 50 nm.

Pour éviter qu'une protéine interagisse avec la surface de la puce entre deux nanoleviers d'ADN, une solution dite de passivation est indispensable pour combler les espaces vides. Cette solution est utilisée à chaque début d'expérience. Pour le stockage des puces, les nanoleviers simple brin sont hybridés avec des ADN complémentaires afin d'augmenter leur stabilité et celle des fluorophores. Après l'étape de passivation, une étape de régénération est effectuée pour utiliser de nouveau la puce. La régénération permet de passer de l'état hybridé à l'état simple brin pour pouvoir ensuite hybrider l'ADN d'intérêt. Pour ce faire, une solution de régénération à pH basique, permet de rompre l'hybridation du simple brin avec le complémentaire. La régénération de la puce supprime alors tout ligand attaché au nanolevier ou tout analyte potentiellement accroché à ce ligand. Cette régénération s'établit en deux étapes successives : tout d'abord une étape de dénaturation des nanoleviers d'ADN pour obtenir une forme simple brin et laisser place ensuite à l'hybridation des nanoleviers double brin.

C'est ce qu'on appelle alors la fonctionnalisation de la puce. Il est nécessaire d'effectuer cette fonctionnalisation pour vérifier avant toute expérience, que l'hybridation de l'ADN complémentaire (cNLA48 ou cNLB48 en fonction de la séquence) soit correcte. En fonction de l'état simple ou double brin de l'ADN présent sur la puce, la fluorescence est différente car ces

deux formes présentent des propriétés de commutation (switch) différentes. De par sa flexibilité, l'ADN simple brin ne présente qu'une faible amplitude de fluorescence en mode dynamique (Figure 13).

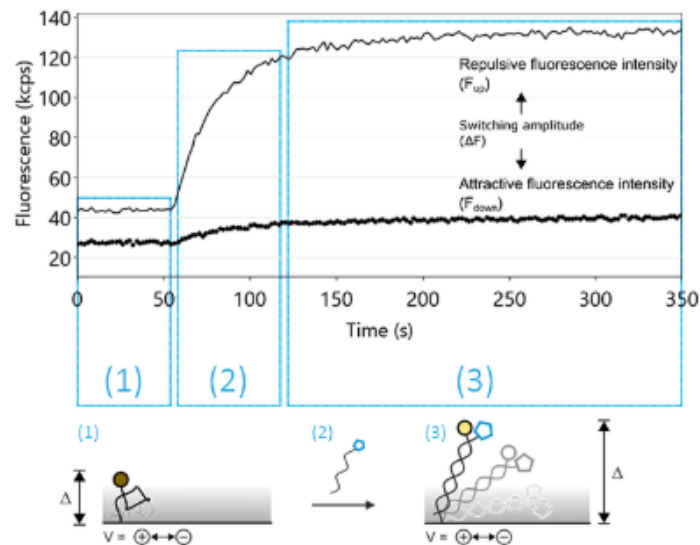


Figure 13 : Profil de switchSENSE décrivant en (1) l'état simple brin de l'ADN sur la puce après étape de régénération, en (2) le début de fonctionnalisation de la puce avec état d'hybridation de l'ADN simple brin en double brin (début de la séparation entre F_{down} – ligne épaisse – et F_{up} – ligne fine – respectivement la fluorescence correspondant à l'attraction de l'ADN et à sa répulsion), en (3) la fonctionnalisation complète de la puce (<https://www.dynamic-biosensors.com/>)

On parle ensuite de fonctionnalisation de la puce. Cette étape correspond à l'état double brin du nanolevier avec hybridation de l'ADN complémentaire à la séquence (A ou B) (Figure 14). Quand la fonctionnalisation de la puce est correcte, alors toute mesure autre peut être démarrée (cinétique, titration, mesure de taille : *sizing* ...).

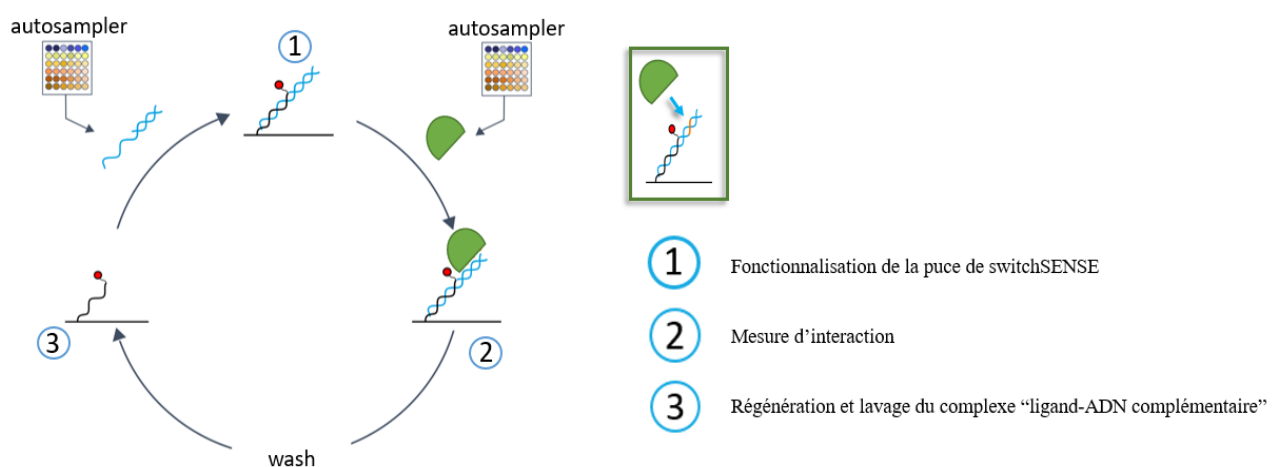


Figure 14 : Fonctionnalisation de la puce de switchSENSE (1), mesure d'interaction de type protéine-ADN après injection de la protéine d'intérêt (2) et régénération durant laquelle le brin complémentaire et la protéine fixée sont dissociés (3)

Caractéristiques de l'appareil (Figure 15) :

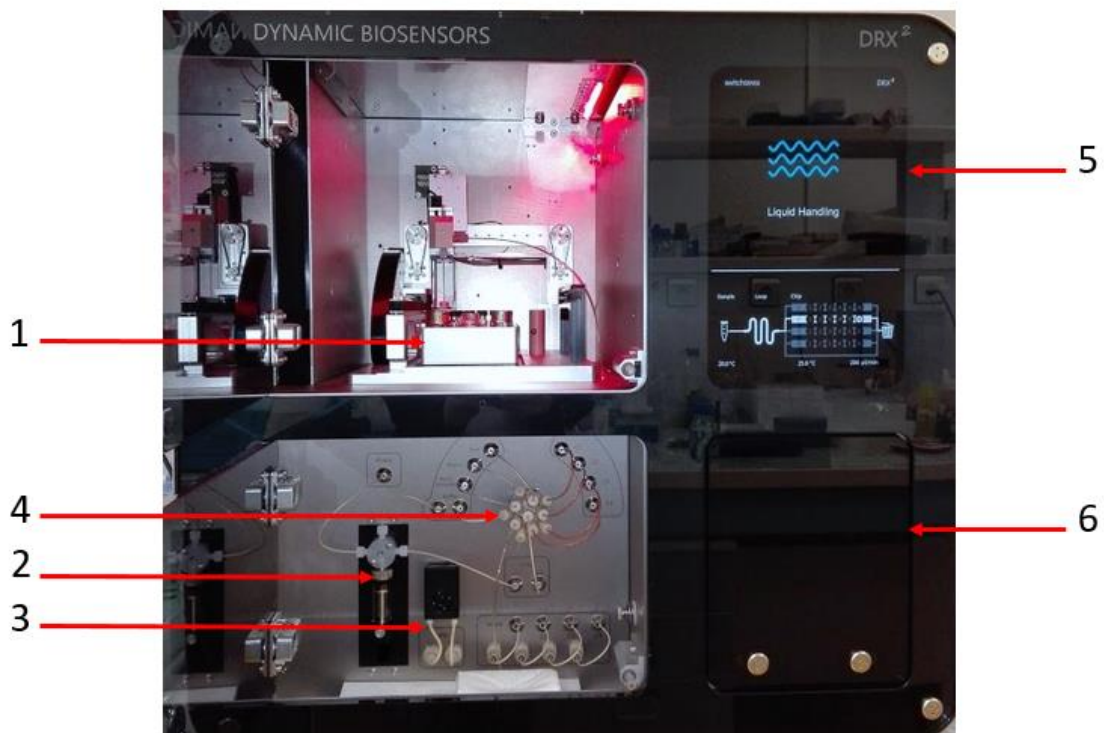


Figure 15 : Photo du DRX² prise à l'I2BC, CEA Saclay

Automate d'échantillons (1) :

- Nombre d'échantillons allant de 48 à 96 (dans ce cas, plaque de 96 puits)
- Contrôle de température : de 8 °C à 40 °C
- Disposition des échantillons à tester et des solutions « eau », « passivation », « régénération »

Seringue intégrée de microfluidique (2) :

- Volume de 5 mL
- Aspiration des échantillons
- Envoi des échantillons vers la puce
- Débit allant jusqu'à 4000 $\mu\text{L}/\text{min}$

Pompe péristaltique (3) :

- Utilisation lorsque l'analyte se dissocie de manière lente
- Volume limité au $V_{\text{RunningBuffer}}$
- Débit allant jusqu'à 4000 $\mu\text{L}/\text{min}$

Valve d'injection (4) :

- Servant d'intermédiaire entre l'automate d'échantillons et la seringue/puce

Ecran de lecture d'électro-optique en temps réel (5) :

- Comptage de photons uniques corrélés dans le temps
- Temps d'opération allant jusqu'à 72 h

Compartiment de la puce (6) :

- Température de mesure de la puce allant de 8 °C à 75 °C
- Système de tiroir automatisé

Premier objectif

Un de mes objectifs de thèse était d'étudier par switchSENSE, différentes interactions de type protéine-ADN pour un projet industriel. En partenariat avec Sanofi, nous voulions savoir si cette technique innovante de biophysique pouvait être appliquée dans le cadre de la recherche de candidats médicaments. J'ai travaillé sur un projet d'oncologie. Quatre constructions protéiques, deux ADN m'ont été confiés afin de réaliser des mesures d'interaction protéine-ADN par switchSENSE. Dans un deuxième temps, je devais tester plusieurs fragments sur les interactions protéine-ADN pour déterminer leur effet d'inhibition. Ces molécules considérées comme inhibitrices avaient été identifiées auparavant chez Sanofi par différentes méthodes de biophysique telles que la SPR, la DSF et la MST. Le but était de pouvoir évaluer les valeurs ajoutées du switchSENSE et de savoir si celle-ci pouvait être utilisée pour du criblage en recherche de médicaments.

PARTIE II :

Etude d'une interaction protéine-ADN, Ku70/80-ADN, au cœur de la voie NHEJ

Plusieurs dates ont marqué l'histoire de la biologie notamment la découverte de la structure en double hélice de l'ADN (Acide DésoxyriboNucléique) en 1953. A l'époque, sa composition chimique est déjà connue mais son assemblage reste encore source d'interrogations. On sait que la molécule d'ADN est composée d'un enchaînement « base azotée – désoxyribose – phosphate » mais l'on essaie de comprendre comment sont assemblées dans l'espace ces longues chaînes. L'ADN est décrite alors par rayons X comme un acide nucléique possédant premièrement trois chaînes distinctes s'entrelaçant autour d'un axe vers lequel s'orientent les phosphates (Pauling et Corey 1953). Mais ces hypothèses sont très vite remises en cause par deux personnages emblématiques que sont le généticien biochimiste Watson et le biophysicien Crick. En 1962, ils obtiennent le prix Nobel de médecine pour la découverte de la structure de l'ADN. Celle-ci est alors décrite comme une double hélice bicaténaire formée par un enchaînement de quatre nucléotides : Cytosine, Guanine, Adénine et Thymine (Watson et Crick 1953). Ces deux jeunes scientifiques bouleversent alors à l'époque le monde de la Biologie et de la génétique. On connaissait désormais la structure de la molécule à l'origine de l'hérédité. Les nucléotides sont décrits comme des bases azotées liées par une fonction 1'-N-glycosidique à un désoxyribose qui porte un phosphate en position 5'. Qui dit hérédité, dit transmission de génération en génération. Pour qu'un organisme soit viable, alors l'information génétique doit être transmise de manière intégrale et correcte durant la descendance. C'est l'enchaînement des nucléotides qui définit l'information génomique et si cet enchaînement est perturbé, alors l'information est perdue ou biaisée. L'ADN étant un polymère, elle présente de nombreuses liaisons très facilement clivables pouvant nuire à l'intégrité génomique. Lors de la synthèse de la molécule d'ADN, des milliers de nucléotides sont liés de manière covalente. Pour chaque complexe sucre-base et sucre-phosphate, la liaison est clivable. Parfois, des mutations ou aberrations chromosomiques apparaissent suite au maintien dans le génome de certaines lésions ou dommages de l'ADN. Dans ce cas, ces aberrations sont susceptibles d'être maintenues dans le génome et peuvent être transmises aux générations futures. Très souvent, aucune conséquence n'est retrouvée au niveau du phénotype ; on dit que ces mutations sont silencieuses. Elles ne causent donc aucun trouble dans le métabolisme. Mais parfois, ces

dommages causent certains dysfonctionnements dans l'organisme, comme l'apoptose des cellules, l'apparition de tumeurs ou de certains cancers. Une surveillance étroite du maintien de l'intégrité génomique au niveau de l'ADN est mise en place pour limiter ces dommages dans les organismes vivants (Hoeijmakers 2001). Plusieurs types d'agents causent des lésions de l'ADN : agents exogènes (rayonnements, agents environnementaux de natures chimie ou physique respectivement les radicaux libres de l'oxygène ou les radiations ultraviolettes) et agents endogènes (métabolisme énergétique). Les liaisons covalentes sont des cibles des agents génotoxiques. Dans ce manuscrit, on s'intéresse principalement aux dommages appelés cassures double-brin de l'ADN (CDB). Cette partie de l'introduction s'intéressera dans un premier temps à la formation des CDB et à la réparation de ces CDB. Nous nous intéresserons ensuite au rôle essentiel d'une protéine hétéromérique Ku70/80 dans une des voies de réparation de l'ADN chez les mammifères (voie NHEJ).

A) Formation des CDB et réparation de ces cassures

1) Les origines diverses des CDB

Les cassures double-brin sont les cassures les plus toxiques de l'ADN. En effet, elles ne permettent pas la poursuite de la vie cellulaire lors des étapes de mitose ou de réplication. Pour éviter l'apoptose de la cellule, ce dommage doit être réparé (Roos et Kaina 2006). Parfois, un dysfonctionnement au niveau cellulaire entraîne une tumorigénisation de la cellule ; celle-ci n'est donc pas conduite en apoptose et c'est ce genre de lésion qui est dangereuse pour le génome.

Trois grandes sortes de CDB sont retrouvées. Elles diffèrent selon leur origine. Le point commun entre ces cassures est qu'elles proviennent de deux cassures simple-brin. La liaison phosphodiester est alors rompue entre deux nucléotides présents sur les deux bras d'un chromosome. La rupture se fait dans un espace maximum de 250 paires de bases *in vitro*. De nombreux facteurs peuvent induire ces cassures notamment les facteurs exogènes tels que les radiations ionisantes ou encore les facteurs endogènes tels que les espèces réactives de l'oxygène (ROS) ou encore certaines enzymes ayant un effet sur l'ADN (ligases, topoisomérases, nucléases, polymérases ...). Les cassures de type exogène représentent seulement une petite partie des lésions. Elles représentent environ 10 à 50 CDB par cellule et

par jour (Haber 1999) sur un total de 75 000 lésions de l'ADN par cellule et par jour (Lindahl et Nyberg 1972).

Ces différentes formes de CDB sont données sur le schéma suivant (Figure 16). Nous ne rentrerons pas dans les détails des différentes formes de CDB.

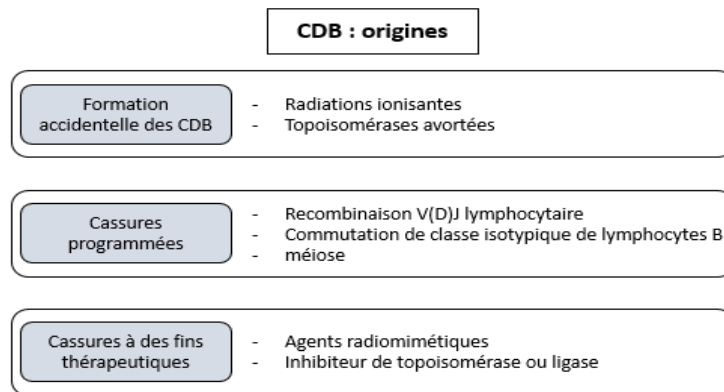


Figure 16 : Origines des différentes cassures double brin de l'ADN

2) Formation des CDB lors d'un traitement thérapeutique

Les CDB de l'ADN sont très souvent létales pour les cellules. En effet, si ces cassures ne sont pas prises en charge par les voies de réparation, alors ces cellules sont conduites directement à l'apoptose. Or, parmi les nombreux traitements anti-cancer qui existent à l'heure actuelle, certains ont pour but de provoquer des CDB pour tuer les cellules tumorales d'un patient. Les cellules cancéreuses, de par leur métabolisme, sont davantage sensibles à de nombreux stress notamment aux stress génotoxiques (Luo, Solimini, et Elledge 2009). Le but ici est de créer un maximum de CDB pour saturer le système de réparation des cellules et les conduire directement en apoptose. Bien sûr, ce genre de thérapie n'est pas sans effet secondaire car il va toucher aussi des cellules saines. On cherche à trouver de nouveaux traitements tout aussi efficaces tout en étant moins agressifs pour l'organisme.

Un des premiers traitements utilisés est l'utilisation des rayonnements radio-ionisants et plus particulièrement les rayons X. L'objectif est d'irradier de manière relativement précise la zone contenant les cellules tumorales. Un second moyen est l'utilisation de molécules dont l'action sera d'imiter l'effet des rayonnements ionisants (c'est ce qu'on appelle les drogues radiomimétiques). Ceux-ci sont isolés de certaines plantes et induisent des CDB de l'ADN en générant des radicaux libres des sucres. Quelques agents radiomimétiques sont connus tels que

la néocarzinostatine, la calicheamicine, la C-1027 ou encore la bléomycine (Povirk 1996). Un sucre de l'ADN est attaqué par le composé-métal à travers une attaque radicalaire (Figure 17A). Un carbone est alors déshydrogéné et il est ensuite oxydé par l'oxygène présent dans le moyau. Cette nouvelle forme oxydée et très réactive induit la rupture de la liaison N-glycosidique qui relie la base et le sucre. Une seconde attaque est alors effectuée sur le brin complémentaire de l'ADN (Schultz et Dervan, 1983 ; Povirk 1996). Ce type de drogue radiomimétique a une spécificité pour certaines séquences d'ADN et *in vitro*, ils sont plus efficaces que les rayonnements ionisants X (Figure 17). Une dose de 1 nM de calicheamicine est plus destructrice qu'une dose de 6 Gy (Elmroth et al. 2003).

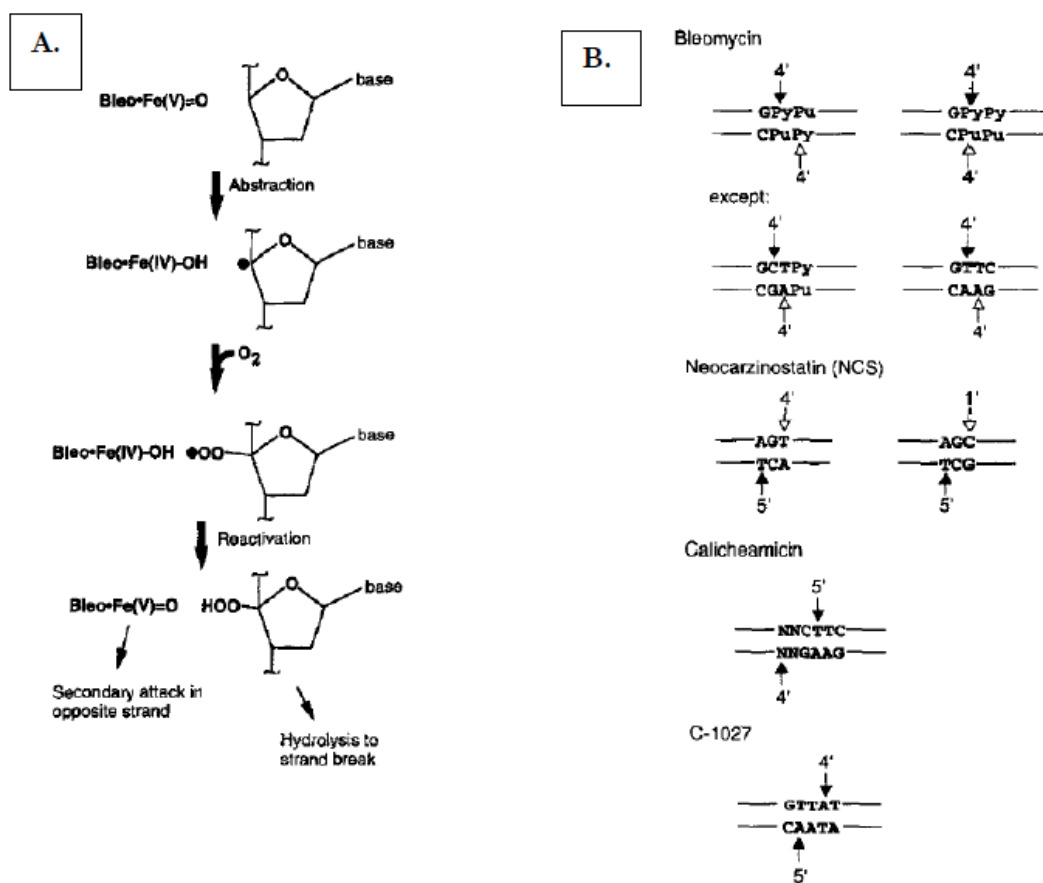


Figure 17 : A. Mécanisme d'action de la bléomycine. Un ion métallique est chélaté et le complexe réagit avec une molécule d'oxygène (O₂) pour la modifier en ion superoxyde (radical). Cet ion induit des CDB de l'ADN. B. Préférence des drogues radiomimétiques pour certaines séquences d'ADN. L'action de ces drogues se réalise à chaque fois sur les deux brins de l'ADN

Les inhibiteurs de ligases ou de topoisomérases peuvent également former des CDB de l'ADN. Ces deux enzymes forment lors de leur cycle catalytique des intermédiaires covalents avec

l'ADN appelés également adduits enzyme-ADN. C'est une troisième source de CDB décrits ci-dessous.

Exemples d'inhibiteurs de ligases

Les ligases sont des enzymes capables d'effectuer l'étape de ligation notamment lors de la réparation de l'ADN en fin de processus. Les ligases sont de bonnes cibles thérapeutiques pour développer de nouvelles molécules capables d'inhiber certaines pathologies notamment les cancers. Des études sur ces ligases ont permis de schématiser leur cycle catalytique (Figure 18) afin de pouvoir élucider de nouvelles petites molécules susceptibles de stopper la réaction de réparation de l'ADN.

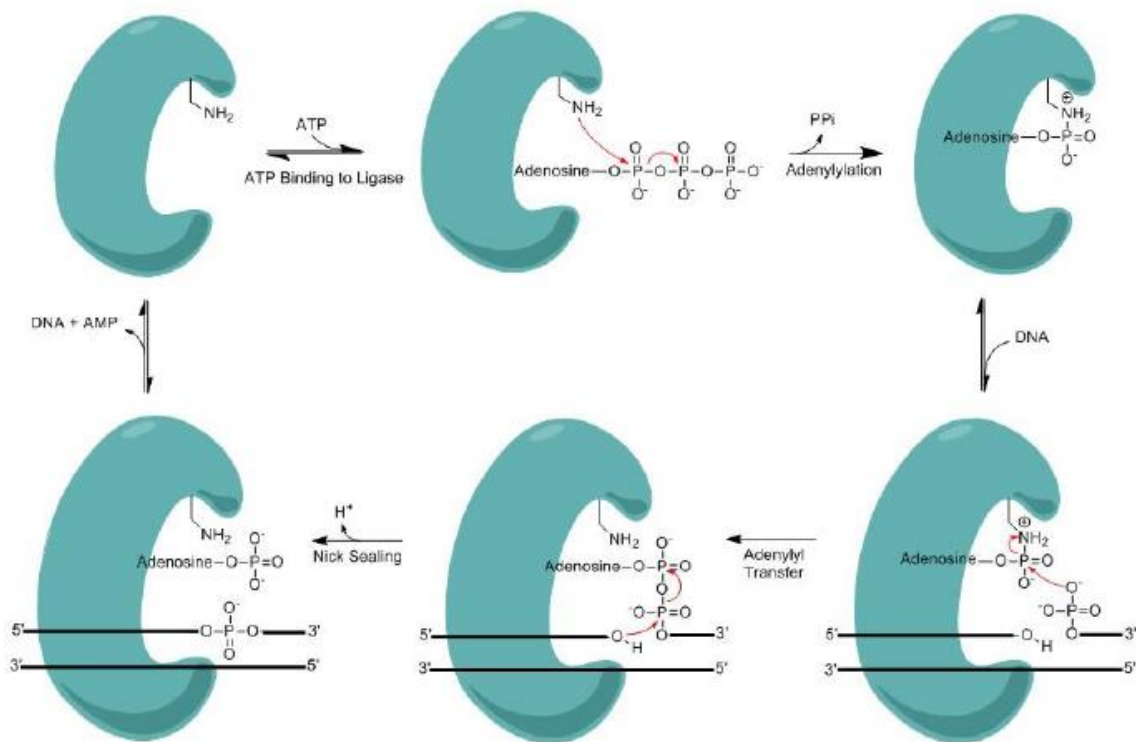


Figure 18 : Cycle catalytique des ligases humaines

Un complexe AMP-ligase covalent est formé lors de la première étape du cycle catalytique des ligases, générant la forme active de la ligase. Grâce à une attaque nucléophile de l'extrémité 5'-phosphate du complexe AMP-ligase, une adénylation est formée sur l'extrémité 5'. Une dernière étape consiste en une nouvelle attaque nucléophile sur l'extrémité 5'-adénylée par l'extrémité 3'-OH. Une liaison phosphodiester est donc créée avec relargage de la molécule d'AMP (Figure 18). Pour obtenir une ligase de nouveau active et efficace, celle-ci doit être de

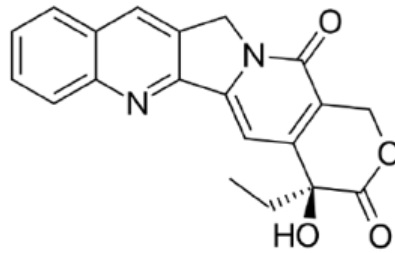
nouveau adénylée à la fin du cycle catalytique. L'étape préférentielle à bloquer lors de ce cycle pour générer des inhibiteurs est l'étape d'adénylation. Pour démarrer cette recherche, la structure de la Ligase I humaine en complexe avec l'AMP a été utilisée comme départ (Pascal et al. 2004). En 2008, des approches *in silico* ont été réalisées depuis l'intermédiaire réactionnel (Chen et al. 2008). Les études réalisées ont été faites à partir de l'homologie des trois Ligases I, II et III permettant de ce fait de pouvoir tester une dizaine de molécules potentiellement capables de diminuer ou d'arrêter leur activité. Suite à cette étape, différents inhibiteurs de la ligase ont été découverts *in vitro* (Chen et al. 2008). Un autre inhibiteur SCR7 a été proposé contre la Ligase IV (Srivastava et al. 2012) et plus récemment, c'est un dérivé de synthèse du SCR7 qui a été proposé comme produit réellement à l'origine des propriétés inhibitrices attribuées au départ à la molécule SCR7 (Greco et al. 2016).

Exemples d'inhibiteurs de topoisomérases

Les topoisomérases sont des enzymes capables de couper l'ADN simple ou double brin. Elles sont essentielles lors de plusieurs étapes de la vie cellulaire notamment la réplication, la transcription mais aussi lors de la séparation des chromosomes, Le mécanisme d'action de ces enzymes n'est pas détaillé ici mais il est bien connu depuis les années 90 (Roca. 1995). Plusieurs agents chimiques capables de bloquer la fonction de religation des extrémités de l'ADN ont été découverts et leurs structures ont été caractérisées. Les exemples que nous pouvons citer sont premièrement l'étoposide venant former un complexe ternaire stable avec l'ADN clivé et la protéine TOP2 (code PDB : 3QX3). C'est un inhibiteur des topoisomérases II (Maanen et al. 1988; Hande 1998; Gordaliza et al. 2004). L'inhibiteur de topoisomérase sert ainsi à bloquer la réplication, la transcription et conduit de ce fait à la mort de la lignée cellulaire. Le deuxième exemple bloque le cycle catalytique de la topoisomérase I : c'est une molécule de camptothécine (Figure 19A). Celle-ci a été découverte suite à un criblage de petites molécules ayant des propriétés anticancéreuses (Figure 19A) (Wall et al. 1966). La camptothécine est capable de lier très fortement le complexe Top1-ADN grâce à un réseau de liaisons hydrogènes avec la protéine et à des interactions π - π avec l'ADN (Figure 19B) (Redinbo 1998). Des CDB sont produites par les dérivés de la molécule de camptothécine.

A l'heure actuelle, beaucoup de travaux se concentrent autour de la recherche de nouveaux inhibiteurs pour des ligases et des topoisomérases.

A.



B.

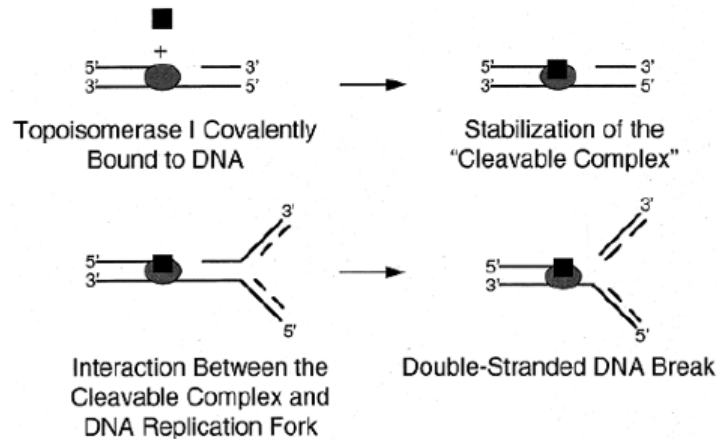


Figure 19 : **A.** Structure chimique de la camptothécine. Pour former le complexe covalent Top1-ADN, les fonctions cétones de la molécule interagissent avec la protéine et les autres fonctions de la molécule forment des liaisons π - π avec l'ADN. Le complexe est stable et indissociable. **B.** Mécanisme d'action de la camptothécine lors du cycle de la topoisomérase I et formation d'une CDB

3) La voie c-NHEJ chez les mammifères

Le génome est sans cesse soumis à de nombreux stress génotoxiques. Ces stress conduisent à la formation de CBD provoquant l'apoptose des cellules ou à la formation de tumeurs si celles-ci ne sont pas prises en charges ou mal prises en charge. Pour pallier à cette situation, divers mécanismes de réparation de l'ADN sont mis en place par l'organisme. Selon la phase du cycle cellulaire, une voie différente est activée. En effet, lorsque que la cellule est en phase S ou G2, un chromosome homologue est présent et peut servir de modèle pour la réparation de l'ADN clivé. Dans cette phase du cycle, la voie de recombinaison homologue (HR pour *Homologous Recombination* en anglais) peut être empruntée (Shrivastav, De Haro, et Nickoloff 2008). Pour toutes les phases du cycle cellulaire notamment en phase G0 ou G1, la réparation se fait chez les vertébrés principalement par la voie NHEJ classique car les chromosomes homologues ne sont pas présents.

La voie NHEJ permet la ligation directe des extrémités de l'ADN sans avoir besoin de chromosome homologue. Deux voies de réparation NHEJ sont retrouvées : la voie classique dite c-NHEJ et la voie alternative dite alt-NHEJ ou MMEJ (pour *Microhomologous Mediated End Joining* en anglais). Mon projet académique s'intéresse à la voie c-NHEJ (Figure 22). La voie majoritaire retrouvée chez la levure est la voie de recombinaison homologue (HR). Les trois étapes décrivant la voie c-NHEJ sont détaillées ci-dessous (Figure 22).

Reconnaissance des extrémités d'une cassure double-brin

La CDB est rapidement reconnue par un premier complexe protéique, l'hétérodimère Ku70/80 (notée Ku). La reconnaissance des extrémités de l'ADN se fait en moins de 5 min (Reynolds et al. 2012). La DNA-PKcs qui est une apo-enzyme et faisant partie de la famille des kinases PI3K, est alors recrutée suite à la fixation des hétérodimères sur les extrémités de la CDB. Le complexe DNA-PK est alors formé, constitué de la protéine Ku et de la DNA-PKcs. Ensemble, ils permettent de stabiliser les deux extrémités de la CDB à proximité pour former ce qu'on appelle la synapse du NHEJ. Le rapprochement des extrémités des CDB, limite l'action des nucléases et le phénomène de résection des extrémités (Dyan et Yoo 1998; Weterings et van Gent 2004). Le complexe DNA-PK possède une activité d'autophosphorylation qui agit en trans cad sur une autre molécule de DNA-PKcs. Cette activité est indispensable pour que la voie NHEJ fonctionne correctement (Chan 2002).

Maturation des extrémités de la CDB

L'étape suivante est de rendre chimiquement compatibles les extrémités de l'ADN en vue de la ligation (les extrémités doivent être alors sous la forme 3'-OH et 5'-P) La maturation des extrémités est un processus qui va mettre en jeu un plus ou moins grand nombre d'activités enzymatiques selon la nature chimique des extrémités. Si la CDB présente des bouts francs et des extrémités 3'-OH et 5'-P, alors cette étape n'est pas nécessaire. La ligation peut alors avoir lieu directement. Si des phosphoglycolates (induits par des radiations ionisantes par exemple), des adduits peptide-ADN (induits par une interruption de la topoisomérase) ou d'autres types d'extrémités non pris en charge par la Ligase4 sont présents (Figure 22), alors l'étape de maturation des extrémités est nécessaire et des enzymes de maturation sont recrutées pour obtenir des extrémités de type 3'-OH et 5'-P. On parle ici de nucléases telles que APLF ou Artemis (famille des nucléases), des polymérases de la famille X, ATM (famille des kinases), PPP6R1 (famille des phosphatases). C'est ensemble que ces différentes enzymes travaillent pour mener à bien leur mission de transformer les extrémités pour qu'elles soient religables. Parfois, il arrive que des nucléotides soient clivés au niveau des extrémités. Des polymérases

qui sont dites peu fidèles et appartenant à la famille des PolX permettent alors la synthèse de nucléotides manquants. La maturation des extrémités peut s'accompagner de l'introduction de nouvelles séquences à la jonction des CDB en fin de voie NHEJ. La diversité des extrémités de la CDB est une source de variabilité au niveau des séquences d'ADN avant la CDB et après la réparation (Bétermier, Bertrand, et Lopez 2014).

Ligation des extrémités

Après avoir obtenu des extrémités de l'ADN de type 3'-OH et 5'-P, une troisième étape consiste à réaliser la ligation de celles-ci. Pour ce faire, un complexe est recruté. Celui-ci est composé de trois facteurs : la ligase IV qui vient établir la ligation des extrémités, la protéine XRCC4 (pour *X-Ray sensitivity Cross Complementing factor 4* en anglais) qui vient stabiliser la ligase 4 et enfin la protéine XLF (pour *XRCC4 Like Factor* en anglais) qui vient en tant que co-facteur du complexe Lig4-XRCC4.

Il a été montré que *in vitro*, les protéines XLF et XRCC4 forment des filaments (Figure 20) (Ropars et al. 2011). Plusieurs années plus tard, des expériences de microscopie à super-résolution ont confirmé l'existence de ces filaments *in cellulo* (Figure 21) (Reid et al. 2015). L'activité de la ligase 4 est de plus stimulée par la protéine XLF (Ahnesorg, Smith, et Jackson 2006; Riballo et al. 2009). Enfin, un autre facteur est découvert par une étude par homologie de séquence (Ochi, Wu, et Blundell 2014) et par un crible d'interactants de la DNA-PKcs par spectrométrie de masse (Craxton et al. 2015), il s'agit du facteur PAXX (pour *PARalogue of XRCC4 and XLF* en anglais). Cette protéine interagit avec le premier partenaire du NHEJ Ku70/80 de par la sous-unité Ku70 de l'hétérodimère (Tadi et al. 2016).

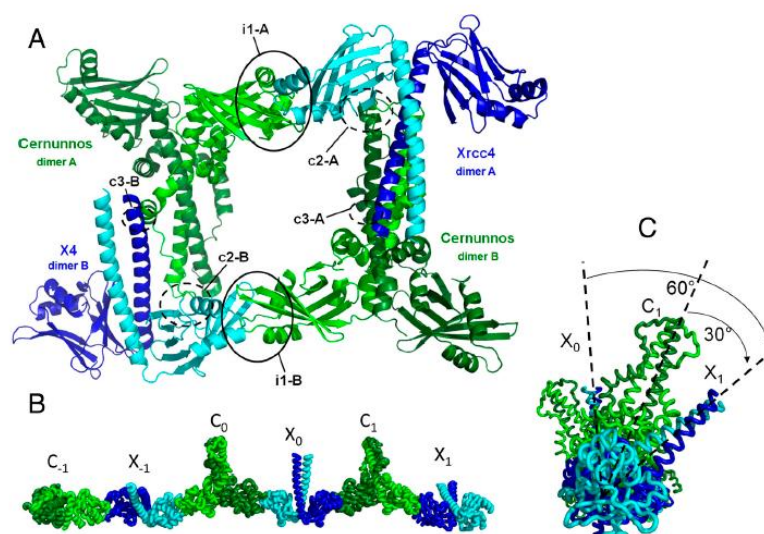


Figure 20 : Structure cristalline du complexe XRCC4/XLF. **A.** L'unité asymétrique contient deux complexes comprenant un homodimère XRCC4 (x4) (bleu foncé et bleu clair) et un homodimère XLF (vert foncé et vert clair). **B.** Structure cristalline en filaments du complexe XRCC4/XLF résolue par V. Ropars au laboratoire. Interaction à travers leur domaine globulaire N-terminal (Ropars et al. 2011)

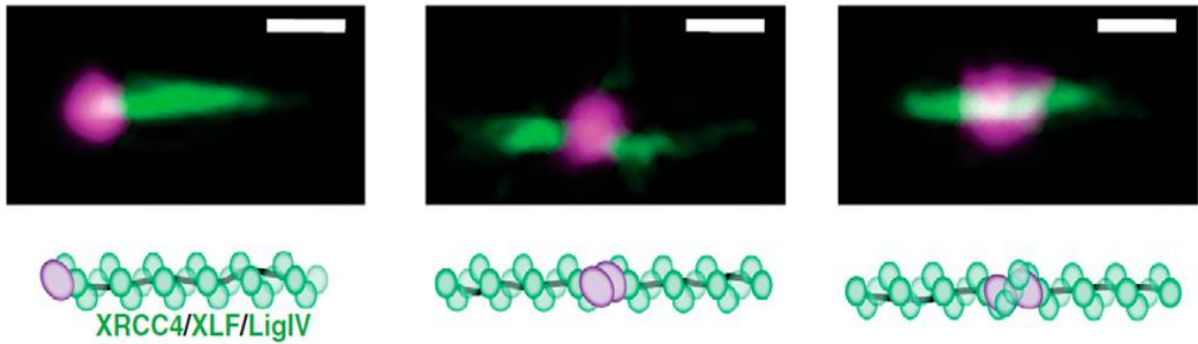


Figure 21 : Résultats de microscopie à super-résolution des filaments de Ligase4-XRCC4-XLF avec chaque schéma correspondant au-dessous. A gauche : forme chenille, au centre : forme papillon et à droite : forme recouvrement. La protéine Ku est en violet sur les extrémités de l'ADN et les protéines XRCC4, XLF et Ligase 4 en vert (Reid et al. 2015)

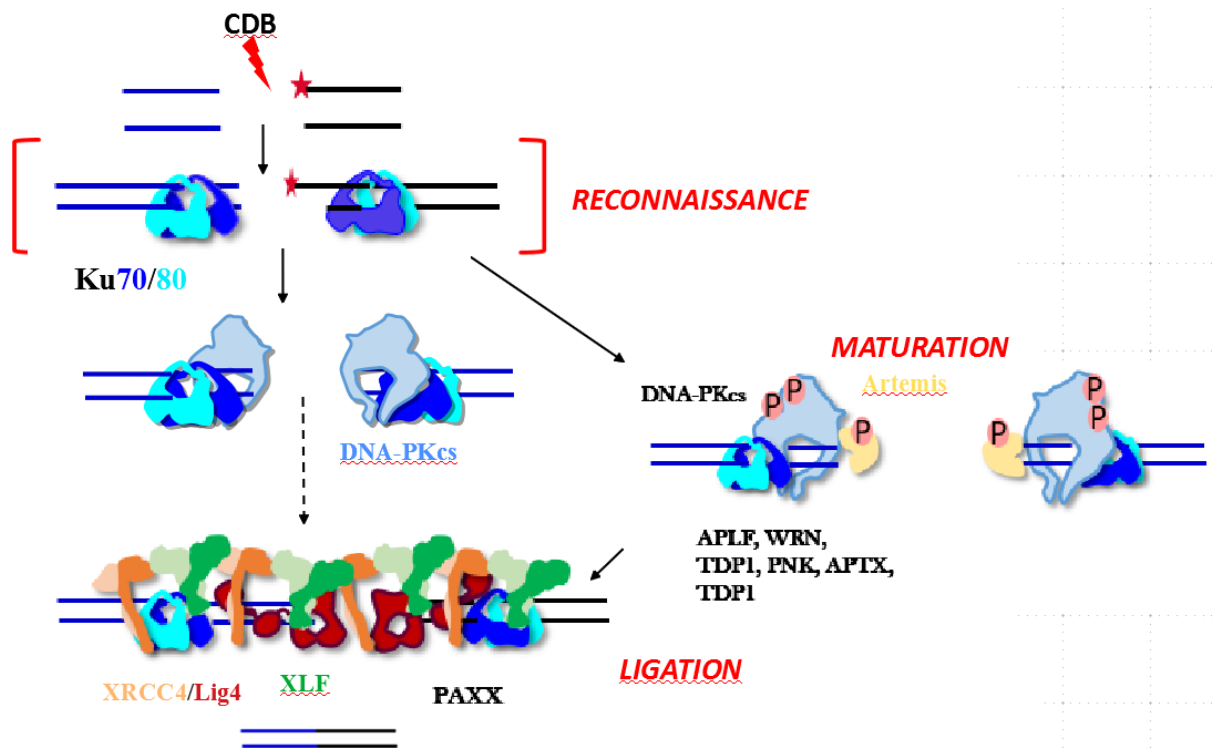


Figure 22 : Schéma du mécanisme général de la voie c-NHEJ chez les mammifères. Cette voie met en jeu une première étape de reconnaissance par le complexe DNA-PK, puis étape de maturation avec recrutement des autres facteurs de maturation pour enfin arriver à l'étape de ligation des extrémités de l'ADN pour la reformation d'une double hélice d'ADN continue (adapté de Chang et al. 2017)

Parmi les différents partenaires du c-NHEJ (abrégié ensuite par NHEJ), un facteur nous intéresse particulièrement dans cette étude ; il s'agit de la protéine Ku qui est considérée comme un facteur central pour cette voie de réparation.

B) Ku70/80 : une protéine indispensable à la voie de réparation c-NHEJ

La protéine Ku a été découverte il y a une quarantaine d'années suite à des études de patients japonais qui étaient atteints de maladies auto-immunes. A l'époque, elle est considérée comme une cible majeure d'auto-antigènes dans les syndromes de sclérodermose. Peu de temps après, on se rend compte qu'elle est impliquée également dans d'autres maladies notamment dans le lupus érythromateux et que cet hétérodimère est principalement retrouvé dans le noyau des cellules (Mimori et al. 1981). De plus, son affinité très forte pour l'ADN est très vite observée et mesurée (K_D d'environ 5 nM) (Blier et al. 1993). Enfin les deux sous-unités protéiques Ku70 et Ku80 permettent de compléter des mutations de deux gènes appelés respectivement XRCC6 et XRCC5. Les cellules déficientes en Ku70 ou Ku80 sont fortement radiosensibles.

1) Structure de l'hétérodimère Ku70/80

La structure cristalline de Ku a été décrite en 2001 par l'équipe de Walker comme un hétérodimère formé de deux sous-unités : Ku70 constituée de 70 kDa et Ku80 constituée de 86 kDa (Walker, Corpina, et Goldberg 2001). Ces deux monomères possèdent des homologies structurales que l'on peut décrire ci-après (Figure 23) :

- Un domaine N-terminal appelé von Willebrand A (vWA) constitué de feuillets β et d'hélices α ,
- Un domaine central constitué d'hélices α , de feuillets β et d'un tonneau β formé par un grand feuillet β enroulé sur lui-même formant un cylindre. Cette structure en trois parties permet la reconnaissance de l'ADN par l'anneau préformé et permet également l'hétérodimérisation de la protéine,
- Un domaine C-terminal constitué par des hélices α différencié pour les deux sous-unités. Ku70 possède un domaine nommé SAP pour SAF-A/B, Acinus and PIAS. Ku80 possède un domaine CTD permettant la reconnaissance de la DNA-PKcs

La structure a été résolue pour Ku seule ou en complexe avec l'ADN respectivement à 2.7 Å (code PDB : 1JEQ) et 2.5 Å (code PDB : 1JEY) (Walker, Corpina, et Goldberg 2001). Ils montrent alors que Ku interagit avec l'ADN de manière séquence indépendante et de par l'anneau préformé par les deux sous-unités. La protéine interagit avec l'ADN sur 14 pb. Cette

forte affinité entre Ku et l'ADN et à de nombreuses interactions électrostatiques. L'ADN, chargé négativement, interagit avec deux surfaces chargées positivement de chaque sous-unité de Ku. Des feuillets β et des hélices α forment la surface positive avec exposition de résidus basiques tels que des lysines qui interagissent avec les groupements phosphate de l'ADN. Dans la structure Ku70/Ku80/ADN (code pdb 1JEY, Walker 2001), c'est un ADN tige-boucle (hairpin) qui a été utilisé avec une partie bicaténaire de 14 pb. La protéine Ku est orientée de façon à avoir la sous-unité Ku80 orientée vers la boucle et la sous-unité Ku70 orientée vers l'extrémité accessible. Ce type d'ADN a permis d'orienter l'hétérodimère. La structure cristalline et des expériences d'empreintes (*footprinting*) ont permis de proposer une orientation de Ku sur la cassure. Ku80 est décrit vers l'intérieur de la cassure et Ku70 vers l'extérieur de la cassure.

Les domaines N-terminaux vWA présentent des repliements de « Rossman fold ». Un feuillet central de 5 brins β est entouré de 6 hélices α . Cette forme de repliement est retrouvée notamment chez des protéines extracellulaires de cellules sanguines, et chez de nombreuses protéines intracellulaires. Ce domaine vWA a un rôle proposé d'interaction avec certains partenaires du NHEJ ; notamment avec la protéine APLF (pour *Aprataxin PNKP Like Factor* en anglais), la protéine XLF et la protéine WRN et plus récemment le facteur MRI/Cyren (Shirodkar et al. 2013; Grundy et al. 2014).

Si l'on compare les deux sous-unités Ku70 et Ku80 (Figure 23), on observe la plus forte divergence au niveau de leurs domaines C-terminaux. Tout d'abord, le domaine C-terminal de Ku70 possède un domaine SAP de 5 kDa et replié sous une structure de type hélice-boucle-hélice. Celui-ci se trouve lié au restant de la protéine par une longue charnière flexible de 50 résidus. Ce domaine est observé dans la structure cristalline 1JEQ de la protéine Ku cristallisée sans ADN et contenant la région C-terminale de Ku70.

Le domaine C-terminal de Ku80, est un domaine de plus grande taille (15 kDa), constitué d'une partie flexible non-structurée (résidus 571 à 595), d'une région ordonnée en hélice α (résidus 595 à 684). Cette partie C-terminale a été résolue par RMN (code PDB : 1RW2) (Zhang et al. 2004). La région C-terminale de Ku80 est impliquée dans l'interaction avec la DNA-PKcs par ses 12 derniers résidus (Dyran et Yoo 1998).

Ces régions C-terminales sont en fait les plus divergentes entre Ku70 et Ku80 et sont des cibles pour des modifications post-traductionnelles qui donnent lieu à des régulations du recrutement de Ku.

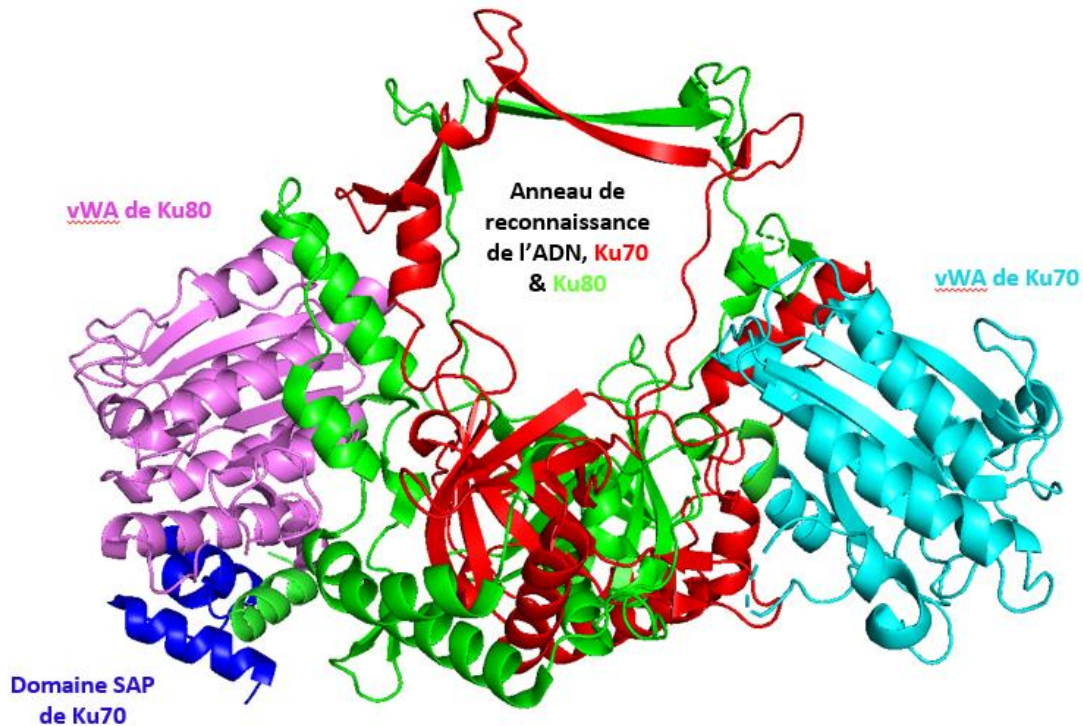


Figure 23 : Structure cristalline de l'hétérodimère Ku70/80 chez l'homme. Le code couleur est donné par la représentation schématique ci-dessous (Figure 24). L'anneau dans lequel vient s'insérer l'ADN est représenté par une surface interne positive (rouge et verte) formée par les deux sous-unités de la protéine. On remarque que la partie C-terminale de Ku80 n'est pas présente sur la structure car cette partie est flexible (code PDB Ku seule : 1JEQ)

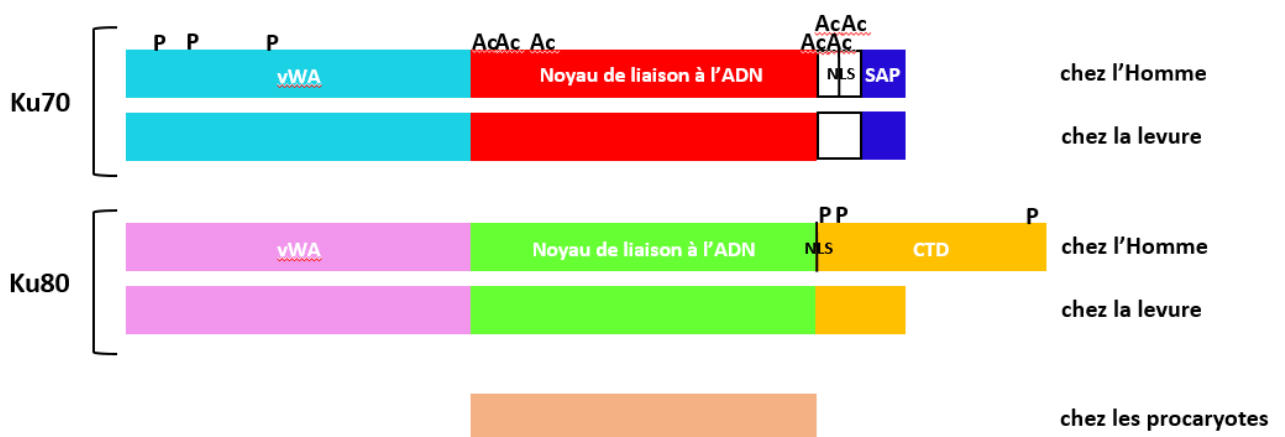


Figure 24 : Représentation schématique des différents domaines de Ku70/80 : comparaison entre les modèles humains, levure et procaryote. Chez l'Homme et la levure, plusieurs domaines sont conservés : le domaine N-terminal de vWA, le domaine central de dimérisation et domaine d'interaction avec l'ADN. Pour le domaine C-terminal, quelques différences sont observées : chez Ku70, on trouve un domaine SAP avec chez l'Homme des modifications post-traductionnelles de type acétylation (Ac) et phosphorylation (P) ; chez Ku80, on ne trouve pas de région CTD pour la levure car la levure n'a pas de DNA-PKcs. On observe que seul le domaine de dimérisation est présent chez les procaryotes

2) *Reconnaissance des extrémités de l'ADN*

Des études ont montré que l'hétérodimère Ku70/80 est retrouvé principalement dans le noyau de la cellule mais est présent dans le cytoplasme. Lors d'expériences de micro-irradiations, des CDB ont été générées au niveau de stries effectuées par micro-irradiations. Ces expériences ont ainsi permis d'observer une co-localisation de la protéine Ku et du variant d'histone γ -H2AX (Britton, Coates, et Jackson 2013). Dans cette étude, il est proposé qu'une ou deux molécules de Ku sont recrutées au niveau de chaque CDB. Nous savons grâce aux expériences d'irradiation, que le temps d'association de Ku aux CDB est de l'ordre de 5 sec (Mari et al. 2006). Ku est capable d'interagir avec diverses extrémités d'ADN double brin pour les protéger des nucléases et de démarrer la réparation (Arosio et al. 2002). Des expériences d'anisotropie de fluorescence ont montré que l'interaction de Ku avec l'ADN est sensible à la force ionique de l'environnement. En effet, plus la concentration en sels est élevée, plus l'affinité entre Ku et l'ADN diminue en accord avec l'interaction fortement électrostatique montrée par la structure cristallographique (Arosio et al. 2004).

Ku est considéré comme une plateforme centrale de recrutement des facteurs du NHEJ de par le nombre d'interactions protéine-protéine directes qu'il effectue avec les facteurs de ligation, de maturation et les polymérases du NHEJ. Trois protéines de cette voie de réparation présentent un motif peptidique conservé : APLF, XLF et WRN. Ce motif d'interaction est appelé KBM (pour *Ku-Binding Motif* en anglais). Ce motif commun aux trois protéines est constitué de résidus basiques et hydrophobes principalement. Ku peut aussi interagir avec d'autres partenaires ne possédant aucun motif d'interactions identifié à ce jour. Pour le complexe XRCC4-LigIV, il a été proposé qu'il existait une interaction directe faible entre l'hétérodimère et le complexe lié à la présence d'ADN (Costantini et al. 2007). Dans la prochaine partie, nous détaillons quelques partenaires de la protéine Ku.

3) *Interaction de Ku avec des partenaires du NHEJ*

Interaction Ku – DNA-PKcs

La DNA-PKcs est une enzyme constituée de trois domaines majeurs et structuraux. De par sa taille d'environ 470 kDa, et ses 4128 résidus, cela fait d'elle la plus grande kinase humaine. Son Nter (résidus 1 à 2801) est formé par une répétition d'un motif que l'on appelle HEAT

(pour *Huntingtin Elongation factor 3, regulatory subunit A of PP2A, TOR1* en anglais) qui forme un anneau ouvert (Chirgadze et al. 2017). Le code PDB de cette enzyme en complexe avec des versions peptidiques de la protéine Ku est 5LUQ. La structure a été résolue à 4,3 Å. Seule la partie N-terminale forme un anneau. Des analyses ont montré que la DNA-PKcs était capable d'interagir fortement avec l'ADN (K_D d'environ 3 nM) mais que lorsque la protéine Ku était déjà présente sur la CDB, alors cette interaction est renforcée (K_D d'environ 35 pM). En présence d'ADN et de Ku, une stimulation de l'activité kinase à sérine/thréonine du complexe Ku-DNA-PKcs ou appelé DNA-PK. Une autophosphorylation est donc initiée sur un total de 15 sites bien spécifiques de l'enzyme (Jette et Lees-Miller 2015) puis des phosphorylations d'autres facteurs du NHEJ notamment Ku. A l'heure actuelle, l'autophosphorylation seule est connue comme essentielle au bon enchaînement de réparation d'une CDB. Aujourd'hui, on sait que la formation de la synapse du NHEJ est due à l'association de la DNA-PKcs et du complexe Ku-ADN est une étape également importante pour la formation de la synapse entre les deux extrémités des CDB. Cette étape permet ainsi la protection et le rapprochement des extrémités de l'ADN lésées. Récemment la structure par cryoEM d'un complexe DNA-PKcs-Ku-ADN a été résolu à 6.6Å de résolution (Yin et al. 2017). Cette structure met en évidence l'interaction du complexe DNA-PK avec un ADN de 30bp à travers un tunnel continu allant de Ku vers DNA-PKcs.

Interaction Ku – APLF

De par des recherches sur des collections de données de séquences proches de la kinase PNK (pour *PolyNucleotide Kinase* en anglais) et de séquences protéiques possédant un domaine FHA (pour *ForkHead Associated domain* en anglais), le partenaire APLF du NHEJ a été découvert. C'est une enzyme à activité à la fois endo et exo-nucléase (Kanno et al. 2007) qui vient interagir sur le NHEJ lorsque les extrémités de l'ADN ne sont pas religables de suite et qu'il faille une étape de maturation. Cette enzyme est capable de cliver les sites abasiques et ainsi de retirer la base azotée qui relie le dommage à la molécule d'ADN. Cette activité est possible grâce à ses domaines C-terminaux à doigts de zinc (ZnF) qui sont indispensables (Kanno et al. 2007). Une étude a montré que la DNA-PKcs était recrutée à la fois au niveau des CDB mais aussi des cassures simple brin (CSB) grâce à son domaine FHA respectivement par les protéines phosphorylées XRCC4 et XRCC1 (Kanno et al. 2007). A propos de l'étude des CDB, nous savons depuis récemment que la phosphorylation de XRCC4 est importante pour son interaction avec le domaine FHA de APLF. C'est suite à son état de phosphorylation, que le partenaire XRCC4 peut être recruté au niveau des CDB (code PDB : 5E50) (Figure 25A) (Cherry et al.

2015). Lors des études sur le recrutement de la protéine APLF sur les CDB, il a été montré qu'une zone particulière de l'hétérodimère Ku80 était en interaction avec APLF (Kanno et al. 2007). Cette zone est un motif d'un total de 19 résidus qui serait nécessaire à l'interaction entre Ku80 et APLF. Ce motif est noté KBM (pour *Ku-Binding Motif* en anglais) et il est localisé dans la région non-repliée de la nucléase (Grundy et al. 2012; Shirodkar et al. 2013). Lors de l'étape de maturation des extrémités de l'ADN, c'est ce motif qui est responsable du recrutement de APLF au niveau des extrémités lésées par la CDB. Des expériences de mutations systématique des résidus hydrophobes, situés en surface et conservés de Ku80 ont permis d'identifier des acides aminés de Ku80 impliqués dans l'interaction avec le KBM de APLF. Ils sont situés dans une même région du domaine vWA de Ku80 (leucine 68, tyrosine 74 et isoleucine 112) (Figure 25B) (Grundy et al. 2012).

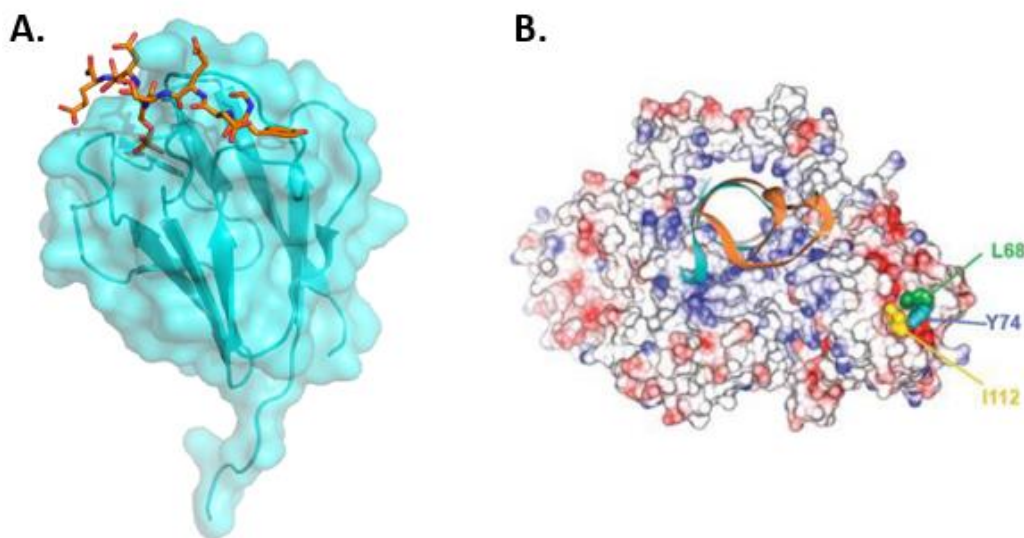


Figure 25 : **A.** Structure cristalline du domaine FHA de APLF avec le peptide 229-YDEpSpTDEES-237 de XRCC4 (pS pour phosphoSérine et pT pour phosphoThréonine). **B.** Position des résidus du vWA de Ku80 identifiés par mutagenèse dirigés comme impliqués avec APLF (d'après xxx)

Récemment, la structure de Ku (version tronquée des parties C-terminales flexibles) en complexe avec un ADN de type tige-boucle et un peptide de 18 résidus du A-KBM (pour APLF *Ku-Binding Motif*) (résidus 174-191) a été résolue (Figure 26A) (Nemoz et al. 2018). L'hypothèse précédente est affirmée car APLF interagit de manière directe avec Ku. Son A-KBM est en interaction avec la périphérie du domaine de vWA de Ku80. De plus le peptide montre une conformation étendue dans un site délimité par deux hélices ($\alpha 4$ et $\alpha 5$) et une boucle que l'on retrouve entre les feuillets β B et C de Ku80 (Nemoz et al. 2018). Des mesures de microcalorimétrie ont montré que ce A-KBM a affinité avec un K_D de 23 ± 2 nM pour le complexe Ku-ADN et que cette affinité est légère plus faible sans ADN (Figure 26B) (Nemoz et al. 2018).

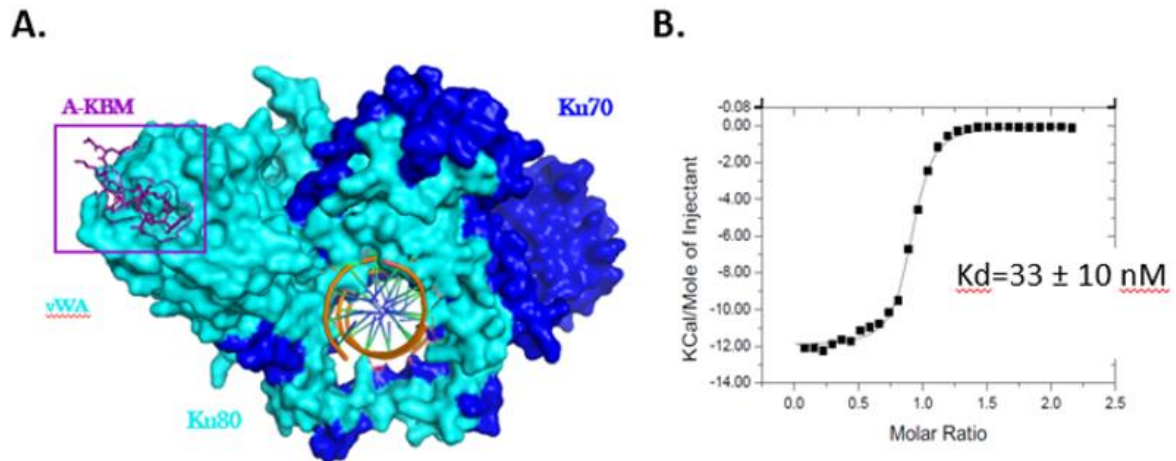


Figure 26 : **A.** Vue globale de la structure cristalline du complexe KuCC-hDNA-pAPLF. Le peptide pAPLF du motif A-KBM (en violet) se lie à la périphérie du domaine de vWA de Ku80 (bleu cyan). **B.** Mesure de microcalorimétrie de l’affinité entre Ku et APLF sans ADN. Cette affinité est légèrement plus forte en présence de la molécule d’ADN (isotherme non montré)

Il a été montré que dans des cellules de type A549, où l’expression de la nucléase APLF est altérée, la réparation des extrémités de l’ADN est diminuée. Il a été proposé que la fonction de APLF serait de maintenir une certaine concentration locale du complexe LigaseIV/XRCC4 à proximité des extrémités d’ADN lésées(Grundy et al. 2012).

Interaction Ku – XLF

La protéine XLF appelée aussi Cernunnos a été identifiée indépendamment par les laboratoires de JP de Villartay (Institut Imagine, Paris) et S Jackson (FmedSci, Cambridge). Le premier laboratoire est parti de cellules de patients SCID (pour *Severe Combined Immuno Deficiency* en anglais). Lors d’un criblage d’une collection d’ADNc, un gène codant pour une protéine de 299 résidus correspondant à Cernunnos, a été identifié comme permettant la complémentation des cellules de ces patients (Buck et al. 2006). Suite à cette découverte, la protéine XLF (*XRCC4 Like Factor*) a été identifiée cette fois à l’aide d’un criblage double-hybride conçu pour identifier les partenaires de XRCC4. Les deux laboratoires ont montré à l’aide de logiciel de prédiction de structure que ce nouveau facteur XLF/Cernunnos présentait une structure probablement proche de celle de XRCC4 (Ahnesorg, Smith, et Jackson 2006). Les structures cristallographiques résolues quelques années après ont confirmé cette proposition et ont montré, que la région N-terminale (résidus 1-224) de la protéine XLF adopte un repliement similaire à celui de la protéine XRCC4 (Figure 27) (codes PDB : 2R9A et 2QM4) (Andres et al., 2007; Li et al. 2008).

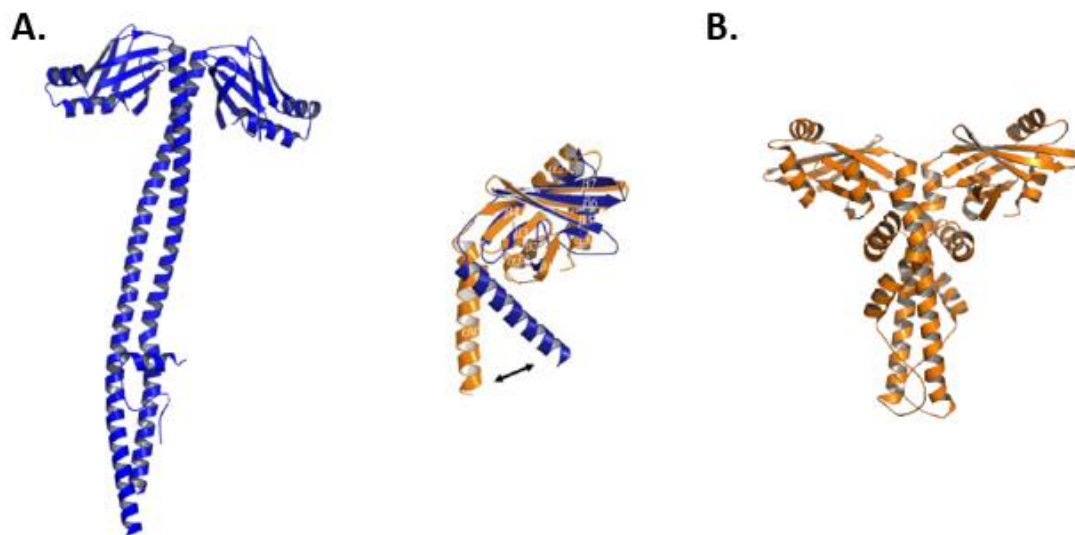


Figure 27 **A.** Structure cristalline de la protéine XRCC4. **B.** Structure cristalline de la protéine XLF résolue à 2,5 Å (orange) montre une similarité de structure avec la protéine XRCC4 (bleue). La superposition des deux structures (centre) montre un décalage dans l'angle du départ du *coiled-coil* avec une variation d'angle de 35 ° supérieure pour XLF (Andres et al., 2007)

XLF ne possède aucune activité enzymatique mais comme la protéine XRCC4, elle est en interaction avec différentes protéines du NHEJ. Comme indiqué plus haut, elle interagit par sa tête globulaire N-terminale avec XRCC4 et est capable de former des filaments avec cette protéine. Il a été montré que XLF était recruté par Ku au niveau des CDB et que ce recrutement de XLF aux CDB était stabilisé par la présence de XRCC4 (Yano et al. 2011). Il a été ensuite démontré que XLF possédait un motif situé en Cter (notamment ses derniers résidus) et semblable au motif de APLF impliqué dans son interaction avec Ku (Grundy et al. 2012). Des gels retards ont montré que le motif X-KBM de nature peptidique pouvait interagir à lui seul avec le complexe Ku-ADN. L'équipe de Caldecott a alors proposé par des mesures d'anisotropie et de recrutement après micro-irradiation que les deux motifs de XLF et APLF avaient des affinités différentes avec la protéine Ku. Ces motifs ont alors été nommés « X-KBM et A-KBM » (Rulten & Grundy, 2017).

Récemment, la structure de Ku (version tronquée des parties C-terminales flexibles) en complexe avec un ADN de type tige-boucle et un peptide de 19 résidus du X-KBM (pour XLF *Ku-Binding Motif*) (résidus 281-299) a été résolue (Figure 28A) (Nemoz et al. 2018). Lorsque le X-KBM est en interaction avec Ku, une ouverture du domaine de vWA est observée (Figure 28A). Cette rotation du vWA de Ku80 permet la formation d'une cavité entre le vWA et le restant de l'hétérodimère qui permet au X-KBM de venir se loger. La taille de la cavité formée est d'environ 20 Å de largeur, 30 Å de longueur et 20 Å de hauteur. En comparaison avec la structure 1JEY résolue par Walker, cette ouverture du vWA est très nette. En microcalorimétrie,

des mesures ont montré que le motif X-KBM était moins afin pour Ku que le motif A-KBM, avec une affinité de l'ordre de quelques micromolaires contre 20-30nM pour APLF (Figure 28B) (Nemoz et al. 2018).

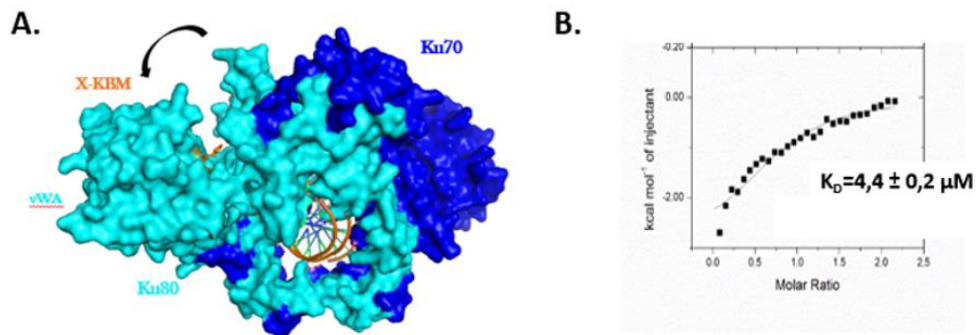


Figure 28 : **A.** Vue globale de la structure cristalline du complexe KuCC-hDNA-pXLF. Le peptide pXLF du motif X-KBM (orange) se loge dans la cavité formée par l'ouverture du domaine de vWA de Ku80 (bleu cyan). **B.** Mesure de microcalorimétrie montrant que l'affinité entre Ku et XLF dépend peu de la présence de l'ADN (isotherme non montré)

Interaction Ku – IP6

La DNA-PKcs possède un domaine catalytique en Cter qui est semblable à ce qui est retrouvé dans la famille des kinases phosphatidyl inositols 3 (PI 3) (Hartley et al. 1995). Des études *in vitro* et *in vivo* ont montré, dans les années 2000, qu'un facteur additionnel semblait participer à la voie NHEJ. Ce facteur a été initialement appelé SFA (pour *Stimulating Factor A* en anglais). Il a été caractérisé après de nombreuses étapes de fractionation et de purification d'extraits cellulaires par le laboratoire de S West. (Les A Hanakahi et al. 2000). Après avoir montré que ce n'était ni une protéine, ni ADN, ni un ARN, et l'avoir purifié par différentes colonnes, ce laboratoire a finalement caractérisé par RMN, la molécule stimulante comme un inositol hexakisphosphate (IP6) (Figure 30A). Ces auteurs ont montré alors que l'IP6 permet en effet une certaine stimulation de la voie NHEJ de par son interaction avec la DNA-PK (complexe DNA-PK-Ku70-Ku80). Il a été montré que le nombre de phosphates, six, a un rôle dans cet effet de stimulation (Figure 29) (Les A Hanakahi et al. 2000). Des études supplémentaires ont alors montré que l'IP6 interagit avec la protéine Ku et que cette interaction est indépendante de l'interaction Ku-ADN (L. A. Hanakahi 2002).

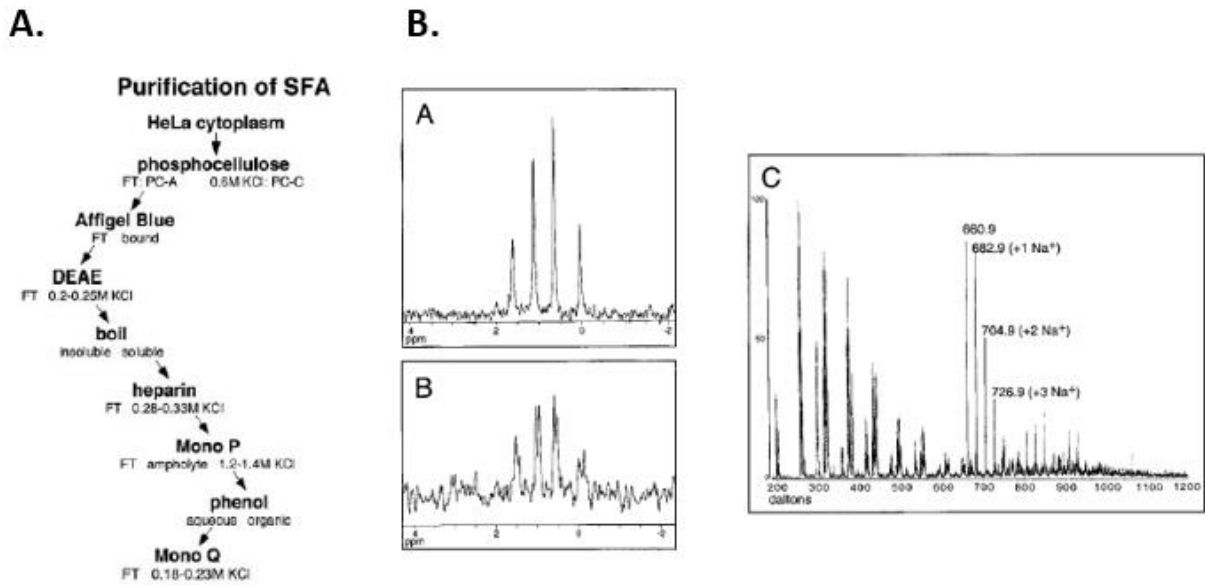


Figure 29 : A. Schéma des étapes chromatographiques de purification du facteur SFA depuis des extraits cellulaires cytoplasmiques de HeLa. B. Spectres du phosphore découplés des protons indiquant 4 pics proches de la référence d'acide phosphorique à 1,4909 ppm, 0,5554 ppm et 20, 0850 ppm suggérant la structure de l'inositol hexakisphosphate (IP6) (Les A Hanakahi et al. 2000)

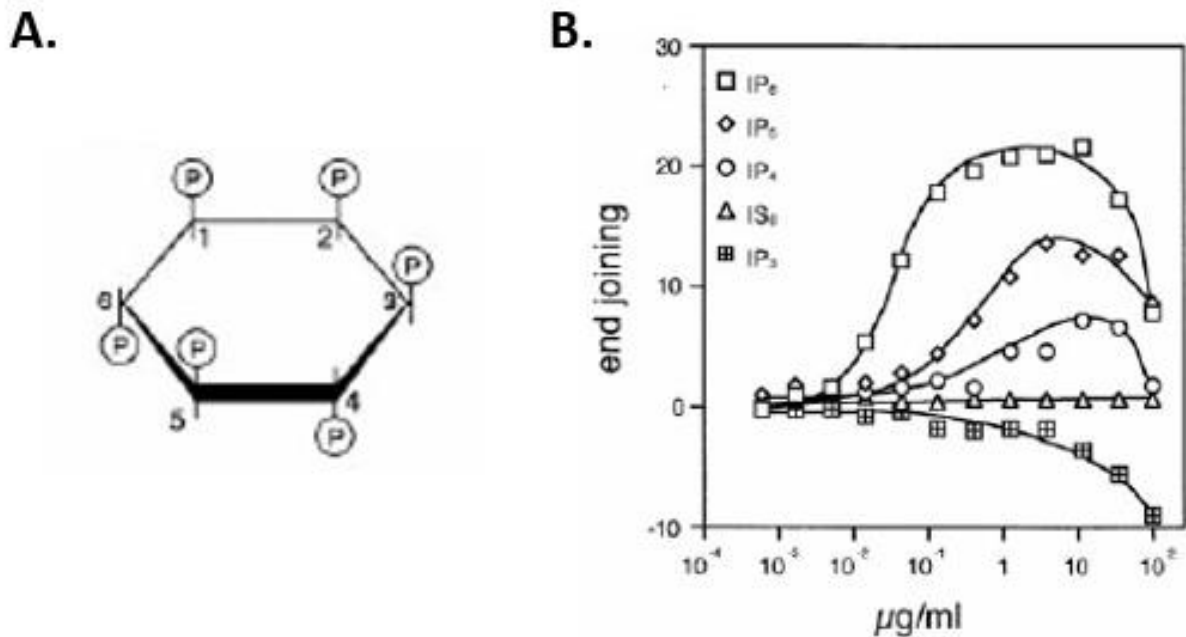


Figure 30 : A. représentation de la molécule d'IP6. B. Effet de l'IP6 sur la voie NHEJ. L'IP6 a un effet stimulateur sur le NHEJ par rapport à IP5, IP4 et IP3 dans un essai de mesures de ligation des extrémités de cassures double-brin (Les A Hanakahi et al. 2000)

A partir de l'analyse de la réparation de charges positives et négatives sur la structure de Ku70/80, (Figure 31) il a été proposé un site potentiel de liaison de l'IP6 sur l'hétérodimère Ku70/80. Ce site contient notamment les résidus lysine 238 et lysine 239 appartenant à la sous-unité Ku80 et lysine 357 appartenant à la sous-unité Ku70 (Cheung, Salerno, et Hanakahi 2008).

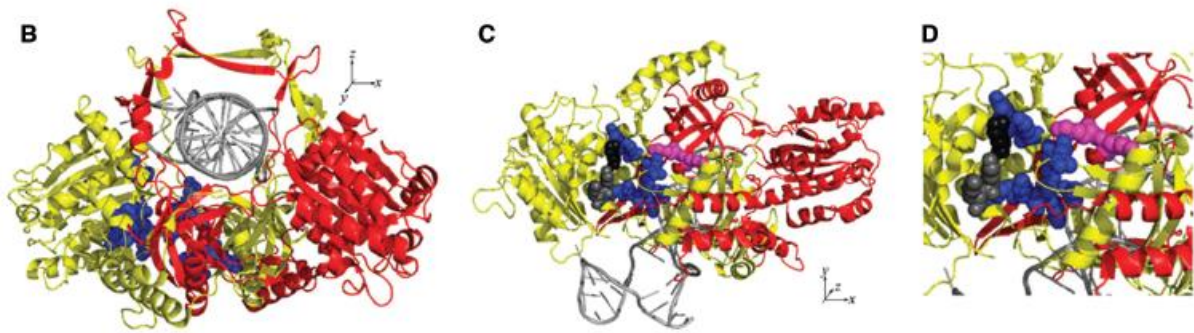


Figure 31 : Identification d'un site de liaison potentiel de l'IP6 sur Ku. La structure utilisée est 1JEY ; Ku70 en rouge, Ku80 en jaune et ADN en gris. Les acides aminés conservés sont colorés en bleu et localisation à l'interface des deux sous-unités de Ku70/80 (Cheung, Salerno, et Hanakahi 2008)

L'équipe de Hanakahi montre *in vitro*, que Ku a une spécificité d'interaction pour l'IP6 et que le nombre de phosphates présents sur l'inositol module la stimulation par l'IP6 de la voie NHEJ. (Cheung, Salerno, et Hanakahi 2008).

Second objectif

Au cours de ma thèse, un second volet a porté sur la caractérisation d'autres interactions de type protéine-ADN impliquées dans la voie NHEJ. La première étape a été de purifier en grande quantité l'hétérodimère Ku humain. J'avais pour objectif d'étudier par cristallographie aux rayons X et par différentes techniques de biophysique, l'interaction entre Ku et l'ADN en présence ou en absence de la petite molécule d'IP6. Nous voulions obtenir la structure cristallographique de Ku-ADN en présence d'IP6 pour décrire le site d'interaction de l'IP6 proposé comme une molécule stimulatrice de la voie NHEJ. Nous voulions aussi quantifier et préciser les paramètres cinétiques et thermodynamiques de cette interaction entre Ku et IP6 en présence et en absence d'ADN en combinant diverses approches biophysiques. Le switchSENSE a également été utilisé pour le projet sur le NHEJ. J'ai également été impliquée au début de ma thèse dans la caractérisation d'interactions entre Ku et les motifs X-KBM et A-KBM respectivement les motifs présents dans les protéines XLF et APLF. Je me suis ensuite intéressée à Ku-pXLF en présence d'IP6 afin de savoir si cet inositol phosphate avait un rôle de stimulation au sein de cette interaction.

Cette thèse CIFRE avait pour objectif de me confronter à diverses techniques de biophysique et de biologie structurale pour étudier notamment des interactions de type protéine-ADN. Le but était de pouvoir faire le lien entre la recherche fondamentale et les applications industrielles.

La suite du manuscrit concerne une partie intitulée « Matériel et Méthodes », une partie « Résultats » et enfin une partie « Discussion et perspectives ».

MATERIEL & METHODES

PARTIE I :

Etudes structurales et fonctionnelles des interactions de Ku70/80 avec des facteurs du NHEJ et des petites molécules

A) A propos de l'hétérodimère Ku70/80

La protéine Ku70/80 est un hétérodimère formé de deux sous-unités : Ku70 (retrouvée aussi sous le nom de XRCC6) composée de 608 acides aminés pour une masse moléculaire de 69484 Da et Ku80 composée de 732 acides aminés pour une masse de 84076 Da. Cet hétérodimère est en interaction avec la double hélice d'ADN de manière assez unique. En effet, les deux sous-unités forment un anneau venant encercler l'ADN. Cette interaction est indépendante de la séquence en ADN. La structure cristallographique de Ku avec un ADN possédant une structure dite de tige boucle (hDNA) a été résolue en 2001 (Walker et al, 2001). Les acides aminés 34 à 609 et 6 à 565 respectivement de Ku70 et Ku80 ont été modélisés. Les codes PDB des deux structures Ku seule et Ku en complexe avec l'ADN sont respectivement 1JEQ et 1JEY. Quelques informations sont désormais disponibles quant aux différents domaines de Ku. En effet, Ku70 possède un domaine von Willebrand (vWA) dans sa partie N-terminale, un domaine central qui permet à la fois la dimérisation avec Ku80 et l'interaction avec l'ADN et un domaine C-terminal, appelé domaine SAP. La sous-unité Ku80 est composée de manière homologue à la sous-unité Ku70 avec en N-terminal un domaine vWA, en partie centrale un domaine à la fois de dimérisation avec Ku70 et d'interaction avec l'ADN. S'ensuit environ 70 acides aminés prédits comme intrinsèquement désordonnés et un domaine Cter en hélices α qui interagit avec DNA-PKcs (Figure 32). La région désordonnée de 70aa crée une flexibilité entre le cœur de Ku70/80 et le domaine Cter de Ku80 en hélices α (code PDB : 1RW2) (Zhang et al, 2004).

Les expériences décrites dans ce manuscrit ont été réalisées essentiellement avec une construction réduite de la protéine Ku70/80. En effet, elles ont été faites pour la plupart avec une construction Ku70C/80C (appelée par la suite KuCC). Cette construction est raccourcie de ses extrémités C-terminales (65 acides aminés ôtés en Cter de Ku70 (cad Ku70 1-544) et 58 acides aminés ôtés en Cter de Ku80 (cad Ku80 1-541)). Ce choix a été fait afin d'essayer de

faciliter la cristallisation de l'hétérodimère en complexe avec différents ligands (petite molécule (IP6), ADN hairpin ou peptides dérivés de protéines partenaires). Supprimer les parties C-terminales de Ku a pour objectif de limiter la présence de parties flexibles de l'hétérodimère et d'améliorer ainsi sa cristallisation seule ou en complexe avec des ligands. Quelques expériences ont également été établies avec une construction légèrement différente de celle de KuCC, c'est-à-dire avec la construction KuSC pour laquelle 23 acides aminés supplémentaires ont été supprimés en Nter de Ku70 sur la base de KuCC. Ces 23 acides aminés ne sont pas visibles dans les premières structures de Ku et semblent donc présenter une forte flexibilité.

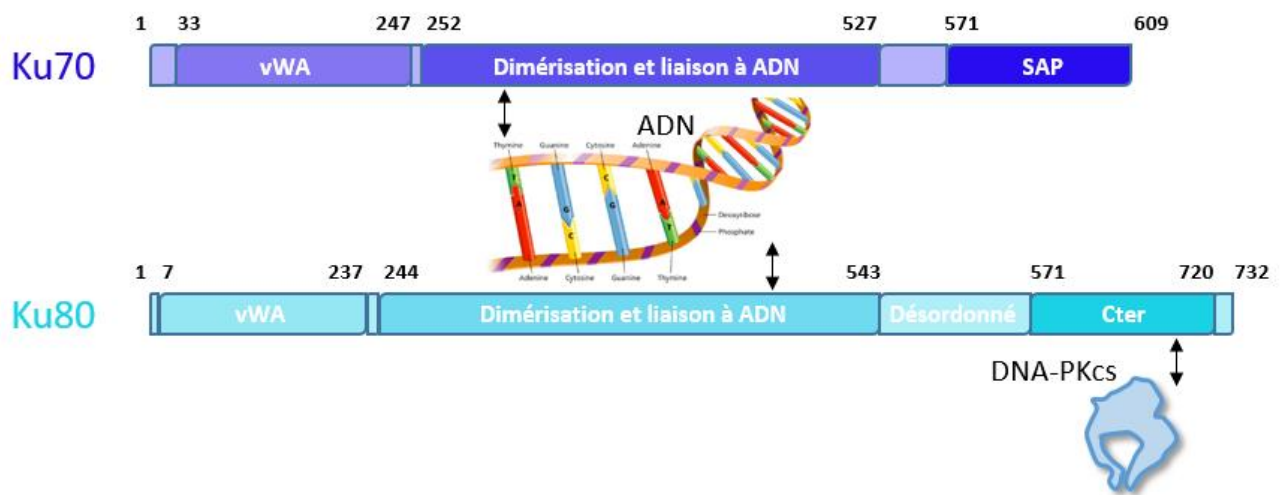


Figure 32 : Représentation schématique des différents domaines de l'hétérodimère Ku70/80 de la voie NHEJ chez l'humain.

SEQUENCE Ku70 (PDBID : 1JEQ : A)

>sp|P12956|XRCC6_HUMAN X-ray repair cross-complementing protein 6 OS=Homo sapiens OX=9606 GN=XRCC6 PE=1 SV=2

1 - 544 ⇔ Ku70CC

24 - 544 ⇔ Ku70SC

MSGWESYYKTEGDEEAEEEEQEENLEASGDYKYSGRDSLIFLVDASKAMFESQSEDELTPFDMSIQCIQSVYISKI
 ISSDRDLLAVVFYGTEDKDKNSVNFKNIIYVLQELDNPGAKRILELDQFKGQQGQKRFQDMMGHGSDYSLSEVLWVC
 ANLFSDVQFKMSHKRIMLFTNEDNPHGNDSAKASRARTKAGDLRDTGIFLDMHLKPKGGFDISLFYRDIISIAE
 DEDLRVHFEESKLEDLLRKVRAKETRKRALSRKLNKNDIVISVGIYNLVQKALKPPPILKLYRETNEPVKTKT
 RTFNTSTGGLLLPSDTKRSQIYGSRQIILEKEETEELKRFDDPGLMLMGFKPLVLLKHHYLRPSLFVYPEESLV
 IGSSTLFSALLIKCLEKEVAALCRYTPRRNIPPYFVALVPQEEELDDQKIQVTPPGFQLVFLPFADDKRKMPFTE
 KIMATPEQVGMKAIVEKLRFTYRSDFSFPVLLQQHFRNLEALALDLMEPEQAVDLTLPKVEAMNKRLGSLVDEF
 KELVYPPDYNPEGKVTKRKHDNEGSGSKRPKVEYSEELKTHISKGTGLGKFTVPMLKEACRAYGLKSGLKKQELL
 EALTKHFQD

SEQUENCE Ku80 (PDBID : 1JEQ : B)

>sp|P13010|XRCC5_HUMAN X-ray repair cross-complementing protein 5 OS=Homo sapiens OX=9606 GN=XRCC5 PE=1 SV=3

1 - 551 10His TEV Nter ⇔ Ku80CC

1 - 551 10His TEV Nter ⇔ Ku80SC

MHHHHHHHHHHMENLYFQGVRSGNKAAVVLCMDVGFMTSNSIPGIESPFQAKKVIITMFVQRQVFAENKDEIALV
 LFGTDGTDNPLSGGDQYQNITVHRHMLPLPDFLLEDIESKIQPGSQADFLDALIVSMDVIQHETIGKKFEKRHI
 EIFTDLSSRFKSKQLDIIHSLKKCDISLQFFLPFSLGKEDGSGDRGDGPFRLGGHGSPFPLKGITTEQQKEGLEI
 VKMVMISLEGEDGLDEIYSFSESLRKLCVFKKIERHSIHWPCRLTIGSNLSIRIAAYKSILQERVKKTWTVVDAK
 TLKKEDIQKETVYCLNDDDETEVLKEDIIQGFRYGSDIVPFPSKVDEEQMKYKSEGKCFSVLGFCKSSQVQRFFM
 GNQVLKVF AARDDEAAAVALSSLIHALDDLDMVAIVRYAYDKRANPQVGVAFPHIKHNYECLVYVQLPFMEDLRQ
 YMFSSLNKSKKYAPTEAQLNAVDALIDSMSLAKKDEKDTLEDLFPPTKI PNPRFQRLFQCLLHRALHPREPLPP
 IQQHIWNMLNPPAEVTTKSQIPLSKIKTLFPLIEAKKKDQVTAQEIFQDNHEDGPTAKKLKTEQGGAHFSVSSLA
 EGSVTSVGSVNPAENFRVLVKQKASFEENASQLINHIEQFLDTNETPYFMKSIDCIRAFREEA IKFSEEQRFN
 FLKALQEKVEIKQLNHFEIVVQDGITLITKEEASGSSVTAEAKKFLAPKDKPSGDTAAVFEEGGD VDDLLDMI

	<u>KuFL</u>	<u>KuCC</u>	<u>KuSC</u>
MW (Da)	154 771	130 804	125 184
ϵ (cm ⁻¹ .M ⁻¹)	85 150	75 180	66 700

Figure 33 : Tableau récapitulatif des masses moléculaires et epsilons molaires des constructions de Ku70/80

B) Préparation des échantillons protéiques

Une plateforme d'expression Multibac de protéines en cellules d'insectes infectées par un baculovirus a été développée au laboratoire depuis 2012 (Figure 34) avec l'aide de l'équipe collaborative d'Imre Berger (EMBL, Grenoble puis Université Bristol, UK). Durant ma thèse, j'ai été formée sur les techniques de clonage et production en cellule d'insecte. J'ai réalisé plusieurs fois les expériences afin de me familiariser avec l'expression en cellules d'insectes. Cependant, la majorité des clonages, des virus et des cultures ont été réalisés principalement par la plateforme du laboratoire dirigée par Virginie Ropars. Mon travail de thèse a démarré principalement avec la purification des protéines d'intérêt (Ku et XLF) pour les expériences de biophysique et de biologie structurale.

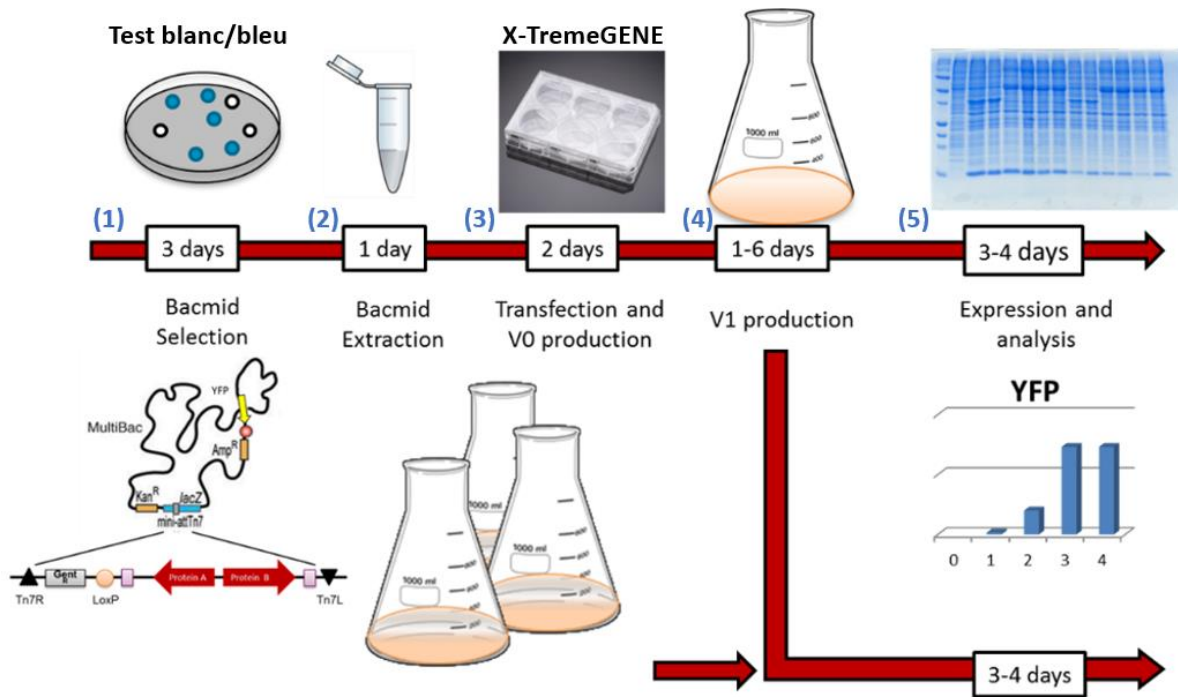


Figure 34 : Schéma du processus d'expression de protéines recombinantes en cellules d'insectes. Description des étapes du processus avec production du virus et des protéines. La technique MultiBac permet de suivre le niveau d'expression dans les cellules d'insectes au cours du temps par une technique d'analyse du phénotype cellulaire assistée par ordinateur (FACS pour *Fluorescence-Assisted Cell Sorting* en anglais) grâce à une protéine YFP insérée dans le bacmide.

1) Clonage de Ku70/80 et XLF dans des vecteurs pFL

Cloner est le terme utilisé pour désigner l'insertion d'un gène d'intérêt (fragment d'ADN) dans un vecteur ou plasmide. Ce vecteur est ensuite recombiné avec le matériel génétique du virus, appelé bacmide grâce au transposon Tn7 et au plasmide transposase Helper présent dans la souche bactérienne DH10EmBacY (Fig 34). L'insertion réussie de notre gène d'intérêt dans le bacmide interrompt le gène de la β -galactosidase. Ainsi, à l'aide d'un test blanc/bleu, seules les colonies blanches possédant notre gène d'intérêt sont sélectionnées (étape 1 Fig 34). Après sélection, le bacmide est extrait (étape 2 Fig 34) et inséré dans des cellules d'insecte en utilisant le transfectant X-TremeGENE (Roche) permettant le démarrage de l'infection et la génération de premiers virus, appelés V0 (étape 3 Fig 34). Ces premiers virus sont ensuite cultivés pour donner une seconde source de virus appelée V1, stockée pendant une année à 4 °C et permettant de faire les cultures en grands volumes (étape 4 Fig 34).

Pour les protéines Ku et XLF, le protocole utilisé est le même. Je détaille dans ce paragraphe la méthode pour Ku70/80. Des gènes synthétiques sont commandés chez Genscript et permettent

d'obtenir le vecteur codant pour l'hétérodimère Ku70/80. La compagnie a développé depuis quelques années un algorithme appelé OptimumGene. Les paramètres optimisés ici au travers des mutations silencieuses sont notamment l'adaptation aux codons les plus fréquents au sein de l'hôte (http://www.genscript.com/codon_opt.html); ceci via des mutations silencieuses. De plus, la sous-unité Ku80 présente dans son domaine en Nter, une séquence de 10 histidines suivie d'un site de coupure par la protéase TEV (Tobacco Etch Virus). La protéase TEV avec un Tag 6His est produite au laboratoire. Les séquences codantes des deux sous-unités Ku70/80 sont clonées dans un vecteur pFL grâce à des enzymes de restriction XhoI, NheI et BamHI, HindIII conduisant respectivement à l'obtention du vecteur pFL Ku70_80_opt. Afin d'obtenir la version tronquée de Ku70/80, Ku70C/80C, la PCR est utilisée pour introduire des codons STOP dans la séquence des protéines aux endroits clés dans les domaines C-terminaux flexibles de Ku70 et Ku80. Pour la construction KuSC, la PCR permet d'amplifier la région d'intérêt sur Ku70 (résidus 24 à 544) pour obtenir l'ADN Ku 70 SC avec aux extrémités 5' un site XhoI et en 3' NheI. L'ADN est digéré par ces deux enzymes. L'ADN de Ku70C est ensuite ôté sur le plasmide pFL-KuCC par digestion enzymatique (NheI-XhoI). Ce plasmide ne contenant plus que Ku80C est purifié puis une ligation est faite entre KuSC et pFL-Ku80C pour obtenir pFL-KuSC. Ce plasmide est ensuite recombiné avec le bacmide.

2) *Expression des protéines en cellules d'insectes*

Pour exprimer Ku et XLF en cellules d'insectes, des cultures de cellules Sf21 de 600 mL, contenant une concentration de cellules de $0,5 \cdot 10^6$ cell.mL⁻¹, sont infectées par une solution de virus V1. Une méthode d'infection avec un faible taux de virus ou multiplicité d'infection (MOI) est réalisée. Une MOI de $5 \cdot 10^{-3}$ est utilisée selon les optimisations réalisées dans le laboratoire de Imre Berger et selon notre propre expérience au laboratoire. Cette MOI basse permet aux cellules d'insectes de doubler deux fois de suite avant de s'arrêter de croître suite à l'infection. Lorsque le doublement du nombre de cellules s'arrête (DPA), les virus commencent à produire notre gène d'intérêt et le gène rapporteur YFP contenus dans le bacmide. La fluorescence de la YFP est mesurée par FACS; c'est une bonne indication du niveau d'expression des protéines d'intérêt dans les cellules d'insectes (Figure 34). Lorsqu'un plateau de fluorescence est atteint, la culture est arrêtée. Cela correspond en général à 1 jour après la date d'arrêt de prolifération (DPA pour *Date of Proliferation Arrest* en anglais). Cette étape marque le début de la mort cellulaire. Les cellules sont alors récoltées, centrifugées une première fois 4 min, 1 800 rpm, 4 °C dans des tubes de 800 mL puis une seconde fois 4 min,

1 000 rpm, 4 °C dans des falcons de 50 mL. Les culots cellulaires sont ensuite lavés avec un tampon PBS puis congelés dans de l'azote liquide pour être stockés à – 80 °C par culot correspondant à 600 mL de culture.

3) Purification de l'hétérodimère Ku70/80 et de la protéine XLF

Les protocoles de purification des deux protéines sont semblables quant aux étapes utilisées. Seuls les tampons diffèrent légèrement. Le protocole suivant est le protocole détaillé de la protéine Ku70/80. Deux tableaux récapitulatifs des tampons utilisés pour chaque protéine sont insérés à la suite de ce paragraphe afin de les comparer les tampons.

- **Du culot cellulaire au lysat soluble**

Les deux sous-unités Ku70 et Ku80 sont tronquées de leur partie C-terminale afin de ne conserver que les parties protéiques structurées. Le culot protéique correspondant à 600 mL de culture de cellules est décongelé sur glace et repris dans 100 mL de tampon de lyse NiNTA-Ku (20 mM Tris pH 8,0, 150 mM KCl, 850 mM NaCl, 10 mM β -mercaptoéthanol, 40 mM imidazole + 1 pastille cocktail d'inhibiteurs de protéases ROCHE EDTA-free 1X). Une fois que le culot est remis en suspension dans le tampon de lyse, 20 μ L d'échantillon sont prélevés pour l'analyse qui suivra par gel d'électrophorèse SDS-PAGE (cet échantillon est noté ET pour « extrait total »). Afin de libérer le contenu cellulaire et la protéine d'intérêt, les membranes cellulaires sont lysées par sonication (méthode par ultrasons) : 6 x 30 sec ; 30 sec on / 30 sec off ; 2 400 Joules au total. Pour garder la température aux alentours de 4 °C, le lysat est placé dans un mélange glace / glace carbonique / éthanol 20 %. Lorsque l'étape de sonication est terminée, 1 mM MgCl₂ et de la benzonase (MerckTM, 3,5 μ L d'enzyme à 30 UI/uL) sont ajoutés à l'extrait cellulaire et le mélange est maintenu sur glace 30 min. L'action de la benzonase est de dégrader progressivement l'ADN en solution. Celle-ci est active en présence de MgCl₂. Après ce temps d'incubation sur glace, l'extrait cellulaire est centrifugé à 20 000 rpm (équivalent à 44 800 g) pendant 30 min à une température de 4 °C afin de séparer culot (C) et surnageant (S). Le culot est repris dans 20 mL de SDS 2 % et forme la fraction insoluble du lysat C ; on garde 20 μ L d'échantillon pour déposer sur gel. Le surnageant, lui, est filtré sur filtre 0,22 μ m (SartoriusTM filtre pour seringue MiniSart NML 0,22 μ m, stérile) afin d'éliminer un maximum de précipité.

- **Première étape de purification : étape Ni-NTA**

Un volume de 10 mL de phase de résine Ni-NTA (Macherey) est équilibré avec trois fois 20 mL H₂O et deux fois 10 mL de tampon de lyse. Le lysat S est alors déposé sur les billes de nickel et l'ensemble est mis 1 h sur une roue à 4 °C. Après sédimentation totale de la résine, la fraction non-retenue sur cette dernière (appelée FT pour *Flow-Through* en anglais) est mise de côté et 20 µL d'échantillon sont mis de côté pour le gel SDS. La résine est alors lavée avec une fois 50 mL de tampon de lyse (20 µL sont conservés pour le gel, fraction = lavage 1), une seconde et une troisième fois avec 50 mL de tampon de lavage. Cette étape permet de se débarrasser de certaines protéines retenues de manière non-spécifique. La présence de 1 M KCl/NaCl permet de se séparer d'une fraction d'ADN potentiellement liée à la protéine d'intérêt Ku. Afin d'éluer ensuite l'hétérodimère, un volume de 30 mL correspondant à 3 volumes de colonne (CV pour *Column Volume* en anglais) de tampon d'élution (20 mM Tris pH 8,0, 150 mM NaCl, 300 mM imidazole, 10 mM β-mercaptoéthanol) sont passés et des fractions d'élution de 8 mL sont récoltés. Un gel d'électrophorèse SDS-PAGE 10 % est réalisé : 20 µL d'échantillon de chaque étape (Extrait Total, Culot, Surnageant, Lavage et éluions) sont mélangés avec 10 µL de bleu de Laemmli 3X, vortexés et incubés 5 min à 95 °C. Les fractions d'élutions sont dosées par Bradford, analysées par spectrophotométrie UV et regroupées en fonction du gel SDS-PAGE. Les fractions comprenant notre protéine d'intérêt Ku sont regroupées et insérées dans un boudin de dialyse Spectra/Por de 12/14 kDa pour être dialysées une fois sur la nuit (ON) puis une fois deux heures contre un tampon QA-Ku (20 mM Tris pH 8,0, 50 mM NaCl, 50 mM KCl, 10 mM β-mercaptoéthanol) à 4 °C.

- **Seconde étape de purification : étape ResourceQ**

Le dialysat est repris et filtré de nouveau sur 0,22 µm. Les tampons QA-Ku (20 mM Tris pH 8,0), 50 mM NaCl, 50 mM KCl, 10 mM β-mercaptoéthanol) et QB-Ku (20 mM Tris pH 8,0, 850 mM NaCl, 150 mM KCl, 10 mM β-mercaptoéthanol) sont également filtrés sur 0,22 µm avant l'équilibration de la colonne ResourceQ échangeuse d'anions (GE Healthcare) d'un volume de 6 mL par le tampon QA-Ku. Suite à l'équilibration de la colonne par 5 CV en QA-Ku, 5 CV en QB-Ku et de nouveau 5 CV en QA-Ku, le dialysat est injecté sur celle-ci et 10 CV de tampon de QA-Ku sont passés pour le lavage. Est ensuite appliqué un gradient de sels de 0 à 40 % durant 12 CV à 2 mL.min⁻¹. Les fractions de FT (lors de l'accrochage de la protéine sur la colonne), lavage (lors des 10 CV de tampon de dialyse QA-Ku) et d'élution de la protéine d'intérêt Ku (lors du gradient de concentration en sels) sont collectées.

Nous observons alors un ou deux pics d'absorbance à une longueur d'onde de 280 nm selon les purifications. Il a été montré par spectrométrie de masse que les deux pics contenaient exactement la même composition en Ku. Une fois l'élution terminée, la colonne est nettoyée grâce au passage successif de 3 CV de tampon QB-Ku, 5 CV de 2 M NaCl pour supprimer tous les contaminants restants, 10 CV de QA-Ku puis 10 CV d'eau MilliPore pour la conservation de la colonne sur une courte période (1 semaine) ou 20 % éthanol sur une plus longue durée. Les fractions correspondantes au(x) pic(s) d'absorbance à 280 nm sont déposées sur gel d'électrophorèse SDS-PAGE 10 %. Les fractions d'intérêt sont rassemblées et leur spectre UV est mesuré (Nanodrop ND-1000). La concentration en Ku des fractions est déduite à partir du coefficient d'extinction molaire de Ku70 et Ku80. Le ratio de l'absorbance à 280 nm sur celle mesurée à 260 nm permet d'évaluer la présence d'ADN résiduel (ratio de 2 pour une protéine pure et de 0,5 pour un ADN pur). La solution protéique est ensuite insérée dans un boudin de dialyse Spectra/Por de 12/14 kDa et dialysée une fois sur la nuit (ON) puis une fois deux heures contre un tampon de congélation (20 mM Tris pH 8,0, 150 mM NaCl, 5 mM β -mercaptoéthanol, 5 % v/v glycérol) à 4 °C.

Lorsque la dialyse est terminée, le spectre UV de la protéine est mesuré pour évaluer la perte éventuelle de petits fragments d'ADN lors de la dialyse. L'hétérodimère Ku est aliquoté par volumes de 1 mL et 200 μ L (pour les expériences nécessitant moins de matériel) à une concentration d'environ 1 mg.mL⁻¹. Les aliquots sont congelés dans de l'azote liquide avant de stocker la protéine à - 80 °C pour une utilisation future.

ETAPES	TAMPON		MOLARITE (mM)		COMPOSE CHIMIQUE	
	Ku70/80	XLF	Ku70/80	XLF	Ku70/80	XLF
<u>Ni-NTA</u>	Lavage = Lyse	Lavage = Lyse	20	20	Tris pH 8.0 KCl NaCl β -mercaptoéthanol Imidazole 1 pastille cocktail inhibiteur de protéases ROCHE EDTA-free	Tris pH 8.0 KCl NaCl β -mercaptoéthanol Imidazole Glycérol ½ pastille cocktail inhibiteur de protéases ROCHE EDTA-free
			150	50		
			850	50		
			10	10		
			40	25		
			10 % v/v			
	Elution	Elution	20	20	Tris pH 8.0 NaCl β -mercaptoéthanol imidazole	Tris pH 8.0 NaCl KCl β -mercaptoéthanol imidazole Glycérol
			150	50		
			10	10		
			10	500		
			300	10 % v/v		
Dialyse n°1	QA-Ku	QA-XLF	20	20	Tris pH 8.0 NaCl KCl β -mercaptoéthanol	Tris pH 8.0 NaCl KCl β -mercaptoéthanol EDTA
			50	50		
			50	50		
			10	10		
				1		
<u>ResourceQ</u>	Lavage QA-Ku	Lavage QA-XLF	20	20	Tris pH 8.0 NaCl KCl β -mercaptoéthanol	Tris pH 8.0 NaCl KCl β -mercaptoéthanol EDTA
			50	50		
			50	50		
			10	10		
				1		
	Elution QB-Ku	Elution QB-XLF	20	20	Tris pH 8.0 NaCl KCl β -mercaptoéthanol	Tris pH 8.0 NaCl KCl β -mercaptoéthanol EDTA
			850	850		
			150	50		
			10	10		
				1		
Dialyse n°2	Congélation	Congélation	20	20	Tris pH 8.0 NaCl β -mercaptoéthanol Glycérol	Tris pH 8.0 NaCl β -mercaptoéthanol Glycérol
			150	150		
			5	5		
			5 % v/v	5 % v/v		

Figure 35 : Tableau récapitulatif de la composition des tampons utilisés pour les purifications de Ku70/80 et XLF

4) Caractérisation des échantillons

- **Electrophorèse sur gel SDS-PAGE**

Afin d'évaluer la pureté des protéines après purification en deux étapes, des micro-volumes d'échantillon à différentes étapes de la purification ont été conservés. Des gels avec 10 % Sodium Dodécyl Sulfate (SDS) sont effectués pour être soumis à une électrophorèse (SDS-PAGE). Ceci permet de séparer par poids moléculaire les différentes protéines contenues dans chaque fraction et de visualiser les éventuels contaminants.

- **Analyse par gel filtration**

Afin de voir plus précisément la pureté des échantillons protéiques, ceux-ci sont passés sur colonne de gel filtration (Superose 6 increase 10/300 GL, GE Healthcare ; 24 mL). Cette analyse est une chromatographie d'exclusion de taille, c'est-à-dire que les protéines les plus grosses sont éluées en premier et les plus petites en dernière position car elles entrent dans les billes de la phase. La colonne est tout d'abord calibrée avec plusieurs marqueurs protéiques de forme globulaire dont les tailles sont indiquées ci-dessous. Un tampon 20 mM Tris pH 8,0, 75

mM NaCl, 75 mM KCl est utilisé. La quantité d'échantillon purifié injectée est de 100 µg pour un volume de 100 µL. L'injection est réalisée après avoir équilibré la colonne avec le tampon cette fois dans le tampon 20 mM Tris pH 8,0, 150 mM NaCl, 5 mM β-mercaptoéthanol. Des fractions de 0,5 mL sont collectées avec un débit de 0,5 mL.min⁻¹.

Les marqueurs de taille utilisés sont :

- Bleu Dextran (M > 2 MDa)
- Thyroglobuline (M = 670 kDa)
- Aldolase (M = 158 kDa)
- Ovalbumine (M = 43 kDa)
- Ribonucléase (M = 13,7 kDa)

C) Préparation des ligands

1) Peptides pXLF et pAPLF

Nous avons étudié les interactions entre Ku70/80 et des peptides contenant un motif de fixation à Ku appelé KBM (pour *Ku Binding Motif* en anglais). Le peptide pXLF correspond au peptide de la protéine XLF et pAPLF au peptide de la protéine APLF. Ces deux peptides contiennent les KBM de XLF et APLF. Ces protéines sont des facteurs du NHEJ qui interviennent respectivement lors de la ligation directe des extrémités de l'ADN ou lors de la maturation des extrémités avant ligation. Ces peptides ont été utilisés lors d'expériences de biochimie, de biologie structurale ou de mesures de biophysiques. Ils ont été synthétisés par voie chimique en phase solide par l'entreprise Genecust. Ils parviennent au laboratoire purifiés et sous forme de poudre. Une analyse par composition en acides aminés, est réalisée au sein du service du JOLIOT/Simopro du CEA pour mesurer la concentration de chacun des peptides resuspendus dans de l'eau. Les peptides sont congelés à - 20 °C par petites fractions.

SEQUENCE XLF (PDBID : 2R9A avec 5His en Cter)

>sp|Q9H9Q4|NHEJ1_HUMAN Non-homologous end-joining factor 1 OS=Homo sapiens
OX=9606 GN=NHEJ1 PE=1 SV=1

281 - 299 ⇔ séquence X-KBM

MEELEQGLLMQPWAWLQLAENSLAKVFITKQGYALLVSDLQQVWHEQVDTSVVVSQRAKELNKRLTAPPAAFLCH
LDNLLRPLLKDAHPSEATFSCDCVADALILRVRSELSGLPFYWNFHCMLASPSLVSQHLIRPLMGMSLALQCQV
RELATLLHMKDLEIQDYQESGATLIRDRLKTEPFENSFLEQFMIEKLPACSIGDGKPFVMNLQDLYMAVTTQE
VQVGQKHQAGADPHTSNSASLQGIDSQCVNQPEQLVSSAPTLSPAPEKESTGTSGPLQRPQLSKVKRKKPRGLFS

SEQUENCE APLF

>sp|Q8IW19|APLF_HUMAN Aprataxin and PNK-like factor OS=Homo sapiens OX=9606
GN=APLF PE=1 SV=1

174 - 192 ⇔ séquence A-KBM

MSGGFELQPRDGGPRVALAPGETVIGRGPLLGITDKRVSRRAHILEVAGGQLRIKPIHTNPCFYQSSEKSQLLLPL
KPNLWCYLNPGDSFSLLDVKYIFRILSIPSEVEMQCTLRNSQVLDEDNINLNETPKSPVINLPHETTASQLEGST
EIAKTQMTPTNSVSVFLGENRDCNKQQPILAEKRKILPTWMLAEHLSDQNLVPAISGGNVIQSGKKEEICKDKSQ
LNTTQQGRRQLISSGSSSENTSAEQDTGEECKNTDQEESTISSKEMPOFSFSAITLSNTEMNNIKTNAQRNKLPIEE
LGKVSXKHKIATKRTPHKEDEAMSCSENCSSAQGDSLQDESQGSSESSNPNPETLHAKATDSVLQSGSEGNKVK
RTSCMYGANCYRKNPVHFQHFHSHPGSDSYGGVQIVGQDETDDRPECPYGPSCYRKNPQHKIEYRHNTLPVRNVLD
EDNDNVGQPNEYDLNDSFLDDEEEDYEPTDEDSWEPGKEDEEKEDVEELLKEAKRFMKRK

2) Petites molécules

A cours de cette thèse, j'ai également étudié l'interaction entre Ku et une petite molécule d'inositol phosphate (IP6). L'IP6 est commandée chez Sigma où elle arrive sous forme de poudre. Son nom commercial est « *phytic acid sodium salt hydrate* ». Cette petite molécule est solubilisée et hautement concentrée à 9 mM dans de l'eau ou du tampon basique (20 mM Tris pH 8,0, 150 mM NaCl). Des molécules proches de l'IP6 sont utilisées également pour évaluer la spécificité de l'interaction. Il s'agit de l'IP3 (3 phosphates), l'IP4 (4 phosphates) et l'IP5 (5 phosphates), commandées également chez Sigma (Figure 36).

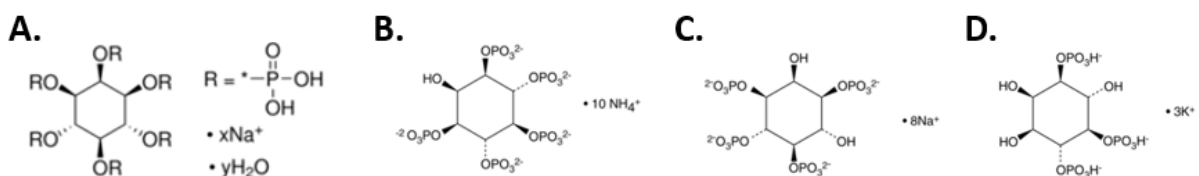


Figure 36 : Représentation chimique des différents inositols utilisés pour l'évaluation de la spécificité d'interaction avec Ku et l'influence du nombre de phosphates. **A.** IP6 ; **B.** IP5 ; **C.** IP4 ; **D.** IP3

3) Oligonucléotides

Les oligonucléotides sont commandés chez Sigma où ils sont synthétisés et purifiés. Ils arrivent sous forme de poudre et d'ADN simple brin au laboratoire où ils sont dissous dans l'eau ou du tampon (20 mM Tris pH 8,0, 150 mM NaCl). La concentration est alors mesurée par absorbance UV à une longueur d'onde de 260 nm. Une étape d'hybridation des oligonucléotides simples brins est alors ensuite réalisée pour les expériences de biochimie, de biologie structurale et de biophysique. Dans cette étude, quatre duplex d'ADN différents sont testés : l'ADN *hairpin*

(noté hDNA et composé d'un oligonucléotide simple brin de 34 pb et un de 21 pb), l'ADN 18 pb, l'ADN 15 pb et l'ADN 42 pb. Chaque oligonucléotide simple brin est hybridé avec son complémentaire.

Pour cette étape d'hybridation, la concentration des ADN simple brin à hybrider et resuspendus dans l'eau est estimée par mesure d'absorbance à 260 nm sur un spectrophotomètre UV (Nanodrop ND-100). La concentration de chaque ADN simple brin est calculée grâce à leur epsilon molaire théorique. Les deux ADN simple brin complémentaires sont mélangés dans un ratio molaire 1 : 1. Après mélange des deux simples brins, l'échantillon est chauffé à 95 °C pendant 5 min sous agitation dans un bécher d'1 L au bain-marie. L'agitation et le chauffage sont arrêtés et l'ADN hybridé est alors laissé toute une nuit à refroidir jusqu'à température ambiante. Ceci permet un refroidissement très lent de 95 °C à 20 °C. Une mesure d'absorbance est alors effectuée pour vérifier la concentration de l'ADN hybridé. L'ADN est stocké à – 20 °C.

D) Interactions Ku-ADN et ligands

1) Mesures d'interactions par microcalorimétrie

La microcalorimétrie est une méthode de mesure des interactions des molécules qui suit les micro-dégagements et absorption de chaleur consécutifs à une interaction entre une molécule A et une molécule B. La microcalorimétrie requiert un tampon strictement identique pour le ligand (injecté dans la cellule de mesure via la seringue d'injection) et le récepteur (se trouvant la cellule de mesure). Ce même tampon servira de référence dans la cellule de référence. Il est préférable de travailler en absence de glycérol. La protéine d'intérêt est stockée en fin de purification dans du tampon contenant 5 % de glycérol pour sa congélation. Pour s'en acquitter, celle-ci est dialysée une fois toute une nuit puis deux heures à 4 °C dans le tampon suivant : 20 mM Tris pH 8,0, 150 mM NaCl, 5 mM β -mercaptoéthanol. Il est possible, si nécessaire, de faire une dialyse trois fois 2 heures contre ce tampon.

Cette méthode de mesure est assez consommatrice en matériel notamment pour des K_D dans la gamme du micromolaire. L'appareil VP-ITC est utilisée pour ces expériences. La concentration du récepteur qui est mis dans la cellule de mesure, doit être 10 à 50 fois supérieure à la valeur du K_D attendue et le volume de récepteur utilisé est de 2,2 mL. C'est ce qu'on appelle la règle du facteur c de Wieseman (Figure 38). Ce paramètre critique permet de déterminer la forme de

l'isotherme, en fonction du K_A attendu et la concentration en récepteur mis dans la cellule de mesure ($[M_{Tot}]_0$) et la stœchiométrie de l'interaction. Ce coefficient c correspond au produit de la constante K d'association, de la concentration totale du récepteur dans la cellule en début d'expérience et du paramètre n de stœchiométrie :

$$c = K_A \times [M_{Tot}]_0 \times N$$

Le ligand est généralement préparé à une concentration 10 fois plus élevée que le récepteur et un volume de 600 μL est chargé dans la seringue de l'appareil. On considère premièrement l'expérience suivante : IP6 (petite molécule) vs KuCC où l'on attend un K_D autour de 50 nM. Après dialyse, la protéine Ku (considérée comme le récepteur) est concentrée à 1 μM (correspondant à 0,13 $\text{mg}\cdot\text{mL}^{-1}$) sur un concentrateur Amicon® Ultra-4 30 kDa. Nous utilisons pour une mesure, un volume de 2,2 mL (soit 0,3 mg) afin d'être légèrement en excès pour remplir complètement la cellule de mesure (volume de 1,4 mL). L'IP6 (ligand) est alors préparé à 10 μM dans le tampon de dialyse pour un volume de 600 μL afin d'être aspiré dans la seringue d'injection. Avant d'être insérés dans la cellule de mesure et dans la seringue d'injection, les deux échantillons sont dégazés et pré-équilibrés pendant environ 10 min à une température proche de la température de mesure souhaitée (ici 25 °C) grâce à un système ThermoVac™ (Malvern). Après dégazage des échantillons, l'IP6 est aspiré dans la seringue d'injection. Le port de remplissage est ensuite fermé et le système est purgé deux fois de suite. Pendant ce temps, Ku est insérée dans la cellule de mesure via une seringue (Hamilton™ seringue Gastight™ série 1000).

On considère dans un deuxième temps l'étude des interactions entre Ku ou Ku-ADN et les peptides pXLF et pAPLF. Ces mesures ont été faites afin de déterminer l'affinité entre l'hétérodimère et ces deux peptides dérivés respectivement des facteurs XLF et APLF du NHEJ. La protéine est dialysée contre le même tampon (20 mM Tris pH 8,0, 150 mM NaCl, 5 mM β -mercaptoéthanol). Les peptides et l'ADN tige-boucle (appelé hDNA pour hairpin DNA) sont préparés à de fortes concentrations de stockage autour du milli molaire car ils sont très solubles. Après dialyse, la protéine Ku (considérée comme le récepteur) est concentrée de 2 à 20 μM (de 0,26 à 2,6 $\text{mg}\cdot\text{mL}^{-1}$) sur un concentrateur Amicon® Ultra-4 30 kDa. Les mesures avec une concentration de 20 μM de Ku consomment 5,2 mg d'hétérodimère par courbe, limitant le nombre de mesures que l'on peut réaliser à cette concentration. Les peptides (considérés comme les ligands) sont alors préparés de 20 à 200 μM dans le tampon de dialyse pour un volume de 600 μL afin d'être aspirés dans la seringue d'injection.

Pour chaque étude (interaction Ku-IP6 ou Ku-peptides), les mesures sont faites avec dans la cellule de mesure Ku seule ou en complexe avec l'ADN tige-boucle.

Quand la purge est terminée, la seringue d'injection est alors insérée dans la cellule de mesure et l'expérience de microcalorimétrie peut alors démarrer après avoir défini les paramètres de la mesure. Nous effectuons 30 injections de 10 μL avec un retour à l'équilibre en 240 secondes.

Paramètres expérimentaux dans « ITC controls »

- Nombre total d'injections : 30
- Température de la cellule ($^{\circ}\text{C}$) : 25
- Puissance de référence ($\mu\text{Cal/sec}$) : 10
- Délai initial (sec) : 120
- Vitesse d'agitation : 310

Paramètres d'injection dans « ITC controls »

- Volume (μL) : 10 (2 pour la première injection)
- Durée (sec) : 20 (4 pour la première injection)
- Espacement (sec) : 240
- Période de filtrage (sec) : 10
-

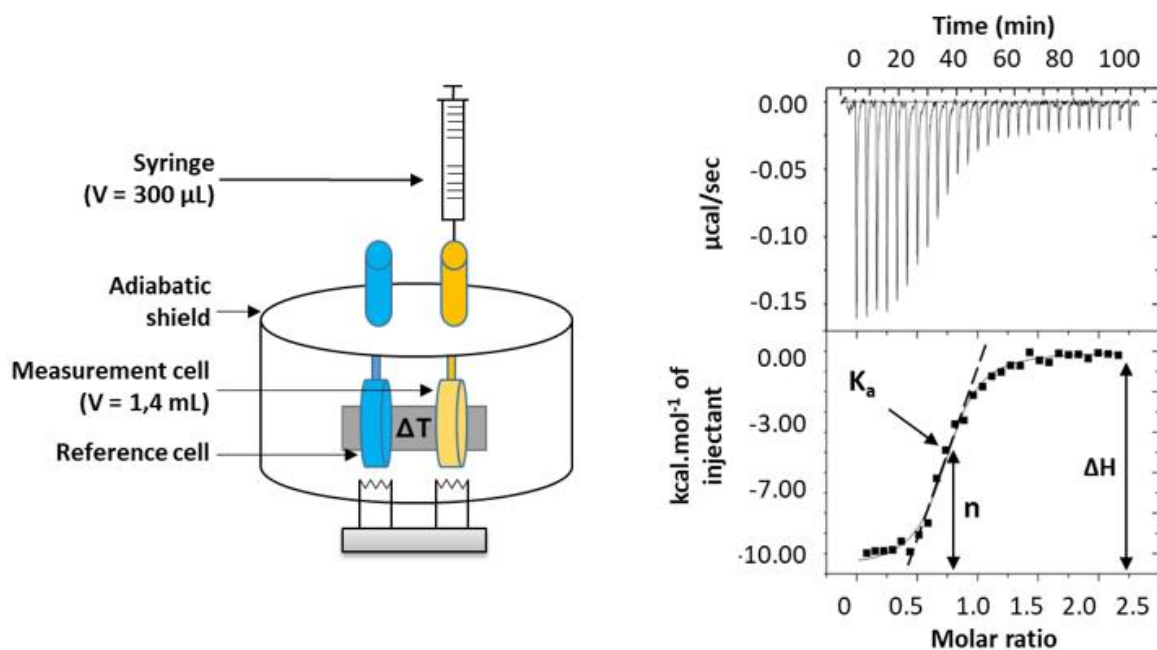


Figure 37 : A. Schéma décrivant le dispositif de mesure de microcalorimétrie du VP-ITC. L'appareil se compose d'une chambre où se situent à la fois une cellule de référence (dans laquelle est disposé le tampon dans lequel est mesurée l'interaction) et d'une cellule de mesure (dans laquelle est disposé le récepteur). Le ligand est injecté à intervalles de temps réguliers dans la cellule de mesure grâce à la seringue d'injection. B. Thermogramme (haut) et isotherme de titration (bas). L'isotherme permet de caractériser la stœchiométrie du complexe (N), la variation d'enthalpie (ΔH) et la constante d'association à l'équilibre (K_a) (Gontier et al, 2018)

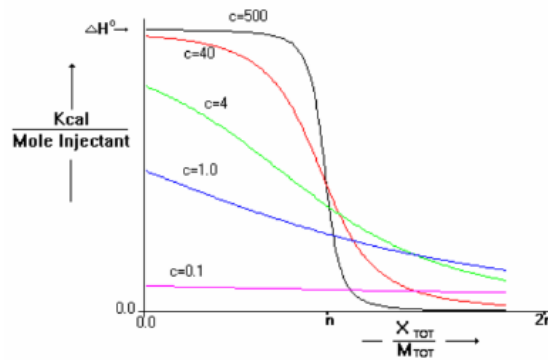


Figure 38 : Evolution de la forme des isothermes de titration pour des valeurs c différentes. La valeur c est définie par la formule $c = n K_A [MTot]$ (n : stœchiométrie, $K_A = [ML] / [M][L]$ à l'équilibre, $[MTot]$: concentration totale du récepteur dans la cellule en début d'expérience)

Durant toute l'expérience, la même quantité de ligand est injectée. Au départ, la plupart des récepteurs sont libres, les pics ont une même intensité correspondant à l'interaction de toutes les molécules de ligand avec le récepteur. A la fin de titration, le ligand est en excès, il n'y a plus de récepteurs libres. Les pics résiduels correspondent à la dilution du ligand dans le tampon. Entre les deux, les pics diminuent progressivement avec un point d'inflexion quand la moitié des récepteurs est saturée. L'analyse des résultats de microcalorimétrie est faite à l'aide du logiciel Origin 7.0 (Malvern). L'analyse du thermogramme consiste à intégrer les aires sous chacun des pics. Les variations d'enthalpie $\Delta H^\circ / \text{kcal.mol}^{-1}$ sont déduites à chaque injection. Ces valeurs sont reportées en fonction du ratio molaire ligand/récepteur. L'isotherme de titration permet de caractériser la variation d'enthalpie de la réaction ($\Delta H / \text{kcal.mol}^{-1}$), la stœchiométrie du complexe étudié ($N / \text{unité absolue}$), la constante thermodynamique d'association ($K_A / \text{L.mol}^{-1}$) (Figure 37). Grâce à une formule liant plusieurs paramètres thermodynamiques, il est possible de retrouver la variation d'enthalpie libre du système étudié ($\Delta G^\circ / \text{kcal.mol}^{-1}$) ou encore la variation d'entropie ($\Delta S^\circ / \text{kcal.mol}^{-1}.\text{deg}^{-1}$).

La formule est la suivante :

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ \quad (1)$$

$$\Delta G^\circ = -RT \ln(K_a) \quad (2)$$

$$K_a = 1/K_d$$

$$\text{d'où } \Delta G^\circ = RT \ln(K_d) \quad (3)$$

T (Kelvin) \Leftrightarrow température

R ($8.314 \text{ J.mol}^{-1}.\text{K}^{-1}$ à $P = 1 \text{ bar}$) \Leftrightarrow constante des gaz parfaits

2) Mesure d'interactions par nanoleviers d'ADN (technologie du switchSENSE)

Le switchSENSE est une méthode de mesure des interactions des molécules qui permet notamment de suivre la vitesse de basculement (*switch*) de nanoleviers d'ADN marqués avec

une sonde de fluorescence. Le principe de cette méthode a été décrite dans la partie « Introduction ». Toutes les expériences de switchSENSE sont réalisées sur un DRX 2400. Cet instrument appelé DRX² permet de suivre la fluorescence de deux fluorophores différents rouge et vert en parallèle. Les expériences sont réalisées à l'aide d'un circuit micro fluïdique sur une puce (*chip*). La puce utilisée ici est notamment la MPC2-48-2-G1R1-S possédant sur chaque emplacement (spot) d'environ 1 mm² des leviers d'ADN simple brin auxquels sont attachés covalamment un fluorophore rouge et d'autres leviers d'ADN fonctionnalisés avec un fluorophore vert et des électrodes de mesure. Un nombre proche de brins d'ADN contenant des fluorophores rouges et verts sont greffés. La protéine Ku d'intérêt étant dans un tampon classique (20 mM Tris pH 8,0, 150 mM NaCl, 5 % glycérol, 5 mM β -mercaptoéthanol) suite à l'étape de purification, celle-ci est tout d'abord préparée dans un tampon semblable sans glycérol et sans agent réducteur : le tampon TE140 DBS (Dynamic Biosensors GmbH ; Planegg, Allemagne). Ce tampon est constitué de 10 mM Tris pH 7,4, 140 mM NaCl, 5 % Tween 20, 50 μ M EGTA, 50 μ M EDTA. Pour ce faire, Ku est successivement diluée et/ou concentrée trois fois sur concentrateur Amicon Ultra 4, 30 K préalablement lavé avec le tampon. Des mesures avec une gamme de concentration en Ku ont permis de montrer qu'une concentration de 250nM est suffisante pour mesurer une cinétique. Un ADN simple brin de 48 bases, complémentaire de l'ADN simple brin greffé sur la puce (appelé NLB48) et couplé au fluorophore est injecté pour former un ADN double brin de 48 pb. Cet ADN complémentaire est appelé cNLB48. Ku est alors injecté et son association sur l'ADN double brin et sa dissociation sont suivies respectivement pendant 1 min et 6 min. Le débit optimal utilisé pour cette expérience est de 200 μ L.min⁻¹ pour l'association et la dissociation. Une gamme de concentrations en Ku est réalisée (allant de 2 μ M à 62,5 nM) avec un facteur de dilution 2 entre chaque concentration. Afin d'assurer la dissociation complète du complexe Ku-ADN, une étape de régénération est réalisée entre chaque concentration permettant de revenir à l'état initial de simple brin sur la puce. Les résultats sont analysés avec le logiciel switchANALYSIS.

3) *High-Throughput Differential Scanning Fluorimetry (HT-DSF)*

Le HT-DSF est une méthode qui permet notamment d'effectuer des cribles d'interactions entre des banques de petites molécules et une protéine donnée. La méthode mesure la température de dénaturation de la protéine (T_m pour *melting Temperature* en anglais) (Figure 39). Celle-ci est généralement influencée par la fixation d'un ligand avec des variations fréquemment supérieures à 5 °C en cas de fixation. La technique de *Thermal Shift Assay* encore appelée TSA

est utilisée pour identifier dans une banque de petites molécules de Sanofi des ligands de la protéine KuCC. Un crible haut-débit utilisant l'émission de fluorescence est effectué ; la fluorescence provenant d'un fluorophore (SYPRO orange) venant se fixer sur la partie hydrophobe de la protéine lorsque celle-ci se dénature (Figure 39). Les conditions optimales de tampon, la concentration de la protéine et la concentration optimale du fluorophore sont déterminées en premier lieu. On procède ensuite par étapes afin d'identifier notre sélection de ligands potentiellement intéressants.

- **Faisabilité**

La première étape de cette étude est appelée « faisabilité ». La protéine est préparée à trois concentrations différentes : 0,25, 0,35 et 0,45 mg.mL⁻¹ avec à chaque fois trois concentrations différentes en SYPRO orange X5, X10 et X15 dans le tampon suivant : 20 mM Tris pH 8,0, 150 mM NaCl. Un total de neuf conditions différentes sont testées. La plaque est mise dans un thermocycleur à PCR (CFX384 Touch Real-Time PCR Detection system). Les échantillons sont chauffés de 20 °C à 90 °C suivant un gradient de 0,5 °C par minute.

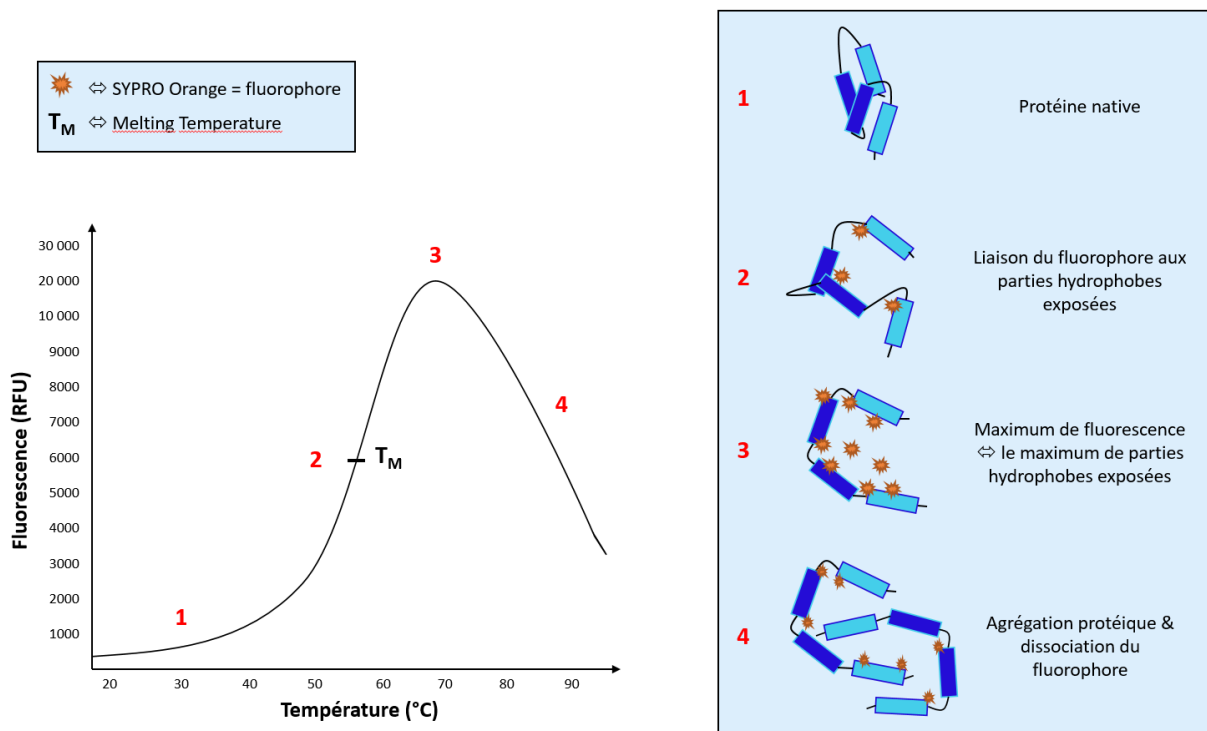


Figure 39 : Profil de fluorescence par méthode de *Differential Scanning Fluorimetry* avec le SYPRO Orange. A gauche : la courbe de stabilité protéique (fluorescence en fonction de la température). La température de dénaturation T_m est définie comme le point d'inflexion de la première transition. A droite : schéma des différents états de la protéine en fonction de la température. La diminution de la fluorescence à haute température après dénaturation est liée à l'agrégation de la protéine dénaturée

- **Fragments testés à une seule concentration**

Suite à cette étape de faisabilité, un crible de fragments d'une collection Sanofi est réalisé sur la protéine KuCC. Cette collection de fragments a été caractérisée par RMN auparavant et est composée de 987 composés dont certains sont des composants dessinés par Sanofi et dont la compagnie a la propriété intellectuelle. Ces fragments sont testés à 1 mM sur la protéine KuCC à 0,45 mg/mL (correspondant à 3,4 μ M) dans le tampon classique de biophysique (noté TB par la suite) 20 mM Tris pH 8,0, 150 mM NaCl. Des mesures contrôles sont effectuées avec l'anhydrase carbonique II (CAII) qui sert de contrôle négatif et qui n'interagit pas avec les fragments et une solution composée uniquement de tampon. Dans chacune des trois solutions (KuCC, CAII et tampon), le fluorophore SYPRO orange est utilisé à une concentration X15. Les conditions de concentrations expérimentales pour la protéine KuCC et pour le SYPRO orange ont été choisies lors de l'étape de faisabilité (voir partie résultats). Les fragments étant dans 100 % DMSO au départ, on effectue un témoin : KuCC + 1 % DMSO, correspondant à la concentration finale en DMSO lorsque Ku est mis en présence de ligands. Etant donné le nombre de fragments et de solutions à préparer, deux robots sont utilisés pour réaliser ces expériences. Un premier robot automate (TECAN, Fluent®), distribue les trois solutions préparées et disposées sur une plaque source de 96 puits à fond rond (Dutscher) vers la plaque DSF 384 puits à fond rond (Dutscher). Pour une question de simplicité de lecture de plaque lors de l'analyse, plusieurs colonnes sont réalisées : une colonne « blanc » (tampon), une colonne « CAII » et des colonnes « KuCC ». Un deuxième robot automate (MOSQUITO, ttplabtech) pipette chaque fragment à 100 mM au départ pour les disposer sur la plaque DSF contenant déjà la protéine KuCC. Le volume pipeté est de 100 nL pour un volume final sur la plaque DSF de 10 μ L. Le DMSO est déposé également dans les colonnes « blanc », « CAII » et « KuCC seule ». En effet, une colonne protéine seule a été réalisée avec le TECAN lors de l'étape précédente afin de visualiser le comportement de la protéine en présence de 1 % DMSO.

- **Fragments testés sur une gamme de concentrations : doses-réponses**

Suite au criblage de la protéine avec les fragments à une concentration de 1 mM, les fragments présentant l'effet le plus important sont sélectionnés pour être testés sur une gamme de concentrations. Grâce à cette gamme de concentrations et aux mesures effectuées, la concentration des fragments stabilisant la protéine est déterminée. Un premier classement de ces fragments est établi en terme de force d'interaction potentielle. Un nombre de 49 composés

est sélectionné. Ces fragments sont alors étudiés dans une gamme de concentrations allant de 2500 μM à 4,9 μM avec un facteur de dilution de 2. Pour ce faire, le MOSQUITO est utilisé afin de réaliser les dilutions sérielles des 49 fragments RMN sur une première plaque DSF. Suite aux dilutions sérielles, la même stratégie est utilisée que précédemment quant à l'utilisation du TECAN. Celui-ci va permettre de déposer de la plaque source 96 puits préalablement préparée des trois solutions « tampon », « CAII » et « KuCC » vers la plaque DSF finale n°2. Une dernière étape permet enfin de déposer la gamme de concentrations en fragments sur la protéine KuCC, de la plaque DSF n°1 vers la plaque DSF n°2 (finale) grâce au MOSQUITO.

4) Mesures d'interactions et de compétitions par RMN

La RMN est une méthode robuste pour étudier qualitativement et quantitativement les interactions protéine-ligand à l'échelle atomique et pour caractériser de nouvelles molécules bioactives. Cette caractérisation peut se faire soit à partir de techniques RMN où l'on observe les résonances associées au ligand (LO-NMR) ou à partir de techniques RMN où l'on observe les résonances associées à la protéine (PO-NMR). Les techniques basées sur l'observation du ligand peuvent être divisées en deux catégories. La première catégorie permet de suivre directement des changements de paramètres RMN de résonance associées au ligand tels que le temps de relaxation, le déplacement chimique et la largeur des raies. La seconde catégorie permet l'observation de la liaison du ligand indirectement de par l'effet NOE. Cette catégorie d'expérience tire profit de l'échange rapide entre le ligand libre et le ligand lié et démontre une grande sensibilité tout en utilisant peu de protéine non-marquée. Ces techniques ont cependant besoin d'expériences complémentaires pour identifier le site de liaison du ligand à la protéine telles les expériences de compétition avec des ligands dont le site de liaison est connu et/ou des expériences de mutagenèse dirigée de la protéine étudiée.

Le Water-LOGSY (pour *Water-Ligand Observation with Gradient Spectroscopy* en anglais) (Dalvit 2009) et le STD (pour *Saturation Transfer Difference* en anglais) (Mayer et Meyer, 2001) sont les deux types d'expériences les plus utilisées basées sur l'observation du ligand. Elles utilisent toutes les deux les transferts d'effets NOEs (transfert de magnétisation) entre les protons de la protéine et ceux du ligand transitoirement lié à la protéine.

Dans une première étape, deux tampons deutérés sont préparés et filtrés sur 0,22 μm : 10 % et 100 % D_2O (20 mM Tris-d11 pH 8,0, 150 mM NaCl Suprapur, 1 mM TCEP-d16)

respectivement pour les expériences de water-LOGSY et de STD. Le tampon de la protéine Ku est alors changé en tampon 100 % deutéré en concentrant / diluant plusieurs fois la protéine sur concentrateur Amicon Ultra-15, cut-off de 50 kDa, 5000 g, 5 min à 4 °C.

En pratique, pour réaliser des expériences de Water-LOGSY et de STD, des tests de contrôle qualité (QC pour *Quality Control* en anglais) doivent être effectués pour vérifier la pureté du composé, sa solubilité, sa compatibilité chimique, sa concentration et son état d'agrégation. Ces contrôles sont faits à la fois sur le proton et le fluor (si les molécules étudiées possèdent un ou plusieurs atomes de fluor). Dans notre étude, on ne parle que de protons mais on utilise cependant la sonde fluor pour vérifier si Ku interagit avec elle. Les concentrations des différents composés sont mesurées grâce à la présence de deux sondes ajoutées dans le tampon (une référence proton : le DSS qui permet également de calibrer les spectres et une sonde fluor : REF¹⁹F (Figure 40)) dont les concentrations sont précisément connues. Pour gagner en sensibilité dans cette phase, on travaille avec des tubes RMN d'un diamètre de 5 mm, pouvant contenir 600 µL de solution. Ces tubes ont une taille compatible avec le passeur d'échantillons SampleJet de l'aimant 600 MHz (Bruker, SampleJet). Dans cette phase préliminaire de QC, différents tubes sont préparés. Les tubes 1 et 2 contiennent les tampons 10 % et 100 % D₂O ; le tube 3, 1 µM Ku seule ; le tube 4, 1 mM IP6 seule ; les tubes 5 et 6, 1 mM composé 2 (RA14917148A) (Figure 41) seul et 1 mM composé 3 (EX00045853) (Figure 41) seul. Tous les tubes contiennent le tampon 10 % D₂O, 100 µM DSS + 500 µM ¹⁹F. Lors des expériences de QC, nous nous sommes aperçus que la sonde REF¹⁹F interagissait avec la protéine Ku. Pour la suite des expériences RMN, les tampons 10 % et 100 % D₂O utilisés sont préparés avec 100 µM DSS uniquement.

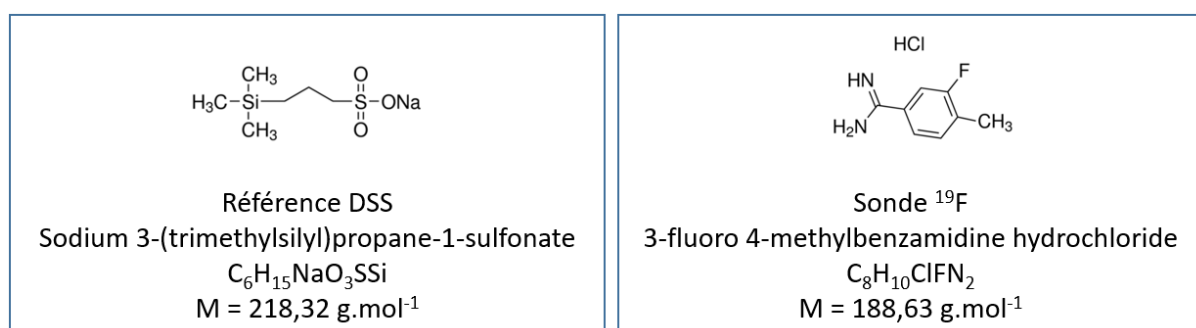


Figure 40 : Formules chimiques et caractéristiques de la référence DSS (gauche) et de la sonde ¹⁹F (droite)

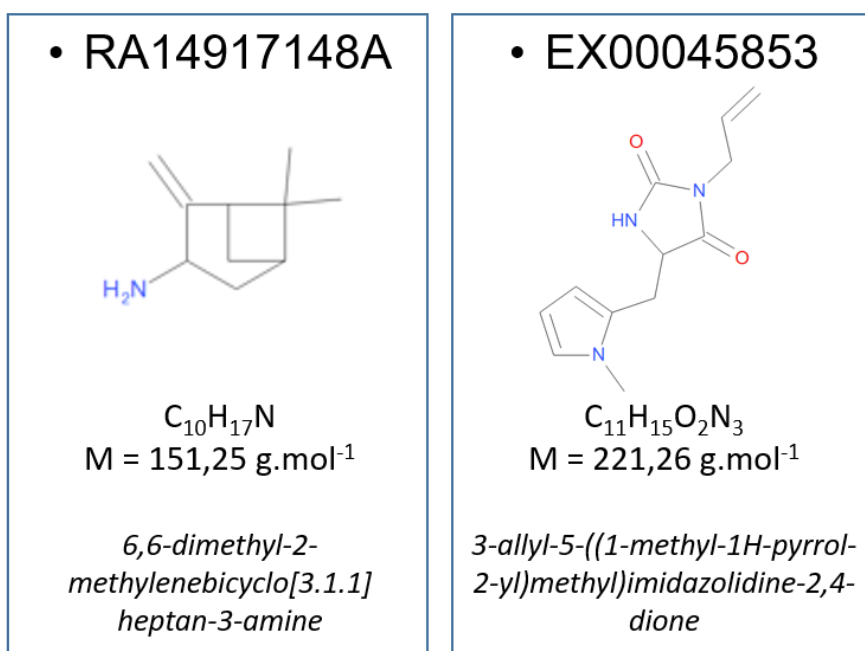


Figure 41 : Structures et données chimiques des deux fragments publics retenus après étude de HT-DSF sur KuCC

Pour les expériences de WL et STD qui suivent, deux tampons deutérés ((20 mM Tris-d11 pH 8,0, 150 mM NaCl Suprapur, 1 mM TCEP-d16) sont d'abord préparés et filtrés sur 0,22 µm : un premier tampon contient 10 % D₂O (WL) et l'autre tampon contient 100 % D₂O (STD). Le tampon de la protéine Ku est ensuite changé en tampon 100 % deutéré en concentrant et diluant plusieurs fois la protéine sur concentrateur Amicon Ultra-15, cut-off de 50 kDa, 5000 g, 5 min à 4 °C.

Des tests d'interactions entre protéine, ADN tige-boucle (*hairpin*) et composés (IP6, composé RA14917148A et composé EX00045853) sont ensuite réalisés grâce au WL et au STD suivis de tests de compétitions :

- Protéine, RA14917148A et l'IP6
- Protéine, EX00045853 et l'IP6
- Protéine, RA14917148A (+ hDNA) et IP6
- Protéine, EX00045853 (+hDNA) et IP6

La protéine est concentrée à 4 ou 6 µM, les composés à 500 µM ou 1 mM, l'ADN à 10 ou 20 µM et l'IP6 à 150 ou 300 µM. Lors d'une expérience de WL, les spectres protons et « WL » sont enregistrés avec un nombre de scans (NS) respectifs de 512 et 256. Lors d'une expérience de STD, les spectres protons et STD sont tous deux enregistrés avec un nombre de scans de 512. De plus, pour ces expériences de WL et STD, les échantillons d'un volume de 200 µL sont préparés dans des tubes RMN de diamètre 3 mm et de longueur 103,5 mm.

Expériences Water-LOGSY :

Ce sont les protons de l'eau qui sont excités et la magnétisation (inversée) est transférée des protons de l'eau transitoirement liée à la protéine vers les protons du ligand lié à la protéine via différents mécanismes.

On travaille dans un tampon 10 % deutéré. Le transfert d'aimantation s'effectue via l'eau à l'interface entre le récepteur et le ligand. L'interaction sera détectée grâce à ces molécules d'eau. On applique alors la saturation sur la fréquence de résonance de l'eau. Dans le schéma ci-dessous, le ligand (*binder*) est en rouge (Figure 42). Le premier spectre ^1H visualisé en WL montre les ligands à l'état libre, le deuxième spectre ^1H visualisé en WL montre les ligands à l'état lié avec la protéine cible. On observe suite à l'irradiation de la protéine, que le ligand rouge est un ligand par le fait que le signal proton négatif du ligand rouge à l'état libre du départ, devient positif lorsque le ligand est à l'état lié.

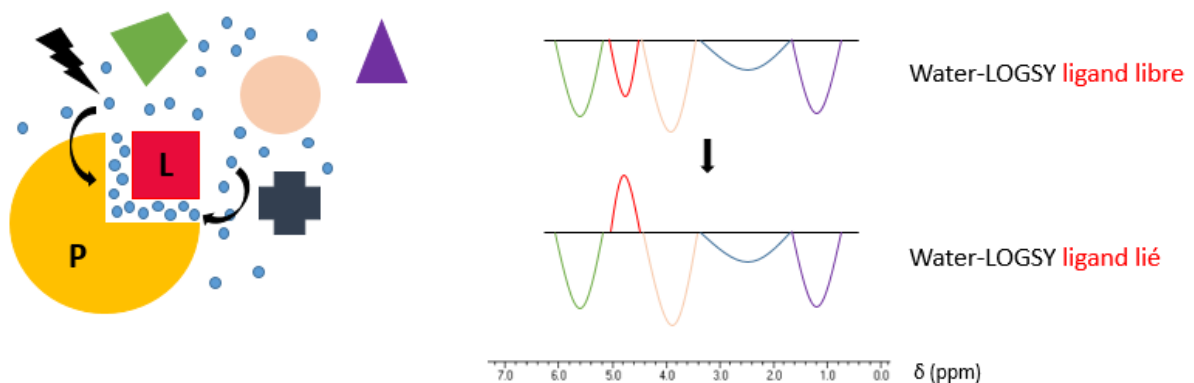


Figure 42 : Schéma décrivant une expérience RMN Water-LOGSY

Expériences de STD :

Dans le cas du STD, une région du spectre proton de la protéine est saturée (« on » résonance) de manière sélective. Cette région contient peu de résonance liée à la protéine et aucune résonance liée au ligand (en général on cible la région aliphatique située entre 0 et 1 ppm). Cette saturation sélective est ensuite transférée à la protéine dans son entier par diffusion de spin (NOEs, intramoléculaires) mais également vers le ligand. Une deuxième expérience dans laquelle une région éloignée de la fenêtre spectrale de la protéine est sélectivement saturée (« off » résonance). La différence entre les expériences « on » et « off » résonance permet d'identifier le transfert de magnétisation des protons de la protéine vers ceux du ligand.

On travaille dans un tampon cette fois 100 % deutéré car les molécules d'eau ne sont plus utiles pour visualiser le transfert d'aimantation ; les protons sont donc tous deutérés. On irradie de manière assez longue (2 sec) les protons du récepteur protéique notamment les groupements méthyles $-CH_3$. La diffusion de spins va permettre le transfert d'irradiation du récepteur protéique vers le ligand. Lors d'une expérience STD, le regard est porté sur le ligand mais l'on conserve tout de même le signal de la protéine. Cette expérience permet de savoir si le ligand est un ligand de la protéine. Si c'est le cas, nous aurons accès au site de fixation de ce ligand et éventuellement à son orientation. Sur les schémas ci-dessous, l'on observe trois spectres 1H différents (Figure 43). Le premier spectre montre la résonance « off » du ligand (on n'irradie ni sur la protéine ni sur les ligands), c'est un « blanc » servant de témoin négatif. Le deuxième spectre montre la résonance « on » du ligand (on irradie la protéine mais on n'irradie pas les ligands). Le dernier spectre 1H est la différence du spectre « off » moins le spectre « on ». Lorsqu'il y a fixation du ligand sur la protéine, alors le signal 1H du ligand en résonance « on » va diminuer et l'on observe un pic d'intensité beaucoup plus faible pour le spectre « différence » : c'est ce qu'on observe pour le signal rouge.

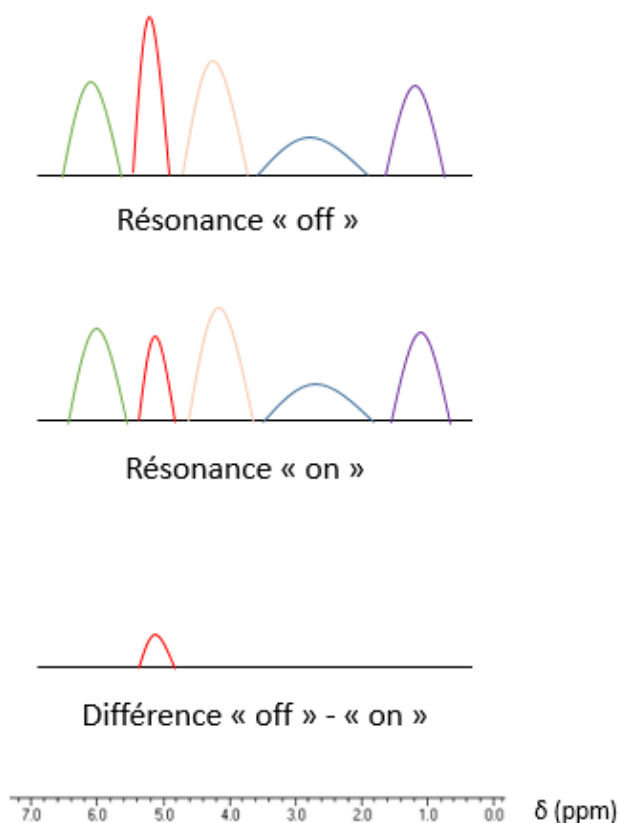


Figure 43 : Schéma décrivant une expérience de RMN-STD

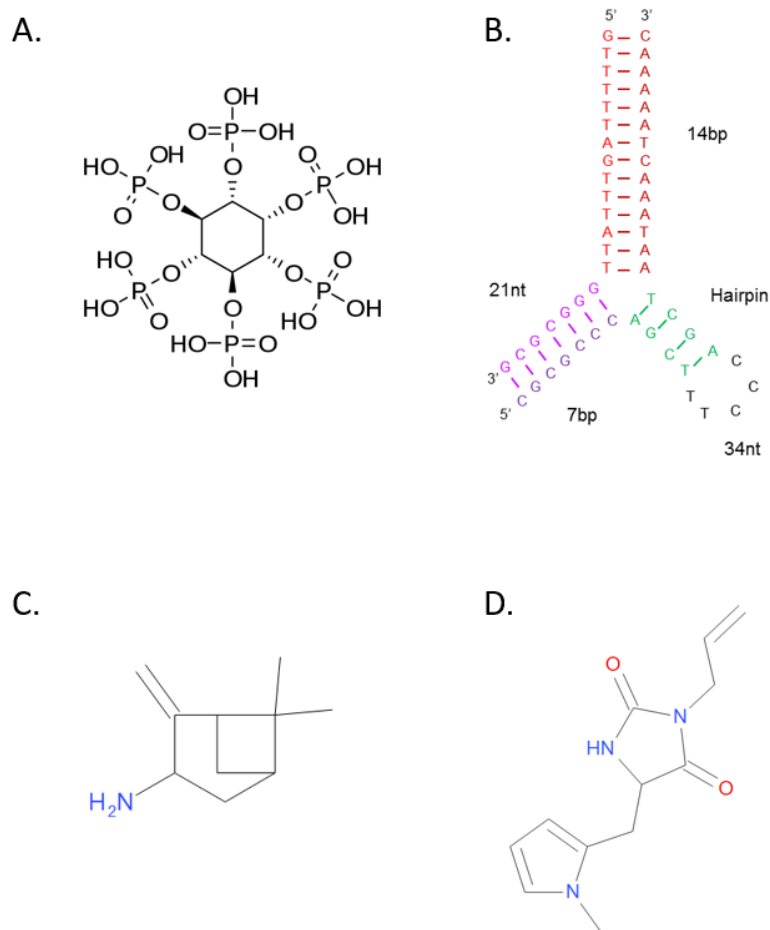


Figure 44 : Formules chimiques des différents composés utilisés lors des expériences de water-LOGSY et STD.
A. Inositol hexakisphosphate (IP6). **B.** ADN hairpin. **C.** Composé 2 (RA14917148A). **D.** Composé 3 (EX00045853)

E) Cristallographie des protéines

Dans le but de localiser la petite molécule d'inositol phosphate au sein du complexe Ku-ADN, j'ai mené une étude cristallographique. Celle-ci permet d'étudier par diffraction de rayons X la structure de complexes à l'échelle atomique (Figure 45). La protéine choisie pour la cristallisation est la construction réduite KuCC car ses deux parties C-terminales de Ku70 et Ku80 sont supprimées ; la flexibilité est alors réduite ce qui favorise la cristallisation. Quant à l'ADN, nous décidons de travailler avec un ADN tige-boucle (noté par la suite hDNA pour *hairpin*) permettant d'orienter Ku. Un ratio protéine : ADN de 1 : 1 est utilisé. Chaque protéine se conduit de manière différente lors des expériences de cristallisation selon le pH, la température, la concentration en sels, le précipitant utilisé, la présence ou non d'agent réducteur et/ou d'additifs. Pour ce faire, des échantillons de ce complexe KuCC-hDNA-IP6 ont été envoyés dans un premier temps sur la plateforme de cristallisation HTX (EMBL, Grenoble)

permettant un crible d'une grande quantité de conditions différentes. Cette plateforme de cristallisation haut-débit à laquelle ont été envoyés les échantillons, est située à Grenoble à l'EMBL (J Marquèz ; <https://htxlab.embl.fr/>).

Les kits commerciaux utilisés pour ce crible sont les suivants :

- Classics Suite : produits chimiques les plus populaires en cristallographie des protéines (Jancarik & Kim)
- JCSG+ : kit de cristallographie combinant conditions classiques et modernes incluant PEG, gamme en sels et solvants organiques (Newmann et al, 2005)
- PACT premier : combinaisons de différentes tailles de PEG et sels ou tailles de PEG et pH (Newmann et al, 2005)
- PEGs-I : solutions de PEGs à pH4 contre des sels de basse concentration
- Salt Grid : conditions chaotropes modifiant la structure tridimensionnelle des macromolécules biologiques
- Wizard I&II : conditions aléatoires de cristallisation reprenant un mix des conditions précédentes

Le complexe est préparé à trois concentrations différentes en KuCC : 150 μL d'échantillon préparé à 5 $\text{mg}\cdot\text{mL}^{-1}$, 8 $\text{mg}\cdot\text{mL}^{-1}$ et 12 $\text{mg}\cdot\text{mL}^{-1}$. De plus, afin de s'assurer que l'ADN et l'IP6 soient bien en complexe avec la protéine Ku, on choisit de mettre l'hDNA et l'IP6 légèrement en excès (1,2 fois plus concentrés par rapport à la concentration de KuCC en molarité). Le robot permet d'établir les tests de cristallisation pour chaque kit (cités ci-dessus). Il va donc réaliser pour chaque complexe KuCC-hDNA-IP6 et pour chaque kit de cristallisation 96 conditions différentes. Il faut savoir que la technique utilisée sur la plateforme HTX est la diffusion de vapeur en goutte assise (Figure 46A).

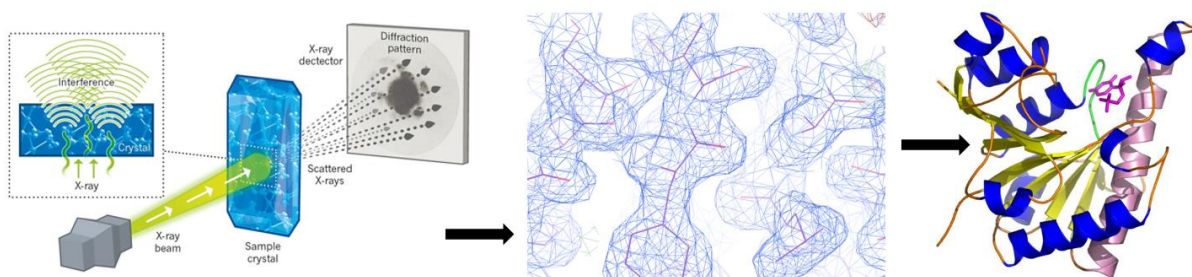


Figure 45 : Différentes étapes de la cristallographie des macromolécules. A gauche : le faisceau de rayons X venant frapper le cristal et génère des tâches de diffraction. Au milieu : un exemple de carte de densité électronique après collection des images de diffraction et détermination du problème des phases. A droite : structure macromoléculaire résolue

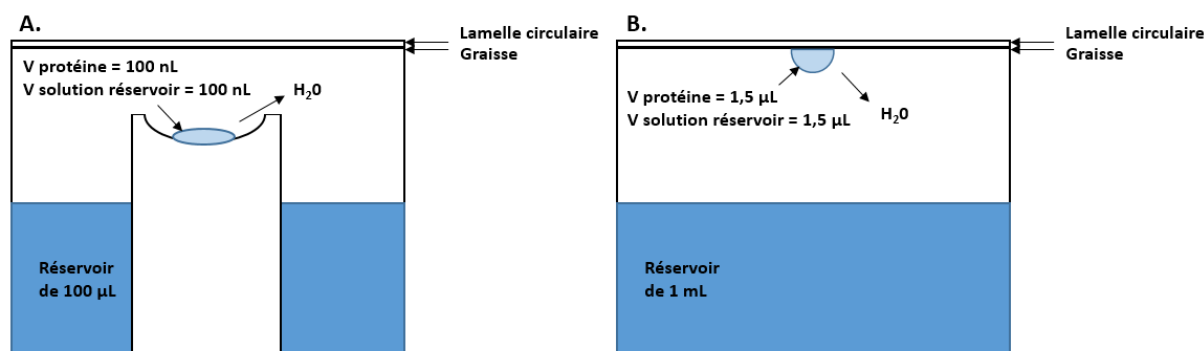


Figure 46 : Schéma expérimental des techniques de diffusion vapeur en goutte assise sur la plateforme de cristallisation HTX (A) et en goutte suspendue au laboratoire (B). Les gouttes réalisées par le robot HTX ont un volume final de 200 nL et celles faites au laboratoire un volume final de 3 µL pour permettre une croissance plus importante du cristal mais pour faciliter sa récupération et sa congélation

Lors d'une expérience de cristallisation par diffusion de vapeur, les composés volatiles peuvent alors diffuser entre la goutte et la solution du réservoir. Les gouttes sont faites à 20 °C et composées d'un volume équivalent ($V = 100$ nL) en échantillon et en solution du réservoir pour obtenir un volume final de 200 nL. Les plaques 96 puits sont alors recouvertes d'un adhésif afin de fermer hermétiquement chaque compartiment. Les boîtes de cristallisation sont alors déplacées et stockées à 4 °C et un autre robot prend la relève pour l'étape de visualisation des gouttes. Il est alors possible d'observer en direct et à distance les boîtes grâce au robot « Formulatrix Rockimager » et à un serveur HTX développé par la plateforme aux temps J+1, J+3, J+7, J+15, J+33, J+61 et J+87. Les photos des différents essais de cristallisation permettent de définir les états possibles allant de la goutte claire avec de la protéine soluble à la présence de précipités protéiques en passant par la présence de cristaux dans la goutte. En cas de succès, les pistes de cristallisation les plus prometteuses sont identifiées et reproduites au laboratoire. Afin de faciliter la manipulation des cristaux et d'optimiser leur taille, les volumes mélangés en complexe et solution du réservoir sont de 1,5 µL. L'apparition des premiers cristaux varie pour ce projet. En effet, d'une condition à une autre le temps varie de 1 jour à 15 jours. Les cristaux de protéines se forment dans une zone du diagramme de solubilité qui se situe entre les zones métastables et de précipitation de la protéine. Les premiers germes cristallins apparaissent dans cette région et croissent (Figure 47). Différents essais de cristallisation sont réalisés au laboratoire pour optimiser les conditions de cristallisation et obtenir des cristaux de tailles suffisantes, monocristallins et sans défauts. Pour cela, des essais sont effectués en variant notamment la concentration en agent précipitant, la concentration en protéine, la température, le pH et la présence d'additifs. A l'issue de cette étape, les cristaux sont trempés 15 sec dans

une solution cryoprotectante. Celle-ci permet la protection des cristaux contre la glace lors de leur congélation. Les solutions cryoprotectantes sont généralement constituées de la solution de cristallisation utilisée pour réaliser le cristal à laquelle est ajoutée un agent cryogénique. Il est primordial de tester plusieurs cryoprotectants pour maximiser les chances de diffraction du complexe et affiner la résolution. Dans mon projet, il m'est arrivé d'obtenir plusieurs cristaux avec une solution de cristallisation constituée de 24 % de PEG 3350 (pour *Polyethylene Glycol* en anglais) et dans ce cas le cristal n'a pas besoin obligatoirement d'agent cryoprotecteur supplémentaire. En effet, lorsque la concentration en PEG dépasse 21 %, le cristal peut être capturé dans une goutte de solution mère et congelé directement dans l'azote liquide. J'ai testé sans et avec différents agents cryoprotectants tels que 7 %, 10 %, 15 % et 20 % v/v glycérol ou éthylène glycol, 10 % et 20 % v/v sucrose). Après trempage rapide dans la solution cryoprotectante, le cristal est plongé très rapidement dans l'azote liquide pour être congelé. Celui-ci est peut alors être conservé dans un conteneur rempli d'azote liquide jusqu'au moment de la collecte de données.

Criblage de conditions de cristallisation Sanofi sur Ku pour une optimisation des conditions

Les techniques d'ensemencement et de coupure à la trypsine ont été utilisées pour réaliser un criblage sur des collections de conditions de cristallisation Sanofi. Premièrement, des stocks de cristaux du complexe KuCC-hDNA-IP6 ayant diffracté ont été réalisés. La technique consiste à broyer des cristaux du complexe grâce à des billes (25 % PEG 3350, 100 mM BTP pH 8,5, 150 mM Na₂SO₄) pour les stocker à 4 °C pour une utilisation directe ou future. Un total de quatre criblages a été effectué avec des conditions Sanofi. Le complexe KuCC-hDNA-IP6 est formé de nouveau à une concentration de 10 mg.mL⁻¹ en protéine KuCC. Pour quatre plaques de 96 conditions chacune, un volume de 80 µL de complexe a été mélangé à un volume de trypsine. Le ratio trypsine : complexe est de 1 : 500 en concentration molaire. Pour chaque plaque de conditions de cristallisation, trois gouttes de 0,2 µL sont réalisées :

- $V_{\text{Solution de cristallisation}} = 0,1 \mu\text{L}$ et $V_{\text{complexe seul}} = 0,1 \mu\text{L}$
- $V_{\text{Solution de cristallisation}} = 0,1 \mu\text{L}$, $V_{\text{complexe seul et semence du complexe (seed)}} = 0,1 \mu\text{L}$
- $V_{\text{Solution de cristallisation}} = 0,1 \mu\text{L}$, $V_{\text{complexe seul coupé à la trypsine}} = 0,1 \mu\text{L}$

La visualisation des gouttes est réalisée à 20 °C chez Sanofi grâce au robot Formulatrix.

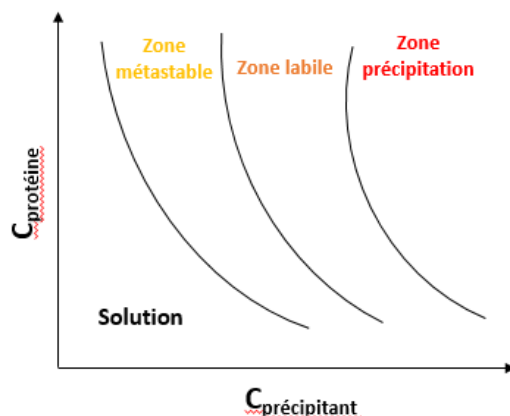


Figure 47 : Diagramme de phase d'une protéine : courbes de transition des différents états de solubilité de la protéine. Courbes représentées sur une représentation avec la concentration en protéine en ordonnée et la concentration en agent précipitant en abscisse

Des clichés de diffraction aux cartes de densité électronique :

Les cristaux sont ensuite passés sur deux lignes du synchrotron SOLEIL situé à Gif-Sur-Yvette : « Proxima-1 » et « Proxima-2A ». Ces deux lignes sont toutes deux utilisées pour étudier les complexes macromoléculaires biologiques. La ligne PX1 est une ligne adaptée aux cristaux de grande taille car le faisceau est plus large (principalement $40 \times 20 \mu\text{m}^2$). Ces cristaux sont moins sensibles aux dommages liés aux rayons X. La ligne PX2A possède un faisceau plus étroit et est davantage adapté à l'étude de petits cristaux. Des tests ont été réalisés sur les deux lignes de lumière car Ku cristallise sous forme de bâtonnets avec une section de petite dimension ($10 \mu\text{m} \times 20 \mu\text{m}$) et une troisième dimension plus longue ($100 \mu\text{m}$). Les cristaux sont placés dans un réceptacle (*puk*) compatible avec les robots qui positionnent chaque cristal sur le goniomètre grâce à un serveur de commande à distance depuis la cabine de pilotage. Les cristaux sont alors positionnés sous un flux d'azote à une température de 100 K et centrés sur la position faisceau de rayons X. Le flux de rayons X traverse ensuite le cristal pour obtenir des clichés de diffraction du complexe macromoléculaire protéique. Les détecteurs utilisés sont pour PX1 et PX2A respectivement les détecteurs PILATUS 6M (Dectris Ltd.) et EIGER X 9M (Dectris Ltd.) qui permettent d'enregistrer un ensemble d'images avec les tâches de diffraction en faisant tourner le cristal autour d'un axe de rotation (par exemple 3 600 images avec une oscillation de $0,1^\circ$). Après l'obtention de premiers clichés de diffraction, les tâches de diffraction de ces premières images sont intégrées grâce au logiciel XDS (Kabsch, 2010). Un script « xdsme » a été écrit par Pierre Legrand ; scientifique sur la ligne de lumière PX1 (<https://github.com/legrandp/xdsme/blob/master/PKG-INFO>). Grâce à ce script, plusieurs informations sont caractérisées :

- La maille cristalline
- Le groupe d'espace de la maille cristalline
- La résolution du jeu de données
- Facteurs de structure et intensité des pics

Ces informations nous permettent de collecter un jeu de données complet qui tient compte notamment de la symétrie du cristal, des paramètres de maille et de son pouvoir de diffraction.

Phasage et affinement :

Le phasage permet de calculer à partir des jeux de données de diffraction les premières cartes de densité électroniques. L'étape de diffraction permet de mesurer uniquement les index des tâches ainsi que leurs intensités. L'information qui est contenue dans la phase de l'onde diffractée est perdue. Afin de connaître les phases des facteurs de structure du cristal, plusieurs techniques sont possibles. Nous pouvons réaliser un dérivé du cristal contenant des atomes lourds. Cette technique permet de faire varier les modules de facteurs de structures en incorporant un ou plusieurs atome(s) lourd(s) : on parle alors de remplacement isomorphe simple ou multiple (respectivement SIR/MIR) ou utilisation de signal anomal simple ou multiple (respectivement SAD/MAD). Dans ces cas, ces atomes sont positionnés spécifiquement dans l'unité asymétrique.. Il est possible également de réaliser le phasage par remplacement moléculaire. Celle-ci permet grâce à une structure déjà résolue de déterminer la nouvelle structure cristallographique. C'est cette seconde technique qui a été utilisée dans le projet Ku-ADN-IP6 en utilisant le modèle 1JEY de la PDB (Walker, Corpina, et Goldberg 2001).

Une première carte de densité est tout d'abord calculée par données expérimentales. Le calcul se fait par la formule suivante :

$$P(\mathbf{u}) = \frac{1}{V} \sum_{\mathbf{S}} |F(\mathbf{S})|^2 \cdot \cos[2\pi \mathbf{u} \cdot \mathbf{S}]$$

Ce calcul est identique à celui de la transformée de Fourier mais uniquement sur les intensités des facteurs de structure. Grâce à celui-ci, nous sommes capables de tracer une carte 3D spécifique représentant la maille cristalline. Un modèle 3D du complexe macromoléculaire est ensuite édité avec la formule du coefficient de Matthews :

$$Vm = \frac{V}{nxM}$$

Avec V_m le coefficient de Matthews (entre 1,68 et 3,53 Å³), V le volume de la maille cristalline, M le poids moléculaire de la macromolécule et n le nombre de molécules de la maille

Une carte de Patterson est alors créée et celle-ci est mise en relation et comparée avec la carte des données expérimentales. Le modèle est parfait seulement si les deux cartes sont superposables. Il faut procéder généralement étape par étape en appliquant des fonctions de rotation / translation pour obtenir un modèle parfait. Lorsque celui-ci est parfait, alors nous pouvons calculer les phases de notre cristal.

Le modèle issu du remplacement moléculaire est alors confronté aux données expérimentales pour voir si des corrections doivent être faites. Pour ce faire, nous avons utilisé le logiciel BUSTER (Bricogne et al, 2016). C'est l'affinement du modèle.

A partir de notre modèle, ce logiciel calcule les facteurs de structure calculés (notés F_{calc}) et les compare aux facteurs de structure observés (notés F_{obs}). Ces derniers sont dérivés des données expérimentales. Lorsque le calcul est correct, les valeurs de R_{work} et R_{free} se stabilisent. On dit que le modèle est concordant avec les données expérimentales. Le logiciel calcule alors deux cartes de densités électroniques :

- La carte $2F_{obs}-F_{calc}$: elle correspond notamment à la densité électronique autour du modèle
- La carte $F_{obs}-F_{calc}$: elle correspond notamment à des atomes non encore inclus dans le modèle (pics positifs) et à des atomes incorrectement positionnés (pics négatifs)

Le but est ensuite d'améliorer le modèle. La première stratégie utilisée est la modification de l'emplacement des résidus (protéiques ou nucléiques) dans l'espace réel. Si un pic se trouve dans un pic négatif et qu'un pic positif se trouve à proximité, alors le résidu est déplacé dans la densité positive. Les contraintes de stéréochimie doivent alors être respectées. Si une densité positive est perçue à un endroit où ne se trouve aucun résidu, alors c'est que l'on doit placer un objet supplémentaire dans le cristal. Cet objet doit donc être déterminé. C'est ici le cas de la densité de la petite molécule d'IP6. La dernière étape est d'affiner la structure de manière encore plus pointue grâce aux tenseurs TLS (pour *Translation, Libration and Screw* en anglais). Ils permettent par des opérations mathématiques de caractériser l'état de désordre et les différentes

vibrations moléculaires de segments indépendants dans le cristal. Cette fonction est notamment pratique dans les cas de résolution moyenne ou basse, elle permet d'obtenir un modèle décrivant au mieux les données expérimentales. Les valeurs R_{work} et R_{free} doivent diminuer avec l'affinement du modèle. Pour ce faire, cette fonction d'affinement est répétée jusqu'à obtenir cette condition. Ainsi ils permettent quantitativement de savoir l'état de ressemblance entre le modèle et les données expérimentales. Plus ce modèle est en accord, plus ces valeurs calculées diminuent. On note que pour le calcul du R_{free} , seulement 5 % du jeu de données complet est utilisé. Si cette valeur de R_{free} devient inférieure à la valeur du R_{work} , alors c'est que le modèle d'étude est très certainement biaisé et qu'il y a eu un sur affinement. Elles sont ensuite comparées avec les valeurs obtenues dans la PDB pour pouvoir valider notre structure cristallographique.

PARTIE II :

Interactions protéine-ADN en *drug discovery*

L'objectif majeur du projet de mesures des interactions sur le projet de Sanofi était d'évaluer l'apport du switchSENSE par rapport à d'autres mesures biophysiques effectuées dans l'entreprise à différentes étapes de la recherche de candidats médicaments pour inhiber une interaction protéine-ADN. Ma thèse a ainsi porté sur un complexe protéine-ADN et des inhibiteurs de cette interaction. Pour des raisons de confidentialité, nous appellerons les différentes protéines ou constructions de protéine étudiées (protéines A, B, C et D), les différents ADN (ADN 1 et ADN 2).

A) Mesures d'interactions par la technologie du switchSENSE

1) Oligonucléotides, puces et modes de mesure utilisés

Pour mesurer les interactions entre les protéines cibles A, B, C et D et les ADN 1 et 2, nous avons choisi d'utiliser un ADN complémentaire au simple brin présent sur la puce qui présente en 3' de cette séquence complémentaire la séquence de l'ADN qui nous intéresse (Figure 48). Nous utilisons alors un ADN simple brin complémentaire de la séquence d'intérêt pour reconstituer un nanolevier avec notre séquence cible d'ADN à l'extrémité du nanolevier classique du switchSENSE (Figure 48).

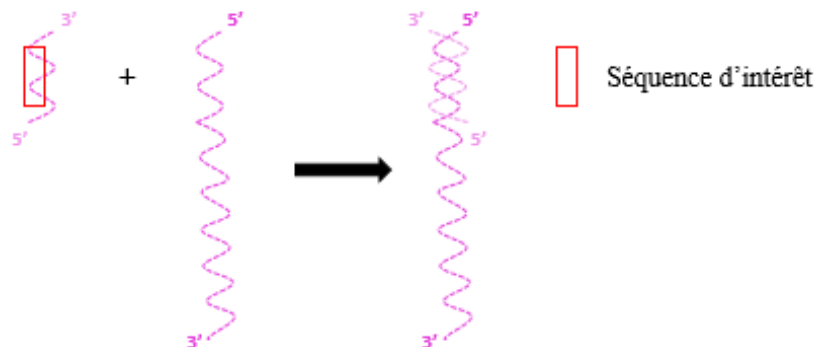


Figure 48 : Représentation schématique de l'étape de pré-hybridation de la séquence d'ADN d'intérêt pour la liaison de la protéine X (projet Sanofi) avant hybridation de la puce

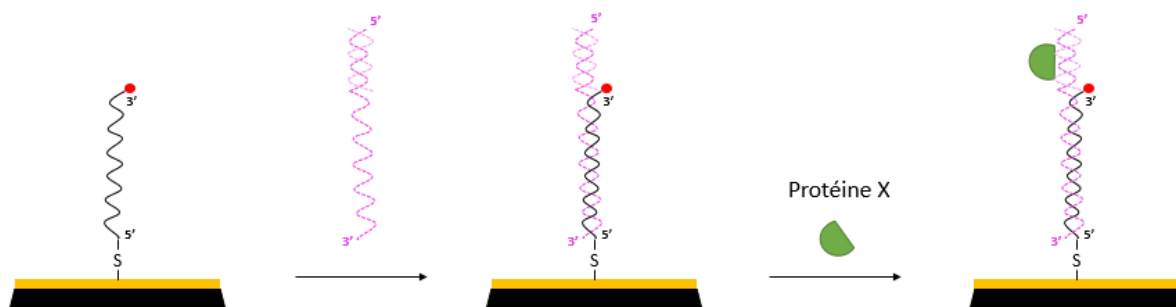
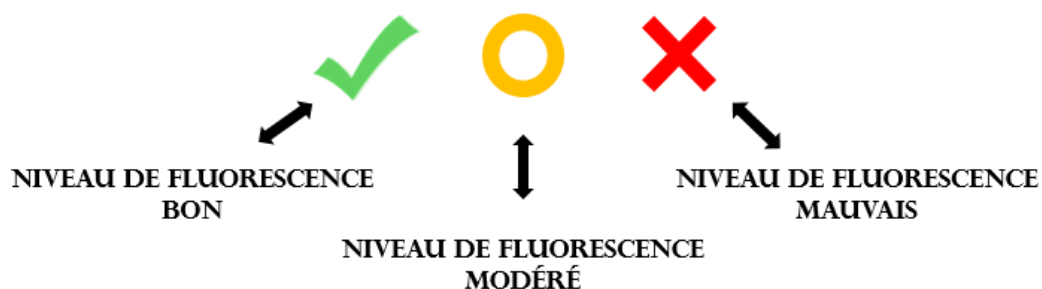


Figure 49 : Fonctionnalisation de la puce MPC-48-2-R1-S avec un ADN Y pré-hybridé présentant la séquence d'intérêt pour la liaison de la protéine X

2) Mise en place de la puce

Afin de vérifier la fonctionnalité d'une puce, il est indispensable de vérifier l'état des électrodes de chaque canal. Chaque puce provenant du réfrigérateur est déposée à température ambiante au minimum 30 min avant son utilisation. La température de la puce augmente alors de manière très progressive. Grâce au logiciel switchCONTROL (voir ses autres fonctionnalités ci-après), la puce est ensuite insérée dans le tiroir dédié et l'on choisit le canal à tester. Un statut du niveau de fluorescence de chaque électrode du canal est effectué. Trois logos permettent de déterminer à première vue l'état du niveau de fluorescence.



On estime qu'une électrode est en bon état quand sa fluorescence « up » est supérieure à 100 kcps. La fluorescence « up » correspond à la fluorescence observée quand le potentiel sur la puce est négatif et que les nanoleviers sont relevés (Figure 50). Il faut également que son amplitude de fluorescence soit supérieure à 40 kcps. L'amplitude de fluorescence correspond à la différence de fluorescence entre les deux positions des nanoleviers, basses et hautes, pour respectivement des potentiels positifs et négatifs sur la puce. On notera que pour des

applications spéciales telles que des expériences à faible densité, ces valeurs peuvent être inférieures au détriment de la stabilité du signal et de l'augmentation du bruit.

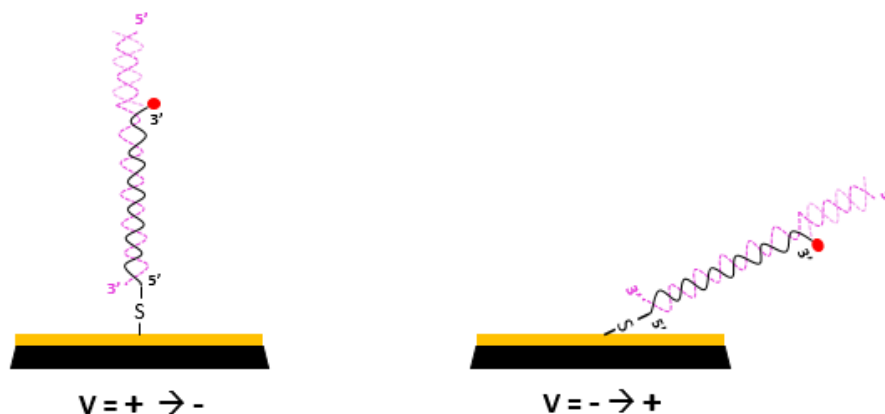


Figure 50 : Représentation schématique de l'état de position sur la puce des nanoleviers d'ADN. A gauche : fluorescence « up » quand potentiel électrique négatif (répulsion) ; à droite : fluorescence « down » quand potentiel électrique positif (attraction)

3) Les logiciels du switchSENSE

Une expérience de switchSENSE se réalise en trois temps. Un premier logiciel de paramétrage de l'expérience en elle-même (**switchBUILD**), un deuxième logiciel permettant la lecture du script détaillé de l'expérience, de sa conversion en diverses commandes et de visualisation en temps réel (**switchCONTROL**), un troisième logiciel d'analyse des données (**switchANALYSIS**). Ci-dessous, est détaillé chaque logiciel avec sa version à la date à laquelle ce mémoire est rédigé. De plus, je décris ci-dessous les expériences de cinétiques, titrations et mesures de tailles (*sizing*) pour une protéine (protéine B) et un ADN (ADN 1). Ils servent d'exemples modèles pour les autres tests réalisés durant cette thèse.

- Etape 1 : **switchBUILD** ⇔ construire et designer

Version : switchBUILD 1.12.0.21



Grâce à ce premier logiciel, les expériences de cinétiques, de mesure de tailles et de titration sont construites avec des paramètres qui ont à chaque fois été optimisés pas à pas. Le mode « cinétique » est choisi ici. Nous choisissons une puce de type MPC-48-2-R1-S. Ces puces présentent deux électrodes (E1 et E2) de référence possédant uniquement un ADN simple brin avec la séquence A greffée sur la puce et quatre autres électrodes (E3, E4, E5 et E6) possédant uniquement la séquence B greffée sur la puce.

Nous pouvons ainsi vérifier facilement la spécificité d'interaction de chaque construction de notre protéine pour les deux ADN d'intérêt (ADN 1 et ADN 2). Une étape de passivation est ajoutée. La passivation est une solution permettant de combler les espaces vides de 50 nm entre chaque levier d'ADN présent sur la puce et éviter qu'une protéine ou ligand vienne interagir de façon non-spécifique.

L'immobilisation choisie est l'hybridation du conjugué et le mode statique. Le ligand correspond à l'ADN complémentaire de NLB48 et présentant la séquence ADN 1 à une extrémité (il est noté cNLB48-ADN1). L'analyte correspond à la protéine B de poids moléculaire 11 kDa. Après plusieurs essais préliminaires et une concentration de 1 μ M en protéine B testée au départ, un premier K_D est estimé à 550 nM. On choisit de lancer une gamme de concentrations en protéine B allant de 3 μ M à 23.4 nM avec un facteur de dilution 2. La première phase d'association est lancée sur 40 sec et la phase de dissociation sur 1 min. Deux options sont possibles : réaliser une expérience en *one-shot* et dans ce cas la protéine est dissociée de l'ADN d'intérêt directement après la phase d'association. Cette option est conseillée lorsque la dissociation est très rapide. La seconde option la plus classique est d'inclure un *one-shot* pour qu'il n'y ait pas d'arrêt entre la phase d'association et la phase de dissociation. C'est l'option qui a été la plus concluante pour cette construction protéique. Pour la gamme de concentrations, nous décidons de régénérer le canal entre chaque phase « association-dissociation ». L'oligonucléotide complémentaire cNLB48-ADN1 contenant la séquence ADN 1 toute protéine B, non encore dissociée, sont ainsi dissociés.

Lorsque la cinétique est paramétrée, une étape de *standby* est insérée. Celle-ci consiste à régénérer la puce et hybrider les ADN complémentaires des NLA48 et NLB48. En effet, il est important lors du stockage que les oligonucléotides présents sur la puce soient sous forme double brin. L'expérience est alors paramétrée et peut être enregistrée par date dans le dossier de l'utilisateur.

- Etape 2 : **switchCONTROL** \Leftrightarrow conversion des scripts et lecture

Version : switchCONTROL 1.6.20



Après avoir créé un premier fichier switchBUILD, celui-ci est ouvert dans le logiciel switchCONTROL. Ainsi, le fichier de commandes de la cinétique peut être ouvert et lu dans les détails. Ce schéma est expliqué dans le chapitre « introduction ». Grâce à ce logiciel, il est possible de personnaliser davantage l'expérience souhaitée en rééditant les commandes (modification des temps ou des volumes). Après avoir

lancé l'expérience, la fluorescence émise est visible durant toutes les étapes grâce à l'onglet « Plots ».

- Etape 3 : **switchANALYSIS** ⇔ analyse

Version : switchANALYSIS 1.9.0.31



Il est possible d'ouvrir les données brutes lorsque la cinétique est terminée. Pour s'assurer que les résultats que l'on observe soient corrects, une première étape primordiale est de vérifier que l'hybridation de l'ADN 1 *overhang* est bien réalisée. Si l'état d'hybridation est correct, alors il est possible de visualiser les données de cinétique. Etant donné, nous travaillons avec une puce ne possédant que le fluorophore rouge, nous choisissons de filtrer les données en ne conservant que les données du fluorophore rouge. Ceci permet d'alléger le nombre de données et de faciliter la lecture des résultats. Un ajustement des données est appliqué afin d'obtenir les paramètres cinétiques (k_{on} / k_{off}) et d'affinité (K_D) de l'expérience.

B) Caractérisation des interactions protéine – ADN

1) Informations sur les échantillons protéiques

L'étude porte sur un projet de recherches de candidats médicaments en oncologie et sur l'inhibition d'une interaction protéine – ADN. Les expériences de switchSENSE sont réalisées sur quatre protéines différentes : trois constructions d'une même cible et une construction d'une protéine de la même famille mais de taille plus importante. Les échantillons proviennent de quatre entreprises différentes (Figure 51).

Protéine « X »	MW (Da)	ϵ ($\text{cm}^{-1} \cdot \text{M}^{-1}$)	Tampon protéique de stockage	Provenance échantillon
Protéine A	8 900	7450	50 mM Tris pH 7.5, 150 mM NaCl, 5 % glycérol, 2 mM DTT	Selvita
Protéine B	11 000	9120	20 mM Tris pH 8.0, 150 mM NaCl, 1 mM TCEP	Selvita
Protéine C	15 090	12 990	20 mM Tris pH 8.0, 150 mM NaCl, 1 mM TCEP	Proteos
Protéine D	71 057	64 750	20 mM HEPES pH 7.4, 100 mM NaCl, 20 % glycérol, 2 mM DTT	Sanofi

Figure 51 : Tableau décrivant les caractéristiques des différentes protéines. Un code couleur est appliqué

2) Informations sur les oligonucléotides utilisés

Afin de réaliser les expériences de switchSENSE, quatre ADN simples brins sont commandés chez IDT (pour Integrated DNA Technologies). A ces ADN simples brins (Tableau ci-dessous), est ajouté l'ADN présent sur la puce utilisée de type MPC-48-2-R1-S.

ADN « Y »	MW (Da)	ϵ (cm-1.M-1)	Nom donné à l'ADN	Provenance échantillon
ADN chip 48 pb	14 886,62	474 700	NLB48	DBS
ADN1	20 163,02	641 800	ADN1-cNLB	IDT
ADN1	5 542,6	164 300	c-ADN1	IDT
ADN2	21 486,81	661 300	ADN2-cNLB	IDT
ADN2	6 690,39	223 700	c-ADN2	IDT

Figure 52 : Tableau décrivant les caractéristiques des différents ADN. Un code couleur est appliqué

Le NLB48 présente 48 pb et est accroché covalamment sur la puce de mesure par son extrémité 5'. A l'extrémité 3' de cet ADN, se trouve le fluorophore rouge. Les deux ADN 1 et 2 doivent ainsi être hybridés afin de préformer un ADN à bout débordant (noté ADN *overhang* par la suite). Cette étape d'hybridation est appelée étape de pré-hybridation car elle précède l'étape d'hybridation de l'ADN *overhang* sur la puce. Les stocks initiaux de DNA1-cNLB/c-DNA1 et DNA2-cNLB/c-DNA2 à 100 mM sont dilués à 1 mM avec du tampon TE4 (10 mM Tris-HCl pH 7,4, 40 mM NaCl, 0,05 % Tween20, 50 μ M EDTA, 50 μ M EGTA). La concentration finale souhaitée de l'ADN *overhang* à hybrider sur la puce est de 500 nM. Pour chaque ADN à pré-hybrider (ADN 1 et ADN 2), des volumes de 500 μ L de séquence d'intérêt et de séquence complémentaire sont mélangés au vortex et chauffés à 85 ° C, 700 rpm pendant deux minutes (ThermoMixer, eppendorf). Les oligonucléotides pré-hybridés sont alors déposés à température ambiante ON pour ensuite les stocker à 4 °C. L'absorbance est mesurée à 260 nm (NanoDrop 2000, Thermo Fischer).

C) Inhibition des interactions protéine – ADN

Cette étude doit permettre d'évaluer le potentiel du switchSENSE en tant que nouvelle méthode biophysique pour la validation de candidats médicaments. J'ai ainsi étudié les interactions

protéine – ADN des protéines (A, B, C et D) et des ADN (1 et 2) par switchSENSE. L'objectif était de mesurer des paramètres cinétiques de ces interactions, de comparer ces mesures aux mesures effectuées par d'autres approches et d'évaluer les avantages et inconvénients de cette méthode par rapport à d'autres techniques utilisées par Sanofi (DSF / ITC / MST / SPR). Mon étude a porté ensuite sur la caractérisation de différents inhibiteurs potentiels identifiés par Sanofi contre ces interactions protéine – ADN. L'objectif était d'identifier les petites molécules capables d'inhiber les interactions protéine – ADN ciblées en contrôlant que ces inhibiteurs n'interagissent pas directement avec l'ADN. La méthode du switchSENSE préconise d'avoir des tampons identiques pour le tampon d'équilibration du canal (*running buffer*) et pour le tampon protéique. J'ai choisi de travailler avec une concentration finale de 4 % DMSO et d'avoir une protéine concentrée à 1 μM et réfléchir à une stratégie possible pour réunir l'intégralité des conditions suivantes :

- Avoir 4 % DMSO dans le tampon d'équilibration
- Avoir 4 % DMSO dans le tampon protéique contenant le ligand
- Que la protéine soit concentrée à 1 μM
- Que la gamme de concentration en ligand soit correcte

1) *Mise en place d'une stratégie de caractérisation*

Une quinzaine de petites molécules fournies par Sanofi IDD m'ont été délivrées en début d'étude pour une analyse en switchSENSE. L'équipe de Sanofi a effectué des mesures en amont sur plusieurs techniques connues de biophysique. Pour certaines molécules, il n'y avait que très peu de résultats et certains inhibiteurs ont été abandonnés en cours d'étude. Le but était ici de tester le pouvoir des inhibiteurs sur les huit couples protéine – ADN. Des premiers tests de cinétique ont permis de voir si chaque inhibiteur avait une interaction spécifique avec la protéine en question et s'il n'interagissait pas avec l'ADN seul.

2) *Titrations*

Après avoir réalisé des premiers tests cinétiques pour caractériser l'inhibition de l'interaction protéine – ADN, j'ai effectué des titrations du ligand. L'affinité du ligand pour la protéine est alors calculée avec une concentration en protéine fixée à 1 μM et une gamme de concentrations en ligand.

RESULTATS

Mes travaux de thèse sont centrés sur une caractérisation des interactions biomoléculaires par différentes méthodes de biophysique et de biologie structurale. Il s'agit notamment d'une étude entre la protéine Ku et plusieurs partenaires du NHEJ mais aussi d'une étude pour un projet d'oncologie sur lequel j'ai travaillé pour le service « *Integrated Drug Discovery* » de Sanofi en utilisant une technologie nouvelle : le switchSENSE. Ces méthodes *in vitro* nécessitent notamment des protéines recombinantes et des expériences biochimiques.

PARTIE I :

L'IP6 : une molécule stimulatrice du NHEJ

A) Expression et purification des facteurs du NHEJ

1) Expression des protéines Ku et XLF en cellules d'insecte

La construction des vecteurs des protéines Ku et XLF a été utilisée pour produire les protéines recombinantes Ku et XLF dans un système d'expression en cellules d'insectes (vecteurs MultiBac fournis grâce à une collaboration avec l'équipe du Dr Imre Berger EMBL Grenoble, actuellement à l'Université de Bristol). J'avais donc en ma possession des plasmides codant pour les protéines entières Ku, pour les versions tronquées de Ku70/80 appelées KuCC (Ku70 1-544 et Ku80 1-551) et KuSC (Ku70 24-544 et Ku80 1-551) ainsi que pour la protéine XLF.

La construction des protéines contient un marqueur 10 His en Nter du Ku80 et de XLF pour faciliter la purification de celles-ci à l'aide d'une résine Ni-NTA. Les vecteurs sont tout d'abord synthétisés, puis extraits et purifiés et sont recombinés avec le matériel génétique du virus à l'aide d'un test blanc/bleu. Le bacmide généré est alors utilisé pour infecter une première génération de cellules d'insectes par transfection. Celles-ci permettent d'obtenir le lot de virus V_0 qui sera ensuite séparé des cellules, récolté et centrifugé pour infecter une seconde génération de cellules. Ces virus sont dits amplifiés et produisent le stock appelé cette fois V_1 qui est conservé une année à 4 °C.

Les protéines Ku et XLF sont exprimées dans des cellules d'insectes de type SF21 ou Hi5. Nous suivons la fluorescence de la protéine YFP présente dans le bacmide lors de la culture (Figure 53). Cette protéine étant sous le même rapporteur que nos protéines, celle-ci donne une bonne indication de l'évolution de la production de nos protéines d'intérêt au cours du temps. L'analyse des gels SDS des extraits totaux (ET) montre que XLF et les hétérodimères Ku sont bien surexprimés.

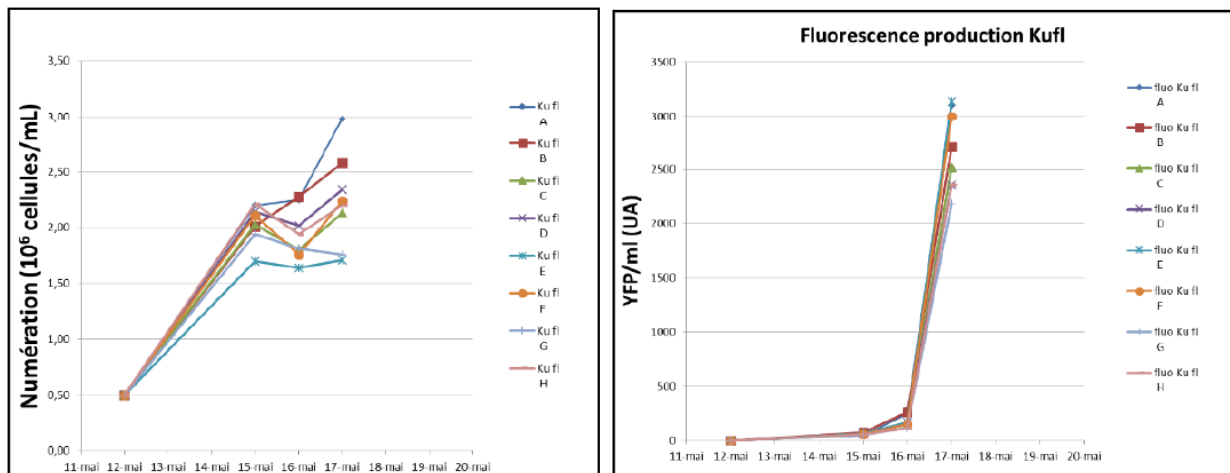


Figure 53 : **A gauche :** numération des cellules permettant d'avoir un avis sur la croissance cellulaire après infection par le virus portant l'ORF de la protéine entière Ku70/80. L'évolution du nombre de cellules montre un arrêt de la croissance cellulaire à J+3 après infection lorsque toute la population est infectée. On a à ce moment-là environ 4 fois plus de cellules. **A droite :** fluorescence des cellules infectées en unités arbitraires (UA). La fluorescence émise par la YFP est un rapporteur indirect de l'expression des protéines d'intérêt (ici Ku). La fluorescence augmente de manière forte à J+3 après infection. Les cellules sont alors récoltées à J+4

2) Purification de la protéine Ku70/80 ; version tronquée KuCC

Les trois constructions de la protéine Ku sont exprimées en cellules d'insecte. La protéine Ku possède sur la région N-terminale un tag Histidine (10 His) et un site de coupure à la protéase TEV. Des premiers tests de purification ont été réalisés sur des culots de 200 mL de culture afin d'optimiser le protocole de purification. Une fois que celui-ci fut optimisé, alors nous avons pu utiliser des culots correspondant à des volumes de culture plus importants de l'ordre de 600 mL.

Après lyse bactérienne, une première étape sur résine Ni-NTA (Qiagen) est effectuée. Toutes les fractions de la purification sont déposées sur gel SDS-PAGE 10 %. Les bandes protéiques sont révélées au bleu de Coomassie. Deux bandes à environ 60 kDa et 70 kDa sont observées et correspondent respectivement aux deux sous-unités Ku70 et Ku80 de KuCC de masse

théorique respectivement 62,542 kDa et 65,262 kDa. Des contaminants sont observés dans les fractions les plus concentrées (Figure 54).

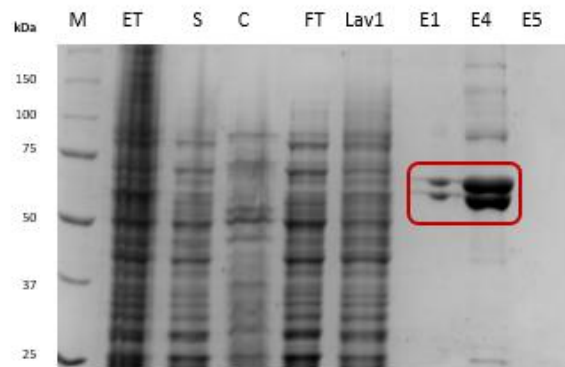


Figure 54 : Gel SDS-PAGE 10 % après la première étape de purification de KuCC : passage sur une résine Ni-NTA. La protéine est surexprimée dans l'extrait total (ET) et le surnageant ou fraction soluble (S). On visualise KuCC (Ku70 à environ 60 kDa et Ku80 à environ 70 kDa) dans la zone encadrée en rouge des fractions d'éluion E1 à E4

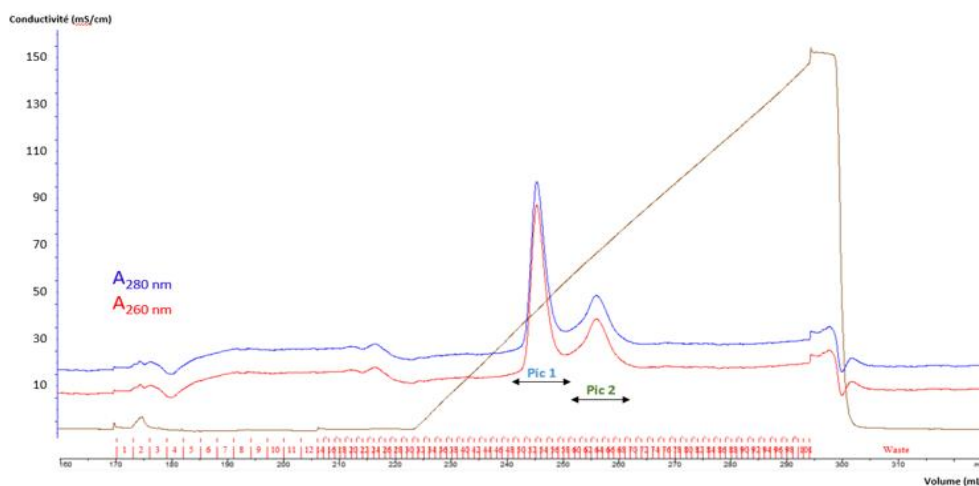


Figure 55 : Seconde étape de purification de KuCC : après passage sur une colonne échangeuse d'anions ResourceQ de 6 mL. **En haut :** chromatogramme de l'étape de ResourceQ. Signal d'absorbance de l'UV à 280 nm (bleu) et à 260 nm (rouge); conductivité (mS/cm) (marron). On observe deux pics (pics 1 et 2). **En bas :** gel SDS-PAGE 10 % après Resource Q. La protéine est présente dans les deux pics (pics 1 et 2) avec un meilleur degré de pureté. Dans les fractions les plus chargées, on observe des contaminants sont observés à 100 et 150 kDa.

Une seconde étape de purification est effectuée par échangeuse d'anions. Les fractions éluées de la colonne de nickel contenant KuCC sont regroupées, dialysées contre un tampon QA-Ku à faible force ionique. Un spectre UV de la protéine indique que l'échantillon est contaminé avec de l'ADN (ratio $A_{260/280} = 0.8$). KuCC est passée sur une colonne ResourceQ échangeuse d'anions de 6 mL et éluée avec un gradient de sels (Figure 55). Les échantillons correspondant aux fractions d'éluion (pic 1 \Leftrightarrow 150 mM sels et 2 \Leftrightarrow 300 mM sels) sont déposés sur gel SDS-PAGE 10 %.

Nous avons cherché à identifier l'origine de la présence de deux pics à cette étape. Nous avons regardé la présence de phosphorylation sur les deux sous-unités en utilisant le test *proQ Diamond Phosphoprotein Gel Stain*. Cet essai consiste en la coloration des protéines phosphorylées sur gel d'acrylamide. D'autres protéines du NHEJ ont été utilisées en comparaison. Le test indique un fort taux de phosphorylation sur la protéine XRCC4 et aucun signal sur KuCC et XLF. Nous avons alors effectué une analyse par spectrométrie de masse sur ces échantillons. Nous avons pu mettre en évidence une acétylation sur le NH2 terminal de chacune des deux sous-unités. Cette acétylation est présente de façon équivalente dans les échantillons issus du pic 1 ou du pic 2. Les fractions de chacun de ces deux pics sont regroupés. Les fractions sont dialysées contre le tampon de congélation contenant 5 % de glycérol. Après deux bains de dialyse, un spectre UV est réalisé. Nous obtenons classiquement une masse en KuCC d'environ 24 mg pour 600 mL de culture ; soit un rendement de 40 mg.L⁻¹ de culture. Les échantillons sont ensuite congelés à - 80 °C. Les protocoles de purification et les résultats obtenus pour la protéine entière KuFL et pour KuSC sont identiques. Seuls les poids moléculaires diffèrent entre ces trois constructions. Nous pouvons obtenir pour ces dernières des rendements allant jusqu'à 48 mg/L de culture. La purification de la version tronquée KuCC a été détaillée car c'est cette construction qui a été la plus utilisée lors de ma thèse. On note que lorsque la fluorescence correspondant au culot dépasse les 1500 UA lors de la production, un volume de 600 mL de culture est suffisant pour purifier 15 mg voire plus de protéines.

3) Purification de la protéine XLF

J'ai été associée au début de ma thèse à la finalisation des travaux de l'article du précédent doctorant Clément Nemoz (Nemoz et al. 2018). Dans ce contexte, j'ai purifié un partenaire de Ku70/80, la protéine XLF qui possède un motif de liaison à Ku, appelé KBM. J'ai purifié XLF dans sa version entière mais aussi deux versions possédant une mutation sur la motif KBM (XLF L297W et XLF L297E).

Ces trois variants de XLF présentent une étiquette 10 His et sont exprimés en cellules d'insecte Hi5. La purification est réalisée entièrement à 4 °C. Nous décrivons la purification pour un culot de cellules d'insectes correspondant à 400 mL de culture. La première étape de la purification est l'étape de dépôt de la fraction soluble sur 10 mL de résine Ni-NTA (Qiagen). Cette étape de chromatographie d'affinité est la même que celle pour la protéine Ku mais avec des tampons différents (détaillés dans la partie « Matériel et Méthodes »). Les fractions d'élution sont rassemblées (Figure 56) et dialysées contre le tampon QA-XLF de faible salinité.

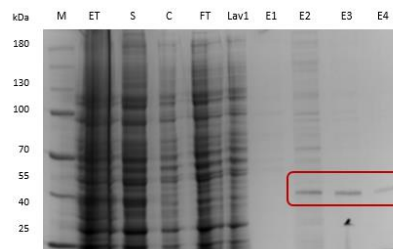


Figure 56 : Gel SDS-PAGE 10 % après la première étape de purification de XLF : passage sur une résine Ni-NTA. La protéine est surexprimée dans l'extrait total (ET) et le surnageant ou fraction soluble (S). On visualise XLF (bande unique à environ 40 kDa) dans la zone encadrée en rouge des fractions d'élution E2 à E5.

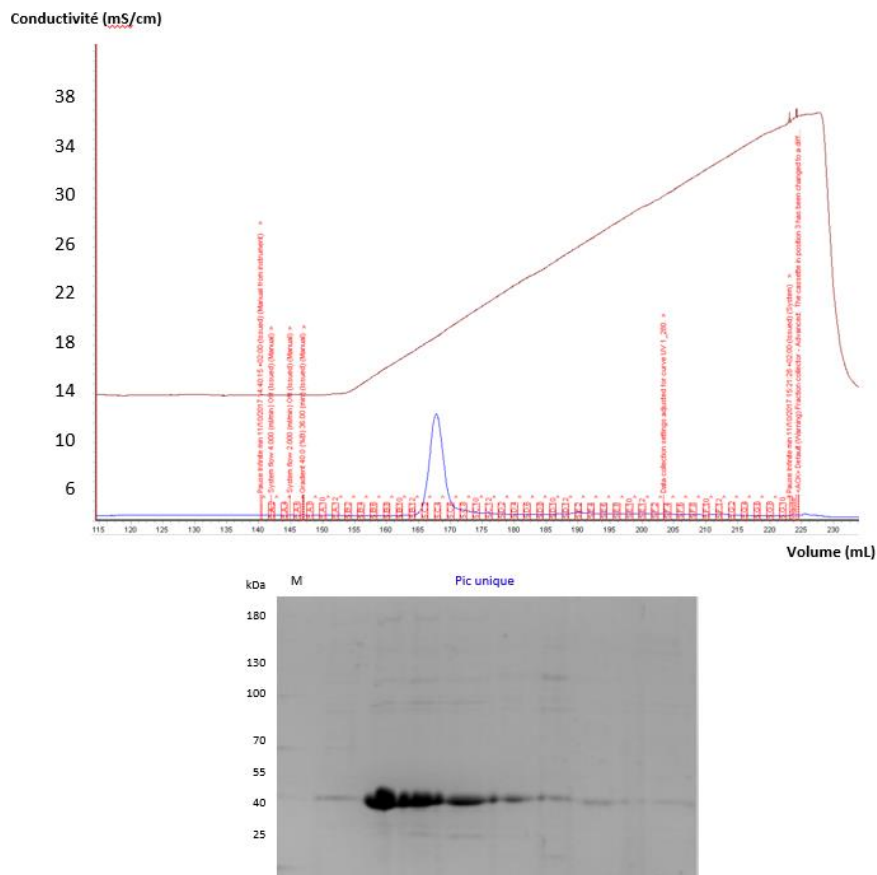


Figure 57 : Seconde étape de purification de XLF : après passage sur une colonne échangeuse d'anions ResourceQ de 6 mL. **En haut :** chromatogramme de l'étape de ResourceQ. Signal d'absorbance de l'UV à 280 nm (bleu); conductivité (mS/cm) (marron). On observe un pic unique correspondant à XLF. **En bas :** gel SDS-PAGE 10 % après ResourceQ

Après dialyse, un spectre UV est réalisé et une mesure de l'absorbance à 280 nm permet d'évaluer la quantité de protéine avant l'étape de passage sur la colonne échangeuse d'anions de 6 mL. Après dépôt de la protéine XLF sur la colonne ResourceQ, celle-ci est éluée avec un gradient de sels. Nous obtenons un seul pic pour XLF qui migre sur gel SDS-PAGE 10 %, vers 40 kDa en accord avec son poids moléculaire de 34 kDa et sa forme allongée due à la présence d'un *coiled-coil* et d'une région désordonnée en Cter de la protéine (Figure 57). Les fractions d'éluion sont regroupées pour être dialysées contre le tampon de congélation contenant 10 % de glycérol. Un spectre UV est réalisé en fin de purification. Nous obtenons classiquement une masse de 19 mg de protéine XLF pour 400 mL de culture soit un rendement de 47 mg de protéines par litre de culture. Les échantillons sont ensuite congelés à $-80\text{ }^{\circ}\text{C}$. La purification des mutants étiquetés 10 His est identique et les rendements semblables.

4) Analyse par chromatographie d'exclusion de taille des protéines purifiées

Une analyse par chromatographie d'exclusion de taille (ou SEC pour *Size Exclusion Chromatography* en anglais) permet à la fois de vérifier l'état de pureté d'un échantillon, si la macromolécule est agrégée et d'obtenir une estimation du poids moléculaire apparent de l'échantillon après calibration de la colonne. Cette analyse est effectuée avant et après congélation pour estimer l'impact de l'étape de congélation/décongélation sur les protéines. Après calibration de la colonne Superose6 increase 10/300 (GE Healthcare) de 24 mL avec des marqueurs de poids moléculaires standards (liste dans la partie « Matériel et Méthodes »), et injection des échantillons, je récupère le chromatogramme d'analyse SEC (Figure 58).

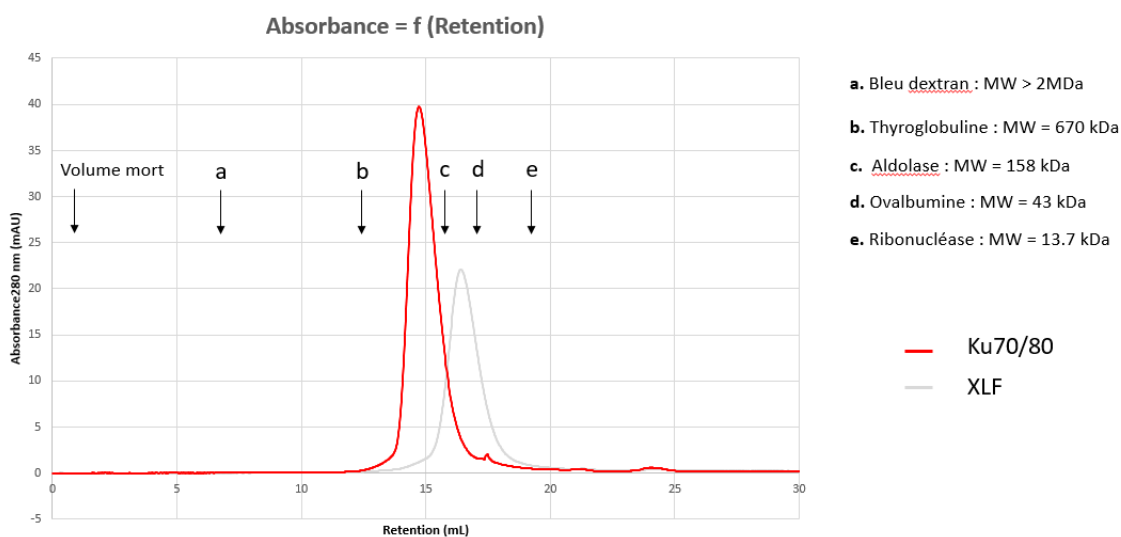


Figure 58 : Chromatogramme de l'analyse par SEC des protéines Ku70/80 et XLF. La colonne Superose6 increase 30/100 de 24 mL est calibrée avec le bleu dextran et 4 marqueurs de poids moléculaire connus

Nous mesurons pour chaque marqueur de poids moléculaire, son volume d'élution. Une gamme étalon est alors tracée et nous permet de déterminer par régression linéaire la taille moléculaire des protéines Ku et XLF purifiées en fonction du coefficient de partage K_{av} :

$$K_{av} = (V_e - V_m) / (V_t - V_m)$$

J'obtiens alors une adéquation entre les masses moléculaires théoriques et expérimentales des deux protéines (tableau ci-dessous). De plus, chacune est éluée en un pic unique ce qui montre que les échantillons sont homogènes et qu'il n'y a pas d'agrégation significative suite à leur congélation. En effet, on n'observe pas d'absorbance à 280 nm au niveau du volume mort. L'erreur expérimentale (%) est calculée grâce à la formule suivante :

$$(|M_{théo} - M_{exp}| / M_{théo}) \times 100$$

Les résultats de la chromatographie d'exclusion de taille (sur Superose6 increase 10/300 de 24 mL) pour KuFL et XLF sont donnés ci-dessous :

	V_e (mL)	K_{av}	logM	$M_{théo}$ (Da)	M_{app} (Da)	Ecart ($M_{app} - M_{théo}$)
Ku70/80	14,70	0,44	5,30	155 000	198 000	+ 28 %
XLF (dimère)	16,45	0,55	4,87	69 000	74 000	+ 6 %

B) Structure cristallographique du complexe

Ku70/Ku80/hDNA en présence de l'inositol phosphate (IP6)

1) Essais de cristallisation Ku-IP6 / Ku-ADN-IP6

Afin de connaître la localisation de l'inositol phosphate au sein de l'hétérodimère seul ou en complexe avec l'ADN, j'ai effectué des tests de cristallisation dans le but d'obtenir une structure cristallographique des complexes Ku-IP6, Ku-ADN-IP6 et Ku-ADN-pXLF/pAPLF-IP6. La première étape de ce travail a été de préparer les échantillons et de les envoyer à cristalliser sur la plateforme de cristallisation HTX située à l'EMBL, Grenoble pour un criblage initial sur 576 conditions de cristallisation différentes. Plusieurs complexes préparés à des concentrations de 5, 8 et 12 mg.mL⁻¹ ont été envoyés avec l'IP6 mais aussi avec les petites molécules IP5 et IP4. Pour chacun de ces complexes, l'ADN *hairpin* (hDNA) et la construction KuCC ont été choisis.

L'ADN *hairpin* contenant une boucle de 13 nucléotides (formant 4 pb) et une partie bicaténaire de 14 pb (Figure 59), permet d'orienter Ku et de favoriser la présence d'un seul complexe avec l'ADN. La construction KuCC ne présente pas les régions C-terminales et les charnières flexibles qui relient ces domaines au reste de la protéine. Ce retrait des régions Cter a pour objectif d'augmenter les chances de succès de l'étape de cristallisation. L'ADN et les petites molécules sont mis en excès de 1,2 fois la concentration molaire en protéine Ku (quelle que soit la construction utilisée) afin de limiter la concentration de protéine Ku libre. Cette première étape m'a permis d'obtenir des pistes de cristallisation parmi 576 conditions (6 x 96 conditions détaillées dans la partie « Matériel et Méthodes »). Des gouttes d'un volume de 0,2 μ L (0,1 μ L de complexe protéique mélangés à 0,1 μ L de solution de cristallisation) ont été faites avec la technique de cristallisation par diffusion en phase vapeur en géométrie de goutte assise. Les boîtes de cristallisation sont stockées à l'EMBL à 4 °C et les images prises régulièrement par le robot de visualisation Formulatrix, sont observées à distance depuis le laboratoire. Des images sont prises de manière automatique et régulière de J+1 à J+87 et visualisables grâce au logiciel CRIMS qui a été développé par la plateforme HTX. Huit conditions de cristallisation intéressantes contenant à chaque fois du PolyEthylène Glycol (PEG) 3350 comme agent précipitant et une solution de BisTris Propane (BTP) tamponnée entre pH 6,5 et pH 8,5, ont été observées à partir du crible et reproduites au laboratoire à différentes températures (4 °C, 17 °C et 20 °C) à la fois en géométrie de goutte assise mais aussi en géométrie de gouttes suspendues. Les sels contenus dans ces solutions de cristallisation sont variés : Na Malonate, Na Citrate tribasic dihydrate, Na Tartrate, Na Sulfate ou encore Na Formate. Après reproduction des cristaux au laboratoire en géométrie de goutte suspendue, des bâtonnets et aiguilles ont été obtenus avec les conditions suivantes : 20 % PEG 3350, 100 mM BTP pH 8,5, 0,2 M Na Sulfate et 20 % PEG 3350, 100 mM BTP pH 7,5, 0,2 M Na Malonate respectivement à 20 °C et 4 °C pour le complexe KuCC-hDNA-IP6 (Figure 60).

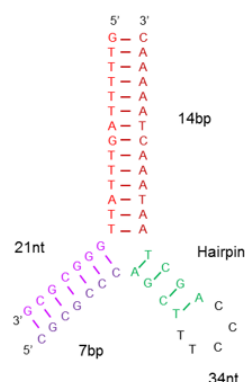


Figure 59 : Structure de l'ADN tige-boucle ou hDNA contenant une boucle de 13 nucléotides (formant 4 pb) et une partie bicaténaire de 14 pb

La taille des cristaux est compatible avec une analyse de diffraction par rayons X. Ils mesurent pour la première condition, environ $50\ \mu\text{m} \times 30\ \mu\text{m} \times 200\ \mu\text{m}$ et pour la seconde condition, environ $20\ \mu\text{m} \times 20\ \mu\text{m} \times 60\ \mu\text{m}$. Les cristaux de la première condition sont plus délicats à manipuler car ils sont dits « maclés » (cristaux qui poussent dans différentes directions à partir d'une même base). Ils présentent fréquemment une partie centrale bien organisée depuis laquelle des extensions se sont développées dans différentes directions. Ces imperfections sont généralement au détriment de la qualité de la diffraction du cristal par la suite car elles créent un désordre. Les meilleurs résultats en diffraction sont souvent obtenus en cassant certaines « branches » de ces cristaux avant de les pêcher et de les congeler dans l'azote liquide. Il arrive très souvent que les « branches » de ces cristaux soient insérées dans la base compacte du cristal. Il est alors impossible de les séparer de leur base.

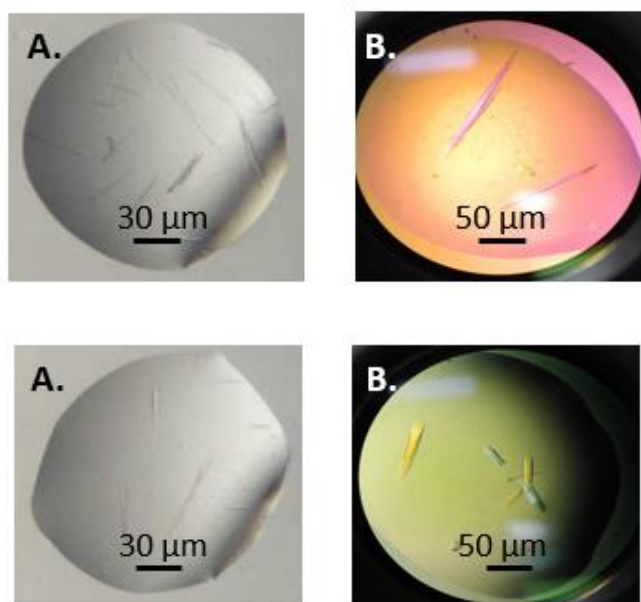


Figure 60 : Résultats de cristallisation du complexe KuCC-hDNA-IP6 à $8\ \text{mg.mL}^{-1}$. **En haut** : pour les conditions 20 % PEG 3350, 100 mM BTP pH 8,5, 0,2 M Na Sulfate à $20\ ^\circ\text{C}$ et **en bas** : pour les conditions 20 % PEG 3350, 100 mM BTP pH 7,5, 0,2 M Na Malonate à $4\ ^\circ\text{C}$. **En A** : les cristaux apparus lors du crible sur la plateforme HTX. **En B** : les cristaux reproduits au laboratoire en géométrie de goutte suspendue

2) Optimisation des conditions de cristallisation et de diffraction de KuCC-hDNA-IP6

Afin d'optimiser la cristallisation du complexe KuCC-hDNA-IP6 et d'améliorer sa diffraction, plusieurs stratégies ont été testées. La première stratégie a consisté tout d'abord à élargir plusieurs paramètres : concentration en PEG 3350, concentration en BTP, établir une gamme de pH plus étroite entre pH 6,5 et pH 8,5, modifier la taille en PEG de PEG 400 à PEG 6000,

température de cristallisation et enfin concentration en sels Na Malonate et Na Sulfate. Une trentaine de boîtes de 24 puits chacune ont été réalisées correspondant à un total de 720 gouttes pour environ 6 mg de protéine KuCC purifiée.

La condition suivante à 20 °C : 25 % PEG 3350, 110 mM BTP pH 8,5, 150 mM Na Sulfate a permis d'obtenir de beaux cristaux anisotropes en forme d'aiguilles au bout de 2 jours (50 µm x 30 µm x 200 µm). Ces cristaux sont pêchés après avoir attendu la fin de leur croissance et sont congelés dans l'azote liquide. Mais avant congélation dans l'azote liquide, pour empêcher toute apparition de glace due à la présence d'eau dans la solution de cristallisation, ces cristaux sont trempés dans une solution cryo-protectrice. Celle-ci est constituée de la solution de cristallisation avec ajout d'un cryo-protectant.

La seconde stratégie d'optimisation a été de tester différents cryo-protectants (sucres, glycérol et éthylène glycol) afin de permettre d'augmenter la résolution du jeu de données lors de l'analyse au synchrotron. Les cristaux sont donc pêchés avec une boucle en nylon ou une plume de 0.1 mm et déposés environ 15 sec dans une goutte contenant un cryo-protectant avant d'être congelés dans l'azote liquide. Etant donné le pourcentage élevé en PEG 3350, il est conseillé de tester le pouvoir de diffraction d'un cristal sans ajout de cryo-protectant. Pour la condition cryo-protectrice de 7 % d'éthylène glycol, la meilleure résolution obtenue est de 5,8 Å (sur la ligne de lumière PX2 au synchrotron SOLEIL). Cette résolution est trop basse pour pouvoir visualiser l'IP6.

Nous avons ainsi choisi de réaliser deux autres cribles utilisant des conditions de cristallisation proposées par Sanofi et contenant 384 conditions de cristallisation. Le premier crible est effectué en mettant des semences de cristaux de Ku70/Ku80/hDNA/IP6 dans toutes les gouttes. Le second est effectué en utilisant du KuCC digéré à la trypsine pour ôter certaines régions non-structurées. Le crible avec Ku protéolysé n'a pas permis d'identifier de nouvelles conditions de cristallisation. Le crible avec les semences de cristaux a donné quelques pistes avec des cristaux présentant une nouvelle forme arrondie pour certains et sous forme de bâtonnets pour d'autres. Les expériences de diffraction ont été réalisées sur la ligne PX2 car les cristaux étaient petits mais ils n'ont malheureusement pas diffracté à une résolution inférieure à 8 Å.

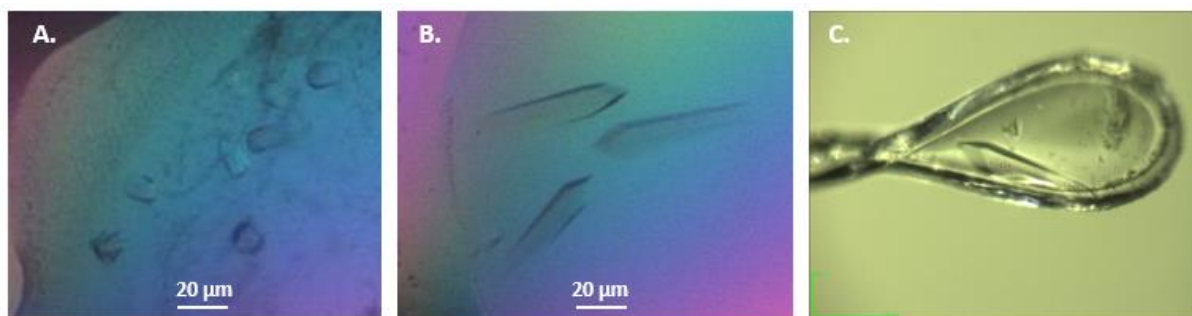


Figure 61 : Résultats de cristallisation du complexe KuCC-hDNA-IP6 après ensemencement et utilisation du crible Sanofi. **En A** : les cristaux sont de forme arrondie pour une apparition à J+3 avec 25% PEG 4000, 100 mM Tris pH 8.5 et 26 % glycérol comme cryo-protectant. **En B** : les cristaux sont de forme bâtonnet pour une apparition à J+2 avec 20 % PEG 3350, 100 mM Na Malonate et 26 % glycérol comme cryo-protectant. **En C** : photo d'une boucle nylon de 0.1 mm et d'un cristal pêché depuis la condition décrite en B

Voyant que les différentes approches n'apportaient rien de plus quant à la résolution de diffraction du cristal, des additifs ont finalement été testés. C'est l'ajout de 15 mM de β -mercaptoéthanol qui a été retenu pour l'amélioration de la diffraction.

3) *Diffraction et traitement des données collectées pour le complexe KuCC-hDNA-IP6*

De nombreux cristaux ont été testés sur les deux lignes de lumière Proxima-2A et Proxima-1. Ceux-ci ont diffracté de 9,8 Å à 4,1 Å. Les données concernant le complexe KuCC-hDNA-IP6 ont été collectées sur la ligne de lumière Proxima-1 au synchrotron SOLEIL. Après montage sur un goniomètre de 3 axes, le cristal, maintenu sous flux d'azote, est traversé par un faisceau de rayons X à 12,5 keV pour collecter les images de diffraction sur un détecteur Pilatus 6M (Dectris Ltd) à raison de 25 images par seconde. La meilleure résolution obtenue est de 4,1 Å pour ce complexe. Une stratégie de collecte a été choisie pour enregistrer 1800 clichés de diffraction sur une plage de 180 ° tous les 0,1 °. Le jeu de données est traité grâce au logiciel « xdsme » développé par Wolfgang Kabsh (<https://www.onlinelibrary.wiley.com/iucr/doi/10.1107/S0907444909047337>) avec son interface développée par Pierre Legrand (<https://github.com/legrandp/xdsme>) et sa mise à l'échelle est effectuée grâce au script XSCALE. Nous obtenons les paramètres suivants :

Groupe d'espace \Leftrightarrow type de groupe d'espace selon les tables internationales de cristallographie

Dimensions a, b et c (Å) \Leftrightarrow longueur des axes de la maille

Dimensions α , β et γ (°) \Leftrightarrow angles entre les trois axes de la maille

Résolution (Å) \Leftrightarrow résolution de diffraction

R_{merge} \Leftrightarrow quantification de dispersion des tâches de diffraction équivalentes

R_{meas} \Leftrightarrow quantification de la qualité de mesure d'intensité des tâches individuelles

R_{pim} \Leftrightarrow quantification de la qualité de mesure d'intensité moyenne

I/ σ (I) \Leftrightarrow signal / bruit

CC_{1/2} \Leftrightarrow coefficient de corrélation entre deux moitiés de jeu lorsque celui-ci est coupé en deux de manière aléatoire

Complétude (%) \Leftrightarrow évaluation du nombre de réflexions mesurées en comparaison avec le nombre de réflexions attendues dans la symétrie donnée du cristal

Redondance \Leftrightarrow la répétition d'enregistrement de chaque tâche de diffraction

Facteur B Wilson (Å²) \Leftrightarrow facteur proportionnel à l'agitation thermique au sein du cristal

Les cristaux sont anisotropes. Ils ne présentent pas le même pouvoir de diffraction dans les différentes directions sur le détecteur. Les tâches de diffraction sont traitées en tenant compte de cette anisotropie. Cette stratégie permet d'améliorer le rapport signal sur bruit global du jeu de données en ne prenant pas en considération des tâches de diffraction trop faibles dans certaines directions. Nous utilisons le serveur STARANISO (pour *Anisotropy of the Diffraction Limit and Bayesian Estimation of Structure Amplitudes* en anglais) : <http://stارانiso.globalphasing.org/cgi-bin/stارانiso.cgi> qui permet d'effectuer le traitement de données anisotropes. On ne récupère donc que les régions de l'espace réciproque contenant les informations importantes de diffraction significatives. Sans traitement anisotrope, la résolution était de 4,1 Å ; après son utilisation, la résolution est de 3,9 Å. Le gain grâce au serveur est de 0,2 Å. Nous obtenons un groupe d'espace P3₂21.

4) Phasage et affinement

Le phasage permet de calculer la carte de densité électronique correspondant aux différentes tâches de diffraction en déterminant la phase de chaque tâche de diffraction. Nous utilisons la

technique de Remplacement Moléculaire (RM) qui permet de positionner une structure de macromolécule proche de celle présente dans le cristal et d'utiliser les phases correspondant à ce modèle une fois positionné. Nous utilisons comme modèle la structure cristallographique du complexe Ku avec l'ADN *hairpin* (code PDB : 1JEY) (Walker, Corpina, et Goldberg 2001) auquel nous retirons le domaine C-terminal de Ku70 non-présent dans notre structure. Le logiciel Molrep de CCP4 arrive à positionner des copies de la structure modèle 1JEY (Figure 62) dans l'unité asymétrique du cristal. Le remplacement moléculaire effectue une première fonction de rotation. La fonction de Patterson des données expérimentales est comparée à la fonction de Patterson du modèle dans une orientation donnée. Les meilleures solutions angulaires trouvées avec la fonction de rotation sont utilisées pour calculer des fonctions de translation qui positionnent le modèle en différents endroits de la maille. Le coefficient de Matthews du cristal suggère la présence de 1 molécule de Ku dans l'unité asymétrique. Cette étape permet de positionner plusieurs molécules de Ku avec un indice de confiance élevé. Nous effectuons alors un affinement en corps rigide avec le logiciel BUSTER. La résolution étant assez basse, la stratégie utilisée a été d'affiner la structure par blocs rigides de domaines protéiques sans chercher à affiner les chaînes latérales des acides aminés. Les chaînes protéiques possèdent des densités satisfaisantes au niveau de la chaîne principale et nous n'observons pas de superpositions entre-elles qui pourraient suggérer des erreurs lors du remplacement moléculaire. Nous analysons alors la carte de densité électronique afin d'identifier la présence éventuelle de l'inositol phosphate (IP6) au sein de la structure KuCC-hDNA. Nous identifions dans les deux cartes de densité électronique générées « 2Fobs-Fcalc » et « Fobs-Fcalc », une densité résiduelle à l'interface des sous-unités de Ku70 et Ku80 (Figure 62C). De par sa forme et son intensité, ce signal semble correspondre à l'IP6. Cet inositol phosphate est localisé dans une poche proche des résidus basiques. Certains de ces résidus avaient notamment été identifiés par mutagenèse dirigée comme importants pour la fixation de cette petite molécule. Cet IP6 est construit avec le logiciel Coot (Crystallographic Object-Oriented Toolkit) puis le complexe KuCC/hDNA/IP6 est affiné par corps rigide avec BUSTER (Figure 62D). Les facteurs d'accord entre les facteurs de structure expérimentaux et le modèle sont améliorés entre le remplacement moléculaire et le dernier affinement en corps rigide (Rwork initial de 34.1 % et final de 27.5 %, Rfree initial de 35.8 % et final de 33.4 %).

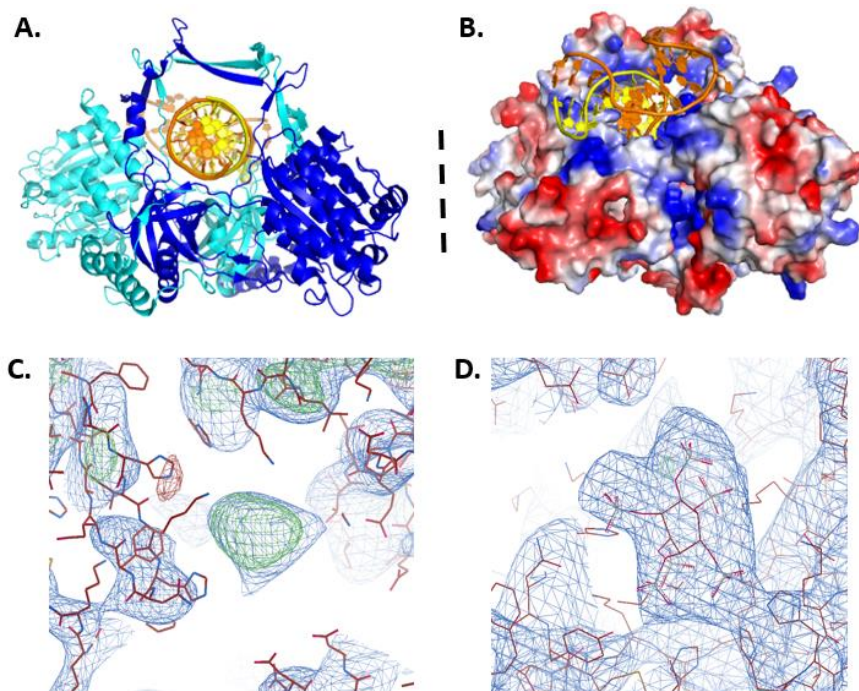


Figure 62 : **A.** Structure cristallographique (PDB : 1JEY) du complexe Ku70/80 avec l'ADN en épingle à cheveux (*hairpin*) (Walker, 2001) ; la sous-unité Ku70 (bleu foncé), la sous-unité Ku80 (bleu cyan) et l'ADN *hairpin*. (jaune orangé). La construction utilisée ici est la construction KuC possédant seulement la région C-terminale de Ku70 supprimée. **B.** Répartition des charges sur les deux sous-unités de Ku70/80. **C.** Carte de densité électronique de KuCC-hDNA-IP6 après le premier cycle d'affinement en corps rigide avec le logiciel BUSTER ; en vert les zones résiduelles observées notamment à l'interface de Ku70 et Ku80 représentant la localisation de l'IP6 (Fo-Fc positif). **D.** Zoom sur la densité résiduelle observée en C et avec la présence de l'IP6 après trois affinements avec BUSTER.

Le modèle final est ainsi obtenu (Figure 63).

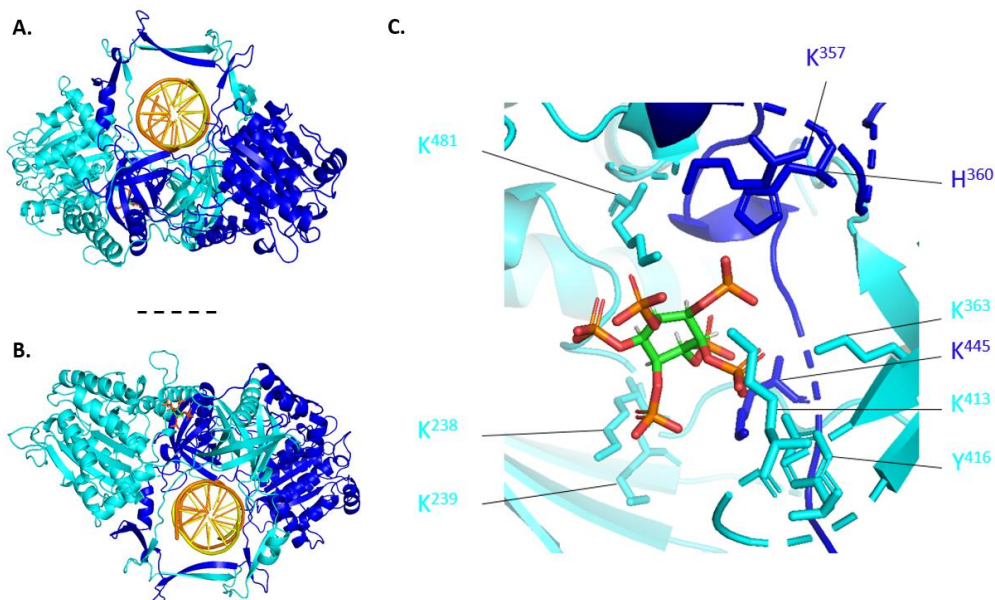


Figure 63 : Modèle final du complexe KuCC-hDNA-IP6 dans sa globalité et zoom sur la localisation de l'IP6. **A.** Modèle final global du complexe KuCC-hDNA-IP6, représentation en rubans des sous-unités Ku70 et Ku80 et en stick de l'IP6 grâce au logiciel Pymol. **B.** Rotation de 180 ° du modèle final global du complexe KuCC-hDNA-IP6 vu en A. **C.** Zoom sur la localisation de l'IP6 à l'interface des deux sous-unités de Ku70 et Ku80 et indication des résidus impliqués dans l'interaction.

5) Analyse des données cristallographiques du complexe KuCC-hDNA-IP6

Malgré la résolution faible de la structure (3,9 Å), la petite molécule d'inositol phosphate est positionnée sans ambiguïté à l'interface des deux sous-unités de l'hétérodimère. Certains acides aminés proposés dans une étude précédente comme impliqués dans la fixation de l'IP6 sont confirmés (K³⁵⁷ sur Ku70 et K²³⁸ et K²³⁹ sur Ku80). De nouveaux résidus impliqués dans l'interaction Ku-IP6 sont validés en plus des résidus décrits dans la littérature : on retrouve les lysines 357 et 445 ainsi que l'histidine 360 sur Ku70 et les lysines 238, 239 363, 413 ainsi que la tyrosine 416 sur Ku80. Certains proposés ne sont pas au contact de l'IP6 mais peuvent avoir un rôle indirect (exemple de la lysine 481). Il n'est pas surprenant que l'IP6 étant chargé négativement, soit en interaction avec la plupart des résidus chargés positifs (acides aminés basiques : His, Lys). Seul le résidu Tyr416 impliqué dans l'interaction avec l'inositol n'est pas chargé.

Afin d'analyser la structure entière de ce complexe allant des données globales aux détails moléculaires, certains paramètres sont traités. Tout d'abord la structure de KuCC-hDNA est comparée à celle de 1JEY (Walker, Corpina, et Goldberg 2001). En comparaison avec la structure 1JEY, la valeur de RMSD calculée nous donne ainsi la distance entre les C α équivalents de chaque structure (1JEY utilisée comme modèle et nos données expérimentales de KuCC-hDNA-IP6). Cette valeur signifie que ces C α sont espacés de moins de 1 Å.

L'interaction de l'IP6 avec le complexe Ku-ADN n'induit donc aucun changement de conformation ni d'ouverture de l'hétérodimère. Nous analysons la conservation en séquence des résidus impliqués dans l'interaction entre Ku et l'IP6. L'utilisation du serveur ConSurf (<http://consurf.tau.ac.il/2016/>) indique que les résidus K³⁵⁷, K⁴⁴⁵ et H³⁶⁰ de Ku70 et K²³⁸, K²³⁹, K³⁶³, K⁴¹³ et Y⁴¹⁶ de Ku80 sont conservés, et que seul le résidu K⁴⁸¹ de Ku80 l'est moins (Figure 64). De plus, l'IP6 est situé au niveau de l'hélice formée par les lysines K²³⁸ et K²³⁹ de Ku80, au niveau des boucles entre feuillets bêta et à une distance de l'ADN de 17 Å.

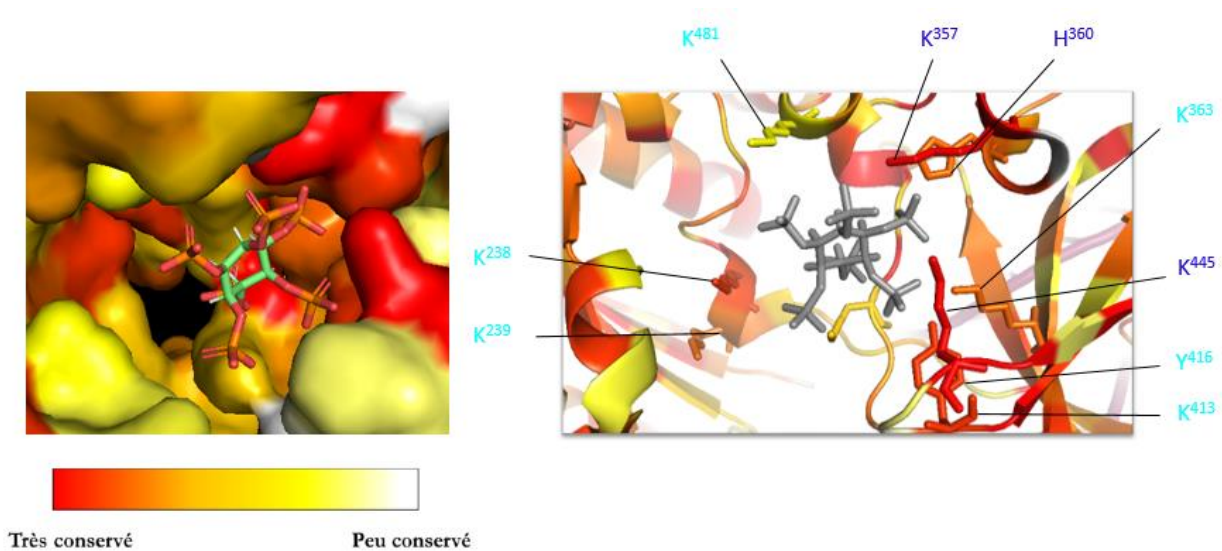


Figure 64 : Représentation de la conservation des résidus de surface de Ku70 et de Ku80 impliqués dans l'interaction avec la petite molécule d'inositol phosphate (IP6). Les acides aminés les plus conservés sont colorés en rouge et cette conservation diminue avec la couleur allant vers le jaune puis le blanc

Nom du complexe	KuCC-hDNA-IP6
Collecte des données	
Groupe d'espace	P3 ₂ 21
a, b, c (Å)	121,254 121,254 320,782
α , β , γ (°)	90 90 120
Résolution (Å)	47,74 – 3,94
Rmerge	0,160
Rmeas	0,171
Rpim	0,059
I/ σ (I)	7,8
CC1/2	0,995
Complétude ellipsoïdale (%)	85,9

6) Comparaison des structures Ku-ADN-IP6 et Ku-ADN-pXLF

La structure de KuCC-hDNA-pXLF a été résolue au laboratoire (Nemoz et al. 2018). Cette structure cristallographique montre une déviation du domaine vWA de Ku80 par rapport à la structure 1JEY (Walker, Corpina, et Goldberg 2001) (Figure 65). Le peptide pXLF (notamment les résidus ₂₉₂KKGRGLFS₂₉₉) est localisé dans une poche délimitée par les hélices $\alpha 6'$ et $\alpha 7$ de la sous-unité Ku80. Cette interaction entre pXLF et Ku est notamment due à une interaction

polaire puisque le Cter de pXLF est en interaction avec la lysine 233 de Ku80. Enfin la sérine 299 de pXLF forme des interactions électrostatiques avec l'arginine 44 et les lysines 238/239 de Ku80.

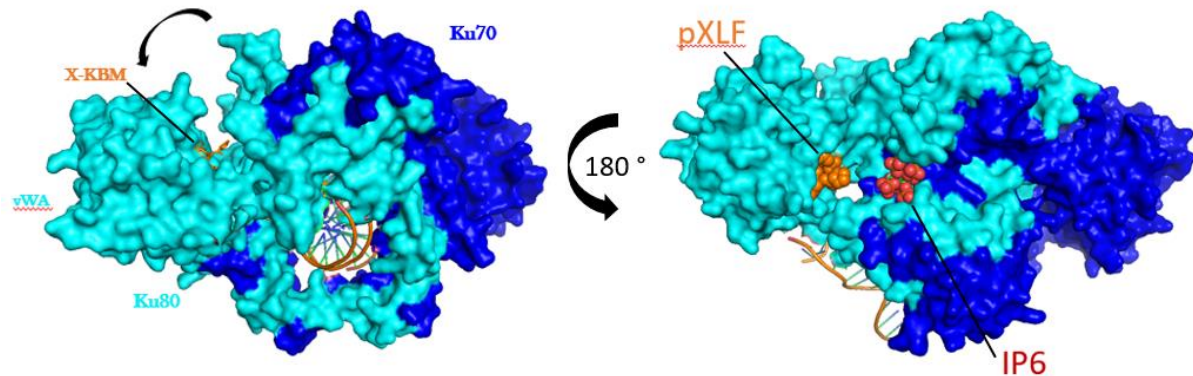


Figure 65 : **A gauche :** structure cristallographique du complexe KuCC-hDNA-pXLF résolue au laboratoire (Nemoz et al, NSMB, 2018). Afin de ne pas cacher la fixation du peptide pXLF, l'ADN *hairpin* est tronqué ; vue globale de Ku en surface (Ku70 en bleu foncé, Ku80 en bleu cyan et le peptide pXLF en orange). Visualisation de l'ouverture du domaine vWA de Ku80. **A droite :** structure cristallographique (tournée de 180 ° par rapport à la figure de gauche) du complexe KuCC-hDNA-IP6 résolue au laboratoire (Gontier et al, en préparation) avec position du peptide pXLF situé dans la même poche et non loin de l'IP6.

L'IP6 interagit à l'interface des deux sous-unités de Ku et on ne retrouve pas l'ouverture du domaine vWA de Ku80 comme dans la structure avec le peptide pXLF. En superposant les deux structures KuCC-hDNA-pXLF et KuCC-hDNA-IP6, on s'aperçoit que pXLF et l'IP6, se retrouvent en interaction avec Ku dans deux poches situées à proximité (Figure 65). Les deux structures cristallographiques montrent notamment deux résidus de Ku (K^{238} et K^{239}) qui interagissent à la fois avec pXLF et avec l'IP6. Ces résultats suggèrent une influence potentielle voire une compétition entre ces deux interactions.

C) Caractérisation biophysique de l'interaction entre Ku et l'IP6

Analyse par microcalorimétrie

La microcalorimétrie permet de mesurer les dégagements ou absorptions de chaleur consécutifs à des interactions protéine-ADN et d'en extraire la constante de dissociation (K_D), la stoechiométrie de la réaction (N) et l'enthalpie (ΔH). Des mesures d'interaction ont été réalisées en utilisant trois constructions de Ku (KuFL / KuCC / KuSC) et trois constructions en ADN (hDNA / ADN 18 pb / ADN 42 pb). L'ADN *hairpin* est l'ADN utilisé en cristallographie, l'ADN 18 pb interagit avec un hétérodimère et l'ADN 42 pb interagit avec deux molécules de

Ku d'après des expériences précédentes faites au laboratoire (Tadi et al. 2016; Nemoz et al. 2018). Pour ces expériences, la protéine Ku (à 1 μM) est mise dans la cellule de mesure et l'ADN (à 10 μM) dans la seringue d'injection. Cette interaction entre Ku et l'ADN est endothermique montrant une absorption de chaleur de faible amplitude (Figure 66). Les résultats obtenus sont regroupés dans le tableau récapitulatif ci-après.

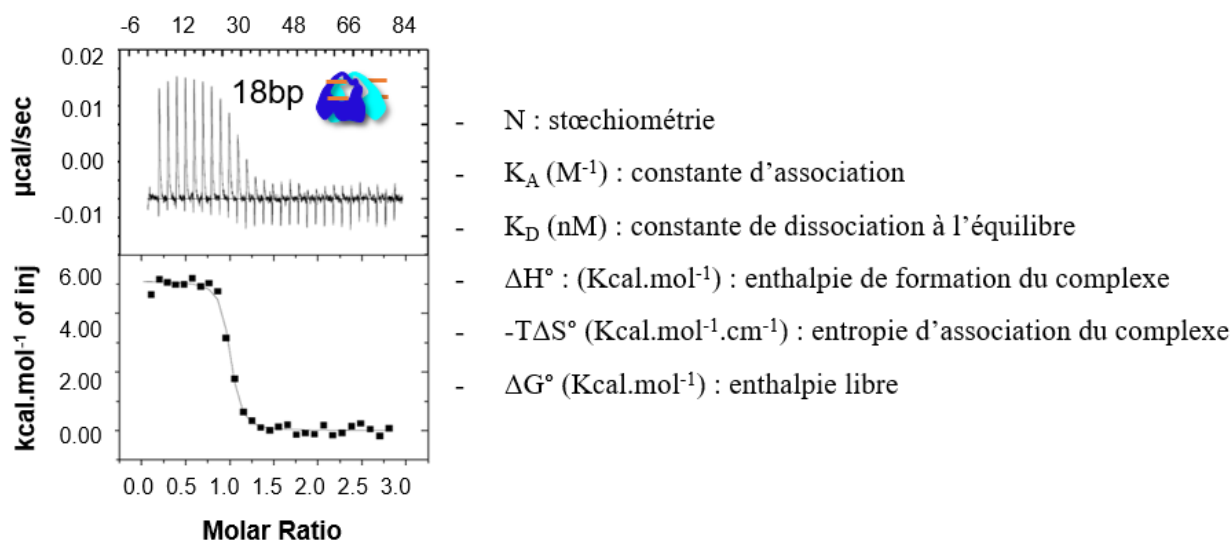


Figure 66 : Illustration d'un exemple de résultat d'ITC d'une interaction KuFL-ADN18 (Gontier et al, MiMB, 2018). **En haut** : thermogramme montrant des pics positifs témoignant d'une interaction de type endothermique. **En bas** : isotherme de titration montrant le ΔH° extrapolé pour un ratio molaire de 0, le K_D est déduit de la pente de la courbe au point d'inflexion et la stœchiométrie N correspond au ratio molaire au point d'inflexion de la courbe

KuFL est la forme complète de Ku70/80. KuCC est une version réduite en Cter de Ku70 et Ku80 et que KuSC est une version de KuCC avec une vingtaine de résidus en moins en Nter du Ku70. Les mesures effectuées montrent que la délétion des régions C-terminales de Ku et les premiers acides aminés de Ku70 n'a pas d'influence significative sur l'ADN hairpin sur lequel Ku interagit avec 14 paires de bases d'après la structure 3D. Ces délétions ont peu d'influence sur la fixation d'un ADN de 18 paires de bases (18 pb) (la mesure doit être faite sur un 18 pb pour KuSC mais celle faite sur un 21 pb montre des constantes semblables). Les 20 premiers acides aminés de Ku70 ne sont pas ordonnés dans le cristal et présentent une distribution de charge fortement négative avec 12 acides aminés acides. Cette région de Ku semble peu interférer avec la fixation de l'ADN.

Cellule de mesure	Seringue d'injection	N	K_D (nM)	ΔH^* (Kcal.mol ⁻¹)	$-T\Delta S^*$ (Kcal.mol ⁻¹)	ΔG^* (Kcal.mol ⁻¹)
KuFL	hDNA	0,99 ± 0,03	3,0 ± 0,2	9,1 ± 0,2	-19,2 ± 0,5	-10,1 ± 0,4
KuFL	ADN18	1 ± 0,1	3,5 ± 0,8	5,1 ± 0,1	-16,6 ± 0,2	-11,5 ± 0,2
KuFL	ADN42	0,47 ± 0,1	0,3 ± 0,2	16,7 ± 0,3	-30 ± 0,1	-12,8 ± 0,9
KuCC	hDNA	0,97 ± 0,02	2,9 ± 0,8	11,8 ± 0,3	-23,4 ± 1,6	-11,6 ± 0,3
KuCC	ADN18	0,98 ± 0,02	3,4 ± 0,4	6,5 ± 0,2	-21,4 ± 0,3	-11,2 ± 0,1
KuCC	ADN42	0,4 ± 0,01	2 ± 0,1	29,6 ± 0,3	-38,4 ± 2	-9,2 ± 0,2
KuSC	hDNA	0,9 ± 0,1	3,1 ± 0,01	9,9 ± 0,4	-21,2 ± 2,8	-10,8 ± 0,3
KuCC	ADN21	0,84 ± 0,01	3,03 ± 2,6	9,5 ± 0,07	-21,2 ± 0,7	-11,6 ± 0,5

La stœchiométrie observée est de 1 : 1 avec l'ADN *hairpin*, l'ADN 18 pb ou 21pb. Les mesures effectuées avec un ADN 42 pb indiquent une constante de dissociation et une stœchiométrie deux fois plus faibles. La variation de stœchiométrie est en bon accord avec la place occupée par Ku sur l'ADN. Le gain d'un facteur 2 en affinité représente un gain faible d'énergie libre d'interaction qui peut correspondre à une liaison faible additionnelle ou à un léger effet de stabilisation entre deux molécules de Ku.

Les valeurs entropiques et enthalpiques retrouvées pour ces différentes expériences indiquent que ces réactions sont conduites par une entropie favorable et une enthalpie défavorable.

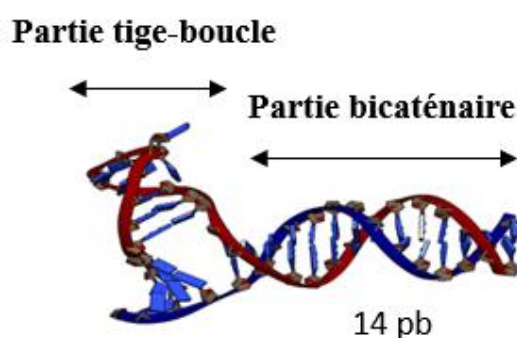


Figure 67 : Structure de l'ADN hairpin (Walker et al, 2001). Une partie tige-boucle (à gauche) et une partie bicaténaire linéaire (à droite) d'une longueur de 14 pb

Nous avons alors analysé l'interaction de l'IP6 avec Ku et avec le complexe Ku-ADN. L'interaction avec Ku seule met en évidence un K_D de 38 nM avec une stœchiométrie de 1 : 1.

L'interaction est conduite par l'enthalpie. Nous avons alors réalisé les expériences en mettant dans la cellule de mesure le complexe Ku-ADN et dans la seringue d'injection, la petite molécule d'inositol phosphate. Nous mesurons un K_D de 50 nM entre le complexe Ku-ADN et l'IP6 très proche de l'affinité de l'IP6 pour Ku seule. La présence de l'ADN sur Ku a peu d'impact sur la fixation de l'IP6. Nous avons ensuite comparé ces résultats avec des mesures où Ku est co-incubé dans la cellule de mesure avec un excès d'IP6 et l'ADN, soit un K_D environ 10 fois inférieur au K_D obtenu pour l'interaction Ku-ADN. L'IP6 stimule de façon significative l'affinité entre Ku et l'ADN. Ces mesures et la structure cristalline apportent un éclairage nouveau sur la stimulation de la voie NHEJ par l'interaction de Ku et l'IP6 (L. A. Hanakahi 2002; Cheung, Salerno, et Hanakahi 2008). Elles suggèrent que l'IP6 permettrait un recrutement plus fort de Ku aux extrémités de l'ADN et des cassures double-brin en augmentant son affinité pour l'ADN. Cependant, le facteur c ici est trop important et la réaction est endothermique. Il semblerait donc judicieux de répéter les expériences en diminuant soit la concentration en Ku soit en baissant la température à 10 °C.

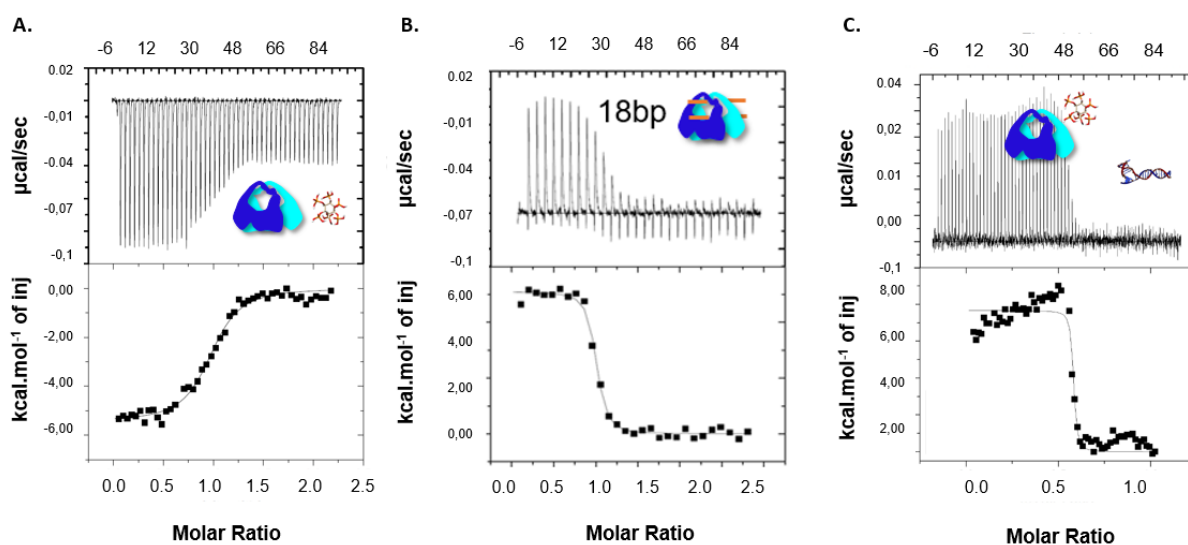


Figure 68 : Illustrations de résultats d'ITC (Gontier et al, 2019) ; en haut : thermogramme montrant des pics négatifs ou positifs témoignant respectivement d'interaction de type exothermique ou endothermique, en bas : isothermes de titration **A.** Interaction KuFL-IP6 ; 70 injections de 5 μ L. **B.** Interaction KuFL-ADN18 ; 30 injections de 10 μ L. **C.** Interaction KuCC/IP6-hDNA ; 70 injections de 2 μ L ; un *fit* avec un seul site a été réalisé. Une étude est en cours pour réaliser un *fit* à deux sites et tenir compte du début de l'isotherme

Cellule de mesure	Seringue d'injection	N	K _D (nM)	ΔH° (Kcal.mol ⁻¹)	-TΔS° (Kcal.mol ⁻¹)	ΔG° (Kcal.mol ⁻¹)
KuFL	IP6	0,8 ± 0,1	50 ± 0,5	-6,9 ± 0,2	-2 ± 0,3	-9,5 ± 0,2
KuCC	IP6	0,8 ± 0,3	27 ± 0,2	-6,2 ± 0,5	-3 ± 0,1	-9 ± 0,4
KuFL-hDNA	IP6	0,9 ± 0,1	21 ± 0,3	-4,2 ± 0,1	-6,1 ± 0,1	-10,5 ± 0,2
KuCC-hDNA	IP6	0,7 ± 0,1	50 ± 0,7	-6 ± 0,1	-4 ± 0,2	-10 ± 0,2
KuSC	IP6	0,4 ± 0,2	4 ± 0,5	-12 ± 0,6	-2,8 ± 0,1	-11,4 ± 0,6
KuCC-IP6	hDNA	0,7 ± 0,2	0,6 ± 0,1	7,7 ± 0,1	-19 ± 0,2	-12,2 ± 0,7

Nous avons alors souhaité analyser ces interactions par une autre méthode biophysique pour confirmer ou nuancer ces dernières observations.

Analyse par switchSENSE

La technologie du switchSENSE permet la caractérisation des interactions moléculaires notamment de type protéine-ADN. Nous avons en premier mesuré l'interaction entre Ku et un ADN de 48 pb greffé sur la puce qui sert de nanolevier. Nous avons mesuré l'interaction en mode statique avec le levier d'ADN relevé en présence d'un potentiel négatif sur la puce.

L'hétérodimère Ku est injecté avec une gamme de concentrations allant de 2 μM à 15,6 nM avec un facteur de dilution de 2. Les résultats sont analysés avec le logiciel switchANALYSIS. Dans une première étape, un ADN complémentaire de 48 nucléotides est injecté sur la puce et vient s'hybrider sur l'ADN simple brin quantitativement par augmentation de la fluorescence lors du passage de l'ADN simple brin à l'ADN double brin (Figure 69). La puce est alors fonctionnalisée et prête à recevoir la protéine Ku. Les cinétiques d'interaction aux différentes concentrations de Ku sont faites dans le tampon TE140 avec un débit de 200 μL.min⁻¹ pour l'association et la dissociation. Le temps d'association du Ku sur l'ADN est de 1 min et le temps de dissociation de 5 min. Le K_D obtenu pour l'interaction KuCC-ADN48pb est de 27,3 ± 0,9 nM en utilisant un ajustement (fit) sur 2 courbes. Les paramètres observés sont un k_{ON} de 2,78 ± 0,08 10⁵ M⁻¹.s⁻¹ et un k_{OFF} de 7,57 ± 0,14 10⁻³ s⁻¹ (Figure 70A). Ces valeurs sont à comparer au K_D de 2 μM mesuré par ITC entre KuCC et un ADN de 42 pb. Ces différences peuvent avoir plusieurs origines. La fixation de l'ADN sur une surface peut limiter l'accessibilité de Ku qui, en solution a la possibilité d'entrer par les deux extrémités. La présence d'un fluorophore à

l'extrémité de l'ADN peut également représenter une gêne stérique pour l'entrée de Ku. Les mesures de K_D , k_{ON} et k_{OFF} sont effectuées en présence de KuCC co-incubé avec la molécule d'IP6 et un ratio 1 : 5 entre Ku et l'IP6. Les gammes de concentration choisies en KuCC-IP6 sont données ci-dessous :

Concentrations (nM)	
KuCC	IP6
960	4800
480	2400
240	1200
120	600
60	300

Après chaque cinétique, une étape de régénération est effectuée. Elle permet de dissocier l'ADN complémentaire et les molécules de Ku encore fixées. Une étape d'hybridation de l'ADN double brin est effectuée et contrôlée avant chaque nouvelle concentration en KuCC-IP6.

L'affinité entre KuCC et l'ADN est environ 10 fois plus forte en présence qu'en absence d'IP6. En effet, nous mesurons un K_D de 240 ± 8 pM avec un k_{ON} de $3,14 \pm 0,05 \cdot 10^6 \text{ M}^{-1} \cdot \text{s}^{-1}$ et un k_{OFF} de $7,52 \pm 0,22 \cdot 10^{-4} \text{ s}^{-1}$. Les temps d'association et de dissociation sont respectivement de 40 sec et 10 min si l'on observe les plateaux de saturation du signal de fluorescence. Le gain en affinité provient d'une cinétique d'association plus rapide et d'une cinétique de dissociation plus lente. Le k_{ON} observé en présence de l'inositol phosphate est environ 10 fois plus élevé qu'en absence de celui-ci et la valeur de k_{OFF} est environ 10 fois plus faible. Afin de s'assurer que le signal de fluorescence observé soit bien dû à l'interaction de KuCC avec l'ADN, une cinétique est effectuée en injectant l'IP6 seul en excès sur l'ADN double brin. Aucun signal de fluorescence n'est mesuré en absence de protéine KuCC (Figure 70C). Ces données suggèrent que l'IP6 permettrait un recrutement plus rapide de la protéine Ku sur l'ADN et stabiliserait l'interaction. La structure cristalline suggère que l'IP6 pourrait permettre de neutraliser une région basique de Ku située à proximité de l'entrée de l'anneau et de modifier la distribution de charges électrostatiques dans cette région pour favoriser l'attraction vers le site principal à savoir l'anneau basique de Ku. Nous observons une variation supérieure sur les K_D avec ou sans IP6, supérieure d'un facteur 10 entre la calorimétrie et le switchSENSE. Cette variation peut

provenir du fait que le switchSENSE est une technique où l'ADN est fixé sur une surface alors qu'il est en solution pour l'ITC. La calorimétrie est effectuée en solution, sans marquage fluorescent sur l'ADN et sur des temps de mesures d'environ 120 min. Le switchSENSE utilise un ADN greffé sur une surface, marqué avec une sonde de fluorescence et sur des temps de mesures d'association de 5 min et de dissociation de 10 min.

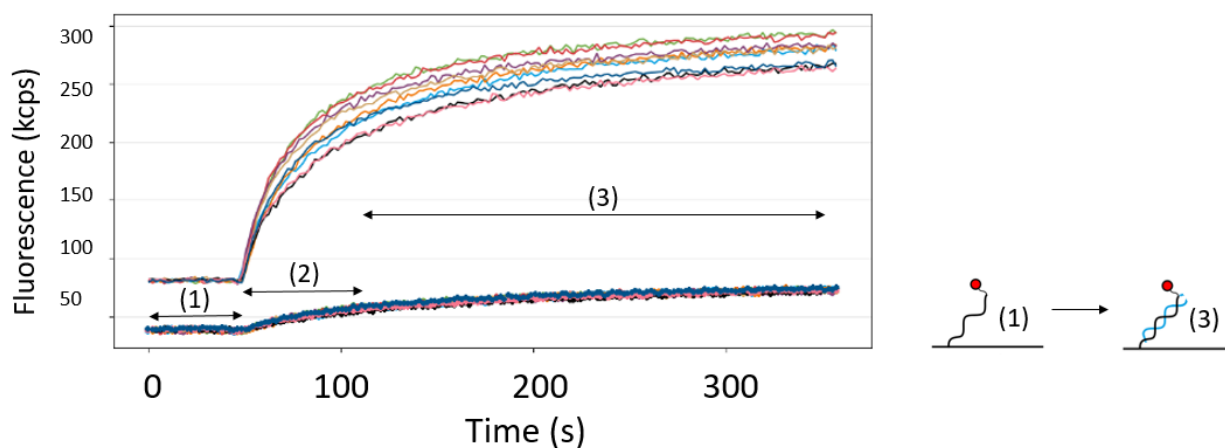


Figure 69 : Courbes de fluorescence obtenues lors 8 hybridations successives de l'ADN complémentaire de 48 pb sur la puce MPC2-48-2-R1-S. Les différentes couleurs à l'état d'hybridation après chaque régénération et entre chaque concentration en Ku. (1) : Profil de fluorescence lorsque l'ADN est simple brin ; (2) : Profil de fluorescence lorsque l'ADN 48 pb de la puce commence à être hybridé : c'est le début de la fonctionnalisation de la puce ; (3) : Apparition du plateau de fluorescence : hybridation complète de l'ADN 48 pb présente sur la puce

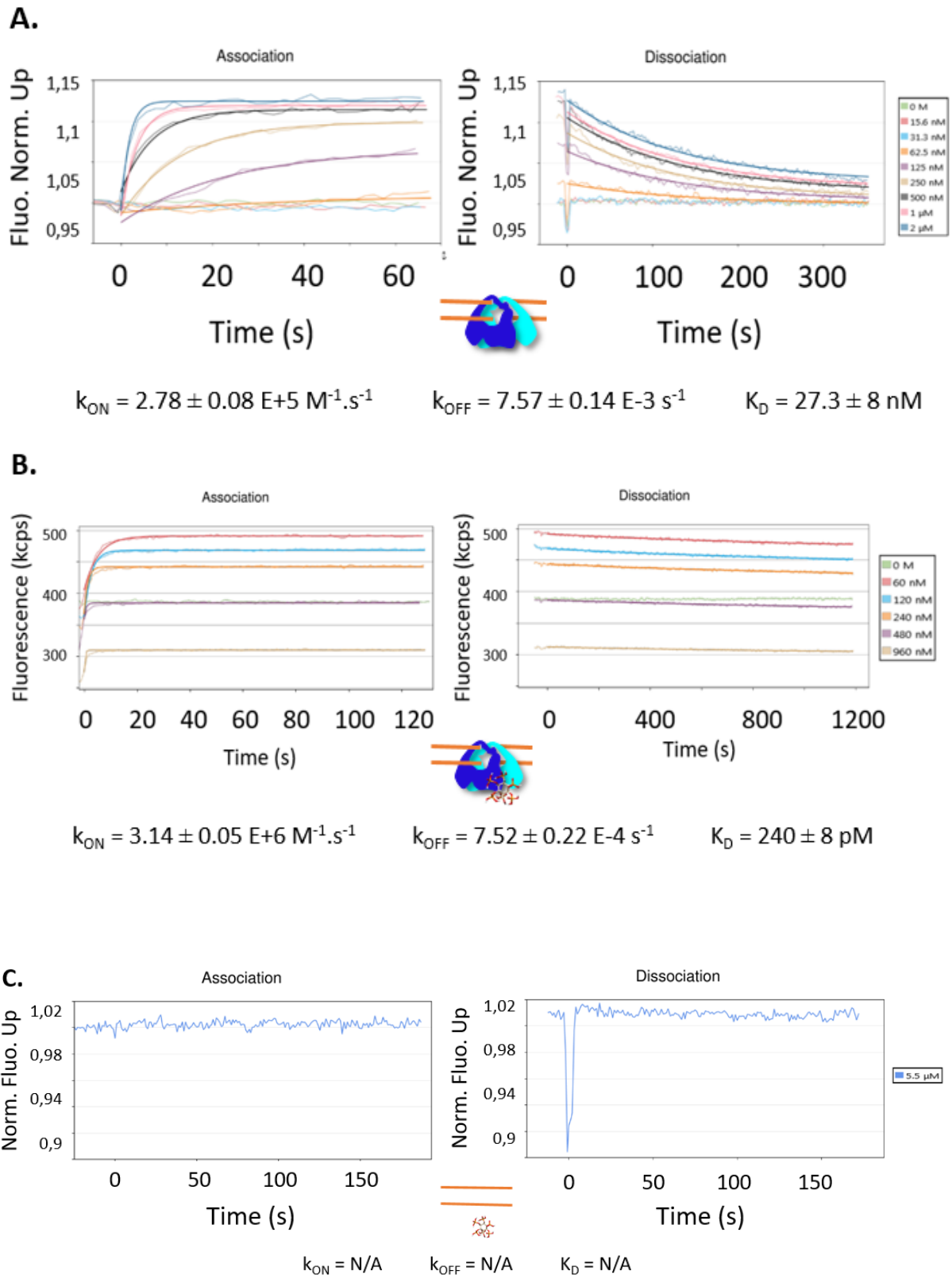


Figure 70 : Analyses de cinétique par switchSENSE grâce au logiciel switchANALYSIS. **A.** Caractérisation de l'interaction KuCC-ADN48 en mode statique. La dissociation est effectuée sur 1 200 s en utilisant une pompe péristaltique montée sur le DRX². **B.** Caractérisation de l'interaction KuCC-ADN48 en mode statique en présence d'IP6 incubé avec la protéine Ku. **C.** Caractérisation de l'interaction IP6-ADN48 en mode statique

Analyse par HT-DSF

La technique de DSF permet de suivre par fluorescence la température de dénaturation de la protéine Ku en suivant son état de déploiement (*unfolding*) suite à l'application d'un gradient de température. Nous avons montré par calorimétrie et switchSENSE que l'IP6 permettait d'augmenter l'affinité entre Ku et l'ADN d'un facteur respectivement 10 et 100. Après avoir réalisé des tests préliminaires de faisabilité (détails dans la partie « Matériel et Méthodes »), la protéine est concentrée à $0,45 \text{ mg.mL}^{-1}$ et la sonde fluorescente de SYPRO orange est utilisée à une concentration X15. Deux températures de dénaturation sont observées avec la protéine seule. Un premier T_m est observé à $30,65 \text{ }^\circ\text{C}$ (T_{m1}) et un second T_m est observé à $49,48 \text{ }^\circ\text{C}$ (T_{m2}) (Figure 71). La structure 3D de l'hétérodimère Ku70/Ku80 suggère que ces deux T_m pourraient correspondre à la dénaturation des domaines vWA de Ku70 et Ku80 et à la dénaturation des domaines de dimérisation.

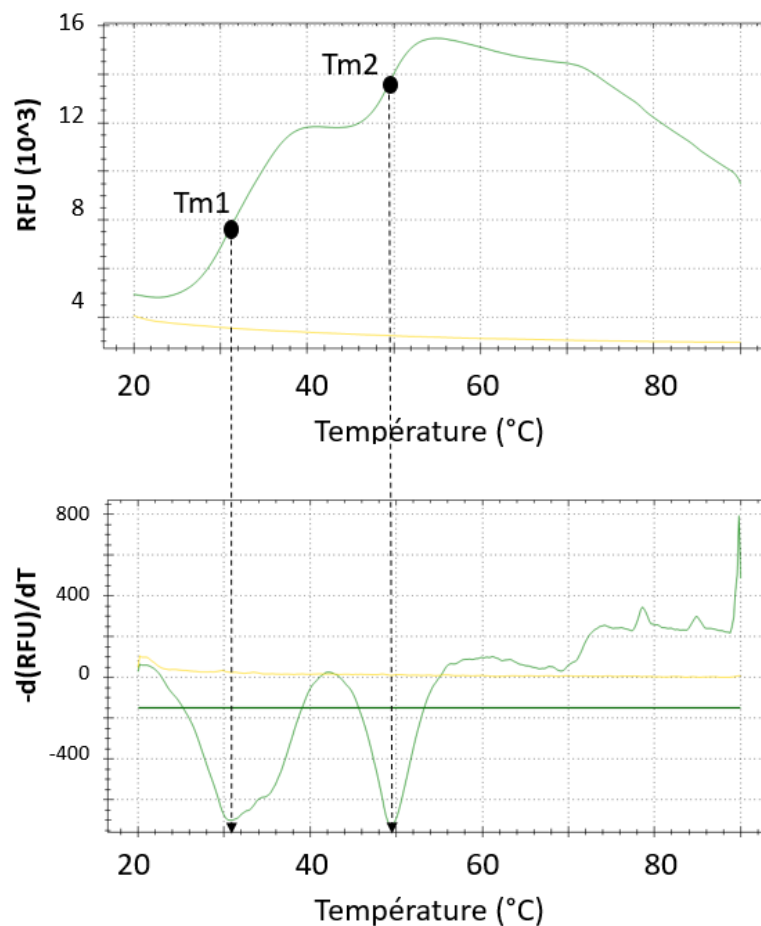


Figure 71 : DSF : signaux de fluorescence émis par la sonde fluorescente de SYPRO orange à X15 lorsque la protéine KuCC est chauffée de $20 \text{ }^\circ\text{C}$ à $90 \text{ }^\circ\text{C}$ selon un gradient de température de $0,5 \text{ }^\circ\text{C} / \text{min}$. **En haut** : signal de fluorescence et **en bas** : dérivée du signal de fluorescence

Après ajout de l'inositol phosphate (IP6) à différentes concentrations (de 5,9 nM à 3 μ M selon un facteur de dilution de 2), un déplacement de T_{m1} est observé à partir d'une concentration en IP6 d'environ 1,5 μ M. Ce T_{m1} est déplacé vers des valeurs supérieures. Les ΔT_m observés vont de + 9,1 $^{\circ}$ C pour une concentration en IP6 à 3 μ M ($T_{m1Ku-IP6} = 39,8$ $^{\circ}$ C) à + 0,1 $^{\circ}$ C pour une concentration en IP6 à 1,5 μ M ($T_{m1Ku-IP6} = 30,8$ $^{\circ}$ C) (Figure 72). Aucun déplacement de T_{m2} n'est observé. L'IP6 semble donc être une molécule stabilisatrice de l'hétérodimère Ku.

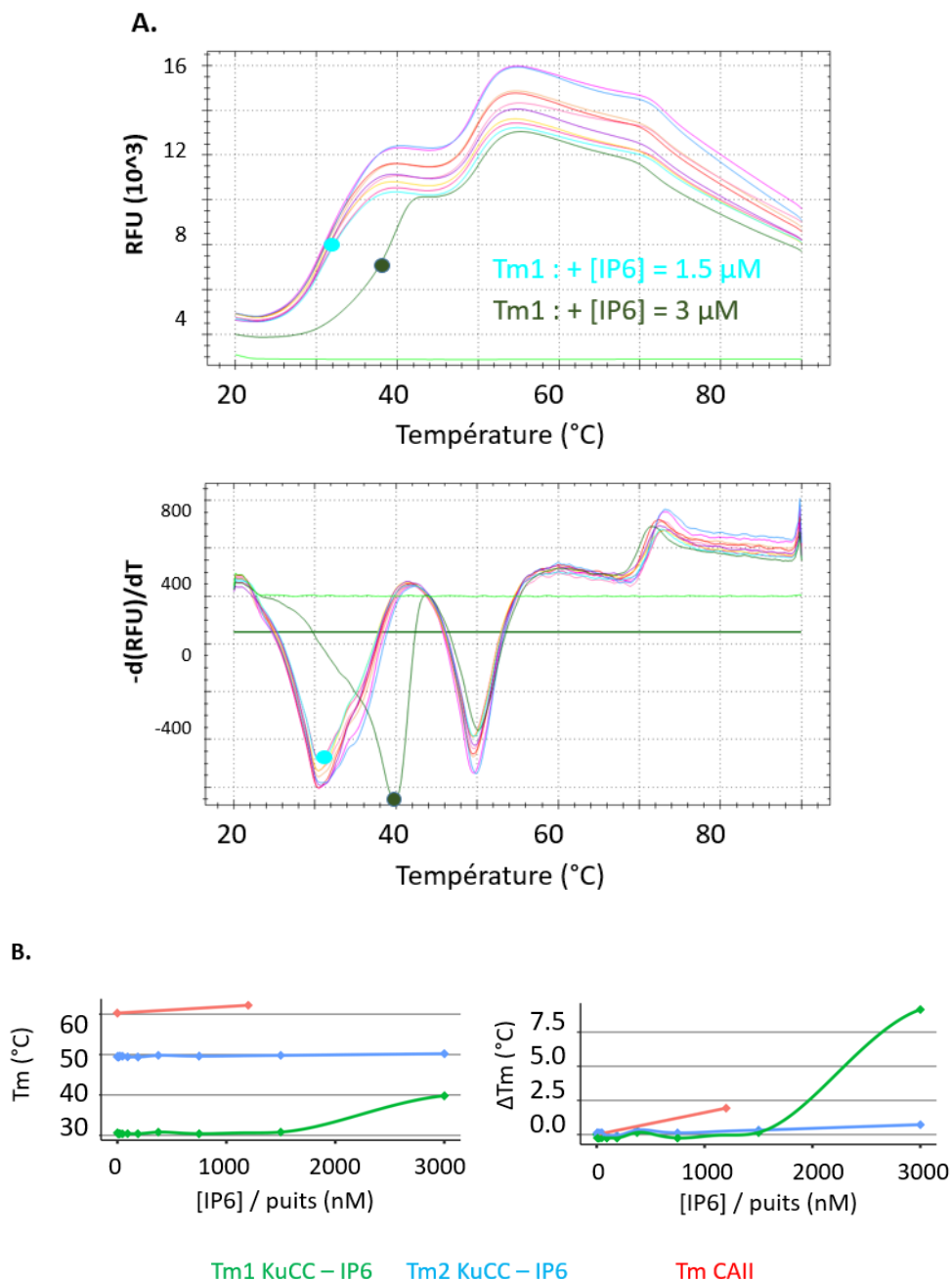


Figure 72 : Résultats de DSF pour le complexe KuCC-IP6 avec IP6 à différentes concentrations. **A.** Signaux de fluorescence émis par le fluorophore de SYPRO orange à X15 lorsque la protéine KuCC est chauffée de 20 $^{\circ}$ C à 90 $^{\circ}$ C selon un gradient de température de 0,5 $^{\circ}$ C / min en présence d'IP6 ; En haut et en bas respectivement avant et après dérivation de la fluorescence. **B.** Résultats de DSF pour KuCC-IP6 sous forme de courbe des T_m ou des ΔT_m en fonction de la concentration en IP6 par puits (nM)

Analyse par sm-TIRF

Une collaboration avec l'équipe de Eli Rothenberg (Université de New York) a été établie afin de réaliser de expériences de sm-TIRF (pour *single-molecule Total Internal Reflection Fluorescence* en anglais). Le sm-TIRF est une méthode de microscopie à fluorescence permettant de visualiser des molécules fluorescentes localisées à proximité d'un interface verre/échantillon. Pour ces expériences, la protéine Ku est marquée avec une sonde fluorescente Cy5. Le Cy5 est excité directement avec le laser rouge. L'arrivée de Ku sur l'ADN est observée avec apparition de fluorescence près de la surface de verre (Figure 73). En absence d'IP6, Ku interagit avec l'ADN mais s'en dissocie très rapidement. En présence d'IP6 à 100 nM, la protéine interagit environ 2 fois plus longtemps aux extrémités de l'ADN double brin. Ces résultats rejoignent ceux de calorimétrie où un gain d'un facteur 10 a été mesuré. Ils sont en accord avec le switchSENSE où nous avons vu que le temps d'association de Ku sur l'ADN en présence d'IP6 était plus court et le temps de dissociation plus long même si l'effet observé sur le switchSENSE était plus marqué.

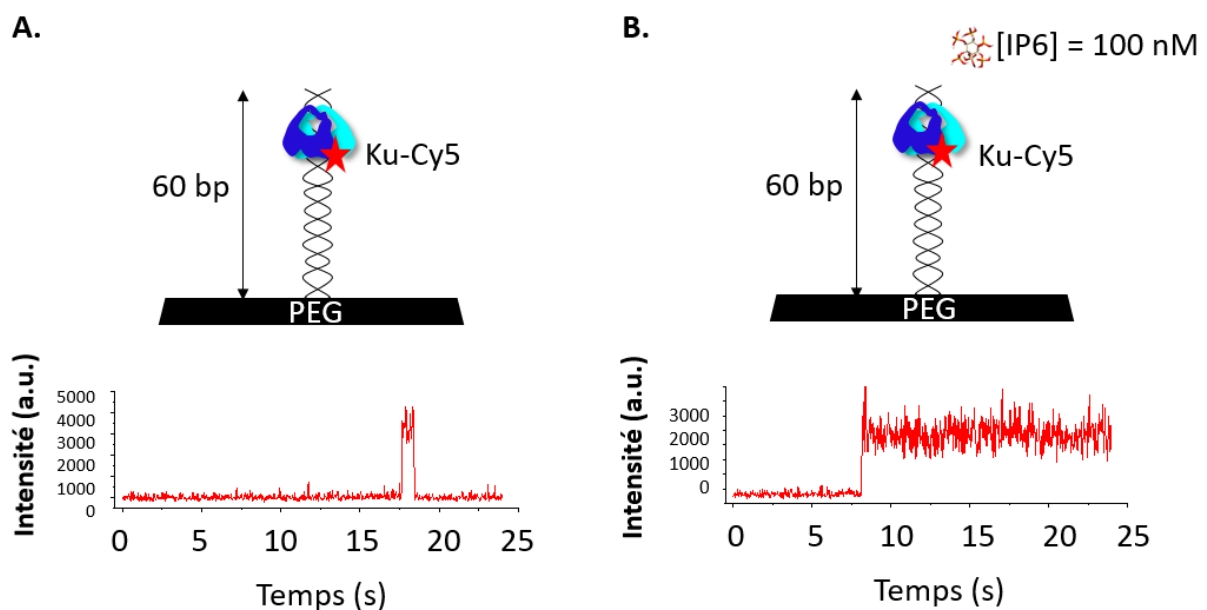


Figure 73 : Résultats d'expériences de sm-TIRF ; visualisation de l'intensité de fluorescence émise par la sonde fluorescente Cy5 accrochée à la protéine Ku en fonction des temps de traces. L'ADN double brin d'une longueur de 60 pb est immobilisé sur une surface PEG. **A.** Exemple de trace de fluorescence observée en absence d'IP6. **B.** Exemple de trace de fluorescence observée en présence d'IP6

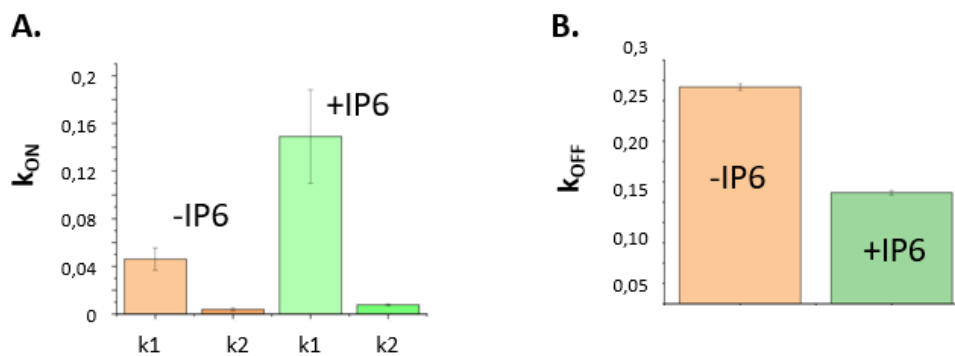


Figure 74 : Histogrammes représentant les valeurs de k_{ON} ($M^{-1}.s^{-1}$) et de k_{OFF} (s^{-1}) de la protéine Ku sur l'ADN double brin en présence ou en absence de 100 nM d'IP6. **A.** Représentation du k_{ON} ($M^{-1}.s^{-1}$). **B.** Représentation du k_{OFF} (s^{-1})

Les hypothèses de départ sont donc confirmées ici : l'IP6 permettrait un recrutement plus rapide du partenaire Ku aux cassures double brin et stimulerait son interaction en freinant la dissociation du complexe Ku-ADN.

d) Interactions simultanées de l'IP6 et du peptide pXLF

Les structures cristallographiques KuCC-hDNA-pXLF et KuCC-hDNA-IP6 montrent que le peptide pXLF et la molécule d'IP6 sont localisés sur des sites situés à proximité sur Ku80. L'interaction de pXLF avec Ku, s'accompagne d'une ouverture du domaine vWA de Ku80. Nous nous sommes demandés si l'IP6 a un rôle dans l'interaction entre la protéine XLF et la protéine Ku.

Analyse par HT-DSF

Les conditions d'expériences sont les mêmes concernant la concentration en protéine Ku (0.45 mg/mL) et en SYPRO orange (X15).

Après ajout du peptide pXLF à différentes concentrations (de 5,9 nM à 3 μ M selon un facteur de dilution de 2), un déplacement de T_{m1} est observé à partir d'une concentration en pXLF d'environ 1,5 μ M. Ce T_{m1} est déplacé vers des valeurs supérieures (de + 0,8 $^{\circ}$ C) pour une concentration en pXLF à 3 μ M (Figure 75). Un déplacement de T_{m2} est observé également même si celui-ci est moindre que pour le T_{m1} . Le peptide pXLF stabilise l'hétérodimère Ku mais de façon moindre que l'IP6 (+ 9,1 $^{\circ}$ C à 3 μ M en IP6).

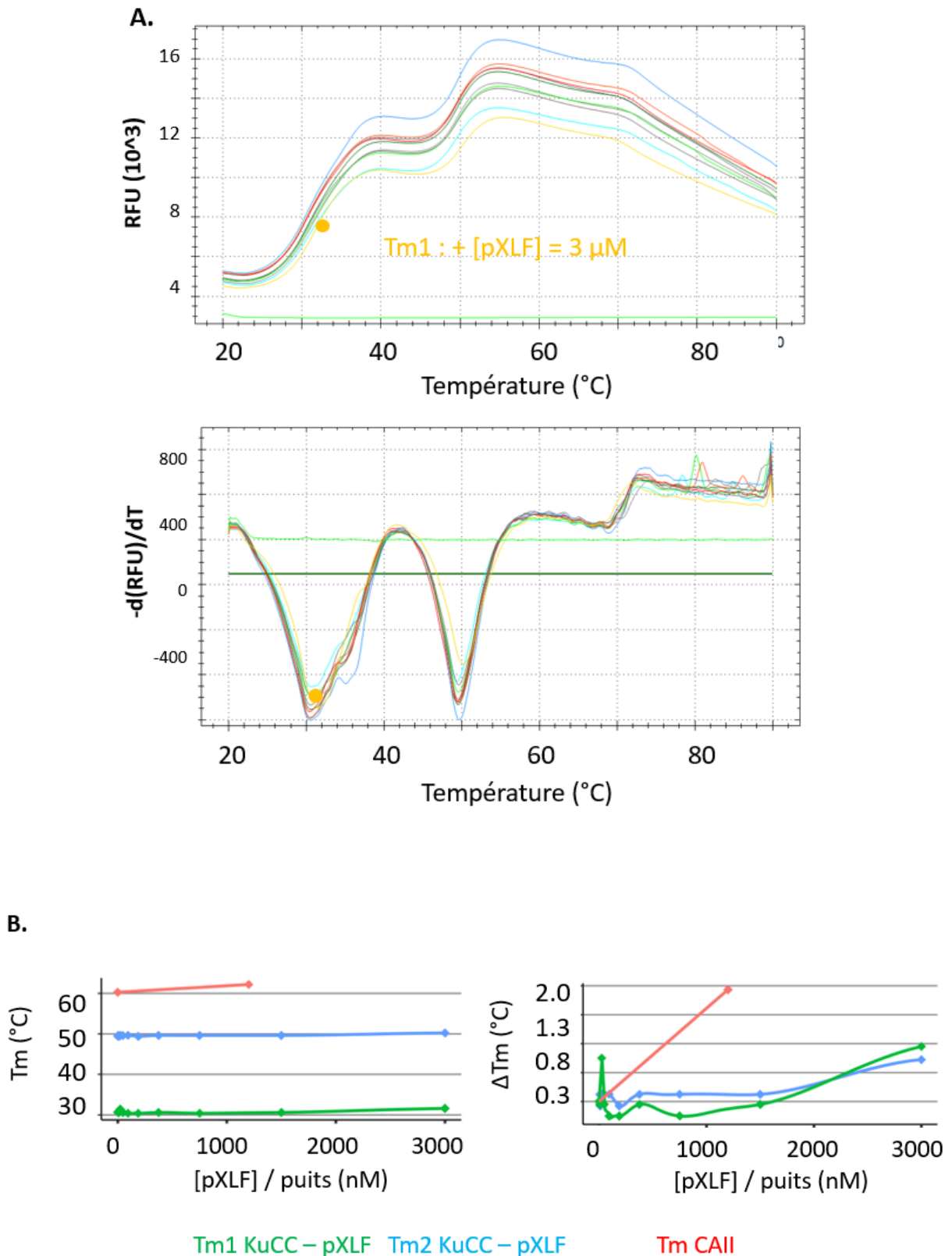


Figure 75 : Résultats de DSF pour le complexe KuCC-pXLF avec pXLF à différentes concentrations (de . **A.** Signaux de fluorescence émis par le fluorophore de SYPRO orange lorsque la protéine KuCC est chauffée de 20 °C à 90 °C selon un gradient de température de 0,5 °C / min en présence de pXLF ; En haut et en bas respectivement la fluorescence et dérivée de la fluorescence. **B.** Résultats de DSF pour KuCC-pXLF sous forme de courbe des T_m ou des ΔT_m en fonction de la concentration en pXLF par puits (nM)

Analyse par microcalorimétrie

Lors de ma participation sur la première étude (Nemoz et al. 2018), j'ai effectué des mesures de calorimétrie pour connaître les propriétés cinétiques et thermodynamiques de l'interaction entre Ku et la protéine XLF ou des peptides dérivés de XLF et Ku et des peptides dérivés d'un autre partenaire la protéine APLF. Dans cette étude, les mesures de calorimétrie sont combinées à d'autres techniques (gel retard et switchSENSE).

Dans ce contexte, j'ai étudié deux mutants (L297E et L297W) qui correspondent à la mutation d'un des acides aminés principaux du KBM de XLF, la Leucine 297. Le peptide pXLF (le X-KBM) a été utilisé dans un premier temps à la place de la protéine entière. La protéine XLF est dans la seringue d'injection à une concentration dix fois supérieure à la protéine Ku (avec ou sans ADN) mise dans la cellule de mesure afin de permettre la titration. En absence d'ADN, la réaction est de type exothermique (dégagement de chaleur) et conduite par l'enthalpie ($\Delta H = -9,0 \text{ kcal.mol}^{-1}$). L'affinité mesurée pour ce complexe est de l'ordre du micromolaire, donc assez faible ($K_D = 1,0 \pm 0,1 \text{ } \mu\text{M}$). L'affinité entre le peptide pXLF contenant le X-KBM et la protéine Ku (avec ou sans ADN) est de $4,4 \pm 0,2 \text{ } \mu\text{M}$. L'affinité est donc légèrement meilleure pour la protéine même si un facteur 4,4 correspond à une énergie libre additionnelle assez faible (une liaison hydrogène a une énergie entre $0,8 \text{ kcal.mol}^{-1}$ et $1,3 \text{ kcal.mol}^{-1}$). En présence d'ADN, la valeur d'affinité ne change pas. Ces résultats suggèrent que la fixation de l'ADN n'a pas d'influence significative sur la fixation du peptide pXLF et l'ouverture consécutive du domaine vWA situé à $20 \text{ } \text{Å}$ de l'ADN par Ku70/Ku80.

Nous avons effectué ces mesures par switchSENSE en mode statique et comparé à la protéine sauvage. Nous avons mesuré pour XLF(wt) un K_D de $0,19 \pm 0,07 \text{ } \mu\text{M}$ (avec un k_{ON} de $4,7 \pm 1,7 \text{ } 10^5 \text{ M}^{-1}.\text{s}^{-1}$ et un k_{OFF} de $0,09 \pm 0,004 \text{ s}^{-1}$). L'affinité observée est environ 10 fois plus forte que celle mesurée par ITC avec un ADN plus petit. Ceci pourrait provenir d'interactions additionnelles de XLF avec l'ADN comme cela avait été observé pour un autre partenaire de Ku ; PAXX (Tadi et al. 2016). Nous obtenons des valeurs de K_D pour L297E et L297W respectivement de $0,98 \pm 0,15 \text{ } \mu\text{M}$ et de $0,45 \pm 0,26 \text{ } \mu\text{M}$. Ces deux mutants présentent donc des affinités respectivement 2,3 fois et 5,1 fois plus faibles que la protéine sauvage. Les mesures d'affinité avec les peptides ont été mesurées par calorimétrie. Dans ce cas, nous observons des différences plus marquées avec un peptide pXLF(L297E) qui ne présente plus d'interaction avec Ku et un peptide pXLF(L297W) qui présente une affinité meilleure que le peptide sauvage. Les peptides ne résument donc pas à eux seuls les propriétés d'interaction de XLF avec Ku.

Nous avons alors réalisé des expériences de compétition entre le peptide pXLF et l'IP6 pour évaluer si la fixation d'une des molécules pouvait avoir une influence sur l'autre. L'affinité mesurée de IP6 contre Ku incubé avec un excès de pXLF (dans la cellule de mesure) est de 132 ± 19 nM. Cette affinité est légèrement plus faible que celle mesurée entre IP6 et Ku seul ($K_D = 38 \pm 0,1$ nM). La différence entre les deux est toutefois faible. En parallèle, nous avons injecté la molécule d'IP6 contre pXLF dans la cellule. Nous n'observons pas de signal d'interaction. (Figure 76).

En parallèle, j'ai voulu tester s'il était possible de voir en ITC, la formation d'une jonction entre deux complexes Ku-ADN avec un ADN18 bp. Il est proposé dans la littérature, que Ku pourrait faire cette jonction et dans d'autres articles que cette jonction ne se ferait que si DNA-PKcs est à l'endroit où certains facteurs du NHEJ interviennent en amont (Lig4-XRCC4-XLF ou les PolX). Nous n'observons en ITC aucune interaction entre deux molécules Ku-ADN18 lorsque nous mettons en interaction Ku-ADN18 (dans la cellule de mesure et dans la seringue d'injection). Dans les conditions expérimentales choisies (complexe à $10 \mu\text{M}$ dans la cellule et à $100 \mu\text{M}$ dans la seringue) cela suggère que l'interaction si elle existe, a probablement un K_D supérieur à $10 \mu\text{M}$ ou que l'interaction se fait à enthalpie nulle (pas de dégagement de chaleur à la température de mesure choisie).

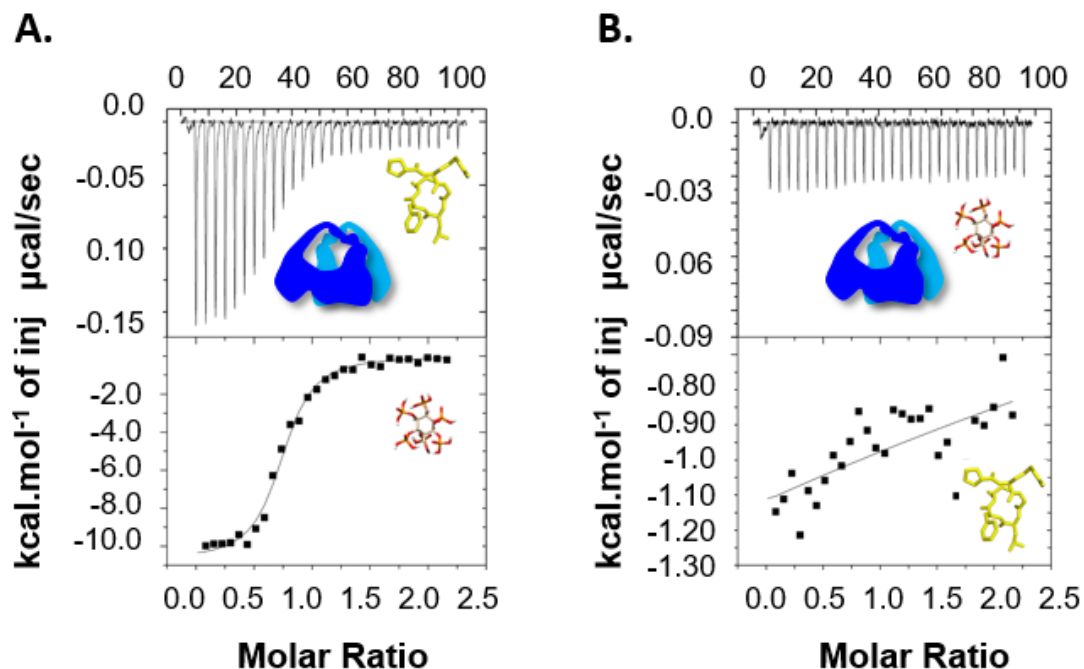


Figure 76 : Illustrations de résultats d'ITC avec compétitions entre l'IP6 et pXLF. **En haut :** thermogramme montrant des pics négatifs témoignant d'une interaction de type exothermique. **En bas :** isotherme de titration avec des paramètres proches de ceux observés pour l'interaction Ku-IP6. **A.** Interaction KuCC/pXLF-IP6. **B.** Interaction KuFL/IP6-pXLF

Analyse par sm-FRET

L'équipe de Eli Rothenberg a réalisé d'autres expériences de sm-FRET (pour *single molecule Fluorescence Resonance Energy Transfer* en anglais) pour visualiser l'appariement de deux complexes Ku-ADN (Figure 77). Un des deux complexes est immobilisé sur une surface PEG et l'autre complexe en solution avec d'autres partenaires du NHEJ (XLF, MRI, XRCC4, LigIV et IP6). Les ADN linéaires sont marqués avec un fluorophore Cy5 (pour l'ADN immobilisé) et avec un Cy3 (pour l'ADN en solution).

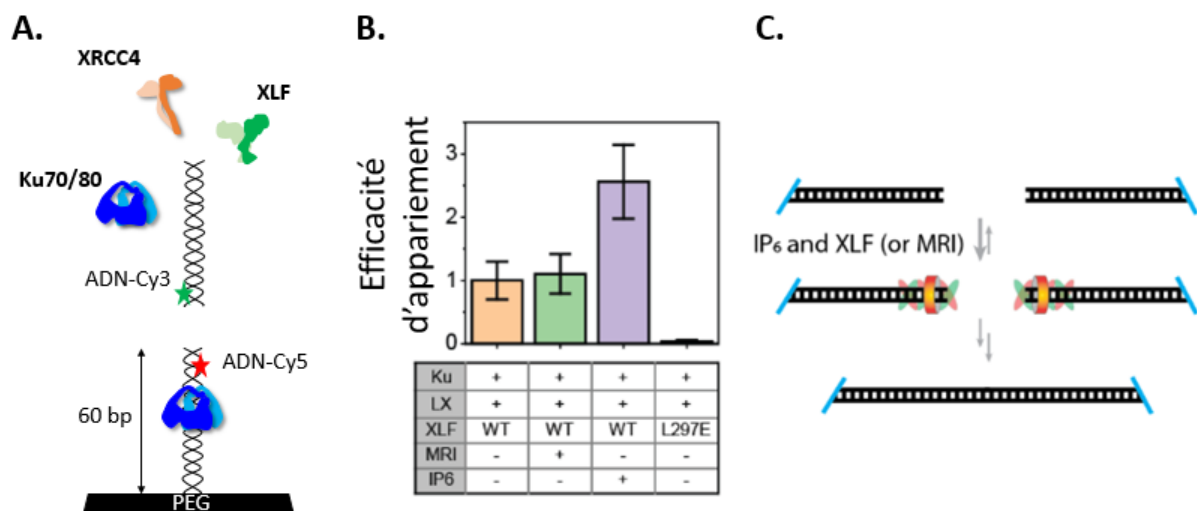


Figure 77 : Résultats d'expériences de sm-FRET ; visualisation de l'intensité de fluorescence émise par la sonde acceptrice accrochée à l'ADN en solution lorsque les deux ADN sont rapprochés et qu'il y a appariement. **A.** Schéma explicatif de l'expérience de sm-FRET. **B.** Efficacité d'appariement des extrémités d'ADN. **C.** Schéma hypothétique de l'appariement des extrémités et rôle de l'IP6

C'est la fluorescence émise par la sonde acceptrice fixée sur l'ADN en solution qui est mesurée. En présence de Ku, du complexe Lig4-XRCC4 et de XLF, on observe un signal de fluorescence lié à la fixation de l'ADN-Cy3 à proximité de l'ADN fixé sur le support. Lorsque que l'on ajoute un partenaire de Ku, MRI, on ne voit pas d'effet additionnel. En revanche, l'ajout d'inositol phosphate (IP6) augmente l'efficacité de l'appariement des extrémités d'ADN d'un facteur 2,5. Ces résultats viennent compléter ceux obtenus précédemment. Ils confirment un effet stimulateur de l'IP6. Ils suggèrent en plus un rôle de cette molécule dans la synapse faite par Ku, l'ADN et le complexe Lig4-XRCC4 et XLF.

PARTIE II :

Stratégies d'inhibition du NHEJ : un croisement de méthodes

Le projet de thèse CIFRE s'est inscrit dans une collaboration étroite entre le service « *Bio-Structure and Biophysics* » localisé à Vitry-Sur-Seine chez Sanofi et l'I2BC au CEA de Saclay. Un volet non prévu initialement dans le projet de thèse a émergé au bout d'une année sur l'utilisation du savoir-faire de Sanofi en terme de crible sur la cible thérapeutique étudiée au CEA, la protéine Ku de la voie du NHEJ. Une étude de criblage de 987 fragments a été réalisée sur la protéine en parallèle de l'étude sur l'IP6 et des autres partenaires de Ku. Le but était d'identifier les fragments susceptibles d'interagir avec Ku. Ces fragments seront ensuite cristallisés avec Ku pour déterminer le site d'interaction avec cette protéine. Cette approche par criblage de fragments permet d'obtenir à terme, un composé de plus forte affinité pour la protéine Ku et ainsi d'en faire un nouvel inhibiteur du NHEJ.

A) Un criblage d'une collection de fragments Sanofi sur Ku par DSF

La notion de *Fragment-Based Drug Discovery* FBDD (Scott et al. 2012) est utilisée dans cette étude pour cribler la protéine Ku, et, identifier des inhibiteurs potentiels de la voie NHEJ à l'aide d'une collection de fragments Sanofi. La technique HT-DSF (Renaud et al. 2016) est couramment utilisée car cette approche est rapide, peu coûteuse et permet de cribler des composés de bas poids moléculaires. La construction protéique utilisée est celle de KuCC car c'est la seule qui a donné des cristaux au laboratoire.

L'étape de faisabilité a été réalisée avec la protéine KuCC réalisant des mesures avec l'IP6 et le peptide pXLF. La concentration optimale de KuCC est de 3,4 μ M. Les fragments ont été testés à une concentration de 1 mM. Dans cette partie, les informations sur certains fragments ne seront pas divulguées pour des raisons de propriété intellectuelle. La protéine KuCC possède deux températures de dénaturation T_{m1} et T_{m2} . La première température correspond à l'ouverture de l'hétérodimère des deux sous-unités Ku70 et Ku80. La seconde température est due à la dissociation et à la dénaturation de ces deux sous-unités. Nous avons choisi de nous

intéresser au T_{m1} . La première étape de criblage a permis d'obtenir 49 composés déplaçant le $T_{m1} + 0,5$ °C. Les composés entraînant une élévation du T_{m1} sont considérés comme stabilisateurs. L'aptitude d'un composé à stabiliser thermiquement une protéine est considérée comme une indication indirecte de sa capacité à interagir avec celle-ci. Les fragments sont classés en fonction de leur effet sur le T_{m1} et de la concentration minimale pour laquelle ils montrent cet effet. Le ΔT_{m1} le plus élevé est de + 7 °C. Dans un second temps, les 49 composés sélectionnés à une concentration de 1 mM ont été criblés par HT-DSF sur la protéine KuCC avec une gamme de concentrations de 2,5 mM à 4,9 μ M avec un facteur de dilution de 2 (9 valeurs mesurées). Cette étape est appelée mesure par « doses-réponse » ou « titration ». A l'issue de cette étape, 15 fragments ont été sélectionnés et classés en fonction de la concentration minimale pour laquelle ils stabilisent le plus la protéine (Figure 79). On observe plusieurs profils à l'issue de ce classement (Figure 78)

Profil n° 1 : Le ΔT_{m1} augmente en fonction de la concentration en fragment. Un plateau avec une valeur maximale du ΔT_{m1} . On parle d'effet stabilisateur

Profil n° 2 : Le ΔT_{m1} augmente puis diminue. Ceci indique un probable problème d'agrégation

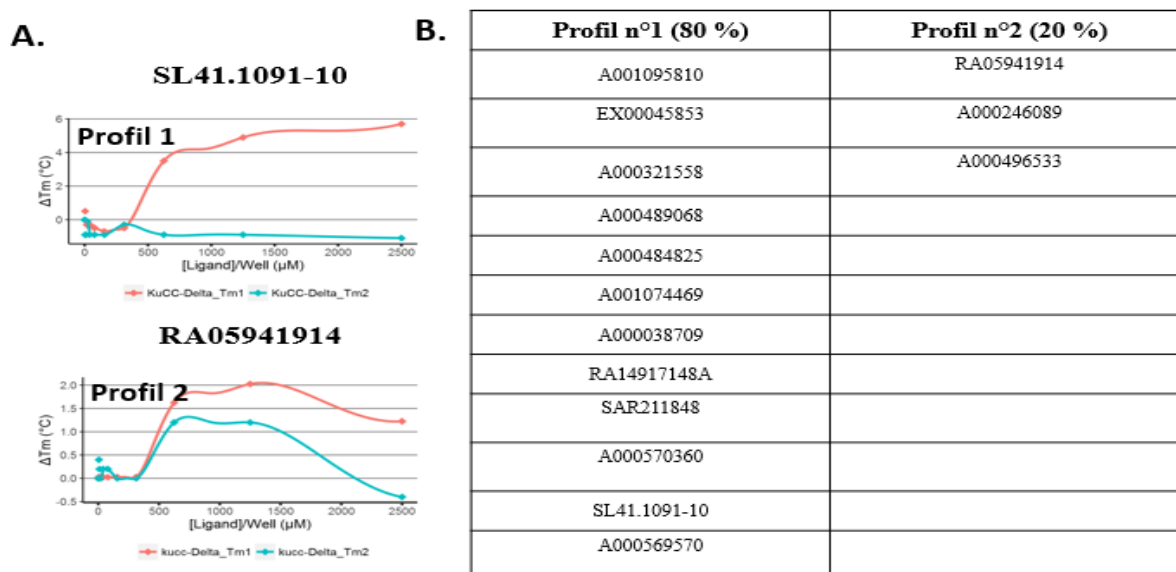


Figure 78 : Résultats des doses-réponse en HT-DSF suite au criblage des 15 fragments sélectionnés par DSF. **A.** Les deux types de profils observés après doses-réponse. **B.** Tableau résumant le nombre de fragments appartenant à chaque profil

Deux fragments « RA14917148A » et « EX00045853 » sur les 12 fragments stabilisateurs, étaient des molécules publiques. Ces deux fragments ont respectivement déplacé le T_{m1} de KuCC de + 1,5 °C (pour une concentration minimale en « RA14917148A » de 1,25 mM) et de + 3 °C (pour une concentration minimale en « EX00045853 » de 1,9 mM).

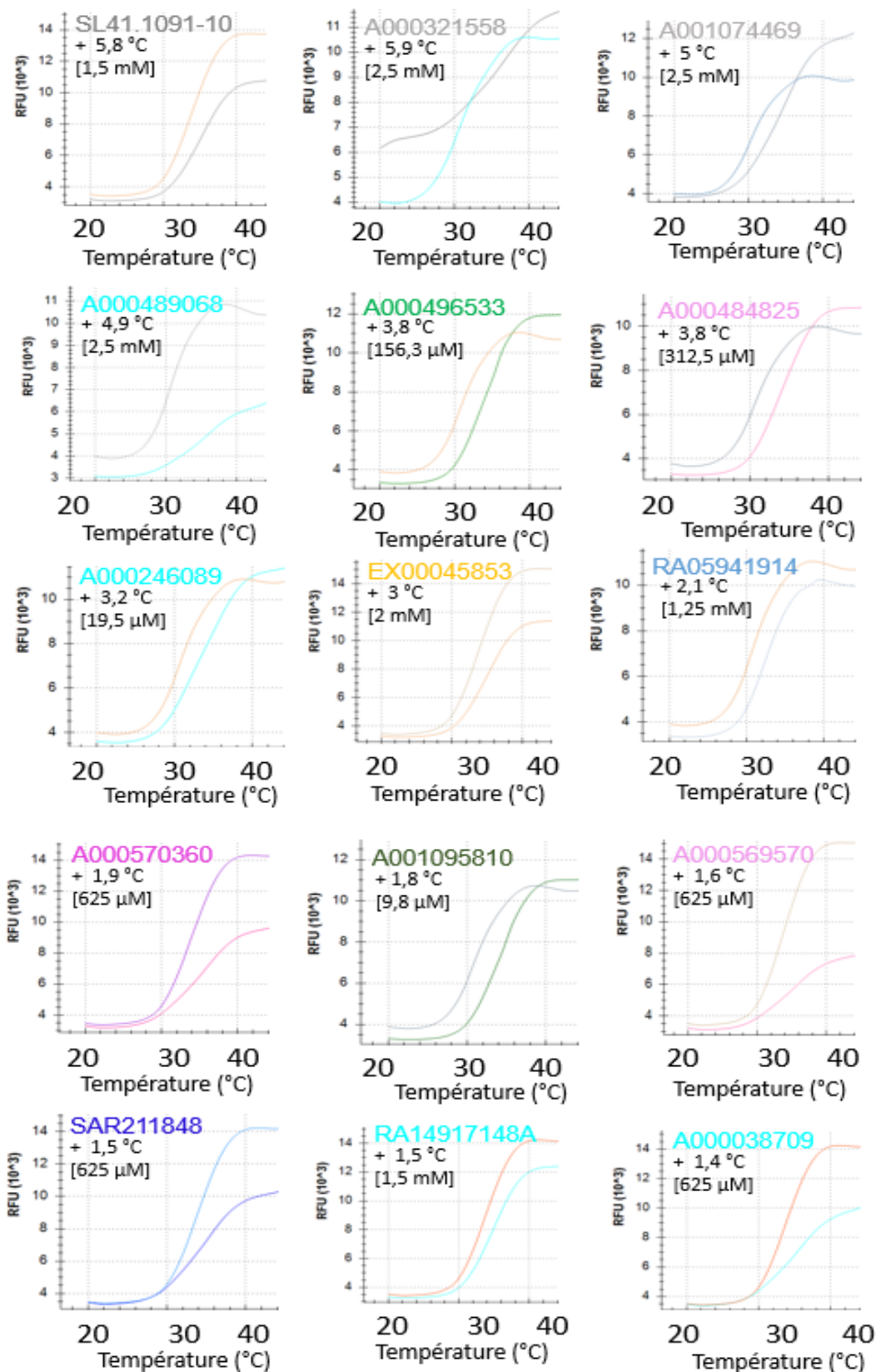


Figure 79 : Ces doses-réponses sont classées par effet stabilisateur décroissant pour une concentration minimale en fragment. Effet stabilisateur de chacun des 15 fragments sélectionnés suite aux doses-réponses avec la concentration du fragment à laquelle l'effet est le plus important. Visualisation du T_{m1} (une couleur est associée à chaque fragment testé, l'autre couleur est la courbe correspondant à la protéine seule avec 1 % de DMSO).

B) Etude d'interactions et de compétitions par RMN

Nous avons utilisé la RMN qui permet d'étudier les interactions protéine-ligand. Les expériences de type Water-LOGSY (Dalvit 2009) et STD (Mayer et Meyer, 2001) ont été utilisées.

Analyse du contrôle qualité des composés

Dans un premier temps, une analyse de QC a été réalisée sur le composé IP6 et les deux fragments pour vérifier leur intégrité structurale, leur solubilité, leur pureté ainsi que leur état d'agrégation en milieu aqueux en absence de protéine Ku. Des études d'interactions et de compétitions peuvent ensuite être réalisées sous plusieurs conditions. Les résonances des protons (ppm) doivent correspondre aux déplacements chimiques des protons de la structure du composé étudié, le composé doit être soluble en milieu aqueux, il doit être pur à plus de 95 % et enfin il ne doit pas s'agréger. La sonde fluor à une concentration de 500 μM et la référence DSS à une concentration de 100 μM , sont utilisées comme double contrôle respectivement pour quantifier et calibrer le spectre ^1H .

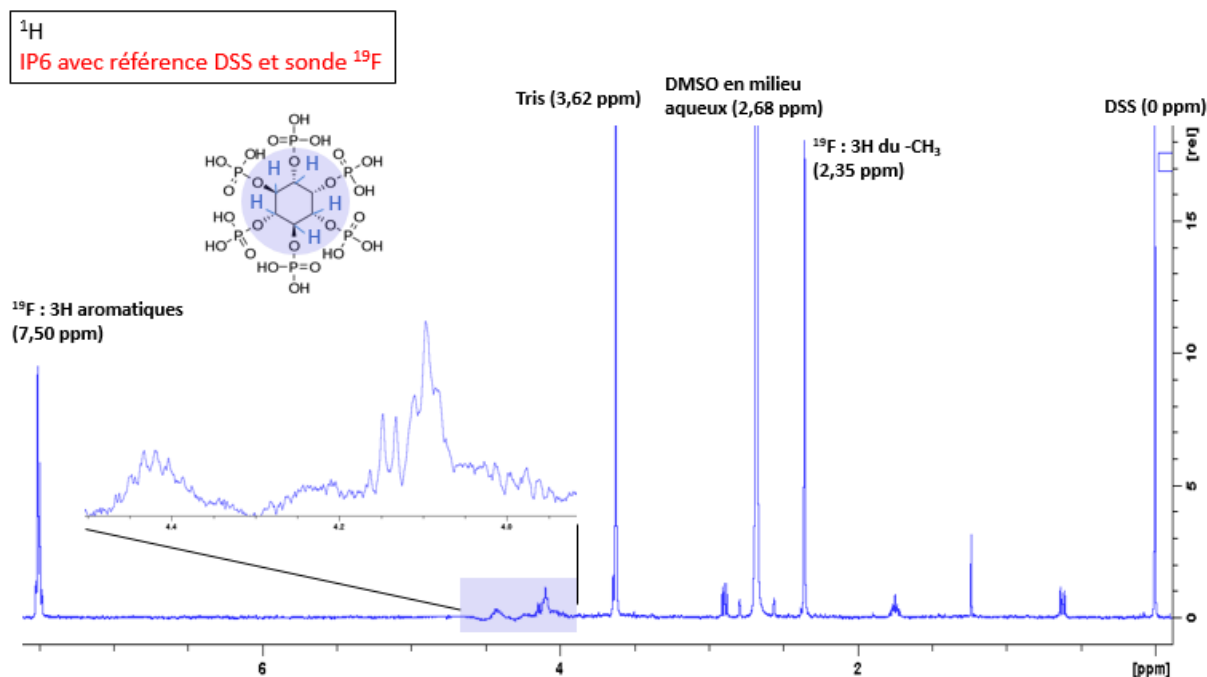


Figure 80 : Spectre ^1H pour le contrôle qualité de la molécule d'IP6, en présence de la sonde ^{19}F et de la référence DSS

Le spectre ^1H (Figure 80) montre des pics de résonance entre 4,2 ppm et 4,4 ppm correspondant aux hydrogènes du cycle inositol dans la zone bleue. Le spectre ^1H (Figure 81) montre des pics

de résonance principalement entre 0,8 ppm et 1,3 ppm, entre 1,5 ppm et 2,7 ppm et vers 4,2 ppm correspondant respectivement aux protons des deux méthyles (-CH₃), et aux protons présents sur la zone bleue du fragment « RA14917148A ». Le spectres ¹H (Figure 82) montre des résonances à environ 2,7 ppm, vers 6,4 ppm et vers 7,3 ppm correspondant respectivement aux protons du -CH₃, aux deux protons de la fonction alcène (en rouge) et aux protons du cycle pyrrole (en vert). Une zone bleue aux alentours de 4,8 ppm correspondant au signal de l'eau, masque les protons des zones colorées en bleu.

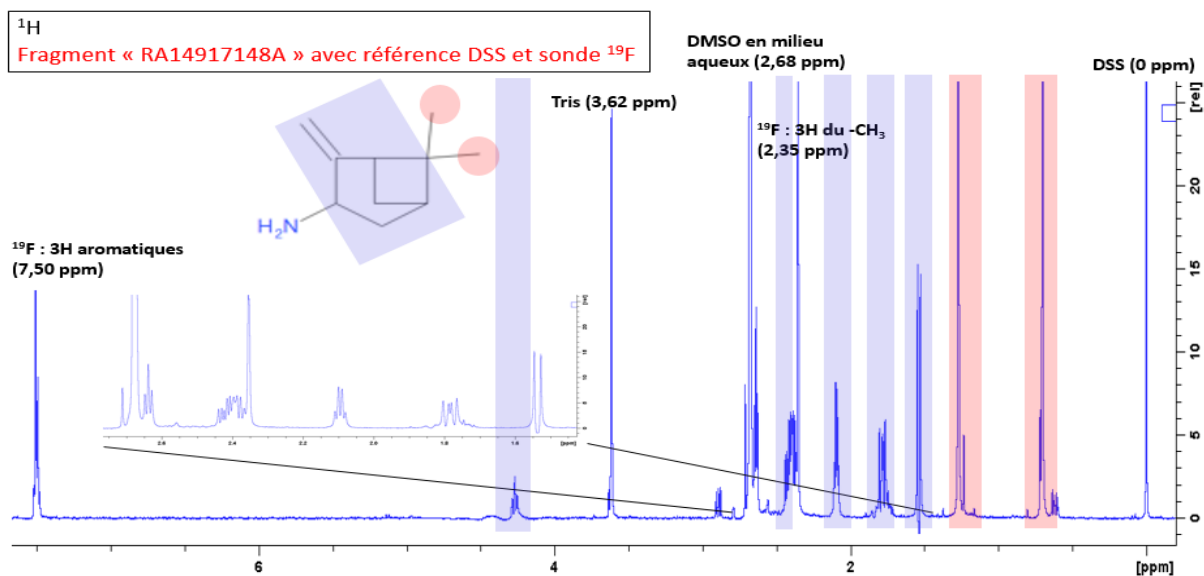


Figure 81 : Spectre ¹H pour le contrôle qualité du fragment « RA14917148A », en présence de la sonde ¹⁹F et de la référence DSS

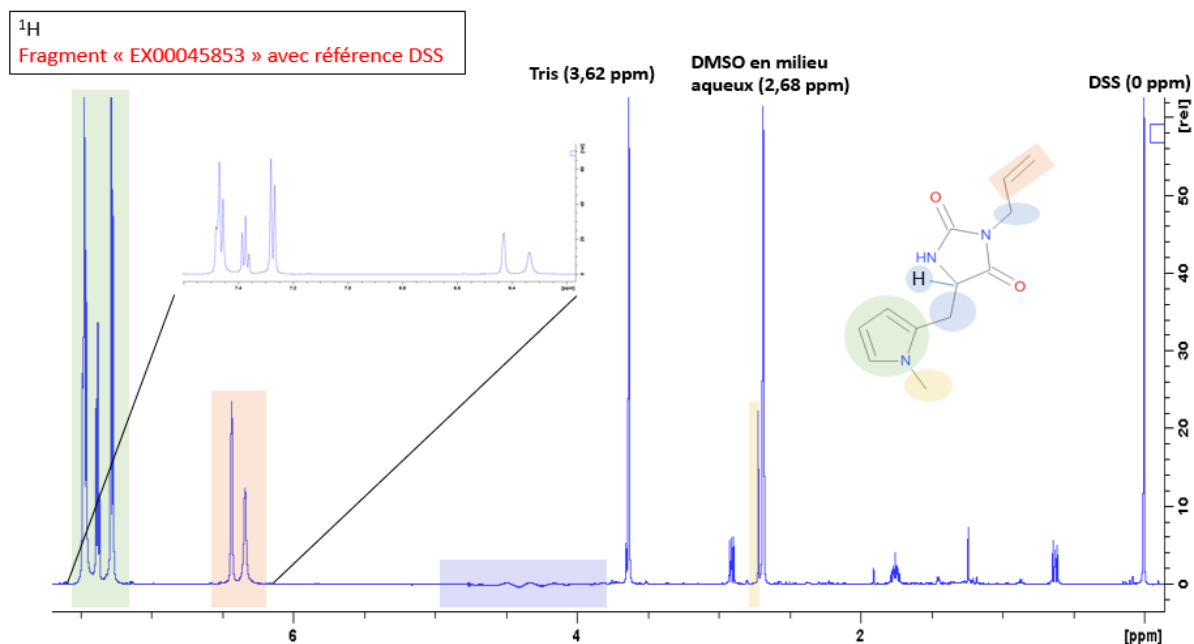


Figure 82 : Spectre ¹H pour le contrôle qualité du fragment « EX00045853 » en présence de la référence DSS

La pureté et la solubilité des composés sont observées sur les spectres ^1H et l'agrégation des différents composés est observée sur les expériences de WL (Figure 83).

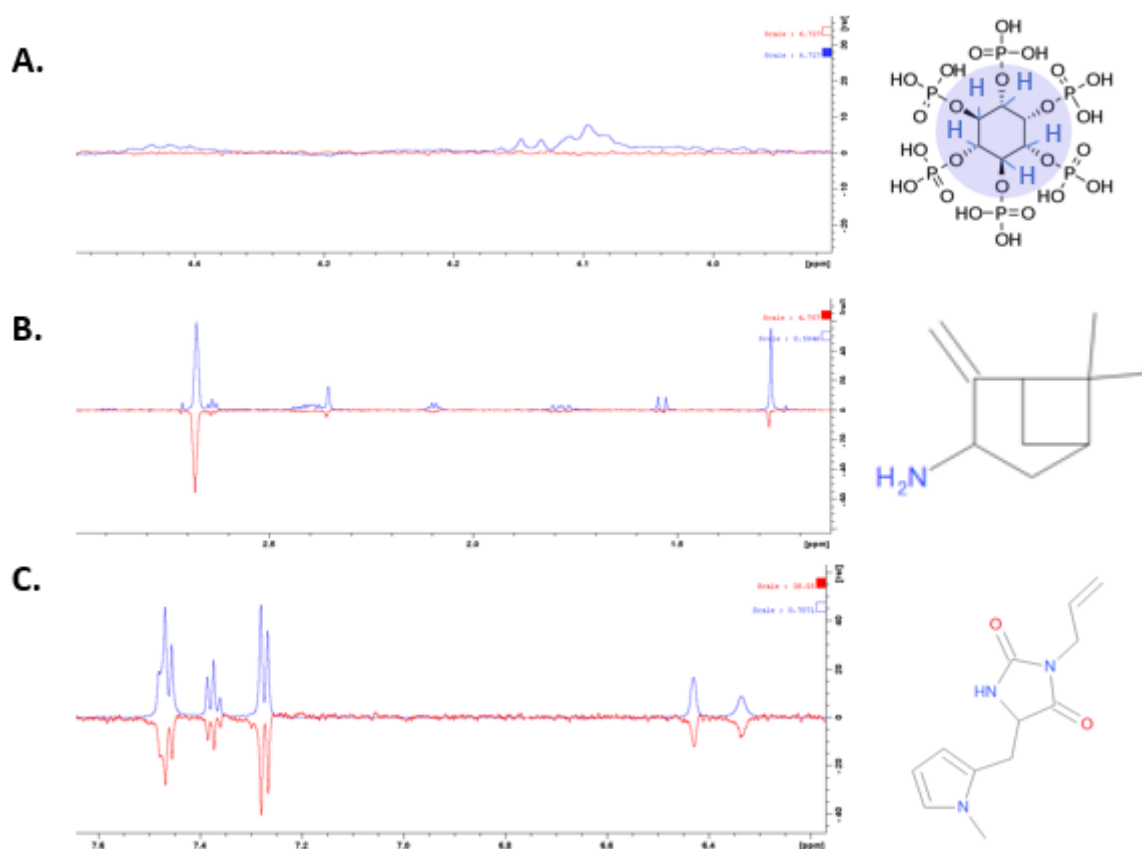


Figure 83 : Superposition des spectres ^1H et WL de l'IP6 (A), du fragment « RA14917148A » (B) et du fragment « EX00045853 » (C).

Voici un tableau récapitulatif des résultats de QC pour les trois composés :

	IP6	RA14917148A	EX00045853
Pureté > 95 %	Ok	OK	OK
Solubilité	Ok	OK	OK
Agrégation	Concentration trop faible en IP6 (1 mM) pour se prononcer	OK, pics négatifs en WL	OK, pics négatifs en WL

Cette analyse de QC permet de dire que les deux fragments peuvent être étudiés en présence de protéine KuCC pour réaliser les expériences d'interaction. L'IP6 ne permet pas cette étude car il n'est pas assez concentré et nous ne savons pas si la molécule est pure à plus de 95 %.

Interactions de type fragment-protéine

La stabilité de KuCC est avant tout vérifiée en réalisant un spectre UV à J=0 et J+1 de la protéine concentrée et en superposant les spectres STD de la protéine diluée à J=0 et J+1. Nous observons une légère baisse d'intensité des pics ^1H correspondant à la protéine KuCC. Ceci signifie que KuCC est légèrement diluée à J+1. Il faut donc préparer les tubes RMN de manière extemporanée.

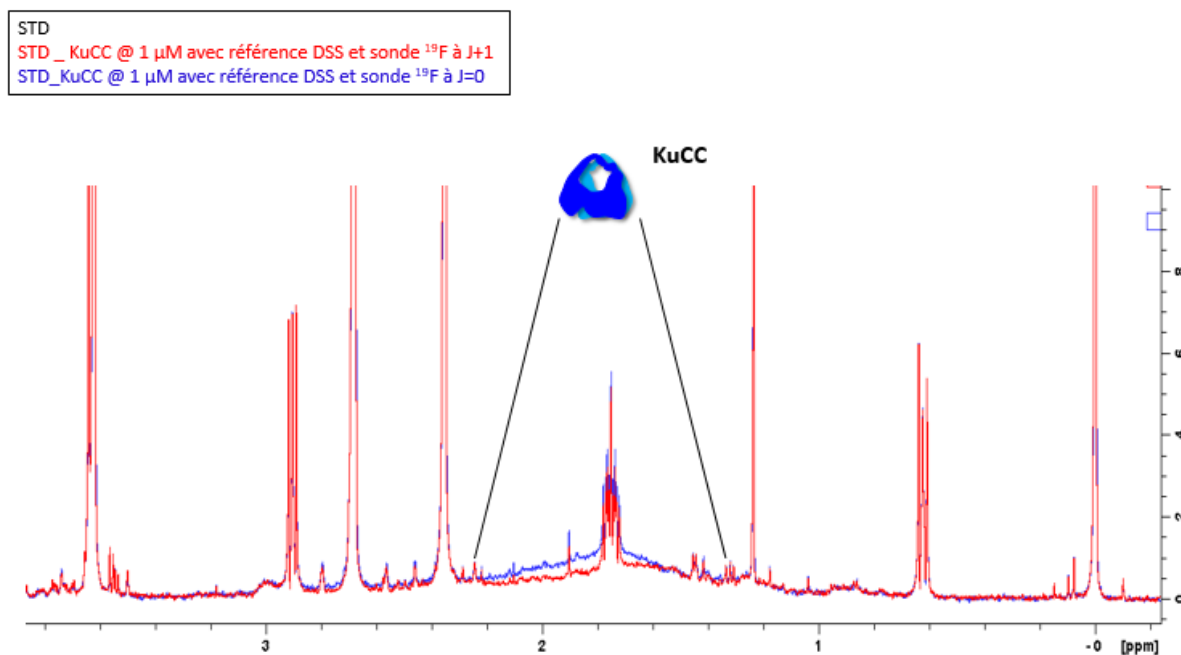


Figure 84 : Superposition des spectres STD de la protéine KuCC à 1 μM à J=0 (bleu) et J+1 (rouge)

Après avoir vérifié la stabilité de KuCC, c'est la concentration optimale en protéine qui a été déduite de la superposition des spectres STD de KuCC à 1, 2, 4 et 6 μM . Nous avons décidé de travailler à une concentration en KuCC de 6 μM car les pics étaient plus intenses. Pour réaliser les expériences d'interactions entre la protéine KuCC et les fragments, il a donc fallu préparer les tubes extemporanément et à une concentration de 6 μM .

Lorsque nous observons le spectre WL de la protéine KuCC, nous constatons des pics positifs au niveau des résonances 2,35 ppm et 7,50 ppm correspondant respectivement aux résonances protons du $-\text{CH}_3$ et aux protons aromatiques de la sonde ^{19}F (Figure 85). La sonde ^{19}F interagit donc avec la protéine KuCC. Pour la suite des expériences, nous avons donc décidé de travailler uniquement avec la référence DSS.

WL
KuCC @ 1 μ M avec référence DSS et sonde ^{19}F

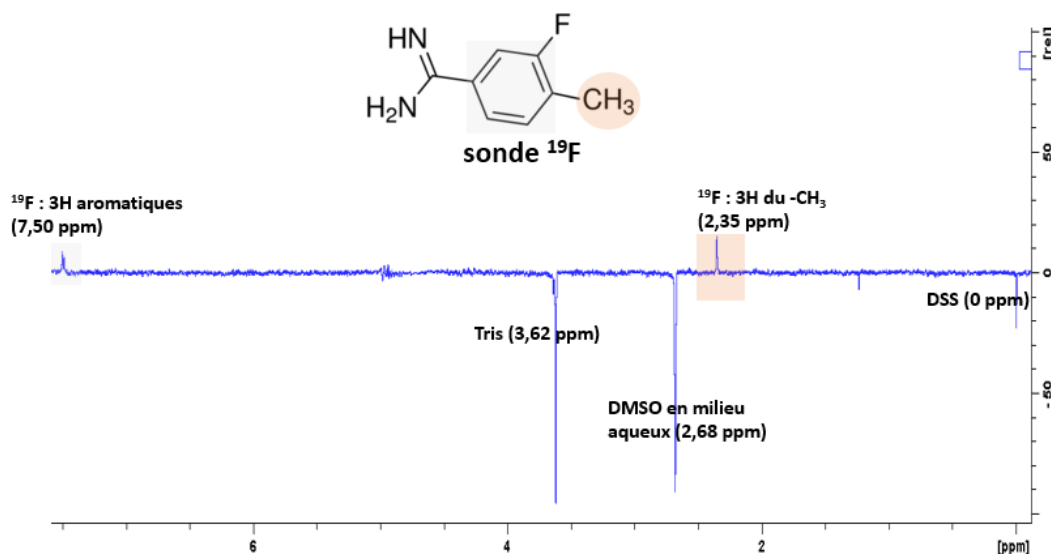


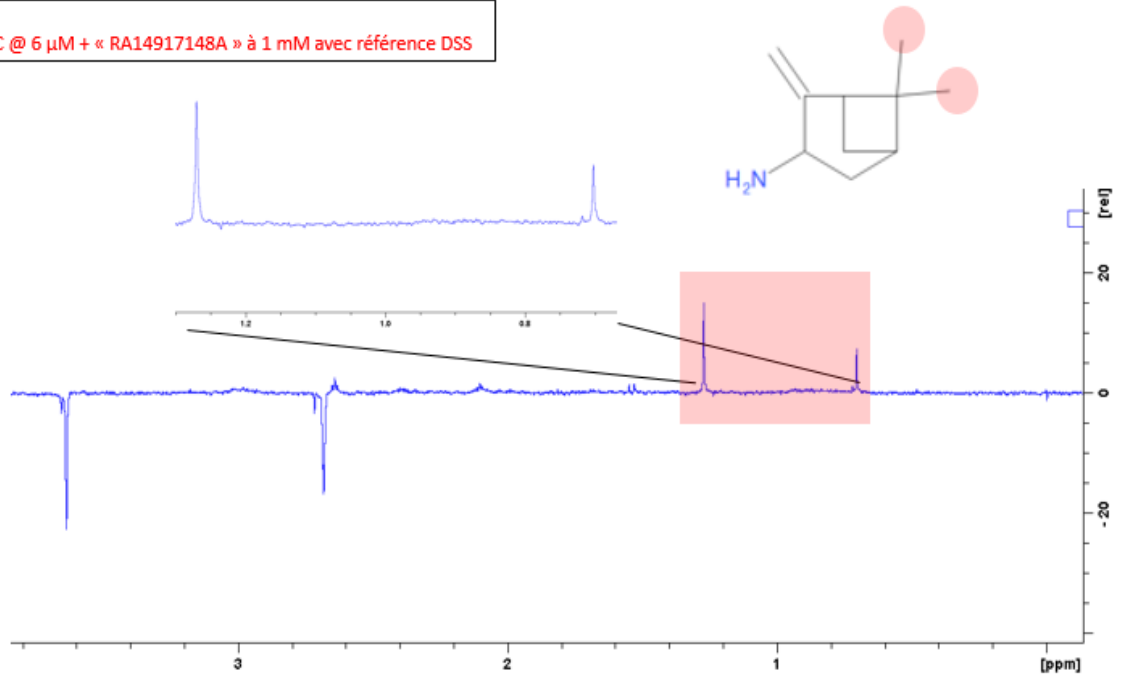
Figure 85 : Spectre WL de la protéine KuCC à 1 μ M dans du tampon 100 % D_2O

Pour visualiser les interactions entre fragments et protéine KuCC, nous avons visualisé les spectres WL de la protéine à 6 μM avec « RA14917148A » à 1 mM et de la protéine à 6 μM avec le fragment « EX00045853 ». Nous constatons que la protéine interagit avec les deux fragments car nous observons des pics positifs en WL aux résonances correspondant aux protons des deux composés. Premièrement, Ku interagit avec les protons des deux groupements $-\text{CH}_3$ (zone encadrée en rouge) du fragment « RA14917148A ». Ku interagit également avec les protons des zones vertes et rouges du fragment « EX00045853 ».

Analyses de compétitions

Suite à ces expériences d'interaction, nous nous sommes intéressés à des expériences de compétition entre la molécule d'IP6 et ces deux composés désormais ligands de Ku. Les mesures de microcalorimétrie (voir partie « Résultats ») ont montré une affinité forte entre la molécule d'IP6 et la protéine Ku ($K_D = 50$ nM). Nous nous sommes donc demandé si cette petite molécule avait le pouvoir d'entrer en compétition avec les deux fragments. Les spectres de chaque composé en absence de protéine (bleu), de chaque composé en présence de protéine (rouge) et de chaque composé en présence de protéine et en présence de 300 μM d'IP6 (vert) ont été superposés à la fois pour les expériences de STD et pour les expériences de WL (Figure 87).

WL
KuCC @ 6 μ M + « RA14917148A » à 1 mM avec référence DSS



WL
KuCC @ 6 μ M + « EX00045853 » à 1 mM avec référence DSS

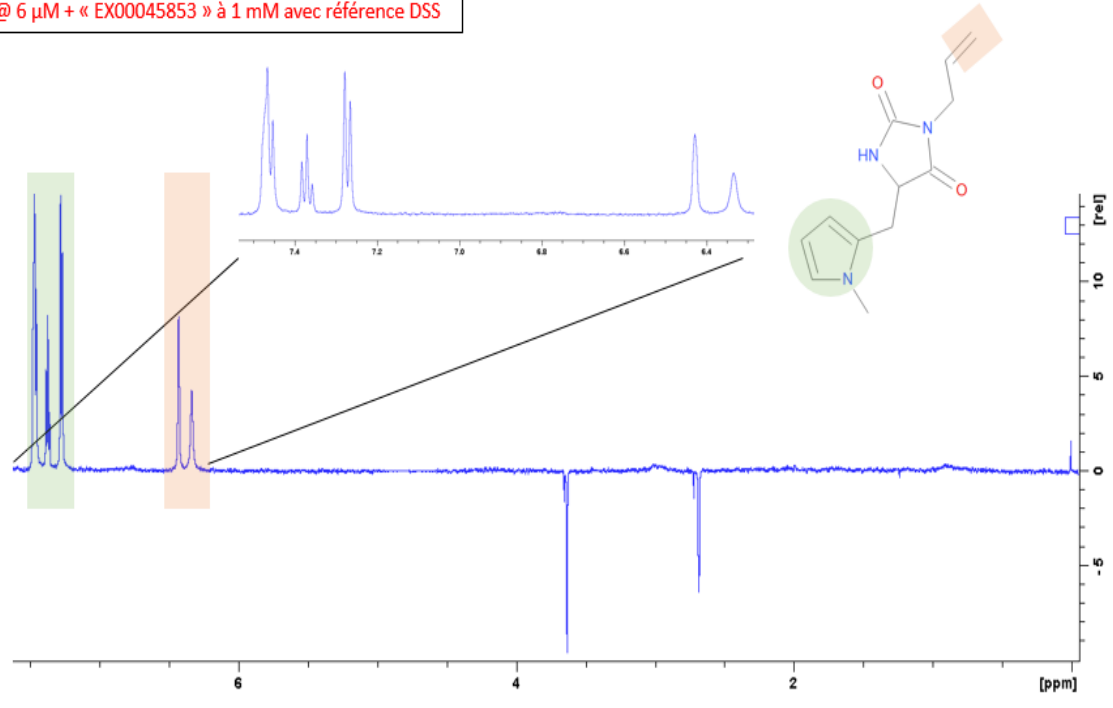


Figure 86 : Expériences de WL pour observer les interactions entre KuCC et le fragment « RA14917148A » (en haut) et entre KuCC et le fragment « EX00045853 » (en bas)

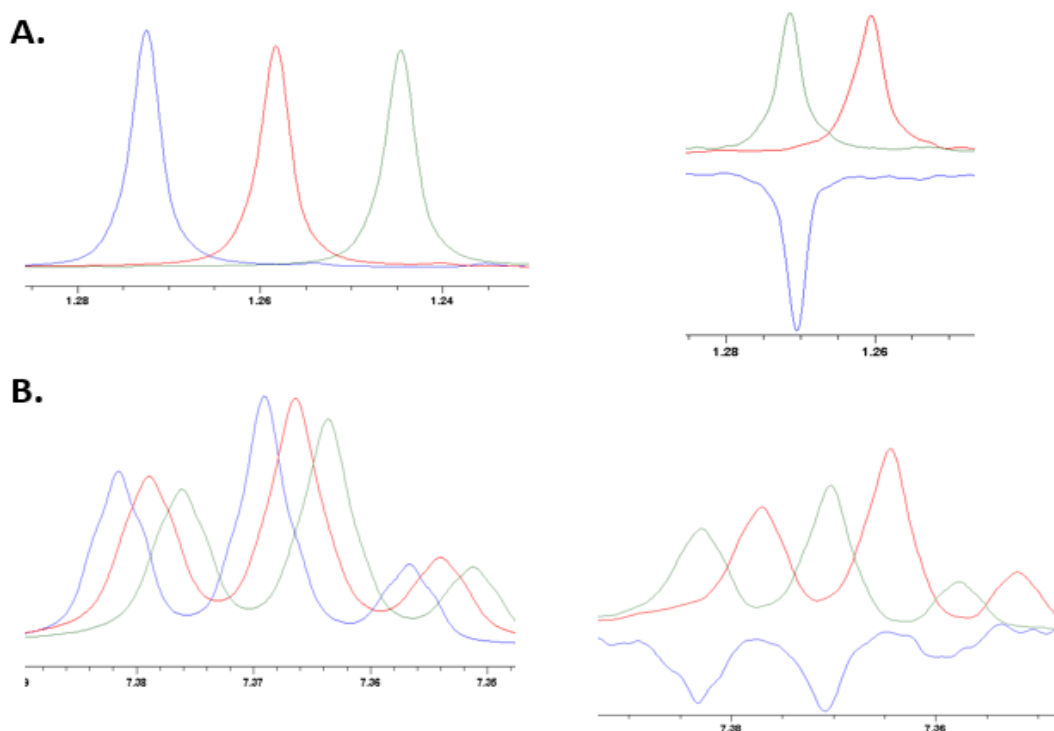
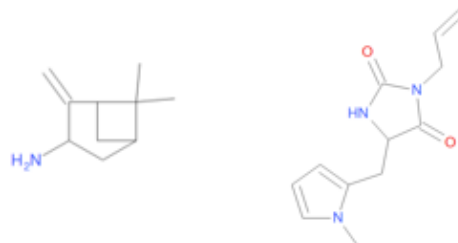


Figure 87 : Expériences de compétitions entre l’IP6 et les fragments. **A.** Spectres concernant la compétition IP6 et fragment « RA14917148A ». A gauche : les spectres STD du fragment seul (bleu), du fragment avec protéine Ku (rouge) et du fragment avec la protéine Ku et l’IP6 (vert) et à droite le même code couleur pour les spectres WL. **B.** Spectres concernant la compétition IP6 et fragment « EX00045853 ». A gauche : les spectres STD du fragment seul (bleu), du fragment avec protéine Ku (rouge) et du fragment avec la protéine Ku et l’IP6 (vert) et à droite le même code couleur pour les spectres WL

La protéine KuCC interagit avec le fragment « RA14917148A » et nous observons une légère compétition avec l’IP6 car les pics ne diminuent que très peu en intensité lorsque l’IP6 est ajouté à 300 μ M. La protéine KuCC interagit avec le fragment « EX00045853 » et nous observons cette fois une réelle compétition avec l’IP6. En effet, les pics diminuent en intensité lorsque l’IP6 est ajouté à 300 μ M. Cet effet est davantage observé pour le fragment « EX00045853 ». De plus, les résultats de DSF ont montré que le fragment « EX00045853 » était plus stabilisateur de la protéine Ku en comparaison avec le fragment « RA14917148A ». Le Tm1 est déplacé de + 3 °C à une concentration minimale de 2 mM en « EX00045853 » et il est déplacé de + 1,5 °C à une concentration minimale de 1,5 mM en « RA14917148A ». Les deux interprétations sont concordantes. Enfin, la petite molécule d’IP6 a une affinité forte pour la protéine Ku et est capable de déplacer le fragment « EX00045853 ». L’IP6 et le fragment « EX00045853 » sont donc situés sur le même site de la protéine Ku.

Tableau récapitulatif des expériences d’interactions et de compétitions entre la protéine Ku et les deux fragments :

		RA14917148A	EX00045853
Interaction avec KuCC	WL STD	Oui Oui	Oui Oui
Compétition avec IP6	WL STD	Non Faible	Oui Oui



PARTIE III :

Caractérisation d'inhibiteurs d'une interaction protéine-ADN dans un projet de Recherche Sanofi de candidats médicaments

Le second volet principal de ma thèse a porté sur l'utilisation de la nouvelle technologie de switchSENSE implantée au laboratoire (I2BC, CEA Saclay). Le projet sur lequel j'ai travaillé s'inscrit dans un projet Universel d'oncologie. Le premier objectif était de savoir si cette nouvelle technologie pouvait être utilisée pour ce projet dont je ne pourrai détailler les molécules en jeu pour des raisons de confidentialité. Il s'agissait d'évaluer si la technologie permettait de mesurer des interactions de type protéine-ADN, étudiées par ailleurs au sein du service par d'autres techniques biophysiques. Le troisième objectif était de caractériser de nouveaux inhibiteurs et de pouvoir mesurer leur efficacité *in vitro* en parallèle de mesures faites chez Sanofi. La finalité de ce projet était donc de comparer les atouts et inconvénients du switchSENSE par rapport aux méthodes plus anciennes comme la SPR, la DSF ou la MST.

A) Caractérisation des interactions protéine-ADN

Au cours de la première année de thèse, le service « *Integrated Drug Discovery* » a reçu l'appareil DRX² durant une semaine lors d'une démonstration ayant eu lieu sur le site de Vitry-Sur-Seine. Cette semaine a permis à l'équipe de voir ce qui était possible de réaliser comme types d'expériences et de réaliser quelques tests sur divers projets. De mon côté, j'avais en ma possession quatre constructions protéiques et deux ADN de séquences différentes. Pour une question de confidentialité, les protéines sont notées A, B, C et D, les ADN sont notés ADN1 et ADN2 et les inhibiteurs 275A, 9689, 4931, 1278 et 5764.

Le switchSENSE permet la caractérisation des interactions moléculaires notamment de type protéine-ADN. Deux sortes de puces ont été testées pour ce projet : la puce de type « MPC2-48-2-G1R1-S » et la puce de type « MPC-2-48-R1-S ». Il s'avère que c'est la puce de type « MPC-48-2-R1-S » standard présentant deux électrodes de référence (1 et 2) avec une séquence A en ADN, quatre électrodes de mesure (3 à 6) avec une séquence B en ADN qui a été utilisée dans cette étude. En effet, la puce de type « MPC2-48-2-G1R1-S » a montré de

manière aléatoire des interactions non-spécifiques entre la séquence A et les protéines. Sur les 6 électrodes, sont greffés des milliers d'ADN simple brin de 48 nucléotides présentant une sonde fluorescente rouge en 3'. On hybride sur les électrodes de mesure un ADN mixte (Figure 88) comprenant la séquence complémentaire plus la séquence d'intérêt. On montre avant toute expérience de cinétique que la protéine (A, B, C ou D) présente bien une interaction spécifique avec la séquence en ADN d'intérêt (Figure 88). Les résultats pour l'ADN1 ou l'ADN2 sont semblables, on ne montre donc qu'un profil obtenu (voir la partie « Matériel et Méthodes » pour les détails). Les mesures d'interactions ont été réalisées tout d'abord en mode statique puis en mode dynamique pour certaines expériences pour vérifier que ces deux modes permettent d'obtenir les mêmes valeurs cinétiques.

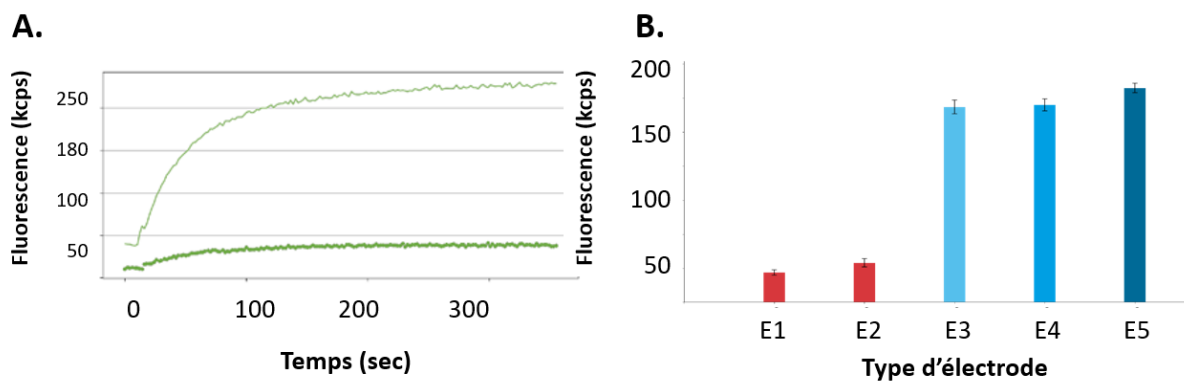


Figure 88 : Résultats préliminaires de switchSENSE avant toute cinétique. **A.** Hybridation de l'ADN complémentaire présentant la séquence en ADN d'intérêt ; augmentation de la fluorescence de l'ADN double brin jusqu'à obtention d'un plateau de saturation de la fluorescence. **B.** Profil d'émission de fluorescence obtenue après injection d'une protéine sur les différentes électrodes de la puce MPC-48-2-R1-S (E1 et E2 : électrodes de référence et E3 à E6 : électrodes de mesure). La protéine (profil valable pour les quatre protéines et les deux ADN) se lie de manière spécifique à la séquence en ADN d'intérêt présente sur les électrodes de mesure

Une première étape de cinétique à une seule concentration en protéine (1 μ M) permet d'observer une courbe d'association et de dissociation de la protéine sur l'ADN. Cette courbe permet de choisir une gamme de concentrations en protéine pour réaliser plusieurs cinétiques. Le test n° 1 (Figure 89A) a permis de dire que le temps d'association pour la cinétique en mode statique était correct puisque l'on aperçoit un plateau de saturation. Cependant, le temps de dissociation est trop court ; la protéine présentée ici se dissocie de l'ADN sur une durée supérieure à la limite choisie (150 s). On améliore l'analyse en utilisant une dissociation sur 1 200 s qui permet un retour de la fluorescence au niveau de départ (fluorescence normalisée de 0,95).

Les constantes de dissociation (K_D) sont du même ordre de grandeur (en A : $K_D = 539 \pm 95$ nM et en B : $K_D = 246 \pm 12$ nM) mais la valeur obtenue en B a une meilleure précision car la dissociation est estimée plus précisément. Les valeurs de k_{ON} sont du même ordre de grandeur avec une erreur semblable ($1,47 \pm 0,06 \cdot 10^4 \text{ M}^{-1} \cdot \text{s}^{-1}$ pour A et $1,85 \pm 0,06 \cdot 10^4 \text{ M}^{-1} \cdot \text{s}^{-1}$ pour B). En revanche, les valeurs de k_{OFF} présentent une plus grande différence ($7,92 \pm 1,36 \cdot 10^{-3} \text{ s}^{-1}$ en A et $4,54 \pm 0,16 \cdot 10^{-3} \text{ s}^{-1}$ en B). L'erreur sur la mesure du k_{OFF} est plus importante en A qu'en B (respectivement de 17 % et de 3 %). Le test n°2 (Figure 89B) montre un profil satisfaisant pour démarrer une cinétique sur une gamme de concentrations en protéine.

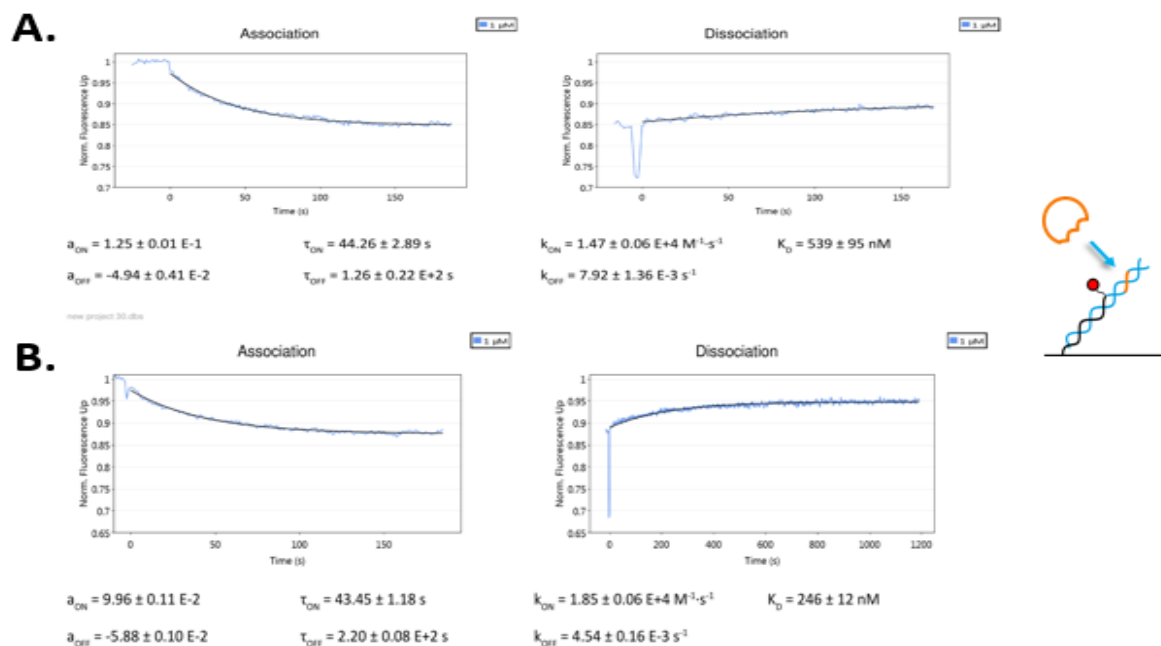


Figure 89 : Exemple d'analyses d'une optimisation de mesure cinétique d'interaction protéine – ADN après fonctionnalisation de la puce MPC-48-2-R1-S avec un ADN complémentaire débordant en mode statique. **A.** Test n°1 à 1 µM en protéine contre un ADN, dissociation mesurée sur 150 s, débit 200 µl/min **B.** Test n°2 à 1 µM en protéine contre un ADN et une dissociation mesurée sur 1 200 s, débit 100 µl/min

Un effet de transport de masse (*mass transport* en anglais) a également été observé lors des premières expériences (données non montrées). Lors des premières expériences d'association, une droite linéaire décrit la liaison de la protéine à l'ADN au lieu d'une courbe exponentielle. Ce phénomène est observé lorsque l'association de la protéine est très rapide. Dans ce cas, le débit en protéine est augmenté afin de limiter cet effet. Lors de l'événement de liaison à la surface du capteur, la concentration en analyte au-dessus de la surface est réduite. Ainsi, une zone d'appauvrissement est observée au-dessus de la surface. La taille de cette zone et la vitesse à laquelle elle se forme dépendent principalement du taux d'association de l'interaction en question, de la densité de surface du ligand et de la surface du capteur. La limitation du transport de masse se produit par exemple pour des interactions avec des taux d'association très rapides

mesurés sur des surfaces à haute densité et des débits trop lents en protéines injectées. Cet événement conduit à un épuisement très marqué de l'analyte, avec pour résultat que l'association se produit à un taux apparent qui est plus lent que le taux d'association réel (Squires, Messinger, et Manalis 2008).

Le tampon protéique et le tampon de dissociation (*running buffer*) doivent être identiques. La protéine est concentrée puis diluée plusieurs fois sur concentrateur Amicon Ultra.

Interactions protéines-ADN1

Les résultats sont détaillés pour la protéine B et l'ADN1. Les résultats concernant les interactions des protéines A, C et D avec l'ADN1 sont résumés.

Un ADN complémentaire est injecté sur la puce. Il présente la séquence complémentaire de 48 nucléotides pour s'hybrider sur les électrodes 3 à 6 et une séquence de 18 pb correspondant à la séquence d'ADN1 cible à tester en cinétique pour les quatre constructions protéiques. La pré-hybridation de l'ADN1 est analysée sur le logiciel switchANALYSIS par vérification de l'augmentation de la fluorescence de l'ADN double brin. Avant de démarrer l'étude par switchSENSE, nous avons une idée de l'affinité entre la protéine B et l'ADN1. Le K_D avait été mesuré aux alentours du μM chez Sanofi. Ceci nous a permis de nous mettre dans une gamme de concentrations adéquate sinon nous aurions démarré à des concentrations plus faibles et serions montés progressivement jusqu'à visualiser un effet. Dans cette étude, les deux ADN (1 et 2) présentent une incision (*nick*) une fois hybridés générant une certaine flexibilité. Il est préférable dans ce cas de réaliser des mesures en mode statique dans un premier temps. Le mode statique permet de mesurer les variations de fluorescence au niveau de la sonde en fonction de l'environnement autour du fluorophore et notamment de l'interaction d'une protéine à proximité.

Parmi les tests effectués, une gamme de concentrations en protéine B (de 3 μM à 23 nM avec un facteur de dilution de 2) a été effectuée dans du tampon DBS TE140. Un débit de 200 $\mu\text{L}/\text{min}$ a été appliqué pour un temps d'association de 40 sec et un temps de dissociation de 50 sec. La protéine B a un k_{ON} de $2,6 \pm 0,4 \cdot 10^5 \text{ M}^{-1} \cdot \text{s}^{-1}$, un k_{OFF} de $1,3 \pm 0,1 \cdot 10^{-1} \text{ s}^{-1}$ et un K_D de $510 \pm 30 \text{ nM}$ pour l'ADN1 en mode statique (Figure 90A). Les ajustements obtenus (*fits*) sont corrects et les erreurs également pour les trois courbes correspondant à des concentrations en protéine de 3 μM à 750 nM. Cette protéine s'associe et se dissocie très rapidement de l'ADN1. J'ai voulu savoir si le mode dynamique (cette fois par mouvement des leviers de l'ordre de 10 μs) était

applicable avec des ADN de type débordants et plus longs que 48 pb (longueur des ADN simples brins présents sur la puce).

J'ai donc réalisé un test en mode dynamique pour une gamme de concentrations en protéine B allant de 3 μM à 375 nM avec un facteur de dilution de 2. Après plusieurs optimisations des paramètres de cinétique, le débit est conservé à 200 $\mu\text{L}/\text{min}$ et les temps d'association et de dissociation sont choisis respectivement à 30 sec et 90 sec. On mesure pour la protéine B un k_{ON} de $3,7 \pm 0,7 \cdot 10^5 \text{ M}^{-1} \cdot \text{s}^{-1}$, un k_{OFF} de $1,6 \pm 0,2 \cdot 10^{-1} \text{ s}^{-1}$ et un K_{D} de $440 \pm 90 \text{ nM}$ pour l'ADN1 en mode dynamique (Figure 90B). La concentration de 375 nM en protéine B est ajustable avec le logiciel switchANALYSIS contrairement au mode statique. Les valeurs obtenues pour le mode statique et le mode dynamique sont proches.

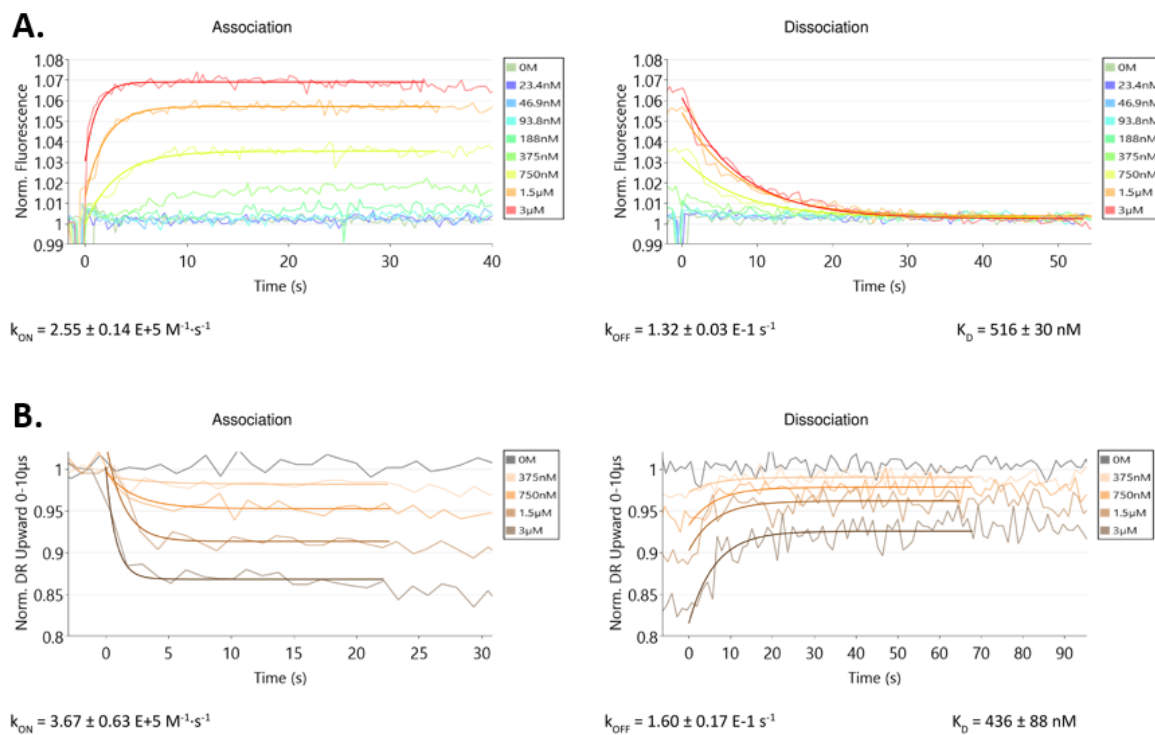


Figure 90 : Analyses d'interactions par switchSENSE de la protéine B sur l'ADN 1. **A.** Caractérisation de l'interaction ProtB-ADN1 en mode statique ; fluorescence normalisée. **B.** Caractérisation de l'interaction ProtB-ADN1 en mode dynamique ; fluorescence normalisée

Interactions protéines-ADN2

Les résultats sont détaillés pour la protéine A et l'ADN2. Les résultats concernant les interactions des protéines B, C et D avec l'ADN2 sont résumés. L'ADN2 a nécessité comme l'ADN1 une pré-hybridation au préalable avant la fonctionnalisation de la puce MPC-48-2-R1-S.

Un ADN complémentaire est injecté sur la puce. Il présente la séquence complémentaire de 48 nucléotides plus une séquence de 22 pb cible à la suite, à tester en cinétique avec les quatre constructions protéiques. La pré-hybridation de l'ADN2 est analysée sur le logiciel switchANALYSIS par vérification de l'augmentation de la fluorescence de l'ADN double brin. Avant de démarrer l'étude par switchSENSE, nous n'avions pas d'idée de l'affinité entre la protéine A et l'ADN2. L'expérience de cinétique est réalisée en plusieurs étapes. La protéine de 8,5 kDa a tout d'abord été concentrée à 500 nM dans un tampon DBS TE140 pour être associée à l'ADN2 pendant 1 min et dissociée pendant 4 min à un débit de 100 $\mu\text{L}/\text{min}$ soit des conditions proches de celles utilisées pour l'ADN1. Ces premières courbes suggèrent que la concentration en protéine est trop faible. La concentration en protéine A a été augmentée à 2 μM , le débit a été augmenté à 200 $\mu\text{L}/\text{min}$ ainsi que les temps d'association et de dissociation à respectivement 4 min et 10 min. Le k_{ON} obtenu est de $3,6 \pm 0,2 \cdot 10^3 \text{ M}^{-1} \cdot \text{s}^{-1}$, le k_{OFF} de $5,3 \pm 0,2 \cdot 10^{-3} \text{ s}^{-1}$ pour un K_{D} de $1,5 \pm 0,1 \mu\text{M}$. Les erreurs calculées sont satisfaisantes. Nous avons cependant choisi d'augmenter un peu la concentration en protéine (6 μM au lieu de 2 μM) et d'augmenter un peu le temps de dissociation pour obtenir un retour du signal de fluorescence à sa valeur initiale (Figure 91A). Avec ces nouveaux paramètres, un k_{ON} de $1 \pm 0,1 \cdot 10^3 \text{ M}^{-1} \cdot \text{s}^{-1}$, un k_{OFF} de $1,7 \pm 0,1 \cdot 10^{-3} \text{ s}^{-1}$ et un K_{D} de $1,6 \pm 0,1 \mu\text{M}$ sont obtenus. Les vitesses auxquelles la protéine s'associe et se dissocie de l'ADN2 sont du même ordre de grandeur bien que les erreurs calculées soient plus faibles (Figure 91B).

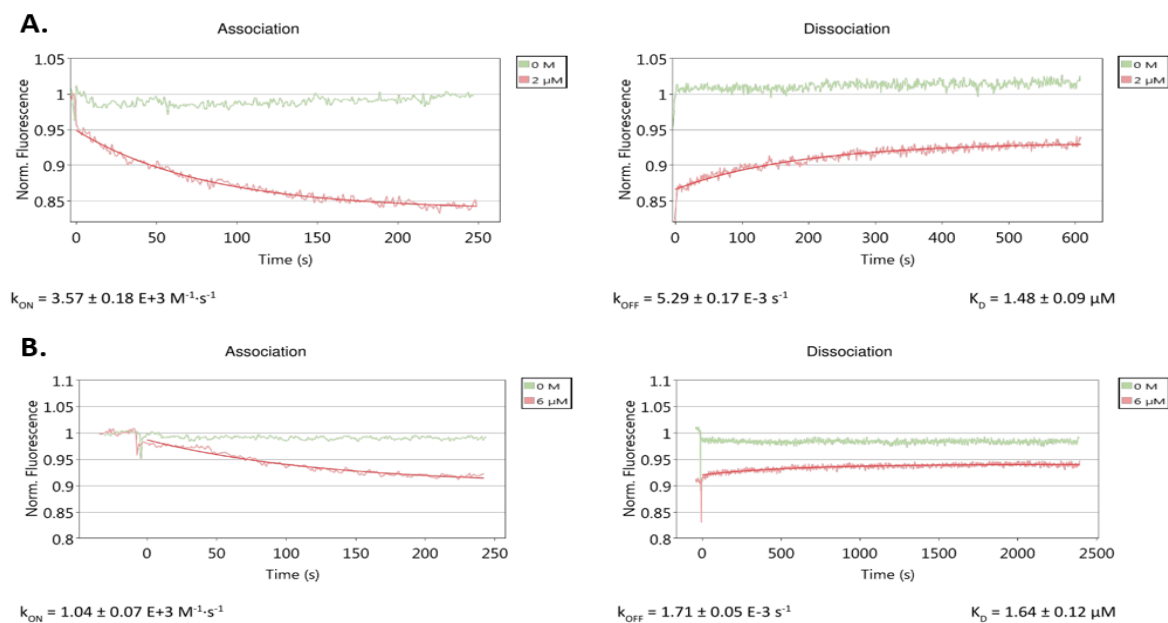


Figure 91 : Analyses des résultats préliminaires d'interactions par switchSENSE de la protéine A sur l'ADN 2. **A.** Caractérisation de l'interaction ProtA-ADN2 en mode statique lors d'un troisième test ; fluorescence normalisée. **B.** Caractérisation de l'interaction ProtA-ADN2 en mode statique lors d'un quatrième test ; fluorescence normalisée

Après avoir trouvé les conditions optimales pour les mesures de cinétique entre la protéine A et l'ADN 2 en mode statique, une gamme de concentrations en protéine A a été réalisée. La fonctionnalisation de la puce MPC-48-2-R1-S a été effectuée. Nous notons qu'à chaque nouvelle hybridation de l'ADN2 sur la puce (pour chaque nouvelle concentration en protéine A), le signal de fluorescence diminue. De ce fait, l'amplitude entre la fluorescence *up* et *down* diminue, ce qui signifie que la puce s'abîme au cours du temps. Après avoir réalisé une gamme de concentrations en protéine A de 12 μM à 375 nM avec un facteur de dilution de 2, seules les concentrations 12 μM , 6 μM , 3 μM et 1,5 μM sont ajustables par switchANALYSIS. Les courbes de fluorescence normalisées correspondant aux concentrations 750 nM et 375 nM sont supprimées. Un k_{ON} de $2,5 \pm 0,02 \cdot 10^3 \text{ M}^{-1} \cdot \text{s}^{-1}$, un k_{OFF} de $3,4 \pm 0,1 \cdot 10^{-3} \text{ s}^{-1}$ et un K_D de $1,3 \pm 0,03 \mu\text{M}$ sont obtenus pour l'interaction protéineA-ADN2. La valeur d'affinité obtenue est identique avec celle obtenue pour une cinétique à une seule concentration en protéine A (Figure 91B) mais l'erreur calculée est plus faible. D'autre part, les valeurs correspondent à la moyenne d'ajustements faits sur ces 4 expériences indépendantes (4 concentrations en protéine). Un fit global avait été testé car il est habituellement plus robuste mais les résultats étaient plus corrects en effectuant une moyenne des 4 expériences indépendantes.

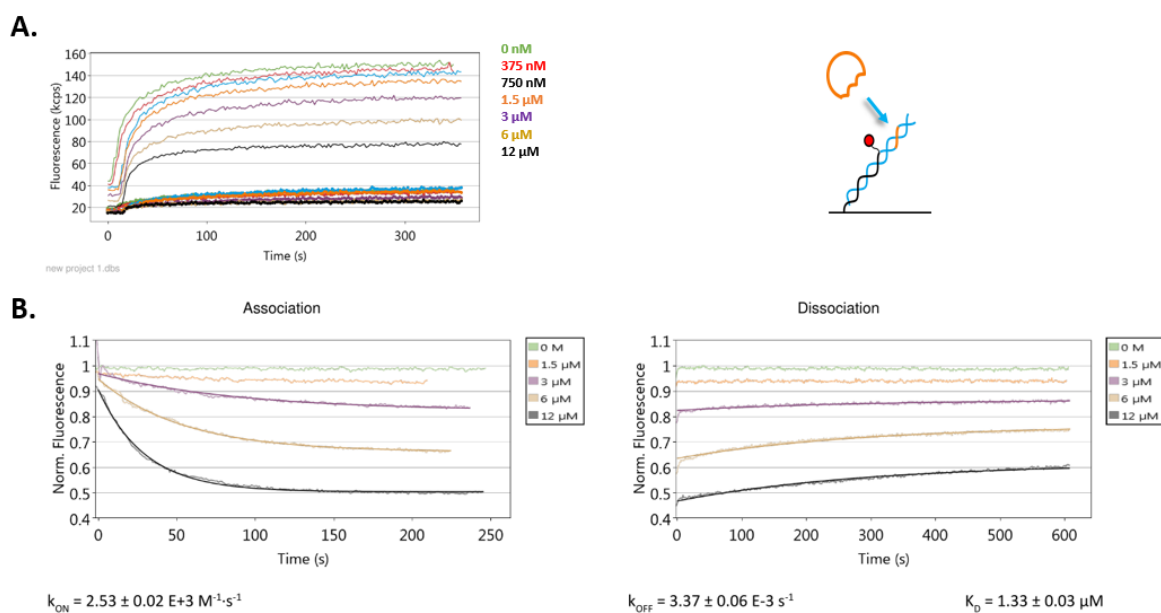


Figure 92 : Analyses de résultats de switchSENSE en mode statique ; cinétique d'interaction pour la protéine A et l'ADN 2. **A.** Fonctionnalisation de la puce MPC-48-2-R1-S avec l'ADN 2 à 500 nM. Le profil obtenu montre que les nanoleviers d'ADN sont fonctionnels après chaque régénération de la puce et avant chaque nouvelle concentration en protéine A. **B.** Caractérisation de l'interaction protéineA-ADN2 en mode statique avec différentes concentrations en protéine A

Bilan des résultats de caractérisation protéine-ADN

Les Figure 93 et 94 présentent le bilan de l'ensemble des mesures effectuées avec les protéines A, B, C et D sur les ADN1 et ADN2. Les K_D mesurés varient entre 19 μM (protéine C vs ADN1) et 85 nM (protéine C vs ADN2). Sept interactions sur huit (K_D , k_{ON} , k_{OFF}) ont pu être mesurées. Les protéines interagissent très peu avec les contrôles. Certaines interactions ont pu être mesurées plus précisément. Des mesures supplémentaires seraient nécessaires pour affiner les mesures effectuées entre la protéine B et l'ADN2. Il faudrait notamment essayer d'augmenter davantage le temps de dissociation grâce à la pompe péristaltique. La tendance globale qui ressort est que la protéine de plus petite taille (protéine A) est la moins affine et que la construction de 15 kDa (protéine C) est la plus affine pour l'ADN.

Protéine	MW (kDa)	k_{ON} ($\text{M}^{-1} \cdot \text{s}^{-1}$)	k_{OFF} (s^{-1})	K_D (μM)
A	8.5	-	-	-
B	11	$3,11 \pm 0,38 \cdot 10^5$	$1,46 \pm 0,1 \cdot 10^{-1}$	$476 \pm 59 \cdot 10^{-3}$
C	15	$0,9 \pm 0,2 \cdot 10^3$	$1,29 \pm 0,07 \cdot 10^{-2}$	19 ± 5
D	71	$2 \pm 0,1 \cdot 10^4$	$3,48 \pm 0,12 \cdot 10^{-3}$	$176 \pm 10 \cdot 10^{-3}$

Figure 93 : Récapitulatif des résultats obtenus par switchSENSE pour des expériences de cinétique concernant les quatre constructions protéiques et l'ADN1

Protéine	MW (kDa)	k_{ON} ($\text{M}^{-1} \cdot \text{s}^{-1}$)	k_{OFF} (s^{-1})	K_D (μM)
A	8.5	$1,78 \pm 0,04 \cdot 10^3$	$2,54 \pm 0,05 \cdot 10^{-3}$	$1,5 \pm 0,07$
B	11	$1,95 \pm 1,1 \cdot 10^3$	$1,2 \pm 1,75 \cdot 10^{-2}$	7 ± 4
C	15	$1,07 \pm 0,05 \cdot 10^4$	$9,05 \pm 0,07 \cdot 10^{-4}$	$85 \pm 4 \cdot 10^{-3}$
D	71	$1,91 \pm 0,14 \cdot 10^4$	$3,15 \pm 0,09 \cdot 10^{-3}$	$192 \pm 28 \cdot 10^{-3}$

Figure 94 : Récapitulatif des résultats obtenus par switchSENSE pour des expériences de cinétique concernant les quatre constructions protéiques et l'ADN2

B) Caractérisation d'inhibiteurs par titrations

Au cours de ma thèse, plusieurs composés identifiés comme candidats intéressants pour empêcher les interactions protéine-ADN étudiées, m'ont été confiés pour les tester en switchSENSE. Le but était de caractériser par switchSENSE l'inhibition de l'interaction entre certaines des protéines A, B, C et D avec les ADN1 et 2. L'objectif était de caractériser des

inhibiteurs liant la protéine et non l'ADN. J'ai travaillé sur quelques inhibiteurs pour lesquels j'avais des données provenant de mesures chez Sanofi par SPR, MST ou DSF.

Le mode statique est choisi pour cette étude. Chaque protéine est concentrée à 1 μM , chaque ADN est hybridé à 500 nM sur l'ADN 48 nt simple brin de la puce. Les composés testés étaient tous dans 100 % de DMSO au départ. Il a fallu tout d'abord trouver une stratégie pour que le tampon protéique, le tampon de dissociation (*running buffer*) et le composé en question soient tous à 4 % de DMSO.

Nous prenons dans cette partie un exemple décrit dans les détails et nous résumons les constantes d'inhibition obtenus dans un tableau récapitulatif.

Afin de démarrer l'étude de titration de manière efficace, j'ai commencé par vérifier la spécificité de liaison de chaque inhibiteur. L'hybridation de l'ADN d'intérêt sur la puce MPC-48-2-R1-S a été réalisée avant chaque début de titration. Dans l'exemple cité ci-dessous, 4 inhibiteurs sont testés sur l'interaction protéineB-ADN1. Ces inhibiteurs sont appelés 925A, 352A, 275A et 610. Pour connaître la concentration à laquelle l'analyte est injecté, j'ai établi une gamme de concentrations en adéquation avec la valeur de K_D obtenue pour ces inhibiteurs chez Sanofi grâce aux autres méthodes de biophysique. Après différents tests préliminaires, il s'est avéré que la meilleure stratégie était d'établir une gamme en inhibiteur ayant pour concentration minimale au moins 3 fois la valeur de K_D attendue et pour concentration maximale 10 fois la valeur de la concentration minimale de la gamme. Le facteur de dilution optimal est préconisé par le logiciel switchANALYSIS. Après cinétique des 4 inhibiteurs à la plus haute concentration de la gamme (ici 150 μM) sur l'ADN d'intérêt (ici l'exemple concerné est l'ADN1) à 100 $\mu\text{L}/\text{min}$ pendant un temps d'association de 50 sec et un temps de dissociation de 100 sec, nous observons que les composés 925A et 352A se lient à l'ADN : la fluorescence normalisée diminue à 10 sec d'association et on obtient une extinction de la fluorescence (*quenching*). Les composés 275A et 610 montrent un saut de fluorescence mais ne semblent pas interagir avec l'ADN double brin (Figure 95A). Les composés 925A et 352A sont éliminés de l'étude. Un premier test de cinétique en mode statique est réalisé dans du tampon TE140 en co-incubant la protéine B à 1 μM avec différentes concentrations en composés 275A et 610. Nous montrons ici les résultats obtenus avec le composé 275A.

Nous effectuons dans un premier temps, des cinétiques en co-incubant la protéine B à 1 μM avec des concentrations croissantes en composé 275A (de 161 nM à 150 μM). La présence du composé 275A mène à une diminution du plateau de signal de fluorescence lors de l'association

de la protéine B pour l'ADN 1. Plus la concentration en inhibiteur augmente, plus l'association de la protéine B sur l'ADN 1 diminue (Figure 95B).

Nous effectuons alors des expériences de titrations où la fluorescence n'est pas mesurée en cinétique mais à l'équilibre lorsque la fluorescence atteint un plateau. La protéine B à 1 μM est co-incubée avec une gamme de concentration en inhibiteur (allant de 600 μM à 934 nM) avec un facteur de dilution de 1,8. La gamme de titration est effectuée en partant de la plus haute concentration en inhibiteur jusqu'à la plus basse. La valeur de la constante d'inhibition de l'interaction K_i entre l'inhibiteur 275A et la protéine B est de $44 \pm 1 \mu\text{M}$ (Figure 96). La même expérience permet de mesurer un K_i de l'inhibiteur 610 pour l'interaction protéine B-ADN de $30 \pm 5 \mu\text{M}$. Ces données ont cependant montré que les signaux de fluorescence étaient les mêmes entre les électrodes de référence et les électrodes de mesure. Ce composé est donc éliminé également.

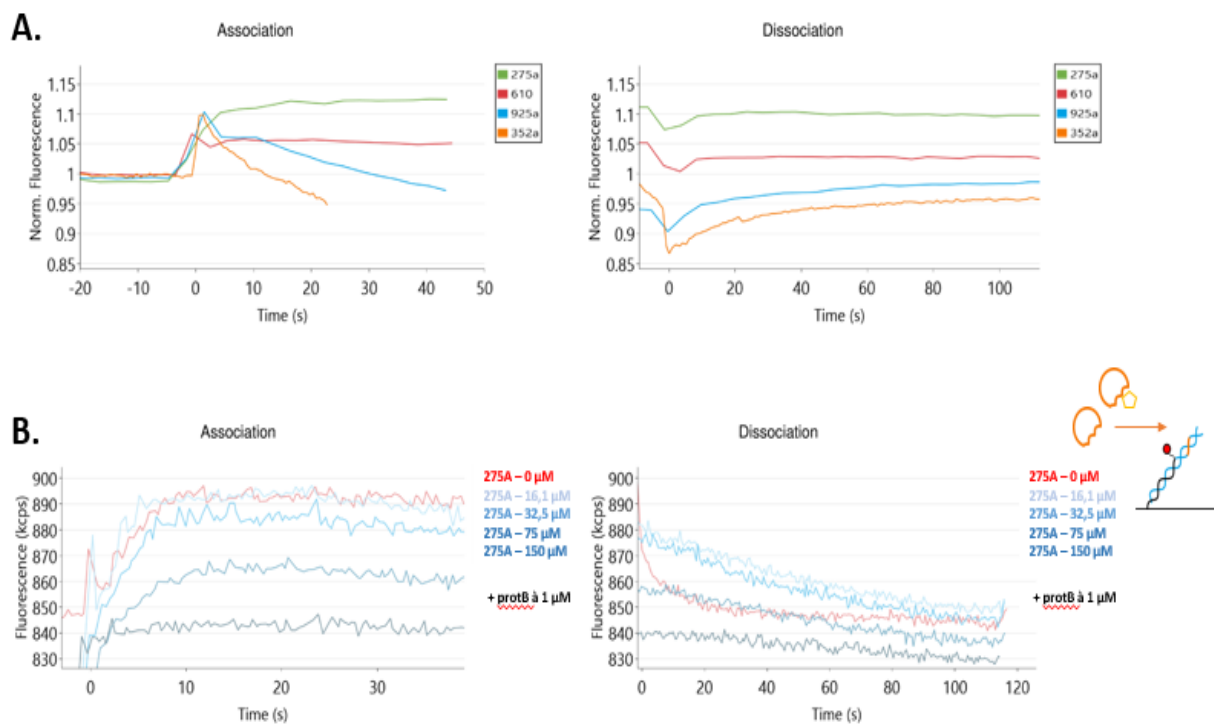


Figure 95 : Résultats de switchSENSE ; caractérisation de l'inhibition de l'interaction protéineB-ADN1. **A.** Test de cinétique de 4 inhibiteurs sur l'ADN d'intérêt double brin pour visualiser leurs interactions éventuelles avec l'ADN1. **B.** Cinétique de la protéine B co-incubée avec le composé 275A à différentes concentrations en composé

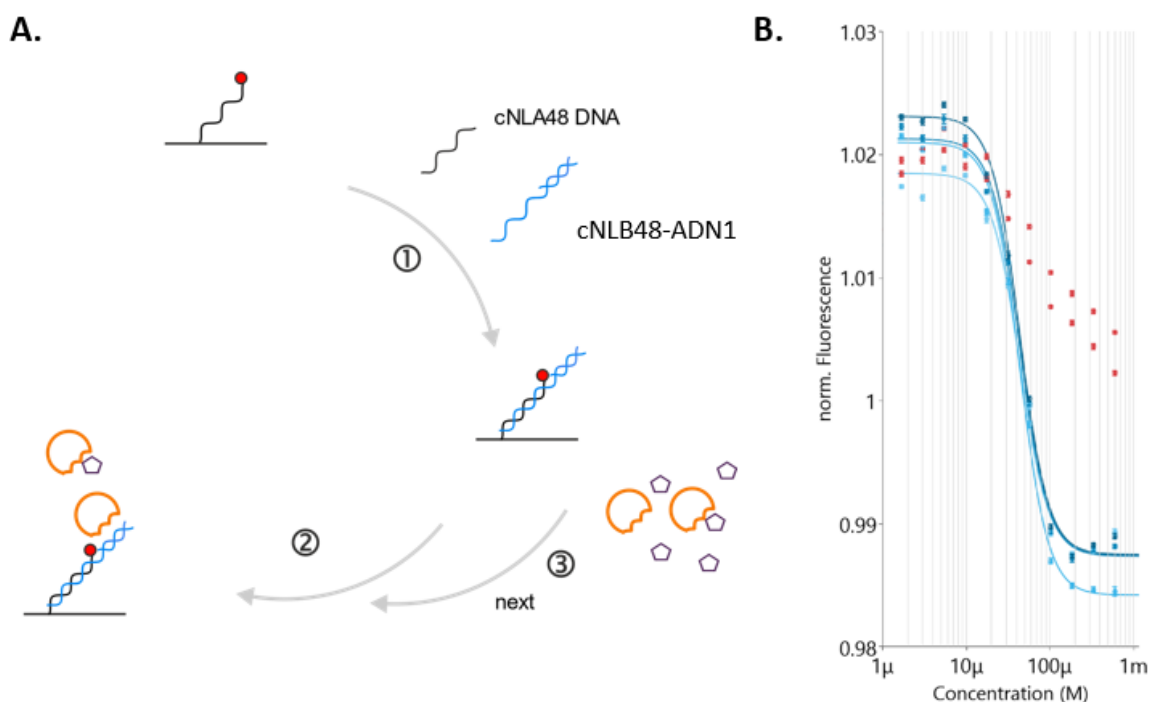


Figure 96 : **A.** Schéma de l'expérience de titration réalisée en switchSENSE pour une concentration en protéine B fixée à 1 μ M co-incubée avec un composé 275A à 12 concentrations allant de 600 μ M à 934 nM avec un facteur de dilution de 1.8. **B.** Résultats de titration en switchSENSE avec en bleu les résultats obtenus sur 4 électrodes de mesure contenant l'ADN1 et en rouge sur deux électrodes de référence

Le tableau suivant résume tous les résultats concernant la caractérisation d'inhibiteurs des interactions protéineB-ADN1.

Interaction Protéine B – ADN 1		
Composé	K_D protéine B - composé	Commentaire
275A	$43,9 \pm 1,2 \mu$ M	Se lie à l'ADN de manière non-spécifique à partir de 100 μ M
9689	$104 \pm 38 \mu$ M	
4931	$317 \pm 85 \mu$ M	
1278	$123 \pm 46 \mu$	
5764	-	Données non-ajustables

CONCLUSIONS ET PERSPECTIVES

PARTIE I :

A) Caractérisation des interactions protéine-ADN pour un projet industriel d'oncologie

Un volet de ma thèse s'est inscrit dans un projet industriel en collaboration avec Sanofi. Une CIFRE a été mise en place entre le service « Bio-Structure and Biophysics » de Sanofi et l'« Institut de Biologie Intégrative de la cellule » au CEA de Saclay pour évaluer la valeur ajoutée d'une nouvelle approche biophysique, le switchSENSE, lors de mesures des interactions des protéines. Mon rôle a été premièrement de valider la technique comme possible pour l'étude d'interactions de type protéine-ADN sur un projet d'oncologie. Des tests préliminaires ont dans un premier temps été réalisés. J'ai montré également que les deux modes statique et dynamique pouvaient être appliqués car j'ai obtenu des valeurs de constantes de dissociation (K_D) et de constantes cinétiques (k_{ON} et k_{OFF}) semblables entre ces deux modes. J'ai réussi à caractériser sept interactions sur les huit à étudier (quatre constructions protéiques et deux ADN).

B) Caractérisation d'inhibition des interactions protéine-ADN

Après avoir caractérisé les interactions de type protéine-ADN sur le projet SANOFI, je devais analyser l'interaction de plusieurs inhibiteurs évalués en parallèle chez Sanofi. Ces inhibiteurs sont plus précisément des fragments qui ont été au préalable testés par d'autres méthodes robustes de biophysique telles que la SPR, la MST et la DSF. J'ai obtenu des résultats de switchSENSE en accord avec les autres techniques biophysiques. Ces derniers résultats ne peuvent cependant pas être publiés pour une raison de confidentialité pour l'instant. Certains fragments ont montré une inhibition d'interaction. Pour d'autres, j'ai pu montrer une interaction non-spécifique soit avec l'ADN 48 pb hybridé sur la puce (ADN de référence) soit avec l'ADN1, un des ADN présentant une séquence d'intérêt. Le but était de trouver de nouveaux candidats inhibant l'interaction protéine-ADN mais interagissant avec la protéine. Plusieurs constantes d'inhibition ont été mesurées. D'autres inhibiteurs doivent encore être testés notamment pour l'interaction protéine-ADN2. Cette méthode a montré une reproductibilité des expériences, une rapidité et une consommation faible en échantillons. Les valeurs obtenues sont

en bon accord avec celles mesurées en parallèle sur certains de ces inhibiteurs par Sanofi. Cette technique pourrait potentiellement s'insérer dans un service comme celui de Sanofi. Il faut cependant noter que de nouvelles méthodes de mesure des interactions des protéines sont régulièrement proposées et qu'un service comme « Bio-Structure and Biophysics » de Sanofi ou même un institut comme l'I2BC ne peuvent investir dans toutes ces techniques. Ces approches demandent en effet des ingénieurs dédiés pour celles-ci et s'accompagnent d'un budget significatif tant en investissement qu'en maintenance. Pour ces raisons, les laboratoires industriels ou académiques investissent dans de nouvelles approches que si ces dernières ont une très forte valeur ajoutée.

Ce projet suggère que la technique switchSENSE est une technique intéressante dans le contexte de la validation d'inhibiteurs d'une interaction protéine-ADN. Les nombreuses discussions de notre laboratoire et d'autres laboratoires académiques avec la compagnie Dynamic Biosensors (DBS) les ont amenés à faire évoluer leur technologie. Nous avons en effet observé une perte assez rapide de fluorescence des puces. Celles-ci étaient proposées comme pouvant être régénérées environ 200 fois. Nous avons observé un nombre beaucoup plus faible d'utilisations comme d'autres laboratoires (en moyenne 20 fois). DBS a récemment tenu compte de ces retours et a proposé une puce avec un ADN simple brin sans fluorophore (*black chip*). L'ADN fluorescent est apporté dans un premier temps en étant hybridé à son ADN complémentaire, l'ensemble ayant une séquence complémentaire de l'ADN greffé sur la puce. Ces puces ont une durée de vie plus longue car la fluorescence est apportée indépendamment de la puce. Dans le même temps, DBS a proposé également une puce avec des ADN origamis (Rothenmund 2006) qui permet de mesurer des interactions de type protéine-ADN pour des protéines de très haut poids moléculaire.

PARTIE II :

A) Caractérisations biochimique et structurale du complexe KuCC-hDNA-IP6

Un autre volet de ma thèse s'est inscrit dans l'étude d'interactions de type protéine-ADN ou de type protéine-petites molécules du NHEJ. Afin de caractériser l'interaction entre Ku et

différents de ses partenaires, j'ai été formée premièrement en culture de cellules d'insecte permettant d'exprimer plusieurs protéines recombinantes (Ku70/Ku80 et XLF). Il a fallu pour les études biophysiques et de biologie structurale, purifier en grande quantité le complexe Ku et la protéine XLF à partir des culots de cellules d'insecte. Une quarantaine de milligrammes de chaque protéine ont été purifiées pour un volume de 600 mL de culture. Différentes versions de l'hétérodimère Ku ont été produites : KuFL, KuCC (version tronquée des deux domaines C-terminaux flexibles) et KuSC (version de KuCC avec une vingtaine de résidus supprimés en Nter de Ku70). Lors des purifications, nous observons deux pics d'éluion correspondant à la protéine Ku. Des analyses de spectrométrie de masse ont montré que la protéine Ku était acétylée de façon équivalente au sein de ces deux pics. Des études supplémentaires doivent être réalisées pour essayer de comprendre l'origine de ces deux pics que nous n'observons pas systématiquement. La construction KuCC est la seule qui a été cristallisée et qui a permis d'obtenir la structure de KuCC-hDNA-IP6 à une résolution de 3,9 Å. Nous avons utilisé pour cette étude par rayons X, l'ADN tige-boucle utilisé en 2001 pour l'obtention de la structure 1JEY et 1JEQ. Différentes stratégies ont été essayées tout au long de la thèse pour essayer d'optimiser les cristaux du complexe KuCC-hDNA-IP6 et leur diffraction: (i) variation des conditions de cristallisation, (ii) variation des solutions de cryo-protection pour la congélation, (iii) essais d'ensemencement ou de coupure à la trypsine chez Sanofi. L'ensemble de ces essais n'a pas permis d'obtenir de jeux de données à une résolution meilleure que 3,9 Å. Les travaux précédents au laboratoire sur des complexes KuCC-hDNA-pXLF ou -pAPLF avaient permis d'avoir des jeux de données respectivement à 2,8 Å et 3 Å (Nemoz et al. 2018). Des essais sont actuellement en cours au laboratoire pour faire diffuser dans des cristaux de KuCC-hDNA-pXLF la molécules d'IP6. Ils pourraient permettre d'améliorer la résolution d'une structure en présence d'IP6 et d'évaluer l'influence éventuellement du motif XLF sur la fixation de IP6.

La localisation du site de fixation a été caractérisée à l'interface des deux sous-unités de Ku et les résidus de Ku impliqués dans l'interaction Ku-IP6 ont été mis en évidence. Cette étude vient compléter et préciser des études précédentes de modélisation qui avaient mis en évidence le rôle de 3 acides aminés sur 9 acides aminés caractérisés dans cette thèse. Cette modélisation avait également proposé un rôle de 2 acides aminés qui ne sont pas au contact de l'IP6. L'inositol phosphate est à une distance de 20 Å de la molécule d'ADN en double hélice. Enfin, en superposant les structures de KuCC-hDNA-pXLF et KuCC-hDNA-IP6, nous nous sommes aperçus que l'inositol phosphate était localisé dans un site proche de celui de pXLF.

B) Etude biophysique de l'interaction Ku-IP6 et Ku-ADN-IP6

Une partie de ma thèse était centrée sur l'étude d'interactions de type protéine-ADN du NHEJ ou de type protéine-petites molécules par différentes méthodes notamment de biophysique. Dans le but de confirmer, premièrement, les résultats de la littérature, je me suis intéressée à l'étude de l'interaction entre Ku et la petite molécule d'inositol phosphate IP6. Ku est très affine pour l'ADN avec un K_D de l'ordre de 5 nM. L'interaction entre Ku et la molécule d'ADN est de plus indépendante de sa séquence. Nous voulions comprendre comment la voie NHEJ était stimulée à partir de l'interaction entre l'IP6 et l'hétérodimère Ku comme cela avait été proposé. Pour ce faire, des mesures de microcalorimétrie ont permis notamment de mesurer les valeurs thermodynamiques de différentes interactions impliquant l'IP6. Nous avons pu mesurer une affinité forte entre Ku et l'IP6 avec un K_D d'environ 50 nM. L'affinité entre IP6 et Ku est indépendante de la présence d'ADN au sein de Ku. Ce résultat est concordant avec le fait que l'IP6 soit localisé en surface de Ku à une distance de l'ADN de 20 Å. L'effet de stimulation a été mesuré en co-incubant l'IP6 avec Ku et en injectant l'ADN. En présence d'IP6, la molécule d'ADN est 10 fois plus affine avec l'hétérodimère. Ces résultats suggèrent que la stimulation du NHEJ par l'inositol pourrait provenir d'une interaction plus forte de Ku pour les extrémités des cassures de l'ADN. Les expériences ont été réalisées principalement avec la construction KuCC cristallisée. Cependant, des mesures avec KuSC ont suggéré que les premiers résidus de la sous-unité de Ku70, avaient une influence sur l'interaction Ku-IP6. La région N-terminale de Ku70 est fortement chargée en acides aminés acides (10 Glu et 2 Asp sur 20 résidus). L'influence de cette région sur la fixation de l'IP6 pourrait venir d'une occupation partielle de la région N-terminale de Ku70 sur le patch basique où se fixe l'IP6.

Les expériences de microcalorimétrie entre le partenaire Ku et les partenaires APLF et XLF du NHEJ ont permis d'évaluer leurs affinités respectives. Le peptide du A-KMB interagit de manière plus forte avec Ku (K_D d'environ 20 nM) par rapport au peptide du X-KBM (K_D d'environ 4 μ M). Ces mesures ont permis de mieux comprendre le mode de fixation de ces deux motifs sur Ku et de mesurer l'affinité de mutants de séparation de XLF qui ont été utilisés pour les mesures de recrutement après micro-irradiations dans les cellules.

Ma thèse étant inscrite dans une optique d'utilisation de diverses techniques anciennes comme innovantes de biophysique, j'ai décidé de poursuivre l'étude sur l'effet de l'IP6 sur le complexe Ku-ADN par switchSENSE. Ces mesures d'interactions de Ku sur l'ADN en présence ou en absence d'IP6 par switchSENSE, ont été réalisées avec l'aide de l'équipe de Dynamic

Biosensors. Les mesures effectuées indiquent une affinité plus forte de Ku pour l'ADN en présence d'IP6 en accord avec les mesures réalisées en microcalorimétrie. Les paramètres de cinétique ont de plus permis de dire qu'en présence de cette petite molécule, Ku est recrutée plus rapidement sur l'ADN et dissociée plus lentement. Ceci suggère que Ku serait recrutée au niveau des extrémités des cassures double-brin plus rapidement lorsque l'IP6 est présent et que la protéine se dissocierait plus lentement. Ces deux variations iraient dans le sens d'une voie NHEJ plus efficace. Une collaboration avec l'équipe de Eli Rothenberg, a permis de réaliser des expériences à la fois de sm-TIRF et de sm-FRET. Les résultats de sm-TIRF ont confirmé ceux de switchSENSE à savoir que l'association de Ku sur l'ADN est très rapide en présence d'IP6 et la protéine reste liée à l'ADN sur un temps très long.

A l'heure actuelle, la radiothérapie est inefficace pour guérir des patients atteints de cancers. Une des raisons est l'apparition de cellules tumorales résistantes qui réparent efficacement les cassures double-brin générées par radiothérapie par la voie NHEJ. Le laboratoire propose d'utiliser la molécule d'IP6 comme plateforme pour greffer des fonctions chimiques qui transformeraient cette molécule stimulatrice en une molécule inhibitrice. Il s'agirait notamment d'étendre cette molécule vers l'entrée de l'anneau de Ku en utilisant les informations issues de la structure cristalline avec l'IP6.

Afin de permettre une avancée plus rapide dans la recherche de nouveaux inhibiteurs du NHEJ, j'ai mené en parallèle par DSF, une étude de criblage d'une collection de fragments Sanofi sur Ku. Le but était de pouvoir trouver des fragments localisés sur des sites proches sur Ku et d'établir une liaison entre-eux pour les rendre plus affins. Cette étude a permis de sélectionner 49 composés stabilisateurs de la protéine Ku. Après mesure des doses-réponses, une sélection de 15 composés a été faite pour en conserver ensuite deux qui appartiennent au domaine public. Des études d'interaction entre Ku et les deux fragments ont montré que ces deux composés interagissaient avec la protéine. Lors des expériences de compétition avec l'IP6, nous avons montré notamment que le composé « EX00045853 » était déplacé par l'IP6 et qu'il était donc situé sur le même site de Ku70/80. Il serait donc intéressant de réaliser des tests de cristallisation avec ce composé, soit par ensemencement, soit par trempage pour déterminer la structure du complexe Ku-ADN- EX00045853 ou de Ku-ADN- EX00045853-IP6. Une des limites actuelles de cette approche par fragments sur Ku est la résolution moyenne des cristaux généralement obtenus. Pour continuer sur l'approche fragment, de nouvelles constructions de Ku seraient nécessaires pour obtenir de façon régulière des cristaux diffractant à une résolution meilleure que 3,9 Å. En ce qui concerne les 13 autres composés, ils pourront également être étudiés au

travers d'un contrat entre l'I2BC et Sanofi afin de pouvoir utiliser dans ce projet ces composés qui sont propriétés de Sanofi.

ANNEXES

Interactions measurements of protein-DNA complexes by Isothermal Titration Calorimetry (ITC) and Microscale Thermophoresis (MST)

Amandine Gontier^{1,2}, Paloma F. Varela^{1,2}, Clement Nemoz¹, Magali Aumont-Nicaise¹, Michel Desmadril¹, Jean Baptiste Charbonnier^{1,3}

1 Institute for Integrative Biology of the Cell (I2BC), Institute Joliot, CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91198, Gif-sur-Yvette Cedex, France

2 These authors contribute equally to the study

3 Corresponding author: jb.charbonnier@i2bc.paris-saclay.fr

Abstract

Interactions between protein complexes and DNA are central regulators of the cell life. They control the activation and inactivation of a large set of nuclear processes including transcription, replication, recombination, repair and chromosome structures. In the literature, protein-DNA interactions are characterized by a large spectrum of approaches including large scale studies and *in cellulo* analyses. Biophysical approaches with purified materials enable to evaluate if these interactions are direct or not. They provide quantitative information on the strength and the specificity of the interactions between a protein complex and different DNA substrates. Isothermal titration calorimetry (ITC) and Microscale Thermophoresis (MST) are widely used and complementary methods to quantitatively measure protein-DNA interactions. We present here protocols to analyse the interaction between a DNA repair complex, Ku70-Ku80 (Ku) (150kDa) and DNA substrates. ITC is a label free method performed with both partners in solution. It enables to determine the dissociation constant (K_d), the enthalpy (ΔH) and the stoichiometry N of an interaction. MST enables to measure the K_d with the protein or the DNA labelled with a fluorescent probe. We report the information obtained on Ku-DNA interactions with ITC and MST and discuss advantages and drawbacks of both methods.

Key words: nucleo-protein complexes, thermodynamic parameters, microcalorimetry, fluorescence, thermophoresis, double strand break repair, NHEJ

1. Introduction

The molecular mechanisms in the nucleus are regulated by intricate networks of protein-protein and protein-DNA interactions. The word “interaction” itself is used in the literature for many purposes. It defines associations characterized by large scale studies with two-hybrid or tap-tag but also by *in cellulo* proximity ligation assay (PLA) or co-immunoprecipitation, by genetic analyses as well as by biophysical approaches and structural studies of complexes [1, 2]. This large set of methods leads in some case to controversy. At opposite, these multiple analyses enable, in many other cases, a complete evaluation of the relevance of a given protein complex. Protein interaction measurements in the cellular context enable a characterization of the interactions in the presence of potential competitors and of post-translational modifications at different steps of the cell cycle. Biophysical approaches nicely complement these data by confirming or not the direct interaction with purified proteins and ligands. These measurements enable to characterize the strength, the specificity and the stoichiometry of the interactions in a quantitative manner. Interactions measured *in cellulo* and not confirmed *in vitro* can also be informative of the role of an additional partner or the need for post-translational modifications absent on the purified proteins.

A large panel of biophysical approaches enable nowadays to perform quality controls and interaction analyses on macromolecular samples (some of these are presented in other chapters of this issue). Each methodology offers diverse information with specific requirements on sample preparation. We present here protocols for isothermal titration calorimetry (ITC) and microscale thermophoresis (MST) analyses of the interaction between a protein complex and its DNA substrates. The complex is formed by the heterodimer Ku70-Ku80 (Ku). Ku is a core factor of the main double-strand break (DSB) repair pathway in human, called non-homologous end-joining (NHEJ). Ku recognizes the DSB ends through its ring shape structure [3] and it has a high affinity, in the nanomolar range, for small double strand DNA (15bp minimum). The K_d of the Ku-DNA interaction were initially characterized by gel shift assays (EMSA) [4] and fluorescent anisotropy analyses [5]. Ku can bind DNA in a sequence independent manner and *in vitro* one Ku molecule binds every 15-20bp of DNA [3]. Interestingly, a super resolution microscopy study showed that only one or two Ku molecules binds to DSBs ends in a cellular context suggesting a limited threading of Ku from the DNA DSB ends *in cellulo* [6]. Ku plays also a central role for the recruitment of the enzymatic activities (nucleases, polymerases, ligases) that process and ligate the DSB ends [7]. Our laboratory recently described at the molecular levels the mechanism of recruitment by Ku of some of these downstream NHEJ factors [8]. Here, we report the measurements by ITC of the thermodynamic parameters of the interaction between Ku and different DNA substrates and by MST the same interaction by using

DNA labelled with a fluorescent probe. We previously used ITC to characterize protein-protein and protein-DNA interactions involved in the NHEJ pathway [9-11]

Isothermal titration calorimetry measures the heat generated or consumed upon the interaction of two molecules, called hereafter the receptor and the ligand [12, 13]. The molecular interaction between the receptor and the ligand can be defined by the following equation: $\Delta G^\circ = -RT \ln K_a = RT \ln K_d = \Delta H - T \Delta S$, where ΔG° (kcal.M⁻¹) is the standard Gibbs free energy, R (kcal.M⁻¹) is the gas constant, T (K) is the temperature, K_a (M⁻¹) is the equilibrium association constant and K_d (M) is the equilibrium dissociation constant. The second part indicates that the free energy is the sum of an enthalpy term ΔH (kcal.M⁻¹) and an entropic term $-T\Delta S$ (kcal.M⁻¹). [Figure 1](#) represents a schematic representation of the calorimeter. A constant power is applied to the reference cell. A very small temperature difference is maintained between the sample and reference cells thanks to a feedback circuit that applies a variable power to the sample cell. The injection of a small volume of ligand loaded in the syringe will lead to heat exchange when the ligand interacts with the receptor. Exothermic reactions generate heat and induce a decrease in the feedback power. Endothermic reactions increase the feedback power. The instrument changes the power applied to the reference cell according to the nature of the interaction. If the reaction is exothermic, it will reduce the power and we observe a negative peak in the thermogram ([Figure 1B, top](#)). At opposite, if the reaction is endothermic, it will increase the power and we observe a positive peak in the thermogram (as in the Ku-DNA example presented later). The isotherm of titration ([Figure 1B, bottom](#)) is calculated by integrating each peak of the thermogram and by reporting the calculated values that correspond to the enthalpy of the reaction versus the ratio of ligand/receptor molecules. The curve that best fit these points gives, in the case of a single binding site, the ΔH , K_d and stoichiometry of the interaction ([Figure 1B, bottom](#)). ITC is widely used since little optimisation is needed and the interaction is followed without labelling the partners. Both partners are in solution at opposite to some methods where one partner is linked to a surface (SPR, BLI, SwitchSense). Its main drawback is the important quantity of materials often needed to have detectable heat exchange.

Microscale thermophoresis (MST) is a more recent approach [14, 15]. It enables to measure the constant affinity (K_d) between a receptor and a ligand. This method can use either the intrinsic fluorescence of the protein (Monolith NT.LabelFree) or the fluorescence signal (Monolith NT.115) of a probe linked to the receptor or to the ligand. A temperature gradient is applied overtime by an IR laser on a capillary containing the fluorescent receptor alone or with

increasing concentration of the ligand (Figure 2A). The laser induces a local heat on the capillary and the diffusion of the fluorescent receptor away from the irradiated zone. The laser is turned off and the receptor diffuses again to the depleted zone. The diffusion rate of the receptor is highly sensitive to its hydration. The formation of the receptor-ligand complex modifies locally the hydration, size and charge of the receptor that can lead to detectable variation on the diffusion of the fluorescent receptor in the capillary. The fluorescence curves (called MST traces) are measured in presence of ligand concentrations varying between 1/10 and 10-fold of the expected K_d (Figure 2B). The dissociation constant (K_d) is estimated by fitting with a single site model, the normalized fluorescence against the ligand concentration (Figure 2C). MST is widely used because it requires small amount of material (typically in the 100ng-1 μ g range), it is really easy to set up and results can be obtained in a few hours.

2. Materials

2.1 Measurement equipment for microcalorimetry

1. MicroCal VP-ITC

The VP-ITC microcalorimeter (Malvern) is composed of two compartments or cells: one reference cell filled with water or buffer and one measurement cell (volume of 1,4ml) filled with the first partner called the receptor. Both are maintained at equal temperature and are positioned in an adiabatic environment. The ligand is injected with small injections (5-10 μ l) in the measurement cell. If the ligand interacts with the receptor, in most cases, some heat are formed or absorbed due to the interaction that is detected by the device which compensates this heat exchange by respectively reducing or increasing the temperature (Figure 1). In some rare case, at a given temperature, an interaction can happen without heat exchange ($\Delta H=0$). In that case, changing the temperature should unmask the interaction, if it exists, since the enthalpy varies with the temperature.

Another widely used device is the ITC-200 microcalorimeter that has a smaller cell (measurement cell of 200 μ L) and needs 40 μ l of ligand for the syringe. In our hand for some projects like Ku-DNA interaction, presented below, the VP-ITC gave nicer thermograms with a better signal/noise ratio. In other projects, ITC-200 proved to be more interesting in particular in terms of material used.

2. ThermoVacTM (Malvern). This device enables to degas the buffers and samples before use and to pre-equilibrate to a temperature close to the measurement temperature. It is also used as pump for cleaning procedures

3. VPViewerTM and OriginTM softwares.

VPViewer is used to command the VP-ITC and follow the experiment. Origin is used to analyse the raw data called thermograms.

2.2 Buffer and samples preparation (VP-ITC)

1. All buffers are prepared with deionized water ($18 \text{ M}\Omega\text{-cm}^{-1}$ at $25 \text{ }^\circ\text{C}$).
2. Prepare 2 L of dialysis buffer (D) of protein Ku70-Ku80 (Buffer D: 20 mM Tris-HCl pH 8.8, 150 mM NaCl, 5 mM β -mercaptoethanol).
3. Dialysis Membrane molecular weight cut-off (MWCO) of 30 kDa (Spectra/Por® Molecular porous membrane tubing).
4. Amicon® Ultra-4 30 kDa for protein concentration.
5. Ku70-Ku80 complex purified at $1 \text{ }\mu\text{M}$ (0.15 mg/ml). Volume 2,3ml for measurement cell plus void volume
6. DNA 18 bp or 42bp DNA at $10 \text{ }\mu\text{M}$. Volume 600 μL to fill the syringe

2.3 Measurement equipment: Microscale Thermophoresis (MST)

The Monolith NT.115 Blue/Green (equipped with two leds, Led 1 Excitation. 470nm-Emission. 520nm/ Led 2 Excitation 550nm-Emission. 600nm, Nanotemper Technologies) is equipped with the standard sample tray to process automatically up to 16 samples per run in 10min. Capillary volumes are 3-5 μl and the labelled molecule are classically used at concentration as low as 5-10nM. Control software guides step-by-step planning and assay setup. Binding values are obtained instantly and results are exportable. Affinity Analysis software enables advanced data analysis like merging, grouping or comparing data sets.

2.4 Labelling kits (MST)

Nanotemper labelling kits contain fluorescence dyes that are suited for MST measurements. The protocols are optimized to ensure a good labelling efficiency and effective removal of free dye and about a 1:1 ratio of labelled protein to dye. Two coupling chemistries are available, one for primary amines (NHS for e.g. lysine labelling) and one for thiol groups (Maleimide for cysteine labelling). 2nd generation dyes are now available (tris-NTA) but only for the instrument with red led. Proteins with a molecular weight of $> 5\text{kDa}$ can be labelled.

2.5 Capillaries (MST)

All capillaries are made of glass with infrared laser quality. Different types of capillaries are available (standards, hydrophobic and premium). Standard Treated Capillaries are physically treated to ensure high quality surface properties. Hydrophobic and Premium Coated Capillaries are covalently coated with a dense brush of specially designed polymers to avoid molecule adsorption.

2.6 Buffers (MST)

Labelling and MST buffers are available with the labelling kits and the capillaries. The composition of the MST buffer is 50mM Tris-HCl pH 7.4, 150mM NaCl, 10mM MgCl₂, 0.05 % Tween-20.

3. Methods

3.1 Methods for ITC

1. The Ku70-Ku80 heterodimer (receptor) is dialyzed against buffer D (D : 20 mM Tris pH 8.0, 150 mM NaCl, 5 mM β -mercaptoethanol). The dialysis buffer is changed 3 times and each dialysis is performed during two hours minimum against a volume of buffer corresponding to minimum 30 times the volume of receptor sample. Some buffer (about 50ml) corresponding to the last dialysis buffer surrounding the sample is kept as reference buffer. It will be used to rinse the measurement cells and the syringe and to adjust Ku and DNA samples to the concentration chosen.

2. Preparation of the receptor (Ku complex) for the measurement cell: Rule of factor c is applied to define the concentration of the receptor in the sample cell. c value is determined by the formula, $c = N \times K_a \times [Ro]$, where N is the stoichiometry of the reaction (number of ligand molecule bound by molecule of receptor), K_a is the association constant (M^{-1}) and $[Ro]$ the initial concentration of the receptor in the sample cell (M). The aim is to have a c value between 10 and 100 to obtain an S-form titration curve with two plateaus ([Figure 1B, bottom](#)). In our example Ku-DNA, with an expected K_d of about 10nM and N of 1, the $[Ro]$ should be between 100nM and 1 μ M. We chose a concentration of 1 μ M. It is classically obtained at 0.5 μ M at the end of the purification from insect cells pellets. A first step is an affinity purification in batch with Ni-NTA beads (Thermo Fisher Scientific) (batch volume of 20mL for 600ml of insect cells culture). A second step is an anion exchange, Resource-Q of 6mL (GE Healthcare Life Science). Ku is concentrated to 1 μ M using Ultra-4 (Amicon). The concentration is estimated using a Nanodrop 2000 (Thermo Fisher Scientific) and UV absorption at 280nm.

3. Preparation of the ligand (DNA substrates). The DNA is classically prepared at 500 μ M. The two complementary strands are mixed at a temperature of 95 °C for 5 min and cooled down overnight at room temperature. The concentration is estimated using a Nanodrop 2000 (Thermo Fisher Scientific) and UV absorption at 260 nm. It is possible to use the Mini-Q (Sigma-Aldrich) to check if the DNA is correctly hybridized in double-strand and the proportion of remaining single strand DNA. The ligand is prepared at a concentration equal to 10-fold the receptor concentration (i.e 10 μ M here). DNA is diluted to 10 μ M using the last dialysis buffer to limit large heats of dilution due to buffer differences (background heats).

4. The receptor (volume 2.5ml) and the DNA (volume 600 μ l) are degassed and heated for 10 min to the measurement temperature (here 20 °C) using the ThermoVac.

5. The receptor (Ku70-Ku80) is deposited in the measurement cell with a 2.5ml glass syringe. The sample is put slowly to avoid bubbles in the cell. The injection is stopped when the liquid comes above a visible level.

6. The ligand (DNA) is loaded on the injection syringe (in open position) thanks to a 10ml plastic syringe connected to the top of the ITC syringe. The syringe is then purged and filled twice to remove bubbles.

7. We analysed the interactions of Ku with a 18bp DNA (Figure 3B) and a 42bp DNA (Figure 3C) and observed two endothermic reactions (positive peaks). The interaction with the 18bp DNA shows a stoichiometry of 0.97 molecules of 18bp DNA bound to one Ku, a K_d of 3.5 ± 0.8 , and an entropy driven reaction with a positive unfavourable enthalpy. The interaction with the 42bp DNA gives a stoichiometry of 0.48 molecule of 42bp DNA bound to one Ku in good agreement with the expected positioning of two Ku molecules on such DNA. A tighter affinity is measured with a K_d of 0.4 ± 0.2 nM and a reaction also driven by the entropy.

3.2 Methods for MST

1. MST is an easy to handle method but to ensure optimal results a few rules need to be respect. First, always spin down the stocks for 5min at 10,000-20,000xg to remove large aggregates, which is a main source of noise. Prepare always volumes of at less 20 μ l of sample in the smallest micro reaction tubes to avoid surface absorption.

2. Preliminary assays with MST. First, we check the optimal concentration of the fluorescence molecules. Ideal fluorescent intensities are between 200 and 1,500 counts. Secondly, we check the capillary type that is optimal for our study (standards, hydrophobic and premium). The aim

is to obtain narrow, shoulder-free, regular and symmetrical fluorescent peaks. Finally, we check the optimal buffer composition. BSA and detergents can be used to improve the homogeneity of the samples. pH and/or ionic strength are also parameters that can be adjusted if needed. Reducing agents freshly prepared can also be used when needed.

3. Preliminary binding assays by MST. We will call receptor the fluorescently labelled molecules and ligand the non-labelled one hereafter. The highest concentration of the ligand used should be high enough to saturate the receptor. It is recommended to use a highest concentration of ligand equal to 20-fold the expected K_d value. In the first assays, we compare the fluorescent signals of the receptor in the absence and in the presence of the higher concentration available of the ligand. The DNA solution (vol 10 μ l of [DNA-Fam5] at 10 nM) is mixed with 10 μ l buffer or with 10 μ l of Ku at 250 nM concentration. We analysed fluorescence intensity and homogeneity, no aggregation on MST traces and the signal to noise ratio by repeating the measurement 4-fold. The signal is defined as the response amplitude (difference of F_{norm} between receptor alone and receptor-ligand complex) and the noise is the standard deviation of errors between replicates. A signal-to-noise ratio less than 5 suggests a too weak signal to measure a K_d , a ratio of 5 or more indicates a quite favourable case for a K_d measurement and a ratio higher than 12 reflects optimal conditions to measure a K_d with MST

4. Determination of K_d by MST. Prepare 16 small micro-tubes and transfer 20 μ l of the highest concentration of the ligand (unlabelled molecule) to the first one. Fill 10 μ l of the optimal assay buffers into tubes 2-16. Transfer 10 μ l of tube 1 to tube 2 and mix well by pipetting up and down several times. Do the same until reach tube 16 to obtain a serial dilution. Remove 10 μ l from the tube 16 after mixing. Add 10 μ l of labelled molecule to each tubes and mix well by pipetting up and down several times. Incubate if needed before loading into the capillaries.

5. Assays with Ku labelled and DNA not labelled. Ku is labelled with NT495 final concentration 20nM. Titrant DNA 18bp is prepared at concentration from 123nM – 3.9pM by serial dilution. The buffer used is MST buffer and the capillaries are Premium capillaries. Measurements are performed at 22°C, at 40% excitation and 40% or 60% MST power. The signal to noise ratio is 3 in these conditions and does not enable to measure a reliable K_d (Figure 4A). The excitation power of 40% was enough to have around 350 fluorescence counts. Increasing the MST power increases the signal, but also the noise of the data.

6. Assays with DNA labelled and Ku not labelled. DNA is purchased with FAM5 in 5' and the fluorescent probe is separated to the 18bp by 3 nucleotides (Sigma Aldrich, HPLC 95% pure). The stock DNA is prepared at 231 μ M with the same hybridation protocol as for ITC (see

above). The final concentration of DNA in MST assays is 10nM. Ku is prepared by serial dilution from 1750nM-80pM. The buffer used is the MST buffer. Standard capillaries are used as well as low binding tubes and tips. Measurements are performed at 60%, 80% (data not shown) and 100% excitation and 40% MST power, 22°C of temperature (Figure 4B-D). Signal to noise at 60%, 80% and 100% are 5.8, 8.7 and 13.2, respectively. The Kd is poorly defined at 60% excitation power, Kd of 100 ± 500 nM (Figure 4B). We measured a better Kd of 10 ± 9 nM at 80% (standard error of regression of 2.30 and a reduced X2 is 1.04). At 100% excitation power, we measured the best Kd of 10 ± 4 nM (standard error of regression is 1.17 and reduced X² is 0.87) (Figure 4C-D)

4. Notes

Notes ITC

1-ITC Make sure before each experiment that measurement cell and syringe are clean. In our laboratory, the user clean the sample cell and the syringe respectively with 14% decon and rinse extensively with 1L and 250 ml of water respectively the sample cell and the syringe. An ITC measurement, called water against water, is then performed with 20 injection of water against water. The signal should be small ($DP < 0.01 \mu\text{cal}\cdot\text{s}^{-1}$) and regular. The next user can then safely use the device for his measurements. When needed, stronger cleaning procedures should be used (see recommendations of the manufacturer (Malvern)). The Hastelloy alloy that covers the ITC cells is compatible with most buffers except strong acids. The ITC instrument should be preferentially installed in a room with a control temperature (about 20°C).

2-ITC The buffer composition can influence the ITC signals. Reducing agents like DTT cause erratic baseline variations and other artefacts. TCEP is recommended over β -mercaptoethanol and at concentration ≤ 1 mM. The ionisation properties of the buffer can have an influence on the values measured by ITC in particular when the interface involves charges exchange. In that case, it is recommended to compare the values obtained with different buffers and at different ionic strengths (for example [16]).

3-ITC If no heat exchange is observed during an interaction, it can happen that the reaction exists but has a null enthalpy at the temperature chosen. To evaluate this possibility, the ITC measurement can be performed at a different temperature (for example 10°C lower or higher).

4-ITC The ITC gives access to a stoichiometry N measurement. In many reaction, a N of 1 is expected, one molecule of ligand bound by molecule of receptor. Even, in this case, a value of 1 is rarely measured for several possible reasons: (i) the presence of unfolded/inactive receptors

in the sample cell will lead to a lower N value, proportional to the fraction of inactive receptor. The presence of unfolded/inactive ligands in the syringe will lead to a higher value N value, proportional to the fraction of inactive ligand.

5-ITC In parallel to the measurement of the receptor-ligand interaction, it is important to perform the ITC reaction with the ligand in the syringe at the same concentration and buffer in place of receptor in the cell. This control will determine if the signals observed are not due to a dissociation of the ligand when it is introduced in the sample cell. This can happen in particular when the ligand is a protein than can form oligomers.

6-ITC Kinetic information can be recovered from ITC measurement using a method called kinITC that can applied for some interaction [17]

7-ITC advantages: (i) Fast analytical method (2h-3h), (ii) Non-destructive, the complex receptor-ligand can be re-purified on a size exclusion chromatography or anions/cations exchange to remove the ligand excess (iii) no labelling, (iv) no immobilization, (v) no molecular weight limitation

8-ITC limitations: (i) consume important quantities of protein and DNA in particular for K_d in the μM range (ii) difficult to detect directly K_d below nanomolar range; under nanomolar range displacement titration can be done [18]

Notes-MST

1-MST It is recommended to use low binding tubes and tips. For the serial dilution, better results are obtained without changing the tips from tube to tube. However when adding the labelling molecule, it is recommended to change the tip at each tube. Avoid making bubbles by pipetting carefully and never vortex these small volumes to avoid denaturation of samples. Handle capillaries with care and never touch at the centre where the measurement will be done.

2-MST An SDS denaturation test (SD-Test) has to be performed to distinguish between fluorescence changes caused by an interaction and those caused by non-specific effects, e.g. loss of protein due to aggregation or adsorption to labware. This test consists in the denaturation of all proteins contained in the sample using a mix of SDS and DTT and heating to 95°C . In this context, the receptor-ligand interaction is disrupted. After denaturation, the change of fluorescence should not be observed in absence or in presence of increasing concentration of ligands. If addition of ligands induces fluorescence change with denatured protein, this suggests some interference of the ligand and some new experimental conditions should be researched.

3-MST If no enough fluorescence is observed we can increase the concentration of the labelled receptor or increase the led power. A different strategy of labelling can be used. In cases of absorption to the capillary adding detergents or other amphiphilic polymer additive will help.

4-MST Under certain conditions, MST can also be used to obtain the stoichiometry of the interaction. For that we need to know precisely the concentration of receptor and ligand as well as the K_d of their interaction. In this case we need a much higher concentration of the receptor at least 20-fold than the K_d .

5-MST It can be worth to label the other molecule and repeat the experiment in the reverse way to confirm the result.

6-MST A positive and a negative control is advisable to confirm that there is no unspecific interaction or that the conditions of the assay are optimal

7-MST advantages (i) the method is easy and very fast to perform (ii) it consumes very few quantities of materials (iii) it is an immobilization free method (iv) there is minimal contamination of the sample (this method is entirely optical and contact-free) (v) it is able to measure complex mixtures (i.e. cell lysates, serum, detergents, liposomes) (vi) it allows wide size range for ligands (ions to MDa complexes).

8- MST limitations (i) one of the partners should be labelled with possible interference with the interaction sites (ii) the theory behind the measurement principle is less obvious to our opinion than for other biophysical approaches (iii) hydrophobic fluorescent labelling may cause non-specific binding (iv) no kinetic or thermodynamic information is obtained (v) highly sensitive to any change in molecular properties.

Figure Legends

Figure 1: (A) Schematic representation of an isothermal titration calorimeter. The reference and measurement cells are positioned in an adiabatic shield. A constant power is applied to the reference cell and a variable power is applied to the measurement cell to maintain a small constant temperature difference between the two cells thanks to a feedback circuit. Most interactions between a receptor in the cell and a small volume of ligand injected by the syringe generate heat (exothermic reaction). This triggers a decrease in the feedback power as represented in (B, top panel). The interaction can in some cases consumes heat and this leads to an increase the feedback power (endothermic reaction, as described later with Ku and DNA). The integrals of the peaks from the top panel enable to calculate the enthalpy exchange at each injection. It is reported versus the ratio of receptor and ligand. The fit of this curve is called the

isotherm of titration and it gives access to the enthalpy ΔH , the association constant K_a (and thus the $K_d=1/K_a$) and to the stoichiometry N .

Figure 2: (A) Schematic representation of microscale thermophoresis optic system. MST is measured in capillaries with a small sample volume of around $\sim 4\mu\text{l}$. The fluorescence within the capillary is excited and detected through an objective. A focused IR-Laser is used to locally heat a defined sample volume. Thermophoresis of fluorescent molecules through the temperature gradient is detected by the objective (B) Typical binding experiment. The thermophoretic movement of a fluorescent molecule (blue trace; “unbound”) changes upon binding to a non-fluorescent ligand (red trace; “bound”), resulting in different traces. For analysis, the change in thermophoresis is expressed as the change in the normalized fluorescence (F_{norm}), which is defined as $F_1(\text{hot})/F_0(\text{cold})$ (F -values correspond to average fluorescence values between defined areas marked by the red and blue cursors, respectively). (C) Titration of the non-fluorescent ligand results in a gradual change in thermophoresis, which is plotted as F_{norm} versus ligand concentration. The resulting binding curve can be fitted with a single binding site model and enables to define the K_d .

Figure 3: ITC measurements between Ku and DNA. (A) Crystal structure of Ku-DNA complex with a 14bp DNA. Ku has a ring shape structure. The DNA present in the crystal structure has an additional hairpin part that has been removed for clarity. A schema is presented of the two interactions analysed with Ku and either an 18bp DNA or 42bp DNA. (C) ITC measurement between Ku and an 18bp DNA. The interaction is endothermic with positive peaks on the thermogram. The heat exchange observed in this interaction is very weak ($<0,02 \mu\text{cal}/\text{sec}$) and underline the high sensitivity of the VP-ITC. The isotherm of titration enables to define ΔH , K_d and N of the reaction, reported in Table below. The affinity is in good agreement with the values reported in the literature by EMSA and anisotropy fluorescence. One Ku molecule binds to a 18bp DNA. The reaction is entropy driven with an unfavourable enthalpy ($\Delta H= + 5.1 \text{ kcal.mol}^{-1}$) compensated by a larger favourable entropy ($-T\Delta S= - 16.2 \text{ kcal.mol}^{-1}$) (D) ITC measurement between Ku and a 42bp hairpin. ΔH , K_d and N of the reaction are reported in Table below. Interestingly, ITC enable to characterize a stoichiometry of two Ku molecules bound on the 42bp DNA. The first injections of the interaction with the 42bp DNA show some drift of the baseline.

Figure 4: MST measurements between Ku70-Ku80 and a 18bp DNA. (A) Ku is labelled with NT-495 fluorescent probe. The signal to noise ratio is too weak and no Kd can be measured at 40% or 60% MST power (B) the 18bp DNA is labelled in 5' with Fam5 on one strand of the duplex. The measurements are performed at 60%, 80% (not shown) and 100% excitation power. The signal to noise ratio and the fit are better defined at 100% excitation (D) than at 60 % excitation. The MST traces observed at 100% excitation are shown in (C).

Acknowledgment

JBC is supported by ARC program (SLS220120605310), ANR (ANR-12-SVSE8-012), INCA DomRep (PLBIO 2012-280), CEFIPRA grant 5203C and by the French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INBS-05. AG is supported by a CIFRE PhD fellowship with Sanofi. We thank Pierre Soule from Nanotemper technologies for his availability and all the fruitful discussion. The experiments were performed on the Platform PIM (Platform for measurements of Interactions of Macromolecules) (<https://www.i2bc.paris-saclay.fr/spip.php?article280>)

References

1. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, *et al.* (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae* [see comments]. *Nature* 403(6770):623-627.
2. Gavin AC, Bosche M, Krause R, Grandi P, Marzioch M, Bauer A, *et al.* (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415(6868):141-147.
3. Walker JR, Corpina RA, Goldberg J (2001) Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* 412(6847):607-614.
4. Blier PR, Griffith AJ, Craft J, Hardin JA (1993) Binding of Ku protein to DNA. Measurement of affinity for ends and demonstration of binding to nicks. *J Biol Chem* 268(10):7594-7601.
5. Arosio D, Costantini S, Kong Y, Vindigni A (2004) Fluorescence anisotropy studies on the Ku-DNA interaction: anion and cation effects. (Translated from eng) *J Biol Chem* 279(41):42826-42835 (in eng).
6. Britton S, Coates J, Jackson SP (2013) A new method for high-resolution imaging of Ku foci to decipher mechanisms of DNA double-strand break repair. (Translated from eng) *J Cell Biol* 202(3):579-595 (in eng).

7. Chang HHY, Pannunzio NR, Adachi N, Lieber MR (2017) Non-homologous DNA end joining and alternative pathways to double-strand break repair. (Translated from eng) *Nat Rev Mol Cell Biol* 18(8):495-506 (in eng).
8. Nemoz C, Ropars V, Frit P, al, Calsou P, Charbonnier JB (2018) XLF and APLF bind to Ku80 on two remote sites to ensure DNA repair by non-homologous end-joining. *Nature structural & molecular biology*.
9. Tadi SK, Tellier-Lebegue C, Nemoz C, Drevet P, Audebert S, Roy S, *et al.* (2016) PAXX Is an Accessory c-NHEJ Factor that Associates with Ku70 and Has Overlapping Functions with XLF. (Translated from eng) *Cell Rep* 17(2):541-555 (in eng).
10. Ropars V, Drevet P, Legrand P, Bacconnais S, Amram J, Faure G, *et al.* (2011) Structural characterization of filaments formed by human Xrcc4-Cernunnos/XLF complex involved in nonhomologous DNA end-joining. (Translated from eng) *Proc Natl Acad Sci U S A* 108(31):12663-12668 (in eng).
11. Malivert L, Ropars V, Nunez M, Drevet P, Miron S, Faure G, *et al.* (2010) Delineation of the Xrcc4-interacting region in the globular head domain of cernunnos/XLF. (Translated from eng) *J Biol Chem* 285(34):26475-26483 (in eng).
12. Holdgate GA (2001) Making cool drugs hot: isothermal titration calorimetry as a tool to study binding energetics. *Biotechniques* 31(1):164-170.
13. Krell T (2008) Microcalorimetry: a response to challenges in modern biotechnology. (Translated from eng) *Microb Biotechnol* 1(2):126-136 (in eng).
14. Jerabek-Willemsen M, Wienken CJ, Braun D, Baaske P, Duhr S (2011) Molecular interaction studies using microscale thermophoresis. (Translated from eng) *Assay Drug Dev Technol* 9(4):342-353 (in eng).
15. Asmari M, Ratih R, Alhazmi HA, El Deeb S (2018) Thermophoresis for characterizing biomolecular interaction. (Translated from eng) *Methods* 146:107-119 (in eng).
16. Czarny B, Stura EA, Devel L, Vera L, Cassar-Lajeunesse E, Beau F, *et al.* (2013) Molecular determinants of a selective matrix metalloprotease-12 inhibitor: insights from crystallography and thermodynamic studies. (Translated from eng) *J Med Chem* 56(3):1149-1159 (in eng).
17. Dumas P, Ennifar E, Da Veiga C, Bec G, Palau W, Di Primo C, *et al.* (2016) Extending ITC to Kinetics with kinITC. (Translated from eng) *Methods Enzymol* 567:157-180 (in eng).

18. Velazquez-Campoy A, Freire E (2006) Isothermal titration calorimetry to determine association constants for high-affinity ligands. (Translated from eng) *Nat Protoc* 1(1):186-191 (in eng).

Figure 1 :

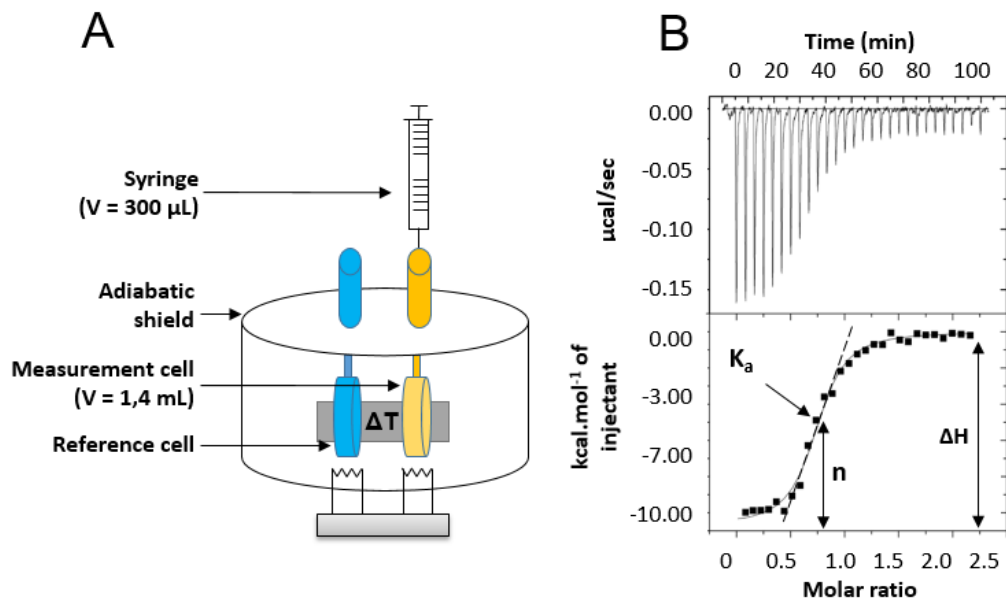


Figure 2

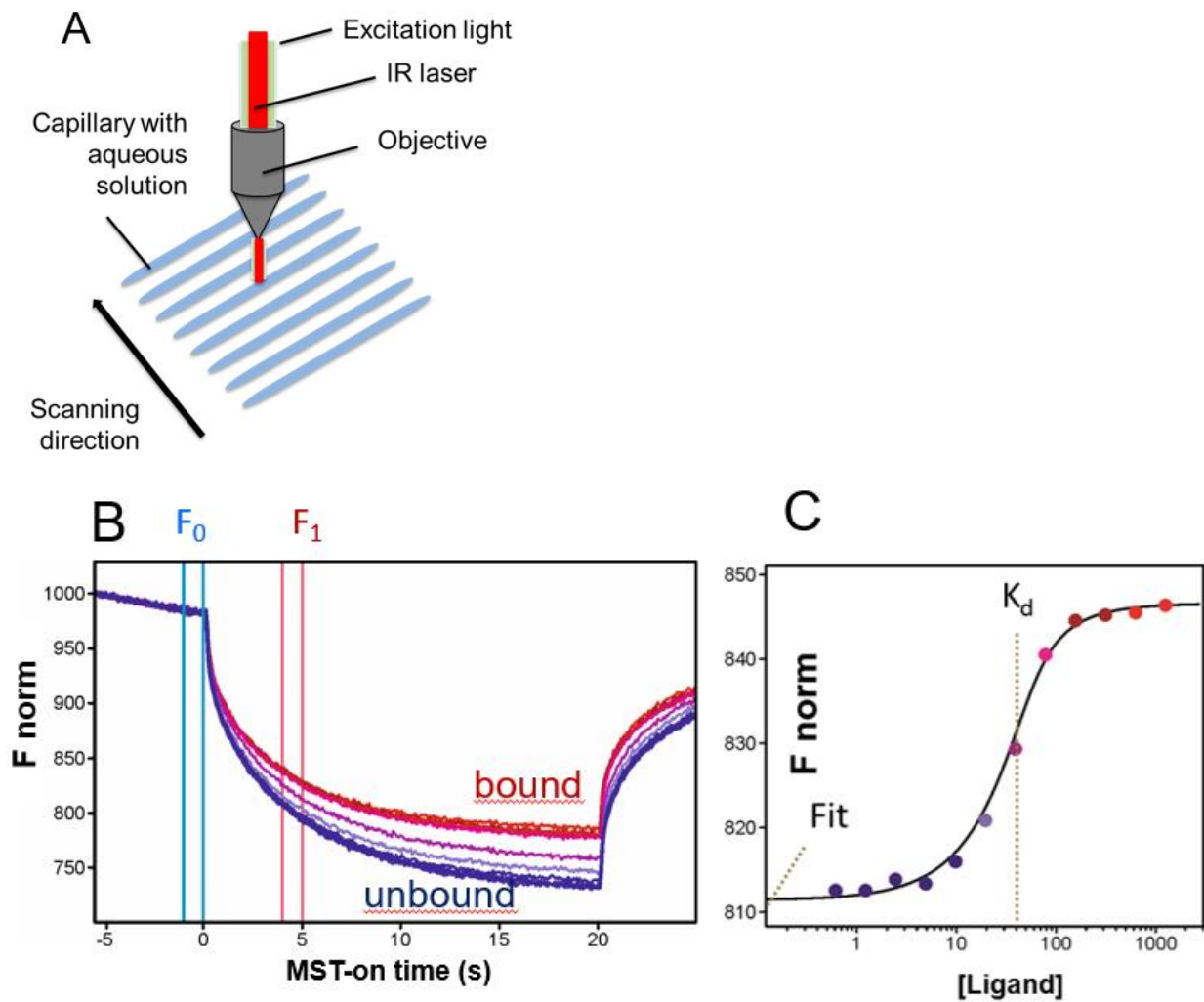
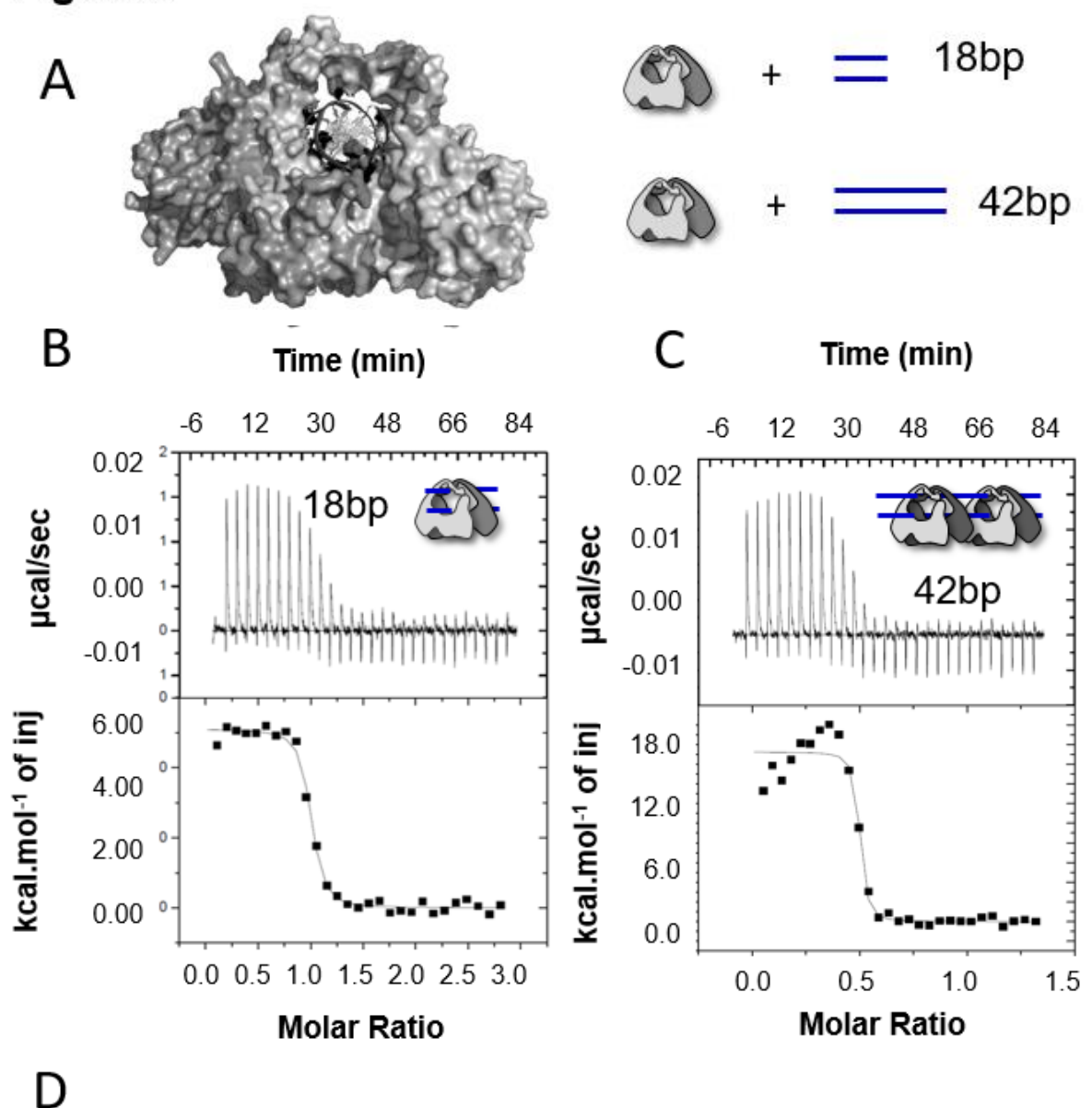
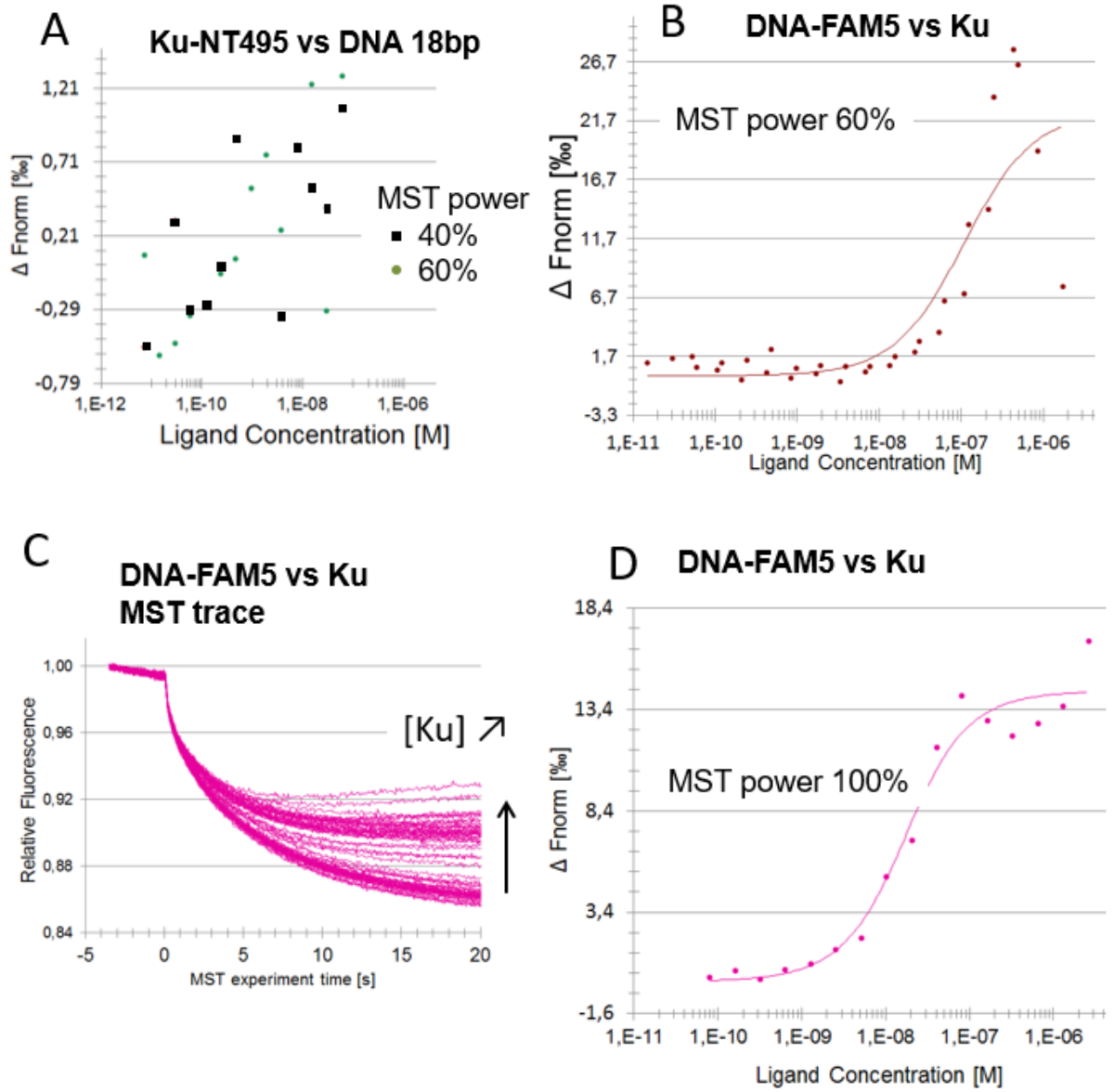


Figure 3



DNA	N	K_D (nM)	ΔH° ($\text{kcal}\cdot\text{mol}^{-1}$)	$-T\Delta S^\circ$ ($\text{kcal}\cdot\text{mol}^{-1}$)	ΔG° ($\text{kcal}\cdot\text{mol}^{-1}$)
18bp	0.97 ± 0.1	3.5 ± 0.8	5.1 ± 0.1	-16.6 ± 0.2	-11.5 ± 0.2
42bp	0.48 ± 0.1	0.4 ± 0.2	17.3 ± 0.5	-30.1 ± 1.4	-12.8 ± 0.9

Figure 4





HHS Public Access

Author manuscript

Nat Struct Mol Biol. Author manuscript; available in PMC 2018 November 13.

Published in final edited form as:

Nat Struct Mol Biol. 2018 October ; 25(10): 971–980. doi:10.1038/s41594-018-0133-6.

XLF and APLF bind to Ku80 on two remote sites to ensure DNA repair by non-homologous end-joining

Clement Nemoz^{#1}, Virginie Ropars^{#1}, Philippe Frit^{#2,3}, Amandine Gontier¹, Pascal Drevet¹, Jinchao Yu¹, Raphaël Guerois¹, Aurelien Pitois¹, Audrey Comte¹, Christine Delteil^{2,3}, Nadia Barboule^{2,3}, Pierre Legrand⁴, Sonia Baconnais⁵, Yandong Yin⁶, Satish Tadi⁷, Emeline Barbet-Massin⁸, Imre Berger⁹, Eric Le Cam⁵, Mauro Modesti⁷, Eli Rothenberg⁶, Patrick Calsou^{2,3,§}, and Jean-Baptiste Charbonnier^{1,§}

¹Institute for Integrative Biology of the Cell (I2BC), Institute Joliot, CEA, CNRS, Univ.Paris-Sud, Université Paris-Saclay, 91198, Gif-sur-Yvette cedex, France

²Institut de Pharmacologie et Biologie Structurale, IPBS, Université de Toulouse, CNRS, UPS, Toulouse, France

³Equipe Labellisée Ligue Contre le Cancer 2013 and 2018

⁴Synchrotron Soleil, L'Orme des Merisiers, Saint-Aubin - BP 48, 91192 Gif-sur-Yvette Cedex, France.

⁵Signalisations, Noyaux et Innovations en Cancérologie, UMR 8126, CNRS, Université Paris-Sud, Gustave Roussy, Université Paris-Saclay, 94805 Villejuif, France.

⁶New York University School of Medicine, Perlmutter Cancer Center, USA

⁷Cancer Research Center of Marseille, CNRS UMR7258, Inserm U1068, Institut Paoli-Calmettes, Aix-Marseille Université UM105, Marseille, France.

⁸Dynamic Biosensors GmbH, Martinsried, Germany

⁹BrisSynBio Centre, The School of Biochemistry, Faculty of Biomedical Sciences, University of Bristol, Bristol, United Kingdom.

These authors contributed equally to this work.

Abstract

§Corresponding authors: jb.charbonnier@cea.fr and calsou@ipbs.fr.

AUTHORS CONTRIBUTIONS

P.C. and J-B.C conceived this study. C.N, V.R, A.G, A.P, A.C, S.B expressed and purified recombinant proteins with help from P.D, E.L.C and I.B. C.N and V.R produced crystals and collected crystallographic data with help from P.L. C.N, V.R, P.L and J-B.C carried out the crystallographic analysis and interpreted the results. C.N, A.G, S.B, E.L.C and J-B.C designed, performed, and analyzed microcalorimetry and biophysical experiments. E.B-M, A.C and J-B.C designed, performed, and analyzed switchSENSE experiments. S.T and M.M designed, performed, and analyzed electromobility shift assays. P.F and P.C designed and constructed vectors and cell lines for life cell imaging and radiosensitivity. P.F, C.D, N.B and P.C designed, performed and analyzed western blot analyses of Ku variants, life cell imaging, DNA repair and radiosensitivity experiments. Y.Y and E.R designed, performed and analyzed the super resolution microscopy. J.Y and R.G performed bioinformatic analyses. All the authors discussed the data. C.N, P.C and J-B.C wrote the manuscript with input from V.R, P.F, P.D, R.G, M.M, E.B-M, Y.Y and E.R.

ACCESSION CODES. Crystal structures are deposited at the pdb with the following codes Ku-DNA-pAPLF (6ERF), Ku-DNA-pXLF (6ERH) and Ku-DNA-pXLFshort (6ERG)

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

The Ku70-Ku80 (Ku) heterodimer binds rapidly and tightly to ends of DNA double-strand breaks and recruits several factors of the Non-Homologous End Joining (NHEJ) pathway through molecular mechanisms that remain unclear. Here, we describe the crystal structures of the Ku-binding motifs (KBM) of the NHEJ proteins APLF (A-KBM) and XLF (X-KBM) bound to a Ku-DNA complex. The two KBMs motifs bind on remote sites of Ku80 α/β domain. The X-KBM occupies an internal pocket formed after an unprecedented large outward rotation of the Ku80 α/β domain. We reveal independent recruitment at laser-irradiated sites of the APLF-interacting protein XRCC4 and of XLF through the respective binding of A- and X-KBMs to Ku80. Finally, we show that mutations on the X-KBM and A KBM binding sites in Ku80 compromises efficiency and accuracy of end-joining and cellular radiosensitivity. A- and X-KBMs may represent two initial anchorage points necessary to build the NHEJ intricate interactions network.

Keywords

DSB repair; NHEJ; X-ray crystallography; induced fit; laser micro-irradiation; super resolution microscopy; switchSENSE; microcalorimetry

INTRODUCTION

In mammals the majority of DNA double-strand breaks (DSBs) is repaired by the non-homologous end-joining (NHEJ) pathway^{1,2}. The Ku70-Ku80 heterodimer (Ku) rapidly and tightly interacts as a preformed ring with DSBs extremities in a non-sequence specific manner³. Ku serves as a hub for the recruitment of several NHEJ factors^{4,5}. Among them, Ku recruits the XRCC4-LIG4-XLF ligation complex through interaction with XRCC4-LIG4^{6,7} and with XLF, the latter relying on a Ku-binding motif (KBM) localized at the XLF extreme C-terminus (thereafter named X-KBM)^{8,9} (Figure 1a and Supplementary Figure 1). The ligation complex organizes into filaments both *in vitro* and in cells¹⁰⁻¹⁴.

Interestingly, Ku interacts also with a number of accessory NHEJ factors. The APTX and PNKP-like factor (APLF) binds poly(ADP)-ribosylated proteins near DSBs sites^{15,16}, and has been reported to have nuclease activity^{16,17}. APLF tightly interacts with Ku through a KBM (thereafter named A-KBM) that is located in its central region^{9,18} (Figure 1a and Supplementary Figure 1). This interaction has been mapped to the periphery of the Ku80 von Willebrand A domain (vWA)⁵. Ku-APLF interaction was shown to facilitate recruitment of the APLF-partner XRCC4 at damaged sites⁹ and was proposed to stabilize the assembly of NHEJ factors around the DSB¹⁹. Notably, an A-KBM-like domain is present at the N-terminus of a recently identified inhibitor of the NHEJ pathway, CYREN(MRI), that also interacts with Ku80²⁰ (Figure 1a). Ku also associates with the Werner syndrome protein (WRN) that is involved in many aspects of DNA metabolism including NHEJ²¹. Two motifs in the C-terminus of WRN cooperate for interaction with Ku, one being A-KBM like, and the other resembling the X-KBM present on XLF (Figure 1a). In addition, we and others showed recently that PAXX (Paralog of XRCC4 and XLF) interacts with the Ku70 subunit through a third type motif that is located in its C-terminus^{22,23} (Figure 1a). Despite identification of KBMs in several NHEJ factors, their respective contribution to the efficiency of DSB repair is not fully understood. For example, the puzzling observations that

KBM deletion in XLF or APLF depletion in human cells lead to null or intermediate repair defect deserve further investigations^{9,24–27}.

The interactome of Ku thus defines a large ensemble of motifs and proteins that could potentially compete or act synergistically. However, despite important structural and biophysical studies on NHEJ complexes²⁸, the absence of high resolution structures of Ku-KBMs complexes limits our understanding of the roles and specificity of the different molecular interactions in the recruitment of NHEJ factors to DSBs. Mapping KBM-binding sites in structures is also needed to clarify potential competition of all the Ku interacting factors on limited positions (Supplementary figure 1).

Here, we establish the structural and functional basis of Ku unique modes of interaction with two factors within the NHEJ repair pathway. We describe the first crystal structures of Ku70-Ku80-DNA complex in interaction with the A-KBM (KBM of APLF) and with the X-KBM (KBM of XLF), revealing that the two KBMs occupy remote interaction sites on the Ku80 vWA domain. The functional context of our structural data was determined using cell based assays to visualize the recruitment of wild-type and mutant X-KBM motifs or XLF proteins as well as of the APLF-partner XRCC4 to DSBs sites induced by micro-irradiation in wild-type or mutant Ku backgrounds. Our data provide new mechanistic insights on the function of XLF and APLF in the NHEJ process.

RESULTS

The KBM of APLF tightly interacts with a highly conserved site of Ku80 vWA domain.

The APLF factor contains a conserved Ku binding motif (A-KBM, aa 179–192) (Figure 1a). We co-crystallized a 18-mer A-KBM peptide (aa 174–191) with a Ku form deleted for the C-terminal regions (Ku_{cc}), and an hairpin DNA (hDNA)³ (Supplementary Figure 2a). In the crystal structure at 3.0 Å resolution (Table 1), the A-KBM peptide is well defined and is positioned at the periphery of the vWA domain of Ku80 (aa T61⁸⁰-C156⁸⁰) (thereafter, superscript ⁸⁰ stands for Ku80) (Figure 1b, and Supplementary Figure 2b). It is located at more than 50 Å from the DNA binding site of Ku80. The A-KBM adopts an extended conformation in a pocket delineated by the helices α 4 and α 5 and the loop located between the β -strands B and C of Ku80 (Figure 1c). The hydrophobic part of the A-KBM, located in the C-terminal part of the motif, is composed by the amino acids I₁₈₅LPTWML₁₉₁ and is buried in a hydrophobic pocket formed by the Ku80 residues L68⁸⁰, I112⁸⁰, M115⁸⁰, I149⁸⁰ and I150⁸⁰ (Figure 1d). The N-terminal part of the A-KBM contained a patch of three consecutive basic residues and an acid residue (E₁₈₁RKR₁₈₄ in human sequence) (Figure 1e). It forms salt bridges and charged hydrogen bonds with respectively the side chains of D106⁸⁰, D109⁸⁰, Q73⁸⁰ and S145⁸⁰, and the main chain of K144⁸⁰ and S143⁸⁰.

Isothermal Titration Calorimetry (ITC) showed that the interaction between A-KBM and Ku had a K_d of 33 ± 10 nM (Table 2, and Supplementary Figure 2e). We measured a nanomolar K_d for the interaction between Ku and a 18bp DNA as already reported²⁹ and found a similar K_d for the interaction between the A-KBM and Ku alone or Ku bound to a 18bp DNA (Table 2). The interaction of the A-KBM with Ku_{cc} showed K_d and enthalpy values similar to full-length Ku (KuFL). Thus, the core heterodimeric region of Ku is likely

sufficient for the interaction with the A-KBM. Notably, these affinities are stronger than the ones previously reported by fluorescence polarization with a labelled A-KBM peptide (Kd of 580 nM)⁹.

Mapping the conservation rate of the residues at the surface of Ku80 shows that this pocket is the main conserved pocket together with the DNA binding pocket (Figure 1f). The residues L68⁸⁰, Y74⁸⁰ and I112⁸⁰ make tight interactions with the hydrophobic part of the A-KBM motif. Mutations at these positions were reported to greatly reduce or disrupt the interactions with APLF in yeast two-hybrid experiments or in EMSA⁹. We produced the Ku I112R mutant that, as expected had no residual interaction with the A-KBM motif by ITC (Table 2).

X-KBM creates an outward rotation of the vWA domain and a large groove in Ku80.

We then determined the crystal structures of the Ku70-Ku80-hDNA complex bound with peptides derived from the XLF X-KBM ((L281^X to S299^X) and (S287^X-S299^X) peptides) (Supplementary Figure 2c). The crystal structures at 2.8 and 2.9 Å resolution (Table 1) show an unprecedented large outward rotation of the Ku80 vWA (Figure 2a, b). This movement forms a large groove between the Ku80 vWA and the rest of the Ku heterodimer. We therefore termed this conformation the open state of Ku, in contrast to the closed state observed in the three other crystal structures reported (Ku alone (1JEQ), Ku-hDNA (1JEY) and Ku-hDNA-A-KBM (this study)). The conformational change of Ku80 vWA does not affect Ku interaction with the duplex DNA (Supplementary Figure 2d)

The X-KBM is located on the Ku80 vWA face of the newly created groove in a pocket delineated by 4 strands (β_A , β_D , β_E and β_E') and 3 helices (α_2 , α_7 , α_7'). The motif is positioned closer to the DNA (at 12Å) than the A-KBM, though not in direct contact. We can model eight residues of the X-KBM (₂₉₂KKPRGLFS₂₉₉) in the crystal obtained with the 19mer or 14mer X-KBM. These residues are the last eight residues of the XLF sequence. The residues ₂₉₆GLFS₂₉₉ of the X-KBM occupy a hydrophobic pocket delineated by Ku80 residues L12⁸⁰, V37⁸⁰, F41⁸⁰, F135⁸⁰, F164⁸⁰, Y225⁸⁰, and L234⁸⁰ (Figure 2c). In the closed state, these Ku residues are buried and mediate intramolecular contacts (Figure 2d). The X-KBM may thus stabilize a transient open conformation of Ku80 in equilibrium with the closed state under basal conditions. In addition, we observed that the Ku80 vWA opening comes along with important secondary structure changes in the linker region (R232⁸⁰-E241⁸⁰) that separates the vWA and the rest of the Ku heterodimer (Figure 2c, d).

ITC measurements showed a moderate affinity of X-KBM for Ku, with a Kd of $4.4 \pm 0.2 \mu\text{M}$ (about 200-fold weaker than the affinity of the A-KBM) (Table 2, and Supplementary Figure 2f). We observed similar affinities and thermodynamic parameters for the X-KBM with a Ku-DNA complex or with Ku_{cc} (Table 2). Notably, this micromolar interaction was not detected in a previous study using fluorescence polarization with labelled peptides⁹. We also measured similar affinities for the interactions between XLF homodimer and Ku alone or Ku bound to a 18bp DNA (Table 2). These data show that in absence of DNA or with a short DNA protected by Ku ring, XLF and its X-KBM interact similarly with Ku.

EMSA analyses confirmed an interaction between XLF and Ku complexed with a 50bp DNA in the μM range (Supplementary Figure 3a–b). Competitions experiments showed that the X-KBM peptide (pXLF) competes in the μM range with the XLF protein and that the A-KBM peptide (pAPLF) does not displace XLF, supporting remote sites of interactions (Figure 2e and Supplementary Figure 3c, d). Also, the C-terminus of PAXX (pPAXX) does not compete with XLF binding, in agreement with previous studies that report an interaction between PAXX C-terminus and Ku70 subunit^{22,23} (Supplementary Figure 3e).

The outward rotation of the vWA domain of Ku80 is mediated by Glu133.

To evaluate if the outward rotation of Ku80 observed in the crystal structure with the X-KBM peptide was present in solution, we performed SAXS analyses. The SAXS data with the A-KBM peptide were comparable with those of the Ku-hDNA sample without any peptide (respective R_g of 53.1Å and 53.7) (Figure 2f). We measured an increase of the R_g to 59.0Å in presence of the X-KBM, corresponding to the opening of Ku80 observed in the crystal (Figure 2f). Comparison of the D_{max} values provided further corroboration for the opening of the Ku molecule with an increase of 30 Å.

The acid residue E133⁸⁰ buried in Ku80 is well positioned to act as a spring facilitating the Ku80 opening (Figure 2c, d). The glutamate E133⁸⁰ is buried in the closed conformation of the Ku80 vWA and its pKa value is estimated by the PDB2PQR-2.0 server³⁰ at a value of 9.1, far from the normal pKa of 4.5 for a glutamate in solvent. The outward rotation of Ku80 vWA should be energetically facilitated by the solvation of this Glu following the displacement of Ku80 residues V236⁸⁰, F237⁸⁰ and I240⁸⁰ away from E133⁸⁰ carboxylate function (Figure 2d). Multiple sequence alignments show that E133⁸⁰ position and the residues surrounding are well conserved in mammalian and saurian and that Ku70 has no equivalent buried acidic residue at this position (Supplementary Figure 4a–b).

Molecular bases of the specificities of A-KBM and X-BKM motifs binding to Ku80.

The A-KBM and X-KBM present sequence similarities with a basic patch in their N-terminus followed by a hydrophobic patch⁹ (Supplementary Figure 1). Comparison of the crystal structures of Ku bound to these two motifs suggests that the high affinity of the A-KBM relies on the tryptophan W189^A in place of the Leu297^X in X-KBM. Thus, we used ITC to measure the interaction of the X-KBM motifs with the mutation L297W (LW) or with a non-conservative L297E mutation (LE). The (LW) peptide has a K_d of $0.12 \pm 0.03 \mu\text{M}$, an interaction 40-fold tighter than wild-type X-KBM (Table 2 and Supplementary figure 2g). The LE mutant presents no detectable interaction with Ku (Table 2). Competition experiments with Ku saturated with the A-KBM showed that the X-KBM (L297W) no longer interacts with Ku, suggesting that the sole L297W mutation is able to redirect the X-KBM towards the A-KBM binding site on Ku80 (Table 2).

Then live cell imaging was used to monitor the recruitment of CFP-fused A-KBM and X-KBM fragments to DSBs sites induced with laser micro-irradiation. Under conditions of similar damage yield (Supplementary Figure 5e), the A-KBM motif transfected in U2OS cells was strongly recruited but the W189G mutation impaired both its nuclear localization and recruitment to laser sites (Figure 3a–b), as reported^{9,18}. X-KBM live recruitment was

impaired by L297E mutation but not L297W mutation (Figure 3c), corroborating ITC data. We then used U2OS cells expressing an inducible shRNA against Ku80³¹ (Supplementary Figure 5a) that were complemented with wild-type or I112R mutant Ku80 (Supplementary Figure 5c–d). The I112R mutation impaired A-KBM recruitment, as expected, but not that of X-KBM (Figure 3d, e). Notably, I112R Ku80 mutant specifically lowered the recruitment of L297W X-KBM (compare Figures 3c and 3f). Conversely, APLF knock-down boosted the recruitment of the LW mutant peptide above that of wild-type X-KBM (Figure 3g and Supplementary Figure 5b for control of shAPLF efficiency). Together, these data in cells support that the LW mutation redirects the X-KBM fragment to the APLF-binding site in Ku80 and point out the W189 residue as a key determinant for APLF specific interaction with Ku80.

X-KBM mutations impair XLF recruitment and XRCC4-XLF filament stability.

We then investigated the properties of the interaction between full-length XLF protein and Ku. We first used the SwitchSENSE approach³² in which oligonucleotide nanolevers labelled with a fluorescence probe are bound to a gold surface (Supplementary Figure 3f). Ku bound onto 48bp DNA nanolevers with a K_d in the nM range as already reported²⁹ and a long dissociation time (Supplementary Figure 3g). Wild-type XLF onto the Ku-DNA complex showed a rapid k_{on} ($4.7 \pm 1.7 \cdot 10^5 \text{ M}^{-1}\text{s}^{-1}$) followed by a rapid dissociation (k_{off} = $0.09 \pm 0.004\text{s}^{-1}$) and a corresponding K_d of $0.19 \pm 0.07 \mu\text{M}$ (Figure 4a). This affinity is about 10 fold stronger than the one measured by ITC with a smaller DNA and may reflect additional interactions of XLF with DNA emerging from Ku ring as observed with PAXX²². LW and LE mutants showed a 2.3- and 5.1-fold weaker affinity than WT protein, respectively (K_d of $0.45 \pm 0.26 \mu\text{M}$ for LW and K_d of $0.98 \pm 0.15 \mu\text{M}$ for LE) (Supplementary Figure 3h). As compared with ITC and recruitment data with X-KBM LW mutant peptide, this suggests that the LW mutation cannot redirect the XLF protein to the APLF binding site in Ku80.

We then measured the recruitment of the CFP-tagged full-length XLF protein (CFP-XLF) expressed in human XLF-defective BuS cells (SV40T-transformed, telomerase immortalized fibroblasts derived from the XLF-deficient P2 patient - homozygous C622T nonsense mutant (R178X))³³ (Figure 4b). Wild-type XLF showed a rapid mobilization to irradiated nuclear sites, as reported³⁴. Both LE and LW mutations impaired to various extents XLF protein recruitment (Figure 4b), again indicating that LW mutation cannot redirect the XLF protein to the APLF binding site in Ku80. In addition, we observed an important reduction of the recruitment of G296W, S299E or F298G XLF mutants in the extreme C-terminus (Supplementary Figure 5f–h). We also questioned the contribution of APLF-XRCC4 complex to XLF recruitment by expressing L115D XLF mutant unable to interact with XRCC4^{35,36}. We measured an efficient recruitment of L115D XLF to laser-induced DSBs that was insensitive to APLF knock-down but impaired with the XLF L115D/L297E XLF double mutant (Supplementary Figure 5i). Together, these data support a major role for Ku80 interaction with L297 and extreme C-terminal residues for XLF recruitment at DSBs in cells.

Multi-color super-resolution localization microscopy (STORM) allows characterizing formation of XRCC4-XLF filaments close to Ku foci and DSBs ends¹⁴. On DSBs induction with the radiomimetic drug neocarzinostatin, extended XLF filaments close to Ku80 foci were observed in XLF complemented BuS cells, whereas cells harboring (LE) and (LW) mutants showed slightly smaller and more punctuated XLF structures (Figure 4c and Figure 4d–e for quantification). In contrast to these data, evaluation of the effect of mutations in the X-KBM on cell radiosensitivity showed that (LW) or (LE) XLF mutants were associated with respectively no or minor radiosensitization, compared to high radiosensitivity of BuS cells and full restoration of radioresistance on expression of wild-type XLF (Figure 4f).

APLF and XLF binding to Ku80 promote DSB repair and cell survival to IR.

Finally, we questioned the discrepancy between subnormal cell survival and defective XLF recruitment and filaments formation associated with mutations in X-KBM. The outward rotation in Ku80 upon X-KBM binding more likely relies on E133⁸⁰, the equivalent position of which in Ku70 is a methionine (M167) (Supplementary Figure 4b). Therefore, we designed E133M mutant in Ku80 and Q162E mutation that may alter the charge environment of E133 (Supplementary Figure 5c). E133M or Q162E mutation in Ku80 negatively impacted X-KBM recruitment but not that of control A-KBM (Supplementary Figures 6a and 6b). This result further supports the independent binding sites on Ku80 of the A-KBM and X-KBM motifs and validates E133 and Q162 positions in Ku80 as key residues for X-KBM binding.

Then, mutations in the Ku80 binding sites for X- and A-KBM were combined in the I112R/E133M Ku80 double mutant (Supplementary Figure 5d). This combination of mutations clearly impaired the recruitment of both CFP-(X- and A-KBM) peptides (Figures 5a and 5b) while the recruitment of the mutant CFP-Ku fusions was preserved (Supplementary Figure 6c–d). We also measured the recruitment of the APLF partner XRCC4 and of XLF, expressed as CFP-tagged full-length proteins (Figure 5c–d). Strikingly, only the double E133M/I112R mutation strongly but not completely impaired the recruitment of both proteins (Figure 5c–d). This result supports the independent binding of APLF-XRCC4 and XLF proteins on Ku80.

Ku80 mutants expressed in U2OS cells had no detectable impact on XLF filaments (Supplementary Figure 6e–f) but lowered end-joining efficiency on a linear plasmid transfected in U2OS cells (Figure 5e). To assess repair accuracy, we used an assay in which GFP expression from a cassette integrated in cells relied on the loss of a DNA fragment between two cut sites (modified from³⁷) (Figure 5f), reported to be favoured upon NHEJ inhibition^{38,39}. Strikingly, we observed an increase in fragment loss with the three Ku80 mutants (Figure 5f), indicating that loose APLF-XRCC4 or XLF interactions with Ku80 promote genomic instability. Finally, the combination of E113M and I112R Ku80 mutations had a strong impact on cell radiosensitivity with an additive effect of both mutations (Figure 5g). These results indicate that APLF-XRCC4 and XLF interactions with Ku80 cooperate to promote cell survival to IR.

DISCUSSION

Here, we show that each of the A- and X-KBM binds Ku independently of other APLF or XLF protein domains. This is in contrast with the absence of binding reported for the isolated X-KBM in recent studies with fluorescent polarization^{9,40} that may be due to lower sensitivity, steric hindrance by the fluorescence probe or differences in the Ku constructs used. Therefore, the present results rule out the conclusion that Ku-XLF interaction necessarily needs XLF binding to DNA (even though DNA could stabilise the interaction, as deduced from our switchSENSE data) or to DNA-PKcs⁴⁰.

Despite the A-KBM and X-KBM motifs have important sequence similarities (Supplementary figure 1), they target different regions of Ku80 with different modes of actions, a more rigid one for the A-KBM and an induced fit one for the X-KBM. In contrast to the L297W X-KBM peptide, the (LW) full-length XLF mutant protein cannot be redirected to the APLF-binding site on Ku80. This may be explained by steric constraints either intrinsic to the protein structure or due to interactions with other partners in the NHEJ complex. It was reported that a X-KBM in WRN protein functions cooperatively with an A-KBM located upstream (1403–1412) in binding Ku complexes⁴⁰ (Figure 1a). Our structural data allow proposing the first model of Ku interacting with the tandem sequence of the WRN A-KBM and X-KBM, including the central linker (Supplementary Figure 4c). This model now awaits validation by the crystal structure of the tandem motif of WRN on Ku.

XRCC4 and XLF organize into filaments both *in vitro* and in cells^{10–14} and can also assemble as sliding sleeves-like structures on broken DNA *in vitro* even without Ku⁴¹. In this study, cells containing single or double mutations on the A-KBM and/or X-KBM binding sites of Ku80 show intact filaments while mutations on the X-KBM of XLF induce a slight reduction in the filaments size. These data suggest that interactions additional to Ku80-XLF may participate in filament formation, like the XLF C-terminal DNA-binding domain that spans the X-KBM⁴².

How is NHEJ ligation complex assembled at broken DNA? APLF FHA domain interacts with XRCC4^{17,26} and XLF establishes head to head contact with XRCC4^{10–13}. Here, we show that the two remote APLF- and XLF-KBM binding sites in Ku80 promote the independent recruitment of XRCC4 and XLF at broken DNA and that disruption of XRCC4-XLF interaction (through L115D mutation) does not compromise XLF recruitment. Destabilisation of APLF or XLF interactions with Ku80 impairs repair efficiency and cell survival to DSBs and also favours genome instability associated with distal end-joining. These effects are most likely explained by loose assembly of the NHEJ apparatus at break ends. Therefore, we propose a model in which APLF and XLF KBMs represent two initial anchorage points for the rapid and independent recruitment of APLF-XRCC4 and XLF on Ku (Figure 5h). After initial recruitment, interactions additional to Ku80-XLF contacts may stabilize XLF at DSB sites (Figure 5h): our recruitment data show a stable interaction of XLF at DSB sites in cells that differs from the rapid dissociation from Ku-DNA observed *in vitro* by switchSENSE; also, although XRCC4 is dispensable for XLF initial recruitment, it has been shown to stabilize XLF at damaged sites³⁴. Thus, additional contacts stabilizing XLF may include XLF-DNA^{34,42,43} and/or XLF-XRCC4 interactions, including filaments

formation^{10–14} and/or interaction with the LIG4 BRCT1 domain⁴⁴. In addition, after XRCC4 recruitment through APLF binding to Ku, XRCC4 may also stabilize in turn APLF at damaged sites since FHA mutants of APLF that do not interact with XRCC4 show a reduced retention after laser micro-irradiation¹⁸. Moreover, links with DNA-PK may further properly stabilize and/or position the NHEJ ligation complex at DSBs (Figure 5h): Ku directly interacts with the XRCC4-LIG4 complex⁷ through either XRCC4⁴⁵ or LIG4⁶ and DNA-PKs also directly contacts XRCC4^{46–48}. Intimate links between the DNA-PK and ligation complexes are illustrated by the requirement of an intact XLF-XRCC4-LIG4 complex to ensure optimal DNA ends synapsis^{49,50}.

The intricate network linking end-recognition and ligation NHEJ complexes may allow compensation of partially defective individual components. Indeed, we found that separate or even combined Ku80 mutations do not completely abolish XRCC4 and XLF recruitment and do not radiosensitize cells as much as XLF complete defect that abolishes all XLF functions in NHEJ⁵¹. In that view, only a mild, if any, repair defect in human cells has been associated with X-KBM deletion²⁷ or APLF depletion^{9,24–26}: in case of individual absence of XLF C-terminal tail or of APLF, the other intact partner would still bind to Ku80 and be able, although with a slower kinetics, to recruit the other components to achieve ligation.

Finally, our present study adds a new aspect to the DNA-PK-ligation complex interaction network that is the swing of the Ku80 vWA domain upon XLF binding (Figure 5h). From our structural and mutagenesis approaches, we propose that the outward swing of the vWA domain of Ku80 is mainly dependent on the conserved acidic residue E133⁸⁰. Although this swing does not impact the A-KBM binding site nor the DNA binding domain, it exposes a large groove between the vWA and the ring domain of Ku80 that may non-exclusively reinforce XRCC4-LIG4 and/or XLF interaction with Ku or attract yet unknown components. Interestingly, XLF deficiency impacts on the activity of the Ku partner terminal deoxynucleotidyl transferase during V(D)J recombination⁵². Mutagenesis studies on the conserved positions that delineated the unmasked surface of this groove will help to define precisely the role of this swing in the NHEJ reaction.

In conclusion, the present data further substantiate the emerging model that several NHEJ factors bearing a limited repertoire of KBMs recognize a limited number of KBM-binding sites on Ku¹⁶. The complex regulation of Ku sites occupancy by NHEJ factors during the repair process deserves further investigations.

ONLINE METHODS

DNA preparation

The 500 bp linear dsDNA molecules were amplified from the 3516–4016 region of pBR322 plasmid with biotinylated primers (5'-bGGATCTCAACAGCGGTA-3' and 5' bCTTTATCCGCCTCCATCC-3'). DNA fragments were purified on a MiniQ anion exchange column with a chromatography SMART system (GE Healthcare), ethanol precipitated and resuspended in a 10 mM Tris-HCl, pH 7.5, 1 mM EDTA buffer.

Oligonucleotides for ITC and EMSA

CN1 34bp : CGCGCCAGCTTTCCAGCTAATAAACTAAAAAC

CN2 21bp : GTTTTTAGTTTATTGGGCGCG

CN3 18bp up : GTTATCCGAGCGTGAGAC

CN4 18bp down : GTCTCACGCTCGGATAAC

NLB48 : TAG TCG TAA GCT GAT ATG GCT GAT TAG TCG GAA GCA TCG AAC
GCT GAT

MM 50bp up : FAM-

TAAATGCCAATGCTGCTGATACGTACTCGGACTGATTTCGGAAGTAAACG

MM 50bp do :

CGTTACAGTTCCGAATCAGTCCGAGTACGTATCAGCAGCATTGGCATTTA

Peptides and Proteins: The synthetic peptides containing the KBM and XLM motifs were purchased from Genecust at 95% purity, and the concentrations of the stock peptide solutions were determined by amino acid composition. The oligonucleotides used for ITC and crystallization were synthesized by Sigma-Aldrich and Eurogentec.

The full length Ku70(1–609)/Ku80(1–732) heterodimer and a truncated version of the heterodimer deleted Ku70(1–544)/Ku80(1–551) were cloned in the Multibac vectors with a 10-His tag and a TEV site on the Ku80 N-terminus⁵³. Each plasmid was integrated in a Yellow Green Protein (YFP) containing bacmid by transformation in EMBACY *E.coli* stain (kind gift from Imre Berger, Bristol University). The resulting recombinant bacmids were used to transfect Sf21 insect cells giving the V0 virus generation. After amplification, stocks of viruses were titrated by the dilution limit method using YFP as marker for infected cells and Mac Grady table. Production was initiated in Sf21 cells culture by infection with baculovirus at MOI of 5×10^{-3} . Insect cells were collected 5–6 days after the infection (3–4 days after the proliferation arrest). Cells were sonicated and the supernatant was incubated with Benzonase (300 units for 30 min at 4°C). The Ku heterodimer was purified on a NiNTA-Agarose affinity column (Protino, Macherey Nagel) with a 1M NaCl wash step to remove DNA excess. The eluted Ku was then bound onto an anion exchange column (Resource Q, GE Healthcare) equilibrate with buffer Q (20 mM Tris pH 8.0, 50 mM NaCl, 50 mM KCl, 10 mM β -mercaptoethanol). Final yield of the Ku heterodimer was typically 35 mg of purified heterodimer by liter of culture. The full length 10-His tagged XLF protein was produced in insect cells with a similar protocols with a yield of 50mg of purified protein by L of culture. A truncated version of XLF(1–224) was also produced in *E. coli*

Crystallization of the Ku70/Ku80/hDNA complexes with A-KBM and X-KBM peptides

The DNA used in this study is the hairpin DNA previously used by Walker et al³. It was obtained using HPLC-purified oligonucleotide of 34 and 21 nucleotides (see oligonucleotides above). The oligonucleotides were annealed and added in 1.1-fold molar excess to Ku heterodimer. The peptides containing the A-KBM motif (18mer, 174–192), the

long X-KBM motif (19mer, 281–299) or the short X-KBM motif (13mer 287–299) were added respectively with a 1.1, 2 and 2 fold excess. Crystallization screenings on the Ku_{CC}-hDNA-peptides were performed on the HTX platform (EMBL, Grenoble) with an automatic visualization at 4°C. The crystals of Ku_{CC}-hDNA-pAPLF were reproduced and optimized in the laboratory at 20 °C using the sitting drop method by mixing 1.5µL of the 20 mg/mL Ku-DNA-peptides complexes with 1.5 µL of the solution containing 13% polyethylene glycol (PEG) 3350, 150 mM NaNO₃, and 100 mM Bis-Tris-Propan (pH 6.5). The crystals (100×150×1000µm) grew in 5–6 hours and were frozen in a solution of the mother liquor with 20% glycerol. The Ku_{CC}-hDNA-pXLF and Ku_{CC}-DNA-pXLFs complexes were crystallized at 20 °C by mixing 1.5µL of the 7 mg/mL complex solution with 1.5 µL of a solution containing 18% polyethylene glycol (PEG) 3350, 150 mM Na₂SO₄, and 100 mM Bis-Tris-Propane (pH 8,5). The crystals (100×200×50µm) grew in 5–6 days and were frozen with 20% glycerol.

Determination of the crystal structures

Diffraction data were collected at the Proxima 1 and Proxima 2 beamlines at the *synchrotron* SOLEIL. The datasets were indexed and integrated using the XDS package⁵⁴, the XDSME package (XDS Made Easier, <https://github.com/legrandp/xdsme>) and the CCP4 suite⁵⁵. The crystals present a highly anisotropic diffraction (between 2.85Å and 4.25Å resolution according to the axes). The anisotropy of Kucc-hDNA-pAPLF crystals was treated with the STARANISO program (<http://staraniso.globalphasing.org/>). The software performs an anisotropic cut-off of merged intensity data, a Bayesian estimation of the structure amplitudes, and applies an anisotropic correction to the data. The structure of the Ku70/80-hDNA-pAPLF was determined by molecular replacement with the program MOLREP using the structure of Ku70/80 (pdb 1JEY) without the DNA coordinates³. Four molecules of Ku were consecutively positioned. Electron density for the hDNA was clearly visible in the position previously reported by Walker. Refinement was performed using BUSTER⁵⁶ and PHENIX⁵⁷. The models were built with Coot⁵⁸. After DNA building, an electron density was visible on the Ku80 vWA near the Ku80 amino acids identified by Grundy et al by mutagenesis. The final statistics are presented below. In the final model, the following regions of Ku70/Ku80 are not visible: Ku70 1–33, 535–544 and Ku80 1–5, 170–181, 190–191, 543–551. The quality of the model was assessed using Molprobity⁵⁹.

The crystal structures of the Ku70/80-hDNA complexed with the long and short pXLF were solved by molecular replacement. Firstly, the coordinates of Ku70/80 and the hairpin DNA present in pdb 1JEY were used as model. We positioned two molecules in the asymmetric unit with clear electron density except on the vWA region of Ku80 (region aa 6 to 242). We then performed the molecular replacement with the same coordinates deleted of the Ku80 vWA region. The electron density for the secondary structure elements of the Ku80 vWA was clearly visible though at a different position than in 1JEY. A second molecular replacement step was performed to position the vWA domain of the two Ku80 molecules in the asymmetric unit. The structure of the Ku70/80-DNA-short pXLF was solved in a similar manner. An electron density was visible in an internal position of Ku80 vWA that could be attributed to pXLF. The final statistics are presented in Table1. In the final model with the long pXLF, the following regions of Ku70/Ku80 are not visible (Ku70 1–34, 535–554 and

Ku80 543–544). In the final model with the short pXLF, the following regions of Ku70/Ku80 are not visible (Ku70 1–34; 535–554 and Ku80 171–194, 300–301, 542–544).

An additional electron density was observed in the long and short pXLF complex structures located close to the extreme N-terminus of Ku80 and to the hairpin DNA. This electron density was successfully modeled with the first missing residues of Ku80, the TEV site sequence preceding (ENLYFQG) and seven histidines from the 10-His tag. To evaluate the influence of the tag present on the N-terminus of Ku80 on the Ku-XLF interaction, we prepared a digested form of Ku with the TEV protease (Ku^{TEV}) and measured by ITC its interaction with the X-KBM. The ITC shows that thermodynamic parameters of the Ku^{TEV} are similar to the Ku and that the tag does not significantly influence the binding of the XLF motif.

Small-angle Xray scattering: Several data of Ku_{cc}-hDNA complex with a protein-DNA ratio (1:1.2) were collected alone or in presence of pXLF or pAPLF peptides. Data were collected on the SWING beamline (SOLEIL synchrotron) at a 1.8m sample-detector distance. The complexes were prepared at 1.0, 3.0 and 5mg/mL and spun for 10 minutes at 13000 rpm prior to SAXS analysis to eliminate aggregates. Volumes of 40μL of each sample and buffer (20 mM Tris pH 8.0, 150mM NaCl, 5mM β-mercaptoethanol) were injected into the SAXS capillary cell and collected continuously, with a frame duration of 0.5 s and a dead time between frames of 0.5 s. Data reduction to absolute units, frame averaging and subtraction were done using FOXTROT⁶⁰, a dedicated home-made application. All subsequent data processing and analysis steps were carried out with PRIMUS and other programs of the ATSAS suite⁶¹. The program GNOM⁶² was used to compute the pair-distance distribution functions, $P(r)$.

Isothermal titration Calorimetry (ITC)

Interactions between Ku70/80 wild-type and Ku70/80cc and the different peptides containing the KBM and XLM motifs were determined by isothermal titration calorimetry (ITC) using a VP-ITC calorimeter (Malvern). Prior to measurements, all solutions were degassed under vacuum. The reaction cell of the ITC (volume 1.8 mL) was loaded with Ku heterodimers alone or complexed with DNA or peptides for competition experiments. Proteins were extensively dialyzed against buffer I (20 mM Tris, pH 8.0, 150 mM NaCl, and 5 mM β-mercaptoethanol). Peptides and DNA were prepared at high concentrations. The syringe (290 μL) was filled with the different peptides at concentration between 20 μM to 200 μM. The Ku heterodimer present in the cell was titrated by automatic injections of 6–10μL of the different peptides. Enthalpy ΔH (in kcal.mol⁻¹), stoichiometry of the reaction N , and association constant K_a (in M⁻¹) were obtained by nonlinear least-squares fitting of the experimental data using the single set of independent binding sites model of the Origin software provided with the instrument. The free energy of binding (ΔG) and the entropy (ΔS) were determined using the classical thermodynamic formula, $\Delta G = -RT \ln(K_a)$ and $\Delta G = \Delta H - T \Delta S$. All binding experiments were performed in duplicate or triplicate at 25°C. Control experiments were performed with peptides injected into the buffer to evaluate the heat of the dilution.

EMSA: Binding reactions (10 μ L) were performed by incubating the annealed oligonucleotides (oligonucleotides used in this study are listed in Table Sxx and indicated in the figure legends) at a final concentration of 25 nM, with the indicated final concentrations of proteins in 75 mM KCl, 10 mM Tris (pH 7.5), 0.5 mM EDTA, 0.5 mM DTT, 0.5 mg/mL acetylated-BSA, and 5% glycerol. Reactions were incubated at room temperature for 1 hr and fractionated by 6% PAGE (29%/1% [w/v] Acrylamide:Bis-acrylamide) in 0.53 standard Tris-borate-EDTA (TBE) buffer at 80 V for 45 min to 1 hr. After electrophoresis, DNA was visualized using a ChemiDoc MP imaging system (Bio-Rad), either by direct detection of the fluorescently labeled DNA (FAM) or after staining with 0.2 mg/mL EtBr. Data were processed and quantified with the Image Lab software version 5.2.1 (Bio-Rad).

switchSENSE measurements: All switchSENSE measurements were carried out on a DRX 2400 instrument, using a multi-purpose 48bp chip (both Dynamic Biosensors GmbH; Planegg, Germany). The sample and running buffer was Tris 10mM pH 7.4, 140mM NaCl, 0.05% Tween20, 50 μ M EDTA, 50 μ M EGTA. In all kinetics experiments, complementary DNA to the strand immobilized on the chip was first hybridized on the measurement electrodes. On top of the 48bp complementary to the chip, this DNA carried a 32bp-overhang hybridized with its 32bp complementary strand. Second, the Ku70/80 protein was associated at 100nM for 3min (reaching saturation) to the immobilized dsDNA. Finally, the association and dissociation of the XLF protein was measured using triplicates of each concentration, at a flow rate of 2mL/min. For each mutant, the surface was not regenerated in between concentrations – Ku does not dissociate in the timescale of the XLF measurement, and XLF fully dissociates for each concentration. The electrodes were only regenerated between the measurements of different mutants of XLF. Kinetics values were determined using Origin software.

Cell lines and cell culture

U2OS human osteosarcoma cells and immortalized BuS cells (derived from an XLF-deficient patient, gift from Jean-Pierre de Villartay, Institut Imagine, Paris, France) were grown in DMEM and RPMI, respectively. Media were supplemented with 10% fetal calf serum (Eurobio), 125 U/ml penicillin and 125 μ g/ml streptomycin. Cells were maintained at 37°C in a 5% CO₂ humidified incubator. Cell lines were tested negative for mycoplasma by PCR. All culture media and antibiotics were from Invitrogen. When necessary (conditional expression of shRNA against Ku80), doxycyclin (Sigma-Aldrich) was added to the medium at a 4 μ g/ml final concentration.

Expression vectors

See Supplementary Data Set 1

Cell transfection and transduction

Production of lentiviral particles in HEK-293T cells and transduction of U2OS and BuS cells were performed as previously described³¹. Transduced cells were used as populations without clonal selection, except when indicated otherwise.

Plasmid recircularization assay

U2OS cell populations expressing wild-type or mutated Ku80 protein were seeded to near confluence in 6-well plates and incubated overnight at 37°C. Cells were then transfected with 1 µg BamHI-linearized pEGFP-N1 plasmid (Clontech) and 1 µg pmCherry-C1-3NLS circular plasmid as a transfection control (gift from Dyche Mullins (Addgene #58476),⁶³). Cells were split 24 h later, incubated at 37°C for two additional days and analyzed by flow cytometry on a Fortessa X-20 (BD Biosciences). For each cell population, the integrated GFP fluorescence signal was normalized to mCherry signal. End-joining activity was set to 100% for Ku80-WT expressing cells. Results were plotted as the mean values of four independent experiments ±s.d.

Distal End-Joining assay

U2OS-EJ5 cells (a kind gift from Jeremy Stark, City of Hope, Duarte, USA) containing one integrated copy of a GFP reporter cassette which allows to measure rejoining of two tandem I-SceI cut sites separated by a ~1.8-kb insert, were modified as follows. The cells were first transduced with lentiviruses produced from pLV-tTR-KRAB and pLV3-Tet-RFP ISceI-GRLBD to allow inducible expression and nuclear translocation of the I-SceI yeast meganuclease. The resulting cells were then transduced with lentiviruses prepared from pLVTHM2-shKu80. A positive clone enabling doxycyclin-dependent conditional knockdown of Ku80 expression and subsequent cell death was isolated. These cells were further transduced with lentiviruses produced from pLV3-HA-Ku80-shR-(WT or mutants) plasmids to replace, in the presence of doxycyclin, endogenous Ku80 expression by expression of the various Ku80 constructs described in the study. To perform Distal End-Joining assay, the different U2OS-EJ5 modified cell populations were seeded onto 6-well plates and incubated at 37°C for 24 h. Dexamethasone (Sigma-Aldrich) was then added to a final concentration of 250 nM. Cells were washed 24 h later, further incubated at 37°C for two to three days and analyzed by flow cytometry on a Fortessa X-20 analyzer (BD Biosciences). The fraction of GFP-positive cells was measured and normalized to 100% for Ku80-WT expressing cells. Results were plotted as the mean values of seven independent experiments ±s.d.

Ionizing irradiation and cell survival analysis

Three to six thousand cells per well were seeded in 6-well plates. Plated cells were exposed 24 h later to various doses of X-ray using a Faxitron RX-650 device (130 kV, 5 mA, dose rate 0.5 Gy.min⁻¹). Six to seven days later, cells were washed with PBS, stained 10 min with crystal violet (0.1% aqueous solution). Stained cells were extensively washed with water and plates were air dried. Staining was dissolved with 10% acetic acid solution and absorption was measured at 570 nm (Ultrospec-3000 spectrophotometer, Pharmacia Biotech). Results were plotted as mean values of 5–10 independent experiments ±s.d. using Microsoft Excel software.

Live-cell microscopy and micro-irradiation

U2OS or BuS cells were seeded in 35-mm glass-bottom culture dishes (MatTek) two days prior laser irradiation. Experiments were carried out with a Zeiss LSM-710 confocal laser

scanning microscope equipped with a coherent chameleon Vision-II tunable laser (690–1080 nm), a 40X/1.3 oil immersion objective and a heated environmental chamber set at 37°C in 5% CO₂ atmosphere. ECFP was excited using biphotonic laser at 800 nm (1.5% of maximum power). Confocal image series were recorded with a frame size of 512×512 pixels. Nuclei micro-irradiation was carried out at 800 nm at 20% of maximum power (mean max power was 3070 mW) in rectangle of 15 μm² area during 50 μs pixel dwell time. Before and after micro-irradiation, confocal image series of one mid z-section were recorded at 1.94 s time interval (typically 9 pre-irradiation and 40–45 post-irradiation frames). For evaluation of the recruitment kinetics, fluorescence intensities of the irradiated region were corrected for total nuclear loss of fluorescence over the time course and normalized to the pre-irradiation value. Data from micro-irradiation of individual cells obtained in several independent experiments performed on different days were averaged, analyzed and displayed using PRISM software. *p* values at last time point were calculated using unpaired Student's t-test.

Protein extraction and immunoblotting

Sub-confluent cells from 60-mm culture dishes were harvested and washed with PBS. Pellets were resuspended in 100 μl of lysis buffer (50 mM HEPES.KOH pH 7.5, 450 mM NaCl, 1% Tritin-X100, 1 mM EDTA, 1 mM DTT, protease-phosphatase Halt Inhibitor cocktail (Pierce, Thermo Scientific)) and processed by four freeze/thaw cycles. After centrifugation at 14000 g for 10 min, protein concentration was measured in the supernatant with the Bradford assay (Bio-Rad). Proteins from 50 μg of cell extracts were separated in 4–15% Mini-Protean TGX precast polyacrylamide gels (Bio-Rad) and transferred to PVDF membrane (Millipore). Immunoblotting analysis was performed with the following antibodies: anti-Ku70 monoclonal antibody (clone N3H10 from NeoMarkers), anti-Ku80 monoclonal antibody (clone 111 from Thermo Fisher Scientific), anti-DNA-PKcs monoclonal antibody (clone 18.2 from Abcam), anti-XLF polyclonal antibody (Bethyl Laboratories), anti-APLF polyclonal antibody (SK3595, kind gift from K.W. Caldecott, University of Sussex, Brighton, UK, see²⁵), anti-β-Actin monoclonal antibody (clone AC-15 from Ambion).

Cell culture and preparation for super-resolution imaging

BuS cells were grown in RPMI medium with 10% FBS and 100 U/mL Penicillin-Streptomycin. For super-resolution imaging, cells were cultured on glass coverslips for 24 hours, followed by serum starvation for 48 hours. The cells were then released into full medium for 4 hours so that most of the cells were in G1 phase.

The synchronized cells were then washed twice with PBS, and permeabilized with 0.5% Triton X-100 in CSK buffer (10 mM Hepes, 200 mM Sucrose, 100 mM NaCl, and 3 mM MgCl₂, pH=7.4) for 10 minutes⁶⁴. Cells were then fixed with paraformaldehyde (4%) for 20 mins, and blocked in block solution (2% glycine, 2% BSA, 0.2% gelatin, and 50 mM NH₄Cl in PBS) overnight at 4°C.

Fixed cells were then immunostained with validated monoclonal antibodies: XLF was stained with XLF-antibody (3D6, NBP2–03275, NOVUS) for 1 hour at room temperature,

followed by Alexa Fluor 647 conjugated goat-anti-mouse 2nd antibody staining (ab 150115, abcam) for 30 minutes at room temperature. Cells were then stained with Alexa Fluor conjugated anti-Ku80 antibody (EPR3467, ab202659, abcam) for 1 hour at room temperature. Cells were then mounted onto microscope glass slide, and imaged with freshly mixed imaging buffer (1 mg/mL glucose oxidase, 0.02 mg/mL catalase, 10% glucose, and 100 mM cycteanube (MEA)).

Microscope and Single-Molecule Localization imaging

Super-resolution imaging was performed on a custom-built optical imaging platform based on a Leica DMI 300 inverse microscope, equipped with a 488 nm (OBIS, Coherent) and a 639 nm laser line (MRL-FN-639-800, CNI). Laser lines were reflected into an HCX PL APO 63X NA=1.47 OIL CORR TIRF Objective (Zeiss) by a penta-edged dichroic beam splitter (FF408/504/581/667/762-Di01-22×29), and the emitted fluorescence was further extended by a 2X lens tube (Diagnostic Instruments), filtered by single-band filters (Semrock FF01-531/40 and FF01-676/37 for Alexa Fluor 488 and Alexa Fluor 647, respectively), and collected onto a sCMOS camera (Prim95B, Photometrics). A 405 nm Laser line (MDL-III-405-150, CNI) was also equipped to reactivate Alexa Fluor 647 fluorophores.

For super-resolution imaging, the 488 and 639 laser lines were adjusted to ~ 1.0 and 1.5 kW/cm², and a Highly Inclined and Laminated Optical sheet (HILO) illumination mode for sample excitation. Alexa Fluor 488 and Alexa Fluor 647 were sequentially excited and their emitted fluorescence was also sequentially collected by switching the single-band filters in a filter wheel. The emitted photons were collected onto a sCMOS camera (Photometrics Prime 95B), and a minimum of 2000 frames at 33 Hz were recorded for each image stack.

Reconstruction from Single-Molecule Localization image to super-resolution image

Each image of the image stack was first of all box-filtered with the box size of 4 times of the FWHM of a 2D gaussian point spread function (PSF). Considering the patterned noise for sCMOS camera construction, each pixel was weighted by the inverse of its noise variance during filtering. The low-pass filtered image was then extracted as the background from the raw image. The local maximums of the image were then recognized and segmented into single PSFs for 2D gaussian fit.

2D gaussian fit of each individual PSF was performed by GPU using the Maximum Likelihood Estimation (MLE) algorithm. Briefly, the likelihood function of each pixel was constructed by convolving the Poisson distribution of the shot noise governed by the photons emitted from fluorophores nearby, and the gaussian distribution of the readout noise of the camera itself⁶⁵. The offsets, variance, and analog-to-digital conversion factor of each pixel of the camera was calibrated beforehand. The position, amplitude, sigma, and background of each individual 2D gaussian PSF were addressed by maximizing the likelihood function of each pixel given its readout digital value. The fitting accuracy was estimated by Cramér-Rao lower bound (CRLB) and the accuracy of Alexa Fluor 488 and Alexa Fluor 647 in this work are ~ 17 and 13 nm, respectively.

Alignment of images of different colors

Mapping of the two colors was performed using a polynomial mapping algorithm. Briefly, a spatially separated and broad-spectrum fluorescent bead (TetraSpec, ThermoFisher) slide was imaged on both color channels. Mass centers of each same bead but in different channels were recorded pairwise. The coefficient of each term in a 2nd order polynomial function was optimized by training the polynomial regression using all the recorded bead's centers. The optimized polynomial function was then used for two color channels alignment. We note that choosing the proper order of polynomial function for optimization depends on the number of training beads, and higher order but not enough trainees would lead to overfit.

Pair-Correlation analysis

Coordinates localized within ~ 2.5 times of the averaged localization uncertainty, and from consecutive frames were considered as artificial blinking and grouped as one coordinate. This coordinate was calculated by taking the $1/\text{var}$ weighted average of all the coordinates within the group, where var is the localization uncertainty of each coordinate. The grouped coordinates were then rendered onto a pixelized image of 5 nm/pixel, and the pair-correlation was performed on this image by series 2D Fourier and inverse Fourier transfers⁶⁶. The correlation was then fitted into two correlation terms: the correlation among coordinates within localization uncertainty, and that among the coordinates that form a cluster/molecular assemble/filament. Considering the size of XLF 'clusters' is not randomly distributed, we fitted the second term into a normal distribution and interpret the fitted sigma as the apparent radius of the averaged XLF 'cluster' radius (Figure XC) across the image.

Data availability

Crystal structures are deposited at the pdb with the following codes Ku-DNA-pAPLF (6ERF), Ku-DNA-pXLF (6ERH) and Ku-DNA-pXLFshort (6ERG) Source data for figures 3, 4 and 5 are available with the paper online Other data that support the findings of this study are available from the corresponding author upon reasonable request.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

J-B.C is supported by ARC program (SLS220120605310), ANR (ANR-12-SVSE8-012), INCA DomRep (PLBIO 2012-280), CEFIPRA grant 5203C and by the French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INBS-05. Work in E.R's laboratory is supported by National Institutes of Health Grants CA187612, GM108119, and the American Cancer Society RSG DMC-16-241-01-DMC. P.C's team is supported by the Ligue Nationale Contre Le Cancer (Equipe labellisée 2013 and 2018) and Electricité de France (EDF, Conseil de Radioprotection). J-B.C, M.M and P.C are supported by ANR (CE12 2017 NHEJLIG4 grant). We thank K.W. Caldecott (University of Sussex, Brighton, UK) for the gift of anti-APLF antibody and J.M. Stark (City of Hope, Duarte, USA) for the gift of U2OS cells engineered with the integrated distal-end-joining reporter. We would like to thank the Imaging Core Facility TRI-IPBS, in particular S. Mazeres and R. Poincloux for maintenance of the live-cell microscopy equipment and E. Näser for the maintenance of the flow-cytometers. Flow-cytometry equipment was with the financial support from ITMO Cancer Aviesan (Alliance Nationale Pour les Sciences de la Vie et de la Santé, National Alliance for Life Science and Health) within the framework of Cancer Plan. We thank J-P. de Villartay and F. Theillet for careful reading of the manuscript.

REFERENCES

1. Lieber MR The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. *Annu Rev Biochem* 79, 181–211 (2010). [PubMed: 20192759]
2. Chang HHY, Pannunzio NR, Adachi N & Lieber MR Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol* 18, 495–506 (2017). [PubMed: 28512351]
3. Walker JR, Corpina RA & Goldberg J Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* 412, 607–14 (2001). [PubMed: 11493912]
4. Grundy GJ, Moulding HA, Caldecott KW & Rulten SL One ring to bring them all-The role of Ku in mammalian non-homologous end joining. *DNA Repair (Amst)* 17, 30–8 (2014). [PubMed: 24680220]
5. Rulten SL & Grundy GJ Non-homologous end joining: Common interaction sites and exchange of multiple factors in the DNA repair process. *Bioessays* 39(2017).
6. Costantini S, Woodbine L, Andreoli L, Jeggo PA & Vindigni A Interaction of the Ku heterodimer with the DNA ligase IV/Xrcc4 complex and its regulation by DNA-PK. *DNA Repair (Amst)* 6, 712–22 (2007). [PubMed: 17241822]
7. Nick McElhinny SA, Snowden CM, McCarville J & Ramsden DA Ku recruits the XRCC4-ligase IV complex to DNA ends. *Mol Cell Biol* 20, 2996–3003 (2000). [PubMed: 10757784]
8. Yano KI, Morotomi-Yano K, Lee KJ & Chen DJ Functional significance of the interaction with Ku in DNA double-strand break recognition of XLF. *FEBS Lett* 585, 841–6 (2011). [PubMed: 21349273]
9. Grundy GJ et al. APLF promotes the assembly and activity of non-homologous end joining protein complexes. *Embo J* 32, 112–25 (2013). [PubMed: 23178593]
10. Ropars V et al. Structural characterization of filaments formed by human Xrcc4-Cernunnos/XLF complex involved in nonhomologous DNA end-joining. *Proc Natl Acad Sci U S A* 108, 12663–8 (2011). [PubMed: 21768349]
11. Hammel M et al. XRCC4 interactions with XRCC4-like factor (XLF) create an extended grooved scaffold for DNA ligation and double-strand break repair. *J Biol Chem* 286, 32638–50 (2011). [PubMed: 21775435]
12. Andres SN et al. A human XRCC4-XLF complex bridges DNA. *Nucleic Acids Res* 40, 1868–78 (2012). [PubMed: 22287571]
13. Wu Q et al. Non-homologous end-joining partners in a helical dance: structural studies of XLF-XRCC4 interactions. *Biochem Soc Trans* 39, 1387–92 (2011). [PubMed: 21936820]
14. Reid DA et al. Organization and dynamics of the nonhomologous end-joining machinery during DNA double-strand break repair. *Proc Natl Acad Sci U S A* 112, E2575–84 (2015). [PubMed: 25941401]
15. Ahel I et al. Poly(ADP-ribose)-binding zinc finger motifs in DNA repair/checkpoint proteins. *Nature* 451, 81–5 (2008). [PubMed: 18172500]
16. Rulten SL, Cortes-Ledesma F, Guo L, Iles NJ & Caldecott KW APLF (C2orf13) is a novel component of poly(ADP-ribose) signaling in mammalian cells. *Mol Cell Biol* 28, 4620–8 (2008). [PubMed: 18474613]
17. Kanno S et al. A novel human AP endonuclease with conserved zinc-finger-like motifs involved in DNA strand break responses. *Embo J* 26, 2094–103 (2007). [PubMed: 17396150]
18. Shirodkar P, Fenton AL, Meng L & Koch CA Identification and functional characterization of a Ku-binding motif in Aprataxin Polynucleotide kinase/phosphatase-Like Factor (APLF). *J Biol Chem* 288, 19604–13 (2013). [PubMed: 23689425]
19. Hammel M et al. An Intrinsically Disordered APLF Links Ku, DNA-PKcs, and XRCC4-DNA Ligase IV in an Extended Flexible Non-homologous End Joining Complex. *J Biol Chem* 291, 26987–27006 (2016). [PubMed: 27875301]
20. Arnoult N et al. Regulation of DNA repair pathway choice in S and G2 phases by the NHEJ inhibitor CYREN. *Nature* 549, 548–552 (2017). [PubMed: 28959974]

21. Rossi ML, Ghosh AK & Bohr VA Roles of Werner syndrome protein in protection of genome integrity. *DNA Repair (Amst)* 9, 331–44 (2010). [PubMed: 20075015]
22. Tadi SK et al. PAXX Is an Accessory c-NHEJ Factor that Associates with Ku70 and Has Overlapping Functions with XLF. *Cell Rep* 17, 541–555 (2016). [PubMed: 27705800]
23. Ochi T et al. DNA repair. PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair. *Science* 347, 185–8 (2015). [PubMed: 25574025]
24. Bekker-Jensen S et al. Human Xip1 (C2ORF13) is a novel regulator of cellular responses to DNA strand breaks. *J Biol Chem* 282, 19638–43 (2007). [PubMed: 17507382]
25. Iles N, Rulten S, El-Khamisy SF & Caldecott KW APLF (C2orf13) is a novel human protein involved in the cellular response to chromosomal DNA strand breaks. *Mol Cell Biol* 27, 3793–803 (2007). [PubMed: 17353262]
26. Macrae CJ, McCulloch RD, Ylanko J, Durocher D & Koch CA APLF (C2orf13) facilitates nonhomologous end-joining and undergoes ATM-dependent hyperphosphorylation following ionizing radiation. *DNA Repair (Amst)* 7, 292–302 (2008). [PubMed: 18077224]
27. Malivert L et al. The C-terminal domain of Cernunnos/XLF is dispensable for DNA repair in vivo. *Mol Cell Biol* 29, 1116–22 (2009). [PubMed: 19103754]
28. Brosey CA, Ahmed Z, Lees-Miller SP & Tainer JA What Combined Measurements From Structures and Imaging Tell Us About DNA Damage Responses. *Methods Enzymol* 592, 417–455 (2017). [PubMed: 28668129]
29. Blier PR, Griffith AJ, Craft J & Hardin JA Binding of Ku protein to DNA. Measurement of affinity for ends and demonstration of binding to nicks. *J Biol Chem* 268, 7594–601 (1993). [PubMed: 8463290]
30. Dolinsky TJ et al. PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res* 35, W522–5 (2007). [PubMed: 17488841]
31. Cheng Q et al. Ku counteracts mobilization of PARP1 and MRN in chromatin damaged with DNA double-strand breaks. *Nucleic Acids Res* 39, 9605–9619 (2011). [PubMed: 21880593]
32. Langer A et al. Protein analysis by time-resolved measurements with an electro-switchable DNA chip. *Nat Commun* 4, 2099 (2013). [PubMed: 23839273]
33. Buck D et al. Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell* 124, 287–99 (2006). [PubMed: 16439204]
34. Yano KI & Chen DJ Live cell imaging of XLF and XRCC4 reveals a novel view of protein assembly in the non-homologous end-joining pathway. *Cell Cycle* 7, 1321–5 (2008). [PubMed: 18418068]
35. Hammel M, Yu Y, Fang S, Lees-Miller SP & Tainer JA XLF Regulates Filament Architecture of the XRCC4.Ligase IV Complex. *Structure* 18, 1431–42 (2010). [PubMed: 21070942]
36. Malivert L et al. Delineation of the XRCC4 interacting region in the globular head domain of cernunnos/XLF. *J Biol Chem* 285, 26475–83 (2010). [PubMed: 20558749]
37. Bennardo N, Cheng A, Huang N & Stark JM Alternative-NHEJ Is a Mechanistically Distinct Pathway of Mammalian Chromosome Break Repair. *PLoS Genet* 4, e1000110 (2008). [PubMed: 18584027]
38. Bennardo N, Gunn A, Cheng A, Hasty P & Stark JM Limiting the persistence of a chromosome break diminishes its mutagenic potential. *PLoS Genet* 5, e1000683 (2009). [PubMed: 19834534]
39. Gunn A, Bennardo N, Cheng A & Stark JM Correct end use during end joining of multiple chromosomal double-strand breaks is influenced by repair protein RAD50, DNA-dependent protein kinase DNA-PKcs, and transcription context. *J Biol Chem* 286, 42470–82 (2011). [PubMed: 22027841]
40. Grundy GJ et al. The Ku-binding motif is a conserved module for recruitment and stimulation of non-homologous end-joining proteins. *Nat Commun* 7, 11242 (2016). [PubMed: 27063109]
41. Brouwer I et al. Sliding sleeves of XRCC4-XLF bridge DNA and connect fragments of broken DNA. *Nature* 535, 566–9 (2016). [PubMed: 27437582]
42. Andres SN, Modesti M, Tsai CJ, Chu G & Junop MS Crystal Structure of Human XLF: A Twist in Nonhomologous DNA End-Joining. *Mol Cell* 28, 1093–101 (2007). [PubMed: 18158905]

43. Lu H, Pannicke U, Schwarz K & Lieber MR Length-dependent binding of human XLF to DNA and stimulation of XRCC4: DNA ligase IV activity. *J Biol Chem* 282, 11155–62 (2007). [PubMed: 17317666]
44. Wu PY et al. Interplay between cernunnos-XLF and NHEJ proteins at DNA ends in the cell. *J Biol Chem* 282, 31937–43 (2007). [PubMed: 17720816]
45. Mari PO et al. Dynamic assembly of end-joining complexes requires interaction between Ku70/80 and XRCC4. *Proc Natl Acad Sci U S A* 103, 18597–602 (2006). [PubMed: 17124166]
46. Hsu HL, Yannone SM & Chen DJ Defining interactions between DNA-PK and ligase IV/XRCC4. *DNA Repair (Amst)* 1, 225–35 (2002). [PubMed: 12509254]
47. Leber R, Wise TW, Mizuta R & Meek K The XRCC4 gene product is a target for and interacts with the DNA-dependent protein kinase. *J Biol Chem* 273, 1794–801 (1998). [PubMed: 9430729]
48. Wang YG, Nnakwe C, Lane WS, Modesti M & Frank KM Phosphorylation and regulation of DNA ligase IV stability by DNA-dependent protein kinase. *J Biol Chem* 279, 37282–90 (2004). [PubMed: 15194694]
49. Cottarel J et al. A noncatalytic function of the ligation complex during nonhomologous end joining. *J Cell Biol* 200, 173–86 (2013). [PubMed: 23337116]
50. Graham TG, Walter JC & Loparo JJ Two-Stage Synapsis of DNA Ends during Non-homologous End Joining. *Mol Cell* 61, 850–8 (2016). [PubMed: 26990988]
51. Menon V & Povirk LF XLF/Cernunnos: An important but puzzling participant in the nonhomologous end joining DNA repair pathway. *DNA Repair (Amst)* 58, 29–37 (2017). [PubMed: 28846869]
52. Jspeert H et al. XLF deficiency results in reduced N-nucleotide addition during V(D)J recombination. *Blood* 128, 650–9 (2016). [PubMed: 27281794]

METHODS-ONLY-REFERENCES

53. Bieniossek C, Imasaki T, Takagi Y & Berger I MultiBac: expanding the research toolbox for multiprotein complexes. *Trends Biochem Sci* 37, 49–57 (2012). [PubMed: 22154230]
54. Kabsch W XDS. *Acta Crystallogr D Biol Crystallogr* 66, 125–32 (2010). [PubMed: 20124692]
55. Winn MD et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67, 235–42 (2011). [PubMed: 21460441]
56. Smart OS et al. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr D Biol Crystallogr* 68, 368–80 (2012). [PubMed: 22505257]
57. Afonine PV et al. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68, 352–67 (2012). [PubMed: 22505256]
58. Emsley P, Lohkamp B, Scott WG & Cowtan K Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66, 486–501 (2010). [PubMed: 20383002]
59. Chen VB et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66, 12–21 (2010). [PubMed: 20057044]
60. Perez J & Nishino Y Advances in X-ray scattering: from solution SAXS to achievements with coherent beams. *Curr Opin Struct Biol* 22, 670–8 (2012). [PubMed: 22954648]
61. Petoukhov MV et al. New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr* 45, 342–350 (2012). [PubMed: 25484842]
62. Svergun DI Small-angle X-ray and neutron scattering as a tool for structural systems biology. *Biol Chem* 391, 737–43 (2010). [PubMed: 20482320]
63. Belin BJ, Lee T & Mullins RD DNA damage induces nuclear actin filament assembly by Formin-2 and Spire-(1/2) that promotes efficient DNA repair. *Elife* 4, e07735 (2015). [PubMed: 26287480]
64. Britton S, Coates J & Jackson SP A new method for high-resolution imaging of Ku foci to decipher mechanisms of DNA double-strand break repair. *J Cell Biol* 202, 579–95 (2013). [PubMed: 23897892]

65. Huang F et al. Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms. *Nat Methods* 10, 653–8 (2013). [PubMed: 23708387]
66. Veatch SL et al. Correlation functions quantify super-resolution images and estimate apparent clustering due to over-counting. *PLoS ONE* 7, e31457 (2012). [PubMed: 22384026]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

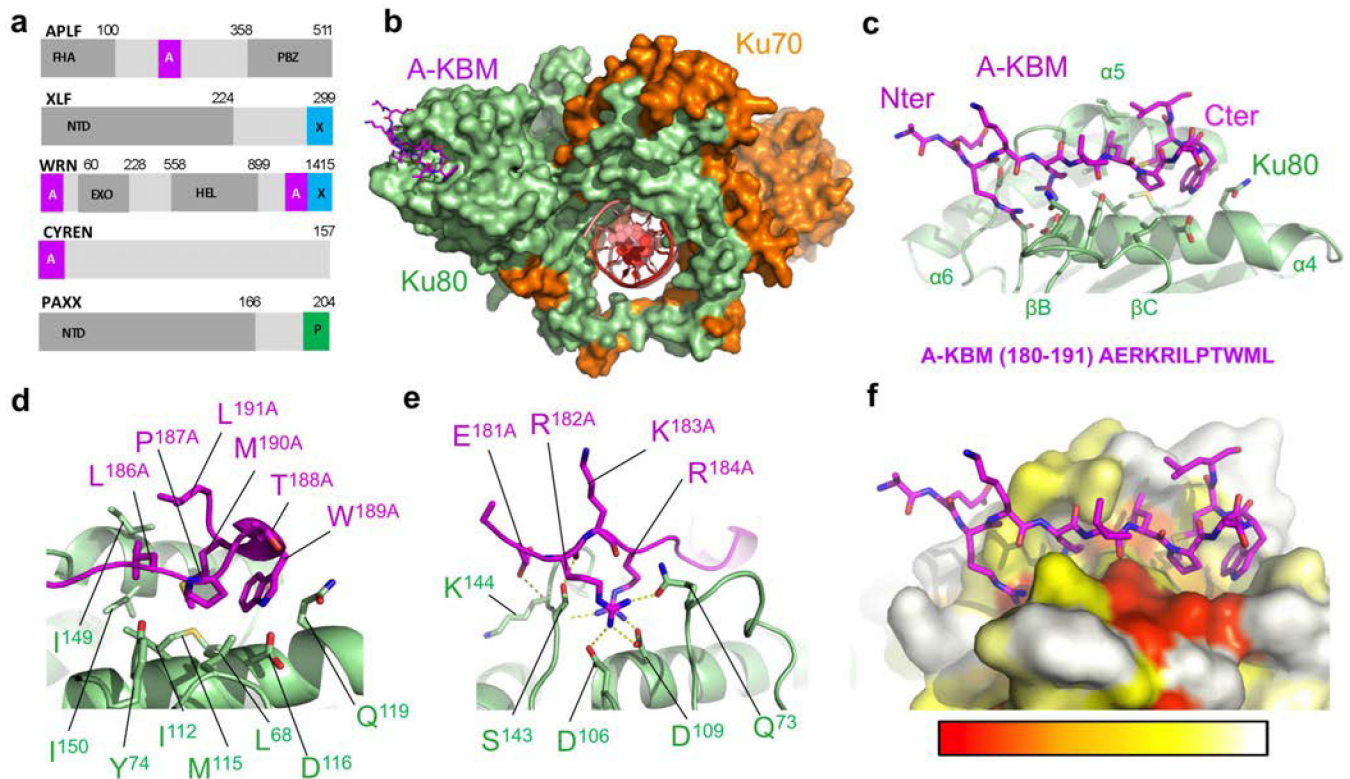


Figure 1. Crystal structure of the APLF KBM (A-KBM) bound to the Ku80 vWA domain. **(a)** Positions of the A-KBM (magenta) and X-KBM (blue) motifs in APLF, XLF, WRN and CYREN. The C-terminal domain of PAXX contains a P-KBM that interacts with Ku70 subunit. NTD: N-terminal domain. **(b)** Overall view of the quaternary complex Ku70/Ku80/hDNA/(APLF peptide). The A-KBM (magenta) binds at the periphery of the Ku80 (light green) vWA domain. The Ku70 subunit and hDNA are represented respectively in orange and red. The hairpin part of the DNA has been removed for clarity. **(c)** The N-terminal part of the A-KBM motif has an extended conformation whereas the C-terminal residues form a turn. **(d-e)** Zoom of the interactions made by (d) the hydrophobic patch and (e) the basic patch of the A-KBM. **(f)** The A-KBM binding site is delineated by conserved residues of Ku80 vWA domain. The binding site is represented in surface mode with amino acids colored according to their conservation rate: red (highly conserved) to white (not conserved)). The conservation rate was measured using sequences of metazoan Ku80. The orientation is the same as in (c).

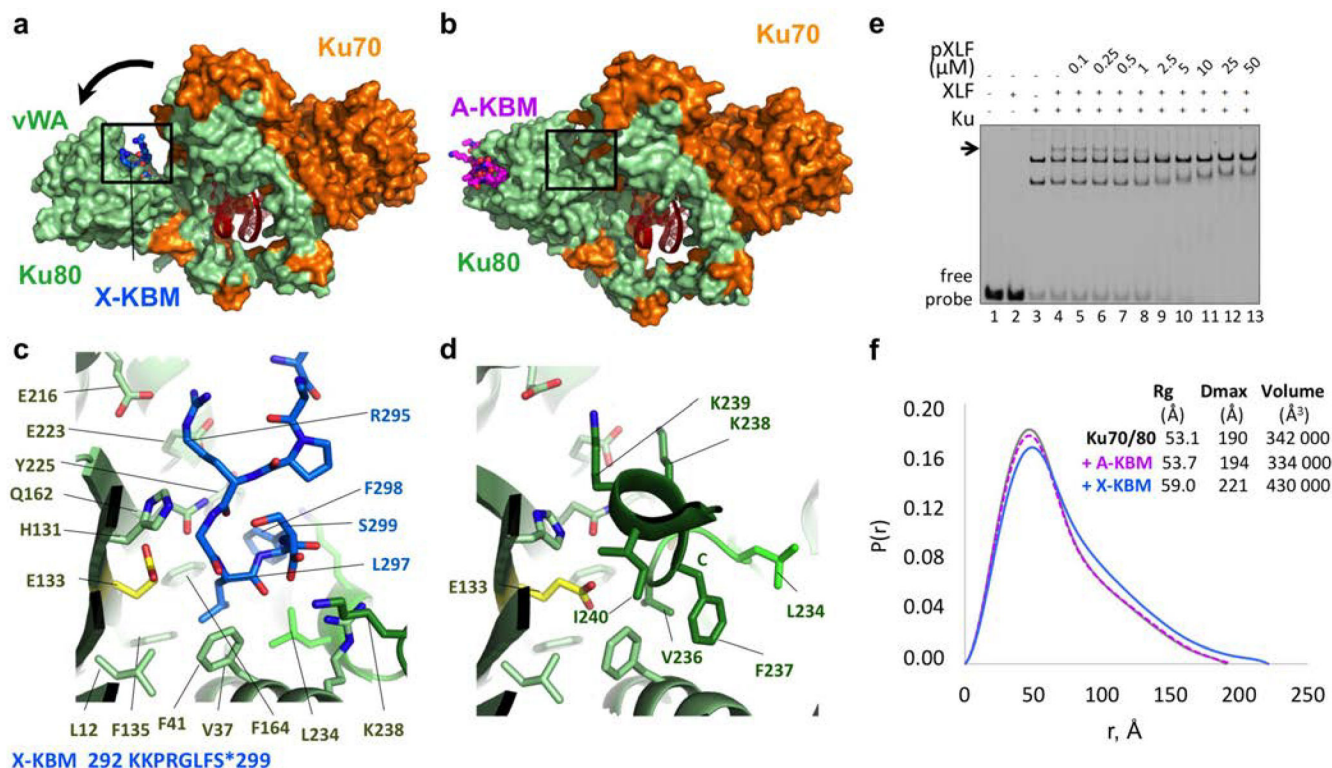


Figure 2. Crystal structure of the XPLF KBM (X-KBM) bound to the Ku80.

(a) Crystal structure of the quaternary complex Ku70-Ku80-DNA-(X-KBM peptide). The X-KBM (blue) binds in an internal site of the Ku80 subunit created upon an outward rotation of the vWA domain. The Ku80 vWA opening creates a large groove between the Ku80 vWA and the rest of the heterodimer. (b) The crystal structure of Ku70/Ku80/DNA in presence of the A-KBM is shown with the same orientation. (c-d) Comparison of the X-KBM binding site in presence of X-KBM (c) or A-KBM (d) peptides. The X-KBM interacts with Ku80 residues involved in Ku intramolecular contacts in the closed state of Ku observed with the A-KBM or with no peptide. The last GLFS residues of the X-KBM interact with the bottom of the groove formed in the open state. The glutamic acid presents an atypical hydrophobic environment and could be at the origin of the vWA instability. The X-KBM residues occupy the position of the helix 236–241 of Ku80 in the closed conformation and some X-KBM side chains (R295^X, L297^X and F298^X) mimic the intramolecular interactions made by Ku80 residues with the vWA domain. (e) Gel shift assay with XLF and Ku in presence of a 50bp DNA with a FAM in 5' and competition with pXLF containing the X-KBM motif. The arrow indicates the XLF Ku-DNA complex. Uncropped gel image is shown in Supplementary Data Set 1. (f) The pair distributions P(r) obtained in solution by SAXS analysis indicates an opening of the Ku70/Ku80/DNA complex with higher Dmax and Rg in presence of the X-KBM (blue line) compared to the Ku/DNA complex without peptide (grey line) and to the A-KBM complex (magenta line). Values deduced from SAXS analysis are reported beside the curves.

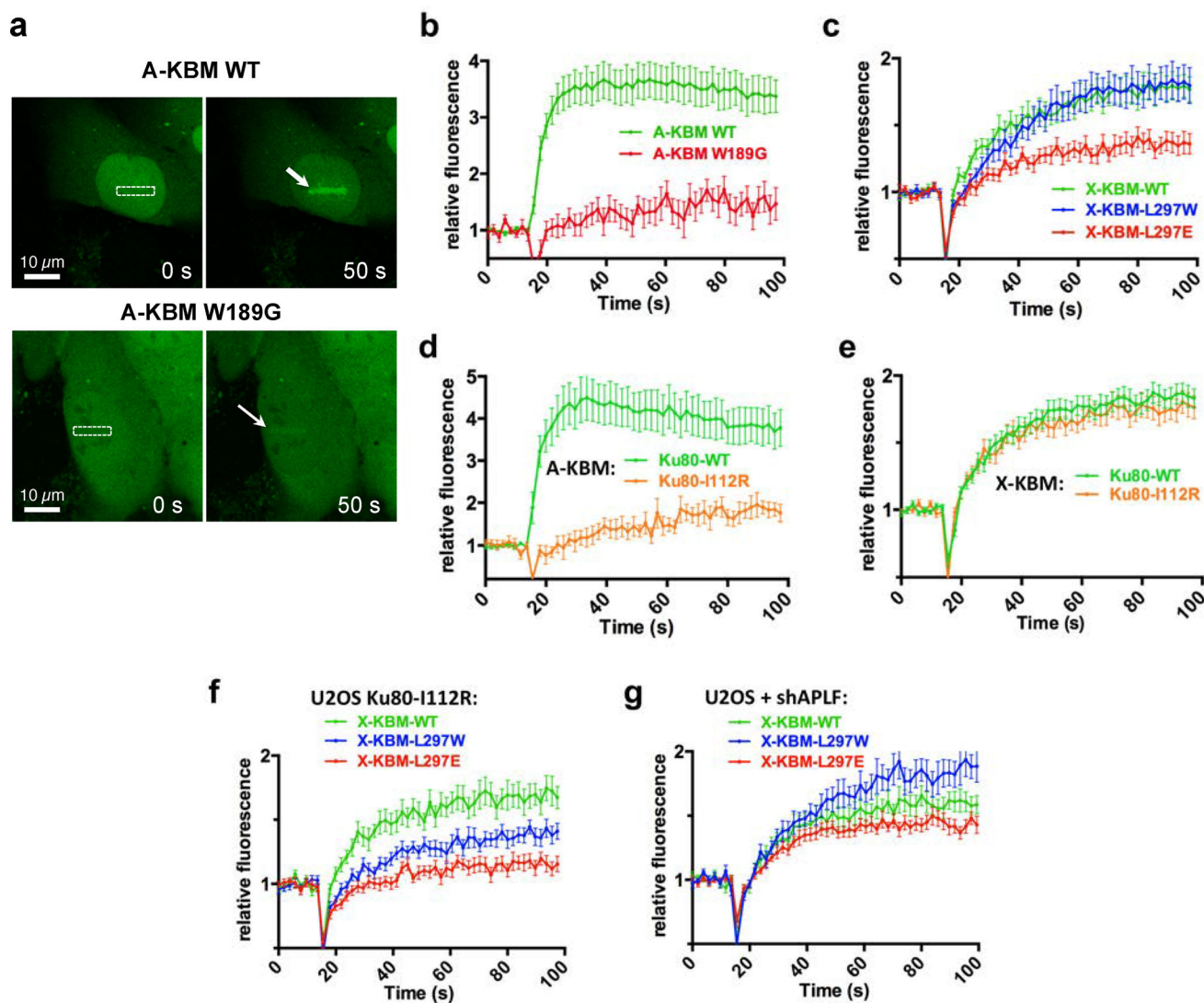


Figure 3. Life cell imaging of A-KBM and X-KBM recruitment after nuclear micro-irradiation. (a) Wild-type (WT) and mutant CFP-(A-KBM) behaviour at 0 s and 50 s after laser nuclear micro-irradiation. The white rectangle and arrows mark irradiated areas. Magnification: X40. (b) Dynamics of wild-type and mutant CFP-(A-KBM) at laser-induced damage sites in U2OS cells. Mean values of relative fluorescence with s.e.m. were calculated from data obtained in several individual cells: $n=23$ and 19 cells for WT and mutant A-KBM, respectively). p values at last time point were calculated using unpaired two-tailed t -test: WT vs W189G $p<0.0001$. (c) Dynamics of wild-type and mutant CFP-(X-KBM) at laser-damaged sites as in b). $n=27$, 21 , and 24 cells for WT, L297E and L297W X-KBM, respectively. p values at last time point : WT vs L297W $p=0.8574$; WT vs L297E $p=0.0021$. (d-e) Dynamics of CFP-(A-KBM) (d) and (X-KBM) (e) at laser damaged sites in cells expressing wild-type or I122R mutant Ku80 as in b). $n=20$, 13 cells for A-KBM in WT or I122R Ku80, and $n=48$ and 39 cells for X-KBM in WT or I122R Ku80, respectively. p values at last time point : (d) WT vs I112R $p=0.0002$; (e) WT vs I112R $p=0.5692$ (f-g) Dynamics of wild-type and mutant CFP-(X-KBM) at laser-damaged sites in cells expressing

I112R mutant Ku80 (f) or treated with a shAPLF (g) as in b). n=26, 28, and 21 cells for WT, L297E and L297W X-KBM in (f), and n=15 cells for each of WT, L297E and L297W X-KBM in (g). *p* values at last time point : (f) WT vs L297W *p*=0.023; WT vs L297E *p*<0.0001; (g) WT vs L297W *p*=0.0144; WT vs L297E *p*=0.2654.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

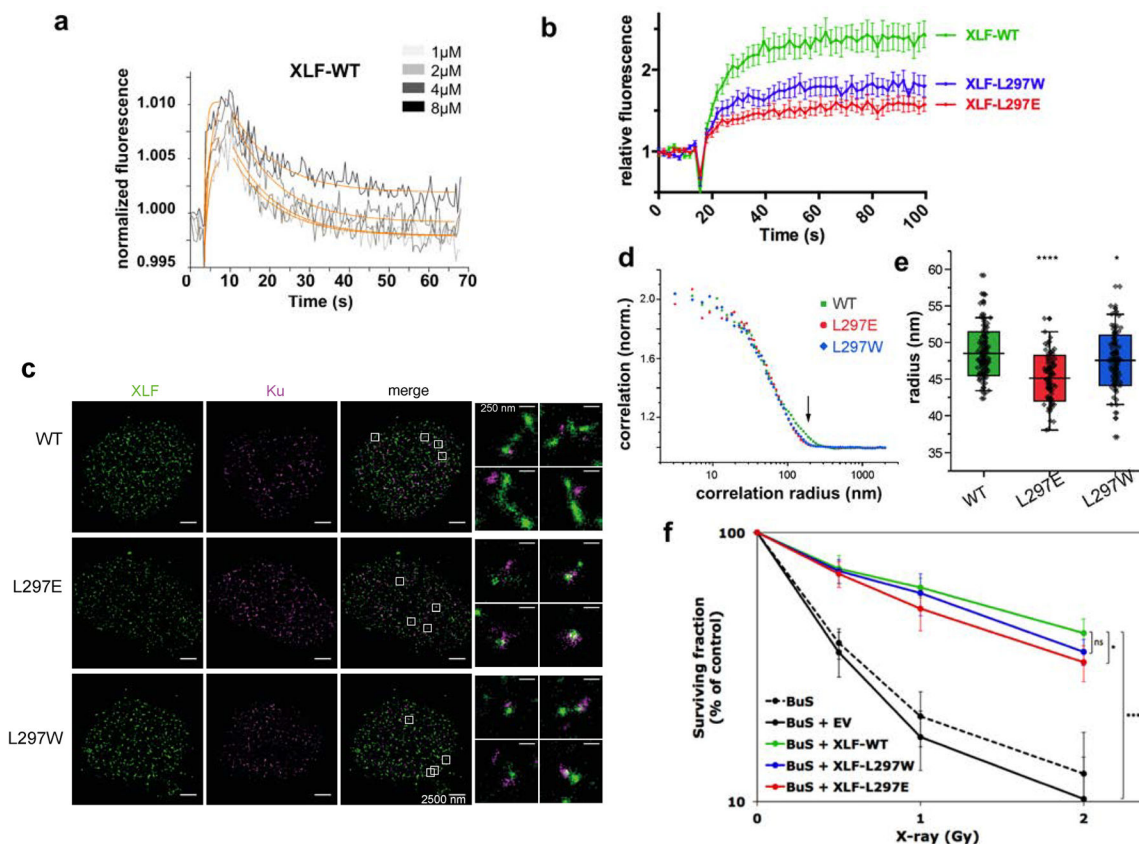


Figure 4. Biophysical and cellular analyses of XLF mutants in X-KBM.

(a) SwitchSENSE kinetic analysis of the WT XLF interaction with Ku-DNA complexes. Solid grey lines represent raw data (from 1 to 8 μM; light grey to dark grey; averages of triplicates). Global fitting was performed, following a single-exponential function (solid orange lines) yielding kinetic rate constants; $k_{ON} = 4.7 \pm 1.7 \cdot 10^5 \text{ M}^{-1}\text{s}^{-1}$ and $k_{OFF} = 9.1 \pm 0.4 \cdot 10^{-2}\text{s}^{-1}$ for XLF(wt). (b) Dynamics of wild-type and mutant CFP-XLF at laser-damaged sites in BuS cells as in Figure 3b. $n=20$ cells for WT, L297E and L297W XLF. p values at last time point : WT vs L297W $p=0.0093$; WT vs L297E $p<0.0001$. (c) Representative super-resolution images of WT, L297E mutant, and L297W mutant BuS nucleus, with XLF and Ku displayed in green and magenta, respectively (scale = 2500 nm). Right: zoomed-in areas (scale = 250 nm). (d) Representative pair correlation function calculated from the $8 \times 8 \mu\text{m}^2$ center square of one XLF nucleus image of WT (green), L297E (red), and L297W (blue) mutants. WT XLF shows bigger correlation radius (arrow). (e) Statistics of XLF foci size. Each plot represents the average XLF foci size (indicated as radius translated from the correlation radius) in one nucleus. Box's height displays the s.d. with the mean value labelled in the middle. $n=116, 95, 104$ nuclei for WT, L297E, and L297W. The two-sample unpaired t -test between WT and L297E is $p=10^{-13}$ while that between WT and L297W is $p=0.03$. (f) Cell survival of BuS cells complemented with vector (EV) or WT or mutated XLF. y axis is log scale. Error bars represent s.d., $n=5$ to 6 independent experiments. p values were calculated using unpaired two-tailed t -test: WT vs EV $p=1.788\text{e-}06$; WT vs LW $p=0.068$; WT vs LE $p=0.021$. ($*p<0.05$, $**p<0.01$, $***p<0.001$).

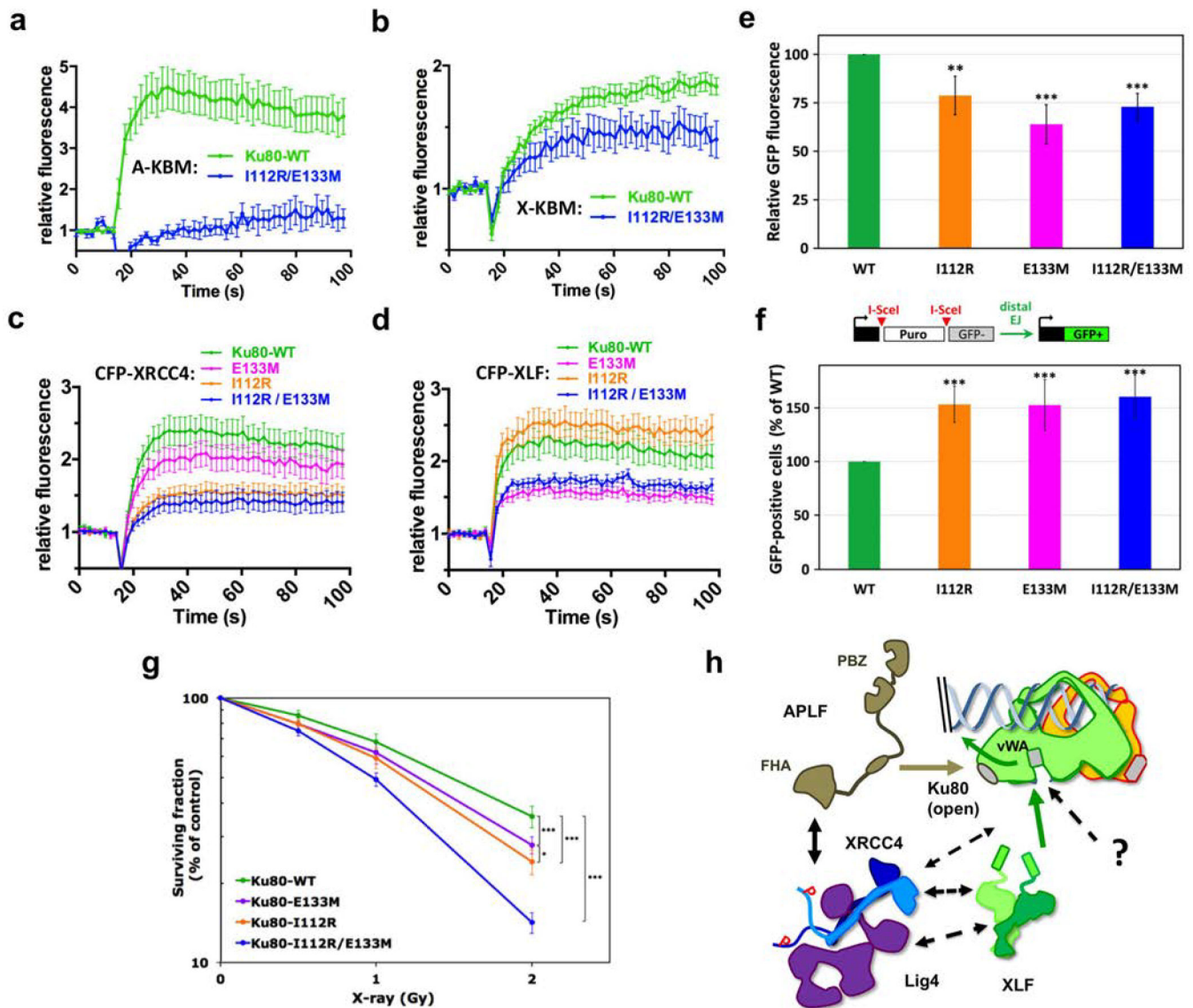


Figure 5. Effects of Ku80 mutations in APLF and XLF binding sites.

(a-b) Dynamics of CFP-(A-KBM) (a) and (X-KBM) (b) at laser damaged sites, in U2OS cells expressing wild-type or I112R/E133M mutant Ku80. $n=20$ and 9 cells in (a) and $n=48$ and 11 cells in (b) for WT and mutant Ku80. p values at last time point: (a) WT vs I112R/E133M $p=0.001$; (b) WT vs I112R/E133M $p=0.0111$. (c-d) Dynamics of CFP-XRCC4 (c) and XLF (d) in cells expressing wild-type, I112R, E133M or I112R/E133M mutant Ku80. $n=38$, 27, 28, and 24 cells for WT, E133M, I112R and I112R/E133M Ku80 conditions in (c) and $n=24$, 26, 20 and 23 cells for I112R, WT, I112R/E133M and E133M Ku80 conditions in (d). p values at last time point: (c) WT vs E133M $p=0.532$; WT vs I112R $p=0.0133$; WT vs I112R/E133M $p=0.0048$; (d) WT vs I112R $p=0.246$; WT vs E133M $p=0.0048$; WT vs I112R/E133M $p=0.0248$. (e) End-joining activity in U2OS cells expressing mutated or WT Ku80. Error bars represent s.d., $n=4$ independent experiments. p values (unpaired two-tailed t -test): WT vs E133M $p=0.0004$; WT vs I112R $p=0.0052$; WT vs I112R/E133M $p=0.0002$. (f) Distal end-joining in U2OS cells containing mutated or WT Ku80. Error bars represent

s.d., n=7 independent experiments. *p* values (unpaired two-tailed *t*-test): WT vs E133M $p=7.49 \times 10^{-5}$; WT vs I112R $p=2.21 \times 10^{-6}$; WT vs I112R/E133M $p=4.05 \times 10^{-6}$. **(g)** Survival of U2OS cells expressing WT or mutated Ku80. y axis is log scale. Error bars represent s.d., n=7 to 10 independent experiments. *p* (unpaired two-tailed *t*-test): WT vs I112R $p=1.47 \times 10^{-6}$; WT vs E133M $p=6.32 \times 10^{-5}$; WT vs I112R/E133M $p=2.52 \times 10^{-13}$; I112R vs E133M $p=0.011$. Significant *p*-values are indicated as follows: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. **(h)** Model for APLF and XLF KBMs function during NHEJ.

Table 1

Data collection and refinement statistics

	Ku-pAPLF (PDB 6ERF)	Ku-pXLF (PDB 6ERH)	Ku-pXLFs (PDB 6ERG)
Data collection			
Space group	P ₁	P2 ₁	P2 ₁
Cell dimensions :			
<i>a, b, c</i> (Å)	98.9, 140.8, 150.3	111.8, 118.9, 128.2	111.7, 114.3, 127.2
α, β, γ (°)	68.6, 80.8, 81.2	90.0, 93.1, 90.0	90.0, 93.1, 90.0
Resolution (Å)	49.5–3.0 (3.28–3.0)	49.01–2.8 (3.16–2.8)	50–2.9 (3.09–2.9)
Anisotropy resolution limits (Å) [§]	2.9, 3.9, 3.3	2.7, 4.4, 3.4	2.8, 3.8, 3.0
Resolution limit overall (Å) [§]	3.15	3.11	3.0
<i>R</i> _{merge}	0.056 (0.79)	0.164 (1.35)	0.171 (2.48)
<i>R</i> _{meas}	0.072 (0.93)	0.182 (1.47)	0.182 (2.56)
<i>R</i> _{rim}	0.051 (0.66)	0.094 (0.76)	0.068 (0.94)
<i>I</i> σ (<i>I</i>)	11.8 (1.3)	9.1 (1.5)	9.7 (1.0)
<i>CC</i> _{1/2}	0.99 (0.60)	0.997 (0.60)	0.996 (0.51)
Completeness (spherical, %) [§]	66.5 (14.3)	51.3 (8.5)	72.9 (21.0)
Completeness (ellipsoidal, %) [§]	92.1 (70.1)	92.4 (69.6)	94.6 (68.5)
Redundancy	3.4 (3.5)	7.1 (6.9)	14.0 (14.2)
Refinement			
Resolution (Å)	49.47–3.01	49.38–2.8	48.76–2.9
No. reflections	90993	41644	51644
<i>R</i> _{work} / <i>R</i> _{free}	0.209/0.227	0.225/0.252	0.218/0.244
No. Atoms	35410	18810	18989
Protein	33225	16649	16821
DNA	2178	2118	2133
SO ₄ ²⁻	0	10	15
Water	7	33	20
B factors (Å ²)			

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

	Ku-pAPLF (PDB 6ERF)	Ku-pXLF (PDB 6ERH)	Ku-pXLFs (PDB 6ERG)
Protein	112	76	94
DNA	201	109	126
R.m.s. deviations			
Bond lengths (Å)	0.008	0.007	0.008
Bond angles (°)	0.96	0.95	0.95

* Values in parentheses are for highest-resolution shell.

§ Values from STARANISO, Global Phasing Ltd.

Table 2 :

Interactions measured by microcalorimetry between Ku70/Ku80, A-KBM, X-KBM and XLF

	Protein in the measurement cell	Ligand in the syringe	K _d (μM)	H (kcal.M ⁻¹)	Remarks
1	Ku _{FL}	pAPLF	0.033 ± 0.01	-13.3 ± 0.1	A-KBM (APLF 174–191)
2	Ku _{FL}	DNA 18bp	0.0041 ± 0.0007	+5.3 ± 0.2	DNA18bp versus Ku
3	Ku _{FL} +DNA-18bp	pAPLF	0.023 ± 0.002	-16.0 ± 0.4	in presence of DNA _{18bp}
4	Ku _{CC}	pAPLF	0.020 ± 0.002	-18.4 ± 0.7	Ku without Cter domains
5	Ku _{FL} I112R	pAPLF	NI ^(a)	NI	Ku80 mutant on APLF site
6	Ku _{FL}	pXLF	4.4 ± 0.2	-2.8 ± 0.2	X-KBM (XLF 281–299)
7	Ku _{FL} +DNA-18bp	pXLF	2.4 ± 0.1	-8.1 ± 0.6	in presence of DNA _{18bp}
8	Ku _{CC}	pXLF	2.2 ± 0.9	-3.1 ± 1.2	Ku without Cter domains
9	Ku _{FL}	XLF	1.0 ± 0.1	-9.0 ^(b)	XLF versus Ku
10	Ku _{FL} +DNA-18bp	XLF	2.35 ± 0.1	-8.1 ± 0.6	XLF versus Ku/DNA _{18bp}
11	Ku _{FL}	pXLF(LW)	0.12 ± 0.03	-12.1 ± 2.0	X-KBM (L297W)
12	Ku _{FL}	pXLF(LE)	NI	NI	X-KBM (L297E)
13	Ku _{FL} +pAPLF	pXLF(LW)	NI	NI	L297W in presence of A-KBM

^(a)NI means no interaction;^(b)The enthalpy value was deduced from the first injection point in absence of lower plateau; The thermograms and isotherms of titration of the experiments corresponding to lines 1, 6 and 11 are presented in Supplementary Figure 2e–g. Interactions in lines 1, 4, 6, 8, 9, 11 were measured in triplicate, and the mean value with standard deviation is reported. Interactions in lines 2, 3, 4, 6, 10 were measured in duplicate, and the mean value with variation between min and max values is reported.

BIBLIOGRAPHIE

- Ahnesorg, Peter, Philippa Smith, et Stephen P. Jackson. 2006. « XLF Interacts with the XRCC4-DNA Ligase IV Complex to Promote DNA Nonhomologous End-Joining ». *Cell* 124 (2): 301-13. <https://doi.org/10.1016/j.cell.2005.12.031>.
- Andres, Sara N, Mauro Modesti, Chun J Tsai, Gilbert Chu, et Murray S Junop. s. d. « Crystal Structure of Human XLF: A Twist in Nonhomologous DNA End-Joining ». *Molecular Cell*, 9.
- Arosio, Daniele, Silvia Costantini, Yong Kong, et Alessandro Vindigni. 2004. « Fluorescence Anisotropy Studies on the Ku-DNA Interaction: ANION AND CATION EFFECTS ». *Journal of Biological Chemistry* 279 (41): 42826-35. <https://doi.org/10.1074/jbc.M406529200>.
- Arosio, Daniele, Sheng Cui, Claudia Ortega, Miroslav Chovanec, Stefania Di Marco, Giancarlo Baldini, Arturo Falaschi, et Alessandro Vindigni. 2002. « Studies on the Mode of Ku Interaction with DNA ». *Journal of Biological Chemistry* 277 (12): 9741-48. <https://doi.org/10.1074/jbc.M111916200>.
- Bétermier, Mireille, Pascale Bertrand, et Bernard S. Lopez. 2014. « Is Non-Homologous End-Joining Really an Inherently Error-Prone Process? » Édité par Sue Jinks-Robertson. *PLoS Genetics* 10 (1): e1004086. <https://doi.org/10.1371/journal.pgen.1004086>.
- Boyd, Susan M., Andrew P. Turnbull, et Björn Walse. 2012. « Fragment Library Design Considerations: Fragment Library Design Considerations ». *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2 (6): 868-85. <https://doi.org/10.1002/wcms.1098>.
- Britton, Sébastien, Julia Coates, et Stephen P. Jackson. 2013. « A New Method for High-Resolution Imaging of Ku Foci to Decipher Mechanisms of DNA Double-Strand Break Repair ». *The Journal of Cell Biology* 202 (3): 579-95. <https://doi.org/10.1083/jcb.201303073>.
- Buck, Dietke, Laurent Malivert, Régina de Chasseval, Anne Barraud, Marie-Claude Fondanèche, Ozden Sanal, Alessandro Plebani, et al. 2006. « Cernunnos, a Novel Nonhomologous End-Joining Factor, Is Mutated in Human Immunodeficiency with Microcephaly ». *Cell* 124 (2): 287-99. <https://doi.org/10.1016/j.cell.2005.12.030>.
- Cala, Olivier, Florence Guillière, et Isabelle Krimm. 2014. « NMR-Based Analysis of Protein–Ligand Interactions ». *Analytical and Bioanalytical Chemistry* 406 (4): 943-56. <https://doi.org/10.1007/s00216-013-6931-0>.
- Chan, D. W. 2002. « Autophosphorylation of the DNA-Dependent Protein Kinase Catalytic Subunit Is Required for Rejoining of DNA Double-Strand Breaks ». *Genes & Development* 16 (18): 2333-38. <https://doi.org/10.1101/gad.1015202>.
- Chang, Howard H. Y., Nicholas R. Pannunzio, Noritaka Adachi, et Michael R. Lieber. 2017. « Non-homologous DNA end joining and alternative pathways to double-strand break repair ». *Nature Reviews Molecular Cell Biology* 18 (mai): 495.

- Chen, X., S. Zhong, X. Zhu, B. Dziegielewska, T. Ellenberger, G. M. Wilson, A. D. MacKerell, et A. E. Tomkinson. 2008. « Rational Design of Human DNA Ligase Inhibitors That Target Cellular DNA Replication and Repair ». *Cancer Research* 68 (9): 3169-77. <https://doi.org/10.1158/0008-5472.CAN-07-6636>.
- Cherry, Amy L., Timothy J. Nott, Geoffrey Kelly, Stuart L. Rulten, Keith W. Caldecott, et Stephen J. Smerdon. 2015. « Versatility in Phospho-Dependent Molecular Recognition of the XRCC1 and XRCC4 DNA-Damage Scaffolds by Aprataxin-Family FHA Domains ». *DNA Repair* 35 (novembre): 116-25. <https://doi.org/10.1016/j.dnarep.2015.10.002>.
- Cheung, Joyce C.Y., Brenda Salerno, et Les A. Hanakahi. 2008. « Evidence for an Inositol Hexakisphosphate-Dependent Role for Ku in Mammalian Nonhomologous End Joining That Is Independent of Its Role in the DNA-Dependent Protein Kinase ». *Nucleic Acids Research* 36 (17): 5713-26. <https://doi.org/10.1093/nar/gkn572>.
- Chirgadze, Dimitri Y., David B. Ascher, Tom L. Blundell, et Bancinyane L. Sibanda. 2017. « DNA-PKcs, Allostery, and DNA Double-Strand Break Repair ». In *Methods in Enzymology*, 592:145-57. Elsevier. <https://doi.org/10.1016/bs.mie.2017.04.001>.
- Cléry, Antoine, Thibault J.M. Sohier, Thomas Welte, Andreas Langer, et Frédéric H.T. Allain. 2017. « SwitchSENSE: A New Technology to Study Protein-RNA Interactions ». *Methods* 118-119 (avril): 137-45. <https://doi.org/10.1016/j.ymeth.2017.03.004>.
- Costantini, Silvia, Lisa Woodbine, Lucia Andreoli, Penny A. Jeggo, et Alessandro Vindigni. 2007. « Interaction of the Ku Heterodimer with the DNA Ligase IV/Xrcc4 Complex and Its Regulation by DNA-PK ». *DNA Repair* 6 (6): 712-22. <https://doi.org/10.1016/j.dnarep.2006.12.007>.
- Craxton, A, J Somers, D Munnur, R Jukes-Jones, K Cain, et M Malewicz. 2015. « XLS (C9orf142) Is a New Component of Mammalian DNA Double-Stranded Break Repair ». *Cell Death & Differentiation* 22 (6): 890-97. <https://doi.org/10.1038/cdd.2015.22>.
- Dalvit, Claudio. 2009. « NMR Methods in Fragment Screening: Theory and a Comparison with Other Biophysical Techniques ». *Drug Discovery Today* 14 (21-22): 1051-57. <https://doi.org/10.1016/j.drudis.2009.07.013>.
- Dynan, W. S., et S. Yoo. 1998. « Interaction of Ku Protein and DNA-Dependent Protein Kinase Catalytic Subunit with Nucleic Acids ». *Nucleic Acids Research* 26 (7): 1551-59. <https://doi.org/10.1093/nar/26.7.1551>.
- Elmroth, Kecke, Jonas Nygren, Susanne Mårtensson, Ismail Hassan Ismail, et Ola Hammarsten. 2003. « Cleavage of Cellular DNA by Calicheamicin □ ». *DNA Repair*, 12.
- Englebienne, Patrick, Anne Van Hoonacker, et Michel Verhas. 2003. « Surface Plasmon Resonance: Principles, Methods and Applications in Biomedical Sciences ». *Spectroscopy* 17 (2-3): 255-73. <https://doi.org/10.1155/2003/372913>.

- Erlanson, Daniel A. 2011. « Introduction to Fragment-Based Drug Discovery ». In *Fragment-Based Drug Discovery and X-Ray Crystallography*, édité par Thomas G. Davies et Marko Hyvönen, 317:1-32. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/128_2011_180.
- Ferenczy, György G., et György M. Keserü. 2010. « Thermodynamics Guided Lead Discovery and Optimization ». *Drug Discovery Today* 15 (21-22): 919-32. <https://doi.org/10.1016/j.drudis.2010.08.013>.
- Gashaw, Isabella, Peter Ellinghaus, Anette Sommer, et Khusru Asadullah. 2011. « What Makes a Good Drug Target? ». *Drug Discovery Today* 16 (23-24): 1037-43. <https://doi.org/10.1016/j.drudis.2011.09.007>.
- Gordaliza, M, P.A García, J.M Miguel del Corral, M.A Castro, et M.A Gómez-Zurita. 2004. « Podophyllotoxin: Distribution, Sources, Applications and New Cytotoxic Derivatives ». *Toxicon* 44 (4): 441-59. <https://doi.org/10.1016/j.toxicon.2004.05.008>.
- Greco, George E., Yoshihiro Matsumoto, Rhys C. Brooks, Zhengfei Lu, Michael R. Lieber, et Alan E. Tomkinson. 2016. « SCR7 Is Neither a Selective nor a Potent Inhibitor of Human DNA Ligase IV ». *DNA Repair* 43 (juillet): 18-23. <https://doi.org/10.1016/j.dnarep.2016.04.004>.
- Grundy, Gabrielle J., Hayley A. Moulding, Keith W. Caldecott, et Stuart L. Rulten. 2014. « One Ring to Bring Them All—The Role of Ku in Mammalian Non-Homologous End Joining ». *DNA Repair* 17 (mai): 30-38. <https://doi.org/10.1016/j.dnarep.2014.02.019>.
- Grundy, Gabrielle J, Stuart L Rulten, Zhihong Zeng, Raquel Arribas-Bosacoma, Natasha Iles, Katie Manley, Antony Oliver, et Keith W Caldecott. 2012. « APLF Promotes the Assembly and Activity of Non-Homologous End Joining Protein Complexes ». *The EMBO Journal* 32 (1): 112-25. <https://doi.org/10.1038/emboj.2012.304>.
- Haber, James E. 1999. « DNA Recombination: The Replication Connection ». *Trends in Biochemical Sciences* 24 (7): 271-75. [https://doi.org/10.1016/S0968-0004\(99\)01413-9](https://doi.org/10.1016/S0968-0004(99)01413-9).
- Hanahan, Douglas, et Robert A Weinberg. 2000. « The Hallmarks of Cancer ». *Cell* 100 (1): 57-70. [https://doi.org/10.1016/S0092-8674\(00\)81683-9](https://doi.org/10.1016/S0092-8674(00)81683-9).
- Hanakahi, L. A. 2002. « Specific Interaction of IP6 with Human Ku70/80, the DNA-Binding Subunit of DNA-PK ». *The EMBO Journal* 21 (8): 2038-44. <https://doi.org/10.1093/emboj/21.8.2038>.
- Hanakahi, Les A, Michael Bartlet-Jones, Claire Chappell, Darryl Pappin, et Stephen C West. 2000. « Binding of Inositol Phosphate to DNA-PK and Stimulation of Double-Strand Break Repair ». *Cell* 102 (6): 721-29. [https://doi.org/10.1016/S0092-8674\(00\)00061-1](https://doi.org/10.1016/S0092-8674(00)00061-1).
- Hande, K.R. 1998. « Etoposide: Four Decades of Development of a Topoisomerase II Inhibitor ». *European Journal of Cancer* 34 (10): 1514-21. [https://doi.org/10.1016/S0959-8049\(98\)00228-7](https://doi.org/10.1016/S0959-8049(98)00228-7).

- Hartley, Katharine O, David Gell, Graeme C.M Smith, Hong Zhang, Nullin Divecha, Margery A Connelly, Arie Admon, Susan P Lees-Miller, Carl W Anderson, et Stephen P Jackson. 1995. « DNA-Dependent Protein Kinase Catalytic Subunit: A Relative of Phosphatidylinositol 3-Kinase and the Ataxia Telangiectasia Gene Product ». *Cell* 82 (5): 849-56. [https://doi.org/10.1016/0092-8674\(95\)90482-4](https://doi.org/10.1016/0092-8674(95)90482-4).
- Jain, Kewal. 2004. « RNAi and siRNA in target validation ». *Drug discovery today* 9 (mai): 307-9. [https://doi.org/10.1016/S1359-6446\(04\)03050-8](https://doi.org/10.1016/S1359-6446(04)03050-8).
- Jette, Nicholas, et Susan P. Lees-Miller. 2015. « The DNA-Dependent Protein Kinase: A Multifunctional Protein Kinase with Roles in DNA Double Strand Break Repair and Mitosis ». *Progress in Biophysics and Molecular Biology* 117 (2-3): 194-205. <https://doi.org/10.1016/j.pbiomolbio.2014.12.003>.
- Kanno, Shin-ichiro, Hiroyuki Kuzuoka, Shigeru Sasao, Zehui Hong, Li Lan, Satoshi Nakajima, et Akira Yasui. 2007. « A Novel Human AP Endonuclease with Conserved Zinc-Finger-like Motifs Involved in DNA Strand Break Responses ». *The EMBO Journal* 26 (8): 2094-2103. <https://doi.org/10.1038/sj.emboj.7601663>.
- Krell, Tino. 2008. « Microcalorimetry: A Response to Challenges in Modern Biotechnology ». *Microbial Biotechnology* 1 (2): 126-36. <https://doi.org/10.1111/j.1751-7915.2007.00013.x>.
- Krimm, Isabelle. 2015. « Le criblage de fragments: Une voie prometteuse pour la conception de médicaments ». *médecine/sciences* 31 (2): 197-202. <https://doi.org/10.1051/medsci/20153102017>.
- Kroemer, Guido, et Jacques Pouyssegur. 2008. « Tumor Cell Metabolism: Cancer's Achilles' Heel ». *Cancer Cell* 13 (6): 472-82. <https://doi.org/10.1016/j.ccr.2008.05.005>.
- Langer, Andreas, Paul A. Hampel, Wolfgang Kaiser, Jelena Knezevic, Thomas Welte, Valentina Villa, Makiko Maruyama, et al. 2013. « Protein Analysis by Time-Resolved Measurements with an Electro-Switchable DNA Chip ». *Nature Communications* 4 (1): 2099. <https://doi.org/10.1038/ncomms3099>.
- Li, Yi, Dimitri Y Chirgadze, Victor M Bolanos-Garcia, Bancinyane L Sibanda, Owen R Davies, Peter Ahnesorg, Stephen P Jackson, et Tom L Blundell. 2008. « Crystal Structure of Human XLF/Cernunnos Reveals Unexpected Differences from XRCC4 with Implications for NHEJ ». *The EMBO Journal* 27 (1): 290-300. <https://doi.org/10.1038/sj.emboj.7601942>.
- Lindahl, Tomas, et Barbro Nyberg. 1972. « Rate of Depurination of Native Deoxyribonucleic Acid ». *Biochemistry* 11 (19): 3610-18. <https://doi.org/10.1021/bi00769a018>.
- Lindsay, Mark A. 2003. « Target discovery ». *Nature Reviews Drug Discovery* 2 (10): 831-38. <https://doi.org/10.1038/nrd1202>.
- Lo, Mei-Chu, Ann Aulabaugh, Guixian Jin, Rebecca Cowling, Jonathan Bard, Michael Malamas, et George Ellestad. 2004. « Evaluation of Fluorescence-Based Thermal Shift

- Assays for Hit Identification in Drug Discovery ». *Analytical Biochemistry* 332 (1): 153-59. <https://doi.org/10.1016/j.ab.2004.04.031>.
- Luo, Ji, Nicole L. Solimini, et Stephen J. Elledge. 2009. « Principles of Cancer Therapy: Oncogene and Non-Oncogene Addiction ». *Cell* 136 (5): 823-37. <https://doi.org/10.1016/j.cell.2009.02.024>.
- Maanen, J. M. S. van, J. Retel, J. de Vries, et H. M. Pinedo. 1988. « Mechanism of Action of Antitumor Drug Etoposide: A Review ». *JNCI Journal of the National Cancer Institute* 80 (19): 1526-33. <https://doi.org/10.1093/jnci/80.19.1526>.
- Mari, P.-O., B. I. Florea, S. P. Persengiev, N. S. Verkaik, H. T. Bruggenwirth, M. Modesti, G. Giglia-Mari, et al. 2006. « Dynamic Assembly of End-Joining Complexes Requires Interaction between Ku70/80 and XRCC4 ». *Proceedings of the National Academy of Sciences* 103 (49): 18597-602. <https://doi.org/10.1073/pnas.0609061103>.
- Mayer, Moriz, et Bernd Meyer. s. d. « Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy », 5.
- Mimori, T, M Akizuki, H Yamagata, S Inada, S Yoshida, et M Homma. 1981. « Characterization of a High Molecular Weight Acidic Nuclear Protein Recognized by Autoantibodies in Sera from Patients with Polymyositis-Scleroderma Overlap. » *Journal of Clinical Investigation* 68 (3): 611-20. <https://doi.org/10.1172/JCI110295>.
- Mullard, Asher. 2012. « Protein–Protein Interaction Inhibitors Get into the Groove ». *Nature Reviews Drug Discovery* 11 (3): 173-75. <https://doi.org/10.1038/nrd3680>.
- Någren, Kjell. 2003. « PET and Knockout Mice in Drug Discovery ». *Drug Discovery Today* 8 (19): 876. [https://doi.org/10.1016/S1359-6446\(03\)02765-X](https://doi.org/10.1016/S1359-6446(03)02765-X).
- Nemoz, Clement, Virginie Ropars, Philippe Frit, Amandine Gontier, Pascal Drevet, Jinchao Yu, Raphaël Guerois, et al. 2018. « XLF and APLF Bind Ku80 at Two Remote Sites to Ensure DNA Repair by Non-Homologous End Joining ». *Nature Structural & Molecular Biology* 25 (10): 971-80. <https://doi.org/10.1038/s41594-018-0133-6>.
- Newman, Janet, David Egan, Thomas S. Walter, Ran Meged, Ian Berry, Marouane Ben Jelloul, Joel L. Sussman, David I. Stuart, et Anastassis Perrakis. 2005. « Towards Rationalization of Crystallization Screening for Small- to Medium-Sized Academic Laboratories: The PACT/JCSG+ Strategy ». *Acta Crystallographica Section D Biological Crystallography* 61 (10): 1426-31. <https://doi.org/10.1107/S0907444905024984>.
- Niesen, Frank H, Helena Berglund, et Masoud Vedadi. 2007. « The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability ». *Nature Protocols* 2 (septembre): 2212.
- Ochi, Takashi, Qian Wu, et Tom L. Blundell. 2014. « The Spatial Organization of Non-Homologous End Joining: From Bridging to End Joining ». *DNA Repair* 17 (mai): 98-109. <https://doi.org/10.1016/j.dnarep.2014.02.010>.

- Pascal, John M., Patrick J. O'Brien, Alan E. Tomkinson, et Tom Ellenberger. 2004. « Human DNA Ligase I Completely Encircles and Partially Unwinds Nicked DNA ». *Nature* 432 (7016): 473-78. <https://doi.org/10.1038/nature03082>.
- Pauling, L., et R. B. Corey. 1953. « A Proposed Structure For The Nucleic Acids ». *Proceedings of the National Academy of Sciences* 39 (2): 84-97. <https://doi.org/10.1073/pnas.39.2.84>.
- Povirk, Lawrence F. 1996. « DNA Damage and Mutagenesis by Radiomimetic DNA-Cleaving Agents: Bleomycin, Neocarzinostatin and Other Eneidyne ». *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 355 (1-2): 71-89. [https://doi.org/10.1016/0027-5107\(96\)00023-1](https://doi.org/10.1016/0027-5107(96)00023-1).
- Redinbo, M. R. 1998. « Crystal Structures of Human Topoisomerase I in Covalent and Noncovalent Complexes with DNA ». *Science* 279 (5356): 1504-13. <https://doi.org/10.1126/science.279.5356.1504>.
- Reid, Dylan A., Sarah Keegan, Alejandra Leo-Macias, Go Watanabe, Natasha T. Strande, Howard H. Chang, Betul Akgol Oksuz, et al. 2015. « Organization and Dynamics of the Nonhomologous End-Joining Machinery during DNA Double-Strand Break Repair ». *Proceedings of the National Academy of Sciences* 112 (20): E2575-84. <https://doi.org/10.1073/pnas.1420115112>.
- Renaud, Jean-Paul, Chun-wa Chung, U. Helena Danielson, Ursula Egner, Michael Hennig, Roderick E. Hubbard, et Herbert Nar. 2016. « Biophysics in Drug Discovery: Impact, Challenges and Opportunities ». *Nature Reviews Drug Discovery* 15 (10): 679-98. <https://doi.org/10.1038/nrd.2016.123>.
- Reynolds, Pamela, Jennifer A. Anderson, Jane V. Harper, Mark A. Hill, Stanley W. Botchway, Anthony W. Parker, et Peter O'Neill. 2012. « The Dynamics of Ku70/80 and DNA-PKcs at DSBs Induced by Ionizing Radiation Is Dependent on the Complexity of Damage ». *Nucleic Acids Research* 40 (21): 10821-31. <https://doi.org/10.1093/nar/gks879>.
- Riballo, Enriqueta, Lisa Woodbine, Thomas Stiff, Sarah A. Walker, Aaron A. Goodarzi, et Penny A. Jeggo. 2009. « XLF-Cernunnos Promotes DNA Ligase IV–XRCC4 Re-Adenylation Following Ligation ». *Nucleic Acids Research* 37 (2): 482-92. <https://doi.org/10.1093/nar/gkn957>.
- Roos, Wynand P., et Bernd Kaina. 2006. « DNA Damage-Induced Cell Death by Apoptosis ». *Trends in Molecular Medicine* 12 (9): 440-50. <https://doi.org/10.1016/j.molmed.2006.07.007>.
- Ropars, V., P. Drevet, P. Legrand, S. Baconnais, J. Amram, G. Faure, J. A. Marquez, et al. 2011. « Structural Characterization of Filaments Formed by Human Xrcc4-Cernunnos/XLF Complex Involved in Nonhomologous DNA End-Joining ». *Proceedings of the National Academy of Sciences* 108 (31): 12663-68. <https://doi.org/10.1073/pnas.1100758108>.

- Schultes, Sabine, Chris de Graaf, Eric E.J. Haaksma, Iwan J.P. de Esch, Rob Leurs, et Oliver Krämer. 2010. « Ligand Efficiency as a Guide in Fragment Hit Selection and Optimization ». *Drug Discovery Today: Technologies* 7 (3): e157-62. <https://doi.org/10.1016/j.ddtec.2010.11.003>.
- Schultz, Peter G, et Peter B Dervan. s. d. « Sequence-Specific Double-Strand Cleavage of DNA by Penta-N- Methylpyrrolocarboxamide-EDTA*Fe(II) », 4.
- Scott, Duncan E., Anthony G. Coyne, Sean A. Hudson, et Chris Abell. 2012. « Fragment-Based Approaches in Drug Discovery and Chemical Biology ». *Biochemistry* 51 (25): 4990-5003. <https://doi.org/10.1021/bi3005126>.
- Seidel, Susanne A. I., Christoph J. Wienken, Sandra Geissler, Moran Jerabek-Willemsen, Stefan Duhr, Alwin Reiter, Dirk Trauner, Dieter Braun, et Philipp Baaske. 2012. « Label-Free Microscale Thermophoresis Discriminates Sites and Affinity of Protein-Ligand Binding ». *Angewandte Chemie International Edition* 51 (42): 10656-59. <https://doi.org/10.1002/anie.201204268>.
- Seidel, Susanne A.I., Patricia M. Dijkman, Wendy A. Lea, Geert van den Bogaart, Moran Jerabek-Willemsen, Ana Lazic, Jeremiah S. Joseph, et al. 2013. « Microscale Thermophoresis Quantifies Biomolecular Interactions under Previously Challenging Conditions ». *Methods* 59 (3): 301-15. <https://doi.org/10.1016/j.ymeth.2012.12.005>.
- Shirodkar, Purnata, Amanda L. Fenton, Li Meng, et C. Anne Koch. 2013. « Identification and Functional Characterization of a Ku-Binding Motif in Aprataxin Polynucleotide Kinase/Phosphatase-like Factor (APLF) ». *Journal of Biological Chemistry* 288 (27): 19604-13. <https://doi.org/10.1074/jbc.M112.440388>.
- Shrivastav, Meena, Leyma P De Haro, et Jac A Nickoloff. 2008. « Regulation of DNA Double-Strand Break Repair Pathway Choice ». *Cell Research* 18 (1): 134-47. <https://doi.org/10.1038/cr.2007.111>.
- Shuker, Suzanne B, Philip J Hajduk, Robert P Meadows, et Stephen W Fesik. 1996. « Discovering High-Affinity Ligands for Proteins »: 274: 5.
- Squires, Todd M, Robert J Messinger, et Scott R Manalis. 2008. « Making It Stick: Convection, Reaction and Diffusion in Surface-Based Biosensors ». *Nature Biotechnology* 26 (4): 417-26. <https://doi.org/10.1038/nbt1388>.
- Srivastava, Mrinal, Mridula Nambiar, Sheetal Sharma, Subhas S. Karki, G. Goldsmith, Mahesh Hegde, Sujeet Kumar, et al. 2012. « An Inhibitor of Nonhomologous End-Joining Abrogates Double-Strand Break Repair and Impedes Cancer Progression ». *Cell* 151 (7): 1474-87. <https://doi.org/10.1016/j.cell.2012.11.054>.
- Tadi, Satish K., Carine Tellier-Lebègue, Clément Nemoz, Pascal Drevet, Stéphane Audebert, Sunetra Roy, Katheryn Meek, Jean-Baptiste Charbonnier, et Mauro Modesti. 2016. « PAXX Is an Accessory C-NHEJ Factor That Associates with Ku70 and Has Overlapping Functions with XLF ». *Cell Reports* 17 (2): 541-55. <https://doi.org/10.1016/j.celrep.2016.09.026>.

- Tsai, J., J. T. Lee, W. Wang, J. Zhang, H. Cho, S. Mamo, R. Bremer, et al. 2008. « Discovery of a Selective Inhibitor of Oncogenic B-Raf Kinase with Potent Antimelanoma Activity ». *Proceedings of the National Academy of Sciences* 105 (8): 3041-46. <https://doi.org/10.1073/pnas.0711741105>.
- Valkov, Eugene, Tim Sharpe, May Marsh, Sandra Greive, et Marko Hyvönen. 2012. « Targeting Protein–Protein Interactions and Fragment-Based Drug Discovery ». In *Fragment-Based Drug Discovery and X-Ray Crystallography*, édité par Thomas G. Davies et Marko Hyvönen, 145-79. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/128_2011_265.
- Walker, John R., Richard A. Corpina, et Jonathan Goldberg. 2001. « Structure of the Ku Heterodimer Bound to DNA and Its Implications for Double-Strand Break Repair ». *Nature* 412 (6847): 607-14. <https://doi.org/10.1038/35088000>.
- Wall, Monroe E., M. C. Wani, C. E. Cook, Keith H. Palmer, A. T. McPhail, et G. A. Sim. 1966. « Plant Antitumor Agents. I. The Isolation and Structure of Camptothecin, a Novel Alkaloidal Leukemia and Tumor Inhibitor from *Camptotheca Acuminata*^{1,2} ». *Journal of the American Chemical Society* 88 (16): 3888-90. <https://doi.org/10.1021/ja00968a057>.
- Watson, J. D., et F. H. C. Crick. 1953. « THE STRUCTURE OF DNA ». *Cold Spring Harbor Symposia on Quantitative Biology* 18 (0): 123-31. <https://doi.org/10.1101/SQB.1953.018.01.020>.
- Weterings, Eric, et Dik C. van Gent. 2004. « The Mechanism of Non-Homologous End-Joining: A Synopsis of Synapsis ». *DNA Repair* 3 (11): 1425-35. <https://doi.org/10.1016/j.dnarep.2004.06.003>.
- Wyss, Daniel F., Yu-Sen Wang, Hugh L. Eaton, Corey Strickland, Johannes H. Voigt, Zhaoning Zhu, et Andrew W. Stamford. 2012. « Combining NMR and X-ray Crystallography in Fragment-Based Drug Discovery: Discovery of Highly Potent and Selective BACE-1 Inhibitors ». In *Fragment-Based Drug Discovery and X-Ray Crystallography*, édité par Thomas G. Davies et Marko Hyvönen, 83-114. Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/128_2011_183.
- Yano, Ken-ichi, Keiko Morotomi-Yano, Kyung-Jong Lee, et David J. Chen. 2011. « Functional Significance of the Interaction with Ku in DNA Double-Strand Break Recognition of XLF ». *FEBS Letters* 585 (6): 841-46. <https://doi.org/10.1016/j.febslet.2011.02.020>.
- Yin, Xiaotong, Mengjie Liu, Yuan Tian, Jiawei Wang, et Yanhui Xu. 2017. « Cryo-EM Structure of Human DNA-PK Holoenzyme ». *Cell Research* 27 (11): 1341-50. <https://doi.org/10.1038/cr.2017.110>.
- Zhang, Ziming, Weidong Hu, Leticia Cano, Terry D Lee, David J Chen, et Yuan Chen. 2004. « Solution Structure of the C-Terminal Domain of Ku80 Suggests Important Sites for Protein-Protein Interactions ». *Structure* 12 (3): 495-502. <https://doi.org/10.1016/j.str.2004.02.007>.

Titre : Étude d'interactions protéine-ADN : interactions Ku70/Ku80-ADN au sein de la voie NHEJ et caractérisation d'inhibiteurs d'une interaction protéine-ADN cible dans un contexte de recherche de médicaments

Mots clés : NHEJ, Ku70/80, switchSENSE, biophysique, cristallographie

Pour qu'un organisme soit viable, il est essentiel que la molécule d'ADN soit dupliquée et transmise aux générations filles dans son intégralité avec le moins d'altérations possibles. Des cassures double-brin (CDB) peuvent apparaître suite à des stress endogènes ou exogènes et constituent les dommages de l'ADN les plus toxiques. La radiothérapie a pour but générer de manière localisée des CDBs dans les cellules tumorales et de déclencher leur apoptose. Un enjeu majeur actuel est de comprendre le mécanisme de résistance de certaines tumeurs à ces CDBs et de caractériser de nouveaux inhibiteurs du NHEJ pour potentialiser l'effet de la radiothérapie. La voie NHEJ débute par la reconnaissance des deux extrémités de l'ADN d'une CDB par une protéine hétérodimérique Ku70/80 (Ku). Celle-ci sert de plateforme de recrutement aux autres facteurs de la voie NHEJ impliqués dans la maturation (*processing*) et la ligation des extrémités de l'ADN. Le premier volet de cette thèse a porté sur la caractérisation de sites d'interactions protéine-protéine de Ku et sur la caractérisation de nouvelles molécules comme médicaments potentiels dirigés contre Ku. Les sites d'interactions de deux partenaires majeurs de Ku, XLF et APLF, ont été caractérisés par calorimétrie en utilisant des peptides contenant le motif d'interaction de ces protéines appelés KBM (pour Ku-Binding Motifs) ou les protéines elles-mêmes. La structure cristallographique d'un complexe Ku-ADN avec une petite molécule, l'IP6, a été résolue. L'IP6 présente une forte affinité pour Ku avec un K_D de 50 nM. Nous avons montré que cette molécule stimule la fixation de Ku sur l'ADN par trois méthodes biophysiques (calorimétrie, switchSENSE et sm-TIRF). Enfin, un criblage haut débit d'une chimiothèque de 987 molécules fragments de Sanofi a été réalisé *in vitro* par DSF et validé par RMN sur la protéine Ku. Ce crible a permis d'identifier 15 fragments susceptibles d'interagir avec Ku. Ces travaux ouvrent la voie pour la conception rationnelle d'inhibiteurs dérivés des peptides pXLF, pAPLF, de la molécule IP6 ou des fragments identifiés par le crible en greffant sur ces plateformes de nouvelles fonctions chimiques pour améliorer encore l'affinité et bloquer la voie NHEJ. Le second volet de cette thèse a porté sur des interactions de type protéine-ADN étudiées par Sanofi dans un projet de découverte de nouvelles drogues en oncologie et l'analyse de la valeur ajoutée d'une nouvelle approche le switchSENSE implémentée au CEA de Saclay. Nous avons analysé 4 constructions d'une protéine interagissant avec deux ADN et l'effet de 15 molécules inhibitrices identifiées en amont dans le service « Integrated Drug Discovery » de Sanofi. Les expériences de switchSENSE ont permis de caractériser les K_D des 4 constructions pour l'ADN et de mesurer des constantes d'inhibition pour les inhibiteurs. Ces mesures sont en bon accord avec les mesures effectuées à Sanofi par SPR, MST et RMN et permettent de positionner le switchSENSE comme une approche prometteuse dans la validation de nouveaux inhibiteurs en particulier pour les interactions protéine-ADN.

Titre : Protein-DNA interaction study: Ku70 / Ku80-DNA interactions within the NHEJ pathway and characterization of inhibitors of a target protein-DNA interaction in a drug research context

Key words : NHEJ, Ku70/80, switchSENSE, biophysics, crystallography

The maintenance of DNA integrity is central in all organisms to ensure successful life and reproduction. The DNA should be duplicated and transmitted to next generation cells in its entirety with a minimal amount of alteration. Double-strand breaks (DSBs) are the most toxic DNA damages and can originate from endogenous or exogenous stresses. In anticancer treatments, radiotherapy aims at generating DSBs locally on tumor cells and thus induce their apoptosis. current major challenge nowadays is to understand the mechanism of resistance of some tumors to these treatments and to characterize new inhibitors of NHEJ to potentiate the effect of radiotherapy. The NHEJ pathway begins with the recognition of CDBs extremities by a Ku70/80 (Ku) heterodimeric protein. Ku serves as a recruitment platform for the other factors of the NHEJ pathway involved in the processing and ligation of the DNA ends. The first part of this thesis focuses on the characterization of Ku-protein interaction sites and on the characterization of new molecules as potential drugs against Ku. The interaction sites of two major Ku partners, XLF and APLF, were characterized by calorimetry using peptides containing the interaction motif of these proteins, called KBM (for Ku-Binding Motifs) or the proteins themselves. The crystallographic structure of a Ku-DNA complex with a small molecule, IP6, has been determined. The IP6 has a strong affinity for Ku with a K_D of 50 nM. We showed that this molecule stimulates Ku binding on DNA by three biophysical methods (calorimetry, switchSENSE and sm-TIRF). Finally, a high throughput screening of library of Sanofi fragments of 987 molecules was performed *in vitro* by DSF and validated by NMR on the Ku protein and identified 15 fragments that could interact with Ku. This work paves the way for the rational design of inhibitors derived from the pXLF peptides, pAPLF, IP6 molecules or fragments by grafting on these platforms new chemical functions to further improve the affinity and inhibit the NHEJ pathway. The second part of this thesis focused on protein-DNA interactions studied by Sanofi in a discovery drugs project in oncology and the analysis of the added value of a new approach switchSENSE implemented at CEA Saclay. We analyzed 4 constructs of a protein interacting with two DNAs and the effect of 15 inhibitory molecules identified upstream in Sanofi's « Integrated Drug Discovery » service. The switchSENSE experiments made it possible to characterize the K_D of the 4 constructs for DNA and to measure inhibition constants for the inhibitors. These measurements are in good agreement with the measurements made at Sanofi by SPR, MST and NMR. This study positions the switchSENSE as a promising approach in the validation of new inhibitors, in particular for protein-DNA interactions.

Université Paris-Saclay

Espace Technologique / Immeuble Discovery

Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France