



HAL
open science

Outils moléculaires pour l'étude des G-quadruplex au sein du génome

Joël Lefebvre

► **To cite this version:**

Joël Lefebvre. Outils moléculaires pour l'étude des G-quadruplex au sein du génome. Chimie organique. Université Paris Saclay (COMUE), 2017. Français. NNT : 2017SACLS536 . tel-01693928

HAL Id: tel-01693928

<https://theses.hal.science/tel-01693928>

Submitted on 26 Jan 2018

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.

Outils moléculaires pour l'étude des G-quadruplex au sein du génom

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'Université Paris Sud

École doctorale n°571 Sciences chimiques : molécules, matériaux,
instrumentation et biosystèmes (2MIB)
Spécialité de doctorat: Chimie

Thèse présentée et soutenue à Orsay, le 15 décembre 2017, par

Joël Lefebvre

Composition du Jury :

Jean-François Riou Professeur, MNHN, SIG (UMR 7196 - U 1154)	Président du jury
Geneviève Pratviel Directrice de recherche, CNRS, LCC (UPR 8241)	Rapporteur
Thomas Lavergne Chargé de recherche, CNRS, I2BM (UMR CNRS-UGA 5250)	Rapporteur
Joelle Dubois Directrice de recherche, CNRS, ICSN (UPR 2301)	Examinatrice
Marie-Paule Teulade-Fichou Directrice de recherche, CNRS, CMIB (UMR 9187 - U 1196)	Directrice de thèse
Florence Mahuteau-Betzer Chargée de recherche, CNRS, CMIB (UMR 9187 - U 1196)	Directrice de thèse

Remerciements

Pour commencer ce manuscrit, je souhaiterais remercier toutes les personnes qui m'ont permis de mener à bien ce travail de doctorat.

Tout d'abord, **Marie-Paule Teulade-Fichou**, directrice de l'unité, avec qui les échanges scientifiques sont toujours très enrichissants et **Florence Mahuteau-Betzer**, qui m'a suivi et soutenu au quotidien aussi bien dans les bons moments que dans ceux un peu plus délicats. Mes deux co-directrices de thèse m'ont apporté toute leur confiance pendant ces trois années et j'espère leur avoir rendu comme elles s'y attendaient. La bonne humeur quotidienne de **Florence** est vraiment très agréable et communicative.

Je souhaiterais également remercier **Sophie Bombard** et **Florent Poyer** qui, chacun dans leur domaine, m'ont formé et accompagné tout au long des expériences que j'ai effectué avec eux. **Sophie** pour toutes les manipulations de radioactivité et de gels d'électrophorèse et **Florent** pour toute sa formation sur la culture cellulaire et l'imagerie.

Je remercie **Geneviève Pratviel** et **Thomas Lavergne** d'avoir accepté d'être rapporteurs et **Jean-François Riou** et **Joelle Dubois** d'avoir accepté d'être examinateurs de ce travail. Merci de prendre de votre temps pour évaluer mes travaux.

Au cours de ces trois années passées au centre de recherche de l'Institut Curie, j'ai rencontré de très belles personnes que je souhaite également remercier et tout particulièrement tous les membres de l'unité **UMR9187 – U1196** dans laquelle j'ai travaillé :

- la team G4 avec laquelle les discussions scientifiques ont été très enrichissantes et m'ont permis de mener à bien ce travail. Dans cette équipe, je tiens à remercier **Elodie** qui m'a précédé dans le rôle de doctorant, qui m'a formé à toutes les techniques de biophysique et m'a permis une très bonne intégration au sein de ce groupe, **Daniela** avec qui nous essayons de valoriser les données de biophysique, **Claire** qui a réalisé un gros travail pour étoffer le nombre de composés et de séquences étudiés par les méthodes de biophysique et avec qui on peut également discuter facilement, **Corinne** qui a réalisé un gros travail de synthèse et **Abhijit** qui n'a pas le même projet mais avec qui les échanges sont toujours très intéressants. J'ajoute également **Anton** avec qui nous avons mis en place l'expérience de FRET en plaque et qui est toujours de bon conseil pour des questions scientifiques. Je souhaite la bienvenue au nouvel arrivant, **Michael**, qui sera très bien entouré pour réaliser de belles choses en trois ans.

- la team kinase composée de **Sandrine**, la chef de cette équipe, **Tom** avec qui j'ai partagé le « grand » bureau pendant deux ans et avec qui j'ai passé beaucoup de bons moments, **Jessy** qui est une fille super gentille avec qui on peut discuter aisément, **Marius** le chanteur et « râleur » du labo qui sait mettre l'ambiance et enfin **Julie** qui vient de rejoindre cette équipe pour une thèse après 6 premiers mois de stage.

- la team Abivax composée de **Pauline**, la sportive, avec qui on peut discuter aussi bien de sciences que de tout et de rien et de **Romain** avec qui les discussions sont très enrichissantes et qui donne de très bons conseils.

- la team « sondes fluorescentes » composée de **Delphine** qui est toujours une oreille attentive et qui a toujours le sourire. Ce fut fort agréable d'être son voisin de bureau pendant 6 mois. Il y a également **Laura** que je souhaite remercier pour sa gentillesse et sa bonne humeur.

- la team LC-MS qui a eu beaucoup de travail pendant ces trois ans et ce n'est pas fini... composée de **Pauline, Claire, Romain et Delphine**.

- la team « Anton » composée de **Coralie** à qui je souhaite bon courage pour les deux ans de thèse qu'il lui reste.

- la team biologie composée de **Frédérique** avec qui j'ai eu des discussions très intéressantes dans le L1, **Florent** et **Sophie** de nouveau, **Deep** et **Linda**.

- la team « gestionnaire » composée de **Nathalie** et **Sandrine** qui m'ont rendu les choses plus faciles pour les côtés administratifs.

J'espère n'avoir oublié personne, je tiens également à remercier **Sergio** de m'avoir appris à utiliser image J de manière plus approfondie, **Marie-Noëlle** de m'avoir formé à la microscopie confocale et tous les autres membres de l'Institut Curie qui ont rendu agréable mon passage ici.

Un second grand merci à **Romain** qui, en plus d'être un super collègue, m'a intégré facilement à l'équipe de foot des bleus. De très bons moments avec toute l'équipe les mercredis midis qui permettaient de faire une coupure au milieu de la semaine. De plus, au laboratoire, c'est la seule personne avec qui je pouvais partager ma passion du football pour discuter des matchs de la veille ou à venir.

Ensuite, je remercie fortement toute ma famille. **Mes parents** qui me suivent et m'encouragent depuis toujours. Mon frère, **Jimmy**, ma sœur, **Françoise**, ma belle-sœur, **Coralie**, mon beau-frère, **Gregory**, mes neveux, **Alexis, Maxence et Théo**, et mes nièces, **Clara et Ophélie**, qui, à chaque fois

que je les vois, me donnent le sourire et me permettent d'être encore plus fort. Je tiens également à remercier mes beaux-parents, **Eric** et **Catherine**, qui me soutiennent à chaque fois que je les vois.

Enfin, la dernière personne que je souhaite remercier chaleureusement est la personne qui me supporte au quotidien, **Alexia**. Merci d'être là pour moi, de me motiver, m'encourager et me soutenir en toute circonstance.

Sans vous tous, ce travail de thèse n'aurait pas été possible. Vraiment MERCI !

Table des matières

Abréviations	8
Introduction.....	9
I. Structures des acides nucléiques.....	10
1) Structure primaire	10
2) Structures secondaires	11
3) Structure <i>in cellulo</i>	15
II. Les G-quadruplex.....	16
1) Structure, topologie et orientation	16
2) Localisation et rôles dans le génome	19
a) Les télomères	19
b) Les promoteurs de gènes	21
c) Les minisatellites	22
III. Ligands de référence de G-quadruplex	23
1) Interaction avec les G-quartets	24
2) Interaction avec le sillon.....	29
3) Interaction avec le canal ionique	30
4) Les complexes métalliques.....	30
5) Cas particuliers : les ligands fluorescents.....	33
IV. Méthodes d'analyses des interactions ligand/G-quadruplex.....	34
1) Analyses semi-quantitatives.....	35
a) FRET-melting.....	35
b) Fluorescence Intercalator Displacement (FID)	37
c) Dichroïsme circulaire	40
2) Analyses quantitatives.....	42
a) Résonance plasmon de surface	42
b) Titration calorimétrique isotherme.....	43
c) Résonance magnétique nucléaire	45
d) Spectrométrie de masse.....	46
e) Diffraction de rayons X	46
3) Analyses biochimiques	47
a) Inhibition de polymérase	47
b) Influence des hélicases sur le déroulement des G-quadruplex	47

Objectifs du travail de doctorat	49
Chapitre 1 : Synthèse de ligands de G-quadruplex « clickables »	52
I. La chimie « click » comme stratégie de ciblage.....	53
1) Principe.....	53
2) Cycloaddition d'un azoture et d'un alcyne catalysée au cuivre (CuAAC).....	54
a) Mécanisme de réaction.....	55
b) Utilisation de chélatants du cuivre.....	57
3) Cycloaddition d'un azoture et d'un alcyne contraint (SPAAC)	58
4) Autres réactions de chimie « click » sans cuivre	60
II. Synthèse de dérivés du PhenDC ₃	61
1) Précurseur commun à tous les dérivés du PhenDC ₃	62
2) Précurseur de ligands bi-fonctionnels.....	65
III. Synthèse de dérivés du PDC	69
Chapitre 2 : Comparaison par analyses biophysiques <i>in vitro</i> de ligands de G-quadruplex	72
I. Expérience de FRET-melting	73
1) Séquences d'ADN utilisées	73
2) Résultats	76
a) Ligands de référence	76
b) Dérivés du PhenDC ₃	80
c) Dérivés du PDC	88
3) Conclusions.....	92
II. Expérience de FID	93
1) Séquences d'ADN utilisées	93
2) Résultats	96
a) Ligands de référence	96
b) Dérivés du PhenDC ₃	103
c) Dérivés du PDC	107
3) Conclusions.....	111
III. Comparaison de ligands en combinant les deux méthodes.....	111
1) Ligands de références.....	112
2) Dérivés du PhenDC ₃ et du PDC.....	118
IV. Corrélation entre l'expérience de FRET-melting et l'instabilité génétique en levures	123
1) Minisatellite CEB25.....	124
2) Minisatellite CEB1.....	128

3) Conclusion	130
Chapitre 3 : Visualisation de ligands de G-quadruplex en cellules de mammifère.....	132
I. Les G-quadruplex observés <i>in vivo</i>	133
1) Imagerie cellulaire	133
a) Immunofluorescence.....	133
b) Etudes de la localisation de ligands de G-quadruplex.....	136
2) Etude de l'instabilité génétique chez la levure	141
3) Séquençage d'ADN	143
II. La chimie bioorthogonale	147
1) Principe.....	147
2) Rappel des réactions de chimie « click » utilisées <i>in cellulo</i>	148
III. La magie de la chimie « click » en cellules.....	149
1) Etude <i>in vitro</i>	149
2) Etude en cellules fixées	156
3) Etude en cellules vivantes	171
IV. Analyse de temps de vie de fluorescence	177
Conclusion générale et perspectives.....	182
Chapitre 4 : Partie expérimentale	187
I. Synthèses.....	188
1) Matériel et méthodes.....	188
2) Modes opératoires	188
II. Expériences de biophysique	196
1) Oligonucléotides.....	196
2) Expériences de FRET-melting	197
3) Expériences de FID	197
4) Dichroïsme circulaire	198
5) Temps de vie de fluorescence	198
III. Culture cellulaire et imagerie cellulaire.....	198
1) Culture cellulaire	198
2) Protocole en cellules fixées avec incubation en cellules vivantes ou fixées	198
3) Protocole en cellules vivantes	199
4) Imagerie cellulaire	199
IV. Expériences d'extraction supportée par billes magnétiques	200
1) Principe de l'extraction supportée par billes magnétiques recouvertes de streptavidine	200

2) Mode opératoire	201
a) Préparation des oligonucléotides.....	201
b) Réaction avec les billes magnétiques recouvertes de streptavidine	201
c) Réaction avec les billes magnétiques recouvertes de fonction -DBCO.....	203

Abréviations

A :	adénine
ADN :	acide désoxyribonucléique
ARN :	acide ribonucléique
C :	cytosine
DBCO :	dibenzoazacyclooctyne
DCC :	<i>N,N'</i> -dicyclohexylcarbodiimide
DCM :	dichlorométhane
DMF :	<i>N,N</i> -diméthylformamide
EDCi :	1-éthyl-3-(3-diméthylaminopropyl)carbodiimide
FID :	<i>fluorescence intercalator displacement</i> , déplacement d'une sonde fluorescente intercalée dans une biomolécule
FRET :	<i>Förster resonance energy transfer</i> , transfert d'énergie par résonance d'effet Förster
G :	guanine
G4 :	G-quadruplex
HOBt :	hydroxybenzotriazole
HPLC :	<i>high performance liquid chromatography</i> , chromatographie liquide haute performance
HRMS :	spectrométrie de masse haute résolution
kb :	kilo-bases, soit 10^3 bases d'acides nucléiques
LC-MS :	chromatographie liquide couplée à la spectrométrie de masse
MS :	spectrométrie de masse
pb :	paire de base, unité de longueur d'un brin d'acides nucléiques
PBS :	<i>phosphate buffered saline</i> , tampon phosphate salin
PCR :	<i>polymerase chain reaction</i> , réaction de polymérisation en chaîne
PDC :	pyridinedicarboxamide, nommé aussi 306A
PEG :	polyéthylène glycol
PhenDC :	phenanthrolinedicarboxamide
RMN :	spectroscopie par résonance magnétique nucléaire
T :	thymine
tpy :	terpyridine
tppy :	tolylterpyridine
U :	Uracile
WT :	<i>wild type</i> , forme native d'une séquence d'acides nucléiques

Introduction

I. Structures des acides nucléiques

Les acides nucléiques existent sous deux différentes structures. La première est l'acide désoxyribonucléique (ADN) qui sert principalement de support à l'information génétique et la seconde est l'acide ribonucléique (ARN) qui joue plusieurs rôles dont celui de messenger de l'information génétique. En effet, l'ARN messenger permet de transmettre l'information d'un gène à l'extérieur du noyau de la cellule (pour les eucaryotes) où seront synthétisées les protéines correspondantes. L'ARN ribosomique ou guide servent quant à eux de catalyseur ou de vecteur pour les enzymes ce qui permet aux protéines d'atteindre des séquences spécifiques de l'ADN. L'ARN ribosomique est également un régulateur de l'expression génétique en bloquant par exemple le mécanisme de traduction par le ribosome. Toute cellule vivante contient ces deux types d'acides nucléiques contrairement au virus qui ne contiennent que l'un des deux.

1) Structure primaire

L'ADN et l'ARN sont constitués d'un enchainement de nucléotides liés entre eux par des liaisons phosphodiester. Un nucléotide est quant à lui formé d'un sucre (ose à cinq carbones), d'un phosphate et d'une base azotée (figure 1). Le brin d'acide nucléique correspond à un polymère réticulé avec une alternance de sucre et de phosphate qui forme le squelette et les bases azotées greffées sont les réticulations.

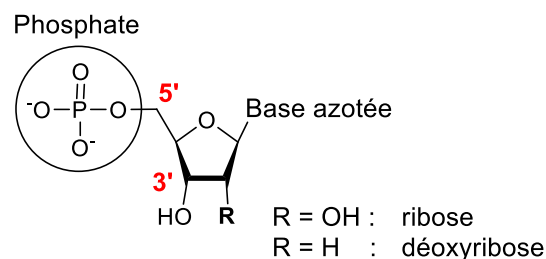


Figure 1 : Structure d'un nucléotide.

Il existe cinq bases azotées principales dans l'ADN et l'ARN (figure 2). Celles-ci constituent l'alphabet du patrimoine génétique. Il s'agit de l'adénine (A), la thymine (T), la cytosine (C) et la guanine (G) pour l'ADN et l'uracile (U) remplace la thymine (T) sur les brins d'ARN. La structure primaire des acides nucléiques correspond donc à l'assemblage de nucléotides. L'originalité du brin se fait uniquement par la succession de bases présentes sur le squelette glycoposphate aussi appelée séquence d'ADN ou séquence d'ARN. Ces séquences représentent un message codé portant les informations génétiques de l'être vivant.

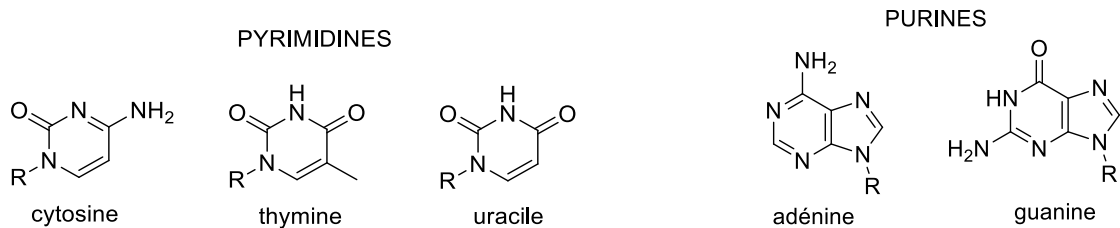


Figure 2 : Structures des bases azotées usuelles : cytosine, thymine, uracile, adénine et guanine.

2) Structures secondaires

Les bases azotées précédemment présentées peuvent s'auto-associer par des liaisons non-covalentes. En effet, une adénine et une thymine peuvent s'associer en formant deux liaisons hydrogène entre elles. De même, une cytosine et une guanine peuvent, elles, former trois liaisons hydrogène (figure 3). Ce réseau de liaisons hydrogène entre deux différentes bases est appelé appariement de Watson-Crick et permet à deux séquences d'acides nucléiques complémentaires de s'associer.

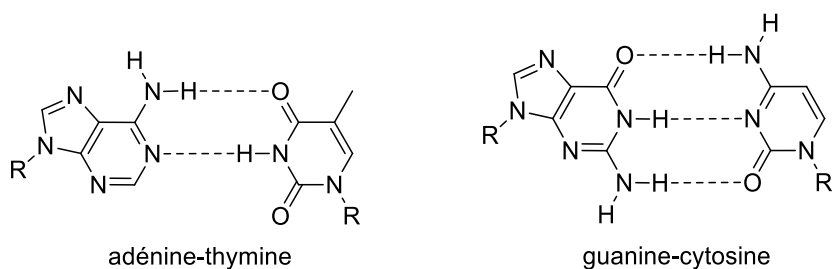


Figure 3 : Appariement de Watson-Crick.

Lorsque deux brins d'ADN complémentaires s'associent par cet appariement, cela forme une hélice appelée hélice double-brin ou ADN duplex. Il s'agit de la structure secondaire d'acide nucléique la plus connue et la plus représentée. Les liaisons hydrogène formées entre une purine (A ou G) et une pyrimidine (C, T ou U) opposée permettent de stabiliser cette structure. Les brins d'acides nucléiques ont une orientation par rapport à la géométrie du sucre et, en général, on note la séquence d'acides nucléiques de 5' vers 3'. Ceux impliqués dans la formation d'une hélice double-brin sont antiparallèles, c'est-à-dire que l'extrémité 5' d'un brin est en contact avec l'extrémité 3' de l'autre. La taille d'une molécule d'acides nucléiques double-brin peut varier de mille paires de bases (unité de taille des acides nucléiques) pour les génomes les plus courts (les virus) à plus de six cents milliards pour le génome le plus long actuellement observé (l'amibe *Polychaos dubium*). Chez l'Homme, le génome est constitué d'environ 3,4 milliards de paires de bases qui codent le fonctionnement de notre organisme et sa complexité.

L'hélice double brin existe sous différentes conformations. La plus classique et surtout la plus formée sous conditions physiologiques est l'hélice B qui consiste en une double-hélice droite effectuant un

tour complet tous les dix nucléotides. Cette conformation a été découverte en 1953 par Watson et Crick^[1]. L'ADN-Z^[2] et l'ADN-A^[3] sont les deux autres structures en hélice formées par l'ADN (figure 4). L'ADN-A est une hélice plus large et plus compacte que l'ADN-B qui a été identifié dans les milieux déshydratés. Il se pourrait que cette conformation se forme pour les doubles hélices hybrides ADN-ARN. L'ADN-Z est une double-hélice gauche dont le squelette présente une conformation en « zigzag ». Cette structure possède deux sillons de même taille. L'ADN-Z se forme à partir de l'ADN-B dans les régions du génome riches en séquences répétées d(GC), pendant la transcription des gènes lorsque la force ionique augmente (au-delà de 2,5 M NaCl par exemple). Cette forme permettrait de réduire les contraintes de super-enroulement au niveau des sites de transcription. L'ADN-Z est une forme transitoire *in vivo* car elle est relaxée en ADN-B en fin de transcription.

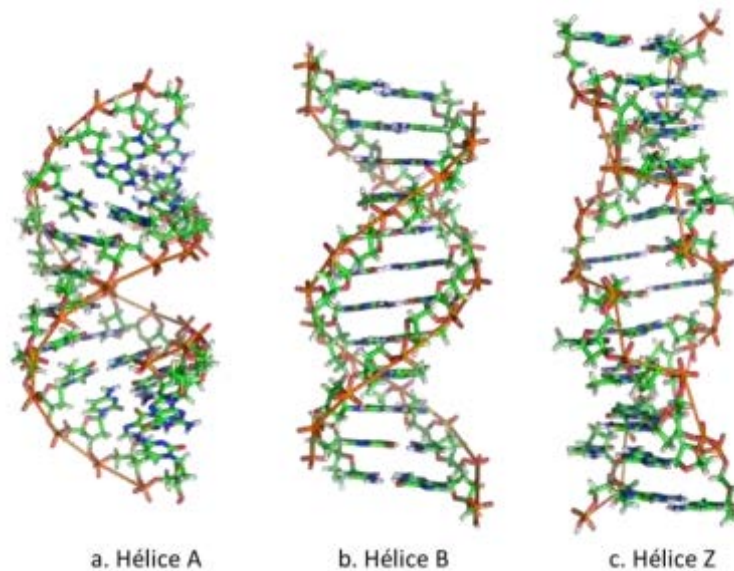


Figure 4 : Représentation des hélices d'ADN a) ADN-A, b) ADN-B et c) ADN-Z

D'autres conformations d'ADN existent différentes de la double hélice. L'ADN cruciforme ou plus connu sous le nom de jonctions de Holliday^[4] (figure 5) qui se forme lors de la recombinaison. Le repliement en épingle à cheveux ou *hairpin*^[5] (figure 5) qui joue un rôle important dans la conformation et la stabilité de l'ARN. Ces deux conformations se forment dans les régions qui possèdent des répétitions en miroir de segments d'acides nucléiques par appariement intra-brin. La jonction en Y^[6] (figure 5) est également une structure secondaire qui se forme entre trois brins d'ADN complémentaires.

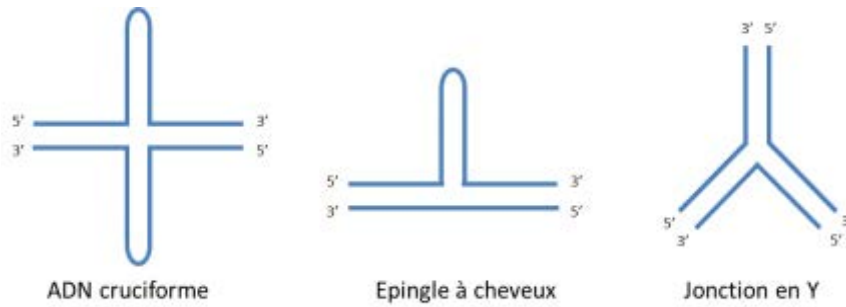


Figure 5 : Représentation schématique des conformations en cruciforme, en épingle à cheveux et jonction en Y.

Une autre structure à trois brins d'ADN existe, il s'agit de l'ADN-h ou triplex^[7] (figure 6a). Un brin riche en base G, A ou T vient enrouler une hélice double-brin B en formant des appariements G.GC, A.AT ou T.AT.^[8] Cette structure triplex a une influence sur la transcription et la régulations de gènes.

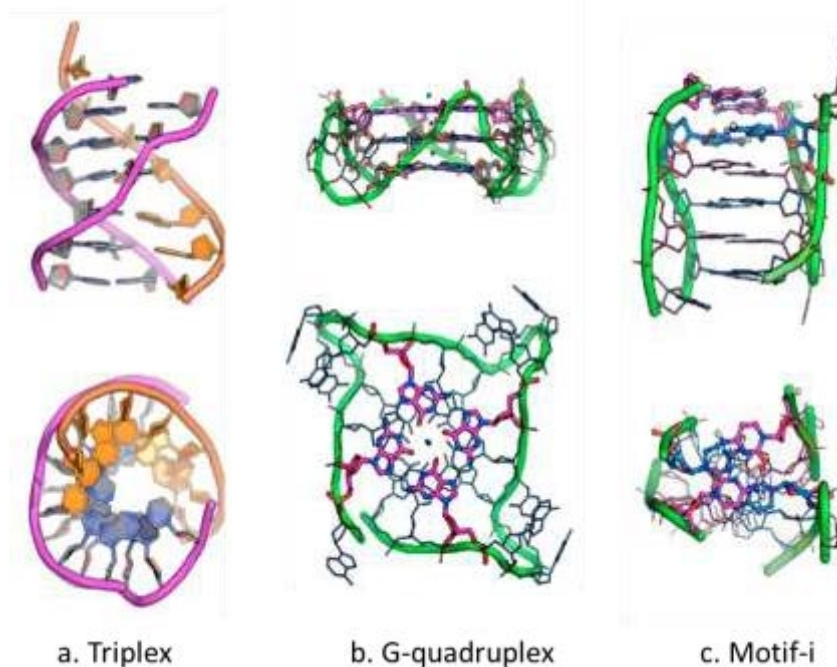


Figure 6: Représentation des structures secondaires a) en triplex, b) en G-quadruplex et c) en i-motif vue de dessus (ligne du haut) et le long de l'axe de leur hélice (ligne du bas)

Dans les zones riches en guanines, le repliement d'un brin d'acides nucléiques autour d'un cation monovalent tel que le potassium (le plus présent en conditions physiologiques) ou le sodium peut permettre la formation d'une structure secondaire d'ADN, le G-quadruplex^[9] (figure 6b). Il s'agit de la structure ciblée dans le cadre de cette thèse. Elle repose sur l'appariement de Hoogsteen qui diverge de l'appariement « classique ». En effet, les guanines entre elles se disposent de façon à former un plan de quatre guanines à l'aide de liaisons hydrogène (figure 7), ce plan est appelé quartet. L'empilement d'au moins deux quartets intercalés par un cation octacoordiné par les huit

guanines adjacentes forme le G-quadruplex. Cette structure est à la fois stable et dynamique au sein du vivant.

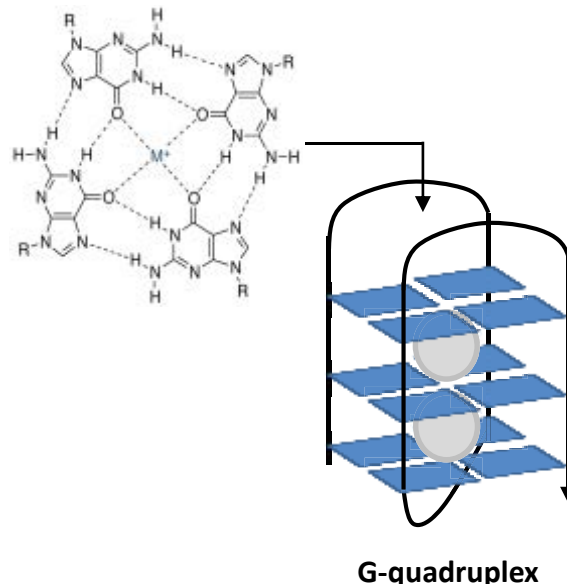


Figure 7: Représentation d'un G-quadruplex

Le brin complémentaire à celui riche en guanines est par conséquent riche en cytosines et est non apparié à la structure G-quadruplex. Ce brin peut également former une structure secondaire particulière : le i-motif (ou ADN-i)^[10] (figure 6c). A pH acide, un appariement entre une cytosine et une cytosine protonée C-C⁺ peut également se former par liaisons hydrogène (figure 8).

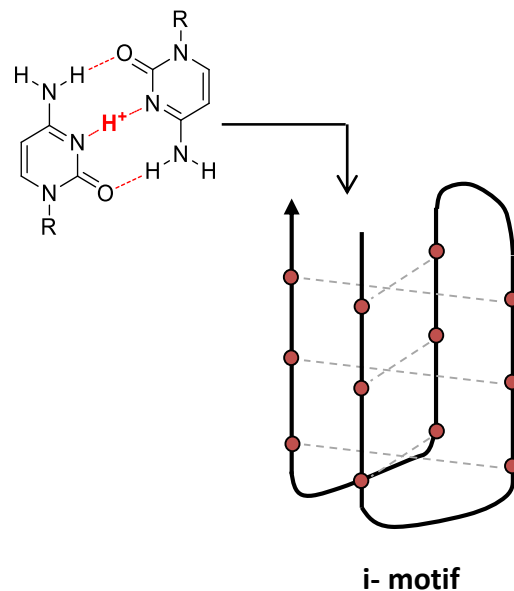


Figure 8: Représentation d'un i-motif

3) Structure *in cellulo*

Les organismes procaryotes, comme les bactéries, ne possèdent pas de noyau cellulaire pour contenir l'ADN. L'information génétique est généralement organisée sous forme de chromosome circulaire unique, naviguant dans le cytoplasme. Pour les organismes eucaryotes, l'ADN se trouve limité majoritairement au noyau de la cellule. Il est associé à de nombreuses protéines, en particulier aux histones, chargées positivement, lui permettant de former la chromatine par interaction électrostatique. Au niveau cellulaire, l'ADN s'arrange de façon à être enroulé sous forme double-brin autour de huit histones pour former l'unité de base de la chromatine, le nucléosome. L'ensemble des nucléosomes s'enroule également pour former une structure encore plus compacte, le chromosome (figure 9).

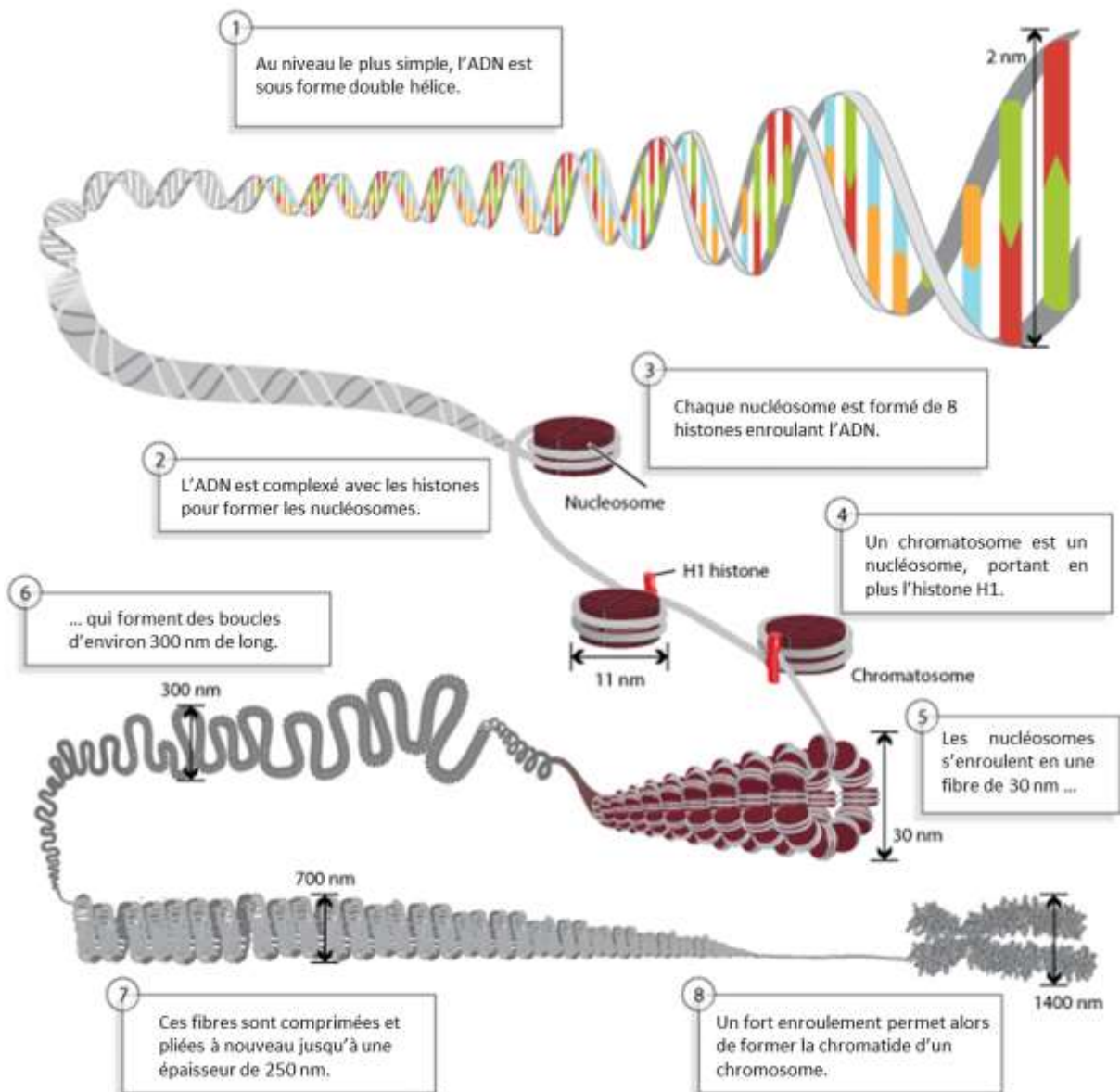


Figure 9 : L'ADN, de la double hélice au chromosome^[11].

L'accès à l'information génétique est ainsi contrôlé et le volume de stockage du génome est condensé dans la chromatine. Cette dernière peut être soit sous forme ouverte et permettre l'accès au mécanisme transcriptionnel, soit sous forme condensée ce qui va empêcher l'expression des gènes. Ces formes sont régulées par des mécanismes d'épigénétique au niveau des histones. L'acétylation des histones provoque l'ouverture de la chromatine et leur méthylation permet de la refermer. L'ARN quant à lui se déplace du noyau au cytoplasme. Après transcription d'un brin d'ADN dans les nucléoles, épissage et traduction, l'ARN peut sortir du noyau par les pores nucléaires et contribuer à la synthèse ribosomale de protéines.

II. Les G-quadruplex

1) Structure, topologie et orientation

Une solution concentrée en acide guanylique est connue pour former un gel depuis plus de cent ans ^[12] mais ce n'est qu'en 1962 que les scientifiques Gellert et Davies ont mis en évidence une structure hélicoïdale libre en solution, la structure G-quadruplex, par diffraction des rayons X d'un gel d'acide 5'-guanylique (5'-GMP).^[13] Ils ont proposé une structure pouvant expliquer les assemblages observés, formée de plans de quatre guanines, appelée G-quartet (figure 10a). Ce dernier est l'unité de base d'une structure G-quadruplex. En effet, pour former cette structure, l'empilement d'au moins deux quartets autour d'un cation monovalent octacoordiné aux huit guanines est nécessaire. Il existe également des G-quadruplex avec une infinité de G-quartets appelés G-wire.^[14] Plus tard, il a été montré que les guanines impliquées peuvent provenir d'un ou plusieurs brins d'ADN différents par repliement de ceux-ci ^[15] (figure 10b).

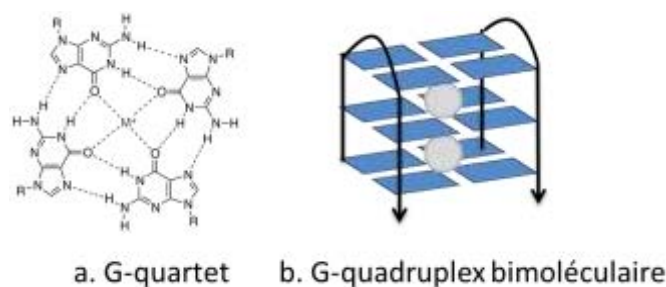


Figure 10 : a) Appariement de Hoogsteen pour former un G-quartet et b) empilement de trois G-quartet pour former un G-quadruplex à partir de deux brins d'ADN.

Il est donc évident que la structure G-quadruplex possède un très grand polymorphisme.^[16] Un à quatre brins d'acides nucléiques peuvent être impliqués dans sa formation, ceci définit la molécularité du G-quadruplex (figure 11).

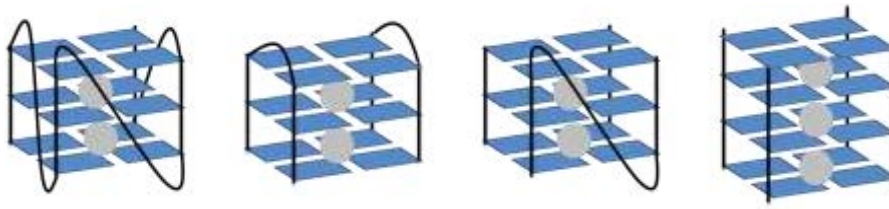


Figure 11 : Représentation schématique de G-quadruplex intramoléculaire, bimoléculaire, trimoléculaire et tétramoléculaire (de gauche à droite).

Les G-quadruplex intramoléculaires sont eux-mêmes polymorphes. En effet, l'orientation du brin d'ADN modifie la conformation (figure 12). Si les quatre piliers sont orientés dans la même direction, il s'agit d'un G-quadruplex parallèle. Lorsque deux piliers sont dans une direction et les deux autres sont dans la direction opposée, il s'agit de G-quadruplex antiparallèles et enfin lorsque trois piliers sont orientés dans un sens et le quatrième est orienté à l'opposé alors il s'agit de G-quadruplex hybride.

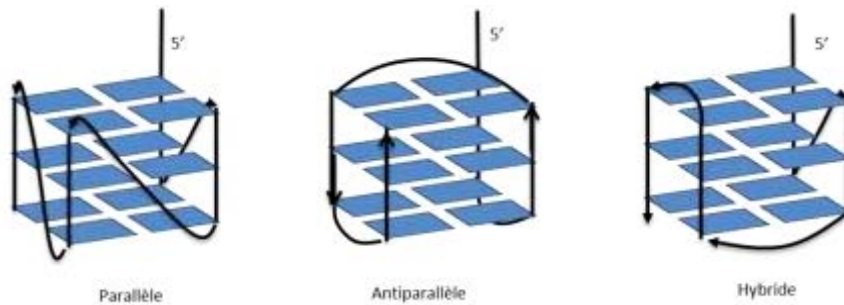


Figure 12 : Représentation schématique de G-quadruplex parallèle, antiparallèle et hybride.

Un autre polymorphisme est induit par l'orientation des guanines : *syn* ou *anti*. Les guanines des quadruplex parallèles sont toutes orientées en *anti* ou en *syn*, ce qui implique que les quatre piliers sont bien parallèles entre eux. Dans les structures antiparallèles et hybrides, les guanines sont en mélange de conformation *syn* et *anti*, ce qui permet qu'au moins un des piliers soit opposé aux autres (figure 13).

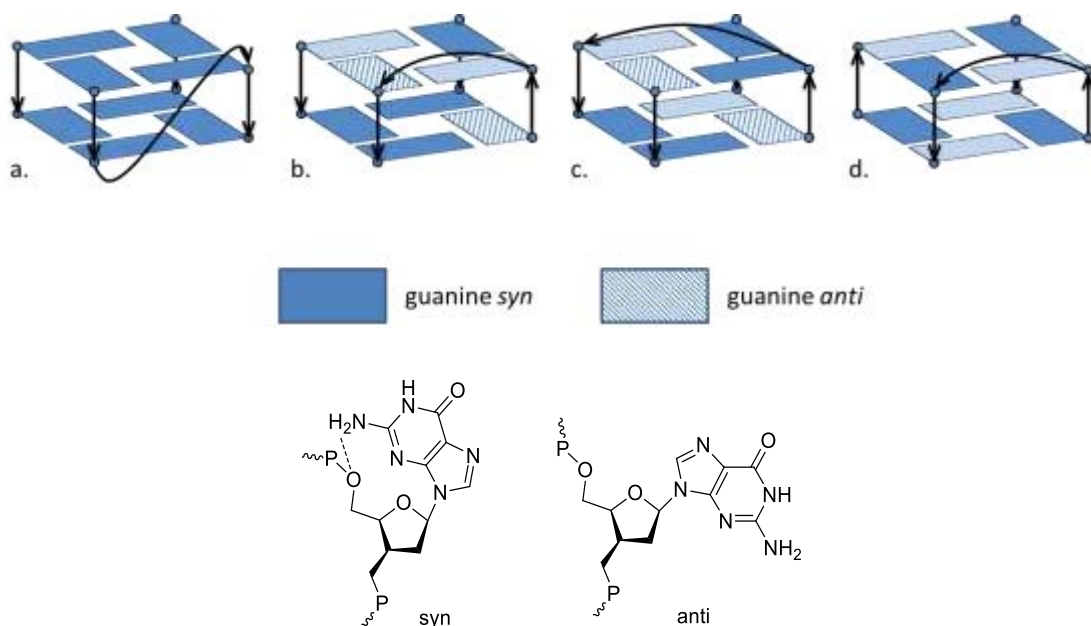


Figure 13 : Orientation *syn* ou *anti* des guanines dans les structures G-quadruplex

Par ailleurs, les boucles formées par les séquences résiduelles reliant deux piliers de la structure peuvent être classées en trois catégories. Les boucles externes (ou en forme d'hélice) connectent deux piliers parallèles adjacents (Figure 13 – a). Les boucles latérales relient des piliers antiparallèles adjacents et sont formées d'en général trois nucléotides ou plus (Figure 13 – b et d). Les boucles diagonales connectent quant à elles deux piliers antiparallèles opposés diagonalement sur un quartet et sont composés de deux nucléotides au minimum (Figure 13 – c). De plus, les boucles sont de tailles et de compositions différentes pour chaque G-quadruplex.

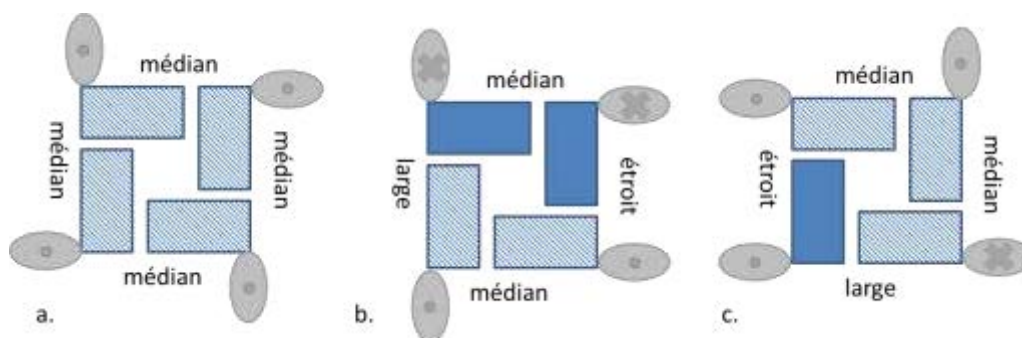


Figure 14 : Représentation schématique de l'orientation de chaque pilier et de la taille des sillons formés. Schéma adapté de [16].

La structure G-quadruplex possède quatre sillons qui ont des caractéristiques structurales différentes des sillons de l'ADN duplex. Trois types de sillons peuvent être formés entre deux guanines adjacentes : deux guanines *syn* ou deux guanines *anti* forment un sillon médian (Figure 14 – a), une guanine *syn* perpendiculaire à une guanine *anti* forme un sillon étroit et une guanine *anti* perpendiculaire à une guanine *syn* forme un sillon large (Figure 14 – b et c).

2) Localisation et rôles dans le génome

Les G-quadruplex sont des structures transitoires qui se forment lorsque l'acide nucléique est sous forme simple brin. C'est pourquoi ils ne peuvent se former que lors de phases très précises du cycle cellulaire. L'ADN est principalement sous forme simple brin lors de la réplication, la transcription et au niveau de l'extrémité des télomères (figure 15).

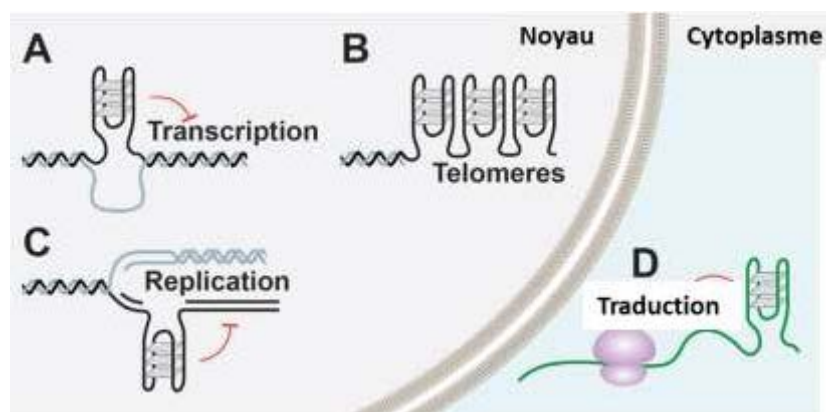


Figure 15 : Exemples de structures G-quadruplex au sein d'une cellule. Schéma adapté de [17].

Tout d'abord, afin de détecter les G-quadruplex potentiels présents dans le génome, différents algorithmes ont été développés. Les deux premiers, datant de 2005, sont le Quadparser [18] selon lequel plus de 300 000 séquences du génome humain sont capable de former des G-quadruplex *in vivo* et l'algorithme utilisé par le groupe de Neidle [19] qui prend en compte des boucles composées de une à sept bases. Enfin, le plus récent est le G4Hunter [20] qui prend en compte la taille des boucles et la richesse en guanine mais aussi en cytosine de la séquence. Les analyses bio-informatiques ont montré que ces séquences de G-quadruplex potentiels (PQS, pour *potential quadruplex sequence*) ne sont pas distribuées aléatoirement dans le génome.[21] Les PQS sont majoritairement présents au niveau des télomères, des régions des séquences de longueur variable répétées en tandem (VNTR pour *variable number tandem repeat* ou minisatellite), des régions codant pour certaines parties des anticorps (chaînes lourdes) et des promoteurs de gènes. Environ 40 % de tous les promoteurs vont contenir au moins une PQS. Par ailleurs, ce type d'analyse a montré que les séquences de G-quadruplex potentiels sont sous-représentées dans les régions codantes, suggérant par exemple un rôle de la structure G-quadruplex dans les mécanismes de transcription plutôt que sur le transcrit lui-même.

a) Les télomères

La première séquence d'ADN à avoir été suspectée de former un G-quadruplex est la séquence télomérique qui possède de nombreuses répétitions de groupe de trois guanines successives

(TTAGGG)_n.^[22] Depuis, il a été montré par des analyses biochimiques et structurales que de nombreuses autres séquences d'ADN ou d'ARN, qui contiennent des répétitions de guanines (G-tract) séparées par d'autres bases, forment spontanément des G-quadruplex *in vitro*.

Les télomères sont des complexes nucléoprotéiques présents à chaque extrémité des chromosomes chez les eucaryotes. Ils ont pour fonction de protéger les chromosomes de la dégradation, de la fusion ou de la recombinaison. Les extrémités télomériques garantissent l'intégrité et la stabilité des chromosomes. Chez les animaux vertébrés, l'ADN télomérique est une séquence double-brin composée du motif TTAGGG, répété en tandem.^[23] Cette séquence est non-codante et sa longueur varie selon les espèces. La longueur moyenne des télomères humains varie de 10 à 15 kilobases (kb). Les extrémités des chromosomes sont donc composées d'un brin particulièrement riche en guanines (le brin G) et d'un brin complémentaire particulièrement riche en cytosine (le brin C). Par ailleurs, une autre caractéristique intéressante des télomères est de posséder une extension simple brin à l'extrémité 3' du brin G, le G-overhang (Figure 16). Cette extension, d'environ 130-210 bases chez l'Homme, est présente sur tous les télomères des eucaryotes et possède un rôle structural important.^[24]

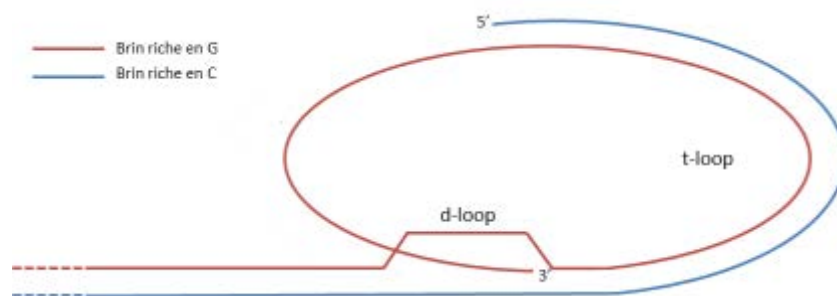


Figure 16 : Représentation schématique de la structure télomérique en lasso.

Il a été montré que la formation de G-quadruplex inhibe l'action de la télomérase *in vitro* et pourrait donc agir comme régulateur négatif de la régulation des télomères.^[25] La télomérase est une enzyme qui permet de conserver la longueur du chromosome pendant la réplication de l'ADN en ajoutant le télomère à chaque extrémité. Cependant, plus récemment, il a été montré que la sous-unité Est1p de la télomérase de levures pouvait promouvoir la formation de G-quadruplex *in vitro*. Or, la séquence télomérique de cellules déficientes en Est1p continue de se raccourcir, ce qui suggère que les G-quadruplex auraient également un rôle de régulateur positif dans le maintien de la longueur des télomères.^[26] Dans les deux cas, la présence de G-quadruplex a une influence sur la durée de vie des cellules.

b) Les promoteurs de gènes

Les séquences promotrices sont des régions de l'ADN situées en amont d'un gène et sont indispensables à la transcription de l'ADN en ARN (figure 17). Elles permettent de réguler la transcription des gènes. L'ARN polymérase se fixe initialement sur la région promotrice avant de commencer à synthétiser l'ARN.

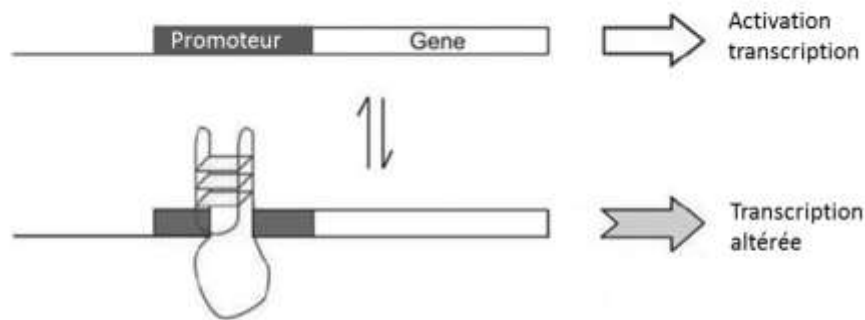


Figure 17 : Représentation schématique de la présence de G4 au sein d'un promoteur.

Les séquences de G-quadruplex potentiels sont très présentes au niveau des promoteurs des oncogènes,^[27] ils le sont moins au niveau des gènes suppresseurs de tumeurs.^[28] Il est donc possible de suggérer un rôle aux structures quadruplex dans la régulation de la transcription, en particulier des oncogènes, et d'établir un lien entre G-quadruplex et cancer. En effet, les promoteurs des gènes codant pour les protéines responsables de l'acquisition de six des altérations essentielles au basculement vers un phénotype tumoral ont été identifiés comme pouvant former une structure G-quadruplex^[29] (figure 18). Ces gènes sont :

- c-myc, c-kit et RAS ; responsables de l'autosuffisance en facteurs de croissance
- pRb ; pour l'insensibilité aux inhibiteurs de croissance
- Bcl-2 ; pour la capacité à éviter l'apoptose
- VEGF-A ; pour le développement de l'angiogénèse
- hTERT ; pour la capacité proliférative illimitée
- PDGF-A ; pour la capacité à former des métastases.

La formation d'un quadruplex au niveau de ces promoteurs peut physiquement empêcher la fixation de la machinerie de transcription de la cellule, inhiber la lecture de l'oncogène et donc la production des protéines associées.^[27a]

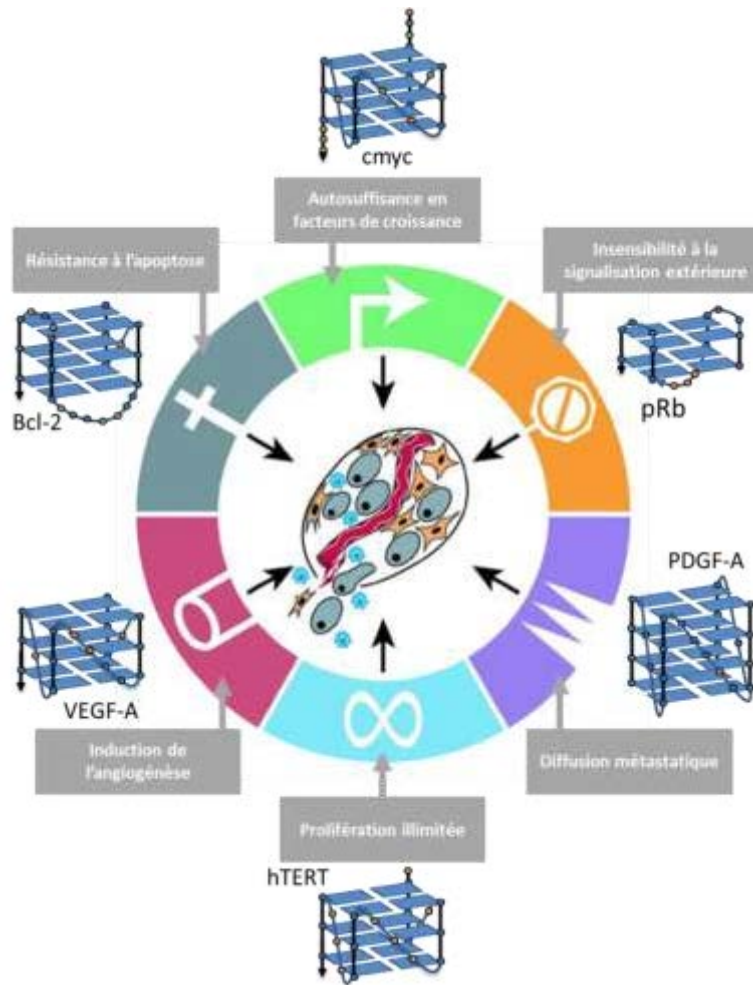


Figure 18 : Structures des G-quadruplex associés aux altérations essentielles découvertes à ce jour au développement du phénotype tumoral. D'après [30].

c) Les minisatellites

Enfin, d'autres séquences de l'ADN sont concernées par la formation de cette structure secondaire. Les minisatellites [31] sont des séquences du génome, répétées en tandem, dont la taille du motif unitaire est entre 10 et 60 nucléotides (figure 19). Elles sont présentes chez toutes les espèces. On les retrouve dans plus de mille localisations différentes dans le génome humain. Il est très fréquent d'observer des erreurs de réplication de ces minisatellites par glissement de réplication ce qui entraîne des variations interindividuelles par rapport au nombre de répétition. Cette variabilité est notamment utilisée pour les empreintes génétiques de chaque individu. Ces minisatellites sont généralement riches en guanines et cytosines d'où la très forte probabilité de présence des G-quadruplex et se situent principalement à proximité des télomères ou au niveau du centromère. Nous nous sommes intéressés à ces structures dans le cadre d'une collaboration avec l'équipe d'Alain Nicolas [32] à l'Institut Curie à Paris.



Figure 19 : Représentation schématique de la présence de G4 au sein de minisatellite.

III. Ligands de référence de G-quadruplex

Un ligand de G-quadruplex ne doit pas uniquement interagir fortement avec sa cible mais également présenter une grande sélectivité par rapport aux autres structures d'acides nucléiques et notamment le duplex. La stabilisation de cette structure par une petite molécule s'effectue principalement par empilement π - π mais le G-quadruplex peut également être stabilisé par liaisons hydrogène et par interactions électrostatiques.^[33] Les ligands sont, par conséquent, très souvent plans avec de larges surfaces aromatiques afin de recouvrir au maximum le quartet externe pour que l'interaction soit optimale. Cette large surface aromatique permet d'éviter aux ligands de pénétrer dans les sillons du double brin et induit ainsi une sélectivité. Les interactions électrostatiques entre les molécules chargées positivement et le squelette de l'acide nucléique (liaison avec les phosphates chargés négativement) n'apportent aucune sélectivité mais aident à la stabilisation tout comme les liaisons hydrogène qu'il peut y avoir entre les molécules et les bases des boucles ou encore le sillon. Le canal ionique dans lequel résident les cations monovalents (potassium, sodium par exemple) peut également être la cible de ligands^[34] (figure 20).

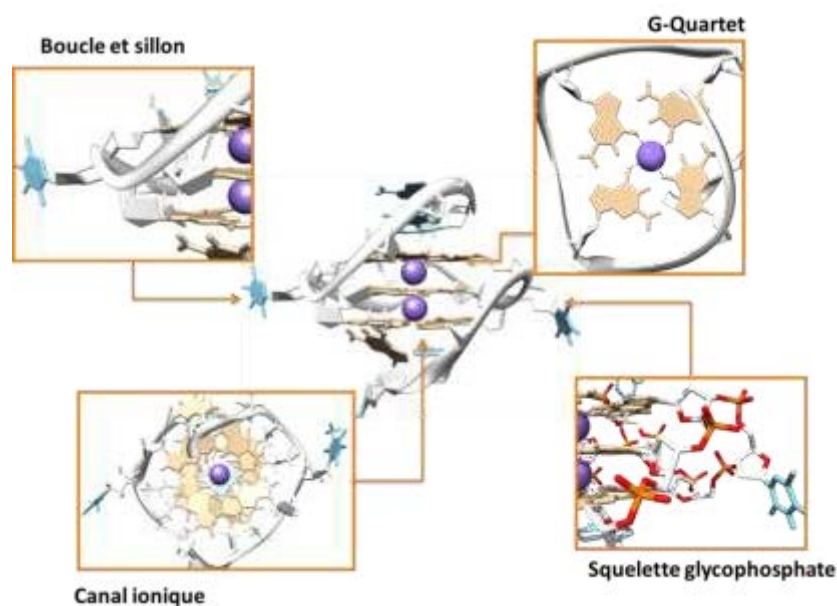


Figure 20 : Caractéristiques structurales d'un G-quadruplex (c-myc, PDB : 1XAV) ciblées par des ligands.

De très nombreuses familles de molécules différentes (macrocycle, complexe métallique, dicarboxamide) peuvent donc interagir avec les G-quadruplex. Les ligands peuvent être classés en catégories suivant leur nature cationique (protonable *in situ*, amine aromatique quaternarisée, complexe métallique...). Je vais vous présenter une liste non exhaustive de ligands classés par modes d'interactions et/ou familles structurales.

1) Interaction avec les G-quartets

La conception de ligands sélectifs de G-quadruplex est basée sur la présence d'une surface aromatique étendue pouvant interagir avec le quartet externe. Mais cela n'est pas suffisant. En effet, une molécule avec une large surface aromatique seule risque de ne pas être assez soluble en milieu aqueux pour être utilisé en solution ou *in vivo*. C'est pourquoi les différents ligands possèdent soit des chaînes protonables *in situ* soit des amines de la surface aromatique quaternarisées.

Dans le premier cas, les chaînes servent à apporter de la solubilité à la molécule et peuvent permettre de stabiliser encore plus les G-quadruplex avec des interactions électrostatiques supplémentaires. Il y a une vingtaine d'années, Neidle et Hurley ont ainsi conçu un dérivé d'une amido-anthraquinone ciblant le quadruplex et inhibant l'activité de la télomérase.^[35] Cependant la sélectivité de ces molécules pour le G-quadruplex n'est pas totale. En effet, elles interagissent également avec l'ADN duplex, c'est pourquoi ils ont continué à optimiser la molécule en faisant notamment des modifications sur le corps aromatique jusqu'à obtenir un ligand sélectif, **BSU6039**,^[36] avec deux chaînes latérales. Finalement, **BRACO-19** portant trois chaînes latérales est, aujourd'hui, un composé de référence (figure 21).^[37] En plus d'un empilement π - π efficace sur les

quartets externes, il a été montré, à l'aide de cristaux de complexes ligands-ADN, que **BRACO-19** est capable d'interagir avec les sillons de la structure grâce à ses trois chaînes latérales protonables en conditions physiologiques.^[38]

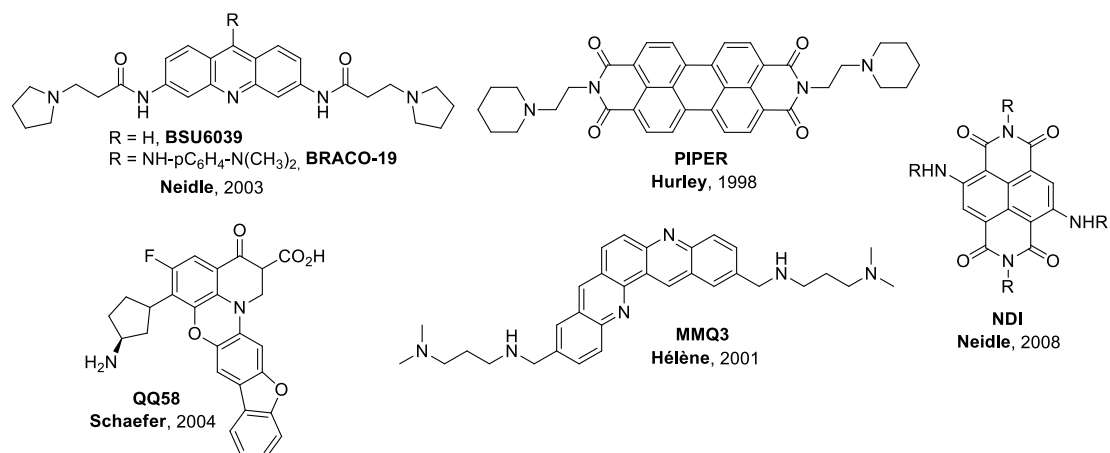


Figure 21 : Exemples de ligands protonables *in situ*.

Par la suite, de nombreux ligands avec une surface aromatique élargie se sont développés tout en conservant le même mode d'interaction et la/les chaîne(s) latérale(s) protonable(s). Il s'agit du pérylène diimide **PIPER**,^[39] de la fluoroquinolone **QQ58**^[40] ou la quinacridine **MMQ3**^[41] (figure 21). L'équipe de Neidle a également développé une famille de composés dérivés du naphthalène diimide **NDI**^[42] (figure 21), encore largement utilisée dans les études actuelles.^[43] Ces composés continuent d'être optimisés encore de nos jours avec la formation de macrocycles contenant une naphthalène diimide (NDI)^[44] pour stabiliser la séquence télomérique et la séquence du promoteur c-kit2. Ce qui nous amène à tous les composés macrocycliques qui ont été développés tel que **BOQ1**, un macrocycle comportant deux unités quinacridine semblable au **MMQ3**.^[45] Parmi les ligands neutres stabilisant les G-quadruplex, il existe également des macrocycles. Ce n'est pas la catégorie de ligands la plus importante mais elle contient un des ligands les plus connus et utilisés, notamment dans les études *in vitro*, le ligand **télomestatine**. Ce composé naturel macrocyclique, extrait de la bactérie *Streptomyces anulatus* en 2001^[46] présente une très bonne affinité pour l'ADN quadruplex^[47] et une sélectivité par rapport à la forme duplex. Cette molécule présente d'excellentes propriétés d'inhibition de la télomérase qui sont dues au recouvrement optimal des quartets externes par la molécule (figure 23b). La principale difficulté est la production de ce ligand à grande échelle par synthèse totale,^[48] bien que celle-ci ait été optimisée ces dernières années.^[49] L'équipe de Nagasawa a ensuite conçu la synthèse de macrocycles dérivés d'**hexaoxazole** en s'inspirant du composé **télomestatine**^[50] (figure 22). Ces nouveaux macrocycles présentent la particularité d'être sélectifs pour la forme antiparallèle de la séquence télomérique. Cela a été montré par des expériences de dichroïsme circulaire avec le déplacement de l'équilibre pour la forme antiparallèle en présence du

ligand. Il serait intéressant de confirmer ces données sur d'autres séquences formant des G-quadruplex antiparallèles comme la séquence 21CTA.

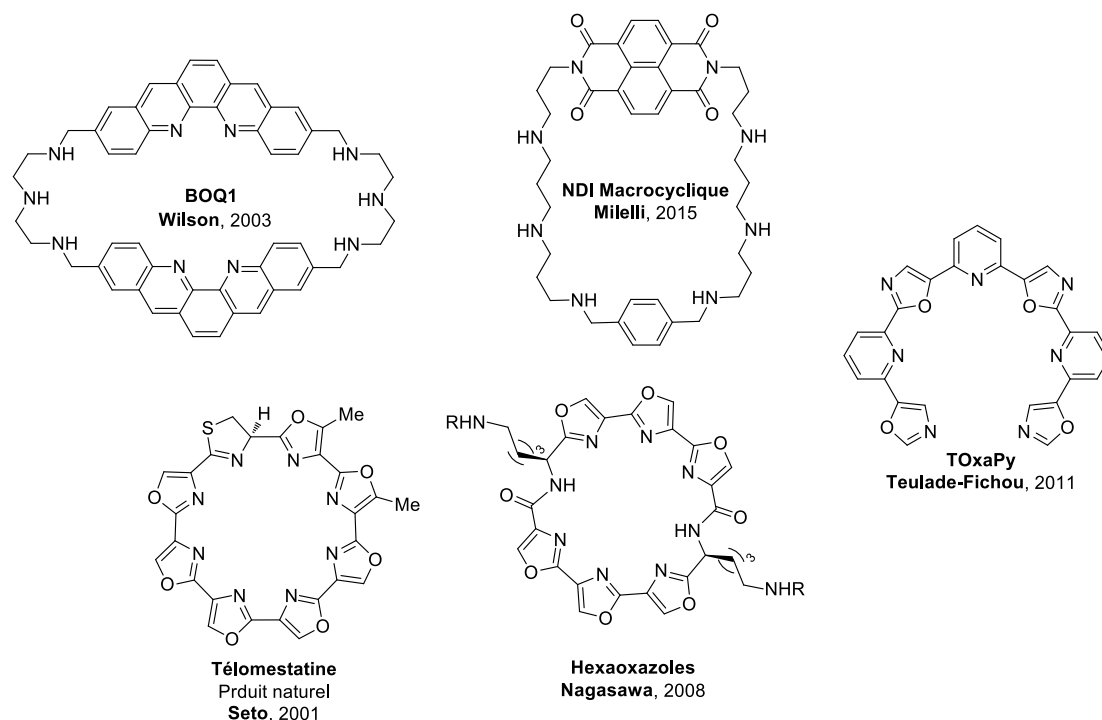


Figure 22 : Structures de ligands de G-quadruplex macrocycliques et analogue.

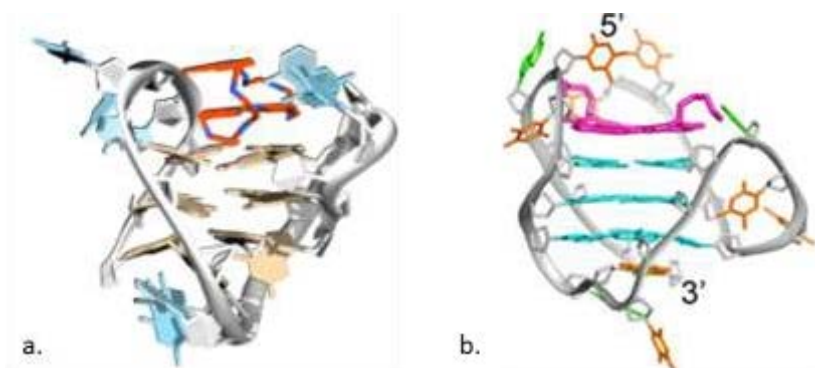


Figure 23 : a. Modélisation moléculaire de l'interaction BOQ1/c-myc^[51] et b. Structure déterminée par RMN PDB (2MB3) d'un dérivé hexaoxazole avec la forme hybride 3+1 de la séquence télomérique^[52].

Par ailleurs, des analogues non macrocycliques neutres ont été développés par la suite et, en particulier, le composé **TOxaPy** (figure 22) afin de simplifier les voies de synthèse tout en conservant la sélectivité du motif pour certaines structures G-quadruplex.^[53]

La dernière principale catégorie de ligands stabilisant les G-quadruplex est les ligands N-méthylés sur la surface aromatique. Le design de ces molécules permet une meilleure solubilité dans l'eau et de renforcer l'interaction π - π avec le quartet en appauvrissant la densité électronique du cœur aromatique. Il est important de noter que les chaînes latérales peuvent également apporter des

interactions électrostatiques favorables et que ce n'est pas un point rédhibitoire pour obtenir des ligands capables de stabiliser les G-quadruplex. Cependant, dans cette catégorie de ligands, l'absence de chaînes permet d'avoir un encombrement minimal pour établir un empilement idéal sans interaction annexe. Dans cette catégorie, on peut distinguer plusieurs familles de molécules. La première est les porphyrines dont l'exemple le plus connu et qui est très utilisé notamment pour les tests *in vitro* est le ligand **TMPyP4** ^[54] (figure 24) Cette molécule possède une très bonne affinité pour les G-quadruplex, inhibe la télomérase et peut réguler l'expression d'oncogènes comme c-myc et KRAS. L'affinité a principalement été mesurée par des expériences de FRET-melting sur la séquence télomérique en milieu potassium et sodium qui ont montré une bonne stabilisation du G-quadruplex par le ligand.^[55] Cependant, le petit bémol (et pas des moindres) est qu'elle n'est pas très sélective vis-à-vis de l'ADN duplex.^[55-56] Il est tout de même encore très utilisé comme ligand de référence.^[57] De nombreux ligands structurellement proches ont été développés dont des porphyrazines dont la plus connue est le composé **3,4-TMPyPz** ^[58] et la porphyrine **TQMP** (figure 24) développée par le groupe de Tan ^[59] qui a déplacé la charge sur des amines benzyliques. Ces deux molécules conservent une bonne affinité pour la structure G-quadruplex et ont une bien meilleure sélectivité par rapport à l'ADN duplex que **TMPyP4**. Enfin, la molécule conçue à partir de la structure porphyrine qui est la plus affine et la plus sélective pour les G-quadruplex est le composé **Se2SAP** (figure 24) développé par le groupe de Hurley.^[60] Il s'agit d'une des premières molécules (avec le composé **télomestatine**) déplaçant un équilibre vers une conformation préférentielle d'un G-quadruplex, la forme hybride sur la séquence télomérique.^[61]

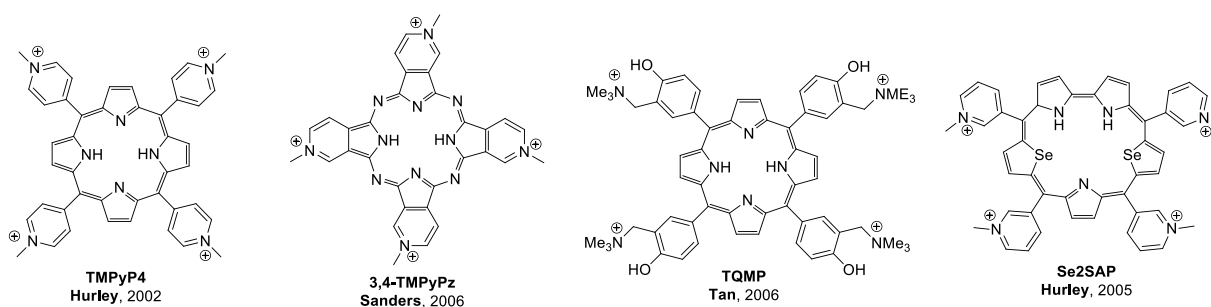


Figure 24 : Différentes structures porphyrines stabilisant les G-quadruplex.

A la même période, une série de composés de très grandes affinités et sélectivités a été développée par l'équipe de Riou autour de dérivés de bisquinolinium. Un ligand se révèle particulièrement intéressant, le ligand **360A** (figure 25), qui possède un cœur pyridodicarboxamide.^[62] En effet, il permet d'inhiber la télomérase et induit l'apoptose de différentes lignées cellulaires.^[63] De plus, ce composé **360A**, aussi appelé **PDC**, est capable de s'accumuler au niveau de l'extrémité télomérique des chromosomes ce qui permet d'appuyer l'hypothèse concernant la formation de structures G-

quadruplex au niveau des télomères.^[64] Cette molécule favorise également la formation de G-quadruplex tétramoléculaire.^[65] L'avantage de ces composés est leur simplicité d'obtention avec une synthèse rapide et efficace. Quelques années plus tard, un analogue du ligand **PDC** équipé de chaînes latérales protonables *in situ* a été développé par l'équipe de Balasubramanian, il s'agit du composé **pyridostatine (PDS, figure 25)**.^[66] Cet analogue et ses dérivés^[67] sont de très bons stabilisants de G-quadruplex et sont très utilisés comme ligands de références dans de nombreuses études.^[68] Lors de ces études, un autre ligand de référence est très souvent utilisé en comparaison, **PhenDC₃**.^[69] Il a été développé par l'équipe de Teulade-Fichou à partir du **PDC**, en élargissant le cœur aromatique, passant d'une pyridine à une phénanthroline (figure 25). La méthylation en position 6 de la quinoline donne le composé **PhenDC₆** qui est également un très bon ligand stabilisant la structure quadruplex (figure 25). La famille **PhenDC** possède une excellente affinité et surtout une excellente sélectivité par rapport à l'ADN duplex. Le recouvrement du composé **PhenDC₃** sur le quartet est optimal comme cela a été montré par spectrométrie RMN (figure 26).^[70] Cette molécule est une des plus utilisées à travers le monde pour cibler les G-quadruplex.^[32a, 71]

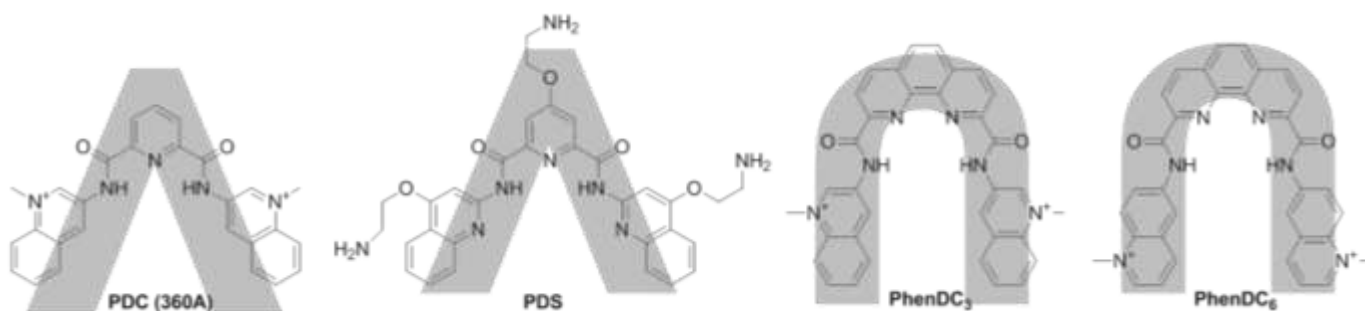


Figure 25 : Structure en V des composés PDC et PDS et structure en U du PhenDC₃ et PhenDC₆.

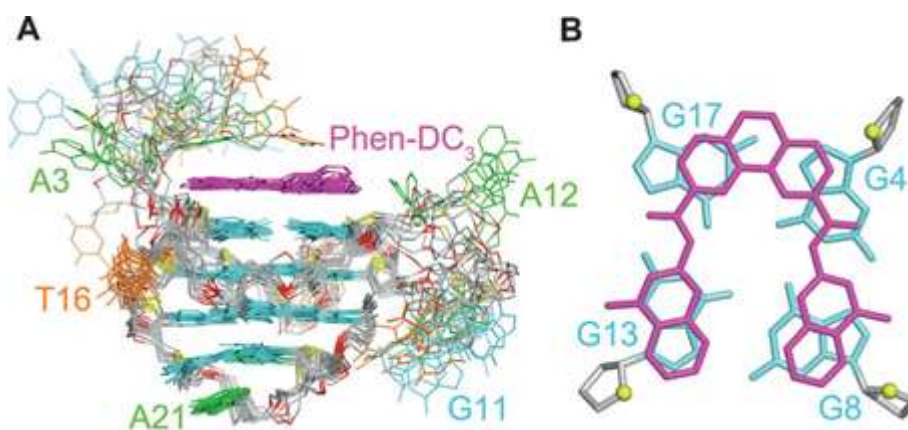


Figure 26 : Structure déterminée par RMN de l'interaction PhenDC₃ avec un G-quadruplex issue de l'oncogène c-myc (séquence Pu24T) A. vue latérale et B. vue de dessus.

Pour terminer sur cette famille de composés, le groupe de Pradeepkumar continue de développer des dérivés (figure 27) en s'inspirant de la structure en U du PhenDC₃ [72] et s'en sert encore aujourd'hui pour étudier l'oncogène c-myc.[73]

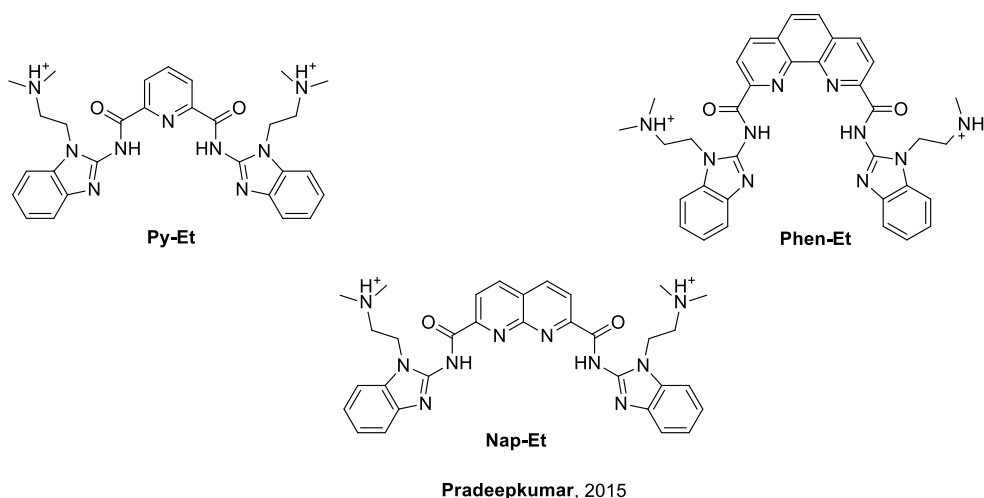


Figure 27 : Structures des composés développés par l'équipe de Pradeepkumar.[72]

2) Interaction avec le sillon

C'est un défi encore aujourd'hui de trouver un ligand interagissant avec les sillons de l'ADN G-quadruplex sélectivement. En effet, les différences structurales entre les sillons d'ADN duplex et quadruplex sont significatives. La reconnaissance exclusive de ces sillons permettrait l'obtention d'une sélectivité très élevée pour un G-quadruplex en particulier et serait un nouveau moyen de stabiliser ces structures.[74] La **Distamycine A** (figure 28) est le premier ligand pour lequel l'interaction avec les sillons de quadruplex a été formellement prouvée. Cette molécule se lie avec une stœchiométrie 4:1 à l'ADN quadruplex tétramoléculaire d(TGGGGT). Ce ligand se place dans les sillons tout en établissant des liaisons hydrogène avec les guanines. Malheureusement, il a également une forte affinité pour l'ADN double hélice et notamment le petit sillon des régions riches en adénine et thymine qu'il stabilise avec une stœchiométrie 2:1.[75]

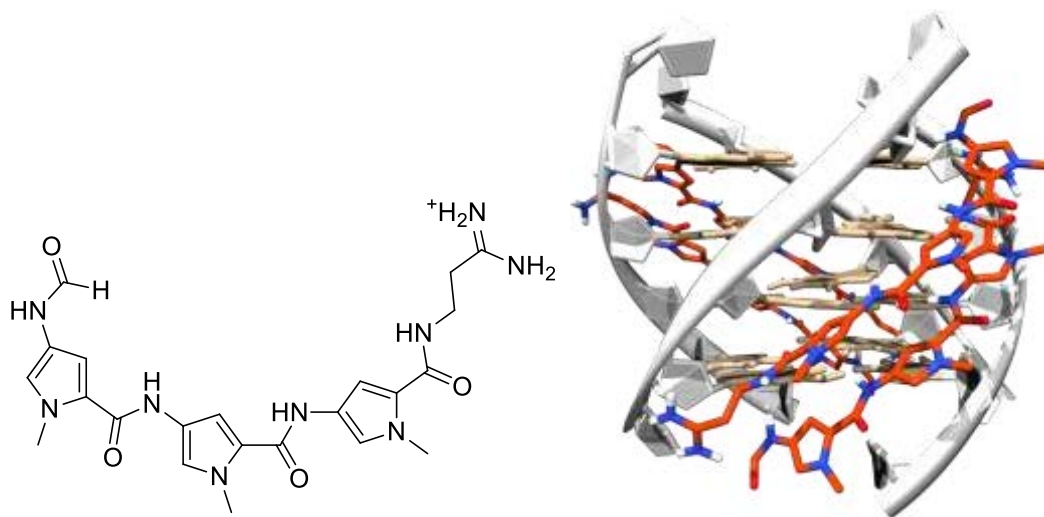


Figure 28 : Structure de la Distamycine A (gauche) et structure déterminée par RMN (PDB : 2JT7) de la molécule sur le G-quadruplex tétramoléculaire d(TGGGGT) (droite).

3) Interaction avec le canal ionique

Le canal ionique du G-quadruplex peut également être une cible afin de stabiliser cette structure, notamment par une molécule protonable *in situ* (figure 29). La structure est basée sur l'hypothèse que la fonction anthracène, empilée sur les guanines du quartet, et les ammoniums comme cations induisent la structuration d'un G-quadruplex par liaisons hydrogène et interactions cation-dipôle, imitant le canal ionique central. Cependant, ces interactions sont très peu observées et pour ce genre de composés l'interaction π - π reste la plus importante pour stabiliser les structures G-quadruplex.^[34]

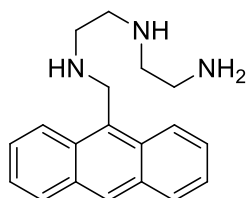


Figure 29 : Structure du ligand ciblant le canal ionique central du G-quadruplex.

4) Les complexes métalliques

Certains ligands métalliques sont connus pour se lier par liaison de coordination avec l'ADN. La molécule la plus connue est bien évidemment le composé **cis-platine**.^[76] En fonction du métal utilisé, la géométrie des complexes est différente (le platine forme des complexes plan carré, le cuivre forme des complexes avec une géométrie pyramidale, le zinc forme des complexes tétraédriques...). Le potentiel de coordination métallique varie également en fonction du métal. Si le cis-platine et ses analogues sont majoritairement utilisés sur l'ADN duplex, ils peuvent également former des adduits de platination avec les G-quadruplex sans aucune sélectivité.^[77]

Afin de trouver de nouveaux ligands de G-quadruplex sélectifs, le groupe de Vilar a commencé à développer des dérivés de **Ni-salphen** (figure 30) qui montrent une affinité similaire à celle du composé **BRACO-19** dans l'expérience de FRET-melting et inhibent l'activité de la télomérase.^[78] Par la suite, des complexes de platine ont été synthétisés autour de structure phénanthroline et pyridine.^[79] Toutes ces molécules sont optimisées pour recouvrir au mieux le quartet externe du G-quadruplex et ainsi induire de la sélectivité vis-à-vis du duplex.

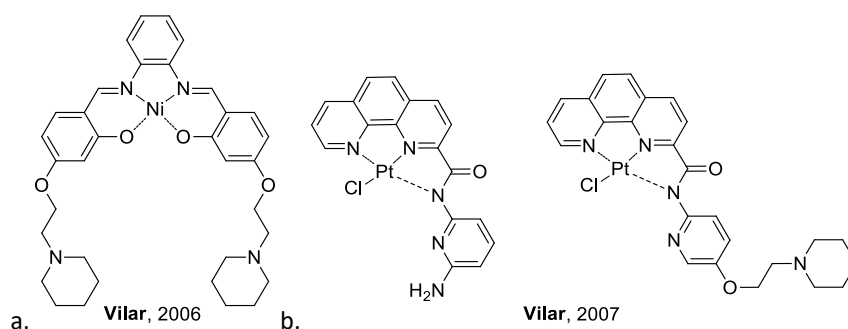


Figure 30 : Ligands métalliques de G-quadruplex a. dérivé de Ni-salphen et b. dérivés de Pt-phénanthroline.

Un complexe métallique qui serait capable de cibler efficacement et sélectivement la structure G-quadruplex peut présenter un double avantage. En effet, cela permettrait à la fois d'avoir une synthèse et une géométrie modulable mais également de former un adduit métallique avec les nucléotides. En particulier, les complexes de platine dérivés de la terpyridine (figure 31) développés par l'équipe de Teulade-Fichou sont capables de former des liens covalents avec l'ADN quadruplex.^[80] Cette interaction est majoritairement dirigée par l'empilement π - π de la surface aromatique plane du ligand sur les quartets extérieurs du G-quadruplex. Introduit en 1931 par Morgan et Burstall,^[81] le motif terpyridine, complexé par de nombreux métaux, est largement utilisé dans de nombreux domaines : photochimie, chimie supramoléculaire, catalyse... Ce motif est facilement fonctionnalisable et ses propriétés de métallation de l'ADN sont particulièrement intéressantes.^[82]

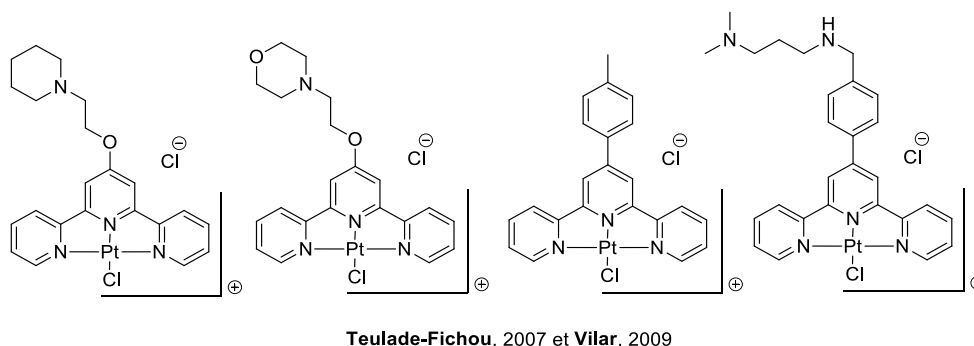


Figure 31 : Ligands métalliques autour de la structure Pt-terpyridine.

Les complexes Pt-terpyridine ont été mis au point et développés en 2007 ^[83] par l'équipe de Teulade-Fichou puis repris par l'équipe de Vilar en 2009 ^[84]. L'ajout de courtes chaînes protonables *in situ* sur le motif terpyridine augmente l'affinité pour les G-quadruplex. Cependant, ces composés restent également très affins pour le duplex à cause de leur petite taille et leur faculté à s'intercaler dans le duplex. La fonctionnalisation de la pyridine centrale par un groupement tolyle permet d'élargir la surface aromatique. Par conséquent, l'interaction π - π avec le quartet externe est renforcée et la sélectivité par rapport à l'ADN double brin est améliorée. Les interactions des complexes de tolylterpyridine (ttpy) au cuivre, palladium ou platine avec l'ADN G-quadruplex ont été très étudiées au sein de l'équipe de Teulade-Fichou.^[85] En particulier, le mécanisme de métallation de l'ADN (figure 32) a été déterminé par une étude par spectrométrie de masse des complexes au palladium et au platine.^[86] Ces complexes passent par une forme *aqua* avant de réagir avec les bases de l'ADN.

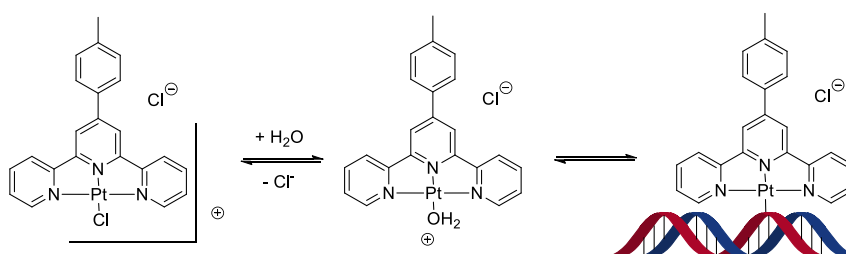


Figure 32 : Schéma du mécanisme de platination du composé Pt-ttpy sur l'ADN.

Il a également été montré, par une étude sur gel d'électrophorèse, que les sites de platination du complexe **Pt-ttpy** sur la structure G-quadruplex de la séquence télomérique sont les adénines présentes dans les boucles.^[80] Le complexe **Pt-tpy** permet, pour sa part, une platination externe sans reconnaissance spécifique de la structure (figure 33).

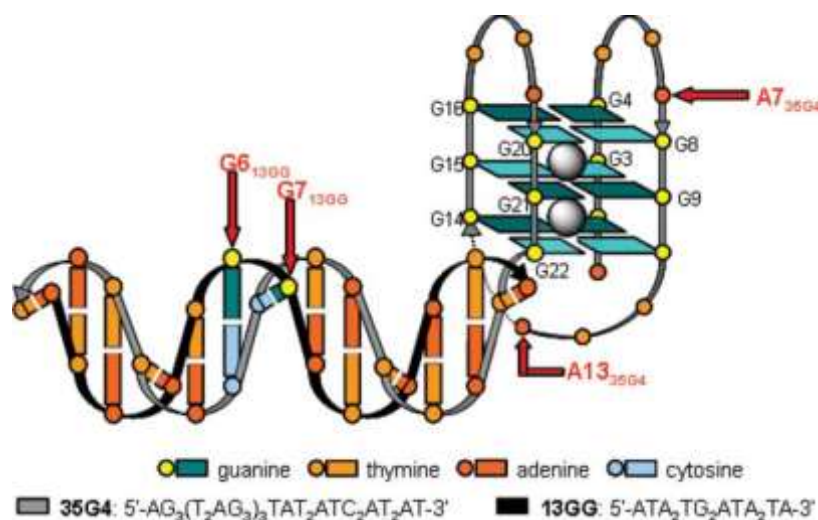


Figure 33 : Sites de platination de Pt-ttpy et Pt-tpy sur la séquence télomérique humaine. Les flèches rouges indiquent les bases platinées.

Les composés triarylpyridines (ou TAP), dérivés de terpyridine, possèdent une affinité intéressante pour les G-quadruplex en l'absence de tout centre métallique.^[87] *A contrario*, les terpyridines sans centre métallique n'ont aucune affinité pour les structures G-quadruplex. Ces ligands étendus interagissent manifestement avec le quartet mais aussi avec les boucles et les sillons. Ces composés provoquent la résolution des structures G-quadruplex en les déstabilisant.^[87b]

Enfin, un grand nombre de ligands macrocycliques possédant un centre métallique ont également été développés,^[88] en particulier des métalloporphyrines de nickel ou de manganèse qui étaient déjà connues pour leurs interactions avec l'ADN duplex. Par ailleurs, les porphyrines ayant des substituants cationiques en *meso* sur leur macrocycle possèdent une affinité plus forte avec la structure G-quadruplex.

Pour terminer avec les ligands métalliques, il existe de nombreux complexes de ruthénium (figure 34) qui permettent de cibler les G-quadruplex. L'avantage de ces composés est leur fluorescence. Certains ligands sont éteints en solution et s'allument en présence du G-quadruplex ou vont même être sélectif d'une structure G-quadruplex par rapport à une autre.^[89]

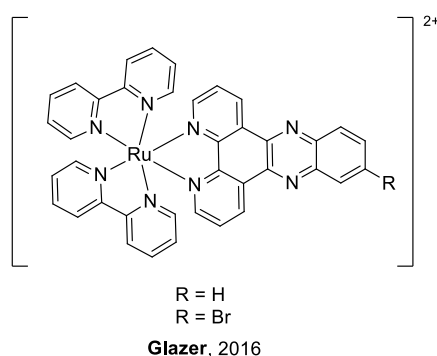


Figure 34 : Exemple de complexe de ruthénium.

5) Cas particuliers : les ligands fluorescents

L'avantage des ligands fluorescents est qu'ils peuvent être observés en milieu cellulaire par microscopie (confocale ou épifluorescence) ce qui permet d'avoir une méthode directe de localisation des molécules. Le premier composé fluorescent à avoir été développé est **RHPS4** (figure 35) qui a été et est encore aujourd'hui utilisé comme ligand de référence dans différentes études.^[90] Il a été montré qu'il inhibe la télomérase et a des effets anti-prolifératifs au sein de cellules tumorales du cerveau.^[91] Les effets biologiques du composé en cellules ont été observés par microscopie par des techniques d'immunomarquage.^[92] En revanche, la molécule elle-même n'a pas été imagée malgré sa fluorescence. *A contrario*, le ligand **Quarfloxine** (ou **CX-3543**, figure 35) a été observé directement en cellules et une localisation nucléolaire a été observée.^[93] Cette molécule a

été très étudiée, brevetée en 2007 (brevet WO 2007/137000A2) et proposée comme candidat médicament allant jusqu'en phase clinique (Arrêt en phase 1). Récemment encore, elle a été utilisée pour montrer les dommages à l'ADN et sa mortalité sélective de cellules tumorales déficientes en BRCA1/2.^[94] Ensuite, un composé qui n'est pas fluorescent lorsqu'il est dans sa forme libre mais qui le devient en présence de G-quadruplex a été mis au point, il s'agit du **N-TASQ**^[95] (figure 35). Cette famille de composé a été développée par le groupe de Monchaud.^[96] Ce sont des ligands possédant une bonne affinité et une sélectivité pour différentes structures G-quadruplex par rapport à l'ADN duplex. Le ligand N-TASQ a été observé dans le cytoplasme de cellules vivantes.^[95] Enfin, le composé utilisé par l'équipe de Vilar en cellules vivantes est **DAOTA-M2**^[97] (figure 35), ce composé fluorescent est très intéressant car il a un temps de vie de fluorescence différent en présence d'ADN duplex et d'ADN quadruplex. Le point négatif est que la molécule n'est pas sélective pour la structure G-quadruplex. Ce ligand a été observé dans le noyau, le cytoplasme et les mitochondries de cellules vivantes humaines.^[97]

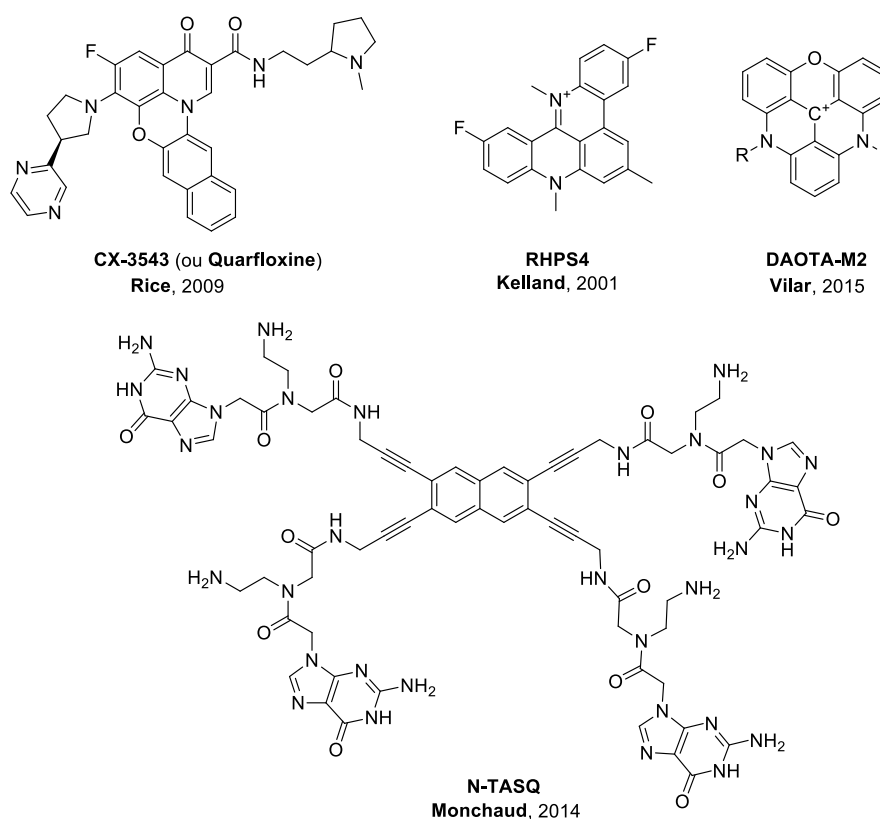


Figure 35 : Structure de ligands de G-quadruplex fluorescents.

IV. Méthodes d'analyses des interactions ligand/G-quadruplex

Les méthodes utilisées pour étudier les interactions ligand/G-quadruplex sont d'une grande variété. Ces méthodes peuvent être simples et faciles à mettre en œuvre (évaluation d'affinité) ou plus

sophistiquées (informations structurales, thermodynamiques, cinétiques, stœchiométriques...).[98] Toutes ces techniques d'analyse sont adaptées de celles utilisées pour l'ADN duplex. Elles doivent permettre à la fois de déterminer l'affinité du composé mais également sa sélectivité par rapport au double brin. Le grand polymorphisme des structures quadruplex rend encore plus difficile l'étude des interactions. Il est donc intéressant de connaître l'affinité des ligands sur différentes structures G-quadruplex. Afin d'éviter les biais d'une méthode en particulier, il est recommandé d'utiliser une combinaison de techniques d'analyse pour essayer de comprendre l'effet de stabilisation de la structure quadruplex ou encore le mode de fixation du ligand.

1) Analyses semi-quantitatives

Les méthodes développées dans cette partie sont des méthodes qui permettent principalement le classement des ligands par rapport à un oligonucléotide évalué.

a) FRET-melting

L'expérience de FRET-melting est basée sur une fusion de l'ADN suivie par un transfert d'énergie par résonance de type Förster (figure 36). Cette méthode a été largement développée par l'équipe de Mergny en utilisant des quadruplex fonctionnalisés par deux fluorophores, un donneur de fluorescence (la fluorescéine FAM) en 5' de l'oligonucléotide et un accepteur de fluorescence (une rhodamine, le TAMRA) en 3' de la séquence d'acides nucléiques.[99]

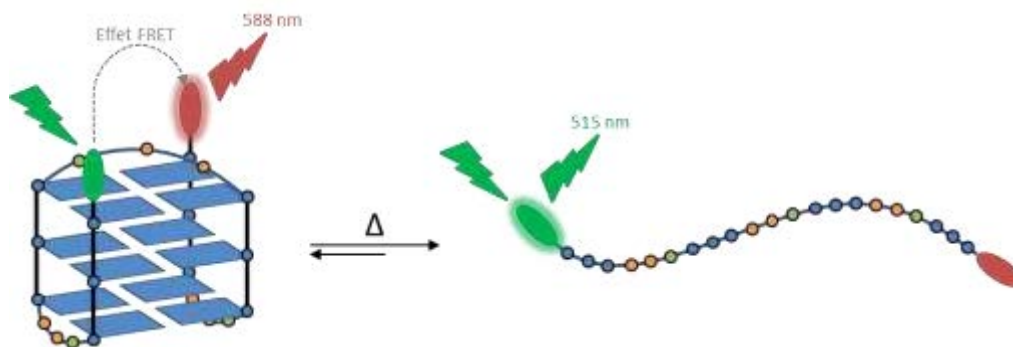


Figure 36 : Principe de l'expérience de FRET-melting.

Lorsque la fluorescéine est excitée à 488 nm et que l'ADN est sous forme G-quadruplex, l'énergie engendrée est transférée par effet FRET à la rhodamine proche ce qui lui donne une forte intensité de fluorescence. Lors de la dénaturation thermique de l'ADN, le quadruplex s'ouvre et l'oligonucléotide sous forme simple brin est complètement déroulé, par conséquent, les deux fluorophores sont très éloignés l'un de l'autre et l'effet FRET n'est plus observé. La fluorescence du donneur est alors restaurée. En suivant la fluorescence du donneur en fonction de la température, une courbe de fusion de l'acide nucléique étudié peut donc être obtenue (figure 37).

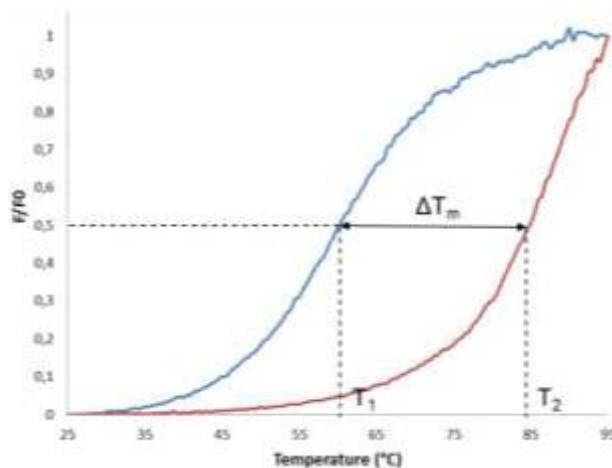


Figure 37 : Exemple de courbes de fusions d'un oligonucléotide G-quadruplex seul (en bleu) et en présence d'un ligand qui le stabilise (en rouge).

Cette expérience permet également de mesurer l'affinité d'un ligand pour le quadruplex évalué. En effet, l'ajout d'une petite molécule affine à cette structure va la stabiliser et donc retarder la dénaturation de l'ADN. La température de fusion obtenue (T_2) pourra être comparée à la valeur T_1 de l'ADN seul. La valeur $\Delta T_m = T_2 - T_1$, mesurée à une valeur de fluorescence $F/F_0 = 0,5$ où F_0 est la valeur de fluorescence maximale (à $T=95$ °C), renseigne donc sur l'affinité du ligand (figure 37). Plus les valeurs de ΔT_m sont grandes, plus la molécule est affine à la structure G-quadruplex. En général, un ligand est considéré comme ayant une bonne affinité pour un ΔT_m supérieur à 10 °C.

L'expérience de FRET-melting est utilisée en routine au laboratoire pour évaluer les nouveaux composés synthétisés. C'est une méthode rapide, fiable et facile à mettre en œuvre avec un équipement standard (fluorimètre thermo-contrôlé). Devant le grand nombre de molécules à évaluer et le nombre d'oligonucléotides croissant, nous avons mis en place une expérience à haut débit en utilisant des plaques 96 puits et un appareil de qPCR pour suivre la dénaturation thermique de l'ADN en mesurant la fluorescence en temps réel. Les valeurs de ΔT_m dépendent directement de la stabilité de l'oligonucléotide testé et des conditions de concentration et de salinité dans lesquelles est réalisée l'expérience. En général, les conditions utilisées au laboratoire sont les suivantes : 0,2 μM d'ADN, 1 μM de ligand soit 5 équivalents et un tampon constitué de cacodylate de lithium (10 mM) et de sels de potassium (KCl, 1 ou 10 mM). La sélectivité est également évaluée par ce test en ajoutant dans le milieu de l'ADN duplex (en général au laboratoire du ds26 à 3 ou 10 μM) non marqué ce qui nous permet de constater s'il y a un effet sur la valeur de ΔT_m . Il est également important de savoir que le marquage d'un oligonucléotide par deux fluorophores peut éventuellement perturber les interactions avec certains ligands et que ceux-ci peuvent aussi interférer avec les fluorophores.

b) Fluorescence Intercalator Displacement (FID)

L'expérience de FID (*Fluorescence Intercalator Displacement*) est une technique adaptée pour les structures G-quadruplex par l'équipe de Teulade-Fichou en 2006. Elle permet d'évaluer l'affinité des molécules pour diverses séquences d'acides nucléiques via le déplacement d'une sonde *on-off*.^[56a, 100] La sonde est éteinte en solution et devient fluorescente une fois fixée sur l'ADN. L'ajout progressif d'un ligand affiné pour l'ADN quadruplex engendre une diminution de la fluorescence ce qui signifie que la sonde est déplacée par le ligand (figure 38).

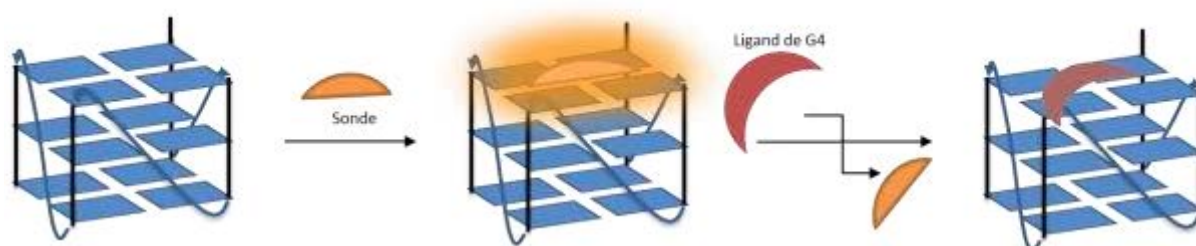


Figure 38 : Principe de l'expérience de FID.

Pour cette expérience, différentes sondes peuvent être utilisées comme fluorophore. Elles doivent présenter une forte exaltation de fluorescence en présence d'acides nucléiques. Elles possèdent une affinité de l'ordre du micromolaire pour différentes structures d'ADN ce qui permet d'évaluer à la fois l'affinité mais également la sélectivité si le ligand déplace la sonde de l'ADN duplex. La sonde principalement utilisée est le **Thiazole Orange (TO)**, figure 39) qui a une affinité à la fois pour le duplex^[101] et le G-quadruplex.^[100a] Deux autres sondes peuvent être utilisées au cas où il y ait un recouvrement spectroscopique entre le ligand et la sonde, une sonde décalée vers le rouge qui est le **TO-PRO3**^[102] et peut notamment servir pour évaluer les composés de type porphyrines et une sonde décalée vers le bleu qui est le **Hoechst 33258**^[103] (figure 39). Ces deux sondes interagissent à la fois avec le duplex et avec le quadruplex comme le **TO**.

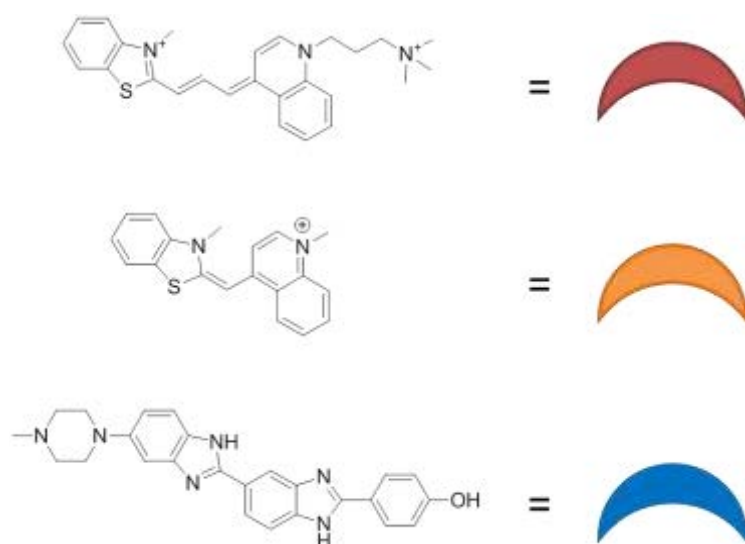


Figure 39 : Structures des sondes utilisées pour l'expérience de FID : TO-PRO3 (rouge), TO (orange), Hoechst 33258 (bleu).

Le thiazole orange possède un axe de rotation dans sa structure à l'état libre qui lui permet une libre rotation, une fois fixée sur sa cible la conformation de la molécule est plane et génère la fluorescence (figure 40).



Figure 40 : Structure du thiazole orange à l'état libre et fixé sur sa cible (conformation plane).

On représente la diminution de fluorescence observée qui correspond au déplacement de la sonde en fonction de la concentration de ligand ajouté. Dans le cas du **TO**, on note TOD (Thiazole Orange Displacement) cette grandeur et est exprimée par :

$$TOD (\%) = 100 \times \left(1 - \frac{F}{F_0}\right)$$

Avec F la valeur de fluorescence observée et F_0 la fluorescence initiale.

La sélectivité inter-oligonucléotide peut également être mesurée en estimant l'affinité sur diverses séquences, de façon séparée. Pour cela, il faut avoir une bonne connaissance des interactions sonde fluorescente / ADN, notamment en termes de stœchiométrie, d'affinité et d'exaltation de fluorescence.

Les avantages du FID sont nombreux. Il est possible d'utiliser différentes sondes, cette technique permet également de travailler à une température constante tout le long de l'expérience et utilise

des oligonucléotides non-marqués. Elle est facile et rapide à mettre en œuvre et est utilisée en routine au laboratoire. Il est possible de robotiser ce test ce qui est très utile pour faire du haut débit. Par ce procédé nous obtenons une estimation de l'affinité de manière semi-quantitative. Le FID est une titration indirecte puisqu'elle se fait via le déplacement d'une sonde fluorescente ce qui implique un équilibre chimique supplémentaire (figure 41).

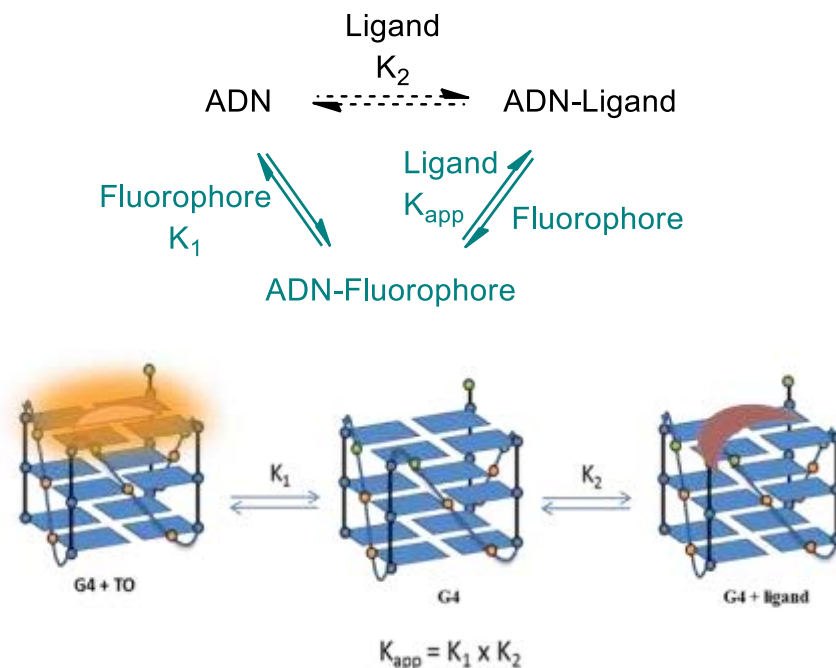


Figure 41 : Equilibres impliqués lors de la titration par FID

En général, les conditions d'utilisation de ce test sont les suivantes : 0,25 μM d'ADN, 0,5 μM de sonde pour l'étude d'ADN G-quadruplex et 1 μM de sonde pour l'étude d'ADN duplex et un ajout de ligand à évaluer entre 0 et 2,5 μM . On représente schématiquement la sonde interagissant sur le quartet externe de la structure G-quadruplex cependant il n'y a aucune certitude quant au mode d'interaction de celle-ci sur le G4.

La compétition ligand/fluorophore avec l'ADN est la limite de ce test car les constantes d'affinité ne peuvent pas nécessairement être mesurées (sauf si $K_2 \gg K_1$). Cependant, il permet d'effectuer un classement des ligands en fonction de leur affinité pour la séquence étudiée. L'affinité peut alors être exprimée de deux manières différentes :

- par une valeur de déplacement du fluorophore à une concentration donnée, on prend arbitrairement pour valeur 1 μM (4 équivalents de ligands par rapport à l'ADN) qui correspond à une valeur pour laquelle les très bons ligands de G4 ont déjà atteint (ou presque) le plateau de saturation (méthode principalement utilisée dans ce manuscrit, figure 42).

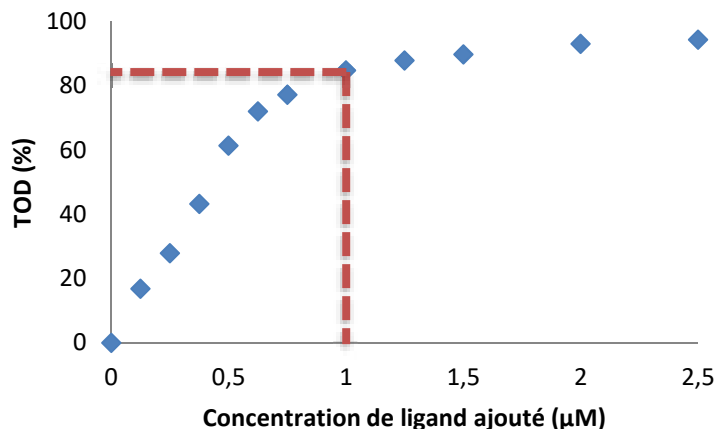


Figure 42 : Exemple de détermination du déplacement de TO à 1 μM

- par une valeur de concentration en ligand pour laquelle 50% de la sonde fluorescente a été déplacée (figure 43).

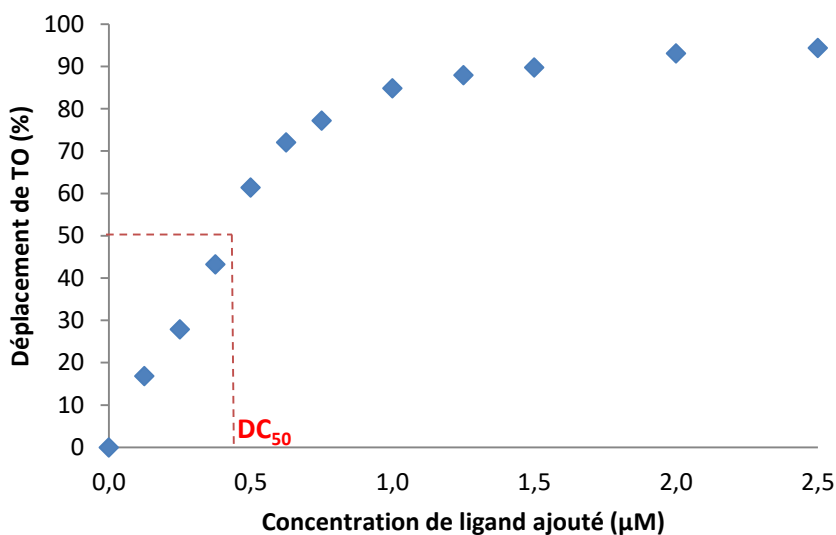


Figure 43 : Exemple de détermination de la valeur de DC50.

c) Dichroïsme circulaire

Le dichroïsme circulaire est basé sur l'absorption différentielle de la lumière polarisée circulairement à gauche et à droite. Tout matériau absorbant de manière différente une lumière polarisée circulaire gauche ou circulaire droite présente un dichroïsme circulaire. L'une des deux composantes de la polarisation sera absorbée plus rapidement que l'autre. Cette technique d'analyse s'est beaucoup développée pour l'étude de structures secondaires de l'ADN comme les G-quadruplex ou de protéines. C'est une méthode non-destructive (l'ADN conserve sa structure G-quadruplex tout au long de l'expérience) qui permet d'étudier les changements de conformations des G-quadruplex en présence de ligands ou encore leur stabilité dans des conditions salines particulières. En général, il

est nécessaire d'utiliser des concentrations en ADN supérieures à 2,5 μM pour avoir des spectres avec une bonne résolution.

Il est possible de reconnaître une structure secondaire d'ADN grâce à l'allure de son spectre de dichroïsme circulaire. La forme duplex en hélice β va par exemple se distinguer par un maximum d'absorbance à 280 nm et un minimum à 250 nm. Les séquences d'oligonucléotides adoptant une structure G-quadruplex présentent également un spectre de dichroïsme circulaire particulier (figure 44). S'il s'agit d'un quadruplex parallèle, il présentera un minimum local d'absorbance à 240 nm et un maximum à 260 nm (dit de type I), s'il s'agit d'un antiparallèle il y a deux groupes différents, le premier possède des maxima à 245 et 290 nm et un minimum à 265 nm (dit de type II)^[104] et le deuxième possède des maxima à 290 et 260 nm et un minimum à 245 nm^[105]. Pour la séquence télomérique humaine 22AG qui possède un haut polymorphisme (quatre structures coexistantes : parallèle, antiparallèle et hybrides), le spectre présente un maximum à 290 nm et un épaulement vers 260-270 nm.

Acronyme	Séquence (5' -> 3')	Tampon	Conformation
22AG	AGGGTTAGGGTTAGGGTTAGGG	K ⁺	polymorphe
22AG	AGGGTTAGGGTTAGGGTTAGGG	Na ⁺	antiparallèle
Pu22	TGAGGGTGGGTAGGGTGGGTAA	K ⁺	parallèle

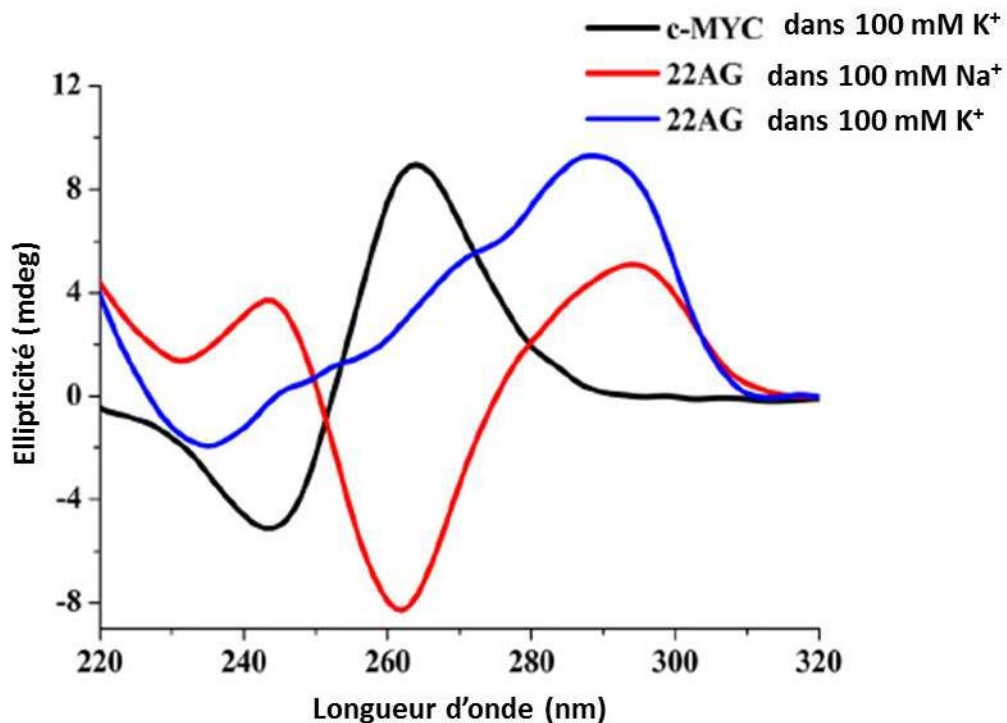


Figure 44 : Spectres de dichroïsme circulaire des séquences 22AG et C-MYC, d'après ^[106].

Des informations sur les sites d'interactions ligand/ADN G4 peuvent également être obtenues. En effet, une molécule non-chirale qui se lie à la biomolécule chirale comme l'ADN induit un signal de dichroïsme circulaire dans une zone du spectre où le ligand absorbe.^[107]

2) Analyses quantitatives

a) Résonance plasmon de surface

La résonance plasmon de surface (SPR, *Surface Plasmon Resonance*) est une technique physique permettant l'étude des interactions entre un ligand et une cible adsorbée à la surface d'une couche métallique, souvent en or. Le système de détection SPR mesure la variation de l'indice de réfraction près de la couche métallique lorsque le ligand se fixe sur sa cible et permet la détermination de constantes d'affinité. Une fois le ligand fixé sur sa cible, l'angle de réflexion du rayon envoyé sur la puce change et peut être analysé (figure 45).

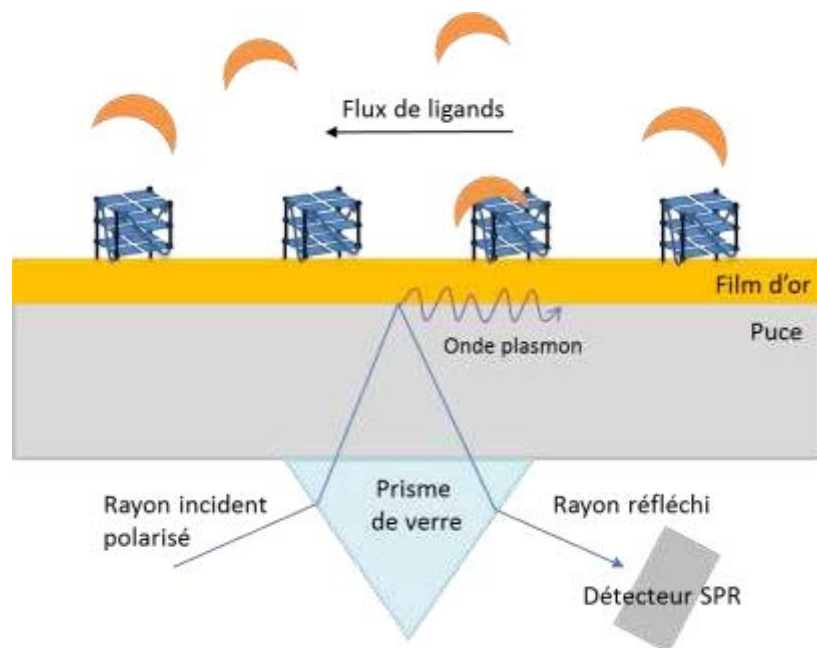


Figure 45 : Principe de la technique de SPR pour les G-quadruplex.^[108]

Les séquences d'acides nucléiques formant une structure G-quadruplex peuvent être immobilisées sur ces surfaces métalliques pour étudier leurs interactions avec des ligands. La technique se limite principalement aux G-quadruplex intramoléculaires afin d'éviter la dissociation potentielle des structures intermoléculaires pendant les étapes de traitement de la puce, propres à cette technique d'évaluation. En général, les oligonucléotides couplés à une biotine sont immobilisés sur une surface recouverte de streptavidine.^[109] De nombreux ligands ont pu être évalués par cette méthode dont des porphyrines et la télomestatine.^[61, 110]

Afin d'élargir la méthode à des quadruplex intermoléculaires, des squelettes peptidiques macrocycliques ont été développés par l'équipe d'Eric Defrancq.^[111] Le quadruplex est formé par l'association de quatre oligonucléotides liés au squelette peptidique lui-même fonctionnalisé par une biotine ce qui permet d'immobiliser l'assemblage sur la surface de la puce. Récemment, de nouvelles structures ont été développées sur ce principe pour avoir une large gamme de quadruplex (figure 46) et même une *hairpin* ^[112] ce qui a permis d'obtenir des valeurs de constante de dissociation pour des ligands de référence comme PhenDC₃ et PDS. Par exemple, sur le G-quadruplex mimant la séquence télomérique humaine, le composé PhenDC₃ a une constante de dissociation de 10⁻⁹ M alors que PDS a une constante de 10⁻⁸ M.^[112c]

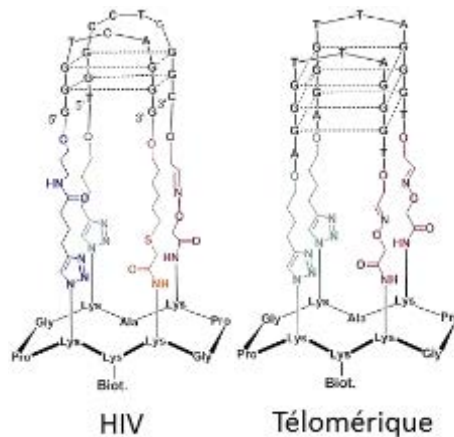


Figure 46 : Exemples de structures de quadruplex immobilisés pour la SPR.

Une étude d'imagerie par SPR a permis le criblage simultané de ligands sur la structure G-quadruplex issue de la séquence télomérique humaine et sur de l'ADN duplex ou simple brin.^[113]

Cette méthode possède l'avantage de pouvoir nous donner des valeurs de constante de dissociation d'un ligand avec une structure G-quadruplex et permet également de quantifier des différences entre structure G-quadruplex pour un même ligand. Cependant, il est important de noter que ces structures sont immobilisées sur une surface ce qui ne reflète pas le comportement de la structure *in vivo*. De plus, cette immobilisation ajoute une gêne stérique plus importante et seul un quartet externe est accessible par le ligand.

b) Titration calorimétrique isotherme

La technique de titration calorimétrique isotherme (ITC, *Isothermal Titration Calorimetry*) permet de déterminer les paramètres thermodynamiques d'une interaction en solution, généralement entre un ligand et des biomolécules ou entre deux biomolécules comme l'ADN quadruplex ^[114] ou des protéines.^[115] L'ITC est une méthode d'analyse quantitative qui donne directement accès à la

variation d'enthalpie (ΔH) et d'entropie (ΔS), à la stœchiométrie de l'interaction étudiée en solution et à la valeur de la constante d'affinité (K_a).

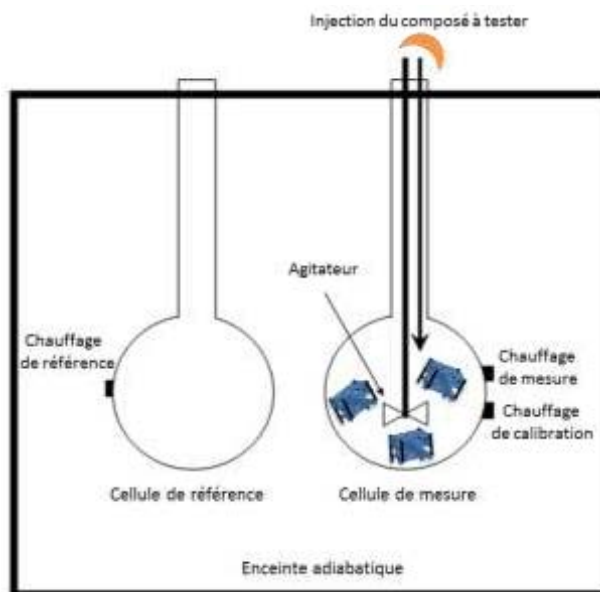


Figure 47 : Principe de l'expérience d'ITC pour l'interaction ligand/G-quadruplex.

Cette expérience est réalisée dans un calorimètre contenant deux cellules identiques incluses dans une enceinte adiabatique. En comparant la cellule de référence contenant du tampon avec la cellule de mesure contenant la biomolécule à étudier, les différences de température peuvent être évaluées. La cellule de référence est dans un premier temps, soumise à une puissance constante (< 1 mW), déclenchant le chauffage de calibration dans la cellule de mesure. Ensuite, le ligand à étudier est ajouté progressivement dans la cellule de mesure provoquant une variation de température du mélange (figure 47). Les mesures sont faites en évaluant la quantité de puissance nécessaire afin de maintenir les deux cellules à la même température en fonction du temps. A partir de ces mesures il est possible de déterminer la valeur d'enthalpie libre et donc de la constante d'affinité de la petite molécule, comme cela a été fait pour les composés PhenDC₃ et PDC-Br sur la séquence télomérique avec des valeurs de constante d'affinité de l'ordre de 10^6 mol⁻¹.^[116] Cette technique permet également de montrer qu'il n'y a qu'un site de fixation pour ces deux ligands sur l'oligonucléotide alors que le composé TMPyP4 en possède quatre.

Cette technique nous permet donc d'obtenir toutes les données thermodynamiques d'interaction entre le ligand et l'ADN cependant l'expérience s'effectue à des concentrations de ligands et d'ADN très élevées (respectivement 700 μ M et 20 μ M). De plus, c'est une méthode qui demande une expertise très poussée pour mettre au point l'expérience et que celle-ci soit reproductible.

c) Résonance magnétique nucléaire

La RMN est de plus en plus utilisée pour étudier les G-quadruplex en solution ainsi que leur dynamique et leurs interactions avec des ligands. La technique d'acquisition est cependant longue, délicate et nécessite de travailler à des concentrations élevées, de l'ordre du millimolaire. Le G-quartet possède une signature particulière en RMN. Le proton situé sur l'azote N1 des guanines, noté H1, possède un déplacement chimique autour de 13-14 ppm lorsqu'il est impliqué dans une liaison hydrogène de Watson-Crick et lorsqu'il est impliqué dans une liaison hydrogène de Hoogsteen il a un déplacement chimique autour de 10-12 ppm. Ces protons sont beaucoup moins échangeables avec le solvant lorsqu'ils sont impliqués dans un G-quartet. Les signaux RMN sont alors beaucoup plus fins pour une structure G-quadruplex.^[117] Le nombre de signaux des protons H1 correspond au nombre de guanines impliqués dans la structure. Cela permet d'évaluer le nombre de conformères présents en solution, comme par exemple avec la séquence télomérique du Tetrahymena.^[118]

Cependant, l'interprétation est très souvent difficile dû au large recouvrement des signaux lors de la présence de plusieurs conformations différentes en solution. La détermination par RMN d'une structure G-quadruplex nécessite la formation d'une espèce stable et majoritaire. Généralement, des légères modifications sont apportées dans les séquences d'acides nucléiques étudiées pour les rendre monomorphe et stable. Ces mutations sont réalisées au niveau des séquences flanquantes ou au niveau des nucléotides des boucles afin d'éviter d'affecter la topologie principale de la structure.

De nombreuses structures d'ADN quadruplex ont été résolues par RMN comme celles issues de la séquence télomérique humaine 22AG,^[119] de l'oncogène c-myc,^[120] des promoteurs ckit2^[121] et Bcl2^[122] et des minisatellites CEB1^[32b] et CEB25.^[123] Cette méthode d'analyse a également permis d'établir les interactions existantes entre une structure quadruplex et un ligand comme **PhenDC₃**^[70] (figure 48), un dérivé de **télomestatine**^[52] ou encore **TMPyP4**.^[124] Le spectre NOESY établie pour le complexe PhenDC₃-Pu24T a permis de montrer les interactions intermoléculaires entre le quartet externe du G-quadruplex de la séquence Pu24T et les protons des amides du ligand (figure 48). De plus, le marquage au carbone 13 (¹³C) a aussi permis de déterminer la position des quinoliniums sur les guanines par HSQC.

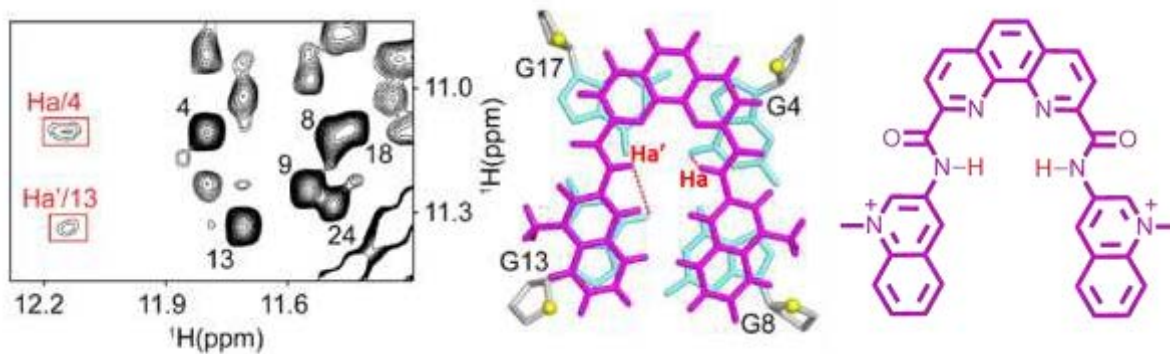


Figure 48 : Spectre NOESY montrant les interactions intermoléculaires entre le quartet de la séquence d'ADN Pu24T et les protons des amides du composé PhenDC₃. D'après [70].

Cette méthode nous permet donc d'avoir une idée de l'interaction entre le ligand et le G-quadruplex ce que ne nous donne pas les autres méthodes. Elle permet également d'avoir des données structurales précises des oligonucléotides formés en G4. Cependant, comme évoqué précédemment, cette technique nécessite de travailler à concentration élevée, de l'ordre du millimolaire, d'avoir un matériel adéquat et de ne pas avoir de mélange en solution pour pouvoir déterminer la structure. Une expertise dans le domaine est indispensable pour interpréter les spectres.

d) Spectrométrie de masse

La spectrométrie de masse à ionisation par électronébuliseur (ESI-MS) est une technique habituelle pour étudier les interactions non-covalentes. La facilité de formation des ions en fait une technique efficace. Les protéines et les acides nucléiques vont se charger positivement au niveau des acides aminés ou négativement au niveau des liaisons phosphodiesters.^[125] Cette méthode a été employée très fréquemment pour étudier les interactions de petites molécules avec les G-quadruplex.^[126] Ces analyses sont cohérentes avec des comportements observés en solution bien qu'elles soient réalisées en phase gazeuse. De plus, elles ne nécessitent qu'une faible quantité de matière. La spectrométrie de masse peut permettre de déterminer les constantes d'affinité et quantifier le ratio des espèces mais, pour cela, il est nécessaire d'utiliser des modèles mathématiques.^[127] Cette méthode d'analyse sert également à déterminer la structure globale du G-quadruplex étudié, la nature covalente ou non de l'interaction et la cinétique de formation du complexe.^[128] Pour cela, une grande maîtrise de la technique est nécessaire.

e) Diffractométrie de rayons X

La diffractométrie de rayons X est une technique très précise pour étudier la topologie d'un G-quadruplex à travers son cristal.^[119] Elle permet également, à l'échelle atomique, d'avoir des précisions sur les interactions ligand/G4 lorsqu'un monocristal est obtenu.^[36, 38, 124] Cette méthode

reflète donc un état condensé de la matière ce qui ne représente pas obligatoirement ce qu'il se passe en solution. Un cristal est soumis à des forces d'empilement qui peuvent avoir un effet sur les structures existant en solution, on ne peut pas appréhender les équilibres et le polymorphisme d'un G-quadruplex par cette analyse.

3) Analyses biochimiques

a) Inhibition de polymérases

En plus des méthodes biophysiques, des méthodes biochimiques étudiant l'inhibition de la télomérase ont été développées. La méthode la plus utilisée, et considérée comme une référence, est le test TRAP (pour *telomere repeat amplification protocol*). Ce test comporte deux étapes : la première permet l'élongation d'un fragment télomérique synthétique par la télomérase issue de l'extrait cellulaire étudié, et le second conduit à l'analyse des produits d'élongation de l'enzyme, par amplification PCR.^[129] Il existe plusieurs techniques de détection de ces produits d'amplification qui vont permettre une mesure semi-quantitative ou quantitative de l'activité télomérase initiale. Si la présence de structures G-quadruplex dans l'extrait cellulaire permet une inhibition significative de l'activité de la télomérase, celle-ci pourra être détectée par le test TRAP. Cependant, le test est susceptible de montrer de nombreux faux positifs, en présence de ligands de G-quadruplex pouvant inhiber l'enzyme polymérase *Taq* de l'amplification PCR.^[130]

La formation d'une structure G-quadruplex peut également être détectée au sein d'un brin plus long d'ADN grâce au test d'arrêt de la polymérase (ou *Polymerase Stop Assay*).^[131] Les polymérases sont des enzymes capables de synthétiser un brin de nucléotides, ADN ou ARN, en utilisant le brin complémentaire comme matrice et en assemblant des nucléotides triphosphates dans le sens 5'-3'. La présence d'une structure G-quadruplex sur le brin d'ADN servant de matrice va provoquer un arrêt inopiné de la polymérase et un raccourcissement du brin synthétisé. Le test de *Polymerase Stop Assay* développé par Weitzmann en 1996 a permis de détecter la formation de ces structures et de mettre au point la séparation de différentes conformations sur gel d'électrophorèse.

b) Influence des hélicases sur le déroulement des G-quadruplex

Les hélicases sont des enzymes qui représentent le moteur moléculaire des cellules. Elles sont capables de se coupler aux triphosphates de nucléosides hydrolysés afin de dérouler des acides polynucléiques tels que l'ADN et l'ARN. Il est estimé à 95 le nombre d'hélicases présumées codées par le génome humain, 31 hélicases d'ADN et 64 hélicases d'ARN.^[132] Le déroulement de l'ADN ou l'ARN double hélice par des hélicases intervient très souvent avant un mécanisme cellulaire. C'est pourquoi les hélicases sont impliquées dans pratiquement tous les métabolismes cellulaires comme

la réplication, la réparation d'acides nucléiques, la recombinaison, la transcription, la ségrégation des chromosomes et la maintenance des télomères.^[133] Elles permettent généralement la formation d'acides nucléiques simples brins mais elles peuvent également être capables de dérouler des formes secondaires des acides nucléiques comme les G-quadruplex.

Ces dernières années, la recherche d'hélicases capables de dérouler sélectivement les G-quadruplex a été très développée.^[134] Afin de les reconnaître, des expériences d'inhibition de l'activité de ces protéines ont été réalisées en présence de ligands stabilisant les structures G-quadruplex.^[135] Un second test a été développé très récemment dans l'équipe de Mergny pour montrer l'influence d'un ligand sur le déroulement du G-quadruplex par l'hélicase Pif1^[136] ou RHAU.^[137] Le principe est donc de mettre en présence un acide nucléique ayant une région simple brin, une région formant un G-quadruplex et une région duplex. Lorsque le duplex est formé, la fluorescence du fluorophore qui est sur le simple brin complémentaire est quenché. La première réaction consiste à lier l'hélicase à la partie 5' de l'oligonucléotide qui est en simple brin. De l'ATP est alors ajoutée ce qui entraîne le déroulement de l'acide nucléique par l'hélicase de la partie 5' vers le bout 3' en passant par le G4 en premier puis par le duplex. Une fois le duplex déroulé, la fluorescence du brin séparé est restaurée car il n'est plus quenché par le dabcyl (en violet sur la figure 49) porté par le brin complémentaire. En suivant la fluorescence, il est possible de savoir le pourcentage d'oligonucléotides déroulés. Pour évaluer des ligands, les petites molécules sont ajoutées avant d'ajouter l'ATP afin qu'il stabilise les G-quadruplex et ainsi rendre l'activité de l'hélicase faible voire même nulle. Par cette méthode, des ligands comme PhenDC₃ et pyridostatine (PDS) ont encore montré leur grande affinité pour les G-quadruplex.

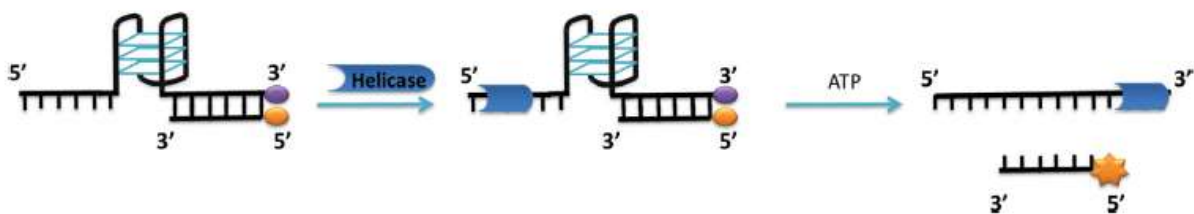


Figure 49 : Principe de déroulement d'un oligonucléotide par une hélicase.^[136]

Objectifs du travail de doctorat

Comme décrit dans l'introduction, la formation de structures G-quadruplex est impliquée dans de nombreux mécanismes au sein de la cellule : réplication, transcription, épissage et traduction. Toutes ces étapes du cycle cellulaire sont influencées par la présence de G-quadruplex. De plus, ces structures jouent également un rôle dans la destructuration des télomères par leur présence dans ces régions du génome. Ainsi, l'architecture de l'ADN peut aussi être affectée au niveau de la chromatine et une instabilité génétique et épigénétique importantes peuvent être provoquées. De nombreuses études ont été développées afin de cibler et localiser les G-quadruplex au sein du génome humain. L'utilisation d'outils fonctionnels nous a donné des indications importantes de leur présence *in vivo* cependant des doutes peuvent subsister. En effet, toutes les méthodes utilisées sont des stratégies indirectes (anticorps, marqueurs de dommages à l'ADN ou encore modification d'un génome pour montrer l'instabilité génétique). De plus, les études d'imagerie ont surtout été réalisées en cellules fixées ce qui ne présage pas toujours de ce qu'il se passe en cellules vivantes. Le panel d'outils moléculaires existant pour cibler les G-quadruplex n'est, aujourd'hui, pas suffisamment étoffé. Cela ne nous permet donc pas de cibler le peu de structures existantes au sein du génome à cause de la fugacité de celles-ci.

C'est pourquoi, plusieurs ligands fonctionnels ont été développés afin d'avoir des outils multifonctionnels susceptibles d'être utilisés dans des milieux biologiques complexes et de prouver l'existence de G-quadruplex dans ces milieux (figure 50). Parmi eux, des composés possédant un groupement photoactivable ont été synthétisés afin de lier de manière covalente le ligand et la cible G-quadruplex. Des composés biotinylés ont également été préparés pour immobiliser le ligand sur une surface telle que des billes recouvertes de streptavidine, des puces SPR ou encore des colonnes d'affinité ce qui permet à la fois l'étude de l'interaction G4/ligand et l'isolement de structures G-quadruplex à partir d'un milieu complexe.

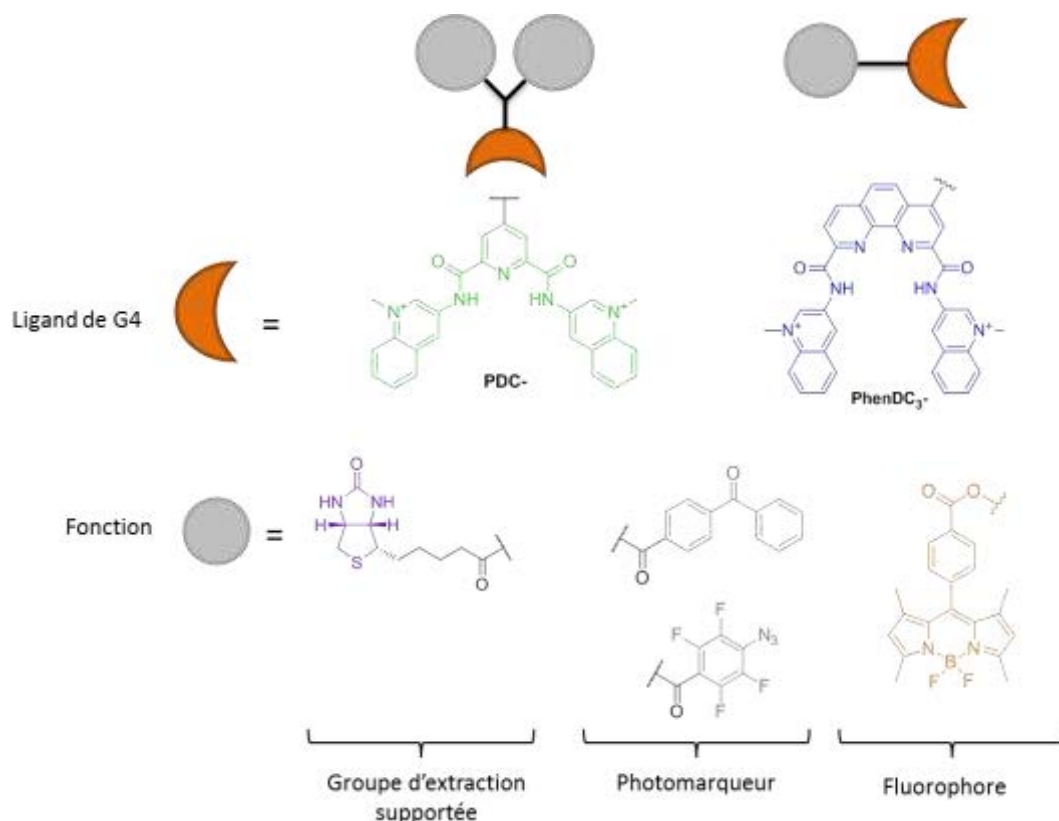


Figure 50 : Conception des outils moléculaires : Représentation schématique des composés bi- et tri-fonctionnels et structure des ligands et fonctions utilisés.

Une des limitations de la fonctionnalisation peut être la perte d'affinité ou même de la sélectivité des ligands pour les G-quadruplex. Deux approches seront donc explorées : d'une part, une fonctionnalisation avant la reconnaissance du ligand pour la cible, cette stratégie sera étudiée avec les molécules détaillées précédemment (figure 50) ; d'autre part, une fonctionnalisation *in situ*. Le ligand est positionné sur la cible et est capable de réagir par réaction bioorthogonale avec la fonction adaptée à l'expérience. La synthèse de ces nouveaux dérivés sera décrite. L'avantage de cette seconde stratégie est la diversité de fonctions que l'on peut utiliser. En effet, il est possible de fonctionnaliser par un fluorophore pour faire de l'imagerie cellulaire (figure 51), par un groupement biotine pour immobiliser les structures G-quadruplex sur une surface recouverte de streptavidine (figure 51) ou encore par des billes recouvertes de fonctions « clickables » pour avoir un lien covalent entre la molécule et les billes.

Un des défis majeurs est que ces composés substitués conservent leur affinité et leur sélectivité pour les G-quadruplex. Nous évaluerons donc ces dérivés par des méthodes biophysiques.

L'utilisation de la chimie bioorthogonale nous a principalement permis d'étudier le comportement de ligands de G-quadruplex au sein de cellules cancéreuses humaines fixées et vivantes et de se rendre compte que la localisation *in cellulo* n'est pas si simple à interpréter. En effet, la localisation d'un

ligand peut être biaisée par la réaction utilisée pour le fonctionnaliser. *A contrario*, la stratégie de chimie « click » ne nous a pas apportés de probantes conclusions pour l'immobilisation de structures G-quadruplex sur des surfaces.

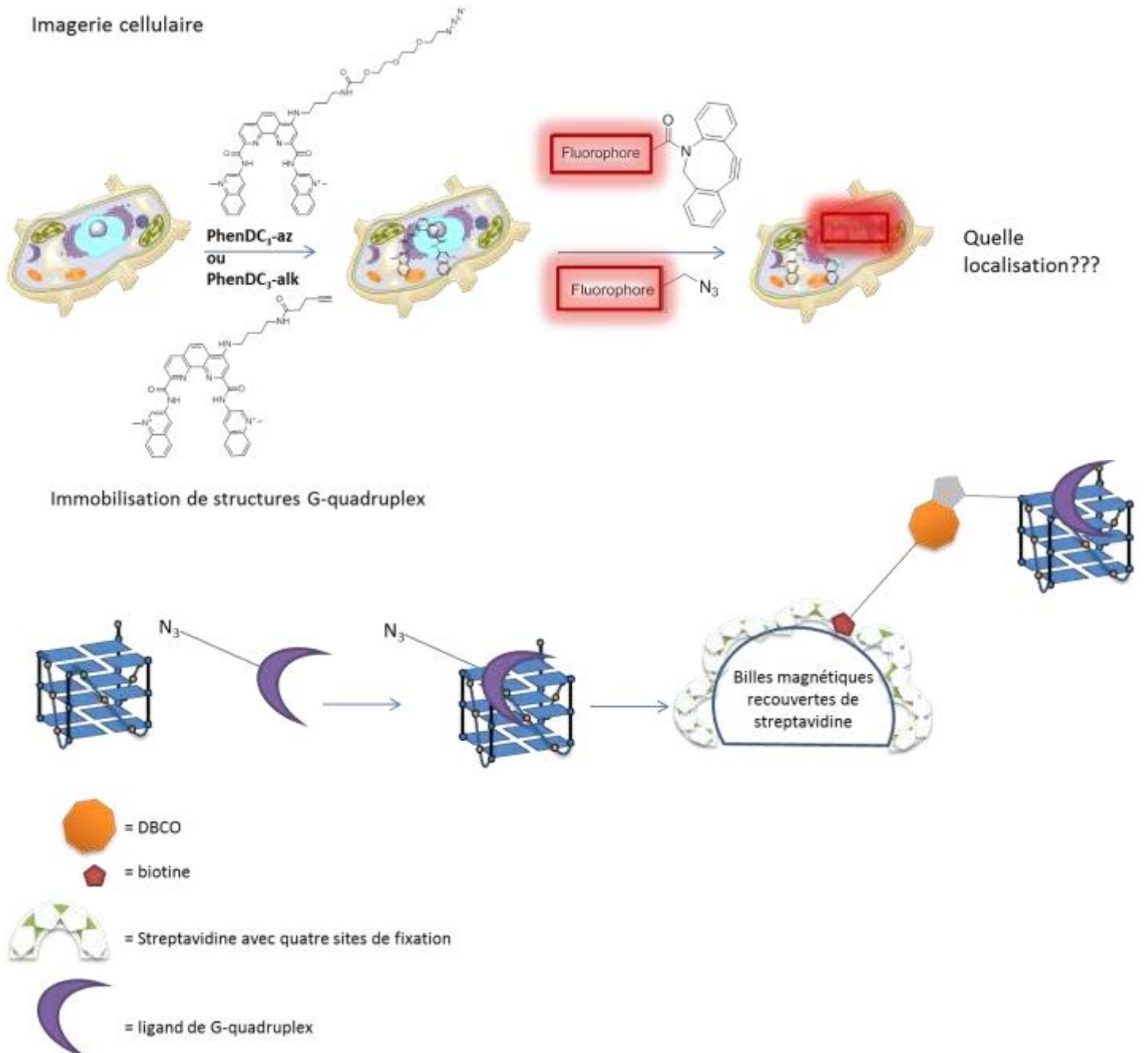


Figure 51 : Stratégie de visualisation de ligand de G-quadruplex *in cellulo* (en haut) et d'immobilisation de G-quadruplex sur des surfaces (en bas).

Chapitre 1 : Synthèse de ligands de G- quadruplex « clickables »

Comme évoqué précédemment, beaucoup de ligands existent pour stabiliser les G-quadruplex. Cependant, excepté les ligands fluorescents capables d'être observés *in cellulo*, peu de ces ligands ne sont fonctionnalisés afin d'être étudiés dans des milieux complexes. Au laboratoire, le panel de ligands développé a permis de trouver deux des principaux ligands capables d'interagir avec les G-quadruplex, PhenDC₃ et PDC (ou 360A). Ces ligands n'étant pas fluorescents, il a été décidé de les fonctionnaliser afin de les étudier en milieu biologique. Les dérivés bi- et tri-fonctionnels ont déjà été synthétisés au laboratoire (thèse d'Elodie Morel) et nous nous sommes intéressés au développement de dérivés alcyne et azoture qui peuvent réagir par réaction de chimie « click » pour être fonctionnalisé *in situ*.

Dans ce chapitre, la chimie « click » va donc être introduite puis la synthèse des ligands « clickables » va être décrite.

I. La chimie « click » comme stratégie de ciblage

1) Principe

La chimie « click » est un concept qui a été mis au point au début des années 2000.^[138] Il s'agit de concevoir d'une nouvelle façon la construction moléculaire afin d'obtenir des structures originales différentes des pharmacophores classiques. C'est une nouvelle approche de synthèse organique pour construire des molécules de manière simple et rapide à partir de petites unités moléculaires via des liaisons carbone-carbone (C-C) ou carbone-hétéroatome (C-X-C). Une réaction de « chimie click » doit répondre à différents critères pour avoir cette dénomination (figure 52). Tout d'abord, elle doit avoir de très bons rendements à partir d'un grand nombre de substrats différents, être facilement mise en œuvre et facile à purifier par des méthodes non chromatographiques comme la recristallisation. La réaction doit être insensible aux conditions de réactions (comme l'oxygène et l'eau) et aux conditions physiologiques ce qui signifie que la réaction doit se réaliser sans précaution particulière à prendre. Elle doit également être stéréospécifique et pour finir, elle ne doit pas utiliser de solvants organiques.

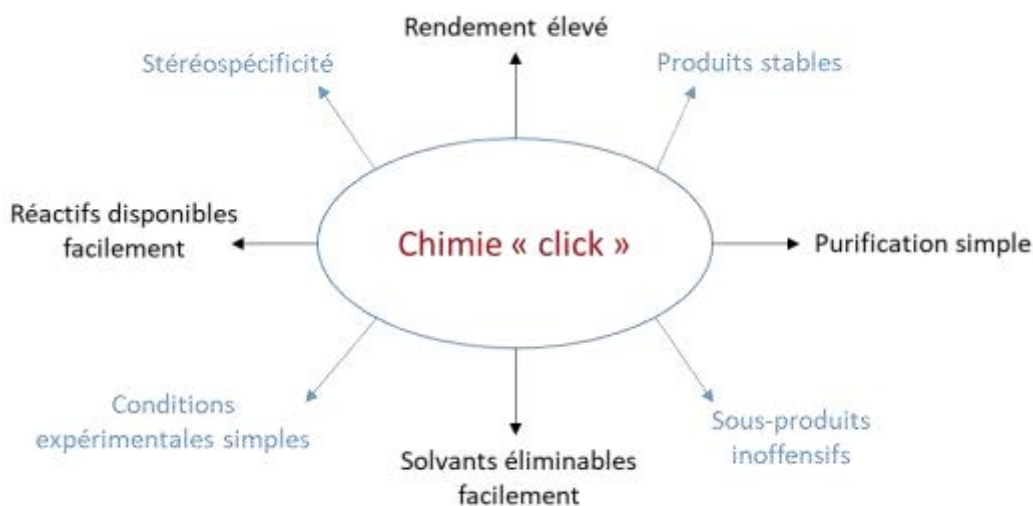


Figure 52 : Principe de la chimie "click".

De nombreuses réactions entrent dans ces critères et peuvent donc être considérées comme des réactions de chimie « click ». Elles nécessitent en général d'une forte enthalpie de réaction (> 20 kcal/mol) afin de conduire rapidement à un produit unique. Les réactions répondant à tous ces critères selon Sharpless sont, par exemple, les cycloadditions d'espèces insaturées (1,3-dipolaire, Diels-Alder), certaines substitutions nucléophiles (ouvertures d'hétérocycles) ou encore les additions sur des liaisons carbone-carbone insaturées (époxydation, dihydroxylation, additions de Michael...). Une des réactions la plus développée en tant que réaction de chimie « click » est la cycloaddition 1,3-dipolaire catalysée par le cuivre (I) entre un alcyne et un azoture (CuAAC). Elle a été découverte simultanément par les équipes de Meldal et Sharpless ^[139] en 2002 juste après l'introduction du principe de chimie « click ». C'est une réaction très simple à mettre en œuvre. Les différentes recherches effectuées sur le site pubmed nous montrent que, parmi les 6120 publications répondant au champs « click chemistry », 797 répondent au champ « copper catalyzed » ce qui représente environ 13 % des publications. Cela signifie que parmi le grand nombre de réactions de chimie « click », cette réaction fait partie des plus étudiées. Par la suite, le développement de cette même réaction sans utilisation de cuivre a été très étudié notamment par le groupe de Bertozzi.^[140]

2) Cycloaddition d'un azoture et d'un alcyne catalysée au cuivre (CuAAC)

Cette réaction est une réaction de Huisgen catalysée par le cuivre.^[141] La réaction de Huisgen est la cycloaddition 1,3-dipolaire entre un dipolarophile (alcène, alcyne, carbonyle, nitrile) et un composé 1,3-dipolaire (azoture, oxyde de nitrile, diazo-alcane) qui conduit à un hétérocycle à cinq atomes, un triazole ou un tétrazole.

Cette réaction de cycloaddition fut initialement découverte par Michael à la fin du XIX^{ème} siècle puis développée par Huisgen à partir de 1963. Elle requiert des températures élevées et n'est pas toujours stéréospécifique. La réaction spécifique entre un alcyne et un azoture conduit à un mélange d'isomère 1,4 et 1,5 du cycle triazole (figure 53).^[142]

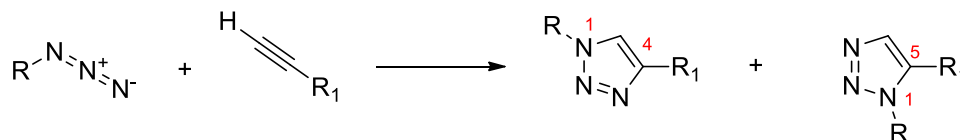


Figure 53 : Cycloaddition 1,3 dipolaire azoture-alcyne de Huisgen.

En 2001, cette réaction est alors considérée comme prometteuse pour être une réaction de chimie « click » malgré l'absence de régiosélectivité qui est à améliorer. Elle évolue par la suite en étant activée thermiquement par l'équipe de Sharpless pour la synthèse de composés tétrazoles à partir d'azoture et de nitrile (figure 54).^[143]

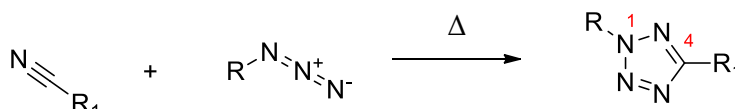


Figure 54 : Formation d'un composé tétrazole par cycloaddition 1,3-dipolaire entre un azoture et un nitrile.

La même année, l'équipe de Meldal utilisent des sels de cuivre (I) pour catalyser la réaction de cycloaddition entre un azoture et un alcyne^[139a] ce qui permet la synthèse, à température ambiante, de composés 1,2,3-triazoles substitués en position 1 et 4 de manière régiosélective : il s'agit du début de la réaction de CuAAC (pour *Cu-catalyzed Alkyne-Azide Cycloaddition*). Cette réaction était également en développement dans l'équipe de Sharpless qui a rapidement proposé un mécanisme.^[139b]

a) Mécanisme de réaction

Le cycle catalytique proposé par le groupe de Sharpless en 2002 commence par la formation d'un complexe cuivre-acétylure. Après ajout de l'azoture, un complexe cuivre-acétylure-azoture est formé puis il y a une cyclisation intramoléculaire pour former un métallo-cycle intermédiaire constitué de six atomes. Le triazole est alors formé par réarrangement puis la scission de la liaison carbone est cuivre permet la régénération du catalyseur et l'obtention du triazole non complexé (figure 55). Une étude théorique par DFT réalisée par la même équipe a permis de confirmer l'hypothèse de la formation de l'intermédiaire Cu-acétylure.^[144]

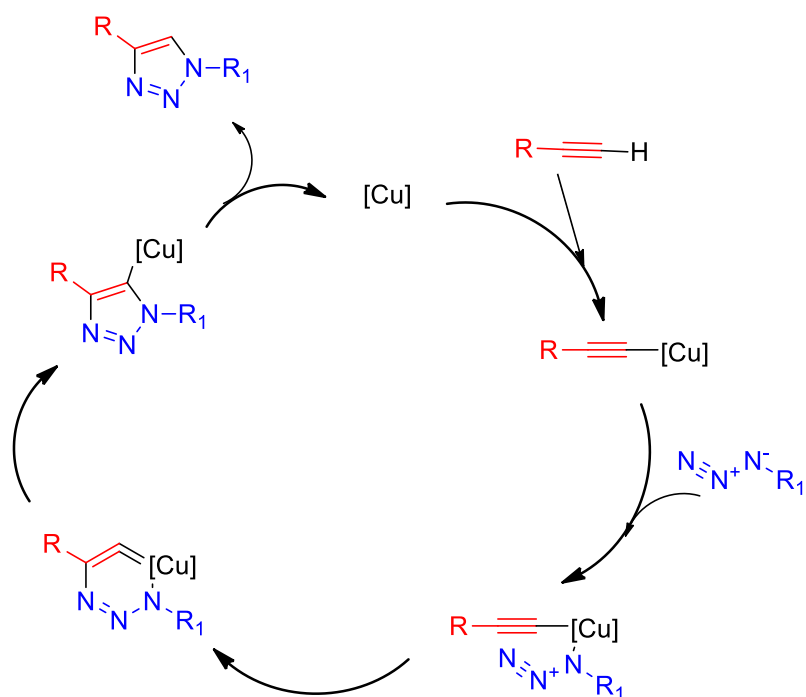


Figure 55 : Mécanisme de réaction de la réaction de CuAAC.

Ce mécanisme est précisé peu après par l'équipe de Finn ^[145] qui montre la participation d'un deuxième atome de cuivre dans le cycle catalytique (figure 56). Ce deuxième cuivre jouerait un rôle dans l'activation de l'azoture.

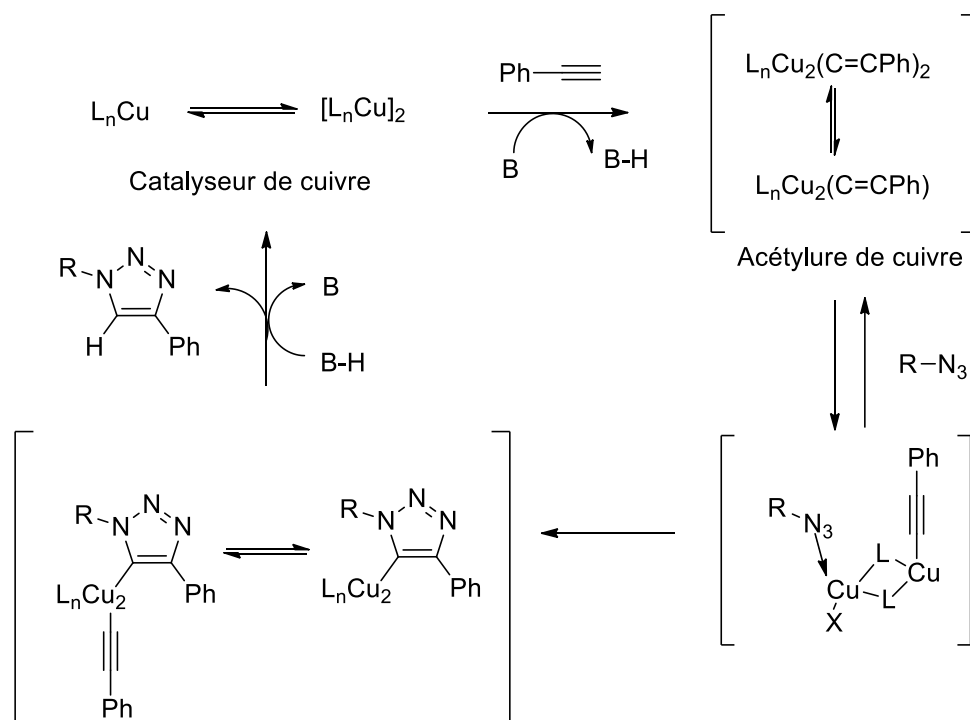


Figure 56 : Cycle catalytique mettant en jeu deux centres métalliques. D'après ^[145].

Plus récemment, en 2013, l'équipe de Finn a encore précisé le mécanisme en montrant les différentes liaisons créées entre le cuivre et l'alcyne. En effet, ils ont montré que les deux atomes de cuivre impliqués dans le mécanisme forment deux types de liaisons différentes, une liaison σ et une liaison π (figure 57).^[146]

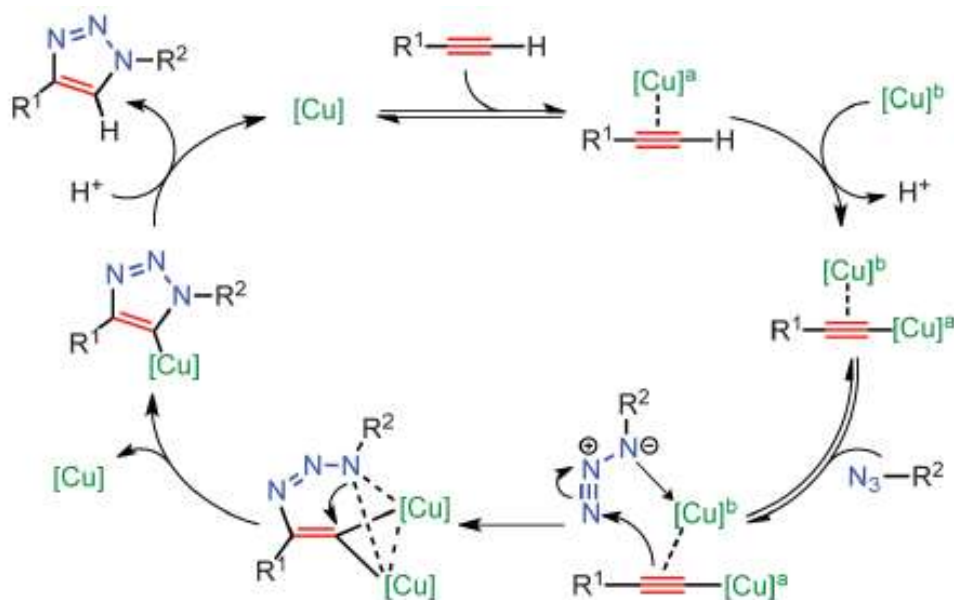


Figure 57 : Implication de liaisons σ et π entre le cuivre et l'alcyne dans le mécanisme réactionnel.

b) Utilisation de chélatants du cuivre

Afin d'améliorer la cinétique de réaction et de réduire la quantité de cuivre (I) qui est un élément toxique pour les cellules humaines à partir d'une certaine concentration,^[147] des ligands chélatants du cuivre ont été développés. Ces ligands vont stabiliser le cuivre (I) ce qui va permettre d'accroître la vitesse de réaction et améliorer les rendements (figure 58).^[148]

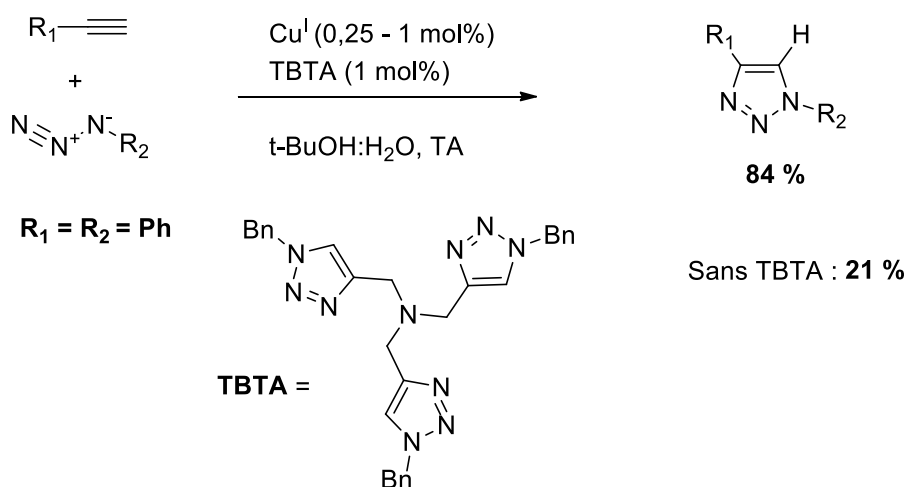


Figure 58 : Utilisation d'un ligand du cuivre (I) pour la réaction de CuAAC. D'après ^[148b].

Les ligands les plus performants pour stabiliser le cuivre (I) s'avèrent être des ligands oligotriazoles dérivés de composés type propargylamine et en particulier la tris(benzyltriazolylméthyle)amine (TBTA, figure 59). Ce ligand est très souvent utilisé dans des réactions de chimie « click » en milieu biologique ^[149] et parfois sans agent réducteur. ^[150]

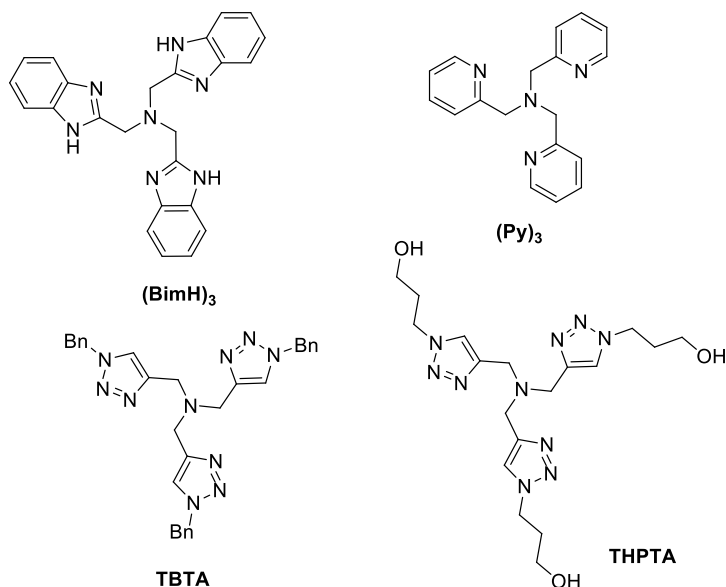


Figure 59 : Exemples de ligands du cuivre.

3) Cycloaddition d'un azoture et d'un alcyne contraint (SPAAC)

L'utilisation de catalyseurs à base de métaux de transitions comme le cuivre peut être problématique dans certains cas et notamment lorsque la cycloaddition envisagée doit se dérouler *in vivo*. Les ions cuivre sont en effet toxiques pour les organismes vivants au-delà d'un seuil de tolérance très faible. ^[151] Ainsi, des réactions alternatives de chimie « click » ont été développées visant à conserver l'efficacité et la régiosélectivité de la réaction catalysée au cuivre tout en s'affranchissant du catalyseur métallique. ^[152] Pour cela, les travaux développés visent à remplacer les alcynes terminaux par d'autres composés capable de réagir efficacement avec les azotures. ^[153]

Ainsi, l'équipe de Bertozzi s'est intéressée à la réaction de cycloaddition [3+2] entre des dérivés azotures et des cyclooctynes substituées. Cela montre la possibilité de s'affranchir du cuivre pour réussir à modifier de manière sélective une molécule en milieu biologique. ^[154] Bien que plus lente que la réaction catalysée au cuivre et non régiosélective, cette réaction présente les caractéristiques de chémosélectivité et biocompatibilité requises pour une réaction de chimie « click ». Cette réaction a été très étudiée pour le marquage de cellules vivantes notamment. ^[155] Elle est connue sous le nom de SPAAC (pour *Strain-Promoted Alkyne-Azide Cycloaddition*).

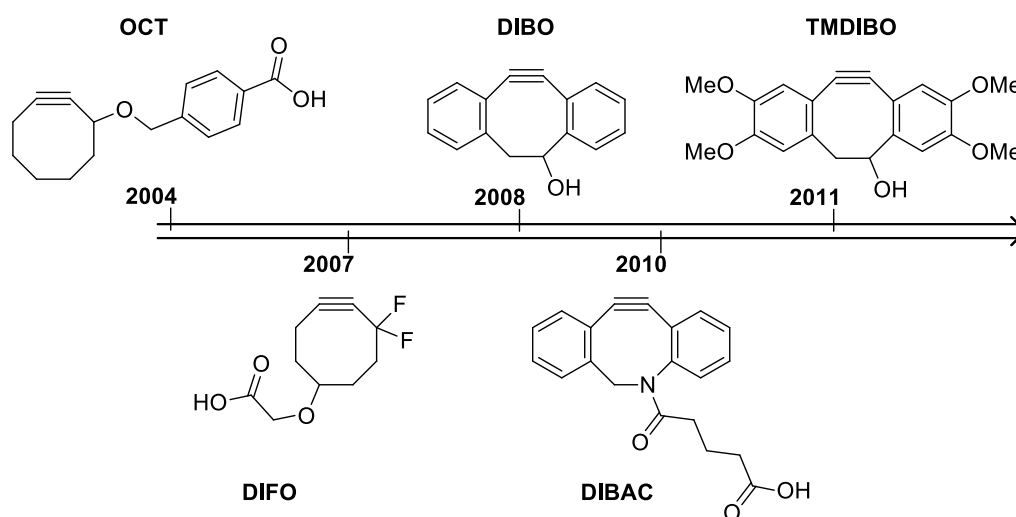


Figure 60 : Structures des différentes cyclooctynes développées pour améliorer la réaction de SPAAC. D'après [156].

Afin d'améliorer la cinétique de la réaction, de nombreuses cyclooctynes substituées ont été développées (figure 60).^[156-157] Cependant, la structure de l'azoture et le solvant utilisé peuvent également influencer la cinétique de la réaction. Il est devenu classique de fonctionnaliser des biomolécules avec une cyclooctyne (beaucoup sont aujourd'hui commerciales) afin de les utiliser *in vivo* et réaliser la réaction de chimie click *in situ*.

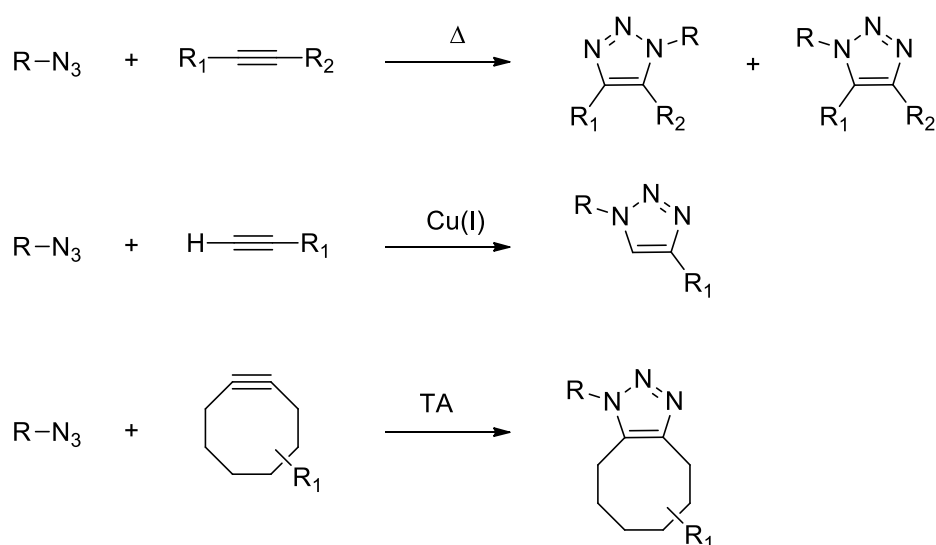


Figure 61 : Evolution de la réaction de Huisgen.

Cette réaction s'est beaucoup améliorée au fil des années (figure 61) pour être beaucoup plus efficace, elle est considérée aujourd'hui comme une réaction très intéressante pour une application *in vivo* de la chimie « click ».

4) Autres réactions de chimie « click » sans cuivre

Il existe un grand nombre d'autres réactions qui ne nécessitent pas de cuivre ni de catalyseur et considérées comme de la chimie « click ».^[156, 158]

Parmi elle, une réaction de rétro-Diels-Alder entre un oxanorborniadiène et un azoture (figure 62) qui donne lieu à la formation d'un triazole stable avec des azotes en position 1, 2 et 3. Cette réaction est plus rapide que la cycloaddition de Huisgen entre un azoture et un alcyne à température ambiante cependant, la réaction a un inconvénient majeur qui est non-négligeable pour la stratégie de chimie « click » qui est la formation de furane en tant que sous-produit.

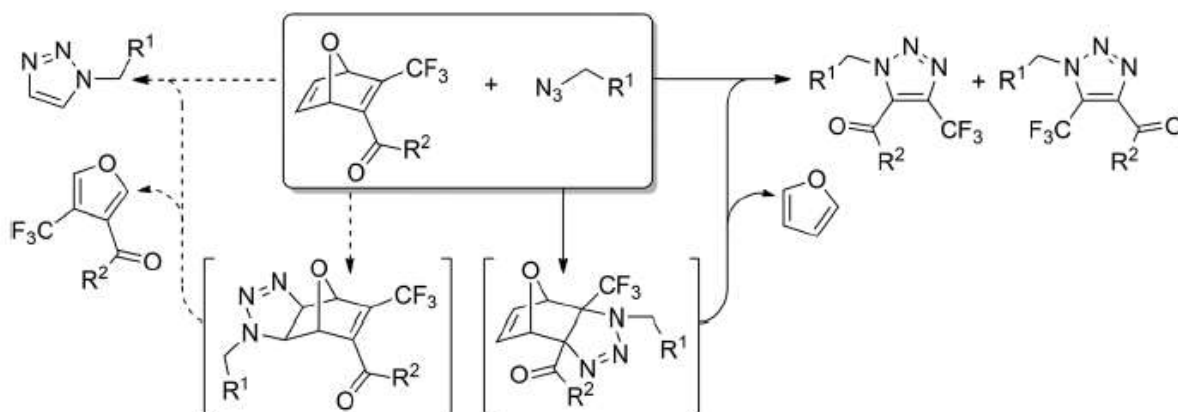


Figure 62 : Mécanisme de réaction de la réaction de rétro-Diels-Alder avec formation de furane. D'après ^[158b].

Une autre réaction de cycloaddition de Huisgen pour fonctionnaliser des polymères biocompatibles a été développée par les équipes de Lutz et Heaney.^[159] Il s'agit d'une réaction entre un alcyne et un oxyde de nitrile qui mène à la formation d'un cycle isoxazole (figure 63). Le principal avantage de cette réaction est qu'elle a lieu à température ambiante. Elle ne nécessite aucun catalyseur et a des rendements très élevés. De plus, elle est hautement régiosélective et favorise la formation de l'isomère 1,4 de l'isoxazole. Cette réaction peut donc être une alternative intéressante pour de nouvelles applications de chimie « click » sans utilisation de cuivre cependant les oxydes de nitrile n'étant pas très stables, cela peut être un inconvénient à son utilisation *in vivo*.

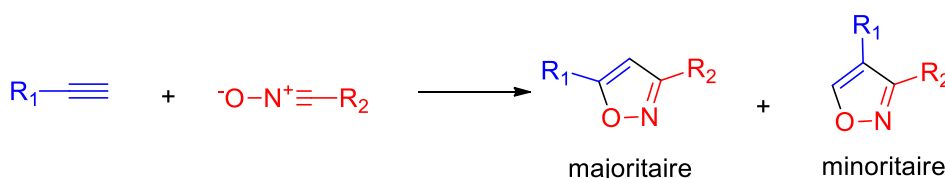


Figure 63 : Schéma réactionnel de la cycloaddition entre un alcyne et un oxyde de nitrile. D'après ^[159].

Enfin, une dernière réaction va être développée dans ce paragraphe même s'il existe bien évidemment beaucoup d'autres réactions, il s'agit de la réaction d'une tétrazine avec un alcène.^[160] Cette réaction de cycloaddition [4+2] n'est pas issue de la réaction de Huisgen et permet d'obtenir

par réaction de Diels-Alder inverse (diène pauvre en électron et diénophile riche en électron) une grande variété de structures (figure 64).

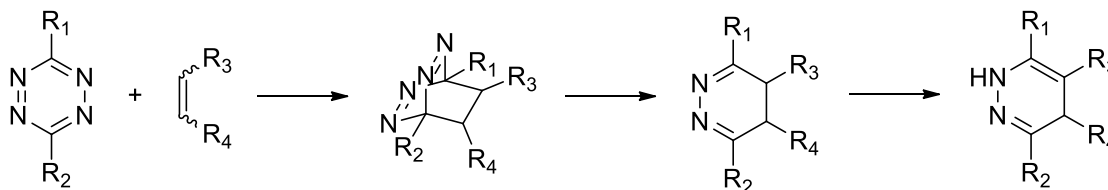


Figure 64 : Schéma réactionnel de la réaction d'une tétrazine avec un alcène.^[158b]

Les groupes de Hilderbrand et Fox ont ensuite étudié cette réaction pour des applications en milieu cellulaire. L'équipe de Hilderbrand a montré que les norborènes réagissent avec les tétrazines avec une vitesse de réaction similaire ^[161] aux réactions les plus rapides de SPAAC.^[162] Cette réaction a ainsi permis le marquage de cellules du cancer du sein. L'équipe de Fox a quant à elle montré la très grande vitesse de réaction des tétrazines avec les composés possédant une *trans*-cyclooctène, beaucoup plus rapide que la réaction de SPAAC.^[163] Cette réaction possède tout de même des inconvénients comme la difficulté de synthèse de ces tétrazines et leur stabilité.^[164] De plus, très peu de dérivés tétrazines ne sont, pour l'heure, disponibles commercialement.

II. Synthèse de dérivés du PhenDC₃

Le composé PhenDC₃ est la molécule référence, développée au laboratoire, pour étudier les G-quadruplex. A ce jour, elle fait partie des molécules ayant une très grande affinité pour ces structures. Ce composé permet de grandement les stabiliser.

Suite à une collaboration avec l'équipe du professeur Anh Tuan Phan à Singapour, une structure RMN du complexe ADN/ligand a été établie.^[70] De là, nous pouvons constater un recouvrement quasi parfait du quartet externe du G-quadruplex par le ligand. Cette molécule n'est que très faiblement fluorescente ainsi il n'est pas possible de suivre son évolution *in cellulo*. Il a donc été décidé de dériver la molécule pour y ajouter une fonction fluorescente. Elle sera ajoutée en position 4 de la phénanthroline car d'après la structure RMN (figure 65), il s'agit de la position qui devrait le moins gêner l'interaction ligand-G-quadruplex.

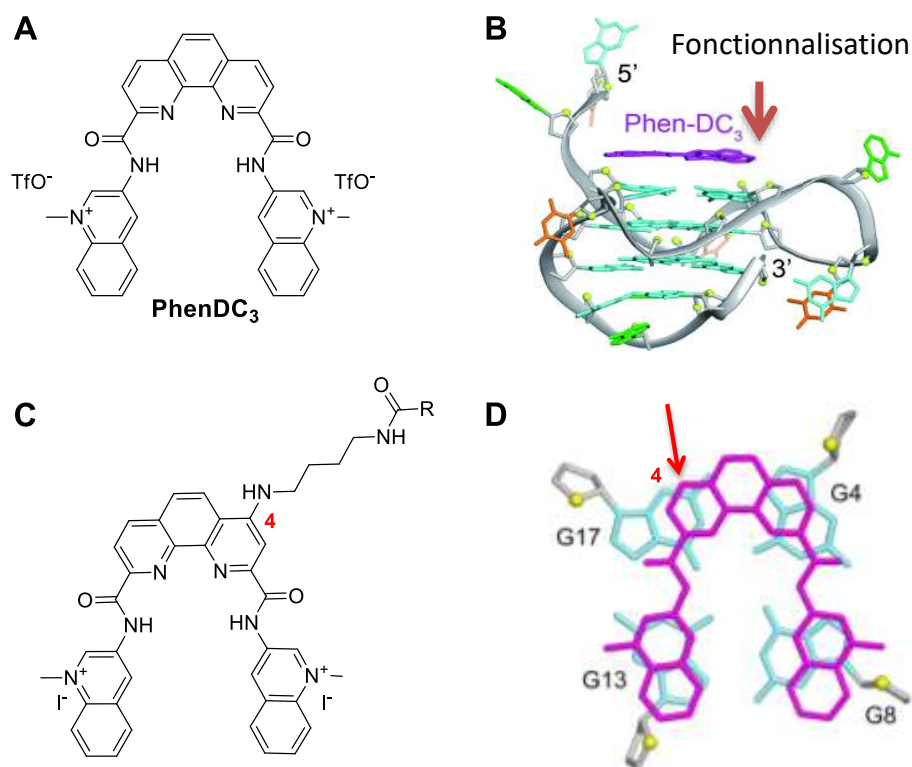


Figure 65 : Structure du composé PhenDC₃ (A) et de ses dérivés (C) et structure RMN du complexe PhenDC₃/Pu24T (B) vue latérale, (D) vue du dessus.

La stratégie envisagée est donc de synthétiser des composés bi- et tri-fonctionnels à partir du composé PhenDC₃ après l'avoir lié au G-quadruplex. Pour cela, nous utiliserons la chimie « click » et plus particulièrement la cycloaddition 1-3 dipolaire de Huisgen. Une simple fonction alcyne ou azoture doit être substituée sur la phénanthroline du composé PhenDC₃.

1) Précurseur commun à tous les dérivés du PhenDC₃

Pour synthétiser ces composés, il a fallu tout d'abord réaliser la synthèse du composé PhenDC₃-Cl qui est le précurseur commun de tous les composés fonctionnalisés (figure 66).^[165]

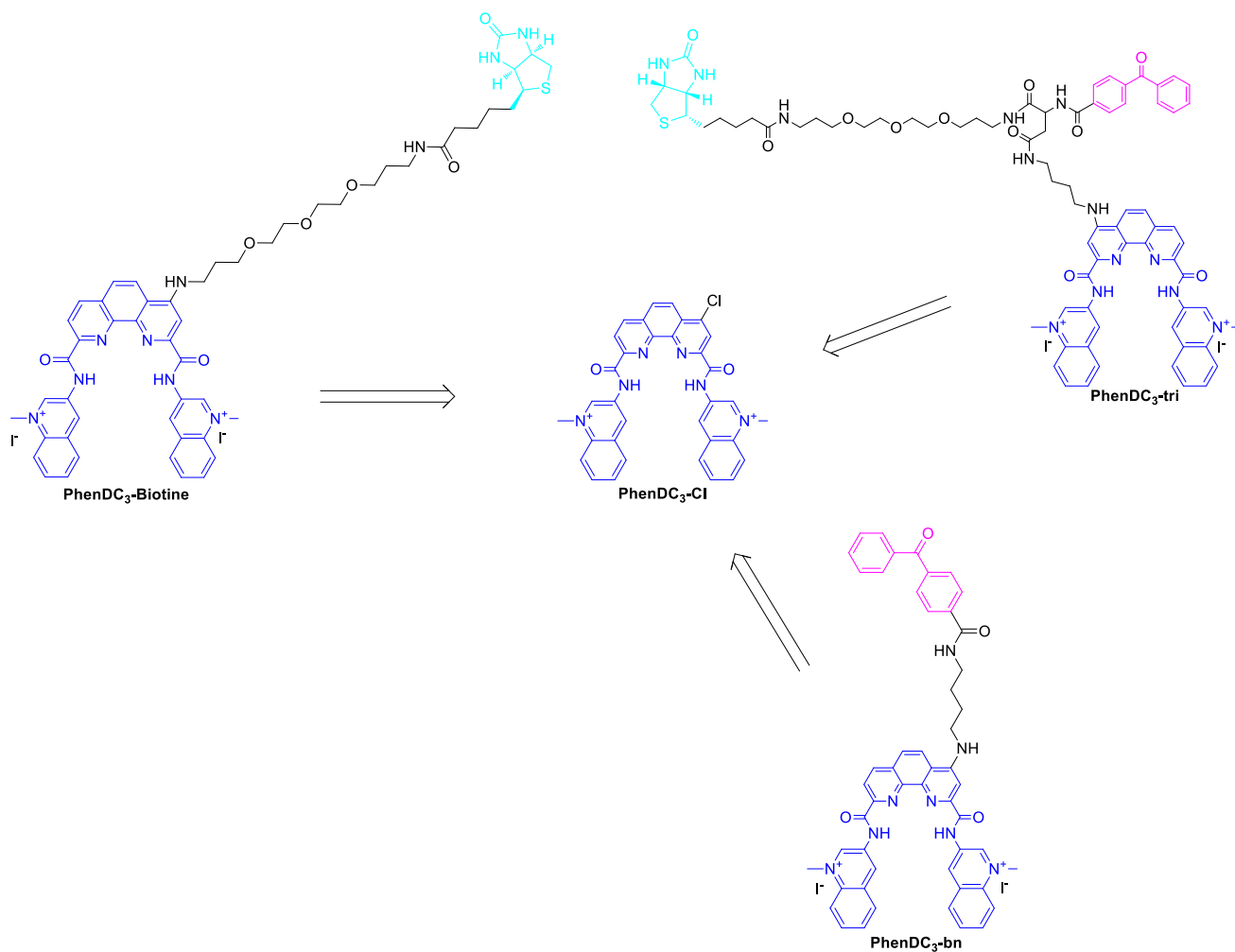


Figure 66 : Structures des dérivés du PhenDC₃ et du précurseur commun.

Nous avons repris la synthèse utilisée au laboratoire pour l'obtention des dérivés bi- et tri-fonctionnels. Celle-ci repose sur une adaptation de la synthèse de l'équipe de Ulven ^[165] qui permet d'obtenir des phénanthrolines symétriques di-substituées (figure 67). Ce noyau phénanthroline est formé par réarrangement et cyclisation du composé synthétisé après réaction du triméthylorthoacétate, de l'acide de Meldrum et du 1,2-diaminobenzène.

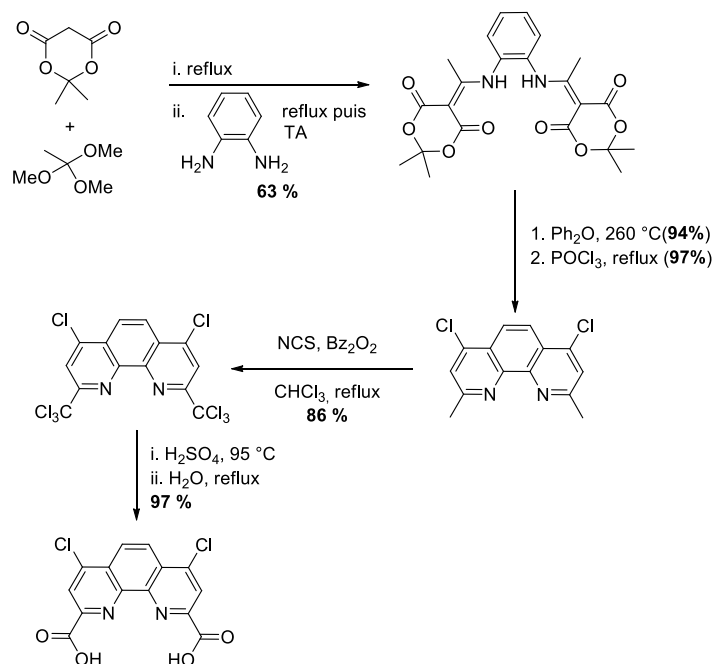


Figure 67 : Synthèse du groupe de Ulven.^[165]

La principale modification est la synthèse d'un ligand non-symétrique avec une seule fonctionnalisation en position 4. Pour cela, le substrat avant formation de la phénanthroline fonctionnalisé est différent de celui utilisé par le groupe d'Ulven. Dans notre cas, une quinoléine est alors utilisée pour la synthèse du cœur phénanthroline (figure 68).

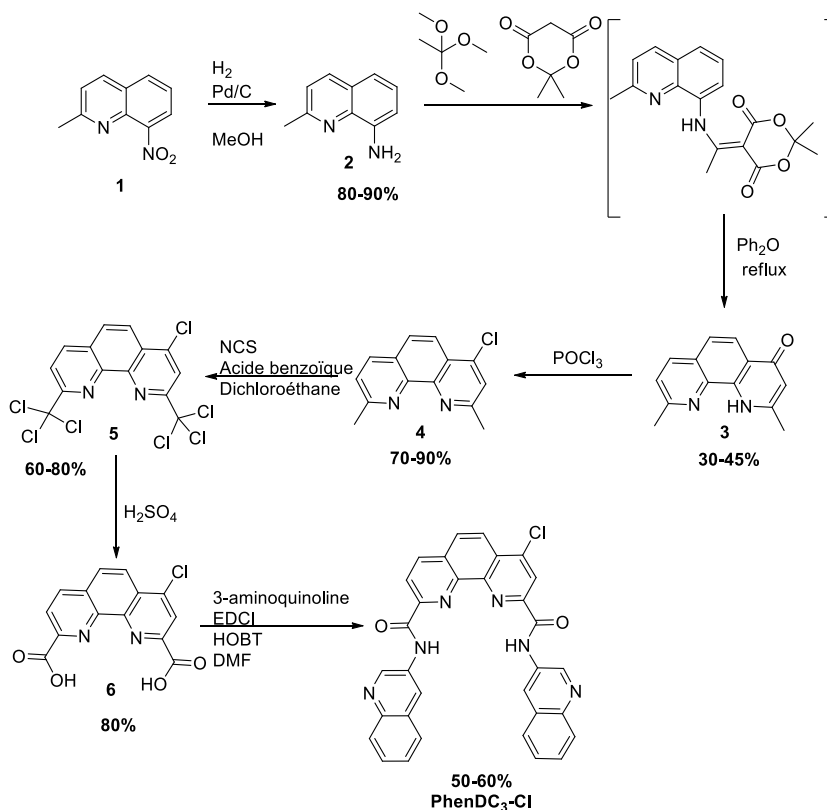


Figure 68 : Schéma de synthèse du composé "PhenDC3-Cl"

Afin d'obtenir ce dérivé, différentes étapes sont nécessaires. Tout d'abord, la réduction de la fonction nitro en amine du composé 2-méthyl-8-nitroquinoline (**1**) permet d'obtenir le composé 2-méthylquinoline-8-amine (**2**) avec de très bons rendements (> 80 %). La condensation du composé **2** avec le triméthylorthoacétate et l'acide de Meldrum suivie de la cyclisation et décarboxylation par reflux dans l'éther diphenylique du composé fraîchement synthétisé permet d'obtenir le composé 2,9-diméthyl-1,10-phénanthroline-4-one **3**. Cependant, le rendement de cette étape de cyclisation pour la formation de la phénanthroline s'est avérée faible (30%), une alternative à celle-ci a alors été recherchée. Une autre méthode de formation de la phénanthroline a été décrite par le groupe de Chen en utilisant une aminoquinoline substituée en position deux par un groupement methoxycarbonyle et en position quatre par un groupement iso-butoxy (figure 69).^[166]

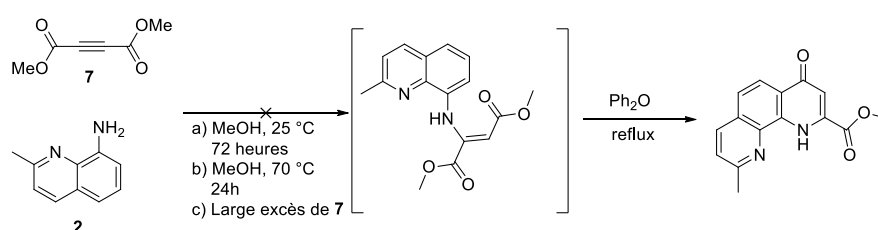


Figure 69 : Alternative exploitée pour la formation de la phénanthroline avec quelques conditions opératoires utilisées.^[166]

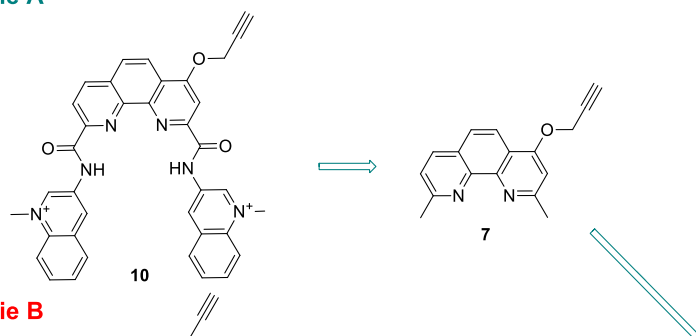
Cependant, le produit attendu n'est observé ni par LC-MS ni par RMN malgré différentes tentatives en modifiant les conditions opératoires. Les produits de départ sont toujours observés que la réaction ait lieu à température ambiante ou à 70 °C. C'est pourquoi la stratégie de synthèse est conservée telle qu'elle existait auparavant. Une fois le composé **3** obtenu, la fonctionnalisation de la phénanthroline en position 4 s'effectue en présence de trichlorure de phosphoryle pour la substituer par un chlore avec des rendements variant entre 70 et 90 %. Il est nécessaire, par la suite, de coupler les quinoléines sur la phénanthroline pour conserver le corps du PhenDC₃. Pour cela, les deux fonctions méthyles en position 2 et 9 de la phénanthroline du composé **4** réagissent en présence de N-chloro-succinimide afin d'obtenir le composé **5** hexachloré avec de bons rendements entre 60 et 80 %. L'ajout d'acide sulfurique sur ce dernier permet la formation du diacide 4-chloro-1,10-phénanthroline-2,9-dicarboxylique (**6**) qui est ensuite couplé aux quinoléines pour former **PhenDC₃-Cl** avec des rendements de 50 à 60 %.

2) Précurseur de ligands bi-fonctionnels

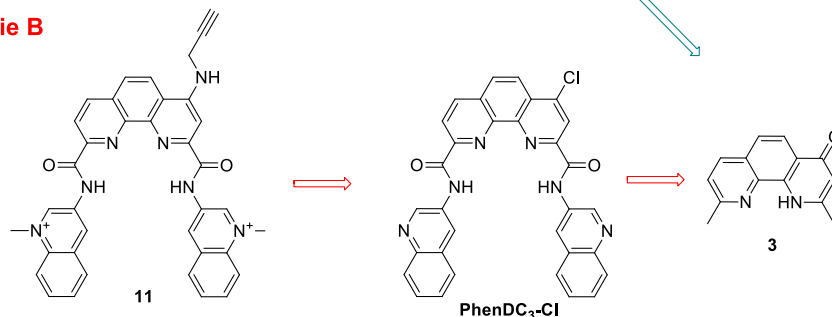
La fonctionnalisation de la phénanthroline a été étudiée sur différents intermédiaires : tout d'abord, voie **A**, sur le composé **2,9-diméthyl-1,10-phénanthroline-4-one (3)** lorsque la phénanthroline est tout juste formée (figure 70A), ensuite, voie **B**, sur le composé **PhenDC₃-Cl** avec substitution de la phénanthroline par la fonction souhaitée (figure 70B) et enfin, voie **C**, toujours sur le composé

PhenDC₃-Cl mais en deux étapes, une première permettant de substituer la phénanthroline par une diamine et une seconde permettant de coupler cette nouvelle amine avec un acide par couplage peptidique (figure 70C).

Stratégie A



Stratégie B



Stratégie C

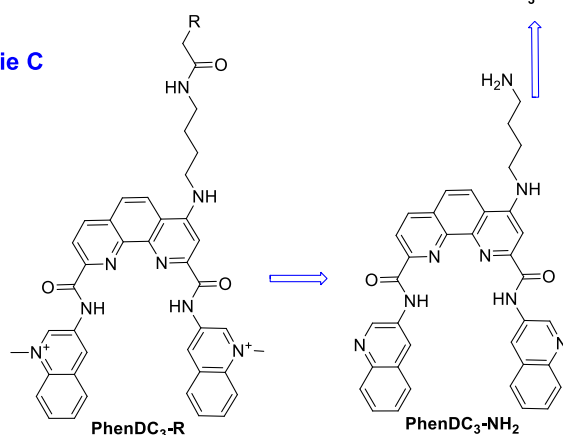


Figure 70 : Schéma rétro-synthétique des trois stratégies de fonctionnalisation.

La première stratégie de synthèse (figure 71) permet la fonctionnalisation de la phénanthroline avec un bon rendement (72 %). L'étape de chloration des groupements méthyles s'effectue avec des rendements d'environ 50 %. Cependant, l'étape d'acidification en présence de l'alcyne est impossible ce qui ne nous permet pas d'obtenir le composé final attendu. Il a donc fallu s'orienter sur une autre voie de synthèse.

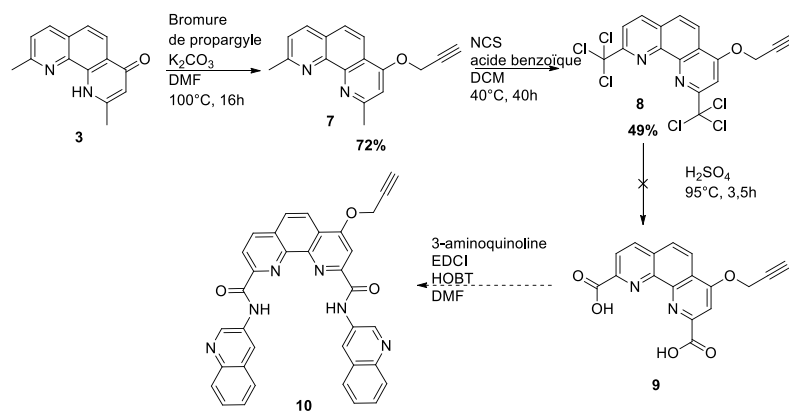


Figure 71 : Première stratégie de fonctionnalisation

Ainsi, le deuxième schéma de synthèse (figure 72) propose la fonctionnalisation du composé **PhenDC₃-Cl** par la propargylamine.

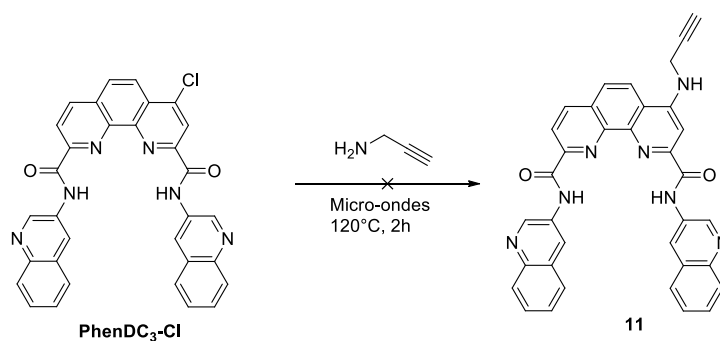


Figure 72 : Deuxième stratégie de fonctionnalisation

Cette méthode n'a pas été concluante car il est difficile de substituer le chlore en position 4 de la phénanthroline par une fonction amine. Cette difficulté peut s'expliquer par l'insolubilité du composé **PhenDC₃-Cl**. De plus, même lorsque la propargylamine est utilisée comme solvant, une faible conversion est observée par suivi HPLC.

La troisième stratégie (figure 73) consiste à introduire le composé 1,4-diaminobutane sur le composé chloré. Cette fois-ci, la réaction aux micro-ondes, comme décrit dans la littérature, ^[165, 167] en présence du composé 1,4-diaminobutane comme solvant permet l'obtention du composé avec un rendement de 70 %. Il s'agit donc de la stratégie retenue.

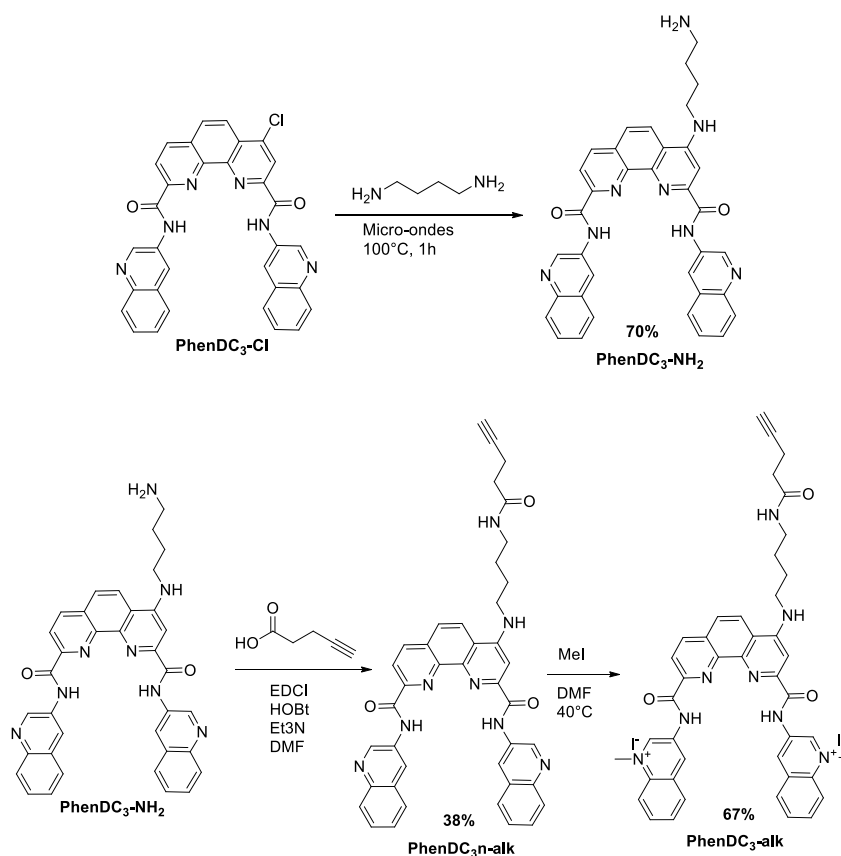


Figure 73 : Troisième stratégie de fonctionnalisation

Cette troisième stratégie est celle utilisée pour synthétiser les ligands bi- et tri-fonctionnels déjà existants au laboratoire. Le chlore en position 4 de la phénanthroline est substitué par une diamine pour former le composé **PhendC₃-NH₂** avec un bon rendement de 70 %. Ensuite, cette amine réagit avec un acide fonctionnalisé par couplage peptidique. C'est ainsi qu'il est possible de fonctionnaliser le composé PhendC₃ par tout type de fonction (ici, un alcyne). Les rendements varient en fonction de l'acide couplé entre 38 et 69% respectivement pour l'alcyne (figure 8) et l'azoture (figure 9). La dernière étape consiste bien évidemment à méthyler les unités quinoléines en présence d'iodure de méthyle avec des rendements de 67 %.

Le composé PhendC₃-az (figure 74) possède une chaîne plus longue que le composé alcyne car l'acide-azoture utilisé est commercial et disponible chez un fournisseur classique (TCI chemicals) à un prix raisonnable. Il serait également intéressant de préparer l'azoture avec la même longueur de chaîne que celle de l'alcyne mais, par manque de temps, cela n'a pas encore été réalisé. Par la suite, les résultats *in vitro* et *in cellulo* montreront que la longueur de la chaîne ne modifie pas le comportement de la molécule, c'est pourquoi il n'y a pas eu d'étude précise de l'effet de la chaîne. Par ailleurs, pour la réaction de SPAAC, un dérivé possédant une cyclooctyne peut être envisageable cependant, par manque de temps et au vu des résultats des tests biophysiques effectués sur les dérivés bi- et tri-fonctionnels du PhendC₃ (voir chapitre 2) un tel dérivé n'a pas été synthétisé.

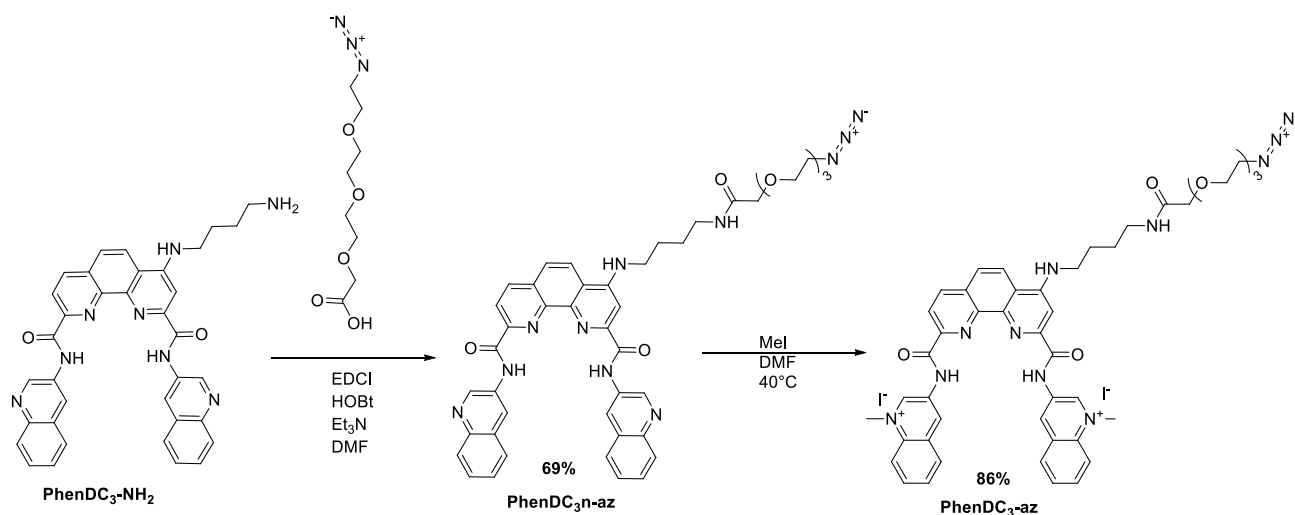


Figure 74 : Synthèse du composé PhenDC₃-az.

III. Synthèse de dérivés du PDC

Contrairement aux dérivés du PhenDC₃ qui n'ont pas été décrits, mis à part les ligands « clickables » suite à ce travail de thèse,^[168] différents dérivés du composé PDC (ou 360A) ont été publiés (quelques exemples figure 75). En effet, deux dérivés couplés à un fluorophore (le thiazole orange) ont été étudiés, PDC-L-TO et PDC-M-TO,^[169] afin de créer des ligands de G-quadruplex qui sont fluorescents uniquement lorsqu'ils sont sur leur cible. Il existe également un autre dérivé possédant une fonction fluorescente Bodipy qui a été synthétisé et étudié sur des billes magnétiques.^[170] Cette molécule a également été étudiée en cellules humaines (résultats non publiés) mais la localisation semble être dirigée par le fluorophore. Le fluorophore seul se localise dans les vésicules de sécrétion et lorsque celui-ci est couplé à un ligand de G-quadruplex comme PDC, il se situe toujours dans ces vésicules. Ensuite, des dérivés photoactivables avec des longueurs de chaînes plus ou moins longues entre la fonction photoactivable et la pyridine de PDC, PDC-benzophénone ou PDC-N3, ont été préparés.^[171] Ces composés ont permis de montrer sur quelle base de la boucle du G-quadruplex se liait de manière covalente le ligand et l'intérêt de la longueur de l'espaceur. Enfin, un dérivé biotinylé a également été étudié en collaboration avec l'équipe de Sugiyama pour de l'imagerie par AFM.^[172] Ce composé a également montré ses limites avec l'expérience Selex utilisée par l'équipe de Mergny qui n'a pas permis d'identifier des séquences d'ADN se liant à ce ligand.^[173] Ces dérivés conservent globalement une bonne sélectivité pour les G-quadruplex et que certains d'entre eux ont perdu légèrement en affinité. Nous avons alors envisagé une nouvelle stratégie utilisant la chimie bioorthogonale en synthétisant un dérivé alcyne.

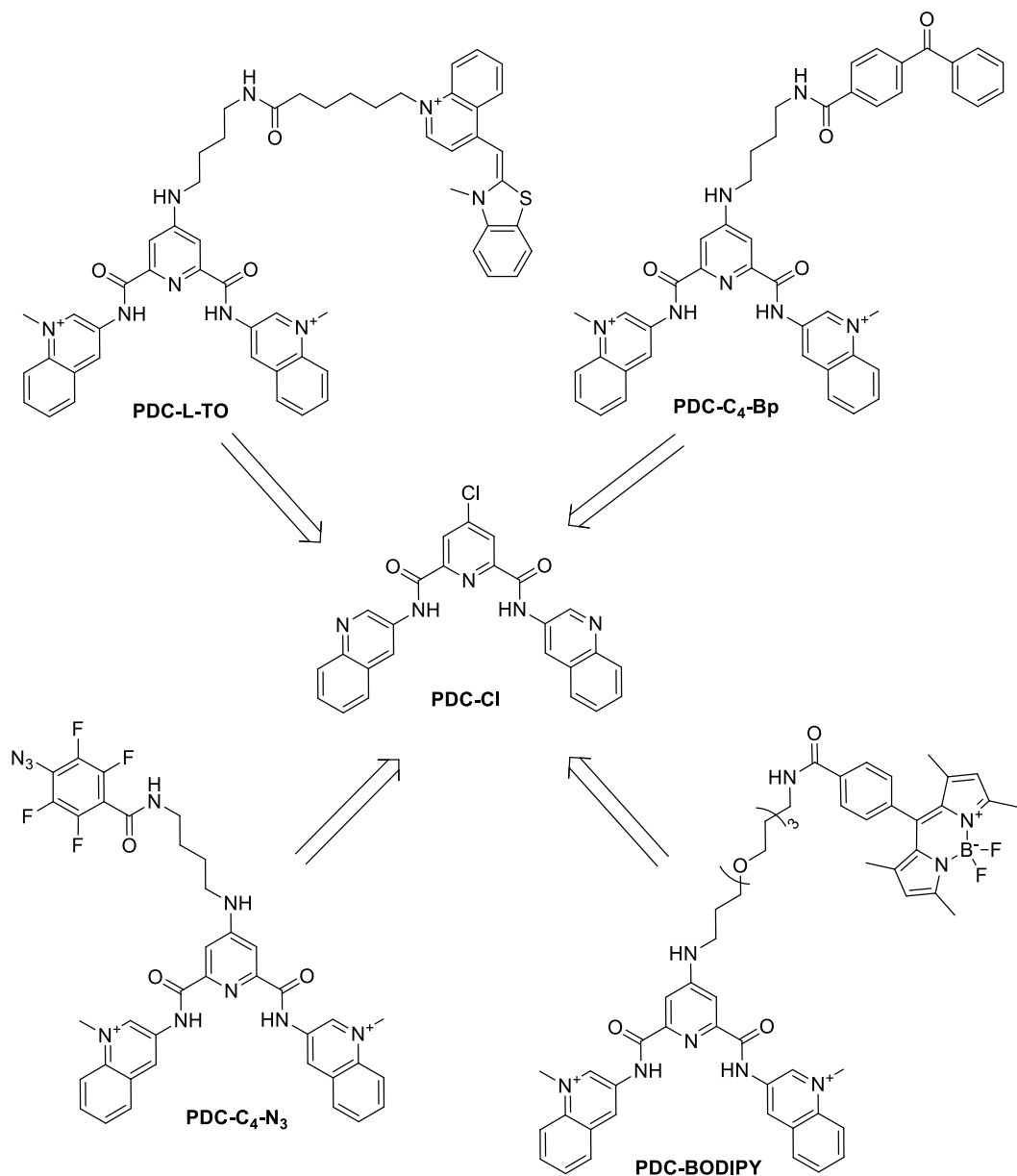


Figure 75 : Structures des dérivés du PDC et du précurseur commun.

Ainsi, pour nombre de ces dérivés, la synthèse est identique jusqu'au composé PDC-Cl. Ensuite, comme pour les dérivés du PhenDC₃, le chlore porté en position 4 de la pyridine est substitué par une diamine. La fonction envisagée est alors ajoutée par couplage peptidique. La méthylation des quinoléines est encore une fois la dernière étape de synthèse (figure 76).

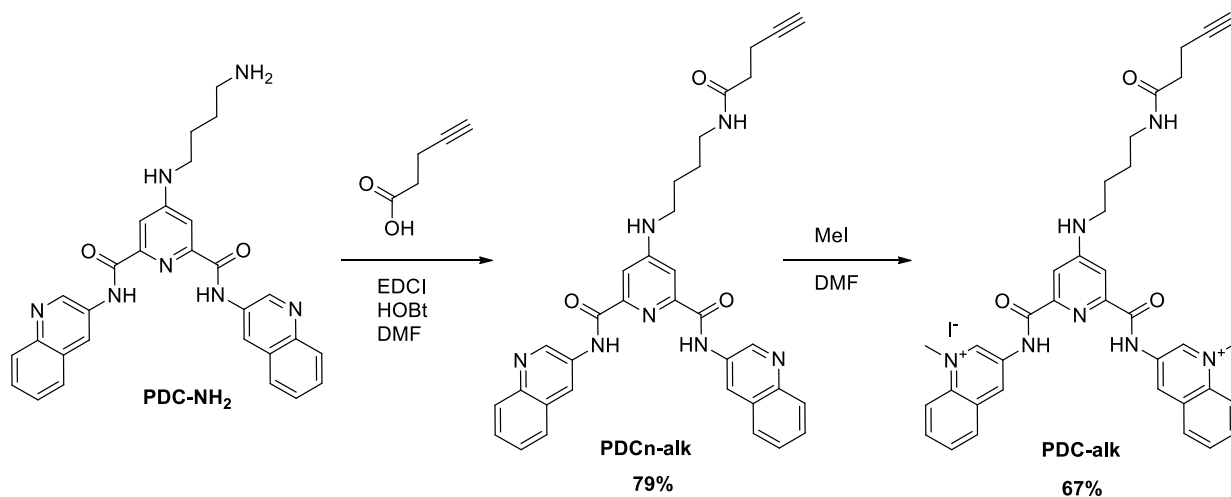


Figure 76 : Synthèse du PDC-alk.

Ainsi, le composé PDCn-alk est obtenu avec un bon rendement de couplage de 79 %. Celui-ci est alors méthylé pour obtenir PDC-alk avec un rendement de 67 %.

Tous ces nouveaux composés synthétisés ainsi que les ligands bi- et tri-fonctionnels vont pouvoir être évalués par les deux méthodes biophysiques utilisées en routine, l'expérience de FRET-melting et l'expérience de FID.

Chapitre 2 : Comparaison par analyses biophysiques *in vitro* de ligands de G- quadraplex

Dans ce chapitre, les molécules synthétisées et une sélection de ligands de référence vont être analysées par deux tests biophysiques. Ces deux évaluations sont effectuées en routine, il s'agit de l'expérience de FRET-melting et l'expérience de FID. Les résultats nous permettront d'effectuer une classification des ligands en fonction de leur affinité pour chaque séquence d'ADN G-quadruplex étudiée. De plus, nous couplerons ces données avec des valeurs de constante de dissociation obtenues par SPR dans le groupe d'Eric Defrancq afin d'établir une corrélation entre la constante de dissociation et les données biophysiques. Par ailleurs, une seconde corrélation a été explorée entre la valeur de température de fusion d'un oligonucléotide en présence de ligand et la valeur d'instabilité génétique générée au sein des levures par la présence du même ligand. Ces travaux de recherche ont été menés en collaboration avec l'équipe d'Alain Nicolas à l'Institut Curie à Paris.

I. Expérience de FRET-melting

1) Séquences d'ADN utilisées

Les séquences d'ADN utilisées sont marquées en 5' par une fluorescéine (FAM) et en 3' par une rhodamine (TAMRA) (figure 77).

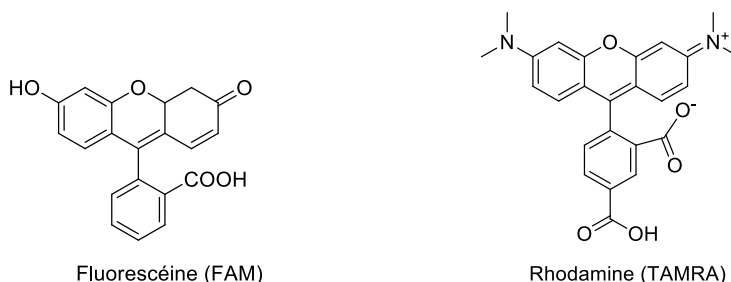


Figure 77 : Fluorophores utilisés pour l'expérience de FRET-melting.

Nous nous sommes intéressés principalement à huit séquences (figure 78) pour l'étude des ligands avec les G-quadruplex : la séquence télomérique (22AG ou F21T) qui a été la première étudiée et dont la structure RMN est connue ;^[119] deux séquences modifiées de l'oncogène c-myc (Myc22 ou F-myc-T et Pu24T) possédant également une structure RMN définie ;^[70, 174] trois séquences du minisatellite CEB25 ont été choisies : la séquence native possédant une grande boucle (CEB25-WT (L191)) et deux séquences modifiées possédant des boucles d'une ou deux bases (CEB25-L111(T) et CEB25-L121(TT) respectivement) ; une séquence du minisatellite CEB1 a également été étudiée. Les connaissances *in vivo* acquises en collaboration avec l'équipe d'Alain Nicolas sur ces minisatellites nous ont incités à étudier leur comportement par une analyse biophysique.^[32a, 32c] De plus, leurs structures ont été définies par RMN.^[32b, 123] L'équipe d'Alain Nicolas a montré que les G-quadruplex à grandes boucles ne se formaient pas *in vivo* contrairement aux G-quadruplex à boucles réduites (< 4

bases). Enfin, l'étude a également été faite sur une autre séquence à longue boucle (13 bases), le promoteur du gène BCL2 dont la structure RMN est également connue.^[175]

Acronyme	Séquence 5' -> 3'
F-21-T	Fam-GGGTTAGGGTTAGGGTTAGGG-Tam
F-Pu24T-T	Fam-TGAGGGTGGTGGAGGGTGGGAAGG-Tam
F-myc-T	Fam-TTGAGGGTGGGTAGGGTGGGTAA-Tam
F-CEB25-WT (L191)-T	Fam-AAAGGGTGGGTGTAAGTGTGGGTGGGT -Tam
F-CEB25-L111(T)-T	Fam-AAAGGGTGGGTGGGTGGGT -Tam
F-CEB25-L121(TT)-T	Fam-AAAGGGTGGGTGGGTGGGT -Tam
F-CEB1-T	Fam- AGGGGGAGGGAGGGTGGCCTGCGGAGGTCCCTGGGCTG -Tam
F-BCL2 NMR-T	Fam-AGGGCGGGCGCTTTAGGAAAAGGGCGGGA-Tam
F-BCL2-WT-T	Fam-AGGGCGGGCGGGAGGAAGGGCGGGA-Tam

Figure 78 : Séquences d'ADN utilisés pour l'évaluation par FRET-melting (en rouge les guanines impliquées dans la formation du G-quadruplex).

Afin de s'assurer que l'ajout de fluorophore en 5' et 3' sur les séquences d'oligonucléotides n'affecte pas la formation des G-quadruplex, nous vérifions la signature de leur structure G-quadruplex par dichroïsme circulaire (figure 79).

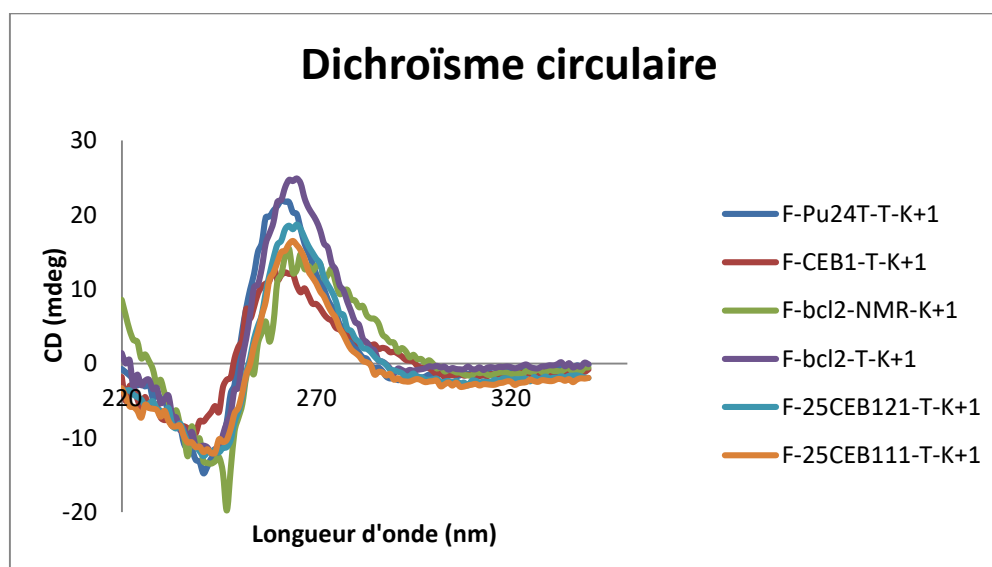


Figure 79 : Spectres de dichroïsme circulaires des séquences modifiées pour l'expérience de FRET-melting.

Nous pouvons constater que l'ajout de fluorophores sur les séquences ne modifie pas leur structure G-quadruplex et que les séquences, dont les spectres sont présentés sur ce graphe (figure 79), ont une conformation parallèle. De plus, toutes ces expériences ont été effectuées dans un tampon K⁺1

qui correspond à une concentration de potassium de 1 mM dans une solution de cacodylate de lithium (10 mM), très inférieures aux conditions physiologiques (~100 mM de potassium). Le tampon est celui utilisé pour les expériences de FRET-melting par la suite. Les spectres des autres séquences ne sont pas représentés ici car ils sont déjà publiés par le groupe de J-L. Mergny.^[176] Excepté la séquence télomérique humaine F21T qui est polymorphe, toutes les autres séquences étudiées sont parallèles.

L'ADN que l'on veut évaluer avec nos composés est placé en présence du tampon avant d'être hybridé à la concentration de l'expérience. Le choix du tampon dépend également de l'ADN. Si un G-quadruplex possède une température de fusion très élevée alors il est très stable *in vitro*. Il sera donc difficile de le stabiliser de manière significative par ajout d'un ligand affiné. Afin d'avoir une température de fusion plus faible, il est nécessaire de se placer dans des conditions moins stabilisantes en diminuant la concentration en potassium du tampon utilisé pour l'expérience.

Les tampons utilisés en FRET-melting sont le K⁺10 (10 mM d'ions potassium (KCl) en plus de 10 mM de cacodylate de lithium et 90 mM de LiCl) pour le 22AG et K⁺1 (1 mM d'ions potassium (KCl), 99 mM de LiCl, 10 mM de cacodylate de lithium) pour les autres séquences étudiées. L'utilisation de ces tampons permet d'avoir des températures de fusion comprises entre 50 et 72 °C. La plupart des séquences formant des G-quadruplex possède des températures de fusion autour de 50 °C ce qui permet de comparer les séquences entre elles. Seules les structures très stables comme les séquences modifiées du minisatellite CEB25 ont des valeurs de température de fusion supérieures à 60 °C.

Il est important de noter que la valeur de ΔT_m dépend du complexe ligand/ADN c'est-à-dire à la fois de la valeur de la température de fusion de l'oligonucléotide seul et du ligand plus ou moins affiné pour la cible. La température de fusion varie en fonction de la stabilité de l'oligonucléotide, le CEB25-WT (L191) possédant une grande boucle est moins stable que le CEB25-L111(T) (boucle de 1 base) et possède donc une température de fusion plus faible. Le CEB25-L121(TT) (boucle de 2 bases) se situant évidemment entre les deux (figure 80).

Acronyme	Tampon	T _m (°C)
F-21-T	K ⁺ 10	51,0
F-Pu24T-T	K ⁺ 1	53,6
F-myc21-T	K ⁺ 1	49,6
F-CEB25-WT (L191)-T	K ⁺ 1	55,6
F-CEB25-L111(T)-T	K ⁺ 1	72,4
F-CEB25-L121(TT)-T	K ⁺ 1	65,8
F-ceb1-T	K ⁺ 1	54,6
F-bcl2-WT-T	K ⁺ 1	59,2

Figure 80 : Température de fusion des oligonucléotides seuls

2) Résultats

a) Ligands de référence

Dans le premier chapitre, j'ai effectué une liste non exhaustive de ligands de référence. Parmi eux, beaucoup de ces ligands ont été synthétisés et développés au sein du laboratoire (la famille PhenDC, Pt-ttpy, MMQ12...). Dans ce paragraphe, je vais présenter les résultats de l'expérience de FRET-melting de ces ligands ainsi que de ligands de référence tels que les composés pyridostatine (PDS), PDC et RHPS4 (figure 81).

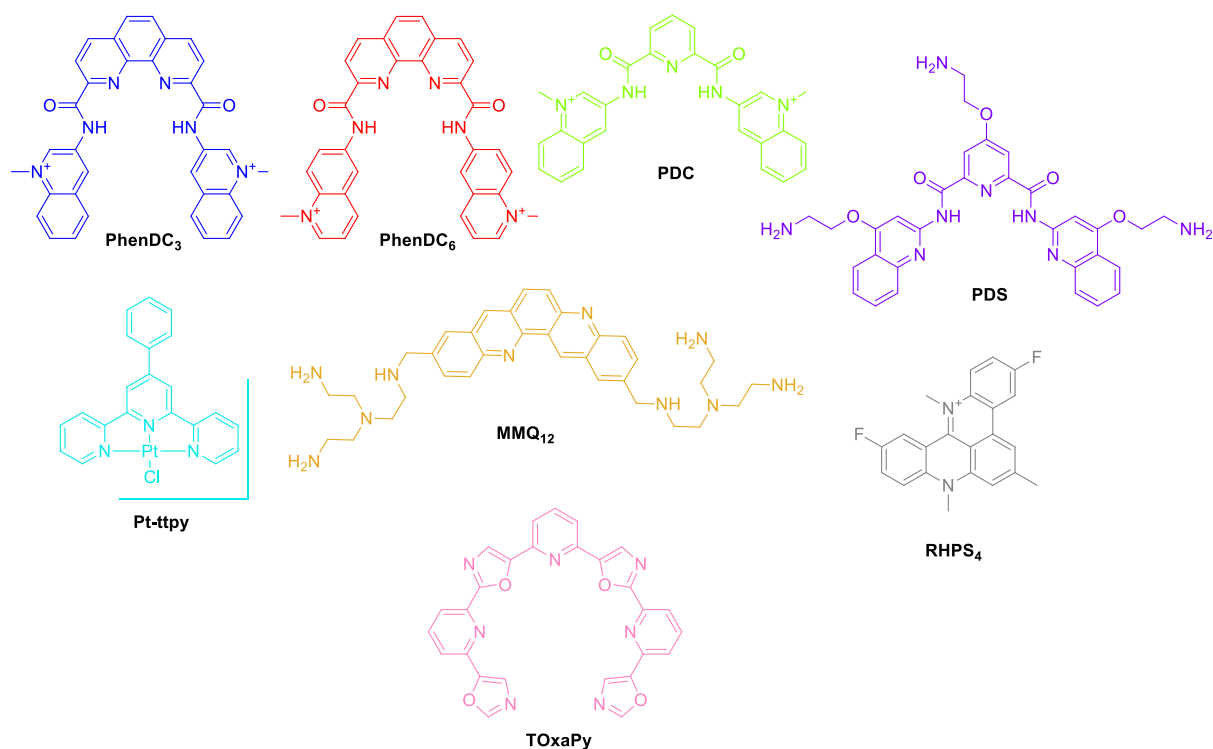


Figure 81 : Structure des composés de référence évalués.

Pour évaluer ces composés, quatre séquences d'oligonucléotides ont été choisies :

- La séquence télomérique humaine (F-21-T) qui présente un grand polymorphisme
- la séquence modifiée de l'oncogène c-myc (F-Pu24T-T) qui a servi à établir la structure RMN du PhenDC₃
- la séquence du minisatellite à grande boucle (9 bases) CEB25 qui a été très étudiée au sein de levures par l'équipe d'Alain Nicolas.^[32c, 123]
- une séquence modifiée avec une boucle courte (une base) du minisatellite CEB25 (CEB25-L111(T)).

Sur cette base, une échelle est établie pour classer les ligands. Tous les composés ayant une valeur de ΔT_m supérieure à 20 °C sont considérés comme des ligands de forte affinité pour les G-quadruplex, les composés ayant un ΔT_m compris entre 10 et 20 °C sont des ligands d'affinité moyenne à élevée. Enfin, les composés ayant une valeur de ΔT_m inférieure à 10 °C sont considérés comme de faibles stabilisants de G-quadruplex ayant une affinité moyenne à faible dans les conditions opératoires choisies.

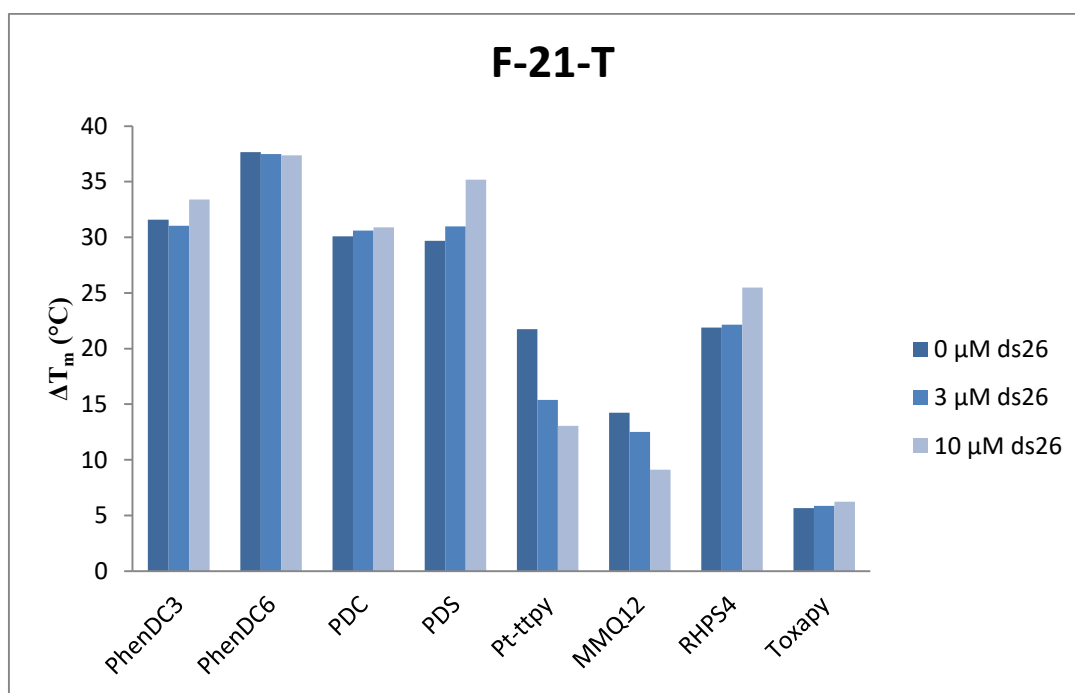


Figure 82 : Résultats de FRET-melting sur la séquence télomérique F-21-T.

Sur ces premiers résultats de FRET-melting sur la séquence télomérique (Figure 82), nous pouvons constater que sur les huit ligands testés, six sont sélectifs pour le G-quadruplex par rapport à l'ADN duplex. Les composés Pt-ttpty et MMQ12 sont moins sélectifs avec une diminution des valeurs de ΔT_m en présence d'ADN duplex. En revanche, ils possèdent une bonne affinité pour le G-quadruplex lorsqu'il n'y a pas de compétition ($\Delta T_m > 10$ °C). De plus, tous les composés sont de très bons ligands capables de stabiliser la séquence télomérique excepté le composé ToxaPy qui est le moins affiné pour

cette séquence. Tous les autres composés ont un effet stabilisant élevé pour cet oligonucléotide, il est donc difficile d'effectuer un classement des ligands avec la séquence télomérique.

Par la suite, la sélectivité ne sera plus développée car elle est identique sur toutes les séquences étudiées. A savoir que les composés Pt-ttpty et MMQ12 sont moins sélectifs et interagissent également avec l'ADN duplex alors que les 6 autres ligands conservent leur haute sélectivité pour toutes les structures G-quadruplex.

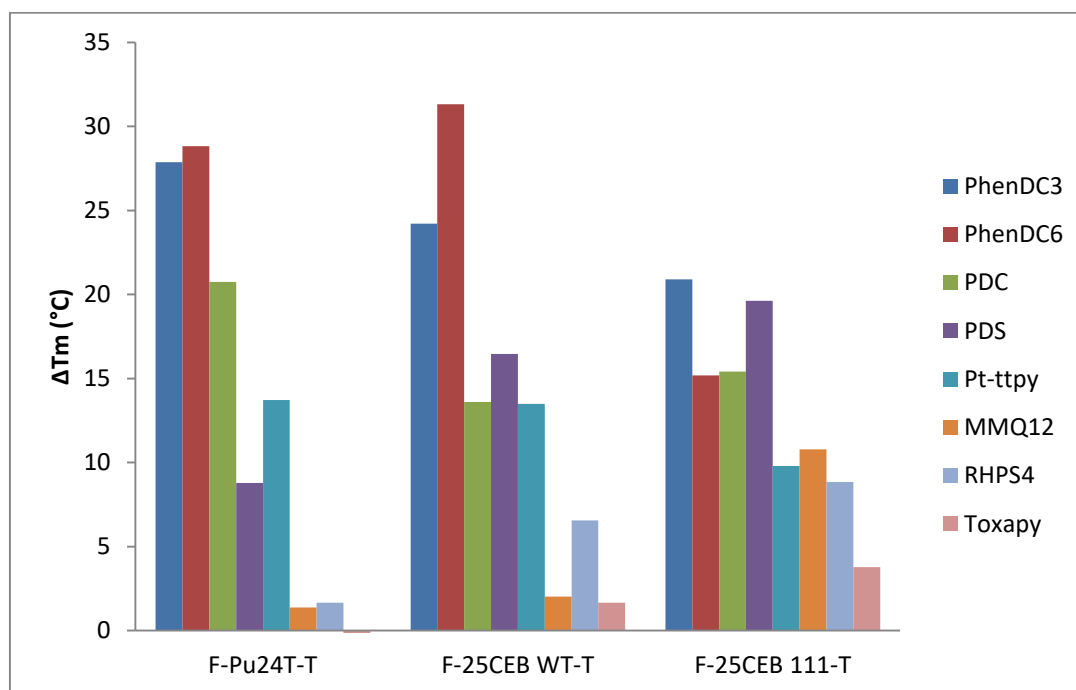


Figure 83 : Résultats de FRET-melting sur les séquences Pu24T, CEB25-WT (L191) et CEB25-L111(T).

Les composés RHPS4, MMQ12 et ToxaPy ne stabilisent pas la séquence F-Pu24T-T, avec une valeur de ΔT_m inférieure à 2 °C (figure 83) ce qui montre l'absence d'affinité pour cette séquence. Les composés PDS et Pt-ttpty ont une valeur de ΔT_m d'environ 10 °C ce qui montre une bonne stabilisation de ce G-quadruplex cependant ils le stabilisent beaucoup moins que les composés PDC, et la famille PhenDC ($\Delta T_m > 20$ °C). Cette séquence permet d'établir une classification en fonction de l'affinité des ligands pour l'oligonucléotide. Enfin, les deux séquences issues du minisatellite CEB25 nous permettent de montrer l'influence de la boucle sur la stabilisation de l'oligonucléotide. La séquence à longue boucle (CEB25-WT (L191), $T_m \sim 55$ °C) est moins stable que la séquence à boucle courte (CEB25-L111(T), $T_m \sim 70$ °C). En s'appuyant sur les résultats obtenus avec la séquence à boucle longue (CEB25-WT (L191)), les ligands sont classés en fonction de leur affinité. Les composés de la famille PhenDC sont toujours d'excellents ligands capables de stabiliser cette séquence ($\Delta T_m > 24$ °C). Les composés PDC, PDS et Pt-ttpty stabilisent fortement le minisatellite CEB25-WT (L191) (12 °C $< \Delta T_m < 17$ °C) alors que le composé RHPS4 ne le stabilise que légèrement ($\Delta T_m \sim 7$ °C). Les ligands MMQ12

et Toxapy ne sont pas de bons ligands pour ce quadruplex ($\Delta T_m < 2 \text{ }^\circ\text{C}$). Sur la séquence CEB25-L111(T), le classement des ligands est cette fois-ci plus compliqué car la valeur de ΔT_m ne peut excéder $20 \text{ }^\circ\text{C}$. En effet, au-delà de cette température ($\sim 90 \text{ }^\circ\text{C}$) l'échantillon est porté à ébullition. La différence de ΔT_m entre ligands n'est pas suffisamment importante pour avoir un classement précis sur cette séquence déjà très stable naturellement dans nos conditions opératoires. On note néanmoins que le composé ToxaPy n'est pas un ligand permettant la stabilisation de G-quadruplex ($\Delta T_m < 4 \text{ }^\circ\text{C}$) dans ces conditions. Les composés Pt-ttpty, MMQ12 et RHPS4 sont des ligands d'affinité modérée avec des valeurs de ΔT_m d'environ $10 \text{ }^\circ\text{C}$ et enfin, les ligands de forte affinité correspondent à la famille PhenDC et les composés PDC et PDS qui ont un $\Delta T_m > 15 \text{ }^\circ\text{C}$.

A l'aide des résultats sur ces quatre séquences, il est possible de faire un bilan général et d'établir un classement des composés par affinité pour les G-quadruplex. En effet, la famille de composés PhenDC permet de stabiliser toutes les séquences de manière équivalente, indépendamment de la structure ou la taille de la boucle. Ce sont d'excellents stabilisants de G-quadruplex ($\Delta T_m > 20 \text{ }^\circ\text{C}$). La seconde catégorie concerne les ligands qui stabilisent fortement les G-quadruplex mais dont la valeur de ΔT_m varie également en fonction de la séquence. Il s'agit du composé PDC avec des valeurs de ΔT_m toujours supérieures à $15 \text{ }^\circ\text{C}$ mais qui est affecté par la taille de la boucle ou encore par la structure particulière du Pu24T (figure 84). Un autre composé pourrait être dans cette catégorie mais son affinité est très impactée sur la séquence Pu24T, il s'agit du composé pyridostatine (PDS) qui entre donc dans une troisième catégorie avec un grand effet de la structure du G-quadruplex sur l'affinité.

La séquence Pu24T possédant une guanine isolée venant former le troisième quartet, seul un quartet externe est accessible pour les ligands. Cela n'a pas d'influence pour les composés de la famille PhenDC car on sait que le composé PhenDC₃ a une stœchiométrie 1:1 avec cet ADN.^[70]

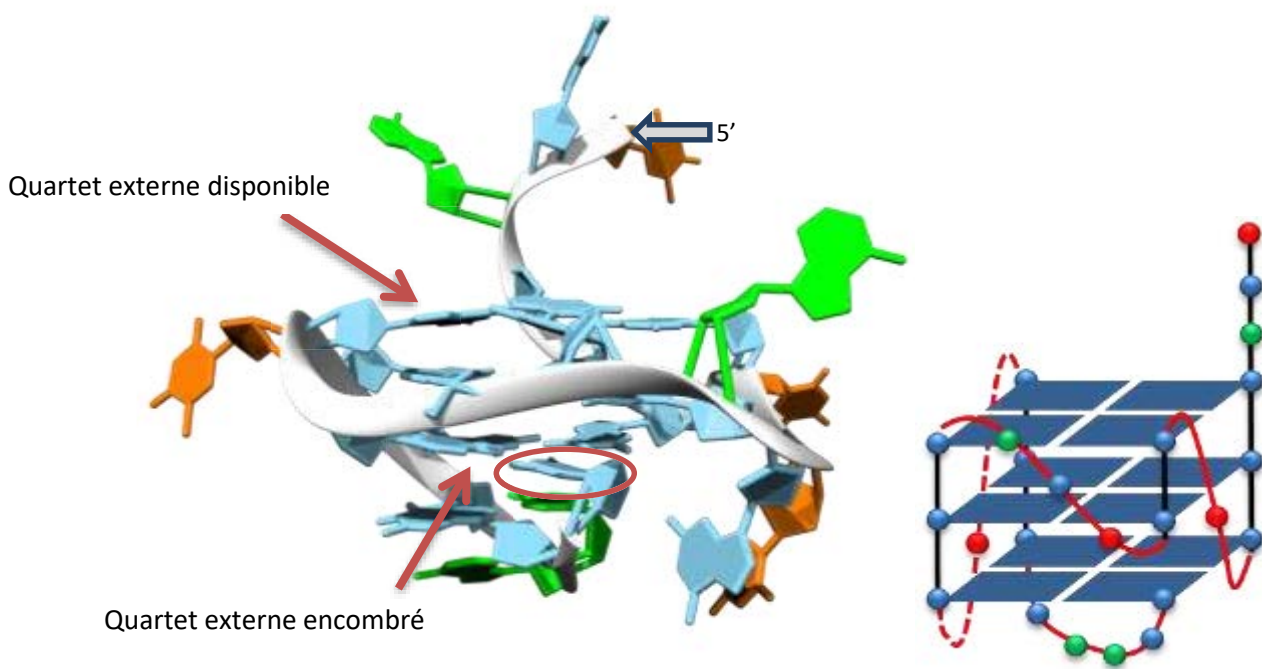


Figure 84 : Structure définie par RMN (PDB : 2MGN) de la séquence Pu24T (TGAGGGTGGTGAGGGGGGGAACG bases soulignés = boucles).^[70]

La quatrième catégorie peut être constituée des composés sélectifs, comme dans les catégories précédentes, mais ayant une plus faible affinité hormis pour la séquence télomérique et les G-quadruplex très stables comme le CEB25-L111(T). Ces composés sont séquences-dépendants. Dans cette catégorie se situe le composé RHPS4. Ensuite, une catégorie pour les ligands de faible affinité comme le composé ToxaPy. Enfin, une catégorie de bons ligands capables de stabiliser les G-quadruplex mais ayant également une affinité pour d'autres formes d'ADN comme le duplex, il s'agit des ligands Pt-ttpty et MMQ12.

Pour conclure, les composés PhenDC₃ et PDC ont montré de très bonnes caractéristiques pour stabiliser les G-quadruplex ainsi ils sont donc des ligands de choix à fonctionnaliser pour les utiliser dans différentes applications.

b) Dérivés du PhenDC₃

De nombreux composés dérivés du ligand PhenDC₃ ont été développés au laboratoire (figure 85). Dans un premier temps, des ligands bi-fonctionnels équipés d'un groupement photoactivable (PhenDC₃-bn) ou d'une biotine (PhenDC₃-biot) ont été synthétisés pour, respectivement, créer un lien covalent entre la molécule et l'ADN ou pouvoir extraire de l'ADN dans un milieu complexe. Ensuite, un ligand tri-fonctionnel (PhenDC₃-tri) possédant les deux fonctions précédentes a été mis au point. Un ligand très proche du composé PhenDC₃ possédant une chaîne aminée sur les

quinoléines au lieu d'un azote quaternaire a également été synthétisé (PhenDC₃-pyro). Enfin, trois composés « clickables » (PhenDC₃-alk, PhenDC₃-az et PhenDC₃-alk-bn) ont été développés.

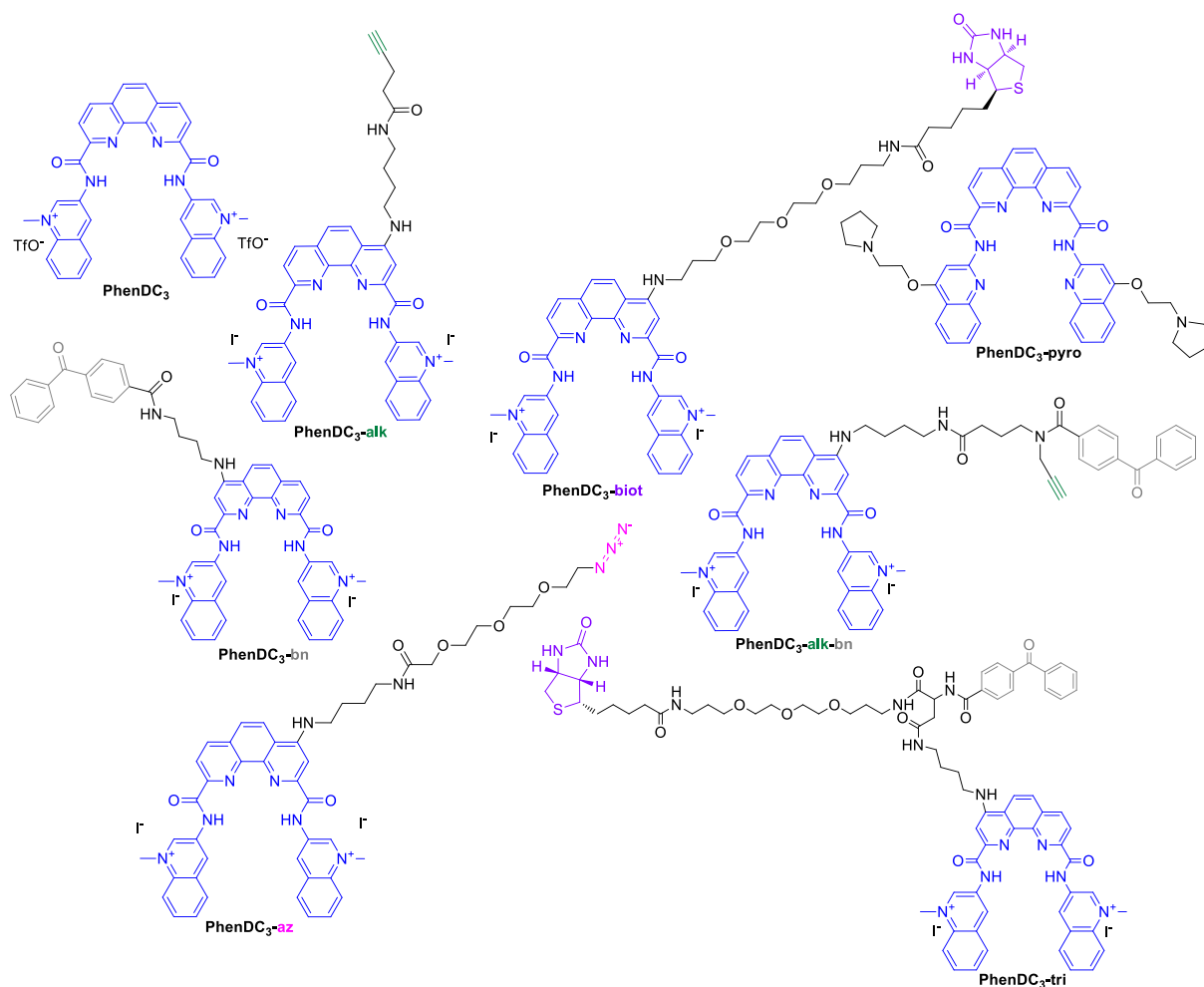


Figure 85: Structures des dérivés PhenDC₃.

Séquence télomérique humaine (F-21-T)

Comme pour les ligands de référence, la sélectivité de ces ligands ne sera illustrée que sur une seule séquence d'ADN, la séquence télomérique humaine. Cependant, toutes les mesures de température de fusion en présence de compétiteurs ont été effectuées sur toutes les séquences d'ADN étudiées et se révèlent analogues à celles mesurées pour la séquence télomérique humaine.

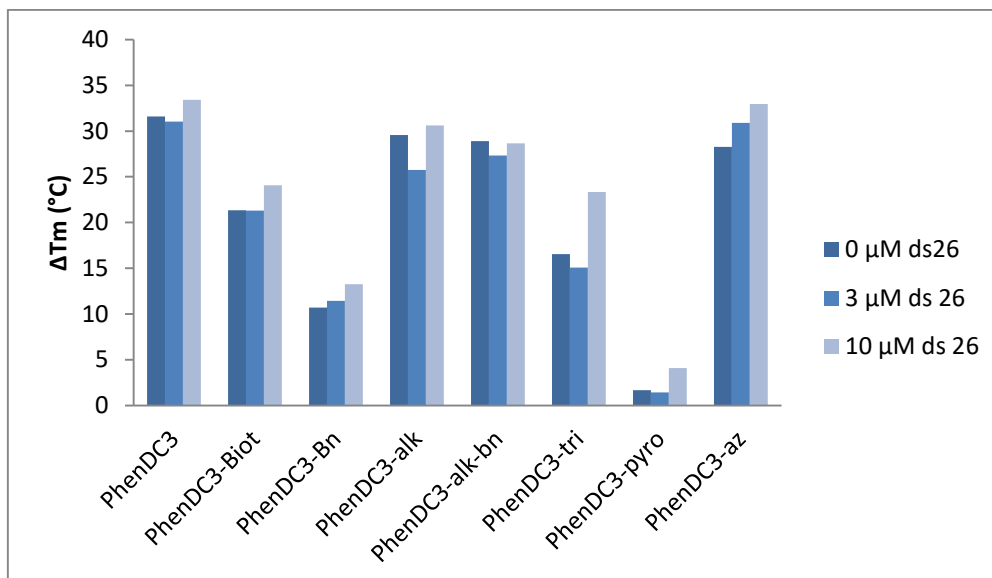


Figure 86 : Résultats de FRET-melting des dérivés PhenDC₃ sur la séquence télomérique humaine F-21-T.

Tous les dérivés du PhenDC₃ développés au laboratoire conservent la sélectivité pour l'ADN G-quadruplex par rapport à l'ADN duplex (figure 86). Concernant l'affinité, on constate que certains ligands sont beaucoup plus affectés que d'autres par leur fonctionnalisation. Le composé PhenDC₃-pyro ne stabilise quasiment pas la séquence télomérique. Il semblerait que l'azote quaternaire sur les quinoléines apporte une meilleure reconnaissance et que l'ajout de chaînes aminées diminue l'affinité. De plus, les groupements pyrrolidines peuvent également induire une gêne stérique plus importante qu'une amine primaire comme cela peut être le cas avec le composé PDS. Les composés fonctionnalisés par une biotine, une benzophénone ou les deux perdent de leur affinité avec des diminutions de 10 à 20 °C de ΔT_m alors que les composés azoture et alcyne, qui apportent moins de gêne stérique, conservent globalement la même affinité que le composé de départ, le ligand PhenDC₃, sur cette séquence.

Séquences de l'oncogène c-myc

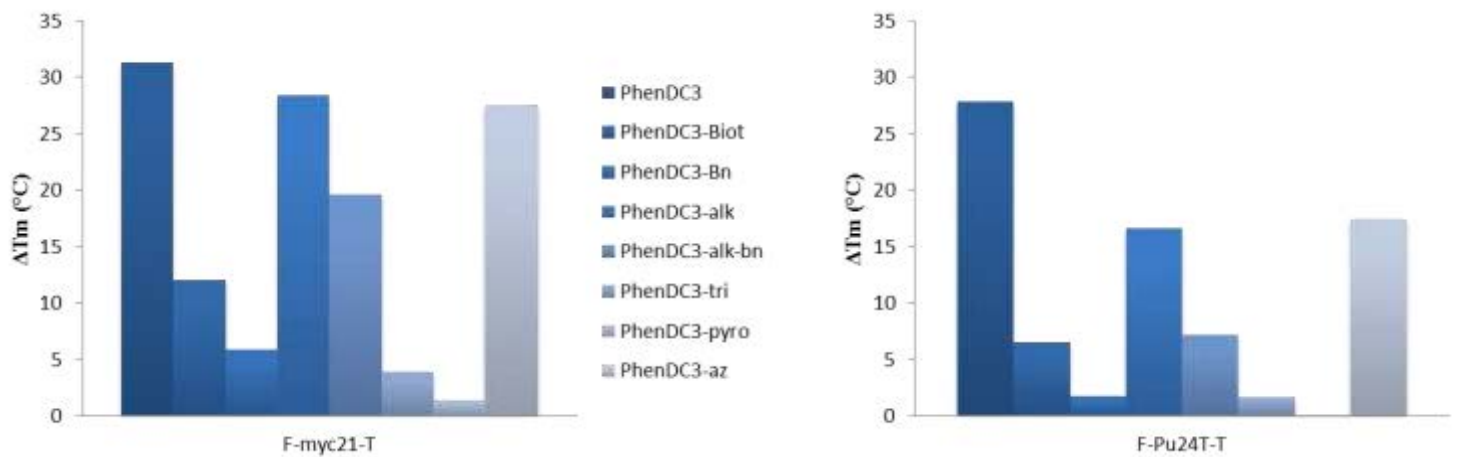


Figure 87 : Résultats de FRET-melting sur deux séquences modifiées de l'oncogène c-myc.

La tendance est exactement la même (figure 87) entre ces deux séquences modifiées de l'oncogène c-myc malgré deux structures de G-quadruplex différentes : la séquence Pu24T, que l'on a vu précédemment (figure 88), avec une guanine isolée qui vient compléter la triade pour former le dernier quartet et la séquence myc22 qui a une structure plus compacte et les deux quartets externes sont disponibles (figure 88).

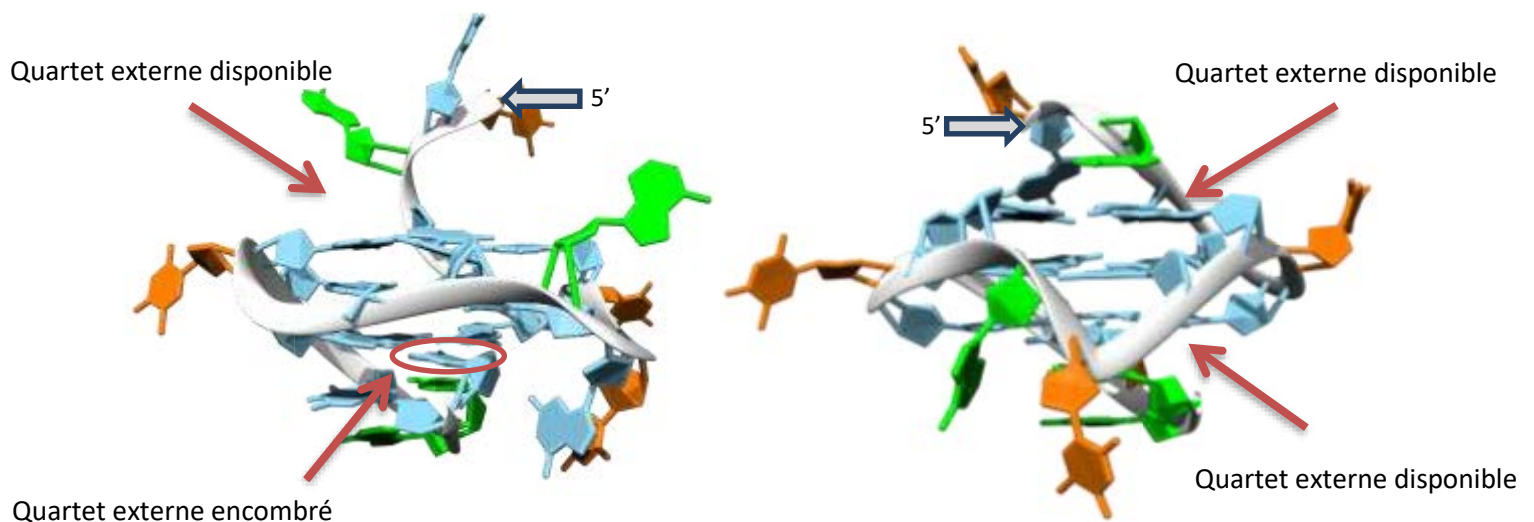


Figure 88 : Structure RMN de la séquence Pu24T (PDB : 2MGN, à gauche)^[70] et de la séquence Myc22^[174] (PDB : 1XAV, à droite) de l'oncogène c-myc.

On remarque ainsi que pour des valeurs de température de fusion proches (49,6 °C pour F-myc21-T contre 53,6 °C pour F-Pu24T-T) pour l'oligonucléotide seul, les valeurs de ΔT_m sont très différentes excepté pour le composé PhenDC₃ non fonctionnalisé. Les températures de fusion en présence de ligands sont plus élevées pour la séquence Myc22 que pour la séquence Pu24T. La légère différence de température de fusion d'environ 4 °C qui existe pour l'oligonucléotide seul ne peut expliquer, à

elle seule, les différences de ΔT_m en présence de ligands. Cela peut également s'expliquer par la disponibilité des quartets et la dynamique de la structure formée. Les composés azoture (-az) et alcyne (-alk) sont les composés qui conservent la meilleure affinité pour ces G-quadruplex. D'après ces trois premières séquences G-quadruplex étudiées, il est déjà possible de dire que l'utilisation de composés alcyne ou azoture est la stratégie à utiliser pour ne pas impacter l'affinité. En effet, la fonctionnalisation de la phénanthroline en position 4 par des fonctions stériquement encombrantes comme la benzophénone ou la biotine entraîne une forte perte d'affinité et encore plus forte pour le composé tri-fonctionnel. Le composé PhenDC₃-pyro stabilise très faiblement les G-quadruplex.

Séquences du minisatellite CEB25

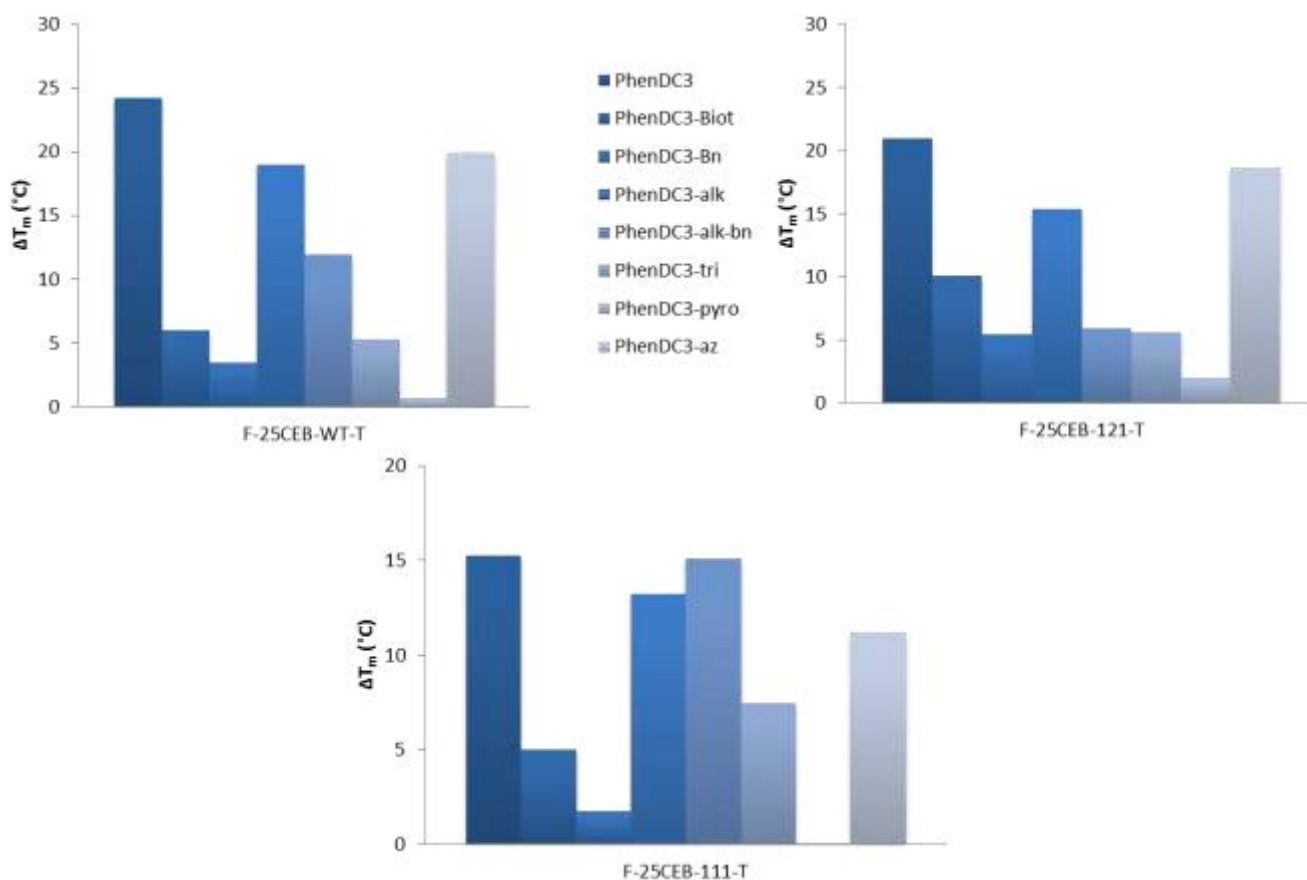


Figure 89 : Résultats de FRET-melting sur les différentes séquences du minisatellite CEB25.

La taille des boucles au sein d'un G-quadruplex a une influence sur la stabilité de l'oligonucléotide, qui se constate par des différences de température de fusion très différentes (55 °C pour CEB25-WT (L191), 64 °C pour CEB25-L121(TT) et 72°C pour CEB25-L111(T)), mais également sur sa stabilisation par un ligand (figure 89). En effet, on peut constater que tous les ligands, excepté le composé fonctionnalisé par une benzophénone et le composé -pyro, sont capables de stabiliser les oligonucléotides très stables, possédant une petite boucle (une base). Ces résultats nous permettent de montrer que l'ajout d'un groupement encombrant tel que la benzophénone à l'extrémité d'une

chaîne aliphatique courte (quatre carbones, figure 90) affecte grandement l'affinité du ligand pour la cible. De plus, la fonction benzophénone peut interagir par liaisons hydrogène avec les boucles des structures G-quadruplex mais également par empilement sur les quartets. L'ajout d'une fonction, qui est également encombrante (la biotine, figure 90), à l'extrémité d'une chaîne PEG affecte également l'affinité mais de manière moins conséquente (figure 89). L'effet de chaîne peut être favorable comme cela a été montré dans la littérature par le groupe de Unrau avec le composé thiazole orange-biotine qui a cristallisé en présence d'un G-quadruplex et a permis une étude aux rayons X.^[177] La fonction biotine s'empile sur le quartet et interagit avec les bases des boucles. La chaîne quant à elle, s'empile également sur le quartet externe de l'ARN G-quadruplex Mango.

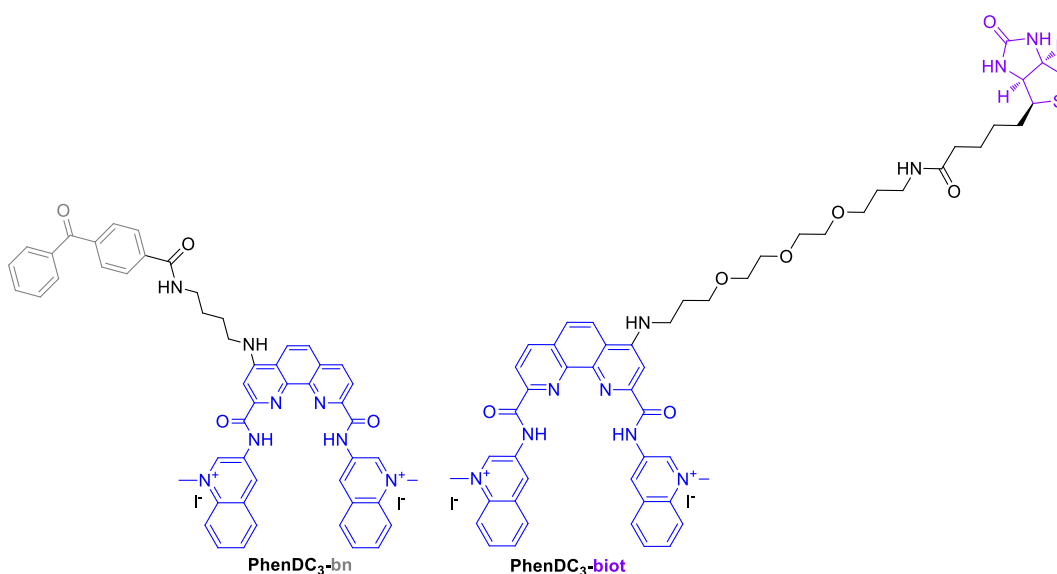


Figure 90 : Structures des bi-fonctionnels biotine et benzophénone.

Le composé tri-fonctionnel (figure 91) conserve à peu près la même affinité que le bi-fonctionnel biotine. Sur ce composé, la fonction benzophénone est à l'extrémité d'une chaîne de 7 carbones ce qui montre que la distance entre la phénanthroline et la fonction a un effet sur l'affinité. Les composés fonctionnalisés semblent être affectés lorsque la longueur de l'espaceur est courte (4 carbones) et semblent être moins affectés lorsque cet espaceur est allongé malgré un encombrement stérique plus important (PhenDC₃-tri).

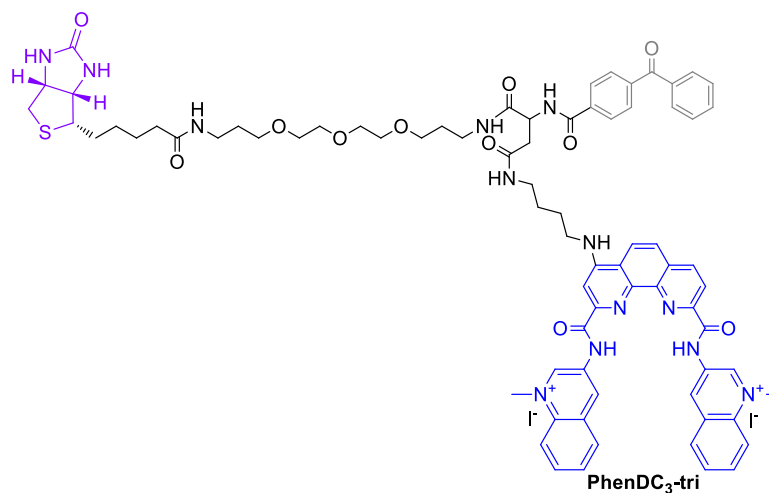


Figure 91 : Structure du dérivé tri-fonctionnel.

Les composés azoture et alcyne (figure 92), qu'ils soient fonctionnalisés à l'extrémité d'une chaîne aliphatique de 7 carbones (alcyne) ou à l'extrémité d'une chaîne PEG (azoture) en plus d'une chaîne de quatre carbones, conservent une affinité élevée sur les différentes séquences étudiées.

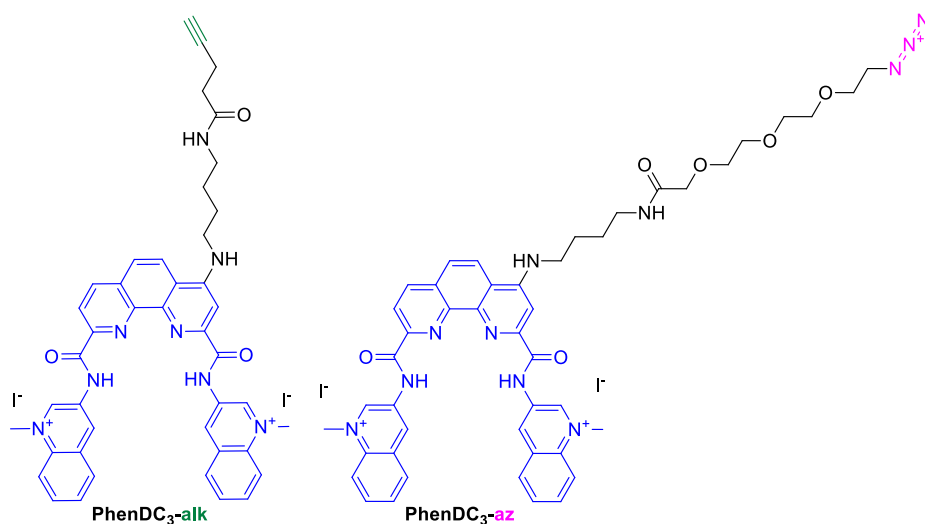


Figure 92 : Structure des composés PhenDC₃-alk et PhenDC₃-az.

La taille de la boucle apporte de la dynamique supplémentaire au G-quadruplex ce qui peut empêcher le ligand de se positionner sur le quartet de manière optimale. D'après les structures RMN définies (figure 93), il est évident que la boucle influence la bonne reconnaissance de la structure par un ligand et peut encombrer stériquement le quadruplex.

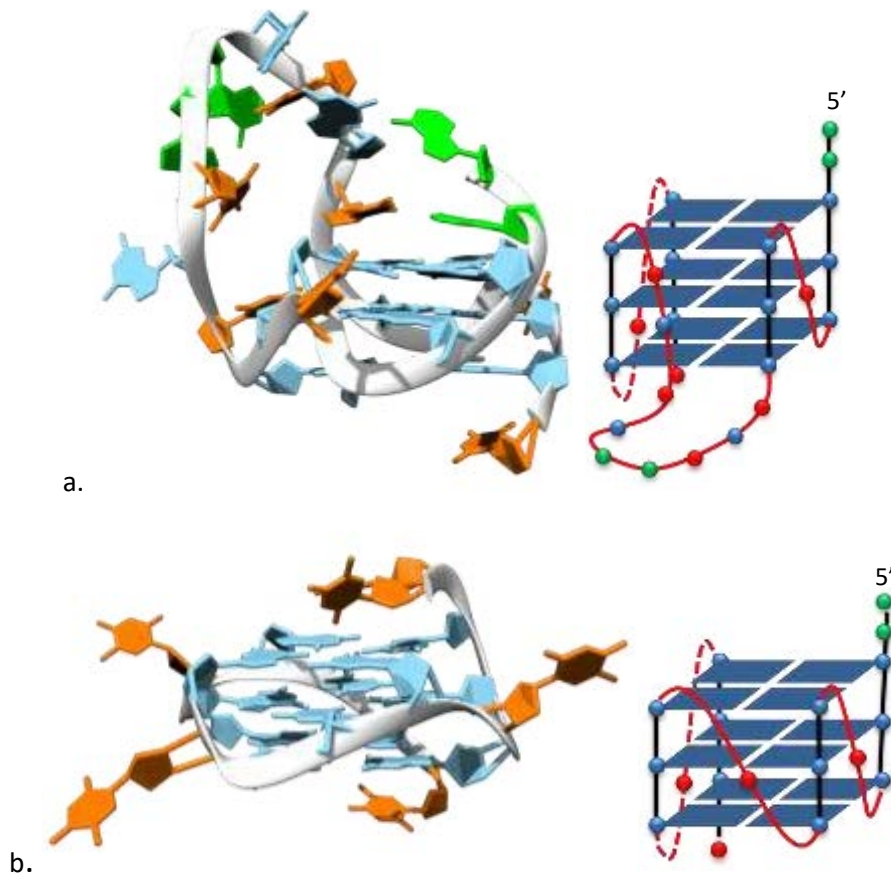


Figure 93 : Structure RMN des séquences a. CEB25-WT (L191) ^[123] (PDB : 2LPW) et b. T95-2T (PDB : 2LK7) ressemblant à CEB25-L111(T).^[178]

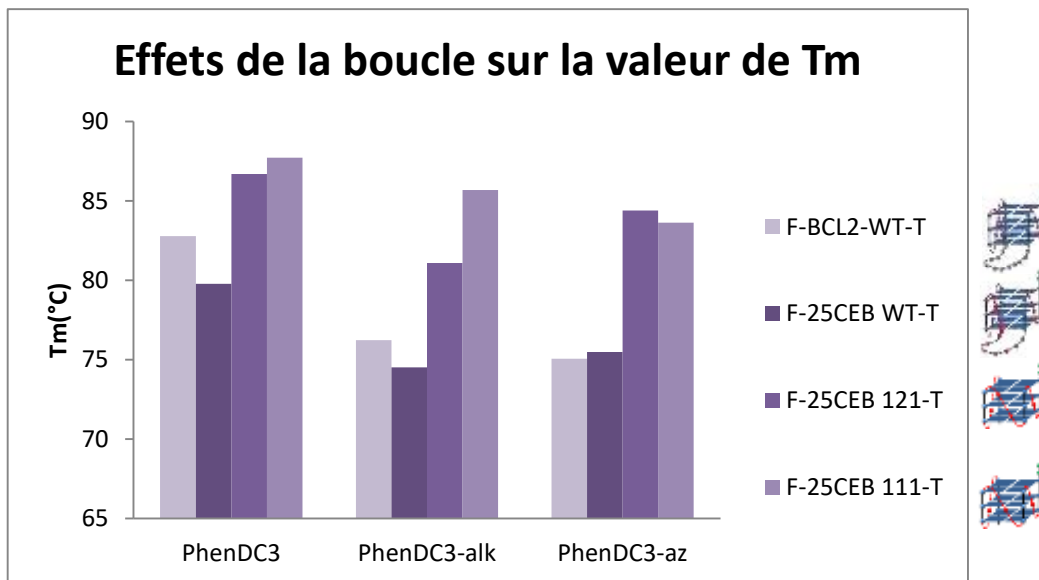


Figure 94 : Effets de la boucle sur la valeur de température de fusion du complexe ligand/ADN.

La taille de la boucle centrale influe sur la stabilité des structures G-quadruplex. En effet, les G4 à longues boucles sont moins stables et ont donc des températures de fusion plus basses ($T_m \sim 55$ °C pour CEB25-WT (L191) et $T_m \sim 59$ °C pour BCL2 WT) que les G4 à boucles courtes (65 °C pour CEB25-

L121(TT) et 72 °C pour CEB25-L111(T)). En présence de ligand, la stabilisation de l'oligonucléotide est également dépendante de la taille de la boucle. Plus la boucle est longue moins la valeur de température de fusion est grande. Pour confirmer cette tendance, nous avons évalué la séquence du promoteur de gène BCL2 qui possède une boucle centrale de 13 bases (figure 94).^[175b] La conclusion reste valable, la longue boucle donne une valeur de température de fusion plus faible en présence de ligand que lorsque les boucles sont courtes. Les oligonucléotides à longues boucles atteignent une valeur de température de fusion entre 75 et 80 °C alors que les G-quadruplex plus stables atteignent des températures de fusion d'environ 85 °C.

c) Dérivés du PDC

Après les dérivés du composé PhenDC₃ qui sont, pour certains, très affectés par la fonctionnalisation et qui perdent en affinité, nous allons nous intéresser à un autre composé référence, PDC (ou 360A), et à ses dérivés (figure 95). Ce composé a été fonctionnalisé de la même manière que le ligand PhenDC₃ c'est-à-dire avec une biotine (PDC-biot), une benzophénone^[171] (PDC-bn), les deux en même temps (PDC-tri) et un composé « clickable » qui est l'alcyne (PDC-alk).^[168]

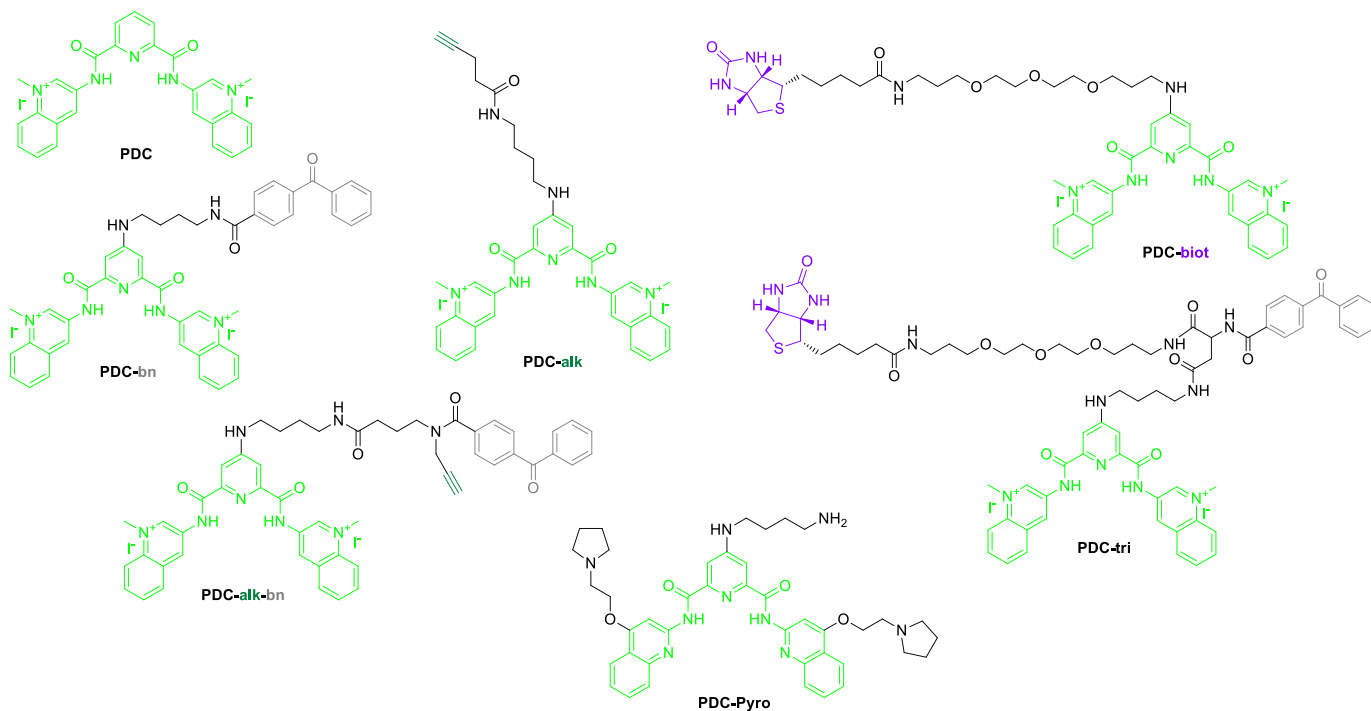


Figure 95 : Structure du PDC et ses dérivés.

Séquence télomérique humaine

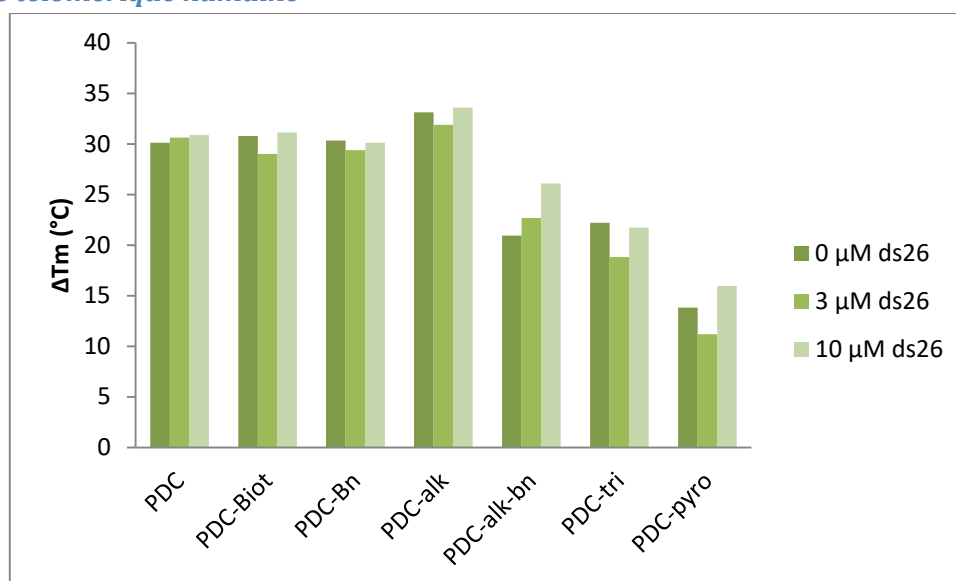


Figure 96 : Résultats de FRET-melting sur la séquence télomérique humaine F-21-T.

Comme pour les dérivés du PhenDC₃, la sélectivité des dérivés du PDC n'est pas affectée par la fonctionnalisation (figure 96). Concernant l'affinité, on constate que tous les ligands sont capables de stabiliser fortement la séquence télomérique. En effet, ils ont tous une valeur de ΔT_m supérieures à 15 $^{\circ}$ C. Cependant, on constate tout de même que le composé tri-fonctionnel, donc très encombré stériquement, et le composé « clickable » possédant une benzophénone induisent des valeurs de ΔT_m inférieures de dix degrés ce qui implique une perte d'affinité. Le dérivé qui stabilise le moins la séquence télomérique humaine est, comme pour les dérivés du PhenDC₃, le dérivé PDC-pyro.

Cependant, il stabilise tout de même de 15 °C cette séquence ce qui montre une affinité pour la cible. Les dérivés alcyne, benzophénone et biotine ne sont pas affectés.

Séquences de l'oncogène c-myc

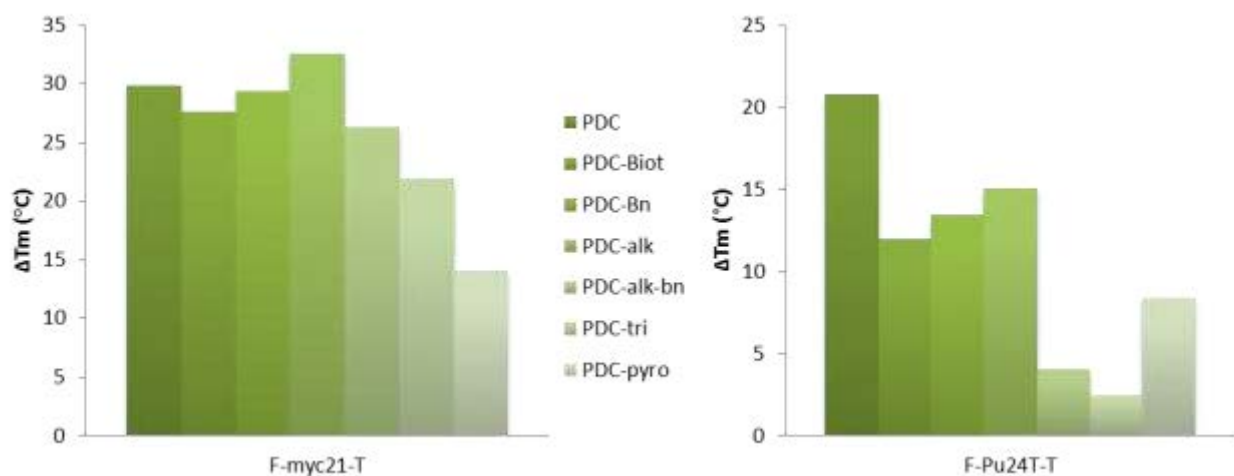


Figure 97 : Résultats de FRET-melting des dérivés du PDC sur les deux séquences modifiées de l'oncogène c-myc.

Les résultats obtenus sur les deux séquences de l'oncogène c-myc sont très intéressants (figure 97). En effet, ils permettent de confirmer que la fonctionnalisation du PDC n'affecte que très peu voir même pas du tout son affinité et cela permet de garder la même tendance que précédemment à savoir que les composés les plus encombrés (PDC-tri et PDC-alk-bn, figure 98) sont moins affins que le composé de référence PDC.

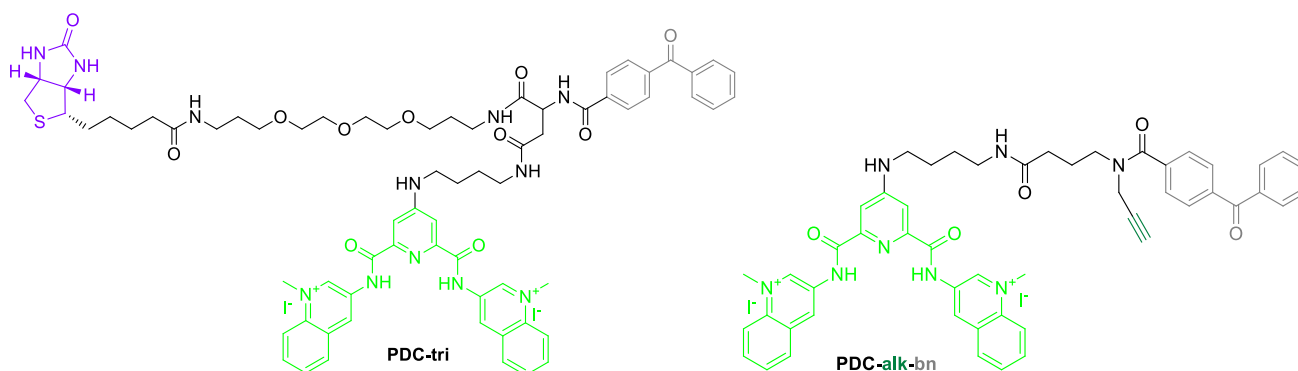


Figure 98 : Structures des composés PDC-tri et PDC-alk-bn.

Le composé –pyro reste un bon ligand mais le fait que la charge ne soit pas sur les quinoléines influe sur son affinité. Ce composé conserve des valeurs de ΔT_m proches quelle que soit la séquence G-quadruplex étudiée. Les chaînes latérales apportent une dynamique à la molécule et peuvent induire un clash stérique avec les boucles des G-quadruplex.

La structure particulière du quadruplex Pu24T permet d'avoir une meilleure discrimination entre ligands avec le composé PDC-tri qui ne stabilise quasiment pas la structure (ΔT_m d'environ 4 °C). Même si la fonctionnalisation n'affecte pas autant les ligands que pour la famille PhenDC₃, une légère

diminution d'affinité est observée pour les composés biotine, benzophénone et alcyne que l'on n'observait pas sur les autres séquences étudiées comme la séquence télomérique F-21-T.

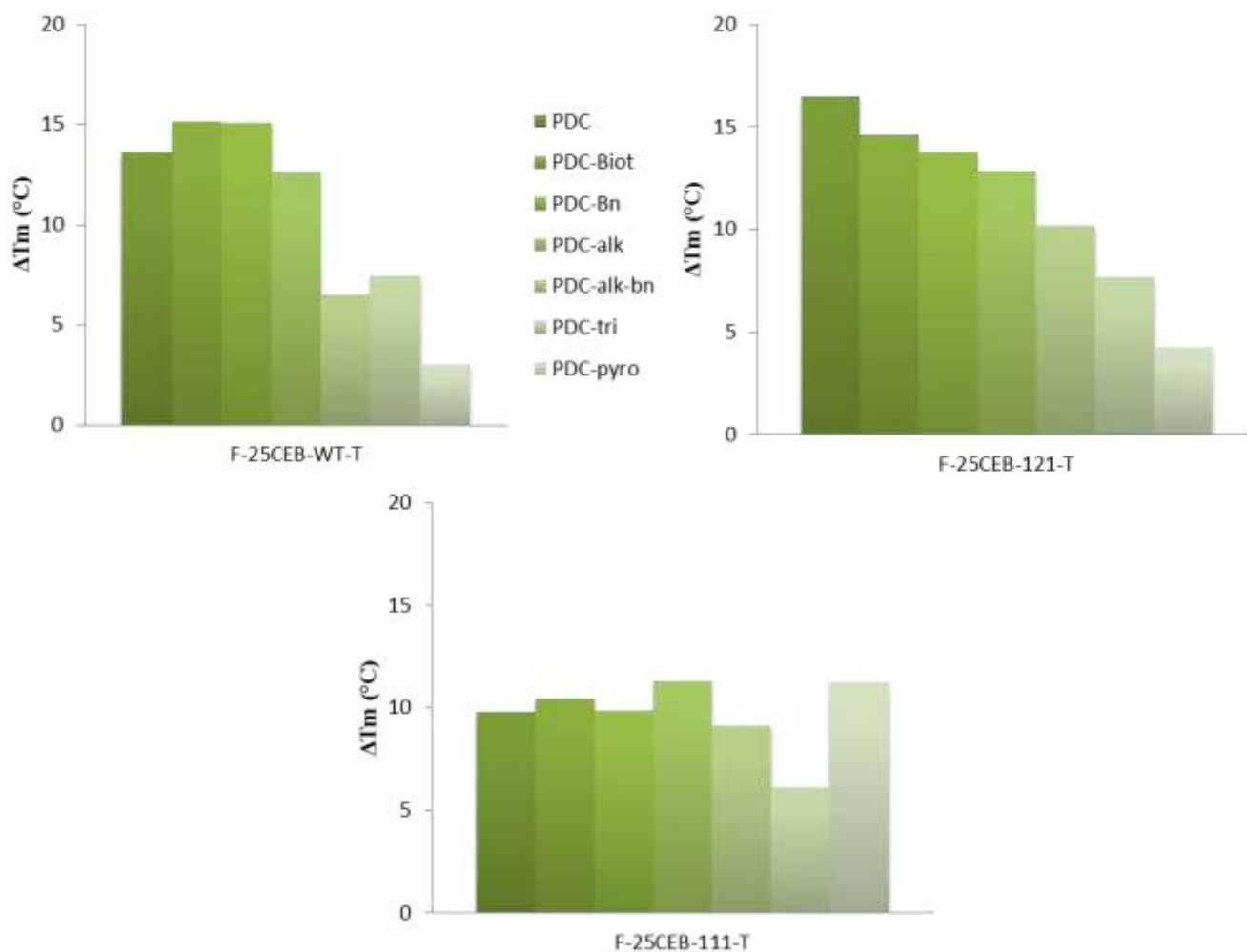


Figure 99 : Résultats de FRET-melting sur les différentes séquences du minisatellite CEB25.

La longueur de la boucle a un effet sur la stabilisation d'un G-quadruplex par un ligand comme montré précédemment et on le constate encore ici avec les dérivés du PDC (figure 99). Le minisatellite CEB25-WT (L191) qui possède une boucle de neuf nucléotides n'est pas stabilisé de manière équivalente avec les différents dérivés. La taille de la boucle induit une gêne stérique ainsi les composés les plus encombrés sont gênés pour stabiliser cette structure. A contrario, lorsque la taille de la boucle est réduite, il n'y a plus de gêne stérique au niveau de l'oligonucléotide et les ligands peuvent donc tous stabiliser la structure G-quadruplex. Ainsi, tous les dérivés du composé PDC ont approximativement la même valeur de ΔT_m sur la séquence CEB25-L111(T) alors qu'ils ont des valeurs de ΔT_m allant de 2 à 15 °C pour la séquence CEB25-WT (L191). Ces résultats nous permettent de montrer que les dérivés du composés PDC sont très robustes et que l'affinité pour les

G-quadruplex va dépendre à la fois de l'encombrement stérique sur le composé mais également celui du G-quadruplex.

3) Conclusions

Pour résumer tous les résultats présentés dans le paragraphe précédent, on peut constater que, parmi les ligands de référence évalués, seule la famille PhenDC conserve une affinité élevée pour tous les G-quadruplex. Les ligands PDC et PDS sont également de très bons ligands avec une affinité légèrement plus faible sur certaines séquences (Pu24T par exemple). Les composés MMQ12 et Pt-ttpty sont des ligands de G-quadruplex capables de stabiliser ces structures cependant ils sont moins sélectifs et interagissent également avec l'ADN duplex. Le composé RHPS4 est un ligand sélectif qui stabilise fortement la séquence télomérique humaine et très peu toutes les autres séquences étudiées. Enfin, le composé ToxaPy ne stabilise que très faiblement la séquence télomérique ($\Delta T_m \sim 5^\circ\text{C}$) dans nos conditions opératoires.

Ensuite, les dérivés du composé PhenDC₃ ont montré une forte perte d'affinité lorsque le composé est substitué par une chaîne carbonée courte (quatre carbones) et fonctionnalisée à son extrémité par un groupement induisant une gêne stérique importante comme la benzophénone. Plus généralement, le fait d'ajouter une chaîne sur la phénanthroline affecte l'affinité ce qui se remarque le plus sur la séquence Pu24T. Les composés alcyne et azoture conservent tout de même une très bonne affinité pour les G-quadruplex malgré une légère diminution de ΔT_m sur certaines structures. Ces résultats valident notre hypothèse de départ qui est d'utiliser la chimie « click » pour identifier la cible et fonctionnaliser le composé *in situ* dans les différentes applications envisagées.

Enfin, les dérivés du PDC ont montré qu'ils conservaient tous une affinité élevée pour les G-quadruplex. Seul le composé très encombré par deux fonctions (biotine et benzophénone) a une affinité plus faible que les autres ligands. On peut également constater que la charge sur les quinoléines est très importante pour l'interaction avec le G-quadruplex ce qui est également le cas pour le dérivé PDC-pyro. On peut supposer que la charge sur les quinoléines permet d'améliorer l'empilement du ligand sur le G-quartet pour stabiliser la structure ce qui suggère que le composé PDS, qui est très proche structuralement du PDC, pourrait avoir un mode d'interaction différent avec le quadruplex pour pouvoir autant les stabiliser. L'interaction électrostatique entre les chaînes aminées et la structure d'ADN doit être très importante.

Afin de confirmer ces classements sur les différents ligands et avoir des informations supplémentaires sur le mode d'interaction, l'étude de ces mêmes ligands a été effectuée par l'expérience de FID.

II. Expérience de FID

1) Séquences d'ADN utilisées

Les séquences utilisées (figure 100) pour cette expérience biophysique sont les mêmes que celles étudiées par FRET-melting. Cependant, l'avantage de ce test est que les séquences ne sont pas modifiées par l'ajout de fluorophores aux extrémités des brins d'ADN.

Acronyme	Séquence 5' -> 3'	Topologie	Rôle biologique
22AG	AGGGTTAGGGTTAGGGTTAGGG	Polymorphe	Séquence télomérique humaine
Pu24T	TGAGGGTGGTGAGGGTGGGGAAGG	Parallèle	Séquence modifiée de l'oncogène c-myc
myc22	TGAGGGTGGGTAGGGTGGGTAA	Parallèle	Séquence modifiée de l'oncogène c-myc
CEB25-WT (L191)	AAGGGTGGGTGTAAGTGTGGGTGGGT	Parallèle	Minisatellite
CEB25-L111(T)	AAGGGTGGGTGGGTGGGT	Parallèle	Séquence modifiée du minisatellite CEB25
CEB25-L121(TT)	AAGGGTGGGTGGGTGGGT	Parallèle	Séquence modifiée du minisatellite CEB25
ceb1-WT	AGGGGGAGGGAGGGTGGCCTGCGGAGGTCCTGGGCTG	Parallèle	Minisatellite
bcl2-WT	AGGGGCGGGCGCGGGAGGAAGGGGGCGGGA	Parallèle	Promoteur du gène bcl2
ds26	CAATCGGATCGAATTCGATCCGATTG	Double hélice + épingle à cheveux (minoritaire)	

Figure 100 : Séquences d'ADN utilisés pour l'expérience de FID.

Pour effectuer l'expérience de FID dans les meilleures conditions, il est indispensable de connaître l'interaction de la sonde avec la séquence d'ADN à étudier. Dans ce sens, nous avons déterminé la stœchiométrie du complexe thiazole orange/ADN par la méthode de Job sur certaines des séquences étudiées (figure 101).^[179] Cette méthode consiste à faire varier la fraction molaire d'un des deux composés du complexe sans modifier la concentration totale dans le milieu (ici, 2 μ M). On représente alors l'intensité de fluorescence du complexe en fonction de la fraction molaire d'un des deux partenaires (ici l'ADN).

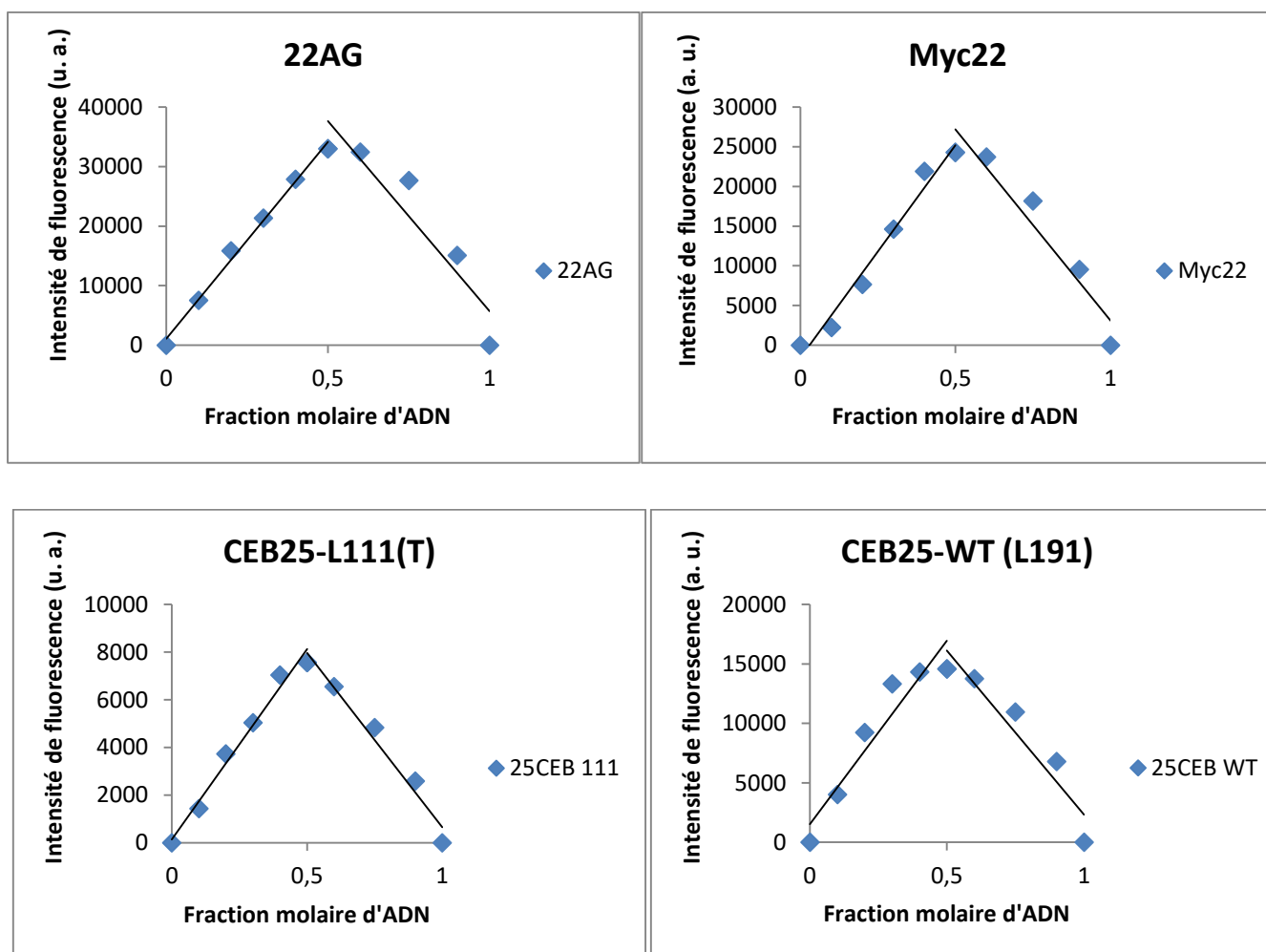


Figure 101 : Détermination de la stœchiométrie du thiazole orange avec les G-quadruplex étudiés par la méthode de Job. (Concentration totale de la sonde et de l'oligonucléotide égale à 2 μ M)

D'après les valeurs obtenues par cette méthode (figure 101), on constate que les droites de tendance prises pour des fractions molaires entre 0 et 0,5 et des fractions molaires entre 0,5 et 1 se croisent, pour toutes les séquences, à une fraction molaire en ADN d'environ 0,5 c'est-à-dire une stœchiométrie 1:1.

Ensuite, toujours dans le but d'avoir le maximum d'information sur l'interaction ADN-TO, les titrations de la sonde fluorescente sur chaque oligonucléotide ont été effectuées afin de connaître la constante d'association et l'exaltation de fluorescence résultante par structure d'ADN (figure 102).

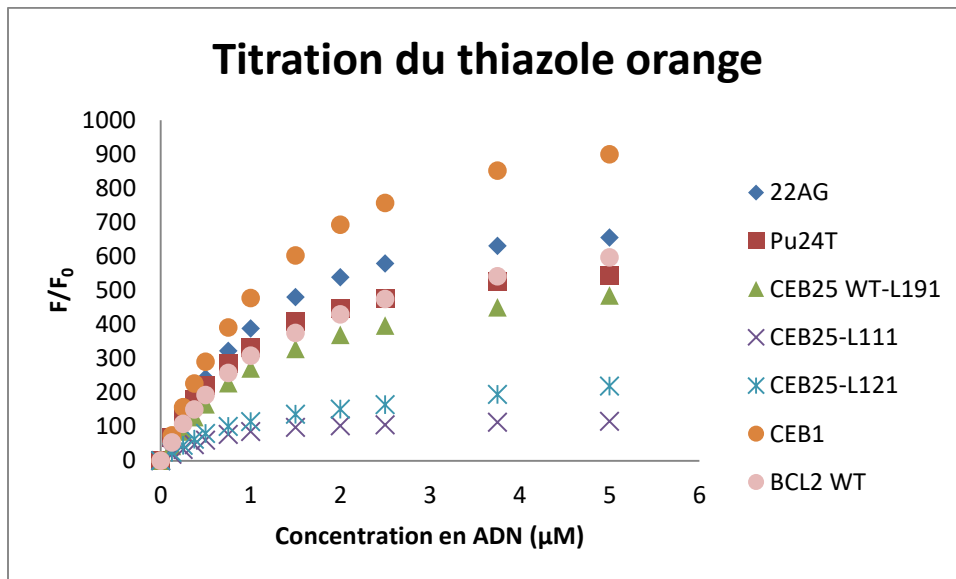


Figure 102 : Titration du thiazole orange (0,5 µM) par les différentes séquences G-quadruplex.

Tout d'abord, l'exaltation de fluorescence du thiazole orange n'est pas la même pour toutes les séquences d'ADN ce qui montre que l'arrangement de la sonde sur le quadruplex n'est pas le même pour chaque oligonucléotide. La planéarité de la sonde qui induit la fluorescence doit être plus ou moins optimale selon l'oligonucléotide (figure 40, page 38). Les exaltations de fluorescence varient de 100 pour la séquence CEB25-L111 à 900 pour le minisatellite CEB1 ce qui montre une grande différence d'exaltation (figures 102 et 103). Il est tout de même intéressant de remarquer que les deux séquences qui montrent les plus fortes exaltations sont celles qui forment des G-quadruplex polymorphes. Ensuite, les valeurs de constantes d'affinité calculées par ces données de fluorescence montrent également que la sonde fluorescente n'a pas la même affinité pour toutes les séquences étudiées. Le modèle utilisé pour évaluer ces constantes d'affinité est celui d'une stœchiométrie 1:1 comme cela a été démontré précédemment. Une variation d'un log est observée, les valeurs du logarithme de la constante d'affinité varient de 5,6 pour le minisatellite modifié CEB25-L121(TT) à 6,3 pour la séquence modifiée du minisatellite CEB25-L111 (figure 103). La valeur de constante d'affinité n'a pas de lien visible avec l'exaltation de fluorescence qui dépend vraiment de son mode d'interaction plutôt que de son affinité.

Fluorophore	ADN	Modèle	R ²	log Ka (M ⁻¹)	F/F ₀
TO	22AG	1:1	0,99971	6,0	650
	Pu24T	1:1	0,99979	6,1	540
	CEB25-L111(T)	1:1	0,9986	6,3	110
	CEB25-L121(TT)	1:1	0,99624	5,6	220
	CEB25-WT (L191)	1:1	0,99918	5,9	490
	BCL 2 WT	1:1	0,99939	5,8	600
	CEB1 WT	1:1	0,99986	5,9	900
	MYC22	1:1	0,99649	6,2	600

Figure 103 : Tableau récapitulatif des constantes d'affinité du thiazole orange pour les différents G-quadruplex.

2) Résultats

Dans ce paragraphe, seules les données de FID sont traitées sans interprétation commune avec l'expérience de FRET-melting.

a) Ligands de référence

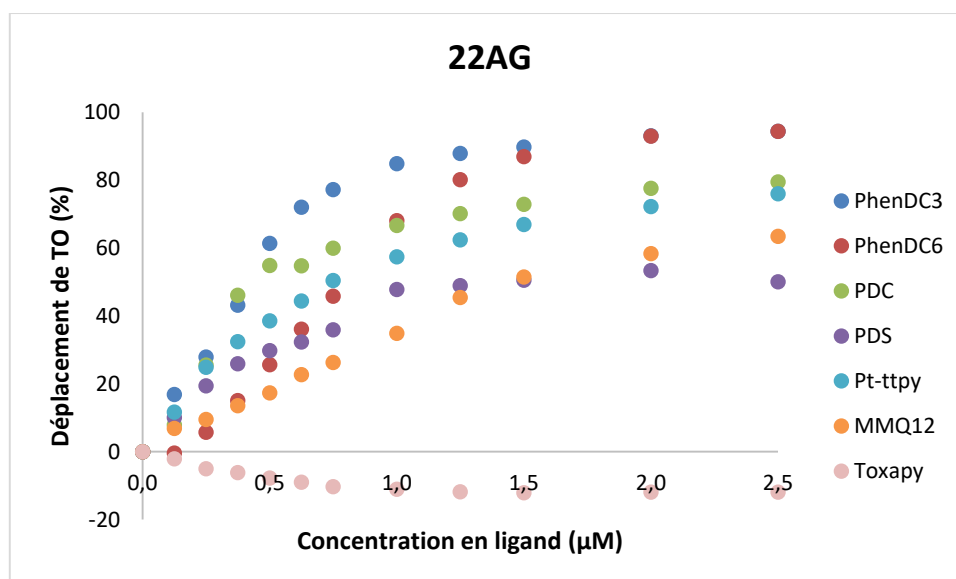


Figure 104 : Résultats de l'expérience de FID sur la séquence télomérique humaine 22AG (concentration TO = 0,5 µM, concentration ADN = 0,25 µM).

Les courbes de déplacement de thiazole orange en fonction de la concentration en ligand ajouté sont obtenues. Il est important de noter que le composé RHPS4 est un ligand fluorescent non compatible avec la sonde thiazole orange (figure 105) car ils absorbent à la longueur d'onde du laser (480 nm). C'est pourquoi sa titration a été effectuée avec la sonde fluorescente TO-PRO-3.^[180]

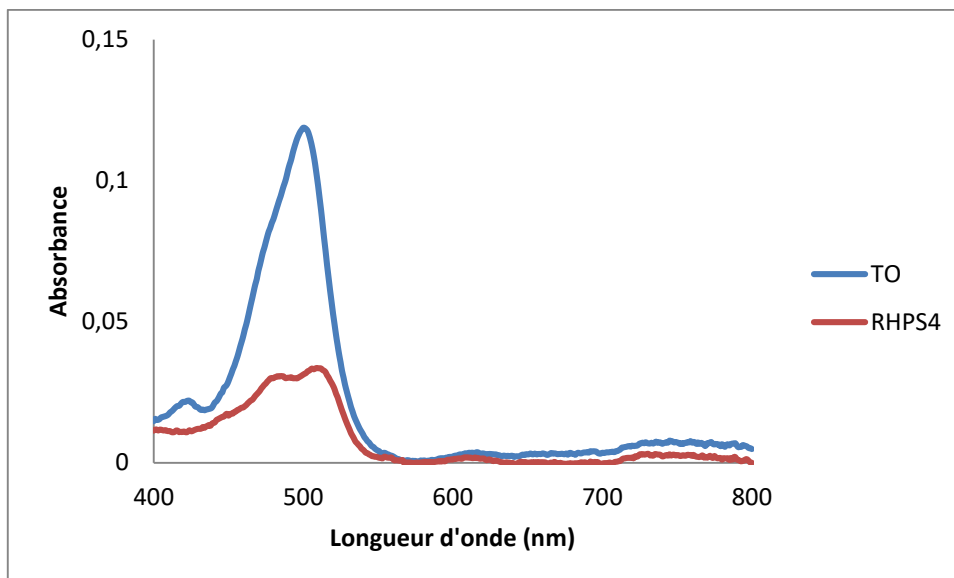


Figure 105 : Spectre UV-visible des composés thiazole orange (en bleu) et RHPS4 (en rouge).

Il n'est donc pas aisé de comparer les résultats de ce ligand avec les autres. Cette sonde TO-PRO-3 présente une exaltation de fluorescence globalement beaucoup plus faible que celle du thiazole orange variant de 93 à 150 pour les séquences étudiées.

La titration des ligands de référence sur la séquence télomérique nous montre une assez grande diversité de déplacement. On propose de classer les ligands en différentes catégories en fonction du déplacement de la sonde :

- les ligands de forte affinité avec un déplacement de la sonde supérieur à 90 % à 10 équivalents de ligands (2,5 μ M) par rapport à l'ADN.
- les ligands de bonne affinité avec un déplacement de la sonde compris entre 60 et 80 % à 10 équivalents de ligands (2,5 μ M).
- les ligands d'affinité modérée pour des déplacements compris entre 40 et 60 % à 10 équivalents (2,5 μ M).
- les ligands peu affins avec des déplacements inférieurs à 40 % à 10 équivalents de ligands (2,5 μ M).

On constate que l'on peut, avec la séquence d'ADN 22AG, classer les ligands en fonction du déplacement de TO (figure 104). La famille PhenDC atteint un déplacement supérieur à 90 % ce qui montre leur très grande affinité pour le G-quadruplex. Sachant que le composé PhenDC₃ se positionne sur le quartet externe avec une interaction d'empilement π - π et déplace à plus de 90% le TO, on peut supposer que le thiazole orange se positionne également sur le quartet externe ou du moins, qu'il interagit avec le quartet externe du G-quadruplex. La deuxième catégorie de ligands contiendrait les composés PDC et Pt-ttpy qui déplacent jusqu'à environ 75-80 % le thiazole orange ce

qui montre également une très bonne affinité. Ensuite, la troisième catégorie avec les composés PDS et MMQ12 qui ne déplacent pas plus de 60 % de la sonde. Ils sont capables de stabiliser fortement les séquences G-quadruplex par l'expérience de FRET-melting mais l'absence d'un plus grand déplacement peut nous interroger sur le mode d'interaction du ligand. Enfin, la dernière catégorie concerne le composé Toxapy qui ne déplace pas du tout la sonde ce qui signifie soit qu'elle n'interagit pas avec la séquence quadruplex, soit que le site d'interaction n'est pas le même que la sonde et ainsi les deux composés peuvent être sur la cible au même moment. Les valeurs de déplacement négatives n'ont pas vraiment de sens. On peut proposer que l'excès de thiazole orange que l'on utilise pour saturer le G-quadruplex est en équilibre avec une seconde position sur la cible ce qui expliquerait une intensité de fluorescence plus élevée ou, qu'en présence de ligand, la sonde est relocalisée sur un autre site de fixation induisant une augmentation de la fluorescence (complexe ternaire).

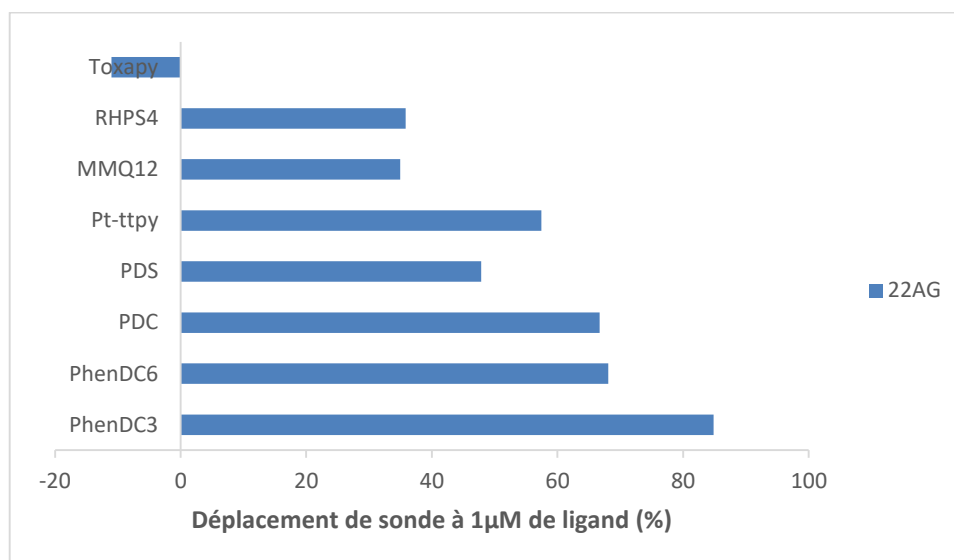


Figure 106 : Résultats de l'expérience de FID à 1 μM de ligand.

Une manière plus simple de représenter les résultats de cette expérience est de prendre la valeur de déplacement de la sonde à une concentration de ligand fixée (figure 106). Nous avons choisi de prendre une concentration de 1 μM car cela correspond à quatre équivalents de ligands par rapport à l'ADN ce qui se rapproche des conditions de FRET-melting où nous utilisons cinq équivalents de ligands. De plus, il s'agit de la concentration de ligand utilisé dans l'expérience de FRET-melting et pour laquelle les ligands atteignent, en général, le plateau de déplacement de la sonde.

On constate avec cette représentation que les catégories décrites précédemment sont toujours valables et que la lecture est simplifiée. Il convient pour les différentes catégories établies de modifier les valeurs de déplacement à 1 μM :

- la première catégorie de forte affinité se caractérise par des ligands ayant un déplacement de sonde supérieur à 60 %.
- la deuxième catégorie pour les ligands ayant un déplacement compris entre 40 et 60 % ce qui constitue la classe de ligand de forte affinité.
- la troisième catégorie pour des déplacements compris entre 20 et 40 % sont les ligands à affinité modérée.
- la quatrième catégorie pour les ligands ayant un déplacement inférieur à 20 % qui sont considérés comme peu affins.

On constate, par ailleurs, que le ligand PhenDC₆ a un déplacement plus faible que le composé PhenDC₃ à 1 μ M alors que ce n'est pas le cas à 2,5 μ M. On peut donc supposer que le composé PhenDC₆ se fixe plus lentement que le composé PhenDC₃ sur la cible ce qui expliquerait cette différence. L'effet cinétique de l'interaction ligand/ADN est minimisé en présence d'excès de PhenDC₆.

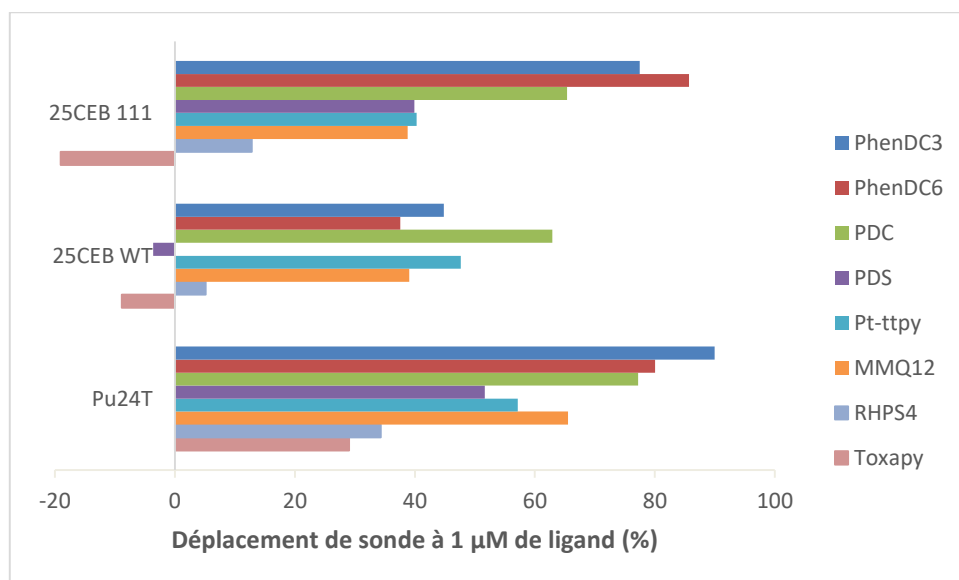


Figure 107 : Résultats de l'expérience de FID des ligands de référence sur différents G-quadruplex.

Avec les résultats de l'expérience de FID sur les mêmes séquences que celles étudiées par FRET-melting, nous pouvons constater que tous les ligands de référence n'ont pas la même affinité ou pas le même mode de fixation que la sonde sur les oligonucléotides (figure 107). En effet, sur la séquence Pu24T, le classement des ligands établi pour la séquence télomérique n'est plus valable. Pour cette séquence, trois catégories sont définies : une première avec des ligands déplaçant fortement (> 80%) la sonde avec la famille PhenDC et le composé PDC, une deuxième avec les composés PDS, Pt-ttpty et MMQ12 qui déplace entre 50 et 60% de la sonde et enfin la dernière avec les composés RHPS4 et Toxapy qui déplacent moins de 40 % de la sonde ce qui signifie que ce sont les ligands les moins

affins pour la cible ou qu'ils n'ont pas le même site de fixation. Le composé Toxapy déplace une fraction du thiazole orange uniquement sur cette séquence.

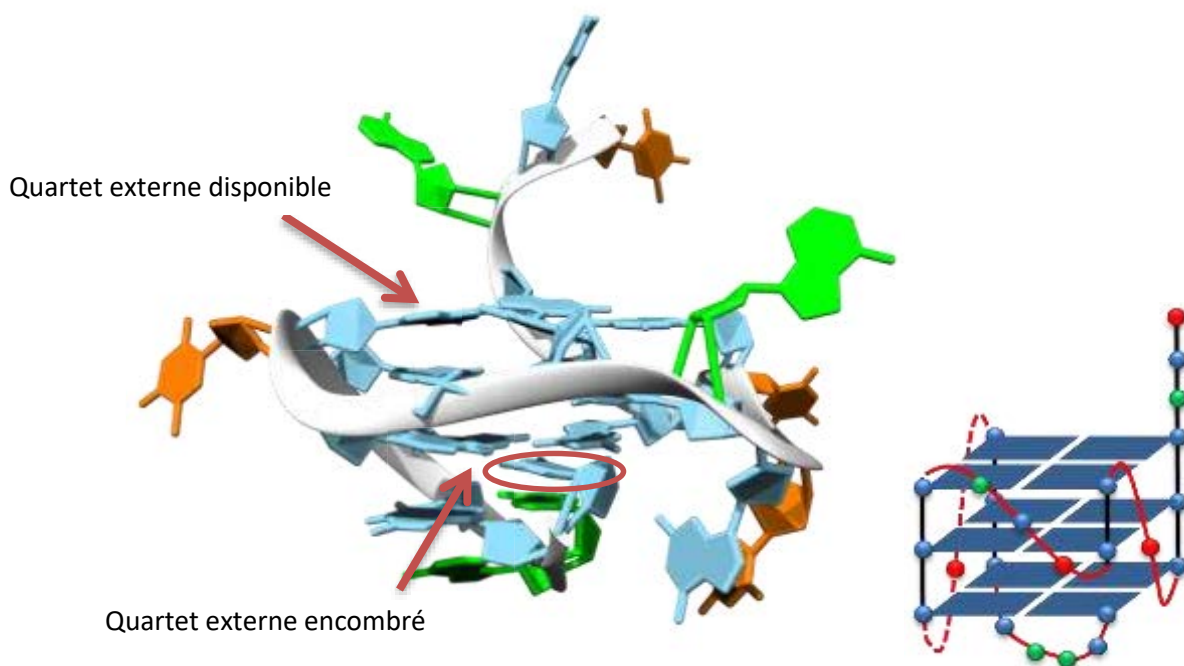


Figure 108 : Structure définie par RMN (PDB : 2MGN) de la séquence Pu24T (TGAGGGTGGTGAGGGTGGGGAACG bases soulignées = boucles).^[70]

On peut supposer que le fait qu'il n'y ait qu'un quartet externe disponible sur cette séquence Pu 24T (figure 108) ait une influence sur le déplacement de la sonde. On a supposé auparavant que le thiazole orange avait un site de fixation sur un quartet externe ou proche de celui-ci car le composé PhenDC₃ le déplace quasiment à 100 %. Le clash stérique généré par la présence du ligand sur le G-quadruplex ne doit pas permettre à la sonde d'être également présente sur la structure pour former un complexe ternaire ce qui pouvait être possible pour les autres séquences étudiées. C'est pourquoi tous les ligands sont capables de déplacer la sonde.

Sur la séquence à longue boucle, le CEB25-WT (L191), on constate qu'il n'y a que deux catégories de ligands, ceux qui peuvent déplacer la sonde et ceux qui ne déplacent que légèrement la sonde. Les composés RHPS4, PDS et Toxapy sont donc dans cette deuxième catégorie et les autres composés déplacent tous environ 40 % de thiazole orange. Il est tout de même important de montrer les courbes de titration pour cette séquence car la boucle semble induire un « retard » de déplacement pour certains ligands comme la famille PhenDC, les composés PDC et Pt-ttpty qui finissent tout de même par avoir un déplacement supérieur à 75 % à de forte concentration en ligand (figure 109). La dynamique de la boucle génère un encombrement stérique supplémentaire. Ainsi, le ligand, à faible concentration, a plus de difficulté pour atteindre son site de fixation cependant, en présence d'un

excès de ligand, l'équilibre est déplacé ce qui permet au ligand de stabiliser la structure G-quadruplex et par conséquent déplacer la sonde fluorescente.

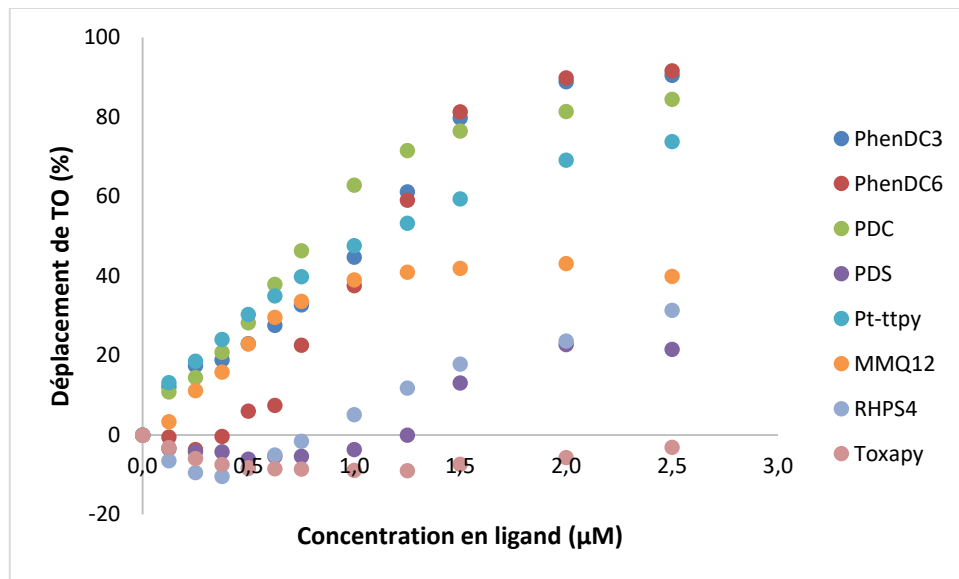


Figure 109 : Courbes de titration des ligands sur la séquence du minisatellite CEB25-WT (L191).

Pour la séquence modifiée du minisatellite CEB25 avec une boucle courte (CEB25-L111(T)), la classification établie sur la séquence télomérique peut être reprise cependant les ligands de chaque catégorie ne sont plus les mêmes. Les meilleurs ligands avec un déplacement supérieur à 80 % restent la famille de composé PhenDC et le ligand PDC classés dans la première catégorie, ensuite, les composés MMQ12 et Pt-ttpty qui déplacent d'environ 50 % la sonde, le composé RHPS4 qui déplace très peu la sonde et enfin le composé Toxapy qui encore une fois a une valeur de déplacement négative. Le composé PDS n'est pas classé car, malgré sa valeur de déplacement positive à 1 μM, sa courbe de titration a une allure très particulière (figure 110) avec un signal de fluorescence instable.

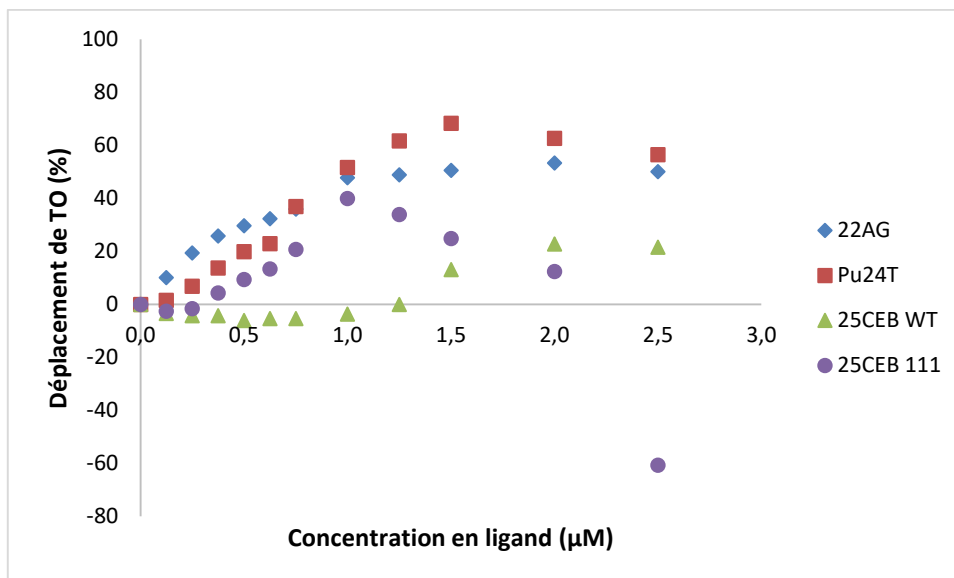


Figure 110 : Courbe de titration du composé PDS sur les différentes séquences étudiées.

On constate finalement que toutes les courbes de titration du composé pyridostatine (PDS) n'ont pas une allure classique (figure 110). En effet, sur toutes les séquences, lorsque la concentration en ligand est élevée le plateau de déplacement n'est pas atteint et le signal de fluorescence a tendance à augmenter de nouveau et induit donc une diminution du déplacement de thiazole orange. Cela est encore plus flagrant sur la séquence très stable CEB25-L111(T). Sur toutes les séquences, le déplacement de la sonde ne dépasse pas les 60 % ce qui montre que le ligand ne peut pas déplacer la totalité du thiazole orange et la diminution de déplacement à haute concentration signifie que le thiazole orange est capable d'exalter de nouveau sa fluorescence. Deux raisons peuvent expliquer ce renouveau d'exaltation, d'une part, le ligand pyridostatine après s'être fixé sur son site d'interaction, en déplaçant en partie la sonde, n'empêche pas la sonde de se repositionner sur le G-quadruplex. D'autre part, la fixation du ligand sur le G-quadruplex pourrait créer un nouveau site de fixation pour la sonde dans lequel son exaltation est possible.

b) Dérivés du PhenDC₃

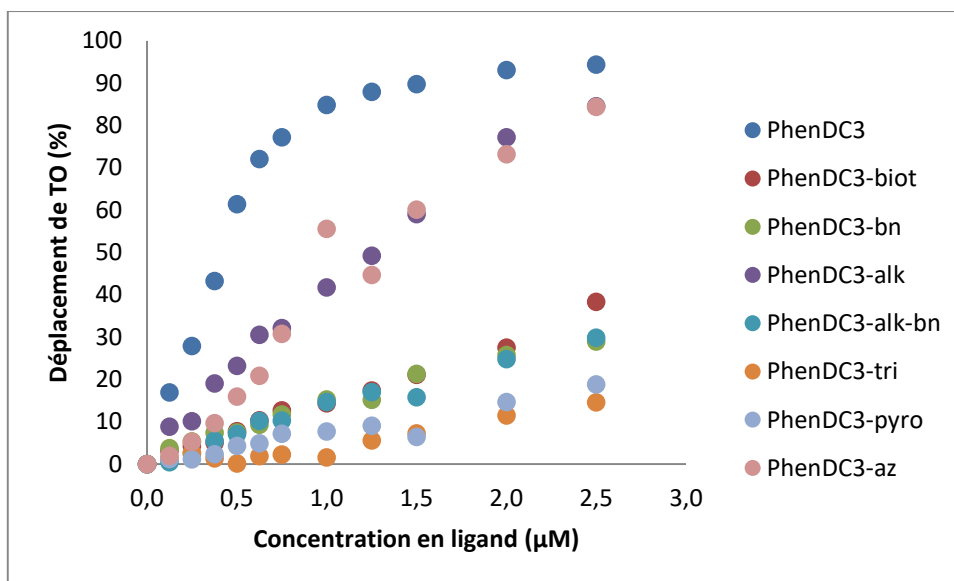


Figure 111 : Résultats de l'expérience de FID sur la séquence télomérique humaine 22AG.

A première vue, trois catégories se distinguent sur les dérivés du PhenDC₃ (figure 111). Tout d'abord, le composé de référence qui déplace plus de 90 % de la sonde. Ensuite, les composés « clickables » - alk et -az qui atteignent un déplacement supérieur à 80 % mais qui ont un « retard » à faible concentration. Cela peut s'expliquer par la chaîne introduite sur la phénanthroline qui doit empêcher le ligand de se positionner aisément sur la cible et déplacer la sonde qui s'y trouve. Cela est encore plus flagrant sur les composés encombrés stériquement qui se classent donc dans la troisième catégorie. En effet, ils ne déplacent pas plus de 40 % de la sonde à 10 équivalents de ligands ce qui montre que la fonctionnalisation affecte l'affinité des composés pour la cible. L'équilibre entre la sonde et le ligand sur la cible est déplacé en faveur de la sonde (figure 112, $K_{A1} \ll K_{A2}$). Dès lors, on peut supposer que les ligands fonctionnalisés et encombrés ont une constante d'affinité de l'ordre de celle du thiazole orange ou légèrement plus faible pour la séquence télomérique humaine.

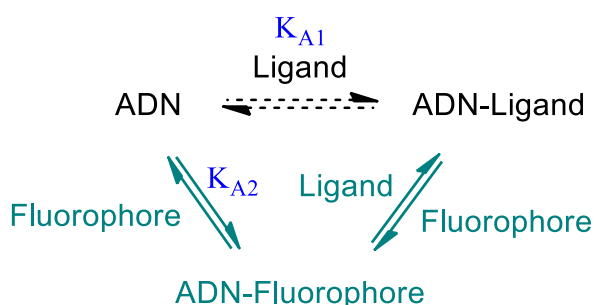


Figure 112 : Equilibre impliqués dans l'expérience de FID.

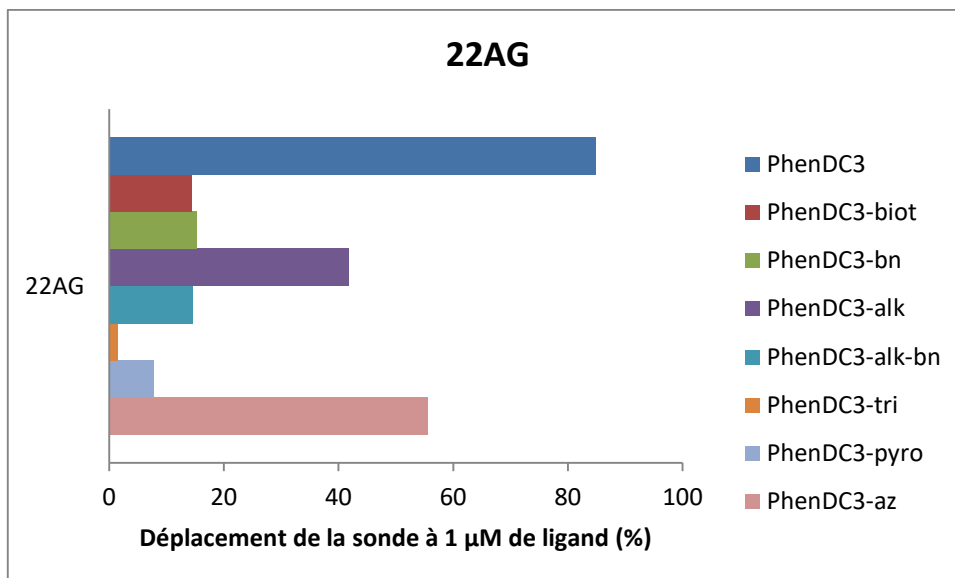


Figure 113 : Résultats de l'expérience de FID à 1 μM de ligand sur la séquence 22AG.

La représentation du déplacement de thiazole orange à 1 μM de ligand (figure 113) confirme la classification établie avec les courbes de titration (figure 111). Cette représentation confirme la forte affinité du composé de référence PhenDC₃ et celles des composés alcyne et azoture.

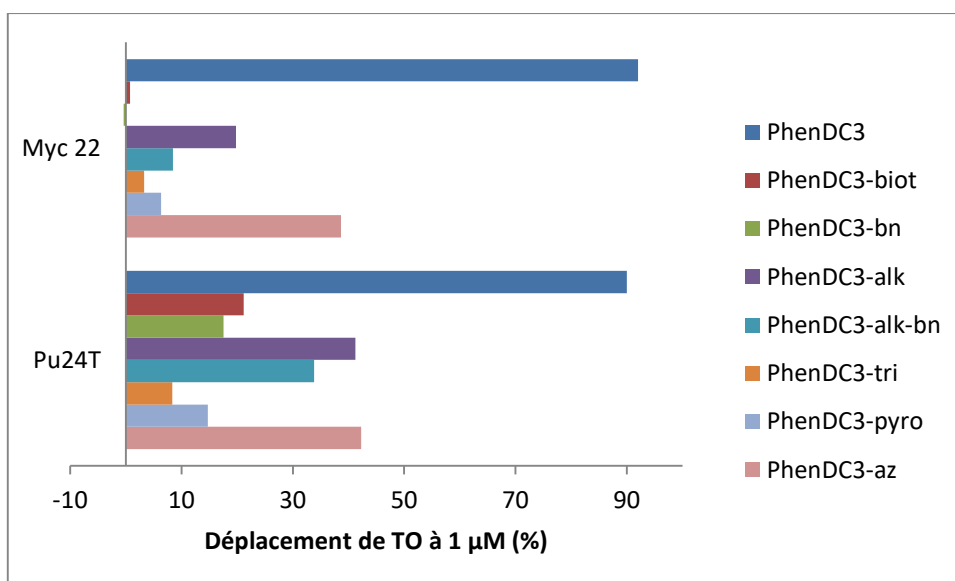


Figure 114 : Résultats de FID à 1 μM de ligands sur les séquences modifiées de l'oncogène c-myc, Myc22 et Pu24T.

Les deux séquences modifiées de l'oncogène c-myc sont très différentes structurellement et cela se ressent sur les résultats de l'expérience de FID (figure 114). En effet, on constate que le déplacement du thiazole orange sur la séquence myc 22 est beaucoup plus difficile que pour la séquence Pu24T pour tous les composés dérivatisés. La structure n'a aucune influence sur le composé de référence PhenDC₃ et le dérivé azoture. Malgré ces résultats, la classification des composés fonctionnalisés restent valable pour les différentes séquences. Les composés azoture et alcyne non-encombrés par

un groupement supplémentaire sont toujours les composés les moins affectés par la fonctionnalisation.

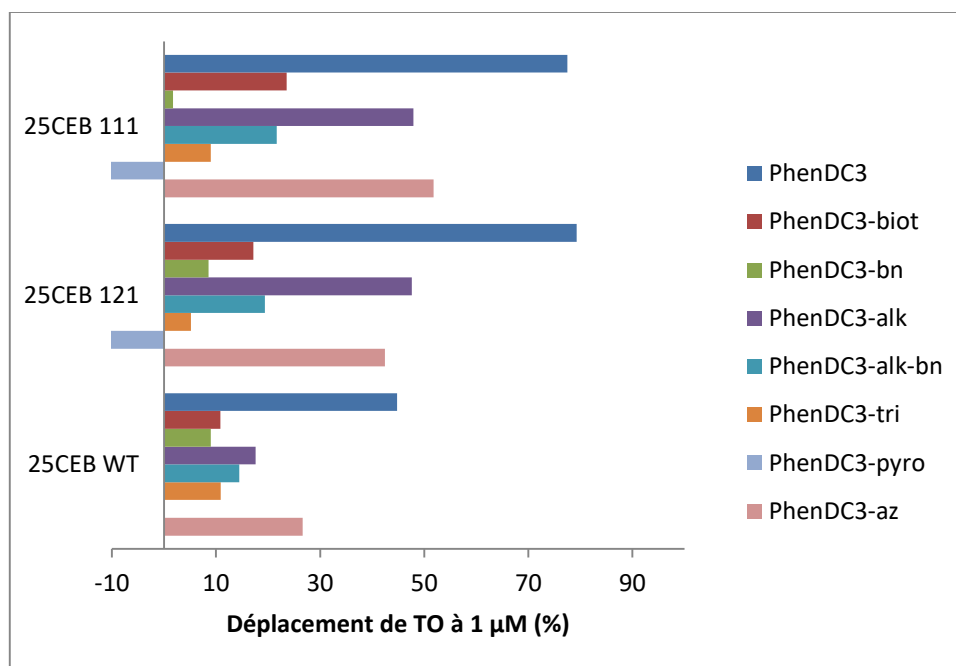


Figure 115 : Résultats de l'expérience de FID à 1 μM de ligand sur les différentes séquences du minisatellite CEB25.

Le minisatellite CEB25 a également été étudié pour constater l'influence de la taille de la boucle centrale sur l'expérience de FID (figure 115). Sur les deux séquences à boucles courtes, la classification des ligands reste la même que précédemment avec le composé de référence qui déplace environ 80 % de la sonde, les composés « clickables » non encombrés qui déplacent environ 50 % de la sonde et enfin les composés fonctionnalisés qui déplacent moins de 30 % du thiazole orange. Ensuite, sur la séquence native à longue boucle (neuf nucléotides), les composés déplacent très peu la sonde à 1 μM de ligand ce qui montre bien que la boucle empêche le ligand de prendre sa position préférentielle et ainsi déplacer la sonde. A plus forte concentration (figure 116), certains ligands (les modifiables *in situ*) ont un déplacement plus élevé. L'ajout d'excès de ligand favorise l'équilibre en faveur du ligand et le déplacement de la sonde fluorescente (figure 111, $K_{A1} \gg K_{A2}$). La dynamique de la boucle doit intervenir dans le positionnement plus ou moins aisé du ligand sur le G-quadruplex et comme la sonde, qui est une petite molécule sans chaîne et qui doit se positionner beaucoup plus facilement, est déjà positionnée sur la structure quadruplex il est encore plus délicat pour le ligand de prendre sa position préférentielle.

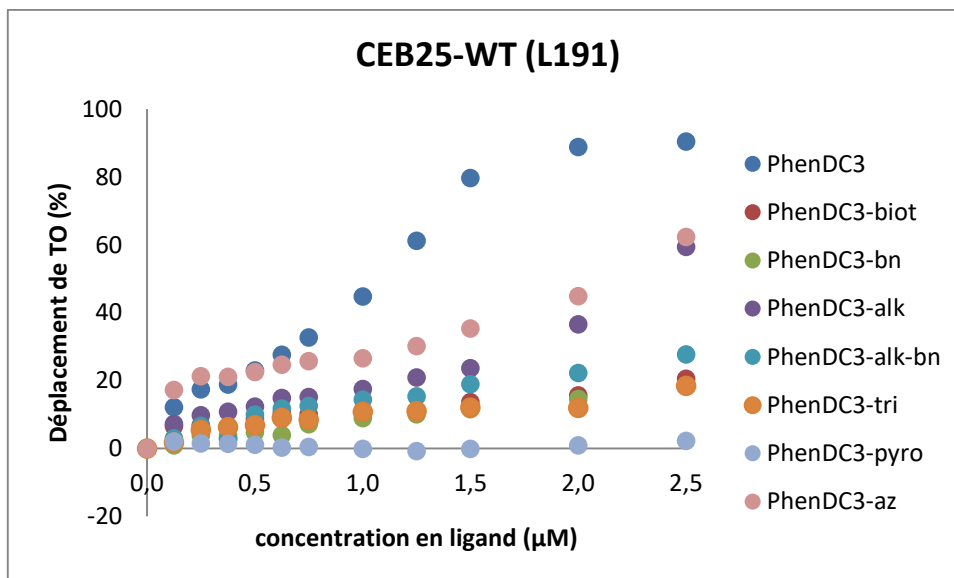


Figure 116 : Courbes de titrations des ligands sur la séquence native du minisatellite CEB25-WT (L191).

La boucle a donc un effet sur l'affinité de plusieurs ligands et on peut constater que plus la boucle est grande, moins le déplacement de la sonde fluorescente est aisé (Figure 115). Cela se confirme par l'étude d'une seconde structure G-quadruplex à longue boucle (12 bases), le BCL2 (figure 117). Sur cet oligonucléotide comme pour le CEB25-WT (L191), le déplacement de la sonde ne dépasse pas 50 % à 1 μM de ligand. Cependant, à 10 équivalents de ligands, PhenDC₃ déplace plus de 80 % de thiazole orange et les composés « clickables » atteignent environ 60 % de déplacement.

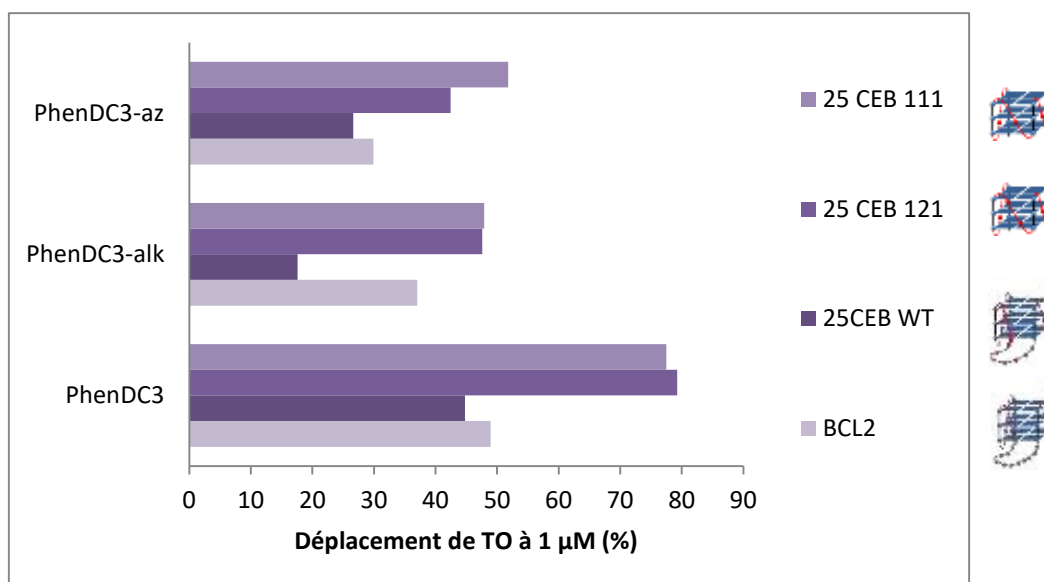


Figure 117 : Influence de la taille de la boucle sur les dérivés de PhenDC₃.

c) Dérivés du PDC

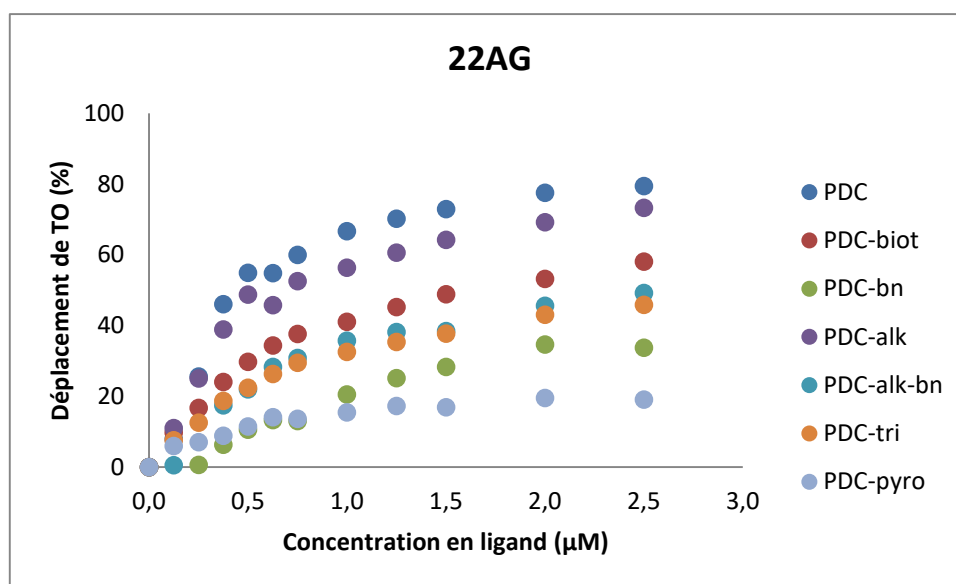


Figure 118 : Courbes de titrations des dérivés du PDC sur la séquence télomérique humaine 22AG.

Le composé de référence PDC et le composé alcyne sont les deux meilleurs ligands pour déplacer la sonde fluorescente (~80% de déplacement) sur la séquence télomérique humaine (figure 118). Les composés fonctionnalisés par une biotine, un alcyne et une benzophénone ou le tri-fonctionnel déplacent environ 50 % de thiazole orange à 10 équivalents de ligands par rapport à la quantité d'ADN. Le composé fonctionnalisé par une benzophénone à l'extrémité d'une chaîne aliphatique de quatre carbones ne déplace pas plus de 35% de la sonde. Le composé PDC-pyro non chargé sur les quinoléines déplace encore moins la sonde fluorescente ce qui suggère que les chaînes protonables in situ peuvent interagir sur un autre site de fixation du G-quadruplex. La classification des ligands sur les dérivés du composé PDC est donc moins évidente que celle établie avec les dérivés du composé PhenDC₃ mais cela confirme les résultats de l'expérience de FRET-melting qui nous ont montré que l'affinité est beaucoup plus affectée pour les composés fonctionnalisés du PhenDC₃ que ceux du PDC.

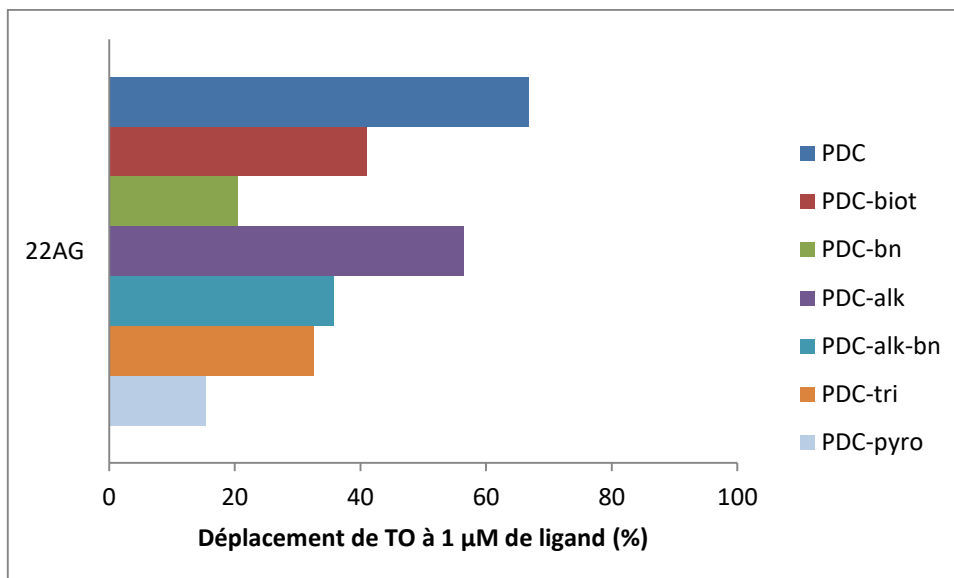


Figure 119 : Résultats de l'expérience de FID à 1 μM de ligand sur la séquence télomérique humaine.

On constate que la représentation du déplacement de la sonde à 1 μM de ligand (Figure 119) nous donne les mêmes informations que les courbes de titration et nous permet d'établir la même classification que précédemment. Plus le groupement est éloigné de la pyridine, plus le ligand est capable de déplacer la sonde fluorescente. Le composé PDC-biotine (figure 120) qui possède un groupement biotine à l'extrémité d'une chaîne PEG est peu affecté (déplacement supérieur à 40 % pour tous les oligonucléotides étudiés, figure 119, 121 et 122) alors que le composé PDC-bn qui possède un groupement benzophénone à l'extrémité d'une chaîne de quatre carbones est fortement affecté (déplacement inférieur à 30 % pour tous les oligonucléotides étudiés, figure 119, 121 et 122). La fonction benzophénone peut être impliquée dans des liaisons hydrogène avec les bases des boucles ou s'empiler sur le quartet et ainsi gêner le bon placement du ligand.^[177] Par ailleurs, le composé ne possédant qu'une chaîne aliphatique, PDC-alk (figure 120), n'est pas affecté ce qui montre que l'encombrement stérique du groupement ajouté influe sur l'affinité du ligand pour la cible.

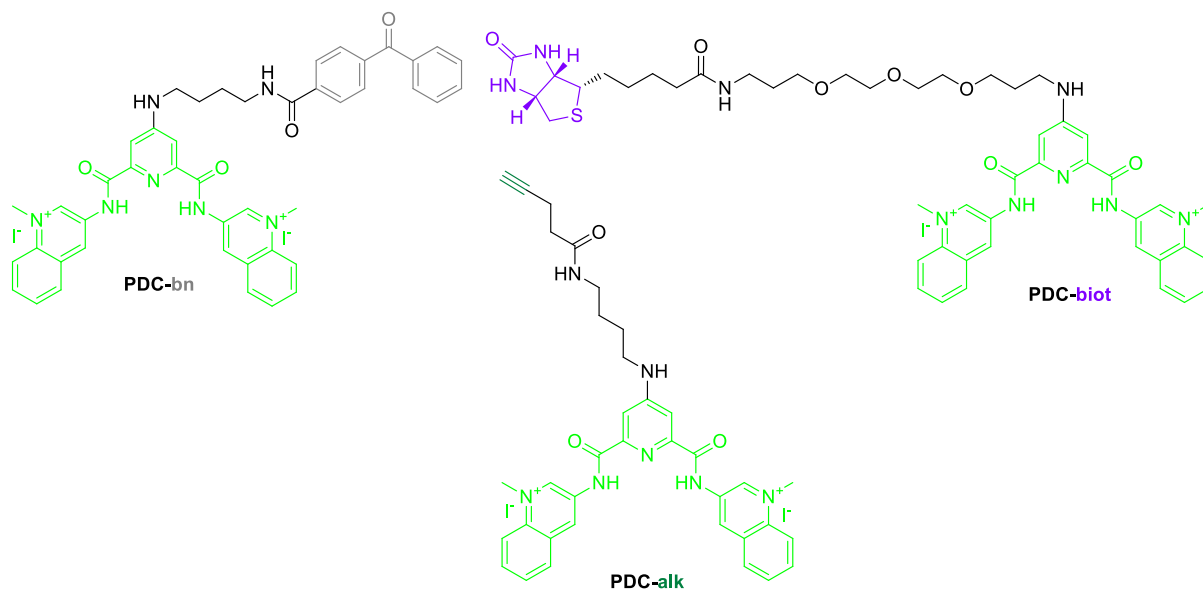


Figure 120 : Structures des composés PDC-bn, PDC-biot et PDC-alk.

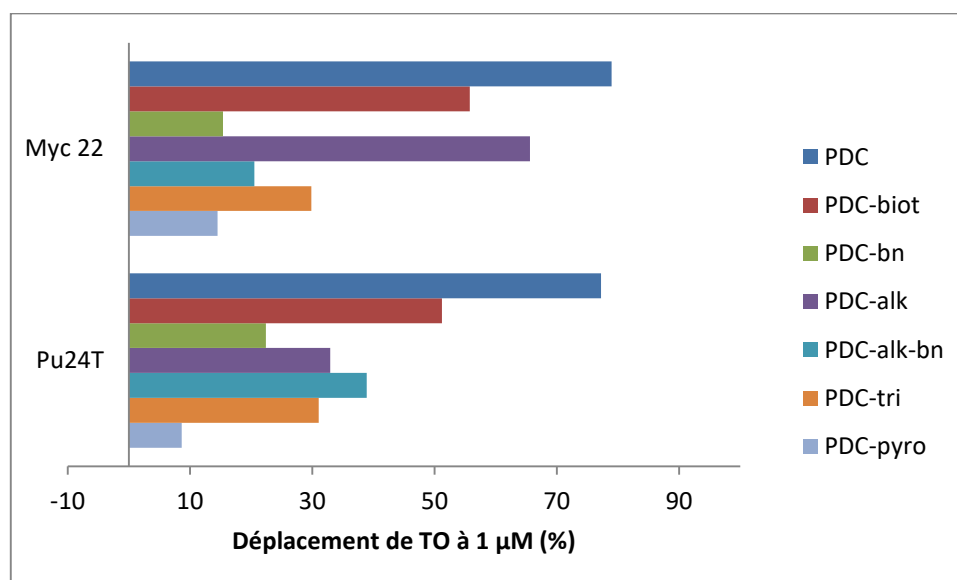


Figure 121 : Résultats de l'expérience de FID à 1 μM de ligand sur les séquences de l'oncogène c-myc.

Sur les séquences modifiées de l'oncogène c-myc (figure 121), le dérivé PDC-biot est le composé fonctionnalisé qui déplace le mieux la sonde fluorescente (> 50 % de déplacement à 1 μM).

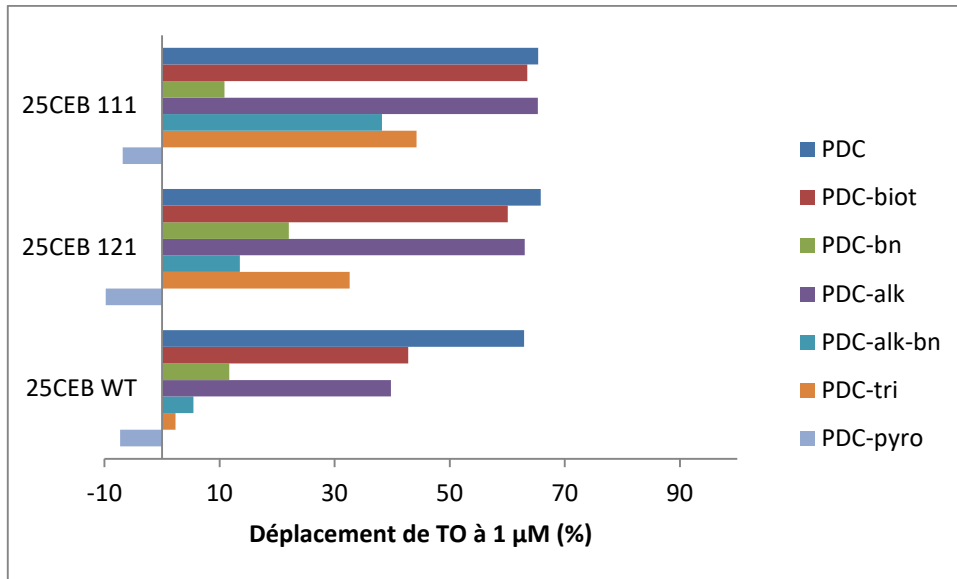


Figure 122 : Résultats de l'expérience de FID à 1 μM de ligand sur les différentes séquences du minisatellite CEB25.

On peut constater que la tendance reste la même sur le minisatellite CEB25 (figure 122). En effet, les composés biotinylé et alcyne sont les deux dérivés les moins affectés par la fonctionnalisation. On note tout de même que lorsque la boucle est longue sur la séquence native, il y a une différence entre le composé de référence PDC et ses dérivés alors que quand la boucle est courte (une ou deux bases), les deux composés les moins affectés ont le même déplacement que la référence.

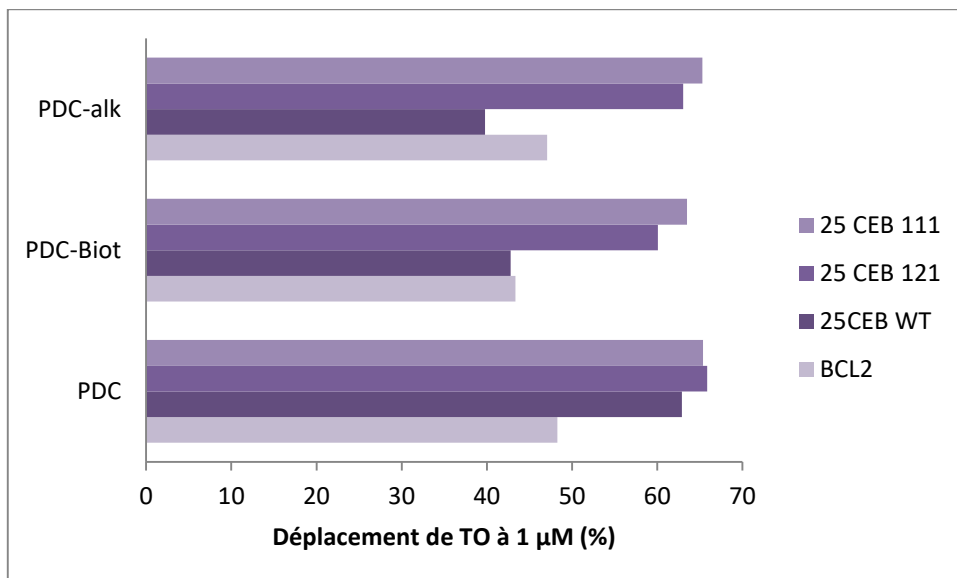


Figure 123 : Influence de la taille de la boucle sur les dérivés de PDC.

Le composé de référence PDC n'a pas l'air impacté par la taille de la boucle sur les différentes séquences du minisatellite CEB25 cependant en présence d'un quadruplex ayant une boucle encore plus longue (12 bases), le BCL 2, on constate un déplacement plus faible de la sonde (figure 123). Le composé fonctionnalisé par une biotine et celui fonctionnalisé par un alcyne conservent de bonnes

affinités pour les différentes séquences avec plus de 40 % de déplacement même sur les G4 à longues boucles. On note tout de même une influence de la boucle sur le déplacement avec un déplacement de thiazole orange plus élevé pour les structures à boucles courtes.

3) Conclusions

En conclusion, on constate que si on prend l'expérience de FID seule, il est très compliqué d'établir un classement des ligands sur différentes séquences d'ADN pour les composés évalués dans cette étude. Bien entendu, la catégorisation par séquence est possible ce qui nous permet d'avoir une estimation de l'affinité des ligands par séquence. Cela signifie que certains ligands sont dépendants des séquences et des structures G-quadruplex comme cela avait été montré par l'expérience de FRET-melting. En particulier, le composé PDS ne déplace pas plus de 60 % de la sonde fluorescente et, à haute concentration en ligand, engendre une forte exaltation du thiazole orange. Il s'agit du seul composé évalué qui possède de telles courbes de titration. Ces résultats peuvent nous interroger sur le mode de fixation de ce ligand qui possède des chaînes latérales aminées qui sont chargées à pH physiologique. Ainsi, les interactions électrostatiques sont fortement impliquées dans l'affinité de ce composé pour les structures G-quadruplex.

L'évaluation des dérivés du PhenDC₃ nous montre que seuls les composés sans gêne stérique ni fonction sont capables de déplacer au moins 50 % de la sonde fluorescente sur tous les oligonucléotides. Les composés fonctionnalisés sont vraiment très affectés et perdent fortement en affinité par rapport à la référence PhenDC₃, cela suggère qu'ils ont une constante d'affinité proche ou inférieure à celle du thiazole orange. Les dérivés du PDC sont quant à eux beaucoup plus robustes notamment le composé biotinylé qui déplace la sonde fluorescente aussi bien que le composé alcyne (fonction alcyne non encombrante). La longueur de l'espaceur entre la pyridine et la fonction semble avoir une grande influence sur l'affinité du ligand. Un espaceur court (chaîne à quatre carbones) avec un groupement encombrant (benzophénone) implique une baisse d'affinité pour la cible. *A contrario*, un long espaceur (PEG) et une fonction peu encombrante (biotine) permet de conserver une haute affinité.

III. Comparaison de ligands en combinant les deux méthodes

Les résultats des deux méthodes biophysiques d'analyse d'interactions ligand/G-quadruplex pris indépendamment l'une par rapport à l'autre nous donne des résultats globalement cohérents avec des tendances générales identiques pour les deux méthodes. Cependant, en combinant les deux méthodes pourrait-on distinguer une tendance différente ou pourrait-on obtenir des informations supplémentaires sur le mode de fixation d'un ligand et sa façon de stabiliser le quadruplex ?

Pour répondre à ces questions, un index a été calculé en multipliant le pourcentage de déplacement de thiazole orange à 1 μM de ligand avec la valeur de température de fusion de l'ADN en présence du ligand.

Figure de mérite : **Index = Déplacement de TO à 1 μM de ligand (%) x T_m (G4+ligand) ($^{\circ}\text{C}$)**

Nous avons choisi de prendre la valeur de température de fusion et non la valeur de ΔT_m pour que l'index soit applicable indépendamment des séquences utilisées. En effet, la valeur de ΔT_m est très variable selon l'oligonucléotide. Lorsque la structure G-quadruplex est stable seule dans le milieu alors elle aura une température de fusion élevée (ex : 72 $^{\circ}\text{C}$ pour la séquence CEB25-L111(T)) et donc des valeurs de ΔT_m ne pouvant pas dépasser les 15 $^{\circ}\text{C}$. *A contrario*, lorsque la séquence d'ADN quadruplex est moins stable ($T_m \sim 55$ $^{\circ}\text{C}$) alors la valeur de ΔT_m peut être supérieure à 30 $^{\circ}\text{C}$. Ainsi, les index ne seraient pas comparables avec les valeurs de ΔT_m .

1) Ligands de références

Avant de commencer à combiner les données des expériences de FRET-melting et de FID, nous allons analyser les différentes valeurs de constante d'affinité existantes pour certains des ligands de ce panel. Pour être cohérent, nous allons utiliser les données d'une même technique, la SPR, avec les données du groupe d'Eric Defrancq publiées cette année.^[112c] Les oligonucléotides (figure 124) sont fixés sur la surface ainsi les structures G-quadruplex formées n'ont qu'un quartet externe disponible pour fixer le ligand.

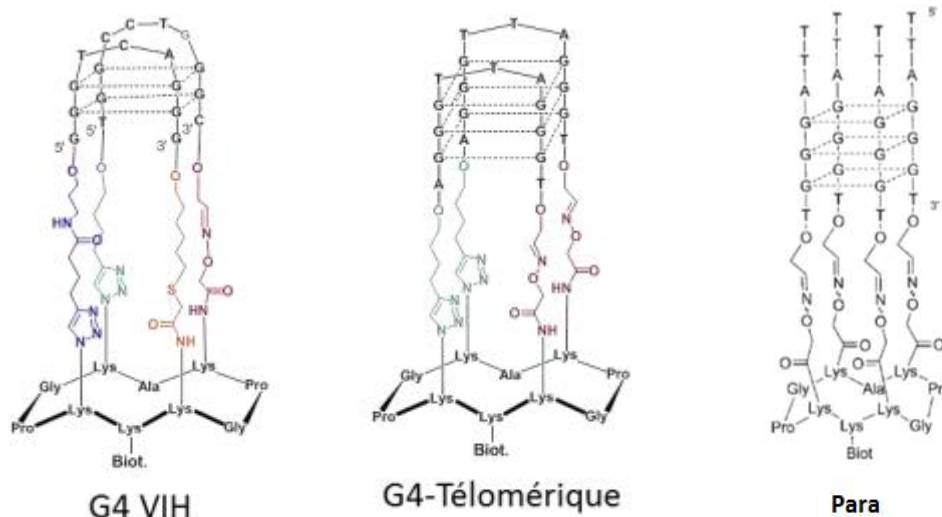


Figure 124 : Structures des oligonucléotides utilisés en SPR par l'équipe d'Eric Defrancq.

Ligands	G4 VIH			G4 télomérique			Para			duplex en épingle		
	k_{on} ($M^{-1}.s^{-1}$)	k_{off} (s^{-1})	K_D (M)	k_{on} ($M^{-1}.s^{-1}$)	k_{off} (s^{-1})	K_D (M)	k_{on} ($M^{-1}.s^{-1}$)	k_{off} (s^{-1})	K_D (M)	k_{on} ($M^{-1}.s^{-1}$)	k_{off} (s^{-1})	K_D (M)
MMQ1	$1,8 \pm 0,4$ $.10^5$	$3,3 \pm 1,1$ $.10^{-1}$	$1,9 \pm 0,8$ $.10^{-6}$	$1,4 \pm 0,4$ $.10^5$	$5,4 \pm 1,4$ $.10^{-1}$	$4,5 \pm 1,9$ $.10^{-6}$	nc	nc	nc	$9,2 \pm 3,3$ $.10^4$	$3,8 \pm 0,6$ $.10^{-1}$	$4,9 \pm 1,3$ $.10^{-6}$
PDS	$6,0 \pm 0,0$ $.10^4$	$1,5 \pm 0,4$ $.10^{-3}$	$2,4 \pm 0,7$ $.10^{-9}$	$2,1 \pm 1,2$ $.10^5$	$1,9 \pm 0,3$ $.10^{-3}$	$1,2 \pm 0,6$ $.10^{-8}$	$1,4 \pm 0,7$ $.10^5$	$2,0 \pm 1,0$ $.10^{-4}$	$1,8 \pm 0,2$ $.10^{-9}$	nd	nd	nd
PhenDC ₃	$2,6 \pm 1,5$ $.10^7$	$5,3 \pm 2,9$ $.10^{-2}$	$2,1 \pm 0,1$ $.10^{-9}$	$4,3 \pm 1,5$ $.10^7$	$9,6 \pm 2,4$ $.10^{-2}$	$2,4 \pm 0,3$ $.10^{-9}$	$5,9 \pm 0,5$ $.10^7$	$1,1 \pm 0,3$ $.10^{-1}$	$1,8 \pm 0,4$ $.10^{-9}$	nd	nd	nd

Figure 125 : Résultats de SPR obtenus par le groupe d'Eric Defrancq. D'après [112c], nc = non communiqué et nd = non déterminable.

Parmi les huit composés de référence que nous avons évalués par les deux méthodes biophysiques précédemment décrites, des valeurs de constante de dissociation ne sont disponibles que pour deux composés, PhenDC₃ et PDS (figure 124). Un composé de la même famille que le composé MMQ12 (figure 125) a également été évalué ce qui nous permettra d'établir le classement des composés en fonction des valeurs de constante de dissociation. Le composé MMQ12 devrait avoir une meilleure interaction avec les structures G-quadruplex car, en plus de l'empilement sur le quartet, les amines primaires permettent des interactions électrostatiques qui vont renforcer l'affinité. En fonction de ces trois composés, nous allons essayer d'établir un classement des différents ligands de G-quadruplex.

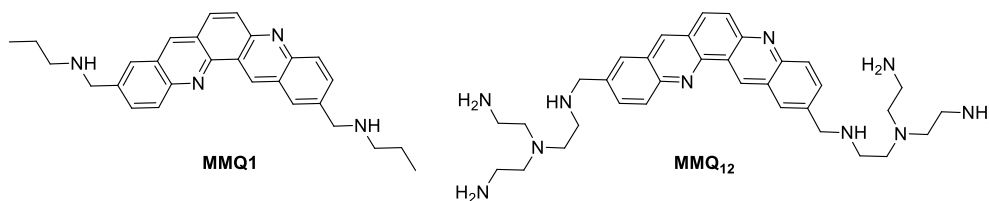


Figure 126 : Structures des composés MMQ1 et MMQ12.

Les résultats de SPR (figure 124) nous confirment que les composés MMQ sont des ligands de G-quadruplex avec une constante de dissociation de l'ordre du micromolaire ($K_D = 4,5.10^{-6}$ M pour le G4 télomérique) et que cette famille n'est pas sélective, elle peut également interagir avec le duplex avec une constante de dissociation du même ordre de grandeur ($K_D = 4,9.10^{-6}$ M pour le duplex en épingle). Le ligand PhenDC₃ est le composé le plus affiné pour tout type de G-quadruplex avec une constante de dissociation de l'ordre du nanomolaire ($K_D = 2,4.10^{-9}$ M pour le G4 télomérique, $K_D = 2,1.10^{-9}$ M pour le G4 du VIH et $K_D = 1,8.10^{-9}$ M pour le G4 para) alors que celle du composé PDS varie d'un ordre de grandeur en fonction de l'oligonucléotide ($K_D = 1,2.10^{-8}$ M pour le G4 télomérique, $K_D = 2,4.10^{-8}$ M pour le G4 du VIH et $K_D = 1,8.10^{-9}$ M pour le G4 para). Cela montre bien que la molécule pyridostatine est dépendante de la séquence étudiée malgré une affinité élevée pour les G-

quadruplex. Ces deux composés n'ont aucune affinité pour le duplex et la valeur de constante de dissociation n'a pas pu être déterminée par cette méthode.

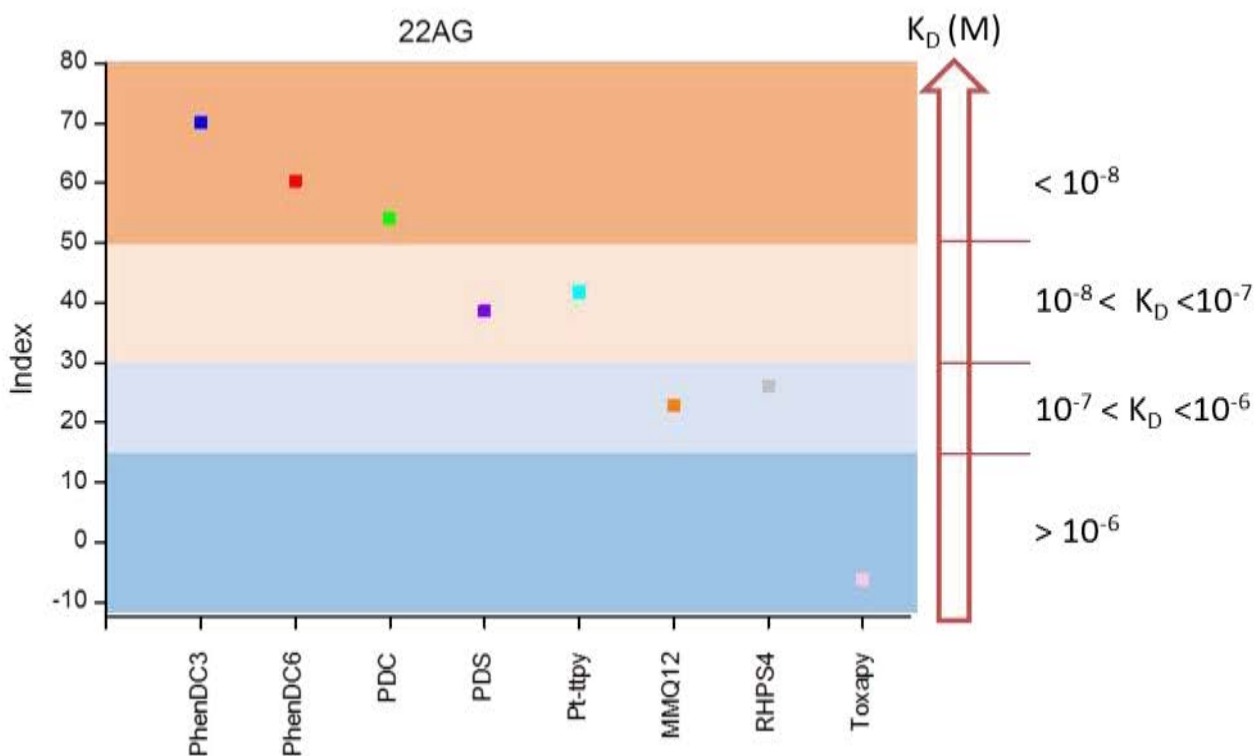


Figure 127 : Représentation de l'index issu des deux tests biophysiques en fonction des ligands de G4 sur la séquence télomérique humaine 22AG.

La première séquence d'ADN G-quadruplex étudiée est la séquence télomérique humaine. Quatre catégories ont été définies suite aux résultats obtenus par cette représentation graphique des index (figure 127) :

- la première pour les index supérieurs à 50 (en orange sur les figures) avec les composés PDC et la famille PhenDC.
- la deuxième catégorie avec les composés PDS et Pt-ttpy pour des index compris entre 30 et 50 (en beige sur les figures).
- la troisième catégorie pour des index compris entre 15 et 30 (en gris clair sur les figures) avec les composés MMQ12 et RHPS4.
- la quatrième (en bleu sur les figures) comprenant le composé Toxapy pour des index inférieurs à 15.

En faisant une comparaison entre ces catégories et les valeurs de constantes de dissociation connues, on peut estimer que la première catégorie correspond à une constante de dissociation de l'ordre du nanomolaire, la seconde une constante de l'ordre de la dizaine à la centaine de nanomolaire, la troisième de l'ordre du micromolaire et enfin la dernière une constante supérieure au micromolaire. De plus, nous savons que la constante d'affinité du thiazole orange sur cette

séquence est de l'ordre du micromolaire ($K_A = 2.10^6 \text{ mol}^{-1}$)^[100b] ainsi, le fait que l'index soit négatif pour le composé Toxapy montre qu'il a une constante de dissociation plus élevée ce qui est cohérent avec la quatrième catégorie évoquée.

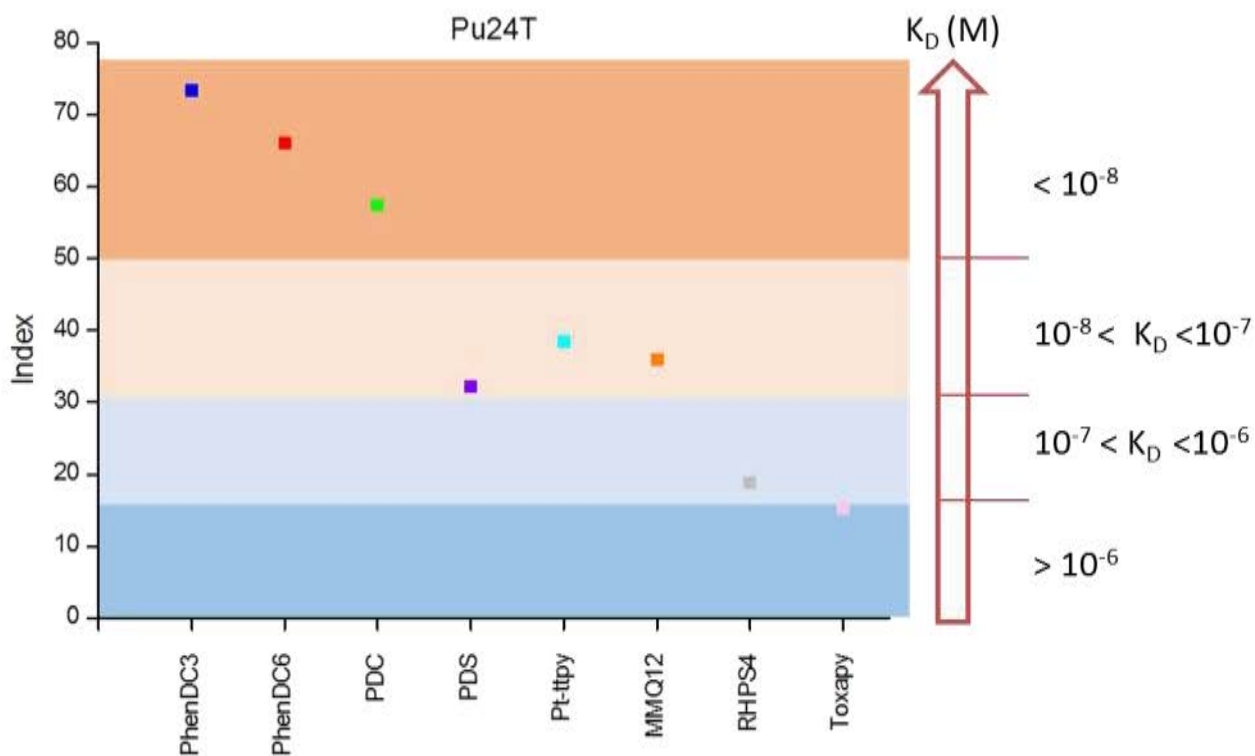


Figure 128 : Représentation de l'index issu des deux tests biophysiques en fonction des ligands de G4 sur la séquence modifiée de l'oncogène c-myc, Pu24T.

Pour la séquence Pu24T (figure 128), on constate que les molécules sont classées dans trois des quatre catégories présentées pour la séquence 22AG. PDC et la famille PhenDC restent à un index supérieur à 50, on peut donc supposer que ces molécules ont une constante de dissociation de l'ordre du nanomolaire. Les ligands qui ont un index entre 30 et 50, comprenant PDS, MMQ12 et Pt-ttpy, sont dans la deuxième catégorie et correspondent à une constante de l'ordre de la dizaine à la centaine de nanomolaire. Enfin, avec un index entre 15 et 30, les ligands RHPS4 et Toxapy ont une constante de l'ordre du micromolaire pour cette séquence. Ces résultats montrent la dépendance de certains ligands dits « de référence » pour les oligonucléotides.

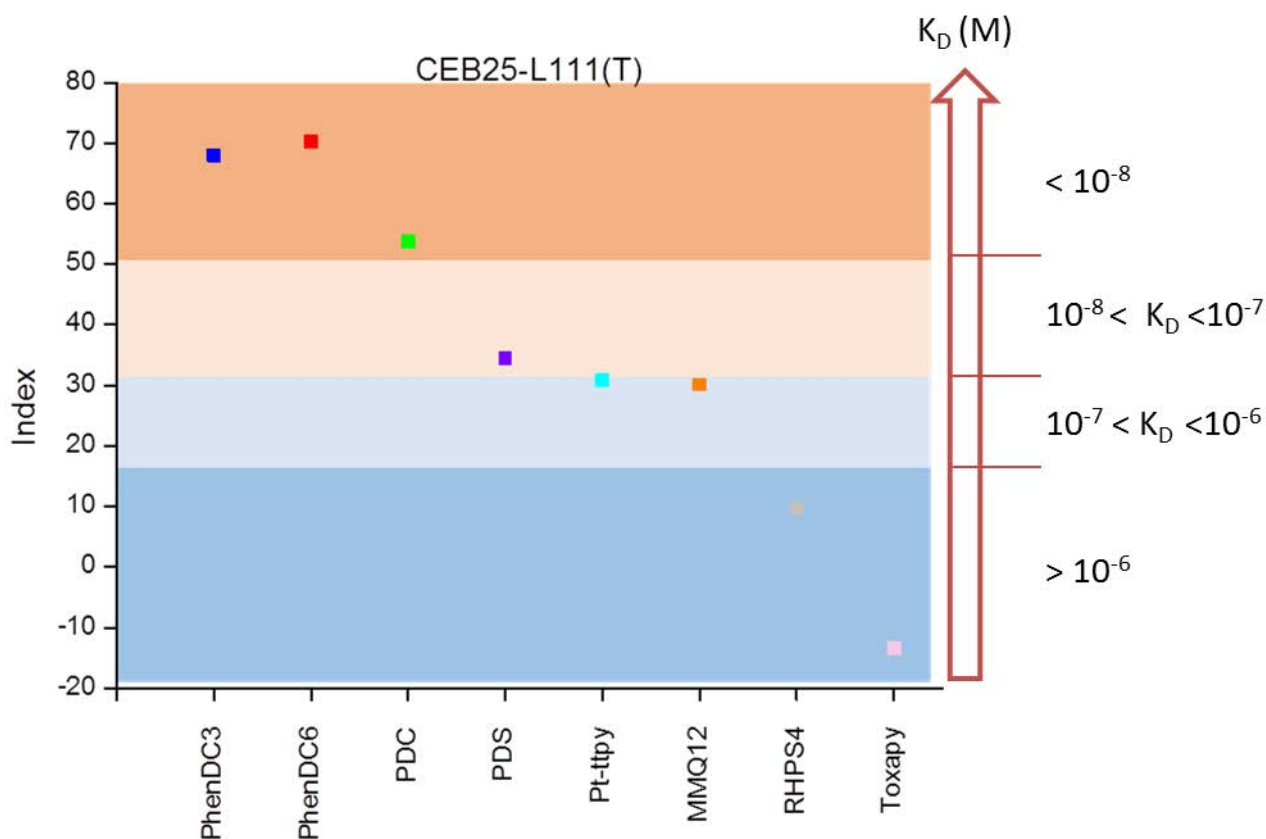


Figure 129 : Représentation de l'index issu des deux tests biophysiques en fonction des ligands de G4 sur la séquence modifiée du minisatellite CEB25-L111(T).

L'étude du minisatellite CEB25 montre que lorsque la boucle centrale est réduite (un nucléotide), le classement des ligands est à peu près identique (figure 129) à celui réalisé pour la séquence télomérique figure 127) avec les trois meilleurs ligands (PhenDC₃, PhenDC₆ et PDC) dans la première catégorie. Suivraient, ensuite, trois ligands (PDS, MMQ12 et PDS) dans la deuxième catégorie, les composés RHPS4 et Toxapy sont dans la quatrième et dernière catégorie avec des index inférieurs à 15.

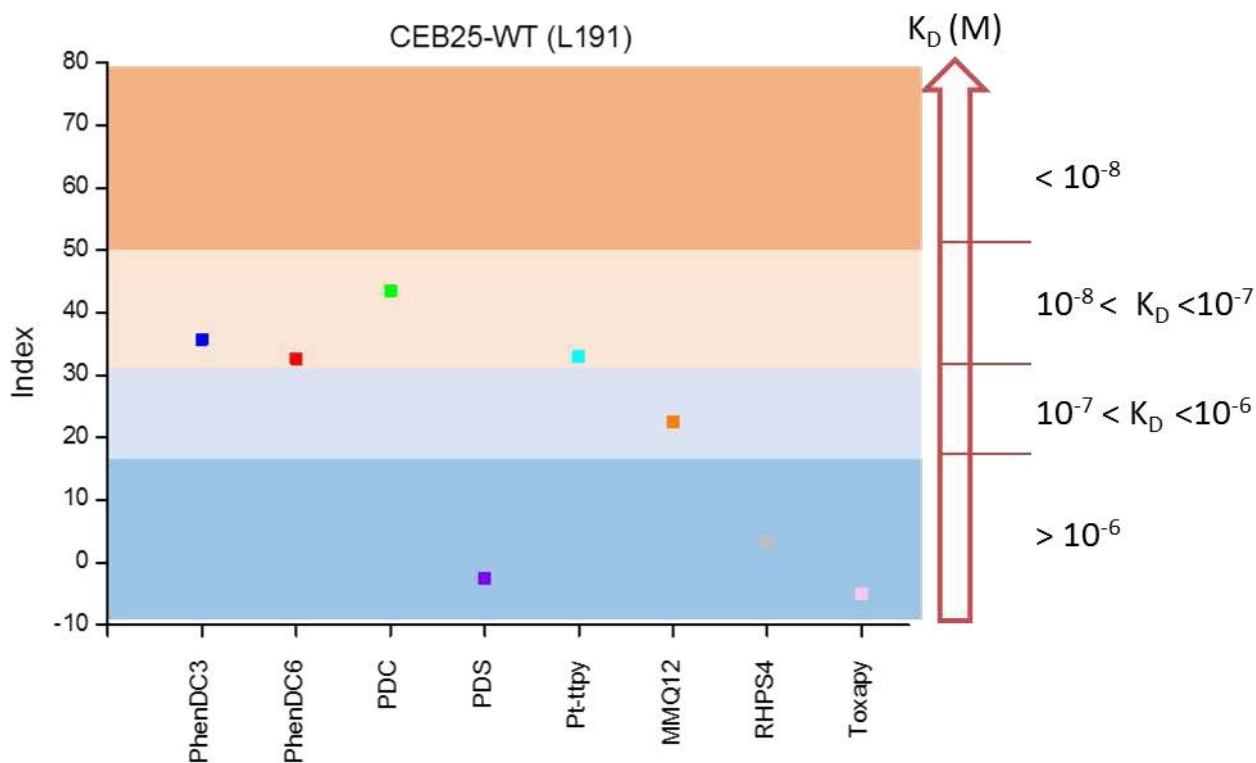


Figure 130 : Représentation de l'index issu des deux tests biophysiques en fonction des ligands de G4 sur la séquence native du minisatellite CEB25.

Sur la séquence native du minisatellite CEB25 (figure 130) comprenant une grande boucle, la dynamique de la structure est modifiée. C'est également sur cet oligonucléotide qu'il a été montré que sa formation en G-quadruplex *in vivo* est moins probable.^[32c] On constate d'ailleurs, avec cette représentation des index, que tous les ligands sont affectés par cette structure à longue boucle. En effet, aucun ligand ne possède d'index supérieur à 50 ce qui signifie qu'aucun ligand, parmi ceux évalués, n'aurait de constante d'affinité de l'ordre du nanomolaire comme peut l'avoir le ligand PhenDC₃ sur d'autres séquences G-quadruplex. Cependant, quatre ligands (PDC, PhenDC₃, PhenDC₆ et Pt-ttpty) entrent dans la deuxième catégorie c'est-à-dire une constante de dissociation de l'ordre de la dizaine à la centaine de nanomolaire. Cela signifie que ces composés conservent une très bonne affinité *in vitro* mais cela peut également vouloir dire que la dynamique de la boucle est tellement importante qu'aucun ligand n'est capable de suffisamment stabiliser ce G4. Ce résultat est cohérent avec les données *in vivo* qui ne montrent pas de phénotype G-quadruplex pour cet oligonucléotide dans l'essai fait au sein des levures.^[32c] *In vitro*, le G4 peut se former car toutes les conditions sont réunies pour qu'il se forme (milieu riche en potassium, conditions de formation favorable à l'ADN quadruplex...) et une structure RMN a même été établie.^[123]

Pour conclure, nous avons donc établi quatre différentes catégories que nous fixons en fonction de l'indice obtenu par les méthodes biophysiques utilisées. Ces catégories ne sont pas dépendantes de

l'oligonucléotide ni du ligand ainsi, en fonction des résultats nous pouvons classer les ligands dans ces différentes catégories. Ces résultats sont basés à la fois sur les données des expériences de FID et de FRET-melting obtenus. La corrélation entre ces données et celles de SPR nous permet de relier des données semi-quantitatives à des valeurs de constantes de dissociation. Par cette valeur d'index calculée, nous avons pu constater que les ligands étudiés sont tous plus ou moins affectés par la présence d'une longue boucle dans la structure d'un G-quadruplex. Nous avons également pu constater que certains ligands dépendent grandement de l'oligonucléotide et changent de catégorie selon l'oligonucléotide étudié. Cependant, hormis le composé Toxapy, tous ces ligands semblent afficher une constante de dissociation comprise entre le micromolaire et le nanomolaire. Par ailleurs, cette représentation en index ne donne pas d'information par rapport à la sélectivité des ligands. La même représentation en prenant comme valeur de température de fusion celle de l'oligonucléotide en présence du ligand et de l'ADN compétiteur (ds26 à 10 μ M) ne montre pas de différence suffisamment importante pour exprimer la sélectivité.

2) Dérivés du PhenDC₃ et du PDC

Nous avons également utilisé la représentation des index pour les ligands dérivés des composés PhenDC₃ et PDC.

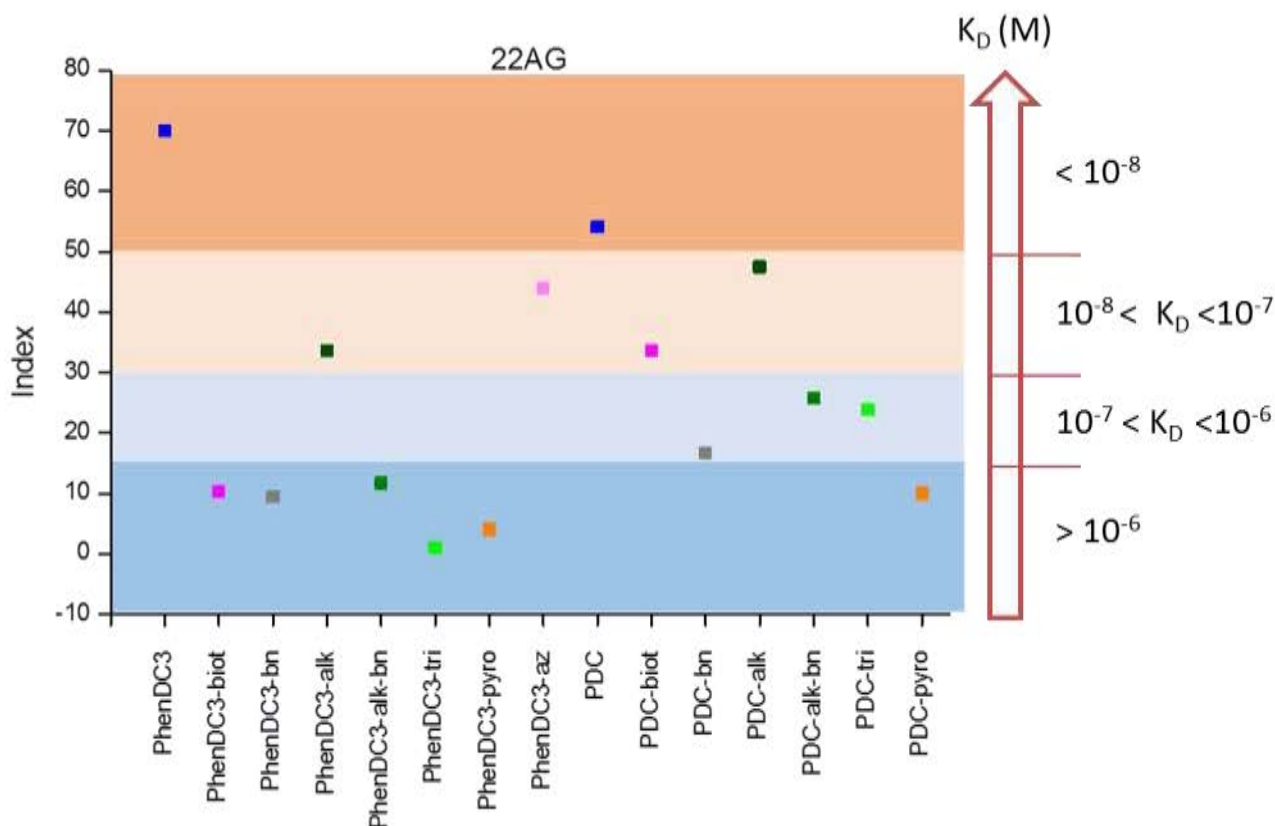


Figure 131 : Représentation de l'index issu des deux tests biophysiques en fonction des ligands de G4 sur la séquence télomérique humaine 22AG.

En prenant en compte la même classification que celle établie pour les ligands de référence, on constate donc que les composés PhenDC₃ et PDC appartiennent à la première catégorie sur la séquence 22AG (figure 131). Les composés alcynes et azoture du PhenDC₃ et du PDC et le dérivé biotine du PDC appartiennent à la deuxième catégorie avec un index compris entre 30 et 50. Tous les autres dérivés du PhenDC₃ ont un index inférieur à 15. A contrario, les dérivés du PDC, excepté PDC-pyro qui est dans la dernière catégorie, sont tous dans la troisième catégorie avec un index compris entre 15 et 30 ce qui correspondrait à une constante de dissociation de l'ordre du micromolaire.

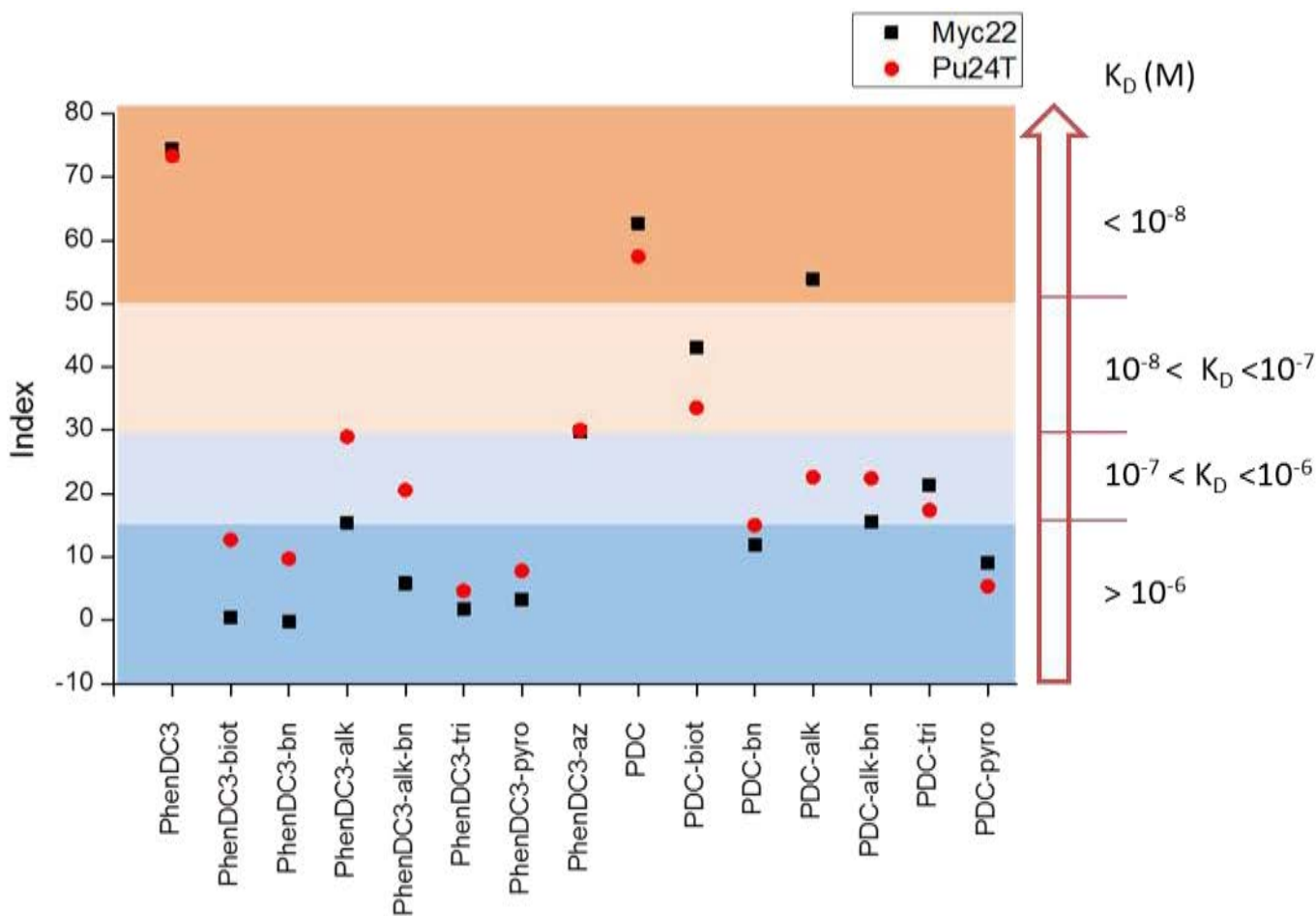


Figure 132 : Représentation de l'index issu des deux tests biophysiques en fonction des ligands de G4 sur les séquences modifiées de l'oncogène c-myc.

Il est intéressant de constater que, pour les dérivés du PhenDC₃, l'index est supérieur sur la séquence Pu24T par rapport à Myc22 alors que, pour les dérivés du PDC, l'index est globalement plus élevé sur la séquence Myc22 par rapport à Pu24T (figure 132). Cela montre une plus grande affinité de la famille PhenDC pour la séquence Pu24T ce qui est cohérent avec la structure RMN qui a été établie.^[70] Le classement des ligands est globalement le même que pour la séquence 22AG hormis les composés alcynes qui ont des comportements différents en fonction de la séquence étudiée. Le dérivé alcyne du PhenDC₃ appartient à la deuxième catégorie avec un index d'environ 30 sur la séquence Pu24T et à la troisième catégorie pour la séquence Myc22 avec un index de 17. Le composé

alcyne-benzophénone est également dans deux catégories différentes suivant de la séquence, il appartient à la troisième catégorie pour la séquence Pu24T et dans la dernière classe pour la séquence Myc22. Enfin, le dernier composé à être dans deux catégories différentes est le ligand PDC-alk qui est dans la première catégorie avec un index supérieur à 50 pour le Myc22 et étonnamment, dans la troisième catégorie sur la séquence Pu24T avec un index de 20. Ces différences de catégorie pour un même ligand montrent la très grande influence de la structure mais également que l'affinité des ligands dépend aussi grandement de la structure G-quadruplexe étudiée. En effet, pour des séquences très proches, les structures sont très différentes et impliquent des différences d'affinité avec les ligands (figure 133 et 134).

Acronyme	Séquence	PDB
Pu27	TGGGGAGGGTGGGGAGGGTGGGGAAGG	
Myc 22	<u>T</u> GAGGGTGGG <u>T</u> AGGGTGGG <u>T</u> AA	1XAV
Pu24T	TGAGGGTGG <u>T</u> GAGGGTGGGGAAGG	2MGN

Figure 133 : Différence entre les séquences Pu24T et Myc 22 par rapport à la séquence native Pu27.

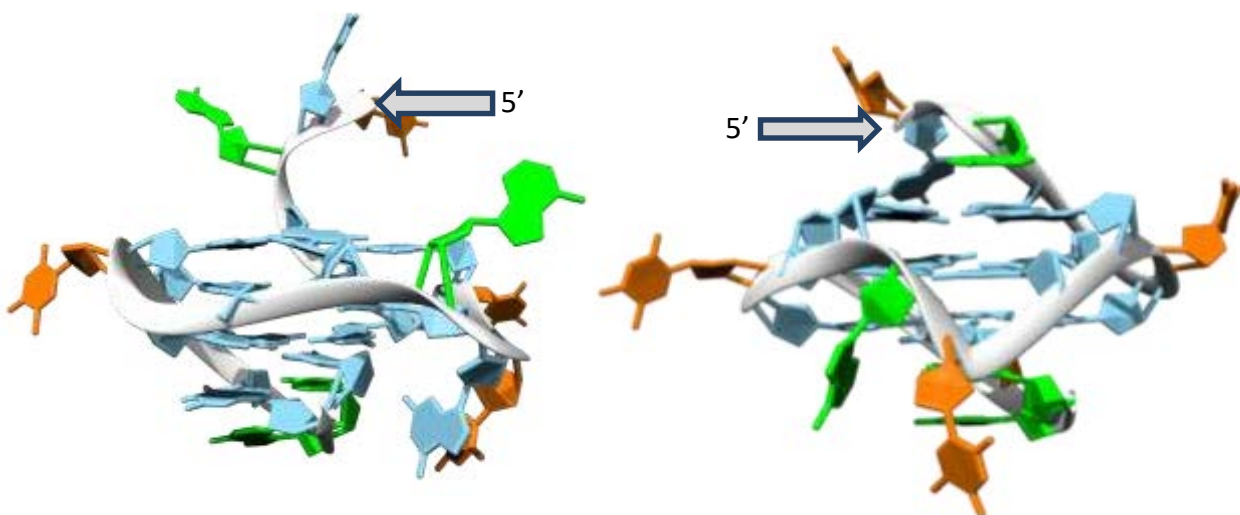


Figure 134 : Structure RMN de la séquence Pu24T (PDB : 2MGN, à gauche)^[70] et de la séquence Myc22 ^[174] (PDB : 1XAV, à droite) de l'oncogène c-myc.

On peut donc s'interroger sur l'intérêt d'utiliser des séquences modifiées pour les tests *in vitro* qui ne reflètent pas la séquence native existante *in vivo*. Nous savons cependant que ces séquences modifiées sont très importantes pour essayer de comprendre leur comportement *in vivo*. En effet, les modifications apportées permettent de travailler avec des structures G-quadruplex non-polymorphes ce qui facilite l'étude de l'interaction avec le ligand.

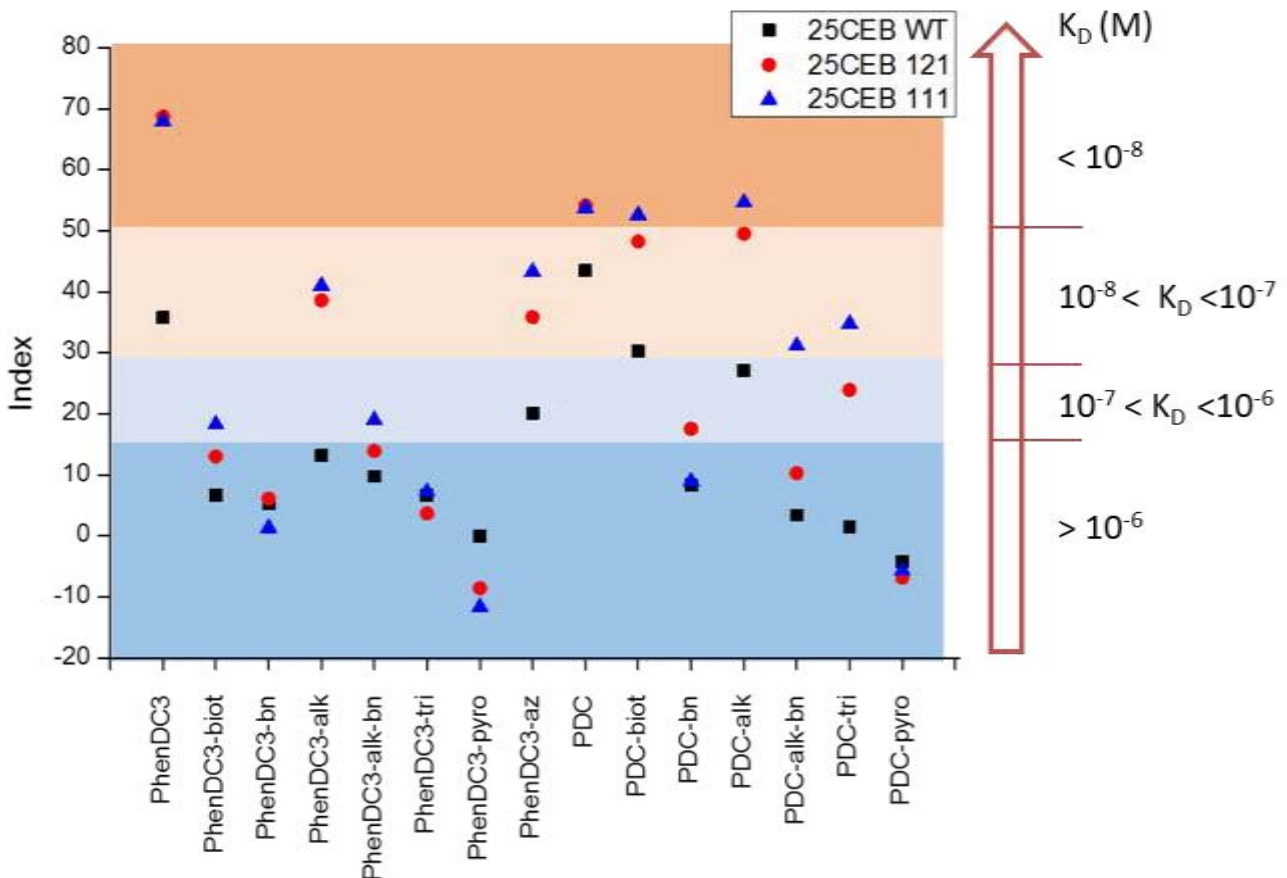


Figure 135 : Représentation de l'index issu des deux tests biophysiques en fonction des ligands de G4 sur les séquences native et modifiées du minisatellite CEB25.

Les séquences modifiées du minisatellite CEB25 nous permettent d'analyser l'effet de la taille de la boucle d'une structure G-quadruplex. En s'appuyant sur les résultats des index sur ces différentes séquences (figure 135), on peut constater que plus la boucle est longue plus l'index est bas ce qui signifie que les ligands sont moins affins pour les séquences à longues boucles. Tous les ligands, qui conservent une bonne affinité pour les G-quadruplex malgré leur fonctionnalisation (ligands de référence et composés azoture et alcyne principalement), remontent au classement des ligands lorsque la boucle du G-quadruplex se réduit. Les composés PDC et PhenDC₃ passent de la deuxième catégorie à la première lorsque la boucle diminue. Les composés PDC-alk, PDC-biot passent de la troisième catégorie à la première ce qui montre que la présence de chaînes latérales a une grande influence lorsque la boucle du G4 est longue. Les composés PhenDC₃-alk et PhenDC₃-az passent de la troisième catégorie à la deuxième. Cela signifie que ces ligands auraient des constantes de dissociation 10 à 100 fois plus grandes lorsque la taille de la boucle augmente si on se réfère à la classification des ligands établie « arbitrairement ».

Pour conclure, le point le plus important relaté ici est que la fonctionnalisation du composé PhenDC₃ affecte très fortement l'affinité du ligand pour sa cible. Cependant, lorsqu'on ajoute uniquement une

chaîne carboné (ou PEG) avec une fonction alcyne ou azoture, la perte d'affinité est moins grande même si elle est toujours présente. Pour les dérivés PDC, seuls les dérivés très encombrés (tri-fonctionnel, alcyne-benzophénone) ou fonctionnalisé avec un espaceur court (benzophénone) sont affectés par la fonctionnalisation. Les composés biotine et alcyne ne sont que très peu affectés et ne perdent quasiment pas en affinité. La seule différence structurale entre PDC et PhenDC₃ est la pyridine qui est remplacée par la phénanthroline. La phénanthroline aide à avoir une surface aromatique plus large qui va permettre un excellent recouvrement du quartet. Cependant, une fois fonctionnalisée par une chaîne aliphatique, la reconnaissance n'est plus aussi idéale car l'affinité est diminuée.

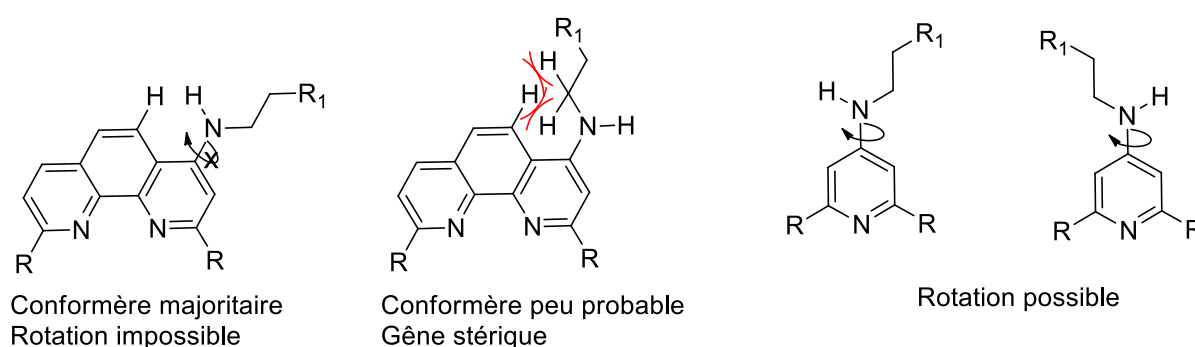


Figure 136 : Gêne stérique induite par la fonctionnalisation de la phénanthroline.

Sur la phénanthroline, seul un conformère n'induit pas de gêne stérique (figure 136). Cela implique que le mouvement de la chaîne est limité, il y a une rigidification de la structure. Ainsi, lors de l'interaction avec le G4, PhenDC₃ est plus rigide que PDC. Il n'y a pas d'adaptation possible à l'environnement sur le composé PhenDC₃ ce qui induit une perte d'affinité. De plus, la chaîne va également avoir tendance à se déplacer dans les plans perpendiculaires à celui de la phénanthroline ce qui peut induire une gêne importante pour l'empilement du ligand sur le quartet du G-quadruplex. Ce problème n'est pas constaté en présence des dérivés du PDC car la chaîne sur la pyridine n'a aucun problème de gêne stérique et est libre de mouvement dans le plan de la pyridine. Cette flexibilité permet à la molécule de s'adapter à l'environnement et influence donc moins l'affinité.

Les fonctions peuvent également avoir une influence sur la reconnaissance de la cible. En effet, la benzophénone peut, à la fois, interagir par liaisons hydrogène avec les boucles des G-quadruplex mais également s'empiler sur le quartet externe. La biotine, comme cela a été constaté par l'équipe de Unrau,^[177] peut aussi s'empiler sur le quartet et interagir avec les bases des boucles. Les chaînes peuvent interagir avec les quartets et les bases des boucles ou encore créer des clashes stériques avec la dynamique des boucles des structures G-quadruplex. Toutes ces interactions peuvent nuire à la bonne reconnaissance et au bon positionnement du motif PDC ou PhenDC sur la cible.

IV. Corrélation entre l'expérience de FRET-melting et l'instabilité génétique en levures

Ce paragraphe fait suite à une collaboration effectuée avec le groupe d'Alain Nicolas qui travaille sur le site de Paris de l'Institut Curie. L'objectif est de savoir s'il est possible de faire une corrélation entre un test *in vitro* (l'expérience de FRET-melting) et un test *in vivo* (instabilité génétique au sein de levures) pour l'étude de différents ligands. Cela permettrait de trouver les molécules idéales pour passer *in vivo*. Pour cela, nous avons étudié, d'une part, le minisatellite CEB25 et les différentes séquences modifiées issues de ce minisatellite avec des tailles de boucles différentes et des compositions de boucles différentes. D'autre part, le minisatellite CEB1 a également été étudié. Les deux ligands étudiés sont PhenDC₃ et PDC-pyro qui sont les deux composés donnant des résultats intéressants en levures.

Les études d'instabilité génétique ont été effectuées par Frédéric Samazan puis Xiaojie Cui au sein de l'équipe d'Alain Nicolas. Le principe du test est d'introduire des séquences répétées (minisatellites humains, oncogène c-myc...) pouvant former des cluster de G-quadruplex au sein des levures *S. Cerevisiae* qui ont un génome ne possédant que peu de répétition de guanines et repérer s'il y a une instabilité génétique générée par la stabilisation de ces séquences (figure 137).^[32c]

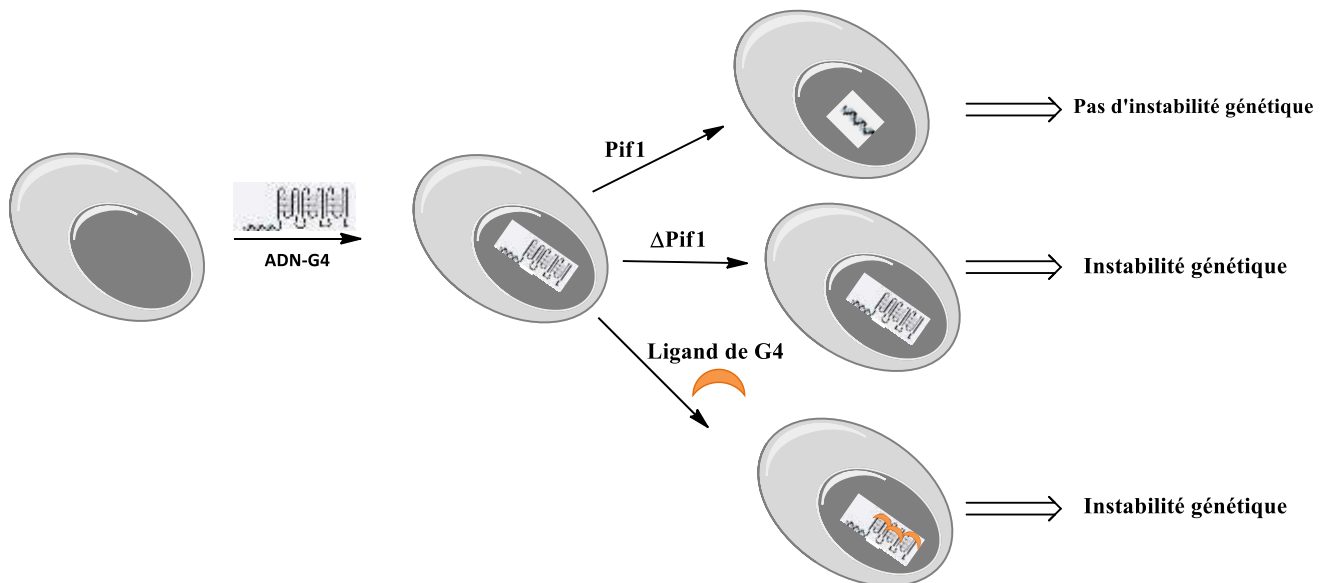


Figure 137 : Schéma de l'expérience de mesure d'instabilité génétique au sein des levures.

Une corrélation entre la température de fusion de l'oligonucléotide seul et l'instabilité génétique a été établie mais peut-on établir une corrélation entre la température de fusion d'un oligonucléotide stabilisé par un ligand et l'instabilité génétique générée par ce ligand *in vivo* ?

1) Minisatellite CEB25

Pour essayer d'établir cette corrélation, nous sommes partis des résultats existants sur le minisatellite CEB25 publiés en 2015 (figure 138).^[32c]

Allèle	Séquence du motif	Instabilité génétique levures non traitées (%)	Instabilité génétique levures + PhenDC ₃ (%)	G4 T ^{UV} (°C)
CEB25-WT (L191)	AAGGGTGGGTGTAAGTGTGGGTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	0,0	55,1
CEB25-L171	AAGGGTGGGTAAGTGTGGGTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	0,0	61
CEB25-L151	AAGGGTGGGAGTGTGGGTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,5	0,0	59,7
CEB25 -L141(TTTT)	AAGGGTGGGTTTTGGGTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	2,1	63,9
CEB25-L131(TGT)	AAGGGTGGGTGTGGGTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,5	0,0	61,9
CEB25 L311(TTT)	AAGGGTTTGGGTGGGTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	3,0	1,6	65,8
CEB25-L131(TTT)	AAGGGTGGGTTTTGGGTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	2,1	63,3
CEB25-L113(TTT)	AAGGGTGGGTGGGTTTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	1,0	63,1
CEB25-L211(TT)	AAGGGTTGGGTGGGTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,3	11,5	68,6
CEB25-L121(TT)	AAGGGTGGGTTGGGTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	16,5	67,9
CEB25-L112(TT)	AAGGGTGGGTGGGTTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	3,4	68,4
CEB25-L121(AA)	AAGGGTGGGAAGGGTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	7,9	65,8
CEB25-L221(TT)	AAGGGTTGGGTTGGGTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	9,4	61,1
CEB25-L212(TT)	AAGGGTTGGGTGGGTTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	3,9	62,1
CEB25-L122(TT)	AAGGGTGGGTTGGGTTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	1,0	6,3	61,5
CEB25-L222(TT)	AAGGGTTGGGTTGGGTTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	1,4	54,9
CEB25-L222(AA)	AAGGGAAGGGAAGGGAAGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	0,0	<40
CEB25-L111(T)	AAGGGTGGGTGGGTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	21,9	73,4
CEB25-L111(T)-G30A	AAGGGTGGGTGGGTGGGT GTGAGTGTGAGTGTGGAGGTAGATGT	0,0	22,9	73,4
CEB25 L111(T)-G12T	AAGGGTGGGTGTGTGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	0,0	<40
CEB25-L111(C)	AAGGGCGGGCGGGCGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	1,0	21,9	74,7
CEB25-L111(A)	AAGGGAGGGAGGGAGGGT GTGAGTGTGGGTGTGGAGGTAGATGT	0,0	1,0	56,5

Figure 138 : Résultats d'instabilité génétique au sein des levures des différentes séquences CEB25 et leur température de fusion par UV-melting (en rouge, les séquences qui vont être étudiées par FRET-melting).

Nous avons sélectionné dix séquences différentes à évaluer par FRET-melting. Parmi elles, certaines génèrent de l'instabilité *in vivo*. Pour le test de FRET-melting, nous limitons la taille de la séquence au G-quadruplex sans la séquence flanquante qui sépare les répétitions de G-quadruplex dans le test d'instabilité génétique (en rouge dans la figure 138). La première chose à regarder est la différence de température de fusion de l'oligonucléotide seul entre le test d'UV-melting et de FRET-melting (figure 139).

Allèles	G4 T ^{UV} (°C)	G4 T ^{FRET} (°C)
CEB25-WT (L191)	55,1	55,6
CEB25-L171	61	60,3
CEB25-L151	59,7	58
CEB25 -L141(TTTT)	63,9	60,6
CEB25-L131(TGT)	61,9	60,7
CEB25 L311(TTT)	65,8	61,1
CEB25-L131(TTT)	63,3	59,8
CEB25-L121(TT)	67,9	65,8
CEB25-L121(AA)	65,8	60,1
CEB25-L111(T)	73,4	72,4

Figure 139 : Comparaison de température de fusion mesurée par UV-melting ou FRET-melting.

Nous pouvons constater que les températures de fusion sont du même ordre de grandeur par les deux méthodes utilisées (figure 139). Les oligonucléotides marqués par des fluorophores pour le FRET-melting sont, globalement, légèrement déstabilisés. Il est donc cohérent d'utiliser ces séquences plus courtes que celles utilisées pour le test *in vivo* pour conserver un effet FRET plus important en conformation G-quadruplex.

Nous pouvons alors évaluer l'effet de ligands sur les oligonucléotides. Les deux ligands sélectionnés sont PhenDC₃ et PDC-pyro (figure 140) qui ont montré des instabilités génétiques au sein des levures sur certaines séquences.

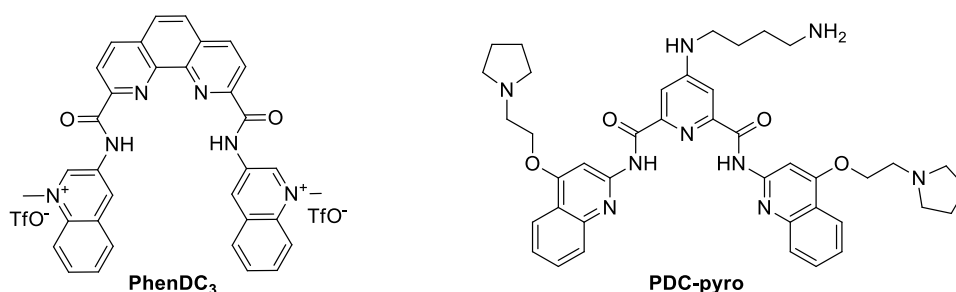


Figure 140 : Structure des composés PhenDC3 et PDC-pyro.

Allèles	FRET G4 Tm (°C)	FRET G4 + 5eq PhenDC ₃ Tm (°C)	FRET G4 + 1 eq PhenDC ₃ Tm (°C)	FRET G4 + 5 eq PDC-pyro Tm (°C)
CEB25-WT (L191)	55,6	79,8	65,2	58,6
CEB25-L171	60,3	78,4	67,9	67,9
CEB25-L151	58	77,8	67,8	66,5
CEB25 -L141(TTTT)	60,6	78,0	70,4	68,9
CEB25-L131(TGT)	60,7	79,6	71,2	69
CEB25 L311(TTT)	61,1	79,4	68,4	67
CEB25-L131(TTT)	59,8	81,1	73,8	67,5
CEB25-L121(TT)	65,8	86,7	78,3	70
CEB25-L121(AA)	60,1	79,3	70,9	65,9
CEB25-L111(T)	72,4	87,7	80,1	83,7

Figure 141 : Résultats de FRET-melting sur les différents allèles du CEB25.

Lorsque l'on utilise les conditions classiques de FRET-melting, à savoir 5 équivalents de ligands par rapport à l'oligonucléotide, il est impossible de discriminer une séquence par rapport à une autre avec le composé PhenDC₃. Peu importe la séquence, il la stabilise (figure 141). Afin de trouver des différences d'affinité entre le ligand PhenDC₃ et les oligonucléotides G-quadruplex, l'expérience de FRET-melting a été réalisée en présence d'un seul équivalent de ligand par rapport à l'ADN (figure 141). On constate que les valeurs de température de fusion du complexe ligand-ADN sont plus basses. Il y a maintenant des différences entre les températures de fusion mesurées. Par ailleurs, PDC-pyro, qui est un ligand moins affiné que PhenDC₃, est toujours testé à 5 équivalents. On constate d'ailleurs que le composé PhenDC₃ à un équivalent et le composé PDC-pyro à cinq équivalents ont des valeurs de température de fusion de la même grandeur ce qui montre encore une fois la très grande affinité de PhenDC₃. Maintenant que les résultats de FRET-melting sont obtenus, est-il possible d'établir une corrélation avec les études d'instabilité génétique faites en levures ?

Allèle	Instabilité levures non traitées (%)	Instabilité PhenDC ₃ (%)	Instabilité PDC-pyro (%)	FRET G4 T _m (°C)	FRET G4 + 1 eq PhenDC ₃ T _m (°C)	FRET G4 + 5 eq PDC-pyro T _m (°C)
CEB25-WT (L191)	0	0	nc	55,6	65,2	58,6
CEB25-L171	0	0	nc	60,3	67,9	67,9
CEB25-L151	0,5	0	nc	58,0	67,8	66,5
CEB25 -L141(TTTT)	0	2,1	nc	60,6	70,4	68,9
CEB25 -L131(TGT)	0,5	0	nc	60,7	71,2	69,0
CEB25 L311(TTT)	3	1,6	nc	61,1	68,4	67,0
CEB25-L131(TTT)	0	2,1	nc	59,8	73,8	67,5
CEB25-L121(TT)	0	16,5	6,8	65,8	78,3	70,0
CEB25-L121(AA)	0	7,9	nc	60,1	70,9	65,9
CEB25-L111(T)	0	21,9	8,3	72,4	80,1	83,7

Figure 142 : Résultats des expériences *in vivo* et *in vitro* sur les allèles du CEB25. nc = non caractérisé.

Les résultats d'instabilité génétique sont ceux de l'article EMBO Journal de 2015 cité précédemment,^[32c] excepté les valeurs en présence du PDC-pyro qui proviennent d'études complémentaires (figure 142). Ce composé a été choisi suite à un criblage de molécules : il génère une instabilité génétique en présence d'allèles du CEB25 et pas en présence du CEB1. Les données de FRET-melting ont été obtenues dans les conditions classiques de l'expérience en présence du tampon K⁺1. On constate que plus la boucle du G-quadruplex est grande moins cela génère d'instabilité génétique ce qui signifie que le G-quadruplex n'est pas stabilisé au sein des levures ni par la déplétion en hélicase ni par le ligand. Soit l'oligonucléotide ne forme pas de G-quadruplex *in vivo*, soit le G-quadruplex est tellement dynamique que le ligand n'a pas le temps de se fixer sur celui-ci. La boucle doit certainement avoir un effet très important dans la dynamique de formation et de déstructuration si le quadruplex se forme. Cela est cohérent avec les valeurs de température de fusion des oligonucléotides seuls, plus la boucle est grande, plus la température de fusion est basse. En présence d'un équivalent de composé PhenDC₃, on constate que toutes les séquences qui génèrent de l'instabilité génétique (>2%) ont une température de fusion supérieure à 70 °C. On peut donc estimer que pour des valeurs de température de fusion supérieures à 70 °C en présence d'un ligand, ce même ligand génèrera une instabilité génétique *in vivo* (figure 143). Cette corrélation est confirmée par les deux valeurs d'instabilité génétique avec le composé PDC-pyro. En effet, sur ces deux séquences les valeurs de température de fusion en présence du ligand sont supérieures à 70 °C.

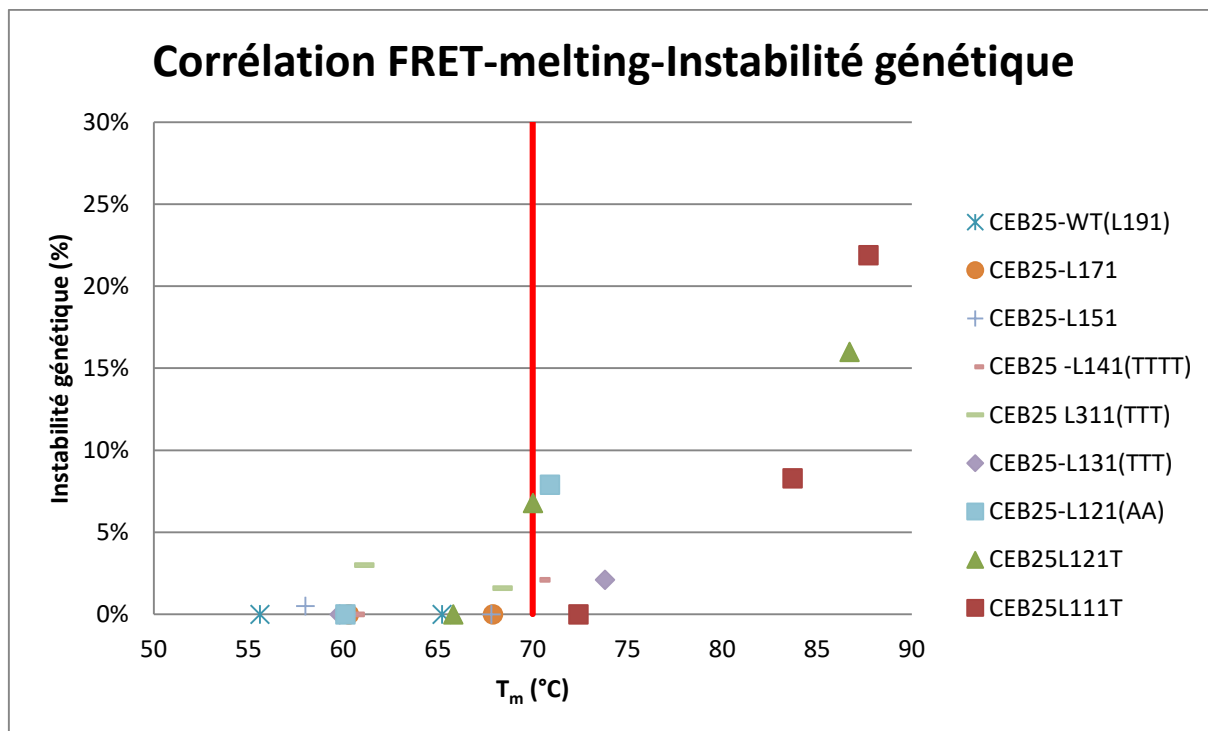


Figure 143 : Graphique représentant l'instabilité génétique en fonction de la température de fusion des oligonucléotides avec ou sans ligand (PhenDC₃ et PDC-pyro).

2) Minisatellite CEB1

Suite à ces résultats, nous avons voulu confirmer que la corrélation trouvée pouvait s'appliquer pour une autre séquence formant un G-quadruplex. Différentes séquences modifiées du CEB1 ont donc été utilisées tout d'abord pour le test *in vivo* (figure 144).

Allèles	Séquences	Instabilité levures seules (%)	Instabilité PhenDC3 (%)
G3T	TGTGCTGAGGGGGGAGGGAGGGTGGCCTGCGGAGGTCCC	1,0	8,9
G(9,10)T	TGGGCTGATTGGGGAGGGAGGGTGGCCTGCGGAGGTCCC	0,0	8,3
G(9-11)T	TGGGCTGATTTGGGAGGGAGGGTGGCCTGCGGAGGTCCC	0,0	11,0
G11T	TGGGCTGAGGTGGGAGGGAGGGTGGCCTGCGGAGGTCCC	0,5	7,8
G12T	TGGGCTGAGGGTGGAGGGAGGGTGGCCTGCGGAGGTCCC	0,0	19,7
G(12,25)T	TGGGCTGAGGGTGGAGGGAGGGTGTCTGCGGAGGTCCC	0,0	17,7
G(24,25)T	TGGGCTGAGGGGGGAGGGAGGGTTTCCTGCGGAGGTCCC	0,0	3,6
G3T, G(24, 25)T	TGTGCTGAGGGGGGAGGGAGGGTTTCCTGCGGAGGTCCC	0,0	3,7
A19T	TGGGCTGAGGGGGGAGGGTGGGTGGCCTGCGGAGGTCCC	0,0	29,7
A19TT	TGGGCTGAGGGGGGAGGGTTGGGTGGCCTGCGGAGGTCCC	1,0	4,6

Figure 144 : Résultats d'instabilité génétique au sein des levures des différentes séquences issues du minisatellite CEB1.

Les différentes séquences modifiées du CEB1 introduites seules au sein des levures n'induisent pas d'instabilité génétique. En présence du composé PhenDC₃, nous pouvons constater de l'instabilité génétique induite pour toutes les séquences étudiées ce qui est cohérent avec la formation de G-quadruplex au sein des levures qui sont stabilisés par le ligand. Il est maintenant intéressant de mesurer les températures de fusion par l'expérience de FRET-melting des différents oligonucléotides avec et sans ligand (figure 145).

Allèles	T _m (°C)	PhenDC ₃ (1 eq)	PDC-pyro (5 eq)	Instabilité PhenDC ₃ (%)
G3T	62,7	65,1	68,1	8,9
G(9,10)T	61,4	74,6	73,2	8,3
G(9-11)T	61,8	70,7	73,9	11,0
G11T	61,5	70,8	68,3	7,8
G12T	59,3	65,6	67,0	19,7
G(12,25)T	59,4	67,3	65,3	17,7
G(24,25)T	62,9	71,5	71,9	3,6
G3T, G(24, 25)T	61,6	63,0	65,1	3,7
A19T	63,6	74,3	71,9	29,7
A19TT	62,8	68,3	65,8	4,6

Figure 145 : Résultats de l'expérience de FRET-melting sur les allèles du minisatellite CEB1.

Au vu des modifications effectuées au sein des séquences G12T et G(12,25)T, celles-ci ne devraient pas induire une si grande instabilité génétique en présence du composé PhenDC₃ (>15%) car d'après les études RMN faites par l'équipe de Anh Tuan Phan, quatre conformations différentes de G-quadruplex peuvent exister et toutes nécessitent la présence de la guanine en position 12 (figure 146). On peut constater par l'expérience de FRET-melting, qu'il s'agit des deux séquences les moins stables (T_m le plus bas) du panel et qu'aucun des deux ligands utilisés ne les stabilise grandement ($\Delta T_m < 10$ °C).

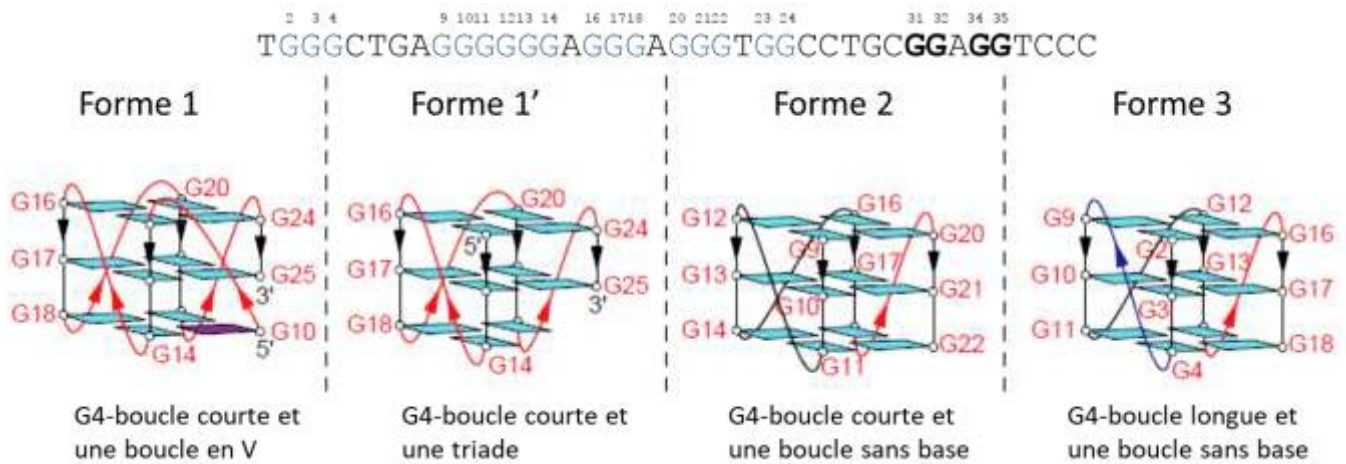


Figure 146 : Structures définies par RMN pour le minisatellite CEB1.

L'interprétation de manière globale est plus compliquée à réaliser sur ce minisatellite. En effet, la corrélation établie précédemment semble s'appliquer mais de nombreux cas particuliers sont présents. Cela peut s'expliquer par le grand polymorphisme du minisatellite CEB1 (figure 146). De manière générale, les séquences possédant une instabilité génétique inférieure à 5% ont une valeur de température de fusion, en présence de ligand, inférieure à 70 °C. Les séquences possédant une instabilité génétique supérieure à 5% ont une valeur de température de fusion supérieure à 70 °C en présence du ligand PhenDC₃. Les résultats sont exactement les mêmes en présence du ligand PDC-pyro excepté pour la séquence **G11T**, il serait donc intéressant de vérifier *in vivo* si PDC-pyro induit de l'instabilité génétique avec cette séquence. Aucune expérience n'a été effectuée avec le ligand PDC-pyro sur ces séquences. Suite à ces résultats de FRET-melting, l'expérience *in vivo* prend de l'importance pour confirmer ou non notre corrélation. Il serait bien de tester trois séquences différentes, la séquence G11T comme cité précédemment, une séquence qui n'induirait pas d'instabilité génétique (par exemple **G3T,G(24,25)T**) et une séquence qui induirait de l'instabilité génétique (par exemple **G9-11T**).

3) Conclusion

Cette étude de comparaison entre une expérience *in vivo* et une expérience *in vitro* nous a permis d'établir une corrélation qui semble être cohérente. En effet, lorsqu'un ligand stabilise une séquence d'ADN pouvant former un G-quadruplex de manière à avoir une température de fusion supérieure à 70 °C alors ce même ligand générera très probablement de l'instabilité génétique *in vivo*.

Avant d'arriver à cette corrélation, il a été nécessaire de savoir comment représenter les valeurs les unes par rapport aux autres pour faire ressortir les conclusions. Pour l'expérience de FRET-melting, il est plus courant de représenter les valeurs de ΔT_m que les valeurs de température de fusion. Cependant, dans ce cas, cela n'a pas de cohérence car la valeur de ΔT_m va dépendre de la valeur de

température de fusion de l'oligonucléotide seul et donc la stabilisation maximale va pouvoir varier entre 15 °C pour un oligonucléotide très stable (par exemple le CEB25111 qui a une température de fusion supérieure à 70 °C) et 35 °C pour un oligonucléotide moins stable (par exemple le CEB25-WT (L191) qui a une température de fusion d'environ 55 °C). La manipulation se faisant dans un tampon aqueux, il n'est pas possible de mesurer des températures de fusion supérieures à 90 °C. On s'affranchit donc de l'oligonucléotide en prenant uniquement la valeur de température de fusion. Il est bien évidemment important d'utiliser les mêmes conditions pour effectuer le test de FRET-melting lorsqu'on utilise des séquences modifiées issues d'un acide nucléique particulier (même tampon, même concentration).

Chapitre 3 : Visualisation de ligands de G-quadruplex en cellules de mammifère

Une fois que l'affinité et la sélectivité des ligands pour les G-quadruplex ont été démontrées *in vitro* par les expériences de biophysique, il est alors intéressant de connaître le comportement de ces molécules *in vivo*. Afin d'apporter plus de précisions, nous considérons dans ce manuscrit que les expériences *in vitro* concernent toutes les expériences de biophysique, les réactions de chimie click effectuées sans présence de cellule ni d'environnement biologique et les expériences de temps de vie de fluorescence. Les expériences *in vivo* concernent pour leur part toutes les expériences effectuées au sein de cellules cancéreuses humaines (fixées ou vivantes).

Dans ce chapitre, un état de l'art des études *in vivo* va être développé suivie d'une brève présentation de la chimie bioorthogonale. Ensuite, les résultats obtenus *in cellulo* vont être présentés. Pour finir, des études *in vitro* de temps de vie de fluorescence vont également être discutées.

I. Les G-quadruplex observés *in vivo*

Les G-quadruplex ont beaucoup été étudiés *in vitro* cependant leur existence *in vivo* demeurait une inconnue importante à résoudre. Ce n'est que ces dix dernières années que les premières études ont été effectuées pour essayer de démontrer leur existence en milieu cellulaire. Pour cela différentes stratégies ont été envisagées : l'imagerie cellulaire qui comprend l'immunofluorescence et la localisation de ligands en milieu cellulaire, l'étude d'instabilité génétique, ou encore plus récemment le séquençage.

1) Imagerie cellulaire

a) Immunofluorescence

L'immunofluorescence est une méthode de marquage de biomolécules faisant appel à des anticorps couplés à des fluorophores. Cette méthode peut être utilisée selon deux techniques différentes : l'immunofluorescence directe qui utilise un seul anticorps couplé à un fluorophore capable de cibler spécifiquement la biomolécule à observer et l'immunofluorescence indirecte qui utilise au moins deux anticorps, un premier capable de cibler la biomolécule et un second couplé à un fluorophore qui reconnaît spécifiquement le premier anticorps. Cette technique permet donc de détecter et cibler spécifiquement une biomolécule. Elle peut donc être très utile pour le marquage de structure telle que les G-quadruplex. La détection de ces structures *in vivo* est un défi majeur encore aujourd'hui. Quelques anticorps ont été développés, ils se lient spécifiquement au quadruplex *in vitro* et permettent un marquage par immunofluorescence en cellules.^[181]

L'utilisation d'anticorps pour la reconnaissance de G-quadruplex n'est pas nouvelle. En effet, l'équipe de Schaffitzel a été la première à le réaliser à partir d'un anticorps généré *in vitro*, le Sty49 qui est

spécifique du G-quadruplex télomérique du cilié *Stylonychia lemnae*.^[182] Par la suite, en utilisant le même anticorps, les équipes de Rhodes et Lipps ont montré qu'au niveau des extrémités des télomères de ce cilié, deux protéines prennent part à la gestion de la formation des G-quadruplex *in vivo*.^[183] Plus récemment, un nouvel anticorps a été développé par l'équipe de Balasubramanian. Il s'agit du BG4 qui a une très forte affinité pour les structures G-quadruplex.^[184] L'utilisation de cet anticorps au sein de cellules HeLa a permis de le visualiser sur des chromosomes en métaphase ce qui devrait correspondre à des structures G-quadruplex si l'anticorps est uniquement spécifique des G-quadruplex *in cellulo* (figure 147). La colocalisation de l'anticorps BG4 avec TRF2, une des protéines du complexe shelterin, montrent l'implication de G-quadruplex aux télomères.

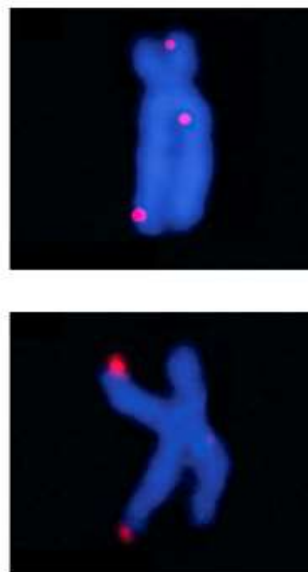


Figure 147 : Immunofluorescence de chromosomes en métaphase de cellules HeLa montrant la localisation de l'anticorps BG4 au niveau des télomères et du centromère.^[184]

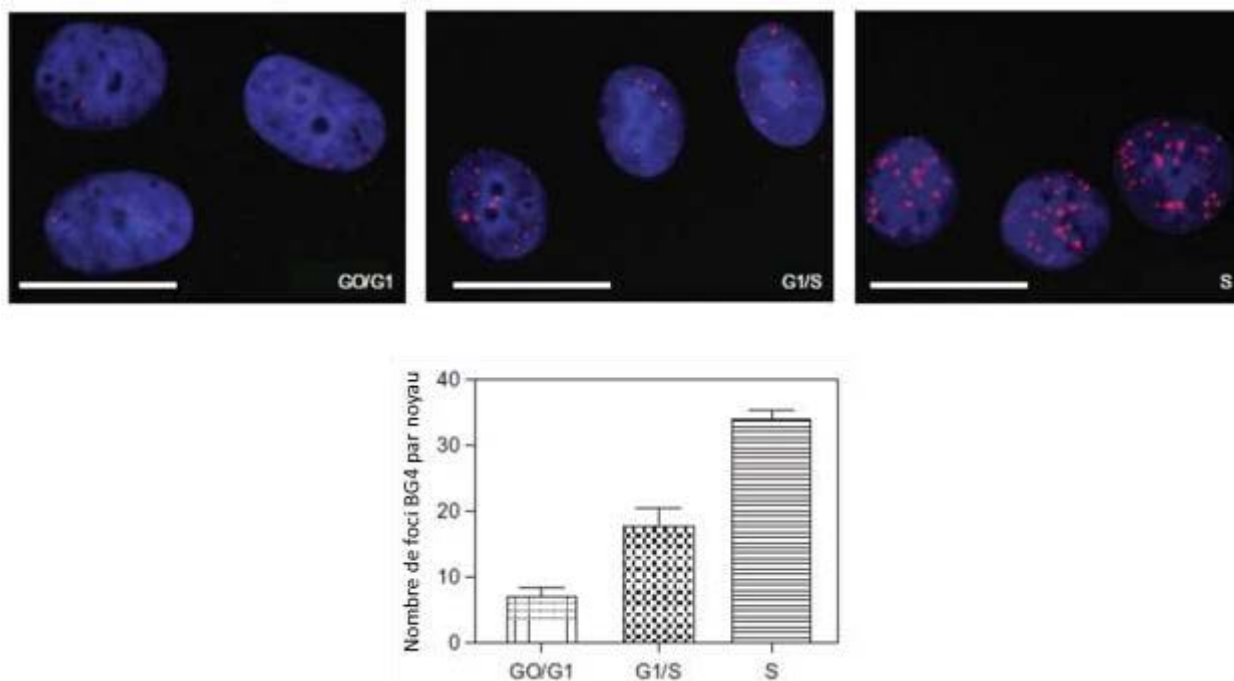


Figure 148 : Observation et quantification des foci présents dans les cultures de cellules HeLa, au cours des phases G0/G1, G1/S et S du cycle cellulaire. D'après [184].

Une des expériences réalisées dans cette étude montre l'augmentation du nombre de foci en fonction des différentes phases du cycle cellulaire. Pour cela, les cellules sont synchronisées, autrement dit, elles sont toutes placées dans la même phase du cycle cellulaire. Ensuite, le nombre de foci est quantifié en fonction de la phase dans laquelle se trouvent les cellules permettant ainsi de suivre l'évolution du nombre de structure G-quadruplex en fonction des phases du cycle cellulaire. Le maximum de foci est observé dans la phase S, la phase de réplication des cellules (Figure 148). Le nombre de foci est également très fortement augmenté lorsque les cellules sont traitées, au préalable, avec un ligand de G-quadruplex comme PDS ce qui montre l'effet G-quadruplex au sein des cellules.^[184] Le groupe de Lansdorp a également développé un anticorps 1H6 qui détecte les structures G-quadruplex intermoléculaires.^[185] Cependant, très récemment, ce même groupe a constaté que cet anticorps n'était pas spécifique uniquement des structures G-quadruplex et qu'il se liait à des zones de simple brin d'ADN riche en thymine.^[186] Un dernier anticorps a été découvert par le groupe de Huang, il permettrait de cibler sélectivement les quadruplex parallèles et en particuliers ceux de la séquence télomérique humaine.^[187]

Un des principaux inconvénients de cette technique d'immunofluorescence est le manque d'informations quant à la sélectivité des anticorps. En effet, il est difficile aujourd'hui d'affirmer que ceux-ci reconnaissent les structures G-quadruplex et uniquement ces structures au sein de milieux biologiques complexes. Des interactions avec des protéines ou d'autres structures secondaires d'acides nucléiques peuvent également exister. Par ailleurs, l'utilisation d'anticorps nécessite de

travailler avec des cellules fixées et perméabilisées pour leur permettre de pénétrer au sein des cellules. Cela signifie que nous ne sommes plus dans des conditions de cellules vivantes à proprement parler, tous les états de la cellule sont figés dans l'état du moment de la fixation. Cette fixation se fait très souvent par des moyens chimiques (utilisation de para-formaldéhyde) qui peuvent entraîner des modifications de localisation aussi bien des protéines à l'intérieur de la cellule que des éventuels molécules à observer.^[188]

b) Etudes de la localisation de ligands de G-quadruplex

De nombreux ligands ont été développés afin de stabiliser les structures G-quadruplex. Dans un premier temps, toutes les études ont été effectuées *in vitro* pour montrer leur efficacité. Ensuite, afin de découvrir si les structures G4 peuvent être ciblées et surtout localisées *in cellulo*, des petites molécules ont été testées en milieu cellulaire. Le premier ligand de G-quadruplex à avoir été incubé en milieu cellulaire afin de connaître sa localisation est la Quarfloxin qui est une molécule fluorescente.^[93a]

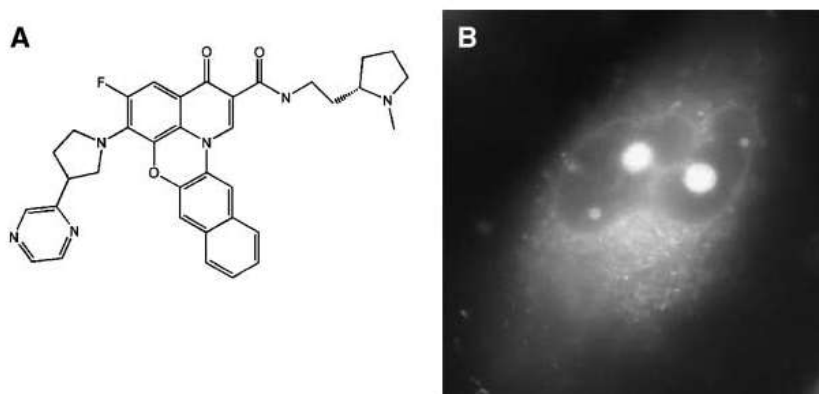


Figure 149 : A. Structure de la Quarfloxin et B. Visualisation de la Quarfloxin en cellules cancéreuses A549 (cancer du poumon) fixées. D'après ^[93a].

Le ligand a été incubé en cellules vivantes et visualisé dans les nucléoles de cellules cancéreuses après fixation de ces dernières (figure 149). Cette localisation dans les nucléoles où se déroule la biogénèse des ribosomes a été une découverte très intéressante car il s'agit d'un compartiment de la cellule multifonction qui contient à la fois de l'ADN, de l'ARN et des protéines. C'est une zone très dense située à l'intérieur du noyau et ne possédant pas de membranes physiques la séparant du nucléoplasme. Cette visualisation dans les nucléoles est cohérente avec la potentielle existence de nombreux G-quadruplex dans l'ADN ribosomal qui est particulièrement riche en guanine et cytosine mais également avec la haute activité de transcription des nucléoles impliquant des structures d'ADN et d'ARN non canoniques qui constitue une cible idéale pour les ligands de G-quadruplex. Il est cependant difficile de savoir ce qui conduit une petite molécule dans les nucléoles à cause de la large

diversité de molécules qui a été localisée dans ce sous-compartiment du noyau.^[189] Certains composés qui interagissent avec l'ADN comme les médicaments anticancéreux l'Actinomycine D et la Camptothécine ou les molécules fluorescentes intercalantes de l'ADN (thiazole orange, iodure de propidium) sont connus pour se distribuer en partie dans les nucléoles mais aussi dans le nucléoplasme. D'autres molécules sont distribuées exclusivement dans les nucléoles mais il n'y a pas de rationnel existant sur les structures capables de se localiser à cet endroit.^[190] Alors que les forces électrostatiques et l'affinité pour les acides nucléiques, en particulier l'ARN, semblent conduire les petites molécules dans les nucléoles, la spécificité et la pertinence de cette localisation en terme de reconnaissance de structure particulière d'acides nucléiques restent une interrogation importante. De plus, dans le cas de ligands de G-quadruplex, la signification de l'accumulation dans les nucléoles reste également un point important à résoudre. Pourtant, nombreuses sont les études qui depuis ont montré la présence de ligands de G-quadruplex au sein des nucléoles mais aucune d'entre elles n'expliquent clairement pourquoi cette localisation.^[68a, 191] La chimie « click » a également déjà été utilisée pour cibler les G-quadruplex par le groupe de Balasubramanian.^[68a, 192] Il a ainsi été montré une co-localisation de la molécule (PDS- α) avec l'hélicase Pif1 connue pour se lier et dérouler les G-quadruplex pendant la réplication^[193] et l'absence de co-localisation avec TRF1 qui montre que la molécule se lie à des sites de l'ADN non-téломériques. Il est également important de relever le marquage nucléolaire du composé PDS- α qui est très présent sur les images même si cela n'est quasiment pas évoqué dans la publication (figure 150).

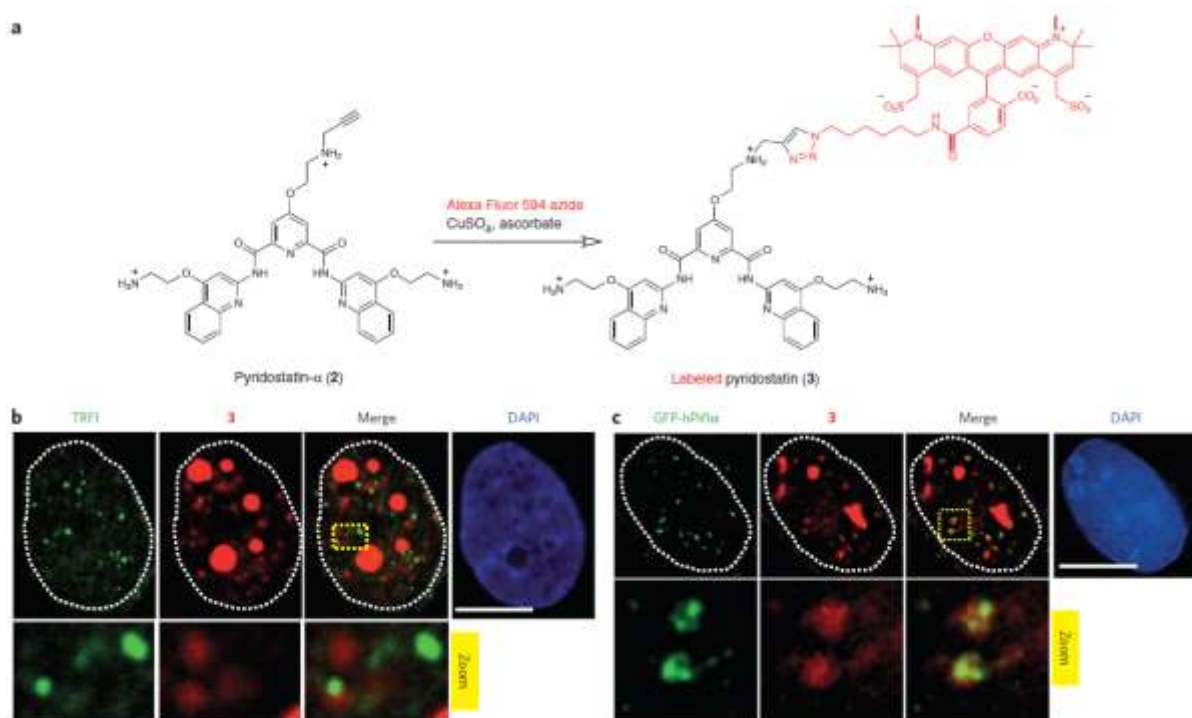


Figure 150 : a. Réaction de CuAAC de PDS- α avec le fluorophore AlexaFluor 594-azide en cellules fixées, b. images de microscopie confocale de la molécule marquée dans les cellules MRC-5-SV40 fixées pour la co-localisation avec TRF1 et c. les cellules U2OS fixées pour la co-localisation avec Pif1. D'après ^[68a].

Cependant, l'imagerie est effectuée en cellules fixées. Il n'y a donc plus de dynamique au sein de la cellule, tout est immobilisé. Il n'est pas certain que la fixation chimique des cellules ne joue pas un rôle dans la localisation des ligands ni que cette fixation ne perméabilise pas les cellules et permette ainsi une meilleure pénétration cellulaire. Par ailleurs, il a été montré dans l'équipe de Teulade-Fichou que certaines molécules ayant une affinité pour de l'ADN pouvaient avoir une localisation différente en cellules fixées et en cellules vivantes.^[194] En effet, l'étude de dérivés de la triphénylamine, ayant une affinité pour l'ADN duplex et interagissant dans le petit sillon de l'ADN, a montré qu'en cellules fixées les molécules TP-2Py et TP-2Bzim se localisent dans le noyau alors qu'en cellules vivantes elles se retrouvent dans les mitochondries. L'illumination continue de ces composés en cellules déclenche l'apoptose avec une relocalisation dans le noyau (figure 151).

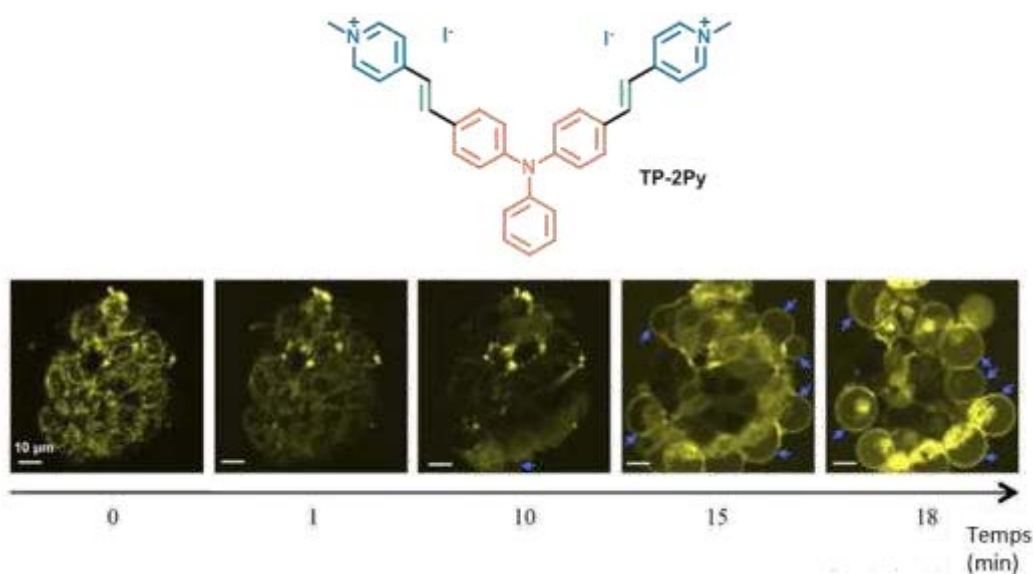


Figure 151 : Structure de la molécule TP-2Py et visualisation en continue de la molécule au sein de cellules MCF7.

Certaines études ont été menées à la fois en cellules fixées et en cellules vivantes. C'est le cas du groupe de Monchaud qui a développé un composé fluorescent uniquement lorsqu'il est fixé aux G-quadruplex, la molécule N-TASQ.^[95, 191b] L'étude en cellules vivantes (figure 152),^[95] qui a été publiée avant l'étude en cellules fixées (figure 153),^[191b] montre que la molécule se situe dans le cytoplasme ce qui suggère que ce soit un ligand de G-quadruplex d'ARN.

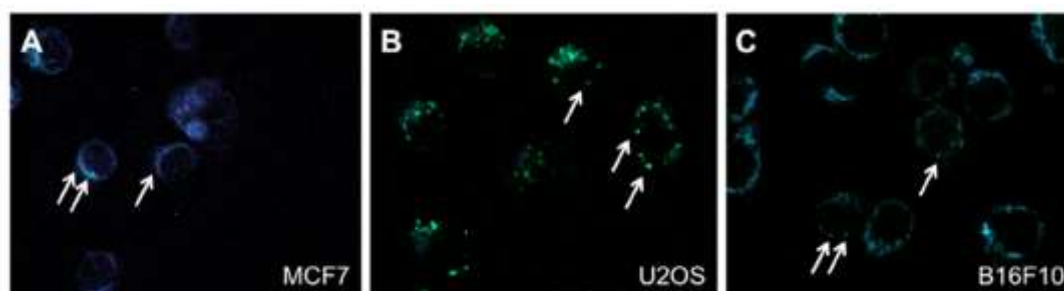
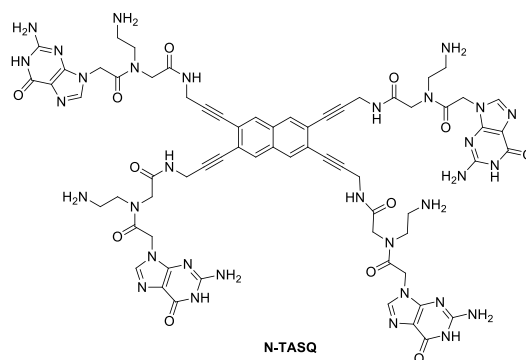


Figure 152 : Images de cellules vivantes incubées avec la molécule N-TASQ pendant 48h.^[95]

Cependant, par la suite, l'étude en cellules fixées montre la molécule à la fois dans le cytoplasme mais également dans le noyau (foci dans le nucléoplasme et marquage des nucléoles). Ainsi, la molécule est également un ligand d'ADN G-quadruplex mais seulement lorsque les cellules sont fixées que ce soit par du *para*-formaldéhyde ou du méthanol. En fonction de la méthode de fixation, la molécule et l'anticorps BG4 n'ont pas la même localisation. En effet, l'anticorps est dans le cytoplasme lorsque le *para*-formaldéhyde et le triton sont utilisés pour fixer et perméabiliser les cellules alors qu'il est dans le cytoplasme et les nucléoles lorsque le méthanol est utilisé pour fixer les cellules. Le ligand a lui aussi des différences de localisation. En incubant le ligand en cellules vivantes, il se trouve dans le cytoplasme. Lorsque l'incubation se fait après une fixation au méthanol, il y a de nombreux foci dans le nucléoplasme et surtout une forte intensité de fluorescence dans les nucléoles. Enfin, lorsque l'incubation se fait après une fixation au PFA suivie d'une perméabilisation au triton, le ligand est dans le cytoplasme et les nucléoles avec une intensité de fluorescence équivalente. Ces résultats montrent l'influence de la fixation sur la localisation des ligands. La nature de la fixation chimique peut aussi influencer comme le montre les expériences réalisées avec l'anticorps BG4 (figure 153, Alexa).

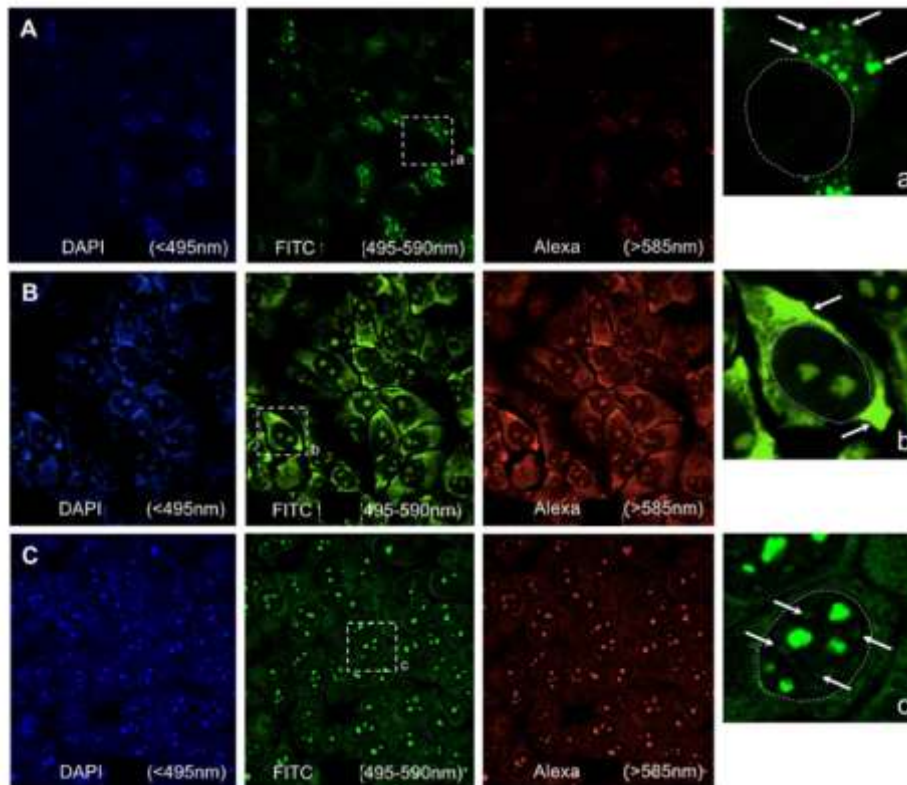
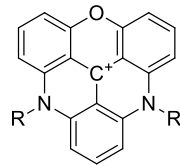


Figure 153 : A. Images de cellules MCF7 incubées en cellules vivantes pendant 24h avant fixation (PFA-triton), B. Cellules MCF7 fixées (PFA-Triton) puis incubées avec le N-TASQ et C. Cellules MCF7 fixées (MeOH) puis incubées avec N-TASQ.^[191b]

Une autre méthode a été développée par l'équipe de Vilar. Il s'agit d'une application au G-quadruplex de l'analyse de temps de vie de fluorescence en cellules vivantes. Il n'y a ainsi pas de problèmes dû à la fixation des cellules. En effet, la molécule DAOTA-M2 a été utilisée pour étudier la différence de temps de vie de fluorescence en présence de G-quadruplex, en présence d'ADN duplex ou d'ARN.^[97] Ainsi, la molécule est localisée dans le nucléoplasme, les mitochondries et les lysosomes (figure 154). Il est important de savoir que la molécule DAOTA-M2 n'est pas spécifique des G-quadruplex et interagit également avec l'ADN duplex. Dans cette étude, il a été montré que la molécule a un temps de vie de fluorescence différent en présence de G-quadruplex qu'en présence d'ARN ou d'ADN duplex mais la différence est très (ou trop) faible entre l'interaction avec un G4 et celle avec de l'ARN pour être complètement certain de l'interprétation.



DAOTA-M2

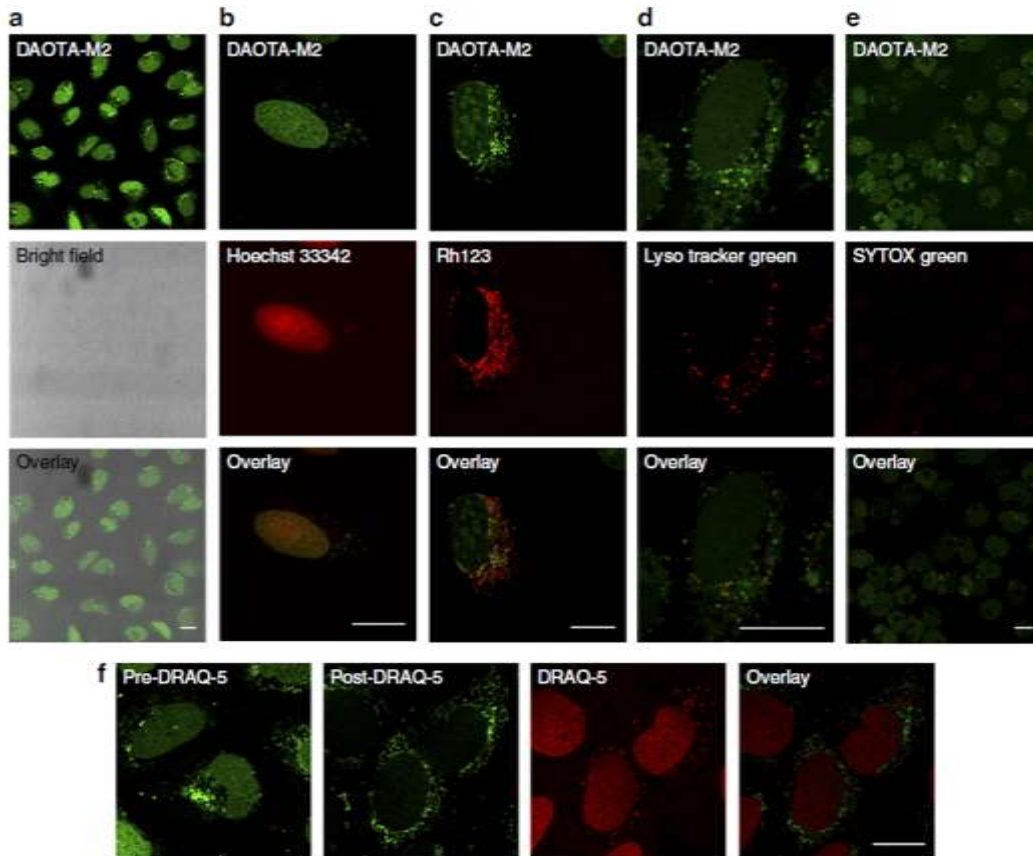


Figure 154 : Images de microscopie confocale des cellules U2OS incubées avec le DAOTA-M2. a) Marquage général observé en combinaison avec l'image en transmission en champ clair, b) à f) Co-localisation de la molécule avec b) Hoechst 33342, c) Rh123, d) Lyso Tracker vert DND-26, e) SYTOX vert et f) DRAQ-5.^[97]

2) Etude de l'instabilité génétique chez la levure

L'instabilité génétique résulte d'une modification de l'ADN. Elle est exprimée par une élévation ou une contraction du génome à la suite d'une cassure dans le double-brin. Cette instabilité génétique générée *in vivo* peut être une preuve de l'existence de G-quadruplex en milieu biologique. En effet, lorsque l'expression d'hélicases connues pour interagir avec les G-quadruplex est supprimée ou lorsqu'un ligand capable de stabiliser ces structures est introduit dans le milieu, une instabilité génétique est générée *in vivo*.^[195] Cela a été montré dans la levure *Saccharomyces cerevisiae* par l'équipe d'Alain Nicolas, en collaboration avec notre équipe, sur les minisatellites humains CEB1 et CEB25 capables de former des G-quadruplex.^[32a, 32c] Cette souche de levure possède un génome très pauvre en répétition de guanines et donc très peu de séquences pouvant former potentiellement des

G-quadruplex (PQS pour *Potential Quadruplex Sequence*). Il s'agit dès lors d'un modèle de choix pour l'introduction de séquences G-quadruplex humaines pour étudier l'effet de ces structures *in vivo*. Les minisatellites sont des séquences de répétition en tandem (VNTR pour *Variable Number Tandem Repeat*) contenus dans le génome humain et principalement dans les zones sub-téломériques. Ces minisatellites sont composés de 10 à 100 paires de bases. Pour les répétitions inférieures à 10 paires de bases, on parle de microsatellites. Ces séquences répétées sont impliquées dans la gestion de l'instabilité génétique d'une cellule et dans de nombreuses maladies telles que l'épilepsie et le diabète.^[196] De plus, les minisatellites jouent très probablement un rôle capital dans la structuration de l'ADN en chromatine.^[197]

Une très grande instabilité génétique est observée lorsque le minisatellite CEB1 est introduit au sein de levures ayant subies une délétion de l'hélicase Pif1, capable de dérouler les G4. Cela se traduit par de nombreuses contractions et extensions de la longueur du motif.^[198] Ce résultat montre la capacité du minisatellite à se former en G-quadruplex au sein des levures car aucune instabilité n'est observée lorsque des séquences contrôles G-mut (les séquences successives de G sont mutés par un C au milieu) sont introduites.^[199] L'effet *in vivo* du ligand PhenDC₃ est également démontré lors de cette étude ce qui montre la bonne pénétration des membranes cellulaires. Les levures sont traitées avec le composé qui va stabiliser les structures G-quadruplex et ainsi induire un effet identique à la délétion de Pif1, on parle alors d'effet de phénocopie.

Enfin, toujours dans l'équipe d'Alain Nicolas, une série de mutations du motif du minisatellite CEB25 a été effectuée afin de montrer l'effet de la taille et de la composition d'une boucle dans la formation *in vivo* de G-quadruplex. Les résultats obtenus permettent d'expliquer que les structures G4 ayant une boucle composée de moins de 4 nucléotides se forment dans les levures alors que les longues boucles (> 5 nucléotides) ne se forment pas et n'induisent donc pas d'instabilité génétique que ce soit en présence de ligand ou de l'hélicase Pif1.^[200] La structure du minisatellite n'est pas affectée par la diminution de la taille des boucles cependant la structure est beaucoup plus stable avec une valeur de température de fusion très élevée pour les boucles composées d'un ou deux nucléotides. Ainsi, la formation *in vivo* de ces structures en G-quadruplex est favorisée et par conséquent l'instabilité génétique est grande en présence de ligand ou de l'hélicase. *A contrario*, Les minisatellites à grande boucle, comme l'est la séquence native du CEB25 (9 bases), est beaucoup moins stable (température de fusion plus faible) et ne se forment pas *in vivo*. Ils n'induisent donc pas d'instabilité génétique qu'il y ait la présence d'un ligand de G4 ou une hélicase (Figure 155).

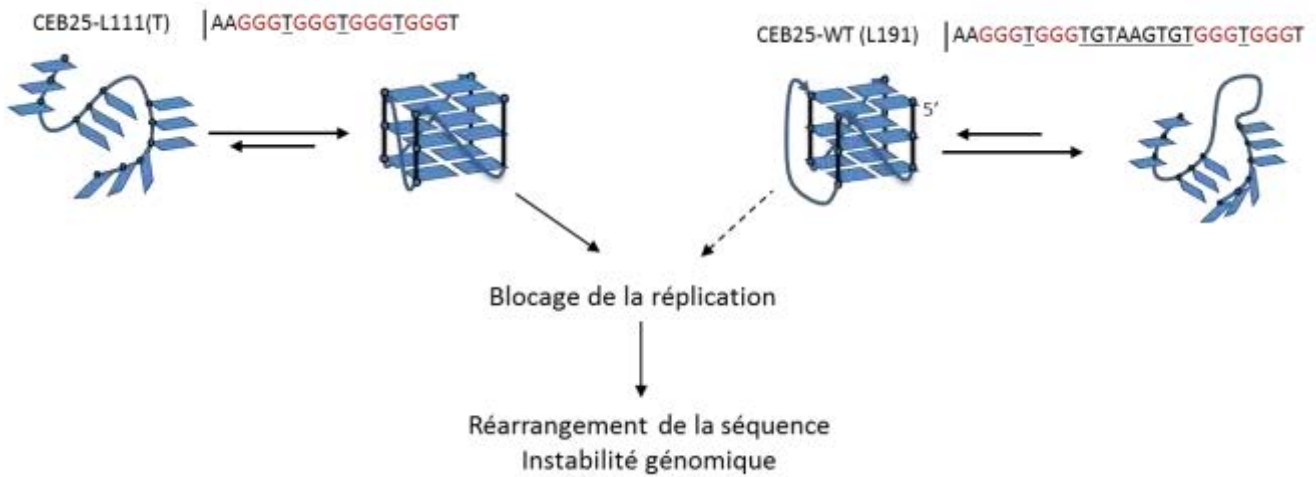
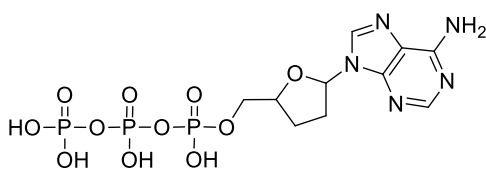
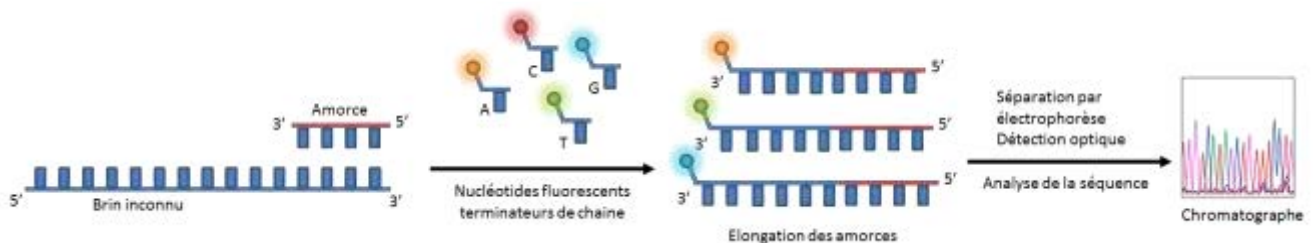


Figure 155 : Représentation schématique de l'arrangement des minisatellites en structures G-quadruplex à boucles courtes ou longues, et leur implication sur l'instabilité génomique.^[200]

3) Séquençage d'ADN

Le séquençage de l'ADN est apparu en 1977 par deux techniques différentes décrites cette année-là. Une technique *in vitro* a été développée par l'équipe de Sanger (Figure 156).^[201] Elle consiste à fonctionnaliser des didéoxynucléotides (dd-nucléotide) par un fluorophore différent pour chaque base. Ensuite, ces dd-nucléotides sont insérés à l'extrémité 3' des fragments d'ADN synthétisés par une enzyme (polymérase) lors de la réplication du brin inconnu.



L'introduction de didéoxynucléotides (dd-adénine représenté ci-contre), équipés avec un fluorophore différent pour chaque base, à l'extrémité de l'amorce à chaque étape de réplication permet, après séparation des fragments de différentes tailles, d'analyser la séquence du brin inconnu directement par détection optique.

Figure 156 : Méthode de séquençage de Sanger.

La seconde technique a été développée par l'équipe de Maxam et Gilbert.^[202] Elle consiste à effectuer des transformations chimiques qui vont engendrer des modifications des bases et ainsi créer des cassures du brin d'ADN à analyser lorsqu'il y a des modifications (figure 157). L'utilisation d'ADN radiomarqué au phosphore 32 (³²P) permet la détection des fragments d'ADN sur gel d'électrophorèse par lecture de la radioactivité. Ces deux techniques sont à l'origine du développement des premières techniques automatisées.

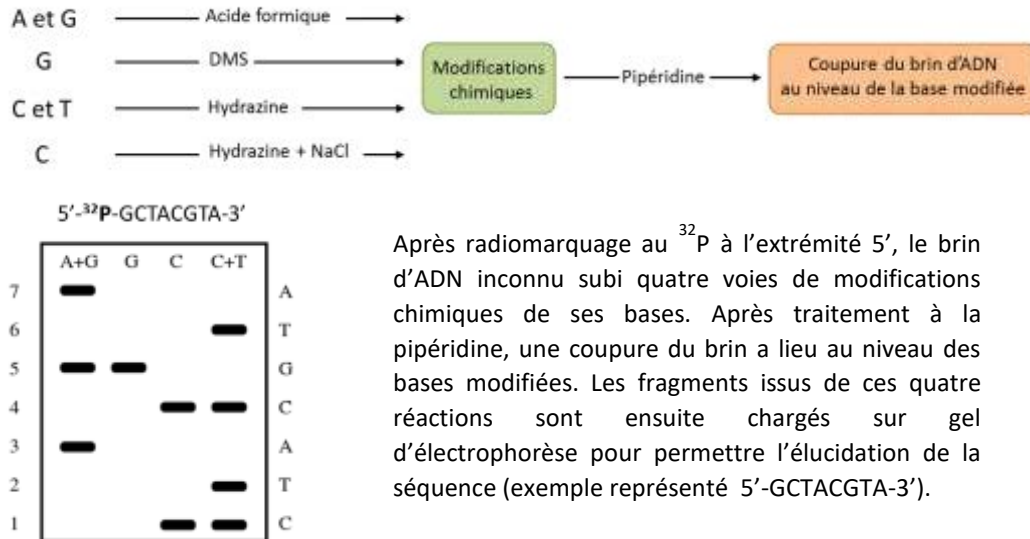


Figure 157 : Méthodes de séquençage de Maxam et Gilbert.

Depuis 2007, de nouvelles techniques de séquençage ont été développées et ont permis de grandes avancées dans ce domaine avec, par exemple, le premier génome humain séquencé en 2008. Ces méthodes, appelée de nouvelle génération (NGS pour *next generation sequencing*), ont permis de comprendre le mécanisme génétique de nombreuses pathologies. Elles sont basées sur le principe suivant : une fragmentation de l'ADN est effectuée puis les fragments sont amplifiés pour avoir suffisamment de matière pour l'étape de séquençage en elle-même. Le séquençage peut être réalisé par différentes techniques qui se résument en trois méthodes (figure 158) : le séquençage par synthèse, le séquençage base par base et le séquençage par ligation.^[203]

Tout d'abord, le séquençage par synthèse (SBS, *sequencing by synthesis*) est une technique basée sur la technique de Sanger. Le principe est d'incorporer un nucléotide fluorescent à l'extrémité du brin d'ADN polymérisé afin de détecter le signal de fluorescence et reconnaître les bases. Ensuite, le séquençage base par base (SMS, *single-molecule sequencing*) est une méthode basée sur l'ajout de nucléotides fluorescents au système. Cela permet donc de décoder l'ADN en lisant la fluorescence base par base. Enfin, le séquençage par ligation (SBL, *sequencing by ligation*) est une méthode qui utilise l'enzyme ADN ligase qui permet d'assembler des fragments d'ADN entre eux et permet ainsi de connaître la position de chaque base dans le brin d'ADN à analyser. Une fois l'expérience de séquençage terminée, il est nécessaire de traiter les données par des méthodes de bioinformatique afin de connaître les séquences analysées et potentiellement leur rôle dans le génome.

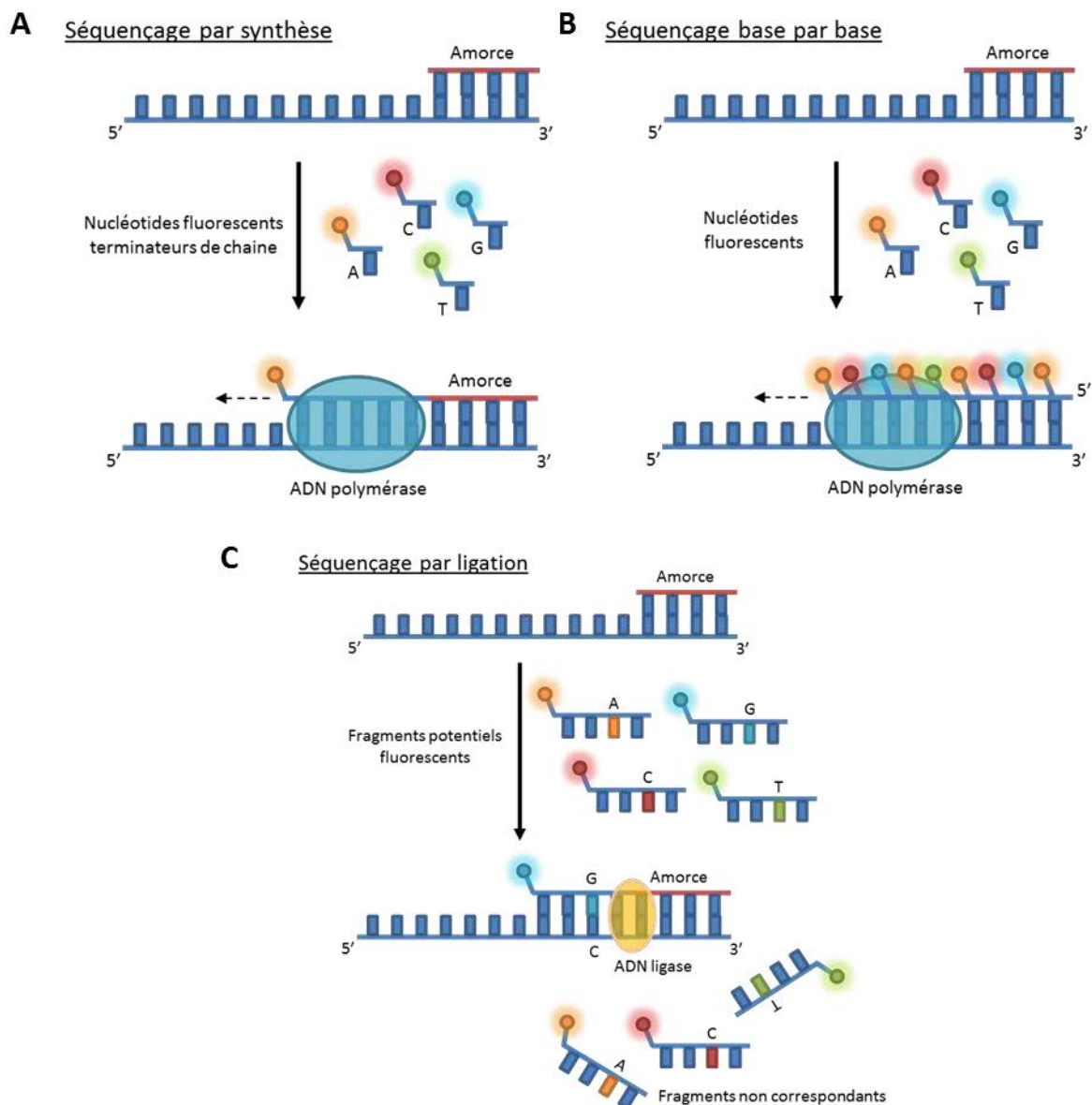


Figure 158 : Représentation schématique des trois principales techniques de séquençage de l'ADN.

L'équipe de Balasubramanian a ensuite développé une technique qui permet le séquençage des structures G-quadruplex au niveau du génome, le G4-seq.^[204] Le principe est basé sur la technique de *Polymerase stop assay* qui a été utilisée pour le séquençage de longs fragments d'ADN génomique comportant des structures G-quadruplex. Le fragment d'ADN à étudier est séquençé une première fois sans condition particulière (Figure 159 – Lecture 1). Les produits alors formés lors du séquençage par polymérisation sont analysés puis éliminés. Ensuite, le brin d'ADN est de nouveau séquençé dans des conditions favorables de formation de G-quadruplex (tampon riche en potassium ou ligand stabilisant ces structures comme PhenDC₃ ou PDS). La présence de G-quadruplex stoppe l'action de l'enzyme polymérase donc la polymérisation s'arrête au niveau des sites G-quadruplex ce qui permet de cartographier les G4 au sein d'un génome (Figure 159 – Lecture 2).

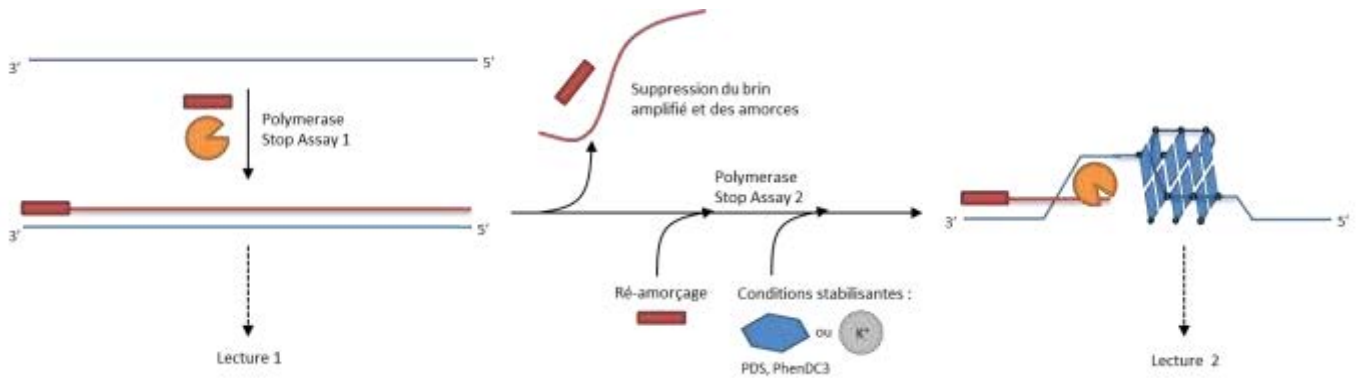


Figure 159 : Représentation schématique de l'expérience de séquençage (G4-seq) développée par l'équipe de Balasubramanian.^[204]

Etonnamment, parmi les 700 000 séquences isolées par la technique de G4-seq à partir de l'ADN génomique issu du lymphocyte B humain, environ 30 % seulement correspond aux séquences potentielles pouvant former des G-quadruplex (PQS) prédites par les algorithmes informatiques sous la forme $G_3+N_xG_3+N_xG_3+N_xG_3+$ ($1 < x < 7$).^[18] Les séquences G4 définies par des séquences de trois ou quatre guanines successives et des boucles comprenant entre un et sept nucléotides sont appelées G-quadruplex canoniques. Des études structurales plus récentes ont montré que des structures G-quadruplex non-canoniques peuvent également exister avec des boucles très longues (> 7 nucléotides)^[205] ou ayant des interruptions plus ou moins longues dans les enchainements de guanines.^[206] Ces séquences canoniques et non-canoniques d'acides nucléiques G-quadruplex sont présentes aussi bien dans l'ADN que l'ARN. Elles ont principalement été localisées au sein des promoteurs de gènes comme les précédentes études le suspectaient mais également dans les régions d'épissage, les exons, introns et dans les régions non-traduites (3'-UTR et 5'-UTR). Encore plus récemment, une étude utilisant l'immunoprécipitation de la chromatine à l'aide de l'anticorps BG4 pour établir une cartographie des structures G-quadruplex endogènes au sein de kératinocytes épidermiques normaux (NHEKs) et de kératinocytes précancéreux (HacaT) a été développée.^[207] Dans cette étude, environ 10000 G-quadruplex ont été détectés dans les cellules cancéreuses alors que seulement 1000 environ ont été détectés dans les cellules normales ce qui représente à peine 1 % de tous les G-quadruplex identifiés par G4-seq et prédit par les algorithmes de détection de G4. De plus, les séquences identifiées correspondent à celles qui sont hautement transcrites. Cela suggère que la formation de G-quadruplex est faible dans l'ADN génomique sous forme chromatine, probablement dû aux protéines qui contrôlent la formation de structures d'ADN et aux partenaires de la chromatine. De plus, avec la dynamique des structures G-quadruplex, il est difficile de savoir à quel moment et en quelle proportion elles se forment.

Ces structures G-quadruplex semblent être liées à la mécanique biologique, leur nombre au sein du génome et leurs séquences sont encore difficiles à prédire. Les méthodes d'analyses informatiques et

in vitro ont permis de mieux comprendre ces structures mais encore peu d'outils sont disponibles pour comprendre directement le rôle des G4 *in vivo*.

II. La chimie bioorthogonale

1) Principe

La chimie bioorthogonale a émergé dans les années 2000 comme stratégie pour l'étude des dynamiques biomoléculaires et de leurs fonctions dans les systèmes vivants.^[153, 154b, 154c, 208] Comparé à l'approche non-covalente, la chimie bioorthogonale se base sur l'attachement covalent et spécifique de la petite molécule à la biomolécule d'intérêt. Par conséquent, elle offre plusieurs avantages : tout d'abord, c'est une stratégie applicable à toutes les classes de biomolécules, des acides nucléiques aux protéines en passant par les lipides et les sucres. Ensuite, la structure de la petite molécule n'a aucune restriction si ce n'est de posséder une fonction bioorthogonale réactive *in situ* ce qui permet une très large diversité de structure. Enfin, cette stratégie permet d'avoir un système intelligent qui permet à la fois de cibler une biomolécule unique mais également de cibler une classe de biomolécules dans un génome. Les propriétés principales de la chimie bioorthogonale sont les suivants : tout d'abord, la spécificité, les fonctions bioorthogonales doivent réagir uniquement entre elles et être complètement inerte vis-à-vis des autres biomolécules présentes. Ensuite, l'irréversibilité, les molécules formées sont stables et la réaction est irréversible. Enfin, la biocompatibilité, la réaction a lieu en milieu aqueux dans les conditions physiologiques de pH et de température sans générer de sous-produit toxique pour l'organisme. Grace à ces caractéristiques uniques, la chimie bioorthogonale a été utilisée avec succès dans de nombreuses études fonctionnelles de protéines allant de la visualisation à l'expression de celles-ci.^[154c, 209] Cette stratégie peut également être utilisée pour créer un lien covalent entre deux petites molécules. En effet, cela permet d'ajouter un ligand capable de cibler une biomolécule sans le modifier au préalable puis de fonctionnaliser ce ligand par une sonde fluorescente, par exemple, *in situ*. Ainsi, cela permet de localiser le ligand au sein du milieu biologique et s'il a reconnu sa cible de localiser également la biomolécule. Il s'agit de la stratégie que nous avons utilisée pour localiser les ligands de G-quadruplex en cellules. Dans ce chapitre, un bref récapitulatif des deux réactions de chimie « click » utilisées va être rappelé puis l'état de l'art de la visualisation de G-quadruplex *in vivo* va être étudiée et enfin, les résultats obtenus *in vivo* de nos ligands vont être dévoilés.

2) Rappel des réactions de chimie « click » utilisées *in cellulo*

Les deux réactions de chimie « click » développées lors du chapitre 1 sont des réactions bioorthogonales ce qui signifie qu'elles peuvent avoir lieu en milieu biologique sans crainte de réagir avec d'autres substituants au sein de la cellule. La réaction de CuAAC a beaucoup été utilisée en cellules cependant elle possède un inconvénient majeur qui est la toxicité cellulaire du cuivre qui est capable de générer des espèces réactives de l'oxygène (ROS pour *reactive oxygen species*) toxiques pour la cellule par activation d'oxygène moléculaire.^[151] Ainsi, son utilisation n'est possible qu'en cellules fixées. En effet, alors qu'en solution la réaction s'effectue avec des quantités catalytiques de cuivre, en milieu cellulaire, la réaction est plus lente à cause de la complexité de l'environnement et de la faible concentration des partenaires de la réaction de chimie « click ». Par conséquent, un large excès de sels de cuivre est communément utilisé pour accélérer la réaction ce qui est le problème majeur de la réaction de CuAAC. Des améliorations importantes ont été développées par l'utilisation de différents ligands capable de chélater le cuivre et le stabiliser dans sa forme d'oxydation Cu (I).^[148a] Plus récemment, des ligands azotures ont été développés afin de chélater le cuivre (I).^[210] En effet, l'utilisation de ligands combinant dans la même structure la fonction réactive azoture et des groupes N-coordinants permet de grandement diminuer la quantité de cuivre nécessaire (figure 160).

Réaction classique CuAAC

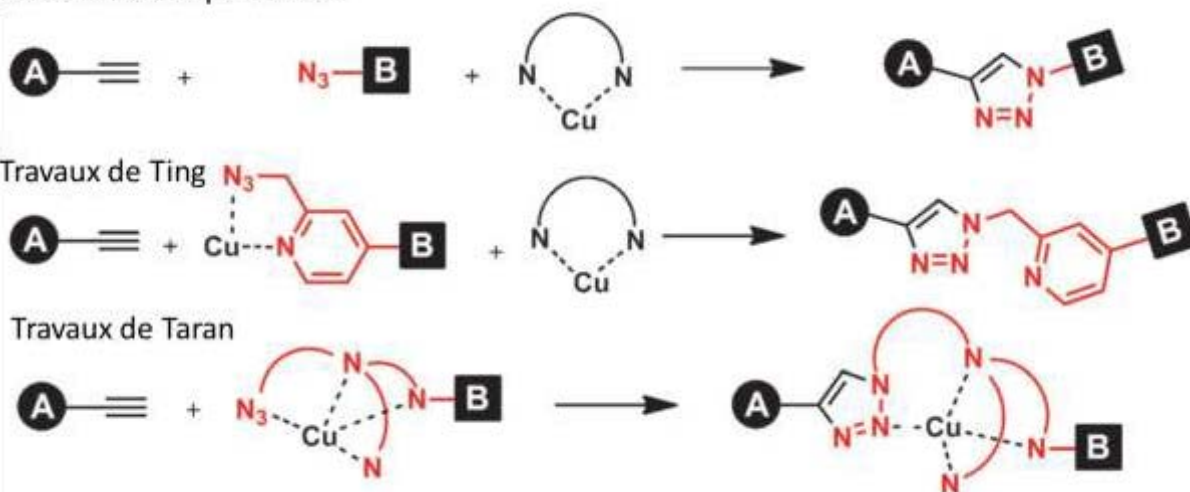


Figure 160 : Stratégies de développement pour avoir une réaction CuAAC rapide. D'après ^[210a].

Néanmoins, malgré le concept qui est très élégant, ces approches sont tout de même limitées par un travail de synthèse laborieux et la toxicité de certains chélatants.^[211] Ainsi, le protocole standard de la réaction de CuAAC en milieu cellulaire est toujours basé sur l'utilisation d'un large excès de sulfate de cuivre (1 mM) et son application se résume à l'imagerie de cellules fixées ou un marquage de la

surface des cellules vivantes.^[212] Pour cette raison, le développement de la cycloaddition sans cuivre en utilisant des alcynes contraints a représenté une avancée majeure pour l'application de la chimie « click » dans l'imagerie de cellules vivantes.^[154b, 213] Un des points importants encore à améliorer est la pénétration cellulaire des cyclooctynes. En effet, les cyclooctynes conjuguées à des fluorophores ont été développées et commercialisées pour, dans un premier temps, marquer les membranes cellulaires cependant, pour la visualisation de molécules dans des organelles à l'intérieur des cellules le problème de pénétration cellulaire est encore à résoudre.^[214]

III. La magie de la chimie « click » en cellules

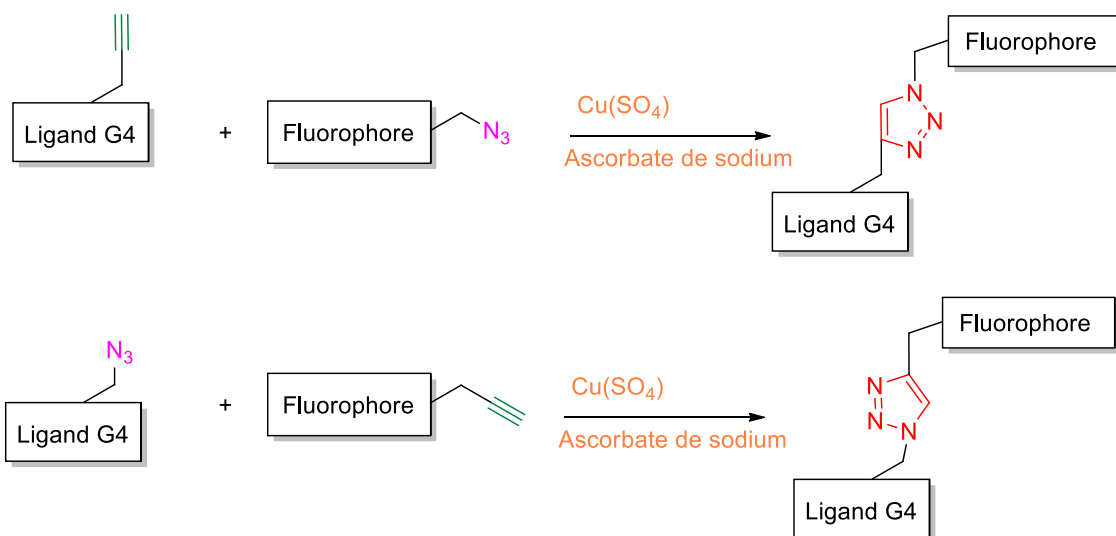
De nombreux éléments de cette partie ont fait l'objet d'une publication publiée cette année (2017) dans *Angewandte Chemie International Edition* et intitulée : *Copper-alkyne complexation responsible for the nucleolar localization of quadruplex nucleic acid drugs labeled by click reactions.*^[168]

Dans cette partie, les résultats obtenus par chimie « click » vont être développés. Tout d'abord, *in vitro* pour montrer que la réaction a bien lieu dans un tampon aqueux et à des faibles concentrations comme cela est le cas *in cellulo*. Ensuite, les résultats obtenus en cellules fixées seront développés et, pour finir, les études effectuées en cellules vivantes seront reportées.

1) Etude *in vitro*

Avant toute application au sein des cellules de mammifère, nous avons vérifié que les réactions de chimie « click » se déroulent correctement et dans des temps raisonnables pour l'application (figure 161). Les réactions de CuAAC et de SPAAC ont été réalisées dans des eppendorfs en utilisant des concentrations en composés azoture, alcyne ou DBCO supérieures (50 μM de ligand de G4 et 100 μM de fluorophores dans le tampon PBS) à ce qu'il se fait au sein des cellules (5 μM de ligand de G4 et 20 μM de fluorophore) afin de pouvoir suivre l'évolution par HPLC/MS alors que les quantités de sel de cuivre (1 mM) et d'ascorbate de sodium (100 mM) ne sont pas modifiées pour la réaction de CuAAC. Dans tous les cas, un excès de fluorophore est utilisé peu importe la fonction portée (alcyne, azoture ou DBCO) afin d'être certain que le fluorophore puisse pénétrer au sein des cellules et du noyau alors qu'il n'est incubé qu'une heure au maximum. Par ailleurs, les conditions de la réaction CuAAC en cellules nécessitent un large excès de sels de cuivre et d'ascorbate de sodium pour avoir la certitude que la réaction puisse se dérouler.

CuAAC



SPAAC

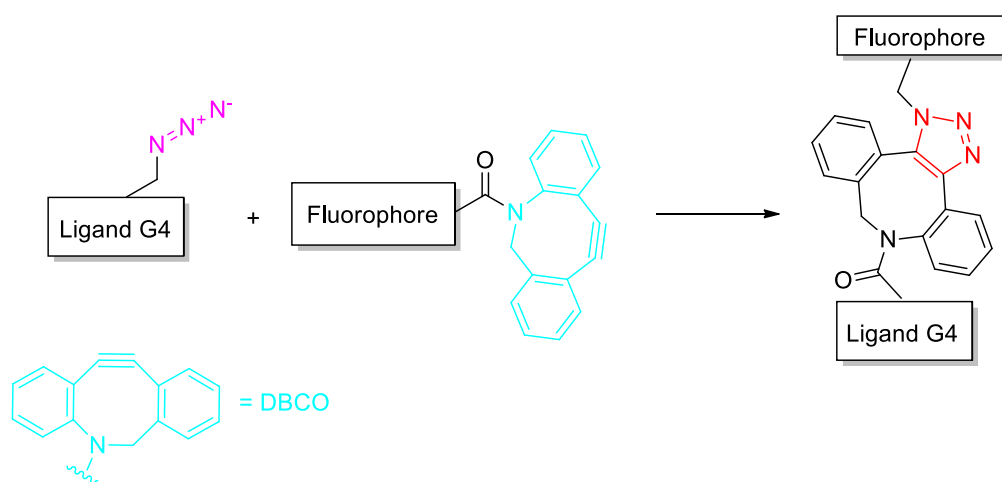


Figure 161 : Stratégie de chimie "click".

Pour la réaction de SPAAC, nous nous sommes orientés sur les dérivés DBCO (ou DIBAC pour dibenzozacyclooctyne) car il s'agit d'une des cyclooctynes les plus réactive. De plus, commercialement, il n'existe quasiment que des dérivés de fluorophores possédant un groupement DBCO pour effectuer la réaction de SPAAC.

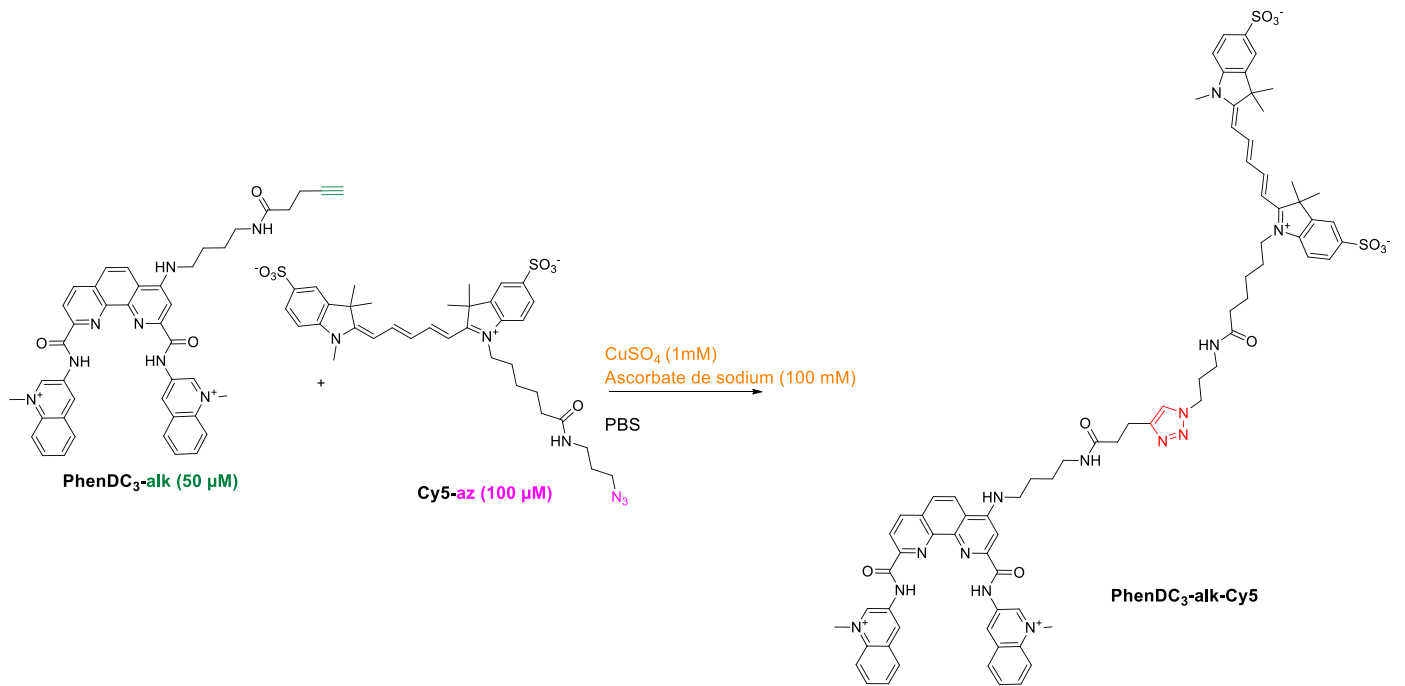


Figure 162 : Réaction de CuAAC entre Cy5-az et PhenDC₃-alk.

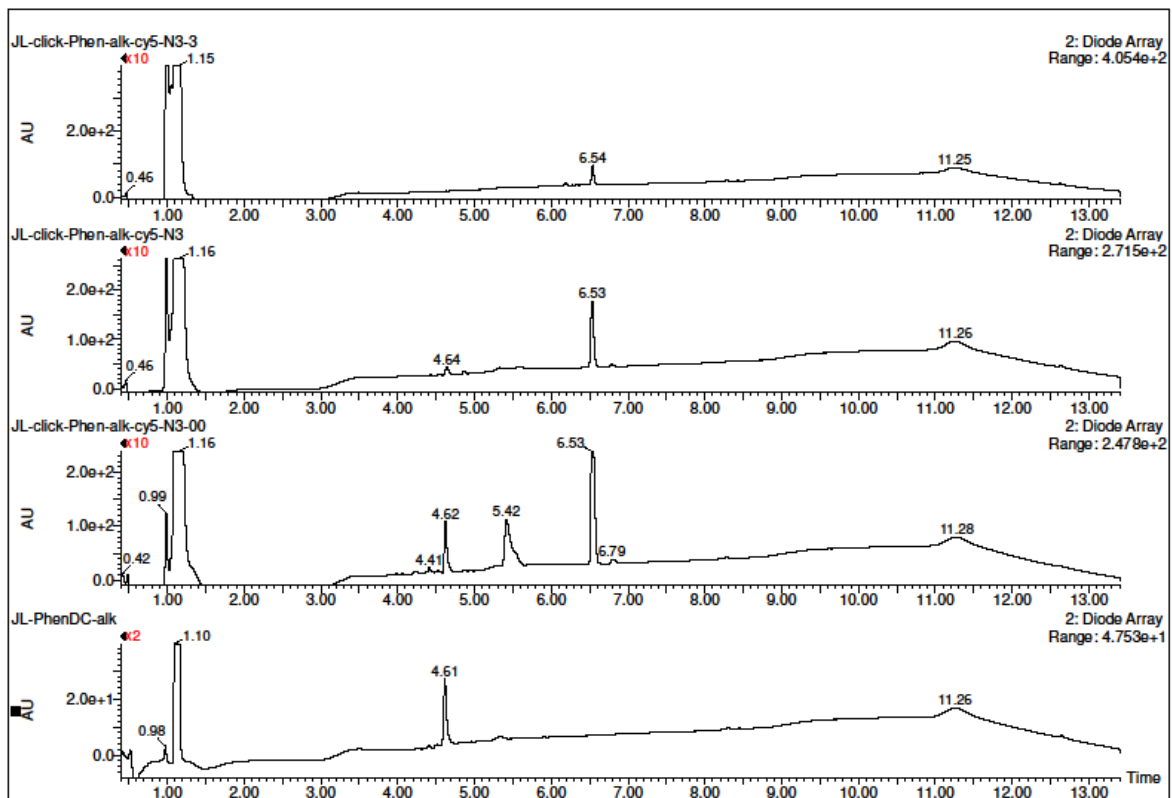


Figure 163 : Suivi de réaction de chimie "click" entre PhenDC₃-alk (50 μM, en bas) et Cyanine 5-az (100 μM) au bout de 10 minutes, 30 min et 1h (de bas en haut) en présence de sulfate de cuivre (1 mM) et ascorbate de sodium (10 mM) .

La réaction entre le dérivé alcyne du PhenDC₃ et le dérivé azoture de la Cyanine 5 est très rapide (figure 162 et 163), le produit de réaction est observable dès 10 minutes de réaction (à 5,42 min).

Cependant, la réaction ayant lieu à une concentration assez élevée (50 μM de PhenDC₃-alk et 100 μM de Cy5-az contre 5 μM et 20 μM respectivement en cellules), on constate que le produit de réaction disparaît après 30 min et par conséquent 1h. Cela s'explique par le fait que le produit final précipite dans l'épendorf alors que cela n'a jamais été observé lors de l'application en milieu cellulaire. Il est donc très important de conserver une concentration suffisamment faible pour ne pas faire précipiter le produit final et en même temps suffisamment élevée pour pouvoir observer de la fluorescence en microscopie confocale. On peut tout de même noter que le produit de départ PhenDC₃-alk n'est plus présent au bout d'une heure de réaction ce qui montre que le réactif a réagi complètement. La réaction est donc complète en une heure.

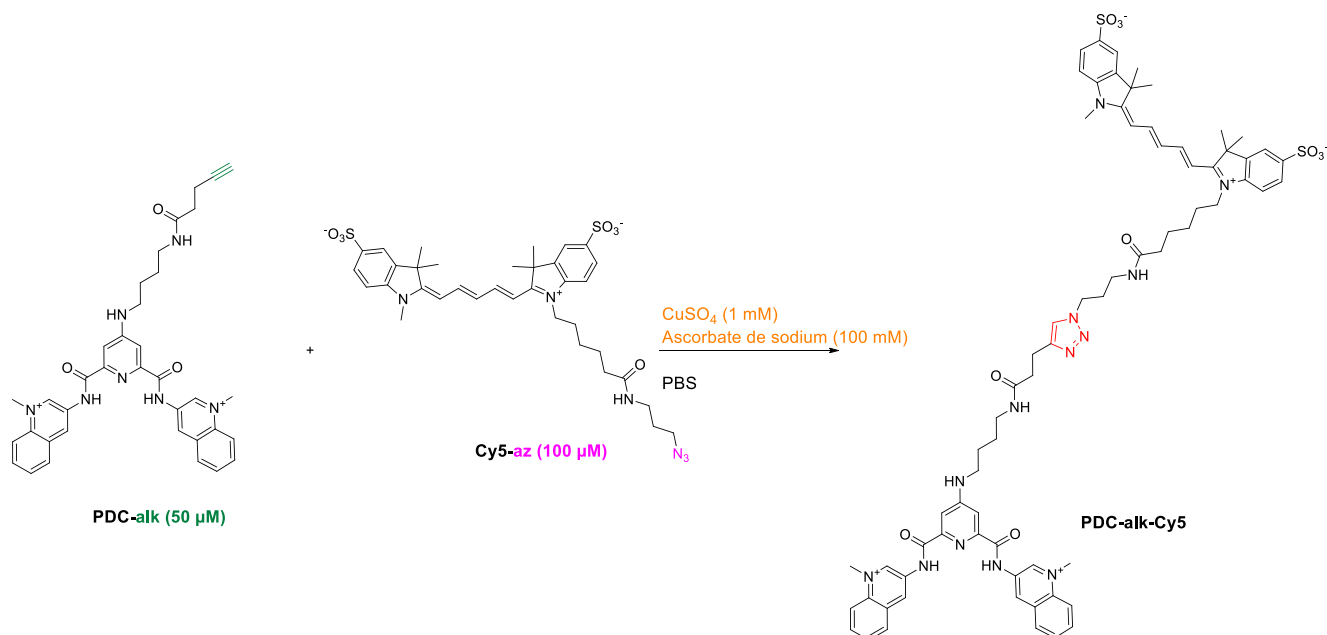


Figure 164 : Réaction de CuAAC entre PDC-alk et Cy5-az.

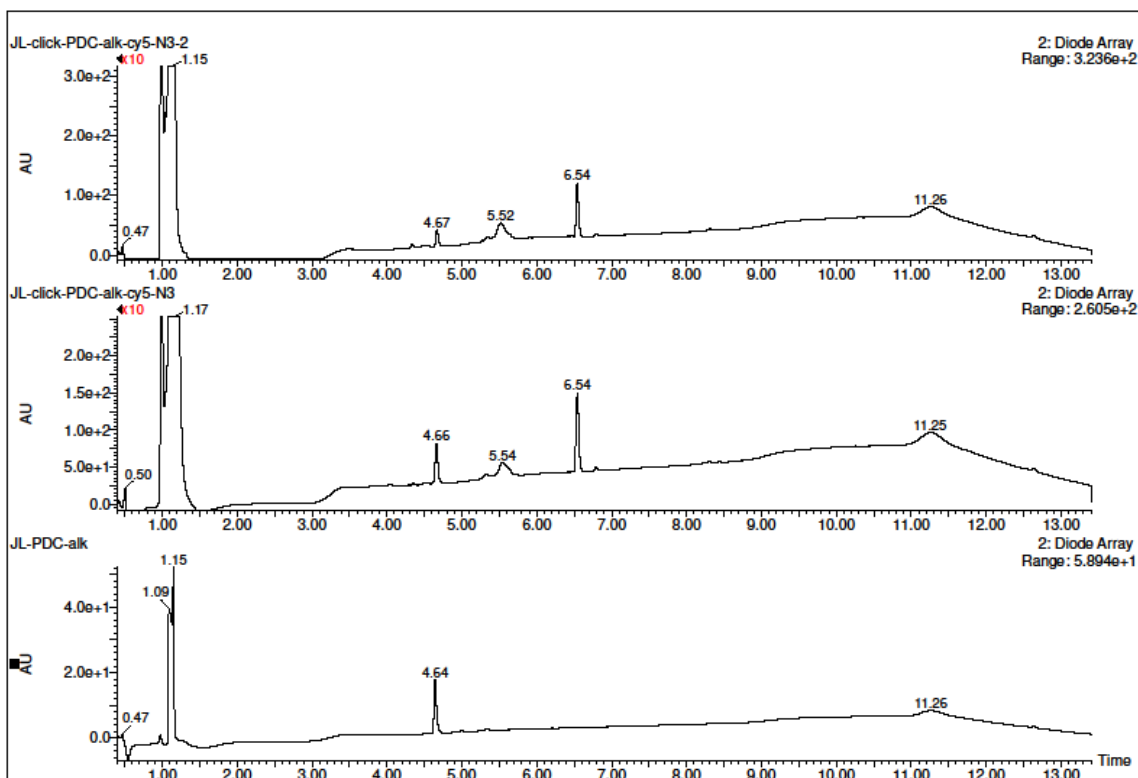


Figure 165 : Suivi de réaction par HPLC de la réaction de chimie "click" entre PDC-alk (50 μ M, en bas) et Cy5-az (100 μ M) après 30 min (au milieu) et 1h (en haut) de réaction en présence de sulfate de cuivre (1 mM) et d'ascorbate de sodium (10 mM).

La réaction est plus longue pour le dérivé alcyne du PDC (figure 164 et 165). En effet, après une heure de réaction, il reste du produit de départ cependant plus de 90% a été consommé par la réaction. Cela signifie que ce temps de réaction est suffisant pour observer la nouvelle molécule formée au sein des cellules.

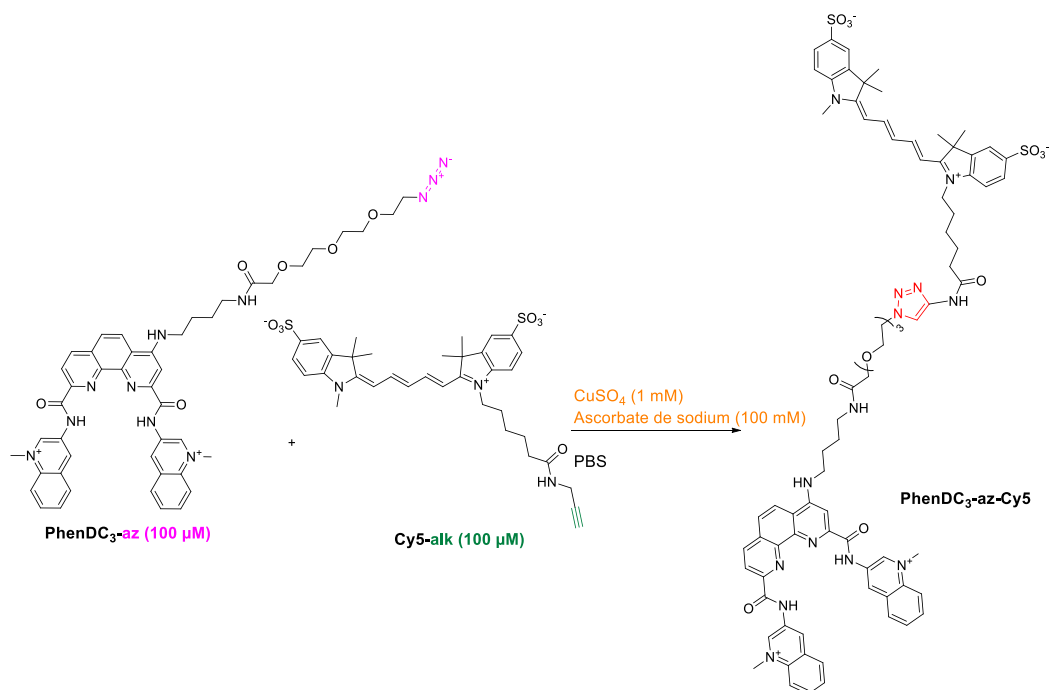


Figure 166 : Réaction de CuAAC entre PhendC₃-az et Cy5-alk.

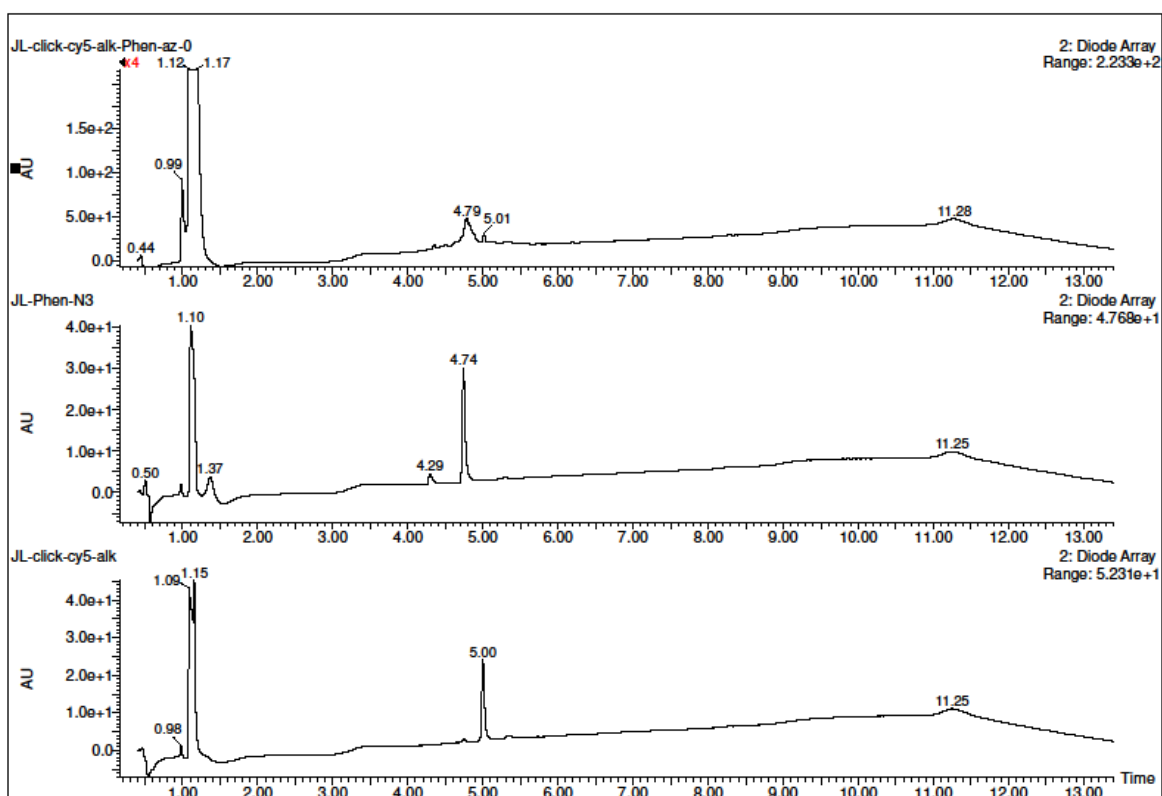


Figure 167 : Suivi de réaction par HPLC de la réaction entre PhendC₃-az (100 μM, au milieu) et Cy5-alk (100 μM, en bas) après 20 min de réaction (en haut) en présence de sulfate de cuivre (1 mM) et d'ascorbate de sodium (100 mM).

Lorsqu'on intervertit les partenaires de click, à savoir la fonction azoture sur le composé PhendC₃ et l'alcyne sur le fluorophore, la réaction est complète et encore plus rapide (figure 166 et 167). Après 20 minutes de réaction, le composé de départ n'est plus observable par HPLC ce qui signifie qu'il a

été consommé. De plus, le produit de la réaction est observable et lui ne précipite qu'après 1h de réaction à fortes concentrations de réactifs.

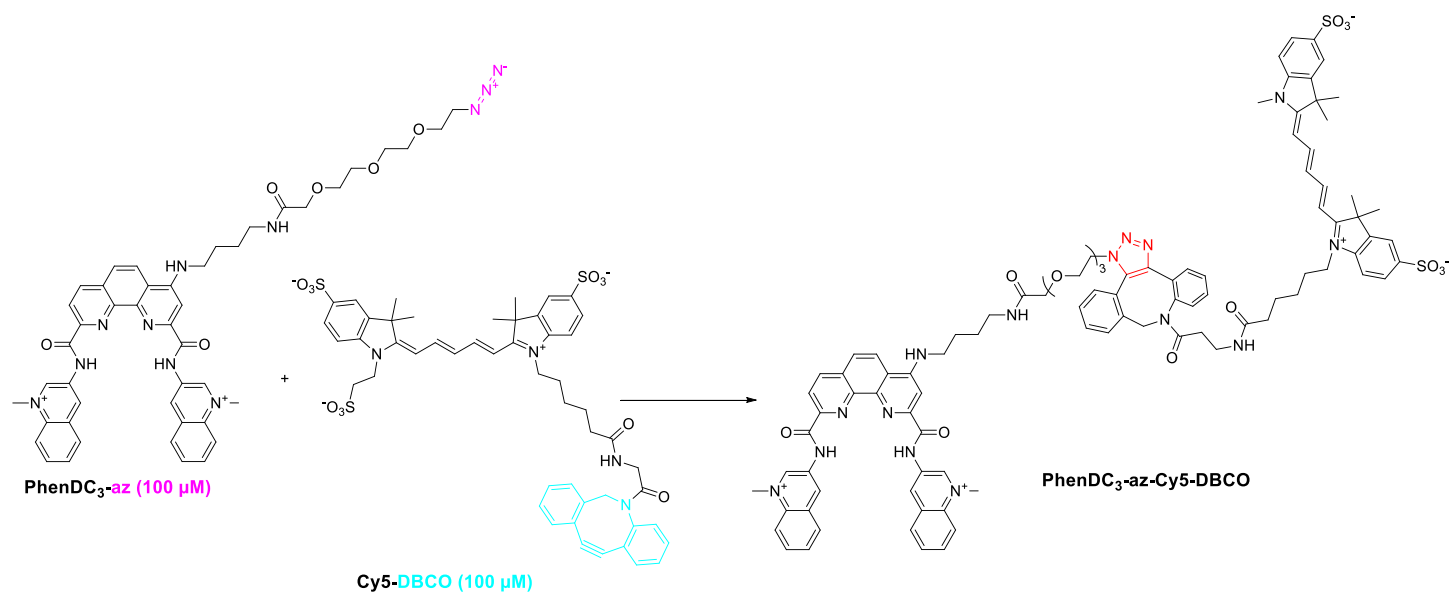


Figure 168 : Réaction de SPAAC entre PhenDC₃-az et Cy5-DBCO.

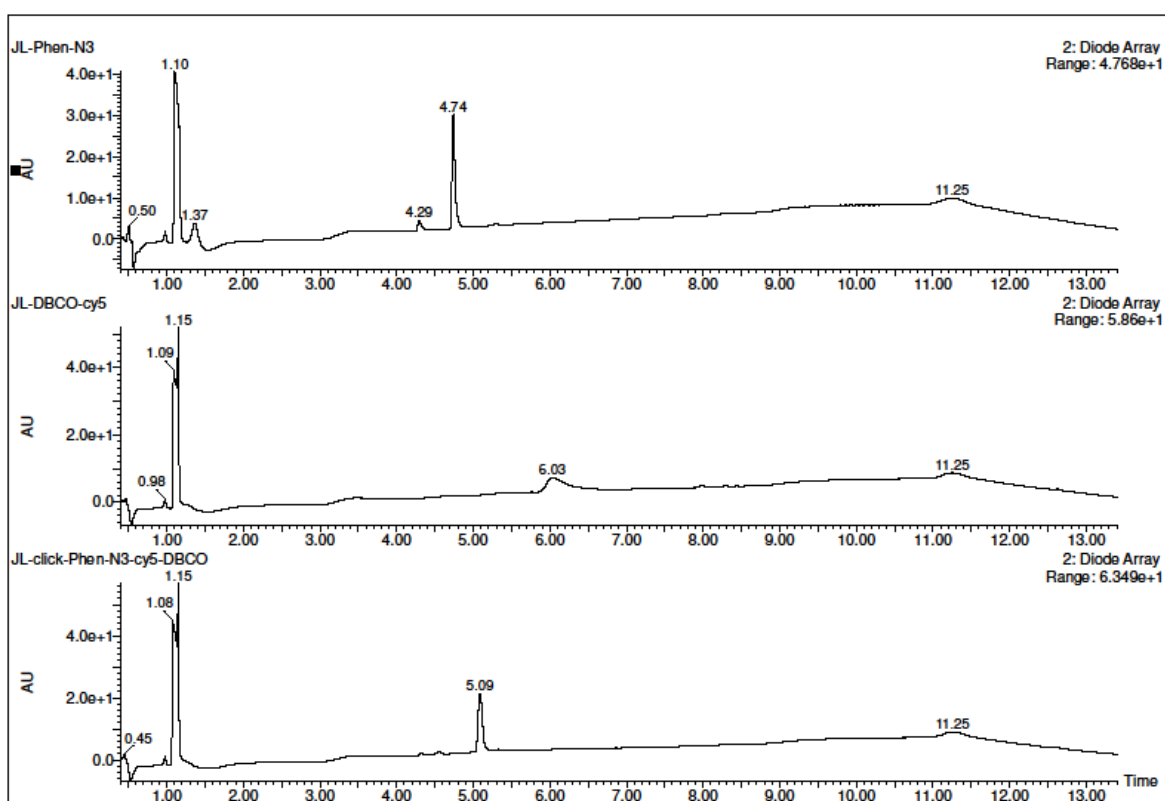


Figure 169 : Suivi de réaction par HPLC de la réaction entre PhenDC₃-az (100 μM, en haut) et Cy5-DBCO (100 μM, au milieu) après 30 min de réaction (en bas) dans le PBS.

Enfin, la réaction de SPAAC entre PhenDC₃-az et la cyclooctyne portée par le fluorophore est complète en 30 minutes (figure 168 et 169). De plus, le produit de réaction est toujours soluble dans le PBS et est donc observable par HPLC à cette concentration.

Le suivi *in vitro* des différentes réactions de chimie « click » avec les différents partenaires était un point crucial pour pouvoir la réaliser en milieu cellulaire. Il a donc été montré que les réactions se déroulent quelles que soient la fonction portée par le ligand (alcyne ou azoture) et que la réaction est complète en une heure de réaction. Il a donc été décidé que pour faire la réaction de chimie « click » *in cellulo*, une durée d'une heure va être utilisée.

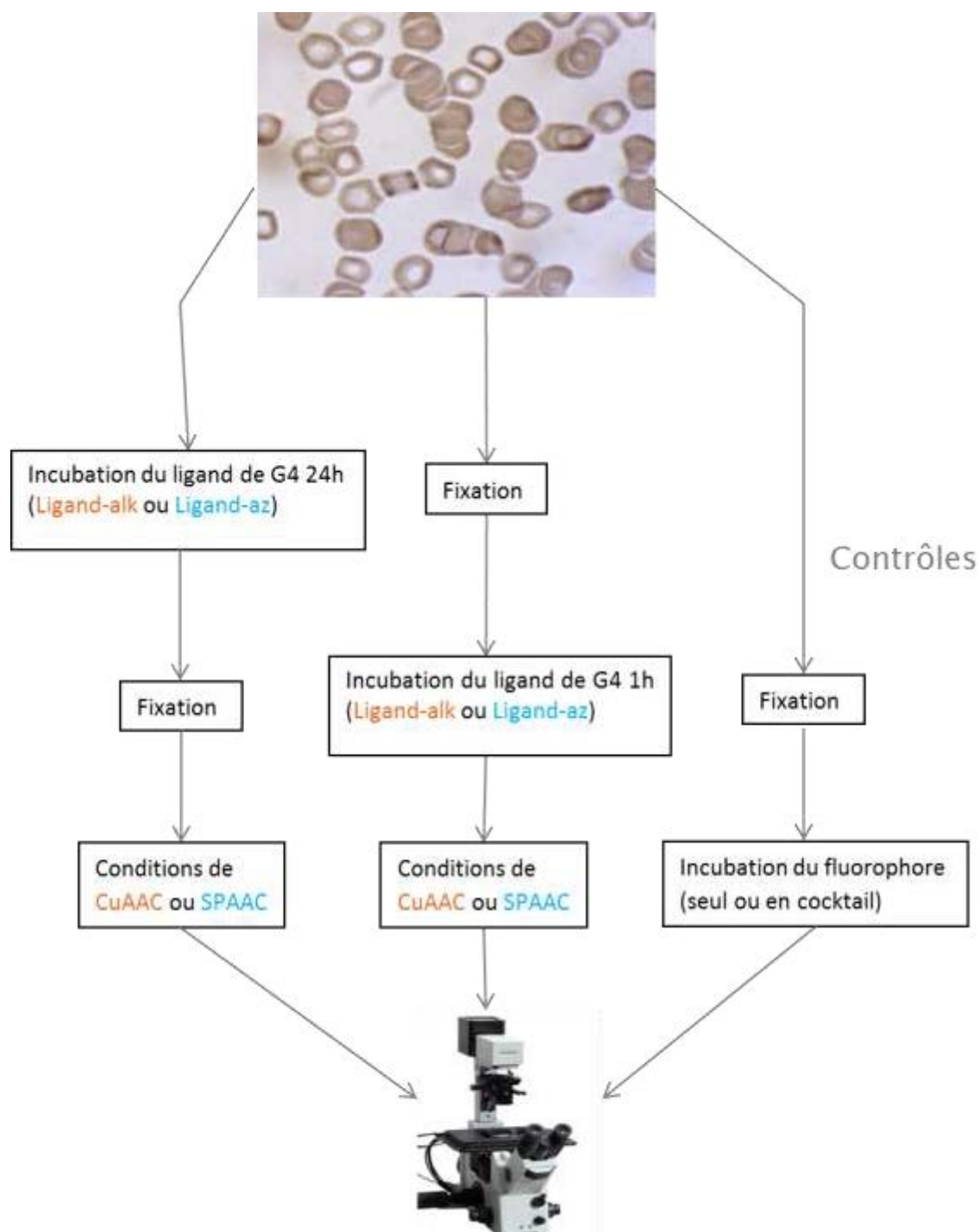
2) Etude en cellules fixées



1 : Nucléole, 2 : Noyau, 3 : Ribosome, 4 : Vésicule, 5 : Réticulum endoplasmique rugueux, 6 : Appareil de Golgi, 7 : Cytosquelette, 8 : Réticulum endoplasmique lisse, 9 : Mitochondries, 10 : Vacuole, 11 : Cytosol, 12 : Lysosome, 13 : Centrosome et 14 : Membrane plasmique.

Figure 170 : Structure type d'une cellule eucaryote.

L'objectif est donc de localiser les dérivés des composés PhenDC₃ et PDC dans les cellules (figure 170). Nous allons alors les fonctionnaliser avec un fluorophore *in situ* puis les observer au microscope confocal au sein de cellules fixées. Les lignées HT29 (cellules du cancer du côlon) et MDA-MB231 (cellules du cancer du sein) ont été utilisées. La fixation cellulaire chimique utilisée est une technique d'ajout de para-formaldéhyde sur les cellules en culture et est suivie d'une étape de perméabilisation par du Triton. Si l'imagerie en cellules fixées permet une manipulation plus aisée, certaines informations fonctionnelles peuvent être perdues lors de la fixation, puisque tous les transports actifs de la cellule vivante sont arrêtés et certaines membranes sont rendues perméables, permettant une relocalisation des métabolites.



Cocktail réactionnel de chimie « click » :

Ligand de G-quadruplex incubé à une concentration de 5 μM .

Conditions **CuAAC**:

Fluorophore : 20 μM

Sulfate de cuivre (II) : 1 mM

Ascorbate de sodium : 100 mM

Tampon : PBS

Conditions **SPAAC** :

Fluorophore : 10 μM

Tampon : PBS

Figure 171 : Stratégie de chimie "click" en cellules fixées.

Deux principales stratégies ont été envisagées pour étudier la localisation des ligands en cellules fixées (figure 171). D'une part, les ligands de G-quadruplex sont incubés en cellules vivantes pendant

24h ce qui correspond à la durée d'un cycle cellulaire complet pour les cellules des lignées étudiées. Ensuite, les cellules sont fixées par un traitement au para-formaldéhyde et perméabilisées avec du Triton avant incubation du cocktail réactionnel pendant une heure. Après différents lavages, la lame est montée et prête à être analysée au microscope confocal. D'autre part, les cellules sont d'abord fixées et perméabilisées avant incubation pendant une heure du ligand de G-quadruplex. Cette durée est suffisante car, en cellules fixées et perméabilisées, la dynamique de la cellule est totalement arrêtée ce qui permet à la petite molécule d'atteindre sa cible rapidement. Il est important de rappeler que la structure G-quadruplex est fugace ce qui veut dire que la présence de ces structures au moment de la fixation chimique n'est pas assurée. Après lavages, le cocktail réactionnel est également incubé pendant une heure et enfin la lame peut être montée. En parallèle, des contrôles sont bien évidemment effectués pour pouvoir comparer le fluorophore seul et le fluorophore en présence de ligands.

Deux fluorophores ont été choisis (figure 172). Il s'agit d'une coumarine azoture qui émet une exaltation de fluorescence plus importante lorsqu'elle a réagi avec un alcyne, on appelle cela une réaction fluorogénique. Cela peut nous permettre de savoir si la réaction a bien lieu en milieu cellulaire. Le second fluorophore étudié est la cyanine 5 qui émet dans le rouge ($\lambda_{\max} = 646 \text{ nm}$) ce qui nous permet de faire de la co-localisation avec du DAPI (4',6-diamidino-2-phénylindole, $\lambda_{\max} = 358 \text{ nm}$) ou du Hoechst 33342 ($\lambda_{\max} = 350 \text{ nm}$) qui marquent le noyau à des longueurs d'onde dans le bleu.

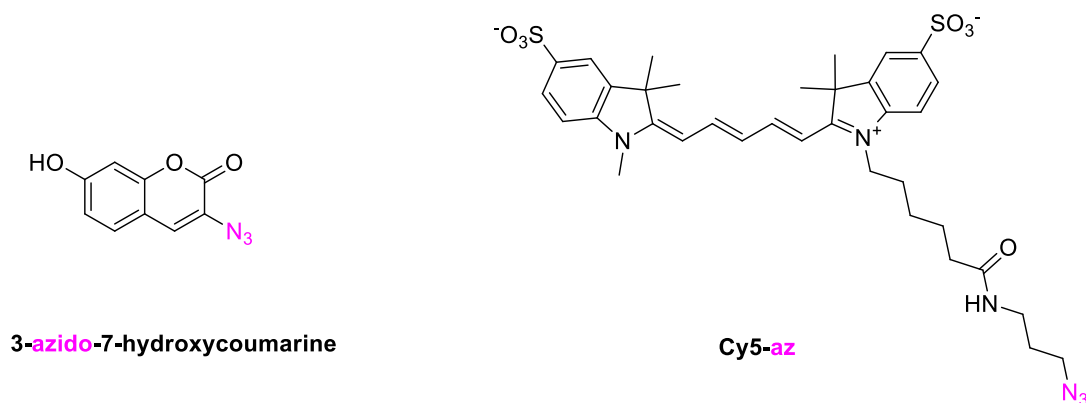


Figure 172 : Structure du 3-azido-7-hydroxycoumarine et de la cyanine 5-azoture.

Les résultats obtenus avec la coumarine comme fluorophore sont difficilement interprétables. En effet, que l'expérience se fasse en présence de ligand de G-quadruplex (figure 174 et 175) ou non (figure 173), la localisation du fluorophore est la même. Il est donc impossible de savoir si la réaction de chimie « click » a eu lieu ou non. Par ailleurs, l'intensité de fluorescence n'est pas modifiée alors soit la réaction n'a pas lieu soit la différence de fluorescence entre le composé seul et le composé « clické » n'est pas suffisamment importante pour l'observer par microscopie confocale. Très peu

d'exemples existent dans la littérature en utilisant la réaction de CuAAC en cellules cependant en utilisant la réaction de SPAAC le groupe de Paira a montré le caractère fluorogénique de la réaction.^[215]

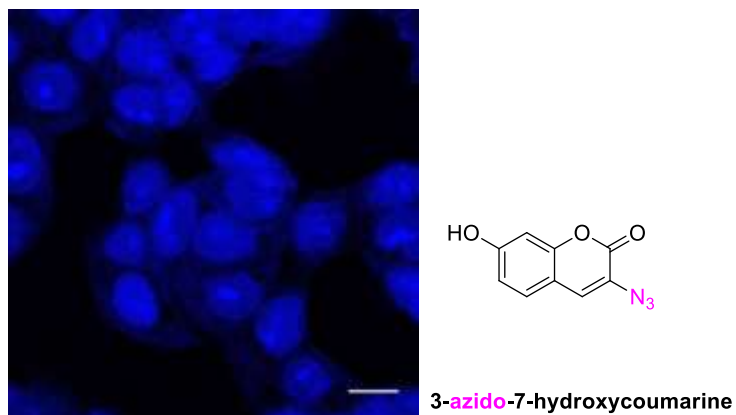


Figure 173 : Coumarine seule ; Barre d'échelle=10 μm

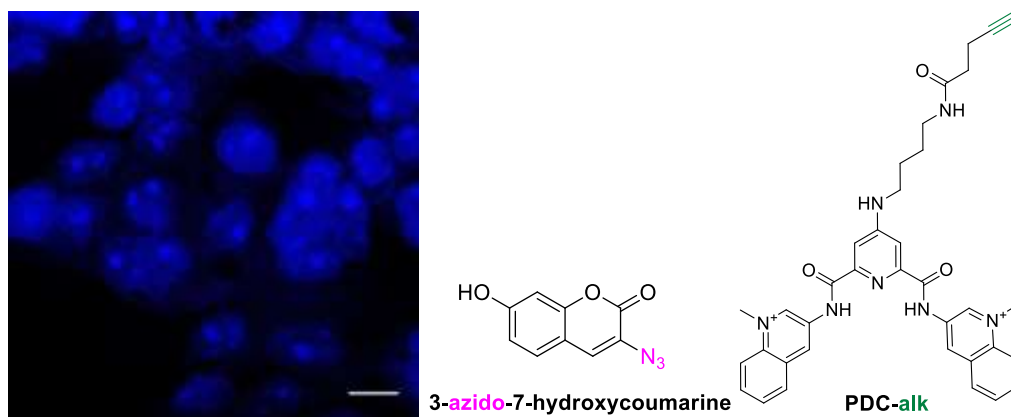


Figure 174 : PDC-alk « clické » à la Coumarine ; Barre d'échelle=10 μm

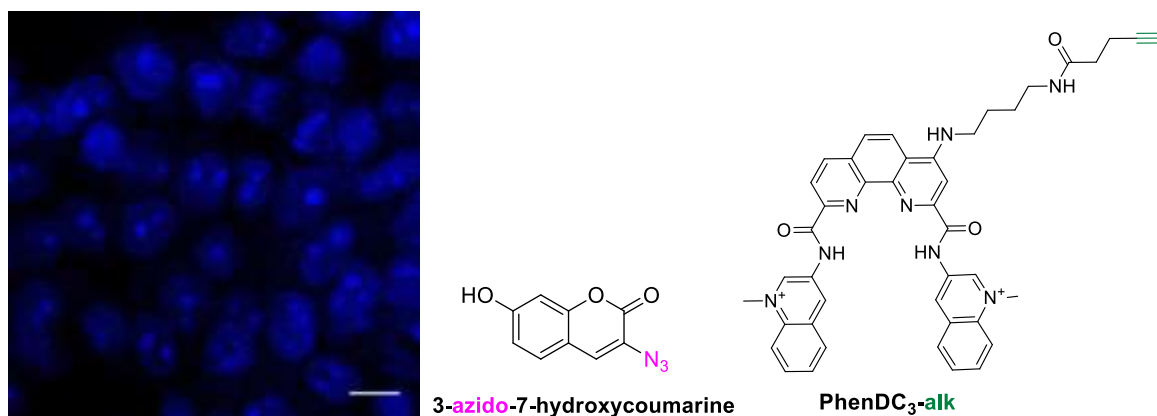


Figure 175 : PhenDC₃-alk « clické » à la Coumarine ; Barre d'échelle=10 μm

Nous nous sommes alors intéressés à la cyanine 5. L'avantage de ce fluorophore est que tous ses dérivés modifiables *in situ* (alcyne, azoture et DBCO, figure 176) sont commerciaux.

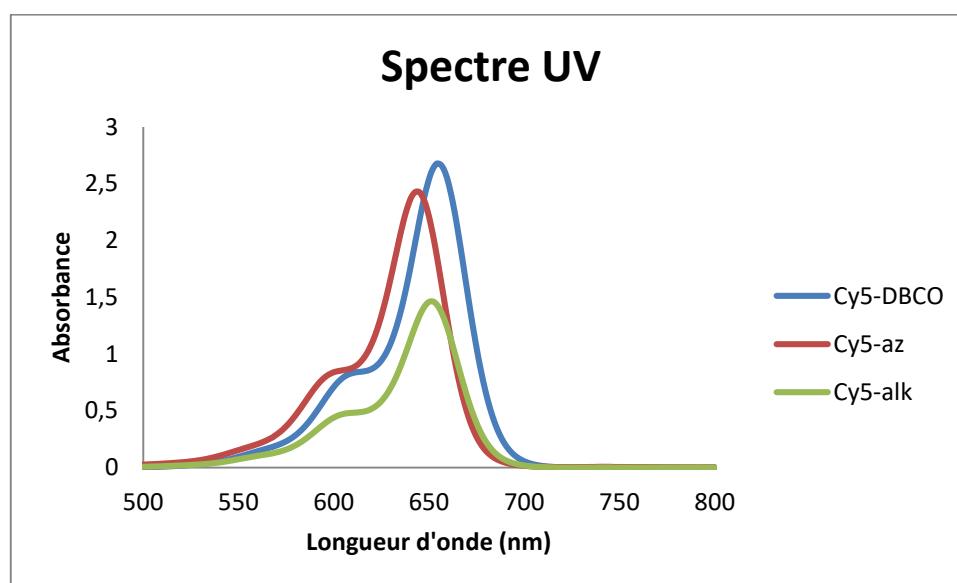
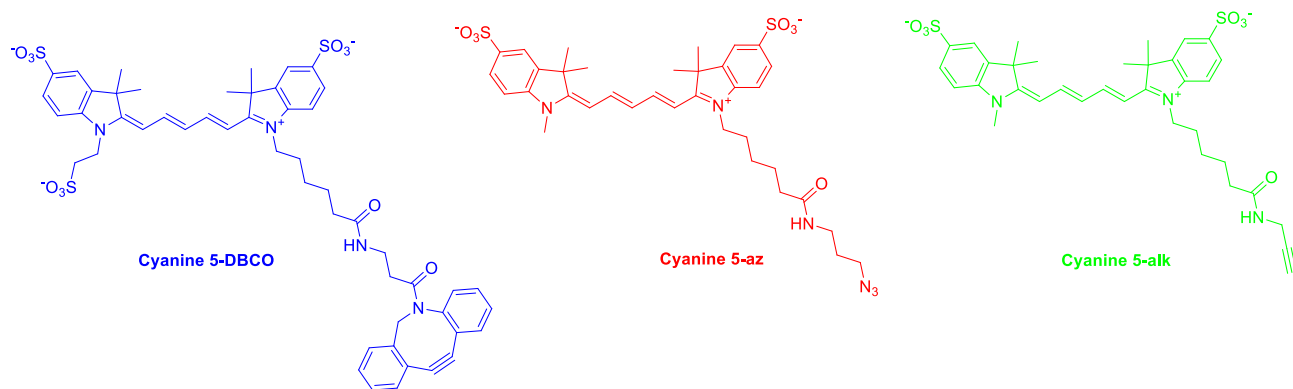


Figure 176 : Structure des dérivés "clickables" de la cyanine 5 et leur spectre UV à 10 μM.

Pour commencer, le dérivé azoture a été utilisé seul, en présence de cuivre et en présence de ligands de G-quadruplex (PhenDC₃-alk et PDC-alk). Le fluorophore seul est éliminé après lavage alors qu'en présence de ligands qui ont été incubés en cellules vivantes pendant 24h, il y a un marquage nucléolaire et dans le cytoplasme. Le marquage nucléolaire est en accord avec la littérature (figure 177) ^[68a] et c'est une zone de la cellule où les acides nucléiques présents sont riches en guanine ce qui en fait une cible pour nos ligands de G-quadruplex. Il est important de noter que la présence de ligands pendant 24h avec les cellules n'induit pas de mort cellulaire. Ces ligands ne sont pas cytotoxiques à cette concentration (5 μM) et les cellules, comme on peut le voir sur les images de microscopie, sont en très bon état.

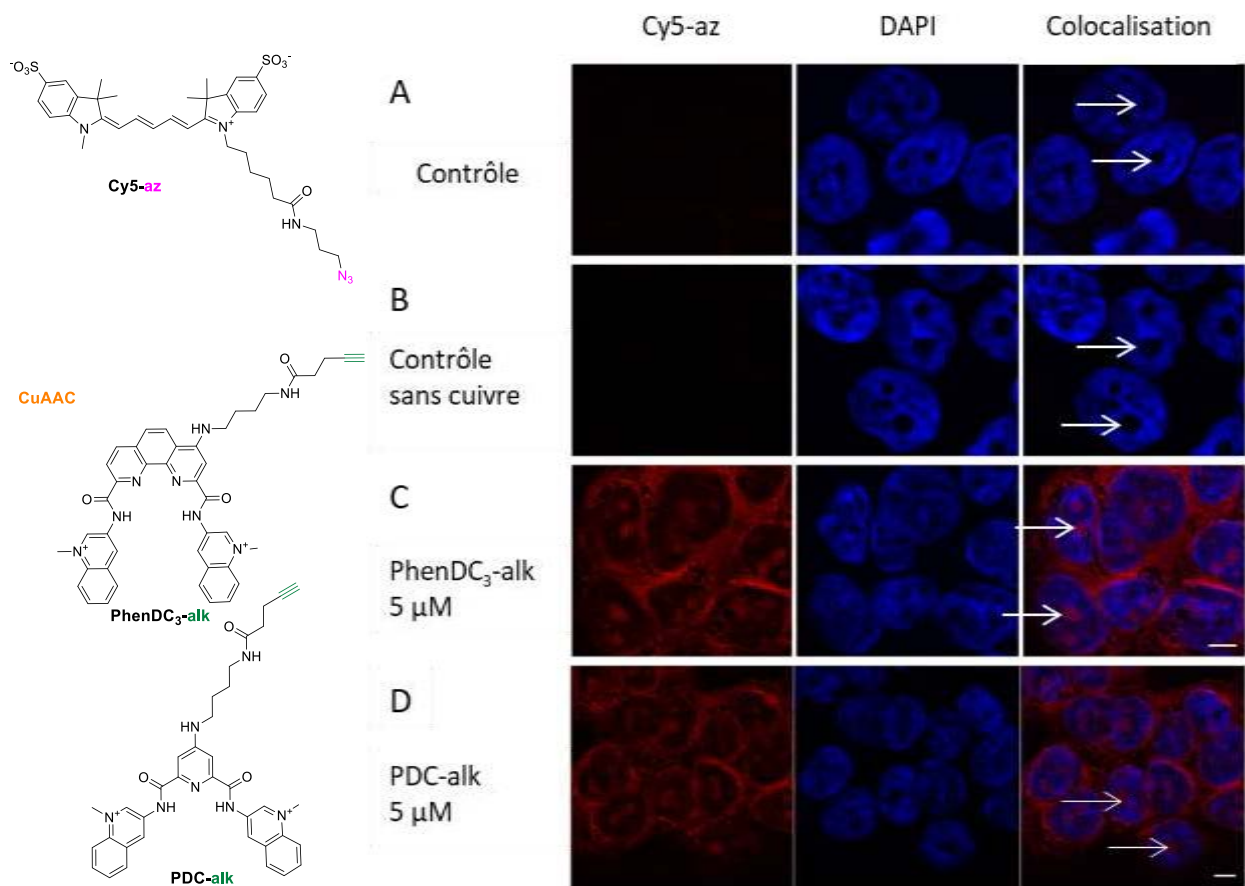
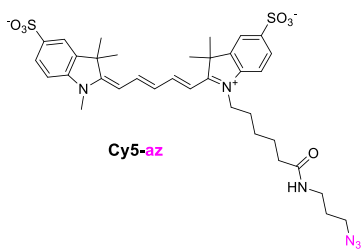


Figure 177 : Images de microscopie confocale des cellules HT29 fixées. **A.** Contrôle du fluorophore dans le cocktail réactionnel CuAAC, **B.** Contrôle du fluorophore seul, **C.** Réaction de CuAAC avec PhenDC₃-alk et **D.** Réaction de CuAAC avec PDC-alk. Barre d'échelle = 5 μm. Les flèches blanches montrent des nucléoles.

Les manipulations contrôles ont bien sûr été réalisées à la même concentration de fluorophore pour pouvoir comparer les images. De plus, le même gain est appliqué à chaque set d'images.

Lorsque nous augmentons le gain des expériences contrôles (figure 178), nous pouvons constater que le fluorophore seul est dans le cytoplasme mais ne pénètre pas le noyau. Ceci est une preuve supplémentaire que la localisation nucléolaire est bien due à la présence des ligands de G-quadruplex.



Contrôle

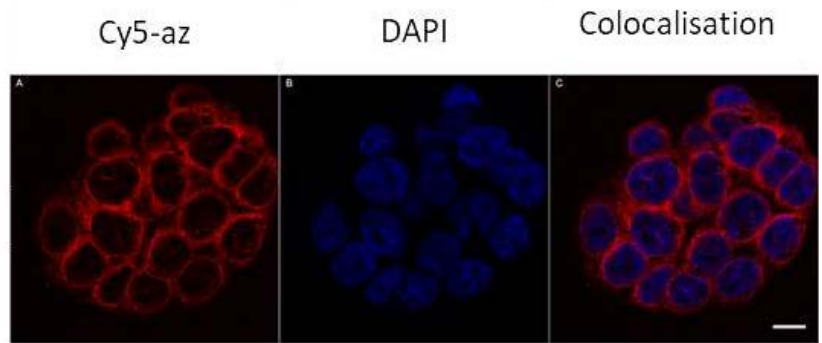
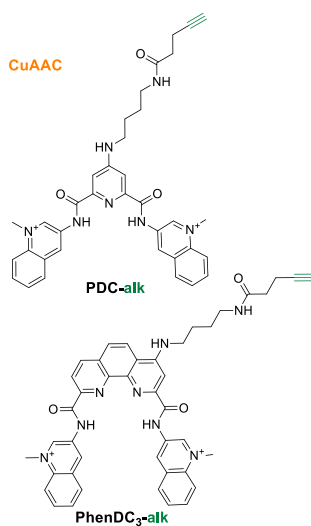


Figure 178 : Images de microscopie confocale de cellules HT29 fixées en présence de la Cy5-az à un gain très élevé. Barre d'échelle = 10 μm .

En parallèle, la même expérience a été effectuée en incubant les ligands une heure en cellules fixées (figure 179). Les résultats de cette expérience nous montrent que nous obtenons la même localisation en incubant le ligand en cellules vivantes ou en cellules fixées ce qui signifie que la fixation n'a pas d'effet de relocalisation du ligand. L'incubation des ligands en cellules vivantes est la méthode la plus intéressante. Les G-quadruplex sont des structures dynamiques qui se forment uniquement pendant des temps très court et à des moments bien précis du cycle cellulaire comme la transcription par exemple. Ainsi pour que le ligand puisse accéder à des structures G-quadruplex au cours du cycle cellulaire et puisse les stabiliser, il est nécessaire d'avoir un temps d'incubation long pour pouvoir les observer par la suite. C'est pourquoi, cette méthode a été utilisée pour la suite des expériences réalisées.



A

PDC-alk
5 μM

B

PhenDC₃-alk
5 μM

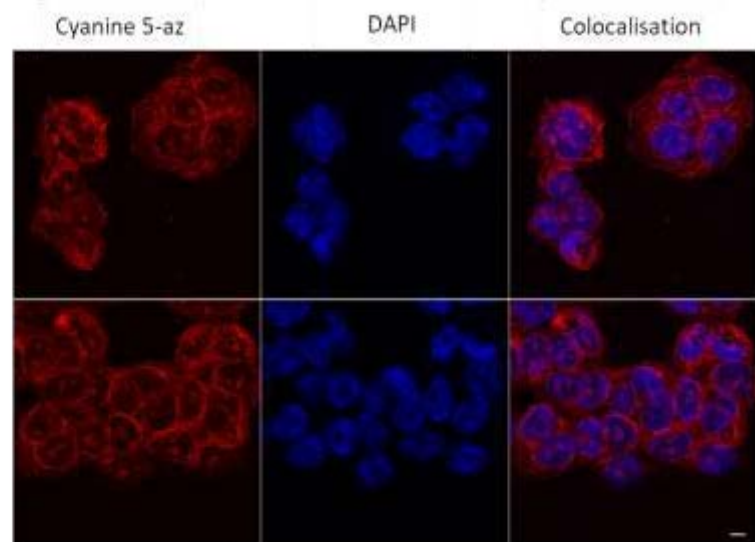


Figure 179 : Images de microscopie confocale de cellules HT29 fixées et incubées avec A. le ligand PDC-alk pendant une heure après fixation et B. le ligand PhenDC₃-alk pendant une heure après fixation. Barre d'échelle 5 μm .

Pour compléter cette étude, nous avons ensuite étudié la réaction de chimie « click » en changeant les partenaires. La fonction azoture est cette fois-ci sur le composé PhenDC₃ alors que sur le

fluorophore il y a une fonction dibenzo-azacyclooctyne (DBCO) qui permet de réaliser la réaction de SPAAC sans utilisation de sels de cuivre (figure 180).

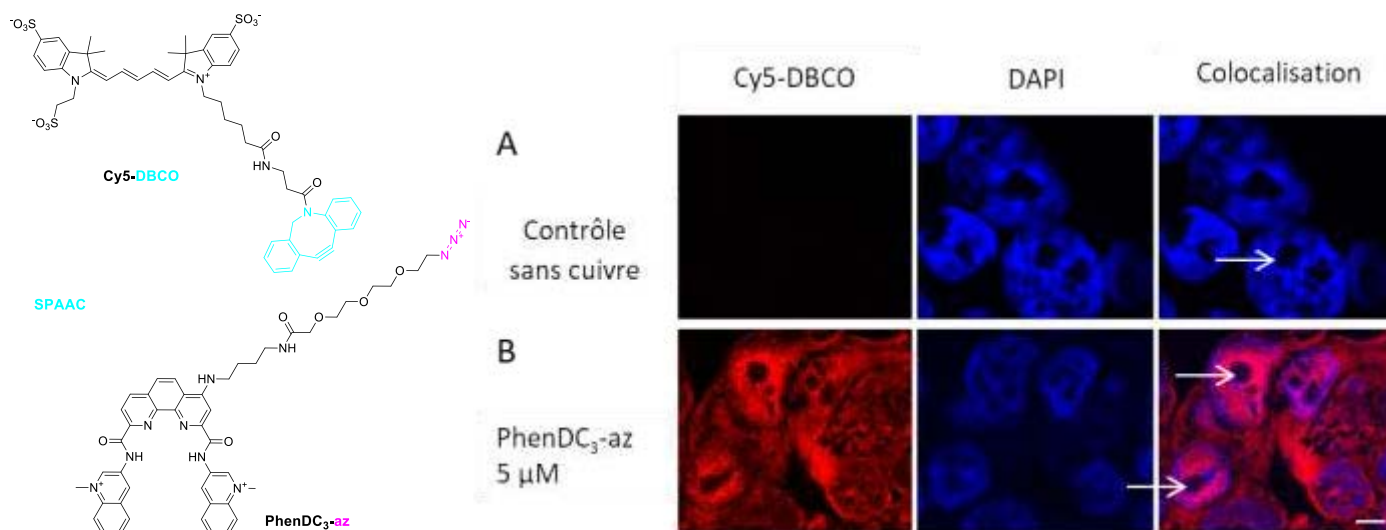


Figure 180 : Images de microscopie confocale des cellules HT29 fixées. A. Contrôle du fluorophore Cy5-DBCO seul et B. Réaction de SPAAC avec PhenDC₃-az. Barre d'échelle = 5 µm. Les flèches blanches montrent des nucléoles.

Encore une fois, le fluorophore seul est complètement lavé lors des différentes étapes de lavages cependant nous pouvons constater un résultat très surprenant en présence du ligand de G-quadruplex. En effet, nous nous attendions à avoir un marquage nucléolaire important car le dérivé azoture est très proche d'un point de vue structure au dérivé alcyne. Or, le marquage est principalement dans le noyau (nucléoplasme) sans marquage des nucléoles ainsi que dans le cytoplasme. On peut supposer que ce marquage cytoplasmique est dû à une accumulation de produits de couplage entre le ligand et le fluorophore à l'extérieur du noyau et, qu'une fois couplé, le produit est bloqué par la membrane nucléaire et ne pénètre pas dans le noyau. Enfin, il n'est pas étonnant d'observer un marquage nucléaire car notre ligand est un ligand d'ADN G-quadruplex.

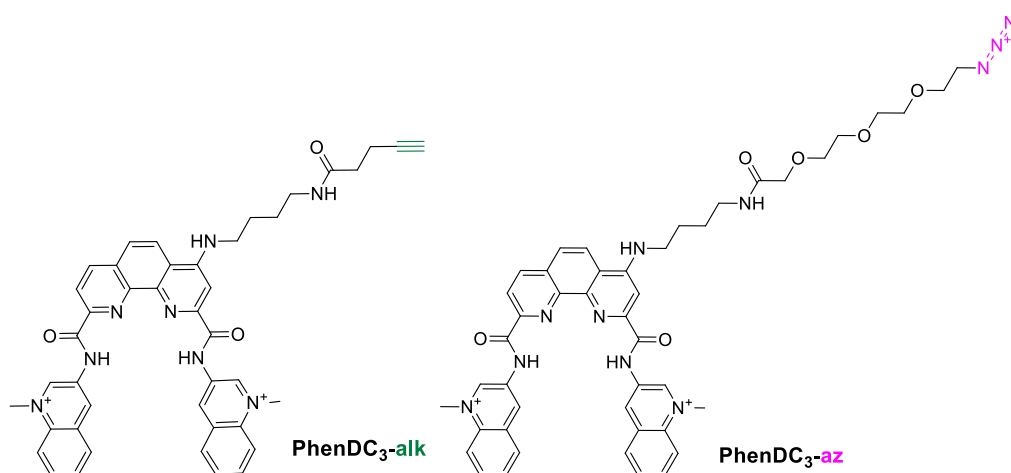


Figure 181 : Structures des composés PhenDC₃-alk et PhenDC₃-az.

Cependant, la différence de marquage entre les deux composés PhenDC₃-alk et PhenDC₃-az (figure 181) n'est pas cohérente. L'un est dans les nucléoles sans le nucléoplasme alors que l'autre est dans le nucléoplasme sans les nucléoles. La seule différence de structure est la taille de la chaîne où est branchée la fonction réactive. Cela ne suffit pas à modifier autant la localisation d'une molécule en milieu cellulaire. Nous avons alors cherché à comprendre cette différence de localisation.

Hormis la différence de structure des deux composés, la réaction de chimie « click » utilisée est également différente. Nous avons utilisé la réaction catalysée par des sels de cuivre pour le composé alcyne (CuAAC) alors qu'on n'a pas utilisé de cuivre pour le composé azoture (SPAAC). Nous avons donc étudié le rôle du cuivre. Pour cela, nous avons reproduit la même expérience que précédemment avec PhenDC₃-az et Cy5-DBCO mais cette fois-ci en présence des sels de cuivre bien qu'ils ne soient pas nécessaires (figure 182).

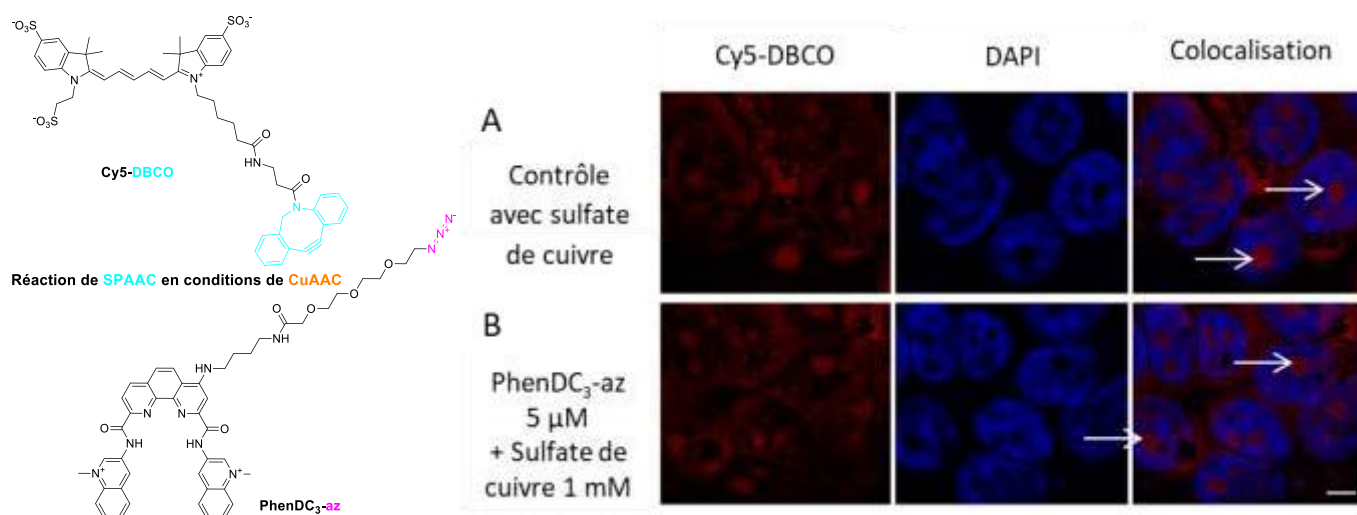


Figure 182 : Images de microscopie confocale des cellules HT29 fixées. A. Contrôle du fluorophore Cy5-DBCO en présence de sulfate de cuivre (1 mM) et d'ascorbate de sodium (100 mM) et B. Réaction de SPAAC avec PhenDC₃-az en présence de sulfate de cuivre (1 mM) et d'ascorbate de sodium (100 mM). Barre d'échelle = 5 μm. Les flèches blanches montrent des nucléoles.

L'expérience contrôle nous montre que même le fluorophore en l'absence de ligand présente un marquage nucléolaire en présence de cuivre (I) (figure 182A) alors qu'il était éliminé en absence de cuivre (figure 180A). De plus, le ligand qui avait une localisation nucléoplasmique (figure 180B) se retrouve maintenant dans les nucléoles sans le nucléoplasme (figure 182B). On a donc confirmation de notre hypothèse, le cuivre a une grande influence sur la localisation des ligands (figure 183).

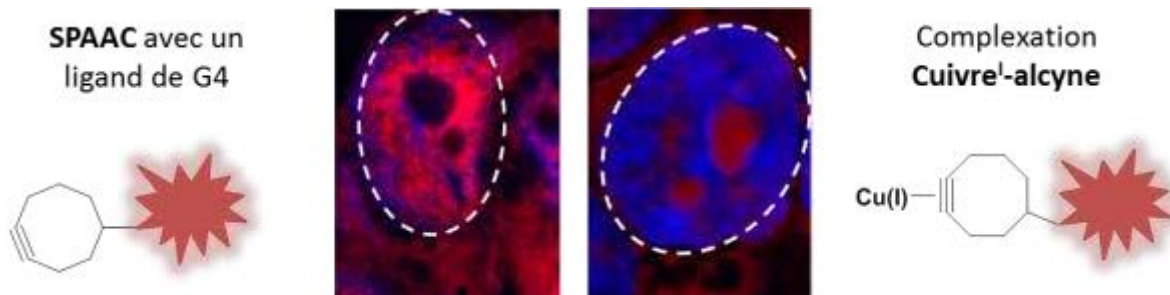


Figure 183 : Influence du cuivre sur la localisation de ligand.

Au vu des différents résultats obtenus jusqu'ici, on peut supposer que cette localisation est due au fait que le cuivre (I) complexe le composé possédant une fonction alcyne et le dirige vers les nucléoles. C'est pourquoi les deux dérivés alcyne du PDC et du PhenDC₃ sont observés dans les nucléoles ainsi que le fluorophore Cy5-DBCO en présence de cuivre. *A contrario*, le composé Cy5-az n'est pas localisé dans les nucléoles même en présence de cuivre (figure 177) ce qui montre bien l'importance de la fonction alcyne pour que la molécule soit dirigée dans les nucléoles.

Afin de confirmer ce résultat, nous avons effectué un grand nombre de contrôles. Tout d'abord, nous avons utilisé le dérivé alcyne du fluorophore, Cy5-alk (figure 184), seul et en présence du PhenDC₃-az pour avoir une idée de la localisation du ligand par la réaction catalysée au cuivre (CuAAC). On constate, comme pour le composé Cy5-DBCO, que le fluorophore est lavé lorsqu'il est introduit seul dans les cellules mais montre un marquage nucléolaire lorsqu'on y ajoute du sulfate de cuivre et de l'ascorbate de sodium. Par conséquent, lorsque le ligand est également incubé, le marquage nucléolaire est aussi observé. Cela confirme l'effet du cuivre sur la localisation nucléolaire. On obtient donc deux marquages distincts avec la même molécule, PhenDC₃-az, lorsqu'elle est en présence du fluorophore alcyne (réaction de CuAAC) ou du fluorophore cyclooctyne (réaction de SPAAC). La seule différence entre ces deux réactions est encore une fois l'utilisation du cuivre.

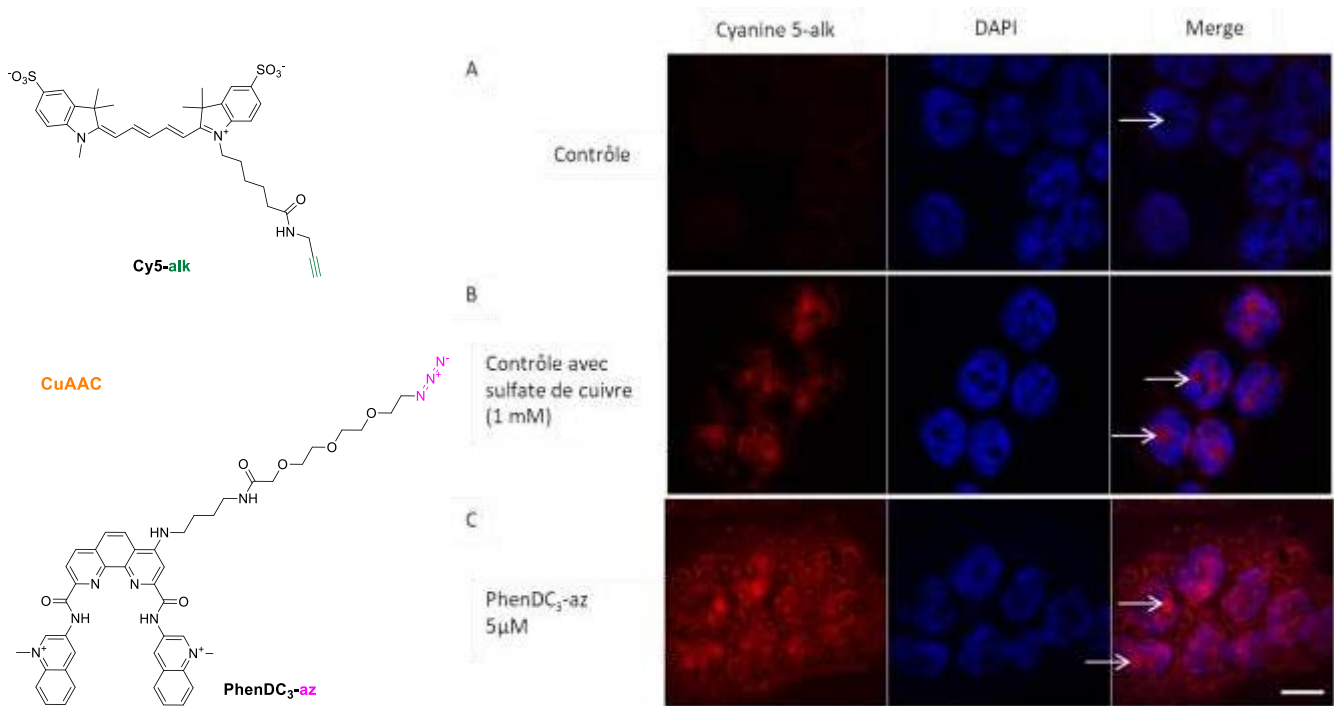


Figure 184 : Images de microscopie confocale des cellules HT29 fixées. A. Contrôle du fluorophore Cy5-alk seul, B. Cy5-alk en présence de sulfate de cuivre (1 mM) et d'ascorbate de sodium (100 mM) et C. Réaction de CuAAC avec PhenDC₃-az en présence de sulfate de cuivre (1 mM) et d'ascorbate de sodium. Barre d'échelle = 10 μm. Les flèches blanches montrent des nucléoles.

On peut tout de même noter qu'il y a également un marquage du nucléoplasme mais plus léger que les nucléoles (figure 184C). La molécule PhenDC₃-az, en l'absence de cuivre, marque le nucléoplasme ce qui est donc cohérent avec les résultats observés.

Afin de s'assurer qu'il ne s'agit pas d'un artefact dû au fluorophore, nous avons testé différentes familles de fluorophore. Après la cyanine, nous avons testé le TAMRA-DBCO qui possède donc une fonction dibenzoazacyclooctyne (figure 185) et l'AlexaFluor 488-alk qui possède une fonction alcyne (figure 186).

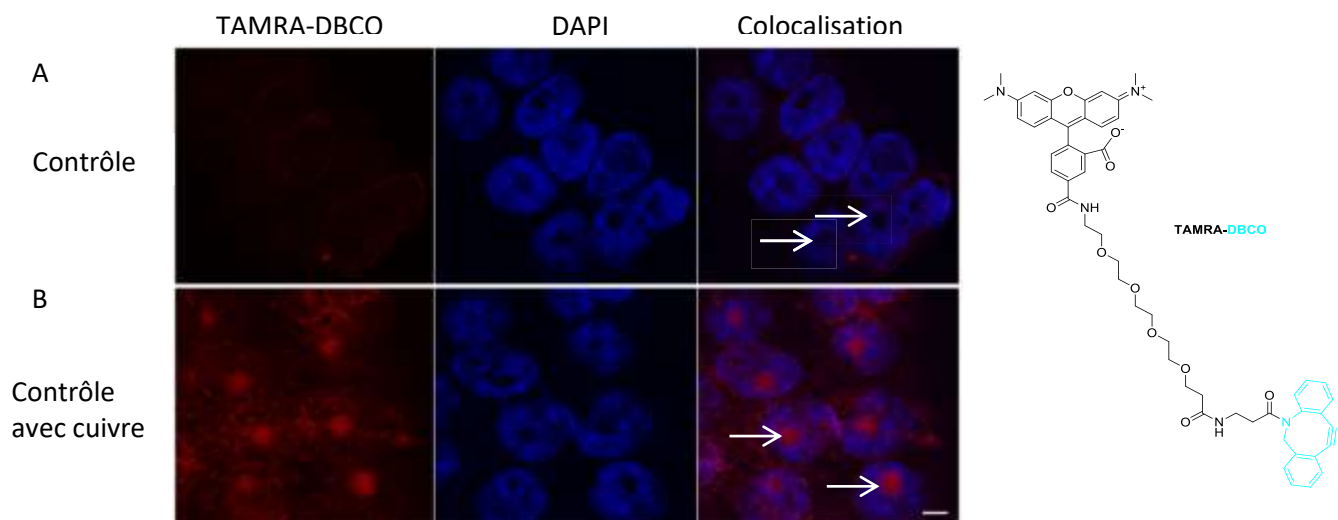


Figure 185 : Images de microscopie confocale des cellules HT29 fixées. A. Contrôle du fluorophore TAMRA-DBCO seul et B. TAMRA-DBCO en présence de sulfate de cuivre (1 mM) et d'ascorbate de sodium (100 mM). Barre d'échelle = 5 μ m.

Les résultats de ce contrôle nous montrent encore une fois que le cuivre a une très grande influence sur la localisation des ligands au sein de cellules de mammifères. Cela nous montre également que cette localisation dans les nucléoles n'est pas dû à un fluorophore particulier (la cyanine) mais que c'est bien la présence de la cyclooctyne et du cuivre qui engendre cette localisation nucléolaire.

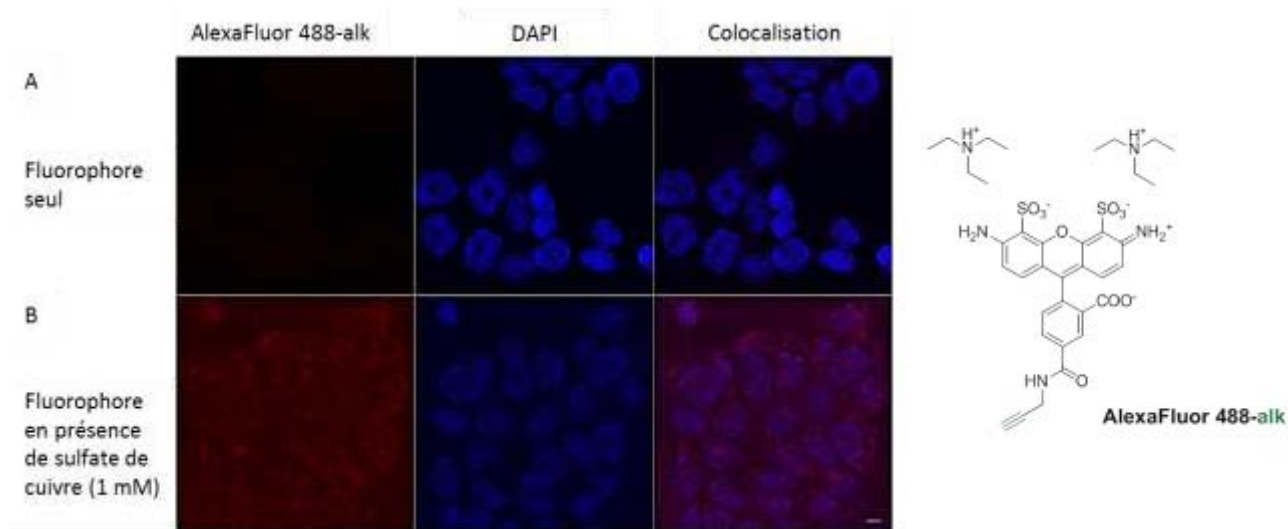


Figure 186 : Images de microscopie confocale des cellules HT29 fixées. A. Contrôle du fluorophore AlexaFluor 488-alk seul et B. AlexaFluor 488-alk en présence de sulfate de cuivre (1 mM) et d'ascorbate de sodium (100 mM). Barre d'échelle = 5 μ m.

Pour ce contrôle, les images ne sont pas très nettes, en effet, pour ce composé nous utilisons un laser différent moins puissant que celui utilisé précédemment pour la cyanine ou le TAMRA. Pour éviter d'endommager les cellules, nous n'avons pas augmenté la puissance du laser mais uniquement le gain pour obtenir plus de signal, le problème est que lorsque le gain est très élevé, nous perdons en qualité d'images. Cependant, malgré son absence de netteté, nous pouvons clairement constater

que l'ajout de sels de cuivre au milieu cellulaire entraîne une localisation du fluorophore dans les nucléoles (figure 186B). Encore une fois, le complexe Cu(I)-alcyne se forme et se dirige dans les nucléoles. La présence d'une triple liaison, qu'elle soit terminale en bout de chaîne aliphatique ou contrainte dans une cyclooctyne, suffit à ce que le cuivre puisse former un complexe et dirige la molécule directement dans les nucléoles.

Afin d'être complet au niveau des contrôles, nous avons refait toutes les expériences avec le fluorophore Cy5-DBCO sur une autre lignée cellulaire, MDA-MB 231, du cancer du sein (figure 187).

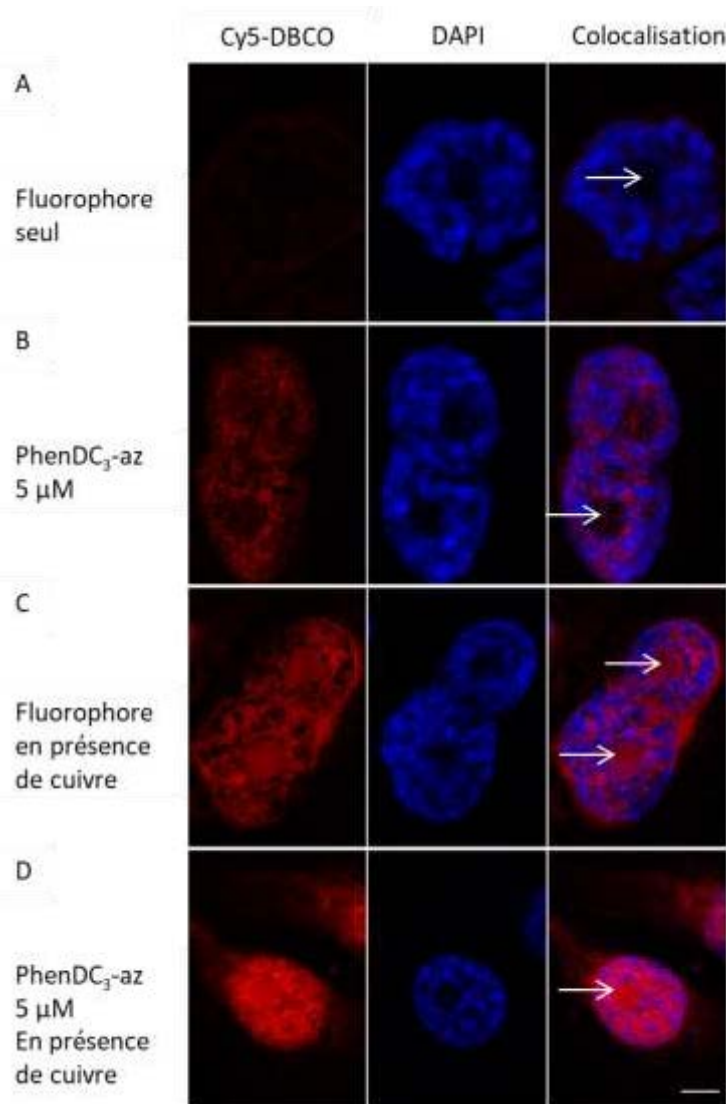


Figure 187 : Images de microscopie confocale des cellules MDA-MB 231 fixées. A. Contrôle du fluorophore seul, B. PhenDC₃-az (5 μM) « clické » à la Cy5-DBCO (10 μM), C. Cy5-DBCO en présence de sulfate de cuivre (1 mM) et d'ascorbate de sodium (100 mM) et D. PhenDC₃-az (5 μM) « clické » à la Cy5-DBCO (10 μM) en présence de sulfate de cuivre (1 mM) et d'ascorbate de sodium (100 mM). Barre d'échelle = 5 μm. Les flèches blanches montrent des nucléoles.

On constate que malgré le changement de lignée cellulaire, la présence de cuivre est suffisante pour diriger le complexe dans les nucléoles (figure 187C et D). Le fluorophore seul est éliminé par lavages (figure 187A), la fluorescence résultante est très faible et ne se situe que dans le cytoplasme.

Lorsqu'on effectue la réaction de SPAAC avec PhenDC₃-az, il y a un marquage du nucléoplasme sans les nucléoles (figure 187B). Une fois que des sels de cuivre sont ajoutés au fluorophore, nous constatons immédiatement un marquage nucléolaire. De même pour la réaction de SPAAC en présence de cuivre, la localisation de la molécule n'est plus uniquement nucléoplasmique mais également nucléolaire.

Tous les contrôles effectués confirment notre hypothèse de départ, le cuivre (I) complexe les composés possédant une triple liaison (alcyne libre et cyclooctyne) et les dirige directement dans les nucléoles. Ce résultat n'est pas dépendant de la structure du fluorophore ni de la lignée cellulaire testée.

Un point reste à éclaircir. En effet, il est compliqué d'expliquer pourquoi le cuivre montre une telle attirance pour les nucléoles. Il est vrai que c'est une organelle de la cellule riche en ADN, ARN et protéines et donc une zone riche en hétérocycle azoté, potentiellement ligand du cuivre. Dans la littérature, le groupe de Li a montré que le cuivre avait une influence sur l'évolution des nucléoles dans les cellules.^[147] Dans ce papier, les auteurs décrivent l'effet du cuivre (II) sur les cellules cependant nous avons montré également au cours des différents contrôles effectués que la présence d'ascorbate de sodium n'était pas nécessaire pour former le complexe et diriger les composés dans les nucléoles (figure 188) ainsi nous pouvons donc penser que le cuivre (II) introduit en large excès dans les cellules est en partie réduit au sein même de la cellule par les différentes réductases présentes.^[216]

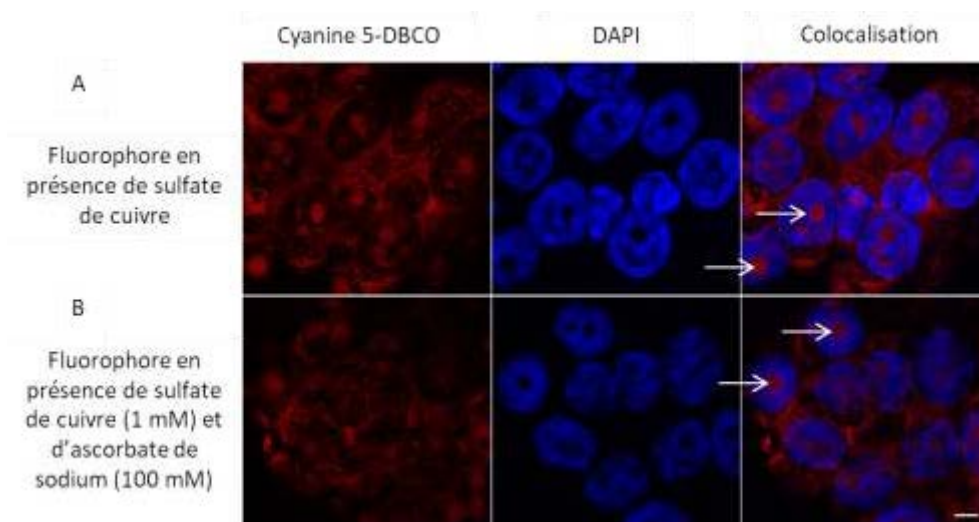


Figure 188 : Images de microscopie confocale des cellules HT29 fixées. A. Contrôle du fluorophore Cy5-DBCO (10 μ M) en présence de sulfate de cuivre (1 mM) et B. Contrôle du fluorophore (10 μ M) en présence de sulfate de cuivre (1 mM) et d'ascorbate de sodium (100 mM).

Afin de valider complètement l'hypothèse selon laquelle un complexe cuivre (I)-alcyne est impliqué dans la localisation des ligands dans les nucléoles de cellules de mammifères, nous avons réalisé une

titration des différents fluorophores dérivés de la Cyanine 5 (alcyne, azoture et DBCO) par des sels de cuivre (figure 189).

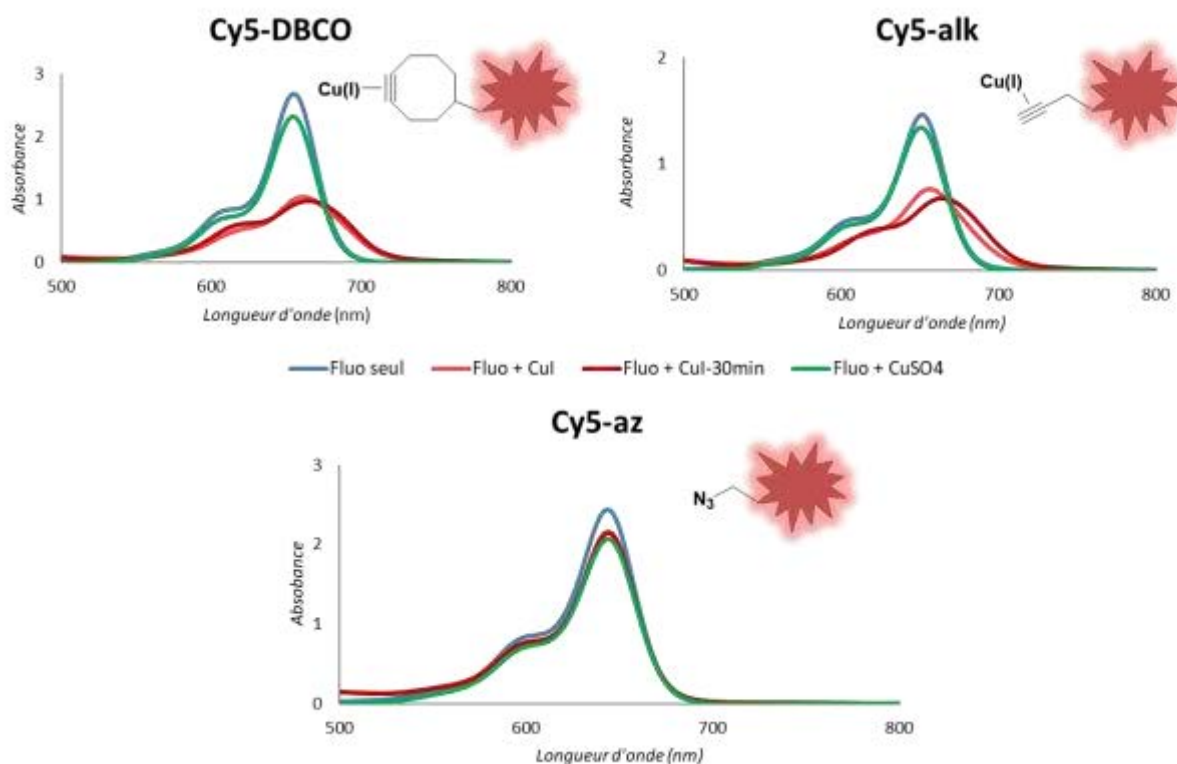


Figure 189 : Spectroscopie UV-visible des différents dérivés de la cyanine 5 seuls (11 μM) et en présence de cuivre (I) (CuI) ou de cuivre (II) (CuSO₄) (10 μM de Cy5 et 1 mM de Cu).

Les spectres UV-visible des dérivés de la cyanine 5 ont, tous les trois, le même spectre avec des maxima d'absorption légèrement décalés en fonction de la fonction présente sur le fluorophore. On peut constater qu'en présence de sulfate de cuivre, autrement dit de cuivre (II), le spectre des trois composés n'est absolument pas modifié (figure 189, courbes vertes), la légère diminution d'absorbance est dû à une légère dilution (11 μM composé seul et 10 μM pour les composés en présence de cuivre) lors de la titration. En présence de cuivre (I), l'iodure de cuivre, on peut remarquer que les composés possédant un alcyne ou une cyclooctyne forment un complexe avec le cuivre (I) (figure 189, courbes rouges). En effet, la titration du fluorophore Cy5-DBCO montre un déplacement d'absorbance, la valeur d'absorbance chute grandement passant d'environ 2,5 à moins de 1. Pour la cyclooctyne, il a été montré par le groupe de Hosoya que le cuivre peut la complexer et ainsi permettre une réaction de chimie « click » régiosélective en présence d'un alcyne libre et d'une cyclooctyne (figure 190).^[217]

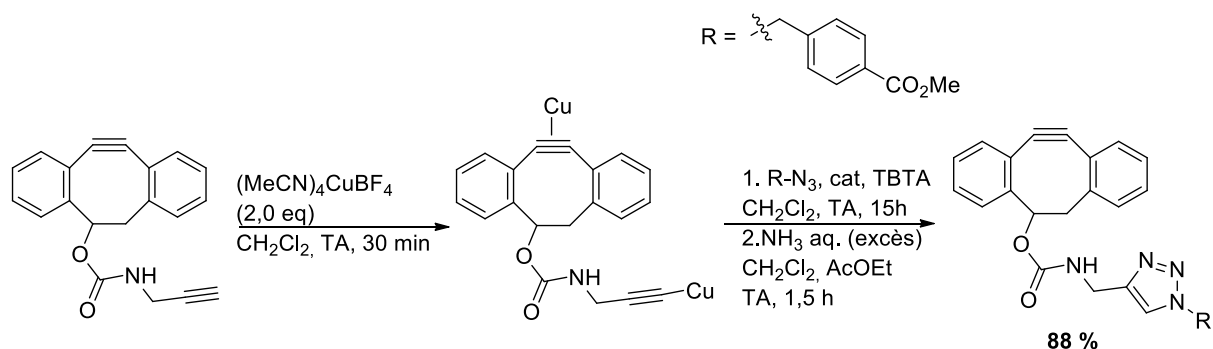


Figure 190 : Réaction régiosélective d'après le groupe de Hosoya.^[217]

La titration du composé Cy5-alk (figure 189) montre également une baisse d'absorbance passant de 1,5 à 0,7 et au bout de 30 minutes un décalage vers des longueurs d'onde plus élevées est également observé. Ce décalage montre que la formation du complexe du cuivre (I) avec l'alkyne terminal possède deux formes différentes. En effet, d'après le mécanisme de la réaction de cycloaddition d'un azoture et d'un alkyne catalysée par le cuivre, le cuivre se coordine à l'alkyne à la fois par liaison σ et par liaison π . La diminution d'absorbance peut donc être due à une forme du complexe et le décalage vers les plus hautes longueurs d'onde d'une seconde forme ou d'un complexe comportant deux cuivres sur l'alkyne possédant les deux liaisons évoquées précédemment (figure 191).

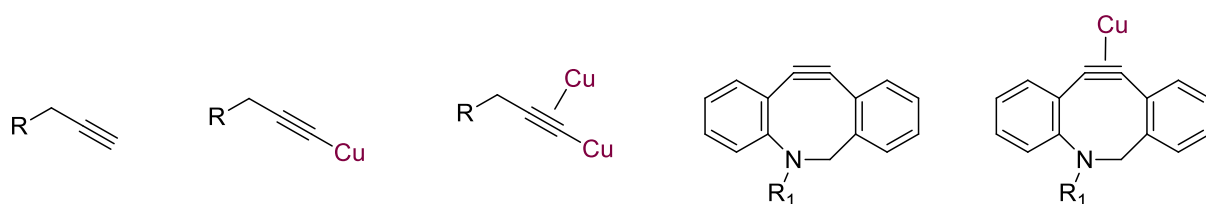


Figure 191 : Structures probablement formés d'un alkyne et une cyclooctyne en présence de cuivre (I).

Pour conclure, nous avons réussi à mettre en évidence un biais important de la réaction CuAAC en milieu cellulaire. En effet, le cuivre (I) se complexe aux triples liaisons présentes sur les différents composés et dirige ces complexes directement dans les nucléoles. Cela fonctionne aussi bien pour le ligand de G-quadruplex que pour le fluorophore. Un point reste encore difficile à résoudre, il s'agit de la localisation de ces composés. Le cuivre est attiré par cette organelle certainement car c'est une zone dense et riche en acides nucléiques et protéines mais il n'existe pas aujourd'hui d'explication précise de ce phénomène.

3) Etude en cellules vivantes

Les expériences en cellules vivantes ont été effectuées au sein des cellules du cancer du côlon HT29 et celles du cancer du poumon A549. La seule réaction possible parmi les deux utilisées

précédemment est la réaction de SPAAC. Nous utiliserons donc le composé PhenDC₃-az comme ligand de G-quadruplex et les fluorophores seront équipés d'une fonction DBCO.

Les deux points essentiels pour réaliser des expériences en cellules vivantes sont que les molécules, que ce soit le ligand de G-quadruplex ou le fluorophore, ne soient pas toxiques pour la cellule et que celles-ci pénètrent en cellules sans perméabilisation préalable. Nous savons déjà que les ligands de G4 ne sont pas toxiques pour les cellules car nous les avons laissés entre 24 et 48h en cellules vivantes à des concentrations variant entre 1 et 20 μM lors des expériences précédentes dans lesquelles les cellules étaient ensuite fixées. Concernant les fluorophores, nous nous sommes intéressés, tout d'abord, à la cyanine 5 et particulièrement à son dérivé DBCO qui a déjà été utilisé auparavant (figure 192).

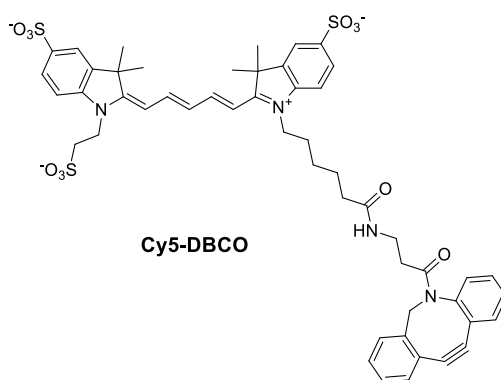


Figure 192 : Structure de la Cy5-DBCO.

Cette molécule s'est avérée toxique pour les cellules. En effet, lorsqu'elle est incubée pendant une heure à température ambiante, les cellules se décollent et sont éliminées lors des rinçages ce qui signifie que les cellules sont mortes. De même, le fluorophore en présence des cellules pendant trente minutes à 37°C tue également les cellules. C'est pourquoi nous faisons notre réaction de SPAAC à température ambiante pendant 30 minutes.

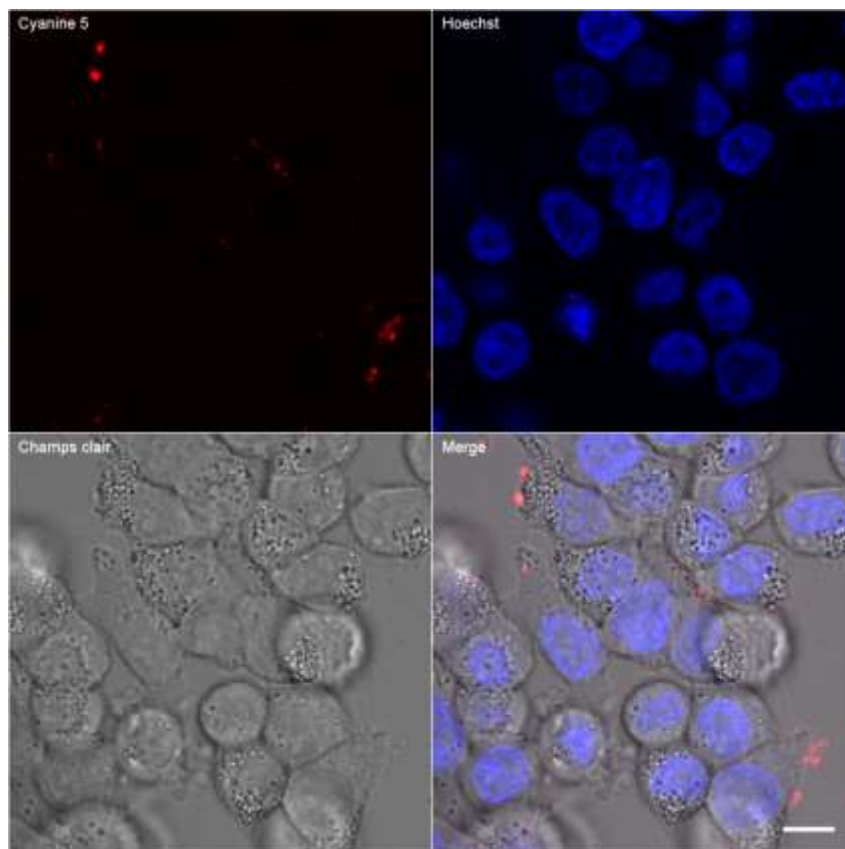


Figure 193 : Images de microscopie confocale de la Cy5-DBCO (rouge) et du Hoechst 33342 (bleu) en cellules HT29 vivantes. Echelle = 10 μm .

Nous pouvons constater (figure 193) que le fluorophore ne pénètre pas la membrane cellulaire et se retrouve ainsi coller à l'extérieur de la membrane. Ce fluorophore n'est donc pas adéquat pour observer notre ligand en cellule. Il est donc nécessaire de trouver un fluorophore fonctionnalisé par une cyclooctyne capable de pénétrer au sein des cellules. Dans la littérature, il a été montré que la Cyanine 5.5 pouvait pénétrer en cellules même en étant dérivé par une fonction DBCO.^[218] Cependant, il est à noter que le fluorophore est lié au cis-platine par une réaction de SPAAC avant d'être introduit au sein des cellules. Ainsi, la pénétration cellulaire du fluorophore seul n'est pas garantie. Nous en avons malheureusement eu la preuve avec, une nouvelle fois, un fluorophore bloqué sur les membranes cellulaires et ne pénétrant pas (figure 194).



Figure 194 : Images de microscopie confocale de la Cy5.5-DBCO (rouge) en cellules HT29 vivantes. Echelle = 10 μ m.

Nous nous sommes alors orientés sur le fluorophore TAMRA pour lequel une étude complète a été effectuée et a montré que le dérivé azoture était plus perméant que le dérivé DBCO.^[214] Cependant, ce dernier pénètre tout de même en cellules vivantes ce qui est un point très important que nous avons confirmé lors de nos expériences (figure 195). De plus, ce composé n'a pas de toxicité pour les cellules lorsqu'elles y sont exposées pendant plus d'une heure.

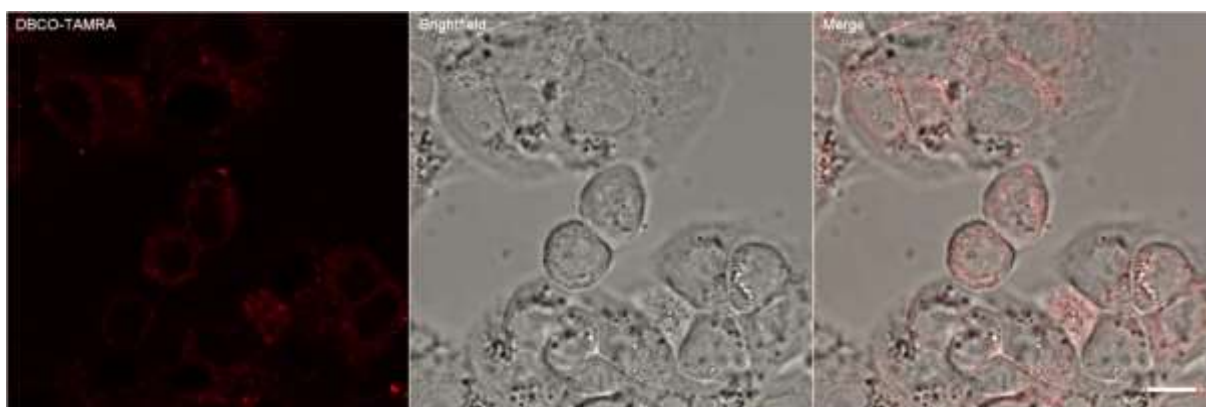


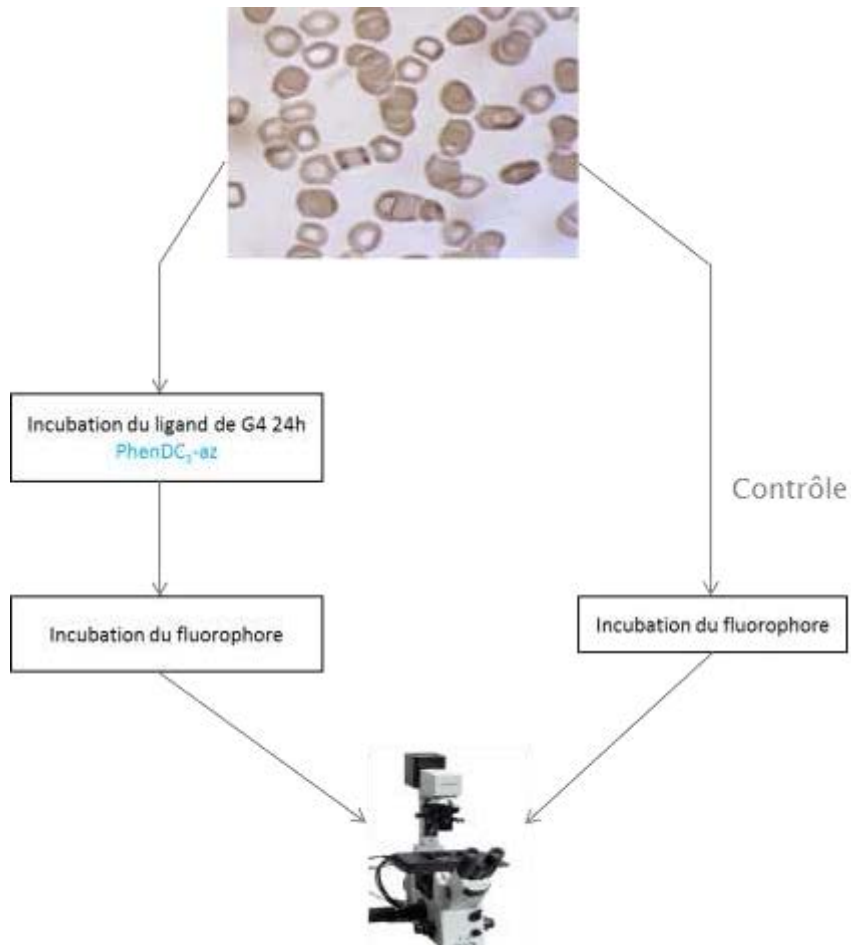
Figure 195 : Images de microscopie confocale du TAMRA-DBCO (rouge) en cellules HT29 vivantes. Echelle = 10 μ m.

Le fluorophore est donc présent dans le cytoplasme et ne pénètre pas à l'intérieur du noyau. Devant la difficulté de pénétration cellulaire, nous nous sommes interrogés sur la lignée cellulaire à utiliser pour les expériences en cellules vivantes. En effet, au préalable nous avons utilisé la lignée humaine HT29 issue du cancer du côlon qui possède des membranes peu perméantes et qui sont même difficiles à transfecter en utilisant un agent de transfection comme la Lipofectamine® (figure 196). Nous avons alors décidé de travailler au sein de cellules du cancer du poumon plus perméantes, la lignée A549. Toutes les expériences en présence de ligand de G-quadruplex en cellules vivantes ont été réalisées au sein de cette lignée cellulaire et les images en champ clair ne sont pas exploitables pour cette lignée cellulaire.

Lignée cellulaire	Efficacité de transfection Lipofectamine 3000	Origine
HT 29	< 30 %	Cancer du côlon humain
A549	51-79 %	Cancer du poumon humain

Figure 196 : Efficacité de la transfection en fonction de la lignée cellulaire d'après <https://www.thermofisher.com/fr/fr/home/brands/product-brand/lipofectamine/lipofectamine-3000.html#validated>

Le protocole utilisé est très simple. Le ligand de G-quadruplex est incubé à une concentration de 5 μM pendant 24h dans du milieu de culture puis le milieu est retiré et le fluorophore est ajouté pendant une heure à une concentration de 10 μM dans du milieu de culture. Une fois la solution retirée, du milieu de culture sain est ajouté et la lame est directement observée au microscope confocal. L'expérience contrôle sans incubation de ligand est également réalisée pour connaître l'impact du ligand de G-quadruplex sur la localisation (figure 197).



Ligand de G-quadruplex incubé à 5 μM et fluorophore-DBCO incubé à 10 μM .

Figure 197 : Stratégie de chimie « click » en cellules vivantes.

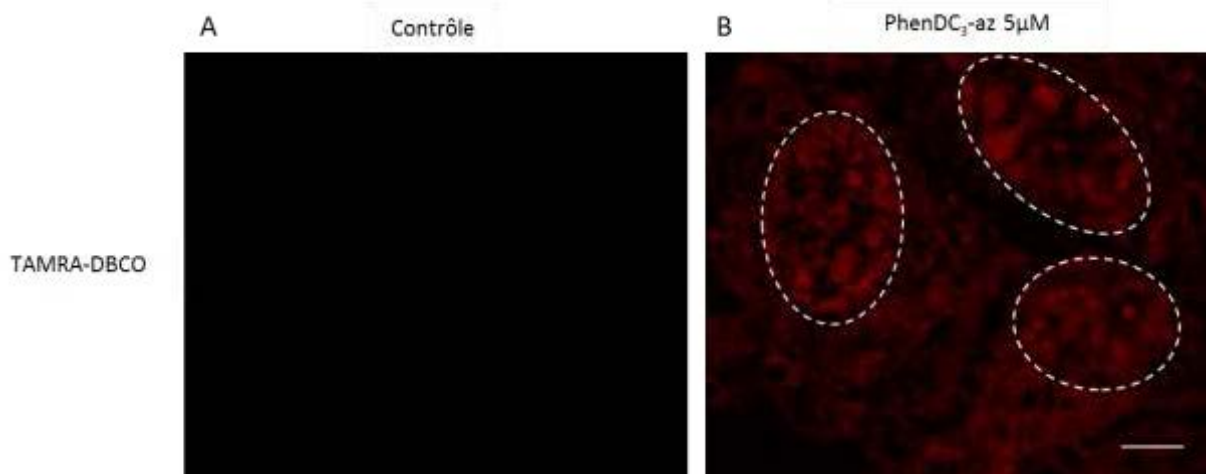


Figure 198 : Images de microscopie confocale du TAMRA-DBCO (rouge) A. seul et B. en présence de ligand PhenDC₃-az en cellules A549 vivantes. Echelle = 5 μm . Les ellipses en pointillés délimitent les noyaux des cellules.

On peut donc constater que le fluorophore n'est plus présent après lavage. En revanche, en présence de ligand de G-quadruplex, un marquage cytoplasmique et des foci plus intenses à l'intérieur du noyau sont observés (figure 198). On observe également des zones sombres à l'intérieur du noyau

qui correspondent aux nucléoles. On remarque donc que la localisation du ligand en cellules vivantes est similaire à la localisation en cellules fixées (figure 180B) c'est-à-dire une localisation nucléoplasmique et cytoplasmique du ligand PhenDC₃-az. Il est important de noter qu'aucune étape de rinçage n'a été effectuée après l'incubation du fluorophore, seulement un changement de milieu. Lorsqu'on effectue des étapes de rinçage, la molécule fonctionnalisée n'est plus visible à l'intérieur de la cellule, elle est évacuée des cellules. De même, lorsqu'une éventuelle co-localisation est envisagée avec un marqueur de noyau par exemple (Hoechst 33342), la molécule n'est plus observée dans les cellules, elle est probablement déplacée par la sonde fluorescente. Il y a une compétition entre le ligand et la sonde. Il s'agit donc d'un problème majeur pour savoir dans quelles organelles précisément se trouve le ligand.

Pour conclure sur ces expériences en cellules vivantes, il est important de noter que les fluorophores dérivés par un groupement DBCO ont des difficultés de pénétration cellulaire en fonction de la lignée cellulaire utilisée. Cela reste un défi pour le chimiste de trouver une cyclooctyne réactive qui pénètre en cellules vivantes. Une manière de s'affranchir de ce problème de pénétration des fluorophores serait d'échanger les partenaires de la réaction de SPAAC. La fonction DBCO pourrait être installée sur le ligand de G-quadruplex et l'azoture serait sur le fluorophore. Ainsi, le fluorophore n'aurait plus de problème de pénétration. Cependant, pour le ligand de G-quadruplex, il risque d'y avoir un problème de perte d'affinité pour la cible *in vitro* sur le composé PhenDC₃ et peut être également de pénétration cellulaire. Le composé idéal pour essayer cette stratégie est le ligand PDC qui porterait donc un groupement DBCO.

Par ailleurs, afin de connaître plus précisément la localisation du ligand de G-quadruplex au sein des cellules, il serait intéressant d'utiliser des fluorophores qui induisent des temps de vie de fluorescence différents lorsqu'ils sont fonctionnalisés par réaction de chimie « click » ou encore mieux lorsqu'ils sont en présence d'ADN duplex ou de structures G-quadruplex comme le groupe de Vilar l'a fait.^[97] L'avantage de notre système est que nous utilisons des ligands spécifiques des G-quadruplex contrairement à ce qui a été fait, cela induira forcément une différence de temps de vie de fluorescence plus importante en présence de G4 qu'en présence de l'ADN duplex.

Quelques expériences *in vitro* ont été réalisées dans ce sens et mériteraient, à mon avis, d'être poursuivies au vu des résultats intéressants.

IV. Analyse de temps de vie de fluorescence

La durée de vie de fluorescence d'un fluorophore est un paramètre qui peut mettre en évidence des interactions entre le fluorophore et son environnement biologique. Dans notre cas, nous avons essayé de savoir si le complexe PhenDC₃-az-G-quadruplex et PhenDC₃-az seul présentaient des

différences de temps de vie. Nous avons également essayé de faire réagir notre ligand avec une cyclooctyne fluorescente, céto-DMDIBO (figure 199), développée dans l'équipe de Joëlle Dubois (thèse de Vida Terzic). Ces mesures ont été effectuées au laboratoire de Rachel Méallet-Renault à l'ISMO (Institut des Sciences Moléculaires d'Orsay). Il s'agit de résultats préliminaires mais qui peuvent être intéressants pour le développement d'un nouveau projet de cyclooctynes fluorescentes pour le marquage de ligands en milieu cellulaire.

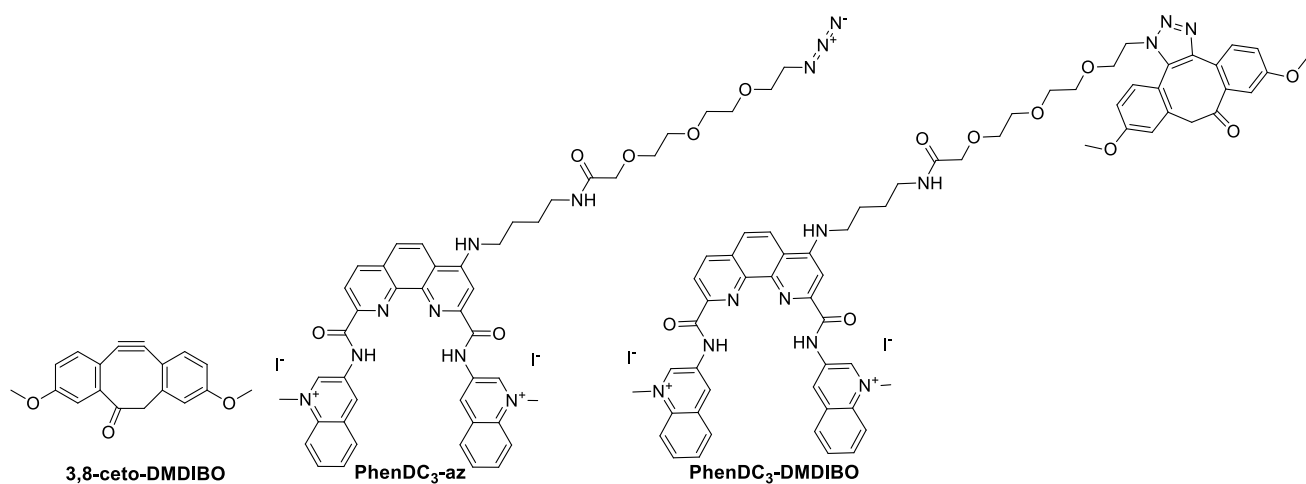


Figure 199 : Ligands et fluorophore utilisés pour mesurer les temps de vie de fluorescence.

Spectres UV

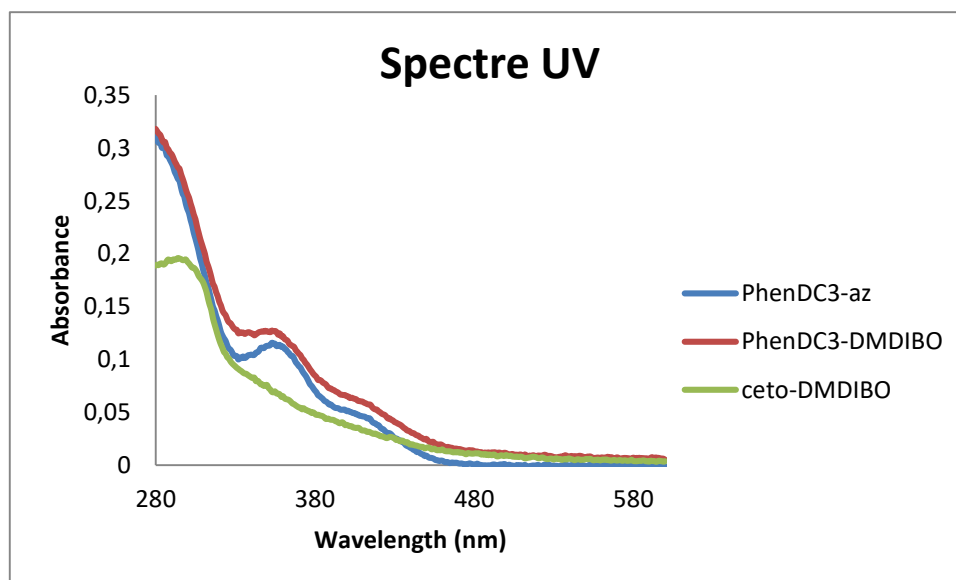


Figure 200 : Spectres UV des composés utilisés pour la mesure de temps de vie de fluorescence ; PhenDC3-az (10 μ M), PhenDC3-DMDIBO (10 μ M) et le céto-DMDIBO (20 μ M).

D'après les spectres UV des trois différents composés (figure 200), on peut constater que la valeur d'absorbance n'est pas nulle à 370 nm qui est la valeur du laser du montage utilisé pour mesurer le temps de vie de fluorescence.

Spectres de fluorescence

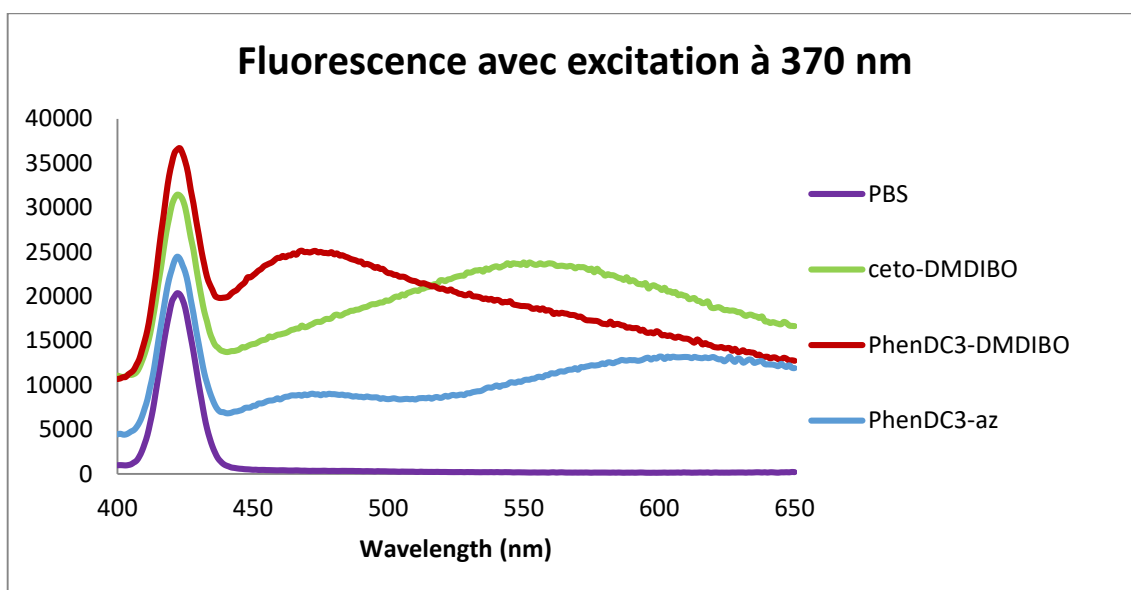


Figure 201 : Spectres de fluorescence des composés avec une excitation à 370 nm dans le PBS.

Malgré une excitation à 370 nm, nous pouvons constater que les spectres d'émission de fluorescence sont différents entre le fluorophore seul et cliqué (figure 201). Le pic à 420 nm est lié au RAMAN de l'eau. Le composé **PhenDC₃-az** est donc très faiblement fluorescent, cela ne nous permet pas de le suivre par microscopie de fluorescence. *A contrario*, une fois conjugué **PhenDC₃-DMDIBO** est suffisamment fluorescent pour être observé au microscope, le principal inconvénient est son excitation à 370 nm. De plus, à cette longueur d'onde, l'auto-fluorescence des cellules peut être un souci dans la bonne visualisation du composé. La toxicité du laser pour les cellules lors d'une analyse prolongée peut également engendrer des soucis de visualisation. C'est pourquoi le développement de cyclooctynes fluorescentes ayant une longueur d'onde d'excitation plus élevée peut être intéressant.

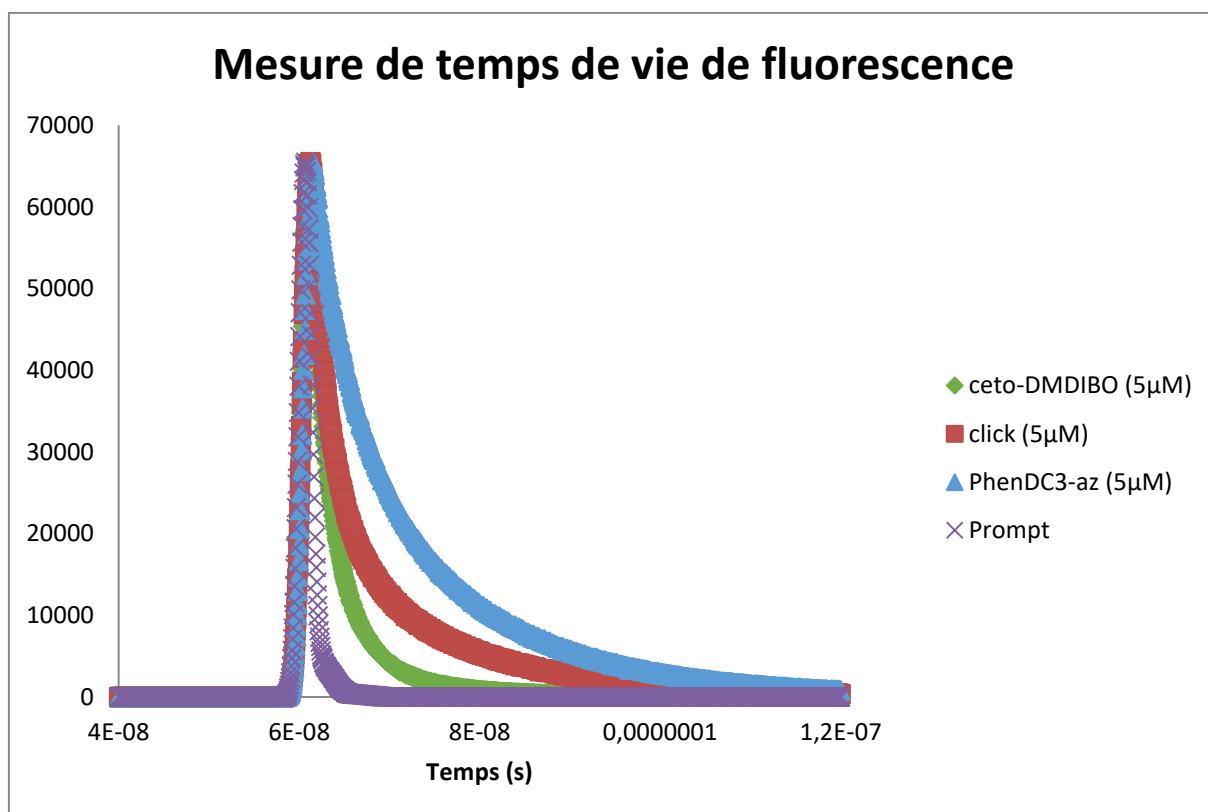


Figure 202 : Représentation des déclin de fluorescence pour mesurer le temps de vie de fluorescence.

La représentation des déclin (figure 202) nous montre bien que le temps de vie de fluorescence de la molécule seule ou conjuguée n'est pas la même. Ce point est très important car il permettrait dans un milieu biologique plus complexe de savoir si la réaction SPAAC a bien eu lieu. Le temps de vie est le temps moyen pour lequel un fluorophore reste dans son état excité après l'excitation. Cela peut donc être noté comme le temps moyen dans l'état excité. En utilisant une formule simplifiée pour calculer le temps de vie moyen de fluorescence d'après le livre de Lakowicz,^[219] on obtient les valeurs suivantes :

$$\text{Durée de vie moyenne} = \frac{\int t \cdot I(t) \cdot dt}{\int I(t) \cdot dt}$$

Déclin	Durée de vie moyenne (ns)
Ceto-DMDIBO	4,34
PhenDC ₃ -DMDIBO	10,34
PhenDC ₃ -az	12,15

Figure 203 : Temps de vie des composés cliquables et cliqué.

On constate une grande différence de temps de vie de fluorescence entre le ceto-DMDIBO « libre » et le composé conjugué au ligand de G-quadruplex (figure 203). La différence est moins grande entre PhenDC₃-az et le composé cliqué mais il y a tout de même une différence de deux nanosecondes ce qui est détectable par mesure de temps de vie de fluorescence en microscopie. Le deuxième point

important à regarder est l'influence de l'environnement sur la fluorescence des composés. En effet, en présence d'ADN G-quadruplex ou d'ADN duplex, le temps de vie de fluorescence peut être modifié. Nous avons alors utilisé différentes séquences d'ADN en présence de ligand et mesuré les temps de vie de fluorescence (Figure 57). Les séquences d'ADN utilisées sont le 22AG qui est la séquence télomérique), le CEB25-WT (L191) et CEB25-L111(T) qui sont issues du minisatellite CEB25, le Pu27 qui est la séquence native de l'oncogène c-myc et le ds26 qui est un ADN duplex.

Déclin	Durée de vie moyenne (ns)
PhenDC3-az	12,15
PhenDC3-az_22AG	9,54
PhenDC3-az_CEB25111	9,82
PhenDC3-az_CEB25WT	8,46
PhenDC3-az_DS26	12,19

Déclin	Durée de vie moyenne (ns)
Ceto-DMDIBO	4,34
CetoDMDIBO_22AG	4,27
CetoDMDIBO_CEB25111	4,59
CetoDMDIBO_CEB25WT	4,26
CetoDMDIBO_DS26	4,69

Déclin	Durée de vie moyenne (ns)
PhenDC3-DMDIBO	10,34
PhenDC3-DMDIBO_22AG	8,98
PhenDC3-DMDIBO_CEB25111	9,82
PhenDC3-DMDIBO_CEB25WT	9,22
PhenDC3-DMDIBO_PU27	9,21
PhenDC3-DMDIBO_DS26	10,53

Figure 204 : Temps de vie de fluorescence en présence d'ADN.

Si nous prenons les temps de vie de fluorescence ligand par ligand (figure 204), on peut constater tout d'abord que le temps de vie de PhenDC₃-az n'est pas modifié en présence de ds26. En revanche, son temps de vie en présence d'ADN G-quadruplex est plus court ce qui montre bien l'interaction du ligand pour les G-quadruplex et sa sélectivité vis-à-vis de l'ADN duplex. Ensuite, la présence d'ADN, qu'il soit duplex ou quadruplex, ne modifie en rien le temps de vie de fluorescence du fluorophore seul. Enfin, le composé cliqué est également influencé par la présence de G-quadruplex avec un temps de vie de 8 à 9 nanosecondes alors que le composé libre ou en présence d'ADN duplex a un temps de vie de fluorescence supérieur à 10 nanosecondes. Ces résultats préliminaires nous permettent d'être optimistes pour la suite du projet qui devrait se développer *in cellulo*.

Conclusion générale et perspectives

Les travaux réalisés pendant ces trois ans de doctorat avaient pour objectif principal l'étude des G-quadruplex au sein de milieux biologiques complexes. Pour cela, la synthèse d'outils fonctionnels a permis d'envisager différentes stratégies. Ces outils ont été conçus à partir de ligands très bons stabilisants de G-quadruplex, PhenDC₃ et PDC (figure 205). Des ligands fonctionnalisés ont été préparés : les molécules possèdent soit un agent alkylant, soit une biotine, soit les deux fonctions. L'agent alkylant permet de créer un lien covalent entre la molécule et la cible alors que la biotine permet d'immobiliser des structures G-quadruplex sur des surfaces par interactions entre la biotine et la streptavidine qui peut être sur des billes, des colonnes d'affinité ou encore sur des puces SPR.

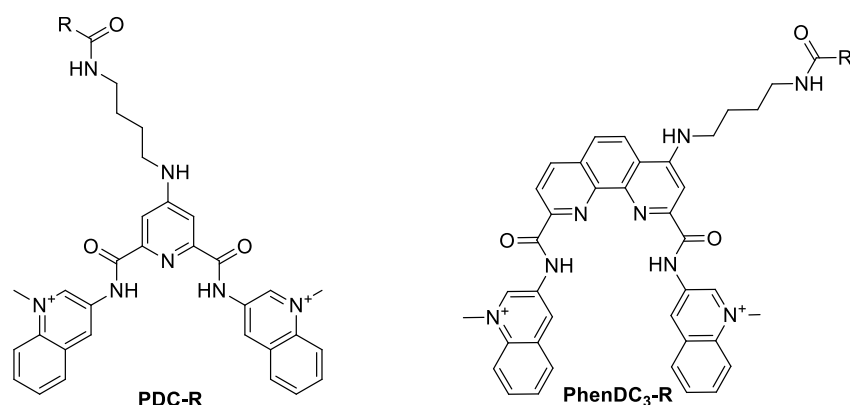


Figure 205 : Structure des dérivés PDC et PhenDC₃.

D'autre part, des ligands à fonctionnaliser *in situ* ont également été synthétisés. Les molécules possèdent une fonction alcyne ou une fonction azoture afin de réagir par cycloaddition catalysée par le cuivre entre un alcyne et un azoture (CuAAC) ou par cycloaddition d'un alcyne contraint et d'un azoture (SPAAC). Ces réactions bioorthogonales permettent d'ajouter une fonction aux ligands une fois qu'ils sont fixés sur leur cible.

Tous ces composés ont ensuite été analysés par deux méthodes biophysiques, l'expérience de FRET-melting et l'expérience de FID. Ces deux techniques nous ont permis de montrer l'influence d'un ligand sur la stabilisation d'une structure G-quadruplex, d'apporter des informations sur le mode de fixation d'un ligand mais également de montrer l'effet de la fonctionnalisation sur les composés PhenDC₃ et PDC. Les dérivés de PDC sont beaucoup plus robustes et leur affinité n'est que très peu modifiée alors que les dérivés de PhenDC₃ sont très fortement affectés et perdent en affinité dès qu'une fonction encombrante est ajoutée. Les dérivés fonctionnalisables *in situ* du PDC et du PhenDC₃ conservent une très bonne affinité pour toutes les structures G-quadruplex étudiées. Tous ces composés ont un point commun : ils conservent la sélectivité pour les structures G-quadruplex

vis-à-vis des structures duplex. Ces deux méthodes d'analyse utilisées en routine nous ont également permis de faire un classement de certains ligands de référence. Ainsi, PhenDC₃ s'est révélé être le meilleur ligand sur toutes les structures G-quadruplex étudiées et PDS a montré une très bonne affinité pour les G4 mais une question se pose par rapport à son mode de fixation car il ne déplace que peu la sonde fluorescente. C'est pourquoi il est important de combiner les deux expériences afin d'avoir le maximum d'informations sur l'interaction ligand/G-quadruplex. L'expérience de FRET-melting permet d'avoir des informations sur la stabilisation de la structure alors que l'expérience de FID apporte des informations sur le mode de fixation du ligand en déplaçant la sonde fluorescente. Nous avons alors mis en place une valeur d'index qui associe les deux techniques et qui permet de classer tous les ligands les uns par rapport aux autres. De plus, même s'il faudrait avoir plus de données, nous avons essayé de relier les valeurs de cet index avec des constantes de dissociation mesurées par SPR. Cela permet de classer tout ligand nouvellement synthétisé.

Par ailleurs, l'expérience de FRET-melting a également été utilisée afin d'établir une corrélation entre la valeur de température de fusion en présence d'un ligand (*in vitro*) et les mesures d'instabilité génétique au sein des levures induite par ce même ligand (*in vivo*). Ce travail de recherche effectué en collaboration avec l'équipe d'Alain Nicolas a montré l'existence d'une valeur de température de fusion à partir de laquelle un ligand induit de l'instabilité génétique dans les levures. Cette valeur nécessite d'être confirmée par d'autres expériences notamment sur des séquences du minisatellite CEB1.

Ensuite, les ligands modifiables par réactions bioorthogonales ont été utilisés au sein de cellules cancéreuses humaines (figure 206). Après avoir vérifié la bonne réactivité des composés, nous avons pu localiser les différentes molécules dans le noyau et le cytoplasme des cellules. Cela prouve la bonne pénétration des différentes membranes cellulaires. L'imagerie en cellules fixées a permis de montrer un biais important de la réaction CuAAC et plus précisément l'influence du cuivre (I) sur la localisation des ligands de G-quadruplex. En effet, le cuivre se coordine aux alcynes et dirige les molécules dans les nucléoles des cellules. L'utilisation de la réaction SPAAC a permis de montrer les composés dans le nucléoplasme aussi bien en cellules fixées qu'en cellules vivantes. Nous avons également pu constater la difficulté de travailler en cellules vivantes. En effet, les fluorophores fonctionnalisés par un groupement DBCO ont des difficultés de pénétration au sein des cellules et la colocalisation des ligands avec des marqueurs de noyau par exemple est très délicate. L'ajout de ces marqueurs déplace les ligands et ces derniers ne sont plus observables en imagerie. Il serait donc intéressant d'avoir une alternative à cette colocalisation en utilisant la technique de mesure de temps de vie de fluorescence. Cette méthode permettrait de savoir où se situe la molécule qui a réagi avec le fluorophore et si elle est liée à une structure d'ADN ou d'ARN particulier. Nous avons

effectué quelques tests préliminaires *in vitro* en utilisant une cyclooctyne fluorescente et montré que le temps de vie de fluorescence est différent lorsque le fluorophore est libre ou « clické » avec le ligand de G4. De plus, ce temps de vie est également modifié lorsqu'il est en présence de G-quadruplex alors qu'il est le même en présence de duplex. Par la suite, il faudrait effectuer les tests en présence du fluorophore TAMRA-DBCO qui n'ont pas été fait car nous n'avions pas le laser adéquat. Si la tendance se confirme, les mesures de temps de vie de fluorescence *in cellulo* seront à réaliser. Cela permettrait de savoir si la localisation des ligands est due à un effet G-quadruplex. D'autres expériences ont été envisagées au sein des cellules. D'une part, la localisation de ligands sur des chromosomes en métaphase pourrait être étudiée afin de savoir sur quelle zone du chromosome se situent les ligands. D'autre part, l'introduction de séquences G-quadruplex marquées par un fluorophore pourrait également être explorée et la colocalisation des ligands avec ces séquences montrerait l'effet G-quadruplex sur la localisation des composés. Cela pourrait également montrer la formation de structures G-quadruplex dans un milieu cellulaire.

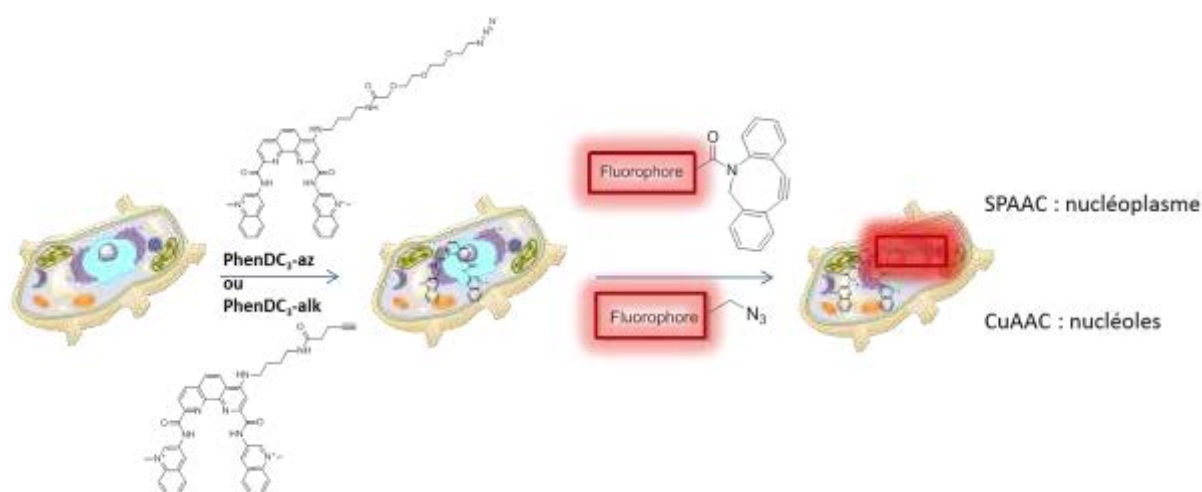


Figure 206 : Imagerie cellulaire des ligands de G-quadruplex.

Enfin, une autre application a été étudiée de manière très préliminaire, c'est pourquoi cela n'est pas détaillé dans ce manuscrit. Il s'agit de l'immobilisation de structures G-quadruplex avec des billes magnétiques. Ces billes peuvent être recouvertes de streptavidine et ainsi interagir avec un ligand biotinylé ou recouvertes de fonctions « clickables » telles qu'un azoture ou une cyclooctyne pour réagir avec la fonction complémentaire par réaction de CuAAC ou SPAAC. L'ADN utilisé dans ces expériences est radiomarqué au phosphore 32 (^{32}P) afin de pouvoir compter la radioactivité et estimer la quantité d'ADN extrait. L'utilisation de ligands bi-fonctionnels biotinylés pour l'immobilisation de G-quadruplex sur des billes recouvertes de streptavidine a été très explorée lors de la thèse d'Elodie Morel avec des résultats très encourageants. En effet, le système qui a été mis au point permet d'immobiliser plus de 60 % des structures G-quadruplex sur les billes (figure 207). Nous

avons alors envisagé d'améliorer cette expérience en utilisant les molécules « clickables » qui ont une meilleure affinité pour les G-quadruplex et qui doivent permettre une meilleure immobilisation de ces structures. Le principe est de mettre le ligand en présence des G-quadruplex puis ajouter la biotine par réaction SPAAC ou CuAAC (figure 207). Plusieurs modes opératoires ont été testés mais aucun n'a été concluant. Nous avons essayé de pré-saturer les billes avec la molécule biotine-DBCO avant de réagir avec le ligand. Nous avons également testé en faisant réagir le ligand et la biotine avant d'ajouter les billes. Cependant, le système ne permet pas d'immobiliser plus de 20 % de G-quadruplex. Il serait nécessaire d'optimiser le système en travaillant sur les quantités de ligands, de biotine et de billes pour le rendre fonctionnel. Malheureusement, par manque de temps, cette optimisation n'a pas été faite. Par ailleurs, des billes recouvertes de fonctions DBCO ont également été utilisées afin de créer un lien covalent entre les billes et le ligand cependant ce système (non optimisé) n'a pas permis d'extraire plus d'ADN qu'avec les ligands bi-fonctionnels.

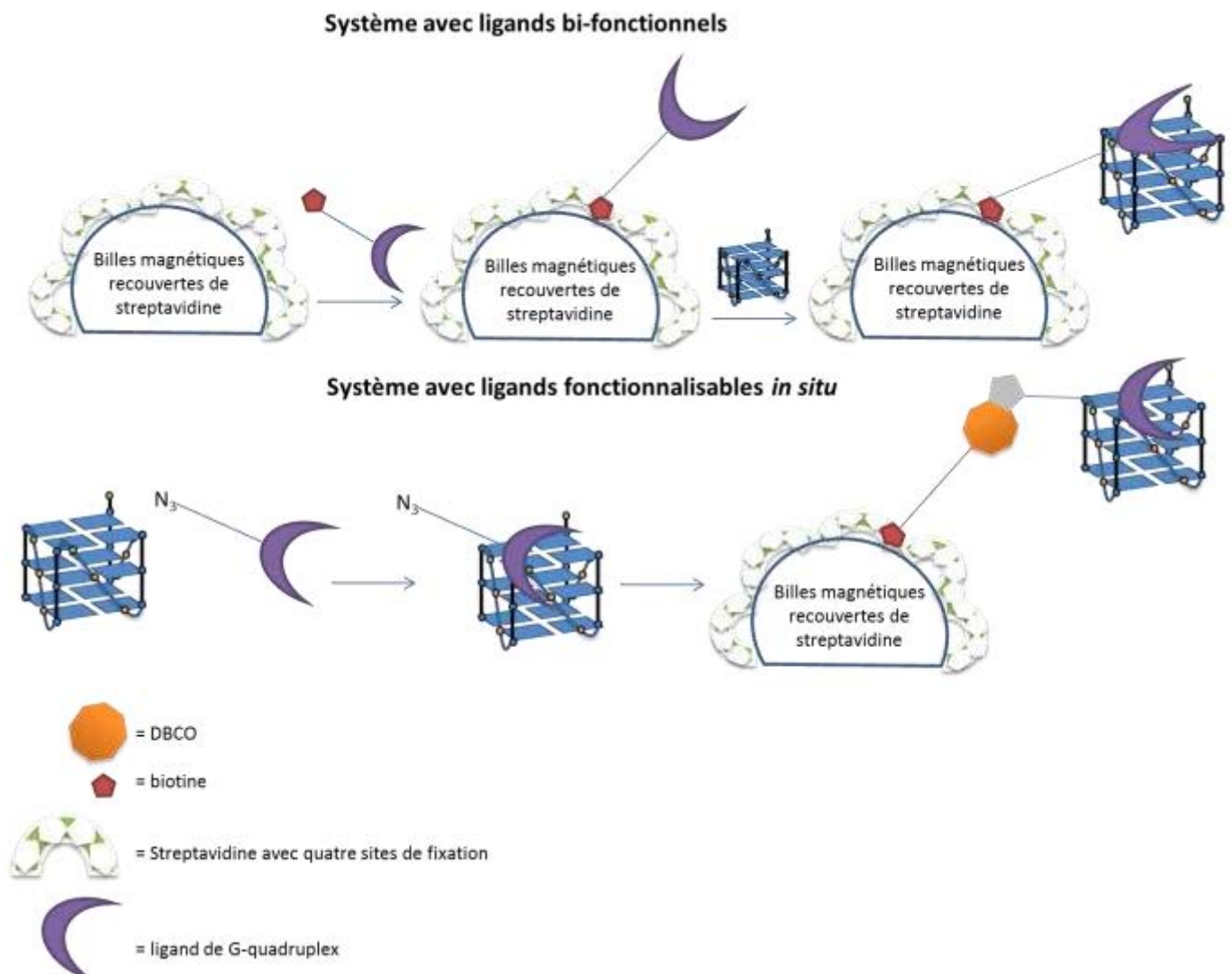


Figure 207 : Stratégies d'immobilisation de G-quadruplex sur des billes magnétiques.

Toutes ces expériences ont montré le fort potentiel ainsi que les limitations de l'utilisation de la chimie « click ». L'utilisation de la chimie bioorthogonale n'améliore pas toujours le système comme on a pu le voir brièvement avec les expériences d'extraction supportée de G-quadruplex. Cependant, elle permet de conserver des structures proches de la molécule de référence ce qui garde l'affinité et la sélectivité des nouveaux ligands pour la cible. Plusieurs applications sont compatibles avec cette stratégie et notamment les expériences *in cellulose*. Il est très important, lors de ces expériences, de faire tous les contrôles afin de s'assurer que la localisation de molécules n'est pas due à un biais de la manipulation.

Chapitre 4 : Partie expérimentale

I. Synthèses

Cette section est rédigée en anglais afin de faciliter la rédaction de publications.

1) Matériel et méthodes

^1H NMR spectra were recorded at 25°C on a Bruker Avance 300 using TMS as internal standard. Deuterated CDCl_3 and $[\text{D}_6]\text{DMSO}$ were purchased from SDS. The following abbreviations are used: singlet (s), doublet (d), triplet (t) and multiplet (m). Low resolution mass spectrometry (ESI-MS) was recorded on a micromass ZQ 2000 (Waters). High resolution mass spectrometry (ESI-MS) and ^{13}C NMR spectra were provided by I.C.S.N. (Institut de Chimie des Substances Naturelles, Gif-sur-Yvette). Reagents and chemicals were purchased from Sigma-aldrich, Acros, TCI or Alfa-aesar.

2) Modes opératoires

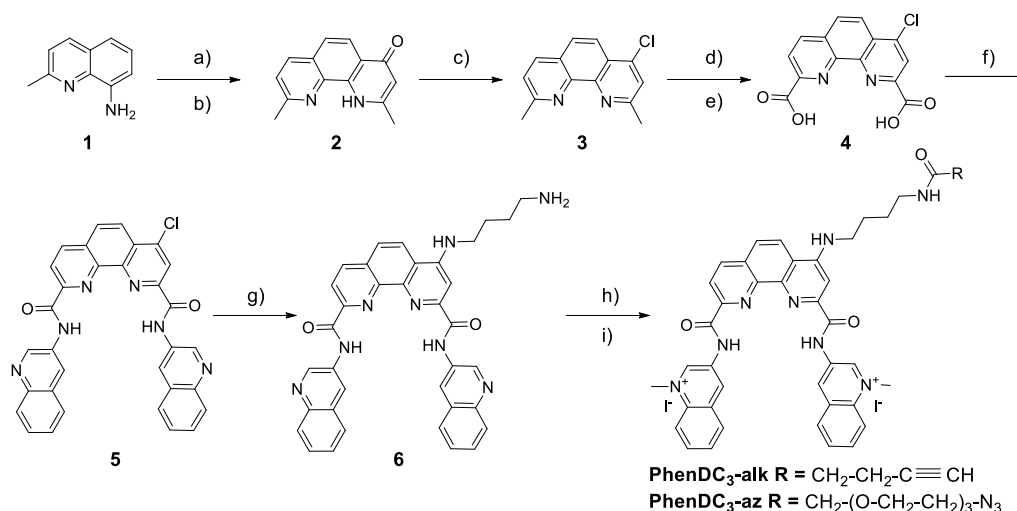
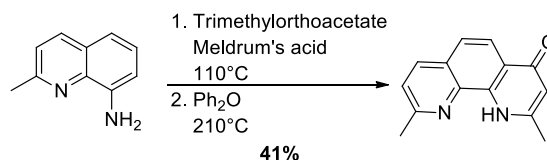


Figure 208 : Synthesis of “Clickable” derivatives of PhenDC₃. a) Trimethyl orthoacetate, Meldrum’s acid, 110 °C, r.t. 18 h; b) Diphenylether, 230 °C, 1 h; c) POCl_3 90 °C; d) NCS, benzoic acid, DCE, 70 °C, 24 h; e) H_2SO_4 , 95 °C, 3 h; f) 3-aminoquinoline, EDCI, HOBT, DMF, 0 °C, r.t. 18 h; g) 1,4-diaminobutane, 100 °C, 30 min, MW; h) for PhenDC₃-alk: 4-pentynoic acid, EDCI, Et₃N, HOBT, DMF, r.t., 18 h; for PhenDC₃-az: 2-[2-[2-(2-azidoethoxy)ethoxy]ethoxy]acetic acid, EDCI, Et₃N, HOBT, DMF, r.t., 18 h; i) CH_3I , DMF, 40 °C, 24 h.



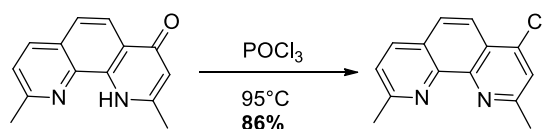
2,9-dimethyl-1,4-dihydro-1,10-phenantrolin-4-one

Trimethylorthoacetate (24.2 mL, 189 mmol, 30 eq) and Meldrum's acid (1.37g, 9.48 mmol, 1.5eq) were brought to a gentle reflux for 15 min, at 110 °C. The resulting yellow solution was cooled down and 2-methylquinolin-8-amine (1.0 g, 6.32 mmol, 1.0 eq) was added. The reaction mixture was heated at reflux for 2 h and left under stirring at room temperature for 16 h. The solvent was removed under vacuum and the red oil obtained was heated at 230 °C for 1 h in diphenylether. After cooling down to 50 °C, petroleum ether was added and a dark red powder was obtained after filtration. The product was purified by column chromatography (DCM/ethanol 90/10) affording a dark red solid (585 mg, 41%).

¹H NMR (300 MHz, CDCl₃): δ (ppm) 10.15 (m, 1H), 8.29 (d, *J* = 9.0 Hz, 1H), 8.12 (d, *J* = 8.5 Hz, 1H), 7.58 (d, *J* = 9.0 Hz, 1H), 7.47 (d, *J* = 8.5 Hz, 1H), 6.36 (s, 1H), 2.80 (s, 3H), 2.56 (s, 3H).

¹³C NMR (75 MHz, CDCl₃): δ (ppm) 178.6, 158.1, 147.2, 137.9, 136.4, 136.3, 127.0, 124.3, 123.3, 122.1, 121.4, 112.4, 25.0, 20.4.

LR-MS (ESI-MS) *m/z* = 224.9 [M+H]⁺



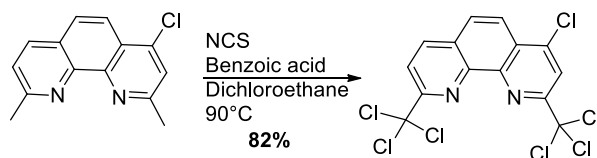
4-chloro-2,9-dimethyl-1,10-phenantrolin-4-one

In a round-bottomed flask, phosphoryl chloride (8 mL, 85.8 mmol, 43 eq) was stirred under argon atmosphere before addition of 2,9-dimethyl-1,4-dihydro-1,10-phenantrolin-4-one (450 mg, 2.0 mmol, 1.0 eq). The reaction mixture was stirred at 90 °C for 3 h 30. While still hot, the reaction mixture was carefully added to ice (30 g) and water (10 mL). After 15 min stirring, chloroform (20 mL) was added and the resulting two-layer system was brought to pH 13-14 thanks to NaOH solution addition. The aqueous layer was extracted by DCM (4 x 100 mL). The combined organic layers were washed by NaOH solution, dried over MgSO₄ and concentrated under vacuum to afford the desired product (420 mg, 86 %) as light tan crystals.

¹H NMR (300 MHz, CDCl₃): δ (ppm) 8.17 (d, *J* = 9.0 Hz, 1H), 8.16 (d, *J* = 8.5 Hz, 1H), 7.81 (d, *J* = 9.0 Hz, 1H), 7.61 (s, 1H), 7.54 (d, *J* = 8.5 Hz, 1H), 2.96 (s, 3H), 2.93 (s, 3H).

¹³C NMR (75 MHz, CDCl₃): δ (ppm) 160.0, 159.2, 146.4, 144.9, 142.6, 136.4, 126.9, 126.4, 124.1, 123.6, 121.1, 25.9, 25.7.

LR-MS (ESI-MS) m/z = 265.2 [M+Na]⁺



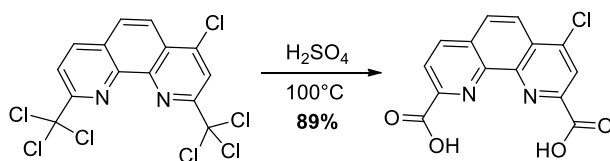
4-chloro-2,9-bis(trichloromethyl)-1,10-phenantroline

4-chloro-2,9-dimethyl-1,10-phenantroline (260 mg, 1.07 mmol, 1.0 eq) was dissolved in dichloroethane (20 mL) before addition of N-chlorosuccinimide (1.03 g, 7.71 mmol, 7.2 eq) and benzoic acid (3 mg, 0.02 mmol, 0.02 eq). The reaction mixture was heated at reflux, at 70 °C for one day. After cooling down to room temperature, the mixture was filtered and the filtrate was concentrated under vacuum. The product was taken up in DCM (50 mL) and washed by saturated aqueous Na₂CO₃ solution three times. The combined organic layers were dried over MgSO₄ and concentrated under vacuum. The desired product (393 mg, 82 %) was obtained after purification by column chromatography (cyclohexane/DCM 60/40).

¹H NMR (300 MHz, CDCl₃): δ (ppm) 8.50-8.35 (m, 4H), 8.07 (d, *J* = 9.0 Hz, 1H).

¹³C NMR (75 MHz, DMSO-*d*⁶): δ (ppm) 158.5, 157.9, 144.5, 144.4, 143.1, 138.4, 129.3, 128.6, 127.5, 123.7, 121.2, 120.8, 98.1, 97.3.

LR-MS (ESI-MS) m/z = 446.9 [M+H]⁺



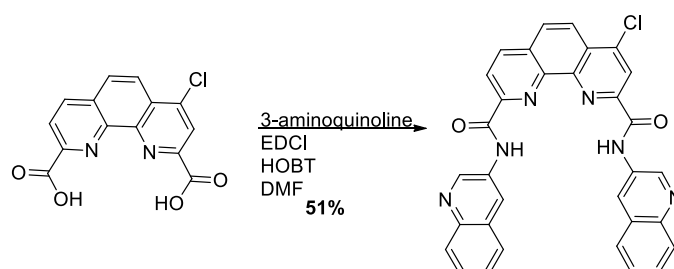
4-chloro-1,10-phenantroline-2,9-dicarboxylic acid

Sulfuric acid (1 mL) and 4-chloro-2,9-bis(trichloromethyl)-1,10-phenantroline (200 mg, 0.44 mmol, 1.0 eq) are introduced in a round-bottomed flask and stirred at 95 °C for 3 h. The reaction mixture was then cooled down to room temperature, before careful addition of water (300 μL). The resulting solution was refluxed for 1 h and a large amount of water was added. The precipitate was filtered, washed by water and dried by diethylether, affording the desired compound (120 mg, 89 %).

¹H NMR (300 MHz, DMSO-*d*⁶): δ (ppm) 8.79 (d, *J* = 8.4 Hz, 1H), 8.50 (s, 1H), 8.44 (d, *J* = 8.4 Hz, 1H), 8.41 (d, *J* = 9.0 Hz, 1H), 8.36 (d, *J* = 9.0 Hz, 1H).

^{13}C NMR (75 MHz, DMSO- d_6): δ (ppm) 166.1, 165.3, 149.0, 148.4, 145.9, 144.4, 142.9, 138.4, 130.3, 130.0, 127.8, 124.0, 123.5, 123.2.

LR-MS (ESI-MS) $m/z = 303.1$ $[\text{M}+\text{H}]^+$

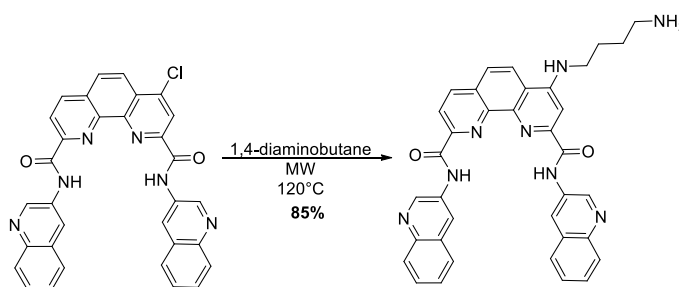


PhenDC₃n-Cl

3-aminoquinoline (440 mg, 3.05 mmol, 2.15 eq) and HOBT (394 mg, 2.91 mmol, 2 eq) were successively added to a solution of 4-chloro-1,10-phenanthroline-2,9-dicarboxylic acid (430 mg, 1.42 mmol, 1.0 eq) suspended in DMF (25 mL), under an argon atmosphere. The reaction mixture was then cooled to 0°C (ice bath) before addition of EDCI (495 mg, 3.19 mmol, 2.2 eq). The reaction mixture was slowly allowed to reach room temperature over 1 h and was stirred overnight. The pale yellow precipitate obtained was filtrated, washed with water and diethyl ether to afford PhenDC₃n-Cl (103 mg, 51 %).

^1H NMR (300 MHz, DMSO- d_6 at 80°C): δ (ppm) 11.66 (s, 2H), 9.64 (s, 2H), 9.03 (m, 2H), 8.89 (d, $J = 8.5$ Hz, 1H), 8.69 (m, 2H), 8.43 (q, $J = 9.0$ Hz, 2H), 8.07 (t, $J = 9.0$ Hz, 4H), 7.74 (t, $J = 7.2$ Hz, 2H), 7.65 (t, $J = 7.2$ Hz, 2H).

LR-MS (ESI-MS) $m/z = 555.5$ $[\text{M}+\text{H}]^+$



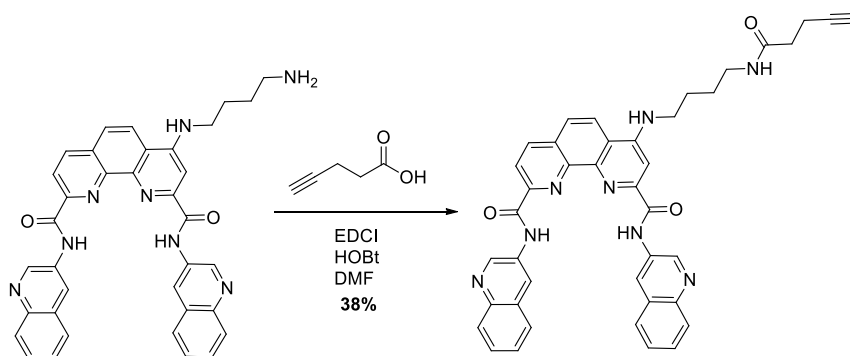
PhenDC₃n-NH₂

A sealed vial containing PhenDC₃n-Cl (100 mg, 554 mmol, 1.0 eq) and butane-1,4-diamine (1.70 mL, 88 mmol, 100 eq) under argon was heated up to 120°C for 1h in microwave to give an homogeneous clear orange mixture. The expected compound was precipitated upon addition of acetonitrile (2-5

mL). The solid was then filtered, washed with acetonitrile and diethyl ether to afford PhenDC₃n-NH₂ (93 mg, 85 %).

¹H NMR (300 MHz, DMSO-d₆): δ (ppm) 11.79 (bs, 1H), 9.67 (s, 2H), 9.13 (d, J = 9.3 Hz, 2H), 8.79 (d, J = 8.3 Hz, 1H), 8.59 (d, J = 8.3 Hz, 1H), 8.55 (d, J = 9.3 Hz, 1H), 8.09 (m, 5H), 7.92 (bs, 1H), 7.69 (m, 5H), 3.47 (m, 2H), 2.69 (m, 2H), 1.81 (m, 2H), 1.58 (m, 2H)

LR-MS (ESI-MS) m/z= 607 [M+H]⁺

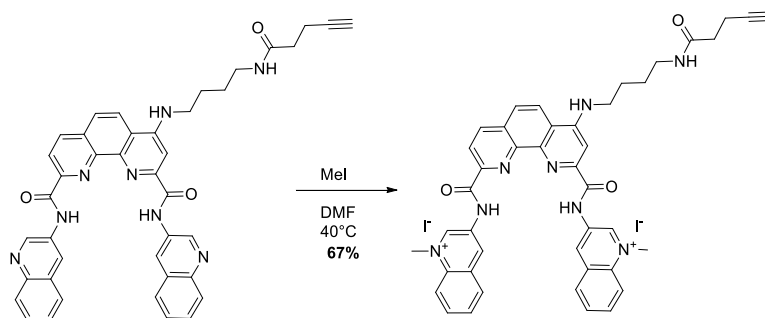


PhenDC₃n-alk

4-pentynoic acid (0.1 mmol, 1.3 eq) was dissolved in DMF (3 mL) with PhenDC₃n-NH₂ (0.08 mmol, 1 eq), HOBt (0.04 mmol, 0.5 eq) and EDCI (0.1 mmol, 1.3 eq). Triethylamine (0.1 mmol, 1.3 eq) was added and the reaction mixture was stirred overnight at room temperature, with protection from light. DMF was then removed under vacuum and the crude mixture was purified by flash chromatography (DCM/ethanol/TEA 90/9/1) to afford the PhenDC₃n-alk compound as a yellow powder (38 %).

¹H NMR (300 MHz, DMSO-d₆): δ (ppm) 11.82 (s, 1H), 11.79 (s, 1H), 9.67 (s, 2H), 9.13 (d, J = 9.0 Hz, 1H), 8.77 (d, J = 8.5 Hz, 1H), 8.59 (d, J = 8.5 Hz, 1H), 8.56 (d, J = 9.0 Hz, 1H), 8.10 (m, 5H), 7.97 (m, 1H), 7.72 (m, 5H), 3.50 (m, 2H), 3.18 (m, 3H), 2.75 (m, 1H), 2.29 (m, 3H), 1.80 (m, 2H), 1.62 (m, 2H).

LR-MS (ESI-MS) m/z= 688.16 [M+H]⁺



PhenDC₃-alk

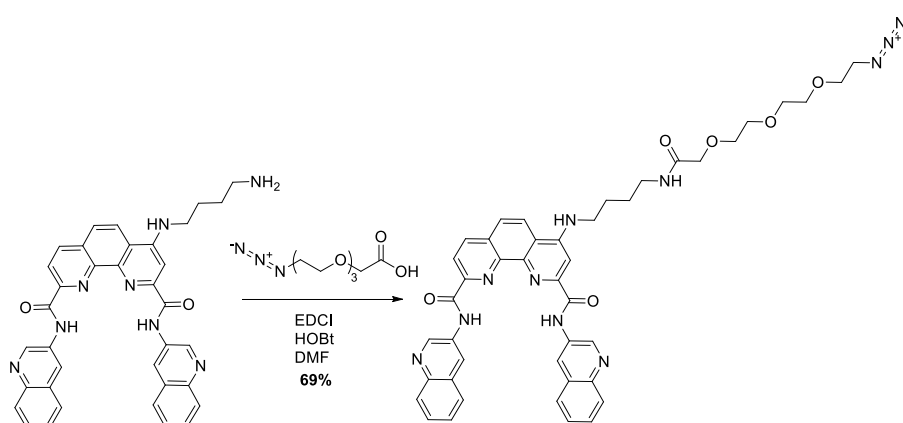
A solution of the starting compound PhenDC₃n-alk (36 mg, 0.07 mmol, 1.0 eq) in DMF (2 mL) was heated at 40 °C before methyl iodide (104 μL, 1.67 mmol, 45 eq) was added dropwise. The reaction was performed under argon atmosphere for 24h. The solvent was evaporated to dryness and the expected compound was taken from ethanol as a yellow solid (67%).

¹H NMR (300 MHz, DMSO-d₆): δ (ppm) 12.12 (s, 2H), 10.30 (s, 2H), 9.89 (s, 2H), 8.88 (d, J = 8.5 Hz, 1H), 8.68 (d, J = 8.5 Hz, 1H), 8.65-8.45 (m, 5H), 8.30-8.20 (m, 3H), 8.08 (m, 2H), 7.98 (m, 1H), 7.69 (s, 1H), 4.73 (s, 6H), 3.48 (under water, 2H), 3.18 (m, 2H), 2.76 (s, 1H), 2.37 (m, 2H), 2.29 (m, 2H), 1.80 (m, 2H), 1.62 (m, 2H).

¹³C NMR (500MHz, DMSO-d₆): δ (ppm) 170.2, 164.4, 163.7, 152.3, 148.5, 147.9, 145.8, 145.7, 144.2, 138.8, 135.5, 135.4, 134.7, 134.4, 133.8, 133.7, 132.8, 132.8, 130.9, 130.3, 130.2, 129.9, 129.9, 129.1, 125.2, 122.9, 121.5, 119.5, 119.2, 99.5, 83.8, 71.3, 46.0, 45.8, 42.5, 38.2, 34.3, 27.0, 25.1, 14.3

LR-MS (ESI-MS) m/z = 358.1 [(M-2I)/2+H]⁺

HR-MS (ESI+) m/z = 843.2268 calculated for C₄₃H₄₀N₈O₃I; found, 843.2266



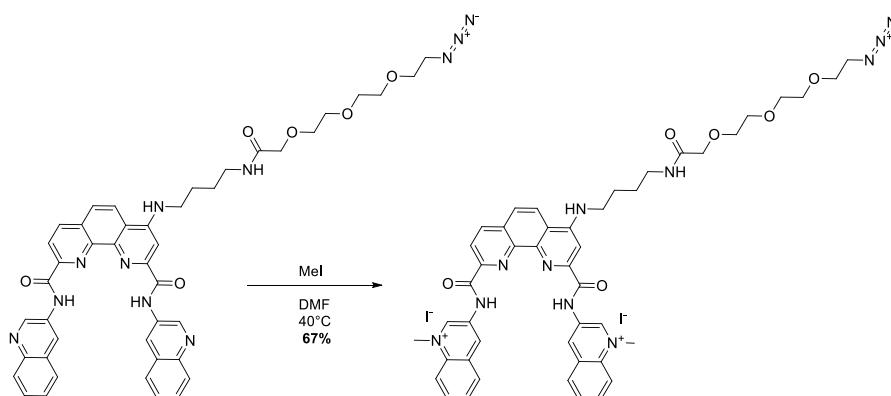
PhenDC₃n-az

2-[2-[2-(2-azidoethoxy)ethoxy]ethoxy]acetic acid (0.06 mmol, 1.2 eq) was dissolved in DMF (3 mL) with PhenDC₃n-NH₂ (0.05 mmol, 1 eq), HOBt (0.02 mmol, 0.5 eq) and EDCI (0.07 mmol, 1.5 eq).

Triethylamine (0.07 mmol, 1.5 eq) was added and the reaction mixture was stirred overnight at room temperature, with protection from light. DMF was then removed under vacuum and the crude mixture was purified by flash chromatography (DCM/ethanol/TEA 90/9/1) to afford the PhenDC₃n-azide compound as a yellow powder (69 %).

¹H NMR (300 MHz, DMSO-d₆): δ (ppm) 11.83 (s, 1H), 11.79 (s, 1H), 9.68 (t, J = 2.2 Hz, 2H), 9.13 (dd, J = 2.2 Hz and J = 9.0 Hz, 2H), 8.79 (d, J = 8.5 Hz, 1H), 8.60 (d, J = 8.5 Hz, 1H), 8.56 (d, J = 9.0 Hz, 1H), 8.11 (m, 5H), 7.70 (m, 6H), 7.64 (s, 1H), 3.90 (s, 2H), 3.55 (m, 10H), 3.50 (m, 2H), 3.26 (m, 3H), 1.79 (m, 2H), 1.67 (m, 2H).

LR-MS (ESI-MS) m/z = 823.4 [M+H]⁺



PhenDC₃-az

A solution of the starting compound PhenDC₃n-az (20 mg, 0.02 mmol, 1.0 eq) in DMF (2 mL) was heated at 40 °C before methyl iodide (151 μL, 2.43 mmol, 100 eq) was added dropwise. The reaction was performed under argon atmosphere for 24h. The solvent was evaporated to dryness and the expected compound was taken from ethanol as a yellow solid (67%).

¹H NMR (300 MHz, DMSO-d₆): δ (ppm) 12.11 (s, 2H), 10.30 (s, 2H), 9.88 (s, 2H), 8.88 (d, J = 8.3 Hz, 1H), 8.69 (d, J = 8.3 Hz, 1H), 8.62 (d, J = 9.2 Hz, 1H), 8.53 (m, 4H), 8.25 (m, 2H), 8.21 (d, J = 9.0 Hz, 1H), 8.08 (m, 2H), 7.95 (s, 1H), 7.76 (t, J = 5.84 Hz, 1H), 7.70 (s, 1H), 4.73 (s, 3H), 4.72 (s, 3H), 3.89 (s, 2H), 3.58 (m, 6H), 3.55 (m, 4H), 3.37 (m, 2H), 3.25 (m, 2H), 3.11 (m, 1H), 1.80 (m, 2H), 1.67 (m, 2H), 1.18 (m, 1H).

¹³C NMR (500MHz, DMSO-d₆): δ (ppm) 169.2, 163.7, 147.9, 145.7, 145.7, 145.6, 138.8, 135.5, 135.4, 134.7, 134.5, 133.9, 133.8, 132.8, 130.9, 130.2, 129.9, 129.9, 129.1, 122.9, 121.6, 119.2, 99.6, 70.2, 70.0, 69.7, 69.6, 69.5, 69.2, 49.9, 46.0, 42.6, 37.7, 27.0, 25.1

LR-MS (ESI-MS) m/z = 425.7 [M-2I/2+H]⁺

HR-MS (ESI+) m/z = 978.2912 calculated for C₄₆H₄₉N₁₁O₆; found, 978.2913

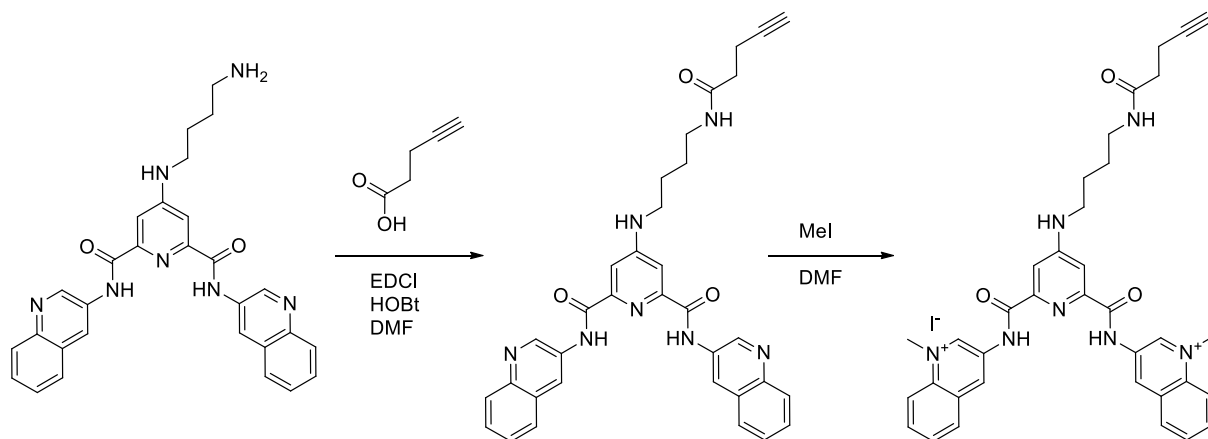
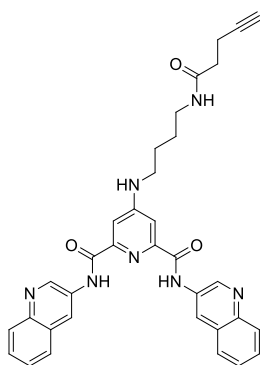


Figure 209 : Synthesis of compound PDC-alk.

La synthèse du PDC-NH₂ a été décrite dans la littérature^[171].

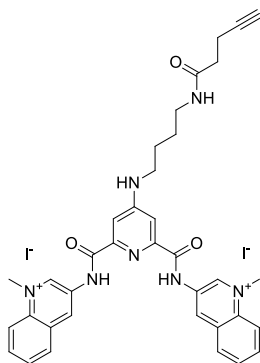


PDCn-alk

PDC-NH₂ (400 mg, 0.79 mmol, 1eq), 4-pentynoic acid (116 mg, 1.19 mmol, 0.5 eq), EDCI (183 mg, 1.18 mmol, 1.49 eq) and HOBt (17 mg, 0.12 mmol, 0.15 eq) were introduced in a 25 mL flask and dissolved in 8 mL of DMF. The resulting mixture was stirred at RT overnight and the reaction was complete. The product was precipitated by adding water and filtered. Then it was washed with water and dried with ether and the resulting product (white powder) was dried once again under reduced pressure.

¹H NMR (300 MHz, DMSO-d₆): δ (ppm) 11.29 (s, 2H), 9.36 (m, 2H), 8.97 (s, 2H), 8.05 (d, J = 8.4Hz, 4H), 7.93 (t, J = 5.5 Hz, 1H), 7.73 (td, J = 1.0 Hz and J = 8.0 Hz, 2H), 7.64 (td, J = 1.0 Hz and J = 8.0 Hz, 2H), 7.53 (m, 2H), 3.26 (q, J = 6.0 Hz, 2H), 3.13 (q, J = 6.1 Hz, 2H), 2.76 (t, J = 2.5 Hz, 1H), 2.36 (m, 2H), 2.28 (m, 2H), 1.59 (m, 4H)

LR-MS (ESI-MS) m/z = 586.6 [M+H]⁺



PDC-alk

A solution of the starting compound PDCn-alk (70 mg, 0.12mmol, 1.0 eq) in DMF (2 mL) was heated at 40 °C before methyl iodide (370 μ L, 5.98mmol, 50eq) was added dropwise. The reaction was performed under argon atmosphere for 24h. The solvent was evaporated to dryness and the expected compound was taken from ethanol as a yellow solid (67%).

$^1\text{H NMR}$ (300 MHz, DMSO- d_6): δ (ppm) 11.73 (s, 2H), 10.11 (s, 2H), 9.64 (s, 2H), 8.56 (d, $J = 8.5$ Hz, 4H), 8.24 (t, $J = 8.5$ Hz, 2H), 8.09 (t, $J = 8.5$ Hz, 2H), 7.94 (m, 1H), 7.71 (m, 1H), 7.60 (bs, 1H), 4.78 (s, 6H), 3.28 (m, 2H), 3.13 (m, 2H), 2.76 (t, $J = 2.8$ Hz, 1H), 2.36 (m, 2H), 2.28 (m, 2H), 1.59 (m, 4H)

$^{13}\text{C NMR}$ (500MHz, DMSO- d_6): δ (ppm) 170.2, 163.2, 156.6, 144.7, 135.7, 134.2, 134.0, 132.4, 130.4, 129.9, 129.2, 119.2, 83.8, 71.3, 46.2, 41.7, 38.0, 34.3, 26.7, 25.4, 14.3

LR-MS (ESI-MS) $m/z = 616$ $[\text{M}+\text{H}]^+$; $m/z = 308$ $[\text{M}/2+\text{H}]^+$

HR-MS (ESI+): $m/z = 742.2003$ calculated for $\text{C}_{36}\text{H}_{37}\text{N}_7\text{O}_3\text{I}$; found, 742.1995

II. Expériences de biophysique

1) Oligonucléotides

Oligonucleotides were synthesized and purified by Eurogentec and stored at -20 °C as a 200 μ M aqueous solution. Concentrations were determined by ultraviolet absorption using the extinction coefficients provided by the manufacturer after heating the DNA 5 minutes at 90 °C. UV-Vis absorption spectra were recorded on an Agilent Cary 300 spectrophotometer.

DNA folding:

- For G-quadruplex DNA, DNA was heated at 95°C for 5 minutes and then was cooled in ice for 30 minutes.
- For duplex DNA, DNA was heated at 95°C for 5 minutes and then let the DNA return at room temperature slowly (a night or 6 hours).

2) Expériences de FRET-melting

A fast real-time PCR apparatus (7900HT) was used to carry out the FRET melting experiments in 96-well plates. The excitation and detection wavelengths were 488 nm and 516 nm, respectively. After an initial incubation at 25 °C for 5 min, the temperature was increased by 0.5 °C every minute until 95 °C was reached. Each experimental condition was tested in duplicate. The mid-transition T_m was determined from normalized curves. The experiments were performed with samples containing 0.2 μ M oligonucleotide, 10 mM lithium cacodylate (pH 7.2), and 10 mM KCl and 90 mM LiCl for F-21-T and 1 mM KCl and 99 mM LiCl for others. For measurements in the presence of a ligand, the concentration of ligand was 1 μ M. The stabilisation induced by a ligand is reported in terms of the difference in the mid-transition temperature with and without the ligand (ΔT_m).

3) Expériences de FID

The High-Throughput-G4-FID assay is performed in 96-well plates (Corning®, Thermofischer). Every ligand was tested on a line of the microplate, at least in duplicate. The plate was filled with the buffer K^+100 (10 mM Cacoli and 100 mM KCl, from column A to H: 190 μ L to 90 μ L), a solution with the pre-folded DNA containing the fluorescent probe (10 μ L) and a solution of the ligand to test (from A to H: 0 to 100 μ L, i.e. 0, 0.125, 0.25, 0.375, 0.5, 0.625, 0.75, 1.0, 1.25, 1.5, 2.0 and 2.5 μ M). Each well was filled with 200 μ L of liquid. After 5 minutes of shaking, fluorescence is measured with a microplate reader (FLUOstar Omega®, BMG Labtech) using the following parameters; position delay: 0.5 s, 20 flashes per well, emission/excitation filters for TO: 485/520 nm, gain adjusted at 80% of the fluorescence from the most fluorescent well. The percentage of displacement (TOD) was calculated from the fluorescence intensity (F) and the fluorescent from the fluorescent probe bound to DNA without added ligand (F_0).

$$TOD (\%) = 100 \times \left(1 - \frac{F}{F_0}\right)$$

The TOD was then plotted as a function of the concentration of added ligand. The DNA affinity was evaluated by the concentration of ligand required to decrease the fluorescence of the probe by 50 %, noted DC_{50} , and determined after non-linear fitting of the displacement curve. Another method to determine the DNA affinity is to take the value of displacement at 1 μ M of ligand. This method is useful when ligands do not displace 50 % of the fluorescent probe.

4) Dichroïsme circulaire

CD spectra were recorded on a Jasco-J815 spectropolarimeter with a scanning speed of 500 nm/min, a response time of 1 s, 1 nm of data pitch and 1 nm of bandwidth. The concentration of DNA for the measurement was 2.5 μ M.

5) Temps de vie de fluorescence

Lifetime fluorescence was performed on DeltaTime TCSPC lifetime plug-in (Horiba®) with a 370 nm laser-diode source and emission was read at 556 nm. Solutions of molecule were prepared at 5 μ M.

III. Culture cellulaire et imagerie cellulaire

1) Culture cellulaire

HT-29 (ATCC® HTB-38™) and A549 (ATCC® CCL-185™) cell lines were purchased from the American Type Culture Collection (ATCC, Manassas, VA, USA,) and cultured in Dulbecco's modified Eagle's medium GlutaMAX™ (DMEM, Life Technologies, Saint-Aubin, France) supplemented with 10% fetal calf serum (FCS, South America origin, BioWhittaker® Lonza, Verviers, Belgium) and 1% Penicillin/Streptomycin (Life Technologies) in humidified atmosphere under 5% CO₂ in air at 37 °C.

MDA-MB-231 cells (ATCC® HTB-26™) were purchased from the American Type Culture Collection (ATCC, Manassas, VA, USA,) and cultured in Dulbecco's modified Eagle's/F12 (1:1) medium GlutaMAX™ (DMEM/F12 (1:1), Life Technologies, Saint-Aubin, France) supplemented with 10% FCS (BioWhittaker® Lonza) and 1% Penicillin/Streptomycin (Life Technologies) in humidified atmosphere under 5% CO₂ in air at 37 °C.

Cells were split twice a week. To split cells, these were washed with Phosphate Buffer Solution (PBS, Life Technologies) and 2mL of TrypLE Express Enzyme (Life technologies) was added for 5-10 minutes at 37°C. Cells were collected in warm medium (addition of 8mL) and centrifuged for five minutes at 2000 rpm, the supernatant was discarded and cells re-suspended in medium and seeded at a density of 6-7 million cells in a T75 culturing flask for HT-29 cells and 1 million cells in a T75 culturing flask for MDA-MB231 and A549 cells. Cells were counted using a MaxiZ™ counter (Orflo®, Ketchum, ID, United States of America).

2) Protocole en cellules fixées avec incubation en cellules vivantes ou fixées

Cells were plated in 2-well LabTek® II chamber slide (Thermofischer, Villebon-sur-Yvette, France) at a density of two or three hundred thousand cells per well. Cells were incubated for 24h at 37°C and 5%

CO₂ to allow them to attach. Ligands were added at 5 μM in PBS for 24h in live cells. Cells were washed two times with PBS and then fixed with 1mL of 2% paraformaldehyde (PFA, Electron Microscopy Sciences, Hatfield, PA, United States of America) in PBS for 15 minutes. They were washed two times with 3% bovine serum albumin (BSA, Sigma, Saint-Quentin-Fallavier, France) in PBS for 5 minutes each. They were also permeabilized with 0.5% (v/v) of Triton X-100 (Sigma) in PBS for 20 minutes. They were washed again two times with 3% BSA in PBS. Fluorophore (20 μM for Cy5-az (Lumiprobe) and Cy5-alk (Lumiprobe) and 10 μM for Cy5-DBCO (Sigma)) was added in presence or not of the copper (1 mM) and sodium ascorbate (100 mM) for 1h in the dark at room temperature. Cells were washed two times and slides were mounted with Prolong[®]-DAPI (Life Technologies) and a coverglass (n° 1.5 thickness).

In fixed cells, ligands were added for 1h after the step of permeabilization.

3) Protocole en cellules vivantes

Cells were plated in 2-well LabTek[®] II chambered coverglass (n°1.5 thickness, Thermofischer, Villebon-sur-Yvette, France) at a density of two or three hundred thousand cells per well. Cells were incubated for 24h at 37°C and 5% CO₂ to allow them to attach. Ligands were added at 10 μM in medium for 24h. Cells were washed once with 1% FCS in PBS. Fluorophore (10 μM for TAMRA-DBCO (Click-Chemistry Tools)) was added for 1h in the dark at room temperature. Medium was changed by and images were recorded immediately after the incubation of fluorophore.

4) Imagerie cellulaire

Images of fields were acquired on a Leica confocal SP5 system equipped with a Plan APO CS 63.0x OIL UV (NA 1.4) objective and coupled with a 405 nm diode laser (observation of DAPI or coumarin or Hoechst 33342), a 488 nm argon laser (observation of AlexaFluor 488-alk), a 561 nm diode-pumped solid state (DPSS) laser (observation of TAMRA) and a 633 nm helium gas laser (observation of Cyanine 5). The images were visualized and processed using the Image J software.

IV. Expériences d'extraction supportée par billes magnétiques

1) Principe de l'extraction supportée par billes magnétiques recouvertes de streptavidine

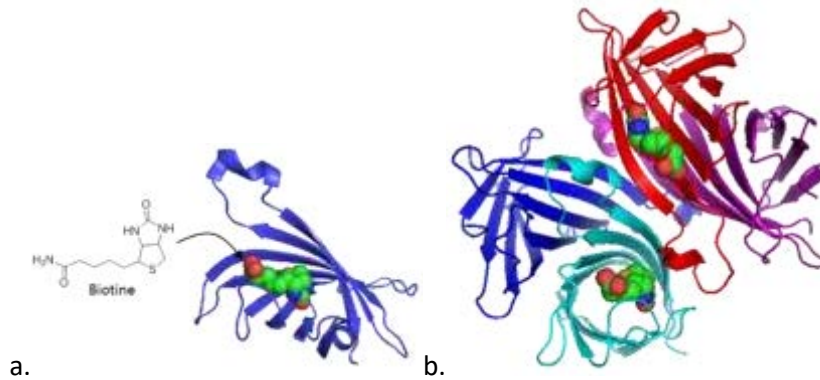


Figure 210 : Représentation de a. l'interaction biotine/streptavidine et b. du tétramère formé par la streptavidine.

La streptavidine est une protéine, issue de la bactérie *Streptomyces avidinii*, qui s'organise en tétramères et possède une affinité particulièrement forte pour la biotine (figure 210). La constante d'association du complexe streptavidine/biotine est l'une des interactions non covalentes naturelles les plus fortes ($K_a = 10^{14}$ mol/L). En particulier, le complexe streptavidine/biotine résiste à la plupart des solvants organiques, aux détergents, à l'action des enzymes protéolytiques et aux conditions extrêmes de température ou de pH. L'utilisation de billes magnétiques recouvertes de streptavidine va permettre l'extraction supportée de tout composé portant un groupement biotine (figure 211). Après formation du complexe streptavidine/biotine, les billes magnétiques sont isolées du reste de la solution grâce à un aimant adapté. Elles peuvent ensuite être lavées afin d'y éliminer toute fixation non spécifique et traitées pour récupérer les composés fixés sur leur surface.

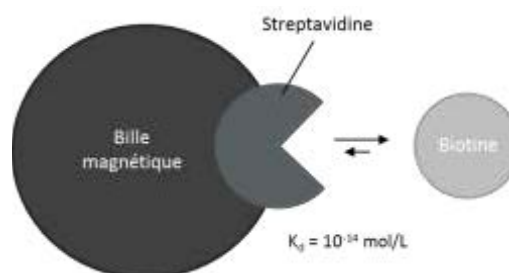


Figure 211 : Représentation schématique du fonctionnement des billes recouvertes de streptavidine.

2) Mode opératoire

a) Préparation des oligonucléotides

Les oligonucléotides (Eurogentec) sont radiomarqués au ^{32}P à l'extrémité 5' grâce à l'action de l'enzyme T4-polymérase kinase en présence d'adénosine 5'-triphosphate [$\gamma\text{-}^{32}\text{P}$] (ATP* - Helmer Perkin) (figure 212).

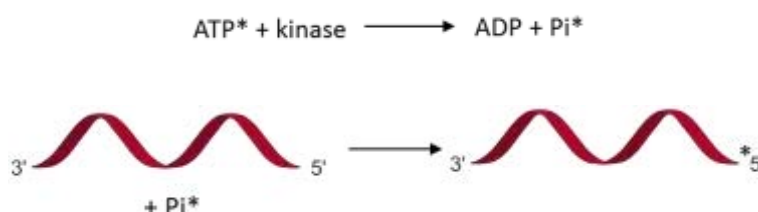


Figure 212 : Mécanisme de radiomarquage de l'ADN.

Un mélange contenant 1 μL d'oligonucléotide à 20 μM , 1 μL de tampon Tris 10x, 2 μL d'enzyme, 3 μL d'ATP* et 2 μL d'eau est incubé 30 min à 37 °C. Après ajout de 8 μL de bleu de dépôt, le mélange est purifié sur gel d'acrylamide 20 % (1 h à 20 W puis 2 h à 30 W). Les bandes d'ADN radiomarqué sont découpées du gel, éluées une nuit dans 200 μL d'une solution de NaCl 0,15 M et précipitées en présence d'éthanol glacé. Ce précipité est ensuite séché et stocké à -20 °C comme culot d'ADN radiomarqué.

b) Réaction avec les billes magnétiques recouvertes de streptavidine

L'ADN quadruplex radiomarqué ^{32}P est hybridé à une concentration de 2 μM dans le tampon K^+100 (CacoLi 10 mM, KCl 100 mM). L'ADN est ensuite réparti dans chaque tube, à une concentration de 1 μM pour un volume total de 200 μL , en présence ou non de ligand, et sa radioactivité est comptée une première fois. Ce comptage initial constitue notre 100 % de radioactivité.

Préparation des billes : 50 μL de solution commerciale de billes Ademtech sont prélevés et dilués dans 350 μL de tampon K^+100 (Cacodylate de lithium 10 mM, KCl 100 mM, pH 7,4). Les billes sont lavées par 3 x 400 μL de tampon K^+100 grâce à un support magnétique adéquat.

Pré-saturation des billes : si l'expérience le nécessite, les billes fraîchement lavées sont incubées pendant 1h avec une solution de molécule biotinylé (de 1 à 6 μM , dans 200 μL). Elles sont ensuite lavées par 2 x 200 μL de tampon K^+100 afin de retirer le ligand non fixé.

Réaction sur billes pré-saturées : L'ADN et le ligand sont ajoutés sur les billes fraîchement préparées. Le mélange est agité 3 h sur un agitateur mécanique à température ambiante pour permettre la réaction de SPAAC. Le premier surnageant est retiré et les billes sont ensuite lavées par 3 x 200 µL de tampon K⁺100 afin de retirer tout l'ADN fixé de façon non spécifique.

Réaction avec billes non pré-saturées : L'ADN et le ligand sont ajoutés dans une solution contenant du DBCO-biotine (1:1 Ligand:DBCO-Biotine) pendant une heure pour que la réaction de SPAAC ait lieu. Ensuite le surnageant est retiré et placé en présence des billes recouvertes de streptavidine pendant 2h. Le premier surnageant est retiré et les billes sont ensuite lavées par 3 x 200 µL de tampon K⁺100 afin de retirer tout l'ADN fixé de façon non spécifique.

Décrochage des billes : (figure 213)

1) **En brisant la liaison ADN G-quadruplex-ligand :** Après le dernier lavage, 200 µL d'eau sont ajoutés sur les billes et le mélange est chauffé à 90 °C pendant 30 min. Le surnageant est retiré à chaud afin de récupérer l'ADN quadruplex dénaturé. Ce type de décrochage à l'eau chaude permet de briser le complexe ligand-ADN formé.

2) **En brisant la liaison biotine-streptavidine :** Un décrochage au formamide est réalisé. 200 µL d'un mélange à base de formamide (950 µl formamide, 30 µL EDTA 0,5 M et 20 µL H₂O) sont ajoutés et les billes sont chauffées à 90 °C pendant 30 °C. Le surnageant est retiré après chauffage.

3) Un décrochage au sel LiCl (8M) est également réalisé. La solution est agitée 30 min à température ambiante puis le surnageant est retiré.

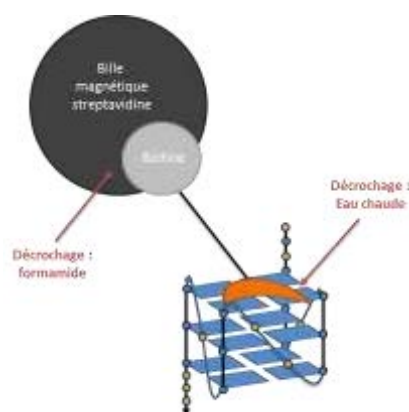


Figure 213 : Représentation schématique de décrochage de l'oligonucléotide.

c) Réaction avec les billes magnétiques recouvertes de fonction -DBCO

Afin de s'affranchir d'une étape par rapport à l'expérience précédente, nous avons testé l'expérience en présence de billes magnétiques recouvertes de fonction -DBCO. Les étapes de préparation des billes et de décrochage de l'ADN restent identiques.

Réaction sur billes pré-saturées : Les billes fraîchement préparées sont mises en présence du ligand fonctionnalisé par un azoture pendant une heure. Le surnageant est alors retiré puis l'ADN est ajouté pendant deux heures. Le premier surnageant est retiré et les billes sont ensuite lavées par 3 x 200 µL de tampon K⁺100 afin de retirer tout l'ADN fixé de façon non spécifique.

Réaction avec billes non pré-saturées : L'ADN et le ligand sont ajoutés sur les billes pendant trois heures pour que la réaction de SPAAC ait lieu. Le premier surnageant est retiré et les billes sont ensuite lavées par 3 x 200 µL de tampon K⁺100 afin de retirer tout l'ADN fixé de façon non spécifique.

Références:

- [1] J. D. Watson, F. H. C. Crick, *Nature* **1953**, *171*, 737-738.
- [2] A. H. J. Wang, G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel, A. Rich, *Nature* **1979**, *282*, 680-686.
- [3] W. Fuller, W. H. F. Wilkins, H. R. Wilson, L. D. Hamilton, *J Mol Biol* **1965**, *12*, 60-IN69.
- [4] B. F. Eichman, M. Ortiz-Lombardía, J. Aymamí, M. Coll, P. S. Ho, *J Mol Biol* **2002**, *320*, 1037-1051.
- [5] A. P. Ryskov, R. B. Church, G. Bajszar, G. P. Georgiev, *Molecular Biology Reports* **1973**, *1*, 119-122.
- [6] T. R. Broker, A. H. Doermann, *Annual Review of Genetics* **1975**, *9*, 213-244.
- [7] J. S. Lee, G. D. Burkholder, L. J. Latimer, B. L. Haug, R. P. Braun, *Nucleic Acids Res* **1987**, *15*, 1047-1061.
- [8] I. Radhakrishnan, D. J. Patel, *Structure* **1993**, *1*, 135-152.
- [9] J. T. Davis, *Angew Chem Int Ed Engl* **2004**, *43*, 668-698.
- [10] J. L. Leroy, M. Guéron, J. L. Mergny, C. Hélène, *Nucleic Acids Res* **1994**, *22*, 1600-1606.
- [11] A. Annunziato, *Nature Education* **2008**, *1*, 26.
- [12] a) I. Bang, *Biochemische Zeitschrift* **1910**, *26*, 293-311; b) G. Wu, in *Guanine Quartets : Structure and Application*, RSC publishing ed., Lea Spindler, Wolfgang Fritzsche, **2012**, pp. 3-14.
- [13] M. Gellert, M. N. Lipsett, D. R. Davies, *Proc Natl Acad Sci USA* **1962**, *48*, 2013-2018.
- [14] T. C. Marsh, J. Vesenka, E. Henderson, *Nucleic Acids Res* **1995**, *23*, 696-700.
- [15] D. Sen, W. Gilbert, *Nature* **1988**, *334*, 364-366.
- [16] S. Zhang, Y. Wu, W. Zhang, *ChemMedChem* **2014**, *9*, 899-911.
- [17] D. Rhodes, H. J. Lipps, *Nucleic Acids Res* **2015**, *43*, 8627-8637.
- [18] J. L. Huppert, S. Balasubramanian, *Nucleic Acids Res* **2005**, *33*, 2908-2916.
- [19] A. K. Todd, M. Johnston, S. Neidle, *Nucleic Acids Res* **2005**, *33*, 2901-2907.
- [20] A. Bedrat, L. Lacroix, J. L. Mergny, *Nucleic Acids Res* **2016**, *44*, 1746-1759.
- [21] a) J. L. Huppert, *Nucleic Acids Research* **2005**, *33*, 2908-2916; b) H. M. Wong, O. Stegle, S. Rodgers, J. L. Huppert, *Journal of Nucleic Acids* **2010**, *2010*, 1-6.
- [22] a) H. J. Lipps, W. Gruissem, D. M. Prescott, *Proc Natl Acad Sci USA* **1982**, *79*, 2495-2499; b) W. I. Sundquist, A. Klug, *Nature* **1989**, *342*, 825-829.
- [23] R. K. Moyzis, J. M. Buckingham, L. S. Cram, M. Dani, L. L. Deaven, M. D. Jones, J. Meyne, R. L. Ratliff, J. R. Wu, *Proc Natl Acad Sci USA* **1988**, *85*, 6622-6626.
- [24] V. L. Makarov, Y. Hirose, J. P. Langmore, *Cell* **1997**, *88*, 657-666.
- [25] A. M. Zahler, J. R. Williamson, T. R. Cech, D. M. Prescott, *Nature* **1991**, *350*, 718-720.
- [26] M. L. Zhang, X. J. Tong, X. H. Fu, B. O. Zhou, J. Wang, X. H. Liao, Q. J. Li, N. Shen, J. Ding, J. Q. Zhou, *Nat Struct Mol Biol* **2010**, *17*, 202-209.
- [27] a) A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, *Proceedings of the National Academy of Sciences* **2002**, *99*, 11593-11598; b) Y. Qin, L. H. Hurley, *Biochimie* **2008**, *90*, 1149-1171.
- [28] J. Eddy, *Nucleic Acids Research* **2006**, *34*, 3887-3896.
- [29] D. Hanahan, R. A. Weinberg, *Cell* **2000**, *100*, 57-70.
- [30] T. A. Brooks, S. Kendrick, L. Hurley, *FEBS Journal* **2010**, *277*, 3459-3469.
- [31] G. Vergnaud, F. Denoeud, *Genome Research* **2000**, *10*, 899-907.
- [32] a) A. Piazza, J. B. Boule, J. Lopes, K. Mingo, E. Largy, M. P. Teulade-Fichou, A. Nicolas, *Nucleic Acids Res* **2010**, *38*, 4337-4348; b) M. Adrian, D. J. Ang, C. J. Lech, B. Heddi, A. Nicolas, A. T. Phan, *J Am Chem Soc* **2014**, *136*, 6297-6305; c) A. Piazza, M. Adrian, F. Samazan, B. Heddi, F. Hamon, A. Serero, J. Lopes, M. P. Teulade-Fichou, A. T. Phan, A. Nicolas, *EMBO J* **2015**, *34*, 1718-1734.
- [33] D. Monchaud, M. P. Teulade-Fichou, *Org Biomol Chem* **2008**, *6*, 627-636.

- [34] R. Rodriguez, G. D. Pantoş, D. P. N. Gonçalves, J. K. M. Sanders, S. Balasubramanian, *Angew Chem Int Edit* **2007**, *46*, 5405-5407.
- [35] D. Sun, B. Thompson, B. E. Cathers, M. Salazar, S. M. Kerwin, J. O. Trent, T. C. Jenkins, S. Neidle, L. H. Hurley, *J Med Chem* **1997**, *40*, 2113-2116.
- [36] S. M. Haider, G. N. Parkinson, S. Neidle, *J Mol Biol* **2003**, *326*, 117-125.
- [37] a) R. J. Harrison, J. Cuesta, G. Chessari, M. A. Read, S. K. Basra, A. P. Reszka, J. Morrell, S. M. Gowan, C. M. Incles, F. A. Tanious, W. D. Wilson, L. R. Kelland, S. Neidle, *J Med Chem* **2003**, *46*, 4463-4476; b) C. M. Schultes, B. Guyen, J. Cuesta, S. Neidle, *Bioorg Med Chem Lett* **2004**, *14*, 4347-4351.
- [38] N. H. Campbell, G. N. Parkinson, A. P. Reszka, S. Neidle, *J Am Chem Soc* **2008**, *130*, 6722-6724.
- [39] O. Y. Fedoroff, M. Salazar, H. Han, V. V. Chemeris, S. M. Kerwin, L. H. Hurley, *Biochemistry* **1998**, *37*, 12367-12374.
- [40] A. K. Mehta, Y. Shayo, H. Vankayalapati, L. H. Hurley, J. Schaefer, *Biochemistry* **2004**, *43*, 11953-11958.
- [41] J.-L. Mergny, L. Lacroix, M.-P. Teulade-Fichou, C. Hounsou, L. Guittat, M. Hoarau, P. B. Arimondo, J.-P. Vigneron, J.-M. Lehn, J.-F. Riou, T. Garestier, C. Hélène, *Proc Natl Acad Sci USA* **2001**, *98*, 3062-3067.
- [42] F. Cuenca, O. Greciano, M. Gunaratnam, S. Haider, D. Munnur, R. Nanjunda, W. D. Wilson, S. Neidle, *Bioorg Med Chem Lett* **2008**, *18*, 1668-1673.
- [43] a) M. Micco, G. W. Collie, A. G. Dale, S. A. Ohnmacht, I. Pazitna, M. Gunaratnam, A. P. Reszka, S. Neidle, *J Med Chem* **2013**, *56*, 2959-2974; b) S. A. Ohnmacht, C. Marchetti, M. Gunaratnam, R. J. Besser, S. M. Haider, G. Di Vita, H. L. Lowe, M. Mellinas-Gomez, S. Diocou, M. Robson, J. Sponer, B. Islam, R. B. Pedley, J. A. Hartley, S. Neidle, *Sci Rep* **2015**, *5*, 11385.
- [44] C. Marchetti, A. Minarini, V. Tumiatti, F. Moraca, L. Parrotta, S. Alcaro, R. Rigo, C. Sissi, M. Gunaratnam, S. A. Ohnmacht, S. Neidle, A. Milelli, *Bioorg Med Chem* **2015**, *23*, 3819-3830.
- [45] M.-P. Teulade-Fichou, C. Carrasco, L. Guittat, C. Bailly, P. Alberti, J.-L. Mergny, A. David, J.-M. Lehn, W. D. Wilson, *J Am Chem Soc* **2003**, *125*, 4732-4740.
- [46] K. Shin-ya, K. Wierzba, K.-i. Matsuo, T. Ohtani, Y. Yamada, K. Furihata, Y. Hayakawa, H. Seto, *J Am Chem Soc* **2001**, *123*, 1262-1263.
- [47] M.-Y. Kim, H. Vankayalapati, K. Shin-ya, K. Wierzba, L. H. Hurley, *J Am Chem Soc* **2002**, *124*, 2098-2099.
- [48] T. Doi, M. Yoshida, K. Shin-ya, T. Takahashi, *Org Lett* **2006**, *8*, 4165-4167.
- [49] T. Doi, K. Shibata, M. Yoshida, M. Takagi, M. Tera, K. Nagasawa, K. Shin-ya, T. Takahashi, *Org Biomol Chem* **2011**, *9*, 387-393.
- [50] M. Tera, H. Ishizuka, M. Takagi, M. Suganuma, K. Shin-ya, K. Nagasawa, *Angew Chem Int Edit* **2008**, *47*, 5557-5560.
- [51] V. Gabelica, E. Shammel Baker, M.-P. Teulade-Fichou, E. De Pauw, M. T. Bowers, *J Am Chem Soc* **2007**, *129*, 895-904.
- [52] W. J. Chung, B. Heddi, M. Tera, K. Iida, K. Nagasawa, A. T. Phan, *J Am Chem Soc* **2013**, *135*, 13495-13501.
- [53] a) F. Hamon, E. Largy, A. Guédin-Beaurepaire, M. Rouchon-Dagois, A. Sidibe, D. Monchaud, J.-L. Mergny, J.-F. Riou, C.-H. Nguyen, M.-P. Teulade-Fichou, *Angew Chem Int Edit* **2011**, *50*, 8745-8749; b) M. Petenzi, D. Verga, E. Largy, F. Hamon, F. Doria, M. P. Teulade-Fichou, A. Guédin, J. L. Mergny, M. Mella, M. Freccero, *Chemistry - A European Journal* **2012**, *18*, 14487-14496.
- [54] a) A. Siddiqui-Jain, C. L. Grand, D. J. Bearss, L. H. Hurley, *Proc Natl Acad Sci USA* **2002**, *99*, 11593-11598; b) D. Sun, K. Guo, J. J. Rusche, L. H. Hurley, *Nucleic Acids Res* **2005**, *33*, 6070-6080; c) J. Seenisamy, E. M. Rezler, T. J. Powell, D. Tye, V. Gokhale, C. S. Joshi, A. Siddiqui-Jain, L. H. Hurley, *J Am Chem Soc* **2004**, *126*, 8702-8709; d) D. J. Cashman, R. Buscaglia, M. W. Freyer, J. Dettler, L. H. Hurley, E. A. Lewis, *Journal of Molecular Modeling* **2008**, *14*, 93-101.
- [55] A. De Cian, L. Guittat, K. Shin-ya, J.-F. Riou, J.-L. Mergny, *Nucleic Acids Symposium Series* **2005**, *49*, 235-236.

- [56] a) D. Monchaud, C. Allain, M. P. Teulade-Fichou, *Bioorg Med Chem Lett* **2006**, *16*, 4842-4845; b) J. Ren, J. B. Chaires, *Biochemistry* **1999**, *38*, 16067-16075.
- [57] a) Y. Li, J. Syed, Y. Suzuki, S. Asamitsu, N. Shioda, T. Wada, H. Sugiyama, *Chembiochem* **2016**, *17*, 928-935; b) K. Kulikov, S. Nozinovic, S. Kath-Schorr, *Chembiochem* **2017**, *18*, 969-973.
- [58] D. P. N. Goncalves, R. Rodriguez, S. Balasubramanian, J. K. M. Sanders, *Chem Commun* **2006**, 4685-4687.
- [59] P. Wang, L. Ren, H. He, F. Liang, X. Zhou, Z. Tan, *Chembiochem* **2006**, *7*, 1155-1159.
- [60] J. Seenisamy, S. Bashyam, V. Gokhale, H. Vankayalapati, D. Sun, A. Siddiqui-Jain, N. Streiner, K. Shin-ya, E. White, W. D. Wilson, L. H. Hurley, *J Am Chem Soc* **2005**, *127*, 2944-2959.
- [61] E. M. Rezler, J. Seenisamy, S. Bashyam, M.-Y. Kim, E. White, W. D. Wilson, L. H. Hurley, *J Am Chem Soc* **2005**, *127*, 9439-9447.
- [62] T. Lemarteleur, D. Gomez, R. Paterski, E. Mandine, P. Mailliet, J. F. Riou, *Biochem Biophys Res Commun* **2004**, *323*, 802-808.
- [63] G. Pennarun, C. Granotier, L. R. Gauthier, D. Gomez, F. Hoffschir, E. Mandine, J. F. Riou, J. L. Mergny, P. Mailliet, F. D. Boussin, *Oncogene* **2005**, *24*, 2917-2928.
- [64] C. Granotier, G. Pennarun, L. Riou, F. Hoffschir, L. R. Gauthier, A. De Cian, D. Gomez, E. Mandine, J. F. Riou, J. L. Mergny, P. Mailliet, B. Dutrillaux, F. D. Boussin, *Nucleic Acids Res* **2005**, *33*, 4182-4190.
- [65] A. De Cian, J. L. Mergny, *Nucleic Acids Res* **2007**, *35*, 2483-2493.
- [66] R. Rodriguez, S. Müller, J. A. Yeoman, C. Trentesaux, J.-F. Riou, S. Balasubramanian, *J Am Chem Soc* **2008**, *130*, 15758-15759.
- [67] S. Muller, D. A. Sanders, M. Di Antonio, S. Matsis, J. F. Riou, R. Rodriguez, S. Balasubramanian, *Org Biomol Chem* **2012**, *10*, 6537-6546.
- [68] a) R. Rodriguez, K. M. Miller, J. V. Forment, C. R. Bradshaw, M. Nikan, S. Britton, T. Oelschlaegel, B. Xhemalce, S. Balasubramanian, S. P. Jackson, *Nat Chem Biol* **2012**, *8*, 301-310; b) A. B. Sahakyan, P. Murat, C. Mayer, S. Balasubramanian, *Nat Struct Mol Biol* **2017**; c) J. Zimmer, E. M. Tacconi, C. Folio, S. Badie, M. Porru, K. Klare, M. Tumiaty, E. Markkanen, S. Halder, A. Ryan, S. P. Jackson, K. Ramadan, S. G. Kuznetsov, A. Biroccio, J. E. Sale, M. Tarsounas, *Mol Cell* **2016**, *61*, 449-460.
- [69] A. De Cian, E. DeLemos, J.-L. Mergny, M.-P. Teulade-Fichou, D. Monchaud, *J Am Chem Soc* **2007**, *129*, 1856-1857.
- [70] W. J. Chung, B. Heddi, F. Hamon, M. P. Teulade-Fichou, A. T. Phan, *Angew Chem Int Ed Engl* **2014**, *53*, 999-1002.
- [71] L. Lacroix, A. Séosse, J.-L. Mergny, *Nucleic Acids Res* **2011**, *39*, e21-e21.
- [72] V. Dhamodharan, S. Harikrishna, A. C. Bhasikuttan, P. I. Pradeepkumar, *ACS Chem Biol* **2015**, *10*, 821-833.
- [73] S. Harikrishna, S. Kotaru, P. I. Pradeepkumar, *Mol Biosyst* **2017**.
- [74] L. Martino, A. Virno, B. Pagano, A. Virgilio, S. Di Micco, A. Galeone, C. Giancola, G. Bifulco, L. Mayol, A. Randazzo, *J Am Chem Soc* **2007**, *129*, 16048-16056.
- [75] a) J. G. Pelton, D. E. Wemmer, *J Am Chem Soc* **1990**, *112*, 1393-1399; b) R. Baliga, D. M. Crothers, *Proc Natl Acad Sci USA* **2000**, *97*, 7814-7818.
- [76] D. Wang, S. J. Lippard, *Nat Rev Drug Discov* **2005**, *4*, 307-320.
- [77] a) S. Teletchea, T. Skauge, E. Sletten, J. Kozelka, *Chemistry* **2009**, *15*, 12320-12337; b) P. Heringova, J. Kasparkova, V. Brabec, *J Biol Inorg Chem* **2009**, *14*, 959-968.
- [78] J. E. Reed, A. A. Arnal, S. Neidle, R. Vilar, *J Am Chem Soc* **2006**, *128*, 5992-5993.
- [79] a) J. E. Reed, S. Neidle, R. Vilar, *Chem Commun* **2007**, 4366-4368; b) J. E. Reed, A. J. White, S. Neidle, R. Vilar, *Dalton Trans* **2009**, 2558-2568.
- [80] H. Bertrand, S. Bombard, D. Monchaud, E. Talbot, A. Guedin, J.-L. Mergny, R. Grunert, P. J. Bednarski, M.-P. Teulade-Fichou, *Org Biomol Chem* **2009**, *7*, 2864-2871.
- [81] G. T. Morgan, F. H. Burstall, *Journal of the Chemical Society (Resumed)* **1932**, 20-30.
- [82] S. D. Cummings, *Coordination Chemistry Reviews* **2009**, *253*, 449-478.

- [83] H. Bertrand, D. Monchaud, A. De Cian, R. Guillot, J.-L. Mergny, M.-P. Teulade-Fichou, *Org Biomol Chem* **2007**, *5*, 2555.
- [84] K. Suntharalingam, A. J. White, R. Vilar, *Inorg Chem* **2009**, *48*, 9427-9435.
- [85] E. Morel, F. Poyer, L. Vaslin, S. Bombard, M.-P. Teulade-Fichou, *Inorganica Chimica Acta* **2016**, *452*, 152-158.
- [86] E. Largy, F. Hamon, F. Rosu, V. Gabelica, E. De Pauw, A. Guedin, J. L. Mergny, M. P. Teulade-Fichou, *Chemistry* **2011**, *17*, 13274-13283.
- [87] a) Z. A. Waller, P. S. Shirude, R. Rodriguez, S. Balasubramanian, *Chem Commun* **2008**, 1467-1469; b) Z. A. Waller, S. A. Sewitz, S. T. Hsu, S. Balasubramanian, *J Am Chem Soc* **2009**, *131*, 12628-12633.
- [88] a) E. Izbiccka, R. T. Wheelhouse, E. Raymond, K. K. Davidson, R. A. Lawrence, D. Sun, B. E. Windle, L. H. Hurley, D. D. Von Hoff, *Cancer Research* **1999**, *59*, 639-644; b) D.-F. Shi, R. T. Wheelhouse, D. Sun, L. H. Hurley, *Journal of Medicinal Chemistry* **2001**, *44*, 4509-4523.
- [89] E. Wachter, D. Moya, S. Parkin, E. C. Glazer, *Chemistry* **2016**, *22*, 550-559.
- [90] S. M. Gowan, R. Heald, M. F. G. Stevens, L. R. Kelland, *Molecular Pharmacology* **2001**, *60*, 981.
- [91] S. Lagah, I. L. Tan, P. Radhakrishnan, R. A. Hirst, J. H. Ward, C. O'Callaghan, S. J. Smith, M. F. G. Stevens, R. G. Grundy, R. Rahman, *PLoS One* **2014**, *9*, e86187.
- [92] C. Leonetti, S. Amodei, C. Angelo, A. Rizzo, B. Benassi, A. Antonelli, R. Elli, M. F. G. Stevens, M. Incalci, G. Zupi, A. Biroccio, *Molecular Pharmacology* **2004**, *66*, 1138.
- [93] a) D. Drygin, A. Siddiqui-Jain, S. O'Brien, M. Schwaebe, A. Lin, J. Bliesath, C. B. Ho, C. Proffitt, K. Trent, J. P. Whitten, J. K. Lim, D. Von Hoff, K. Anderes, W. G. Rice, *Cancer Res* **2009**, *69*, 7653-7661; b) T. A. Brooks, L. H. Hurley, *Genes Cancer* **2010**, *1*, 641-649.
- [94] H. Xu, M. Di Antonio, S. McKinney, V. Mathew, B. Ho, N. J. O'Neil, N. D. Santos, J. Silvester, V. Wei, J. Garcia, F. Kabeer, D. Lai, P. Soriano, J. Banath, D. S. Chiu, D. Yap, D. D. Le, F. B. Ye, A. Zhang, K. Thu, J. Soong, S. C. Lin, A. H. Tsai, T. Osako, T. Algara, D. N. Saunders, J. Wong, J. Xian, M. B. Bally, J. D. Brenton, G. W. Brown, S. P. Shah, D. Cescon, T. W. Mak, C. Caldas, P. C. Stirling, P. Hieter, S. Balasubramanian, S. Aparicio, *Nat Commun* **2017**, *8*, 14432.
- [95] A. Laguerre, K. Hukezalie, P. Winckler, F. Katranji, G. Chanteloup, M. Pirrotta, J. M. Perrier-Cornet, J. M. Wong, D. Monchaud, *J Am Chem Soc* **2015**, *137*, 8521-8525.
- [96] A. Laguerre, L. Stefan, M. Larrouy, D. Genest, J. Novotna, M. Pirrotta, D. Monchaud, *J Am Chem Soc* **2014**, *136*, 12406-12414.
- [97] A. Shivalingam, M. A. Izquierdo, A. L. Marois, A. Vysniauskas, K. Suhling, M. K. Kuimova, R. Vilar, *Nat Commun* **2015**, *6*, 8178.
- [98] P. Murat, Y. Singh, E. Defrancq, *Chem Soc Rev* **2011**, *40*, 5293-5307.
- [99] a) J.-L. Mergny, J.-C. Maurizot, *ChemBioChem* **2001**, *2*, 124-132; b) J. L. Mergny, L. Lacroix, M. P. Teulade-Fichou, C. Hounsou, L. Guittat, M. Hoarau, P. B. Arimondo, J. P. Vigneron, J. M. Lehn, J. F. Riou, T. Garestier, C. Helene, *Proc Natl Acad Sci USA* **2001**, *98*, 3062-3067.
- [100] a) D. Monchaud, C. Allain, M. P. Teulade-Fichou, *Nucleosides, Nucleotides and Nucleic Acids* **2007**, *26*, 1585-1588; b) D. Monchaud, C. Allain, H. Bertrand, N. Smargiasso, F. Rosu, V. Gabelica, A. De Cian, J. L. Mergny, M. P. Teulade-Fichou, *Biochimie* **2008**, *90*, 1207-1223.
- [101] D. L. Boger, B. E. Fink, S. R. Brunette, W. C. Tse, M. P. Hedrick, *J Am Chem Soc* **2001**, *123*, 5878-5891.
- [102] E. J. Wood, *Biochemical Education* **1994**, *22*, 83-83.
- [103] C. Bailly, P. Colson, J. P. Hénichart, C. Houssier, *Nucleic Acids Res* **1993**, *21*, 3705-3709.
- [104] S. Masiero, R. Trotta, S. Pieraccini, S. De Tito, R. Perone, A. Randazzo, G. P. Spada, *Org Biomol Chem* **2010**, *8*, 2683-2692.
- [105] A. I. Karsisiotis, N. M. Hessari, E. Novellino, G. P. Spada, A. Randazzo, M. Webba da Silva, *Angew Chem Int Ed Engl* **2011**, *50*, 10645-10648.
- [106] J. Carvalho, J. A. Queiroz, C. Cruz, *Journal of Chemical Education* **2017**.
- [107] J. Dash, P. S. Shirude, S. T. Hsu, S. Balasubramanian, *J Am Chem Soc* **2008**, *130*, 15950-15956.
- [108] M. A. Cooper, *Nat Rev Drug Discov* **2002**, *1*, 515-528.

- [109] J. E. Redman, *Methods* **2007**, *43*, 302-312.
- [110] D. P. N. Gonçalves, S. Ladame, S. Balasubramanian, J. K. M. Sanders, *Org Biomol Chem* **2006**, *4*, 3337-3342.
- [111] P. Murat, R. Bonnet, A. Van der Heyden, N. Spinelli, P. Labbe, D. Monchaud, M. P. Teulade-Fichou, P. Dumy, E. Defrancq, *Chemistry* **2010**, *16*, 6106-6114.
- [112] a) L. Bonnat, J. Dejeu, H. Bonnet, B. Gennaro, O. Jarjayes, F. Thomas, T. Lavergne, E. Defrancq, *Chemistry* **2016**, *22*, 3139-3147; b) R. Bonnet, T. Lavergne, B. Gennaro, N. Spinelli, E. Defrancq, *Chem Commun* **2015**, *51*, 4850-4853; c) L. Bonnat, L. Bar, B. Gennaro, H. Bonnet, O. Jarjayes, F. Thomas, J. Dejeu, E. Defrancq, T. Lavergne, *Chemistry* **2017**, *23*, 5602-5613.
- [113] F. Pillet, C. Romera, E. Trévisiol, S. Bellon, M.-P. Teulade-Fichou, J.-M. François, G. Pratviel, V. A. Leberre, *Sensors and Actuators B: Chemical* **2011**, *157*, 304-309.
- [114] B. Pagano, C. A. Mattia, C. Giancola, *Int J Mol Sci* **2009**, *10*, 2935-2957.
- [115] M. M. Pierce, C. S. Raman, B. T. Nall, *Methods* **1999**, *19*, 213-221.
- [116] M. Boncina, C. Podlipnik, I. Piantanida, J. Eilmes, M. P. Teulade-Fichou, G. Vesnaver, J. Lah, *Nucleic Acids Res* **2015**, *43*, 10376-10386.
- [117] M. Adrian, B. Heddi, A. T. Phan, *Methods* **2012**, *57*, 11-24.
- [118] A. T. Phan, Y. S. Modi, D. J. Patel, *J Mol Biol* **2004**, *338*, 93-102.
- [119] G. N. Parkinson, M. P. H. Lee, S. Neidle, *Nature* **2002**, *417*, 876-880.
- [120] A. T. Phan, V. Kuryavyi, H. Y. Gaw, D. J. Patel, *Nat Chem Biol* **2005**, *1*, 167-173.
- [121] V. Kuryavyi, A. T. Phan, D. J. Patel, *Nucleic Acids Res* **2010**, *38*, 6757-6773.
- [122] M. I. Onyshchenko, T. I. Gaynutdinov, E. A. Englund, D. H. Appella, R. D. Neumann, I. G. Panyutin, *Nucleic Acids Res* **2009**, *37*, 7570-7580.
- [123] S. Amrane, M. Adrian, B. Heddi, A. Serero, A. Nicolas, J. L. Mergny, A. T. Phan, *J Am Chem Soc* **2012**, *134*, 5807-5816.
- [124] G. N. Parkinson, R. Ghosh, S. Neidle, *Biochemistry* **2007**, *46*, 2390-2397.
- [125] F. Rusconi, *Manuel de spectrométrie de masse à l'usage des biochimistes*, Lavoisier.
- [126] a) F. Rosu, E. De Pauw, V. Gabelica, *Biochimie* **2008**, *90*, 1074-1087; b) A. Marchand, D. Strzelecka, V. Gabelica, *Chemistry* **2016**, *22*, 9551-9555; c) M. J. Lecours, A. Marchand, A. Anwar, C. Guetta, W. S. Hopkins, V. Gabelica, *Biochim Biophys Acta* **2017**, *1861*, 1353-1361.
- [127] V. Gabelica, F. Rosu, E. De Pauw, *Anal Chem* **2009**, *81*, 6708-6715.
- [128] F. Rosu, V. Gabelica, H. Poncelet, E. De Pauw, *Nucleic Acids Res* **2010**, *38*, 5217-5225.
- [129] N. Kim, M. Piatyszek, K. Prowse, C. Harley, M. West, P. Ho, G. Coviello, W. Wright, S. Weinrich, J. Shay, *Science* **1994**, *266*, 2011-2015.
- [130] A. De Cian, G. Cristofari, P. Reichenbach, E. De Lemos, D. Monchaud, M.-P. Teulade-Fichou, K. Shin-ya, L. Lacroix, J. Lingner, J.-L. Mergny, *Proceedings of the National Academy of Sciences* **2007**, *104*, 17347-17352.
- [131] M. N. Weitzmann, K. J. Woodford, K. Usdin, *Journal of Biological Chemistry* **1996**, *271*, 20958-20964.
- [132] P. Umate, N. Tuteja, R. Tuteja, *Commun Integr Biol* **2011**, *4*, 118-137.
- [133] R. M. Brosh, Jr., *Nat Rev Cancer* **2013**, *13*, 542-558.
- [134] O. Mendoza, A. Bourdoncle, J. B. Boule, R. M. Brosh, Jr., J. L. Mergny, *Nucleic Acids Res* **2016**, *44*, 1989-2006.
- [135] S. K. Bharti, J. A. Sommers, F. George, J. Kuper, F. Hamon, K. Shin-ya, M. P. Teulade-Fichou, C. Kisker, R. M. Brosh, Jr., *J Biol Chem* **2013**, *288*, 28217-28229.
- [136] O. Mendoza, N. M. Gueddouda, J. B. Boule, A. Bourdoncle, J. L. Mergny, *Nucleic Acids Res* **2015**, *43*, e71.
- [137] N. M. Gueddouda, O. Mendoza, D. Gomez, A. Bourdoncle, J. L. Mergny, *Biochim Biophys Acta* **2017**, *1861*, 1382-1388.
- [138] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angew Chem Int Edit* **2001**, *40*, 2004-2021.
- [139] a) C. W. Tornøe, C. Christensen, M. Meldal, *J Org Chem* **2002**, *67*, 3057-3064; b) V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angew Chem Int Edit* **2002**, *41*, 2596-2599.

- [140] N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J Am Chem Soc* **2004**, *126*, 15046-15047.
- [141] R. Huisgen, *Angewandte Chemie International Edition in English* **1963**, *2*, 565-598.
- [142] a) R. Huisgen, in *1, 3-Dipolar Cycloaddition Chemistry*, Wiley, **1984**, pp. 1-176; b) Z.-X. Wang, H.-L. Qin, *Chem Commun* **2003**, 2450.
- [143] a) Z. P. Demko, K. B. Sharpless, *Angew Chem Int Ed* **2002**, *41*, 2113-2116; b) Z. P. Demko, K. B. Sharpless, *Angew Chem Int Ed* **2002**, *41*, 2110-2113.
- [144] F. Himo, T. Lovell, R. Hilgraf, V. V. Rostovtsev, L. Noodleman, K. B. Sharpless, V. V. Fokin, *J Am Chem Soc* **2005**, *127*, 210-216.
- [145] V. O. Rodionov, V. V. Fokin, M. G. Finn, *Angew Chem Int Ed Engl* **2005**, *44*, 2210-2215.
- [146] B. T. Worrell, J. A. Malik, V. V. Fokin, *Science* **2013**, *340*, 457.
- [147] W. Jiang, D. Liu, H. Li, *Science of The Total Environment* **2000**, *256*, 59-65.
- [148] a) V. O. Rodionov, S. I. Presolski, D. Díaz Díaz, V. V. Fokin, M. G. Finn, *J Am Chem Soc* **2007**, *129*, 12705-12712; b) T. R. Chan, R. Hilgraf, K. B. Sharpless, V. V. Fokin, *Org Lett* **2004**, *6*, 2853-2855.
- [149] a) A. J. Link, D. A. Tirrell, *J Am Chem Soc* **2003**, *125*, 11164-11165; b) A. J. Link, M. K. S. Vink, D. A. Tirrell, *J Am Chem Soc* **2004**, *126*, 10598-10602; c) Q. Wang, T. R. Chan, R. Hilgraf, V. V. Fokin, K. B. Sharpless, M. G. Finn, *J Am Chem Soc* **2003**, *125*, 3192-3193.
- [150] A. E. Speers, G. C. Adam, B. F. Cravatt, *J Am Chem Soc* **2003**, *125*, 4686-4687.
- [151] D. C. Kennedy, C. S. McKay, M. C. B. Legault, D. C. Danielson, J. A. Blake, A. F. Pegoraro, A. Stolorow, Z. Mester, J. P. Pezacki, *J Am Chem Soc* **2011**, *133*, 17993-18001.
- [152] J. F. Lutz, *Angew Chem Int Ed Engl* **2008**, *47*, 2182-2184.
- [153] N. J. Agard, J. M. Baskin, J. A. Prescher, A. Lo, C. R. Bertozzi, *ACS Chem Biol* **2006**, *1*, 644-648.
- [154] a) J. M. Baskin, J. A. Prescher, S. T. Laughlin, N. J. Agard, P. V. Chang, I. A. Miller, A. Lo, J. A. Codelli, C. R. Bertozzi, *Proc Natl Acad Sci U S A* **2007**, *104*, 16793-16797; b) P. Shieh, C. R. Bertozzi, *Org Biomol Chem* **2014**, *12*, 9307-9320; c) E. M. Sletten, C. R. Bertozzi, *Acc Chem Res* **2011**, *44*, 666-676.
- [155] a) X. Ning, J. Guo, M. A. Wolfert, G. J. Boons, *Angew Chem Int Ed Engl* **2008**, *47*, 2253-2255; b) K. E. Beatty, J. D. Fisk, B. P. Smart, Y. Y. Lu, J. Szychowski, M. J. Hangauer, J. M. Baskin, C. R. Bertozzi, D. A. Tirrell, *Chembiochem* **2010**, *11*, 2092-2095.
- [156] J. Dommerholt, F. P. Rutjes, F. L. van Delft, *Top Curr Chem (Cham)* **2016**, *374*, 16.
- [157] J. Dommerholt, O. van Rooijen, A. Borrmann, C. F. Guerra, F. M. Bickelhaupt, F. L. van Delft, *Nat Commun* **2014**, *5*, 5378.
- [158] a) S. S. van Berkel, A. T. Dirks, M. F. Debets, F. L. van Delft, J. J. Cornelissen, R. J. Nolte, F. P. Rutjes, *Chembiochem* **2007**, *8*, 1504-1508; b) M. F. Debets, J. C. van Hest, F. P. Rutjes, *Org Biomol Chem* **2013**, *11*, 6439-6455.
- [159] I. Singh, Z. Zarafshani, J.-F. Lutz, F. Heaney, *Macromolecules* **2009**, *42*, 5411-5413.
- [160] a) F. Thalhammer, U. Wallfahrer, J. Sauer, *Tetrahedron Letters* **1990**, *31*, 6851-6854; b) J. Sauer, D. K. Heldmann, J. Hetzenegger, J. Krauthan, H. Sichert, J. Schuster, *European Journal of Organic Chemistry* **1998**, *1998*, 2885-2896.
- [161] N. K. Devaraj, R. Weissleder, S. A. Hilderbrand, *Bioconjug Chem* **2008**, *19*, 2297-2299.
- [162] C. G. Gordon, J. L. Mackey, J. C. Jewett, E. M. Sletten, K. N. Houk, C. R. Bertozzi, *J Am Chem Soc* **2012**, *134*, 9199-9208.
- [163] a) M. L. Blackman, M. Royzen, J. M. Fox, *J Am Chem Soc* **2008**, *130*, 13518-13519; b) M. T. Taylor, M. L. Blackman, O. Dmitrenko, J. M. Fox, *J Am Chem Soc* **2011**, *133*, 9646-9649.
- [164] B. L. Oliveira, Z. Guo, G. J. L. Bernardes, *Chem Soc Rev* **2017**, *46*, 4895-4950.
- [165] A. F. Larsen, T. Ulven, *Org Lett* **2011**, *13*, 3546-3548.
- [166] Z.-Q. Hu, H.-Y. Hu, C.-F. Chen, *J Org Chem* **2006**, *71*, 1131-1138.
- [167] A. F. Larsen, M. C. Nielsen, T. Ulven, *Chemistry* **2012**, *18*, 10892-10902.
- [168] J. Lefebvre, C. Guetta, F. Poyer, F. Mahuteau-Betzer, M. P. Teulade-Fichou, *Angew Chem Int Ed Engl* **2017**, *38*, 11365-11369.
- [169] P. Yang, A. De Cian, M. P. Teulade-Fichou, J. L. Mergny, D. Monchaud, *Angew Chem Int Ed* **2009**, *48*, 2188-2191.

- [170] E. Largy, F. Hamon, M.-P. Teulade-Fichou, *Methods* **2012**, *57*, 129-137.
- [171] D. Verga, F. Hamon, F. Poyer, S. Bombard, M. P. Teulade-Fichou, *Angew Chem Int Ed Engl* **2014**, *53*, 994-998.
- [172] A. Rajendran, M. Endo, K. Hidaka, M. P. Teulade-Fichou, J. L. Mergny, H. Sugiyama, *Chem Commun* **2015**, *51*, 9181-9184.
- [173] A. Renaud de la Faverie, F. Hamon, C. Di Primo, E. Largy, E. Dausse, L. Delauriere, C. Landras-Guetta, J. J. Toulme, M. P. Teulade-Fichou, J. L. Mergny, *Biochimie* **2011**, *93*, 1357-1367.
- [174] A. Ambrus, D. Chen, J. Dai, R. A. Jones, D. Yang, *Biochemistry* **2005**, *44*, 2048-2058.
- [175] a) J. Dai, D. Chen, R. A. Jones, L. H. Hurley, D. Yang, *Nucleic Acids Res* **2006**, *34*, 5133-5144; b) P. Agrawal, C. Lin, R. I. Mathad, M. Carver, D. Yang, *J Am Chem Soc* **2014**, *136*, 1750-1753.
- [176] A. De Rache, J. L. Mergny, *Biochimie* **2015**, *115*, 194-202.
- [177] R. J. Trachman, 3rd, N. A. Demeshkina, M. W. L. Lau, S. S. S. Panchapakesan, S. C. Y. Jeng, P. J. Unrau, A. R. Ferre-D'Amare, *Nat Chem Biol* **2017**, *13*, 807-813.
- [178] N. Q. Do, A. T. Phan, *Chemistry* **2012**, *18*, 14752-14759.
- [179] R. Sahai, G. L. Loper, S. H. Lin, H. Eyring, *Proc Natl Acad Sci U S A* **1974**, *71*, 1499-1503.
- [180] E. Largy, F. Hamon, M. P. Teulade-Fichou, *Anal Bioanal Chem* **2011**, *400*, 3419-3427.
- [181] P. Murat, S. Balasubramanian, *Current Opinion in Genetics & Development* **2014**, *25*, 22-29.
- [182] C. Schaffitzel, I. Berger, J. Postberg, J. Hanes, H. J. Lipps, A. Pluckthun, *Proceedings of the National Academy of Sciences* **2001**, *98*, 8572-8577.
- [183] K. Paeschke, T. Simonsson, J. Postberg, D. Rhodes, H. J. Lipps, *Nature Structural Molecular Biology* **2005**, *12*, 847-854.
- [184] G. Biffi, D. Tannahill, J. McCafferty, S. Balasubramanian, *Nat Chem* **2013**, *5*, 182-186.
- [185] a) A. Henderson, Y. Wu, Y. C. Huang, E. A. Chavez, J. Platt, F. B. Johnson, R. M. Brosh, Jr., D. Sen, P. M. Lansdorp, *Nucleic Acids Res* **2014**, *42*, 860-869; b) R. F. Hoffmann, Y. M. Moshkin, S. Mouton, N. A. Grzeschik, R. D. Kalicharan, J. Kuipers, A. H. Wolters, K. Nishida, A. V. Romashchenko, J. Postberg, H. Lipps, E. Berezikov, O. C. Sibon, B. N. Giepmans, P. M. Lansdorp, *Nucleic Acids Res* **2016**, *44*, 152-163.
- [186] H. G. Kazemier, K. Paeschke, P. M. Lansdorp, *Nucleic Acids Res* **2017**, *45*, 5913-5919.
- [187] H. Y. Liu, Q. Zhao, T. P. Zhang, Y. Wu, Y. X. Xiong, S. K. Wang, Y. L. Ge, J. H. He, P. Lv, T. M. Ou, J. H. Tan, D. Li, L. Q. Gu, J. Ren, Y. Zhao, Z. S. Huang, *Cell Chem Biol* **2016**.
- [188] U. Schnell, F. Dijk, K. A. Sjollem, B. N. Giepmans, *Nat Methods* **2012**, *9*, 152-158.
- [189] E. Largy, A. Granzhan, F. Hamon, D. Verga, M. P. Teulade-Fichou, *Top Curr Chem* **2013**, *330*, 111-177.
- [190] a) A. J. Pickard, U. Bierbach, *ChemMedChem* **2013**, *8*, 1441-1449; b) K. Burger, B. Muhl, T. Harasim, M. Rohmoser, A. Malamoussi, M. Orban, M. Kellner, A. Gruber-Eber, E. Kremmer, M. Holzels, D. Eick, *J Biol Chem* **2010**, *285*, 12416-12425.
- [191] a) I. Lubitz, D. Zikich, A. Kotlyar, *Biochemistry* **2010**, *49*, 3567-3574; b) A. Laguerre, J. M. Wong, D. Monchaud, *Sci Rep* **2016**, *6*, 32141; c) F. Doria, M. Nadai, M. Zuffo, R. Perrone, M. Freccero, S. N. Richter, *Chem Commun* **2017**, *53*, 2268-2271.
- [192] M. Di Antonio, G. Biffi, A. Mariani, E. A. Raiber, R. Rodriguez, S. Balasubramanian, *Angew Chem Int Ed Engl* **2012**, *51*, 11073-11078.
- [193] C. Ribeyre, J. Lopes, J. B. Boule, A. Piazza, A. Guedin, V. A. Zakian, J. L. Mergny, A. Nicolas, *PLoS Genet* **2009**, *5*, e1000475.
- [194] R. Chennoufi, H. Bougherara, N. Gagey-Eilstein, B. Dumat, E. Henry, F. Subra, F. Mahuteau-Betzer, P. Tauc, M. P. Teulade-Fichou, E. Deprez, *Chem Commun* **2015**, *51*, 14881-14884.
- [195] a) E. Kruisselbrink, V. Guryev, K. Brouwer, D. B. Pontier, E. Cuppen, M. Tijsterman, *Current Biology* **2008**, *18*, 900-905; b) M. Di Antonio, R. Rodriguez, S. Balasubramanian, *Methods* **2012**, *57*, 84-92; c) J.-B. Vannier, V. Pavicic-Kaltenbrunner, Mark I. R. Petalcorin, H. Ding, Simon J. Boulton, *Cell* **2012**, *149*, 795-806; d) W. Koole, R. van Schendel, A. E. Karambelas, J. T. van Heteren, K. L. Okihara, M. Tijsterman, *Nat Commun* **2014**, *5*.
- [196] N. A. Di Prospero, K. H. Fischbeck, *Nat Rev Genet* **2005**, *6*, 756-767.

- [197] S. Amrane, M. Adrian, B. Heddi, A. Serero, A. Nicolas, J.-L. Mergny, A. T. Phan, *J Am Chem Soc* **2012**, *134*, 5807-5816.
- [198] K. Paeschke, M. L. Bochman, P. D. Garcia, P. Cejka, K. L. Friedman, S. C. Kowalczykowski, V. A. Zakian, *Nature* **2013**, *497*, 458-462.
- [199] a) A. Piazza, J. B. Boule, J. Lopes, K. Mingo, E. Largy, M.-P. Teulade-Fichou, A. Nicolas, *Nucleic Acids Research* **2010**, *38*, 4337-4348; b) J. Lopes, A. Piazza, R. Bermejo, B. Kriegsman, A. Colosio, M.-P. Teulade-Fichou, M. Foiani, A. Nicolas, *The EMBO Journal* **2011**, *30*, 4033-4046.
- [200] A. Piazza, M. Adrian, F. Samazan, B. Heddi, F. Hamon, A. Serero, J. Lopes, M.-P. Teulade-Fichou, A. T. Phan, A. Nicolas, *The EMBO Journal* **2015**, *34*, 1718-1734.
- [201] F. Sanger, S. Nicklen, A. R. Coulson, *Proceedings of the National Academy of Sciences of the United States of America* **1977**, *74*, 5463-5467.
- [202] A. M. Maxam, W. Gilbert, *Proceedings of the National Academy of Sciences* **1977**, *74*, 560-564.
- [203] C. Anandhakumar, S. Kizaki, T. Bando, G. N. Pandian, H. Sugiyama, *ChemBioChem* **2015**, *16*, 20-38.
- [204] V. S. Chambers, G. Marsico, J. M. Boutell, M. Di Antonio, G. P. Smith, S. Balasubramanian, *Nat Biotech* **2015**, advance online publication.
- [205] a) A. Guedin, J. Gros, P. Alberti, J. L. Mergny, *Nucleic Acids Research* **2010**, *38*, 7858-7868; b) S. L. Palumbo, S. W. Ebbinghaus, L. H. Hurley, *J Am Chem Soc* **2009**, *131*, 10878-10891.
- [206] a) V. T. Mukundan, A. T. Phan, *J Am Chem Soc* **2013**, *135*, 5017-5028; b) B. De Nicola, C. J. Lech, B. Heddi, S. Regmi, I. Frasson, R. Perrone, S. N. Richter, A. T. Phan, *Nucleic Acids Res* **2016**, *44*, 6442-6451.
- [207] R. Hansel-Hertsch, D. Beraldi, S. V. Lensing, G. Marsico, K. Zyner, A. Parry, M. Di Antonio, J. Pike, H. Kimura, M. Narita, D. Tannahill, S. Balasubramanian, *Nat Genet* **2016**, *48*, 1267-1272.
- [208] K. T. Barglow, B. F. Cravatt, *Nat Methods* **2007**, *4*, 822-827.
- [209] a) Y. Takaoka, A. Ojida, I. Hamachi, *Angew Chem Int Ed Engl* **2013**, *52*, 4088-4106; b) R. K. Lim, Q. Lin, *Chem Commun* **2010**, *46*, 1589-1600.
- [210] a) V. Bevilacqua, M. King, M. Chaumontet, M. Nothisen, S. Gabillet, D. Buisson, C. Puente, A. Wagner, F. Taran, *Angew Chem Int Ed Engl* **2014**, *53*, 5872-5876; b) C. Uttamapinant, A. Tangpeerachaikul, S. Grecian, S. Clarke, U. Singh, P. Slade, K. R. Gee, A. Y. Ting, *Angew Chem Int Edit* **2012**, *51*, 5852-5856.
- [211] M. Fatfat, R. A. Merhi, O. Rahal, D. A. Stoyanovsky, A. Zaki, H. Haidar, V. E. Kagan, H. Galimuntasib, K. Machaca, *BMC Cancer* **2014**, *14*, 527.
- [212] J. Mas Pons, A. Dumont, G. Sautejeau, E. Fugier, A. Baron, S. Dukan, B. Vauzeilles, *Angew Chem Int Ed Engl* **2014**, *53*, 1275-1278.
- [213] I. Nikic, J. H. Kang, G. E. Girona, I. V. Aramburu, E. A. Lemke, *Nat Protoc* **2015**, *10*, 780-791.
- [214] H. E. Murrey, J. C. Judkins, C. W. Am Ende, T. E. Ballard, Y. Fang, K. Riccardi, L. Di, E. R. Guilmette, J. W. Schwartz, J. M. Fox, D. S. Johnson, *J Am Chem Soc* **2015**, *137*, 11461-11475.
- [215] M. V. Bharathi, M. Chhabra, P. Paira, *Bioorg Med Chem Lett* **2015**, *25*, 5737-5742.
- [216] M. L. Turski, D. J. Thiele, *J Biol Chem* **2009**, *284*, 717-721.
- [217] S. Yoshida, Y. Hatakeyama, K. Johmoto, H. Uekusa, T. Hosoya, *J Am Chem Soc* **2014**, *136*, 13590-13593.
- [218] R. K. Pathak, C. D. McNitt, V. V. Popik, S. Dhar, *Chemistry* **2014**, *20*, 6861-6865.
- [219] J. R. Lakowicz, *Principles of fluorescence spectroscopy*, Springer, **2006**.

Titre : Outils moléculaires pour l'étude des G-quadruplex au sein du génome

Mots clés : G-quadruplex – Ligands – Interactions – Imagerie cellulaire

Résumé : L'acide désoxyribonucléique se structure chez les êtres vivants de différentes façons. La plus connue est sa forme double hélice mais de nombreuses autres structures secondaires existent et notamment les G-quadruplex. Il s'agit d'une structure basée sur le repliement d'un brin d'ADN possédant des répétitions de guanines. L'association de quatre guanines entre elles par liaisons hydrogène forme un plan appelé G-quartet. Ce réseau de liaisons hydrogène est appelé appariement de Hoogsteen. L'empilement d'au moins deux quartets autour d'un cation monovalent comme le potassium ou le sodium constitue la structure G-quadruplex. Ces structures ont été très étudiées lors des vingt dernières années et il a été montré qu'elles sont impliquées dans de nombreux mécanismes biologiques tels que la réplication, la transcription, la traduction et également le maintien des télomères. La présence des G-quadruplex peut provoquer une instabilité importante aussi bien génétique qu'épigénétique. C'est pourquoi de nombreuses méthodes ont été développées afin de localiser et comprendre le rôle de ces structures *in vivo*. Pour cela, un large panel d'outils moléculaires a été utilisé cependant il est encore difficile, à partir de ce panel, d'apporter une réponse à toutes les questions sur l'implication des G-quadruplex au niveau du génome. Lors de ce travail de thèse, nous avons alors développés de nouvelles molécules capables de cibler sélectivement les G-quadruplex au sein d'un milieu biologique complexe à partir de deux ligands PDC et PhenDC₃ affins et sélectifs pour les structures G-quadruplex.

Sur la base de molécules de référence que sont PhenDC₃ et PDC, de nombreux ligands ont été mis au point. D'une part, des ligands

fonctionnalisés avec une biotine et/ou un groupement photoactivable ont été synthétisés afin de capturer et d'extraire des structures G-quadruplex dans un milieu biologique. D'autre part, des dérivés capables d'être fonctionnalisés *in cellulo* par l'utilisation de chimie bioorthogonale ont également été obtenus. Ceci permet d'ajouter une fonction (fluorescente ou biotine...) après que le dérivé ait interagi avec sa cible cellulaire. L'ensemble des composés a été évalué par des techniques biophysiques, l'expérience de FRET-melting et l'expérience de FID, afin de mesurer leur affinité pour différentes structures G-quadruplex et leur sélectivité. Nous avons proposé une relation entre les deux expériences afin d'avoir un classement de ligands le plus approprié pour les G-quadruplex.

Un des objectifs majeurs de ce travail était de localiser les ligands de G-quadruplex au sein de cellules cancéreuses humaines. Dans un premier temps, toute une étude au sein de cellules fixées a été réalisée en utilisant deux réactions de chimie « click », la réaction de cycloaddition d'un azoture et d'un alcyne catalysée par le cuivre (CuAAC) et la réaction de cycloaddition d'une cyclooctyne et d'un azoture (SPAAC). L'étude s'est, dans un second temps, poursuivie au sein de cellules vivantes en utilisant uniquement la réaction SPAAC à cause de la toxicité *in cellulo* du cuivre.

Ces composés ont également été testés pour l'extraction de G-quadruplex à l'aide de billes magnétiques recouvertes d'une fonction cyclooctyne. Cependant, les résultats observés, lors de cette étude préliminaire, n'ont pas été concluants et demandent une mise au point pour optimiser le système.

Title: Molecular tools for the study of G-quadruplex in the genome

Keywords: G-quadruplex – Ligands – Interactions – Cellular imaging

Abstract: Deoxyribonucleic acid has different structures in human beings. The most known is the double helix but a lot of secondary structures exist and particularly G-quadruplex. It consists of guanine-rich nucleic acid sequences. The association of four guanines through hydrogen bonds forms a plan called G-quartet. This set of hydrogen bonds is called Hoogsteen base pairs. The stacking of at least two quartets around a monovalent cation like potassium or sodium establishes the G-quadruplex. These structures have been much studied over the past twenty years. They are involved in numerous biological mechanisms like replication, transcription, translation and also telomere maintenance. G-quadruplex presence can cause an important genetic as well as epigenetic instability. That is why many methods have been developed in order to localize these structures and to understand their role *in vivo*. To this end, a broad panel of molecular tools has been used. However, it is still difficult to bring an answer to all the questions about the involvement of G-quadruplex at the genomic level with this panel. In this thesis work, we developed new molecular tools able to target selectively G-quadruplex in a complex biological medium from two benchmark ligands, PhenDC₃ and PDC, which have very good affinity and selectivity for G-quadruplex.

On the one hand, functionalized ligands have been synthesized with

a biotin and/or a photoactivatable group in order to trap and pull-down G-quadruplex in various cellular contexts. On the other hand, derivative compounds which are able to be functionalized *in cellulo* by bioorthogonal reactions have been obtained. Once the compound interacts with its cellular target, a function (fluorophore or biotin) can be added through an orthogonal reaction. The new panel of compounds has been evaluated by biophysical techniques, FRET-melting experiment and FID assay, in order to determine their affinity to G-quadruplex and their selectivity. We proposed a relation between the two biophysical experiments in order to have a good ranking of ligands for G-quadruplex structures.

One of the most important objectives of this work was to localize G-quadruplex ligands in human cancer cells. First, a complete study in fixed cells has been performed using two reactions of click chemistry: reaction of copper-catalyzed-alkyne-azide-cycloaddition (CuAAC) and reaction of strain-promoted alkyne-azide cycloaddition (SPAAC). Secondly, the study has been pursued in living cells using SPAAC reaction because of the toxicity of copper in cells.

These compounds have also been used to extract G-quadruplex from biological systems with cyclooctyne-coated magnetic beads. However, results obtained in this preliminary study are not decisive so it could be interesting to optimize the system before concluding.

