



UNIVERSITE DE BOURGOGNE

UFR Sciences et Techniques

Département de Chimie

Institut de Chimie Moléculaire de l'Université de Bourgogne

THESE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE DE BOURGOGNE

Mention : Chimie

par

Loïc STEFAN

Soutenue le 04 décembre 2013

Template-Assembled Synthetic G-Quartets (TASQ) hydrosolubles : du ligand de quadruplexes d'ADN et d'ARN à la plateforme catalytique

devant la commission d'examen

S. Ladame
R. Rodriguez
J.-L. Mergny
E. Defrancq
F. Denat
D. Monchaud

Joint Lecturer, Imperial College of London
Chargé de Recherche, CNRS
Directeur de Recherche, INSERM
Professeur, Université de Grenoble
Professeur, Université de Bourgogne
Chargé de Recherche, CNRS

Rapporteur
Rapporteur
Examineur
Examineur
Codirecteur de thèse
Directeur de thèse

« Il y aura toujours dans la foule un crétin qui, sous prétexte qu'il ne comprend pas,
décrétera qu'il n'y a rien à comprendre »

Amélie Nothomb

« Pour la chimie, les cieux sont grand ouverts, parce que si elle est une science, elle est aussi un art. Par la beauté de ses objets, évidemment, mais aussi par son essence même, par son aptitude à inventer le futur et à se recréer soi-même, sans fin.

Comme l'artiste, le chimiste grave dans la matière les produits de son imagination créatrice. La pierre, les sons, les mots ne contiennent pas les œuvres que le sculpteur, le compositeur, l'écrivain en tirent. De même, le chimiste crée des molécules originales, des matériaux nouveaux et des propriétés nouvelles à partir des éléments fournis par la nature, en fait des mondes entièrement nouveaux, qui n'existaient pas avant d'avoir été mis en forme par les mains du chimiste, comme la matière est mise en forme par la main de l'artiste.

La chimie possède effectivement cette puissance créatrice, ainsi que l'a affirmé Marcelin Berthelot: "la chimie crée son objet". Elle ne se contente pas de fabriquer des objets, elle crée son propre objet qui ne préexiste pas, mais est inventé progressivement, qui n'a pas à être simplement découvert, mais qui est à créer. »

Jean-Marie Lehn
Prix Nobel de Chimie, 1987.

REMERCIEMENTS

Ce travail de recherche a été réalisé à l'Institut de Chimie Moléculaire de l'Université de Bourgogne, unité mixte de recherche entre l'Université de Bourgogne et le Centre National de la Recherche Scientifique, UMR6302.

Je remercie le CNRS, pour le financement de ma bourse de thèse qui m'a permis de mener cette thèse pendant ces trois années, ainsi que pour m'avoir offert l'incroyable opportunité de vivre une formidable expérience, personnelle et scientifique, auprès de 34 lauréats du prix Nobel de tous horizons au cours du Lindau Nobel Meeting de 2013.

Je tiens tout d'abord à exprimer ma profonde reconnaissance au Dr. Sylvain Ladame, lecturer, d'avoir traversé la Manche depuis l'Imperial College of London et d'avoir accepté de juger ce travail, ainsi qu'au Dr. Raphaël Rodriguez, chargé de recherche à l'ICSN de Gif-sur-Yvette, pour m'avoir accordé de son temps et d'avoir accepté d'être rapporteur de cette thèse.

Je remercie très sincèrement le Dr. Jean-Louis Mergny, directeur de recherche à l'Institut Européen de Chimie Biologie de Pessac, pour sa contribution à ce projet de recherche, pour son accueil au sein de son laboratoire au cours du mois de février 2013, ainsi que pour l'honneur qu'il me fait en acceptant d'examiner ce travail, réalisé dans l'autre grande région du vin français.

J'exprime ma profonde gratitude au Pr. Eric Defrancq, professeur à l'Université de Grenoble, pour sa collaboration à ce travail de thèse, pour son accueil chaleureux au sein du laboratoire I2BM en avril 2013, et pour me faire l'immense plaisir de participer à l'examen de ce mémoire.

Je remercie le Pr. F. Denat, directeur de l'ICMUB, pour avoir accepté la codirection de cette thèse et pour m'avoir fait confiance, depuis le début, pour mener à bien ce travail de recherche sur une thématique nouvelle et originale au sein du laboratoire.

Je tiens particulièrement à remercier le Dr. David Monchaud, chargé de recherche et directeur de cette thèse, pour ces incroyables années passées. Je le remercie pour m'avoir donné toute sa confiance ainsi qu'une immense liberté pour réaliser ce projet de recherche qui m'a passionné depuis le début. Son approche scientifique, sa culture bibliographique, sa pédagogie et nos longues discussions le soir dans le laboratoire m'ont permis de me forger ma personnalité scientifique, et je

lui en suis extrêmement reconnaissant. En parallèle de la recherche, nos bavardages philosophiques, littéraires et musicaux souvent accompagnés d'un café ont réussi à créer, au-delà du cercle professionnel, une réelle amitié. Je n'aurai pu imaginer un meilleur encadrement de thèse.

Je remercie également les Dr. Nicolas Spinelli et Thomas Lavergne, pour m'avoir accueilli au sein du Département de Chimie Moléculaire de l'Université de Grenoble où j'ai eu la chance de m'initier à la technique de SPR grâce à l'aide et à la bonne humeur du Dr. Enora Prado.

Au cours de mon passage bordelais à l'IECB, j'ai eu la chance d'être encadré par Aurore Guédin que je remercie pour m'avoir accordé beaucoup de son temps pour m'aider à percer les secrets de l'électrophorèse sur gel. J'associe également mes remerciements à tout le laboratoire pour leur accueil chaleureux, et je souhaite en particulier une bonne continuation à Amandine Renaud de la Faverie.

J'envoie également mes remerciements au Dr. Richard Decréau, chaire d'excellence à l'ICMUB, pour nos nombreuses discussions scientifiques très enrichissantes partagées dans une bonne humeur communicative. Bien qu'il m'ait souvent confondu avec Lorànt Deutsch, sa passion pour la numismatique et son appétit pour la gaufre MgSO_4 font de lui le président virtuel de cette thèse.

Je remercie le Dr. Romain Haudecoeur, pour nos nombreuses discussions autour du Gringo et pour avoir toujours su regrouper autour de lui tous les ingrédients du succès auxquels je n'oublie pas d'associer Gaëlle et Mr. Chou.

Je n'oublie pas de remercier chaleureusement le Dr. David Genest pour cette dernière année passée ensemble au laboratoire, pour sa culture internet parallèle, son imitation de Bourvil, son réveil matinal, sa relecture incroyablement méticuleuse de ce mémoire, et son soutien lors de la rédaction de cette thèse. Je le remercie pour sa bonne humeur et lui souhaite le meilleur pour la suite.

Je remercie bien sûr le Dr. Manuel Larrouy, pour son accent du sud et sa jovialité communicative, ainsi que pour ses fabuleux sushi et sa machine à expresso georgelocloonesque. J'y associe bien évidemment Stéphanie pour tous ces moments passés ensemble.

Je tiens particulièrement à remercier le Dr. Alejandra Sornosa Ten avec qui j'ai partagé mon bureau tout au long de la rédaction de cette thèse et qui a toujours réussi à amener tout le soleil et la

chaleur espagnols pour me mettre de bonne humeur. Son savoir-faire inimitable pour la sangria aux dosages plus qu'approximatifs nous a toujours fait passer de merveilleuses soirées.

Je n'oublie pas de remercier le Dr. Victor Goncalves pour sa disponibilité, sa pédagogie, son enthousiasme, sa piscine de ballons, son amour pour les petits chats, et pour les nombreux group meeting dans son appartement.

Je remercie également tous les stagiaires que j'ai eu la chance d'encadrer au cours de cette thèse, et plus particulièrement Julie, Alexandre, Marine ainsi que Pauline à qui je souhaite une belle suite scientifique en région parisienne.

Je tiens à remercier Marie-Jo Penouilh, pour son temps passé à chercher à ioniser mes échantillons d'ADN dans l'ESI, et pour s'être toujours rendue disponible avec bonne humeur.

Je remercie bien entendu Marc Pirrotta, le papa des étudiants, et le Dr. PDG Claire Bernhard avec qui j'ai pu partager les chemins tortueux de la création d'entreprise.

Je souhaite remercier particulièrement le Dr. Maryline Beyler pour nos nombreux ciné & resto ainsi que le Dr. Nicolas Weibel pour son immense culture musicale.

J'associe également à mes remerciements tous les membres de l'ICMUB, permanents ou non, qui ont su rendre mon quotidien agréable. Merci à Pauline D., Antoine, Guillaume P., Stéphane, Semra, Guillaume O., Matthieu, Bertrand, Yann, Biquet, Clément, Trong, Pawel, Emilie, Margot, Sophie, Ewen, Christine, Frédéric, Claire, Stéphanie, Alla, Justine, Nicolas M., Mylène du privé, ainsi qu'Aurélien à qui je souhaite bon courage.

Je n'oublie pas d'associer à mes remerciements Jean-Luc Vorng, pour notre amitié, son calme, nos escapades britanniques et écossaises ainsi que pour son soutien tout au long de ces années d'études.

Bien évidemment, j'associe ce travail de thèse à toutes les autres personnes non chimistes qui m'ont entouré depuis le début de cette aventure et qui ont rendu ma vie personnelle agréable.

Je remercie tout d'abord les « anciens » : Marielle et sa petite famille, Flo et Aurélie, Alexandre, Amélia à qui je souhaite une belle réussite, et bien sûr Jessica D. R. qui a toujours été pour moi un exemple de dépassement de soi et de courage dont je reste admiratif. Son soutien, sa

cuisine et sa créativité ainsi que l'air de la campagne m'ont toujours permis de me ressourcer et de ne pas perdre de vue l'essentiel, tout comme les barbecues de Doudou !

Je remercie également Thomas pour tous nos voyages inoubliables aux quatre coins du monde, et pour tous ceux qu'il nous reste à faire, et pour avoir découvert, ensemble, le merveilleux musée du sac (plus de 4000 !) d'Amsterdam.

Je tiens également à remercier mes amis de toujours, Denis et Sylvain avec qui j'ai partagé tant de discussions et de bons moments, ainsi que Camille, pour sa connaissance astronomique et poétique.

J'adresse une pensée particulière à Cathy et à Marie pour m'avoir apporté avec leurs notes de guitare et de violon certains de mes plus beaux moments, assis autour du piano.

Enfin, je remercie les membres de ma famille qui ont toujours cru en moi, en particulier mon grand-père Gérard et mon oncle Patrick et bien sûr, mes parents à qui je dois beaucoup et à qui je dois mon amour du travail bien fait. Je tiens bien sûr à remercier mes deux petits frères Rémi et Nicolas ainsi que ma petite sœur Laura, à qui je souhaite le meilleur et que j'embrasse.

Pour terminer, j'envoie mes derniers remerciements en direction de la Bretagne tout d'abord à Jessica pour son immense culture littéraire et artistique ainsi que pour les nombreuses discussions-débats autour d'un bon verre de vin. En dernier lieu, je tiens à remercier Jordy pour sa confiance, son soutien sans faille, son enthousiasme et son humanité dont il a su faire preuve tout au long de mes études et qui a toujours été là pour partager à la fois mes réussites et surtout les moments de doutes. Alors que ma thèse s'achève la sienne commence, et je lui souhaite le meilleur pour toutes ces belles années à venir.

SOMMAIRE

ABREVIATIONS	1
INTRODUCTION GENERALE.....	5
Chapitre 1 – DE L’ADN CELLULAIRE AUX ASSEMBLAGES SUPRAMOLECULAIRES DE NUCLEOBASES	11
I. Introduction.....	11
II. Prérequis : biochimie élémentaire	13
II.1 La cellule eucaryote et le cycle cellulaire	13
II.2 Rôles et fonctions de l’ADN et de l’ARN	16
III. Les structure de l’ADN et de l’ARN	24
III.1 L’histoire de la double hélice d’ADN	24
III.2 Structures usuelles d’ADN et d’ARN	26
III.3 Structures non-usuelles.....	33
III.4 Bilan des structures d’ADN et d’ARN	44
IV. De l’ADN naturel à la molécule supramoléculaire	44
IV.1 Les nucléobases comme motifs supramoléculaires	45
IV.2 Auto-assemblages de tétrades de guanines synthétiques.....	51
V. Les tétrades de guanines intramoléculaires, ou TASQ.....	57
V.1 Etat de l’art.....	57
V.2 Les TASQ développés et étudiés au cours de ces travaux de thèse.....	61
Chapitre 2 – LES TASQ HYDROSOLUBLES COMME LIGANDS DE QUADRUPLPLEXES.....	65
VI. Contexte biologique	65
VI.1 Les quadruplexes présents sur l’ADN	65
VI.2 Les quadruplexes présents sur l’ARN	73
VI.3 Intérêts du ciblage des quadruplexes d’ADN et d’ARN et apologie de la recherche de ligands sélectifs	76
VII. Développement de nouveaux ligands biomimétiques : les TASQ.....	79
VII.1 Modes d’interaction des quadruplexes	79
VII.2 Aperçu des ligands dans la littérature	81
VII.3 Concept des TASQ	84
VIII. La première série de TASQ : les DOTASQ.....	85
VIII.1 Concept et synthèse.....	85
VIII.2 Analyses préliminaires de la structuration par RMN ¹ H.....	87
VIII.3 Evaluation par FRET-melting.....	89
VIII.4 Conclusion	94
IX. Modification du <i>template</i> : le PorphySQ.....	95
IX.1 Etudes préliminaires par modélisation moléculaire	96
IX.2 Synthèse du PorphySQ.....	96
IX.3 Etudes biophysiques	97
IX.4 Conclusion.....	100
X. Le ^{PNA}DOTASQ comme meilleur TASQ hydrosoluble.....	100
X.1 Synthèse du ^{PNA} DOTASQ.....	102
X.2 Etudes par RMN ¹ H	102

X.3	Analyses par spectrométrie de masse ESI-MS	105
X.4	Expériences de FID (<i>fluorescence intercalator displacement</i>)	107
X.5	Etude du ^{PNA} DOTASQ par FRET-melting.....	109
X.6	Bilan pour le ^{PNA} DOTASQ	117
XI.	Conclusions et perspectives	118
Chapitre 3 – LES TASQ COMME PLATEFORME POUR LA CATALYSE PSEUDO-ENZYMATIQUE		125
XII.	Les oligonucléotides catalytiques : ribozymes et DNazymes	126
XIII.	De la <i>horseradish peroxidase</i> aux quadruplexes d'ADN	128
XIII.1	Découverte des propriétés DNazyme des quadruplexes d'ADN.....	128
XIII.2	Applications des systèmes DNazyme quadruplexes, mimes de réactions de peroxydation	132
XIV.	Des quadruplexes d'ADN aux TASQ hydrosolubles.....	135
XIV.1	Première preuve de concept avec les DOTASQ.....	135
XIV.2	Amélioration catalytique avec le ^{PNA} DOTASQ.....	143
XIV.3	Optimisation de l'efficacité catalytique des réactions de TASQzyme.....	145
XV.	Les TASQ hydrosolubles comme « boosters » des réactions de DNazyme.....	153
XV.1	Elucidation de l'activité DNazyme des quadruplexes multimériques.....	155
XV.2	Création de sites artificiels de haute activité.....	158
XVI.	Conclusion.....	162
CONCLUSION GENERALE.....		167
EXPERIMENTAL PART		171
XVII.	Materials and chemicals.....	171
XVIII.	Chemical synthesis.....	172
XVIII.1	Synthesis of DOTASQ-C ₁	172
XVIII.2	Synthesis of DOTASQ-C ₅	175
XVIII.3	Synthesis of Tb.DOTASQ-C ₁ and Tb.DOTASQ-C ₅	179
XVIII.4	Synthesis of ^{PNA} DOTASQ.....	180
XIX.	Biophysical experiments	183
XIX.1	Preparation of the DNA and RNA structures	183
XIX.2	FRET-melting experiments	185
XIX.3	G4-FID experiments.....	185
XIX.4	ESI-MS experiments	186
XIX.5	¹ H NMR studies	187
XX.	DNazyme and TASQzyme experiments.....	188
XX.1	Experiments carried out in 1mL-quartz cuvettes	188
XX.2	Experiments carried out in a 96-well plate	189
BIBLIOGRAPHIE		192
VALORISATION		206

ABREVIATIONS

A : adénine

Å : angström

ABTS : 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)

ADN ou **DNA** : acide désoxyribonucléique

ADP-N-P : adénosine 5'-(β,γ -imido)-triphosphate

ALT : *alternative lengthening of telomeres*

ARN ou **RNA** : acide ribonucléique

ARNm : ARN messenger

ARNr : ARN ribosomal

ARNt : ARN de transfert

ATP : adénosine 5'-triphosphate

Boc : *tert*-butoxycarbonyle

C : cytosine

c-di-GMP : diguanosine monophosphate cyclique

COSY : *correlation spectroscopy*

CTP : cytosine 5'-triphosphate

DAB : 3,3'-diaminobenzidine

DABCO : 1,4-diazacyclo[2.2.2]octane

dAMP : désoxyadénosine monophosphate

DDQ : 2,3-dichloro-5,6-dicyano-1,4-benzoquinone

dGMP : désoxyguanosine monophosphate

DMF : diméthylformamide

DMSO : diméthylsulfoxyde

DOTA : acide 1,4,7,11-tétraazacyclododécane-*N,N',N'',N'''*-tétraacétique

dTMP : désoxythymidine monophosphate

EDCI : *N*-(3-diméthylaminopropyl)-*N'*-éthylcarbodiimide hydrochloride

EF-TEM : *energy filtered transmission electron microscopy*

ELISA : *enzyme-linked immunosorbent assay*

ESI : *electrospray ionization*

FAM : 6-carboxyfluorescein

FeTPPS : 5,10,15,20-tetrakis(4-sulfonatophenyl)porphyrinato iron(III)

FID : *fluorescent intercalator displacement*
FRET : *fluorescence resonance energy transfer*
FWJ : *four-way junction*
G : *guanine*
GTP : *guanosine 5'-triphosphate*
Hemin : *chloroproporphyrin IX iron(III)*
HEPES : *4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid*
HMQC : *heteronuclear single-quantum correlation*
HOBt : *1-hydroxybenzotriazole*
HPLC : *high performance liquid chromatography*
HRMS : *high resolution mass spectrometry*
HRP : *horseradish peroxidase*
IEX-HPLC : *ion exchange HPLC*
IRES : *internal ribosome entry site*
m7G : *7-méthylguanine*
MALDI : *matrix-assisted laser desorption ionization*
MeOH : *méthanol*
MS : *mass spectroscopy*
NHEJ : *non-homologous end-joining*
NOESY : *nuclear overhauser effect spectroscopy*
NTP : *nucléoside 5'-triphosphate*
ORF : *open reading frame*
PAGE : *polyacrylamide gel electrophoresis*
PNA : *peptide nucleic acid*
RAFT : *regioselectively addressable functionalized template*
RCM : *ring closing metathesis*
RE : *réticulum endoplasmique*
RISC : *RNA-induced silencing complex*
RMN ou **NMR** : *résonance magnétique nucléaire*
RP-HPLC : *reversed-phase HPLC*
SELEX : *systematic evolution of ligands by exponential enrichment*
SPR : *surface plasmon resonance*
SSB : *single strand binding*
STM : *scanning tunneling microscope*
T : *thymine*

TACN : 1,4,7-triazacyclononane
TAMRA : *tetramethyl-6-carboxyrhodamine*
TASP : *template-assembled synthetic proteins*
TASQ : *template-assembled synthetic G-quartet*
TDS : *thermal difference spectra*
TEOS : *tetraethylorthosilicate*
TERRA : *telomeric repeat-containing RNA*
TFO : *triplex-forming oligonucleotide*
TMB : 3,3',5,5'-tetramethylbenzidine
TMPyP4 : *tetrakis(1-methyl-4-pyridyl)porphyrin*
TO : thiazole orange
TOF : *time of flight*
TTP : thymidine 5'-triphosphate
TTS : *triplex target site*
TWJ : *three-way junction*
U : uracile
UTP : uridine 5'-triphosphate
UTR : *untranslated region*
VIH : virus de l'immunodéficience humaine
Z ou **Cbz** : carboxybenzyle

INTRODUCTION GENERALE

Les assemblages supramoléculaires sont des édifices composés de plusieurs molécules dont la stabilité globale repose sur des interactions intermoléculaires non covalentes de nature électrostatique ou de Van der Waals, auxquelles s'ajoutent également des liaisons hydrogènes, qui jouent un rôle prépondérant dans la chimie du vivant. En effet, la biochimie cellulaire se base à la fois sur la chimie moléculaire pour synthétiser une large gamme de biomolécules complexes indispensables au bon fonctionnement de la cellule, et également sur la chimie supramoléculaire qui lui offre la possibilité de créer temporairement des assemblages non figés capables de s'assembler et de se dissocier en fonction des besoins. Ce domaine des interactions non covalentes est à l'origine de la reconnaissance d'une enzyme par son substrat, de l'association antigène et anticorps, de la fixation d'une protéine avec une séquence d'ADN (ex : histones) ou d'ARN (ex : ribosome), ou encore de l'appariement des brins complémentaires d'ADN qui conduisent à la formation de la double hélice.^[1-3]

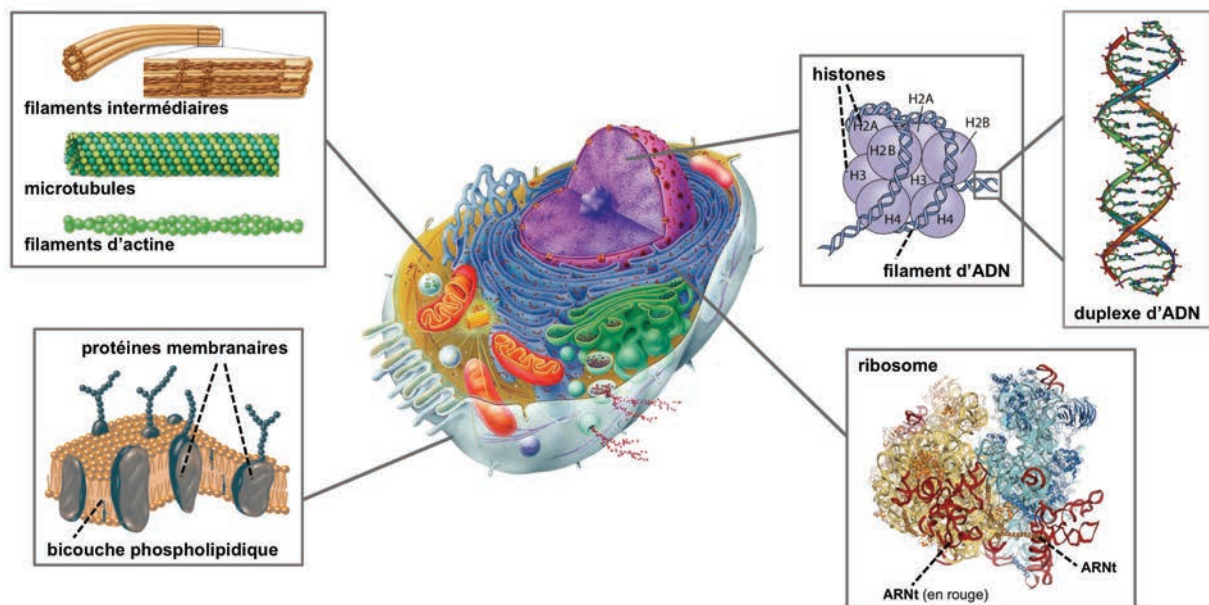


Figure 1 – Différents exemples de structures supramoléculaires présentes au sein d'une cellule eucaryote.

Le choix de ce type d'assemblage par la Nature n'est pas anodin puisqu'il permet de construire des édifices complexes évolutifs au cours du développement cellulaire à partir d'un nombre restreint de briques élémentaires (Figure 1). Dans le cas des acides nucléiques, seulement cinq unités nucléotidiques primitives suffisent : l'adénine, la cytosine et la guanine qui sont communes à l'ADN et à l'ARN et auxquelles s'ajoutent la thymine pour l'ADN et l'uracile pour l'ARN. Ces nucléobases ont une fonction de stockage, de transmission et de réplication de l'information

génétique, directement liée à leur propriété d'auto-assemblage rigoureux. Alors que les appariements entre bases complémentaires reposent sur des liaisons hydrogènes, leur faible énergie est contrebalancée par leur nombre au sein des chromosomes qui renferment de 45 à 245 millions de bases chez l'homme.^[4, 5]

Il est alors assez simple de comprendre l'intérêt du chimiste supramoléculaire à étudier ces phénomènes dans le but de comprendre davantage leurs propriétés, mais également de les détourner pour développer de nouvelles applications spécifiques. En effet, la construction d'architectures supramoléculaires nécessite une communication entre les briques moléculaires et le choix des nucléobases comme lien non covalent est particulièrement judicieux. Ainsi, plusieurs modèles ont été développés sur ce principe et trouvent des applications diverses dans le domaine des polymères, de la chimie combinatoire ou encore du transport d'ions à travers des liposomes.^[6-9]

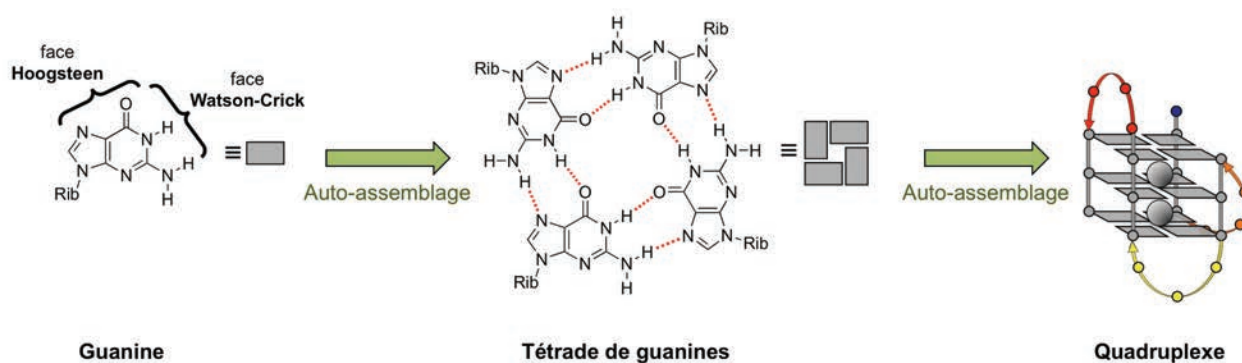


Figure 2 – Représentation d'une guanine pouvant s'auto-assembler en tétrades de guanines qui, dans certaines conditions, s'empilent pour former un quadruplexe.

Dans ce contexte, nous nous sommes intéressés à la construction de molécules renfermant un seul type de nucléobases et plus exactement quatre guanines, connues pour leur capacité à s'auto-assembler en une tétrade de guanines. Egalement appelée G-quartet, elle est l'élément constitutif majeur d'un quadruplexe d'ADN ou d'ARN (Figure 2). Appartenant à la catégorie des structures non-usuelles, c'est-à-dire divergentes du duplexe, le quadruplexe est au cœur d'un travail de recherche particulièrement intense depuis une vingtaine d'années. Son rôle au sein du génome en fait une cible de choix pour le développement de nouvelles thérapies anticancéreuses et explique l'engouement des scientifiques pour la synthèse et l'étude de ligands spécifiques de ces structures.^[10-13]

Loin de n'être qu'une cible biologique d'intérêt, le quadruplexe d'ADN est également devenu à la fin des années 1990 une plateforme de choix pour des réactions catalytiques. Jouant le rôle de catalyseur aussi bien pour des réactions de métallation que de peroxydation, son application actuelle

en nanotechnologie est grandissante et constitue un volet important et prometteur dans le domaine des DNAzymes qui regroupe les multiples utilisations de l'ADN.^[14-16]

En partant du rôle biologique de l'ADN et de l'ARN au cours du cycle cellulaire, nous introduirons au cours du premier chapitre les différentes structures usuelles et non-usuelles qu'ils adoptent en fonction de leurs séquences et du milieu. Ces observations seront le point de départ pour le développement de systèmes supramoléculaires biomimétiques. Parmi eux, l'emploi des guanines a été particulièrement étudié et s'explique par leur forte modularité due à leur nombre important de sites disponibles pour la formation de liaisons hydrogènes. Leur association sous forme d'une tétrade de guanines synthétique est largement reportée dans la littérature au départ de monomères renfermant une ou deux guanines.

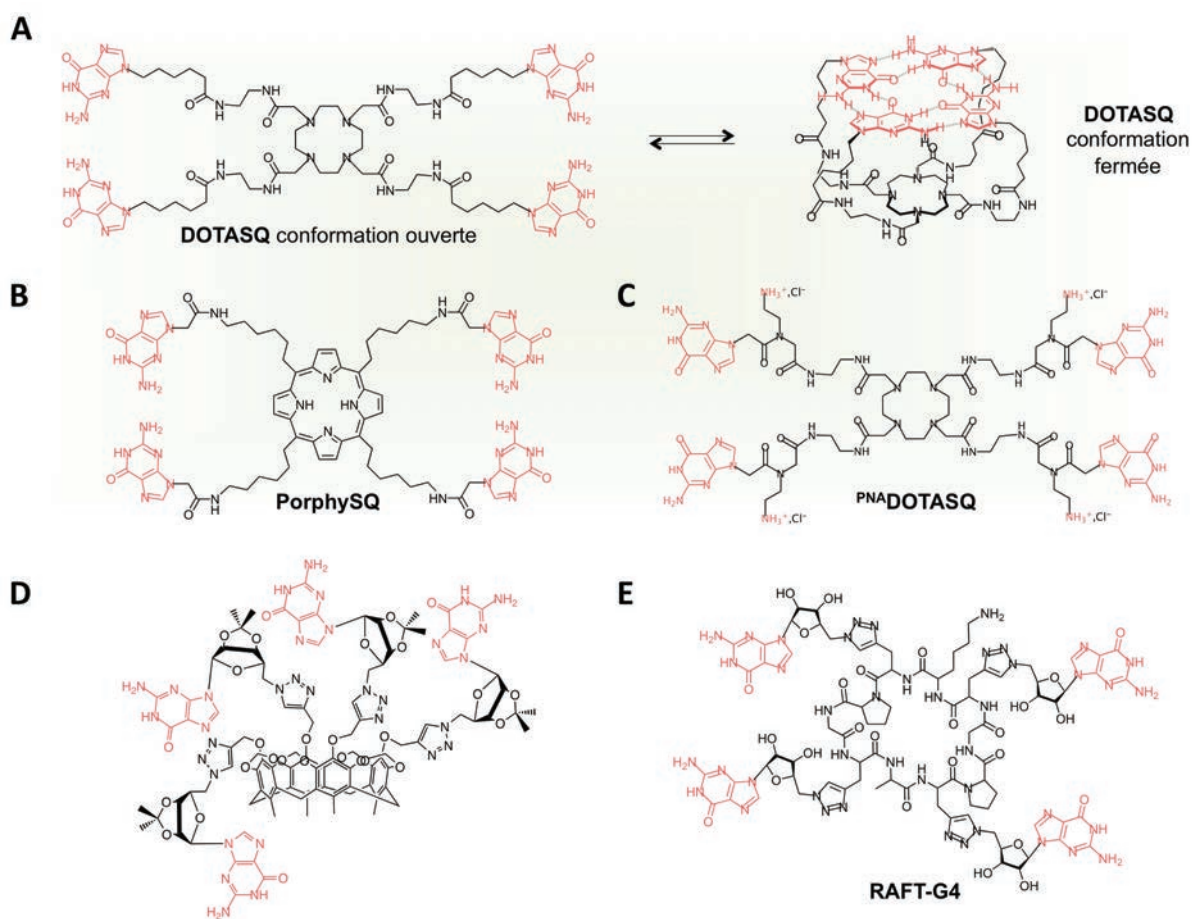


Figure 3 – Structures chimiques du DOTASQ (A), du PorphySQ (B) et du ^{PNA}DOTASQ (C) développés au laboratoire. Premier TASQ lipophile mis au point par l'équipe de J. Sherman (D) et RAFT-G4 issu du groupe de E. Defrancq (E).

Chaque tétrade nécessitant quatre guanines, il semblait intéressant d'avoir directement ces quatre nucléobases au sein d'une même molécule qui pourrait ainsi former intramoléculairement sa tétrade de guanines. Ce concept a été introduit en 2008 par les travaux de J. Sherman qui développa ainsi la première série de TASQ (pour *template-assembled synthetic G-quartet*) au départ d'un cœur

calixarène lipophile (Figure 3D) qui lui permis de valider l'approche, mais qui limitait fortement leur application.^[17] Pour pallier ce problème, nous avons conçu et étudié les tout premiers modèles hydrosolubles construits autour d'un cœur polyazamacrocyclique (DOTASQ et ^{PNA}DOTASQ, Figure 3A et C),^[18-20] ou porphyrinique (PorphySQ, Figure 3B)^[21] auxquels s'ajoute l'étude d'une troisième série reposant sur une armature cyclopeptidique développée par l'équipe de E. Defrancq (RAFT-G4, Figure 3E).^[22]

Ainsi, la formation de la tétrade de guanines sur des TASQ hydrosolubles a permis leur application selon deux axes de recherche (Figure 4). Le premier repose sur leur utilisation en tant que ligands biomimétiques de quadruplexes dont l'interaction est basée sur une reconnaissance spécifique entre deux tétrades de guanines, l'une naturelle (provenant du quadruplexe d'ADN ou d'ARN), l'autre synthétique (formée de manière intramoléculaire au sein d'un TASQ).^[18-21] L'intérêt biologique et thérapeutique de cibler des quadruplexes d'ADN et/ou d'ARN ainsi que les évaluations biophysiques entreprises sur ces composés seront discutés dans la seconde partie de cette thèse. Nous démontrerons ainsi l'avantage d'une telle approche en terme de stabilisation et surtout de sélectivité de ces composés pour leur cible biologique.

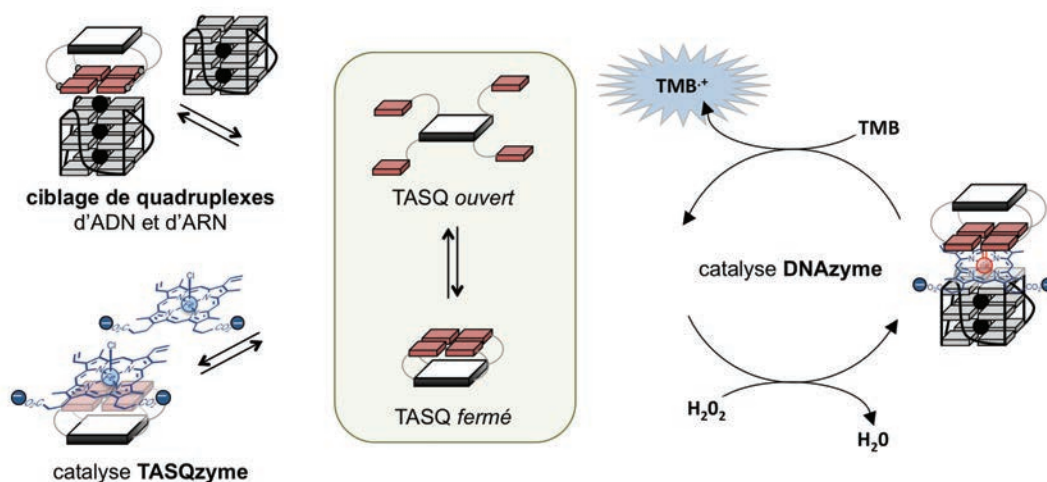


Figure 4 – Les TASQ hydrosolubles présentent de multiples applications développées au cours de ce travail de thèse et seront détaillées au long des chapitres 2 (ciblage de quadruplexes) et 3 (catalyse DNAzyme et TASQzyme).

Le second consiste en l'utilisation de ces molécules biomimétiques en tant que catalyseur dans des réactions de DNAzyme en lieu et place d'un quadruplexe. En effet, les TASQ, constitués d'une simple tétrade de guanines, peuvent être considérés comme un modèle réduit et simplifié d'un quadruplexe d'ADN et peuvent donc potentiellement les substituer dans un contexte *in vitro*.^[19, 23-25] Nous verrons que cette modification, qui entraîne une variation de l'activité catalytique globale,

offre une bien meilleure modularité ainsi qu'une plus grande simplicité d'utilisation que leurs homologues oligonucléotidiques.

Enfin, une conclusion générale de ces travaux résumera l'apport de ce travail de recherche et ouvrira sur les multiples perspectives qu'offrent les TASQ hydrosolubles, aussi bien dans le domaine des ligands de quadruplexes d'ADN et d'ARN que dans la mise au point d'outils biotechnologiques.

En espérant que ce manuscrit soit le plus didactique possible.

CHAPITRE 1 – DE L'ADN CELLULAIRE AUX ASSEMBLAGES SUPRAMOLECULAIRES DE NUCLEOBASES

I. Introduction

Fondée au cours de la première moitié du XIX^{ème} siècle suite aux travaux de F. Wöhler sur la synthèse de l'urée, la chimie moléculaire se définit comme l'étude et l'élaboration de structures moléculaires qui résultent de la rupture et/ou de la formation de liaisons covalentes entre atomes. N'ayant cessé de s'améliorer depuis, elle permet aujourd'hui d'accéder à des systèmes moléculaires complexes grâce à des stratégies de synthèses toujours plus sélectives, efficaces et contrôlées. Son impact sur notre quotidien est indéniable et s'applique aussi bien au développement de nouveaux matériaux pour le génie civil ou les nouvelles technologies, qu'à la mise au point de composés aux propriétés innovantes pour la cosmétique, la pétrochimie, la médecine, etc. La liste est bien trop longue pour être énumérée ici et ne cesse de s'allonger grâce au travail quotidien des chimistes aussi bien en recherche fondamentale qu'appliquée.

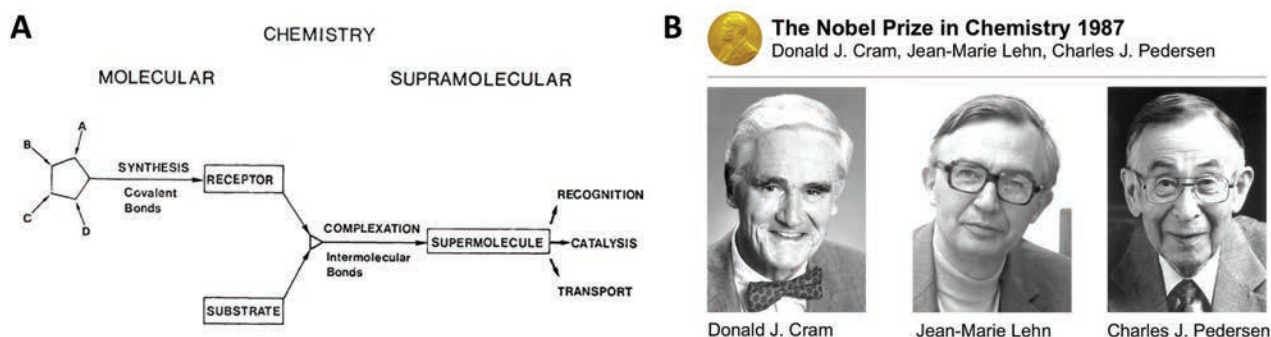


Figure 5 – Premier concept de la chimie supramoléculaire proposé par J.-M. Lehn en 1978 (A) qui partage en 1987 le prix Nobel de Chimie avec D. Cram et C. Pedersen (B).

Basée sur des interactions électrostatiques, de Van der Waals ou hydrophobes ainsi que sur des liaisons hydrogènes plus faibles que les liaisons covalentes, la chimie supramoléculaire est loin de n'être qu'un concept de laboratoire. En effet, la Nature les utilise depuis toujours pour concevoir des assemblages complexes entre biomolécules dont le processus d'association et de dissociation évolue en fonction du temps et du milieu (comme la séparation des deux brins d'ADN complémentaires qui se produit au moment de la réplication et à l'aide d'hélicases). La chimie supramoléculaire régit donc la reconnaissance sélective d'un substrat par son enzyme, d'un antigène par son anticorps, d'une protéine par l'ADN (comme les histones qui régulent la compaction de

l'ADN) ou l'ARN (complexe ribosomique pour la synthèse de protéines) ou encore l'association entre nucléobases aussi bien pour l'ADN que l'ARN.^[1-3]

Cependant, l'élaboration d'édifices complexes ne réside pas uniquement dans l'assemblage de briques moléculaires par réaction chimique ; en effet, cet assemblage peut naître d'interactions non covalentes intermoléculaires. Ce domaine des associations intermoléculaires est baptisé « chimie supramoléculaire » dont le terme et le concept ont été introduits en 1978 par J.-M. Lehn (Figure 5A).^[26] Précurseur de la chimie hôte-invité et de la reconnaissance moléculaire avec D. J. Cram et C. J. Pedersen, ils partageront tous les trois le Prix Nobel de Chimie de 1987 (Figure 5B).^[27-29] Depuis lors, un véritable engouement est né pour cette nouvelle discipline qui s'explique sans aucun doute par son extraordinaire modularité permettant de concevoir des édifices aux multiples propriétés et applicables à plusieurs domaines de recherche, comme la catalyse, les polymères, la médecine, etc.^[3, 30-33]

En s'axant sur les acides nucléiques qui sont le cœur des travaux de recherche de cette thèse, nous commencerons par un bref rappel des étapes clefs de la vie de la cellule faisant intervenir l'ADN et l'ARN. Nous étudierons ainsi leurs rôles et leurs fonctions au cours des processus de réplication et de transcription. Loin de n'être limités qu'à la célèbre double hélice où deux brins s'associent selon leurs bases complémentaires (*i.e.* cytosine-guanine et adénine-thymine (ADN) ou adénine-uracile (ARN)), nous verrons que l'ADN et l'ARN sont très polymorphiques et peuvent adopter en fonction de leur séquence et des conditions du milieu différentes structures dites non-usuelles (Figure 6).^[34]

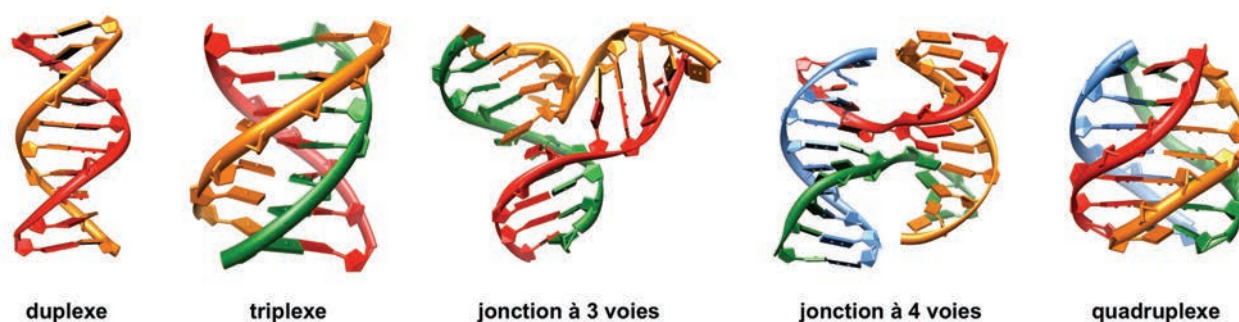


Figure 6 – Au-delà de la structure duplexe, l'ADN et l'ARN peuvent former sous certaines conditions des structures non-usuelles.

Ces différents arrangements témoignent de la modularité des nucléobases et sont une source d'inspiration pour les chimistes qui peuvent alors concevoir des édifices complexes dont les briques moléculaires qui les constituent s'assemblent entre elles à l'aide d'interactions entre nucléobases. Certaines de ces architectures supramoléculaires appliquées à différents domaines de recherche

seront présentées en quatrième partie de ce chapitre avant de nous concentrer sur les assemblages mettant en jeu la formation d'une tétrade de guanines, naturellement présente au sein des quadruplexes d'ADN et d'ARN. En partant de la découverte par Gellert en 1962 des premiers auto-assemblages de guanines,^[35] nous énumérerons les travaux entrepris depuis lors par plusieurs équipes pour nous focaliser finalement sur les molécules renfermant quatre guanines, et plus particulièrement les TASQ.

Nous présenterons alors les différentes séries hydrosolubles développées au laboratoire et dont les propriétés d'interaction avec l'ADN et de catalyse feront l'objet des chapitres suivants.

II. Prérequis : biochimie élémentaire

Avant de nous intéresser strictement aux structures de l'ADN et de l'ARN, nous proposons dans un premier temps de les recontextualiser dans leur milieu biologique : la cellule. En s'intéressant uniquement aux organismes eucaryotes, nous présenterons les rôles et fonctions des acides nucléiques permettant d'une part d'introduire le vocabulaire propre à la biochimie et d'autre part, de constituer un socle solide pour la compréhension des stratégies d'interactions relatives au ciblage de quadruplexes d'ADN qui seront présentées chapitre 2.

II.1 La cellule eucaryote et le cycle cellulaire

II.1.A Organisation de la cellule eucaryote

Tout être vivant est constitué d'unités biologiques élémentaires appelées cellules qui, selon l'organisme étudié, peuvent être de deux types différents qui se distinguent par la présence ou non d'un noyau.

Les cellules procaryotes, d'une taille variant de 1 à 5µm, n'en possèdent pas et confinent donc leur ADN, support de l'information génétique, sous une forme circulaire dans le cytoplasme. Elles sont observées chez les bactéries et les archées.

A l'inverse, les cellules eucaryotes ont une taille bien supérieure, de 10 à 100µm, et disposent d'un noyau (Figure 7). Séparées du milieu extérieur par la membrane plasmique, il est possible de distinguer deux grands domaines :

- d'une part **le cytoplasme** qui renferme le cytosol, phase liquide qui entoure tous les organites indispensables au bon fonctionnement de la cellule. Alors que certains sont présents en un seul exemplaire comme le réticulum endoplasmique rugueux et lisse ainsi que l'appareil de Golgi,

d'autres le sont en plusieurs copies. C'est le cas des mitochondries, véritables « centrales électriques » qui ont pour rôle de transformer en présence de dioxygène les nutriments en ATP (pour adénosine triphosphate), source d'énergie directement utilisable par la cellule. Autres constituants notables, les ARN autocatalytiques, ou ribozymes, présentent à l'instar des enzymes une activité catalytique et ont très largement inspiré les chimistes qui les ont adaptés à l'ADN pour créer le domaine du DNAzyme qui sera le sujet du troisième chapitre de cette thèse.

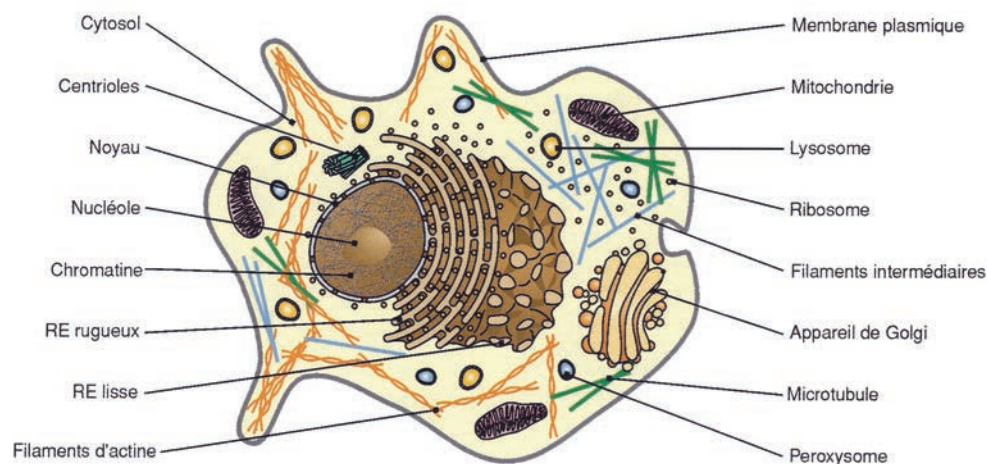


Figure 7 – Structure d'une cellule eucaryote mammifère type (RE pour réticulum endoplasmique).

- d'autre part **le noyau**, séparé du cytoplasme par l'enveloppe nucléaire composée de deux membranes externes et internes permettant des échanges avec le cytoplasme ainsi qu'avec le réticulum endoplasmique voisin. Il renferme l'ADN qui s'associe à des protéines, les histones, pour former de longs filaments désordonnés de chromatine qui ne se condensent sous la forme d'un chromosome qu'avant la division cellulaire (c'est-à-dire la mitose ou la méiose). Egalement présent, le nucléole, dont le nombre varie en fonction du cycle cellulaire, a pour fonction la transcription des ARN ribosomiques (ARNr), utiles pour la formation des ribosomes qui entrent en jeu dans la synthèse des protéines et qui seront abordés plus loin (voir II.2.B).^[4, 5]

II.1.B Le cycle cellulaire

Les cellules sont des systèmes dynamiques en perpétuelle évolution au cours du temps. Leur prolifération, c'est-à-dire leur croissance et leur division, est contrôlée par le cycle cellulaire qui à partir d'une cellule mère produit à chaque cycle deux cellules filles ayant rigoureusement la même information génétique (Figure 8).

Ce processus capable de recopier et de transmettre à l'identique son ADN s'effectue en quatre phases principales se déroulant dans cet ordre :

- **la phase G_1** : correspond à la phase de croissance de la cellule pendant laquelle elle va synthétiser les protéines indispensables à sa fonction au sein de l'organisme. Des facteurs de croissance contrôlent cette phase au cours de laquelle l'ADN est présent sous forme de simples chromatides.

- **la phase S** : appelée également phase de réplication, elle permet la copie de chaque chromatide en deux chromatides sœurs identiques.

- **la phase G_2** : contrôle que la réplication de l'ADN s'est bien déroulée. Lorsque ce n'est pas le cas, un processus de réparation post-répllicative corrige les erreurs avant d'entrer dans la phase suivante.

- **phase M** : la mitose se déroule en cinq étapes majeures qui sont la prophase, la prométaphase, la métaphase, l'anaphase et la télophase qui ne seront pas détaillées ici. Leurs rôles sont la séparation et la distribution aux deux extrémités de la cellule mère de chacune des deux chromatides qui constituent à ce stade chacun des chromosomes. Suit alors la cytokinèse qui divise la cellule mère en deux cellules filles rigoureusement identiques.

A cet instant, elles vont chacune à leur tour entrer dans un nouveau cycle cellulaire depuis la phase G_1 .

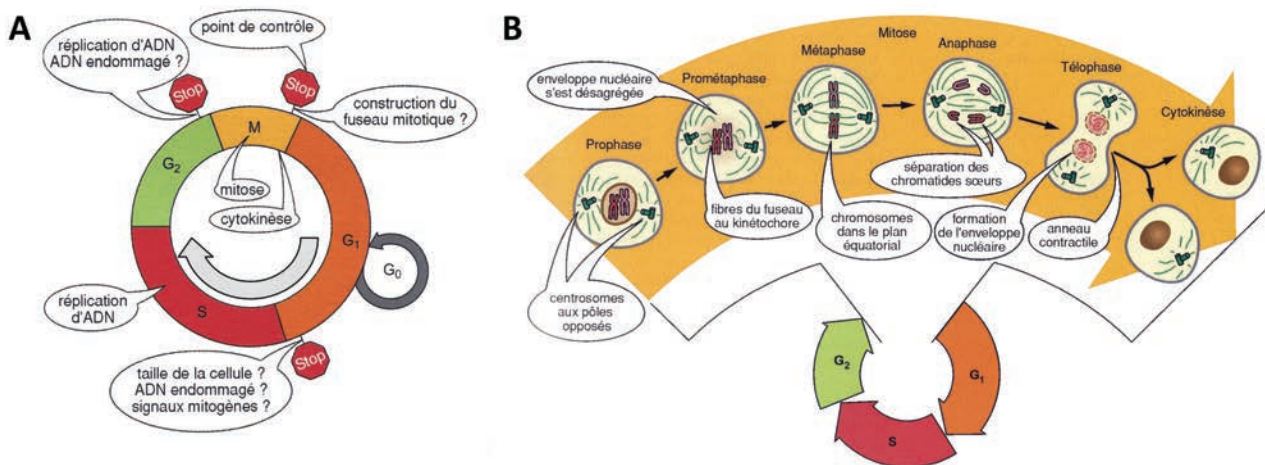


Figure 8 – Cycle cellulaire avec représentation des points de contrôle (A) et détail de la mitose (B).

Notons que dans certains cas, les cellules en phase G_1 ne continuent pas le cycle mais entrent en phase d'attente G_0 , ou quiescence, au cours de laquelle la prolifération cellulaire s'arrête. Elle peut être entraînée par des facteurs environnementaux, ou être une propriété intrinsèque de la cellule, comme pour les neurones par exemple.^[4, 5]

Le cycle cellulaire permet donc la prolifération d'une cellule au sein de l'organisme et doit de surcroît s'assurer qu'aucune erreur n'est commise. Pour cela, la cellule a mis en place différents mécanismes qui sont des points de contrôles entre chacune des étapes du cycle cellulaire. Ils reposent entre autres sur des protéines régulatrices de la famille des cyclines qui s'associent ou non à des protéines kinases cyclines dépendantes, ou Cdk. A titre d'exemple, le point de contrôle G₁/S qui vérifie si la cellule a atteint une taille suffisante et si l'ADN n'est pas endommagé dépend des facteurs cyclines A et E et des Cdk1 et 2, alors que le passage G₂/M qui contrôle si la duplication est bien complète et si l'ADN ne porte pas d'erreur à réparer fait intervenir la cycline B et A et Cdk1 (Figure 8A). Beaucoup d'autres points de contrôle existent et sont là pour garantir la conservation de l'intégrité génétique.^[36-39]

II.2 Rôles et fonctions de l'ADN et de l'ARN

L'ADN et l'ARN, dont les structures seront discutées en détail partie III, jouent un rôle primordial dans la conservation et l'expression de l'information génétique d'un organisme. En effet, l'ADN doit à la fois être copié et transmis aux cellules filles lors de la réplication (ou phase S) avec une haute fidélité, mais doit également être exprimé grâce au processus de transcription et traduction qui permet la biosynthèse de protéines vitales pour la cellule, en passant par des intermédiaires clefs que sont les ARN.

II.2.A Réplication de l'ADN

Présent sous la forme de la célèbre double hélice, l'ADN est une biomolécule de deux brins complémentaires construits à partir de seulement quatre briques élémentaires que sont les nucléotides. Ceux-ci sont toujours constitués d'un résidu phosphate et d'un monosaccharide et se distinguent par l'une des nucléobases présente: l'adénine (A), la thymine (T), la cytosine (C) et la guanine (G). Par un processus de reconnaissance spécifique faisant intervenir des liaisons hydrogènes, les deux brins vont toujours s'associer selon les appariements adénine/thymine et guanine/cytosine. Nous verrons par la suite que ces brins ont une orientation particulière notée 5'-3' et que l'appariement se fait spécifiquement en antiparallèle, c'est-à-dire avec un brin dans le sens 5'-3' et l'autre dans le sens opposé 3'-5' (Figure 9A et B).

La conservation du patrimoine génétique d'un organisme implique une transmission rigoureuse de l'ADN d'une cellule mère aux cellules filles des générations suivantes. Ce processus de réplication est associé à plusieurs protéines clefs, dont les ADN polymérases.

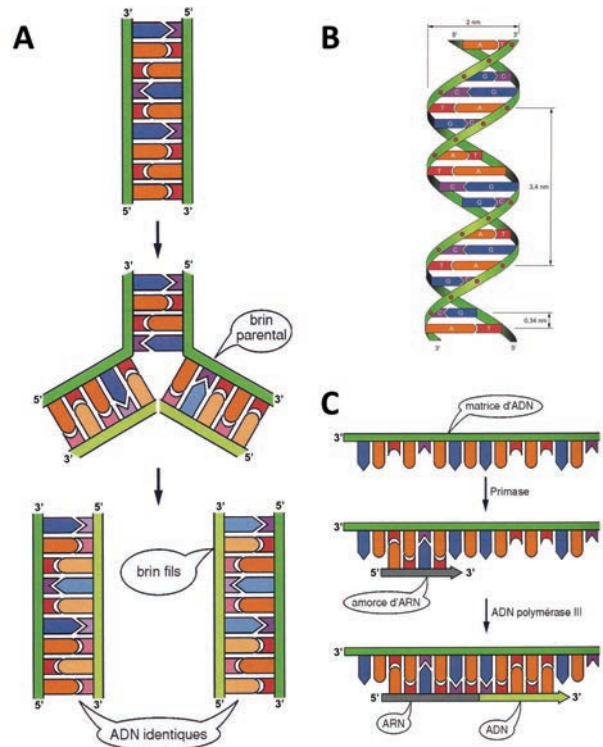


Figure 9 – Schéma de la réplication semi-conservative (A) dont la structure duplexe (B) a été simplifiée pour plus de clarté. Etape de synthèse des amorces d'ARN et action des polymérases au cours de la réplication (C).

La réplication nécessite dans un premier lieu l'ouverture des brins complémentaires qui est initiée au niveau des origines de réplifications sur une faible portion de 10 à 20 bases grâce à l'action des protéines de liaison à l'ADN double brin. L'hélicase peut alors venir se fixer et poursuivre le travail de séparation sur le reste du chromosome. Ce processus fait apparaître des bases non appariées qui sont alors stabilisées en s'associant aux protéines SSB (pour *single strand binding protein*) pour éviter leur dégradation et empêcher le réappariement.

Cette ouverture donne à l'ADN une forme de Y caractéristique appelée fourche de réplication. Les primases, qui jouent le rôle de précurseurs, vont ensuite synthétiser le long de chaque brin de courtes amorces d'ARN sur lesquelles les ADN polymérases vont pouvoir commencer leur travail (Figure 9C). Leur rôle est alors d'allonger l'amorce d'ARN en synthétisant un nouveau brin fille d'ADN complémentaire au brin matrice mère.^[4, 5, 40]

Cependant, ce processus n'est pas si linéaire : en effet, l'ADN polymérase ne peut lire le brin matrice que dans le sens 3'-5'. Or, les deux brins matrices sont antiparallèles et chacun va conduire à un mécanisme de transcription différent :

- la synthèse du **brin précoce** (ou *leading strand*) se fait le long du brin mère de sens 3'-5'. Elle a donc lieu dans le sens favorable pour la polymérase c'est-à-dire en direction de la fourche de réplication et ne nécessite donc qu'une amorce d'ARN : il s'agit du cas le plus simple (Figure 10A).

- la synthèse du **brin tardif** (ou *lagging strand*) est plus complexe puisqu'elle s'effectue dans le sens opposé à la progression de la fourche. Pour pallier le problème d'orientation défavorable du brin 5'-3', la cellule utilise comme astuce la formation d'une boucle qui crée un « retour » du brin en amont de la polymérase. Ainsi, il se trouve momentanément orienté dans le bon sens et sa transcription peut être effectuée (Figure 10A).

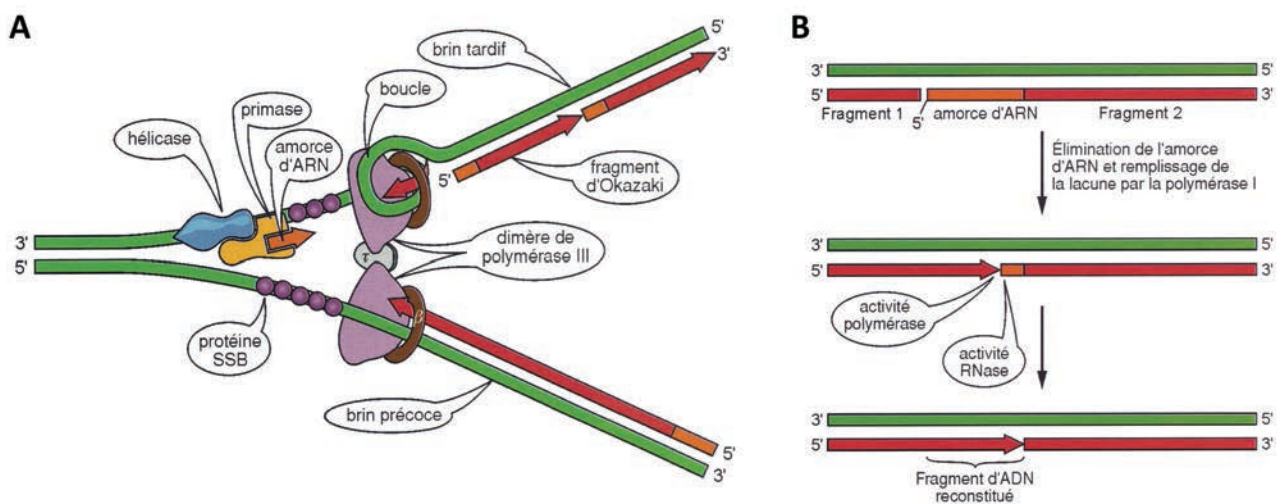


Figure 10 – Synthèse des brins filles au niveau de la fourche de réplication (A) et achèvement du brin tardif (B).

En procédant ainsi, il apparaît que la synthèse ne peut pas être continue puisqu'à intervalles réguliers, la polymérase bute sur du double brin d'ADN nouvellement synthétisé. A cet instant, elle se dissocie et glisse le long du brin jusqu'à un nouveau site d'amorce d'ARN en amont où elle recommence son travail, créant une succession de petits fragments d'environ 150 bases nommés fragments d'Okazaki. [41, 42]

L'assemblage de tous ces fragments en un brin continu d'ADN s'effectue dans un premier temps par la suppression puis le remplacement des amorces d'ARN par leurs équivalents ADN, grâce à l'action d'autres polymérases. Enfin, les ligases attachent de manière covalente tous les segments d'ADN entre eux (Figure 10B). [4, 5, 40]

Par tout ce processus de réplication, les deux brins matrices d'ADN mères ont été copiés en deux brins filles rigoureusement identiques dont le taux d'erreur incroyablement faible, de l'ordre

de 10^{-10} , est dû à un système très précis de relecture et de correction au cours de la réplication mais également ultérieurement (réparation des mésappariements, système Mut, recombinaison, etc.).^[5, 43]

Remarquons toutefois qu'une toute petite partie du génome n'est pas répliquée : les extrémités télomériques en 3' du brin mère (Figure 11A). En effet, au cours de la synthèse du brin tardif, une fois que l'amorce ARN située à l'extrémité 3' du brin matrice est supprimée, il n'est pas possible de la remplacer par son homologue ADN puisque que la polymérase qui en a la tâche n'a plus d'amorce pour initier son travail dans le sens 3'-5' (puisque'il s'agit ici de la dernière amorce). Ainsi, ce phénomène conduit à un raccourcissement progressif des chromosomes au niveau des télomères au fur et à mesure des divisions cellulaires et porte le nom de « problème de fin de réplication ».^[44] Sachant que la fourche de réplication travaille dans les deux sens depuis l'une des 30000 origines de réplifications afin d'accélérer cette phase S où plus de 3,2 milliard de bases doivent être synthétisées chez l'homme, les deux brins filles sont donc touchés (Figure 11B).

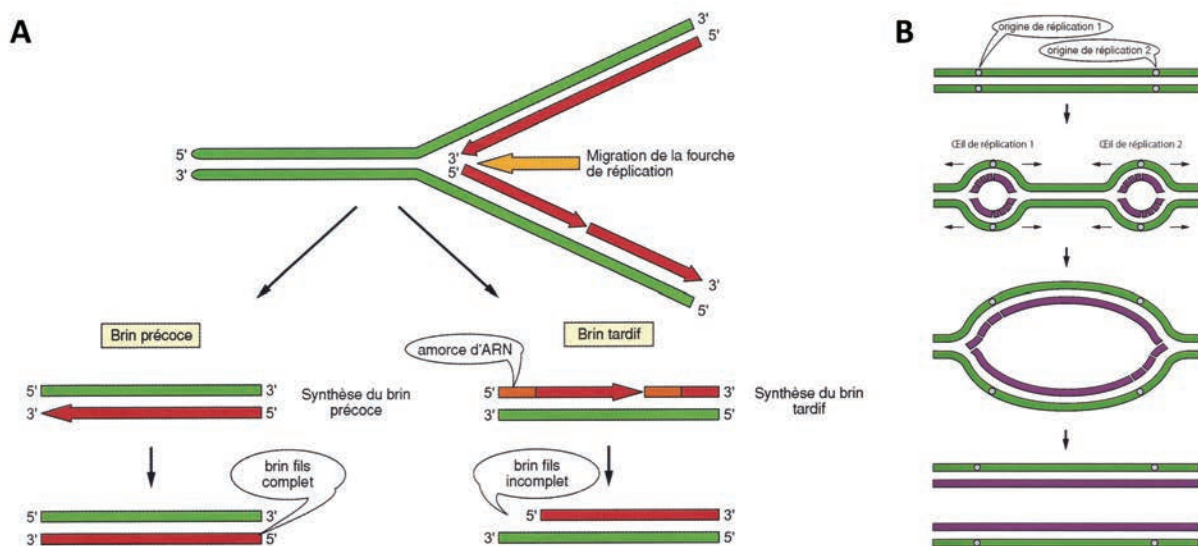


Figure 11 – Problème de fin de réplication au niveau du brin tardif (A) et représentation des multiples origines de réplifications (B).

Ce qui semble n'être en apparence qu'un détail a une incidence majeure sur la vie d'une cellule puisqu'une fois que les télomères atteignent une taille critique appelée limite de Hayflick,^[45] la cellule entre en sénescence puis meurt.

Notons également que par ce processus, l'extrémité 3' des télomères est sous forme simple brin sur une séquence de 150 à 200 bases. Composée d'une répétition hexanucléotidique riche en guanines capable se structurer sous la forme d'un quadruplexe d'ADN, elle s'associe au complexe protéinique *shelterin* et se replie ainsi sur elle-même pour former une boucle appelée t-loop.^[12, 46] Nous reviendrons sur ce point plus loin dans le manuscrit (voir VI.1.A.b).

II.2.B Transcription de l'ADN et traduction de l'ARNm

Le second rôle que nous allons étudier maintenant est l'expression de l'ADN en ARN puis de l'ARNm en protéines : ces étapes sont appelées respectivement la transcription et la traduction (Figure 12A et B).

La transcription se déroule dans le noyau des cellules eucaryotes à l'aide de l'ARN polymérase qui, pour un segment d'ADN duplexe défini, choisira de ne travailler que sur l'un des deux brins en général. La séquence d'ARN alors synthétisée est complémentaire du brin matrice, ou en d'autres termes identique à celle de l'autre brin d'ADN, aux différences près que la thymine (T) se voit remplacée par l'uracile (U).

Dans un premier temps, l'ARN polymérase se lie à l'ADN à l'aide d'autres protéines que sont les facteurs de transcription, au niveau de sites spécifiques nommés promoteurs. Il est intéressant de noter que certaines séquences consensus sont présentes dans plusieurs de ces sites, comme les boîtes TATA, GC, CAT.^[47]

L'ARN polymérase déroule alors l'ADN et synthétise le brin complémentaire d'ARN à partir de NTP (pour nucléotides triphosphates), toujours dans la direction 3'-5' (tout comme le fait l'ADN polymérase au cours de la réplication). La synthèse s'achève au niveau d'une séquence ou d'un facteur de terminaison où la polymérase se détache de l'ADN libérant ainsi l'ARN précurseur qui à ce stade n'est pas encore fonctionnel et nécessitera une phase de maturation.

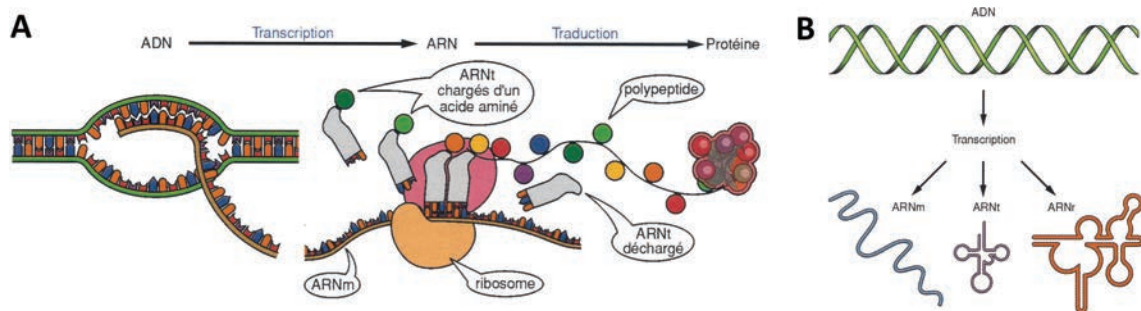


Figure 12 – La transcription et la traduction sont les étapes clés de l'expression génétique (A) ; au cours de la transcription trois types d'ARN sont synthétisés remplissant chacun des fonctions distinctes (B).

La nature et la fonction de l'ARN synthétisé sont dépendantes du couple séquence d'ADN d'origine et ARN polymérase spécifique dont il est issu. Ainsi :

- l'ARN polymérase I synthétise les grands ARN ribosomiques (ARNr) ;
- l'ARN polymérase II synthétise les ARN pré-messagers (préARNm) ;
- l'ARN polymérase III synthétise les petits ARN de transferts (ARNt), ribosomiques (ARNr), ou les petits ARN nucléaires (ARNsn).

Chacun de ces types d'ARN (Figure 12B) occupe une place essentielle au cours de la vie cellulaire et en particulier pendant la traduction.^[4, 5]

Les ARN messagers (ARNm)

Les ARN pré-messagers issus de la transcription directe de l'ADN par l'ARN polymérase II (Figure 13B) sont composés de deux extrémités 3' et 5' non codantes notées 3'-UTR et 5'-UTR (UTR pour *untranslated region*) qui encadrent le cadre ouvert de lecture ou ORF (*open reading frame*) qui à l'inverse des deux autres sera traduit en protéine. Nous verrons plus loin que ces régions non codantes peuvent être reliées à la formation d'un quadruplexe d'ARN (voir VI.2.B)

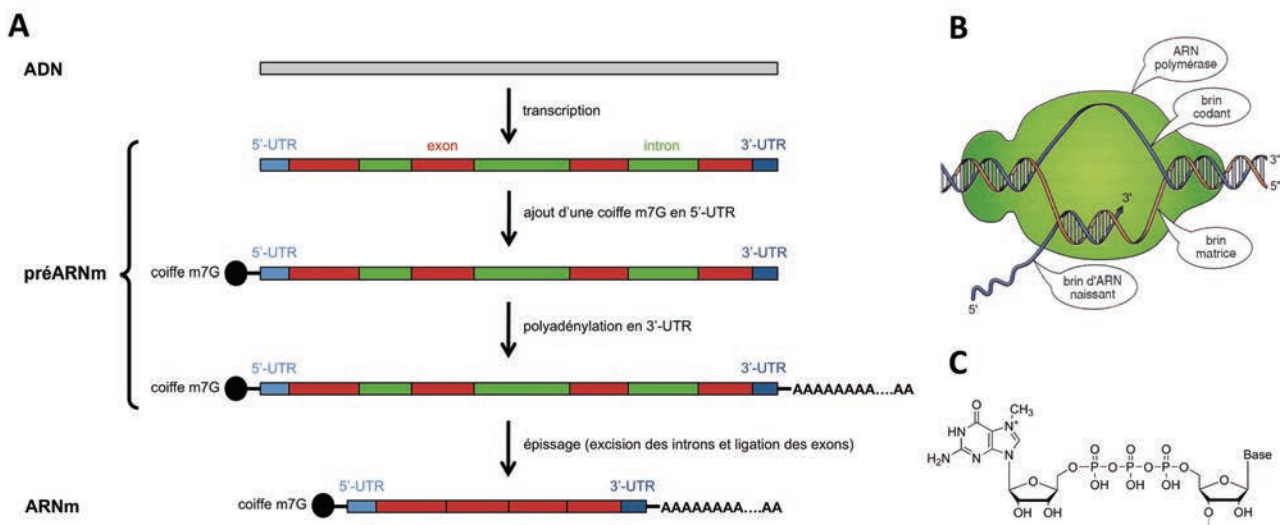


Figure 13 – Processus de maturation des ARNm (A) passant par l'ajout d'une coiffe 7-méthylguanosine (C). Synthèse de l'ARNm depuis l'ADN par l'action de l'ARN polymérase II (en vert, B).

Leur maturation (Figure 13A) commence par l'ajout d'une coiffe de type GTP, en général une 7-méthyl-guanine (Figure 13C), fixée à l'extrémité 5' par une liaison 5'-5' triphosphate. Elle servira par la suite de signal de reconnaissance pour le ribosome mais a aussi pour fonction de protéger l'ARN de la dégradation par les exonucléases une fois dans le cytoplasme. L'autre extrémité subit un clivage par une endonucléase suivi d'une polyadénylation (jusqu'à 250 unités). Arrive alors l'épissage de l'ORF pendant lequel les introns, séquences très majoritairement non codantes, vont être éliminés pour ne garder que les exons. A ce stade, l'ARN messager est mature et peut donc quitter le noyau cellulaire en direction du cytoplasme où il sera traduit en protéine par les ribosomes en suivant les règles de traductions imposées par le code génétique.^[4, 5]

Les ARN ribosomiques (ARNr)

Ils forment, en collaboration avec les protéines ribosomiques, des complexes moléculaires appelés ribosomes sur lesquels s'effectue la traduction de l'ARNm. Les longs ARNr précurseurs

synthétisés par l'ARN polymérase I subissent une maturation, comme l'unité 45S de 13000 nucléotides, qui consiste en leur sédimentation en plusieurs fragments d'ARNr de tailles plus petites (18S, 5,8S et 28S, Figure 14A, B). Les ARNr issus de la transcription par l'ARN polymérase III comme 5S sont eux plus petits et directement fonctionnels.

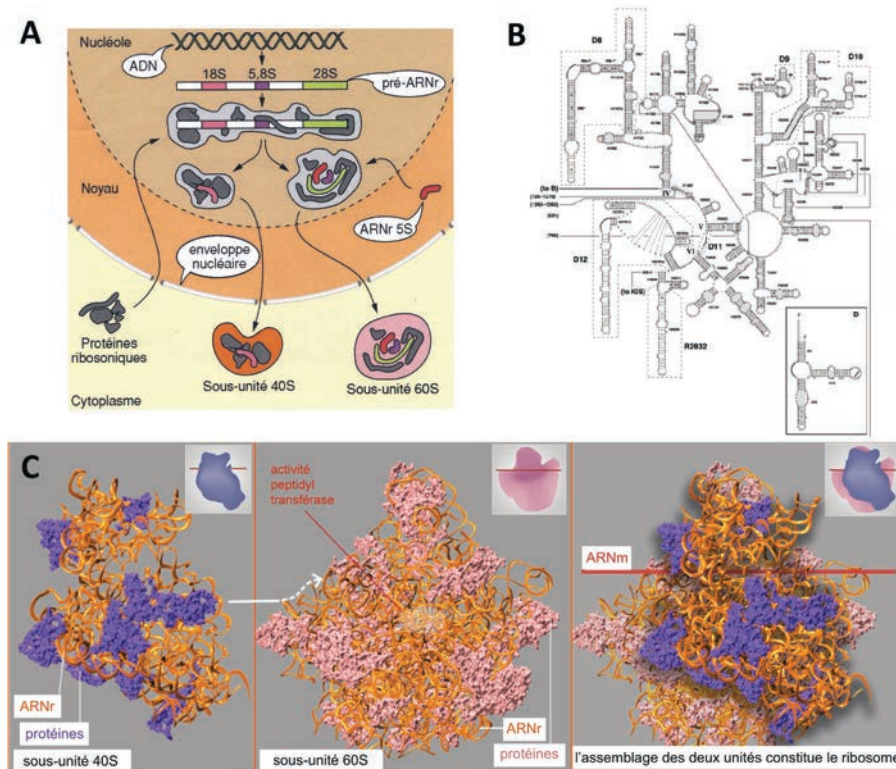


Figure 14 – Assemblage des ribosomes dans le nucléole avant de répartir pour le cytoplasme (A). Structure secondaire de l'ARNr (18S + 5,8S + 28S + 5S) chez l'abeille (B) et assemblage des sous-unités 40S et 60S (C).

Ainsi, deux sous-unités ribosomiques 40S et 60S se forment dans le noyau avant de partir pour le cytoplasme où elles s'assembleront pour former le ribosome opérationnel (Figure 14C).^[4, 5]

Les ARN de transfert (ARNt)

D'environ 70 à 100 nucléotides, leur maturation diffère des précédentes et passe par l'ajout d'une séquence $5'CCA3'$ à leur extrémité 3' qui sert de site de fixation pour les acides aminés (Figure 15A). Après les avoir transportés au niveau du ribosome, l'ARNt va se charger de lire à l'aide de son anticodon l'ARNm par séquences de trois nucléotides, c'est-à-dire par codon, auxquelles il fera correspondre un et un seul acide aminé (Figure 15C).

Cette correspondance est définie par le code génétique (Figure 15B) et la succession de codons d'un ARNm entraîne donc la formation d'une chaîne polypeptidique spécifique ; cette synthèse se termine dès l'apparition d'un des trois codons stop. A cet instant, l'ARNm ainsi que la protéine nouvellement synthétisée se détachent du ribosome.^[4, 5]

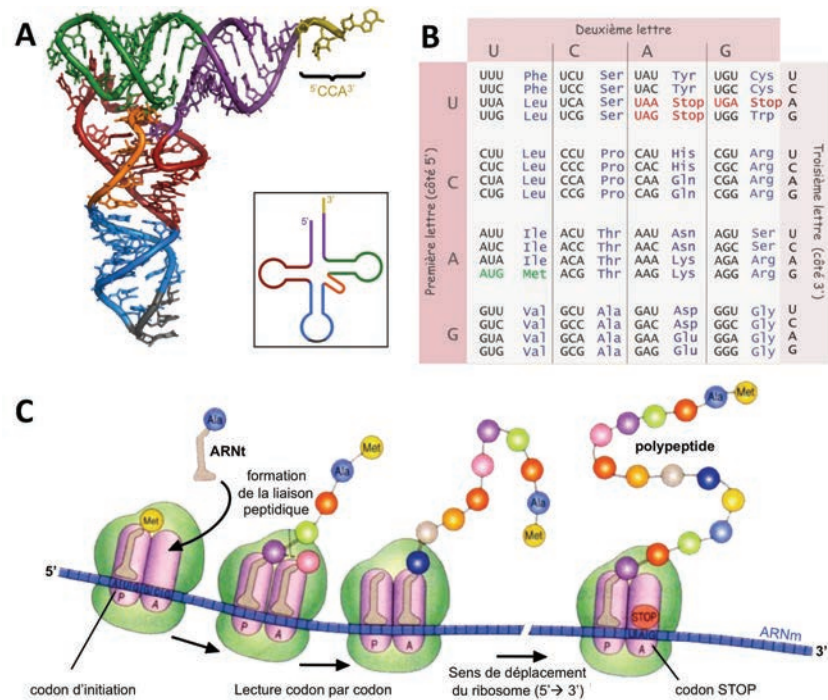


Figure 15 – Structure tridimensionnelle et schématique d'un ARNt (**A**), tableau du code génétique (**B**), et mécanisme de synthèse des protéines à l'aide des ribosomes (en vert, **C**).

Les autres ARN

De par leurs différentes fonctions, les ARN ne sont pas tous présents en même proportion dans la cellule. En effet, les ARNr qui ont un rôle purement fonctionnel en représentent 80%, contre 15% pour les ARNt. A l'inverse, les ARNm qui une fois produits peuvent être utilisés plusieurs fois par les ribosomes sont minoritaires à 5%.

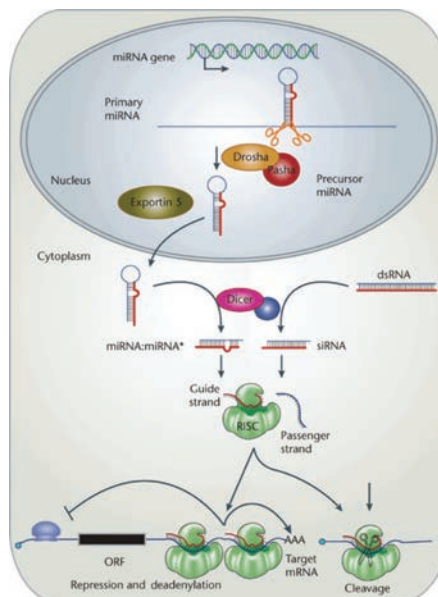


Figure 16 - Mécanisme d'action des microARN

Il existe néanmoins une quatrième classe d'ARN : les petits ARN nucléaires. Certains d'entre eux interviennent lors des étapes de maturation des autres ARN comme les petits ARN cytoplasmiques (ARNsc) ou les petits ARN nucléolaires (ARNsno) alors que d'autres sont des régulateurs de l'expression génique comme les ARN interférents (ARNi). Ces derniers incluent les petits ARN interférents (ARNsi) et les microARN (miARN) qui après des étapes de maturation s'associent à un complexe protéinique nommé RISC (RNA-induced silencing complex) avant d'interagir avec d'autres ARN (Figure 16). S'il s'agit d'un ARNm, un appariement parfait avec le complexe RISC conduira à sa destruction, alors qu'un appariement incomplet inhibera sa traduction.^[48, 49]

Il semble assez clair aujourd'hui que leur rôle de régulateur est mis à mal dans plusieurs maladies, dont certains cancers.^[50-52]

III. Les structure de l'ADN et de l'ARN

Jusqu'alors, nous avons présenté les brins d'ADN et d'ARN comme des rails linéaires qui pouvaient s'associer entre eux selon leurs bases complémentaires ; cependant, les structures qu'ils adoptent sont loin d'être si schématiques mais présentent des structures secondaires complexes et tridimensionnelles. La découverte de la double hélice d'ADN, forme la plus répandue dans le génome et véritable icône de la biologie du vivant pour le grand public, a eu un impact retentissant dans le monde de la recherche et marqua un tournant dans le domaine de la biologie moléculaire. Nous verrons que cette structure usuelle est l'une des nombreuses possibilités offertes par la grande modularité des nucléobases d'ADN et d'ARN qui permet de donner naissance à une multitude de structures non-usuelles.

III.1 L'histoire de la double hélice d'ADN

Cette année 2013 marque le 60^{ème} anniversaire de la publication dans *Nature* de la structure de la double hélice d'ADN par James Watson et Francis Crick.^[53] Leur article est directement suivi par ceux de Maurice Wilkins^[54] puis de Rosalind Franklin^[55] qui contiennent tous deux des résultats d'analyse par rayons X de cristaux d'ADN.

Avant eux, d'autres scientifiques avaient tenté d'élucider la structure de l'ADN,^[56] dont S. Furberg en 1952^[57] ou encore L. Pauling en 1953.^[58] Ce dernier, qui avait découvert deux années auparavant l'hélice alpha des protéines,^[59] proposa un modèle de triple hélice dont l'axe interne était formé par les phosphates et dont les nucléobases pointaient vers l'extérieur.

A cette époque, les connaissances scientifiques sur l'ADN étaient nombreuses mais fragmentaires ; la composition chimique était connue depuis longtemps grâce aux travaux de P. Levene de 1930 sur les désoxyriboses,^[60, 61] ou encore ceux de A. Kossel entre 1885 et 1901 sur les bases azotées.^[62, 63] D'autres résultats plus récents publiés par E. Chargaff indiquaient que les différents génomes étudiés présentaient un pourcentage d'adénine et de thymine identique, tout comme celui de guanine et de cytosine.^[64, 65]

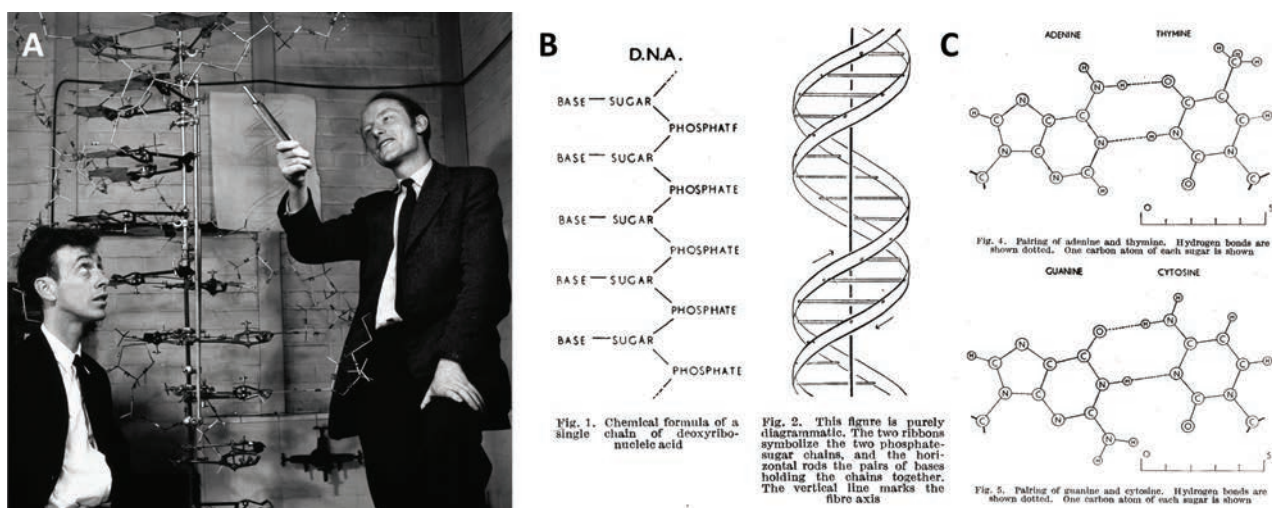


Figure 17 – J. Watson (gauche) et F. Crick (droite) devant leur modèle d'hélice (A), et premiers schémas publiés de la structure de l'ADN (B) ainsi que de l'appariement des bases (C, notons l'omission d'une liaison H pour la paire guanine-cytosine).

L'ingéniosité de J. Watson et F. Crick (Figure 17A) réside dans leur capacité à avoir su compiler les données existantes de la littérature en s'appuyant sur des clichés de diffraction des rayons X de cristaux d'ADN d'une qualité exceptionnelle qu'ils ont en leur possession. Tous sont obtenus par M. Wilkins et R. Franklin également chercheurs au King's College of London parmi lesquels se trouve la fameuse « photo 51 », obtenue fin 1952 et publiée dans *Nature* le 25 avril 1953 (Figure 18).^[55]

Bien que l'histoire soit sujette à controverse, cette photo déterminante de R. Franklin ainsi que les conclusions qu'elle en tirait se trouvaient dans un rapport pour le Medical Research Council que J. Watson et F. Crick se seraient procurés sans l'autorisation de l'auteur.^[66, 67]

Quoiqu'il en soit, la forme de X obtenue suppose une symétrie de type C_2 , typique d'une association de deux chaînes en hélice dans des directions opposées. Il en sera également extrait que les bases sont perpendiculaires à l'axe de rotation et espacées entre elles de $3,4\text{\AA}$; dix de ces bases forment le pas complet de l'hélice de $34\text{\AA}$ d'un diamètre de $10\text{\AA}$.^[53]

Le premier modèle supposé, non publié, était construit à l'envers avec les nucléobases à l'extérieur et les phosphates à l'intérieur ; de l'aveu de J. Watson,^[68] c'est R. Franklin qui rejeta cette hypothèse à cause des répulsions électrostatiques et des problèmes de solubilité qu'elle

engendrerait. De là naîtra l'idée d'association des bases complémentaires entre elles qui forment ainsi l'axe de rotation de l'hélice. Le modèle définitif de l'hélice B sera publié en avril 1953 et sera complété par deux articles en mai et en juillet qui portent sur l'appariement des bases (Figure 17B, C)^[69] ainsi que sur l'étude de l'hélice A.^[70]

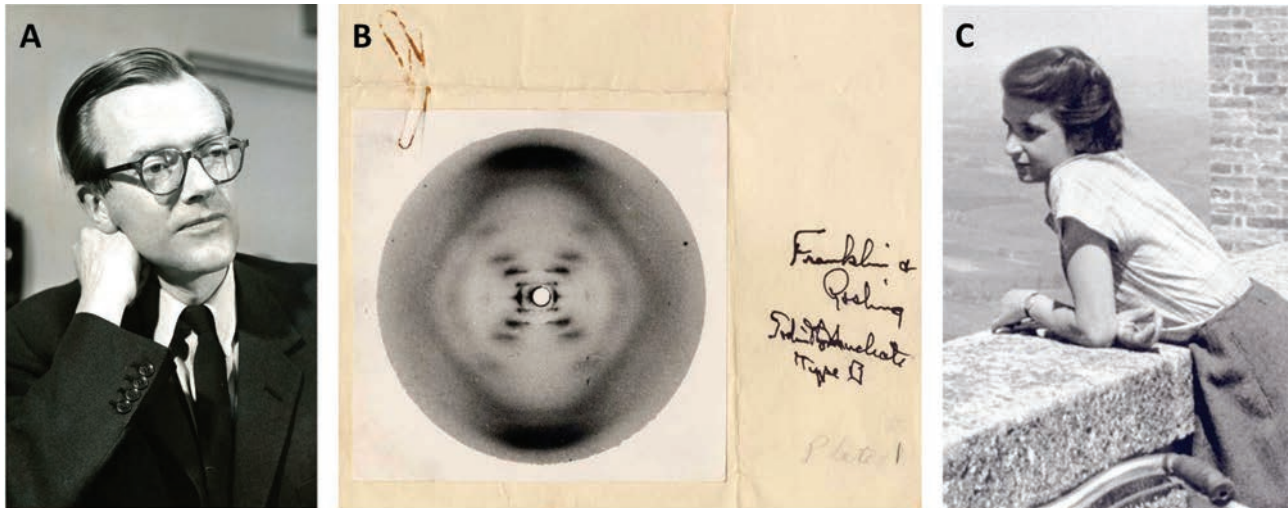


Figure 18 – M. Wilkins, troisième lauréat du prix Nobel de 1962 avec J. Watson et F. Crick (A). Original de la célèbre « photo 51 » (B) signé par son auteur R. Franklin (C).

Ces travaux seront récompensés en 1962 par le prix Nobel de Physiologie ou de Médecine attribué à James Watson, Francis Crick et Maurice Wilkins. Ne pouvant être décerné à titre posthume, il ne put être remis à Rosalind Franklin qui décéda quelques années plus tôt, en 1958.

III.2 Structures usuelles d'ADN et d'ARN

III.2.A Il y a la règle...

Les brins d'ADN et d'ARN sont dotés d'une composition proche et sont issus de la polymérisation d'unités élémentaires appelées nucléotides composées de trois éléments principaux:

- **la nucléobase.** Elle peut se présenter sous la forme d'une purine ou d'une pyrimidine, constituées respectivement de deux hétérocycles à quatre atomes d'azotes ou d'un hétérocycle à deux atomes d'azote. Les purines, l'adénine (A) et la guanine (G), sont communes à l'ADN et à l'ARN alors que dans les cas des pyrimidines, seule la cytosine (C) l'est. En effet, la thymine (T) complète les bases pyrimidiques pour l'ADN alors que sa forme déméthylée, l'uracile (U), est présente au sein de l'ARN (Figure 19A).

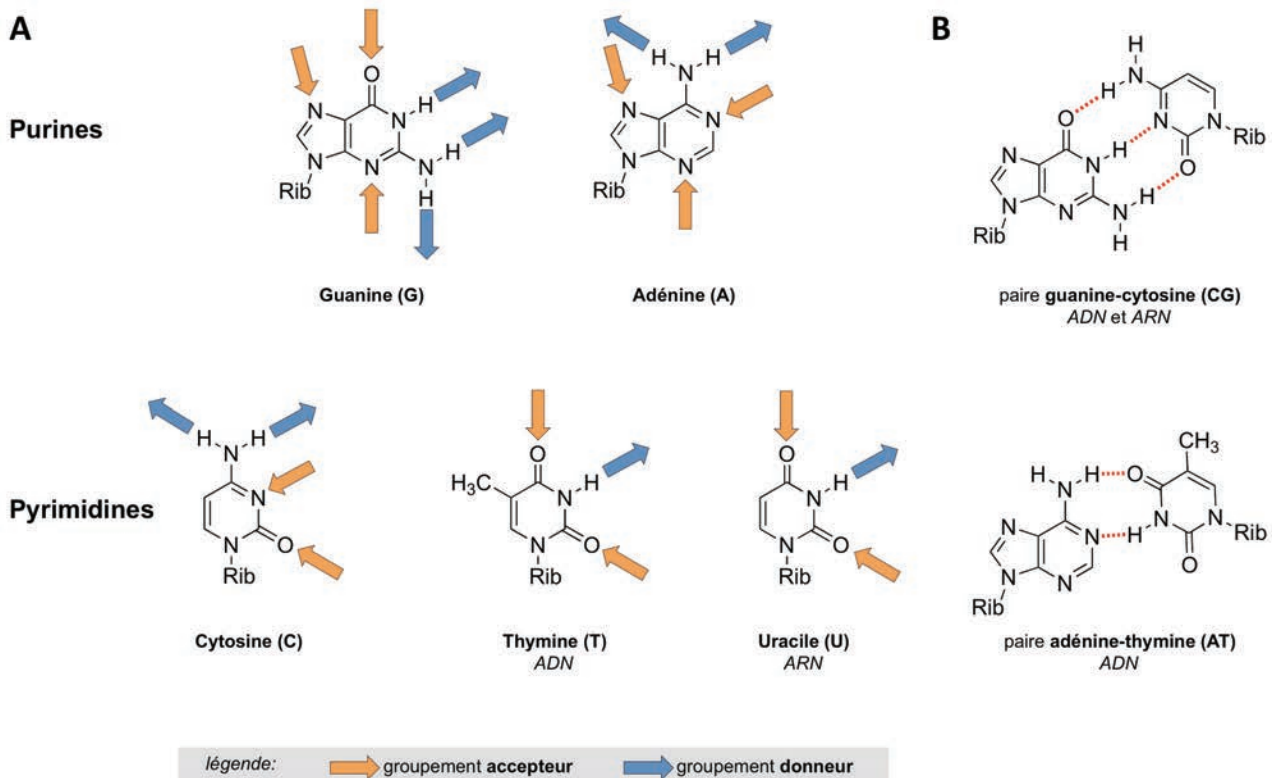


Figure 19 – Représentation des nucléobases de l'ADN et de l'ARN ainsi que de leurs groupements donneurs et accepteurs de liaisons hydrogènes (A). Appariements classiques des bases CG et AT (identique à AU) (B).

Chacune de ces nucléobases possède de nombreux sites donneurs ou accepteurs de liaisons hydrogènes (Figure 19A). Elles peuvent s'apparier entre elles selon la reconnaissance spécifique d'une purine pour une pyrimidine, à savoir l'adénine avec la thymine et la guanine avec la cytosine, en créant respectivement 2 et 3 liaisons hydrogènes *via* leurs faces Watson-Crick (Figure 19B). Ce phénomène est à l'origine de l'association de deux brins complémentaires sous forme d'une double hélice et constitue l'élément novateur du modèle proposé par J. Watson et F. Crick.^[53, 69]

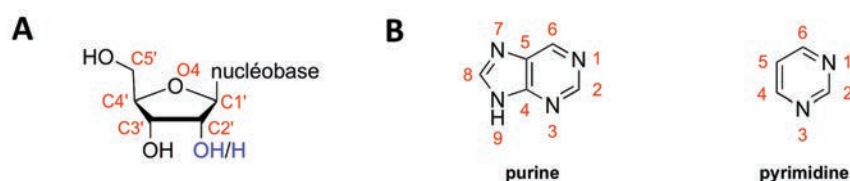


Figure 20 – Nomenclatures des riboses et des désoxyriboses (A) ainsi que des purines et des pyrimidines (B).

- **le sucre**, un furanose, est présent sous sa forme ribose pour l'ARN qui est déshydroxylé en position C2' (voir nomenclature Figure 20A) dans le cas de l'ADN pour former un désoxyribose (Figure 21). A titre d'anecdote, le terme ribose dérive de l'acronyme de « Rockefeller Institute for Biochemistry » où sa découverte a été réalisée par P. Levene en 1909.^[71] Qu'il s'agisse de la forme ribose ou désoxyribose, le furanose est relié à la nucléobase par une liaison *N*-glycosidique en position C1' et porte un hydroxyle en position C3' ainsi qu'un groupement phosphate en C5'.

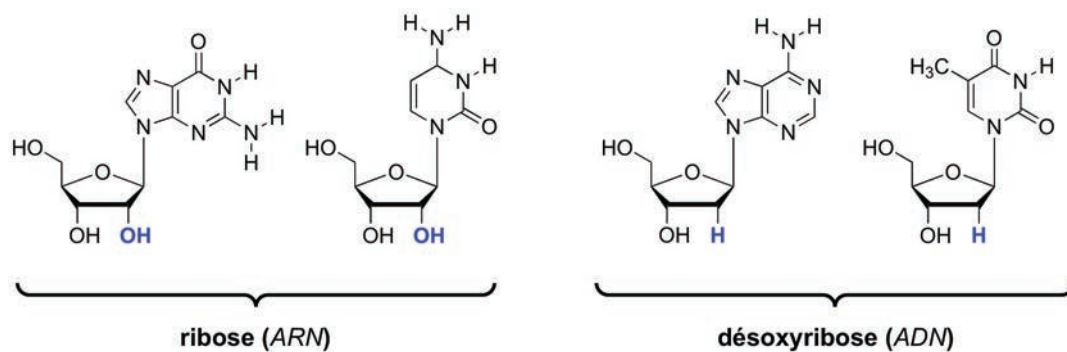


Figure 21 – Différences entre le ribose et le désoxyribose constituant l'ARN et l'ADN, respectivement.

- **le groupement phosphate**, peut contenir de une à trois unités phosphates. Il constitue avec la fonction alcool la seconde fonction réactive des nucléotides. Sous leurs formes triphosphates, ils vont pouvoir réagir entre eux par une réaction d'estérification en chaîne conduisant à la biopolymérisation et donc à la synthèse de polynucléotides, c'est-à-dire de brins d'ADN et d'ARN (Figure 22).

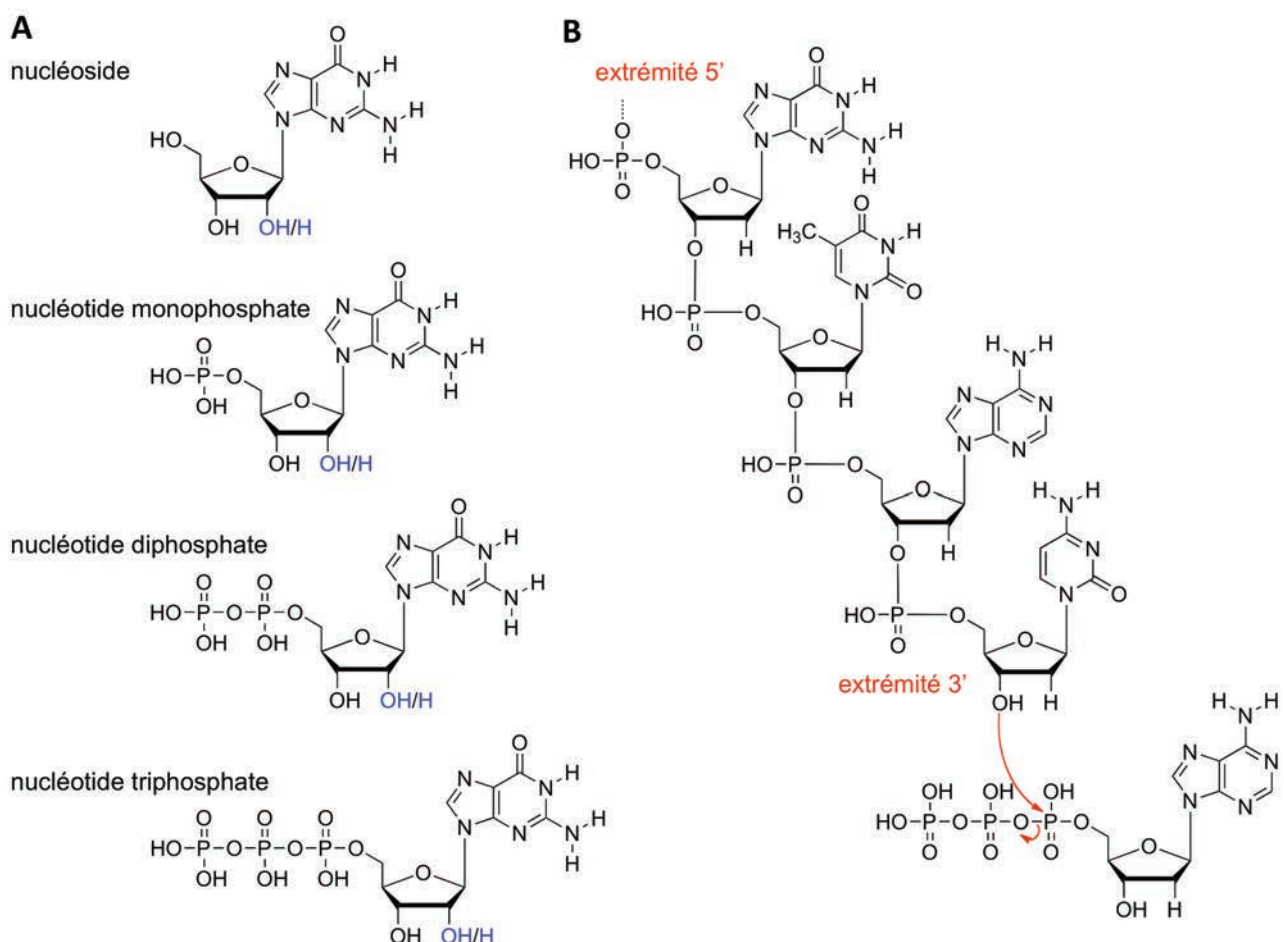


Figure 22 – Alors que les nucléosides en sont dépourvus, les nucléotides peuvent contenir de un à trois groupements phosphates (A). Mécanisme de polymérisation synthétisant une nouvelle liaison phosphodiester avec à chaque étape consommation d'un nucléotide triphosphate et élimination d'un pyrophosphate (B).

Ce processus qui se réalise dans la direction 5'-3'-5'-3' induit une orientation du brin, décisive lors de plusieurs phénomènes biologiques, comme lors de la réplication (voir II.2.A) mais surtout pour l'appariement des brins complémentaires qui se produit toujours de manière antiparallèle (c'est-à-dire avec un brin 5'-3' faisant face à un autre brin orienté 3'-5').

Evidemment, les nucléotides ne sont pas des systèmes figés et peuvent se présenter sous plusieurs conformations. La libre rotation de la liaison *N*-glycosidique (notée χ) peut conduire à l'orientation de la nucléobase soit vers l'extérieur du cycle furane (*anti*), soit vers l'intérieur (*syn*) (Figure 23A). Pour des raisons d'encombrement stérique, la position *anti* est largement majoritaire dans la plupart des édifices alors que la position *syn* ne se forme que dans certaines conditions comme au sein des quadruplexes d'ADN^[72] ou de certaines régions clefs de l'ARN.^[73]

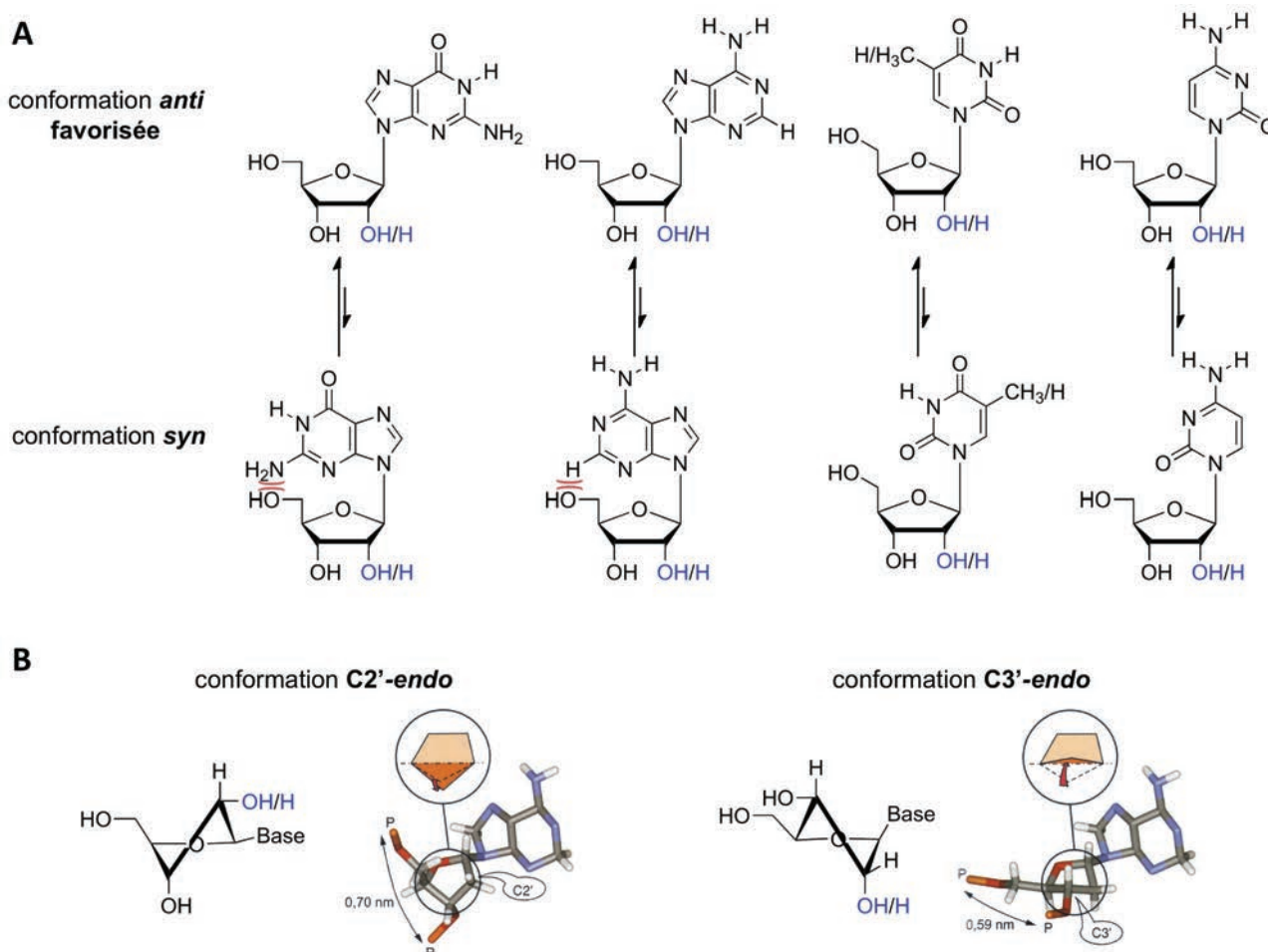


Figure 23 – Les nucléobases peuvent se placer selon deux conformations, bien que la forme *anti* soit très largement majoritaire (A). De même, le sucre adopte selon les conditions la forme C2'-endo ou C3'-endo (B).

Alors qu'il est souvent représenté comme plan, le furanose existe sous la forme C2'-endo ou C3'-endo (Figure 23B) qui traduit la position des carbones C2' et C3' par rapport au plan formé par les liaisons (C4'-O4-C1'). Les deux conformations peuvent être adoptées par les désoxyriboses de

l'ADN alors que la C3'-*endo* est prévalente pour le ribose, à cause de la présence de sa fonction hydroxyle en C2', proche de la nucléobase.^[74]

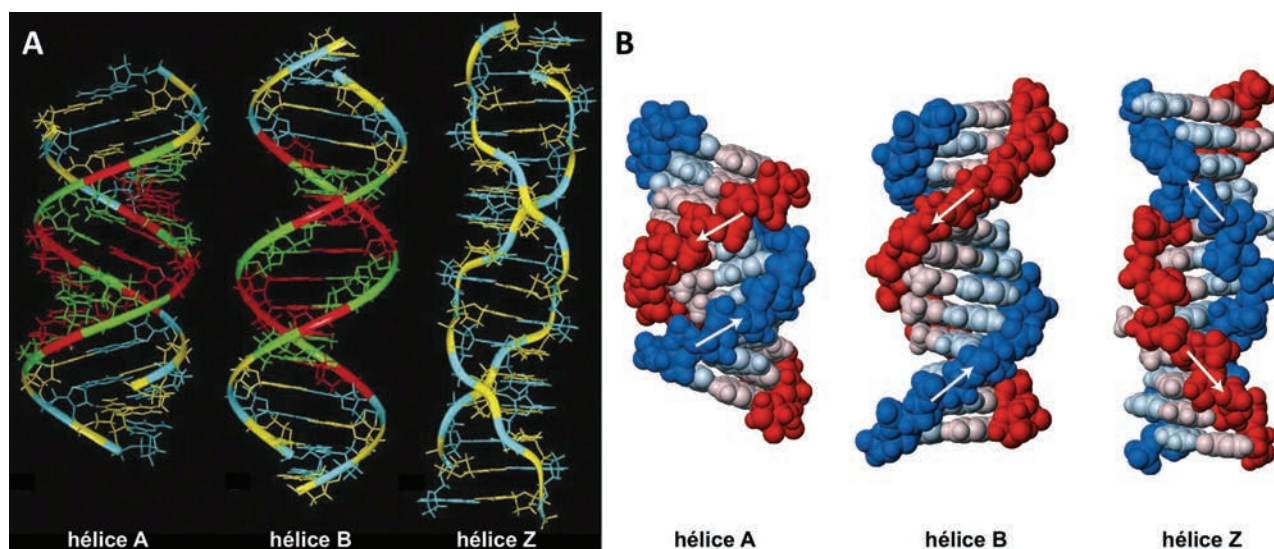


Figure 24 – Structure cristallographique des hélices A, B et Z de l'ADN (A) et représentation tridimensionnelle faisant apparaître l'orientation des brins (flèches dans le sens 5'-3') et les sillons (B).

Les différentes conformations possibles influent sur la morphologie de l'hélice duplexe adoptée par les polynucléotides. De plus, l'appariement des bases ne s'effectue pas exactement au niveau de l'axe de rotation pour la simple raison que les purines et les pyrimidines ne sont pas de même dimension ; l'association n'est donc pas symétrique et va produire pour chaque période de l'hélice deux types de sillons, le grand et le petit, qui s'alternent.

Tableau 1 - Caractéristiques des hélices A, B et Z de l'ADN.

Paramètres	hélice A	hélice B	hélice Z
Sens de l'hélice	droite	droite	gauche
Pas (Å)	28	34	45
Bases par tour	11	10	12
Distance entre bases (Å)	2,6	3,4	3,7
Angle de rotation (° par base)	32,7	34	-30
Diamètre moyen (Å)	23	20	18
Conformation du désoxyribose	C3'- <i>endo</i>	C2'- <i>endo</i>	C2'- <i>endo</i> : pyrimidine C3'- <i>endo</i> : purine
Orientation de la base	<i>anti</i>	<i>anti</i>	<i>anti</i> : pyrimidine <i>syn</i> : purine
Forme globale	courte et large	longue et fine	très étirée et fine

Bien que la forme B soit prépondérante au sein du génome, d'autres structures peuvent se créer (Figure 24) : la forme A est obtenue dans des conditions d'hydratation moindre et semble présenter des répétitions de séquences caractéristiques d(CCGG)_n ; la forme Z, pour « zigzag », est souvent

affectée lors d'alternances entre purines et pyrimidines de type $d(CG)_n$ ^[75] Les caractéristiques de ces trois hélices sont répertoriées dans le Tableau 1.^[5, 34]

Néanmoins, d'autres assemblages d'ADN duplexes, bien que rares, existent parmi lesquels nous citerons la forme B', C, D, E, K, T ou encore W.^[34]

III.2.B ... et toutes ses exceptions.

Le cas de l'ARN est quant à lui un peu particulier. Remplissant des fonctions très distinctes de celles de l'ADN qui a besoin de ses deux brins pour assurer la conservation et la transmission de l'information génétique au cours de la réplication semi-conservative, l'ARN est lui présent dans la cellule sous sa forme simple brin. Sa taille beaucoup plus courte que celle de l'ADN favorise sa structuration de manière intramoléculaire. La présence de séquences complémentaires peut induire localement la formation d'une hélice duplexe dont les propriétés sont proches de celles de l'hélice A formée par l'ADN, avec une conformation C3'-*endo* pour le ribose et l'orientation *anti* de la base (Figure 24 et Tableau 1).^[74]

Néanmoins certains duplexes ARN/ARN bimoléculaires se forment, comme lors de l'association de microARN avec leur cible ARN messager et sont, à séquence identique, plus stables que leurs homologues ADN.^[76-78]

Des hélices hybrides peuvent également voir le jour par l'association d'un brin d'ADN avec un brin d'ARN complémentaire.^[79] Leur morphologie est dépendante des séquences étudiées et se rapproche selon les cas d'une hélice B, ou A, ou d'un intermédiaire entre ces deux formes.^[80] Leur présence au sein du génome est multiple puisqu'elles se forment temporairement au cours de la réplication (fragments d'Okazaki), pendant la phase de transcription ou encore lors de la transcription inverse par des rétrovirus ou par la télomérase.^[80] Cette dernière joue un rôle pivot dans l'immortalisation des cellules cancéreuses qui sera discuté chapitre 2 de cette thèse.

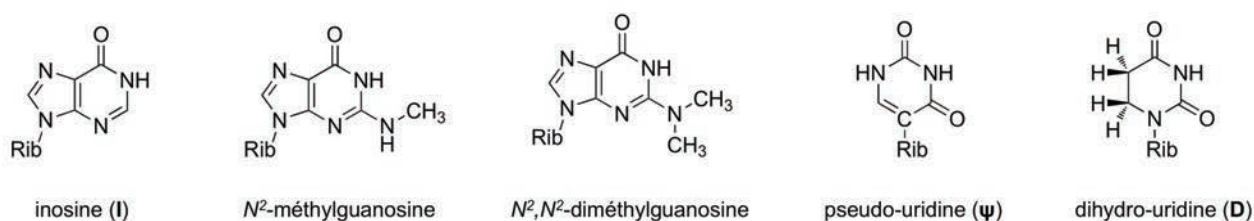


Figure 25 – Structures chimiques de quelques bases modifiées présentes au sein de l'ARNt.

Une autre source de modification provient de la nature même des nucléotides qui ne sont pas limités aux cinq A, T/U, C, G. En effet, des modifications chimiques peuvent être observées soit au niveau du sucre, soit au niveau de la base. Ainsi, alors que seulement quelques variations ont été reportées pour l'ADN,^[81-83] leur nombre atteint 144 pour l'ARN, dont plus de 80 uniquement pour les ARN de transfert (Figure 25).^[84] Pour ces derniers, la transformation s'effectue à la fin de leur maturation et produit le plus fréquemment de l'inosine (I) par désamination de l'adénosine. Très souvent placée en position 5' de l'anticodon de l'ARNt, elle s'apparie aussi bien avec l'uracile, la cytosine ou l'adénine en position 3' de l'ARNm, et permet donc à un seul ARNt de reconnaître plusieurs codons, expliquant *in fine* l'une des origines de redondance du code génétique (Figure 26A).^[85]

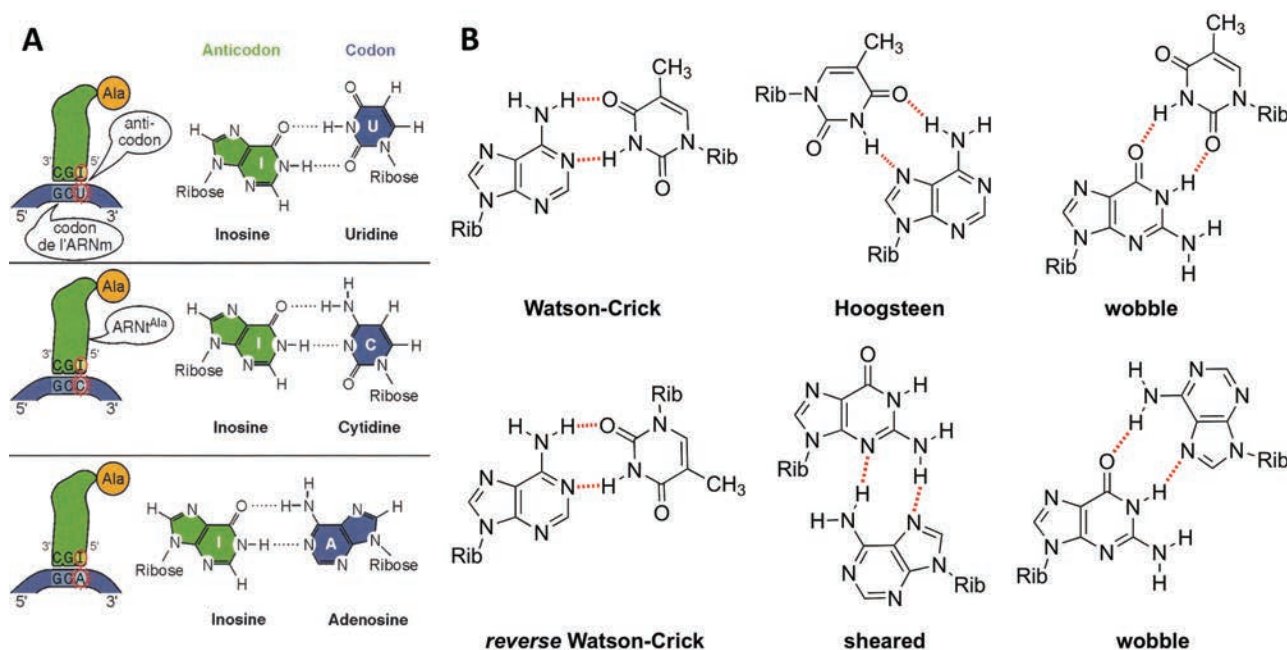


Figure 26 – Appariements de bases imprécis entre l'anticodon de l'ARNt et le codon de l'ARNm (**A**) et présentation des différents appariements possibles entre bases (**B**).

Cette adaptabilité, ou imprécision, provient des multiples possibilités dont disposent les nucléobases pour interagir entre elles par liaisons hydrogènes. De cette manière, des appariements non Watson-Crick, appelés *wobble*, *sheared* ou *Hoogsteen* selon le mode d'interaction, permettent d'augmenter le nombre de couples issus d'un minimum de deux liaisons hydrogènes à vingt-huit (Figure 26B).^[9, 86]

Toutes ces exceptions sur l'assemblage de nucléobases deux par deux témoignent du caractère polymorphique de l'ADN et de l'ARN. Mais les oligonucléotides sont-ils contraints à ne se structurer que sous une forme d'hélice plus ou moins classique ? Nous aborderons justement au cours de la partie suivante les différents édifices tridimensionnels qui peuvent naître de l'association

de plusieurs duplexes, comme les jonctions à 3 et 4 voies ou les *i-motifs* et nous présenterons également les triplexes et les quadruplexes dans lesquels plusieurs nucléobases peuvent se placer dans un même plan pour interagir entre elles.

III.3 Structures non-usuelles

III.3.A De la polymorphie de l'ARN

Au cours d'une revue parue dans *Angew. Chem. Int. Ed.* consacrée aux motifs tertiaires de l'ARN, J. Doudna introduit leurs structures comme « s'apparentant beaucoup plus aux protéines qu'à celles de leurs homologues chimiques d'ADN ».^[87] Cette remarque révèle le degré de complexité des assemblages d'ARN qui, en fonction de leur rôle et de leur séquence, adoptent sur un seul brin plusieurs niveaux d'organisation. En d'autres termes, de multiples structures non-usuelles se forment en différents endroits sur l'oligonucléotide et conduisent à l'élaboration d'édifices complexes.^[88]

Certains de ces arrangements sont communs aux deux acides nucléiques (voir III.3.B), alors que d'autres sont propres à l'ARN (Figure 27) :

- les **mésappariements** (ou *mismatches*) qui apparaissent le long d'une double hélice pour former des bosses (ou *bulges*) ou bien des « boucles internes » selon qu'ils apparaissent sur un seul brin ou en face à face sur deux brins, respectivement.^[87] Tolérés sur l'ARN, ils sont rapidement résolus sur l'ADN par un processus de réparation des mésappariements (ou *mismatch repair*), indispensable à la conservation de l'information génétique au fil des divisions cellulaires.^[43]

- les **épingles à cheveux** (ou *hairpins*) forment une boucle qui relie deux séquences complémentaires duplexes. Il s'agit de la structure la plus fréquente au sein de l'ARN et jouent un rôle majeur au cours de plusieurs processus biologiques puisqu'elles sont impliquées dans la régulation de l'expression génique, offrent également un site de fixation pour les protéines ou sont encore le substrat de réactions enzymatiques.^[89]

- le **stacking coaxial** qui permet de former une double hélice à partir de deux duplexes d'ARN localisés en deux endroits distincts à l'aide d'interactions de π -stacking à l'interface.^[90, 91]

- les **pseudo-nœuds** sont issus de l'interaction entre une séquence contenue dans une *hairpin* avec celle d'un simple brin non contigu.^[92] Ils forment le centre catalytique de certains ribozymes et sont présents dans la transcriptase inverse télomérase.^[93]

- les **interactions boucle-boucle** entre séquences complémentaires contenues dans chacune des deux *hairpins* ont une fonction antisens qui régule l'activité de l'ARN cible, et par conséquent

l'activité biologique qui en découle. Leur existence dans le génome des rétrovirus est également fondamentale. ^[94]

- les **tétraboucles** (ou *tetraloops*) sont une réorganisation complexe qui s'opère depuis quelques séquences spécifiques telles que UNCG ou GNRA (avec N= A, U, C ou G, et R = A ou G). Associées à un récepteur, elles sont fréquemment présentes au cœur des ARN catalytiques : les ribozymes. ^[95, 96]

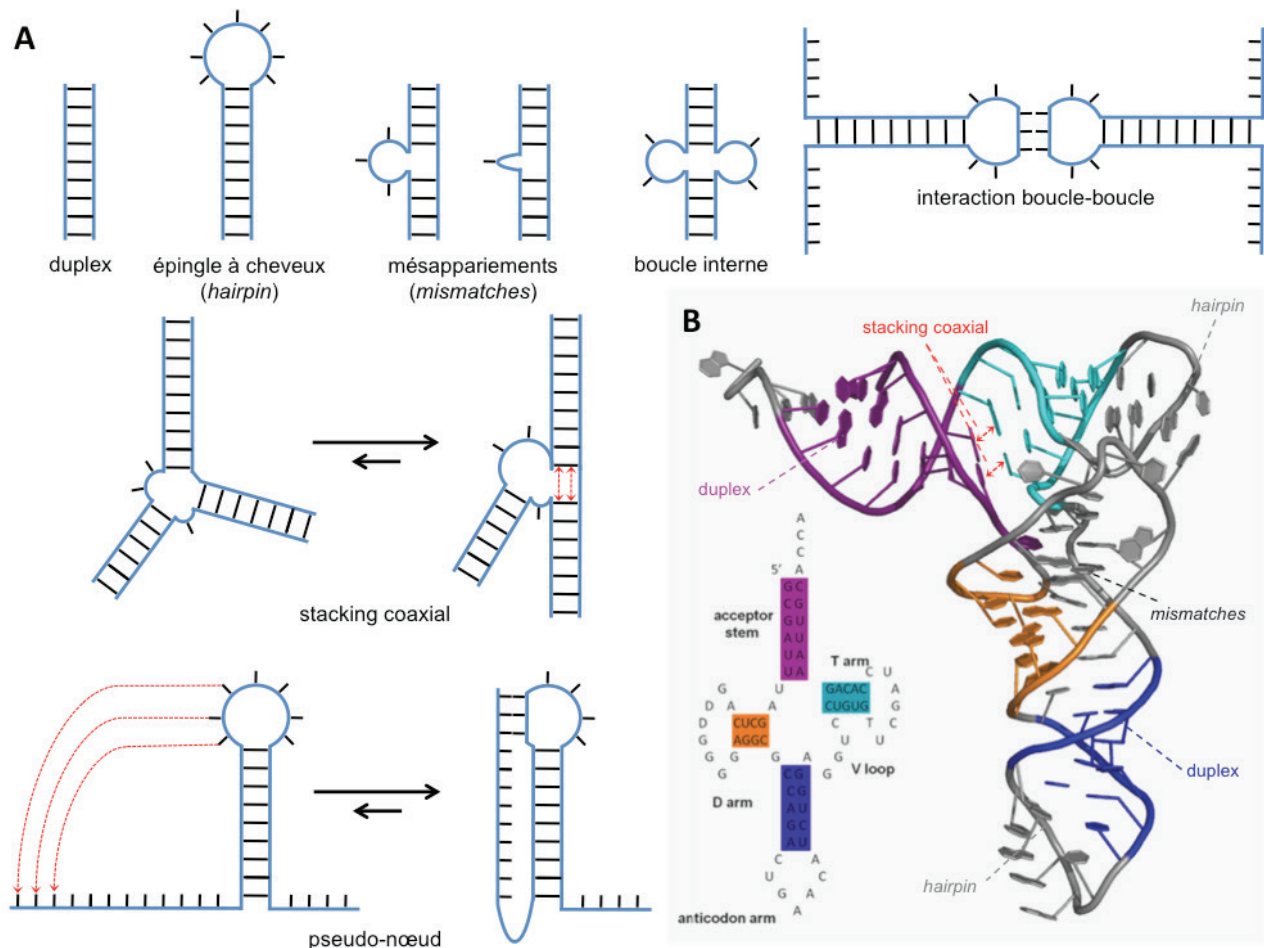


Figure 27 – Structures secondaires adoptées par l'ARN (A) souvent présentes simultanément sur une même séquence, comme au sein de l'ARNt (B).

Tous ces exemples (Figure 27) illustrent bien le fort caractère polymorphique de l'ARN qui, finalement, n'adoptera que très rarement une forme duplexe unique. Dans ces conditions, le terme structure tertiaire est préféré face au terme « structure non-usuelle », réservé à l'ADN, qui dans une très large majorité se trouve sous forme d'une double hélice.

Néanmoins, dans certaines conditions de séquences et de milieu, des structures tridimensionnelles peuvent voir le jour dans des régions clefs du génome et sont en général communes à l'ADN et l'ARN.

III.3.B Principales structures non-usuelles

III.3.B.a Les triplexes

Les triplexes naissent de l'insertion d'une chaîne polynucléotidique dans le grand sillon d'un duplexe. Alors que ce dernier renferme deux brins antiparallèles mettant en jeu des liaisons hydrogènes de type Watson-Crick, le 3^{ème} brin nommé TFO (pour *triplex-forming oligonucleotide*) interagit avec l'une des deux chaînes par des interactions de type *Hoogsteen* ou *reverse-Hoogsteen*, selon qu'il est orienté de manière antiparallèle ou parallèle, respectivement (Figure 28). Cet ajout n'apporte qu'une très faible modification de la conformation de l'hélice initiale.^[97-100]

Se formant aussi bien avec l'ADN qu'avec l'ARN ou un mélange des deux, les triplexes nécessitent la présence de longues séquences répétées d'un même nucléotide, comme par exemple d'un duplexe formé de chaînes poly(A) et poly(U) dans lequel s'insère une autre chaîne poly(U).^[101]

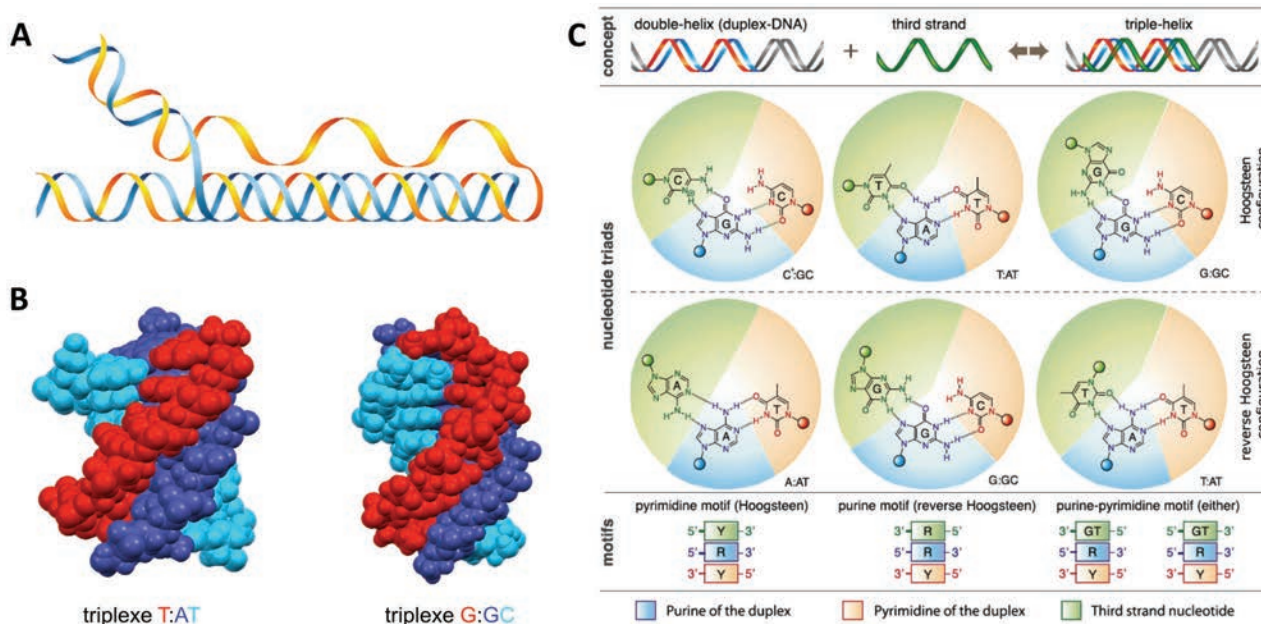


Figure 28 – Formation d'un triplex intramoléculaire (A) et structures RMN des triplexes T:AT et G:GC (B). Triplets de nucléotides s'assemblant selon deux conformations : *Hoogsteen* et *reverse-Hoogsteen* (C, dans le cas de l'ARN, la thymine est remplacée par l'uracile).

L'existence de ces structures *in vivo* n'a jamais été démontrée mais plusieurs études *in vitro* et des analyses bioinformatiques supposent leur implication dans une variété de mécanismes incluant la régulation de la transcription, l'organisation des chromatines, ou encore la réparation de l'ADN. Ces sites cibles pouvant accueillir un TFO sont nommés TTS (*triplex target site*) et sont estimés à 3,5 millions pour le génome humain.^[97]

De telles interprétations ont donc ouvert la voie au développement de petites séquences TFO utilisées entre autres comme outil biotechnologique pour l'évaluation de la translocation des protéines sur l'ADN ou de l'activité des topoisomérases.^[102, 103] Notons également leur application naissante en tant qu'agent thérapeutique contrôlant l'expression génique par régulation de séquences promotrices cibles,^[104] comme reporté récemment sur l'oncogène C-myc.^[105]

III.3.B.b Les jonctions à trois voies

Les jonctions à trois voies, ou TWJ pour *Three-Way Junctions*, résultent de la convergence en un point de jonction de trois hélices duplexes issues d'un appariement complet ou partiel. Dans ce dernier cas, la présence d'une bosse (*bulge*) aux abords de la cavité centrale est observée et influe sur la conformation de la structure. En général, pour minimiser les répulsions électrostatiques induites par les nombreux groupements phosphates des nucléotides, les bras tendent à s'écarter le plus loin possible les uns des autres en formant trois angles approximativement égaux (d'environ 60°, voir Figure 29C),^[106] donnant à la TWJ une forme symétrique en étoile (ou *extended*, Figure 29A).^[107, 108] De plus, pour relâcher de la tension et atténuer la gêne stérique, quelques bases se désappariaient au niveau du point de jonction, formant ainsi une cavité hydrophobe d'un diamètre de 12Å et offrant à l'édifice une meilleure flexibilité.^[109]

Notons que cette conformation évolue lors de la présence d'un *bulge* et d'une certaine concentration en ion bivalent, comme Mg^{2+} ou Ca^{2+} (>1mM), vers une forme *Y-shaped* où deux duplexes parfaitement appariés interagissent entre eux par stacking coaxial pour former une seule hélice. Le troisième bras reste indépendant et forme avec les deux autres un angle plus ou moins prononcé selon la taille de la bosse (Figure 29B).^[107, 108, 110]

Dans un contexte biologique, ces structures se forment lors de séquences répétées de tri-, tétra- ou pentanucléotides qui permettent au brin d'ADN de glisser pour former une *hairpin* dont la base est une jonction à trois voies. Ces glissements se forment au cours de la réplication sur le brin retard au niveau du brin mère ou du brin fille, et conduisent respectivement au raccourcissement (sous-réplication) c'est-à-dire à une perte d'information génétique, ou à l'élongation du brin nouvellement synthétisé (Figure 29D). Une longueur excessive de ces répétitions entraîne de fortes modifications et se trouve liée au développement de maladies, comme le syndrome de l'X fragile ((CGG)_n), l'ataxie de Friedreich ((GAA)_n) ou la dystrophie myotonique de Steinert ((CTG)_n).^[111, 112]

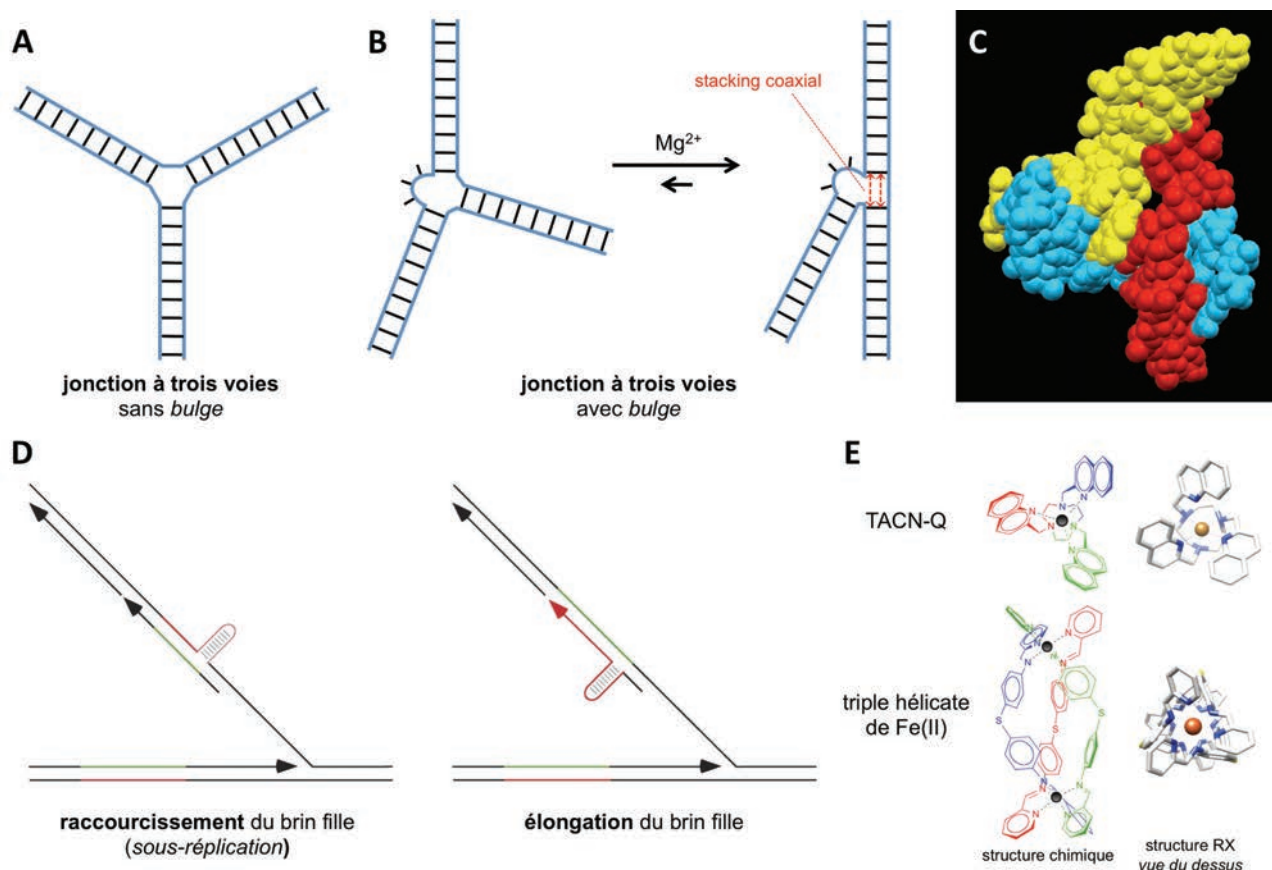


Figure 29 – Schémas d’une jonction à trois voies (TWJ) parfaitement appariée (A) ou présentant un bulge (B), et structure RMN d’une TWJ d’ADN (C). La formation d’une TWJ sur le brin mère ou fille au cours de la réplication altère la longueur du brin fille (D). Présentation des deux seuls modèles de ligands de TWJ connus à ce jour (E).

Les jonctions d’ADN à trois voies sont également estimées par certaines équipes de recherche comme des modèles simplifiés de fourches de réplication.^[113, 114]

L’arrangement en triple hélice est également adopté par l’ARN au sein de plusieurs ribozymes, de certains ARNr (5S et 16S)^[115] ou encore de la télomérase^[116] et jouent un rôle pivot dans l’internalisation de virus tels que celui de l’hépatite C.^[117]

Pour leurs multiples implications biologiques, le développement de ligands spécifiques de ces structures semble particulièrement pertinent mais reste naissant puisqu’à ce jour seules deux séries sont référencées ;^[118] la première a été mise au point par l’équipe de M. Hannon^[99, 119, 120] alors que la seconde, construite autour de dérivés TACN, a été synthétisée et étudiée au cours de cette thèse (Figure 29E), mais ne sera pas détaillée dans ce manuscrit.^[121] Notons que ces molécules présentent toutes deux une symétrie de type C_3 identique à celle des jonctions à trois voies dites *extended*.

En outre, leur structure particulière a également été employée dans le développement d’outils moléculaires,^[15, 16, 122] dont des aptamères pour la détection de cocaïne^[123-126] ou de stéroïdes.^[127]

III.3.B.c Les jonctions à quatre voies

A l'instar des TWJ, les jonctions à quatre voies, ou FWJ pour *Four-Way Junctions*, se forment à l'intersection de 4 hélices d'ADN ou d'ARN (Figure 30A, B et C).^[128] Proposées en 1964, elles sont également connues sous le nom de jonctions de Holliday, en référence à son découvreur.^[129]

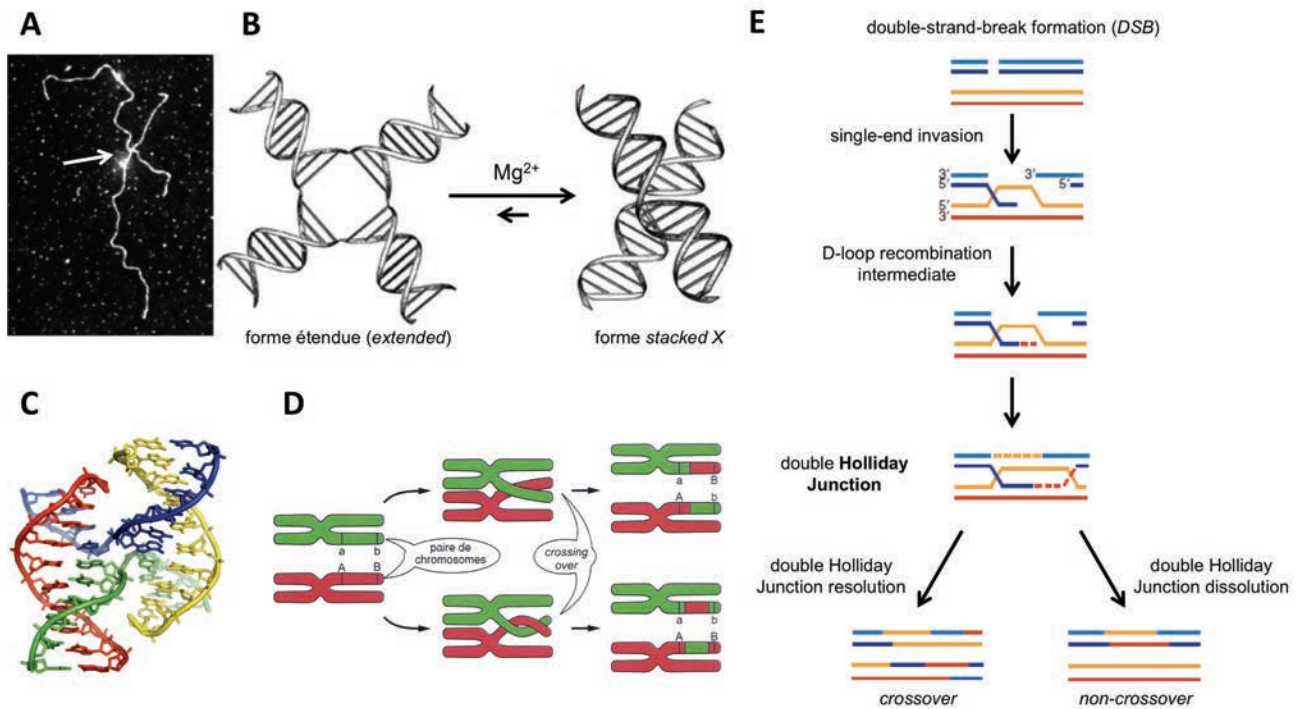


Figure 30 – Photomicrographie d'une jonction à quatre voies (A), équilibre entre les formes *extended* et *stacked X* (B) et structure cristallographique d'une FWJ (C). Ces jonctions entrent en jeu dans les recombinaisons homologues permettant l'échange de chromatides lors de la méiose (D) ou lors de la réparation de cassures doubles brins (E).

La structure de ces édifices est très dépendante de la concentration en cations bivalents (ex : Mg^{2+} , Ca^{2+}) présents dans le milieu (Figure 30B).^[108]

En effet, dans le cas de l'ADN, leur absence conduit à la formation d'une jonction étendue (ou *extended*) où les quatre bras duplexes sont indépendants et se rejoignent en un point de jonction de géométrie carrée. Cette conformation permet de minimiser les répulsions électrostatiques engendrées par les nombreux groupements phosphates qui s'amassent autour de la cavité centrale.^[130] A l'inverse, une concentration en ion magnésium supérieure à 100 μM neutralise ces charges et entraîne un rapprochement des bras qui peuvent alors interagir deux à deux par stacking coaxial. La forme globale s'apparente alors à un X d'un angle d'environ 60° et se nomme *stacked X structure*. Deux conformations dites parallèles et antiparallèles existent et leur répartition dépend des séquences de nucléotides mises en jeu.^[107, 108]

En raison de leur différence de taille et de profondeur de leur sillon, les hélices A d'ARN forment toujours une jonction de type *stacked X* dont la morphologie évolue en fonction de la

quantité d'ions bivalents présente, passant d'une conformation parallèle (sans cations) à antiparallèle ($>0,5\text{mM}$), puis complètement ouverte avec un angle de jonction maximal de 90° .^[107, 108]

Sous leur forme ADN, ces structures sont impliquées dans la recombinaison homologue qui est un processus clef de la vie cellulaire. Elle participe d'une part à la diversité génétique en permettant l'échange de segments homologues entre chromatides au cours de la méiose (Figure 30D),^[131-134] et permet d'autre part de réparer les cassures doubles brins qui pourraient apparaître sur le génome à la suite de radiations ionisantes ou au contact d'un composé chimique toxique (Figure 30E).^[135-137] Ce mécanisme de réparation qui a lieu au cours de la phase S et G1 du cycle cellulaire est primordial pour la conservation du matériel génétique et son dysfonctionnement peut être impliqué dans l'oncogenèse.^[138-140] Les FWJ deviennent alors des cibles d'intérêt potentielles actuellement à l'étude par différents groupes.^[118, 141-147]

La recombinaison homologue participe également au maintien ou à l'augmentation de la longueur des télomères dans un contexte tumoral. En effet, ce phénomène est soit contrôlé par la télomérase (environ 85% des lignées cancéreuses),^[148] soit par mécanisme totalement indépendant nommé ALT (pour *alternative lengthening of telomeres*, dans 15% des cas), faisant intervenir une jonction de Holliday.^[149]

Notons également leur participation à la formation de la boucle télomérique (ou t-loop) où les FWJ, quelle que soit leur séquence, s'associent à l'une des protéines de la shelterin, la TRF2.^[150-152]

Dans le cas de l'ARN, leur rôle est principalement structural et influe sur l'activité de certains ribozymes.^[108, 153]

III.3.B.d Les *i*-motifs

Les *i*-motifs se forment au départ d'une séquence riche en cytosines qui dans certaines conditions se replie sur elle-même en plaçant deux paires de brins de manière orthogonale et antiparallèle. Chacune de ces paires résulte de l'association parallèle de deux brins à l'aide de liaisons hydrogènes entre des cytosines natives et hémiprotonées. Le nombre de cytosines se faisant face évolue classiquement entre 2 et 4 (Figure 31).^[154-156]

Avec un pKa compris entre 4,2 et 4,8 en fonction de la température et de la force ionique,^[157] la protonation des cytosines implique donc que la formation des *i*-motifs est dépendante d'un milieu

légèrement acide, avec une gamme de pH optimale évaluée entre 4,2 et 5,2.^[155] Cette propriété soulève donc la question de leur pertinence dans un contexte biologique ; cependant, plusieurs études font état de leur présence à pH 7,^[158] favorisée notamment par des interactions entre les nucléobases des boucles^[159, 160] ou des conditions de *supercoiling*.^[161, 162]

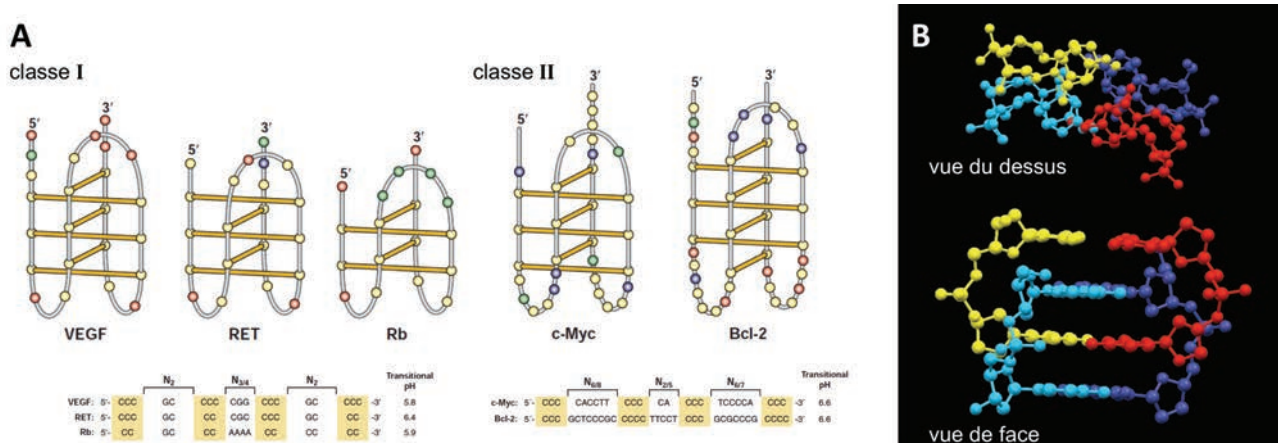


Figure 31 – Les *i-motifs* présentent un nombre de cytosines (sphères jaunes) variable et sont répartis en deux classes selon la longueur de leur boucle (**A**). Partie de la structure cristallographique d'un *i-motif* tétramoléculaire d(ACCCT) (**B**).

En effet, lorsqu'ils sont issus d'un assemblage intramoléculaire, ces édifices sont dotés de trois boucles plus ou moins longues (Figure 31A) qui leurs valent d'être discriminés en deux catégories notées classe I et classe II, respectivement, les boucles plus longues conduisant à une stabilité globale plus grande.^[159] En plus de la forme monomérique, notons l'existence *in vitro* de modèles dimériques^[163, 164] et tétramériques.^[165-168]

Alors que leur rôle au sein du génome n'est pas clairement élucidé, il apparaît que les séquences formant potentiellement des *i-motifs* sont présentes dans des régions clefs régulatrices, comme par exemple au niveau des promoteurs d'oncogènes RET,^[169] Rb,^[170] Bcl2,^[160] VEGF ou encore Myc^[161] et Kit ainsi qu'au niveau des télomères,^[171] et constituent des cibles thérapeutiques intéressantes.^[172]

Leur présence est liée à celle de séquences riches en guanines sur le brin d'ADN complémentaire qui lui, pourra alors former un G-quadruplexe (voir partie suivante).

Au-delà de leur aspect biologique, ces structures non-usuelles sont employées pour diverses utilisations biotechnologiques^[15] parmi lesquelles peuvent être citées le développement de portes logiques à base d'ADN,^[173, 174] ou encore comme outil moléculaire pour les nanotechnologies^[175, 176] ou bien, plus récemment, pour le suivi des variations de pH intracellulaires.^[177]

Enfin, nous n'avons pas évoqué le cas de l'ARN qui pour sa part forme des *i-motifs* beaucoup moins stables que leur homologues ADN et sont par conséquent peu représentés dans la littérature scientifique.^[157, 178]

III.3.B.e Les G-quadruplexes

Les quadruplexes se forment à partir de séquences nucléotidiques riches en guanines et sont sans nul doute la structure non-usuelle la plus étudiée avec près de 3500 références bibliographiques, dont plus de 1000 depuis 2011 (selon Web of Knowledge).

Comme présenté Figure 19 et Figure 32A, les guanines ont la capacité d'interagir selon leurs faces Watson-Crick, qui entrent en jeu dans les liaisons G-C de la double hélice, et Hoogsteen. Ainsi, en utilisant simultanément ces deux modes, quatre guanines peuvent s'auto-assembler dans un même plan pour former une tétrade de guanines, ou G-quartet, favorisée par la création d'un réseau de huit liaisons hydrogènes.^[12, 13]

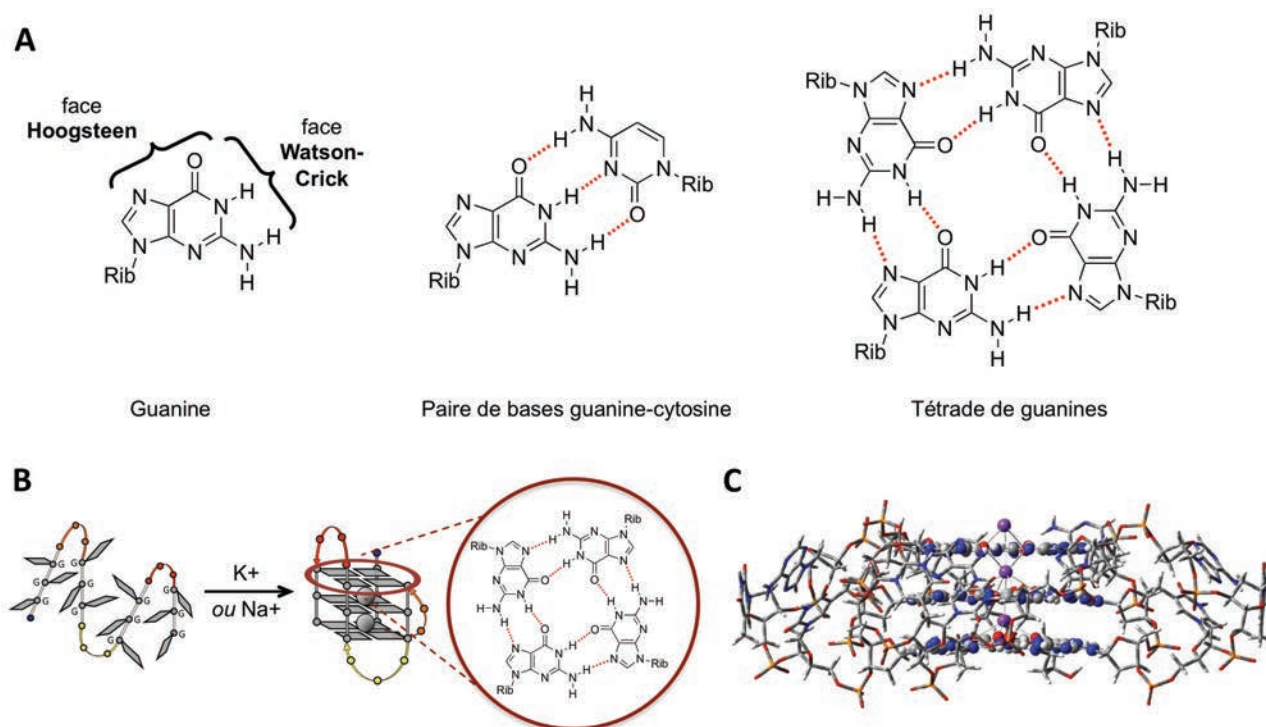


Figure 32 – Une guanine présente deux faces principales : la face Watson-Crick permet la formation d'une paire de bases GC alors que la face Hoogsteen entre également en jeu dans la formation d'une tétrade de guanines (A). Formation d'un quadruplexe d'ADN à partir d'une séquence présentant des guanines consécutives (B) et structure d'un quadruplexe obtenu par modélisation moléculaire (C).

Lorsqu'une séquence d'ADN ou d'ARN présente des motifs répétés de guanines consécutives, elle se replie pour donner naissance à un quadruplexe qui résulte de l'empilement des tétrades de guanines. Ces dernières s'auto-stabilisent à l'aide d'interactions de type π -stacking et également par

le piégeage d'un cation alcalin, principalement Na^+ ou K^+ en milieu physiologique (Figure 32B et C).

Le nombre de tétrades qui constitue un quadruplexe est variable et se situe dans une large majorité entre deux (ex : TBA) et cinq (ex : $(\text{TG}_5\text{T})_4$),^[12, 13] bien que des structures beaucoup plus grandes contenant jusqu'à 4000 guanines existent et s'organisent sous forme de nanofils de plusieurs centaines de nanomètres, visibles par AFM.^[6, 179-183]

Tout comme les *i-motifs*, ces édifices se forment de manière intramoléculaire dans un contexte biologique au départ d'un seul brin, alors qu'*in vitro* plusieurs modèles ont été développés au départ de deux,^[184, 185] trois,^[186] et quatre brins.^[187, 188]

En fonction de l'orientation des brins entre eux (*i.e.* parallèles ou antiparallèles) et de la position de la guanine par rapport au cycle pentose (*i.e.* *syn* ou *anti*), trois conformations types de quadruplexes existent^[72, 74, 189, 190] et sont illustrées Figure 33A :

- la **conformation antiparallèle** (ou *basket type*) : deux brins sont orientés 3'-5' alors que les deux autres le sont en sens inverse 5'-3' ; les bases sont en position *syn* (en rouge) et *anti* (en jaune).
- la **conformation parallèle** : tous les brins sont orientés dans la même direction ; toutes les bases sont en *anti*.
- la **conformation hybride** (nommée *hybrid type* ou [3+1]) : porte bien son nom puisqu'il s'agit d'un compromis entre les deux précédentes formes avec trois brins dans la même direction et un autre en antiparallèle ; les bases sont présentes en position *syn* et *anti*. Notons qu'il existe deux conformations hybrides en équilibre (hybrid-1 et hybrid-2).

Selon les conditions et les séquences étudiées, un quadruplexe d'ADN choisira préférentiellement l'une ou l'autre de ces conformations alors que l'ARN, à cause de la fonction hydroxyle supplémentaire du ribose qui impose à la guanine de se placer en *anti*, adoptera logiquement la conformation parallèle où toutes les bases sont en *anti*.^[191, 192]

Les nucléobases qui ne sont pas impliquées dans la formation des tétrades de guanines composent les boucles qui selon leur morphologie sont nommées boucles latérales (ou *edgewise*), diagonales, et *chain reverse* (ou *propeller*) (Figure 33B).^[72, 74, 189, 190] La nature de ces boucles et leur longueur influent sur la conformation du quadruplexe ainsi que sur sa stabilité. Dans une certaine mesure, des boucles courtes favorisent une forme parallèle et induisent un quadruplexe plus

stable ; leur élongation entraîne à l'inverse un déplacement vers une conformation hybride et une diminution de stabilité.^[193]

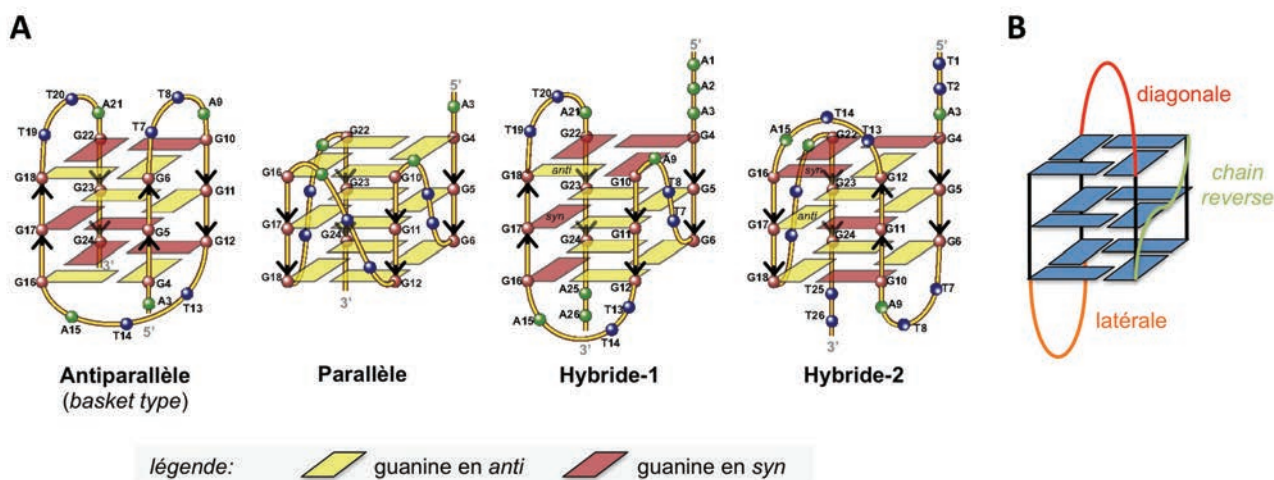


Figure 33 – Présentation des différentes conformations adoptées par un quadruplexe (conformation de la base et orientation des brins, (A) et dénomination des boucles (B).

Cette stabilité est également corrélée au nombre de tétrades qui constitue le quadruplexe : plus il est élevé, plus le quadruplexe est stable.^[194]

Cependant, tous ces facteurs sont interdépendants et il est donc important de considérer que l'énergie globale d'un système est fonction de plusieurs paramètres à la fois.

L'intérêt suscité par les quadruplexes est multiple et grandissant. Leur présence au sein du génome dans des régions régulatrices clefs telles que les télomères^[12, 195, 196] ou les promoteurs d'oncogènes^[10, 154] sous leur forme ADN ou encore au niveau des séquences 5'-UTR des ARNm^[197, 198] ou au niveau des TERRA^[199-201] en fait une cible biologique prometteuse pour le développement de nouvelles thérapies antitumorales. Ce champ d'application explique pourquoi depuis plus d'une quinzaine d'années, les scientifiques cherchent à développer des composés chimiques, ou ligands,^[11, 192] de plus en plus sélectifs de ces structures dont certains, en association avec d'autres anticancéreux,^[196] valident ce type d'approche même si le chemin à parcourir reste encore long. Nous reviendrons plus en détail sur cet aspect des G-quadruplexes au cours du second chapitre de ce manuscrit où nous étudierons également *in vitro* de nouveaux ligands biomimétiques synthétisés au laboratoire.

Le domaine de la biotechnologie s'est également emparé de l'édifice pour en faire un outil adaptable pour la détection de cations^[202, 203] ou de protéines,^[204, 205] mais également pour le développement de portes logiques^[206] ou de tests biophysiques telles que la détection de la

télomérase.^[207] Ces différents exemples appartiennent au domaine du DNzyme en pleine expansion depuis quelques années et qui sera discuté au chapitre 3 de cette thèse.

III.4 Bilan des structures d'ADN et d'ARN

Tout au long de cette partie III, nous avons illustré l'extraordinaire modularité de l'ADN et de l'ARN. Loin de ne se résumer qu'à la célèbre double hélice découverte il y a maintenant 60 ans, les nucléobases ont la propriété de s'apparier entre elles deux par deux pour former bien évidemment un duplexe, mais aussi dans certaines conditions une jonction à trois voies, une jonction de Holliday, une *hairpin*, etc. La présence de multiples sites donneurs et accepteurs de liaisons hydrogènes au sein de chaque nucléobase leur permet d'interagir par plusieurs faces à la fois, leur conférant la possibilité de former des édifices où elles se placent à trois ou quatre dans un même plan, comme pour les triplexes ou les quadruplexes, respectivement.

En présentant pour chacun des édifices non-usuels leur rôle biologique réel ou supposé, il advient qu'en plus de leur fonction structurale importante pour les ARN, ils sont présents dans des régions clés du génome, principalement au niveau de séquences régulatrices, comme les télomères ou les promoteurs d'oncogènes.

L'existence de telles architectures à un si haut degré de complexité n'est paradoxalement régie que par la présence de liaisons faibles. Ces liaisons inter- ou intramoléculaires permettent à la Nature de créer des assemblages stables (ex : appariement de brins d'ADN complémentaires sous forme duplexe qui protège l'information génétique) qui seront au besoin dissociés (ex : l'ADN duplexe s'ouvre au cours de la réplication ou de la transcription) puis réassemblés par la suite pour revenir à l'état initial.

Cette suite d'événements ne peut être régie que par des interactions non covalentes. Ces propriétés qui permettent à un système moléculaire d'évoluer, de « respirer » en fonction des conditions du milieu sont celles de la chimie supramoléculaire.

IV. De l'ADN naturel à la molécule supramoléculaire

L'introduction du terme « supramoléculaire » par J.-M. Lehn en 1978^[26] sera suivie quelques années plus tard par le couronnement de cette chimie innovante par le prix Nobel de 1987 (Figure 5).^[27-29] La multitude de travaux s'appuyant sur ce concept témoigne des nombreux avantages qui peuvent être additionnés à la chimie moléculaire classique pour en décupler les propriétés et les possibilités.

Pouvant aussi bien être présents dans la Nature qu'issus de l'ingéniosité des chimistes, les systèmes supramoléculaires peuvent parfois tirer le meilleur de chacun de ces mondes. En effet, l'utilisation de nucléobases pour leurs propriétés d'interactions et de reconnaissances spécifiques est judicieuse puisqu'elle confère à la molécule un premier état où les bases sont indépendantes, et un second état où elles s'auto-assemblent.

Afin d'illustrer cette idée, nous présenterons au long de cette partie différents systèmes supramoléculaires de la littérature allant du petit mime de duplexe à deux paires de bases jusqu'au développement de polymères. Précisons que nous ne parlerons pas ici des longues séquences d'oligonucléotides fonctionnalisées par un groupement organique mais bien de molécules de synthèses fonctionnalisées par des nucléobases. En effet, dans le premier cas, la supramolécularité est directement amenée par la biomolécule et ne relève pas de la conception d'une nouvelle architecture élaborée par la chimie de synthèse.

Le cas particulier des auto-assemblages de tétrades de guanines, constituant l'objet d'étude de ce travail de thèse sera introduit par une approche bibliographique et nous conduira à la présentation des molécules de type TASQ hydrosoluble élaborées au laboratoire et dont les multiples applications seront discutées au long des chapitres 2 et 3.

IV.1 Les nucléobases comme motifs supramoléculaires

Le défi de reproduire par synthèse un duplexe où deux bases complémentaires s'apparient a été relevé par l'équipe de J. Sessler.^[8] Le premier modèle contenait une guanine et une cytosine séparées par un pont diéthylamide lui même relié à une chaîne polyéthylène glycol (Figure 34A). L'interaction entre deux molécules devait conduire à la formation d'un duplexe à deux paires de bases, mais la trop grande flexibilité du système n'était pas favorable à l'auto-association.^[208] Pour pallier ce problème, des armatures plus rigides 1,8-diéthynylantracène et 4,6-diéthylidibenzofurane ont été choisies et offrent de meilleurs résultats, surtout dans le cas du dérivé anthracène où l'orientation des nucléobases est parallèle, exactement comme au sein d'une hélice d'ADN (Figure 34C).^[209, 210]

Notons qu'ici les bases sont accompagnées de leur ribose sous une forme protégée soit par des groupements acétyles plus favorables à la formation de l'auto-assemblage, soit par des *tert*-butyldiméthylsilyles (TBDMS) dont le rôle est d'augmenter la solubilité de l'édifice dans les solvants organiques. L'approche est également validée par l'utilisation du couple A-U qui curieusement, présente de meilleurs résultats que l'équivalent G-C.^[209, 210]

Une architecture similaire a été publiée par l'équipe de G. Gokel dans laquelle le *template* sélectionné est un amine éther couronne diaza-18-crown-6 renfermant une adénine et une thymine (Figure 34B) et dont l'association sous forme de dimère a été démontrée par RMN ^1H .^[211]

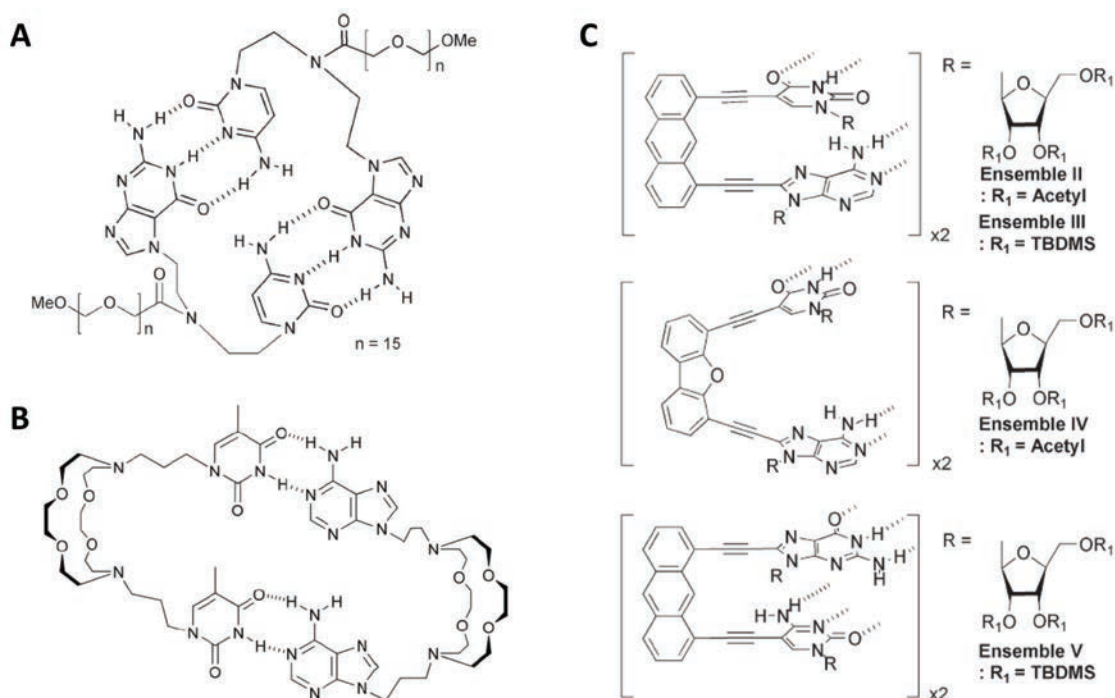


Figure 34 – Structures dimériques mimant un petit duplexe d'ADN à deux paires de bases construites sur une armature flexible (A, B) ou rigide (C).

La construction de portes logiques peut aussi se faire grâce à la chimie supramoléculaire des nucléobases. Comme le démontrent H. Baytekin et E. Akkaya, la fluorescence du 4',6-diamidino-2-phenylindole (ou DAPI) subit une diminution ainsi qu'un déplacement vers les plus basses longueurs d'ondes uniquement en présence simultanée de dAMP et de dTMP (Figure 35). Cet assemblage requiert la présence de trois composants et est régi par des interactions de type électrostatique entre les groupements amidines du colorant et les phosphates des nucléotides, ainsi que par deux liaisons hydrogènes Watson-Crick entre l'adénine et la thymine.^[9, 212]

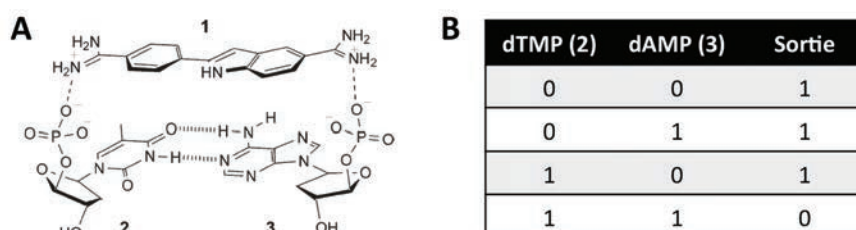


Figure 35 – Edifice supramoléculaire à base de DAPI (A) fonctionnant comme une porte logique NON-ET (B).

L'intégration dans des systèmes de transfert d'énergie non covalents comme celui présenté Figure 36A permet de relier deux porphyrines de zinc jouant le rôle de photorécepteurs à une

porphyrine non métallée réceptrice. Dans ce cas précis, lorsque l'excitation conduit à un état triplet, le transfert d'énergie s'opère au travers des liaisons hydrogènes formées par le couple guanine-cytosine^[8, 213]. Plusieurs édifices équivalents ont été mis au point en modifiant l'accepteur d'électrons en utilisant un fullerène C₆₀ par exemple (Figure 36B).^[214]

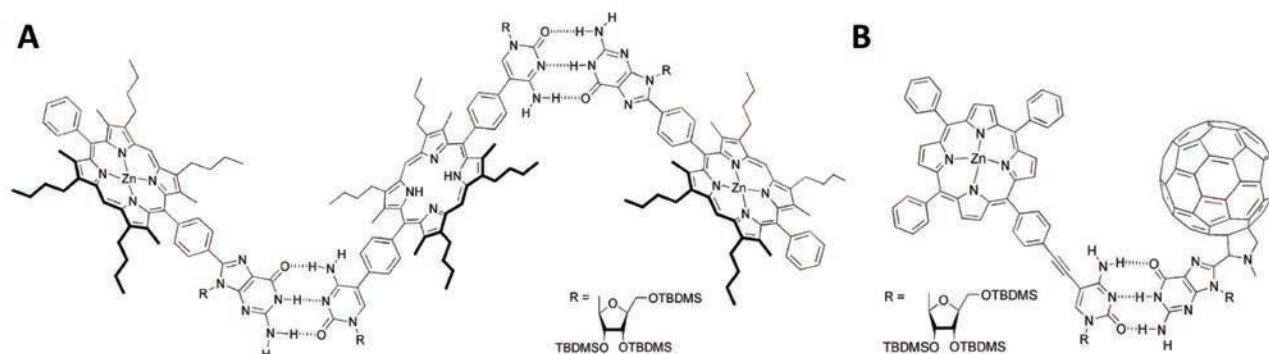


Figure 36 – Systèmes de transfert d'énergie non covalents basés sur des interactions guanine-cytosine dont l'accepteur est soit une porphyrine libre (**A**) soit un fullerène (**B**).

La synthèse d'une sapphyrine fonctionnalisée par un bras cytosine permet la reconnaissance et le transport sélectif du nucléotide dGMP grâce aux interactions entre le phosphate et un pyrrole protoné de la sapphyrine, ainsi que par l'appariement des nucléobases complémentaires (Figure 37A). Le complexe qui se forme spécifiquement à pH neutre, traverse une membrane lipophile CH₂Cl₂ avant de relarguer son nucléotide en arrivant dans le troisième milieu aqueux basique.^[215] Cet élégant concept a par la suite été amélioré en remplaçant le *template* par un calixpyrrole fonctionnalisé en position β .^[216]

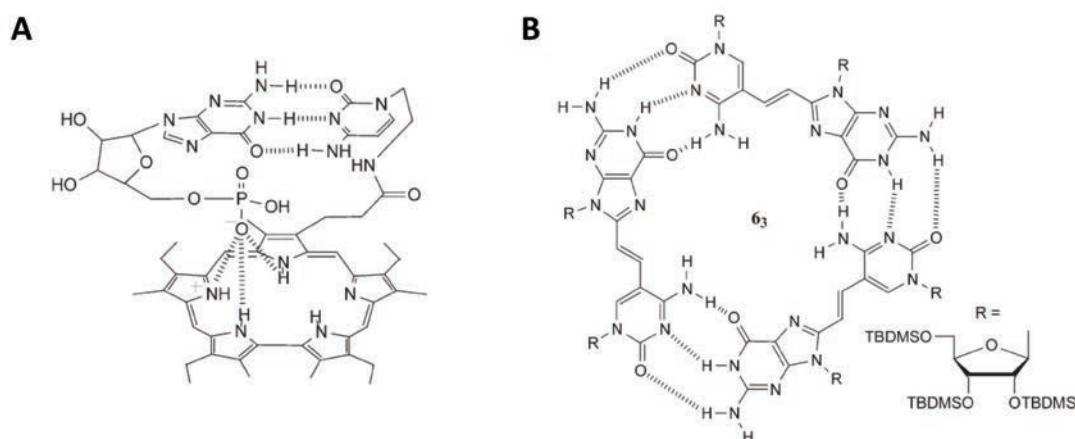


Figure 37 – Auto-assemblage de trimères GC dans un même plan (**A**) et système sapphyrinique permettant le transport sélectif des dGMP (**B**).

La fonctionnalisation d'un alcène en *trans* par une guanine et une cytosine conduit à un arrangement non classique de trimères dans un même plan, s'auto-stabilisant par un réseau de neuf liaisons hydrogènes entre trois paires G-C (Figure 37B).^[8, 9, 217]

La chimie organométallique est également représentée dans ces assemblages supramoléculaires. Un article de C. White *et al.* fait état de l'emploi d'une adénine fixée sur un motif 2,2'-bipyridine complexant un ruthénium (II).^[218] Deux autres ligands 2,2'-bipyridine (Figure 38B) ou 5,5'-di-*tert*-butyl-2,2'-bipyridine (Figure 38A) bidentates complètent la sphère de coordination du métal hexavalent.

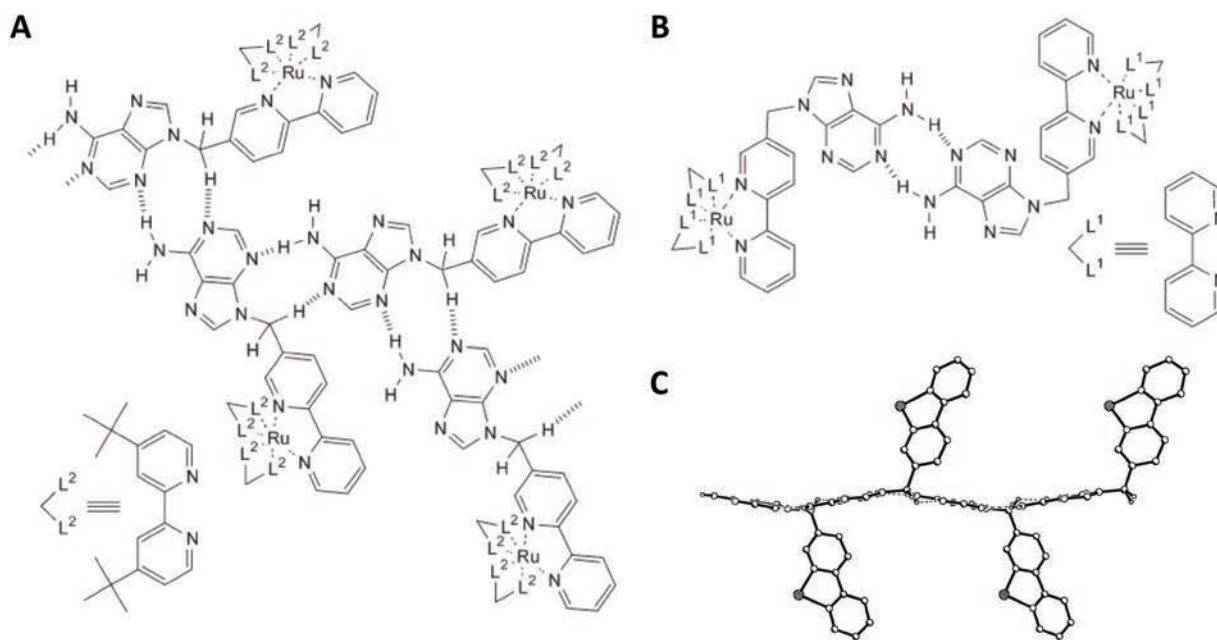


Figure 38 – Représentation des différents arrangements de monomères d'adénine en fonction des substituants présents sur le ruthénium (**A**, **B**) et structure cristallographique du complexe A (**C**, les groupements bipyridines ont été omis pour plus de clarté).

Dans les deux cas les cristaux obtenus et analysés par diffraction des rayons X montrent des arrangements très différents : la présence de bipyridines simples entraîne la formation de dimères résultant d'interactions non-usuelles entre deux adénines *via* leurs faces Watson-Crick, alors que des bipyridines encombrées par des *tert*-butyles conduisent au développement d'un réseau de liaisons hydrogènes assuré par les adénines et les protons méthyléniques présents entre la nucléobase et l'hétérocycle. La structure obtenue par diffraction des rayons X (Figure 38C) révèle que les adénines forment un plan auquel les bipyridines adjacentes sont orthogonales.^[9, 218]

Le platine peut quant à lui se lier directement à trois 9-méthylguanines par l'atome d'azote en position 7 et à un NH₃ pour compléter sa valence. Un dimère peut alors se former (Figure 39A) en disposant avec un léger décalage quatre guanines dans un même plan de façon à placer en face d'un site accepteur (*i.e.* les oxygènes des carbonyles) un site donneur (*i.e.* protons des amines ou amides). Après l'ajout d'un excès de 1-méthylcytosine dans une solution à pH 5,7 suivi d'une cristallisation, les analyses permettent de résoudre la structure présentée Figure 39B.^[219]

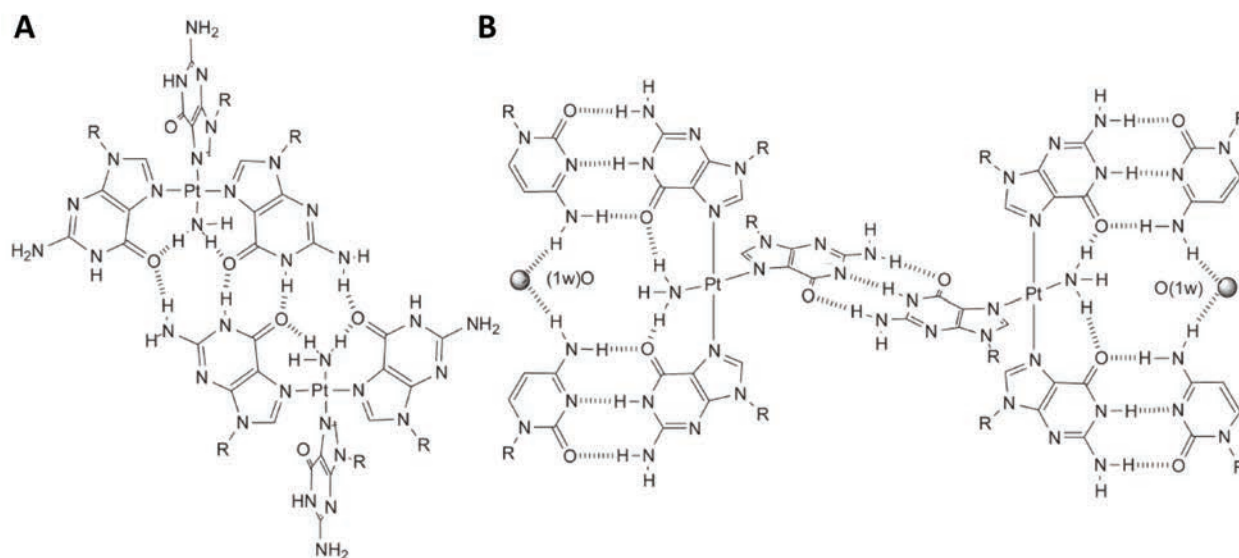


Figure 39 – Assemblage de complexes de Pt-guanine en l'absence (A) ou en présence de 1-méthylcytosine (B).

Cette dernière est constituée de deux sous-unités monomériques reliées entre elles par des interactions Watson-Crick entre une guanine native et une guanine déprotonée au niveau de sa fonction amide. Remarquons que cette perte de proton est particulièrement curieuse dans cette gamme de pH sachant que le pKa de cette fonction est estimé entre 7,8 et 8 lorsque la guanine est fonctionnalisée par un platine en position 7. Chacune de ces sous-unités résulte de deux associations Watson-Crick entre deux guanines et deux cytosines auxquelles s'ajoutent des interactions de type Hoogsteen entre les guanines et l'ammoniaque ainsi que la chélation d'une molécule d'eau entre les amines des deux cytosines. Notons que les deux plans (GGCC) sont parallèles entre eux et orthogonaux au plan contenant les deux guanines, formant ainsi un édifice aux dimensions de 28,4Å x 14,3Å.^[9, 219]

Un autre grand domaine dans lequel la chimie supramoléculaire s'est immiscée est celui des polymères où les nucléobases sont principalement utilisées pour ajouter de la réticulation et modifier l'organisation du matériau à l'aide d'interactions électrostatiques entre les chaînes polymérisées. Le choix d'un couple de nucléobases tel que l'adénine et la thymine pour fonctionnaliser un polymère linéaire à chacune de ses extrémités conduit à l'obtention de trois polymères distincts : A-n-A, T-n-T, A-n-T (avec n = chaîne polymère, Figure 40A). La formation du matériau peut alors se faire soit par mélange statistique des deux premiers (hétéro-assemblage, Figure 40B), soit par l'utilisation du dernier seul (homo-assemblage) qui présente toutefois l'inconvénient de pouvoir se cycliser intramoléculairement selon la longueur de la chaîne.^[8, 9, 220]

Dans ce cas, des nanofibres uniformes de 15 à 30nm de largeur sur plusieurs dizaines de µm de longueur sont obtenues (Figure 40C) et se structurent d'autant plus volontiers que les interactions de

π -stacking entre purines et entre pyrimidines ainsi que les liaisons hydrogènes entre les amides des chaînes adjacences stabilisent l'édifice.^[220]

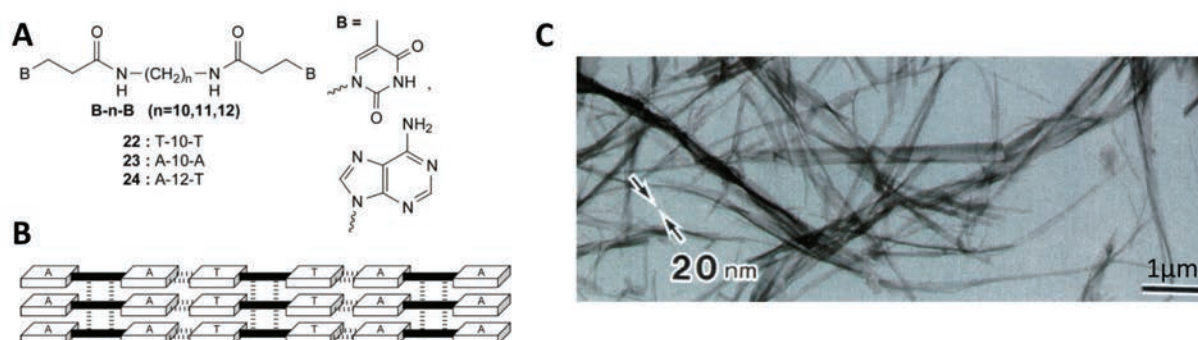


Figure 40 – Différents monomères fonctionnalisés par deux nucléobases (**A**), schéma de l'hétéro-assemblage supposé (**B**) et nanofibres formées par l'homo-assemblage A-10-T analysées par EF-TEM (**C**).

Le dernier exemple que nous aborderons ici concerne l'emploi des nucléobases lors des procédés sol-gel. Elles sont greffées sur des monomères de type orthosilicate de tétraéthyle, ou TEOS, afin d'obtenir des matériaux mésoporeux au sein desquels les chaînes siloxanes obtenues après condensation sont reliées entre elles par des appariements entre bases complémentaires ou par un réseau moins ordonné de liaisons hydrogènes (Figure 41).^[7, 221]

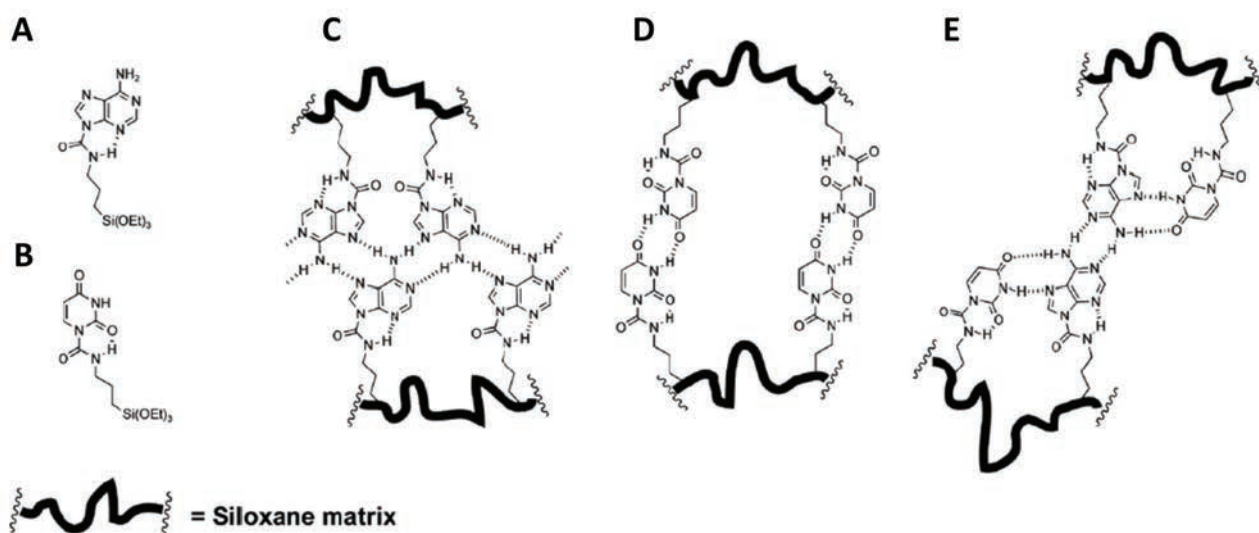


Figure 41 – Monomères adénine-siloxane (**A**) et uracile-siloxane (**B**) permettant d’obtenir les polymères homomériques (**C**, **D**) ou hétéromériques (**E**).

Ces quelques exemples proposés ici montrent la richesse des assemblages supramoléculaires bâtis sur des associations entre nucléobases. Nous allons à présent nous intéresser aux assemblages faisant intervenir uniquement des guanines qui, comme nous allons le voir, adoptent

préférentiellement un motif tétrade de guanines où quatre guanines s'auto-assemblent dans un même plan en se stabilisant par un réseau de huit liaisons hydrogènes Hoogsteen et Watson-Crick.

IV.2 Auto-assemblages de tétrades de guanines synthétiques

Dès 1910, I. Bang observe que des solutions concentrées en guanosine monophosphate forment des gels et il faudra attendre 1962 pour que M. Gellert, M. Lipsett et D. Davies reproduisent ces expériences à partir de 5'-GMP en solution dans 200mM NaCl. Après séchage, les fibres obtenues sont analysées par diffraction des rayons X et les auteurs proposent la première structure de tétrade de guanines où quatre guanines s'organisent dans un même plan et interagissent entre elles par un réseau de huit liaisons hydrogènes (Figure 42).^[35]

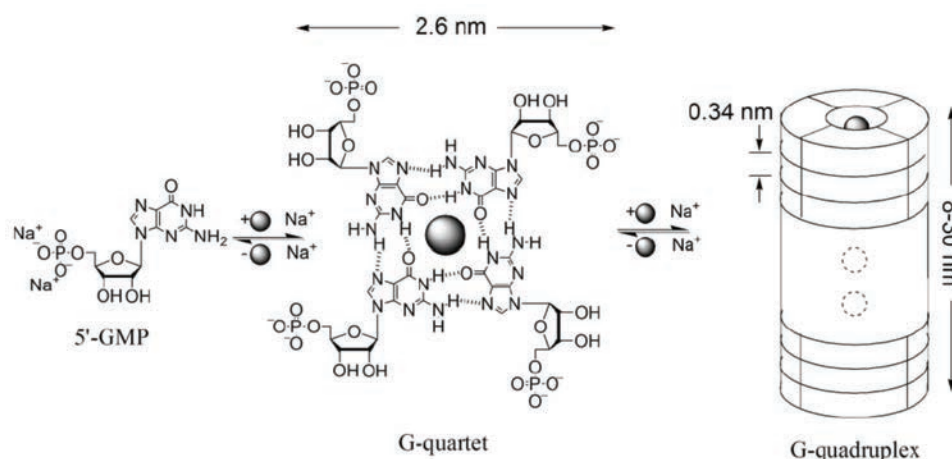


Figure 42 – Auto-assemblage de 5'-GMP en G-quartets qui s'empilent à leur tour pour former un G-quadruplexe cylindrique.

Une étude plus récente et par conséquent plus complète révèle que ces fibres peuvent atteindre de 8 à 30nm de longueur, correspondant à un nombre de tétrades empilées de 24 à 87, respectivement (Figure 42). La stabilité globale est régie par les liaisons hydrogènes entre les sites donneurs (*i.e.* NH en position 1 et NH₂ en position 2) et accepteurs (*i.e.* N en position 3, N en position 7 et O du carbonyle en position 6) des guanines, ainsi que par les interactions de π -stacking entre les tétrades et par des interactions électrostatiques entre le cation piégé entre deux tétrades et les huit carbonyles avoisinants.^[222]

Pour mieux comprendre l'organisation fine de tels édifices, plusieurs monomères ont été synthétisés avec comme caractéristique commune la suppression du groupement phosphate et la fonctionnalisation du sucre par des substituants lipophiles. L'idée était de rendre l'édifice plus hydrophobe afin de faciliter sa solubilité dans des solvants organiques apolaires tels que le

dichlorométhane ou le chloroforme dans le but d'optimiser les analyses RMN ainsi que la cristallisation (Figure 43).

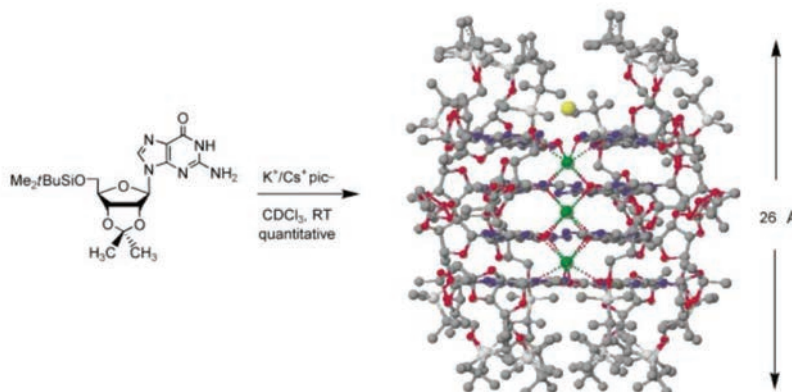


Figure 43 – Structure cristallographique issue de l'auto-assemblage de seize unités guanosines lipophiles formant un quadruplexe à quatre tétrades.

La protection des fonctions alcools en 5' et 3' du désoxyribose par des esters d'alkyles à neuf carbones conduit à la formation de structures à deux tétrades entre lesquelles l'ion potassium est séquestré (Figure 44A),^[223] alors que l'utilisation d'un *tert*-butyldiméthylsilanol en 5' et d'un isopropylidène en 2'-3' favorise l'empilement de quatre tétrades de guanines distantes de 3,3 à 3,4 Å (Figure 43).^[224]

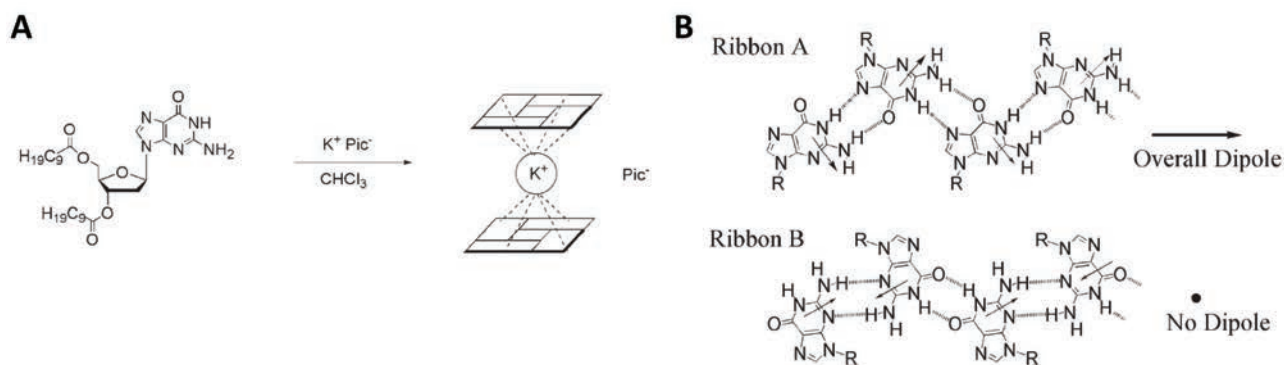


Figure 44 – Monomère guanosine lipophile formant un quadruplexe à deux tétrades (A) ou des *ribbons* (B) selon qu'il est en présence ou en l'absence d'ions K^+ , respectivement.

Il est intéressant de noter que la présence de cations soutient la formation de tétrades et par extension de quadruplexes tandis que leur absence conduit à l'association des guanines entre elles dans un même plan, en interaction par tous leurs sites donneurs et accepteurs de liaisons hydrogènes.^[225] Deux orientations sont alors possibles dont l'une induit un dipôle qui permet son utilisation en électronique moléculaire (Figure 44B).^[226]

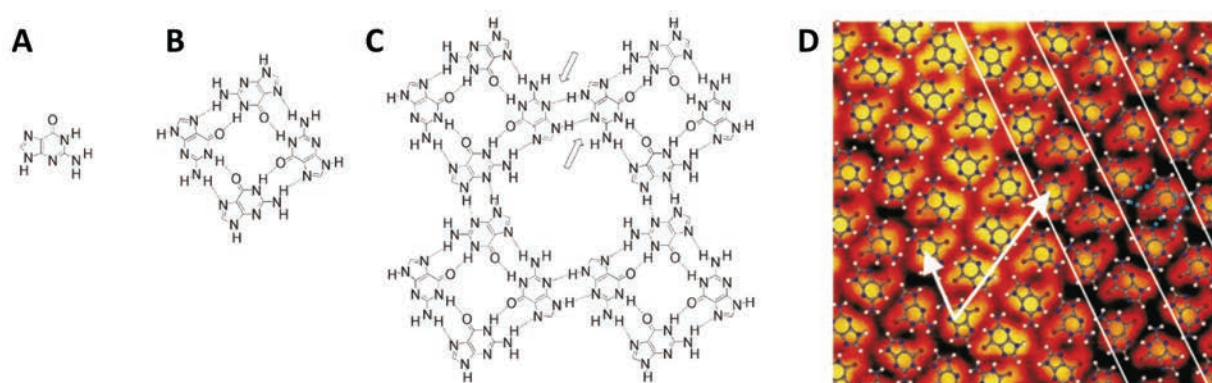


Figure 45 – Guanine (A), tétrade de guanines (B) et réseau formé par l'interaction des tétrades de guanines entre elles (C), observables sur surface d'Au(111) par STM (D).

Les guanines seules, c'est-à-dire non substituées en position 9, peuvent également s'auto-assembler comme le démontre le groupe de F. Besenbacher par microscopie à effet tunnel après dépôt sur une surface de Au(111). En plus de former des tétrades *via* les faces Watson-Crick et Hoogsteen, la libération de la troisième face permet aux tétrades d'interagir entre elles en formant un réseau de 8 guanines favorisé par la présence de 16 liaisons hydrogènes (Figure 45).^[227]

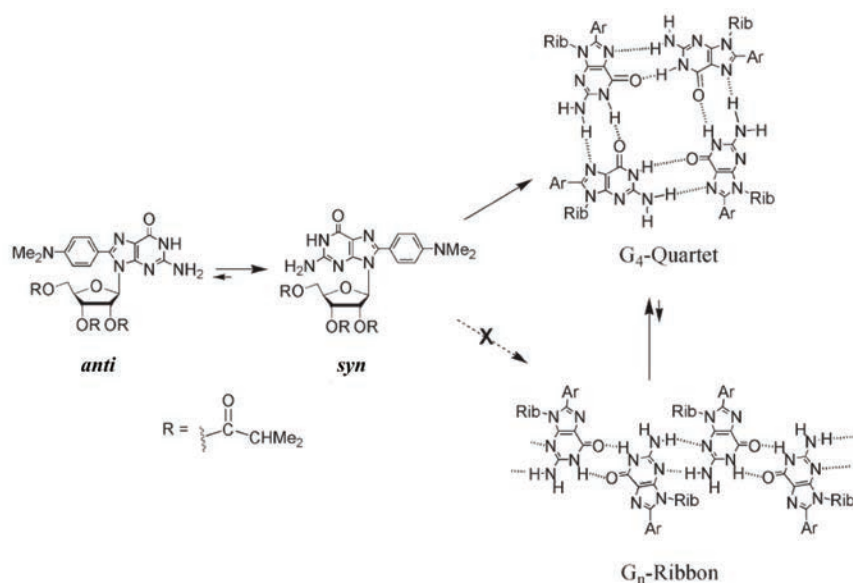


Figure 46 – La présence d'une contrainte stérique portée par la guanine empêche la formation de *ribbon*.

A l'inverse, l'ajout de groupements encombrants tels qu'un *N,N'*-diméthylaniline en position 8 de la guanine la contraint à ne s'orienter qu'en position *syn* à cause de la gêne stérique engendrée par le substituant, mais également de manière à placer l'amine primaire de la nucléobase à proximité du carbonyle du groupement protecteur du ribose (Figure 46). Les deux azotes N3 et N9 ne sont alors plus disponibles et la formation de *ribbon* devient impossible. Notons qu'ici l'auto-assemblage des guanines en tétrade peut se faire en l'absence de cations comme l'attestent les résultats des expériences de RMN ^1H .^[228]

Exactement comme avec les appariements entre nucléobases complémentaires (voir IV.1), l'utilisation de guanines pour le développement de polymères supramoléculaires offre de nombreuses possibilités. Le premier exemple proposé ici a été mis au point par l'équipe de J.-M Lehn qui s'est proposée de greffer deux bras guanines aux deux extrémités d'un polyéthylène glycol linéaire (Figure 47A). L'idée particulièrement ingénieuse provient de l'utilisation d'un cryptant [2.2.2] renfermant un ion potassium à pH neutre. Par addition d'acide dans le milieu, les amines du cryptant se protonent et ne peuvent plus chélater l'ion K^+ qui est alors libéré dans la solution où il est aussitôt capté par les guanines qui s'auto-assemblent en tétrades, créant ainsi un gel à l'état macroscopique par polymérisation supramoléculaire. Pour revenir à l'état initial, il suffit d'augmenter le pH par ajout de base pour déprotoner les amines et par conséquent déplacer le potassium vers le cryptant ce qui a pour effet de dissocier les tétrades et par conséquent le gel. Ainsi un système sol-gel dynamique dépendant du pH a été développé (Figure 47B).^[229]

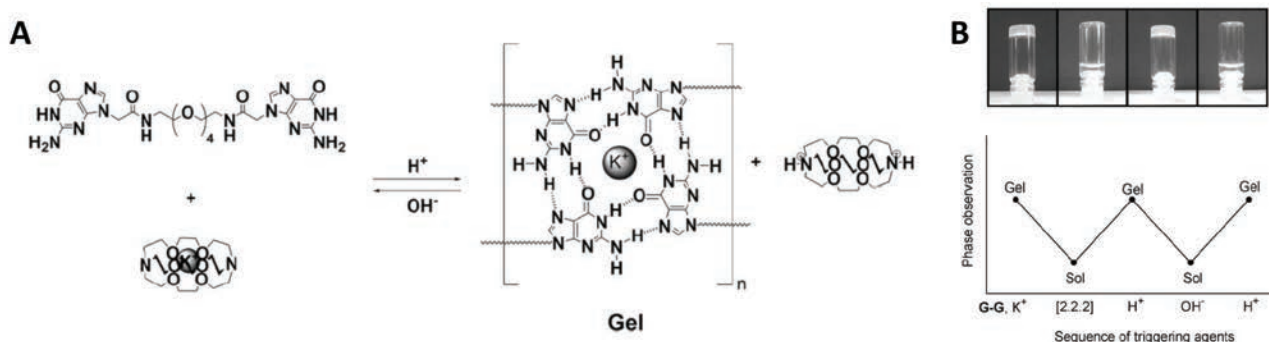


Figure 47 – L'utilisation d'un cryptant avec un monomère diguanine permet la formation réversible de polymères en fonction du pH du milieu (A). Variation de l'équilibre sol-gel (B).

Une autre application très élégante a été conçue par l'équipe de J. Davis au départ d'une guanosine dont la position 5' du ribose protégé est fonctionnalisée par un bisallyloxybenzoyl (Figure 48). Le choix des alcènes offre la possibilité de lier de façon covalente les nucléosides appartenant à une même tétrade ou bien à une tétrade adjacente par métathèse de fermeture de cycle, ou RCM. La structure résultante, élucidée par RMN, spectrométrie de masse et dichroïsme circulaire, se présente sous la forme d'un quadruplexe à quatre tétrades aux dimensions de 26Å x 30Å x 30Å. Elle peut être incorporée dans la membrane phospholipidique d'un liposome où elle joue alors le rôle de transporteur transmembranaire de sodium.^[230]

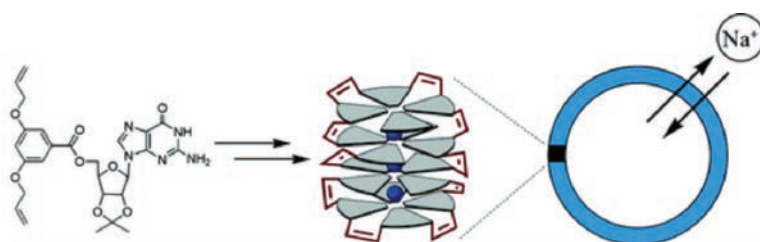


Figure 48 – Modèle de transporteur transmembranaire obtenu par métathèse.

D'autres canaux ioniques issus de l'auto-assemblage d'unités monomériques construites autour d'un cœur dilithocolamide reliant deux bras guanosines permettent de rendre compte de l'importance du *template* choisi (Figure 49). En effet, en version *bis*-urée, l'édifice tridimensionnel obtenu est actif comme transport transmembranaire dès 0,05mM alors que le dérivé *bis*-carbamate l'est au delà de 30mM. Cette différence d'efficacité est directement liée à l'auto-assemblage thermodynamiquement et cinétiquement plus favorable pour le monomère *bis*-urée.^[231, 232]

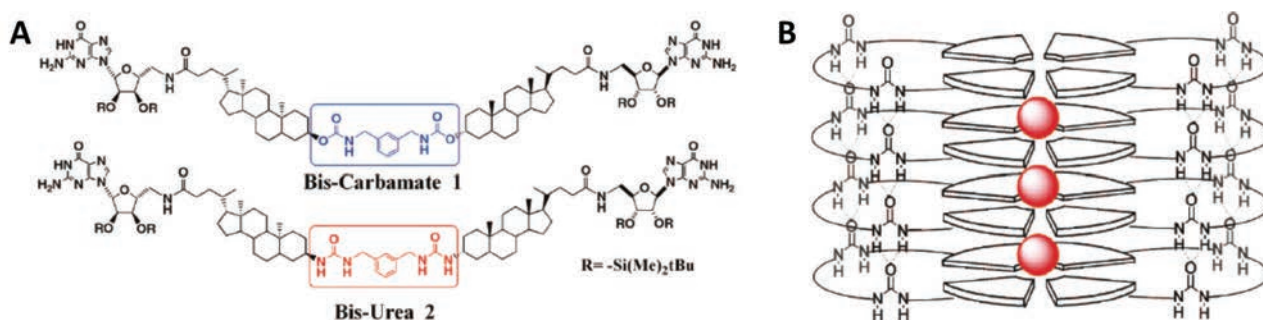


Figure 49 – Monomères diguanosines (A) utilisés pour la formation de canaux ioniques (B).

La formation d'un film polymérique macroscopique depuis le monomère diguanine présenté Figure 50 peut renfermer une architecture de type *ribbon* ou tétrade selon l'absence ou la présence de KCl, respectivement.

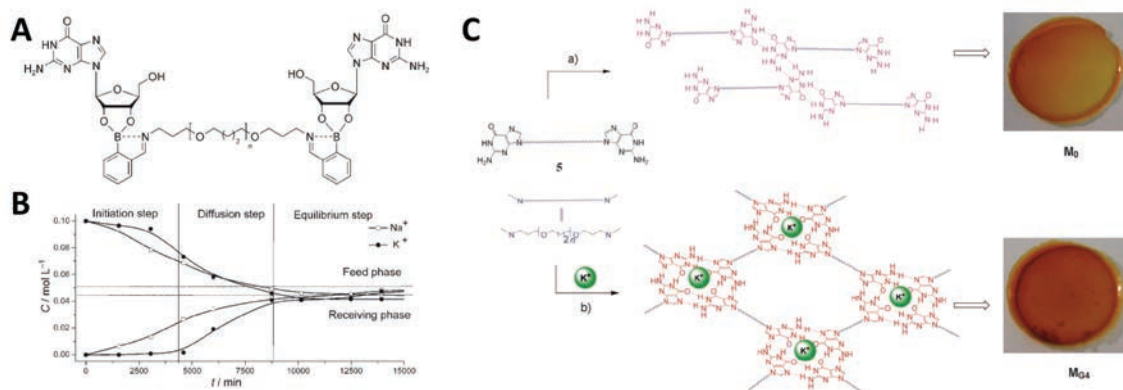


Figure 50 – Monomère diguanosine (A) se structurant en *ribbon* (C, haut) ou en tétrade (C, bas) en l'absence ou en présence de K⁺, respectivement, et évolution des concentrations en Na⁺ et K⁺ lors de l'utilisation en membrane échangeuse d'ions (B).

L'emploi de ce dernier comme membrane de transport d'ions s'avère efficace pour les échanges de K⁺ et de Na⁺, bien que la stabilisation nécessaire soit longue puisqu'elle nécessite environ 8500min, soit 6 jours.^[233]

En marge des systèmes présentés jusqu'alors, l'équipe de H. Sintim s'est intéressée à l'utilisation de la diguanosine monophosphate cyclique, ou c-di-GMP, d'origine naturelle puisqu'elle est présente chez les bactéries comme messenger secondaire.^[234] La subtilité provient ici

de l'ajout nécessaire et indispensable de certaines molécules aromatiques spécifiques telles que le thiazole orange ou la proflavine qui en stabilisant les tétrades par des interactions de π -stacking favorisent leur formation (Figure 51).^[235, 236] Son utilisation récente comme catalyseur de réactions de peroxydation de type DNAzyme montre une efficacité similaire à celle d'un quadruplexe d'ADN naturel.^[237, 238]

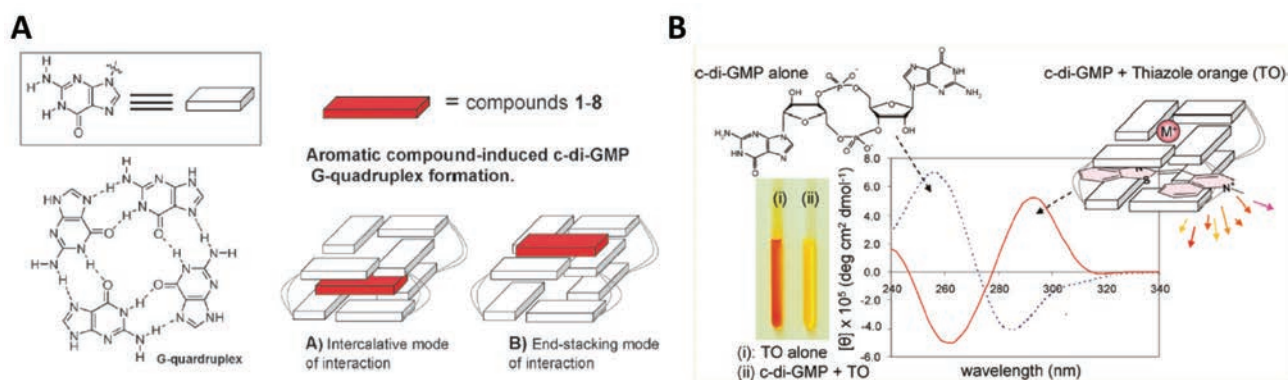


Figure 51 – Principe d'auto-assemblage des c-di-GMP en présence d'un composé aromatique (A) et application comme sonde fluorescence en présence de thiazole orange (B).

Cette petite liste d'édifices supramoléculaires régis par la formation d'une tétrade de guanines est loin d'être exhaustive et ne reflète qu'une partie de la créativité sans limite des chimistes qui les utilisent aussi bien pour le développement de systèmes énantiosélectifs contrôlés par le cation^[239] qu'en chimie combinatoire,^[240] en électronique moléculaire ou bien pour la formation de cristaux liquides.^[9] Néanmoins, pour plus de détails et d'exemples, plusieurs revues très complètes ont été publiées sur le sujet depuis une dizaine d'années.^[6, 7, 9, 181, 241]

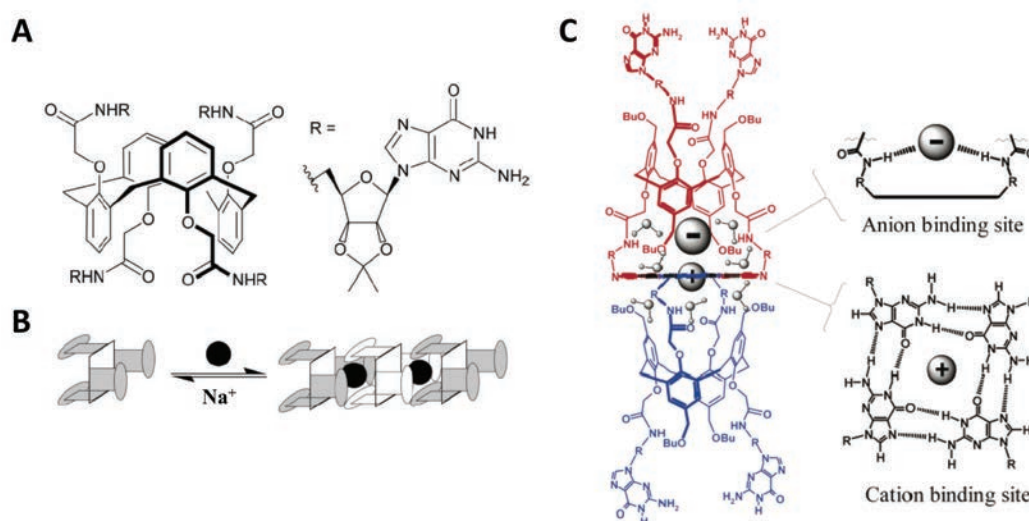


Figure 52 – Structure d'un calixarène à quatre tétrades alternées (A) et représentation schématique de l'assemblage intermoléculaire (B) utilisable pour le piégeage de paires d'ions (C).

Remarquons que jusqu'à présent, tous les monomères utilisés n'étaient fonctionnalisés que par une ou deux guanines. Seulement quelques exemples à quatre guanines sont représentés dans la littérature comme les dérivés 1,3-alternate calix[4]arenes développés par l'équipe de J. Davis en 2000 et 2003 (Figure 52A).^[242, 243] L'auto-assemblage intermoléculaire est ici privilégié par l'orientation des cœurs phénoliques et conduit soit à la formation de dimères permettant le piégeage de paires d'ions (avec le cation capté par les carbonyles de la tétrade et l'anion piégé entre le cœur calixarène et la charge positive du cation, Figure 52B, C),^[242] soit à la formation de polymères qui s'organisent en nanotubes.^[243]

A notre connaissance, il s'agit des seuls édifices à quatre guanines dont la formation de la tétrade résulte d'une association bimoléculaire. A la vue de ces systèmes, il apparaît évident que la réorientation des bras guanines du même côté du *template* pourrait conduire à la formation d'une tétrade intramoléculaire cette fois-ci. Ces composés existent et sont nommés TASQ.

V. Les tétrades de guanines intramoléculaires, ou TASQ

V.1 Etat de l'art

Le terme TASQ, pour *Template-Assembled Synthetic G-Quartet*, a été introduit pour la première fois par J. Sherman en 2008 et définit les composés capables de former de manière intramoléculaire une tétrade de guanines au départ des quatre guanines qui les composent.^[17] Ce concept naît sans doute de la convergence des travaux antérieurs de J. Davis dont nous parlions plus tôt dans ce manuscrit avec les siens. A cette époque, il s'intéressait en effet à des édifices TASP (pour *Template-Assembled Synthetic Protein*) au sein desquels quatre séquences peptidiques sont greffées sur un cœur calixarène afin d'étudier et de comprendre les interactions qui gouvernent les relations entre la structure tertiaire d'une protéine et sa séquence primaire.^[244]

Ainsi, le premier TASQ est construit autour d'un cœur calixarène fonctionnalisé par quatre bras guanines (Figure 53A) et dont la formation de la tétrade intramoléculaire est démontrée par RMN ¹H, NOESY, COSY et HMQC. Parmi les indices, l'utilisation des deux solvants [D₆]DMSO et CDCl₃ qui empêchent ou favorisent la formation de la tétrade, respectivement, permet d'observer le déplacement du proton imino en position 1 des guanines ; la présence d'un pic à 11,6ppm dans le chloroforme implique que ce proton est engagé dans une liaison hydrogène, ou plus précisément dans le réseau de liaisons hydrogènes de la tétrade (Figure 53B, C).

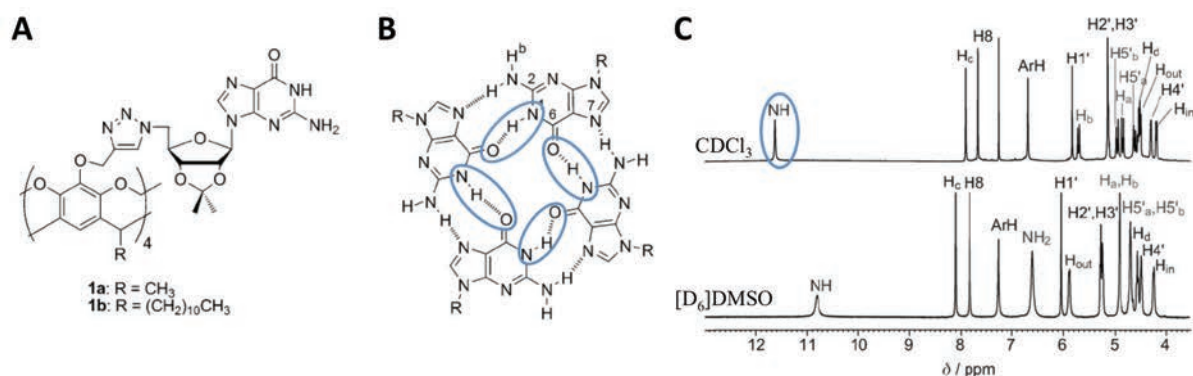


Figure 53 – Structures de la première série de TASQ (A) et mise en évidence de la formation de la tétrade intramoléculaire (B) par RMN ^1H (C).

Notons qu'ici l'auto-assemblage ne nécessite pas la présence de cations alcalins (*i.e.* Na^+ ou K^+) puisque toutes ces expériences ont été menées en l'absence de sel.^[17]

La cristallisation d'un de ces TASQ dans un mélange chloroforme et éthanol en présence cette fois de sodium permet d'obtenir la structure présentée Figure 54A et constitue une preuve formelle de la formation de la tétrade de guanines intramoléculaire. Il s'agit d'un dimère de deux TASQ formant un petit quadruplexe à deux tétrades distantes de 3,1Å et piégeant l'ion Na^+ à l'aide des huit groupements carbonyles pointant vers le centre de la cavité. Il n'existe pas de plan de symétrie entre les monomères mais uniquement un axe de symétrie C_2 parallèle aux tétrades et passant par le cation central. Ainsi, les guanines ne se recouvrent que par leur cycle à cinq chaînons en alignant le carbone électropositif en position 8 avec l'azote électronégatif en position 9 afin de minimiser l'énergie globale du système (Figure 54B).^[245]

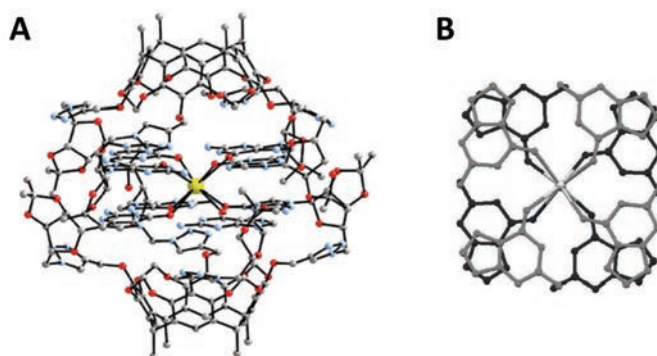


Figure 54 – Structure cristallographique de deux TASQ en face à face chélatant un ion sodium (en jaune) (A) et de la superposition des deux tétrades (B).

Entre temps, cette même équipe étudia l'effet de différents cations sur ces TASQ et conclut que les cations Na^+ et Sr^{2+} favorisaient la forme monomérique alors que des ions plus gros comme le Cs^{2+} favorisaient la forme dimérique et que l'utilisation de K^+ entraînait un équilibre entre ces deux formes. Bien que les résultats de ces travaux soient en contradiction avec la structure cristallographique obtenue trois ans plus tard, gardons à l'esprit que les conditions expérimentales

ne sont pas les mêmes puisque dans ce cas-ci, les analyses par RMN et par dichroïsme circulaire ont été réalisées en solution et non à l'état solide.^[246]

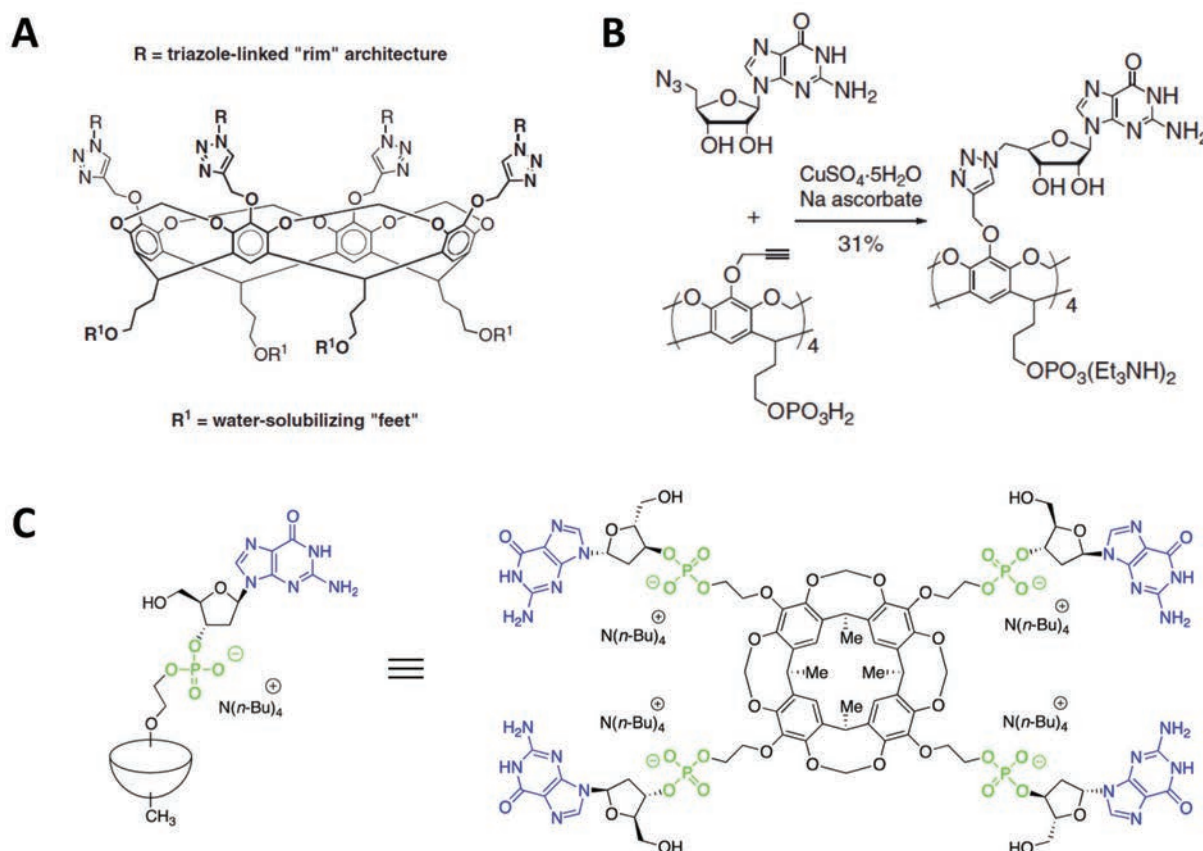


Figure 55 – Stratégie envisagée (A) par le groupe de J. Sherman pour la synthèse du premier TASQ hydrosoluble construit autour d'un cœur calixarène (B). Seconde molécule publiée en 2013 et incorporant quatre bras guanosines 3'-monophosphates (C).

Jusqu'alors, tous ces composés calixarènes avaient l'avantage d'être lipophiles et de pouvoir être étudiés plus facilement, surtout en RMN dans différents solvants organiques. Néanmoins le champ d'application de tels édifices est limité et explique pourquoi en 2011, un dérivé hydrosoluble a été synthétisé en substituant les chaînes alkyles des ponts méthylènes par des groupements phosphates ainsi qu'en laissant les fonctions alcools du ribose libres (Figure 55A et B).^[247] Seulement, cette molécule devient alors beaucoup plus complexe à étudier dans l'eau ce qui explique pourquoi aucune preuve de la formation intramoléculaire de la tétrade n'a été publiée. Ce pas sera franchi en juillet 2013 où cette même équipe reporte la mise au point d'un nouveau TASQ hydrosoluble au sein duquel les quatre bras guanines sont des guanosines 3'-monophosphates (Figure 55C).^[248] Toujours construite autour d'un calixarène, cette molécule hydrosoluble permet l'auto-assemblage des guanines sous forme d'une tétrade visible par des expériences de RMN ^1H dans l'eau ainsi que par dichroïsme circulaire. Les auteurs l'utilisent ainsi comme support pour l'étude de ligands de quadruplexes par dichroïsme circulaire et par fluorescence ; au vu des résultats

présentés, cette application nous semble toutefois limitée, et est d'ailleurs qualifiée d'« anticipée » en conclusion de l'article.

Il semble donc que le choix du *template* soit primordial pour la mise au point d'un TASQ « opérationnel » et que la balance entre l'applicabilité liée à l'hydrosolubilité et l'analyse de la formation de la tétrade facilitée dans le cas d'une molécule hydrophobe soit difficile à équilibrer.

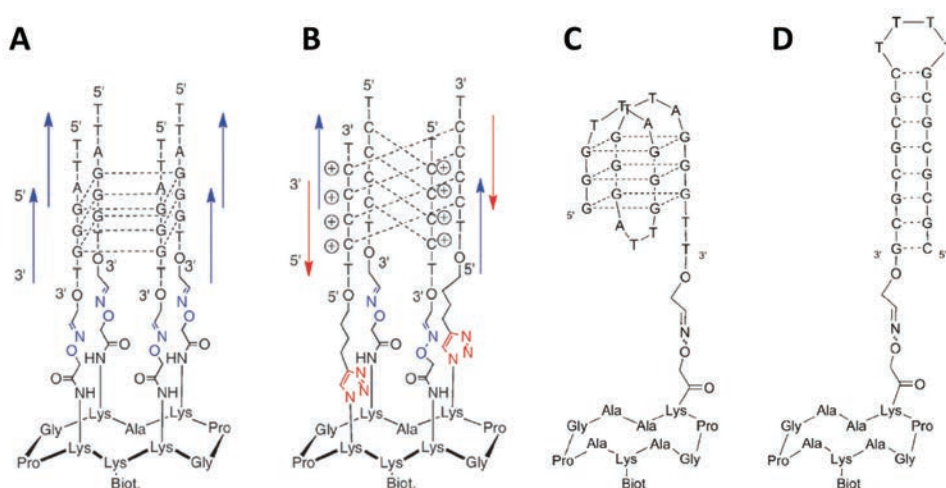


Figure 56 – Différents modèles de RAFT fonctionnalisés par un quadruplexe tétramérique (A), un *i-motif* (B), un quadruplexe intramoléculaire (C) et un duplexe *hairpin* (D).

L'équipe de E. Defrancq s'y est justement risquée. Au départ, elle s'intéressait à l'adaptation des TASP pour l'ADN, c'est-à-dire à greffer des oligonucléotides sur un *template*. Leur choix s'est porté sur un cyclodécapeptide nommé RAFT (pour *regioselectively addressable functionalized template*), fonctionnalisé au départ par quatre brins d($5'T_2AG_3T^{3'}$) pour former un quadruplexe tétramoléculaire parallèle à 3 tétrades,^[249] puis par un brin d($5'G_3(T_2AG_3)_3T_2^{3'}$) de séquence télomérique humaine pour former un quadruplexe intramoléculaire, ainsi que par un duplexe *hairpin* (Figure 56A, C, D).^[250] Plus récemment, une nouvelle stratégie de synthèse a permis l'addition de deux brins parallèles (5'-3') et deux antiparallèles (3'-5') pour former un *i-motif* (Figure 56B),^[251] et est également à l'étude pour la formation de quadruplexes antiparallèles.

Plus qu'une preuve de concept, ces édifices supramoléculaires présentent l'avantage de pouvoir être greffés à la surface d'une puce pour SPR à l'aide de la biotine fixée sur le RAFT. Ces puces fonctionnalisées permettent alors d'effectuer des criblages de ligands potentiels de quadruplexes d'ADN pour en définir leur affinité et leur sélectivité *via* l'obtention de données cinétiques et thermodynamiques.^[250]

Cependant, sommes-nous dans ce cas précis en présence de TASQ ? La définition veut que la formation de la tétrade soit dirigée par le *template* (terme « *template-assembled* »), ce qui à notre

sens renvoie à l'idée que sans lui, la tétrade ne peut se former. Or ici, toutes les séquences d'oligonucléotides utilisées sont connues pour leur capacité à se structurer seules en solution et écartent donc ces structures de cette catégorie.

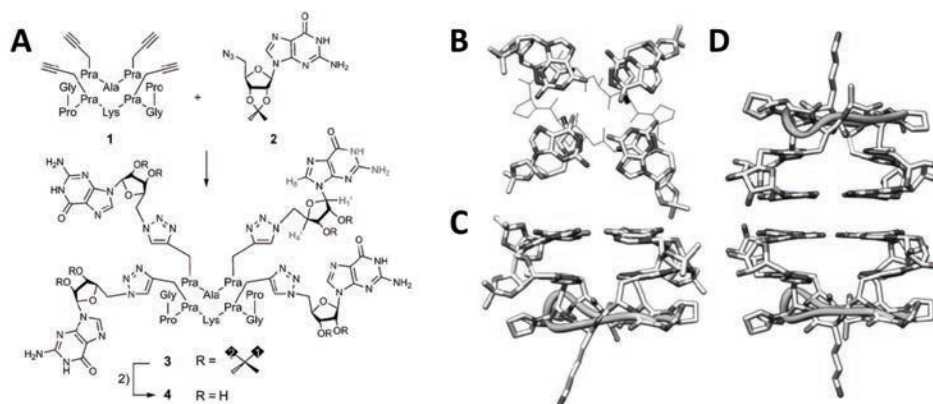


Figure 57 - TASQ hydrosoluble construit sur un RAFT (A) et structure élucidée par RMN ^1H (vue du dessus (B), de coté (C) avant déprotection (composé 3) et dimère envisagé pour le composé 4 (D)).

En revanche, un autre modèle à une tétrade de guanines également construit sur un cœur RAFT est, sans aucune ambiguïté, un TASQ (Figure 57A). Publiée en 2011, son étude complète par RMN et par dichroïsme circulaire tend à prouver que la formation de la tétrade est favorisée par des interactions de π -stacking supplémentaires entre les guanines et les triazoles qui se trouvent à proximité. Il est proposé que la molécule aurait une tendance à dimériser pour former des interactions entre deux tétrades (Figure 57B, C, D), exactement comme dans la structure cristallographique des TASQ-calixarènes de J. Sherman *et al.*^[22]

Nous reparlerons de ce composé un peu plus loin dans cette thèse (voir XIV.3.B) puisque nous avons eu la chance de pouvoir l'étudier dans des réactions de peroxydation grâce à une collaboration particulièrement stimulante avec l'équipe du Pr. E. Defrancq de l'Université Joseph Fourier de Grenoble.

V.2 Les TASQ développés et étudiés au cours de ces travaux de thèse

La synthèse de TASQ hydrosolubles commence au laboratoire en 2010 et a pour objectif initial de développer de nouveaux ligands hydrosolubles pour le ciblage de quadruplexes d'ADN dans un contexte général de recherche de nouveaux agents antitumoraux au mode d'action innovant et toujours plus sélectif.

A ce moment, la preuve de l'auto-assemblage de monomères à une ou deux guanines sous la forme de tétrades n'était plus à faire et le développement récent de la première série de TASQ lipophile par l'équipe de J. Sherman nous confortait dans cette approche.^[17]

Dès lors, le challenge était de concevoir une nouvelle série de molécules solubles dans l'eau à pH physiologique, condition indispensable pour des applications *in vitro* et à plus long terme, *in vivo*. Notre attention s'est alors focalisée sur le choix du *template* et l'utilisation de polyazamacrocycles nous paraissait particulièrement judicieuse. Le cyclène a alors été sélectionné pour son excellente hydrosolubilité qui lui vaut d'être employé comme agent de contraste en imagerie médicale sous sa forme tétraacide acétique DOTA (dont le DOTAREM®).^[252] Cet exemple conduit au second avantage qui est la possibilité d'insérer un métal soit pour des raisons de modification de conformation du cycle, soit pour ajouter une propriété supplémentaire aux TASQ, telle la luminescence.

Bien qu'elle ne soit pas essentielle, la symétrie C_4 identique à celle d'une tétrade de guanines semblait également optimale.

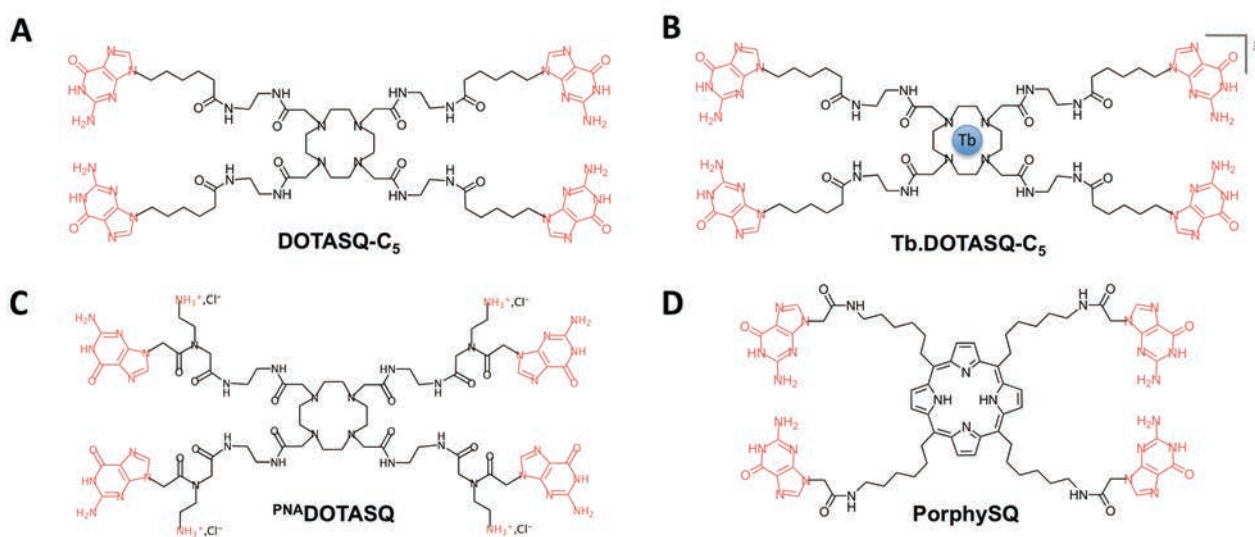


Figure 58 – Structure chimique des TASQ développés au cours de cette thèse et qui feront l'objet des chapitres suivants.

Ainsi, deux séries ont été développées au laboratoire autour d'un cœur cyclène ; la première se nomme DOTASQ et présente deux longueurs de bras alkyle à un ou cinq carbones conduisant respectivement aux DOTASQ- C_1 et - C_5 (Figure 58A). Les équivalents métallés au terbium, les Tb.DOTASQ- C_1 et Tb.DOTASQ- C_5 (Figure 58B), ont permis une première amélioration.^[20]

Un pas supplémentaire est franchi par la substitution des bras alkyles-guanines par des dérivés de PNA-guanines (PNA pour *Peptide Nucleic Acid*) pour conduire au ^{PNA}DOTASQ (Figure 58C) qui est à ce jour le meilleur TASQ hydrosoluble existant pour les applications en tant que ligand de quadruplexes ainsi qu'en tant que catalyseur de réactions de peroxydation.^[18, 19]

En parallèle des polyazacycloalcanes, le PorphySQ (Figure 58) formé sur un cœur porphyrinique a été mis au point. L'idée sous-jacente était d'avoir une communication entre le *template* et la tétrade de guanines ; en effet, la porphyrine étant une molécule aromatique, elle peut former des interactions de π -stacking avec la tétrade de guanines qui en la stabilisant, favorise sa formation.^[21]

Les détails plus précis de ces structures ainsi que les raisons successives de leurs optimisations seront apportés au fil de ce manuscrit, principalement au cours du chapitre 2.

Intéressons nous à présent aux deux domaines dans lesquels ces TASQ ont été appliqués, en commençant par le ciblage de quadruplexes d'ADN (chapitre 2) pour finir par leur rôle comme catalyseurs de réactions de type DNAzyme (chapitre 3).

CHAPITRE 2 – LES TASQ HYDROSOLUBLES COMME LIGANDS DE QUADRUPLEXES

Au long de ce chapitre sera présenté la première application des TASQ hydrosolubles développés au laboratoire comme ligands sélectifs de quadruplexes. Pour comprendre l'intérêt de cibler de telles structures, nous aborderons tout d'abord leur présence et leurs fonctions biologiques aussi bien au niveau de l'ADN que de l'ARN. Puis, en expliquant l'intérêt du ciblage de ces édifices, un bref résumé des ligands de la littérature permettra de poser les bases des stratégies d'interaction et nous conduira à la présentation de notre approche biomimétique à l'origine du concept des TASQ. Les résultats obtenus sur les trois composés DOTASQ, PorphySQ et ^{PNA}DOTASQ seront discutés dans cet ordre, correspondant au cheminement logique de l'amélioration de ces composés.

Enfin, après avoir résumé ce domaine d'applications, nous proposerons de multiples perspectives démontrant la richesse de notre approche bioinspirée.

VI. Contexte biologique

VI.1 Les quadruplexes présents sur l'ADN

VI.1.A Les régions télomériques de l'ADN

VI.1.A.a Approche historique

Les travaux de L. Hayflick au cours des années 1960 ont montré que les cellules somatiques saines n'étaient capables de se diviser qu'un certain nombre de fois (environ 50) ;^[45] passée cette limite, appelée aujourd'hui « limite de Hayflick », elles cessent leur division et entrent dans un processus d'apoptose conduisant à la mort cellulaire.^[12]

Ce phénomène, décrypté parallèlement par J. Watson^[253] et A. Olovnikov^[254-256] en 1971, s'explique par l'érosion chromosomique naturelle des parties terminales des chromosomes : les télomères. En effet, au cours de chaque cycle cellulaire, le génome est rigoureusement répliqué en

phase S, sauf au niveau des régions télomériques (voir II.2.A) qui subissent un raccourcissement progressif, nommé « problème de fin de réplication », jusqu'à atteindre une taille critique où la cellule arrête de se développer. Les télomères fonctionnent donc comme une horloge mitotique.^[12]

Alors en stage postdoctoral à l'Université de Yale dans l'équipe de J. Gale, E. Blackburn met en évidence pour la première fois en 1978 chez le protozoaire cilié *Tetrahymena* que l'extrémité 3' du télomère est simple brin, riche en guanine, et formé par la répétition du motif hexanucléotidique d(TTGGGG)_n.^[257, 258] Quelques années plus tard, en 1987, ses recherches sur plusieurs microorganismes lui permirent d'observer que ces parties simples brins renfermaient des guanines en conformation *syn* et avaient la capacité de former une nouvelle structure intramoléculaire, qu'elle attribua à un motif *hairpin* contenant des paires de bases guanine-guanine.^[259] Enfin, en 1989, W. Sundquist démontre que la structure réellement adoptée est un quadruplexe d'ADN antiparallèle.^[260]

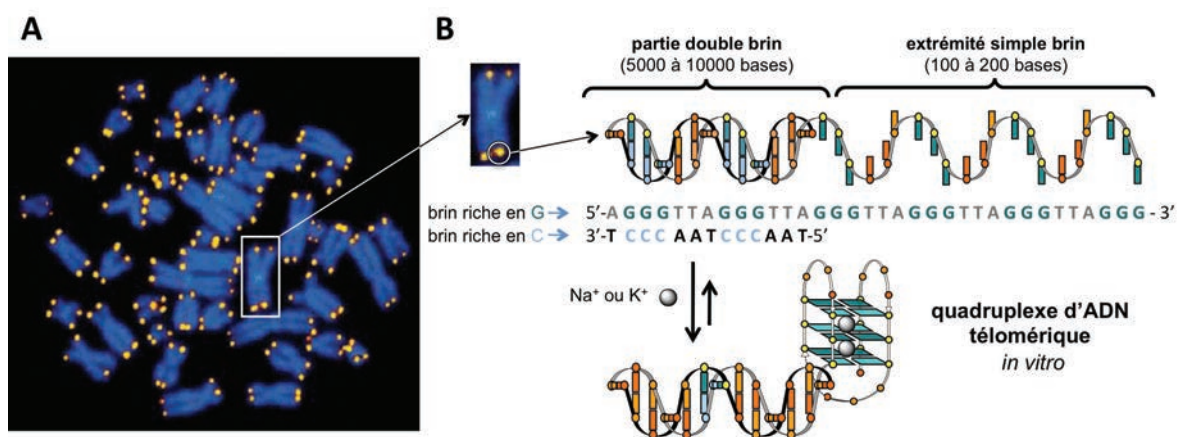


Figure 59 – Mise en évidence des télomères (en jaune) sur les chromosomes humains (en bleu) par fluorescence (A). Représentation schématique de l'extrémité télomérique humaine avec la partie simple brin riche en guanines en équilibre avec la forme quadruplexe (B).

Entre temps, la séquence télomérique humaine d(TTAGGG)_n est élucidée par R. Moyzis en 1988^[261] et s'avère être commune à tous les vertébrés. Différant chez les autres organismes, leurs séquences présentent toutefois la caractéristique commune d'être riches en guanines.^[72, 262]

Les télomères humains (Figure 59A) sont donc constitués d'une longue partie double brin, de 5000 à 10000 paires de bases issues de l'appariement du brin riche en guanines (ou brin G) en 3' avec le brin C de séquence complémentaire d(AATCCC)_n en 5', se terminant par le simple brin d(TTAGGG)_n en 3' (appelé *G-overhang*) sur 100 à 200 bases en moyenne (Figure 59B, haut).^[12, 13, 195] La première preuve de sa structuration *in vitro* en quadruplexe d'ADN est obtenue par Y. Wang et D. Patel par analyse RMN en 1993^[263] et a très largement été confirmée depuis (Figure 59B, bas).^[189, 195]

Néanmoins, il faudra encore attendre le début du XXI^{ème} siècle pour voir apparaître les premières preuves *in vivo* de leur existence au sein d'organismes eucaryotes unicellulaires, les *Stylonychia*, où la présence de protéines régulant la formation (TEBP α et TEBP β) et la déstructuration (TEBP β phosphorylé) des quadruplexes télomériques a également été mise en évidence.^[264-266]

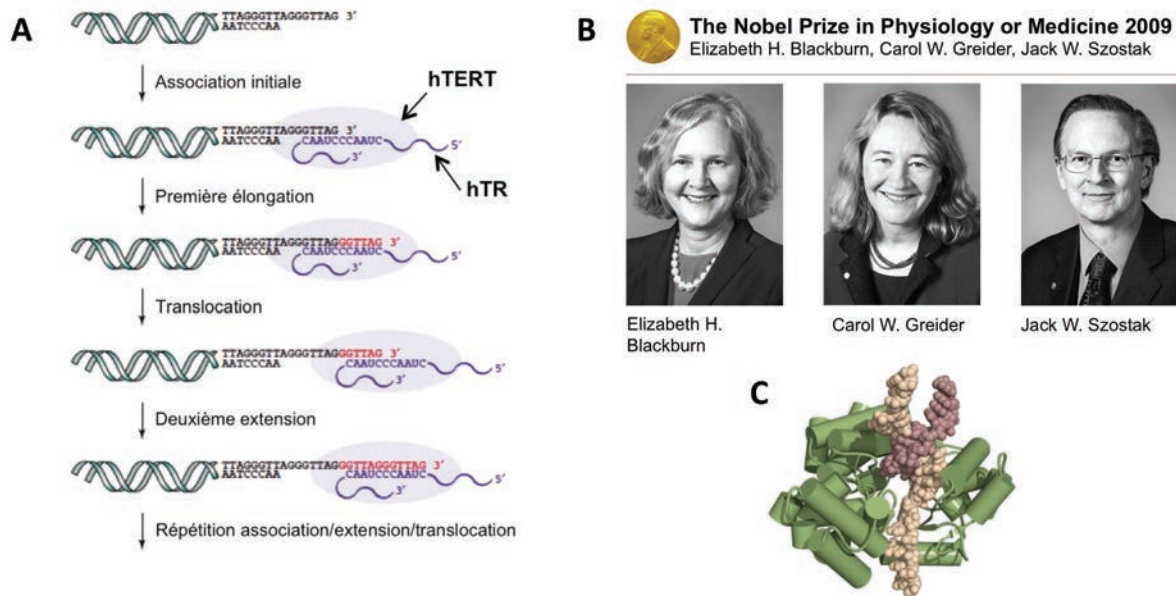


Figure 60 – Représentation schématique du mécanisme d'action de la télomérase (A), structure cristallographique de la télomérase en interaction avec un court brin d'ADN télomérique (en beige, C) et lauréats du prix Nobel de physiologie ou de médecine de 2009 (B).

L'engouement des scientifiques et plus particulièrement des chimistes pour ces structures naît au cours des années 1990 et est étroitement lié aux travaux menés en parallèle sur la télomérase. Cette enzyme, découverte en 1985 par C. Greider et E. Blackburn^[267], a une activité transcriptase inverse et permet à la cellule de contourner l'érosion chromosomique naturelle en synthétisant du brin d'ADN télomérique perdu à chaque division cellulaire.^[93, 268] Constituée d'une partie protéinique hTERT et d'un ARN contenant une séquence complémentaire à celle de l'ADN télomérique *G-overhang* lui permettant de s'y lier pour amorcer la synthèse (Figure 60A et C), la télomérase a été mise à l'honneur en 2009 en attribuant à ses découvreurs E. Blackburn,^[257] C. Greider^[269] et J. Szostak^[270] le prix Nobel de physiologie ou de médecine (Figure 60B).

Une observation particulièrement importante réalisée par l'équipe de J. Shay publiée dans la revue *Science* en 1994 montre que la télomérase, inactive dans les cellules somatiques saines, est fortement réactivée dans une vaste majorité des cellules tumorales (sur 85% des lignées cellulaires testées),^[148] faisant ainsi de cette enzyme un marqueur de cancer à fort potentiel thérapeutique.^[271, 272] De plus, l'article de Zahler *et al.* de 1991 conclut que la télomérase a besoin de l'ADN

télomérique *G-overhang* sous sa forme simple brin, et que sa structuration sous forme d'un quadruplexe d'ADN inhibe son action.^[273]

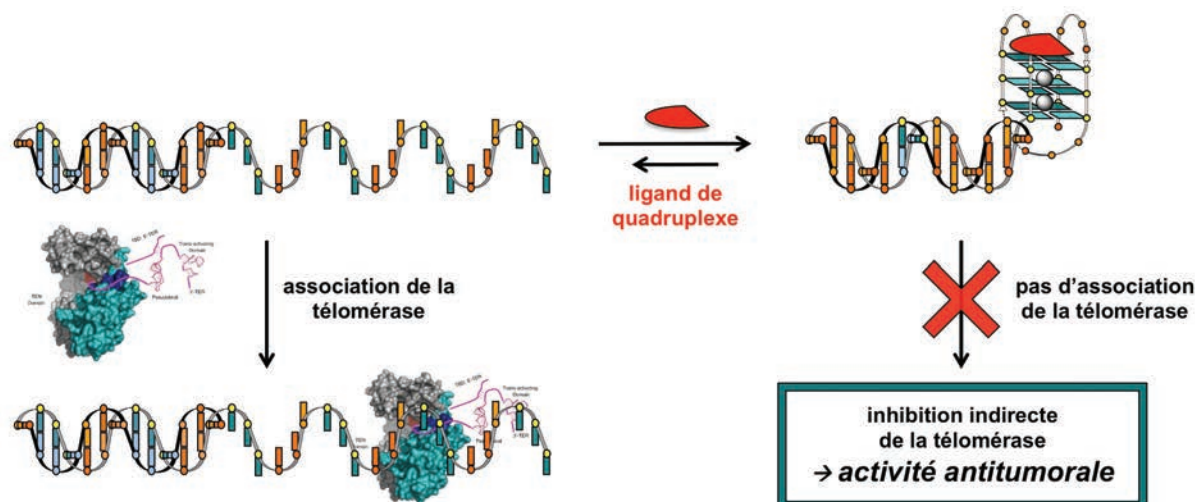


Figure 61 – Représentation schématique de la théorie initiale : inhibition indirecte de la télomérase par un ligand de quadruplexe.

Ainsi le croisement de ces deux observations permet de construire une première stratégie antitumorale basée sur l'inhibition de la télomérase, responsable de l'immortalisation des cellules cancéreuses, en cherchant à déplacer la forme simple brin télomérique linéaire vers sa forme quadruplexe (Figure 61). Le premier exemple de cette approche a été reporté par Sun *et al.* en 1997 avec l'utilisation *in vitro* d'une molécule stabilisant le quadruplexe d'ADN, un dérivé d'anthraquinone, et confirme la validité de l'approche.^[274] De là naîtra le développement d'une chimiothèque importante de molécules, appelées ligands, capables d'interagir sélectivement et le plus fortement possible avec les quadruplexes d'ADN.^[11, 275, 276]

VI.1.A.b La complexité des télomères

L'approche initiale particulièrement novatrice et élégante a permis d'ouvrir la voie à l'étude des télomères aussi bien du point de vue de leur structure que de leur rôle biologique et de développer de nombreuses molécules spécifiques des quadruplexes d'ADN depuis plus de 15 ans.

La recherche n'ayant cessé de s'intensifier depuis, il apparaît aujourd'hui que la vision du simple brin télomérique comme l'équilibre entre une forme linéaire reconnue par la télomérase et une forme fermée inhibant son action est très éloignée de la réalité.

En effet, la partie terminale des télomères s'accompagne d'un complexe protéinique nommé shelterin, qui protège l'extrémité 3' simple brin des chromosomes de la dégradation par des

exonucléases, mais également de la recombinaison non homologue (NHEJ pour *non-homologous end-joining*).^[12, 46, 277] Ce complexe est formé par l'assemblage de trois protéines qui interagissent directement avec l'ADN (Pot1, TRF1 et TRF2) et sont interconnectées entre elles par trois autres protéines (TIN2, TPP1 et Rap1) (Figure 62A, B et F). Alors que TRF1 et TRF2 (pour *telomeric repeat-binding factor 1* et 2) se fixent à l'ADN double brin, Pot1 reconnaît spécifiquement la séquence simple brin d(5'(T)TAGGGTTAG)3') à laquelle elle s'accroche (Figure 62B) et il est suggéré que cette interaction empêche la formation d'un quadruplexe.^[278]

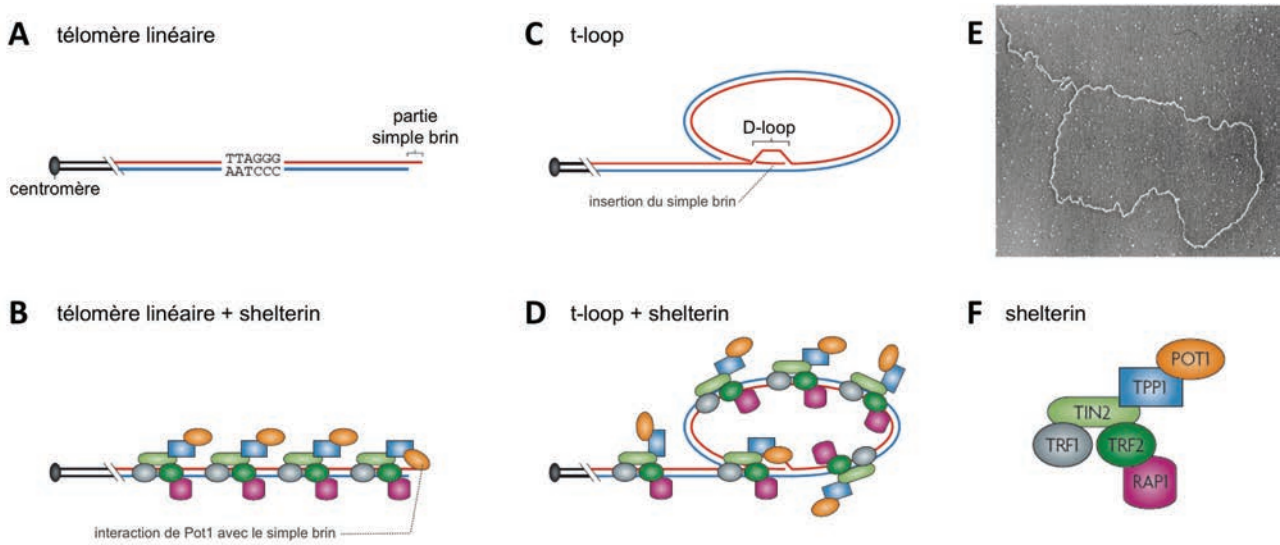


Figure 62 – Représentation du télomère humain sous sa forme linéaire ou t-loop en l'absence (A et C) ou en présence (B et D) du complexe de shelterin (F). Vue par microscopie électronique d'une t-loop formée à partir de l'ADN télomérique de cellules HeLa (E).

L'intégrité du complexe protéinique de *shelterin* est indispensable à la stabilité globale des télomères et par extension à la stabilité chromosomique (ce rôle de protection est appelé *capping*). Ainsi, l'utilisation de molécules connues pour stabiliser le quadruplexe d'ADN *in vitro* telles la télomestatine,^[279, 280] le BRACO-19^[281] ou la pyridostatine,^[282] va empêcher *in cellulo* Pot1 de se fixer à sa cible : elle se retrouve délocalisée dans le cytoplasme. Les télomères ne sont alors plus protégés convenablement (*uncapping*) et il est globalement observé une dégradation rapide des régions télomériques, rapidement létale pour la cellule.

D'autres protéines y sont également associées comme les hélicases WRN et BLM chez l'homme, utiles à l'ouverture de l'ADN duplexe et également capables de déstructurer les quadruplexes *in vitro* ; une mutation des gènes codant pour ces protéines pivots est à l'origine des syndromes de Werner et de Bloom, respectivement.^[283]

Parallèlement, il est proposé que chaque télomère se replie sur lui-même en formant une boucle, ou t-loop (Figure 62C et D), observée pour la première fois *in vitro* par microscopie

électronique en 1999 (Figure 62E) et nécessitant uniquement la présence de TRF2.^[284] Dans cette structure, la partie terminale simple brin s'insère entre les deux brins du duplexe télomérique pour interagir avec le brin C qui lui est complémentaire, formant ainsi localement une D-loop (Figure 62C).^[12, 195, 277]

Un modèle proposé par Xu *et al.* suggère que l'interaction entre les brins G conduit à la formation d'un quadruplexe qui verrouille la t-loop sous sa forme fermée.^[285]

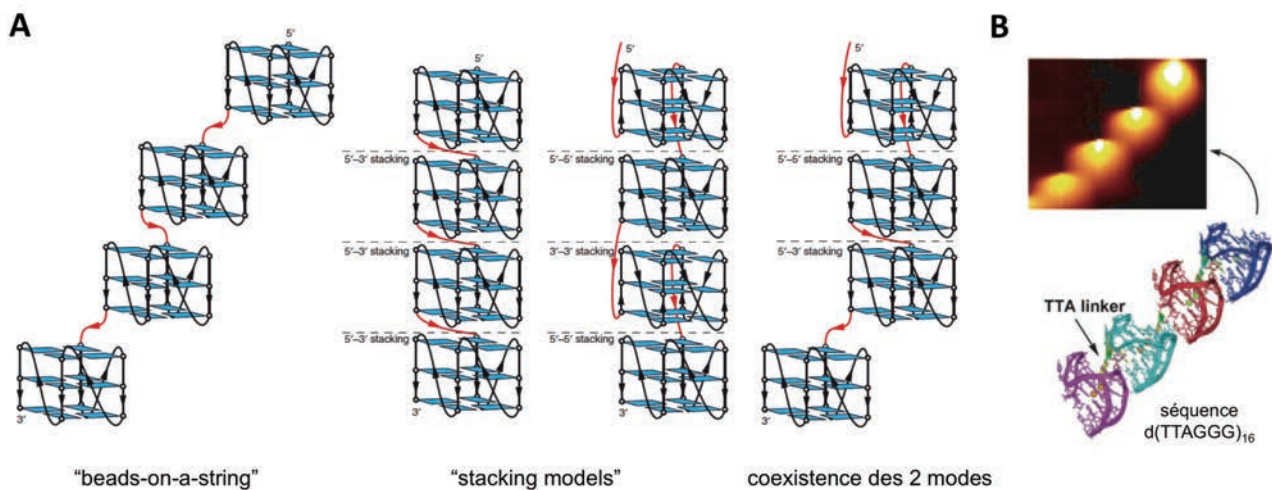


Figure 63 – Représentations des différents modes pouvant être adoptés par les multimères d'ADN (A), et image d'un mode « stacking » obtenue par AFM sur la séquence $d(T_2AG_3)_{16}$ (B).

La proposition d'une autre structure part du constat que la plupart des modèles *in vitro* développés pour l'étude de ligands de quadruplexes utilisent un quadruplexe discret d'environ 20 à 24 bases. Or, le *G-overhang*, avec une longueur de 100 à 200 bases peut donc potentiellement renfermer jusqu'à huit quadruplexes à la suite. Deux hypothèses principales sont alors possibles :^[286] soit les quadruplexes se forment de manière indépendante le long du brin sans interagir entre eux, soit au contraire ils s'auto-stabilisent par des interactions de π -stacking entre les tétrades extérieures de chaque quadruplexe ; ces modèles portent le nom de « beads-on-a-string »^[287, 288] et de « stacking model »,^[289-291] respectivement (Figure 63). La coexistence de ces deux modes au sein d'un même brin est possible et dépend entre autres de la longueur de celui-ci, comme nous le verrons plus tard dans ce chapitre (voir XV).^[23]

Pour conclure, la complexité des extrémités télomériques ne permet pas aujourd'hui d'avoir un modèle parfaitement clair et univoque, mais au vu des avancées spectaculaires de ces deux dernières décennies, il y a fort à parier que leur structure et leurs fonctions se préciseront encore davantage dans les prochaines années. Alors que la majorité des ligands de quadruplexe testés induisent bien une inhibition indirecte de l'activité de la télomérase, l'origine des effets cellulaires

observés avec ces composés s'avère beaucoup plus complexe et leur mode d'action reste encore à comprendre.^[195, 292, 293]

La bibliographie relative à la formation de quadruplexes *in vitro* laisse présager que leur existence au sein des télomères est fortement probable, d'autant plus que de nouvelles méthodes sont mises au point afin de visualiser directement et/ou indirectement leur présence *in cellulo*. Parmi elles, notons l'utilisation d'un ligand tritié permettant un suivi par autoradiographie,^[294] l'ajout d'une fonction alcyne vrai sur un ligand permettant *in situ* de réaliser une réaction de « click chemistry » avec un colorant fluorescent,^[295, 296] ou bien plus récemment l'utilisation d'anticorps reconnaissant spécifiquement les quadruplexes et détectables par la suite par addition d'un second puis d'un troisième anticorps marqué avec une sonde fluorescente.^[297, 298]

VI.1.B Les régions promotrices de gènes

En 2005, deux études bioinformatiques indépendantes menées par les équipes de S. Neidle et S. Balasubramanian révèlent qu'au sein du génome plus de 350000 séquences riches en guanines sont capables de se structurer en quadruplexes d'ADN.^[299, 300]

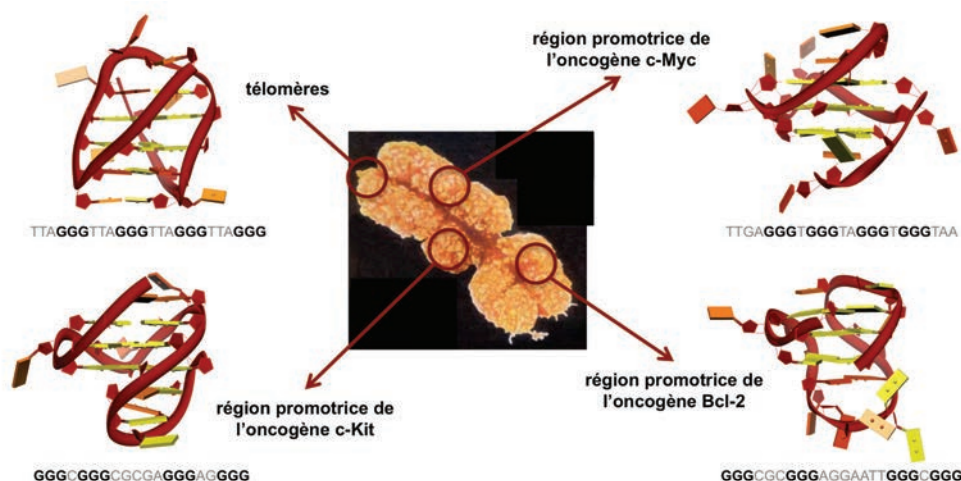


Figure 64 – Représentation de la diversité de localisation des séquences et des structures de quadruplexes au sein du génome humain.

De façon intéressante, ces séquences ne sont pas réparties de manière aléatoire mais sont au contraire enrichies au niveau des régions promotrices d'oncogènes (c'est-à-dire dont la surexpression favorise le développement des cancers, telles C-kit, C-myc, Bcl2, K-ras, voir Figure 64) et de gènes régulateurs (impliqués dans l'expression d'autres gènes *via* l'ARN issu de sa transcription ou *via* la protéine issue de sa traduction). A l'inverse, ces structures non-usuelles

seraient sous-exprimées dans les promoteurs de gènes constitutifs domestiques (ou *housekeeping*, qui s'expriment de la même manière dans toutes les cellules et ce, sans mécanisme de régulation) et de gènes suppresseurs de tumeurs (régulant la croissance cellulaire).^[10, 13, 154, 301]

Ainsi, les gènes assurant le bon fonctionnement cellulaire sont pauvres en quadruplexe alors que ceux accroissant l'instabilité chromosomique, et plus particulièrement les oncogènes, en sont riches.^[302] Cette observation laisse alors supposer un rôle de régulation contrôlé par la formation et la destruction de cette architecture.

Publié en 2000 dans la revue *Cell*, l'article de D. Hanahan et R. Weinberg (qui fait aujourd'hui autorité avec plus de 10000 citations) regroupait pour la première fois sous forme de six caractéristiques principales les différentes transformations physiologiques observées au sein des cellules tumorales.^[303, 304] En partant de ce constat, T. Brooks *et al.* remarquent qu'à chacune de ces six catégories correspond une ou plusieurs protéines codées par un gène dont le promoteur contient une séquence riche en guanines capable de former un quadruplexe d'ADN, dont les structures et les effets cellulaires sont représentés Figure 65.^[154]

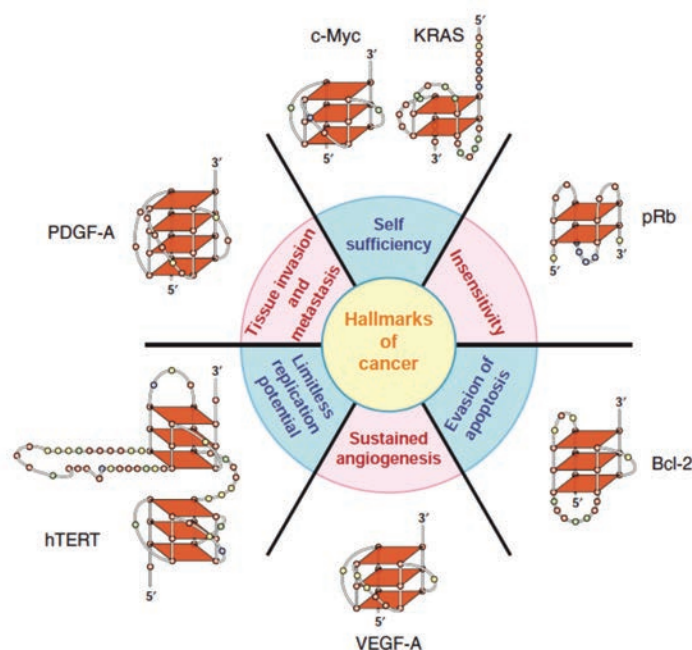


Figure 65 – Correspondance entre les six caractéristiques principales d'une cellule cancéreuse et les promoteurs d'oncogènes capables de former des quadruplexes d'ADN.

Les régions oncogéniques constituent donc, en plus des zones télomériques, une seconde cible particulièrement intéressante et prometteuse pour le développement d'une nouvelle stratégie anticancéreuse. En stabilisant ou en favorisant la formation des quadruplexes d'ADN dans ces régions, il serait alors possible de contrôler l'expression des gènes en limitant la production d'ARN et par conséquent de protéines oncogéniques.^[10, 13]

A ce jour, plusieurs travaux confirment la validité de cette approche, comme l'utilisation d'un dérivé tétra-guanidinium de phthalocyanine qui diminue la production d'ARNm issue de C-myc,^[305] ou encore la pyridostatine qui réduit la transcription du gène PVT1^[306] et qui inhibe également à près de 95% celle du gène SRC.^[296]

VI.2 Les quadruplexes présents sur l'ARN

VI.2.A Les TERRA

Pendant longtemps, les télomères ont été considérés comme des régions silencieuses du génome, c'est-à-dire non transcrites en ARN. Or, depuis 2007, de nombreux travaux démontrent clairement que l'ADN télomérique est bien transcrit en ARN, non codant, nommé TERRA pour *telomeric repeat-containing RNA* et renfermant environ 200 bases chez l'homme (Figure 66).^[200] Cette séquence r(UUAGGG)_n issue du brin C de l'ADN présente une forte similitude avec celle du brin G, avec seulement un remplacement des thymines par des uraciles, qui ne diffèrent que d'un groupement méthyle (voir III.2.A), et bien sûr la présence de la fonction hydroxyle sur le ribose en position 2' par rapport au désoxyribose.

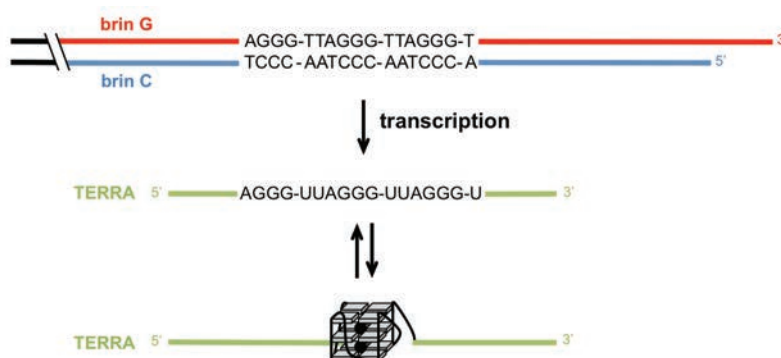


Figure 66 – Représentation schématique de la transcription des télomères en TERRA de séquence r(UUAGGG)_n lui permettant *in vitro* de se structurer en quadruplexe.

Grâce à leur richesse en guanines, les TERRA sont capables *in vitro* de se structurer sous la forme d'un quadruplexe plus stable que son homologue ADN (Figure 66). Sa conformation est parallèle pour les raisons évoquées partie III.3.B.e et ce, quel que soit l'ion alcalin présent (Na⁺ ou K⁺).^[12, 307-309]

L'utilisation de longues séquences multimériques montre que les quadruplexes d'ARN ont tendance à se stacker entre eux par deux ou trois unités maximum, formant ainsi une structure où

coexistent les modèles « stacking » et « beads-on-a-string » (Figure 67) ;^[287, 310, 311] ces résultats sont similaires à ceux que nous avons obtenus sur des multimères d'ADN étudiés par la méthode de DNazyme (voir XV.1).^[23]

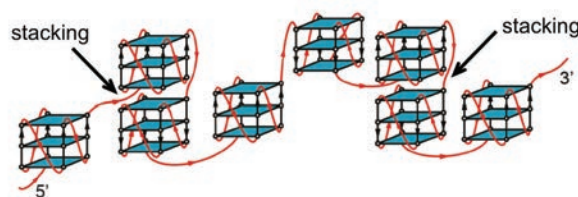


Figure 67 – Représentation schématique de la structure multimérique d'un long TERRA où coexistent les formes « beads-on-a-string » et « stacked ».

Chez l'homme, les TERRA sont synthétisés majoritairement par l'ARN polymérase II, bien que d'autres contributions soient supposées, puis subissent une maturation consistant en l'ajout d'une coiffe m7G en 5' et, dans 7% des cas, en l'ajout d'une queue poly(A) en 3'. Une partie des TERRA non polyadénylés s'associe à l'ADN, et particulièrement au niveau des télomères.^[200, 312]

Il est proposé que ces associations jouent un rôle lors de la réplication de l'ADN, *via* par exemple l'appariement des TERRA avec le brin C de l'ADN qui leur est complémentaire, ou encore leur interaction avec le brin G pour former un quadruplexe hybride ADN/ARN qui empêcherait l'action de la télomérase. Leur interaction avec les protéines TRF1 et TRF2 de la shelterin, responsables de la protection des télomères et de la stabilité chromosomique, a également été reportée.^[200, 313, 314]

Notons que, de manière intéressante, des tests *in vitro* ont montré que les TERRA inhibent de manière directe l'action de la télomérase en s'associant à sa partie protéinique hTERT et non avec la séquence hTR qui lui est pourtant complémentaire.^[201, 315, 316] En outre, il apparaît que plus les télomères sont longs, plus la concentration nucléaire en TERRA est élevée ; à l'inverse, dans le cas des cellules tumorales où la télomérase est surexprimée par rapport aux cellules saines, leur concentration y est faible.^[312, 316]

Cependant récemment, ce concept a été pour la première fois mis à mal par une étude de A. Smirnova *et al.* parue en mai 2013 qui conclut à la non corrélation entre la longueur des télomères et la concentration en TERRA.^[317]

Ceci démontre bien à quel point les connaissances scientifiques sur ces séquences d'ARN sont encore émergentes et demandent à être étendues pour comprendre précisément leurs fonctions biologiques. Néanmoins, les TERRA sont un exemple supplémentaire d'acides nucléiques capables de se structurer en quadruplexes et dont le ciblage par un ligand sélectif permettrait de contribuer davantage à la compréhension de leur rôle *in cellulo*.

VI.2.B Les régions 5'-UTR des ARN

En 2007 et 2008, deux études bioinformatiques sur le transcriptome humain ont mis en évidence que les séquences d'ARN capables de se structurer en quadruplexe sont particulièrement enrichies au niveau des régions 5'-UTR des ARNm et sont présentes sur 6% à 8% de la totalité des gènes.^[198, 318]

Les séquences 5'-UTR se situent en amont du cadre ouvert de lecture (ou *ORF* en anglais) et sont indispensables à l'initiation de la traduction de l'ARNm en protéine. Dans un fonctionnement classique (dit *cap-dependant*), ce processus commence au niveau de la coiffe (ou *cap*) m⁷G fixée à l'extrémité 5' par l'assemblage de la petite sous-unité ribosomale 40S avec plusieurs protéines spécifiques, nommées facteurs d'initiation, et un ARNt. Ce complexe va ensuite se déplacer le long de la séquence 5'-UTR, d'une longueur moyenne de 250 bases chez l'homme, jusqu'à atteindre le codon d'initiation AUG. A cet instant, les facteurs d'initiation sont libérés et la sous-unité 60S est recrutée pour former avec la sous-unité 40S le ribosome fonctionnel ; la synthèse du polypeptide codé par l'ORF peut alors commencer.^[5, 197]

Pour que cette étape d'initiation se réalise convenablement, il est indispensable que la séquence 5'-UTR soit lisse, c'est-à-dire sous sa forme linéaire simple brin, pour permettre l'assemblage et la progression du complexe. En effet, des structures secondaires de l'ARN telles que les *hairpins*^[319] ou les quadruplexes^[318, 320, 321] peuvent bloquer ce processus et empêcher *in fine* la production de protéines (Figure 68).

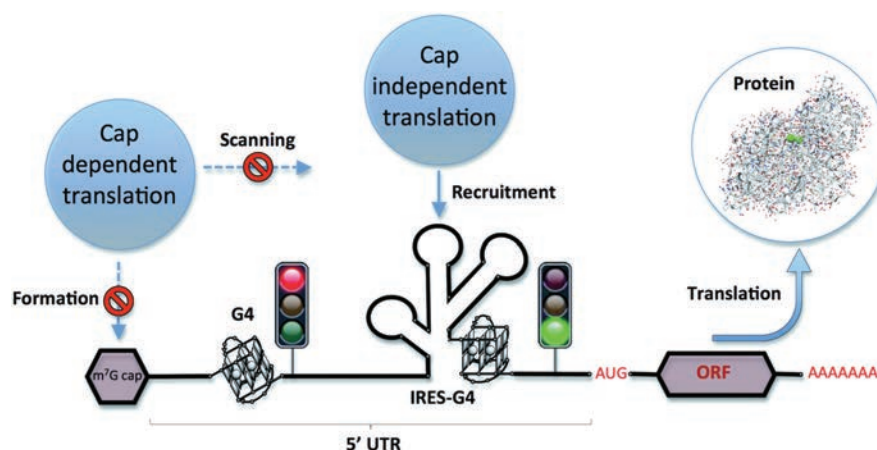


Figure 68 – Illustration des possibles rôles de la formation de quadruplexes d'ARN au niveau de la région 5'-UTR pouvant entraîner une inhibition ou une favorisation de la traduction, selon que le mécanisme est *cap-dependant* ou *cap-independant*, respectivement

Il apparaît alors que la maîtrise de l'équilibre entre la forme simple brin et sa structure secondaire, dont la forme quadruplexe, permettrait de contrôler l'expression génique. Au sein de la

cellule, ce sont des protéines de régulation telles FMRP et Pat1 (évaluées *in vitro*) qui stabilisent la forme quadruplexe alors que les hélicases antagonistes RHAU et DHX9 (évaluées *in vitro*) cherchent au contraire à lisser le brin.^[197]

Sur cette base, il semble donc intéressant dans le cas d'une cellule tumorale de bloquer ou de ralentir la production d'oncoprotéines en stabilisant le quadruplexe qui se forme sur la région 5'-UTR de son ARNm. Quelques exemples de la littérature tendent à valider cette démarche, parmi lesquels nous pouvons citer l'effet de la pyridostatine sur le proto-oncogène NRAS qui se traduit *in vitro* par une diminution de son expression d'environ 50%,^[322] ou encore le composé PhenDC₃ inhibant la traduction de TRF2 à près de 30% *in vitro*^[323] et montrant de bons résultats *in cellulo*.^[324]

En parallèle, un second mécanisme de traduction dit *cap-independant* permet à certains ARNm de recruter le ribosome directement au niveau du codon d'initiation où il se fixe à un site d'entrée nommée IRES (pour *internal ribosome-entry site*).^[197] Certains de ces sites renferment une séquence capable de se structurer en quadruplexe et, contrairement au cas précédent, sa formation favorise ici la traduction du gène (Figure 68), comme pour le facteur de croissance des fibroblastes FGF-2^[325] ou le facteur de croissance de l'endothélium vasculaire VEGF.^[326]

Ainsi, puisque la formation d'un quadruplexe dans les régions 5'-UTR peut soit favoriser, soit inhiber la production de protéines en fonction du mécanisme d'initiation de l'ARN, l'utilisation de ligands spécifiques de ces structures comme agents thérapeutiques potentiels semble actuellement complexe mais reste toutefois à explorer.^[314]

VI.3 Intérêts du ciblage des quadruplexes d'ADN et d'ARN et apologie de la recherche de ligands sélectifs

Depuis près d'une vingtaine d'années où les quadruplexes sont au cœur d'un réel engouement scientifique tant pour les biologistes que les chimistes, cette recherche multidisciplinaire n'a cessé de s'amplifier et de se complexifier, à tel point que la connaissance actuelle ne permet pas de conclure de manière univoque sur leur rôle au sein de la cellule ni d'attester formellement de l'intérêt de leur ciblage.

Néanmoins, en s'éloignant de la théorie initiale qui supposait que la stabilisation des quadruplexes par de petites molécules inhibait l'action de la télomérase, enzyme surexprimée

uniquement dans un contexte tumoral, les découvertes réalisées ont ouvert un champ d'application des ligands encore plus large.

D'un point de vue structural, les multiples études biophysiques montrent l'extraordinaire polymorphisme de ces édifices tridimensionnels qui en présentant une structure globale commune, varient en terme de nombre de tétrades, de longueur et de configuration de boucles, de conformation de bases, etc.^[10, 12, 13, 286] Quant à leurs fonctions biologiques, la répartition des nombreuses séquences aptes à former des quadruplexes est loin d'être aléatoire et tend à démontrer qu'ils jouent un rôle clef dans la régulation de l'expression génique, que ce soit au niveau transcriptionnel (ADN promoteur d'oncogènes ou de proto-oncogènes),^[10, 154] traductionnel (ARN des régions 5'-UTR)^[197] ou plus précisément dans la région télomérique, zone pivot de la stabilité chromosomique, via l'ADN télomérique *G-overhang* et l'ARN TERRA (Figure 69).^[12, 195, 293]

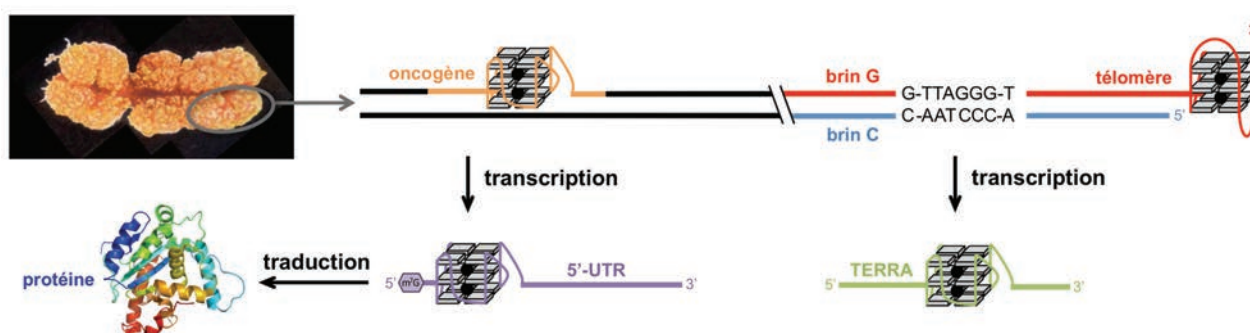


Figure 69 – Représentation schématique des multiples localisations des quadruplexes d'ADN et d'ARN.

Alors qu'il est encore trop tôt pour définir précisément les modes d'action d'un ligand une fois dans la cellule, des études *in vivo* préliminaires sur des ligands tels que le BRACO-19, la quarfloxin, RHPS4 ou encore AS1410 tendent à montrer leur efficacité sur des xénogreffes, seuls ou en combinaison avec d'autres agents anticancéreux.^[196]

Il apparaît aujourd'hui assez clairement que ces molécules, interagissant avec les quadruplexes comme attendu ou par une autre voie, créent *in cellulo* un désordre lié à une instabilité chromosomique empêchant la cellule de se diviser et par conséquent de proliférer. Au cours du cycle cellulaire, le génome est constamment inspecté lors de nombreux points de contrôles (ou *checkpoints*, voir II.1.B) au cours desquels les erreurs sont méticuleusement réparées avant de permettre à la cellule de continuer son cycle. La création d'erreurs stables sur l'ADN comme la formation d'un adduit quadruplexe/ligand impossible à résoudre contraindra la cellule à arrêter sa division puis à mourir.^[37]

Plus récemment, les quadruplexes capables de se former au sein de séquences d'ARN des régions 5'-UTR des gènes ou des TERRA sont devenus de nouvelles cibles potentielles puisque leur ciblage peut, dans le cas des ARNm, modifier leur expression, c'est-à-dire modifier la

production de protéines (dont les oncoprotéines dans le cas d'oncogènes). Seulement, selon le mécanisme d'initiation mis en jeu (*cap-dependant* ou *cap-independant*), l'effet produit peut être en faveur ou en défaveur de la synthèse peptidique, rendant ainsi cette approche complexe dans un contexte cellulaire global.^[197] Pour les TERRA, la relation entre leur concentration et leur effet sur l'inhibition de la télomérase est depuis peu remise en doute et nécessite d'être approfondie.^[200, 317]

Avec le nombre de cibles quadruplexes potentielles au sein d'une même cellule, un même ligand aura donc la possibilité de se fixer sur l'ADN ou l'ARN, engendrant différents effets. Néanmoins, il apparaît aux premiers abords qu'une interaction avec l'ARN sera privilégiée face à l'ADN et ce, pour deux raisons principales. La première repose sur le fait que l'ARN est présent dans le cytoplasme et n'impose donc au ligand de ne franchir qu'une seule barrière, la membrane cytoplasmique, alors que pour atteindre l'ADN, il lui faudra traverser le cytoplasme puis passer la membrane nucléaire. La seconde est due à l'accessibilité des oligonucléotides : en effet, l'ADN est majoritairement présent sous une forme double brin condensée autour des histones alors que l'ARN est lui sous une forme simple brin plus ou moins repliée sur elle-même pour former une structure secondaire beaucoup plus accessible. L'ARN apparaît donc intuitivement beaucoup plus accessible que l'ADN bien qu'il souffre néanmoins d'une durée de vie plus courte résultant d'une stabilité chimique moindre.^[327]

Cependant, de récentes études menées *in cellulo* avec la pyridostatine montrent au contraire que la molécule se fixe bien sur l'ADN et préférentiellement sur des régions non-téломériques lorsqu'elle est présente à faible concentration, puis au niveau des télomères à concentration plus élevée.^[296, 297]

Pour conclure, il est évident qu'à l'heure actuelle les ligands de quadruplexes ont encore du chemin à parcourir avant d'être appliqués cliniquement comme nouveaux agents thérapeutiques, mais l'utilisation actuelle de molécules anticancéreuses ciblant l'ADN duplexe comme le cis-platine, la doxorubicine ou la camptothécine valide indéniablement ce type d'approche,^[328] et conforte ce nouvel axe de recherche.

Toutefois, les ligands de quadruplexes ne doivent pas être réduits à leur seule application antitumorale potentielle. En effet, ils constituent de formidables outils moléculaires innovants permettant de décrypter *in vitro* et *in cellulo* l'impact de la stabilisation des quadruplexes, c'est-à-dire l'impact de la modification de l'équilibre entre la forme simple brin et quadruplexe d'une séquence d'ADN ou d'ARN et permettent ainsi de parfaire les connaissances scientifiques fondamentales relatives à ces structures tridimensionnelles.

Bien entendu, de tels outils moléculaires doivent présenter certaines propriétés indispensables comme une forte affinité pour les quadruplexes mais surtout une excellente sélectivité vis-à-vis des autres formes existantes, et principalement la double hélice d'ADN. De plus, la synthèse de composés hydrosolubles est primordiale et s'avère, comme nous le verrons, un point relativement complexe.

VII. Développement de nouveaux ligands biomimétiques : les TASQ

VII.1 Modes d'interaction des quadruplexes

Comme nous l'avons vu, les quadruplexes d'ADN ou d'ARN peuvent se former au départ de nombreuses séquences riches en guanines, ce qui leur confère un fort caractère polymorphique. Néanmoins, ce type d'édifice possède quelle que soit sa structure trois éléments principaux caractéristiques qui sont autant de cibles potentielles pour créer des interactions avec un ligand (Figure 70A). Le premier et le plus important est la présence des tétrades de guanines qui s'auto-stabilisent par des interactions de π -stacking ; les tétrades extérieures sont particulièrement disponibles et forment des surfaces planes hydrophobes idéales pour des interactions avec des molécules organiques planes riches en électrons, comme des composés aromatiques. Le second élément caractéristique est l'existence de boucles dont la séquence, la conformation et la longueur varient en fonction du quadruplexe. Le dernier élément se forme quant à lui entre les armatures sucre-phosphate portant les nucléobases créant ainsi un sillon apte à former de nombreuses interactions électrostatiques avec un ligand.^[11, 192, 276, 329]

Pour développer un ligand de quadruplexe sélectif, c'est-à-dire qui n'interagit pas avec d'autres formes d'ADN présentes dans le génome comme le duplexe, il est indispensable de distinguer les éléments structuraux qui lui sont propres de ceux inhérents à d'autres édifices. Ainsi, le ciblage des sillons ne paraît pas judicieux puisqu'ils sont une caractéristique majeure de la double hélice d'ADN avec la présence d'un petit et d'un grand sillon (qui sont toutefois de nature et de taille différentes de ceux d'un quadruplexe) ; quelques ligands de sillons de quadruplexes sélectifs font toutefois exception, comme le TOxaPy.^[330] L'interaction avec les boucles ne semble également pas optimale puisqu'elle résulterait d'une reconnaissance de séquence et non de structure qui aurait pour conséquence la fixation aléatoire à d'autres structures d'ADN ou d'ARN renfermant cette même séquence.

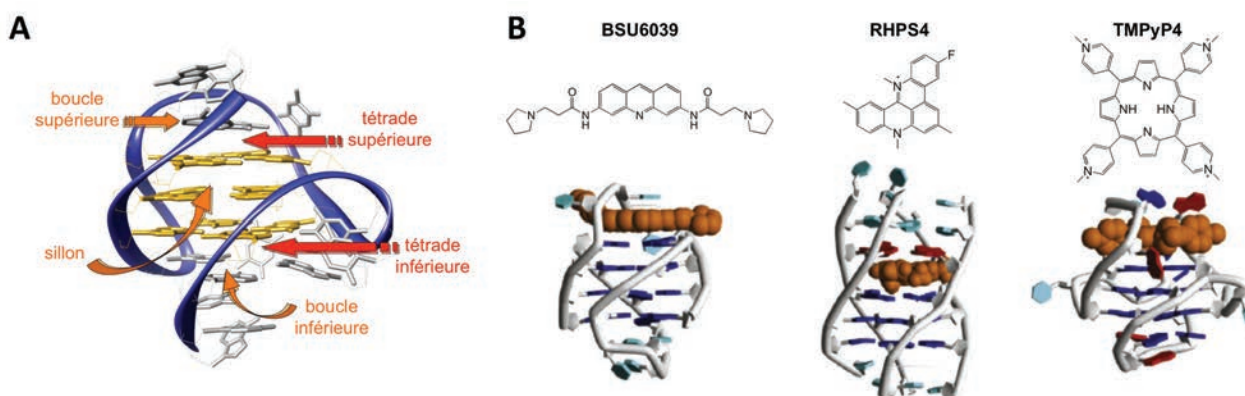


Figure 70 – Représentation schématique des différents sites d'interaction d'un quadruplexe d'ADN (A) et structure cristallographique (BSU6039) ou RMN (RHPS4 et TMPyP4) montrant l'interaction du ligand (en brun) avec une tétrade de guanines (en bleu foncé) extérieure (B).

L'approche la plus intéressante est donc clairement le ciblage des tétrades extérieures accessibles des quadruplexes, caractéristiques propres à ces édifices. De nombreux ligands ont été développés depuis la fin des années 1990 et reposent en très grande majorité sur la présence d'une large surface aromatique leur permettant d'interagir avec l'une de ces tétrades par interactions de π -stacking (Figure 70B).^[192] En complément, l'ajout d'un substituant peut permettre une seconde reconnaissance avec le sillon ou les boucles permettant d'ancrer un peu plus le ligand sur sa cible et d'augmenter ainsi son efficacité.

Cependant les interactions sélectives avec les tétrades nécessitent de grandes surfaces hydrophobes qui sont un obstacle à l'hydrosolubilité, et donc aux applications, des ligands. Ainsi, de nombreuses possibilités ont été testées comme l'ajout de bras PEG, de sucres ou plus fréquemment de fonctions amines protonables en milieu physiologique. Cette présence de charges peut conduire à une augmentation de la stabilisation du quadruplexe par interactions avec des groupements phosphates à proximité (chargés négativement), mais peut également provoquer une perte de sélectivité par des interactions purement électrostatiques de même nature avec la forme duplexe. De plus, pour des applications *in cellulo*, la charge influe sur le transport transmembranaire et peut donc s'avérer cruciale sur son efficacité globale puisqu'une molécule, même très active une fois sur un quadruplexe *in vitro*, doit dans un premier temps pouvoir l'atteindre *in vivo*.

La conception de ligands affins, sélectifs et hydrosolubles se révèle donc complexe et justifie leur pléthore au sein de la littérature scientifique.

VII.2 Aperçu des ligands dans la littérature

Actuellement, plus de 800 ligands de quadruplexes sont référencés dans la littérature scientifique et sont depuis peu accessibles sur une base de données en ligne, gratuite, qui propose pour bon nombre de composés un résumé détaillé des différentes analyses effectuées (FRET-melting, dichroïsme circulaire, analyses UV-visible, cytotoxicité, etc.) et des données structurales cristallographiques ou obtenues par RMN.^[331]

Plusieurs revues récentes regroupent également la majorité de ces composés,^[13, 275, 276] dont la revue très complète publiée dans *Current Pharmaceutical Design* par le groupe de H.-J. Park en mai 2012.^[11]

Pour ces raisons, nous nous contenterons de ne présenter ici que les grandes classes de ligands que nous illustrerons par les exemples les plus représentatifs :

- **les dérivés porphyriniques** (Figure 71) possèdent une large surface aromatique très hydrophobe contrebalancée dans le cas de la TMPyP4, exemple le plus emblématique de cette catégorie, par la présence des quatre substituants méthyl-pyridinium. La TMPyP4 interagit avec les quadruplexes de manière non spécifique et a pour inconvénient majeur de n'être absolument pas sélective puisqu'elle interagit très fortement avec les duplexes par intercalation d'un de ses pyridinium.^[11] Au sein du laboratoire, elle est d'ailleurs utilisée comme contrôle pour la plupart de nos expériences.

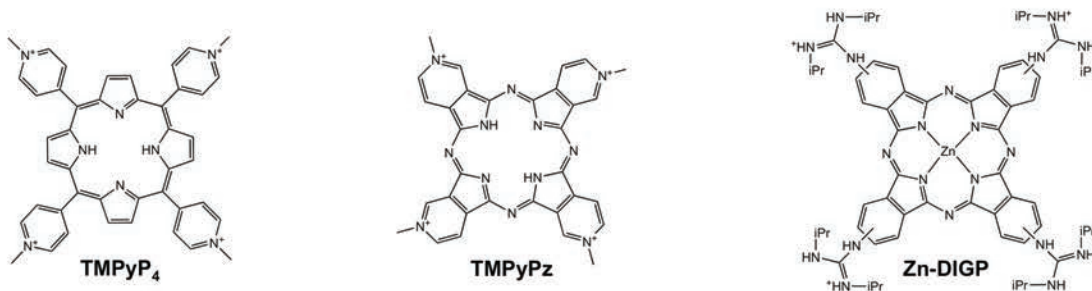


Figure 71 – Exemples de dérivés porphyriniques utilisés comme ligands de quadruplexes.

L'utilisation d'une porphyrazine telle TMPyPz permet un gain de sélectivité de l'ordre d'un facteur 30^[332] alors que des dérivés de phthalocyanines comme Zn-DIGP offrent une très bonne affinité pour la séquence ADN de l'oncogène C-myc ainsi qu'une excellente sélectivité vis-à-vis du duplexe.^[305]

- **les dérivés d'acridine** (Figure 72) comme le BRACO-19,^[333, 334] la quinacridine MMQ₃^[335] ou le AS1410^[196] possèdent de bonnes propriétés d'affinité et de sélectivité. L'ajout de deux chaînes tétra-peptidiques interagissant avec les boucles et les sillons du quadruplexe offre un second mode de reconnaissance bénéfique.^[336-338]

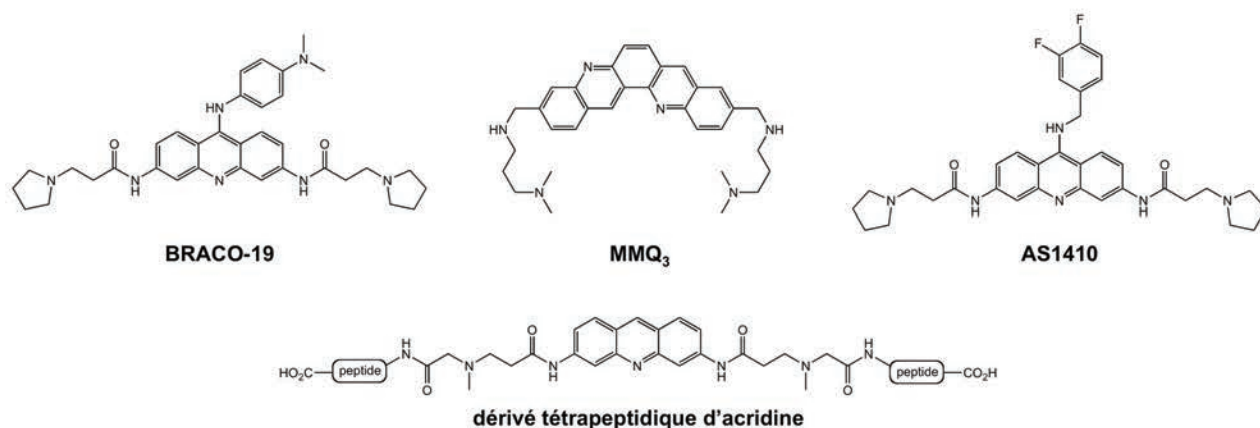


Figure 72 - Exemples de dérivés d'acridine utilisés comme ligands de quadruplexes.

- **les composés aromatiques condensés** (Figure 73) : certains dérivent de la fonctionnalisation d'un cœur polyaromatique tels le pérylène (PIPER) ou le naphthalène diimide (NDI).^[339] La quarfloxine, mise au point par le laboratoire Cylene Pharmaceuticals sur la base des travaux de l'équipe de L. Hurley, a atteint la phase clinique II et présente un large spectre d'utilisation pour le traitement de tumeurs ainsi qu'une activité sélective.^[10]

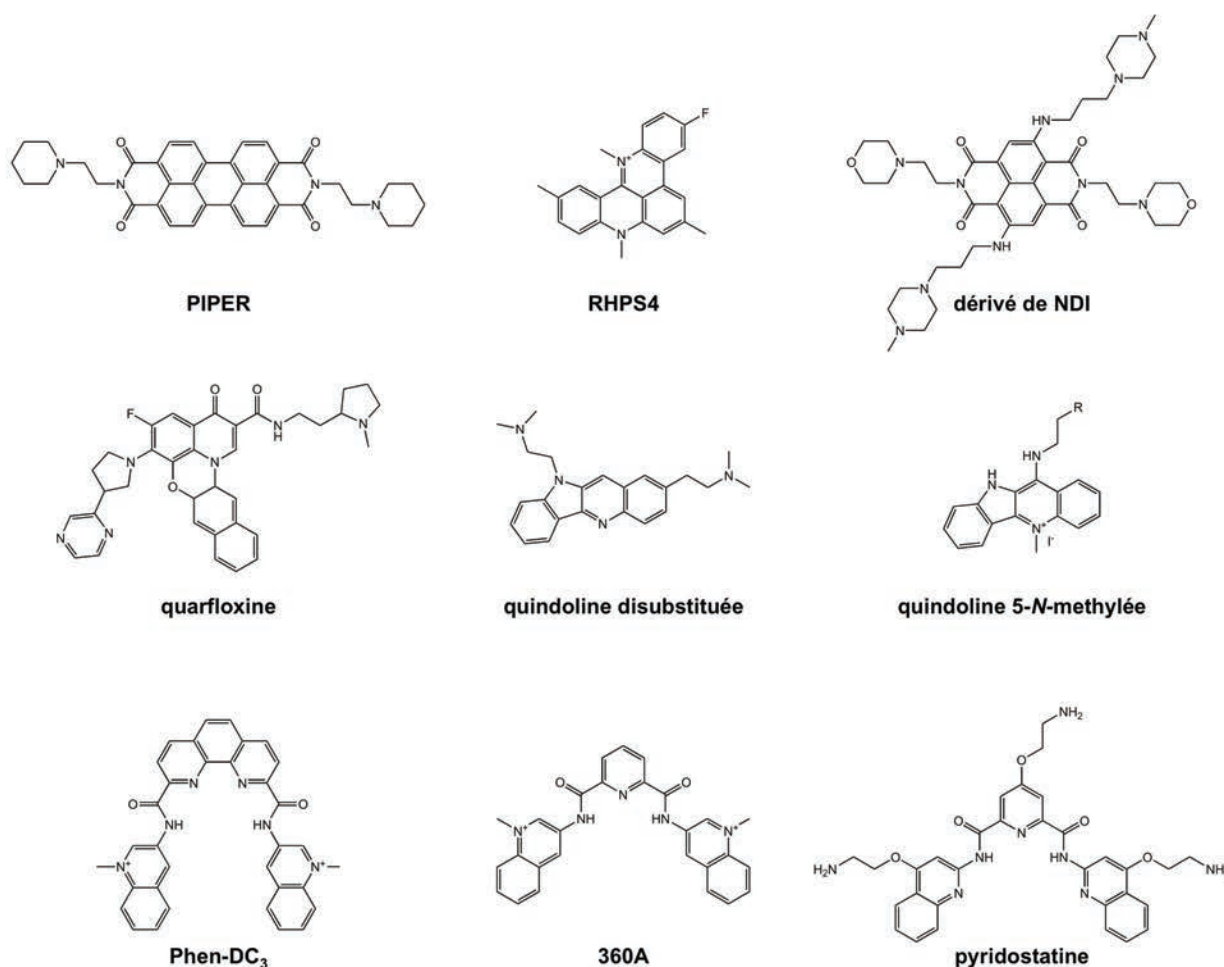


Figure 73 - Exemples de composés aromatiques condensés utilisés comme ligands de quadruplexes.

De nombreux dérivés de quindoline ont également été reportés et offrent une efficacité accrue lorsque la fonctionnalisation se situe en position 11. De manière intéressante, la quaternarisation de l'azote du groupement pyridine amène une charge cationique favorisant les interactions avec les quadruplexes antiparallèles en comparaison à son homologue neutre.^[340-345] Les bisquinolines possèdent deux bras aromatiques reliés entre eux par un pont urée, pyridine, triazine, etc., et ont de bonnes à d'excellentes affinités et sélectivités grâce à leur adaptation structurale, comme le Phen-DC3, le 360A ou plus récemment la pyridostatine.^[295-297]

- **les composés macrocycliques** (Figure 74) : la télomestatine, un composé naturel isolé en 2001 de *Streptomyces Annulatus*, est l'un des meilleurs ligands actuels et a été une source d'inspiration pour l'élaboration d'autres molécules, comme HXDL beaucoup plus hydrosoluble avec ses deux chaînes alkyl-amines.^[11]

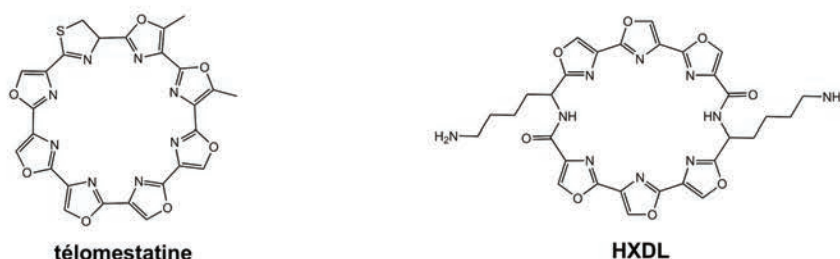


Figure 74 - Exemples de composés macrocycliques utilisés comme ligands de quadruplexes.

- **les composés organométalliques** (Figure 75) portent une charge cationique sur le métal favorable à leur interaction avec les oligonucléotides chargés négativement. La métallation de porphyrines par du Cu(II), du Zn(II) ou du Mn(III) augmente leur sélectivité par rapport à leur forme base libre grâce au placement du métal au centre de la tétrade de guanines avec laquelle il peut interagir au niveau des carbonyles. Le Ni-salphen^[346] offre quant à lui une très forte stabilisation et une bonne sélectivité tandis que le Pt-MPQ^[347] forme une sorte de pince et interagit simultanément avec les deux tétrades accessibles d'une part par interaction non covalente « classique » avec le dérivé de quinacridine et d'autre part grâce à la platination de la seconde tétrade accessible.

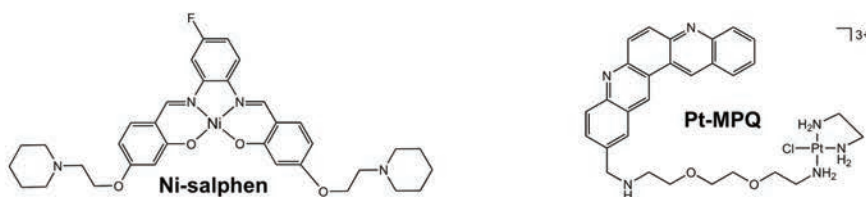


Figure 75 - Exemples de composés organométalliques utilisés comme ligands de quadruplexes.

- les **ligands de quadruplexes multimériques** (Figure 76) sont pour le moment peu nombreux et seulement deux exemples ont été reportés : celui de l'hélicène M1 dont l'énantiomère P1 correspondant se trouve être beaucoup moins efficace,^[348] et l'azatrux.^[349]

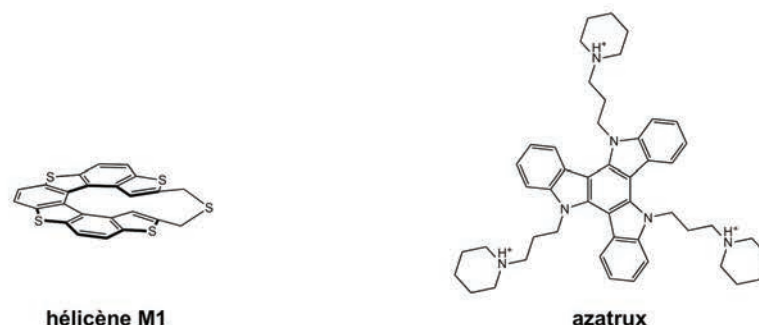


Figure 76 - Exemples de ligands de quadruplexes multimériques.

Tous deux interagissent à l'interface entre deux quadruplexes d'un même multimère.

VII.3 Concept des TASQ

L'idée originale de développer les premières molécules de type TASQ hydrosolubles pour le ciblage de quadruplexes est née de l'observation qu'un quadruplexe est d'autant plus stable que le nombre de tétrades qui le constitue est élevé. En effet, un quadruplexe à deux tétrades empilées (comme TBA pour *thrombin binding aptamer* de séquence d(G₂T₂G₂TG₂T₂G₂)) est moins stable qu'un quadruplexe à trois tétrades (comme 22AG de séquence télomérique humaine d(AG₃T₂AG₃T₂AG₃T₂AG₃)) lui même également moins stable qu'un quadruplexe à quatre tétrades (comme la séquence télomérique du protozoaire cilié *Oxytricha* d(T₄G₄T₄G₄T₄G₄)) avec des température de fusion T_m évaluées par UV-melting de 50°C, 63°C et 84°C dans 100mM K⁺, respectivement (Figure 77).

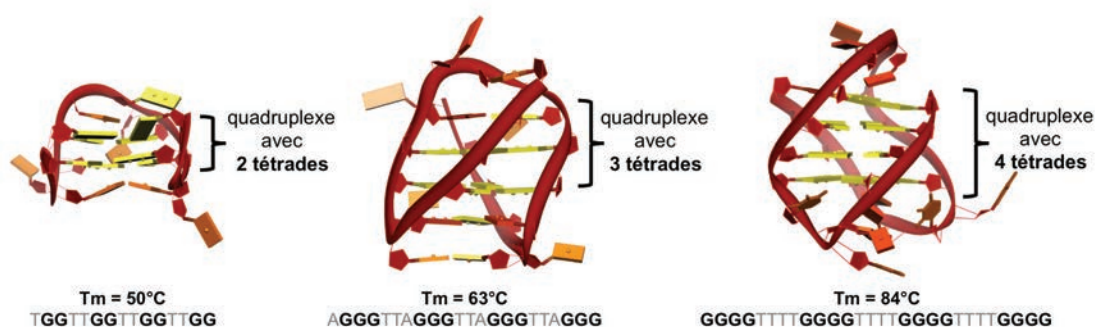


Figure 77 – Evolution de la stabilité des quadruplexes en fonction du nombre de tétrade.

En suivant cette logique, l'ajout d'une tétrade synthétique supplémentaire à un quadruplexe d'ADN ou d'ARN devrait conduire à l'augmentation globale de sa stabilité. Ce processus de reconnaissance biomimétique basé sur l'interaction entre la tétrade synthétique d'un TASQ et une tétrade naturelle d'un quadruplexe semble favorisé par des interactions de π -stacking et par la chélation entre les deux feuillets de tétrades d'un cation alcalin (Na^+ ou K^+) (Figure 78).

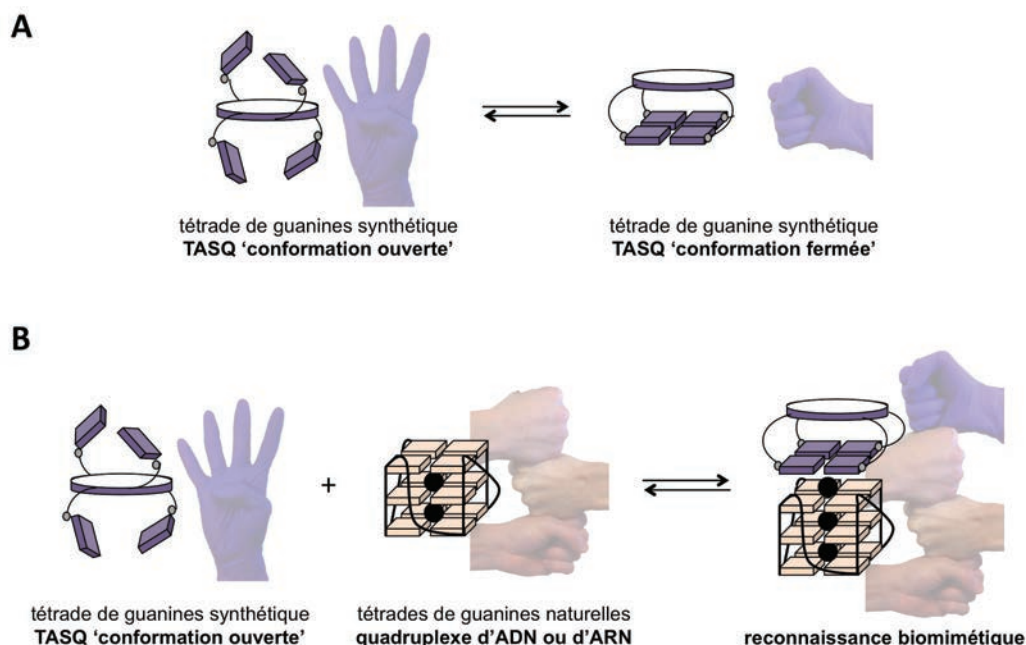


Figure 78 – Equilibre entre les conformations ouvertes et fermées d'un TASQ (A) et concept de l'utilisation des TASQ comme ligands de quadruplexes basés sur une reconnaissance biomimétique (B).

Sur cette base, nous avons conçu une première série de TASQ construite autour d'un cœur cyclène et de bras alkyles guanines (les DOTASQ) qui a permis de valider l'approche et de développer deux nouveaux composés pour lesquels nous avons fait varier la nature du *template* (le PorphyrSQ) ou la nature des bras ($^{\text{PNA}}$ DOTASQ). La conception de ces molécules bioinspirées ainsi que les évaluations biophysiques de leurs propriétés d'interaction avec les quadruplexes d'ADN et d'ARN font l'objet des parties VIII, IX et X de ce chapitre.

VIII. La première série de TASQ : les DOTASQ

VIII.1 Concept et synthèse

Peu de temps avant le début de ce projet, la notion de tétrade synthétique avait été introduite pour la première fois par J. Sherman avec la mise au point de la synthèse d'une série de TASQ

lipophiles construits autour d'un motif calixarène sur lequel quatre résidus guanosines étaient greffés (Figure 79A).^[17] Ces TASQ présentaient l'inconvénient majeur de ne pas être solubles en milieu aqueux, limitant ainsi fortement leurs applications, mais avaient permis de valider l'approche selon laquelle quatre guanines pouvaient former intramoléculairement une tétrade (voir V.1).

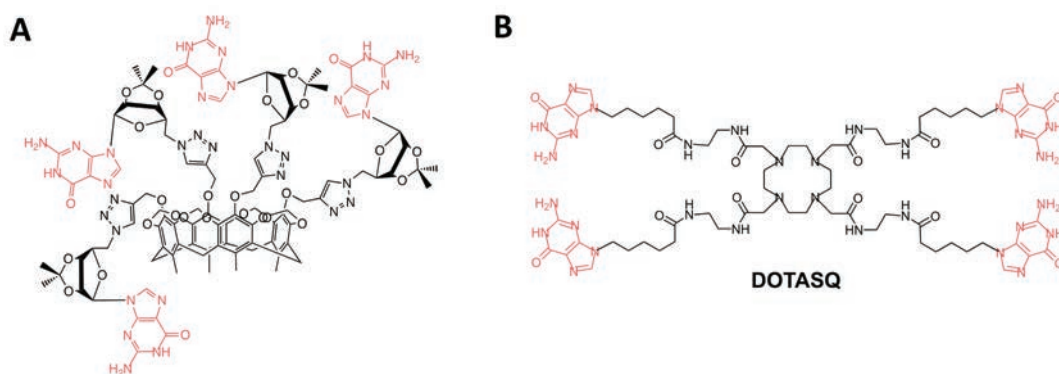


Figure 79 – Structure chimique du premier TASQ lipophile construit autour d'un cœur calixarène (A) et premier TASQ hydrosoluble développé au cours de cette thèse : le DOTASQ (B).

Pour favoriser l'hydrosolubilité de ce type de composé, nous nous sommes d'abord focalisés sur le choix du *template*, élément primordial pour la construction de tels édifices. Rapidement, notre choix s'est porté sur un polyazamacrocycle, le DOTA (pour acide 1,4,7,10-tétraazacyclododécane-*N,N',N'',N'''*-tétraacétique) connu pour ses excellentes propriétés de solubilité en milieu aqueux et également pour sa capacité de chélation des métaux qui lui valent d'être utilisé, entre autre, comme agent de contraste en IRM sous forme de complexe de Gd(III).^[252, 350] De plus, sa symétrie C_4 identique à celle d'une tétrade de guanines semblait optimale pour favoriser sa formation intramoléculaire.

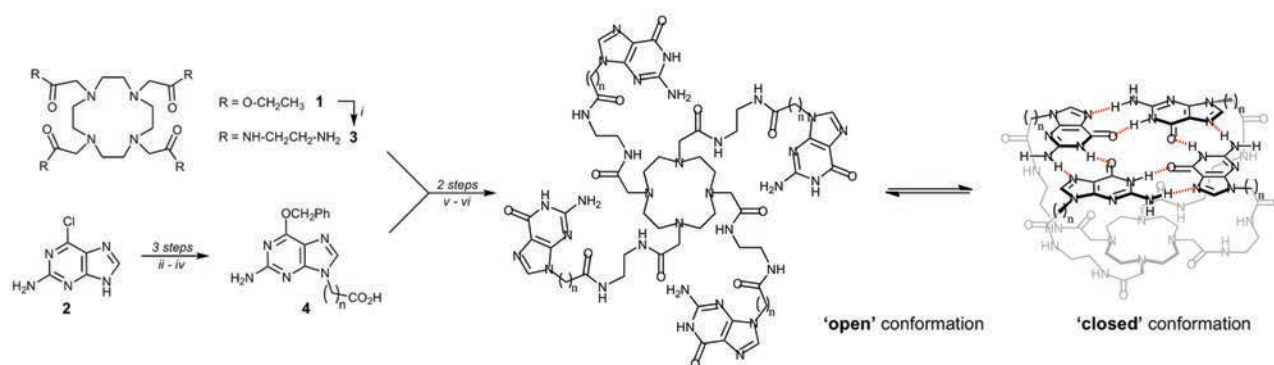


Figure 80 – Synthèse, structure et représentation schématique des deux conformations possibles des DOTASQ- C_1 ($n=1$) et DOTASQ- C_5 ($n=5$).

Ainsi, la première série DOTASQ (pour *DOTA-templated synthetic G-quartet*, Figure 79B) a été synthétisée en 6 étapes convergentes au départ du DOTA-Et **1** et de la 2-amino-6-chloropurine **2** (Figure 80). En adaptant les travaux de R. Condom relatifs à la synthèse de PNA,^[351] deux purines fonctionnalisées par une chaîne à un et à cinq carbones ont été obtenues et ont conduit *in fine* à

l'obtention des composés DOTASQ-C₁ et DOTASQ-C₅, respectivement. Le choix de ces deux longueurs de bras est issu d'une étude préliminaire par modélisation moléculaire par optimisation complète de géométrie, sans contraintes de symétrie et avec un champ de force MM3 ou semi-empirique AM1 sous le logiciel CAChe Worksystem Pro[®]. Elle nous a permis de mettre en évidence que dans les deux cas, la longueur des bras était favorable à la formation de la tétrade intramoléculaire de guanines, avec toutefois de plus fortes tensions dans le cas du DOTASQ-C₁.

Alors que la solubilité des intermédiaires de réaction s'avère parfois difficile et très différente en fonction de la longueur de la chaîne alkyle, les DOTASQ finaux offrent eux une bonne solubilité dans les tampons aqueux à pH neutre, confortant notre choix judicieux de *template*, ainsi que dans des solvants fortement dissociant comme le DMSO.

VIII.2 Analyses préliminaires de la structuration par RMN ¹H

Afin de démontrer que les DOTASQ existent sous deux conformations en équilibre, l'une dite 'ouverte' où les guanines sont indépendantes, et l'autre 'fermée' au sein de laquelle la tétrade de guanines se forme, des études par RMN ¹H ont été entreprises sur le DOTASQ-C₅ par S. Amrane à l'IECB sous la direction du Dr. J.-L. Mergny.

La méthode employée ici dérive directement de celle utilisée pour l'étude de la structuration des quadruplexes d'ADN. Elle consiste à suivre, en milieu tamponné, l'évolution des signaux des protons de l'amide (de l'acide imine) en position 1 de la guanine qui sont impliqués dans le réseau de liaisons hydrogènes formé par la tétrade. Ainsi, lorsque les guanines sont indépendantes, les protons sont échangeables et ne sont pas visibles en RMN ¹H alors que dès que la tétrade se forme, la vitesse d'échange est très fortement réduite et ils deviennent alors observable autour de 11ppm (Figure 81A).^[352]

Dans notre cas, à 293K, le signal obtenu est très faible, signifiant que les échanges avec le solvant sont rapides ; la descente en température à 283K fait apparaître un signal large centré à 10,8ppm indiquant que les protons sont impliqués dans des liaisons hydrogènes, démontrant ainsi l'interaction entre les guanines et, probablement, la formation de la tétrade (Figure 81B).

Une seconde étude a été menée en présence cette fois-ci d'un quadruplexe d'ADN à trois tétrades naturelles de séquence G₃TG₃T₂G₃TG₃. Puisque chacune de ces tétrades renferme quatre

protons amides, le nombre total de signaux obtenu est donc de douze pour le quadruplexe complet (Figure 81C, haut).

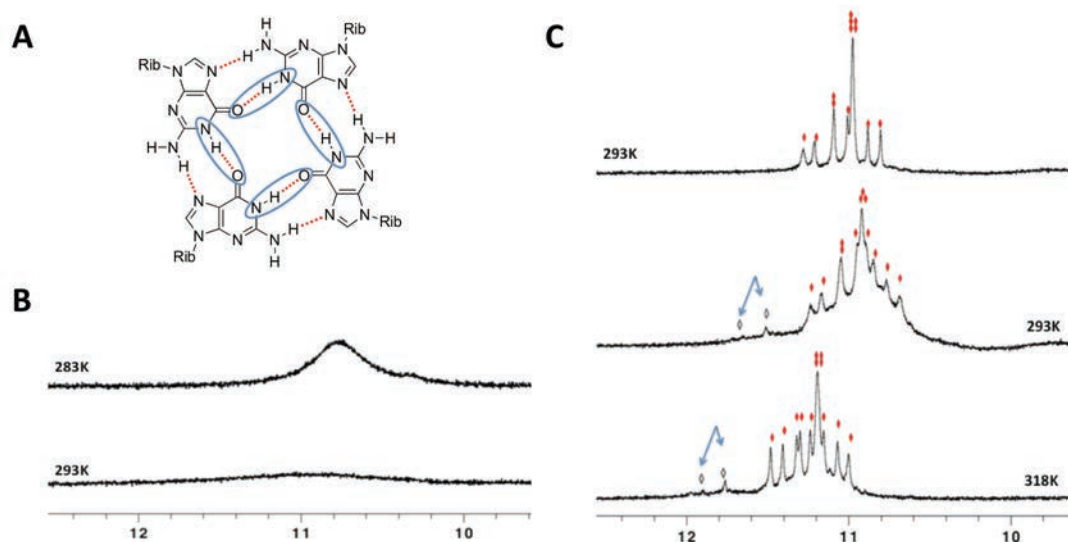


Figure 81 – Tétrade de guanines et protons des amides (cercle bleu) dont le déplacement chimique est compris entre 10 et 12ppm (A). Parties des spectres RMN ^1H du DOTASQ- C_5 en solution (à 3,2mM dans un tampon à 20mM Kpi, 70mM KCl, pH 7,0) (B) et de $\text{G}_3\text{TG}_3\text{T}_2\text{G}_3\text{TG}_3$ seul (haut), en présence de 1 équivalent de DOTASQ- C_5 à 293K (milieu) puis à 318K (bas) dans une solution tampon à 4mM Kpi, 14mM KCl, pH 7,0 (C).

L'ajout d'un équivalent de DOTASQ- C_5 conduit à un élargissement et à un déplacement des signaux qui est une conséquence caractéristique d'une interaction entre un ligand et l'ADN. De plus, l'apparition de deux nouveaux pics entre 11,4 et 11,7ppm attribués aux protons amides de la tétrade synthétique du DOTASQ confortent l'hypothèse que l'interaction a bien lieu grâce à la formation intramoléculaire de la tétrade (Figure 81C, milieu). Néanmoins, les intégrations révèlent que seulement 20% du ligand se trouve sous sa conformation fermée (contre 80% en conformation ouverte inactive).

L'augmentation de la température de l'échantillon de 293K à 318K entraîne classiquement un affinement et un déplacement des pics du quadruplexe d'ADN, et il est intéressant de noter que les signaux du DOTASQ- C_5 sont toujours présents, confirmant l'interaction par une reconnaissance biomimétique (Figure 81C, bas).

Les mêmes expériences ont été réalisées sur un autre quadruplexe de séquence $\text{G}_3\text{TG}_3\text{T}_3\text{G}_3\text{TG}_3$. Les spectres obtenus offrent une moins bonne résolution mais permettent toutefois d'observer que là encore, l'ajout de ligand modifie les signaux de l'ADN et fait apparaître de nouveaux pics induits par la formation de la tétrade synthétique.

VIII.3 Evaluation par FRET-melting

Afin de quantifier l'efficacité de nos composés en tant que ligands de quadruplexes, nous avons procédé à une série d'évaluations biophysiques. Les premières reposaient sur la méthode d'UV-melting^[194, 353] où l'évolution de l'absorbance des guanines d'un quadruplexe est mesurée en fonction de la température en présence et en l'absence du ligand à tester. Cependant, cette technique s'est avérée inadaptée dans notre cas puisque les DOTASQ, composés de quatre guanines, absorbaient également la lumière entre 250 et 300nm et interféraient donc avec la mesure.

Pour pallier ce problème, la méthode de FRET-melting a été envisagée car elle permet de se placer dans une gamme de longueurs d'ondes située entre 480 et 520nm, grâce à l'utilisation de fluorophores greffés sur l'ADN.

VIII.3.A Principe de la méthode de FRET-melting

La méthode de FRET-melting développée par l'équipe du Dr. J.-L. Mergny^[354] repose sur l'utilisation d'un oligonucléotide pouvant former une structure secondaire (duplexe, quadruplexe, jonction à trois voies ou à quatre voies, etc.)^[355] marquée par deux sondes, classiquement la 6-carboxyfluoresceine (notée F ou FAM) et la tétraméthylrhodamine (T ou TAMRA) entre lesquelles il existe un phénomène de FRET (*fluorescence resonance energy transfer*), conséquence d'un recouvrement spectral entre l'émission d'une sonde et l'absorbance de l'autre.^[356]

Dans le cas d'un quadruplexe d'ADN structuré à 20°C, les deux sondes sont proches dans l'espace et l'émission de fluorescence de la fluorescéine est directement réabsorbée par la rhodamine ; la valeur de l'émission de la fluorescéine à 515nm est donc très faible. Lors de l'augmentation de la température du milieu, le quadruplexe d'ADN dont la structure n'est maintenue que par un réseau de liaisons hydrogènes va se déstructurer, éloignant ainsi les deux sondes et restituant donc progressivement l'émission de la fluorescéine dans le milieu.

Les courbes de l'évolution de cette fluorescence (normalisées entre 0 et 1, Figure 82) en fonction de la température sont caractéristiques de chaque quadruplexe et permettent de déterminer la valeur de $T_{1/2}$ appelée température de fusion (ou de *melting*) correspondant à la température à laquelle la moitié de l'échantillon est sous une forme structurée, l'autre moitié sous une forme déstructurée (correspondant donc à la température pour laquelle la fluorescence normalisée est de 0,5). Ainsi, plus un quadruplexe est stable, plus l'énergie thermique nécessaire pour le déstructurer est grande et par conséquent, plus la valeur de $T_{1/2}$ est élevée.

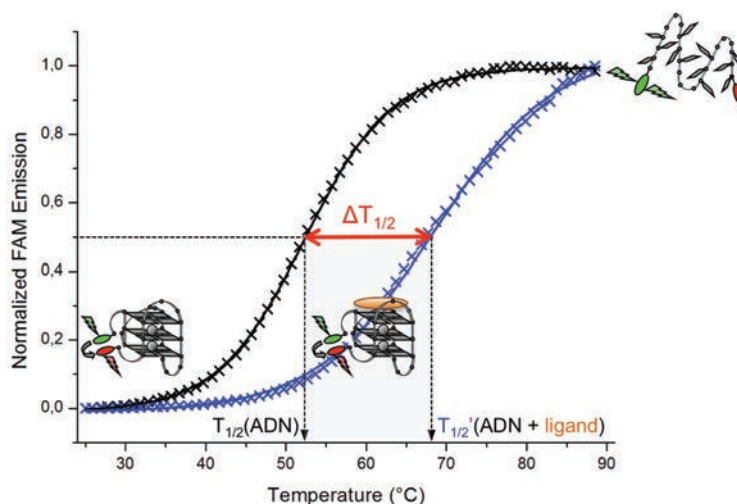


Figure 82 - Principe de la méthode de FRET-melting et courbes-types obtenues pour un quadruplexe marqué par FAM (cercle vert) et TAMRA (cercle rouge) en l'absence (courbe noire) ou en présence (courbe bleue) de ligand (en orange).

La méthode de FRET-melting permet alors de mesurer l'affinité apparente d'un ligand pour un quadruplexe d'ADN en comparant le $T_{1/2}$ de l'ADN seul au $T_{1/2}'$ obtenu en présence de ligand. Plus la valeur $\Delta T_{1/2} = T_{1/2}' - T_{1/2}$ est élevée, plus la stabilisation est forte et donc plus le ligand est efficace (Figure 82).

La sélectivité peut également être évaluée par l'ajout de quantités croissantes d'ADN compétiteur non marqué (ex : duplexe d'ADN) dans un mélange d'ADN marqué (ex : quadruplexe) et de ligand. Si la variation du $\Delta T_{1/2}$ est nulle ou faible, cela signifie que le ligand reste préférentiellement sur l'ADN marqué et est donc sélectif ; en revanche, une variation importante du $\Delta T_{1/2}$ témoigne d'une meilleure affinité du ligand pour l'ADN non marqué compétiteur par rapport à l'ADN marqué d'intérêt : il n'est donc pas sélectif.

Notons qu'une attention toute particulière doit être portée sur le choix du compétiteur qui doit posséder une température de fusion T_m (évaluée par UV-melting) nettement supérieure à celle de l'ADN marqué (classiquement imposée comme supérieure de 20°C) afin de s'assurer que le compétiteur soit encore parfaitement structuré dans la gamme de température de stabilisation du ligand.

VIII.3.B Résultats de FRET-melting des DOTASQ

Les mesures de FRET-melting ont donc été effectuées sur la séquence télomérique humaine F21T (FAM-d(G₃(T₂AG₃)₃)-TAMRA, $T_{1/2} = 52^\circ\text{C}$) seule, en présence de 5 équivalents de DOTASQ-C₁ ou DOTASQ-C₅, puis toujours en présence de DOTASQ avec successivement 15 et

50 équivalents de duplexe compétiteur ds26 (de séquence auto-complémentaire d(CA₂TCG₂ATCGA₂T₂CGATC₂GAT₂G), $T_m = 79^\circ\text{C}$).

Les résultats obtenus (Figure 83) sont dans un premier temps décevants, surtout pour le DOTASQ-C₁ qui montre une stabilisation nulle à faible ($\Delta T_{1/2} = 0,0^\circ\text{C}$ et $1,8^\circ\text{C}$ en Na⁺ et K⁺, respectivement), légèrement plus élevée dans le cas du DOTASQ-C₅ ($\Delta T_{1/2} = 2,5^\circ\text{C}$ et $5,3^\circ\text{C}$ en Na⁺ et K⁺, respectivement).

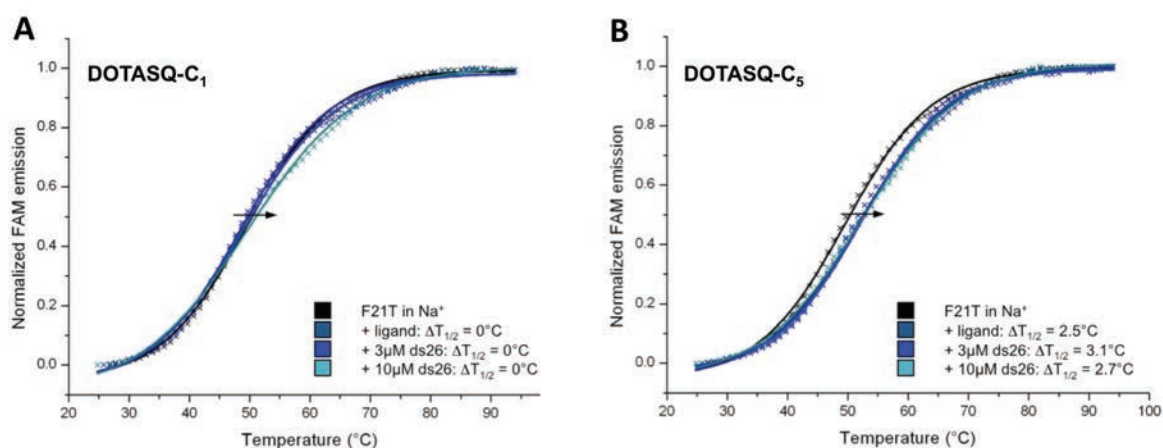


Figure 83 – Résultats de FRET-melting obtenus pour le DOTASQ-C₁ (A) et le DOTASQ-C₅ (B) en milieu sodium.

Ces faibles affinités tendent à montrer que la formation de la tétrade de guanines intramoléculaire permettant l'interaction avec les quadruplexes n'est pas favorisée par l'architecture même des DOTASQ, ce qui est cohérent avec les résultats préliminaires obtenus par RMN ¹H où seul 20% du DOTASQ-C₅ était présent sous sa conformation 'fermée' active. Ainsi, deux hypothèses peuvent être proposées : soit les bras et le *template* sont trop flexibles et diminuent la probabilité de rencontre des quatre guanines dans un même plan, soit au contraire ils sont trop rigides et empêchent les quatre guanines d'être simultanément du même côté du *template*, condition *sine qua non* à la formation de la tétrade.

Dans notre cas, le cœur cyclène est un macrocycle offrant une grande flexibilité et les bras guanines fonctionnalisant les amines du cycle possèdent par conséquent un fort degré de liberté.

L'idée était donc de réorganiser la structure des DOTASQ en forçant les quatre bras-guanines à se trouver du même côté du cycle DOTA, favorisant par la même la formation de la tétrade de guanines. Pour cela nous avons opté pour l'insertion d'un métal dans la cavité : le terbium.

VIII.3.C Etude des Tb.DOTASQ

Parmi les différents métaux envisagés, le terbium a finalement été sélectionné pour sa forte coordination, sa forte stabilité dans les cycles DOTA et pour ses propriétés de luminescence qui pourraient permettre à long terme une détection *in cellulo* de ces composés.

Ainsi, le métal se coordine aux quatre atomes d'azote du macrocycle et aux quatre carbonyles des bras en formant virtuellement deux plans parallèles.^[252] Les guanines possèdent alors un degré de liberté moindre et l'équilibre entre la conformation 'ouverte' et 'fermée' devrait se déplacer en direction de cette dernière, formant une succession de trois couches « NOG » stables (Figure 84).

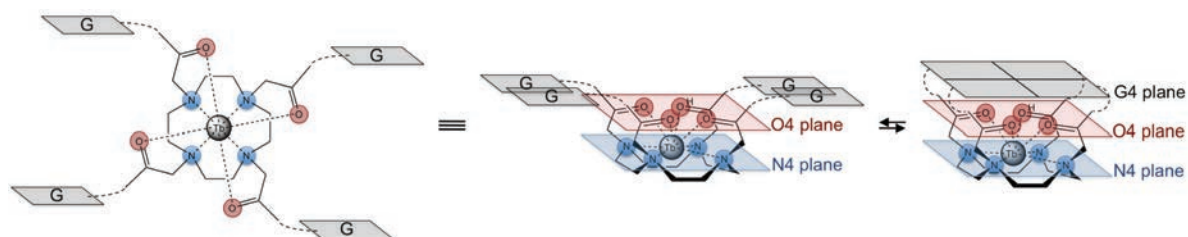


Figure 84 – Représentation schématisée des répercussions structurales de l'insertion d'un métal dans la cavité du cyclène des DOTASQ.

L'insertion du terbium dans les deux DOTASQ se fait par addition d'un équivalent de $\text{Tb}(\text{NO}_3)_3 \cdot \text{H}_2\text{O}$ dans un mélange MeOH/tampon cacodylate de lithium (20mM à pH 7,2) laissé sous agitation pendant trois jours à 55°C, conduisant aux Tb.DOTASQ-C₁ et Tb.DOTASQ-C₅.

Les mêmes expériences de FRET-melting (Figure 85) permettent d'observer sur les séries terbium une augmentation drastique de la stabilisation de F21T avec une efficacité pour le Tb.DOTASQ-C₁ ($\Delta T_{1/2} = 9,8^\circ\text{C}$ (Na^+) et $10,6^\circ\text{C}$ (K^+)) proche de celle du Tb.DOTASQ-C₅ ($\Delta T_{1/2} = 10,5^\circ\text{C}$ (Na^+) et $9,4^\circ\text{C}$ (K^+)).

L'addition de duplexe compétiteur ds26 permet de calculer le facteur de sélectivité $^{\text{FRET}}S$ défini selon :

$$^{\text{FRET}}S = \frac{^{\text{FRET}}\Delta T_{1/2} + \text{ds26}}{^{\text{FRET}}\Delta T_{1/2}}$$

Pour les deux composés, les résultats démontrent une excellente sélectivité avec des valeurs $^{\text{FRET}}S$ toujours supérieures à 90% et ce, quel que soit le milieu ionique (K^+ ou Na^+) utilisé. Un léger avantage est à noter pour le Tb.DOTASQ-C₅ dont la plus faible valeur obtenue est de 97,6% contre 93% pour le Tb.DOTASQ-C₁.

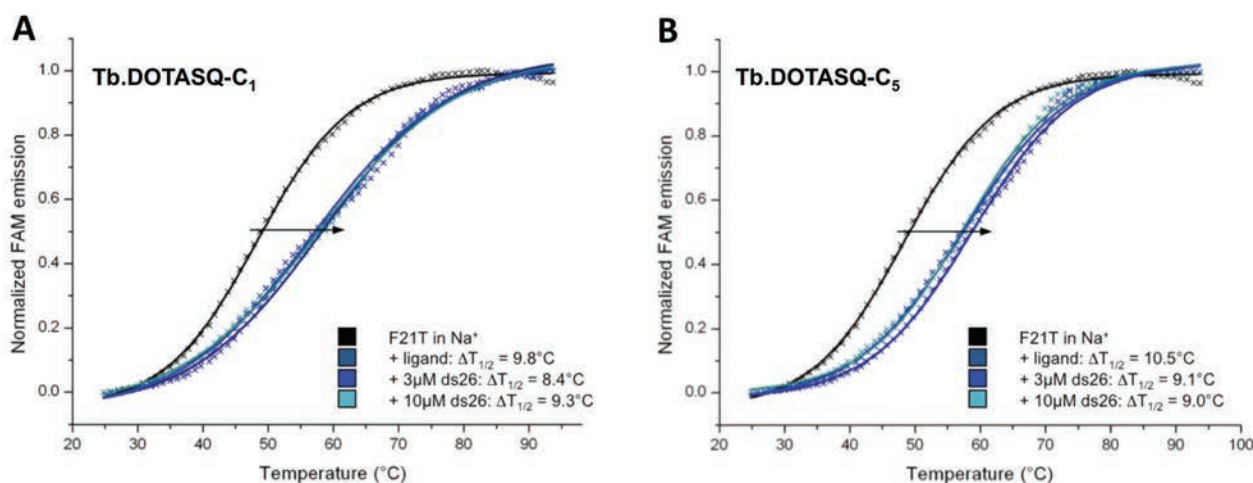


Figure 85 - Résultats de FRET-melting obtenus pour le Tb.DOTASQ-C₁ (A) et le Tb.DOTASQ-C₅ (B) en milieu sodium.

L'augmentation de la charge globale de ces édifices, associée à une contrainte géométrique induites par la présence du terbium, apportent un gain drastique en terme d'efficacité tout en présentant une sélectivité exemplaire. De par leur conception même, les Tb.DOTASQ ne contiennent pas de substituant particulièrement propice aux interactions avec le duplexe d'ADN, mise à part les guanines elles-mêmes. L'apport de trois charges positives aurait pu occasionner une certaine affinité pour la double hélice ce qui n'est pas le cas au vu de ces résultats de FRET-melting. Les Tb.DOTASQ sont donc à la fois des composés affins et sélectifs des quadruplexes d'ADN.

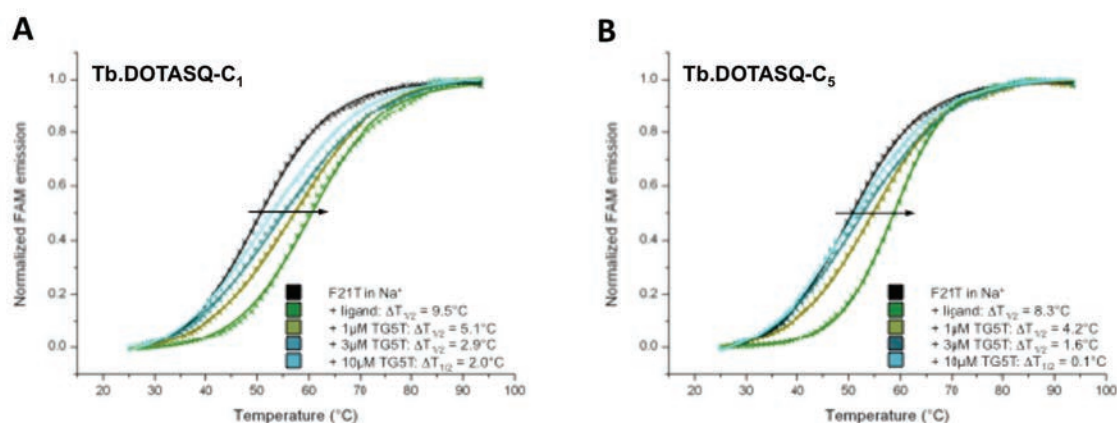


Figure 86 - Résultats de FRET-melting obtenus pour le Tb.DOTASQ-C₁ (A) et le Tb.DOTASQ-C₅ (B) en milieu sodium avec [TG₅T]₄ comme compétiteur.

Pour s'assurer du mode d'interaction de ces ligands avec les quadruplexes, une nouvelle série d'expériences est réalisée, cette fois-ci en présence d'un quadruplexe tétramoléculaire compétiteur [TG₅T]₄ ($T_m > 90^\circ\text{C}$) sans boucle offrant deux tétrades extérieures particulièrement disponibles. Pour les deux Tb.DOTASQ (Figure 86), la diminution du $\Delta T_{1/2}$ est importante jusqu'à atteindre 0,1°C à 50 équivalents de compétiteur, venant appuyer les conclusions obtenues par RMN ^1H selon

lesquelles l'interaction entre les (Tb.)DOTASQ et les quadruplexes se fait bien par reconnaissance entre les tétrades de guanines synthétiques et naturelles.

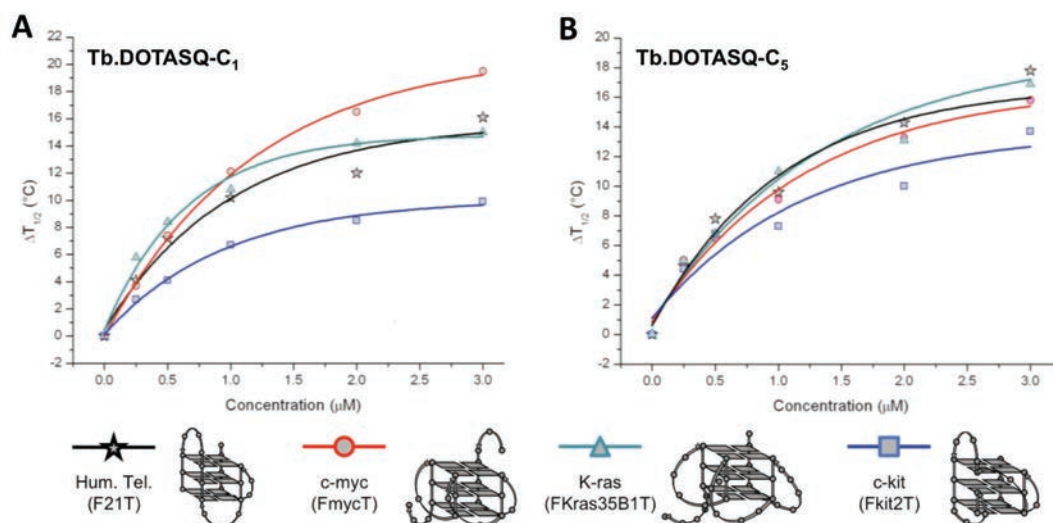


Figure 87 – Evolution de la stabilisation $\Delta T_{1/2}$ de quatre quadruplexes d'ADN en présence de quantités croissantes de Tb.DOTASQ-C₁ (A) et de Tb.DOTASQ-C₅ (B).

Enfin, d'autres quadruplexes de séquences oncogéniques équipés de sondes fluorescentes ont été étudiés (FmycT, FkitT et FKrasT, Figure 87) : leur stabilisation est effective avec les deux Tb.DOTASQ, et confirme la validité de la conception de ces composés qui, interagissant par les tétrades extérieures par un processus biomimétique, sont des ligands « universels » de quadruplexes.

VIII.4 Conclusion

Les DOTASQ sont les premiers modèles de TASQ hydrosolubles dont les expériences préliminaires de RMN ^1H tendent à montrer la formation de la tétrade de guanines lorsqu'ils sont seuls en solution et surtout, lorsqu'ils interagissent avec un quadruplexe d'ADN. Néanmoins, leur efficacité en terme de stabilisation est moindre à cause d'une trop grande liberté des bras guanines qui est défavorable à la formation de la tétrade. Ce problème est pallié par l'insertion d'un Tb(III) dans la cavité du macrocycle qui en le rigidifiant impose aux guanines de se placer du même côté du *template* ; les guanines sont par conséquent plus aptes à s'auto-assembler. Cette modification se traduit par une forte augmentation de la stabilisation des quadruplexes qui, en fonction de la séquence étudiée, varie de 6°C à 12°C. Ces valeurs sont loin de celles des meilleurs ligands de la littérature (à titre de comparaison $\Delta T_{1/2}(\text{360A}) = 21^\circ\text{C}$, $\Delta T_{1/2}(\text{télomestatine}) = 24^\circ\text{C}$ et $\Delta T_{1/2}(\text{BRACO-19}) = 27^\circ\text{C}$ sur la séquence F21T),^[276] mais offrent l'avantage d'une très forte sélectivité qui est le point crucial du développement de ligands. Cette propriété est directement liée

au mode d'action complètement innovant fondé sur une reconnaissance biomimétique entre une tétrade synthétique (DOTASQ) et une tétrade naturelle (quadruplexe).

Ces premiers prototypes valident l'approche bioinspirée et ouvrent la voie à l'élaboration de nouveaux TASQ.^[20] La première optimisation a ainsi consisté en la recherche d'un nouveau *template* qui par ses propriétés intrinsèques favoriserait la formation de la tétrade intramoléculaire.

IX. Modification du *template* : le PorphySQ

Pour sélectionner un nouveau *template*, nous sommes partis du constat qu'une partie de la stabilité des quadruplexes d'ADN naturels provient des interactions de π -stacking entre les feuillets consécutifs de tétrades de guanines qui les constituent.^[6, 181] De plus, la grande majorité des ligands développés à ce jour sont construits autour de grandes surfaces aromatiques également capables de créer ce type d'interaction avec l'une des tétrades accessibles des quadruplexes (voir VII.1 et VII.2). Ainsi, la formation de ce type d'interaction au sein même d'un TASQ favoriserait l'auto-assemblage des guanines et offrirait donc potentiellement un gain en terme de stabilisation.

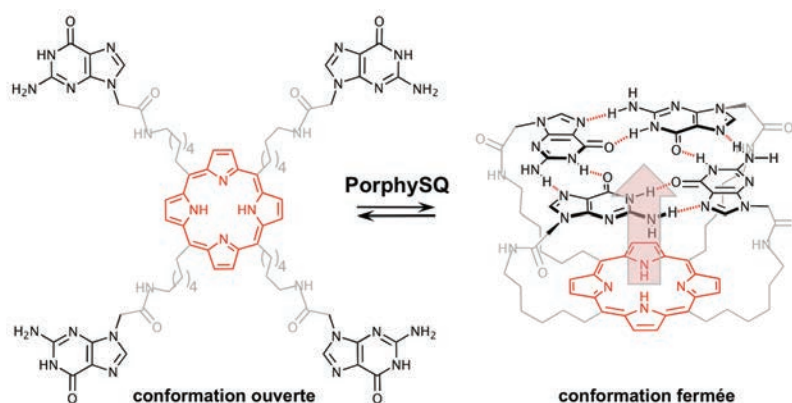


Figure 88 – Structure, équilibre conformationnel du PorphySQ et représentation des interactions de π -stacking entre le cycle porphyrinique et la tétrade de guanines (à droite, flèche rouge).

Il nous a alors semblé particulièrement pertinent d'utiliser en remplacement du cycle DOTA un motif porphyrinique, connu pour favoriser le π -stacking avec les tétrades de guanines comme l'attestent les multiples études entreprises sur la TMPyP4.^[276, 357, 358] Cette dernière offre de bons résultats en terme d'affinité pour les quadruplexes même si elle reconnaît toutes les autres structures secondaires d'ADN.^[355]

Ainsi, nous avons mis au point le PorphySQ (Figure 88) dont la synthèse et les études d'interactions seront discutées ci-après.

IX.1 Etudes préliminaires par modélisation moléculaire

Le remplacement du cycle DOTA par une porphyrine, augmentant la taille du cycle de 12 à 16 atomes, nécessite l'optimisation de la longueur des bras le reliant aux guanines. Un bras plus court que celui du DOTASQ-C₅ (le meilleur candidat DOTASQ) a été choisi afin de minimiser l'hydrophobie induite par les chaînes alkyles, passant ainsi de 12 à 9 atomes (dont sept méthylènes).

Avant d'entreprendre la synthèse du PorphySQ, des études de minimisation d'énergie ont été entreprises par le Dr. P. Richard à l'ICMUB en utilisant la méthode semi-empirique de l'Hamiltonien RM1.

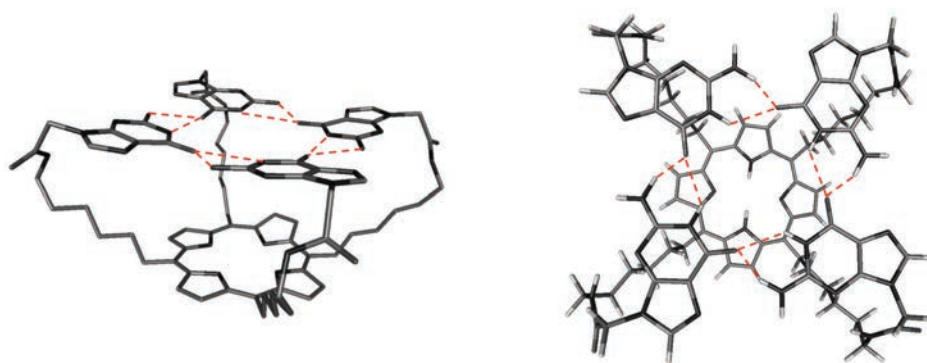


Figure 89 – Vue latérale (gauche) et supérieure (droite) de la structure du PorphySQ obtenue par modélisation moléculaire. Les liaisons hydrogènes sont représentées en pointillés rouges.

Les résultats obtenus *in silico* (Figure 89) montrent la présence de la tétrade intramoléculaire plane, située à 6,5Å au-dessus du cycle porphyrinique et dont la configuration est similaire aux études B3LYP DFT effectuées sur quatre guanines non substituées et publiées par L. Gu et J. Leszczynski en 2000.^[359] Notons que le réseau de liaisons hydrogènes, obtenu ici en l'absence de solvant et de cations alcalins (Na⁺ ou K⁺) diffère de celui classiquement obtenu pour une tétrade de guanines en milieu physiologique où les purines subissent une légère rotation due à l'interaction entre le carbonyle en position 6 et le cation qui se place au centre de la tétrade. Ici, les liaisons se créent uniquement entre le carbonyle et le proton de l'amine en position 2 ainsi qu'avec celui de l'amide en position 1. Cette étude préliminaire valide donc le concept de ce nouveau TASQ et approuve le choix de la longueur des bras.

IX.2 Synthèse du PorphySQ

La synthèse, présentée Figure 90, a été réalisée par le Dr. H.-J. Xu au sein de l'équipe du Pr. C. P. Gros à l'ICMUB, spécialisée dans le développement de dérivés porphyriniques.

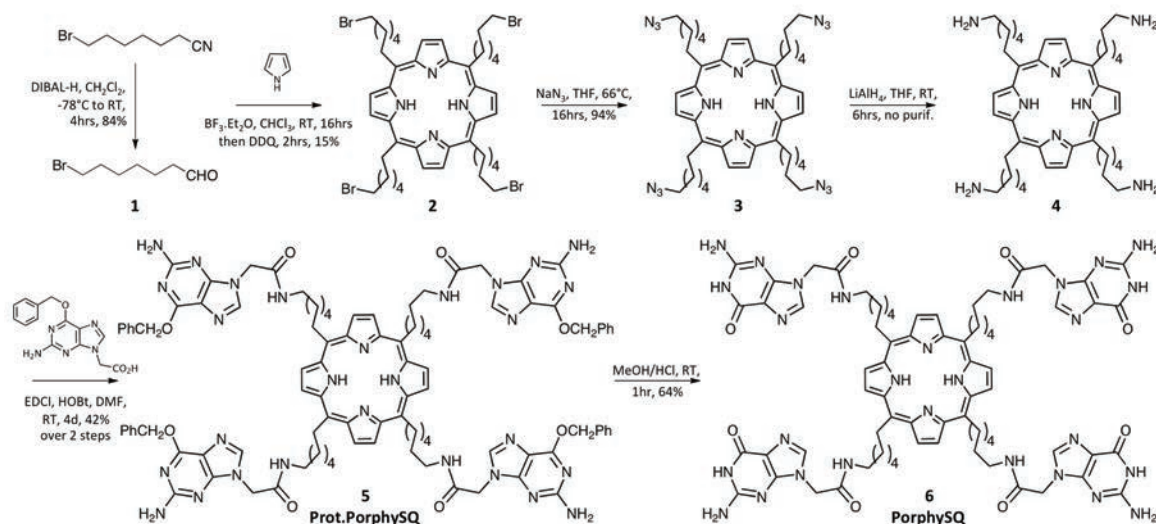


Figure 90 – Schéma réactionnel de synthèse du PorphSQ

Brièvement, le produit résultant de la réduction du 7-bromoheptanitrile **1** par le DIBAL-H, obtenu avec un rendement de 84%, est condensé sur du pyrrole en présence de quantités catalytiques de trifluorure de bore étherate. L'intermédiaire obtenu est alors oxydé *in situ* en présence de DDQ, permettant d'obtenir la porphyrine tétra-(6-bromohexyle) **2** avec un rendement satisfaisant pour ce type de réaction (15,5%). Ce composé est alors traité par de l'azoture de sodium conduisant à l'azoture **3** correspondant, directement réduit par LiAlH_4 pour former la porphyrine tétra-(6-aminohexyle) **4**. La fonctionnalisation des amines primaires se fait par couplage peptidique classique avec l'acide 2-(2-amino-(6-benzyloxy)-9H-purin-9-yl)acétique utilisé pour la synthèse du DOTASQ- C_1 en présence d'EDCI et de HOBt. Les groupements benzyloxy du **Prot.PorphSQ** obtenu avec un rendement de 42% sont déprotégés en présence de méthanol saturé en acide chlorhydrique pour conduire au composé cible **PorphSQ** avec un haut degré de pureté et un bon rendement chimique (64%).

IX.3 Etudes biophysiques

IX.3.A Evaluation du gain d'affinité du PorphSQ par rapport aux DOTASQ

Les propriétés d'interaction du PorphSQ en tant que ligand de quadruplexes ont été évaluées au laboratoire par des expériences de FRET-melting.

Une première série de mesures a été réalisée sur le quadruplexe F21T, de séquence identique à celle du simple brin télomérique humain, en présence ou non de 25 équivalents de ligands (Figure

91A. *N.B.* : à 5éq, la valeur $\Delta T_{1/2}$ mesurée pour le PorphySQ étant faible et par conséquent peu exploitable, le choix de se placer à 25éq a été fait afin d'obtenir une valeur significative et reproductible).

Après avoir mis en place tout le protocole expérimental en plaque 96 puits pour l'évaluation par FRET-melting au sein du laboratoire de l'ICMUB, les premiers résultats obtenus sur le DOTASQ-C₅, utilisés ici à titre de comparaison, diffèrent légèrement de ceux obtenus précédemment à l'IECB avec une stabilisation de 1,7°C. L'utilisation du PorphySQ s'avère néanmoins plus efficace avec une $\Delta T_{1/2} = 6,5^\circ\text{C}$. Afin de vérifier que la modification du *template* confère toujours à la molécule le même type d'interaction *via* la tétrade de guanines synthétique intramoléculaire, plusieurs mesures de contrôle ont été effectuées dans les mêmes conditions expérimentales.

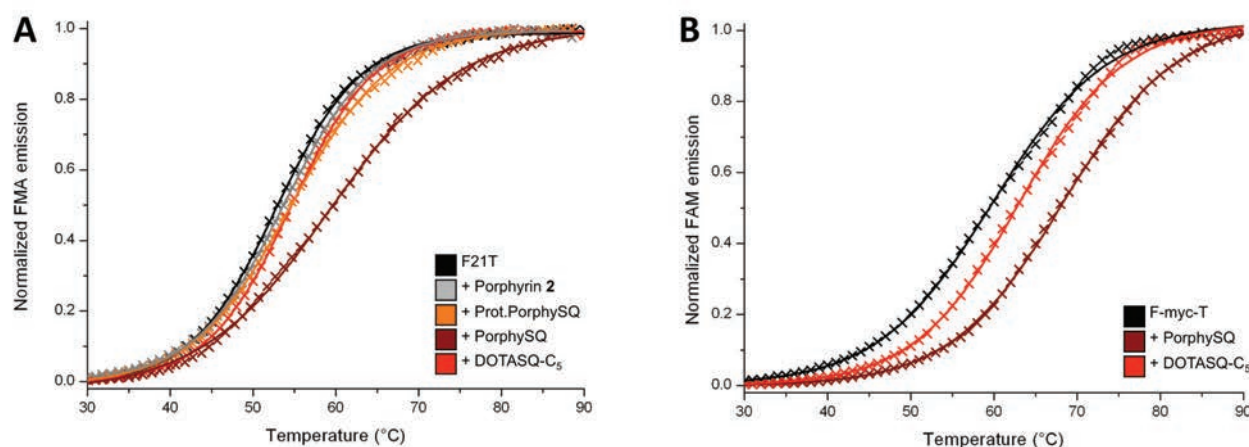


Figure 91 – Résultats de FRET-melting obtenus pour la séquence télomérique humaine F21T (A) et pour une région oncogène (F-myc-T, B) à 25 équivalents de ligand.

Dans un premier temps, pour s'assurer que la stabilisation ne provient pas uniquement du motif porphyrinique, les courbes obtenues sur la porphyrine neutre 2 mettent en évidence qu'en l'absence de bras guanines l'affinité pour les quadruplexes est très faible ($\Delta T_{1/2}$ de seulement 1,6°C).

Dans un second temps, l'utilisation de la molécule Prot.PorphySQ, semblable au PorphySQ mais dont les guanines sont protégées par un groupement benzyloxy ne montre également qu'une très faible stabilisation de 0,8°C. Cette protection du carbonyle de l'amide empêche la formation de la tétrade de guanines en supprimant les liaisons hydrogènes, ce qui explique la chute drastique du $\Delta T_{1/2}$, consolidant l'hypothèse selon laquelle les propriétés d'interaction du PorphySQ proviennent de sa capacité à former une tétrade de guanines intramoléculaire.

Une nouvelle série d'expériences de FRET-melting est réalisée sur la séquence F-myc-T (FAM-d(GAG₃TG₄AG₃TG₄A₂G)-TAMRA) (Figure 91B), présente au sein du génome humain au

niveau de la région promotrice de l'oncogène C-myc, et confirme la meilleure affinité du PorphySQ pour les quadruplexes d'ADN par rapport au DOTASQ-C₅, avec respectivement des $\Delta T_{1/2}$ de 8,6°C contre 3,0°C dans ces conditions expérimentales. De par sa conformation parallèle, le F-myc-T possède des tétrades extérieures beaucoup plus accessibles que celles du F21T, de conformation hybride, et expliquent sa meilleure stabilisation par les ligands.

Ces expériences confirment notre approche de ligands capables d'interagir avec les quadruplexes d'ADN par une reconnaissance entre tétrades synthétiques et naturelles. Comme attendu, l'utilisation du cycle porphyrinique comme socle pour la conception d'un nouveau type de TASQ permet de promouvoir l'auto-assemblage des guanines et augmente substantiellement l'affinité pour les quadruplexes d'ADN.

IX.3.B Evaluation de la sélectivité du PorphySQ

Toujours en se basant sur la méthode de FRET-melting, des quantités croissantes de ligands compétiteurs non marqués sont ajoutées aux solutions contenant à la fois F21T et 25 équivalents de PorphySQ (Figure 92).

L'utilisation de duplexe compétiteur ds26 n'influe guère sur la température de fusion qui évolue de 6,5°C à 6,6°C puis 5,8°C à 0, 15 et 50 équivalents de motif duplexe, respectivement. Ceci nous permet de déduire la très bonne sélectivité du PorphySQ pour les quadruplexes d'ADN, avec un facteur de sélectivité $S > 90\%$, ainsi que son inaptitude à interagir avec la double hélice d'ADN.

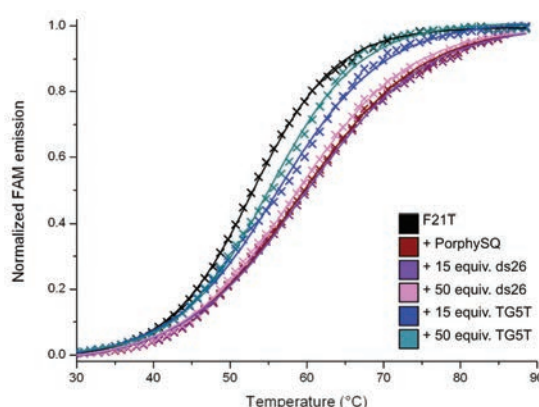


Figure 92 - Résultats de FRET-melting obtenus en présence de compétiteurs ds26 et TG5T (correspondant à $[TG_5T]_4$) sur la stabilisation de F21T par 25 équivalents de PorphySQ.

Pour confirmer le mode d'interaction du PorphySQ *via* la tétrade accessible d'un quadruplexe d'ADN, des quantités croissantes de quadruplexe tétramoléculaire $[TG_5T]_4$ non marqué sont ajoutées au système F21T plus 25 équivalents de ligand. Comme nous l'avons vu précédemment,

les tétrades extérieures de cet édifice sont ici extrêmement accessibles puisque dépourvues de boucles, contrairement au F21T. Ainsi, en présence des deux quadruplexes, le PorphySQ interagit préférentiellement avec $[TG_5T]_4$ tétramoléculaire induisant une diminution de la stabilisation de F21T de 6,5°C à 3,5°C puis 2,3°C à 0, 15 et 50 équivalents, respectivement.

Tous ces arguments confirment donc que la reconnaissance entre le PorphySQ et les quadruplexes se fait bien par une reconnaissance biomimétique entre les tétrades naturelles et synthétiques.

IX.4 Conclusion

Après le premier prototype DOTASQ, nous avons cherché à optimiser encore l'architecture des TASQ hydrosolubles en incorporant un *template* capable de promouvoir la formation de la tétrade intramoléculaire. L'utilisation d'un cœur porphyrinique en lieu et place du macrocycle DOTA le permet, grâce à des interactions de π -stacking identiques à celles rencontrées lors des interactions entre de nombreux ligands de la littérature et les tétrades accessible des quadruplexes. Cette modification se traduit par l'augmentation de l'affinité de ces molécules pour les quadruplexes dont la stabilisation s'avère meilleure tout en conservant l'excellente sélectivité du modèle initial.

Néanmoins, la présence d'une porphyrine a fortement réduit l'hydrosolubilité de la molécule et constitue un frein important quant à l'utilisation de ce *template*. La substitution des bras alkyles par des bras PEG a été envisagée mais la synthèse n'a pour le moment pas abouti. Une autre série de PorphySQ renfermant des amines pendantes protonables a récemment été testée et sera discutée en conclusion générale de ce chapitre (voir XI).

Nous sommes alors reparti du DOTASQ au sein duquel cette fois-ci, la nature des bras a été modifiée.

X. Le ^{PNA}DOTASQ comme meilleur TASQ hydrosoluble

La modulation de la nature des bras du DOTASQ doit permettre d'optimiser la formation intramoléculaire de la tétrade de guanines sans perdre les bonnes propriétés d'hydrosolubilité.

Ainsi, notre choix s'est porté sur l'utilisation de bras PNA-guanines (pour *peptide nucleic acid*) en lieu et place des bras alkyles-guanines en raison de trois avantages principaux :

- les interactions entre séquences PNA et ADN (ou ARN) sont connues et s'avèrent particulièrement fortes.^[360]

- l'introduction d'une charge cationique par bras par le biais de la fonction amine primaire protonée en conditions physiologiques (pH 7,2) constituerait une couronne de charges positives autour de la tétrade synthétique, accroissant l'affinité du ligand pour un quadruplexe naturel *via* des interactions électrostatiques avec les nombreux groupements phosphates chargés négativement. Cet apport de charge devrait également augmenter l'hydrosolubilité globale mais pourrait aussi se traduire par une perte de sélectivité à cause d'interactions aléatoires avec d'autres structures d'ADN ou d'ARN comme le duplexe.

- les PNA font l'objet de nombreuses applications *in vitro* prometteuses pour le diagnostic génétique et également comme drogues pour le développement de stratégies antisens visant l'ARNm ou les miARN. Cet emploi des PNA est en effet possible grâce à leur excellente résistance aux conditions intracellulaires (enzymes de dégradation, etc.).^[360]

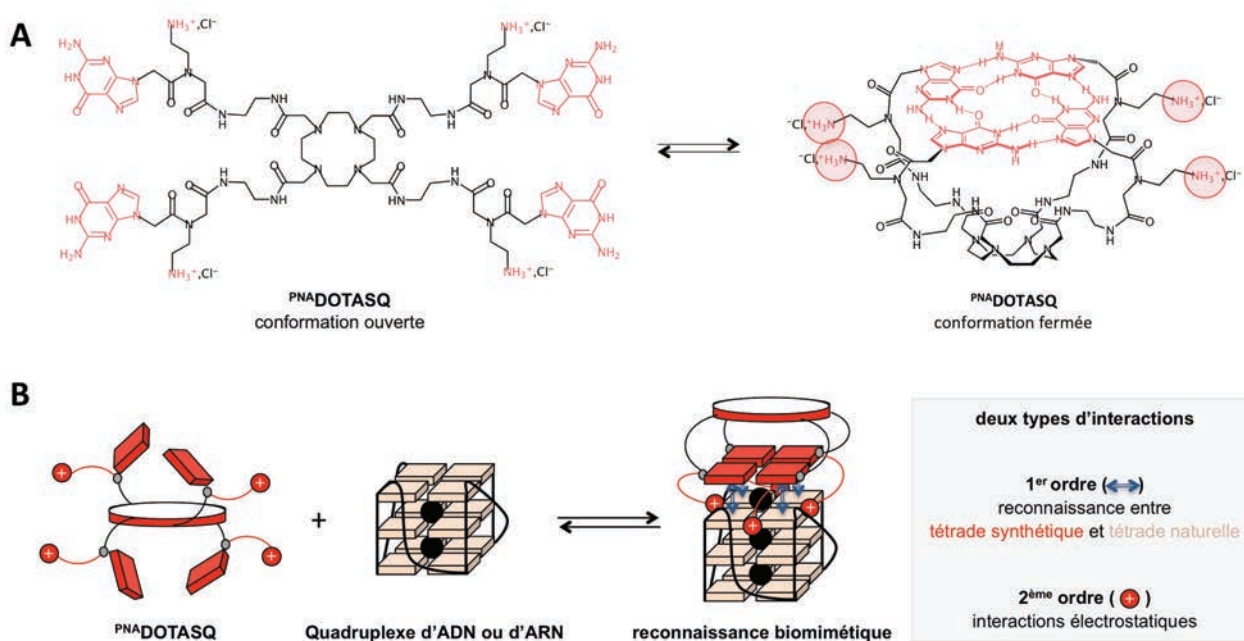


Figure 93 – Structure et équilibre conformationnel du ^{PNA}DOTASQ (A). Principe de reconnaissance biomimétique du ^{PNA}DOTASQ selon deux modes d'interactions complémentaires (B).

Le ^{PNA}DOTASQ (Figure 93A) devrait donc offrir deux niveaux d'interactions complémentaires : la reconnaissance entre tétrades par π -stacking (principe des TASQ), accompagnée d'interactions électrostatiques entre les charges négatives de l'ADN (ou de l'ARN) et les quatre charges positives des amines pendantes des PNA (Figure 93B). Nous nous attendons donc à une augmentation drastique de l'affinité et de la sélectivité des ^{PNA}DOTASQ pour leurs cibles quadruplexes dont les résultats sont présentés ci-après.

X.1 Synthèse du ^{PNA}DOTASQ

La synthèse du ^{PNA}DOTASQ s'effectue en quatre étapes convergentes offrant chacune de très bons rendements, supérieurs à 89% (Figure 94). L'élaboration du *template* est identique à celle du DOTASQ et consiste en la condensation du **DOTA-Et** en présence d'éthylène diamine. Parallèlement, le **Boc-PNA-G(Z)-OH** commercial est déprotégé au niveau du groupement protecteur carboxybenzyle (noté Cbz ou Z) par hydrogénation catalytique au Pd/C.

Le couplage entre le cœur DOTA et les bras PNA se fait en conditions classiques HOBt et EDCI (déjà utilisées pour le DOTASQ et le PorphyrSQ) pour conduire au **Prot.^{PNA}DOTASQ** dont le groupement Boc est finalement éliminé par un traitement au méthanol saturé en acide chlorhydrique afin d'obtenir le ^{PNA}DOTASQ cible.

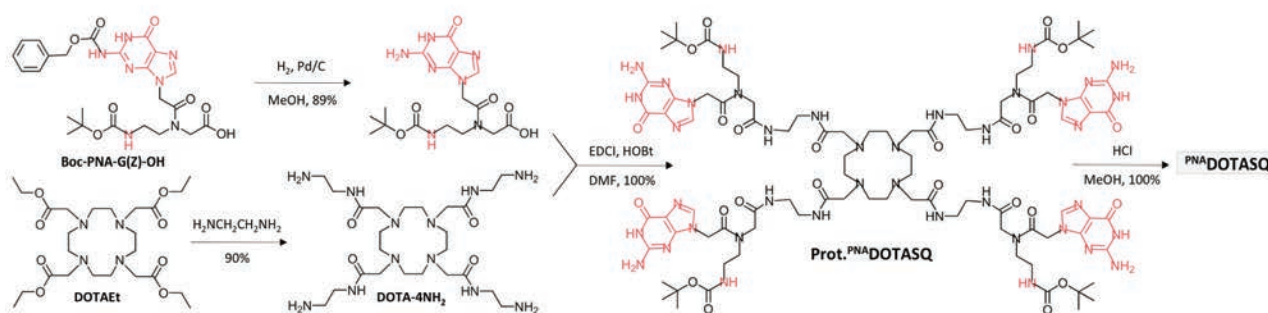


Figure 94 – Schéma réactionnel de synthèse du ^{PNA}DOTASQ

Afin de s'assurer de l'hydrosolubilité de ce composé, propriété clef pour les applications à venir, des mesures d'absorbance sont effectuées par spectroscopie UV-visible sur une gamme de concentrations permettant de conclure que dans l'eau, le ^{PNA}DOTASQ est soluble au-delà de 1,5mM et suit la loi de Beer-Lambert.^[18]

X.2 Etudes par RMN ¹H

Comme pour le DOTASQ, nous avons dans un premier temps évalué le mode d'interaction du ^{PNA}DOTASQ par des analyses par RMN ¹H, menées cette fois-ci au sein du laboratoire.

En nous inspirant des travaux de Y. Kato *et al.* de 2005,^[361] la séquence [T₂AG₃T]₄ a été choisie pour son absence de boucles et sa symétrie S₄ qui permettent d'obtenir un spectre relativement simple à exploiter grâce à l'équivalence des quatre protons amides d'une même tétrade, conduisant à la présence de trois signaux à 10,90ppm, 11,11ppm et 11,53ppm. Il est intéressant de noter que fréquemment les protons de la tétrade centrale sont moins visibles sur le spectre que les autres, et notre expérience ne déroge pas à la règle (Figure 95).

L'ajout de quantités croissantes de PNA DOTASQ, de 0,5 à 2,0 éq, entraîne la diminution croissante de l'intensité des pics correspondants aux protons du quadruplexe au profit de l'apparition de trois nouveaux signaux (flèches rouges, Figure 95), plus blindés, correspondant aux protons des tétrades de l'ADN quadruplexe en interaction avec le ligand.^[335, 362] Ces modifications témoignent donc d'une interaction réelle entre le PNA DOTASQ et le quadruplexe d'ADN tétramoléculaire et soutiennent notre approche. Notons la disparition quasi complète de l'ADN seul en solution à 2 éq et surtout l'apparition progressive d'un pic centré à 11,43 ppm (flèche rouge avec étoile, Figure 95) qui, au regard des résultats obtenus sur le DOTASQ et sur le RAFT-G4 par l'équipe de E. Defrancq,^[22] peut être attribué aux protons des amides de la tétrade de guanines intramoléculaire du PNA DOTASQ. Ces résultats soutiennent donc notre hypothèse d'une reconnaissance biomimétique, puisque le TASQ semble refermer sa tétrade pour interagir avec le quadruplexe.

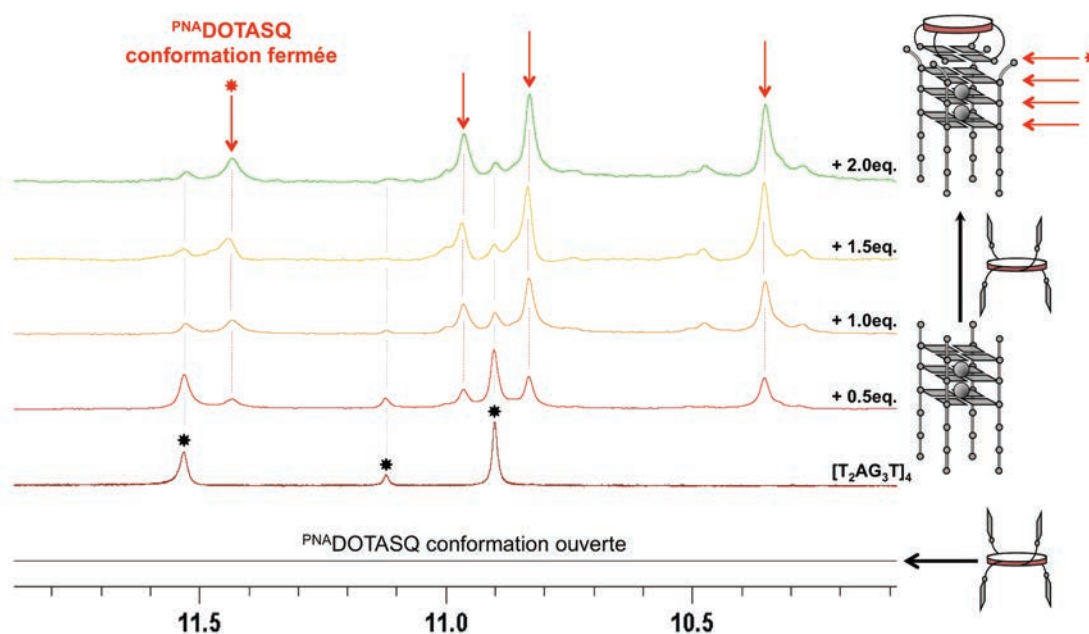


Figure 95 – Spectres obtenus par RMN ^1H dans une solution de Caco.K 10mM (50mM K^+ , pH 7,2) à 10% D_2O pour le PNA DOTASQ seul (en noir), $[\text{T}_2\text{AG}_3\text{T}]_4$ seul (en marron) ou après ajout de quantités croissantes de PNA DOTASQ (de 0,5 à 2 équivalents).

Contrairement aux DOTASQ (voir VIII.2), aucun signal dans la région de 11 à 13 ppm n'a pu être mis en évidence pour le PNA DOTASQ seul en solution à différentes températures. La présence de fonctions amines chargées peut être un frein au déplacement de l'équilibre de la conformation ouverte vers la conformation fermée à cause d'effets de répulsions électroniques.

Cette observation tend à montrer que ce nouveau ligand ne referme sa tétrade qu'au contact d'un quadruplexe par reconnaissance biomimétique entre tétrades, favorisée par des interactions

électrostatiques entre le ruban phosphate et les amines des PNA. Ce type d'interaction original jamais décrit pour un ligand de quadruplexe permet de qualifier le ^{PNA}DOTASQ par le terme anglo-saxon de « smart ligand ».

Une mesure contrôle sur la même séquence mais en présence de la porphyrine TMPyP4 fait apparaître un déplacement des pics vers les hauts champs, preuve de l'interaction molécule/ADN, sans jamais obtenir un signal vers 11,4ppm, confirmant notre proposition d'attribution de ce pic aux protons amides de la tétrade synthétique du ^{PNA}DOTASQ.^[18]

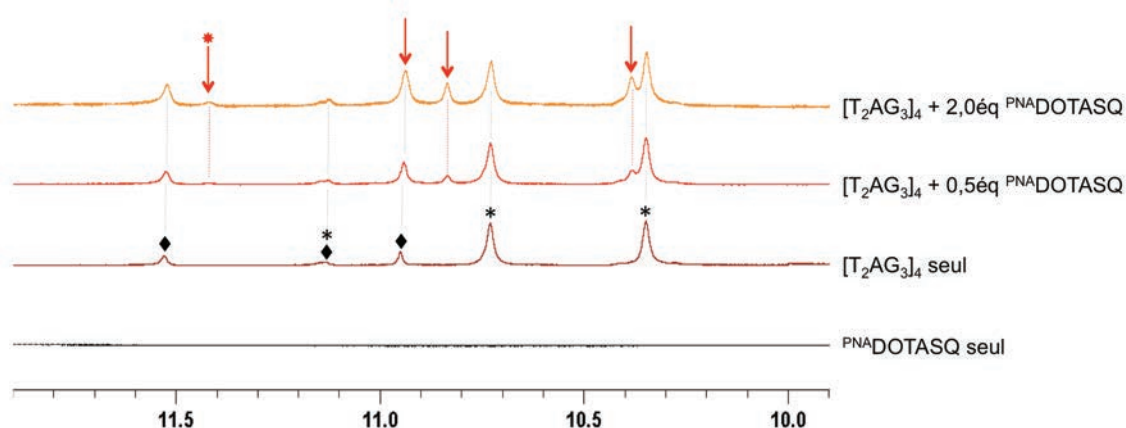


Figure 96 - Spectres obtenus par RMN ¹H dans une solution de Caco.K 10mM (10mM K⁺, pH 7,2) à 10% D₂O pour le ^{PNA}DOTASQ seul (en noir), [T₂AG₃]₄ seul (en marron) ou après ajout de quantités croissantes de ^{PNA}DOTASQ (à 0,5 et 2 équivalents). Les signaux correspondants aux formes monomériques et dimériques sont symbolisés par les losanges et les étoiles, respectivement.

Pour confirmer ces résultats, une seconde étude est menée sur un autre quadruplexe tétramoléculaire dépourvu de thymine en position 3', [T₂AG₃]₄, qui en conditions riches en K⁺ (300mM) va majoritairement se présenter sous une forme dimérique issue d'interactions de π -stacking entre les deux tétrades externes de deux quadruplexes. A des concentrations plus faibles en K⁺ (10mM, Figure 96), le quadruplexe est présent à la fois sous forme dimérique et monomérique.^[361, 363, 364] L'ajout de quantités croissantes de ^{PNA}DOTASQ (Figure 96) conduit à l'apparition de trois nouveaux signaux correspondant aux trois tétrades du quadruplexe monomérique en interaction avec le ligand auxquels s'ajoute le signal caractéristique de la tétrade synthétique intramoléculaire du ^{PNA}DOTASQ à 11,4ppm.

Une dernière expérience d'évaluation de la sélectivité du TASQ est réalisée sur le duplexe ds12, présentant un centre de symétrie au milieu de sa séquence (Figure 97B). Son spectre RMN ¹H caractéristique comprend alors cinq signaux correspondant aux protons amides des thymine impliquées dans les paires de bases T-A (à 13,67 et 13,51ppm) et des guanines mises en jeu dans les paires G-C (à 12,95, 12,78 et 12,58ppm, Figure 97A).

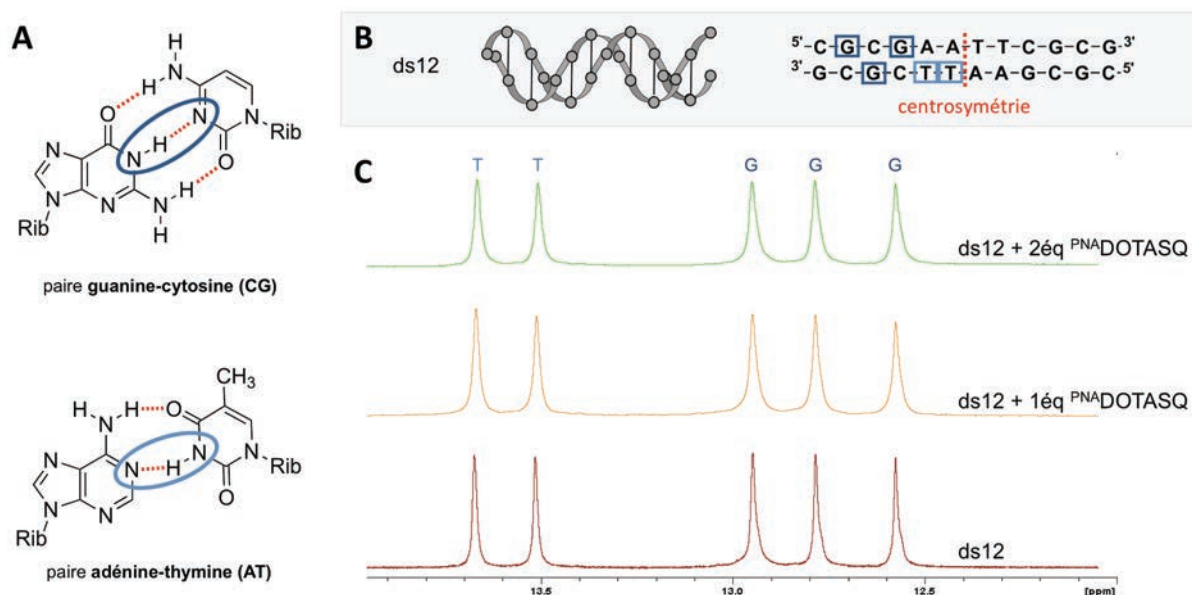


Figure 97 – Représentation des protons amides des paires de bases CG et AT visibles au delà de 12ppm en RMN ^1H (A) et représentation schématique du duplexe ds12 utilisé pour l'étude (B). Spectres obtenus par RMN ^1H dans une solution de Caco.K 50mM (300mM K^+ , pH 7,2) à 10% D_2O pour le duplexe ds12 seul ou en présence de $^{\text{PNA}}$ DOTASQ (C).

Classiquement, les bases G-C aux extrémités ne sont pas appariées, expliquant leur absence sur le spectre.^[365] L'ajout de quantités croissantes de $^{\text{PNA}}$ DOTASQ n'entraîne aucune modification et constitue par conséquent une première preuve de leur sélectivité (Figure 97C).

X.3 Analyses par spectrométrie de masse ESI-MS

La spectrométrie de masse par ionisation électrospray (ESI-MS) repose sur une ionisation douce limitant la fragmentation des échantillons analysés et permettant d'accéder directement au pic moléculaire. Elle est largement employée pour l'analyse de biomolécules mais également pour des études d'interactions non covalentes. Ainsi, son adaptation aux quadruplexes d'ADN permet de déterminer la présence ou non d'une interaction entre un ligand potentiel et l'oligonucléotide, mais permet également d'évaluer sa stœchiométrie ainsi que, par un modèle mathématique, ses constantes d'affinités.^[366, 367]

Les résultats, bien que préliminaires, de l'ajout de 2éq de $^{\text{PNA}}$ DOTASQ sur le quadruplexe tétramoléculaire $[\text{TG}_4\text{T}]_4$ (Figure 98A) montrent la présence d'un pic intense de l'ADN seul ($z=5$ et $z=4$), et d'un pic beaucoup plus faible ($z=5$) correspond à l'ADN + 1éq de $^{\text{PNA}}$ DOTASQ. Bien évidemment, les conditions expérimentales sont ici loin d'être optimales mais permettent toutefois de manière qualitative de confirmer l'existence d'une interaction entre notre nouveau TASQ et la structure quadruplexe.

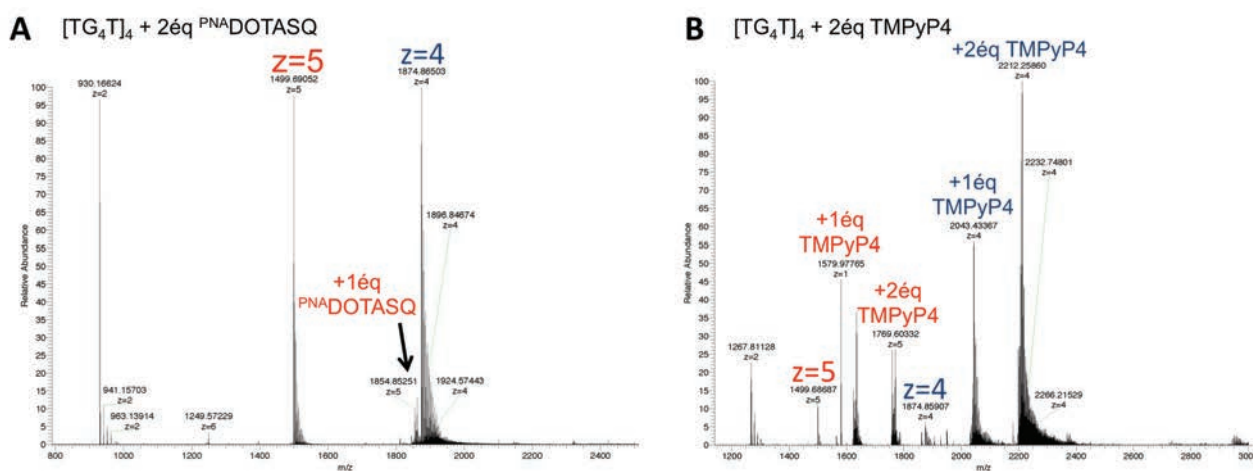


Figure 98 – Spectres obtenus par ESI-MS (ionisation négative) avec $[TG_4T]_4$ en présence de 2éq de $PNA DOTASQ$ (A) ou de $TMPyP4$ (B), dans un tampon acétate d’ammonium (100mM, pH 7,1).

Parallèlement, en utilisant les mêmes conditions expérimentales mais cette fois-ci en présence de la molécule contrôle $TMPyP4$ (Figure 98B), il est clairement observé la présence majoritaire de l’ADN en interaction avec le ligand avec une stœchiométrie de 1:1 et 1:2. Ceci nous permet dans un sens de valider notre protocole expérimental mais révèle aussi que l’affinité du $PNA DOTASQ$ est moindre comparée à celle de la $TMPyP4$ qui interagit suffisamment fortement pour ne pas se « décrocher » de l’ADN lors de l’ionisation.

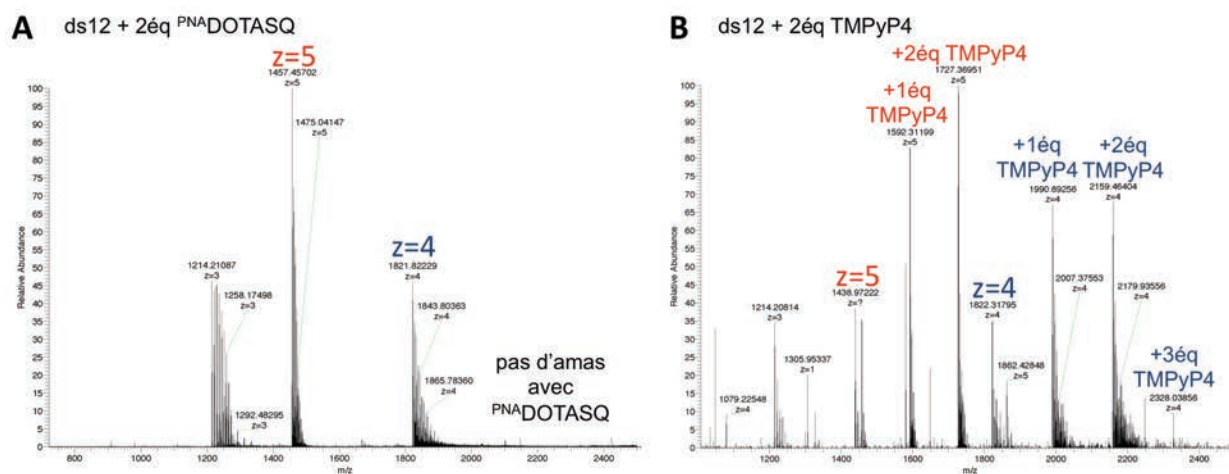


Figure 99 - Spectres obtenus par ESI-MS (ionisation négative) avec le duplexe $ds12$ en présence de 2éq de $PNA DOTASQ$ (A) ou de $TMPyP4$ (B), dans un tampon acétate d’ammonium (100mM, pH 7,1).

Comme pour les études par RMN, le duplexe $ds12$ est également utilisé et les spectrogrammes obtenus mettent en évidence que le $PNA DOTASQ$ n’interagit pas avec cette structure (Figure 99A) et s’avère donc sélectif des quadruplexes d’ADN alors que la $TMPyP4$, elle, se lie très largement et ce, jusqu’à trois ligands par duplexe (Figure 99B).

Nous sommes toutefois conscients que ces résultats obtenus par ESI-MS sont encore préliminaires et optimisables, mais ils permettent néanmoins de mettre en évidence l’existence

d'une interaction sélective du PNA DOTASQ avec le quadruplexe d'ADN, et confirment les observations de RMN 1H .

Afin de valider ces observations, nous nous sommes tournés vers deux méthodes classiques d'analyse biophysique par fluorescence que sont le FID (isotherme) et le FRET-melting.

X.4 Expériences de FID (*fluorescence intercalator displacement*)

La méthode de FID (ou *fluorescence intercalators displacement*) repose sur l'utilisation d'une sonde dite *light-up*^[368] qui, seule en solution, est idéalement peu ou pas fluorescente alors qu'elle le devient une fois en interaction avec une structure d'ADN (Figure 100A). Dans le cas des quadruplexes, les cyanines TO (thiazole orange, Figure 100A) et TO-PRO3 (Figure 100B) sont les plus largement utilisées et dans notre cas, les études ont été menées avec cette dernière ($\lambda_{ex}=630nm$).

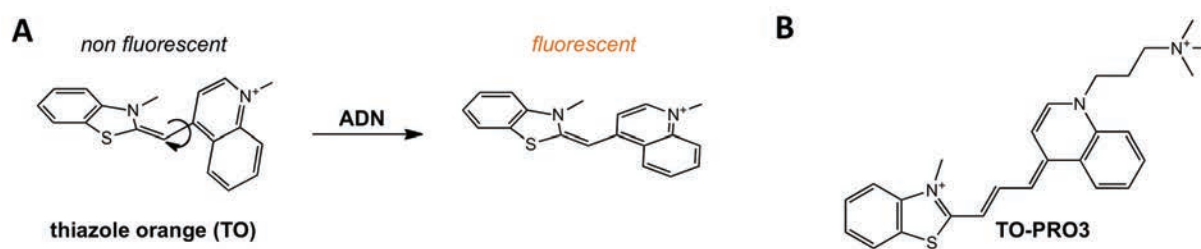


Figure 100 – Structure chimique et équilibre conformationnel du thiazole orange en présence d'ADN (A), et structure chimique du TO-PRO3 (B).

Sous leur forme libre, leur faible fluorescence provient de la libre rotation entre les deux hétérocycles ; une fois en interaction avec l'ADN, la conformation se fige sous une forme coplanaire qui rétablit l'aromaticité du système et par conséquent régénère la fluorescence de la sonde.^[369-372]

Lorsque deux équivalents de TO-PRO3 sont mis en présence d'un équivalent de quadruplexe d'ADN tel que le 22AG (séquence télomérique humaine), l'émission mesurée est maximale (Figure 101A, courbe noire). L'ajout de quantités croissantes de PNA DOTASQ compétiteur chasse la sonde de la tétrade accessible du quadruplexe qui se retrouve par conséquent dans le milieu (où sa conformation redevient libre), et entraîne une chute de la fluorescence globale. Cette expérience démontre que le PNA DOTASQ a donc une affinité pour le quadruplexe avec lequel il interagit au niveau de la tétrade accessible, ce qui conforte le principe d'interaction biomimétique.

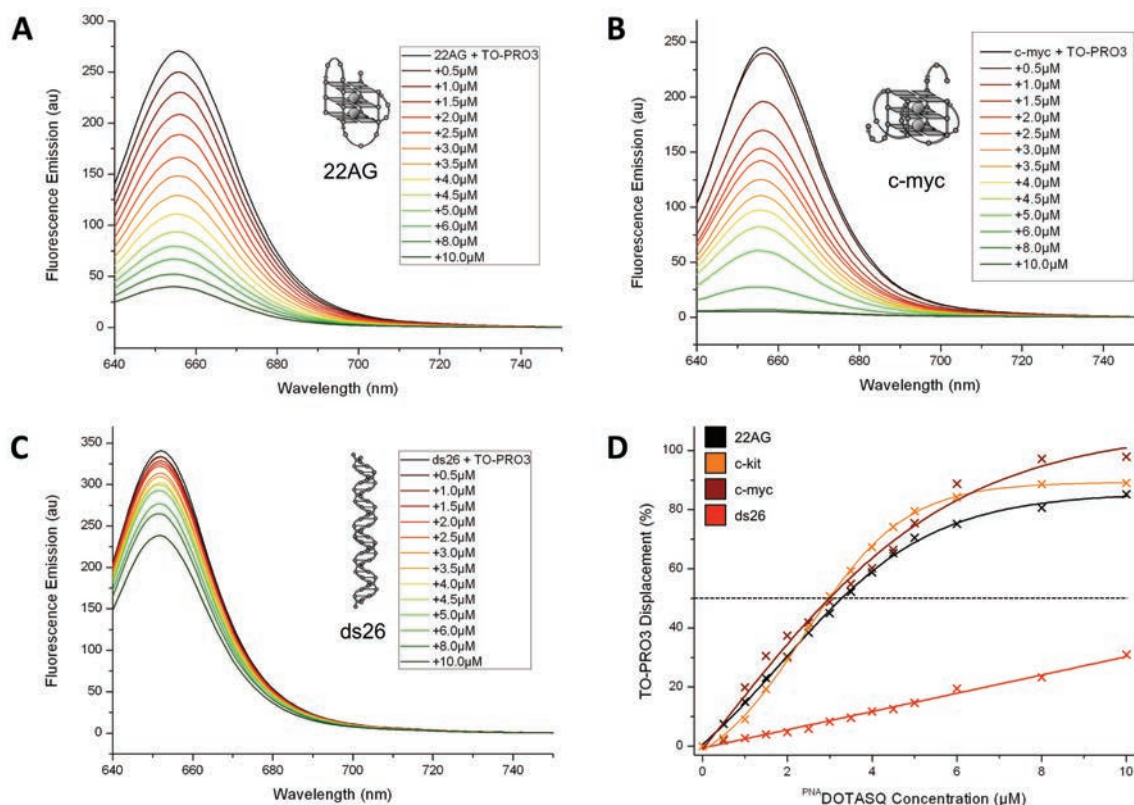


Figure 101 – Résultats des expériences de FID réalisées avec le TO-PRO3 en présence de 22AG (A), c-myc (B) et ds26 (C) et de quantités croissantes de PNA DOTASQ (de 0,5 à 10 μM, [ADN]=2 μM). Evolution du déplacement de la sonde fluorescente TO-PRO3 en fonction de la concentration en PNA DOTASQ (D).

Des expériences similaires sont réalisées sur les séquences oncogéniques C-myc (Figure 101B) et C-kit dont les résultats sont regroupés sur le graphique Figure 101D où le déplacement de la sonde (en %) est reporté en fonction de la concentration de PNA DOTASQ selon la formule :

$$\text{TO - PRO3 displacement (\%)} = 100 - 100 \times \frac{F}{F_0}$$

avec F fluorescence maximale mesurée pour chaque courbe et F_0 , fluorescence de l'ADN en présence de TO-PRO3 (avant ajout de ligand compétiteur).^[370]

Les valeurs clefs DC_{50} , correspondant à la concentration à laquelle la moitié de la sonde a été déplacée, sont d'environ 3 μM pour les trois structures quadruplexes (ou plus exactement 3,2 μM pour le 22AG et 2,9 μM pour C-kit et C-myc).

L'évaluation de la sélectivité par la méthode de FID se fait par l'utilisation de 3éq de TO-PRO3 en présence d'un équivalent du duplexe ds26. L'ajout de quantités croissantes de PNA DOTASQ (Figure 101C et D) entraîne une diminution moindre dont la DC_{50} est évaluée par extrapolation à environ 15 μM, soit cinq fois plus que pour les quadruplexes, démontrant donc une bonne sélectivité de notre composé. Notons également que l'aspect linéaire de la courbe obtenue pour le duplexe

ds26 témoigne d'une interaction non spécifique aléatoire probablement de nature électrostatique due à l'interaction des phosphates de l'ADN avec les amines primaires des bras PNA.

X.5 Etude du ^{PNA}DOTASQ par FRET-melting

Comme pour le DOTASQ et le PorphyrSQ, nous avons effectué tout un panel de mesures par FRET-melting qui est une méthode de choix pour l'évaluation des paramètres d'affinité et de sélectivité des ligands pour une structure non-usuelle oligonucléotidique. Dans un premier temps, nous nous sommes intéressés à l'étude d'interaction sur l'ADN puis à la fin de cette partie, nous présenterons l'adaptation de cette méthode aux quadruplexes d'ARN.

X.5.A Interaction du ^{PNA}DOTASQ avec les quadruplexes d'ADN

Classiquement, nous avons mesuré l'impact de l'ajout de concentrations croissantes de ^{PNA}DOTASQ sur la stabilisation des structures F21T (séquence télomérique humaine, Figure 102A), C-myc et C-kit (séquences oncogéniques).

Les résultats montrent une forte stabilisation de ces trois séquences avec, pour 5 équivalents de ligand, des valeurs $\Delta T_{1/2}$ de 12,5°C, 12,3°C et 15,2°C pour F21T, C-myc et C-kit, respectivement (Figure 102C), faisant de cette molécule le meilleur modèle de TASQ hydrosoluble développé à ce jour. Alors que la stabilisation atteint plus de 21°C à 15éq sur F21T, le résultat le plus important est sans nul doute l'excellente sélectivité de ce composé qui présente une stabilisation sur le duplexe F-duplex-T, double hélice doublement marquée de 10 paires de bases renfermant une *hairpin* de six thymines, pratiquement nulle avec un $\Delta T_{1/2}$ ne dépassant pas 0,5°C à 5 équivalents (Figure 102B et C).

Cette forte sélectivité est confirmée par des expériences supplémentaires mettant en jeu F21T en présence de 5éq de ^{PNA}DOTASQ plus 15 ou 50éq de duplexe ds26 non marqué compétiteur et permettent d'obtenir respectivement les facteurs de sélectivité $^{FRET}S=84\%$ et 74% (Figure 102D).

Ces données confirment à leur tour les fortes affinités et sélectivités du ^{PNA}DOTASQ pour les quadruplexes d'ADN et ce, quelles que soient leurs structures. Néanmoins, pour valider que son mode d'interaction biomimétique nécessite à la fois la présence des quatre guanines et des fonctions amines pendantes des bras PNA qui permettent à la fois une reconnaissance spécifique tétrade/tétrade et une interaction électrostatique supplémentaire jouant le rôle d' « agrafe », deux composés contrôles ont été utilisés.

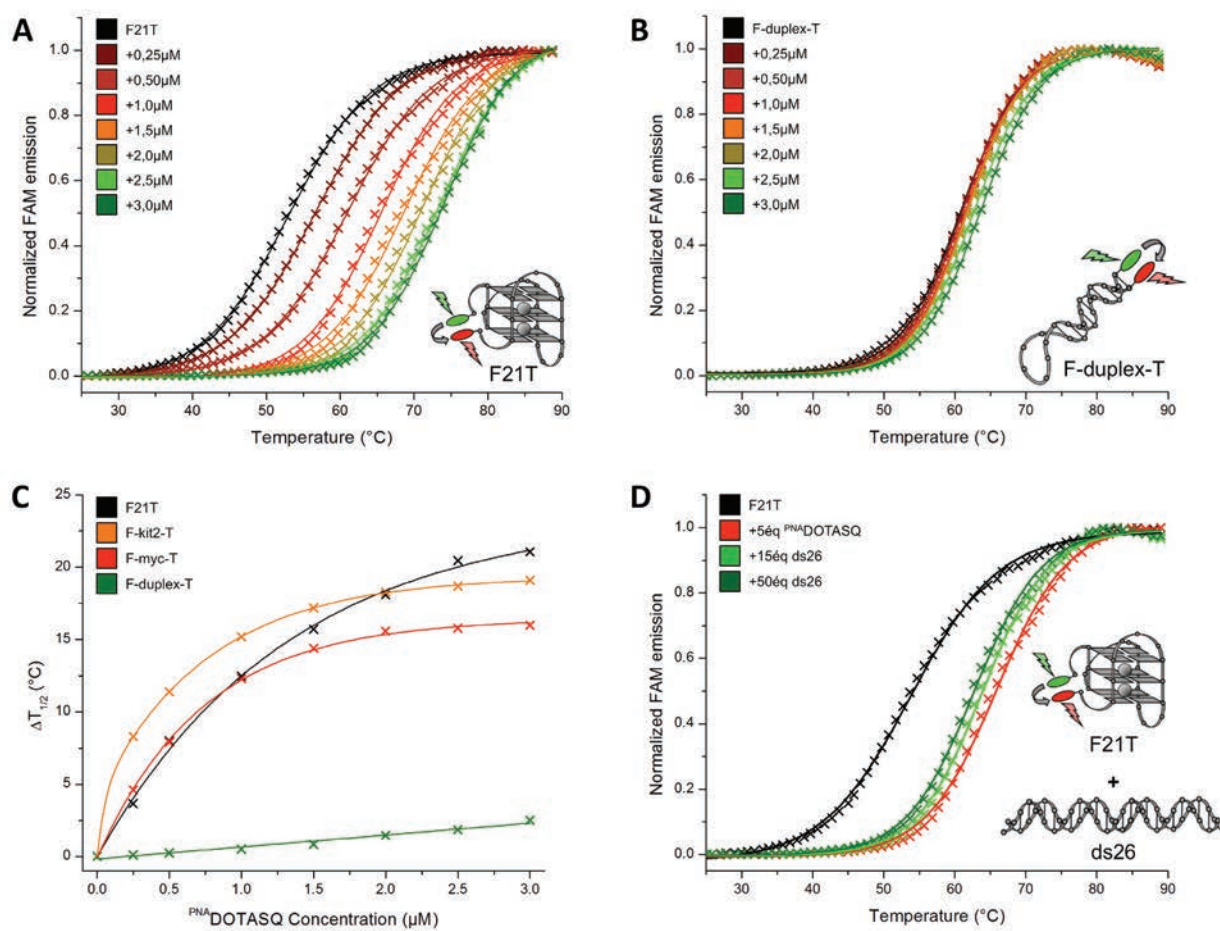


Figure 102 – Courbes de FRET-melting obtenues sur F21T (A) et F-duplex-T (B) en présence de quantités croissantes de PNA DOTASQ (de 0,25 μM à 3,0 μM, [ADN]=0,2 μM). Variation du $\Delta T_{1/2}$ en fonction de la concentration en PNA DOTASQ (C) et résultats de la compétition avec le duplexe ds26 (D).

Le premier de ces composés contrôles, le Prot. PNA DOTASQ (Figure 94) est capable de former une tétrade de guanine mais est dépourvu des fonctions amines, protégées par un groupement *boc* alors que le second, le DOTA-4NH₂ (Figure 94) ne possède pas de guanine mais offre quatre amines primaires terminales protonables.

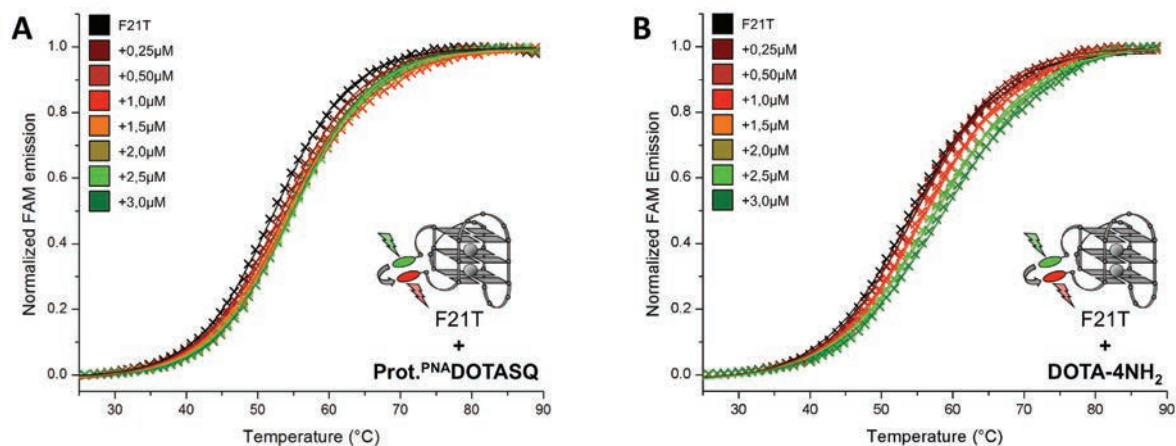


Figure 103 – Résultats de FRET-melting obtenus pour une gamme de concentration des contrôles négatifs Prot. PNA DOTASQ (A) et DOTA-4NH₂ (B) sur F21T (0,2 μM).

Les gammes effectuées sur F21T révèlent une stabilisation très faible par le Prot.^{PNA}DOTASQ (Figure 103A) qui, en l'absence des charges cationiques, s'apparente au premier modèle DOTASQ avec une température de stabilisation proche de $\Delta T_{1/2}=2,2^{\circ}\text{C}$ à 5éq, six fois plus faible que celle du ^{PNA}DOTASQ. Ceci permet de conforter l'hypothèse précédemment émise selon laquelle la simple présence de quatre guanines au sein d'une même molécule n'est pas suffisante à la bonne formation de la tétrade intramoléculaire. La suppression des guanines sur le DOTA-4NH₂ induit également une stabilisation extrêmement faible de $\Delta T_{1/2}=1,9^{\circ}\text{C}$ à 5éq (Figure 103B), preuve que la simple présence de charges n'est pas suffisante pour avoir un ligand performant.

Le ^{PNA}DOTASQ profite donc d'un effet synergique entre ces deux constituants majeurs de sa structure qui, pris séparément, stabilisent le quadruplexe à hauteur de 2°C seulement, alors que la présence simultanée des deux augmente drastiquement cette valeur à $12-15^{\circ}\text{C}$.

X.5.B Influence de l'insertion d'un terbium : le Tb.^{PNA}DOTASQ

Alors que le DOTASQ offrait de faibles effets sur la stabilisation des quadruplexes à cause, entre autres, d'une trop grande flexibilité des bras induite par un polyazamacrocyclé trop flexible, l'insertion d'un Tb(III) dans sa cavité bloquait la conformation du *template* en projetant les quatre bras du même côté, permettant ainsi d'augmenter fortement son efficacité.

En accord avec nos attentes, il apparaît qu'une approche similaire n'est absolument pas bénéfique pour le ^{PNA}DOTASQ. En suivant la même méthodologie que pour le DOTASQ, à savoir l'ajout du ^{PNA}DOTASQ dans 1,1 équivalent de Tb(NO₃)₃.6H₂O dans un mélange méthanol et tampon Caco.Li 20mM pH 7,2, le Tb.^{PNA}DOTASQ correspondant est obtenu avec un rendement de 78%.

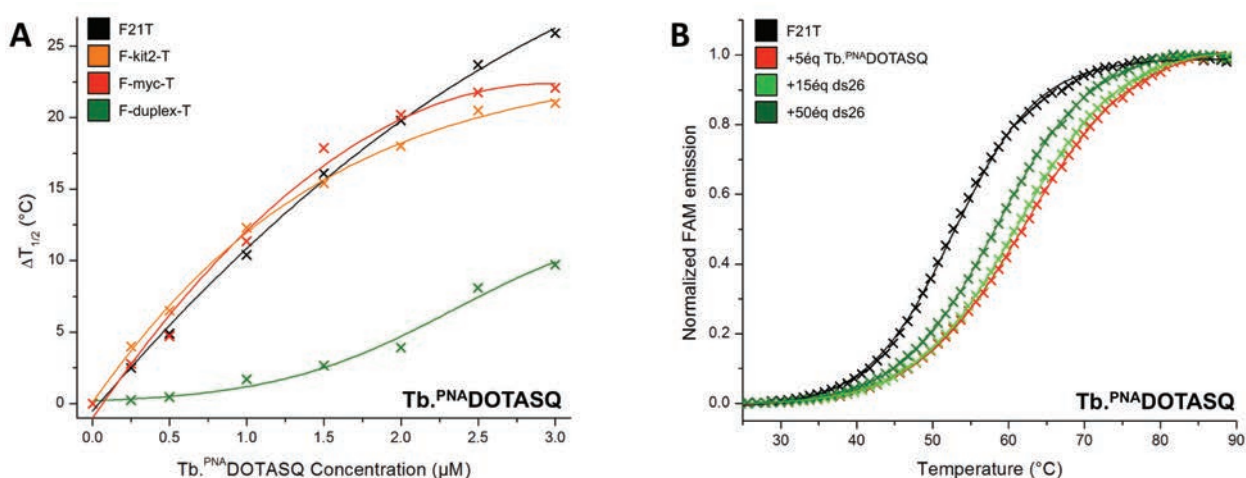


Figure 104 – Influence de la concentration en Tb.^{PNA}DOTASQ sur la valeur $\Delta T_{1/2}$ de stabilisation de F21T, F-myc-T, F-kit-T et F-duplex-T (A). Résultats de FRET-melting sur F21T et Tb.^{PNA}DOTASQ en présence de compétiteurs duplexes ds26 (B).

Comme le montrent les courbes Figure 104A, l'insertion du terbium n'accentue pas le pouvoir stabilisant de ce composé qui reste sensiblement proche du ^{PNA}DOTASQ ($\Delta T_{1/2}$ =10,4 et 12,5°C avec et sans Tb(III) à 5éq, respectivement). En revanche, l'augmentation de la charge positive globale du composé de 4 à 7 entraîne une perte de sélectivité qui se traduit par de fortes interactions aléatoires électrostatiques avec le duplexe avec un $\Delta T_{1/2}$ de 9,7°C pour la forme métallée contre 2,5°C pour la forme libre à 15éq. Les expériences complémentaires de compétition mettant en jeu F21T et le duplexe ds26 non marqué (Figure 104B) permettent de calculer le facteur de sélectivité ^{FRET}S qui chute à 60% pour 50éq de duplexe.

Ainsi, l'insertion d'un métal dans le but de rigidifier le cycle s'avère inefficace en terme de stabilisation, démontrant que de par son architecture intrinsèque, le ^{PNA}DOTASQ est parfaitement capable de refermer sa tétrade synthétique au contact de la tétrade naturelle d'un quadruplexe. Cette modification, qui entraîne une augmentation de la charge cationique globale, a au contraire un effet négatif puisqu'elle diminue significativement la sélectivité du composé, point clef du développement de ligands de quadruplexes.

Cette étape de métallation nous a donc permis de confirmer que la présence de bras PNA-guanines au sein du ^{PNA}DOTASQ suffit à lui conférer une qualité de « smart ligand », indépendamment de la conformation du *template* : la formation de la tétrade synthétique intramoléculaire est donc uniquement régie par un processus de reconnaissance biomimétique.

X.5.C Détermination du BONDS-index

Comme nous avons pu le voir très largement au cours du premier chapitre de cette thèse, les structures non-usuelles d'ADN présentes dans la cellule sont nombreuses et forment toutes des cibles potentielles pour les ligands développés par différentes équipes de recherche. Afin d'évaluer les préférences d'interactions d'un composé pour une structure visée, nous avons mis en place au cours de ces travaux de thèse un protocole expérimental basé sur la méthode de FRET-melting en plaque 96 puits. Il permet de quantifier l'affinité d'un nouveau composé pour onze séquences d'ADN se structurant selon les quatre structures secondaires les plus représentatives, à savoir : les quadruplexes (GQ), les duplexes (DS), les jonctions à trois voies (ou TWJ, voir III.3.B.b) et les jonctions à quatre voies (FWJ, voir III.3.B.c). Ainsi, le calcul du BONDS-index (pour *branched and other non-canonical DNA selectivity*) ou BI s'effectue de la manière suivante :

$$^{GQ}BI = \frac{^{GQ}\Delta T_{1/2}}{^{GQ}\Delta T_{1/2} + ^{DS}\Delta T_{1/2} + ^{TWJ}\Delta T_{1/2} + ^{FWJ}\Delta T_{1/2}}$$

Sans détailler ici la mise au point de la méthode (pour une description complète, voir référence ^[355]), l'essai sur le ^{PNA}DOTASQ confirme une forte stabilisation du quadruplexe, une très faible interaction avec le duplexe et la jonction à trois voies, mais révèle une interaction non négligeable avec la jonction d'ADN à quatre voies, conduisant à $^{GQ}BI=0,66$.^[19] Puisque le ligand ne reconnaît pas le duplexe, il semble probable que l'interaction s'effectue au niveau du point de jonction central de la jonction d'ADN à quatre voies, même si aucune preuve ne nous permet actuellement de le vérifier. Des études sont actuellement en cours afin d'éclaircir ce point, notamment par électrophorèse sur gel.

X.5.D Interaction du ^{PNA}DOTASQ avec les quadruplexes d'ARN

X.5.D.a Etudes préliminaires

Alors que depuis une quinzaine d'années, de nombreuses études sur les quadruplexes d'ADN ont permis de développer une multitude de ligands et de méthodes d'analyses biophysiques, très peu de travaux s'intéressent aujourd'hui aux quadruplexes d'ARN. Présents au niveau des régions 5'-UTR des ARNm ou au sein des TERRA, les quadruplexes d'ARN constituent de nouvelles cibles prometteuses potentiellement plus accessibles que leurs homologues ADN de par leur localisation (voir VI.2).

Néanmoins, le coût engendré par la synthèse de l'ARN, bien supérieur à celui de l'ADN, peut constituer un frein à cette recherche. Dans le cas d'ARN doublement marqués par les sondes FAM et TAMRA en vue de réaliser des expériences de FRET-melting, les demandes sont tellement faibles qu'ils sont préparés à façon par les fournisseurs d'oligonucléotides.

Un article récent de M. Di Antonio *et al.* paru dans *Angew. Chem. Int. Ed.* en 2012 présentait le premier ligand sélectif de quadruplexes d'ARN, un dérivé de la pyridostatine, par une méthodologie de FRET-melting sur laquelle nous reviendrons ci-après.^[373] La synthèse d'un ligand capable de reconnaître sélectivement un quadruplexe d'ARN et non d'ADN est une quête particulièrement complexe de par leurs faibles différences structurales, mais présente un formidable intérêt pour la compréhension des mécanismes cellulaires régis par les quadruplexes d'ARN.

Dans ce contexte, nous avons voulu évaluer la capacité du ^{PNA}DOTASQ à stabiliser les quadruplexes d'ARN. Notre choix s'est porté sur trois séquences :

- **L-TERRA** (FAM-r(^{5'}G₃(U₂AG₃)₃^{3'})-TAMRA) correspondant aux quadruplexes des *telomeric repeat-containing RNA*.
- **L-TRF2** (FAM-r(^{5'}CG₃AG₃CG₄AG₃C^{3'})-TAMRA) présente dans la région 5'-UTR du gène codant pour la protéine TRF2 constituant la shelterin et répondant à un mécanisme *cap-dependant* (voir VI.2.B).^[197]
- **L-VEGF** (FAM-r(^{5'}G₂AG₂AG₄AG₂AG₂A^{3'})-TAMRA) présente dans la région 5'-UTR du gène codant pour le facteur de croissance de l'endothélium vasculaire et répondant à un mécanisme *cap-independant* (voir VI.2.B).^[197]

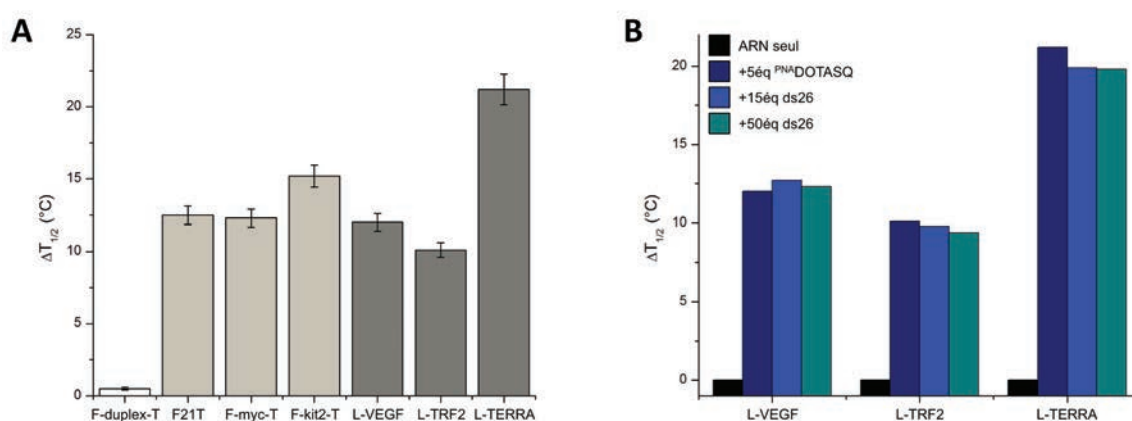


Figure 105 – Résultats de FRET-melting obtenus pour F-duplex-T (en blanc), trois quadruplexes d'ADN (en gris clair) et trois quadruplexes d'ARN (en gris foncé) en présence de 5éq de ^{PNA}DOTASQ (A). Diagrammes récapitulatifs des expériences de compétitions réalisées par FRET-melting sur les quadruplexes d'ARN en présence de 5éq de ^{PNA}DOTASQ et présence ou en l'absence de quantités croissantes de duplexe ds26 (B).

Les température de fusion $T_{1/2}$ de chacune de ces structures sont de 57,0°C, 66,9°C, et 62,7°C, respectivement, dans un tampon Caco.Li à 1mM K⁺, pH 7,2. L'ajout de 5éq de ^{PNA}DOTASQ conduit à une stabilisation $\Delta T_{1/2}$ de 21,1°C, 10,1°C et 12,0°C, respectivement (Figure 105A).

Globalement, ces résultats confirment que le ^{PNA}DOTASQ interagit fortement avec les quadruplexes quelle que soit leur nature, confirmant ainsi l'intérêt de son architecture. De façon intéressante, le $\Delta T_{1/2}$ obtenu sur F21T, équivalent ADN du L-TERRA, était de 12,5°C contre 21,1°C (Figure 105A), témoignant d'une bien meilleure affinité de notre ligand pour l'ARN que pour l'ADN. Cette différence peut provenir de la modification de la conformation du quadruplexe (hybride pour l'ADN et parallèle pour l'ARN), de l'accessibilité de la tétrade extérieure (présence de boucles latérales pour l'ADN et uniquement *chain-reverse* pour l'ARN) ou encore d'une interaction avec la fonction hydroxyle en position 2' sur le ribose de l'ARN, absent sur l'ADN.

Une série d'expériences classiques de compétitions a été menée sur ces trois structures quadruplexes d'ARN en présence de 5éq de ^{PNA}DOTASQ avec 15 ou 50éq de duplexe d'ADN ds26 et valide encore l'excellente sélectivité de ce composé avec un facteur de sélectivité ^{FRET}S>95% (Figure 105B). Notons que pour ce type d'expérience, la température de fusion du ds26, évaluée à

$T_m=79^\circ\text{C}$ par UV-melting, est supérieure à la $T_{1/2}$ du L-TERRA en présence de 5éq de ligand ($^{L-TERRA}T_{1/2}(+5\text{éq}^{PNA}\text{DOTASQ})=57+21,1=78,1^\circ\text{C}$).

X.5.D.b Notion de préférence ADN/ARN

Les expériences de FRET-melting ont mis en évidence une stabilisation plus forte des quadruplexes d'ARN que d'ADN. Pour le vérifier, nous nous sommes inspirés des travaux de M. Di Antonio *et al.* sur la pyridostatine au cours desquels, en partant d'un système L-TERRA + 5éq de ligand, des quantités croissantes de H-Telo (correspondant à la séquence télomérique humaine notée 22AG dans ce manuscrit) sont ajoutées.^[373] Or, les températures de fusion de chacune de ces deux structures sont quasi identiques et rendent cette méthodologie caduque. Avec une stabilisation $\Delta T_{1/2}$ d'environ 35°C , l'adduit L-TERRA + pyridostatine présente une température de fusion de l'ordre de $T_{1/2}=50+35=85^\circ\text{C}$ très supérieure à celle du compétiteur 22AG qui se situe autour de $50-55^\circ\text{C}$. Ainsi, au moment de la fusion de l'adduit, le compétiteur s'est déjà déstructuré, il ne peut donc pas assurer sa fonction et invalide leurs conclusions.

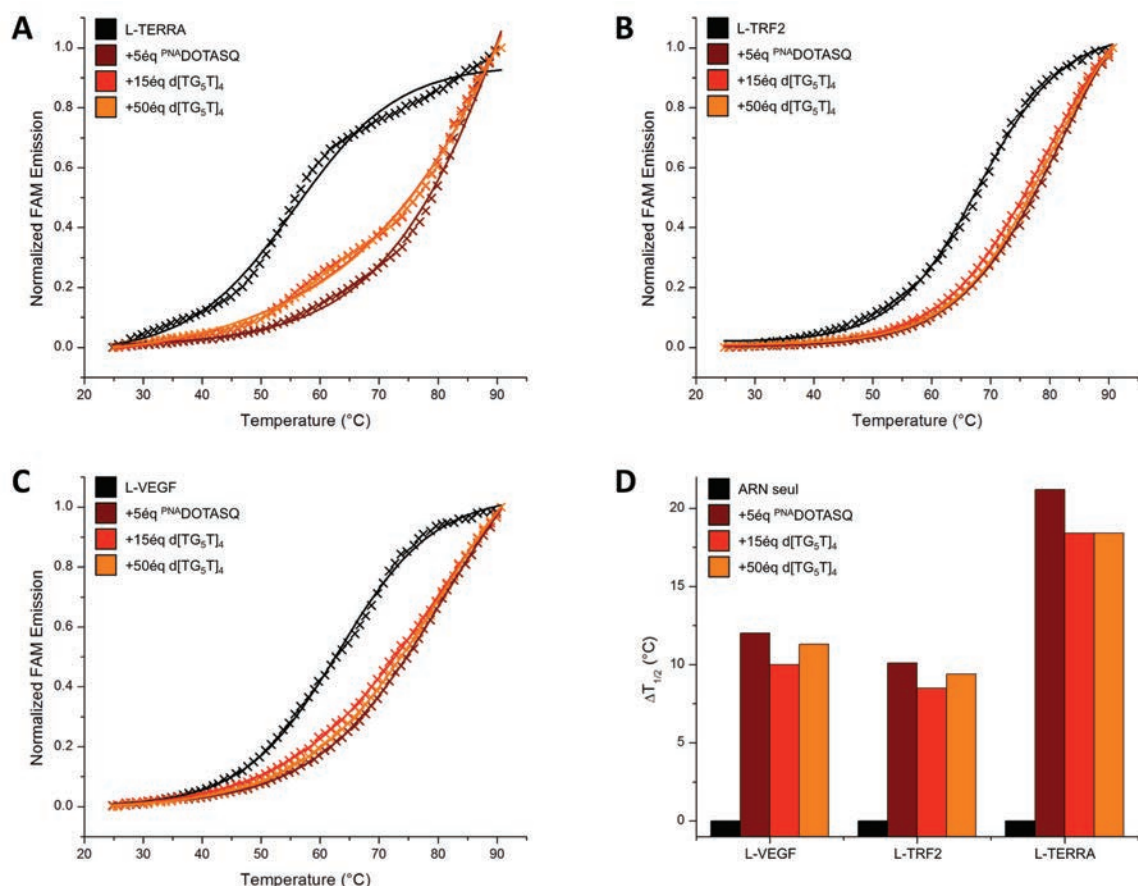


Figure 106 – Courbes de FRET-melting réalisées sur le ^{PNA}DOTASQ au départ de quadruplexes d'ARN (L-TERRA (A), L-TRF2 (B) et L-VEGF (C)) doublement marqués et en présence de quadruplexes d'ADN d[TG₅T]₄ non-marqués (A-C). Diagramme récapitulatif des températures de stabilisation $\Delta T_{1/2}$ obtenues (D).

Pour résoudre ce problème, l'utilisation d'un quadruplexe tétramoléculaire d'ADN compétiteur très stable comme $[TG_5T]_4$ ($T_m > 90^\circ\text{C}$) est indispensable pour ce type d'expérience.^[187]

Nous avons donc réalisé les expériences sur L-TERRA, L-TRF2 ou L-VEGF en présence de 5éq de $^{PNA}DOTASQ$ avec 15 ou 50éq de $[TG_5T]_4$ (Figure 106) et d'une manière intéressante, le $\Delta T_{1/2}$ varie peu après l'ajout de compétiteur et permet d'obtenir des facteur de sélectivité pour l'ARN de $^{ARN}S(^{PNA}DOTASQ)=87, 93$ et 94% , respectivement.

Pour valider ces résultats la compétition inverse est réalisée, c'est-à-dire en utilisant le quadruplexe d'ADN marqué F21T en présence de 5éq de $^{PNA}DOTASQ$ sur lequel le quadruplexe tétramoléculaire $[UG_5U]_4$ (équivalent de $[TG_5T]_4$ en ARN) est ajouté à 15 ou 50éq (Figure 107A). Le déplacement des courbes observé ici témoigne d'une préférence d'affinité du ligand pour l'ARN par rapport à l'ADN, avec un facteur de sélectivité pour l'ADN de $^{ADN}S(^{PNA}DOTASQ)=37\%$.

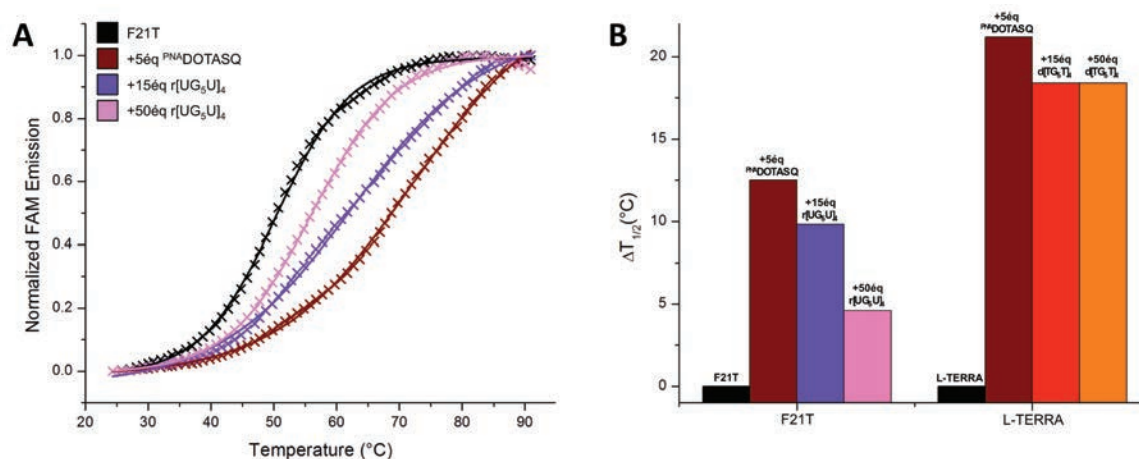


Figure 107 – Résultats des expériences de FRET-melting sur le quadruplexe d'ADN F21T, seul ou en présence de 5éq de $^{PNA}DOTASQ$, avec comme compétiteur le quadruplexe tétramoléculaire d'ARN $r[UG_5U]_4$ à 15 ou 50éq (A).

Diagrammes comparatifs des deux types de compétitions quadruplexes d'ADN/quadruplexes d'ARN réalisées sur le $^{PNA}DOTASQ$ (B).

Ainsi, lorsque le $^{PNA}DOTAQ$ est mis simultanément en présence de quadruplexe d'ADN et d'ARN, l'interaction préférentielle se fera au niveau de l'ARN comme le montrent les expériences de compétitions, pour des raisons qui restent toutefois à élucider. Néanmoins, une attention toute particulière doit être portée sur la sémantique ; en effet, un ligand tel que la pyridostatine ne peut être qualifié de sélectif des quadruplexes d'ARN puisque ce terme implique une reconnaissance spécifique pour les structures ARN sans interagir avec l'ADN. De ce fait, l'emploi du qualificatif « interaction préférentielle » paraît beaucoup plus prudente et qualifie de manière plus précise les propriétés d'un tel ligand.

X.6 Bilan pour le ^{PNA}DOTASQ

Le ^{PNA}DOTASQ est le dernier ligand biomimétique développé au cours de cette thèse et présente une évolution importante en terme de propriétés d'affinités et de sélectivités pour les quadruplexes par rapport à ses précurseurs DOTASQ et PorphySQ.

L'exceptionnelle hydrosolubilité de ce composé nous a permis d'utiliser de nombreuses méthodes d'analyses parmi lesquelles la RMN ¹H, la spectrométrie de masse ESI-MS, le FRET-melting et le FID. D'autres études biophysiques telles que la SPR (*surface plasmon resonance*)^[250] en collaboration avec l'équipe du Pr. E. Defrancq à l'Université Joseph Fournier de Grenoble ont également été entreprises mais nécessitent d'être confirmées, tout comme les analyses par électrophorèse sur gel ou par NOESY que nous développons actuellement au laboratoire.

Le mode d'interaction du ^{PNA}DOTASQ par une reconnaissance entre sa tétrade synthétique et la tétrade naturelle d'un quadruplexe semble clairement démontré par les études RMN ¹H où apparaissent dans différentes conditions expérimentales un signal centré à 11,4ppm attribué aux protons amides du réseau de liaisons hydrogènes de la tétrade de guanines intramoléculaire. De manière intéressante, l'absence de ce signal lorsque le composé est laissé seul en solution tend à démontrer que la tétrade synthétique ne se forme pas spontanément et nécessite d'être au contact d'un quadruplexe. Cette interaction est donc régie par un processus biomimétique de reconnaissance entre tétrades.

La présence des bras PNA, et plus particulièrement des quatre amines pendantes protonées à pH physiologique, joue un rôle supplémentaire d'« agrafe » offrant un second mode d'interaction au ^{PNA}DOTASQ de type électrostatique avec les groupements phosphates du squelette des ADN ou des ARN. Cet apport se traduit par une augmentation drastique de l'affinité du composé, avec des $\Delta T_{1/2}$ de 12 à 15°C pour l'ADN montant jusqu'à 21°C pour l'ARN tout en conservant une excellente sélectivité, comme tous les autres TASQ développés au laboratoire.

La mise en place d'une méthodologie rigoureuse pour l'évaluation de la préférence d'affinité d'un ligand pour les quadruplexes d'ADN ou d'ARN démontre clairement que pour le cas du ^{PNA}DOTASQ, l'interaction préférentielle s'effectue avec les structures ARN. Bien que plusieurs hypothèses soient avancées, comme la différence de conformation, d'accessibilité des tétrades ou de la nature du ribose, il serait particulièrement intéressant d'évaluer les ligands de quadruplexes d'ADN de la littérature les plus emblématiques afin de rationaliser cette préférence et vérifier si elle est dépendante des TASQ ou au contraire générale.

XI. Conclusions et perspectives

Depuis plus de quinze ans, les quadruplexes sont au cœur d'un formidable travail de recherche pluridisciplinaire qui a permis de comprendre leur morphologie mais également de décrypter en partie leurs multiples fonctions biologiques. D'une manière intéressante, il a été mis en lumière que les séquences capables de se structurer sous la forme de quadruplexes d'ADN ne sont pas réparties de manière aléatoire au sein du génome mais sont, au contraire, particulièrement enrichies au niveau des télomères et des promoteurs d'oncogènes ou de gènes de régulations.^[10, 13, 154, 299-301] Les régions télomériques qui assurent la protection du chromosome, c'est-à-dire sa stabilité, jouent un rôle clef dans l'immortalisation des cellules cancéreuses en association avec une enzyme, la télomérase, silencieuse dans les cellules saines et surexprimée dans un contexte tumoral.^[12, 195, 293] Les promoteurs d'oncogènes recrutent quant à eux les ARN polymérases chargées de la transcription du gène en ARN qui est ensuite traduit en oncoprotéine, impliquée dans l'une des six grandes caractéristiques de la cellule cancéreuse définies par D. Hanahan et R. Weinberg.^[154, 303, 304]

De plus, les quadruplexes peuvent également se former au départ de l'ARN comme dans le cas des TERRA, issus de la transcription du brin C des télomères et impliqués dans le processus d'inhibition de la télomérase.^[200, 201, 312] Les régions 5'-UTR des ARNm, dont la fonction est de recruter la petite sous-unité du ribosome indispensable à la synthèse de la protéine codée par l'ARNm, se voient également enrichies en séquences riches en guanines capables de se structurer en quadruplexes d'ARN.^[197]

Il apparaît ainsi que les quadruplexes d'ADN et d'ARN entrent en jeu dans des processus clefs responsables de la stabilité chromosomique (ADN télomérique et TERRA) et de la régulation de l'expression génique (au niveau transcriptionnel avec les promoteurs d'ADN oncogéniques, et au niveau traductionnel au sein des régions 5'-UTR des ARNm). Les études réalisées *in vitro* et *in cellulo* tendent à montrer que le maintien de ces différentes fonctions repose sur l'équilibre entre la forme simple brin et quadruplexe de la séquence riche en guanines. Il devient alors intéressant dans un contexte où la cellule n'assure plus une régulation raisonnée, c'est-à-dire une cellule cancéreuse, de déplacer cet équilibre vers la forme quadruplexe très généralement liée à une inhibition de la fonction. Pour cela, de nombreux ligands de quadruplexes ont été développés et démontrent l'intérêt de cette approche aussi bien sur le ciblage des zones télomériques (en créant une instabilité chromosomique)^[195, 292, 293] que des promoteurs d'oncogènes (avec une diminution de la transcription en ARNm)^[296, 305, 306] ou encore des régions 5'-UTR (avec une diminution de la traduction en protéines).^[322-324]

Tandis que la plupart des ligands actuels sont construits autour de larges surfaces aromatiques permettant des interactions de π -stacking avec les quadruplexes,^[11, 192, 276] nous avons conçu et étudié *in vitro* trois nouvelles séries de molécules de type TASQ au mode d'action complètement innovant. Basés sur une reconnaissance spécifique entre une tétrade de guanines synthétique (TASQ) et naturelle (quadruplexe), ces ligands biomimétiques hydrosolubles offrent d'excellents résultats en terme de sélectivité.

Le premier modèle baptisé DOTASQ (Figure 108A), construit autour d'un cœur cyclène et de quatre bras alkyles-guanines, nous a permis de valider notre approche en démontrant par RMN ^1H la formation intramoléculaire de la tétrade de guanines lorsque le ligand est seul en solution ou en présence de quadruplexes d'ADN. Néanmoins, les expériences de FRET-melting se sont dans un premier temps avérées peu concluantes avec des stabilisations $\Delta T_{1/2}$ n'excédant pas 5,6°C (F21T) pour le DOTASQ-C₅, meilleur que le DOTASQ-C₁ aux bras plus courts. Cette faible affinité est attribuée à une trop grande flexibilité du polyazamacrocycle qui offre un trop grand degré de liberté aux quatre guanines, les empêchant de se rapprocher pour former la tétrade. Ainsi, l'insertion d'un Tb(III) dans la cavité du macrocycle contraint les quatre bras à se placer à proximité et favorise donc la formation de la tétrade qui se traduit par une augmentation drastique des stabilisations avec $\Delta T_{1/2}$ =10,5°C (F21T) pour le DOTASQ-C₅.^[20]

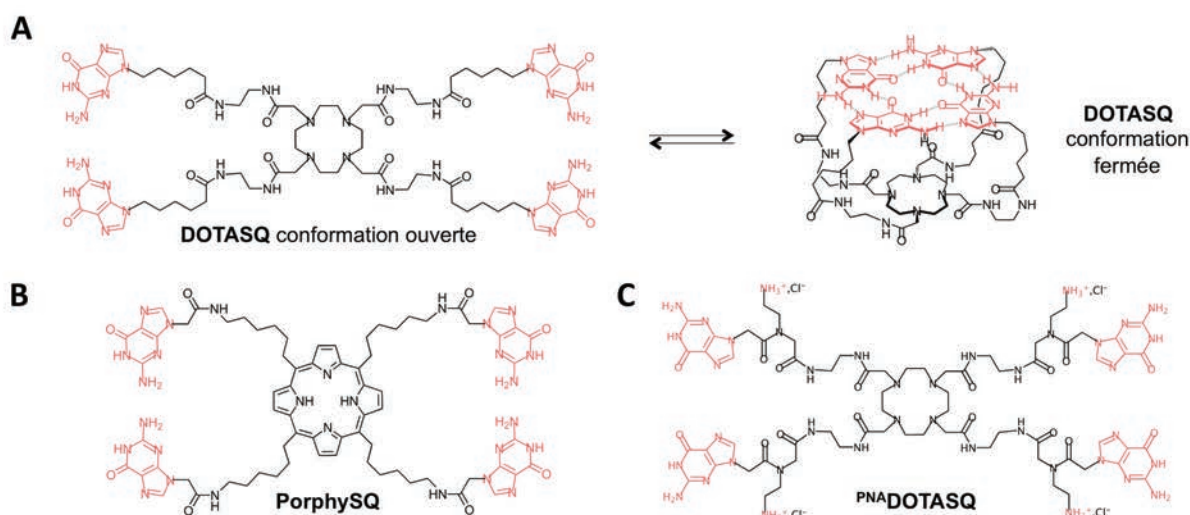


Figure 108 – Structures chimiques des trois TASQ hydrosolubles développés au cours de ce travail de thèse et étudiés comme ligands de quadruplexes.

Afin d'optimiser encore davantage ces premiers résultats, nous avons cherché à remplacer le *template* par un motif capable de promouvoir à lui seul la formation de la tétrade intramoléculaire. Notre choix s'est alors porté sur un cœur porphyrinique, connu pour former des interactions de π -stacking avec les tétrades (Figure 109), et nous a permis de synthétiser le PorphySQ (Figure

108B). Ce nouveau TASQ présente une amélioration face aux DOTASQ d'un facteur trois environ. Néanmoins, l'avantage induit par la porphyrine est contrebalancé par sa solubilité moindre qui ne nous permet pas d'effectuer d'autres études biophysiques où les concentrations aqueuses nécessaires sont supérieures à la dizaine de micromolaires.^[21]

Nous sommes alors repartis du DOTASQ auquel nous avons cette fois-ci modifié la nature des bras. Ainsi, le remplacement des bras alkyles-guanines par des bras PNA-guanines conduit à la troisième série de TASQ développée au cours de cette thèse, le ^{PNA}DOTASQ (Figure 108C). Offrant une solubilité exemplaire dans les tampons à pH 7,2, nous avons pu réaliser de nombreuses expériences dont la RMN ¹H qui a démontré une interaction sélective entre notre composé et plusieurs quadruplexes d'ADN, mais a surtout révélé un mode d'action biomimétique innovant. En effet, alors que seul en solution le ^{PNA}DOTASQ est présent sous une conformation ouverte où les quatre guanines sont indépendantes, l'ajout de quadruplexes d'ADN entraîne la fermeture de la tétrade intramoléculaire, visible par l'apparition d'un signal caractéristique à 11,4ppm. La fermeture de la tétrade synthétique est la condition *sine qua non* à son interaction biomimétique avec la tétrade naturelle d'un quadruplexe.

Les expériences complémentaires par ESI-MS, FID et FRET-melting conduisent à la démonstration d'une interaction hautement sélective du ^{PNA}DOTASQ pour les quadruplexes d'ADN quelle que soit leur nature. La valeur clef $\Delta T_{1/2}$ obtenue sur F21T est de 12,5°C faisant de cette molécule le meilleur TASQ hydrosoluble développé à ce jour. En plus d'une reconnaissance tétrade/tétrade semblable à celle du DOTASQ ou du PorphySQ, le ^{PNA}DOTASQ offre un second mode d'interaction grâce à ses fonctions amines qui jouent un rôle d'« agrafe » en se liant de manière électrostatique aux groupements phosphates de l'oligonucléotide (Figure 109).^[18]

Afin de compléter cette étude, nous avons voulu évaluer l'affinité de notre ligand pour les quadruplexes d'ARN. Après la mise en place d'une méthodologie expérimentale *ad hoc* de FRET-melting par compétition, il a été mis en évidence que les interactions avec les quadruplexes d'ARN sont favorisées au détriment de celles avec l'ADN. Ainsi, le ^{PNA}DOTASQ présente une interaction préférentielle avec l'ARN, qui demande encore à être rationalisée par des tests complémentaires.^[19]

Ce panel d'analyses nous a donc permis de confirmer le mode d'interaction des TASQ hydrosolubles dont le ^{PNA}DOTASQ est le meilleur représentant actuel. Sa forte affinité et son excellente sélectivité en font un parfait candidat pour des études *in cellulo* qui permettront dans un premier temps de vérifier l'incorporation de la molécule au sein des cellules et sa localisation (cytoplasme ou noyau) puis de déterminer leur mode d'action (interactions avec l'ARN, l'ADN

télomérique ou oncogénique, et impact sur le plan transcriptionnel et traductionnel) et leur efficacité sur les cellules tumorales par rapport aux cellules saines. Ainsi, une première étude a été réalisée sur le ^{PNA}DOTASQ par H. Lévaïque du groupe de R. Rodriguez à l'ICSN de Gif-sur-Yvette. Les résultats préliminaires montrent une cytotoxicité sur la lignée cellulaire de fibrosarcome HT1080 avec une valeur IC_{50} de 18 μ M correcte mais modeste puisqu'à titre de comparaison, le ligand pyridostatine présente une IC_{50} sur cette même lignée de 1 μ M. Néanmoins, ce test préliminaire nous conforte sur l'activité cellulaire du TASQ pour lequel une collaboration vient d'être créée avec le Dr. J. Wong de l'Université de British Columbia de Vancouver afin d'évaluer plus précisément le mode d'action de ce composé. Parmi les analyses envisagées, une première étape cherchera à évaluer l'efficacité d'un traitement au ^{PNA}DOTASQ sur les lignées cellulaires de cancer du sein MCF-7 et MDA-MB-231 en comparaison à une lignée saine de cellules épithéliales mammaires humaines (HMEC). Puis, des études de co-localisation des quadruplexes *in cellulo* seront menées selon la méthodologie publiée par l'équipe de S. Balasubramanian,^[297] reposant sur l'utilisation d'anticorps spécifiques des structures quadruplexes. Des mesures par cytométrie de flux permettront ensuite de quantifier l'interaction entre le TASQ et les quadruplexes et seront suivies par l'analyse du transcriptome afin de déterminer son impact sur l'expression génique. Cette méthodologie innovante publiée récemment par R. Rodriguez *et al.*^[296] sera complétée par une expérience de ChIP-seq (pour *chromatin immunoprecipitation sequencing*) qui permettra de remonter à la séquence et à la nature du gène ciblé.

En parallèle, de nouveaux ligands de type TASQ sont actuellement en développement au sein du laboratoire et leur conception s'inspire des différentes conclusions émanant de ce travail de thèse. Nous avons en effet démontré l'importance de l'utilisation de bras PNA-guanines sur la reconnaissance sélective des quadruplexes d'ADN et d'ARN et mis en évidence l'intérêt d'un cœur aromatique pour la formation de la tétrade de guanines intramoléculaire. L'association de ces deux fonctions au sein de la même molécule permettra de conduire aux TASQ de troisième génération (Figure 109) qui devraient offrir un gain d'efficacité global pouvant être allié à l'ajout d'une nouvelle propriété physique. En effet, si le *template* choisi possède une propriété intrinsèque de fluorescence, le TASQ final sera une molécule aux propriétés de fluorescence dont l'intensité devrait varier en fonction de l'association ou non des guanines sous forme d'une tétrade. Le suivi de ces composés *in cellulo* sera envisageable et apportera des indications majeures sur leur mode d'interaction ainsi que sur leur comportement (passage des membranes, dégradation, etc.).

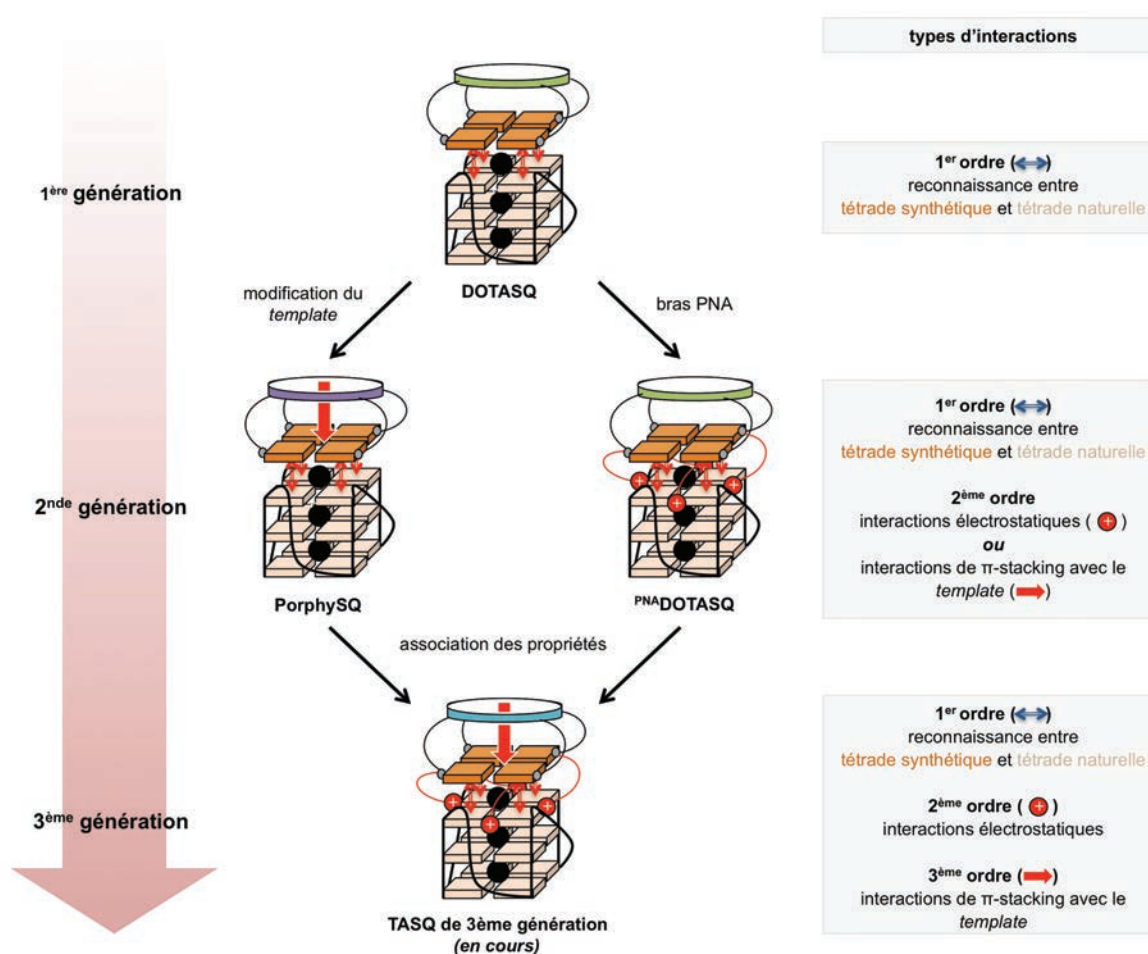


Figure 109 – Représentation schématique des trois générations de TASQ hydrosolubles et des types d'interactions qui les définissent. Seuls les TASQ de 1^{ère} et 2^{ème} génération ont été développés au cours de ce travail de thèse.

Parmi les nouvelles structures envisagées, le ^{PNA}PorphySQ (Figure 110A) dérivant directement du PorphySQ a été synthétisé par A. Laguerre et se trouve actuellement en cours d'études biophysiques. En parallèle, une série de composés aux cœurs polyaromatiques a vu le jour grâce au travail du Dr. D. Genest au sein de notre équipe de recherche. Le premier de ces composés, le NDISQ (Figure 110B), est construit autour d'un naphthalène diimide dont de nombreux résultats comme ligand de quadruplexe ont été reportés, notamment par l'équipe de S. Neidle et S. Balasubramanian.^[295, 339, 374]

Une seconde molécule organisée autour d'un dérivé naphthalène, le NaphtoTASQ (Figure 110C), semble présenter après quelques études préliminaires une très forte affinité pour les quadruplexes ($\Delta T_{1/2} \approx 17^\circ\text{C}$ à 5 μeq sur F21T) mais souffre en revanche d'une sélectivité moindre ($^{\text{FRET}}S < 50\%$ à 50 μeq ds26).

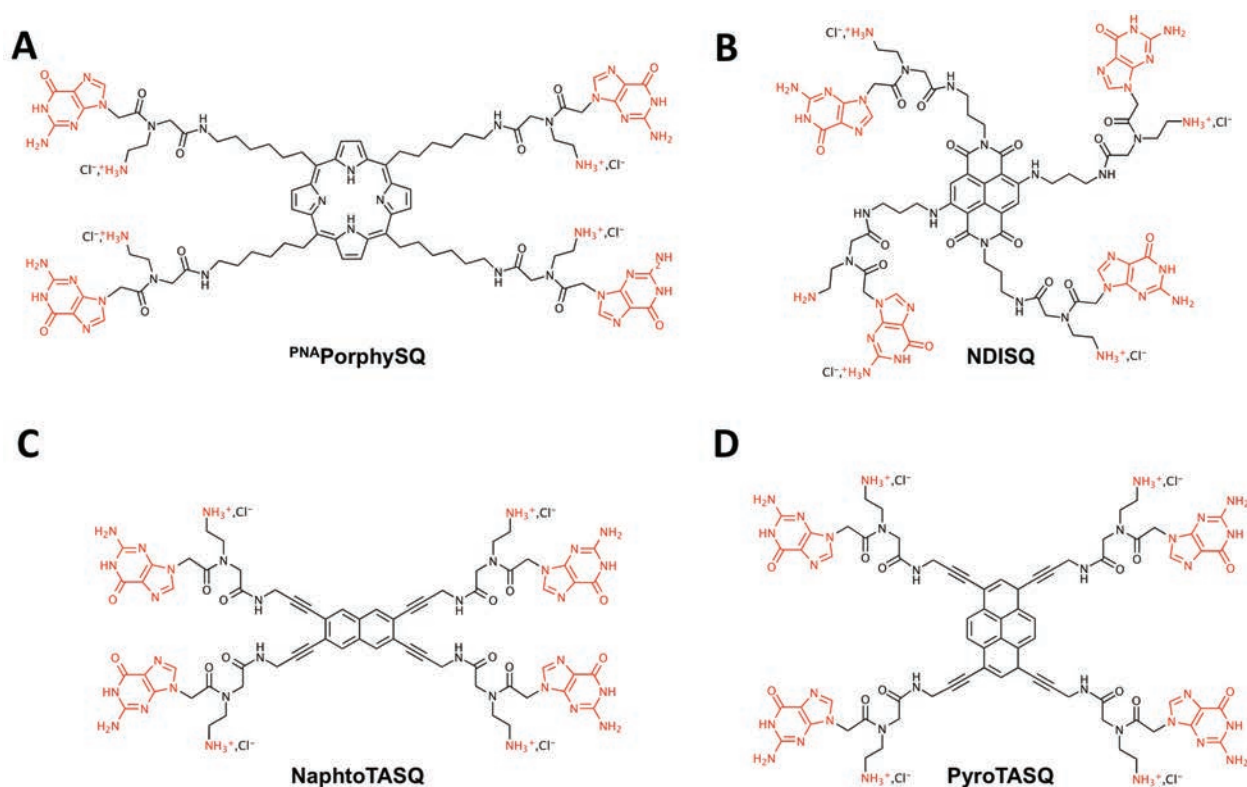


Figure 110 – Structures chimiques des TASQ hydrosolubles de 3^{ème} génération actuellement en développement au sein du laboratoire.

Plus étendu, le cœur pyrène permet quant à lui d'obtenir une affinité toute aussi importante alliée à une excellente sélectivité ($FRET^S > 95\%$) qui font du PyroTASQ (Figure 110D) un TASQ particulièrement prometteur dont les études biophysiques plus complètes sont en cours. L'introduction de quatre fonctions alcynes conjuguées au cœur pyrène conduit à un effet bathochrome favorable, permettant d'éloigner la longueur d'onde d'excitation du PyroTASQ ($\lambda_{ex} \approx 440\text{nm}$) de la zone d'absorbance de l'ADN. Cette propriété pourra, nous l'espérons, permettre de suivre directement ce composé *in cellulo* par mesures de fluorescences afin de suivre son évolution et mécanisme d'action.

