MUSEUM NATIONAL



D'HISTOIRE NATURELLE

Ecole Doctorale Sciences de la Nature et de l'Homme – ED 227

Année 2015

N°attribué par la bibliothèque

THESE

Pour obtenir le grade de

DOCTEUR DU MUSEUM NATIONAL D'HISTOIRE NATURELLE

Spécialité : Sciences Moléculaires et Cellulaires de la Biologie

Présentée et soutenue publiquement par

Karine Nozeret

Le 03 décembre 2015

Nouveaux outils chimiques pour visualiser des séquences d'ADN répétées dans les cellules vivantes

Sous la direction de : Monsieur Boutorine, Alexandre, Professeur

JURY :

M. Teulade-Fichou, Marie-Paule	Directeur de Recherche, Institut Curie, Paris (075)	Président
M. Boutorine, Alexandre	Professeur, Muséum national d'histoire naturelle, Paris (075)	Directeur de Thèse
M. Riou, Jean-François	Professeur, Muséum national d'histoire naturelle, Paris (075)	Examinateur
M. Mignet, Nathalie	Directeur de Recherche CNRS, Unité de Technologies Chimiques et Biologiques	Rapporteur
	pour la Santé, Paris (075)	
M. Constant, Jean-François	Chargé de Recherche CNRS, Département de Chimie Moléculaire, Grenoble (038)	Rapporteur
M. Jullien, Ludovic	Professeur, Département de Chimie de l'Ecole Normale Supérieure, Paris (075)	Examinateur

Á ma famille et mes amis,

Remerciements

J'aimerai tout d'abord remercier les Drs Nathalie Mignet et Jean-François Constant d'avoir accepté d'être rapporteurs de cette thèse. Je remercie également les Drs Marie-Paule Teulade-Fichou et Ludovic Jullien, ainsi que le Pr Jean-François Riou d'avoir accepté d'être membres du jury et d'évaluer ce travail.

Je remercie les Drs Carine Giovannangeli et les Prs Jean-François Riou et Tula Saison-Behmoaras pour m'avoir accueillie au sein de leur laboratoire.

Ce travail a été encadré par un comité de thèse composé par Marie-Paule Teulade-Fichou, Chahrazade El Amri, Tula Saison-Behmoaras et Anne-Laure Guieysse-Peugeot donc merci à vous, mesdames, pour votre aide, vos critiques et votre soutien pendant ces trois années de thèse.

Mes remerciements s'adressent tout particulièrement au Pr Alexandre Boutorine (Sacha) pour avoir dirigé et encadré ce travail de thèse. Merci très sincèrement pour ma formation qui a commencée dès mon master, votre disponibilité, vos conseils, pour les libertés d'action et surtout d'expression !!! Et, bien sûr, pour avoir supporté mon caractère et ma maniaquerie et merci pour la confiance que vous m'avez accordée tout au long de ces années de thèse. Sans oublier merci pour la découverte de Novossibirsk, la rencontre et l'accueil de votre sœur, de Lana et de nos collaborateurs russes (les Drs Alya Venyaminova et Darya Novopashina, Svetlana, Olga,...) et merci pour la vodka russe ©.

Je tiens également à remercier le Dr Christophe Escudé pour m'avoir accueillie au sein de son équipe.

Un grand merci à Patrick Maillet, sans qui, je n'aurai pas poursuivi en thèse (avec l'appui de Jean-François). Merci pour tous tes conseils jusqu'à la relecture de ce manuscrit, ton soutien, ta disponibilité, ta bonne humeur dans le bureau et surtout pour avoir compris mon caractère !

Je voudrais particulièrement remercier François Loll qui m'a formée aux manips cellulaires et qui m'a supportée toutes ces longues heures devant le microscope !!! Merci pour ton aide, ton soutien, ta bonne humeur, tes blagues carambar et le sacrifice de cet été pour relire mon manuscrit. Un GRAND merci au Dr Anthony Bugaut, pour ton soutien

1

pendant la rédaction des articles, pour ta disponibilité sans faille surtout cet été pour la relecture de ce manuscrit, pour les pauses clopes, pour ton aide quand le moral était au plus bas et quand les machines à Tm étaient capricieuses et surtout les déjeuners « entre filles ». Je n'oublierai pas mon chimiste préféré allias Gildas Mouta Cardoso qui a su se moquer de mes lacunes en chimie (puisque je ne suis pas chimiste de formation) mais qui a tout particulièrement permis que je travaille dans de bonnes conditions (blagues pourries, gentillesse, bonne humeur, taquineries, pain d'épices). Merci pour la synthèse des fluorophores et merci pour tous tes conseils !!! Nous formions une équipe de choc pour la maintenance des HPLC et pour mettre de la couleur au labo même sur la verrerie. Merci également à Julien Cochennec « l'ex-informaticien » pour ton soutien et ta disponibilité.

Merci à mes cOUpines du labo (les « B...... du labo ») Hind Ghezraoui, Marine Charpentier, Charlotte Boix et Armêl Millet pour leur soutien, leur bonne humeur, les déjeuners « entre filles », les soirées gin-fizz et gâteaux au rhum, les chorégraphies du L2, et tant d'autres choses que je ne peux pas citer dans ce manuscrit. Merci à vous les filles et bon courage après mon départ !!!

Ce travail n'aurait pas pu être réalisé dans d'aussi bonnes conditions sans le soutien, les conseils et l'attention que m'ont portés les membres du laboratoire au cours de ces années de thèse. Merci à Patrizia Alberti (mon équivalent psychorigide) pour sa confiance en ce qui concerne l'utilisation des appareils de physico-chimie et surtout des machines à Tm très capricieuses et notamment merci d'avoir remis en marche « l'ancêtre » pour que je puisse finir mes manips en juin, pour le dichroïsme circulaire et les bonnes rigolades et bien évidemment pour ton francotalien que j'adore. Merci à Anne De Cian et Erika Brunet, pour vos critiques constructives, votre bonne humeur et votre écoute. Merci à Loïc Ponger « le bioinformaticien », qui s'est cassé la tête sur les séquences et merci pour tes tartes à la praline. Merci à Loïc Perrouault pour m'avoir tenue compagnie pendant mes expériences de FISH, pour tes connaissances sur les gels et le prêt de tes peignes, pour le saumon fumé et ta bonne humeur. Merci à Judith Lopes pour ta compréhension en ce qui concerne la disponibilité du PSM, pour tes questions en réunion d'équipe et pour les discussions. Merci à Carole Saintomé pour ta gentillesse et ton aide en ce qui concerne la « quête » de ma salle de soutenance. Merci à Emmanuelle Delagoutte pour tes bons conseils et nos petites discussions matinales. Merci à Jean-Baptiste Boulé pour les pauses clopes et les rigolades. Merci à Arul Marie et Lionel Dubost pour les spectres de masse. Merci aux ex-thésards Céline, Jean,

Lorena et Vivien que j'ai eu la chance de côtoyer et aux étudiants et thésards qui ont encore du boulot : Marion, Lauriane, Astrid, Pauline, Loélia, Frédérick, Sheldon et Gabriel. Merci aux secrétaires Corinne, Fara et Patricia pour votre gentillesse et votre disponibilité. Merci à Gilles « notre homme à tout faire » pour ta disponibilité. Merci à tous les appareils du labo qui ont su jouer avec mes nerfs en tombant en panne quand j'en avais besoin. Et merci à tous les autres membres du labo pour leur accueil et leur gentillesse.

Et pour finir comme on dit « le meilleur pour la fin », je ne remercierai jamais assez mes parents et ma sœur qui ont toujours « trop » cru en moi et qui m'ont soutenue et supportée aussi bien moralement que financièrement pendant toutes ces années d'études et désolée pour mon stress permanent et ma mauvaise humeur pendant la rédaction de ce manuscrit. Merci à mes grands-parents pour leurs éternelles questions : Mais qu'est-ce que tu fais exactement ? Tu comptes travailler quand ? Merci à mon Daco et mon mimie pour m'avoir tenue compagnie pendant toutes ces heures de rédaction. Un grand merci à mes amis, sans qui, je n'aurai pas tenu le coup : les frères Blin pour nos repas du mercredi, les vacances et bien sûr votre soutien; Titi pour ta bonne humeur et si je deviens bibliothécaire bref...; Sosso pour ton soutien et toutes ces soirées pour se changer les idées; Cyril et Axelle pour les petits repas du mardi quand c'était possible et votre écoute; Tif mon acolyte et nos soirées hebdomadaires ; Medmed pour ta bonne humeur ; Ro pour tes encouragements et ton intérêt à mon projet ; Vesna, Jimmy et les loulous pour faire le parrainage de mon filleul au moment de ma rédaction ; Pedro et Sophie pour votre soutien et surtout les vacances à la Réunion qui m'ont permis de souffler pendant ma thèse (et pour votre mariage juste avant ma soutenance). Et si par hasard j'avais omis de remercier quelqu'un, il comprendra qu'avec le stress l'attention peut être altérée.

Résumé

Mots clés : [Séquences d'ADN répétées centromériques ; Ligands du petit sillon de l'ADN ; Synthèse peptidique sur support solide ; « chimie-click » ; Reconnaissance séquence-spécifique de l'ADN ; Sondes fluorescentes ; Cellules fixées et vivantes ; Imagerie cellulaire]

L'imagerie de l'ADN dans des cellules vivantes par la microscopie de fluorescence requiert habituellement des approches transgéniques qui modifient le génome. Des polyamides synthétiques composés par des monomères N-méthylpyrrole (Py) et Nméthylimidazole (Im), pouvant interagir spécifiquement dans le petit sillon d'une séquence d'ADN double brin, représentent une approche intéressante pour l'imagerie de fluorescence dans des cellules, sans nécessiter de modification des génomes. Plusieurs polyamides en « épingle à cheveux » et deux tandems « tête à tête » ciblant des séquences répétées centromériques chez la souris et l'homme ont été synthétisés. Avant et après le couplage à divers fluorophores, l'interaction de ces polyamides avec des duplexes d'ADN synthétiques a été étudiée, in vitro, par des méthodes physico-chimiques : dénaturation thermique de l'ADN, retard sur gel d'acrylamide, dichroïsme circulaire et spectroscopie de fluorescence. Une grande variabilité des affinités et des propriétés de fluorescence a été observée, ce qui suggère que la capacité de ces molécules à être utilisées comme sondes fluorescentes ne repose pas simplement sur les règles de reconnaissances établies, mais aussi sur d'autres facteurs structuraux. Chaque sonde doit être testée individuellement avant son utilisation pour les applications cellulaires. Chez la souris, la validation de la spécificité des sondes fluorescentes a été effectuée sur des cellules fixées et les résultats obtenus sur deux lignées de cellules murines vivantes ont permis de valider le concept d'utiliser des sondes fluorescentes, à base de polyamides, pour cibler des séquences spécifiques d'ADN. Pour la première fois, des outils chimiques synthétiques ont permis de visualiser les chromocentres qui sont des structures nucléaires formées par les séquences d'ADN répétées péricentromériques dans des cellules murines vivantes. Chez l'homme, les études cellulaires ont révélé un marquage non spécifique de l'ADN et n'ont pas permis de mettre en évidence des structures apparentées aux chromocentres murins dans des cellules fixées et vivantes, malgré une bonne pénétration des sondes. Cela suggère, d'une part que des modifications structurales des sondes fluorescentes sont nécessaires et d'autre part, que l'accessibilité des séquences d'ADN centromériques est potentiellement différente et donc que l'organisation de ces séquences chez l'homme et la souris ne peut pas être étudiée de la même façon.

Abstract

Title : New chemical tools to visualize repeated DNA sequences in live cells

Keywords : [Centromeric repeated DNA sequences ; DNA minor groove binders ; Peptidic solid phase synthesis ; « click-chemistry » ; Sequence-specific DNA recognition ; Fluorescent probes ; Fixed and live cells ; Cell imaging]

DNA imaging in live cells by fluorescence microscopy usually requires transgenic approaches that modify the genome. Synthetic N-methylpyrrole-N-methylimidazole (Py-Im) polyamides that bind specifically to the minor groove of double-stranded DNA represent an attractive approach for in cell imaging that does not necessitate modification of the genome. Several hairpin polyamides and two « head to head » tandem polyamides that target murine and human centromeric repeated sequences have been synthesized. Before and after coupling to various fluorophores, their interaction with synthetic DNA duplexes have been studied, in *vitro*, by physico-chemical methods : thermal denaturation ; gel-shift assay, circular dichroism and fluorescence spectroscopy. A great variability in affinities and fluorescence properties has been observed, suggesting that the capacity of these molecules to be used as fluorescent probes does not only rely on the established recognition rules, but also on other structural factors. Individual testing of each probe is needed before cellular applications. In the murine model, the specificity of the fluorescent probes was validated in fixed cells and the results obtained on two murine cell lines have validated the concept of use the fluorescent probes based on polyamides to target specific DNA sequences. For the first time, chemical synthetic tools have lead to the live cell visualization of nuclear substructures formed by murine repeated DNA sequences (chromocenters). In the human model, the studies in fixed and live cells have shown a non specific DNA staining but not similar structures as murine chromocenters, despite a good cellular uptake. These results suggest that structural modifications of the fluorescent probes are needed and that the accessibility of centromeric DNA sequences is potentially different, so the sequence organizations in human and mouse cannot be studied in the same way.

Table des matières

Remerciements	1
Résumé	4
Abstract :	5
Table des matières	6
Table des illustrations	11
I able des tableaux	13
Liste des abreviations	15 10
Chanitre I – Introduction	23
A - L'ADN	23
A.1 Structure de la double hélice et appariements des bases de l'ADN	23
A.2 Différents degrés de compaction de l'ADN dans la cellule	25
A.2.1 Le nucléosome et les histones	26
A.2.2 Les niveaux de compaction supérieurs	27
B - Les régions centromériques et les séquences répétées	29
B.1 Les régions centromériques	29
B.1.1 Le centromère et les régions péricentromériques : définitions et généralités	29
B.1.2 Le centromère et les régions péricentromériques : structure et fonction	31
B.2 Les deux modèles de séquences répétées étudiés dans ce travail	32
B.2.1 Chez la souris :	32
B.2.1.1 Les séquences satellites mineurs et majeurs	32
B.2.1.2 L'association des régions péricentromériques murines : les chromocentr	es
	34
B.2.2 Chez l'homme	35
B.2.2.1 Les séquences satellites	35
B.2.2.2 L'association des régions centromériques humaines	37
B.3 Pourquoi et comment étudier la structure et la dynamique de la chromatine au ni	veau
des séquences d'ADN répétées ?	37
C - Les outils pour visualiser des séquences d'ADN double brin	40
C.1 Les oligonucléotides modifiés	41
C.1.1 Les TFOS (Triplex Forming Origonucleondes)	41
C.1.2 Les modifications chimiques des oligonucléotides	42
C.2 Les protéines modifiées	44
C.3 Les polyamides ou MGBs (Minor Groove Binders) fluorescents	48
D - Les polyallilles ou MODS	33
D 2 Vers de nouveaux ligands avec des monomères alternatifs	55 58
D.2.1 Les monomères β-alanines	58
D.2.2 Le coude γ	59
D.2.3 Les monomères Hp (hydroxypyrrole)	59
D.2.4 Les monomères dérivés	60

D.2.5 Les analogues des paires de monomères	. 61
D.3 La synthèse et les optimisations des polyamides D.3.1 Les différents types de synthèse	. 63
D.3.2 L'optimisation des polyamides	66
D.3.2.1 Les polyamides cycliques	. 66
D.3.2.2 Les polyamides longs et les tandems de polyamides	. 67
D.4 L'utilisation des polyamides comme régulateurs de l'expression des gènes D.4.1 L'inhibition de la transcription par des polyamides	. 69 . 69
D.4.2 L'extinction de gènes via l'alkylation de l'ADN par des conjugués de polyamides	72
D.4.3 L'activation transcriptionnelle par l'intermédiaire des polyamides	76
D.5 Les accessibilités cellulaire et nucléaire des polyamides E – Le projet de thèse	. 78
Chapitre II - Résultats	. 87
A - Les résultats obtenus avec le modèle murin	. 87
A.1 Syntheses et caracterisations des MGBs et des sondes nuorescentes murins	. 89
A.1.1.1 Les MGBs murins standard	. 89
A.1.1.2 Les tandems de MGBs murins	. 91
A.1.2 Les sondes fluorescentes murines	. 93
A.2 Études physico-chimiques de l'interaction des MGBs murins et des sondes fluorescentes murines avec l'ADN <i>in vitro</i>	. 96
A.2.2 Dénaturation thermique de l'ADN suivie par spectroscopie UV	. 98
A.2.2.1 Dénaturation thermique des duplexes d'ADN en présence des MGBs murins	99
A.2.2.1.1 Avec les MGBs murins standard	. 99
A.2.2.1.2 Avec les tandems de MGBs murins	100
A.2.2.2 Dénaturation thermique des duplexes d'ADN en présence des sondes fluorescentes murines	103
A.2.2.3 Conclusions des expériences de dénaturation thermique avec les MGBs murins	104
A.2.3 Électrophorèses en conditions non dénaturantes (« Gel-retard »)	105
A.2.3.1 « Gels-retard » des duplexes d'ADN en présence des MGBs murins	105
A.2.3.1.1 Avec les MGBs murins standard	105
A.2.3.1.2 Avec les tandems de MGB murins	107
A.2.3.2 « Gels-retard » des duplexes d'ADN en présence des sondes fluorescente murines	es 109
A.2.3.3 Conclusions des expériences de « gel-retard » avec les MGBs murins	109
A.2.4 Dichroïsme circulaire	110
A.2.4.1 Interaction ADN / MGBs murins	110

A.2.4.1.1 Avec les MGBs murins standard	110
A.2.4.1.2 Avec les tandems de MGBs murins	112
A.2.4.2 Interaction ADN / sondes fluorescentes murines	114
A.2.4.3 Conclusions des expériences de dichroïsme circulaire	114
A.2.5 Spectroscopie de fluorescence	115
A.2.5.1 Interaction ADN / sondes fluorescentes murines	115
A.2.5.2 Transfert par résonance d'énergie de fluorescence (FRET) des sondes fluorescentes murines	117
A.2.5.3 Conclusions des expériences de fluorescence	118
A.3 Études de l'interaction des sondes fluorescentes murines avec l'ADN <i>in cellulo</i> A.3.1 Cellules murines vivantes) 119 119
A.3.1.1 Problèmes de pénétration nucléaire	119
A.3.1.2 Changements de fluorophores	121
A.3.1.3 Contrôles	123
A.3.2 Cellules murines fixées	125
A.3.3 Conclusions des expériences cellulaires	127
 A.4 Conclusions générales des résultats obtenus avec le modèle murin B - Les résultats obtenus avec le modèle humain B.1 Synthèses et caractérisations des MGBs et des sondes fluorescentes humains 	128 130 130
B 1 1 Les MGBs humains	
B.1.1 Les MGBs humains B.1.2 Les sondes fluorescentes humaines	130
 B.1.1 Les MGBs humains B.1.2 Les sondes fluorescentes humaines B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescent 	130 132 es
 B.1.1 Les MGBs humains B.1.2 Les sondes fluorescentes humaines B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescent humains avec l'ADN <i>in vitro</i> 	130 132 es 133
 B.1.1 Les MGBs humains B.1.2 Les sondes fluorescentes humaines B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescent humains avec l'ADN <i>in vitro</i> B.2.1 Dénaturation thermique de l'ADN suivie par spectroscopie UV 	130 132 es 133 133
 B.1.1 Les MGBs humains B.1.2 Les sondes fluorescentes humaines B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescent humains avec l'ADN <i>in vitro</i> B.2.1 Dénaturation thermique de l'ADN suivie par spectroscopie UV B.2.1.1 Dénaturation thermique des duplexes d'ADN en présence des MGBs humains 	130 132 es 133 133
 B.1.1 Les MGBs humains B.1.2 Les sondes fluorescentes humaines B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescent humains avec l'ADN <i>in vitro</i> B.2.1 Dénaturation thermique de l'ADN suivie par spectroscopie UV B.2.1.1 Dénaturation thermique des duplexes d'ADN en présence des MGBs humains B.2.1.2 Dénaturation thermique des duplexes d'ADN en présence des sondes fluorescentes humaines 	130 132 es 133 133 133
 B.1.1 Les MGBs humains B.1.2 Les sondes fluorescentes humaines B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescent humains avec l'ADN <i>in vitro</i> B.2.1 Dénaturation thermique de l'ADN suivie par spectroscopie UV B.2.1.1 Dénaturation thermique des duplexes d'ADN en présence des MGBs humains B.2.1.2 Dénaturation thermique des duplexes d'ADN en présence des sondes fluorescentes humaines B.2.1.2 Dénaturation thermique des duplexes d'ADN en présence des sondes fluorescentes humaines 	130 132 es 133 133 133 135
 B.1.1 Les MGBs humains B.1.2 Les sondes fluorescentes humaines B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescent humains avec l'ADN <i>in vitro</i> B.2.1 Dénaturation thermique de l'ADN suivie par spectroscopie UV B.2.1.1 Dénaturation thermique des duplexes d'ADN en présence des MGBs humains B.2.1.2 Dénaturation thermique des duplexes d'ADN en présence des sondes fluorescentes humaines B.2.1.3 Conclusions des expériences de dénaturation thermique B.2.2 Électrophorèses en conditions non dénaturantes (« Gel-retard ») 	130 132 es 133 133 133 135 136 137
 B.1.1 Les MGBs humains B.1.2 Les sondes fluorescentes humaines B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescent humains avec l'ADN <i>in vitro</i> B.2.1 Dénaturation thermique de l'ADN suivie par spectroscopie UV B.2.1.1 Dénaturation thermique des duplexes d'ADN en présence des MGBs humains B.2.1.2 Dénaturation thermique des duplexes d'ADN en présence des sondes fluorescentes humaines B.2.1.3 Conclusions des expériences de dénaturation thermique B.2.2 Électrophorèses en conditions non dénaturantes (« Gel-retard ») B.2.2 1 « Gels-retard » des duplexes d'ADN en présence des MGBs humains 	130 132 es 133 133 133 135 136 137 137
 B.1.1 Les MGBs humains B.1.2 Les sondes fluorescentes humaines B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescent humains avec l'ADN <i>in vitro</i> B.2.1 Dénaturation thermique de l'ADN suivie par spectroscopie UV B.2.1.1 Dénaturation thermique des duplexes d'ADN en présence des MGBs humains B.2.1.2 Dénaturation thermique des duplexes d'ADN en présence des sondes fluorescentes humaines B.2.1.3 Conclusions des expériences de dénaturation thermique B.2.2 Électrophorèses en conditions non dénaturantes (« Gel-retard ») B.2.2.1 « Gels-retard » des duplexes d'ADN en présence des MGBs humains	130 132 es 133 133 133 135 136 137 137 137
 B.1.1 Les MGBs humains B.1.2 Les sondes fluorescentes humaines B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescent humains avec l'ADN <i>in vitro</i> B.2.1 Dénaturation thermique de l'ADN suivie par spectroscopie UV B.2.1.1 Dénaturation thermique des duplexes d'ADN en présence des MGBs humains B.2.1.2 Dénaturation thermique des duplexes d'ADN en présence des sondes fluorescentes humaines B.2.1.3 Conclusions des expériences de dénaturation thermique B.2.2 Électrophorèses en conditions non dénaturantes (« Gel-retard ») B.2.2.1 « Gels-retard » des duplexes d'ADN en présence des MGBs humains B.2.2.2 Conclusions des expériences de « gel-retard » 	130 132 es 133 133 133 133 135 136 137 137 140 140
 B.1.1 Les MGBs humains B.1.2 Les sondes fluorescentes humaines B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescent humains avec l'ADN <i>in vitro</i> B.2.1 Dénaturation thermique de l'ADN suivie par spectroscopie UV B.2.1.1 Dénaturation thermique des duplexes d'ADN en présence des MGBs humains B.2.1.2 Dénaturation thermique des duplexes d'ADN en présence des sondes fluorescentes humaines B.2.1.3 Conclusions des expériences de dénaturation thermique B.2.2 Électrophorèses en conditions non dénaturantes (« Gel-retard ») B.2.2.1 « Gels-retard » des duplexes d'ADN en présence des MGBs humains B.2.2.2 Conclusions des expériences de « gel-retard » B.2.3 Dichroïsme circulaire B.2.3 L Interaction ADN / MGBs humains 	130 132 es 133 133 133 133 135 136 137 137 140 140 140
 B.1.1 Les MGBs humains B.1.2 Les sondes fluorescentes humaines B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescent humains avec l'ADN <i>in vitro</i> B.2.1 Dénaturation thermique de l'ADN suivie par spectroscopie UV B.2.1.1 Dénaturation thermique des duplexes d'ADN en présence des MGBs humains B.2.1.2 Dénaturation thermique des duplexes d'ADN en présence des sondes fluorescentes humaines B.2.1.3 Conclusions des expériences de dénaturation thermique B.2.2 Électrophorèses en conditions non dénaturantes (« Gel-retard ») B.2.2.1 « Gels-retard » des duplexes d'ADN en présence des MGBs humains B.2.2.2 Conclusions des expériences de « gel-retard » B.2.3 Dichroïsme circulaire B.2.3.1 Interaction ADN / MGBs humains B.2.3.2 Interaction ADN / sondes fluorescentes humaines 	130 132 es 133 133 133 133 135 136 137 137 137 140 140 140 140
 B.1.1 Les MGBs humains	130 132 es 133 133 133 133 135 136 137 137 137 140 140 143 143
 B.1.1 Les MGBs humains	130 132 es 133 133 133 133 135 136 137 137 140 140 140 143 143 144
 B.1.1 Les MGBs humains	130 132 es 133 133 133 133 135 136 137 137 140 140 140 143 143 144
 B.1.1 Les MGBs humains B.1.2 Les sondes fluorescentes humaines B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescent humains avec l'ADN <i>in vitro</i> B.2.1 Dénaturation thermique de l'ADN suivie par spectroscopie UV B.2.1.1 Dénaturation thermique des duplexes d'ADN en présence des MGBs humains B.2.1.2 Dénaturation thermique des duplexes d'ADN en présence des sondes fluorescentes humaines B.2.1.3 Conclusions des expériences de dénaturation thermique B.2.2 Électrophorèses en conditions non dénaturantes (« Gel-retard ») B.2.2.1 « Gels-retard » des duplexes d'ADN en présence des MGBs humains B.2.3 Dichroïsme circulaire B.2.3.1 Interaction ADN / MGBs humains. B.2.3.2 Interaction ADN / sondes fluorescentes humaines B.2.4 Spectroscopie de fluorescence B.2.4.1 Interaction ADN / sondes fluorescentes humaines B.2.4.2 Conclusions des expériences de fluorescence 	130 132 es 133 133 133 133 135 135 136 137 137 140 140 140 143 144 144 144

B.3.2 Cellules fixées 148 B.3.3 Conclusions des expériences cellulaires 149 B.4 Conclusions générales des résultats obtenus avec le modèle humain 150 Chapitre III - Conclusions, discussions et perspectives 155 A - Conclusions et discussions 155 La formation des complexes ADN / MGB 155 Le synthon β-ala 155 Le synthon β-ala 156 Le choix du fluorophore pour la construction de sondes fluorescentes 156 Le choix du fluorophore pour la construction de sondes fluorescentes 157 B - Perspectives 159 Le fluorophore 159 Le solyamides 160 La séquence cible d'ADN 162 Une autre utilité des molécules synthétisées 163 Chapitre IV - Matériels et méthodes 167 A - Réactifs 167 B - Tampons 168 C - Appareils 169 D - Synthèse des polyamides 173 F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard ») 174 G - Expériences de spectroscopie 176 G.1 Absorbance 176 G.2
B.3.3 Conclusions des expériences cellulaires 149 B.4 Conclusions générales des résultats obtenus avec le modèle humain 150 Chapitre III - Conclusions, discussions et perspectives 155 A - Conclusions et discussions 155 La formation des complexes ADN / MGB 155 Le synthon β-ala 155 La spécificité des MGBs 156 Le choix du fluorophore pour la construction de sondes fluorescentes 156 La pénétration cellulaire et le marquage nucléaire 157 B - Perspectives 159 Le plance cible d'ADN 160 La séquence cible d'ADN 162 Une autre utilité des molécules synthétisées 163 Chapitre IV - Matériels et méthodes 167 A – Réactifs 167 B – Tampons 168 C – Appareils 169 D – Synthèse des polyamides 173 F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard ») 174 G – Expériences de spectroscopie 176 G.1 Absorbance 176 G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sonde fluorescent de fluorescence 177
B.4 Conclusions générales des résultats obtenus avec le modèle humain 150 Chapitre III - Conclusions, discussions et perspectives 155 A - Conclusions et discussions 155 La formation des complexes ADN / MGB 155 Le synthon β-ala 155 Le synthon β-ala 155 Le synthon β-ala 156 Le choix du fluorophore pour la construction de sondes fluorescentes 156 La pénétration cellulaire et le marquage nucléaire 157 B - Perspectives 159 Le fluorophore 159 Les polyamides 160 La séquence cible d'ADN 162 Une autre utilité des molécules synthétisées 163 Chapitre IV - Matériels et méthodes 167 A - Réactifs 167 B - Tampons 168 C - Appareils 169 D - Synthèse des polyamides 170 E - Euctrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard ») Marquage fluorescent des polyamides 176 G.1 Absorbance 176 G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sonde 177 G.3 Spectres dif
Chapitre III - Conclusions, discussions et perspectives 155 A - Conclusions et discussions 155 La formation des complexes ADN / MGB 155 Le synthon β-ala 155 Le synthon β-ala 155 Le synthon β-ala 156 Le choix du fluorophore pour la construction de sondes fluorescentes 156 La pénétration cellulaire et le marquage nucléaire 157 B - Perspectives 159 Le fluorophore 159 Les polyamides 160 La séquence cible d'ADN 162 Une autre utilité des molécules synthétisées 163 Chapitre IV - Matériels et méthodes 167 A - Réactifs 167 B - Tampons 168 C - Appareils 169 D - Synthèse des polyamides 173 F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard ») 174 G - Expériences de spectroscopie 176 G.1 Absorbance 176 G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sonde 177 G.3 Spectres différentiels de dichroïsme circulaire 178 G.4 Spectres de f
La formation des complexes ADN / MGB. 155 La softinité des MGBs 155 La spécificité des MGBs 156 Le choix du fluorophore pour la construction de sondes fluorescentes 156 La pénétration cellulaire et le marquage nucléaire. 157 B - Perspectives 159 Le fluorophore 159 Les polyamides. 160 La séquence cible d'ADN 162 Une autre utilité des molécules synthétisées 163 Chapitre IV - Matériels et méthodes 167 A - Réactifs 167 B - Tampons 168 C - Appareils 169 D - Synthèse des polyamides 170 E - Marquage fluorescent des polyamides 173 F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard ») 174 G - Expériences de spectroscopie 176 G.1 Absorbance 177 G.3 Spectres différentiels de dichroïsme circulaire 178 G.4 Spectres de fluorescence 180 H - Études Cellulaires 181 H 1.1 Culture cellulaire 181 H 1.1 Cellules murines NIH-3T3 et MEF 181
Le synthon β-ala 155 La spécificité des MGBs 156 Le choix du fluorophore pour la construction de sondes fluorescentes 156 La pénétration cellulaire et le marquage nucléaire 157 B - Perspectives 159 Le fluorophore 159 Le spolyamides 160 La séquence cible d'ADN 162 Une autre utilité des molécules synthétisées 163 Chapitre IV - Matériels et méthodes 167 A - Réactifs 167 B - Tampons 168 C - Appareils 169 D - Synthèse des polyamides 170 E - Marquage fluorescent des polyamides 173 F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard ») 174 G - Expériences de spectroscopie 176 G.1 Absorbance 176 G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sonde fluorescente 177 G.3 Spectres différentiels de dichroïsme circulaire 178 G.4 Spectres de fluorescence 180 H - Études Cellulaires 181 H 1.1 Culture cellulaire 181 H 1.1 Cellules murine
La spécificité des MGBs
Le choix du fluorophore pour la construction de sondes fluorescentes
La pénétration cellulaire et le marquage nucléaire. 157 B - Perspectives 159 Le fluorophore 159 Les polyamides 160 La séquence cible d'ADN 162 Une autre utilité des molécules synthétisées 163 Chapitre IV - Matériels et méthodes 167 A - Réactifs 167 B - Tampons 168 C - Appareils 169 D - Synthèse des polyamides 170 E - Marquage fluorescent des polyamides 173 F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard ») 174 G - Expériences de spectroscopie 176 G.1 Absorbance 177 G.3 Spectres différentiels de dichroïsme circulaire 177 G.3 Spectres différentiels de dichroïsme circulaire 178 G.4 Spectres de fluorescence 180 H - Études Cellulaires 181 H 1.1 Cellules murines NIH-3T3 et MEF 181 H 1.2 Cellules humaines RPE-1 182
B - Perspectives 159 Le fluorophore 159 Les polyamides 160 La séquence cible d'ADN 162 Une autre utilité des molécules synthétisées 163 Chapitre IV - Matériels et méthodes 167 A - Réactifs 167 B - Tampons 168 C - Appareils 169 D - Synthèse des polyamides 170 E - Marquage fluorescent des polyamides 173 F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard ») 174 G - Expériences de spectroscopie 176 G.1 Absorbance 177 G.3 Spectres différentiels de dichroïsme circulaire 177 G.3 Spectres différentiels de dichroïsme circulaire 180 H - Études Cellulaires 181 H.1 Culture cellulaire 181 H 1.1 Cellules murines NIH-3T3 et MEF 181 H 1.2 Cellules humaines RPE-1 182
Le fluorophore 159 Les polyamides 160 La séquence cible d'ADN 162 Une autre utilité des molécules synthétisées 163 Chapitre IV - Matériels et méthodes 167 A - Réactifs 167 B - Tampons 168 C - Appareils 169 D - Synthèse des polyamides 170 E - Marquage fluorescent des polyamides 173 F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard ») 174 G - Expériences de spectroscopie 176 G.1 Absorbance 176 G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sonde 177 G.3 Spectres différentiels de dichroïsme circulaire 178 G.4 Spectres de fluorescence 180 H - Études Cellulaires 181 H.1 Culture cellulaire 181 H 1.1 Cellules murines NIH-3T3 et MEF 181 H.1.2 Cellules humaines RPE-1 182
Les polyamides. 160 La séquence cible d'ADN 162 Une autre utilité des molécules synthétisées 163 Chapitre IV - Matériels et méthodes 167 A – Réactifs 167 B – Tampons 168 C - Appareils 169 D - Synthèse des polyamides 170 E - Marquage fluorescent des polyamides 173 F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard ») 174 G – Expériences de spectroscopie 176 G.1 Absorbance 176 G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sonde 177 G.3 Spectres différentiels de dichroïsme circulaire 178 G.4 Spectres de fluorescence 180 H – Études Cellulaires 181 H.1 Culture cellulaire 181 H 1.1 Cellules murines NIH-3T3 et MEF 181 H.1.2 Cellules humaines RPE-1 182
La séquence cible d'ADN
Une autre utilité des molécules synthétisées163Chapitre IV - Matériels et méthodes167A - Réactifs167B - Tampons168C - Appareils169D - Synthèse des polyamides170E - Marquage fluorescent des polyamides173F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard »)174G - Expériences de spectroscopie176G.1 Absorbance176G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sondefluorescente177G.3 Spectres différentiels de dichroïsme circulaire178G.4 Spectres de fluorescence180H - Études Cellulaires181H.1 Culture cellulaire181H 1.1 Cellules murines NIH-3T3 et MEF181H.1.2 Cellules humaines RPE-1182
Chapitre IV - Matériels et méthodes 167 A - Réactifs 167 B - Tampons 168 C - Appareils 169 D - Synthèse des polyamides 170 E - Marquage fluorescent des polyamides 173 F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard ») 174 G - Expériences de spectroscopie 176 G.1 Absorbance 176 G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sonde 177 G.3 Spectres différentiels de dichroïsme circulaire 178 G.4 Spectres de fluorescence 180 H - Études Cellulaires 181 H.1 Culture cellulaire 181 H 1.1 Cellules murines NIH-3T3 et MEF 181 H.1.2 Cellules humaines RPE-1 182
A - Réactifs167B - Tampons168C - Appareils169D - Synthèse des polyamides170E - Marquage fluorescent des polyamides173F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard »)174G - Expériences de spectroscopie176G.1 Absorbance176G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sondefluorescente177G.3 Spectres différentiels de dichroïsme circulaire178G.4 Spectres de fluorescence180H - Études Cellulaires181H.1 Culture cellulaire181H 1.1 Cellules murines NIH-3T3 et MEF181H.1.2 Cellules humaines RPE-1182
B – Tampons 168 C - Appareils 169 D - Synthèse des polyamides 170 E - Marquage fluorescent des polyamides 173 F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard ») 174 G - Expériences de spectroscopie 176 G.1 Absorbance 176 G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sonde 177 G.3 Spectres différentiels de dichroïsme circulaire 178 G.4 Spectres de fluorescence 180 H – Études Cellulaires 181 H.1 Culture cellulaire 181 H 1.1 Cellules murines NIH-3T3 et MEF 181 H.1.2 Cellules humaines RPE-1 182
C - Appareils
D - Synthese des polyamides
E - Marquage nuorescent des polyamides 173 F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de « gel-retard ») 174 G - Expériences de spectroscopie 176 G.1 Absorbance 176 G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sonde 177 G.3 Spectres différentiels de dichroïsme circulaire 178 G.4 Spectres de fluorescence 180 H – Études Cellulaires 181 H.1 Culture cellulaire 181 H 1.1 Cellules murines NIH-3T3 et MEF 181 H.1.2 Cellules humaines RPE-1 182
1 - Electrophonese sur ger de poryaer ynamide non denaturant (experiences de « ger-retard »)
G – Expériences de spectroscopie 176 G.1 Absorbance 176 G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sonde 177 G.3 Spectres différentiels de dichroïsme circulaire 178 G.4 Spectres de fluorescence 180 H – Études Cellulaires 181 H.1 Culture cellulaire 181 H 1.1 Cellules murines NIH-3T3 et MEF 181 H.1.2 Cellules humaines RPE-1 182
G.1 Absorbance176G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sonde177fluorescente177G.3 Spectres différentiels de dichroïsme circulaire178G.4 Spectres de fluorescence180H – Études Cellulaires181H.1 Culture cellulaire181H 1.1 Cellules murines NIH-3T3 et MEF181H.1.2 Cellules humaines RPE-1182
G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sonde fluorescente 177 G.3 Spectres différentiels de dichroïsme circulaire 178 G.4 Spectres de fluorescence 180 H – Études Cellulaires 181 H.1 Culture cellulaire 181 H 1.1 Cellules murines NIH-3T3 et MEF 181 H.1.2 Cellules humaines RPE-1 182
fluorescente
G.3 Spectres différentiels de dichroïsme circulaire
 G.4 Spectres de fluorescence
H – Etudes Cellulaires
H 1.1 Cellules murines NIH-3T3 et MEF
H.1.2 Cellules humaines RPE-1
H.1.3 Cellules humaines GM12878
H 2 Marquage in situ des cellules fixées 183
H.3 Acquisition des images par microscopie
Liste des publications scientifiques co-signées
Annexes
Annexes
Annexes
Annexes 191 Annexe 1 : Structures en « épingle à cheveux antiparallèle » et caractéristiques chimiques des MGBs murins 191 Annexe 2 : Caractérisation des MGBs murins 195

Annexe 4 : Caractérisation des tandems de polyamides murins
Annexe 5 : Structures et caractéristiques chimiques des différents fluorophores utilisés . 199
Annexe 6 : Caractéristiques des différentes sondes murines synthétisées à partir des MGBs
standard
Annexe 7 : Caractérisation des sondes murines construites à partir des MGBs standard 202
Annexe 8 : Caractéristiques des sondes fluorescentes murines construites à partir des
tandems murins
Annexe 9 : Courbes de fusion du Duplexe 1 et ses dérivées premières en présence et en
absence du polyamide murin F1
Annexe 10 : Courbes de fusion du Duplexe 1 et ses dérivées premières en présence et en
absence du tandem murin T1
Annexe 11 : Exemples de « gel-retard » avec le duplexe contrôle et les MGBs murins 205
Annexe 12 : Spectres de dichroïsme circulaire différentiels des complexes ADN / tandem
de polyamide murin avec les duplexes courts mutés sur un site
Annexe 13 : Spectres de dichroïsme circulaire différentiels des complexes ADN / sonde
fluorescente murine
Annexe 14 : Expériences de FRET entre les sondes murines F1-Cy3 et F2β-Cy5 en
présence du duplexe cible d'ADN
Annexe 15 : Structures en « épingle à cheveux antiparallèle » et caractéristiques chimiques
des MGBs humains
Annexe 16 : Caractérisation des MGBs humains212
Annexe 17 : Caractéristiques des différentes sondes humaines synthétisées à partir des
MGBs
Annexe 18 : Caractérisation des sondes humaines
Annexe 19 : Spectres de dichroïsme circulaire différentiels des complexes ADN / MGB
humain
Annexe 20 : Spectres de dichroïsme circulaire différentiels des complexes ADN / sonde
fluorescente humaine
Annexe 21 : Expériences de colocalisation sur des cellules humaines RPE-1 fixées 217
Bibliographie

Table des illustrations

Figure 1 : Nucléobases, appariements canoniques et non canoniques	. 24
Figure 2 : Les différents niveaux de compaction de la chromatine	. 26
Figure 3 : Organisation des régions centromériques de différents organismes	. 30
Figure 4 : Organisation des régions centromériques murines : le centromère et les régions	
péricentromériques	. 33
Figure 5 : Association des régions péricentromériques chez la souris	. 35
Figure 6 : Organisation des régions centromériques humaines : les séquences satellites	. 36
Figure 7 : La triple hélice	. 42
Figure 8 : Les différentes modifications chimiques des oligonucléotides	. 43
Figure 9 : Les différentes étapes nécessaires à la visualisation de l'ADN par une protéine d	le
fusion	. 45
Figure 10 : Représentation schématique d'un motif à doigt de zinc Cys ₂ His ₂	. 46
Figure 11 : Représentation schématique d'une protéine TALE	. 47
Figure 12 : Le système CRISPR/Cas9 pour la visualisation de séquences d'ADN	. 48
Figure 13 : Exemples de sondes fluorescentes à base de polyamides	. 50
Figure 14 : L'environnement électronique du petit sillon de l'ADN	. 54
Figure 15 : Structures et modes d'interaction des petites molécules dans le petit sillon de la	1
double hélice d'ADN	. 56
Figure 16 : Règles de reconnaissance dérivées d'un complexe polyamide : ADN (2:1)	. 57
Figure 17: Représentation schématique d'un complexe polyamide/ADN	. 60
Figure 18 : Structures des monomères, paires de monomères et de leurs dérivés	. 62
Figure 19 : Modèle schématique d'un polyamide en « épingle à cheveux » antiparallèle en	
complexe avec le duplexe cible d'ADN	. 64
Figure 20 : Les différents variants des queues C-term des polyamides	. 65
Figure 21 : Exemples de polyamides cycliques	. 00
Figure 22 : Representation schematique des tandems et du trimère de polyamides	. 00
Figure 25 : Les polyamides comme des agents arkylants	. 73
contrôlable	75
Figure 25 • Exemples de facteurs de transcription artificiels concus avec des polyamides	. 73 77
Figure 26 : Exemples de facteurs de transcription artificiers conçus avec des poryanides	
dexaméthasone	80
Figure 27 : Représentation schématique des différents duplexes utilisés et des polyamides	. 00
murins synthétisés	. 89
Figure 28 : Schéma de la synthèse peptidique manuelle des polyamides sur support solide .	. 91
Figure 29 : Schéma de la synthèse et structures des tandems de polyamides	. 92
Figure 30 : Méthodes de couplages des fluorophores aux MGBs	. 94
Figure 31 : Formules des tandems fluorescents murins	. 95
Figure 32 : Image de l'expérience « d'arrêt de retard sur gel » pour étudier la cinétique de	
formation d'un complexe ADN / MGB	. 97
Figure 33 : Courbes de fusion des duplexes d'ADN et leurs dérivées premières en présence	e et
en absence du polyamide murin F4 β	100
Figure 34 : Courbes de fusion des duplexes d'ADN et leurs dérivées premières en présence	e et
en absence du tandem T2 1	101
Figure 35 : Représentation schématique des duplexes courts construits pour les études de	
l'interaction du tandem T2 avec l'ADN1	102
Figure 36 : Images des « gels-retard » des complexes ADN / MGB murins 1	106

Figure 37 : Images des « gels-retard » des complexes ADN / tandem murins
Figure 41 : Expériences de FRET entre les sondes murines F1-Cy3 et F4β-Cy5 en présence du duplexe cible d'ADN 118 Figure 42 : Images de cellules murines vivantes NIH-3T3 traitées par F1-Cy3 et F4β-Cy5 120 120 Figure 43 : Images de cellules murines vivantes NIH-3T3 traitées par les sondes marquées par la fluorescéine et son analogue isothiocyanate (FITC) 122 Figure 44 : Images de cellules murines vivantes traitées par des sondes couplées à la FITC 124
Figure 45 : Expériences de colocalisation sur des cellules murines NIH-3T3 fixées
Figure 55 : Expériences de colocalisation sur des cellules humaines RPE-1 fixées

Table des tableaux

Tableau 1 : Règles de reconnaissance pour cibler les différentes paires de bases de l'ADN avec les Py-Im polyamides
Tableau 2 : Augmentations de la température de dissociation du Duplexe 2 en présence de 6 équivalents des différents polyamides murins déterminées par dénaturation thermique suivie par UV ($\Delta T_m \pm 0.5^{\circ}$ C)
Tableau 3 : Augmentations de la température de dissociation des duplexes d'ADN enprésence de 6 équivalents des tandems de polyamides murins déterminées par dénaturationthermique suivie par UV ($\Delta T_m \pm 0.5^{\circ}$ C)103
Tableau 4 : Augmentations de la température de dissociation du Duplexe 2 en présence de 6 équivalents des sondes fluorescentes murines composées par des polyamides standard, déterminées par dénaturation thermique suivie par UV ($\Delta T_m \pm 0.5^{\circ}$ C)104
Tableau 5 : Constantes de dissociation apparentes ($K_{d,app}$) des complexes entre le Duplexe 2 et les différents polyamides, déterminées par des expériences de retard sur gel de polyacrylamide. Les augmentations de la température de fusion du Duplexe 2 en présence de 6 équivalents des MGBs sont données pour rappel ($\Delta T_m \pm 0.5^{\circ}$ C) 107
Tableau 6 : Constantes de dissociation apparentes ($K_{d,app}$) des complexes entre les duplexes d'ADN et les tandems de polyamides murins déterminées par des expériences de retard sur gel de polyacrylamide et les augmentations de la température de fusion des duplexes en présence de 6 équivalents des tandems sont données pour rappel ($\Delta T_m \pm 0.5^{\circ}$ C) 109
Tableau 7 : Augmentations des intensités de fluorescence des sondes murines $(2 \ \mu M)$ en présence des duplexes d'ADN $(2 \ \mu M)$ cible (Duplexe 1) et contrôle (Duplexe 3) 117
Tableau 8 : Augmentations de la température de dissociation des duplexes cible (Duplexe 7) et contrôle (Duplexe 3) en présence de 6 équivalents des différents polyamides humains déterminées par dénaturation thermique suivie par UV ($\Delta T_m \pm 0.5^{\circ}$ C)
Tableau 9 : Augmentations de la température de dissociation des duplexes cible (Duplexe 7) et contrôle (Duplexe 3) en présence de 6 équivalents des différentes sondes humaines déterminées par dénaturation thermique suivie par UV ($\Delta T_m \pm 0.5^{\circ}$ C)
Tableau 10 : Constantes de dissociation apparentes ($K_{d,app}$) des complexes entre le Duplexe 7 et les différents polyamides humains, déterminées par des expériences de retard sur gel de polyacrylamide. Les augmentations de la température de fusion du Duplexe 7 en présence de 6 équivalents des MGBs sont données pour rappel ($\Delta T_m \pm 0.5^{\circ}$ C)
Tableau 11 : Augmentations des intensités de fluorescence des sondes humaines $(2 \ \mu M)$ en présence des duplexes d'ADN $(2 \ \mu M)$ cible (Duplexe 7) et contrôle (Duplexe 3) 145
Tableau 12 : Liste des oligonucléotides utilisés dans ce manuscrit

Tableau 13 : Les différents réactifs pour la synthèse du MGB F4β de séquence Nterm-γ2- Py8-Py7-Py6-Py5-γ1-Py4-Py3-βala2-Im1-β-Dp-Cterm
Tableau 14 : Les différents filtres de microscopie utilisés pour l'acquisition des images 185
Tableau 15 : Caractéristiques des MGBs murins fonctionnalisés en N-term et des tandems murins
Tableau 16 : Noms, structures, masses molaires (g/mol) et longueurs d'onde d'excitation (λ_{ex}, nm) et d'émission (λ_{em}, nm) des différents fluorophores
Tableau 17 : Caractéristiques des sondes murines composées par des MGBs standard 201
Tableau 18 : Caractéristiques des tandems fluorescents murins
Tableau 19 : Caractéristiques des sondes humaines synthétisées à partir des MGBs

Liste des abréviations

Α	Adénine
ADN db	Acide DésoxyriboNucléique double brin
ADN	Acide DésoxyriboNucléique
ARN	Acide RiboNucléique
ARNm	Acide RiboNucléique messager
ARNT	Aryl hydrocarbon Receptor Nuclear Translocator
Boc	tert-Butoxycarbonyle
Boc-βala-PAM	Résine 4-hydroxyméthyle-Phényl Acétamido Méthyle estérifiée par une Boc- β -alanine
С	Cytosine
Cas9	CRISPR-associated nuclease (nucléase associée au système CRISPR)
CDE	Centromeric DNA Element (élément d'ADN centromérique)
CenpA	Centromeric protein A (protéine centromérique A)
CFP	Cyan Fluorescent Protein (protéine fluorescente cyan)
ChiP	Chromatin immunoPrecipitation (immunoprécipitation de la chromatine)
CMV	CytoMégaloVirus
COMBO-FISH	COMBinatorial Oligo Fluorescence In Situ Hybridization (hybridation de fluorescence in situ d'une combinatoire d'oligonucléotides)
	Clustered Degularly Interspeced Delindromic Deposts (courtes
CRISPR	répétitions en palindrome regroupées et régulièrement espacées)
CRISPR Ct	répétitions en palindrome regroupées et régulièrement espacées) Chlorothiophène
CRISPR Ct C-term	répétitions en palindrome regroupées et régulièrement espacées) Chlorothiophène C-terminal(e)
CRISPR Ct C-term CTGF	répétitions en palindrome regroupées et régulièrement espacées) Chlorothiophène C-terminal(e) Connective Tissue Growth Factor (facteur de croissance du tissu conjonctif)
CRISPR Ct C-term CTGF Cy3	répétitions en palindrome regroupées et régulièrement espacées) Chlorothiophène C-terminal(e) Connective Tissue Growth Factor (facteur de croissance du tissu conjonctif) Cyanine 3
CRISPR Ct C-term CTGF Cy3 Cy3-ITC	répétitions en palindrome regroupées et régulièrement espacées) Chlorothiophène C-terminal(e) Connective Tissue Growth Factor (facteur de croissance du tissu conjonctif) Cyanine 3 Cyanine 3 fonctionnalisée par un groupement IsoThioCyanate
CRISPR Ct C-term CTGF Cy3 Cy3-ITC Cy3-sulf	répétitions en palindrome regroupées et régulièrement espacées) Chlorothiophène C-terminal(e) Connective Tissue Growth Factor (facteur de croissance du tissu conjonctif) Cyanine 3 Cyanine 3 fonctionnalisée par un groupement IsoThioCyanate Cyanine 3 possédant des groupements sulfate
CRISPR Ct C-term CTGF Cy3 Cy3-ITC Cy3-sulf Cy5	répétitions en palindrome regroupées et régulièrement espacées) Chlorothiophène C-terminal(e) Connective Tissue Growth Factor (facteur de croissance du tissu conjonctif) Cyanine 3 Cyanine 3 fonctionnalisée par un groupement IsoThioCyanate Cyanine 3 possédant des groupements sulfate Cyanine 5
CRISPR Ct C-term CTGF Cy3 Cy3-ITC Cy3-sulf Cy5 Cy5-sulf	répétitions en palindrome regroupées et régulièrement espacées) Chlorothiophène C-terminal(e) Connective Tissue Growth Factor (facteur de croissance du tissu conjonctif) Cyanine 3 Cyanine 3 fonctionnalisée par un groupement IsoThioCyanate Cyanine 3 possédant des groupements sulfate Cyanine 5 Cyanine 5 possédant des groupements sulfate
CRISPR Ct C-term CTGF Cy3 Cy3-ITC Cy3-sulf Cy5 Cy5-sulf Cys	répétitions en palindrome regroupées et régulièrement espacées) Chlorothiophène C-terminal(e) Connective Tissue Growth Factor (facteur de croissance du tissu conjonctif) Cyanine 3 Cyanine 3 fonctionnalisée par un groupement IsoThioCyanate Cyanine 3 possédant des groupements sulfate Cyanine 5 Cyanine 5 possédant des groupements sulfate Cystéine
CRISPR Ct C-term CTGF Cy3 Cy3-ITC Cy3-sulf Cy5 Cy5-sulf Cy5 Sulf Cy5 Sulf Cy5	répétitions en palindrome regroupées et régulièrement espacées) Chlorothiophène C-terminal(e) Connective Tissue Growth Factor (facteur de croissance du tissu conjonctif) Cyanine 3 Cyanine 3 fonctionnalisée par un groupement IsoThioCyanate Cyanine 3 possédant des groupements sulfate Cyanine 5 Cyanine 5 possédant des groupements sulfate Cystéine 4'-6'-DiAmidino-2-PhénylIndole
CRISPR Ct C-term CTGF Cy3 Cy3-ITC Cy3-sulf Cy5 Cy5-sulf Cy5 DAPI DBD	répétitions en palindrome regroupées et régulièrement espacées) Chlorothiophène C-terminal(e) Connective Tissue Growth Factor (facteur de croissance du tissu conjonctif) Cyanine 3 Cyanine 3 fonctionnalisée par un groupement IsoThioCyanate Cyanine 3 possédant des groupements sulfate Cyanine 5 Cyanine 5 possédant des groupements sulfate Cystéine 4'-6'-DiAmidino-2-PhényIIndole DNA Binding Domain (domaine de liaison à l'ADN)

N, N'-DiCyclohexylCarbodiimide
N, N-DiIsopropylEthylAmine
Dulbecco's Modified Eagle Medium
DiMéthylFormamide
N, N-DiMéthylaminoPropylAmine
Epstein Barr Virus (Virus Epstein Barr)
EthyleneDiamineTetraacetic Acid (Acide Ethylène Diamine Tétraacétique)
Electrophoretic Mobility Shift Assay (retard sur gel)
Équivalents
Fluorescent In Situ Hybridization (hybridation de fluorescence in situ)
Fluorescéine-5-IsoThioCyanate
5-carboxyfluorescéine
Fluorénylméthyloxycarbonyle
Fluorescence Resonance Energy Transfer (Transfert par Résonance d'Energie de Fluorescence)
Facteurs de Transcription
Guanine
Green Fluorescent Protein (protéine fluorescente verte)
Glucocorticoïd Receptor Ligand Binding Domain (domaine de liaison du ligand du récepteur au glucocorticoïde)
O-(7-azabenzotriazol-1-yl)-1, 1, 3, 3-tétraméthyluronium- hexafluorophosphate
Acide chlorhydrique
Histone DéACétylase
Hypoxia Inducible Factor 1 (facteur 1 induit par l'hypoxie)
Histidine
1-HydrOxy-7-Azabenzotriazole
N-HydrOxyBenzotriazole
High-Order Repeat (répétitions d'ordre supérieur)
Hydroxypyrrole
Heterochromatin Protein 1 (protéine de l'hétérochromatine)
High Performance Liquid Chromatography (chromatographie en phase liquide à haute performance)
2-Hydroxypropyl-β-cyclodextrine
Hypoxia Response Element (élément de réponse à l'hypoxie)
Heat shock protein 90 (protéine de choc thermique 90)

Hx	2-(4-méthoxyphényl)-6-carboxamide
Im	N-méthylImidazole
ITC	IsoThioCyanate
kb	kilobases
K _{d,app}	Constante de dissociation apparente
kDa	kiloDalton
LNA	Locked Nucleic Acid (acide nucléique bloqué)
Μ	Molaire
MajSat	Séquence de la répétition satellite majeur
Mb	Mégabases
mdeg	milli-degrés
MGB	Minor Groove Binder (ligand du petit sillon)
mM	milliMolaire
MM14	Coumarine MM14
nm	nanomètre
N-term	N-terminal(e)
OG-ITC	Oregon Green 488 fonctionnalisé par un groupement IsoThioCyanate
OG-NHS	Oregon Green 488 sous forme d'ester activé (ester N- HydroxySuccinimidyle)
pb	paire de bases
PBS	Phosphate Buffered Saline (tampon phosphate salin)
PCR	Polymerase Chain Reaction (réaction en chaîne par polymérase)
PFA	ParaFormAldéhyde
Ру	N-méthylPyrrole
RMN	Résonance Magnétique Nucléaire
SAHA	SuberoylAnilide Hydroxamic Acid (acide suberoylanilide hydroxamique)
SDS	Sodium DodecylSulfate (dodécylsulfate de sodium)
seco-CBI	seco-1-Chlorométhyle-5-hydroxy-1,2-dihydro-3H-Benz[e]Indole
SM	Spectrométrie de Masse
SPR	Surface Plamon Resonance (résonance plasmonique de surface)
SVF	Sérum de Veau Fœtal
Т	Thymine
TALE	Transcription Activator Like Effector (effecteur de type activateur de transcription)

TALEN	Transcription Activator Like Effector Nuclease (nucléase effectrice de type activateur de transcription)
TAMRA	5-carboxy-TétrAMéthylRhodAmine
TBE	Tampon Tris Borate EDTA
TEMED	N, N, N', N'-TEtraMéthylEthylèneDiamine
TFA	TriFluoroacetic Acid (acide trifluoroacétique)
TFO	Triplex Forming Oligonucleotide (oligonucléotide formant un triplexe)
TGFβ	Transforming Growth Factor β (facteur de croissance transformant β)
TISH	Triplex In Situ Hybridization (hybridation de triplexe in situ)
T _m	Température de fusion
TMR	TétraMéthylRhodamine
TMR-ITC	TétraMéthylRhodamine fonctionnalisée par un groupement IsoThioCyanate
TMR-NHS	TétraMéthylRhodamine sous forme d'ester activé (ester N- HydroxySuccinimidyle)
ТО	Thiazole Orange
TPW	Solution de déprotection (TFA 92,5% / phénol 5% / eau 2,5%)
Tr	Temps de rétention sur HPLC
U.R.	Unité Relative
VEGF	Vascular Endothelial Growth Factor (facteur de croissance vasculaire endothéliale)
VIH	Virus de l'Immunodéficience Humaine
YFP	Yellow Fluorescent Protein (protéine fluorescente jaune)
β	β-alanine
γ	GABA (acide γ-aminobutyrique)
$\gamma^{\rm NH}$	DABA (acide α,γ-diaminobutyrique)
λem	Longueur d'onde d'émission
λex	Longueur d'onde d'excitation
μL	Microlitre
μΜ	Micromolaire

Les objectifs

Chez la plupart des organismes vivants, l'ADN ou acide désoxyribonucléique est le support de l'information génétique. Sa structure en double hélice et ses niveaux de compaction dans la cellule jouent un rôle majeur dans le transfert de cette information de génération en génération. Pendant la division cellulaire, l'assemblage du fuseau mitotique est essentiel à la bonne ségrégation des chromatides sœurs afin que les cellules filles héritent de l'information génétique nécessaire à leur développement et soient viables. Ce fuseau s'assemble au niveau de la constriction primaire des chromosomes qui est appelée « centromère ». Malgré un rôle fonctionnel conservé (assemblage du fuseau mitotique) dans toutes les espèces, les séquences présentes dans cette région ne sont pas conservées : il s'agit du « paradoxe du centromère ».

Au niveau des régions centromériques, les séquences d'ADN sont répétées et connues pour avoir une organisation nucléaire dynamique mais la signification de cette organisation est encore mal connue. Chez la souris, une répétition particulière est la séquence satellite majeur MajSat qui est située au niveau des régions péricentromériques des chromosomes. Elle a la capacité de s'associer en interphase pour former des « chromocentres » qui sont des structures visibles par microscopie de fluorescence dans les cellules murines fixées quand l'ADN est marqué par le DAPI. Chez l'homme, ce sont les séquences α -satellites, situées au niveau des régions centromériques, qui sont les plus abondantes.

L'organisation nucléaire de l'ADN est un facteur important dans le contrôle de l'expression génique et deux aspects sont à retenir pour étudier cette organisation : l'aspect régulation de l'expression des gènes et l'aspect visualisation. C'est sur ce dernier aspect que porte ce travail de thèse. En effet, la possibilité de visualiser des séquences spécifiques d'ADN dans des cellules vivantes permet d'observer, en temps réel, la dynamique de l'ADN dans le noyau en utilisant la microscopie de fluorescence. Mais comment visualiser des séquences d'ADN spécifiques directement dans des cellules vivantes ? À cette fin, différents outils comme les oligonucléotides modifiés ou les protéines modifiées ont été utilisés mais ils possèdent chacun leurs limites. Au cours de ces dernières années, il a été développé un autre type d'outil moléculaire capable de reconnaitre des séquences spécifiques d'ADN : les polyamides N-méthylpyrrole et N-méthylimidazole, aussi appelés « ligands du petit sillon de

l'ADN » ou « MGBs » (Minor Groove Binders) qui peuvent interagir avec des séquences spécifiques d'ADN double brin au sein du petit sillon.

L'objectif de ces travaux a été le développement de polyamides pour étudier l'organisation et la dynamique des séquences répétées centromériques dans des cellules vivantes murines et humaines par la microscopie de fluorescence, sans les fixer ni les perméabiliser. Les questions posées étaient : est-il possible de visualiser les chromocentres murins directement dans des cellules vivantes en utilisant des sondes fluorescentes composées par des polyamides ? Et chez l'homme, est-il possible de visualiser les séquences α -satellite et existe-t-il des structures apparentées aux chromocentres murins dans les cellules vivantes?

Pour atteindre ce but, la synthèse, la purification et la caractérisation de sondes fluorescentes hybrides composées de polyamides couplés à des fluorophores ont été effectuées pour cibler des répétitions centromériques murines et humaines. La synthèse de ces polyamides, en suivant les règles de reconnaissances établies par Dervan a été entreprise en optimisant, d'une part le motif de liaison à l'ADN (polyamides standard ou tandems de polyamides ou longs polyamides) et d'autre part le signal fluorescent émis (utilisation de fluorophores capables de moduler leur spectre de fluorescence en fonction de l'interaction avec l'ADN ou permettant un effet FRET (Transfert par Résonance d'Energie de Fluorescence)).

L'ensemble de ces travaux est reporté dans le deuxième chapitre de ce manuscrit. Avant d'aborder cette partie, une introduction générale est également proposée. Elle contient une revue bibliographique non-exhaustive de l'état de la science dans les différents domaines ayant trait à cette thèse. Tout d'abord, après quelques généralités sur l'ADN et ses propriétés, les centromères et les séquences répétées des deux modèles utilisés seront décrits. Ensuite, les outils actuels pour pouvoir visualiser des séquences d'ADN dans les cellules seront présentés. Enfin, les polyamides qui constituent les molécules de choix de ces travaux seront décrits en détail.

Chapitre I - Introduction

Chapitre I – Introduction

Chapitre I – Introduction	
A - L'ADN	
A.1 Structure de la double hélice et appariements des bases de l'ADN	
A.2 Différents degrés de compaction de l'ADN dans la cellule	
B - Les régions centromériques et les séquences répétées	
B.1 Les régions centromériques	
B.2 Les deux modèles de séquences répétées étudiés dans ce travail	
B.3 Pourquoi et comment étudier la structure et la dynamique de la chromatine au	ı niveau
des séquences d'ADN répétées ?	
C - Les outils pour visualiser des séquences d'ADN double brin	
C.1 Les oligonucléotides modifiés	
C.2 Les protéines modifiées	
C.3 Les polyamides ou MGBs (Minor Groove Binders) fluorescents	
D - Les polyamides ou MGBs	
D.1 L'histoire des polyamides et l'établissement des règles de reconnaissance	
D.2 Vers de nouveaux ligands avec des monomères alternatifs	
D.3 La synthèse et les optimisations des polyamides	
D.4 L'utilisation des polyamides comme régulateurs de l'expression des gènes	
D.5 Les accessibilités cellulaire et nucléaire des polyamides	
E – Le projet de thèse	

<u>A - L'ADN</u>

A.1 Structure de la double hélice et appariements des bases de l'ADN

L'acide désoxyribonucléique (ADN) qui constitue le support de l'information génétique est une macromolécule biologique présente dans chaque cellule de la plupart des êtres vivants et localisée dans le noyau des cellules eucaryotes ou dans le cytoplasme des cellules procaryotes. Son squelette est composé par la succession de sucres 2-désoxyribose connectés par des liaisons phosphodiesters qui s'établissent entre les positions 3' et 5' des sucres. Une des 4 bases, adénine (A), cytosine (C), guanine (G) ou thymine (T), est attachée au sucre en position 1'. A et G sont des bases puriques et C et T des bases pyrimidiques. Sa structure proposée par James Watson et Francis Crick en 1953 (Franklin & Gosling 1953;

Watson & Crick 1953a; Watson & Crick 1953b) est une hélice droite formée par deux brins antiparallèles, appariés par 3 liaisons hydrogène entre les bases G/C et 2 liaisons hydrogène entre les bases A/T (**Figures 1A et B**). Cette structure hélicoïdale est stabilisée par des interactions d'empilement de type π entre les bases. Nommée ADN de forme B, elle correspond à la forme canonique retrouvée majoritairement dans les cellules.



Figure 1 : Nucléobases, appariements canoniques et non canoniques

(A) Formules chimiques des 4 bases constituant l'ADN : A et G - les bases puriques et T et C - les bases pyrimidiques. (B) Schéma de la double hélice d'ADN (Montgomery 2013) et appariements de type Watson-Crick entre A/T et G/C. (C) En noir, exemples d'appariements de type Hoogsteen entre $A \times T$, $G \times C^+$, permettant la formation de triple hélices d'ADN.

Chapitre I – Introduction L'ADN

Le type d'appariements entre les bases nucléotidiques dans la forme B (appariements dits Watson-Crick) n'est pourtant pas la seule conformation pouvant être adoptée géométriquement par les bases nucléotidiques (Donohue 1956). En effet, chaque nucléobase possède de multiples sites accepteurs et donneurs de liaisons hydrogène. Ainsi, quatre ans après la découverte de la structure de la double hélice, G. Felsenfeld, D. Davis et A. Rich décrivaient une structure à 3 brins (Felsenfeld & Rich 1957). La formation d'une paire non canonique, entre la 1-méthylthymine et la 9-méthyladénine pouvant expliquer la formation de ce triplexe, fut démontrée par diffraction des rayons X en 1959 par Karst Hoogsteen (Figure 1C) (Hoogsteen 1959; Hoogsteen 1963). Les appariements dits « Hoogsteen » seront ensuite attribués, par analogie, aux liaisons s'effectuant sur le côté non « Watson-Crick » de la nucléobase (en noir, Figure 1C).

La structure en double hélice de l'ADN est caractérisée par la présence le long du squelette sucre-phosphate de deux sillons : le grand sillon qui est large et peu profond et le petit sillon qui est étroit et profond. L'environnement chimique d'une séquence d'ADN donnée dans chacun des sillons est différent et forme la base de reconnaissance pour les ligands de l'ADN. Ainsi, la double hélice d'ADN peut être reconnue par des oligonucléotides grâce à des appariements de type Hoogsteen ou par des intercalants de l'ADN tels que le bromure d'éthidium, grâce à l'empilement des noyaux aromatiques sur les paires de bases, ou encore par des molécules capables d'interagir avec l'ADN au sein des petit et grand sillons, comme le Hoechst 33258 et le DAPI (4',6-diamidino-2-phénylindole) qui sont des marqueurs de l'ADN souvent utilisés dans les tests cellulaires pour visualiser les noyaux des cellules.

A.2 Différents degrés de compaction de l'ADN dans la cellule

La chromatine correspond à l'association entre l'ADN et des petites protéines de quelques kDa appelées histones. Cette organisation permet la compaction d'un polymère d'ADN d'environ 2 mètres, dans un espace nucléaire dont les dimensions ne dépassent pas quelques micromètres dans le cas de l'homme. La chromatine est très organisée dans le noyau. Cette organisation résulte de différents niveaux hiérarchiques de compaction qui sont présentés dans la **Figure 2**. Le premier niveau correspond à l'enroulement de l'ADN autour des histones formant le nucléofilament (ou « fibre de 11 nm » de diamètre) constitué de nucléosomes. Les nucléosomes interagissent entre eux pour former une fibre de 30 nm (de

diamètre) qui s'organise en domaines de l'ordre de la mégabase (Mb). Enfin, le dernier niveau de compaction est atteint avec le chromosome mitotique.



Figure 2 : Les différents niveaux de compaction de la chromatine

L'ADN s'enroule autour d'un octamère d'histones (représenté schématiquement par le disque rose avec les queues N-terminales des histones) pour former le nucléosome. La chaine de nucléosomes constitue le premier niveau de compaction de la chromatine et est appelée « fibre de 11 nm » ou « collier de perles ». Le niveau supérieur est appelé « fibre de 30 nm ». La condensation maximale est atteinte avec le chromosome mitotique (image modifiée de Genetika®).

A.2.1 Le nucléosome et les histones

L'unité de base de la chromatine est le nucléosome qui est une structure d'un diamètre de 10 nm et d'une hauteur de 5,5 nm. Chez l'homme, il est constitué par l'enroulement de 147 paires de bases (pb) d'ADN (longueur variable selon les espèces) autour d'un octamère d'histones H3, H4, H2A et H2B (Olins & Olins 1974; Olins & Olins 2003). Chaque octamère est composé de deux dimères H3-H4 agencés en tétramère stable au centre du nucléosome encadré par un dimère H2A-H2B de chaque côté ; la structure du nucléosome a pu être caractérisée de manière très précise par cristallographie aux rayons X (Luger et al. 1997). Riches en résidus basiques, les histones sont chargées positivement ce qui permet des interactions de type électrostatique avec l'ADN. La présence d'une queue N-terminale (N-term) non-structurée d'une trentaine d'acides aminés qui émerge du nucléosome (Figure 2) permet aux histones de subir différents types de modifications post-traductionnelles (méthylation, acétylation, phosphorylation, ubiquitination, glycosylation, sumoylation...) sur des résidus ciblés (Arginine R, Lysine K, Sérine S...) (Campos & Reinberg 2009). Ces modifications permettent de réguler les interactions protéines/ADN, d'influer sur la

compaction de la fibre chromatinienne ou de moduler le recrutement de facteurs protéiques au sein de la chromatine. Les conséquences fonctionnelles de ces modifications dépendent du type de groupement chimique apposé, du type de résidu ciblé et de sa position le long de la queue N-terminale. Par exemple, la tri-méthylation de l'histone H3 sur la lysine 9, appelée H₃K₉me₃ est une marque associée à un état chromatinien compact et considéré comme transcriptionnellement silencieux. Les classes d'histones H2A, H2B et H3 comprennent plusieurs sous-types appelés « variants d'histones » et retrouvés dans des contextes chromatiniens particuliers. À titre d'exemple, le variant d'histone CenpA remplace l'histone H3 au niveau des centromères (les détails sur le centromère sont donnés au § Chapitre I - B.1.1) des chromosomes et joue un rôle essentiel dans l'assemblage du kinétochore qui est un complexe protéique permettant l'assemblage du fuseau mitotique aux chromosomes pendant la mitose.

A.2.2 Les niveaux de compaction supérieurs

L'enchaînement des nucléosomes via les interactions entre histones constitue la fibre de 11 nm (de diamètre) qui s'enroule ensuite sur elle-même pour former la fibre de 30 nm (de diamètre). Cependant, la structure tridimensionnelle précise et le mécanisme exact de formation de cette fibre sont encore débattus (Fussner et al. 2012). La compaction de la chromatine atteint un niveau maximum, en mitose, avec la formation des chromosomes métaphasiques (Figure 2). La découverte des chromosomes par Flemming remonte à plus d'un siècle quand il observa, juste avant la division cellulaire, la formation de corps marqués par un colorant basique de l'ADN (Flemming 1965). Ces structures très condensées sont caractérisées par l'organisation de la fibre de 30 nm en boucles sur un squelette protéique. Les différents niveaux de compaction permettent la définition de domaines chromatiniens en interphase : l'euchromatine et l'hétérochromatine, que l'on peut distinguer dans le noyau après l'incorporation d'intercalants de l'ADN ou en microscopie électronique (Bierne et al. 2009). L'euchromatine constitue une forme relâchée et contient des régions riches en gènes ; elle est active transcriptionnellement. En revanche, l'hétérochromatine est une structure beaucoup plus compacte, pauvre en gènes mais riche en régions répétées. On distingue deux types de régions hétérochromatiniennes. L'hétérochromatine facultative est une zone condensée, inactive en transcription mais qui peut potentiellement se décondenser et devenir transcriptionnellement active dans différents contextes; alors que l'hétérochromatine constitutive, retrouvée notamment au niveau des télomères et des centromères des

chromosomes, intervient dans le maintien de l'intégrité du génome et reste condensée tout au long de la vie de la cellule.

L'hétérochromatine constitutive, qui ne change pas « d'état », est composée principalement de séquences répétées en tandem. Au niveau des régions centromériques, ces séquences appelées aussi « séquences satellites » sont très abondantes et peuvent représenter jusqu'à 10% des génomes. Or, le caractère répété de ces séquences est à l'origine de difficultés techniques car elles se prêtent mal aux techniques classiques de biologie moléculaire. De plus, il est difficile de les positionner les unes par rapport aux autres et donc de les intégrer dans les séquences assemblées des génomes. Cependant, de nombreuses études ont permis de caractériser les séquences de différentes unités de répétitions chez un grand nombre d'organismes (Wu & Manuelidis 1980; Manuelidis 1981). Ainsi, l'identification de plusieurs familles de répétitions, qui se distinguent par leur longueur et leur composition, a été possible. Les séquences satellites diffèrent souvent entre les espèces, même si celles-ci sont relativement proches d'un point de vue phylogénétique, apparaissant ainsi comme une composante originale des génomes susceptible d'évoluer rapidement. La diversité de ce matériel génétique, dans une région associée à une fonction aussi conservée que celle du centromère, est une énigme posée par l'évolution et qu'il est d'autant plus intéressant de résoudre qu'on sait aujourd'hui que les séquences satellites sont transcrites (Rudert et al. 1995; Eymery et al. 2009).

Dans ce manuscrit, nous nous limiterons à la description des séquences répétées présentes au niveau des régions centromériques (centromères et régions péricentromériques).

B - Les régions centromériques et les séquences répétées

B.1 Les régions centromériques

B.1.1 Le centromère et les régions péricentromériques : définitions et généralités

C'est en 1936 que Darlington utilise le terme centromère pour définir la constriction primaire des chromosomes mitotiques (Gonçalves dos Santos Silva et al. 2008). Le terme centromère vient du grec « kentron » qui signifie central et « meros » pour partie ; en référence à la position généralement centrale du centromère sur les chromosomes métaphasiques, bien que chez la souris, qui possède des chromosomes acrocentriques, ce ne soit pas le cas. Le centromère est un domaine chromosomique spécialisé sur lequel s'assemble le kinétochore et joue donc un rôle crucial dans la ségrégation des chromatides sœurs (Cheeseman & Desai 2008). Les régions flanquantes du centromère, appelées régions péricentromériques, sont impliquées dans la cohésion des chromatides sœurs. Pour simplifier, dans ce manuscrit, le terme de régions centromériques désignera le centromère et ses régions associées (péricentromériques).

Au niveau des régions centromériques, les séquences d'ADN sont très divergentes entre les espèces (**Figure 3**). Le centromère le plus simple, suffisant pour faciliter l'assemblage du kinétochore, est trouvé chez la levure bourgeonnante *Saccharomyces cerevisiae*. Il est constitué de trois séquences appelées CDEI, CDEII et CDEIII (Centromeric DNA Element), riches en A/T, qui s'étendent sur une longueur de 125 pb (**Figure 3A**). Contrairement à ce centromère ponctuel, la plupart des eucaryotes (y compris la levure à fission *Schizosaccharomyces pombe*) possèdent des centromères régionaux (Pluta et al. 1995). Ces régions sont composées de séquences d'ADN répétées en tandem ; la séquence et la taille de l'unité de répétition (35-110 Kb chez *Schizosaccharomyces pombe* (**Figures 3B**) contre 0,3-5 Mb chez l'homme (**Figure 3E**)) ainsi que le nombre de répétitions sont variables entre les espèces. Parfois, plusieurs familles de séquences sont trouvées sur les différents chromosomes voire sur un même chromosome. À titre d'exemple, chez la souris, deux familles principales de répétitions ont été décrites : les satellites mineurs de 120 pb au niveau des centromères et les satellites majeurs de 234 pb au niveau des péricentromères (**Figure 3D**). Ces séquences sont présentes sur tous les chromosomes à l'exception du chromosome Y (Pardue & Gall 1970). Chez l'homme, les répétitions α-satellites de 171 pb sont les plus abondantes. Les autres familles de séquences peuvent différer entre les chromosomes (Figure 3E) (les détails sur les modèles murins et humains seront présentés au § Chapitre I - B.2).





Les régions centromériques et péricentromériques de différents organismes présentent des différences de taille et de séquence d'ADN. A/ Les régions centromériques de S.cerevisiae sont composées de trois domaines CDEI, CDEII et CDEIII. B/ Chez S.pombe, un noyau central unique (Cnt) est flanqué de répétitions (ImrL, ImrR, dg, dh). C/ Les régions centromériques de la drosophile sont formées par deux blocs de répétitions AATAT et AAGAG entrecoupés par des éléments transposables. D/ Chez la souris, les satellites mineurs composent le centromère et les satellites majeurs - les régions péricentromériques. E/ Les régions centromériques humaines sont principalement constituées de répétitions α -satellites (adapté de (Rudd 2005)).

À l'heure actuelle, les connaissances sur les séquences des régions centromériques, chez des espèces où elles sont très étendues, sont limitées du fait de leur caractère répété avec un degré d'homogénéité élevé. En effet, l'assemblage de séquences répétées de l'ordre d'une centaine de milliers de copies reste très compliqué (Rudd & Willard 2004). Les cartes des génomes présentent donc souvent un intervalle « centromeric gap » non assemblé qui correspond à la région centromérique sur chaque chromosome.

B.1.2 Le centromère et les régions péricentromériques : structure et fonction

Au cours de l'évolution, la conservation d'une fonction d'un locus est souvent corrélée avec la conservation de la structure. En revanche, une énigme persiste en ce qui concerne le centromère. En effet, le rôle joué par le centromère dans l'assemblage du kinétochore et la division cellulaire est très conservé au cours de l'évolution, alors que les séquences d'ADN sont très divergentes entre les espèces. Cette singularité est connue sous le nom de « paradoxe du centromère ». Cependant, deux caractéristiques sont conservées au cours de l'évolution : la présence du variant spécifique de l'histone H3 (CenpA, parfois nommé CenH3) au niveau de tous les centromères (Sullivan et al. 1994) et le caractère répété et riche en A/T des séquences. CenpA est la seule protéine centromérique totalement conservée au cours de l'évolution. Des expériences de délétions ont montré que la déplétion de la protéine CenpA provoquait des problèmes pour l'assemblage du kinétochore et la bonne ségrégation des chromatides sœurs au cours de la mitose (Howman et al. 2000).

Au niveau des péricentromères, l'absence de la protéine CenpA et l'enrichissement de la marque répressive H₃K₉me₃ (Rice et al. 2003) sont caractéristiques. Une bonne cohésion des chromatides sœurs est permise grâce à l'interaction de la marque H₃K₉me₃ avec des complexes de cohésines et la protéine HP1 (Sullivan & Karpen 2004; Guenatri et al. 2004). Cette cohésion est essentielle à l'attachement correct des chromosomes aux microtubules du fuseau mitotique pendant la division cellulaire (Amor et al. 2004). L'importance du rôle de la marque H₃K₉me₃ a été décrite chez l'homme par une étude montrant que des cellules tumorales, présentant une forte instabilité chromosomique, avaient souvent une diminution de la marque H₃K₉me₃. De plus, des cellules dans lesquelles la quantité de cette marque était diminuée artificiellement, par la surexpression d'une histone déméthylase (JMJD2B), présentaient une forte instabilité chromosomique (Slee et al. 2012). Ces observations suggèrent que l'hétérochromatine péricentromérique joue un rôle dans la fonction du centromère.

Il est possible que des transcrits aident à la fonction des séquences répétées. En effet, des transcrits issus des régions centromériques ont été observés dans de nombreux organismes comme la levure *S.pombe*, les plantes, la souris et l'homme. Ces transcrits s'accumulent dans différents contextes physiologiques, notamment au cours du cycle cellulaire, de la

différentiation, du développement, en réponse à des stress et dans des pathologies telles que les cancers (Rudert et al. 1995; Eymery et al. 2009). Leur accumulation forcée dans des cellules de souris mène à des problèmes de ségrégation des chromatides sœurs ; ce qui suggère un rôle potentiel de ces transcrits dans la fonction du centromère (Bouzinba-Segard et al. 2006). Ces transcrits semblent aussi être un composant essentiel de la structure des régions centromériques et sont impliqués dans l'association de la protéine HP1 à l'hétérochromatine péricentromérique (Maison et al. 2002). Cependant, les mécanismes sous-jacents sont encore peu compris.

B.2 Les deux modèles de séquences répétées étudiés dans ce travail

Seulement 2% des génomes humain et murin code pour des protéines. Le reste (98%) est composé principalement d'introns, de promoteurs et de séquences répétées. Environ la moitié de leurs génomes est composée de séquences répétées d'ADN. Parmi ces séquences, on trouve les répétitions dispersées et les répétitions en tandem. Les séquences répétées en tandem ou « séquences satellites » occupent environ 10% des génomes, humain et murin, et peuvent atteindre plusieurs millions de bases. Elles sont caractérisées par la présence de multiples copies d'une unité de répétition. Le nom « ADN satellite » vient du fait que ces séquences d'ADN, en raison de leur nature répétée, ont une composition en nucléotides fortement biaisée par rapport au reste du génome. Par conséquent, elles sédimentent différemment après centrifugation dans un gradient de césium et forment une bande secondaire dite « satellite » (Corneo et al. 1967).

B.2.1 Chez la souris :

B.2.1.1 Les séquences satellites mineurs et majeurs

À l'exception du chromosome Y, tous les centromères des chromosomes murins portent les mêmes séquences satellites. Deux types de séquences d'ADN répétées bien définies sont présents au niveau des régions centromériques : les satellites mineurs et majeurs (**Figure 4**). Le centromère murin est caractérisé par la présence des séquences satellites mineurs d'une longueur de 120 pb. Cette unité est répétée plus de 2500 fois sur chaque chromosome. Les régions péricentromériques sont composées par les séquences satellites majeurs « MajSat », qui sont les plus abondantes. Leur motif de 234 pb est répété sur une
longueur de plus de 6 Mb sur les chromosomes murins (Manuelidis 1981; Wong & Rattner 1988; Guenatri et al. 2004). Une distribution inégale des bases adénines et thymines, dans la séquence, est à l'origine d'une distinction entre les deux brins de l'ADN : on parle d'un brin riche en purines (ou riche en A) et l'autre, riche en pyrimidines (ou riche en T).



Figure 4 : Organisation des régions centromériques murines: le centromère et les régions péricentromériques

Les séquences satellites mineurs (vert) sont localisées au niveau du centromère alors que les séquences satellites majeurs (rose) entourent le centromère pour former le péricentromère.

Au niveau des régions centromériques, des modifications épigénétiques différentes permettent aussi une distinction des satellites mineurs et majeurs. Les satellites mineurs composent la plateforme de l'assemblage du kinétochore avec la présence du variant d'histone CenpA, alors que les satellites majeurs sont associés aux protéines de l'hétérochromatine (comme par exemple la protéine HP1) et jouent un rôle important dans la cohésion des chromatides. En effet, une perte de l'histone méthyltransférase Suv39h, qui conduit à une perte de la marque H₃K₉me₃ au niveau des séquences MajSat, empêche le recrutement de la protéine HP1; ce qui engendre un défaut de cohésion des chromatides sœurs et donc des problèmes de ségrégation des chromatides au cours de la division cellulaire (Eymery et al. 2009). Un autre élément clé des différences entre les deux types de séquences est leur réplication asynchrone pendant la phase S qui est également observée chez la levure *S.pombe* et chez la mouche *D.melanogaster*. Alors que les satellites mineurs se répliquent tardivement en phase S, les satellites majeurs se répliquent en milieu de phase S. Cette régulation temporelle de la réplication au niveau de ces séquences semble être un facteur

supplémentaire contribuant à l'établissement d'un profil, d'organisation et de modifications, différent pour ces deux types de séquences ; elles jouent donc des rôles essentiels mais différents dans l'organisation et la fonction du centromère (Guenatri et al. 2004).

Les répétitions MajSat sont très abondantes dans les cellules de souris, portent des marques épigénétiques spécifiques et s'associent, en interphase, pour former des « chromocentres » (voir § Chapitre I - B.2.1.2). De plus, un avantage supplémentaire de travailler avec ce type de séquences est qu'il n'y a pas de variabilité interchromosomique puisque tous les chromosomes portent cette même unité de répétition.

B.2.1.2 L'association des régions péricentromériques murines : les chromocentres

En 1908, le botaniste italien Pasquale Baccarini observait des structures denses dans le noyau en interphase de plantes, qu'il nomma « chromocentri » (chromocentres). Chez de nombreuses espèces telles que la drosophile, la souris ainsi que de nombreux mammifères, ces structures ont été observées dans le noyau interphasique par un simple marquage de l'ADN (Jost et al. 2012).

En interphase, le marquage de l'ADN par des colorants, comme le DAPI ou le Hoechst, permet de visualiser les régions centromériques murines dans le noyau. Elles correspondent à des foyers, aux contours nets, marqués plus intensément que le reste du noyau et qui sont appelés chromocentres. Ce contraste peut être expliqué par la grande affinité de ces colorants de l'ADN pour les paires de bases A/T et, comme les séquences satellites murines sont plus riches en A/T que le reste du génome, alors elles seront plus marquées. De plus, ces régions sont plus condensées que le reste du génome (Fussner et al. 2012) ; ce qui augmente la concentration locale de colorants. Des expériences d'Hybridation de Fluorescence In Situ (FISH) utilisant des sondes dirigées contre les satellites mineurs et majeurs ont montré que ces foyers correspondent à l'association des péricentromères murins portés par plusieurs chromosomes et que les centromères se localisent en périphérie de ces foyers pendant l'interphase (**Figure 5**). D'autre part, des marques épigénétiques caractéristiques de l'hétérochromatine sont enrichies au niveau de ces foyers comme H₃K₉me₃ ou encore la protéine HP1 (Guenatri et al. 2004).



Figure 5 : Association des régions péricentromériques chez la souris

Organisation spatiale des péricentromères (satellites majeurs, dégradé de rouges) et centromères (satellites mineurs, dégradé de verts) murins (inspiré de (Guenatri et al. 2004)). En interphase, les péricentromères portés par plusieurs chromosomes s'associent pour former des foyers appelés chromocentres. Les centromères se localisent en périphérie de ces foyers sous forme de points.

L'association des régions centromériques est un phénomène dynamique et variable puisque la taille et le nombre des chromocentres varient entre les types cellulaires et selon l'état de différenciation des cellules (Almouzni & Probst 2011). Les différences les plus marquées sont observées entre les cellules sphériques (lymphocytes, macrophages, myotubes) qui présentent peu de chromocentres mais de grande taille et les cellules aplaties (fibroblastes, myoblastes) dont les chromocentres sont plus nombreux et plus petits (Mayer et al. 2005). La signification fonctionnelle de cette association est encore peu claire, mais il est suggéré qu'elle peut servir à compartimenter l'hétérochromatine, ce qui permettrait une accumulation de facteurs épigénétiques et une maintenance de l'état hétérochromatinien (Almouzni & Probst 2011). Il a aussi été observé que l'association de loci d'euchromatine avec les chromocentres corrélaient avec leur répression; ce qui suggère que les chromocentres pourraient participer aux mécanismes de régulation de l'expression des gènes (Brown et al. 1997; Roldán et al. 2005). Ainsi, les chromocentres sont des foyers d'hétérochromatine dont le nombre et la position sont variables. Leur état d'hétérochromatine suggère qu'ils peuvent avoir une fonction de répresseur transcriptionnelle ; par conséquent, un gène situé à proximité des chromocentres peut voir son expression inhibée.

B.2.2 Chez l'homme

B.2.2.1 Les séquences satellites

Les régions centromériques humaines sont composées de plusieurs familles de séquences satellites. Les séquences α -satellites, qui sont répétées sur des longueurs pouvant atteindre plusieurs Mb, constituent la famille retrouvée majoritairement (Manuelidis 1978;

Wu & Manuelidis 1980) ; à celles-ci s'ajoutent les séquences β -/ γ -satellites et les satellites 1, 2 et 3 (**Figure 6** ; (Lee et al. 1997)). Au sein de chaque région centromérique, les unités répétées d' α -satellites peuvent être distinguées en deux types : les monomériques qui sont divergentes et qui présentent peu d'homologie locale, et les HORs (Higher-Order Repeats) qui sont des unités de répétitions multi-monomériques. Les HORs sont répétées en tandem et présentent une très forte homologie au sein d'un même ensemble de répétitions (97-100%) (Alexandrov et al. 1993) ; alors qu'il n'y a environ que 60-80% d'homologie en moyenne entre deux monomères. Une ou plusieurs HORs, répétées sur des longueurs de l'ordre de la Mb, se trouvent au niveau de chaque région centromérique humaine et le variant d'histone CenpA est assemblé sur l'une d'entre elles pour former le centromère (Alexandrov et al. 2001).



Figure 6 : Organisation des régions centromériques humaines : les séquences satellites

Chaque petite flèche verte représente un monomère d' α -satellite simple. Au niveau des centromères, les α -satellites sont organisés en HORs représentées par les flèches blanches. Ces HORs sont des unités de répétition composées de plusieurs monomères. Au niveau des régions péricentromériques, on trouve des monomères d' α -satellites ainsi que d'autres types de séquences satellites.

Chez l'homme, il existe des régions centromériques à forte similarité (Alcobia et al. 2000; Hayden 2012) et des régions centromériques présentant des variations de séquences entre les différents chromosomes. En effet, les HORs peuvent se distinguer par la composition des monomères, le nombre de répétitions et les autres familles de séquences satellites qui ne sont pas réparties de façon homogène sur tous les chromosomes (Lee et al. 1997).

B.2.2.2 L'association des régions centromériques humaines

Chez l'homme, le marquage de l'ADN par le DAPI ne permet pas d'observer le même type de foyers de chromatine péricentromérique que chez la souris. Cette différence peut s'expliquer par le fait que les séquences satellites humaines sont moins riches en A/T que les séquences murines. Des approches faisant intervenir un marquage indirect des régions centromériques, par immunocytochimie en utilisant un sérum ciblant plusieurs protéines centromériques ou un anticorps dirigé contre CenpA, ont permis de montrer que les régions centromériques portées par différents chromosomes humains s'associent dans le noyau en interphase (Gonçalves dos Santos Silva et al. 2008). La détection d'un nombre de foyers centromériques plus faible que le nombre de chromosomes de la cellule est généralement interprétée comme l'association de plusieurs centromères. Plusieurs études ont montré que les associations observées varient au cours du cycle cellulaire, entre les types cellulaires et au cours de la différenciation (Alcobia et al. 2000; Gonçalves dos Santos Silva et al. 2008). Au sein de notre laboratoire, des études sur des cellules humaines fixées ont permis de décrire deux aspects de l'organisation spatiale des régions centromériques humaines : d'une part, elles se positionnent préférentiellement soit en périphérie du noyau soit au bord des nucléoles et d'autre part, les régions péricentromériques portées par plusieurs chromosomes s'associent pour former des foyers d'hétérochromatine comme chez la souris et de nombreux autres organismes. Chez l'homme, l'observation de structures apparentées aux chromocentres murins a été effectuée uniquement sur des cellules fixées (Ollion 2014).

B.3 Pourquoi et comment étudier la structure et la dynamique de la chromatine au niveau des séquences d'ADN répétées ?

La compréhension du contrôle des fonctions nucléaires par l'organisation dynamique de la chromatine est devenue le centre d'intérêt de recherches intensives. Le remodelage tridimensionnel de la chromatine dans le noyau est reconnu comme étant déterminant pour la régulation du génome incluant la régulation de la transcription des gènes, le développement ou encore la différenciation cellulaire. De plus, des anomalies de l'organisation nucléaire mènent à plusieurs maladies (par exemple la mutation de la cohésine qui est une protéine jouant un rôle dans la cohésion des chromatides sœurs peut être à l'origine de problèmes de croissance pré- et post-natales) (Misteli 2010). Deux aspects sont à retenir pour étudier l'organisation dynamique de la chromatine : l'aspect régulation transcriptionnelle des gènes et

Chapitre I – Introduction Générale Les régions centromériques et les séquences répétées

l'aspect visualisation de l'ADN db dans des cellules vivantes. L'observation de cellules vivantes par microscopie de fluorescence a permis des avancées majeures dans ce domaine. Cependant, les recherches restent limitées par le manque d'outils disponibles pour la visualisation spécifique de loci génomiques dans les cellules vivantes, malgré le développement de technologies récentes qui ont permis de commencer à explorer les rôles biologiques du mouvement de la chromatine dans le noyau. L'organisation spatio-temporelle des génomes dans le noyau joue un rôle clé dans les fonctions du génome, mais celui-ci est encore mal compris. Utilisées pour permettre une meilleure compréhension du rôle fonctionnel et des mécanismes régulant l'architecture du génome, les études sur cellules fixées ont leurs limites car la fixation des cellules peut fortement altérer l'architecture 3D du noyau, la morphologie du noyau ou de la chromatine et la localisation des biomolécules ; d'où l'utilité de développer des sondes fluorescentes qui permettraient d'accéder à la dynamique de l'organisation des séquences d'ADN dans des cellules vivantes pour aborder la question de la relation entre structure / architecture nucléaire et fonction du génome.

« L'état » d'hétérochromatine joue un rôle critique pour la fonction du centromère. Cependant la contribution respective des différentes séquences répétées trouvées dans ces régions reste encore méconnue. Les séquences d'ADN répétées, localisées au niveau des régions centromériques, sont connues pour avoir une organisation nucléaire dynamique et spécifique, mais l'importance de cette organisation n'est pas complètement comprise. Les séquences d'ADN satellites, qui représentent le composant principal de la chromatine centromérique, sont parmi les premières régions pour lesquelles une organisation nucléaire non aléatoire a été observée. Les régions centromériques sont des structures mobiles qui peuvent s'associer à la lamina ou au nucléole dans les cellules humaines et de souris, et leur distribution dans le noyau interphasique est modifiée pendant le cycle et la différentiation cellulaire (Pluta et al. 1995; Wiblin et al. 2005). De nombreuses études, chez la souris, ont montré que ces régions péricentromériques sont soumises à des réorganisations majeures pendant la différentiation cellulaire (Terranova et al. 2005; Almouzni & Probst 2011; Aguirre-lavin et al. 2012) et qu'elles peuvent former des foyers d'hétérochromatine capables d'induire une répression transcriptionnelle des gènes situés à proximité (Brown et al. 1997).

Une séquence d'ADN répétée particulière est la séquence murine MajSat de 234 pb qui est présente en milliers de copies dans les régions péricentromériques de tous les chromosomes. Dans les noyaux en interphase, ces régions s'agrègent pour former des foyers, appelés chromocentres, visibles après le marquage de l'ADN par un colorant tel que le DAPI ou par la technique de FISH (Guenatri et al. 2004). C'est pourquoi la séquence d'ADN MajSat est un modèle idéal afin de tester l'applicabilité de sondes fluorescentes pour l'imagerie des séquences d'ADN en visualisant les chromocentres directement dans des cellules vivantes.

Pour le modèle humain, la séquence satellite la plus abondante est la séquence α satellite de 171 pb qui est présente en millions de copies sur toutes les régions centromériques des chromosomes humains. Elle s'organise en structures d'ordre supérieur, les HORs. Il est connu que les régions centromériques de différents chromosomes peuvent s'associer au bord nucléaire et aux nucléoles (Haaf & Michael 1991). La séquence α -satellite peut être utilisée comme modèle, afin de voir si les outils fluorescents conçus se comportent comme ceux utilisés chez la souris, en ce qui concerne la pénétration cellulaire et l'accessibilité à l'ADN. Ainsi, il serait possible de vérifier si dans les cellules humaines vivantes, d'éventuelles structures s'apparentant aux chromocentres murins sont observables.

<u>C - Les outils pour visualiser des séquences d'ADN double</u> brin

Les différents mécanismes de reconnaissance et de liaison à des séquences d'ADN db spécifiques ont été exploités pour la conception de nouveaux outils. Le premier mécanisme étant la formation d'une triple hélice en utilisant un oligonucléotide simple brin qui a la capacité de se lier dans le grand sillon de l'ADN de façon séquence-spécifique (Sun & Hélène 1993). Ensuite, il y a différents mécanismes permettant aux protéines naturelles d'interagir et de se fixer spécifiquement à des séquences d'ADN. Plusieurs tentatives ont été décrites pour utiliser artificiellement ces mécanismes, comme par exemples l'utilisation de motifs à doigt de zinc ou encore le recours à des systèmes plus récents : TALE (Transcription Activating Like Effector) et CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats). Pour finir, un autre type de molécules, inspiré de la structure de molécules naturelles et qui ont la capacité d'interagir avec l'ADN au sein du petit sillon, s'appellent les polyamides N-méthylpyrrole et N-méthylimidazole (Dervan & Edelson 2003; Boutorine et al. 2013). C'est sur ces dernières molécules que notre choix s'est porté pour la conception de nouveaux outils afin de pouvoir visualiser des séquences d'ADN répétées dans les cellules vivantes (les détails sur les polyamides sont donnés au § Chapitre I - D).

La conception de sondes fluorescentes pour visualiser des séquences d'ADN dans des cellules vivantes doit satisfaire à plusieurs exigences comme :

- ✓ une haute spécificité et affinité pour l'ADN cible
- une pénétration facile dans les cellules vivantes et une capacité à trouver la séquence cible dans les cellules
- ✓ une perturbation minimale du fonctionnement cellulaire
- ✓ une cytotoxicité limitée
- ✓ une stabilité dans les conditions physiologiques (résistance aux nucléases par exemple)
- ✓ une détection facile
- ✓ une modulation du spectre de fluorescence de la sonde lors de son interaction avec l'ADN cible
- ✓ un rapport « signal/bruit de fond » exploitable

C.1 Les oligonucléotides modifiés

La découverte des propriétés structurales de l'ADN, notamment la formation de paires de bases complémentaires, a rapidement amené au développement de nouvelles techniques expérimentales pour l'étude des acides nucléiques (Southern 1975). À titre d'exemples, le Northern blot permet la détection des ARNs, le Southern blot celle des ADN ou encore la PCR (Polymerase Chain Reaction) qui permet la détection et l'amplification d'une séquence d'ADN donnée. Ces techniques sont principalement basées sur le processus d'hybridation entre un acide nucléique et un oligonucléotide de séquence complémentaire. Un oligonucléotide est une petite séquence d'acide nucléique simple brin synthétisé chimiquement et qui peut s'hybrider spécifiquement à une séquence cible. Les oligonucléotides composés uniquement par des nucléotides ou des ribonucléotides naturels sont peu stables dans l'environnement cellulaire où ils sont rapidement reconnus et dégradés par les nucléases. De plus, ils pénètrent mal dans les cellules. C'est pourquoi des analogues chimiques des oligonucléotides ont été développés.

C.1.1 Les TFOs (Triplex Forming Oligonucleotides)

La formation d'une triple hélice est possible lorsqu'un oligonucléotide se lie à un duplexe d'acide nucléique. Il a été montré que le 3^{ème} brin nucléotidique se lie dans le grand sillon du duplexe en s'hybridant avec l'hélice Watson-Crick polypurique / polypyrimidique de façon séquence-spécifique (Felsenfeld 1957; Sun & Hélène 1993; Faria & Giovannangeli 2001). Les TFOs (Triplex Forming Oligonucleotides) peuvent reconnaitre des séquences de plus de 10 - 12 pb via des liaisons hydrogène de type Hoogsteen (Figure 7) ou Hoogsteen inverse. La capacité à former une liaison hydrogène supplémentaire par appariement de type Hoogsteen, requiert un pH acide afin de protoner une cytosine. La synthèse des oligonucléotides est simple et ils sont disponibles commercialement. Cependant, dans des conditions physiologiques, les triple hélices sont souvent instables (accessibilité limitée, dégradation par les nucléases, pH neutre, mésappariements) (Sun & Hélène 1993; Fox 2000; Casey & Glazer 2001; Knauert & Glazer 2001; Duca et al. 2008). De plus, la pénétration intracellulaire des acides nucléiques nécessite le recours à des agents de transfection ou des méthodes de perméabilisation cellulaire. Malgré ces contraintes techniques et en utilisant au mieux leurs atouts, leur marquage par des fluorophores a été développé (Mergny et al. 1994; Grimm et al. 2002; Daele et al. 2008; Renard et al. 2008). Les TFOs fluorescents ont été

Chapitre I – Introduction Générale Les outils pour visualiser des séquences d'ADN double brin

utilisés pour détecter des séquences d'ADN dans des cellules par des techniques de microscopie de fluorescence comme le TISH (Triplexe In Situ Hybridization) (Johnson III & Fresco 1999) et le COMBO-FISH (Schwarz-Finsterle et al. 2007) qui sont basées sur la technique de FISH mais, qui ne nécessitent pas d'étape de dénaturation thermique de l'ADN. Cependant, ces travaux ont été réalisés sur des cellules fixées afin de pouvoir éliminer l'excès d'oligonucléotides fluorescents non liés par des lavages et ainsi améliorer le signal du bruit de fond avant l'acquisition d'images.



Figure 7 : La triple hélice

(A) La triple hélice d'ADN : le brin polypurique en rouge, le brin pyrimidique en cyan et le 3^{eme} brin polypyrimidique en jaune, qui est dans le grand sillon du duplexe d'ADN. (B) Exemples de triplets de type Hoogsteen entre T/A×T, C/G×C⁺. WC : appariements de type Watson-Crick et H : appariements de type Hoogsteen (adapté de (Boutorine et al. 2013)).

C.1.2 Les modifications chimiques des oligonucléotides

Pour améliorer l'affinité, la spécificité, la stabilité, la résistance aux nucléases et aux conditions physiologiques et la pénétration cellulaire des TFOs, diverses stratégies ont été utilisées. Elles reposent sur des modifications chimiques des bases, du « squelette phosphate » ou encore des sucres (**Figure 8**).



Figure 8: Les différentes modifications chimiques des oligonucléotides

A/ Exemple de modification de la base avec la 5-méthylcytosine (en rouge, le groupement méthyle porté par le carbone 5 de la cytosine). B/ Exemple de modification du « squelette phosphate » avec le PNA (entre les crochets bleus, le lien peptidique du PNA qui remplace le lien phosphodiester des acides nucléiques). C/ Exemple de modification du sucre avec le LNA (en orange, le pont méthylène entre l'hydroxyle en position 2' et le carbone en C4').

L'idée du recours à des modifications portant sur les bases est de stabiliser la structure du triplexe. Par exemple, la contrainte relative au pH, afin d'obtenir une cytosine protonée dans des conditions physiologiques, a pu être améliorée par l'utilisation de la 5-méthylcytosine (**Figure 8A**) (Uil et al. 2003; Duca et al. 2008).

En ce qui concerne les modifications du « squelette phosphate », elles ont notamment pour objectifs de changer les propriétés électrostatiques du squelette phosphodiester chargé négativement, d'améliorer la résistance aux nucléases et la pénétration cellulaire. Ainsi, le développement des PNAs (Peptide Nucleic Acids), qui ont un squelette peptique non chargé, a permis d'améliorer les problèmes de répulsions électrostatiques entre les groupements phosphate chargés négativement (**Figure 8B**) (Uil et al. 2003; Duca et al. 2008). Marqués par un fluorophore, ces PNAs ont pu être utilisés pour visualiser les télomères dans des cellules humaines et de souris (Molenaar et al. 2003; Boutorine et al. 2013). Pour finir, les modifications portant sur les sucres sont basées principalement sur des substitutions sur le carbone en position 2'. Un des exemples le plus remarquable est celui des LNAs (Locked Nucleic Acids) qui sont des analogues d'acides nucléiques portant un pont méthylène entre l'hydroxyle en position 2' et le carbone en C4', ce lien covalent réduit la flexibilité du ribose en le bloquant en conformation C3'-endo (**Figure 8C**). La structure des LNAs rend ces oligonucléotides modifiés très affins pour leur brin complémentaire (**Prakash** 2011). Le squelette phosphate étant similaire à celui de l'ADN, l'incorporation d'un ou plusieurs monomères de LNA au sein d'un oligonucléotide modifié est possible. À titre d'exemple, des séquences mixtes ADN / LNA dans des oligonucléotides modifiés ont été utilisés, après leur couplage a des fluorophores, comme sondes fluorescentes pour visualiser des séquences d'ADN répétées dans des expériences de FISH (Silahtaroglu et al. 2004; Ollion 2014).

Les oligonucléotides naturels ou modifiés représentent une approche intéressante pour la construction de sondes fluorescentes spécifiques de séquences cibles d'ADN. Cependant, la meilleure reconnaissance de séquences riches en purines et les problèmes de pénétration cellulaire, rendent leur utilisation limitée aux études sur cellules fixées ou au recours à des systèmes de vectorisation complexes.

C.2 Les protéines modifiées

Un autre type d'approche pour visualiser l'ADN db dans des cellules est le recours à des constructions d'ingénierie génétique pour exprimer différents peptides ou protéines, capables d'interagir spécifiquement avec l'ADN, fusionnés à des protéines fluorescentes comme la GFP (Green Fluorescent Protein). Après les étapes de fusion de gènes dans des plasmides, de transfection et d'expression des gènes fusionnés, la visualisation de séquences d'ADN db par microscopie de fluorescence est rendue possible (**Figure 9**, adaptée de (Boutorine et al. 2013)).



Figure 9 : Les différentes étapes nécessaires à la visualisation de l'ADN par une protéine de fusion

La structure de protéines naturelles capables d'interagir avec l'ADN db de façon séquence-spécifique est une réelle source d'inspiration pour la recherche. En effet, les facteurs de transcription naturels ont une structure modulaire composée de deux domaines fonctionnels : le domaine effecteur et le domaine de liaison à l'ADN (DBD pour DNA Binding Domain). Ce dernier est à l'origine de diverses constructions d'ingénierie génétique afin de créer des protéines de fusion (sondes fluorescentes ou nucléases artificielles) capables d'interagir de façon séquence-spécifique avec l'ADN db (Garvie & Wolberger 2001; Tachikawa & Briggs 2006; Ghosh et al. 2006; Rodriguez-Marinez et al. 2014). Il y a quatre types de motifs structuraux majeurs des DBDs : hélice-tour-hélice, fermeture à leucine, hélice-boucle-hélice et doigt de zinc. Chez l'homme, le motif à doigt de zinc est le plus représenté (Figure 10). Le motif Cys₂His₂ (Cys-Cystéine et His-Histidine) est caractéristique des protéines à doigt de zinc et est composé de 30 acides aminés assemblés en repliement de type ββα stabilisé par un ion zinc. Chaque motif peut reconnaitre et interagir avec une séquence d'ADN de 3 pb mais plusieurs motifs peuvent être assemblés afin de reconnaitre une séquence plus longue (Uil et al. 2003; Klug 2010; Urnov et al. 2010; Rodriguez-Marinez et al. 2014). Par exemple, chez la plante Arabidopsis thaliana et la souris, des sondes fluorescentes ciblant des répétitions centromériques ont été construites en fusionnant la GFP à un doigt de zinc polydactyle qui contient 3 à 6 motifs pour des expériences d'imagerie cellulaire (Lindhout et al. 2007).



Figure 10 : Représentation schématique d'un motif à doigt de zinc Cys₂His₂

Exemple d'un motif à doigt de zinc Cys₂His₂ ayant une structure secondaire de type $\beta\beta\alpha$. La localisation des résidus conservés est représentée par des cercles avec un C (Cystéine) ou un H (Histidine) à l'intérieur. Ces résidus coordonnent l'ion Zn²⁺ qui est représenté par une sphère noire. Les carrés montrent les résidus impliqués dans la liaison spécifique à l'ADN et les nombres à l'intérieur des carrés symbolisent leur position par rapport au premier acide aminé de l'hélice α (Uil et al. 2003).

Récemment, le modèle des protéines à doigt de zinc pour construire des protéines de fusion capables d'interagir avec l'ADN a été remplacé par de nouvelles protéines spécifiques de séquences d'ADN appelées les TALEs (Transcription Activating Like Effectors). Les TALEs sont des protéines naturelles (facteurs de transcription) sécrétées par Xanthomonas spp qui est une bactérie pathogène des plantes. Elle les sécrète lors de l'infection de différentes espèces de plantes. Les TALEs peuvent se lier à des séquences d'ADN (promoteurs) de la plante hôte et activer l'expression de gènes chez celle-ci afin de faciliter l'infection par la bactérie. Elles reconnaissent les séquences d'ADN des plantes grâce à un domaine central de répétitions en tandem de 34 acides aminés (Figure 11A). Les deux acides aminés en position 12 et 13 de chaque répétition sont hypervariables et reconnaissent une paire de bases dans la séquence d'ADN cible. Ainsi, la nature des acides aminés en position 12 et 13 détermine le code de reconnaissance (Figure 11B) (Boch et al. 2009; Nga-Sze Mak et al. 2012). Cette propriété des TALEs a été utilisée pour la construction de nucléases artificielles (TALEN pour Transcription Activating Like Effector Nuclease) en fusionnant les TALEs avec le domaine catalytique de la nucléase Fok1. Les TALENs sont capables de modifier l'ADN génomique en induisant des coupures (mutagénèse dirigée) (Christian et al. 2010; Miller et al. 2011; Bedell et al. 2012). La structure cristallographique d'un complexe ADN-TALE a été élucidée (Nga-Sze Mak et al. 2012). Et très récemment, la construction de

sondes fluorescentes en fusionnant les TALEs à des protéines fluorescentes, a été réalisée pour cibler et visualiser des séquences répétées centromériques et télomériques chez l'homme et la souris par microscopie de fluorescence (Ma et al. 2013; Miyanari et al. 2013; Miyanari 2014; Thanisch et al. 2014).



Figure 11 : Représentation schématique d'une protéine TALE

A/ La TALE contient un domaine central de 34 acides aminés qui est répété en tandem, un signal d'adressage au noyau (NLS) et un domaine d'activation (AD). En rouge, les acides aminés hypervariables en position 12 et 13. B/ Code de reconnaissance des TALEs (adapté de (Boch et al. 2009)).

Encore plus récemment a été découvert chez *Streptococcus pyogenes*, le système CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) qui est un mécanisme de défense immunitaire adaptatif chez les bactéries et les archées utilisé pour la dégradation du matériel génétique qui leur est étranger. Ce système utilise une nucléase Cas9 (CRISPR-associated) qui, complexée à un ARN guide (ARNg) complémentaire d'une séquence cible, coupe l'ADN de façon spécifique en amont d'une séquence PAM (Protospacer Adaptor Motif) localisée sur l'ADN génomique (Mojica et al. 2005; Bolotin et al. 2005).

Ce nouvel outil est devenu un outil puissant pour l'ingénierie des génomes (Mali et al. 2013). Fusionné à la GFP et en inactivant la nucléase Cas9 (**Figure 12**), ce système a permis la construction de sondes fluorescentes pour des séquences centromériques et télomériques chez l'homme et la souris (Chen et al. 2013; Anton et al. 2014).



Figure 12 : Le système CRISPR/Cas9 pour la visualisation de séquences d'ADN

Deux plasmides, l'un contenant la séquence pour l'ARNg (ARN guide) et l'autre celle de la Cas9 inactive fusionnée à la GFP, sont co-transfectés dans les cellules puis l'ARNg guide la Cas9-GFP au niveau des séquences cibles d'ADN, ce qui permet l'accumulation de protéines fluorescentes au niveau des séquences d'intérêt (adapté de (Chen et al. 2013)).

Les protéines modifiées ont permis de développer des sondes fluorescentes spécifiques de séquences cibles d'ADN. Cependant, certaines contraintes à leur utilisation persistent comme la difficulté à construire des lignées cellulaires stables exprimant les protéines d'intérêt, l'étape de transfection qui peut être cytotoxique pour certains types cellulaires et les constructions génétiques complexes pour permettre l'expression des protéines artificielles dans les cellules qui, de plus, peuvent modifier le génome.

C.3 Les polyamides ou MGBs (Minor Groove Binders)

fluorescents

Le dernier type d'approche décrit dans ce manuscrit pour visualiser l'ADN db est l'utilisation de petites molécules synthétiques composées par des monomères Nméthylpyrrole (Py) et N-méthylimidazole (Im). Leur structure a été inspirée de molécules naturelles capables d'interagir avec la double hélice d'ADN au sein du petit sillon. Pour simplifier, dans ce manuscrit les termes polyamides et MGBs (Minor Groove Binders) feront référence à des polyamides composés par des monomères Py et Im. Compte tenu de leur capacité à interagir avec l'ADN au sein du petit sillon, les MGBs ont été identifiés comme de bons candidats pour la construction de sondes fluorescentes reconnaissant spécifiquement des séquences d'ADN (voir § Chapitre I - D pour la description des polyamides).

Chapitre I – Introduction Générale Les outils pour visualiser des séquences d'ADN double brin

Un grand nombre de polyamides conjugués à des molécules fluorescentes a été caractérisé pour des études allant de la vérification de leur interaction avec des séquences d'ADN modèles jusqu'à la détermination de leur localisation cellulaire (Hsu & Dervan 2008; Vaijavanthi et al. 2012). Le groupe de Dervan a publié des résultats montrant que des polyamides conjugués à la TétraMéthylRhodamine (TMR) via un monomère interne Py (Figure 13A) peut produire un signal fluorescent spécifique en présence du duplexe cible d'ADN. Ces conjugués fluorescents sont largement « éteints » en solution ; cependant, lorsqu'ils interagissent avec l'ADN cible, une augmentation de l'intensité de fluorescence supérieure à 10 fois est observée ; alors qu'avec un duplexe d'ADN contrôle (présentant des mésappariements), le signal observé est atténué (Rucker et al. 2003; Rucker et al. 2004). Le même effet d'augmentation spécifique de l'intensité de fluorescence ainsi qu'un effet FRET (Transfert par Résonance d'Energie de Fluorescence), ont été observés avec des polyamides couplés à des fluorophores de types cyanine (Cy3 et Cy5) lors de leur interaction avec un nucléosome (Han et al. 2014). Un effet FRET est observable quand deux fluorophores (donneur et accepteur) sont situés à proximité l'un de l'autre. Suite à l'excitation du fluorophore donneur, il y a un transfert d'énergie vers le fluorophore accepteur qui, excité, va émettre un signal de fluorescence à une longueur d'onde différente. Sur le nucléosome, la possibilité que des polyamides soit situés proches l'un de l'autre, avait déjà été vérifiée et les structures cristallographiques de complexes nucléosomes-polyamides avaient été publiées (Gottesfeld et al. 2001; Suto et al. 2003). La détection de séquences d'ADN spécifiques, par fluorescence, a été étendue en utilisant des fluorophores intercalants (Fechter et al. 2005). En couplant le fluorophore thiazole orange (TO) en C-term d'un polyamide (Figure 13B), une augmentation (1000 fois) de la fluorescence a été observée uniquement en présence du duplexe cible d'ADN. De plus, le conjugué polyamide-TO est capable de perturber le duplexe d'ADN en le déroulant d'un angle de 8°.



Figure 13 : Exemples de sondes fluorescentes à base de polyamides

A/ Modèle schématique de l'interaction entre un polyamide couplé à la tétraméthylrhodamine (TMR) et sa séquence cible d'ADN. Le conjugué polyamide-TMR est faiblement fluorescent en absence de l'ADN mais une augmentation spécifique de la fluorescence, représentée en jaune, est observée en présence de la séquence cible d'ADN. Les Py et Im sont représentés par les boules blanches et noires respectivement, la β -alanine par un losange blanc et le demi-cercle avec un + représente la Dp (diméthylaminopropylamine) (Rucker et al. 2003). B/ Représentation d'un conjugué polyamide-thiazole orange (TO) dans le petit sillon de l'ADN db. Les liaisons hydrogène sont représentées par des lignes pointillées, les cercles avec deux points représentent la paire d'électrons libres des N-3 (azote en position 3) des purines et O-2 (oxygène en position 2) des pyrimidines et les cercles avec un H représentent le groupement amine exocyclique des guanines. L'émission de la fluorescence est représentée en vert (Fechter et al. 2005). C/ Structure d'un polyamide fluorescent ayant intégré des « dimères fusionnés » (en gras) (Chenoweth et al. 2007).

Une autre stratégie pour construire une sonde fluorescente a été d'intégrer la partie fluorescente de la sonde directement dans le « corps » du polyamide comme module de reconnaissance (Chenoweth et al. 2007). Excités à 340 nm, des polyamides ayant incorporés des « dimères fusionnés » (**Figure 13C**) ont également montré une augmentation du signal fluorescent en présence de l'ADN cible. Dans le même contexte, un polyamide contenant un dérivé du Hoechst 33258, Hx (2-(4-méthoxyphényl)-6-carboxamide aussi appelé p-anisylbenzimidazolecarboxamide) a été développé. En ciblant les paires de bases A/T, Hx

émet une lumière bleue après une excitation à 322 nm et permet la mesure directe de l'interaction du polyamide avec l'ADN (Chavda et al. 2011) ; l'utilisation de groupement pyrène a également été largement étudié pour construire des sondes fluorescentes à base de polyamides (Vaijayanthi et al. 2013).

Plusieurs études ont utilisé des conjugués polyamide-fluorophore pour étudier la localisation de loci chromosomiques spécifiques dans des cellules de Drosophile et de mammifère et sur des étalements chromosomiques (Janssen, Durussel, et al. 2000; Janssen, Cuvier, et al. 2000; Maeshima et al. 2001; Gygi et al. 2002). Le groupe de Laemmli a synthétisé des polyamides poly-pyrroles ciblés vers les séquences riches en A et T des satellites I et III de la *Drosophile*. Aussi, un polyamide composé de monomères Im et β -ala, ciblé vers le satellite V riche en bases puriques (A et G) des chromosomes polytènes de la Drosophile, a été développé (Janssen, Cuvier, et al. 2000). Couplés à la fluorescéine, ces conjugués ont été utilisés pour confirmer leur pénétration nucléaire et vérifier le marquage des chromosomes. Dans les noyaux de cellules fixées de Drosophile, chacun des polyamides a été capable de marquer sa séquence cible. De plus, par une augmentation de la sensibilité aux enzymes de restriction et à la topoisomérase II, il a été montré que le polyamide poly-pyrrole était capable d'induire l'ouverture de domaines chromatiniens (Janssen, Durussel, et al. 2000). Ce même groupe a synthétisé des tandems de polyamides couplés au fluorophore Texas Red, ciblés vers les séquences télomériques d'insectes et de vertébrés. Toujours sur des cellules fixées, la localisation des conjugués fluorescents au niveau de leur cible a été vérifiée par des expériences de colocalisation avec la protéine télomérique TRF1 dans des cellules humaines Hela (Maeshima et al. 2001). Plus récemment, afin de cibler des séquences d'ADN plus longues dans le but d'améliorer la spécificité de ces molécules, des tandems de polyamides ciblant des séquences télomériques ont été couplés à différents fluorophores (TAMRA, Cyanines, Texas Red) afin de pouvoir visualiser les télomères sur des étalements chromosomiques et des cellules fixées (Kawamoto et al. 2013; Hirata et al. 2014; Kawamoto et al. 2015).

L'équipe de Trask a utilisé des polyamides conjugués à la fluorescéine pour différencier des chromosomes par microscopie et cytométrie de flux en ciblant des séquences hétérochromatiennes abondantes au niveau des chromosomes humains 1, 9 et Y (Gygi et al. 2002). En parallèle, des expériences de FISH ont été réalisées avec des oligonucléotides fluorescents afin de vérifier que les profils de marquage obtenus avec les polyamides

conjugués étaient similaires. L'avantage des polyamides est qu'ils ne nécessitent pas l'étape de dénaturation de l'ADN pour marquer les chromosomes. Ainsi, ce type d'études a permis de mettre en évidence le rôle potentiel des polyamides conjugués comme outils de diagnostic (par exemples, le criblage prénatal ou les cancers) (Gygi et al. 2002).

Des études sur la possibilité d'utiliser les polyamides conjugués comme des outils potentiels de diagnostic de maladies présentant des répétitions de triplets ont été publiées (Bando et al. 2007; Fujimoto et al. 2008). À titre d'exemple, la maladie de Huntington est caractérisée par une expansion du nombre de triplets CAG (le nombre de répétitions est de 20 chez un individu sain et supérieur à 40 chez un individu malade). La conjugaison de groupements pyrène à deux résidus pyrrole d'un polyamide par un bras de liaison butyle, a permis la conception d'un conjugué capable de reconnaitre des répétitions CAG avec le profil d'émission de fluorescence d'un excimère à 475 nm directement relié à la concentration de la séquence cible. Pour synthétiser une sonde fluorescente avec une longueur d'onde d'excitation plus longue et un profil d'émission plus stable, la modification du bras de liaison a été réalisée en utilisant un lien alcyne qui est plus rigide qu'un lien butyle. Ainsi la longueur d'onde d'excitation a été décalée vers le rouge et la longueur d'onde d'émission du conjugué était d'environ 495 nm avec une intensité maintenue (85%) pendant 4h (Fujimoto et al. 2008).

D - Les polyamides ou MGBs

D.1 L'histoire des polyamides et l'établissement des règles de reconnaissance

Deux produits naturels, la nétropsine et la distamycine A, sont à l'origine du développement des polyamides. Ces deux petites molécules en forme de croissant, composées par plusieurs résidus N-méthylpyrroles (Py) sont capables d'interagir avec des séquences d'ADN riches en A et T avec des affinités et des spécificités modestes (Arcamone et al. 1964; Patel 1982; Dickerson & Kopka 1985; Kopka et al. 1985a; Kopka et al. 1985b; Zimmer & Wähnert 1986; Pelton & Wemmer 1989). L'analyse des interactions entre ces molécules et l'ADN a permis le développement d'oligomères synthétiques composés de répétitions de monomères N-méthylpyrroles (Py) et N-méthylimidazoles (Im), appelés polyamides ou MGBs. Ces molécules de synthèse sont capables d'interagir avec des séquences d'ADN prédéfinies avec des affinités et des spécificités comparables à celles de protéines qui se lient à l'ADN (Wemmer & Dervan 1997; Dervan & Bürli 1999; Dervan 2001; Best et al. 2003; Melander et al. 2004; Dervan, Doss, et al. 2005).

Dans la partie exposée du petit sillon de la double hélice d'ADN, le plan des paires de bases est perpendiculaire aux molécules qui viennent se lier ; ainsi, un seul côté des paires de bases est accessible (l'autre côté est accessible dans le grand sillon). Ceci implique la présence, au sein de ce petit sillon, d'un environnement électronique unique qui dépend de la séquence nucléotidique (**Figure 14**) et qui peut être reconnu spécifiquement par les molécules.



Figure 14 : L'environnement électronique du petit sillon de l'ADN

A/ Pour les paires de bases A•T et T•A, A dispose d'une paire d'électrons libres alors que T en a deux, et pour les paires de bases G•C et C•G, C dispose d'une paire d'électrons libres et G en possède une ainsi qu'une liaison hydrogène potentielle portée par la fonction amine exocyclique. Les cercles avec deux points représentent la paire d'électrons libres des N-3 (azote en position 3) des purines et O-2 (oxygène en position 2) des pyrimidines, et les cercles avec un H représentent le groupement amine exocyclique des guanines. B/ Schéma de l'environnement électronique des paires de bases Watson-Crick dans le petit sillon de l'ADN pour une séquence 5'-CTAG-3' (adapté de (Montgomery 2013)).

En 1985, l'équipe de Dickerson a publié la structure cristallographique de la nétropsine (**Figure 15A**) en complexe avec l'ADN db (Dickerson & Kopka 1985). La structure indiquait que la nétropsine, qui est composée de deux monomères pyrroles (Py : N-méthylpyrrole), interagit avec l'ADN dans une stœchiométrie 1:1 en déplaçant le réseau de molécules d'H₂O (« the spine of hydration »), le long du site de liaison, dans le petit sillon. Elle se lie via des liaisons hydrogène faisant intervenir le proton de la liaison amide (N-H), l'azote en position 3 (N-3) de l'adénine et l'oxygène en position 2 (O-2) de la thymine dans la séquence riche en A et T et des interactions de van der Waals entre le C-H du Py et le C2 de l'adénine (Kopka et al. 1985b). De cette structure, il a été supposé que le remplacement d'un monomère Py par un monomère par un atome d'azote) peut considérablement étendre le code de reconnaissance de ces oligomères, en permettant la reconnaissance des paires C•G et G•C par la formation d'une liaison hydrogène entre le N2 (azote en position 2) de l'Im et le proton de l'amine exocyclique des guanines (**Figure 15B**) (Kopka et al. 1985b; Dervan 1986).

Deux analogues de la distamycine A, qui est une molécule très similaire à la nétropsine, ont été synthétisés et testés pour leur capacité à interagir avec l'ADN. Le premier était un trimère Py-Im-Py, par Lown et son équipe, qui avait une efficacité et une spécificité moyenne (Lown et al. 1986). Le second était un trimère Im-Py-Py, par Wade et Dervan, qui marqua un tournant dans la conception des pyrrole-imidazole polyamides comme ligands synthétiques de l'ADN (Wade & Dervan 1987; Wade et al. 1992). Basé sur un modèle d'interaction avec une stœchiométrie 1:1, il était supposé que le deuxième analogue pouvait interagir avec une séquence d'ADN G (A/T) (T/A). Cependant, ce ligand pouvait aussi se lier à une séquence G (A/T) C. Ces résultats remettaient en question l'hypothèse du mode d'interaction avec une stœchiométrie 1:1 de ces ligands dans le petit sillon de l'ADN.

Un élément clé pour répondre à ce dilemme a été trouvé par le groupe de Wemmer qui a démontré, par des études de RMN (Résonance Magnétique Nucléaire), qu'à forte concentration de distamycine A, un autre mode d'interaction était possible. La distamycine A est composée de trois monomères Py et les structures cristallographiques et RMN de complexes distamycine A/ADN ont été étudiées en détails (Klevit et al. 1986; Pelton & Wemmer 1988; Pelton & Wemmer 1989; Pelton & Wemmer 1990; Mitra et al. 1999). Son interaction avec l'ADN est plus complexe, que celle de la nétropsine, puisqu'elle peut interagir selon deux steechiométries différentes soit 1:1 soit 2:1 (distamycine : ADN) par des liaisons hydrogène entre le proton de la liaison amide (N-H) de la distamycine A et les paires d'électrons libres des bases de l'ADN exposées dans le petit sillon (Figures 15B et C). Dans la stechiométrie 2:1, les deux molécules de distamycine A se lient à l'ADN de façon antiparallèle, avec une orientation N-term/C-term qui suit le sens 5'-3' de la séquence d'ADN et chacune de leurs liaisons amides forment des liaisons hydrogène favorables avec les paires d'électrons libres présentes du même côté du petit sillon (Figure 15C). Ces informations structurales tirées des structures RMN et cristallographique révèlent une interaction favorable entre l'ADN et la distamycine A permettant à deux molécules de se lier à une même séquence d'ADN dans le petit sillon. Les monomères Py d'une molécule de distamycine A sont « décalés » par rapport à ceux de l'autre molécule de distamycine A afin que le réseau de liaisons hydrogène puisse se faire favorablement.



Figure 15 : Structures et modes d'interaction des petites molécules dans le petit sillon de la double hélice d'ADN

A/ Structures chimiques de la nétropsine et de la distamycine A. B/ et C/ Représentations schématiques de la reconnaissance du petit sillon de l'ADN par la distamycine dans B/ une stœchiométrie 1:1 et C/ une stœchiométrie 2:1. Les liaisons hydrogène sont représentées par des lignes pointillées et les cercles avec deux points représentent la paire d'électrons libres des N-3 des purines et O-2 des pyrimidines et les cercles avec un H représentent le groupement amine exocyclique des guanines. En bas, les cercles vides sont la représentation schématique des Py et les demi-cercles avec un + représentent les groupements guanidine (adapté de (Dervan, Poulin-Kerstein, et al. 2005)).

Inspirées des structures de la nétropsine et de la distamycine, une nouvelle classe d'oligomères hétérocycliques appelés pyrrole-imidazole polyamides (Py-Im) ou MGBs a été développée pour interagir avec une séquence d'ADN cible de manière programmée. Une paire de monomères Py/Py peut reconnaitre les paires de bases A•T ou T•A, et en remplaçant un monomère Py par un monomère Im, les séquences contenant des paires de bases G•C ou C•G peuvent être reconnues (**Figure 16**) (Kielkopf, Baird, et al. 1998).

D'autres études ont montré que la liaison covalente entre l'extrémité N-term d'un oligomère et l'extrémité C-term d'un autre dans le complexe « polyamide : ADN » (2 : 1), permet d'augmenter l'affinité de ces oligomères pour la séquence cible d'ADN et, en même temps, d'étendre la gamme de séquences reconnues. Pour former la liaison covalente entre les deux oligomères, l'insertion d'un acide γ -aminobutyrique (GABA) permet la formation d'un « coude γ » ; ainsi le polyamide se structure dans une conformation « d'épingle à cheveux » qui procure à la molécule une entropie favorable. À titre d'exemple, le couplage d'un trimère

Im-Py-Py à la distamycine via un acide γ -aminobutyrique a permis de cibler une séquence d'ADN 5'-G(A/T)(A/T)-3' avec de meilleures affinité et spécificité que la distamycine seule (Mrksich et al. 1994; Mrksich & Dervan 1994). Il a été montré que ce motif structural en « épingle à cheveux » permettait d'augmenter de 100 à 3600 fois l'affinité du polyamide pour sa séquence cible d'ADN comparé à une interaction de type 2 :1 (**Figure 16A**) (Trauger et al. 1996). De plus, l'insertion du coude γ permet le ciblage de séquences d'ADN nonpalindromiques en incorporant des paires de résidus asymétriques et de façon orientée. Ainsi, le polyamide reconnait la séquence d'ADN de son extrémité N-term à son extrémité C-term dans le sens 5'-3' (White et al. 1997).



Figure 16 : Règles de reconnaissance dérivées d'un complexe polyamide : ADN (2 :1)

A/ Modèle de l'interaction d'un dimère Im-Py-Py antiparallèle avec la séquence d'ADN db 5'-TGTCA-3'. Les liaisons hydrogène potentielles entre les polyamides et l'ADN sont représentées par des lignes pointillées, les cercles avec deux points représentent la paire d'électrons libres des N-3 des purines et O-2 des pyrimidines et les cercles avec un H représentent le groupement amine exocyclique des guanines. B/ Orientation et liaisons hydrogène d'un dimère Im/Py en interaction avec une paire de bases G•C. Les interactions entre le N-H de la liaison amide et la paire d'électrons libres (N-3 des purines et O-2 des pyrimidines) de l'ADN est clairement visible. De plus, l'interaction entre l'azote hétérocyclique de l'Im et le N-H de la fonction amine exocyclique de la guanine est aussi visible. Les liaisons hydrogène entre les paires de bases Watson-Crick sont représentées et les distances sont données en Angström. En bas, les cercles vides et noirs sont la représentation schématique des Py et Im respectivement, et les demi-cercles avec un + représentent la Dp (N,N-diméthylaminopropylamine) (Chenoweth & Dervan 2009).

Finalement, un ensemble de règles d'appariements a donc été décrit. Ainsi, une séquence d'ADN programmée peut être reconnue spécifiquement : une paire N-méthylpyrrole/N-méthylimidazole (Py/Im) cible et reconnait une paire de bases C•G, une

paire Im/Py reconnait une paire G•C, et une paire Py/Py peut reconnaitre les paires de bases A•T ou T•A (**Figure 16**) (White et al. 1996). De plus, l'absence d'interaction avec des duplexes d'ARN synthétiques a été vérifiée par des techniques de retard sur gel ou par dénaturation thermique des acides nucléiques, même si par SPR une faible interaction a été observée, donc ces molécules ont une grande préférence de reconnaissance pour l'ADN (Chenoweth et al. 2013; Iguchi et al. 2013).

D.2 Vers de nouveaux ligands avec des monomères alternatifs

Les études structurales des complexes polyamides/ADN ont permis d'obtenir des informations sur les contraintes structurales de la formation de ces complexes et ainsi de faire des découvertes sur les modifications nécessaires à l'amélioration de la reconnaissance moléculaire.

D.2.1 Les monomères β-alanines

Tout d'abord, un polyamide standard est composé par 4 paires de monomères Py/Im pour des questions de rigidité structurale due au(x) monomère(s) Im. Ainsi, le pas de l'hélice d'ADN est inférieur à celui du polyamide et pour éviter que le polyamide soit trop long pour permettre les interactions entre monomères et paires de bases, il se « courbe » ; ce qui provoque une certaine tension dans la molécule et influe ainsi sur son affinité et sa spécificité (Han et al. 2012). Ce problème a pu être résolu par l'introduction d'un nouvel élément de reconnaissance : le monomère β -alanine (β -ala) (Wang et al. 2001; Wang et al. 2012; Wang et al. 2 al. 2014). Inséré dans un polyamide, le monomère β-alanine a les mêmes propriétés de reconnaissance que le monomère Py ; sa nature flexible permet essentiellement au polyamide de se repositionner pour suivre la courbure et le pas de l'hélice lors de l'interaction et une étude a montré que ce monomère pouvait aussi améliorer l'affinité du polyamide pour sa séquence cible d'ADN (Y.-W. Han et al. 2013). Des polyamides reconnaissant des séquences d'ADN allant jusqu'à 16 pb (longueur d'une séquence potentiellement unique dans le génome humain) ont ainsi pu être synthétisés avec des affinités semblables à celles de protéines capables d'interagir avec l'ADN (de l'ordre du nanomolaire) (Trauger et al. 1998; He et al. 2014).

D.2.2 Le coude γ

Ensuite, pour substituer le monomère γ du coude du polyamide en « épingle à cheveux », le développement d'acides γ -aminobutyrique chiraux substitués (comme le DABA : acide 2-4-diaminobutyrique) permet une meilleure interaction avec l'ajout d'une charge positive sur le coude γ . De plus, cette fonction amine supplémentaire peut être utilisée comme point d'attachement pour des modifications chimiques ou encore pour le couplage d'un deuxième polyamide (voir § Chapitre I - D.3.2.2 pour la description des tandems de polyamides) (Herman et al. 1998; Hsu et al. 2007; Dose et al. 2008).

D.2.3 Les monomères Hp (hydroxypyrrole)

Enfin, l'incorporation de monomères hydroxypyrrole (Hp: N-méthyl-3-hydroxy pyrrole) permet de distinguer les paires de bases A•T des paires T•A (**Tableau 1**) (White et al. 1998; Kielkopf, White, et al. 1998; Kielkopf et al. 2000). Les paires Hp/Py et Py/Hp ciblent les paires T•A et A•T, respectivement. Les monomères Hp sont spécifiques du nucléotide T grâce à l'interaction entre le groupe hydroxyle (O-H) de l'Hp et la paire supplémentaire d'électrons libres des thymines dans le petit sillon (**Figure 17**) (Dervan & Edelson 2003). Cependant, l'utilisation de ces monomères Hp reste limitée par leur instabilité chimique au cours du temps. Leur dégradation prend quelques semaines ou quelques mois lorsqu'ils sont stockés sous forme solide mais seulement quelques jours en solution (Dervan & Edelson 2003; Melander et al. 2004).

Tableau 1 : Règles de reconnaissance pour cibler les différentes paires de bases de l'ADN avec les Py-Im polyamides				
	C•G	G•C	A•T	T•A
Py/Im	+	-	-	-
Im/Py	-	+	-	-
Py/Hp	-	-	+	-
Hp/Py	-	-	-	+
Py/Py	-	-	+	+



Figure 17 : Représentation schématique d'un complexe polyamide/ADN

Vue perspective et schématique du petit sillon de l'ADN. Pour faciliter la visibilité du réseau de liaisons hydrogène et la complémentarité des structures de l'ADN et du polyamide en conformation « d'épingle à cheveux antiparallèle », l'hélice a été représentée sans torsion, avec une courbure diminuée et le squelette phosphate de chacun des brins est représenté par une ligne annotée par les extrémités 5' et 3'. Le groupement R représente la β -Dp (Dervan & Edelson 2003).

D.2.4 Les monomères dérivés

Après avoir établi que la discrimination entre les paires A•T et T•A était possible, de nombreux efforts ont été effectués dans le but de concevoir des monomères qui permettraient de discriminer ces paires de bases sans l'instabilité chimique observée avec les monomères Hp. Dans une étude, plusieurs polyamides composés de 4 paires de monomères ont été synthétisés en incorporant, pour former une paire de reconnaissance avec un monomère Py, un des 8 acides aminés aromatiques suivants: 1-méthyl-1H-pyrazole (Pz), 1H-pyrrole (Nh), 5-méthylthiazole (Nt), 4-méthylthiazole (Th), 3-méthylthiophène (Tn), thiophène (Tp), 3-hydroxythiophène (Ht) et le furane (Fr) (**Figure 18A**). La paire Py/Tn était sélective pour les

paires de bases A•T et ne montrait pas d'interaction avec les paires G•C. Cependant les monomères Th, Fr et Ht empêchaient l'interaction avec l'ADN. Les études de modélisation moléculaires suggéraient que tous ces monomères dérivés induisaient des altérations nuisibles pour la forme et les propriétés électroniques des polyamides ; ce qui implique que la gamme de monomères disponibles pour construire des polyamides est limitée (Marques et al. 2002). Une autre étude a montré la sélectivité pour les paires T•A de dérivés des résidus thiophène (comme le chlorothiophène Ct) lorsqu'ils sont placés à l'extrémité N-term des polyamides (Foister et al. 2003).

D.2.5 Les analogues des paires de monomères

Le Hoechst est un colorant de l'ADN ayant la capacité à se lier dans le petit sillon des séquences riches en A et T ; sa structure est composée de benzimidazoles (Dervan & Edelson 2003). Inspirés par cette molécule et sa capacité à interagir avec l'ADN, des benzimidazoles substitués ont aussi été étudiés pour leur capacité à élargir la gamme de monomères construire des polyamides. L'hydroxybenzimidazole disponibles pour (Hz) et l'imidazopyridine (Ip) ont été fusionnés au Py sans liaison amide intermédiaire (Figure 18B) et introduit dans des polyamides pour remplacer deux monomères (Renneberg & Dervan 2003). Les paires Ip/Py et Py/Ip ont montré une sélectivité pour les paires de bases G•C et C•G respectivement, ce qui indique qu'ils sont des analogues des paires de monomères « parentaux » Im/Py et Py/Im. La paire Hz/Py a montré une bonne sélectivité pour la paire T•A, avec une perte d'affinité par rapport à la paire Hp/Py mais une meilleure stabilité au cours du temps. Ainsi, avec la caractérisation de ces dérivés de benzimidazoles, de nouveaux monomères, pouvant élargir la gamme de monomères disponibles pour la construction des polyamides, ont été développés.





A/ Les monomères « parentaux » Py, Im et Hp sont encadrés et leurs 8 dérivés sont représentés. B/ Structures des paires de monomères usuelles (à gauche) et de leurs dérivées benzimidazoles (à droite).

D.3 La synthèse et les optimisations des polyamides

D.3.1 Les différents types de synthèse

Les premières synthèses de polyamides étaient effectuées en phase liquide et nécessitaient des étapes longues et difficiles qui pouvaient prendre plusieurs mois pour la synthèse et la purification d'un polyamide (Nishiwaki et al. 1988). Une avancée majeure pour faciliter cette procédure a été le développement d'une méthode de synthèse sur support solide reportée par Baird et Dervan (Baird & Dervan 1996). Cette méthode permettait :

- d'une part, de diminuer le nombre d'étapes puisque la synthèse des monomères à l'échelle du gramme ne nécessitait plus de purifications chromatographiques préalables à leur utilisation pour les couplages, permettant ainsi de construire un polyamide,

- d'autre part, l'échelle de temps était considérablement diminuée puisque la synthèse
d'un polyamide se faisait en quelques jours ou semaines plutôt qu'en mois (Baird & Dervan 1996).

D'autres protocoles de synthèse ont été élaborés afin de perfectionner cette approche de synthèse sur support solide ; comme celui de Krutzik et Chamberlin, utilisant des méthodes d'activation de monomères différentes (activation par le 1-HydrOxy-7-Azabenzotriazole (HOAt) plutôt que par le N-HydrOxyBenzotriazole (HOBt)), toujours dans l'optique de rendre la synthèse des polyamides plus simple et plus rapide (Krutzik & Chamberlin 2002a; Krutzik & Chamberlin 2002b) ou encore l'utilisation de monomères protégés par des groupements Fmoc (Fluorénylméthyloxycarbonyle) plutôt que Boc (*tert*-Butoxycarbonyle) (Wurtz et al. 2001).

Les premières synthèses supportées de polyamides reportées utilisaient une résine Boc- β ala-PAM (résine PAM (4-hydroxyméthyle-PhénylAcétamidoMéthyle) estérifiée par une Boc- β -alanine) compatible avec l'utilisation de monomères protégés par des groupements Fmoc et Boc. L'enchainement des monomères sur la résine permet la synthèse d'un polyamide, suivi du clivage de la résine par aminolyse en utilisant de la N,Ndimétylaminopropylamine (DMPA), ce qui procure une queue C-term β -Dp au polyamide (Baird & Dervan 1996). Des études de l'interaction séquence-spécifique ont montré que le coude γ ainsi que la queue C-term (β -Dp) peuvent aussi cibler des paires de bases A•T ou T•A ; ainsi, un polyamide standard composé de 8 monomères Py-Im peut reconnaitre spécifiquement une séquence d'ADN de 6 pb (**Figure 19**) (Swalley et al. 1999).



Figure 19 : Modèle schématique d'un polyamide en « épingle à cheveux » antiparallèle en complexe avec le duplexe cible d'ADN

Le polyamide ImHpPyPy-(R)^{H₂N} γ -ImHpPyPy- β -Dp lié dans le petit sillon de la séquence cible d'ADN 5'-TGTACA-3' selon les règles de reconnaissance des paires de bases Watson-Crick. Les liaisons hydrogène potentielles sont représentées par des lignes pointillées, les cercles avec deux points représentent la paire d'électrons libres des N-3 des purines et O-2 des pyrimidines et les cercles avec un H représentent le groupement amine exocyclique des guanines. La reconnaissance des bases A ou T par la queue C-term β -Dp et par le coude γ permet d'allonger la séquence cible d'ADN reconnue de 2 pb. Ainsi un polyamide composé de 4 paires de monomères peut reconnaitre une séquence cible d'ADN de 6 pb (Dervan, Poulin-Kerstein, et al. 2005).

Pour élargir la gamme de séquences d'ADN reconnues par les polyamides, l'utilisation d'une résine « oxime de Kaiser » (oxime dérivée de la 4-nitrobenzophénone), lors de la synthèse permet d'utiliser plusieurs composants pour le clivage nucléophile du polyamide sur la résine. Ainsi, plusieurs queues C-term (**Figure 20A**) ont été obtenues. L'influence de la queue C-term sur la spécificité des polyamides a ainsi pu être testée. Les polyamides contenant une queue C-term acétamide, plus courte que la queue β -Dp (**Figure 20B**),

montrent une affinité similaire mais avec une reconnaissance des paires de bases A•T et T•A diminuée (Belitsky, Nguyen, et al. 2002).



Figure 20 : Les différents variants des queues C-term des polyamides

A/ Exemple d'un polyamide avec une queue C-term variable (R= NHMe ou NHEt ou OH ou NH₂) obtenue après clivage nucléophile d'une résine Kaiser oxime. B/ Exemple d'un polyamide avec une queue C-term β -Dp obtenue par le clivage aminolytique d'une résine β -ala-PAM (adapté de (Blackledge & Melander 2013)).

Plus récemment, la synthèse sur support solide assistée par micro-ondes est souvent privilégiée car elle permet de former rapidement des produits purs avec de bons rendements. L'équipe de Dervan a reporté, récemment, la première synthèse assistée par micro-ondes d'un Py-Im polyamide contenant un coude DABA pour former « l'épingle à cheveux ». En utilisant une résine « oxime de Kaiser » et la chimie Boc, un polyamide a été synthétisé en diminuant significativement les temps de couplage, de déprotection et d'acétylation des amines, en opérant sous micro-ondes (Puckett et al. 2012).

Bien que le développement des méthodes de synthèse sur support solide a permis de faciliter et de rendre plus rapide la production de quantités nécessaires de polyamides pour des études *in vitro* ou en culture cellulaire, les quantités produites n'étaient pas suffisantes pour des études pilotes sur animaux. Pour surmonter cette difficulté, de nouvelles méthodes de synthèse en phase liquide ont été mises au point ; à titre d'exemple, un polyamide à 4 paires de monomères ciblant une séquence du récepteur aux androgènes a été synthétisé. En utilisant une synthèse convergente, l'équipe de Dervan a exploité les différences de solubilité entre les réactifs et les produits de réaction afin d'isoler le matériel pur par précipitation ; ainsi, le polyamide a été synthétisé à l'échelle de plusieurs grammes sans purification chromatographique et avec une queue C-term modifiable (Chenoweth, Harki & Dervan 2009). Une autre manière de synthètiser les polyamides a été décrite très récemment ; elle utilise à la fois des étapes de synthèse sur support solide et des étapes en phase liquide. Un fragment de 4 monomères a été synthétisé sur support solide en utilisant une nouvelle résine

PAM sans β -alanine, le deuxième bloc de 4 monomères a été synthétisé en phase liquide et pour finir les deux fragments ont été couplés sur le support solide, une étape d'aminolyse libérant ensuite le polyamide de son support (Fallows et al. 2014).

D.3.2 L'optimisation des polyamides

D.3.2.1 Les polyamides cycliques

Des polyamides de nature différente ont été développés avec l'apparition de polyamides cycliques (Baliga et al. 2001; Raskatov, Hargrove, et al. 2012). L'ajout d'un coude γ supplémentaire permet au polyamide en « épingle à cheveux » de se fermer par cyclisation. Dans l'étude de Chenoweth, la synthèse et l'étape de cyclisation du polyamide ciblant le récepteur aux androgènes s'effectue en phase liquide (Chenoweth, Harki, Phillips, et al. 2009), alors que, dans l'étude de Li, la synthèse du polyamide s'effectue sur support solide assistée par micro-ondes tandis que l'étape de cyclisation a lieu en phase liquide (**Figure 21A**) (Li et al. 2013).



Figure 21 : Exemples de polyamides cycliques

A/ Structure d'un polyamide cyclique composé de 4 paires de monomères et pouvant cibler une séquence d'ADN de 6 pb (Li et al. 2013). B/ Structure d'un petit polyamide cyclique capable d'interagir avec un G-quadruplexe (Zhang et al. 2012).

D'après ces travaux, l'affinité du polyamide cyclique est meilleure que celle du même polyamide en « épingle à cheveux ». Un autre type de polyamide cyclique, de petite taille (4 monomères Py et 2 monomères β -ala), ciblant une structure secondaire d'ADN en G-quadruplexe (au niveau du gène c-myb), a été synthétisé en phase liquide. L'interaction avec cette structure non-canonique d'ADN (G-quadruplexe) a été validée par des études de RMN et de spectrométrie (**Figure 21B**) (Zhang et al. 2012).

D.3.2.2 Les polyamides longs et les tandems de polyamides

Une limite à l'utilisation des polyamides standard, composés de 4 paires de monomères, est leur longueur ; en effet, ils ne peuvent cibler que des séquences d'ADN de 5-6 pb. Cependant, des efforts pour contrer ce problème ont été faits. Ainsi, la synthèse de polyamides longs constitués par 6 à 8 paires de monomères ont permis la reconnaissance de séquences d'ADN allant jusqu'à 10 pb. Ils contiennent des monomères β -alanine pour rendre la molécule plus flexible et permettre un bon alignement des paires de monomères pour l'interaction avec les bases de l'ADN (Wang et al. 2010; Edwards et al. 2011; He et al. 2014; Taniguchi et al. 2014; Mishra et al. 2015).

Toujours dans le but de reconnaitre des séquences d'ADN plus longues et donc d'améliorer la spécificité, une autre stratégie a émergé avec la synthèse de tandems de polyamides ; 3 exemples sont présentés dans la **Figure 22**. Le premier exemple est une étude originale présentant la conception d'un tandem de 2 polyamides avec une orientation « tête à queue ». Pour construire ce tandem, deux polyamides fonctionnalisés par des groupements alcyne et azoture ont été synthétisés et conjugués par une réaction de « chimie-click » : la réaction de Huisgen qui est une cycloaddition 1,3-dipolaire. La formation d'un cycle triazole permet de construire un tandem capable de reconnaitre une séquence d'ADN de 10 pb. L'originalité de ce travail est, qu'en plus d'utiliser la « chimie-click » pour conjuguer les deux molécules, la réaction a été faite directement sur la matrice d'ADN au niveau de la séquence cible d'ADN (**Figure 22A**) (Poulin-Kerstien & Dervan 2003).

Le deuxième exemple est celui d'un autre type de tandem construit avec une orientation « tête à tête ». Les deux polyamides composant ce tandem ont été synthétisés sur support solide puis couplés en phase liquide (**Figure 22B**). De plus, le bras de liaison entre les deux polyamides contient une fonction amine qui peut porter des groupements fonctionnels ou encore permettre l'attachement d'un fluorophore (Halby et al. 2005; Boutorine et al. 2007).

Aussi, les travaux du groupe de Sugiyama ont permis de développer de nouveaux tandems composés de 2 polyamides ou même un trimère composé de 3 polyamides de 3 paires de monomères ciblant des séquences d'ADN télomériques humaines de 12 à 18 pb (**Figure 22C**) (Kawamoto et al. 2013; Kawamoto et al. 2015; Guo et al. 2015).



Figure 22 : Représentation schématique des tandems et du trimère de polyamides

A/ Modèle schématique de la réaction de couplage, par une cycloaddition 1,3-dipolaire (réaction de Huisgen), de deux polyamides fonctionnalisés par un groupement azoture (à droite) et un groupement alcyne (à gauche), sur la matrice d'ADN. La formation d'un cycle triazole (pentagone gris) permet de lier les deux polyamides pour former un tandem de polyamides, avec une orientation « tête à queue », capable de reconnaitre une séquence d'ADN de 10 pb (encadrée sur le schéma) (Poulin-Kerstien & Dervan 2003). B/ Modèle schématique d'un tandem de polyamides avec une orientation « tête à tête ». L'accolade représente le bras de liaison entre les deux polyamides (à partir de (Halby et al. 2005; Halby et al. 2007)). C/ Modèle schématique d'un trimère de 3 polyamides avec une orientation « tête à queue ». L'hexagone gris représente la charnière entre les polyamides. Les Py et Im sont représentés par les cercles vides et noirs respectivement, la β -alanine par un losange blanc et le demi-cercle avec un + représente la Dp.

Enfin, un « tridem » de polyamides reconnaissant une séquence d'ADN de 14 pb et composé de 3 éléments (un polyamide cyclique central et 2 polyamides de 3 paires de monomères aux extrémités) a été synthétisé et sa capacité à interagir avec sa séquence cible a été vérifiée (Yamamoto, Bando, Morinaga, et al. 2014).
D.4 L'utilisation des polyamides comme régulateurs de l'expression des gènes

En plus de leur utilisation comme sondes fluorescentes lorsqu'ils sont conjugués à des fluorophores (voir § Chapitre I - C.3), les polyamides ont été utilisés comme des régulateurs de l'expression des gènes seuls ou couplés à des groupes fonctionnels. Chez l'homme, chaque cellule contient environ 20000 gènes dont la transcription est finement régulée par des facteurs de transcription, qui sont des protéines capables d'interagir avec l'ADN db de façon séquence-spécifique afin d'activer ou de réprimer la transcription, donc l'expression de gènes d'intérêt. Plusieurs maladies, dont les cancers, sont le résultat d'aberration(s) dans la régulation de l'expression de gènes spécifiques. C'est pourquoi, du fait de leur capacité à interagir avec des séquences d'ADN db de façon séquence-spécifique et avec de bonnes affinités, les polyamides ont été largement étudiés pour concevoir des outils synthétiques capables d'activer ou de réprimer l'expression de certains gènes. Le but ultime étant de pouvoir corriger des aberrations de l'expression génétique impliquées dans différentes maladies et permettre le développement de nouvelles options thérapeutiques pour de nombreuses maladies.

D.4.1 L'inhibition de la transcription par des polyamides

Une collaboration initiale entre les laboratoires de Dervan et Gottesfeld a permis de décrire les polyamides comme des inhibiteurs potentiels de la transcription via l'inhibition de l'interaction de certains facteurs de transcription avec leur séquence cible d'ADN. En utilisant un polyamide, composé par 8 monomères, ciblant le site de reconnaissance du facteur de transcription TFIIIA, la transcription du gène de l'ARN 5S a été supprimée *in vitro* et dans des cellules de foie de *Xenopus* (Gottesfeld et al. 1997). D'autres études sur la régulation de l'expression des gènes ont montré que des polyamides étaient capables d'inhiber la liaison de facteurs de transcription (comme par exemple Ets-1 ou la protéine virale Tax), et donc de réprimer l'expression de certains gènes et aussi d'empêcher la réplication du virus VIH1 (Virus de l'Immunodéficience Humaine) dans des lymphocytes humains (Dickinson et al. 1998; Dickinson, Trauger, Baird, Dervan, et al. 1999; Lenzmeier et al. 1999; Zhang et al. 2011; Syed et al. 2014).

L'inhibition de la transcription a ensuite été démontrée par des polyamides de 8 monomères ciblant la séquence de l'élément de réponse à l'hypoxie ou HRE (Hypoxia Response Element) (Olenyuk et al. 2004). Dans des conditions d'hypoxie, le facteur HIF-1 (Hypoxia Inducible Factor 1) forme un hétérodimère avec ARNT (Aryl hydrocarbon Receptor Nuclear Translocator). Cet hétérodimère se lie à l'élément HRE et permet le recrutement de co-activateurs transcriptionnels additionnels pour induire la transcription des gènes induits par l'hypoxie, incluant par exemple le gène du facteur de croissance VEGF (Vascular Endothelial Growth Factor). VEGF est un facteur angiogénique qui joue un rôle clé dans des cancers métastasiques. Le polyamide synthétisé porte, en N-term, un monomère 3-chlorothiophène, permettant la reconnaissance spécifique d'une paire T•A (Foister et al. 2003), et est conjugué en C-term à un fluorophore, la fluorescéine, pour les études d'imagerie. Les résultats obtenus par des expériences d'empreintes à la DNase I (« DNase Footprint ») ont validé d'une part l'interaction du polyamide avec la séquence du HRE et d'autre part, sa capacité à empêcher, in vitro, la fixation de l'hétérodimère HIF-1/ARNT à des concentrations nanomolaires. Les expériences cellulaires ont révélé que le traitement par le polyamide inhibit effectivement l'expression de la luciférase quand celle-ci était placée sous le contrôle du promoteur du gène VEGF. Les analyses de l'expression des gènes ont appuyé ces résultats puisque l'expression de l'ARNm du gène VEGF était diminuée d'environ 60% par rapport à celle d'autres gènes induits par l'hypoxie.

À partir de ces résultats prometteurs, Viger et Dervan ont synthétisé une série de polyamides portant plusieurs monomères modifiés (voir § Chapitre I - D.2.4 et D.2.5) (Viger & Dervan 2006). Un nouveau dimère benzimidazole-chlorothiophène a été introduit pour mimer la paire Py/chlorotiophène utilisée dans les études précédentes sur le HRE. À l'exception d'un polyamide composé uniquement par des paires de monomères modifiés, tous les polyamides démontraient des affinités comparables à celles des polyamides parentaux. Cette étude soulignait l'utilité des paires de monomères fusionnés (dérivés de benzimidazoles) pour augmenter la spécificité en maintenant une bonne affinité des polyamides pour leur séquence cible.

Un autre exemple d'inhibition de la transcription par des polyamides est celui du facteur de croissance TGF- β 1 (Transforming Growth Factor- β 1). TGF- β 1 est une cytokine régulatrice ayant des effets pléiotropiques dans la croissance, la prolifération, la différentiation cellulaire ou encore dans l'apoptose. Il est impliqué dans plusieurs maladies

rénales comme les gloméruloscléroses ou les néphrites (Blobe et al. 2000). Dans cette étude, un polyamide long, composé de 6 paires de monomères, a été conçu et synthétisé pour cibler le site de liaison d'un facteur de transcription dans la région promotrice de TGF-\beta1, présentant une homologie de séquence entre l'homme et le rat (Matsuda et al. 2006). L'affinité du polyamide pour sa séquence cible a été déterminée (de l'ordre du nanomolaire) par des expériences de retard sur gel et de résonnance plasmonique de surface (SPR : Surface Plasmon Resonance). Le traitement de cellules humaines HEK-293 avec le ligand a montré une diminution concentration-dépendante de l'activité de la luciférase quand celle-ci est placée sous le contrôle du promoteur de TGF-\beta1. Des expériences, *in vivo*, ont montré que le polyamide fluorescent se localise dans le noyau des glomérules et tubules du néphron 24h après son administration. Des rats nourris par un régime riche en sel, montrent une augmentation de l'expression de TGF-\beta1, de CTGF (Connective Tissue Growth Factor), du collagène de type 1 et de la fibronectine ; ce qui conduit à une augmentation du niveau des protéines urinaires et de l'albumine. Un traitement avec le polyamide a montré une réduction de ces niveaux, aussi bien qu'une diminution des ARNm associés à ces protéines, donc cela suggère que le traitement par le polyamide permet de moduler les dommages rénaux en inhibant le TGF-\u00df1 et ses fonctions en aval (Matsuda et al. 2011). Le même type d'étude a permis de vérifier l'inhibition de l'interaction du facteur de transcription NFkB dans des cellules humaines A549 (cellules du cancer du poumon) (Raskatov, Meier, et al. 2012). Tous ces travaux constituent des preuves de l'intérêt des polyamides comme inhibiteurs potentiels de la transcription des gènes.

Les polyamides ont aussi été étudiés pour leur capacité à être utilisés comme des agents antiviraux (Barrett et al. 2013). L'équipe de Sugiyama a synthétisé des polyamides standard de 8 monomères dirigés contre une séquence très conservée du virus Epstein-Barr (EBV) affectant la reconnaissance de la protéine virale EBNA1. La liaison de cette protéine sur sa séquence cible initie le recrutement de facteurs impliqués dans la réplication du virus à l'origine du processus infectieux (Yasuda et al. 2011). Le traitement de cellules latentes, infectées par le virus EBV, par un des polyamides synthétisés a montré un effet antiprolifératif des cellules positives pour la présence de l'épisome viral. De plus, des expériences de ChIP (ImmunoPrécipitation de la Chromatine) ont permis de confirmer que le polyamide agit en empêchant la liaison de la protéine virale EBNA1 avec sa séquence cible et donc que la réplication virale est affectée par le traitement. Un autre exemple de polyamide capable de diminuer la quantité d'épisomes viraux a été décrit pour le ciblage d'une séquence

riche en A et T présente dans le génome du papillome humain (HPV pour Human PapillomaVirus). Des examens histologiques de tissus épithéliaux, infectés par le virus HPV et testés avec le polyamide synthétisé, ont montré que le traitement par le polyamide permet une perte dose-dépendante d'épisomes viraux en inhibant la transcription (Edwards et al. 2011; Edwards et al. 2013). Ainsi, ces études antivirales et celles du groupe de Dervan sur le VIH (Dickinson et al. 1998; Dickinson, Trauger, Baird, Dervan, et al. 1999; Dickinson, Trauger, Baird, Ghazal, et al. 1999; Coull et al. 2002; Ehley et al. 2002; Dudouet et al. 2003) indiquent un rôle émergeant des polyamides qui peuvent être utilisés, soit comme des outils pour étudier la fonction de certains éléments capables de se lier à l'ADN et ayant trait à la réplication virale, soit directement comme des agents antiviraux.

D.4.2 L'extinction de gènes via l'alkylation de l'ADN par des conjugués de polyamides

Gottesfeld et Dervan avaient observé que les polyamides n'étaient pas capables d'inhiber l'expression d'un gène quand ils se lient directement à des régions codantes de l'ADN (Wurtz & Dervan 2000). Quand l'ARN polymérase commence la transcription, elle ne peut pas être arrêtée par un ligand non covalent de l'ADN. En utilisant les propriétés des polyamides, le groupe de Dervan a développé des outils capables d'effectuer des modifications covalentes sur les séquences d'ADN de façon séquence-spécifique afin d'inhiber la transcription. La création de ligands bifonctionnels en conjuguant des polyamides à des agents alkylants a ouvert la voie pour le développement d'agents alkylants ayant une activité séquence-spécifique pour éteindre sélectivement l'expression de certains gènes. Deux agents alkylants, le chloraminophène ou « chlorambucil » (Figure 23A) (Wurtz & Dervan 2000) et le CC-1065 (Chang & Dervan 2000), ont été couplés à des polyamides au niveau de l' α -amine de l'acide γ - diaminobutyrique du coude γ . Couplés aux polyamides, chacun des agents a alkylé de façon séquence-spécifique l'azote en position 3 (N-3) de l'adénine dans le petit sillon de l'ADN. Sugiyama a développé ensuite un conjugué polyamide-CycloPropylpyrroloIndole (CPI) (Figure 23A), basé sur la partie alkylante de la duocarmycine A, qui a aussi ciblé une adénine située à proximité du site de reconnaissance du polyamide dans le petit sillon de l'ADN (Shinohara et al. 2004). Les études suivantes ont permis de moduler la taille ou la composition du bras de liaison ainsi que sa position pour trouver le conjugué optimal permettant l'alkylation de séquences spécifiques.



Figure 23 : Les polyamides comme des agents alkylants

A/ Exemples de groupements alkylants couplés aux polyamides. B/ Modèle schématique de l'hétérotrimère alkylant pour cibler les régions télomériques humaines. Le losange blanc représente la β -alanine (Kashiwazaki et al. 2009).

Dans une autre étude, l'équipe de Sugiyama s'est intéressée à ces conjugués polyamides-agents alkylants afin de cibler les régions télomériques pour inhiber l'activité de la télomérase. Une suractivation de la télomérase permet d'empêcher le raccourcissement des télomères et la mort cellulaire dans les lignées cellulaires immortalisées et dans 85% des cancers ; aussi le contrôle d'une activité aberrante de la télomérase est une option intéressante pour le traitement de cancers. Un premier conjugué polyamide-CPI a été synthétisé et sa capacité à se lier à la séquence cible et à l'alkyler a été validée (Takahashi et al. 2003). Cependant les tests de croissance cellulaire ont révélé que l'activité du conjugué dépend de la lignée cellulaire utilisée. Un nouveau polyamide a été conjugué au seco-CBI (seco-1-Chlorométhyle-5-hydroxy-1,2-dihydro-3H-Benz[e]Indole) (**Figure 23A**) via un bras de liaison indole (Sasaki et al. 2006). Testé dans les mêmes lignées cellulaires que le précédent conjugué, ce polyamide-seco-CBI a montré une amélioration de l'activité anti-tumorale mais aussi une alkylation de sites sur des séquences non spécifiques.

À partir de ces résultats, l'amélioration de la spécificité de séquence a été réalisée par l'utilisation d'un hétérotrimère composé par deux molécules de distamycine A et d'un

conjugué linéaire polyamide-seco-CBI. Cet hétérotrimère se lie et alkyle la séquence cible d'ADN (Figure 23B) (Kashiwazaki et al. 2009). Cette preuve de concept d'utiliser un hétérotrimère comme agent alkylant a été reprise en couplant le chlorambucil en C-term ou Nterm d'un polyamide linéaire (Kashiwazaki et al. 2010). En utilisant la distamycine A (deux molécules) comme partenaire, il a été montré que la spécificité de séquence est meilleure avec le seco-CBI; cependant, la synthèse du chlorambucil est plus simple et son couplage est possible en N-term ou C-term du polyamide. La position de l'agent alkylant sur le polyamide est importante puisque le couplage en C-term permet d'alkyler le brin d'ADN qui interagit avec le polyamide linéaire alors que le couplage en N-term permet une alkylation du brin opposé. En tenant compte des résultats décrits précédemment, le groupe de Sugiyama a publié plusieurs études sur les polyamides couplés à des agents alkylants. Divers polyamides ou tandems de polyamides unis par des bras de liaison différents ont été couplés au seco-CBI et leur capacité à alkyler des séquences spécifiques a été vérifiée (Yoshidome et al. 2012; Taylor, Kawamoto, et al. 2014; Taylor et al. 2015). Ainsi un tandem alkylant spécifiquement les régions télomériques, de faible cytotoxicité, a été développé (Yamamoto, Bando, Kawamoto, et al. 2014). La synthèse d'un long polyamide couplé au chlorambucil a aussi permis l'alkylation de sites spécifiques et a empêché la progression de l'ARN polymérase ; induisant une inhibition de la transcription (Asamitsu et al. 2014).

Une des applications intéressantes des conjugués polyamides-agents alkylants est l'extinction sélective d'oncogènes en vue du traitement de cancers. À titre d'exemple, le criblage d'une librairie de polyamides conjugués au chlorambucil a permis d'isoler un conjugué capable d'altérer la morphologie de cellules humaines du cancer du côlon et de provoquer l'arrêt de ces cellules en phase G2/M du cycle cellulaire. Les analyses moléculaires ont révélé que les gènes de l'histone H4 étaient alkylés et donc réprimés, amenant à une déplétion de la protéine histone H4 qui induit une perte des nucléosomes (Dickinson et al. 2004; Alvarez et al. 2006; Chou, Farkas, et al. 2008). Testé sur des lignées cellulaires de la leucémie myéloïde chronique, ce même conjugué inhibait leur prolifération et en combinaison avec l'imatinib (inhibiteur de la kinase Abl) il s'est révélé comme un partenaire potentiel d'une bithérapie peu toxique dans des modèles de souris (Chou, O'Hare, et al. 2008). Enfin, compte tenu du rôle pléiotropique des GTPases, la mutation du codon 13 du gène KRAS, à l'origine de plusieurs cancers, en fait une cible idéale pour des molécules capables d'interagir avec des séquences spécifiques et de les alkyler. Couplé au seco-CBI par un lien indole, un long polyamide a permis l'alkylation spécifique de la séquence mutée du codon 13 du gène

KRAS (Taylor, Asamitsu, et al. 2014). Très récemment, le même type d'étude a été réalisée par cette équipe pour cibler la séquence mutée du codon 12 du gène KRAS (Taylor et al. 2015).

Une étude originale publiée par l'équipe de Sugiyama, décrit un conjugué polyamideseco-CBI photo-contrôlable (**Figure 24**) (Park et al. 2011).



Figure 24 : Représentation schématique d'un conjugué polyamide-seco-CBI photocontrôlable

À gauche, le conjugué polyamide-secoCBI est inactivé par la présence d'un groupement protecteur ortho-nitrobenzyle (PG) « photoclivable » sur le groupement hydroxyle de l'agent alkylant seco-CBI. Après une irradiation UV (hv), le groupement protecteur est coupé, laissant place à un agent alkylant actif (à droite). Ainsi, une régulation spatio-temporelle de ce conjugué est possible. La régulation spatiale est faite par le polyamide qui peut trouver sa séquence cible dans le petit sillon de l'ADN db et la régulation temporelle est effectuée par le groupement protecteur « photoclivable ». Les boules bleues et rouges représentent les monomères Im et Py respectivement, et l'ovale violet représente le bras de liaison entre le polyamide et le groupement alkylant seco-CBI (adapté de (Park et al. 2011)).

Au niveau de l'agent alkylant, le groupement protecteur, qui peut être enlevé par irradiation UV, a été développé pour empêcher l'activité d'alkylation du seco-CBI en protégeant le groupement hydroxyle naphtalénique. *In vitro*, après une irradiation UV (365 nm) de 5 minutes, le conjugué est capable d'alkyler sélectivement sa séquence cible d'ADN. Les études cellulaires (lignée humaine A549 du cancer du poumon) ont révélé que le traitement par le conjugué activable par « photoclivage » est toxique pour les cellules cancéreuses. Les traitements par le conjugué sans irradiation ou par l'irradiation seule ne montrent pas d'effets. Les résultats suggèrent que c'est l'alkylation spécifique par le polyamide conjugué activé qui est responsable du changement morphologique des cellules et de la toxicité sur les cellules cancéreuses. Ainsi, un nouveau type d'approche a été développé pour réguler de façon spatio-temporelle les gènes par les conjugués polyamides-agents alkylants.

D.4.3 L'activation transcriptionnelle par l'intermédiaire des polyamides

Les polyamides ont aussi été étudiés pour leur capacité à construire des facteurs de transcription artificiels. Les facteurs de transcription (FTs) sont des protéines modulaires composées généralement de deux domaines : un domaine de liaison à l'ADN (DBD - DNA Binding Domain) et un domaine effecteur qui permet la dimérisation et l'activation ou la répression de la transcription de régions spécifiques. Alors, le remplacement du DBD par un polyamide pourrait permettre la construction de FTs artificiels pouvant cibler des séquences d'ADN spécifiques et ainsi agir sur la régulation de gènes d'intérêt. Dans ce contexte, le premier activateur synthétique de la transcription, basé sur l'utilisation d'un polyamide comme DBD, a été conçu et synthétisé en couplant un polyamide de 8 monomères à deux domaines : l'un permettant la dimérisation (élément de dimérisation du facteur de levure GCN4 formant une structure de superhélice) et l'autre l'activation (peptide riche en glutamates structuré en hélice α) (Figure 25A) (Mapp et al. 2000). Cet activateur synthétique possède une affinité nanomolaire pour sa séquence cible et a permis une augmentation de la transcription des régions palindromiques ciblées dans des extraits nucléaires de levure. Des contrôles (le polyamide avec ou sans l'un des deux domaines ou encore les deux domaines sans le polyamide) ont permis de montrer les rôles indispensables du polyamide et du domaine d'activation ; en revanche, le domaine de dimérisation peut être modifié sans que l'activation de la transcription soit complètement empêchée.

En utilisant la séquence minimale permettant l'activation de VP16 (protéine activatrice virale), la construction d'un autre type d'activateur de la transcription a été développée (Ansari et al. 2001). Des ajustements sur la taille, la nature ou encore la position du bras de liaison entre le polyamide et le domaine d'activation, ont permis le développement d'un activateur de la transcription efficace, *in vitro*, sur une séquence en amont de la TATA box de l'adénovirus (Arora et al. 2002). Cependant, l'activité cellulaire de ces FTs artificiels était limitée par leur perméabilité cellulaire. Un nouveau type d'activateur transcriptionnel synthétique capable de pénétrer dans les cellules a été développé (Xiao et al. 2007). Un polyamide de 8 monomères a été couplé en C-term par un lien poly-éther à un hexapeptide activateur (capable de lier un coactivateur transcriptionnel de mammifère). Pour mesurer l'activation transcriptionnelle dans des cellules vivantes, des cellules Hela ont été transfectées transitoirement avec un plasmide portant le gène rapporteur de la luciférase sous le contrôle d'un promoteur ayant 6 sites de liaison pour le polyamide. Le traitement des cellules par le FT

artificiel a montré une augmentation dose-dépendante de l'expression de la luciférase. Les expériences contrôles indiquaient que le polyamide et le domaine d'activation étaient essentiels à l'activation de l'expression de la luciférase.





A/ Représentation schématique d'un FT artificiel composé d'un polyamide comme domaine de liaison à l'ADN (DBD), d'un domaine de dimérisation et d'un domaine d'activation. Les cercles noirs et vides représentent les monomères Im et Py, respectivement (adapté de (Mapp et al. 2000)). B/ Formule d'un FT artificiel composé d'un polyamide et de l'inhibiteur d'histone déacétylase (SAHA) (Ohtsuki et al. 2009).

Ensuite, l'équipe de Sugiyama a développé une nouvelle approche pour activer la transcription en utilisant un polyamide conjugué à l'acide SuberoylAnilide HydroxAmique (SAHA) qui est un inhibiteur d'histone déacétylase (Ohtsuki et al. 2009). Le but de cette étude était d'induire la transcription de gènes pluripotents requis pour la conversion de fibroblastes somatiques en cellules souches. Une librairie de plusieurs polyamides couplés au SAHA en N-term via un bras de liaison composé par deux monomères β -ala a été synthétisée. Le criblage de ces molécules a été fait en analysant leur capacité à induire l'expression de gènes connus pour leur rôle dans la reprogrammation cellulaire (exemples : Oct-3/4, Nanog, Sox2...) sans induire de toxicité cellulaire. Ce criblage a permis d'identifier un polyamide conjugué au SAHA (**Figure 25B**) qui, par des expériences de ChIP, conduit à une augmentation de l'acétylation des histones dans les promoteurs des gènes Oct-3/4 et Nanog.

De plus, une augmentation du niveau de triméthylation de l'histone H3K4, qui est une marque associée à une chromatine active, a été observée (Pandian, Nakano, et al. 2012). D'autres études du même groupe ont permis d'identifier d'autres polyamides conjugués au SAHA ayant des effets similaires sur d'autres gènes pluripotents (Kashiwazaki et al. 2012; Pandian, Ohtsuki, et al. 2012; L. Han et al. 2013; Saha et al. 2013; Saha et al. 2014; Anandhakumar et al. 2014).

Récemment, pour empêcher la répression de la transcription, des polyamides ont été synthétisés pour cibler une séquence riche en îlots CpG afin d'inhiber leur méthylation par les méthyltransférases (Kang et al. 2014). Une autre approche « compétitive » plus ancienne pour activer la transcription a été décrite. L'interaction d'un polyamide avec un site spécifique d'ADN peut empêcher la liaison de protéines sur ce même site. Ainsi, plusieurs polyamides conçus pour se lier sur le site de liaison de la protéine IE86 (protéine du CytoMégaloVirus CMV) ont été synthétisés. IE86 est une protéine connue pour son activité de répresseur de la transcription d'une séquence promotrice du CMV. L'interaction du polyamide a permis d'empêcher la fixation du répresseur de la transcription IE86 et donc d'activer la transcription de séquences spécifiques (Dickinson, Trauger, Baird, Ghazal, et al. 1999).

D.5 Les accessibilités cellulaire et nucléaire des polyamides

En solution, les polyamides montrent une bonne affinité et une bonne spécificité pour leur séquence cible d'ADN. Cependant, dans des études cellulaires ou animales, ces caractéristiques des MGBs vont dépendre de nombreux facteurs dont notamment de leurs pénétrations cellulaire et nucléaire. Plusieurs études ont montré que la localisation nucléaire dépend du type cellulaire, de la structure du polyamide, de l'identité du fluorophore conjugué ou encore du nombre de monomères imidazoles présents dans le polyamide (Belitsky, Leslie, et al. 2002; Best et al. 2003; Crowley et al. 2003; Edelson et al. 2004; Nickols et al. 2007; Liu & Kodadek 2009; Nishijima et al. 2010; Raskatov et al. 2014). Par exemple, des conjugués polyamide-BODIPY ont été capables de se localiser dans le noyau de cellules humaines vivantes (cellules T primaires humaines) (Belitsky, Leslie, et al. 2002). Cependant, dans d'autres types cellulaires, les conjugués étaient exclus des noyaux et séquestrés dans des lysosomes et d'autres vésicules cytoplasmiques. Ainsi, le fait que la localisation nucléaire varie avec la structure du polyamide conjugué, suggère que la conception du polyamide et le choix du fluorophore utilisé doivent être optimisés.

Chapitre I – Introduction Générale Les polyamides ou MGBs

Les facteurs structuraux pouvant être impliqués dans l'adressage nucléaire des polyamides ont été étudiés par l'équipe de Dervan en synthétisant une série de MGB couplés à la fluorescéine et en testant leur capacité à se localiser dans les noyaux d'un large panel de cellules vivantes de mammifères, incluant 11 lignées de cellules cancéreuses humaines, des cellules transformées de foie humaines et une lignée de cellules leucémiques murines (Best et al. 2003). D'après les résultats obtenus, certains facteurs semblent impliqués dans la localisation des conjugués. Généralement, les polyamides conjugués qui se localisaient dans le noyau étaient composés de 8 monomères et portaient une charge positive, soit sur le bras de liaison entre le polyamide et le fluorophore, soit sur le coude γ . Une année plus tard, une étude similaire a décrit la synthèse d'une librairie de 100 polyamides, conjugués à différents types de fluorophores, et testés pour leur capacité à se localiser dans les noyaux de 13 lignées cellulaires de mammifères (Edelson et al. 2004). Les conjugués ont été conçus avec des variations structurales afin de pouvoir déduire systématiquement les effets de ces modifications sur la localisation. Les paramètres étudiés étaient entre autre : l'influence du poids moléculaire (la taille du polyamide conjugué), la composition en Im et en β -ala, l'ordre des monomères, la charge, le type et la position du fluorophore. Alors que le poids moléculaire et la composition en Im semblaient ne pas influencer la localisation nucléaire des polyamides conjugués, la position des Im avait des conséquences sur l'accès aux noyaux ; cependant, aucune règle générale sur la position des imidazoles n'a pu être déduite de cette étude. Toujours d'après les résultats de cette étude, la présence d'une β -ala au niveau de la queue C-term du polyamide et la nature du bras de liaison entre le fluorophore et le polyamide affectaient la localisation nucléaire; alors que la présence d'un groupement acide 2,4diaminobutyirique acétylé au niveau du coude pouvait augmenter la pénétration dans les noyaux.

En utilisant dans un test cellulaire « rapporteur », des polyamides conjugués à la dexaméthasone (dex) agissant comme des activateurs de la transcription, l'équipe de Kodadek a voulu déterminer les effets de la taille et de la composition du bras de liaison sur la pénétration cellulaire (Liu & Kodadek 2009). Le test rapporteur consistait à traiter, par les polyamides conjugués, des cellules exprimant une protéine de fusion contenant le domaine de liaison à l'ADN (DBD) de Gal4, le domaine de liaison au ligand (ici le ligand est la dex) du récepteur au glucocorticoïde humain (GRLBD) et VP16, qui est le domaine d'activation de l'activateur transcriptionnel du virus de l'herpès (**Figure 26**).



Figure 26 : Schéma du test cellulaire « rapporteur » avec un polyamide conjugué à la dexaméthasone (adapté de (Liu & Kodadek 2009))

Dans sa forme apo (non liée au ligand), la protéine de fusion est séquestrée dans le cytoplasme via son interaction avec la protéine chaperonne Hsp90 (Heat shock protein 90 Kd). L'interaction du domaine GRLBD de la protéine de fusion avec la dexaméthasone couplée aux polyamides permet la libération de Hsp90 et une translocation du complexe dans le noyau, où il active la transcription du gène rapporteur luciférase placé sous le contrôle d'un promoteur Gal4 (**Figure 26**). L'activité dose-dépendante de la luciférase a permis de confirmer certains résultats trouvés précédemment par l'équipe de Dervan. La présence de monomères β -ala pourrait affecter la perméabilité cellulaire alors que le recours à un bras de liaison plus court (éthylènediamine) l'améliore.

Dans le même but, Jacobs et Dervan ont rapporté la synthèse d'une librairie de polyamides avec diverses modifications en C-term pour évaluer leurs effets sur la perméabilité cellulaire (Jacobs & Dervan 2009). L'utilisation d'un bras de liaison oxime entre le polyamide et un groupe aromatique fonctionnel en C-term, a permis d'augmenter le ciblage de la séquence cible d'ADN (élément de réponse au récepteur de l'androgène) dans des cellules humaines LNCap (adénocarcinome épithélial de la prostate). Après avoir exploré les modifications en C-term, les modifications sur le coude γ ont été étudiées pour leurs effets sur la pénétration cellulaire, la biodistribution et l'éventuelle toxicité chez la souris (Meier et al. 2012; Synold et al. 2012; Yang et al. 2013). L'addition de groupements aryle sur le coude γ a permis d'améliorer la localisation nucléaire des polyamides.

Enfin, la solubilité des molécules peut être une barrière aux études sur animal. Ainsi, de récentes études se sont intéressées à améliorer la solubilité des polyamides en solution aqueuse. À titre d'exemple, l'addition de groupements amine ou ammonium sur les monomères pyrroles a été réalisée et a pu permettre une augmentation de la solubilité du polyamide dans l'eau (Babu et al. 2011; Satam, Babu, Chavda, et al. 2012; Satam, Babu, Porte, et al. 2012; Ramos et al. 2013). Une autre stratégie développée par l'équipe de Sugiyama était d'ajouter un groupement polyéthylène glycol en C-term d'un conjugué polyamide-seco-CBI. Une cytotoxicité accrue et une augmentation de la solubilité ont été observées avec ce conjugué (Takagaki et al. 2011). Plutôt que de modifier la structure du polyamide, le groupe de Dervan a augmenté la solubilité en solution aqueuse en utilisant une formulation contenant de la 2-Hydroxypropyl- β -cyclodextrine (Hp- β -cyclodextrine) qui est un carbohydrate utilisé comme agent solubilisant (Hargrove et al. 2012). Chez des souris, en utilisant des polyamides solubilisés par de la Hp- β -cyclodextrine, les conditions de l'injection et la pénétration cellulaire ont été améliorées, probablement par une diminution de la formation d'agrégats dans le plasma.

Les propriétés des polyamides ou MGBs évoquées dans cette introduction en font d'excellents candidats pour la construction de sondes fluorescentes afin de cibler des séquences d'ADN spécifiques. En effet, contrairement aux oligonucléotides modifiés, ils ne sont pas sensibles au pH et ne nécessitent pas une dénaturation de l'ADN pour leur fixation puisqu'ils sont capables d'interagir dans le petit sillon de l'ADN db avec leur séquence cible. De plus, à partir de la séquence cible d'ADN, on peut programmer l'enchainement des monomères afin de synthétiser un polyamide spécifique de cette séquence donc la synthèse est simple et ne requiert pas de constructions génétiques complexes comme celles utilisées pour la conception de protéines modifiées. Pour finir, ils sont capables de pénétrer dans les cellules et de trouver leur séquence cible d'ADN, sans recours à des techniques de transfection pouvant perturber le fonctionnement et la morphologie des cellules. Les polyamides fluorescents répondent bien aux exigences requises à la conception de sondes fluorescentes pour visualiser des séquences d'ADN db.

<u>E – Le projet de thèse</u>

L'organisation nucléaire de l'ADN est un facteur important dans le contrôle de l'expression génique. La possibilité de visualiser des séquences spécifiques d'ADN dans des cellules vivantes permettrait d'observer, en temps réel, la dynamique de l'ADN dans le noyau en utilisant la microscopie de fluorescence. Le but de ce projet était le développement d'outils chimiques qui permettraient d'étudier l'organisation et la dynamique de séquences cibles dans des cellules vivantes, sans les fixer ni les perméabiliser.

En utilisant comme modèles des séquences d'ADN répétées centromériques, murine et humaine, le projet de thèse s'est organisé autour de trois questions principales : Chez la souris, est-il possible de visualiser les chromocentres directement dans des cellules vivantes en utilisant des sondes fluorescentes composées par des polyamides ? Et chez l'homme, est-il possible de visualiser les séquences α -satellite ? Et existe-t-il des structures apparentées aux chromocentres murins dans les cellules vivantes?

Tout d'abord, la synthèse de pyrrole-imidazole polyamides appelés "ligands du petit sillon de l'ADN" ou "MGBs" (Minor Groove Binders) qui sont capables de reconnaître et d'interagir avec des séquences spécifiques d'ADN au sein du petit sillon a été réalisée dans le but de cibler des séquences répétées d'ADN centromérique murin et humain. Concernant la pénétration cellulaire des MGBs, beaucoup de modifications structurales ont été testées mais aucune règle précise ne semble en découler. Par conséquent, la synthèse des polyamides a été entreprise selon la méthode décrite par Krutzik et Chamberlin (Krutzik & Chamberlin 2002a; Krutzik & Chamberlin 2002b), en suivant les règles de reconnaissances établies par Dervan. Ensuite, ces polyamides ont été purifiés et caractérisés.

Puis avant et après leur couplage à des fluorophores, l'interaction de ces polyamides avec des duplexes d'ADN synthétiques a été étudiée, *in vitro*, par des méthodes physicochimiques : dénaturation thermique de l'ADN, retard sur gel d'acrylamide, dichroïsme circulaire et spectroscopie de fluorescence.

Enfin, pour les deux modèles, murin et humain, l'interaction de ces sondes fluorescentes a été étudiée, *in cellulo*, sur des cellules fixées et vivantes.

Chapitre II - Résultats

Chapitre II - Résultats

Chapitre II - Résultats	1
A - Les résultats obtenus avec le modèle murin	1
A.1 Synthèses et caractérisations des MGBs et des sondes fluorescentes murins)
A.2 Études physico-chimiques de l'interaction des MGBs murins et des sondes	
fluorescentes murines avec l'ADN in vitro	5
A.3 Études de l'interaction des sondes fluorescentes murines avec l'ADN in cellulo 119)
A.4 Conclusions générales des résultats obtenus avec le modèle murin 128	;
B - Les résultats obtenus avec le modèle humain)
B.1 Synthèses et caractérisations des MGBs et des sondes fluorescentes humains 130)
B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescentes	
humains avec l'ADN in vitro	5
B.3 Études de l'interaction des sondes fluorescentes humaines avec l'ADN in cellulo 146)
B.4 Conclusions générales des résultats obtenus avec le modèle humain)

A - Les résultats obtenus avec le modèle murin

Les résultats obtenus avec le modèle murin ont fait l'objet de deux publications situées à la fin de cette partie. Un des article est focalisé sur les études physico-chimiques des polyamides et sondes murins (Nozeret, Bonan, et al. 2015), alors que l'autre inclut les résultats cellulaires (Nozeret, Loll, et al. 2015).

L'observation de séquences spécifiques d'ADN repose sur l'utilisation de sondes fluorescentes. Les séquences répétées de l'ADN génomique qui se situent au niveau de loci chromosomiques spécifiques (centromères, télomères) constituent des cibles idéales pour les sondes fluorescentes. Leur concentration locale est élevée, ce qui permet une accumulation du signal fluorescent au niveau d'une même région. Dans cette étude, le choix de la séquence cible s'est tourné vers la répétition péricentromérique murine MajSat de 234 pb répétée des milliers de fois. La séquence consensus de ces répétitions a été déterminée par la méthode de Maxam et Gilbert (Manuelidis 1981).

Les sondes doivent être capables d'interagir avec l'ADN de façon séquence-spécifique et d'émettre un signal de fluorescence. Dans notre cas, les sondes fluorescentes ont été construites en conjuguant des fluorophores avec des polyamides pouvant interagir avec l'ADN db.

Pour la partie fluorescence, un des critères majeurs pour l'imagerie de séquences d'ADN spécifiques dans des cellules vivantes est la possibilité de détecter un signal spécifique produit par les molécules fluorescentes liées à leur cible d'ADN et de pouvoir distinguer ce signal du bruit de fond de fluorescence venant de molécules non-liées. Pour contourner ce problème, plusieurs approches sont possibles. La première est la construction de sondes fluorescentes en utilisant des fluorophores qui peuvent moduler leur spectre de fluorescence en fonction de leur interaction avec l'ADN, en augmentant l'intensité de fluorescence (« light-up probes ») et / ou en décalant leur spectre vers une autre longueur d'onde (« red or blue shift »). D'après la littérature les fluorophores de type cyanine sont de bons candidats (Yarmoluk et al. 1996; Hilal & Taylor 2008). L'autre approche consiste en l'utilisation de deux fluorophores (donneur et accepteur) situés à proximité l'un de l'autre pour observer un effet FRET (Transfert par Résonance d'Energie de Fluorescence) (Su et al. 2013). Suite à l'excitation du fluorophore donneur, il y a un transfert d'énergie vers le fluorophore accepteur qui, excité, émet un signal de fluorescence à une longueur d'onde différente qui est donc distinguable (Boutorine et al. 2013).

Pour la partie polyamide, le critère important pour permettre de visualiser des séquences d'ADN dans des cellules vivantes est l'augmentation de la spécificité de reconnaissance ; c'est-à-dire de la longueur de la séquence reconnue. Deux stratégies ont été suivies. La première était de synthétiser deux sondes à base de polyamides courts ou polyamides standard (composés par 4 paires de monomères) situés à proximité l'un de l'autre sur la séquence cible d'ADN afin de pouvoir observer un effet FRET. La deuxième stratégie était de construire un tandem covalent de polyamides, dans une orientation « tête à tête », en utilisant deux polyamides fonctionnalisés par des groupements alcyne et azoture permettant leur conjugaison par une réaction de « chimie-click », la cycloaddition-1,3 dipolaire de Huisgen.

A.1 Synthèses et caractérisations des MGBs et des sondes

fluorescentes murins

A.1.1 Les MGBs murins

A.1.1.1 Les MGBs murins standard

Le *Duplexe 1* de 30 pb, issu de la répétition murine MajSat (Manuelidis 1981), a été utilisé comme cible modèle pour la synthèse des polyamides et pour les études physicochimiques (**Figure 27**).



Figure 27 : Représentation schématique des différents duplexes utilisés et des polyamides murins synthétisés

Le **Duplexe 1** est le duplexe cible modèle de 30 pb issu de la répétition péricentromérique murine MajSat. Il possède plusieurs sites d'interaction pour chaque MGB (les sites de liaison pour le polyamide **F1** sont encadrés). Les polyamides synthétisés pour cibler cette séquence sont alignés relativement à leurs sites de liaison sur le **Duplexe 1**. Le **Duplexe 2** contient uniquement un site de liaison pour chacun des polyamides (le site de liaison pour le polyamide **F1** est encadré) alors que le **Duplexe 3** n'en contient aucun ; il s'agit du duplexe contrôle (adapté de (Nozeret, Bonan, et al. 2015)).

Chapitre II – Résultats Les résultats obtenus avec le modèle murin

Le polyamide F1 a été construit de son extrémité N-term à son extrémité C-term pour reconnaître le duplexe cible dans le sens $5 \rightarrow 3$ ' du brin riche en A. Dans le but de sélectionner la paire de polyamides optimale pour les expériences de FRET et pour construire un tandem antiparallèle (« tête à tête »), la synthèse de plusieurs versions du deuxième polyamide (appelées F2-F6) capable de reconnaitre le brin riche en T dans le sens $5 \rightarrow 3$ ' a été entreprise. Ces polyamides ciblent des séquences voisines avec un décalage de 1 à 4 paires de bases (Figure 27). Cette orientation antiparallèle des paires de polyamides, sur le duplexe cible, procure une proximité spatiale des deux extrémités N-term des polyamides après la fonctionnalisation de leurs groupements amines terminaux par des modifications chimiques (Figure 28). Un autre avantage de cette proximité est la possibilité, après le marquage des polyamides par des fluorophores, de construire des paires de sondes fluorescentes pour pouvoir étudier un effet FRET potentiel en raison des distances proches entre les fluorophores donneur et accepteur.

La synthèse de ces polyamides a été effectuée manuellement sur support solide en utilisant une résine Boc-B-ala-PAM selon le protocole décrit par Krutzik et Chamberlin (Krutzik & Chamberlin 2002a; Krutzik & Chamberlin 2002b). Les détails de la synthèse sont donnés dans la Figure 28 et dans le Chapitre IV (Matériels et méthodes). Le suivi, étape par étape, de la synthèse des polyamides sur support solide par des tests à la ninhydrine, qui réagit avec les amines primaires en produisant une coloration pourpre de Ruhemann, a mis en évidence une difficulté du couplage d'un Py directement à la suite d'un Im (Baird & Dervan 1996). Pour contrer cette difficulté, deux stratégies ont été suivies. La première est de synthétiser un dimère Py-Im en phase liquide (Baird & Dervan 1996). En optant pour cette stratégie, une optimisation du protocole décrit a été faite en précipitant le produit à l'acide chlorhydrique avec un suivi du pH, en doublant le volume nécessaire au couplage du dimère sur le support solide et en augmentant le temps de couplage à 2h. La deuxième stratégie consiste à remplacer le Py, qui doit suivre un Im, par un synthon β -ala. De plus, le remplacement d'un Py par un synthon β-ala peut améliorer l'affinité d'un polyamide quand celui-ci est dirigé vers des séquences riches en G et C (Turner et al. 1998; Dervan & Bürli 1999; Floreancig et al. 2000). Ainsi, les analogues des polyamides F2, F3 et F4 nommés $F2\beta$, $F3\beta$ et $F4\beta$ respectivement ont été synthétisés. Les formules et caractéristiques chimiques des polyamides synthétisés sont données en Annexe 1. Après leur synthèse, les polyamides ont été analysés par HPLC analytique en phase inverse (Chromatographie en phase Liquide à Haute Performance), caractérisés par spectrométrie de masse (SM) et spectroscopie UV-visible et enfin, purifiés par HPLC préparative en phase inverse. Les chromatogrammes obtenus révèlent que les produits ont une pureté de 80 à 90% à l'issue des synthèses. L'exemple de la caractérisation du MGB *F4* est présenté dans l'Annexe 2.



Figure 28 : Schéma de la synthèse peptidique manuelle des polyamides sur support solide

La synthèse des polyamides s'effectue de leur extrémité C-term à leur extrémité N-term. La première étape consiste à déprotéger la résine Boc- β -ala-PAM en utilisant une solution TPW (TFA 92,5% / phénol 5% / eau 2,5%), ce qui permet de libérer une fonction amine. Les synthons commerciaux portent une fonction acide carboxylique (encadré de droite) et en utilisant des activateurs, tels que le HATU (2-(1H)-7-azabenzotriazol-1-yl)-1, 1, 3, 3-tétraméthyluronium-hexafluorophosphate) (en rose sur le schéma), il y a formation d'un synthon sous forme d'ester activé. En présence de l'amine libre de la résine, l'ester activé réagit pour former une liaison amide qui couple le premier synthon sur la résine. Ensuite, il faut répéter les étapes de déprotection, d'activation des synthons suivants et de couplage pour enchainer le nombre de synthons nécessaires à l'obtention du polyamide désiré. Une fonctionnalisation en N-term des polyamides est possible, directement sur la phase solide, en utilisant des « synthons N-terminaux » (encadré de droite). Pour finir, une étape d'aminolyse par la DMPA (3-(diméthylamino)-1-propylamine) permet de libérer le polyamide de la résine. Dans l'encadré de droite, les groupements protecteurs Boc sont représentés en noir.

A.1.1.2 Les tandems de MGBs murins

Au sein de notre laboratoire, la synthèse de tandems de polyamides avait été effectuée en utilisant différents bras de liaison (Halby et al. 2005; Halby et al. 2007). Les résultats de l'étude de leur interaction avec leur séquence cible d'ADN avaient montré qu'ils possédaient de meilleures affinités et spécificités puisque la séquence reconnue par les tandems était plus longue. Dans ce travail, la synthèse des tandems a été améliorée et simplifiée en utilisant une réaction de « chimie-click » : la cycloaddition 1,3-dipolaire entre un alcyne et un azoture catalysée par le cuivre (ou réaction de Huisgen) (Meldal & Tornøe 2008; El-Sagheer & Brown 2012) (voir Chapitre IV – Matériels et méthodes). Des paires de polyamides modifiés ont été utilisées pour construire ces tandems : soit *F1* et *F2β* modifiés, soit *F1* et *F4β* modifiés (Figure 29A et Annexe 3).



Figure 29 : Schéma de la synthèse et structures des tandems de polyamides

A/ Schéma de la synthèse des tandems de polyamides par conjugaison de deux polyamides (en bleu) fonctionnalisés en N-term par des groupements alcyne (en rose) ou azoture (en marron). La réaction de Huisgen, en présence de cuivre pendant 20 min à 90°C dans un mélange eau 75% / DMSO 25% dans un réacteur à micro-ondes, permet la formation d'un cycle triazole (en vert) qui conjugue les deux polyamides. B/ Structure du tandem **T1** composé par les polyamides **K1** et **K2**β. C/ Structure du tandem **T2** composé par les polyamides **K1** et **K4**β. Le groupement protecteur Boc (en rouge) protège la fonction amine des tandems qui peut être utilisée pour le couplage de groupements chimiques ou de fluorophores.

Les groupements azoture et alcyne ont été ajoutés lors des étapes finales de la synthèse des polyamides, directement sur le support solide. Pour la première paire de polyamides (tandem *T1*), l'acide 2-(Boc-amino)-5-azido-pentanoïque a été attaché à la place de l'acide γ aminobutyrique N-terminal pour obtenir un analogue de F1 qui contient un groupement fonctionnel azoture, appelé K1. L'acide 4,5-pentynoïque a été attaché à l'acide yaminobutyrique N-terminal de $F2\beta$ afin d'obtenir son analogue nommé $K2\beta$ qui contient un groupement fonctionnel alcyne (Figure 29B et Annexe 3). Pour la construction du deuxième tandem T2, le polyamide K1y porte l'acide 2-(Boc-amino)-5-azido-pentanoïque attaché à l'acide γ -aminobutyrique et non à la place de celui-ci, et le polyamide **K4** β est l'analogue de $F4\beta$ avec le groupement additionnel acide 4,5-pentynoïque en N-term (Figure 29C et Annexe 3). Donc le tandem T2 possède un bras de liaison plus long de 5 liaisons entre les deux polyamides afin d'adapter sa structure à la distance plus longue entre les deux polyamides sur la séquence cible d'ADN. La réaction de « chimie-click » entre les deux composants de chaque tandem a été effectuée à 90°C dans un mélange eau 75% / DMSO 25% dans un réacteur à micro-ondes pendant 20 min. Le tandem T2 reste soluble dans le milieu réactionnel et après une précipitation à l'éther, il apparait comme un pic relativement homogène en HPLC analytique (Annexe 4). Dans le cas du tandem T1, un précipité, contenant le tandem presque pur, se forme dans le milieu réactionnel. Les deux produits ont été ensuite purifiés par HPLC préparative et caractérisés par spectrométrie de masse et spectroscopie UV-Visible (Annexes 3 et 4).

A.1.2 Les sondes fluorescentes murines

La construction des sondes fluorescentes murines a été effectuée en couplant aux polyamides murins divers fluorophores (commerciaux ou synthétisés soit directement au sein de notre laboratoire soit par des collaborateurs). Le choix des fluorophores (**Annexe 5**) s'est tourné vers des molécules qui ont la capacité de moduler leur spectre de fluorescence (décalage de la longueur d'onde ou augmentation de l'intensité de fluorescence) en réponse à une interaction avec l'ADN comme la tétraméthylrhodamine (TMR) (Rucker et al. 2003; Rucker et al. 2004), le thiazole orange (TO) (Fechter et al. 2005), la fluorescéine (Hsu & Dervan 2008; Nozeret, Loll, et al. 2015) ou encore les cyanines Cy3 et Cy5 (Foister 2003; Hilal & Taylor 2008). De plus, les cyanines Cy3 et Cy5 peuvent être utilisées comme paire de fluorophores pour des expériences de FRET (Foister 2003; Han et al. 2014).

Chapitre II – Résultats Les résultats obtenus avec le modèle murin

La plupart des fluorophores utilisés possèdent un groupement acide carboxylique et ils sont soit achetés directement sous forme d'esters activés (NHS – ester N-HydroxySuccinimidyle) ou d'isothiocyanate (ITC), soit transformés en esters activés par des agents tels que le HATU ou fonctionnalisés en isothiocyanate en greffant une chaine diamine *tert*-butyl-N-(3-aminopropyl)carbamate, qui après déprotection par du TFA 10 %, libère une fonction amine transformée par la suite en ITC (Munch et al. 2008). Les fluorophores ont été liés à la fonction amine libre présente à l'extrémité N-term des MGBs standard (**Figure 30**) ou à l'amine des synthons acide amino-pentanoïque des tandems après leur déprotection (**Figure 31**).



Figure 30 : Méthodes de couplages des fluorophores aux MGBs

A/ En utilisant la fonction amine libre du polyamide (NH₂ en N-term, exemple du MGB F4 β , en bleu) et des fluorophores sous forme d'esters activés (exemple de l'ester N-succinimidyle de la 5-carboxyfluorescéine, en rouge, à gauche), le couplage permet la formation d'une liaison amide (entourée en vert, à gauche) entre le fluorophore et le polyamide. B/ La fonction amine libre du MGB (NH₂ en N-term, exemple du MGB F4 β , en bleu) peut aussi réagir avec un fluorophore sous forme d'isothiocyanate (exemple de la fluorescéine-5isothiocyanate, en rouge, à droite) pour former une liaison thiourée (entourée en vert, à droite) entre le fluorophore et le polyamide.

En utilisant d'une part l'amine terminale des MGBs et d'autre part des fluorophores sous forme d'esters activés, la formation d'une liaison amide permet la construction d'une sonde fluorescente polyamide-fluorophore. L'exemple de la sonde $F4\beta$ -fluorescéine est donné dans la **Figure 30A**. En ce qui concerne les fluorophores sous forme d'isothiocyanate, une liaison thiourée conjugue le polyamide au fluorophore (l'exemple de la sonde *F4β-FITC* est présenté dans la **Figure 30B**). Il est important de noter que le couplage aux MGBs ne modifie pas les spectres de fluorescence des fluorophores. Un tableau récapitulant les différentes sondes synthétisées à partir des MGBs murins standard et leurs caractéristiques chimiques est donné en **Annexe 6**. À titre d'exemple, la caractérisation de la sonde *F4β-FITC* est donnée en **Annexe 7**. Enfin, les caractéristiques chimiques des deux tandems fluorescents *T1-Cy3* et *T2-FITC* (**Figure 31**) sont présentées en **Annexe 8**.



Figure 31 : Formules des tandems fluorescents murins

A/ Formule du tandem **T1** (en bleu) marqué par le fluorophore Cy3 (en rouge). Après sa déprotection, l'amine présente sur le synthon acide amino-pentanoïque (en marron) réagit avec le fluorophore sous forme d'ester activé pour former une liaison amide (entourée en vert foncé) pour coupler le fluorophore au MGB afin de construire une sonde fluorescente **T1-Cy3**. B/ Formule du tandem **T2** (en bleu) conjugué à la fluorescéine isothiocyanate (FITC, en rouge) par un une liaison thiourée (entouré en vert foncé) afin d'obtenir la sonde **T2-FITC**. Les synthons β-ala sont représentés en noir.

A.2 Études physico-chimiques de l'interaction des MGBs murins et des sondes fluorescentes murines avec l'ADN *in vitro*

Le *Duplexe 1* possède plusieurs sites de liaison pour chacun des polyamides standard et un site de liaison pour les tandems de polyamides (**Figure 27**). Par exemple, le polyamide *F1* peut se lier à 4 sites (2 étant partiellement superposés, les sites sont encadrés dans la **Figure 27**). Dans le but de déterminer les propriétés de fixation de chaque polyamide à sa séquence cible d'ADN, deux autres duplexes ont été utilisés. Le *Duplexe 2* qui n'a qu'un seul site de liaison pour chacun des polyamides et le *Duplexe 3* utilisé comme contrôle qui n'a aucun site de liaison pour les polyamides (**Figure 27**). L'interaction des polyamides et des sondes fluorescentes avec ces différents duplexes a été étudiée par plusieurs méthodes physico-chimiques (dénaturation thermique de l'ADN, gel-retard, dichroïsme circulaire et spectroscopie de fluorescence) afin de sélectionner les candidats avec les meilleures affinités et spécificités pour leur séquences cibles d'ADN en vue des applications cellulaires.

A.2.1 Cinétique de l'interaction ADN / polyamide

La cinétique de formation des complexes entre l'ADN db et les polyamides est une considération importante. À titre d'exemple, la formation de triples hélices est un processus lent qui peut prendre plusieurs heures (Novopashina et al. 2005). Ainsi, avant d'étudier l'interaction ADN / polyamides, il est nécessaire de déterminer la cinétique de fixation des MGBs à leur séquence cible d'ADN afin d'adapter les expériences physico-chimiques à la cinétique de formation des complexes. Dans le but d'estimer le taux de formation du complexe entre le Duplexe 2 (ayant un seul site d'interaction pour chaque MGB) et le MGB $F4\beta$ (Figure 32), la méthode « d'arrêt de retard sur gel » (« stop-in gel retardation ») a été appliquée (Novopashina et al. 2005) (voir Chapitre IV – Matériels et méthodes). Cette méthode repose sur la technique de l'électrophorèse qui est fondée sur le déplacement de molécules chargées (positivement ou négativement) sous l'effet d'un champ électrique. Du fait de leurs caractéristiques propres et en fonction des conditions de l'électrophorèse ces molécules auront des vitesses de migration différentes. Elles vont donc se séparer les unes des autres. Le retard sur gel ou EMSA (Electrophoretic Mobility Shift Assay) débute par l'incubation d'un oligonucléotide marqué (par fluorescence ou radioactivité) avec une molécule supposée interagir avec lui (comme des protéines ou même des extraits de protéines cellulaires). Ensuite, la migration du fragment nucléotidique étudié (seul) et celle du fragment mis en contact avec la molécule susceptible de se fixer dessus sont effectuées sur deux pistes différentes. Alors, si la molécule se fixe effectivement sur le fragment, la migration de ce dernier sera plus lente (retard), et la bande correspondante aura migré moins loin sur le gel (elle est dite retardée). Sinon, les deux bandes seront au même niveau sur les deux pistes.



Figure 32 : Image de l'expérience « d'arrêt de retard sur gel » pour étudier la cinétique de formation d'un complexe ADN / MGB

Électrophorèse sur un gel de polyacrylamide à 20% non dénaturant, dans un tampon TBE standard (89 mM Tris-HCl, 89 mM acide borique, 2 mM EDTA, pH 8,3) à 4°C, des mélanges entre le **Duplexe 2** marqué à la fluorescéine (0,52 μ M) et le polyamide **F4** β (1 μ M). 10 μ L de chaque mélange ont été chargés sur le gel à t₀, t₀+10, t₀+30, t₀+60 et t₀+120 min d'incubation, comme indiqué au-dessus du gel et immédiatement soumis à l'électrophorèse. to correspond au moment où le **Duplexe 2** a été mis en présence de **F4** β et que les premiers échantillons ont été chargés sur le gel. À gauche, les deux traits noirs représentent la position du duplexe d'ADN et le MGB est symbolisé par un point noir. a : les échantillons contrôles qui contiennent uniquement le duplexe d'ADN et b : les mélanges ADN/MGB.

Sur la **Figure 32**, les pistes « a » contiennent le *Duplexe 2* seul et les pistes « b », le *Duplexe 2* en présence du MGB $F4\beta$ à différents temps d'incubation : t₀, t₀+10, t₀+30, t₀+60 et t₀+120 min. Le temps t₀ correspond au moment où le *Duplexe 2* est mis en présence de $F4\beta$ et que les premiers échantillons ont été chargés sur le gel. L'apparition d'une bande retardée, dès le temps t₀, signifie que le complexe *Duplexe 2* : $F4\beta$ (1 : 2) est complètement formé presque immédiatement après l'ajout du polyamide (le temps de chargement était inférieur à 1 minute). Ainsi, la cinétique de fixation du polyamide ne représente pas une étape critique à la formation des complexes ADN / MGB. C'est un avantage supplémentaire des polyamides comparés aux TFOs.

A.2.2 Dénaturation thermique de l'ADN suivie par spectroscopie UV

Après l'étude de la cinétique de formation des complexes ADN / MGB, la question de la spécificité a été étudiée par la technique de dénaturation thermique de l'ADN (ou fusion de l'ADN) qui consiste à dissocier un double brin d'ADN en deux simples brins, en rompant les liaisons hydrogène entre les bases azotées des deux brins complémentaires de l'ADN, sous l'effet d'une augmentation de la température. Contrairement à la dénaturation des protéines, celle de l'ADN est généralement réversible à condition que le retour aux conditions initiales soit suffisamment lent. Les bases se ré-apparient, c'est la renaturation. La dénaturation / renaturation thermique des acides nucléiques peut être facilement suivie par spectroscopie UV à une longueur d'onde fixée (classiquement 260 nm pour les acides nucléiques) grâce à l'existence d'un effet hyperchrome (augmentation de l'absorption dans l'UV pendant la dissociation). En effet, le coefficient d'extinction molaire de la molécule augmente au cours de la transition double brin \rightarrow simple brin. Ceci est dû au changement d'environnement des bases de l'ADN qui passent d'un état empilé dans le duplexe à un état désordonné dans le simple brin. La température à laquelle la moitié des molécules d'ADN est dénaturée est appelée température de fusion (« melting temperature », T_m). Cette température peut être obtenue en calculant la dérivée première de la courbe de fusion et elle correspond à l'extremum de la courbe dérivée. La température de fusion dépend de la séquence du duplexe d'ADN : les séquences riches en paires de bases G•C, impliquant 3 liaisons hydrogène, sont plus stables (T_m plus élevée) et sont donc plus difficiles à dénaturer que les séquences riches en A•T (2 liaisons hydrogène).

À faible température, les MGBs peuvent interagir avec le duplexe d'ADN. Quand la température augmente, le duplexe d'ADN se dissocie et le MGB « sort » du duplexe provoquant un effet hypochromique (diminution de l'absorption dans l'UV) à 335 nm, longueur d'onde d'absorption des MGBs. Ainsi, il est possible d'étudier l'interaction ADN / MGB par des expériences de dénaturation thermique suivie par spectroscopie UV à 260 et 335 nm ; en effet, si la température de fusion du duplexe d'ADN augmente en présence du MGB, c'est qu'il y a une stabilisation du duplexe par le MGB due à la formation d'un complexe ADN / MGB. En déterminant le ΔT_m (T_m du complexe ADN / MGB - T_m du duplexe seul) pour chacun des MGBs, un classement peut être effectué afin de trouver la molécule qui stabilise au mieux le duplexe cible d'ADN.

A.2.2.1 Dénaturation thermique des duplexes d'ADN en présence des MGBs murins

A.2.2.1.1 Avec les MGBs murins standard

Pour étudier l'interaction entre l'ADN et les polyamides par la méthode de dénaturation thermique suivie par des mesures d'absorption en spectroscopie UV, deux paramètres ont été analysés : la stabilisation du duplexe d'ADN due à l'ajout de polyamide (augmentation de la T_m) et les changements spectraux qui accompagnent la fusion des complexes ADN / MGB.

Les profils typiques des courbes de dénaturation thermique suivie par UV du *Duplexe* 1 (duplexe modèle qui possède 4 sites de liaison pour F1) seul et en présence d'un excès molaire (8 fois) du polyamide F1 (2 molécules de MGB par site de liaison) ont été enregistrés (Annexe 9). Afin de pouvoir comparer strictement la dénaturation thermique des différents complexes ADN / MGB, le *Duplexe 2* qui possède un seul site de liaison pour chacun des MGBs a ensuite été utilisé.

L'exemple du profil de dénaturation thermique du *Duplexe 2* en présence et en absence d'un excès molaire (6 fois) du MGB *F4β* est montré dans la **Figure 33**. Sur la dérivée première de la courbe de fusion, un pic négatif à 335 nm (courbe verte, **Figure 33B**) est caractéristique de la dissociation du complexe ADN / MGB. Pendant la formation du complexe, le duplexe est fortement stabilisé : sa température de fusion passe de 49°C (courbes bleues, **Figures 33A** et **B**) à 67°C (courbes rouges, **Figures 33A** et **B**) ($\Delta T_m = 18^{\circ}$ C). De plus, la fusion du complexe ADN / MGB et la dissociation du duplexe d'ADN sont coopératives et s'effectuent à la même température. Les résultats obtenus avec les différents MGBs sont récapitulés dans le **Tableau 2**. Tous les polyamides synthétisés stabilisent le *Duplexe 2* (un seul site d'interaction) mais à des degrés différents. Les meilleures stabilisations ont été obtenues avec les polyamides *F4β* et *F1* ($\Delta T_m = 18^{\circ}$ C et 17°C, respectivement). Cette interaction est plutôt sélective, toutefois en présence des différents MGBs testés, une faible stabilisation du duplexe contrôle *Duplexe 3*, qui n'a pas de site de liaison pour les MGBs ($\Delta T_m \leq 5^{\circ}$ C), a été observée (exemple du MGB *F4β*; **Figures 33C** et **D**).



Figure 33 : Courbes de fusion des duplexes d'ADN et leurs dérivées premières en présence et en absence du polyamide murin F4β

Les expériences de dénaturation thermique ont été effectuées dans un tampon cacodylate de sodium (50 mM) à pH 6,2 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. A/ Courbes de fusion du **Duplexe 2** seul (1,3 μ M) à 260 nm (courbe bleue), en présence de 6 équivalents de **F4** β (7,8 μ M) à 260 nm (courbe rouge) et à 335 nm (courbe verte) et B/ leurs dérivées premières. C/ Courbes de fusion du duplexe contrôle (**Duplexe 3**) seul (1,3 μ M) à 260 nm (courbe bleue), en présence de 6 équivalents de **F4** β (7,8 μ M) à 260 nm (courbe verte) et D/ leurs dérivées premières. Les détails des expériences sont donnés dans le **Chapitre IV – Matériels et méthodes**.

Tableau 2 : Augmentations de la température de dissociation du Duplexe 2 en présence de 6 équivalents des différents polyamides murins déterminées par dénaturation thermique suivie par UV ($\Delta T_m \pm 0,5^{\circ}C$)									
Polyamide	F4β	F1	F5	F2β	F6	F4	F2	F3	F3β
ΔT_m (°C)	18	17	15	12	11	10	9	9	5

A.2.2.1.2 Avec les tandems de MGBs murins

Les expériences de dénaturation thermique avec le *Duplexe 1* (un seul site d'interaction pour chacun des tandems) et les tandems de polyamides T1 et T2 ont montré que

la fusion des complexes s'effectuait sur une large gamme de températures (**Figures 34A** et **B** ; **Annexe 10**) ; reflétant une faible coopérativité de la dénaturation des complexes ADN / tandem et / ou une interaction possible des composants individuels des tandems sur des sites secondaires.



Figure 34 : Courbes de fusion des duplexes d'ADN et leurs dérivées premières en présence et en absence du tandem T2

Les expériences de dénaturation thermique ont été effectuées dans un tampon cacodylate de sodium (50 mM) à pH 6,2 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. A/ Courbes de fusion du **Duplexe 1** seul (1,3 μ M) à 260 nm (courbe bleue), en présence de 6 équivalents de **T2** (7,8 μ M) à 260 nm (courbe rouge) et à 335 nm (courbe verte) et B/ leurs dérivées premières. C/ Courbes de fusion du **Duplexe 4** (sa séquence est donnée dans la **Figure 35**) seul (1,3 μ M) à 260 nm (courbe bleue), en présence de 6 équivalents de **T2** (7,8 μ M) à 260 nm (courbe bleue), en présence de 6 équivalents de **T2** (7,8 μ M) à 260 nm (courbe verte) et D/ leurs dérivées premières. Les détails des expériences sont donnés dans le **Chapitre IV – Matériels et méthodes**.

En effet, les tandems sont construits à partir d'un analogue de F1 (soit K1 pour le tandem T1 soit $K1\gamma$ pour le tandem T2) et comme F1 a 4 sites potentiels d'interaction sur le Duplexe 1 (Figure 27), alors il est envisageable que la composante K1 du tandem interagisse seule avec un autre site de liaison. C'est pourquoi des études approfondies ont été effectuées en utilisant le tandem T2 et des duplexes plus courts (Duplexes 4, 5 et 6) dans le but de minimiser les interactions secondaires (Figure 35).



Figure 35 : Représentation schématique des duplexes courts construits pour les études de l'interaction du tandem T2 avec l'ADN

Le **Duplexe 4** possède un site d'interaction pour le tandem T2 intégral, le **Duplexe 5** est muté sur le site de liaison du composant K1 γ du tandem T2 donc il ne possède que le site de liaison pour la partie K4 β du tandem et enfin le **Duplexe 6** possède uniquement le site d'interaction pour la partie K1 γ car il est muté sur le site de liaison de K4 β . Les paires de bases mutées sont encadrées sur chacun des duplexes.

Avec ces duplexes courts, les courbes de fusion des complexes étaient plus coopératives (**Figures 34C** et **D**). Dans le but de révéler une liaison monodentate possible, chacun des duplexes 5 et 6 possède un site d'interaction pour un seul composant du tandem *T2*. Le deuxième site est muté dans le site de liaison de *K1* γ (*Duplexe 5*) ou *K4* β (*Duplexe 6*) (**Figure 35**). Les résultats obtenus par les expériences de dénaturation thermique présentés

dans le Tableau 3 montrent l'existence d'une interaction du tandem par un seul de ses composant mais la stabilisation du *Duplexe 4* par le tandem *T2* est supérieure ($\Delta T_m = 16^{\circ}C$); en effet, l'interaction intégrale du tandem correspond à l'addition de celles sur les Duplexes 5 et *Duplexes 6* ($\Delta T_m = 9$ et 7°C, respectivement).

présence de 6 équivalents des tandems de polyamides murins déterminées par dénaturation thermique suivie par UV ($\Delta T_m \pm 0.5^{\circ}C$)					
Polyamide	ΔT_m (°C)	Cible d'ADN			
T1	$9\pm4*$	Duplexe 1			
T1	0	Duplexe 3			
T2	$9\pm4*$	Duplexe 1			
T2	0	Duplexe 3			
T2	16	Duplexe 4			
T2	9	Duplexe 5			
T2	7	Duplexe 6			

Tableau 3 · Augmentations de la température de dissociation des duplexes d'ADN en

*Transition large donc peu coopérative.

De plus, aucune stabilisation du Duplexe 3 (duplexe contrôle qui ne possède pas de site d'interaction pour les tandems) n'a été observée (**Tableau 3**). La construction de tandems de polyamides pour augmenter la spécificité en reconnaissant une séquence d'ADN plus longue est donc une approche valable; car une faible stabilisation du duplexe contrôle (Duplexe 3) avait été observée ($\Delta T_m \leq 5^{\circ}C$) dans le cas des MGBs standard (voir pour rappel § Chapitre II - A.2.2.1.1).

A.2.2.2 Dénaturation thermique des duplexes d'ADN en présence des sondes fluorescentes murines

L'interaction entre l'ADN et les sondes fluorescentes composées par des polyamides couplés avec divers fluorophores a également été étudiée par dénaturation thermique des complexes ADN / sonde. Le point de fusion a été déterminé, comme précédemment avec les MGBs, en utilisant le pic négatif des dérivées des courbes de fusion à 335 nm. Pour les sondes composées par des polyamides standard, le Duplexe 2 qui possède un seul site d'interaction pour chacune des sondes a été utilisé et les résultats obtenus sont présentés dans le **Tableau 4**. Les ΔT_m sont très inférieurs à ceux obtenus avec les MGBs non couplés. Ainsi, l'attachement d'un fluorophore aux MGBs, quel qu'il soit, mène à une déstabilisation significative du complexe ADN / MGB. Cette déstabilisation peut être due, au moins en partie, au masquage du groupement amine (qui peut être protoné) qui est utilisé pour le couplage du fluorophore.

Tableau 4 : Augmentations de la température de dissociation du Duplexe 2 en présence de 6 équivalents des sondes fluorescentes murines composées par des polyamides standard, déterminées par dénaturation thermique suivie par UV, dans un tampon cacodylate de sodium (50 mM) à pH 6,2 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂ ($\Delta T_m \pm 0,5^{\circ}$ C)

Sonde	F1-FITC	F4β-FITC	F4β-Cy5	F1-Cy3	F4β-fluo	F4-Cy5	F5-Cy5
$\Delta T_m(^{\circ}C)$	12	9	7	6	6	5	1

En ce qui concerne les tandems de polyamides fluorescents *T1-Cy3* et *T2-FITC*, les résultats obtenus étaient similaires : une forte diminution de la stabilisation des duplexes d'ADN a été observée.

A.2.2.3 Conclusions des expériences de dénaturation thermique avec les MGBs murins

Pour conclure sur ces expériences de dénaturation thermique, les MGBs standard sont capables d'interagir spécifiquement avec leur duplexe cible d'ADN en augmentant leur température de fusion. Cependant, ces augmentations diffèrent fortement (de 5 à 18°C) ; ce qui signifie que des petites variations structurales sur les MGBs peuvent entraîner un changement de leurs propriétés de fixation à l'ADN. Les deux polyamides standard $F4\beta$ et F1 stabilisent le mieux le duplexe cible d'ADN ($\Delta T_m = 18$ et 17°C ± 0,5°C, respectivement). Les tandems de polyamides T1 et T2 stabilisent également les duplexes cibles d'ADN, mais l'utilisation de séquences mutées sur un des sites de liaison pour chacun des composants du tandem a révélé qu'une interaction « monodentate » était possible ; en revanche, elle est moins forte que l'interaction complète avec le tandem. L'absence (dans le cas des tandems de polyamides) ou la faible (pour les MGBs standard) stabilisation du duplexe contrôle a permis de valider la sélectivité de ces molécules. Enfin, l'interaction de l'ADN avec les sondes fluorescentes (polyamide + fluorophore) a été vérifiée, cependant elle est altérée ; probablement en partie à cause du masquage de la fonction amine utilisée pour le couplage des fluorophores.
A.2.3 Électrophorèses en conditions non dénaturantes (« Gel-retard »)

Après avoir vérifié la formation des complexes ADN / MGB, la question de l'affinité des MGBs pour leur séquence cible d'ADN a été étudiée par la technique de retard de migration sur gel de polyacrylamide non dénaturant (« gel-retard »). En utilisant les duplexes cibles marqués à la fluorescéine et des concentrations croissantes de MGBs, leur affinité a pu être estimée (les détails des expériences sont donnés dans le **Chapitre IV – Matériels et méthodes**).

A.2.3.1 « Gels–retard » des duplexes d'ADN en présence des MGBs murins

A.2.3.1.1 Avec les MGBs murins standard

Tout d'abord, aucune interaction entre les MGBs standard et le *Duplexe 3* (contrôle qui ne possède aucun site de liaison pour les MGBs) n'a pu être observée par cette méthode, confirmant encore la sélectivité de ces molécules (**Annexe 11**). Ensuite, avec le *Duplexe 2* qui possède un seul site d'interaction pour chaque MGB murin, les résultats des expériences de « gel-retard » sont en accord avec ceux obtenus par dénaturation thermique (**Tableau 5**).

En dépit de leur parfaite correspondance aux règles de reconnaissance de Dervan, les différents polyamides synthétisés ont montré des différences significatives concernant leurs propriétés d'interaction avec l'ADN cible. Après l'incubation du *Duplexe 2* avec les MGBs *F1*, *F2β*, *F4*, *F4β* et *F5*, la migration des échantillons sur gel a révélé l'apparition d'une bande retardée (**Figure 36**) dont l'intensité augmente en fonction de la concentration en polyamides, prouvant la formation d'un complexe ADN / MGB. Des constantes de dissociation apparentes ($K_{d,app}$) ont pu être déterminées et sont données dans le **Tableau 5**.

En revanche avec les ligands F2, F3, $F3\beta$ et F6, les résultats obtenus par les expériences de « gel-retard » n'ont pas permis de déterminer des $K_{d,app}$; en effet, des traces diffuses (des « smears ») sont présentes entre les bandes (**Figure 36**), suggérant une dissociation des complexes dans le gel.



Figure 36 : Images des « gels-retard » des complexes ADN / MGB murins

Les électrophorèses ont été effectuées sur des gels de polyacrylamide à 20% non dénaturant dans du tampon TBE à 4°C. Le **Duplexe 2** (qui possède un seul site de liaison pour chacun des MGBs) a été incubé en présence des différents MGBs standard murins. Le brin riche en T du duplexe est marqué à la fluorescéine. Les concentrations des polyamides sont indiquées au-dessus de chaque piste des gels et celle du **Duplexe 2** est de 0,52 μ M dans chaque cas. Les schémas sur la droite symbolisent la position du duplexe (deux traits noirs) et des complexes ADN / MGB (deux traits noirs avec un point noir symbolisant le MGB) après la migration. Les bandes retardées détectées prouvent l'existence de complexes ADN / MGB. Les détails des expériences sont donnés dans le **Chapitre IV – Matériels et méthodes**.

Parmi les polyamides dont l'interaction avec l'ADN cible a pu être observée par « gelretard », le MGB *F4β* est celui qui possède la meilleure affinité ($K_{d,app} = 0,15 \ \mu M \pm 0,01 \ \mu M$), et il est à noter que c'est ce même MGB qui stabilise le mieux le duplexe ($\Delta T_m = 18^{\circ}C \pm 0.5$ °C) dans les expériences de dénaturation thermique. Son K_{d,app} est 30 fois supérieur à celui de son analogue aromatique F4 (qui ne possède pas de synthon β -ala) et 9 fois supérieur à celui d'un autre ligand F5 qui peut reconnaitre uniquement des séquences A•T (Tableau 5). Il est également notable que tous les ligands qui stabilisent faiblement le duplexe par dénaturation thermique, et dont le K_{d.app} n'a pas pu être déterminé, ont deux synthons N-méthylimidazole. Pour les ligands dont l'affinité a pu être estimée, seul $F2\beta$ a deux Im. Cependant, il possède aussi une β-ala, qui rend probablement le MGB plus flexible pour faciliter son alignement sur sa séquence cible d'ADN dans le *Duplexe 2*. La différence entre $F2\beta$ (K_{d,app} = 1,29 μ M ± 0,18 μ M et $\Delta T_m = 12^{\circ}C \pm 0.5^{\circ}C$) et *F3β* (K_{d,app} non déterminée et $\Delta T_m = 5^{\circ}C \pm 0.5^{\circ}C$) est très difficile à rationaliser. La seule différence repose sur la position des deux Im, dans le domaine C-term des polyamides, qui pourrait peut-être rendre la molécule trop rigide même après l'insertion d'une β -ala.

de polyacrylamide non dénaturant à 20%, réalisées à 4°C dans du tampon TBE. Les augmentations de la température de fusion du Duplexe 2 en présence de 6 équivalents des MGBs sont données pour rappel ($\Delta T_m \pm 0.5^{\circ}$ C) (voir § Chapitre II - A.2.2.1.1)		
Polyamide	$K_{d,app}\left(\mu M ight)$	ΔT_m (°C)
F 4β	$0,15 \pm 0,01$	18
F1	$0,73 \pm 0,06$	17
F5	$1,3 \pm 0,02$	15
F 2β	$1,\!29 \pm 0,\!18$	12
F6	-	11
F4	$4,4 \pm 0,06$	10
F2	-	9
F3	-	9
F3β	-	5

Tableau 5 : Constantes de dissociation apparentes (K_{d,app}) des complexes entre le Duplexe 2 et les différents polyamides, déterminées par des expériences de retard sur gel

A.2.3.1.2 Avec les tandems de MGB murins

Aucune interaction entre les tandems de polyamides murins et le Duplexe 3 (contrôle qui ne possède aucun site de liaison pour les tandems de MGBs) n'a été observée par cette méthode, confirmant encore la sélectivité de ces molécules (**Figure 37**). Avec le *Duplexe 1* qui possède un seul site d'interaction pour chacun des tandems de MGBs murins, les résultats des expériences de « gel-retard » corrèlent avec les résultats de dénaturation thermique (**Tableau 6**).

Après l'incubation du **Duplexe 1** avec les tandems **T1** et **T2**, la migration des échantillons sur gels a révélé l'apparition d'une bande retardée (**Figure 37**), dont l'intensité augmente en fonction de la concentration en tandems de polyamides, prouvant la formation de complexes ADN / tandem. Les $K_{d,app}$ ont pu être déterminés et sont donnés dans le **Tableau 6**. Au vu des résultats obtenus par ces expériences, le tandem **T2** semble posséder une meilleure affinité pour la séquence cible d'ADN. Des études avec les duplexes **4**,**5** et **6** ont donc été poursuivies afin de déterminer l'affinité du tandem **T2** quand il se lie soit intégralement (par ses deux parties, **Duplexe 4**) à sa séquence cible, soit par un seul de ses composants (pour la description des duplexes, voir § Chapitre II - A.2.2.1.2). L'affinité de **T2** pour le **Duplexe 4** qui contient son site de liaison complet est 30 fois supérieure à celles obtenues avec les **Duplexes 5** et **6** mutés sur l'un des sites de liaison du tandem. Ces résultats sont donc en accord avec ceux obtenus par les expériences de dénaturation thermique.





Les électrophorèses ont été effectuées sur des gels de polyacrylamide à 20% non dénaturant dans du tampon TBE à 4°C. À gauche, le **Duplexe 1** (qui possède un seul site de liaison pour chacun des tandems de MGBs) a été incubé en présence de **T1** ou **T2**. Les bandes retardées détectées prouvent l'existence de complexes ADN / tandem. À droite, le **Duplexe 3** (qui ne possède aucun site de liaison pour chacun des tandems de MGBs) a été incubé en présence de **T1** ou **T2**. Aucune bande retardée n'a été observée donc aucune preuve d'une quelconque interaction avec le duplexe contrôle n'a été mise en évidence. Les concentrations des tandems sont indiquées au-dessus de chaque piste des gels et celles des duplexes d'ADN sont de 0,52 μ M dans chaque cas. Le brin riche en T des duplexes est marqué à la fluorescéine. Les schémas sur la droite de chaque gel symbolisent la position du duplexe (deux traits noirs) et des complexes ADN / tandem (deux traits noirs avec un point noir symbolisant le tandem) après la migration. Les détails des expériences sont donnés dans le **Chapitre IV – Matériels et méthodes**.

Tableau 6 : Constantes de dissociation apparentes (Kd,app) des complexes entre les
duplexes d'ADN et les tandems de polyamides murins déterminées par des expériences
de retard sur gel de polyacrylamide à 20% non dénaturant dans du tampon TBE à 4°C
et les augmentations de la température de fusion des duplexes en présence de 6
équivalents des tandems sont données pour rappel ($\Delta T_m \pm 0.5^{\circ}C$) (voir § Chapitre II -
A.2.2.1.2)

Polyamide	$K_{d,app}$ (μM)	ΔT_m (°C)	Cible d'ADN
T1	$1,93 \pm 0,24$	9 ± 4	Duplexe 1
Τ2	$0,\!76\pm0,\!03$	9 ± 4	Duplexe 1
Τ2	$0,\!19\pm0,\!05$	16	Duplexe 4
Τ2	$5,\!80\pm1,\!57$	9	Duplexe 5
Τ2	$6,83 \pm 1,97$	7	Duplexe 6

A.2.3.2 « Gels-retard » des duplexes d'ADN en présence des sondes fluorescentes murines

Les expériences de dénaturation thermique ont montré que le couplage d'un fluorophore sur un polyamide déstabilisait le complexe ADN / MGB aussi bien dans le cas des MGBs standard qu'avec les tandems. Aucune bande retardée n'a pu être mise en évidence dans les expériences de « gel-retard » des complexes ADN / sondes fluorescentes. Afin d'éliminer l'hypothèse d'un encombrement stérique dû à la présence de la fluorescéine à l'extrémité d'un des brins du duplexe d'ADN, des expériences avec des duplexes marqués radioactivement au phosphate ³²P ont été effectuées, mais là encore, aucune bande retardée n'a été visible (les détails des expériences sont donnés dans le **Chapitre IV – Matériels et méthodes**). La forte déstabilisation des complexes engendre probablement leur dissociation dans les gels donc les K_{d,app} n'ont pas pu être déterminés par cette technique.

A.2.3.3 Conclusions des expériences de « gel-retard » avec les MGBs murins

Pour conclure sur ces expériences de « gel-retard », l'apparition de bandes retardées dont l'intensité augmente en fonction de la concentration de polyamides permet de valider que les MGBs sont capables de former des complexes avec leur duplexe cible d'ADN avec des affinités micromolaires ou submicromolaires. Cependant, les affinités de ces polyamides pour leur séquence cible sont différentes. Les deux polyamides standard $F4\beta$ et F1 sont les plus affins pour le duplexe cible d'ADN (K_{d,app} = 0,15 ± 0,01 µM et 0,73 ± 0,06 µM, respectivement). Ces résultats sont en accord avec ceux obtenus par dénaturation thermique $(\Delta T_m, F4\beta = 18^\circ C \pm 0.5^\circ C$ et $\Delta T_m, F1 = 17^\circ C \pm 0.5^\circ C$). Les tandems de polyamides T1 et T2 forment également des complexes avec les duplexes cibles d'ADN. En revanche, l'utilisation de séquences mutées sur un des sites de liaison pour chacun des composants du tandem a révélé que des interactions « monodentate » étaient possibles ; cependant, le tandem T2 interagit avec sa séquence cible intégrale avec une affinité 30 fois supérieure à celles obtenues avec les séquences mutées sur un site. L'absence de bande retardée sur les gels avec le duplexe contrôle a aussi permis de valider la sélectivité de ces molécules. Enfin, l'interaction de l'ADN avec les sondes fluorescentes n'a pas pu être mise en évidence par cette technique.

A.2.4 Dichroïsme circulaire

Une autre technique utilisée pour étudier l'interaction ADN / MGB a été le dichroïsme circulaire (DC), qui est une méthode de choix pour l'étude de molécules chirales en solution et permet d'obtenir des informations structurales. Le DC est communément utilisé pour étudier les macromolécules biologiques telles que l'ADN : soit pour évaluer les changements de conformation des macromolécules elles-mêmes, soit pour révéler leurs interactions avec de petites molécules, plus particulièrement celles qui sont achirales et dont le signal induit est uniquement dû à leur interaction avec la macromolécule. En ce qui concerne les MGBs, ils sont achiraux et ne sont pas « structurés » en solution. C'est uniquement au contact de l'ADN qu'ils prennent leur forme « d'épingle à cheveux » antiparallèle qui leur permet de reconnaître leur séquence cible. Ainsi, les MGBs seuls n'ont aucun signal de DC et lorsqu'ils interagissent avec l'ADN, alors un signal induit apparaît qui est une signature du mode d'interaction des MGBs au sein du petit sillon de l'ADN.

A.2.4.1 Interaction ADN / MGBs murins

A.2.4.1.1 Avec les MGBs murins standard

L'équipe de Gursky a démontré qu'en présence de séquences cibles d'ADN, les profils des spectres de DC de la bis-nétropsine ont des formes particulières qui dépendent de sa conformation au sein du petit sillon de l'ADN : linéaire monomoléculaire, « épingle à cheveux » parallèle ou antiparallèle, orientation parallèle ou antiparallèle bimoléculaire (Surovaya et al. 2001).

Les expériences de dichroïsme circulaire ont été conduites en titrant une solution de duplexe cible d'ADN (*Duplexe 1* qui possède plusieurs sites d'interaction pour les MGBs) par des solutions des différents MGBs synthétisés. Dans tous les cas, les MGBs se lient à leur ADN cible puisqu'un spectre de DC induit a été obtenu tandis qu'aucun signal de DC n'a été observé avec les MGBs seuls en solution. La présence d'un signal induit large et positif entre 300 et 400 nm est caractéristique d'une interaction entre les MGBs et l'ADN double brin (longueur d'onde d'absorption des MGBs) (Wang et al. 2012). Les spectres différentiels de DC (le spectre de DC de l'ADN a été soustrait) des deux polyamides les plus affins *F1* et *F4* β en complexe avec le *Duplexe 1* sont donnés dans les **Figures 38A** et **B**, respectivement. Ils présentent chacun une large bande entre 300 et 400 nm qui révèle une interaction des polyamides avec le *Duplexe 1*.

En comparaison avec les spectres modèles établis par Gursky, $F4\beta$ a un spectre de DC induit typique de la signature d'un positionnement en « épingle à cheveux » antiparallèle au sein du petit sillon du duplexe cible d'ADN (Figure 38B); alors que F1 a un spectre légèrement différent (Figure 38A). Cependant, ce dernier ressemble plus à un spectre caractéristique du positionnement en « épingle à cheveux » antiparallèle plutôt qu'aux autres orientations (Halby et al. 2007). La différence de forme des spectres entre F1 et F4 β repose probablement sur le remplacement d'un Py par un synthon β -ala dans le dérivé F4 β , ce qui peut changer légèrement l'environnement du complexe ADN / MGB dans le petit sillon.

La saturation des courbes de titration commence après l'ajout de 4 équivalents de *F1* qui a quatre sites de liaison sur la séquence du *Duplexe 1* (Figure 38C) et de 2 équivalents de *F4β* qui a deux sites de liaison sur la séquence du *Duplexe 1* (Figure 38D), ce qui correspond à une stœchiométrie 1 :1.

Les spectres de DC induits des MGBs $F2\beta$ et F5 en interaction avec l'ADN ressemblent à ceux obtenus avec F1. En revanche, pour F4 ayant la plus faible affinité, les spectres ont une forme légèrement différente et l'amplitude des signaux est relativement faible, comme ceux obtenus avec le *Duplexe 3* (contrôle qui n'a aucun site d'interaction pour chacun des MGBs) avec lequel une interaction non-spécifique des MGBs murins a été détectée (**Figures 38E** et **F**). De plus, l'interaction des MGBs avec l'ADN ne provoque pas de changement significatif des signaux de DC de l'ADN ; donc les polyamides ne causent pas de changement majeur.



Figure 38 : Spectres différentiels de dichroïsme circulaire des complexes ADN / MGB

Les expériences de dichroïsme circulaire ont été effectuées dans un tampon Tris-HCl (50 mM) à pH 7,5 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. Spectres de DC différentiels des complexes ADN : MGB, A/ Duplexe 1 : F1, B/ Duplexe 1 : F4 β , E/ Duplexe 3 : F1 et F/ Duplexe 3 : F4 β . Les spectres de DC de l'ADN sont soustraits. Des solutions de duplexes d'ADN à 2,5 μ M sont titrées par des solutions de polyamides à 250 μ M (voir description dans le Chapitre IV – Matériels et méthodes). Les concentrations de polyamides après chaque ajout sont indiquées à droite de chaque figure. Courbes de titration : ellipticité des complexes C/ Duplexe 1 : F1 et D/ Duplexe 1 : F4 β , à 335 nm aux concentrations de polyamides.

A.2.4.1.2 Avec les tandems de MGBs murins

Les spectres de dichroïsme circulaire enregistrés avec les tandems de polyamides et les duplexes cibles d'ADN étaient similaires à ceux obtenus avec *F1* (Figure 38A); ce qui

suggère un positionnement en « épingle à cheveux » antiparallèle des tandems de polyamides au sein du petit sillon de l'ADN. À titre d'exemples, les spectres de DC différentiels de T2 en complexe avec le *Duplexe 4* (qui possède le site d'interaction intégral pour T2) sont donnés dans la Figure 39A.



Figure 39 : Spectres différentiels de dichroïsme circulaire des complexes ADN / tandem

Les expériences de dichroïsme circulaire ont été effectuées dans un tampon Tris-HCl (50 mM) à pH 7,5 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. A/ **Duplexe 4** : **T2** et B/ **Duplexe 3** : **T2**. Les spectres de l'ADN sont soustraits. Des solutions de duplexes d'ADN à 2,5 μ M sont titrées par une solution de tandem de polyamides à 250 μ M (voir description dans le **Chapitre IV** – **Matériels et méthodes**). Les concentrations de **T2** après chaque ajout sont indiquées à droite de chaque figure. C/ Courbes de titration : ellipticité du complexe **Duplexe 4** : **T2** à 334 nm aux concentrations de tandem indiquées.

La courbe de titration du *Duplexe 4* par *T2* (Figure 39C) montre une saturation du duplexe dès l'ajout d'une concentration équimolaire de tandem $(2,5 \mu M)$. De plus, une preuve supplémentaire de la meilleure spécificité du tandem par rapport aux polyamides standard est qu'en présence du *Duplexe 3* (contrôle qui ne possède pas de site d'interaction pour le tandem), aucun spectre induit n'a été observé (Figure 39B). Cependant, avec les duplexes mutés chacun sur un site de liaison pour un des composants du tandem (*Duplexe 5* et *6*), des spectres similaires à ceux obtenus avec le *Duplexe 4* (ayant le site intégral pour le tandem) ont été enregistrés, mais toutefois avec une amplitude plus faible (ellipticité maximale plus faible), et les courbes de titration atteignent une saturation à des concentrations de tandem plus élevées (2 - 3 équivalents) (Annexe 12). Ces résultats appuient l'hypothèse d'une

interaction « monodentate » du tandem de polyamides avec chacun des duplexes d'ADN mutés.

A.2.4.2 Interaction ADN / sondes fluorescentes murines

Les spectres de dichroïsme circulaire enregistrés avec les sondes fluorescentes sont similaires à ceux obtenus avec les MGBs non couplés ; ce qui suggère une conformation en « épingle à cheveux » antiparallèle des sondes fluorescentes dans le petit sillon de l'ADN et donc que le couplage à un fluorophore ne modifie pas le mode d'interaction du polyamide avec l'ADN. Cependant, à l'exception des spectres obtenus avec le tandem *T2-FITC*, les amplitudes des signaux de DC sont plus faibles que celles obtenues avec les polyamides non couplés aux fluorophores. Ceci pourrait refléter la perte d'affinité des sondes pour l'ADN comparée aux MGBs non couplés et la diminution de la stabilité des complexes ADN / sondes fluorescentes. Plusieurs exemples de spectres de DC des sondes fluorescentes lors de leur interaction avec des duplexes d'ADN sont donnés dans l'Annexe 13.

A.2.4.3 Conclusions des expériences de dichroïsme circulaire

Les résultats des expériences de dichroïsme circulaire ont permis de valider les résultats obtenus par les autres techniques et de montrer que les polyamides ne se structurent qu'au contact de leur ADN cible. En effet, les polyamides en solution n'ont pas de spectres de DC et c'est lors de leur interaction avec l'ADN qu'un signal de DC induit apparait. De ces expériences, il apparait donc que des complexes ADN / MGB murins se forment, et les expériences avec un duplexe contrôle montre une faible interaction non-spécifique des polyamides. De plus, en comparaison avec les spectres modèles établis par Gursky, la forme des spectres obtenus apporte des informations sur la structure des complexes ADN / MGB. Les résultats suggèrent un positionnement dans une conformation d' « épingle à cheveux » antiparallèle des MGBs murins au sein du petit sillon de l'ADN. L'interaction de l'ADN avec les tandems de polyamides et les sondes fluorescentes murins a également été vérifiée.

Ainsi, l'ensemble des résultats obtenus par ces études physico-chimiques a révélé que les polyamides F1, $F2\beta$ et $F4\beta$ étaient les meilleurs candidats pour la construction de sondes fluorescentes à base de polyamides. Des expériences de FRET ont été réalisées avec ces molécules (voir § Chapitre II - A.2.5.2) et ils ont aussi été couplés en tandems de MGBs (T1 et T2). Le polyamide F5 qui reconnait des séquences A/T, semblait aussi être un candidat

intéressant mais qui pourrait s'avérer peu spécifique pour des études cellulaires car le génome murin contient d'autres régions riches en A et T. De plus, F1 et F5 sont séparés par une distance assez longue sur le *Duplexe 1*; ce qui pourrait compliquer la construction d'un tandem et la possibilité d'observer un effet FRET.

A.2.5 Spectroscopie de fluorescence

L'étude des sondes fluorescentes murines a été ensuite poursuivie par des expériences de spectroscopie de fluorescence dans le but d'observer une modulation de leur spectre d'émission de fluorescence lors de leur interaction avec leur cible d'ADN. Aussi, en utilisant des paires de sondes donneur / accepteur, un effet FRET a été étudié lorsque les sondes sont situées à proximité l'une de l'autre sur la séquence cible d'ADN.

A.2.5.1 Interaction ADN / sondes fluorescentes murines

Aucun changement des profils d'émission de fluorescence des fluorophores n'a été mis en évidence après leur conjugaison aux différents polyamides. Lors de l'interaction des sondes fluorescentes (polyamide + fluorophore) avec l'ADN cible, des modifications de leur intensité de fluorescence ont été observées. Des études avaient déjà montré des résultats similaires (Rucker et al. 2003; Rucker et al. 2004; Fechter et al. 2005; Hsu & Dervan 2008). Dans notre cas, une forte augmentation séquence-spécifique de l'intensité de fluorescence (36 fois, soit 9 fois supérieure à celle observée en présence du duplexe contrôle) a été observée lors de l'interaction de *F1-FITC* avec le duplexe cible d'ADN (*Duplexe 1*) (Figure 40B et Tableau 7). Pour les MGBs marqués par des fluorophores de type cyanine, l'augmentation de l'intensité de fluorescence était plus faible, mais toujours supérieure en présence du duplexe cible (*Duplexe 1*) comparée à celle avec le duplexe contrôle (*Duplexe 3*). À titre d'exemples, les spectres de fluorescence des sondes *F1-Cy3*, *F4β-Cy5* et *F4β-FITC* seules ou en présence des *Duplexes 1* et 3 sont donnés dans les Figures 40A, C et D. Le Tableau 7 résume l'ensemble des résultats obtenus.



Figure 40 : Spectres de fluorescence des sondes murines en absence et en présence des duplexes d'ADN

Les expériences de fluorescence ont été conduites dans un tampon Tris-HCl (10 mM) à pH 7,5 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. Les concentrations des duplexes d'ADN et des sondes fluorescentes sont de 2 µM. Les spectres de fluorescence des sondes seules sont représentés en bleu, ceux en présence du **Duplexe 3** (contrôle qui ne possède aucun site d'interaction pour chacune des sondes) en rouge et en présence du **Duplexe 1** (cible qui possède plusieurs sites d'interaction pour chacune des sondes) en vert. A/**F1-Cy3**, B/**F1-FITC**, C/**F4β-Cy5** et D/**F4β-FITC** (voir description dans le **Chapitre IV** – **Matériels** et méthodes).

La modulation des spectres d'émission de fluorescence des sondes a été observée avec le *Duplexe 1* (cible, courbes vertes **Figure 40**), mais aussi et plus faiblement avec le *Duplexe* 3 (contrôle, courbes rouges **Figure 40**). Le bruit de fond de fluorescence des sondes en présence du duplexe contrôle peut s'expliquer par une possible interaction non-spécifique des fluorophores avec l'ADN.

Tableau 7 : Augmentations des intensités de fluorescence des sondes murines (2 μ M) en présence des duplexes d'ADN (2 μ M) cible (Duplexe 1) et contrôle (Duplexe 3) dans un tampon Tris-HCl (10 mM) à pH 7,5 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂

Nom	Duplexe 1	Duplexe 3
F1-FITC	× 36,2	× 4
F4β-FITC	× 5	× 2,5
F4β-Cy5	× 3,7	× 1,4
F1-Cy3	$\times 2,2$	× 1,4
F2β-Cy5	× 1,2	× 1,04

En résumé, les résultats obtenus ont pu mettre en évidence une influence de la nature du fluorophore utilisé. Une augmentation plus forte de l'intensité de fluorescence des sondes a été observée avec les MGBs marqués par la FITC. Aussi, en comparant *F1-FITC* et *F4β-FITC*, le type de polyamide semble avoir un impact sur cette augmentation observée. Dans l'ensemble, la sonde *F1-FITC* semble être le candidat le plus prometteur pour pouvoir observer un signal de fluorescence dans des expériences d'imagerie cellulaire.

A.2.5.2 Transfert par résonance d'énergie de fluorescence (FRET) des sondes fluorescentes murines

Les sondes *F1-Cy3* (donneur) et *F2\beta-Cy5* ou *F4\beta-Cy5* (accepteurs) sont susceptibles de former des paires de sondes fluorescentes pour étudier un effet FRET.

Les premières expériences ont été conduites avec la sonde $F2\beta$ -Cy5 en tant qu'accepteur, mais l'effet FRET observé n'était pas optimal puisque l'intensité de la fluorescence émise par Cy5 était faible comparée à celle observée pour Cy3. En effet, l'extinction complète de la fluorescence de F1-Cy3 n'a pas été observée (Annexe 14). Ensuite, le même type d'expérience a été réalisé en remplaçant la sonde $F2\beta$ -Cy5 par la sonde $F4\beta$ -Cy5 (Figure 41).

Cependant, aucune différence significative n'a été mise en évidence entre les intensités de fluorescence (induite par FRET) de Cy5 couplé au MGB $F2\beta$ qui est proche de F1 sur la séquence cible d'ADN ou au MGB $F4\beta$ qui est plus distant de F1. En éloignant les deux sondes sur la séquence cible d'ADN, aucune amélioration de l'effet FRET n'a pu être observée.

La **Figure 41** montre les spectres de fluorescence de *F1-Cy3* seule et en présence de *F4β-Cy5* complexées ou non au *Duplexe 1* (cible d'ADN). En excitant la sonde donneuse à 510 nm, une émission de fluorescence de la sonde acceptrice typique de Cy5 avec un maximum à 660 nm apparait. L'effet FRET observé s'intensifie avec des concentrations croissantes de sonde acceptrice (dégradé de verts, **Figure 41**), de telle sorte qu'un effet plus important a été observé à une concentration de sonde acceptrice supérieure (6 μ M) à une concentration équimolaire de sonde donneuse (*F1-Cy3*, 2 μ M). Cependant l'extinction complète de la sonde *F1-Cy3* n'a pas pu être obtenue.



Figure 41 : Expériences de FRET entre les sondes murines F1-Cy3 et F4 β -Cy5 en présence du duplexe cible d'ADN

Les expériences ont été réalisées dans un tampon Tris-HCl (10 mM) à pH 7,5 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. La longueur d'onde d'excitation était de 510 nm et l'émission a été enregistrée dans la gamme 530 - 750 nm. Les concentrations de F1-Cy3 et du **Duplexe 1** sont de 2 μ M. En bleu, le spectre de fluorescence de la sonde donneuse F1-Cy3 seule. En bleu cyan, le spectre de fluorescence de la sonde acceptrice F4 β -Cy5 seule à 2 μ M. En rouge les sondes F1-Cy3 (donneur) et F4 β -Cy5 (accepteur) à 2 μ M sans l'ADN. En vert, les sondes F1-Cy3 et F4 β -Cy5 en présence du **Duplexe 1**. Le dégradé de verts symbolise l'augmentation de la concentration de F4 β -Cy5 : en vert clair - 2 μ M, en vert foncé - 4 μ M et en vert très foncé – 6 μ M. En excitant la sonde donneuse F1-Cy3 à 510 nm, une émission de fluorescence à 660 nm typique de Cy5 (sonde acceptrice) apparait : c'est l'effet FRET.

A.2.5.3 Conclusions des expériences de fluorescence

Une augmentation de l'intensité de fluorescence des sondes murines a été observée lors de leur interaction avec le duplexe cible d'ADN. Avec le duplexe contrôle, une faible augmentation a été mise en évidence qui pourrait résulter d'une interaction non-spécifique des fluorophores avec l'ADN. D'après les résultats de ces expériences, la sonde *F1-FITC* semble être le meilleur candidat pour des expériences d'imagerie cellulaire. En utilisant les sondes *F1-Cy3* et *F2β-Cy5* ou *F4β-Cy5*, des effets FRET ont pu être observés lorsqu'elles interagissent à proximité l'une de l'autre sur la séquence cible d'ADN. Cependant, ces derniers n'étaient pas optimaux.

A.3 Études de l'interaction des sondes fluorescentes murines avec l'ADN *in cellulo*

Pour finir, les sondes fluorescentes ont été évaluées pour leur capacité à détecter leurs séquences cibles d'ADN dans un contexte cellulaire. Un des critères majeur pour l'imagerie de séquences d'ADN spécifiques, dans des cellules vivantes, est la possibilité de détecter un signal spécifique produit par les molécules fluorescentes liées à leur cible d'ADN et de pouvoir distinguer ce signal du bruit de fond de fluorescence qui peut être dû à la présence de molécules non-liées. L'objectif était de pouvoir visualiser ces structures nucléaires directement dans des cellules vivantes, en utilisant :

- d'une part la capacité des séquences péricentromériques murines à s'associer sous forme de chromocentres, et ainsi de permettre une accumulation de signal fluorescent à des loci spécifiques (due à la présence de multiples répétitions) et,

- d'autre part les propriétés des sondes précédemment décrites qui sont capables d'augmenter leur intensité de fluorescence lors de leur interaction avec la séquence d'ADN spécifique.

A.3.1 Cellules murines vivantes

A.3.1.1 Problèmes de pénétration nucléaire

Tout d'abord, les études ont été conduites avec les sondes marquées par les fluorophores de type cyanine (Cy3 et Cy5) couplés aux MGBs par une liaison amide (pour rappel voir § Chapitre II - A.1.2). Des images de cellules fibroblastiques murines NIH-3T3 vivantes après un traitement de 24h (temps d'incubation nécessaire à l'obtention d'un bon rapport signal / bruit de fond) par 2 μ M (concentrations optimales) des sondes *F1-Cy3* et

 $F4\beta$ -Cy5 sont représentées dans la Figure 42. Une brève étape de lavage par du tampon phosphate salin (PBS) a été effectuée préalablement à la prise d'image en chambre humide à 37°C. Deux problèmes majeurs ont été mis en évidence avec l'utilisation de ces sondes : une absence de pénétration nucléaire et la présence de nombreux agrégats fluorescents dans le cytoplasme, surtout avec la sonde *F1-Cy3* (Figure 42A).



Figure 42 : Images de cellules murines vivantes NIH-3T3 traitées par F1-Cy3 et F4β-Cy5

Images de cellules murines NIH-3T3 vivantes traitées par 2 μ M de F1-Cy3 (A) ou F4 β -Cy5 (B) pendant 24h. 1 : les images de fluorescence et 2 : les superpositions des images avec F1-Cy3 (A/2) ou F4 β -Cy5 (B/2), en rouge et de la lumière blanche, en gris. Barre d'échelle : 10 μ m (voir Chapitre IV - Matériels et méthodes pour les conditions détaillées des expériences).

Pour tenter de résoudre ces problèmes de pénétration, des expériences de nucléofection (transfection nucléaire) et de perméabilisation des cellules par des détergents, à différentes concentrations, comme le Tween 20 ou la saponine ont été menées avec ces mêmes sondes ; mais la viabilité des cellules a été fortement altérée et aucun marquage nucléaire n'a été observé.

Concernant les agrégats fluorescents, quatre types de stratégies ont été suivies : 1) un changement de milieu de culture en éliminant le sérum de veau fœtal (SVF) car celui-ci pourrait provoquer l'agrégation de molécules, 2) la solubilisation des agrégats par l'utilisation de différentes concentrations de Hp- β -cyclodextrine (Gould & Scott 2005; Hargrove et al. 2012) (pour rappel voir § Chapitre I – D.5), 3) l'éloignement du fluorophore de la partie polyamide en allongeant le bras de liaison avec un γ supplémentaire en N-term du polyamide (synthèse de la sonde *F4\beta\gamma-Cy5*; la formule du MGB *F4\beta\gamma* est donnée en Annexe 1), ou

encore 4) différentes gammes de concentrations de sondes ont été testées. Aucune de ces expériences n'a permis d'améliorer la pénétration nucléaire. Cependant la privation de SVF a permis d'éviter la formation d'agrégats fluorescents.

A.3.1.2 Changements de fluorophores

L'hypothèse impliquant le fluorophore comme étant à l'origine des problèmes de pénétration nucléaire a été testée en utilisant tout d'abord des fluorophores tels que TO, S24 ou la coumarine MM14 (les structures des fluorophores sont présentées dans l'**Annexe 3**) qui ont des masses molaires inférieures à celle des cyanines. Ces fluorophores ont été couplés aux MGBs par des liaisons amide ou par « chimie-click » afin de vérifier l'importance du bras de liaison entre le fluorophore et le MGB, mais là encore aucun marquage nucléaire n'a été observé.

Par la suite, l'utilisation de la fluorescéine et de son analogue isothiocyanate (FITC) a révélé que le bras de liaison avait effectivement une influence sur la pénétration cellulaire (**Figure 43**). Le MGB $F4\beta$ a été couplé, d'une part à l'ester N-succinimidyle de la 5-carboxyfluorescéine en formant une liaison amide et, d'autre part à la FITC par une liaison thiourée. Avec la sonde $F4\beta$ -fluorescéine, les résultats obtenus étaient similaires à ceux observés avec les fluorophores de types cyanine (**Figure 43A**).

En revanche, avec $F4\beta$ -FITC (Figure 43B) et F1-FITC (Figure 43C), une pénétration nucléaire a été observée. Les profils de marquage ponctuel sous forme de foyers semblaient indiquer que les sondes F1-FITC et F4 β -FITC marquent les séquences d'ADN répétées péricentromériques murines associées sous forme de chromocentres. Ces résultats ont apporté les premières indications que les sondes fluorescentes à base de polyamides permettent la détection de séquences cibles d'ADN dans des cellules vivantes. Le meilleur rapport signal / bruit de fond a été obtenu avec F1-FITC (Figure 43C). Ces résultats sont en accord avec ceux obtenus lors des expériences de spectroscopie de fluorescence puisqu'une forte augmentation de l'intensité du signal de fluorescence a été observée avec cette sonde lors de son interaction spécifique avec le duplexe cible (voir § Chapitre II - A.2.5.1). De plus, aucune perturbation de la morphologie des cellules ou du rythme de division cellulaire n'a été observée après 5 jours d'incubation. Le traitement par ces sondes fluorescentes n'est donc pas cytotoxique.



Figure 43 : Images de cellules murines vivantes NIH-3T3 traitées par les sondes marquées par la fluorescéine et son analogue isothiocyanate (FITC)

Images de cellules murines NIH-3T3 vivantes traitées par 2 μ M de F4 β -fluo (A) ou F4 β -FITC (B) ou F1-FITC (C et D) pendant 24h. 1 et 3 : les images de fluorescence, 2 et 4 : la superposition des images de F4 β -fluo (A/2) ou F4 β -FITC (B/2) ou F1-FITC (C/2 et 4), en rouge et de la lumière blanche, en gris. D/ Images de fluorescence de F1-FITC prises à différents temps au niveau d'un même champ de cellules. Barre d'échelle : 10 μ m (voir Chapitre IV - Matériels et méthodes pour les conditions détaillées des expériences).

L'identification de l'importance d'une liaison thiourée a conduit à la conception de cyanine sous forme d'isothiocyanate (*F1-Cy3-ITC*). Malheureusement cette sonde ne

pénétrait toujours pas dans les cellules (un tableau récapitulant l'ensemble des sondes fluorescentes synthétisées est donné dans l'Annexe 5).

Cependant, un inconvénient à utiliser la FITC est l'extinction rapide de la fluorescence de la sonde (photoblanchiment) qui est un obstacle à la prise d'images successives et donc à une étude dynamique des cellules. Des efforts ont donc été faits pour améliorer la prise d'images en utilisant un milieu de culture DMEM^{gfp} (Dulbecco's Modified Eagle Medium) qui est déplété de deux vitamines (riboflavine et pyridoxal) et en ajoutant du rutin qui est un flavonoïde de plante servant à éliminer les radicaux libres de l'oxygène (**Figure 43D**) (**Bogdanov et al. 2012**). Ces conditions ont permis d'améliorer légèrement les conditions de prise d'images mais l'extinction du signal était toujours trop rapide et ne permettait pas de suivre en temps réel l'évolution des séquences marquées. Une autre stratégie pour éviter cette photosensibilité de la FITC était d'utiliser d'autres fluorophores de structures similaires mais moins sensibles à la lumière (voir **Annexes 5 et 6**). C'est pourquoi des dérivées de rhodamine ont aussi été utilisés mais sans succès. Seul le fluorophore Oregon Green sous forme d'isothiocyanate (OG-ITC) couplé au MGB *F1 (F1-OG-ITC)* a permis d'obtenir les mêmes résultats que ceux en présence de *F1-FITC*, mais sans aucune amélioration de la photosensibilité.

Des expériences cellulaires similaires ont aussi été tentées avec les tandems de polyamide murins, soit *T1* marqué par Cy3, soit *T2* marqué par la FITC. Mais là aussi aucune pénétration cellulaire n'a pu être observée.

A.3.1.3 Contrôles

Dans le but de vérifier que les profils de marquages nucléaires obtenus n'étaient pas « cellulaire-dépendant », des expériences ont été effectuées dans des cellules MEF (fibroblastes murins). Le même profil de marquage que celui obtenu dans les cellules NIH-3T3 a été obtenu après le traitement des cellules MEF par 2 µM de *F1-FITC* (Figure 44A).

D'autre part, le traitement de cellules NIH-3T3 par le fluorophore FITC seul, non couplé à un polyamide, a permis de vérifier que la FITC n'était pas à l'origine du marquage nucléaire observé. Des expériences avec des MGBs marqués par la FITC qui n'interagissent pas spécifiquement avec la séquence cible d'ADN, *in vitro*, comme dans le cas de *F2-FITC* (**Figure 44B**), ou qui ne possède pas de site d'interaction dans la séquence cible d'ADN (*ctrl*-

FITC : N-term - Py-Py-Py-Py- γ -Im- β -Py-Im - C-term ; **Figure 44C**) ont également été effectuées et aucun marquage nucléaire sous forme de foyers n'a été observé.



Figure 44 : Images de cellules murines vivantes traitées par des sondes couplées à la FITC

Images de cellules murines MEF (A) ou NIH-3T3 (B et C) vivantes traitées par 2 μ M de F1-FITC (A) ou F2-FITC (B) ou ctrl-FITC (C) pendant 24h. 1 et 3 : les images de fluorescence, 2 et 4 : la superposition des images de F1-FITC (A/2 et 4) ou F2-FITC (B/2) ou ctrl-FITC (C/2) en rouge et de la lumière blanche en gris. La sonde F2-FITC (B) a été utilisée comme contrôle négatif car le MGB murin standard F2 n'interagit pas, in vitro, avec la séquence cible d'ADN et la sonde ctrl-FITC (N-term - Py-Py-Py-Py-Py-Im- β -Py-Im - Cterm) n'a pas de site de fixation sur la séquence cible murine. Barre d'échelle : 10 μ m (voir Chapitre IV - Matériels et méthodes pour les conditions détaillées des expériences).

A.3.2 Cellules murines fixées

Dans le but de valider la localisation spécifique des sondes fluorescentes au niveau des séquences d'ADN répétées péricentromériques murines, des expériences de colocalisation entre la sonde *F1-FITC* et des marqueurs connus de ces régions ont été effectuées sur des cellules NIH-3T3 fixées (les détails des expériences sont donnés dans le **Chapitre IV-Matériels et méthodes**). Le marquage des cellules fixées par 2 μ M de *F1-FITC* pendant 30 min révèle un profil de marquage ponctuel (images de gauche, **Figure 45**) similaire à celui observé dans les cellules vivantes. Un marquage additionnel avec le DAPI, qui permet de visualiser les chromocentres murins compte tenu de la richesse de ces séquences en paires A/T, montre une superposition importante des signaux de *F1-FITC* et du DAPI (**Figure 45A**). 87% du signal de *F1-FITC* colocalise avec le signal du DAPI sur n = 19 noyaux (coefficient de Pearson = 0,82). Les mêmes résultats ont été obtenus avec l'histone H₃K₉me₃ (**Figure 45B**) qui est enrichie au niveau des chromocentres murins (Rice et al. 2003) (pour rappel voir § Chapitre I – B.2.1.1). 97% du signal de *F1-FITC* colocalise avec le signal de H₃K₉me₃ sur n = 16 noyaux (coefficient de Pearson = 0,88).

La localisation de *F1-FITC* au niveau des séquences cibles a également été vérifiée par des expériences de FISH en utilisant une sonde ADN / LNA (oligonucléotide « Sat Maj » ciblant une séquence d'ADN péricentromérique murine, la séquence de cette sonde est donnée dans le **Chapitre IV – Matériels et méthodes**) (**Figure 45C**) et la superposition des deux signaux valide la bonne localisation de la sonde *F1-FITC*. 98% du signal de « Sat Maj » colocalise avec le signal de *F1-FITC* sur n = 10 noyaux (coefficient de Pearson = 0,65). Le signal fluorescent obtenu avec la sonde *F1-FITC* est donc majoritairement détecté au niveau des régions péricentromériques, ce qui indique qu'elle lie spécifiquement les séquences d'ADN péricentromériques murines.

En revanche, le signal fluorescent de la protéine centromérique CenpB entoure les foyers marqués par « Sat Maj » (et donc *F1-FITC*), et le profil du marquage ne se superpose pas à ceux décrits précédemment ; ce qui prouve que *F1-FITC* serait une sonde majoritairement péricentromérique (coefficient de Pearson = 0,28).

Le rapport signal / bruit de fond est aussi en faveur de *F1-FITC* ; ce qui est en accord avec les résultats de spectroscopie de fluorescence (voir § Chapitre II - A.2.5.1).



Figure 45 : Expériences de colocalisation sur des cellules murines NIH-3T3 fixées

Images de cellules NIH-3T3 fixées traitées par 2 µM de F1-FITC pendant 30 min et les autres marqueurs connus des chromocentres tels que le DAPI (A) ou la marque $H_3K_9me_3$ (B) ou « Sat maj » (C) et des centromères avec CenpB (C). A/ Expérience de marquage fluorescent de la sonde F1-FITC (à gauche) et du DAPI (au milieu). À droite, la superposition des signaux de F1-FITC (en vert) et du DAPI (en rouge). Les profils de l'intensité de ces signaux sont donnés dans l'encadré de droite. B/ Expériences de marquage fluorescent de la sonde F1-FITC (à gauche) et d'immunomarquage fluorescent de $H_3K_9me_3$ (au milieu). À droite, la superposition des signaux de **F1-FITC** (en vert) et de $H_3K_9me_3$ (en rouge). Les profils de l'intensité de ces signaux sont donnés dans l'encadré de droite. C/ Expériences de marquage fluorescent par F1-FITC (à gauche) et d'immunoFISH avec la sonde « Sat Maj » oligonucléotidique péricentromérique contenant des LNAs (au milieu) et de la protéine centromérique CenpB (à droite). En-dessous, à gauche la superposition des signaux de F1-FITC (en vert) et de « Sat maj » (en rouge) et à droite, celle des signaux de « Sat Maj » (en rouge) et de CenpB (en blanc). Les profils de l'intensité de ces signaux sont donnés dans les encadrés du bas (les détails des expériences sont donnés dans le Chapitre IV - Matériels et méthodes).

Dans les cellules fixées, des profils de marquage similaires à celui de *F1-FITC* ont été obtenus avec les sondes *F4β-FITC*, *F4β-fluo*, *F1-Cy3*, *F4β-Cy5*, *F1\gamma^{(R)NH2}-MM14* et *T1-*

FITC. De plus, en utilisant les mêmes sondes contrôles que pour les études sur cellules vivantes, aucun marquage ponctuel n'a pu être observé (**Figure 46**).



Figure 46 : Images de cellules NIH-3T3 fixées traitées par le DAPI et la sonde murine contrôle F2-FITC

Images de cellules murines NIH-3T3 fixées traitées par 2 μ M de **F2-FITC** pendant 30 min et du DAPI. À gauche, le marquage fluorescent de la sonde contrôle **F2-FITC**, au milieu celui du DAPI et à droite, la superposition des signaux de **F2-FITC** (en vert) et du DAPI (en bleu). Les profils de l'intensité de ces signaux sont donnés dans l'encadré de droite (les détails des expériences sont donnés dans le **Chapitre IV – Matériels et méthodes**).

A.3.3 Conclusions des expériences cellulaires

L'interaction de sondes fluorescentes composées par des polyamides couplés à des fluorophores avec l'ADN dans un contexte génomique a été vérifiée sur des cellules murines NIH-3T3 et MEF fixées et vivantes. Les études sur cellules fixées ont permis de valider la localisation des sondes testées au niveau des régions péricentromériques murines ; en effet, un marquage ponctuel typique des chromocentres a été observé. Ce profil est superposable à ceux observés en présence des marqueurs connus des chromocentres tels que le DAPI, la marque $H_3K_9me_3$ et une sonde LNA « Sat Maj ». En ce qui concerne la détection de la protéine centromériques observés. Des profils de marquage similaires ont été obtenus avec les sondes *F1-FITC*, *F4\beta-FITC* et *F1-OG-ITC* après le traitement de cellules NIH-3T3 vivantes. Ainsi, trois sondes fluorescentes composées par des polyamides ont permis la détection spécifique des structures nucléaires formées par les séquences d'ADN répétées péricentromériques murines dans des cellules vivantes. Cependant, un problème d'extinction rapide du signal de fluorescence persiste et reste à améliorer pour étudier plus en détails la dynamique de ces structures.

A.4 Conclusions générales des résultats obtenus avec le modèle murin

Dans ces travaux sur le modèle murin, nous avons étudié de manière approfondie l'interaction entre une série de polyamides et de leur conjugués fluorescents avec une séquence d'ADN qui est représentative des régions péricentromériques des chromosomes murins. Cette étude a tout d'abord impliqué des méthodes physico-chimiques pour optimiser la conception de sondes fluorescentes à bases de polyamides. Afin d'améliorer la spécificité en ajustant au mieux la longueur de la séquence reconnue par les polyamides, une nouvelle méthode de synthèse de tandems a été élaborée en utilisant une réaction de « chimie-click » afin de construire simplement un tandem « tête à tête » de MGBs standard.

Nous avons montré que la formation des complexes ADN / polyamide est un processus rapide et que la cinétique de formation des complexes n'est donc pas un facteur critique. Une conclusion importante de ces expériences est que la conception de sondes fluorescentes à base de polyamides en respectant les règles de reconnaissances établies est nécessaire mais pas suffisante pour obtenir des sondes efficaces, affines et spécifiques. Les propriétés de l'interaction entre l'ADN et des MGBs très similaires ciblant des séquences voisines sur le duplexe cible d'ADN sont différentes en terme d'affinité et semblent reposer sur des facteurs structuraux qui ne sont pas totalement compris à ce jour. À titre d'exemple, nous avons vu que l'insertion d'un synthon β -ala à la place d'un Py suivant un Im a un effet positif sur l'affinité du MGB *F4* β . Cependant le couplage d'un fluorophore en N-term d'un polyamide affecte significativement l'affinité des polyamides pour leur séquence cible d'ADN. Ces résultats nous démontrent qu'il est important de valider, *in vitro*, l'interaction de chaque sonde avec sa séquence cible d'ADN par des méthodes physico-chimiques avant d'envisager des applications cellulaires.

En utilisant les techniques spectroscopiques de dénaturation thermique et de dichroïsme circulaire, nous avons observé une faible interaction non-spécifique entre les polyamides standard et un duplexe contrôle, qui ne possède pas de sites d'interaction. Toutefois, nous avons montré que les tandems « tête à tête » composés par deux polyamides unis par un bras de liaison flexible sont une bonne alternative pour augmenter la spécificité (exemple du tandem T2). L'inconvénient de ce type de molécule vient de l'interaction

indépendante de ses deux composants avec leur séquence cible d'ADN, mais avec une affinité plus faible.

Finalement, les MGBs avec les meilleures affinités et spécificités ont été couplés avec différents fluorophores, et des études de spectroscopie de fluorescence ont permis d'observer un faible effet FRET non optimal, mais surtout une augmentation de l'intensité de fluorescence des sondes lors de leur interaction spécifique avec le duplexe cible d'ADN. Par la suite, l'étude de ces sondes dans des cellules murines fixées et vivantes ont permis de mettre en évidence le rôle crucial du choix du fluorophore utilisé en ce qui concerne la pénétration cellulaire. La détection spécifique des chromocentres via les séquences d'ADN répétées péricentromériques murines a été possible dans des cellules fixées avec l'ensemble des sondes testées (sauf avec les sondes contrôles). Et plus important, trois sondes ont aussi permis de visualiser les séquences cibles directement dans des cellules murines vivantes. Ainsi, ce travail a permis de montrer, pour la première fois, que la stratégie d'avoir recours à des polyamides standard, ligands de l'ADN péricentromérique, couplés à des fluorophores peut être utilisée pour la visualisation de structures nucléaires formées par les séquences répétées dans des cellules vivantes.

B - Les résultats obtenus avec le modèle humain

Chez l'homme, le choix de la séquence cible s'est tourné vers la répétition centromérique α -satellite de 171 pb qui est la plus abondante puisqu'elle est présente en millions de copies sur les régions centromériques de tous les chromosomes humains (Wu & Manuelidis 1980). Les objectifs étaient de vérifier si dans les cellules humaines vivantes, les séquences α -satellites centromériques étaient observables en utilisant des sondes fluorescentes à base de polyamides et si d'éventuelles structures s'apparentant aux chromocentres murins étaient identifiables. Dans ce but et en tenant compte des résultats obtenus avec le modèle murin, deux stratégies ont été suivies : la synthèse de polyamides standard reconnaissant des séquences d'ADN de 5 - 6 pb et la synthèse de polyamides longs composés par 6 à 8 paires de monomères (pour rappel voir § Chapitre I - D.3.2.2).

B.1 Synthèses et caractérisations des MGBs et des sondes fluorescentes humains

B.1.1 Les MGBs humains

Le *Duplexe* 7 de 26 pb, issu d'une répétition α -satellite humaine a été utilisé comme cible modèle pour la synthèse des polyamides humains et pour les études physico-chimiques *in vitro* (**Figure 47**).

Les polyamides standard H1, Hb1 et Hbb1 ont été construits de leur extrémité N-term à leur extrémité C-term pour reconnaître le duplexe cible dans le sens $5' \rightarrow 3'$. Afin d'augmenter la spécificité pour leur séquence cible d'ADN, des polyamides longs de 6 à 8 paires de monomères ont aussi été synthétisés (Hb2, Hbb2, Hb3, Hbb3, Hb4 et Hbb4) (**Figure 47**).

La synthèse de ces polyamides a été effectuée comme précédemment (voir § Chapitre II - A.1) pour le modèle souris en utilisant une résine Boc- β -ala-PAM selon le protocole décrit par Krutzik et Chamberlin (Krutzik & Chamberlin 2002a; Krutzik & Chamberlin 2002b). Les détails de la synthèse sont donnés dans la **Figure 28** et dans le **Chapitre IV** (**Matériels et méthodes**).



Figure 47 : Représentation schématique des deux duplexes utilisés et des polyamides humains synthétisés

Le **Duplexe 7** est le duplexe cible modèle de 26 pb issu d'une répétition centromérique α satellite humaine. Il possède un site d'interaction pour chaque MGB (le site de liaison pour les polyamides **H1**, **Hb1** ou **Hbb1** est encadré). Les polyamides synthétisés pour cibler cette séquence sont alignés relativement à leurs sites de liaison sur le **Duplexe 7**. Le **Duplexe 3** utilisé comme duplexe contrôle ne contient aucun site de liaison pour chacun des polyamides humains.

En ce qui concerne la synthèse des polyamides longs, le remplacement d'un Py qui doit suivre un Im par un synthon β -ala a été effectué. En effet, les polyamides standard sont composés par 4 paires de monomères Py / Im pour des questions de rigidité structurale due à la présence de monomère(s) Im. Pour compenser cette rigidité structurale, des monomères β -ala ont été ajoutés à la place des Py. Ces changements n'affectent pas la reconnaissance des paires de base de l'ADN. La nature flexible du synthon β -ala permet essentiellement au polyamide de se repositionner pour suivre la courbure et le pas de l'hélice d'ADN lors de l'interaction. Les résultats des études physico-chimiques avec le MGB murin *F4* β ont montré

que ce monomère pouvait aussi améliorer l'affinité du polyamide pour sa séquence cible d'ADN (voir II - A.2).

Ainsi, deux séries de polyamides ont été conçues : la première composée par les MGBs *H1* (sans β -ala), *Hb1*, *Hb2*, *Hb3* et *Hb4* dont le nombre de monomères β -ala varie mais dont la position sur la partie C-term des MGBs est constante ; et la deuxième composée par les MGBs *Hbb1*, *Hbb2*, *Hbb3* et *Hbb4* dont le nombre et la position des synthons β -ala varient (**Figure 47**). Les formules et les caractéristiques chimiques des différents polyamides humains synthétisés sont données en **Annexe 15**. Lors de la synthèse des MGBs longs, des difficultés ont été rencontrées qui sont probablement liées à la capacité limitée de la résine utilisée. Pour les polyamides *Hb3* et *Hbb3*, la synthèse a pu être réalisée en augmentant les temps et les volumes de couplage. Cependant pour *Hb4* et *Hbb4*, malgré plusieurs tentatives notamment en utilisant des trimères de Py, préparés préalablement, afin de limiter le nombre d'étapes de couplage, les polyamides ont toujours été obtenus, avec des rendements très faibles, contaminés par d'autres produits inséparables par HPLC. De ce fait, les études de l'interaction entre l'ADN et les MGBs n'ont pas pu être poursuivies avec *Hb4* et *Hbb4*.

Après leur synthèse, les polyamides ont été analysés par HPLC analytique en phase inverse, caractérisés par spectrométrie de masse (SM) et par spectroscopie UV-visible et enfin, purifiés par HPLC préparative en phase inverse (l'exemple de la caractérisation du MGB humain *Hb1* est présenté dans l'**Annexe 16**).

B.1.2 Les sondes fluorescentes humaines

Compte tenu des résultats obtenus avec le modèle murin, le fluorophore utilisé pour la construction des sondes fluorescentes humaines a été la FITC. En utilisant, d'une part l'amine terminale des MGBs et d'autre part le fluorophore sous forme d'isothiocyanate, la formation d'une liaison thiourée a permis la construction d'une sonde fluorescente polyamide-fluorophore (voir pour rappel la **Figure 30**). Un tableau récapitulant les caractéristiques chimiques des différentes sondes fluorescentes humaines est donné en **Annexe 17** (à titre d'exemple, la caractérisation de la sonde *Hb1-FITC* est donnée en **Annexe 18**).

B.2 Études physico-chimiques de l'interaction des MGBs et des sondes fluorescentes humains avec l'ADN *in vitro*

B.2.1 Dénaturation thermique de l'ADN suivie par spectroscopie UV

B.2.1.1 Dénaturation thermique des duplexes d'ADN en présence des MGBs humains

Comme pour le modèle murin, l'interaction entre l'ADN et les polyamides a été étudiée par la méthode de dénaturation thermique suivie par spectroscopie UV. Les profils typiques des courbes de dénaturation thermique du *Duplexe* 7 (duplexe cible qui possède un site de liaison pour chacun des MGBs humains) seul et en présence d'un excès molaire (6 équivalents) d'un polyamide, *Hbb1*, sont donnés dans la **Figure 48A**.

La formation d'un complexe avec le MGB conduit à une forte stabilisation du duplexe : sa température de fusion passe de 64°C (courbes bleues, **Figures 48A** et **B**) à 78°C en présence de *Hbb1* (courbes rouges, **Figures 48A** et **B**) ($\Delta T_m = 14^{\circ}C \pm 0,5^{\circ}C$). Les résultats obtenus avec les différents MGBs humains sont donnés dans le **Tableau 8**. Tous les polyamides synthétisés stabilisent le *Duplexe 7* mais à des degrés différents. Les meilleures stabilisations ont été obtenues avec les polyamides *Hb1* et *Hb2* ($\Delta T_m \ge 20^{\circ}C$) qui possèdent des synthons β -ala uniquement sur la partie C-term des polyamides.

Tableau 8 : Augmentations de la température de dissociation des duplexes cible
(Duplexe 7) et contrôle (Duplexe 3) en présence de 6 équivalents des différents
polyamides humains déterminées par dénaturation thermique suivie par UV ($\Delta T_m \pm$
0,5°C)

Nom	ΔT_m (°C) du Duplexe 7	ΔT_m (°C) du Duplexe 3
Hb1	$\geq 20*$	4
Hb2	$\geq 20*$	4,5
Hbb2	16	3
Hb3	15	2
Hbb1	14	3
H1	12,3	2
Hbb3	12	1

*Le spectrophotomètre utilisé pour ces expériences ne permettait pas d'enregistrer des températures supérieures à 83°C.

Cette interaction est plutôt sélective ; toutefois, en présence des différents MGBs testés, une faible stabilisation du duplexe contrôle *Duplexe 3*, qui n'a pas de site de liaison pour les MGBs ($\Delta T_m \leq 5^{\circ}C$), a été observée (exemple du MGB *Hbb1* ($\Delta T_m = 3^{\circ}C \pm 0,5^{\circ}C$) ; Figures 48C et D).



Figure 48 : Courbes de fusion des duplexes d'ADN et leurs dérivées premières en présence et en absence du polyamide humain Hbb1

Les expériences de dénaturation thermique ont été effectuées dans un tampon cacodylate de sodium (50 mM) à pH 6,2 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. A/ Courbes de fusion du **Duplexe 7** seul (1,3 μ M) à 260 nm (courbe bleue) et en présence de 6 équivalents de **Hbb1** (7,8 μ M) à 260 nm (courbe rouge) et à 335 nm (courbe verte) et B/ leurs dérivées premières. C/ Courbes de fusion du duplexe contrôle (**Duplexe 3**) seul (1,3 μ M) à 260 nm (courbe bleue) et en présence de 6 équivalents de **Hbb1** (7,8 μ M) à 260 nm (courbe touge) et à 335 nm (courbe verte) et D/ leurs dérivées premières. Les détails des expériences sont donnés dans le **Chapitre IV – Matériels et méthodes**.

En comparaison avec *Hb1*, le ΔT_m obtenu avec le MGB humain standard *H1*, qui ne diffère que par l'absence d'un synthon β -ala est nettement inférieur (12°C). Ce résultat coïncide avec les résultats obtenus avec le modèle murin : l'insertion d'un synthon β -ala peut améliorer les propriétés de l'interaction ADN / MGB. De plus, en comparant *Hb1* avec *Hbb1* qui possède un monomère β -ala supplémentaire sur la partie N-term du polyamide, la position et / ou le nombre des synthons β -ala semble aussi influencer les propriétés de l'interaction.

B.2.1.2 Dénaturation thermique des duplexes d'ADN en présence des sondes fluorescentes humaines

L'interaction entre l'ADN et les sondes fluorescentes composées par les polyamides humains couplés à la FITC a également été étudiée par dénaturation thermique des complexes ADN / sonde. Des profils typiques des courbes de dénaturation thermique du *Duplexe* 7 (duplexe cible) ou du *Duplexe* 3 (duplexe contrôle) seuls et en présence d'un excès molaire (6 équivalents) de la sonde *Hbb1-FITC* ont été enregistrés (Figure 49). En présence de *Hbb1-FITC*, la T_m du *Duplexe* 7 passe de 64°C à 68,5°C (Δ T_m = 4,5°C ± 0,5°C, Figures 49 A et B) alors qu'en présence du *Duplexe* 3, le Δ T_m est ≤1°C (± 0,5°C, Figures 49 C et D). Les Δ T_m obtenus avec les différentes sondes humaines sont donnés dans le Tableau 9 et ils sont inférieurs (environ 10°C de différence) à ceux obtenus avec les MGBs non couplés. Ces résultats sont similaires à ceux obtenus avec le modèle murin.

humaines déterminées par dénaturation thermique suivie par UV ($\Delta T_m \pm 0.5^{\circ}C$)		
Nom	ΔT_m (°C) du Duplexe 7	ΔT_m (°C) du Duplexe 3
Hb2-FITC	9	≤1
Hb1-FITC	5,5	≤ 1
Hbb1-FITC	4,5	≤ 1
Hb3-FITC	4	≤ 1
Hbb2-FITC	3,5	≤ 1
H1-FITC	2,2	≤ 1
Hbb3-FITC	2	≤ 1

Tableau 9 : Augmentations de la température de dissociation des duplexes cible (Duplexe 7) et contrôle (Duplexe 3) en présence de 6 équivalents des différentes sondes humaines déterminées par dénaturation thermique suivie par UV ($\Delta T_m \pm 0.5^{\circ}$ C)

Ainsi, le couplage d'un fluorophore aux MGBs, quel qu'il soit, mène à une déstabilisation importante du complexe ADN / MGB. Cette déstabilisation peut être due, au moins en partie, au masquage du groupement amine (qui peut être protoné) qui est utilisé pour le couplage du fluorophore.

La sonde humaine qui stabilise le mieux le duplexe cible (*Duplexe 7*) est la sonde *Hb2-FITC* ($\Delta T_m = 9^{\circ}C \pm 0.5^{\circ}C$).



Figure 49 : Courbes de fusion des duplexes d'ADN et leurs dérivées premières en présence et en absence de la sonde humaine Hbb1-FITC

Les expériences de dénaturation thermique ont été effectuées dans un tampon cacodylate de sodium (50 mM) à pH 6,2 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. A/ Courbes de fusion du **Duplexe 7** seul (1,3 μ M) à 260 nm (courbe bleue) et en présence de 6 équivalents de **Hbb1-FITC** (7,8 μ M) à 260 nm (courbe rouge) et à 335 nm (courbe verte) et B/ leurs dérivées premières. C/ Courbes de fusion du duplexe contrôle (**Duplexe 3**) seul (1,3 μ M) à 260 nm (courbe rouge) et à 335 nm (courbe bleue) et en présence de 6 équivalents de **Hbb1-FITC** (7,8 μ M) à 260 nm (courbe verte) et D/ leurs dérivées premières. Les détails des expériences sont donnés dans le **Chapitre IV – Matériels et méthodes**.

B.2.1.3 Conclusions des expériences de dénaturation thermique

Pour conclure sur ces expériences de dénaturation thermique, les MGBs humains standard et longs sont capables d'interagir spécifiquement avec leur duplexe cible d'ADN en augmentant leur température de fusion. Cependant, ces augmentations varient fortement ; ce qui signifie que des petites variations structurales sur les MGBs peuvent entraîner un changement de leurs propriétés de fixation à l'ADN. Les deux polyamides *Hb1* et *Hb2* stabilisent le mieux le duplexe cible d'ADN ($\Delta T_m \ge 20^{\circ}C \pm 0.5^{\circ}C$). En comparant les différentes structures des MGBs, l'introduction de synthon(s) β -ala dans la partie C-term du polyamide (*Hb1*, *Hb2* et *Hb3*) semble favorable, alors que l'addition d'un synthon β -ala supplémentaire dans la partie N-term (*Hbb1*, *Hbb2* et *Hbb3*) semble défavorable. La faible stabilisation du duplexe contrôle a permis de valider la sélectivité de ces molécules. Enfin, l'interaction de l'ADN avec les sondes fluorescentes (polyamide + fluorophore) a été vérifiée. Elle est cependant fortement altérée, probablement en partie à cause du masquage de la fonction amine utilisée pour le couplage des fluorophores.

B.2.2 Électrophorèses en conditions non dénaturantes (« Gel-retard »)

B.2.2.1 « Gels-retard » des duplexes d'ADN en présence des MGBs humains

Tout d'abord, aucune interaction entre les MGBs humains et le *Duplexe 3* (contrôle qui ne possède aucun site de liaison pour les MGBs) n'a pu être observée par « gel-retard », confirmant encore la sélectivité de ces molécules.

En dépit de leur parfaite correspondance aux règles de reconnaissance de Dervan, les polyamides synthétisés ont montré des différences concernant leurs propriétés d'interaction avec l'ADN cible. Après l'incubation du *Duplexe 7* avec les MGBs humains *H1*, *Hb1*, *Hbb1*, *Hbb2* et *Hbb3*, la migration des échantillons sur gel a révélé l'apparition d'une bande retardée (**Figure 50**) dont l'intensité augmente en fonction de la concentration en polyamides, prouvant la formation d'un complexe ADN / MGB. Les constantes de dissociation apparentes (K_{d,app}) ont pu être estimées et sont données dans le **Tableau 10**.

En revanche avec les ligands *Hb2* et *Hb3*, l'apparition d'une deuxième bande retardée suggère la formation d'un complexe ADN / MGB de structure différente (**Figure 50**). En effet, un second mécanisme d'interaction avec l'ADN, représenté sur la **Figure 51**, est envisageable. Deux molécules de MGBs dans une conformation ouverte pourraient s'insérer dans le duplexe d'ADN. De plus, en comparant *Hb2* et *Hbb2* ou *Hb3* et *Hbb3*, la seule différence repose sur la position des synthons β -ala ; ainsi *Hb2* et *Hb3* pourraient être «plus flexibles » puisque tous les synthons β -ala sont situés du même côté de « l'épingle à cheveux » ; ce qui favoriserait peut être son ouverture.



Figure 50 : Images des « gels-retard » des complexes ADN / MGBs humains

Les électrophorèses ont été effectuées sur des gels de polyacrylamide à 20% non dénaturant dans du tampon TBE à 4°C. Le **Duplexe 7** (qui possède un seul site de liaison pour chacun des MGBs) a été incubé en présence des différents MGBs humains. Le brin riche en T du duplexe est marqué à la fluorescéine. Les concentrations des polyamides sont indiquées audessus de chaque piste des gels et celle du **Duplexe 7** est indiquée à gauche en gris. Les schémas sur la droite symbolisent la position du duplexe (deux traits noirs) et des complexes ADN / MGB (deux traits noirs avec un ou deux points noirs symbolisant les MGBs) après la migration. Les bandes retardées détectées prouvent l'existence de complexes ADN / MGB. Les détails des expériences sont donnés dans le **Chapitre IV – Matériels et méthodes**.

Toutefois, c'est le MGB *Hb2* qui possède la meilleure affinité ($K_{d,app} = 0,07 \ \mu M \pm 0,004 \ \mu M$), et il est remarquable que c'est aussi un des MGBs qui stabilise le mieux le duplexe d'ADN ($\Delta T_m \ge 20^{\circ}C \pm 0,5 \ ^{\circ}C$). Il est également notable que les longs MGBs (*Hb2*, *Hbb2*, *Hb3* et *Hbb3*) sont plus affins que les MGBs standard (*H1*, *Hb1* et *Hbb1*) pour leurs séquences cibles d'ADN alors qu'ils ne stabilisent pas plus le duplexe par dénaturation thermique (Tableau 10). Le chauffage des longs MGBs entraîne probablement leur déstructuration et, du fait de leur longueur, ces MGBs mettraient plus longtemps à se conformer de manière à pouvoir interagir avec l'ADN double brin. En comparant deux à deux les structures des MGBs (*Hb1 / Hbb1* ou *Hb2 / Hbb2* ou *Hb3 / Hbb3*), ceux qui possèdent les synthons β -ala dans leur partie C-term (*Hb1*, *Hb2* et *Hb3*) ont toujours une meilleure affinité

et stabilisent mieux le duplexe dans les expériences de dénaturation thermique par rapport à leur homologue qui, soit ne possède pas de β -ala (*H1*), soit possède aussi un β -ala dans la partie N-term.

Tableau 10 : Constantes de dissociation apparentes (K_{d,app}) des complexes entre le Duplexe 7 et les différents polyamides humains, déterminées par des expériences de retard sur gel de polyacrylamide non dénaturant à 20%, réalisées à 4°C dans du tampon TBE. Les augmentations de la température de fusion du Duplexe 7 en présence de 6 équivalents des MGBs sont données pour rappel ($\Delta T_m \pm 0,5^{\circ}C$) (voir § Chapitre II -B.2.1.1)

Polyamide	$K_{d,app}\left(\mu M\right)$	ΔT_m (°C)
Hb2	$0,07 \pm 0,004*$	≥ 20
Hbb2	$0,085 \pm 0,005$	16
Hb3	$0,09 \pm 0,01*$	15
Hbb3	$0,11 \pm 0,02$	12
Hb1	$0,\!17\pm0,\!02$	≥20
Hbb1	$0,\!38\pm0,\!01$	14
H1	$0,\!43 \pm 0,\!015$	12,3

*Les K_{d,app} ont été estimés par rapport à la première bande retardée.

Duplexe 7 + Hb2 en "épingle à cheveux"





Le **Duplexe 7** est le duplexe cible modèle de 26 pb issu de la répétition centromérique humaine α -satellite. Il possède un site d'interaction pour le MGB humain Hb2. En haut, l'interaction entre le **Duplexe 7** et le MGB Hb2 dans une conformation « d'épingle à cheveux » antiparallèle selon les règles de reconnaissance établies. En bas, l'interaction hypothétique de deux MGBs Hb2 dans une conformation ouverte. Enfin, comme pour le modèle murin, l'interaction des sondes fluorescentes humaines avec l'ADN n'a pas pu être vérifiée probablement en raison de la dissociation des complexes dans le gel.

B.2.2.2 Conclusions des expériences de « gel-retard »

Pour conclure sur ces expériences de « gel-retard », les MGBs humains standard et longs sont capables de former des complexes avec leur duplexe cible d'ADN avec des affinités micromolaires ou submicromolaires. Cependant, les affinités de ces polyamides pour leur séquence cible varient beaucoup. Le polyamide long Hb2 est le plus affin, dans des expériences de gel-retard, pour le duplexe cible d'ADN ($K_{d,app} = 0.07 \ \mu M \pm 0.004 \ \mu M$); ce résultat est en accord avec celui obtenu par dénaturation thermique (ΔT_m , Hb2 $\geq 20^{\circ}C \pm$ 0,5°C). Les longs MGBs (Hb2, Hbb2, Hb3 et Hbb3) sont plus affins, dans des expériences de gel-retard, que les MGBs standard (H1, Hb1 et Hbb1) pour leurs séquences cibles d'ADN, alors qu'ils ne conduisent pas à une augmentation plus importante de la température de fusion du duplexe. De plus, les MGBs qui possèdent les synthons β -ala dans leur partie C-term (*Hb1*, Hb2 et Hb3) ont toujours une meilleure affinité et stabilisent le mieux le duplexe dans les expériences de dénaturation thermique par rapport à leur homologue qui, soit ne possède pas de β -ala (*H1*), soit possède aussi un β -ala dans leur partie N-term. Le nombre et la position de synthon(s) β-ala semble être un facteur important pour la stabilité des complexes ADN / MGB. L'apparition d'une deuxième bande retardée avec les MGBs Hb2 et Hb3 suggère la formation d'un complexe ADN : MGB avec une stœchiométrie (1 : 2). L'absence de bande retardée sur les gels avec le duplexe contrôle a aussi permis de valider la sélectivité de ces molécules. Enfin, l'interaction de l'ADN avec les sondes fluorescentes n'a pas pu être mise en évidence par cette technique.

B.2.3 Dichroïsme circulaire

B.2.3.1 Interaction ADN / MGBs humains

Suite aux résultats obtenus par les expériences de « gel-retard », l'hypothèse d'un mode d'interaction particulier entre les MGBs *Hb2* ou *Hb3* et le *Duplexe* 7 a été étudiée par dichroïsme circulaire. Les expériences ont été conduites en titrant une solution de duplexe cible d'ADN (*Duplexe* 7 qui possède un site d'interaction pour chaque MGB) par des solutions des différents MGBs humains synthétisés. Dans tous les cas, les MGBs se lient à
leur ADN cible et un spectre de DC induit a été obtenu. En effet, les polyamides en solution n'ont pas de spectres de DC et c'est au contact de l'ADN qu'ils se structurent et qu'un signal de DC induit apparait. Les spectres différentiels de DC des polyamides *H1*, *Hb1*, *Hb2* et *Hb3* en complexe avec le *Duplexe* 7 sont donnés dans les **Figures 52A**, **B**, **E** et **F**, respectivement. Ils présentent tous une large bande entre 300 et 400 nm qui révèle une interaction des polyamides avec le *Duplexe* 7 (pour rappel voir § Chapitre II - A.2.4). En comparaison avec les spectres modèles établis par Gursky (Surovaya et al. 2001), les spectres de DC induits ressemblent à la signature d'un positionnement en « épingle à cheveux » antiparallèle au sein du petit sillon du duplexe cible d'ADN.

La différence de forme des spectres entre *H1* (Figure 52A) et *Hb1* (Figure 52B) repose probablement sur la perte d'un hétérocycle dans le dérivé *Hb1* ou à l'introduction du synthon β -ala qui peut changer légèrement l'environnement du complexe ADN / MGB dans le petit sillon. En ce qui concerne les autres MGBs, leurs spectres de DC induits sont similaires à ceux obtenus avec *Hb1*, à l'exception de ceux de *Hb2* et *Hb3*. Pour *Hb2* (Figure 52E) et *Hb3* (Figure 52F), la seule différence avec les spectres de *Hb1* (Figure 52B), est la présence d'une double bosse entre 320 et 360 nm. Cette observation conforte l'idée d'une interaction dans une conformation légèrement différente mais ne permet pas de valider l'hypothèse d'une interaction de deux MGBs « ouverts » avec le duplexe d'ADN.

De plus, la saturation des courbes de titration est atteinte seulement après l'ajout d'environ 4 équivalents de chacun des MGBs (**Figures 52C**, **D**, **G** et **H**) sauf pour *Hb3* le plus long des MGBs. Toutefois, il est difficile d'établir une relation entre cette différence de stœchiométrie apparente et l'hypothèse d'une interaction dans une conformation ouverte des MGBs *Hb2* et *Hb3* avec le *Duplexe 7*.

Comme pour le modèle murin, les spectres obtenus avec le *Duplexe 3* (contrôle qui n'a aucun site d'interaction pour chacun des MGBs humains) ont une forme légèrement différente avec une amplitude plus faible ; ce qui semble indiquer une faible interaction non-spécifique des MGBs humains avec la cible d'ADN (Annexe 19). D'autre part, l'interaction des MGBs avec l'ADN ne provoque pas de changement significatif des signaux de DC de l'ADN.



Figure 52 : Spectres différentiels de dichroïsme circulaire des complexes ADN / MGB

Les expériences de dichroïsme circulaire ont été effectuées dans un tampon Tris-HCl (50 mM) à pH 7,5 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. Spectres de DC différentiels des complexes ADN : MGB, A/ Duplexe 7 : H1, B/ Duplexe 7 : Hb1, E/ Duplexe 7 : Hb2 et F/ Duplexe 7 : Hb3. Les spectres de DC de l'ADN sont soustraits. Des solutions de duplexe d'ADN à 2,5 µM sont titrées par des solutions de polyamides à 250 µM (voir description dans le Chapitre IV – Matériels et méthodes). Les concentrations de polyamides après chaque ajout sont indiquées à droite de chaque figure. Courbes de titration : ellipticité des complexes C/ Duplexe 7 : H1, D/ Duplexe 7 : Hb1, G/ Duplexe 7 : Hb2 et H/ Duplexe 7 : Hb3, à 335 nm aux concentrations de polyamides indiquées.

B.2.3.2 Interaction ADN / sondes fluorescentes humaines

Les spectres de dichroïsme circulaire enregistrés avec les sondes fluorescentes composées par les différents polyamides humains couplés à la FITC sont similaires à ceux obtenus avec les MGBs non couplés ; ce qui suggère que le couplage à un fluorophore ne modifie pas le mode d'interaction des polyamides avec l'ADN. Cependant, les amplitudes des signaux de DC sont plus faibles que celles obtenues avec les polyamides non couplés aux fluorophores ; ce qui est en accord avec la perte d'affinité pour l'ADN cible des sondes fluorescentes en comparaison avec les MGBs non couplés et donc à la diminution de la stabilité des complexes ADN / sonde fluorescente. Plusieurs exemples de spectres de DC des sondes fluorescentes lors de leur interaction avec des duplexes d'ADN sont donnés dans l'**Annexe 20**.

B.2.3.3 Conclusions des expériences de dichroïsme circulaire

Les résultats des expériences de dichroïsme circulaire ont permis de valider les résultats obtenus par les autres techniques. De ces expériences de DC, il apparait donc que des complexes entre l'ADN cible et les différents MGBs humains synthétisés (seuls ou couplés avec un fluorophore) se forment, et les expériences avec un duplexe contrôle ne montre qu'une faible interaction non-spécifique de ces polyamides. Les différentes formes des spectres de DC suggèrent des différences subtiles dans les propriétés de l'interaction entre l'ADN et les MGBs. En comparaison avec les spectres modèles établis par Gursky, la forme des spectres obtenus apporte quelques indications sur la structure des complexes ADN / MGB. Les résultats suggèrent un positionnement dans une conformation d' « épingle à cheveux » antiparallèle des MGBs humains au sein du petit sillon de l'ADN. Cependant, aucune conclusion ne peut être faite en ce qui concerne l'interaction dans une conformation ouverte des MGBs Hb2 et Hb3.

Ainsi, l'ensemble des résultats précédents a révélé que les polyamides longs pouvaient être une bonne alternative pour la construction de sondes fluorescentes ; en effet, les affinités estimées de *Hb2*, *Hbb2*, *Hb3* et *Hbb3* sont supérieures à celles des polyamides standard *H1*, *Hb1* et *Hbb1*. De plus, l'introduction de synthon(s) β -ala du côté C-term des polyamides semble favorable (*Hb1*, *Hb2* et *Hb3*). Tous ces polyamides humains ont été couplés à la FITC afin de vérifier si une modification du spectre de fluorescence des sondes s'effectuait lors de leur interaction avec l'ADN.

B.2.4 Spectroscopie de fluorescence

B.2.4.1 Interaction ADN / sondes fluorescentes humaines

Dans cette étude, aucun changement des profils d'émission de fluorescence de la FITC après son couplage aux différents polyamides humains n'a été observé. En revanche, l'intensité de fluorescence des sondes (polyamide + fluorophore) change en présence de l'ADN cible. Une augmentation séquence-spécifique de l'intensité de fluorescence (17 fois, environ 4 fois supérieure à celle observée en présence du duplexe contrôle) a été observée lors de l'interaction de *Hb1-FITC* avec le *Duplexe* 7 (duplexe cible d'ADN) (**Figure 53** et **Tableau 11**).



Figure 53 : Spectres de fluorescence de la sonde humaine Hb1-FITC en absence et en présence des duplexes d'ADN

Les expériences de fluorescence ont été conduites dans un tampon Tris-HCl (10 mM) à pH 7,5 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. Les concentrations des duplexes d'ADN et de la sonde fluorescente sont de 2 μ M. Le spectre de fluorescence de la sonde seule est représenté en bleu, celui en présence du **Duplexe 3** (contrôle qui ne possède aucun site d'interaction pour la sonde) en rouge et celui en présence du **Duplexe 7** (cible qui possède un site d'interaction pour la sonde) en vert.

Pour les autres MGBs, l'augmentation de l'intensité de fluorescence était plus faible, mais toujours supérieure en présence du duplexe cible (*Duplexe 7*) qu'en présence du duplexe contrôle (*Duplexe 3*). L'ensemble des résultats obtenus avec les sondes fluorescentes humaines est donné dans le **Tableau 11**. La modulation des spectres d'émission de fluorescence des sondes a été observée en présence du *Duplexe 7* (cible, courbe verte, **Figure 53**), mais aussi plus faiblement avec le *Duplexe 3* (contrôle, courbe rouge, **Figure 53**). Le bruit de fond de fluorescence des sondes en présence du duplexe contrôle est sans doute dû à une interaction non-spécifique du fluorophore avec l'ADN.

Tableau 11 : Augmentations des intensités de fluorescence des sondes humaines (2 μ M) en présence des duplexes d'ADN (2 μ M) cible (Duplexe 7) et contrôle (Duplexe 3) dans un tampon Tris-HCl (10 mM) à pH 7,5 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂

Nom	Duplexe 7	Duplexe 3
Hb1-FITC	× 17	× 4,5
Hb2-FITC	× 9,5	× 3,5
Hb3-FITC	× 3	× 1,1
Hbb2-FITC	× 2,9	× 1,2
H1-FITC	× 1,9	× 1,06
Hbb3-FITC	× 1,5	× 1,06
Hbb1-FITC	× 1,4	× 1,1

En résumé, les résultats obtenus ont pu mettre en évidence une influence de la nature du MGB. Encore une fois, l'insertion de synthon(s) β -ala du côté C-term du polyamide (*Hb1*, *Hb2* et *Hb3*) semble favorable. En effet, c'est avec ces mêmes MGBs couplés à la FITC que les augmentations de l'intensité de fluorescence sont les plus fortes (**Tableau 11**). Ainsi les résultats obtenus par la spectroscopie de fluorescence sont en accord avec les résultats précédents.

B.2.4.2 Conclusions des expériences de fluorescence

Une augmentation de l'intensité de fluorescence des sondes humaines a été observée lors de leur interaction avec le duplexe cible d'ADN. Avec le duplexe contrôle, une faible augmentation a été mise en évidence qui pourrait résulter d'une interaction non-spécifique du fluorophore avec l'ADN. D'après les résultats de ces expériences, les sondes *Hb1-FITC* et *Hb2-FITC* semblent être les meilleurs candidats pour des expériences d'imagerie cellulaire.

B.3 Études de l'interaction des sondes fluorescentes humaines avec l'ADN *in cellulo*

Selon des études précédentes menées sur des cellules fixées au sein de notre laboratoire (Ollion 2014), les séquences α -satellite centromériques humaines seraient capables de s'associer pour former des structures nucléaires apparentées aux chromocentres murins. Les sondes fluorescentes, précédemment décrites dont les propriétés de reconnaissance pour une séquence modèle ont été évaluées *in vitro*, ont ensuite été testées pour leurs performances comme outils de visualisation de l'ADN centromérique humain dans des cellules en cultures.

B.3.1 Cellules vivantes

Pour réaliser ces expériences, des cellules fibroblastiques humaines RPE-1 vivantes ont été traitées pendant 24h (temps d'incubation nécessaire à l'obtention d'un bon rapport signal / bruit de fond) par 2 μ M (concentration optimale) de l'ensemble des sondes fluorescentes humaines synthétisées. Aucune perturbation de la morphologie des cellules ou du rythme de division cellulaire n'a été observée après 24h d'incubation. Le traitement par ces sondes fluorescentes n'est donc pas cytotoxique. Une brève étape de lavage par du tampon phosphate salin (PBS) a été effectuée préalablement à la prise d'image en chambre humide à 37°C. Trois types de résultats ont été obtenus et sont présentés dans la **Figure 54**.

1) Tout d'abord, les sondes *Hb1-FITC* (Figure 54A) et *Hbb1-FITC* ont conduit au même type de profil de marquage : les noyaux sont uniformément marqués sans distinction de structures nucléaires particulières. De plus, sur la Figure 54A, la présence d'une cellule vivante en division suggère que le traitement par les sondes fluorescentes ne perturbe pas le fonctionnement cellulaire et la visualisation des chromosomes montre que seulement l'ADN est marqué par ces sondes.

2) Ensuite, avec la sonde *H1-FITC* (qui ne possède pas de synthon β -ala), un faible signal nucléaire et aussi la présence de petits agrégats fluorescents ont été observés (**Figure 54B**). Ce résultat montre l'importance de l'insertion d'un synthon β -ala pour faciliter la pénétration cellulaire d'un polyamide qui possède deux synthons Im.

3) Enfin, avec les sondes fluorescentes composées par les longs polyamides (*Hb2-FITC* (Figure 54C), *Hbb2-FITC*, *Hb3-FITC* et *Hbb3-FITC*) aucune pénétration cellulaire

n'a été observée. De plus, la présence de nombreux agrégats fluorescents suggère qu'en raison de leur taille, ces sondes pourraient s'agréger ce qui empêcherait probablement le passage de la membrane cellulaire.



Figure 54 : Images de cellules RPE-1 vivantes traitées par les sondes fluorescentes humaines

Images de cellules humaines RPE-1 vivantes traitées par 2 μ M de Hb1-FITC (A) ou H1-FITC (B) ou Hb2-FITC (C) pendant 24h. 1 : les images de fluorescence et 2 : les superpositions des images de fluorescence avec Hb1-FITC (A/2) ou H1-FITC (B/2) ou Hb2-FITC (C/2), en rouge et de la lumière blanche, en gris. Barre d'échelle : 10 μ m (voir Chapitre IV - Matériels et méthodes pour les conditions détaillées des expériences).

Ces résultats semblent indiquer que dans des cellules humaines vivantes, les sondes fluorescentes, à base de polyamides, préparées au cours de ce travail ne permettent pas la détection de structures nucléaires s'apparentant à des chromocentres, comme celles observées dans le modèle murin.

B.3.2 Cellules fixées

Dans le but de vérifier la localisation potentielle des sondes fluorescentes au niveau des séquences d'ADN répétées centromériques humaines (α -satellite), des expériences de colocalisation entre la sonde *Hb1-FITC* et un marqueur de ces régions ont aussi été effectuées sur des cellules RPE-1 fixées (les détails des expériences sont donnés dans le **Chapitre IV-Matériels et méthodes**). Le traitement des cellules fixées par 2 μ M de *Hb1-FITC* pendant 30 min révèle un marquage de tous les noyaux (**Figure 55**), similaire à celui observé dans les cellules vivantes. De plus, un marquage avec le DAPI, un marqueur générique de l'ADN, montre une superposition significative des signaux de *Hb1-FITC* et du DAPI (**Figure 55**, à gauche, coefficient de Pearson = 0,986).



Figure 55 : Expériences de colocalisation sur des cellules humaines RPE-1 fixées

Images de cellules RPE-1 fixées traitées par 2 μ M de **Hb1-FITC** pendant 30 min et le DAPI (à gauche) ou la sonde « PCY » (à droite) qui est un marqueur des séquences α -satellite humaines. En haut à gauche, l'expérience de marquage fluorescent de la sonde **Hb1-FITC** et du DAPI. En bas à gauche, la superposition des signaux de **Hb1-FITC** (en vert) et du DAPI (en bleu). Les profils de l'intensité de ces signaux sont donnés dans l'encadré de gauche. En haut à droite, l'expérience de marquage fluorescent par **Hb1-FITC** et de FISH avec la sonde « PCY » oligonucléotidique centromérique composée par des LNAs. En bas à droite, la superposition des signaux de **Hb1-FITC** (en vert) et de « PCY » (en rouge). Les profils de l'intensité de ces signaux sont donnés dans l'encadré de droite (les détails des expériences et la séquence de la sonde PCY sont donnés dans le **Chapitre IV** – **Matériels et méthodes**). La localisation potentielle de *Hb1-FITC* au niveau des séquences cibles a ensuite été étudiée par des expériences de FISH en utilisant une sonde LNA (oligonucléotide « PCY » composé par des LNAs et ciblant les séquences d'ADN α -satellite centromériques humaines ; la séquence est donnée dans le **Chapitre IV – Matériels et méthodes**) (**Figure 55**, à droite). Un profil de marquage ponctuel a été obtenu avec PCY. L'absence de superposition des signaux obtenus avec *Hb1-FITC* et PCY suggère qu'avec le modèle humain, les sondes fluorescentes construites dans ce travail ne permettent pas de localiser les séquences répétées centromériques de manière sélective. En effet, le signal fluorescent obtenu avec *Hb1-FITC* est détecté dans tous les noyaux, et donc cette sonde fluorescente marque l'ADN en général mais pas de façon séquence-spécifique.

Dans les cellules fixées, des profils de marquage similaires à celui de *Hb1-FITC* ont été obtenus avec les sondes *H1-FITC*, *Hbb1-FITC*, *Hb2-FITC* (Annexe 21), *Hbb2-FITC*, *Hb3-FITC* et *Hbb3-FITC*. Dans les cellules vivantes, le faible marquage (dans le cas de *H1-FITC*) ou son absence (dans les cas de *Hb2-FITC*, *Hbb2-FITC*, *Hb3-FITC* et *Hbb3-FITC*) est dû à un problème de pénétration cellulaire. Les structures de ces sondes ne semblent pas optimales pour permettre le passage des membranes des cellules humaines.

De même, sur des cellules humaines fixées GM12878, les mêmes types de marquage ont été observés.

B.3.3 Conclusions des expériences cellulaires

L'interaction des sondes fluorescentes avec l'ADN dans un contexte génomique a été étudiée sur des cellules humaines RPE-1 fixées et vivantes. Les études sur cellules fixées n'ont pas permis de valider la localisation des sondes testées au niveau des régions centromériques humaines ; en effet, un marquage ponctuel a été observé avec un marqueur spécifique de ces régions (PCY) mais pas avec les sondes fluorescentes synthétisées. Le profil de marquage obtenu est superposable à celui du DAPI ce qui indique que les sondes marquent l'ADN uniformément. Après le traitement de cellules RPE-1 vivantes, un marquage de tous les noyaux a été obtenu avec les sondes fluorescentes composées uniquement par des polyamides standard. Cependant la présence de cellules vivantes en division indique que le traitement par ces sondes ne perturbe pas le fonctionnement cellulaire et de plus, la visualisation des chromosomes suggère que seul l'ADN est marqué. Ces sondes fluorescentes composées par des polyamides n'ont cependant pas permis la détection spécifique de structures nucléaires formées par les séquences d'ADN répétées centromériques humaines dans les cellules fixées et vivantes.

B.4 Conclusions générales des résultats obtenus avec le modèle humain

Dans ces travaux réalisés dans des modèles humains, nous avons étudié de manière approfondie l'interaction entre une série de polyamides et leurs conjugués fluorescents avec une séquence d'ADN représentative des régions centromériques des chromosomes humains. Cette étude a tout d'abord impliqué des méthodes physico-chimiques pour évaluer l'importance de modifications structurales sur l'interaction entre les polyamides et l'ADN. Afin d'améliorer la spécificité de l'interaction de ces sondes, la longueur de la séquence reconnue par les polyamides a été ajustée et la synthèse de longs polyamides a été réalisée après le développement d'une procédure de synthèse optimisée.

Nous avons montré que des complexes ADN / polyamide se forment. Cependant, nous avons observé que les propriétés de l'interaction entre l'ADN et des MGBs de tailles variables diffèrent fortement en termes d'affinité. Une conclusion importante de ces expériences physico-chimiques est que l'insertion d'un synthon β -ala dans la partie C-term des polyamides est favorable pour l'interaction avec l'ADN. À titre d'exemples, nous avons vu que le polyamide standard *H1* (sans β -ala) est moins affin et stabilise moins le duplexe cible d'ADN que *Hb1* (qui possède un synthon β -ala dans sa partie C-term). En comparant *Hb1* et *Hbb1*, il apparaît que seule l'introduction de synthon(s) β -ala du côté C-term des polyamides est favorable. La synthèse de polyamides longs semble également une bonne alternative pour concevoir des polyamides plus affins (voir *Hb2* versus *Hb1*), mais des différences dans le mode d'interaction de ces longs polyamides ont pu être observées. Ces différences restent à être étudiées en détails par des méthodes physico-chimiques et de modélisation moléculaire afin d'être mieux caractérisées.

Comme précédemment pour le modèle murin, nous avons aussi mis en évidence que le couplage d'un fluorophore en N-term des polyamides affecte significativement leur affinité pour leur séquence cible d'ADN. Ces résultats démontrent qu'il est important de valider, *in vitro*, l'interaction de chaque sonde avec sa séquence cible d'ADN par des méthodes physico-chimiques avant d'envisager des applications cellulaires.

Finalement des études de spectroscopie de fluorescence ont permis d'observer une augmentation de l'intensité de fluorescence des sondes lors de leur interaction spécifique avec le duplexe cible d'ADN.

Par la suite, les études de ces sondes dans des cellules humaines fixées et vivantes, n'ont pas permis de détecter spécifiquement les séquences d'ADN répétées centromériques humaines, contrairement aux modèles murins. En effet, aucun marquage « séquencespécifique » ni à fortiori aucune structure apparentée aux chromocentres murins n'a été visible avec ces sondes. En ce qui concerne la pénétration cellulaire, l'importance de l'insertion d'un synthon β -ala a été mise en évidence en comparant *H1* et *Hb1*. De plus avec les longs MGBs, la présence de nombreux agrégats fluorescents dans les cellules vivantes suggère qu'une interaction intermoléculaire de ces sondes serait envisageable.

Ainsi, ce travail a permis de montrer que les polyamides longs peuvent être une bonne alternative aux polyamides standard pour la construction de sondes fluorescentes. Ces sondes sont plus affines. Cependant, des optimisations de la structure du polyamide ou de la nature de la sonde restent nécessaires afin de permettre une pénétration efficace des sondes dans les cellules. De plus, le choix de la séquence cible n'est peut-être pas optimal et des études bioinformatiques approfondies restent à réaliser afin de sélectionner la séquence cible idéale. Néanmoins, un marquage de l'ADN a pu être observé avec les sondes fluorescentes et dans les cellules vivantes, la présence de cellules en division indique que le traitement par les MGBs fluorescents ne perturbe pas le fonctionnement des cellules et la possibilité de visualiser les chromosomes suggère que seul l'ADN est marqué.

Chapitre III – Conclusions, discussions et perspectives

Chapitre III - Conclusions, discussions et perspectives

<u>A – Conclusions et discussions</u>

L'ensemble de ces travaux de thèse a permis la synthèse de plusieurs polyamides (MGBs ou ligands du petit sillon de l'ADN) qui ont été utilisés, après leur couplage à différents fluorophores, comme sondes fluorescentes afin de visualiser des séquences répétées d'ADN directement dans des cellules.

La formation des complexes ADN / MGB

Les études physico-chimiques de l'interaction des polyamides avec l'ADN db par les méthodes de dénaturation thermique et de « gel-retard » ont permis de valider la formation des complexes ADN / MGB et celles de dichroïsme circulaire ont apporté des informations sur la structure probable de ces complexes. La comparaison des spectres différentiels de DC induits des MGBs en complexe avec l'ADN avec des spectres modèles suggère que les polyamides se positionnent dans le petit sillon de l'ADN dans une conformation « d'épingle à cheveux » antiparallèle. Les différentes formes des spectres de DC suggèrent des différences subtiles dans les propriétés de l'interaction entre l'ADN et les MGBs. Les résultats obtenus par ces différentes méthodes ont révélé que l'interaction entre l'ADN et des polyamides ayant des structures très semblables peut être très variable et doit être étudiée préalablement à l'utilisation de ces polyamides comme sondes fluorescentes pour des applications cellulaires.

Le synthon β-ala

Plusieurs facteurs semblent cruciaux pour améliorer l'affinité de ces molécules pour leur séquence cible d'ADN. À titre d'exemples, les études avec le modèle murin nous ont appris que des MGBs ayant deux synthons Im successifs sont peu affins pour leur séquence cible d'ADN (exemples des polyamides murins *F2* et *F3*, **Figure 27**). Aussi, la synthèse du polyamide murin *F4β* a montré l'intérêt de remplacer le monomère Py situé directement à la suite d'un Im par un synthon β-ala. En effet, l'affinité de ce dernier (*F4β*, K_{d,app} = 0,15 ± 0,01 μ M) s'est révélée bien meilleure que celle de son analogue sans β-ala (*F4*, K_{d,app} = 4,4 ± 0,06 μ M). De plus, grâce aux études sur le modèle humain, l'importance de ce synthon β-ala pour la pénétration cellulaire a été mise en évidence en comparant le MGB *H1*, qui ne pénètre pas ou très peu dans les cellules, avec ses analogues *Hb1* et *Hbb1*, qui possèdent des synthons β ala et marquent l'ensemble des noyaux. Enfin, en comparant les MGBs humains *Hb1* et *Hbb1* avec leurs analogues plus longs, une précision sur la position de ce synthon a été observée : c'est dans la partie C-term des polyamides que le synthon β -ala semble favorable (meilleure affinité et augmentation de l'intensité de fluorescence des sondes plus forte).

La spécificité des MGBs

Parmi les trois stratégies suivies pour augmenter la spécificité des polyamides, celle de synthétiser des polyamides longs semble la plus prometteuse. En effet, malgré la possibilité d'observer un effet FRET avec les paires de sondes fluorescentes murines F1-Cy3 / F2β-Cy5 ou F1-Cy3 / F4 β -Cy5, celui-ci n'était pas optimal puisque l'extinction totale de la sonde donneuse (F1-Cy3) n'a pas été possible et l'intensité maximale de l'émission de la sonde acceptrice n'a pas été atteinte. Ensuite, avec la construction des tandems de polyamides murins en utilisant la « chimie-click » de Huisgen, aucune amélioration de l'affinité n'a été visible et une interaction parasite « monodentate » des composantes individuelles de ces tandems a été observée par les différentes techniques physico-chimiques, cependant avec des affinités plus faibles. La synthèse de polyamides humains longs a été optimisée et ces polyamides ont montré de meilleures affinités que les polyamides courts standard (Hb2, Kd.app = 0,07 \pm 0,004 μ M versus *Hb1*, K_{d.app} = 0,17 \pm 0,02 μ M). Toutefois, l'apparition d'une deuxième bande retardée lors des expériences de « gel-retard » suggère une interaction dans une conformation ouverte des MGBs longs Hb2 et Hb3 avec l'ADN. Ces derniers possèdent deux synthons β-ala du même côté de « l'épingle à cheveux », ce qui pourrait en favoriser l'ouverture.

Le choix du fluorophore pour la construction de sondes fluorescentes

En parallèle des études de l'interaction ADN / MGB, le couplage des MGBs à différents fluorophores a été effectué. L'étude de l'interaction de ces sondes (MGB + fluorophore) avec l'ADN a révélé, dans tous les cas, que le couplage d'un fluorophore à un polyamide déstabilise le complexe formé avec l'ADN. En effet, dans les expériences de dénaturation thermique, des différences en ΔT_m d'environ 10°C ont été observées (*F4β*, $\Delta T_m = 18 \pm 0.5$ °C versus *F4β-FITC*, $\Delta T_m = 6 \pm 0.5$ °C). De plus, les spectres de DC induits des sondes en complexes avec l'ADN ont des amplitudes plus faibles et aucune bande retardée

n'a pu être observée par la méthode de « gel-retard ». Cette perte d'affinité des sondes par rapport aux MGBs seuls est probablement liée au masquage de la fonction amine (qui peut être protonée) lors du couplage aux fluorophores. En revanche dans des expériences de spectroscopie de fluorescence, une augmentation de l'intensité de fluorescence des sondes a été observée lors de leur interaction spécifique avec le duplexe cible d'ADN (*F1-FITC* : augmentation de 36,2 fois avec le duplexe cible et de 4 fois avec un duplexe contrôle).

Les résultats de ces études physico-chimiques ont permis de mettre en évidence les candidats les plus prometteurs pour les applications cellulaires : F1-FITC et $F4\beta$ -FITC pour le modèle murin et Hb1-FITC et Hb2-FITC pour le modèle humain.

La pénétration cellulaire et le marquage nucléaire

Des cellules murines et humaines ont été traitées par 2 µM de chacune des sondes synthétisées. Tout d'abord, seules les sondes marquées par les fluorophores FITC ou OG-ITC ont montré une capacité à pénétrer dans les cellules vivantes. La nature du fluorophore ainsi que celle de son bras de liaison au polyamide influence dramatiquement la pénétration cellulaire. Ainsi, le MGB $F4\beta$, couplé à la fluorescéine par une liaison amide ne rentre pas dans les cellules alors que F4β-FITC, dont le fluorophore est couplé par une liaison thiourée, pénètre dans les cellules. Des profils de marquage nucléaire ponctuels ont été observés dans des cellules murines vivantes avec les sondes F1-FITC, F1-OG-ITC et F4 β -FITC, et des profils de marquage similaires ont été observés dans des cellules fixées. Par la suite, des expériences de colocalisation ont été réalisées pour valider la localisation des sondes fluorescentes au niveau de leur séquence cible d'ADN, c'est-à-dire au niveau des régions péricentromériques murines qui sont capables de s'associer pour former des structures nucléaires appelées « les chromocentres ». Les signaux fluorescents du DAPI, de l'oligonucléotide « Sat Maj » et de la marque H₃K₉me₃, connus pour marquer les chromocentres murins, coïncident bien avec ceux obtenus avec les sondes MGBs. En revanche, le signal fluorescent de la protéine centromérique CenpB entoure les foyers marqués par « Sat Maj » et F1-FITC. Le profil du marquage obtenu ne se superpose pas à ceux décrits précédemment ; ce qui indique que F1-FITC est une sonde majoritairement péricentromérique. En conclusion, dans cette étude, des sondes, à base de polyamides, marquées par divers fluorophores ont permis la détection des séquences d'ADN répétées péricentromériques murines dans des cellules fixées et trois sondes fluorescentes ont permis

d'observer ces mêmes séquences directement dans les cellules vivantes. Mais, l'extinction rapide du signal fluorescent obtenu ne permet pas d'étudier la dynamique des séquences répétées dans les cellules vivantes et une optimisation ultime de ces sondes demeure donc nécessaire. En ce qui concerne le modèle humain, le traitement des cellules vivantes par les différentes sondes synthétisées a conduit à trois types de marquage distincts. Le premier, obtenu avec *Hb1-FITC* et *Hbb1-FITC*, consiste en un marquage uniforme des noyaux sans révéler de structures nucléaires particulières et la visualisation des chromosomes, dans les cellules vivantes en division, montre que seulement l'ADN est marqué par ces sondes. Le deuxième, obtenu avec *H1-FITC*, montre un signal nucléaire de très faible intensité, mettant en évidence, là encore, l'importance de l'insertion du synthon β -ala pour favoriser la pénétration dans les cellules quand les polyamides contiennent deux synthons Im. Le troisième, obtenu avec les sondes composées par les longs polyamides (Hb2-FITC, Hbb2-FITC, Hb3-FITC et Hbb3-FITC), ne présente aucun marquage nucléaire. De plus, la présence d'agrégats fluorescents suggère que ces longs polyamides ont tendance à interagir entre eux. Enfin, dans des cellules humaines fixées, un marquage nucléaire uniforme a été observé avec l'ensemble des sondes humaines synthétisées. Un marquage additionnel avec le DAPI, qui est un marqueur générique de l'ADN, a montré une superposition des signaux des sondes synthétisées et du DAPI, alors qu'avec une sonde LNA, ciblant les séquences asatellites centromériques humaines, un profil de marquage ponctuel non superposable à ceux décrits précédemment a été obtenu. Les sondes humaines marquent donc tout l'ADN sans reconnaissance spécifique des séquences d'intérêt. Pour conclure, les études sur le modèle humain ont apporté des informations sur l'importance des modifications structurales des MGBs afin de permettre la construction de sondes fluorescentes avec des meilleures propriétés d'affinité et de spécificité. En effet, les longs polyamides semblent être une bonne alternative mais des optimisations sont nécessaires pour permettre leur pénétration dans les cellules vivantes. Aussi, malgré leur manque de spécificité, les sondes fluorescentes humaines composées par les polyamides standards peuvent être utilisées dans les cellules vivantes pour marquer l'ADN sans perturber le fonctionnement des cellules.

Ainsi, ce travail de thèse a permis de valider l'utilisation des polyamides fluorescents comme des nouveaux outils pour visualiser des séquences répétées d'ADN dans les cellules. Des améliorations restent cependant à apporter à ces sondes afin de pouvoir étudier la dynamique des séquences répétées d'ADN en interphase et pendant la division cellulaire. Ces outils pourraient alors révéler des aspects importants de la biologie des séquences répétées des chromosomes.

B - Perspectives

Les travaux présentés dans cette thèse s'inscrivent dans une perspective d'identification et de développement de nouveaux outils pour l'étude de la dynamique des séquences répétées d'ADN, et plus particulièrement des régions centromériques, dans des cellules vivantes. Ils ont porté sur la synthèse de molécules qui sont des polyamides composés par des monomères N-méthylpyrrole et N-méthylimidazole (également appelés MGBs) et leur couplage à des fluorophores. Ils ont notamment permis de mieux comprendre les paramètres pouvant influer sur les propriétés de l'interaction entre les polyamides et l'ADN. L'ensemble des résultats obtenus au cours de ces travaux de thèse permet de proposer un certain nombre d'optimisations structurales qui pourraient permettre d'améliorer les performances de ces sondes fluorescentes comme outils de visualisation des séquences d'ADN répétées dans les cellules vivantes.

Le fluorophore

Au cours des expériences de traitement des cellules murines vivantes par différentes sondes fluorescentes, nous avons pu constater l'importance de la nature du fluorophore (dans notre cas avec une structure similaire à celle de la FITC) et d'une liaison thiourée entre la partie polyamide et le fluorophore pour obtenir une bonne pénétration nucléaire des sondes.

Toutefois, l'extinction de fluorescence des sondes utilisées est trop rapide pour permettre une prise d'images successives. Récemment, des travaux du groupe de Merkx ont démontré la faisabilité pratique de coupler des peptides et des protéines fluorescentes telles que la YFP (Yellow Fluorescent Protein) et la CFP (Cyan Fluorescent Protein) avec les pyrrole-imidazole polyamides en utilisant une réaction de « ligation chimique native » (Janssen et al. 2015). Même si les résultats obtenus dans cette étude ne sont pas optimaux (perte de l'affinité des polyamides pour leur séquence cible), cette approche originale mérite d'être considérée. Une stratégie analogue pourrait être envisagée dans le but d'améliorer la durée de vie de fluorescence des sondes synthétisées dans ce travail et ainsi, elles pourraient être utilisées afin d'étudier la dynamique des séquences répétées au cours du cycle cellulaire.

D'autre part, une diminution de l'affinité des MGBs a été observée après leur couplage à un fluorophore. Cet effet néfaste pourrait être contrecarré par le déplacement du site de couplage du fluorophore sur le polyamide. En effet, le changement de position du fluorophore, comme par exemple sur le coude γ , permettrait de libérer la fonction amine Nterminale des MGBs qui pourrait alors être protonée, et ainsi favoriserait l'interaction avec la cible d'ADN. Une autre approche envisageable serait d'utiliser la réaction de « chimie-click » avec (Besanceney-Webler et al. 2011) ou sans cuivre (Baskin et al. 2007) pour coupler le fluorophore au polyamide, *in situ*, directement dans les cellules. À titre d'exemple, dans un travail de Rodriguez, Miller et al, la pyridostatine, qui est un ligand des structures Gquadruplexes de l'ADN, a été fonctionnalisée par un groupement alcyne afin d'être couplée in cellulo au fluorophore Alexa fluor 594 (fonctionnalisé par un groupement azide) en utilisant un kit « Click-iT » disponible commercialement (Invitrogen) qui permet de réaliser la « chimie-click » directement dans des cellules (Rodriguez et al. 2012). En appliquant une telle approche aux MGBs, l'affinité des polyamides seuls étant supérieure à celles de leurs conjugués fluorescents, on peut supposer qu'ils reconnaitraient mieux leur cible dans les cellules ; l'autre avantage étant que les MGBs ne souffrent pas de problèmes de pénétration cellulaire. De plus, l'utilisation de fluorophores dont l'extinction de fluorescence est moins rapide, tels que les cyanines (par exemples Cy3 et Cy5), pourrait alors devenir possible. À plus long terme, en utilisant ces sondes couplées à leurs fluorophores directement dans les cellules, des expériences de FRET pourraient être envisageables. Elles permettraient de cibler des séquences plus longues en utilisant une paire de sondes fluorescentes, avec pour conséquence d'améliorer la spécificité des sondes en évitant les problèmes d'agrégation potentielle rencontrés avec les longs polyamides (pour rappel voir le § Chapitre II – B.3.1).

Les polyamides

En ce qui concerne les MGBs, plusieurs possibilités peuvent être proposées dans l'optique d'améliorer leurs propriétés de reconnaissance de l'ADN.

Tout d'abord, pour les tandems de polyamides, des études de modélisation moléculaire pourraient permettre de comprendre si une structure alternative des tandems, dans une orientation « tête à queue » plutôt que « tête à tête » serait plus favorable à l'interaction avec l'ADN db. Aussi, ces études structurales approfondies pourraient apporter des informations sur le bras de liaison entre les deux parties du tandem de polyamides : doit-il être plus rigide ou plus flexible afin de permettre un bon positionnement du tandem au sein du petit sillon de la double hélice d'ADN ? Récemment, l'équipe de Sugiyama a publié des travaux sur les répétitions télomériques humaines (TTAGGG)ⁿ, dans lesquels un trimère « tête à queue » composé par trois blocs de polyamides en « épingle à cheveux » (**Figure 56B**) pouvant reconnaître une séquence d'ADN de 18 pb a été conçu (Kawamoto et al. 2015). La spécificité du trimère est nettement améliorée par rapport à celle obtenue avec un tandem composé de deux polyamides (**Figure 56A**) décrit précédemment par cette même équipe (Kawamoto et al. 2013). La comparaison du marquage fluorescent des télomères dans des cellules humaines fixées avec le trimère couplé au TAMRA par rapport à celui obtenu avec le tandem fluorescent a révélé une forte diminution du bruit de fond (**Figures 56A** et **B**). En s'inspirant de ces travaux, une telle approche pourrait aussi être envisagée pour cibler les séquences d'ADN répétées centromériques.



Figure 56 : Représentation schématique du tandem et du trimère de polyamides ciblant les répétitions télomériques humaines

A/ À gauche, le modèle schématique du tandem fluorescent de 2 polyamides, couplé au TAMRA, avec une orientation « tête à queue » et ciblant les répétitions télomériques humaines. À droite, l'image de fluorescence d'un noyau d'une cellule fixée humaine Hela 1.3 marqué par ce tandem fluorescent (10 nM). B/ À gauche, le modèle schématique du trimère fluorescent de 3 polyamides couplé au TAMRA, avec une orientation « tête à queue » et ciblant les répétitions télomériques humaines. À droite, l'image de fluorescent (10 nM). B/ À gauche, le modèle schématique du trimère fluorescent de 3 polyamides couplé au TAMRA, avec une orientation « tête à queue » et ciblant les répétitions télomériques humaines. À droite, l'image de fluorescence d'un noyau d'une cellule fixée humaine Hela 1.3 marqué par ce trimère fluorescent (50 nM). L'hexagone gris représente la charnière entre les polyamides. Les Py et Im sont représentés par les cercles vides et noirs respectivement, la β -alanine par un losange blanc et l'étoile jaune représente le fluorophore TAMRA. Barre d'échelle : 5 µm (les images des cellules sont tirées de (Kawamoto et al. 2015)).

De plus, nous avons pu constater grâce aux résultats des expériences de dénaturation thermique avec le tandem T2 que l'interaction avec sa séquence cible entière conduisait à une

stabilisation thermique du duplexe cible d'ADN additive et non coopérative. En effet, l'augmentation des températures de fusion obtenues avec les duplexes mutés sur chacun des sites de fixation étaient de 9°C et 7°C, alors que le ΔT_m du duplexe contenant le site intégral pour le tandem *T2* était de 16°C (pour rappel voir § Chapitre II – A.2.2.1.2). En changeant le bras de liaison du tandem, par exemple en le rendant moins flexible, il serait sans doute possible d'augmenter la coopérativité entre les deux sous-unités et ainsi d'améliorer les propriétés de l'interaction de ces tandems avec l'ADN.

Enfin en ce qui concerne les longs MGBs, l'ouverture de « l'épingle à cheveux » (pour rappel voir la **Figure 51** et le § Chapitre II – B.2.2.1), même si elle reste encore hypothétique, pourrait être empêchée en synthétisant des polyamides analogues cycliques (pour rappel voir § Chapitre I – D.3.2.1). De plus, l'ajout d'une charge positive sur les coudes γ pourrait favoriser l'interaction avec l'ADN (Li et al. 2013).

Au-delà des améliorations structurales des polyamides, le choix de la séquence cible dans les régions centromériques pourrait également être revu.

La séquence cible d'ADN

Avec le modèle murin, les chromocentres qui correspondent à l'agrégation des séquences péricentromériques murines ont pu être observés directement dans des cellules fixées et vivantes avec les sondes fluorescentes à base de polyamides. En revanche avec le modèle humain, en utilisant le même type de molécules, un marquage nucléaire uniforme de l'ADN a été observé sans distinction de structures nucléaires particulières apparentées aux chromocentres murins. Trois hypothèses peuvent être émises en ce qui concerne les différences observées entre les modèles murin et humain en termes de structures nucléaires révélées par les sondes MGBs.

La première hypothèse est que l'accessibilité des séquences d'ADN est différente dans les deux espèces. Il est possible que les séquences d'ADN centromériques soient plus condensées chez l'homme par rapport à la souris, ou encore que la présence de protéines centromériques spécifiques empêche les polyamides fluorescents de se fixer. La poursuite de cette étude avec l'utilisation de sondes MGBs plus affines (voir ci-dessus pour les modifications structurales des polyamides qui seraient susceptibles d'améliorer l'affinité de ces molécules) permettrait sans doute d'apporter des éclaircissements sur ces questions qui restent ouvertes.

La deuxième hypothèse est que le nombre de sites de fixation des polyamides par séquence cible serait plus important chez la souris, permettant ainsi de détecter un signal chez cette espèce mais pas chez l'homme. En tenant compte des données existantes, il y aurait en moyenne 19,5 sites de fixation par répétition murine contre seulement 4,8 sites de fixation par répétition humaine (communication personnelle de Loïc Ponger). Cependant, à l'heure actuelle, le nombre exact de répétitions au niveau de ces séquences n'est pas connu car ces régions centromériques ne sont encore que partiellement séquencées à cause des difficultés techniques liées à leur caractère répétitif. De plus, il est à noter que les répétitions des régions centromériques sont très conservées chez la souris ; alors que chez l'homme, ces régions sont plus polymorphes ; ce qui pourrait influer sur le nombre de sites de fixation disponible (Vissel & Choo 1989; Lee et al. 1997).

Enfin, la troisième hypothèse est que l'enrichissement des séquences cibles au niveau des régions centromériques par rapport au reste du génome, qui permet de distinguer le signal du bruit de fond, soit différent. Une analyse « bio-informatique » des séquences disponibles pour les génomes de ces deux espèces (communication personnelle de Loïc Ponger) montre un enrichissement de 3 chez la souris ; c'est-à-dire qu'au niveau des régions centromériques, le motif d'intérêt est retrouvé toutes les 10 pb alors qu'il est retrouvé toutes les 31 pb dans le reste du génome. Chez l'homme, cet enrichissement est de 1,5 car le motif est retrouvé toutes les 33 pb au niveau des régions centromériques alors qu'il est retrouvé toutes les 50 pb dans le reste du génome. Une telle différence d'enrichissement entre les deux espèces pourrait –elle expliquer les différences de comportement des MGBs fluorescents ?

Une autre utilité des molécules synthétisées

La dynamique des séquences répétées peut être étudiée sous deux aspects : leur organisation nucléaire ou leur transcription. Au-delà des travaux qui ont été réalisés pendant cette thèse et pour lesquels les polyamides synthétisés ont été employés comme outils de visualisation, il serait également intéressant d'utiliser ces molécules afin d'étudier leurs effets sur la transcription de ces séquences répétées (pour rappel voir le § Chapitre I – B.1.2). En effet, seuls ou couplés à des protéines comme des histones déméthylases (afin de voir l'effet sur la marque $H_3K_9me_3$) ou des déacétylases ou bien à des groupements fonctionnels tels que

des agents alkylants (comme par exemples le chlorambucil (Wurtz & Dervan 2000) ou des complexes de platine) ou des nucléases chimiques (comme la bipyridine) (Simon et al. 2008), ils pourraient avoir une influence positive ou négative sur l'expression transcriptionnelle. Ces dérèglements induits par les polyamides pourraient permettre d'étudier les mécanismes impliqués et les effets biologiques associés à la transcription des régions centromériques. On sait notamment que des transcrits non codants issus des régions centromériques sont impliqués dans l'assemblage de la chromatine (Eymery et al. 2009) (pour rappel voir le § Chapitre I – B.1.2). Leur fixation pourrait aussi entraîner une absence d'interaction ou un déplacement de protéines ou de facteurs à l'origine de modifications potentielles dans le fonctionnement cellulaire. Ainsi, des informations complémentaires importantes sur la fonction de ces séquences répétées pourraient être obtenues.

Chapitre IV – Matériels et méthodes

Chapitre IV - Matériels et méthodes

Chapitre IV - Matériels et méthodes	
A – Réactifs	167
B – Tampons	168
C - Appareils	
D - Synthèse des polyamides	
E - Marquage fluorescent des polyamides	
F - Électrophorèse sur gel de polyacrylamide non dénaturant (expériences de «	gel-retard »)
G – Expériences de spectroscopie	
G.1 Absorbance	176
G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sono	le
fluorescente	177
G.3 Spectres différentiels de dichroïsme circulaire	178
G.4 Spectres de fluorescence	
H – Études Cellulaires	
H.1 Culture cellulaire	
H.2 Marquage in situ des cellules fixées	
H.3 Acquisition des images par microscopie	

<u>A – Réactifs</u>

Les monomères utilisés pour la synthèse des polyamides, l'acide 4-(Boc-amino)-1méthyl-1H-pyrrole-2-carboxylique, l'acide 4-(Boc-amino)-1-méthyl-1H-imidazole-2carboxylique, la N-Boc-β-alanine, l'acide N-Boc-γ-aminobutyrique et l'acide 2-(Boc-amino)-5-azido-pentanoïque sont achetés chez Bachem (Germany) ou Polypeptide Laboratories (France). Les agents activateurs, le HATU ((2-(1H)-7-Azabenzotriazol-1-yl)-1, 1, 3, 3tétraméthyluronium-hexafluorophosphate) et l'HOAt (1-Hydroxy-7-azabenzotriazole) viennent de chez « Molekula » (UK). La résine Boc-β-ala-PAM (Boc-β-alanine-(4carbonylaminométhyl)-benzyl-ester-copoly-(styrène-divinyl-benzène)) vient de chez Merck. Les esters N-hydroxysuccinimide des fluorophores Cy3 (NHS-Cy3), Cy5 (NHS-Cy5) et de la rhodamine (NHS-Rhod) et la fluorescéine isothiocyanate sont achetés chez FluoProbes et Santa Cruz. La coumarine MM14 a été donnée par l'équipe de Marie-Paule Teulade-Fichou, Institut Curie, Orsay, France. Les fluorophores non commerciaux S24 ((E)-1-(4carboxybutyl)-4-(4-(diméthylamino)styryl)pyridinium iodide) (Barni et al. 1986) et TO ((Z)- $1-(5-carboxypentyl)-4-((3-m\acute{e}thylbenzo[d]thiazol-2(3H)-ylidene)m\acute{e}thyl)quinolinium$

perchlorate) (Yarmoluk et al. 1996) ont été donnés par Serguiy M. Yarmoluk de l'Institut de Biologie Moléculaire et de Génétique (National Academy of Sciences of Ukraine). Leurs synthèses sont décrites dans les publications mentionnées ci-dessus. Les autres réactifs (sels ou monomères) viennent des sociétés Sigma-Aldrich, Acros, Merck, Fluka et les solvants sont achetés chez Carlo Erba. Ils sont utilisés sans purification préalable. Le THPTA (Tris(HydroxyPropyl)TriazolméthylAmine) a été donné par le professeur T. Brown (Université de Southampton, UK) ou acheté chez Sigma Aldrich.

Les oligonucléotides utilisés sont synthétisés par la société Eurogenetec (Seraing, Belgique). Délivrés sous forme lyophilisée, ils sont remis en suspension dans de l'eau bidistillée. Leur concentration est déterminée par spectroscopie d'absorbance à 260 nm en utilisant les coefficients d'extinction molaire (ε) donnés dans le **Tableau 12**. Toutes les solutions stocks d'oligonucléotides sont ensuite conservées à -20°C dans de l'eau bi-distillée.

Tableau 12 : Liste des oligonucléotides utilisés dans ce manuscrit						
Nom	Longueur (pb)	ϵ (L.mol ⁻¹ .cm ⁻¹)	Séquence			
Duplexe 1	30	326400 261000	5'-GAG-GAA-AAC-TGA-AAA-AGG-TGG-AAA-ATT-TAG-3' 3'-CTC-CTT-TTG-ACT-TTT-TCC-ACC-TTT-TAA-ATC-5'			
Duplexe 2	18	196900 152500	5'-AAC-TGA-AAA-AGG-TGG-AAG-3' 3'-TTG-ACT-TTT-TCC-ACC-TTC-5'			
Duplexe 3	25	263200 208900	5'-GGG-TGA-GGA-GGA-GGG-AGA-TCG-GTC-A-3' 3'-CCC-ACT-CCT-CCC-TCT-AGC-CAG-T-5'			
Duplexe 4	15	169100 128800	5'-AAA-AAG-GTG-GAA-AAT-3' 3'-TTT-TTC-CAC-CTT-TTA-5'			
Duplexe 5	15	168400 127900	5'-AAA-AAG-GTG-GAA-GAT-3' 3'-TTT-TTC-CAC-CTT-CTA-5'			
Duplexe 6	15	168400 127900	5'-AAG-AAG-GTG-GAA-AAT-3' 3'-TTC-TTC-CAC-CTT-TTA-5'			
Duplexe 7	26	256700 267900	5'-TAA-TCA-ATT-TGT-GAT-GTG-TCA-TTA-TG-3' 3'-ATT-AGT-TAA-ACA-CTA-CAC-AGT-AAT-AC-5'			

B – Tampons

Pour les expériences de spectroscopie, des tampons qui n'absorbent pas ou peu dans le domaine de longueurs d'onde étudiées, entre 200 et 700 nm et qui n'ont pas de propriétés de fluorescences notables, sont utilisés (comme le Tris-HCl à pH 7,5). Les expériences de dénaturation thermique nécessitent des tampons dont le pH varie peu avec la température, ce qui est le cas du tampon cacodylate à pH 6,2 utilisé pour ces expériences. Tous les tampons utilisés pour les expériences de spectroscopie contiennent 50 mM de NaCl, 50 mM de KCl et

1 mM de MgCl₂. Un tampon standard TBE 1X (89 mM Tris-HCl, 89 mM acide borique, 2 mM EDTA, à pH 8,3) est utilisé pour les expériences d'électrophorèse.

<u>C</u> - Appareils

Les analyses et purifications par chromatographie liquide haute performance en phase inverse sont effectuées sur des HPLC analytique et préparative de marque Agilent Technologies 1200 pilotées par le programme Agilent ChemStation. Les gradients utilisés sont linéaires 0 – 100% d'acétonitrile / 0,1% TFA (acide trifluoroacétique). L'analyse des produits est effectuée sur une colonne analytique Waters XBridge C18 5µm 4,6×250 mm pendant 25 min à un débit de 2 mL / min et leur purification est effectuée sur une colonne préparative Waters XBridge C18 5µm 19×250 mm pendant 22 min à un débit de 10 mL / min. Les analyses de spectrométrie de masse Q-TOF par électrospray (ESI Q-TOF) sont effectuées sur un instrument Q-Star (Applied Biosystems, Courtaboeuf, France) en mode positif par l'équipe de la plateforme de spectrométrie de masse du Muséum National d'Histoire Naturelle. Les échantillons sont dilués dans du méthanol. Les spectres UV-visible et les courbes de dénaturation thermique sont enregistrés sur des spectrophotomètres Uvikon XL UV-visible (Secomam, Nova Analytic Compagny) avec un passeur thermostaté de 10 échantillons et 2 références. Pour les mesures d'absorbance dans l'UV à l'aide de ces spectrophotomètres, des cuves en quartz (Hellma, France) de parcours optique 1 cm et de contenance 500 - 600 µL sont utilisées. Ces cuves sont installées dans un support mobile thermostaté piloté par le logiciel ThermAlys (DudoTec GmbH). Un thermostat Julabo F25, piloté par le logiciel Easy Temp Professional est utilisé pour gérer et enregistrer les températures. Les gels de polyacrylamide sont analysés par un scanner fluorescent Typhoon 9410 (GE Healthcare) avec une longueur d'onde d'excitation de 532 nm et un filtre d'émission à 526 nm. Les spectres de dichroïsme circulaire sont enregistrés sur un spectropolarimètre Jasco J810, à lampe de Xénon (150 W refroidie par l'air), purgé par de l'azote afin d'enlever l'ozone pour préserver les systèmes optiques et équipé d'un passeur de 6 échantillons, thermostaté par effet Peltier. Les spectres de fluorescence sont enregistrés sur un fluorimètre SPEX Fluorolog II (Jobin-Yvon-Horiba), à lampe de Xénon (450 W), possédant deux monochromateurs pour l'émission. Des cuves en quartz transparentes et de trajet optique 1 cm sont utilisées. Les réactions de « chimie-click » sont effectuées dans un réacteur à micro-ondes Monowave 300 (Anton Paar).

D - Synthèse des polyamides

La synthèse de type peptidique des polyamides N-méthylpyrrole / N-méthylimidazole est effectuée manuellement sur un support solide en utilisant la méthode décrite par Krutzik et Chamberlin (Krutzik & Chamberlin 2002a; Krutzik & Chamberlin 2002b). Les différentes étapes sont détaillées ci-après.

La résine Boc- β -ala-PAM : En utilisant 100 mg de résine de capacité théorique $\geq 0,5$ mmol/g, la capacité minimum est de 0,05 mmoles, et compte tenu des rendements obtenus après les différentes synthèses, la capacité estimée de la résine dépasse 0,06 mmoles ; c'est pourquoi les calculs ci-dessous ont été effectués en considérant la capacité de la résine à 0,06 mmoles.

Tableau 13 : Les différents réactifs pour la synthèse du MGB F4β de séquence Nterm-
γ2-Py8-Py7-Py6-Py5-γ1-Py4-Py3-βala2-Im1-β-Dp-Cterm

Nom des	Masse Molaire	Quantité	Nombre d'équivalents par	Masses pesées
réactifs	(g / mol)	(mmoles)	rapport à la résine	(mg)
Résine	-	0,06	-	100
Im1	241,3	0,24	4	57,9
DCC	206,3	0,24	4	49,5
HOAT	136,1	0,24	4	32,7
βala2	189,2	0,48	8	90,8
HATU ₂	380,2	0,432	7,2	164,2
Py3	240	0,24	4	57,6
HATU ₃	380,2	0, 216	3,6	82,1
Py4	240	0,24	4	57,6
HATU ₄	380,2	0, 216	3,6	82,1
γ1	203,2	0,48	8	97,5
$HATU_{\gamma 1}$	380,2	0,432	7,2	164,2
Py5	240	0,24	4	57,6
HATU ₅	380,2	0, 216	3,6	82,1
Руб	240	0,24	4	57,6
HATU ₆	380,2	0, 216	3,6	82,1
Py7	240	0,24	4	57,6
HATU ₇	380,2	0, 216	3,6	82,1
Py8	240	0,24	4	57,6
HATU ₈	380,2	0, 216	3,6	82,1
γ2	203,2	0,48	8	97,5
$HATU_{\gamma 2}$	380,2	0,432	7,2	164,2

Procédure générale :

Préparation de la résine : À température ambiante, 100 mg de résine Boc-β-ala-PAM sont déposés dans une colonne avec un filtre en verre de porosité 4. Deux lavages avec du solvant DCM (DiChloroMéthane) sont effectués puis la résine est laissée 15 min dans ce même solvant pour qu'elle gonfle. Une agitation avec une baguette métallique en acier inoxydable est nécessaire pour éliminer les bulles qui peuvent se former. Ensuite, le solvant est élué plus rapidement sous pression d'azote. Puis la résine est lavée 2 fois par du DMF (DiMéthylFormamide) et laissée 15 min dans ce même solvant. Pour finir, la résine est rincée 3 fois par du DCM.

Déprotection de la résine : Pour libérer la fonction amine protégée par un groupement Boc sur la résine Boc- β -ala-PAM, un volume de solution TPW (TFA 92,5 % / phénol 5 % / eau 2,5%, préparée au préalable et pipetée à l'aide d'une pipette en plastique) suffisant pour recouvrir la résine est ajouté dans la colonne. La résine est laissée 1 min dans la solution TPW puis 1 min et enfin 3 min (en remplaçant à chaque fois la solution TPW). Pour finir, la résine est rincée 3 fois par du DCM (premier lavage par du DCM car le TFA est miscible dans le DCM mais pas dans le DMF) et 3 fois par du DMF (solvant utilisé pour le couplage des synthons sur la résine).

Activation du premier synthon et son couplage sur la résine : L'activation des synthons Py, γ et β -ala s'effectue en présence de HATU dans un mélange de 600 μ L de DMF : NMP (N-Méthyl-2-Pyrrolidone) (1 :1). L'activation des synthons Im et du dimère Py-Im s'effectue en présence de HOAt et de DCC dans un mélange DMF : NMP (1 :1). Le HATU est l'agent de couplage utilisé pour activer les synthons Py, β -ala et γ et le HOAt est l'agent de couplage utilisé en association avec le DCC pour activer les synthons Im ou le dimère Py-Im. Pour finir l'activation des différents synthons, 12 équivalents de DIEA (N, N-DiIsopropylEthylAmine) sont ajoutés (129,3 mg/mmol ; densité = 0,742). Une incubation de 10 min maximum est préalable à l'ajout du mélange sur la résine. La DIEA est une base permettant de générer des conditions basiques pour le couplage en déprotonant complètement les fonctions amines utilisées pour les différents couplages.

En ce qui concerne le synthon Im et le dimère Py-Im, l'activation à température ambiante dure 2 h. Le synthon Im est ajouté au HOAt dilué dans 300 μ L de NMP et au DCC dilué dans 300 μ L de DMF. Le dimère Py-Im est dissout dans 600 μ L de DMF sec (doublement des volumes pour faciliter la dissolution) et le DCC et l'HOAt dissouts dans 600 μ L de NMP sont ensuite ajoutés. Une étape de centrifugation de 3 min à 4000 rpm et une filtration (filtre de porosité 4) du précipité d'urée sont nécessaires avant l'ajout de la DIEA.

Fixation des synthons : Le couplage du synthon sur la résine est effectué pendant 25 min sous agitation par l'azote (ou 2 h pour le dimère). Ensuite, la colonne est vidée et la résine est rincée 3 fois par du DMF et 3 fois par du DCM.

Déprotection du premier synthon par TPW : La même procédure que celle utilisée pour la déprotection de la résine par la solution de TPW est suivie, sauf pour la déprotection des synthons Im qui se fait en deux étapes : la première dure 1min et la deuxième 10 min. Enfin la résine est lavée 3 fois par du DCM et 3 fois par du DMF.

Activation et couplage des synthons suivants puis déprotection : Les mêmes étapes de déprotection et de couplage sont alternées jusqu'à la fixation du dernier synthon sur la résine.

Fin de la synthèse et lavage de la résine : La résine est lavée 2 fois par du DMF puis une série de lavages est effectuée : 3 fois DMF, 3 fois DCM, 2 fois DMF, 2 fois DCM, 3 fois MeOH (Méthanol), 3 fois DCM, 3 fois MeOH, 2 fois DCM et 5 fois éther.

Récupération de la résine : En utilisant une pipette Pasteur, la résine dans l'éther est récupérée, l'éther est enlevé puis la résine est séchée sous-vide.

Aminolyse : L'étape finale, qui consiste à décrocher le polyamide synthétisé de la résine, est effectuée pendant 14 – 16 h voir toute la nuit à 55°C dans un bain thermostaté en ajoutant 0,75 mL de DMPA (N, N-DiMéthylaminoPropylAmine). Pour finir, la résine est filtrée puis rincée par 500 μ L de DMPA et un volume équivalent de DCM. Une précipitation au diéthyléther est effectuée et le produit est conservé à - 20°C. Une faible quantité du produit obtenu est prélevée et dissoute dans du MeOH afin d'effectuer des analyses par HPLC analytique (gradients linéaires 0 – 100% d'acétonitrile / 0,1% TFA pendant 25 min à un débit de 2 mL / min) et par spectrométrie de masse pour vérifier la pureté du produit et le caractériser (voir **Annexes 1** et **2**). Ensuite, les produits synthétisés sont purifiés par HPLC préparative (gradients linéaires 0 – 100% d'acétonitrile / 0,1% TFA pendant 22 min à un débit de 10 mL / min).

Une réaction de « chimie-click » (cycloaddition 1,3 dipolaire de Huisgen) catalysée par le cuivre a été utilisée pour coupler en tandem deux polyamides en « épingle à cheveux » (Meldal & Tornøe 2008). Des polyamides fonctionnalisés en N-term par des groupements alcyne ou azoture sont synthétisés sur support solide en utilisant de l'acide 2-(Boc-amino)-5azido-pentanoïque ou de l'acide 4,5-pentynoïque comme synthons N-terminaux. Ils sont dissouts en quantité équimolaire dans 500 µL de DMSO à une concentration de 10 - 15 mM. Cette solution est mélangée avec des solutions aqueuses de sulfate de cuivre (300 µL, 50 mM), d'ascorbate de sodium (300 µL, 50 mM), de THPTA (300 µL, 50 mM), du tampon acétate de triéthylammonium (300 µL, 1M), du DMSO (250 µL) et complétée par de l'eau pour obtenir un volume final de 3 mL. Le mélange est ensuite placé dans le réacteur à microonde sous irradiation à 90°C pendant 20 min sous agitation magnétique. Si le tandem précipite pendant la réaction, ce précipité est collecté par centrifugation pendant 5 min à 4000 rpm, dissout dans du DMSO (200 µL) et dilué dans l'eau (250 µL). Les solutions contenant les produits synthétisés sont traitées par une résine de chélation CHELEX (400 µéquiv. / mL, BioRad) dans le but d'enlever les ions de cuivre. Un excès molaire de résine (3 fois) est ajouté aux solutions qui sont agitées quelques minutes et centrifugées. L'eau contenue dans le surnageant est enlevée par évaporation sous vide et les produits finaux sont ensuite purifiés par HPLC dans les mêmes conditions que celles décrites pour les polyamides.

E - Marquage fluorescent des polyamides

La fonction amine N-terminale des polyamides ou l'amine libre des tandems de polyamides sont couplées soit aux esters N-hydroxysuccinimide, soit aux groupements isothiocyanate des fluorophores en respectant les protocoles donnés par les fournisseurs. 2 mg $(1,2 - 1,5 \mu mol)$ des polyamides dissouts dans du DMF sec $(70 - 100 \mu L)$ et 2 - 4 équivalents de fluorophores sous formes d'esters activés ou d'isothiocyanate dans du DMF sec $(70 - 100 \mu L)$ et 2 - 4 équivalents de fluorophores sous formes d'esters activés ou d'isothiocyanate dans du DMF sec $(70 - 100 \mu L)$ sont mélangés ensemble avec de la DIEA (12 équiv.). Après 2 - 3h d'incubation à température ambiante à l'abri de la lumière, le mélange est ajouté goutte à goutte sous agitation dans du diéthyléther (8 volumes). Le précipité est séparé par centrifugation (3 min ; 12000 rpm), rincé par du diéthyléther et séché sous vide. Le produit est ensuite dissout dans 100 μ L de DMSO et purifié par HPLC préparative dans les mêmes conditions que celles utilisées pour les polyamides.

Les fluorophores sous forme d'acide carboxylique sont activés par du HATU. Les polyamides (2 mg ; $1,2 - 1,5 \mu mol$) sont dissouts dans du DMF sec (200 μ L) et ensuite, le fluorophore (3 équiv.), le HATU (2,8 équiv.) et la DIEA (12 équiv.) sont ajoutés. Le mélange est agité pendant 2 – 3h à température ambiante à l'abri de la lumière. Le produit est ensuite précipité puis purifié par HPLC préparative.

La caractérisation des conjugués (polyamide + fluorophore) est effectuée par spectrométrie de masse et spectroscopie UV-visible (voir **Annexes 7** et **18**).

Pour les expériences biophysiques, les polyamides et les sondes fluorescentes sont dissouts dans du MeOH ou du DMSO.

<u>F - Électrophorèse sur gel de polyacrylamide non</u> <u>dénaturant (expériences de « gel-retard »)</u>

L'électrophorèse de gels de polyacrylamide en condition non dénaturante (native) permet d'observer d'éventuels complexes formés par les duplexes d'ADN et les polyamides. Cette méthode est fondée sur le déplacement de molécules chargées (positivement ou négativement) sous l'effet d'un champ électrique. Du fait de leurs caractéristiques propres et en fonction des conditions de l'électrophorèse, ces molécules auront des vitesses de migration différentes ; elles vont donc se séparer les unes des autres.

La méthode décrite pour les triples hélices a été utilisée avec quelques modifications (Boutorine & Escudé 2007). À la place du troisième brin oligonucléotidique, des solutions stocks de polyamides ou de sondes fluorescentes $(1 - 50 \ \mu\text{M})$ sont préparées dans du MeOH ou du DMSO. Des gammes de solutions à des concentrations différentes de polyamides (seuls ou couplés à des fluorophores) sont préparées dans du tampon TBE 1X standard à pH 8,3 en présence des duplexes d'ADN (15 – 520 nM, chaque concentration est indiquée dans les légendes des figures) dont un brin est marqué par la fluorescéine à l'extrémité 3' ou radioactivement par du phosphate ³²P à l'extrémité 5'. Le marquage radioactif des oligonucléotides (0,5 μ M dans 16 μ L d'eau) avec une extrémité 5'OH libre est effectué dans un volume final de 20 μ L par 0,5 μ L de T4 polynucléotide kinase (10 unités / mL, New England Biolabs (NEB)) dans 2 μ L de tampon kinase (10 X à pH 7,6 contenant 70 mM Tris-HCl, 10 mM MgCl₂ et 5 mM Dithiothréitol) en présence de 1,5 μ L de γ -³²P-ATP (3000 Ci /

mmol, Amersham Biosciences). L'excès de γ^{-32} P-ATP est retiré lors d'une étape de purification sur une microcolonne chromatographique (Bio-Spin). Les oligonucléotides marqués radioactivement ont été utilisés lors des expériences de « gel-retard » avec les sondes fluorescentes pour éliminer l'hypothèse d'un encombrement stérique dû à la présence de la fluorescéine (voir **p.109**).

Un mélange de glycérol, de bleu de bromophénol et de xylène cyanol (jusqu'à 8% de glycérol et 0,06% des indicateurs) est ajouté aux échantillons, contenant l'ADN et / ou les polyamides (volume final de 20 µL), qui sont chauffés à 90°C pendant 3 min, refroidis lentement puis stockés à 4°C jusqu'à leur utilisation. Un aliquote de chaque échantillon (10 µL) est chargé en chambre froide sur les gels de polyacrylamide non dénaturant à 20% dans du tampon TBE 1X (épaisseur de 0,5 mm, rapport 19:1 acrylamide / bis-acrylamide; polymérisé par 0,06 % de TEMED (N, N, N', N'-Tétraméthyléthylènediamine) et 0,6% de persulfate d'ammonium). Ensuite, l'étape de migration électrophorétique est effectuée dans du tampon de migration (TBE 1X) à 4 - 8°C à une puissance contrôlée de 5 Watt. Après la migration, les gels sont analysés sur le scanner Typhoon. Pour les gels radioactifs, un séchage préalable (transfert des gels sur du papier Whatman 3MM) suivie d'une exposition sur un écran radiosensible (GE Healthcare) est nécessaire pour l'analyse des gels. Ensuite, l'écran est numérisé à l'aide d'un scanner Typhoon. Les images générées sont analysées en utilisant le logiciel ImageJ pour déterminer les pourcentages de conversion en complexe dans chacun des cas. Les constantes de dissociation apparentes (K_{d,app}) sont estimées selon le protocole décrit pour les triples hélices (Novopashina et al. 2005; Boutorine & Escudé 2007).

Les expériences de cinétique sont effectuées en utilisant la méthode « d'arrêt de retard sur gel » en chambre froide à 4 - 8 °C (Novopashina et al. 2005). Un mélange (100 µL) contenant le duplexe d'ADN marqué à la fluorescéine (0,52 µM) et le polyamide (1 µM) dans du tampon TBE 1X avec 5% de glycérol, du bleu de bromophénol et du xylène cyanol est préparé comme décrit ci-dessus, si ce n'est que le polyamide est ajouté juste avant le premier chargement. Le premier échantillon (10 µL) est chargé immédiatement sans incubation puis des aliquotes de 10 µL sont pris à des intervalles de temps 10, 30, 60 et 120 min d'incubation après l'ajout du polyamide et chargés dans les puits libres d'un même gel non dénaturant (en parallèle des échantillons contrôles sans polyamide). Après chaque chargement, les échantillons sont soumis à la migration électrophorétique. Pour finir, le gel est scanné sur un scanner Typhoon et analysé par le logiciel ImageJ.

<u>G – Expériences de spectroscopie</u>

G.1 Absorbance

La spectroscopie ultraviolet-visible est une technique de spectroscopie mettant en jeu les photons dont les longueurs d'onde sont dans le domaine de l'ultraviolet (200 nm – 400 nm), du visible (400 nm – 750 nm) ou du proche infrarouge (750 nm - 1400 nm). Soumises à un rayonnement dans cette gamme de longueurs d'onde, les molécules absorbent la lumière et sont susceptibles de subir une ou plusieurs transition(s) électronique(s). Cette technique d'analyse est souvent utilisée pour déterminer la concentration d'une entité chimique en solution, en utilisant la Loi de Beer-Lambert :

$$A_{\lambda} = -\log_{10} I / I_0 = \varepsilon_{\lambda} l C$$

où I / I_0 est la transmittance de la solution (sans unité), A_{λ} est l'absorbance ou densité optique à une longueur d'onde λ (sans unité), ε_{λ} est le coefficient d'extinction molaire (en L / mol / cm), l est la longueur du trajet optique dans la solution traversée (elle correspond à l'épaisseur de la cuve utilisée (en cm)) et C est la concentration molaire de la solution (en mol / L). Le spectrophotomètre UV-visible mesure l'intensité de la lumière I passant au travers d'un échantillon et la compare à I_0 , l'intensité de la lumière passant dans un échantillon de référence contenant le même solvant que celui utilisé pour l'échantillon dans une cuve identique. En mesurant l'absorbance des différentes solutions contenant les molécules d'intérêt et comme l et ε_{λ} sont connus, cette équation a permis de déterminer les différentes concentrations de nos solutions. Les ε_{λ} des polyamides sont calculés en multipliant le nombre de noyaux aromatiques contenus dans chaque molécule par 8333 (à titre d'exemple, le MGB FI contient 8 noyaux aromatiques donc son ε_{λ} est 66664 L / mol / cm).

L'absorbance des acides nucléiques entre 200 et 300 nm est exclusivement due aux systèmes insaturés conjugués et aux fonctions carbonyles présents dans les bases puriques et pyrimidiques. En effet, la contribution du squelette phosphate n'est mesurable qu'à des longueurs d'onde inférieures à 190 nm. Le maximum d'absorption dépend donc de la séquence, de la structure et de la composition en bases. Il est compris entre 254 nm pour des séquences riches en pyrimidines. En ce qui
concerne les polyamides, les hétérocycles pyrrole et imidazole confèrent à ces molécule leur capacité à absorber à des longueurs d'onde entre 280 et 340 nm.

G.2 Dénaturation thermique des complexes ADN / polyamide ou ADN / sonde fluorescente

L'étude des interactions entre des acides nucléiques et des ligands est rendue possible par le changement des propriétés d'absorbance de l'acide nucléique et / ou du ligand. Cela nécessite néanmoins une absorption des ligands dans une gamme de longueurs d'ondes différente de celle de l'ADN, au-delà de 300 nm par exemple. Ce qui est le cas des polyamides.

Les spectrophotomètres utilisés possèdent un double faisceau qui permet la soustraction, lors de l'acquisition, du spectre du solvant ou du tampon de référence. Une première étape consiste à réaliser une ligne de base avec 2 cuves contenant la solution de référence, permettant ainsi l'initialisation et la mise à zéro du signal entre les deux charriots. Cette ligne de base est généralement réalisée de 900 à 190 nm, à une vitesse de 500 nm / min, avec un pas de 1 nm. Le spectre d'absorbance de la cuve contenant l'échantillon est ensuite mesuré entre 240 et 400 nm pour les acides nucléiques, et jusqu'à 700 nm pour les sondes fluorescentes à une vitesse de 500 nm / min et un pas de 1 nm.

Les expériences de dénaturation thermique suivie par spectroscopie d'absorbance permettent de mesurer la stabilité d'une structure en fonction de la température. En effet, la variation des propriétés d'absorbance de l'échantillon, à des longueurs d'onde bien choisies (voir ci-dessous), peut témoigner d'un changement conformationnel de la molécule. En fonction de sa structure (double ou simple brin), la dénaturation thermique d'un ADN structuré peut s'accompagner d'une hyperchromicité (augmentation de l'absorbance) alors qu'une hypochromicité (diminution de l'absorbance) peut être visible pendant sa renaturation. Dans le cas des polyamides, une hypochromicité s'exprime à une longueur d'onde bien spécifique de 335 nm. Les spectrophotomètres sont reliés à des thermostats qui permettent la régulation thermique des passeurs d'échantillons, c'est-à-dire le chauffage et le refroidissement à des gradients de températures bien définis $(0,2^{\circ}C / min)$. La concentration en acides nucléiques doit être choisie pour obtenir une variation du signal d'absorbance de l'ordre de 0,1 ce qui est généralement le cas en se plaçant à une absorbance à 260 nm comprise entre 0,7 et 1,2. Le programme ThermAlys permet l'enregistrement à intervalle régulier de l'absorbance des échantillons (9 au maximum), avec un temps d'intégration de 1 s, et de la température relevée par une sonde immergée dans une cuve. Les données sont ensuite exportées sous forme de tableur (Excel). Les variations d'absorbance à une longueur d'onde de référence (680 nm dans notre cas), où aucune variation dépendante de la structure étudiée n'est attendue, sont soustraites aux absorbances mesurées aux autres longueurs d'onde.

Les courbes de dénaturation thermique obtenues présentent une transition qui se déroule sur une gamme restreinte de températures. Ce type de transition marque le passage de la forme associée de la molécule à basse température vers sa forme dissociée à haute température. Il est possible de déduire de ces courbes sigmoïdales la température de demitransition T_m, qui correspond à la température pour laquelle 50% des molécules ont été dénaturées. Pour faciliter la détermination des T_m, les dérivées premières des courbes de fusion par rapport à la température sont calculées et les températures de fusion correspondant aux extremums de ces courbes sont déterminées.

Le protocole utilisé est aussi similaire à celui décrit pour les analyses des triples hélices (Boutorine & Escudé 2007). De 3 à 9 échantillons (500 µL chacun dans 50 mM de tampon cacodylate de sodium à pH 6,2 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂) sont préparés dans des cuves en quartz à faces optiques parallèles, transparentes et de trajet optique 1 cm. Chaque échantillon contient 1,3 µM de duplexe d'ADN et des concentrations croissantes de polyamides de 0 à 8 équivalents molaires par rapport aux duplexes d'ADN. Les solutions sont couvertes par quelques gouttes d'huile minérale et les cuves sont fermées hermétiquement par des bouchons en téflon pour prévenir l'évaporation. Les cuves contenant les échantillons sont ensuite déposées dans le porteur de cuves, thermostaté, du spectrophotomètre. Les échantillons sont chauffés et les densités optiques à 260, 335 et 680 nm sont mesurées toutes les 5 min pour chaque échantillon. Deux cycles de mesures sont effectués : la dénaturation pendant le chauffage de 20 à 95°C et la renaturation pendant le refroidissement de 95 à 20°C. Les données sont ensuite exportées et traitées en utilisant le logiciel Microsoft Excel.

G.3 Spectres différentiels de dichroïsme circulaire

Le dichroïsme circulaire (DC) est une méthode de choix pour l'étude de molécules chirales en solutions, et donc une méthode particulièrement attractive pour l'étude de systèmes biologiques. Le DC est communément utilisé pour étudier les macromolécules biologiques soit pour évaluer les changements de conformation de celles-ci, soit pour révéler leurs interactions avec de petites molécules, plus particulièrement celles qui sont achirales et dont le signal de DC induit est uniquement dû à leur interaction avec la macromolécule.

Le signal de DC d'une molécule est défini par la différence d'absorption entre la lumière polarisée circulairement à gauche Ag et celle polarisée à droite Ad.

$$DC = Ag - Ad$$

En pratique, le spectropolarimètre possède un modulateur photoélastique qui produit alternativement une lumière polarisée circulairement à droite puis à gauche avec une fréquence de changement de 50 Hz. Lorsque la lumière, d'intensité constante, traverse un échantillon chiral, une fluctuation d'intensité en phase avec le modulateur de fréquence apparaît. Les photons transmis sont détectés par un photomultiplicateur qui produit un courant dont l'amplitude dépend du nombre de photons incidents, ce qui génère le signal DC (positif ou négatif en fonction de la phase du courant alternatif piloté par le modulateur).

Entre 200 et 300 nm, les spectres de DC des acides nucléiques résultent des transitions dues aux couplages entre les bases, lorsqu'elles sont empilées dans une structure chirale (par exemple de type hélicoïdale). Le signal reflète l'orientation des bases entre elles. Les polyamides sont, très généralement, des molécules achirales qui ne présentent aucun spectre de DC en solution. Mais, au contact de l'ADN, ils se structurent et acquièrent un signal de DC induit, entre 300 et 400 nm, qui est caractéristique de leur interaction.

Le spectropolarimètre utilisé, Jasco J-810, est un appareil monofaisceau. Il est donc nécessaire d'enregistrer une ligne de base correspondant au tampon sans molécule (ou avec l'ADN pour la soustraction de son spectre) lors de toutes les mesures réalisées. Il est d'ailleurs préférable d'effectuer le spectre du tampon dans la même cuve en quartz que la mesure elle-même, chaque cuve pouvant avoir un signal de DC intrinsèque. L'acquisition des spectres de DC est réalisée à l'aide du logiciel Spectra Manager développé par Jasco. Le balayage pour les polyamides est généralement effectué de 400 à 240 nm à une vitesse de 200 nm / s avec un pas de 1 nm, une largeur de bande de 1 nm et un temps de réponse de 1 s. Le signal de DC est généralement faible et les spectres résultant bien plus « bruités » que des spectres d'absorbance ; il est donc préférable d'accumuler plusieurs spectres (3 en moyenne)

pour limiter le bruit de fond. Le passeur d'échantillon est thermostaté à 20°C par un effet Peltier piloté directement par le logiciel Spectra Manager. Une sonde de température externe, immergée dans une cuve, permet de mesurer la température réelle de la cuve.

Des solutions de duplexes d'ADN (500 μ L, 2,5 μ M) dans un tampon Tris-HCl à pH 7,5 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂, sont placées dans des cuves en quartz à faces optiques parallèles, transparentes et de trajet optique 1 cm. Les échantillons contenant les solutions de duplexe d'ADN sont titrées par des solutions de polyamides dissouts dans du MeOH ou du DMSO à 250 μ M. Après chaque ajout de la solution de polyamides (1 - 2 μ L), les spectres de dichroïsme circulaire sont enregistrés trois fois dans une gamme de longueur d'onde 240 – 400 nm. Les soustractions de la ligne de base et du spectre de DC de l'ADN sont effectuées manuellement sur le logiciel Spectra Analysis fourni par Jasco ou après exportation des données brutes sous Excel.

G.4 Spectres de fluorescence

La fluorescence est un outil puissant pour l'analyse des propriétés dynamiques de fluorophores d'intérêt biologique. En effet, la molécule reste pendant un temps suffisamment long dans un état excité pour que son émission de fluorescence reflète son environnement ; il est donc possible de caractériser une interaction entre des molécules par l'étude des modifications des propriétés de fluorescence d'une de ces substances. À titre d'exemple, l'utilisation du phénomène de Transfert par Résonance d'Energie de Fluorescence (FRET) entre deux fluorophores résulte du transfert d'énergie d'excitation d'un donneur de fluorescence vers un accepteur et se produit par le biais d'un couplage très faible entre le moment de transition d'émission du donneur et le moment de transition d'absorption de l'accepteur.

Dans notre cas, la fluorescence des sondes synthétisées à base de polyamides est exaltée lorsqu'elles interagissent avec l'ADN, ce qui les rend intéressantes pour des applications de détection de l'ADN par microscopie de fluorescence.

Des solutions aqueuses de sondes fluorescentes à base de polyamides $(1 - 3 \mu M)$ dans un tampon Tris-HCl à pH 7,5 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂ sont placées dans des cuves en quartz transparentes et de trajet optique 1 cm. Des solutions de duplexes d'ADN sont ajoutées pour obtenir une concentration finale de 2 μ M et des spectres de fluorescence sont enregistrés avant et après chaque ajout. Les longueurs d'onde d'excitation sont 480 nm (FITC), 510 nm (Cy3) et 610 nm (Cy5) et les spectres sont enregistrés dans des intervalles de longueurs d'onde 500 - 650 nm (FITC), 530 - 750 nm (Cy3) et 630 - 750 nm (Cy5). Les fentes du spectrofluorimètre sont de 0,5 mm en excitation et de 2,5 mm en émission.

Pour les expériences de FRET, une solution de sonde *F1-Cy3* (donneur) à 2µM dans le même tampon que celui décrit ci-dessus est placée dans une cuve en quartz identique à celles précédemment décrites. Les spectres de fluorescence sont enregistrés plusieurs fois puis les sondes (accepteurs) *F2β-Cy5* ou *F4β-Cy5* sont ajoutées pour obtenir une concentration finale de 2 µM et les spectres sont également enregistrés. Ensuite, l'ajout du duplexe d'ADN à 2 µM (concentration finale) est effectué et les spectres sont enregistrés plusieurs fois. Pour finir, une titration avec des concentrations croissantes de sondes acceptrices est effectuée. Les spectres de fluorescence sont enregistrés dans la gamme de longueurs du donneur (Cy3) et de l'accepteur (Cy5) 530 – 750 nm après l'excitation du donneur (Cy3) à 510 nm.

Les intensités de fluorescence obtenues au cours des différentes expériences sont exportées, traitées (Excel) et normalisées entre 0 et 1 afin de permettre la comparaison des différents résultats.

<u>H – Études Cellulaires</u>

H.1 Culture cellulaire

H 1.1 Cellules murines NIH-3T3 et MEF

Les lignées cellulaires murines NIH-3T3 ou MEF sont cultivées dans des flasques T75 contenant 10 mL de milieu liquide (Dulbecco's modified Eagles's medium (DMEM) supplémenté par 10 % de sérum de veau fœtal (SVF - Eurobio)), sous atmosphère contrôlée (37°C, 5 % CO₂). Tous les 2 ou 3 jours, les cellules sont décollées par un traitement de 5 min à la trypsine à 37°C puis diluées et remises en suspension dans du milieu frais.

Pour les expériences sur cellules fixées, 10^5 cellules, remises en suspension dans 300 μ L de milieu de culture, sont ensemencées sur des lamelles 22×22 mm (Thermo Fisher Scientific Menzel), d'épaisseur moyenne de 0,17 ± 0,01 mm, nettoyées au préalable dans un

plasma cleaner Femto v4 (Diener electronic GmbH) et déposées dans des plaques 6 puits. Après 2h, 2 mL de milieu frais sont ajoutés dans chaque puits et, afin que les cellules adhèrent, elles sont incubées dans une étuve à 37°C jusqu'au lendemain pour obtenir une confluence d'environ 70%. Les lamelles sont ensuite lavées trois fois par du PBS 1X, fixées par une solution de paraformaldéhyde (PFA) à 4% dans du PBS 1X pendant 10 min à température ambiante et finalement lavées trois fois par du PBS 1X. Une fois fixées, les cellules sont stockées à 4°C dans du PBS 1X pendant un mois maximum.

Pour les expériences sur cellules vivantes, 3×10^4 cellules, remises en suspension dans 200 µL de milieu de culture, sont ensemencées dans des micro-lamelles de verres quadrillées « ibiTreat » (Ibidi, Biovalley) puis placées dans une étuve à atmosphère contrôlée. Après 24h, le milieu de culture est remplacé par une solution de traitement contenant 199 µL de milieu frais et 1 µL d'une solution de polyamide fluorescent à 400 µM dans du DMSO, pour avoir une concentration finale de polyamide de 2µM, puis les cellules sont incubées à 37°C (5% CO₂) pendant 24h. Avant l'acquisition des images par microscopie de fluorescence, les cellules sont lavées brièvement deux fois par du PBS 1X pour enlever l'excès de molécules fluorescentes et le PBS 1X est remplacé, sous conditions stériles, par un milieu de culture modifié pour l'imagerie, le DMEM^{gfp} (Euromedex) supplémenté par 16 µM de Rutin (Sigma Aldrich) qui est un flavonoïde de plante servant à éliminer les radicaux libres de l'oxygène (Bogdanov et al. 2012). Une incubation de 30 min à 37°C (5% de CO₂) dans ces conditions est effectuée avant la prise d'images.

H.1.2 Cellules humaines RPE-1

La lignée cellulaire humaine RPE-1 a été cultivée dans des flasques T75 contenant 10 mL de milieu liquide (Roswell Park Memorial Institute medium (RPMI)) supplémenté par 10% de SVF, sous atmosphère contrôlée. Tous les 2 ou 3 jours, les cellules sont décollées par un traitement de 5 min à la trypsine à 37°C puis diluées et remises en suspension dans du milieu frais.

Pour les expériences sur cellules fixées, les mêmes étapes que pour les cellules murines sont suivies mais le nombre de cellules utilisées est de 5×10^4 ; de même que pour les expériences sur cellules vivantes où le nombre de cellules est de $1,5 \times 10^4$.

H.1.3 Cellules humaines GM12878

La lignée cellulaire humaine GM12878 est cultivée en suspension dans des flasques T25 contenant 20 mL de milieu liquide (RPMI) supplémenté par 15% de SVF, sous atmosphère contrôlée. Tous les 2 jours, les cellules sont diluées et remises en suspension dans du milieu frais.

La fixation des cellules en suspension diffère légèrement : 2×10^6 cellules sont reprises dans 300 µL de PBS 1X et déposées sur une lamelle de verre préalablement nettoyée et traitée avec un revêtement en poly-L-lysine 0,1%. L'adhésion des cellules s'effectue en 5 min. Elles sont ensuite traitées avec 1 mL de PBS 0,3X pendant 1 min, puis fixées pendant 10 min dans une solution de PFA à 4% diluée dans du PBS 0,3X. Les cellules sont enfin lavées par 3 bains de PBS 1X et conservées dans du PBS 1X au maximum 1 mois à 4°C.

H.2 Marquage in situ des cellules fixées

Dans le but de faciliter les manipulations le jour des expériences, les lamelles sont fixées sur des lames de verre 26×76 mm un utilisant de la colle « rubber cement » (BHV). Trois types d'expériences ont été effectués.

Pour la visualisation combinée du DAPI et des MGBs fluorescents, le marquage par les polyamides fluorescents est effectué en premier en incubant les lamelles dans le noir pendant 30 min, à température ambiante dans 50 μ L d'une solution de polyamide à 2 μ M. Ensuite, les cellules sont lavées brièvement par du PBS 1X avant une étape de post-fixation qui consiste à incuber les cellules dans 100 μ L d'une solution de PFA à 2% dans du PBS 1X pendant 10 min. Puis trois rinçages dans une solution de Triton X–100 0,05% dans du PBS 1X sont effectués. Le marquage de l'ADN est réalisé par l'addition d'une solution de DAPI à 70 nM (Sigma-Aldrich) pendant 5 min à température ambiante. Après une brève étape de lavage dans du PBS 1X, les lamelles sont montées donc déposées sur une lame après l'ajout d'une goutte de milieu de montage PPD8 préservant la fluorescence (Sigma-Aldrich, solution à 1mg / mL dans un mélange de PBS 1X / glycérol 90% à pH 8 ajusté par un tampon carbonate / bicarbonate). Enfin les lamelles sont scellées avec du vernis, conservées à 4°C à court terme et à -20°C à long terme.

Pour la visualisation combinée des MGBs fluorescents et des sondes FISH, les cellules sont perméabilisées avec une solution de Triton X-100 à 0,5% dans du PBS 1X, lavées trois fois par une solution de Triton X-100 à 0,05% dans du PBS 1X, traitées par une solution de 0,1 N HCl pendant 2 min, rincées deux fois, équilibrées 5 min dans du SSC 2X et traitées par une solution de SSC 2X / 50% de formamide déionisé à température ambiante pendant au moins 30 min avant l'hybridation. Les sondes FISH proviennent de la société Eurogenetec (Belgique) et sont des oligonucléotides marqués en 3' par le fluorophore Cy3. Elles sont composées soit par des séquences mixtes ADN / LNA pour cibler la répétition péricentromérique satellite majeure murine («Sat Maj » 5'gACgACtTGaAAaATgACgAAaTC -3', les lettres minuscules représentent les modifications LNAs), soit uniquement par des LNAs pour cibler la répétition centromérique α-satellite humaine (« PCY » 5'- tttgtgatgtgtg- 3'). Les solutions d'hybridation sont préparées en diluant les sondes oligonucléotidiques à une concentration finale de 1 µM dans une solution d'hybridation contenant 50% de formamide déionisé, de la solution de Denhardt 1X (réactifs de chez Sigma Aldrich), du sulfate de dextran 10%, 0,1% de dodécylsulfate de sodium (SDS) et diluée dans un tampon SSC 2X à pH 5,6. Sur le côté des lamelles, 45 µL de la solution d'hybridation sont déposés. Les lamelles sont chauffées en atmosphère humide pendant 3 min à 85°C, puis lentement refroidies à 37°C à une vitesse de 1°C / s pour renaturer l'ADN. Ceci est effectué dans un appareil PCR (MJ Research). Chaque lamelle est lavée deux fois dans du tampon SSC 2X à 63°C et une fois dans du SSC 2X à température ambiante. Le marquage par les polyamides fluorescents est ensuite effectué comme décrit précédemment. Les lamelles sont enfin montées avec du PPD8 sans marquage préalable de l'ADN.

Pour la visualisation combinée des MGBs fluorescents et des protéines (immunocytochimie), les cellules sont perméabilisées avec une solution de Triton X-100 à 0,1% dans du PBS 1X pendant 5 min à température ambiante puis lavées deux fois dans du PBS 1X. Après une étape de saturation dans une solution de « blocking reagent » à 1,5% (Roche Applied Science ; poudre, à laquelle est ajoutée 0,05% de Tween 20, et diluée dans du SSC 2X à pH 6,3), l'immunomarquage est effectué en incubant les anticorps primaires pendant 90 min à température ambiante en chambre humide, en lavant trois fois avec la solution de Triton X-100 à 0,1% dans du PBS 1X, en incubant les anticorps secondaires 45 min à température ambiante en chambre humide dans le noir puis en lavant trois fois par la solution de Triton X-100 à 0,1% dans du PBS 1X.

Tous les anticorps sont dilués dans une solution de saturation contenant 1% de « blocking reagent ». Les anticorps primaires utilisés sont les suivants :

- CenpB (clone H65, 1/200, polyclonal de lapin, Santa Cruz) pour la détection des centromères
- H₃K₉me₃ (1/400, polyclonal de lapin, Upstate, Millipore) pour la détection des régions péricentromériques murines

Les anticorps secondaires sont dilués à 1/200 dans la solution de saturation et sont les suivants :

- Anticorps anti-lapin produit chez la chèvre et conjugué au Cy3 (Jackson ImmunoResearch)
- Anticorps anti-lapin produit chez la chèvre et conjugué au Cy5 (Jackson ImmunoResearch)

Dans le cas des expériences d'immuno-FISH, une expérience d'immunocytochimie est suivie d'une expérience de FISH, après une étape intermédiaire de post-fixation (PFA 2% / PBS 1X pendant 10 min puis trois fois PBS 1X / 0,1% Triton X-100). Pour la détection simultanée des sondes FISH et des MGBs fluorescents, l'expérience de FISH est effectuée préalablement au marquage par les MGBs fluorescents.

H.3 Acquisition des images par microscopie

Les acquisitions d'images sont réalisées un utilisant un microscope inversé à épifluorescence de marque Zeiss (Axio Observer Z1), équipé d'un objectif à immersion à huile plan-apochromat 63x 1,4 NA et des filtres suivants (**Tableau 14**) :

Tableau 14 : Les différents filtres de microscopie utilisés pour l'acquisition des images					
Fluorophore/Colorant	Excitation	Dichroïque	Emission	Source Lumineuse	
DAPI	G365	FT395	BP445/50	LED : 365 nm	
FITC	BP470/40	FT495	BP525/50	LED : 470 nm	
Cy3	BP546/10	FF555	BP583/22	HXP120 Metal Halide	
Cy5	BP643/20	FF660	BP684/24	LED : 625 nm	

L'huile d'immersion utilisée a pour indice de réfraction 1,518 à 23°C. Le logiciel ZEN 2012 est utilisé pour l'acquisition et le traitement des images. Un filtre gaussien (1 :1 :1) est appliqué à toutes les images pour lisser le signal (Bolte & Cordelières 2006). Les quantifications et les mesures de Pearson ont été réalisées sur les images originales par François Loll.

Le coefficient de Pearson est obtenu à l'aide du module JACoP pour ImageJ. Les différents canaux sont comparés deux à deux, pixel par pixel, au niveau de leurs intensités respectives. Le résultat est compris entre 0 (pas de corrélation) et 1 (corrélation parfaite).

Liste des publications scientifiques co-signées

- Nozeret K, Loll F, Escudé C, Boutorine A. Polyamide Fluorescent Probes for Visualization of Repeated DNA Sequences in Living Cells. *ChemBioChem.* 2015 ; 16 (4) : 549-554.
- Nozeret K, Bonan M, Yarmoluk S.M, Novopashina D.S, Boutorine A.S. Synthesis of mouse centromere-targeted polyamides and physico-chemical studies of their interaction with the target double-stranded DNA. *Bioorganic & Medicinal Chemistry*. 2015; 123 (17): 15932-15945.
- Boutorine A.S, Novopashina D.S, Krasheninina, O.A, Nozeret K, Venyaminova A.G. Fluorescent Probes for Nucleic Acid Visualization in Fixed and Live Cells. *Molecules*. 2013; 18: 15357-15397.

Annexes

Annexes

Annexe 1 : Structures en « épingle à cheveux antiparallèle » et caractéristiques chimiques des MGBs murins

✓ MGB F1 :

 $\begin{array}{l} MGB \ F1: C_{63}H_{80}N_{22}O_{11}\\ M=1320,64 \ g/mol\\ SM \ (ESI+): M=1320,68 \ g/mol\\ Temps \ de \ rétention \ sur \ HPLC \ analytique \ Tr=24,3 \ min\\ Rendement \ final \ estimé: 82 \ \% \end{array}$

✓ MGB F2 :

 $\label{eq:mgb} \begin{array}{l} MGB \ F2: C_{62}H_{79}N_{23}O_{11} \\ M = 1321,63 \ g/mol \\ SM \ (ESI+): M=1321,64 \ g/mol \\ Temps \ de \ rétention \ sur \ HPLC \ analytique: Tr = 12,73 \ min \\ Rendement \ final \ estimé: 74 \ \% \end{array}$

✓ MGB F2 β :

MGB F2 β : C₅₉H₇₈N₂₂O₁₁

M = 1270,62 g/mol SM (ESI+) : M= 1270,66 g/mol Temps de rétention sur HPLC analytique : Tr = 20,8 min Rendement quantitatif par rapport à la capacité théorique de la résine

Le synthon β *interne est représenté en noir.*

✓ MGB F3 :

Hal

 $\begin{array}{l} MGB~F3:C_{62}H_{79}N_{23}O_{11}\\ M=1321,63~g/mol\\ SM~(ESI+):M=1321,63~g/mol\\ Temps de rétention sur HPLC analytique:Tr=12,76~min\\ Rendement final estimé:24~\% \end{array}$

✓ MGB F3 β :

MGB F3 β : C₅₉H₇₈N₂₂O₁₁ M = 1270,62 g/mol SM (ESI+) : M= 1270,66 g/mol Temps de rétention sur HPLC analytique : Tr = 12,6 min

Rendement final estimé : 63 %

Le synthon β *interne est représenté en noir.*

✓ MGB F4 :

 $\begin{array}{l} MGB \ F4: C_{63}H_{80}N_{22}O_{11}\\ M=1320,64 \ g/mol\\ SM \ (ESI+): M=1320,67 \ g/mol\\ Temps \ de \ rétention \ sur \ HPLC \ analytique: Tr=12,76 \ min\\ Rendement \ final \ estimé: 67 \ \% \end{array}$

✓ MGB F4 β :

 $\begin{array}{l} MGB \ F4\beta: C_{60}H_{79}N_{21}O_{11}\\ M=1269,63 \ g/mol\\ SM\ (ESI+): M=1269,66 \ g/mol\\ Temps\ de\ rétention\ sur\ HPLC\ analytique: Tr=12,69\ min\\ Rendement\ final\ estimé: 60\ \% \end{array}$

Le synthon β *interne est représenté en noir.*

✓ MGB F4 $\beta\gamma$: MGB F4 $\beta\gamma$: C₆₄H₈₆N₂₂O₁₂ M = 1354,68 g/mol SM (ESI+) : M= 1354,68 g/mol Temps de rétention sur HPLC analytique : Tr = 12,63 min Rendement final estimé : 49 %

Le synthon β interne est représenté en noir et le synthon γ supplémentaire en rouge.

✓ MGB F5 :

 $\begin{array}{l} MGB \ F5: C_{64}H_{81}N_{21}O_{11}\\ M=1248, 61 \ g/mol\\ SM \ (ESI+): M=1248, 62 \ g/mol\\ Temps \ de \ rétention \ sur \ HPLC \ analytique : Tr=12, 75 \ min\\ Rendement \ final \ estimé : 90 \ \% \end{array}$

✓ MGB F6 :

H

 $\begin{array}{c} MGB \ F6: C_{62}H_{79}N_{23}O_{11}\\ M=1321,63 \ g/mol\\ SM \ (ESI+): M=1321,64 \ g/mol\\ Temps \ de \ rétention \ sur \ HPLC \ analytique: Tr=12,56 \ min\\ Rendement \ final \ estimé: 76 \ \% \end{array}$

Annexe 2 : Caractérisation des MGBs murins

✓ Exemple du chromatogramme analytique du MGB murin F4 brut (avec un temps de rétention Tr = 12,76 min) :



✓ Exemple du spectre de masse du MGB murin F4 (M = 1320,67 g/mol) :



✓ Exemple du spectre d'absorbance UV-visible du MGB murin F4 (absorbance caractéristique des MGBs entre 280 et 340 nm) :



Les détails sur les conditions des analyses sont donnés dans le **Chapitre IV** (**Matériel** et Méthodes).

Annexe 3 : Formules et caractéristiques des MGBs murins fonctionnalisés en N-term et des deux tandems de polyamides murins



✓ Formules des MGBs modifiés composant les tandems et des deux tandems de polyamides dans une orientation « tête à tête »

K1 et K1 γ sont les analogues de F1 protégés par un groupement Boc et fonctionnalisés en N-term par un azoture. Ils diffèrent par la longueur du lien N-term. K2 β et K4 β sont les analogues de F2 β et F4 β , respectivement. Ils portent des groupements alcynes en N-term. T1 et T2 sont des tandems antiparallèles de polyamides, déprotégés et construits par une réaction de cycloaddition 1,3-dipolaire catalysée par le cuivre entre K1 et K2 β (T1) ou K1 γ et K4 β (T2), respectivement (adapté de (Nozeret, Bonan, et al. 2015)).

tandems murins					
Nom	Formule brute	Masse molaire	Tr UDI C analytique	Masse molaire	
		calculée	(min)	obtenue par SM	
		(g/mol)		(g/mol)	
Boc-K1	C69H89N25O13	1475,71	26,44*	1475,73	
Boc-K1γ	C73H96N26O14	1560,76	11,77 [‡]	1560,76	
Κ2β	C64H82N22O12	1350,65	23,75*	1350,66	
Κ4β	C65H83N21O12	1349,65	$10,71^{\dagger}$	1349,66	
T1	C128H163N47O23	2726,3	23,72*	2726,36	
T2	C133H171N47O24	2810,36	13,44 [‡]	2810,39	

Tableau 15 : Caractéristiques des MGBs murins fonctionnalisés en N-term et des tandems murins

*Pour l'analyse des composants de T1 et du tandem T1, un gradient d'HPLC différent a été utilisé (0 – 100% acétonitrile en 45 min).

⁺Pour l'analyse des autres produits, le gradient standard décrit dans le **Chapitre IV** (**Matériels et Méthodes**) a été utilisé.

Annexe 4 : Caractérisation des tandems de polyamides murins

✓ Exemple du chromatogramme analytique du tandem T2 (K1γK4β) brut (avec un temps de rétention Tr = 13,44 min) :



✓ Exemple du spectre de masse du tandem T2 (M = 2810,39 g/mol) :



✓ Exemple du spectre d'absorbance UV-visible du tandem T2 (absorbance caractéristique des MGBs entre 280 et 340 nm) :



Les détails sur les conditions des analyses sont donnés dans le **Chapitre IV** (**Matériel** et Méthodes).

Tableau 16 : Noms, structures, masses molaires (g/mol) et longueurs d'onde d'excitation (λ_{ex}, nm) et d'émission (λ_{em}, nm) des différents fluorophores					
Nom	Structure	Masse molaire	λ_{ex}	λ_{em}	
Coumarine MM-14 (MM14)	HO O O	203,2	350	495	
Cyanine3-NHS ester (Cy3)		554,7	550	570	
Cyanine3-sulfate-NHS ester (Cy3-sulf)	$^{-O_{3}S}$	726,84	550	570	
Cyanine3-isothiocyanate (Cy3-ITC)		555,32 C=S	550	570	
Cyanine5-NHS ester (Cy5)		580,74	649	670	
Cyanine5-sulfate-NHS ester (Cy5-sulf)		752,87	649	670	
Fluorescéine-ITC (FITC)	O O HO HO N=C=S	391,4	494	521	

Annexe 5 : Structures et caractéristiques chimiques des différents fluorophores utilisés

Nom	Structure	Masse molaire	λ_{ex}	λ_{em}
Fluorescéine-NHS ester (fluo)		473,39	494	521
Oregon Green 488-NHS (OG-NHS)	$HO \xrightarrow{O} \xrightarrow{O} \xrightarrow{F} \xrightarrow{O} \xrightarrow{O} \xrightarrow{F} \xrightarrow{O} \xrightarrow{O} \xrightarrow{O} \xrightarrow{F} \xrightarrow{O} \xrightarrow{O} \xrightarrow{O} \xrightarrow{O} \xrightarrow{O} \xrightarrow{O} \xrightarrow{O} O$	509,38	496	524
Oregon Green 488-ITC (OG-ITC)	$HO \xrightarrow{O} \xrightarrow{O} \xrightarrow{F} \xrightarrow{O} \xrightarrow{F} \xrightarrow{O} \xrightarrow{O} \xrightarrow{F} \xrightarrow{O} \xrightarrow{O} \xrightarrow{O} \xrightarrow{F} \xrightarrow{O} \xrightarrow{O} \xrightarrow{O} \xrightarrow{F} \xrightarrow{F} \xrightarrow{O} \xrightarrow{F} \xrightarrow{F} \xrightarrow{O} \xrightarrow{F} \xrightarrow{F} \xrightarrow{F} \xrightarrow{F} \xrightarrow{F} \xrightarrow{F} \xrightarrow{F} F$	510,07	496	524
S24	N-CO OH	767,33	450	495
Tétraméthylrhodamine- NHS ester (TMR-NHS)		528,18	529	596
Tétraméthylrhodamine- ITC (TMR-ITC)	$ \overset{N_{+}^{+}}{\underset{O}{\overset{O}}} \overset{O}{\underset{O}{\overset{N_{+}}{\overset{O}}}} \overset{O}{\underset{O}{\overset{O}}} \overset{N_{+}}{\underset{O}{\overset{O}}} \overset{O}{\underset{O}{\overset{N_{+}}{\overset{O}}}} \overset{O}{\underset{O}{\overset{O}}} \overset{N_{+}}{\underset{O}{\overset{O}}} \overset{O}{\underset{O}{\overset{O}}} \overset{O}{\underset{O}} \overset{O}{\underset{O}}} \overset{O}{\underset{O}} \overset{O}{\underset{O}} \overset{O}}{} \overset{O}{\underset{O}} \overset{O}}{} \overset{O}} \overset{O}}{} \overset{O}{} \overset{O}} \overset{O}{} \overset{O}} \overset{O}{} \overset{O}} \overset{O}}{} \overset{O}} \overset{O}}{} \overset{O}} \overset{O}}{} \overset{O}} \overset{O}}{} \overset{O}} \overset{O}}{} \bullet$	528,18	529	596
TO		504,11	505	530
TO-azoture	S N ⁺ N ⁻ N ₃	324,13	505	530

Tableau 17 : Caractéristiques des sondes murines composées par des MGBs standard					
Nom	Masse molaire calculée (g/mol)		Tr HPLC	Masse molaire obtenue	
			analytique (min)	par SM (g/mol)	
F1-Cy3	$[C_{93}H_{115}N_{24}O_{12}]^+$	1759,91	14,35	1758,89*	
F1-Cy3-sulf	$[C_{94}H_{115}N_{24}O_{18}S_2]^{-1}$	1931,83	20,8	1932,86*	
F1-Cy3-ITC	$[C_{97}H_{123}N_{26}O_{12}S]^+$	1875,95	13,4	1875,9	
$F1\gamma^{(R)NH2}$ -Cy3	$[C_{93}H_{116}N_{25}O_{12}]^+$	1774,92	13,53	1773,94*	
$F1\gamma^{(R)NH2}$ -MM14	C77H90N26O15	1619,71	10,06	1619,65	
$F1\gamma^{(R)NH2}$ -TO	$[C_{85}H_{104}N_{28}O_{12}S]^+$	1740,81	10	1741,02	
F1-TO	$[C_{87}H_{103}N_{24}O_{12}S]^+$	1707,79	14,3	1708,05	
F1-S24	$[C_{83}H_{103}N_{24}O_{12}]^+$	1627,82	13,34	1627,78	
F1-FITC	C ₈₄ H ₉₁ N ₂₃ O ₁₆ S	1709,67	13,76	1709,67	
F1-TMR-NHS	$[C_{88}H_{101}N_{24}O_{15}]^+$	1733,79	12,24	1732,79*	
F1-TMR-ITC	$C_{92}H_{108}N_{26}O_{15}S$	1848,82	10,7	1848,84	
F1-OG-NHS	$C_{84}H_{88}F_2N_{22}O_{17}$	1714,67	13,25	1714,68	
F1-OG-ITC	$C_{88}H_{96}F_2N_{24}O_{17}S$	1830,71	13,25	1830,70	
F2-FITC	C ₈₃ H ₉₀ N ₂₄ O ₁₆ S	1710,67	12,52	1710,68	
F2β-Cy5-sulf	$[C_{92}H_{117}N_{24}O_{18}S_2]^{-1}$	1909,84	23,8	1908,89*	
F3-FITC	C ₈₃ H ₉₀ N ₂₄ O ₁₆ S	1710,67	12,51	1710,67	
F4-Cy5	$[C_{95}H_{117}N_{24}O_{12}]^+$	1785,93	14,69	1784,93*	
F4β-Cy5	$[C_{92}H_{116}N_{23}O_{12}]^+$	1734,92	14, 61	1733,92*	
F4βγ-Cy5	$[C_{96}H_{123}N_{24}O_{13}]^+$	1819,97	14,68	1818,97*	
F4βγ-Cy5-sulf	$[C_{97}H_{123}N_{24}O_{19}S_2]^{-1}$	1991,88	14,7	1992,88*	
F4β-fluo	C ₈₁ H ₈₉ N ₂₁ O ₁₇	1627,67	13,84	1627,67	
F4β-FITC	$C_{81}H_{90}N_{22}O_{19}S$	1658,66	13,2	1658,65	
F5-Cy5	$[C_{93}H_{113}N_{22}O_{11}]^+$	1713,9	14,66	1712,9*	
F6-Cy5	$[C_{94}H_{116}N_{25}O_{12}]^+$	1786,92	14,72	1785,93*	

Annexe 6 : Caractéristiques des différentes sondes murines synthétisées à partir des MGBs standard

*La différence de 1 fréquemment observée est due aux échanges de protons dans les conditions d'analyse de la spectrométrie de masse.

Annexe 7 : Caractérisation des sondes murines construites à partir des MGBs standard

 ✓ Exemple du chromatogramme analytique de la sonde murine F4β-FITC brute (avec un temps de rétention Tr = 13,2 min) :



✓ Exemple du spectre de masse de la sonde murine F4 β -FITC (M = 1658,65 g/mol) :



 ✓ Exemple du spectre d'absorbance UV-visible de la sonde murine F4β-FITC (absorbances caractéristiques : MGB (280-340 nm) et FITC (400-530 nm)) :



Les détails sur les conditions des analyses sont donnés dans le **Chapitre IV** (**Matériel** et Méthodes).

Annexe 8 : Caractéristiques des sondes fluorescentes murines construites à partir des tandems murins

Tableau 18 : Caractéristiques des tandems fluorescents murins						
Nom	Masse molaire calculée (g/mol)		Tr HPLC	Masse molaire obtenue		
INOIII			analytique (min)	par SM (g/mol)		
T1-Cy3	[C158H198N49O24] ⁺	3165,58	28,78	3164,58*		
T2-FITC	C154H182N48O29S	3199,40	14,28	3199,42		

*La différence de 1 fréquemment observée est due aux échanges de protons dans les conditions d'analyse de la spectrométrie de masse.

Annexe 9 : Courbes de fusion du Duplexe 1 et ses dérivées premières en présence et en absence du polyamide murin F1



Les expériences de dénaturation thermique ont été effectuées dans un tampon cacodylate de sodium (50 mM) à pH 6,2 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. Un excès (8 fois) de MGB F1 est mis en présence du **Duplexe 1** (1,3 μ M) qui possède 4 sites de liaison pour F1 (soit 2 molécules de MGB par site de liaison). A/ Courbes de fusion du **Duplexe 1** seul à 260 nm (courbe bleue) et en présence de F1 à 260 nm (courbe rouge) et à 335 nm (courbe verte) et B/ leurs dérivées premières. La température de fusion du **Duplexe 1** augmente de 57°C à 81°C ($\Delta T_m = 24$ °C) en présence de F1.

Annexe 10 : Courbes de fusion du Duplexe 1 et ses dérivées premières en présence et en absence du tandem murin T1



Les expériences de dénaturation thermique ont été effectuées dans un tampon cacodylate de sodium (50 mM) à pH 6,2 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. Un excès (6 fois) de tandem **T1** est mis en présence du **Duplexe 1** (1,3 μ M) qui possède un site de liaison pour **T1**. A/ Courbes de fusion du **Duplexe 1** seul à 260 nm (courbe bleue) et en présence de **T1** à 260 nm (courbe rouge) et à 335 nm (courbe verte) et B/ leurs dérivées premières. La température de fusion du **Duplexe 1** augmente mais le ΔT_m est difficile à définir compte tenu de la largeur des pics des courbes rouge et verte.



Annexe 11 : Exemples de « gel-retard » avec le duplexe contrôle et les MGBs murins

Les électrophorèses ont été effectuées sur des gels de polyacrylamide à 20% non dénaturant dans du tampon TBE à 4°C. Le **Duplexe 3** (contrôle qui ne possède aucun site de liaison pour chacun des MGBs) a été incubé en présence de différents MGBs murins. Le brin riche en T du duplexe est marqué à la fluorescéine. Les concentrations des polyamides sont indiquées audessus de chaque piste des gels et celle du **Duplexe 3** est de 0,52 μ M dans chaque cas. Les schémas sur la droite symbolisent la position du duplexe après la migration. Aucune bande retardée n'a été observée donc aucune preuve d'une quelconque interaction avec le duplexe contrôle n'a été mise en évidence. Annexe 12 : Spectres de dichroïsme circulaire différentiels des complexes ADN / tandem de polyamide murin avec les duplexes courts mutés sur un site



Les expériences de dichroïsme circulaire ont été effectuées dans un tampon Tris-HCl (50 mM) avec 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. Spectres de DC différentiels des complexes ADN : tandem, A/ **Duplexe 5** : **T2** et B/ **Duplexe 6** : **T2**. Les spectres de DC de l'ADN sont soustraits. 2,5 µM des solutions de duplexes d'ADN sont titrés par une solution de tandem à 250 µM (voir description dans le Chapitre IV – Matériels et méthodes). Les concentrations de tandem de polyamides après chaque ajout sont indiquées à droite de chaque figure. Courbes de titration : ellipticité des complexes C/ **Duplexe 5** : **T2** et D/ **Duplexe 6** : **T2**, à 333 nm aux concentrations de polyamides indiquées.





Les expériences de dichroïsme circulaire ont été effectuées dans un tampon Tris-HCl (50 mM) avec 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. Spectres de DC différentiels des complexes ADN : sonde fluorescente, A/ Duplexe 2 : F1-Cy3, B/ Duplexe 2 : F1-FITC, C/ Duplexe 2 : F4 β -FITC et D/ Duplexe 1 : T2-FITC. Les spectres de DC de l'ADN sont soustraits. 2,5 μ M des solutions de duplexes d'ADN sont titrés par des solutions de sondes fluorescentes à 250 μ M (voir description dans le Chapitre IV – Matériels et méthodes). Les concentrations de sondes après chaque ajout sont indiquées à droite de chaque figure.

Annexe 14 : Expériences de FRET entre les sondes murines F1-Cy3 et F2β-Cy5 en présence du duplexe cible d'ADN



Les expériences ont été réalisées dans un tampon Tris-HCl (10 mM) avec 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂ à pH 7,5. La longueur d'onde d'excitation était de 510 nm et l'émission a été enregistrée dans la gamme 530 - 750 nm. Les concentrations de F1-Cy3 et du Duplexe 1 sont de 2 μ M. En bleu, le spectre de fluorescence de la sonde donneuse F1-Cy3 seule. En rouge les sondes donneuse F1-Cy3 et acceptrice F2 β -Cy5 (2 μ M) sans l'ADN. En vert, les sondes F1-Cy3 et F2 β -Cy5 en présence du Duplexe 1. Le dégradé de verts symbolise l'augmentation de la concentration de F2 β -Cy5 : en vert clair - 2 μ M, en vert foncé - 4 μ M et en vert très foncé - 6 μ M. En excitant la sonde donneuse F1-Cy3 à 510 nm, une émission de fluorescence à 660 nm typique de Cy5 (sonde acceptrice) apparait : c'est l'effet FRET.

Annexe 15 : Structures en « épingle à cheveux antiparallèle » et caractéristiques chimiques des MGBs humains



✓ MGB Hb2 :



Temps de rétention sur HPLC analytique Tr = 12,52 min Rendement final estimé : 22 %

✓ MGB Hbb3 :



Les synthons β internes sont représentés en noir.

Annexe 16 : Caractérisation des MGBs humains

✓ Exemple du chromatogramme analytique du MGB humain Hb1 brut (avec un temps de rétention Tr = 12,02 min) :



✓ Exemple du spectre de masse du MGB humain Hb1 (M = 1270,62 g/mol) :



✓ Exemple du spectre d'absorbance UV-visible du MGB humain Hb1 (absorbance caractéristique des MGBs entre 280 et 340 nm) :



Les détails sur les conditions des analyses sont donnés dans le **Chapitre IV** (**Matériel** et Méthodes).
Tableau 19 : Caractéristiques des sondes humaines synthétisées à partir des MGBs				
Nom	Formule brute	Masse molaire calculée (g/mol)	Tr HPLC analytique (min)	Masse molaire obtenue par SM (g/mol)
H1-FITC	$C_{83}H_{90}N_{24}O_{16}S$	1710,67	13,20	1710,67
Hb1-FITC	$C_{80}H_{89}N_{23}O_{16}S$	1659,66	13,29	1659,68
Hbb1-FITC	C77H88N22O16S	1608,65	13,03	1608,65
Hb2-FITC	$C_{100}H_{111}N_{31}O_{20}S$	2097,83	13,48	2097,84
Hbb2-FITC	$C_{100}H_{111}N_{31}O_{20}S$	2097,83	13,18	2097,83
Hb3-FITC	$C_{112}H_{123}N_{35}O_{22}S$	2341,93	13,53	2341,96
Hbb3-FITC	$C_{112}H_{123}N_{35}O_{22}S$	2341,93	13,32	2341,95

Annexe 17 : Caractéristiques des différentes sondes humaines synthétisées à partir des MGBs

Annexe 18 : Caractérisation des sondes humaines

 ✓ Exemple du chromatogramme analytique de la sonde humaine Hb1-FITC brute (avec un temps de rétention Tr = 13,29 min) :



✓ Exemple du spectre de masse de la sonde Hb1-FITC (M = 1659,68 g/mol) :



 ✓ Exemple du spectre d'absorbance UV-visible de la sonde humaine Hb1-FITC (absorbances caractéristiques : MGB (280-340 nm) et FITC (400-530 nm)) :



Les détails sur les conditions des analyses sont donnés dans le **Chapitre IV** (**Matériel** et Méthodes).





Les expériences de dichroïsme circulaire ont été effectuées dans un tampon Tris-HCl (50 mM) à pH 7,5 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. Spectres de DC différentiels des complexes ADN : MGB, A/ Duplexe 3 : H1, B/ Duplexe 3 : Hb1, C/ Duplexe 3 : Hb2 et D/ Duplexe 3 : Hb3. Les spectres de DC de l'ADN sont soustraits. Des solutions de duplexe d'ADN à 2,5 μ M sont titrées par des solutions de polyamides à 250 μ M (voir description dans le Chapitre IV – Matériels et méthodes). Les concentrations de polyamides après chaque ajout sont indiquées à droite de chaque figure. Pour faciliter la comparaison, les échelles sont similaires à celles présentées dans la Figure 52 (voir p.142) pour les expériences de dichroïsme circulaire des complexes entre le Duplexe 7 et les différents MGBs humains.

Annexe 20 : Spectres de dichroïsme circulaire différentiels des complexes ADN / sonde fluorescente humaine



Les expériences de dichroïsme circulaire ont été effectuées dans un tampon Tris-HCl (50 mM) à pH 7,5 contenant 50 mM NaCl, 50 mM KCl et 1 mM MgCl₂. Spectres de DC différentiels des complexes ADN : MGB, A/ Duplexe 7 : H1-FITC, B/ Duplexe 3 : H1-FITC, C/ Duplexe 7 : Hb1-FITC, D/ Duplexe 3 : Hb1-FITC, E/ Duplexe 7 : Hb2-FITC, F/ Duplexe 3 : Hb2-FITC, G/ Duplexe 7 : Hb3-FITC et H/ Duplexe 3 : Hb3-FITC. Les spectres de DC de l'ADN sont soustraits. Des solutions des duplexes d'ADN à 2,5 μ M sont titrées par des solutions de polyamides à 250 μ M (voir description dans le Chapitre IV – Matériels et méthodes). Les concentrations de polyamides après chaque ajout sont indiquées à droite de chaque figure.



Annexe 21 : Expériences de colocalisation sur des cellules humaines RPE-1 fixées

Images de cellules RPE-1 fixées traitées par 2 μ M de **Hb2-FITC** pendant 30 min et le DAPI ou la sonde « PCY » qui est un marqueur des séquences α -satellite humaines. En haut à gauche, l'expérience de marquage fluorescent de la sonde **Hb2-FITC** et du DAPI. En bas à gauche, la superposition des signaux de **Hb2-FITC** (en vert) et du DAPI (en bleu). Les profils de l'intensité de ces signaux sont donnés dans l'encadré de gauche. En haut à droite, l'expérience de marquage fluorescent par **Hb2-FITC** et de FISH avec la sonde « PCY » oligonucléotidique centromérique contenant des LNAs. En bas à droite, la superposition des signaux de **Hb2-FITC** (en vert) et de « PCY » (en rouge). Les profils de l'intensité de ces signaux sont donnés dans l'encadré de droite (les détails des expériences sont donnés dans le **Chapitre IV – Matériels et méthodes**).

Bibliographie

Bibliographie

- Aguirre-Lavin, T., Adenot, P., Bonnet-Garnier, A., Lehmann, G., Fleurot, R., Boulesteix, C., Debey, P. and Beaujean, N. (2012) 3D-FISH analysis of embryonic nuclei in mouse highlights several abrupt changes of nuclear organization during preimplantation development. *BMC Dev. Biol.*, 12 : 30 (1-19).
- Alcobia, I., Dila, R. and Parreira, L. (2000) Spatial associations of centromeres in the nuclei of hematopoietic cells : evidence for cell-type-specific organizational patterns. *Blood*, 95, 1608-1615.
- Alexandrov, I., Kazakov, A. and Tumeneva, I. (2001) Alpha-satellite DNA of primates : old and new families. *Chromosoma*, 110, 253-266.
- Alexandrov, I.A., Medvedev, L.I., Mashkova, T.D., Kisselev, L.L., Romanova, L.Y. and Yurov, Y.B. (1993) Definition of a new alpha satellite suprachromosomal family characterized by monomeric organization. *Nucl. Acids Res.*, 21, 2209-2215.
- Almouzni, G. and Probst, A.V. (2011) Heterochromatin maintenance and establishment : Lessons from the mouse pericentromere. *Nucleus*, 2, 332-338.
- Alvarez, D., Chou, J.C., Latella, L., Zeitlin, S.G., Ku, S., Puri, P.L., Dervan, P.B. and Gottesfeld, J.M. (2006) A two-hit mechanism for pre-mitotic arrest of cancer cell proliferation by a polyamide-alkylator conjugate. *Cell Cycle*, 5, 1537-1548.
- Amor, D.J., Kalitsis, P., Sumer, H. and Choo, K.H.A. (2004) Building the centromere : from foundation proteins to 3D organization. *Trends Cell. Biol.*, 14, 359-368.
- Anandhakumar, C., Li, Y., Kizaki, S., Pandian, G.N., Hashiya, K., Bando, T. and Sugiyama, H. (2014) Next-generation sequencing studies guide the design of pyrrole-imidazole polyamides with improved binding specificity by the addition of β-alanine. *ChemBiochem*, 15, 2647-2651.
- Ansari, A.Z., Mapp, A.K., Nguyen, D.H., Dervan, P.B. and Ptashne, M. (2001) Towards a minimal motif for artificial transcriptional activators. *Chem. Biol.*, 8, 583-592.
- Anton, T., Bultmann, S., Leonhardt, H. and Markaki, Y. (2014) Visualization of specific DNA sequences in living mouse embryonic stem cells with a programmable fluorescent CRISPR / Cas system. *Nucleus*, 5, 163-172.
- Arcamone, F., Penco, S., Orezzi, P., Nicolella, V. and Pirelli, A. (1964) Structure and synthesis of distamycin A. *Nature*, 203, 1064-1065.
- Arora, P.S., Ansari, A.Z., Best, T.P., Ptashne, M. and Dervan, P.B. (2002) Design of artificial transcriptional activators with rigid poly-L-proline linkers. J. Amer. Chem. Soc., 124, 13067-13071.
- Asamitsu, S., Kawamoto, Y., Hashiya, F., Hashiya, K., Yamamoto, M., Kizaki, S., Bando, T. and Sugiyama, H. (2014) Sequence-specific DNA alkylation and transcriptional inhibition by long-chain hairpin pyrrole-imidazole polyamide-chlorambucil conjugates targeting CAG/CTG trinucleotide repeats. *Bioorg. Med. Chem.*, 22, 4646-4657.

- Babu, B., Liu, Y., Plaunt, A., Riddering, C., Ogilvie, R., Westrate, L., Davis, R., Ferguson, A., Mackay, H., Rice, T., Chavda, S., Wilson, D., Lin, S., Kiakos, K., Hartley, J.A. and Lee, M. (2011) Design, synthesis and DNA binding properties of orthogonally positioned diamino containing polyamide f-IPI. *Biochem. Biophys. Res. Commun.*, 404, 848-852.
- Baird, E.E. and Dervan, P.B. (1996) Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. J. Amer. Chem. Soc., 118, 6141-6146.
- Baliga, R., Baird, E.E., Herman, D.M., Melander, C., Dervan, P.B. and Crothers, D.M. (2001) Kinetic consequences of covalent linkage of DNA binding polyamides. *Biochemistry*, 40, 3-8.
- Bando, T., Fujimoto, J., Minoshima, M., Shinohara, K.I., Sasaki, S., Kashiwazaki, G., Mizumura, M. and Sugiyama, H. (2007) Detection of CAG repeat DNA sequences by pyrene-functionalized pyrrole-imidazole polyamides. *Bioorg. Med. Chem.*, 15, 6937-6942.
- Barni, E., Savarino, P., Larovere, R. and Viscardi, G. (1986) Long chain heterocyclic dyes. Part one. Hydrophobic structures. *J. Heterocycl. Chem.*, 23, 209-221.
- Barrett, M.P., Gemmell, C.G. and Suckling, C.J. (2013) Minor groove binders as antiinfective agents. *Pharmacol. Ther.*, 139, 12-23.
- Baskin, J.M., Prescher, J.A., Laughlin, S.T., Agard, N.J., Chang, P.V., Miller, I.A., Lo, A., Codelli, J.A. and Bertozzi, C.R. (2007) Copper-free click chemistry for dynamic in vivo imaging. *Proc. Natl. Acad. Sci. U. S. A.*, 104, 16793-16797.
- Bedell, V.M., Wang, Y., Campbell, J.M., Poshusta, T.L., Starker, C.G., Krug II, R.G., Tan, W., Penheiter, S.G., Ma, A.C., Leung, A.Y.H., Fahrenkrug, S.C., Carlson, D.F., Voytas, D.F., Clark, K.J., Essner, J.J. and Ekker, S.C. (2012) *In vivo* genome editing using a high-efficiency TALEN system. *Nature*, 490, 114-118.
- Belitsky, J.M., Leslie, S.J., Arora, P.S., Beerman, T.A. and Dervan, P.B. (2002) Cellular uptake of N-methylpyrrole/N-methylimidazole polyamide-dye conjugates. *Bioorg. Med. Chem.*, 10, 3313-3318.
- Belitsky, J.M., Nguyen, D.H., Wurtz, N.R. and Dervan, P.B. (2002) Solid-phase synthesis of DNA binding polyamides on oxime resin. *Bioorg. Med. Chem.*, 10, 2767-2774.
- Besanceney-Webler, C., Jiang, H., Zheng, T., Feng, L., Soriano del Amo, D., Wang, W., Klivansky, L.M., Marlow, F.L., Liu, Y. and Wu, P. (2011) Increasing the efficacy of bioorthogonal click reactions for bioconjugation : A comparative study. *Angew. Chem. Int. Ed.*, 50, 8051-8056.
- Best, T.P., Edelson, B.S., Nickols, N.G. and Dervan, P.B. (2003) Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc. Natl. Acad. Sci. U. S. A.*, 100, 12063-12068.
- Bierne, H., Tham, T.N., Batsche, E., Dumay, A., Leguillou, M., Kernéis-Golsteyn, S., Regnault, B., Seeler, J.S., Muchardt, C., Feunteun, J. and Cossart, P. (2009) Human BAHD1 promotes heterochromatic gene silencing. *Proc. Natl. Acad. Sci. U. S. A.*, 106, 13826-13831.
- Blackledge, M.S. and Melander, C. (2013) Programmable DNA-binding Small Molecules. *Bioorg. Med. Chem.*, 21, 6101-6114.

- Blobe, G.C., Schiemann, W.P. and Lodish, H.F. (2000) Role of transforming growth factor β in human disease. *New Engl. J. Med.*, 342, 1350-1358.
- Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. and Bonas, U. (2009) Breaking the code of DNA binding specificity of TAL-Type III effectors. *Science*, 326, 1509-1512.
- Bogdanov, A.M., Kudryavtseva, E.I. and Lukyanov, K.A. (2012) Anti-fading media for live cell GFP imaging. *PLoS ONE*, 7, e53004.
- Bolotin, A., Quinquis, B., Sorokin, A. and Ehrlich, S.D. (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology*, 151, 2551-2561.
- Bolte, S. and Cordelières, F.P. (2006) A guided tour into subcellular colocalisation analysis in light microscopy. *J. Microsc.*, 224, 213-232.
- Boutorine, A. and Escudé, C. (2007), Biophysical analysis of triple-helix formation. *Curr. Prot. Nucleic Acid Chem.*, 29, 7.12.1-7.12.16.
- Boutorine, A.S., Halby, L., Simon, P., Perrouault, L., Giovannangeli, C., Gursky, G.V., Surovaya, A.N., Grokhovsky, S.L., Ryabinin, V.A. and Sinyakov, A.N. (2007) Sequence-specific recognition of double-stranded DNA by synthetic minor groove binder conjugates. Toward the construction of artificial site-specific deoxyribonucleases. *Nucleos. Nucleot. Nucl. Acids*, 26, 1559-1563.
- Boutorine, A.S., Novopashina, D.S., Krasheninina, O.A., Nozeret, K. and Venyaminova, A.G. (2013) Fluorescent probes for nucleic acid visualization in fixed and live cells. *Molecules*, 18, 15357-15397.
- Bouzinba-Segard, H., Guais, A. and Francastel, C. (2006) Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. *Proc. Natl. Acad. Sci. U. S. A.*, 103, 8709-8714.
- Brown, K.E., Guest, S.S., Smale, S.T., Hahm, K., Merkenschlager, M. and Fisher, A.G. (1997) Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell*, 91, 845-854.
- Campos, E.I. and Reinberg, D. (2009) Histones : annotating chromatin. *Annu. Rev. Genet.*, 43, 559-599.
- Casey, B.P. and Glazer, P.M. (2001) Gene targeting via triple-helix formation. *Prog. Nucleic Acid Res. Mol. Biol.*, 67, 163-192.
- Chang, A.Y. and Dervan, P.B. (2000) Strand selective cleavage of DNA by diastereomers of hairpin polyamide-seco-CBI conjugates. *J. Amer. Chem. Soc.*, 122, 4856-4864.
- Chavda, S., Liu, Y., Babu, B., Davis, R., Sielaff, A., Ruprich, J., Westrate, L., Tronrud, C., Ferguson, A., Franks, A., Tzou, S., Adkins, C., Rice, T., Mackay, H., Kluza, J., Tahir, S.A., Lin, S., Kiakos, K., Bruce, C.D., Wilson, W.D., Hartley, J.A. and Lee, M. (2011) Hx, a novel fluorescent, minor groove and sequence specific recognition element : design, synthesis, and DNA binding properties of p-anisylbenzimidazole-imidazole/pyrrole-containing polyamides. *Biochemistry*, 50, 3127-3136.
- Cheeseman, I.M. and Desai, A. (2008) Molecular architecture of the kinetochore microtubule interface. *Nat. Rev. Mol. Cell. Biol.*, 9, 33-46.

- Chen, B., Gilbert, L.A., Cimini, B.A., Schnitzbauer, J., Zhang, W., Li, G., Park, J., Blackburn, E.H., Weissman, J.S., Qi, L.S. and Huang, B. (2013) Dynamic imaging of genomic loci in living human cells by an optimized CRISPR / Cas system. *Cell*, 155, 1479-1491.
- Chenoweth, D.M. and Dervan, P.B. (2009) Allosteric modulation of DNA by small molecules. *Proc. Natl. Acad. Sci. U. S. A.*, 106, 13175-13179.
- Chenoweth, D.M., Harki, D.A. and Dervan, P.B. (2009) Solution-phase synthesis of pyrroleimidazole polyamides. J. Amer. Chem. Soc., 131, 7175-7181.
- Chenoweth, D.M., Harki, D.A., Phillips, J.W., Dose, C. and Dervan, P.B. (2009) Cyclic pyrrole-imidazole polyamides targeted to the androgen response element. *J. Amer. Chem. Soc.*, 131, 7175-7181.
- Chenoweth, D.M., Meier, J.L. and Dervan, P.B. (2013) Pyrrole-imidazole polyamides distinguish between double-helical DNA and RNA. *Angew. Chem. Int. Ed.*, 52, 415-418.
- Chenoweth, D.M., Viger, A. and Dervan, P.B. (2007) Fluorescent sequence-specific dsDNA binding oligomers. J. Amer. Chem. Soc., 129, 2216-2217.
- Chou, C.J., O'Hare, T., Lefebvre, S., Alvarez, D., Tyner, J.W., Eide, C.A., Druker, B.J. and Gottesfeld, J.M. (2008) Growth arrest of BCR-ABL positive cells with a sequence-specific polyamide-chlorambucil conjugate. *PLoS ONE*, 3, e3593.
- Chou, J.C., Farkas, M.E., Tsai, S.M., Alvarez, D., Dervan, P.B. and Gottesfeld, J.M. (2008) Small molecules targeting Histone H4 as potential therapeutics for chronic myelogenous leukemia. *Mol. Cancer Ther.*, 7, 769-778.
- Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J. and Voytas, D.F. (2010) Targeting DNA double-strand breaks with TAL Effector Nucleases. *Genetics*, 186, 757-761.
- Corneo, G., Genelli, E. and Polli, E. (1967) A satellite DNA isolated from human tissues. J. *Mol. Biol.*, 23, 619-622.
- Coull, J., He, G., Melander, C., Rucker, V.C., Dervan, P.B. and Margolis D.M. (2002) Targeted derepression of the human immunodeficiency virus type 1 long terminal repeat by pyrrole-imidazole polyamides. *J. Virol.*, 76, 12349-12354.
- Crowley, K.S., Phillion, D.P., Woodard, S.S., Schweitzer, B.A., Singh, M., Shabany, H., Burnette, B., Hippenmeyer, P., Heitmeier, M. and Bashkin, J.K. (2003) Controlling the intracellular localization of fluorescent polyamide analogues in cultured cells. *Bioorg. Med. Chem. Lett.*, 13, 1565-1570.
- Daele, I.V., Bomholt, N., Filichev, V.V., Calenbergh, S.V. and Pedersen, E.B. (2008) Triplex formation by pyrene-labelled probes for nucleic acid detection in fluorescence assays. *ChemBioChem*, 9, 791-801.
- Dervan, P.B. (1986) Design of sequence-specific DNA-binding molecules. *Science*, 232, 464-471.
- Dervan, P.B. (2001) Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.*, 9, 2215-2235.
- Dervan, P.B. and Bürli, R.W. (1999) Sequence-specific DNA recognition by polyamides. *Curr. Opin. Chem. Biol.*, 3, 688-693.

- Dervan, P.B. and Edelson, B.S. (2003) Recognition of the DNA minor groove by pyrroleimidazole polyamides. *Curr. Opin. Struct. Biol.*, 13, 284-299.
- Dervan, P.B., Doss, R.M. and Marques, M.A. (2005) Programmable DNA binding oligomers for control of transcription. *Curr. Med. Chem. Anti-Cancer Agents*, 5, 373-387.
- Dervan, P.B., Poulin-Kerstein, A.T., Fechter, E.J. and Edelson, B.S. (2005) Regulation of gene expression by synthetic ligands. *DNA Binders Rel. Subj.*, 253, 2-28.
- Dickerson, R.E. and Kopka, M.L. (1985) Nuclear Overhauser data and stereochemical considerations suggest that netropsin binds symmetrically within the minor groove of poly(dA)·poly(dT), forming hydrogen bonds with both strands of the double helix. *J. Biomol. Struct. Dynam.*, 3, 423-431.
- Dickinson, L.A., Burnett, R., Melander, C., Edelson, B.S., Arora, P.S., Dervan, P.B. and Gottesfeld, J.M. (2004) Arresting cancer proliferation by small-molecule gene regulation. *Chem. Biol.*, 11, 1583-1594.
- Dickinson, L.A., Gulizia, R.J., Trauger, J.W., Baird, E.E., Mosier, D.E., Gottesfeld, J.M. and Dervan, P.B. (1998) Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. *Proc. Natl. Acad. Sci. U. S. A.*, 95, 12890-12895.
- Dickinson, L.A., Trauger, J.W., Baird, E.E., Dervan, P.B., Graves, B.J. and Gottesfeld, J.M. (1999) Inhibition of Ets-1 DNA binding and ternary complex formation between Ets-1, NF-κB, and DNA by a designed DNA-binding ligand. *J. Biol. Chem.*, 274, 12765-12773.
- Dickinson, L.A., Trauger, J.W., Baird, E.E., Ghazal, P., Dervan, P.B. and Gottesfeld, J.M. (1999) Anti-repression of RNA polymerase II transcription by pyrrole-imidazole polyamides. *Biochemistry*, 38, 10801-10807.
- Donohue, J. (1956) Hydrogen-bonded helical configurations of polynucleotides. *Proc. Natl. Acad. Sci. U. S. A.*, 42, 60-65.
- Dose, C., Farkas, M.E., Chenoweth, D.M. and Dervan, P.B. (2008) Next generation hairpin polyamides with (R)-3,4-diaminobutyric acid turn unit. *J. Amer. Chem. Soc.*, 130, 6859-6866.
- Duca, M., Vekhoff, P., Oussedik, K., Halby, L. and Arimondo, P.B. (2008) The triple helix : 50 years later, the outcome. *Nucl. Acids Res.*, 36, 5123-5138.
- Dudouet, B., Burnet, R., Dickinson, L.A., Wood, M.R., Melander, C., Belitsky, J.M., Edelson, B.S., Wurtz, N., Briehn, C., Dervan, P.B. and Gottesfeld, J.M. (2003) Accessibility of nuclear chromatin by DNA binding polyamides. *Chem. Biol.*, 10, 859-867.
- Edelson, B.S., Best, T.P., Olenyuk, B., Nickols, N.G., Doss, R.M., Foister, S., Heckel, A. and Dervan, P.B. (2004) Influence of structural variation on nuclear localization of DNAbinding polyamide-fluorophore conjugates. *Nucl. Acids Res.*, 32, 2802-2818.
- Edwards, T.G., Koeller, K.J., Slomczynska, U., Fok, K., Helmus, M., Bashkin, J.K. and Fisher, C. (2011) HPV episome levels are potently decreased by pyrrole-imidazole polyamides. *Antiviral Res.*, 91, 177-186.
- Edwards, T.G., Vidmar, T.J., Koeller, K., Bashkin, J.K. and Fisher, C. (2013) DNA damage repair genes controlling human papillomavirus (HPV) episome levels under conditions of stability and extreme instability. *PLoS ONE*, 8, e75406.

- Ehley, J.A., Melander, C., Herman, D., Baird, E.E., Ferguson, H.A., Goodrich, J.A., Dervan, P.B. and Gottesfeld, J.M. (2002) Promoter scanning for transcription inhibition with DNA-binding polyamides. *Mol. Cell. Biol.*, 22, 1723-1733.
- El-Sagheer, A.H. and Brown, T. (2012) Click nucleic acid ligation : applications in biology and nanotechnology. *Acc. Chem. Res.*, 45, 1258-1267.
- Eymery, A., Callanan, M. and Vourc'H, C. (2009) The secret message of heterochromatin : new insights into the mechanisms and function of centromeric and pericentric repeat sequence transcription. *Int. J. Dev. Biol.*, 53, 259-268.
- Fallows, A.J., Singh, I., Dondi, R., Cullis, P.M. and Burley, G.A. (2014) Highly efficient synthesis of DNA-binding polyamides using a convergent fragment-based approach. *Org. Lett.*, 16, 4654-4657.
- Faria, M. and Giovannangeli, C. (2001) Triplex-forming molecules : from concepts to applications. J. Gene Med., 3, 299-310.
- Fechter, E.J., Olenyuk, B. and Dervan, P.B. (2005) Sequence-specific fluorescence detection of DNA by polyamide-thiazole orange conjugates. *J. Amer. Chem. Soc.*, 127, 16685-16691.
- Felsenfeld, G. (1957) Formation of a three stranded polynucleotide molecule. J. Amer. Chem. Soc., 79, 2023-2024.
- Felsenfeld, G. and Rich, A. (1957) Studies on the formation of two- and three-stranded polyribonucleotides. *Biochim. Biophys. Acta*, 26, 457-468.
- Flemming, W. (1965) Contributions to the knowledge of the cell and its vital processes. J. *Cell Biol.*, 25, 3-69.
- Floreancig, P.E., Swalley, S.E., Trauger, J.W. and Dervan, P.B. (2000) Recognition of the minor groove of DNA by hairpin polyamides containing α-substituted-β-amino acids. J. Amer. Chem. Soc., 122, 6342-6350.
- Foister, S. (2003) Shape selective Recognition of the DNA Minor Groove by Hairpin Polyamides. Ph.D. thesis, Chapter IV, 109-127, California Institute of Technology, Pasadena, California.
- Foister, S., Marques, M.A., Doss, R.M. and Dervan, P.B. (2003) Shape selective recognition of T·A base pairs by hairpin polyamides containing N-terminal 3-methoxy (and 3-chloro) thiophene residues. *Bioorg. Med. Chem.*, 11, 4333-4340.
- Fox, K.R. (2000) Targeting DNA with Triplexes. Curr. Med. Chem., 7, 17-37.
- Franklin, S.E. and Gosling, R.G. (1953) Molecular configuration in sodium thymonucleate. *Nature*, 171, 740-741.
- Fujimoto, J., Bando, T., Minoshima, M., Uchida, S., Iwasaki, M., Shinohara, K.I. and Sugiyama, H. (2008) Detection of triplet repeat sequences in the double-stranded DNA using pyrene-functionalized pyrrole-imidazole polyamides with rigid linkers. *Bioorg. Med. Chem.*, 16, 5899-5907.
- Fussner, E., Strauss, M., Djuric, U., Li, R., Ahmed, K., Hart, M., Ellis, J. and Bazett-Jones, D.P. (2012) Open and closed domains in the mouse genome are configured as 10-nm chromatin fibres. *EMBO Rep.*, 13, 992-996.
- Garvie, C.W. and Wolberger, C. (2001) Recognition of specific DNA sequences. *Mol. Cell*, 8, 937-946.

- Ghosh, I., Stains, C.I., Ooi, A.T. and Segal, D.J. (2006) Direct detection of double-stranded DNA : molecular methods and applications for DNA diagnostics. *Mol. Biosyst.*, 5, 551-560.
- Gonçalves dos Santos Silva, A., Sarkar, R., Harizanova, J., Guffei, A., Mowat, M., Garini, Y. and Mai, S. (2008) Centromeres in cell division, evolution, nuclear organization and disease. J. Cell. Biochem., 104, 2040-2058.
- Gottesfeld, J.M., Melander, C., Suto, R.K., Raviol, H., Luger, K. and Dervan, P.B. (2001) Sequence-specific recognition of DNA in the nucleosome by pyrrole-imidazole polyamides. *J. Mol. Biol.*, 309, 615-629.
- Gottesfeld, J.M., Neely, L., Trauger, J.W., Baird, E.E. and Dervan, P.B. (1997) Regulation of gene expression by small molecules. *Nature*, 387, 202-205.
- Gould, S. and Scott, R.C. (2005) 2-Hydroxypropyl-β-cyclodextrin (HP-β-CD) : a toxicology review. *Food Chem. Toxicol.*, 43, 1451-1459.
- Grimm, G.N., Boutorine, A.S., Lincoln, P., Nordén, B. and Hélène, C. (2002) Formation of DNA triple helices by an oligonucleotide conjugated to a fluorescent ruthenium complex. *ChemBioChem*, 3, 324-331.
- Guenatri, M., Bailly, D., Maison, C. and Almouzni, G. (2004) Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. *J Cell Biol.*, 166, 493-505.
- Guo, C., Kawamoto, Y., Asamitsu, S., Sawatani, Y., Hashiya, K., Bando, T. and Sugiyama, H. (2015) Rational design of specific binding hairpin Py–Im polyamides targeting human telomere sequences. *Bioorg. Med. Chem.*, 23, 855-860.
- Gygi, M.P., Ferguson, M.D., Mefford, H.C., Lund, K.P., O'Day, C., Zhou, P., Friedman, C., van den Engh, G., Stolowitz, M.L. and Trask, B.J. (2002) Use of fluorescent sequence-specific polyamides to discriminate human chromosomes by microscopy and flow cytometry. *Nucl. Acids Res.*, 30, 2790-2799.
- Haaf, T. and Michael, A. (1991) Chromosome topology in mammalian interphase nuclei. *Exp. Cell Res.*, 192, 325-332.
- Halby, L., Ryabinin, V.A., Sinyakov, A.N. and Boutorine, A.S. (2005) Functionalized headto-head hairpin polyamides : synthesis, double-stranded DNA-binding activity and affinity. *Bioorg. Med. Chem. Lett.*, 15, 3720-3724.
- Halby, L., Ryabinin, V.A., Sinyakov, A.N., Novopashina, D.S., Venyaminova, A.G., Grokhovsky, S.L., Surovaya, A.N., Gursky, G.V. and Boutorine, A.S. (2007) Head-tohead bis-hairpin polyamide minor groove binders and their conjugates with triplexforming oligonucleotides : studies of interaction with target double-stranded DNA. J. Biomol. Struct. Dynam., 25, 61-76.
- Han, L., Pandian, G.N., Junetha, S., Sato, S., Anandhakumar, C., Taniguchi, J., Saha, A., Bando, T., Nagase, H. and Sugiyama, H. (2013) A synthetic small molecule for targeted transcriptional activation of germ cell genes in a human somatic cell. *Angew. Chem. Int. Ed.*, 52, 13410-13413.
- Han, Y.W., Kashiwazaki, G., Morinaga, H., Matsumoto, T., Hashiya, K., Bando, T., Harada, Y. and Sugiyama, H. (2013) Effect of single pyrrole replacement with β-alanine on DNA binding affinity and sequence specificity of hairpin pyrrole/imidazole polyamides targeting 5'-GCGC-3'. *Bioorg. Med. Chem.*, 21, 5436-5441.

- Han, Y.-W., Matsumoto, T., Yokota, H., Kashiwazaki, G., Morinaga, H., Hashiya, K., Bando, T., Harada, Y. and Sugiyama, H. (2012) Binding of hairpin pyrrole and imidazole polyamides to DNA : relationship between torsion angle and association rate constants. *Nucl. Acids Res.*, 40, 11510-11517.
- Han, Y.-W., Tsunaka, Y., Yokota, H., Matsumoto, T., Kashiwazaki, G., Morinaga, H., Hashiya, K., Bando, T., Sugiyama, H. and Harada, Y. (2014) Construction and characterization of Cy3- or Cy5-conjugated hairpin pyrrole–imidazole polyamides binding to DNA in the nucleosome. *Biomater.Sci.*, 2, 297-307.
- Hargrove, A.E., Raskatov, J.A., Meier, J.L., Montgomery, D.C. and Dervan, P.B. (2012) Characterization and solubilization of pyrrole-imidazole polyamide aggregates. *J. Med. Chem.*, 55, 5425-5432.
- Hayden, K.E. (2012) Human centromere genomics : now it's personal. *Chromosome Res.*, 20, 621-633.
- He, G., Vasilieva, E., Harris, G.D., Koeller, K.J., Bashkin, J.K. and Dupureur, C.M. (2014) Binding studies of a large antiviral polyamide to a natural HPV sequence. *Biochimie*, 102, 83-91.
- Herman, D.M., Baird, E.E. and Dervan, P.B. (1998) Stereochemical control of the DNA binding affinity, sequence specificity, and orientation preference of chiral hairpin polyamides in the minor groove. *J. Amer. Chem. Soc.*, 120, 1382-1391.
- Hilal, H. and Taylor, J.A. (2008) Cyanine dyes for the detection of double stranded DNA. *J. Biochem. Biophys. Meth.*, 70, 1104-1108.
- Hirata, A., Nokihara, K., Kawamoto, Y., Bando, T., Sasaki, A. and Ide, S. (2014) Structural evaluation of tandem hairpin pyrrole–imidazole. *J. Amer. Chem. Soc.*, 136, 11546-11554.
- Hoogsteen, K. (1959) The structure of crystals containing a hydrogen-bonded complex of 1methylthymine and 9-methyladenine. *Acta Cryst.*, 12, 822-823.
- Hoogsteen, K. (1963) The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Cryst.*, 16, 907-916.
- Howman, E.V., Fowler, K.J., Newson, A.J., Redward, S., Macdonald, A.C., Kalitsis, P. and Choo, K.H.A. (2000) Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. *Proc. Natl. Acad. Sci. U. S. A.*, 97, 1148-1153.
- Hsu, C.F. and Dervan, P.B. (2008) Quantitating the concentration of Py-Im polyamide-fluorescein conjugates in live cells. *Bioorg. Med. Chem. Lett.*, 18, 5851-5855.
- Hsu, C.F., Phillips, J.W., Trauger, J.W., Farkas, M.E., Belitsky, J.M., Heckel, A., Olenyuk, B.Z., Puckett, J.W., Wang, C.C.C. and Dervan, P.B. (2007) Completion of a programmable DNA-binding small molecule library. *Tetrahedron*, 63, 6146-6151.
- Iguchi, A., Fukuda, N., Takahashi, T., Watanabe, T., Matsuda, H., Nagase, H., Bando, T., Sugiyama, H. and Shimizu, K. (2013) RNA binding properties of novel gene silencing pyrrole-imidazole polyamides. *Biol. Pharm. Bull.*, 36, 1152-1158.
- Jacobs, C.S. and Dervan, P.B. (2009) Modifications at the C-terminus to improve pyrroleimidazole polyamide activity in cell culture. *J. Med. Chem.*, 52, 7380-7388.
- Janssen, B.M.G., van Ommeren, S.P.F.I. and Merkx, M. (2015) Efficient synthesis of peptide and protein functionalized pyrrole-imidazole polyamides using native chemical ligation. *Int. J. Mol. Sci.*, 16, 12631-12647.

- Janssen, S., Cuvier, O., Müller, M. and Laemmli, U.K. (2000) Specific gain- and loss-offunction phenotypes induced by satellite-specific DNA-binding drugs fed to Drosophila melanogaster. *Mol. Cell*, 6, 1013-1024.
- Janssen, S., Durussel, T. and Laemmli, U.K. (2000) Chromatin opening of DNA satellites by targeted sequence-specific drugs. *Mol. Cell*, 6, 999-1011.
- Johnson III, M.D. and Fresco, J.R. (1999) Third-strand in situ hybridization (TISH) to nondenatured metaphase spreads and interphase nuclei. *Chromosoma*, 108, 181-189.
- Jost, K.L., Bertulat, B. and Cardoso, M.C. (2012) Heterochromatin and gene positioning : inside, outside, any side ? *Chromosoma*, 121, 555-563.
- Kang, J.S., Meier, J.L. and Dervan, P.B. (2014) Design of sequence-specific DNA binding molecules for DNA methyltransferase inhibition. *J. Amer. Chem. Soc.*, 136, 3687-3694.
- Kashiwazaki, G., Bando, T., Shinohara, K.I., Minoshima, M., Kumamoto, H., Nishijima, S. and Sugiyama, H. (2010) Alkylation of a human telomere sequence by heterotrimeric chlorambucil PI polyamide conjugates. *Bioorg. Med. Chem.*, 18, 2887-2893.
- Kashiwazaki, G., Bando, T., Shinohara, K.I., Minoshima, M., Nishijima, S. and Sugiyama, H. (2009) Cooperative alkylation of double-strand human telomere repeat sequences by PI polyamides with 11-base-pair recognition based on a heterotrimeric design. *Bioorg. Med. Chem.*, 17, 1393-1397.
- Kashiwazaki, G., Bando, T., Yoshidome, T., Masui, S., Takagaki, T., Hashiya, K., Pandian, G.N., Yasuoka, J., Akiyoshi, K. and Sugiyama, H. (2012) Synthesis and biological properties of highly sequence-specific-alkylating N-methylpyrrole-N-methylimidazole polyamide conjugates. J. Med. Chem., 55, 2057-2066.
- Kawamoto, Y., Bando, T., Kamada, F., Li, Y., Hashiya, K., Maeshima, K. and Sugiyama, H. (2013) Development of a new method for synthesis of tandem hairpin pyrrole-imidazole polyamide probes targeting human telomeres. J. Amer. Chem. Soc., 135, 16468-16477.
- Kawamoto, Y., Sasaki, A., Hashiya, K., Ide, S., Bando, T., Maeshima, K. and Sugiyama, H. (2015) Tandem trimer pyrrole–imidazole polyamide probes targeting 18 base pairs in human telomere sequences. *Chem. Sci.*, 6, 2307-2312.
- Kielkopf, C.L., Baird, E.E., Dervan, P.B. and Rees, D.C. (1998) Structural basis for G•C recognition in the DNA minor groove. *Nat. Struct. Biol.*, 5, 104-109.
- Kielkopf, C.L., Bremer, R.E., White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E., Dervan, P.B. and Rees, D.C. (2000) Structural effects of DNA sequence on T.A recognition by hydroxypyrrole/pyrrole pairs in the minor groove. J. Mol. Biol., 295, 557-567.
- Kielkopf, C.L., White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E., Dervan, P.B. and Rees, D.C. (1998) A structural basis for recognition of A·T and T·A base pairs in the minor groove of B-DNA. *Science*, 282, 111-115.
- Klevit, R.E., Wemmer, D.E. and Reid, B.R. (1986) 1H NMR studies on the interaction between distamycin A and a symmetrical DNA dodecamer. *Biochemistry*, 25, 3296-3303.
- Klug, A. (2010) The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annu. Rev. Biochem.*, 79, 213-231.
- Knauert, M.P. and Glazer, P.M. (2001) Triplex forming oligonucleotides : sequence-specific tools for gene targeting. *Hum. Mol. Genet.*, 10, 2243-2252.

- Kopka, M.L., Yoon, C., Goodsell, D., Pjura, P. and Dickerson, R.E. (1985a) Binding of an antitumor drug to DNA netropsin and C-G-C-G-A-A-T-T-BrC-G-C-G. J. Mol. Biol., 183, 553-563.
- Kopka, M.L., Yoon, C., Goodsell, D., Pjura, P. and Dickerson, R.E. (1985b) The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proc. Natl. Acad. Sci. U. S.* A., 82, 1376-1380.
- Krutzik, P.O. and Chamberlin, A.R. (2002a) Rapid solid-phase synthesis of DNA-binding pyrrole-imidazole polyamides. *Bioorg. Med. Chem. Lett.*, 12, 2129-2132.
- Krutzik, P.O. and Chamberlin, A.R. (2002b) Synthesis of DNA-Binding Polyamides. Robust solid-phase methods for coupling heterocyclic aromatic amino acids. *Methods Mol. Biol.*, 201, 77-92.
- Lee, C., Wevrick, R., Fisher, R.B., Ferguson-Smith, M.A. and Lin, C.C. (1997) Human centromeric DNAs. *Hum. Genet.*, 100, 291-304.
- Lenzmeier, B.A., Baird, E.E., Dervan, P.B. and Nyborg, J.K. (1999) The tax protein-DNA interaction is essential for HTLV-I transactivation in vitro. *J. Mol. Biol.*, 291, 731-744.
- Li, B.C., Montgomery, D.C., Puckett, J.W. and Dervan, P.B. (2013) Synthesis of cyclic Py-Im polyamide libraries. *J. Org. Chem.*, 78, 124-133.
- Lindhout, B.I., Fransz, P., Tessadori, F., Meckel, T., Hooykaas, P.J.J. and Zaal, B.J.V.D. (2007) Live cell imaging of repetitive DNA sequences via GFP-tagged polydactyl zinc finger proteins. *Nucl. Acids Res.*, 35, e107.
- Liu, B. and Kodadek, T. (2009) Investigation of the relative cellular permeability of DNAbinding pyrrole-imidazole polyamides. *J. Med. Chem.*, 52, 4604-4612.
- Lown, J.W., Krowicki, K., Bhat, U.G., Skorobogaty, A., Ward, B. and Dabrowiak, J.C. (1986) Molecular recognition between oligopeptides and nucleic acids : novel imidazole-containing oligopeptides related to netropsin that exhibit altered DNA sequence specificity. *Biochemistry*, 25, 7408-7416.
- Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, 389, 251-260.
- Ma, H., Reyes-Gutierrez, P. and Pederson, T. (2013) Visualization of repetitive DNA sequences in human chromosomes with transcription activator-like effectors. *Proc. Natl. Acad. Sci. U. S. A.*, 110, 21048-21053.
- Maeshima, K., Janssen, S. and Laemmli, U.K. (2001) Specific targeting of insect and vertebrate telomeres with pyrrole and imidazole polyamides. *EMBO J.*, 20, 3218-3228.
- Maison, C., Bailly, D., Peters, A.H.F.M., Quivy, J.-P., Roche, D., Lachner, M., Jenuwein, T. and Almouzni, G. (2002) Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat. Genet.*, 30, 329-334.
- Mali, P., Esvelt, K.M. and Church, G.M. (2013) Cas9 as a versatile tool for engineering biology. *Nat. Methods*, 10, 957-963.
- Manuelidis, L. (1978) Chromosomal localization of complex and simple repeated human DNAs. *Chromosoma*, 66, 23-32.
- Manuelidis, L. (1981) Consensus sequence of mouse satellite DNA indicates it is derived from tandem 116 basepair repeats. *FEBS Lett.*, 129, 25-28.

- Mapp, A.K., Ansari, A.Z., Ptashne, M. and Dervan, P.B. (2000) Activation of gene expression by small molecule transcription factors. *Proc. Natl. Acad. Sci. U. S. A.*, 97, 3930-3935.
- Marques, M.A., Doss, R.M., Urbach, A.R. and Dervan, P.B. (2002) Toward an understanding of the chemical etiology for DNA minor-groove recognition by polyamides. *Helv. Chim. Acta*, 85, 4485-4517.
- Matsuda, H., Fukuda, N., Ueno, T., Katakawa, M., Wang, X., Watanabe, T., Matsui, S.-I., Aoyama, T., Saito, K., Bando, T. Matsumoto, Y., Nagase, H., Matsumoto, K. and Sugiyama, H. (2011) Transcriptional inhibition of progressive renal disease by gene silencing pyrrole-imidazole polyamide targeting of the transforming growth factor-β1 promoter. *Kidney Int.*, 79, 46-56.
- Matsuda, H., Fukuda, N., Ueno, T., Tahira, Y., Ayame, H., Zhang, W., Bando, T., Sugiyama, H., Saito, S., Matsumoto, K., Mugishima, H. and Serie, K. (2006) Development of gene silencing pyrrole-imidazole polyamide targeting the TGF-β1 promoter for treatment of progressive renal diseases. *J. Am. Soc. Nephrol.*, 17, 422-432.
- Mayer, R., Brero, A., Hase, J.V., Schroeder, T., Cremer, T. and Dietzel, S. (2005) Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. *BMC Cell Biol.*, 6 : 44 (1-22).
- Meier, J.L., Montgomery, D.C. and Dervan, P.B. (2012) Enhancing the cellular uptake of Py-Im polyamides through next-generation aryl turns. *Nucl. Acids Res.*, 40, 2345-2356.
- Melander, C., Burnett, R. and Gottesfeld, J.M. (2004) Regulation of gene expression with pyrrole–imidazole polyamides. *J. Biotechnol.*, 112, 195-220.
- Meldal, M. and Tornøe, C.W. (2008) Cu-catalyzed azide-alkyne cycloaddition. *Chem. Rev.*, 108, 2952-3015.
- Mergny, J.-L., Boutorine, A.S., Garestier, T., Belloc, F., Rougée, M., Bulychev, N.V., Koshkin, A.A., Bourson, J., Lebedev, A.V., Valeur, B., Thuong, N.T. and Hélène, C. (1994) Fluorescence energy transfer structures as a probe for nucleic acid structures and sequences. *Nucl. Acids Res.*, 22, 920-928.
- Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J., Dulay, G.P., Hua, K.L., Ankoudina, I., Cost, G.J., Urnov, F.D., Zhang, H.S., Holmes, M.C., Zhang, L., Gregory, P.D. and Rebar, E.J. (2011) A TALE nuclease architecture for efficient genome editing. *Nat. Biotechnol.*, 29, 143-150.
- Mishra, R., Watanabe, T., Kimura, M.T., Koshikawa, N., Ikeda, M., Uekusa, S., Kawashima, H., Wang, X., Igarashi, J., Choudhury, D., Grandori, C., Kemp, C.J., Ohira, M., Verma, N.K., Kobayashi, Y., Takeuchi, J., Koshinaga, T., Nemoto, N., Fukuda, N., Soma, M., Kusafuka, T., Fujiwara, K. and Nagase, H. (2015) Identification of a novel E-box binding pyrrole-imidazole polyamide inhibiting MYC-driven cell proliferation. *Cancer Sci.*, 106, 421-429.
- Misteli, T. (2010) Higher-order genome organization in human disease. *Cold Spring Harb. Perspect. Biol.*, 2, a000794.
- Mitra, S.N., Wahl, M.C. and Sundaralingam, M. (1999) Structure of the side-by-side binding of distamycin to d(GTATATAC). *Acta Cryst.*, 55, 602-609.
- Miyanari, Y. (2014) TAL effector-mediated genome visualization (TGV). *Methods*, 69, 198-204.

- Miyanari, Y., Ziegler-Birling, C. and Torres-Padilla, M.-E. (2013) Live visualization of chromatin dynamics with fluorescent TALEs. *Nat. Struct. Mol. Biol.*, 20, 1321-1325.
- Mojica, F.J.M., Diez-Villasenor, C., Garcia-Martinez, J. and Soria, E. (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.*, 60, 174-182.
- Molenaar, C., Wiesmeijer, K., Verwoerd, N.P., Khazen, S., Eils, R., Tanke, H.J. and Dirks, R.W. (2003) Visualizing telomere dynamics in living mammalian cells using PNA probes. *EMBO J.*, 22, 6631-6641.
- Montgomery, D.C. (2013) Improving the biological activity of pyrrole-imidazole polyamides. Ph.D. thesis, California Institute of Technology, Pasadena, California.
- Mrksich, M. and Dervan, P.B. (1994) Design of a covalent peptide heterodimer for sequencespecific recognition in the minor groove of double-helical DNA. J. Amer. Chem. Soc., 116, 3663-3664.
- Mrksich, M., Parks, M.E. and Dervan, P.B. (1994) Hairpin peptide motif. A new class of oligopeptides for sequence-specific recognition in the minor groove of double-helical DNA. J. Amer. Chem. Soc., 116, 7983-7988.
- Munch, H., Hansen, J.S., Pittelkow, M., Christensen, J.B. and Boas, U. (2008) A new efficient synthesis of isothiocyanates from amines using di-tert-butyl dicarbonate. *Tetrahedron Lett.*, 49, 3117-3119.
- Nga-Sze Mak, A., Bradley, P., Cernadas, R.A., Bogdanove, A.J. and Stoddard, B.L. (2012) The crystal structure of TAL Effector PthXo1 bound to its DNA target. *Science*, 335, 716-719.
- Nickols, N.G., Jacobs, C.S., Farkas, M.E. and Dervan, P.B. (2007) Improved nuclear localization of DNA-binding polyamides. *Nucl. Acids Res.*, 35, 363-370.
- Nishijima, S., Shinohara, K.-I., Bando, T., Minoshima, M., Kashiwazaki, G. and Sugiyama, H. (2010) Cell permeability of Py-Im-polyamide-fluorescein conjugates : Influence of molecular size and Py/Im content. *Bioorg. Med. Chem.*, 18, 978-983.
- Nishiwaki, E., Tanaka, S., Lee, H. and Shibuya, M. (1988) Efficient synthesis of oligo-Nmethylpyrrolecarboxamides and related compounds. *Heterocycles*, 27, 1945-1952.
- Novopashina, D.S., Sinyakov, A.N., Ryabinin, V.A., Venyaminova, A.G., Halby, L., Sun, J.-S. and Boutorine, A.S. (2005) Sequence-specific conjugates of oligo(2'-Omethylribonucleotides) and hairpin oligocarboxamide minor-groove binders: design, synthesis, and binding studies with double-stranded DNA. *Chem. Biodivers.*, 2, 936-952.
- Nozeret, K., Bonan, M., Yarmoluk, S.M., Novopashina, D.S. and Boutorine, A.S. (2015) Synthesis of mouse centromere-targeted polyamides and physico-chemical studies of their interaction with the target double-stranded DNA. *Bioorg. Med. Chem.*, 23, 5932-5945.
- Nozeret, K., Loll, F., Escudé, C. and Boutorine, A.S. (2015) Polyamide fluorescent probes for visualization of repeated DNA sequences in living cells. *ChemBioChem*, 16, 549-554.
- Ohtsuki, A., Kimura, M.T., Minoshima, M., Suzuki, T., Ikeda, M., Bando, T., Nagase, H., Shinohara, K.I. and Sugiyama, H. (2009) Synthesis and properties of PI polyamide-SAHA conjugate. *Tetrahedron Lett.*, 50, 7288-7292.

- Olenyuk, B.Z., Zhang, G.-J., Klco, J.M., Nickols, N.G., Kaelin, W.G. and Dervan, P.B. (2004) Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc. Natl. Acad. Sci. U. S. A.*, 101, 16768-16773.
- Olins, A.L. and Olins, D.E. (1974) Spheroid chromatin units (v bodies). Science, 183, 330-332.
- Olins, D.E. and Olins, A.L. (2003) Chromatin history : our view from the bridge. *Nat. Rev. Mol. Cell. Biol.*, 4, 809-814.
- Ollion, J. (2014) Dynamique de l'organisation nucléaire des séquences d'ADN répétées centromériques humaines au cours du cycle cellulaire. Thèse de doctorat, Muséum National d'Histoire Naturelle, Paris.
- Pandian, G.N., Nakano, Y., Sato, S., Morinaga, H., Bando, T., Nagase, H. and Sugiyama, H. (2012) A synthetic small molecule for rapid induction of multiple pluripotency genes in mouse embryonic fibroblasts. *Sci. Rep.*, 2 : 544 (1-8).
- Pandian, G.N., Ohtsuki, A., Bando, T., Sato, S., Hashiya, K. and Sugiyama, H. (2012) Development of programmable small DNA-binding molecules with epigenetic activity for induction of core pluripotency genes. *Bioorg. Med. Chem.*, 20, 2656-2660.
- Pardue, M.L. and Gall, J.G. (1970) Chromosomal localization of mouse satellite DNA. *Science*, 168, 1356-1358.
- Park, S., Bando, T., Shinohara, K.I., Nishijima, S. and Sugiyama, H. (2011) Photocontrollable sequence-specific DNA alkylation by a pyrrole-imidazole polyamide seco-CBI conjugate. *Bioconjug. Chem.*, 22, 120-124.
- Patel, D.J. (1982) Antibiotic-DNA interactions : Intermolecular nuclear Overhauser effects in the netropsin-d (C-G-C-G-A-A-T-T-C-G-C-G) complex in solution. *Proc. Natl. Acad. Sci. U. S. A.*, 79, 6424-6428.
- Pelton, J.G. and Wemmer, D.E. (1988) Structural modeling of the distamycin Ad(CGCGAATTCGCG)₂ complex using 2D NMR and molecular mechanics. *Biochemistry*, 27, 8088-8096.
- Pelton, J.G. and Wemmer, D.E. (1989) Structural characterization of a 2:1 distamycin A.d(CGCAAATTGGC) complex by two-dimensional NMR. *Proc. Natl. Acad. Sci. U.* S. A., 86, 5723-5727.
- Pelton, J.G. and Wemmer, D.E. (1990) Binding modes of distamycin A with d(CGCAAATTTGCG)₂ determined by two-dimensional NMR. *J. Amer. Chem. Soc.*, 112, 1393-1399.
- Pluta, A.F., Mackay, A.M., Ainsztein, A.M., Goldberg, G. and Earnshaw, W.C. (1995) The centromere : Hub of chromosomal activities. *Science*, 270, 1591-1594.
- Poulin-Kerstien, A.T. and Dervan, P.B. (2003) DNA-templated dimerization of hairpin polyamides. J. Amer. Chem. Soc., 125, 15811-15821.
- Prakash, T. (2011) An overview of sugar-modified oligonucleotides for antisense therapeutics. *Chem. Biodivers.*, 8, 1616-1641.
- Puckett, J.W., Green, J.T. and Dervan, P.B. (2012) Microwave assisted synthesis of Py-Im polyamides. *Org. Lett.*, 14, 2774-2777.

- Ramos, J.P., Babu, B., Chavda, S., Liu, Y., Plaunt, A., Ferguson, A., Savagian, M., Lee, M., Tzou, S., Lin, S., Kiakos, K., Wang, S., Lee, M., Hartley, J.A. and Wilson, W.D. (2013) Affinity and kinetic modulation of polyamide-DNA interactions by N-modification of the heterocycles. *Biopolymers*, 99, 497-507.
- Raskatov, J.A., Hargrove, A.E., So, A.Y. and Dervan, P.B. (2012) Pharmacokinetics of Py-Im polyamides depend on architecture : cyclic versus linear. *J. Amer. Chem. Soc.*, 134, 7995-7999.
- Raskatov, J.A., Meier, J.L., Puckett, J.W., Yang, F., Ramakrishnan, P. and Dervan, P.B. (2012) Modulation of NF-kB-dependent gene transcription using programmable DNA minor groove binders. *Proc. Natl. Acad. Sci. U. S. A.*, 109, 1023-1028.
- Raskatov, J.A., Szablowski, J.O. and Dervan, P.B. (2014) Tumor xenograft uptake of a pyrrole–imidazole (Py-Im) polyamide varies as a function of cell line grafted. *J. Med. Chem.*, 57, 8471-8476.
- Renard, B.-L., Lartia, R. and Asseline, U. (2008) Targeting DNA with « light-up » pyrimidine triple-helical forming oligonucleotides conjugated to stabilizing fluorophores (LU-TFOs). Org. Biomol. Chem., 6, 4413-4425.
- Renneberg, D. and Dervan, P.B. (2003) Imidazopyridine / pyrrole and hydroxybenzimidazole / pyrrole pairs for DNA minor groove recognition. *J. Amer. Chem. Soc.*, 125, 5707-5716.
- Rice, J.C., Briggs, S.D., Ueberheide, B., Barber, C.M., Shabanowitz, J., Hunt, D.F., Shinkai,Y. and Allis, C.D. (2003) Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol. Cell*, 12, 1591-1598.
- Rodriguez, R., Miller, K.M., Forment, J.V., Bradshaw, C.R., Nikan, M., Britton, S., Oelschlaegel, T., Xhemalce, B., Balasubramanian, S. and Jackson, S.P. (2012) Small molecule-induced DNA damage identifies alternative DNA structures in human genes. *Nat. Chem. Biol.*, 8, 301-310.
- Rodriguez-Marinez, J.A., Peterson-Kaufman, K.J. and Ansari, A.Z. (2014) Small-molecule regulators that mimic transcription factors. *Biochim. Biophys. Acta*, 1799, 768-774.
- Roldán, E., Fuxa, M., Chong, W., Martinez, D., Novatchkova, M., Busslinger, M. and Skok, J.A. (2005) Locus « decontraction » and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. *Nat. Immunol.*, 6, 31-41.
- Rucker, V.C., Dunn, A.R., Sharma, S., Dervan, P.B. and Gray, H.B. (2004) Mechanism of sequence-specific fluorescent detection of DNA by N-methyl-imidazole, N-methyl-pyrrole, and β -alanine linked polyamides. *J. Phys. Chem. B*, 108, 7490-7494.
- Rucker, V.C., Foister, S., Melander, C. and Dervan, P.B. (2003) Sequence specific fluorescence detection of double strand DNA. J. Amer. Chem. Soc., 125, 1195-1202.
- Rudd, K.M. (2005) Organization, evolution and function of alpha satellite DNA at human centromeres. Ph.D. thesis, Department of Genetic, Case Western Reserve University.
- Rudd, M.K. and Willard, H.F. (2004) Analysis of the centromeric regions of the human genome assembly. *Trends Genet.*, 20, 529-533.
- Rudert, F., Bronner, S., Garnier, J.M. and Dollé, P. (1995) Transcripts from opposite strands of γ satellite DNA are differentially expressed during mouse development. *Mamm. Genome*, 6, 76-83.

- Saha, A., Pandian, G.N., Sato, S., Taniguchi, J., Hashiya, K., Bando, T. and Sugiyama, H. (2013) Synthesis and biological evaluation of a targeted DNA-binding transcriptional activator with HDAC8 inhibitory activity. *Bioorg. Med. Chem.*, 21, 4201-4209.
- Saha, A., Pandian, G.N., Sato, S., Taniguchi, J., Kawamoto, Y., Hashiya, K., Bando, T. and Sugiyama, H. (2014) Chemical modification of a synthetic small molecule boosts its biological efficacy against pluripotency genes in mouse fibroblasts. *ChemMedChem*, 9, 2374-2380.
- Sasaki, S., Bando, T., Minoshima, M., Shimizu, T., Shinohara, K.I., Takaoka, T. and Sugiyama, H. (2006) Sequence-specific alkylation of double-strand human telomere repeat sequence by pyrrole-imidazole polyamides with indole linkers. *J. Amer. Chem. Soc.*, 128, 12162-12168.
- Satam, V., Babu, B., Chavda, S., Savagian, M., Sjoholm, R., Tzou, S., Ramos, J., Liu, Y., Kiakos, K., Lin, S., Wilson, D.W., Hartley, J.A. and Lee, M. (2012) Novel diamino imidazole and pyrrole-containing polyamides : Synthesis and DNA binding studies of mono- and diamino-phenyl-ImPy*Im polyamides designed to target 5'-ACGCGT-3'. *Bioorg. Med. Chem.*, 20, 693-701.
- Satam, V., Babu, B., Porte, A., Savagian, M., Lee, M., Smeltzer, T., Liu, Y., Ramos, J., Wilson, W.D., Lin, S., Kiakos, K., Hartley, J.A. and Lee, M. (2012) Synthesis and DNA binding properties of 1-(3-aminopropyl)-imidazole-containing triamide f-Im*PyIm : a novel diamino polyamide designed to target 5'-ACGCGT-3'. *Bioorg. Med. Chem. Lett.*, 22, 5898-5902.
- Schwarz-Finsterle, J., Stein, S., Grobmann, C., Schmitt, E., Luba, T., Rechavi, G., Amariglio, N., Cremer, C. and Hausmann, M. (2007) Comparison of triple helical COMBO-FISH and standard FISH by means of quantitative microscopic image analysis of abl/bcr positions in cell nuclei. J. Biochem. Biophys. Meth., 70, 397-406.
- Shinohara, K.-I., Narita, A., Oyoshi, T., Bando, T., Teraoka, H. and Sugiyama, H. (2004) Sequence-specific gene silencing in mammalian cells by alkylating pyrrole-imidazole polyamides. J. Amer. Chem. Soc., 126, 5113-5118.
- Silahtaroglu, A., Pfundheller, H., Koshkin, A., Tommerup, N. and Kauppinen, S. (2004) LNA-modified oligonucleotides are highly efficient as FISH probes. *Cytogenet*. *Genome Res.*, 107, 32-37.
- Simon, P., Cannata, F., Perrouault, L., Halby, L., Concordet, J.-P., Boutorine, A., Ryabinin, V., Sinyakov, A. and Giovannangeli, C. (2008) Sequence-specific DNA cleavage mediated by bipyridine polyamide conjugates. *Nucl. Acids Res.*, 36, 3531-3538.
- Slee, R.B., Steiner, C.M., Herbert, B.-S., Vance, G.H., Hickey, R.J., Schwarz, T., Christan, S., Radovich, M., Schneider, B.P., Schindelhauer, D. and Grimes, B.R. (2012) Cancerassociated alteration of pericentromeric heterochromatin may contribute to chromosome instability. *Oncogene*, 31, 3244-3253.
- Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol., 98, 503-517.
- Su, W., Bagshaw, C.R. and Burley, G.A. (2013) Addressable and unidirectional energy transfer along a DNA three-way junction programmed by pyrrole-imidazole polyamides. *Sci. Rep.*, 3 : 1883.

- Sullivan, B.A. and Karpen, G.H. (2004) Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat. Struct. Mol. Biol.*, 11, 1076-1083.
- Sullivan, K.F., Hechenberger, M. and Masri, K. (1994) Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. *J. Cell Biol.*, 127, 581-592.
- Sun, J.-S. and Hélène, C. (1993) Oligonucleotide-directed triple-helix formation. *Curr. Opin. Struct. Biol.*, 3, 345-356.
- Surovaya, A.N., Burckhardt, G., Grokhovsky, S.L., Birch-Hirschfeld, E., Nikitin, A.M., Fritzsche, H., Zimmer, C. and Gursky, G.V. (2001) Binding of bis-linked netropsin derivatives in the parallel-stranded hairpin form to DNA. *J. Biomol. Struct. Dynam.*, 18, 689-701.
- Suto, R.K., Edayathumangalam, R.S., White, C.L., Melander, C., Gottesfeld, J.M., Dervan, P.B. and Luger, K. (2003) Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. J. Mol. Biol., 326, 371-380.
- Swalley, S.E., Baird, E.E. and Dervan, P.B. (1999) Effects of γ -turn and β -tail amino acids on sequence-specific recognition of DNA by hairpin polyamides. *J. Amer. Chem. Soc.*, 121, 1113-1120.
- Syed, J., Pandian, G.N., Sato, S., Taniguchi, J., Chandran, A., Hashiya, K., Bando, T. and Sugiyama, H. (2014) Targeted suppression of EVI1 oncogene expression by sequence-specific pyrrole-imidazole polyamide. *Chem. Biol.*, 21, 1370-1380.
- Synold, T.W., Xi, B., Wu, J., Yen, Y., Li, B.C., Yang, F., Phillips, J.W., Nickols, N.G. and Dervan, P.B. (2012) Single-dose pharmacokinetic and toxicity analysis of pyrroleimidazole polyamides in mice. *Cancer Chemother. Pharmacol.*, 70, 617-625.
- Tachikawa, K. and Briggs, S.P. (2006) Targeting the human genome. *Curr. Opin. Biotechnol.*, 17, 659-665.
- Takagaki, T., Bando, T., Kitano, M., Hashiya, K., Kashiwazaki, G. and Sugiyama, H. (2011) Evaluation of PI polyamide conjugates with eight-base pair recognition and improvement of the aqueous solubility by PEGylation. *Bioorg. Med. Chem.*, 19, 5896-5902.
- Takahashi, R., Bando, T. and Sugiyama, H. (2003) Specific alkylation of human telomere repeats by hairpin pyrrole-imidazole polyamide. *Bioorg. Med. Chem.*, 11, 2503-2509.
- Taniguchi, M., Fujiwara, K., Nakai, Y., Ozaki, T., Koshikawa, N., Toshio, K., Kataba, M., Oguni, A., Matsuda, H., Yoshida, Y., Tokuhashi, Y., Fukuda, N., Ueno, T., Som, M. and Nagase, H. (2014) Inhibition of malignant phenotypes of human osteosarcoma cells by a gene silencer, a pyrrole-imidazole polyamide, which targets an E-box motif. *FEBS Open Bio*, 4, 328-334.
- Taylor, R.D., Asamitsu, S., Takenaka, T., Yamamoto, M., Hashiya, K., Kawamoto, Y., Bando, T., Nagase, H. and Sugiyama, H. (2014) Sequence-specific DNA alkylation targeting for Kras codon 13 mutation by pyrrole-imidazole polyamide seco-CBI conjugates. *Chem. Eur. J.*, 20, 1310-1317.
- Taylor, R.D., Chandran, A., Kashiwazaki, G., Hashiya, K., Bando, T., Nagase, H. and Sugiyama, H. (2015) Selective targeting of the KRAS codon 12 mutation sequence by pyrrole-imidazole polyamide seco -CBI conjugates. *Chem. Eur. J.*, 21, 14996-15003.

- Taylor, R.D., Kawamoto, Y., Hashiya, K., Bando, T. and Sugiyama, H. (2014) Sequencespecific DNA alkylation by tandem Py-Im polyamide conjugates. *Chem. Asian J.*, 9, 2527-2533.
- Terranova, R., Sauer, S., Merkenschlager, M. and Fisher, A.G. (2005) The reorganisation of constitutive heterochromatin in differentiating muscle requires HDAC activity. *Exp. Cell Res.*, 310, 344-356.
- Thanisch, K., Schneider, K., Morbitzer, R., Solovei, I., Lahaye, T., Bultmann, S. and Leonhardt, H. (2014) Targeting and tracing of specific DNA sequences with dTALEs in living cells. *Nucl. Acids Res.*, 42, e38.
- Trauger, J.W., Baird, E.E. and Dervan, P.B. (1996) Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature*, 382, 559-561.
- Trauger, J.W., Baird, E.E. and Dervan, P.B. (1998) Recognition of 16 base pairs in the minor groove of DNA by a pyrrole-imidazole polyamide dimer. *J. Amer. Chem. Soc.*, 120, 3534-3535.
- Turner, J.M., Swalley, S.E., Baird, E.E. and Dervan, P.B. (1998) Aliphatic/aromatic amino acid pairings for polyamide recognition in the minor groove of DNA. *J. Amer. Chem. Soc.*, 120, 6219-6226.
- Uil, T.G., Haisma, H.J. and Rots, M.G. (2003) Therapeutic modulation of endogenous gene function by agents with designed DNA-sequence specificities. *Nucl. Acids Res.*, 31, 6064-6078.
- Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S. and Gregory, P.D. (2010) Genome editing with engineered zinc finger nucleases. *Nature Rev. Genetics*, 11, 636-646.
- Vaijayanthi, T., Bando, T., Hashiya, K., Pandian, G.N. and Sugiyama, H. (2013) Design of a new fluorescent probe : pyrrole/imidazole hairpin polyamides with pyrene conjugation at their γ-turn. *Bioorg. Med. Chem.*, 21, 852-855.
- Vaijayanthi, T., Bando, T., Pandian, G.N. and Sugiyama, H. (2012) Progress and prospects of pyrrole-imidazole polyamide-fluorophore conjugates as sequence-selective DNA probes. *ChemBiochem*, 13, 2170-2185.
- Viger, A. and Dervan, P.B. (2006) Exploring the limits of benzimidazole DNA-binding oligomers for the hypoxia inducible factor (HIF) site. *Bioorg. Med. Chem.*, 14, 8539-8549.
- Vissel, B. and Choo, K.H. (1989) Mouse major (γ) satellite DNA is highly conserved and organized into extremely long tandem arrays : implications for recombination between nonhomologous chromosomes. *Genomics*, 5, 407-414.
- Wade, W.S. and Dervan, P.B. (1987) Alteration of the sequence specificity of distamycin on DNA by replacement of an N-methylpyrrolecarboxamide with pyridine-2-carboxamide. J. Amer. Chem. Soc., 109, 1574-1575.
- Wade, W.S., Mrksich, M. and Dervan, P.B. (1992) Design of peptides that bind in the minor groove of DNA at 5'-(A,T)G(A,T)C(A,T)-3' sequences by a dimeric side-by-side motif. *J. Amer. Chem. Soc.*, 114, 8783-8794.
- Wang, C.C.C., Ellervik, U. and Dervan, P.B. (2001) Expanding the recognition of the minor groove of DNA by incorporation of β -alanine in hairpin polyamides. *Bioorg. Med. Chem.*, 9, 653-657.

- Wang, S., Aston, K., Koeller, K.J., Harris, G.D., Rath, N.P., Bashkin, J.K. and Wilson, W.D. (2014) Modulation of DNA-polyamide interaction by β-alanine substitutions : a study of positional effects on binding affinity, kinetics and thermodynamics. *Org. Biomol. Chem.*, 12, 7523-7536.
- Wang, S., Nanjunda, R., Aston, K., Bashkin, J.K. and Wilson, W.D. (2012) Correlation of local effects of DNA sequence and position of β-alanine inserts with polyamide-DNA complex binding affinities and kinetics. *Biochemistry*, 51, 9796-9806.
- Wang, X., Nagase, H., Watanabe, T., Nobusue, H., Suzuki, T., Asami, Y., Shinojima, Y., Kawashima, H., Takagi, K., Mishra, R., Igarashi, J., Kimura, M., Takayama, F., Fukuda, N. and Sugiyama, H. (2010) Inhibition of MMP-9 transcription and suppression of tumor metastasis by pyrrole-imidazole polyamide. *Cancer Sci.*, 101, 759-766.
- Watson, J.D. and Crick, F.H. (1953a) Molecular structure of nucleic acids ; a structure for desoxyribose nucleic acid. *Nature*, 171, 737-738.
- Watson, J.D. and Crick, F.H. (1953b) The structure of DNA. *Cold Spring Harb. Symp. Quant. Biol.*, 18, 123-131.
- Wemmer, D.E. and Dervan, P.B. (1997) Targeting the minor groove of DNA. Curr. Opin. Struct. Biol., 7, 355-361.
- White, S., Baird, E.E. and Dervan, P.B. (1996) Effects of the A.T/T.A degeneracy of pyrrole-imidazole polyamide recognition in the minor groove of DNA. *Biochemistry*, 35, 12532-12537.
- White, S., Baird, E.E. and Dervan, P.B. (1997) Orientation preferences of pyrrole-imidazole polyamides in the minor groove of DNA. J. Amer. Chem. Soc., 119, 8756-8765.
- White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E. and Dervan, P.B. (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*, 391, 468-471.
- Wiblin, A.E., Cui, W., Clark, A.J. and Bickmore, W.A. (2005) Distinctive nuclear organisation of centromeres and regions involved in pluripotency in human embryonic stem cells. *J. Cell Sci.*, 118, 3861-3868.
- Wong, A.K.C. and Rattner, J. (1988) Sequence organization and cytological localization of the minor satellite of mouse. *Nucl. Acids Res.*, 16, 11645-11661.
- Wu, J.C. and Manuelidis, L. (1980) Sequence definition and organization of a human repeated DNA. *J. Mol. Biol.*, 142, 363-386.
- Wurtz, N.R. and Dervan, P.B. (2000) Sequence specific alkylation of DNA by hairpin pyrrole-imidazole polyamide conjugates. *Chem. Biol.*, 7, 153-161.
- Wurtz, N.R., Turner, J.M., Baird, E.E. and Dervan, P.B. (2001) Fmoc solid phase synthesis of polyamides containing pyrrole and imidazole amino acids. *Org. Lett.*, 3, 1201-1203.
- Xiao, X., Yu, P., Lim, H.-S., Sikder, D. and Kodadek, T. (2007) A cell-permeable synthetic transcription factor mimic. *Angew. Chem. Int. Ed.*, 46, 2865-2868.
- Yamamoto, M., Bando, T., Kawamoto, Y., Taylor, R.D., Hashiya, K. and Sugiyama, H. (2014) Specific alkylation of human telomere repeat sequences by a tandem-hairpin motif of pyrrole-imidazole polyamides with indole-seco-CBI. *Bioconjug. Chem.*, 25, 552-559.

- Yamamoto, M., Bando, T., Morinaga, H., Kawamoto, Y., Hashiya, K. and Sugiyama, H. (2014) Sequence-specific DNA recognition by cyclic pyrrole-imidazole cysteinederived polyamide dimers. *Chem. Eur. J.*, 20, 752-759.
- Yang, F., Nickols, N.G., Li, B.C., Szablowski, J.O., Hamilton, S.R., Meier, J.L., Wang, C.-M. and Dervan, P.B. (2013) Animal toxicity of hairpin pyrrole-imidazole polyamides varies with the turn unit. *J. Med. Chem.*, 56, 7449-7457.
- Yarmoluk, S.M., Zhyvoloup, A.N., Koval'ska, V.B., Klimenko, I.V., Kukharenko, A.P., Kovtun, Y.P. and Slominsky, Y.L. (1996) Interaction of cyanine dyes with nucleic acids. I. Studies on monomethyne cyanine dyes as possible fluorescent probes for the detection of nucleic acids. *Biopolymers and Cell*, 12, 69-74.
- Yasuda, A., Noguchi, K., Minoshima, M., Kashiwazaki, G., Kanda, T., Katayama, K., Mitsuhashi, J., Bando, T., Sugiyama, H. and Sugimoto, Y. (2011) DNA ligand designed to antagonize EBNA1 represses Epstein-Barr virus-induced immortalization. *Cancer Sci.*, 102, 2221-2230.
- Yoshidome, T., Endo, M., Kashiwazaki, G., Hidaka, K., Bando, T. and Sugiyama, H. (2012) Sequence-selective single-molecule alkylation with a pyrrole-imidazole polyamide visualized in a DNA nanoscaffold. *J. Amer. Chem. Soc.*, 134, 4654-4660.
- Zhang, Q., Cui, X., Lin, S., Zhou, J. and Yuan, G. (2012) Convenient method for the synthesis of a flexible cyclic polyamide for selective targeting of c-myb G-quadruplex DNA. *Org. Lett.*, 14, 6126-6129.
- Zhang, Y., Sicot, G., Cui, X., Vogel, M., Wuertzer, C.A., Lezon-Geyda, K., Wheeler, J., Harki, D.A., Muzikar, K.A., Stolper, D.A., Dervan, P.B. and Perkins, A.S. (2011) Targeting a DNA binding motif of the EVI1 protein by a pyrrole-imidazole polyamide. *Biochemistry*, 50, 10431-10441.
- Zimmer, C. and Wähnert, U. (1986) Non intercalating DNA binding ligands : specificity of the interaction and their use as tools in biophysical, biochemical and biological investigations of the genetic material. *Prog. Biophys. Molec. Biol*, 47, 31-112.