THÈSE

Pour obtenir le grade de Docteur

Délivré par l'**Ecole Nationale Supérieure de Chimie de Montpellier**

Préparée au sein de l'école doctorale
Sciences chimiques Balard
Et de l'unité de recherche
IBMM – Glycochimie et reconnaissance moléculaire

Spécialité : **Ingénierie Moléculaire**

Présentée par **Delphine PAOLANTONI**

AUTO-ASSEMBLAGES DE SYSTEMES DE RECONNAISSANCE DE L'ADN BASES SUR DES LIGANDS GUANIDINIUMS

Soutenue le 10 décembre 2014 devant le jury composé de

M. Rachid BENHIDA, DR CNRS, Université de Nice	Président du jury Rapporteur
M. Pascal DUMY, Professeur, ENSCM	Directeur de thèse
M. Jean-Serge REMY, DR CNRS, Université de Strasbourg	Rapporteur
M. Sébastien ULRICH, CR CNRS, ENSCM	Examinateur

A Nanou,

Parce que tu es toujours là.

A ma famille.

TABLE DES MATIERES

REMERCIEMENTS 1

ABREVIATIONS 3

CHAPITRE 1 : GENERALITES – BIBLIOGRAPHIE 5

I. L'ADN ET SA RECONNAISSANCE 7

A. STRUCTURE CHIMIQUE DE L'ADN 7

B. LES DIFFERENTS MODES DE RECONNAISSANCE DE L'ADN 18

1. Reconnaissance biologique 18
2. Reconnaissance par des ligands synthétiques 19
   - *a. Interaction covalente* 19
   - *b. Intercalation* 20
   - *c. Interaction dans le grand sillon* 22
   - *d. Interactions dans le petit sillon* 24
   - *e. Interaction avec le squelette phosphodiester* 25

C. AUTO-ASSEMBLAGES MATRICES PAR L'ADN 27

1. Auto-assemblages par reconnaissance des bases 28
2. Auto-assemblages par interaction dans les sillons 39
3. Auto-assemblages par interactions électrostatiques (squelette phosphodiester) 40

II. METHODES D'IDENTIFICATION DE LIGANDS DE L'ADN 42

A. APPROCHE DITE RATIONNELLE 43

B. APPROCHE COMBINATOIRE DYNAMIQUE 45

1. Principe de la Chimie Combinatoire Dynamique 45
2. Avantages, inconvénients et limites 46
3. Les différentes liaisons réversibles employées 48
   - *a. Critères de choix et exemples des liaisons réversibles employées dans les bibliothèques* 48
   - *b. Liaison imine* 52
   - *c. Liaison (acyl)hydrazone* 53
   - *d. Liaison disulfure* 54
   - *e. Liaison de coordination métal - ligand* 55
4. Application à la reconnaissance d'ADN 56
5. Génération de matériaux dynamiques 61

III. PROJET DE RECHERCHE 63

A. BUT ET APPLICATIONS 63

B. STRATEGIE 64

C. CONCEPTION DES COMPOSES : CHOIX DU GUANIDINIUM COMME GROUPEMENT DE RECONNAISSANCE 64

RÉFÉRENCES BIBLIOGRAPHIQUES 66

CHAPITRE 2 : ETUDES DES INTERACTIONS PETITES MOLECULES / ADN PAR L'APPROCHE RATIONNELLE 73

INTRODUCTION – RAPPELS BIBLIOGRAPHIQUES 75

I. SYNTHESE DES DIFFERENTS COMPOSES 77

II. ANALYSES SPECTROSCOPIQUES 80

A. GENERALITES ET INTERETS DES TECHNIQUES EMPLOYEES 80

1. Spectroscopies UV-Visible et de fluorescence 80
2. Dichroïsme circulaire 83

*B. AUTO-ASSEMBLAGE DU COMPOSE **GUABIPY** AVEC UN OLIGOTHYMINE dT_N EN SOLUTION.* 85

C. EFFET DE LA TEMPERATURE SUR LES AUTO-ASSEMBLAGES *88*
D. EFFET DE LA LONGUEUR DE LA MATRICE SIMPLE BRIN D'ADN *89*
*E. AUTO-ASSEMBLAGE DU COMPOSE **GUABINAPH** AVEC UN OLIGOTHYMINE dT_N EN SOLUTION* *94*
III. ANALYSES EN SPECTROMETRIE DE MASSE MALDI-TOF 96
CONCLUSIONS 99
REFERENCES BIBLIOGRAPHIQUES 100

CHAPITRE 3 : FORMATION DE POLYMERES DYNAMIQUES COVALENTS POUR LA TRANSFECTION **105**

INTRODUCTION – RAPPELS BIBLIOGRAPHIQUES 107
I. SYNTHESE DES MONOMERES 110
A. SYNTHESE DES MONOMERES BISHYDRAZIDES *110*
1. Synthèse du monomère éthylèneimine bishydrazide **EI-Hyd** 110
2. Synthèse du monomère éthylèneglycol bishydrazide **EG-Hyd** 113
B. SYNTHESE DES MONOMERES BISALDEHYDES *114*
1. Synthèse du monomère éthylène glycol bisaldéhyde **EG-Ald** 114
2. Synthèse du monomère guanidinium bisaldéhyde **Gua-Ald** 114
II. REACTIONS DE POLYMERISATION ET CARACTERISATION 118
A. REACTION MODELE *118*
B. REACTION DE POLYCONDENSATION *120*
1. Suivi par RMN ^{1}H 120
2. Caractérisation par DOSY 122
3. Caractérisation par MALDI-ToF et LC-MS 124
III. DEGRADATION DES POLYMERES EN MILIEU ACIDE 127
A. DEGRADATION PAR REACTION D'ECHANGE *127*
1. Réaction modèle 127
2. Dégradation du polymère **Poly(Gua-EG)** 129
B. DEGRADATION PAR HYDROLYSE *131*
IV. ETUDES DE FLUORESCENCE (ESSAIS DE DEPLACEMENT DU BROMURE D'ETHIDIUM) 133
V. RETARD SUR GEL EN PRESENCE D'ADN 136
CONCLUSIONS 137
REFERENCES BIBLIOGRAPHIQUES 138

CHAPITRE 4 : ETUDES DES INTERACTIONS GUANIDINIUMS / ADN PAR CHIMIE COMBINATOIRE DYNAMIQUE **145**

INTRODUCTION – RAPPELS BIBLIOGRAPHIQUES 147
I. MISE AU POINT DE L'AMINATION REDUCTRICE DANS L'EAU 150
A. CONDITIONS DE REFERENCE ET CHOIX DE L'AMINE POUR LA MISE AU POINT DES CONDITIONS D'AMINATION REDUCTRICE. *151*
B. INFLUENCE DU pH ET DE LA NATURE DU TAMPON *153*
C. INFLUENCE DU NOMBRE D'EQUIVALENTS D'AMINE *157*
*D. INFLUENCE DE LA CONCENTRATION EN GUANIDINIUM **GUA-ALD*** *158*
E. INFLUENCE DE LA TEMPERATURE *160*
F. INFLUENCE D'UN PALIER D'EQUILIBRATION AVANT REDUCTION *161*
G. ETUDE DE LA REACTIVITE DES MONOAMINES *163*
H. CONDITIONS CHOISIES POUR LES PREMIERES BIBLIOTHEQUES ET MISE AU POINT DES METHODES HPLC / LC-MS *164*
II. ESSAIS AVEC dT_{10} ET dT_{40} EN SOLUTION 166
*A. BIBLIOTHEQUES AVEC UNE CONCENTRATION EN COMPOSE **GUA-ALD** DE 1 mM* *166*

1. Difficultés de quantification des constituants des bibliothèques en présence d'ADN 166
2. Influence du solvant 168
3. Influence de la longueur de l'oligonucléotide et de sa stœchiométrie 169
4. Influence de la température 170
*B. BIBLIOTHEQUES AVEC UNE CONCENTRATION EN COMPOSE **GUA-ALD** DE 10 MM 172*
III. ESSAIS AVEC ADN/DT SUR CELLULOSE 176
A. CONDITIONS ET METHODE DE GENERATION DES BIBLIOTHEQUES 176
B. BIBLIOTHEQUES AVEC DES MONOAMINES 180
1. Bibliothèques avec une seule monoamine 180
a. Premières bibliothèques avec la benzylamine dans l'eau 180
b. Rôle compétiteur du tampon phosphate 183
c. Différences de rétention en fonction de la nature des monoamines présentes 186
2. Compétition avec plusieurs monoamines 191
C. BIBLIOTHEQUES AVEC DES DIAMINES 196
D. BIBLIOTHEQUES EN PRESENCE DE TRIAZINES 202
1. Bibliothèques avec les triazines diamines **TN^n** 202
2. Bibliothèques avec la triazine bisaldéhyde **TO** 205
IV. CONFIRMATION DE LA METHODE AVEC LA SPECTROMETRIE DE MASSE (MALDI-TOF) 208
CONCLUSIONS 211
RÉFÉRENCES BIBLIOGRAPHIQUES 212

CONCLUSION GENERALE 215

EXPERIMENTAL PART 221
I. GENERAL METHODS AND MATERIAL 223
II. SYNTHESES AND CHARACTERIZATION 226
A. GUANIDINIUM COMPOUNDS 226
1. Guanidiniums for the rational approach 226
2. Guanidinium for polymers and dynamic combinatorial libraries 228
B. SYNTHESIS OF HYDRAZIDE AND SECOND BISALDEHYDE COMPOUNDS (DR. CAMILLE BOUILLON) 230
C. GENERAL PROCEDURE FOR THE SYNTHESES OF OLIGONUCLEOTIDE dT_{10} AND dT_{40} 234
III. SPECTROSCOPIC STUDIES FOR THE RATIONAL APPROACH (COLLABORATION WITH DR. JENIFER RUBIO-MAGNIETO AND DR. MATHIEU SURIN) 234
IV. SMALL MOLECULES-DNA COMPLEXES DETECTION BY MALDI-TOF MASS SPECTROMETRY (ANALYSES PERFORMED BY DR. SONIA CANTEL) 236
V. POLYMERS FORMATION, DEGRADATION AND COMPLEXATION WITH DNA 237
A. SINGLE ACYLHYDRAZONE MODEL REACTION (STUDIED BY DR. CAMILLE BOUILLON) 237
B. FORMATION AND DEGRADATION OF POLYMERS (STUDIED BY DR. CAMILLE BOUILLON) 238
C. ETHIDIUM BROMIDE DISPLACEMENT ASSAY (DR. SÉBASTIEN ULRICH) 238
D. GEL RETARDATION ASSAY (DR. YANNICK BESSIN) 239
VI. DYNAMIC COMBINATORIAL LIBRARIES GENERATION 239
*A. GENERAL METHODS FOR LIBRARIES GENERATION WITH GUANIDINIUM **GUA-ALD** AND COMMERCIAL AMINES/DIAMINES 240*
1. With dT_n in solution 240
2. With DNA on solid support 240
*B. GENERAL METHODS FOR LIBRARY GENERATION WITH GUANIDINIUM **GUA-ALD** AND TRIAZINES **TN^N** 242*
*C. GENERAL METHOD FOR LIBRARY GENERATION WITH GUANIDINIUM **GUA-ALD**, BISALDEHYDE TRIAZINE **TO** AND COMMERCIAL DIAMINES 243*
D. ISOLATION OF LIBRARIES'CONSTITUENTS FOR MALDI-TOF ANALYSES IN PRESENCE OF dT_{10} 245
REFERENCES 246

REMERCIEMENTS

En premier lieu, je tiens à remercier le Pr. Pascal Dumy, mon directeur de thèse, pour m'avoir donné l'opportunité de réaliser ces travaux de thèse, ainsi que mon encadrant de thèse, le Dr. Sébastien Ulrich, pour son aide et ses conseils au cours de ces trois années, malgré les difficultés.

Je remercie également les Dr. Rachid Benhida et Dr. Jean-Serge Remy, qui ont accepté de juger ce manuscrit en tant que rapporteurs.

Un grand merci aux différents collaborateurs de mon projet de recherche sans qui les résultats présentés ci-après auraient été bien maigres. Au sein du laboratoire, ma gratitude va ainsi aux Dr. Camille Bouillon pour l'étude sur les polymères acylhydrazones et Dr. Arnaud Gilles pour la synthèse des triazines bisamines. Je remercie également le Dr. Sonia Cantel, de l'IBMM, pour toutes les analyses de spectrométrie MALDI-ToF sur les complexes petites molécules – ADN, et les Dr. Jenifer Rubio Magnieto et Dr. Mathieu Surin, de l'université de Mons, en Belgique, pour toutes les études de spectroscopie UV et de dichroïsme circulaire.

Je remercie les membres de l'équipe GMR pour leur aide au cours des deux années passées à Montpellier. Merci à Yannick pour sa bonne humeur et sa compréhension, à Eline et Renaud, mes compagnons de bureau pour m'avoir supportée dans les bons comme dans les mauvais jours. Merci à Joanna, Simona, Irene, April et Léa pour leur compagnie : merci pour vos grains de folie quand on en avait terriblement besoin (ou pas ?).

Je tiens également à remercier les membres de l'équipe I2BM, à Grenoble, où j'ai commencé ma thèse. J'ai une pensée particulière pour Romaric et Emilie, les aînés pleins de sagesse, ainsi que Morgane et les garçons pour les bons moments passés ensemble.

Enfin, je remercie ma famille pour son soutien inconditionnel dans les joies comme les épreuves et les galères. Merci à mes parents pour leur patience et leur soutien sans faille qui m'ont donné la force d'aller jusqu'au bout.

ABREVIATIONS

A	Adénine
Ac_2O	Anhydride acétique
ADN	Acide désoxyribonucléique
ARN	Acide ribonucléique
BnOH	Alcool benzylique
Boc	*tert*-ButOxyCarbonyle
Boc_2O	Dicarbonate de di-*tert*-butyle
BOP	Hexafluorophosphate de (benzotriazol-1-yloxy)tris(dimethylamino)phosphonium
C	Cytosine
CD	Dichroïsme Circulaire (« Circular Dichroism »)
CE	Excès de Charge
CE_{50}	Excès de charge nécessaire à une diminution de 50% de la fluorescence (test de déplacement du bromure d'éthidium)
DAPI	4',6-diamidino-2-phénylindole
DCC	*N,N'*-DiCyclohéxylCarbodiimide
DIPEA	*N,N*-DiIsoPropylEthylAmine
DMAP	4-(DiMéthylAmino)Pyridine
DMF	DiMéthylFormamide
DMSO	DiMéthylSulfOxyde
DNPH	2,4-DiNitroPhénylHydrazine
DOSY	Diffusion-Ordered SpectroscopY
E/M	Ratio émission de l'Excimère sur l'émission du Monomère
EDC	*N*-(3-Diméthylaminopropyl)-N'-éthylcarbodiimide
EDTA	Acide Ethylène Diamine Tétra Acétique
Eq.	Equivalent molaire
ESI	Ionisation ElectroSpray
Et_3N	Triéthylamine
G	Guanine
HBTU	Hexafluorophosphate de *N,N,N',N'*-tétraméthyl-*O*-(1*H*-benzotriazol-1-yl)uronium
HEPES	Acide 4-(2-HydroxyEthyl)Pipérazine-1-EthaneSulfonique
HOBt	1-HydrOxyBenzotriazole
HPLC	Chromatographie Liquide Haute Performance (« High-Performance Liquid Chromatography »)
HR	Haute Résolution (« High resolution »)
I	Inosine
IR	Infra-Rouge
K_{Ass}	Constante d'association
LC-MS	Chromatographie Liquide – Spectrométrie de Masse (« Liquid Chromatography – Mass Spectroscopy »)
MALDI	Désorption-Ionisation laser Assistée par Matrice (« Matrix-Assisted Laser Desorption/Ionisation »)
MeOH	Méthanol
PEG	PolyEthylène Glycol
PEI	PolyEthylènImine
PNA	Acide Nucléique Peptidique (« Peptide Nucleic Acid »)

PVDF	PolyFluorure de VinyliDène (« PolyVinyliDene Fluoride »)
R_f	Rapport frontal
R_h	Rayon hydrodynamique
RMN, NMR	Résonance Magnétique Nucléaire (« Nuclear Magnetic Resonance »)
RP	Phase inverse (« Reverse phase »)
$S_{A/R}$	Sélectivité pour l'Amination réductrice par rapport à la Réduction directe
SM, MS	Spectrométrie de masse
T	Thymine
$t_{1/2}$	Temps de demi-vie
TA	Température Ambiante
TAE	Tris, Acétate et EDTA
TE	Tris-EDTA
TFA	Acide TriFluoroAcétique
TFO	Oligonucléotides générateurs de triplexes (« Triplex Forming Oligonucleotides »)
THF	TétraHydroFurane
TIS	TriIsopropylSilane
TLC	Chromatographie sur Couche Mince (« Thin Layer Chromatography »)
T_m	Température de demi-fusion (« Melting Temperature »)
ToF	Temps de Vol (« Time Of Flight »)
t_R	Temps de rétention
U	Uracile
UV	UltraViolet
v/v	Rapport en volume
Vis	Visible

CHAPITRE 1 :

GENERALITES – BIBLIOGRAPHIE

I. L'ADN ET SA RECONNAISSANCE

Les acides nucléiques, sous forme d'acide désoxyribonucléique (ADN) ou d'acide ribonucléique (ARN), sont les supports de l'information génétique de tous les êtres vivants. Ils constituent la base du génome de chaque individu.

A. STRUCTURE CHIMIQUE DE L'ADN

L'ADN est une macromolécule composée d'unités nucléotidiques (Figure 1.1). Un nucléotide monophosphate est composé d'une base azotée (adénine A, cytosine C, guanine G ou thymine T ainsi que les bases méthylées comme la 5-méthylcytosine m^5C et la 5-hydroxyméthylcytosine hm^5C impliquées dans l'épigénome) liée à un sucre (le 2'-désoxyribose), lui-même lié à un groupement phosphate qui permet l'assemblage des unités nucléotidiques par liaisons 5'-3' phosphodiesters.

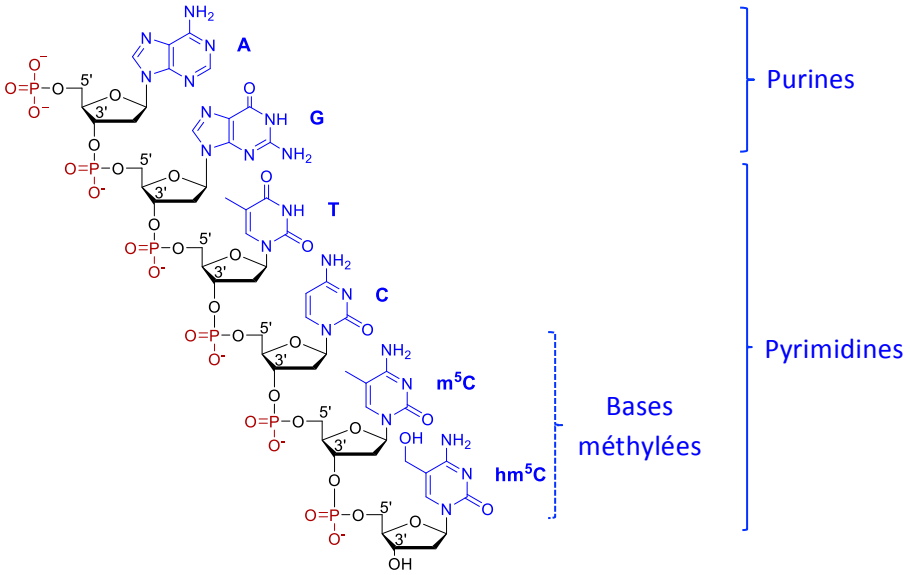


Figure 1.1 : Structure d'un simple brin d'ADN $^{5'}d(AGTCm^5Chm^5C)^{3'}$. Bases azotées en bleu, 2'-désoxyribose en noir et phosphodiester en rouge.

L'existence de quatre bases nucléiques différentes permet le codage de l'information génétique suivant une séquence particulière à chaque individu. Les bases sont complémentaires deux à deux selon les fonctions donneuses et accepteuses de liaisons hydrogène qu'elles comportent. Ces différentes fonctions définissent en effet l'agencement et le nombre de liaisons hydrogène que chacune des bases est susceptible de former. L'adénine

et la thymine peuvent ainsi former deux liaisons hydrogène entre elles, tandis que la cytosine et la guanine peuvent en former trois (Figure 1.2).

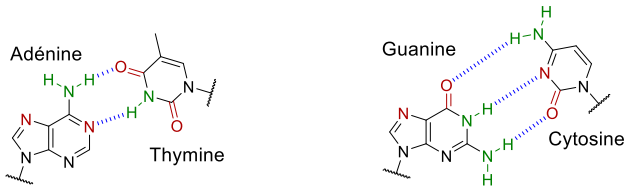


Figure 1.2 : Appariement des bases adénine – thymine (gauche) et cytosine – guanine (droite) selon Watson et Crick. Les fonctions donneuses de liaisons hydrogène sont en vert, les fonctions accepteuses de liaisons hydrogène sont en rouge et les liaisons hydrogène formées sont en bleu.

Un polynucléotide naturel apparenté à l'ADN est l'acide ribonucléique (ARN). L'ARN est composé de nucléotides de la même manière qu'un brin d'ADN à la différence que ses nucléotides contiennent le ribose à la place du 2'-désoxyribose, ce qui rend l'ARN plus instable. La thymine T y est remplacée par la base uracile U (Figure 1.3). D'autres bases peuvent également être présentes dans des brins d'ARN en très faibles proportions, comme l'hypoxanthine (dont le nucléoside correspondant est l'inosine I) apparentée à l'adénine (Figure 1.3).

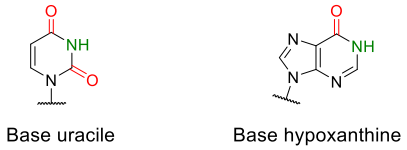


Figure 1.3 : Structure chimique des bases uracile U et hypoxanthine I, présentes en différentes proportions dans l'ARN.

Entre autres fonctions, l'ARN est notamment généralement utilisé par les cellules comme copies de gènes codants pour la synthèse de protéines lors des phénomènes de transcription et de traduction. Lors du processus de transcription, un brin d'ARN est apparié à un brin d'ADN par reconnaissance des bases complémentaires au cours d'un état transitoire dans lequel l'ARN polymérase ouvre la double hélice d'ADN sur une longueur d'une quinzaine de base pour permettre la copie de la séquence d'ADN à exprimer.

En biologie, contrairement à l'ARN, la présence d'ADN natif sous forme de simples brins est limitée. La structure la plus connue de l'ADN, proposée pour la première fois par Watson et Crick en 1953 (Figure 1.4), est celle d'une double hélice formée par hybridation de deux brins d'ADN de séquence complémentaire.[1] Les deux brins d'ADN sont antiparallèles l'un par rapport à l'autre suivant l'axe de la double hélice (la direction de leur axe 5' – 3' est opposée). Les squelettes sucres – phosphodiesters se trouvent repoussés à l'extérieur de la double hélice tandis que les bases des deux brins s'apparient par liaisons hydrogène à

l'intérieur de celle-ci selon une organisation qui est également assistée par empilement π-π entre bases.

Malgré les différences de tailles entre les bases purines (deux cycles aromatiques) et pyrimidines (un cycle aromatique), la géométrie de la double hélice est maintenue par des paires de bases de taille équivalente puisqu'une purine (adénine ou guanine) reconnaît une pyrimidine (thymine ou cytosine). Les paires de bases formées sont également organisées selon des plans perpendiculaires à l'axe de la double hélice, donc selon des plans parallèles entre elles, ce qui permet à des interactions d'empilement entre les cycles aromatiques des bases de stabiliser encore davantage la structure en double hélice. L'intérieur de la double hélice est hydrophobe tandis que l'extérieur est hydrophile et chargé négativement en raison de la présence des phosphodiesters.

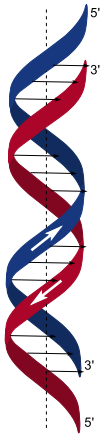


Figure 1.4 : Représentation schématique de la structure en double hélice d'un double brin d'ADN proposée par Watson et Crick en 1953. Les rubans bleus et rouges représentent les squelettes sucres - phosphodiesters des deux brins d'ADN, les traits pleins noirs représentent les paires de bases et l'axe en pointillé symbolise l'axe de la double hélice.

Différents paramètres (séquence, hydratation, environnement) ont une influence sur la géométrie d'une double hélice d'ADN, ce qui conduit à l'existence de plusieurs formes de duplexes d'ADN : les formes A, B et Z sont les trois formes les plus connues (Figure 1.5), mais de nombreuses autres formes ont été générées synthétiquement et décrites dans la littérature.[2]

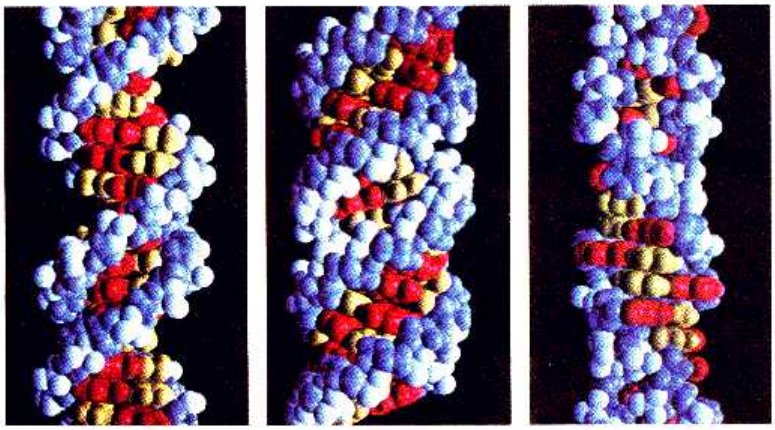


Figure 1.5 : Différentes formes de duplexes d'ADN. De gauche à droite : formes A, B et Z.

La forme B, décrite par Watson et Crick, est la plus fréquemment observée des trois. Cette structure tertiaire est une double hélice orientée vers la droite. Les données de diffraction X indiquent que les paires de bases sont espacées de 3,4 Å selon l'axe de la double hélice. L'hélice faisant un tour complet tous les 34 Å (rotation de 36° entre deux sucres adjacents), il y a donc 10 paires de bases par tour d'hélice.

A l'extérieur de cette structure, deux sillons de taille différente, connus comme les petit et grand sillons, sont présents entre les squelettes sucres – phosphodiesters des deux brins car les liaisons osidiques des paires de bases ne sont pas orientées de manière diamétralement opposées. Le grand sillon est plus large que le petit : cette différence de taille permet une « lecture » des bases par différentes protéines selon le sillon sans déranger l'organisation de la double hélice. Cette structure constitue la forme canonique de l'ADN et est la plus représentée biologiquement.

La forme A, présente pour un taux d'hydratation de l'ADN plus faible, a également été observée par diffraction X pour la première fois en 1953[3] mais l'interprétation des données étant plus difficile, cette structure n'a pas été la première explicitée.[4] Comme la forme B, la forme A de l'ADN est une double hélice composée de deux brins antiparallèles dont la rotation est orientée vers la droite. Cependant, la forme A est plus large et plus courte que la forme B. Avec 11 bases par tour d'hélice, la forme A est aussi plus compacte. Le grand sillon est plus étroit et plus profond que dans la forme B tandis que le petit sillon est plus large et plat.[5]

Les différences structurelles entre les formes A et B sont essentiellement dues à la conformation différente des unités 2'-désoxyriboses dans les deux formes. En effet, alors que les sucres ont une conformation C2'-*endo* dans le cas de la forme B (Figure 1.6), ils adoptent

dans la forme A une conformation C3'-*endo* (Figure 1.6) qui entraîne une inclinaison des paires de bases de 19° par rapport à la normale de l'axe de la double hélice. Cette forme particulière de la double hélice est induite par les polymérases impliquées dans la synthèse biologique de l'ADN et permet à l'enzyme de contrôler la fidélité de la réplication effectuée.[6]

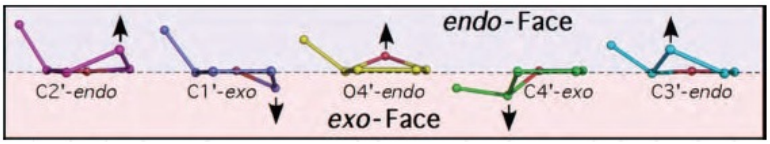


Figure 1.6 : Plissement du sucre. Les faces endo (au-dessus) et exo (en dessous) du cycle à 5 atomes du furanose sont montrées de part et d'autre du plan de référence, la base du nucléotide étant au-dessus de ce plan. La transformation des sucres est illustrée dans l'ordre des conformations C2'-*endo* à C3'-*endo*. Les flèches indiquent l'atome plissé et la direction du plissement. Reproduit à partir de la référence [6].

Caractérisée pour la première fois avec la structure cristalline de l'hexanucléotide d(CGCGCG) en 1979,[7] la forme Z est également une double hélice de deux brins antiparallèles d'ADN mais dont la rotation est orientée vers la gauche contrairement aux formes A et B déjà décrites.[4] Cette forme a été nommée ainsi car lorsque la double hélice est vue de côté, les bases semblent zigzaguer. La double hélice est plus étroite et plus longue que pour les formes A et B. Cette structure particulière est rencontrée pour de courts doubles brins d'ADN de séquence alternant purines et pyrimidines.

Les phosphodiesters étant plus proches les uns des autres avec cette forme Z, l'environnement doit permettre de minimiser la répulsion électrostatique qui a lieu (importante concentration en sels par exemple). Un des rôles biologiques de la forme Z est de réduire les tensions de torsion dans le cas d'ADN surenroulé, c'est à dire lorsque la double hélice est « compactée » sur elle-même pour gagner de l'espace dans la cellule.[6] Cette forme peut aussi être impliquée dans la régulation de gènes ainsi que dans l'édition d'ARN, mais elle peut également être responsable de phénomènes d'instabilité génomique.[6]

Les principales différences structurelles entre les trois formes peuvent être regroupées dans le tableau suivant (Tableau 1.1).

Tableau 1.1 : Données structurelles des différentes formes B, A et Z de duplexes d'ADN cristallisés.[7–9]

	Forme B	**Forme A**	**Forme Z**
Sens de l'hélice	Droite	Droite	Gauche
Diamètre de l'hélice	20 Å	23 Å	18 Å
Unités de répétition	1	1	2
Rotation de l'hélice par paire de bases	35,9°	32,7°	-30° en moyenne*
Paires de bases par tour d'hélice	10	11	12
Inclinaison normale des paires de bases vers l'axe de l'hélice	-6°	+20°	7°
Hauteur d'hélice par paires de bases	3,4 Å	2,6 Å	3,8 Å
Pas d'hélice (longueur d'hélice par tour)	34 Å	28 Å	45 Å
Liaison glycosidique	Anti	Anti	Py : anti Pu : syn
Plissement du sucre	C2'-endo	C3'-endo	Py : C2'-*endo* Pu : C3'-*endo*
Grand sillon	Large et profond	Etroit et profond	Aplati
Petit sillon	Etroit et profond	Large et superficiel	Etroit et profond

Py : pyrimidines ; Pu : Purines. * : -9° pour les successions pyrimidines – purines, -51° pour les successions purines – pyrimidines.

Des triplexes d'ADN peuvent également être observés.[5,10] Ces triples hélices ont été observées pour la première fois entre un duplexe poly(dA)·poly(U) et un simple brin de poly(U), quatre ans seulement après le postulat de la structure en double hélice de l'ADN par Watson et Crick.[11] Dans ce cas, les appariements de bases selon Watson et Crick (Figure 1.2) sont insuffisants pour expliquer les interactions entre trois brins d'ADN ou d'ARN. Des liaisons hydrogène supplémentaires, connues sous le nom d'appariements de Hoogsteen,[12,13] peuvent être formées sur la face des purines exposées dans le grand sillon du duplexe sans déranger les appariements de Watson et Crick déjà formés (Figure 1.7).

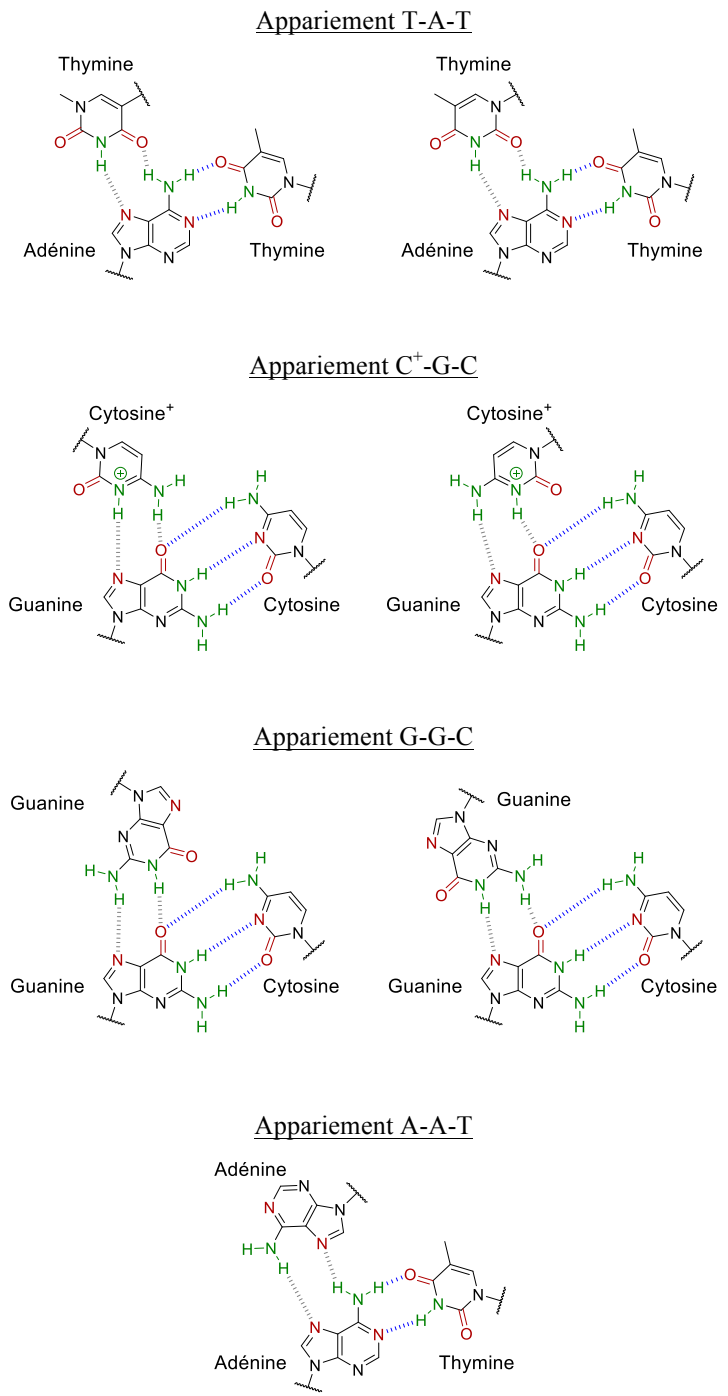


Figure 1.7: Appariements des bases selon Watson et Crick (bleu) et selon Hoogsteen (gris) rencontrés dans une triple hélice d'ADN. A gauche, appariements de Hoogsteen directs (les sucres des deux bases impliquées sont du même côté de l'axe de la paire de base). A droite, appariements de Hoogsteen inverses (les sucres des deux bases impliquées ne sont pas du même côté de l'axe de la paire de base). L'appariement A-A-T contient un appariement de Hoogsteen inverse. Les fonctions donneuses de liaisons hydrogène sont en vert, les fonctions accepteuses de liaisons hydrogène sont en rouge et les liaisons hydrogène formées sont en bleu (Watson et Crick) et en gris (Hoogsteen).

La génération de triplexes nécessite qu'un des différents brins soit homopurique quand les deux autres seront homopyrimidiques ou inversement, qu'un des brins soit

homopyrimidique et les deux autres homopuriques. Cela permet de générer une double hélice *via* des appariements de Watson et Crick pour deux brins de nature différente, tandis que le troisième brin s'enroule dans le grand sillon de la double hélice formée par appariements de Hoogsteen avec le brin homopurine de la double hélice.

Les triplexes peuvent être intermoléculaires ou intramoléculaires. Dans ce dernier cas, ils peuvent correspondre à une structure de l'ADN connue en tant que forme H.[2,5,6] Les triplexes d'ADN sont à l'heure actuelle connus pour avoir un rôle dans la régulation de l'expression des gènes, des structures intramoléculaires ayant été identifiées dans de multiples promoteurs de gènes.[6] Cette structure peut également être formée suite à une double lésion de l'ADN.[6]

Des structures impliquant quatre brins d'ADN peuvent également être formées. Ces quadruplexes peuvent être, au fur et à mesure que l'on s'éloigne de plus en plus du modèle proposé par Watson et Crick, des jonctions de Holliday (forme HJ), des G-quadruplexes (forme G) ou des i-motifs (forme I ou C-quadruplexes).[6]

Les jonctions de Holliday sont des structures intermédiaires permettant l'échange de brins entre deux duplexes d'ADN : elles consistent en quatre doubles hélices de forme B liées entre elles par des « simples brins croisants » (Figure 1.8). Ce type de structure a été proposé en 1964 par Robin Holliday pour expliquer une recombinaison génétique observée entre deux ADN homologues.[14] Par la suite, l'organisation spatiale des quatre brins d'ADN impliqués dans ce genre d'assemblage a été déterminée avec précision par des études cristallographiques[15,16] et est corrélée avec des résultats obtenus en solution.[17] D'un point de vue biologique, les jonctions de Holliday sont utilisées dans les processus de recombinaison et réplication de l'ADN (lors de sa synthèse ou sa réparation) ainsi que pour l'intégration de génomes viraux.[6]

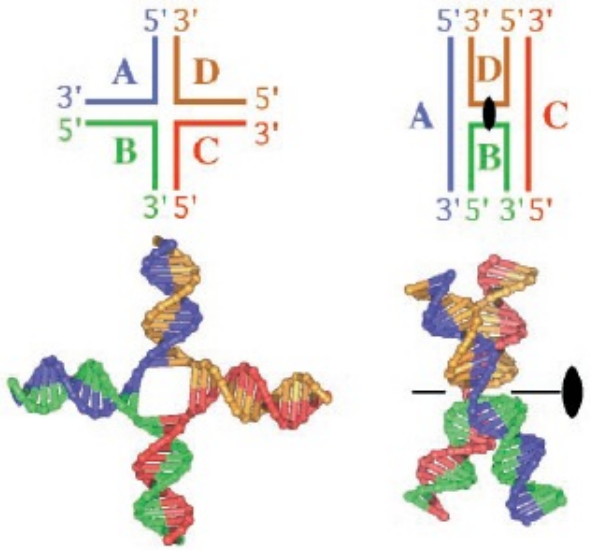


Figure 1.8 : Représentations des conformations des jonctions de Holliday.[16] A gauche : association de quatre brins d'ADN A (bleu), B (vert), C (rouge) et D (jaune) pour former une jonction avec quatre bras double brins (haut) dans une géométrie planaire carrée (forme X étendue, bas). A droite : les mêmes brins associés pour former la structure aplatie en X de la jonction (haut), avec les paires de bras empilés co-axialement comme des doubles-hélices liées par une symétrie d'axe 2 (bas).

Les G-quadruplexes sont formés dans les zones riches en guanines contigües.[18] Les guanines peuvent en effet s'organiser en tétrades autour d'un cation métallique monovalent (Figure 1.9) et deux tétrades ou plus peuvent s'empiler les unes sur les autres pour former le G-quadruplexe. Plusieurs topologies d'empilement sont possibles à partir d'un, deux ou quatre brins d'ADN différents (Figure 1.10).

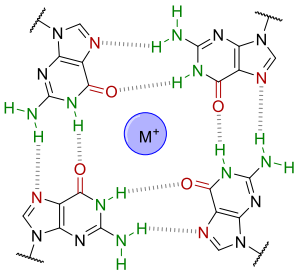


Figure 1.9 : Tétrades de guanines autour d'un cation métallique M^+ (généralement potassium ou calcium).

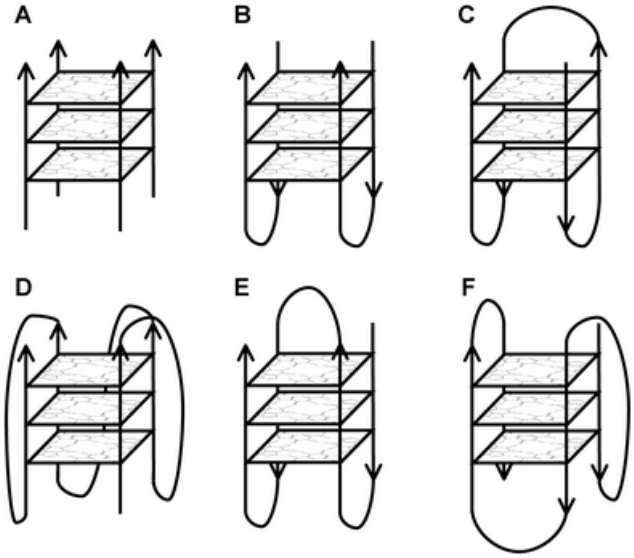


Figure 1.10 : Différentes topologies d'empilement des G-quadruplexes.[18] A : Structure tétramoléculaire avec la totalité des brins parallèles. B : Structure bimoléculaire antiparallèle avec les brins adjacents parallèles. C : Structure unimoléculaire antiparallèlle avec alternance de brins adjacents antiparallèles. D : Structure unimoléculaire parallèle avec trois boucles inversées de double hélice. E. Structure unimoléculaire antiparallèle avec les brins adjacents parallèles et une boucle en diagonale. F : Structure unimoléculaire mixte avec trois brins parallèles et un brin antiparallèle. Les trois structures D, E et F ont été observées dans les répétitions télomériques humaines.

Dans les cellules eucaryotes, les G-quadruplexes sont présents sur les télomères, partie non codante de l'ADN à l'extrémité des chromosomes, ce qui permet de différencier la fin d'un chromosome d'une double rupture d'ADN pour laquelle un processus de réparation du double brin existe.[18] La structure en G-quadruplexes permet donc d'éviter qu'un processus de réparation ou de recombinaison de l'ADN n'ait lieu à un endroit du chromosome non approprié. D'autre part, à chaque cycle de division cellulaire, les télomères se trouvent raccourcis et servent en quelque sorte d'horloge biologique : lorsque le télomère devient trop court, la cellule ne peut plus se diviser (elle entre alors en sénescence).

La télomèrase, exprimée dans les cellules à division multiple comme les cellules souches ou les globules blancs, permet de contrer ce processus de raccourcissement des télomères lorsqu'elle est exprimée. Or, cette enzyme est souvent « détournée » par les cellules cancéreuses pour acquérir leur immortalité. Il a été montré *in vitro* qu'une structure en G-quadruplexe peut inhiber l'activité enzymatique de la télomérase,[19] ce qui pourrait présenter un intérêt thérapeutique si cette propriété était retrouvée *in vivo* chez l'humain. Les G-quadruplexes ont également été observés dans des portions internes du génome, notamment sur des promoteurs de gènes (dont de nombreux oncogènes), et sont dans ce cas-là impliqués dans la régulation de la transcription de ces gènes.[18] Ces observations permettent d'envisager

d'utiliser les G-quadruplexes comme potentielles cibles thérapeutiques, notamment pour de nouveaux produits anti-cancéreux.[20]

Les i-motifs, ou forme I de l'ADN, sont complémentaires aux G-quadruplexes.[6] En effet, extruder une structure de G-quadruplexe intramoléculaire à partir d'une double hélice implique également la sortie d'un simple brin de séquence complémentaire dC du duplexe initial. Comme les G-quadruplexes, les i-motifs peuvent être formés à partir d'un, deux ou quatre brins d'ADN, mais leur séquence est cette fois-ci riche en cytosine et l'appariement des bases ne se fait pas en tétrade comme pour les guanines.

Dans cette structure, décrite pour la première fois de manière détaillée en 1993,[21] les cytosines sont appariées deux par deux (Figure 1.11) et organisées en deux duplexes de brins parallèles intercalés (d'où le nom de i-motifs) en tête à queue (Figure 1.12). Sur les paires de cytosines impliquées dans chacun des deux duplexes, une des deux doit être protonée pour permettre l'appariement, ce qui rend l'édifice sensible au pH du milieu dans lequel il se trouve : les i-motifs sont d'ailleurs stables pour des pH acides à légèrement neutres.

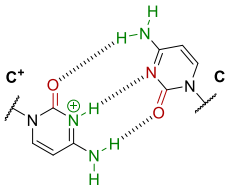


Figure 1.11 : Appariement de deux cytosine (dont une protonée) dans le cas d'un duplexe intercalé au sein d'un i-motif. Les fonctions donneuses de liaisons hydrogène sont en vert, les fonctions accepteuses de liaisons hydrogène sont en rouge.

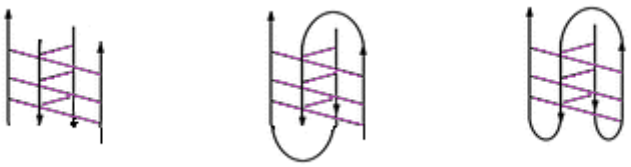


Figure 1.12: Topologies de i-motifs. De gauche à droite : i-motifs à quatre, deux et un brins. Les paires de bases C^+-C sont symbolisées par les traits violets transversaux.

Comme les G-quadruplexes, les i-motifs sont présents sur les télomères des chromosomes et peuvent être reconnus par certaines protéines.[22] Ils ont également été identifiés dans les centromères et sur les promoteurs de certains gènes (notamment des oncogènes comme pour les G-quadruplexes).[22,23] Cependant, le rôle biologique de ces structures n'est pas clairement défini à l'heure actuelle bien que la localisation des i-motifs au sein d'un génome laisse penser qu'ils puissent être impliqués dans la régulation des gènes.

Dans la suite du chapitre, l'ADN sera considéré comme une double hélice de forme B sauf indications contraires.

B. LES DIFFERENTS MODES DE RECONNAISSANCE DE L'ADN

1. Reconnaissance biologique

La reconnaissance de l'ADN au cours de différents processus biologiques est primordiale pour la bonne transmission de l'information génétique et le bon fonctionnement d'une cellule. Lors de la réplication d'ADN (au cours de la multiplication cellulaire) ou la transcription de gènes pour la synthèse de protéines, l'ADN est ainsi reconnu par différentes enzymes. Un certain nombre de structures protéines – ADN ont été déterminées par cristallographie aux rayons X ou RMN et sont décrites dans la littérature (près de 3000 sont reportées dans la base de données NDB).[24]

Les interactions mises en jeu, multiples et complexes, ne seront pas discutées en détail ici. Il est toutefois possible de souligner que la reconnaissance de l'ADN par les protéines se fait généralement au niveau du grand sillon de la double hélice : le site de reconnaissance de la protéine entre ainsi en contact direct avec les bases et interagit également avec le squelette sucre – phosphodiester de l'ADN.[24] Une reconnaissance de l'ADN au niveau du petit sillon est possible bien que généralement peu favorisée par l'étroitesse du sillon.[24]

Les interactions protéine – ADN sont systématiquement non covalentes et peuvent être des liaisons hydrogène, des interactions électrostatiques (ponts salins) ou des interactions hydrophobes. Les interactions avec les bases de l'ADN permettent à la reconnaissance d'être spécifique pour une séquence particulière.

La reconnaissance biologique de l'ADN repose donc sur la chimie supramoléculaire de par la nature non covalente des interactions impliquées. La force des interactions est faible et le phénomène de multivalence est prépondérant pour apporter de la stabilité aux complexes protéine - ADN.

2. Reconnaissance par des ligands synthétiques

Divers ligands synthétiques, de natures et de tailles variées, sont connus pour interagir avec l'ADN. Il est important d'étudier la nature de ces interactions car elles permettent de mieux comprendre les mécanismes à l'origine de la reconnaissance de l'ADN de façon à pouvoir concevoir de futurs ligands synthétiques plus efficaces dans leurs applications.

Cependant, l'environnement aqueux dans lequel se trouve l'ADN peut représenter un challenge pour sa reconnaissance par des ligands synthétiques.[25] En effet, l'eau est un milieu compétitif, bien qu'elle puisse dans certains cas servir d'intermédiaire dans une liaison entre un ligand et sa cible. De plus, par sa nature de polyanion, l'ADN peut former des liaisons hydrogène avec les molécules d'eau qui l'entoure et pour le reconnaître, les ligands synthétiques doivent contourner cette importante hydratation.

Un ligand synthétique peut interagir avec l'ADN selon différents modes de reconnaissance. Une double hélice d'ADN peut ainsi être reconnue de manière non covalente par :

- intercalation au niveau des paires de bases ;
- interaction dans le grand sillon ;
- interaction dans le petit sillon ;
- ou interaction avec le squelette phosphodiester.

La reconnaissance de certains composés pour l'ADN peut également conduire à la formation de liaisons covalentes.

Les ligands reconnaissant l'ADN ayant fait l'objet de nombreuses revues,[24,26–28] seuls quelques exemples significatifs de ligands d'ADN sous sa forme de duplexe seront présentés ici, sans être exhaustifs.

a. Interaction covalente

Des composés se liant de manière covalente à l'ADN peuvent être responsables de mutations génétiques, mais ont également été développés pour « réguler » les phénomènes de réplication et / ou de transcription dans une optique médicale. Ces recherches ont notamment conduit au développement d'agents alkylants pour les traitements anticancéreux.

La liaison covalente avec l'ADN est généralement créée au niveau des bases de manière à déranger l'appariement avec la base complémentaire et engendrer des changements conformationnels de l'ADN pour qu'il ne soit plus « lisible » par les enzymes. Selon sa

structure, un même composé peut interagir de manière covalente en une ou plusieurs positions de l'ADN (Figure 1.13) : cela conduit à la formation d'adduits dits mono-alkylés (une liaison covalente avec l'ADN) ou « réticulés » (deux liaisons covalentes ou plus).

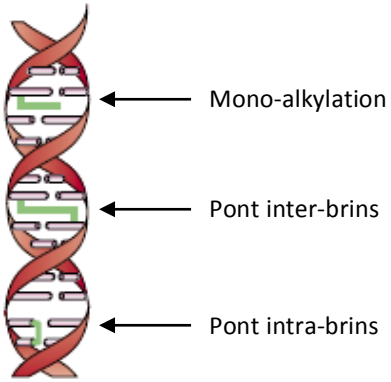


Figure 1.13 : Représentation schématique des adduits à l'ADN possibles lors de la reconnaissance covalente de l'ADN.

Parmi les exemples connus d'agents alkylants, on trouve des composés apparentés au tristement célèbre gaz moutarde comme la chlorméthine ou la cyclophosphamide, et des composés dérivés comme le cisplatine ou le carboplatine (Figure 1.14). Ces agents alkylants réagissent préférentiellement sur les guanines.

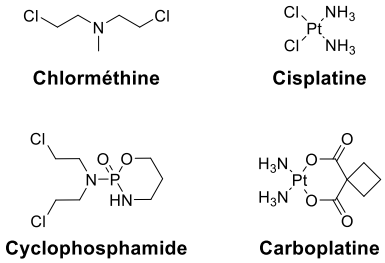


Figure 1.14 : Structures chimiques de quelques agents alkylants.

b. Intercalation

Dans ce mode de reconnaissance, des composés s'insèrent entre deux paires de bases sans en déranger l'appariement. Des cycles aromatiques portés par les molécules intercalantes permettent un empilement de type π-π avec les cycles aromatiques des bases.

On distingue deux types d'intercalation : l'intercalation classique et l'intercalation enfilée. La différence entre ces deux genres d'intercalation réside dans la manière dans les potentielles chaînes latérales à la partie intercalée vont interagir dans les petit et grand sillons de la double hélice. Un intercalant comprenant une ou plusieurs chaînes latérales interagissant dans un seul des sillons sera considéré comme un intercalant classique, tandis qu'une

molécule s'intercalant de manière dite enfilée aura des interactions dans les deux sillons. Ce dernier type d'intercalant a généralement deux chaînes latérales opposées par rapport au système aromatique intercalant : lors de l'intercalation, ces chaînes latérales se trouvent ainsi naturellement dans le petit sillon pour la première et à l'opposé, dans le grand sillon pour la seconde. Des composés bis-[29] et poly-intercalants[30,31] existent également : ces composés comportent plusieurs systèmes aromatiques susceptibles de s'intercaler, liés entre eux par des espaceurs.

Font partie des intercalants classiques des composés tels que : le bromure d'éthidium utilisé comme marqueur de l'ADN, la quinacrine connue pour ses propriétés antimalariales, ou encore les anthracyclines aux propriétés anti-cancéreuses comme la doxorubicine (Figure 1.15). A ces exemples peut être ajouté celui du ditercalinium, qui est un bis-intercalant.

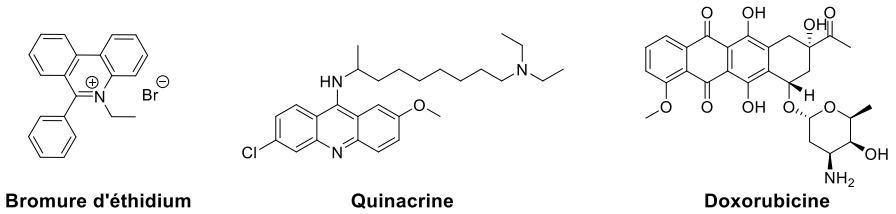


Figure 1.15 : Structures chimiques de quelques intercalants classiques.

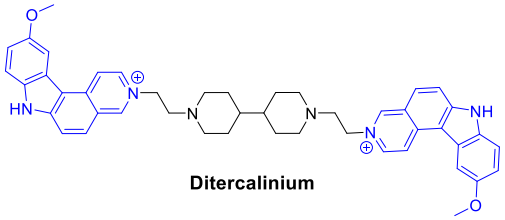


Figure 1.16 : Structure chimique d'un bis-intercalant, le ditercalinium. Les systèmes aromatiques en bleu constituent les parties intercalantes de la molécule tandis que l'espaceur est en noir.

Des exemples significatifs d'intercalants enfilés sont la nogalomycine (produit naturel) ou les dérivés du naphtaldiimide (Figure 1.17).

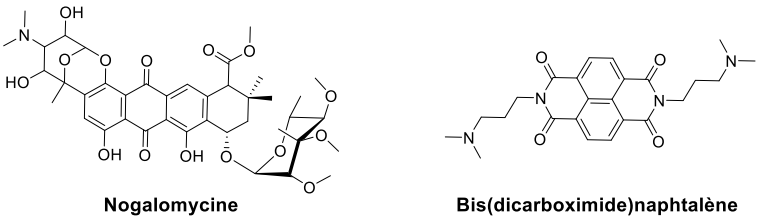


Figure 1.17 : Structures chimiques d'intercalants enfilés.

c. Interaction dans le grand sillon

Comme les protéines, des ligands synthétiques peuvent interagir dans le grand sillon d'une double hélice d'ADN. Cependant, le nombre de ces ligands est plutôt limité en comparaison du grand nombre de ligands connus pour le petit sillon (voir d.). Cette observation s'explique par le fait que les dimensions du grand sillon ne se prêteraient pas particulièrement à sa reconnaissance sélective et spécifique par de petites molécules.

Toutefois, des exemples de petites molécules interagissant dans le grand sillon de l'ADN sont connus : c'est le cas de la spermine, molécule naturelle impliquée dans le métabolisme cellulaire chez les eucaryotes, ainsi que du vert de méthyle[32] utilisé comme colorant pour microscopie biologique (Figure 1.18). Des calixarènes dimériques (Figure 1.19) ont également été conçus récemment pour interagir avec l'ADN et malgré l'absence à l'heure actuelle de données directes obtenues par cristallisation ou RMN, cette famille de composés reconnaitrait le grand sillon de l'ADN.[33]

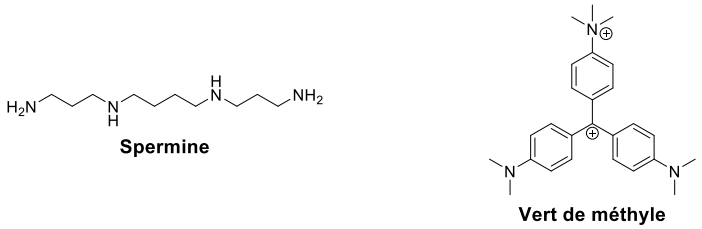


Figure 1.18 : Structures chimiques de petites molécules interagissant dans le grand sillon de l'ADN.

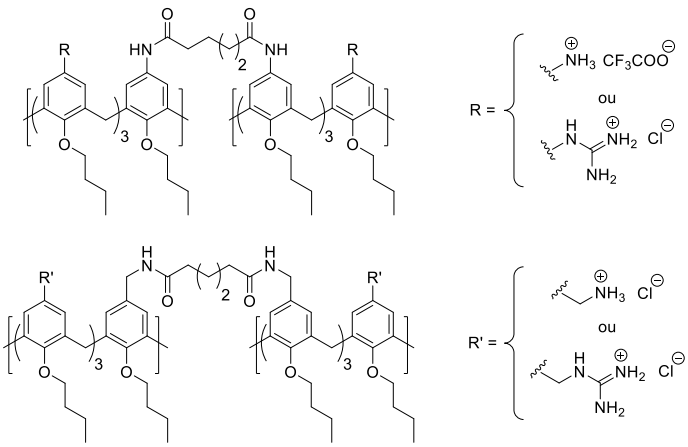


Figure 1.19 : Structures chimiques des calixarènes dimériques capables d'interagir dans le grand sillon.[33]

La majorité des ligands du grand sillon appartiennent le plus souvent à la classe des oligonucléotides générateurs de triplexes (TFO pour « Triplex Forming Oligonucleotides », oligonucléotides naturels ou modifiés chimiquement), ou à la classe des acides nucléiques peptidiques (PNA pour « Peptide Nucleic Acid », Figure 1.20).[28]

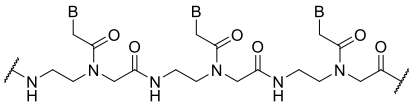


Figure 1.20 : Structure chimique générale d'un acide peptido-nucléique (PNA). B représente les bases azotées naturelles A, C, G, T ou U ; ou des bases synthétiques.

Les TFO interagissent dans le grand sillon de la double hélice en générant une structure de triple hélice canonique telle que décrite précédemment par appariement de Hoogsteen avec les purines de la double hélice (voir I.A.1). La séquence de l'ADN a dans ce cas-là une influence importante sur la bonne reconnaissance du duplexe par l'oligonucléotide synthétique.

L'hypothèse que des PNA pourraient interagir avec un duplexe d'ADN pour former une triple hélice a constitué le point de départ des essais de reconnaissance de l'ADN par un PNA (Figure 1.21 A).[34] Toutefois, bien qu'une triple hélice conventionnelle ait été observée dans certains cas, il s'est rapidement avéré qu'un autre mode de reconnaissance était privilégié : il s'agit de l'invasion du triplexe qui est une forme alternative d'interaction avec le grand sillon de l'ADN puisque celui-ci n'est que partiellement impliqué. Dans ce mode d'interaction, la double hélice d'ADN est ouverte pour permettre à deux brins de PNA d'interagir avec un des brins d'ADN (*via* appariement de Watson – Crick pour l'un des brins de PNA et appariement de Hoogsteen pour l'autre) tandis que le second brin d'ADN reste non apparié (Figure 1.21 B). D'autres modes d'interaction PNA – ADN sont également possibles mais n'ont pas lieu au niveau du grand sillon de la double hélice d'ADN (Figure 1.21 C et D).

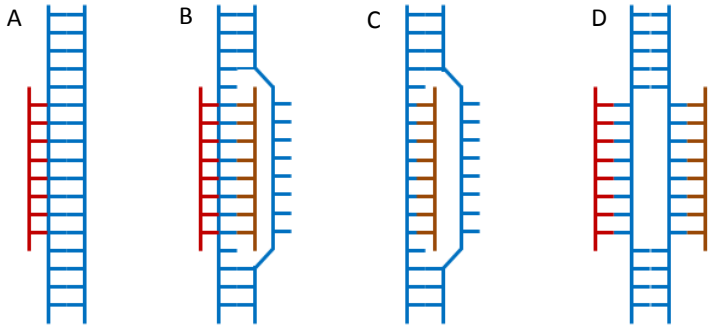


Figure 1.21 : Différents modes d'interaction d'un PNA (brins rouge et orange) avec un double brin d'ADN (bleu). A : triplexe ; B : invasion du triplexe (mode d'interaction le plus stable) ; C : invasion du duplexe ; D : double invasion du duplexe (si les PNA contiennent des bases non naturelles).[34]

d. Interactions dans le petit sillon

De nombreuses petites molécules sont connues pour interagir dans le petit sillon dont les dimensions semblent plus attractives pour ce type de composés. Par rapport à l'intercalation, l'interaction dans le petit sillon implique généralement une meilleure affinité et une plus grande spécificité de séquence.[28] Les interactions impliquées sont non covalentes : interactions électrostatiques et de van der Waals, interactions hydrophobes et liaisons hydrogène sont rencontrées. Les ligands ont généralement une structure courbée pour suivre la courbe du petit sillon.

Parmi les composés susceptibles d'interagir dans le petit sillon de l'ADN, on trouve des produits naturels comme la netropsine et la distamycine A ou des marqueurs d'ADN comme le composé Hoechst 33258 ou le 4',6-diamidino-2-phénylindole, plus connu par son abréviation DAPI (Figure 1.22).[26]

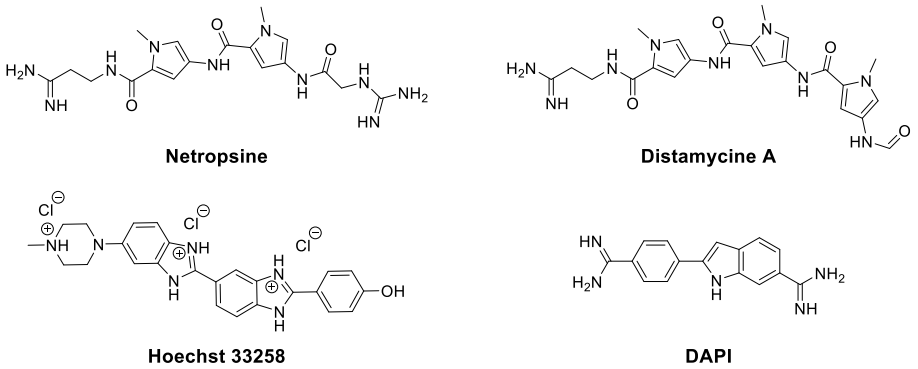


Figure 1.22 : Structures chimiques de quelques ligands du petit sillon de l'ADN.

Pour augmenter la sélectivité de la reconnaissance pour une séquence de bases particulière, allonger le site de reconnaissance des ligands est nécessaire puisque la sélectivité augmente avec le nombre de paires de bases reconnues. Pour concevoir de tels ligands, une des stratégies est de lier entre eux des motifs reconnaissant le petit sillon de l'ADN avec les espaceurs appropriés. Cette méthode, détaillée plus loin, a notamment été utilisée par le groupe de Dervan pour construire des polyamides analogues de la distamycine capables d'interagir avec des séquences spécifiques de bases par reconnaissance *via* le petit sillon de l'ADN (Figure 1.23).[35]

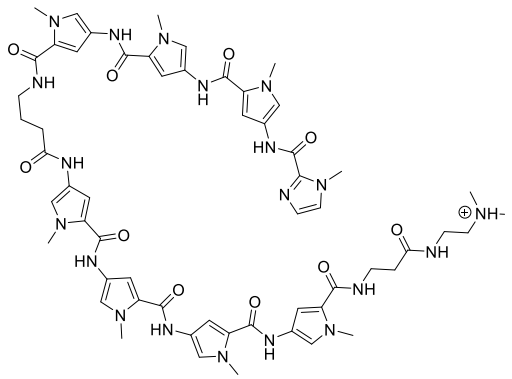
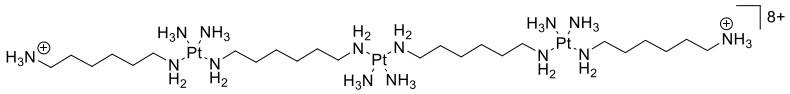


Figure 1.23 : Structure chimique d'un polyamide en épingle à cheveux synthétisé par le groupe de Dervan pour la reconnaissance de l'ADN *via* le petit sillon de la double hélice.[35]

e. Interaction avec le squelette phosphodiester

Assez peu de composés synthétiques sont connus pour reconnaître exclusivement le squelette phosphodiester des brins d'ADN bien que les charges portées par celui-ci soient connues pour limiter la pénétration cellulaire des oligonucléotides synthétiques et puissent être facilement utilisées pour la reconnaissance.

Le premier exemple de ligand synthétique reconnaissant exclusivement le squelette phosphodiester n'a été décrit qu'en 2006 par le groupe de Farrell :[36] il s'agit d'un complexe de platine(II) (**triplatinNC**) comportant trois atomes de platine liés par des espaceurs aminés hydrophobes (Figure 1.24). Les atomes de platine de ce composé ne peuvent pas interagir de manière covalente avec l'ADN comme le cisplatine car ils ne portent pas de groupements labiles tels que le chlore.



TriplatinNC

Figure 1.24 : Structure chimique du premier ligand connu du squelette phosphodiester, le **triplatinNC**.

Cristallisé en présence d'un dodécamère de Dickerson-Drew (DDD, d(CGCGAATTCGCG)$_2$), le **triplatinNC** ne s'intercale pas, ni n'interagit dans les sillons du double brin de conformation B. La structure cristallographique obtenue par diffraction des rayons X montre que le **triplatinNC** interagit sélectivement, *via* ses fonctions amines *cis*, avec les atomes d'oxygènes portés par les phosphodiesters de l'ADN par liaisons hydrogène (complexes NH···O···NH) : ce type d'interaction a été appelé pince à phosphate (Figure 1.25).

Figure 1.25 : Comparaison entre la fourche à arginine et la pince à phosphate.

Le composé **triplatinNC** peut s'organiser suivant le squelette phosphodiester d'un même brin d'ADN (Figure 1.26), ou en enjambant un sillon par interaction avec des phosphodiesters des deux brins d'ADN.

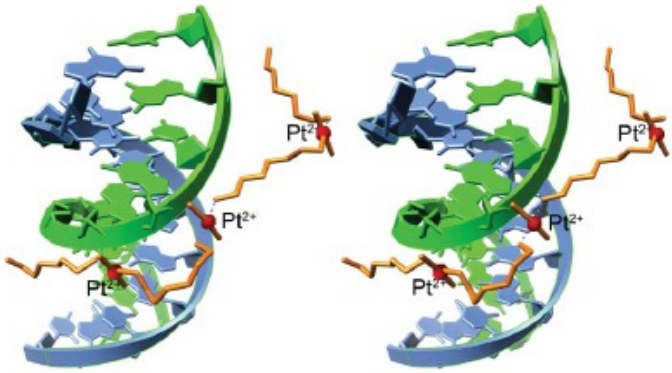


Figure 1.26 : Vue du triplatinNC suivant le squelette phosphodiester d'un brin d'ADN (PDB 2DYW).[37]

Par la suite, un second composé à base de platine, le **triplatinNC-A** (Figure 1.27), a été cristallisé avec un mode de reconnaissance de l'ADN identique.[38] La similarité du mode de reconnaissance des deux composés a également été vérifiée en solution.

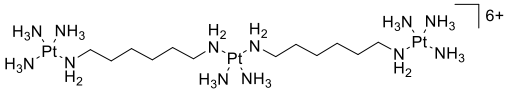


Figure 1.27 : Structure chimique du composé triplatinNC-A.

Lorsqu'ils se complexent avec un double brin d'ADN, ces composés stabilisent les duplexes de manière proportionnelle à la charge globale qu'ils portent. Ils peuvent également induire des changements conformationnels de l'ADN (transitions B → A et B → Z).[39]

L'activité biologique des deux composés a été évaluée et comparée à celle de complexes de platine interagissant de manière covalente avec l'ADN, dont le cisplatine.[40,41] De manière très intéressante, l'absorption cellulaire des composés **triplatinNC** et **triplatin NC-A** est supérieure à celle du cisplatine et augmente avec leur charge globale. Cette plus grande absorption cellulaire permet au **triplatinNC** d'avoir une cytotoxicité comparable à celle du cisplatine dans certaines lignés cellulaires.

Des ligands connus pour reconnaître les oxoanions pourraient reconnaître l'ADN *via* ses phosphodiesters de manière similaire à celle des composés à base de platine. Cette

hypothèse est notamment avancée pour des ligands isophtalanilides et dipicolinanilides (Figure 1.28) qui ont démontré leur capacité à transformer un ADN plasmidique.[42]

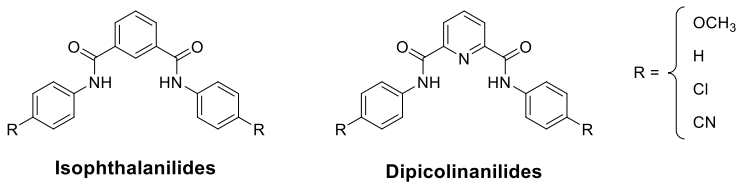


Figure 1.28 : Structures chimiques des dérivés isophthalanilides et dipicolinanilides ayant montré une capacité à transformer un ADN plasmidique et dont l'interaction avec l'ADN se ferait au niveau du squelette phosphodiester.[42]

A l'heure actuelle, aucune démonstration directe de ce type d'interaction n'a encore été publiée pour ces composés. Toutefois, une transformation optimale a été obtenue avec un ratio de 1 pour 1 entre les ligands et le nombre de phosphodiester de l'ADN, ce qui vient corréler l'hypothèse d'une interaction des petites molécules avec l'ADN.

C. AUTO-ASSEMBLAGES MATRICES PAR L'ADN

Puisque le terme d'auto-assemblage n'est pas formalisé, des définitions aussi diverses que ses domaines d'application sont utilisées.[43] Un auto-assemblage peut être défini comme l'organisation réversible de composés indépendants en motifs ou structures sans intervention extérieure. La structure finale d'un auto-assemblage peut être initiée par un composé en particulier, jouant le rôle de matrice pour organiser tous les autres composés présents.

Dans ce cadre, les propriétés de reconnaissance moléculaire de l'ADN en font une bonne matrice pour la formation de nanostructures dont la géométrie, la taille et l'organisation sont contrôlées de manière précise.

Selon la définition d'un auto-assemblage présenté plus haut, une double hélice d'ADN peut elle-même être considérée comme un auto-assemblage. L'organisation précise apportée par les appariements entre bases des polynucléotides ont conduit à la création de structures bi- ou tri-dimensionnelles à base de brins d'ADN, naturels ou modifiés de manière covalente, et au développement des nanotechnologies en ADN.[44] La reconnaissance entre oligonucléotides, généralement modifiés de manière covalente,[45] est à l'origine de nombres d'auto-assemblages pour des applications aussi diverses que la catalyse[46] ou les puces à ADN. Pour des raisons de clarté, ces types d'auto-assemblages ne seront pas évoqués ici. Seuls seront présentés des

exemples d'auto-assemblages impliquant un oligo- ou polynucléotide matrice et des composés d'autre nature.

1. Auto-assemblages par reconnaissance des bases

La reconnaissance des bases de l'ADN par appariement de Watson et Crick peut permettre d'organiser un système de composés initialement désordonné. Des composés amphiphiles ont ainsi été conçus pour reconnaître des bases comme la thymine ou l'adénine d'un brin d'ADN.

Des nanofibres ont pu être observées suite à l'auto-assemblage d'oligo(dA) avec des molécules bolaamphiphiles comportant des thymidines phosphorylées sur leurs têtes polaires (Figure 1.29).[47–52] L'interaction entre les molécules bolaamphiphiles et les brins d'oligo(dA) se fait donc par reconnaissance entre les thymines des bolaamphiphiles et les adénines des oligonucléotides matrices.

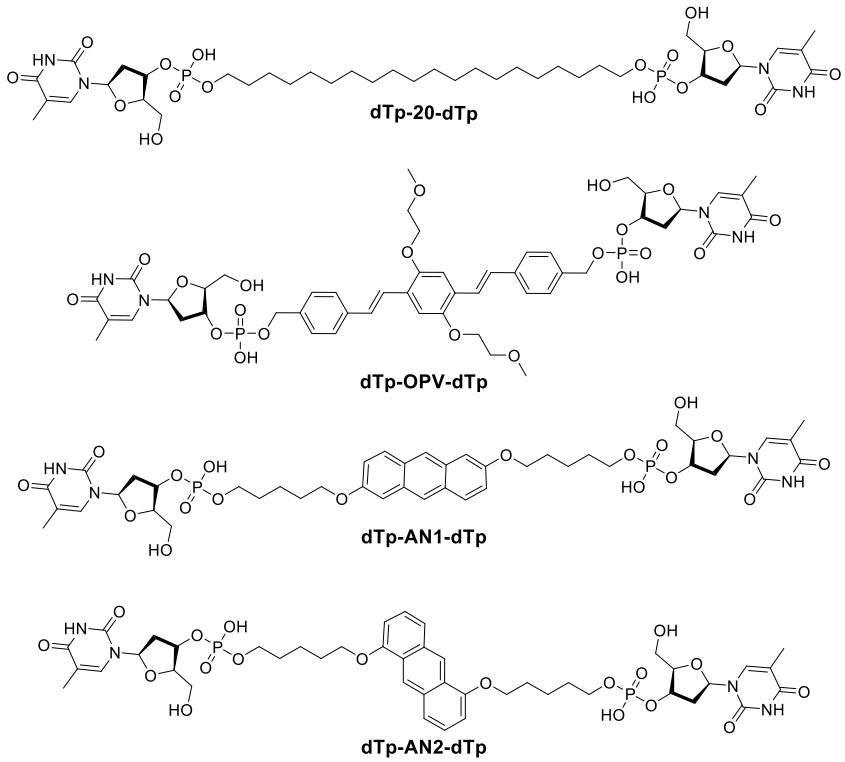


Figure 1.29 : Structures chimiques de molécules bolaamphiphiles connues pour former des nanofibres en présence d'oligo(dA).

Le composé **dTp-20-dTp** a été le premier synthétisé et étudié, d'abord seul,[53,54] puis en présence de différents oligonucléotides[47,49,50] utilisés comme matrices. Il s'est révélé capable de former des hydrogels dans les deux cas, mais avec des caractéristiques physiques

différentes selon qu'un oligonucléotide matriciel est présent ou non. En absence d'oligonucléotides, l'hydrogel formé spontanément se compose d'un réseau de nanofibres de 10 à 30 nm de diamètre et de plusieurs micromètres de long.[53] Sa formation repose sur des liaisons hydrogène entre l'eau et les désoxyriboses, des interactions hydrophobes entre les chaînes carbonées et des empilements π-π des thymines terminales. La formation de l'hydrogel est dépendante de la température ainsi que du pH de l'eau ou du tampon utilisé. Il a également été observé que des microfibres peuvent être générées à partir des nanofibres constituant l'hydrogel en augmentant la concentration de chlorure de sodium.[54]

L'ajout d'oligonucléotides matrices permet d'organiser les nanostructures existantes. Ces oligonucléotides matrices doivent comporter des adénines pour qu'une reconnaissance moléculaire se fasse par appariement de Watson et Crick avec les thymines des têtes polaires du composé **dTp-20-dTp**. Dans un premier temps, les oligonucléotides matrices utilisés ont été des oligo(dA) de longueurs variables (entre 2 et 40 bases) avec un ratio d'une adénine par thymine, ce qui a permis de mettre en relation les différences structurales observées au sein des hydrogels avec la longueur des oligo(dA) utilisés.[47] Ainsi, pour des oligo(dA) comportant moins de dix bases, le réseau de nanofibres demeure similaire à ce qui est observé en absence de matrice. Par contre, pour des longueurs de 10 à 40 bases, des nanofibres en hélice, de 7 à 8 nm de diamètre et de plusieurs centaines de nanomètres de long, ont été observées. Ces dernières ont une longueur et un pas d'hélice proportionnels à la longueur de la matrice utilisée. Les auto-assemblages obtenus sont formés par les composés **dTp-20-dTp** empilés les uns sur les autres entre deux brins d'oligo(dA) (Figure 1.30). L'hélicité des nanofibres peut également être induite en absence de matrice par un rayonnement UV entraînant la photodimérisation de thymines terminales.[49] Toutefois, les nanofibres obtenues de cette manière diffèrent des nanofibres matricées par des oligo(dA) au niveau des dimensions. Il est intéressant de noter que le processus de photodimérisation s'est révélé être réversible et inhibé en présence de dA_6.

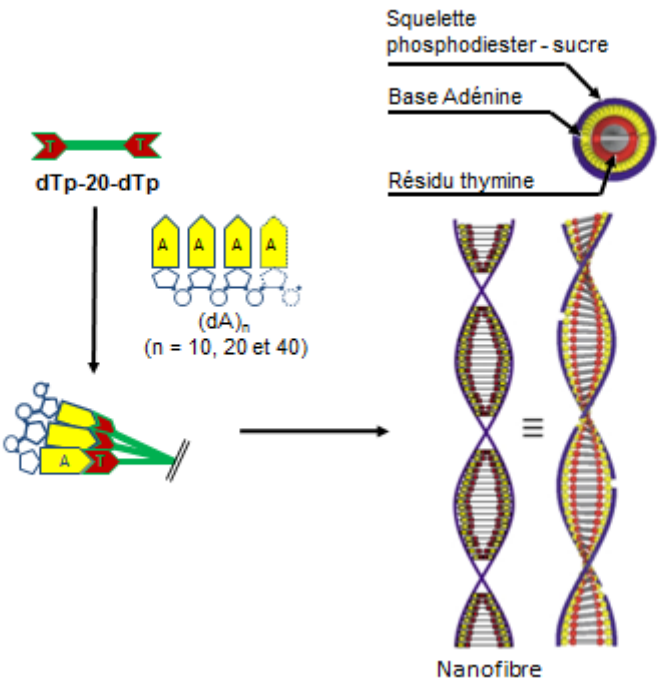


Figure 1.30 : Organisation des auto-assemblages en hélice formés par reconnaissance du composé bolaamphiphile **dTp-20-dTp** par une matrice oligo(dA). Les thymines et adénines sont représentées respectivement en rouge et jaune. Adapté de [47].

Par la suite, l'auto-assemblage du composé **dTp-20-dTp** a également été décrit en présence d'oligonucléotides hétéropolymériques.[50] Pour se faire, le composé **dTp-20-dTp** est mélangé avec deux oligonucléotides matriciels comportant chacun une suite de 20 adénines pour interagir avec les thymines des têtes polaires et une séquence pauvre en adénine (Figure 1.31). Ces séquences pauvres en adénine des deux oligonucléotides matriciels ne sont pas complémentaires entre elles, ce qui empêche toute hybridation intermoléculaire. Lorsque les séquences pauvres en adénines des deux oligonucléotides matrices ne sont pas hybridées avec un oligonucléotide cible complémentaire à ces séquences, les auto-assemblages forment des objets sphériques. Par contre, lorsque les séquences non matricielles sont hybridées par ajout d'un troisième oligonucléotide, les structures initiales constituées d'empilements des bolaamphiphiles **dTp-20-dTp** sont allongées et des nanofibres en hélice sont observées. Ces nanofibres ont un diamètre de 7 à 8 nm, comme pour les auto-assemblages générés avec les oligo(dA), et des longueurs comprises entre 300 nm et 3 µm.

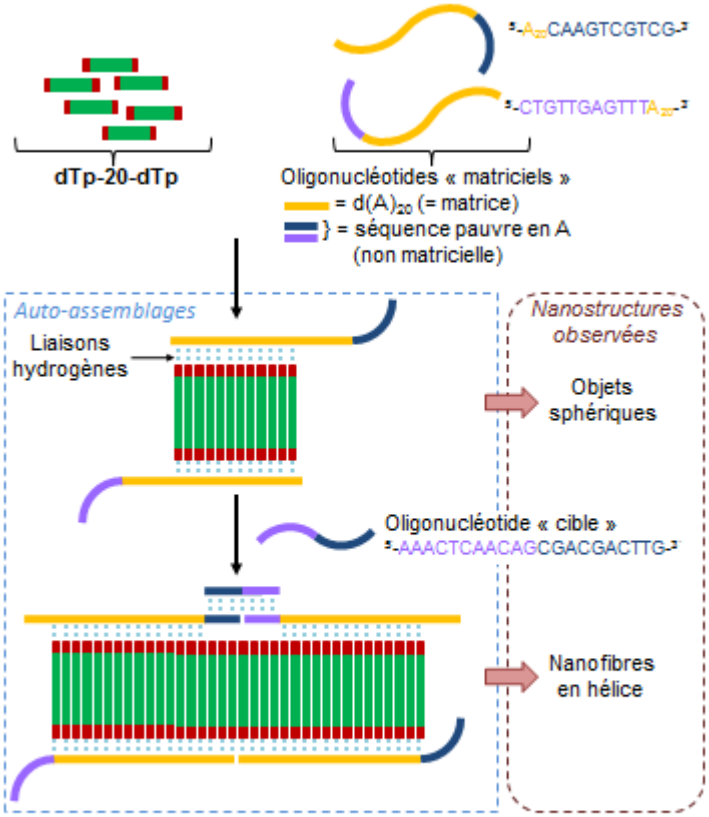


Figure 1.31 : Formation de nanostructures sphériques par auto-assemblage de composés **dTp-20-dTp** avec deux oligonucléotides hétéropolymériques, et de nanofibres en hélice suite à l'hybridation de l'oligonucléotide cible complémentaire des séquences hétéropolymériques des oligonucléotides matriciels. Adapté de [50].

En termes de conception des composés bolaamphiphiles, des variations ont ensuite été introduites au niveau de leur chaîne hydrophobe. Des systèmes aromatiques ont ainsi remplacés la simple chaîne carbonée du composé **dTp-20-dTp** dans le but de permettre des empilements basés sur des interactions π-π au sein des auto-assemblages, ce qui pourrait augmenter leur stabilité.

L'espaceur hydrophobe du composé **dTp-OPV-dTp** (Figure 1.29) est ainsi un oligo(p-phénylènevinylène), connu pour ses propriétés de conduction électronique.[48] Comme il est généralement difficile d'obtenir des structures bien définies d'assemblages d'oligo(p-phénylènevinylène), l'utilisation d'une matrice comme un brin d'ADN pourrait permettre d'organiser le système. Dans le cas du composé **dTp-OPV-dTp**, c'est effectivement le cas puisque que des empilements en hélice droite ont été observés en présence d'une matrice dA_{20}. En présence d'un oligo(dT) non complémentaire, seuls des agrégats amorphes ont été observés, confirmant ainsi l'importance de la reconnaissance des thymines des têtes polaires dans le processus d'auto-assemblage. Les dimensions des nanofibres observées en présence d'oligo(dA) se sont avérées différentes en fonction du ratio de thymine par adénine utilisé : le

diamètre des hélices est ainsi de 6,4 nm pour un ratio T:A de 1:1 alors qu'il n'est que de 5,1 nm pour un ratio T:A de 2:1. Ces différences sont dues à une organisation structurelle des auto-assemblages légèrement différente en fonction du ratio T:A : de manière générale, seule une des deux thymines du composé **dTp-OPV-dTp** est reconnue par un brin d'oligo(dA) avec un ratio de 2:1 alors que toutes les thymines interagissent avec un brin d'ADN pour le ratio 1:1 (deux brins de part et d'autre des composés bolaamphiphiles comme sur la Figure 1.30).

Des groupements anthracènes, connus pour leurs propriétés optiques et électroniques, ont également été introduits sur l'espaceur hydrophobe des composés bolaamphiphiles : les composés **dTp-AN1-dTp** et **dTp-AN2-dTp** (Figure 1.29) ont ainsi été synthétisés et leurs propriétés optiques comparées en présence et en absence de matrice dA_{20}.[51,52] En effet, des assemblages supramoléculaires de molécules aromatiques comme l'anthracène dans lesquels les molécules s'empilent parallèlement face à face (agrégat H, Figure 1.32 gauche) ou avec un décalage des plans en tête à queue (agrégat J, Figure 1.32 droite) modifient les propriétés optiques initiales des molécules : un déplacement bathochrome (vers les plus grandes longueurs d'onde) des bandes spectrales par rapport à la molécule seule est caractéristique d'un agrégat J tandis qu'un déplacement hypsochrome (vers les plus faibles longueurs d'onde) indique un agrégat H.

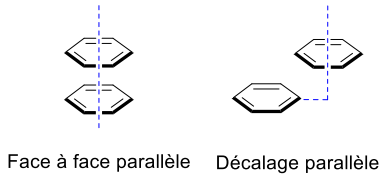


Figure 1.32 : Illustration des géométries d'empilement impliquées dans l'observation d'agrégats H et J.

Le composé **dTp-AN1-dTp** forme ainsi des agrégats J en présence de matrice dA_{20} (ratio A:T de 2:1) contrairement au composé **dTp-AN2-dTp** dont l'organisation au sein de l'auto-assemblage ne permet pas un empilement correct en tête à queue des anthracènes pour former l'agrégat J.[51] Bien qu'en termes de nanostructure, des nanofibres en hélice aient été observées avec le bolaamphiphile **dTp-AN1-dTp** à la fois en absence et en présence de dA_{20}, l'organisation moléculaire au sein des auto-assemblages demeure différente puisque les agrégats J n'ont été observés qu'en présence de matrice. D'ailleurs, les dimensions des nanostructures observées diffèrent selon qu'un oligonucléotide matriciel est présent ou non. Un transfert d'énergie a également été mis en évidence au sein de l'auto-assemblage formé du composé **dTp-AN1-dTp** et la matrice dA_{20}.[52]

D'autres composés comportant une ou deux bases naturelles de l'ADN, sous forme nucléosidique ou non, ont également démontré leur capacité à former des auto-assemblages en présence d'oligonucléotide complémentaire (Figure 1.33).[55–58]

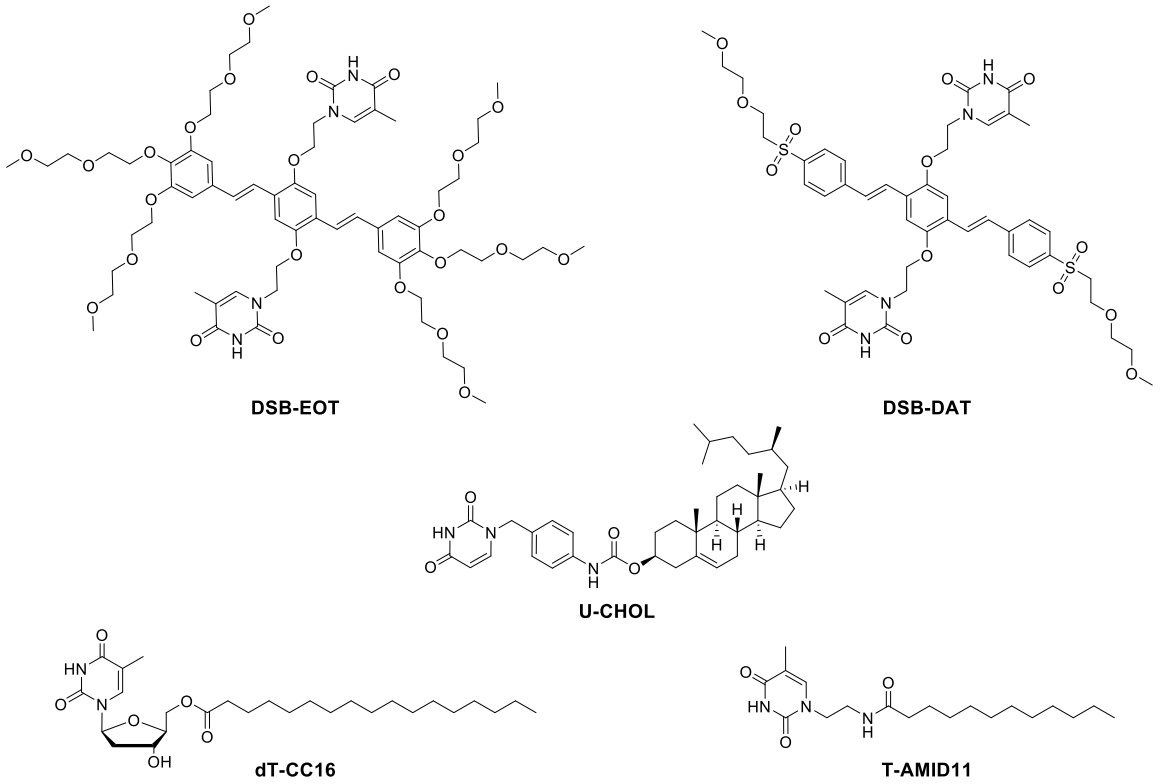


Figure 1.33 : Structures chimiques de molécules comportant des bases thymine ou uracile, et connues pour former des auto-assemblages avec des oligo(dA).

En présence d'un oligo(dA), les composés **DSB-EOT**[55] et **DSB-DAT**[56] reconnaissent les adénines grâce à leurs thymines. Les interactions π-π des groupements distyrylbenzènes centraux et des thymines terminales apportent de la stabilité aux auto-assemblages formés. En présence de dA_{20}, des nanofibres de plusieurs centaines de nanomètres sont observées (diamètre d'environ 6-7 nm en fonction du composé utilisé) ce qui n'est pas le cas pour les composés seuls ou en présence d'oligonucléotides non matriciels comme un dT_{20}. Il est possible de noter que contrairement aux auto-assemblages décrits précédemment, les nanofibres formées n'adoptent pas une structure en hélice bien qu'elles puissent être légèrement distordues.

Contrairement aux composés précédents, les molécules **U-CHOL**, **dT-CC16** et **T-AMID11** n'ont été construites qu'à partir d'une seule base naturelle au lieu de deux bases terminales rencontrées jusqu'à présent. L'auto-assemblage ne va donc pas s'organiser de la même façon en présence de matrice. D'autre part, la formation de ces auto-assemblages a été testée dans des solvants organiques comme le *n*-butanol et le benzène, ce qui nécessite

d'échanger les ions potassiums des polynucléotides utilisés comme matrice par des surfactants cationiques pour les solubiliser.[57]

Le composé **U-CHOL** comporte une base uracile comme site de reconnaissance des matrices ADN. Testé dans le *n*-butanol, ce composé interagit avec les adénines d'un poly(dA) comme attendu, mais également avec les cytosines d'un poly(dC), conduisant à la formation d'un organogel de stabilité comparable quel que soit le polynucléotide ajouté. Toutefois, malgré cette interaction possible avec les cytosines, les nanostructures obtenues ne sont pas les mêmes en fonction du polynucléotide utilisé. Ainsi, les agrégats sphériques formés par le composé seul deviennent des nanofibres en présence de polynucléotides dont les caractéristiques dépendent de la nature du polynucléotide utilisé. En présence de poly(dA), les nanofibres forment des hélices, comparables à des doubles hélices d'ADN, qui ne sont pas observées avec le poly(dC), car les interactions avec l'ADN ne sont alors pas suffisamment fortes pour induire cette structure particulière.

Les composés **dT-CC16** et **T-AMID11** seuls se sont révélés incapables de former des gels stables dans des solvants organiques polaires. Il a donc été choisi de tester la formation d'auto-assemblages avec des polynucléotides directement dans le benzène, qui est un solvant organique apolaire, en solubilisant les polynucléotides par échange de leurs contre-ions cationiques avec un surfactant cationique (le bromure de didodécyldiméthylammonium). Les gels obtenus en présence de poly(dA) présentent des différences claires en terme de stabilité et de morphologie par rapport aux gels obtenus pour les composés seuls ou en présence de poly(dC). Alors que pour les composés seuls, une nanostructure en forme de plaques est observée, des nanofibres sont visibles en présence de poly(dA). Ces résultats démontrent pour la première fois qu'il est possible de générer des auto-assemblages matricés par des polynucléotides dans des solvants organiques, ce qui élargit considérablement les champs de leurs applications potentielles.

Des nucléolipides tels que le composé **diC8PA** ont également montré leur capacité à interagir avec un polynucléotide complémentaire (Figure 1.34).[58] De par son caractère amphiphile prononcé, le composé **diC8PA** s'organise en micelles globulaires comportant la base adénine vers l'extérieur en milieu aqueux : la tête polaire de la molécule peut donc interagir avec une matrice polynucléotidique à l'extérieur des micelles. Les micelles formées sont en effet capables de reconnaître une matrice poly(U), ce qui conduit à la formation d'un réseau hexagonal intercalant poly(U) et cylindres micellaires de **diC8PA** (Figure 1.34 B).

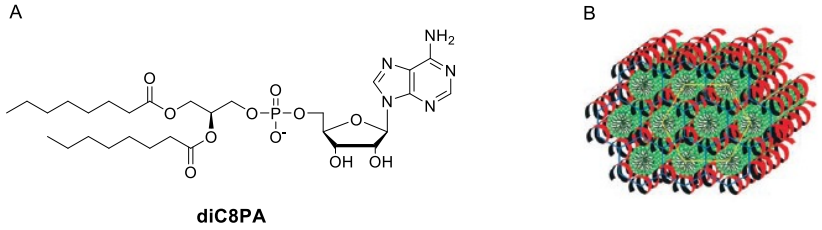


Figure 1.34 : Structure chimique du nucléolipide **diC8PA** (A) et représentation schématique de la nanostructure obtenue avec la matrice poly(U) en rouge et les micelles de nucléolipides en vert (B).[57]

Des composés comportant des bases modifiées peuvent s'auto-assembler avec une matrice oligonucléotide. C'est le cas des composés **Py≡dU** qui comportent un chromophore pyrène lié à une base uracil *via* un espaceur éthynyle (Figure 1.35).[59] Alors qu'il était déjà connu que de telles bases modifiées introduites dans un oligonucléotide s'hybridaient avec un brin naturel complémentaire, ces bases « monovalentes » se sont également révélées capables de s'organiser par empilement le long d'un dA_{17} complémentaire de manière à obtenir un auto-assemblage stable. Les propriétés optiques des auto-assemblages obtenus avec les bases monovalentes sont similaires à celles de leur équivalent oligonucléotide modifié à l'exception de la chiralité des auto-assemblages : alors qu'avec l'oligonucléotide comportant les bases modifiées, l'auto-assemblage adopte la chiralité en hélice droite de l'oligonucléotide matriciel, l'auto-assemblage non covalent conserve la chiralité en hélice gauche des agrégats de **Py≡dU** non matricés.

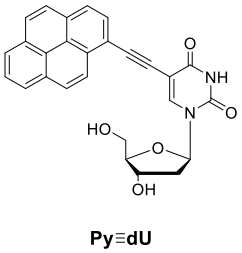


Figure 1.35 : Structure chimique de la base modifiée **Py≡dU** capable de s'empiler le long d'un oligo(dA).

Des auto-assemblages peuvent également être générés par reconnaissance des bases de la matrice ADN par des composés synthétiques ne comportant aucune base naturelle au sein de leur structure. C'est par exemple le cas des diaminotriazines et diaminopurines qui peuvent interagir avec les thymines grâce à leur motif donneur – accepteur – donneur d'hydrogène similaire à celui de l'adénine (Figure 1.36).[60–64]

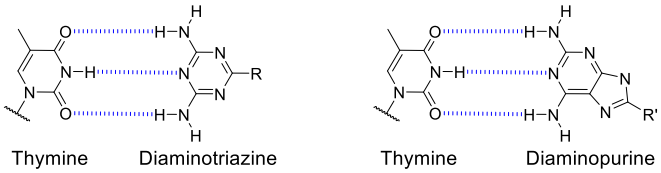


Figure 1.36 : Appariements de Watson et Crick entre une thymine et une diaminotriazine (gauche) ou une diaminopurine (droite), composés synthétiques possédant un motif donneur – accepteur – donneur d'hydrogène similaire à celui de l'adénine, base naturelle complémentaire de la thymine.

Les triazines **TN1** et **TOPV** (Figure 1.37) ont été les premières étudiées en présence de dT_{40}.[60,61] Il a été démontré que les triazines **TN1** interagissent spécifiquement avec les thymines et que cette interaction est totalement réversible.[60]

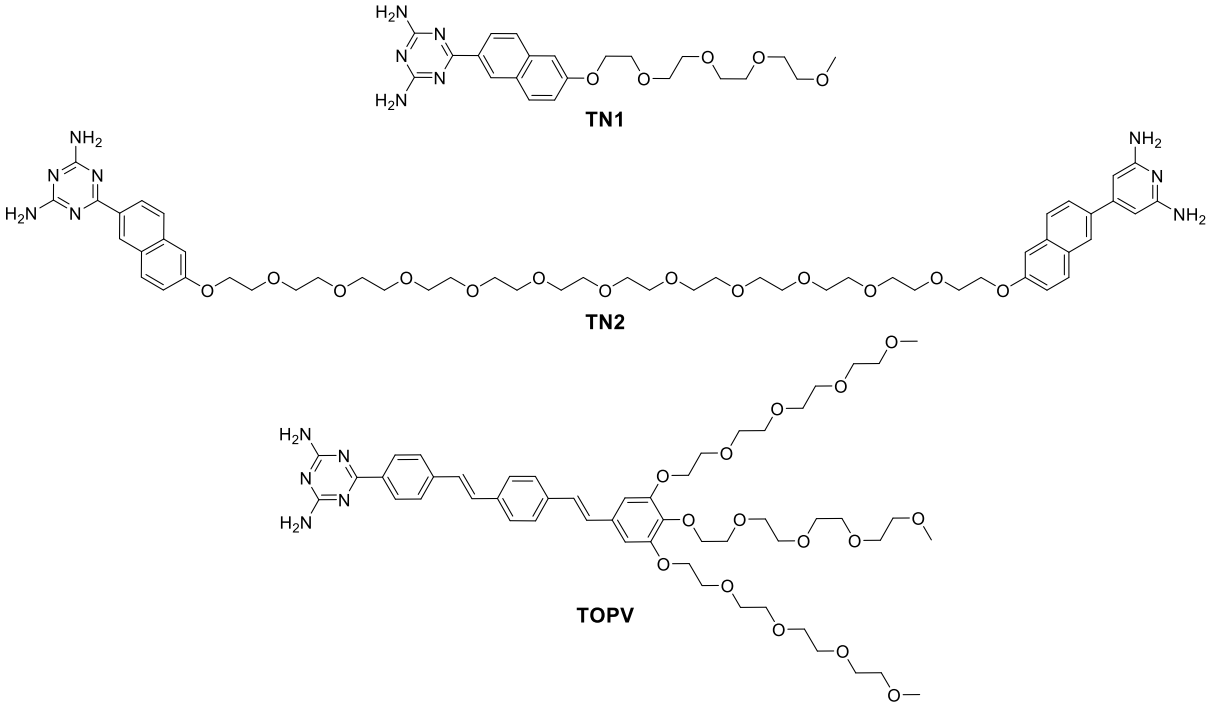


Figure 1.37 : Structures chimiques des diaminotriazines **TN1**, **TN2** et **TOPV** dont les auto-assemblages avec des oligo(dT) ont été étudiés.

D'autre part, l'auto-assemblage des triazines **TN1** avec un dT_{40} adopte la chiralité de la matrice, ce qui conduit à la formation d'une hélice droite comme pour un duplexe dA-dT de forme B. Cependant, l'auto-assemblage **TN1**-dT_{40} est moins stable que le duplexe dA_{40}-dT_{40}. En terme de nanostructures, l'auto-assemblage **TN1**-dT_n adopte une forme de fil enroulé dont la longueur s'échelonne entre 100 et 700 nm.[61]

Comme pour la triazine **TN1**, la triazine **TOPV** adopte une conformation en hélice droite suite à l'interaction avec un dT_{40}.[60] La stabilité de l'auto-assemblage **TOPV**-dT_{40} est plus importante que pour le complexe **TN1**-dT_{40} (température d'élongation T_e plus élevée), probablement en raison des interactions π-π plus étendues. Ces interactions supplémentaires

entre les triazines **TOPV** expliqueraient pourquoi les nanostructures obtenues avec celles-ci en présence de dT_n sont plus longues (de l'ordre du micromètre) que celles des auto-assemblages **TN1**-dT_n.[61] Des simulations de modélisation moléculaires ont par ailleurs montré que lorsque les interactions entre ligands dans le complexe sont plus importantes, la force des liaisons hydrogène entre ligands et matrice diminue et les interactions d'empilement des ligands ont une plus grande influence sur la structure de la matrice.

Dans le but d'augmenter la stabilité des auto-assemblages formés entre les triazines **TN1** et les oligo(dT) matriciels, la triazine divalente **TN2** a été conçue à partir de deux diaminotriazines chacune substituée par un naphtalène et liées entre elles par un espaceur éthylèneglycol long et flexible, ce qui n'induit aucune pré-organisation des diaminotriazines pour reconnaître les thymines des matrices oligo(dT).[63] Il a été effectivement démontré que l'auto-assemblage **TN2**-dT_{40} est plus stable que son équivalent **TN1**-dT_{40} (température d'élongation plus élevée). D'autre part, contrairement aux composés bolaamphiphiles décrits précédemment, une titration de la triazine divalente **TN2** par un dT_{40} met en évidence une stœchiométrie de 20 pour 1. Ce ratio indique que la triazine **TN2** n'interagit pas avec deux brins d'oligonucléotides différents *via* ses deux triazines, mais sur deux thymines consécutives d'un même brin d'oligo(dT).

La diaminopurine **P** (Figure 1.38) a été conçue avec un système aromatique plus étendu que celui de la triazine **TN1** dans le but initial d'augmenter l'efficacité de la reconnaissance des thymines, mais cela s'est avéré contre-productif puisque l'auto-assemblage **P**-dT_{40} est globalement moins stable que son équivalent **TN1**-dT_{40}.[64]

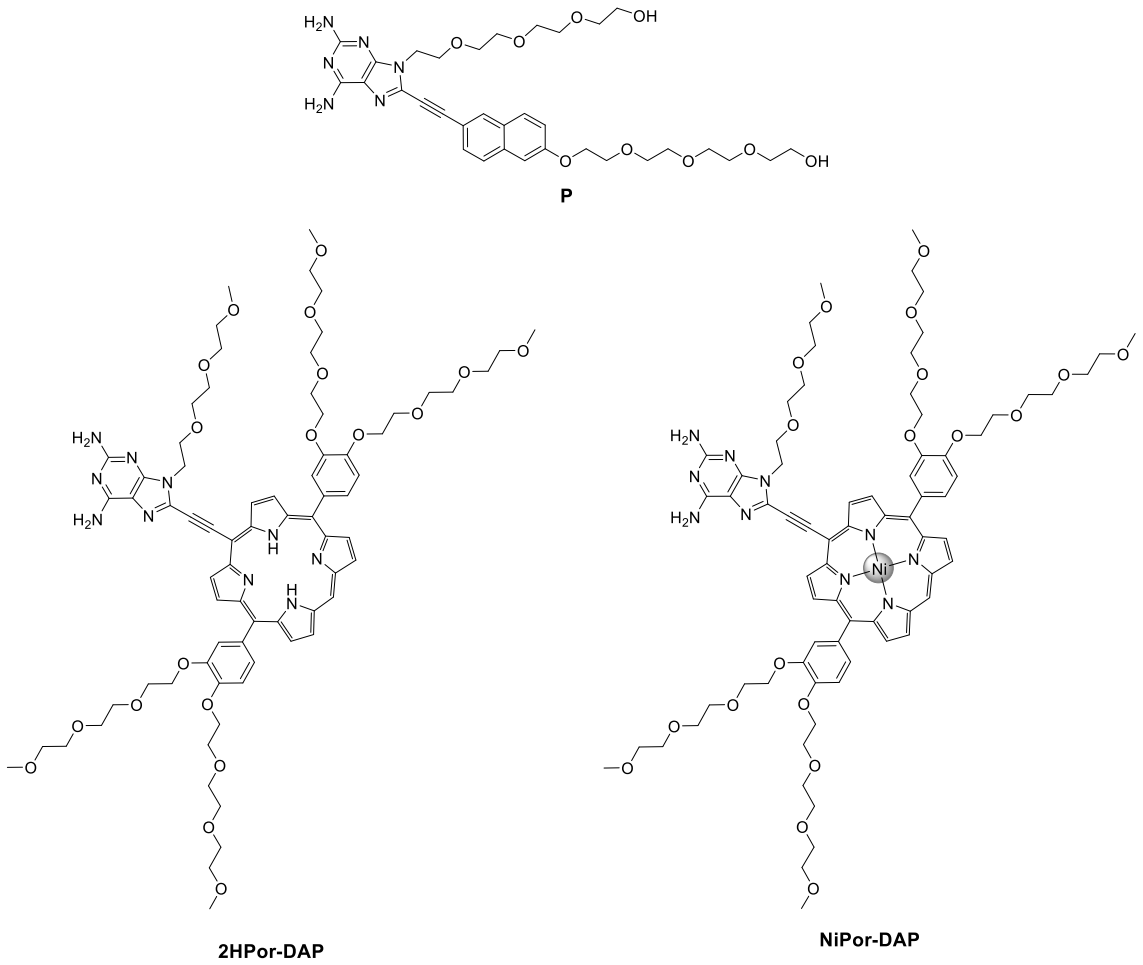


Figure 1.38 : Structures chimiques des diaminopurines **P**, **2HPor-DAP** et **NiPor-DAP** dont les auto-assemblages avec des oligo(dT) ont été étudiés.

Cependant, une propriété intéressante de l'auto-assemblage **P**-dT_{40} a été mise en évidence : en effet, alors qu'à pH 7, cet auto-assemblage adopte une hélicité droite, cette hélicité s'inverse et devient gauche à pH 3, suite à la protonation de la diaminopurine **P**. Si l'inversion d'hélicité était connue pour les différentes formes d'ADN naturelles (formes A et B d'hélicité droite, forme Z d'hélicité gauche), il s'agit d'une première pour des auto-assemblages matricés par l'ADN. Ainsi, bien que l'auto-assemblage soit moins stable avec la diaminopurine **P** qu'avec la triazine **TN1**, la potentielle inversion d'hélicité demeure une caractéristique intéressante pour la conception de nouvelles nanostructures.

Des diaminopurines fonctionnalisées par des porphyrines comportant ou non un ion métallique **2HPor-DAP** et **NiPor-DAP** forment également des auto-assemblages en hélice en présence d'une matrice oligo(dT) comme un dT_{40}.[65,66] L'auto-assemblage du composé **NiPor-DAP** avec une matrice dT_{40} présente d'ailleurs en microscopie une nanostructure en fibres hélicoïdales de quelques micromètres de longueur. La chiralité de l'hélice peut être contrôlée par la vitesse d'hybridation, le ratio chromophore - matrice et la force ionique de la solution

lors de l'organisation de l'auto-assemblage. Ainsi, pour les deux composés **2HPor-DAP** et **NiPorDAP**, une vitesse d'annelage faible induit la formation d'une hélice droite tandis qu'une vitesse plus importante conduit à la formation d'un auto-assemblage en hélice gauche. Par contre, le sens des hélices auto-assemblées est peu sensible aux variations de température et de pH (chiralité préservée jusqu'à plus de 85°C et pour des pH compris entre 3 et 12). Avec ces propriétés structurelles, optiques et cette stabilité en température et pH, la conception de nanomatériaux optiques chiraux peut être envisagée à partir de ces porphyrines.

2. Auto-assemblages par interaction dans les sillons

Reconnaître les bases de l'ADN n'est pas la seule façon de former des auto-assemblages matricés par l'ADN. Des interactions au niveau des sillons peuvent également engendrer des auto-assemblages, bien que la sélectivité de la reconnaissance soit *a priori* moindre.

Les cyanines, utilisées comme colorants, sont connues pour être des ligands du petit sillon. Ces composés forment facilement des agrégats de plusieurs molécules dont les propriétés optiques diffèrent sensiblement des monomères. Parmi ces molécules, les dérivés benzothiazoles sont connus pour former des dimères face à face dans l'eau à des concentrations de l'ordre du micromolaire. Les composés **DiSC$_2$(5)** ou **DiSC$_{3+}$(5)** font partie de cette dernière catégorie et ont démontré leur capacité à former des auto-assemblages sous forme de dimères avec des duplexes d'ADN (Figure 1.39).[67,68] Alors que la majorité des composés aromatiques rencontrés dans des auto-assemblages s'organisent par empilement *via* des interactions π-π, des dimères individuels s'insèrent ici dans le petit sillon du duplexe d'ADN et s'y agrègent côte à côte de manière coopérative (Figure 1.40).

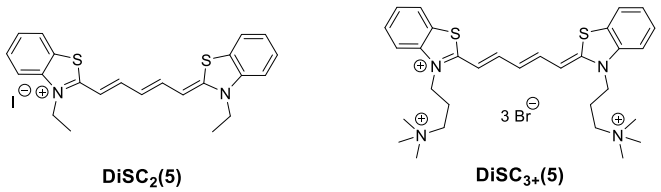


Figure 1.39 : Structures chimiques des cyanines **DiSC$_2$(5)** et **DiSC$_{3+}$(5)**.

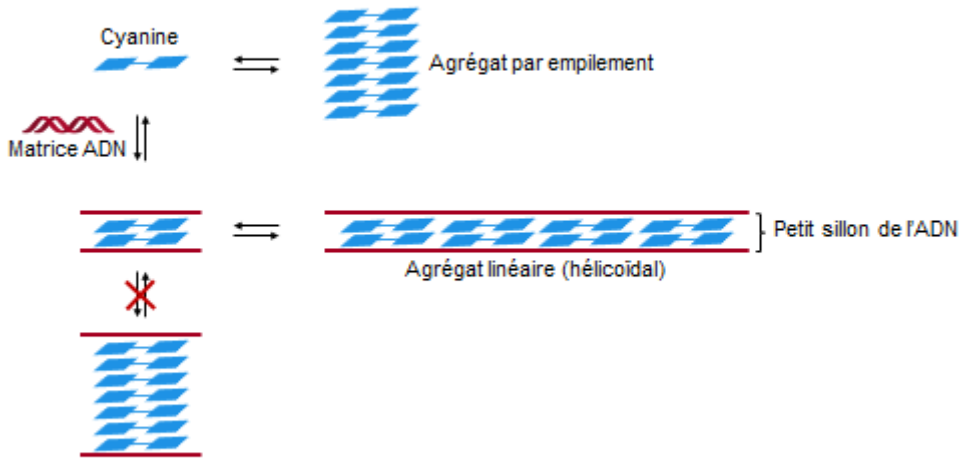


Figure 1.40 : Organisation des agrégats des cyanines **DiSC$_2$(5)** et **DiSC$_{3+}$(5)** en présence d'une matrice ADN.

Le composé **DiSC$_2$(5)** s'insère dans les petits sillons contenant des alternances dA-dT ou dI-dC sous forme de dimère, alors qu'il ne s'insère que sous forme de monomère pour des alternances dG-dC : la géométrie des petits sillons concernés explique cette différence.[67] Des agrégats H en hélice droite ont été observés avec des [poly(dA-dT)]$_2$ et [poly(dI-dC)]$_2$. La longueur de ces agrégats est contrôlée par le duplexe d'ADN : l'agrégation cesse lorsque le petit sillon de la matrice ADN est plein.

En présence de duplexe [poly(dI-dC)]$_2$, le composé **DiSC$_{3+}$(5)** forme également des agrégats de dimères en hélice droite, bien que leur stabilité soit moindre en raison des répulsions électrostatiques plus importantes.[68] Contrairement au composé **DiSC$_2$(5)**, le type d'agrégat formé n'est pas exclusivement un agrégat H mais peut également être un agrégat J sous certaines conditions. La formation de tel ou tel agrégat dépend fortement de paramètres comme la longueur de la matrice ADN, la concentration du composé **DiSC$_{3+}$(5)** utilisée, la température ou encore la force ionique du milieu.

3. Auto-assemblages par interactions électrostatiques (squelette phosphodiester)

Les auto-assemblages reposant sur la reconnaissance des charges du squelette phosphodiester de la matrice ADN ne présentent aucune sélectivité pour une séquence particulière de la matrice.

Des mimes de capside de virus ont été conçus à partir de différents peptides portant une charge positive.[69] Les trois composés utilisés comme monomères comportent une spermine chargée positivement (pour interagir avec l'ADN), un peptide connu pour s'agréger selon une structure en faisceau torsadé d'hélice, et une longue chaîne poly(éthylèneglycol)

permettant la solubilité dans l'eau (Figure 1.41). Ces molécules s'assemblent pour former des nanostructures heptamériques induites par la présence du peptide.

En présence d'un ADN double brin linéaire ou plasmidique, un auto-assemblage peptides-ADN se forme. Si le ratio peptides-ADN est suffisant, les charges négatives du squelette phosphodiester sont neutralisées par les charges positives des spermines terminales des peptides : il semble donc vraisemblable que l'interaction peptides-ADN se fasse au niveau du squelette phosphodiester de l'ADN matrice.

Les nanostructures des auto-assemblages obtenus avec un ADN plasmidique (pbr322) diffèrent selon les chaînes PEG des ligands peptidiques. Ainsi, en absence de chaîne PEG (**Sp-PEP-NH_2**), seule une masse amorphe est observée. Avec une chaîne PEG_{2000} (**Sp-PEP-PEG_{2000}**), un ensemble hétérogène de nanostructures est visible, parmi lesquelles des complexes linéaires (en baguettes), branchés ou toroïdaux. Avec la plus longue chaîne PEG_{5000} (**Sp-PEP-PEG_{5000}**), des nanostructures bien plus homogènes sont observées, à la fois dans leurs dimensions et leurs morphologies : il s'agit de filaments dont la longueur correspond à la longueur des ADN plasmidiques surenroulés utilisés aux incertitudes de mesure près. Cette tendance est similaire lorsqu'un ADN linéaire est utilisé. De nombreuses applications biologiques de ces auto-assemblages sont envisageables, en vectorisation notamment, mais aucun exemple n'est encore décrit dans la littérature à l'heure actuelle.

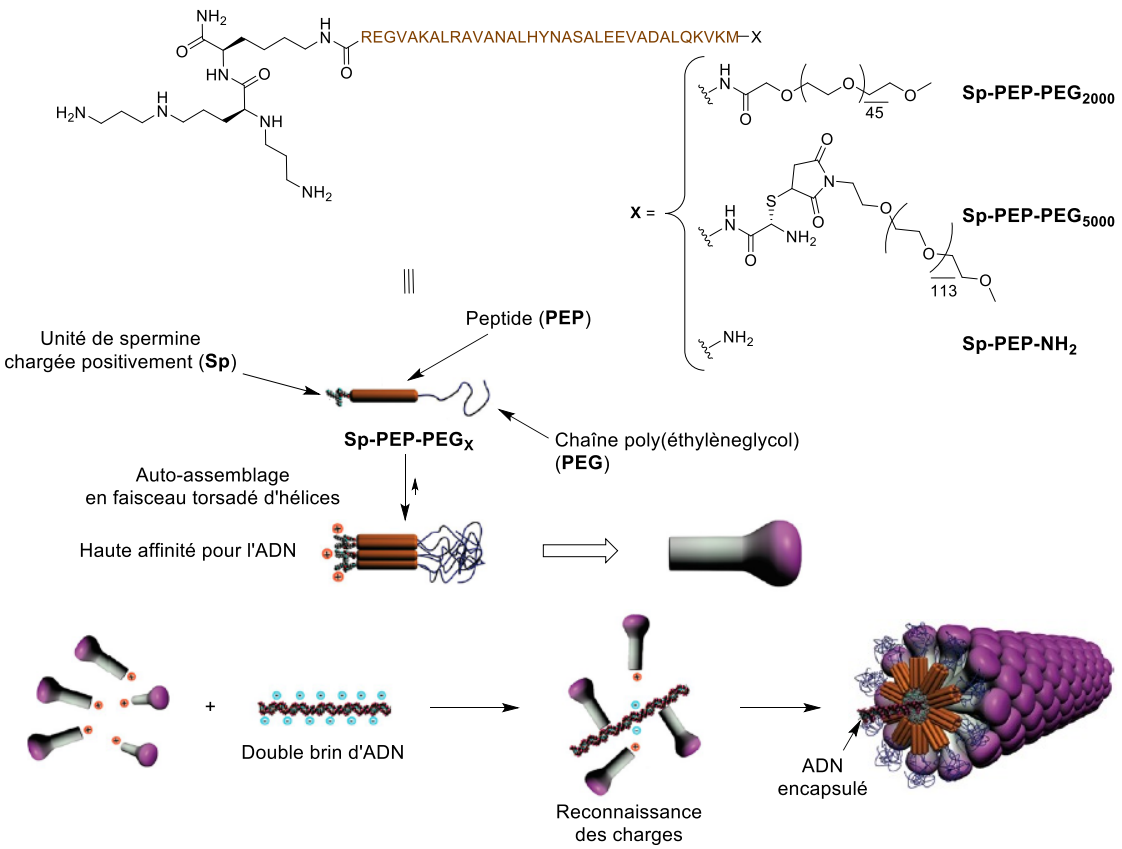


Figure 1.41 : Représentation schématique illustrant la formation d'auto-assemblages mimes de capsides de virus à partir de peptides substitués par un groupement spermine et une chaîne poly(éthylèneglycol) et de matrices ADN. Adapté de [69].

II. METHODES D'IDENTIFICATION DE LIGANDS DE L'ADN

Un ligand, naturel ou synthétique, reconnaît le plus souvent sa cible (ou son récepteur) par des interactions non covalentes. La majorité des modes de reconnaissance de l'ADN explicités plus haut repose d'ailleurs sur des interactions supramoléculaires.

La conception de nouveaux ligands de l'ADN n'a pas toujours suivie une démarche rationnalisée. Beaucoup d'entre eux ont en effet été découverts par hasard. Toutefois, notre connaissance de l'ADN s'améliorant, plusieurs ligands ont été identifiés en utilisant des approches dite rationnelle ou combinatoire dynamique.

A. APPROCHE DITE RATIONNELLE

L'approche dite rationnelle est une méthode d'identification des ligands de type « relation structure – activité » très utilisée en chimie médicinale. Elle consiste à synthétiser différents composés, puis à procéder à des tests de reconnaissance de ces composés en tant que ligand (dans notre cas). A la suite de ces premiers tests, un composé de tête (« lead compound »), présentant la meilleure affinité ou reconnaissance pour l'ADN, peut être identifié et sa structure est ensuite optimisée en variant les différents substituants et en évaluant l'effet de chacun sur l'affinité ou reconnaissance pour l'ADN.

Cette approche a été utilisée par le groupe de Dervan pour concevoir un système synthétique capable de distinguer des séquences données au sein d'un duplexe d'ADN par reconnaissance au sein du petit sillon. La reconnaissance sélective de séquence de bases par des molécules naturelles comme la netropsine et la distamycine A a constitué le point de départ de leurs nombreuses études.[70] L'insertion de dimère de distamycine A dans le petit sillon de l'ADN est fréquente et ses dérivés synthétiques s'insèrent également facilement sous forme de dimère dans le petit sillon. Ces observations ont conduit à la synthèse de peptides en épingles à cheveux, mimes des dimères de polyamides linéaires observés mais dont la spécificité pour une séquence donnée est plus importante (Figure 1.42).[71]

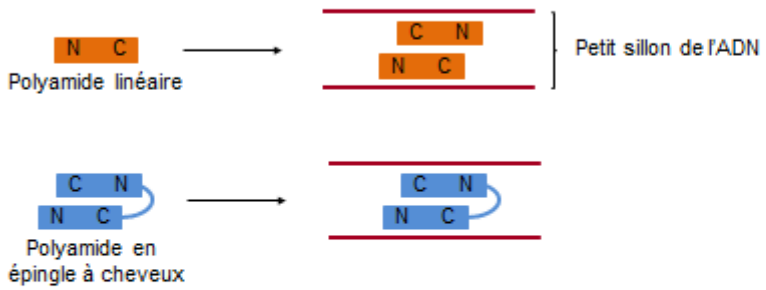


Figure 1.42 : Insertion de polyamides linéaire ou en épingle à cheveux dans le petit sillon de l'ADN. Les lettres N et C désigne les fonctions amines et acides carboxyliques « terminales » des polyamides considérés.

La forme particulière en épingle à cheveux de ces polyamides est obtenue en liant deux polyamides constitués d'une succession d'acides aminés aromatiques synthétiques (Figure 1.43) par des espaceurs comme la β-alanine ou l'acide γ-aminobutyrique (GABA) qui induisent un repliement de la molécule finale sur elle-même.

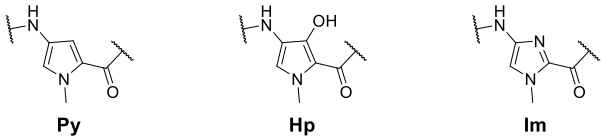


Figure 1.43 : Structures chimiques des acides aminés synthétiques pyrole **Py**, 3-hydroxypyrrole **Hp** et imidazole **Im** utilisés pour coder les séquences de bases à reconnaître par le ligand dans le petit sillon.

Des variations structurelles au sein des différents polyamides en épingle testés ont permis d'observer que les agencements face à face de trois acides aminés aromatiques **Py**, **Hp** et **Im** utilisés déterminaient les sélectivités de séquence observées (Tableau 1.2) : ces règles simples permettent de concevoir des composés polyamides pour cibler directement une séquence d'ADN donnée.[35,72,73] Les interactions à l'origine des spécificités de reconnaissance observées ont pu être déterminées.[74,75]

Tableau 1.2 : Disposition face à face des acides aminés synthétiques **Py**, **Hp** et **Im** codant pour la reconnaissance de paires de bases par interaction dans le petit sillon. Reconnaissance favorable +, défavorable -.

Paires face à face	G-C	C-G	T-A	A-T
Im/Py	+	-	-	-
Py/Im	-	+	-	-
Hp/Py	-	-	+	-
Py/Hp	-	-	-	+

La méthode utilisée par Dervan et son équipe constitue un exemple très complet d'approche rationnelle utilisée pour identifier des ligands de l'ADN. Les résultats peuvent être appliqués pour concevoir des ligands d'ADN reconnaissant des séquences de l'ordre d'une dizaine de bases (Figure 1.44).[76]

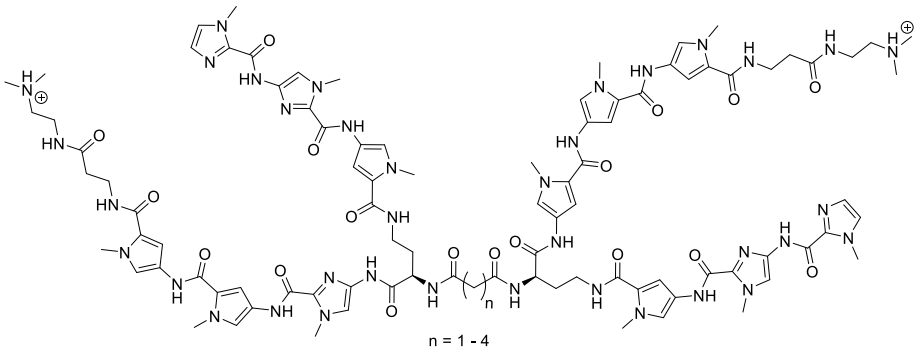


Figure 1.44 : Exemples de ligands polyamides testés par l'équipe de Dervan pour reconnaître une séquence d'une dizaine de base.[76]

Cette approche « relation structure – activité » peut également être utilisée pour optimiser certaines caractéristiques des ligands telles que leur capacité de transfection. Par

exemple, les structures des composés isophtalamides et dipicolinamides qui ont démontré pour la première fois la capacité de transfection des petites molécules, ont été optimisés avec cette méthode.[42]

B. APPROCHE COMBINATOIRE DYNAMIQUE

Conceptualisée dans les années 1990, la Chimie Combinatoire Dynamique s'est révélée être une méthode de choix pour étudier la reconnaissance de biomolécules.[77–87] Reposant sur la chimie supramoléculaire, cette approche a été particulièrement fructueuse pour identifier de nouveaux ligands de l'ADN.[88]

1. Principe de la Chimie Combinatoire Dynamique

La Chimie Combinatoire Dynamique repose sur la formation de liaisons réversibles entre différents fragments pour former les différents constituants d'une bibliothèque sous contrôle thermodynamique. Déterminée par les équilibres thermodynamiques, la composition d'une bibliothèque combinatoire dynamique est sensible aux variations internes ou externes du milieu dans lequel elle se trouve. Cette capacité des bibliothèques combinatoires dynamiques à s'adapter peut être utilisée pour sélectionner les espèces les plus stables parmi un mélange donné.

Ainsi, l'ajout d'une cible capable d'interagir de façon supramoléculaire avec les différents constituants peut entraîner un changement constitutionnel au sein de la bibliothèque (Figure 1.45). En accord avec le principe de Le Châtelier, les équilibres vont alors être déplacés vers la formation du ou des composés les plus stables thermodynamiquement en présence de cible. Ces composés ne sont pas forcément les plus stables en absence de cible et vont donc généralement être amplifiés (leur concentration dans la bibliothèque augmente par rapport à celle des autres constituants). En comparant les constitutions de la bibliothèque avant et après ajout de la cible, il est donc possible de déterminer le ou les constituants qui ont la meilleure reconnaissance pour la cible. Contrairement à des approches plus conventionnelles, la dimension dynamique de la méthode induit la sélection du ou des meilleurs ligands par la cible elle-même.

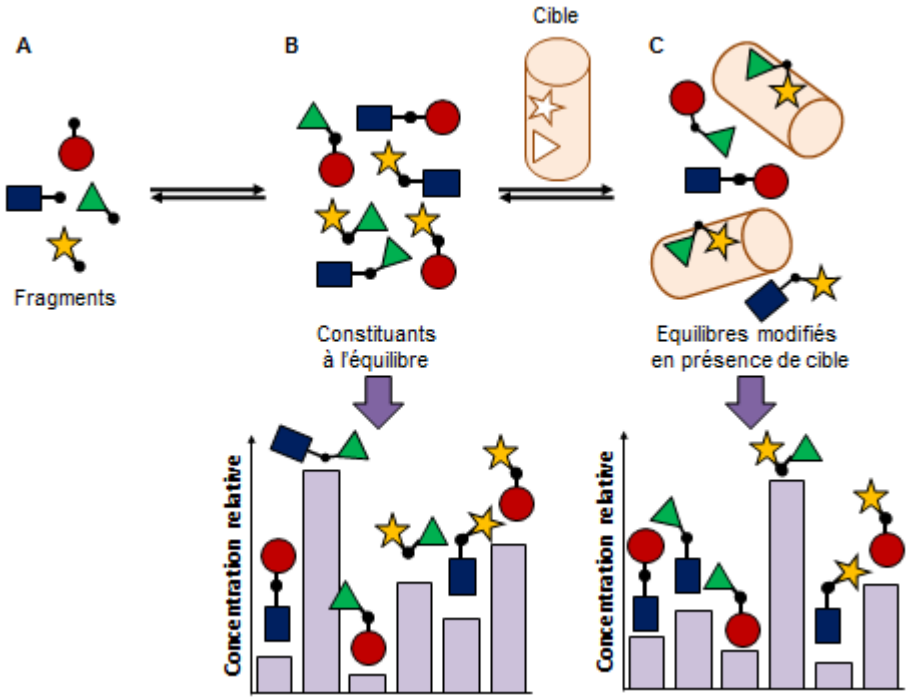


Figure 1.45 : Principe de la Chimie Combinatoire Dynamique. Les fragments fonctionnalisés (A) sont mélangés pour générer par liaisons réversibles les constituants d'une bibliothèque combinatoire dynamique (B) ; l'association sélective du meilleur ligand avec la cible ajoutée entraîne l'amplification de ce constituant dans la bibliothèque (C), mise en évidence par l'augmentation de la concentration relative du composé sélectionné.

En plus de permettre l'identification de nouveaux ligands des biomolécules, la Chimie Combinatoire Dynamique est également appliquée pour concevoir des récepteurs (ou ligands) synthétiques[89] ou développer de nouveaux catalyseurs.[90] Elle peut également être utilisée pour concevoir de nouveaux polymères sensibles à certains stimuli, internes ou externes, comme un processus de reconnaissance moléculaire.[91,92] Des applications en détection analytique sont aussi possibles.[93]

2. Avantages, inconvénients et limites

La Chimie Combinatoire Dynamique diffère des approches conventionnelles par son caractère dynamique, présent dans la réversibilité des liaisons impliquées dans la génération des constituants des bibliothèques. C'est ce qui fait à la fois sa force et sa faiblesse.

Alors qu'une approche combinatoire usuelle « statique » nécessitent de synthétiser et purifier toute une bibliothèque de composés avant de pouvoir tester individuellement leur reconnaissance pour la cible, la Chimie Combinatoire Dynamique évite ces étapes préliminaires et peut permettre un gain de temps considérable. La reconnaissance de la cible

par un constituant au sein d'une bibliothèque combinatoire dynamique entraîne en effet l'amplification de celui-ci et les composés sélectionnés peuvent éventuellement être directement isolés des bibliothèques.[94]

Au contraire de la chimie combinatoire classique, des constituants d'architecture complexe, tels que des macrocycles ou des cages, peuvent être facilement générés.[94] En effet, alors que leur formation n'est généralement pas cinétiquement favorisée, ces constituants complexes pourront plus aisément être sélectionnés par reconnaissance moléculaire et amplifiés au sein des bibliothèques combinatoires dynamiques. Les architectures obtenues dépendront de la conception des différents fragments utilisés (nombre de liaisons réversibles, géométrie) ainsi que de la cible utilisée.

En opposition à la conception *in silico* de nouveaux ligands pour une molécule donnée, très utilisée notamment avec les protéines, il n'est pas nécessaire de connaître la structure précise de la cible considérée pour utiliser la Chimie Combinatoire Dynamique. Il reste cependant nécessaire de concevoir les fragments de base des bibliothèques de façon à ce qu'ils puissent interagir avec la cible, ou aucune amplification ne pourra être observée.

La Chimie Combinatoire Dynamique présente également des inconvénients par rapport à d'autres méthodes et des limites quant à son utilisation.

Un des premiers points de difficultés de la méthode est l'analyse des bibliothèques. Les principales méthodes utilisées sont l'HPLC et la LC-MS, mais la spectrométrie de masse ou la RMN ont également prouvé leur potentiel.[95] Si l'amélioration de techniques d'analyses comme l'HPLC a conduit à l'émergence de la Chimie Combinatoire Dynamique, il peut en effet demeurer difficile d'analyser et quantifier correctement les constituants présents dans les bibliothèques formées, notamment lorsqu'un grand nombre de composés est présent dans le mélange. Pour cette raison, la majorité des bibliothèques combinatoires dynamiques comportent un nombre limité de constituants et se cantonnent à apporter une preuve de concept. Toutefois, un exemple de bibliothèque contenant plus de 9000 constituants a démontré que l'analyse par LC-MS d'une bibliothèque de cette taille était possible.[96] Pour éviter les problèmes d'analyses avec les bibliothèques de grande taille, il est également possible d'utiliser les séparations de phases pour réduire le nombre de constituants des bibliothèques à analyser. En fonction de la méthode utilisée, il est en effet possible d'analyser les constituants sélectionnés par la cible ou, à l'inverse, ceux qui ne sont pas sélectionnés. Pour parvenir à une telle séparation, la préparation de bibliothèques biphasiques[97,98] et

l'immobilisation de cibles[99–101] ou de fragments[102] sur support solide ont déjà fait leurs preuves.

Les bibliothèques combinatoires dynamiques peuvent également être biaisées en fonction des conditions dans lesquelles elles sont générées. Le constituant ayant la plus forte interaction avec la cible peut ne pas être amplifié si cette interaction n'est pas plus importante que celles de tous les autres constituants présents (qui ne sont que des ligands faibles de la cible) : idéalement, les constituants devraient être isoénergétiques dans la bibliothèque. Les amplifications peuvent aussi être trop faibles pour être détectées lorsque le constituant amplifié par la cible est déjà largement présent dans la bibliothèque en absence de cible. Ces biais peuvent en général être éliminés en utilisant des conditions plus adéquates pour générer les bibliothèques, notamment par correction de la réactivité des différents fragments en modulant les nombres d'équivalents introduits.

De plus, la constitution des bibliothèques peut aussi être biaisée si un constituant n'est pas suffisamment soluble dans le solvant utilisé, ce qui peut entraîner un piégeage thermodynamique ou cinétique du constituant en question et un déplacement de tous les équilibres : la bibliothèque n'est alors plus sous contrôle thermodynamique.

Enfin, l'utilisation de liaisons réversibles pour générer les constituants des bibliothèques peut induire une instabilité de ceux-ci, vis-à-vis de la méthode d'analyse des bibliothèques ou lors de leur isolement par exemple. Il faut pouvoir figer les échanges au sein des bibliothèques si besoin est. D'autre part, la liaison réversible utilisée dans les bibliothèques ne doit pas non plus interférer avec la cible, ce qui peut limiter les choix envisageables en fonction de la cible utilisée.

3. Les différentes liaisons réversibles employées

a. Critères de choix et exemples des liaisons réversibles employées dans les bibliothèques

Les liaisons réversibles employées pour générer les bibliothèques doivent répondre à un certain nombre de critères. En premier lieu, les réactions doivent être réversibles sur une échelle de temps raisonnable et compatibles avec les conditions expérimentales du processus de sélection, incluant les fonctions portées par les fragments et la cible, le solvant et le pH. Les conditions de la réaction réversible doivent également être douces (température, pression,

concentration) pour ne pas interférer avec le processus de reconnaissance de la cible par interactions non covalentes. Comme mentionné dans le point précédent, les constituants obtenus doivent être solubles pour ne pas constituer un piège thermodynamique ou cinétique : si la solubilité des fragments utilisés constitue un premier impératif, la réaction réversible utilisée doit également garantir la solubilité de tous les constituants formés dans la bibliothèque. De même, il doit également être possible de figer les équilibres d'échange pour permettre l'analyse et / ou l'isolement des constituants.

Plusieurs types de liaisons réversibles sont utilisables, qu'elles soient covalentes ou non covalentes. Bien que leurs équilibres soient atteints plus lentement qu'avec des liaisons non covalentes (liaisons hydrogène ou de coordination), ce sont les liaisons covalentes qui sont les plus utilisées. En effet, les constituants obtenus sont plus robustes et il est plus facile de figer les équilibres d'échange.

Les différentes liaisons réversibles utilisées en Chimie Combinatoire Dynamique sont variées (Figure 1.46) et leurs propriétés respectives, qui ne seront pas détaillées pour chacune ici, permettent de générer des bibliothèques dans des conditions très différentes.[80,103]

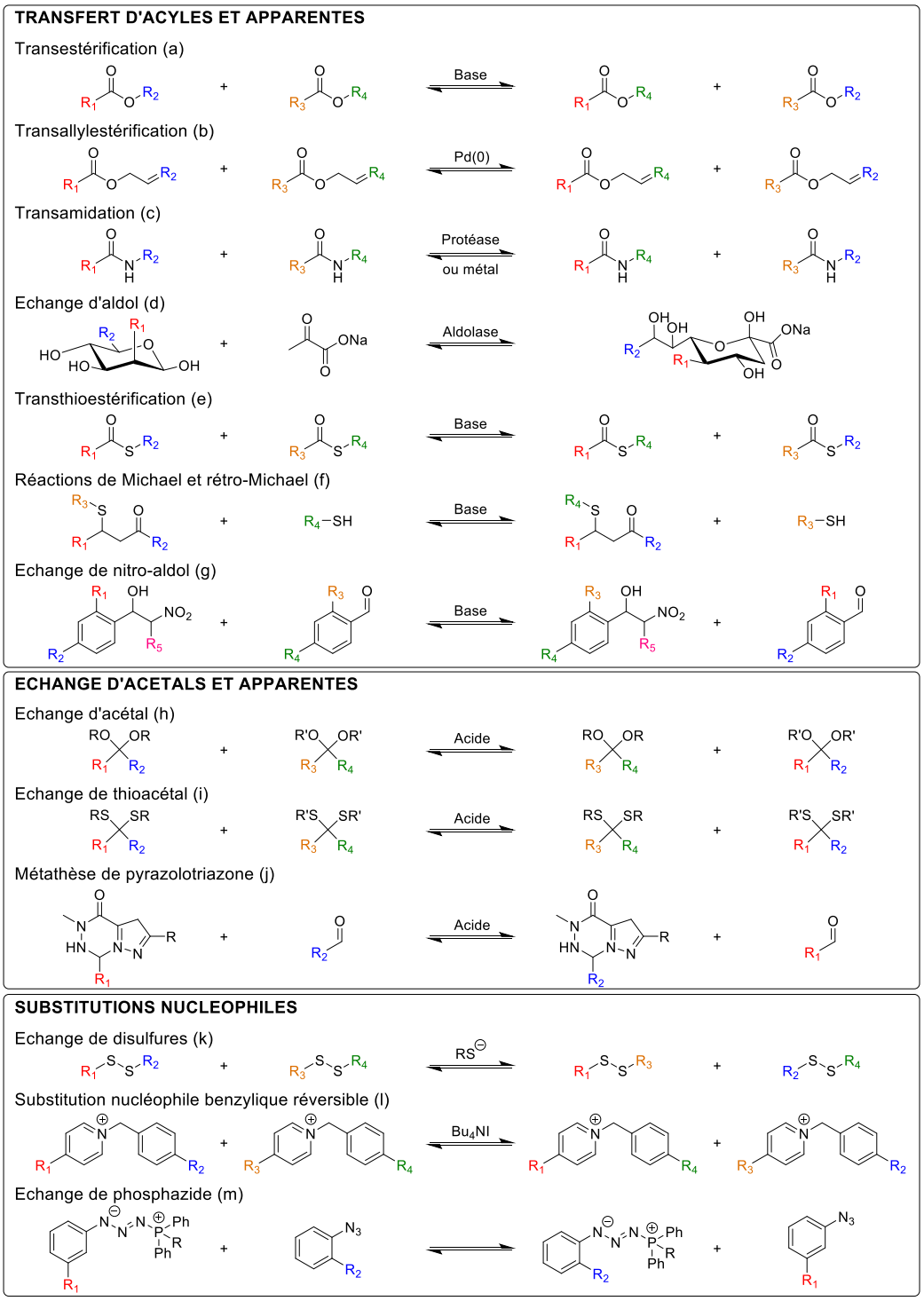


Figure 1.46 : Exemples de liaisons réversibles utilisées en Chimie Combinatoire Dynamique. Adapté de [103].

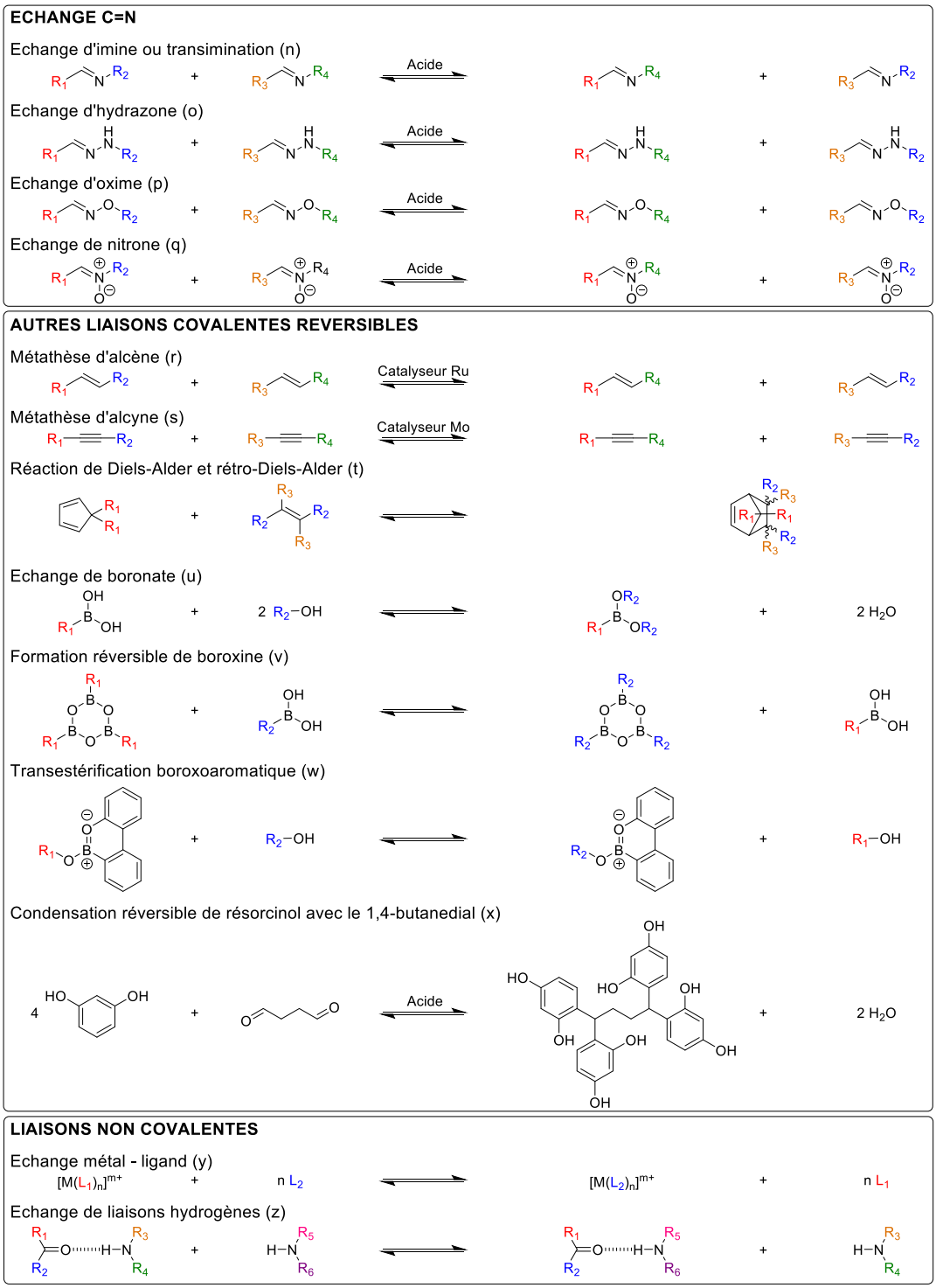


Figure 1.46 (suite) : Exemples de liaisons réversibles utilisées en Chimie Combinatoire Dynamique. Adapté de [103].

Parmi ces différentes liaisons réversibles, seul un nombre limité d'entre elles sont utilisables pour générer des bibliothèques combinatoires dynamiques en présence de biomolécules, pour des raisons de compatibilités avec la cible et des conditions dans laquelle cette dernière est stable. Généralement, l'utilisation de biomolécules se fait en milieu aqueux à un pH auquel la biomolécule est connue pour être stable (souvent compris entre 5 et 8).

Les liaisons principalement utilisées sont celles basées sur l'échange d'imines (Figure 1.46 n) et de leurs dérivées hydrazones (Figure 1.46 o), l'échange de disulfures (Figure 1.46) et l'échange métal – ligands (Figure 1.46 y).[104] Ces quatre liaisons vont être présentées succinctement par la suite.

L'utilisation d'autres liaisons réversibles comme la transthiolestérification (Figure 1.46 e), les réactions de Michael et rétro-Michael (Figure 1.46 f), la métathèse d'alcène (Figure 1.46 r) ou les échanges d'aldols (Figure 1.46 d) a également été démontrée dans des bibliothèques combinatoires dynamiques comportant des biomolécules, mais est beaucoup moins répandue que les précédentes.[104] Pour cette raison, leurs caractéristiques particulières ne seront pas détaillées ici.

b. Liaison imine

Les conditions nécessaires à la transimination sont des conditions douces compatibles avec l'utilisation de biomolécules. En plus de pouvoir générer les bibliothèques en milieu organique, il est en effet également possible de les préparer en tampons aqueux au pH compris entre 5 et 8,5 à température ambiante. La transimination est orthogonale à d'autres liaisons réversibles, dont la liaison de coordination métallique qui peut ainsi être également utilisée dans une même bibliothèque combinatoire dynamique pour diversifier les structures obtenues.[105]

Cependant, les liaisons imines sont sensibles à l'hydrolyse : les imines ne sont ainsi formées qu'en faible proportion dans les bibliothèques préparées en présence d'eau (ce qui est typiquement le cas en présence de biomolécules) et sont généralement instables lors des analyses chromatographiques lorsque les éluants contiennent de l'eau.[80,106] Cette sensibilité vis-à-vis de l'hydrolyse constitue le principal inconvénient de l'utilisation des liaisons imines comme liaison réversible au sein des bibliothèques combinatoires dynamiques. Cependant, il est possible de figer les équilibres en réduisant les imines formées par un réducteur spécifique comme le cyanoborohydrure de sodium (Figure 1.47).[80]

Figure 1.47 : Formation et figeage des liaisons imines par amination réductrice avec le cyanoborohydrure de sodium ($NaBH_3CN$).

Cette amination réductrice, non réversible, piège cinétiquement les imines.[80,107] Elle a pour avantage de former des amines stables en milieu aqueux (donc stables vis-à-vis des méthodes HPLC), et d'augmenter la proportion de produits d'amination grâce au déplacement de l'équilibre vers ces produits. Cependant, elle présente également des limites.[80,89] Il est ainsi nécessaire de s'assurer que les constituants de la bibliothèque ont la même réactivité vis-à-vis de l'amination réductrice pour que la distribution des amines soit représentative de la distribution des imines. Il faut également vérifier que la réduction directe des aldéhydes en alcools n'influe pas sur le processus de sélection. D'autre part, la réduction entraînant des changements en termes de géométrie et de distribution électronique, l'amine obtenue peut ne pas présenter les mêmes capacités de reconnaissance de la cible que l'imine qui est amplifiée dans la bibliothèque. Bien qu'une corrélation entre amine isolée et imine amplifiée en termes de reconnaissance de la cible donnée ait été mise en évidence dans plusieurs cas, l'inverse – l'absence de corrélation – a également été observé[108] et la possibilité qu'il n'y ait pas toujours de corrélation doit être gardée à l'esprit.

c. Liaison (acyl)hydrazone

Les liaisons hydrazones sont des liaisons réversibles apparentées aux imines qui sont également utilisées dans les bibliothèques contenant des biomolécules. Elles présentent l'avantage d'être plus stables par rapport à l'hydrolyse que les imines et il est généralement inutile de les réduire pour analyser les bibliothèques ou isoler les produits amplifiés.[80] Lorsque les hydrazines de départ ne portent pas de groupements électroattracteurs comme un acyle en α, les hydrazones sont même généralement trop stables pour permettre des échanges au sein de bibliothèques combinatoires dynamiques (Figure 1.48). Cette augmentation de la stabilité est due à l'effet α dont l'importance dépend des substituants de l'azote de la liaison imine : par mésomérie, ces substituants peuvent défavoriser l'attaque de l'eau sur le carbone de la liaison imine et ainsi limiter l'hydrolyse. Les hétéroatomes en α de l'azote sp^2 réduisent également son caractère basique, d'où sa moindre sensibilité à l'hydrolyse.

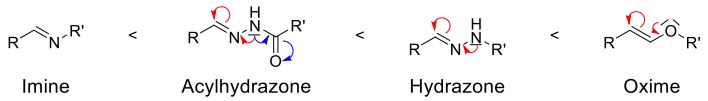


Figure 1.48 : Ordre de stabilité entre imine et dérivés apparentés vis-à-vis de l'hydrolyse. Les flèches rouges et bleues représentent les effets mésomères responsables des différences de stabilités observées.

La plupart des bibliothèques combinatoires dynamiques utilisant la liaison hydrazone est en fait basée sur des composés acylhydrazones. Généralement, hormis lorsqu'elles sont substituées par des groupements fortement électroattracteurs,[109] la formation et l'hydrolyse des acylhydrazones sont très lentes à pH neutre ou légèrement basique tandis qu'elles sont très rapides à pH acide : le contrôle du pH permet donc de contrôler les échanges.[86] *A priori*, les conditions nécessaires à ces échanges d'acylhydrazones ne conviennent pas toujours aux biomolécules, le pH devant généralement être inférieur à 4 pour que la réaction d'échange ait lieu rapidement. Cependant, l'ajout d'un catalyseur comme l'aniline permet à l'échange des acylhydrazones de s'effectuer à un pH proche de 6 (Figure 1.49)[110] et est compatible avec des bibliothèques combinatoires dynamiques en présence de biomolécules.[111,112]

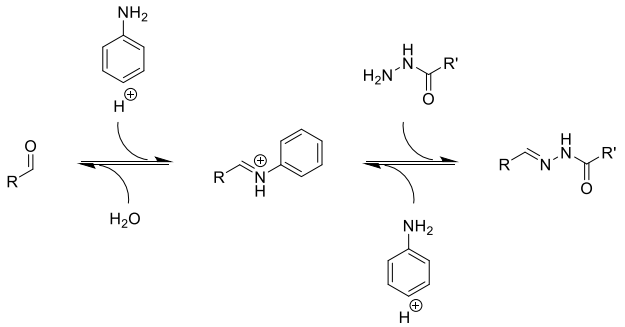


Figure 1.49 : Catalyse nucléophile de la formation d'acylhydrazone par l'aniline.

D'autre part, la stabilité des hydrazones à pH légèrement basique permet d'éviter la réduction des liaisons pour figer les équilibres. Les hydrazones peuvent ainsi être directement isolées sans que le figeage des équilibres ne porte atteinte à leur structure ou conformation comme ce peut être le cas avec la réduction des imines.

Par contre, à structure équivalente, les hydrazones sont généralement moins solubles que les imines correspondantes, que ce soit dans l'eau ou les solvants organiques, ce qui peut constituer un inconvénient majeur à leur utilisation dans les bibliothèques combinatoires dynamiques.[86] Cette limitation de solubilité explique pourquoi les exemples de bibliothèques combinatoires dynamiques sont moins nombreux avec les hydrazones qu'avec les imines ou les disulfures.

d. Liaison disulfure

L'échange de disulfures est une des réactions réversibles les plus utilisées en Chimie Combinatoire Dynamique.[80,113] Bien que des bibliothèques puissent être générées en solvant organique, les disulfures sont généralement formés par oxidation de thiols en milieu aqueux

exposé à l'oxygène de l'air et leur échange se fait dans des conditions douces. Cet échange suit le mécanisme d'une S_N2 dans lequel un ion thiolate joue le rôle de nucléophile (Figure 1.50) : il est donc nécessaire d'avoir une quantité au moins catalytique de thiol dans le mélange et de contrôler le pH entre 7 et 9 de façon à ce que ce thiol présent soit sous forme de thiolate (le thiol n'est pas suffisamment nucléophile pour attaquer les disulfures). De ces deux conditions découlent les moyens de figer l'échange de disulfures : il peut être figé par acidification du milieu, ce qui va entraîner la protonation du thiolate « catalyseur », ou bien par consommation de tous les thiolates présents dans le milieu par oxidation en disulfures.

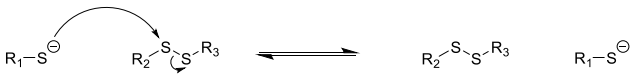


Figure 1.50 : Echange de disulfure par substitution nucléophile.

L'échange des disulfures est une réaction aux conditions idéales pour une utilisation en présence de biomolécules. De plus, contrairement aux imines, les disulfures sont suffisamment stables pour être analysés et isolés directement des bibliothèques.

Toutefois, les temps d'équilibration des bibliothèques sont généralement plus importants que ceux des bibliothèques d'imines ou d'hydrazones, ce qui peut constituer un inconvénient à l'utilisation de cette réaction.[86] D'autre part, il n'est pas forcément aisé d'éviter la formation de composés homotopiques, les thiols pouvant réagir entre eux et générer des composés dimériques.

e. Liaison de coordination métal - ligand

Les liaisons de coordination métal – ligand peuvent également être utilisées en présence de biomolécules. La relative prédictibilité de la direction de ces liaisons a pour avantage de conduire à la formation d'auto-assemblages dont les structures sont bien définies. Toutefois, le processus de sélection dans les bibliothèques combinatoires dynamiques n'est possible qu'à condition que les échanges de ligands se fassent suffisamment rapidement.[80] La vitesse de ces échanges, tout comme la formation des complexes au sein des bibliothèques, dépendent de plusieurs paramètres comme la nature du métal ou encore la structure et la charge des ligands.

Contrairement aux autres liaisons détaillées plus haut, les liaisons de coordination métal – ligand sont de nature non covalente. Elles se trouvent donc être généralement particulièrement labiles dans les bibliothèques combinatoires dynamiques – le métal utilisé devant permettre les échanges réversibles entre constituants -, ce qui rend difficile l'isolement

et la manipulation des constituants des bibliothèques.[80] La fixation des équilibres est donc généralement nécessaire et peut être effectuée en changeant le degré d'oxydation du métal ou en piégeant les constituants de manière covalente ou encore, de manière plus ponctuelle, en cristallisant sélectivement des constituants spécifiques dans la bibliothèque. Comme pour les imines dont la réduction en amines peut avoir des répercussions sur la reconnaissance pour la cible, la fixation des équilibres peut ici aussi faire varier la structure des constituants isolés de ceux amplifiés dans la bibliothèque.

4. Application à la reconnaissance d'ADN

Très rapidement après son émergence, la Chimie Combinatoire Dynamique a été appliquée à la reconnaissance d'acides nucléiques, sous forme d'ADN ou d'ARN.[88]

La preuve de concept de la compatibilité de cette méthode avec une cible ADN a été apportée par les travaux de Miller et son équipe en 1997.[99] Dans cette première approche, une bibliothèque combinatoire dynamique de complexes métal – salicylaldimines a été générée et mise en présence d'une cible ADN double brins immobilisée sur cellulose (Figure 1.51). La cible sur support solide est retirée des bibliothèques avec les composés qui y sont retenus, ce qui entraîne une diminution de ces composés au sein de la bibliothèque analysée, au lieu de l'habituelle amplification recherchée.

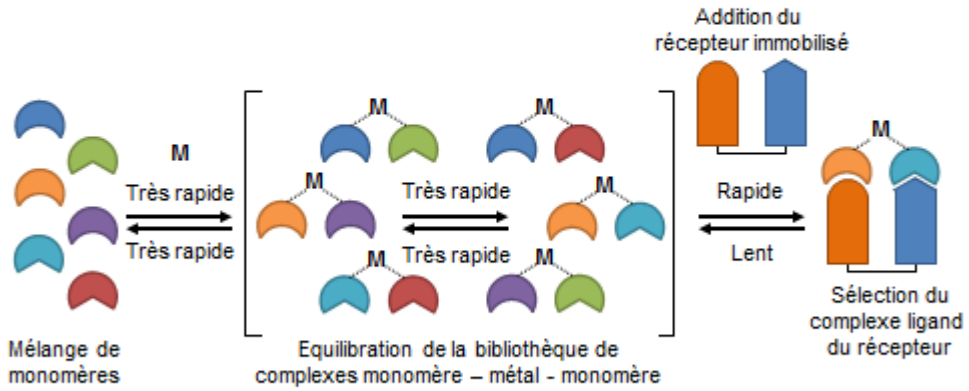


Figure 1.51 : Illustration schématique de la stratégie combinatoire dynamique utilisée par Miller et ses collaborateurs pour sélectionner des complexes métalliques ayant une affinité pour une cible immobilisée. La constitution de la bibliothèque s'équilibre rapidement par complexation avec un métal, puis l'ajout de la cible choisie immobilisée sur support solide conduit à la sélection de constituants ayant une affinité pour celle-ci.[99]

Dans ce premier exemple, la cible choisie est un duplexe de poly(dT·dA) immobilisé sur cellulose. Les bibliothèques ont été générées à partir de six salicylaldimines différentes (Figure 1.52), en présence ou non de chlorure de zinc, et incubées en présence d'ADN sur

cellulose. Après élution des bibliothèques sur la cible sur cellulose, les constituants encore présents dans les mélanges ont été hydrolysés, puis les amines correspondantes ont été « dérivées » avec un excès de chlorure de 2-naphthoyle (pour permettre leur détection UV) et séparées par HPLC. L'analyse des différentes bibliothèques a mis en évidence que deux composés parmi les six salicylaldimines, **Sal2** et **Sal4**, étaient retenus sur l'ADN sur cellulose en présence de zinc bivalent. L'analyse des mêmes bibliothèques éluées sur de la cellulose seule a par la suite mis en évidence que les complexes de zinc à base de salicylaldimines **Sal4** étaient sélectivement retenus par la cellulose : seuls les complexes de salicylaldimine **Sal2** à base de N-méthyl-2-aminoéthyl pyrrolidine sont donc sélectivement retenus sur l'ADN cible. Après titrations UV, il a finalement été déterminé que le complexe homodimérique de cette salicylaldimine **Sal4** avec le zinc (Figure 1.53) était le meilleur ligand parmi les combinaisons possibles et a une constante de dissociation de 1,1 µM avec l'ADN cible.

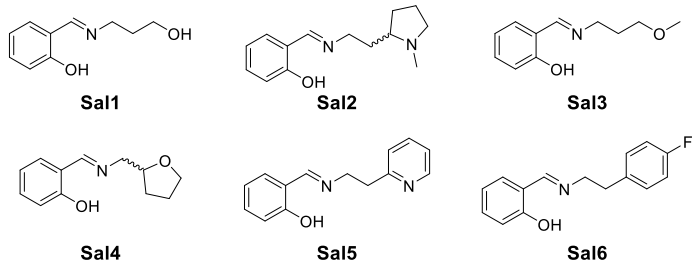


Figure 1.52 : Salicylaldimines synthétisées par l'équipe de Miller et dont l'affinité pour un poly(dT·dA) sur cellulose a été testée en absence et en présence de chlorure de zinc.[99]

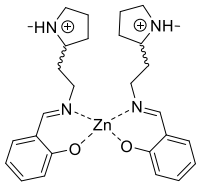


Figure 1.53 : Complexe homodimérique de zinc avec la salicylaldimine **Sal2** identifié par Chimie Combinatoire Dynamique comme ayant la meilleure affinité pour la cible poly(dT·dA) parmi les autres constituants des bibliothèques.

Suite à l'observation que la complexation avec le zinc ne rend pas négligeable l'hydrolyse des salicylaldimines utilisées dans les premières bibliothèques et qui avaient été synthétisées séparément par condensation du salicylaldéhyde avec les amines correspondantes, d'autres bibliothèques combinatoires dynamiques ont été générées directement à partir des réactifs salicylaldéhyde et amines pour former les salicylaldimines *in situ*.[105] En présence de la même cible poly(dT·dA) sur cellulose, la N-méthyl-2-aminoéthyl pyrrolidine est de nouveau sélectionnée comme étant à la base des ligands avec la plus grande affinité pour la cible. Ce nouvel exemple de bibliothèque combinatoire dynamique démontre

qu'une équilibration « multi-étapes » est utilisable avec une cible immobilisée sur support solide.

Ces deux exemples ont apportés la preuve de concept de l'utilisation d'une cible ADN en Chimie Combinatoire Dynamique. Cependant, la réversibilité des liaisons imines et de coordination métallique n'ont pas rendu l'identification particulièrement facile dans ces deux cas. D'autres bibliothèques combinatoires dynamiques dans lesquelles il était plus facile d'identifier les constituants amplifiés ont par la suite été générées avec des cibles ADN.

L'équipe de Balasubramanian a ainsi utilisée la liaison disulfure pour générer des bibliothèques combinatoires dynamiques testées en présence de différentes cibles à base d'acides nucléiques. Une bibliothèque de disulfures polyamides dérivés de la distamicine et proches des composés de l'équipe de Dervan (Figure 1.54) a par exemple été générée en présence d'un oligonucléotide 11-mère double brin riche en séquence A-T ou en présence d'un G-quadruplexe.[114] Les deux cibles ADN sont biotinylées pour être retirées des bibliothèques par immobilisation sur des billes magnétiques portant de la streptavidine.

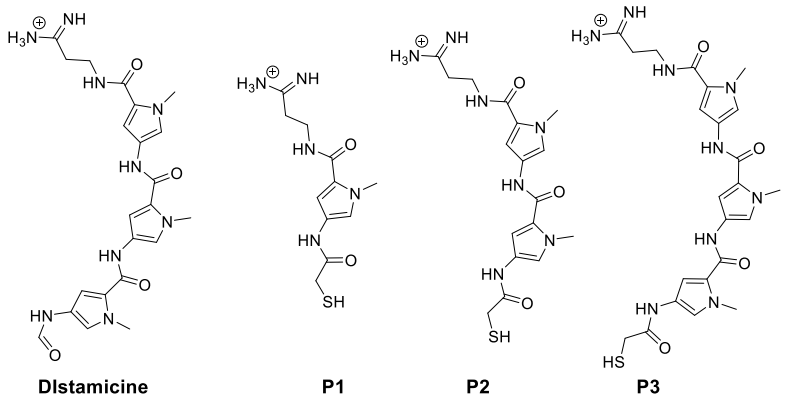


Figure 1.54 : Thiols polyamides utilisés par l'équipe de Balasubramanian pour générer des bibliothèques combinatoires dynamiques en présence de cible d'ADN double brin ou G-quadruplexe.[114]

Avant analyses, les échanges de disulfures sont figés par réduction du pH de 7,4 à 2. Les cibles ADN sont ensuite retirées des bibliothèques avec les constituants qui interagissent avec elles, les oligonucléotides sont dénaturés par chauffage pour séparer cible et constituants ; puis les mélanges de constituants récupérés sont analysés par HPLC et comparés avec la constitution d'une bibliothèque générée en absence de cible (Figure 1.55). Des différences flagrantes sont visibles en fonction de la cible présente (Figure 1.56) : alors que la cible G-quadruplexe n'induit pas ou peu d'amplification par rapport à la bibliothèque témoin, des constituants sont clairement sélectionnés et amplifiés en présence de la cible double brins.

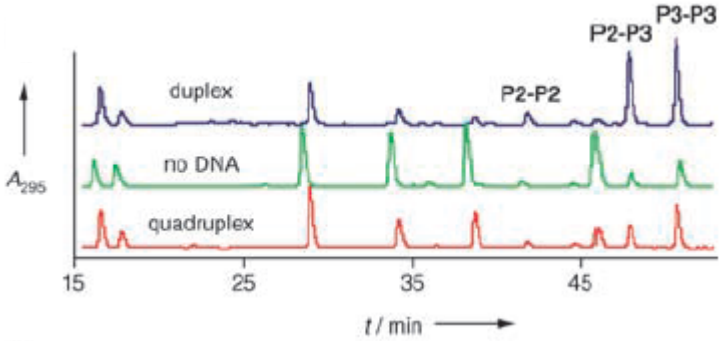


Figure 1.55 : Chromatogrammes HPLC des bibliothèques en absence de cible (vert), en présence de cible G-quadruplexe (rouge) ou de cible double brins (bleu). Bibliothèques après 18 heures d'équilibration générées à partir des thiols **P1**, **P2** et **P3** (100 µM de chaque) dans un tampon Tris (300mM Tris-HCl pH 7,4, 150 mM KCl) contenant de la glutathione réduite (30 mM) et oxidée (7,5 mM), en absence de cible ou en présence de G-quadruplexe (100 µM, rouge) ou de double brins (100 µM, bleu).[114]

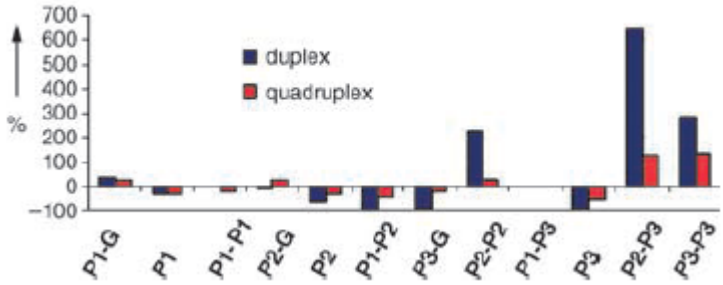


Figure 1.56 : Histogramme des amplifications des constituants observées dans les bibliothèques en présence de cible (double brins en bleu, G-quadruplexes en rouge) par rapport à la bibliothèque témoin en absence de cible.[114]

La sélectivité des constituants des bibliothèques pour l'ADN double brin n'est pas surprenante étant donné qu'une sélectivité similaire avait déjà été observée pour des composés apparentés par d'autres méthodes. Toutefois, cet exemple démontre clairement pour la première fois que des ligands spécifiques d'une structure d'ADN peuvent être identifiés par Chimie Combinatoire Dynamique.

En plus de l'identification de nouveaux ligands de l'ADN, la Chimie Combinatoire Dynamique peut également être utilisée pour introduire des modifications chimiques sur la cible ADN en vue d'augmenter sa stabilité.

Cette propriété a été démontrée pour la première fois par Rayner et ses collaborateurs en 2004.[115] Dans ce premier exemple, une 2'-amino-2'-déoxyuridine est introduite en position terminale 3' d'un hexaoligonucléotide complémentaire de lui-même, le $^{5'}$-d(ACGCGU*)-$^{3'}$: un duplexe d'ADN comportant une fonction amine de part et d'autre du double brin est ainsi

obtenu après hybridation. Cette cible est testée en présence de trois aldéhydes différents, susceptibles de réagir avec les fonctions amines terminales de la cible duplexe (Figure 1.57). Pour analyser les bibliothèques obtenues, les liaisons imines formées ont été réduites par le cyanoborohydrure de sodium ajouté aux bibliothèques.

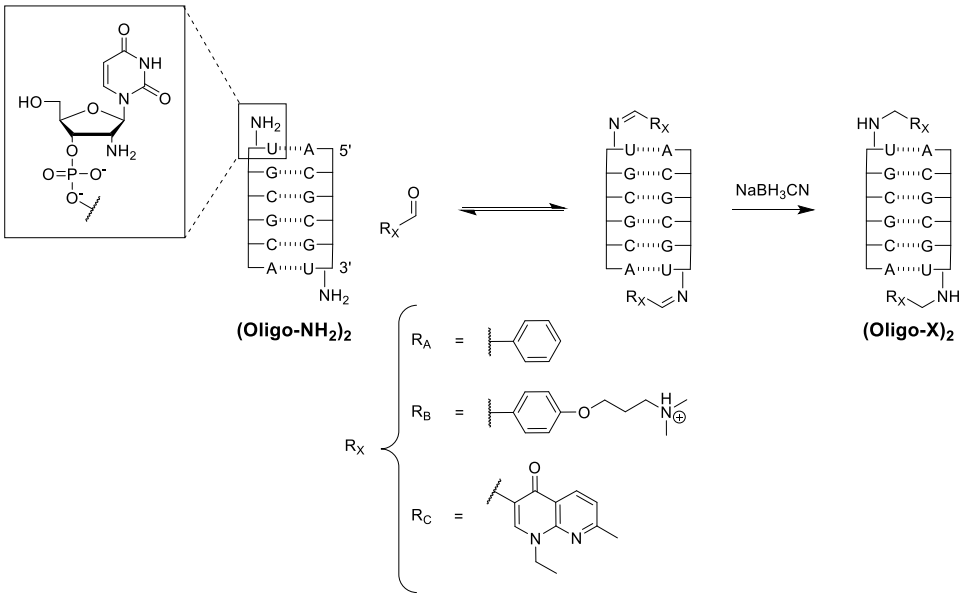


Figure 1.57 : Modifications chimiques d'un oligonucléotide modifié apportées par Chimie Combinatoire Dynamique en vue de trouver le groupement apportant le plus de stabilité au duplexe formé.[115]

Les bibliothèques obtenues sont comparées à une bibliothèque générée en présence d'un oligonucléotide comportant également la base uracile modifiée U* mais non complémentaire de lui-même, donc incapable de s'hybrider et de former un duplexe (le $^{5'}$-d(TTTCGU*)-$^{3'}$). Les autres conditions sont identiques entre les bibliothèques, mais des différences dans leur constitution sont visibles : une amplification du duplexe **(Oligo-C)$_2$** est observée par rapport au simple brin correspondant dans la bibliothèque témoin tandis que la proportion des deux autres duplexes **(Oligo-A)$_2$** et **(Oligo-B)$_2$** a diminué.

La comparaison des températures de demi-fusion (T_m) des duplexes modifiés par Chimie Combinatoire Dynamique avec celle du duplexe réactif **(Oligo-NH$_2$)$_2$** met clairement en évidence le lien entre amplification dans la bibliothèque combinatoire dynamique et stabilisation du duplexe. En effet, le duplexe **(Oligo-C)$_2$** amplifié dans la bibliothèque présente une température de demi-fusion supérieure de plus de 10°C par rapport à celle du duplexe réactif **(Oligo-NH$_2$)$_2$**.

Depuis ce premier exemple de stabilisation d'un double brin d'ADN par introduction de modification chimique par Chimie Combinatoire Dynamique, cette approche a depuis été

élargie à diverses cibles comme des ARN[116] ou des oligonucléotides formant des triplexes,[117] toujours dans le but de trouver des groupements stabilisant la structure de la cible.

5. Génération de matériaux dynamiques

Comme mentionné plus haut, la Chimie Combinatoire Dynamique peut être une méthode de formation des polymères supramoléculaires ou covalents dynamiques. Alors que les polymères supramoléculaires sont assemblés par des liaisons non covalentes, les monomères des polymères covalents dynamiques s'auto-assemblent par de multiples liaisons covalentes réversibles qui leur permettent de s'adapter à leur milieu par réorganisation constitutionnelle (sous réserve d'être dans les conditions de réversibilité des liaisons covalentes).[118–123] Cette adaptation au milieu peut impliquer l'échange de monomères, modifiant la nature même du polymère et ses propriétés. Un bon exemple de l'effet des stimuli externes sur les polymères dynamiques peut être illustré par le comportement de polyimines à base de fluorène (Figure 1.58) dont les propriétés optiques changent suite à la réorganisation des monomères induite par l'addition de zinc^{II} au mélange.[124,125]

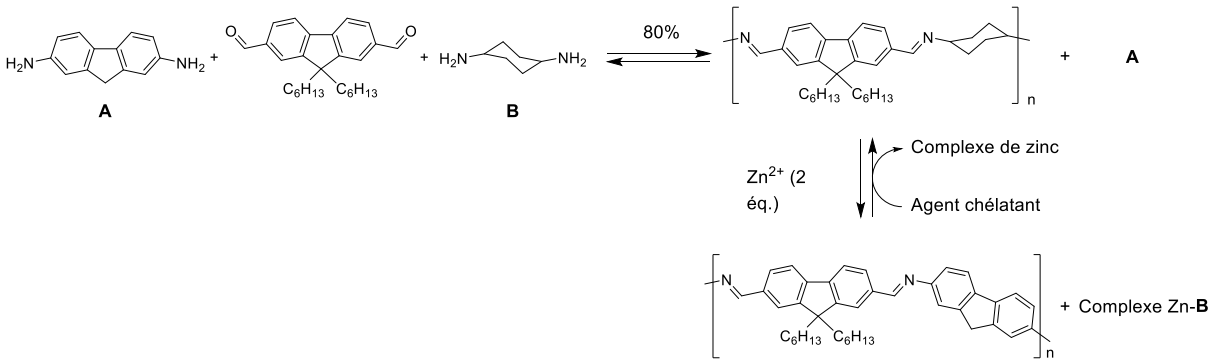


Figure 1.58 : Polyiminofluorènes dynamiques dont l'échange de monomères est induit par la concentration en zinc^{II}. Adapté de [121].

La réversibilité des liaisons impliquées pour lier les différents monomères peut conférer aux polymères des propriétés d'auto-réparation[126] et / ou de dégradation qui font défaut aux polymères non dynamiques. En fonction des différentes propriétés des polymères, des applications dans le domaine des matériaux peuvent être envisagées.

Appliqués à la reconnaissance de biomolécules, les polymères dynamiques peuvent aussi être utilisés pour concevoir des auto-assemblages multivalents dont la constitution et le degré de polymérisation dépendront de la matrice utilisée. L'ADN étant une excellente matrice pour organiser des auto-assemblages et permettre l'édification de nanostructures bien

définies (voir I.C.), une approche de Chimie Combinatoire Dynamique pourrait être utilisée pour concevoir des polymères dynamiques de longueur et de structure connues.

Ainsi, si une matrice ADN est capable d'interagir avec différents monomères, elle pourrait contrôler la longueur, voire l'organisation des polymères obtenus. La synthèse par amination réductrice de polyamines de longueur définie à l'aide d'une matrice dA_8 indique qu'un contrôle de la longueur du « polymère » est possible.[127,128] En effet, les polyamines sont synthétisées à partir de huit monomères ou de quatre dimères comportant respectivement une ou deux thymines comme site de reconnaissance avec la matrice ADN (Figure 1.59) : l'utilisation d'une matrice permet à la polymérisation d'avoir lieu et détermine la longueur de la polyamine obtenue. La reconnaissance par appariement des bases entre la matrice et les monomères est également importante, un défaut d'appariement ne permettant pas la polymérisation.[127]

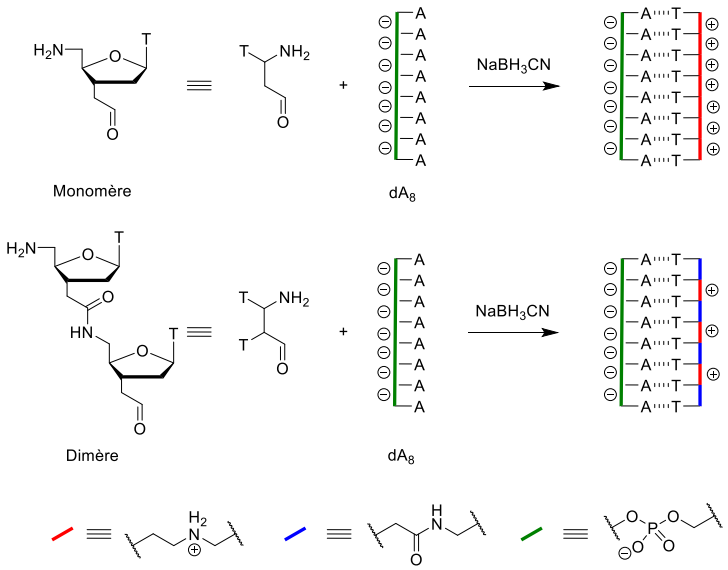


Figure 1.59 : Polymérisation matricée par l'ADN de polyamines de longueur déterminées.[127,128]

Des polymères covalents dynamiques polyacylhydrazones ont également démontré leur adaptation à divers polyanions (polyphosphates ou acide polyaspargique) ainsi que leur capacités d'interactions avec un oligonucléotide dA_{40} (Figure 1.60).[129] Il est intéressant de remarquer que dans le cas présent, les nucléobases présentes sur les polyacylhydrazones ne semblent pas intervenir dans les interactions entre polymère et ADN, les interactions étant principalement électrostatiques. Le polymère avec les bases thymines s'adapte différemment à la présence de poly(A) ou de poly(U) dans le milieu : un hydrogel stable est observé avec le poly(A) alors qu'un précipité apparaît en présence de poly(U).

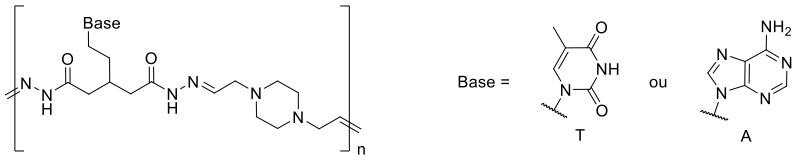


Figure 1.60 : Polymères dynamiques polyacylhydrazones adaptant leur taille à la présence de polyanions.

La reconnaissance spécifique d'une matrice ADN par des fragments est connue pour réaliser des ligations. Cependant, peu de ces réactions de ligation font appel à une liaison covalente réversible pour obtenir des systèmes dynamiques informatifs. L'équipe de Vasseur et Smietana a mis au point ce type de ligation en utilisant l'échange de boronate.[130] Une matrice oligonucléotide peut être utilisée pour rapprocher les deux oligonucléotides entre lesquels la ligation est souhaitée tandis que la réversibilité de la ligation permet au système de s'adapter aux conditions du milieu et de sélectionner les ligands optimaux pour la matrice ADN présente.

III. PROJET DE RECHERCHE

A. BUT ET APPLICATIONS

Le but premier du projet de recherche est de sonder les interactions petites molécules – ADN pour mieux comprendre les processus de reconnaissance de l'ADN. Contrairement aux travaux de Dervan qui utilise des dérivés de composés connus pour interagir dans le petit sillon de l'ADN, notre approche consisterait à utiliser la reconnaissance des phosphodiesters par nos petites molécules comme interaction primaire avec l'ADN pour ensuite sonder les interactions secondaires des différents substituants de nos petites molécules pour l'ADN. Les résultats obtenus pourraient par la suite être appliqués à la conception de nouveaux ligands de l'ADN.

En parallèle, des polymères covalents dynamiques seront générés pour étudier l'interaction de ces systèmes multivalents avec l'ADN ainsi que leur dégradation contrôlée en fonction du pH. Les monomères utilisés sont conçus pour interagir avec le squelette phosphodiester de l'ADN indépendamment de la séquence d'ADN considérée.

A moyen et long termes, il est possible d'envisager la conception de nouveaux vecteurs universels d'ADN, aussi bien à partir des petites molécules que des polymères

dynamiques. Actuellement, les vecteurs viraux sont les plus efficaces en termes de transport mais restent potentiellement toxiques, ce qui limite leur usage en thérapies géniques.[131]

B. STRATEGIE

Comme mentionné précédemment, notre projet de recherche utiliserait des petites molécules capables d'interagir en premier lieu avec les phosphodiesters de l'ADN. Pour cela, les petites molécules utilisées doivent être conçues autour d'un groupement capable de reconnaître les oxoanions. Ce groupement permettrait aux composés finaux de reconnaître les phosphodiesters de l'ADN plutôt que d'interagir avec les bases ou dans les sillons. La reconnaissance des phosphodiesters de l'ADN permettrait aux molécules synthétiques d'interagir avec n'importe quel type d'ADN, quelle que soit sa séquence de bases et sa structure (simples ou doubles brins pourraient être reconnus de la même manière). Des ligands (voir des vecteurs) universels de l'ADN pourraient ainsi être obtenus.

Deux approches, une approche rationnelle et une approche combinatoire dynamique, sont utilisées en parallèle pour sonder les interactions petites molécules – ADN. Ces deux approches ayant chacune leurs points forts et leurs limites, différentes informations pourront être obtenues avec l'une ou l'autre approche.

L'approche dynamique est prolongée par la conception de polymères dynamiques covalents capables d'interagir avec l'ADN et dont la réversibilité constituerait un caractère important pour la dégradation du polymère dans certaines conditions, ce qui permettrait alors de libérer l'ADN complexé.

C. CONCEPTION DES COMPOSES : CHOIX DU GUANIDINIUM COMME GROUPEMENT DE RECONNAISSANCE

Les composés utilisés dans le projet sont conçus autour du groupement de reconnaissance des phosphodiesters, auquel différentes substitutions peuvent ensuite être introduites de part et d'autre, en fonction des propriétés recherchées et des besoins du projet (Figure 1.61).

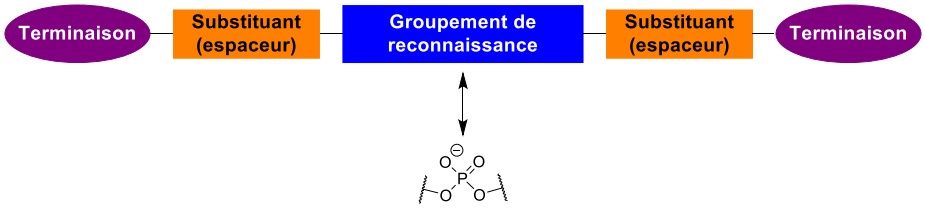


Figure 1.61 : Conception générale des composés synthétiques utilisés dans notre projet.

Devant le nombre très important de groupements connus pour interagir avec les molécules phosphorylées,[132] donc susceptibles d'interagir avec les phosphodiesters de l'ADN, un choix s'impose, qui doit prendre en compte les spécificités du projet.

Plusieurs contraintes sont en effet à prendre en compte, la première étant la compatibilité des groupements portés par les molécules synthétiques avec l'ADN et son « milieu d'utilisation ». Il s'agit en l'occurrence de concevoir des composés solubles et stables en milieux aqueux tout en demeurant capables de former des interactions supramoléculaires avec l'ADN (un ordre de grandeur de la force des interactions possibles est donné par la constante d'association du groupement de reconnaissance avec le phosphate).

La préparation des composés, en particulier le nombre d'étape de synthèse, les rendements, la mise en œuvre et la « pré-validation » de la synthèse (description du protocole et des analyses, reproductibilité), est un autre critère général d'importance, applicable aux groupements de reconnaissance des phosphodiesters comme aux substituants introduits sur celui-ci. Le groupement de reconnaissance doit également pouvoir être aisément fonctionnalisé par différents substituants, notamment pour pouvoir introduire les fonctions réactives impliquées dans les liaisons réversibles à la base des bibliothèques combinatoires dynamiques ou les substituants aromatiques utilisés pour l'approche rationnelle.

Les méthodes utilisées peuvent également induire des contraintes particulières à prendre en compte dès la conception des composés. L'utilisation de la Chimie Combinatoire Dynamique génère notamment des contraintes qui diffèrent selon la liaison réversible choisie pour préparer les bibliothèques. Pour l'étude des interactions petites molécules - ADN, la liaison imine a été choisie, malgré son instabilité en milieu aqueux, en raison de sa compatibilité avec l'ADN et du grand nombre d'amines commerciales. Quant au développement des polymères covalents dynamiques pour la reconnaissance d'ADN, la liaison hydrazone a été retenue en raison du contrôle de son hydrolyse en fonction du pH. Le choix de ces liaisons implique que le groupement de reconnaissance choisi soit compatible avec les fonctions imines et acylhydrazones formées ainsi qu'avec les fonctions aldéhydes et amines réactives avec lesquelles il ne doit pas réagir. Il doit aussi être compatible avec une

réduction par le cyanoborohydrure de sodium utilisé pour figer les équilibres des bibliothèques d'imines. L'analyse des bibliothèques étant effectuée par LC-MS, les composés doivent également être compatibles avec cette méthode d'analyse.

Plusieurs groupements reconnaissant les phosphates sont susceptibles de respecter les différents critères du cahier des charges du projet. Parmi eux, le groupement guanidinium est particulièrement approprié : il est en effet déjà connu pour sa capacité à interagir avec des oligonucléotides,[133] est soluble et stable dans l'eau sur une gamme de pH importante et est compatible avec l'utilisation de la Chimie Combinatoire Dynamique dans les conditions choisies.

RÉFÉRENCES BIBLIOGRAPHIQUES

(1) Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737–738.

(2) Ghosh, A.; Bansal, M. *Acta Cryst. D* **2003**, *59*, 620–626.

(3) Franklin, R. E.; Gosling, R. G. *Nature* **1953**, *172*, 156–157.

(4) Berg, J. M. *Biochemistry*; 5th ed.; W.H. Freeman: New York, 2002.

(5) Rich, A. *Gene* **1993**, *135*, 99–109.

(6) Ho, P. S.; Carter, M. In *DNA Replication-Current Advances*; Seligmann, H., Ed.; InTech, 2011.

(7) Wang, A. H.-J.; Quigley, G. J.; Kolpak, F. J.; Crawford, J. L.; van Boom, J. H.; van der Marel, G.; Rich, A. *Nature* **1979**, *282*, 680–686.

(8) Sinden, R. R. *DNA structure and function*; Academic Press: San Diego, 1994.

(9) Voet, D.; Voet, J. G. *Biochimie*; De Boeck: Bruxelles, 2005.

(10) Gilbert, D. E.; Feigon, J. *Curr. Opin. Struct. Biol.* **1999**, *9*, 305–314.

(11) Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023–2024.

(12) Hoogsteen, K. *Acta Cryst.* **1959**, *12*, 822–823.

(13) Hoogsteen, K. *Acta Cryst.* **1963**, *16*, 907–916.

(14) Holliday, R. *Genetical Research* **1964**, *5*, 282.

(15) Ortiz-Lombardia, M.; Gonzalez, A.; Eritja, R.; Aymami, J.; Azorin, F.; Coll, M. *Nat. Struct. Mol. Biol.* **1999**, *6*, 913–917.

(16) Eichman, B. F.; Vargason, J. M.; Mooers, B. H. M.; Ho, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 3971–3976.

(17) McKinney, S. A.; Declais, A.-C.; Lilley, D. M. J.; Ha, T. *Nat. Struct. Mol. Biol.* **2003**, *10*, 93–97.

(18) Huppert, J. L. *Chem. Soc. Rev.* **2008**, *37*, 1375–1384.

(19) Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M. *Nature* **1991**, *350*, 718–720.

(20) Balasubramanian, S.; Neidle, S. *Curr. Opin. Chem. Biol.* **2009**, *13*, 345–353.

(21) Gehring, K.; Leroy, J.-L.; Guéron, M. *Nature* **1993**, *363*, 561–565.

(22) Guéron, M.; Leroy, J.-L. *Curr. Opin. Struct. Biol.* **2000**, *10*, 326–331.

(23) Guo, K.; Pourpak, A.; Beetz-Rogers, K.; Gokhale, V.; Sun, D.; Hurley, L. H. *J. Am. Chem. Soc.* **2007**, *129*, 10220–10228.

(24) Pazos, E.; Mosquera, J.; Vázquez, M. E.; Mascareñas, J. L. *ChemBioChem* **2011**, *12*, 1958–1973.

(25) Kubik, S. *Chem. Soc. Rev.* **2010**, *39*, 3648–3663.

(26) Neidle, S. *Nat. Prod. Rep.* **2001**, *18*, 291–309.

(27) Strekowski, L.; Wilson, B. *Mutation Res. Fund. Mol. Mech. Mutagen.* **2007**, *623*, 3–13.

(28) Paul, A.; Bhattacharya, S. *Curr. Sci.* **2012**, *102*, 212–231.

(29) Bahr, M.; Gabelica, V.; Granzhan, A.; Teulade-Fichou, M.-P.; Weinhold, E. *Nucleic Acids Res.* **2008**, *36*, 5000–5012.

(30) Holman, G. G.; Zewail-Foote, M.; Smith, A. R.; Johnson, K. A.; Iverson, B. L. *Nat. Chem.* **2011**, *3*, 875–881.

(31) Rhoden Smith, A.; Iverson, B. L. *J. Am. Chem. Soc.* **2013**, *135*, 12783–12789.

(32) Kim, S. K.; Nordén, B. *FEBS Lett.* **1993**, *315*, 61–64.

(33) Hu, W.; Blecking, C.; Kralj, M.; Šuman, L.; Piantanida, I.; Schrader, T. *Chem. Eur. J.* **2012**, *18*, 3589–3597.

(34) Nielsen, P. E. In *Methods in Enzymology*; Jonathan B. Chaires, M. J. W., Ed.; Academic Press, 2001; Vol. Volume 340, pp. 329–340.

(35) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1996**, *382*, 559–561.

(36) Komeda, S.; Moulaei, T.; Woods, K. K.; Chikuma, M.; Farrell, N. P.; Williams, L. D. *J. Am. Chem. Soc.* **2006**, *128*, 16092–16103.

(37) Boer, D. R.; Canals, A.; Coll, M. *Dalton Trans.* **2009**, 399–414.

(38) Komeda, S.; Moulaei, T.; Chikuma, M.; Odani, A.; Kipping, R.; Farrell, N. P.; Williams, L. D. *Nucleic Acids Res.* **2011**, *39*, 325–336.

(39) Qu, Y.; Harris, A.; Hegmans, A.; Petz, A.; Kabolizadeh, P.; Penazova, H.; Farrell, N. *J. Inorg. Biochem.* **2004**, *98*, 1591–1598.

(40) Harris, A. L.; Yang, X.; Hegmans, A.; Povirk, L.; Ryan, J. J.; Kelland, L.; Farrell, N. P. *Inorg. Chem.* **2005**, *44*, 9598–9600.

(41) Harris, A. L.; Ryan, J. J.; Farrell, N. *Mol. Pharmacol.* **2006**, *69*, 666–672.

(42) Atkins, J. L.; Patel, M. B.; Daschbach, M. M.; Meisel, J. W.; Gokel, G. W. *J. Am. Chem. Soc.* **2012**, *134*, 13546–13549.

(43) Whitesides, G. M.; Grzybowski, B. *Science* **2002**, *295*, 2418–2421.

(44) Zhang, F.; Nangreave, J.; Liu, Y.; Yan, H. *J. Am. Chem. Soc.* **2014**, *136*, 11198–11211.

(45) Singh, Y.; Murat, P.; Defrancq, E. *Chem. Soc. Rev.* **2010**, *39*, 2054–2070.

(46) Boersma, A. J.; Megens, R. P.; Feringa, B. L.; Roelfes, G. *Chem. Soc. Rev.* **2010**, *39*, 2083–2092.

(47) Iwaura, R.; Yoshida, K.; Masuda, M.; Ohnishi-Kameyama, M.; Yoshida, M.; Shimizu, T. *Angew. Chem. Int. Ed.* **2003**, *42*, 1009–1012.

(48) Iwaura, R.; Hoeben, F. J. M.; Masuda, M.; Schenning, A. P. H. J.; Meijer, E. W.; Shimizu, T. *J. Am. Chem. Soc.* **2006**, *128*, 13298–13304.

(49) Iwaura, R.; Shimizu, T. *Angew. Chem. Int. Ed.* **2006**, *45*, 4601–4604.

(50) Iwaura, R.; Ohnishi-Kameyama, M.; Shimizu, T. *Chem. Commun.* **2008**, 5770–5772.

(51) Iwaura, R.; Ohnishi-Kameyama, M.; Iizawa, T. *Chem. Eur. J.* **2009**, *15*, 3729–3735.

(52) Iwaura, R.; Yui, H.; Someya, Y.; Ohnishi-Kameyama, M. *Journal of Photochemistry and Photobiology B: Biology* **2014**, *130*, 199–204.

(53) Iwaura, R.; Yoshida, K.; Masuda, M.; Yase, K.; Shimizu, T. *Chem. Mater.* **2002**, *14*, 3047–3053.

(54) Iwaura, R.; Minamikawa, H.; Shimizu, T. *Journal of Colloid and Interface Science* **2004**, *277*, 299–303.

(55) Yang, W.; Xia, P. F.; Wong, M. S. *Org. Lett.* **2010**, *12*, 4018–4021.

(56) Yang, W.; Chen, Y.; Wong, M. S.; Lo, P. K. *Biomacromolecules* **2012**, *13*, 3370–3376.

(57) Numata, M.; Sugiyasu, K.; Kishida, T.; Haraguchi, S.; Fujita, N.; Park, S. M.; Yun, Y. J.; Kim, B. H.; Shinkai, S. *Org. Biomol. Chem.* **2008**, *6*, 712–718.

(58) Banchelli, M.; Berti, D.; Baglioni, P. *Angew. Chem. Int. Ed.* **2007**, *46*, 3070–3073.

(59) Sezi, S.; Wagenknecht, H.-A. *Chem. Commun.* **2013**, *49*, 9257–9259.

(60) Janssen, P. G. A.; Vandenbergh, J.; van Dongen, J. L. J.; Meijer, E. W.; Schenning, A. P. H. J. *J. Am. Chem. Soc.* **2007**, *129*, 6078–6079.

(61) Surin, M.; Janssen, P. G. A.; Lazzaroni, R.; Leclère, P.; Meijer, E. W.; Schenning, A. P. H. J. *Adv. Mat.* **2009**, *21*, 1126–1130.

(62) Janssen, P. G. A.; van Dongen, J. L. J.; Meijer, E. W.; Schenning, A. P. H. J. *Chem. Eur. J.* **2009**, *15*, 352–360.

(63) Janssen, P. G. A.; Brankaert, N. J. M.; Vila, X.; Schenning, A. P. H. J. *Soft Matter* **2010**, *6*, 1494–1502.

(64) Janssen, P. G. A.; Ruiz-Carretero, A.; González-Rodríguez, D.; Meijer, E. W.; Schenning, A. P. H. J. *Angew. Chem. Int. Ed.* **2009**, *48*, 8103–8106.

(65) Sargsyan, G.; Schatz, A. A.; Kubelka, J.; Balaz, M. *Chem. Commun.* **2013**, *49*, 1020–1022.

(66) Sargsyan, G.; Leonard, B. M.; Kubelka, J.; Balaz, M. *Chem. Eur. J.* **2014**, *20*, 1878–1892.

(67) Seifert, J. L.; Connor, R. E.; Kushon, S. A.; Wang, M.; Armitage, B. A. *J. Am. Chem. Soc.* **1999**, *121*, 2987–2995.

(68) Wang, M.; Silva, G. L.; Armitage, B. A. *J. Am. Chem. Soc.* **2000**, *122*, 9977–9986.

(69) Ruff, Y.; Moyer, T.; Newcomb, C. J.; Demeler, B.; Stupp, S. I. *J. Am. Chem. Soc.* **2013**, *135*, 6211–6219.

(70) Wade, W. S.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 8783–8794.

(71) Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983–7988.

(72) White, S.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1997**, *4*, 569–578.

(73) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468–471.

(74) Kielkopf, C. L.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Nat. Struct. Mol. Biol.* **1998**, *5*, 104–109.

(75) Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Science* **1998**, *282*, 111–115.

(76) Weyermann, P.; Dervan, P. B. *J. Am. Chem. Soc.* **2002**, *124*, 6872–6878.

(77) Huc, I.; Lehn, J.-M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2106–2110.

(78) Lehn, J.-M. *Chem. Eur. J.* **1999**, *5*, 2455–2463.

(79) Ramström, O.; Lehn, J.-M. *Nat. Rev. Drug Discov.* **2002**, *1*, 26–36.

(80) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, *106*, 3652–3711.

(81) Ladame, S. *Org. Biomol. Chem.* **2008**, *6*, 219–226.

(82) Miller, B. L. *Dynamic combinatorial chemistry in drug discovery, bioorganic chemistry, and materials science*; Wiley: Hoboken, N.J., 2010.

(83) Reek, J. N. H.; Otto, S. *Dynamic combinatorial chemistry*; Wiley-VCH: Weinheim, 2010.

(84) Cougnon, F. B. L.; Sanders, J. K. M. *Acc. Chem. Res.* **2012**, *45*, 2211–2221.

(85) Li, J.; Nowak, P.; Otto, S. *J. Am. Chem. Soc.* **2013**, *135*, 9222–9239.

(86) Matache, M.; Bogdan, E.; Hădade, N. D. *Chem. Eur. J.* **2014**, *20*, 2106–2131.

(87) Ulrich, S.; Dumy, P. *Chem. Commun.* **2014**, *50*, 5810–5825.

(88) Gareiss, P. C.; Miller, B. L. In *Dynamic Combinatorial Chemistry*; John Wiley & Sons, Inc., 2009; pp. 83–117.

(89) Escalante, A. M.; Furlan, R. L. E. In *Dynamic Combinatorial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, 2010; pp. 49–90.

(90) Breuil, P.-A. R.; Reek, J. N. H. In *Dynamic Combinatorial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, 2010; pp. 91–108.

(91) Wietor, J.-L.; Botterhuis, N. E.; Sijbesma, R. P. In *Dynamic Combinatorial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, 2010; pp. 151–168.

(92) Maeda, T.; Otsuka, H.; Takahara, A. In *Dynamic Combinatorial Chemistry*; John Wiley & Sons, Inc., 2009; pp. 229–260.

(93) Severin, K. In *Dynamic Combinatorial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, 2010; pp. 169–184.

(94) Otto, S.; Furlan, R. L. .; Sanders, J. K. . *Drug Discov. Today* **2002**, *7*, 117–125.

(95) Otto, S. In *Dynamic Combinatorial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, 2010; pp. 23–47.

(96) Ludlow, R. F.; Otto, S. *J. Am. Chem. Soc.* **2008**, *130*, 12218–12219.

(97) Epstein, D. M.; Choudhary, S.; Churchill, M. R.; Keil, K. M.; Eliseev, A. V.; Morrow, J. R. *Inorg. Chem.* **2001**, *40*, 1591–1596.

(98) Perez-Fernandez, R.; Pittelkow, M.; Belenguer, A. M.; Sanders, J. K. M. *Chem. Commun.* **2008**, 1738–1740.

(99) Klekota, B.; Hammond, M. H.; Miller, B. L. *Tetrahedron Lett.* **1997**, *38*, 8639–8642.

(100) Besenius, P.; Cormack, P. A. G.; Liu, J.; Otto, S.; Sanders, J. K. M.; Sherrington, D. C. *Chem. Eur. J.* **2008**, *14*, 9006–9019.

(101) Besenius, P.; Cormack, P. A. G.; Ludlow, R. F.; Otto, S.; Sherrington, D. C. *Org. Biomol. Chem.* **2010**, *8*, 2414–2418.

(102) McNaughton, B. R.; Miller, B. L. *Org. Lett.* **2006**, *8*, 1803–1806.

(103) Beeren, S. R.; Sanders, J. K. M. In *Dynamic Combinatorial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, 2010; pp. 1–22.

(104) Ramström, O.; Amorim, L.; Caraballo, R.; Norberg, O. In *Dynamic Combinatorial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, 2010; pp. 109–150.

(105) Klekota, B.; Miller, B. L. *Tetrahedron* **1999**, *55*, 11687–11697.

(106) Godoy-Alcántar, C.; Yatsimirsky, A. K.; Lehn, J.-M. *J. Phys. Org. Chem.* **2005**, *18*, 979–985.

(107) Meyer, C. D.; Joiner, C. S.; Stoddart, J. F. *Chem. Soc. Rev.* **2007**, *36*, 1705–1723.

(108) Hochgürtel, M.; Kroth, H.; Piecha, D.; Hofmann, M. W.; Nicolau, C.; Krause, S.; Schaaf, O.; Sonnenmoser, G.; Eliseev, A. V. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 3382–3387.

(109) Nguyen and Ivan Huc, R. *Chem. Commun.* **2003**, 942–943.

(110) Dirksen, A.; Dirksen, S.; Hackeng, T. M.; Dawson, P. E. *J. Am. Chem. Soc.* **2006**, *128*, 15602–15603.

(111) Bhat, V. T.; Caniard, A. M.; Luksch, T.; Brenk, R.; Campopiano, D. J.; Greaney, M. F. *Nat. Chem.* **2010**, *2*, 490–497.

(112) Clipson, A. J.; Bhat, V. T.; McNae, I.; Caniard, A. M.; Campopiano, D. J.; Greaney, M. F. *Chem. Eur. J.* **2012**, *18*, 10562–10570.

(113) Miller, B. L. In *Dynamic Combinatorial Chemistry*; John Wiley & Sons, Inc., 2009; pp. 1–42.

(114) Ladame, S.; Whitney, A. M.; Balasubramanian, S. *Angew. Chem. Int. Ed.* **2005**, *44*, 5736–5739.

(115) Bugaut, A.; Toulmé, J.-J.; Rayner, B. *Angew. Chem. Int. Ed.* **2004**, *43*, 3144–3147.

(116) Bugaut, A.; Bathany, K.; Schmitter, J.-M.; Rayner, B. *Tetrahedron Lett.* **2005**, *46*, 687–690.

(117) Azéma, L.; Bathany, K.; Rayner, B. *ChemBioChem* **2010**, *11*, 2513–2516.

(118) Rowan, S. J.; Cantrill, S. J.; Cousins, G. R. L.; Sanders, J. K. M.; Stoddart, J. F. *Angew. Chem. Int. Ed.* **2002**, *41*, 898–952.

(119) Skene, W. G.; Lehn, J.-M. P. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 8270–8275.

(120) Lehn, J.-M. *Prog. Polym. Sci.* **2005**, *30*, 814–831.

(121) Maeda, T.; Otsuka, H.; Takahara, A. *Prog. Polym. Sci.* **2009**, *34*, 581–604.

(122) Lehn, J.-M. *Aust. J. Chem.* **2010**, *63*, 611–623.

(123) Wojtecki, R. J.; Meador, M. A.; Rowan, S. J. *Nat. Mater.* **2011**, *10*, 14–27.

(124) Giuseppone, N.; Lehn, J.-M. *J. Am. Chem. Soc.* **2004**, *126*, 11448–11449.

(125) Giuseppone, N.; Fuks, G.; Lehn, J.-M. *Chem. Eur. J.* **2006**, *12*, 1723–1735.

(126) Wei, Z.; Yang, J. H.; Zhou, J.; Xu, F.; Zrinyi, M.; Dussault, P. H.; Osada, Y.; Chen, Y. M. *Chem. Soc. Rev.* **2014**.

(127) Li, X.; Zhan, Z.-Y. J.; Knipe, R.; Lynn, D. G. *J. Am. Chem. Soc.* **2002**, *124*, 746–747.

(128) Li, X.; Lynn, D. G. *Angew. Chem. Int. Ed.* **2002**, *41*, 4567–4569.

(129) Sreenivasachary, N.; Hickman, D. T.; Sarazin, D.; Lehn, J.-M. *Chem. Eur. J.* **2006**, *12*, 8581–8588.

(130) Martin, A. R.; Barvik, I.; Luvino, D.; Smietana, M.; Vasseur, J.-J. *Angew. Chem. Int. Ed.* **2011**, *50*, 4193–4196.

(131) Lehrman, S. *Nature* **1999**, *401*, 517–518.

(132) Hargrove, A. E.; Nieto, S.; Zhang, T.; Sessler, J. L.; Anslyn, E. V. *Chem. Rev.* **2011**, *111*, 6603–6782.

(133) Ohara, K.; Smietana, M.; Vasseur, J.-J. *J. Am. Soc. Mass Spect.* **2006**, *17*, 283–291.

CHAPITRE 2 :

ETUDES DES INTERACTIONS PETITES MOLECULES / ADN

PAR L'APPROCHE RATIONNELLE

INTRODUCTION – RAPPELS BIBLIOGRAPHIQUES

Les auto-assemblages supramoléculaires offrent une méthode attractive pour préparer des (bio)matériaux fonctionnels par une approche ascendante.[1] La technologie de l'origami d'ADN illustre bien la complexité structurale qui peut être atteinte par des assemblages spontanés d'ADN et des procédés de pliage.[2] D'un autre côté, les oligonucléotides peuvent aussi servir de matrices pour générer des auto-assemblages matricés par l'ADN *via* des interactions non covalentes de petites molécules avec l'ADN. Cette approche peut conduire à des auto-assemblages biomoléculaires fonctionnels avec des propriétés catalytiques[3,4] ou optiques[5,6] uniques. De plus, l'assemblage programmé de multiples petites molécules, identiques ou différentes, sur une seule matrice peut être utile pour générer des systèmes de reconnaissance multivalents homotopiques ou hétérotopiques.[7–9] Alternativement, la matrice oligonucléotidique peut également être utilisée pour la synthèse supramoléculaire sans enzyme dirigée par l'ADN de polymères[10–15] ou de matériaux poreux à base de silice.[16]

Une progression importante a récemment été atteinte dans ce domaine par l'identification de composés qui s'auto-assemblent avec des oligonucléotides simples brins en solvants aqueux par des liaisons hydrogène avec les nucléobases. Ces composés incluent des nucléobases fonctionnalisées telles que les thymines,[17–21] mais aussi des analogues de nucléobases artificiels tels que les diaminotriazines,[22–25] et les complexes cyclènes qui interagissent tous les deux avec les bases thymines d'oligonucléotides simples brins.[26] Les interactions secondaires comme l'empilement ont été identifiées comme étant cruciales pour stabiliser ces auto-assemblages non covalents. Il est important de rappeler que ces interactions d'empilement sont également essentielles, en conjonction avec les liaisons hydrogène, pour l'hybridation des oligonucléotides et la formation de la double hélice d'ADN.[27] Les interactions d'empilement peuvent aussi être utilisées en tant que seule interaction non covalente pour conduire à la formation de complexes par intercalation.[28]

En plus de la reconnaissance des nucléobases et de l'intercalation, l'interaction avec le squelette phosphodiester polyanionique a récemment été reconnue comme un nouveau mode de liaison possible.[29,30] Bien que cette approche ne donne pas accès à des auto-assemblages hétérotopiques programmés par la séquence, elle peut être utile pour des applications en délivrance de gènes. En effet, de petites molécules ligands d'anions dérivées de l'isophthalamide et la dipicolinamide ont récemment démontrés leur efficacité en tant que vecteurs pour le transport d'ADN plasmidique dans *Escherichia coli*.[31] Dans cet unique

exemple, il a été suggéré, mais non démontré, que les interactions de type empilement pourraient jouer un rôle important dans la stabilisation de l'auto-assemblage. Lorsqu'une meilleure compréhension des effets sous-jacents conduisant à ces processus d'auto-assemblage sera atteinte, les petites molécules capables d'interagir avec les oligonucléotides par interactions non covalentes avec le squelette phosphodiester pourraient révéler un fort potentiel en tant que vecteurs universels pour des applications en délivrance de gènes.

Dans ce chapitre sont reportées la conception et la synthèse de composés dérivés de guanidinium dont la capacité à s'auto-assembler avec de simples brins d'ADN en milieu aqueux a été démontrée par une combinaison d'études spectroscopiques et de spectrométrie de masse. De plus, le rôle des interactions d'empilement est étudié et leur importance dans la stabilisation de ce type d'auto-assemblage est démontrée.

Le groupement guanidinium est bien connu pour interagir avec les oxoanions par interactions électrostatiques et liaisons hydrogène.[32–45] Par exemple, dans le contexte de la reconnaissance d'oligonucléotides, la reconnaissance de l'ARN TAR par la protéine TAT du VIH se fait par des interactions guanidinium – phosphodiester.[46,47] Cependant, malgré le nombre significatif de systèmes basés sur les guanidiniums utilisés pour des applications en vectorisation et pour la complexation d'oligonucléotides,[48–62] la base structurelle des auto-assemblages entre les composés guanidiniums synthétiques et les oligonucléotides reste peu comprise. Un guanidinium fonctionnalisé par un pyrène s'est montré capable d'interagir efficacement avec un anion pyrophosphate dans le méthanol, formant un complexe 2:1 par interactions guanidinium – phosphate (K_{11} = 1,0.10^4 M^{-1} et K_{21} = 1,2.10^8 M^{-2}) avec l'assistance des interactions d'empilement des pyrènes, induisant la formation d'un excimère.[63] Plus récemment, Vasseur et ses collaborateurs ont démontré, par spectrométrie de masse MALDI-ToF, que ce composé guanidinium mono-fonctionnalisé par un pyrène interagit sélectivement avec les groupements phosphodiesters d'un ADN simple brin.[64] De nouveau, l'importance des interactions d'empilement est mise en évidence. Toutefois, peu de choses sont connues à propos du comportement de cet auto-assemblage matricé par l'ADN en solution.

Dans ce contexte, des composés guanidiniums bisfonctionnalisés portant des substituants aromatiques de différentes surfaces – en l'occurrence des groupements phényles, naphtyles et pyrènes – ont été conçus pour étudier leur capacité d'auto-assemblage avec l'ADN en solution et caractériser le rôle précis des interactions d'empilement (Figure 2.). Un

groupement méthyle a été introduit en tant qu'espaceur entre le groupement guanidinium central et les substituants aromatiques pour apporter de la flexibilité et permettre aux aromatiques d'adopter des conformations coplanaires tout en respectant la fonction guanidinium. De tels composés guanidiniums bisfonctionnalisés ne sont à l'heure actuelle pas décrits dans la littérature. Pourtant, l'introduction d'un second substituant pourrait augmenter la stabilité des auto-assemblages formés avec l'ADN en comparaison de ceux formés avec des composés équivalents monofonctionnalisés, lorsque de tels auto-assemblages existent.

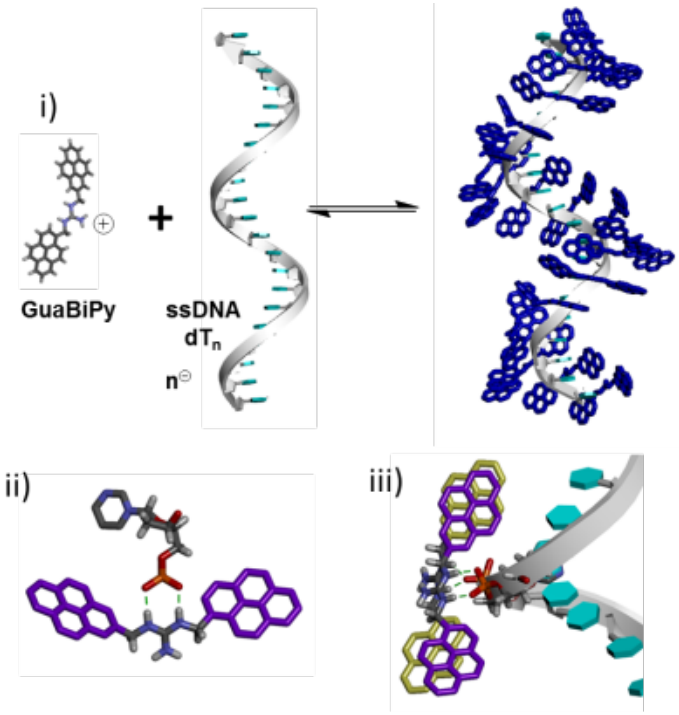


Figure 2.1 : Représentation schématique de l'auto-assemblage matricé par l'ADN avec des composés guanidiniums bisfonctionnalisés (exemple avec le composé bispyrène **GuaBiPy**) (i) et les différents types d'interactions intermoléculaires : pont salin (ii) et empilement π-π avec les groupements pyrènes de deux composés **GuaBiPy** présentés de différentes couleurs (iii).

I. SYNTHESE DES DIFFERENTS COMPOSES

Le composé guanidinium bisphényle **GuaBiPhe** est synthétisé par réaction directe de la benzylamine sur le bromure de cyanogène avec un rendement de 54% (Schéma 2.1) selon un mode opératoire décrit dans la littérature avec un composé similaire.[65]

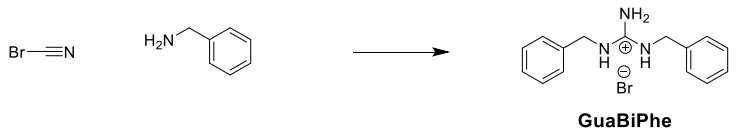


Schéma 2.1 : Synthèse du composé **GuaBiPhe**. Réactifs et conditions : acétonitrile, reflux, 54%.

Cette méthode de préparation des guanidiniums a également été testée avec la 1-pyrèneméthymamine dans le but de former le composé guanidinium bispyrène **GuaBiPy** (Schéma 2.2). Malheureusement, dans les différentes conditions testées, la triazine **1** a été formée plus facilement que le produit **GuaBiPy** souhaité : en effet, lors des suivis réactionnels par LC-MS, il a été observé que la triazine **1** se formait à température ambiante alors qu'il était nécessaire de chauffer pour former le guanidinium **GuaBiPy** (Figure 2.2). De plus, lorsque ce produit **GuaBiPy** a pu être isolé avec cette méthode de synthèse, il était contaminé par de la 1-pyrèneméthylamine résiduelle dont il était difficile de se débarrasser.

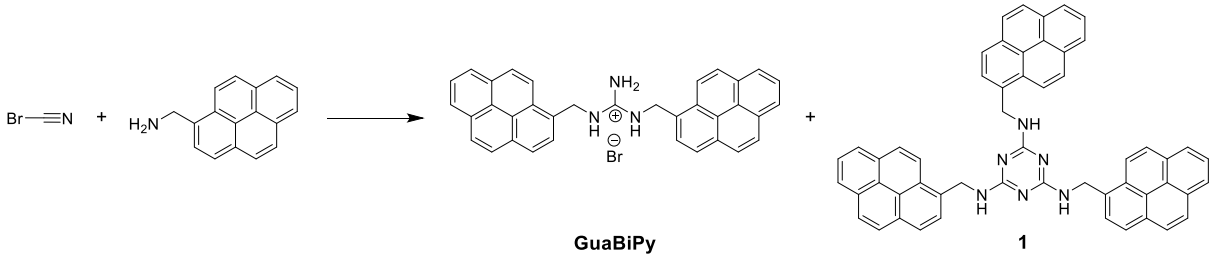


Schéma 2.2 : Préparation du composé **GuaBiPy** *via* la voie « bromure de cyanogène » et formation de la triazine **1**. Différents réactifs et conditions testés : CH_2Cl_2, reflux ; CH_2Cl_2/DMF, 50°C ; Et_3N, CH_2Cl_2, reflux ; Et_3N, sans solvant, 45°C ; Et_3N, 50°C ; Et_3N, DMF, 50°C.

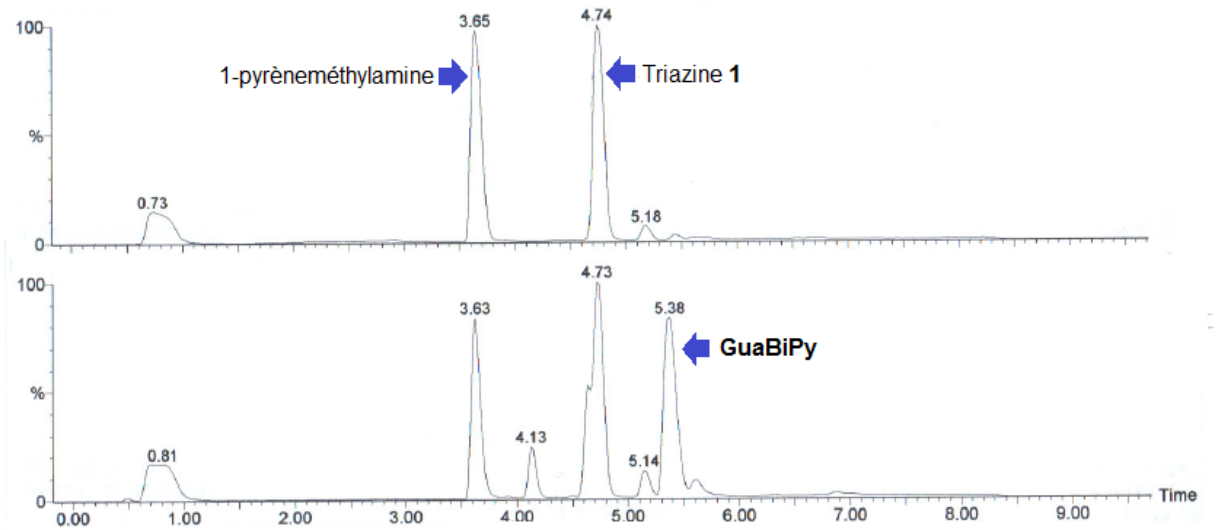


Figure 2.2 : Suivi LC-MS de l'apparition de la triazine **1** après 1 heure 30 à température ambiante (haut), puis de la formation du composé **GuaBiPy** après une nuit à 50°C (bas), lors de la réaction de la 1-pyrèneméthylamine avec le bromure de cyanogène. Méthode III (LC-MS 1).

Pour résoudre ces problèmes, une seconde voie de synthèse, passant par un intermédiaire « bis-imidazole » **2,** a été mise en œuvre pour synthétiser le composé **GuaBiPy** sans former de triazine **1** (Schéma 2.3).[66,67] L'intermédiaire **2** a été obtenu par réaction directe d'imidazole sur le bromure de cyanogène avec un rendement de 51%. La 1-pyrèneméthylamine a ensuite substitué deux fois cet intermédiaire **2** pour donner le produit **GuaBiPy** avec un rendement de 24%.

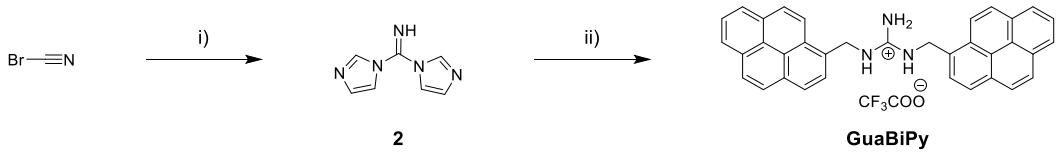


Schéma 2.3 : Synthèse du composé **GuaBiPy** *via* l'intermédiaire « bis-imidazole » **2**. Réactifs et conditions : i) imidazole, CH_2Cl_2, reflux, 51% ; ii) 1-pyrèneméthylamine, TFA, sans solvant, 105°C, 24%.

Le produit **GuaBiPy** a été isolé par simple filtration après reprise du milieu réactionnel dans du dichlorométhane. Malheureusement, une petite proportion d'impuretés non identifiées (visibles en RMN ^{1}H, inférieures à 5% de l'aire totale en HPLC) était également présente avec le produit. Le produit **GuaBiPy** étant insoluble dans la plupart des solvants, les méthodes de purification envisageables étaient restreintes. De nombreux essais de purification par recristallisation ou trituration ont été effectués mais n'ont pas permis d'éliminer les impuretés présentes. Le produit a donc été utilisé tel quel par la suite malgré la présence de ces impuretés résiduelles.

La méthode de synthèse *via* l'intermédiaire « bis-imidazole » **2** a ensuite été adaptée pour obtenir le composé guanidinium bisnaphtyle **GuaBiNaph** (Schéma 2.4). L'intermédiaire **2** a été doublement substitué par la 1-naphtylméthylamine pour obtenir le produit **GuaBiNaph** souhaité avec un rendement de 31%.

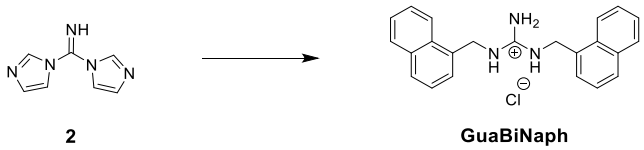


Schéma 2.4 : Synthèse du composé **GuaBiNaph** *via* l'intermédiaire « bis-imidazole » **2**. Réactifs et conditions : 1-naphtylméthylamine, TFA, sans solvant, 105°C, 31%.

A l'issue de cette double substitution, le produit **GuaBiNaph** était soluble dans le dichlorométhane et un lavage avec une solution aqueuse de chlorure d'ammonium saturée a été effectué pour éliminer l'imidazole présent dans le mélange. Le produit obtenu a ensuite été purifié par chromatographie flash sur gel de silice, puis recristallisé, la chromatographie sur

gel s'étant révélée insuffisante pour éliminer toutes les impuretés présentes. Le produit **GuaBiNaph** a ainsi été obtenu pur sans difficultés.

II. ANALYSES SPECTROSCOPIQUES

A. GENERALITES ET INTERETS DES TECHNIQUES EMPLOYEES

1. Spectroscopies UV-Visible et de fluorescence

Les méthodes de spectroscopie UV-Visible et de fluorescence reposent sur le fait que la lumière, en tant que rayonnement électromagnétique, peut transmettre de l'énergie à toute molécule qu'elle rencontre.[68,69] Dans le domaine de l'UV-Visible (longueurs d'onde comprises entre 200 et 800 nm), l'énergie absorbée par une molécule peut induire une transition électronique, c'est à dire le passage d'électrons de valence d'un état fondamental à un état excité (Figure 2.3). Lorsqu'une telle transition électronique a lieu, le retour à l'état fondamental peut ensuite se faire de trois manières : dissipation thermique, fluorescence ou phosphorescence. Les spectroscopies UV-Vis et de fluorescence sont donc complémentaires l'une de l'autre puisque que la première mesure l'absorption du rayonnement tandis que la seconde mesure l'émission d'énergie liée à la fluorescence.

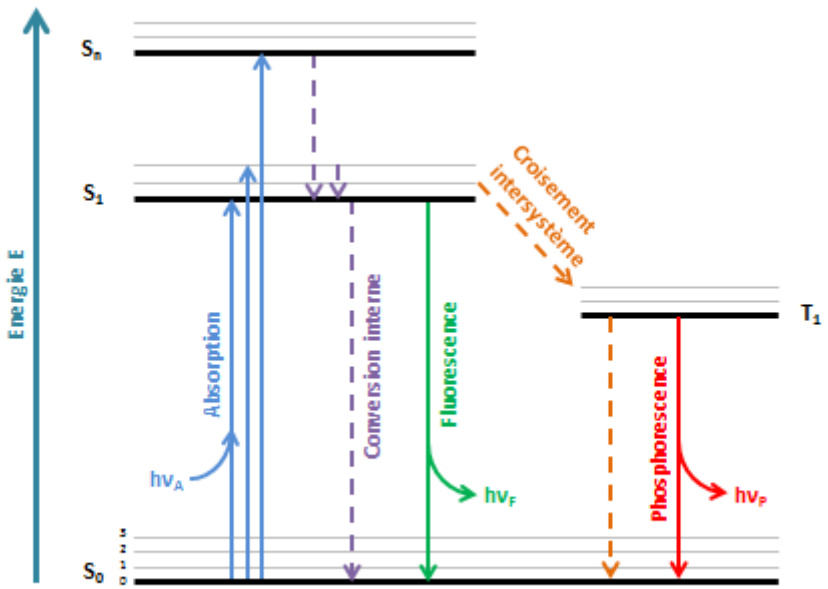


Figure 2.3 : Diagramme de Jablonski illustrant les transitions électroniques entre état fondamental et états excités lors des phénomènes d'absorption, de fluorescence et de phosphorescence. Les états singulets et triplets

sont respectivement notés S et T. Les indices 0, 1, …, n indiquent les états fondamental et excités d'énergie croissante. Les lignes grises représentent les états vibrationnels de chaque état électronique. Les processus radiatifs sont indiqués par des flèches pleines tandis que les processus non-radiatifs sont indiqués par des flèches en pointillés.

En fonction de leur structure électronique, toutes les molécules ne sont pas susceptibles d'absorber la lumière dans le domaine UV-Visible. En effet, l'énergie absorbée, dépendante de la fréquence ν du rayonnement et donc de sa longueur d'onde λ (Équation 2.1), doit être suffisante pour induire la transition électronique.

Équation 2.1

$$E = h\nu = h\frac{c}{\lambda}$$

Avec :

- E énergie d'un photon (en J),
- h constante de Planck (~ 6,63 x 10^{-34} J.s^{-1}),
- ν fréquence de l'onde (en Hz),
- c vitesse de la lumière dans le vide (~ 3 x 10^{8} m.s^{-1}),
- λ longueur d'onde (en m).

Une transition de type σ – σ* nécessite une énergie trop importante pour qu'un rayonnement dans le domaine de l'UV soit suffisant pour que ce type de transition soit observable. En revanche, les transitions de type π – π* ou n – π*, nécessitant moins d'énergie, sont observables dans ce domaine du spectre électromagnétique : c'est pourquoi les composés insaturés absorbent dans le domaine UV.

Concernant la fluorescence, elle n'est pas toujours détectée car la dissipation thermique peut être plus rapide. Dans le cas des composés guanidiniums bisfonctionnalisés étudiés, ils absorbent tous en UV en raison de leurs substituants aromatiques mais la fluorescence n'est bien visible qu'avec les substituants napthyle et pyrène. C'est pourquoi seuls les composés **GuaBiNaph** et **GuaBiPy** ont été étudiés par spectroscopie.

Différents paramètres influent sur les propriétés d'absorption UV-Vis et / ou de fluorescence d'un composé. Par exemple, l'absorbance mesurée en spectroscopie UV est directement proportionnelle à la concentration de l'échantillon selon la loi de Beer – Lambert

(Équation 2.2). Le solvant ou le pH de la solution sont également des paramètres susceptibles d'entraîner des variations importantes des propriétés optiques.

Équation 2.2

$$A = -\log\left(\frac{I}{I_0}\right) = \varepsilon \ell C$$

Avec :

- A absorbance (sans unité),
- I et I_0 intensité de la lumière respectivement incidente et sortante de l'échantillon (le ratio $\frac{I}{I_0}$ est la transmittance T, sans unité),
- ε coefficient d'extinction molaire ($m^2.mol^{-1}$),
- ℓ longueur du trajet optique (cm),
- C concentration ($mol.L^{-1}$).

La fluorescence d'un composé peut également varier en fonction de son organisation. Cette technique permet en effet d'observer la formation d'excimères, qui sont des dimères formés suite à l'excitation des molécules par la lumière.[70] De tels complexes sont susceptibles de se former lorsque les deux molécules qui le constituent sont en étroite interaction entre elles, comme dans le cas d'un empilement. Par exemple, le pyrène est bien connu pour former des excimères (Figure 2.4).[71] L'étude de la fluorescence du composé **GuaBiPy** est donc particulièrement intéressante pour obtenir des informations sur l'arrangement dans l'espace de ce composé, puisque que le phénomène d'empilement, s'il a lieu, sera immédiatement visible sur le spectre d'émission de fluorescence.

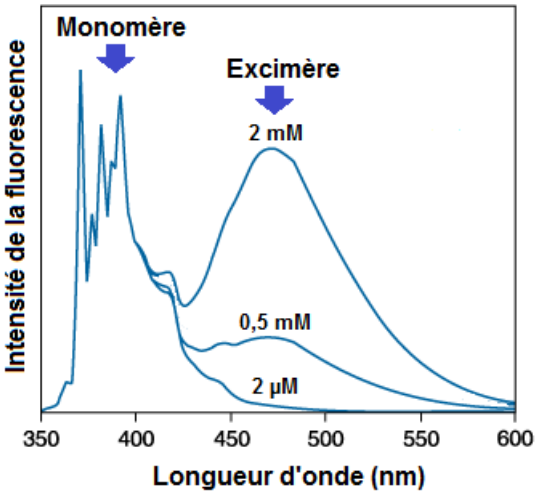


Figure 2.4 : Formation de l'excimère du pyrène dans l'éthanol.[72] Les spectres ont été normalisés sur le pic à 371,5 nm du monomère et sont essentiellement identiques en dessous de 400 nm après normalisation. Le ratio monomère sur excimère (371,5 nm / 470 nm) est dépendant de la concentration en pyrène.

2. Dichroïsme circulaire

Le phénomène de dichroïsme circulaire est lié au fait qu'une molécule chirale absorbe différemment la lumière selon que cette dernière est polarisée circulairement à gauche ou à droite.[68,73] Cette différence d'absorbance ΔA est mesurée par la spectroscopie de dichroïsme circulaire (Équation 2.3).

Équation 2.3

$$\Delta A = A_{PCG} - A_{PCD}$$

Avec :

- ΔA différence d'absorbance,
- A_{PCD} absorbance de la lumière polarisée circulairement à droite,
- A_{PCG} absorbance de la lumière polarisée circulairement à gauche.

En tenant compte de la concentration de l'échantillon C et de la longueur du trajet optique ℓ, il est possible de remonter à la valeur du dichroïsme circulaire molaire Δε (Équation 2.5) avec la loi de Beer – Lambert suivant l'Équation 2.4.

Équation 2.4

$$\Delta A = (\varepsilon_{PCG} - \varepsilon_{PCD})\, C\ell$$

Soit :

Équation 2.5

$$\Delta\varepsilon = \varepsilon_{PCG} - \varepsilon_{PCD}$$

Avec :

- ε_{PCG} coefficient d'extinction molaire pour la lumière polarisée circulairement à gauche ($m^2.mol^{-1}$ ou $L.mol^{-1}.cm^{-1}$),
- ε_{PCD} coefficient d'extinction molaire pour la lumière polarisée circulairement à droite ($m^2.mol^{-1}$ ou $L.mol^{-1}.cm^{-1}$),
- C concentration molaire ($mol.L^{-1}$),
- ℓ longueur du trajet optique (cm).

Pour des raisons historiques, le dichroïsme circulaire peut être exprimé en degrés d'ellipticité θ qui se rapporte à la polarisation elliptique générée lorsqu'un rayon de lumière polarisé linéairement est partiellement absorbé. L'expression du dichroïsme circulaire par une valeur d'ellipticité est un héritage de la polarimétrie à partir de laquelle la spectroscopie de dichroïsme circulaire a été mise au point. La conversion entre le dichroïsme circulaire et l'ellipticité θ est donnée par l'Équation 2.6.

Équation 2.6

$$\Delta A = \frac{\theta}{32{,}982}$$

Avec : θ ellipticité (millidegré m° ou 1/1000 de degré).

Le dichroïsme circulaire est particulièrement intéressant pour étudier la conformation structurale de macromolécules chirales comme l'ADN[74] ou les protéines[75] bien que cette technique ne donne pas d'informations au niveau atomique. En effet, les spectres obtenus permettent de discriminer entre les différents types de structures secondaires ou tertiaires des macromolécules. Les différentes conformations A, B ou Z d'un ADN sont ainsi différenciées par dichroïsme circulaire et l'usage de cette technique de spectroscopie peut également apporter des informations précieuses quand des changements structuraux ont lieu lors d'interactions avec des petites molécules.[76]

La spectroscopie de dichroïsme circulaire présente également des avantages pratiques par rapport aux autres techniques, comme la diffraction X ou la RMN, susceptibles de fournir de telles informations structurales sur l'ADN :[74]

- faible quantité de produit nécessaire et faible concentration ;
- taille des composés non limitée ;
- possibilité de tester un grande nombre de conditions différentes et d'ajouter différents agents sans perturber l'observation du dichroïsme circulaire ;
- analyse en solution ou en film (corrélation possible avec d'autres techniques où l'analyse ne peut pas se faire en solution) ;
- rapide et de coût relativement abordable.

Dans notre cas, l'usage de la spectroscopie de dichroïsme circulaire va permettre de déterminer si les petites molécules synthétisées **GuaBiNaph** et **GuaBiPy** interagissent avec l'ADN en entraînant un changement conformationnel de celui-ci.

*B. AUTO-ASSEMBLAGE DU COMPOSE **GuaBiPy** AVEC UN OLIGOTHYMINE dT_n EN SOLUTION.*

Le composé **GuaBiPy** a été étudié en premier car il est le plus proche de la sonde guanidinium mono-fonctionnalisée dont la capacité à interagir avec un pyrophosphate[63] et un simple brin d'ADN[64] a été démontrée. De la même manière, les oligothymines dT_n ont été choisies comme matrices puisque la sélectivité de l'interaction guanidinium – phosphodiester a été démontrée sur un simple brin d'oligothymine.[64]

Une titration par fluorescence montre que l'émission de l'excimère (472 nm) est favorisée au dépend de l'émission du monomère (395 nm) au cours de l'addition du dT_{40} (Figure 2.5). Cet effet est clairement visible lorsque le ratio de l'émission de l'excimère sur l'émission du monomère (E/M) est tracé graphiquement (Figure 2.5). Contrairement au composé **GuaBiPy** seul qui montre un ratio E/M presque constant sur la gamme de concentration utilisée, le ratio E/M augmente graduellement avec l'addition de dT_{40} jusqu'à atteindre un maximum pour une stœchiométrie de 30 **GuaBiPy** par brin de dT_{40}, soit un ratio N/P de 0,75 (ratio molaire des charges positives portées par les groupements guanidiniums, N, sur les charges négatives formelles présentes sur le simple brin d'ADN, P). L'addition supplémentaire de **GuaBiPy** conduit à une diminution très importante du ratio E/M jusqu'à

atteindre des valeurs similaires à celles du composé **GuaBiPy** seul après l'addition de plus de 50 équivalents de **GuaBiPy** par brin de dT_{40}.

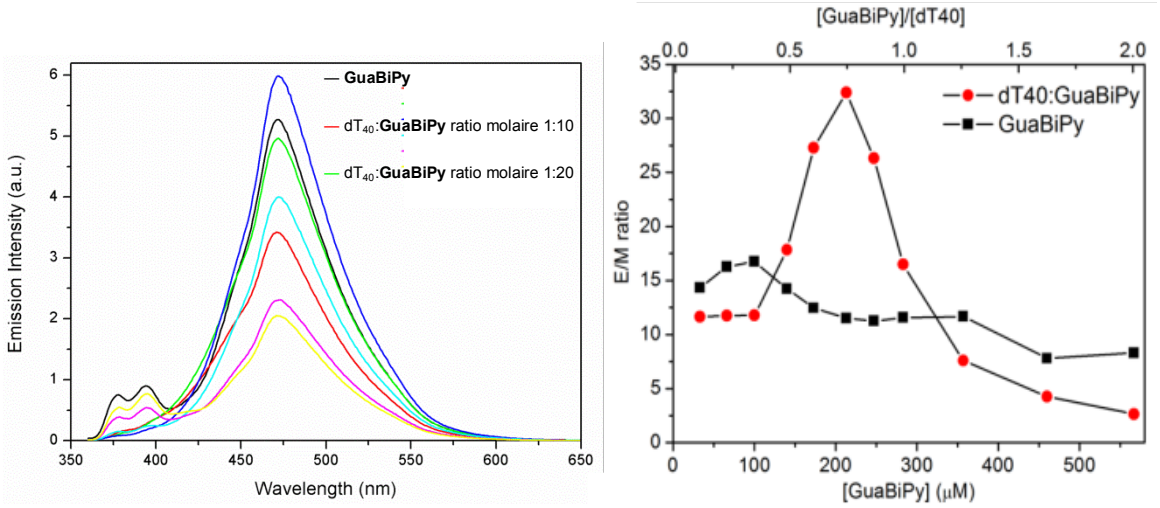


Figure 2.5 : Spectres de fluorescence de la titration du dT_{40} par le composé **GuaBiPy** (gauche) et tracé du ratio de l'émission de l'excimère (472 nm) sur l'émission du monomère (395 nm) E/M (droite). Mesures effectuées à –5°C dans le tampon Tris-EDTA (TE ; pH 7,4). Concentration en dT_{40} de 7,10 mM. Concentration de la solution mère de **GuaBiPy** de 10 mM dans le DMSO.

Ces résultats indiquent que le dT_{40} agit comme une matrice et organise le composé **GuaBiPy** d'une manière qui favorise la formation d'excimère. La spectroscopie de dichroïsme circulaire apporte des informations supplémentaires qui soutiennent cette conclusion. Alors que le composé **GuaBiPy** et le dT_{40} sont silencieux dans la gamme 300 – 400 nm, un signal de dichroïsme circulaire induit est observé lors de l'addition de **GuaBiPy** à une solution de dT_{40}, démontrant ainsi que le composé **GuaBiPy** s'auto-assemble avec le dT_{40} et adopte un arrangement chiral (Figure 2.6 et Figure 2.7).

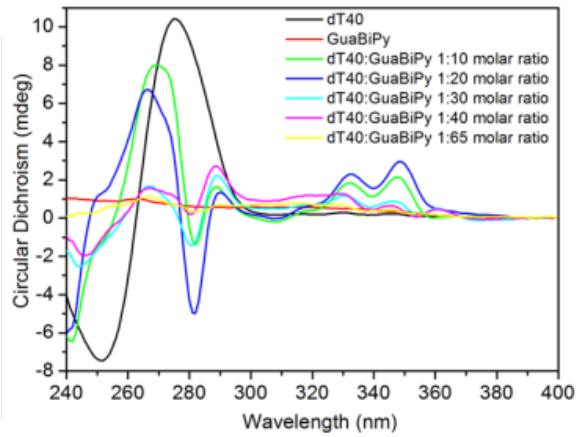


Figure 2.6 : Titration du dT_{40} par le composé **GuaBiPy**, suivi par spectroscopie de dichroïsme circulaire. Mesures effectuées à -5°C dans un tampon TE (pH 7,4). Concentration en dT_{40} de 7,10 mM. Concentration de la solution mère de **GuaBiPy** de 10 mM dans le DMSO.

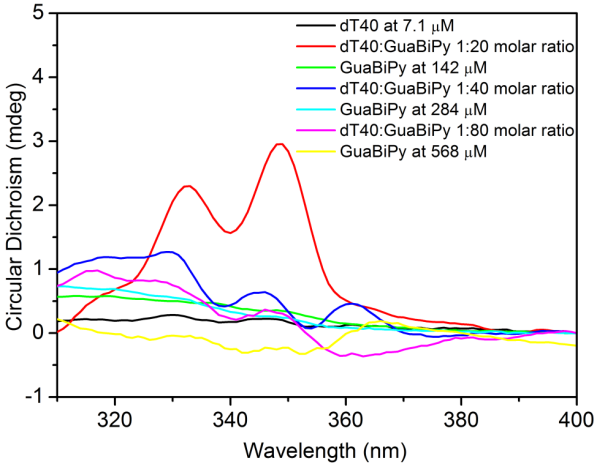


Figure 2.7 : Spectres de dichroïsme circulaire du composé **GuaBiPy** pure et du mélange dT_{40}:**GuaBiPy** à différents ratios molaires.

Pris ensemble, ces résultats indiquent que jusqu'à 30 **GuaBiPy** peuvent s'auto-assembler sur un brin de dT_{40} dans ces conditions. Puisque l'émission de l'excimère est promue par l'ADN simple brin, le système auto-assemblé implique probablement l'empilement du composé **GuaBiPy** le long du squelette phosphodiester du simple brin d'ADN. Cette observation est en accord avec le résultat précédemment reporté indiquant que les composés guanidiniums interagissent sélectivement avec les résidus phosphodiesters du simple brin d'oligothymine.[64] Cette hypothèse est renforcée par le fait que le ratio E/M est bien plus faible lorsque l'expérience d'auto-assemblage est effectuée en présence de tampon phosphate plutôt qu'en présence de tampon Tris-EDTA (TE) (Figure 2.8). En effet, dans ce cas, les ions phosphates peuvent jouer le rôle de compétiteur et conduire à la dissociation du complexe si les interactions guanidinium – phosphodiester gouvernent le processus d'auto-assemblage.

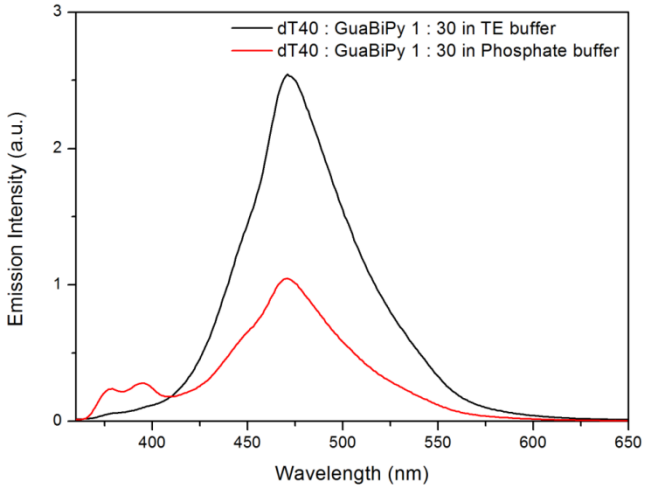


Figure 2.8 : Spectres d'émission (à λ_{exc} = 348 nm) d'un mélange de dT_{40}:**GuaBiPy** à un ratio molaire de 1:30 dans un tampon TE (ligne noire) et dans un tampon phosphate (ligne rouge) à -5 ºC lors du second cycle de chauffage (H-C). [dT_{40}] ~ 5.7 µM.

L'addition d'un excès de **GuaBiPy** semble être nuisible à l'auto-assemblage matricé par l'ADN. Une possible explication pour cette observation peut être le fait qu'un second processus d'auto-assemblage – impliquant peut-être l'intercalation – a lieu après l'addition de 30 équivalents de **GuaBiPy**, ce qui conduit finalement au désassemblage de l'empilement de **GuaBiPy** matricé par l'ADN par distorsion structurale de la matrice ADN. Cette hypothèse peut être soutenue par le fait que le spectre de dichroïsme circulaire indique une perte de chiralité de la structure de la matrice ADN simple brin après addition de plus de 30 équivalents de **GuaBiPy** (Figure 2.6).

C. EFFET DE LA TEMPERATURE SUR LES AUTO-ASSEMBLAGES

Des cycles de chauffage / refroidissement avec des températures variant de -5°C à 80°C ont été effectués sur les complexes **GuaBiPy**:dT_{40} (stœchiométrie de 30 pour 1) et suivis en mesurant le ratio E/M. Ce ratio E/M est maximal à basse température et diminue lors du chauffage (Figure 2.9). Puisque le composé **GuaBiPy** ne subit pas de changements d'un tel ordre de grandeur, il est donc possible de conclure que l'effet de matrice est plus fort à basse température.

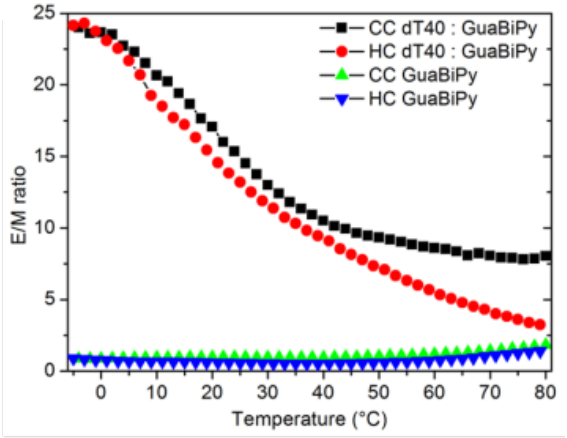


Figure 2.9 : Ratio E/M du complexe **GuaBiPy**:dT_{40} 30:1 et du composé **GuaBiPy** seul, dans le tampon TE, en fonction de la température (CC : cycle de refroidissement ; HC : cycle de chauffage).

Le processus est essentiellement réversible bien qu'un effet de fatigue soit observé après quelques cycles de chauffage / refroidissement (Figure 2.10). De ces expériences, la température de dissociation – la température à laquelle le ratio E/M est à la moitié du maximum – peut être estimée à environ 40°C. Cette température caractérise la robustesse de l'auto-assemblage en fonction de la température.

Une température de dissociation bien plus basse (5 – 10°C) a été déterminée quand cette expérience a été effectuée dans un tampon phosphate au lieu du tampon TE, renforçant encore ainsi l'hypothèse que les interactions guanidinium – phosphodiester gouvernent l'auto-assemblage (Figure 2.10).

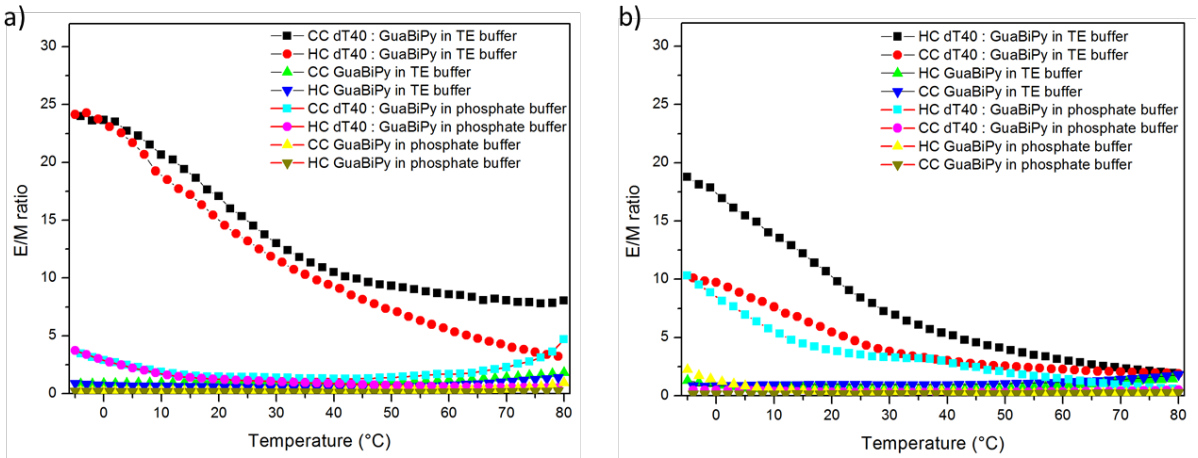


Figure 2.10 : Ratio E/M du complexe **GuaBiPy**:dT_{40} 30:1 et du composé **GuaBiPy** seul, dans les tampons TE ou phosphate, en fonction de la température : a) premier cycle de refroidissement (CC) et second cycle de chauffage (HC); b) troisième cycle de chauffage (HC) et troisième cycle de refroidissement (CC).

D. Effet de la longueur de la matrice simple brin d'ADN

La longueur de la matrice ADN simple brin a précédemment été identifiée comme étant un paramètre critique gouvernant la stabilité d'auto-assemblages triazine – oligothymine.[23] Ce paramètre a donc été étudié sur l'auto-assemblage de **GuaBiPy** avec un dT_{20} et un dT_{10}, simples brins d'ADN plus courts que le dT_{40} utilisé précédemment.

La spectroscopie de dichroïsme circulaire confirme, grâce à l'observation d'un signal de dichroïsme circulaire induit dans la gamme de 320 – 360 nm, qu'un complexe supramoléculaire se forme également lors de l'addition du composé **GuaBiPy** dans une solution de dT_{20} (Figure 2.11). Par contre, aucun signal de dichroïsme circulaire induit n'apparaît en présence de dT_{10} (non présenté).

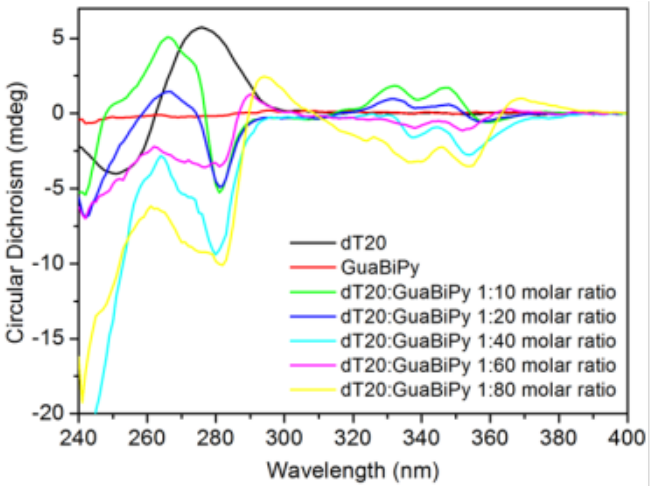


Figure 2.11 : Titration du dT_{20} par le composé **GuaBiPy**, suivi par dichroïsme circulaire. Mesures effectuées à -5°C dans le tampon TE (pH 7,4).

La forme des spectres de dichroïsme circulaire est similaire à un ratio N/P d'environ 0,5 pour les mélanges dT_{40}:**GuaBiPy** et dT_{20}:**GuaBiPy**, tandis que des différences sont observées pour des ratios N/P supérieurs à 2 (Figure 2.12). Cela suggère l'existence d'autres mécanismes d'auto-assemblages à forte concentration de **GuaBiPy**.

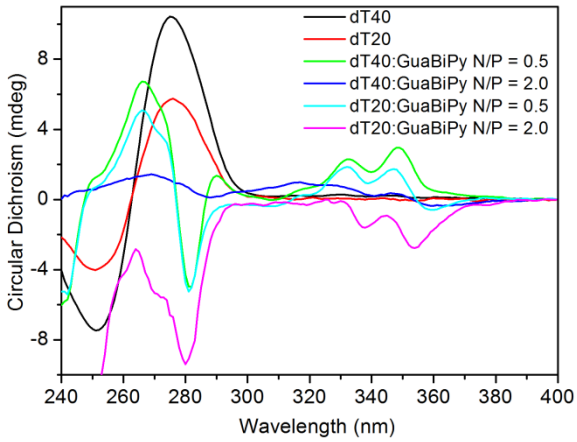


Figure 2.12 : Spectres de dichroïsme circulaire (CD) des mélanges dT_{40}:**GuaBiPy** et dT_{20}:**GuaBiPy** à différents ratios molaires N/P = 0,5 et 2,0 dans le tampon TE à -5 ºC. [dT_n] ~ 5.7 µM.

La titration par spectroscopie UV-Vis montre une augmentation plus faible de l'absorbance avec le dT_{20} qu'avec le dT_{40}, suggérant une affinité plus faible (Figure 2.13). La titration par fluorescence montre également des changements plus faibles avec le dT_{20} qu'avec le dT_{40} (Figure 2.14). Le tracé des ratios E/M indique que le dT_{20} promeut aussi la formation d'excimère mais avec un maximum du ratio E/M pour un N/P de 1,25 (25 **GuaBiPy** par simple brin d'ADN) au lieu de 0,75 (30 **GuaBiPy** par simple brin d'ADN) avec le dT_{40} comme matrice. Avec le dT_{10}, des changements sont visibles en fluorescence, le ratio E/M maximum étant atteint pour un N/P d'environ 0,3 – 0,5 (soit 3 à 5 **GuaBiPy** par dT_{10}).

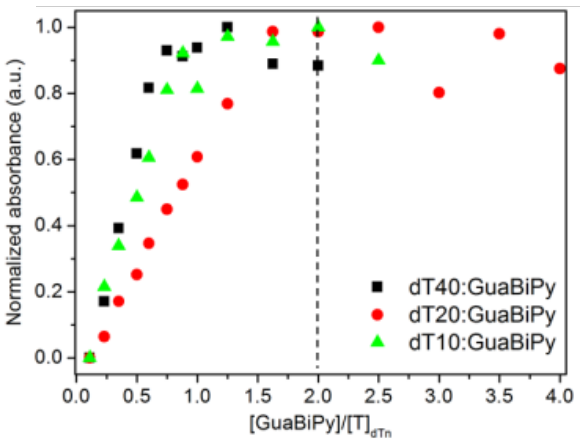


Figure 2.13 : Titration par spectroscopie UV-Vis, normalisée, de dT_n par le composé **GuaBiPy** suivi à 349 nm. La ligne en pointillés représente le ratio N/P = 2. Mesures effectuées à -5°C dans un tampon TE (pH 7,4).

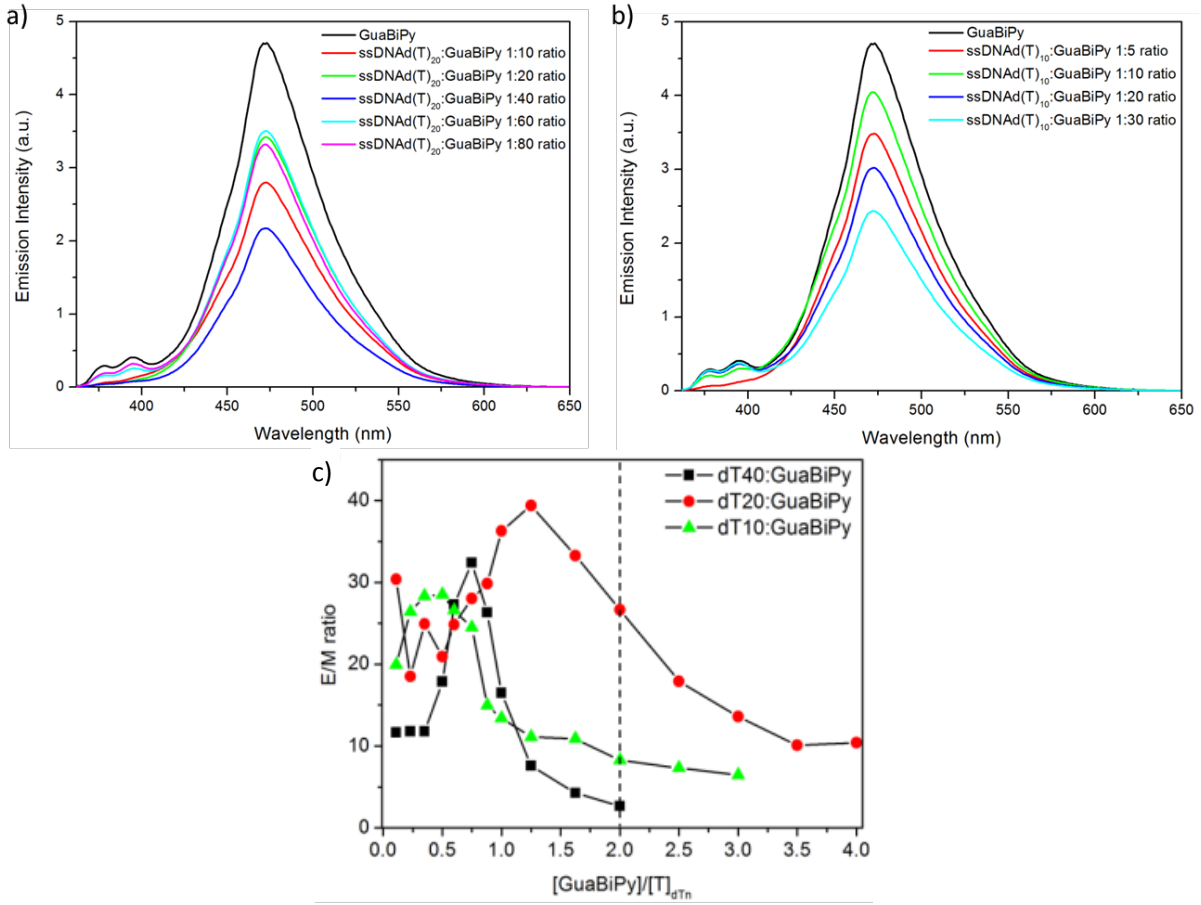


Figure 2.14 : Spectres de fluorescence de la titration par le composé **GuaBiPy** du dT_{20} (a) et du dT_{10} (b) et tracé du ratio de l'émission de l'excimère (472 nm) sur l'émission du monomère (395 nm) E/M (c). Mesures effectuées à -5°C dans un tampon TE (pH 7,4). Concentration en dT_{40} de 7,10 mM. Concentration en dT_{20} de 7,41 mM. Concentration de la solution mère du composé **GuaBiPy** de 10 mM dans le DMSO.

Des cycles de chauffage / refroidissement entre -5°C et 80°C ont été effectués sur les complexes dT_{20}:**GuaBiPy** (N/P de 1,25) et dT_{10}:**GuaBiPy** (N/P de 0,5) et montrent de faibles changements (température de dissociation estimée à moins de 5°C).

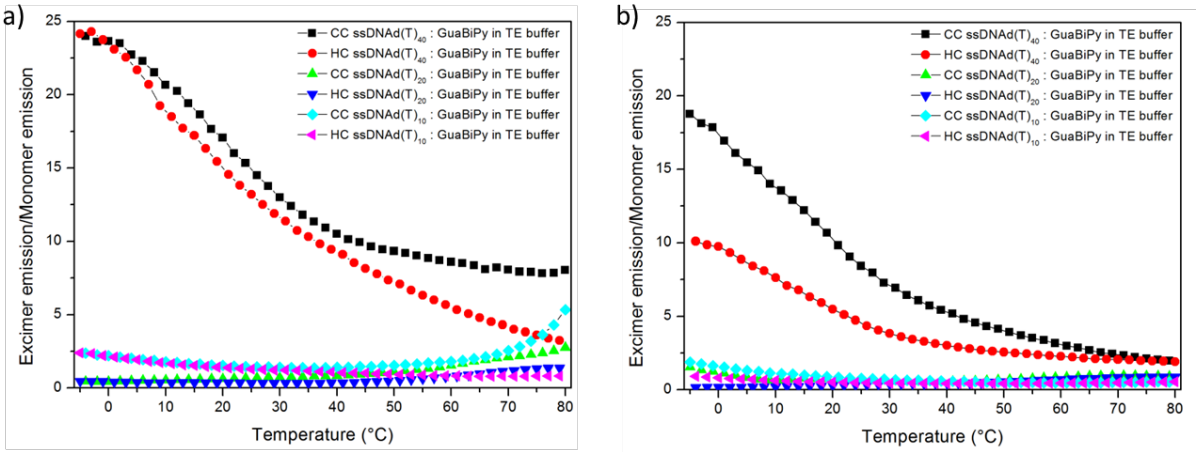


Figure 2.15 : Ratio E/M des complexes dT_n:**GuaBiPy** 1:X dans le tampon TE en fonction de la température. A) premier cycle de refroidissement (CC) et second cycle de chauffage (HC) ; et b) troisièmes cycles de chauffage (HC) et de refroidissement (CC). $[dT_n] \sim 4.6$ µM.

Ces résultats indiquent que le dT_{20} agit également comme une matrice et génère un arrangement chiral du composé **GuaBiPy** le long du simple brin d'ADN mais que les complexes qui en résultent sont bien moins stables qu'avec une matrice dT_{40}.

Pour étudier les effets de la séquence et de la topologie de l'ADN, dR_{43} et le double brin dR_{43}:dR_{rev43}, contenant un mélange de purines et de pyrimidines, ont été utilisés comme matrices. Les différentes techniques spectroscopiques confirment également l'auto-assemblage matricé par l'ADN *via* des « pinces à phosphates » (Figure 2.16).

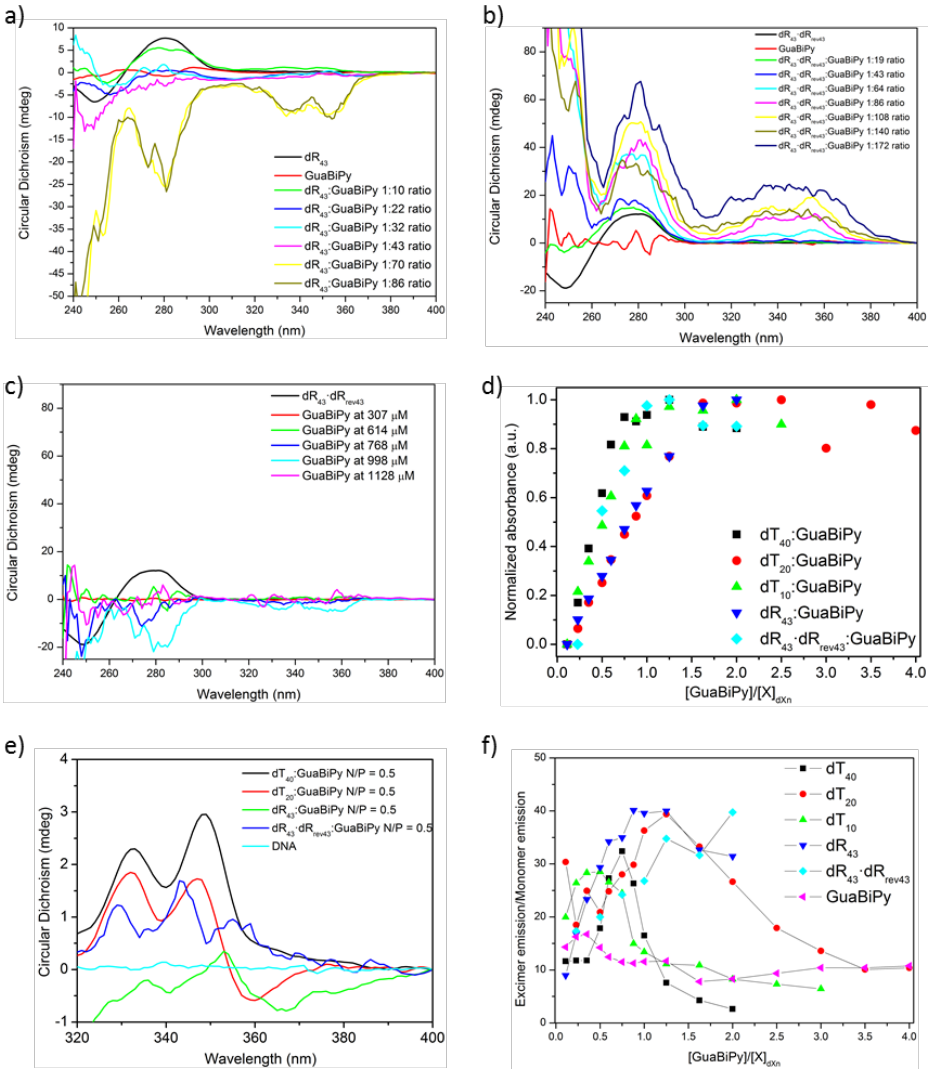


Figure 2.16 : a) Titration spectroscopique de dR43 par le composé **GuaBiPy** dans le tampon TE à -5 ºC. [dR_{43}] ~ 7.1 µM. [**GuaBiPy**] ~ 612 µM; b) titration de dR_{43}:dR_{rev43} par le composé **GuaBiPy** dans le tampon TE à -5 ºC. [dR_{43}:dR_{rev43}] ~ 7.1 µM. [**GuaBiPy**] ~ 614 µM; c) titration du composé **GuaBiPy** in TE buffer at -5 ºC.; d) titration UV-Vis normalisée suivie à 349 nm de xDNAd$(X)_n$ par le **GuaBiPy**; e) spectres de dichroïsme circulaire (CD) de dT_{40}, dT_{20}, dT_{10}, dR_{43} et dR_{43}:dR_{rev43} en présence de **GuaBiPy** à des ratios N/P de 0.5; f) trace du ratio de l'émission de l'excimère (472 nm) / émission du monomère (396 nm) de différents mélanges xDNAd$(X)_n$.

*E. AUTO-ASSEMBLAGE DU COMPOSE **GuaBiNaph** AVEC UN OLIGOTHYMINE dT_n EN SOLUTION*

Une titration par fluorescence, effectuée dans les mêmes conditions que celles utilisées pour étudier l'auto-assemblage du composé **GuaBiPy** matricé par un simple brin d'ADN, montre une augmentation de l'émission de fluorescence, le maximum étant atteint après

l'ajout de 50 équivalents de **GuaBiNaph** par dT_{40} (Figure 2.17). Ce résultat indique que le processus d'auto-assemblage a également lieu entre le composé **GuaBiNaph** et le dT_{40} dans ces conditions. Cependant, l'augmentation plus faible de l'émission de la fluorescence tend à indiquer une stabilité plus faible du complexe comparé au système **GuaBiPy**:dT_{40}.

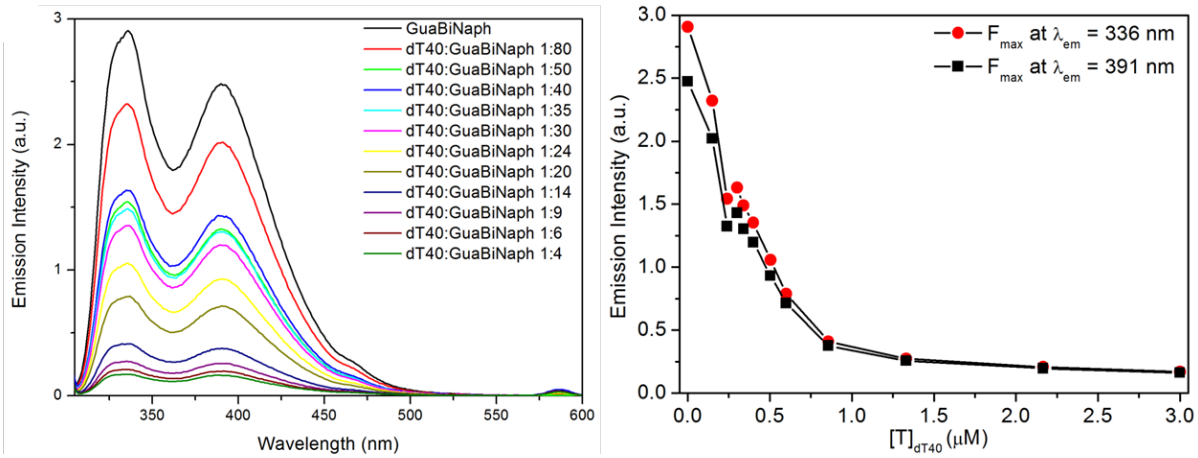


Figure 2.17 : Spectres de fluorescence de mélanges dT_{40}:**GuaBiNaph** (gauche) et intensité de l'émission (à λ_{exc} = 292 nm) du dT_{40} dans une solution de **GuaBiNaph** (droite) à divers ratios molaires. Mesures effectuées à -5°C dans le tampon TE (pH 7,4). [**GuaBiNaph**] = 12 µM. [**GuaBiNaph**] $_{stock}$= 10 mM in DMSO.

Une titration du dT_{40} par le composé **GuaBiNaph** suivie par dichroïsme circulaire montre que l'ajout du composé **GuaBiNaph** entraîne une perte de la chiralité du simple brin d'ADN (Figure 2.18). Le composé **GuaBiNaph** perturbe donc la structure de l'ADN, mais n'adopte pas d'arrangement chiral contrairement à ce qui a été observé précédemment avec le composé **GuaBiPy**.

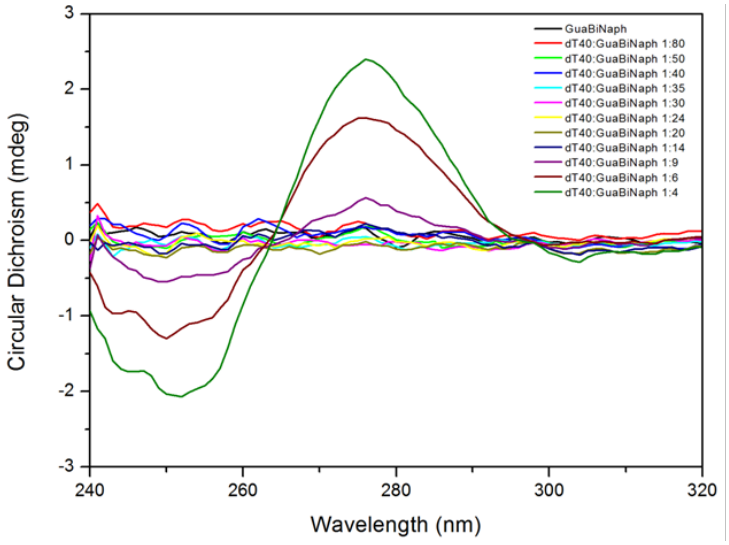


Figure 2.18 : Dichroïsme circulaire de mélanges dT_{40}:**GuaBiNaph** à divers ratios molaires. Mesures effectuées à -5°C dans le tampon TE (pH 7,4). [**GuaBiNaph**] = 12 µM. [**GuaBiNaph**] $_{stock}$= 10 mM in DMSO.

Des cycles de chauffage / refroidissement entre -5°C et 80°C ont été effectués avec le complexe dT_{40}:**GuaBiNaph** (N/P 0,75), mais n'ont pas montré, contrairement au comportement du composé **GuaBiPy** dans les mêmes conditions, de différences significatives par rapport aux changements de fluorescence induit par la température pour le composé **GuaBiNaph** seul (Figure 2.19). Ce résultat indique que le composé **GuaBiNaph** forme des auto-assemblages plus faibles avec un dT_{40} que le composé **GuaBiPy**.

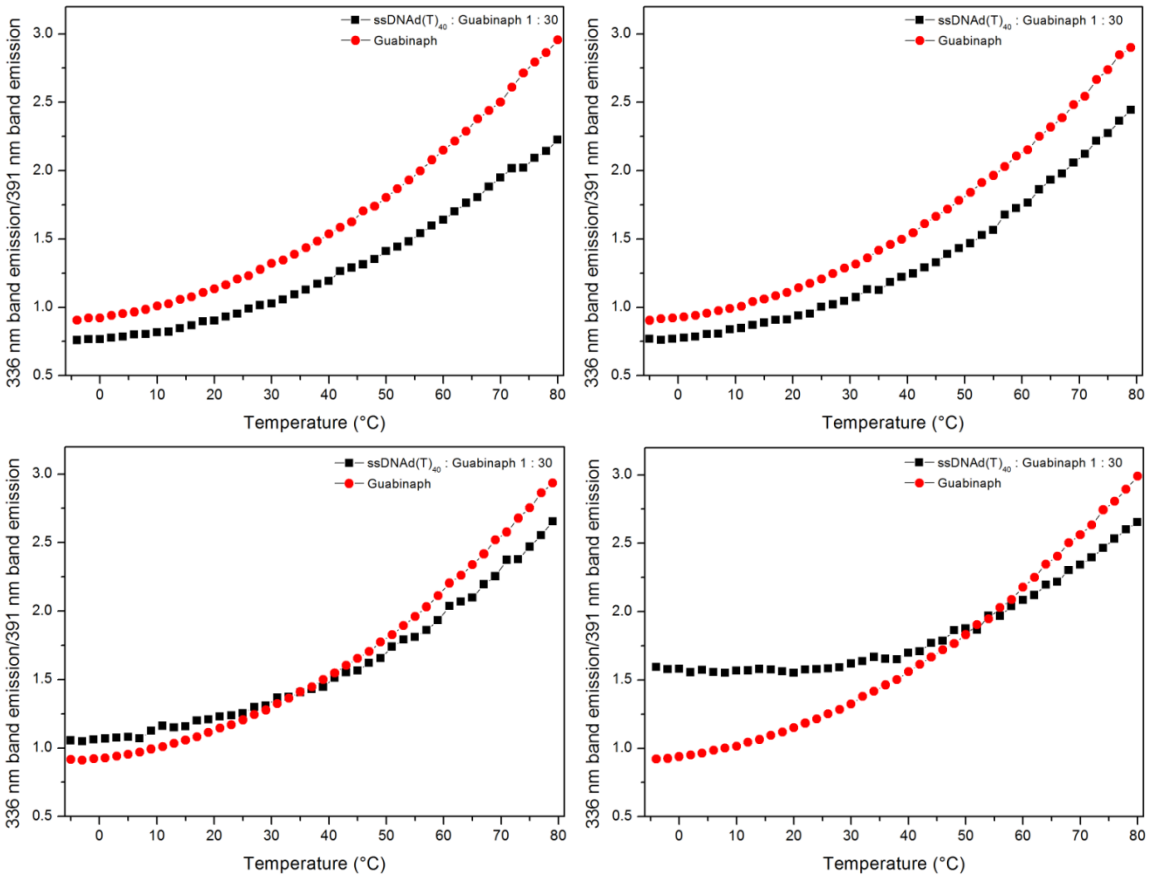


Figure 2.19 : Ratio E/M du complexe dT_{40}:**GuaBiNaph** 1:30 et du composé **GuaBiNaph** seul, dans le tampon TE, en fonction de la température. En haut à gauche : premier cycle de refroidissement ; en haut à droite : second cycle de chauffage ; en bas à gauche : troisième cycle de chauffage ; en bas à droite : troisième cycle de refroidissement. [dT_{40}] ~ 4.6 µM.

III. ANALYSES EN SPECTROMETRIE DE MASSE MALDI-TOF

La spectrométrie de masse MALDI-ToF peut être utilisée pour détecter des auto-assemblages non covalents et a précédemment été utilisée pour identifier des complexes entre des composés guanidiniums et des oligonucléotides.[64] Dans le cas présent, un simple brin de dT_{10} a été utilisé car il était plus facilement détecté par spectrométrie MALDI-ToF.

En analysant des mélanges guanidiniums – dT_{10} avec un ratio de 10 pour 1, il est possible d'observer avec nos composés jusqu'à quatre guanidiniums assemblés sur un simple brin de dT_{10} (Figure 2.20).

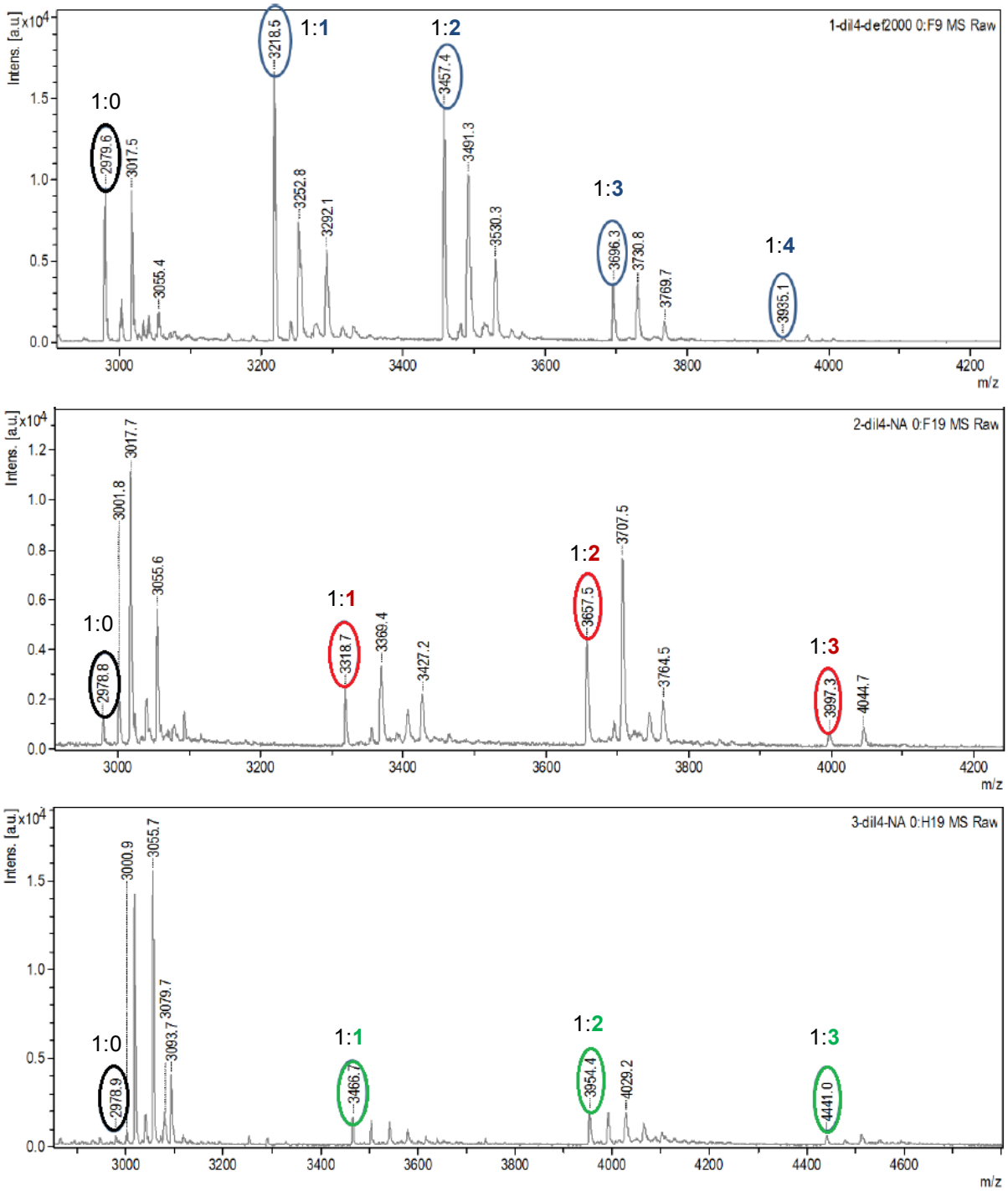


Figure 2.20 : Détection des complexes formés entre les guanidiniums **GuaBiPhe** (haut), **GuaBiNaph** (milieu) et **GuaBiPy** (bas) avec un simple brin de dT_{10} par spectrométrie de masse MALDI-ToF.

Des expériences de compétitions ont ensuite été effectuées pour déterminer la préférence d'affinité relative de l'ADN pour les trois différents composés. Les résultats

montrent la tendance suivante : **GuaBiPhe** < **GuaBiNaph** < **GuaBiPy** (Figure 2.21). A noter qu'un hétérocomplexe constitué d'un composé **GuaBiPhe** et d'un composé **GuaBiPy** est présent dans l'expérience de compétition entre ces deux composés.

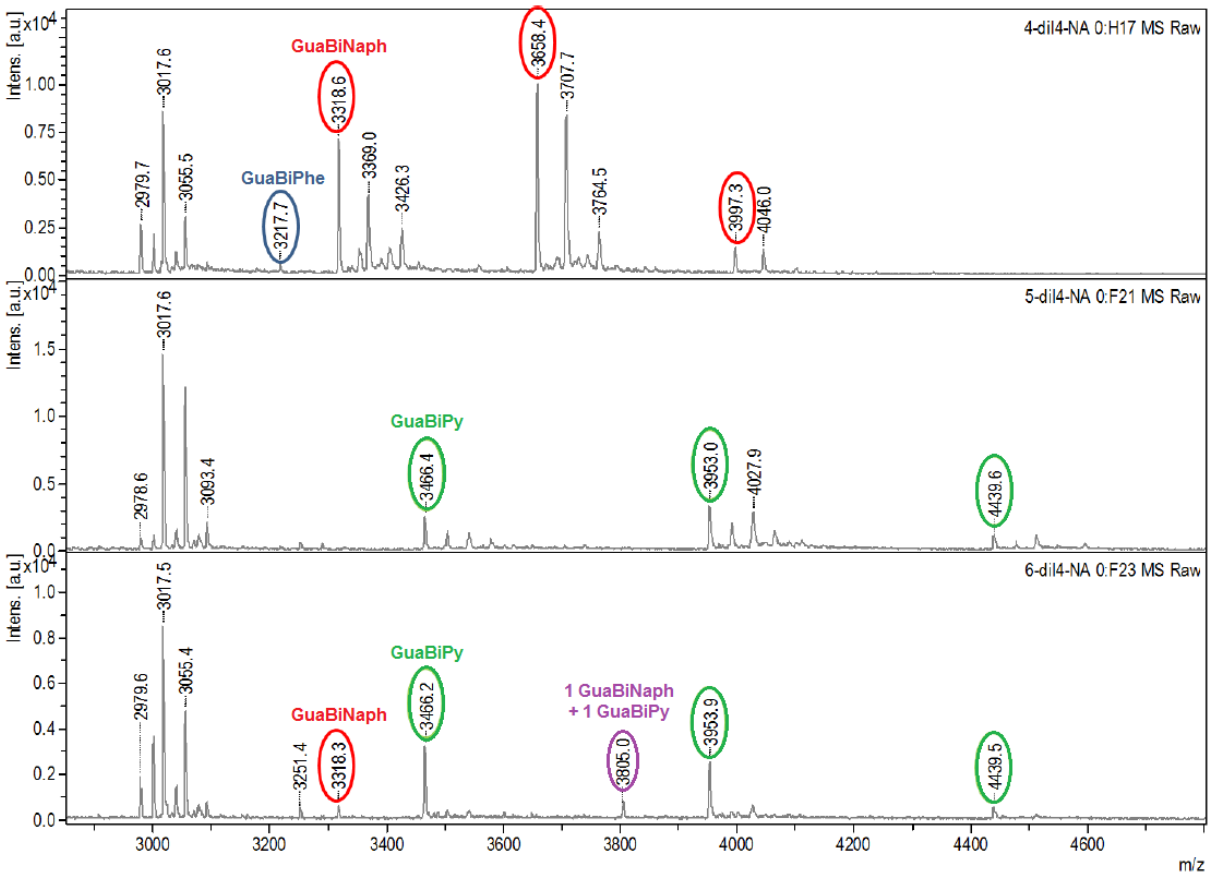


Figure 2.21 : Expériences de compétition en présence de simple brin de dT_{10}, analysées par spectrométrie de masse MALDI-ToF : **GuaBiPhe** *vs.* **GuaBiNaph** (haut), **GuaBiPhe** *vs.* **GuaBiPy** (milieu) et **GuaBiNaph** *vs.* **GuaBiPy** (bas).

Une expérience de compétition où les trois composés sont en compétition pour une matrice dT_{10} confirme la préférence pour le composé **GuaBiPy** (Figure 2.22). L'hétérocomplexe constitué d'un composé **GuaBiNaph** et d'un composé **GuaBiPy** sur un brin de dT_{10} est également observé dans cette expérience de compétition, ce qui confirme que les préférences de sélection observées sont bien dues aux préférences d'affinité plutôt qu'à une discrimination due à l'ionisation.

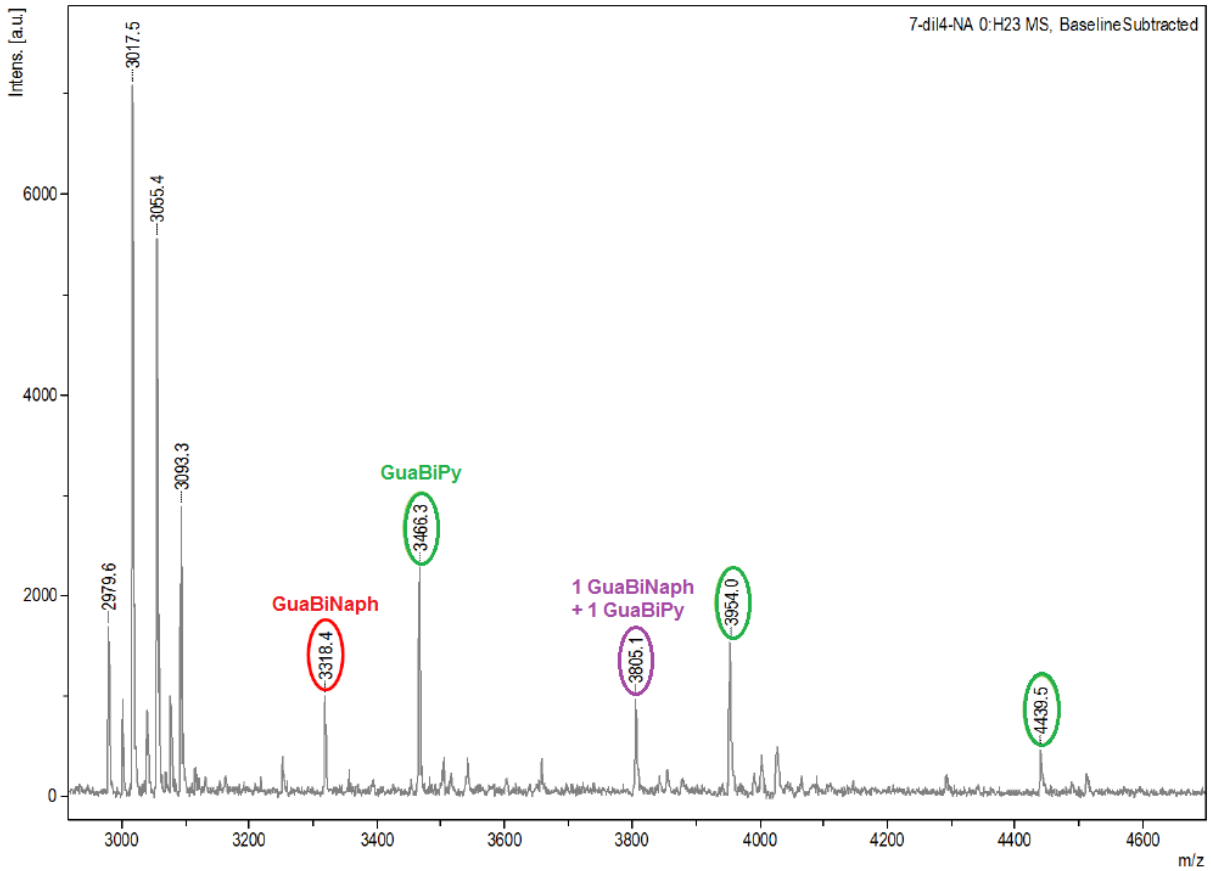


Figure 2.22 : Expérience de compétition avec un simple brin de dT_{10}, analysée par spectrométrie de masse MALDI-ToF : **GuaBiPhe** *vs.* **GuaBiNaph** *vs.* **GuaBiPy**.

Ces résultats montrent que les interactions d'empilement sont importantes pour stabiliser les assemblages matricés par l'ADN de ces composés guanidiniums bisfonctionnalisés sur de simples brins d'ADN.

CONCLUSIONS

La conception et la synthèse de composés guanidiniums bisfonctionnalisés, **GuaBiPhe**, **GuaBiNaph** et **GuaBiPy** a été reportée ici, ainsi que leur étude par une combinaison de techniques analytiques pour observer les auto-assemblages matricés par l'ADN qu'ils génèrent. Les résultats obtenus par les spectroscopies de fluorescence et de dichroïsme circulaire démontrent clairement que le composé **GuaBiPy** est en effet capable de générer des complexes supramoléculaires en présence d'oligothymines simples brins dT_{20} et dT_{40}. La stabilité de ces auto-assemblages est dépendante de la longueur du simple brin d'ADN et de la température. Les effets de tampons indiquent que cette association non covalente est gouvernée par les interactions guanidinium – phosphodiester, suggérant ainsi

que ces guanidiniums s'empilent le long du squelette phosphodiester du simple brin d'ADN. En effet, les analyses par spectrométrie de masse MALDI-ToF d'expériences de compétition entre les guanidiniums bisfonctionnalisés portant des substituants aromatiques de tailles variables démontrent l'importance des interactions d'empilement dans ces auto-assemblages matricés par des simples brins d'ADN.

REFERENCES BIBLIOGRAPHIQUES

(1) Busseron, E.; Ruff, Y.; Moulin, E.; Giuseppone, N. *Nanoscale* **2013**, *5*, 7098–7140.

(2) Aldaye, F. A.; Palmer, A. L.; Sleiman, H. F. *Science* **2008**, *321*, 1795–1799.

(3) Wang, J.; Benedetti, E.; Bethge, L.; Vonhoff, S.; Klussmann, S.; Vasseur, J.-J.; Cossy, J.; Smietana, M.; Arseniyadis, S. *Angew. Chem. Int. Ed.* **2013**, *52*, 11546–11549.

(4) Boersma, A. J.; Megens, R. P.; Feringa, B. L.; Roelfes, G. *Chem. Soc. Rev.* **2010**, *39*, 2083–2092.

(5) Ruiz-Carretero, A.; Janssen, P. G. A.; Kaeser, A.; Schenning, A. P. H. J. *Chem. Commun.* **2011**, *47*, 4340–4347.

(6) Su, W.; Bonnard, V.; Burley, G. A. *Chem. Eur. J.* **2011**, *17*, 7982–7991.

(7) Sadhu, K. K.; Röthlingshöfer, M.; Winssinger, N. *Israel J. Chem.* **2013**, *53*, 75–86.

(8) Englund, E. A.; Wang, D.; Fujigaki, H.; Sakai, H.; Micklitsch, C. M.; Ghirlando, R.; Martin-Manso, G.; Pendrak, M. L.; Roberts, D. D.; Durell, S. R.; Appella, D. H. *Nat Commun* **2012**, *3*, 614.

(9) Liu, R.; Jiang, B.; Yu, H.; Chaput, J. C. *ChemBioChem* **2011**, *12*, 1813–1817.

(10) Watson, S. M. D.; Galindo, M. A.; Horrocks, B. R.; Houlton, A. *J. Am. Chem. Soc.* **2014**, *136*, 6649–6655.

(11) Chen, W.; Schuster, G. B. *J. Am. Chem. Soc.* **2013**, *135*, 4438–4449.

(12) Lin, J.; Surin, M.; Beljonne, D.; Lou, X.; van Dongen, J. L. J.; Schenning, A. P. H. J. *Chem. Sci.* **2012**, *3*, 2732–2736.

(13) Chen, W.; Schuster, G. B. *J. Am. Chem. Soc.* **2011**, *134*, 840–843.

(14) Lo, P. K.; Sleiman, H. F. *J. Am. Chem. Soc.* **2009**, *131*, 4182–4183.

(15) Li, X.; Zhan, Z.-Y. J.; Knipe, R.; Lynn, D. G. *J. Am. Chem. Soc.* **2002**, *124*, 746–747.

(16) Numata, M.; Sugiyasu, K.; Hasegawa, T.; Shinkai, S. *Angew. Chem. Int. Ed.* **2004**, *43*, 3279–3283.

(17) Sargsyan, G.; Leonard, B. M.; Kubelka, J.; Balaz, M. *Chem. Eur. J.* **2014**, *20*, 1878–1892.

(18) Sargsyan, G.; Schatz, A. A.; Kubelka, J.; Balaz, M. *Chem. Commun.* **2013**, *49*, 1020–1022.

(19) Sezi, S.; Wagenknecht, H.-A. *Chem. Commun.* **2013**, *49*, 9257–9259.

(20) Iwaura, R.; Hoeben, F. J. M.; Masuda, M.; Schenning, A. P. H. J.; Meijer, E. W.; Shimizu, T. *J. Am. Chem. Soc.* **2006**, *128*, 13298–13304.

(21) Iwaura, R.; Yoshida, K.; Masuda, M.; Ohnishi-Kameyama, M.; Yoshida, M.; Shimizu, T. *Angew. Chem. Int. Ed.* **2003**, *42*, 1009–1012.

(22) Janssen, P. G. A.; Meeuwenoord, N.; van der Marel, G.; Jabbari-Farouji, S.; van der Schoot, P.; Surin, M.; Tomovic, Z.; Meijer, E. W.; Schenning, A. P. H. J. *Chem. Commun.* **2010**, *46*, 109–111.

(23) Janssen, P. G. A.; Jabbari-Farouji, S.; Surin, M.; Vila, X.; Gielen, J. C.; de Greef, T. F. A.; Vos, M. R. J.; Bomans, P. H. H.; Sommerdijk, N. A. J. M.; Christianen, P. C. M.; Leclère, P.; Lazzaroni, R.; van der Schoot, P.; Meijer, E. W.; Schenning, A. P. H. J. *J. Am. Chem. Soc.* **2008**, *131*, 1222–1231.

(24) Surin, M.; Janssen, P. G. A.; Lazzaroni, R.; Leclère, P.; Meijer, E. W.; Schenning, A. P. H. J. *Adv. Mat.* **2009**, *21*, 1126–1130.

(25) Janssen, P. G. A.; Vandenbergh, J.; van Dongen, J. L. J.; Meijer, E. W.; Schenning, A. P. H. J. *J. Am. Chem. Soc.* **2007**, *129*, 6078–6079.

(26) Nakamura, M.; Okaue, T.; Takada, T.; Yamana, K. *Chem. Eur. J.* **2012**, *18*, 196–201.

(27) Kool, E. T. *Chem. Rev.* **1997**, *97*, 1473–1488.

(28) Takada, T.; Otsuka, Y.; Nakamura, M.; Yamana, K. *Chem. Eur. J.* **2012**, *18*, 9300–9304.

(29) Komeda, S.; Moulaei, T.; Chikuma, M.; Odani, A.; Kipping, R.; Farrell, N. P.; Williams, L. D. *Nucleic Acids Res.* **2011**, *39*, 325–336.

(30) Komeda, S.; Moulaei, T.; Woods, K. K.; Chikuma, M.; Farrell, N. P.; Williams, L. D. *J. Am. Chem. Soc.* **2006**, *128*, 16092–16103.

(31) Atkins, J. L.; Patel, M. B.; Daschbach, M. M.; Meisel, J. W.; Gokel, G. W. *J. Am. Chem. Soc.* **2012**, *134*, 13546–13549.

(32) Kuchelmeister, H. Y.; Schmuck, C. *Chem. Eur. J.* **2011**, *17*, 5311–5318.

(33) Hargrove, A. E.; Nieto, S.; Zhang, T.; Sessler, J. L.; Anslyn, E. V. *Chem. Rev.* **2011**, *111*, 6603–6782.

(34) Rehm, T. H.; Schmuck, C. *Chem. Soc. Rev.* **2010**, *39*, 3597–3611.

(35) Tamaru, S.; Hamachi, I. In *Recognition of Anions*; Vilar, R., Ed.; Structure and Bonding; Springer Berlin Heidelberg, 2008; Vol. 129, pp. 95–125.

(36) Blondeau, P.; Segura, M.; Perez-Fernandez, R.; de Mendoza, J. *Chem. Soc. Rev.* **2007**, *36*, 198–210.

(37) Sessler, J. L.; Gale, P.; Cho, W.-S.; Stoddart, J. F.; Rowan, S. J.; Aida, T.; Rowan, A. E. *Anion Receptor Chemistry*; Monographs in Supramolecular Chemistry; The Royal Society of Chemistry, 2006.

(38) Schmidtchen, F. P. *Coordination Chemistry Reviews* **2006**, *250*, 2918–2928.

(39) Schug, K. A.; Lindner, W. *Chem. Rev.* **2005**, *105*, 67–114.

(40) Kubik, S.; Reyheller, C.; Stüwe, S. *J. Incl. Phenom. Macrocycl. Chem.* **2005**, *52*, 137–187.

(41) Best, M. D.; Tobey, S. L.; Anslyn, E. V. *Coordination Chemistry Reviews* **2003**, *240*, 3–15.

(42) Schmidtchen, F. P.; Berger, M. *Chem. Rev.* **1997**, *97*, 1609–1646.

(43) Seel, C.; Galán, A.; Mendoza, J. In *Supramolecular Chemistry II — Host Design and Molecular Recognition*; Weber, E., Ed.; Topics in Current Chemistry; Springer Berlin Heidelberg, 1995; Vol. 175, pp. 101–132.

(44) Dixon, R. P.; Geib, S. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **1992**, *114*, 365–366.

(45) Echavarren, A.; Galan, A.; Lehn, J. M.; De Mendoza, J. *J. Am. Chem. Soc.* **1989**, *111*, 4994–4995.

(46) Tao, J.; Frankel, A. D. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 2723–2726.

(47) Calnan, B. J.; Tidor, B.; Biancalana, S.; Hudson, D.; Frankel, A. D. *Science* **1991**, *252*, 1167–1171.

(48) Gehin, C.; Montenegro, J.; Bang, E.-K.; Cajaraville, A.; Takayama, S.; Hirose, H.; Futaki, S.; Matile, S.; Riezman, H. *J. Am. Chem. Soc.* **2013**, *135*, 9295–9298.

(49) Geihe, E. I.; Cooley, C. B.; Simon, J. R.; Kiesewetter, M. K.; Edward, J. A.; Hickerson, R. P.; Kaspar, R. L.; Hedrick, J. L.; Waymouth, R. M.; Wender, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 13171–13176.

(50) Qi, R.; Wu, S.; Xiao, H.; Yan, L.; Li, W.; Hu, X.; Huang, Y.; Jing, X. *J. Mater. Chem.* **2012**, *22*, 18915–18922.

(51) Klemm, K.; Radić Stojković, M.; Horvat, G.; Tomišić, V.; Piantanida, I.; Schmuck, C. *Chem. Eur. J.* **2012**, *18*, 1352–1363.

(52) Dix, A. V.; Fischer, L.; Sarrazin, S.; Redgate, C. P. H.; Esko, J. D.; Tor, Y. *ChemBioChem* **2010**, *11*, 2302–2310.

(53) Theodossiou, T. A.; Pantos, A.; Tsogas, I.; Paleos, C. M. *ChemMedChem* **2008**, *3*, 1635–1643.

(54) Pantos, A.; Tsogas, I.; Paleos, C. M. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2008**, *1778*, 811–823.

(55) Wender, P. A.; Galliher, W. C.; Goun, E. A.; Jones, L. R.; Pillow, T. H. *Adv. Drug Delivery Rev.* **2008**, *60*, 452–472.

(56) Tziveleka, L.-A.; Psarra, A.-M. G.; Tsiourvas, D.; Paleos, C. M. *J. Control. Release* **2007**, *117*, 137–146.

(57) Elson-Schwab, L.; Garner, O. B.; Schuksz, M.; Crawford, B. E.; Esko, J. D.; Tor, Y. *J. Biolog. Chem.* **2007**, *282*, 13585–13591.

(58) Goun, E. A.; Pillow, T. H.; Jones, L. R.; Rothbard, J. B.; Wender, P. A. *ChemBioChem* **2006**, *7*, 1497–1515.

(59) Funhoff, A. M.; van Nostrum, C. F.; Lok, M. C.; Fretz, M. M.; Crommelin, D. J. A.; Hennink, W. E. *Bioconjugate Chem.* **2004**, *15*, 1212–1220.

(60) Aissaoui, A.; Martin, B.; Kan, E.; Oudrhiri, N.; Hauchecorne, M.; Vigneron, J.-P.; Lehn, J.-M.; Lehn, P. *J. Med. Chem.* **2004**, *47*, 5210–5223.

(61) Oudrhiri, N.; Vigneron, J.-P.; Peuchmaur, M.; Leclerc, T.; Lehn, J.-M.; Lehn, P. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 1651–1656.

(62) Vigneron, J. P.; Oudrhiri, N.; Fauquet, M.; Vergely, L.; Bradley, J. C.; Basseville, M.; Lehn, P.; Lehn, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 9682–9686.

(63) Nishizawa, S.; Kato, Y.; Teramae, N. *J. Am. Chem. Soc.* **1999**, *121*, 9463–9464.

(64) Ohara, K.; Smietana, M.; Vasseur, J.-J. *J. Am. Soc. Mass Spect.* **2006**, *17*, 283–291.

(65) Lovick, H. M.; Michael, F. E. *Tetrahedron Lett.* **2009**, *50*, 1016–1019.

(66) Wu, Y.-Q.; Hamilton, S. K.; Wilkinson, D. E.; Hamilton, G. S. *J. Org. Chem.* **2002**, *67*, 7553–7556.

(67) Jadhav, V. D.; Schmidtchen, F. P. *J. Org. Chem.* **2008**, *73*, 1077–1087.

(68) Parson, W. W. In *Modern Optical Spectroscopy*; Springer Berlin Heidelberg, 2007; pp. 1–27.

(69) In *Principles of Fluorescence Spectroscopy*; Lakowicz, J., Ed.; Springer US, 2006; pp. 1–26.

(70) In *Principles of Fluorescence Spectroscopy*; Lakowicz, J., Ed.; Springer US, 2006; pp. 205–235.

(71) Goldschmidt, C. R.; Ottolenghi, M. *J. Phys. Chem.* **1970**, *74*, 2041–2042.

(72) Figure adaptée de : http://www.b2b.invitrogen.com/site/us/en/home/References/Molecular-Probes-The-Handbook/Probes-for-Lipids-and-Membranes/Fatty-Acid-Analogs-and-Phospholipids.html (accessed Jul 4, 2014).

(73) Parson, W. W. In *Modern Optical Spectroscopy*; Springer Berlin Heidelberg, 2007; pp. 307–334.

(74) Kypr, J.; Kejnovská, I.; Renčiuk, D.; Vorlíčková, M. *Nucleic Acids Res.* **2009**, *37*, 1713–1725.

(75) Ranjbar, B.; Gill, P. *Chem. Biol. Drug Des.* **2009**, *74*, 101–120.

(76) Chang, Y.-M.; Chen, C. K.-M.; Hou, M.-H. *Int. J. Mol. Sci.* **2012**, *13*, 3394–3413.

CHAPITRE 3 :

FORMATION DE POLYMERES DYNAMIQUES COVALENTS

POUR LA TRANSFECTION

INTRODUCTION – RAPPELS BIBLIOGRAPHIQUES

Les thérapies géniques telles que les technologies antisens[1,2] et d'extinction de gènes [3–5] ouvrent des perspectives thérapeutiques prometteuses pour diverses maladies comme les cancers. Toutefois, les oligonucléotides synthétiques ne peuvent pas traverser efficacement les membranes biologiques en raison des charges négatives qu'ils portent et de la répulsion électrostatique avec les constituants des membranes cellulaires qui en résulte. De plus, ils sont rapidement dégradés par les nucléases endogènes, avec un temps de demi-vie de l'ordre de quelques minutes dans le sérum humain.[6,7] Bien que les vecteurs viraux soient plus efficaces, leurs effets potentiellement inflammatoires, immunogènes et mutagènes limitent de manière importante leur usage et rend nécessaire le développement d'alternatives non virales.[7,8] C'est pourquoi développer un système capable de complexer, transporter et libérer des oligonucléotides thérapeutiques dans les cellules, qui soit à la fois efficace et sûr, est un point clé en thérapie génique.[9]

Les polymères[10–13] et dendrimères cationiques[14] sont bien connus pour leur application en tant que vecteurs non viraux de gènes. En effet, ils sont capables de former des complexes stables avec les oligonucléotides grâce aux multiples interactions électrostatiques formées avec ces derniers. Malgré tout, bien que ces vecteurs artificiels puissent être moins chers et plus faciles à produire, la plupart des études cliniques utilisent encore des vecteurs viraux pour la transfection en raison de la plus grande efficacité de ces derniers.[15] Plusieurs raisons peuvent expliquer la moindre efficacité des vecteurs non viraux. Une première raison peut être l'association permanente des vecteurs non viraux avec les oligonucléotides, ce qui limite la libération intracellulaire des oligonucléotides thérapeutiques. En effet, les polyplexes sont souvent piégés dans les endosomes qui constituent un des principaux obstacles pour les systèmes non viraux de délivrance de gènes.[16,17] Pour cette raison, des matériaux 'intelligents' capables de promouvoir la libération de l'endosome *via* l'effet de « l'éponge à proton »,[18] ou par la dégradation chimique[19,20–26,27–33] ou enzymatique[20,34,35] du vecteur artificiel représentent des solutions de grand intérêt. Une autre caractéristique des polymères et dendrimères capable d'expliquer leur moindre efficacité peut également être leur haut poids moléculaire qui limite leur élimination du corps et peut conduire à des effets toxiques,[36] tandis que leur caractère cationique et leur hydrophilicité conduisent à une association avec les protéines du plasma, d'où un temps de circulation dans le sang plus court.

Puisque la pégylation constitue une stratégie bien développée pour augmenter les stabilité, biocompatibilité et temps de circulation des biomolécules *in vivo*,[37] plusieurs polymères hybrides[38] combinant des moitiés polyamines cationiques et éthylène glycols ont récemment été préparés et testés en tant que cargos de gènes.[39–45] Par exemple, la pégylation des polyéthylènimines (PEI) induit une réduction de la charge superficielle du polymère ainsi qu'une diminution de l'interaction avec les composés sanguins, ce qui réduit donc la cytotoxicité et augmente le temps de circulation dans le sang.[46] Les systèmes hybrides dégradables représentent donc des matériaux prometteurs pour l'application réussie de vecteurs artificiels de délivrance de gènes dans la phase clinique.[27,32,47]

Les polymères covalents dynamiques – des polymères qui s'auto-assemblent à partir de leurs monomères par multiples liaisons covalentes réversibles[48] – sont récemment apparus comme une classe de matériaux dynamiques capables de s'adapter par une réorganisation constitutionnelle.[49–53] Ils peuvent également être considérés comme des matériaux « intelligents » pour la reconnaissance de biomolécules.[54] En effet, l'incorporation répétée de multiples monomères fonctionnalisés en une seule structure peut conduire à la présentation multivalente de plusieurs unités de reconnaissance à l'état polymérique, alors que les monomères seuls ne montreront qu'une faible affinité pour la cible à cause de leur valence réduite.[55–57] Les polymères covalents dynamiques peuvent donc servir de vecteurs multivalents « intelligents » dont la dégradation, suite à un changement physico-chimique, entraîne la libération de l'agent thérapeutique et génère des composés de plus petite taille qui peuvent donc être plus facilement éliminés du corps.[19]

Les polymères covalents dynamiques basés sur les poly-acylhydrazones sont particulièrement intéressants pour ce type d'application car l'hydrolyse des acylhydrazones a lieu dans des conditions légèrement acides, à un pH comparable à celui des endosomes tardifs (5 – 5,5). De plus, les poly-acylhydrazones sont formées par condensation de partenaires complémentaires – bisaldéhydes et bishydrazides – qui permettent la conception de matériaux hybrides. Dans le contexte de la transfection génique, des lipides cationiques[33] et des polymères conjugués qui incorporent une liaison acylhydrazone ont récemment été reportés.[47,58–60] Toutefois, la conception de polymères covalents dynamiques basés sur les poly-acylhydrazones pour la complexation d'ADN n'a à l'heure actuelle pas encore été explorée. Il est important de mentionner que les polymères supramoléculaires, dynamiques par nature, ont été impliqués dans des essais de transfection avec succès. [61]

Dans ce chapitre est reportée la préparation de polymères poly-acylhydrazones hydrides qui alternent au sein de leur chaine principale : des résidus cationiques pour interagir avec des oligonucléotides, et des groupements polyéthylènoxides neutres. Il est également démontré que ces matériaux sont dégradables à pH acides et complexent efficacement l'ADN de thymus de veau et un ADN plasmidique.

Les polymères poly-acylhydrazones hybrides sont conçus pour combiner des résidus cationiques et des groupements polyéthylènoxides neutres. Comme résidus cationiques, nous avons choisi d'utiliser la triamine diéthylène ou le guanidinium. Le choix de la triamine diéthylène a été guidé par l'observation que des clusters monodisperses[62–64] et des systèmes polydisperses[65] caractérisés par des motifs multiples contenants trois amines secondaires ont montrés qu'ils pouvaient complexer efficacement un ADN plasmidique. De plus, la présentation de multiples groupements diaminoéthanes comme chaînes latérales sur un polymère induit une déstabilisation de la membrane endosomale comme dans le cas des polyéthylèneimines (PEI). Les composés guanidiniums sont pour leur part bien connus pour former des ponts salins avec les oxoanions. De nombreux systèmes contenants des guanidiniums, tels que des oligomères, polymères et dendrimères[24,33,66–89] ainsi que des lipides cationiques,[90–93] ont été développés pour des applications en délivrance de médicaments et de gènes. Le triéthylèneglycol a été choisi comme groupement neutre parce qu'il a une taille similaire à celles des résidus cationiques. Tous les monomères ont été fonctionnalisés en position terminale par des aldéhydes et / ou des hydrazides de manière à ce que l'auto-assemblage puisse se faire *via* une étape de polycondensation afin d'obtenir des polymères acylhydrazones hybrides dynamiques alternant des résidus cationiques et des groupements triéthylèneglycol neutres (Figure 3.).

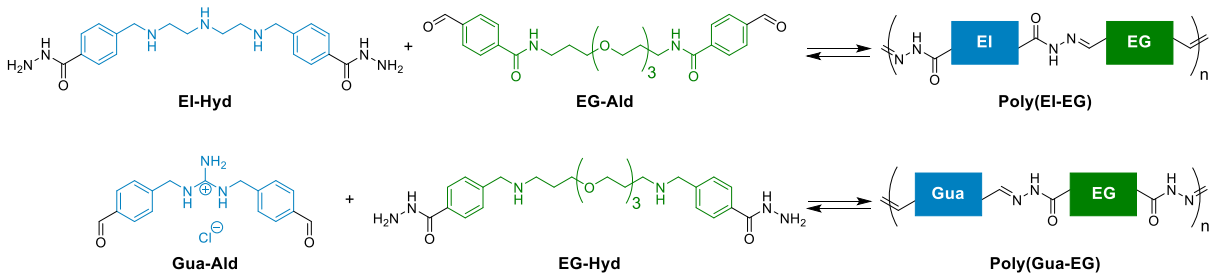


Figure 3.1 : Polymères hybrides dynamiques **Poly(EI-EG)** et **Poly(Gua-EG)** issu de l'auto-assemblage des monomères cationiques **EI-Hyd** et **Gua-Ald** avec, respectivement, les monomères éthylèneglycol **EG-Ald** et **EG-Hyd** *via* la formation d'acylhydrazones.

I. SYNTHESE DES MONOMERES

A. SYNTHESE DES MONOMERES BISHYDRAZIDES

1. Synthèse du monomère éthylèneimine bishydrazide **EI-Hyd**

Dans un premier temps, nous avons envisagé de préparer le monomère éthylèneimine bishydrazide via un couplage peptidique entre une bisamine et un bisacide carboxylique fonctionnalisé en bout de chaîne par un hydrazide préalablement protégé (Figure 3.2). L'acide adipique a été sélectionné car la synthèse du synthon **3** est décrite dans la littérature.[94]

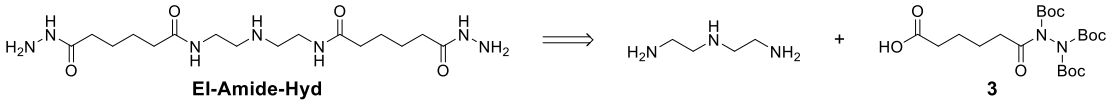


Figure 3.2 : Rétrosynthèse du premier monomère **EI-Amide-Hyd**.

La première partie de la synthèse consiste à obtenir le composé **3** à partir d'acide adipique (Schéma 3.), pour pouvoir ensuite le coupler avec la diéthylène triamine. Le composé **3** est obtenu en quatre étapes.[94] La première étape est la monoprotection de l'acide adipique sous forme d'ester benzylique en présence d'anhydride acétique et d'alcool benzylique. Le composé **4** peut ensuite réagir avec le *tert*-butylcarbazate dans des conditions de couplage classique et une étape de protection totale de l'hydrazide par des groupements *tert*-butoxycarbonyle conduit au composé **6**. Enfin, une dernière étape d'hydrogénation permet de déprotéger l'ester benzylique et d'obtenir le composé **3** avec un rendement global de 27% sur les quatre étapes.

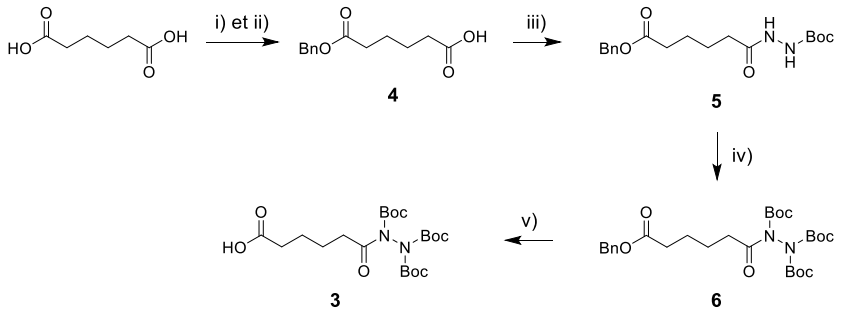


Schéma 3.1 : Synthèse du composé **3** à partir d'acide adipique. Réactifs et conditions: i) Ac_2O, ii) DMAP, pyridine, BnOH / CH_2Cl_2, TA, 36% sur i et ii ; iii) *tert*-butylcarbazate, EDC.HCl, CH_2Cl_2, réaction quantitative ; iv) Boc_2O / DMAP, Et_3N / CH_2Cl_2, 0°C à TA, 71% ; v) H_2, Pd / C 10%, MeOH, TA, 94%.

Dans le but de tester la réactivité de la diéthylène triamine avec un acide carboxylique, la Boc-glycine a été choisie comme acide modèle. Différentes conditions classiques de

couplage (HBTU / DIPEA, EDC.HCl, BOP / DIPEA, DCC / DMAP) ont été testées, mais malheureusement, elles n'ont pas permis de mettre en évidence la faisabilité du couplage sur la diéthylène triamine. Le constat est identique lorsque le synthon **3** a été utilisé. Les composés **7** et **8** n'ont pu être ni détectés, ni isolés (Schéma 3.2).

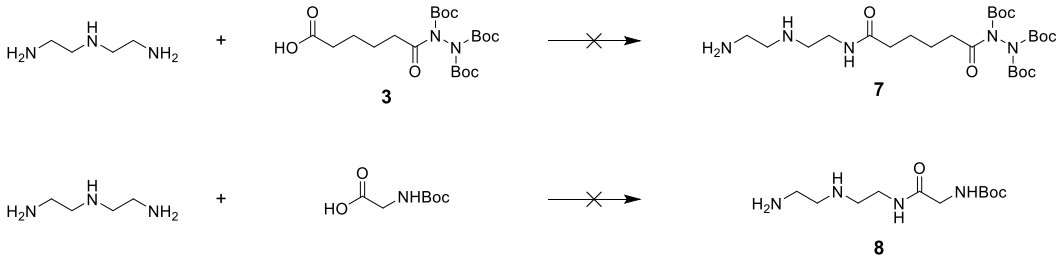


Schéma 3.2 : Tentatives de couplage avec la diéthylène triamine. Différentes conditions testées: HBTU / DIPEA ; EDC.HCl ; BOP / DIPEA ; DCC / DMAP.

La première stratégie de synthèse ne donnant pas de résultats satisfaisants, l'introduction des fonctions hydrazides a été envisagée *via* une addition de Michael entre la diéthylène triamine et le composé **9** (Figure 3.3).

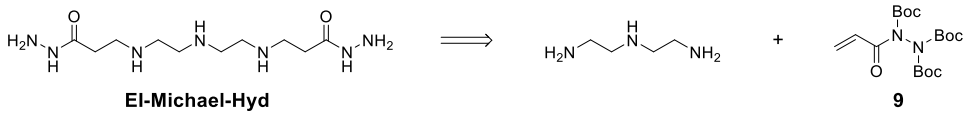


Figure 3.3 : Rétrosynthèse du monomère **EI-Michael-Hyd**.

Avant de commencer la synthèse de l'accepteur de Michael **9**, une réaction modèle de l'addition de Michael a été développée avec l'acrylate de *tert*-butyle et la diéthylène triamine (Schéma 3.3). Cette réaction effectuée avec 1 équivalent d'acrylate permet d'obtenir le composé **10** avec un rendement de 19%. Lorsque 2 équivalents d'acrylate sont utilisés, une double addition sur la diéthylène triamine permet d'obtenir le composé **11** avec un rendement de 10%. La diéthylène triamine étant une polyamine susceptible de réagir *via* ses fonctions amines primaires et secondaires sur l'accepteur **9**, des additions de Michael multiples peuvent avoir lieu, ce qui rend les purifications difficiles et induit des rendements faibles.

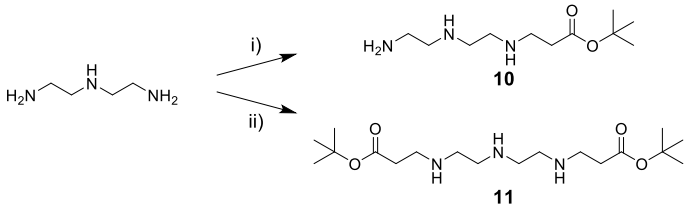


Schéma 3.3 : Addition de Michael modèle entre l'acrylate de *tert*-butyle et la diéthylène triamine. Réactifs et conditions: i) acrylate de *tert*-butyle 1,0 équivalent, MeOH, TA, 19%; ii) acrylate de *tert*-butyle 2,0 équivalents, MeOH, TA, 10%.

L'addition de Michael étant possible avec la diéthylène triamine, la synthèse de l'acrylamide **9** a été effectuée à partir de l'acide acrylique en deux étapes (Schéma 3.4). La première étape consiste à introduire la fonction hydrazide par réaction avec le *tert*-butylcarbazate comme précédemment en présence d'hydrochlorate d'EDC. L'hydrazide monoprotégée du composé **12** obtenu est ensuite totalement protégée pour donner le composé **9**. Le rendement global sur ces deux étapes est de 38%.

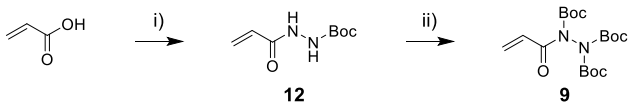


Schéma 3.4 : Synthèse de l'accepteur de Michael **9**. Réactifs et conditions: i) *tert*-butylcarbazate, EDC.HCl, CH_2Cl_2, TA, 49%; ii) Boc_2O, CH_2Cl_2, Et_3N, DMAP, 0°C à TA, 76%.

L'addition de Michael de la diéthylène triamine sur l'accepteur **9** dans les conditions de réactions mises au point avec l'acrylate de *tert*-butyle n'a pas permis d'isoler le produit **13** souhaité (Schéma 3.5). La réaction a donc été testée dans les mêmes conditions avec l'acrylamide mono-protégée **12** : le produit **14** a été détecté en LC-MS dans le mélange réactionnel avec les autres produits d'addition, mais n'a pas été isolé.

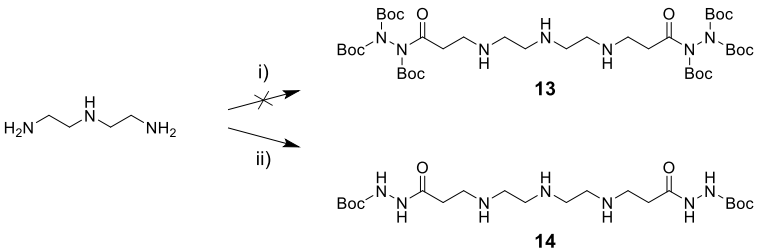


Schéma 3.5 : Addition de Michael de la diéthylène triamine sur les composés **9** et **12**. Réactifs et conditions: i) composé **9**, MeOH, TA; ii) composé **12**, MeOH, TA.

L'addition de Michael ne fonctionnant pas avec le composé **9**, la voie de synthèse est modifiée pour introduire les terminaisons hydrazides après avoir réalisé l'addition de Michael (Schéma 3.6). Pour se faire, la diéthylène triamine réagit sur l'acrylate de méthyle et le composé **15** obtenu est ensuite mis en présence d'hydrazine pour obtenir le produit **EI-Michael-Hyd** souhaité avec un rendement global de seulement 7%. Ce faible rendement est dû aux multiples additions de Michael qui ont toujours lieu au cours de la première étape, générant de nombreux produits qu'il est par ailleurs difficile de séparer.

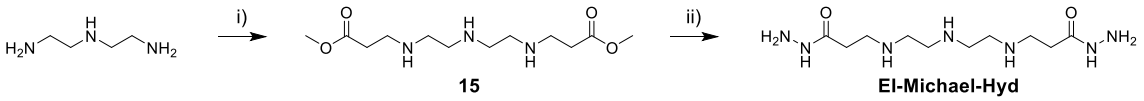


Schéma 3.6 : Synthèse finale du monomère **EI-Michael-Hyd**. Réactifs et conditions: i) acrylate de méthyle, MeOH, TA, 11% ; ii) monohydrate d'hydrazine, reflux, 65%.

Cette dernière méthode de synthèse, bien qu'elle permette d'obtenir le produit souhaité, n'est pas viable pour une montée en échelle, le rendement étant trop faible et les purifications difficiles. Une quatrième stratégie a donc été envisagée (Figure 3.4). Il a été choisi de procéder à la ligation de l'éthylèneimine centrale avec les terminaisons hydrazides par amination réductrice. Pour faciliter à la fois la détection et la purification des produits obtenus, l'espaceur introduit est un groupement benzyle.

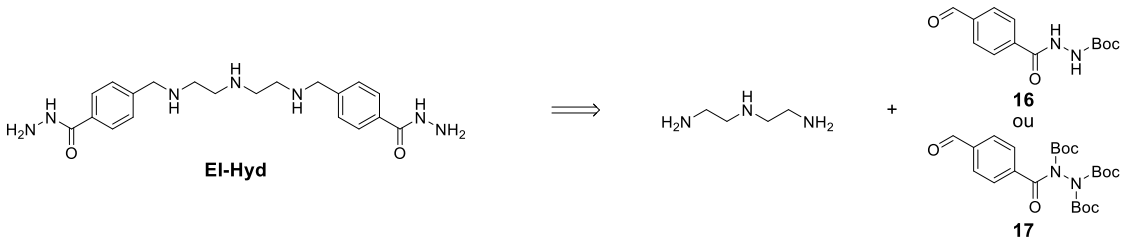


Figure 3.4 : Rétrosynthèse du monomère **EI-Hyd**.

La première étape de cette synthèse consiste à faire réagir l'acide 4-formylbenzoïque avec le chlorure de thionyle et le *tert*-butylcarbazate pour obtenir l'aldéhyde **16** avec un rendement de 55% (Schéma 3.8). Le composé **16** est ensuite engagé dans une réaction d'amination réductrice avec la diéthylène triamine[95,96] pour obtenir le composé N-Boc-bishydrazide **18**. La déprotection des amines par un mélange TFA/TIS/eau (95/2,5/2,5) donne le monomère EthylèneImine bisHydrazide (**EI-Hyd**) souhaité sans difficultés particulières.

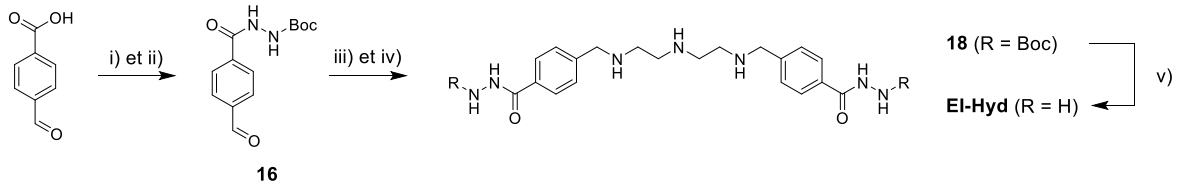


Schéma 3.7 : Synthèse du monomère bishydrazide **EI-Hyd**. Réactifs et conditions : i) $SOCl_2$, toluène, reflux, 12 h ; ii) *t*-butylcarbazate, Et_3N, CH_2Cl_2, 0°C à TA, 48h ; iii) diéthylène triamine, MeOH, 0°C à TA, 48h ; iv) $NaBH_4$, MeOH, 0°C à TA, 24h ; v) TFA/TIS/H_2O (95/2,5/2,5), TA, 48h. Rendements : composé **16** 55%, composé **18** 26%, **EI-Hyd** 80%.

Il faut toutefois préciser que lors de la déprotection finale, une certaine quantité d'amine est protonée par le TFA mais il n'est pas possible de déterminer par simple calcul la quantité finale de sel après lyophilisation. Le rendement est donc déterminé par dosage RMN ^{1}H des produits avec une solution de *tert*-butanol comme étalon interne.

2. Synthèse du monomère éthylèneglycol bishydrazide **EG-Hyd**

Les deux monomères bishydrazides ne diffèrent dans leur structure que par la nature du chaînon diéthylène triamine ou éthylène glycol. Il est donc possible d'utiliser une voie de synthèse similaire pour le monomère **EG-Hyd** en remplaçant la diéthylène triamine par le chaînon éthylène glycol lors de l'amination réductrice.

La première étape de la synthèse est donc commune aux deux monomères et le composé **16** est ensuite engagé dans une réaction d'amination réductrice avec la 4,7,10-trioxa-1,13-tridécanediamine pour obtenir le composé N-Boc-bishydrazide **19** (Schéma 3.8). La déprotection des amines par un mélange TFA/TIS/eau (95/2,5/2,5) donne le monomère Ethylène Glycol bisHydrazide (**EG-Hyd**) souhaité. La détermination du rendement de déprotection finale se fait comme précédemment par dosage RMN ^{1}H.

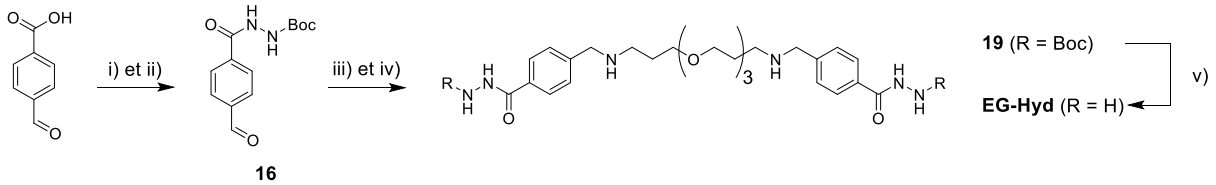


Schéma 3.8 : Synthèse du monomère éthylène glycol bishydrazide **EG-Hyd**. Réactifs et conditions : i) $SOCl_2$, toluène, reflux, 12 h ; ii) *t*-butylcarbazate, Et_3N, CH_2Cl_2, 0°C à TA, 48h ; iii) 4,7,10-trioxa-1,13-tridécanediamine, MeOH, 0°C à TA, 48h ; iv) $NaBH_4$, MeOH, 0°C à TA, 24h ; v) TFA/TIS/H_2O (95/2,5/2,5), TA, 48h. Rendements : composé **16** 55%, composé **19** 22%, **EG-Hyd** 55%.

B. SYNTHESE DES MONOMERES BISALDEHYDES

1. Synthèse du monomère éthylène glycol bisaldéhyde **EG-Ald**

Le monomère Ethylène Glycol bisAldéhyde **EG-Ald** a été préparé par couplage de la 4,7,10-trioxa-1,13-tridécanediamine et de l'acide 4-formylbenzoïque en présence d'EDC et d'HOBt : le produit souhaité est obtenu avec un rendement de 81% (Schéma 3.9).

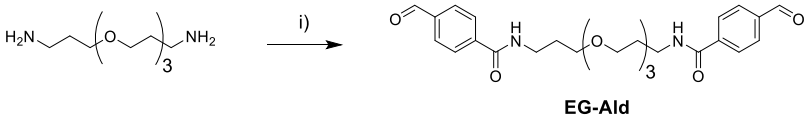


Schéma 3.9 : Synthèse du monomère éthylène glycol bisaldéhyde **EG-Ald**. Réactifs et conditions : i) acide 4-formylbenzoïque, EDC, HOBt, Et_3N, CH_2Cl_2, 0°C à TA, 48h, 81%.

2. Synthèse du monomère guanidinium bisaldéhyde **Gua-Ald**

Le monomère Guanidinium bisAldéhyde **Gua-Ald** est construit depuis sa terminaison aldéhyde vers le groupement guanidinium central. Pour se faire, l'introduction du guanidinium a tout d'abord été envisagée en faisant réagir deux équivalents d'amine sur le bromure de cyanogène (Schéma 3.10), de la même manière que ce qui avait été précédemment effectué pour le guanidinium bisphényle **GuaBiPhe** (*cf.* chapitre 2). Le monomère **Gua-Ald** a donc été préparé en trois étapes à partir du cyanobenzaldéhyde (Schéma 3.10). Les deux premières étapes consistent à protéger l'aldéhyde avec l'orthoformate de triméthyle, puis à réduire le nitrile avec $LiAlH_4$ pour obtenir l'amine **20**.[97,98] Lors de la dernière étape, l'amine **20** réagit sur le bromure de cyanogène, puis la déprotection des aldéhydes est effectuée *in situ* avec un traitement à l'acide chlorhydrique avant la purification du produit final.

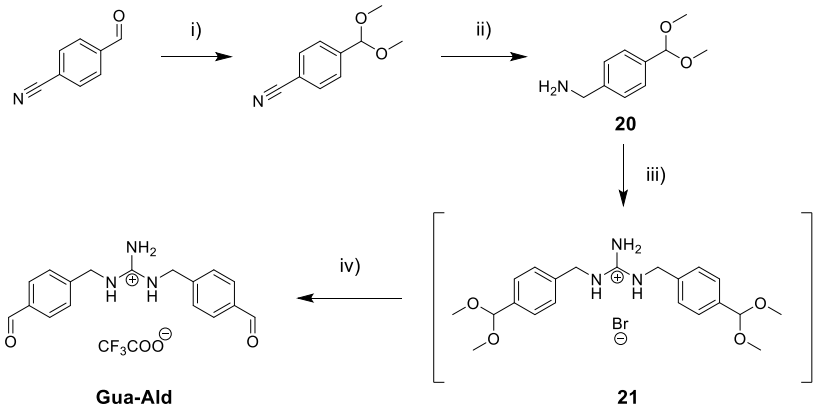


Schéma 3.10 : Synthèse du monomère guanidinium bisaldéhyde (contre-ion CF_3COO^-) *via* la voie 'bromure de cyanogène'. Réactifs et conditions : i) Orthoformate de triméthyle, HCl_{conc}, MeOH, 45°C, 3h30, 98% ; ii) $LiAlH_4$, THF, 0°C à TA, 12h, 93% ; iii) BrCN, DIPEA, CH_2Cl_2, reflux, 48h ; iv) HCl 1 M jusqu'à pH 1, suivi d'une HPLC semi-préparative, 32% (sur iii et iv).

L'amine **20** a été obtenue avec un rendement global de 91% sur les deux étapes. Par contre, la formation du guanidinium avec l'amine **20** et le bromure de cyanogène, non décrite dans la littérature, a été plus laborieuse. En effet, lors de cette étape, il a été observé que bien que les deux réactifs (amine **20** et bromure de cyanogène) aient été introduits en quantité stœchiométrique pour obtenir le guanidinium disubstitué, il restait systématiquement de l'amine **20** résiduelle dans le mélange réactionnel, en quantité plus ou moins importante.

Lors des premiers essais, le guanidinium cétal **21** a été isolé, mais n'a généralement pas été obtenu pur après une purification sur colonne de gel de silice. En effet, l'amine résiduelle **20** peut être présente avec le cétal **21** en sortie de colonne. D'autre part, une purification par chromatographie sur colonne de gel de silice après déprotection des aldéhydes, dans les conditions testées (éluant dichlorométhane – méthanol) n'a pas non plus

permis de séparer correctement le guanidinium bisaldéhyde **Gua-Ald** de l'amine résiduelle lorsque cette dernière était présente dans le mélange déposé sur la colonne.

Différents tests ont été effectués pour trouver une manière de séparer correctement l'amine résiduelle **20** de l'acétal **21** isolé. Un traitement aqueux a d'abord été introduit en fin d'étape d'addition sur le bromure de cyanogène, mais n'a pas été concluant, une partie de l'amine **20** passant dans la phase aqueuse à pH 9 et le guanidinium acétal **21** restant dans la phase organique à pH 14 (sous forme guanidine). Il a également été envisagé de piéger la fonction amine primaire du composé **20** en présence d'anhydride benzoïque **22** afin d'obtenir le composé **23** (Schéma 3.11). Le milieu est extrait par de l'eau, et le composé **23**, plus aromatique que l'amine **20**, reste dans la phase organique. Cette méthode, facile à mettre en œuvre, permet donc de séparer efficacement l'amine **20** du produit **Gua-Ald** (obtenu après acidification de la phase aqueuse d'extraction pour compléter la déprotection des acétals déjà initiée lors de l'extraction et observée par RMN ^{1}H de la phase aqueuse).

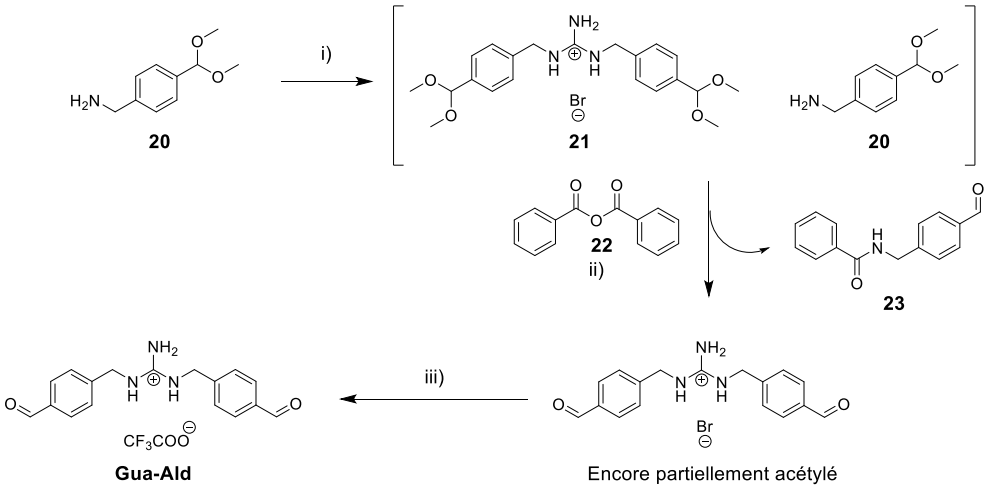


Schéma 3.11 : Procédure envisagée pour neutraliser l'amine **20** résiduelle dans la voie de synthèse du produit **Gua-Ald** *via* le bromure de cyanogène. Réactifs et conditions : i) BrCN, DIPEA, CH_2Cl_2, reflux, 48h ; ii) anhydride benzoïque **22**, CH_2Cl_2, TA, 12h ; iii) HCl 0,1 M, TA, suivi d'une HPLC semi-préparative. Rendement : 23% isolé.

Malheureusement, le produit **Gua-Ald** obtenu avec cette méthode n'était pas tout à fait pur et sa purification par chromatographie sur colonne de gel de silice était *a priori* rendue difficile par la grande polarité du produit. Cette dernière propriété nous a conduits à envisager l'HPLC semi-préparative comme méthode de purification avec laquelle l'amine **20** s'est révélée facilement séparable du produit **Gua-Ald**.

La procédure finale a donc pu être simplifiée : le guanidinium acétal **21** n'est pas isolé et il n'est pas nécessaire de neutraliser l'amine **20** par l'anhydride benzoïque **22**. Le produit

Gua-Ald a ainsi pu être obtenu pur avec un rendement de 32% au cours de l'étape d'addition sur le bromure de cyanogène.

Toutefois, bien que cette première voie de synthèse ait permis d'obtenir le produit **Gua-Ald** voulu, elle ne peut pas être envisagée dans le cadre d'une montée en échelle, pour obtenir suffisamment de produit dans le but de former un polymère avec le produit **EG-Hyd**. L'étape limitante pour cette potentielle montée en échelle se trouve être l'addition de l'amine **20** sur le bromure de cyanogène. D'une part, le suivi réactionnel de cette étape est difficile et le rendement est plutôt faible. D'autre part, la montée en échelle sera difficile à mettre en œuvre s'il reste nécessaire de purifier le produit par HPLC semi-préparative. Pour essayer de contourner ces problèmes, une deuxième voie de synthèse faisant appel à des thiourées et décrite dans la littérature a été testée (Schéma 3.12).[99] A partir de l'amine **20**, le guanidinium bisaldéhyde **Gua-Ald** est obtenu en trois étapes : i) la réaction de l'amine **20** avec l'isothiocyanate d'éthoxycarbonyle conduit à la formation de la thiourée **24** ; ii) cette thiourée est couplée avec un deuxième équivalent d'amine **20** en présence d'EDC pour obtenir la guanidine protégée **25** ; et iii) la déprotection des acétals et du carbamate avec une solution d'acide chlorhydrique permet enfin d'obtenir le guanidinium bisaldéhyde **Gua-Ald** souhaité avec un rendement global, après optimisation, de 51% sur ces trois étapes.

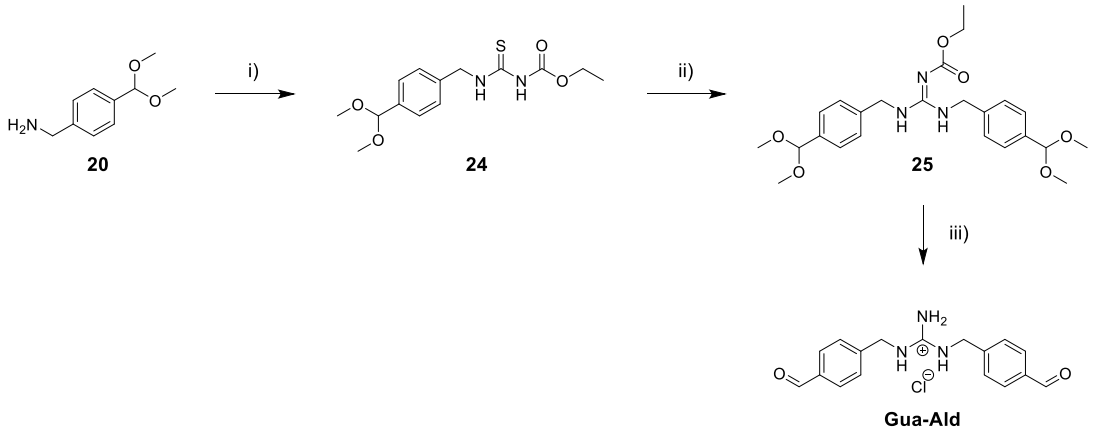


Schéma 3.12 : Synthèse du monomère guanidinium bisaldéhyde **Gua-Ald** *via* la voie thiourée. Réactifs et conditions : i) isothiocyanate d'éthoxycarbonyle, Et_3N, CH_2Cl_2, TA, 2h, 90% ; ii) amine **20**, EDC, Et_3N, DMF, 60°C, 12h, 84% ; iii) HCl 1 M, reflux, 13h30, 86%.

Quelques modifications mineures ont été apportées pour adapter les conditions de synthèse décrites dans la littérature aux propriétés des produits. Ainsi, de la triéthylamine a été ajoutée au mélange réactionnel pour former la thiourée **24** sans déprotéger les fonctions acétals : cet ajout de base n'est pas négligeable en termes de rendement puisque pour une même échelle de réaction, celui-ci passe de 70% à 90% avec la triéthylamine. Lors de la

formation de la guanidine protégée **25**, un chauffage à 60°C a été effectué pour accélérer la cinétique de réaction. La déprotection du carbamate qui s'effectue avec le bromotriméthylsilane dans la publication de référence, a été remplacée par une déprotection par une solution d'acide chlorhydrique, essentiellement pour des raisons de simplicité.

Avec cette nouvelle voie de synthèse, les réactions sont complètes et aucune trace d'amine **20** résiduelle n'est présente dans les produits **24** et **25** lors de leur purification par chromatographie sur colonne de gel de silice. A l'issue de la dernière étape de déprotection, une purification du produit **Gua-Ald** paraissait nécessaire pour enlever les sels et les impuretés pouvant être présents. Une méthode de purification par chromatographie flash sur colonne de gel de silice a donc été mise au point avec un éluant acétonitrile – eau. Tout en assurant la migration d'un produit aussi polaire que le monomère **Gua-Ald**, cet éluant plutôt inhabituel permet également de ne pas protéger de nouveau les aldéhydes lors de la purification, contrairement à un éluant dichlorométhane – méthanol.

Après optimisation, cette synthèse du monomère **Gua-Ald** *via* la voie 'thiourée' a permis une monté en échelle de la centaine de milligrammes à quelques grammes sans difficultés particulières.

II. REACTIONS DE POLYMERISATION ET CARACTERISATION

A. REACTION MODELE

Pour vérifier la chimiosélectivité de la réaction de condensation, une réaction modèle a été effectuée en mélangeant, dans le méthanol et à température ambiante le composé monohydrazide éthylèneimine **26** (issu de la déprotection du monohydrazide protégé isolé comme produit secondaire lors de l'amination réductrice entre le composé **16** et l'éthylènediamine, Schéma 3.7) avec un équivalent de benzaldéhyde (Schéma 3.13). Une analyse HPLC après une heure montre une conversion complète et la formation d'un seul composé dont la masse correspond à un produit de condensation, probablement l'acylhydrazone **27** (Figure 3.5).

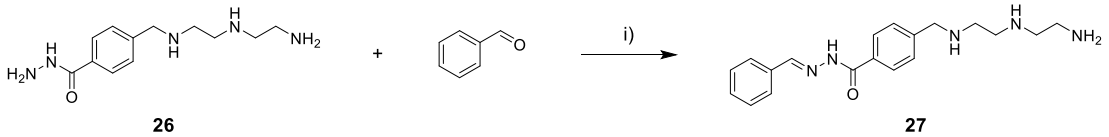


Schéma 3.13 : Ligation entre le composé monohydrazide **26** et le benzaldéhyde. Conditions et réactifs: i) 10 mM dans MeOH, TA, 1h.

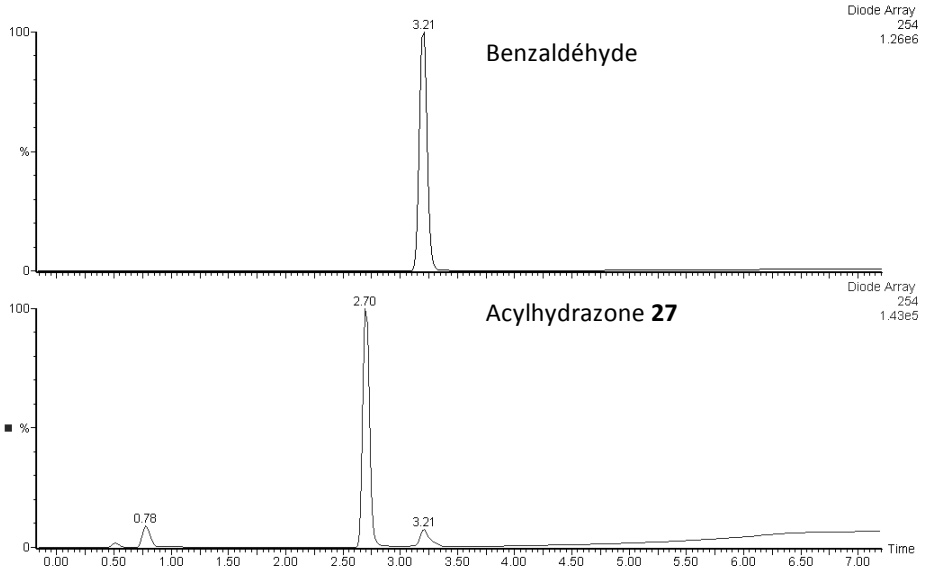


Figure 3.5 : Chromatogrammes HPLC du benzaldéhyde (haut) et du produit de condensation obtenu après une heure de réaction avec le composé modèle **26** et identifié par spectrométrie de masse (m/z = 340,0) (bas). Méthode I (HPLC 2), suivi à 254 nm.

Une deuxième réaction modèle a été faite avec le monomère **Gua-Ald** et la benzhydrazide dans différents solvants pour déterminer le plus approprié pour effectuer la polycondensation (Tableau 3.). Le composé **28** (Schéma 3.14) est détecté dans tous les mélanges réactionnels. Il se trouve que parmi les solvants testés, le méthanol permet de conjuguer cinétique et solubilisation suffisante des produits. D'autres réactions modèles entre le monomère **EG-Ald** et la benzhydrazide ainsi qu'entre les monomères bishydrazides et le benzaldéhyde confirment que le méthanol est le solvant le plus approprié pour effectuer les différentes polycondensations.

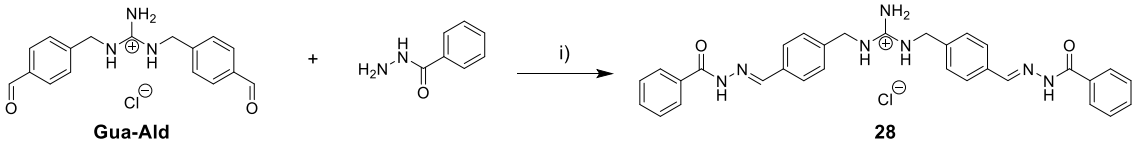


Schéma 3.14 : Ligation entre le composé bisaldéhyde **Gua-Ald** (à 10 mM) et la benzhydrazide (2éq.). Effet du solvant détaillé dans le Tableau 3..

Tableau 3.1 : Effet du solvant sur la condensation modèle entre le monomère **Gua-Ald** à 10 mM et 2 équivalents de benzhydrazide.

Solvant	Observations
Méthanol	Plus que quelques traces de **Gua-Ald** après 6 jours à température ambiante. Apparition d'un précipité au bout de 6 jours.
Eau 1% TFA	Apparition d'un précipité immédiatement, pas de monomère **Gua-Ald** dans le surnageant.
Tampon acétate d'ammonium pH 4	Apparition d'un précipité immédiatement, disparition du monomère **Gua-Ald** du surnageant après 7 heures de réaction et pas d'évolution de la composition du surnageant au-delà de ce temps.
Tampon acétate de sodium pH 4	
Acétonitrile	Apparition d'un précipité au cours des premières 24 heures, toujours des traces du monomère **Gua-Ald** au bout de 10 jours.
DMF	Solution homogène, disparition du monomère **Gua-Ald** au bout de 16 jours.
DMSO	

B. REACTION DE POLYCONDENSATION

1. Suivi par RMN ^{1}H

La réaction de polycondensation a ensuite été effectuée en mélangeant les monomères bisaldéhyde et bishydrazide en quantités stœchiométriques à 200 mM dans le méthanol. Des analyses ont été réalisées après évaporation du solvant et dilution dans le DMSO deutéré – ce solvant étant le meilleur choix pour toutes les combinaisons de monomères sur une large gamme de concentrations –, par spectroscopie RMN ^{1}H (Figure 3.6) et DOSY. Les données RMN ^{1}H montrent la disparition totale de l'aldéhyde de départ (δ = 10 ppm, C*H*O) et l'apparition de signaux caractéristiques des acylhydrazones à δ = 12 ppm (C(O)N*H*N) et δ = 8,5 ppm (C(O)NHNC*H*).

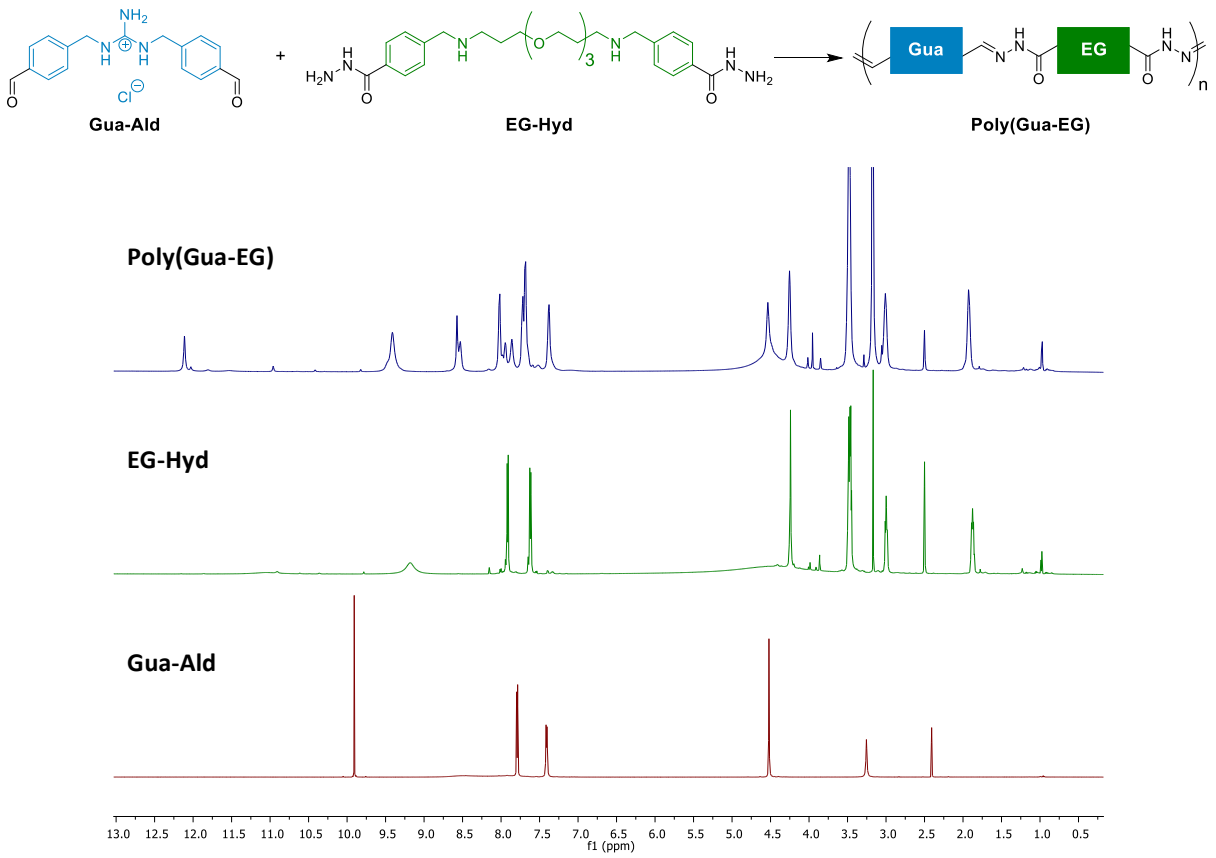


Figure 3.6 : Suivi RMN ^{1}H dans le DMSO deutéré de la réaction de condensation entre les monomères **EG-Hyd** et **Gua-Ald** dans le méthanol à 200 mM après 1 heure.

La formation d'acylhydrazones ayant été démontrée dans ces conditions, l'effet de la concentration sur le résultat de la polycondensation a été testé. En effet, la concentration est connue pour affecter l'équilibre macrocycle – polymère pour des réactions de polymérisation contrôlées thermodynamiquement : les polymères sont favorisés à des concentrations dépassant la concentration effective des monomères tandis que les macrocycles sont prépondérants à des concentrations plus faibles.[100,101] Le caractère flexible de nos monomères devrait contribuer à diminuer leur concentration effective. Les données RMN ^{1}H montre un élargissement des signaux lorsque la concentration augmente (Figure 3.7). A faible concentration (10 mM), le spectre RMN ^{1}H du **Poly(EI-EG)** comporte de multiples signaux d'acylhydrazones, ce qui peut indiquer la formation de différents produits.

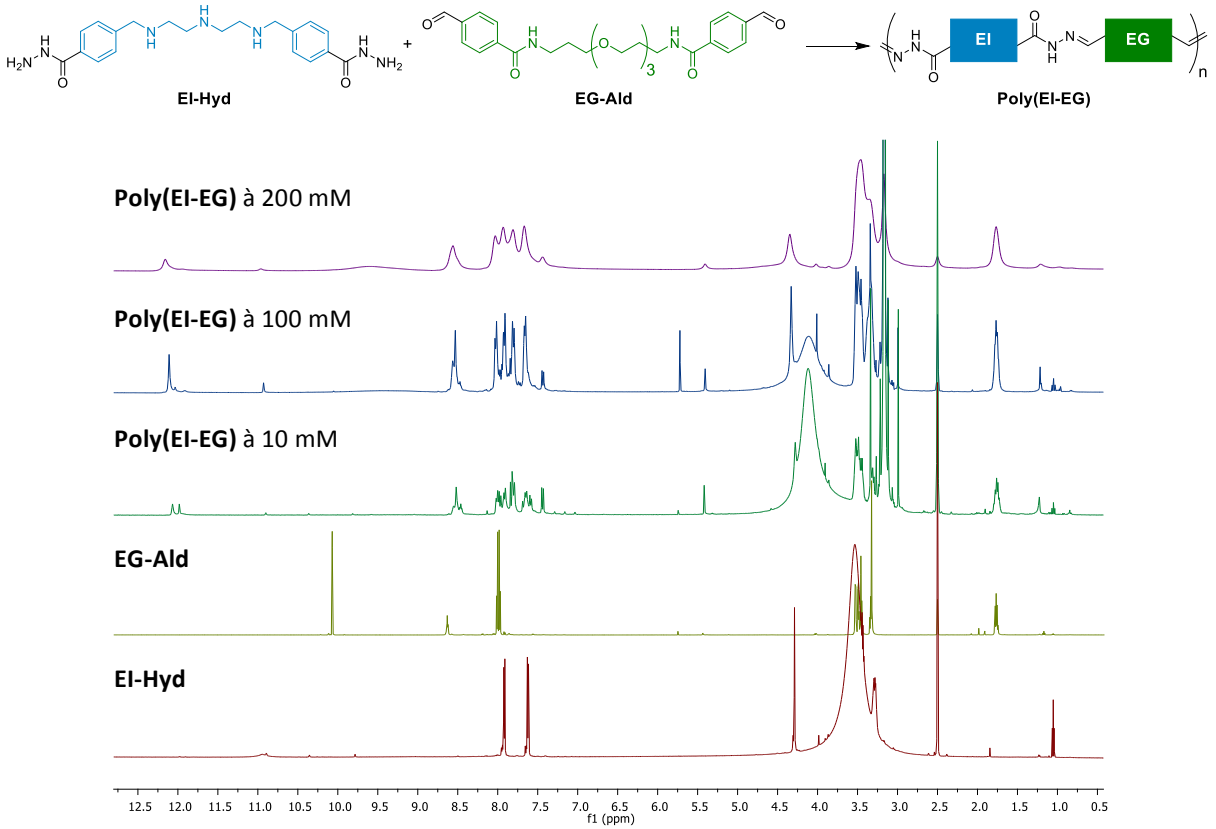


Figure 3.7 : Spectres RMN ^{1}H dans le DMSO deutéré des monomères **EG-Hyd**, **Gua-Ald** et des réactions de polycondensation pour la formation de **Poly(EI-EG)** à différentes concentrations.

2. Caractérisation par DOSY

La RMN DOSY (Diffusion Order SpectroscopY) permet de déterminer le coefficient de diffusion d'espèces chimiques en solution. Le coefficient de diffusion d'une molécule dépendant de sa masse molaire et de propriétés hydrodynamiques comme sa taille, sa forme ou sa charge, il permet de discriminer les espèces en présence dans un mélange. Une fois le coefficient de diffusion D d'une molécule connu, il est possible de remonter au rayon hydrodynamique R_h de celle-ci à l'aide de l'équation de Stockes-Einstein (Équation 3.).

Équation 3.1

$$D = \frac{k_B T}{6\pi\eta R_h}$$

Avec : - D coefficient de diffusion en m².s^{-1}

- k_B constante de Bolzmann (1,3806488.10^{-3} J.K^{-1})

- T température absolue en K
- η viscosité du solvant en $Pa.s^{-1}$
- R_h rayon hydrodynamique en m.

Dans notre cas, la RMN DOSY montre la formation, à faible concentration, d'espèces relativement petites (R_h = 6 – 17 Å) par rapport aux monomères (R_h = 5 – 10 Å), tandis que de beaucoup plus gros objets (R_h = 134 – 157 Å) sont formés à plus haute concentration (Tableau 3.2). Il est intéressant de noter que deux produits de taille différente sont observés quand la polycondensation est réalisée à faible concentration (10 mM, exemple en Figure 3.8), ce qui confirme l'observation de deux types d'acylhydrazones par RMN ^{1}H. La dépendance vis-à-vis de la concentration observée par RMN DOSY est compatible avec la formation de macrocycles à faible concentration et de polymères à haute concentration.

Tableau 3.2 : Caractérisation des réactions de polycondensations par RMN DOSY. Les rayons hydrodynamiques ont été déterminés avec l'équation de Stockes – Einstein.

Composés	Concentration (mM)	Coefficient de diffusion D ($m^2.s^{-1}$)	Rayon hydrodynamique (Å)
Gua-Ald	10	$2,32 \times 10^{-10}$	5
EG-Hyd	10	$1,12 \times 10^{-10}$	10
EG-Ald	10	$1,73 \times 10^{-10}$	6
EI-Hyd	10	$1,19 \times 10^{-10}$	9
Poly(Gua-EG)	200	7×10^{-12}	157
Poly(Gua-EG)	10	$6,4 – 18,6 \times 10^{-11}$	6 – 17
Poly(EI-EG)	200	$8,2 \times 10^{-12}$	134
Poly(EI-EG)	100	$1,4 – 3,27 \times 10^{-11}$	34 – 78
Poly(EI-EG)	10	$1,03 – 1,87 \times 10^{-10}$	6 – 11

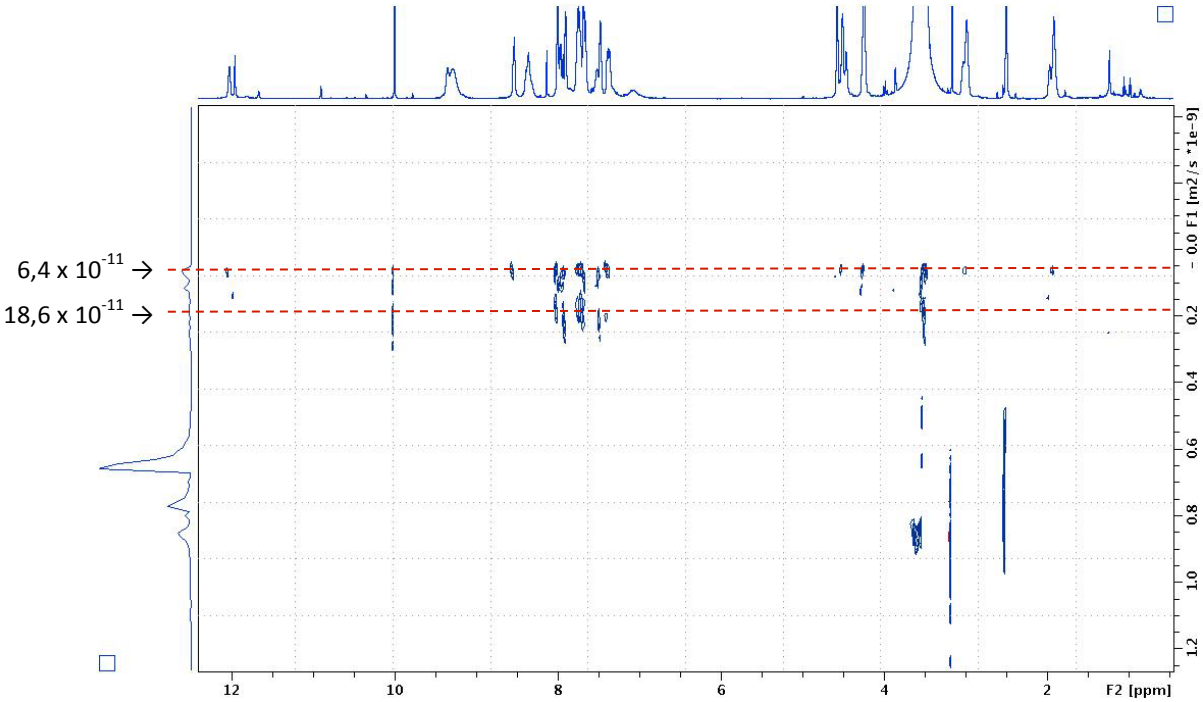


Figure 3.8 : Spectre RMN DOSY (600 MHz) du polymère **Poly(Gua-EG)** à 10 mM dans le DMSO-d6.

De plus, une augmentation importante de la viscosité du milieu réactionnel est aussi observée visuellement, ce qui est en corrélation avec la formation de polymère.

3. Caractérisation par MALDI-ToF et LC-MS

Une analyse de spectrométrie de masse MALDI-ToF confirme la présence d'oligomères à chaîne ouverte ainsi que de macrocycles (Figure 3.9).

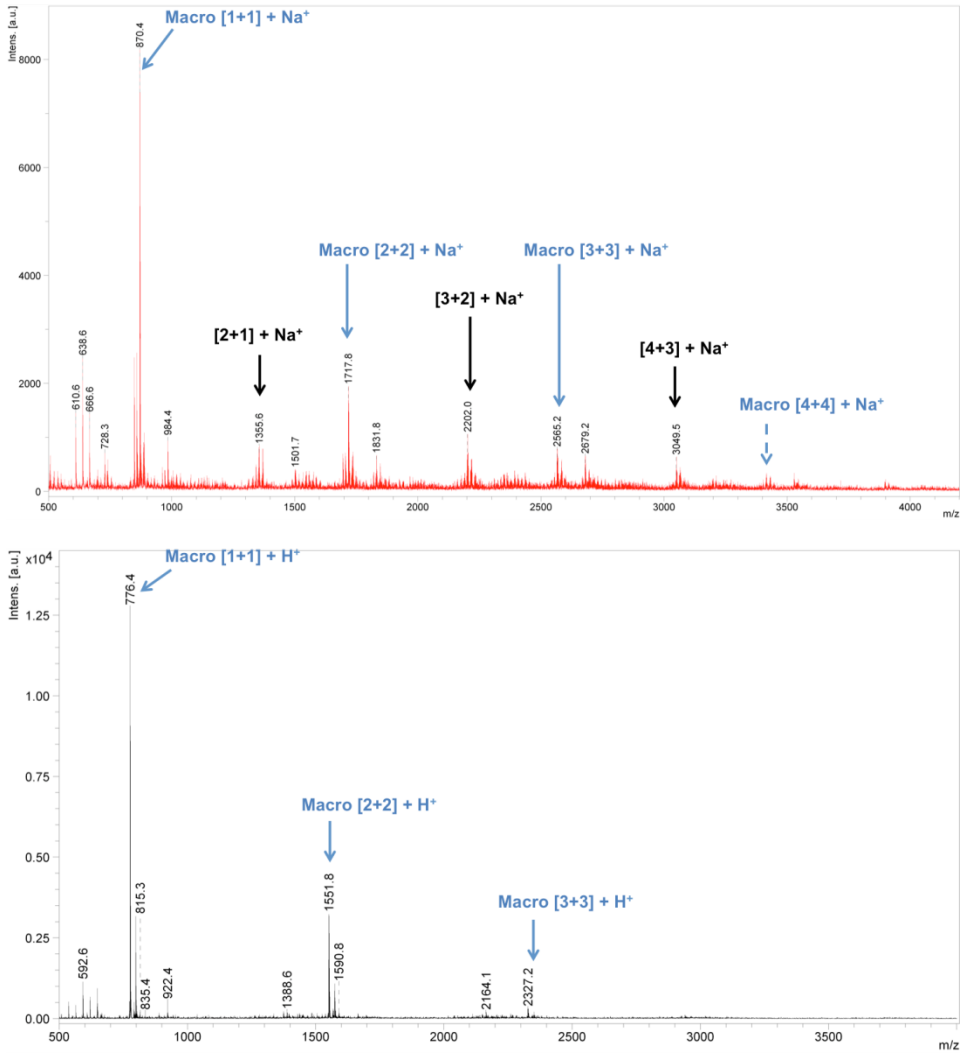


Figure 3.9 : Spectres MALDI-ToF de **Poly(EI-EG)** (haut) et **Poly(Gua-EG)** (bas) (matrice HCCA).

L'analyse LC-MS du produit **Poly(Gua-EG)** met en évidence deux populations de produits : un pic étroit (t_R 2,8 minutes) qui correspond à des macrocycles [1+1] et / ou [2+2], et un large pic (t_R 3,2 minutes) qui peut correspondre aux oligomères (Figure 3.10). Cette interprétation est en effet renforcée par le fait qu'une dilution de 200 mM à 1 mM entraîne la disparition du pic large et l'augmentation du pic étroit, ce qui est compatible avec une conversion des chaînes en cycles sous l'effet de la dilution (Figure 3.11). Considérés ensemble, ces résultats indiquent que les monomères **EI-Hyd** et **Gua-Ald** réagissent avec leur partenaire complémentaire, respectivement **EG-Ald** et **EG-Hyd**, pour générer des polymères covalents dynamiques basés sur les acylhydrazones à haute concentration (typiquement supérieure à 100 mM) tandis que les macrocycles sont les espèces dominantes formées à faible concentration. Le réarrangement, engendré par la dilution, des chaînes polymériques en espèces cycliques plus petites s'effectue en plusieurs heures ($t_{1/2}$ > 50 h, Figure 3.12) ce qui laisse suffisamment de temps au polymère pour s'associer avec sa cible biomoléculaire.

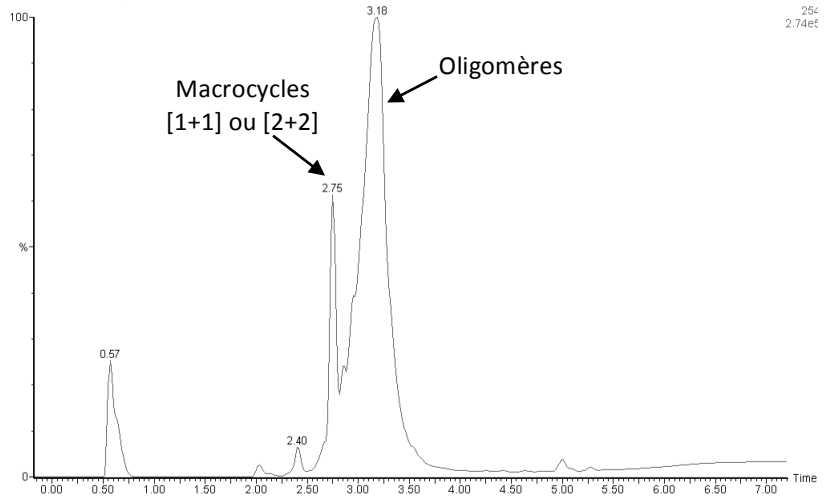


Figure 3.10: Chromatogramme HPLC de **Poly(Gua-EG)** à 1 mM dans le DMSO (injection de 6 µL). Le pic étroit à 2,75 minutes a été identifié en masse comme correspondant aux macrocycles [1+1] et / ou [2+2] tandis que le large pic à 3,18 minutes correspondrait aux oligomères. Méthode I (HPLC 2), suivi à 254 nm.

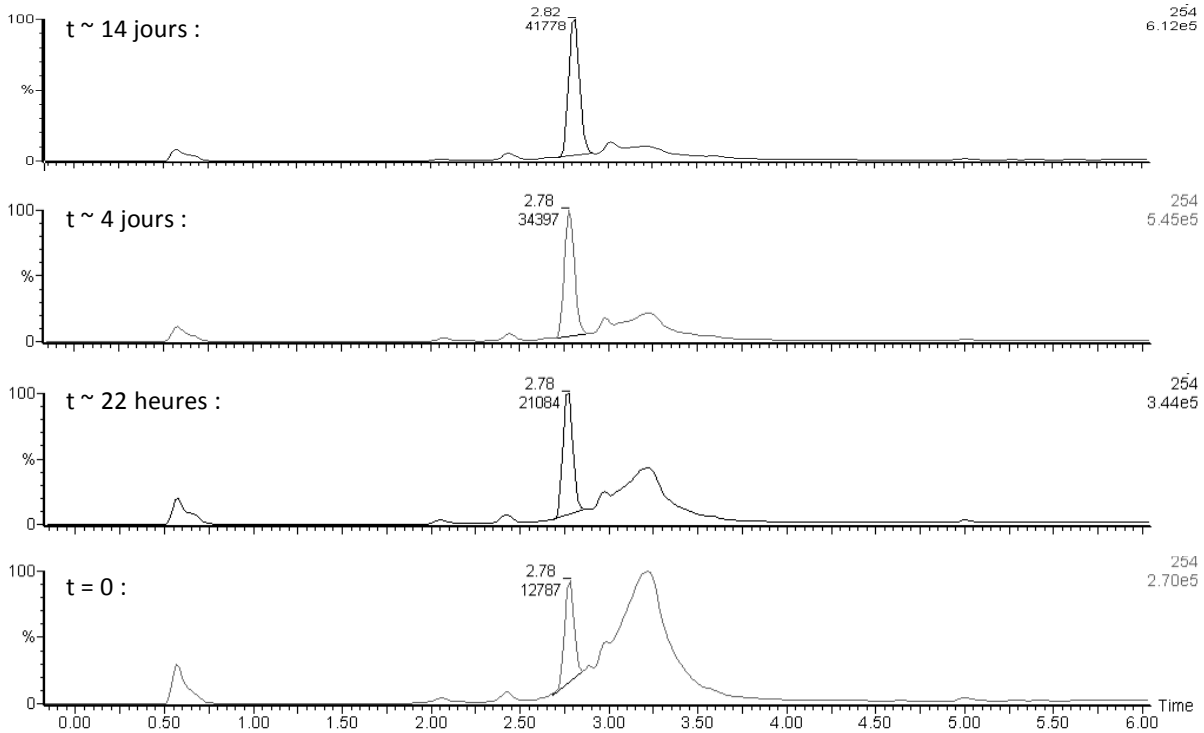


Figure 3.11: Suivi HPLC au cours du temps de la formation de macrocycles **[Gua+EG]** (pic à 2,7 – 2,8 minutes) sous l'effet de la dilution du polymère **Poly(Gua-EG)** de 200 mM à 1 mM dans le DMSO. Sur ces exemples de chromatogrammes, la disparition des espèces oligomériques (pic large à 3,2 – 3,3 minutes) est également observée, mais non quantifiée. Méthode I (HPLC 2), suivi à 254 nm.

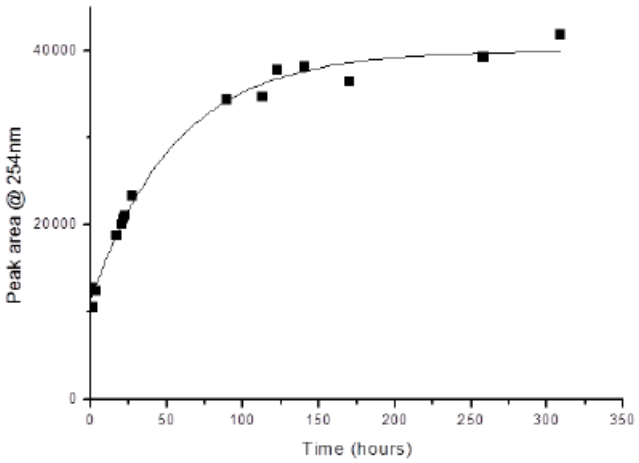


Figure 3.12: Formation au cours du temps de macrocycles (aire du pic à 2,7 - 2,8 minutes) lors de la dilution du **Poly(Gua-EG)** de 200 mM à 1 mM dans le DMSO.

Le degré de polymérisation d'un polymère covalent dynamique est déterminé par la constante d'équilibre de la réaction réversible utilisée pour assembler la chaîne principale, dans le cas présent celle de la formation d'acylhydrazone. Tandis que les résultats décrits ci-dessus indique clairement la formation de chaînes ouvertes à haute concentration, les données obtenues par spectroscopie RMN DOSY et spectrométrie de masse MALDI-ToF tendent plutôt à indiquer la formation d'oligomères en lieu et place de longues chaînes polymériques. Toutefois, les clusters cationiques contenant un nombre limité et plutôt faible d'unités de liaison ont récemment montré leur capacité à complexer efficacement un ADN, soulignant ici les potentialités d'oligomères dynamiques pour une application en transfection.[62–64,102–110]

III. DEGRADATION DES POLYMERES EN MILIEU ACIDE

A. DEGRADATION PAR REACTION D'ECHANGE

1. Réaction modèle

Les poly-acylhydrazones sont connues pour leur échange avec les oxyamines, conduisant à la formation d'oximes.[111] Il a ainsi été vérifié avec le composé modèle **27**, contenant une seule acylhydrazone, que la réaction avec un excès de méthoxyamine (5 équivalents) en solution aqueuse conduit à la formation du composé oxime **29** au détriment de l'acylhydrazone initiale (Schéma 3.15, Figure 3.13 et Figure 3.14).

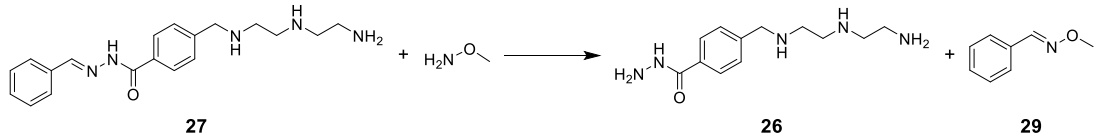


Schéma 3.15 : Dégradation du composé modèle **27** par échange avec la méthoxyamine.

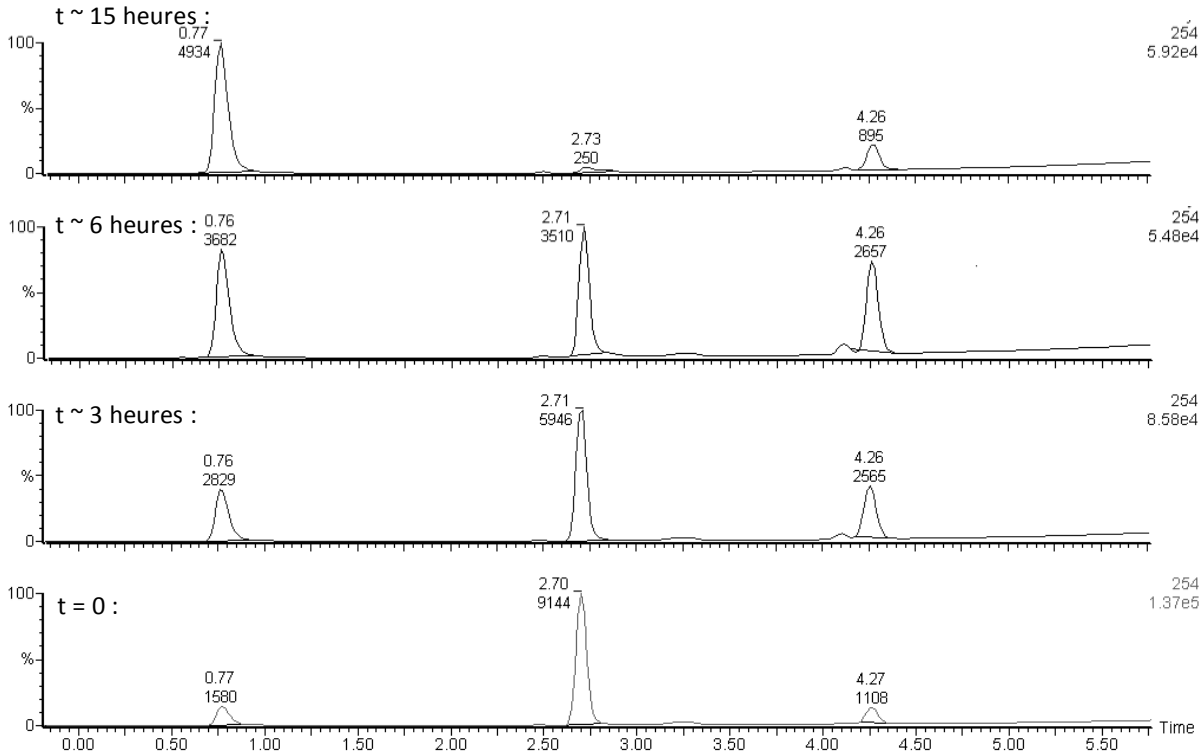


Figure 3.13 : Suivi HPLC au cours du temps de la dégradation du composé modèle **27** en présence de 5,0 équivalents de méthoxyamine à pH 5. Acylhydrazone **27** à t_R 2,7 minutes. Méthode I (HPLC 2), suivi à 254 nm.

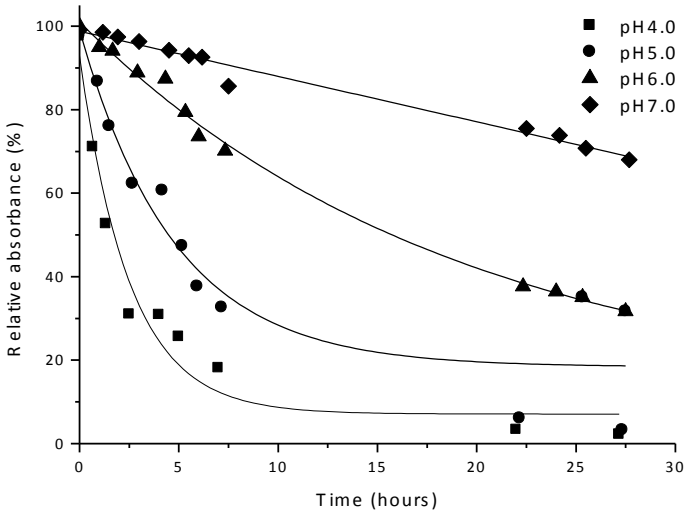


Figure 3.14: Dégradation du composé modèle **27** en présence de 5,0 équivalents de méthoxyamine à différents pH. Suivi HPLC de la disparition de l'acylhydrazone initiale.

2. Dégradation du polymère **Poly(Gua-EG)**

Pour vérifier la nature dynamique des polymères, la stabilité du produit **Poly(Gua-EG)** a ainsi été suivie par HPLC à différents pH en présence de 5 équivalents de méthoxyamine par acylhydrazone (estimée par rapport au nombre de fonctions aldéhydes initiales), la méthoxyamine pouvant agir comme un terminateur de chaîne (Figure 3.15). La formation de la bisoxime **GuaOx** et du bishydrazide **EG-Hyd** suit une cinétique du pseudo premier ordre, confirmant donc que le produit **Poly(Gua-EG)** est capable de subir l'échange de monomères (Figure 3.17 et Figure 3.18). Les résultats obtenus mettent également en évidence que le composé **Poly(Gua-EG)** est relativement stable à pH 7,0 tandis que l'échange avec la méthoxyamine a facilement lieu à pH acide, les temps de demi-vie ($t_{1/2}$) de la formation du guanidinium bisoxime **GuaOx** montrant une forte dépendance au pH ($t_{1/2}$ > 200 heures à pH 7,0 ; $t_{1/2}$ = 54 heures à pH 5,0 ; Figure 3.17).

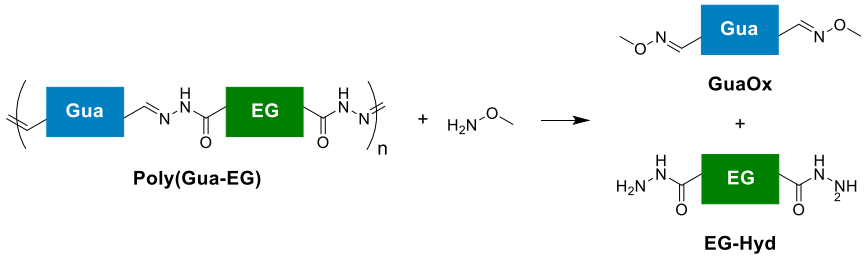


Figure 3.15: Dégradation du polymère **Poly(Gua-EG)** par action de la méthoxyamine.

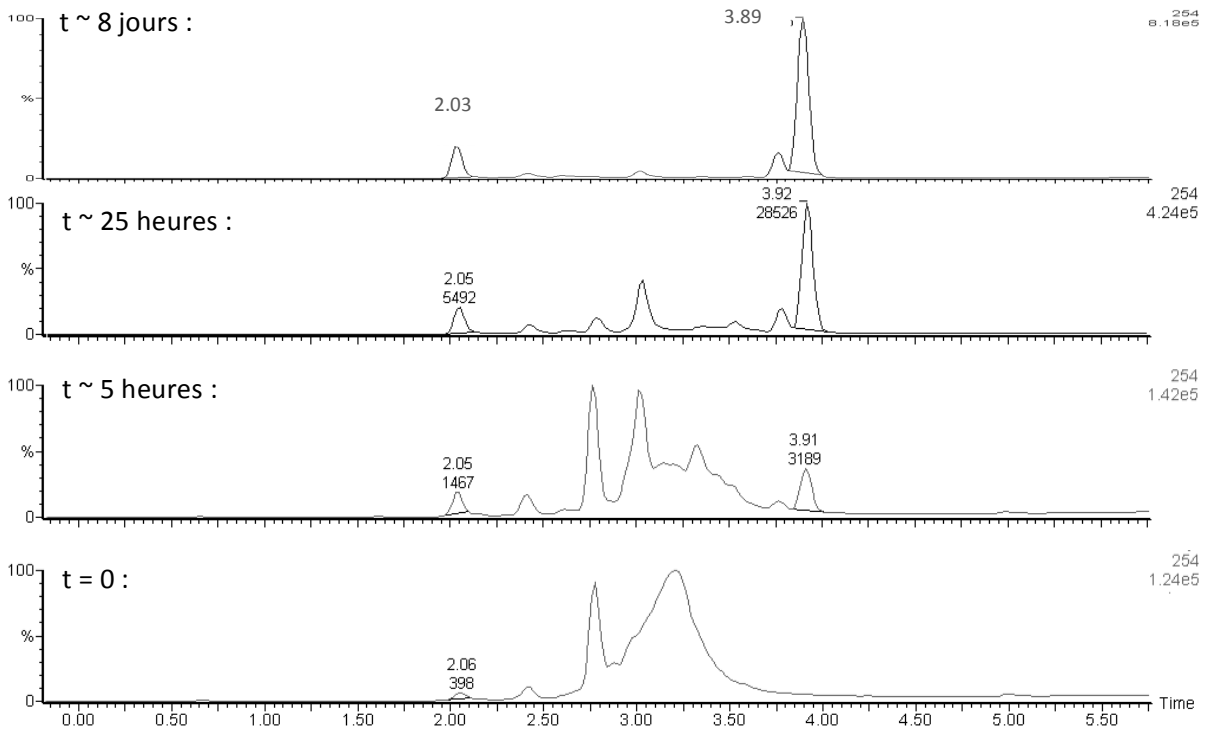
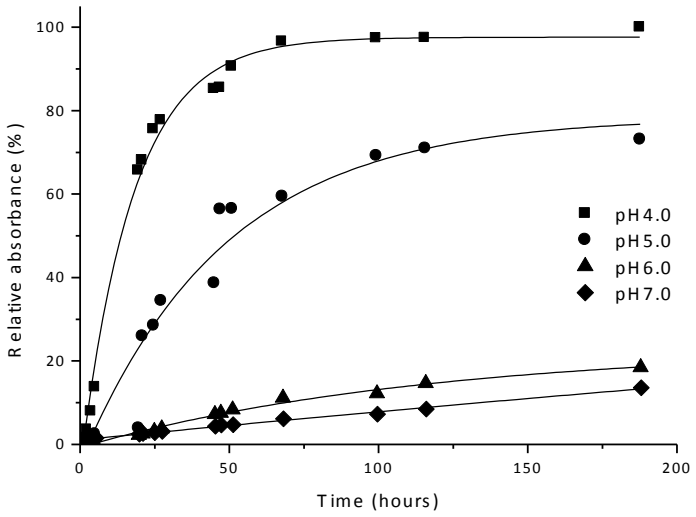


Figure 3.16 : Suivi HPLC de la dégradation du polymère **Poly(Gua-EG)** par échange avec la méthoxyamine à pH 5. Monomère **EG-Hyd** à t_R 2,0 minutes et composé **GuaOx** à t_R 3,9 minutes. Méthode I (HPLC 2), suivi à 254 nm.



pH	$t_{1/2}$
4	8 h
5	54 h
6	> 200 h[a]
7	> 200 h[a]

Figure 3.17: Cinétique et temps de demi-vie de formation du guanidinium bisoxime **GuaOx** à partir du composé **Poly(Gua-EG)** en présence de 5,0 équivalents de méthoxyamine à différents pH. [a] Estimation, temps de demi-vie non atteint.

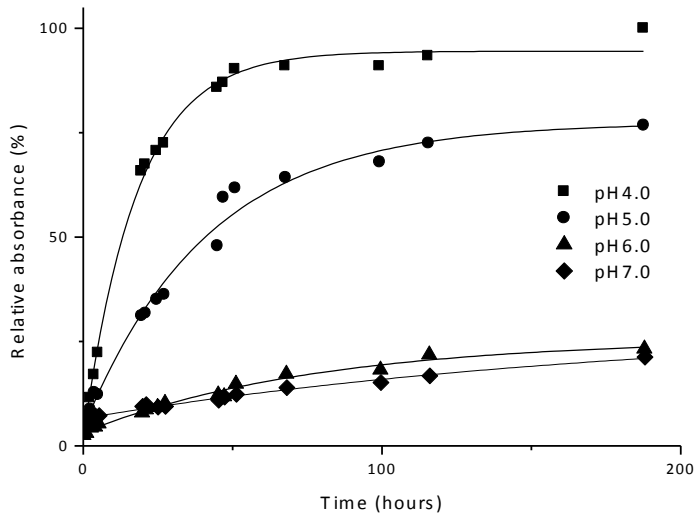


Figure 3.18: Cinétique de formation du bishydrazide **EG-Hyd** à partir du composé **Poly(Gua-EG)** en présence de 5,0 équivalents de méthoxyamine à différents pH.

B. DEGRADATION PAR HYDROLYSE

L'hydrolyse directe du composé modèle **27** contenant une seule acylhydrazone a aussi été testée à différents pH et suivie par HPLC (Figure 3.20). L'hydrolyse est beaucoup plus rapide à pH 5,0 ($t_{1/2}$ = 50 h) qu'à pH 7,0 ($t_{1/2}$ = 200 h), ce qui est très intéressant pour la conception de systèmes de délivrance 'intelligents' qui peuvent promouvoir la sortie de l'endosome par dégradation du cargo.

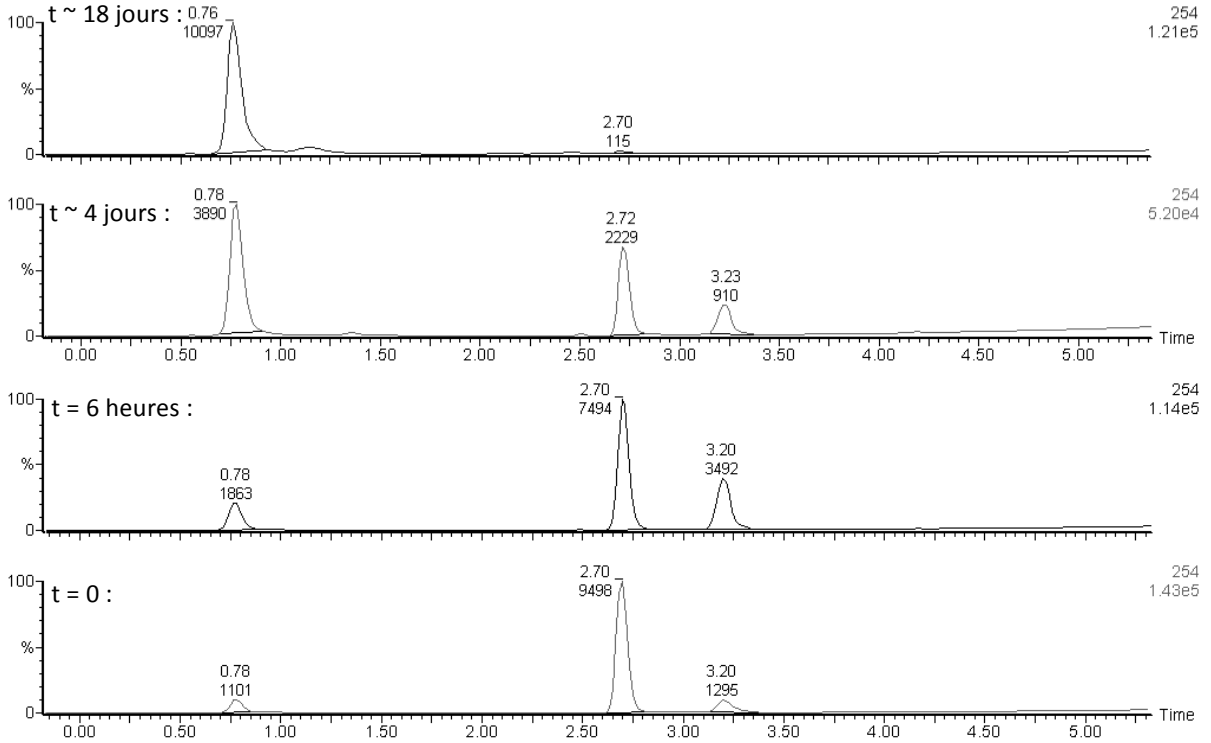


Figure 3.19 : Suivi HPLC au cours du temps de la dégradation par hydrolyse du composé modèle **27** à pH 5. Acylhydrazone **27** à t_R 2,7 minutes, benzaldéhyde à t_R 3,2 minutes.

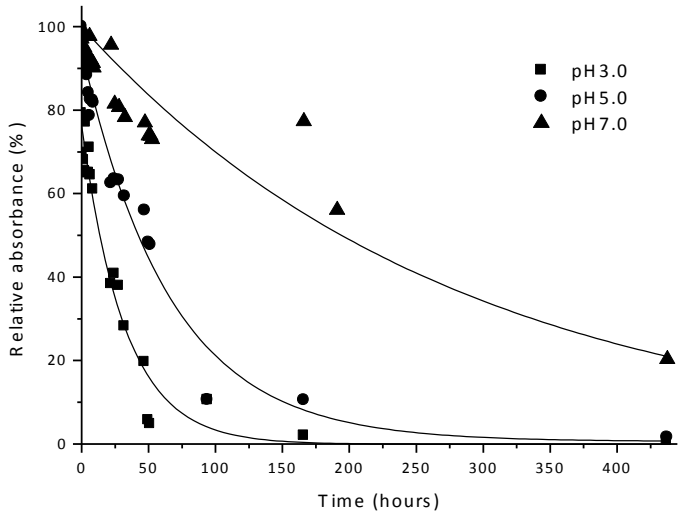


Figure 3.20: Dégradation du composé modèle **27** à différents pH par hydrolyse directe.

De manière similaire, l'hydrolyse du composé **Poly(Gua-EG)** dans des tampons aqueux a été suivie à différents pH. Seule la conversion polymères à macrocycles a été observée, mais les résultats démontrent clairement la dépendance au pH de ce processus, la conversion étant plus rapide à pH 4,0 qu'à pH 7,0.

Dans l'ensemble, ces résultats démontrent que ces polymères dynamiques se réarrangent en espèces plus petites à pH acide par échange de constituants et hydrolyse.

IV. ETUDES DE FLUORESCENCE (ESSAIS DE DEPLACEMENT DU BROMURE D'ETHIDIUM)

La capacité des polymères dynamiques à se lier avec un double brin d'ADN a été évaluée par un essai de déplacement du bromure d'éthidium en présence d'un double brin d'ADN de thymus de veau dans l'eau tamponnée à pH 7,2. La fluorescence du bromure d'éthidium augmente significativement lorsque celui-ci s'intercale dans un double brin d'ADN. Toutefois, la complexation d'un polymère cationique peut entraîner des changements structurels au sein du double brin d'ADN, ce qui peut diminuer son affinité pour le bromure d'éthidium. En conséquence, la fluorescence du bromure d'éthidium peut diminuer au fur et à mesure de la complexation du polymère cationique.

Une solution d'ADN de thymus de veau et de bromure d'éthidium a été dosée avec les composés **Poly(EI-EG)** et **Poly(Gua-EG)** et l'émission de fluorescence a été suivie à 590 nm (Figure 3.21). Les polymères sont ajoutés à partir d'une solution mère concentrée (200 mM) pour éviter une réorganisation constitutionnelle des polymères en macrocycles sous l'effet de la dilution. Les résultats ont été quantifiés en déterminant l'excès de charge (CE) des polymères cationiques par rapport à l'ADN de thymus de veau. Cet excès de charge est défini comme le rapport du nombre nominal de charges positives des polymères sur le nombre de charges négatives présentes sur l'ADN de thymus de veau. Pour le calculer, il a été considéré que toutes les amines sont protonées et que tous les phosphodiesters portent une charge négative. Avec ces deux hypothèses, l'excès de charges est équivalent à un ratio N/P. Les valeurs de CE_{50} correspondent à l'excès de charge nécessaire pour atteindre une diminution de 50% de la fluorescence (Tableau 3.3).

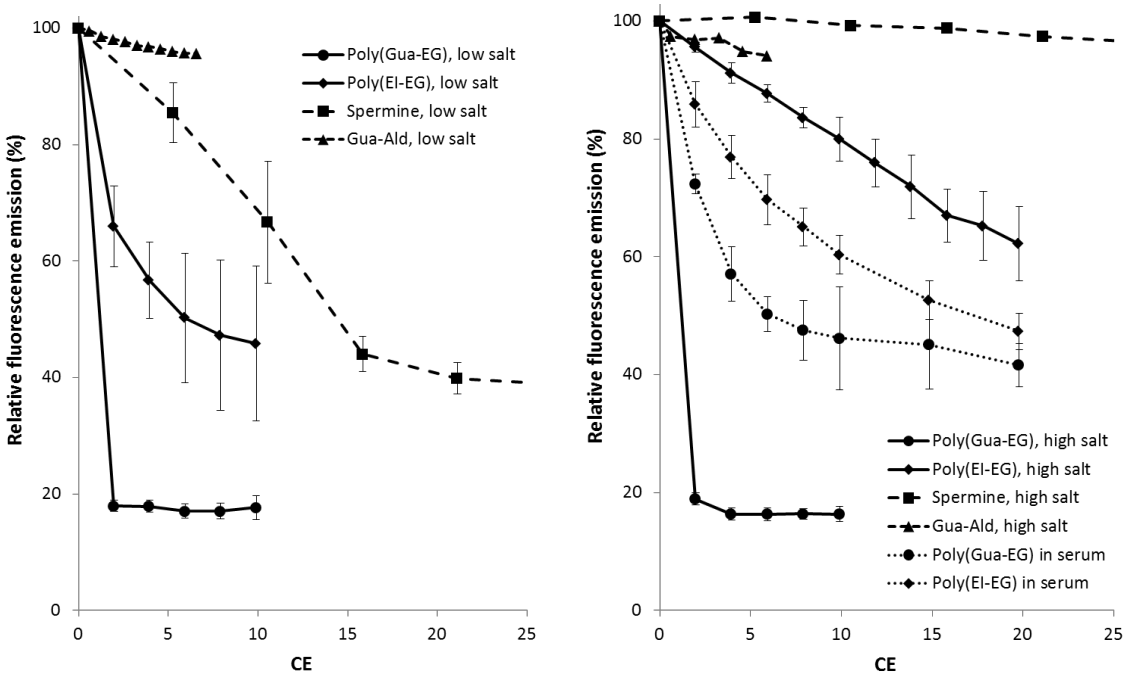


Figure 3.21: Titration par fluorescence d'une solution d'ADN de thymus de veau et de bromure d'éthidium par la spermine, **Poly(EI-EG)**, **Gua-Ald** et **Poly(Gua-EG)**, en présence de 9,4 mM (low salt) ou de 150 mM (high salt) de chlorure de sodium. Les expériences ont été répétées deux ou trois fois, dans l'eau tamponnée (pH = 7,2) ou le sérum de veau fétal. Les barres d'erreur représentent l'écart-type.

Tableau 3.3 : Propriétés d'assemblage à l'ADN de thymus de veau pour la spermine, **Poly(EI-EG)**, **Gua-Ald** et **Poly(Gua-EG)** déterminées par l'essai de déplacement du bromure d'éthidium. [a]

Composés	[NaCl] (mM)	Charge nominale	CE_{50}
Spermine	9,4	4+	14,4
Poly(Gua-EG)	9,4	3+	1,2
Poly(EI-EG)	9,4	3+	6
Gua-Ald	9,4	1+	> 30 [b]
Spermine	150	4+	> 30 [b]
Poly(Gua-EG)	150	3+	1,2
Poly(EI-EG)	150	3+	> 20 [b]
Gua-Ald	150	1+	> 30 [b]

[a] Conditions : expériences effectuée dans un tampon HEPES (100 mM HEPES, 10 µM EDTA, pH 7,2) avec 9,4 ou 150 mM de NaCl. La quantité d'ADN de thymus de veau (0,1 mg) et la concentration finale en bromure d'éthidium sont gardées constantes. [b] Valeur estimée.

La spermine a été utilisée comme témoin car c'est un ligand bien connu de l'ADN double brin.[112] La spermine ne montre pas d'assemblage significatif avec l'ADN de thymus de veau dans les conditions de forte salinité et seulement un assemblage modéré dans les

conditions de faible salinité (CE_{50} = 14,4). Ces résultats sont en accord avec les données de la littérature.[64]

De manière similaire, le composé **Poly(EI-EG)** présente également une dépendance à la salinité typique des polyamines. Toutefois, l'ordre de grandeur des changements de fluorescence est plus important qu'avec la spermine, une valeur de CE_{50} égale à 6 étant atteinte dans les conditions de faible salinité. Cette observation confirme ainsi la nature polymérique du composé **Poly(EI-EG)** et l'avantage apporté par la présentation de multiples triamines diéthylènes au sein de la chaîne principale. Il est important de noter que les PEI atteignent généralement une complexation complète à un ratio N/P supérieur à 4.[113]

Le monomère **Gua-Ald** ne présente presque pas de changements de fluorescence quelques soient les conditions de salinité. Cependant, une rapide décroissance de la fluorescence est observée au cours de l'addition du polymère **Poly(Gua-EG)** (CE_{50} = 1,2). Il est également intéressant de noter qu'aucune dépendance à la salinité n'est observée dans ce cas. Ce résultat confirme que la présence de répétitions multiples d'un groupement guanidinium au sein de la chaîne principale d'un polymère covalent dynamique est suffisante pour atteindre une complexation efficace de l'ADN de thymus de veau. Un essai de déplacement du bromure d'éthidium avec une solution diluée (1 mM) de **Poly(Gua-EG)**, contenant essentiellement des macrocycles (visibles en HPLC), a été effectué pour vérifier qu'aucune réorganisation du polymère par dilution n'est à l'origine de la complexation observée : cette complexation avec l'ADN de thymus de veau étant plus faible qu'avec la solution concentrée (CE_{50} supérieure à 4), il est possible de conclure que ce sont bien les oligomères à chaines ouvertes plutôt que les espèces macrocycliques qui promeuvent la complexation avec l'ADN par interactions multivalentes. Comparé au composé **Poly(EI-EG)**, le composé **Poly(Gua-EG)** présente une valeur de CE_{50} plus faible, indiquant ainsi que les guanidiniums sont plus efficaces que les amines pour promouvoir l'interaction avec l'ADN double brin dans ces matériaux hybrides. De petites molécules présentant des guanidiniums ont en effet été reportées pour être supérieures à leurs équivalents aminés en termes d'interactions avec l'ADN.[114]

La complexation avec l'ADN de thymus de veau a aussi été testée dans le sérum de veau fétal puisqu'il est bien connu que les protéines du plasma peuvent interagir avec les polymères cationiques et déstabiliser le complexe ADN – polymère. Une diminution significative de l'efficacité d'interaction du composé **Poly(Gua-EG)** (CE_{50} = 5,8) est observée tandis que le composé **Poly(EI-EG)** (CE_{50} = 17,1) conserve l'essentiel de son activité. Cette observation peut s'expliquer par le fait que, contrairement au composé

Poly(EI-EG) qui est composé d'éthylènimines de faible basicité, le composé **Poly(Gua-EG)** est probablement plus protoné à pH physiologique (pKa guanidinium ~ 12,5) et pour cette raison, plus sensible à la présence de protéines anioniques.

V. RETARD SUR GEL EN PRESENCE D'ADN

Pour confirmer la capacité des polymères à complexer l'ADN double brin, une électrophorèse sur gel d'agarose a été réalisée : un ADN plasmidique pET-b a été utilisé et différents ratios N/P (de 0,1 à 10) ont été testés. Les résultats mettent en évidence l'absence de déplacement de la bande plasmidique avec la spermine, le monomère **Gua-Ald** ou de manière plus surprenante le polymère **Poly(EI-EG)** (Figure 3.22).

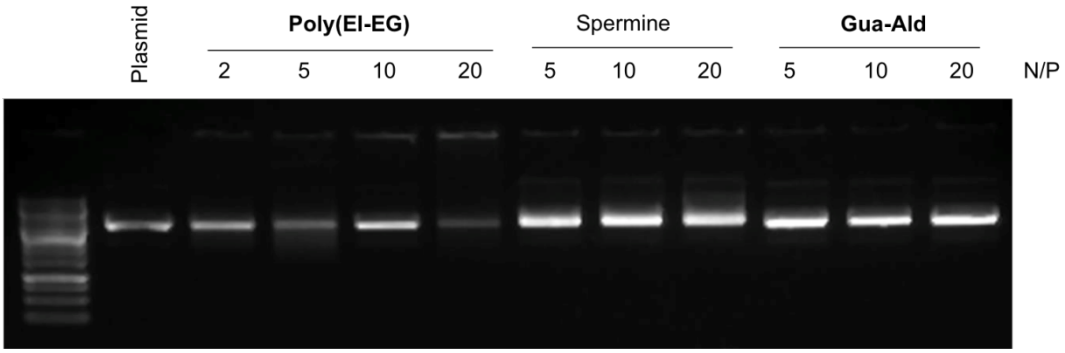


Figure 3.22 : Electrophorèse sur gel de l'ADN plasmidique avec la spermine, **Poly(EI-EG)** et **Gua-Ald**.

Le fait que le polymère **Poly(EI-EG)** ne semble pas complexer le plasmide dans cette expérience peut être dû au pH plus haut du tampon TAE (pH = 8,2). Cela peut en effet réduire le degré de protonation du polymère **Poly(EI-EG)** en comparaison de l'expérience de déplacement du bromure d'éthidium qui était effectuée à pH 7,2.

Cependant, la complexation est détectée par la disparition de la bande plasmidique native avec le polymère **Poly(Gua-EG)** pour des ratios N/P supérieurs à 2 (Figure 3.23). En effet, le traceur utilisé pour visualiser l'ADN, SYBR® Safe, est un intercalant de l'ADN comme le bromure d'éthidium et suite à la complexation de l'ADN par le polymère **Poly(Gua-EG)**, il peut ne pas pouvoir s'intercaler : la bande plasmidique reste donc invisible sur le gel. Ainsi, l'essai de retard sur gel confirme que le polymère covalent dynamique polyguanidinium est plus efficace pour complexer l'ADN double brin que le polymère polyamine.

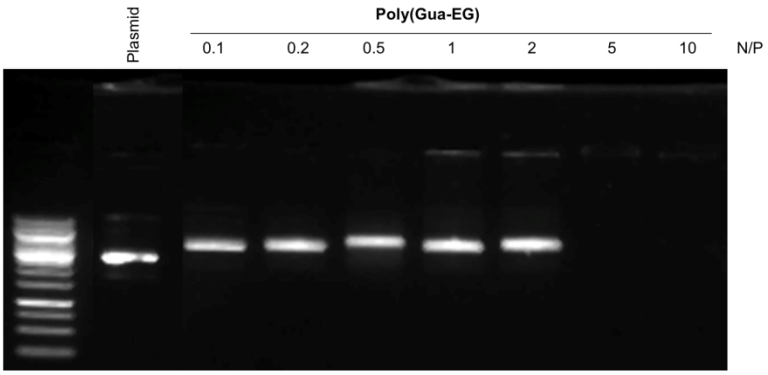


Figure 3.23: Electrophorèse sur gel de l'ADN plasmidique avec le polymère **Poly(Gua-EG)**.

Considéré ensemble, ces résultats démontrent que les matériaux cationiques basés sur les poly-acylhydrazones décrits ici sont capables de complexer efficacement l'ADN double brin dans des tampons aqueux et le sérum biologique grâce à la présentation de multiples unités de reconnaissance le long de leur chaîne principale.

CONCLUSIONS

La conception et la synthèse de deux polymères covalents dynamiques hybrides qui combinent des moitiés cationiques et de courts groupements d'éthylène glycol au sein de leur chaîne principale ont été décrites. Ces matériaux sont capables de s'auto-assembler par une réaction de polycondensation. Ce processus conduit principalement à la formation de macrocycles à basse concentration et d'oligomères à chaîne ouverte à haute concentration, comme l'a mis en évidence la spectroscopie RMN DOSY.

Le caractère dynamique de ces polymères a été mis en évidence à travers les réactions d'échange avec la méthoxyamine et d'hydrolyse directe. Il est intéressant de noter que l'échange et l'hydrolyse ont lieu bien plus vite à pH 5,0 qu'à pH 7,0, ce qui est de première importance pour promouvoir la sortie de l'endosome d'un cargo suite à sa dégradation contrôlée par le pH.

De plus, les essais de déplacement du bromure d'éthidium et de retard sur gel démontrent qu'une complexation efficace de l'ADN double brin a lieu dans les tampons aqueux comme dans le sérum biologique à des ratios N/P comparables à ceux des polymères PEI.

En résumé, ces matériaux hybrides auto-assemblés combinent deux caractéristiques intéressantes pour une application en tant que vecteurs 'intelligents' pour la délivrance de

gènes : une complexation efficace de l'ADN double brin dans les médias biologiques, et une dégradation catalysée en milieu acide qui devrait promouvoir la sortie de l'endosome.

REFERENCES BIBLIOGRAPHIQUES

(1) De Mesmaeker, A.; Haener, R.; Martin, P.; Moser, H. E. *Acc. Chem. Res.* **1995**, *28*, 366–374.

(2) Uhlmann, E.; Peyman, A. *Chem. Rev.* **1990**, *90*, 543–584.

(3) Burnett, J. C.; Rossi, J. J.; Tiemann, K. *Biotechnol. J.* **2011**, *6*, 1130–1146.

(4) Gaynor, J. W.; Campbell, B. J.; Cosstick, R. *Chem. Soc. Rev.* **2010**, *39*, 4169–4184.

(5) Castanotto, D.; Rossi, J. J. *Nature* **2009**, *457*, 426–433.

(6) Kool, E. T. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, *25*, 1–28.

(7) Mintzer, M. A.; Simanek, E. E. *Chem. Rev.* **2008**, *109*, 259–302.

(8) Lehrman, S. *Nature* **1999**, *401*, 517–518.

(9) Whitehead, K. A.; Langer, R.; Anderson, D. G. *Nat Rev Drug Discov* **2009**, *8*, 129–138.

(10) Samal, S. K.; Dash, M.; Van Vlierberghe, S.; Kaplan, D. L.; Chiellini, E.; van Blitterswijk, C.; Moroni, L.; Dubruel, P. *Chem. Soc. Rev.* **2012**, *41*, 7147–7194.

(11) Nguyen, D. N.; Green, J. J.; Chan, J. M.; Langer, R.; Anderson, D. G. *Adv. Mater.* **2009**, *21*, 847–867.

(12) Wong, S. Y.; Pelet, J. M.; Putnam, D. *Polym. Biomed. Appl.* **2007**, *32*, 799–837.

(13) Merdan, T.; Kopeček, J.; Kissel, T. *Polym. Conjug. Cancer Ther.* **2002**, *54*, 715–758.

(14) Liu, X.; Rocchi, P.; Peng, L. *New J. Chem.* **2012**, *36*, 256–263.

(15) Edelstein, M. L.; Abedi, M. R.; Wixon, J. *J. Gene Med.* **2007**, *9*, 833–842.

(16) Kwon, Y. J. *Acc. Chem. Res.* **2011**, *45*, 1077–1088.

(17) Wagner, E. *Acc. Chem. Res.* **2011**, *45*, 1005–1013.

(18) Boussif, O.; Lezoualc'h, F.; Zanta, M. A.; Mergny, M. D.; Scherman, D.; Demeneix, B.; Behr, J. P. *Proc. Natl. Acad. Sci.* **1995**, *92*, 7297–7301.

(19) Barnard, A.; Posocco, P.; Pricl, S.; Calderon, M.; Haag, R.; Hwang, M. E.; Shum, V. W. T.; Pack, D. W.; Smith, D. K. *J. Am. Chem. Soc.* **2011**, *133*, 20288–20300.

(20) Barnard, A.; Posocco, P.; Fermeglia, M.; Tschiche, A.; Calderon, M.; Pricl, S.; Smith, D. K. *Org. Biomol. Chem.* **2014**, *12*, 446–455.

(21) Kozielski, K. L.; Tzeng, S. Y.; Green, J. J. *Chem. Commun.* **2013**, *49*, 5319–5321.

(22) Son, S.; Namgung, R.; Kim, J.; Singha, K.; Kim, W. J. *Acc. Chem. Res.* **2011**, *45*, 1100–1112.

(23) Wang, Y.; Zhang, R.; Xu, N.; Du, F.-S.; Wang, Y.-L.; Tan, Y.-X.; Ji, S.-P.; Liang, D.-H.; Li, Z.-C. *Biomacromolecules* **2010**, *12*, 66–74.

(24) Kim, T.; Lee, M.; Kim, S. W. *Biomaterials* **2010**, *31*, 1798–1804.

(25) Hardy, J. G.; Love, C. S.; Gabrielson, N. P.; Pack, D. W.; Smith, D. K. *Org. Biomol. Chem.* **2009**, *7*, 789–793.

(26) Bauhuber, S.; Hozsa, C.; Breunig, M.; Göpferich, A. *Adv. Mater.* **2009**, *21*, 3286–3306.

(27) Binauld, S.; Stenzel, M. H. *Chem. Commun.* **2013**, *49*, 2082–2102.

(28) Hong, B. J.; Chipre, A. J.; Nguyen, S. T. *J. Am. Chem. Soc.* **2013**, *135*, 17655–17658.

(29) Duan, S.; Yuan, W.; Wu, F.; Jin, T. *Angew. Chem. Int. Ed.* **2012**, *51*, 7938–7941.

(30) Cui, L.; Cohen, J. L.; Chu, C. K.; Wich, P. R.; Kierstead, P. H.; Fréchet, J. M. J. *J. Am. Chem. Soc.* **2012**, *134*, 15840–15848.

(31) Wolff, J. A.; Rozema, D. B. *Mol Ther* **2007**, *16*, 8–15.

(32) Rozema, D. B.; Lewis, D. L.; Wakefield, D. H.; Wong, S. C.; Klein, J. J.; Roesch, P. L.; Bertin, S. L.; Reppen, T. W.; Chu, Q.; Blokhin, A. V.; Hagstrom, J. E.; Wolff, J. A. *Proc. Natl. Acad. Sci.* **2007**, *104*, 12982–12987.

(33) Aissaoui, A.; Martin, B.; Kan, E.; Oudrhiri, N.; Hauchecorne, M.; Vigneron, J.-P.; Lehn, J.-M.; Lehn, P. *J. Med. Chem.* **2004**, *47*, 5210–5223.

(34) Welsh, D. J.; Jones, S. P.; Smith, D. K. *Angew. Chem. Int. Ed.* **2009**, *48*, 4047–4051.

(35) Toita, R.; Kang, J.-H.; Tomiyama, T.; Kim, C. W.; Shiosaki, S.; Niidome, T.; Mori, T.; Katayama, Y. *J. Am. Chem. Soc.* **2012**, *134*, 15410–15417.

(36) Moghimi, S. M.; Symonds, P.; Murray, J. C.; Hunter, A. C.; Debska, G.; Szewczyk, A. *Mol Ther* **2005**, *11*, 990–995.

(37) Joralemon, M. J.; McRae, S.; Emrick, T. *Chem. Commun.* **2010**, *46*, 1377–1393.

(38) Zhang, S.; Zhao, Y.; Zhao, B.; Wang, B. *Bioconjug. Chem.* **2010**, *21*, 1003–1009.

(39) Li, W.; Zhang, P.; Zheng, K.; Hu, Q.; Wang, Y. *J. Mater. Chem. B* **2013**, *1*, 6418–6426.

(40) Weber, N. D.; Merkel, O. M.; Kissel, T.; Muñoz-Fernández, M. Á. *J. Controlled Release* **2012**, *157*, 55–63.

(41) Jager, M.; Schubert, S.; Ochrimenko, S.; Fischer, D.; Schubert, U. S. *Chem. Soc. Rev.* **2012**, *41*, 4755–4767.

(42) Zhao, N.; Roesler, S.; Kissel, T. *Int. J. Pharm.* **2011**, *411*, 197–205.

(43) Glodde, M.; Sirsi, S. R.; Lutz, G. J. *Biomacromolecules* **2005**, *7*, 347–356.

(44) Zhong, Z.; Feijen, J.; Lok, M. C.; Hennink, W. E.; Christensen, L. V.; Yockman, J. W.; Kim, Y.-H.; Kim, S. W. *Biomacromolecules* **2005**, *6*, 3440–3448.

(45) Tang, G. P.; Zeng, J. M.; Gao, S. J.; Ma, Y. X.; Shi, L.; Li, Y.; Too, H.-P.; Wang, S. *Focus Biomater. Sci. Asia* **2003**, *24*, 2351–2362.

(46) Ogris, M.; Brunner, S.; Schüller, S.; Kircheis, R.; Wagner, E. *Gene Ther.* **1999**, *6*, 595–605.

(47) Xiong, M. P.; Bae, Y.; Fukushima, S.; Forrest, M. L.; Nishiyama, N.; Kataoka, K.; Kwon, G. S. *ChemMedChem* **2007**, *2*, 1321–1327.

(48) Rowan, S. J.; Cantrill, S. J.; Cousins, G. R. L.; Sanders, J. K. M.; Stoddart, J. F. *Angew. Chem. Int. Ed.* **2002**, *41*, 898–952.

(49) Wojtecki, R. J.; Meador, M. A.; Rowan, S. J. *Nat Mater* **2011**, *10*, 14–27.

(50) Lehn, J.-M. *Aust. J. Chem.* **2010**, *63*, 611–623.

(51) Maeda, T.; Otsuka, H.; Takahara, A. *Prog. Polym. Sci.* **2009**, *34*, 581–604.

(52) Lehn, J.-M. *Plenary Lect. World Polym. Congr. 40th IUPAC Int. Symp. Macromol.* **2005**, *30*, 814–831.

(53) Skene, W. G.; Lehn, J.-M. P. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 8270–8275.

(54) Ulrich, S.; Dumy, P. *Chem. Commun.* **2014**.

(55) Barnard, A.; Smith, D. K. *Angew. Chem. Int. Ed.* **2012**, *51*, 6572–6581.

(56) Joshi, A.; Vance, D.; Rai, P.; Thiyagarajan, A.; Kane, R. S. *Chem. – Eur. J.* **2008**, *14*, 7738–7747.

(57) Haag, R.; Kratz, F. *Angew. Chem. Int. Ed.* **2006**, *45*, 1198–1215.

(58) Lin, Y.-L.; Jiang, G.; Birrell, L. K.; El-Sayed, M. E. H. *Biomaterials* **2010**, *31*, 7150–7166.

(59) Fella, C.; Walker, G. F.; Ogris, M.; Wagner, E. *Eur. J. Pharm. Sci.* **2008**, *34*, 309–320.

(60) Walker, G. F.; Fella, C.; Pelisek, J.; Fahrmeir, J.; Boeckle, S.; Ogris, M.; Wagner, E. *Mol Ther* **2005**, *11*, 418–425.

(61) Dong, R.; Su, Y.; Yu, S.; Zhou, Y.; Lu, Y.; Zhu, X. *Chem. Commun.* **2013**, *49*, 9845–9847.

(62) Martínez, Á.; Bienvenu, C.; Jiménez Blanco, J. L.; Vierling, P.; Mellet, C. O.; García Fernández, J. M.; Di Giorgio, C. *J. Org. Chem.* **2013**, *78*, 8143–8148.

(63) Hardy, J. G.; Kostiainen, M. A.; Smith, D. K.; Gabrielson, N. P.; Pack, D. W. *Bioconjug. Chem.* **2005**, *17*, 172–178.

(64) Kostiainen, M. A.; Hardy, J. G.; Smith, D. K. *Angew. Chem. Int. Ed.* **2005**, *44*, 2556–2559.

(65) Uchida, H.; Miyata, K.; Oba, M.; Ishii, T.; Suma, T.; Itaka, K.; Nishiyama, N.; Kataoka, K. *J. Am. Chem. Soc.* **2011**, *133*, 15524–15532.

(66) Geihe, E. I.; Cooley, C. B.; Simon, J. R.; Kiesewetter, M. K.; Edward, J. A.; Hickerson, R. P.; Kaspar, R. L.; Hedrick, J. L.; Waymouth, R. M.; Wender, P. A. *Proc. Natl. Acad. Sci.* **2012**, *109*, 13171–13176.

(67) Qi, R.; Wu, S.; Xiao, H.; Yan, L.; Li, W.; Hu, X.; Huang, Y.; Jing, X. *J. Mater. Chem.* **2012**, *22*, 18915–18922.

(68) Klemm, K.; Radić Stojković, M.; Horvat, G.; Tomišić, V.; Piantanida, I.; Schmuck, C. *Chem. – Eur. J.* **2012**, *18*, 1352–1363.

(69) Kuchelmeister, H. Y.; Schmuck, C. *Chem. – Eur. J.* **2011**, *17*, 5311–5318.

(70) Hargrove, A. E.; Nieto, S.; Zhang, T.; Sessler, J. L.; Anslyn, E. V. *Chem. Rev.* **2011**, *111*, 6603–6782.

(71) Rehm, T. H.; Schmuck, C. *Chem. Soc. Rev.* **2010**, *39*, 3597–3611.

(72) Theodossiou, T. A.; Pantos, A.; Tsogas, I.; Paleos, C. M. *ChemMedChem* **2008**, *3*, 1635–1643.

(73) Pantos, A.; Tsogas, I.; Paleos, C. M. *Biochim. Biophys. Acta BBA - Biomembr.* **2008**, *1778*, 811–823.

(74) Wender, P. A.; Galliher, W. C.; Goun, E. A.; Jones, L. R.; Pillow, T. H. *Membr. Permeable Pept. Vectors Chem. Funct. Des. Ther. Appl.* **2008**, *60*, 452–472.

(75) Tamaru, S.; Hamachi, I. In *Recognition of Anions*; Vilar, R., Ed.; Structure and Bonding; Springer Berlin Heidelberg, 2008; Vol. 129, pp. 95–125.

(76) Blondeau, P.; Segura, M.; Perez-Fernandez, R.; de Mendoza, J. *Chem. Soc. Rev.* **2007**, *36*, 198–210.

(77) Tziveleka, L.-A.; Psarra, A.-M. G.; Tsiourvas, D.; Paleos, C. M. *J. Controlled Release* **2007**, *117*, 137–146.

(78) Goun, E. A.; Pillow, T. H.; Jones, L. R.; Rothbard, J. B.; Wender, P. A. *ChemBioChem* **2006**, *7*, 1497–1515.

(79) Sessler, J. L.; Gale, P.; Cho, W.-S.; Stoddart, J. F.; Rowan, S. J.; Aida, T.; Rowan, A. E. *Anion Receptor Chemistry*; Monographs in Supramolecular Chemistry; The Royal Society of Chemistry, 2006.

(80) Schmidtchen, F. P. *Anion Coord. Chem. II* **2006**, *250*, 2918–2928.

(81) Schug, K. A.; Lindner, W. *Chem. Rev.* **2005**, *105*, 67–114.

(82) Kubik, S.; Reyheller, C.; Stüwe, S. *J. Incl. Phenom. Macrocycl. Chem.* **2005**, *52*, 137–187.

(83) Funhoff, A. M.; van Nostrum, C. F.; Lok, M. C.; Fretz, M. M.; Crommelin, D. J. A.; Hennink, W. E. *Bioconjug. Chem.* **2004**, *15*, 1212–1220.

(84) Best, M. D.; Tobey, S. L.; Anslyn, E. V. *35 Years Synth. Anion Recept. Chem. 1968-2003* **2003**, *240*, 3–15.

(85) Oudrhiri, N.; Vigneron, J.-P.; Peuchmaur, M.; Leclerc, T.; Lehn, J.-M.; Lehn, P. *Proc. Natl. Acad. Sci.* **1997**, *94*, 1651–1656.

(86) Schmidtchen, F. P.; Berger, M. *Chem. Rev.* **1997**, *97*, 1609–1646.

(87) Seel, C.; Galán, A.; Mendoza, J. In *Supramolecular Chemistry II — Host Design and Molecular Recognition*; Weber, E., Ed.; Topics in Current Chemistry; Springer Berlin Heidelberg, 1995; Vol. 175, pp. 101–132.

(88) Dixon, R. P.; Geib, S. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **1992**, *114*, 365–366.

(89) Echavarren, A.; Galan, A.; Lehn, J. M.; De Mendoza, J. *J. Am. Chem. Soc.* **1989**, *111*, 4994–4995.

(90) Sen, J.; Chaudhuri, A. *J. Med. Chem.* **2005**, *48*, 812–820.

(91) Bhattacharya, S.; Bajaj, A. *Chem. Commun.* **2009**, 4632–4656.

(92) Mével, M.; Breuzard, G.; Yaouanc, J.-J.; Clément, J.-C.; Lehn, P.; Pichon, C.; Jaffrès, P.-A.; Midoux, P. *ChemBioChem* **2008**, *9*, 1462–1471.

(93) Vigneron, J. P.; Oudrhiri, N.; Fauquet, M.; Vergely, L.; Bradley, J. C.; Basseville, M.; Lehn, P.; Lehn, J. M. *Proc. Natl. Acad. Sci.* **1996**, *93*, 9682–9686.

(94) Matson, J. B.; Stupp, S. I. *Chem. Commun.* **2011**, *47*, 7962–7964.

(95) Belowich, M. E.; Valente, C.; Smaldone, R. A.; Friedman, D. C.; Thiel, J.; Cronin, L.; Stoddart, J. F. *J. Am. Chem. Soc.* **2012**, *134*, 5243–5261.

(96) Sclafani, J. A.; Maranto, M. T.; Sisk, T. M.; Van Arman, S. A. *J. Org. Chem.* **1996**, *61*, 3221–3222.

(97) Casi, G.; Huguenin-Dezot, N.; Zuberbühler, K.; Scheuermann, J.; Neri, D. *J. Am. Chem. Soc.* **2012**, *134*, 5887–5892.

(98) Simpson, M. G.; Pittelkow, M.; Watson, S. P.; Sanders, J. K. M. *Org. Biomol. Chem.* **2010**, *8*, 1181–1187.

(99) Manimala, J. C.; Anslyn, E. V. *Tetrahedron Lett.* **2002**, *43*, 565–567.

(100) Ulrich, S.; Buhler, E.; Lehn, J.-M. *New J. Chem.* **2009**, *33*, 271–292.

(101) Ulrich, S.; Lehn, J.-M. *Angew. Chem. Int. Ed.* **2008**, *47*, 2240–2243.

(102) Bagnacani, V.; Franceschi, V.; Bassi, M.; Lomazzi, M.; Donofrio, G.; Sansone, F.; Casnati, A.; Ungaro, R. *Nat Commun* **2013**, *4*, 1721.

(103) Nierengarten, I.; Nothisen, M.; Sigwalt, D.; Biellmann, T.; Holler, M.; Remy, J.-S.; Nierengarten, J.-F. *Chem. – Eur. J.* **2013**, *19*, 17552–17558.

(104) Bienvenu, C.; Martinez, A.; Jimenez Blanco, J. L.; Di Giorgio, C.; Vierling, P.; Ortiz Mellet, C.; Defaye, J.; Garcia Fernandez, J. M. *Org. Biomol. Chem.* **2012**, *10*, 5570–5581.

(105) Sigwalt, D.; Holler, M.; Iehl, J.; Nierengarten, J.-F.; Nothisen, M.; Morin, E.; Remy, J.-S. *Chem. Commun.* **2011**, *47*, 4640–4642.

(106) Méndez-Ardoy, A.; Guilloteau, N.; Di Giorgio, C.; Vierling, P.; Santoyo-González, F.; Ortiz Mellet, C.; García Fernández, J. M. *J. Org. Chem.* **2011**, *76*, 5882–5894.

(107) Sansone, F.; Baldini, L.; Casnati, A.; Ungaro, R. *New J. Chem.* **2010**, *34*, 2715–2728.

(108) Díaz-Moscoso, A.; Le Gourriérec, L.; Gómez-García, M.; Benito, J. M.; Balbuena, P.; Ortega-Caballero, F.; Guilloteau, N.; Di Giorgio, C.; Vierling, P.; Defaye, J.; Ortiz Mellet, C.; García Fernández, J. M. *Chem. – Eur. J.* **2009**, *15*, 12871–12888.

(109) Diaz-Moscoso, A.; Balbuena, P.; Gomez-Garcia, M.; Ortiz Mellet, C.; Benito, J. M.; Le Gourrierec, L.; Di Giorgio, C.; Vierling, P.; Mazzaglia, A.; Micali, N.; Defaye, J.; Garcia Fernandez, J. M. *Chem. Commun.* **2008**, 2001–2003.

(110) Sansone, F.; Dudič, M.; Donofrio, G.; Rivetti, C.; Baldini, L.; Casnati, A.; Cellai, S.; Ungaro, R. *J. Am. Chem. Soc.* **2006**, *128*, 14528–14536.

(111) Ulrich, S.; Boturyn, D.; Marra, A.; Renaudet, O.; Dumy, P. *Chem. – Eur. J.* **2014**, *20*, 34–41.

(112) Morgan, J. E.; Blankenship, J. W.; Matthews, H. R. *Arch. Biochem. Biophys.* **1986**, *246*, 225–232.

(113) Kircheis, R.; Wightman, L.; Wagner, E. *Triggering Drug Deliv. Syst.* **2001**, *53*, 341–358.

(114) Ohara, K.; Smietana, M.; Restouin, A.; Mollard, S.; Borg, J.-P.; Collette, Y.; Vasseur, J.-J. *J. Med. Chem.* **2007**, *50*, 6465–6475.

CHAPITRE 4 :

ETUDES DES INTERACTIONS GUANIDINIUMS / ADN

PAR CHIMIE COMBINATOIRE DYNAMIQUE

INTRODUCTION – RAPPELS BIBLIOGRAPHIQUES

La Chimie Combinatoire Dynamique[1–11] utilise les liaisons réversibles pour déterminer les constitutions de fragments stabilisées par une cible donnée. Cette méthode évite de longues synthèses et les nombreux essais avec chaque composé synthétisé qu'une approche combinatoire classique nécessite. En effet, il suffit de générer une bibliothèque de constituants à l'aide d'une ou plusieurs liaisons réversibles et de la tester directement en présence de cible. Si les bibliothèques sont générées dans des conditions adéquates (voir chapitre 1), le changement constitutionnel généré par l'interaction des constituants avec la cible conduit à l'amplification des constituants les plus stables en présence de cible, c'est-à-dire ceux qui ont la meilleure reconnaissance pour cette cible.

Une des difficultés majeures de la Chimie Combinatoire Dynamique consiste à trouver une méthode d'analyse appropriée pour pouvoir séparer et quantifier correctement les différents constituants formés au sein de la bibliothèque. Les principales méthodes analytiques utilisées sont l'HPLC et la LC-MS bien que dans certains cas, la spectrométrie de masse ou la RMN puissent également être envisagées.[12] Bien qu'une bibliothèque de plus 9000 constituants ait pu être analysée par LC-MS,[13] le nombre de constituants présents dans une bibliothèque peut rendre difficile leur séparation et / ou leur quantification. Pour contourner ce problème, des bibliothèques biphasiques[14,15] et l'utilisation de cibles[16–18] ou de fragments[19] immobilisés sur support solide peuvent être une solution.

La Chimie Combinatoire Dynamique a déjà été largement utilisée en présence d'acides nucléiques.[20] Les premiers exemples décrits dans la littérature ont permis de discriminer les ligands d'un complexe organométallique en fonction de leur affinité pour un double brin d'ADN immobilisé sur cellulose.[16,21] Une sélectivité dans la reconnaissance d'une structure particulière de l'ADN, entre duplexe et quadruplexe par exemple, a aussi pu être atteinte avec cette méthode.[22–24] Des modifications chimiques ont également pu être introduites à l'aide de liaisons réversibles sur un double brin d'ADN dans le but de le stabiliser.[25] Cependant, aucune des bibliothèques combinatoires dynamiques ciblant les acide nucléiques et décrites dans la littérature à l'heure actuelle ne repose sur une interaction avec le squelette phosphodiester en tant qu'origine de la reconnaissance des constituants de la bibliothèque pour la cible.

Ce type de reconnaissance n'a été que peu étudié par comparaison avec les autres modes d'interactions possibles avec l'ADN et peut être lié à sa mise en évidence bien plus

récente.[26,27] Pourtant, reconnaître le squelette phosphodiester de l'ADN plutôt que les bases pourrait permettre d'atteindre une reconnaissance universelle des brins d'ADN, indépendamment de leur séquence. Sonder les interactions secondaires capables de stabiliser un ligand reconnaissant un phosphodiester de l'ADN présente donc un grand intérêt pour pouvoir concevoir des ligands universels efficaces par la suite, voir des transporteurs d'ADN. En complément de l'approche rationnelle décrite dans le chapitre 2 et qui visait justement à mieux connaître les interactions gouvernant la formation d'auto-assemblages de petites molécules ligands d'anions avec l'ADN, une approche combinatoire dynamique peut permettre de compléter les observations déjà effectuées. En effet, la sélection par association sélective de fragments qu'elle implique peut conduire à l'identification des fragments stabilisant les complexes avec l'ADN. Bien qu'il demeure nécessaire de synthétiser et étudier les composés sélectionnés par cette méthode, elle est néanmoins plus efficace qu'une démarche rationnelle ou combinatoire « statique » pour une première sélection. Par rapport à la démarche utilisée précédemment dans le chapitre 2, cette approche apporte d'autres informations comme l'interaction pour l'ADN de composés dont les propriétés optiques ne permettent pas son étude spectroscopique ou encore le comportement en solution d'un mélange de composés mis en compétition en présence d'ADN et dont les différences d'ionisation ne permettraient pas une analyse en spectrométrie de masse.

Dans ce chapitre, des bibliothèques combinatoires dynamiques de petites molécules organisées autour d'un groupement guanidinium, connu comme ligand d'oxoanions, ont été générées par formation de liaisons imines entre les fragments, en absence et en présence de cibles oligonucléotides. Cette méthodologie permet, dans le cas où des effets de matrices sont observées, d'étudier l'influence des groupements adjacents du guanidinium dans la stabilisation des auto-assemblages constituants – ADN.

Les bibliothèques combinatoires dynamiques ont été conçues pour être construites à partir du monomère **Gua-Ald** dont la synthèse est décrite dans le chapitre précédent et de différentes amines *via* des liaisons imines qui vont permettre d'obtenir des constituants guanidinium bisubstitués en fonction des amines utilisées (Schéma 4.). La présence de cycles aromatiques sur le composé **Gua-Ald** va permettre sa détection et celles de ses dérivés par spectroscopie UV.

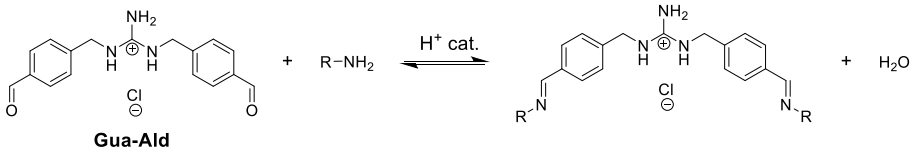


Schéma 4.1 : Formation générale d'imine à partir du composé **Gua-Ald**.

La présence de cycles aromatiques sur le composé **Gua-Ald** va permettre sa détection et celles de ses dérivés par spectroscopie UV (Figure 4.).

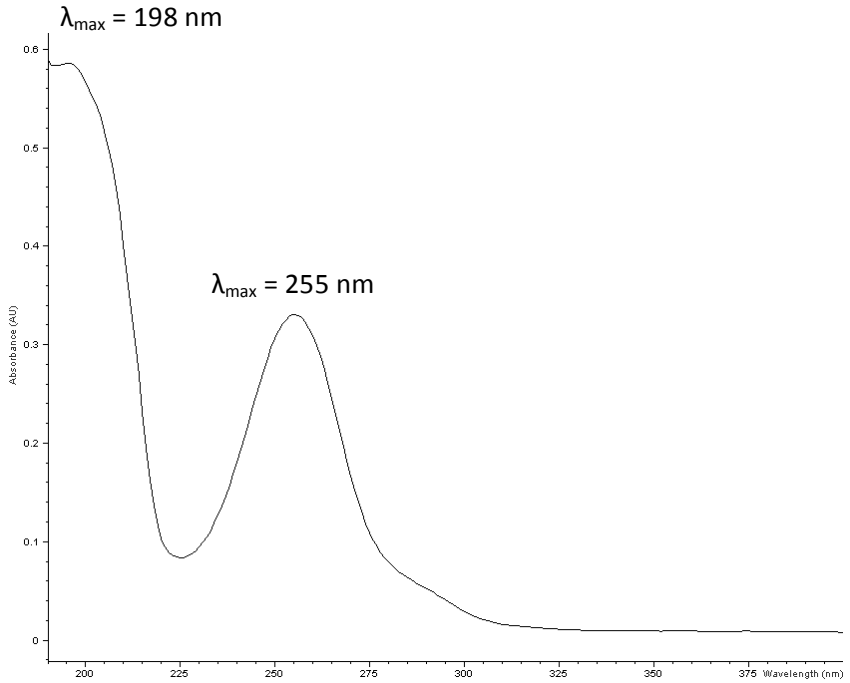


Figure 4.1: Spectre UV (190 – 400 nm) du composé **Gua-Ald** dans l'eau milli-Q.

Le choix de la liaison imine en tant que liaison réversible pour générer ces bibliothèques repose sur le fait que des bibliothèques combinatoires dynamiques ont déjà été générées avec cette liaison en présence d'ADN.[25] De plus, il est important de privilégier la formation de systèmes hétérotopiques pour sonder les interactions secondaires induites par les groupements adjacents du guanidinium : la liaison imine répond bien à ce critère puisqu'elle permet d'éviter la formation de dimères homotopiques contrairement à ce qui est observé avec la liaison disulfure (le produit **Gua-Ald** ne peut se lier avec lui-même, les fonctions aldéhydes ne pouvant pas réagir avec elles-mêmes). Par ailleurs, les amines disponibles commercialement sont nombreuses, ce qui évite des étapes de synthèse supplémentaires.

Toutefois, l'ADN étant soluble dans l'eau, les bibliothèques vont être générées en milieu aqueux, ce qui est une difficulté pour la formation d'imines. En effet, celles-ci ne sont généralement pas (ou peu) stables dans l'eau.[4,28–30] Pour contourner ce problème de stabilité, il est possible de procéder à une amination réductrice pour figer les équilibres.[1] Au cours de

ce type de procédure, du cyanoborohydrure de sodium, un réducteur compatible avec l'eau, est généralement ajouté dans les bibliothèques. En fonction des vitesses des différentes réactions, les bibliothèques peuvent être sous contrôle thermodynamique ou cinétique. En effet, si la vitesse de la réduction est plus lente que celles de la formation d'imine et de la trans-imination, la bibliothèque sera sous contrôle thermodynamique, le système pouvant corriger ses erreurs avant le figeage des équilibres par réduction. *A contrario*, si la réduction a lieu plus rapidement que la formation d'imine et la trans-imination, la bibliothèque sera sous contrôle cinétique et les constituants seront cinétiquement piégés avant que le système ait pu corriger ses erreurs. Dans le cas de ces bibliothèques sous contrôle cinétique, les bibliothèques perdent leur caractère dynamique proprement dit, mais si la cible utilisée induit un effet de matrice et stabilise des constituants par rapport à d'autres, des différences peuvent être observées et donner des indications sur l'affinité relative des différents constituants avec la cible tout comme ce qui serait obtenu avec une bibliothèque sous contrôle thermodynamique.[31] La réduction d'imines par le cyanoborohydrure de sodium étant une technique largement utilisée dans les bibliothèques combinatoires dynamiques basées sur les imines, la vitesse globale de l'amination réductrice a été suivie, mais nous n'avons pas suivi les échanges d'amines dans l'eau dans le cas de nos bibliothèques.

La reconnaissance d'anions par des petites molécules dans l'eau peut également s'avérer difficile, en particulier s'il n'y a pas d'effet multivalent tel que des interactions d'empilement pour stabiliser l'assemblage.[32–34] Dans le cadre de nos bibliothèques combinatoires dynamiques, les substituants vont donc être de première importance. En effet, leur introduction sur le constituant de départ fait varier les propriétés de celui-ci et des interactions supplémentaires peuvent se faire entre constituants et cible ADN ou entre les constituants eux-mêmes. L'affinité des constituants pour la cible peut ainsi être augmentée et les auto-assemblages formés plus stables qu'avec le composant de départ.

La possibilité d'une bisfonctionnalisation des constituants au sein des bibliothèques peut permettre d'augmenter la stabilité des complexes petites molécules – ADN dans l'eau et pourrait également permettre de générer des macrocycles ou des polymères / oligomères dans les bibliothèques par réaction du guanidinium bisaldéhyde **Gua-Ald** avec des diamines ou autres composés bisfonctionnalisés.

I. MISE AU POINT DE L'AMINATION REDUCTRICE DANS L'EAU

Lors de l'amination réductrice en milieu aqueux, la réduction directe des aldéhydes du composé **Gua-Ald** peut également avoir lieu, ce qui entraîne l'apparition de produits secondaires (Schéma 4.2). Dans une bibliothèque générée à partir du composé **Gua-Ald** et d'une seule amine, il y a en effet jusqu'à trois produits différents lorsque les réactions d'amination réductrice et de réduction directe sont complètes. La mise au point de l'étape d'amination réductrice vise donc à réduire la proportion de produit directement réduit puisqu'il ne présente pas d'intérêt pour nos bibliothèques combinatoires dynamiques.

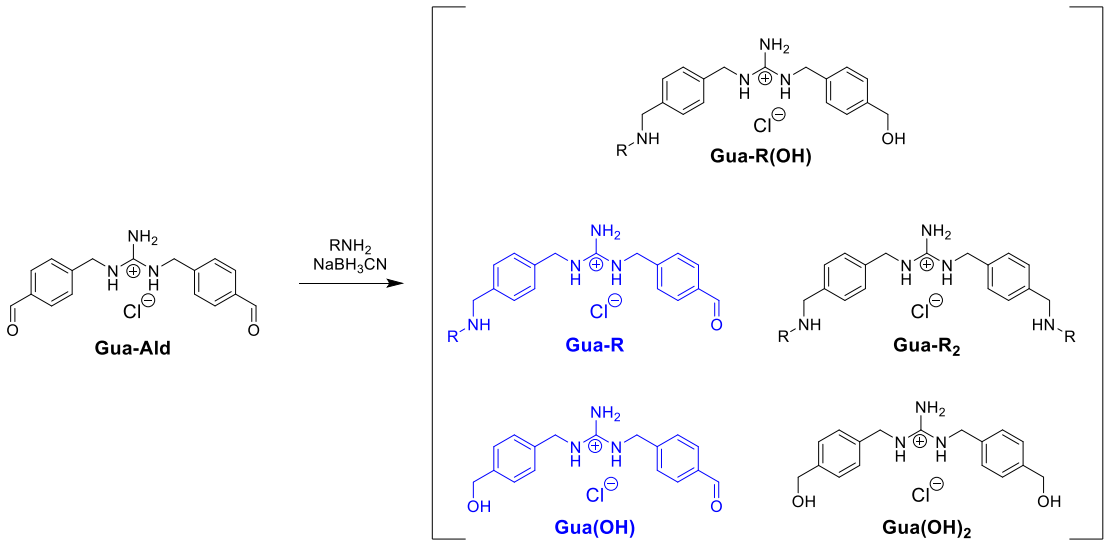


Schéma 4.2 : Constitution générale d'une bibliothèque générée par amination réductrice entre le composé **Gua-Ald** et une seule amine RNH_2. Les composés **Gua-R** et **Gua(OH)** indiqués en bleu sur le schéma sont des intermédiaires comportant encore une fonction aldéhyde qui n'a pas réagi. Les composés **Gua-R$_2$**, **Gua-R(OH)** et **Gua(OH)$_2$**, en noir, sont les composés finaux de la bibliothèque.

A. Conditions de reference et choix de l'amine pour la mise au point des conditions d'amination reductrice.

Les conditions de référence pour la mise au point de l'amination réductrice ont été définies comme étant les suivantes :

- Concentration en composé **Gua-Ald** fixée à 1 mM (concentration usuelle pour la génération de bibliothèques combinatoires dynamiques) ;
- 10 équivalents d'amines (lorsque plusieurs amines seront testées dans une même bibliothèque, il y aura forcément un excès d'amines par rapport aux fonctions aldéhydes ; de plus, cet excès d'amine permet de limiter les biais possibles liés à la consommation d'un fragment dans les bibliothèques) ;

- 10 équivalents de réducteur (cyanoborohydrure de sodium $NaBH_3CN$) ;
- Tampon phosphate pH 6,0 (le tampon permet de fixer le pH pour étudier son influence sur la constitution des bibliothèques et un pH 6,0 est compatible avec l'utilisation d'une cible ADN) ;
- Equilibration à température ambiante.

L'objectif des premières bibliothèques étant de mettre au point l'amination réductrice, elles ont bien évidemment été générées en absence de cible.

Les bibliothèques ont été préparées par dilution de différentes solutions mères de réactifs. En effet, en plus de restreindre les problèmes de reproductibilité, cette méthode est facile à mettre en œuvre pour des petits volumes, ce qui est rendu nécessaire par les faibles quantités disponibles des différents oligonucléotides utilisés comme cibles.

Les premiers tests d'amination réductrice ont été limités à la *n*-butylamine et la benzylamine. Lors de la génération des bibliothèques avec ces deux monoamines, des différences de réactivité ont clairement été observées (Figure 4.2). Ainsi, la proportion de produit de réduction directe est plus importante dans le cas de la *n*-butylamine que pour la benzylamine. Dans les deux cas, une proportion du composé **Gua-Ald** est directement réduite en alcool (produit **Gua(OH)$_2$**).

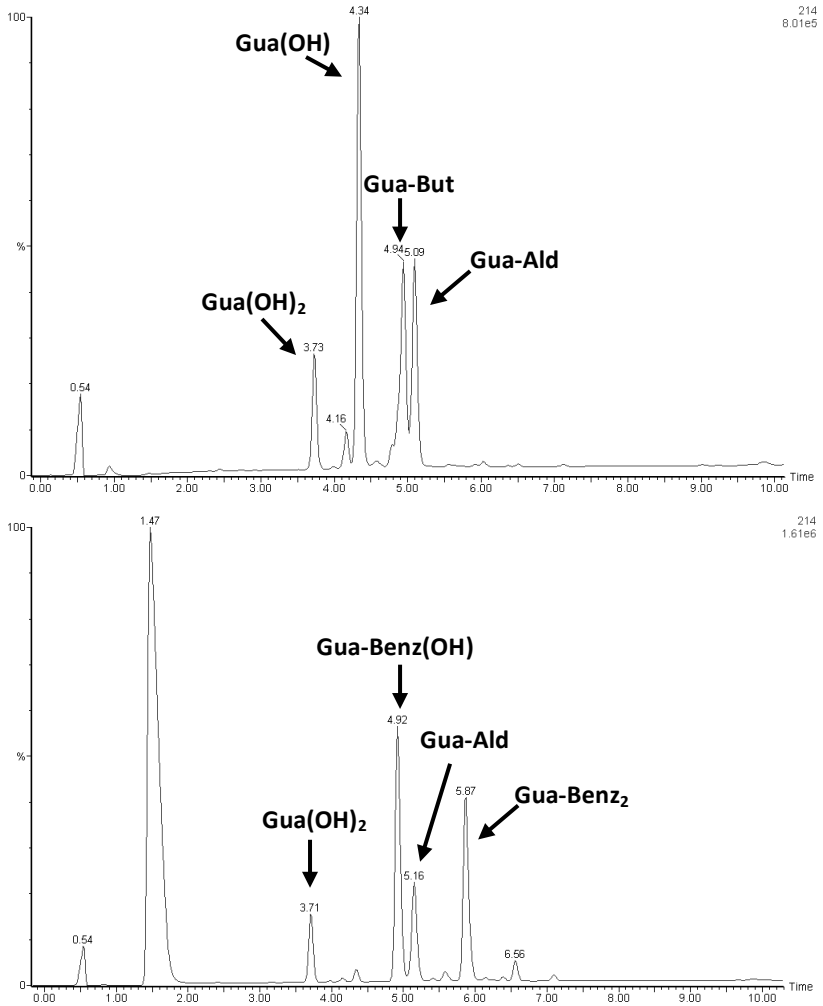


Figure 4.2 : Chromatogrammes LC-MS des bibliothèques préparées avec la *n*-butylamine (But, haut) ou la benzylamine (Benz, bas) dans les conditions de référence. Méthode IV sur la LC-MS 1, suivi à 214 nm.

La benzylamine donnant de meilleurs résultats en termes de réactivité et de cinétique, elle a été utilisée pour étudier l'influence des différents paramètres testés sur la génération des bibliothèques.

B. INFLUENCE DU pH ET DE LA NATURE DU TAMPON

Le pH est un paramètre important pour l'amination réductrice : avec le cyanoborohydrure comme agent de réduction, l'amination réductrice est cinétiquement favorisée à pH 6 – 7 tandis que la réduction directe des aldéhydes est plus rapide à pH 3 – 4.[35] Il est donc particulièrement important d'étudier l'influence des variations du pH sur la constitution des bibliothèques. Trois bibliothèques avec le composé **Gua-Ald** et la benzylamine ont donc été générées en présence de tampon phosphate ou acétate à différents pH, à savoir pH 5 (tampon acétate), pH 6 (tampon phosphate) et pH 7 (tampon phosphate).

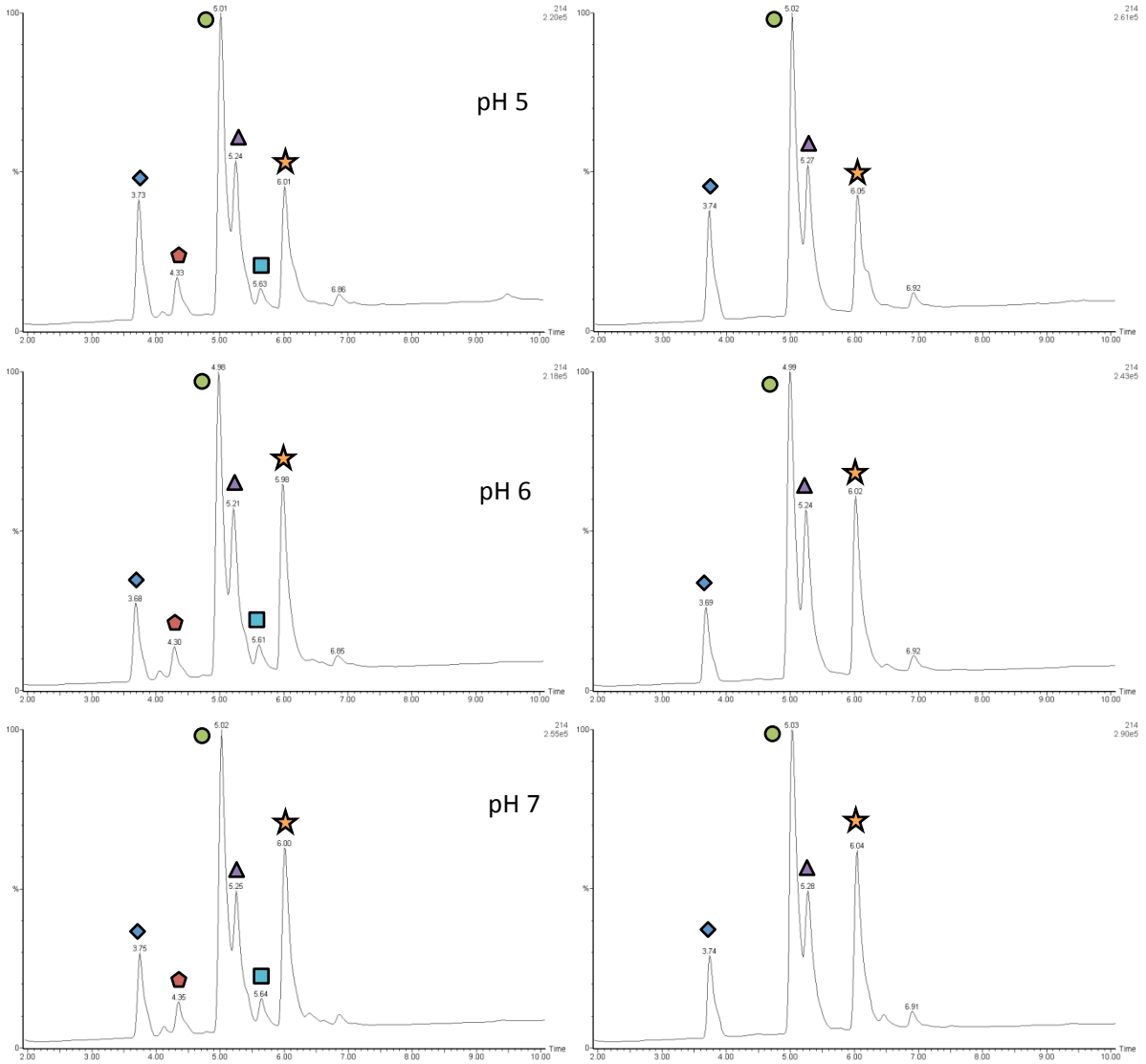


Figure 4.3 : Chromatogrammes LC-MS après 48 heures (gauche) et 7 jours (droite) d'amination réductrice des bibliothèques générées avec la benzylamine (Benz) dans les tampons acétate pH 5 (haut) et phosphate pH 6 (milieu) et pH 7 (bas). Méthode IV sur la LC-MS 1, suivi à 214 nm. Produits : ◆ **Gua(OH)$_2$** ; ⬟ **Gua(OH)** ; ● **Gua-Benz(OH)** ; ▲ **Gua-Ald** ; ■ **Gua-Benz** et ★ **Gua-Benz$_2$**.

Pour faciliter la comparaison des bibliothèques, la répartition des différents composés pouvant varier, la sélectivité de l'amination réductrice par rapport à la réduction directe (Équation 4.) a été calculée dans les trois exemples précédents (Tableau 4.).

Équation 4.1 : Formule générale de la sélectivité pour l'amination réductrice par rapport à la réduction $S_{A/R}$

$$S_{A/R} = \frac{Aire(monoaminés) + Aire(monoaminés\ réduits) + Aire(bisaminés)}{Aire(monoréduits) + Aire(bisréduits)}$$

$$S_{A/R} = \frac{A_{Gua-R} + A_{Gua-R(OH)} + A_{Gua-R_2}}{A_{Gua(OH)} + A_{Gua(OH)_2}}$$

Tableau 4.1 : Sélectivité $S_{A/R}$ à temps comparable en fonction du pH du tampon phosphate utilisé pour préparer les bibliothèques.

	Tampon acétate	**Tampon phosphate**	
	pH 5	**pH 6**	**pH 7**
$S_{A/R}$ après 36 heures	2,5	3,9	3,7
$S_{A/R}$ après 48 heures	3,0	4,9	4,6
$S_{A/R}$ après 7 jours	4,3	7,2	6,5

Dans le cas de la bibliothèque à pH 5, la proportion de produits d'amination réductrice est moins importante qu'à pH 6 ($S_{A/R}$ plus faible). A pH 7, la constitution de la bibliothèque est assez similaire à celle à pH 6. Le pH optimal pour procéder à l'amination réductrice avec un tampon phosphate semble donc être égal à 6.

Pour vérifier si la nature du tampon a une influence sur la constitution des bibliothèques, une bibliothèque a été générée en présence de tampon cacodylate pH 6 et comparée avec la bibliothèque dans le tampon phosphate pH 6. Au même temps de réaction, les chromatogrammes des deux bibliothèques sont similaires : la nature du tampon n'a donc pas d'influence sur la constitution des bibliothèques et la répartition des produits aminés et réduits.

La présence d'un tampon phosphate dans le milieu des bibliothèques, alors que leurs constituants visent à reconnaître les phosphodiesters des brins d'ADN, pourrait réduire, voire anéantir, les interactions possibles entre la cible et nos composés. Pour éviter ce problème, une première possibilité aurait été de changer la nature du tampon. Malheureusement, le choix de tampons usuels utilisables à pH 6 est relativement restreint et ces tampons contiennent le plus souvent des composés susceptibles d'interagir avec le groupement guanidinium. Ainsi, l'ion cacodylate, à la base du deuxième tampon testé, a une géométrie proche de l'ion phosphate (Figure 4.4) : une interaction entre le groupement guanidinium et l'ion cacodylate serait donc possible.

Figure 4.4 : Structure des ions cacodylate et phosphate.

Une seconde possibilité pour contourner ce problème de compétition du tampon avec la cible ADN consiste à ne pas utiliser de tampon. Un essai a été effectué en générant une bibliothèque dans l'eau milli-Q et comparé avec la bibliothèque dans le tampon phosphate à pH 6, l'eau ayant un pH autour de cette valeur (Figure 4.5). Toutefois, dans ce cas-là, les variations du pH ne sont plus limitées par le tampon. Pour éviter une trop importante variation de pH, la solution mère de benzylamine utilisée pour générer la bibliothèque a été préalablement neutralisée à pH 6. Cette neutralisation sera particulièrement importante par la suite pour pouvoir comparer des bibliothèques avec des amines différentes, le pH de leur solution mère pouvant être très différents. Quant au pH de la solution finale, il n'est pas connu avec précision puisque le volume de la solution est trop faible pour permettre l'utilisation d'un pH-mètre. Sa valeur n'a pu être mesurée qu'à l'aide de papier pH tribandes et est comprise entre 5 et 6.

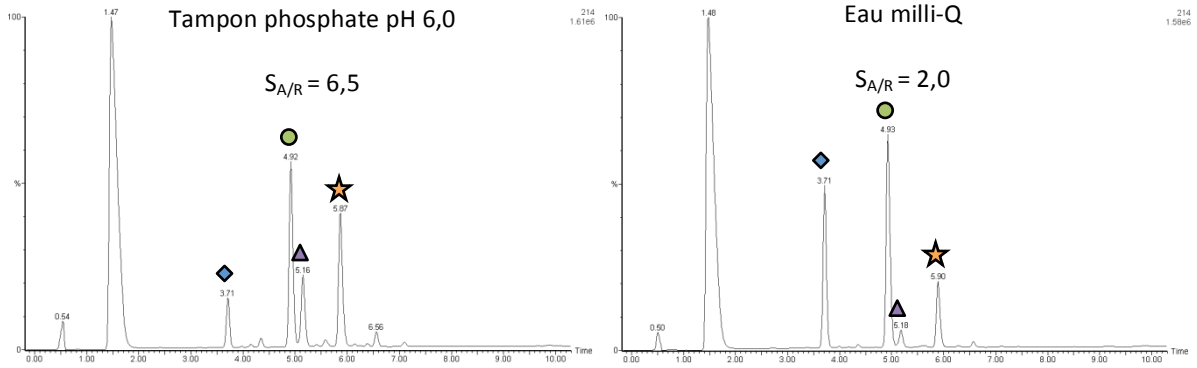


Figure 4.5 : Chromatogrammes LC-MS au bout de 3 jours d'amination réductrice des bibliothèques générées avec la benzylamine (Benz) dans le tampon phosphate pH 6 (gauche) et l'eau milli-Q avec neutralisation préalable de la benzylamine à pH 6 (droite). Méthode IV sur la LC-MS 1, suivi à 214 nm. Produits : ◆ **Gua(OH)$_2$** ; ● **Gua-Benz(OH)** ; ▲ **Gua-Ald** et ★ **Gua-Benz$_2$**.

Pour un même temps de réaction, en présence d'eau milli-Q, les proportions de réactif **Gua-Ald** encore présent et de produit bis-aminé **Gua-Benz$_2$** sont plus faibles par rapport à celles de la bibliothèque de référence et la proportion de produit de réduction simple **Gua(OH)$_2$** est plus importante. Des différences très claires dans l'avancement de la réaction sont observées : l'absence de tampon semble favoriser la réduction directe par rapport à l'amination réductrice.

La comparaison de cette bibliothèque générée dans l'eau milli-Q avec celle générée dans un tampon acétate pH 5 conduit au même constat, bien que le pH de la bibliothèque

générée dans l'eau seule soit compris entre 5 et 6. L'absence de tampon entraîne donc une augmentation de la cinétique de réduction directe, ce qui induit moins de constituants aminés dans les bibliothèques.

C. *INFLUENCE DU NOMBRE D'EQUIVALENTS D'AMINE*

Le nombre d'équivalents d'amine utilisé pour générer une bibliothèque peut modifier la sélectivité amination réductrice sur réduction directe en favorisant ou non la formation d'imines ou iminiums dans le milieu (suite au déplacement de l'équilibre suivant le principe de Le Châtelier). Une bibliothèque a donc été générée en utilisant 100 équivalents de benzylamine plutôt que les 10 équivalents des conditions de référence (Figure 4.6). Sur les chromatogrammes, la proportion de produit bis-aminé **Gua-Benz$_2$** augmente fortement avec le nombre d'équivalents d'amine : l'amination réductrice est donc logiquement plus favorisée avec un nombre d'équivalents élevé d'amine.

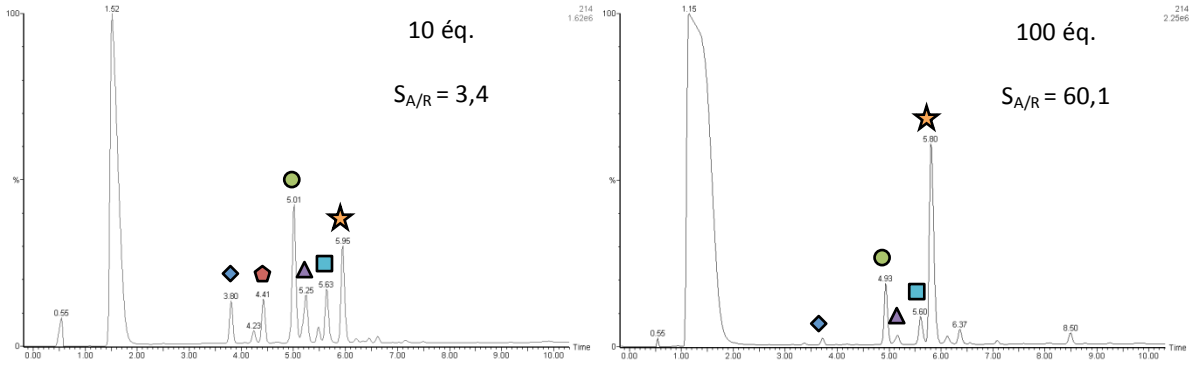


Figure 4.6 : Chromatogrammes LC-MS au bout de 24 heures d'amination réductrice des bibliothèques générées avec 10 équivalents (gauche) et 100 équivalents de benzylamine (droite) dans le tampon phosphate pH 6. Méthode IV sur la LC-MS 1, suivi à 214 nm. Produits : ◆ **Gua(OH)$_2$** ; ⬟ **Gua(OH)** ; ● **Gua-Benz(OH)** ; ▲ **Gua-Ald** ; ■ **Gua-Benz** et ★ **Gua-Benz$_2$**.

Dans le cas où une seule amine est utilisée, l'utilisation d'un nombre d'équivalent d'amine aussi élevé que 100 équivalents ne semble pas appropriée pour générer des bibliothèques combinatoires dynamiques. En effet, la formation d'un produit majoritaire en proportion aussi importante que celle du produit bis-aminé **Gua-Benz$_2$** observée ici avec la benzylamine pourrait rendre difficilement détectable les variations induites par l'effet de matrice de la cible sur la formation des produits d'amination réductrice (à moins que ce constituant majoritaire soit fortement défavorisé en présence de cible). En présence de

plusieurs amines en compétition, un tel nombre d'équivalent d'amines ne devrait pas poser problème si les réactivités des amines sont équivalentes les unes par rapport aux autres.

Des raisons d'ordre pratique s'opposent également à l'usage d'un tel nombre d'équivalents. Dans un premier temps, il faut pouvoir préparer une solution mère suffisamment concentrée pour préparer les bibliothèques par dilution et en gardant la possibilité de générer des bibliothèques avec plusieurs amines. Or, toutes les amines n'ont pas la même solubilité dans l'eau : préparer une solution mère à 1 M dans l'eau pourrait donc se révéler ardu dans certains cas.

Dans un second temps, l'excès d'amine peut poser problème lors des analyses LC-MS. En effet, l'amine seule peut être retenue sur la colonne et visible sur le chromatogramme comme dans le cas de la benzylamine, ce qui entraîne la présence d'un pic supplémentaire qu'il ne sera pas forcément possible d'atténuer en jouant sur la longueur d'onde à laquelle le chromatogramme est considéré (l'exemple de la benzylamine entre dans cette catégorie). Lorsque des bibliothèques avec plusieurs amines seront préparées, il est possible que le pic de l'amine soit proche du pic de l'un des constituants. Dans ce cas, l'usage d'un large excès d'amine pourrait entraîner un mauvais retour à la ligne de base entre les deux pics ou dans le pire des cas, une superposition des pics. Même en modifiant la méthode d'analyse, la marge de manœuvre pourrait se révéler étroite.

*D. INFLUENCE DE LA CONCENTRATION EN GUANIDINIUM **GUA-ALD***

La concentration est un paramètre influant sur la vitesse de réaction, elle influe donc sur la vitesse de génération des bibliothèques. En effet, plus la concentration est élevée, plus la probabilité pour deux molécules de se rencontrer est importante : la probabilité qu'une réaction se produise est donc plus élevée.

La concentration est également un paramètre critique pour les interactions entre constituants de la bibliothèque et la cible. L'existence et la force de telles interactions déterminent les constantes d'association des différents constituants avec la cible.

En considérant l'équilibre d'un constituant A et de la cible B formant un complexe AB (Figure 4.7), la constante d'association est définie par l'Équation 4.2.

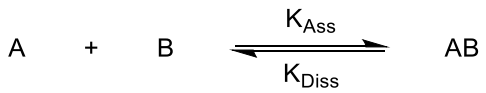


Figure 4.7 : Equilibre de formation d'un complexe AB à partir des réactifs A et B.

Équation 4.2

$$K_{Ass} = \frac{1}{K_{Diss}} = \frac{[AB]}{[A][B]}$$

Si la concentration est trop faible, le complexe ne sera pas formé. Par contre, si la concentration d'un réactif augmente et que la stœchiométrie est conservée (la concentration du deuxième réactif est donc également augmentée), la concentration de complexe augmentera à température et pression constante. Varier la concentration des constituants d'une bibliothèque peut donc favoriser la reconnaissance de la cible et l'apparition de changements constitutionnels.

Une bibliothèque avec la benzylamine a été générée en augmentant la concentration de guanidinium **Gua-Ald** de 1 à 10 mM (Figure 4.8). Pour des raisons pratiques, la stœchiométrie de l'amine a été réduite de 10 à 5 équivalents.

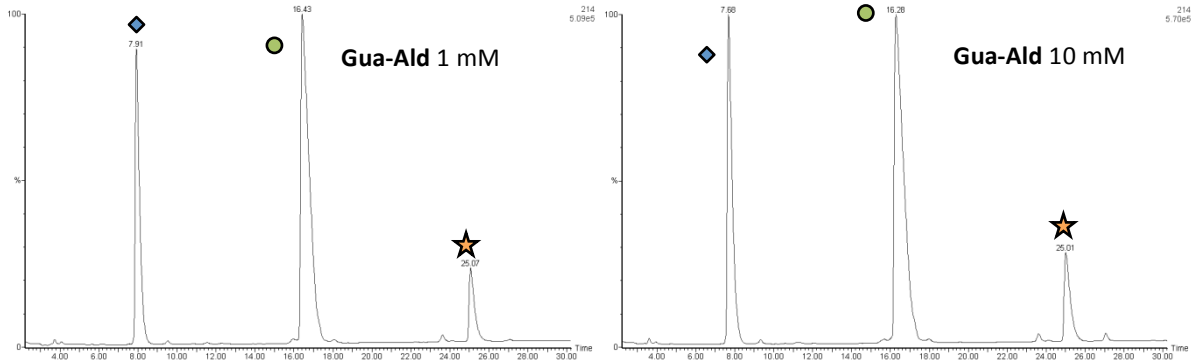


Figure 4.8 : Chromatogrammes LC-MS des bibliothèques générées avec une concentration en **Gua-Ald** de 1 mM (gauche) et 10 mM (droite) dans l'eau à pH 7,0. Méthode V sur la LC-MS 1, suivi à 214 nm. Produits : ◆ **Gua(OH)$_2$** ; ● **Gua-Benz(OH)** et ★ **Gua-Benz$_2$**.

L'amination réductrice est beaucoup plus rapide à 10 mM qu'à 1 mM, le réactif **Gua-Ald** n'étant plus détecté au bout de 13 heures, alors qu'il est encore détecté au bout de ce même temps de réaction à 1 mM. Par contre, l'augmentation de la concentration initiale en guanidinium n'influence pas la répartition entre amination réductrice et réduction directe. En effet, bien qu'une comparaison directe des aires des constituants ne soient pas possible en raison de la différence de concentration en guanidinium entre les deux bibliothèques, le calcul

de la sélectivité $S_{A/R}$ (Équation 4.) pour chacune des bibliothèques montre que la répartition des différents constituants est similaire en fin de réaction quelle que soit la concentration utilisée (Tableau 4.2).

Tableau 4.2 : Sélectivités $S_{A/R}$ en fonction de la concentration initiale en **Gua-Ald**.

Concentration initiale en Gua-Ald	**1 mM**	**10 mM**
$\mathbf{S_{A/R}}$	2,4	2,2

L'analyse de la bibliothèque générée à plus haute concentration et avec 5 équivalents d'amine montre donc qu'une fois l'amination réductrice terminée, la constitution de la bibliothèque est similaire à celle d'une bibliothèque générée à la concentration de référence avec 10 équivalents d'amine. Cela n'exclue cependant pas une influence de la concentration sur la constitution des bibliothèques en présence de cible.

E. INFLUENCE DE LA TEMPERATURE

La température est un paramètre bien connu pour influencer la vitesse de réaction. En réduisant l'agitation thermique, la vitesse de réaction est réduite. Toutefois, la cinétique n'a pas été suivie.

En absence de cible, l'étude de l'influence de la température a pour seul intérêt de déterminer si elle influence la sélectivité entre amination réductrice et réduction directe. Une bibliothèque est donc générée à 4°C et comparée avec une bibliothèque équivalente générée à température ambiante (Figure 4.9). Lors de la génération de cette bibliothèque à 4°C, aucun précipité n'est apparu dans le mélange : aucun constituant n'est donc piégé sous forme solide dans la bibliothèque.

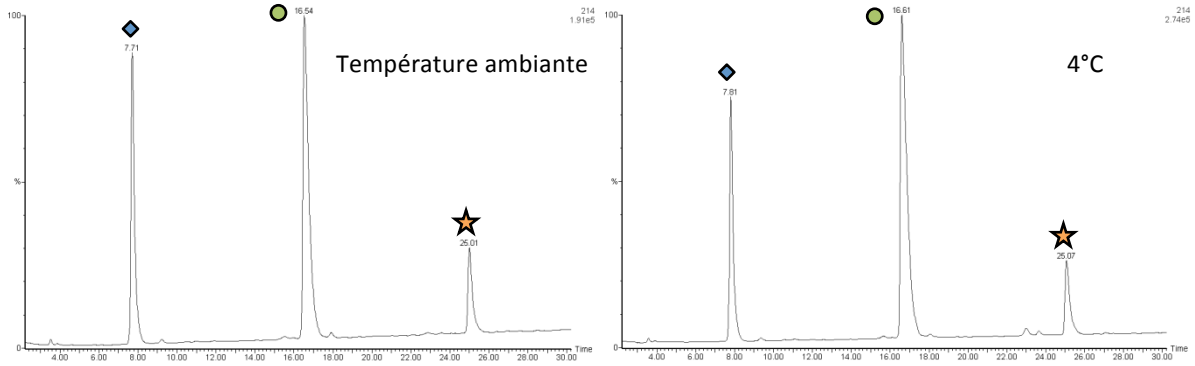


Figure 4.9 : Chromatogrammes LC-MS des bibliothèques générées avec la benzylamine (Benz) à température ambiante (gauche) et à 4°C (droite) dans l'eau à pH 7,0 avec une concentration en **Gua-Ald** de 10 mM. Méthode V sur la LC-MS 1, suivi à 214 nm. Produits : ◆ **Gua(OH)$_2$** ; ● **Gua-Benz(OH)** et ★ **Gua-Benz$_2$**.

A concentration identique, il est possible de comparer directement les aires des différents constituants formés en fin de réaction. De faibles différences sont visibles, correspondant aux erreurs et incertitudes de mesure. La sélectivité $S_{A/R}$ à 4°C est identique à celle observée dans les mêmes conditions à température ambiante (Tableau 4.3).

Tableau 4.3 : Sélectivité $S_{A/R}$ en fonction de la température de réaction dans les différentes conditions testées.

Conditions	4°C	TA
Gua-Ald 1 mM	3,5	3,5
Gua-Ald 10 mM	2,2	2,4

La température n'a donc pas d'influence sur la sélectivité amination sur réduction finale des bibliothèques en absence de cible.

F. INFLUENCE D'UN PALIER D'EQUILIBRATION AVANT REDUCTION

Dans le but de favoriser l'amination réductrice, un temps d'équilibration peut être introduit avant d'ajouter le cyanoborohydrure de sodium dans la bibliothèque : un tel palier d'équilibration pourrait permettre au mélange de s'organiser et à une quantité plus importante d'imines ou iminiums de se former.

Une bibliothèque a été générée en introduisant un palier de plusieurs heures à température ambiante avant d'ajouter le cyanoborohydrure de sodium au mélange. Le palier d'équilibration ne change pas la constitution finale de la bibliothèque par rapport à une

bibliothèque similaire générée en introduisant immédiatement le réducteur après l'ajout des autres réactifs.

En fonction de la stabilité des imines formées, il peut être possible d'analyser directement une bibliothèque d'imines dans un système eau – acétonitrile – 0,1% TFA sans passer par l'étape de réduction.[36] Toutefois, un test effectué avec le guanidinium **Gua-Ald** et la benzylamine dans un milieu à pH 7 (tampon phosphate ou eau) ne nous a pas permis d'observer autre chose que le composé **Gua-Ald** en HPLC. La réduction par le cyanoborohydrure de sodium est donc nécessaire dans le cas de nos bibliothèques.

Des analyses RMN ^{1}H d'un mélange de produit **Gua-Ald** en présence de 10 équivalents d'amines ont été effectuées dans l'eau deutérée pour déterminer si les imines sont formées dans l'eau avant la réduction (pD du mélange non mesuré). Au bout de 24 heures, aucune trace d'imine n'est visible et la fonction aldéhyde du réactif **Gua-Ald** (δ 9,93 ppm) est toujours visible sans que l'intégration du pic n'ait par ailleurs véritablement diminué.

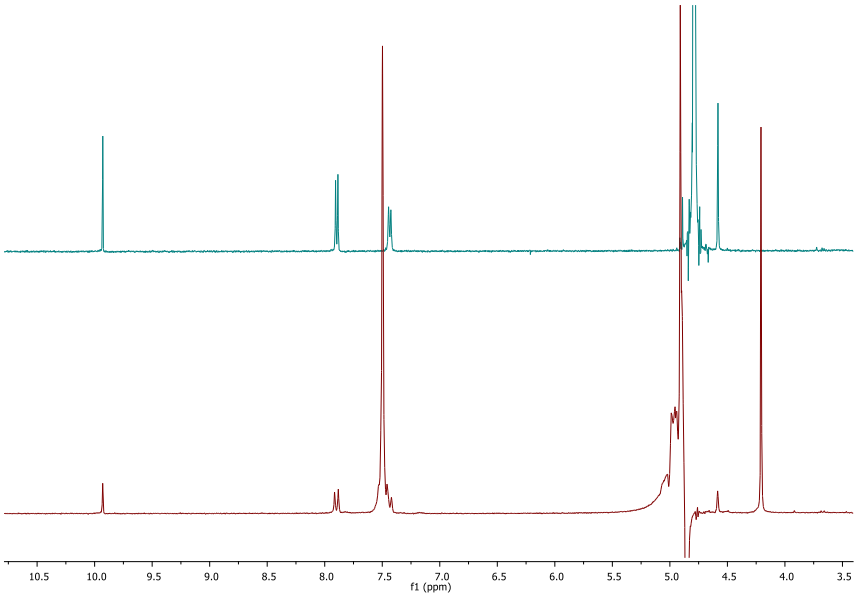
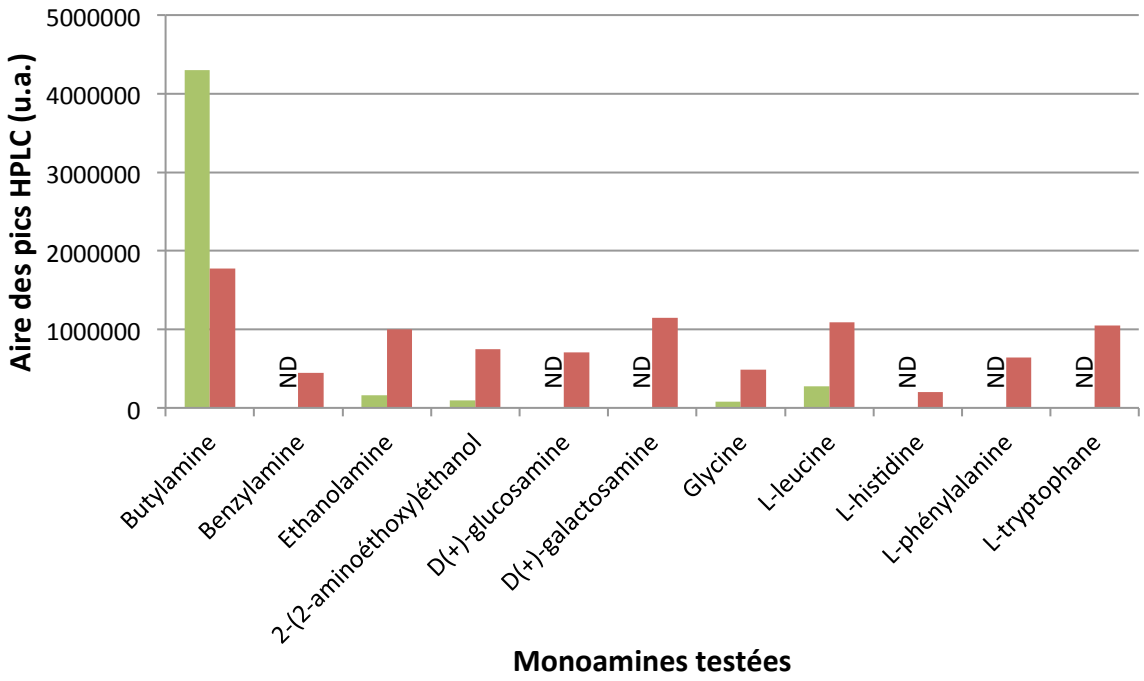


Figure 4.10 : Comparaison des spectres RMN ^{1}H du composé **Gua-Ald** dans D_2O (haut) et du mélange de **Gua-Ald** avec 10 équivalents de benzylamine après 24 heures (bas) dans D_2O/H_2O après présaturation de l'eau (pulse 50).

Pour la génération des bibliothèques, l'ajout d'un temps d'équilibration avant la réduction n'est donc pas nécessaire, la formation d'imines étant l'étape cinétiquement déterminante. Toutefois, la même constitution finale des bibliothèques en absence de cible n'exclut pas des différences possibles en présence de cible. Des différences ne sont pas non plus à exclure en présence de plusieurs amines, leur réactivité pouvant différer (d'où la nécessité de les pondérer au maximum en jouant sur les nombres d'équivalent introduits).

G. ETUDE DE LA REACTIVITE DES MONOAMINES

Différentes monoamines ont été testées séparément en absence de cible dans un tampon phosphate pH 7 plutôt que de l'eau à pH 7 pour éliminer toute possibilité de différences de pH entre les bibliothèques (Graphique 4.). Les différences sont bien marquées entre les bibliothèques et l'étude de la tendance est suffisante pour les premières constatations. En effet, en comparant les aires obtenues pour les produits de réduction **Gua(OH)** et **Gua(OH)$_2$** à un même temps de réaction pour toute la série de monoamines testée, il est facile de repérer celles en présence desquelles la réduction directe est plus importante dans la bibliothèque (ce qui corrèle avec une plus faible réactivité de l'amine vis-à-vis de l'amination réductrice).



Graphique 4.1 : Aire des pics HPLC du réactif **Gua-Ald** et des produits formés, pour un même temps de réaction (4 jours), en fonction des différentes amines utilisées pour générer les bibliothèques. Sont indiqués : en vert : aire du pic du réactif **Gua-Ald** ; en rouge : somme de l'aire des pics des produits réduits **Gua(OH)** et **Gua(OH)$_2$**. Seules les aires des produits de réduction directe **Gua(OH)** et **Gua(OH)$_2$** ainsi que l'aire du réactif **Gua-Ald** résiduel sont présentés car comparables entre les différentes bibliothèques (coefficient d'extinction molaire différents pour les produits aminés, d'où une impossibilité de comparer leur aire tant qu'ils ne sont pas connus).

Sur le graphique précédent, la *n*-butylamine se différencie particulièrement des autres monoamines testées par son absence de réactivité au bout de 4 jours de réaction. Aucun produit d'amination n'est détecté dans la bibliothèque et lors des différents essais effectués par la suite, le produit bis-aminé **Gua-But$_2$** n'a jamais été observé avec la *n*-butylamine contrairement aux autres amines. L'absence de corrélation entre la réactivité et le pK$_a$ des amines, déjà connue dans la littérature,[30] est observée ici. En effet, bien que la *n*-butylamine ait un pK$_a$ plus élevée que celui de la benzylamine (10,60 contre 9,34 dans l'eau[37]), l'amine la plus réactive des deux est la benzylamine. Cette absence de corrélation entre pK$_a$ et réactivité des amines peut être due aux différences des niveaux énergétiques HOMO-LUMO de ces deux amines.[30] Les différences d'hydratation pourraient également avoir une influence mais n'étant pas quantifiées pour chaque amine, il n'est pas possible de confirmer cette hypothèse.

L'étude de bibliothèques générées avec différentes monoamines permet donc de constater des différences de réactivité, mais celles-ci demeurent difficiles à rationaliser.

H. CONDITIONS CHOISIES POUR LES PREMIERES BIBLIOTHEQUES ET MISE AU POINT DES METHODES HPLC / LC-MS

Les conditions choisies pour les premières bibliothèques correspondent aux conditions de références utilisées, à l'exception du tampon phosphate remplacé par de l'eau pour éviter une compétition avec les phosphodiesters de l'ADN lorsque les bibliothèques seront générées en présence de cible :

- Concentration en composé **Gua-Ald** de 1 mM ;
- 10 équivalents d'amine ;
- 10 équivalents de cyanoborohydrure de sodium ajouté immédiatement ;
- Eau à pH 7,0 ;
- Equilibration à température ambiante.

Pour l'analyse des bibliothèques, une méthode HPLC générale a été mise au point pour optimiser au mieux la séparation des différents constituants et pouvoir facilement faire des comparaisons. L'idée est ici d'optimiser au maximum la méthode d'analyse pour ne plus avoir à la changer pour chaque bibliothèque avec différentes amines par la suite.

L'éluant utilisé est un système eau – acétonitrile – 0,1% TFA, usuel pour les peptides mais pas pour les oligonucléotides. Ce système avait déjà été utilisé pour purifier le composé **Gua-Ald** par HPLC semi-préparative (cf. chapitre 3) et il est également utilisé en routine sur

les HPLC et LC-MS du laboratoire. Ce sont pour ces deux raisons que l'analyse des bibliothèques a d'abord été essayée avec ce système d'élution. La séparation et la forme des pics étant correctes dans ce système, il a été conservé pour la suite. Toutefois, il faudra veiller à ce que ce système convienne pour l'analyse des bibliothèques générées en présence d'ADN.

Des bibliothèques ont donc été préparées en absence de cible avec plusieurs amines différentes (Figure 4.11). Le guanidinium **Gua-Ald** étant un composé polaire, son temps de rétention en HPLC phase inverse est relativement faible. La fonctionnalisation par des amines fait varier cette polarité de manière plus ou moins importante : les différents produits obtenus peuvent être proches en HPLC. Pour optimiser leur séparation, le gradient a été diminué et ralenti.

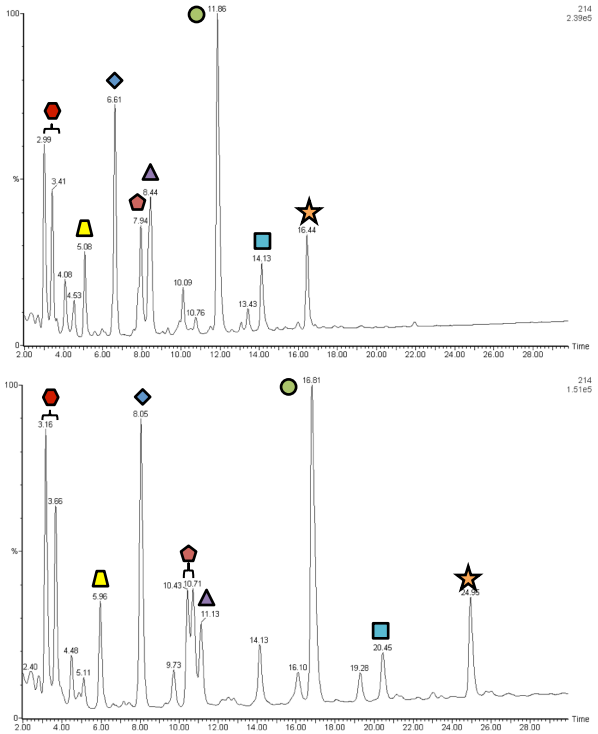


Figure 4.11 : Chromatogrammes LC-MS des bibliothèques générées avec la butylamine (But-NH_2), la benzylamine (Benz), le 2-(2-aminoéthoxy)éthanol (EE) et la D(+)-glucosamine (Gluco) à température ambiante dans l'eau à pH 7,0 avec une concentration en **Gua-Ald** de 1 mM. Méthode IV (haut) et V (bas) sur la LC-MS 1, suivi à 214 nm. Produits : ⬢ **Gua-Gluco(OH)** ; △ **Gua-EE(OH)** ; ◇ **Gua(OH)$_2$** ; ⬠ **Gua(OH)** / **Gua-EE(But)** ; ▲ **Gua-(Benz)(Gluco)** ; ○ **Gua-Benz(OH)** ; □ **Gua-Benz** et ☆ **Gua-Benz$_2$**.

Des analyses en LC-MS (utilisant la méthode mise au point en HPLC) ont permis d'identifier des constituants formés au sein des bibliothèques, notamment des constituants bis-aminés mixtes **Gua-(A)(B)** (constituants bis-aminés par une amine A d'un côté et une amine B de l'autre).

II. ESSAIS AVEC dT_{10} ET dT_{40} EN SOLUTION

Une fois les conditions optimales trouvées pour effectuer l'amination réductrice avec le composé **Gua-Ald** en milieu aqueux, les premières bibliothèques[i] ont été générées en présence de cibles. Les oligonucléotides utilisés, le dT_{10} et le dT_{40}, ont été choisis parce que les seuls sites de reconnaissance connus avec le groupement guanidinium sur ce type d'oligonucléotides sont les phosphodiesters de l'oligonucléotide.[38] De plus, les oligothymines ont déjà été utilisés pour matricer des auto-assemblages de petites molécules.[39–42]

*A. BIBLIOTHEQUES AVEC UNE CONCENTRATION EN COMPOSE **Gua-Ald** DE 1 MM*

Plusieurs bibliothèques ont été générées avec la benzylamine (10 équivalents) en tant que seule amine, en présence de dT_{10} ou de dT_{40}, et systématiquement comparées avec les bibliothèques témoins correspondantes générées en absence de cible. Dans ces différentes bibliothèques, plusieurs paramètres ont été modifiés pour trouver les meilleures conditions pour induire des amplifications claires entre les bibliothèques.

1. Difficultés de quantification des constituants des bibliothèques en présence d'ADN

Les oligonucléotides absorbant en UV, les dT_{10} et dT_{40} utilisés dans les bibliothèques sont visibles en HPLC. Dans le système eau – acétonitrile – 0,1% TFA utilisé en routine sur les appareils HPLC et LC-MS, les pics de ces deux oligonucléotides sont larges et traînants (Figure 4.12). Or, cela peut gêner la quantification de certains pics de produits, en l'occurrence le constituant bis-aminé **Gua-Benz$_2$** formé avec la benzylamine.

[i] Le terme de bibliothèque est utilisé bien qu'une seule amine soit ajoutée et que le nombre de constituants soient limités (trois produits seulement avec une seule amine).

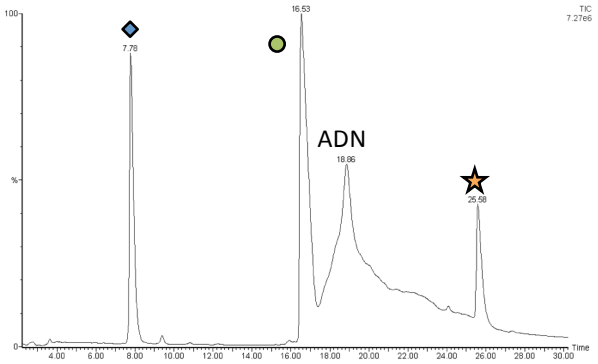


Figure 4.12 : Chromatogramme LC-MS d'une bibliothèque générée avec la benzylamine (Benz) en présence de dT_{10} dans l'eau à pH 7. Méthode V sur la LC-MS 1, suivi UV sur la totalité des longueurs d'ondes de 200 à 400 nm (addition des spectres UV). Produits : ◆ **Gua(OH)$_2$** ; ● **Gua-Benz(OH)** et ★ **Gua-Benz$_2$**.

Les pics d'oligonucléotides dans le système d'éluant eau – acétonitrile – TFA n'ayant pas une forme correcte et équilibrée, des analyses ont été effectuées avec un système d'éluant eau – acétonitrile – acétate de triéthylammonium à pH 7, connu pour être adapté à l'analyse d'oligonucléotide (Figure 4.13). Malheureusement, bien que ce système d'éluant permette d'avoir des pics d'ADN fins et équilibrés, les constituants guanidiniums des bibliothèques ont, pour un gradient d'élution similaire, des temps de rétention beaucoup trop faibles pour permettre l'analyse efficace des bibliothèques combinatoires dynamiques. Le système eau – acétonitrile – TFA a donc été conservé pour effectuer les analyses des bibliothèques.

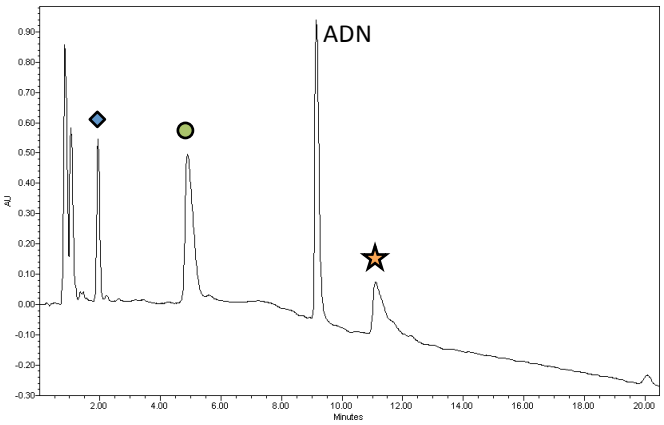


Figure 4.13 : Chromatogramme HPLC de la bibliothèque générée avec la benzylamine (Benz) en présence de dT_{10} dans l'eau à pH 7. Méthode : mode isocratique éluant A''-B'' 100%-0% pendant 2 minutes, puis gradient linéaire en éluant B'' dans l'éluant A'' de 0 à 30% en 20 minutes, puis de 30 à 100% en 4 minutes (HPLC 1). Suivi à 214 nm. Produits : ◆ **Gua(OH)$_2$** ; ● **Gua-Benz(OH)** et ★ **Gua-Benz$_2$**.

Pour limiter le problème de superposition du pic d'ADN avec ceux des constituants, une longueur spécifique des constituants guanidiniums (λ_{max} = 198 nm et 255 nm pour le réactif **Gua-Ald**) a été recherchée pour quantifier ces derniers en minimisant le signal de l'ADN. Malheureusement, le composé **Gua-Ald** n'ayant pas été conçu pour absorber au

minimum d'absorbance de l'ADN (longueur d'onde supérieure ou égale à 300 nm), les constituants guanidiniums n'absorbent pas dans cette gamme de longueur d'onde. Toutefois, une analyse à 214 nm permet de limiter le signal de la cible sur les chromatogrammes des bibliothèques et rend donc la quantification possible, à défaut d'être extrêmement précise (Figure 4.14).

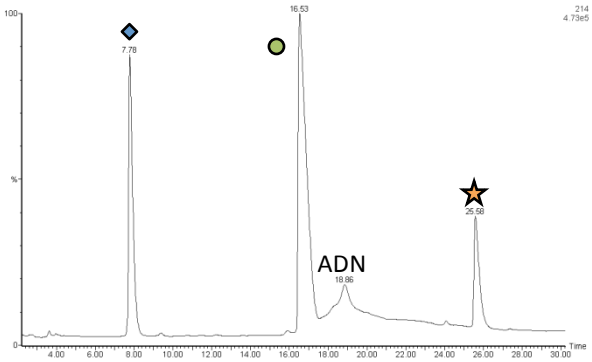


Figure 4.14 : Chromatogramme LC-MS de la bibliothèque générée avec la benzylamine (Benz) en présence de dT_{10} dans l'eau à pH 7. Méthode V sur la LC-MS 1, suivi à 214 nm. Produits : ◆ **Gua(OH)$_2$** ; ● **Gua-Benz(OH)** et ★ **Gua-Benz$_2$**.

Toutes les analyses ultérieures ont donc été effectuées dans le système eau – acétonitrile – TFA et pour les échantillons contenant des oligonucléotides, la longueur d'onde a été fixée à 214 nm pour la quantification des aires des produits.

2. Influence du solvant

Deux bibliothèques ont été générées en présence de dT_{10} dans l'eau à pH 7 et le tampon phosphate pH 7 (100 mM) et comparées avec leur bibliothèques témoins respectives générées en absence de dT_{10} (Équation 4.3, Tableau 4.4).

Équation 4.3 : Quantification des différences d'aire pour un même constituant entre deux bibliothèques.

$$Amplification(constituant\ X) = \frac{Aire\ X(échantillon) - Aire\ X(témoin)}{Aire\ X(témoin)}$$

Tableau 4.4 : Influence du tampon phosphate sur l'amplification des constituants des bibliothèques en présence de dT_{10} par rapport aux bibliothèques correspondantes générées en absence de dT_{10}.

Solvant	**Amplifications en présence de dT_{10}**		
	Gua(OH)$_2$	**Gua-Benz(OH)**	**Gua-Benz$_2$**
Eau pH 7,0	+7%	-16%	-55%
Tampon phosphate pH 7,0	-3%	-8%	-9%

Lorsque le tampon phosphate est utilisé, la présence de dT_{10} n'induit aucun changement de la constitution de la bibliothèque, alors que de faibles variations sont visibles pour les bibliothèques générées dans l'eau à pH 7. L'absence de changement constitutionnel en présence de tampon phosphate pourrait s'expliquer par la possible compétition des ions phosphates du tampon avec les phosphodiesters du dT_{10} pour la reconnaissance des constituants guanidiniums. Toutefois, les changements sont faibles par rapport aux variations induites par les erreurs et incertitudes de mesures pour les produits **Gua(OH)$_2$** et **Gua-Benz(OH)** des bibliothèques générées dans l'eau : il reste nécessaire de démontrer l'existence d'un effet de matrice dans les bibliothèques en présence d'ADN.

Des essais de relargage ont été effectués avec le chlorure de sodium sur ces différentes bibliothèques, mais n'ont pas permis d'observer de différences notables par rapport aux chromatogrammes déjà obtenus.

3. Influence de la longueur de l'oligonucléotide et de sa stœchiométrie

La longueur de l'oligonucléotide peut avoir une influence sur la stabilité des auto-assemblages formés dans les bibliothèques.[40,41] En effet, des interactions secondaires telles que des empilements peuvent stabiliser les constituants le long de la matrice. L'approche rationnelle (chapitre 2) a d'ailleurs démontré que la stabilité des auto-assemblages **GuaBiPy** – ADN dépend de la longueur des matrices oligonucléotides utilisées (pour une séquence identique). Si les bibliothèques sont générées dans les conditions appropriées de manière à ce que l'amplification reflète la stabilité des auto-assemblages formés, des différences pourraient donc être observées en fonction de la longueur de l'oligonucléotide présent.

La stœchiométrie utilisée entre les fragments et la cible introduite influence les amplifications observées dans les bibliothèques.[12] Les conditions optimales doivent être trouvées pour permettre l'observation d'amplifications importantes qui soient également représentatives de l'affinité des constituants pour la cible. En effet, augmenter la concentration de cible peut permettre d'observer des amplifications plus importantes. Toutefois, l'utilisation d'un large excès de cible peut également biaiser les bibliothèques en faveur de constituants n'ayant pas la meilleure affinité pour la cible mais dont la formation permet de diminuer l'enthalpie libre totale du système, les quantités de fragments présents étant alors limitées.

Différentes bibliothèques ont donc été générées en présence de dT_{10} ou de dT_{40} avec des stœchiométries de 1 ou 10 bases T par réactif **Gua-Ald**, soit 1 ou 10 phosphodiesters par constituant guanidinium (Tableau 4.5).

Tableau 4.5 : Influence de la longueur de l'oligonucléotide et de la stœchiométrie en oligonucléotide sur la constitution des bibliothèques générées dans l'eau à pH 7,0.

		Amplifications en présence de cible		
Oligothymine	**Equivalents de bases T par réactif Gua-Ald**	**$Gua(OH)_2$**	**Gua-Benz(OH)**	**$Gua\text{-}Benz_2$**
dT_{10}	1	+7%	-16%	-55%
	10	+8%	-45%	*
dT_{40}	1	-29%	-49%	-62%
	10	-56%	-55%	*

*Produit difficilement discernable car masqué par le pic d'ADN.

Des changements constitutionnels par rapport aux bibliothèques témoins sont visibles dans tous les cas et les observations suivantes peuvent être faites : plus l'oligonucléotide est long, plus les changements sont importants tandis que la stœchiométrie a une influence plus limitée. Toutefois, les variations observées sont relativement faibles.

De plus, alors qu'aucune précipitation n'est visible dans les bibliothèques avec cible, tous les produits diminuent par rapport aux bibliothèques témoins, ce qui n'est pas logique. En effet, lorsque certains produits diminuent dans une bibliothèque combinatoire dynamique, d'autres constituants sont amplifiés, ce qui n'est pas le cas ici. Il serait possible d'expliquer cette tendance si les produits formés, ou au moins certains d'entre eux, étaient retenus sur la cible ADN. Malheureusement, l'ajout de chlorure de sodium comme agent chaotrope dans les différentes bibliothèques n'a pas permis de confirmer cette hypothèse et aucun produit ne semble retenu sur l'ADN. L'ajout d'agents compétiteurs comme des ions phosphates ou d'autres agents chaotropes comme le chlorure de guanidinium tout comme la dénaturation par ajout d'urée n'ont pas été testés.

L'inhibition des réactions de réduction et d'amination réductrice en présence de cible peut également être une explication : bien qu'aucun pic de réactif **Gua-Ald** ne soit visible sur les chromatogrammes des bibliothèques en présence d'ADN, il est possible que du produit soit encore présent dans le mélange mais ne soit pas discernable du pic d'ADN encore visible à 214 nm.

4. Influence de la température

La température est un paramètre important pour la stabilité d'un auto-assemblage avec l'ADN.[40] La stabilité des doubles brins d'ADN est d'ailleurs caractérisée par une température de demi-fusion T_m (melting temperature) qui est la température à laquelle 50% de la totalité des molécules d'ADN d'une séquence donnée est hybridée sous forme de double brin, les 50% restants demeurant dissociés sous forme de simple brin.

D'après les lois de la thermodynamique, toute constante d'équilibre est dépendante de la température (Équation 4.4). Ce constat s'applique évidemment aux constantes d'association caractérisant des phénomènes de reconnaissance tels que ceux sur lesquels repose la génération de bibliothèques combinatoires dynamiques.

Équation 4.4

$$\Delta_r G^0 = -RTln(K)$$

Avec :

- $\Delta_r G^0$ enthalpie libre standard de réaction ($J.mol^{-1}$) ;
- R constante des gaz parfaits (~8,314 $J.mol^{-1}.K^{-1}$) ;
- T température (K) ;
- K constante d'équilibre.

En fonction des paramètres thermodynamiques de l'équilibre, une diminution de la température peut entraîner une augmentation de la constante d'équilibre. Par exemple, un double brin d'ADN se dissocie en simple brins lorsque la température augmente. La plus grande stabilité des auto-assemblages guanidiniums – ADN à faible température a également été démontrée avec l'approche rationnelle présentée précédemment (chapitre 2).

Générer des bibliothèques à plus faible température pourrait permettre d'observer des changements plus importants en présence de cible par une meilleure stabilisation des assemblages formés. Des bibliothèques ont donc été préparées à 4°C et comparées aux bibliothèques générées à température ambiante (Tableau 4.6).

Tableau 4.6 : Amplifications observées pour les différentes bibliothèques générées à 4°C dans l'eau à pH 7,0 (équilibration de 21 heures à 4°C avant ajout du réducteur $NaBH_3CN$).

		Amplification		
Oligothymine	**Equivalents en base**	**Gua(OH)$_2$**	**Gua-Benz(OH)**	**Gua-Benz$_2$**
dT$_{10}$	1	+16%	-7%	-29%
	10	-34%	-41%	-46%
dT$_{40}$	1	-28%	-42%	-72%
	10	-56%	-61%	-78%

Peu de différences sont observées par rapport aux mêmes bibliothèques générées à température ambiante. Par exemple, les amplifications observées pour les bibliothèques en présence de dT_{40} sont identiques à 4°C et à température ambiante. La tendance à la diminution de tous les constituants dans les bibliothèques avec cible est retrouvée. De nouveau, l'ajout d'agents chaotropes comme le chlorure de sodium ou le chlorure de guanidinium n'ont pas permis d'observer le décrochage de constituants retenus sur l'ADN. La diminution de la température ne permet donc pas d'observer de variations plus importantes entre les bibliothèques avec et sans cible.

*B. BIBLIOTHEQUES AVEC UNE CONCENTRATION EN COMPOSE **Gua-Ald** DE 10 MM*

Les premières bibliothèques avec une concentration en réactif **Gua-Ald** de 1 mM n'ayant pas données de résultats probants, une seconde série de bibliothèques est générée en augmentant la concentration de réactif **Gua-Ald** à 10 mM. Pour ne pas préparer de nouvelles solutions d'amines, la stœchiométrie de celles-ci a été réduite à 5 équivalents par composé **Gua-Ald**.

De même que précédemment, plusieurs bibliothèques ont été générées avec la benzylamine comme seule amine, en présence de dT_{10} ou de dT_{40} (1 équivalent en base T par guanidinium), dans l'eau à pH 7,0. Des bibliothèques ont été générées à température ambiante et à 4°C (Tableau 4.7).

Tableau 4.7 : Amplification des constituants dans les bibliothèques avec cible en fonction de la longueur de l'oligonucléotide et de la température de réaction (concentration en guanidinium **Gua-Ald** initiale de 10 mM).

		Amplification		
Oligothymine	**Température**	**Gua(OH)$_2$**	**Gua-Benz(OH)**	**Gua-Benz$_2$**
dT$_{10}$	Ambiante	-15%	-9%	-35%
	4°C	-22%	-41%	-13%
dT$_{40}$	Ambiante	-11%	-12%	+76%
	4°C	-6%	-9%	+69%

Avec le dT$_{10}$, la constitution des bibliothèques change assez peu par rapport à celles des bibliothèques témoins correspondantes, quelle que soit la température à laquelle elles ont été générées, et aucune tendance claire ne se dégage.

Avec le dT$_{40}$, une amplification du produit bis-aminé **Gua-Benz$_2$** est visible à 4°C comme à température ambiante. La quantification de l'aire du pic du produit bis-aminé **Gua-Benz$_2$** n'est pas la plus précise que l'on puisse obtenir (le pic de l'ADN demeurant visible à 214 nm et se superposant assez partiellement au pic mesuré), mais l'amplification calculée est tout de même représentative des phénomènes observés.

Malheureusement, lorsque les bibliothèques ont été générées avec d'autres lots de dT$_{40}$, les amplifications observées dans les premières bibliothèques n'ont plus été obtenues par la suite (Tableau 4.8). Les contradictions observées pour deux lots différents de dT$_{40}$ ne permettent donc pas de conclure sur les essais avec cette cible. Par ailleurs, il n'est pas exclu que ce problème de reproductibilité des résultats soit également rencontré avec les bibliothèques générées en présence de dT$_{10}$, qui n'ont pas été répétées car les variations entre bibliothèques avec cible et témoin étaient beaucoup moins importantes que celles observées avec le dT$_{40}$.

Tableau 4.8 : Amplifications observées pour des bibliothèques générées en présence de dT$_{40}$ en fonction du lot de l'oligonucléotide (**Gua-Ald** 10 mM, benzylamine, eau à pH 7,0).

	Amplification		
Lot de dT$_{40}$	**Gua(OH)$_2$**	**Gua-Benz(OH)**	**Gua-Benz$_2$**
DP-01-011-4-F1	-11%	-12%	+76%
Eurogentec lot 5127733	-23%	-40%	+8%

Le lot de dT$_{40}$ est le seul paramètre qui a été changé entre les deux séries de bibliothèques. Les autres solutions sont identiques et le mode opératoire de la première série

de bibliothèques a été reproduit à la lettre. Les deux lots de dT_{40} incriminés ont donc été analysés par HPLC dans le système eau – acétonitrile – acétate de triéthylammonium pH 7,0 (Figure 4.15). Or, leur temps de rétention respectif diffère à concentration similaire. Cette différence de temps de rétention peut s'expliquer si les deux oligonucléotides ne sont pas de même nature ni de même longueur, ou si l'un des deux s'est dégradé, ou encore si des appariements T-T[43,44] changent la conformation d'un des dT_{40} (ce qui entraîne un repliement des simples brins sur eux-mêmes et change la disponibilité des phosphodiesters pour les guanidiniums des bibliothèques).

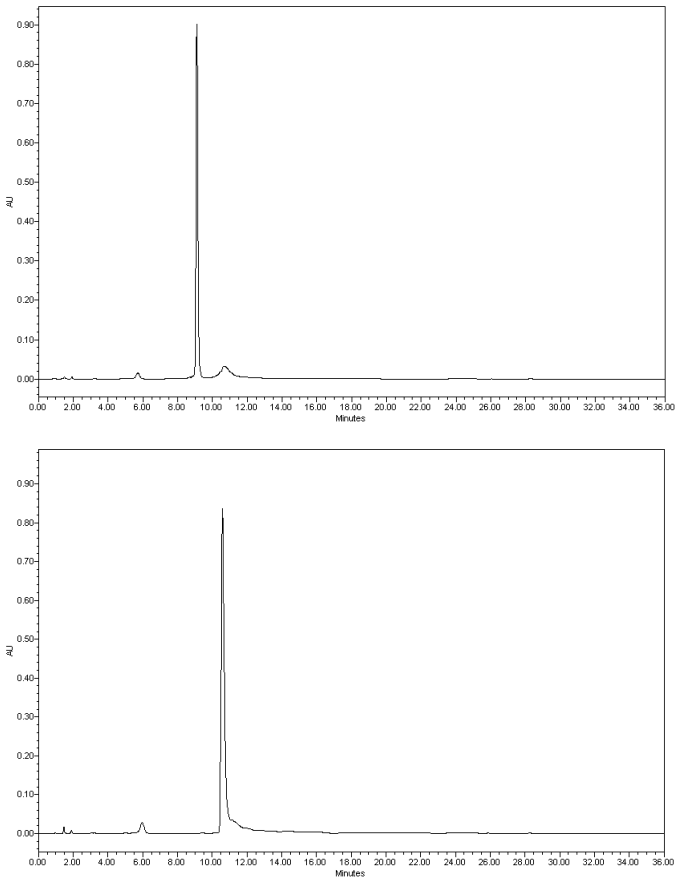


Figure 4.15 : Chromatogrammes HPLC de deux lots de dT_{40}. En haut, le lot DP-01-011-4-F1 avec lequel des amplifications ont été observés dans les bibliothèques ; en bas, le lot Eurogentec 5127733 avec lequel aucune amplification n'a pu être reproduite. Méthode : mode isocratique éluant A''-B'' 100%-0% pendant 2 minutes, puis gradient linéaire en éluant B'' dans l'éluant A'' de 0 à 30% en 20 minutes, puis de 30 à 100% en 4 minutes (HPLC 1), suivi à 254 nm.

Dans le cas où des appariements T-T auraient lieu, un chauffage préalable de l'ADN devrait permettre de déstabiliser ces appariements T-T et de revenir à une conformation de simple brin. L'analyse HPLC du dT_{40} Eurogentec lot 5127733 après un cycle de chauffage – refroidissement a permis d'observer que le profil HPLC de l'oligonucléotide était modifié

sans toutefois correspondre à celui du dT_{40} DP-01-011-4-F1 (Figure 4.16). D'ailleurs, une nouvelle bibliothèque avec la benzylamine, générée en présence du dT_{40} Eurogentec (lot 5127733) préalablement chauffé à 90°C pendant 30 minutes, présente une diminution significative du produit bis-aminé **Gua-Benz$_2$** en présence de dT_{40}, à l'inverse de ce qui a été observé dans les bibliothèques précédentes.

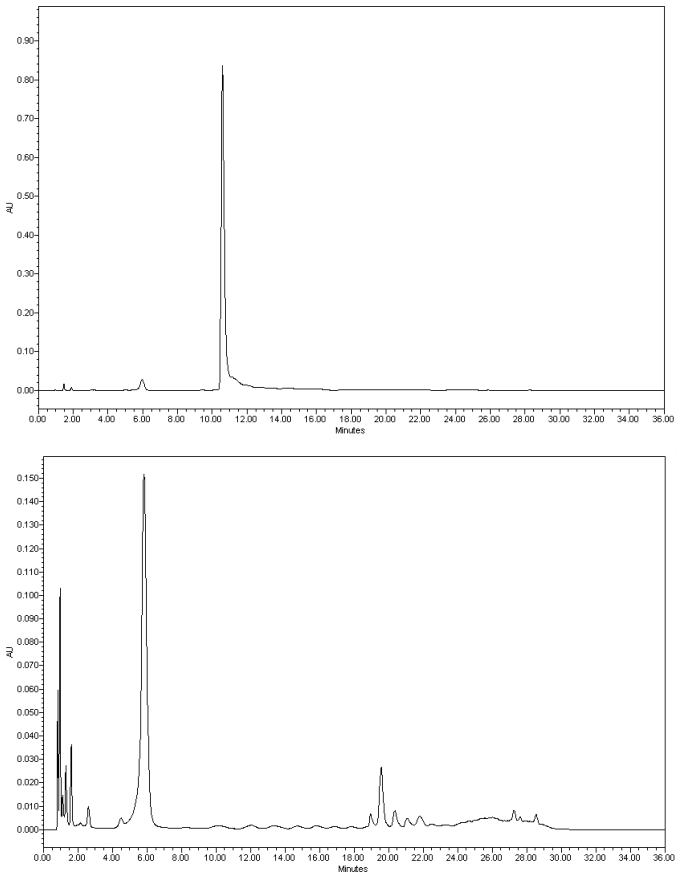


Figure 4.16 : Chromatogrammes HPLC du dT_{40} Eurogentec lot 5127733 avant et après chauffage à 90°C pendant 30 minutes. Méthode : mode isocratique éluant A''-B'' 100%-0% pendant 2 minutes, puis gradient linéaire en éluant B'' dans l'éluant A'' de 0 à 30% en 20 minutes, puis de 30 à 100% en 4 minutes (HPLC 1) ; suivi à 254 nm.

Aux vues des faibles amplifications obtenues et des problèmes de reproductibilité rencontrés, les résultats obtenus sont loin d'être satisfaisants. De plus, la présence d'un pic d'ADN visible en HPLC lors de l'utilisation d'oligonucléotides en solution rend difficile la quantification précise des constituants. Pour résoudre ce problème, il faudrait pouvoir retirer la cible de la bibliothèque avant l'analyse. Pour se faire, plusieurs méthodes sont envisageables. Il est ainsi possible de procéder à une dialyse de la bibliothèque pour séparer l'ADN des petites molécules ou d'utiliser un ADN modifié par de la biotine ou un support solide.

III. ESSAIS AVEC ADN/DT SUR CELLULOSE

L'utilisation d'une cible sur support solide[16–18] a paru intéressante à tester car elle est facile à mettre en œuvre et permet de travailler avec des volumes relativement faibles.

Différents types de supports solides peuvent être utilisés pour immobiliser un ADN. Ce peut être un polymère comme le polystyrène ou la cellulose, ou une bille magnétique. L'important dans le cadre de nos bibliothèques combinatoires dynamiques est d'avoir un support solide inerte vis-à-vis des constituants des bibliothèques et de l'amination réductrice : c'est *a priori* le cas pour les trois supports solides cités plus hauts. Le choix du support solide s'est fait en fonction de la disponibilité commerciale des différents ADN et de leur prix ainsi que des produits et appareillages disponibles au laboratoire. La cellulose a été utilisée car il est facile de se procurer différents ADN immobilisés sur ce support, et aucun appareillage particulier n'est nécessaire.

A. CONDITIONS ET METHODE DE GENERATION DES BIBLIOTHEQUES

Deux types d'ADN sur cellulose ont pu être testés : un ADN de thymus de veau simple brin et un oligo(dT) simple brin. Les bibliothèques générées avec les ADN sur support solide ont systématiquement été comparées aux bibliothèques générées sans cible en solution (témoins) et à celles générées en présence du support solide seul (Figure 4.17). Ces bibliothèques avec la cellulose microcristalline seule permettent de vérifier que le support solide n'a pas (ou peu) d'influence sur la constitution finale de la bibliothèque.

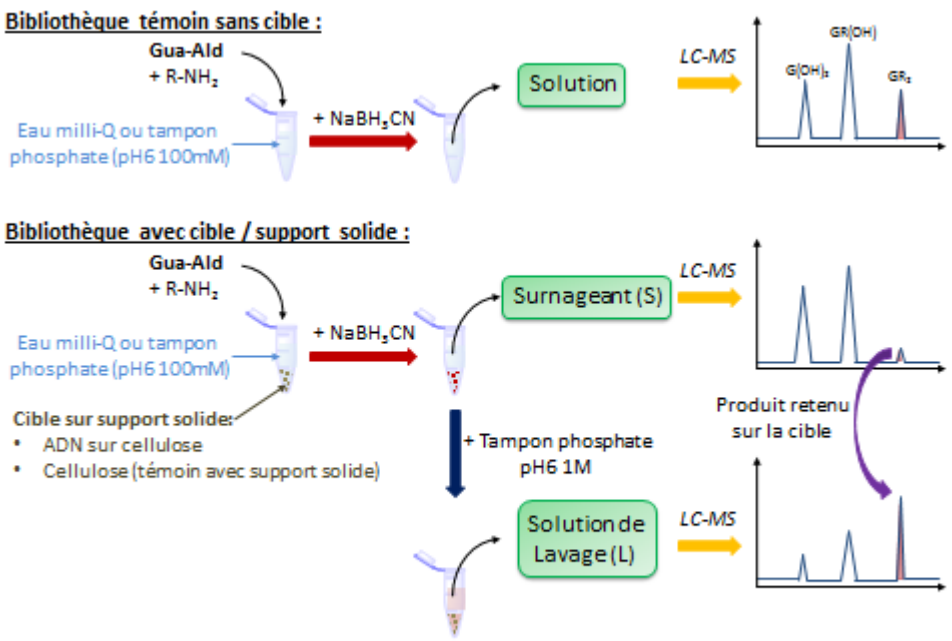


Figure 4.17 : Illustration schématique de la procédure suivie pour générer et analyser les différentes bibliothèques.

Les surnageants des bibliothèques étant analysés en LC-MS, des lavages des supports solides ont également été systématiquement effectués et analysés pour déterminer si des constituants étaient sélectivement retenus sur la cible ou le support solide.

La concentration en réactif **Gua-Ald** a été fixée à 1 mM et les bibliothèques ont été générées dans l'eau milli-Q. Des bibliothèques ont également été générées dans un tampon phosphate pH 6,0 pour déterminer si une compétition entre les ions du tampon et les phosphodiesters de l'ADN avait lieu pour l'interaction avec les constituants guanidiniums.

La quantité d'ADN introduite dans les bibliothèques a quant à elle été estimée à environ 1 équivalent en phosphodiesters, donc 1 équivalent en bases. Il a donc préalablement fallu estimer la quantité d'ADN immobilisée sur la cellulose.

Pour le simple brin d'ADN de thymus de veau sur cellulose, le fournisseur indique la masse d'ADN par rapport à la masse totale (dans notre cas, 4 mg d'ADN par gramme de cellulose). Il suffit d'estimer la masse molaire de l'ADN pour déterminer la quantité d'ADN présente sur la cellulose, et savoir combien peser de cellulose en fonction du volume de la bibliothèque souhaité.

Pour estimer la masse molaire d'un simple brin d'ADN, il est nécessaire de connaître le nombre des différentes bases. Les masses molaires des nucléotides étant connues, le calcul dans le cas général est ensuite aisé (Équation 4.5).

Équation 4.5

$$M_{ADN} = N_A M_{dAMP} + N_T M_{dTMP} + N_G M_{dGMP} + N_C M_{dCMP} - (N_{Tot} - 1) M_{H_2O}$$

Avec :

- M_{ADN} masse molaire estimée de l'ADN (en $g.mol^{-1}$) ;
- N_X nombre de bases X ;
- M_{dXMP} masse molaire du nucléotide dXMP (en $g.mol^{-1}$) ;
- N_{Tot} nombre total de bases ; $N_{Tot} - 1$ représente le nombre de molécules d'eau éliminées lors de la formation des liaisons phosphodiesters entre les oligonucléotides ;
- M_{H2O} masse molaire des molécules d'eau éliminées lors de la formation des liaisons phosphodiesters entre les nucléotides (18 $g.mol^{-1}$).

Dans le cas d'un simple brin d'ADN de thymus de veau, le nombre de bases total (environ 50 000 bases) et la répartition des différentes bases (41,9% de bases G-C et 58,1% de bases A-T) sont connues : cela revient à connaître indirectement le nombre de chacune des bases. Dans ce cas précis, la formule générale de l'Équation 4.5 devient l'Équation 4.6.

Équation 4.6

$$M_{ADN\ thymus\ de\ veau} = N_{Bases}\left(\%_{A-T}^{mol} \times \frac{M_A + M_T}{2} + \%_{G-C}^{mol} \times \frac{M_G + M_C}{2}\right) + M_{H_2O}$$

Avec :

- $M_{ADN\ thymus\ de\ veau}$ masse molaire estimée d'un simple brin d'ADN de thymus de veau (en $g.mol^{-1}$) ;
- N_{Bases} nombre total de bases (~50 000) ;
- $\%_{A-T}^{mol}$ pourcentage molaire de bases A et T (41,9% soit 41,9/100) ;
- $\%_{G-C}^{mol}$ pourcentage molaire de bases G et C (58,1% soit 58,1/100).

La masse molaire d'un simple brin d'ADN de thymus de veau est ainsi estimée à environ 15 445 482 $g.mol^{-1}$. Comme la masse d'ADN par gramme de cellulose est connue, il est ensuite simple d'estimer la quantité d'ADN par gramme de cellulose. Dans le cas de notre ADN de thymus de veau sur cellulose, cette dernière est donc estimée à $2{,}59.10^{-1}$ nmol.

Connaissant la quantité d'ADN immobilisés sur la cellulose, il suffit à présent de l'exprimer en quantité de phosphodiesters immobilisés sur la cellulose, ce qui équivaut à l'exprimer en quantité de bases par gramme de cellulose (Équation 4.7).

Équation 4.7

$$n_{P/g\ de\ cellulose} = n_{ODN(X)/g\ de\ cellulose} \times N_{Bases}$$

Avec :

- $n_{P\ /\ g\ de\ cellulose}$ quantité de phosphodiester par gramme de cellulose (mol/g) ;
- $n_{ODN(X)\ /\ g\ de\ cellulose}$ quantité d'oligonucléotide X par gramme de cellulose (mol/g) ;
- N_{Bases} nombre de bases de l'oligonucléotide X.

Pour l'ADN de thymus de veau sur cellulose, la quantité de phosphodiesters est de 12,95 µmol par gramme de cellulose.

Pour l'oligo(dT) sur cellulose, la masse molaire du simple brin d'ADN peut être estimée de la même manière que pour celle de l'ADN de thymus de veau, mais sa séquence étant connue précisément, le calcul de la masse molaire est effectué à l'aide de calculateurs en ligne.[45,46] L'oligo(dT) comportant un maximum de 20 thymines par brin, la masse molaire maximale est donc de 6022,0 $g.mol^{-1}$.

Par contre, contrairement à l'ADN de thymus de veau sur cellulose, le fournisseur n'indique pas la masse d'oligo(dT) par rapport à la masse de cellulose, mais une capacité d'hydridation CH avec des poly(A), exprimée en absorbance à 260 nm par gramme de cellulose. Cette capacité d'hybridation permet d'estimer la quantité d'oligonucléotides complémentaires qui est retenue par hybridation sur l'oligo(dT) par gramme de cellulose. Malheureusement, la longueur des poly(A) utilisés pour déterminer cette capacité d'hybridation n'est pas connue, ce qui ne permet pas de calculer cette quantité de matière avec précision.

Il n'est donc pas possible d'estimer la quantité d'oligo(dT) par gramme de cellulose. Pour des raisons pratiques, une quantité identique de support solide a donc été utilisée avec l'ADN de thymus de veau sur cellulose et l'oligo(dT) sur cellulose. Toutefois, il est

nécessaire de garder à l'esprit que si la quantité d'ADN de thymus de veau sur cellulose est estimée, la quantité d'oligo(dT) utilisée pour générer les bibliothèques est inconnue.

B. BIBLIOTHEQUES AVEC DES MONOAMINES

1. Bibliothèques avec une seule monoamine

Les premières bibliothèques ont d'abord été générées avec une seule monoamine dans chaque bibliothèque pour étudier le comportement de chacune d'elles vis-à-vis des cibles testées et faire varier différents paramètres.

a. Premières bibliothèques avec la benzylamine dans l'eau

Dans un premier temps, plusieurs bibliothèques ont été générées avec la benzylamine seule (Figure 4.18). Quatre types de bibliothèques différentes ont été préparés : deux avec les cibles oligo(dT) ou ADN de thymus de veau sur cellulose, un troisième avec de la cellulose microcristalline seule et le dernier en tant que témoin en solution, sans support solide.

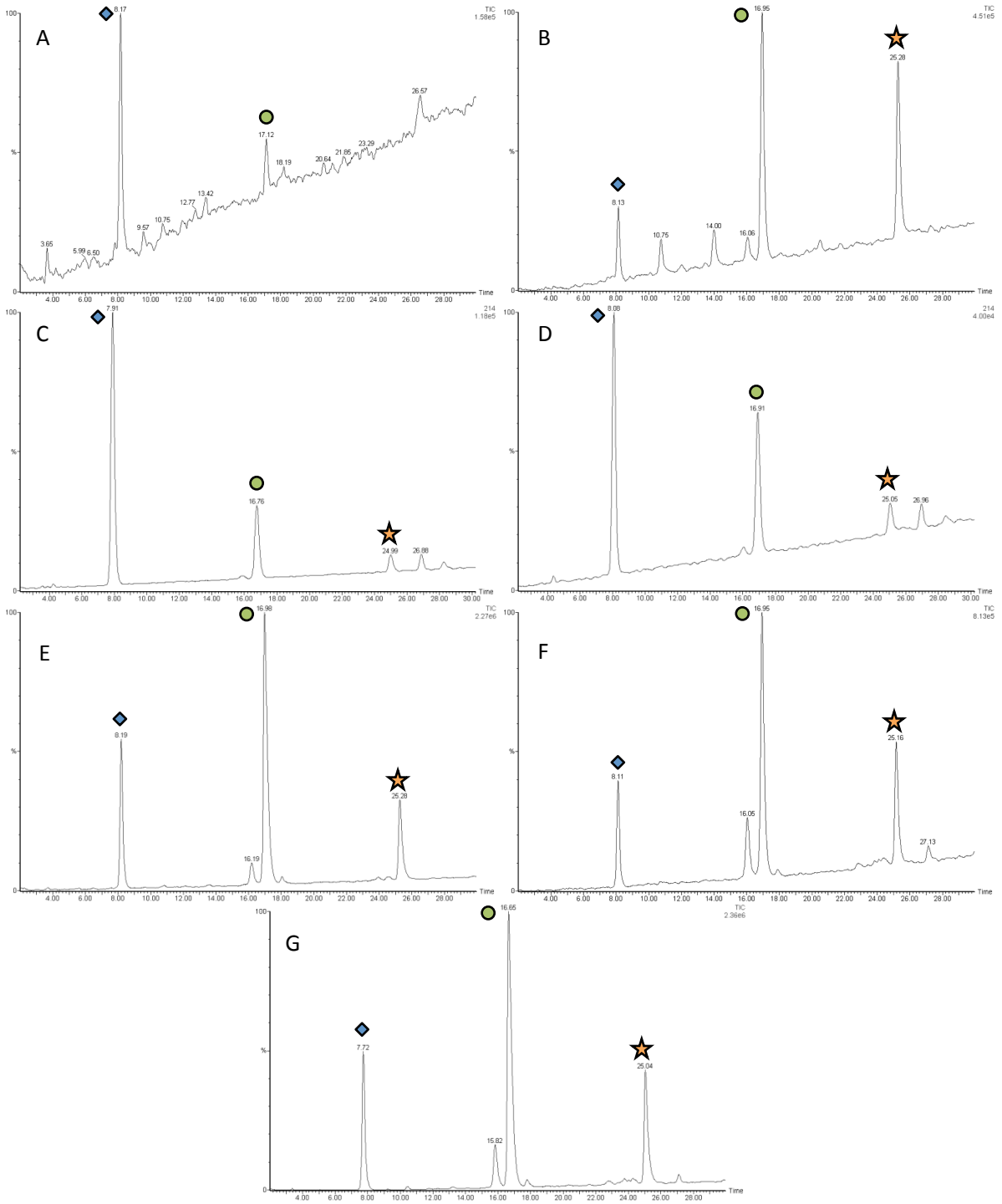
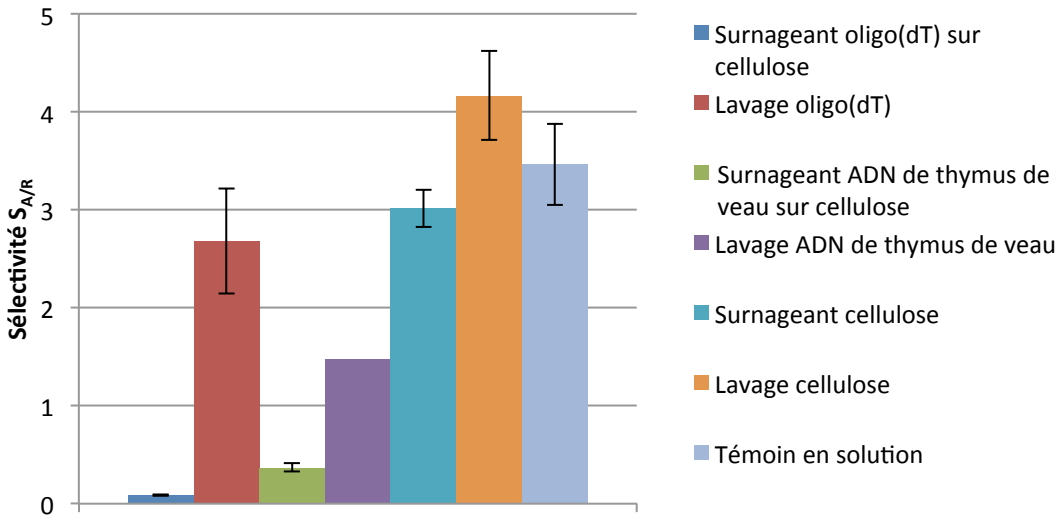


Figure 4.18 : Chromatogrammes LC-MS des surnageants et lavages des bibliothèques générées avec la benzylamine (Benz) en présence d'oligo(dT) sur cellulose dans l'eau milli-Q. A : surnageant de la bibliothèque en présence d'oligo(dT) sur cellulose ; B : lavage de l'oligo(dT) sur cellulose ; C : surnageant de la bibliothèque en présence d'ADN de thymus de veau sur cellulose ; D : lavage de l'ADN de thymus de veau sur cellulose ; E : surnageant de la bibliothèque en présence de cellulose microcristalline ; F : lavage de la cellulose microcristalline ; G : bibliothèque témoin en solution. Les lavages de l'oligo(dT) sur cellulose et de la cellulose ont été effectués avec une solution de tampon phosphate 1 M pH 6,0. Méthode V sur la LC-MS 1, suivi à 214 nm. Produits : ◆ **Gua(OH)$_2$** ; ● **Gua-Benz(OH)** et ★ **Gua-Benz$_2$**.

Des différences flagrantes sont visibles entre les surnageants des deux bibliothèques avec les cibles sur support solide et les bibliothèques témoins avec support solide et en solution. En effet, les produits aminés **Gua-Benz$_2$** et **Gua-Benz(OH)** sont moins présents que le produit de réduction directe **Gua(OH)$_2$** dans les surnageants qui ont été en présence d'ADN. L'analyse des lavages des supports solides avec ou sans cible par un tampon phosphate montrent que les constituants aminés sont formés mais sont retenus sur les cibles ADN. Cette rétention sélective peut être quantifiée en calculant les sélectivités amination sur réduction $S_{A/R}$ (Équation 4.) pour chaque analyse (Graphique 4.2).



Graphique 4.2 : Sélectivité $S_{A/R}$ des surnageants et lavages des bibliothèques générées en présence d'oligo(dT) sur cellulose et d'ADN de thymus de veau sur cellulose, avec la benzylamine dans l'eau milli-Q. Lorsque les bibliothèques ont été répétées deux ou trois fois, les barres d'erreurs représentent les écarts-types.

La similarité des sélectivités $S_{A/R}$ des surnageant et lavage de la bibliothèque générée en présence de cellulose seule avec la sélectivité $S_{A/R}$ de la bibliothèque témoin en solution met en évidence que le support solide seul a peu d'influence sur la constitution finale des bibliothèques et ne retient pas plus les produits aminés **Gua-Benz$_2$** et **Gua-Benz(OH)** que le produit de réduction directe **Gua(OH)$_2$**.

Par contre, d'importantes diminutions des sélectivités $S_{A/R}$ sont observées dans les surnageants des bibliothèques générées en présence de cible sur cellulose, ce qui correspond à la diminution, voire la disparition, des produits aminés **Gua-Benz$_2$** et **Gua-Benz(OH)** qui a été observée sur les chromatogrammes précédents (Figure 4.18). Les valeurs plus importantes des sélectivités $S_{A/R}$ des lavages des supports solides de ces bibliothèques indiquent que la proportion de produits aminés est plus importante dans le lavage que dans le surnageant correspondant. Ces résultats mettent en évidence que les produits aminés **Gua-Benz$_2$** et **Gua-**

Benz(OH) sont sélectivement retenus sur les cibles ADN par rapport au produit de réduction **Gua(OH)$_2$**.

L'importance de la rétention sélective dépend cependant de la nature de la cible. En effet, des différences sont visibles entre les sélectivités $S_{A/R}$ des surnageants des bibliothèques en présence d'oligo(dT) et d'ADN de thymus de veau sur cellulose : la rétention sélective est plus importante sur l'oligo(dT) sur cellulose que sur l'ADN de thymus de veau sur cellulose. De même, la sélectivité $S_{A/R}$ est plus importante dans le lavage de l'oligo(dT) sur cellulose que dans celui de l'ADN de thymus de veau sur cellulose : cette tendance corrèle avec la moindre rétention des produits aminés sur l'ADN de thymus de veau par rapport à l'oligo(dT). Cette différence de rétention peut être due à la quantité d'ADN différente présente dans les bibliothèques avec l'oligo(dT) ou l'ADN de thymus de veau sur cellulose. Les différences de séquences et de longueurs de ces deux oligonucléotides peuvent également jouer.

D'autres modes d'interactions peuvent également coexister avec l'interaction entre le guanidinium et le squelette phosphodiester et expliquer les différences observées entre les deux cibles, notamment dans les lavages. En effet, ceux-ci s'effectuent avec un tampon phosphate qui va jouer un rôle de compétiteur avec les phosphodiesters de l'ADN : les lavages contiendront donc les constituants qui interagissaient avec les phosphodiesters, mais ils ne décrocheront pas forcément les constituants qui interagissent à un autre niveau sur l'ADN (intercalation par exemple). Cela peut également expliquer pourquoi la sélectivité $S_{A/R}$ du lavage de l'ADN de thymus de veau sur cellulose est moins importante que celui du lavage de l'oligo(dT) sur cellulose.

b. Rôle compétiteur du tampon phosphate

Pour confirmer si les constituants des bibliothèques interagissent au niveau des phosphodiesters, les mêmes bibliothèques ont été générées dans un tampon phosphate pH 6,0 à la place de l'eau milli-Q. Comme pour les lavages, le tampon phosphate va être en compétition avec les phosphodiesters des ADN pour l'interaction avec les constituants guanidiniums. Sur les chromatogrammes HPLC (Figure 4.19), des différences par rapport aux bibliothèques avec cible générées dans l'eau milli-Q sont visibles, notamment la présence plus importante du produit **Gua-Benz$_2$** dans le surnageant des bibliothèques avec l'oligo(dT) et l'ADN de thymus de veau sur cellulose.

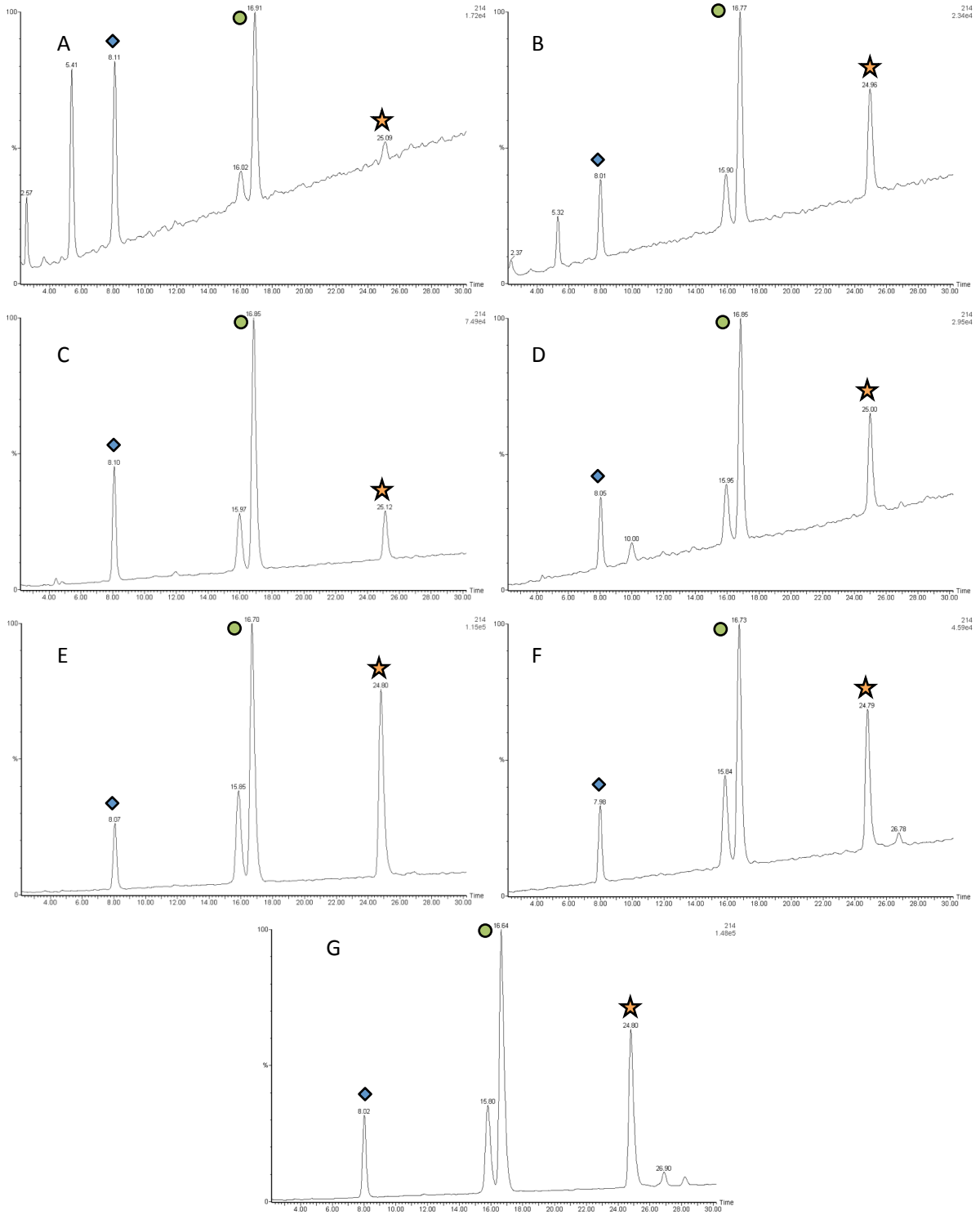


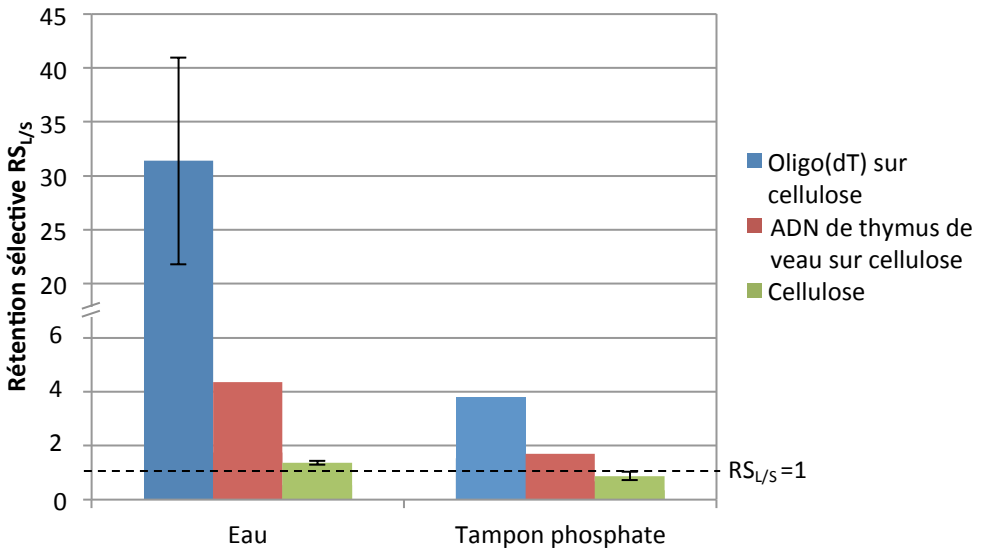
Figure 4.19 : Chromatogrammes LC-MS des surnageants et lavages des bibliothèques générées avec la benzylamine (Benz) en présence d'oligo(dT) sur cellulose dans le tampon phosphate pH 6 (100 mM). A : surnageant de la bibliothèque en présence d'oligo(dT) sur cellulose ; B : lavage de l'oligo(dT) sur cellulose ; C : surnageant de la bibliothèque en présence d'ADN de thymus de veau sur cellulose ; D : lavage de l'ADN de thymus de veau sur cellulose ; E : surnageant de la bibliothèque en présence de cellulose microcristalline ; F : lavage de la cellulose microcristalline ; G : bibliothèque témoin en solution. Les lavages de l'oligo(dT) sur cellulose et de la cellulose ont été effectués avec une solution de tampon phosphate 1 M pH 6,0. Méthode V sur la LC-MS 1, suivi à 214 nm. Produits : ◆ **Gua(OH)$_2$** ; ● **Gua-Benz(OH)** et ★ **Gua-Benz$_2$**.

Pour faciliter la comparaison entre les différentes bibliothèques, la rétention sélective des produits aminés par rapport aux produits de réduction sur la cible ou le support solide peut être calculée (Équation 4.8).

Équation 4.8 : Formule générale de la rétention sélective des produits aminés sur une cible ADN ou la cellulose.

$$RS_{L/S} = \frac{S_{A/R}(lavage)}{S_{A/R}(surnageant)}$$

La comparaison des rétentions sélectives $RS_{L/S}$ entre bibliothèques générées dans l'eau et le tampon phosphate met en évidence la moindre rétention des constituants dans les bibliothèques générées dans le tampon phosphate pH 6,0 (Graphique 4.3). Aucune rétention sélective n'est observée sur la cellulose ($RS_{L/S}$ proche de 1) : le support solide n'induit donc pas de rétention sélective de certains constituants.



Graphique 4.3 : Rétention sélective $RS_{L/S}$ des produits aminés par rapport aux produits de réduction dans les différentes bibliothèques générées avec la benzylamine dans l'eau milli-Q ou le tampon phosphate pH 6,0. Lorsque les bibliothèques ont été répétées deux ou trois fois, les barres d'erreurs représentent l'écart-type.

Lorsque toutes les autres conditions sont identiques, les sélectivités $S_{A/R}$ sont plus faibles dans les bibliothèques générées dans l'eau milli-Q que dans celles générées dans le tampon phosphate pH 6,0. Cette observation s'explique par le fait que l'absence de tampon favorise la réduction directe par rapport à l'amination réductrice, comme cela a été observé précédemment lorsque les différents paramètres influant sur l'amination réductrice ont été étudiés (voir I.). Le calcul des rétentions sélectives sur les cibles et supports solide permet de normaliser ces différences.

Les rétentions sélectives $RS_{L/S}$ diffèrent entre les bibliothèques dans l'eau et le tampon phosphate. Si la rétention sélective $RS_{L/S}$ de la bibliothèque en présence d'oligo(dT) dans l'eau est 23 fois plus importante que celle de la bibliothèque témoin avec la cellulose, cette augmentation n'est plus que de 4 fois pour les mêmes bibliothèques dans le tampon phosphate. La rétention est donc plus faible en présence de tampon phosphate : le tampon joue donc bien le rôle de compétiteur vis-à-vis des phosphodiesters pour l'interaction avec les constituants des bibliothèques.

Comme dans l'eau et pour les mêmes raisons, la rétention est moins importante avec l'ADN de thymus de veau qu'avec l'oligo(dT) sur cellulose dans le tampon phosphate. En termes de rétention, l'intérêt d'utiliser un ADN de thymus de veau est donc moindre. Les bibliothèques seront donc par la suite générée avec l'oligo(dT) sur cellulose comme seule cible.

c. Différences de rétention en fonction de la nature des monoamines présentes

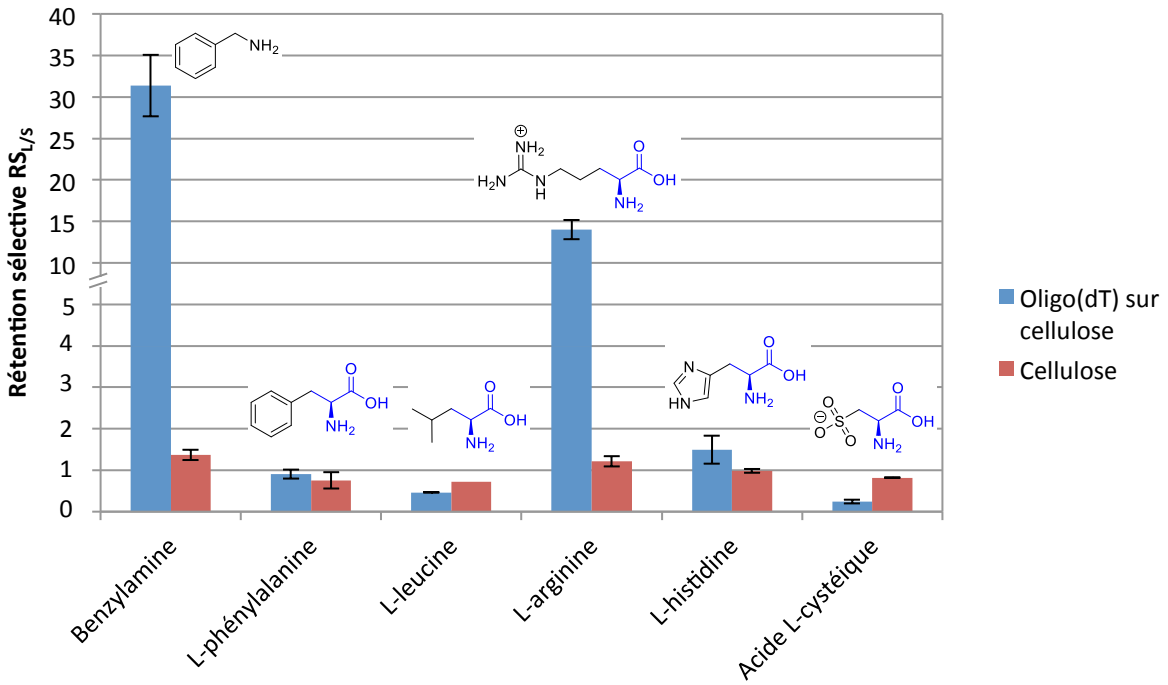
Comme une rétention sélective des constituants aminés a été observée avec la benzylamine, d'autres monoamines aux propriétés différentes (aromatiques, chargées, …) ont été testées séparément. L'idée est ici de voir si les différences qui pourront être observées dans les bibliothèques peuvent être mises en relation avec les différences structurelles des amines et apporter une explication à la rétention sélective observée.

Des monoamines apparentées mais portant des substituants différents ont d'abord été testées et comparées à la benzylamine (Tableau 4.9). Ces amines sont la L-phénylalanine (aromatique), la L-leucine (aliphatique), la L-arginine (chargée positivement), la L-histidine (aromatique avec possibilité de liaison H) et l'acide L-cystéique (chargé négativement).

Tableau 4.9 : Propriétés acido-basiques des différentes monoamines testées dans les bibliothèques combinatoires dynamiques[37] et types de leur substituant.

Amines	Structure	pK_a COOH	pK_a NH_2	Autre pK_a	pI
Benzylamine			9,34		
L-phénylalanine		2,18	9,09		5,48
L-leucine		2,32	9,58		5,97
L-arginine		2,03	9,00	12,10	10,76
L-histidine		1,70	9,09	6,04	7,59
Acide L-cystéique		1,89	8,70	1,30	

Les rétentions sélectives $RS_{L/S}$ des différentes bibliothèques ont été calculées et des tendances différentes sont observées en fonction du substituant présent sur la monoamine (Graphique 4.4).



Graphique 4.4 : Rétention sélective $RS_{L/S}$ des différentes bibliothèques en fonction des monoamines testées. Bibliothèques générées dans l'eau milli-Q.

Les sélectivités $S_{A/R}$ étant calculées à partir des aires des constituants, elles dépendent des coefficients d'extinction molaires des différents produits et ne peuvent donc pas être

comparés directement entre deux monoamines, puisque les produits impliqués ne sont pas les mêmes. Toutefois, ces différences de coefficient d'extinction molaire sont atténuées lors du calcul des rétentions sélectives $RS_{L/S}$, rendant possible le dégagement d'une première tendance.

Quelle que soit l'amine présente, la cellulose n'a pas ou peu d'influence sur la constitution finale des bibliothèques. En effet, les sélectivités $S_{A/R}$ du surnageant et du lavage des bibliothèques générées en présence de cellulose sont similaires à celle de la bibliothèque témoin en solution correspondante (non présentées). Les différences observées dans les bibliothèques avec l'oligo(dT) sur cellulose sont donc toujours dues à l'oligo(dT) et non à la présence de cellulose.

Si les valeurs des rétentions sélectives $RS_{L/S}$ des bibliothèques avec l'oligo(dT) sur cellulose sont comparées entre elles, on remarque qu'il n'y a rétention sélective sur l'oligo(dT) des produits **Gua-R$_2$** et **Gua-R(OH)** que pour la benzylamine et la L-arginine. Pour les L-phénylalanine et L-histidine, les rétentions sélectives sont équivalentes à celles des bibliothèques correspondantes avec la cellulose : l'oligo(dT) ne retient donc pas préférentiellement les produits aminés **Gua-R$_2$** et **Gua-R(OH)** plutôt que le produit de réduction directe **Gua(OH)$_2$**.

Dans le cas des L-leucine et acide L-cystéique, les valeurs des rétentions sélectives sont inférieures à celles des témoins avec la cellulose correspondants. La situation est donc inversée par rapport à ce qui est observée avec la benzylamine : c'est le produit de réduction directe **Gua(OH)$_2$** qui est ici préférentiellement retenu. Cela signifie que la substitution apportée par l'amination réductrice ne stabilise pas l'interaction avec l'ADN.

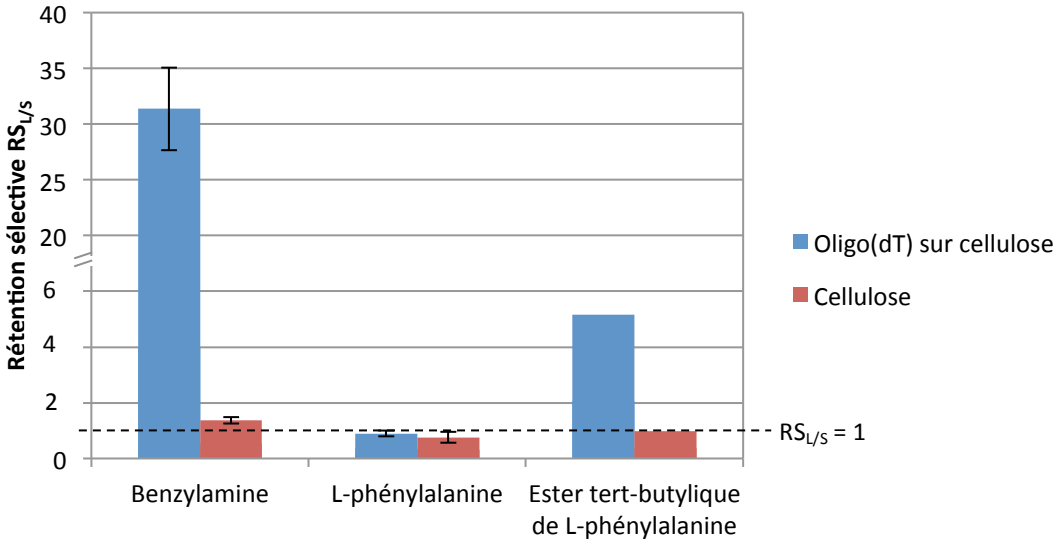
En raisonnant en termes de structures, ces résultats montrent que les groupements cationiques favorisent la rétention sélective sur l'ADN tandis que les groupements aliphatiques ou anioniques ne favorisent *a priori* pas l'interaction avec l'oligonucléotide.

Des différences sont visibles entre la benzylamine et la L-phénylalanine alors qu'elles portent toutes deux un substituant aromatique. En considérant la structure de ces deux amines, il est facile de repérer que leur principale différence se situe au niveau de la fonction acide carboxylique présente sur la L-phénylalanine puisque celle-ci est un acide aminé. Or, cet acide carboxylique est sous forme carboxylate à un pH d'environ 6 (les acides aminés sont sous forme zwitterionique à ce pH) : la L-phénylalanine porte donc une charge négative qui n'est pas présente sur la benzylamine. Comme l'exemple avec l'acide L-cystéique montre que les substituants anioniques défavorisent la rétention sur l'ADN, il est possible que la présence

du carboxylate sur la L-phénylalanine annule le bénéfice apporté par le cycle phénylique qu'elle comporte.

Pour vérifier cette hypothèse, des bibliothèques ont été générées avec des acides aminés dont la fonction carboxylate est protégée par un ester et dans la mesure du possible, avec des amines proches structurellement des acides aminés mais sans leur fonction carboxylate.

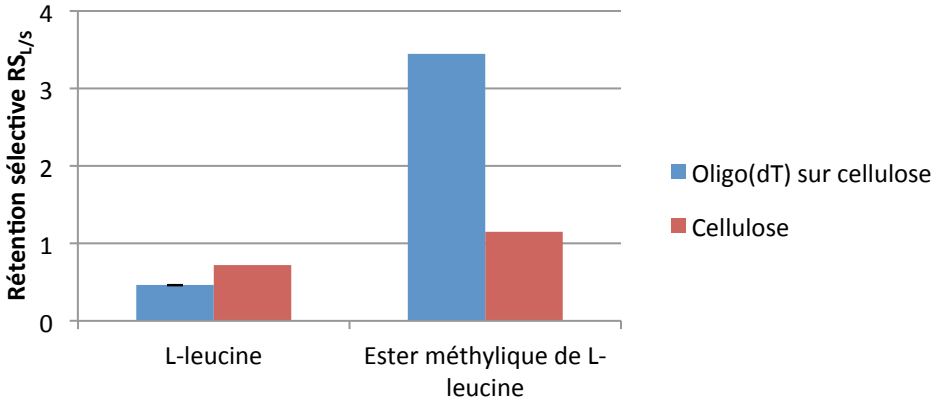
Ainsi, les mêmes bibliothèques ont été générées avec l'ester *tert*-butylique de L-phénylalanine et la rétention sélective $RS_{L/S}$ sur l'oligo(dT) sur cellulose est comparée à celles observées pour la benzylamine et la L-phénylalanine (Graphique 4.5). Bien qu'il soit plus encombré qu'un ester méthylique, l'ester tert-butylique de L-phénylalanine a été utilisé car il était déjà disponible au laboratoire. Avec l'ester *tert*-butylique de L-phénylalanine, la rétention sélective $RS_{L/S}$ sur l'oligo(dT) augmente de manière significative par rapport à celle de la bibliothèque générée avec la L-phénylalanine. Même si elle n'est pas aussi prononcée que celle de la benzylamine, cette tendance indique qu'il y a rétention sélective des produits **Gua-R$_2$** et **Gua-R(OH)** sur l'oligo(dT) dans les bibliothèques générées avec l'ester *tert*-butylique de L-phénylalanine. Ce résultat met en évidence que « masquer » la charge négative de la L-phénylalanine permet de retrouver une rétention sur l'ADN.



Graphique 4.5 : Rétention sélective $RS_{L/S}$ des différentes bibliothèques générées avec la benzylamine, la L-phénylalanine et l'ester *tert*-butylique de L-phénylalanine dans l'eau milli-Q.

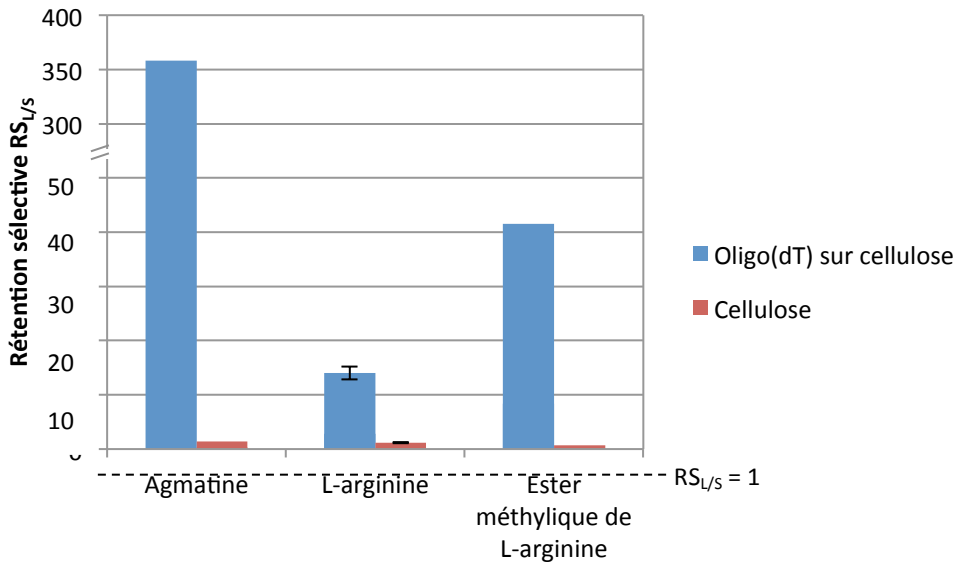
Des bibliothèques ont également été préparées avec l'ester méthylique de L-leucine pour être comparées avec la L-leucine. La comparaison des rétentions sélectives $RS_{L/S}$ met en évidence que la tendance observée précédemment se confirme également avec des amines aliphatiques (Graphique 4.6). Toutefois, l'effet semble moins important qu'avec les amines

aromatiques. Il est possible que les groupements aliphatiques apportent moins de stabilité aux assemblages constituants – ADN que des groupements aromatiques, mais cette hypothèse devra être vérifiée par des expériences de compétition. Il est également possible que la nature des groupements esters soit à l'origine des différences observées.



Graphique 4.6 : Rétention sélective $RS_{L/S}$ des différentes bibliothèques générées avec la L-leucine et l'ester méthylique de L-leucine dans l'eau milli-Q.

Lors de la première comparaison des amines, la L-arginine était sortie du lot puisqu'elle était le seul acide aminé à être sélectivement retenue sur l'ADN. Des bibliothèques ont été préparées avec l'agmatine et l'ester méthylique de L-arginine pour voir si l'absence ou le masquage de la charge négative amplifie la rétention des produits aminés **Gua-R$_2$** et **Gua-R(OH)** sur l'ADN (Graphique 4.7). Avec l'agmatine, les produits aminés sont retenus si fortement sur l'ADN qu'ils ne sont pas détectés dans le surnageant (d'où une estimation grossière de la valeur de la rétention sélective $RS_{L/S}$). Avec l'ester méthylique de L-arginine, la rétention sélective $RS_{L/S}$ sur l'oligo(dT) est plus importante qu'avec la L-arginine : pour une même quantité de produit de réduction directe **Gua(OH)$_2$** (les aires de ce produit sont similaires dans les deux lavages), une proportion plus importante de produits aminés **Gua-R$_2$** et **Gua-R(OH)** est donc récupérée dans le lavage de l'oligo(dT) avec l'ester méthylique. La tendance déduite des différentes sélectivités $S_{A/R}$ semble donc aller dans le sens d'une amplification de la rétention sur l'ADN en absence de charge négative.



Graphique 4.7 : Rétention sélective $RS_{L/S}$ des bibliothèques générées avec l'agmatine, la L-arginine et l'ester méthylique de L-arginine dans l'eau milli-Q.

Considérés ensemble, ces résultats montrent que la nature des amines utilisées dans les bibliothèques joue un rôle important dans une potentielle stabilisation de l'auto-assemblage constituants – ADN, observée par la rétention sélective de certains constituants sur l'oligo(dT) sur cellulose dans plusieurs des bibliothèques générées. Les groupements favorisant la rétention sélective des constituants sur l'ADN sont les groupements cationiques et aromatiques. A l'inverse, la présence de charges négatives défavorise fortement la rétention sur l'ADN et peut entraîner la perte du bénéfice acquis avec un groupement stabilisateur. Cette tendance reste cependant à confirmer par des mises en compétitions.

2. Compétition avec plusieurs monoamines

Générer des bibliothèques contenant le réactif **Gua-Ald** et plusieurs monoamines va permettre de déterminer, parmi les amines présentes, quelles sont celles qui favorisent la stabilisation des constituants avec la cible oligonucléotide. Cependant, il peut être nécessaire de corriger la réactivité des amines pour que tous les constituants puissent être formés. Lorsque la différence de réactivité est trop importante, la formation d'un ou plusieurs produits peut en effet être favorisée par rapport à d'autres. Cette formation privilégiée peut biaiser les bibliothèques et rendre caduque les résultats obtenus.

Par exemple, lors de la génération d'une bibliothèque avec le réactif **Gua-Ald** et un nombre équimolaire de benzylamine et de L-leucine (10 équivalents de chaque amine), le

composé bis-aminé **Gua-Leu$_2$** formé avec la L-leucine n'est pas détecté alors que dans une bibliothèque avec la L-leucine comme seule amine, il était bien présent. Une correction des réactivités de la benzylamine et de la L-leucine sera donc nécessaire dans une bibliothèque les contenant.

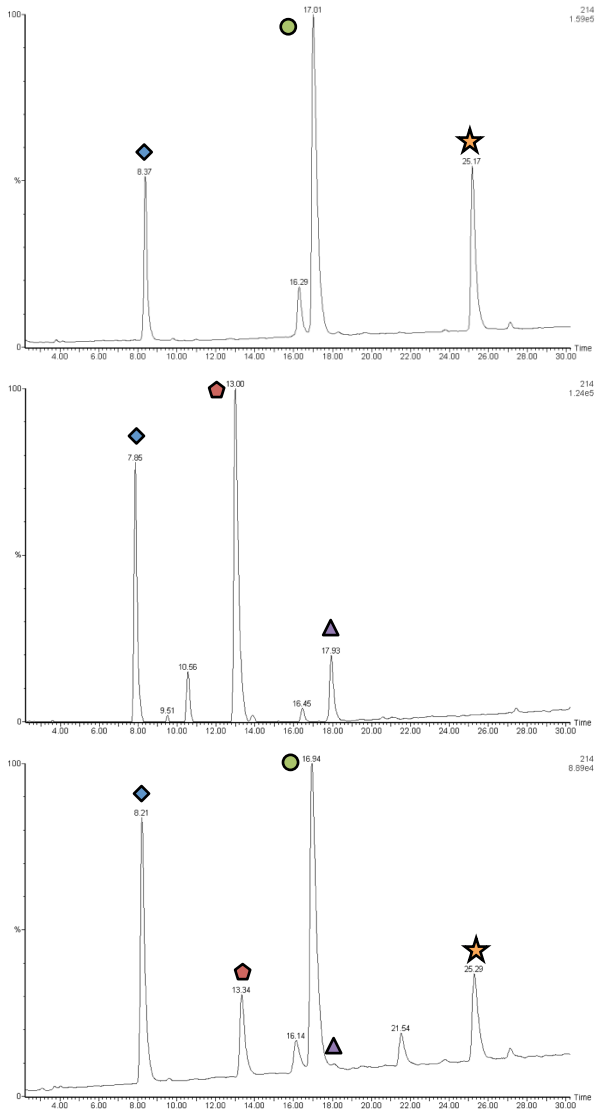


Figure 4.20 : Chromatogrammes LC-MS des bibliothèques générées avec 10 équivalents de benzylamine (haut), 10 équivalents de L-leucine (milieu) et un mélange de 10 équivalents de benzylamine et de 10 équivalent de L-leucine (bas) dans l'eau milli-Q. Méthode V sur la LC-MS 1, suivi à 214 nm. Produits : ◆ **Gua(OH)$_2$** ; ⬟ **Gua-Leu(OH)** ; ● **Gua-Benz(OH)** ; ▲ **Gua-Leu$_2$** et ☆ **Gua-Benz$_2$**.

La correction de la réactivité des amines se fait en jouant sur le nombre d'équivalents d'amine introduit dans la bibliothèque. Diminuer la stœchiométrie d'une amine permet de pondérer sa réactivité et, *a contrario*, l'augmenter doit favoriser la formation d'une quantité plus importante de constituants dérivés de l'amine. Cette correction n'est pas nécessaire lorsque les amines ont une réactivité similaire, mais ce n'est pas toujours le cas dans nos

bibliothèques. En effet, des différences marquées vis-à-vis de l'amination réductrice ont déjà été observées entre les différentes monoamines en absence de cible (voir précédemment I.G).

Comme il n'est pas possible de comparer les aires des constituants fonctionnalisés par les différentes amines entre eux, la réactivité des amines avec le réactif **Gua-Ald** peut être estimée en comparant la proportion de produit de réduction directe **Gua(OH)$_2$** obtenue dans les différentes bibliothèques générées en absence de cible dans des conditions identiques. Plus l'aire du produit **Gua(OH)$_2$** est importante, moins l'amine est réactive dans les conditions utilisées. Dans un premier temps, la comparaison des aires du produit de réduction **Gua(OH)$_2$** observées dans les différentes bibliothèques avec une seule amine peut ainsi aider à estimer le nombre d'équivalents d'amines à utiliser dans les bibliothèques avec mise en compétition pour que les amines présentes dans le mélange aient une réactivité similaire. Cependant, il demeure important de vérifier que les bibliothèques générées ne sont pas biaisées. L'absence d'un produit aminé dans une bibliothèque témoin avec plusieurs amines alors que ce même produit était présent dans la bibliothèque avec une seule amine devra attirer l'attention.

Avec nos cibles sur support solide, ce sont les sélectivités $S_{A/R}$ qui ont été comparées entre bibliothèques pour savoir dans quel sens avait évolué la constitution finale de la bibliothèque en présence de cible. En présence de plusieurs amines dans une même bibliothèque, cette sélectivité perd son intérêt. En effet, dans le cas d'une bibliothèque préparée avec deux amines aux effets contraires sur la rétention des constituants sur la cible, la sélectivité $S_{A/R}$ serait équivalente à celle d'une bibliothèque où la constitution finale n'a pas changée : la sélectivité $S_{A/R}$ ne reflète donc pas toujours les changements survenus dans les bibliothèques en présence de cible.

Des bibliothèques ont été générées avec deux amines au lieu d'une pour confirmer les observations faites précédemment avec les bibliothèques avec une seule amine et, si possible, discriminer les amines en fonction de la force de rétention sélective de leur constituants sur la cible ADN. Une nouvelle sélectivité $S_{A1/A2}$ a été définie pour comparer les bibliothèques générées (Équation 4.9). Comme pour la sélectivité $S_{A/R}$, cette sélectivité $S_{A1/A2}$ n'est pas comparable pour deux bibliothèques ne comportant pas les mêmes amines, mais va permettre des comparaisons entre surnageants, lavages et témoins en solution (Tableau 4.10). Il faut également noter que les produits mixtes **Gua-(A1)(A2)** ne sont pas pris en compte pour le calcul de cette sélectivité.

Équation 4.9 : Formule générale de la sélectivité amine 1 / amine 2 $S_{A1/A2}$.

$$S_{A1/A2} = \frac{Aire[GR(OH)\ amine\ 1] + Aire[GR_2\ amine\ 1]}{Aire\ [GR(OH)\ amine\ 2] + Aire[GR_2\ amine\ 2]} = \frac{1}{S_{A2/A1}}$$

Tableau 4.10 : Sélectivité $S_{A1/A2}$ des différentes bibliothèques générées avec deux monoamines différentes. Les amines 1 et 2 sont indiquées par les abréviations A1 et A2 respectivement.

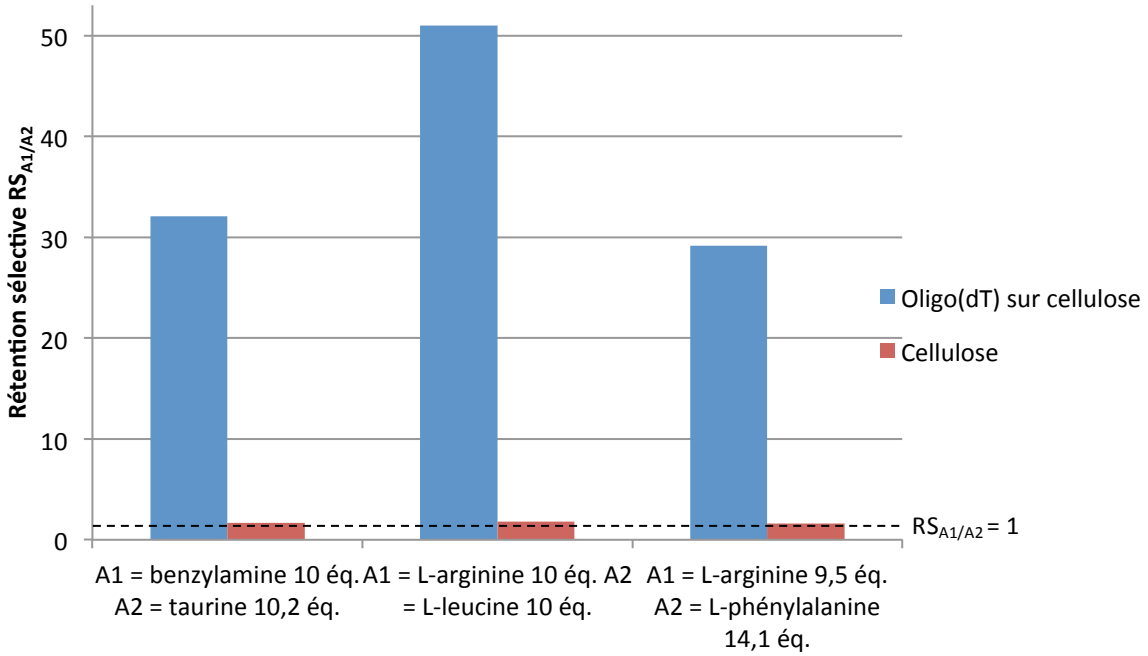
Cible et / ou support solide	**Oligo(dT) sur cellulose**		**Cellulose**		**Aucun**
Phase	**Surnageant**	**Lavage**	**Surnageant**	**Lavage**	**Témoin**
A1 = Benzylamine 10 eq + A2 = taurine 10,2 eq	0,12	3,85	0,92	1,52	1,21
A1 = L-arginine 10 eq + A2 = L-leucine 10 eq	0,06	3,06	0,74	1,35	1,03
A1 = L-arginine 9,5 eq + A2 = L-phénylalanine 14,1 eq	0,06	1,75	0,47	0,75	0,59

Similairement à la rétention sélective des produits aminés par rapport aux produits de réduction $RS_{L/S}$, une rétention sélective des produits de l'amine 1 par rapport aux produits de l'amine 2 $RS_{A1/A2}$ peut être exprimée à partir des sélectivités $S_{A1/A2}$ des surnageants et lavages des bibliothèques en présence de cible ou de support solide (Équation 4.10) et permettre des comparaisons plus aisées.

Équation 4.10 : Formule générale de la rétention sélective des constituants générés à partir d'une amine 1 par rapport à ceux générés à partir d'une amine 2 dans les bibliothèques en présence de cible ou de support solide.

$$RS_{A1/A2} = \frac{S_{A1/A2}(lavage)}{S_{A1/A2}(surnageant)}$$

Cette rétention sélective peut permettre de déterminer quelle amine est à l'origine des constituants les plus retenus sur la cible et a été calculée pour les différentes bibliothèques préparées avec deux amines différentes (Graphique 4.8).



Graphique 4.8 : Rétention sélective des amines $RS_{A1/A2}$ des différentes bibliothèques en fonction de leur composition en amines.

La cellulose n'a pas d'impact sur la répartition finale des constituants en fonction des amines, la sélectivité $S_{A1/A2}$ étant similaire dans le surnageant de la cellulose, le lavage de la cellulose et la bibliothèque témoin en solution. Cette observation n'est pas surprenante puisque la cellulose ne retenait pas sélectivement de constituants dans les bibliothèques générées avec une seule monoamine, mais elle permet de confirmer que la cellulose n'a pas d'influence sur la répartition des différents constituants des bibliothèques.

Des différences de rétention sélective plus ou moins marquées en fonction des mélanges étudiés sont par contre observées entre les bibliothèques avec la cible et les bibliothèques correspondantes contenant uniquement la cellulose.

Ainsi, les constituants dérivés de la benzylamine sont sélectivement retenus sur l'oligo(dT) par rapport aux dérivés de la taurine. De même, les dérivés de la L-arginine sont préférentiellement retenus sur l'ADN plutôt que les dérivés de la L-leucine et la L-phénylalanine. Vu les tendances des rétentions sélectives entre produits aminés et produits de réduction $RS_{L/S}$ qui ont été observées précédemment avec les bibliothèques préparées avec une seule amine, ces résultats n'ont rien de très surprenant mais confirment les observations précédentes. Il est clair que les groupements cationiques favorisent la rétention sélective des constituants sur la cible ADN contrairement aux groupements anioniques qui sont fortement déstabilisants pour une interaction constituants – ADN.

Des bibliothèques ont été générées avec l'agmatine et la taurine, mais des produits peuvent être superposés sur le chromatogramme LC-MS, la méthode HPLC utilisée ne permettant pas une séparation efficace.

Un cinquième mélange d'amines composé de benzylamine (Benz, 10 équivalents) et d'agmatine (Agm, 14 équivalents) a également été utilisé mais les bibliothèques obtenues sont biaisées, le produit **Gua-Agm$_2$** obtenu avec l'agmatine n'est pas présent dans la bibliothèque témoin contrairement au produit bis-aminé avec la benzylamine. Une rétention sélective des produits aminés avait été observée dans les bibliothèques préparées avec chaque amine séparément et il n'était pas possible de conclure quant à une rétention sélective sur l'ADN des constituants dérivés de l'une ou l'autre amine. Une diminution du nombre d'équivalents de benzylamine introduits dans les bibliothèques ne permet d'observer que des traces de produit bis-aminé **Gua-Agm$_2$** dans les bibliothèques témoins, bien que la quantité de produits générés à partir de la benzylamine diminue de manière significative (produit bis-aminé **Gua-Benz$_2$** peu intense avec 3 équivalents de benzylamine). La rétention sélective du produit monoaminé **Gua-Agm(OH)** par rapport aux produits aminés **Gua-Benz(OH)** et **Gua-Benz$_2$** observée dans ces différentes conditions ne peut donc être considérée comme indiscutable vue la persistance des biais présents dans les bibliothèques. Des essais avec une quantité plus importante d'agmatine doivent être réalisés pour éliminer ces biais.

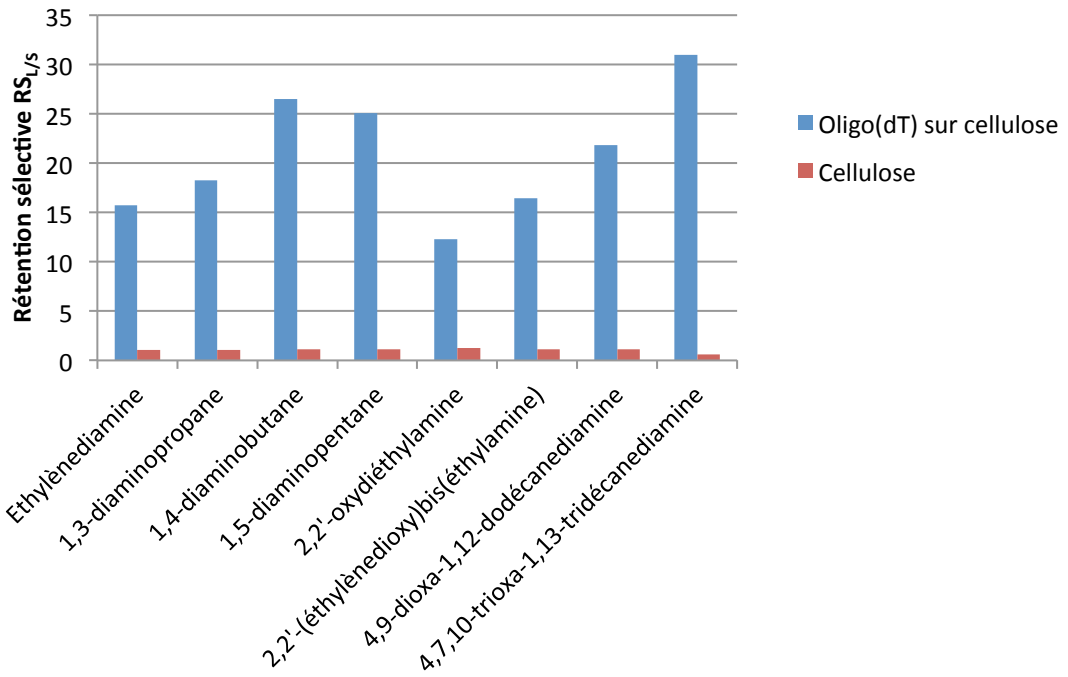
La rétention sélective des différents constituants des bibliothèques sur l'oligo(dT) sur cellulose dépend de leur structure chimique. En rationalisant les différents résultats de mise en compétition, la tendance d'un groupement ou substituant, à stabiliser l'auto-assemblage constituants – ADN suit l'ordre suivant : groupement cationique > aromatique > aliphatique >> anionique.

La rationalisation de la stabilisation d'un auto-assemblage petite molécule – oligonucléotide n'ayant jamais été effectuée et la méthode utilisée pour parvenir à ces résultats n'ayant pas été testée avec un autre modèle connu, il est cependant nécessaire de valider la méthode utilisée en confirmant les résultats obtenus avec une méthode déjà validée comme la spectrométrie de masse MALDI-ToF. Toutefois, ces premiers résultats sont intéressants pour mieux comprendre les interactions petites molécules – phosphodiesters de l'ADN et complètent les résultats du chapitre 2.

C. BIBLIOTHEQUES AVEC DES DIAMINES

Le composé **Gua-Ald** étant un guanidinium bisaldéhyde, il est susceptible de réagir avec des diamines pour former des macrocycles ou des oligomères / polymères.[47] Les conditions utilisées jusqu'à présent pour générer les bibliothèques sont *a priori* peu favorables à la formation de tels objets, sauf si un effet de matrice suffisamment important est induit par la cible. Connaissant la disposition d'un oligonucléotide, il pourrait être possible de favoriser la formation d'oligo-guanidinium le long du squelette phosphodiester de la cible, sous réserve que la taille des « monomères » leur permettent de s'organiser le long du brin d'ADN et d'être liés entre eux par réaction avec une diamine jouant le rôle d'espaceur. Concernant de potentiels constituants macrocycliques, leur formation ne semble pas favorable en raison de l'excès de diamines utilisé lors de la génération des bibliothèques. De plus, la chaîne carbonée des diamines n'interagit *a priori* pas avec les bases de l'ADN, ne permettant pas une pré-organisation des fragments de la bibliothèque autour de la matrice ADN.

Des bibliothèques ont donc été générées avec différentes diamines dans les mêmes conditions que pour les monoamines. Une première vision d'ensemble peut être obtenue en comparant la tendance suivie par les rétentions sélectives $RS_{L/S}$ (Équation 4.8) des différentes bibliothèques (Graphique 4.9). Il est intéressant de remarquer ici que quelle que soit la diamine considérée, les produits aminés sont préférentiellement retenus sur l'ADN plutôt que le produit de réduction directe **Gua(OH)$_2$**. Toutefois, les produits d'amination susceptibles d'être formés sont plus nombreux que dans le cas des monoamines puisqu'il faut y rajouter tous les produits linéaires issus de la « poly-amination » et les macrocycles. La situation est donc plus complexe et la comparaison des sélectivités $S_{A/R}$ n'est pas suffisante pour conclure sur de possibles changements constitutionnels au sein des bibliothèques.



Graphique 4.9 : Rétention sélective $RS_{L/S}$ des différentes bibliothèques en fonction de la diamine présente. Bibliothèques générées dans l'eau milli-Q avec 10 équivalents de diamine.

Pour les diamines dont la chaîne carbonée (pouvant comporter des hétéroatomes) comprend trois atomes ou plus, les constitutions des bibliothèques sont assez similaires et l'exemple avec le 1,5-diaminopentane reflète la tendance générale (Figure 4.21). Dans le cas général, le produit bis-aminé **Gua-R$_2$** est systématiquement retenu sur l'oligo(dT). Bien que ce soit peu visible à la seule vue des chromatogrammes, ce produit est également amplifié en présence de cible.

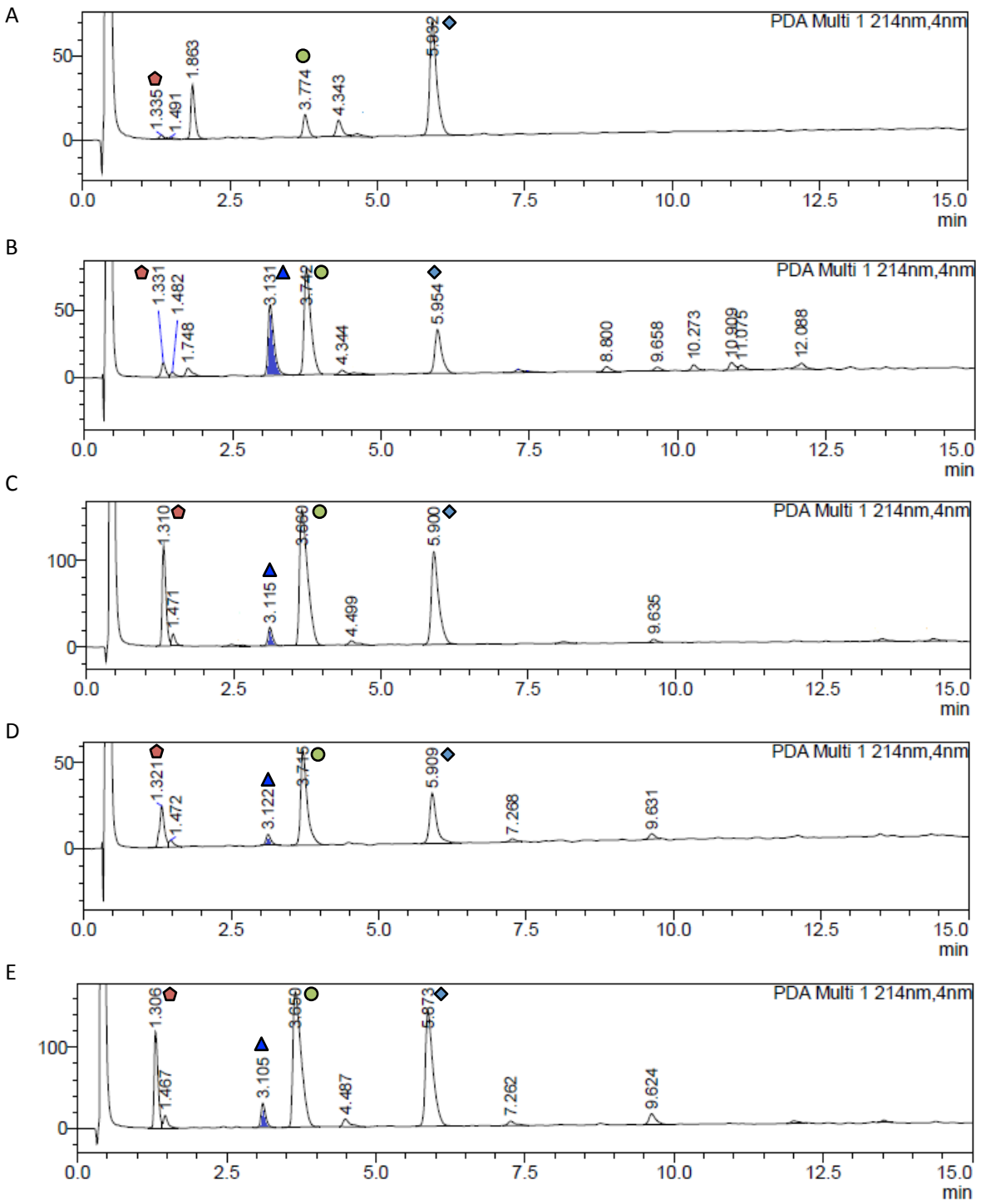


Figure 4.21 : Chromatogrammes LC-MS des surnageants et lavages des bibliothèques générées avec le 1,5-diaminopentane en présence d'oligo(dT) sur cellulose dans l'eau milli-Q. A : surnageant de la bibliothèque en présence d'oligo(dT) sur cellulose ; B : lavage de l'oligo(dT) sur cellulose ; C : surnageant de la bibliothèque en présence de cellulose microcristalline ; D : lavage de la cellulose microcristalline ; E : bibliothèque témoin en solution. Les lavages de l'oligo(dT) sur cellulose et de la cellulose ont été effectués avec une solution de tampon phosphate 1 M pH 6,0. Méthode VIII sur la LC-MS 2, suivi à 214 nm. Produits : ⬟ **[1+1]** ; ▲ **Gua-R$_2$** ; ● **Gua-R(OH)** et ◆ **Gua(OH)$_2$**.

L'amplification du produit **Gua-R$_2$** en présence de l'oligo(dT) a été quantifiée pour les différentes diamines en comparant les aires du produit dans le lavage de l'oligo(dT) sur cellulose et la bibliothèque témoin en solution (Équation 4.3, Tableau 4.11). Pour toutes les diamines dont la chaîne carbonée (hétéroatomes inclus) est de trois atomes ou plus, le produit **Gua-R$_2$** est toujours amplifié mais il est difficile de rationaliser les différents degrés d'amplification observés en fonction de la longueur de chaîne des diamines. Egalement visibles sur les chromatogrammes, les macrocycles **[1+1]** ne sont pas amplifiés. Les macrocycles **[2+2]** ne sont quant à eux pas détectés dans les bibliothèques (ou alors, seulement à l'état de traces).

Tableau 4.11 : Amplification du produit **Gua-R$_2$** dans le lavage de l'oligo(dT) sur cellulose par rapport au témoin en solution. Le produit **Gua-R$_2$** n'est pas détecté dans le surnageant de la bibliothèque avec l'oligo(dT) sur cellulose sauf avec l'éthylènediamine où il est présent à l'état de traces (amplification -98%).

Diamine testée	Amplification de Gua-R$_2$ dans le lavage de l'oligo(dT) par rapport au témoin en solution
Ethylènediamine	-2%
1,3-diaminopropane	+421%
1,4-diaminobutane	+166%
1,5-diaminopentane	+127%
2,2'-oxydiéthylamine	+832%
2,2'-(éthylènedioxy)bis(éthylamine)	+579%
4,9-dioxa-1,12-dodécanediamine	+94%
4,7,10-trioxa-1,13-tridecanediamine	+165%

La constitution des bibliothèques avec l'éthylènediamine, diamine testée la plus courte, ne suit pas la tendance générale comme le met en évidence l'absence d'amplification du produit **Gua-R$_2$** en présence d'oligo(dT) (Tableau 4.11), bien que les produits aminés soient sélectivement retenus sur la cible. Les chromatogrammes LC-MS des bibliothèques préparées avec cette diamine mettent en évidence la présence dans le lavage de l'oligo(dT) sur cellulose d'un produit absent dans tous les autres surnageants et lavages, identifié comme le produit linéaire **Gua$_2$R$_3$** (Figure 4.22). La formation de ce produit linéaire **Gua$_2$R$_3$** uniquement en présence d'oligo(dT) et sa rétention sélective sur l'oligo(dT) démontrent que la cible oligo(dT) sert ici de matrice pour former le produit linéaire.

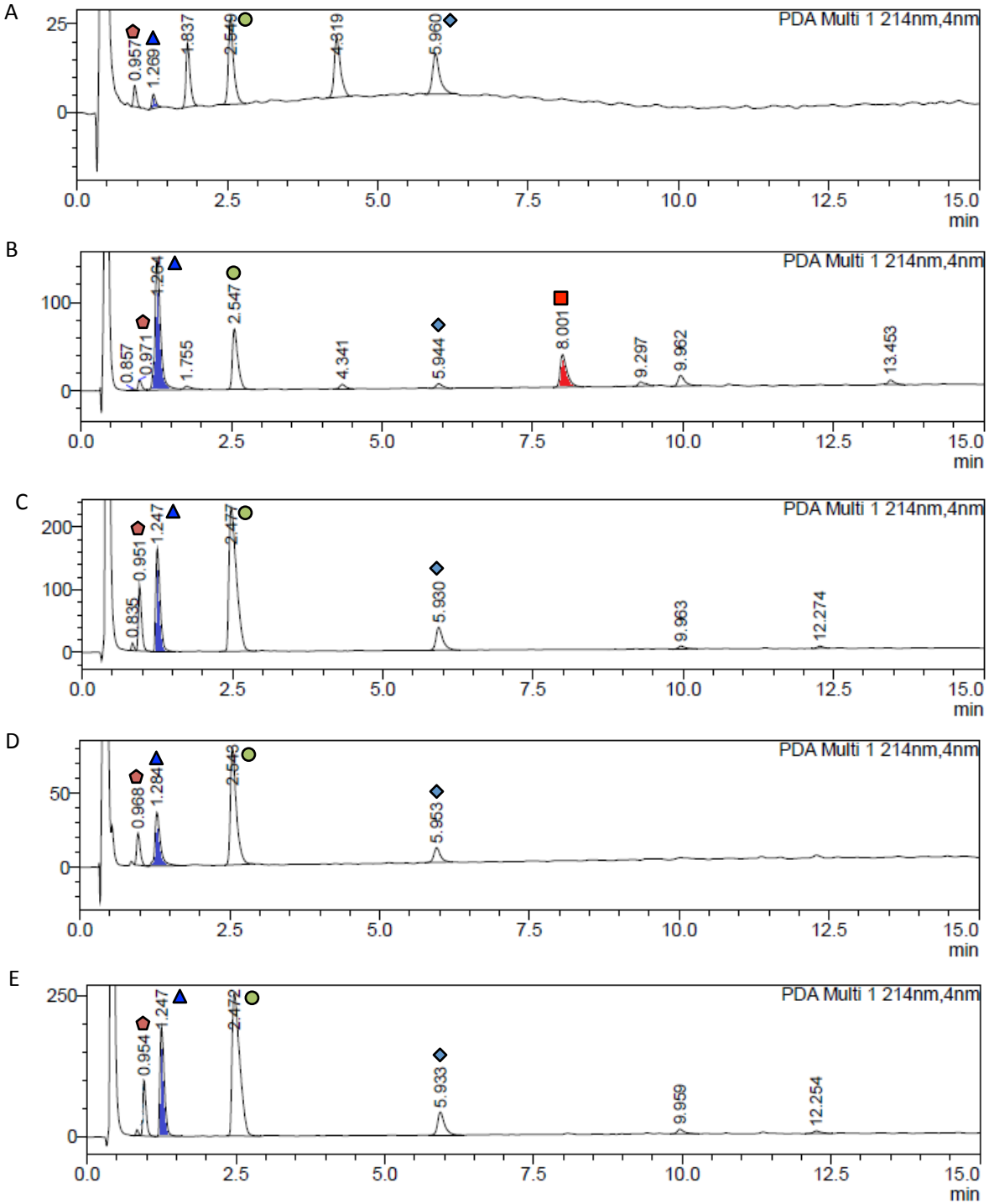


Figure 4.22 : Chromatogrammes LC-MS des surnageants et lavages des bibliothèques générées avec l'éthylènediamine en présence et absence d'oligo(dT) sur cellulose dans l'eau milli-Q. A : surnageant de la bibliothèque en présence d'oligo(dT) sur cellulose ; B : lavage de l'oligo(dT) sur cellulose ; C : surnageant de la bibliothèque en présence de cellulose microcristalline ; D : lavage de la cellulose microcristalline ; E : bibliothèque témoin en solution. Les lavages de l'oligo(dT) sur cellulose et de la cellulose ont été effectués avec une solution de tampon phosphate 1 M pH 6,0. Méthode VIII sur la LC-MS 2, suivi à 214 nm. Produits : ⬟ **[1+1]** ; ▲ **Gua-R$_2$** ; ● **Gua-R(OH)** ; ◆ **Gua(OH)$_2$** et ■ **Gua$_2$R$_3$**.

L'absence de produits linéaires plus longs peut être due au fait que la taille et la structure des différents fragments ne sont pas optimales pour une organisation le long de l'ADN. Cependant, l'observation d'un effet de matrice en présence de cible est un bon début et la génération d'oligo- ou poly-guanidiniums matricée par un brin d'ADN dans les conditions utilisées ne paraît pas impossible.

D. BIBLIOTHEQUES EN PRESENCE DE TRIAZINES

Les diaminotriazines sont connues pour interagir avec les bases thymines des oligonucléotides. Elles sont en effet complémentaires en termes de donneurs et d'accepteurs de liaisons H (Figure 4.23) et ont déjà été utilisées pour préparer des auto-assemblages de triazines avec un simple brin de dT_n.[39]

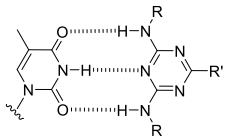
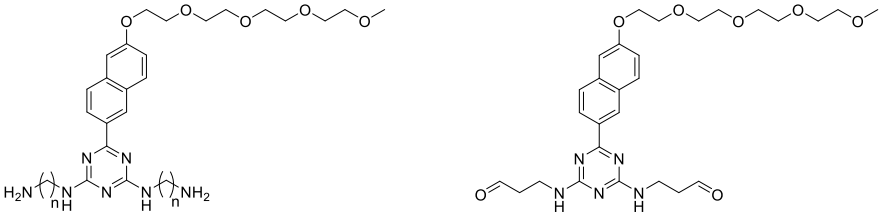


Figure 4.23 : Assemblage par liaisons hydrogène d'une base thymine avec une diaminotrizine.

Des diaminotriazines fonctionnalisées par des fonctions amines ou aldéhydes à l'extrémité des groupements latéraux R ont été synthétisées au laboratoire pour générer des bibliothèques combinatoires dynamiques en présence ou en absence de dT_n (Figure 4.24). En utilisant une cible oligo(dT) sur cellulose, les constituants de ces bibliothèques sont susceptibles de reconnaître les phosphodiesters avec le fragment guanidinium et les thymines avec les triazines.



TN^n **TO**

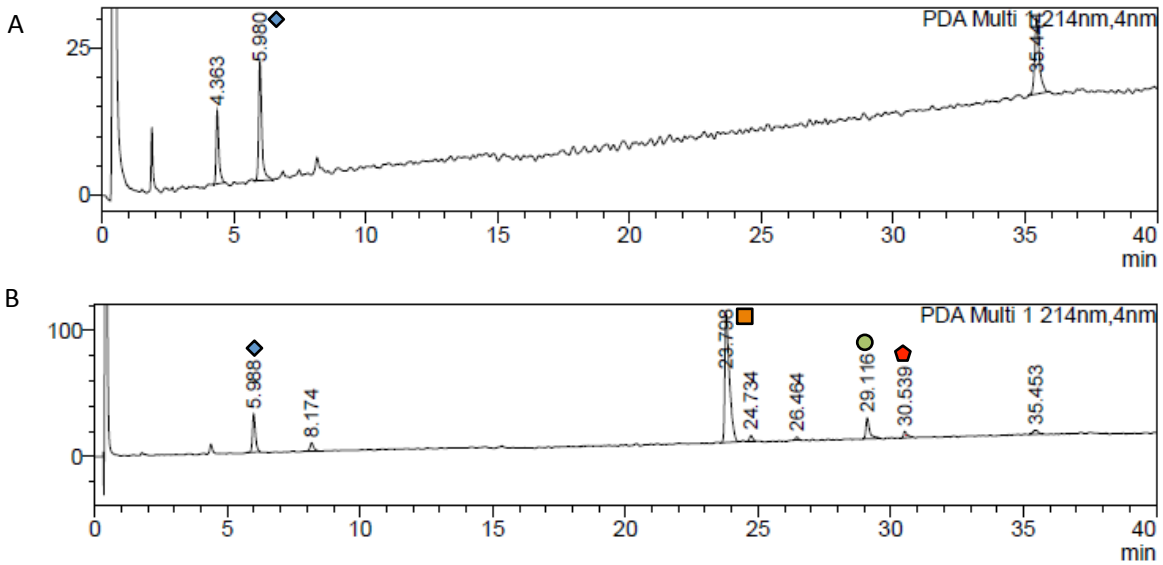
Figure 4.24 : Structure des diaminotriazines fonctionnalisées par des fonctions amines (à gauche, n compris entre 2 et 5) ou aldéhydes (à droite), respectivement notées **TN^n** et **TO**.

1. Bibliothèques avec les triazines diamines **TN^n**

Les triazines **TN^n** ont été utilisées pour générer des bibliothèques en présence de guanidinium bis-aldéhyde **Gua-Ald** à la place des amines commerciales. Les triazines étant bivalentes comme le guanidinium **Gua-Ald**, des oligomères linéaires ou des macrocycles peuvent théoriquement être formés dans les bibliothèques. Toutefois, la longueur des chaînes carbonées latérales des triazines est importante car elle sert « d'espaceur » entre la triazine et le guanidinium et détermine si un macrocycle peut se former autour de l'ADN.

Les triazines reconnaissant les bases de la cible oligo(dT) sur cellulose, elles n'ont pas été introduites en excès par rapport au guanidinium **Gua-Ald** comme les amines commerciales précédentes, mais avec une stœchiométrie estimée à une triazine par base (le ratio **Gua-Ald**:**TN^n** est donc de 1:1). Cette stœchiométrie doit permettre aux composés triazines et guanidinium de s'organiser de part et d'autre d'un simple brin d'oligo(dT) et de cycliser autour si la longueur des chaînes latérales des triazines le permet. Si un trop grand excès de triazines est utilisé par rapport au nombre estimé de bases, les bibliothèques risquent d'être biaisées en présence de cible : en effet, des macrocycles pourraient se former, mais sans qu'une organisation préalable autour de l'ADN ait été nécessaire.

Des bibliothèques ont été préparées avec chaque triazine séparément. La seule cible testée est l'oligo(dT) sur cellulose. Une bibliothèque avec de la cellulose seule et une bibliothèque témoin en solution sont systématiquement générées pour comparaison. La tendance est similaire quelle que soit la triazine **TN^n** testée, illustrée par l'exemple avec la triazine **TN^3** (Figure 4.25).



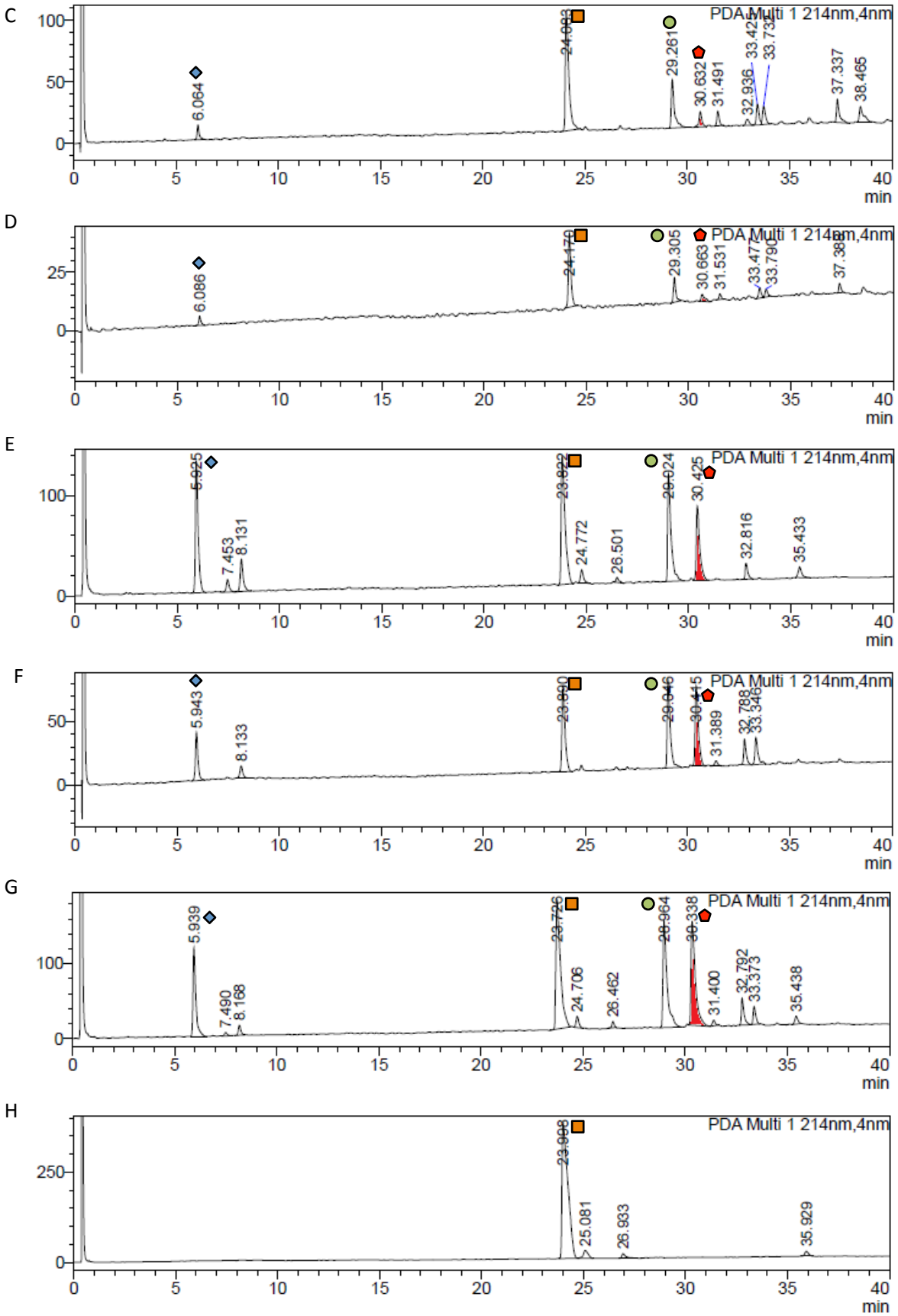


Figure 4.25 : Chromatogrammes LC-MS des surnageants et lavages des bibliothèques générées avec la triazine **TN³** (1 équivalent par **Gua-Ald**) en présence et absence d'oligo(dT) sur cellulose dans l'eau milli-Q. A :

surnageant de la bibliothèque en présence d'oligo(dT) sur cellulose ; B : premier lavage par un tampon phosphate 1 M pH 6,0 de l'oligo(dT) sur cellulose ; C : second lavage de l'oligo(dT) sur cellulose par une solution de guanidine 6 M ; D : troisième lavage de l'oligo(dT) sur cellulose par une solution d'urée 8 M ; E : surnageant de la bibliothèque en présence de cellulose microcristalline ; F : premier lavage de la cellulose microcristalline par un tampon phosphate 1 M pH 6,0 (les chromatogrammes des second et troisième lavages de la cellulose, effectués dans les mêmes conditions que les lavages de l'oligo(dT) sur cellulose, sont omis car ils sont similaires à celui du premier lavage présenté ici) ; G : bibliothèque témoin en solution ; H : solution de triazine **TN^3** seule. Les lavages de l'oligo(dT) sur cellulose et de la cellulose ont été effectués avec une solution de tampon phosphate 1 M pH 6,0. Méthode VIII sur la LC-MS 2, suivi à 214 nm. Produits : ◆ **Gua(OH)$_2$** ; ■ **TN^3** ; ● **Gua-TN^3(OH)** et ⬟ **[1+1]**.

Pour toutes les triazines testées, les macrocycles [1+1] sont formés dans les bibliothèques témoins en solution. Ils sont également présents dans le surnageant et le lavage des bibliothèques avec la cellulose seule. Par contre, ces mêmes macrocycles [1+1] sont systématiquement absents dans le surnageant des bibliothèques avec l'oligo(dT) sur cellulose et sont largement minoritaires dans les lavages de l'oligo(dT) sur cellulose, quelle que soit la solution utilisée pour le lavage (tampon phosphate 1 M pH 6, solution de guanidine 6 M ou solution d'urée 8 M).

Les autres produits retenus sur l'oligo(dT) étant récupérés dans des proportions plus importantes que les macrocycles [1+1] dans les lavages, ces observations mettent en évidence que la formation de ces macrocycles [1+1] est défavorisée en présence d'oligo(dT) sur cellulose. Une longueur de chaîne insuffisante en position latérale sur les triazines peut expliquer ce résultat : deux composés guanidinium et triazine reconnaissent respectivement un phosphodiester et une base thymine proche, mais les fonctions aldéhydes et amines terminales sont trop éloignées l'une de l'autre pour réagir et fermer le macrocycle autour de l'ADN.

2. Bibliothèques avec la triazine bisaldéhyde **TO**

La triazine **TO** a été utilisée pour générer des bibliothèques avec le composé **Gua-Ald** et des diamines. Le niveau de complexité des objets virtuellement formés au sein des bibliothèques augmente de manière importante. Ainsi, former un macrocycle comportant guanidinium et triazine autour d'un simple brin de dT est théoriquement possible, mais nécessite de rassembler quatre fragments, ce qui peut s'avérer difficile si les bibliothèques combinatoires dynamiques ne sont pas conçues de manière optimale. Bien qu'ils ne reposent pas sur la liaison imine, des exemples de macrocycles de cette taille agencés par Chimie Combinatoire Dynamique sont connus.[48–50]

Comme précédemment, les bibliothèques ont été préparées dans l'eau milli-Q et la cible utilisée est l'oligo(dT) sur cellulose. Le ratio **Gua-Ald**:**TO** est de 1:1, ce qui équivaut à un guanidinium et une triazine par base d'oligo(dT). Les diamines n'étant pas impliquées dans la reconnaissance, elles peuvent être introduites en excès et la quantité de 10 équivalents de diamines par rapport au composé **Gua-Ald** est conservée. Avec cet excès d'amines, aucun macrocycle [2+2] n'était détecté dans les bibliothèques préparées avec le guanidinium et les différentes diamines seules. Les conditions ne sont pas les plus favorables pour la formation de ces macrocycles mais comme la réduction directe a également lieu dans les bibliothèques, il est difficile de réduire le nombre d'équivalents de diamines utilisé sans avoir d'impact sur l'amination réductrice.

Chaque diamine a été testée séparément dans les bibliothèques préparées avec le guanidinium **Gua-Ald** et la triazine bis-aldéhyde **TO**. La tendance dans les bibliothèques générées avec chacune des diamines est similaire et peut être illustrée avec l'exemple de la 2,2'-oxydiéthylamine (Figure 4.26).

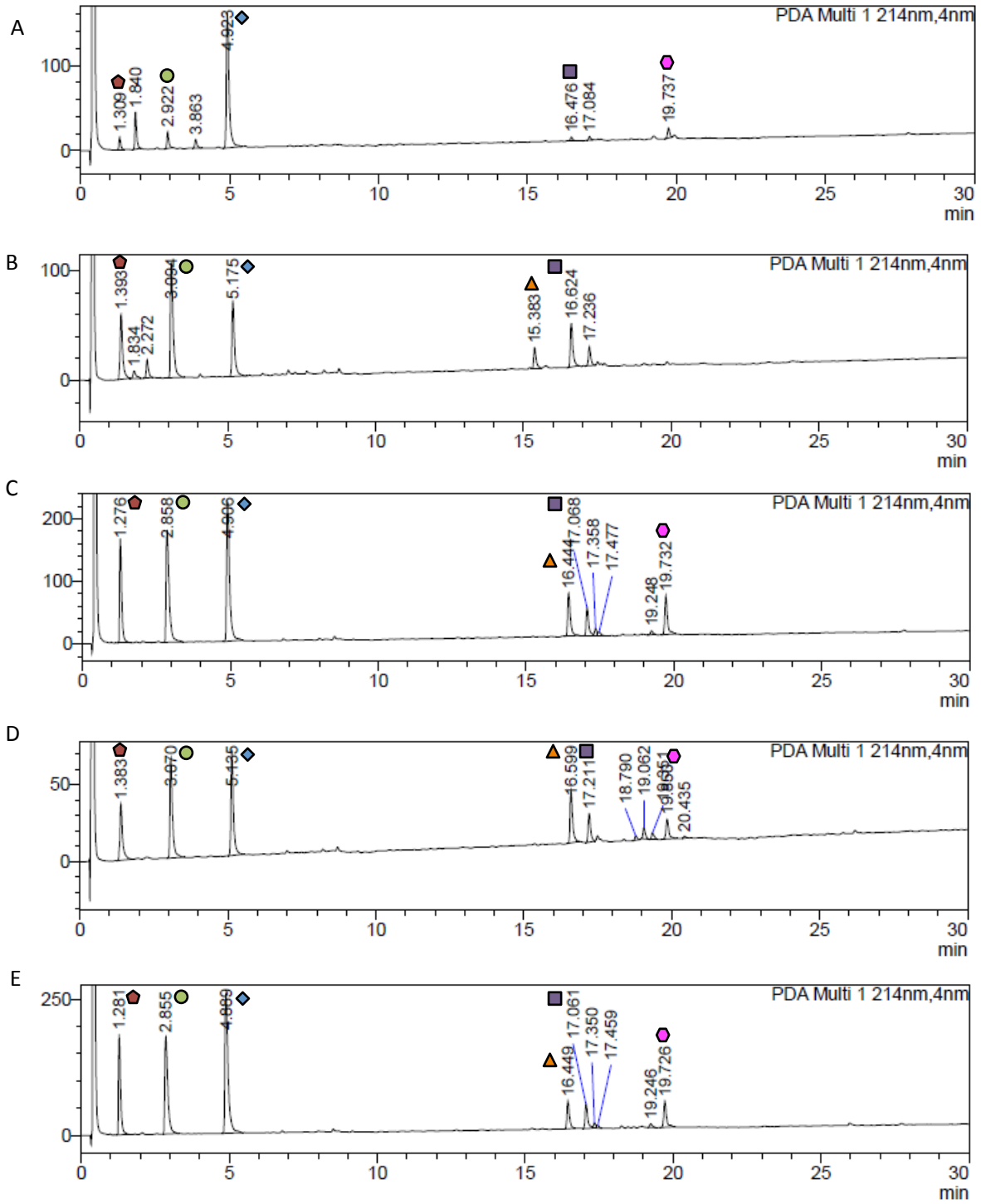


Figure 4.26 : Chromatogrammes LC-MS des surnageants et lavages des bibliothèques générées avec le guanidinium **Gua-Ald**, la triazine bisaldéhyde **TO** et la 2,2'-oxydiéthylamine en présence et absence d'oligo(dT) sur cellulose dans l'eau milli-Q. A : surnageant de la bibliothèque en présence d'oligo(dT) sur cellulose ; B : lavage de l'oligo(dT) sur cellulose ; C : surnageant de la bibliothèque en présence de cellulose microcristalline ; D : lavage de la cellulose microcristalline ; E : bibliothèque témoin en solution. Les lavages de l'oligo(dT) sur cellulose et de la cellulose ont été effectués avec une solution de tampon phosphate 1 M pH 6,0. Méthode IX sur la LC-MS 2, suivi à 214 nm. Produits : ⬟ **[1Gua+1R]** ; ● **Gua-R(OH)** ; ◆ **Gua(OH)$_2$** ; ▲ **[1TO+1R]** ; ■ **TO-R(OH)** et ⬢ **TO(OH)$_2$**.

Malheureusement, aucun macrocycle [2+2] n'est formé dans les bibliothèques en présence de cible (pas de détection dans le surnageant ou le lavage) comme dans les bibliothèques témoins. L'oligo(dT) ne favorise donc pas la formation de ces macrocycles dans les conditions utilisées.

Une analyse plus approfondie de la constitution des bibliothèques témoins en solution met en évidence la présence des constituants issus de la réduction directe **Gua(OH)$_2$** et **TO(OH)$_2$** ainsi que les macrocycles **[1Gua+1R]**. Les produits d'amination linéaires **Gua-R(OH)** sont également rencontrés mais l'absence systématique de constituants **Gua-R$_2$** est assez surprenante puisque ces composés étaient formés dans les bibliothèques préparées avec le guanidinium **Gua-Ald** et les différentes diamines. De plus, ils étaient généralement amplifiés en présence d'oligo(dT) alors qu'ils ne sont même pas visibles dans le cas présent. De même, aucun constituant **TO-R$_2$** n'est présent dans ces bibliothèques témoins alors que le macrocycle **[1TO+1R]** et le produit linéaire **TO-R(OH)** sont présents.

Les mêmes bibliothèques ont été préparées à 4°C, mais des observations identiques sont faites : pour toutes les diamines testées, aucun macrocycle **[1Gua+1TO+2R]** n'est formé et aucune trace de produits bis-aminés **Gua-R$_2$** et **TO-R$_2$** n'est visible quelle que soit la bibliothèque considérée.

Les produits bis-aminés **Gua-R$_2$** et **TO-R$_2$** étant complètement absents des bibliothèques témoins en solution, il est logique que la formation de macrocycles **[1Gua+1TO+2R]** ne soit pas observée. En augmentant le nombre d'équivalents de diamines, la réaction d'amination réductrice sera plus favorisée, même si ce ne sont pas les conditions optimales pour la formation des macrocycles.

IV. CONFIRMATION DE LA METHODE AVEC LA SPECTROMETRIE DE MASSE (MALDI-TOF)

La spectrométrie de masse MALDI-ToF permet sous certaines conditions d'observer des complexes non-covalents (comme dans le chapitre 2). En isolant les composés sélectivement retenus sur la cible oligo(dT) sur cellulose et en effectuant des analyses par MALDI-ToF, il devrait être possible de valider la méthode de préparation des bibliothèques combinatoires dynamiques.

Les composés bis-aminés **Gua-R$_2$** avec la benzylamine (Benz), l'agmatine (Agm) et le 1,4-diaminobutane (C4) ainsi que le guanidinium bisréduit **Gua(OH)$_2$** ont donc été

synthétisés et isolés par HPLC semi-préparative. Le guanidinium bisréduit **Gua(OH)$_2$** n'était pas retenu sur l'ADN et peut donc servir de contrôle négatif. Le composé mono-aminé **Gua-Agm(OH)** formé lors de la synthèse du composé bis-aminé **Gua-Agm$_2$** avec l'agmatine a également été isolé pour comparaison.

De premières analyses ont été effectuées avec les différents composés seuls en présence de dT_{10}, avec une stœchiométrie d'un guanidinium par base (Figure 4.27). Des adduits sont visibles avec les composés bis-aminés **Gua-Benz$_2$**, **Gua-Agm$_2$** et **Gua-C4$_2$** mais pas avec le composé issu de la réduction directe **Gua(OH)$_2$**.

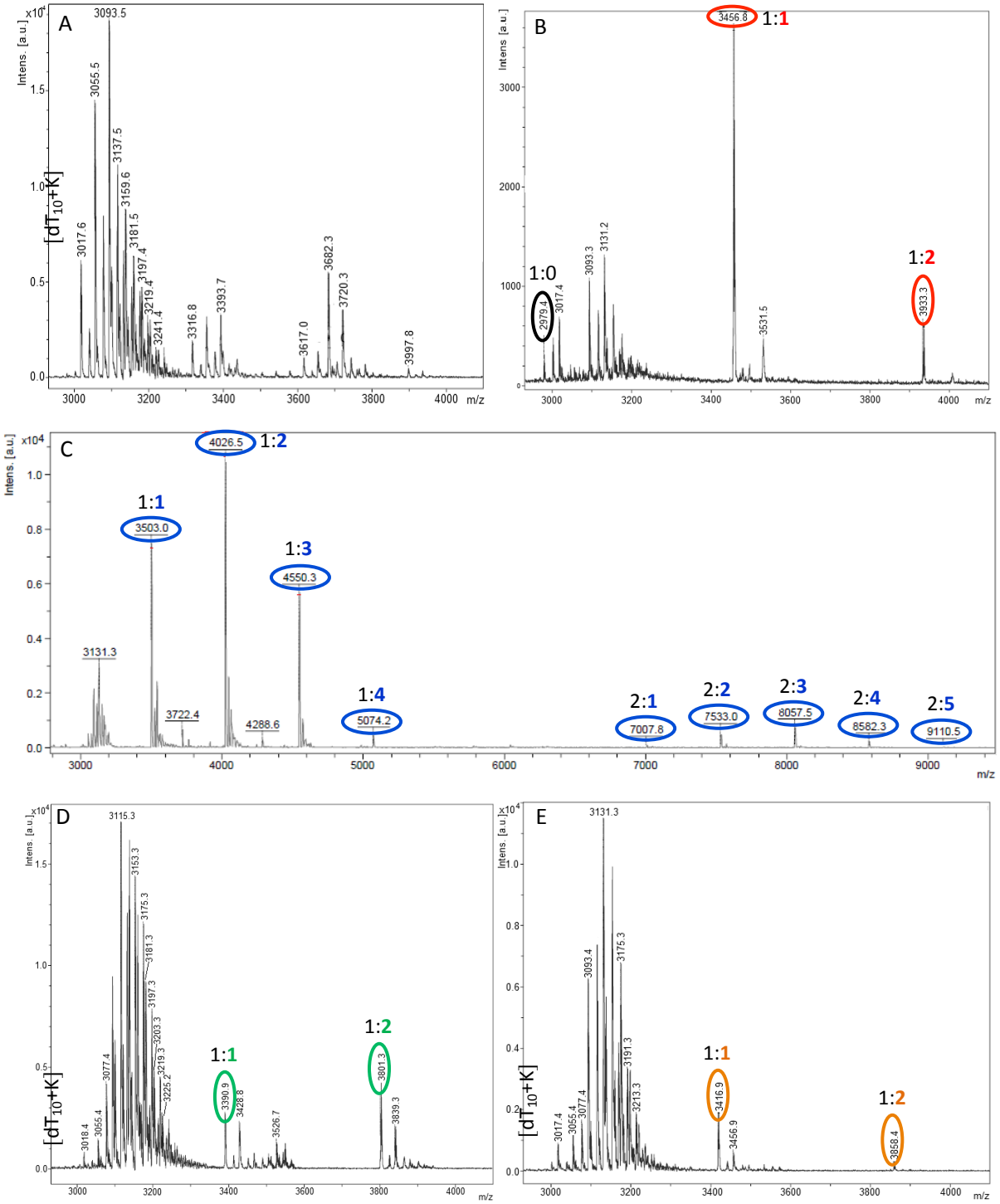


Figure 4.27 : Détection des complexes formés entre un simple brin de dT_{10} et les guanidiums isolés des bibliothèques combinatoires dynamiques par spectrométrie de masse MALDI-ToF (ratio molaire 1:10). A : composé bisréduit **Gua(OH)$_2$** (pas de complexes visibles) ; B : composé bis-aminé avec la benzylamine **Gua-Benz$_2$** ; C : composé bis-aminé avec l'agmatine **Gua-Agm$_2$** ; D : composé monoaminé avec l'agmatine **Gua-Agm(OH)** ; E : composé bis-aminé avec le 1,4-diaminobutane **Gua-C4$_2$**. Les adduits dT_{10} – guanidinium sont indiqués sur les spectres de masse.

Une expérience de mise en compétition a ensuite été effectuée avec les composés bis-aminés de la benzylamine **Gua-Benz$_2$** et l'agmatine **Gua-Agm$_2$** (Figure 4.28). Cette mise en

compétition met en évidence une préférence du dT_{10} pour le composé aromatique **Gua-Benz$_2$** puisque la majorité des adduits observés avec le dT_{10} comporte ce composé (un seul adduit est observé avec le composé **Gua-Agm$_2$**).

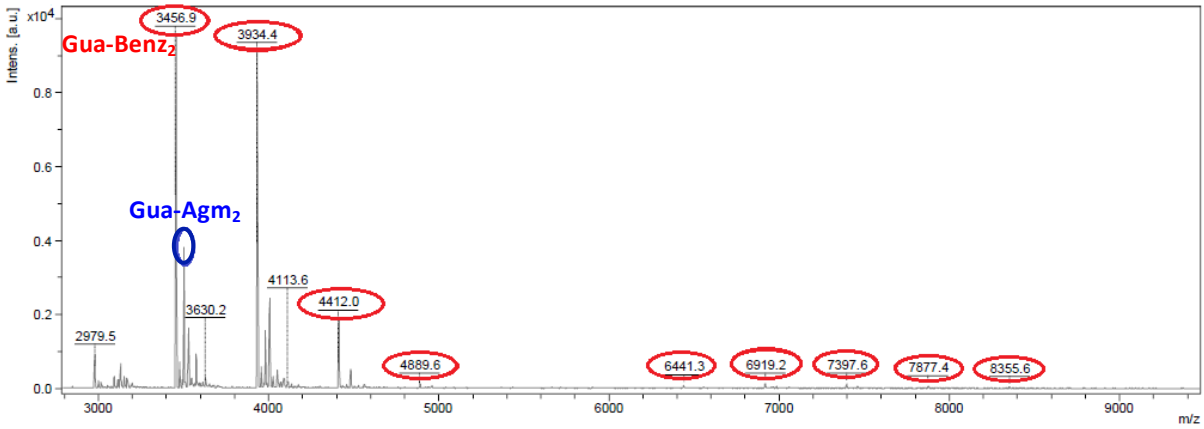


Figure 4.28 : Expérience de compétition entre les composés bis-aminés **Gua-Benz$_2$** et **Gua-Agm$_2$** en présence de simple brin de dT_{10}, analysée par spectrométrie de masse MALDI-ToF. Adduits entre le dT_{10} et **Gua-Benz$_2$** indiqués en rouge et adduit entre le dT_{10} et **Gua-Agm$_2$** indiqués en bleu.

Ces derniers résultats ne sont pas en accord avec la tendance observée lors des mises en compétitions des différentes amines dans les bibliothèques combinatoire dynamiques. En effet, bien que biaisées, les bibliothèques à base de benzylamine et d'agmatine tendaient à indiquer une rétention plus importante des constituants à base d'agmatine. De plus, les bibliothèques générées avec la L-arginine et la L-phénylalanine démontrent clairement la rétention préférentielle des constituants cationiques par rapport aux aromatiques. Il est possible que les différences observées entre les bibliothèques combinatoires dynamiques et les analyses de spectrométrie de masse soient dues au fait que lors des analyses de spectrométrie de masse, le mélange analysé est sous forme gazeuse et les interactions présentes ne sont pas forcément celles prédominantes en solution, observées avec les bibliothèques combinatoires dynamiques.

CONCLUSIONS

Les bibliothèques combinatoires dynamiques générées avec le composé **Gua-Ald** et différentes monoamines en présence d'oligo(dT) sur cellulose ont mis en évidence des différences de rétention des constituants des bibliothèques en fonction des substituants portés. Ces différences de rétention, confirmées par des mises en compétition avec deux amines au

sein des bibliothèques, peuvent être classées selon la tendance suivante, du plus retenu ou moins retenu sur l'oligo(dT) sur cellulose : groupement cationique > aromatique > aliphatique >> anionique.

L'effet compétiteur des ions phosphates vis-à-vis de la reconnaissance des constituants guanidiniums par les phosphodiesters de l'ADN a également été mis en évidence, confirmant une fois de plus l'interaction des guanidiniums avec les phosphodiesters.

Des bibliothèques combinatoires dynamiques générées avec différentes diamines ont mis en évidence, dans les conditions utilisées, l'amplification, en présence d'oligo(dT) sur cellulose, de constituants linéaires dont la nature dépend de la longueur de la diamine, mais pas de macrocycles.

Des bibliothèques générées avec des triazines bis-amines ou bis-aldéhyde n'ont pas permis d'observer la formation de macrocycles avec les guanidiniums en présence d'oligo(dT) sur cellulose, probablement en raison de la conception des différents fragments ne leur permettant pas de s'organiser autour de la matrice de manière à cycliser. La conception des composés semblent donc à revoir pour arriver à cet objectif.

Les informations structurales obtenues à partir des bibliothèques générées avec les mono- et diamines démontrent l'importance des interactions secondaires dans le processus de reconnaissance et d'auto-assemblage avec l'ADN. Ces données viennent compléter les informations obtenues par l'approche rationnelle (chapitre 2) et pourront permettre à l'avenir de concevoir plus aisément de nouveaux ligands de l'ADN.

RÉFÉRENCES BIBLIOGRAPHIQUES

(1) Huc, I.; Lehn, J.-M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 2106–2110.

(2) Lehn, J.-M. *Chem. Eur. J.* **1999**, *5*, 2455–2463.

(3) Ramström, O.; Lehn, J.-M. *Nat. Rev. Drug Discov.* **2002**, *1*, 26–36.

(4) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, *106*, 3652–3711.

(5) Ladame, S. *Org. Biomol. Chem.* **2008**, *6*, 219–226.

(6) Miller, B. L. *Dynamic combinatorial chemistry in drug discovery, bioorganic chemistry, and materials science*; Wiley: Hoboken, N.J., 2010.

(7) Reek, J. N. H.; Otto, S. *Dynamic combinatorial chemistry*; Wiley-VCH: Weinheim, 2010.

(8) Cougnon, F. B. L.; Sanders, J. K. M. *Acc. Chem. Res.* **2012**, *45*, 2211–2221.

(9) Li, J.; Nowak, P.; Otto, S. *J. Am. Chem. Soc.* **2013**, *135*, 9222–9239.

(10) Matache, M.; Bogdan, E.; Hădade, N. D. *Chem. Eur. J.* **2014**, *20*, 2106–2131.

(11) Ulrich, S.; Dumy, P. *Chem. Commun.* **2014**, *50*, 5810–5825.

(12) Otto, S. In *Dynamic Combinatorial Chemistry*; Wiley-VCH Verlag GmbH & Co. KGaA, 2010; pp. 23–47.

(13) Ludlow, R. F.; Otto, S. *J. Am. Chem. Soc.* **2008**, *130*, 12218–12219.

(14) Epstein, D. M.; Choudhary, S.; Churchill, M. R.; Keil, K. M.; Eliseev, A. V.; Morrow, J. R. *Inorg. Chem.* **2001**, *40*, 1591–1596.

(15) Perez-Fernandez, R.; Pittelkow, M.; Belenguer, A. M.; Sanders, J. K. M. *Chem. Commun.* **2008**, 1738–1740.

(16) Klekota, B.; Hammond, M. H.; Miller, B. L. *Tetrahedron Lett.* **1997**, *38*, 8639–8642.

(17) Besenius, P.; Cormack, P. A. G.; Liu, J.; Otto, S.; Sanders, J. K. M.; Sherrington, D. C. *Chem. Eur. J.* **2008**, *14*, 9006–9019.

(18) Besenius, P.; Cormack, P. A. G.; Ludlow, R. F.; Otto, S.; Sherrington, D. C. *Org. Biomol. Chem.* **2010**, *8*, 2414–2418.

(19) McNaughton, B. R.; Miller, B. L. *Org. Lett.* **2006**, *8*, 1803–1806.

(20) Gareiss, P. C.; Miller, B. L. In *Dynamic Combinatorial Chemistry*; John Wiley & Sons, Inc., 2009; pp. 83–117.

(21) Klekota, B.; Miller, B. L. *Tetrahedron* **1999**, *55*, 11687–11697.

(22) Whitney, A. M.; Ladame, S.; Balasubramanian, S. *Angew. Chem. Int. Ed.* **2004**, *43*, 1143–1146.

(23) Ladame, S.; Whitney, A. M.; Balasubramanian, S. *Angew. Chem. Int. Ed.* **2005**, *44*, 5736–5739.

(24) Jantos, K.; Rodriguez, R.; Ladame, S.; Shirude, P. S.; Balasubramanian, S. *J. Am. Chem. Soc.* **2006**, *128*, 13662–13663.

(25) Bugaut, A.; Toulmé, J.-J.; Rayner, B. *Angew. Chem. Int. Ed.* **2004**, *43*, 3144–3147.

(26) Komeda, S.; Moulaei, T.; Woods, K. K.; Chikuma, M.; Farrell, N. P.; Williams, L. D. *J. Am. Chem. Soc.* **2006**, *128*, 16092–16103.

(27) Komeda, S.; Moulaei, T.; Chikuma, M.; Odani, A.; Kipping, R.; Farrell, N. P.; Williams, L. D. *Nucleic Acids Res.* **2011**, *39*, 325–336.

(28) Meyer, C. D.; Joiner, C. S.; Stoddart, J. F. *Chem. Soc. Rev.* **2007**, *36*, 1705–1723.

(29) Belowich, M. E.; Stoddart, J. F. *Chem. Soc. Rev.* **2012**, *41*, 2003–2024.

(30) Godoy-Alcántar, C.; Yatsimirsky, A. K.; Lehn, J.-M. *J. Phys. Org. Chem.* **2005**, *18*, 979–985.

(31) Cheeseman, J. D.; Corbett, A. D.; Gleason, J. L.; Kazlauskas, R. J. *Chem. Eur. J.* **2005**, *11*, 1708–1716.

(32) Kubik, S. *Chem. Soc. Rev.* **2010**, *39*, 3648–3663.

(33) Paolantoni, D.; Rubio-Magnieto, J.; Cantel, S.; Martinez, J.; Dumy, P.; Surin, M.; Ulrich, S. *Chem. Commun.* **2014**, *50*, 14257–14260.

(34) In *Chapitre 2*.

(35) Borch, R. F.; Bernstein, M. D.; Durst, H. D. *J. Am. Chem. Soc.* **1971**, *93*, 2897–2904.

(36) Zameo, S.; Vauzeilles, B.; Beau, J.-M. *Eur. J. Org. Chem.* **2006**, *2006*, 5441–5444.

(37) Lide, D. R. *CRC Handbook of Chemistry and Physics, Internet Version 2005*; CRC Press, Boca Raton, FL, 2005.

(38) Ohara, K.; Smietana, M.; Vasseur, J.-J. *J. Am. Soc. Mass Spect.* **2006**, *17*, 283–291.

(39) Janssen, P. G. A.; Vandenbergh, J.; van Dongen, J. L. J.; Meijer, E. W.; Schenning, A. P. H. J. *J. Am. Chem. Soc.* **2007**, *129*, 6078–6079.

(40) Surin, M.; Janssen, P. G. A.; Lazzaroni, R.; Leclère, P.; Meijer, E. W.; Schenning, A. P. H. J. *Adv. Mat.* **2009**, *21*, 1126–1130.

(41) Janssen, P. G. A.; Jabbari-Farouji, S.; Surin, M.; Vila, X.; Gielen, J. C.; de Greef, T. F. A.; Vos, M. R. J.; Bomans, P. H. H.; Sommerdijk, N. A. J. M.; Christianen, P. C. M.; Leclère, P.; Lazzaroni, R.; van der Schoot, P.; Meijer, E. W.; Schenning, A. P. H. J. *J. Am. Chem. Soc.* **2008**, *131*, 1222–1231.

(42) Janssen, P. G. A.; Ruiz-Carretero, A.; González-Rodríguez, D.; Meijer, E. W.; Schenning, A. P. H. J. *Angew. Chem. Int. Ed.* **2009**, *48*, 8103–8106.

(43) Neidlein, U.; Leumann, C. *Tetrahedron Lett.* **1992**, *33*, 8057–8060.

(44) Tollin, P.; Walker, R. T.; Wilson, H. R. *Nucleic Acids Res.* **1984**, *12*, 8345–8347.

(45) Oligo Analyzer http://eu.idtdna.com/ANALYZER/Applications/OligoAnalyzer/Default.aspx (accessed Jul 15, 2014).

(46) OligoCalc: Oligonucleotide Properties Calculator http://www.basic.northwestern.edu/biotools/oligocalc.html (accessed Jul 15, 2014).

(47) Li, X.; Zhan, Z.-Y. J.; Knipe, R.; Lynn, D. G. *J. Am. Chem. Soc.* **2002**, *124*, 746–747.

(48) De Bruin, B.; Hauwert, P.; Reek, J. N. H. *Angew. Chem. Int. Ed.* **2006**, *45*, 2660–2663.

(49) Hunt, R. A. R.; Otto, S. *Chem. Commun.* **2011**, *47*, 847–858.

(50) Otto, S. *Acc. Chem. Res.* **2012**, *45*, 2200–2210.

CONCLUSION GENERALE

Les travaux de recherche présentés dans ce manuscrit visent à mieux connaître les interactions au sein d'auto-assemblages petites molécules – ADN dans le but de faciliter la conception de futurs ligands ou vecteurs de l'ADN. Pour se faire, des composés guanidiniums bisfonctionnalisés ont été conçus pour interagir avec le squelette phosphodiester de la matrice ADN et sonder les interactions secondaires induites par les différents substituants introduits. Deux approches différentes, rationnelle et combinatoire dynamique, ont été utilisées pour étudier les interactions au sein des auto-assemblages formés. Parallèlement, des polymères covalents dynamiques basés sur les liaisons acylhydrazones ont été conçus pour interagir avec l'ADN : le caractère dynamique de ces polymères est en effet une propriété intéressante pour la conception de vecteurs « intelligents » dont la dégradation peut être contrôlée pour libérer l'ADN transporté.

L'approche rationnelle a consisté à synthétiser trois guanidiniums bisfonctionnalisés avec des groupements aromatiques de taille variable et à les étudier par spectroscopie et spectrométrie de masse, en présence d'ADN. Les résultats obtenus par les spectroscopies de fluorescence et de dichroïsme circulaire démontrent clairement que le composé bispyrène **GuaBiPy** est capable de générer des complexes supramoléculaires en présence d'oligonucléotides. La stabilité de ces auto-assemblages est dépendante de la longueur de l'ADN et de la température.

Les effets de tampons indiquent que cette association non-covalente est gouvernée par les interactions guanidinium – phosphodiester, suggérant ainsi que ces guanidiniums s'empilent le long du squelette phosphodiester de l'ADN. Les analyses par spectrométrie de masse MALDI-ToF d'expériences de compétition entre les guanidiniums bisfonctionnalisés démontrent l'importance des interactions d'empilement dans les auto-assemblages matricés par l'ADN.

L'approche combinatoire dynamique permet de sélectionner *in situ* les ligands dont les complexes avec une cible ADN sont les plus stabilisés sans passer par les étapes de synthèse nécessaires avec l'approche rationnelle ou une approche combinatoire « statique ».

Dans le cadre de cette approche combinatoire dynamique, un guanidinium comportant deux terminaisons aldéhydes a été synthétisé et utilisé pour générer des bibliothèques combinatoires dynamiques basées sur la liaison imine avec différentes amines en présence d'oligo(dT) sur cellulose. Pour des raisons de stabilité des constituants, les bibliothèques ont été figées par amination réductrice.

Des différences de rétention des constituants des bibliothèques ont été mises en évidence en fonction des substituants introduits par les différentes monoamines testées. L'importance de la rétention, du constituant le plus retenu au moins retenu sur l'oligo(dT) sur cellulose, suit la tendance : substituant cationique > aromatique > aliphatique >> anionique.

L'effet compétiteur des ions phosphates vis-à-vis de la reconnaissance des constituants guanidinium par les phosphodiesters de l'ADN a également été mis en évidence avec cette approche, confirmant une fois de plus l'interaction des guanidiniums avec les phosphodiesters.

Des bibliothèques combinatoires dynamiques générées avec différentes diamines ont mis en évidence, dans les conditions utilisées, l'amplification, en présence d'oligo(dT) sur cellulose, de constituants linéaires dont la nature dépend de la longueur de la diamine.

Les informations structurales obtenues démontrent l'importance des interactions secondaires dans le processus de reconnaissance et d'auto-assemblage avec l'ADN. Ces données viennent compléter les informations obtenues par l'approche rationnelle.

Parallèlement à ces études des interactions guanidiniums bisfonctionnalisés – ADN, des polymères covalents dynamiques poly-acylhydrazones combinant des moitiés cationiques et de courts groupements d'éthylène glycol au sein de leur chaîne principale ont été conçus et synthétisés. Ces matériaux sont capables de s'auto-assembler par une réaction de polycondensation. L'auto-assemblage des différents monomères conduit principalement à la formation de macrocycles à basse concentration et d'oligomères à chaîne ouverte à haute concentration.

La dégradation catalysée par l'acide de ces polymères a été mise en évidence dans des solutions tampons aqueuses au pH compris entre 5,0 et 7,0. De plus, une complexation efficace de l'ADN double brin est observée dans les tampons aqueux comme dans le sérum biologique à des ratios N/P comparables à ceux des polymères PEI.

Ces matériaux hybrides auto-assemblés combinent donc deux caractéristiques intéressantes pour une application en tant que vecteurs « intelligents » pour la délivrance de gènes : une complexation efficace de l'ADN double brin dans les médias biologiques, et une dégradation catalysée en milieu acide qui pourrait promouvoir la sortie de l'endosome de l'ADN transporté.

Les études des auto-assemblages guanidiniums – ADN constituent donc un premier pas vers une meilleure connaissance des interactions secondaires stabilisant des petites

molécules reconnaissant le squelette phosphodiester de l'ADN. Les informations collectées pourront ainsi permettre de développer de nouveaux ligands ou transporteurs universels de l'ADN.

Les polymères covalents dynamiques permettent une reconnaissance multivalente de l'ADN, ce qui peut permettre une meilleure stabilisation des complexes par rapport aux auto-assemblages avec les petites molécules seules. La capacité de complexation de l'ADN et la dégradation en milieu acide des polymères auto-assemblés conçus ici en font de bons candidats pour la conception de nouveaux transporteurs d'ADN « intelligents ».

EXPERIMENTAL PART

I. GENERAL METHODS AND MATERIAL

Materials

All reagents and solvents were purchased from commercial sources (Sigma-Aldrich, Alpha Aesar or Fisher Scientific) and used as received. Dry solvents were purchased in anhydrous quality from Fisher Scientific (Acros). For dichloromethane, amylene was the stabiliser. Oligonucleotides were either synthesized (see the oligonucleotide synthesis part below) or purchased from Eurogentec as RP-Cartridge purification or HPLC-RP purification (Ultrapure Gold, >95% purity) in dried format. Oligo(dT)-cellulose and DNA-cellulose single-stranded from calf thymus DNA were purchased from Sigma-Aldrich. Triazines **TN^n** and **TO** were kindly provided by Dr. Arnaud Gilles and Dr. Sébastien Ulrich.

TLC and column chromatography

TLC were performed on silica gel 60 F_{254} plates purchased from Merck with detection under UV light (254 nm) and / or staining with either a ninhydrine solution in ethanol or a 2,4-dinitrophenylhydrazine (DNPH) solution, followed by heating. Flash column chromatography were performed on silica gel 60 (40 – 63 µm) purchased from Merck.

NMR analysis

NMR spectra were recorded on Brucker 400 or 250 MHz instruments at the ENSCM, or on a Bruker Avance III instrument (600 MHz) at the Laboratoire de Mesures Physiques, IBMM-Université de Montpellier 2. All spectra were referenced with respect to the residual solvent peak for the deuterated solvent.[1] Data are reported as follows: chemical shift (δ, in ppm), multiplicity (s for singlet, d for doublet, t for triplet, q for quartet, quint for quintuplet, m for multiplet), coupling constant (J in Hertz) and integration.

Mass Spectrometry analysis

Electrospray mass spectrometry. ESI mass spectrometry analyses were carried out on a Waters Micromass ZQ mass spectrometer (positive mode). High resolution mass spectrometry analyses were carried out at the Laboratoire de Mesures Physiques, IBMM – Université Montpellier 2, and were obtained on a Waters Micromass QTof mass spectrometer (positive mode). For synthesized oligonucleotides, ESI mass spectrometry analyses (negative mode) were carried out at the Plateau de Spectrométrie de Masse, ICGM – Université Joseph Fourier

(Grenoble) and obtained on an Esquire 3000 (Bruker): the multiply charged data produced by the mass spectrometer on the m/z scale were converted to the molecular weight.

MALDI-ToF mass spectrometry. MALDI mass spectra were performed on an Ultraflex III ToF/ToF instrument (Bruker Daltonics, Wissembourg, France) at the Laboratoire de Mesures Physiques, IBMM – Université Montpellier 2.

FT-IR spectroscopy

All IR spectra were measured on a Perkin Elmer Spectrum 100 FT-IR spectrometer equipped with an universal ATR sampling accessory. Wavenumbers $\tilde{\nu}$ are indicated in cm^{-1} and band intensities are reported in brackets as weak (w), medium (m) or strong (s).

HPLC

Analytical RP-HPLC analyses were performed on Waters instruments 515 HPLC pump connected to Waters 2487 dual λ absorbance detector [HPLC 1] or on Waters 2695 HPLC separation module connected to Waters 996 photodiode array detector [HPLC 2]. For analytical analysis, both were equipped with a Nucleosil C18 column (Macherey-Nagel, 300 Å, 5µm, 125 mm x 3 mm). Unless specified otherwise, analyses were performed with the following linear gradients of eluent B (acetonitrile/water/TFA 90%/9.9%/0.1%) in eluent A (water/TFA 99.9%/0.1%).

Method I: 5% B → 100% B in 5 minutes, flow 1 mL/min.

Method II: 5% B → 100% B in 20 minutes, flow 1 mL/min.

Semi-preparative HPLC were performed on HPLC 1 with a Macherey-Nagel VP Nucleodur C18 column (HTech 7 µm, 250 mm x 21 mm). Method details are indicated when appropriate.

Retention time (t_R) are given in minutes.

LC-MS

LC-MS analysis were performed either on Waters instruments 2695 HPLC separation module equipped with a Nucleosil C18 column (Macherey-Nagel, 300 Å, 5 µm 125 mm x 3 mm), connected to Waters 996 photodiode array detector and Waters micromass ZQ mass spectrometer [LC-MS 1] or on Shimadzu LCMS-2020 equipped with a Kinetex C18 column (Phenomenex, 100 Å, 2.6 µm, 75 mm x 3 mm) [LC-MS 2].

For LC-MS 1, eluent A is water/TFA 99.9%/0.1% and eluent B is acetonitrile/water/TFA 90%/9.9%/0.1%. Flow is 1 mL/min and column is heated at 30°C. LC methods follow linear gradient of eluent B in eluent A indicated below:

Method III: 5 % B → 100% B in 5 minutes.

Method IV: 5% B → 100% B in 20 minutes, flow 1 mL/min.

Method V: 5% B → 30% B in 40 minutes, then 30% B → 100% B in 5 minutes.

Method VI: 5% B → 50% B in 40 minutes, then 50% B → 100% B in 5 minutes.

For LC-MS 2, eluent A' is water/TFA 99.9%/0.1% and eluent B' is acetonitrile/TFA 99.9%/0.1%. Flow is 1 mL/min and column is heated at 40°C. LC methods follow linear gradient of eluent B' in eluent A' indicated below:

Method VII: 5 % B' → 100% B' in 5 minutes.

Method VIII: 5% B' → 30% B' in 45 minutes, then 30% B' → 100% B' in 5 minutes.

Method IX: 5% B' → 50% B' in 45 minutes, then 50% B' → 100% B' in 5 minutes.

UV-Vis, circular dichroism and fluorescence

UV-Vis spectroscopy for compound characterization. UV-Vis absorption was recorded using a HP 8453 UV-Vis spectrophotometer. *Fluorescence (ethidium bromide assays).* Fluorescence analyses were carried out on a AF-2500 HITACHI fluorescence spectrophotometer. *Studies of guanidinium – DNA interactions (related to chapter 2).* UV-Vis absorption, circular dichroism, and fluorescence measurements were recorded using a ChirascanPlus CD Spectrophotometer from Applied Photophysics (UK). The measurements were carried out using 2 mm Suprasil quartz cells from Hellma Analytics. Fluorescence measurements were carried out at different temperatures by using a 10 mm quartz cells (1 mL) from Lightpath Optical. The variable temperature spectroscopic experiments were performed using a TC125 Temperature Controller from Quantum Northwestern running on the ChirascanPlus Spectrophotometer. The temperatures were varied from -5ºC to 80ºC at rate of 0.5ºC/min.

Oligonucleotide synthesis

Oligonucleotides were synthesized using β-cyanoethylphosphoramidite chemistry on an ABI 3400 DNA synthesizer at 1 μmol scale. Purity was checked with HPLC 1 in the following conditions (method Oligo): linear gradient of eluent B" (acetonitrile/water 95%/5%) in eluent A" (50 mM triethylammonium acetate pH 7 buffer/acetonitrile 95%/5%) from 100% A" to 30% B" in 20 minutes, flow 1 mL/min, UV monitoring at 260 and 280 nm.

Oligonucleotide sequences

dT_{10}: 5'-TTT TTT TTT T-3'

dT_{20}: 5'-TTT TTT TTT TTT TTT TTT TT-3'

dT_{40}: 5'-TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT T-3'

dR_{43}: 5'-CGT CAC GTA AAT CGG TTA ACA AAT GGC TTT CGA AGC TAG CTT C-3'

$dR_{43}.dR_{rev43}$:

5'-CGT CAC GTA AAT CGG TTA ACA AAT GGC TTT CGA AGC TAG CTT C-3'

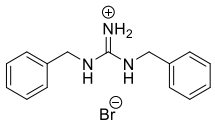
3'-GCA GTG CAT TTA GCC AAT TGT TTA CCG AAA GCT TCG ATC GAA G-5'

II. SYNTHESES AND CHARACTERIZATION

A. GUANIDINIUM COMPOUNDS

1. Guanidiniums for the rational approach

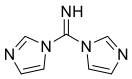
Bis-N,N-benzylguanidinium bromide (**GuaBiPhe**)



The synthetic procedure was adapted from a previous report.[2]

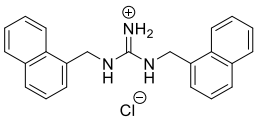
A solution of cyanogen bromide (103 mg, 0.96 mmol) in anhydrous acetonitrile (1 mL) was added dropwise to a solution of benzylamine (0.21 mL, 1.92 mmol) in anhydrous acetonitrile (2.5 mL) cooled at 0°C. The reaction mixture was then refluxed for 14 hours. The suspension was then filtrated and the white solid was washed with acetonitrile and diethyl ether. The filtrate was concentrated *in vacuo* and the resulting oil was crystallized in dichloromethane at -20°C. The solid was filtered, washed with the minimum volume of dichloromethane and dried *in vacuo*, yielding **GuaBiPhe** as white crystals (165 mg, 54% yield). ^{1}H NMR (CD_3OD, 400 MHz): δ 7.39-7.27 (m, 5H, $CH_{(aro)}$), 4.46 0(s, 2H, CH_2NH); ^{13}C NMR (CD_3OD, 63 MHz): δ 157.4, 137.6, 129.8, 128.8, 127.7, 45.9; HR-ESI-MS(+): m/z calcd for $C_{15}H_{18}N_3^+$ ($[M]^+$): 240.1501, found: 240.1500; IR (ATR): $\tilde{\nu}$ = 3323 (w), 3157 (m), 3025 (w), 2876 (w), 1644 (m), 1623 (s), 1595 (m), 1496 (m), 1469 (m), 1441 (m), 1358 (m), 1325 (w), 1245 (w), 1206 (w), 1136 (w), 1103 (w), 1057 (w), 1029 (w), 989 (w), 1029 (w), 989 (w), 961 (w), 903 (w), 875 (w), 750 (m), 740 (m), 720 (m), 702 (s), 673 (m), 621 (w).

Di(1*H*-imidazol-1-yl)methanamine (**2**)



This compound was prepared as previously described and the characterization were conformed to the literature data.[3] To a solution of imidazole (5.52 g, 84.9 mmol) in dry dichloromethane (408 mL) under argon was added a 3M solution of cyanogen bromide in dichloromethane (9.44 mL, 28.3 mmol). The reaction mixture was refluxed for 35 minutes and the formation of a white precipitate was observed. The suspension was then filtrated. The filtrate was concentrated *in vacuo* to 10% of its initial volume and crystallized at -20°C. The resulting suspension was filtrated, washed with cold dichloromethane and dried *in vacuo*, yielding compound **2** as white crystals (3.32 g, 73% yield). ^{1}H NMR (DMSO-*d6*, 400 MHz): δ 10.19 (s, 1H, N*H*), 8.08 (d, J = 20.8 Hz, 2H, CNC*H*CHN), 7.56 (d, J = 29.3 Hz, 2H, CNCHC*H*N), 7.11 (s, 2H, NC*H*N); ^{13}C NMR (DMSO-*d6*, 100 MHz): δ 140.9, 137.4, 129.6, 118.9.

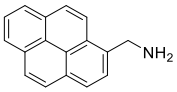
Bis-N,N-(naphtalen-1-ylmethyl)guanidinium chloride (**GuaBiNaph**)



The procedure was adapted from a previous report.[3]

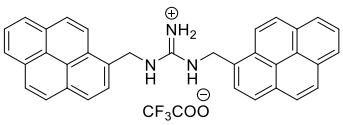
A mixture of di(1*H*-imidazol-1-yl)methanamine (**2**) (203 mg, 1.26 mmol), 1- naphtylmethylamine (0.36 mL, 2.52 mmol) and TFA (74.5 µL, 0.98 mmol) in a screwed-cap tube was stirred at 105°C for 4 hours. Once cooled to room temperature, the obtained yellow-brownish oil was dissolved in dichloromethane (20 mL). The organic solution was washed with saturated ammonium chloride solution (20 mL). The aqueous phase was extracted with dichloromethane (3 x 10 mL). The organic phases were combined and washed with a saturated aqueous ammonium chloride solution, dried over sodium sulfate and concentrated *in vacuo*. The residue was purified by flash chromatography on silica gel (dichloromethane/methanol 100:0 → 90:10, v/v). Further recrystallization in chloroform/acetonitrile yielded the desired **GuaBiNaph** as a white solid (145 mg, 31% yield). LC-MS (LC-MS 1): t_R 4.30 minutes (method III), extracted m/z calcd for $C_{23}H_{22}N_3^+$ ($[M+H]^+$): 340.18, found: 340.62; ^{1}H NMR (CD_3OD, 400 MHz): δ 7.99-7.87 (m, 6H, C*H*$_{(aro)}$), 7.59-7.44 (m, 8H, C*H*$_{(aro)}$), 4.95 (s, 4H, C*H*$_2$NH); ^{13}C NMR (CD_3OD, 63 MHz): δ 157.7, 135.4, 132.7, 132.4, 130.0, 130.0, 127.8, 127.2, 126.5, 126.4, 123.9, 44.6; HR-ESI-MS(+): m/z calcd for $C_{23}H_{22}N_3^+$ ($[M]^+$): 340.1814, found: 340.1815; IR (ATR): $\tilde{\nu}$ = 3290 (m), 3146 (m), 3055 (m), 3014 (m), 1679 (m), 1660 (s), 1639 (s), 1623 (s), 1600 (s), 1513 (m), 1451 (w), 1399 (w), 1362 (w), 1324 (w), 1264 (w), 1216 (w), 1166 (w), 1110 (m), 1030 (w), 1005 (w), 883 (w), 853 (w), 786 (s), 765 (s), 731 (m), 645 (m).

1-pyrenemethylamine



1-pyrenemethylamine was prepared from 1-pyrenemethylammonium chloride as follows:[4] 1-pyrenemethylammonium chloride (500 mg, 1.86 mmol) was suspended in an aqueous 10% sodium hydroxide solution (10 mL). Extraction with dichloromethane (10 mL), drying over sodium sulfate and concentration *in vacuo* afforded the desired compound as a pale yellow solid (431 mg, 99% yield). ^{1}H NMR ($CDCl_3$, 400 MHz): 8.12-8.08 (m, 3H, C*H*$_{(aro)}$), 8.00-7.92 (m, 5H, C*H*$_{(aro)}$), 7.90 (d, J = 7.8 Hz, 1H, C*H*$_{(aro)}$), 4.36 (s, 2H, C*H*$_2$NH).

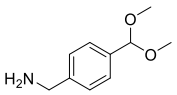
Bis-N,N-(pyren-1-ylmethyl)guanidinium trifluoroacetate (**GuaBiPy**)



The procedure was adapted from a previous report.[3] A mixture of di(1*H*-imidazol-1-yl)methanamine (**2**) (87.8 mg, 0.54 mmol), 1-pyrenemethylamine (**3**) (241.5 mg, 1.13 mmol) and TFA (32.7 µL, 0.42 mmol) in a screwed-cap tube was stirred at 105°C for 4 hours. Once cooled to room temperature, the solidified orange - brown oil was re-suspended in dichloromethane. The solid was filtrated, further washed with dichloromethane and diethyl ether, and dried *in vacuo* to obtain **GuaBiPy** as a powder (19.7 mg, 24% yield). LC-MS (LC-MS 1): t_R 5.34 minutes (method III), 97.8% peak area at 254 nm; extracted m/z calcd for $C_{35}H_{26}N_3^+$ ($[M]^+$): 488.21, found: 488.67; ^{1}H NMR (DMSO-*d6*, 400 MHz): δ 8.34-8.06 (m, 20H, C*H*$_{(aro)}$ and N*H*/N*H*$_2^+$), 7.90 (broad s, 2H, N*H*/N*H*$_2^+$), 5.25 (d, J = 4.7 Hz, 4H, C*H*$_2$NH); ^{13}C NMR (DMSO-*d6*, 63 MHz): δ 156.1, 130.8, 130.5, 130.3, 130.2, 128.0, 127.9, 127.5, 127.4, 126.5, 125.6, 125.5 (2C), 124.8, 124.1, 123.8, 122.8, 43.0; HR-ESI-MS(+): m/z calcd for $C_{35}H_{26}N_3^+$ ($[M+H]^+$): 488.2127, found: 488.2130; IR (ATR): $\tilde{\nu}$ = 3310 (m), 3129 (m), 3041 (m), 1663 (s), 1632 (s), 1489 (w), 1459 (w), 1433 (w), 1416 (w), 1392 (w), 1370 (w), 1341 (w), 1298 (w), 1207 (s), 1172 (s), 1132 (s), 958 (w), 888 (w), 834 (s), 801 (m), 752 (w), 724 (m), 702 (m), 678 (w), 656 (w).

2. Guanidinium for polymers and dynamic combinatorial libraries

4-(dimethoxymethyl)benzylamine (**20**)

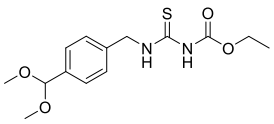


The procedure was adapted from a previous report.[5] To a solution of 4-cyanobenzaldehyde (3.79 g, 28.9 mmol) in dry methanol (50 mL) under argon were added trimethylorthoformate (19 mL, 173.4 mmol) and concentrated aqueous hydrochloric acid (260 µL, 8.67 mmol). The reaction solution was heated at 45°C for 3h30.

The solvent was then removed under reduced pressure and the residue was diluted with ethyl acetate (50 mL), washed with saturated sodium bicarbonate solution (2 × 50 mL), and brine (2 × 50 mL), dried over anhydrous sodium sulfate, and concentrated *in vacuo*. The resulting colourless oil (yield 98%) was used in the next step without further purification. Characterizations were in agreement with published data for 4-(dimethoxymethyl)benzonitrile.[5] R_f 0.50 (petroleum ether/ethyl acetate 8:2, v/v); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.56 (d, 2H, J = 8.1 Hz, C*H*$_{(aro)}$), 7.48 (d, 2H, J = 8.1 Hz, C*H*$_{(aro)}$), 5.34 (s, 1H, C*H*$(OCH_3)_2$), 3.22 (s, 6H, C*H*$_3$).

The following procedure was adapted from a previous report.[6] A solution of the previously obtained 4-(dimethoxymethyl)benzonitrile (505 mg, 2.8 mmol) in anhydrous THF (3.5 mL) was then added dropwise to a stirred suspension of lithium aluminium hydride (215 mg, 5.6 mmol) in anhydrous THF (3.4 mL) under argon at 0°C. The reaction mixture was stirred at room temperature overnight. The mixture was then cooled at 0°C and basified by addition of an aqueous 1 M sodium hydroxide solution (2 mL). The mixture was filtrated on celite and rinsed with ethyl acetate (30 mL). The organic filtrate was washed with a saturated sodium bicarbonate solution (20 mL) and brine (20 mL), dried over anhydrous sodium sulfate, and concentrated *in vacuo* to obtain **4** as a yellow oil (93% yield) which was directly used in the next step without further purification. Characterizations were in agreement with published data for compound **20**.[6] ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.41 (d, 2H, J = 8.1 Hz, C*H*$_{(aro)}$), 7.31 (d, 2H, J = 8.1 Hz, C*H*$_{(aro)}$), 5.38 (s, 1H, C*H*$(OCH_3)_2$), 3.87 (s, 2H, PhC*H*$_2$N), 3.32 (s, 6H, C*H*$_3$).

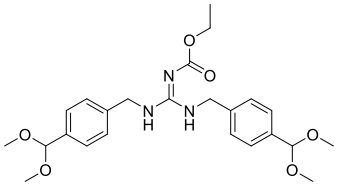
<u>1-ethyloxycarbonyl-3-[4-(dimethoxymethyl)benzyl]thiourea (**24**)</u>



To a solution of ethoxycarbonyl isothiocyanate (50 µL, 0.42 mmol) and triethylamine (70 µL, 0.50 mmol) in dry dichloromethane (1 mL) was added under argon a solution of **20** (83 mg, 0.46 mmol) in dry dichloromethane (1.5 mL). The reaction mixture was stirred at room temperature for 2 hours, and then concentrated *in vacuo.* The crude material was purified by flash chromatography (petroleum ether/ethyl acetate 8:2, v/v) to provide **24** (yield 90%) as a white solid. R_f 0.36 (petroleum ether/ethyl acetate 8:2, v/v); ^{1}H NMR (250 MHz, $CDCl_3$ neutralized by K_2CO_3) δ 9.98 (s broad, 1H, N*H*), 8.60 (s broad, 1H, N*H*), 7.43 (d, 2H, J = 7.5 Hz, C*H*$_{(aro)}$), 7.33 (d, 2H, J = 7.5 Hz, C*H*$_{(aro)}$), 5.38 (s, 1H, C*H*$(OCH_3)_2$), 4.85 (d, 2H, J = 5.0 Hz, PhC*H*$_2$N), 4.18 (q, 2H, J = 7.5 Hz, OC*H*$_2$CH$_3$), 3.30 (s, 6H, OC*H*$_3$), 1.27 (t, 3H, J = 7.5 Hz, OCH$_2$C*H*$_3$); ^{13}C NMR (63

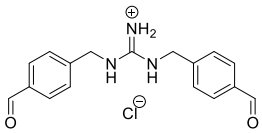
MHz, $CDCl_3$ neutralized by K_2CO_3) δ 179.5, 152.9, 137.8, 136.6, 127.7, 127.3, 102.7, 62.8, 52.7, 49.2, 14.2.

1,3-bis[4-dimethoxymethyl)benzyl]-2-ethoxycarbonyl guanidine (**25**)



To a solution of **24** (51 mg, 0.16 mmol) in anhydrous DMF (1 mL) were added triethylamine (230 µL, 1.6 mmol) and N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (38 mg, 0.19 mmol). The reaction mixture was stirred for 30 minutes at room temperature and then a solution of **20** (30 mg, 0.16 mmol) in anhydrous DMF (900 µL) was added. The reaction mixture was heated at 60°C overnight. The solvent was then removed under reduced pressure and the crude material was purified by flash chromatography (petroleum ether/ethyl acetate 5:5, v/v) to provide **25** (yield 84%) as a colourless oil. R_f 0.36 (petroleum ether/ethyl acetate 5:5, v/v); ^{1}H NMR (250 MHz, $CDCl_3$ neutralized by K_2CO_3) δ 7.37 (d, 4H, J = 7.7 Hz, C*H*$_{(aro)}$), 7.18 (s broad, 4H, C*H*$_{(aro)}$), 5.34 (s, 2H, C*H*$(OCH_3)_2$), 4.43 (s broad, 4H, PhC*H*$_2$NH), 4.11 (q, 2H, J = 7.1 Hz, OC*H*$_2$CH$_3$), 3.29 (s, 12H, OC*H*$_3$), 1.29 (t, 3H, J = 7.1 Hz, OCH$_2$C*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$ neutralized by K_2CO_3) δ 164.6, 160.3, 137.9 (2C), 127.4, 127.1, 102.9, 60.9, 52.8, 45.1, 14.7; HR-ESI-MS(+): m/z calcd for $C_{24}H_{34}N_3O_6^+$ ([M+H]$^+$): 460.2448, found: 460.2450.

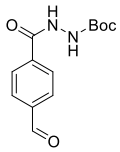
1,3-bis-(4-formylbenzyl)guanidinium chloride (**Gua-Ald**)



Protected guanidine **6** (1.12 g, 2.44 mmol) was suspended in 1 M aqueous hydrochloric acid solution (12 mL) and the solution was heated at reflux for 13.5 hours. The reaction mixture was concentrated *in vacuo* and the crude material was purified by flash chromatography (acetonitrile/water 98:2 → 97:3, v/v) to provide, after freeze-drying, **Gua-Ald** as a white solid (yield 86%). R_f 0.66 (acetonitrile/water 8:2, v/v); HPLC (HPLC 2): t_R 2.84 min (method I); ^{1}H NMR (600 MHz, DMSO-*d6*) δ 10.0 (s, 2H, C*H*O), 7.89 (d, 4H, J = 12.0 Hz, C*H*$_{(aro)}$), 7.50 (d, 4H, J = 12.0 Hz, C*H*$_{(aro)}$), 4.61 (s, 4H, PhC*H*$_2$N); ^{13}C NMR (100 MHz, DMSO-*d6*) δ 192.8, 156.3, 144.4, 135.4, 129.7, 127.6, 43.8; HR-ESI-MS(+): m/z calcd for $C_{17}H_{18}N_3O_2^+$ ([M]$^+$): 296.1399, found: 296.1397; UV molar extinction coefficients ε_{198} = 42226 L.mol^{-1}.cm^{-1}, ε_{255} = 28631 L.mol^{-1}.cm^{-1}.

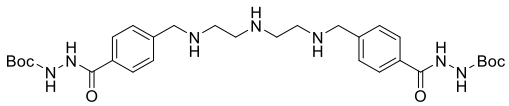
B. Synthesis of hydrazide and second bisaldehyde compounds (Dr. Camille Bouillon)

Tert-butyl 2-(4-formylbenzoyl)hydrazine-1-carboxylate (**16**)



To a solution of 4-formyl benzoic acid (100 mg, 0.67 mmol) in dry toluene (3 mL) under argon was added thionyl chloride (145 µL, 2.01 mmol). The reaction solution was stirred at reflux overnight under argon. The solvent was removed under reduced pressure in dry conditions to afford the crude 4-formyl benzoyl chloride as a yellow solid. Then *tert*-butyl carbazate (88 mg, 0.67 mmol) and triethylamine (139 µL, 1.00 mmol) were added to a stirred solution of the crude acyl chloride in dry dichloromethane (3 mL) cooled at 0°C. The reaction solution was then warmed to room temperature and stirred for 48 hours. After removal of solvent, the residue was diluted with ethyl acetate (15 mL), washed with saturated sodium bicarbonate solution (15 mL), and brine (15 mL), dried over anhydrous sodium sulfate, and concentrated *in vacuo*. The resulting crude material was purified by flash chromatography (petroleum ether/ethyl acetate 7:3 → 5:5, v/v) to provide **16** (yield 55%) as a white solid. R_f 0.31 (petroleum ether/ethyl acetate 6:4, v/v); HPLC (HPLC 2): t_R 3.16 min (method I); ^{1}H NMR (400 MHz, $CDCl_3$) δ 10.04 (s, 1H, C*H*O), 8.93 (s broad, 1H, N*H*), 7.94 (d, 2H, J = 8.0 Hz, C$H_{(aro)}$), 7.87 (d, 2H, J = 8.0 Hz, C$H_{(aro)}$), 6.96 (s broad, 1H, N*H*), 1.48 (s, 9H, C(CH_3)$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ 191.5, 165.8, 156.2, 138.6, 136.8, 129.9, 128.2, 82.6, 28.3; ESI-MS: m/z calcd for $C_{13}H_{17}N_2O_4^+$ ([M+H]$^+$): 265.12, found: 265.24; HR-ESI-MS(+): m/z calcd for $C_8H_9N_2O_2^+$ ([M-Boc+H]$^+$): 165.0664, found: 165.0665.

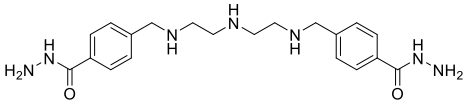
tert-butyl 2-{4-[((2-((2-[((4-{2-(*tert*butoxycarbonyl)hydrazine}benzoyl)methyl)amino] ethyl) amino) ethyl)amino)methyl]benzoyl}hydrazine-1-carboxylate (**18**)



To a solution of diethylene triamine (41 µL, 0.38 mmol) in dry methanol (2 mL) under argon was added at 0°C **16** (200 mg, 0.76 mmol). The reaction mixture was stirred at room temperature for 48 hours. Then sodium borohydride (43 mg, 1.14 mmol) was added to the solution cooled at 0°C. The reaction was warmed to room temperature and stirred for 24 hours. After removal of solvent, the obtained residue was purified by flash chromatography (*iso*-propanol/water/ammonium hydroxide 7:0.1:0.1 → 7:0.25:0.25, v/v/v) to provide **2** (yield 26%) as an oil. R_f 0.39 (*iso*-propanol/water/ammonium hydroxide 7:0.5:0.5, v/v/v); HPLC (HPLC 2): t_R 3.02 min (method I); ^{1}H NMR (400 MHz, CD_3OD) δ 7.82 (d, 4H, J = 8.0 Hz, C$H_{(aro)}$), 7.46 (d, 4H, J = 8.0 Hz, C$H_{(aro)}$), 3.82 (s, 4H, PhCH_2N), 2.72 (s, 8H, NHCH_2CH_2),

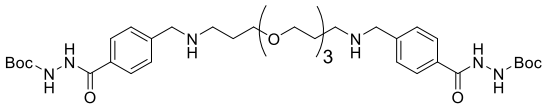
1.50 (s, 18H, C(CH_3)$_3$); ESI-MS(+): m/z calcd for $C_{30}H_{46}N_7O_6^+$ ([M+H]$^+$): 600.35, found: 600.65.

4-{((2-[(2-(((4-[(hydrazine)carbonyl]benzyl)amino)ethyl)amino]ethyl)amino)methyl}-benzhydrazide (**EI-Hyd**)



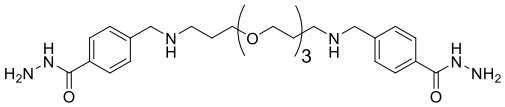
Compound **2** (25 mg, 0.42 mmol) was dissolved in a solution of TFA/water/TIS (95:2.5:2.5, v/v/v) (10 mL) and stirred for 48 hours at room temperature. After removal of 90% of the solvent, diethyl ether was added to the residual solution. The precipitate residue was triturated with diethyl ether and filtered. The material was then lyophilized twice to afford the trifluoroacetate salt **EI-Hyd** as a sticky solid in 80% yield. The yield is calculated by a ^{1}H NMR titration method using *t*-butanol as an internal reference. ^{1}H NMR (600 MHz, DMSO-*d6*) δ 7.92 (d, 4H, J = 8.3 Hz, C$H_{(aro)}$), 7.64 (d, 4H, J = 8.3 Hz, C$H_{(aro)}$), 4.29 (s, 4H, PhCH_2N), 3.29-3.27 (m, 8H, NHCH_2CH_2); ^{13}C NMR (150 MHz, DMSO-*d6*) δ 165.3, 158.8 (TFA), 135.9, 132.1, 130.0, 127.7, 116.0 (TFA), 49.8, 43.2 (2C); HR-ESI-MS(+): m/z calcd for $C_{20}H_{30}N_7O_2^+$ ([M+H]$^+$): 400.2461, found: 400.2456.

tert-butyl 2-{4-[((13-[((4-{2-(*tert*butoxycarbonyl)hydrazine}benzoyl)methyl)amino]-4,7,10-trioxatridecyl)amino)methyl]benzoyl}hydrazine-1-carboxylate (**19**)



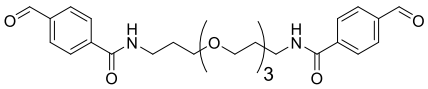
To a solution of 4,7,10-trioxa-1,13-tridecanediamine (200 mg, 0.91 mmol) in dry methanol (5 mL) under argon was added at 0°C **16** (480.10 mg, 1.82 mmol). The reaction solution was stirred at room temperature for 48 hours. Then sodium borohydride (103.3 mg, 2.73 mmol) was added to the solution cooled at 0°C. The reaction was warmed to room temperature and stirred for 24 hours. After removal of solvent, the residue was purified by flash chromatography (*iso*-propanol/water/ammonium hydroxide 7:0.1:0.1 → 7:0.2:0.2, v/v/v) to provide **19** as an oil in 22% yield. R_f 0.27 (*iso-propanol*/water/ammonium hydroxide 7:0.2:0.2, v/v/v); HPLC (HPLC 2): t_R 2.64 min (method I); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.65 (d, 4H, J = 8.0 Hz, C$H_{(aro)}$), 7.24 (d, 4H, J = 8.0 Hz, C$H_{(aro)}$), 3.74 (s, 4H, PhCH_2N), 3.58-3.56 (m, 4H, NHCH$_2$CH$_2$$CH_2$), 3.54-3.51 (m, 8H, OC$H_2CH_2$O), 2.71 (t, 4H, J = 6.0 Hz, NHCH_2), 1.79 (quint, 4H, *J* = 6.0 Hz, NHCH$_2$CH_2CH$_2$), 1.46 (s, 18H, C(CH_3)$_3$); ESI-MS(+): m/z calcd for $C_{36}H_{57}N_6O_9^+$ ([M+H]$^+$): 717.42, found: 717.40.

4-{((13-[(4-((hydrazine)carbonyl)phenyl)amino]-4,7,10-trioxatridecyl)amino)methyl} benzhydrazide (**EG-Hyd**)



Compound **19** (145.6 mg, 0.20 mmol) was dissolved in a solution of TFA/water/TIS (95:2.5:2.5, v/v/v) (8 mL) and stirred for 24 hours at room temperature. After removal of 90% of the solvent, diethyl ether was added to the residue. The precipitate was triturated with diethyl ether and filtered. The crude material was then lyophilized twice to afford the trifluoroacetate salt **EG-Hyd** as a sticky solid in 55% yield. The yield is calculated by a ^{1}H NMR titration method using *t*-butanol as an internal reference. HPLC (HPLC 2): t_R 2.93 min (method II); ^{1}H NMR (600 MHz, DMSO-*d6*) δ 7.92 (d, 4H, J = 8.1 Hz, C*H*$_{(aro)}$), 7.62 (d, 4H, J = 8.1 Hz, C*H*$_{(aro)}$), 4.24 (s, 4H, PhC*H*$_2$N), 3.48-3.45 (m, 12H, NHCH$_2$CH$_2$C*H*$_2$OC*H*$_2$C*H*$_2$O), 3.00 (t, 4H, J = 6.0 Hz, NHC*H*$_2$), 1.88 (quint, 4H, J = 6.0 Hz, NHCH$_2$C*H*$_2$CH$_2$); ^{13}C NMR (150 MHz, DMSO-*d6*) δ 165.3, 158.7 (TFA), 136.2, 131.9, 130.0, 127.7, 116.0 (TFA), 69.7, 69.5, 67.3, 49.6, 44.6, 25.8; HR-ESI-MS(+): m/z calcd for $C_{26}H_{41}N_6O_5^+$ ([M+H]$^+$): 517.3138, found: 517.3135.

N-{13-[2-(4-formylphenyl)acetamido]-4,7,10-trioxatridecyl}-4-(formyl)benzamide (**EG-Ald**)



To a solution of 4,7,10-trioxa-1,13-tridecanediamine (500 mg, 2.27 mmol) and triethylamine (315 µL, 2.27 mmol) in dry dichloromethane (10 mL) under argon was added 4-formyl benzoic acid (681.45 mg, 4.54 mmol) and 1-hydroxybenzotriazole hydrate (920 mg, 6.81 mmol). Then the solution was cooled at 0°C and N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (1.3 g, 6.81 mmol) was added. The reaction was warmed to room temperature and stirred for 48 hours. After removal of solvent, the obtained residue was diluted with ethyl acetate (20 mL), washed with saturated sodium bicarbonate solution (20 mL) and brine (15 mL), dried over anhydrous sodium sulfate, and concentrated *in vacuo*. The crude material was purified by flash chromatography (ethyl acetate/methanol 95:5, v/v) to provide **EG-Ald** (yield 81%) as a white solid. R_f 0.37 (ethyl acetate/methanol 95:5, v/v); HPLC (HPLC 2): t_R 3.65 min (method I); ^{1}H NMR (400 MHz, DMSO-*d6*) δ 10.06 (s, 2H, C*H*O), 8.66 (t, 2H, J = 4.0 Hz, N*H*), 8.01-7.96 (m, 8H, C*H*$_{(aro)}$), 3.54-3.48 (m, 8H, OC*H*$_2$C*H*$_2$O), 3.45 (t, 4H, J = 8.0 Hz, NHCH$_2$CH$_2$C*H*$_2$), 3.35-3.30 (m, 4H, NHC*H*$_2$), 1.76 (quint, 4H, J = 8.0 Hz, NHCH$_2$C*H*$_2$CH$_2$); ^{13}C NMR (100 MHz, DMSO-*d6*) δ 193.0, 165.4, 139.7, 137.7, 129.5, 127.9, 69.8, 69.6, 68.3, 36.9, 29.3; HR-ESI-MS(+): m/z calcd for $C_{26}H_{33}N_2O_7^+$ ([M+H]$^+$): 485.2288, found: 485.2292.

C. General procedure for the syntheses of oligonucleotide dT_{10} and dT_{40}

Oligonucleotides were obtained from automated synthesis on a T-CPG resin at a 1 µmol scale (final detritylation was performed by the synthesizer). After synthesis, cyanoethyl protection groups were removed with a mixture triethylamine – acetonitrile 1/1 v/v at room temperature for 15 hours.[7] Cleavage for the CPG solid support was then performed in aqueous 28% ammonium hydroxide solution for 14 hours at 55°C. Oligonucleotides were finally purified by precipitation in ethanol. Yields were determined by UV spectroscopy at 260 nm.

dT_{10}
HPLC: t_R = 14.7 minutes (method Oligo).
UV quantification: 708 nmol, 71% yield (ε_{260} = 81 600 L.mol^{-1}.cm^{-1}).
ESI-MS(-): m/z calcd for $[dT_{10}\text{-H}]^-$ ($[\text{M-H}]^-$): 2979.0, found: 2977.6

dT_{40}
HPLC: t_R = 15.4 minutes (method Oligo).
UV quantification: 587 nmol, 59% yield (ε_{260} = 324 600 L.mol^{-1}.cm^{-1}).
ESI-MS(-): m/z calcd for $[dT_{40}\text{-H}]^-$ ($[\text{M-H}]^-$): 12104.9, found: 12104.5.

III. Spectroscopic studies for the rational approach (collaboration with Dr. Jenifer Rubio-Magnieto and Dr. Mathieu Surin)

For these spectroscopic experiments, oligonucleotides were exclusively purchased from Eurogentec as HPLC-RP purification (Ultrapure Gold, >95% purity) in dried format.

Preparation of the oligonucleotides
The buffer was prepared by using tris(hydroxymethyl)aminomethane ($(HOCH_2)_3CNH_2$), EDTA (Aldrich, HPLC grade) in MilliQ water. The purity of the oligonucleotide sequences was checked with MALDI-ToF. The oligonucleotides were dissolved in a volume of Tris-EDTA (TE) buffer at a concentration of 100 µM. The solution obtained was centrifuged during 2 minutes at 2000 rpm. 20 µL of this solution were used in order to prepare different aliquots. A solution of 280 µL of MQ water was added to each aliquot in order to obtain a

final volume of 300 µL (1.33 mM Tris buffer and 67 µM EDTA) and the final diluted solution was mixed using a vortex.

Preparation of the DNA-Guanidinium solutions

The concentration of the aliquot of DNA in TE buffer solution (1.33 mM Tris buffer and 67 µM EDTA) was determined by UV-Vis at 25 ºC using the specific extinction coefficients (ε_{260}) of each DNA, which are 81600 $L.mol^{-1}.cm^{-1}$, 162600 $L.mol^{-1}.cm^{-1}$, and 324600 $L.mol^{-1}.cm^{-1}$, 412900 $L.mol^{-1}.cm^{-1}$ and 675815 $L.mol^{-1}.cm^{-1}$ for ssDNAd$(T)_{10}$, ssDNAd$(T)_{20}$, ssDNAd$(T)_{40}$, ssDNAd$(R)_{43}$ and dsDNAd$(R)_{43}$, respectively (see part I for a description of the structure of these DNA oligonucleotides). The guanidinium compounds were dissolved in DMSO with a 10 mM concentration (stock solution). The stock solution of guanidinium was added to the DNA solution and the molar ratio between guanidinium compounds and DNA was adjusted to the DNA concentration. Both compounds were stirred using the vortex at vigorous speed during 2 minutes.

Analyses

*UV-Vis and circular dichroism***.** The spectra were recorded between 240 and 500 nm, with a bandwidth of 1 nm and time per point 1 s.

Fluorescence. The excitation wavelength was set at 348 nm. The spectra were recorded between 355 and 650 nm, with a bandwidth of 10 nm and time per point 0.5 s. The buffered water solvent reference spectra were used as baselines and were automatically subtracted from the CD spectra of the samples.

Variable temperature spectroscopic analyses. The temperature within the quartz cells was determined using a temperature probe. The rate for decreasing the temperature from 80ºC to -5ºC was of 0.5ºC/min in order to allow equilibration. For the heating/cooling cycles, the temperatures were varied from -5ºC to 80ºC using the following conditions:

- First Heating Cycle (H-C): from -5ºC to 80ºC at a rate of 10ºC/min.
- 10 minutes of stabilization at 80ºC.
- First Cooling Cycle (C-C): from 80ºC to -5ºC at a rate of 0.5ºC/min.
- Second Heating Cycle (H-C): from -5ºC to 80ºC at a rate of 0.5ºC/min.
- Second Cooling Cycle (C-C): from 80ºC to -5ºC at a rate of 0.5ºC/min.
- After about 14 hours, the solution was stirred.
- Third Heating Cycle (H-C): from -5ºC to 80ºC at a rate of 0.5ºC/min.
- Third Cooling Cycle (C-C): from 80ºC to -5ºC at a rate of 0.5ºC/min.

IV. SMALL MOLECULES-DNA COMPLEXES DETECTION BY MALDI-TOF MASS SPECTROMETRY (ANALYSES PERFORMED BY DR. SONIA CANTEL)

Oligonucleotides used for MALDI-ToF mass spectrometry analyses were exclusively purchased from Eurogentec as RP-Cartridge purification in dried format.

Sample preparation

Stock solutions of guanidinium were prepared in DMSO at a concentration of 10 mM (**GuaBiPy**) or 100 mM (**GuaBiPhe**, **GuaBiNaph**, **Gua(OH)$_2$**, **Gua-Benz$_2$**, **Gua-Agm$_2$** and **Gua-C4$_2$**). A dT_{10} (Eurogentec, RP-Cartridge purification) stock solution was prepared by dilution with MilliQ water to a concentration of 10 mM. Samples were then prepared by dilution of stock solutions to final concentrations of 1 mM in guanidinium and 0.1 mM in dT_{10} in a mixture of MilliQ water and DMSO (final proportion of DMSO: 1% with **GuaBiPhe**, **Gua(OH)$_2$**, **Gua-Benz$_2$**, **Gua-Agm$_2$** and **Gua-C4$_2$**; 40% with **GuaBiNaph** and 80% with **GuaBiPy**).

Samples with **GuaBiPhe**, **GuaBiNaph** and **GuaBiPy** were further diluted 4 times in MilliQ water prior to analyses. According to the dried droplet procedure, 0.5 µl of a solution of the 4-NA matrix in ethanol (0,1 M) was deposited on the MALDI target (AnchorChip™, Bruker), then mixed with the sample in equal amount. Sample spots were dried at room temperature.

Mass spectrometric analysis conditions for the detection of small molecules – DNA complexes with guanidiniums **GuaBiPhe**, **GuaBiNaph** and **GuabiPy**.

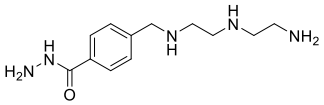
MALDI mass spectra were performed on an Ultraflex III ToF/ToF instrument (Bruker Daltonics, Wissembourg, France) equiped with LIFT capability at the Laboratoire de Mesures Physiques, IBMM – Université Montpellier 2. A pulsed Nd:YAG laser at a wavelength of 355 nm was operated at a frequency of 100 Hz. MS analyses were conducted in positive reflectron ion mode with a pulse ion extraction delay of 30 ns. An acceleration voltage of 25.0 kV (IS1) was applied for a final acceleration of 21.95 kV (IS2). Mass spectra were acquired from at least 150 laser shots, over a mass range from m/z 500 to 5000. A deflection at 2000 Da could be applied. The laser fluence was adjusted for each studied sample above the threshold for generation of molecular ions. Data were acquired with the Flex Control software and processed with the Flex Analysis software. External calibration was systematically performed

with commercial peptide mixture (Calibration peptide standard II) in a linear correction calibration.

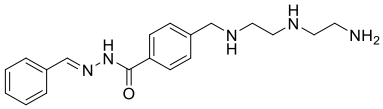
V. POLYMERS FORMATION, DEGRADATION AND COMPLEXATION WITH DNA

A. SINGLE ACYLHYDRAZONE MODEL REACTION (STUDIED BY DR. CAMILLE BOUILLON)

Syntheses of monohydrazide **26** and model compound **27**



The monohydrazide **26** was obtained after deprotection with TFA/water/TIS (95:2.5:2.5, v/v/v) of the protected hydrazide isolated as a side product during the synthesis of compound **18**. ^{1}H NMR (400 MHz, CD_3OD) δ 7.97 (d, 2H, J = 8.0 Hz, C*H*$_{(aro)}$), 7.70 (d, 2H, J = 8.0 Hz, C*H*$_{(aro)}$), 4.39 (s, 2H, PhC*H*$_2$N), 3.57 (s broad, 4H, NHC*H*$_2$C*H*$_2$NH$_2$), 3.47-3.40 (m, 4H, NHC*H*$_2$C*H*$_2$NH).



The model compound **27** was obtained by ligation in methanol of monohydrazide **26** (10 mM) with an equimolar quantity of benzaldehyde (10 mM) at room temperature. LC-MS (LC-MS 1): t_R = 2,70 minutes (method III), extracted m/z calcd for $C_{19}H_{26}N_5O^+$ ([M+H]$^+$): 340.21, found: 340.00.

Degradation of model compound **27** by exchange reaction

The stability of the model compound **27** in presence of methoxyamine was monitored by reverse-phase HPLC (HPLC 2) using method I at different pH (pH=4 100mM acetate buffer, pH=5 100mM acetate buffer, pH=6 100mM phosphate buffer, and pH=7 100mM phosphate buffer). Model compound **27** was diluted in the buffer to provide 1 mM concentration, and methoxyamine was added to the solution to provide 5 mM concentration. The solutions were kept at room temperature and analyzed by HPLC at different time points (6 µL injected). The disappearance of the acylhydrazone compound was monitored at 254 nm and quantified by peak area measurement. Data were fitted with exponential curves.

Degradation of model compound **27** by hydrolysis

The direct hydrolysis experiment was carried out similarly in the absence of methoxyamine at pH 3 (100 mM phosphate buffer), 5 (100 mM acetate buffer) and 7 (100 mM phosphate buffer).

B. Formation and degradation of polymers (studied by Dr. Camille Bouillon)

Polycondensation (formation of polymers **Poly(EI-EG)** and **Poly(Gua-EG)**)

Bishydrazide and bisaldehyde monomers are mixed in stoichiometric amounts at a final concentration of 200 mM in methanol.

Exchange reaction of polymer **Poly(Gua-EG)** with methoxyamine

The stability of **Poly(Gua-EG)** in presence of methoxyamine was monitored by reverse-phase HPLC (HPLC 2) using method II at different pH (pH = 4 100mM acetate buffer, pH = 5 100mM acetate buffer, pH = 6 100mM phosphate buffer, and pH = 7 100mM phosphate buffer). **Poly(Gua-EG)** was diluted in the buffer to a final concentration of 1 mM, and methoxyamine was added to the solution (final concentration 5 mM). The solution was kept at room temperature and analyzed by HPLC at different time points (6 µL injected). The formations of the bisoxime guanidinum **GuaOx** and bishydrazide **EG-Hyd** were monitored at 254 nm and quantified by peak area measurement. Data were fitted with exponential curves.

Hydrolysis of polymer **Poly(Gua-EG)**

The direct hydrolysis experiment was carried out similarly in the absence of methoxyamine at pH 4, 5, 6 and 7.

C. Ethidium bromide displacement assay (Dr. Sébastien Ulrich)

Samples were prepared in 3 mL quartz cuvette. Excitation wavelength was set at 546 nm and emission was measured at 590 mm. The low salt buffer contained 100 mM of HEPES, 10 µM EDTA and 9.4 mM of NaCl. The high salt buffer contained 100 mM of HEPES, 10 µM EDTA and 150 mM of NaCl. Ethidium bromide was dissolved in the buffer to provide 42 µM concentration. After standing 1 min, the fluorescence was measured. Calf-thymus DNA (ctDNA) (Sigma-Aldrich) was added to a final concentration of 50 µM. **Poly(Gua-EG)** or

Poly(EI-EG), spermine and **Gua-Ald** solutions (100-200 mM depending on the compound) were added in microliter portions. The fluorescence measurement was recorded after shaking the solution a few seconds as it was found to be sufficient to reach a stable value. The total added polymer or compound solution did not exceed 5% of the total volume, therefore, corrections were not made for sample dilution.

D. GEL RETARDATION ASSAY (DR. YANNICK BESSIN)

100 ng of 5.7 kilobase pair expression vector, pET-15b (Novagen) were mixed with appropriate amounts of polyamine to achieve the desired N/P ratio and diluted in TAE buffer (Tris – acetate – EDTA pH 8.2) to obtain a final volume of 10 µL. 2 µL of Blue 6X loading dye (Fisher Scientific) was added, after which 12 µL was run on a 0.7% agarose gel (50V). DNA was visualized with SYBER Safe (Life Technologies). The reference for the gel is GeneRuler 1 kb DNA ladder (250 to 10000 bp) from Fischer Scientific.

VI. DYNAMIC COMBINATORIAL LIBRARIES GENERATION

Concentrated solutions of components

All libraries were prepared by dilution of stock solutions of each component in the appropriate solvent. The stock solutions were at least 10 times more concentrated than the final concentration of the component in the library.

Gua-Ald stock solution was prepared in DMSO at a concentration of 100 mM.

Amines stock solutions were prepared by dilution of amine in MilliQ water, neutralization at pH 7 by addition of hydrochloric acid or sodium hydroxide and completion by MilliQ water to obtain a 100 mM solution.

Triazines $\mathbf{TN^n}$ stock solutions were prepared at a concentration of 50 mM in a mixture water/DMSO 50/50. Triazine **TO** stock solution was prepared at a concentration of 100 mM in DMSO.

Oligonucleotide were either synthesized or purchased from Eurogentec as RP-Cartridge purification in dried format. Stock solutions were prepared in MilliQ water at a concentration of 10 mM for dT_{10} and 2.5 mM for dT_{40} (100 mM in T bases for both dT_n stock solutions).

A. General methods for libraries generation with guanidinium ***Gua-Ald*** *and commercial amines/diamines*

1. With dT_n in solution

In an LC-MS vial were successively diluted in water pH 7 or 100 mM phosphate buffer pH 7: the 100 mM stock solution(s) of the chosen amine(s) (10 eq.) and the 100 mM stock solution of guanidinium **Gua-Ald** in DMSO (final concentration of 1 mM). When relevant, stock solution of oligonucleotide in MilliQ water (1 or 10 eq. in nucleotide) was added at this point. After mixing with vortex, an aqueous 100 mM sodium cyanoborohydride solution (10 µL, 10 eq.) was added to the library which was then mixed again with vortex and abandoned to room temperature the necessary time for reductive amination to be completed (no more constituants with aldehyde visible in LC-MS, ~10 days).
The libraries compositions were determined by LC-MS (method V on LC-MS 1 or method VIII on LC-MS 2).

2. With DNA on solid support

<u>Assay with oligo(dT)-cellulose or calf thymus ssDNA-cellulose:</u>
Oligo(dT)-cellulose (23.0 mg) or calf thymus ssDNA-cellulose (23.2 mg, ~1 eq. in T bases) was weighted in a microtube. To the solid were successively added MilliQ water or a 100 mM phosphate buffer at pH 6 (volume needed to reach a final volume of 300 µL), a 100 mM stock solution of guanidinium **Gua-Ald** in DMSO (3 µL, final concentration of 1 mM) and the 100 mM solution(s) of the chosen monoamine(s) (3 x n µL, n eq., function of the nature and number of amines present in the library, Table E.1). After mixing with vortex and light centrifugation, the library was heated at 60°C for 10 minutes in a dry-bath, then allowed to slowly cool back to room temperature.
After a slight centrifugation, an aqueous 100 mM sodium cyanoborohydride solution (30 µL, 10 eq.) was added to the library which was then mixed again with vortex and abandoned to room temperature the necessary time for reductive amination to be completed (no more constituants with aldehyde visible in LC-MS in the corresponding control in solution, ~10 days).

The library was centrifuged and the supernatant was taken with a syringe and directly filtrated through a PVDF 0.45 µm syringe-filter in an LC-MS vial. The composition of the supernatant was then determined by LC-MS (method V on LC-MS 1 or method VIII on LC-MS 2).

Washing of the solid support

To the solid support was added a phosphate buffer 1 M solution (330 µL). After mixing with vortex, the mixture was heated at 60°C for 10 minutes and then immediately centrifuged. The washing solution was taken with a syringe and directly filtrated through a PVDF 0.45 µm syringe-filter in a LC-MS vial for analysis (method V on LC-MS 1 or method VIII on LC-MS 2).

Control with micro-crystallized cellulose

Same procedure than for the assay with oligo(dT)-cellulose with replacement of the oligo(dT)-cellulose by the cellulose (23.0 mg). Washing procedure of solid support is the same.

Control in solution

In a microtube were successively added MilliQ water (volume needed to reach a final volume of 300 µL), a 100 mM stock solution of guanidinium **Gua-Ald** in DMSO (3 µL, final concentration of 1 mM) and the 100 mM stock solution(s) of the chosen monoamine(s) (3n µL, n eq., function of the nature and number of amines present in the library, Table E.1). After mixing with vortex and light centrifugation, the library was heated at 60°C for 10 minutes in a dry-bath, then allowed to slowly cool back to room temperature.

After a slight centrifugation, an aqueous 100 mM sodium cyanoborohydride solution (30 µL, 10 eq.) was added to the library which was then mixed again with vortex and abandoned to room temperature the necessary time for reductive amination to be completed (no more constituants with aldehyde visible in LC-MS, ~10 days).

The library was then directly transferred in a LC-MS vial and its composition determined by LC-MS (method V on LC-MS 1 or method VIII on LC-MS 2).

Table E.1 : Number of equivalents n used in the libraries as a function of the number and nature of amines present in the library.

Total number of amines in the library	Nature of the amines and n*
1	10 eq. for every amine
2	L-arginine 10 eq. + L-leucine 10 eq. L-phenylalanine tert-butyl ester 8 eq. + L-leucine methyl ester 8.8 eq. Benzylamine 10 eq. + taurine 10.2 eq. L-arginine 9.5 eq. + L-phenylalanine 14.1 eq. Agmatine 14.1 eq. + benzylamine 3 or 5 or 7 or 10 eq.

* n is the number of equivalent in comparison to the quantity of **Gua-Ald**.

*B. General methods for library generation with guanidinium **Gua-Ald** and triazines **TN**^{N}*

Assay with oligo(dT)-cellulose or calf thymus ssDNA-cellulose

Oligo(dT)-cellulose (23.0 mg) was weighted in a microtube. To the solid were successively added MilliQ water (volume needed to reach a final volume of 300 µL), a 100 mM stock solution of guanidinium **Gua-Ald** in DMSO (3 µL, final concentration of 1 mM) and the 50 mM solution of the chosen triazine **TN**^{n} (6 µL, 1 eq.). After mixing with vortex and light centrifugation, the library was heated at 60°C for 10 minutes in a dry-bath, then allowed to slowly cool back to room temperature.

After a slight centrifugation, an aqueous 100 mM sodium cyanoborohydride solution (30 µL, 10 eq.) was added to the library which was then mixed again with vortex and abandoned to room temperature the necessary time for reductive amination to be completed (no more constituants with aldehyde visible in LC-MS in the corresponding control in solution, ~10 days).

The library was centrifuged and the supernatant was taken with a syringe and directly filtrated through a PVDF 0.45 µm syringe-filter in a LC-MS vial. The composition of the supernatant was then determined by LC-MS (method V on LC-MS 1 or method VIII on LC-MS 2).

Washing of the solid support

To the solid support was added a phosphate buffer 1 M solution (330 µL). After mixing with vortex, the mixture was heated at 60°C for 10 minutes and then immediately centrifuged. The washing solution was taken with a syringe and directly filtrated through a PVDF 0.45 µm

syringe-filter in a LC-MS vial for analysis (method V on LC-MS 1 or method VIII on LC-MS 2).

Control with micro-crystallized cellulose

Same procedure than for the assay with oligo(dT)-cellulose with replacement of the oligo(dT)-cellulose by the cellulose (23.0 mg). Washing procedure of solid support is the same.

Control in solution

In a microtube were successively added MilliQ water (volume needed to reach a final volume of 300 µL), a 100 mM stock solution of guanidinium **Gua-Ald** in DMSO (3 µL, final concentration of 1 mM) and the 50 mM stock solution of the chosen triazine **TN^n** (6 µL, 1 eq.). After mixing with vortex and light centrifugation, the library was heated at 60°C for 10 minutes in a dry-bath, then allowed to slowly cool back to room temperature.

After a slight centrifugation, an aqueous 100 mM sodium cyanoborohydride solution (30 µL, 10 eq.) was added to the library which was then mixed again with vortex and abandoned to room temperature the necessary time for reductive amination to be completed (no more constituants with aldehyde visible in LC-MS, ~10 days).

The library was then directly transferred in a LC-MS vial and its composition determined by LC-MS (method V on LC-MS 1 or method VIII on LC-MS 2).

*C. General method for library generation with guanidinium **Gua-Ald**, bisaldehyde triazine **TO** and commercial diamines*

Assay with oligo(dT)-cellulose

Oligo(dT)-cellulose (23.0 mg) was weighted in a microtube. To the solid were successively added MilliQ water (volume needed to reach a final volume of 300 µL), a 100 mM stock solution of guanidinium **Gua-Ald** in DMSO (3 µL, final concentration of 1 mM), a 100 mM stock solution of triazine **TO** in DMSO (3 µL, 1 eq.) and the 100 mM solution of the chosen monoamine (30 µL, 10 eq.). After mixing with vortex and light centrifugation, the library was heated at 60°C for 10 minutes in a dry-bath, then allowed to slowly cool back to room temperature.

After a slight centrifugation, an aqueous 100 mM sodium cyanoborohydride solution (30 µL, 10 eq.) was added to the library which was then mixed again with vortex and abandoned to room temperature the necessary time for reductive amination to be completed (no more constituants with aldehyde visible in LC-MS in the corresponding control in solution, ~10 days).

The library was centrifuged and the supernatant was taken with a syringe and directly filtrated through a PVDF 0.45 µm syringe-filter in a LC-MS vial. The composition of the supernatant was then determined by LC-MS (method IX on LC-MS 2).

Washing of the solid support

To the solid support was added a phosphate buffer 1 M solution (330 µL). After mixing with vortex, the mixture was heated at 60°C for 10 minutes and then immediately centrifuged. The washing solution was taken with a syringe and directly filtrated through a PVDF 0.45 µm syringe-filter in a LC-MS vial for analysis (method IX on LC-MS 2).

Control with micro-crystallized cellulose

Same procedure than for the assay with oligo(dT)-cellulose with replacement of the oligo(dT)-cellulose by the cellulose (23.0 mg). Washing procedure of solid support is the same.

Control in solution

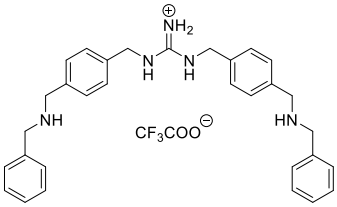
In a microtube were successively added MilliQ water (volume needed to reach a final volume of 300 µL), a 100 mM stock solution of guanidinium **Gua-Ald** in DMSO (3 µL, final concentration of 1 mM), a 100 mM stock solution of triazine **TO** in DMSO (3 µL, 1 eq.) and the 100 mM stock solution of the chosen diamine (30 µL, 10 eq.). After mixing with vortex and light centrifugation, the library was heated at 60°C for 10 minutes in a dry-bath, then allowed to slowly cool back to room temperature.

After a slight centrifugation, an aqueous 100 mM sodium cyanoborohydride solution (30 µL, 10 eq.) was added to the library which was then mixed again with vortex and abandoned to room temperature the necessary time for reductive amination to be completed (no more constituants with aldehyde visible in LC-MS, ~10 days).

The library was then directly transferred in an LC-MS vial and its composition determined by LC-MS (method IX on LC-MS 2).

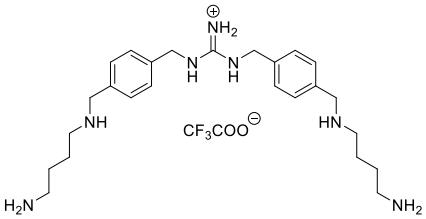
D. Isolation of libraries' constituents for MALDI-ToF analyses in presence of dT_{10}

Isolation of bisaminated product with benzylamine **Gua-Benz$_2$**



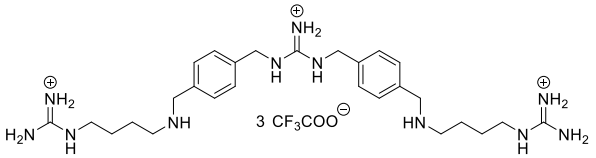
Guanidinium **Gua-Ald** (25.2 mg; 0.08 mmol) was dissolved in methanol (7.6 mL). Benzylamine (17 µL, 0.16 mmol) was added and the solution was stirred at room temperature for 2 hours. Sodium borohydride (6.2 mg, 0.16 mmol) was then added in the solution. After overnight stirring at room temperature, the reaction mixture was quenched by addition of aqueous 1 M hydrochloric acid (4 mL) and concentrated in vacuo. Residue was purified by semi-preparative HPLC (HPLC 1, Macherey-Nagel column VP Nucleodur C18 HTech 7µm 250 mm x 21 mM, linear gradient 5% B → 35% B in 30 minutes, flow 20 mL/min) and lyophilized to provide the product **Gua-Benz$_2$** as a white powder (35.0 mg, 78% yield). LC-MS (LC-MS 2): t_R = 2.41 minutes (method VII), extracted m/z calcd for $C_{31}H_{36}N_5^+$ ($[M]^+$): 478.30, found: 478.25; HR-ESI-MS(+): m/z calcd for $C_{31}H_{36}N_5^+$ ($[M]^+$): 478.2971, found: 478.2957.

Isolation of bisaminated product with 1,4-diaminobutane **Gua-C4$_2$**

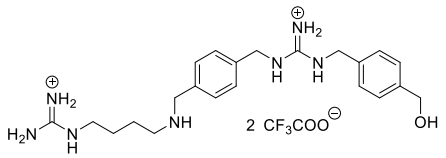


To a solution of 1,4-diaminobutane (16µL, 0.16 mmol) in methanol (4 mL) were added a solution of guanidinium **Gua-Ald** (26.4 mg, 0.08 mmol) in methanol (4 mL) and sodium borohydride (16 mg, 0.42 mmol). After overnight stirring at room temperature, the mixture was quenched by addition of aqueous 1 M hydrochloric acid (4 mL) and concentrated in vacuo. The white-yellowish residue was purified by semi-preparative HPLC (HPLC 1 Macherey-Nagel column VP Nucleodur C18 HTech 7 µm 250 mm x 21 mM, isochratic mode at 5% B for 10 minutes followed by linear gradient 5% B → 15% B in 20 minutes, flow 20 mL/min) and lyophilized to provide the product **Gua-C4$_2$** as a white powder (12.2 mg, 20% yield). LC-MS (LC-MS 2): t_R = 1.49 minutes (method VII), extracted m/z calcd for $C_{25}H_{47}N_7^+$ ($[M]^+$): 440.35, found: 440.25; HR-ESI-MS(+): m/z calcd for $C_{25}H_{42}N_7^+$ ($[M]^+$): 440.3502, found: 440.3499.

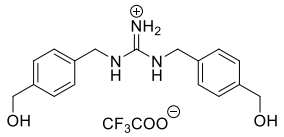
Isolation of bisreduced guanidinium **Gua(OH)$_2$** and aminated products **Gua-Agm(OH)** and **Gua-Agm$_2$** with agmatine



Gua-Agm$_2$



Gua-Agm(OH)



Gua(OH)$_2$

Agmatine sulphate salt (69.1 mg, 0.30 mmol) was dissolved in aqueous 1 M pH 6 phosphate buffer (2 mL). To this solution were successively added a solution of guanidinium **Gua-Ald** (10.5 mg, 0.03 mmol) in aqueous 1 M pH phosphate buffer (1.2 mL), and sodium cyanoborohydride (21.0 mg, 0.33 mmol). After stirring for 45 hours at room temperature, products were separated by semi-preparative HPLC (HPLC 1, Macherey-Nagel column VP Nucleodur C18 HTech 7 µm 250 mm x 21 mM, linear gradient 5% B → 20% B in 30 minutes, flow 20 mL/min) and lyophilized to provide products **Gua-Agm$_2$** (13.5 mg), **Gua-Agm(OH)** (9.9 mg) and **Gua(OH)$_2$** (7.3 mg) as white powders. **Gua-Agm$_2$**: LC-MS (LC-MS 2): t_R = 1.83 minutes (method VII), extracted m/z calcd for $C_{27}H_{46}N_{11}^+$ ([M-2H]$^+$): 524.39, found: 524.25; HR-ESI-MS(+): m/z calcd for $C_{27}H_{46}N_{11}^+$ ([M-2H]$^+$): 524.3938, found: 524.3943; **Gua-Agm(OH)**: LC-MS (LC-MS 2): t_R = 1.78 minutes (method VII), extracted m/z calcd for $C_{22}H_{34}N_7O^+$ ([M-H]$^+$): 412.28, found: 412.20; HR-ESI-MS(+): m/z calcd for $C_{22}H_{34}N_7O^+$ ([M-H]$^+$): 412.2825, found: 412.2825; **Gua(OH)$_2$**: LC-MS (LC-MS 2): t_R = 1.98 minutes (method VII), extracted m/z calcd for $C_{17}H_{22}N_3O_2^+$ ([M-H]$^+$): 300.17, found: 300.15; HR-ESI-MS(+): m/z calcd for $C_{17}H_{22}N_3O_2^+$ ([M]$^+$): 300.1712, found: 300.1711.

REFERENCES

(1) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997**, *62*, 7512–7515.

(2) Lovick, H. M.; Michael, F. E. *Tetrahedron Lett.* **2009**, *50*, 1016–1019.

(3) Jadhav, V. D.; Schmidtchen, F. P. *J. Org. Chem.* **2008**, *73*, 1077–1087.

(4) Sawoo, S.; Dutta, P.; Chakraborty, A.; Mukhopadhyay, R.; Bouloussa, O.; Sarkar, A. *Chem. Commun.* **2008**, 5957–5959.

(5) Casi, G.; Huguenin-Dezot, N.; Zuberbühler, K.; Scheuermann, J.; Neri, D. *J. Am. Chem. Soc.* **2012**, *134*, 5887–5892.

(6) Simpson, M. G.; Pittelkow, M.; Watson, S. P.; Sanders, J. K. M. *Org. Biomol. Chem.* **2010**, *8*, 1181–1187.

(7) Capaldi, D. C.; Gaus, H.; Krotz, A. H.; Arnold, J.; Carty, R. L.; Moore, M. N.; Scozzari, A. N.; Lowery, K.; Cole, D. L.; Ravikumar, V. T. *Org. Process Res. Dev.* **2003**, *7*, 832–838.

Résumé

Le présent travail de recherche vise à mieux connaître les interactions au sein d'auto-assemblages petites molécules – ADN dans le but de faciliter la conception de futurs ligands ou vecteurs de l'ADN. Pour se faire, des composés guanidiniums, sélectionnés pour leur capacité d'interaction avec les oxoanions – incluant les phosphodiesters de l'ADN – ont été bisfonctionnalisés pour sonder l'importance des interactions secondaires dans les auto-assemblages petites molécules - ADN. Deux approches différentes, rationnelle et combinatoire dynamique, ont été utilisées pour étudier les interactions au sein des auto-assemblages formés le long du squelette phosphodiester.
Avec l'approche rationnelle, des guanidiniums bisfonctionnalisés avec des groupements aromatiques de taille variable ont démontré leur capacité à s'auto-assembler avec l'ADN en solution aqueuse *via* la reconnaissance du squelette phosphodiester. Les expériences de compétition montrent l'importance des interactions d'empilement π-π dans la stabilisation de ces auto-assemblages.
Dans le cadre de l'approche combinatoire dynamique, un guanidinium comportant deux terminaisons aldéhyde a été utilisé pour générer, par amination réductrice, des bibliothèques combinatoires dynamiques basées sur la liaison imine. La génération de bibliothèques avec différentes amines en absence et en présence d'oligo(dT) sur cellulose met en évidence des effets de matrice caractérisés par des changements constitutionnels. Les résultats montrent ainsi que les constituants aux groupements cationiques et aromatiques sont sélectivement retenus contrairement aux constituants portant des substituants anioniques.
Les informations structurales obtenues par ces deux approches pourront à l'avenir être utilisées pour la conception de nouveaux ligands de l'ADN.
Parallèlement à ces études des interactions petites molécules – ADN, des polymères covalents dynamiques poly-acylhydrazones combinant des moitiés cationiques et de courts groupements d'éthylène glycol au sein de leur chaîne principale ont été conçus. Ces matériaux, capables de s'auto-assembler sous forme d'oligomères par polycondensation à haute concentration, complexent efficacement l'ADN double brin dans les médias biologiques et sont dégradables en milieu acide. L'association de ces deux caractéristiques les rend donc intéressants pour une application en tant que vecteurs « intelligents » pour la délivrance de gènes.

Mots clés
Auto-assemblage, oligonucléotides, reconnaissance moléculaire, guanidiniums bisfonctionnalisés, Chimie Combinatoire Dynamique, polymères covalents dynamiques.

Summary

The present research work aims to learn more about the interactions inside small-molecules – DNA self-assemblies in order to facilitate the conception of future DNA ligands or vectors. To achieve this goal, guanidinium compounds, selected for their recognized ability to interact with oxoanions – DNA phosphodiesters – were bisfunctionnalized to probe the importance of secondary interactions in the small-molecules – DNA self-assemblies. Two different approaches, rational and dynamic combinatorial, were used to study the interactions inside the self-assemblies formed along the phosphodiester backbone.
With the rational approach, guanidiniums bisfunctionnalized with aromatic groups of varying size demonstrated the ability to self-assemble with DNA in aqueous solution via phosphodiester recognition. The competition experiments show the importance of the π-π stacking interactions in the stabilisation of these self-assemblies.
Within the framework of the dynamic combinatorial approach, a guanidinium bearing two aldehyde terminations was used to generate, by reductive amination, dynamic combinatorial libraries based on the imine bond. The generation of libraries with different amines in absence and in presence of oligo(dT) on cellulose shows templating effects that are characterized by constitutional changes. The results thus show that constituents with cationic or aromatic groups are selectively retained on DNA while those featuring anionic groups are not.
The structural information gained by these two approaches could be used in the future to design new DNA ligands.
In parallel to these studies of the small-molecules – DNA interactions, dynamic covalent polymers combining cationic moieties and short ethylene glycol groups in their principal chain have been designed. These materials are able to self-assemble as oligomeres by polycondensation performed at high concentration and effectively complex double-stranded DNA in biological media and are degradable in acidic conditions. The association of these two characteristics make them interesting for an application as "smart" vectors for gene delivery.

Key words
Self-assembly, oligonucleotides, molecular recognition, bisfunctionnalized guanidiniums, Dynamic Combinatorial Chemistry, dynamic covalent polymers.