# Synthèse d'analogues d'aminoglycosides par voie chimique et ingénierie métabolique : Application à l'étude des ARN par RMN du fluor

Thomas Lombès

▶ **To cite this version:**

Thomas Lombès. Synthèse d'analogues d'aminoglycosides par voie chimique et ingénierie métabolique : Application à l'étude des ARN par RMN du fluor. Autre. Université René Descartes - Paris V, 2012. Français. NNT : 2012PA05T044 . tel-00787257

**HAL Id: tel-00787257**

**https://theses.hal.science/tel-00787257**

Submitted on 11 Feb 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

***UNIVERSITE PARIS DESCARTES***

ECOLE DOCTORALE FRONTIERES DU VIVANT

***Année 2012***

**DOCTORAT DE L'UNIVERSITE PARIS DESCARTES**

Spécialités : Chimie/Sciences du vivant

**Thomas LOMBES**

# Synthèse d'analogues d'aminoglycosides par voie chimique et ingénierie métabolique : Application à l'étude des ARN par RMN du fluor

Thèse dirigée par Frédéric DARDEL et Laurent MICOUIN

soutenue publiquement le 26 octobre 2012

Membres du jury :

| | | |
|---|---|---|
| M. le Docteur | Dominique FOURMY | Rapporteur |
| M. le Docteur | Boris VAUZEILLES | Rapporteur |
| M. le Professeur | Andreas KIRSCHNING | Examinateur |
| M. le Docteur | Jean-Luc PERNODET | Examinateur |
| M. le Professeur | Frédéric DARDEL | Directeur de thèse |
| M. le Docteur | Laurent MICOUIN | Directeur de thèse |

« …Tard, très tard, j'ai découvert la véritable nature de la science, de sa démarche, des hommes qui la produisent. J'ai compris que contrairement à ce que j'avais pu croire, le cheminement de la science ne constitue pas en une suite de conquêtes inéluctables ; qu'elle ne parcourt pas la voie royale de la raison humaine ; qu'elle n'est pas le résultat nécessaire, le produit inévitable d'observations sans appel imposées par l'expérimentation et le raisonnement. J'ai trouvé là un monde de jeu et d'imagination, de manies et d'idées fixes. A ma surprise, ceux qui atteignaient l'inattendu et inventaient le possible, ce n'était pas simplement des hommes de savoir et de méthode. C'était surtout des esprits insolites, des amateurs de difficultés, des êtres à vision saugrenue… »

**François Jacob**

*La statue intérieure*, 1987

# Remerciements

Cette thèse est l'aboutissement d'un travail interdisciplinaire effectué en collaboration entre deux équipes, l'équipe « Sélectivité et Diversité en Synthèse » et l'équipe « Structure des ARN, interactions et anti-infectieux » faisant respectivement partie des unités mixtes de recherche CNRS/Université Paris Descartes 8638 et 8015 de la faculté de pharmacie de Paris. Elle représente pour moi une période très enrichissante, durant laquelle j'ai appris beaucoup tant sur le plan humain que professionnel. J'y ai découvert des individus exceptionnels avec lesquels j'ai été heureux de pouvoir partager ces années. Je veux ici les remercier, même si aucun discours ne saurait être assez fort pour leur exprimer ma reconnaissance.

Je tiens tout d'abord à remercier les membres du jury pour avoir accepté de juger ce travail, les rapporteurs Boris Vauzeilles et Dominique Fourmy, ainsi que les examinateurs Andreas Kirschning et Jean-Luc Pernodet. Merci aussi à mes deux tuteurs de thèse, Matthieu Sollogoub et Ciaran Condon, qui ont su par leurs remarques et conseils faire avancer ce projet de recherche. Je souhaite également remercier chaleureusement la fondation Bettencourt-Schueller pour avoir financé ma thèse.

Je souhaite ensuite remercier mes deux directeurs de thèse, Laurent Micouin et Frédéric Dardel. Laurent tout d'abord, pour avoir guidé, dès mon M2, mes premiers pas dans le monde de la recherche. Tu m'as transmis ta passion pour la science, ta très grande ouverture d'esprit et tu as toujours été présent, dans les bons comme dans les mauvais moments. Laurent, tu es l'exemple du chercheur auquel j'aimerais un jour ressembler, j'espère sincèrement pouvoir un jour arriver à ta hauteur, scientifiquement et humainement. Frédéric, je te remercie pour m'avoir fait confiance et avoir initié, avec patience et enthousiasme, le chimiste que j'étais à la biochimie et à la biologie moléculaire. Je te remercie aussi pour l'autonomie que tu m'as laissée dans la conduite de notre projet ainsi que pour nos discussions éclectiques lors desquelles tu partageais avec moi tes connaissances encyclopédiques.

Vous m'avez fourni des conditions privilégiées pour effectuer ma thèse et m'avez énormément apporté. J'emporte avec moi cet héritage pour le faire vivre et peut-être un jour le transmettre à mon tour.

Un grand merci à Thomas Lecourt et à Serge Turcaud. Thomas, je te remercie énormément pour tous les moments que l'on a partagé, scientifiques et non scientifiques. J'ai adoré travailler avec toi et parler pendant des heures entre deux portes ou autour d'une bière. Serge, merci pour m'avoir accueilli dans ton antre à l'annexe, où ont toujours régné la chaleur humaine et une très bonne ambiance de travail. Merci de m'avoir fait bénéficier de ton expérience et de m'avoir enseigné les trucs et astuces pour faire des miracles avec une HPLC ou une réaction récalcitrante.

Merci également à Carine Tisné pour m'avoir permis de découvrir avec elle la biologie des ARNs et la RMN « biologique ». Je te remercie de m'avoir donné l'opportunité de travailler avec toi sur le sujet fluor. J'ai beaucoup apprécié nos discussions devant la console RMN ou autour d'un café, voire d'un rhum.

Un grand merci à nos collaborateurs sans qui aucun des travaux présentés n'auraient pu être menés à bout. Tout d'abord, merci du fond du cœur à Jean-Luc Pernodet et Sylvie Lautru pour m'avoir accueilli si chaleureusement au laboratoire et avoir été si ouverts et compétents. Merci infiniment pour les bons moments passés à Orsay, Tübingen ou Barcelone. Vous m'avez transmis le virus de la microbiologie des Streptomyces ! J'espère que notre route se recroisera très souvent. Merci à Florence Lorieux et Audrey Vingadassalon pour m'avoir initié aux techniques de modification génétique de ces organismes si particuliers. Un grand merci à Anne Regazzetti pour son aide précieuse en HPLC-MS et pour son enthousiasme communicatif. Merci d'avoir consacré tellement de temps et d'énergie au projet mutasynthèse malgré les difficultés techniques et les conditions parfois difficiles.

Je souhaite ensuite remercier toutes les autres personnes du laboratoire qui ont contribuées à la réalisation de ce travail. Merci beaucoup à Elise Prost pour son expertise en RMN. Analyser des spectres de sucres sans toi n'a pas la même saveur… Un très grand merci à Bili Seijo pour m'avoir appris à faire des clonages et à purifier des protéines, bref à survivre dans un laboratoire de biochimie. Merci à Valery Larue pour m'avoir appris à attribuer les signaux d'un ARN en RMN 2D, à Marjorie Catala pour les purifications d'ARN, à Pascale Leproux pour les masses, et à Franck Brachet pour l'aide avec le robot de cristallisation des protéines. Merci également à Cathy Bonamour du Tartre et à Nicole Mazure, qui m'ont guidé dans les méandres de l'administration universitaire.

J'adresse également mes remerciements à l'ensemble des membres présents et passés des UMR 8638 et 8015 et plus particulièrement, Aurélie (bonne chance pour la suite !), Aurélien, Caroline (je vais regretter nos pauses à l'annexe), Camille, Christelle, Eric, Farouk, Gauthier, Jordi, Ludovic, Marie, Marlène, Mathilde, Mélissa, Morgane, Mounira, Olivier, Paola, Roba, Yuhan, et mes compères de bureau en biologie, Dona et Ehsan.

Je souhaite aussi remercier l'école doctorale Frontières du vivant, qui s'est révélée être un formidable lieu de vie, riche en rencontres et en échanges. Merci aux membres des équipes IGEM 2010 et 2011 et aux membres des comités d'organisation des « Paris Interdisciplinary PhD Symposium » 2009 et 2011. J'y ai vécu des moments très forts et enrichissants. De façon plus générale, je veux exprimer ici ma reconnaissance à l'ensemble des professeurs et des mentors que j'ai eu au cours de mes études à l'ENS ou ailleurs. Ils ont su me guider et me faire découvrir quelques-uns des mystères passionnants du monde qui m'entoure.

Enfin, je tiens à remercier mes amis et ma famille, que j'aime tendrement et sans qui je ne serai pas celui que je suis, avec une mention spéciale pour Anne et Marc, mes parents exceptionnels (mieux que des pélicans), ma grand-mère Pierrette, dont la curiosité n'a d'égale que son ouverture d'esprit, et mes deux petites sœurs, Lucie et Amélie, qui sont souvent plus grandes que moi.

Déborah, merci pour le chemin parcouru ensemble et l'horizon qui nous appartient.

Merci tout simplement.

# Table des matières

**Remerciements** **5**

**INTRODUCTION GENERALE** **15**

**I. CONTEXTE SCIENTIFIQUE** **19**

**I - Les ARN comme cible thérapeutique** **20**
A - Caractéristiques chimiques des ARN 20
B - Structure des ARN 21
C - Rôle Biologique des ARN 25
1) Le dogme central 25
2) ARNm et riboswitches 28
3) ARNr et ribozymes 30
4) ARNt, du décodage du code génétique à l'initiation de la rétrotranscription 34
5) Interférence ARN 36
6) Conclusion : les ARN, cibles thérapeutiques de première importance. 37
D - Les difficultés à cibler les ARN 38

**II - Les aminoglycosides, « ligands universels » d'ARN** **42**
A - Historique de la découverte et utilisation médicale 42
B - Caractéristiques physicochimiques des aminoglycosides 44
C - Interaction avec les ARN : considérations générales 46
D - Activité antibiotique des aminoglycosides 48
1) Interaction avec le ribosome 48
2) Résistance aux aminoglycosides 54
3) Conclusion 55
E - Autres cibles biologiques des aminoglycosides 56
1) Traitement de désordres génétiques 56
a) Translecture 56
b) Motifs d'ARN répétés 58
2) Les aminoglycosides comme molécules antivirales 59
F - Conclusion : l'intérêt de développer des analogues d'aminoglycosides 61

**III - Stratégies de recherche d'analogues d'aminoglycosides** **63**
A - Exploration de la biodiversité microbienne 63
1) Nouvelles méthodes d'isolement et de fermentation de souches bactériennes 63
2) Exploitation des clusters biosynthétiques cryptiques 66

3) Approche métagénomique 67
4) Approche de biosynthèse combinatoire 68
5) Conclusion 70
B - Stratégies de synthèse organique d'analogues d'aminoglycosides 71
1) Stratégies de synthèse totale et d'hémisynthèse 71
a) Aménagement fonctionnel des aminoglycosides 72
b) Changement du nombre de cycle et de leur géométrie globale des aminoglycosides 79
c) Analogues contraints 79
d) Aminoglycosides conjugués 81
2) Approche par fragment et synthèse combinatoire 85
3) Approche chemoenzymatique 91
4) Conclusion 91
C - Une stratégie hydride : la mutasynthèse 92
1) Concept de la mutasynthèse 92
2) Exemples de stratégies de mutasynthése 95
D - Conclusion 98

**IV - Projet de thèse 100**

**II. EXPRESSION HETEROLOGUE D'ENZYMES DE BIOSYNTHESE D'AMINOGLYCOSIDES 103**

**I - Biosynthèse de la néomycine et d'aminoglycosides apparentés 105**

**II - Résultats personnels 109**
A - Stratégies pour la surexpression hétérologue et la purification de protéines recombinantes 109
B - Enzymes de biosynthèse de la néomycine chez S.fradiae : Neo8, Neo11, Neo16, Neo18 113
1) Clonage et essais de surexpression 113
2) Purification de Neo8 sous forme dénaturée et expériences de repliement 116
3) Test de l'activité biologique de Neo8 117
4) Développement d'un anticorps anti-Neo8 119
C - La glycosyltransférase GenM1 de la biosynthèse de la gentamicine chez M. echinospora 123
1) Clonage et tentatives de surexpression 123
2) Protéines chimèriques 124
a) GST-GenM1 124
b) NusA-GenM1 126
c) AAC(6')-Ib-GenM1 128
3) Ingénierie de la protéine GenM1 : conception de GenM1* 130
D - Enzymes de biosynthèse de la butirosine chez B.circulans :BtrM et BtrD 134

**III - Conclusion 136**

**III. STRATEGIE DE MUTASYNTHESE D'ANALOGUES DE NEOMYCINE 139**

**I - Introduction 141**
A - Les travaux fondateurs 141
B - Contexte scientifique actuel 148
C - Objectifs de l'étude 151

**II - Résultats personnels 153**
A - Construction du mutant Δneo6 153
1) Stratégie de clonage du plasmide pFL02-neoApra 154
2) Suppression du gène neo6 par double recombinaison homologue 157
B - Purification et quantification de la production d'aminoglycosides 160
1) Purification par échangeuse d'ions (SPE) 160
2) Analyse des échantillons par HPLC-MS/MS 161
C - Optimisation de la production de néomycine et définition des conditions de mutasynthèse 164
1) Criblage de différents milieux de culture 164
2) Cinétique de production de la néomycine 165
D - Synthèse des mutasynthons 167
1) Métabolites « naturels » 167
a) Synthèse de la désoxystreptamine 167
b) Synthèse de la paromamine et de la néamine 168
2) Métabolites « non-naturels » 169
a) Synthèse de la streptidine et de la streptamine 169
b) Synthèse du *cis*-1,3-diaminocyclopent-4,5-diol C5 170
c) Synthèse de la paromamine avec un azoture en position 6' (paro$N_3$) 171
d) Synthèse de la paromamine fluorée en position 6' (paroF) 172
E - Mutasynthèse 173
1) Complémentation chimique avec les métabolites naturels de la voie de biosynthèse 173
a) Vérification du phénotype des mutants DSM 41550 et Δneo6 par HPLC-MS 173
b) Complémentation avec la désoxystreptamine (DOS) et étude de l'influence de la concentration en DOS sur la régulation de la biosynthèse chez *S.fradiae* 175
c) Complémentation avec la néamine et la ribostamycine 180
d) Complémentation avec la paromamine 181
2) Complémentation avec des mutasynthons non naturels 183
a) Complémentation avec des analogues de DOS 183
b) Complémentation avec des analogues de métabolites plus complexes 188

**III - Conclusion et perspectives 191**

**IV. DEVELOPPEMENT DE SONDES FLUOREES POUR L'ETUDE DES ARN EN RMN $^{19}F$ 193**

**I - La RMN du fluor dans l'étude des ARN 195**

A - L'intérêt de la RMN : donner une information dynamique à l'échelle atomique 195

B - La RMN $^{19}F$ comme outil d'étude des biomolécules 197

1) Les avantages de la RMN $^{19}F$ 197

2) Application de la RMN $^{19}F$ à l'étude des ARN 199

a) Stratégie de marquage de la cible 199

b) Les inconvénients du marquage de la cible 203

c) Utilisation d'une sonde fluorée externe 203

**II - Résultats antérieurs du laboratoire 205**

A - Contexte général : synthèse de ligands d'ARN par fragment 205

B - Le DACPF comme sonde dans l'étude des ARN par RMN $^{19}F$ 207

1) Synthèse du DACPF 207

2) Interaction avec les ARN : le cas de l'ARNt$^{Lys}_3$ 210

3) Le DACPF comme sonde topologique des ARN 214

4) Suivi de changements conformationnels: exemple de la dénaturation d'une tige boucle 215

**III - Résultats personnels 217**

A - Suivi de changements conformationnels: capture conformationnelle d'un aptamère par son ligand 217

B - Evaluation de l'affinité de ligands vis-à-vis d'une cible ARN : exemple de l'ARN 16S23 222

1) La cible : l'ARN 16S23, analogue du site de décodage du ribosome bactérien 222

2) Principe de la mesure d'affinité : compétition avec la sonde fluorée 224

3) Utilisation du DACPF comme ligand compétitif 225

a) Etude de l'interaction du DACPF avec l'ARN 16S23 225

b) Résultats des compétitions avec le DACPF 230

c) Interprétation critique des résultats 234

4) Développement d'une sonde fluorée plus sensible :la paromamine fluorée 235

a) Synthèse de la paromamine fluorée (paroF) 235

b) Etude de l'interaction de paroF avec l'ARN 16S23 et optimisation des conditions expérimentales pour les expériences de compétition. 238

c) Résultats de compétitions avec la paroF 240

**IV - Conclusion 244**

**V - Perspectives 246**

A - Criblage en RMN $^{19}F$ de nouveaux ligands d'ARN 246

B - Etude en RMN $^{19}F$ de la dynamique de systèmes biologiques 246

C - Paromamine fluorée et aptamère, vers un système artificiel de régulation traductionnelle ? 248
D - Développement de nouvelles sondes fluorées : analogues fluorés de la kynuramine 250

**CONCLUSION GENERALE 253**

**V. PARTIE EXPERIMENTALE 257**

**I - Partie expérimentale du chapitre III : Mutasynthèse 258**
A - Production de spores de S.fradiae 258
B - Extraction d'ADN total de Streptomyces : technique Fast-prep 259
C - Conditions PCR 260
1) Amplification du fragment neo5-neo7 260
2) Amplification du gène neo6 260
3) PCR de vérification des clones lors de la double recombinaison 261
D - Conjugaison interspécifique Streptomyces/E.coli 261
E - Test phénotypique de production de néomycine sur milieu solide 263
F - Production de néomycine par fermentation de S.fradiae 263
1) Production de néomycine à petite échelle (fernbach) 263
2) Production de néomycine à grande échelle (fermenteur) 264
3) Expériences de mutasynthèse 264
4) Différents milieux testés 264
G - Purification d'aminoglycosides par colonne échangeuse d'ions (SPE) 267
H - Analyse des analogues d'aminoglycosides purifiés par HPLC-MS/MS 267
1) Chromatographie liquide haute performance (HPLC) 267
2) Spectrométrie de masse (MS) 268
3) Spectres de références 269

**II - Partie expérimentale du chapitre IV : $^{19}$F RMN 275**
A - Préparation des ARN 275
B - Généralités sur les expériences RMN 277
C - Caractérisation des interactions entre les sondes fluorées et leurs cibles ribonucléiques 278
1) Mesure de l'affinité de la sonde fluorée DACPF pour l'ARN 16S23 278
2) Mesure de l'affinité de la sonde fluorée DACPF pour l'aptamère à la néomycine 279
3) Mesure de l'affinité de la sonde fluorée ParoF pour l'ARN 16S23 280
4) Analyse de l'interaction DACPF/ARN 16S23 282
D - Suivi de la capture conformationnelle de l'aptamère par la néomycine 283
E - Expériences de compétition sonde fluorée/ligand/ARN 16S23 ou 16S27 283
1) Conditions expérimentales 283
2) Fit des expériences de compétition 283

3) Spectres RMN des différentes expériences de compétition 286

**III - Partie expérimentale : Synthèse organique 296**

A - Généralités 296

B - Synthèse d'aminoglycosides naturels par dégradation 297

C - Synthèse de la paromamine avec un fluor en 6' paroF 303

D - Synthèse de la paromamine avec un azoture en 6' paroN3 316

E - Synthèse du 1,3-diamino-cyclopentan-4,5-diol C5 321

F - Synthèse du 1,3-diamino-4-fluoro-cyclopentane DACPF 324

G - Synthèse de β aminocétones par la réaction de Mannich 327

**VI. BIBLIOGRAPHIE 332**

**VII. ANNEXES 367**

**I - Résumé de la thèse 368**

**II - Production scientifique 369**

A - Communications orales 369

B - Posters 370

C - Organisation de congrès 370

D - Enseignements 370

E - Publications 371

1) Liste des publications 371

2) Articles en pdf 372

a) Publication 1 372

b) Publication 2 376

c) Publication 3 382

d) Publication 4 387

F - Compétition IGEM 392

1) Projet IGEM Paris 2010 « Every Bacteria Counts » 393

2) Projet IGEM Paris 2011 « Tube or not Tube ? Towards harnessing bacterial nanotubes by and for synthetic biology » 396

**III - Liste des abréviations 397**

**IV - Liste des figures 400**

# Introduction Générale

Les ARN constituent des cibles thérapeutiques extrêmement intéressantes bien qu'encore assez peu exploitées. Outre leur rôle bien connu de support temporaire de l'information génétique, on découvre aujourd'hui leur implication majeure dans de nombreux phénomènes biologiques de première importance : régulation transcriptionelle ou traductionelle, cascades de signalisation cellulaire, catalyse de réactions chimiques, …

De façon cruciale, la propriété des ARN à former des structures tridimensionnelles capables de reconnaitre sélectivement divers partenaires moléculaires permet d'envisager de venir interférer avec les mécanismes complexes dans lesquels ils jouent un rôle clé. La recherche de nouveaux ligands d'ARN est ainsi en plein essor même si les obstacles pour une interaction spécifique avec ces cibles non traditionnelles, polyanioniques et très flexibles, sont encore loin d'être complètement levés. Ces ligands constituent en effet des outils scientifiques extrêmement intéressants pour l'étude du rôle biologique des ARN et pour l'ouverture potentielle de nouvelles voies dans le traitement de maladies infectieuses ou génétiques.

Notre compréhension parcellaire des interactions des petites molécules avec leurs partenaires ARN rend pour le moment difficile une conception rationnelle de ligands d'ARN. Nous pouvons néanmoins nous appuyer sur une famille de produits naturels, les aminoglycosides, souvent décrits comme des « ligands universels » d'ARN pour concevoir de nouveaux ligands. L'architecture de ces produits naturels polysaccharidiques, utilisés depuis plus de 60 ans pour leurs propriétés antibiotiques, constitue en effet un excellent point de départ pour la conception de nouveaux ligands d'ARN.

Le travail exposé ici s'inscrit dans le cadre global de la recherche de nouveaux ligands d'ARN, travaux effectués en collaboration entre les deux unités de recherche dans lesquelles je me suis intégré à la Faculté de Pharmacie. L'objectif du sujet de thèse qui m'a été confié était de contribuer à développer une nouvelle stratégie synthétique pour la conception d'analogues d'aminoglycosides, la mutasynthèse. Nous nous sommes plus particulièrement intéressés à la néomycine et à la voie de biosynthèse de cet aminoglycoside dans la bactérie *Streptomyces fradiae* du groupe des actinomycètes. En s'inspirant de la structure de cette dernière, nous avons par ailleurs synthétisé certains analogues d'aminoglycosides fluorés que nous avons utilisés comme outils d'étude des ARN en RMN du fluor.

L'introduction de ce manuscrit rappelle les rôles biologiques clés de l'ARN ainsi que leurs potentiels comme cible thérapeutique en dépit des difficultés à cibler ces molécules complexes. Les aminoglycosides sont ensuite abordés en tant que ligands de l'ARN et nous nous intéresserons à leurs caractéristiques physico-chimiques, à leur interaction avec les ARN et à leur mode d'action

biologique. Nous exposerons ensuite les différentes stratégies qui ont été mises en place pour créer des analogues d'aminoglycosides afin de tirer parti des architectures de ces molécules optimisées pour la fixation à des cibles nucléiques. Ceci nous permettra d'introduire les objectifs de ce travail de thèse.

Les résultats de ma thèse sont exposés en trois parties. La première est consacrée à l'étude de la voie de biosynthèse de la néomycine chez *S.fradiae* et à nos tentatives d'expression hétérologue chez *E.coli* de certaines des enzymes clés de cette voie de biosynthèse et de voies de biosynthèse apparentées.

La deuxième partie aborde plus spécifiquement la stratégie de mutasynthèse pour la création de nouveaux analogues de néomycine. Après un bref survol des travaux qui ont initié nos recherches, nous nous intéresserons aux progrès accomplis vers la mise en place d'une telle stratégie dans des bactéries *S.fradiae* génétiquement modifiées.

La troisième partie s'intéresse à l'utilisation d'analogues fluorés d'aminoglycosides comme outils d'étude des ARN par RMN du fluor. Après avoir rappelé l'intérêt de la RMN du fluor dans l'étude des ARN et les travaux préliminaires effectués au laboratoire, nous montrerons comment ces sondes fluorées permettent non seulement d'avoir accès à la structure et la dynamique de différents ARN mais aussi de mesurer l'affinité de certains de leur ligands.

# I. Contexte scientifique

# I - Les ARN comme cible thérapeutique

## *A - Caractéristiques chimiques des ARN*

L'ARN, ou acide ribonucléique, est un polymère non ramifié de nucléotides. Chaque nucléotide est constitué d'un sucre, le ribose, sur lequel est greffé une des quatre bases azotées canoniques : l'adénine (A), la guanine (G), la cytosine (C) ou l'uracile (U) (figure I-1).

Les riboses sont liés entre eux par des ponts phosphodiesters qui mettent en jeu l'oxygène 3' du ribose du nucléotide *n* et le phosphate en 5' du nucléotide suivant *n + 1*. Au pH physiologique, les phosphates sont chargés négativement, ce qui fait de l'ARN une molécule fortement polyanionique. L'ARN est chimiquement très proche de l'acide déoxyribonucléique (ADN), support de l'information génétique, où le sucre est un déoxyribose et les bases azotées sont identiques à l'exception de la thymine qui remplace l'uracile.

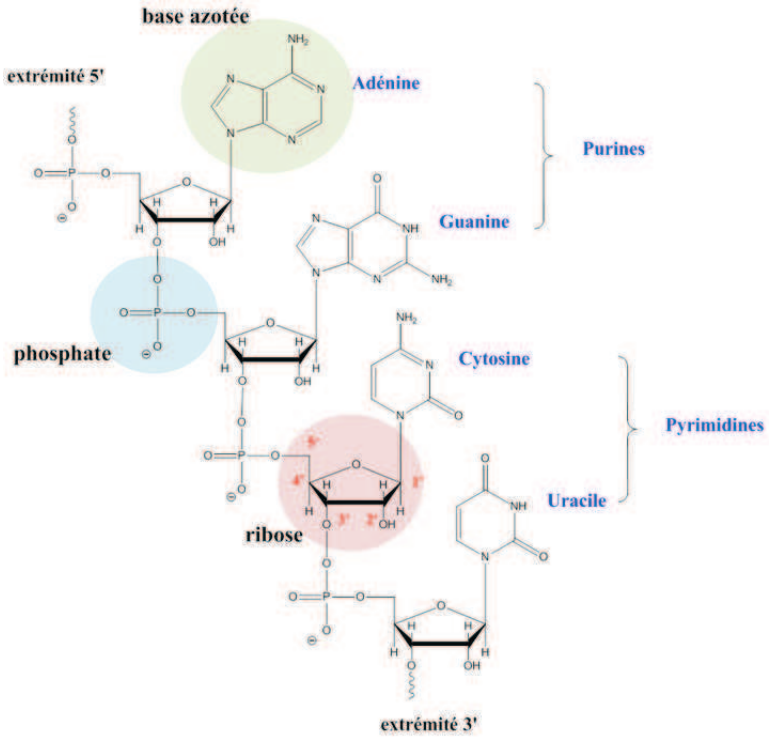



**Figure I-1 Structure primaire des ARN**

Dans certains cas, il peut y avoir modification post-transcriptionelle des nucléotides de l'ARN. Ces modifications sont très variées, allant d'une simple méthylation du ribose à une modification profonde de la base azotée, et peuvent donner plus d'une centaine de nucléotides différents[1,2,3]. On trouve par exemple très fréquemment dans les ARN de transfert des nucléotides pseudouridines qui

résultent de l'isomérisation de l'uracile par des pseudouridine synthases pour former un C-glycoside avec le ribose (figure I-2).

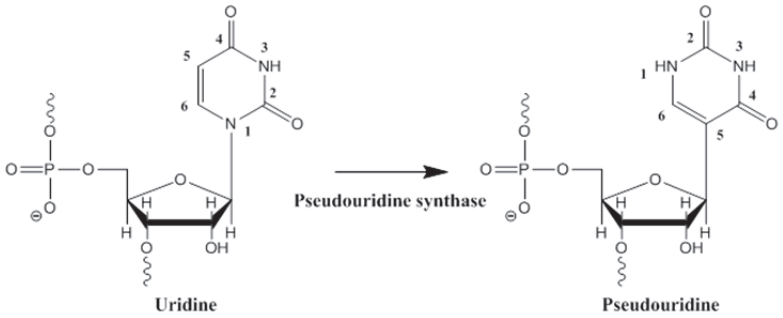



**Figure I-2 Un exemple de modification post transcriptionelle des ARN : synthèse de la pseudouridine**

Contrairement à l'ADN, les ARN sont généralement des polymères relativement courts, d'une taille allant de quelques dizaines à quelques milliers de nucléotides. Ils sont par ailleurs beaucoup plus flexibles et sensibles à l'hydrolyse.

## *B - Structure des ARN*

La structure des ARN présente, comme dans le cas des protéines, trois niveaux d'organisation hiérarchique : structure primaire, structure secondaire et structure tertiaire (figure I-3). Ces niveaux correspondent aux différentes étapes du repliement séquentiel des ARN[4].

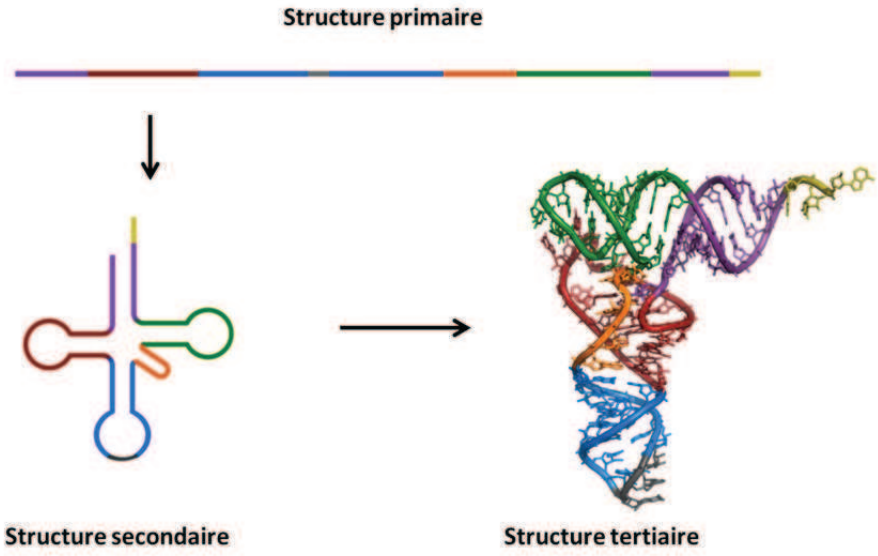



**Figure I-3 Niveaux d'organisation de la structure des ARN**
(adapté de la structure de l'ARNt$^{Phe}$ de la levure[5] ; Protein Data Bank, entrée 1EHZ)

L'enchainement des nucléotides constitue ce qu'on appelle la structure primaire, ou séquence. Elle est directement dérivée de la séquence d'ADN codant, après éventuelles modifications post-transcriptionelles (modification des bases azotées, épissage, …). Cette séquence primaire va dicter le type de structure secondaire qui peut se former à travers des liaisons hydrogène, via des appariements de type Watson-Crick entre bases azotées complémentaires ou bien suite à des interactions non-canoniques entre nucléotides (par exemple appariement bancal (wobble), appariement de Hoogsteen ou interaction base azotée/ribose). On dénombre de nombreux types d'interactions différentes[6], ce qui laisse entrevoir la complexité des structures potentiellement formées (figure I-4).

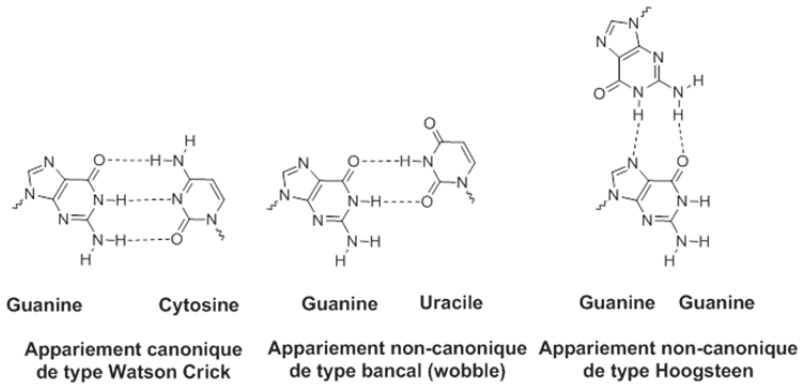



**Figure I-4 Exemples d'appariement des bases de l'ARN**

Comme dans le cas des protéines, les structures secondaires peuvent interagir entre elles, souvent à travers des interactions non-canoniques, stabilisant des repliements compacts pour former des structures tridimensionnelles[7,8]. En raison de la nature très polyanionique de l'ARN, les interactions électrostatiques avec le solvant et les molécules chargées en solution (comme par exemple les ions magnésium divalent $Mg^{2+}$) jouent un rôle prédominant dans la stabilisation de ces structures[9].

Comme son cousin chimique l'ADN, l'ARN double-brin peut former des hélices où les bases azotées sont appariées à travers des interactions de type Watson-Crick. La structure de ces hélices diffère cependant légèrement de celles formées par l'ADN (figure I-5). En effet, si dans l'ADN, le désoxyribose adopte préférentiellement une conformation C2'-endo, le cycle furanose du ribose de l'ARN a lui tendance à adopter une conformation C3'-endo. Cette conformation a pour conséquence un léger élargissement de la double hélice, un décalage de son axe par rapport au plan formé par les bases azotées ainsi qu'un changement du pas de l'hélice. Ceci induit d'autre part une modification de la géométrie des sillons de l'hélice, avec un grand sillon plus profond et étroit que dans le cas de

l'ADN et un petit sillon moins marqué[10]. Ces caractéristiques structurales des doubles hélices d'ARN induisent une forte contrainte stérique et ont tendance à défavoriser les éventuelles interactions de partenaires protéiques ou moléculaires avec le grand sillon d'une double hélice d'ARN.

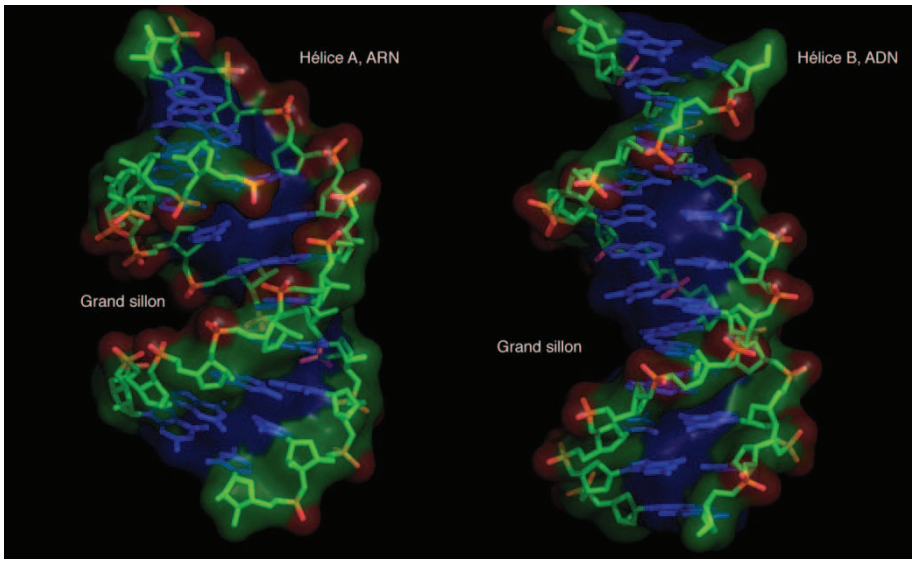



**Figure I-5 Comparaison structurale d'une hélice d'ARN avec une hélice d'ADN**
(source : fdardel ; Protein Data Bank, entrées 1BNA et 1QCU)

D'un point de vue biologique, les ARN ont la particularité d'être la plupart du temps sous forme monocaténaire. Le brin d'ARN peut se replier sur lui-même pour former des structures secondaires, séquestrant certaines séquences en les engageant dans des appariements tout en exposant d'autres séquences nucléotidiques, facilitant leur interaction avec d'autres molécules. Comme illustré dans la figure I-6, on trouve de nombreux motifs secondaires des ARN comme par exemple des tiges boucles, des hernies, des boucles internes, … Les structures secondaires ainsi formées peuvent ensuite conduire à la formation de motifs ou de modules structuraux caractéristiques[11]. Les structures d'ARN peuvent être vues comme des « poupées russes », résultant de combinaisons et d'enchevêtrements de motifs structuraux simples[12].

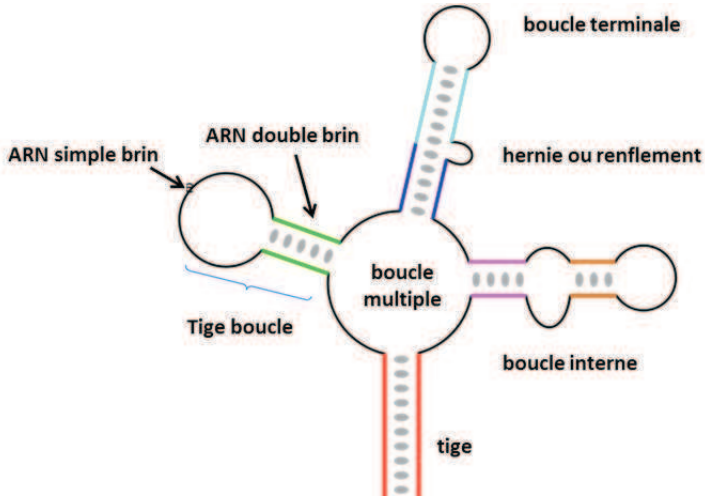



**Figure I-6 Variété des motifs structuraux secondaires de l'ARN**
(adaptée de : wikipédia, article ARN, fdardel)

La combinaison de motifs structuraux donne lieu à des structures tridimensionnelles complexes avec alternance de régions appariées et de régions simple-brin, qui disposent de boucles et de poches susceptibles d'interagir de façon affine et spécifique avec de nombreux partenaires nucléiques, protéiques ou même moléculaires. Cette propriété des ARN a pendant longtemps été ignorée par la communauté scientifique, persuadée, pour reprendre les mots de Jacob et Monod, que « la capacité à former des complexes stéréospécifiques avec des petites molécules semble être un privilège des protéines »[13]. Cette hypothèse était jugée d'autant plus vraisemblable du fait du caractère très polyanionique des ARN et de leur faible diversité chimique par comparaison aux protéines (4 bases azotées possibles seulement contre 20 acides aminés différents). On sait aujourd'hui que les ARN, très proches chimiquement de l'ADN, ont des propriétés structurales paradoxalement assez proches de celles des protéines. Ce changement de paradigme a bouleversé notre compréhension de l'importance biologique des ARN, qui se sont révélés avoir un grand nombre de rôles clé comme nous le verrons par la suite dans la partie I-II.

Au niveau énergétique, il est intéressant de constater que la majorité de l'énergie libre de repliement des ARN est due aux structures secondaires, chaque paire de nucléotides correspondant en moyenne à une énergie d'interaction de -1 à -5 kcal/mol. La structure tertiaire quant à elle ne représente qu'une faible fraction de l'énergie totale de repliement qui, bien qu'elle reste l'objet de controverse, peut néanmoins être évaluée à moins d'une dizaine de kcal/mol seulement[14]. La formation de la structure tertiaire ne constitue ainsi qu'une perturbation des structures secondaires, qui existent et sont stables par elles-mêmes. D'un point de vue biologique, cela signifie qu'il est possible de modifier complétement la structure tertiaire d'un ARN avec un ligand d'une affinité

nanomolaire ou même micromolaire dans les cas favorables. Ces grandeurs thermodynamiques permettent d'expliquer la flexibilité des ARN et leur propension à être en équilibre entre plusieurs conformations, propriété souvent au cœur de leur activité biologique.

## *C - Rôle Biologique des ARN*

### 1) Le dogme central

Lorsqu'en 1944, Avery démontre que l'information génétique est portée par l'ADN[15], la grande question de la biologie moléculaire devient comment passer de cette séquence linéaire d'acide desoxyribonucléique, dont on obtient la structure en 1953[16], à l'ensemble des composés cellulaires nécessaires à la vie. En 1958, Francis Crick publie ce qui allait devenir le « dogme central de la biologie moléculaire ». Il identifie les ARN comme des intermédiaires clés entre l'information génétique portée par l'ADN, et les protéines, responsables notamment des activités enzymatiques nécessaires à la vie[17,18]. Ces travaux donnent lieu à la découverte du processus de transcription, durant lequel l'information portée par l'ADN peut être copiée sur les ARN messagers (ARNm). Ceux-ci vont ensuite pouvoir être décodés par les ARN de transferts (ARNt) lors du processus de traduction pour former les protéines. La traduction est un phénomène hautement contrôlé par une machinerie complexe, le ribosome, constituée par les ARN ribosomiques (ARNr) et par des protéines[19](voir figure I-7). Jusqu'à une période récente, notre compréhension des ARN a été marquée par cet héritage prégnant qui malheureusement, s'il est simplifié, peut conduire à minimiser leur rôle biologique en les considérant uniquement comme des supports temporaires de l'information génétique.

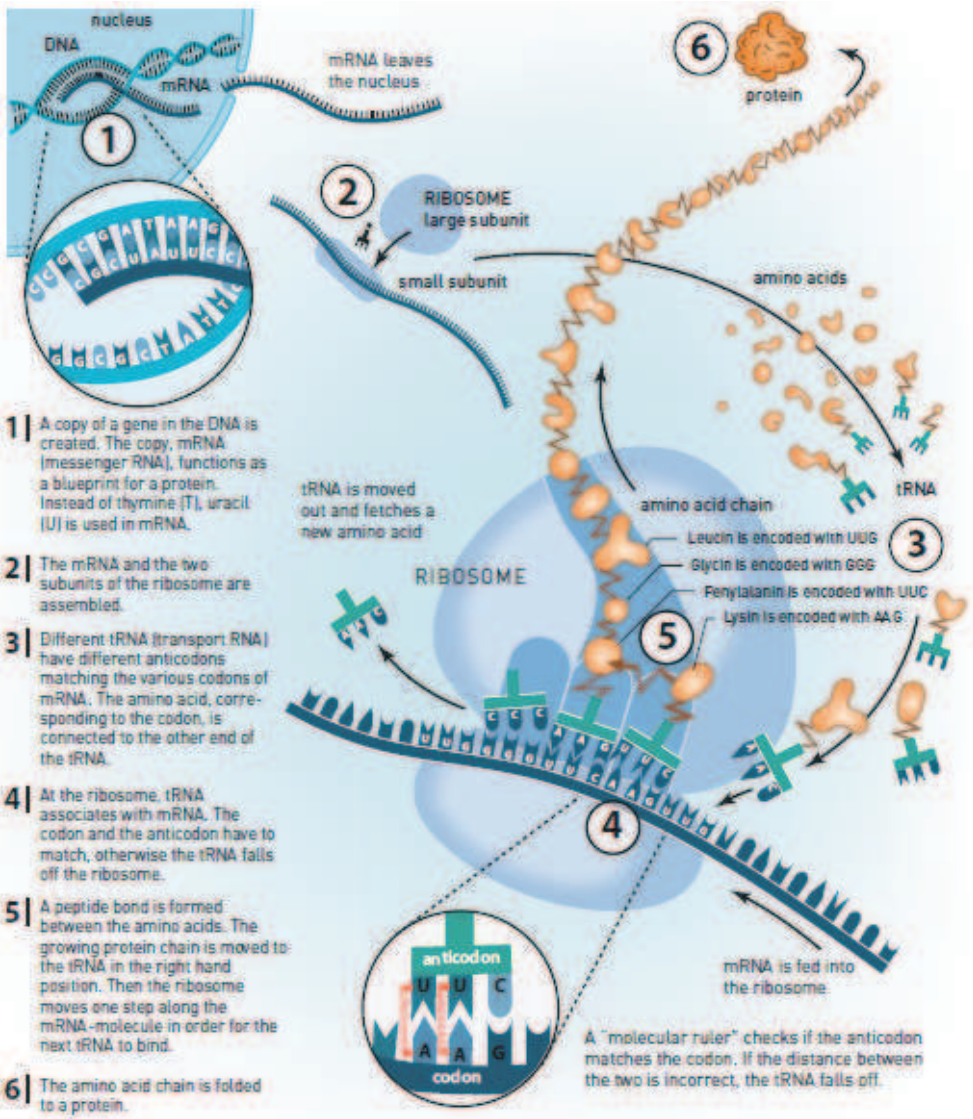



**Figure I-7 Représentation schématique de la transcription et de la traduction**
(source : Ehrenberg, M. Scientific Background on the Nobel Prize in Chemistry 2009. Structure and Function of the Ribosome)

Par exemple, une idée fausse généralement tirée du dogme central de la biologie moléculaire est que, comme il y a transfert d'information de l'ADN aux protéines via les ARNm, l'information génétique contenue dans l'ADN se réduit à des séquences codant pour des protéines. Si cette simplification reste à peu près vraie pour les organismes unicellulaires modèles comme *E.coli* où l'essentiel du génome est traduit, elle montre vite ses limites lorsqu'on cherche à étudier des organismes d'une complexité plus importante. Il a en effet été montré que le ratio séquence non-codante/séquence codante était une fonction croissante de la complexité des organismes étudiés. Ainsi, chez les mammifères, seule une faible fraction du génome est effectivement traduite[20] (figure I-8).

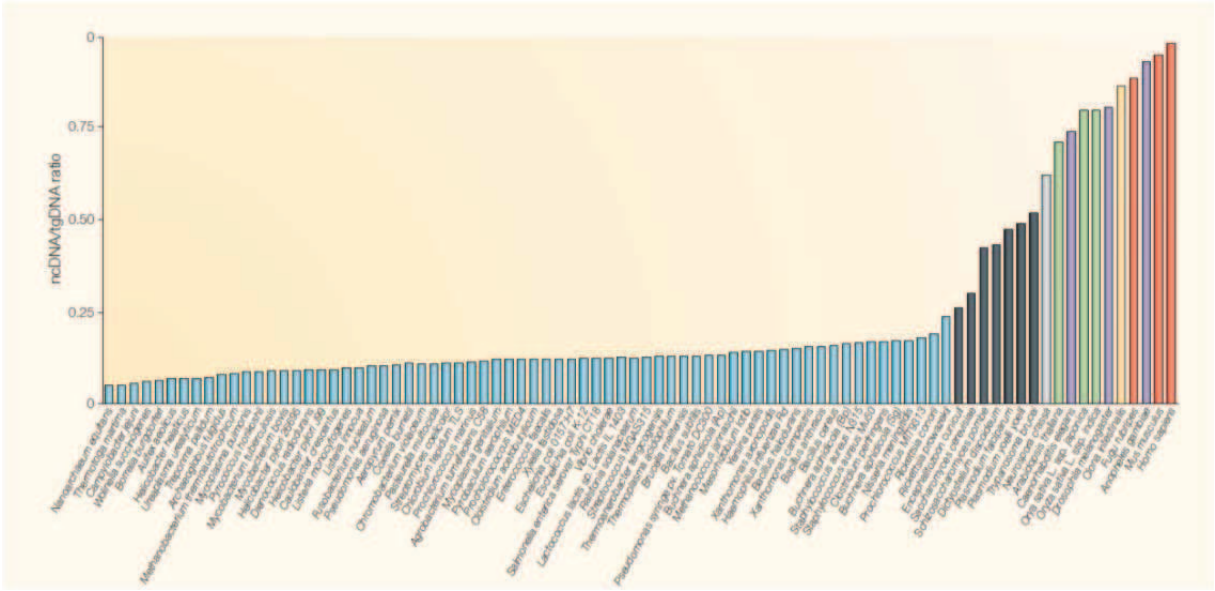



**Figure I-8 Le ratio séquences non-codantes/séquences codantes augmente en fonction de la complexité développementale apparente des organismes**[20]

Par exemple, chez l'homme, seul 2 % de l'ADN total est traduit en protéines. Ce fait surprenant a tout d'abord conduit une partie de la communauté scientifique à considérer les 98 % du génome restant comme du « junk DNA », débris d'ADN inutiles, simple reflet de l'histoire évolutive. Cette hypothèse a été un temps confortée lorsqu'on s'est aperçu que certains de ces fragments n'étaient pas conservés, ne conférant donc pas un avantage évolutif, et qu'on pouvait en supprimer une partie sans impact apparent sur le phénotype d'un organisme modèle, la souris[21]. Des estimations récentes montrent cependant que plus de 50 % du génome est néanmoins transcrit, ce qui suggère que notre compréhension des fonctions de l'ARN est encore très parcellaire. Pourquoi en effet payer un tel prix métabolique s'il ne s'agit que de « junk DNA » inutile ? De plus en plus d'expériences viennent apporter la preuve qu'en dépit de l'incompréhension et du désintérêt initial de la communauté scientifique pour le « junk DNA », ces séquences d'ADN transcrites mais non traduites sont des acteurs clés dans l'évolution et la diversité des génomes(voir par exemple cette très bonne revue sur les éléments transposables[22]). De même, nous savons aujourd'hui que les séquences amont et aval de la séquence codante de l'ARNm sont essentielles et jouent un rôle extrêmement important en termes de régulation et de signalisation.

Pour se donner un ordre d'idée de la richesse de fonctions biologiques des ARN, revenons au dogme central de la biologie moléculaire, et concentrons-nous sur les différents partenaires ARN de la transcription, ARNm, ARNr et ARNt. Ils se sont tous révélés avoir des fonctions biologiques bien plus riches et complexes que celles initialement identifiées. Nous allons exposer rapidement, et sans

viser l'exhaustivité, certaines de ces fonctions et donner des exemples plus généraux des découvertes qui ont modifié notre compréhension de la diversité des fonctions biologiques de l'ARN.

### 2) ARNm et riboswitches

Les ARNm, à travers les différentes conformations qu'ils peuvent prendre, ont un rôle extrêmement important dans la régulation transcriptionelle ou traductionelle, ce qui était jusqu'alors considéré comme étant l'apanage des protéines. Ainsi, certains ARN messagers possèdent, en amont de la séquence codante où on trouve les sites d'interactions essentiels au démarrage de la traduction (comme les sites consensus de Shine-Dalgarno pour les bactéries ou de Kozak pour les eucaryotes) des séquences capables de fixer des petites molécules. Ces séquences, appelée riboswitchs (ou riborégulateurs), ont la particularité de se lier de façon très affine à une molécule cible à travers une région appelée aptamère. La reconnaissance spécifique d'un petit métabolite induit un changement conformationel important de la structure du riboswitch. Ce faisant, il peut y avoir, comme illustré dans la figure I-9, séquestration ou libération des sites de fixation du ribosome en fonction de facteurs extérieurs (présence ou absence d'un cofacteur, concentration d'un métabolite secondaire, changement d'environnement, …). Ce type de contrôle de l'initiation de la traduction à travers un riboswitch est utilisé par certains organismes pour mettre en place un rétrocontrôle direct entre la production d'un métabolite secondaire et l'expression de l'enzyme responsable de sa biosynthèse[23]. Le même type de procédé mettant en jeu des riboswitchs est aussi employé pour le contrôle de la terminaison de la transcription par séquestration ou libération du terminateur dans la séquence en aval de la région codante[24]. La séquence en aval de la séquence codante des ARNm est par ailleurs impliquée directement dans le contrôle de la vitesse de dégradation de l'ARN ainsi que dans sa localisation cellulaire[25,26].

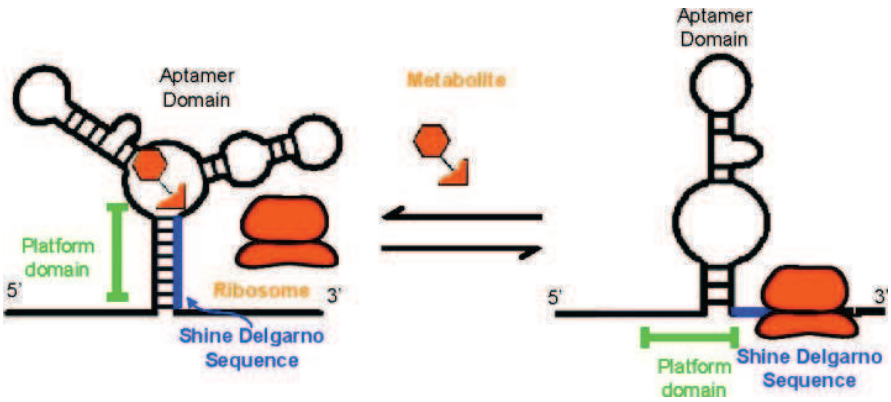



**Figure I-9 Exemple de fonctionnement d'un riboswitch : en présence du métabolite, le ribosome ne peut pas initier la traduction[14]**

Ainsi, contrairement à ce qui était initialement envisagé, les ARN, en dépit de leur faible diversité chimique et leur très forte charge négative, sont capables de reconnaitre des petites molécules et par là, de jouer un rôle important dans la régulation de nombreuses voies de biosynthèse. La variété des molécules reconnues par les riboswitchs est très large, allant d'acides aminés hydrophobes à des ions métalliques en passant par des petites molécules organiques[27]. De façon plus générale, en utilisant des techniques d'évolution dirigée comme la technique SELEX (« systematic evolution of ligands by exponential enrichment »)[28], il est théoriquement possible de sélectionner in vitro des aptamères capables de reconnaitre de façon très affine virtuellement presque n'importe quelle molécule, à l'instar de ce que l'on peut faire avec les anticorps. Il suffit pour cela de partir de banques combinatoires de plusieurs milliards d'oligonucléotides synthétiques et de sélectionner l'aptamère voulu à travers plusieurs cycles alternant sélection par affinité et amplification par réaction en chaîne par polymérase (PCR) (cf figure I-10).

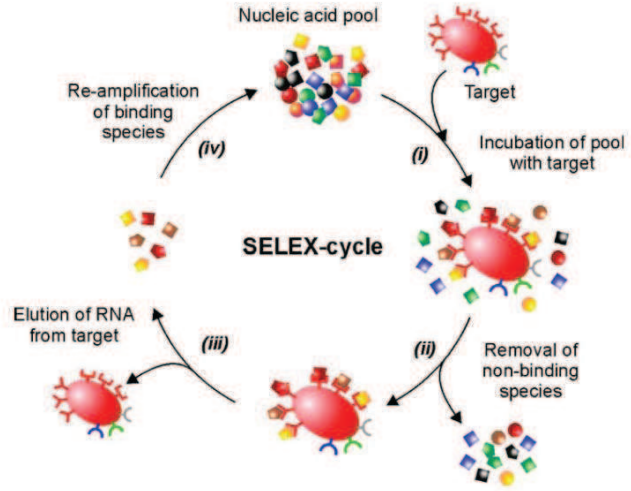



**Figure I-10 Le principe de la méthode SELEX**
(source Univ. Toronto, aptamer Lab,
http://bio349.biota.utoronto.ca/20069/bio349jerry1/background.html)

Cette diversité dans les ligands potentiellement reconnus ne se fait cependant pas au détriment de la spécificité. Il a en effet été montré qu'un aptamère sélectionné in vitro pour se fixer sur la théophylline, était capable de discriminer la théophylline de la caféine, pourtant très proches chimiquement, ne différant que par un groupe méthyle[29].

Les ARNm, et plus particulièrement les régions en amont ou en aval de leur région codante, permettent donc un niveau de régulation hautement efficace, très fin et rapide. Ils sont donc bien plus que de simples supports temporaires de l'information génétique. Ces mécanismes, initialement

sous-estimés, sont présents dans de nombreuses espèces de bactéries, d'archées, de champignons ou de plantes[30]. En particulier, les riboswitchs semblent être très utilisés par les procaryotes pour contrôler diverses voies de biosynthèses de métabolites essentiels, tel les flavines (cofacteurs d'oxydoréduction), les thiamines (impliquées dans le métabolismes des sucres)[31,32] ou bien la méthionine[33]. Certaines études bioinformatiques estiment même à environ 2% la fraction des ARNm régulés par des riboswitchs dans *B. subtilis*[24]. La synthèse de ces métabolites n'ayant pas lieu chez l'homme, ceci fait des riboswitchs de bonnes cibles thérapeutiques potentielles. Interférer avec les fonctions de régulation des riboswitchs conduirait à déstabiliser les voies métaboliques vitales de l'organisme cible, ce qui serait potentiellement létal. La découverte récente, par l'équipe de Breaker en 2009, de composés antibiotiques naturels ciblant spécifiquement les voies de régulation de la biosynthèse des flavines est une excellente preuve de concept de la validité de cette approche[34].

Ce domaine est très prometteur et connait actuellement un développement très rapide. On découvre tous les mois de nouveaux riboswitchs et on commence à comprendre précisément la structure de certains d'entre eux[35] ainsi que leur interaction dynamique avec leur ligand[36] De nombreuses équipes se sont récemment intéressées à la possibilité de concevoir des riboswitchs artificiels en utilisant des méthodes d'évolution in vitro de type SELEX[28]. Il a par exemple été démontré qu'il était possible de réguler la traduction d'un gène rapporteur en insérant en amont de sa séquence codante un aptamère artificiel reconnaissant une petite molécule[37]. Il est également envisagé de détourner les riboswitchs naturels en exploitant leur modularité pour créer divers systèmes artificiels comme des biosenseurs ou des marqueurs[38].

Les ARNm, et plus particulièrement les riboswitchs, en raison de leur rôle clé dans les phénomènes de régulation de la biosynthèse de métabolites essentiels et de signalisation cellulaire, constituent des cibles thérapeutiques de la première importance. Cela justifie l'intérêt de développer de nouveaux ligands d'ARN ciblant spécifiquement ces structures biologiques. Ces ligands sont par ailleurs essentiels au développement de systèmes artificiels utilisant des aptamères ou des riboswitchs.

### 3) ARNr et ribozymes

En 2000, des études de cristallographie menées par les équipes de Steitz, Yonath et Ramakrishnan, récompensés par le prix Nobel de chimie en 2009, ont permis d'analyser très finement la structure des différentes sous-unités du ribosome (figure I-11). Le ribosome bactérien est formé de deux sous-unités, la petite sous-unité 30S et la grande sous-unité 50S, qui sont composées à la fois d'ARN et de protéines ribosomiques. La petite sous-unité a pour rôle de contrôler que seul

l'ARNt correct soit autorisé à incorporer l'acide aminé qu'il porte dans la chaîne peptidique en croissance. La grande sous-unité permet, quant à elle, l'élongation du peptide à travers la catalyse de la formation de la liaison peptidique (voir figure I-7). Il a été montré que les ARN ribosomaux, loin d'être cantonnés à un simple rôle structural, étaient sans doute au cœur de l'activité catalytique de formation de la liaison peptidique[39,40,41]. Cette découverte fondamentale a confirmé que les ARN pouvaient avoir des activités catalytiques à l'instar de ce que l'on connaissait pour les enzymes protéiques.

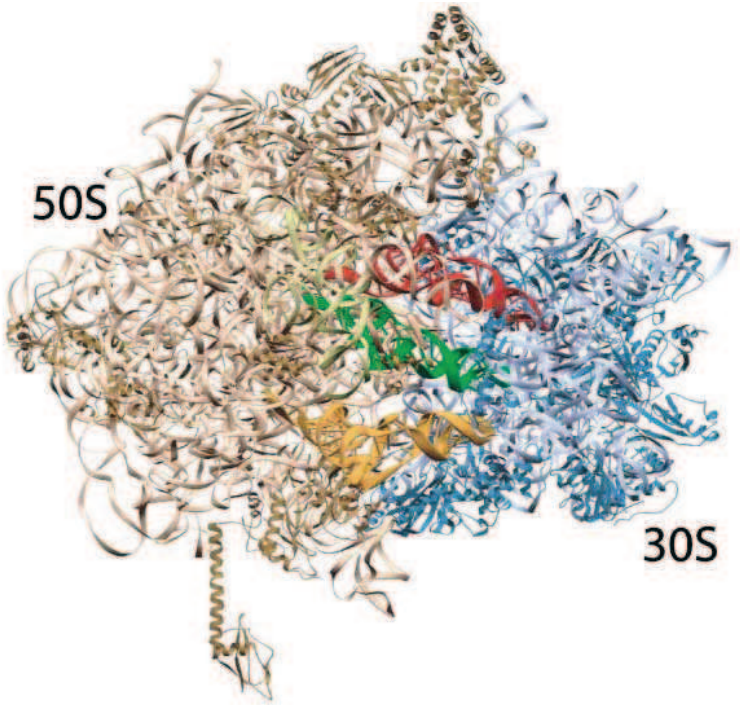



**Figure I-11 Structure du ribosome procaryote avec les deux sous-unités 50S et 30S**
La sous-unité 30S est constituée de l'ARN 16S et de 21 partenaires protéiques. La sous-unité 50S est quant à elle composé des ARN 5S et 23S, ainsi que de 31 partenaires protéiques[19].

Le ribosome est une cible thérapeutique bien connue et représente un excellent exemple de l'importance des ARN en tant que cible thérapeutique. Cibler cette machinerie permet en effet d'interférer avec la synthèse protéique en gênant l'activité catalytique de formation de la liaison peptidique. Les différences entre ribosomes eucaryotes et procaryotes peuvent être exploitées pour développer des molécules tendant à cibler spécifiquement le ribosome procaryote et minimiser la toxicité des éventuels traitements chez l'homme.

Par ailleurs, depuis des millions d'années, les microorganismes se sont battus contre des compétiteurs bactériens pour le contrôle de ressources limitées. La pression de sélection a conduit

au développement de tout un arsenal de molécules naturelles antibiotiques qui constituent des armes essentielles à la survie de ces microorganismes. Les produits naturels antibiotiques ciblant le ribosome, et plus particulièrement les ARNr, sont très nombreux et divers. Aujourd'hui, les substances actives antibiotiques utilisées de façon thérapeutique sont dans leur majorité produites par des microorganismes ou bien issues d'hemi-synthèses à partir d'architectures chimiques naturelles[42,43].

On peut citer par exemple la tétracycline, qui vient inhiber la liaison des ARNt à la sous-unité 30S ou bien le chloramphénicol, qui vient inhiber l'activité peptidyl transférase de la sous-unité 50S. Dans la famille des aminoglycosides, que l'on étudiera en détail plus tard, la kanamycine ou la néomycine, vont agir au niveau du site A de décodage de l'ARN 16S dans la sous-unité 30S, augmentant le taux d'erreur dans la traduction. Comme nous le verrons ultérieurement dans la partie I-II-D, ces molécules interférent avec les propriétés de reconnaissance du site A qui ne va donc pas pouvoir discriminer l'ARNt correspondant au codon exposé[44,45].

Le ribosome n'est pas le seul exemple où un ARN est capable d'agir comme une enzyme. Les travaux pionniers d'Altman au début des années 80 avaient déjà démontré la capacité de certains ARN à catalyser des réactions chimiques. En s'intéressant à la Ribonucléase P (RNAse P), une enzyme de la maturation des ARNt constituée d'un complexe ARN/protéine qui semble être conservée dans l'ensemble du monde vivant, il montre que l'activité catalytique est due uniquement à la partie ARN du complexe[46]. Ceci fait écho aux découvertes de Cech qui s'intéresse à la capacité de certaines séquences d'ARN à se cliver spécifiquement en l'absence de protéines[47]. Leurs travaux, récompensés par le prix Nobel en 1989, démontrent ainsi que certains ARN, qu'on appelle ribozymes, sont capables d'avoir une activité catalytique indépendante, spécifique et aussi efficace que ce que l'on connait avec les enzymes protéiques. Ceci est rendu possible en raison de la capacité des ARN à se replier pour former des structures tertiaires, où il existe des poches susceptibles de se lier spécifiquement à des partenaires moléculaires. Des études récentes de ribozymes par difractions des rayons X, par RMN et microscopie électronique nous donnent des informations quant à l'orientation précise des résidus catalytiques et des cofacteurs éventuels (voir par exemple la résolution de la structure de la RNase P de *T.maritima*[48,49]). Il s'agit d'une première étape dans la tentative de corrélation entre la structure et de la fonction de ces ARN particuliers.

Même si le ribosome et la RNAse P sont encore à ce jour les seuls ribozymes connus possédant un turnover supérieur à 1[1], on découvre de plus en plus de ribozymes qui sont leur propre substrat et qui jouent un rôle important de régulation post-transcriptionelle. On peut citer comme exemple la régulation de la production de glucosamine-6-phosphate chez *B. subtillis* (figure I-12). Le groupe de Breaker a démontré en 2004 que le niveau d'expression de l'enzyme GlmS, responsable de la biosynthèse de la glucosamine-6-phosphate, était directement régulé par le niveau de glucosamine dans le milieu selon un mécanisme de rétrocontrôle négatif d'un type nouveau. En effet, l'ARNm de la protéine GlmS contient, en amont de la séquence codante, un ribozyme capable de se lier spécifiquement à la glucosamine. Lorsque la concentration en glucosamine atteint un certain seuil, le ribozyme est capable de se cliver, rendant impossible la traduction de l'ARNm. Ce mécanisme de régulation post-transcriptionelle, extrêmement économe en moyen et finalement assez simple, permet d'éviter de produire l'enzyme GlmS lorsque qu'il y a déjà suffisamment de glucosamine-6-phosphate dans le milieu. Il ne fait pas intervenir d'autres partenaires que l'ARNm, qui sert à la fois de senseur de concentration et de régulateur de la production enzymatique et bien sûr de support de l'information génétique. De façon très intéressante, les auteurs ont par ailleurs montré que ce système était modulaire et qu'on pouvait donc le détourner en le plaçant en amont de gènes reporteurs[50].

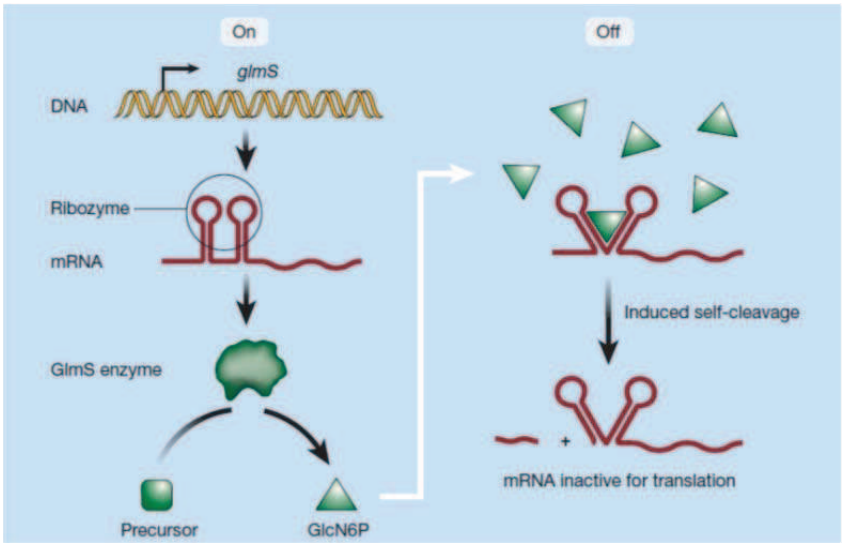



**Figure I-12 Rétrocontrole négatif de la glucosamine 6 phosphate sur la transcription de l'enzyme GlmS à travers un ribozyme** [51]

[1] Des travaux récents de cristallographie tendraient à prouver que le splicéosome, la machinerie chargée d'assurer l'épissage chez les eucaryotes, est aussi un ribozyme possédant un turnover supérieur à 1[378].

Cette diversité de fonctions biologiques rendues possibles par les ARN, et en particulier leur capacité à être à la fois support de l'information génétique (ARNm) mais aussi catalyseur enzymatique (ribozyme), a conduit une partie de la communauté scientifique à proposer l'hypothèse du « RNA world ». Dans cette théorie, l'ARN serait la première des macromolécules biologiques à être apparue dans le monde prébiotique et constituerait le précurseur universel des molécules de la vie [52,53].

Par ailleurs, la capacité des ARN à se lier spécifiquement à des partenaires moléculaires et à avoir des activités catalytiques en fait de potentielles cibles thérapeutiques. Le ribosome est d'ailleurs la cible principale des antibiotiques actuellement sur le marché.

### 4) ARNt, du décodage du code génétique à l'initiation de la rétrotranscription

Comme nous l'avons vu précédemment, lors de la traduction, les ARNt assurent la lecture des codons de l'ARNm par l'intermédiaire de leur anticodon au niveau des sites A et P de la petite sous-unité du ribosome. Ils apportent l'acide aminé fixé à leur extrémité au niveau du centre péptidyl-transférase situé dans la grande sous-unité du ribosome, où a lieu la synthèse de la chaîne polypeptidique[19].

Les ARNt, ainsi que leurs mimes, se sont cependant avérés être impliqués dans de nombreux systèmes biologiques autres que la traduction, comme par exemple la réplication, l'épissage ou la régulation transcriptionelle[54]. Ainsi, les ARNt ne sont pas de simples « décodeurs » de l'information génétique, faisant le lien entre les 64 codons possibles de l'ARNm et une séquence protéique écrite avec 20 acides aminés naturels, mais bien des objets biologiques complexes, dont on commence seulement à deviner tous les rôles. A titre d'exemple, concentrons-nous sur l'$ARNt^{Lys}_{3}$, cible très étudiée dans l'équipe de recherche dans laquelle je me suis intégré.

Au début des années 80, le monde est confronté à l'émergence d'une pandémie d'un type nouveau, le SIDA, qui provoque un affaiblissement très important du système immunitaire des patients touchés. En 1983, on découvre que cette maladie est due au virus de l'immunodéfience humaine (VIH), un rétrovirus qui a la particularité de conserver son information génétique sous forme d'ARN et non d'ADN[55]. Cette particularité, qui avait déjà été observée pour d'autres virus par Temin et Baltimore[56,57], vient enrichir notre compréhension du dogme central de la biologie moléculaire, en démontrant une fois de plus que l'ADN mais aussi l'ARN peuvent être les supports de l'information génétique.

Pour se répliquer, le VIH va détourner la machinerie de biosynthèse de son hôte et intégrer son ARN génomique à l'ADN de son hôte en le convertissant en ADN double-brin à l'issue de la réaction de rétrotranscription. L'enzyme catalysant ce processus, la transcriptase inverse, a besoin d'une amorce venant s'apparier avec l'ARN viral matrice afin d'initier l'incorporation des désoxyribonucléotides complémentaires. Pour cela, le virus détourne de sa fonction première un ARN de transfert cellulaire, l'ARNt$^{Lys}_{3}$, afin de l'utiliser comme amorce de polymérisation[58]. L'ARN viral contient en effet une séquence complémentaire d'une partie de l'ARNt amorce via des appariements Watson-Crick (site de liaison PBS, pour Primer Binding Site). Cette hybridation, qui permet l'initiation de la rétrotranscription et donc la réplication du virus, est aussi dépendante de l'acitivité chaperone d'une petite protéine virale constituée de deux doigts de zinc, la nucléocapside. Le mécanisme de l'initiation de la rétrotranscription inverse a été étudié en détail au laboratoire[59,60,61](figure I-13). Il a été démontré qu'il était possible de viser cette étape précoce dans le cycle de vie du virus à l'aide de petites molécules[62], en développant une approche de synthèse de ligand par fragment, guidée par RMN[63]. Interférer avec la reconnaissance ARN viral/ ARNt$^{Lys}_{3}$ pourrait potentiellement être une nouvelle voie thérapeutique à explorer dans la lutte contre le VIH.

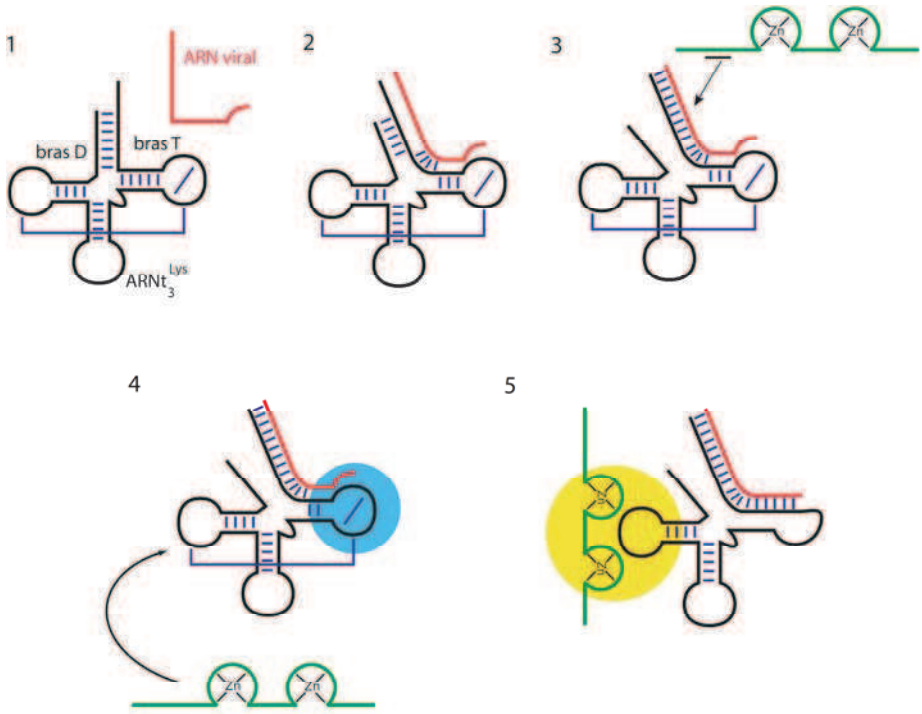



**Figure I-13 L'initiation de la transcription inverse du VIH par hybridation entre le PBS de l'ARN viral (en rouge) et l' ARNt$^{Lys}_{3}$ (en noir), assistée par la nucléocapside (en vert)[61]**

De façon plus générale, on voit l'intérêt qu'il y a à concevoir des molécules capables de cibler spécifiquement une cible ARN, à la fois car ces molécules représentent des outils essentiels à une meilleure compréhension des mécanismes biologiques complexes qui sont en jeu mais aussi car elles permettent potentiellement d'ouvrir de nouvelles pistes thérapeutiques.

### 5) Interférence ARN

La biologie des ARN est extrêmement riche et recèle encore de nombreuses surprises. Nous avons vu précédemment que les ARN, en plus de leur rôles biologiques « canoniques » dans la transcription et la traduction, pouvait catalyser des réactions, être des acteurs clés dans des phénomènes biologiques comme la retrotranscription ou être impliqué dans des cascades complexes de régulation ou de signalisation cellulaire. On découvre chaque année de nouvelles fonctions à des ARN connus, mais aussi de nouveaux ARN ayant des fonctions jusqu'alors inconnues. Par exemple, la découverte de l'interférence ARN il y a une vingtaine d'années a eu un retentissement énorme car elle a complètement bouleversé notre compréhension de la régulation de l'expression des gènes.

En 1998, C. Mello et A. Fire, prix Nobel de médecine en 2006, montrent que l'on peut sélectivement « éteindre » l'expression de gènes chez le ver *C.elegans* en introduisant de manière artificielle des fragments d'ARN double brin, correspondant à la séquence d'une partie du gène d'intérêt[64]. Il a été démontré ultérieurement que ces fragments d'ARN sont pris en charge par une machinerie cellulaire complexe, comprenant l'enzyme Dicer et le complexe RISC (pour "RNA-Induced Silencing Complexe"), qui va dans un premier temps tronçonner cet ARN double brin pour former de petits fragments d'ARN. Ceux-ci vont pouvoir, dans un second temps, interagir de façon spécifique avec l'ARNm du gène cible à travers des interactions de type Watson-Crick, pour, soit initier sa dégradation, soit empêcher sa traduction par divers moyens[65,66]. On a depuis découvert qu'il existe de petits fragments d'ARN naturels, appelés miARN (mi pour micro), allant de 20 à 30 nucléotides. Ceux-ci semblent être au cœur de mécanismes de régulation très fins qui sont communs à de nombreux organismes (invertébrés, plantes, insectes, et même vertébrés comme par exemple les mammifères,…)[67]. Leur importance est énorme : certaines études estiment à environ 50 % le pourcentage d'ARNm régulés par des mécanismes d'interférence ARN chez les vertébrés[68]. De petits fragments d'ARN artificiels, appelés les siARN (si pour small interfering), ont été développés pour exploiter ces mécanismes.

Cette découverte a eu un impact très important sur la communauté scientifique car ce phénomène de régulation par interférence ARN est généralisable à des nombreux organismes et permet de comprendre les voies de régulation complexes jusqu'alors inexpliquées. D'un point de vue

pratique, l'interférence ARN constitue par ailleurs un formidable outil pour étudier la fonction de gènes d'intérêt en réprimant spécifiquement leur expression. Enfin, l'interférence ARN ouvre potentiellement de nouvelles voies thérapeutiques. On peut par exemple imaginer moduler leur activité naturelle en les ciblant à l'aide de petites molécules[69]. Une autre stratégie très étudiée par les entreprises de biotechnologie est d'introduire directement les petits ARN interférents dans les organes malades en utilisant des méthodes de vectorisation[70].

### 6) Conclusion : les ARN, cibles thérapeutiques de première importance.

On l'aura compris l'ARN joue un rôle clé dans de très nombreux processus biologiques complexes, et en particulier dans les organismes multicellulaires comme l'Homme[66]. Outre son rôle d'intermédiaire dans le transfert de l'information génétique, il est un acteur essentiel dans divers phénomènes de régulation, de catalyse, de cascades de signalisation cellulaire ou encore d'epissage. L'ARN constitue donc, en lui-même, une cible thérapeutique de première importance[7169].

Par ailleurs, même si les protéines restent au cœur des stratégies de ciblage de l'industrie pharmaceutique, on découvre aujourd'hui les limitations d'une telle approche qui conduit à un grave déficit dans le nombre de nouveaux médicaments découverts. En effet, s'il est possible de cibler une enzyme ou un récepteur, interférer avec une interaction protéine/protéine ou avec une protéine de structure s'avère assez complexe. Ainsi, certaines estimations considèrent que, chez l'homme, seule une infime partie des protéines peut constituer une cible thérapeutique valable[72]. Ceci est visible lorsqu'on étudie les cibles des médicaments actuellement sur le marché en santé humaine. Seules 207 cibles sont actuellement visées sur les dizaines de milliers potentiellement disponibles chez l'homme[14]. On constate d'autre part un manque flagrant d'innovation dans les cibles thérapeutiques visées puisqu'entre 1989 et 2000, 76 % des substances approuvées par la FDA visaient une cible thérapeutique déjà connue[73]. Ainsi, non seulement une large partie du génome humain reste inexploré du point de vue thérapeutique, mais l'exploitation des cibles thérapeutiques connues est de moins en moins efficace dans la recherche de nouveaux médicaments avec des stratégies classiques de ciblage des protéines.

Cibler les ARN constitue une stratégie pour dépasser ces limitations[74,75]. On peut par exemple envisager de cibler l'ARNm d'une cible protéique identifiée par screening mais où le développement de ligand n'a pas été concluant, ce qui permet d'utiliser les connaissances biologiques acquises durant la phase initiale de screening. Les ligands d'ARNm offrent d'autre part des avantages significatifs puisqu'ils permettent virtuellement, non seulement de stopper ou diminuer la

production de la protéine cible en empêchant la traduction, mais aussi d'augmenter cette production en stabilisant et en augmentant la durée de vie de l'ARNm.

Les ARN, pour leur fonction biologique propre ou bien en tant qu'intermédiaires dans la synthèse protéique, constituent des cibles thérapeutiques extrêmement intéressantes. Malgré leur énorme potentiel, les ARN restent néanmoins encore assez peu exploités en thérapeutique en raison des difficultés inhérentes à leur étude et à leur ciblage spécifique par des petites molécules. La partie qui suit a pour but de rappeler quelles sont ces difficultés et quelle est leur origine. Nous nous intéresserons dans un second temps à une classe de ligands d'ARN très connus, les aminoglycosides, à même de répondre à certaines de ces difficultés.

## *D - Les difficultés à cibler les ARN*

Développer des ligands visant spécifiquement les cibles thérapeutiques non traditionnelles que sont les ARN représente un défi important. Comme nous l'avons rappelé en introduction, les ARN sont capables de former des structures tertiaires, comportant des boucles et des cavités, qui leur permettent de se lier spécifiquement à de petites molécules. Ces structures restent bien moins étudiées et comprises que les structures protéiques en partie en raison de la difficulté à produire des ARN en quantité et pureté suffisante. Ainsi, la première difficulté lorsqu'on cherche à concevoir des ligands d'ARN est l'accessibilité à des quantités suffisantes de cible pour pouvoir en étudier les propriétés structurales et biologiques.

Par ailleurs, les caractéristiques physicochimiques particulières des ARN influencent grandement les stratégies à mettre en place pour potentiellement développer des ligands affins et spécifiques de ces structures.

Tout d'abord, du fait du fort caractère polyanionique des ARN, les ligands d'ARN connus possèdent souvent un caractère cationique très marqué qui peut poser des problèmes de biodisponibilité. Les nombreuses fonctions polaires portées par les bases azotées ne font que renforcer ce biais vers de ligands très polaires. Si on se réfère aux fameuses mais controversées règles de Lipinksi[76], la majorité des ligands d'ARN connus, et en particulier les antibiotiques de la famille des aminoglycosides, constituent des violations flagrantes de cette loi statistique, possédant à la fois une charge trop élevée et une masse moléculaire trop importante. En conséquence, les criblages des vastes librairies de composés disponibles dans l'industrie pharmaceutique donnent des résultats décevants lorsqu'on cherche à cibler un ARN. Ces librairies reflètent en effet l'histoire du développement de l'industrie pharmaceutique. Elles sont majoritairement composées de molécules

dites « drug-like », facilement accessibles par les techniques usuelles de synthèse organique, et sont optimisées pour le criblage de cibles protéiques. Même si on découvre aujourd'hui des ligands d'ARN plus en adéquation avec les critères généralement admis pour des candidats médicament, c'est-à-dire hydrosolubles mais néanmoins capables de traverser passivement les membranes biologiques, l'amélioration du taux de succès des campagnes de criblage passe par une meilleure compréhension de ce qui fait un bon ligand d'ARN[77,78].

La physicochimie particulière des ARN influence donc l'espace chimique des ligands pour inclure des zones non couvertes par les ligands de protéines. Il est essentiel de développer de nouvelles stratégies de synthèse organique à même de fournir des composés plus susceptibles de se lier avec des ARN. Comme nous le verrons par la suite, c'est un des grands enjeux dans la découverte de ligands d'ARN, qui a conduit une partie de la communauté scientifique à développer de nouvelles molécules basées sur l'architecture de ligands connus, les aminoglycosides (voir partie I-II). D'autres stratégies prometteuses mettent en œuvre de nouvelles méthodes de screening, dite « par fragment », afin d'explorer un espace chimique plus vaste et plus adapté aux ARN que ce qui est permis par les librairies traditionnelles (voir partie I-III-B).

Le squelette polyphosphate des ARN induit d'autre part une très forte solvatation de ces biomolécules. De même, les cations mono- ou bivalent présents en solution (comme par exemple $Mg^{2+}$) jouent un rôle essentiel à la fois dans la structure mais aussi dans la fixation d'éventuels ligands. Ceci complique énormément les tentatives de modélisation des structures d'ARN et de screenings virtuels *in silico* car il faut tenir compte de la déstabilisation de la cage de solvatation des ARN par leur ligand, ce qui a des conséquences importantes en termes entropiques et donc en temps de calcul [79]. Certaines études récentes suggèrent même que les molécules d'eau et les cations présents autour de l'ARN jouent un rôle actif dans la fixation des ligands, conférant une très grande plasticité de reconnaissance. Il est par exemple estimé que près d'un tiers des interactions des aminoglycosides avec leur cible l'ARN ribosomique 16S se fait à travers des molécules d'eau[12,80]. Pour concevoir rationnellement un ligand d'ARN, il faut ainsi prendre en compte de multiples partenaires en plus de l'ARN lui-même et de son ligand.

Enfin, les ARN sont des biomolécules assez flexibles et bien souvent, leur activité biologique est dépendante de leur capacité à être en équilibre entre plusieurs conformations[81]. La vision classique où une séquence donne une seule structure qui possède une seule fonction est maintenant remplacée par une vision dynamique de l'ARN en équilibre entre de nombreuses structures qui peuvent être traitées par une machinerie cellulaire complexe pour donner un large échantillon de fonctions biologiques possibles[82]. Il y a ainsi bien souvent capture conformationelle de la cible par

son ligand, ou encore accommodation structurale de la cible lors de la fixation du ligand (induced-fit en anglais)[83,84] (voir figure I-14). Pour compliquer le tout, le ligand lui aussi peut faire l'objet d'un changement conformationel important lors de sa liaison avec sa cible ARN. On parle alors d'accommodation structurale mutuelle. C'est notamment le cas des aminoglycosides comme nous le verrons dans la partie suivante.

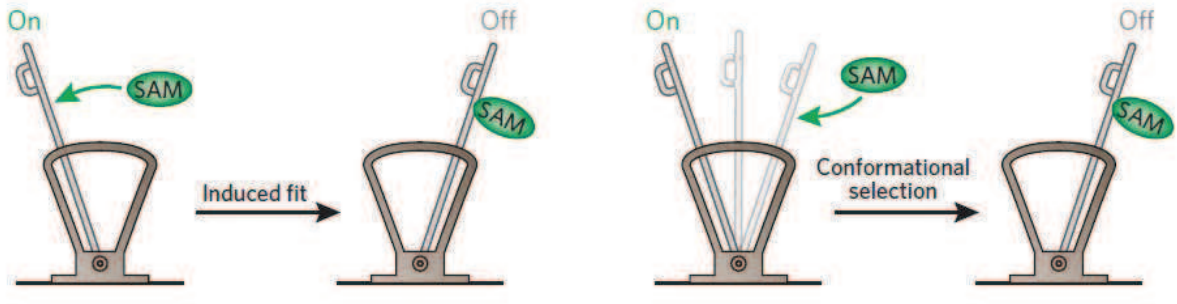



**Figure I-14 Le principe de l'accommodation et de la capture conformationelle d'un ARN par son ligand (en vert)[84].**

Cette souplesse leur confère une nature dynamique qui complique elle aussi les analyses par modélisation et les études structurales. On ne peut en effet considérer une seule structure de la cible pour concevoir un ligand, mais bien une cible en équilibre complexe entre plusieurs structures. La définition rationnelle de ligands doit prendre en compte les perturbations de la dynamique conformationelle des ARN lors de la fixation du ligand et les couts énergétiques associés. Contrairement au cas classique d'une protéine qu'on va chercher à inhiber à l'aide d'un ligand de forte affinité, un ligand d'affinité modérée peut s'avérer suffisant pour interférer avec une cible ARN, un léger changement de la dynamique conformationelle d'un ARN pouvant avoir des conséquences importantes au point de vue biologique[853686]. D'un point de vue pratique, cela signifie qu'il faut être capable de détecter des ligands de relativement faible affinité envers les ARN lors des criblages haut-débit. Cela implique un effort d'adaptation des techniques actuellement utilisées pour la détection de ligands d'ARN, qu'elles fassent appel à la fluorescence, à la RMN ou bien à la spectrométrie de masse.

En conclusion, en dépit des informations que l'on commence à avoir sur la structure et les propriétés physicochimiques des ARN, développer des ligands affins et spécifiques de ces biomolécules reste un défi important. Le développement de méthodes générales pour explorer la liaison spécifique de petites molécules sur l'ARN reste encore modeste. Ceci contraste fortement avec le cas de l'ADN, domaine de recherche plus mature, où de nombreuses molécules ont été

développées pour cibler spécifiquement des fragments d'ADN en fonction de leur structure primaire[87]. Les produits naturels décrits pour se fixer sur les ARN, et en particulier les aminoglycosides qui feront l'objet de la partie I-II, peuvent constituer des sources d'inspiration à même de dépasser les difficultés actuelles.

# II - Les aminoglycosides, « ligands universels » d'ARN

Dans la partie précédente, nous avons vu la richesse des fonctions biologiques des ARN et à quel point il était intéressant de pouvoir les cibler avec des petites molécules. Nous avons également exposé certaines des difficultés à concevoir des petites molécules capables d'interférer spécifiquement avec ces cibles non traditionnelles, polyanioniques et très flexibles.

Si nous découvrons aujourd'hui de plus en plus de nouveaux ligands d'ARN, la famille des aminoglycosides, qui regroupe plusieurs dizaines de produits naturels ou semi-synthétiques, reste la plus étudiée et la plus riche[88]. Leur emploi depuis plus de 60 ans comme molécules antibiotiques interagissant avec les ARN du ribosome constitue une preuve éclatante de la possibilité de cibler des ARN avec de petites molécules. Il a par ailleurs été démontré que les aminoglycosides avaient la capacité de reconnaitre des structures d'ARN très variées, ce qui a conduit une partie de la communauté scientifique à les qualifier de « ligands d'ARN universels » et à les utiliser pour comprendre les multiples facettes des interactions petites molécules/ARN[89]. Leur architecture moléculaire est un très bon point de départ pour la compréhension de ce qui fait un bon ligand d'ARN ainsi que pour la conception de nouveaux ligands, outils d'étude des ARN ou molécules à visée potentiellement thérapeutique[14].

Après un bref survol de l'histoire de la découverte des aminoglycosides ainsi que de leur utilisation médicale en tant qu'antibiotiques, nous allons nous pencher sur les caractéristiques physicochimiques des aminoglycosides, leur mode d'action biologique et les phénomènes émergents de résistance bactérienne. Nous verrons enfin que, si les aminoglycosides ont pour cible principale le ribosome, ils peuvent également interagir avec de nombreuses autres cibles biologiques ce qui renforce d'autant plus l'intérêt du développement de nouveaux analogues fondés sur le squelette des aminoglycosides.

## *A - Historique de la découverte et utilisation médicale*

Les aminoglycosides sont des composés naturels produits par des bactéries des genres *Micromonospora* ou *Streptomyces* appartenant à l'ordre des Actinomycètes. Ces molécules polycationiques et hydrophiles ont un large spectre antibactérien, particulièrement contre les bactéries gram-négatif, peu sensibles aux pénicillines et aux sulfamides. Elles sont majoritairement bactéricides, avec une action rapide et dose-dépendante qui constitue, avec leur pharmacocinétique très prévisible, un de leurs grands avantages thérapeutiques[90]. Les voies de biosynthèse des

aminoglycosides commencent à être connues, ce qui permet d'envisager des approches biotechnologiques pour la synthèse de nouveaux analogues. Dans les chapitres II et III, nous nous pencherons sur la biosynthèse de la néomycine chez *S.fradiae* et sur son détournement potentiel à des fins synthétiques.

**Figure I-15 Une colonie de *Streptomyces*(Streptomyces sp. Acta 1362), producteurs naturels de plus des 2/3 des antibiotiques connus**
(source : Fielder lab http://www.hpfiedler-group.de)

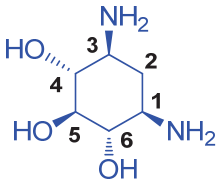
En 1943, la découverte par Selman Waksman et son étudiant Albert Schatz de la Streptomycine, molécule antibiotique naturelle produite par *Streptomyces griseus*[91], a causé une grande excitation au sein de la communauté médicale car ce composé s'est avéré être le premier traitement clinique contre la tuberculose. Ceci valu à Waksman le prix Nobel de médecine en 1952. Dans les années qui suivirent ce succès marquant, de nombreux composés similaires, comme par exemple la néomycine ou la kanamycine, ont été isolés de souches bactériennes de la famille des Actinomycètes[92].

En dépit des problèmes de toxicité rénale et auditive, les aminoglycosides gardent encore aujourd'hui, après plus de 60 ans d'utilisation en tant qu'antibiotiques, une place importante dans l'arsenal thérapeutique moderne. N'étant pas absorbés par voie orale, les aminoglycosides sont principalement utilisés en milieu hospitalier , par exemple en traitement local ou systémique pour lutter contre des infections nosocomiales sévères causées par des bactéries multi-résistantes[88]. Comme pour tous les antibiotiques, l'efficacité thérapeutique des aminoglycosides est menacée par l'émergence de souches bactériennes résistantes[93,94]. Ce problème, lié à la consommation excessive et souvent à mauvais escient des molécules sur le marché, a été identifié comme un des défis majeurs de santé publique par les organisations nationales et internationales. Il est d'autant plus

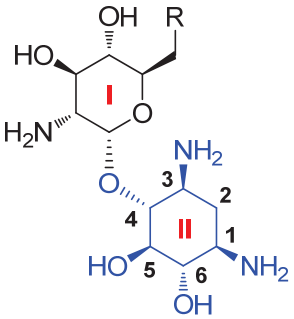
pressant que les maladies infectieuses restent la deuxième cause de mortalité dans le monde[95,96]. Ceci justifie un effort de recherche important pour trouver de nouvelles molécules ayant les mêmes propriétés antibiotiques que les aminoglycosides tout en échappant aux phénomènes de résistance.

## *B - Caractéristiques physicochimiques des aminoglycosides*

Les aminoglycosides sont des oligosaccharides basiques, très polaires et hydrophiles. Au pH physiologique, ils sont chargés positivement en raison de la présence de nombreux groupes ammonium quaternaires. D'un point de vue structural, ils sont la plupart du temps organisés autour d'un cycle central caractéristique de type aminocyclitol sur lequel sont greffés divers sucres aminés. Dans la majorité des cas, le cycle central est la déoxystreptamine (DOS), qui peut être monosubstituée en position 4 comme dans les cas de la néamine ou de l'apramycine, ou bien disubstituée en position 4 et 5 (néomycine, paromomycine, lividomycine, …) ou en position 4 et 6 (kanamycine, tobramycine, …)(figure I-16). Les aminoglycosides sont des molécules comportant de nombreuses fonctions alcool et amine qui peuvent être éventuellement modifiées par méthylation, guanidinylation ou ajout d'une chaine secondaire. La famille des aminoglycosides est relativement large puisqu'elle comprend plusieurs dizaines de molécules naturelles ou semi-synthétiques, certaines structurellement atypiques par rapport à la description ci-dessus. On peut par exemple citer la streptomycine qui possède un cycle central de type streptidine ou bien la spectinomycine qui contient 3 cycles fusionnés.

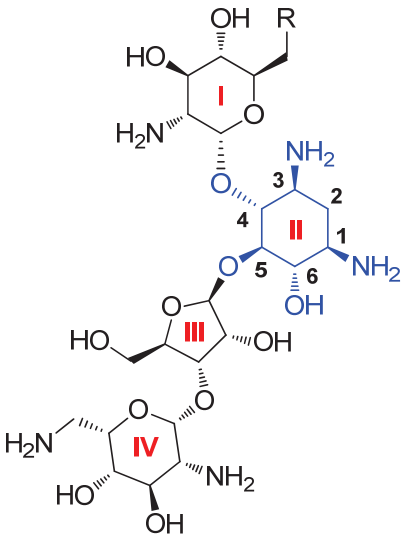



**desoxystreptamine** (**1**, **DOS**), le cycle central conservé dans la plupart des aminoglycosides



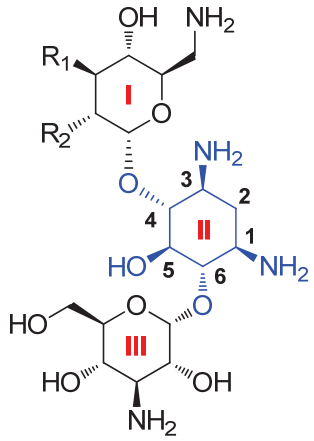

**paromamine** (**2**, R= OH)

**néamine** (**3**, R = $NH_2$)



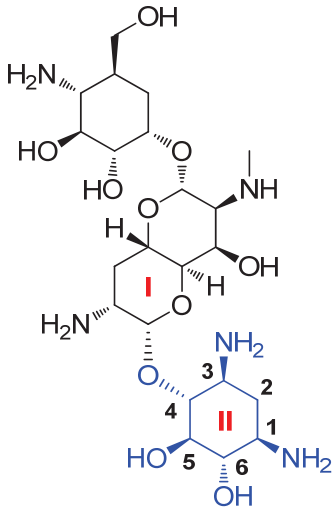
**paromomycine** (**4**, R = OH)

**néomycine** (**5**, R = $NH_2$)

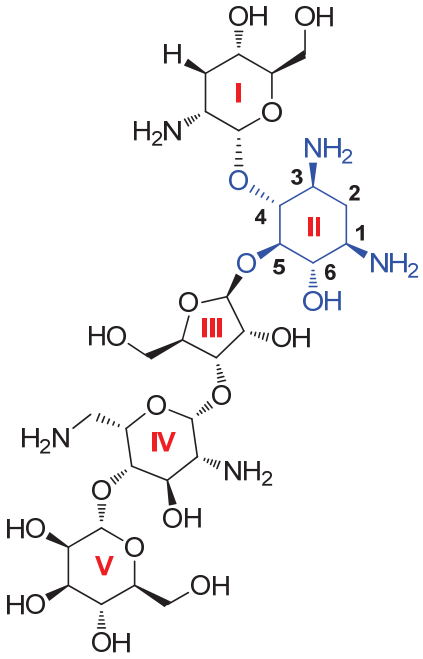


**kanamycine** (**6**, $R_1$= OH, $R_2$=OH)

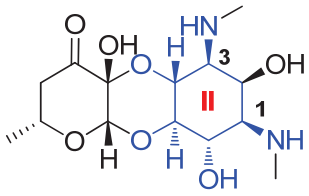
**tobramycine** (**7**, $R_1$ = H, $R_2$=$NH_2$)



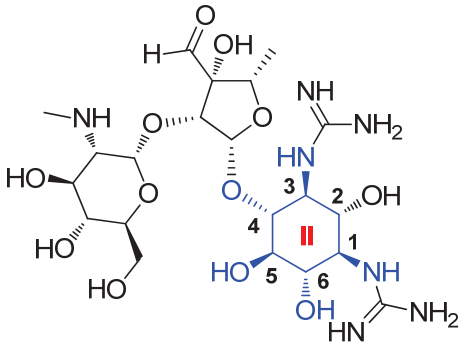
**apramycine 8**



**lividomycine 9**



**spectinomycine 10**



**streptomycine 11**

**Figure I-16 Structure de différents aminoglycosides : le cycle central conservé DOS est indiqué en bleu**

Les aminoglycosides sont des molécules très fonctionnalisées et, en conséquence, leur chimie est connue pour être relativement difficile. A titre d'exemple, le cycle central DOS est un composé *méso* ce qui rend particulièrement délicat les approches de synthèse asymétrique. Ceci a grandement contraint les stratégies d'hémisynthèse de nouveaux ligands d'ARN par fonctionnalisation d'aminoglycosides naturels[97] et a poussé la communauté scientifique à chercher de nouvelles stratégies qui seront exposées dans la partie I-III.

## *C - Interaction avec les ARN : considérations générales*

Les aminoglycosides ont tendance à venir interagir dans les boucles ou les hernies de l'ARN, zones où la structure de l'ARN en hélice est perturbée par des interactions non canoniques, ce qui rend les bases azotées et le squelette ribose/phosphate un peu plus accessibles. De façon intéressante, les interactions ARN/acides nucléique et ARN/protéines semblent suivre les mêmes règles ce qui signifie que les aminoglycosides vont pouvoir entrer en compétition pour les mêmes sites de fixation et donc interférer avec les fonctions biologiques de ces partenaires naturels. Il est intéressant de noter à ce stade que, grâce à la flexibilité conformationelle des liaisons glycosidiques entre les différents aminosucres, les aminoglycosides peuvent adopter différentes conformations pour former des complexes stables avec divers partenaires ARN. Cette flexibilité fait écho à la flexibilité propre des ARN et complique d'autant plus la conception rationnelle de ligands spécifiques d'une cible ARN particulière.

Lorsque l'on compare la structure des aminoglycosides, fortement polycationiques, avec celle des ARN, possédant un squelette polyphosphate chargé négativement, on comprend aisément

l'importance des forces électrostatiques pour décrire les interactions entre ces partenaires. Il est possible d'évaluer cette contribution relative en regardant l'influence d'une variation de la force ionique du solvant sur la constante d'affinité d'une molécule aminoglycosidique pour sa cible ARN. Ainsi, en augmentant progressivement la concentration en NaCl et donc en masquant progressivement les charges, Pilch et ses collègues ont réussi à montrer à l'aide d'expériences de microcalorimètrie et de fluorescence que les forces d'interactions électrostatiques représentaient plus de 50% de l'énergie d'interaction de la paromomycine pour le site A de l'ARNr 16S[98].

L'architecture complexe des aminoglycosides et le réseau d'interactions intramoléculaires permet un positionnement précis des ammoniums quaternaires, selon une géométrie à même d'interagir avec la cible biologique d'intérêt. L'affinité des aminoglycosides pour leur cible est bien souvent corrélée au nombre d'amines présentes ainsi qu'à leur répartition spatiale. En augmentant artificiellement le nombre d'amines portées par une même architecture moléculaire, on obtient bien souvent des analogues plus affins[99]. Des études de modélisation montrent même qu'il est possible de prédire dans une certaine mesure la géométrie des complexes ARN-aminoglycosides en se basant sur le champ electrostatique créé par le squelette polyanionique des ARN, les ammoniums quaternaires ayant tendance à se placer dans les poches les plus chargées négativement, souvent remplies à l'état naturel par des cations divalents comme le magnésium[100].

La basicité des amines semble être finement modulée par la structure même des molécules (methylation, présence d'alcools proximaux, …). On a ainsi montré que les alcools avaient une influence importante sur la basicité des amines proches, pouvant potentiellement baisser leur pKa de plusieurs unités[101]. Les alcools influencent par ailleurs la flexibilité conformationelle des aminoglycosides à travers les liaisons H intramoléculaires qu'ils peuvent potentiellement former.

Cette importance des interactions électrostatiques induit bien souvent un manque de sélectivité des aminoglycosides, qui ont tendance à avoir plusieurs sites de fixation sur un même ARN ou se lient avec des affinités moyennes à faibles à de nombreux ARN (Kd de l'ordre de la dizaine de micromolaires) (à titre d'exemple, voir l'étude de Wong sur les interactions multiples de la néomycine avec le site RRE de l'ARNm de HIV-1[102] ou encore l'étude de Van Boom démontrant la faible sélectivité de différents aminoglycosides pour des aptamères à la tobramycine[103]).

Si les interactions électrostatiques permettent souvent d'expliquer une part importante de l'affinité des aminoglycosides pour leur cible ARN, il ne faut pas pour autant minimiser le rôle des interactions non-ioniques. Ces interactions jouent en effet un rôle de tout premier plan, contribuant à l'affinité globale de la molécule mais souvent aussi permettant d'apporter une certaine sélectivité vis-à-vis de la cible. De nombreuses études structurales mettent par exemple en lumière des

interactions de type empilement hydrophobe entre les sucres des aminoglycosides et les bases azotées de l'ARN. On peut ainsi citer les travaux du groupe de Puglisi qui montrent par RMN que la gentamicine interagit avec le site A de l'ARN 16S à travers un empilement de la zone hydrophobe du site I sur la base G1491[104]. Les structures cristallographiques d'aminoglycosides en complexe avec le ribosome entier ou un fragment du site de décodage confirment l'existence de ce type d'interaction, couplé avec des interactions ioniques et des liaisons H impliquant les amines et les alcools des aminoglycosides avec les bases azotées et le squelette de l'ARN 16S[105,106].

Comme le suggère l'analyse structurale attentive de certains de ces complexes, les aminoglycosides semblent par ailleurs être capables de mimer des contacts semblables à un appariement avec une base azotée non-appariée. On peut voir ce cas de figure lors de l'interaction de l'ARN 16S avec la gentamicine[104] mais aussi avec un aminoglycoside plus exotique comme l'apramycine[107]. Enfin, les molécules d'eau constituent toujours des partenaires clés dans les interactions ARN/petites molécules comme nous l'avons rappelé dans la partie introductive sur les ARN[12]. Les molécules d'eau des cages de solvatation de l'ARN et de l'aminoglycoside vont en effet pouvoir se réorganiser pour écranter plus ou moins fortement les charges, optimiser la géométrie et la distance des contacts intermoléculaires et notamment des liaisons H. Près d'un tiers des contacts entre les aminoglycosides et les ARN mettent en jeu des molécules d'eau[80] ce qui a conduit à développer des outils de docking prenant en compte ces molécules d'eau[108].

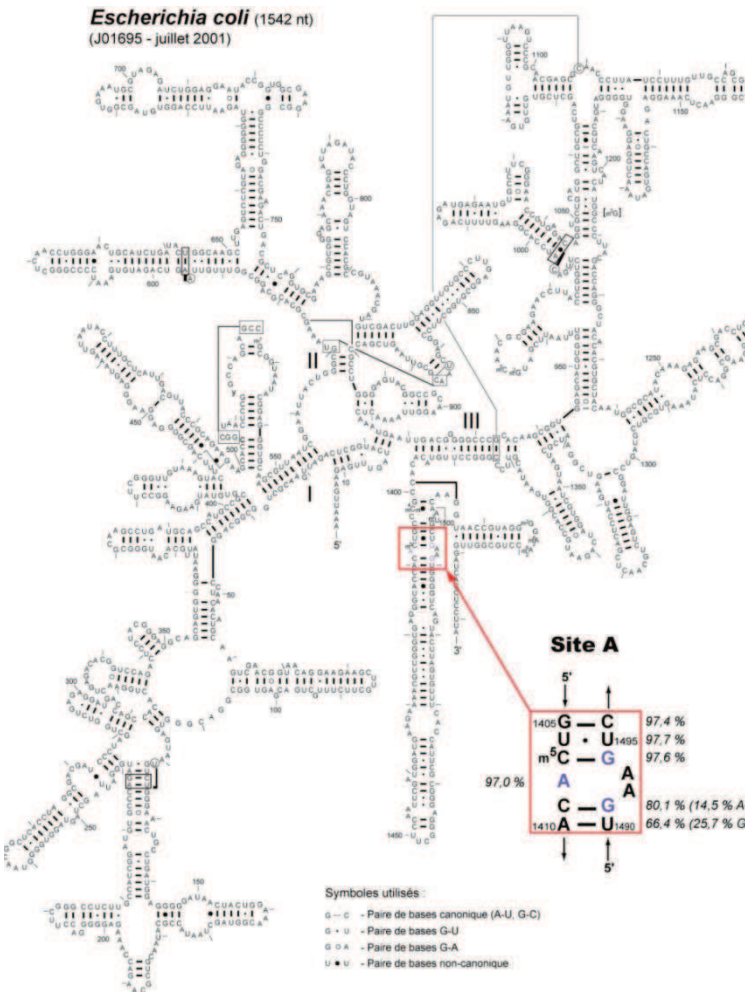
On voit ici l'intérêt qu'il y aurait à pouvoir contrôler précisément cette géométrie et la répartition dans l'espace des différentes fonctions de ces ligands universels d'ARN. Contrôler cette architecture permettrait en effet, dans un premier temps, de mieux comprendre les éléments essentiels à la reconnaissance spécifique d'un ligand par sa cible, puis, dans un second temps, de concevoir des ligands plus efficaces et sélectifs d'une cible particulière[14].

## *D - Activité antibiotique des aminoglycosides*

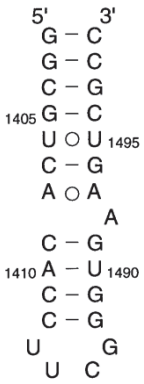
### 1) Interaction avec le ribosome

Comme souvent en médecine, l'utilisation thérapeutique des aminoglycosides a précédé de plusieurs dizaines d'années l'identification de leur cible biologique et la compréhension de leur mode d'action. Si on a assez tôt l'intuition que les aminoglycosides interfèrent avec la fiabilité du décodage de l'information génétique lors de la synthèse protéique[109], ce n'est qu'avec les expériences d'empreinte chimique de Maozed et Noller que l'on comprend que les aminoglycosides se lient sur la sous-unité 30S du ribosome bactérien et plus particulièrement sur l'ARN 16S[110] (voir figure I-17). Les

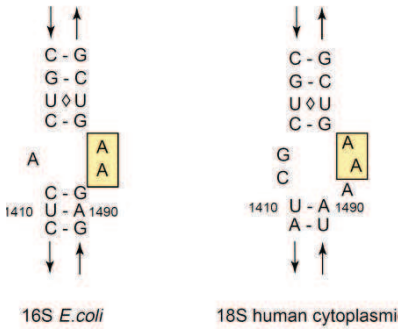
travaux de Yoshisawa, Puglisi et Fourmy permirent de préciser le mode d'interaction des aminoglycosides avec l'ARN 16S en adoptant une approche réductionniste. A l'aide d'expériences RMN, ils montrèrent en effet qu'un petit fragment de 27 nucléotides mimant le site de décodage de l'ARN 16S (site A) était capable d'expliquer la liaison des aminoglycosides avec le ribosome entier[111,112]. La RMN s'est avérée être un outil très puissant pour étudier les changements conformationels induits par la fixation des aminoglycosides sur le site de décodage de l'ARN 16S[45]. Lors d'expériences très astucieuses combinant modifications chimiques ou mutagenèse du site de décodage et purification par affinité à l'aide d'un système biotine-streptavidine, les auteurs ont par ailleurs identifié les bases nécessaires à la reconnaissance et à la fixation des ARNt au site A de l'ARN 16S [113]. Ces informations furent plus tard confirmées par la publication par Ramakrishnan en 2000 de la structure de la sous-unité 30S du ribosome de *Thermus thermophilus,* seule[41] ou en complexe avec divers aminoglycosides [105]. Les travaux cristallographiques du groupe de Westhof contribuèrent de même à mieux comprendre les interactions entre les aminoglycosides et le site A de l'ARNr 16S, ainsi que le géométrie du complexe formé[114,115]. Suite à ces travaux, l'interaction des aminoglycosides avec l'ARN 16S a pu être étudiée par d'autres techniques biophysiques telles la spectrométrie de masse[116] ou la résonnance de plasmon de surface (SPR)[117]. Les perturbations de la dynamique conformationelle observées par RMN ont de même été confirmées par de nombreuses études faisant intervenir d'autres méthodes expérimentales comme la microcalorimétrie[118] ou la fluorescence[119,120].




L'ARN 16S d'*E.coli*



Modèle du site A utilisé par Puglisi *et al.*[44,111]




Comparaison des sites A bactérien et humain[74]

**Figure I-17 L'ARN 16S d'*E.coli***
(source thèse Q. Vicens (2002) d'après le site « comparative RNA »)
(http://www.rna.icmb.utexas.edu/).
Le site A est encadré en rouge. Les pourcentages de covariation associant chaque paire de bases aux deux adénines A1492 et A1493 sont indiquées de part et d'autre du cadre rouge.

Toutes ces observations ont conduit la communauté scientifique à proposer un mécanisme d'interaction entre les aminoglycosides et le site A de décodage de la sous unité 30S du ribosome bactérien.

Durant la synthèse protéique, le ribosome décode l'information portée par l'ARNm à travers la sous-unité 30S avant de catalyser, à l'aide de la sous-unité 50S, l'incorporation séquentielle des acides aminés dans la chaine polypeptidique en croissance. La traduction est un processus très fidèle.

On estime en effet la fréquence d'erreur dans l'incorporation des acides aminés à environ 0.03 % en moyenne chez E.coli[121], et sans doute encore moins chez l'homme[2]. Ceci est rendu possible par la capacité du ribosome à discriminer de façon très fine les changements conformationels du site de décodage de l'ARN 16S lors de la fixation de l'ARNt adapté.

Lors de la reconnaissance du triplet codon de l'ARNm par l'anticodon de l'ARNt, deux résidus adénine conformationellement flexibles (A1492 et A1493), formant une hernie avec l'adénine A1408 dans la structure en hélice du site A, adoptent une conformation extra-hélicoïdale pour interagir avec le petit sillon de l'hélice formée par les sites codon et anticodon. Ces bases permettent de vérifier qu'il y a bien appariement entre les bases du codon et celles de l'anticodon, à travers des interactions de Watson Crick ou des interactions bancales (wobble) dans le cas de la dernière base du codon. Même si le mécanisme de reconnaissance entre l'ARNr et le complexe codon/anticodon est sans doute plus compliqué qu'une paire de liaisons H entre les deux adénines A1492 et A1493 et le squelette de l'ARNm, ces adénines semblent permettre de vérifier la géométrie globale de l'hélice codon/anticodon et de répercuter cette information en diminuant le $k_{off}$ de l'ARNt correctement apparié, induisant des changements conformationels qui sont transmis dans la sous-unité 50S. En l'absence d'interaction spécifique codon/anticodon, il n'y a ainsi pas de passage à la conformation dite « ouverte » du site de décodage où les adénines viennent interagir avec l'hélice codon/anticodon et plus particulièrement avec deux alcools du squelette de l'ARNm. La formation de la liaison peptidique est alors bloquée (voir figure I-18).

[2] Le site Bionumbers (www.bionumbers.hms.harvard.edu) est une compilation intéressante d'un certain nombre de grandeurs caractéristiques de phénomènes biologiques. On y trouve de nombreux taux d'erreur traductionelle pour divers organismes.

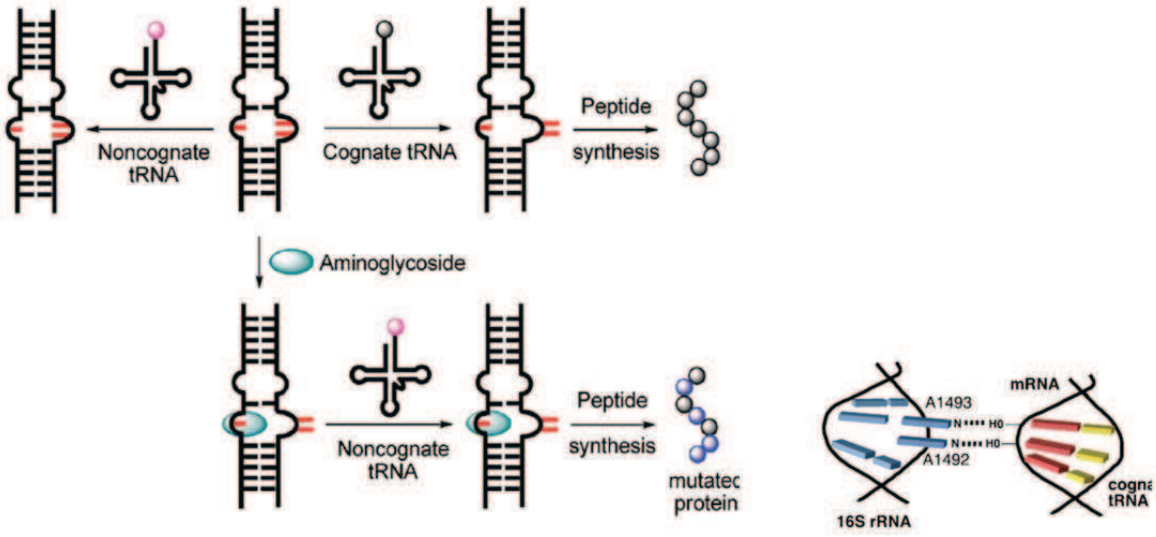


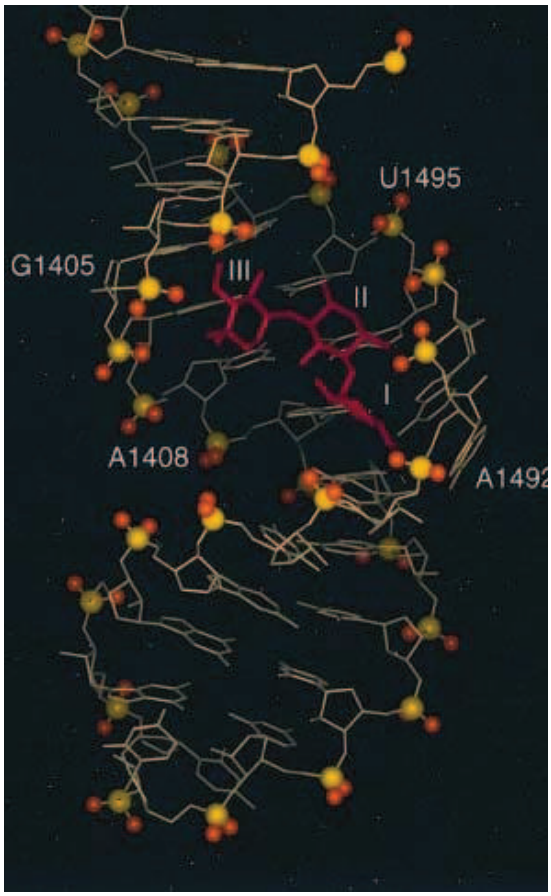

**Figure I-18 Représentation schématique du mécanisme d'action biologique des aminoglycosides.** L'élongation de la chaine polypeptidique se produit quand l'ARNt correct se lie sur le site A de l'ARN 16S, conduisant les adénines 1492 et 1493 (en rouge) à adopter une conformation extra-hélicoïdale. Les aminoglycosides induisent le même changement conformationel indépendement de l'ARNt lié ce qui amène des erreurs dans l'incorporation des acides aminés[86,113].

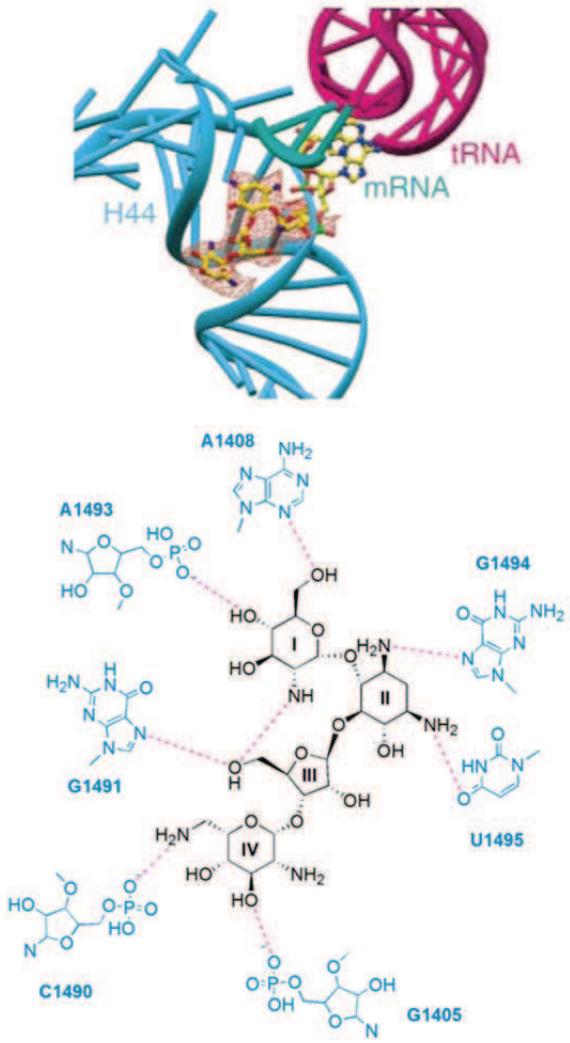
En venant se fixer sur le site de décodage, les aminoglycosides vont stabiliser cette conformation « ouverte », même si un ARNt non adapté vient se fixer en regard du codon de l'ARNm. Les aminoglycosides vont ainsi diminuer l'énergie d'activation nécessaire à l'incorporation à la chaine polypeptidique d'un acide aminé non codé. Le ribosome perd sa capacité à lire correctement le code génétique car la précision de la reconnaissance moléculaire entre codon et anticodon ne peut pas être vérifiée[113,122]. L'augmentation drastique du nombre d'erreurs de traduction qui en résulte conduit à la production de protéines aberrantes et non fonctionnelles, dont l'accumulation induit à terme à la mort cellulaire.

D'un point de vue structural, il a été montré que l'essentiel des interactions des aminoglycosides avec le site A de l'ARN 16S se faisait à travers le cycle central DOS et le cycle I de type glucosamine[88,111]. Les aminoglycosides se lient dans le grand sillon du site de décodage et ont tendance à adopter une conformation coudée.

La DOS interagit avec les paires de bases situées à l'intérieur du grand sillon du site A de l'hélice (hélice 44) par l'intermédiaire des deux fonctions amines en position 1 et 3, alors que le substituant situé en position 4 du cycle I vient se positionner au sein d'une poche formée par la paire non

canonique, A1408 et A1493, et l'adénine non appariée A1492. Ces éléments structuraux semblent suffisant pour assurer la spécificité pour le site A. Les sucres situés en position 5 ou 6 de la DOS, en formant des interactions supplémentaires, contribuent à augmenter l'affinité avec la cible (figure I-19).




Structure par RMN du complexe Gentamicine/site A de l'ARN 16S [104]

Structure par cristallographie des rayons X du complexe Paromomycine/site A [105]

**Figure I-19 Exemple d'interaction d'aminoglycosides avec le site A de l'ARN 16S**

Au niveau biologique, il est intéressant de constater qu'un certain nombre de résidus clés, comme les deux adénines A1492 et A1493, semblent très largement conservés chez les procaryotes et les eucaryotes (voir la figure I-17). Même si l'affinité des aminoglycosides pour le site de décodage

humain est légèrement plus faible que pour celui des bactéries, elle explique en partie la toxicité potentielle de ces molécules lors de traitements prolongés chez l'homme. Le différentiel de toxicité entre procaryotes et eucaryotes s'explique néanmoins avant tout par une différence de biodisponibilité. Les aminoglycosides, fortement polycationiques, ne peuvent pas facilement traverser les membranes cellulaires lipophiles, même si on a montré qu'ils pouvaient interagir avec beaucoup d'éléments de celles-ci (acides phospholipidiques ou protéines transmembranaires par exemple[123,124]. Chez les procaryotes, la diffusion passive est sans doute doublée d'un mécanisme de transport actif efficace qui conduit à l'augmentation drastique de la concentration intracellulaire du principe actif, ce qui augmente sa puissance.

Au vu des informations structurales dont on dispose s'agissant des interactions entre les aminoglycosides et le ribosome, on peut envisager de modifier légèrement la structure de ces antibiotiques pour influencer leurs propriétés biologiques et notamment leur biodisponibilité, leur toxicité ou leur capacité à échapper aux phénomènes de résistance.

### 2) Résistance aux aminoglycosides

Comme c'est la cas pour toutes les molécules antibiotiques, il existe une grande diversité de mécanismes de résistance aux aminoglycosides[125,126,94]. On peut en distinguer trois principaux: la modification de la cible (le ribosome), la réduction de la concentration cytoplasmique en antibiotique et enfin l'inactivation de l'antibiotique par des enzymes de résistance (figure I-20).

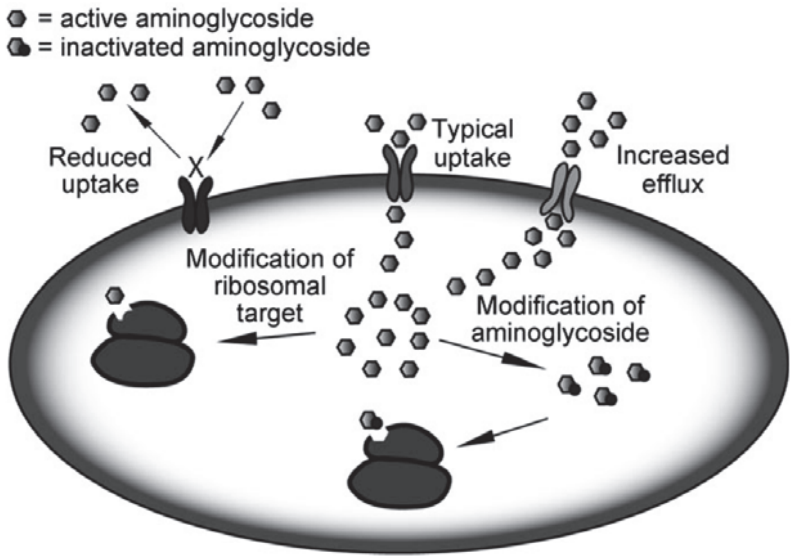



**Figure I-20 Les principaux mécanismes de résistance aux aminoglycosides** (adaptée de [97])

La résistance aux aminoglycosides peut être acquise par modification enzymatique du site A du ribosome de la bactérie. C'est le cas notamment des bactéries naturellement productrices d'aminoglycosides qui ont la particularité d'exprimer des méthylases capables de méthyler l'ARN 16S à des positions critiques pour la fixation des aminoglycosides[127]. On trouve aussi des mutations ponctuelles de l'ARN ribosomique conférant d'importants niveaux de résistance, une seule mutation suffisant parfois à supprimer quasi totalement l'effet des aminoglycosides (par exemple, la transformation de l'adénine A1408 en guanine)[128]. En ce qui concerne la réduction de la concentration de l'antibiotique, on a souvent affaire à des mécanismes non spécifiques qui conduisent à la diminution de la perméabilité des membranes ou à l'augmentation des phénomènes d'efflux[126].

Le mécanisme de résistance bactérien le plus fréquent est néanmoins la modification directe des aminoglycosides par voie enzymatique. On trouve de nombreuses enzymes de résistance, qui vont acétyler, phosphoryler ou transférer un nucléotide sur certaines des fonctions nécessaires à l'interaction de la molécule avec sa cible ARN[90]. En guise d'exemple, on peut citer la *N*-acétyl transférase AAC(6')-Ib11, initialement décrite par Cassin et al en 2003[129], qui inactive presque tous les aminoglycosides connus et confère donc aux bactéries un spectre de résistance particulièrement large (figure I-21). Le gène codant pour cette enzyme étant porté par un intégron mobile, la menace du développement de souches résistantes est particulièrement grande, ce qui a conduit notre équipe à développer une approche originale pour la recherche d'analogues d'aminoglycosides inhibiteurs de cette enzyme[130].

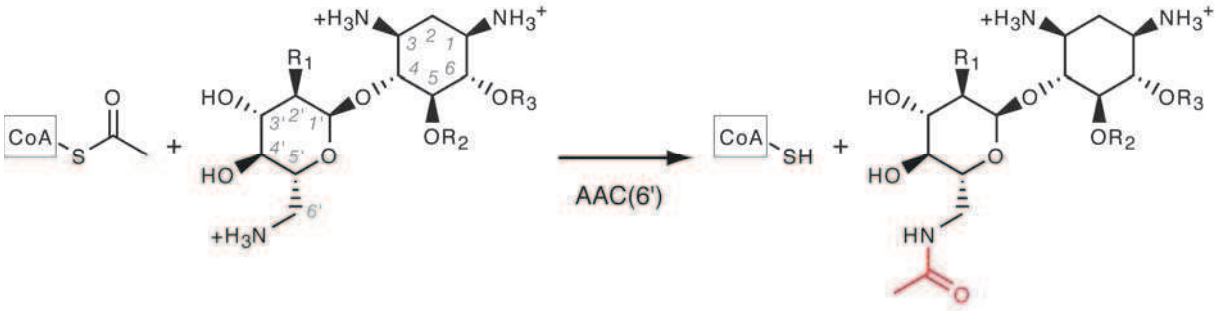



**Figure I-21 Inactivation des aminoglycosides par l'enzyme de résistance AAC(6')-Ib11**

### 3) Conclusion

En conclusion les aminoglycosides comme tous les antibiotiques sont sujets à des stratégies de résistance bactérienne dont l'émergence est un enjeu majeur de santé publique. Par ailleurs, ils pâtissent aussi d'une faible biodisponibilité et d'une spécificité modeste pour le ribosome bactérien.

Une meilleure compréhension des interactions des aminoglycosides avec les ARN ribosomaux humains et bactériens est indispensable pour tenter de dépasser ces limitations. Le développement de nouveaux analogues aminoglycosidiques est une voie essentielle de l'amélioration potentielle des traitements disponibles, ce qui justifie un effort de recherche important pour développer de nouvelles stratégies de synthèse de tels composés.

## *E - Autres cibles biologiques des aminoglycosides*

Bien que la communauté scientifique se soit essentiellement concentrée sur l'activité antibiotique des aminoglycosides, le ribosome bactérien est loin d'être la seule cible de ces très bons ligands d'ARN. On découvre aujourd'hui de nombreuses autres cibles thérapeutiques potentiellement visées par les aminoglycosides. Sans chercher à en faire ici un exposé exhaustif, nous allons présenter dans les quelques lignes qui suivent des exemples de ces interactions et leur potentiel thérapeutique.

### 1) Traitement de désordres génétiques

En raison de leur capacité à interagir spécifiquement avec les ARN, les aminoglycosides peuvent être utilisés en tant qu'outils pour potentiellement traiter les désordres génétiques au niveau transcriptionel. Deux stratégies mettent en jeu les aminoglycosides pour le traitement de désordres génétiques : la translecture de mutations non-sens ou la fixation compétitive sur des motifs d'ARN répétitifs. Pour avoir une vue plus générale sur l'utilisation de petites molécules, et en particulier des aminoglycosides, dans le traitement de maladies génétiques, voir les récentes revues de Dietz[131] et de Clancy[132].

#### a) Translecture

Plus de 20 ans après la découverte de la streptomycine, on a découvert qu'en raison de sa capacité à altérer la fidélité de traduction du ribosome, elle pouvait induire une translecture ( « read-through ») des codons stop, en particulier des codons stop dits prématurés (premature codon terminaison, PTC), c'est-à-dire apparaissant par mutation au milieu d'une séquence codante[133]. Les auteurs ont ainsi constaté qu'en présence d'une faible concentration d'aminoglycosides, il était possible de restaurer chez *E.coli* un phénotype lié à la production d'une protéine entière en dépit de la présence d'un codon stop au milieu de la séquence codante. Cette expérience initiale fut suivie plus de 20 ans après par la preuve que cette propriété des aminoglycosides pouvait aussi être utilisée

*in vivo* dans des cellules de primates pour supprimer le phénotype d'une mutation non-sens[134]. (figure I-22)

Chez l'homme, plus de 1800 maladies génétiques sont liées à des mutations causant l'apparition de codons stop prématurés. On peut citer par exemple la mucoviscidose, due dans certains cas à la production d'un canal ionique de transport transmembranaire du chlore tronqué non fonctionnel. De récents essais en laboratoire ont montré l'intérêt d'utiliser certains aminoglycosides pour améliorer le phénotype des patients atteints. En particulier, la gentamicine a permis d'améliorer significativement le phénotype de cellules de patients cultivées en laboratoire[135]. Des essais cliniques exploratoires ont plus tard confirmé l'intérêt thérapeutique de cette démarche[136].

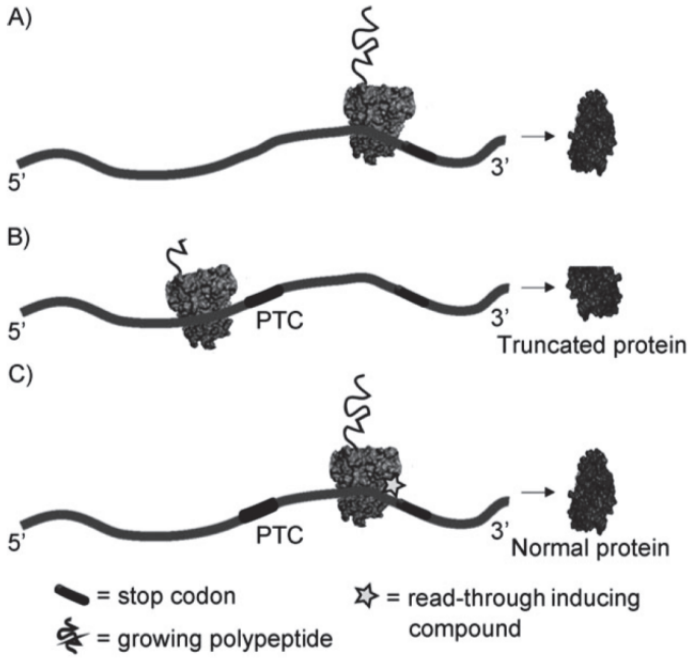



**Figure I-22 Principe général de la translecture induite par les aminoglycosides**[97]
Comparaison d'une traduction normale (A), d'une traduction interrompue par un codon stop prématuré (PTC) qui donne lieu à la production d'une enzyme non fonctionnelle (B) et d'une traduction où le ribosome peut lire à travers le codon stop prématuré grâce à l'emploi d'un aminoglycoside ce qui conduit à la restauration de la production d'une protéine entière (C).

De nombreux groupes de recherche essayent aujourd'hui d'exploiter la propriété des aminoglycosides à induire la translecture de mutations stop comme traitement potentiel de la mucoviscidose mais aussi d'autres maladies dues à des codons stop prématurés. Pour explorer ces possibilités thérapeutiques, il faut être en mesure de comprendre ce qui fait la spécificité d'un ligand de type aminoglycoside pour sa cible ARN. Par ailleurs certains groupes cherchent à créer des

analogues d'aminoglycosides ayant une moindre toxicité et une plus grande efficacité. On peut citer par exemple le groupe de Baasov qui s'est intéressé au traitement potentiel de la maladie de Duchenne ou du syndrome de Hurler [137] à l'aide d'analogues de ribostamycine. En dépit de l'excitation causée à juste titre par ces thérapies potentielles pertinentes pour de nombreuses maladies génétiques humaines, certains problèmes restent à résoudre pour en faire une réalité[131] (il faudrait par exemple augmenter la spécificité des molécules utilisées pour diminuer leur toxicité ou bien concevoir des molécules qui permettraient de diminuer la cinétique de dégradation des ARNm contenant des mutations non sens).

**b) Motifs d'ARN répétés**

Une autre grande famille de maladie génétique est liée à la présence de larges répétitions de petits motifs d'ARN. Ces répétitions sont toxiques car elles vont séquestrer un certain nombre de protéines partenaires essentielles au bon fonctionnement de la cellule. C'est le cas par exemple de la dystrophie myotonique de Steinert, caractérisée par la présence d'un très grand nombre de triplets CTG dans la région amont (3'-UTR) de l'ARNm du gène DMPK[138]. Une grande répétition de ce triplet conduit à la formation d'une tige boucle possédant de nombreuses boucles internes régulièrement espacées, qui constituent des sites de haute affinité pour un régulateur d'épissage nommé MBNL1[139]. La séquestration de celui-ci induit de nombreux dysfonctionnements cellulaires[140]. Une stratégie thérapeutique envisagée est donc d'entrer en compétition avec MBNL1 pour les sites de la tige boucle (voir figure I-23). On peut alors envisager d'exploiter la capacité des aminoglycosides à reconnaitre spécifiquement des poches de partenaires ARN pour concevoir des ligands compétiteurs à l'image de ce qu'a fait le groupe de Disney en 2009[141,142]. En développant des polymères de néamine et de kanamycine, aminoglycosides connus pour se fixer préférentiellement sur des boucles internes de 4 à 6 nucléotides riches en pyrimidines, les auteurs ont pu créer des ligands affins compétiteur de MBNL1 et ont ainsi montré la validité de ce type d'approche.

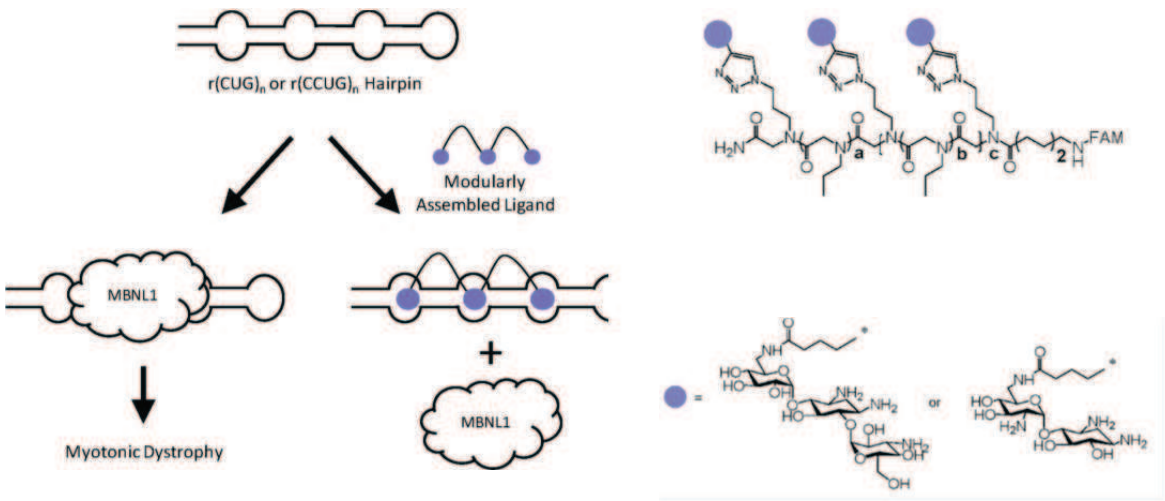



**Figure I-23 Principe général de traitement de la dystrophie myotonique avec un ligand d'ARN**[142]

Ces résultats prometteurs nécessitent de développer des méthodes de vectorisation permettant d'apporter *in vivo* ces molécules à leur cible et de contrôler leur toxicité.

### 2) Les aminoglycosides comme molécules antivirales

L'application des aminoglycosides dans le traitement des infections virales est un domaine de recherche particulièrement dynamique en raison des implications sous-jacentes en santé publique.

De nombreux groupes de recherche se penchent sur les infections causées par le virus du VIH ou de l'hépatite C. Ces rétrovirus doivent, pour se répliquer après infection, rétro-transcrire leur ARN génomique, l'intégrer à l'ADN de l'hôte avant de détourner la machinerie cellulaire de celui-ci pour produire des copies du virus. Ces mécanismes sont finement régulés au niveau transcriptionel et traductionel par des séquences de l'ARN viral. Pouvoir développer des petites molécules venant interférer avec ces mécanismes en ciblant spécifiquement l'ARN viral pourrait constituer une voie thérapeutique viable.

Dans le cas du VIH, de nombreux travaux ont montré que les aminoglycosides pouvaient interférer avec certaines des étapes clé de la maturation du virus (voir les revues[143] et [97]). En particulier, après la rétrotranscription et une fois l'ADN viral intégré à celui de son hôte, le cycle du virus nécessite la transcription du génome viral. Or l'ARN polymérase humaine ne transcrit que très mal ces séquences. La protéine virale Tat sert alors d' « adaptateur » à la polymérase humaine pour faciliter la transcription de ces séquences virales. Pour ce faire, elle interagit avec une zone très

structurée de l'ARNm viral nommée TAR (pour "trans-activating response" element), une tige-boucle de 45 nucléotides comportant une hernie située en amont des séquences codantes. L'interaction Tat-TAR est une étape essentielle du cycle viral du VIH : elle permet en effet de multiplier plusieurs centaines de fois le taux de transcription du génome viral et assure que les transcrits ne sont pas tronqués[144]. Ce rôle central joué par l'activateur de transcription TAR a stimulé les recherches d'inhibiteurs de l'interaction Tat-TAR. En 1997, à la suite d'un screening d'une librairie d'aminoglycosides, l'équipe de Czarnic démontre que la néomycine est capable d'interférer avec la formation du complexe Tat-TAR avec une $IC_{50}$ submicromolaire[145]. Cette découverte a initié la recherche d'analogues potentiellement plus affins et spécifiques, projet qui s'est avéré assez compliqué en raison de la grande flexibilité du fragment d'ARN TAR(voir par exemple le travail de Tor[146,147]). Si la cible biologique TAR est désormais validée, un effort de recherche important reste à faire pour concevoir des analogues d'aminoglycosides pouvant être utilisés en thérapeutique.

Une autre étape importante du cycle viral du VIH peut faire l'objet d'une inhibition par les aminoglycosides. En effet, après transcription du génome viral, l'ARN du HIV doit être correctement exporté du noyau pour être épissé par la machinerie cellulaire. La protéine virale Rev facilite ces étapes en se fixant sur le segment d'ARN RRE (pour "Rev response element"). Lors de cette fixation, Rev polymérise ce qui a deux conséquence : la protection de l'ARN viral vis-à-vis de phénomènes de dégradation et l'exposition de la partie C-terminale de la protéine qui contient un signal d'export vers le cytoplasme[148] [143]. Il a été suggéré qu'empêcher l'interaction Rev-RRE pourrait bloquer la maturation du virus HIV. De façon très intéressante, l'équipe de Green a montré que la néomycine était capable d'inhiber *in vitro* cette interaction[149] avec un Ki de l'ordre du micromolaire[102]. Dans le même article, ces auteurs démontrent l'activité de la néomycine *in vivo* sur des cellules infectées. Cette étude fut suivie par de nombreuses études structurales[150] et biochimiques[151] visant à comprendre plus précisément l'interaction de la néomycine avec l'ARN RRE en vue de créer des analogues plus efficaces. De nombreux projets de recherche sont actuellement en cours pour exploiter cette possibilité.

Le virus de l'hépatite C dépend lui aussi du détournement de la machinerie cellulaire de son hôte pour sa réplication. Par exemple le virus va détourner le ribosome humain pour traduire la polyprotéine virale qui donnera, après clivage par des protéases cellulaires, l'ensemble des protéines nécessaires à la maturation du virus. Pour parvenir à ses fins, le virus vient « capturer » la sous-unité 40S du ribosome eucaryote à l'aide d'une séquence de son ARN, l'IRES (pour "internal ribosome entry site") située en amont des séquences codantes. Cette interaction va permettre l'assemblage du ribosome entier et initier la traduction du génome viral. Le domaine IRES adopte une conformation

en L stabilisée par des ions magnésium[152] qui est essentielle à la reconnaissance de la sous-unité 40S[153]. Modifier cette conformation ou, de façon plus générale, empêcher l'interaction entre l'IRES et le ribosome pourrait constituer une voie thérapeutique intéressante à explorer. En 2005, l'équipe de Seth chez Ibis Therapeutics a développé une stratégie de screening haut débit par spectrométrie de masse pour trouver des ligands de l'IRES du virus de l'hépatite C. Elle a trouvé une famille de ligands du type benzimidazoles[154] assez originale. Il a plus tard été démontré que ces molécules étaient capables d'interférer avec le démarrage de la traduction de l'ARN avec un $EC_{50}$ de l'ordre du micromolaire[155]. Des analogues basés sur la désoxystreptamine, le cycle central des aminoglycosides, développés par la suite dans l'équipe d'Hermann, possèdent les mêmes propriétés et interfèrent avec la reconnaissance IRES/ribosome en induisant un changement conformationel de l'IRES[156].

Même si l'étude de cette cible thérapeutique n'en est qu'à ses débuts, le développement de nouveaux analogues combinant les pharmacophores des deux types de ligands identifiés pourrait être une piste très intéressante pour étudier plus précisément la biologie des IRES qui reste globalement faiblement comprise.

## *F - Conclusion : l'intérêt de développer des analogues d'aminoglycosides*

Au terme de ce bref survol, il est clair que la propriété des aminoglycosides à se lier à des cibles ARN permet potentiellement de cibler de multiples cibles biologiques. Le ribosome bien sûr, qui reste une des cibles principales de cette famille de molécules, mais aussi un certain nombre d'acteurs clé de la maturation de virus comme le VIH ou le virus de l'hépatite C. Les aminoglycosides peuvent d'autre part contribuer à apporter une solution thérapeutique dans un certain nombre de maladies génétiques comme, par exemple, la mucoviscidose. Nous avons choisi de nous limiter à quelques exemples représentatifs des cibles biologiques des aminoglycosides mais il existe de nombreux développements récents très intéressants qui démontrent une fois de plus, l'intérêt de ces structures pseudopolysaccharidiques (voir par exemple l'inhibition de l'activité catalytique de ribozymes par les aminoglycosides[157,158]).

Dans tous ces cas, la versatilité des aminoglycosides, « ligands universels d'ARN », en fait de très bons points de départ pour développer des ligands d'ARN à même d'interférer spécifiquement avec des cibles biologiques très diverses. Même si ils ne sont pas directement utilisés pour leurs vertus potentiellement thérapeutiques, les aminoglycosides constituent des outils de choix pour étudier le rôle biologique des ARN et comprendre comment interagir spécifiquement avec ces biomolécules complexes, fortement chargées et très flexibles. Pour pouvoir répondre à ces enjeux, et pour pouvoir

concevoir des molécules efficaces en dépassant les éventuels problèmes de toxicité ou de résistance, il importe d'être capable de pouvoir altérer l'architecture et les fonctions chimiques des aminoglycosides afin d'explorer un espace chimique le plus large possible.

Les différentes stratégies mises en place pour concevoir de nouveaux analogues d'aminoglycosides seront illustrées par quelques exemples dans la partie suivante. Cela nous amènera à exposer un des objectifs de cette thèse qui est de proposer une nouvelle stratégie dans la synthèse d'analogues d'aminoglycosides, l'approche mutasynthétique.

# III - Stratégies de recherche d'analogues d'aminoglycosides

Poussée par le fort intérêt qu'il y a à trouver de nouveaux analogues d'aminoglycosides, la communauté scientifique a exploré de nombreuses stratégies différentes pour développer des composés basés sur l'architecture de ces molécules complexes. Sans chercher l'exhaustivité, nous allons ici survoler certaines de ces stratégies, qu'elles exploitent la richesse de la biodiversité microbienne, l'efficacité des systèmes enzymatiques ou la créativité et la flexibilité des techniques de synthèse organique moderne. Dans la partie I-III-C, à présenter une stratégie originale pour la synthèse d'analogues d'aminoglycosides, visant à combiner certaines de ces caractéristiques.

## *A - Exploration de la biodiversité microbienne*

### 1) Nouvelles méthodes d'isolement et de fermentation de souches bactériennes

Historiquement, pour trouver de nouveaux aminoglycosides, les chercheurs se sont tournés vers la source des premiers aminoglycosides connus, les bactéries de la famille des actinomycètes. En explorant des biotopes semblables à ceux qui avaient permis d'isoler les premières souches bactériennes productrices, de nombreuses molécules de la famille des aminoglycosides ont été découvertes à l'aide d'analyse et de purification systématiques de l'activité antibactérienne d'extraits des souches isolées. En l'espace d'une vingtaine d'année, on a découvert de cette façon la grande majorité des aminoglycosides connus à ce jour, naturellement produits par des bactéries des familles des *Streptomyces*, *Micronomospora* ou *Bacillus*. Ainsi, la découverte de la streptomycine en 1943 par Waksman[91] fut rapidement suivie par celle de la néomycine (1949)[159], de la kanamycine (1957)[160], puis de nombreuses autres molécules de la même famille (paromomycine (1959), gentamicine (1963), tobramycine (1968), … pour n'en citer que quelques-unes). Durant l' « âge d'or » de la recherche sur les aminoglycosides, entre les années 1950 et 1970, on trouve de nombreuses nouvelles molécules. Ces architectures complexes et très fonctionnalisées sont de plus facilement produites à grande échelle par fermentation des souches productrices.

Cette méthode d'exploitation de la diversité des produits naturels a initialement été extrêmement efficace, dans un domaine bien plus large que celui des seuls aminoglycosides. Pendant plus de 40 ans, les produits naturels ont constitué le cœur de l'industrie pharmaceutique[42] et, en dépit du déclin annoncé de la chimie des substances naturelles, l'essentiel des nouvelles molécules mises sur le marché restent encore à ce jour directement inspirées de produits naturels. Pour donner

un ordre de grandeur, plus de 70 % des petites molécules mises sur le marché dans le domaine des anti-infectieux entre 1981 et 2010 sont des produits soit naturels, soit issus d'hémisynthèse à partir de produits naturels, soit largement inspirées d'un pharmacophore de produit naturel[43]. Dans le domaine des anti-cancéreux, la proportion monte même à presque 80 %[161].

D'un point de vue évolutif, il n'est pas étonnant que les produits naturels constituent d'intéressants points de départ de composés bioactifs, surtout dans le domaine des antibiotiques. En effet, pendant des millions d'années, les microorganismes se sont battus avec des compétiteurs bactériens pour le contrôle de ressources limitées et les molécules antibiotiques ont été des armes vitales dans cette lutte. D'une certaine façon, les architectures moléculaires de ces molécules ont été « prévalidées » par l'évolution pour leur capacité à se lier à des cibles biologiques d'intérêt[162]. Même si le rôle des métabolites secondaires n'est pas toujours clair, qu'ils soient utilisés pour la prédation, la défense, ou bien à des fins de communication, ils représentent une mine de molécules bioactives et constituent donc une source d'inspiration intéressante pour les chimistes[163]. En particulier, les aminoglycosides semblent avoir une architecture optimisée pour venir se fixer, *in vitro* mais surtout *in vivo*, à différentes structures d'ARN comme par exemple le site de décodage du ribosome bactérien.

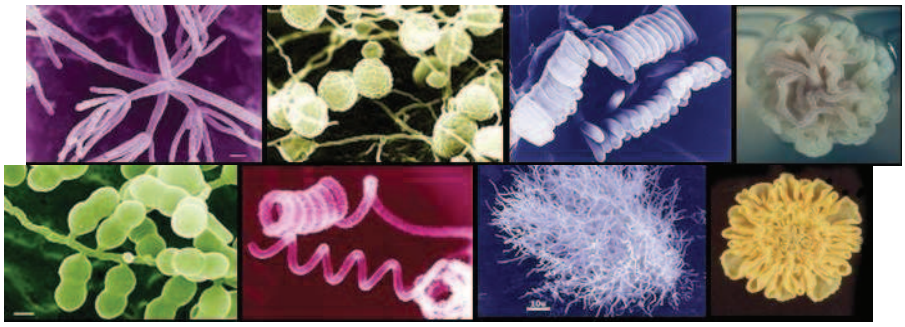


**Figure I-24 Un aperçu de la biodiversité des actinomycètes**
(source Pr. Mervyn Bibb, http://www.jic.ac.uk/staff/mervyn-bibb/)

Dès les années 70, d'inquiétants phénomènes de résistance aux aminoglycosides commencent à apparaitre ce qui induit la recherche de nouveaux analogues [93]. C'est à ce moment que les limites des stratégies de criblage « traditionnelles » deviennent apparentes. Elles ne permettent en effet que de détecter les antibiotiques fréquemment produits et en grande quantité. Ce n'est pas un hasard si la streptomycine, premier aminoglycoside découvert, est aussi un des plus fréquemment produit (près de 1% des actinomycètes du sol en sont des producteurs naturels)[164].

Lorsqu'on cherche à découvrir de nouveaux aminoglycosides, plus rarement produits, il faut être capable d'échantillonner une grande variété de souches, préférentiellement dans des biotopes

encore peu explorés comme par exemple les déserts ou bien les fonds marins [165]. Pour ce faire, de nouvelles stratégies d'isolement de souches et d'analyse haut-débit ont été développées. On peut par exemple citer les travaux de Baltz dans l'entreprise Cubist Pharmaceuticals. En développant une méthode de criblage haut débit combinant robotique et fermentations de multiples souches en parallèle dans des microbilles d'alginate, il a été possible de tester plus de 10 millions de souches différentes par an. Ceci a permis de découvrir des molécules à propriétés antibiotiques produites par les actinomycètes avec des fréquences très faibles (moins d'une bactérie testée sur $10^7$)[166]. Une autre solution est d'utiliser des plaques 96 puits et de cultiver les souches dans une grande variété de milieux nutritifs différents pour tester la production d'antibiotiques[167]. Dans ces expériences, pour éviter de retomber toujours sur des molécules déjà identifiées qui constitueraient des faux positifs, les tests phénotypiques des milieux de culture peuvent être effectués sur des souches résistantes aux antibiotiques usuels.

Ces approches d'exploration semblent prometteuses au vu de notre connaissance de la biodiversité microbienne pour le moment très parcellaire. Une simple poignée de terre contient en effet plus d'une dizaine de milliards de microorganismes dont nous ne connaissons qu'une infime proportion. Selon la technique popularisée par Pace[168], l'analyse par PCR (Polymerase Chain Reaction) des séquences d'ARNr 16S bactériens trouvées dans un seul gramme de terre permet souvent d'identifier plus de 10 000 espèces de bactéries inconnues[169]. Même si le nombre total d'espèces bactériennes est extrêmement difficile à estimer (de $10^7$ à plus de $10^{10}$ selon les estimations[170171]), ce nombre contraste fortement avec les 15000 espèces procaryotes formellement décrites (référence : http://www.bacterio.cict.fr/number.html#total). Ce décalage laisse entrevoir la richesse de biodiversité microbienne qui reste à explorer et les nouveaux aminoglycosides qui pourraient en résulter.

Par ailleurs, en raison de notre faible compréhension de la biologie microbienne, il est estimé que moins de 1 % voire 0.1 % des espèces bactériennes que nous pouvons analyser par microscopie ou par PCR, peuvent effectivement être cultivées en laboratoire [172,173]. Même dans le cas favorable où nous savons cultiver une souche bactérienne, encore faut-il être dans des conditions où il y a production des métabolites secondaires d'intérêt, c'est-à-dire avoir la conjonction des bons facteurs environnementaux et des bons cofacteurs de biosynthèse, qui malheureusement sont tous deux souvent inconnus... C'est ainsi que de nombreux métabolites ne semblent pas pouvoir être produits en laboratoire en dépit d'une machinerie biosynthétique apparemment fonctionnelle. Un des problèmes fréquemment rencontrés est lié à la tendance de la plupart des bactéries à vivre en « communautés microbiennes » où chaque souche est dépendante des autres souches présentes.

Les pratiques courantes de la microbiologie, qui cherchent à isoler les souches pour les cultiver séparément, ne permettent pas d'avoir accès à cette biodiversité. Ceci a conduit certains groupes à développer de nouvelles méthodes de culture permettant de simuler un environnement naturel[174].

Ces ordres de grandeurs vertigineux nous permettent de voir à quel point notre exploration de la biodiversité microbienne est parcellaire. Nous nous rendons compte aujourd'hui que les aminoglycosides isolés et caractérisés ne représentent sans doute qu'un minuscule échantillon de ce qui existe. Il est ainsi estimé que près de 90 % de la diversité chimique des composés produits par la famille des *Streptomyces* reste à ce jour inexplorée[163,175]. Développer de nouvelles stratégies de criblage et de fermentation de souches issues de biotopes inexplorés est un moyen d'avoir accès à de nouveaux analogues d'aminoglycosides. On peut également souligner l'intérêt qu'il y a à revenir à l'étude des souches décrites dans les années 50 en tirant profit des moyens modernes d'analyse des milieux de cultures comme la chromatographie haute performance (HPLC) couplée à la spectrométrie de masse ou à la résonance magnétique nucléaire. Il y a fort à parier que la finesse d'analyse de ces techniques permettrait de découvrir de nouvelles molécules produites par des souches déjà décrites.

### 2) Exploitation des clusters biosynthétiques cryptiques

Les progrès drastiques des méthodes de séquençage et de la génomique ouvrent d'autres voies d'exploration de la biodiversité microbienne[176,163]. On s'aperçoit aujourd'hui que même les souches bactériennes que l'on pensait bien connaitre recèlent encore de nombreuses surprises.

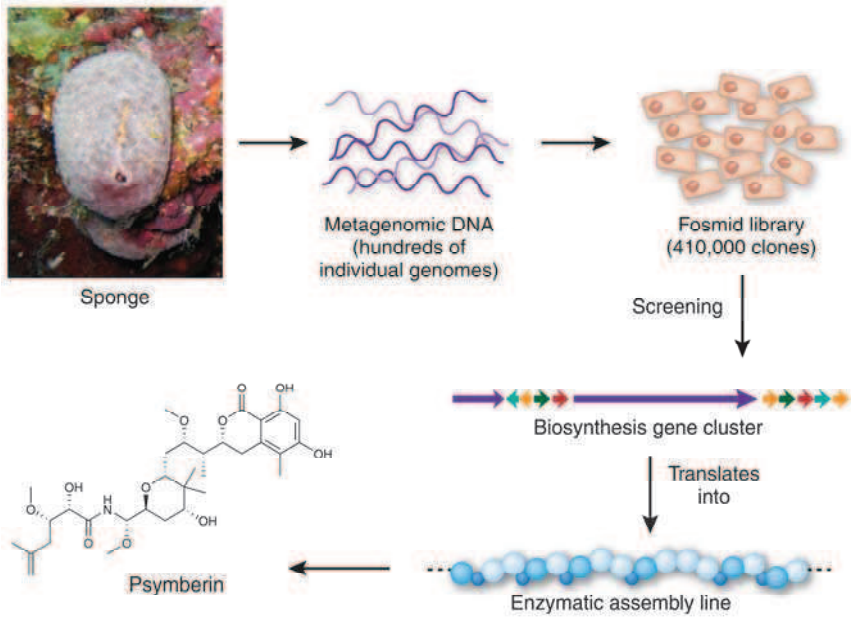
Ainsi, avec le séquençage des souches *S. coelicolor* en 2002[177] et *S. avermitilis* en 2003[178], la communauté scientifique s'est rendue compte que le génome de ces souches contenait beaucoup plus de clusters de biosynthèse de métabolites secondaires que de molécules effectivement identifiées, la majorité de ces clusters, dits « cryptiques », ne semblant pas exprimée dans les conditions expérimentales usuelles[179]. Cette constatation a depuis été confirmée dans de nombreuses autres souches bactériennes (comme par exemple chez *S. erythraea* [180]). La détection de ces clusters a été grandement facilitée par les progrès de la bioinformatique ainsi que par notre compréhension croissante de la génétique et de l'enzymologie des voies de biosynthèse des métabolites secondaires bactériens, notamment de celles des polyketides et des peptides non ribosomaux[181]. Dans les cas favorables, certains outils bioinformatiques permettent même la prédiction des molécules synthétisées et de leurs caractéristiques physicochimiques (voir par exemple le logiciel AntiSMASH (http://antismash.secondarymetabolites.org) développé par l'équipe de Takano[182]).

Une stratégie originale pour trouver de nouveaux analogues d'aminoglycosides consiste à « activer » ces clusters cryptiques, par exemple en changeant leur promoteur ou en le mettant sous le contrôle de régulateurs artificiels qui augmentent leur expression[179,183]. Evidemment, ceci n'est possible qu'à la condition d'avoir des outils génétiques pour modifier directement la souche d'intérêt. Une alternative consiste à exprimer de façon hétérologue le cluster identifié dans une souche adaptée. Pour le moment, cette stratégie a principalement été appliquée avec succès pour des clusters de biosynthèse de terpènes[184], de peptides non ribosomaux[185], de lipopeptides[186] ou de polyketides[187,188]. Son application aux aminoglycosides ne pose a priori aucun problème technique même si notre compréhension des différentes voies de biosynthèse et de leur régulation est pour le moment plus faible.

### 3) Approche métagénomique

Une autre piste très intéressante pour explorer la biodiversité microbienne est la métagénomique[189]. Les progrès des techniques de clonage et de sélection permettent en effet de cloner aléatoirement de larges banques d'ADN, directement isolées d'extraits environnementaux, et/ou de les séquencer avant de procéder à leur analyse fonctionnelle. Par exemple, une expédition scientifique assez médiatisée de Venter en 2007, intitulée « The Sorcerer II Global Ocean Sampling expedition » [190], avait pour but de récolter et séquencer une large bibliothèque d'échantillons d'ADN issus des océans Atlantique et Pacifique en utilisant une technique de séquençage de type « shotgun » analogue à celle utilisée pour le séquençage du génome humain. Ces analyses métagénomiques permettent d'avoir une meilleure vision de la diversité des microorganismes présents dans ces environnements, sans forcément avoir pour autant avoir à les isoler ou les cultiver[191].

On peut aussi potentiellement avoir accès à de nombreux clusters de biosynthèse de souches « non cultivables » que l'on peut essayer d'exprimer de façon hétérologue dans une souche adaptée. Cette méthode reste pour le moment limitée par les types de vecteurs disponibles pour le clonage ainsi que par les possibilités d'analyse parallèle de multiples constructions génétiques. Certains groupes, comme par exemple celui de Piel, ont néanmoins prouvé la validité de ce genre d'approche en développant des stratégies de criblage originales combinant bioinformatique, chimie analytique et génomique afin d'identifier de nouveaux polyketides produit par des actinomycètes vivant en symbiose avec des éponges marines [192]. Dans ce cas précis, l'approche metagénomique a permis de lever un important verrou méthodologique car les souches bactériennes étaient très difficilement isolables et n'étaient pas cultivables en laboratoire (voir figure I-25).




Stratégies de criblage possibles

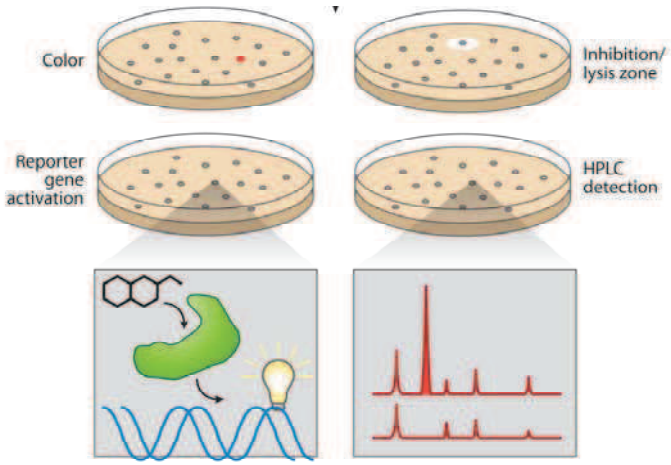



**Figure I-25 Principe de la métagénomique**
(figure adaptée de [189] et [183])

### 4) Approche de biosynthèse combinatoire

L'augmentation de nos connaissances des différentes voies de biosynthèses des aminoglycosides ainsi que l'amélioration des outils d'ingénierie génétique disponibles rendent possible la modification des voies de biosynthèse pour obtenir des produits naturels « non naturels ». Les approches de biosynthèse combinatoire visent à combiner des gènes de différentes voies de biosynthèse pour

concevoir des voies de biosynthèse artificielles, capables de produire des molécules qui seraient difficiles à obtenir en utilisant les approches classiques de chimie organique[193]. On peut effectivement considérer l'ensemble des gènes de biosynthèse identifiés comme une immense « boite à outils » où venir piocher pour constituer des nouvelles voies de biosynthèses qui n'ont pas encore été explorées par l'évolution. Ceci a poussé certains groupes, comme celui de Faulon à Evry, à développer des méthodes d'analyse rétrosynthétique à l'instar de ce que l'on peut faire en chimie organique. Ces programmes permettent de concevoir des voies métaboliques optimales pouvant conduire à la synthèse du produit d'intérêt à partir des bases de données des enzymes biosynthétiques connues[194]. L'intérêt des approches de biosynthèse combinatoire est de permettre une modification profonde du squelette mais aussi des pharmacophores des produits naturels ciblés en reconstruisant dans un microorganisme hôte une voie de biosynthèse artificielle. Créer *de novo* une voie métabolique complète dans un microorganisme permet potentiellement d'éviter de fastidieuses étapes de purification intermédiaire ou la production de produits secondaires qui sont des problèmes inhérents aux stratégies traditionnelles de synthèse organique. Par ailleurs, un des objectifs de la biosynthèse combinatoire peut s'apparenter à celui de la chimie combinatoire : générer de vastes librairies d'analogues à travers la combinaison de divers gènes biosynthétiques, sans forcément nécessiter un design rationnel préexistant.

Pour le moment, ces approches ont été principalement appliquées avec succès aux polyketides synthases (PKS) et aux synthases de peptides non ribosomaux (NRPS) en raison de leur structure modulaire qui facilite grandement l'échange de modules pour former de nouveaux composés non naturels et la prédiction de la structure des composés formés[195,196]. Ces enzymes de biosynthèse de taille impressionnante peuvent en effet être vues comme des chaines d'assemblage où chaque module est en charge d'une modification chimique particulière[197]. Le remplacement, l'insertion ou la suppression d'un module peut se faire sans pour autant changer profondément le fonctionnement de l'ensemble de la voie de biosynthèse. Par ailleurs, afin d'augmenter leur tolérance ou leur efficacité, des stratégies d'évolution dirigée ou d'ingénierie rationnelle des différents modules peuvent être envisagées.

Les stratégies de biosynthèse combinatoire commencent également à être appliquées avec succès à d'autres systèmes biologiques. On peut notamment citer les travaux de Keasling à Berkeley qui a récemment décrit une stratégie d'ingénierie métabolique pour produire de grandes quantités d'acide artémisinique, un précurseur important de l'artemisinine, un anti antipaludéen très utilisé[198]. Pour ce faire, la voie de biosynthèse du mévalonate de *Saccharomyces cerevisiae* a été modifiée, couplée à des étapes de biosynthèse présentes chez *Escherichia coli* et à quatre étapes enzymatiques

provenant de la plante *Artemisia annua*, producteur naturel de l'artémisinine. La souche de levure modifiée qu'ils ont développée est capable de produire de grosses quantités d'acide artémisinique, qui est excrété dans le milieu de culture et peut donc être récupérée avec un processus très simple de purification. Ce résultat scientifique important, qui fut plus tard transposé chez *E.coli*[199], a nécessité un gros travail d'optimisation des différents niveaux d'expression des enzymes correspondantes à chacune des étapes de la voie de biosynthèse artificielle.

En ce qui concerne les stratégies de biosynthèse combinatoire appliquées aux aminosucres, on peut citer les travaux de Salas qui, en exploitant la tolérance de certaines glycotransférases vis-à-vis de substrats non naturels, a réussi à concevoir des analogues de produits naturels antibiotiques[200] ou anticancéreux[201] où la distribution et la nature des substitutions glycosidiques est complétement altérée. Même si elle n'a pour le moment pas donné lieu à de nouvelles molécules, cette stratégie est particulièrement prometteuse pour la biosynthèse combinatoire de nouveaux analogues d'aminoglycosides[202].

### 5) Conclusion

Les aminoglycosides ont initialement été isolés de souches bactériennes du genre Actinomycètes entre les années 1945 et 1970. Suite à cette avancée scientifique exceptionnelle, la communauté scientifique s'est essentiellement concentrée sur des approches de type hémisynthétique ou purement chimique pour concevoir de nouveaux analogues, la chimie des substances naturelles entrant dans un relatif déclin. On redécouvre aujourd'hui l'énorme potentiel qu'il y a à exploiter la biodiversité microbienne pour découvrir de nouveaux aminoglycosides.

En effet, l'isolation et le criblage à haut débit de souches isolées dans des biotopes jusqu'à présent inexplorés nous permettent d'avoir accès à des sources de produits naturels toujours renouvelées, surtout quand ces techniques sont couplées à de stratégies de fermentation innovantes. Le séquençage de génomes entiers étant de plus en plus rapide, on découvre toujours plus de clusters biosynthétiques qu'on peut analyser à l'aide d'outils bioinformatiques en constant développement. Si les approches dites de « genome mining » (métagénomiques ou passant par l'activation de clusters cryptiques) et les approches de biosynthèse combinatoire sont pour le moment surtout appliquées aux polyketides ou aux peptides non ribosomaux, elles commencent à être utilisées dans le but de découvrir de nouveaux aminoglycosides naturels ou non naturels.

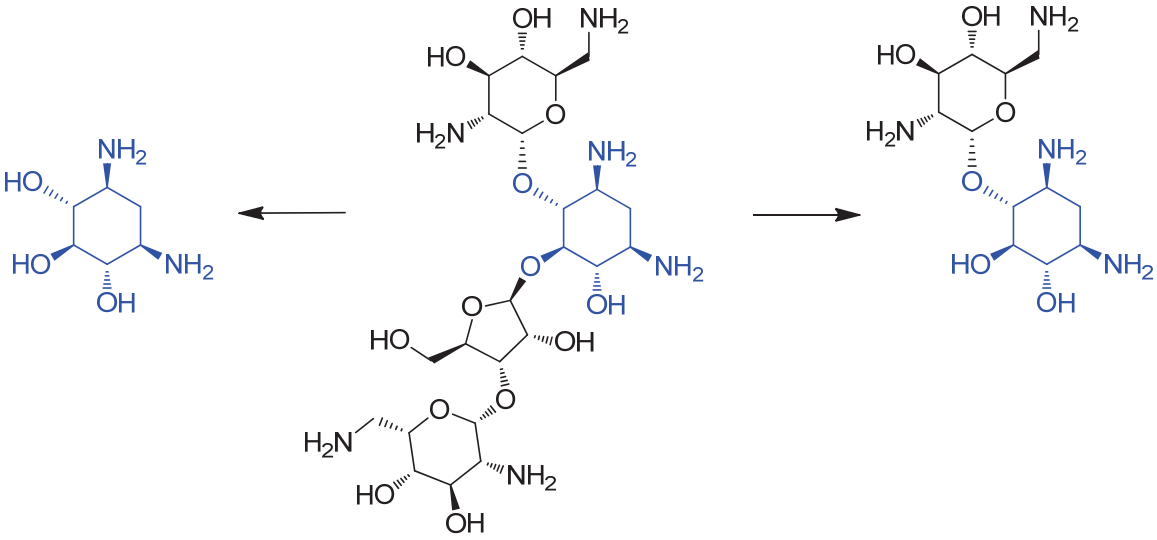
## *B - Stratégies de synthèse organique d'analogues d'aminoglycosides*

Depuis plus de 20 ans, de nombreuses approches ont été envisagées pour concevoir de nouveaux analogues d'aminoglycosides. Certaines ont pour but de simplifier et de modifier le squelette très fonctionnalisé des aminoglycosides pour améliorer leur biodisponibilité et diminuer leur toxicité, d'autres de développer des antibiotiques plus puissants, capables de limiter les phénomènes de résistance bactérienne. D'autres enfin cherchent à exploiter l'architecture particulière des aminoglycosides pour développer des ligands d'ARN plus affins et spécifiques. Pour ce faire, de nombreuses modifications fonctionnelles ont été envisagées : altération de la distribution spatiale des fonctions alcools et amines, modification de la flexibilité conformationelle, dérivatisation avec des chaines latérales variées (chaines aliphatiques ou aromatiques fonctionnalisées, nucléotides, peptides, autres molécules antibiotiques, …). Sans chercher à être exhaustifs, nous allons dans les pages qui suivent passer rapidement en revue certaines de ces stratégies synthétiques.

### 1) Stratégies de synthèse totale et d'hémisynthèse

Bien que la synthèse totale des aminoglycosides soit possible (voir par exemple la synthèse totale de la néomycine par Umezawa[203] ou les travaux de Wong[204]), elle n'est souvent pas raisonnable à grande échelle, ne fournissant que quelques milligrammes de produit final pour plusieurs dizaines de grammes de produit de départ et de nombreuses heures de travail. La synthèse de ces polysaccharides complexes requiert souvent en effet de nombreuses étapes de protections et déprotections orthogonales, et le contrôle de la diastéréosélectivité, en particulier dans les réactions de glycosylation, peut s'avérer compliqué. Les difficultés de la synthèse totale, qui diminuent d'autant le rendement final de ces synthèses totales multi-étapes, ont néanmoins induit un recentrage des efforts de recherche vers des stratégies de modification des composés naturels de type hémisynthèse.

La majorité des stratégies de synthèse mises en place commence ainsi par l'hydrolyse en milieu acide des aminoglycosides commerciaux (obtenus par fermentation) pour récupérer le cycle central deoxystreptamine (DOS) ou bien les pseudodisaccharides néamine ou paromamine, qui constituent bien souvent les motifs minimaux de reconnaissance avec les différents partenaires ARN (voir figure I-26). Ces motifs vont ensuite pouvoir faire l'objet de modifications chimiques variées.



**DOS, 1** **néomycine, 5** **néamine, 3**

**Figure I-26 Un exemple de départ d'hémisynthèse, hydrolyse sélective de la néomycine commerciale pour obtenir la deoxystreptamine ou la néamine.**

### a) Aménagement fonctionnel des aminoglycosides

Parmi les pionniers du domaine, on peut citer Umezawa et Hanessian qui conçoivent, à partir du début des années 70, des stratégies pour modifier sélectivement les architectures aminoglycosidiques. Même pour des modifications apparaissant mineures, il faut bien souvent déployer des trésors d'ingéniosité pour modifier sélectivement une seule fonction de ces molécules extrémement fonctionalisées.

Ainsi, dans le but de développer des aminoglycosides capables de résister aux enzymes de résistance venant phosphoryler les fonctions alcools, Umezawa décrit des stratégies de protection/déprotection pour supprimer une ou plusieurs fonctions alcools des aminoglycosides entiers[205]. Ces travaux conduisent Hanessian à développer des analogues où certaines fonctions alcools sont épimérisées[206], remplacées par des halogènes ou des amines[207], ou bien encore supprimées[208].

Ces premières études ouvrent la voie à de nombreuses stratégies d'aménagement fonctionnel des aminoglycosides que nous ne pouvons pas étudier en détail ici. On peut néanmoins citer les travaux de Vasella qui modifie (déoxygénation, fluoration, bromation, azidation, …) certaines des fonctions alcools portées par la paromomycine dans le but de modifier la basicité relative des amines proches et de potentiellement échapper aux enzymes de résistance venant reconnaitre de façon

spécifique les alcools modifiés[209]. A l'aide d'une approche alternative de « glycodiversification », Chang étudie systématiquement la relation structure-activité d'une large librairie d'analogues de la kanamycine résultant de la glycosylation d'aminosucres modifiés sur la néamine[210] (voir figure I-27). En suivant une approche similaire, Wong et ses collègues synthétisent une série d'analogues de néamine dans laquelle la position des amines autour du cycle glucopyranose varie afin de trouver le profil de substitution optimal pour interférer spécifiquement avec la traduction bactérienne[204].

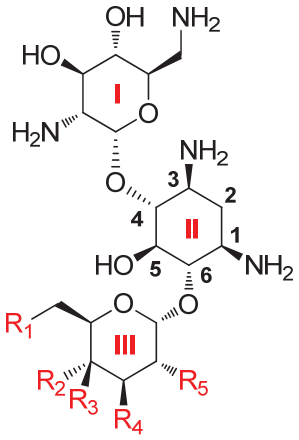
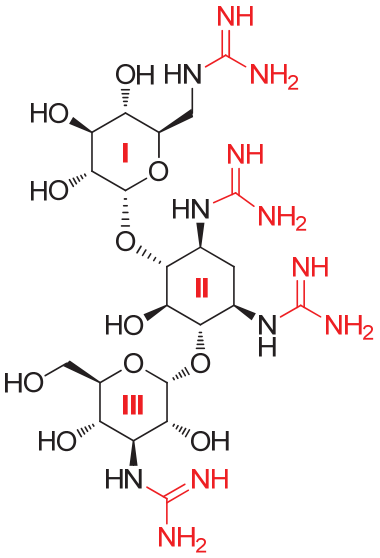



**Figure I-27 Stratégie de glycodiversification pour obtenir des analogues de kanamycine**[210]
Le cycle III est ajouté par glycosylation de la néamine protégée, les groupements R peuvent être des alcools, des amines ou des hydrogènes.

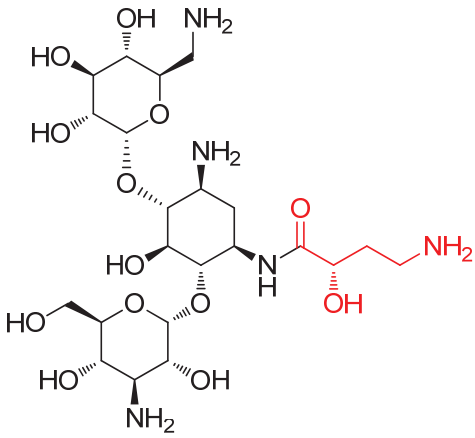
Afin d'étudier l'influence des interactions électrostatiques sur l'affinité des aminoglycosides pour les ARN, Tor étudie l'influence du remplacement d'un alcool par une amine sur diverses architectures d'aminoglycosides[10199]. Réalisant l'importance de la contribution électrostatique, ceci l'amène à concevoir des analogues d'aminoglycosides où l'ionisation des azotes est modulée. Il remplace pour cela les fonctions amines par des fonctions guanidines, plus basiques et de géométrie plane. Les analogues ainsi conçus ont été nommés « guanidinoglycosides » et sont de biens meilleurs inhibiteurs de la réplication du virus HIV que les aminoglycosides correspondants[211] (voir figure I-28). On peut expliquer ces résultats par des considérations électrostatiques mais aussi par des considérations de biodisponibilité, les guanidinoglycosides semblant en effet pouvoir passer les membranes plus facilement que les aminoglycosides[212].



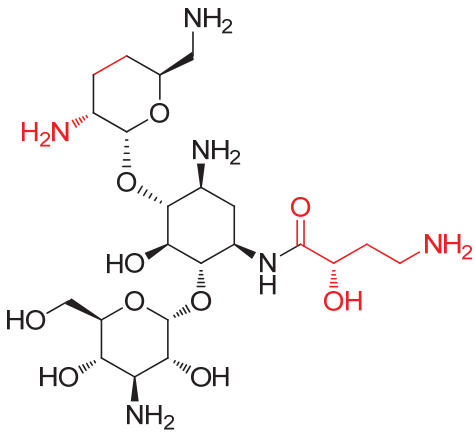

**12**

**Figure I-28 Un exemple de guanidinoglycoside, la kanamycine guanidinylée**

Ces études de type structure/activité ont eu un impact important sur notre compréhension des caractéristiques structurales et physicochimiques nécessaires à la reconnaissance des partenaires protéiques (enzymes de résistance) ou nucléiques (ARNr par exemple). Elles ont ainsi grandement influencé notre compréhension de l'activité biologique des aminoglycosides. En contribuant à raffiner les outils synthétiques à la disposition de la communauté scientifique, elles ont par ailleurs ouvert la voie au développement d'aminoglycosides modifiés par hemisynthèse capables d'échapper aux enzymes de résistance connues à l'époque. Ainsi, deux antibiotiques utilisés en clinique et actifs sur des souches multirésistantes, l'amikacine[213] et l'arbekacine[214], sont issus de la condensation d'acide 4-amino-2-(S)-hydroxy-butanoique (AHBA) sur, respectivement, la kanamycine et la 2'-amino-3'-4'-didesoxykanamycine (figure I-29).



**amikacine**, **13**



**arbekacine**, **14**

**Figure I-29 L'amikacine et l'arbekacine sont des composés semi-synthétiques issus de modifications (en rouge) de la kanamycine naturelle**

A partir du milieu des années 90, les études structurales de complexes d'aminoglycosides avec leur cible ARN ou avec des enzymes de résistance ont permis de guider la modification chimique des aminoglycosides par hemisynthèse.

L'équipe de Mobashery s'est ainsi appuyé sur les structures RMN et cristallographiques du complexe de la néamine avec le site A du ribosome (voir la partie I-II-D) pour concevoir des analogues de néamine possédant une chaine 4-amino-2-hydroxybutyryl sur l'amine 1 de la DOS, comme dans le cas de l'amikacine, ainsi que divers substituants amines aliphatiques sur l'hydroxyle en position 6 (voir figure I-30)[215]. De façon très positive, dans la plupart des cas, ces composés se sont révélés être de mauvais substrats des enzymes de résistance connues, tout en ayant une activité antibiotique plus importante que la néamine.

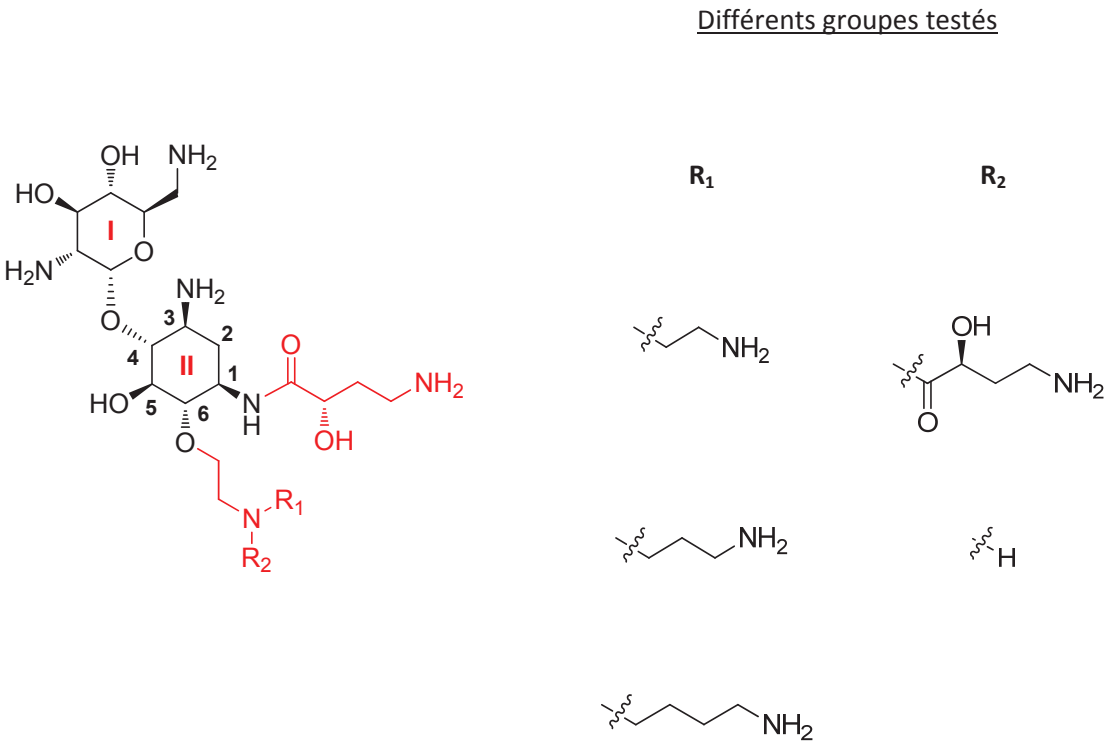



**Figure I-30 Analogues de néamine développés par Mobashery**[215]

Le même groupe a plus tard conçu une stratégie élégante pour lutter contre les phénomènes de résistance causés par des phosphorylases modifiant l'alcool en position 3 du cycle I de la kanamycine (APH(3')). En oxydant sélectivement cet alcool en cétone (en équilibre avec sa forme hydratée), ils ont obtenu un analogue de la kanamycine (composé **15**) capable de contourner l'inactivation enzymatique en éliminant naturellement le phosphate ajouté[216] (voir figure I-31).

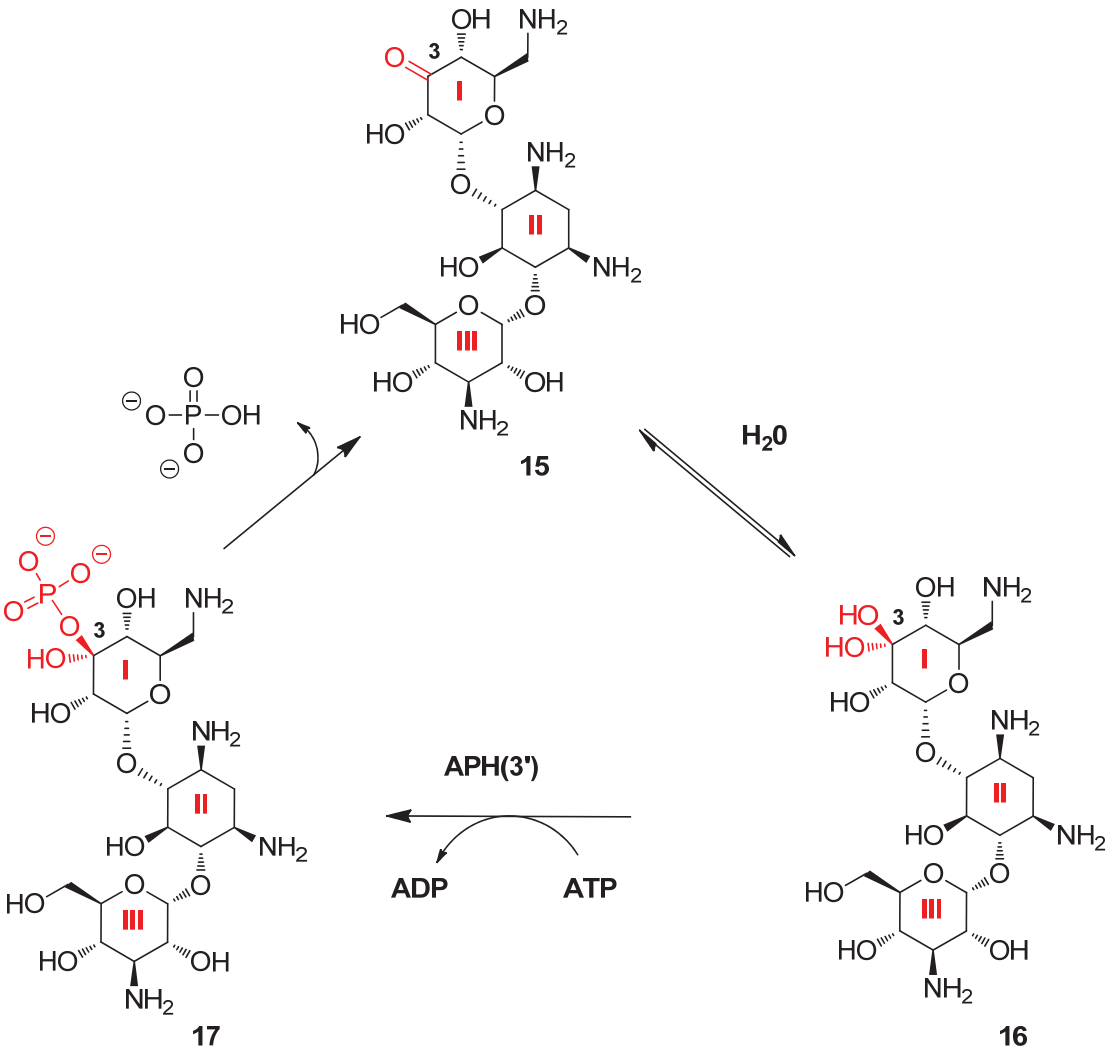



**Figure I-31 Mécanisme de régénération de l'analogue de kanamycine.**
La fonction cétone est en équilibre avec sa forme hydratée. La phosphorylation par l'enzyme de résistance conduit à l'élimination du phosphate et à la régénération de l'antibiotique.

Suivant la même stratégie que celle ayant conduit à la découverte de l'amikacine, les équipes de Westhof et Hanessian ont conçu des analogues de la paromomycine, en substituant l'hydroxyle de la position C6'' par des chaines ethylamine, propylamine ou pyridine. Les études structurales du site A du ribosome bactérien indiquaient en effet l'intérêt de cette position pour créer des ligands plus affins. De façon surprenante, les composés synthétisés se sont avérés avoir un mode de fixation très différent de leur molécule mère, la paromomycine, sans toutefois être de bien meilleurs ligands[217] (voir figure I-32). Le même type d'étude a été effectué sur d'autres aminoglycosides, par exemple en position 5 de la tobramycine[218], sans donner de résultats spectaculaires en termes d'affinité.

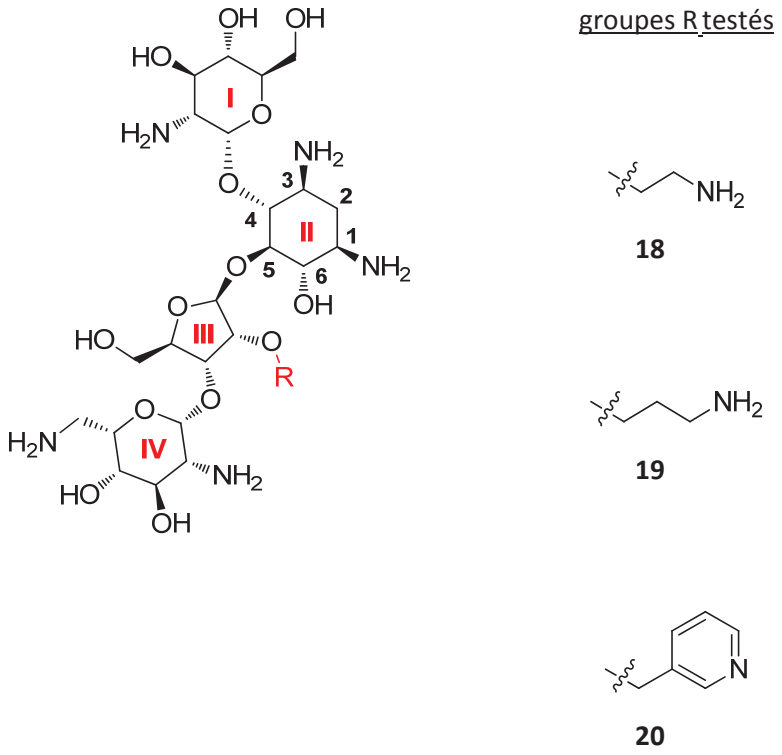



**Figure I-32 Analogues de paromomycine développés par Hanessian *et al*[217]**

Enfin, on peut citer les travaux récents de Vourloumis, qui s'inspire de la structure du cycle central déoxystretamine commun à la plupart des aminoglycosides, pour concevoir des ligands de type spiro-ether, ayant des affinités extrêmement élevées (Kd ≈ nM) pour le site A du ribosome procaryote malgré leur relative simplicité structurale en comparaison avec la néomycine[219,220] (voir figure I-33, composés **21** et **22**).

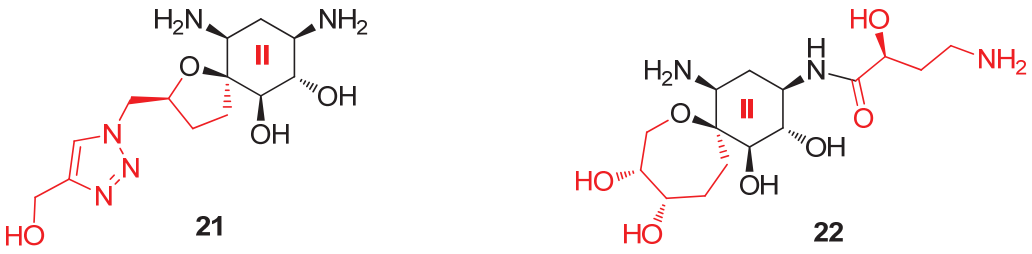
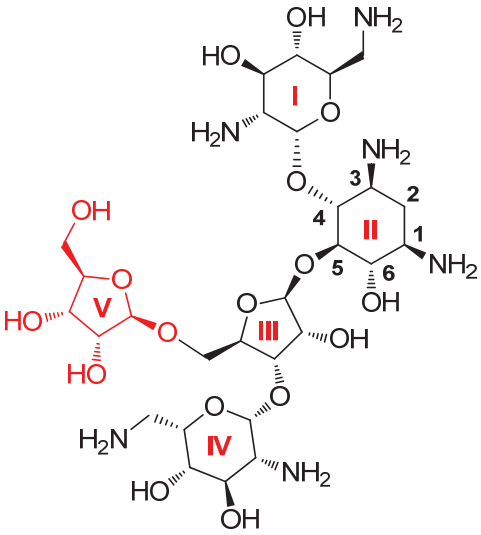


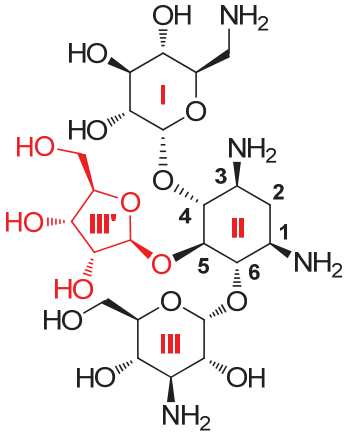

**Figure I-33 Analogues d'aminoglycosides de type spiro-ether développés par l'équipe de Vourloumis**

### b) Changement du nombre de cycle et de leur géométrie globale des aminoglycosides

Une autre stratégie intéressante pour obtenir des analogues non naturels d'aminoglycosides est de modifier le nombre de cycles glycosidiques. Cette approche a été adoptée avec succès par le groupe Baasov qui a synthétisé des analogues de néomycine en ajoutant un sucre supplémentaire sur l'hydroxyle en 5'' du cycle III. Il a ainsi préparé une série de pseudopentasaccharides où la partie néomycine responsable de la sélectivité envers le ribosome est intacte tandis que le cycle supplémentaire est censé augmenter l'affinité. De nombreux cycles furent testés, certains analogues étant plus actifs que la néomycine sur des souches résistantes (composé **23**)[221]. Cette stratégie a aussi été exploitée par Asensio pour concevoir des hybrides ribostamycine/kanamycine très encombrés et peu susceptibles d'être modifiés par des enzymes de résistance (voir le composé **24** par exemple)[222] (voir figure I-34).




**23**




**24**

**Figure I-34 Hybrides d'aminoglycosides obtenus par ajout de cycle glycosidique**

### c) Analogues contraints

Nous avons vu précédemment (partie I-II-C) l'importance des phénomènes d'adaptation conformationelle des aminoglycosides lors de l'interaction avec leur cible. Il a été montré qu'une même molécule pouvait adopter des conformations drastiquement différentes en fonction de la cible à laquelle elle est liée, cette flexibilité étant souvent à l'origine d'une spécificité modeste. De façon très intéressante, les études structurales par RMN ou cristallographie ont démontré que les

conformations reconnues par différentes cibles pouvaient être très différentes. En bloquant un ligand dans la conformation la plus favorable à son interaction avec une cible particulière, on peut espérer améliorer sa spécificité.

Par exemple, lors de sa fixation au site A du ribosome, la néomycine adopte une conformation très compacte où l'hydroxyle en position 5 du cycle III et l'amine en position 2 du cycle I forment une liaison hydrogène intramoléculaire. Cette conformation est très différente de la conformation observée lors de l'interaction de la néomycine avec l'ARN viral TAR, où les deux fonctions sont bien plus éloignées. Partant de cette constatation, les équipes de Tor et Herman ont proposé la synthèse d'un analogue de la néomycine, le composé **25**, résultant de la cyclisation intramoléculaire entre ces deux positions afin d'augmenter la spécificité de la néomycine pour le site A du ribosome[147,223] (voir figure I-35).

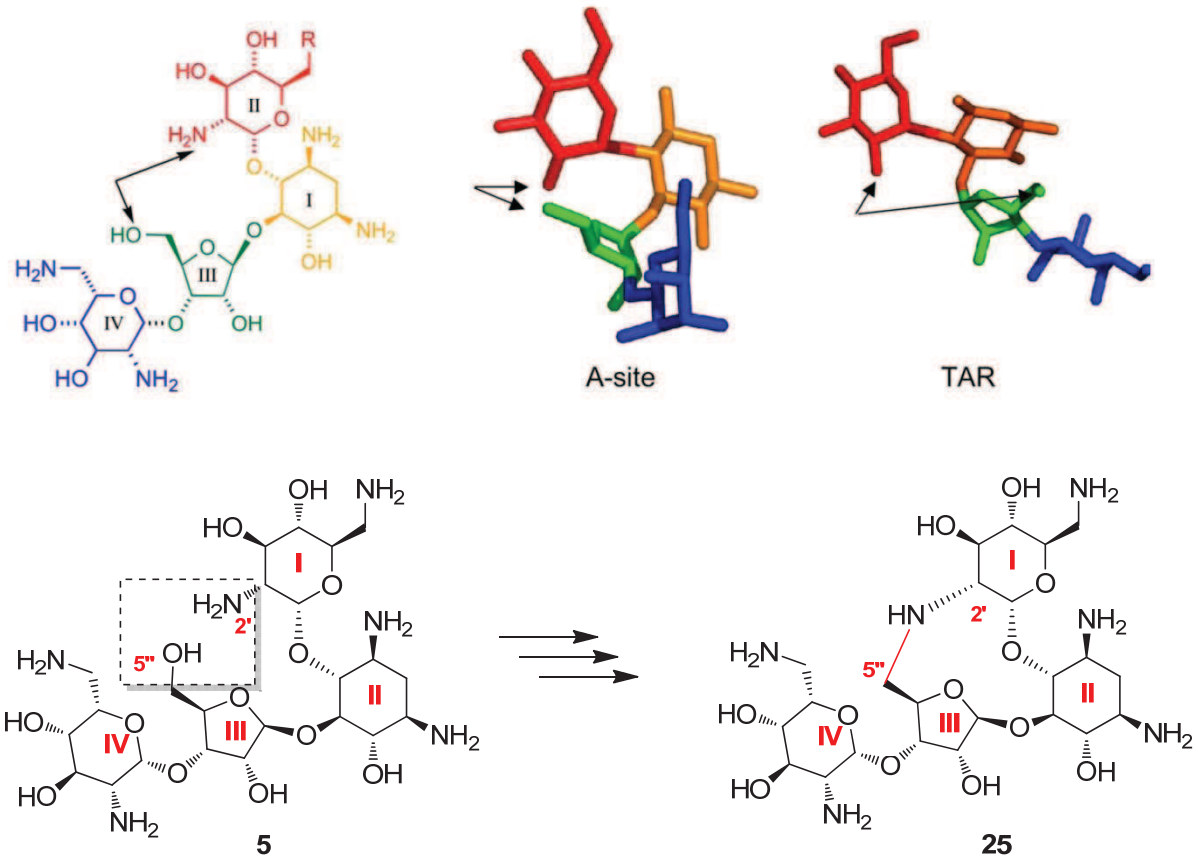



**Figure I-35 Etude structurale de la conformation de la néomycine en complexe avec le site A du ribosome ou l'ARN viral TAR et design de l'analogue contraint**
(figure adaptée de [89])

En dépit de l'ingéniosité de cette démarche, cet analogue s'est révélé avoir environ la même affinité pour le site A que la néomycine, le gain entropique étant compensé par une perte d'enthalpie. Par ailleurs, les auteurs n'ont pas constaté de différences significatives de spécificité pour l'ARN TAR, malgré la conformation a priori défavorable de l'analogue. Ceci illustre une fois de plus la flexibilité des cibles ARN, capables de changer de conformations pour s'adapter à leur ligand.

Si cette stratégie de modification de la flexibilité des aminoglycosides pour améliorer leur spécificité n'a pas pour le moment donné de résultats positifs pour des cibles ARN, il n'en n'est pas de même avec les cibles protéiques. Ainsi, après avoir comparé par RMN les conformations de la néomycine liée au site A du ribosome ou avec certaines enzymes de résistance, le groupe d'Asensio a conçu des analogues contraints qui se sont révélés garder toute leur activité antibiotique tout en étant moins sensibles aux enzymes de résistance[224].

### d) Aminoglycosides conjugués

Pour concevoir des ligands affins et spécifiques d'une cible ARN particulière, une solution est de lier deux ligands connus entre eux. Cette stratégie s'est révélée assez efficace même si elle met fréquemment en lumière notre faible compréhension des caractéristiques d'un bon ligand d'ARN. En effet, les ligands synthétisés ont souvent des sélectivités très différentes de ce qui avait été prédit et il est parfois difficile de préserver leur affinité tout en augmentant leur sélectivité.

#### *i. Conjugués avec des bases nucléiques ou des intercalants*

L'équipe de Tor a synthétisé des hybrides aminoglycosides/bases nucléiques dans l'espoir de concevoir des ligands bifonctionels, combinant reconnaissance par la partie aminoglycosidique de motifs structuraux caractéristiques (boucles, …) et appariement classique entre bases nucléiques[146]. Ils ont ainsi synthétisé des hybrides de paromomycine ou de néomycine, sur lesquels étaient greffés diverses bases azotées en position 5'' ou 6' (voir figure I-36).

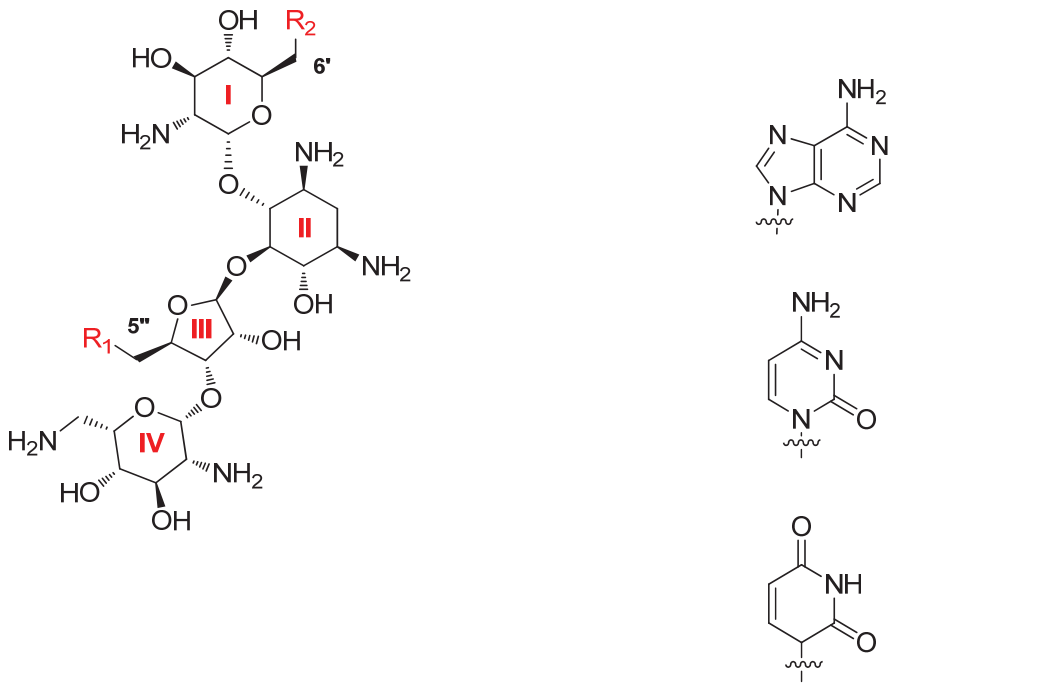



**Figure I-36 Hybrides aminoglycosides/ base nucléique**

Au contraire de ce qui avait été envisagé durant leur conception, ces molécules conjuguées se sont révélées être de plus mauvais ligands du site A que l'aminoglycoside de départ alors que leur affinité pour l'ARN viral TAR était plus grande. Ceci illustre bien la difficulté qu'il y a à prévoir les sélectivités de nouveaux ligands d'ARN.

Suivant la même stratégie, il a été envisagé de coupler les aminoglycosides avec des molécules intercalantes de type polycycliques aromatiques, connues pour se fixer sur les régions duplex de l'ARN. En adaptant la longueur du lien entre ces deux entités chimiques, on peut alors imaginer créer des molécules capables d'interagir spécifiquement à l'interface entre une boucle ciblée par l'aminoglycoside et une hélice ciblée par l'intercalant. L'équipe de Tor a ainsi synthétisé des hybrides aminoglycosides/acridine qui ont été testés pour leur capacité à cibler l'ARN viral RRE, impliqué dans la maturation du VIH (voir partie I)1))[225] (composés **26** et **27**, figure I-37). Certains de ces analogues ont montré une synergie intéressante entre les deux parties de la molécule, résultant dans des affinités de l'ordre du nanomolaire, plus de 100 fois supérieures à celle de l'aminoglycoside de départ[226]. D'autres analogues aminoglycoside/intercalant ont plus tard montré des activités intéressantes vis-à-vis de la partie ARN de la télomérase humaine[227].

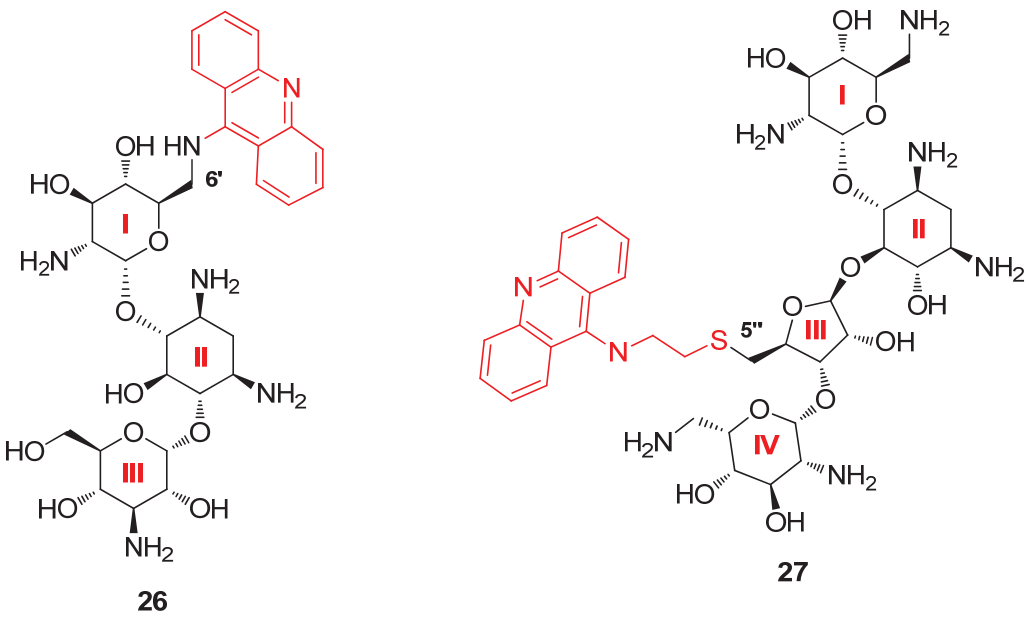
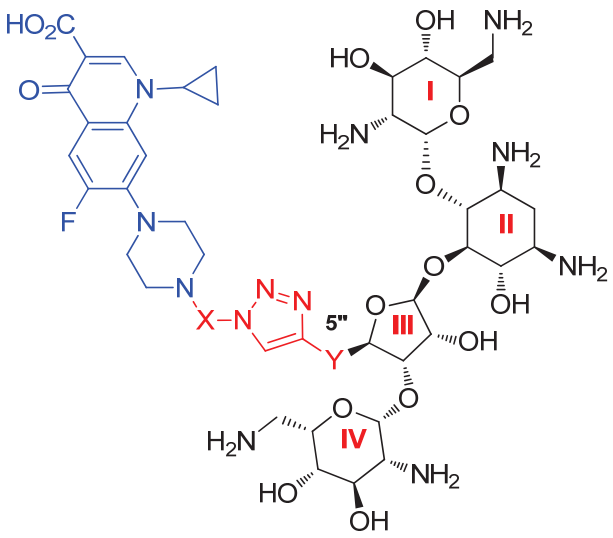



**Figure I-37 Hybrides aminoglycosides/acridine**
à gauche, kanamycine couplée en 6' (**26**); à droite, néomycine couplée en 5'' (**27**)

### *ii. Conjugués avec d'autres antibiotiques*

Baasov et ses collègues ont conçu des hybrides de néomycine liés par chimie click à la ciprofloxacine, un antibiotique de la famille des fluoroquinones (voir figure I-38). Un gros travail d'optimisation a été effectué sur le lien triazole, qui est essentiel au bon positionnement des deux sous-unités antibiotiques de la molécule. De façon très intéressante, ces molécules se sont avérées être de très bons antibiotiques, avec des activités supérieures ou égales aux composés de départ et un mode d'action dual, c'est-à-dire ciblant à la fois le ribosome à travers la partie aminoglycoside et l'ADN gyrase bactérienne qui est la cible des fluoroquinones. Par ailleurs, il a été constaté que l'apparition de résistance bactérienne contre ces molécules était très lente en comparaison de ce qui est observé pour un mélange des deux antibiotiques initiaux[228]. Ceci est sans doute dû à la difficulté de coévoluer deux mécanismes de résistance en même temps.



Quelques-uns des espaceurs essayés

**X** = -$(CH_2)_n$-, n variant de 2 à 6

**X** = -$CH_2$-CH(OH)-$CH_2$-

**X** = -$(CH_2)_2$-O-$(CH_2)_2$-

**X** = -$CH_2$-*m*$C_6H_4$-$CH_2$-

**Y** = -*m*$C_6H_4$-NHCO-

**Y** = -$(CH_2)_2$-NHCO-

**Y** = -$(CH_2)$-O-$(CH_2)$-

**Figure I-38 Hybride neomycine (noir)/cipofloxacine (bleu), le lien triazole est en rouge**

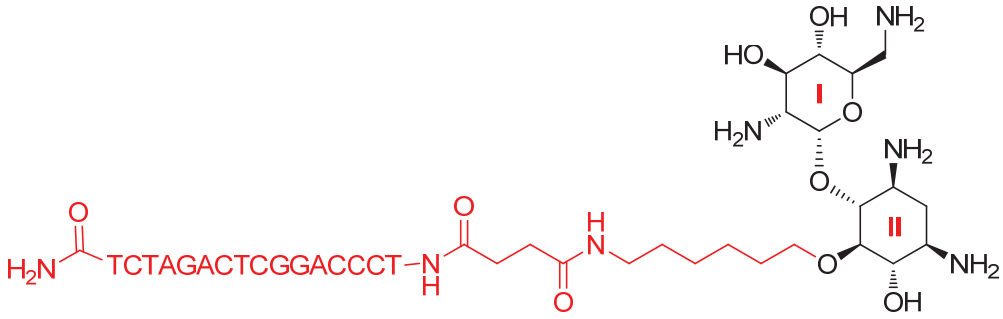
Cette démarche intéressante permit le développement d'autres analogues d'aminoglycosides. Il a par exemple été envisagé de coupler des aminoglycosides, dont la cible est la sous-unité 30S du ribosome, avec le chloramphenicol qui lui vise la sous unité 50S. En adaptant le lien entre les deux parties de la molécule, l'équipe de Yu a obtenu des ligands de très haute affinité et de bonne spécificité pour un certain nombre de cibles ARN[229].

### *iii. Conjugués avec des oligonucléotides antisens*

Les aminoglycosides sont connus pour leur grande affinité envers l'ARN doublée d'une spécificité modeste, souvent cause de leur toxicité. Pour remédier en partie à ce problème, l'équipe de Décout a eu l'idée de coupler les aminoglycosides avec une autre classe de ligands d'ARN, les acides nucléiques peptidiques (PNA)[230]. Ces molécules possèdent les mêmes bases azotées que l'ARN, fixées sur un squelette composé d'unités N-(2-aminoethyl)-glycine reliées entre elles par des liaisons peptidiques. Ces courtes séquences pseudo-nucléiques, lorsqu'elles sont complémentaires de la séquence d'un ARNm cible, vont pouvoir interférer spécifiquement avec l'expression du gène associé. Cette stratégie antisens est néanmoins pour le moment limitée par le manque de solubilité et la faible pénétration cellulaire de ces molécules.

Concevoir une molécule hybride aminoglycoside/PNA confère deux avantages. Cela permet d'une part d'améliorer la spécificité du ligand en combinant reconnaissance de motifs structuraux due à la partie aminoglycoside avec reconnaissance de séquence due à la partie PNA. L'autre avantage de cette stratégie est qu'elle permet, en tirant profit de la très forte solubilité des aminoglycosides, de remédier en partie au problème de solubilité et de pénétration cellulaire des PNA.

L'équipe de Décout a ainsi synthétisé le premier exemple de conjugué PNA-néamine, le composé **28**, qui a montré une bonne inhibition de la réplication du virus VIH-1 via interaction de forte affinité et clivage de l'ARN viral TAR[230] (voir figure I-39). Il a plus tard été démontré que cette stratégie pouvait être utilisée pour cibler d'autres ARN viraux comme par exemple l'ARN RRE[231].




**28**

**Figure I-39 Conjugué néamine-PNA développé par Décout**

### 2) Approche par fragment et synthèse combinatoire

Pour trouver une architecture de départ dans la conception de ligands, une technique traditionnellement utilisée, notamment par l'industrie pharmaceutique, consiste à cribler de très larges bibliothèques de molécules. Cette méthode de criblage conduit couramment à la détection d'un certain nombre de produits de forte affinité pour la cible d'intérêt nommés les « touches » (« hits » en anglais). Dans notre cas, les aminoglycosides constituent une bonne architecture de départ pour concevoir des ligands d'ARN. Les propriétés de ces molécules comme leur affinité, leur spécificité, ou leur pharmacocinétique peuvent ensuite améliorées à l'aide des stratégies précédemment évoquées pour conduire à des « têtes de série » (« lead » en anglais).

La synthèse par fragments s'appuie sur une approche différente du problème. Les ligands d'ARN peuvent en effet être vus comme la combinaison de briques élémentaires, fragments de petite taille et d'affinité modeste pour la cible[232]. Le but du criblage est alors d'identifier ces fragments afin de les combiner entre eux pour développer des ligands puissants ayant l'activité biologique recherchée

(voir figure I-40). Les avantages de cette méthode sont qu'elle permet potentiellement d'explorer un espace de diversité chimique très large avec un faible nombre de fragments et que son taux de découverte d'inhibiteurs est meilleur qu'avec les techniques traditionnelles[233]. En outre, en s'intéressant à des entités de faible masse moléculaire, on s'assure qu'une large partie de la molécule est directement impliquée dans l'interaction avec la cible.

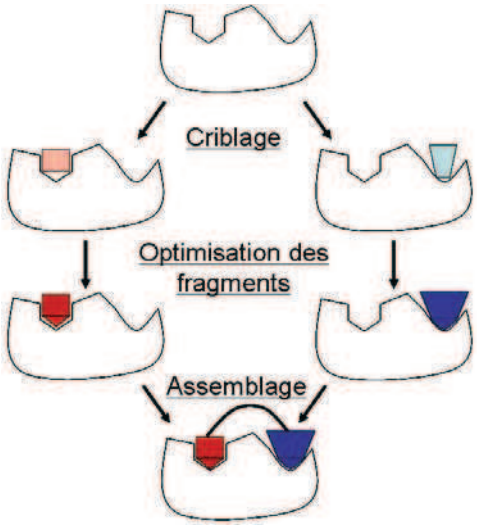



**Figure I-40 Stratégie de synthèse de ligand par fragment**

Notre connaissance des interactions aminoglycosides/ARN permet de guider le design des différents fragments en utilisant des techniques telles la RMN, la diffraction des rayons X, la spectrométrie de masse ou encore la résonnance de plasmon de surface (SPR). Le véritable enjeu est ensuite de parvenir à relier les fragments entre eux afin de créer un ligand à la fois affin et spécifique d'une cible donnée. Dans le cas des aminoglycosides, qui ont souvent plusieurs sites de fixation sur une même cible ARN, cette stratégie a été employée avec succès pour concevoir de nouveaux analogues en combinant des motifs structuraux connus.

Ainsi, en étudiant par SPR l'interaction de la néamine avec un analogue du site A du ribosome procaryote, Wong et ses collègues ont démontré que la stœchiométrie du complexe formé était de 2 pour 1. S'appuyant sur cette constatation, ils ont donc ont créé des dimères de néamine en faisant varier la longueur et la nature du lien entre ces deux fragments afin de créer un ligand « bidentate ». Un gros travail d'optimisation du lien a été nécessaire, celui-ci devant être à la fois suffisamment flexible pour permettre un positionnement favorable à l'interaction avec l'ARN et suffisamment rigide pour tirer profit de l'effet entropique dû à la liaison entre les deux fragments. De nombreux bras espaceurs ont ainsi été testés : lien amide, 1,2-hydroxyamine ou éther, longueur du bras espaceur variant de 2 à 12 atomes, …

Dans les cas favorables, ces dimères se sont révélés avoir une activité biologique plus importante que celle obtenue en doublant la concentration des monomères correspondants, ce qui est indicatif du caractère synergique de leur action. Par ailleurs, de façon très intéressante, ces dimères sont actifs sur des bactéries résistantes aux aminoglycosides[234,235] (voir figure I-41). D'autres dimères de néamine ont plus tard été développés par Décout pour leur capacité à cibler spécifiquement l'ARN viral TAR à des concentrations submicromolaires[236].

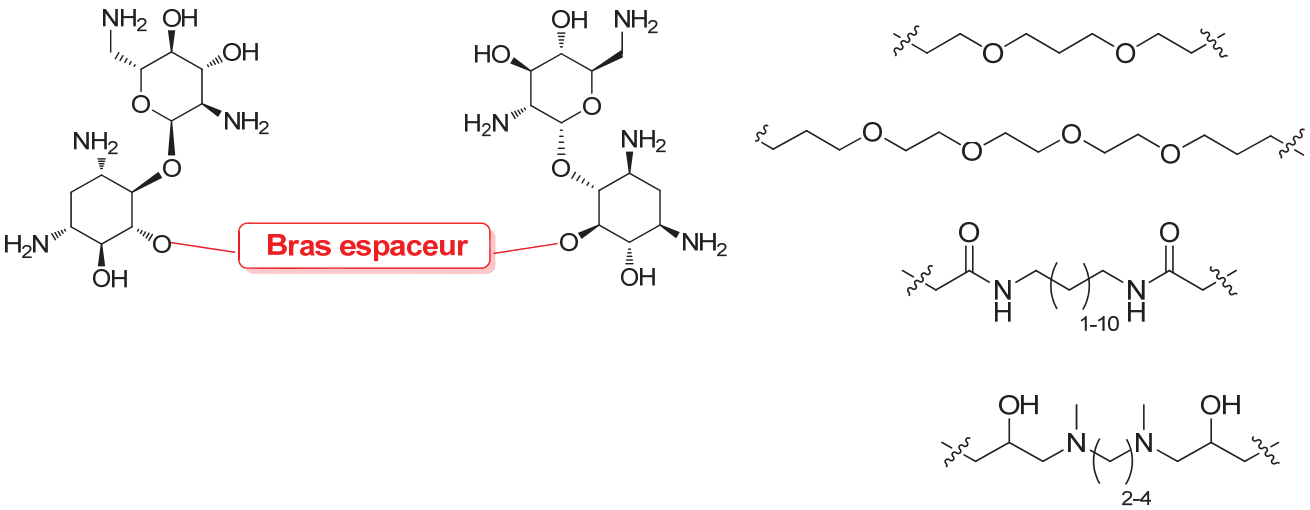



**Figure I-41 Exemples de dimères de néamine synthétisés par Wong**

Cette stratégie a été également mise à profit par les équipes de Tok[237], de Tor[238] et d'Arya[239] pour synthétiser des dimères d'aminoglycosides symétriques ou asymétriques dérivés de la néomycine, de la kanamycine ou de la tobramycine. En testant leur affinité respectivement pour l'ARN viral RRE, le ribozyme « hammerhead » ou l'ARN viral TAR, les auteurs ont constaté eux aussi que les dimères étaient globalement meilleurs que les monomères correspondants à condition d'avoir un lien optimisé pour permettre la fixation des deux monomères dans leurs sites respectifs. Ceci démontre bien l'intérêt de s'appuyer sur la promiscuité des différents sites de fixation des aminoglycosides pour concevoir des ligands puissants en liant plusieurs fragments (ou pharmacophores) entre eux.

En poussant cette stratégie un peu plus loin, on peut même imaginer concevoir des ligands se liant spécifiquement à des motifs non reconnus par les fragments aminoglycosidiques de départ. C'est le but d'une étude menée par l'équipe d'Hergenrother qui cherchait à développer des ligands

pouvant reconnaitre spécifiquement des tiges-boucles d'ARN particulières[240]. Pour ce faire, ils se sont appuyés sur les études de Puglisi qui démontrait que le cycle desoxystreptamine, commun à la plupart des aminoglycosides, avait une affinité de l'ordre du millimolaire pour la tige boucle mimant le site A du ribosome[241]. En utilisant une stratégie de synthèse combinatoire s'appuyant sur la chimie « click », son équipe a synthétisée plus de 100 dimères de DOS différents qui ont ensuite été testés pour leur capacité à reconnaitre spécifiquement des motifs particuliers comme par exemple une tétra- ou une octa-boucle[242]. Deux constructions, les composés **29** et **30**, se sont révélées avoir une affinité ainsi qu'une sélectivité raisonnable pour l'une ou l'autre de ces structures (voir figure I-42).

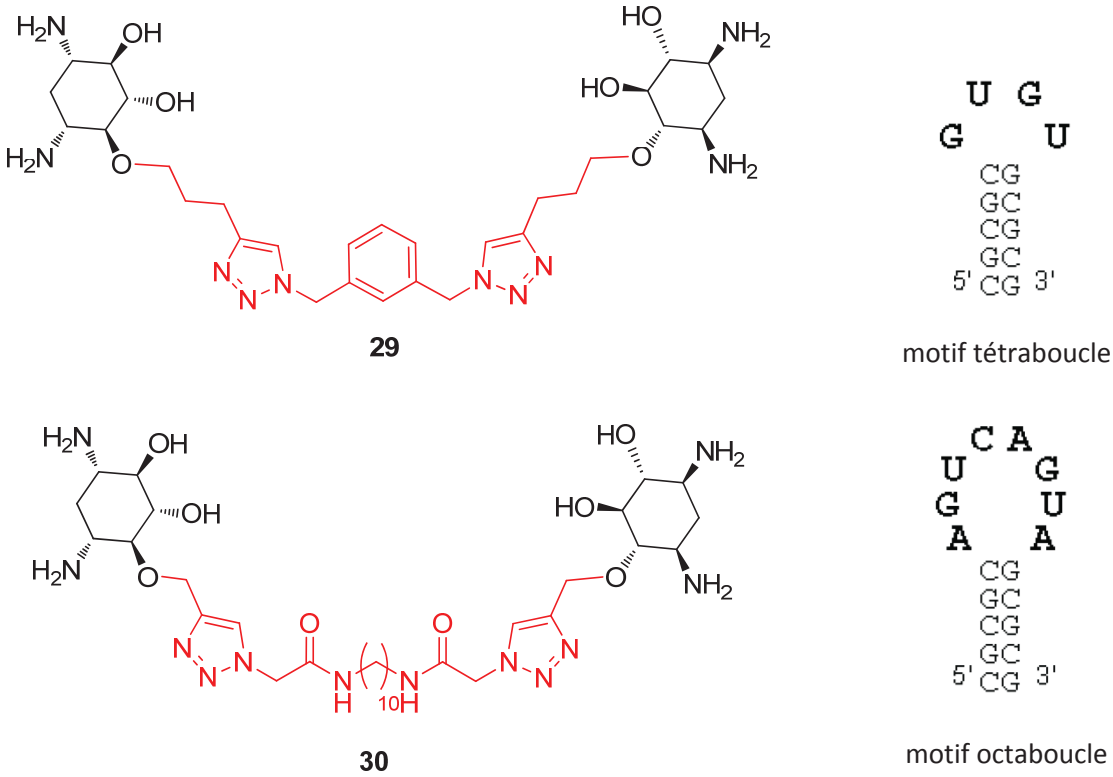



**Figure I-42 Deux dimères de deoxystreptamine en regard des boucles d'ARN auxquelles ils se lient spécifiquement** (adaptée de [242])

On peut noter ici que cette stratégie de synthèse de ligands par fragment peut également être mise en œuvre avec des analogues de DOS. C'est notamment ce que j'ai pu accomplir au laboratoire dans le cadre de mon M2 ayant pour objectif le développement d'inhibiteurs de l'enzyme de résistance AAC(6')-Ib11, une *N*-acétyl transferase venant inactiver l'ensemble des aminoglycosides connus[130]. Après avoir identifié en RMN par expérience de STD (pour « Saturation Transfert Difference ») le fragment diaminocyclopentanol comme étant un mime convenable de la DOS, nous avons synthétisé un inhibiteur résultant de l'accroche de ce fragment sur l'acetylCoA, cofacteur

essentiel au bon fonctionement de l'acetylase (voir figure I-43). La molécule ainsi synthétisée s'est avérée être un bon inhibiteur de l'enzyme de résistance *in vitro* ( Ki de l'ordre du micromolaire).

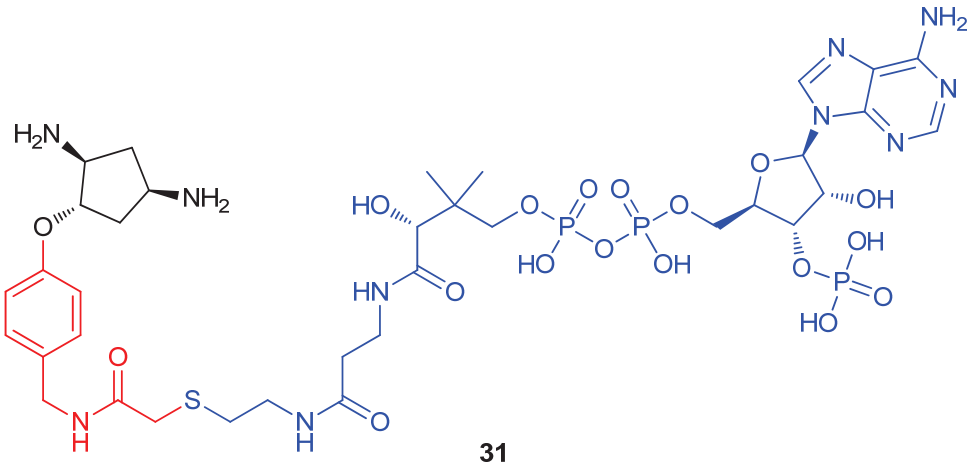



**Figure I-43 Stratégie de synthèse par fragment d'inhibiteur de l'enzyme de résistance AAC(6')-Ib11.**

Le fragment diaminocyclopentanol (en noir) constitue un mime du cycle DOS. Il est lié de façon covalente à l'acetylCoA (en bleu) par un bras espaceur optimisé (en rouge).

De la même façon, l'équipe d'Hermann s'est intéressée au cycle cis-3-5-diamino-piperidinyl comme mime potentiel de la DOS en raison du motif cis-1,3-diamino qu'ils ont en commun. En utilisant ce fragment dans la synthèse d'une famille de 3,5-diamino-piperidinyl triazines, ils ont pu identifier des composés antibactériens mimant les aminoglycosides et ayant une forte affinité pour le site A du ribosome[243] (composé **32**, à gauche dans la figure I-44) . On peut également citer les travaux de Kirschning qui a développé des dimères d'analogues simplifiés d'aminoglycosides qui se sont révélés avoir une activité intéressante vis-à-vis de l'ARN viral TAR[244] (composé **33**, à droite dans la figure I-44).

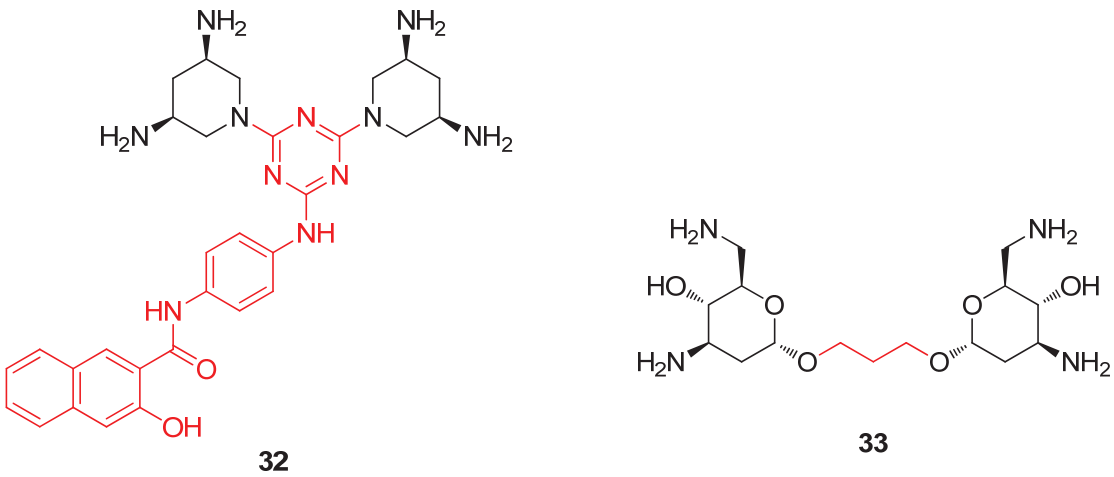



**Figure I-44 Dimères de mimes simplifiés d'aminoglycosides**

Enfin, afin de dépasser certains des problèmes liés aux stratégies de synthèse linéaire, plusieurs stratégies de synthèse combinatoire ont été développées dans le but de rapidement générer de larges bibliothèques d'analogues d'aminoglycosides en combinant de nombreux fragments de façon parrallèle. Celles-ci s'appuient le plus généralement sur la combinaison de motifs structuraux inspirés des aminoglycosides à l'aide de réactions permettant une diversification très importante. On peut citer par exemple les travaux de Westerman en 2005, qui exploite la flexibilité de la réaction multicomposant de Ugi pour synthétiser rapidement, après déprotection des aminosucres, une petite bibliothèque d'analogues d'aminoglycosides comportant un, deux ou trois cycles aminosaccharidiques en faisant réagir en un pot une amine, une cétone, un acide carboxylique et un isonitrile[245] (voir figure I-45).

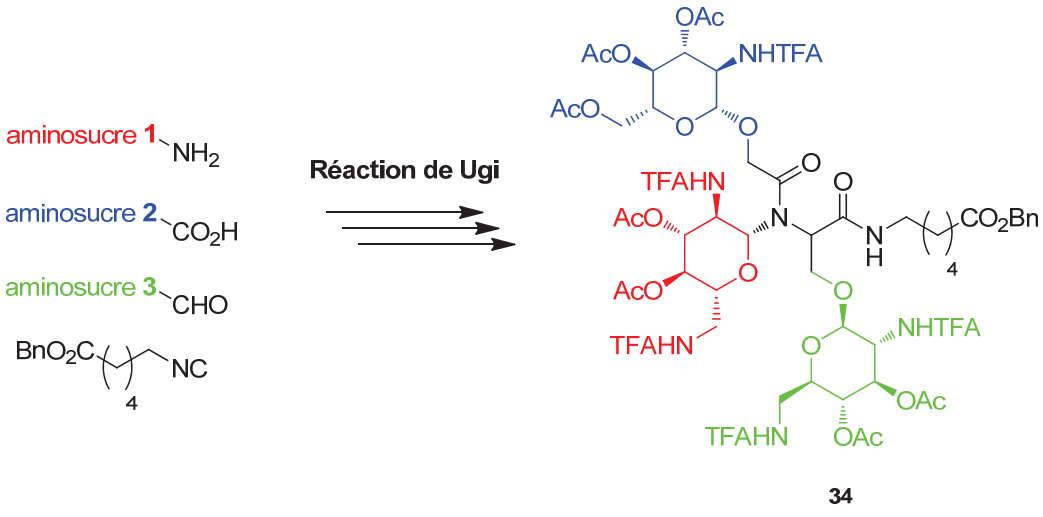



**Figure I-45 Un exemple d'analogue d'aminoglycoside obtenu par réaction multicomposant de type Ugi** (adapté de [245])

### 3) Approche chemoenzymatique

Pour dépasser certaines des limitations de la chimie organique classique (nombreuses protections/déprotections, problèmes de régiosélectivité, …), une idée originale consiste à détourner les enzymes de résistance bactérienne. Celles-ci ont en effet la particularité de pouvoir modifier les aminoglycosides de façon hautement contrôlée et efficace, par exemple en acétylant les fonctions amines dans le cas de la famille des N-acetyltransferases. Il a été montré que ces enzymes avaient une grande tolérance pour le cofacteur d'acylation (l'acetyl-CoA dans le système naturel). En tirant profit de cette tolérance pour des substrats non naturels, l'équipe de Garneau-Tsodikova a synthétisé de nombreux analogues d'aminoglycosides N-acylés en utilisant ces enzymes purifiées en tant que catalyseur et des analogues d'acétylCoA non naturels[246].

Il est également possible de détourner certaines enzymes des voies de biosynthèses des aminoglycosides. En 2008, l'équipe de Spencer à Cambridge a décrit l'emploi d'enzymes de la biosynthèse de la butirosine chez *S. aureus*, BtrH et BtrG, pour synthétiser des analogues non naturels d'aminoglycosides à partir de la kanamycine ou de la néomycine[247]. Cette même approche a été mise à profit par Baasov pour concevoir, avec une stratégie en un pot et sans aucune protection, de nombreux analogues d'aminogycosides en utilisant ces enzymes purifiées en tant que catalyseur (voir figure I-46).

**Figure I-46 Approche chemoenzymatique pour la synthèse d'analogues d'aminoglycosides**

### 4) Conclusion

L'ensemble des exemples décrits précédemment rend honneur à la créativité et à l'ingéniosité des différents groupes qui se sont penchés sur la création d'analogues d'aminoglycosides mais témoigne aussi des difficultés qu'il y a à concevoir rationnellement des molécules bioactives. Notre compréhension parcellaire des interactions petites molécules/ARN rend difficile la mise en place de modèles prédictifs de l'affinité et de la spécificité d'une molécule pour sa cible ribonucléique.

La majorité des stratégies s'appuient ainsi sur des modifications hémisynthétiques du cycle DOS et des bicycles néamine ou paromamine, qui constituent le squelette de la plupart des aminoglycosides. Ces stratégies ne permettent souvent qu'une diversification chimique modeste au

regard des efforts investis car elles mettent en œuvre des modifications chimiques dans les étapes tardives des synthèses, où la forte fonctionnalisation des aminoglycosides peut représenter un frein à ce que l'on peut faire. Par ailleurs, l'amélioration de la diversité chimique ne se fait souvent qu'au détriment de la taille de la molécule.

Les approches de synthèse totale quant à elles restent toujours assez longues et techniquement exigeantes, sans être nécessairement être suffisamment flexibles pour permettre la synthèse d'une large librairie d'analogues. Même si les progrès de la chimie organique rendent maintenant accessibles de nombreux aminosucres, la synthèse de structures aminoglycosidiques innovantes ayant une activité biologique intéressante reste un défi important.

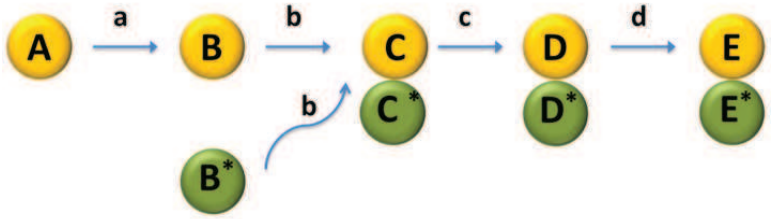
## *C - Une stratégie hydride : la mutasynthèse*

### 1) Concept de la mutasynthèse

Le concept de la mutasynthèse est simple : créer de nouvelles entités chimiques bioactives à partir de l'architecture de produits naturels connus en altérant leurs voies de biosynthèse (voir les revues de Kirschning[193,248] et de Sussmuth[249]). Il s'agit ainsi d'explorer un espace chimique potentiellement non exploité par l'évolution en concevant de petites molécules non naturelles, que l'on appellera mutasynthons, qui peuvent être utilisées par des microorganismes modifiés génétiquement. Ces mutasynthons vont constituer des précurseurs exogènes dans certaines voies de biosynthèse (voir figure I-47). En tirant profit de la richesse et de la versatilité des systèmes biologiques, on peut envisager la création de nouvelles constructions chimiques complexes obtenues à partir de mutasynthons non naturels (comme par exemple des molécules fluorées, des cycles aromatiques ou des hétérocycles, des chaines aliphatiques comportant des groupements tertbutyloxy, …). Cette stratégie est particulièrement intéressante quand ces constructions seraient difficiles à obtenir avec les outils classiques de chimie organique.

Voie de biosynthèse non modifiée

Stratégie de biosynthèse dirigée par les précurseurs (PDB)




Stratégie de mutasynthèse

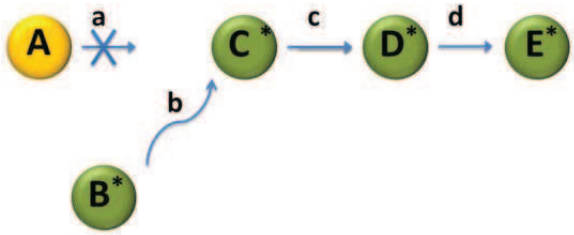



**Figure I-47 Principe des stratégies de PDB et de mutasynthèse** (adaptée de [193])

D'une certaine façon, la mutasynthèse peut être vue comme une stratégie d'hémisynthèse inversée où, au lieu de transformer un produit naturel en fin de biosynthèse, étape souvent assez complexe si on est en présence de molécules très fonctionnalisées, les transformations chimiques sont effectuées en début de biosynthèse et la machinerie biosynthétique de la souche modifiée prend en charge le mutasynthon et effectue les transformations délicates de fin de biosynthèse.

Cette idée, promue par Rinehart à partir du milieu des années 70[250], a été précédée par ce qu'on appelle la « biosynthèse dirigée par les précurseurs » (PDB pour « Precursor-Directed Biosynthesis »). L'idée de cette méthode est tout simplement de concevoir des analogues de métabolites naturels de début de biosynthèse qui sont ensuite tout simplement ajoutés au milieu de culture des microorganismes naturels producteurs (wild-type). Ces analogues peuvent entrer en compétition avec les métabolites naturels et conduire, si la voie de biosynthèse est suffisamment flexible pour les accommoder, à des analogues des produits naturels finaux (figure I-47). Les rendements de cette

méthode sont cependant souvent très faibles parce que les analogues sont la plupart du temps de moins bons substrats que les métabolites avec lesquels ils restent en compétition. De plus les mélanges d'analogues obtenus sont souvent très complexes à purifier. L'avantage de cette méthode est qu'elle ne requiert qu'une compréhension partielle des voies de biosynthèse utilisées et une connaissance minimale des milieux de fermentation adéquats pour favoriser la production de métabolite.

La mutasynthèse représente une amélioration de la PDB. Pour éviter la compétition entre les analogues synthétisés et les métabolites naturels de début de la biosynthèse, les voies de biosynthèse sont modifiées afin d'inactiver la production des métabolites naturels. Initialement, lorsque la biologie moléculaire n'avait pas encore apporté les outils performants actuels de modification génétique, la méthode consistait à provoquer des mutations aléatoires dans le producteur naturel jusqu'à l'obtention d'un organisme avec le phénotype désiré. C'est par exemple le cas dans les travaux pionniers de Rinehart sur la mutasynthèse d'analogues d'aminoglycosides (sur lesquels nous reviendrons dans le chapitre III) où les souches bactériennes étaient cultivées sur un milieu mutagène contenant de la N-methyl-N'-nitro-nitrosoguanidine[251]. Le rendement de la production des analogues est dans la plupart des cas fortement augmenté par l'impossibilité de production du produit naturel de départ sans adjonction de métabolites exogènes. Il est important de noter à cette étape que la stratégie de mutasynthèse n'est possible que dans le cas où le métabolite dont on bloque la production n'est pas vital.

Selon Rinehart[250], on peut décomposer la stratégie de mutasynthèse en 5 étapes : (1) génération d'un mutant incapable de synthétiser un métabolite essentiel dans la biosynthèse du produit naturel cible, (2) synthèse d'une librairie d'analogues de ce métabolite, (3) complémentation chimique de la souche mutée avec ces analogues, (4) purification des analogues du produit naturel final éventuellement formés, (5) évaluation de leur activité biologique. La troisième étape est bien sûr la plus importante mais malheureusement aussi la plus difficile à contrôler expérimentalement. Il faut en effet que la machinerie biosynthétique soit suffisamment flexible pour pouvoir accommoder des analogues des métabolites naturels, mais aussi qu'elle soit suffisamment efficace et sélective pour conduire à la production des produits finaux désirés.

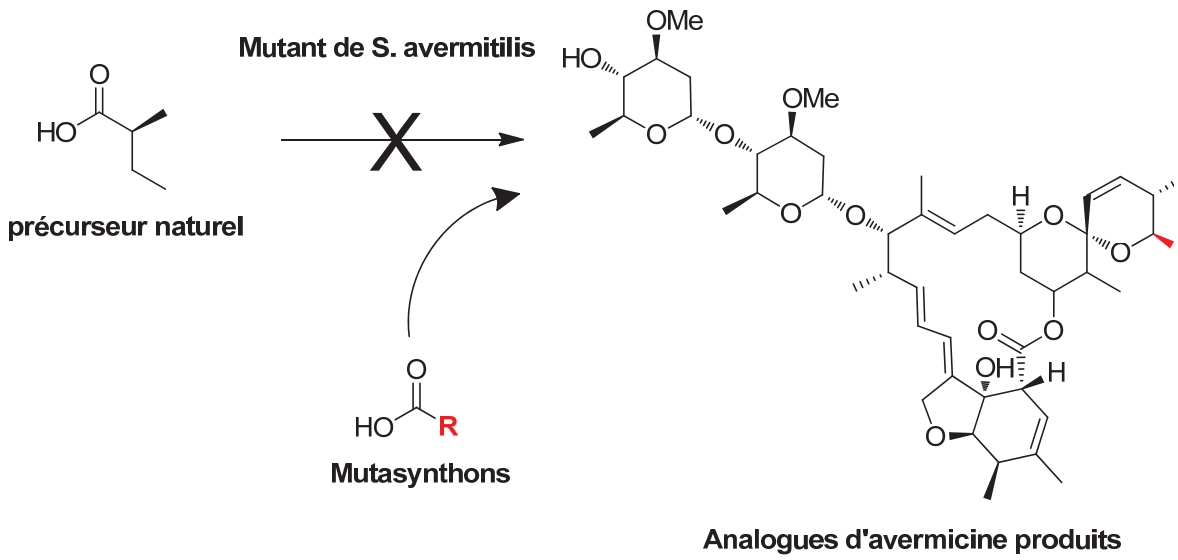
En dépit de leur grand potentiel, les stratégies de mutasynthèse n'ont connu qu'un développement limité pendant les vingt dernières années. Leur essor récent est dû à trois raisons majeures. Tout d'abord, l'amélioration des méthodes de séquençage et d'analyse bio-informatiques a permis l'augmentation exponentielle du nombre de clusters de biosynthèse décrits[252]. Cela a considérablement amélioré notre compréhension de la biochimie et du mode de fonctionnement de

ces clusters. Deuxièmement, le développement continu de nombreux outils de génie génétique permet désormais la modification rationnelle des voies de biosynthèse pour créer des mutants bloqués à une étape spécifique. On dispose par exemple aujourd'hui de nombreuses méthodes pour modifier de façon hautement contrôlée les voies de biosynthèse chez les actinomycètes (voir le livre de référence d'Hopwood sur la génétique des Streptomyces[253] ou par exemple les travaux de Pernodet[254]). Par ailleurs, il est désormais possible d'altérer la spécificité et l'efficacité des différentes enzymes de la voie de biosynthèse en utilisant des techniques d'évolution dirigée[255] ou de mutagénèse[256]. On peut d'autre part concevoir *de novo* des enzymes ayant les propriétés requises pour substituer les enzymes naturelles[257], voire d'exprimer des clusters de biosynthèse entiers ou tronqués dans des organismes optimisés pour la production hétérologue[258]. Enfin, les avancées récentes de la synthèse organique permettent d'envisager la synthèse de librairies de mutasynthons explorant un large espace chimique. Les techniques modernes d'analyse comme la chromatographie liquide haute résolution couplée à la spectrométrie de masse (HPLC-MS/MS) ou encore la RMN à haut champ facilitent par ailleurs le criblage des produits synthétisés, même s'ils sont produits en faible quantité et en mélange complexe.

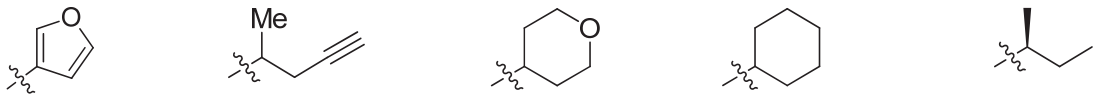
Nous allons donner dans la partie qui suit quelques exemples qui illustrent le potentiel des approches mutasynthétiques dans l'obtention d'analogues de produits naturels difficilement accessibles par les méthodes de chimie organique traditionelle. Les travaux initiaux de Rinehart sur les analogues d'aminoglycosides seront quant à eux plus précisement analysés dans le chapitre III.

### 2) Exemples de stratégies de mutasynthése

La doramectine est la première molécule issue d'une stratégie de mutasynthèse à avoir été mise sur le marché. C'est un analogue de l'avermicine, un macrolide naturel produit par *Streptomyces avermitilis* possédant de fortes propriétés antiparasitaires. Elle a été obtenue en utilisant des bactéries mutées, incapables de synthétiser l'acide carboxylique nécessaire au démarrage de la biosynthèse. En conduisant des expériences de mutasynthèse avec plus de 40 acides carboxyliques différents, il a été possible de créer des analogues de l'avermicine difficilement accessibles par des stratégies traditionnelles d'hemisynthèse. La tolérance structurale de la machinerie biosynthétique vis-à-vis de substrats non naturels a ainsi permis d'accéder à la doramectine avec des rendements convenables[259] (voir figure I-48).




Exemples de mutasynthons acceptés



**35** **36** **37** **Doramectine 38** **Avermectine 39** **(produit naturel)**

**Figure I-48 Mutasynthèse de la doramectine** (adaptée de [193])

On peut également citer les travaux de Leadlay à l'université de Cambridge et de la société Biotica qui développèrent des analogues de la rapamycine, un antibiotique de la famille des polyketides, en utilisant une stratégie de mutasynthèse sur des bactéries *Streptomyces hygroscopicus* incapables de synthétiser l'acide carboxylique de départ de la biosynthèse[260,261]. En testant une grande variété d'analogues de l'acide (4R,5R)-4,5-dihydroxycyclohex-1-enoique, ils ont pu explorer un espace chimique large en faisant varier à la fois la taille du cycle ainsi que la répartition et la nature et la stéréochimie de ses substituants. Ceci a conduit à l'identification d'analogues plus actifs que la rapamycine qui sont actuellement en essai clinique[262]. Cette même stratégie a également été employée avec succès par Kirschning sur d'autres polyketides comme par exemple la geldanamycine[263] ou l'ansamitocine[264]. La mutasynthèse permet par ailleurs la synthèse

d'analogues d'autres produits naturels comme par exemple les sidérophores phénoliques[265] ou les lipopeptides[266].

De façon extrêmement intéressante, les groupes de Sussmuth et Wohlleben ont montré que la mutasynthèse pouvait permettre la modification sélective de plusieurs domaines différents de la même molécule en fonction de la partie de la voie de biosynthèse modifiée et des mutasynthons utilisés. Ainsi, en travaillant sur la mise au point d'analogues de la balhimycine, un glycopeptide antibiotique de la famille de la vancomycine produit par *Amycolatopsis balhimycina*, ils ont pu, lors d'une première série d'expériences, créer des analogues fluorés en position $R_1$[267] (voir figure I-49). Une deuxième série d'expériences utilisant des mutants bloqués à une étape différente de la biosynthèse a permis de modifier sélectivement une autre partie de la molécule en introduisant des groupements alcools ou éther méthyliques (groupes $R_2$) [268].

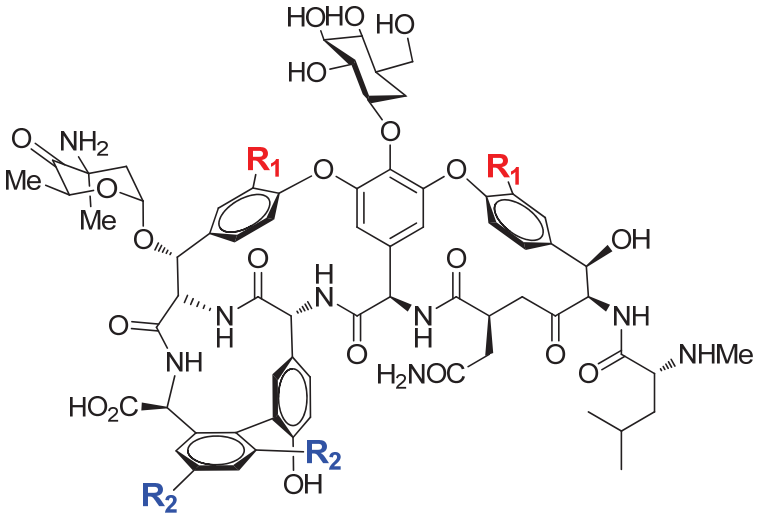
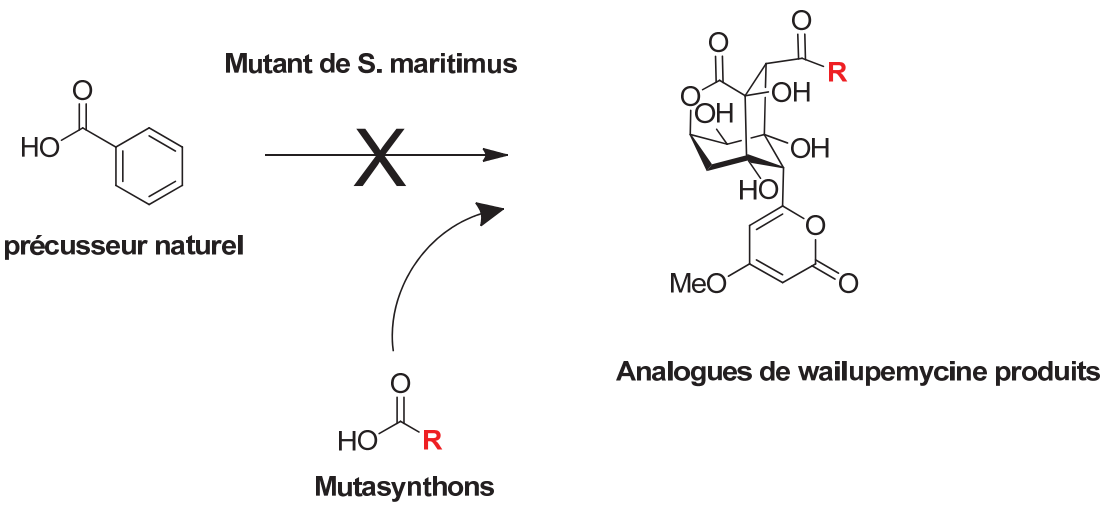



**Figure I-49 Sites de modification de la balhimycine permis par mutasynthèse**

Opter pour une stratégie de mutasynthèse est particulièrement intéressant quand les alternatives synthétiques pour obtenir des analogues ne sont pas raisonnables. C'est le cas par exemple pour la wailupemycine, un produit naturel antibiotique isolé de *Streptomyces maritimus*, dont la structure complique les stratégies de synthèse totale. En utilisant des mutants bloqués dans la synthèse d'acide benzoïque qui constitue une brique de départ de la biosynthèse, Moore et ses collègues ont pu exploiter une stratégie mutasynthétique pour concevoir une petite librairie d'analogues[269] (voir figure I-50)




Exemples de mutasynthons acceptés

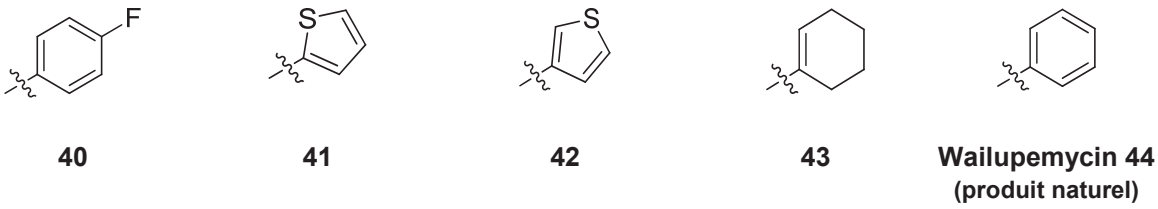



**Figure I-50 Stratégie de mutasynthèse pour la conception d'analogues de la Wailupemycine**[269]

On l'aura compris, la mutasynthèse est une stratégie potentiellement extrêmement puissante pour créer des analogues de produits naturels, spécialement quand ils sont difficilement accessibles à l'aide de stratégies traditionnelles d'hemisynthèse ou de synthèse totale. En raison de leur complexité structurale, les aminoglycosides représenteraient donc des cibles idéales pour mettre en place une stratégie de mutasynthèse.

## *D - Conclusion*

Il n'y a pas de réponse unique ni de stratégie optimale pour la découverte de nouveaux analogues d'aminoglycosides. Le bref survol des différentes stratégies mises en place par la communauté scientifique nous a permis de constater que de nombreuses solutions avaient été proposées, chacune ayant ses avantages propres. Certaines tirent parti de la richesse de la

biodiversité microbienne pour isoler de nouvelles molécules à partir de microorganismes jusqu'alors inexploités. D'autres conçoivent de nouveaux analogues non naturels par une stratégie d'ingénierie génétique des voies de biosynthèse tirant parti de l'efficacité de ces systèmes enzymatiques. Enfin, les progrès de la chimie organique permettent d'envisager le développement d'analogues en utilisant des réactions toujours plus tolérantes et sélectives à travers des stratégies de synthèse totale ou d'hémisynthèse. La figure I-51 représente schématiquement les approches possibles pour le développement d'analogues d'aminoglycosides.

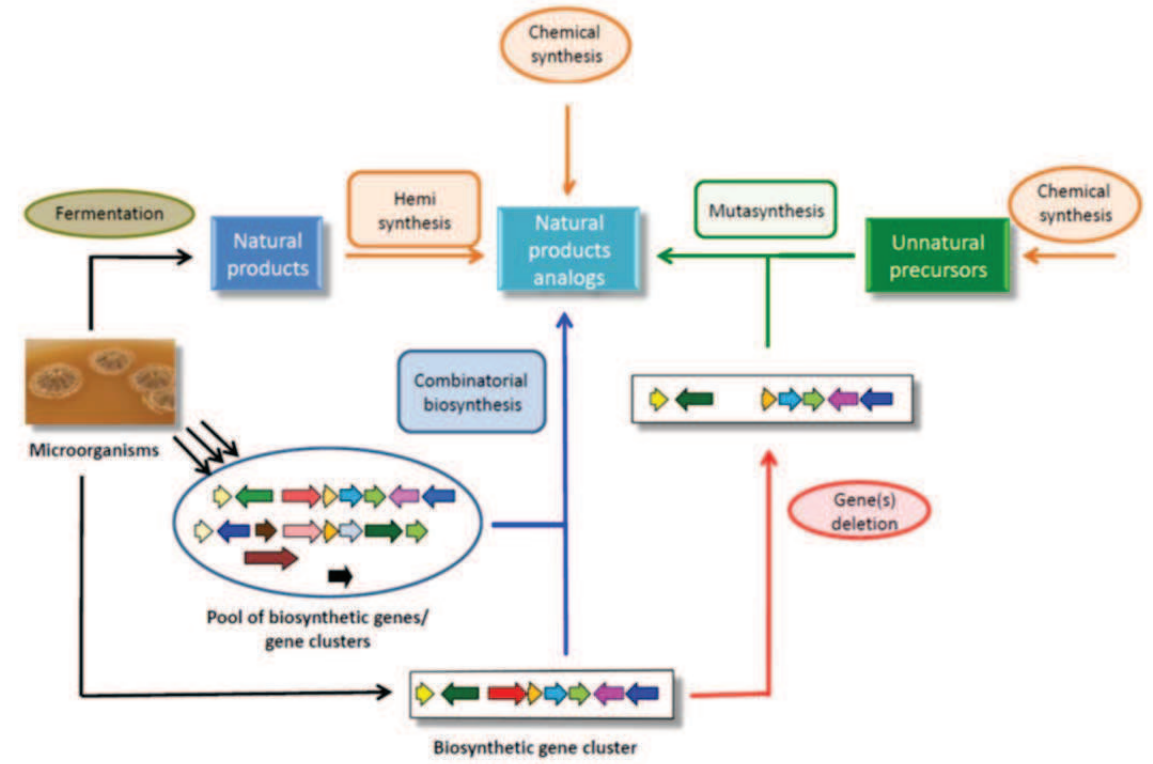



**Figure I-51 Approches pour le développement d'analogues d'aminoglycosides**
(adaptée de [270])

Un des objectifs du sujet de thèse qui m'a été confié visait le développement de la mutasynthèse comme approche alternative pour la synthèse d'analogues d'aminoglycosides. Cette stratégie combine en effet de façon avantageuse les avancées en synthèse organique et les progrès de la biologie moléculaire et de la microbiologie dans la modification des voies biosynthétiques des aminoglycosides. Elle permet ainsi l'exploitation d'un espace de diversité chimique non encore exploré par l'évolution tout en utilisant la richesse, l'efficacité et la versatilité des systèmes biologiques.

# IV - Projet de thèse

Le projet de thèse qui m'a été confié s'inscrit dans la thématique générale de la conception de nouveaux ligands d'ARN basés sur l'architecture des aminoglycosides. Il vise à synthétiser des analogues d'aminoglycosides dans le double but de :

- mettre en place une stratégie de mutasynthèse pour l'obtention d'analogues de la néomycine, un aminoglycoside produit par la bactérie *Streptomyces fradiae*
- développer de nouveaux outils d'étude des ARN par RMN du fluor.

La figure I-52 représente de façon schématique les différents aspects de ce projet de thèse.

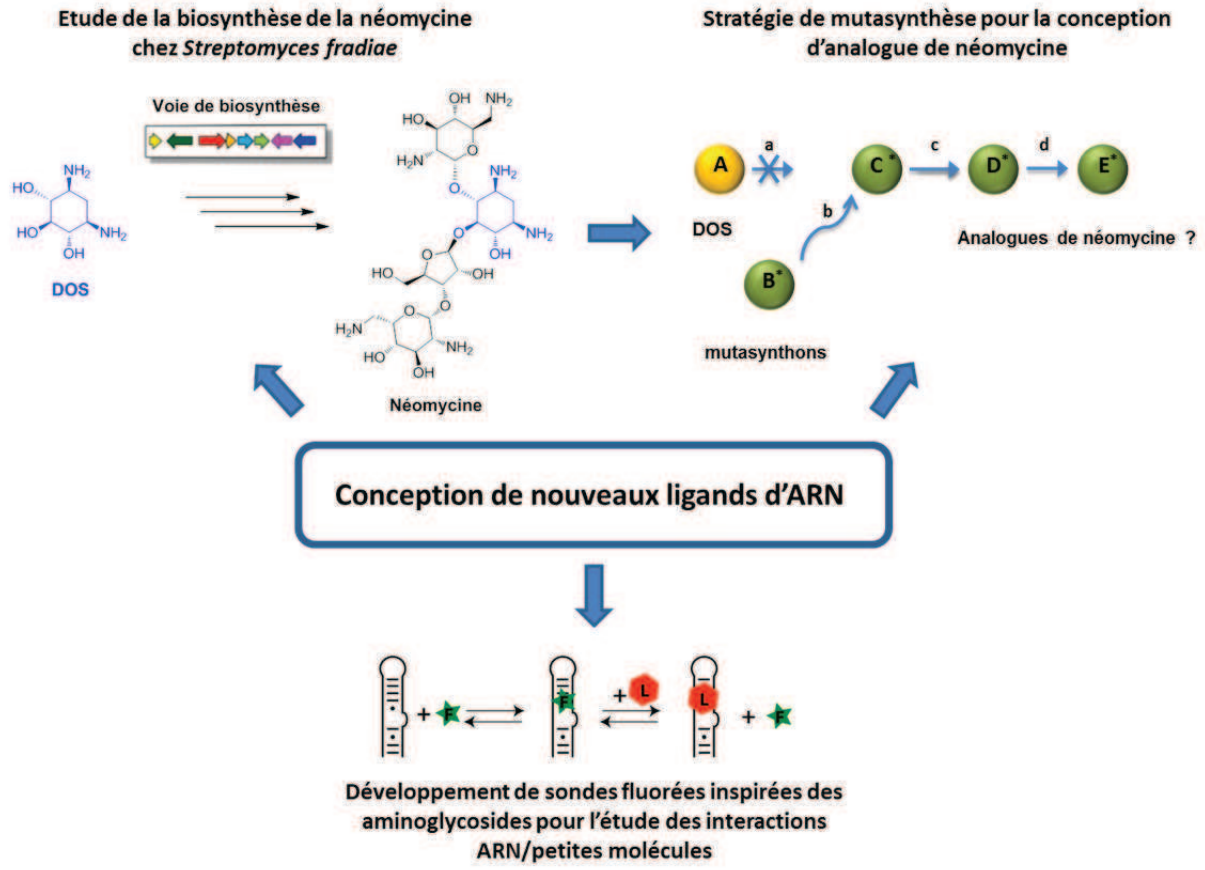



**Figure I-52 Représentation schématique du projet de thèse**

En raison de sa nature très interdisciplinaire à l'interface entre chimie synthétique et analytique, biologie moléculaire et structurale, et microbiologie, ce projet a été effectué en collaboration entre trois équipes, les équipes du Dr Micouin (UMR 8638 CNRS) et du Pr Dardel (UMR 8015 CNRS) à la faculté de Pharmacie de l'université Paris Descartes et l'équipe du Dr Pernodet (UMR 8621 CNRS) à

l'université Paris-Sud d'Orsay, avec le soutien du service de HPLC de la faculté de pharmacie ( A. Regazzetti).

Ce projet s'appuie sur notre compréhension récente des étapes essentielles de la voie de biosynthèse de la néomycine chez *Streptomyces fradiae*. Il repose de façon cruciale sur la disponibilité d'outils de génie génétique permettant la modification de cette voie de biosynthèse et sur la synthèse de molécules pouvant être pris en charge par cette voie modifiée en tant que précurseurs exogènes.

La première partie est consacrée à l'étude de la voie de biosynthèse de la néomycine chez *S.fradiae* et à nos tentatives d'expression hétérologue chez *E.coli* de certaines des enzymes clés de cette voie de biosynthèse et de voies de biosynthèse apparentées.

La deuxième partie aborde plus spécifiquement la stratégie de mutasynthèse pour la création de nouveaux analogues de néomycine. Après un bref survol des travaux qui ont initié nos recherches, nous nous intéresserons aux progrès accomplis vers la mise en place d'une telle stratégie dans des bactéries *S.fradiae* génétiquement modifiées.

La troisième partie décrit l'utilisation d'analogues fluorés d'aminoglycosides comme outils d'étude des ARN par RMN du fluor. Après avoir rappelé l'intérêt de la RMN du fluor dans l'étude des ARN et les travaux préliminaires effectués au laboratoire, nous montrerons comment ces sondes fluorées permettent non seulement d'avoir accès à la structure et la dynamique de différents ARN mais aussi de mesurer l'affinité de certains de leurs ligands.

# II. Expression hétérologue d'enzymes de biosynthèse d'aminoglycosides

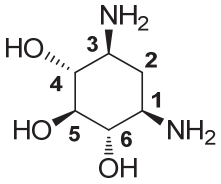
Une connaissance approfondie du fonctionnement de la voie de biosynthèse de la néomycine chez son producteur naturel, *Streptomyces fradiae*, nous semblait essentielle au développement de nouveaux analogues d'aminoglycosides. Notre objectif au début de ma thèse a donc été la production par expression hétérologue chez *Escherichia coli* de grandes quantités de certaines des enzymes clés de cette voie biosynthétique. Nous espérions les obtenir dans un état de solubilité, de pureté et de stabilité compatible avec leur étude biochimique et structurale (RMN et cristallographie). Nous envisagions ensuite d'altérer par mutagénèse la spécificité de certaines de ces enzymes afin de les utiliser dans une voie de biosynthèse modifiée produisant des analogues d'aminoglycosides.

Nous nous sommes particulièrement penchés sur les premières étapes de fonctionnalisation du cycle désoxystreptamine (DOS), cycle central commun à la plupart des aminoglycosides (voir figure II-1 et II-2). Ces étapes sont aussi présentes dans les voies de biosynthèse d'aminoglycosides apparentés à la néomycine. Nous avons donc également travaillé sur des enzymes impliquées dans la biosynthèse de la gentamicine chez *M. echinospora* et dans celle de la butirosine chez *B.circulans*.

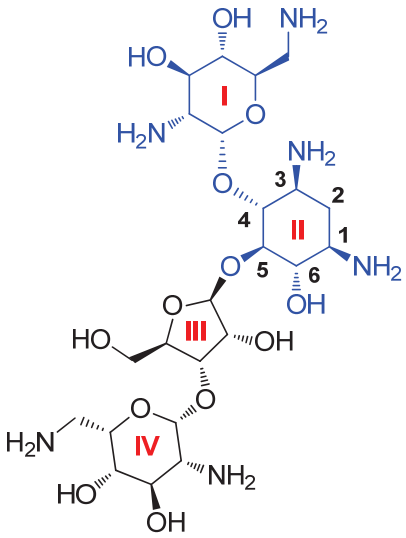
Dans la mesure où ces travaux n'ont pas complètement abouti, nous avons fait le choix de n'en donner dans ce manuscrit qu'un aperçu se concentrant sur la démarche générale sans entrer dans les détails expérimentaux tout en mettant en évidence les difficultés rencontrées. Les résultats obtenus nous amèneront à proposer une hypothèse sur le fonctionnement de la voie de biosynthèse de la néomycine chez *S.fradiae*.

# I - Biosynthèse de la néomycine et d'aminoglycosides apparentés

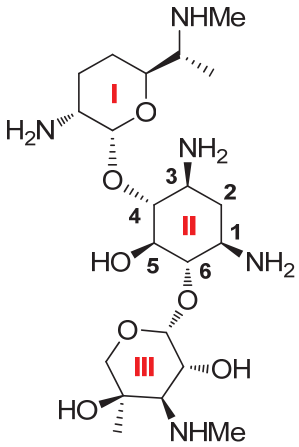
La plupart des aminoglycosides sont organisés autour du cycle central DOS, sur lequel sont greffés divers sucres aminés en position 4 et 5 (néomycine, butyrosine…) ou 4 et 6 (gentamicine…). La biosynthèse de la plupart des aminoglycosides passe par des intermédiaires communs pseudo-dissacharidiques de type néamine, formés par glycosylation de la DOS (voir figure II-1).



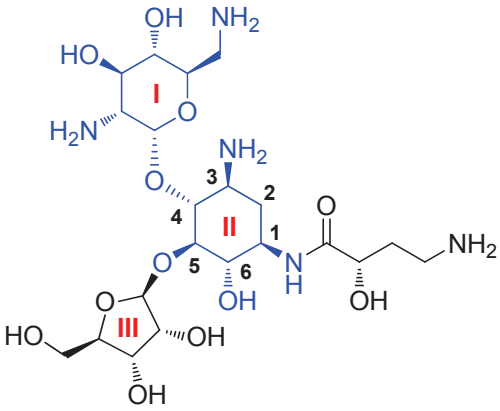
**désoxystreptamine (DOS, 1)**



**néomycine 3**



**gentamicine 45**



**butirosine 46**

**Figure II-53 Structure de la DOS et de différents aminoglycosides**
Le fragment néamine est représenté en bleu. Les premières étapes de la biosynthèse de ces aminoglycosides à partir de la DOS sont communes à de nombreuses souches productrices.

Historiquement, la biosynthèse des aminoglycosides a tout d'abord été étudiée par des expériences de marquage métabolique avec des précurseurs radioactifs ou enrichis en isotopes utilisables en RMN ($^{13}C$, $^{15}N$, $^{2}H$, …), sur des souches sauvages productrices d'antibiotique ou sur des souches mutantes (pour plus d'informations, voir la très bonne revue de Spencer[271]). Plus récemment, les progrès du séquençage ont permis d'obtenir les séquences d'un grand nombre de clusters de gènes de biosynthèse d'aminoglycosides apparentés. Dans le cas de la néomycine, dont le producteur naturel est *Streptomyces fradiae* (référence : ATCC 10745, NCIMB 8233 ou NBRC 12773), on dispose de plusieurs séquences complètes de l'ensemble des gènes impliqués dans la voie de biosynthèse[272,273] (référence EBI (European Bioinformatics Institute) : AJ843080, AJ629247, AJ786317 et AB211959).

Ces informations, couplées à des études bioinformatiques et biochimiques de certaines enzymes clés, ont clarifié l'organisation globale des voies de biosynthèse des aminoglycosides corrigeant certaines des erreurs initiales d'interprétation[274]. En particulier, les groupes de Spencer et Eguchi ont récemment proposé une voie de biosynthèse permettant d'expliquer la formation de la néamine qui serait commune à de nombreux aminoglycosides[275,276] (voir figure II-2). La formation de DOS à partir de glucose-6-phosphate est maintenant très bien comprise[272] mais il reste encore des zones d'ombres concernant le fonctionnement des enzymes responsables de la formation de la néamine à partir de la DOS. La néamine est identifiée comme le point d'embranchement à partir duquel la biosynthèse des aminoglycosides fonctionnalisés en position 4 et 5 diverge de celle des aminoglycosides fonctionnalisés en position 4 et 6[277] (voir figure II-2).

De façon intéressante, si certaines de ces enzymes ont pu être exprimées chez *E.coli* puis étudiées *in vitro* dans les équipes de Spencer et Leadley[278,279], d'Eguchi [280] et de Sohng [281], aucune étude n'a permis de résoudre leur structure. On peut citer en particulier le cas de la glycosyltransferase Neo8 ou celui des glycosyltransférases apparentées, GenM1 ou BtrM, qui n'ont jamais pu être surexprimées ni purifiées pour leur étude *in vitro*. Notre connaissance de ces enzymes se réduit donc à des expériences de complémentation chimique sur des mutants de producteurs naturels et à des études d'activité à partir de lysats cellulaires.

Après avoir effectué quelques essais d'expression hétérologue des enzymes Neo8, Neo16, Neo11 et Neo18, responsables de la biotransformation de la DOS en néamine chez *S.fradiae*, nous nous sommes concentrés sur les deux premières étapes de cette biosynthèse : la glycosylation de la DOS par la *N*-acétyl-glucosamine catalysée par la glycosyltransferase Neo8, et la déacétylation du produit formé par la déacétylase Neo16 aboutissant à la formation de paromamine (voir figure II-2). La première partie de ce chapitre exposera nos essais d'expression hétérologue de ces protéines, nos

travaux de développement d'un outil de suivi de l'activité biologique de Neo8 et de production d'un anticorps reconnaissant spécifiquement cette protéine. Dans la deuxième et la troisième partie, nous nous intéresserons aux homologues de Neo8 et de Neo16 présents dans deux voies de biosynthèse d'aminoglycosides apparentés : les glycosyltransférase GenM1 (biosynthèse de la gentamicine) puis BtrM (biosynthèse de la butirosine) et la déacetylase BrtD (biosynthèse de la butirosine).

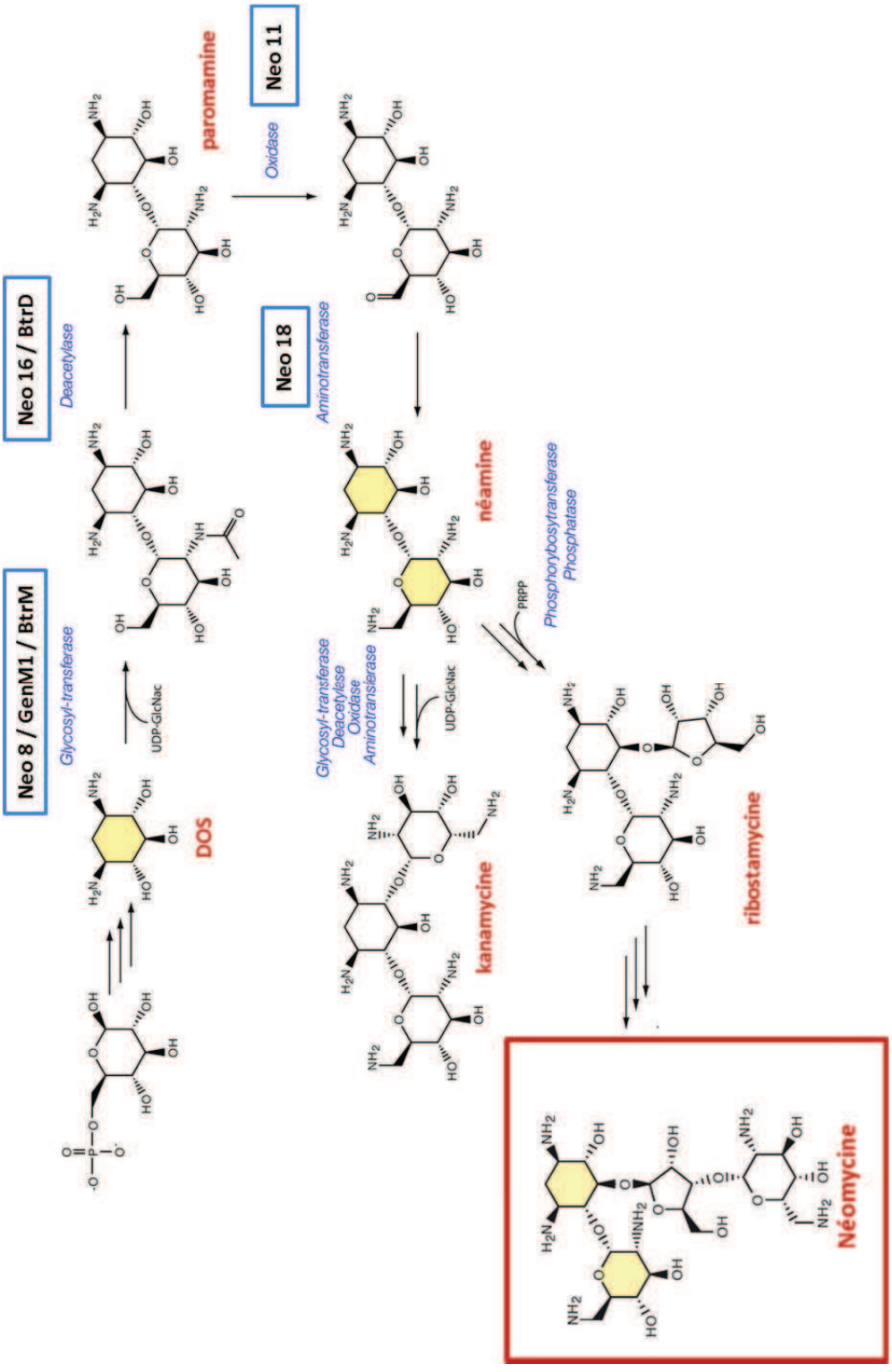



**Figure II-54 Voie de biosynthèse de la néamine à partir du glucose-6-phosphate**

# II - Résultats personnels

## *A - Stratégies employées pour la surexpression hétérologue et la purification de protéines recombinantes*

Il existe de nombreux paramètres sur lesquels on peut jouer lors de l'expression hétérologue d'une protéine chez *E.coli*. Nous allons rapidement passer en revue les stratégies que nous avons mises en place pour essayer d'améliorer la solubilité des protéines exprimées ainsi que les rendements de leur production.

La première question qui se pose lorsque l'on cherche à surexprimer une protéine de façon hétérologue est le système d'expression que l'on va choisir et en particulier le choix du promoteur. Dans notre cas, en dehors de cas explicitement indiqués, nous avons utilisé des plasmides de type pET (pET15b et pET22 notamment) qui utilisent le promoteur du bactériophage T7. Celui-ci permet une surexpression très importante de la protéine placée sous son contrôle lors de l'induction de l'expression par ajout au milieu de culture d'IPTG (isopropyl β-D-1-thiogalactopyranoside, un analogue synthétique non hydrolysable du lactose). Ces systèmes d'expression hétérologues permettant des niveaux de production très élevés sont particulièrement adaptés aux études structurales qui nécessitent de grandes quantités de protéines solubles. Un de leurs inconvénients est qu'ils ne permettent pas de réguler finement la production de protéine recombinante (production « tout ou rien »). Les systèmes d'expression utilisant les promoteurs *lacUV5* ou *ara* permettent un meilleur contrôle de la production de la protéine d'intérêt. Ils montrent cependant des fuites importantes (production de la protéine même en l'absence d'inducteur) et n'ont pas été utilisés dans cette étude. Nous avons par ailleurs réalisé un certain nombre de constructions chimériques de nos protéines d'intérêt, leur ajoutant soit des étiquettes d'affinité (His-tag, …) pour faciliter leur purification, soit des « partenaires » pour augmenter leur production sous forme soluble (par exemple , les protéines NusA ou GST sur lesquelles nous reviendront ultérieurement).

Le deuxième paramètre important dans la surexpression d'une protéine est le choix du génotype de la souche bactérienne que l'on va utiliser. Les différentes souches d'*E.coli* disponibles peuvent en effet être plus ou moins adaptées en fonction des propriétés des protéines que l'on veut produire. Dans notre cas, 4 souches d'*E.coli* différentes ont été essayées dans la plupart de nos essais :

- La souche BL21 (DE3), qui contient dans son génome une copie du gène codant pour l'ARN polymérase du bactériophage T7 sous le contrôle du promoteur lacUV5, inductible par l'IPTG. Cette souche très fréquemment utilisée permet de très hauts niveaux d'expression des protéines à travers l'utilisation de vecteurs d'expression utilisant des promoteurs du

phage T7, inductible à l'IPTG (système pET par exemple). Par ailleurs, cette souche est déficiente pour les protéases Lon et OmpT, ce qui augmente la stabilité des protéines exprimées.

- La souche BL21 (DE3) pLysS, qui est semblable à la souche précédente à la différence qu'elle contient également un plasmide codant le lysozyme T7, un inhibiteur naturel de la T7 ARN polymérase[282]. Ceci permet de limiter l'expression stochastique des gènes sous le contrôle du promoteur T7 même en l'absence d'ITPG (fuite). C'est particulièrement utile lorsque la protéine que l'on veut produire est toxique pour la bactérie productrice et inhibe sa croissance.
- La souche Rosetta 2 (DE3), qui est une souche là encore dérivée de la souche BL21. Elle dispose d'un plasmide codant pour 7 ARN de transfert associés à des codons rarement utilisés chez *E.coli*. Le génome de *S.fradiae* étant un des plus riches en G et C connus (plus de 73 % !), l'usage des codons est complétement différent de celui de *E.coli* où le taux de G et C avoisine les 50 %.
- La souche C43, dérivée de la souche BL21, qui permet de bons niveaux d'expression de protéines très toxiques (génome non complétement caractérisé)[283].

Les conditions d'expression proprement dites ont une influence extrêmement importante sur le rendement en protéine recombinante soluble. Voici quelques-uns des paramètres sur lesquels nous avons joué :

- La température de la culture lors de la croissance bactérienne puis lors de l'induction de l'expression hétérologue : diminuer la température de culture lors de l'induction de la production de protéine recombinante permet de ralentir le métabolisme des bactéries. Une des conséquences intéressantes est que la synthèse protéique est ralentie, ce qui permet d'éventuellement adapter la cinétique de production de la protéine d'intérêt à sa vitesse de repliement. En diminuant la température d'induction, on obtient ainsi parfois moins de protéines exprimées sur forme de corps d'inclusion, même si cela réduit aussi souvent la quantité globale de protéine d'intérêt produite. Lorsque nous avons fait varier ce paramètre, nous avons testé 3 températures différentes : 37 °C (température habituelle), 25 °C et 16°C.
- La concentration en molécule inductrice, dans notre cas, l'IPTG. Même si les systèmes pET que nous avons utilisés ne sont pas très sensibles à la concentration en IPTG, nous avons néanmoins souvent testé plusieurs concentrations en IPTG (de 50 à 500 µM final) toutefois sans jamais trouver de différence significative dans la solubilité des protéines exprimées.

- La durée de l'induction : augmenter cette durée peut augmenter le rendement de production hétérologue mais peut aussi favoriser la dégradation de la protéine d'intérêt. Les temps habituels d'induction varient de 1h à plus de 12h (spécialement pour une induction à basse température).
- Le milieu de culture. Nous avons systématiquement utilisé du milieu LB mais il est connu que le milieu de culture peut avoir une influence sur les rendements en production hétérologue. Dans certains cas qui seront exposés ultérieurement, nous avons ajouté au milieu de culture 2 % d'éthanol. Cet ajout avait pour but de provoquer un stress chimique induisant la surexpression de protéines chaperonnes bactériennes endogènes. Il permet parfois d'augmenter la quantité de protéine d'intérêt soluble. Nous n'avons jamais observé d'effet significatif dans notre cas.

Le rendement en protéine recombinante soluble est aussi grandement influencé par les conditions de lyse des bactéries. Nous avons la plupart du temps utilisé la sonication. Dans certains cas néanmoins (surexpression de neo8 par exemple), nous avons comparé cette méthode de lyse avec d'autres méthodes supposées moins agressives : la lyse chimique (détergent et/ou lysozyme) et la lyse par presse de French, sans jamais toutefois constater de différence significative de concentration en protéines solubles. Le tampon de lyse est également très important et dépend de la protéine d'intérêt : l'objectif est de trouver des conditions dans lesquelles la protéine va être la plus stable et la plus soluble possible. Il n'y a pas vraiment de règles permettant de guider le choix et il faut donc faire varier le pH, le type de sels et leur concentration, le tampon utilisé… On peut également ajouter des inhibiteurs de protéases (PMSF, EDTA, cocktails commerciaux…) pour éviter les phénomènes de dégradation des protéines produites. Les différentes modifications de tampon de lyse ne nous ont jamais apporté d'amélioration significative de la quantité de protéine soluble obtenue.

Enfin, le dernier facteur influençant le rendement final en protéine recombinante est la méthode de purification employée. Il y a toujours un compromis à faire entre quantité et pureté de la protéine recombinante. Pour les études de structure en cristallographie ou en RMN, il faut malheureusement à la fois une grande quantité et une grande pureté de la protéine. Bien évidemment, pour que l'étude structurale ait un intérêt, il faut également s'assurer que la protéine purifiée a gardé son activité biologique.

Au cours de nos tentatives d'expression et de purification des différentes protéines de la voie de biosynthèse de la néomycine ou d'aminoglycosides apparentés, nous avons utilisés 5 types de purification de protéines :

- Purification par affinité : c'est un des moyens les plus efficaces de purification des protéines. En particulier, la plupart des constructions que nous avons obtenues possédaient une étiquette d'affinité poly-histidine qui permet leur purification sur colonne de nickel (Ni-NTA), souvent utilisée par nous comme première étape de purification. Nous avons également utilisé des constructions chimériques avec la GST, qui peuvent être purifiées sur une colonne greffée avec du glutathion.
- Purification par taille : la chromatographie d'exclusion stérique permet de séparer les différentes protéines en fonction de leur volume hydrodynamique. Nous avons couramment utilisé des colonnes de type Superdex (Superdex 200 ou Superdex 75).
- Purification par interaction de charges : dans quelques cas, nous avons utilisé des colonnes de type S (S sépharose) ou Q (Ressource Q ou MonoQ) qui ont la particularité d'être chargées positivement ou négativement. Les protéines d'intérêt sont donc plus ou moins retenues en fonction de leur point isoélectrique.
- Purification par hydrophobicité : nous avons effectué quelques essais de purification sur une colonne de type Phenyl-sépharose séparant les protéines en fonction de leur hydrophobicité.
- Purification par précipitation sélective : nous avons souvent utilisé la précipitation des protéines au sulfate d'ammonium pour permettre leur conservation pendant de longues périodes à 4°C. La propension des protéines à précipiter à différentes concentrations en sels peut être également utilisée dans un protocole de purification ; dans notre cas, la précipitation différentielle au sulfate d'ammonium a uniquement été exploitée lors de l'étude du lysat cellulaire de *S.fradiae* sauvage et au cours de la purification des anticorps anti-Neo8 (voir ultérieurement).

Dans tous les exemples que nous allons exposer, nous avons construit les plasmides d'expression en utilisant les techniques usuelles de la biologie moléculaire[284] qui ne seront pas détaillées. Ces plasmides ont été vérifiés par carte de restriction, par PCR avec des amorces spécifiques et par séquençage pour les plasmides finaux. L'identité des protéines exprimées de façon soluble que nous avons purifiées a été vérifiée par spectrométrie de masse (technique PMF (« Peptide mass fingerprinting », plate-forme de l'institut Cochin). La concentration des échantillons purifiés a été estimée par mesure de l'absorption à 280nm d'une solution diluée, en utilisant le coefficient d'extinction théorique de la protéine (calculé à partir de sa composition).

## *B - Enzymes de biosynthèse de la néomycine chez S.fradiae : Neo8, Neo11, Neo16, Neo18*

### 1) Clonage et essais de surexpression

Nous avons dans une première étape amplifié par PCR à partir du génome de *S.fradiae* sauvage les gènes Neo8, Neo11, Neo16 et Neo18. Pour ce faire, nous avons extrait l'ADN total de la souche suivant un protocole standard fondé sur une lyse par le détergent SDS (sodium dodecyl sulfate) et une digestion par la protéinase K, suivies d'une extraction par le mélange phénol/chloroforme et d'une précipitation de l'ADN génomique par l'éthanol.

Le génome de *S.fradiae* étant très riche en nucléotides G et C, un travail d'optimisation des conditions d'amplification par PCR a été nécessaire (test de plusieurs polymérases, modification de la température d'hybridation, optimisation du tampon de PCR, ajout de formamide…) (voir figure II-3). Pour chaque gène d'intérêt, nous avons choisi deux amorces PCR aux extrémités du gène et nous y avons ajouté en 5' les sites des enzymes de restriction *BamHI* ou *NdeI* pour faciliter le clonage ultérieur des fragments amplifiés.

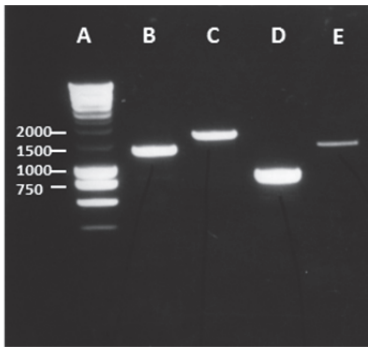



**Figure II-55 Amplification par PCR des gènes neo8 (B), neo11 (C), neo16 (D), neo18 (E).**
Colonne A : marqueur de poids moléculaire (les chiffres indiqués correspondent aux nombres de bases des fragments d'ADN)

Les gènes d'intérêt amplifiés ont été clonés dans le plasmide pET-15b, soit directement par digestion *BamH1*/*NdeI* des fragments amplifiés puis ligation dans un plasmide digéré par les mêmes enzymes de restriction, soit en passant par le système de clonage TOPO (Life technologies).

Une fois en possession de ces constructions, nous avons effectué un criblage rapide des conditions d'expression de chacune de ces protéines en utilisant des conditions classiques de surexpression (milieu LB, induction à 37°C pendant 2 h avec 500 µM d'IPTG dans des souches BL21 (DE3), lyse par sonication). Comme on peut en voir des exemples sur les figures II-4 et II-5, toutes les

protéines étaient surexprimées, malheureusement pas de façon soluble mais uniquement sous forme de corps d'inclusion.

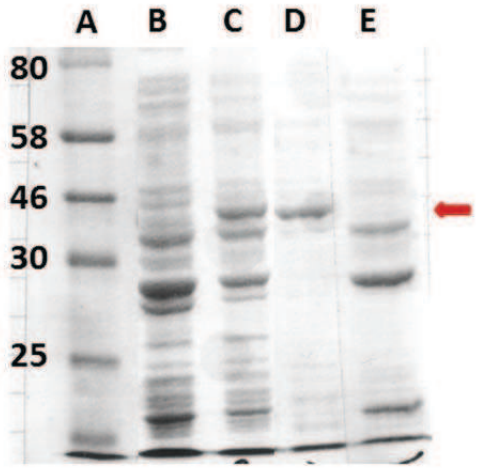



**Figure II-56 Test d'expression de la protéine Neo8 dans la souche BL21 (DE3).**
A : marqueur de poids moléculaire (en kDa) ; B : pas d'induction ; C : extrait total après induction avec 500 µM IPTG final ; D : partie insoluble des protéines (corps d'inclusion) ; E : protéines solubles Sur ce gel de polyacrylamide, en comparant les colonnes B et C, on voit clairement que la protéine Neo8 (M ≈ 42 kDa) est bien surexprimée (flèche rouge), mais uniquement sous forme de corps d'inclusion (colonne D).

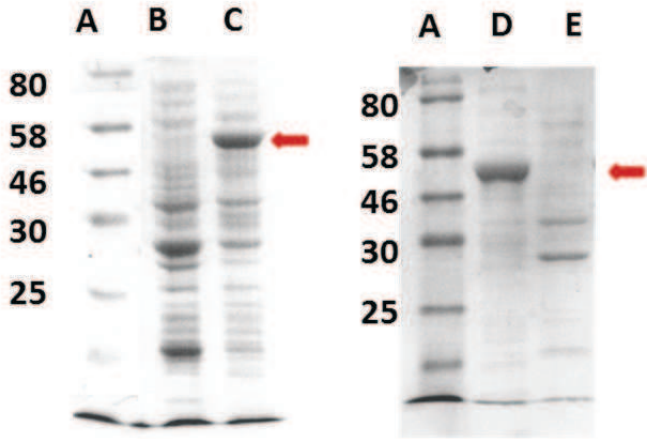



**Figure II-57 Test d'expression de la protéine Neo11 dans la souche BL21 (DE3).**
A : marqueur de poids moléculaire (en kDa) ; B : pas d'induction ; C : extrait total après induction avec 500 µM IPTG final ; D : partie insoluble des protéines (corps d'inclusion) ; E : protéines solubles. En comparant les colonnes B et C, on voit clairement que la protéine Neo11 (M ≈ 58 kDa) est bien surexprimée (flèche rouge), mais uniquement sous forme de corps d'inclusion (colonne D).

A l'issue de ce criblage rapide, nous avons décidé de nous concentrer sur la protéine Neo8 en raison de son intérêt biologique mais aussi parce que nous étions intrigués par les échecs préalables de sa purification en dépit d'essais répétés. L'équipe d'Eguchi a notamment travaillé sur l'enzyme Neo8 sans jamais réussir à la purifier[280] et les mêmes problèmes ont été rencontrés par l'équipe de Spencer et Leadley (communication personnelle).

En suivant les principes généraux exposés dans la partie précédente, nous avons testé un large éventail de conditions de surexpression différentes pour Neo8 :

- changement de la souche bactérienne (BL21(DE3), BL21(DE3) pLysS, Rosetta2, C43)
- changement des conditions de surexpression :
    - température (16°C, 25°C, 37°C)
    - concentration en IPTG (de 50 à 500 µM final)
    - composition du milieu de culture (milieu LB avec ou sans ajout d'ethanol 2 % final)
    - durée d'induction (de 1h à plus de 12h pour les cultures à 16°C)
- changement des méthodes de lyse après induction (sonication, lyse par presse de French, lyse chimique)

Nous avons également testé différents vecteurs d'expression (passage de pET-15b à pET-22, c'est-à-dire passage d'une étiquette poly-histidine N-terminale à une étiquette poly-histidine C-terminale) et tenter de co-exprimer Neo8 avec un plasmide permettant la surexpression des protéines chaperonnes GroES et GroEL, à l'instar de ce qui avait été fait précédemment pour d'autres protéines de *Streptomyces*[280,285].

Pour éviter de passer à côté d'une faible quantité de protéine exprimée de façon soluble non visible à l'œil nu sur gel de polyacrylamide, nous avons effectué des expériences de Western blot avec des anticorps spécifiques des étiquettes poly-histidines présentes sur nos protéines. Ces expériences, plus sensibles qu'un simple gel de polyacrylamide, n'ont pas mis en évidence de fraction soluble des protéines d'intérêt quelque soient les conditions de leur expression.

Devant les difficultés rencontrées dans l'expression hétérologue de Neo8 sous forme soluble chez *E.coli*, nous avons tenté de purifier la protéine en conditions dénaturantes à partir des corps d'inclusion, puis d'essayer de la replier de façon à l'obtenir sous forme soluble et bioactive.

### 2) Purification de Neo8 sous forme dénaturée et expériences de repliement

Une expression sous forme d'agrégats insolubles nommés corps d'inclusion est souvent rencontrée pour les protéines toxiques pour E coli ou incorrectement repliées. Cette expression sous forme de corps d'inclusion s'accompagne souvent d'une perte de leur activité biologique.

Pour obtenir une protéine soluble, il est possible de récupérer les corps d'inclusion et de les dénaturer dans l'urée (8 M). La protéine d'intérêt ainsi dénaturée peut, dans les cas favorables, être repliée par renaturation dans un tampon adéquat. Les avantages de la production en corps d'inclusion sont que les rendements totaux en protéines sont souvent très importants et que les protéines sont protégées des protéases bactériennes. Par ailleurs, la purification de protéine sous forme dénaturée à partir de corps d'inclusion est souvent très simple. Ça a été le cas pour Neo8 : une simple dénaturation des corps d'inclusion dans un tampon contenant de l'urée (Tris 50 mM pH 8, urée 8M, 5 mM DTT) suivi d'une purification sur colonne Ni-NTA selon le protocole conseillé par le fabricant (Invitrogen) a permis d'obtenir de grandes quantités de Neo8 dénaturée pure (voir figure II-6).

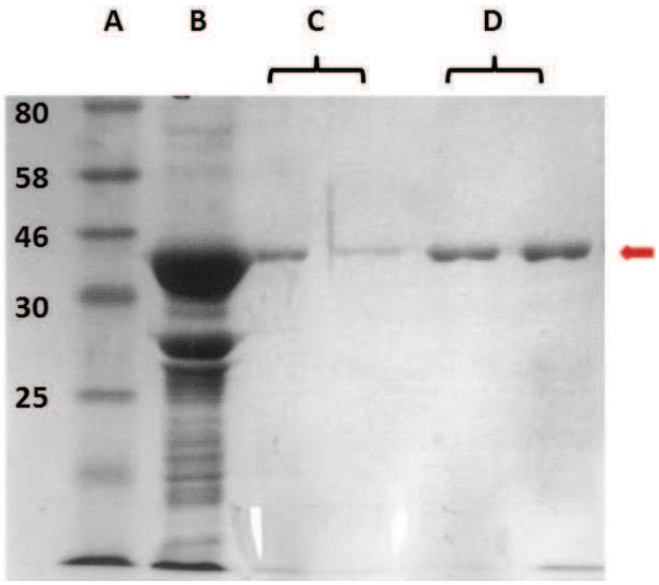



**Figure II-58 Purification par colonne d'affinité (Ni-NTA) de la protéine Neo8 sous forme dénaturée**

A : marqueur de poids moléculaire (en kDa) ; B : corps d'inclusion dénaturé ; C : lavages ; D : élution
La protéine Neo8 (M ≈ 42 kDa) est indiquée par une flèche rouge.

L'étape suivante a été d'essayer de replier la protéine dénaturée en passant du tampon d'élution de colonne contenant de l'urée (urée 8M, 50 mM Tris pH 8, 500 mM imidazole, 500 mM NaCl) à un tampon sans urée. Pour ce faire, nous avons testé deux méthodes différentes :

- La diminution progressive de la concentration en urée par dialyses successives contre un tampon identique à l'exception de la concentration en urée (4M, 2M, 1M, 0.5M, 0.25M, 0M).
- La dilution très rapide d'un petit volume de solution concentrée de protéine Neo8 dans le tampon de l'urée dans un grand volume d'un autre tampon ne contenant pas d'urée (« flash-dilution »).

Dans le premier cas, l'essentiel de la protéine a commencé à précipiter dès la dialyse contre le tampon contenant 2M d'urée (sans doute aux alentours de 3M urée). Nous avons cependant récupéré des traces de protéine soluble en fin de dialyse sans urée. Cette fraction a été concentrée afin de permettre l'évaluation de son activité biologique car seule celle-ci permet de démontrer que le repliement obtenu est correct (voir partie suivante). Malheureusement, la protéine renaturée s'est montrée inactive. Dans le cas de la « flash-dilution », nous n'avons jamais réussi à détecter de protéine soluble malgré les quelques 20 conditions testées (pH de 3.5 à 9.6, tampons et concentration en sels variés).

### 3) Test de l'activité biologique de Neo8

Afin de guider à la fois nos essais d'expression hétérologue et nos expériences de repliement de la protéine Neo8, nous avons voulu développer un test de son activité biologique Nous nous sommes inspirés de travaux récents visant à évaluer par absorption UV l'activité biologique d'une glycosyltransferase de *Pasteurella multocida* utilisant, comme Neo8, l'UDP-*N*-acetylglucosamine (UDP-GlucNAc) comme substrat[286] (voir figure II-2). Notre test repose sur une cascade de réactions enzymatiques couplées mettant en jeu, en plus de la glycosyltransferase Neo8 ou d'une glycosyltransferase apparentée, la pyruvate kinase (Pyr. Kin.) et la lactate déshydrogénase (L-LDH). La décroissance du signal de la forme réduite du nicotinamide adénine dinucléotide (NADH) en UV (340 nm) permet de déduire la cinétique de consommation de l'UDP-*N*-acetylglucosamine et donc d'évaluer l'activité enzymatique de la glycosyltransférase étudiée (voir figure II-7). Alternativement, on peut suivre la disparition du NADH en fluorescence.

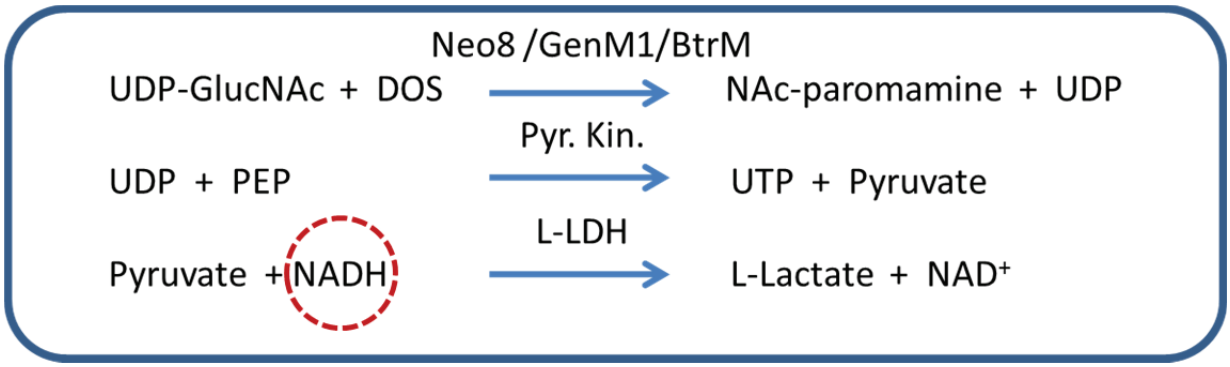



**Figure II-59 Test enzymatique de l'activité biologique de la glycosyltransférase Neo8**
L'activité de la glycosyltransferase testée est corrélée à la décroissance du signal du NADH en UV. PEP désigne le phosphoénol pyruvate.

Le protocole de ce test est simple : au tampon du test d'activité (20 mM Hepes pH 7.5, 5 mM $MgCl_2$, 50 mM KCl, 5 mM DTT, 0.7 mM PEP, 140 µM NADH), on ajoute les deux substrats de la glycosyltransferase, l'UDP-GlucNAc (2 mM) et la DOS (2 mM) ainsi que les 2 enzymes Pyr.Kin. (7.5 U) et L-LDH (15 U). Le mélange est agité dans une cuve UV thermostatée à 25 °C. On peut ensuite ajouter le lot de protéine dont on veut tester l'activité de glycosyltransferase. Pour vérifier que la cascade de réaction enzymatique est bien fonctionnelle, on peut ajouter dans le milieu de l'uridine diphosphate (UDP) qui induit une diminution du signal du NADH. Un des inconvénients de ce test est qu'il ne peut être utile que sur des échantillons suffisamment purs. En effet, il existe beaucoup d'enzymes bactériennes utilisant l'UDP-GlucNAc comme substrat ou conduisant à la formation d'UDP ce qui induit un bruit de fond très important. Une autre difficulté à laquelle nous nous sommes heurtés lors du développement de ce test d'activité est notre manque global de connaissance sur la biochimie de ces protéines. En particulier, on ne connait pas à priori les cofacteurs nécessaires au bon fonctionnement des glycosyltransférases testées. Certaines études ayant montré le rôle important du magnésium dans l'activité de glycosyltransférases apparentées[280], nous avons toujours ajouté 5mM de $MgCl_2$ au tampon de test de l'activité biologique. Il n'est cependant pas possible d'exclure la possibilité d'être passé à côté d'un cofacteur essentiel. De la même façon, si un partenaire protéique est essentiel au bon fonctionnement de la glycosyltransférase testée, notre test n'est pas en mesure de détecter d'activité biologique.

Nous avons par la suite systématiquement utilisé ce test pour l'ensemble des protéines solubles purifiées dont on voulait tester l'activité de glycosyltransférase. Comme nous le verrons par la suite,

nous n'avons malheureusement jamais été en mesure de détecter d'activité biologique suffisamment importante pour se démarquer du bruit de fond.

### 4) Développement d'un anticorps anti-Neo8

Au vu des difficultés rencontrées dans l'expression hétérologue de la protéine Neo8 sous forme soluble, nous sommes revenus vers les expériences d'expression hétérologue précédemment décrites dans des situations similaires. Nous avons alors constaté que nous n'étions évidemment pas les seuls à rencontrer ces difficultés puisqu'aucun groupe n'a à ce jour réussi à surexprimer et purifier Neo8 ou une glycosyltransférase homologue.

L'examen attentif de la littérature nous conduit à proposer une hypothèse d'organisation de la voie de biosynthèse de la néomycine (sans doute valable pour les aminoglycosides apparentés) dans laquelle la glycosyltransférase Neo8 ferait partie d'un complexe biosynthétique multiprotéique à l'instar de ce qui a récemment été démontré pour la biosynthèse des purines chez les eucaryotes[287] ou pour la biosynthèse des polyketides chez les actinomycètes comme les *Streptomyces*[197288] (polykétide synthases de type II). Dans ces assemblages biosynthétiques, il existe un réseau complexe d'interactions entre les différents partenaires protéiques qui influe sur la stabilité à la fois du complexe dans son ensemble et de chacune de ces composantes mais aussi sur l'activité biologique des différentes enzymes. Dans cette hypothèse, l'expression hétérologue de Neo8 sans ses partenaires conduirait inévitablement à une enzyme inactive, exprimée sous forme de corps d'inclusion.

Il existe dans la littérature un faisceau d'observations indirectes venant étayer cette hypothèse. Tout d'abord, de façon particulièrement significative, les conditions d'expression hétérologue des enzymes homologues à Neo8 les plus favorables, où l'activité glycosyltransférase a pu être mise en évidence sans toutefois aller jusqu'à permettre une purification de la protéine d'intérêt, impliquent soit des souches bactériennes proches de *S.fradiae*[289,290] (*S.venezueliae* ou *S.lividans*) où on peut potentiellement trouver les éventuels partenaires protéiques nécessaires (chaperonnes, …), soit une coexpression avec d'autres enzymes de la voie de biosynthèse[281,291].

Par ailleurs, au cours de l'étude de la biosynthèse de la néomycine chez *S.fradiae*, l'équipe de Spencer s'est aperçue que la suppression du gène neo8 perturbait de façon non négligeable les étapes ultérieures de la voie de biosynthèse en dépit d'une complémentation chimique avec des métabolites situés en aval de l'étape supprimée[275]. Cette observation rejoint nos propres travaux sur la délétion d'une enzyme essentielle à la biosynthèse de la DOS, la protéine Neo6 (enzyme Neo6, voir

chapitre III, partie II E 1 b), qui conduit là aussi à une modification de la spécificité et de l'efficacité des étapes biosynthétiques situées en aval de la formation de la DOS. La perturbation de l'ensemble de la voie de biosynthèse après modification d'un des partenaires enzymatique est un autre argument important en faveur de l'existence d'un complexe multienzymatique.

Pour éclaircir cette question, nous avons produit des anticorps polyclonaux spécifiques de Neo8 dans le but de nous permettre d'étudier les éventuels partenaires protéiques de Neo8 en conditions natives par immunoprécipitation ou western blot sur un lysat cellulaire de la souche sauvage *S.fradiae*.

Pour ce faire, nous avons purifié en conditions dénaturantes la protéine Neo8 à l'échelle de 20 mg. Une partie de cet échantillon a été envoyé à l'entreprise Genscript pour immuniser 2 lapins. Nous avons récupéré leur sérum à partir duquel nous avons purifié les anticorps en utilisant un protocole combinant précipitation sélective au sulfate d'ammonium et chromatographie d'affinité sur colonne greffée à la protéine A. Comme on peut le voir sur la figure II-8, les anticorps obtenus ont bien été séparés des autres protéines contenues dans le sérum (albumine notamment).

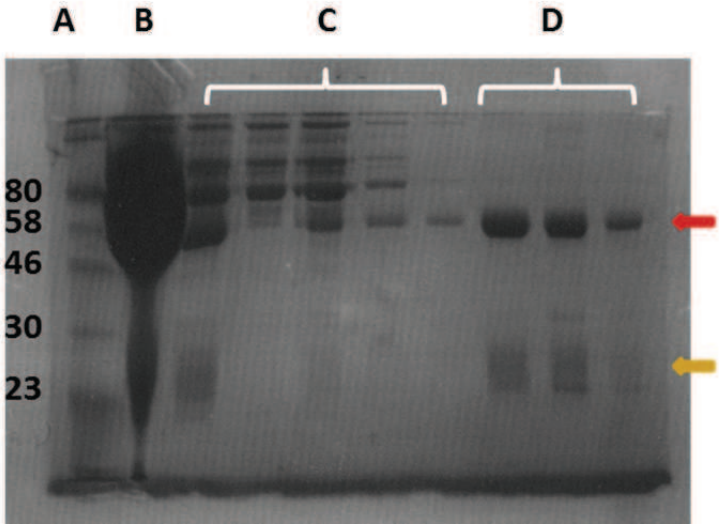



**Figure II-60 Purification des anticorps polyclonaux anti-Neo8 à partir de sérum de lapin par colonne d'affinité à la protéine A.**
A : marqueur de poids moléculaire ; B : sérum total ; C : lavages ; D : fractions d'élution de la colonne d'affinité. On distingue bien les chaines lourdes (flèche rouge) et les chaines légères (flèche jaune) des anticorps.

L'étape suivante a été de vérifier que ces anticorps polyclonaux reconnaissaient bien la protéine Neo8 dénaturée. Nous avons donc effectué une expérience de western blot sur un lot de protéine Neo8 purifiée sous forme dénaturée et mais aussi sur deux lysats cellulaires de BL21(DE3), induites

ou non pour la surexpression de Neo8 sous forme de corps d'inclusion. Les anticorps ont reconnu la protéine Neo8 dans l'échantillon purifié comme on pouvait s'y attendre mais aussi directement dans le lysat cellulaire correspondant à la souche de *E.coli* induite (voir figure II-9).

Nous nous sommes également demandé si ces anticorps seraient capables de reconnaitre la protéine Neo8 directement dans un lysat de *S.fradiae*. Pour tester cette hypothèse, nous avons effectué une culture de *S.fradiae* dans des conditions où il y a production de néomycine c'est-à-dire où la protéine Neo8 est exprimée (voir la partie expérimentale de la partie mutasynthèse). On peut constater dans la figure II-9 qu'on observe bien une bande de faible intensité à une masse correspondant à la protéine Neo8 (flèche rouge). On peut aussi constater la présence de nombreuses bandes supplémentaires comme c'est souvent le cas pour des anticorps polyclonaux obtenus à partir de lapins qui n'étaient pas stériles mais avaient déjà été exposés à des antigènes. La purification des immunoglobulines du sérum n'a aucune raison d'améliorer cet aspect qui nécessiterait par exemple une purification fondée sur la reconnaissance de Neo8 (colonne d'affinité ou purification sur gel préparatif).

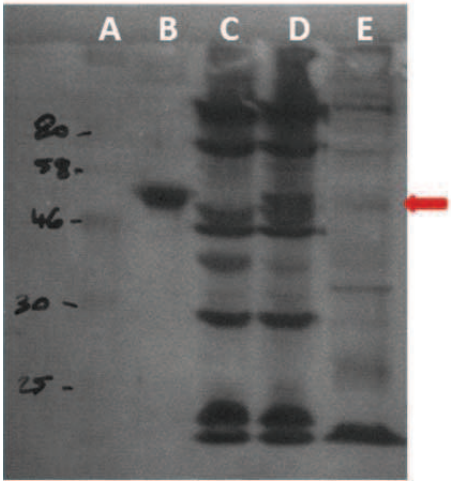



**Figure II-61 Western blot utilisant les anticorps anti-Neo8 purifiés.**
A : marqueur de poids moléculaire ; B : Neo8 purifiée dénaturée ; C : lysat de BL21(DE3) non induite pour la surexpression de Neo8 ; D : lysat de BL21(DE3) induites (production de Neo8 sous forme de corps d'inclusion) ; E : lysat de *S.fradiae* (Neo8 est exprimée car il y a production de néomycine)

Pour pallier à cette difficulté, nous avons réalisé une immunoprécipitation à petite échelle de nos anticorps polyclonaux. Pour cela, nous avons greffé la protéine Neo8 purifiée et dénaturée sur une colonne de sépharose activée par le bromure de cyanogène (conditions recommandées par le fournisseur, GE Healthcare). Cette colonne a ensuite été utilisée pour purifier sélectivement les anticorps dirigés contre Neo8 des autres anticorps du sérum. Ces anticorps immunopurifiés sont

spécifiques et réduisent considérablement le bruit de fond précédemment observé sur un extrait cellulaire total. A ce stade de l'étude, nous disposions donc d'un outil permettant de détecter Neo8 directement dans un lysat de *S.fradiae*. Nous avons voulu utiliser cet outil pour évaluer la taille apparente de Neo8 dans des conditions natives où l'enzyme est impliquée dans la biosynthèse de la néomycine afin d'en déduire la présence de partenaires potentiels. Cette expérience, réalisée par Frédéric Dardel et Paola Frigols-Arroyo, est partie de bactéries *S.fradiae* cultivées dans des conditions où elles produisent de la néomycine, lysées par sonication puis déposées en conditions natives sur une colonne d'exclusion de taille. Les différentes fractions de la colonne ont été analysées par Western Blot en utilisant les anticorps polyclonaux anti-Neo8 immunopurifiés précédemment décrits. Le résultat de cette expérience est présenté dans la figure II-10. On voit assez clairement la présence d'une bande correspondant à Neo8 dans les premières fractions de la colonne, qui correspondent à des objets de taille très importante (M > $10^3$ kDa). Cette observation pourrait être expliquée par l'appartenance de Neo8 à un complexe multienzymatique de taille importante. Elle pourrait cependant aussi être expliquée par une agrégation/dégradation de la protéine Neo8 au cours des étapes de lyse et de purification utilisées. Des expériences complémentaires faisant varier les conditions de lyse et/ou de purification seraient nécessaires pour tester la reproductibilité de cette observation.

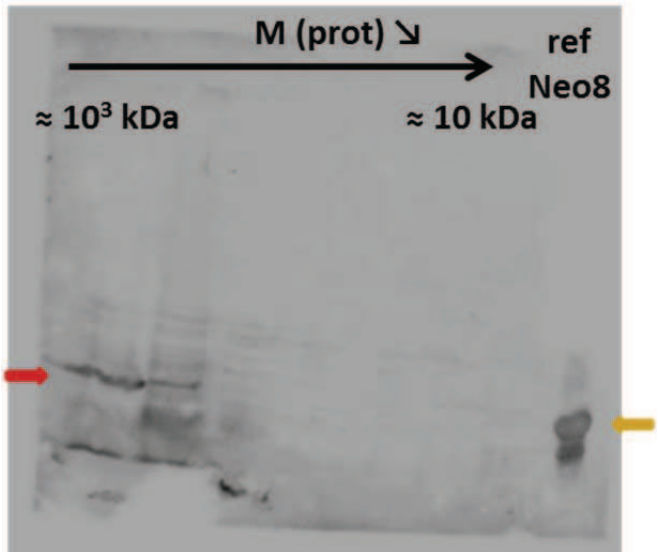



**Figure II-62 Western blot des fractions d'une chromatographie par exclusion de taille sur Superdex 200 d'un lysat cellulaire de *S.fradiae* en conditions natives.**
On distingue une bande correspondant à la taille de Neo8 dans les toutes premières fractions de la chromatographie (flèche rouge), correspondant à des objets de taille très importante (≈ $10^3$ kDa), vraisemblablement des assemblages multiprotéiques. La protéine Neo8 purifiée qui sert de référence est indiquée par une flèche jaune.

En conclusion, les résultats de ces expériences préliminaires sont suffisamment intéressants pour motiver la purification par affinité d'une quantité plus importante d'anticorps anti-Neo8 afin de disposer d'un marquage spécifique de Neo8 qui permettrait 1) de guider par Western Blot la purification de Neo8 à partir d'un lysat cellulaire de *S.fradiae*, 2) de vérifier la taille du (des) partenaire(s) au(x)quel(s) s'associe Neo8 par des expériences d'électrophorèse sur gel en conditions natives, éventuellement suivies d'une deuxième dimension en conditions dénaturantes, et 3) si les anticorps étaient suffisamment affins pour la protéine Neo8 native, d'essayer la purification de Neo8 et de ses éventuels partenaires protéiques par immunoprécipitation sur un lysat cellulaire de *S.fradiae*.

## *C - La glycosyltransférase GenM1 de la biosynthèse de la gentamicine chez M. echinospora*

### 1) Clonage et tentatives de surexpression

En parallèle à nos travaux sur la glycosyltransferase Neo8 de *S.fradiae*, nous nous sommes penchés sur une protéine homologue à Neo8, la glycosyltransferase GenM1 de *M.echinospora*, producteur naturel de la gentamicine. Les premières étapes de la biosynthèse de la gentamicine sont en effet identiques à celles de la néomycine. Ainsi la glycosyltransferase GenM1, qui possède un grand pourcentage d'identité de séquence avec Neo8 (≈ 54 %), catalyse elle aussi la glycosylation de la DOS par la N-acetylglucosamine[289].

Nous avons voulu savoir dans quelle mesure les différences entre Neo8 et GenM1 pourraient être exploitées pour exprimer cette dernière de façon soluble chez *E.coli*. Pour supprimer les éventuelles difficultés liées à une utilisation non optimale des codons du gène *genM1*, nous avons fait synthétiser le gène (entreprise DNA 2.0) après en avoir optimisé l'utilisation des codons pour une expression chez *E.coli*[292]. Le gène *genM1* modifié a ensuite été cloné dans deux plasmides différents, pET-15b et pET-22, de façon à ajouter une étiquette poly-histidine en position N-terminale ou C-terminale. Nous avons ensuite conduit des expériences équivalentes à celles effectuées avec Neo8, jouant sur l'ensemble des paramètres précédemment décrits pour obtenir une protéine exprimée de façon soluble (tests de 4 souches différentes de *E.coli*, variations des conditions de surexpression et

de lyse…)[3]. Malheureusement, là encore, la protéine GenM1 s'est révélée produite très majoritairement sous forme de corps d'inclusion et l'ensemble des conditions d'expression hétérologue testées ne nous a permis d'obtenir qu'un faible pourcentage de protéine soluble, presque indétectable à l'œil nu. Cette très faible expression soluble n'était pas compatible avec les études biochimiques et structurales envisagées mais semblait néanmoins très légèrement meilleure que dans le cas de Neo8. Nous avons donc décidé d'exploiter cette piste en concevant des protéines chimères, où la protéine GenM1 serait greffée à un partenaire protéique censé déplacer l'équilibre vers une forme soluble. Ces travaux feront l'objet de la partie suivante.

### 2) Protéines chimèriques

#### a) GST-GenM1

La construction de protéines chimères, où la protéine d'intérêt est fusionnée avec un partenaire, est une stratégie classique de production de protéines recombinantes difficiles à exprimer de façon soluble. La glutathion-S-transférase (GST) est un des partenaires de fusion les plus utilisés car cette protéine favorise souvent la solubilité de la construction chimérique et permet une purification simple par affinité sur une colonne greffée avec du glutathion. Nous avons donc cloné GenM1 dans le vecteur pGEX-KG, qui permet la fusion de la GST avec la protéine d'intérêt tout en insérant un site de reconnaissance pour la protéase thrombine entre les deux partenaires. Ce plasmide a été transformé dans des *E.coli* BL21(DE3) puis nous avons procédé à des tests d'expression dans des conditions classiques (37 °C, 2h d'induction, 500 µM IPTG).

Une large portion de la protéine chimère est exprimée sous forme de corps d'inclusion mais nous avons été heureux de constater qu'on obtenait une fraction non négligeable de protéine soluble de la taille attendue (M ≈ 73 kDa ≈ M(GenM1) + M(GST) ≈ 46 + 27 kDa). Nous avons purifié cette protéine en utilisant une première purification sur colonne d'affinité greffée au glutathion (référence Sigma G3907) puis une deuxième sur colonne d'exclusion de taille Superdex 200, qui exploitait la tendance de la GST à dimériser et former, dans le cas de la protéine chimère, un objet de volume hydrodynamique assez important pour sortir dans les fractions de tête de la colonne d'exclusion (voir figure II-11).

[3] Ces travaux ont été initiés avant mon arrivée au laboratoire par Bili Seijo et Nathalie Ulryck, tous deux techniciens dans mon équipe de recherche.

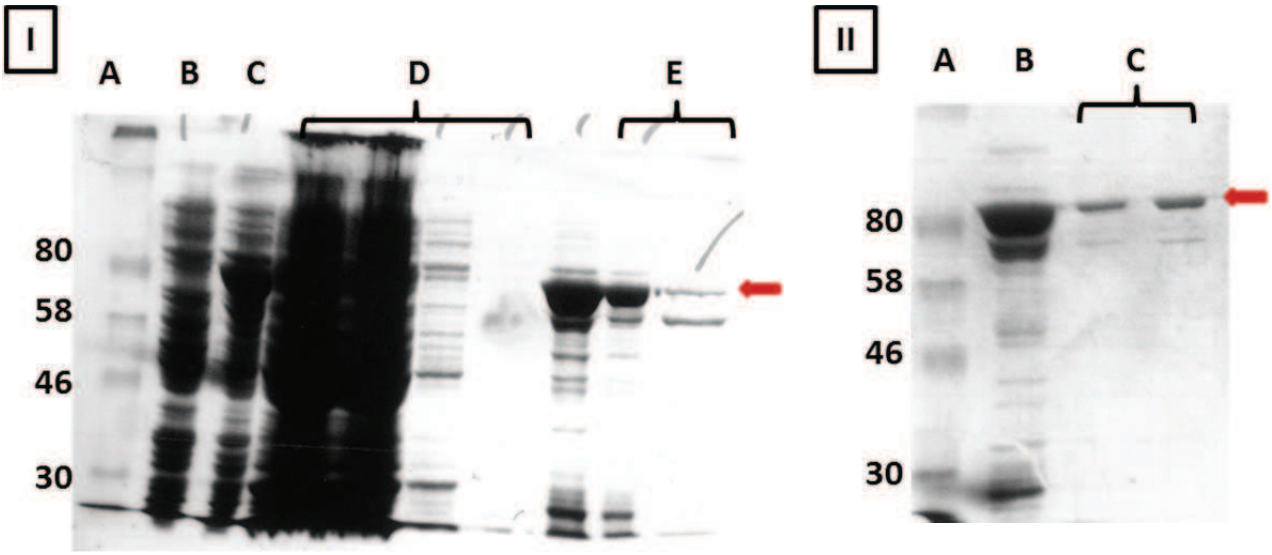



**Figure II-63 Protocole de purification de la protéine chimère GST-GenM1**

Gel I, purification par colonne d'affinité

A : marqueur de poids moléculaire ; B : BL21(DE3) non induites pour l'expression de GST-GenM1; C : BL21(DE3) induites pour l'expression de GST-GenM1; D : lavages ; E : élution de la protéine GST-GenM1 par ajout de glutathion ;

Gel II, purification par colonne d'exclusion stérique

A : marqueur de poids moléculaire ; B : échantillon avant colonne d'exclusion ; C : fractions purifiées de GST-GenM1

Le test d'activité biologique développé précédemment (partie II-II-B-3) n'a cependant pas détecté d'activité glycosyltransférase sur cette protéine chimérique purifiée. Nous avons donc clivé cette construction en utilisant le site thrombine intégré entre la GST et GenM1. Nous avons testé plusieurs conditions de digestion à la thrombine et étudié la cinétique de cette digestion. La figure II-12 montre la digestion de la protéine chimèrique GST-GenM1 par la thrombine à 25 °C dans un tampon contenant : 50 mM NaCl, 20 mM TrisHCl pH 7.8, 1mM DTT, et 1 % glycérol. Comme on peut le voir, la diminution de la bande correspondant à la protéine GST-GenM1 (flèche rouge, M ≈ 73 kDa) est accompagnée de l'apparition progressive des 2 bandes correspondant à GenM1 (M ≈ 46 kDa) et GST (M ≈ 27 kDa) (flèches jaunes). Des bandes plus faibles apparaissent également et correspondent à la digestion non spécifique de la chimère par la thrombine (présence de sites secondaires de digestion à la thrombine dans GenM1). Nous avons pu constater la précipitation assez rapide de la protéine GenM1 détachée de la GST. La coexistence de la précipitation et de la digestion non

spécifique par la thrombine entraîne la disparition totale de la bande de la protéine GenM1 soluble au bout d'environ 5h de digestion (dernière piste du gel présenté en figure II-12).

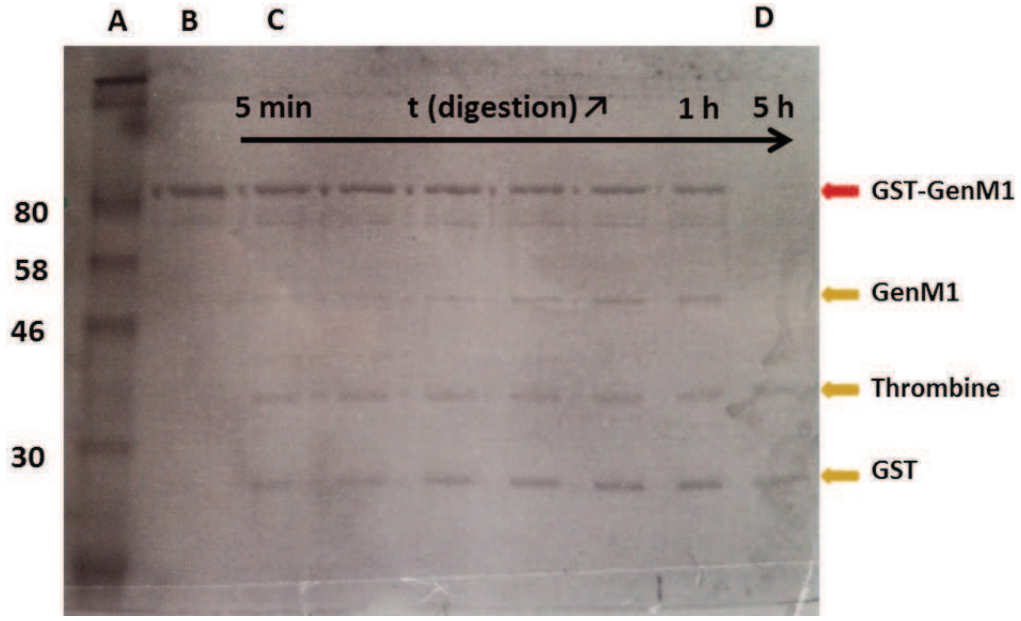



**Figure II-64 Digestion de la protéine GST-GenM1 par la thrombine**
A : marqueur de poids moléculaire ; B : GST-GenM1 purifiée ; C : début de la digestion à la thrombine (5 min) ; D : fin de la digestion (5h)

Dans le but de récupérer un peu de protéine GenM1 soluble, nous avons effectué des digestions incomplètes et testé un certain nombre de protocoles de purification (échangeuses d'ion de type S ou Q, exclusion de phase, purification par affinité ou par hydrophobicité, …) sans toutefois réussir à purifier une quantité significative de GenM1. Aucune des fractions obtenues au cours de ces expériences n'a montré d'activité glycosyltransferase avec notre test biologique.

### b) NusA-GenM1

Les difficultés rencontrées avec la GST nous ont conduit à tester une construction chimérique différente mettant en jeu une protéine de fusion récemment décrite, la protéine NusA[293]. La protéine NusA (pour « N-using substance A ») est un cofacteur essentiel de la transcription bactérienne, jouant un rôle clé dans les mécanismes d'antiterminaison[294]. Les caractéristiques très favorables d'expression cytoplasmique sous forme soluble de cette protéine en font un partenaire de fusion extrêmement intéressant pour la production de protéines insolubles dans les conditions classiques d'expression hétérologue[295]. Nous avons donc testé l'expression de GenM1 sous forme de protéine chimérique GenM1-NusA. Pour ce faire, Bili Seijo, technicien au laboratoire, a amplifié le gène NusA

de *E.coli* par PCR puis construit un plasmide, appellé pG-NusA, dérivé du plasmide d'expression classique pET-15b, permettant d'exprimer une protéine d'intérêt sous forme de fusion avec NusA et avec une étiquette poly-histidine en position C-terminale. Afin de permettre la purification ultérieure de la protéine GenM1, nous avons intégré entre les deux protéines un site de digestion pour la protéase TEV (qui n'a pas comme la thrombine de sites secondaires de digestion non spécifiques dans la séquence de la protéine GenM1).

Une fois en possession de ce plasmide, nous avons y avons intégré le gène GenM1 et avons effectué des tests d'expression. Cette construction chimérique était exprimée sous forme soluble dans des conditions classiques d'expression hétérologue (milieu LB, induction à DO = 0.6 avec 100 µM d'IPTG pendant 2 h à 25 °C). Nous avons ensuite purifié cette protéine par affinité (colonne Ni-NTA) puis par exclusion de taille (colonne Superdex200) (figure II-13). Ce protocole nous a permis d'obtenir un échantillon pur de la protéine NusA-GenM1 (flèche rouge, M ≈ 102 kDa ≈ M(GenM1) + M(NusA) ≈ 46 + 56 kDa) en quantité compatible avec une étude biochimique et/ou structurale. Le test de l'activité biologique de cette construction ne nous a cependant pas permis de mettre en évidence une quelconque activité de type glycosyltransférase.

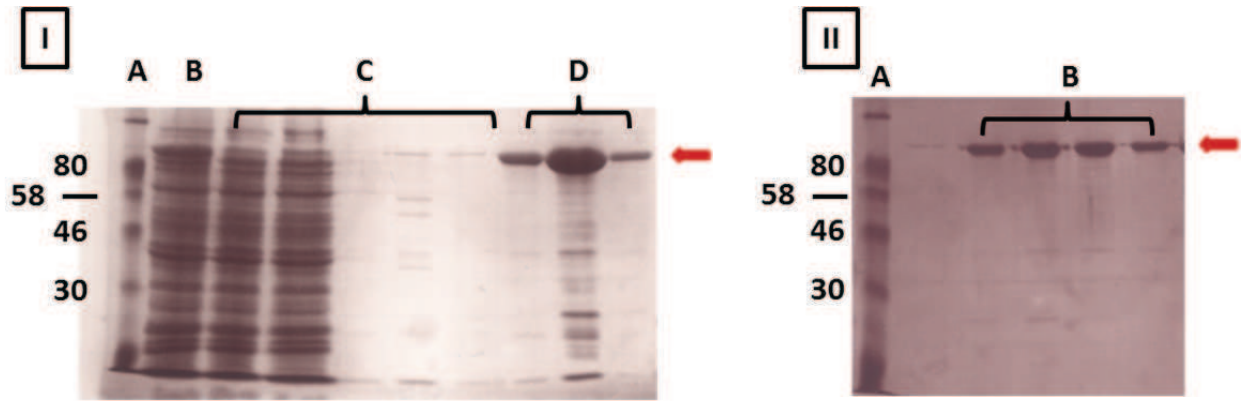



**Figure II-65 Expression et purification de la protéine NusA-GenM1**

Gel I, purification par colonne d'affinité

A : marqueur de poids moléculaire ; B : BL21(DE3) induites pour l'expression de Nusa-GenM1; C : lavages ; D : élution de la protéine NusA-GenM1 (flèche rouge, M ≈ 102 kDa);

Gel II, purification par colonne d'exclusion stérique

A : marqueur de poids moléculaire ; B : fractions purifiées de GST-GenM1

Nous avons alors décidé de cliver cette construction de façon à récupérer la protéine GenM1 seule. L'étape de digestion à la protéase TEV s'est montrée assez délicate. Il n'a en particulier jamais

été possible d'obtenir de digestion complète de la chimère, la réaction s'arrêtant à environ 50 % de conversion malgré les plus de 20 conditions différentes testées (changement du rapport TEV/NusA-GenM1, ajout de surfactant à une concentration supérieure ou inférieure à la concentration micellaire critique pour éviter l'agrégation de la protéine chimère, changement de tampon ou température (16°C, 25°C, 37°C), …) (figure II-14).

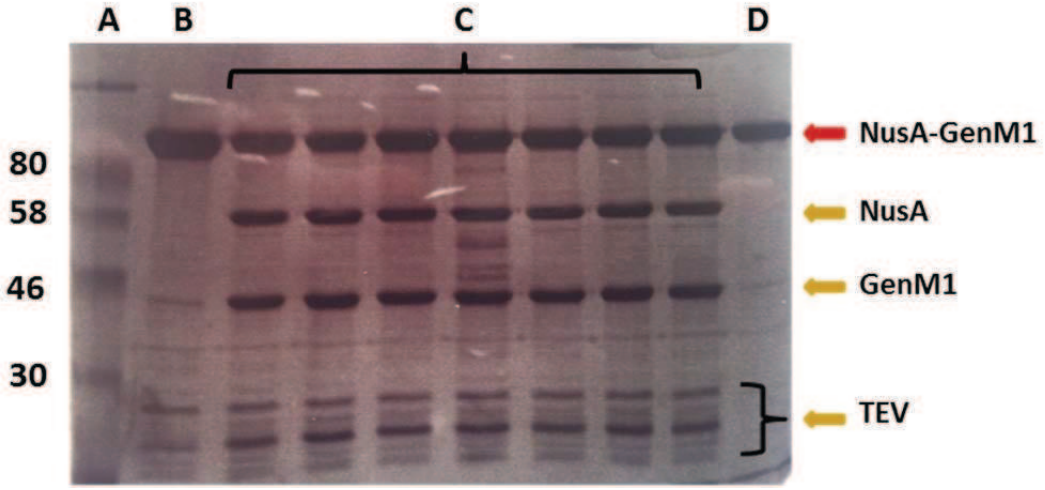



**Figure II-66 Essais de digestion de la protéine NusA-GenM1**
A : marqueur de poids moléculaire ; B : GenM1-NusA purifiée + protéase TEV; C : tests de digestion ; D : protéine NusA-GenM1 seule (flèche rouge, M ≈ 102 kDa)
On voit clairement l'apparition de deux bandes correspondant aux protéines NusA (M ≈56 kDa) et GenM1 (M ≈46 kDa).

Comme observé avec la protéine GST-GenM1, le fragment GenM1 avait tendance à précipiter une fois séparé du fragment correspondant à la protéine NusA avec cependant une cinétique légèrement plus lente. Nos différents essais de purification n'ont pas permis d'obtenir des quantités significatives de protéine GenM1 soluble et aucun des tests d'activité sur des échantillons de NusA-GenM1 partiellement digérés n'ont mis en évidence d'activité de type glycosyltransférase.

### c) AAC(6')-Ib-GenM1

La dernière construction chimérique que nous avons testée pour l'expression de la protéine GenM1 est beaucoup moins classique que les deux précédentes. Nous avons choisi d'utiliser comme partenaire de fusion une enzyme de résistance à la kanamycine précédemment étudiée au laboratoire, la protéine AAC(6')-Ib[130,296]. L'idée sous-jacente était de pouvoir détecter l'expression sous forme soluble de la protéine chimère à travers un criblage phénotypique de résistance à la kanamycine. Par ailleurs, nous souhaitions utiliser l'augmentation graduelle de la concentration en

kanamycine comme pression de sélection pour augmenter l'expression de la protéine chimère sous forme soluble.

Nous avons donc construit un plasmide permettant de produire la protéine chimère AAC(6')-Ib-GenM1 avec une étiquette poly-histidine en position N-terminale. Comme on peut le voir sur la figure II-15, la protéine chimérique est bien surexprimée dans des *E.coli* BL21(DE3) (conditions : milieu LB, 100 µM IPTG, 4 h induction à 37°C).

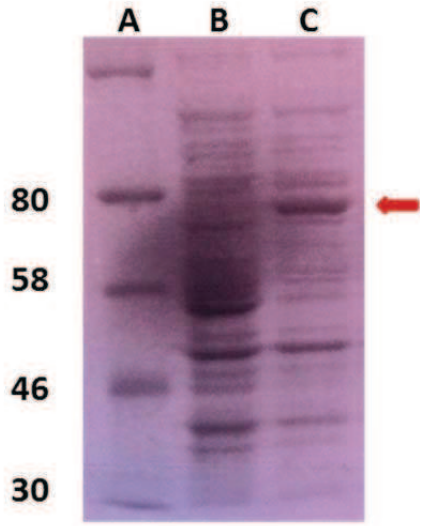



**Figure II-67 Surexpression de la protéine chimère AAC(6')-Ib-GenM1**
A : marqueur de poids moléculaire ; B : BL21(DE3) non induites; C : BL21(DE3) induites pour l'expression de la protéine chimère AAC(6')-Ib-GenM1. On distingue l'apparition d'une bande caractéristique à la taille attendue (flèche rouge, M ≈ 68 kDa ≈ M(GenM1) + M(AAC(6')-Ib) ≈ 46 + 22 kDa)

Les bactéries surexprimant la protéine chimère sont par ailleurs résistantes à la kanamycine (jusqu'à une concentration finale maximale de 300 µg/mL en fonction de la concentration en IPTG) alors que celles qui ne surexpriment par la protéine chimère ne le sont pas (une concentration de 50 µg/mL de kanamycine dans le milieu de culture est létale). Nous avons donc été surpris de constater que seule une très faible fraction de la protéine chimère était effectivement exprimée de façon soluble. Cette quantité n'a pas été modifiée par diverses expériences changeant les paramètres d'expression (concentration en IPTG, température, durée d'induction…). Nous avons alors répété la même expérience en augmentant graduellement la concentration en kanamycine, en utilisant à chaque fois pour ensemencer l'expérience les bactéries qui avaient réussi à pousser avec la concentration précédente. L'objectif de cette expérience d'évolution dirigée était de sélectionner des mutations qui amélioreraient l'expression sous forme soluble de la protéine chimère responsable de la résistance à la kanamycine. Nous avons répété l'expérience chaque jour pendant 15 jours en

augmentant graduellement la concentration en kanamycine jusqu'à des concentrations très élevées (1 mg/mL). A ce stade, les bactéries obtenues étaient capables de survivre même en l'absence d'induction de la surexpression de la protéine chimère AAC(6')-Ib-GenM1 (pas d'IPTG dans le milieu de culture). Elles ne montraient de plus pas de surexpression de la protéine AAC(6')-Ib-GenM1 (analyse des protéines bactériennes sur gel d'acrylamide). Les bactéries semblent donc avoir développé des stratégies alternatives de résistance à la kanamycine au lieu de mutations augmentant la solubilité de la protéine chimère. Ces stratégies pourraient impliquer des mécanismes non-spécifiques mettant en jeu la mutation de leurs pompes d'efflux ou la modification de leur perméabilité membranaire.

Rétrospectivement, si la stratégie d'évolution présente sans doute un intérêt pour obtenir un analogue soluble de GenM1, on peut juger que sa mise en place expérimentale a souffert de deux défauts principaux :

- Premièrement, l'enzyme de résistance AAC(6')-Ib s'est révélée bien trop efficace, une très faible concentration de celle-ci suffisant à conférer une résistance très importante à la kanamycine. Il aurait au contraire fallu utiliser une enzyme très peu efficace de façon à promouvoir une expression très importante de protéine chimère soluble.
- Deuxièmement, pour éviter que l'expérience d'évolution dirigée conduise à l'apparition de mutations en dehors du gène codant la protéine d'intérêt, il aurait fallu systématiquement initier chaque expérience avec un stock de bactéries fraichement transformées avec le plasmide purifié à l'étape précédente.

Il serait intéressant de recommencer cette expérience d'évolution dirigée en prenant en compte ces observations. Bien évidemment, si l'on obtient une protéine soluble analogue à GenM1, il faudra s'assurer qu'elle a bien gardé son activité de glycosyltransférase.

A ce stade de l'étude, nous avons décidé d'abandonner nos tentatives d'expression de GenM1 sous forme de construction chimérique. En effet, même dans les cas favorables avec expression et purification d'une protéine chimère soluble (GST-GenM1 ou NusA-GenM1 par exemple), il n'a jamais été possible de détecter d'activité glycosyltransférase, avec ou sans clivage protéolytique de la protéine chimère.

### 3) Ingénierie de la protéine GenM1 : conception de GenM1*

Nous avons ensuite envisagé de modifier de façon rationnelle GenM1 dans le but d'améliorer son expression hétérologue chez *E.coli*. Cette stratégie d'ingénierie protéique s'appuie de façon essentielle sur le développement récent d'un algorithme d'analyse, le logiciel PROSO

(http://mips.helmholtz-muenchen.de/proso/proso.seam), permettant de prédire avec un intervalle de confiance raisonnable, la propension d'une protéine à être exprimée de façon soluble chez *E.coli* en fonction de sa séquence en acide aminés[297]. Cet algorithme s'appuie sur une analyse statistique de l'ensemble des protéines recombinantes décrites comme solubles ou insolubles

L'objectif était donc de modifier sélectivement un certain nombre d'acides aminés de GenM1 afin d'augmenter sa probabilité d'être exprimée sous forme soluble. La difficulté de cette approche réside dans le fait que ces remplacements peuvent modifier l'activité biologique de la protéine. Pour repérer les acides aminés essentiels qu'il fallait éviter de modifier, nous avons procédé à des alignements de séquence avec des glycosyltransférases analogues à GenM1 et avons identifié les acides aminés conservés. Par ailleurs, Frédéric Dardel a créé un modèle structural de la protéine GenM1 à partir de structures de glycosyltransférases homologues déjà décrites[298,299] (entrées PDB : 2BIS et 3C48). Nous nous sommes ensuite efforcés de ne modifier que les acides aminés non conservés et situés à la surface de ce modèle structural. Dans la mesure du possible, ces acides aminés ont été remplacés par des acides aminés présents dans des glycosyltransférases homologues à GenM1, comme Neo8 (voie de biosynthèse de la néomycine), BtrM (voie de biosynthèse de la butirosine) ou KanF (voie de biosynthèse de la kanamycine). La protéine GenM1 artificielle ainsi conçue, nommée GenM1*, comportait un peu moins de 5 % d'acides aminés modifiés par rapport à GenM1 sauvage (voir figure II-16).

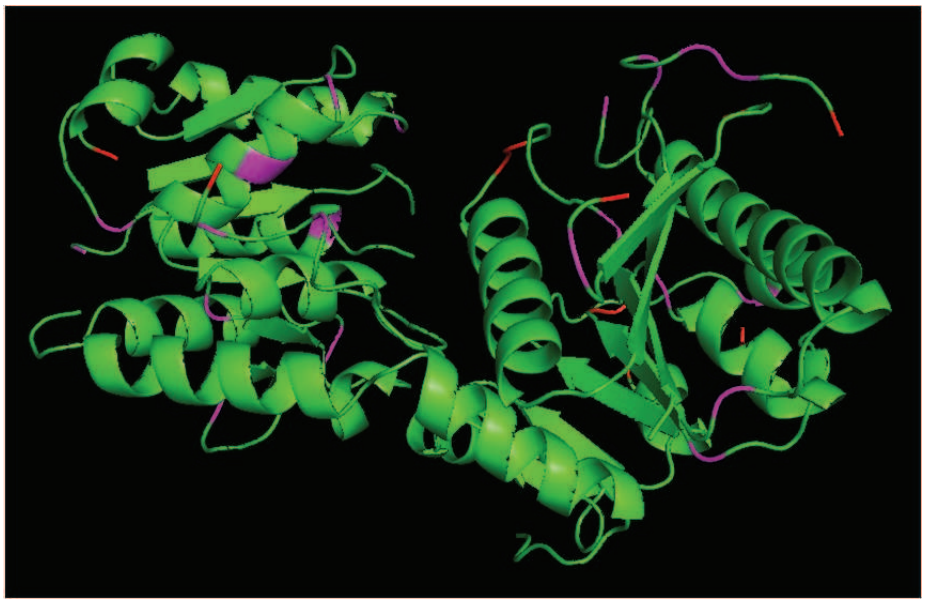


**Figure II-68 Modèle structural de la protéine GenM1*.**
Les acides aminés modifiés par rapport à GenM1 sont indiqués en rose ou en rouge.

Le gène codant pour cette protéine a été synthétisé (entreprise DNA 2.0) en optimisant l'utilisation des codons pour une expression optimale chez *E.coli*[292], puis intégré dans un plasmide pET-15b. Les premiers tests d'expression dans la souche BL21(DE3), présentés dans la figure II-17, ont montré une bonne expression de la protéine (conditions : induction pendant 2h à 37 °C, IPTG 500 µM). La fraction de protéine exprimée de façon soluble, bien qu'encore relativement faible, était plus importante que dans le cas de la protéine GenM1 non modifiée.

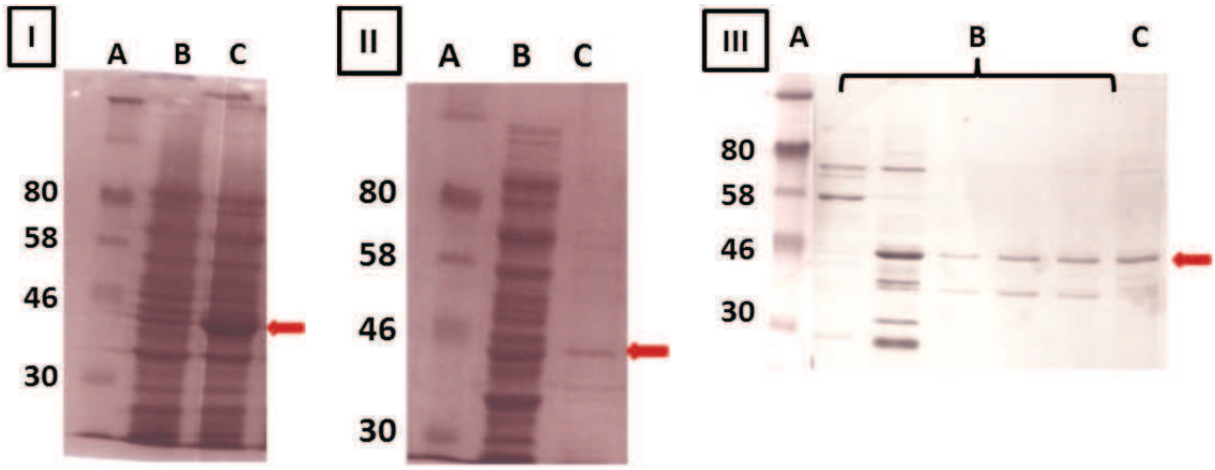



**Figure II-69 Surexpression et purification de la protéine GenM1***

Gel I, surexpression de GenM1*

A : marqueur de poids moléculaire ; B : BL21(DE3) non induites pour l'expression de GenM1*; C : BL21(DE3) induites pour l'expression de GenM1* (flèche rouge, M ≈ 43 kDa).

Gel II, purification par colonne d'affinité (Ni-NTA)

A : marqueur de poids moléculaire ; B : BL21(DE3) induites pour l'expression de GenM1* ; C : élution de la colonne Ni-NTA

Gel III, purification par exclusion de taille

A : marqueur de poids moléculaire ; B : fractions de la colonne d'exclusion; C : GenM1* purifiée

Comme son homologue sauvage, GenM1* montrait une tendance importante à précipiter. Nous avons cependant pu obtenir un échantillon relativement pur de la protéine en deux étapes (colonne Ni-NTA puis colonne d'exclusion). Le test de son activité biologique n'a cependant pas montré d'activité glycosyltransférase. Il est certainement possible que ceci soit l'effet de la modification des acides aminés de la protéine. Nous ne connaissons en effet pas suffisamment la structure et la biochimie des glycosyltransférases de la famille de GenM1 et la prévision de l'effet des mutations est donc extrêmement hasardeuse. Il est également possible que notre test d'activité ne soit pas assez sensible pour détecter de faibles niveaux de glycosylation de la DOS par l'UDP-GlucNAc. Enfin, il est également possible de l'activité glycosyltransférase que nous souhaitions mesurer n'est possible qu'en présence de cofacteurs qui nous sont inconnus.

## *D - Enzymes de biosynthèse de la butirosine chez B.circulans :BtrM et BtrD*

La dernière famille d'enzymes biosynthétiques à laquelle nous nous sommes intéressés étaient les enzymes de la voie de biosynthèse de la butirosine chez *Bacillus circulans*. Cet organisme étant phylogénétiquement plus proche d'*E.coli*, nous espérions produire plus aisément par expression hétérologue la glycosyltransférase BtrM, homologue à Neo8 et la déacetylase BtrD, homologue à Neo16.

Nous avons fait synthétiser ces deux gènes en optimisant l'utilisation des codons pour une expression hétérologue chez E.coli (entreprise DNA 2.0) puis nous les avons intégrés dans le vecteur d'expression pET-15b. Les premiers essais d'expression de BtrM ont une fois de plus conduit à de très faibles quantités de protéine soluble, qui n'ont pas permis de purification convenable. Pour BtrD par contre, nous avons pu suivre un protocole d'expression et de purification précédemment publié par l'équipe de Spencer[278]. Les conditions d'induction décrites (souche BL21(DE3), induction avec 200 µM d'IPTG pendant 16h à 16°C) nous ont permis d'obtenir de bonnes quantités de protéine BtrD soluble. La protéine a ensuite été purifiée très aisément en deux étapes (lyse par sonication, colonne d'affinité Ni-NTA suivie de colonne d'exclusion Superdex 75, rendement final après purification : ≈ 20 mg par litre de culture) (voir figure II-18).

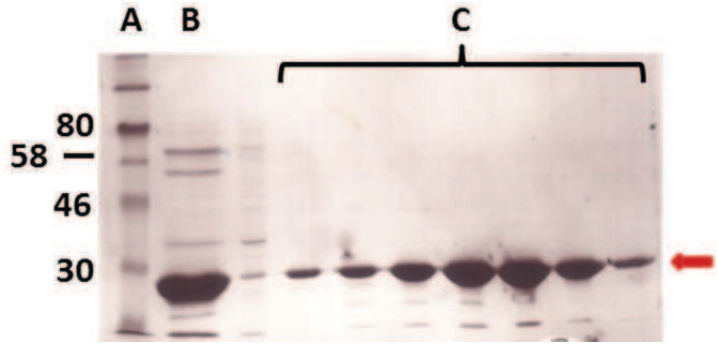



**Figure II-70 Surexpression et purification de la déactetylase BtrD (flèche rouge, M ≈ 32 kDa)**
A : marqueur de poids moléculaire ; B : élution de la colonne Ni-NTA; C : BtrD purifiée à l'issue de la colonne d'exclusion (flèche rouge, M ≈ 32 kDa).

Les échantillons de protéine BtrD purifiée ont été utilisés dans des expériences de cristallogenèse, avec ou sans digestion par la thrombine de l'étiquette polyhistidine (technique de diffusion de vapeur en gouttes suspendues). Avec l'aide de Franck Brachet, responsable du robot de cristallisation dans l'UMR 8015, nous avons pu tester plus de 3000 conditions de cristallisation différentes sans obtenir de cristaux exploitables (kits JCSG Core suite I, II, III et IV (QIAGEN), kits

Cristal screen 1 et 2 (Hampton), kit Memb/fac (Hampton), kit PEG suite et PEGion (Hampton), concentration en BtrD avec ou sans His-tag variant de 3mg/ml à 9 mg/ml, ajout d'UDP-GlucNAc, de DOS ou de paromamine…). Les cristaux obtenus se sont en effet avérés, après analyse au synchrotron SOLEIL, être des cristaux de sels.

# III - Conclusion

La difficulté du projet qui m'a été confié résidait dans l'expression hétérologue de certaines des enzymes clés de la voie de biosynthèse de la néomycine ou d'aminoglycosides apparentés en grande quantité et dans un état de solubilité et de pureté permettant une étude structurale. Le premier obstacle auquel nous nous sommes heurtés était la relative ignorance de la communauté scientifique du mode de fonctionnement de ces différentes voies de biosynthèse et en particulier des éventuels partenaires protéiques ou cofacteurs nécessaires au bon fonctionnement des différentes enzymes.

Nous nous sommes plus particulièrement intéressés à la première étape de la biosynthèse de la néamine à partir de la DOS, une étape commune à de nombreuses voies de biosynthèse. Pour augmenter nos chances de succès nous avons étudié les glycosyltransférases homologues Neo8, GenM1 et BtrM, respectivement impliquées dans la biosynthèse de la néomycine, de la gentamicine et de la butirosine. Leur expression hétérologue dans *E.coli* sous forme soluble s'est avérée extrêmement compliquée et, malgré la très large palette expérimentale utilisée (modification des conditions d'induction et de purification, introduction de partenaires de fusion dans des constructions chimériques, ingénierie protéique, …) nous n'avons jamais pu obtenir de quantité suffisante de protéine à la fois soluble et bioactive (résumé dans la figure II-19).

Pour guider nos études, nous avons développé une méthode permettant un suivi en UV de l'activité biologique de ces glycosyltransférases. Nous avons par ailleurs développé des anticorps polyclonaux spécifiques de la protéine Neo8 qui ont été utilisés pour étudier plus précisément le fonctionnement de cette dernière directement dans des lysats cellulaires du producteur naturel de la néomycine, la bactérie *S.fradiae*.

Les essais d'expression hétérologue des enzymes Neo11, Neo 16 et Neo18 ont eux aussi été marqués par la difficulté d'obtenir des protéines exprimées de façon soluble. Par contre, la protéine BtrD, une déacetylase homologue à Neo16, responsable de la biosynthèse de la butirosine chez *B.circulans*, a pu être exprimée de façon soluble et purifiée dans des quantités compatibles avec une étude structurale. Les pistes de cristallisation obtenues n'ont cependant pas donné les résultats escomptés.

| protéine | clonage | surexpression | purification | renaturation | activité biologique | cristallogenèse |
|---|---|---|---|---|---|---|
| **Neo8** | pEt-15b<br>pET-22 | Corps d'inclusion | **√** (conditions dénaturantes) | **X** | **X** (protéine renaturée) | - |
| **Neo16** | pEt-15b | Corps d'inclusion | - | - | - | - |
| **Neo11** | pEt-15b | Corps d'inclusion | **√** (conditions dénaturantes) | **X** | - | - |
| **Neo18** | pEt-15b | Corps d'inclusion | - | - | - | - |
| **GenM1** | pEt-15b<br>pET-22 | Faible % soluble | - | - | - | - |
| **GST-GenM1** | pGEX-KG | soluble | **√** | - | **X** (protéine entière ou digérée à la thrombine) | - |
| **Nusa-GenM1** | pET-15b modifié : pG-NusA | soluble | **√** | - | **X** (protéine entière ou digérée à la TEV) | - |
| **AAC(6') Ib-GenM1** | pET-15b modifié ptacRimL | Faible % soluble | - | - | - | - |
| **GenM1*** | pEt-15b | % soluble modéré | **√** | - | **X** | - |
| **BtrM** | pEt-15b | Faible % soluble | **X** | - | - | - |
| **BtrD** | pEt-15b | soluble | **√** | - | - | **X** |

**Figure II-19 Tableau récapitulatif des tentatives d'expression hétérologues des différentes enzymes étudiées.**

Les croix (X) représentent des résultats négatifs, les tirets (-) signifient que les expériences n'ont pas été tentées

L'ensemble des expériences présentées dans ce chapitre, ainsi que certaines des observations que nous avons faites sur des souches de *S.fradiae* génétiquement modifiées (exposées dans le chapitre suivant), nous conduisent à émettre l'hypothèse de l'existence de complexes multienzymatiques responsables de la biosynthèse des aminoglycosides apparentés à la néomycine. Cette hypothèse constituerait une explication plausible des difficultés que nous avons rencontrées ainsi que des difficultés exposées dans la littérature par d'autres groupes de recherche. En effet, si de tels complexes existent, étudier un de leurs partenaires protéiques indépendamment des autres serait extrêmement difficile. Au contraire, on peut imaginer qu'une approche visant à évaluer les différents partenaires biosynthétiques possibles puis à les coexprimer de façon hétérologue serait une manière bien plus simple d'avoir accès à ceux-ci.

Au terme de ce chapitre, nous proposons des pistes expérimentales permettant de vérifier cette hypothèse d'organisation de la voie de la néomycine et des aminoglycosides apparentés ainsi que certains des outils nécessaires (test d'activité et anticorps notamment).

Constatant la difficulté qu'il y avait à exprimer de façon hétérologue les différents enzymes clés de la voie de biosynthèse de la néomycine, nous avons à ce stade décidé de travailler directement sur la souche *S.fradiae*. En créant des souches génétiquement modifiées où la voie de biosynthèse est altérée, nous avons mis en place une stratégie de mutasynthèse visant à la production d'analogues de néomycine. Ces travaux seront exposés dans le chapitre suivant.

# III. Stratégie de mutasynthèse d'analogues de néomycine

Comme nous l'avons vu dans la partie introductive, la mutasynthèse vise l'obtention d'analogues de produits naturels en exploitant la transformation de mimes de métabolites par des organismes génétiquement modifiés. Historiquement, cette stratégie a été introduite par l'équipe de Rinehart dans les années 70 afin de produire des analogues de néomycine à partir de mutants du producteur naturel, la bactérie *Streptomyces fradiae*. Cette expérience a constitué le point de départ d'un domaine de recherche florissant qui, au vu des nombreuses publications qu'il suscite, semble aujourd'hui atteindre une certaine maturité scientifique. Presque 40 ans plus tard, nous avons souhaité revisiter les travaux initiaux de Rinehart, en tirant profit à la fois de notre compréhension récente de la voie de biosynthèse de la néomycine chez *S.fradiae* (cf chapitre II) mais aussi des derniers développements méthodologiques en ingénierie génétique, chimie organique et chimie analytique.

Après avoir brièvement rappelé les travaux fondateurs de Rinehart et examiné le contexte scientifique qui justifie notre projet de recherche, nous exposerons nos travaux de façon chronologique en suivant les étapes essentielles d'une stratégie de mutasynthèse telles qu'elles ont été proposées par Rinehart[250] :

(1) Génération par ingénierie génétique d'un mutant de *S.fradiae* incapable de synthétiser la désoxystreptamine (DOS), intermédiaire essentiel dans la biosynthèse de la néomycine

(2) Développement d'une méthode de purification et d'analyse des aminoglycosides

(3) Optimisation des conditions de culture pour la mutasynthèse

(4) Synthèse d'une librairie d'analogues d'aminoglycosides pouvant servir de mutasynthons

(5) Complémentation chimique de la souche mutée avec des métabolites naturels

(6) Expérience de mutasynthèse proprement dite.

Nous procéderons ensuite à une analyse critique des résultats obtenus et proposerons des pistes à explorer pour dépasser certaines des difficultés rencontrées qui nous ont empêchés de conduire ce projet à son terme.

# I - Introduction

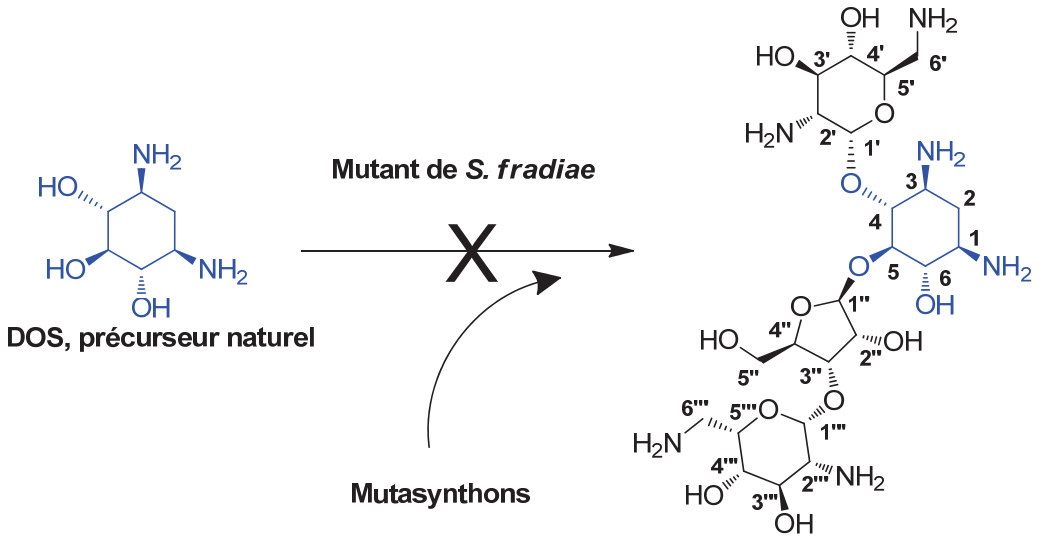
## *A - Les travaux fondateurs*

Dans les années 60, l'élucidation de la structure de certaines molécules antibiotiques de la famille des aminoglycosides permet de souligner la similitude de ces assemblages pseudo-polysaccharidiques et, en particulier, leur organisation autour d'un cycle central commun très conservé, la désoxystreptamine (DOS)[92]. Ceci attise la curiosité de la communauté scientifique qui cherche alors à en déterminer les voies de biosynthèse, principalement à travers des expériences de marquage métabolique à l'aide de précurseurs radioactifs[300] ou enrichis en isotope pour la RMN (comme par exemple le [6-$^{13}$C]-glucose, la [1-$^{13}$C]-glucosamine ou la [$^{15}$N]-glucosamine)[301].

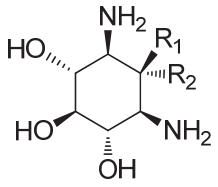
En 1969, au cours de l'étude de la voie de biosynthèse de la néomycine chez son producteur naturel *Streptomyces fradiae*[300], l'équipe de Rinehart décrit une méthode de sélection de souches mutantes incapables de produire de la néomycine sans ajout de désoxystreptamine exogène dans le milieu de culture. Ils isolent une telle souche après de nombreux essais de culture de la souche sauvage sur un milieu contenant de la N-methyl-N'-nitro-N-nitrosoguanidine, composé connu pour générer des mutations ponctuelles aléatoires dans le génome. Plus de 4000 des mutants obtenus ont dû être criblés[251]. Le principe du criblage est simple : les bactéries survivantes après culture en milieu liquide hautement mutagène sont étalées sur un milieu de culture solide contenant de l'agar. Après sporulation, les clones sont isolés puis repiqués sur le même milieu de culture solide dans deux boites de pétri, avec et sans DOS. Les boites sont alors recouvertes de milieu nutritif contenant de l'agar ensemencé avec une bactérie sensible à la néomycine (dans cette étude, *Bacillus subtilis*) qui va servir de rapporteur de la production éventuelle de l'antibiotique. Les souches d'intérêt montrent une zone d'inhibition de la croissance de *Bacillus subtilis* uniquement dans la boite contenant de la DOS.

Rinehart comprend tout de suite qu'en remplaçant la DOS par un analogue suffisamment proche structurellement pour être toléré par la voie de biosynthèse, on peut utiliser ces souches mutées pour obtenir un analogue de néomycine difficilement accessible par les méthodes classiques de synthèse totale ou d'hémisynthèse. Le concept de la mutasynthèse est né. Dans les années qui suivent, plus de 30 analogues de la DOS sont synthétisés puis testés en tant que mutasynthons sur la souche de *S.fradiae* mutée[302,303]. La figure III-1 donne un aperçu de la variété des différents mutasynthons testés : modification d'une ou plusieurs fonctions alcools de la DOS (suppression,

épimérisation, méthylation ou remplacement par un halogène ou une amine), modification d'une ou des deux amines (methylation, acétylation, …), création d'analogue bicyclique rigide, …

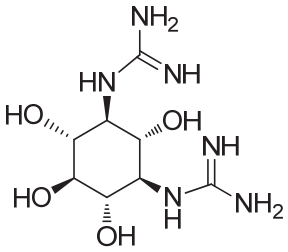



Exemples de mutasynthons testés

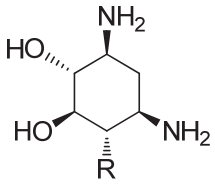


**streptamine 47**
($R_1$ = H, $R_2$ = OH)

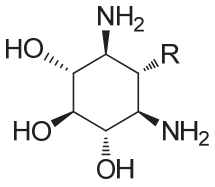
**epi-streptamine 48**
($R_1$ = OH, $R_2$ = H)



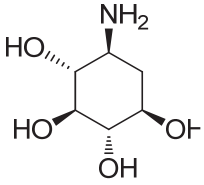
**streptidine 49**



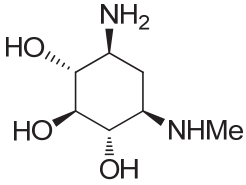
R = H (**50**), OMe (**51**)
ou Br (**52**)



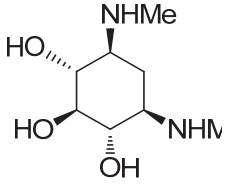
R = Cl (**53**), Br (**54**)
ou $NH_2$ (**55**)



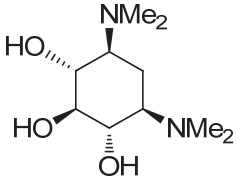
**56**



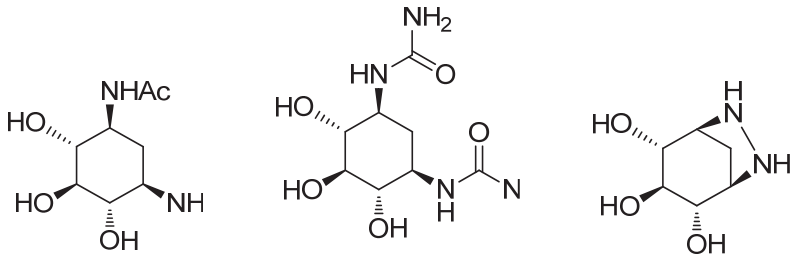
**57**



**58**



**59**




**60** **61** **62**

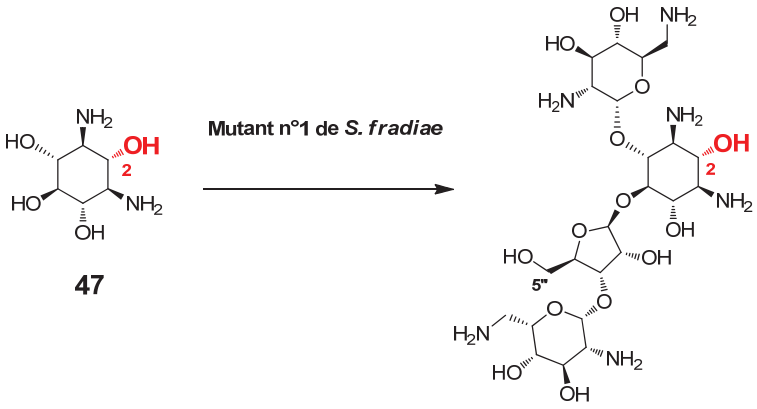
**Figure III-71 les premiers essais de mutasynthèse sur la souche mutante de *S.fradiae* isolée par Rinehart[302,303].**

Sur l'ensemble des molécules testées, seules 3 ont été transformées par la souche mutante. La DOS **1** bien sûr, qui conduit à la formation de néomycine, mais aussi la streptamine **47** et l'épi-streptamine **48**, qui conduisent à des analogues de néomycine hydroxylés en position 2. De plus l'activité biologique de ces analogues s'est révélée plus faible que celle de la néomycine.

Malgré ce bilan mitigé, de nombreuses équipes se sont néanmoins lancées dans des stratégies de mutasynthèse. On peut par exemple citer l'équipe de Pearce qui a testé d'autres mutasynthons sur la souche originale isolée par Rinehart et a démontré que la 2,6-désoxystreptamine pouvait être utilisée comme mutasynthon pour former la 6-déoxynéomycine[304] (voir figure III-2). D'autres équipes ont développé de nouveaux mutants aléatoires de *S.fradiae* ayant une tolérance plus grande vis-à-vis d'autres mutasynthons que la souche mutante originale. C'est le cas de l'équipe de Rinehart qui a utilisé des mutants de *S.fradiae* fournis par le docteur Kimura de l'entreprise TokyoTanabe capables de transformer la mono-N-methyldeoxystreptamine, la streptidine, et la 2-bromo- ou 6-bromostreptamine[250]. Malheureusement, la transformation de ces mutasynthons n'a été que déduite de tests d'activité biologique et les produits obtenus par mutasynthèse n'ont été ni isolés ni caractérisés. (voir figure III-2). D'autres groupes enfin ont étudié des mutants issus de souches apparentées à *S.fradiae,* productrices d'aminoglycosides similaires à la néomycine comme *S. kanamyceticus* (kanamycine) et *S. ribosidificus* (ribostamycine)[305], *S. rimosus* (paromomycine)[306] ou *B.circulans* (butirosine)[307]. Contrairement à ce que suggéraient les premières expériences, on a alors découvert que certaines souches modifiées étaient capables de prendre en charge des mutasynthons pseudodissacharidiques comme la néamine ou la 3',4'-didéoxynéamine[305] (voir figure III-2).

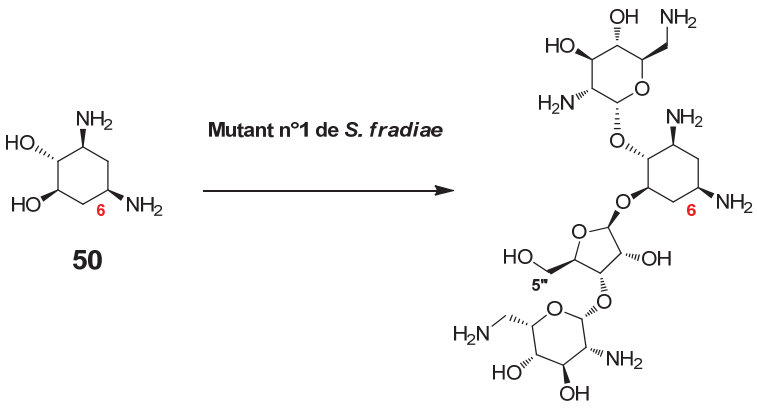
Dans certains cas surprenants, le produit formé par mutasynthèse n'était pas simplement issu de l'incorporation du mutasynthon à l'architecture de l'aminoglycoside de départ mais dérivait de

modifications plus importantes de l'architecture du produit naturel. Ces cas très intéressants, comme par exemple la transformation de l'épi-streptamine **48** en 6'-deamino-6'-hydroxy-2-hydroxy-kanamycine (composé **70**) par une souche mutée de *S.kanamyceticus*[305], montraient que les mutasynthons pouvaient interférer avec les transformations biochimiques situées en aval de leur incorporation dans la voie de biosynthèse. Ceci justifiait l'effort de la caractérisation rigoureuse systématique des produits éventuellement formés par mutasynthèse.



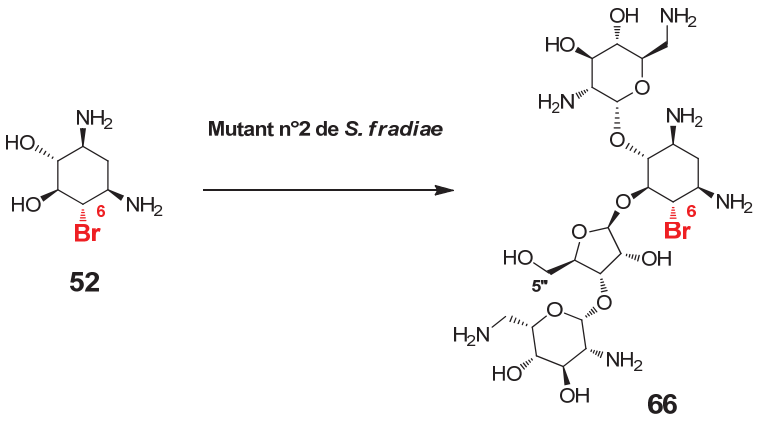

**47** **63**

Référence : [251]

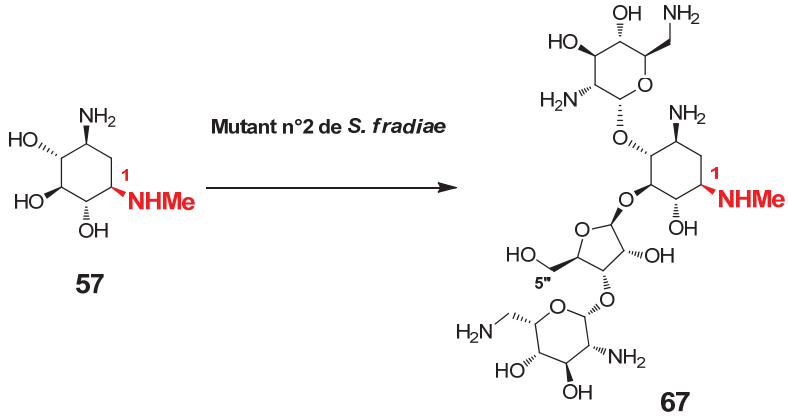



**50** **65**

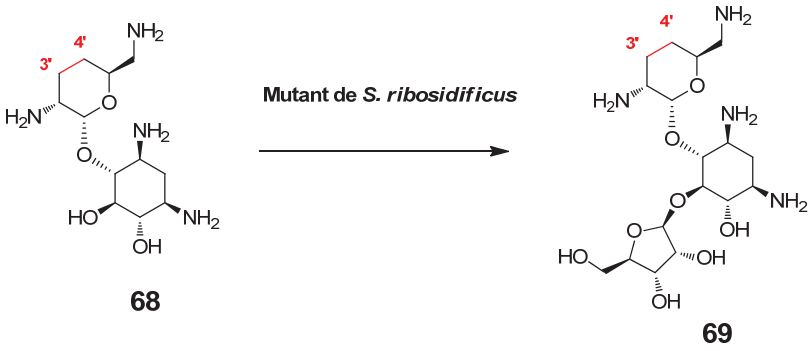
Référence : [304]



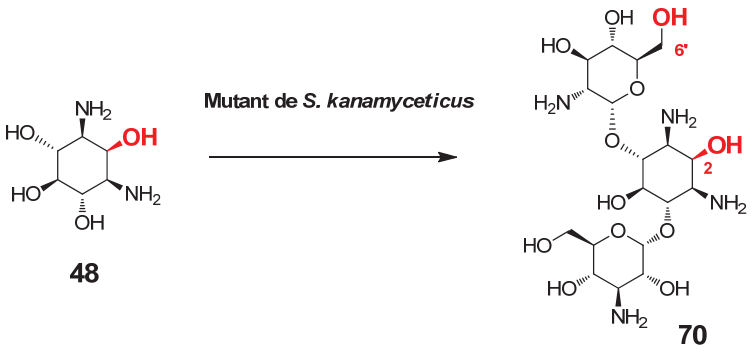

Référence : [250] (produit non caractérisé)




Référence : [250] (produit non caractérisé)




Référence : [305]




Référence : [305]

**Figure III-72 Exemples d'analogues d'aminoglycosides obtenus par mutasynthèse**

En dépit du nombre relativement limité de mutasynthons testés, ces recherches ont étendu la diversité structurale des mutasynthons pris en charge par des voies de biosynthèse modifiées. On peut retenir que la voie de biosynthèse de la néomycine et des aminoglycosides apparentés semble être assez tolérante vis-à-vis de métabolites non naturels. Par contre, le motif *cis*-1,3-diamine de la DOS est absolument essentiel pour la reconnaissance du mutasynthon par la machinerie biosynthétique et pour l'activité biologique de l'éventuel composé formé. Une trop grande modification de ce motif conduit inévitablement à des mutasynthons non pris en charge par les souches modifiées ou à des produits non antibiotiques et donc non détectés par les techniques de criblage utilisées à l'époque. Pour obtenir un analogue de néomycine, il est bien sûr important de conserver les deux hydroxyles en position 4 et 5 de la DOS car ils servent de points d'attache aux unités glycosidiques ajoutées au cours de la biosynthèse. Lorsque l'on utilise un mutasynthon pseudo-dissacharidique, il semble possible d'effectuer un certain nombre de modifications structurales sur le cycle de type glucosamine sans toutefois altérer de façon rédhibitoire l'efficacité de la voie de biosynthèse.

Le bilan des recherches menées au cours des années 70 à 80 permet de souligner un certain nombre de défauts expérimentaux qui ont limité la portée de leurs résultats et donc de la stratégie de mutasynthèse.

Tout d'abord, les souches bactériennes ont toutes été obtenues aléatoirement et donc sans aucun contrôle sur la cible mutée. Ainsi différents mutants de la même souche bactérienne initiale ont montré une tolérance différente vis-à-vis d'une série de mutasynthons (voir notamment les travaux de Rinehart sur deux mutants différents de *S.fradiae*[250]). Les mécanismes gouvernant la

tolérance d'une voie de biosynthèse modifiée vis-à-vis de mutasynthons exogènes restent donc à découvrir ce qui représente un défi scientifique très important.

Ensuite, l'aspect quantitatif de la production d'analogues par mutasynthèse a été peu développé. En particulier, le grand nombre de données contradictoires publiées semblent indiquer une reproductibilité faible des expériences. Par exemple un même mutasynthon a souvent été rapporté comme transformé en un analogue de néomycine par une souche mutée dans un laboratoire mais pas dans un autre (voir par exemple les travaux de Sato sur *S.ribosidificus*[305] qui semblent contredire ceux de Rinehart). Ces contradictions pourraient être expliquées par le faible niveau de détection des métabolites formés que permettaient les techniques d'analyse utilisées. Il est cependant aussi possible qu'elles soient liées à la variabilité de production des métabolites secondaires souvent observée chez les *Streptomyces*[4] (voir la caractérisation précise de la variabilité de la production de novobiocine chez une souche de *Streptomyces* par Heide[308]). On peut par ailleurs supposer que cette variabilité est d'autant plus grande que les protocoles de mutasynthèse mis en œuvre dans les différents laboratoires sont variés.

Enfin, ces travaux n'ont pas été accompagnés de caractérisation rigoureuse des produits finaux éventuellement formés par mutasynthèse. Au minimum cette caractérisation se résumait à un test d'activité biologique (antibiogramme) du surnageant d'une culture d'une souche mutante en présence d'un mutasynthon. Bien souvent, elle consistait en un rapport frontal en chromatographie (Rf), un point de fusion (mp) et un pouvoir rotatoire ($\alpha_D$). Les débuts de la technique MALDI en spectrométrie de masse ont permis quelques caractérisations de la masse des éventuels produits formés. Quelques analogues ont été très bien caractérisés, comme par exemple les « hybridimycines » obtenues par Rinehart[250] ou les analogues de gentamicine obtenus par l'équipe de Daum[309,310], pour lesquels on dispose de spectres RMN $^{13}C$ entièrement attribués des molécules formées ainsi que d'analyses élémentaires qui permettent d'affirmer sans ambiguïté l'existence et la structure des molécules issues de la mutasynthèse. Dans la plupart de ces études, les analogues d'aminoglycosides produits sont purifiés sur résine échangeuse d'ions puis par chromatographie sur papier ou sur couche mince d'alumine, avec ou sans dérivatisation chimique (en général, peracétylation des amines).

[4] Nos collaborateurs ont eux-mêmes été exposés à ce caractère stochastique de la production de métabolites secondaires chez *Streptomyces* lors de leur étude de la voie de biosynthèse de la congomycine chez *Streptomyces ambofaciens*[379].

## *B - Contexte scientifique actuel*

Un certain nombre d'avancées scientifiques récentes permettent de revisiter ces travaux fondateurs afin d'en dépasser certaines des limitations : connaissances accrues des voies de biosynthèse, développement d'outils d'ingénierie génétique chez Streptomyces et nouvelles méthodologies en chimie organique et analytique.

Notre compréhension de la voie de biosynthèse de la néomycine chez *Streptomyces fradiae* s'est considérablement accrue grâce aux travaux des équipes de Spencer[279,275] et d'Eguchi[280,276] (voir chapitre II). Nous disposons maintenant de la séquence de l'ensemble de l'opéron de biosynthèse et d'une fonction bien définie pour la quasi-totalité des gènes décrits (voir figure III-3). De façon particulièrement intéressante, certaines enzymes sont impliquées dans plusieurs étapes de la voie de biosynthèse et acceptent des substrats variés. Ceci laisse présager une spécificité relâchée et donc potentiellement une certaine tolérance vis-à-vis de substrats exogènes qui pourrait être exploitée en mutasynthèse. Cette tolérance semble d'ailleurs être commune à de nombreuses voies de biosynthèse d'aminoglycosides comme le suggèrent les travaux récents de l'équipe de Sonhg[311].

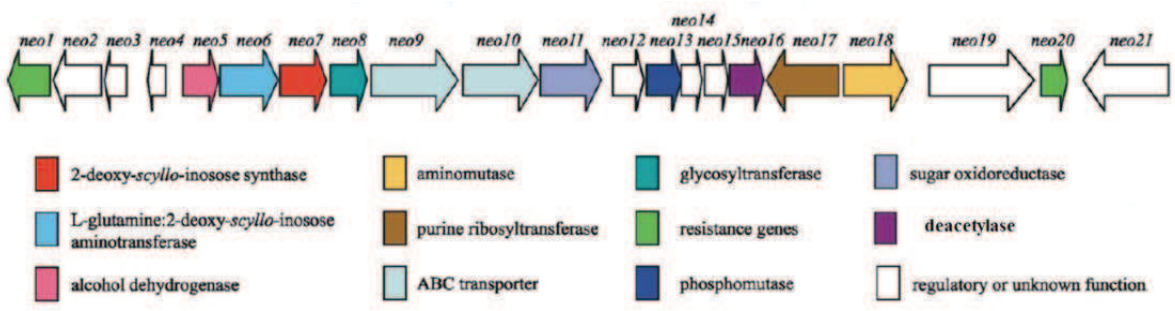



**Figure III-73 L'opéron de biosynthèse de la néomycine chez *S.fradiae*** (figure adaptée de [273])

Etape essentielle dans la mise en place d'une stratégie de mutasynthèse fondée sur la modification rationnelle de la voie de biosynthèse de la néomycine, les travaux d'Eguchi ont prouvé que les gènes neo5, neo6 et neo7 étaient impliqués dans la biosynthèse de la DOS à partir du glucose-6-phosphate[272]. L'expression hétérologue de chacun de ces gènes chez *E.coli* a en effet permis l'apparition de l'activité enzymatique des protéines dans le lysat cellulaire (activité suivie en HPLC) et ceci bien qu'elles soient majoritairement exprimées sur forme de corps d'inclusion. On a ainsi pu en déterminer les différentes étapes catalysées par chacune de ces enzymes, représentées dans la figure III-4.

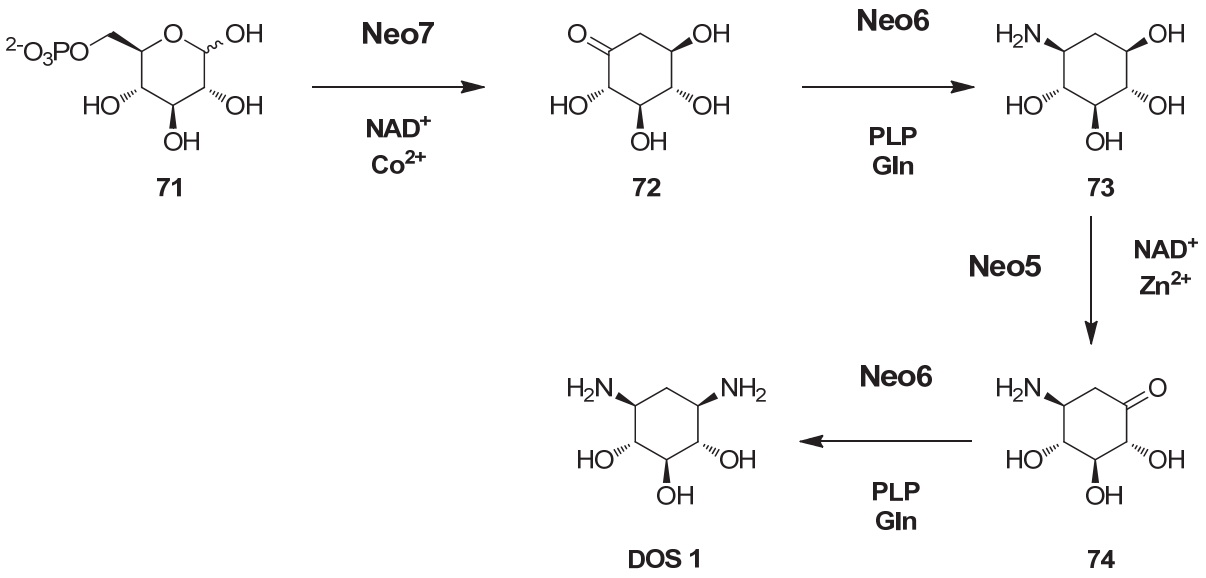



**Figure III-74 Synthèse de la DOS à partir de glucose-6-phosphate**[272]
La biosynthèse implique 2 aminations réductrices catalysées par Neo6 à l'aide du cofacteur pyridoxal phosphate (PLP) et de la glutamine (Gln) qui sert de donneur d'azote. La première et la troisième étapes sont des réactions d'oxydoréductions utilisant le cofacteur nicotinamide adénine dinucléotide ($NAD^+$)

Ces informations nous permettent d'envisager de supprimer sélectivement un des gènes de biosynthèse de la DOS de façon à créer de façon rationnelle et contrôlée un mutant de *S.fradiae* utilisable dans une stratégie de mutasynthèse. Pour ce faire, nous pouvons nous appuyer sur le développement récent de certains outils de modification génétique des *Streptomyces*[253], et en particulier la stratégie de la double recombinaison (voir figure III-5). Ceci nous a conduit à initier une collaboration avec l'équipe de Jean-Luc Pernodet à Orsay qui développe depuis quelques années des outils de modification génétique des *Streptomyces*[254] et qui possède une expertise scientifique dans la production de métabolites secondaires chez les Actinomycètes[312,313].

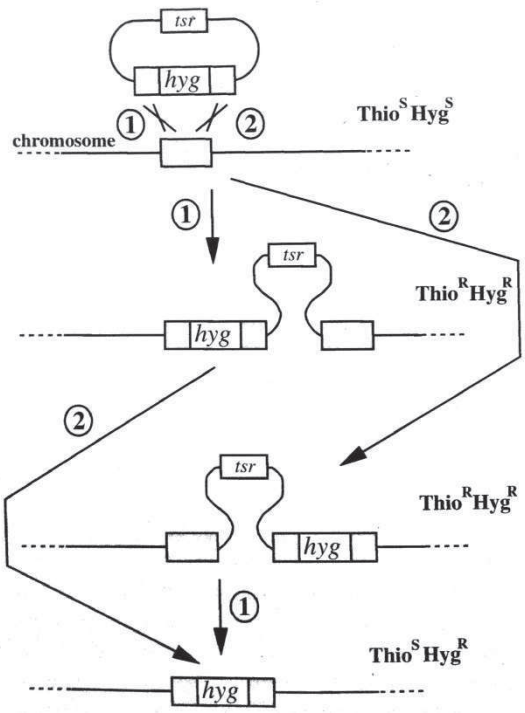



**Figure III-75 Stratégie de délétion d'un gène par double recombinaison**
(figure adaptée de livre de référence d'Hopwood[253])
Le plasmide utilisé comporte une cassette de résistance à l'hygromycine (hyg), flanquée de deux parties homologues au gène d'intérêt et d'une cassette de résistance au thiostrepton (tsr). Il s'intégre au chromosome de la souche sauvage par recombinaison homologue. Un événement de double recombinaison conduit à l'insertion de la cassette Hyg au milieu de la séquence cible. Ces souches ainsi obtenues sont résistantes à l'hygromycine mais sensibles au thiostrepton dont la cassette est perdue lors de la deuxième recombinaison. D'autres combinaisons d'antibiotiques peuvent être utilisées à la condition que la souche que l'on veut modifier n'y soit pas naturellement résistante.

Les nouvelles techniques d'analyse des métabolites produits par les souches mutées facilitent aussi grandement la mise en place d'une stratégie de mutasynthèse. Contrairement aux premières expériences de mutasynthèse effectuées dans les années 70, nous disposons maintenant d'appareils de chromatographie en phase liquide à haute performance couplés à la spectrométrie de masse (HPLC-MS/MS) qui permettent un criblage moyen débit des différentes expériences de mutasynthèse effectuées et une caractérisation plus facile des éventuels produits formés. Pour appliquer ces techniques à notre projet, nous avons ainsi initié une collaboration avec le laboratoire de Chimie-Toxicologie analytique et cellulaire de la faculté de Pharmacie de Paris (Anne Regazzetti et Olivier Laprévote).

Enfin, notre équipe a développé au cours de ces dernières années des méthodologies de synthèse d'analogues de désoxystreptamine à 5 ou 6 chainons reposant sur la désymétrisation de

l'hydrazine bicyclique **75** (voir [314,315,316] et travaux non publiés d'Aurélie Blond). Ceci nous permet d'envisager la synthèse d'une petite librairie de molécules qui pourront être testées en tant que mutasynthons (voir figure III-6). Les expériences préliminaires effectuées par Sato dans les années 70 laissaient entrevoir la possibilité d'utiliser des pseudo-dissacharides en tant que mutasynthons[305]. Les méthodologies classiques de la chimie des sucres nous permettent d'envisager de synthétiser ces analogues qui sont plus avancés dans la voie de biosynthèse de la néomycine. Il est en particulier envisageable de synthétiser des analogues de paromamine modifiés en position 6' (voir figure III-7).

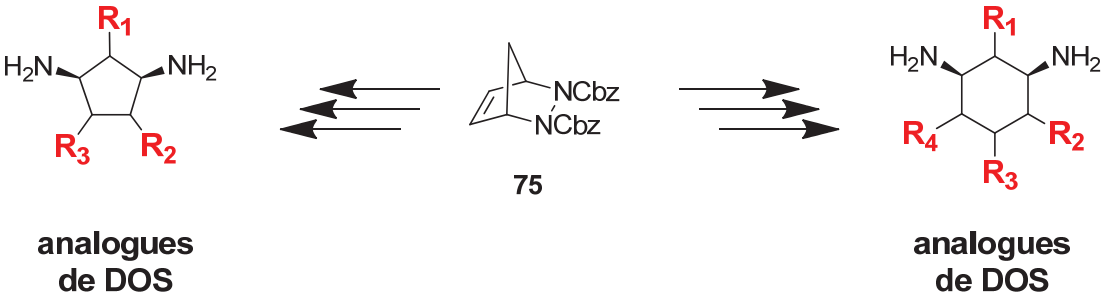



**Figure III-76 Synthèse d'analogues de DOS par désymétrisation d'hydrazine bicycliques**

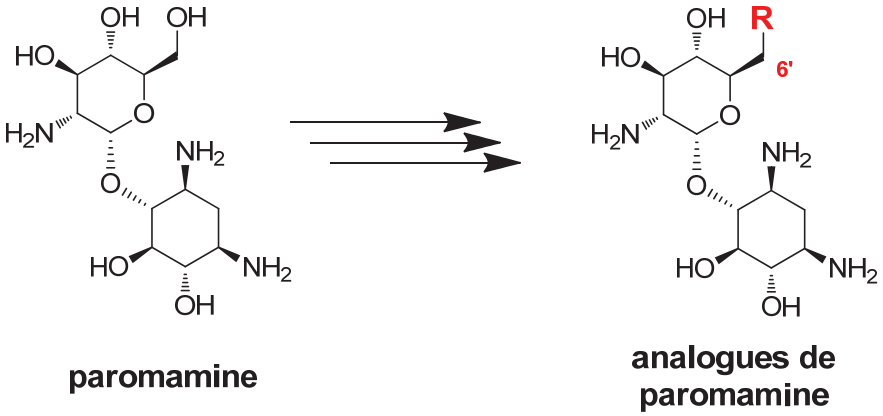



**Figure III-77 Synthèse d'analogues de paromamine modifiés en position 6'**

## *C - Objectifs de l'étude*

Notre projet était de mettre en place une stratégie de mutasynthèse pour obtenir des analogues de néomycine à partir d'une souche de *S.fradiae* incapable de synthétiser de la DOS.

Nous avons choisi de construire une souche de *S.fradiae* où le gène neo6, essentiel à la biosynthèse de la DOS, est supprimé. En effet, ce gène, situé entre les gènes neo5 et neo7 eux-mêmes impliqués dans la biosynthèse de la DOS, peut être supprimé par double recombinaison de

façon contrôlée avec un risque minimal de perturbation des autres étapes de la biosynthèse. Une inhibition même partielle de l'activité de ce gène aurait des conséquences importantes sur la biosynthèse, neo6 étant impliqué à deux reprises dans la transformation de la DOS à partir du glucose-6-phosphate. Le développement de cette souche de *S.fradiae*, que l'on appellera Δneo6 par la suite, sera exposé dans la partie III-II-A. Les résultats d'expériences de mutasynthèse obtenus avec Δneo6 seront comparés avec ceux obtenus avec la souche mutante originale obtenue par mutagénèse aléatoire[251] disponible commercialement (référence DSM 41550 qui servira pour désigner cette souche par la suite).

Une fois en possession de ces deux souches, notre objectif a été de mettre en place une méthodologie de purification et d'analyse des aminoglycosides à partir de milieu de culture (partie III-II-B). Nous avons pour cela développé une méthode associant purification par extraction liquide-solide sur résine échangeuse d'ions (SPE) et analyse par HPLC-MS/MS. Il nous a alors été possible d'optimiser les conditions expérimentales requises à la mutasynthèse à partir de l'analyse quantitative de la production de néomycine par la souche *S.fradiae* sauvage ou par les souches mutées DSM 41550 ou Δneo6 (partie III-II-C). L'objectif suivant a été de mettre en place la synthèse d'analogues d'aminoglycosides (partie III-II-D) puis à les tester dans des expériences de mutasynthèse (partie III-II-E).

La figure III-8 représente schématiquement les différentes étapes de cette étude.

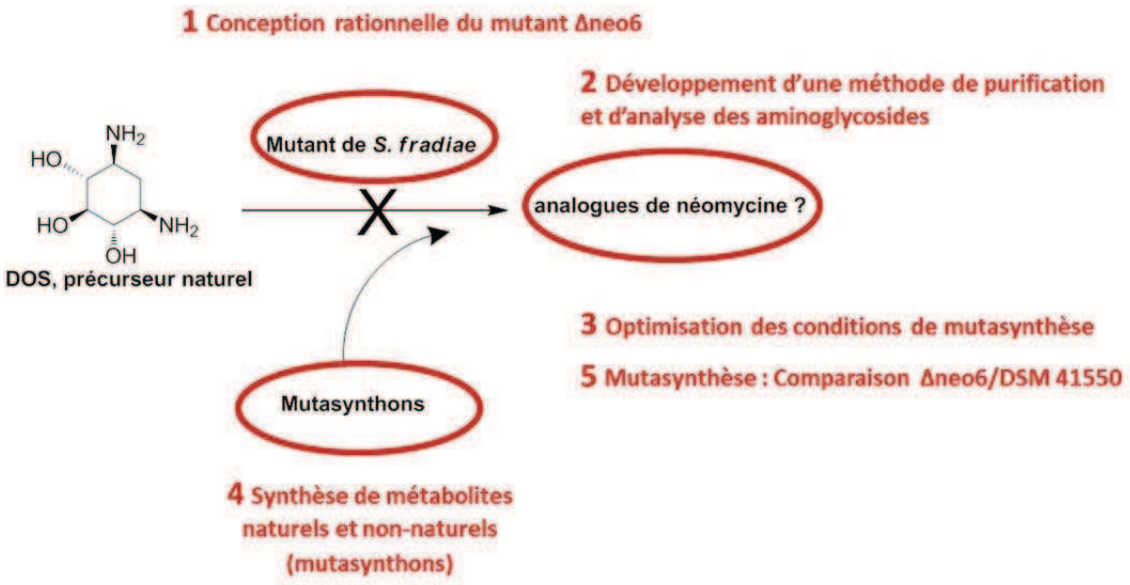



**Figure III-78 Mutasynthèse d'analogues de néomycine : les différentes étapes de l'étude**

# II - Résultats personnels

## *A - Construction du mutant Δneo6*

La première étape de ce projet de recherche a été de concevoir un mutant de *S.fradiae* compatible avec la mutasynthèse d'analogues de néomycine. Nous sommes partis de la souche de *S.fradiae* sauvage ATCC 10745. Comme rappelé précédemment, nous avons choisi de supprimer le gène neo6, responsable de la biosynthèse de la DOS à l'aide d'une stratégie de double recombinaison. Cette stratégie nécessite l'introduction dans *S.fradiae* d'un plasmide spécifique, que l'on appellera pFL02-neoApra, contenant une cassette de résistance à l'apramycine, notée Apra, flanquée de grandes régions homologues aux régions de l'opéron de biosynthèse de la néomycine situées de part et d'autre du gène neo6 qui vont être recombinées[253] (voir figure III-9). Ce plasmide, échangé par conjugaison interspécifique entre *E.coli* S17.1 et *S.fradiae*, contient également une origine de réplication oriT et une cassette de résistance à l'hygromycine, notée Hygro. Il peut être répliqué chez *E.coli* mais pas chez *S.fradiae* ce qui permet de vérifier que les souches de *S.fradiae* résistantes à l'apramycine mais pas à l'hygromycine sont bien issues de l'intégration de ce plasmide à leur génome par double recombinaison homologue.

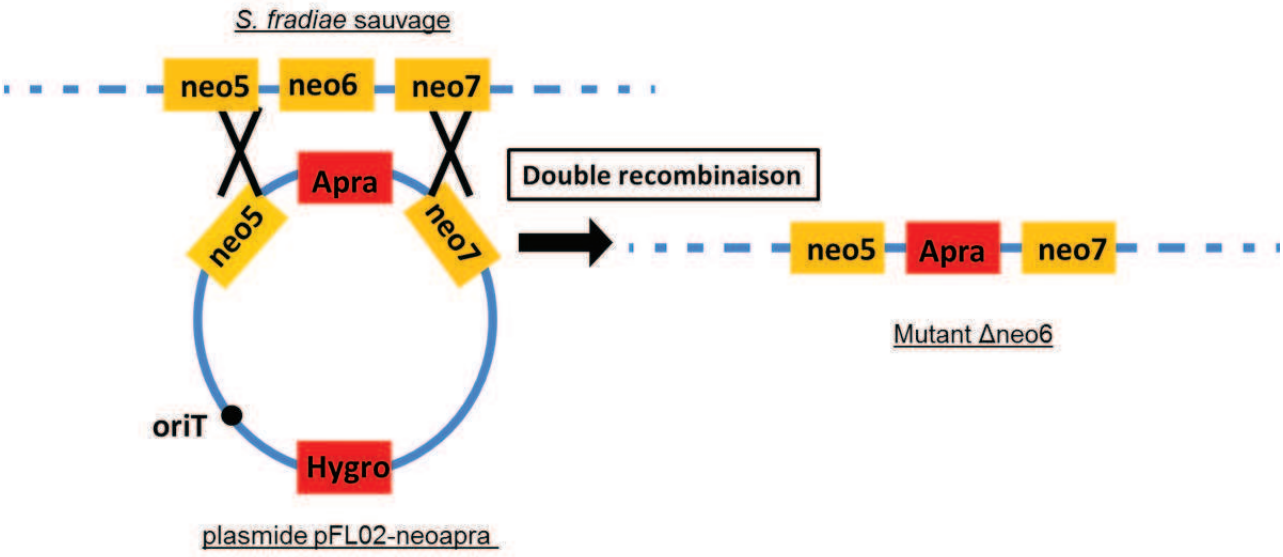



**Figure III-79 Principe de la suppression de neo6 par double recombinaison homologue**

La première étape sur laquelle nous nous sommes penchés était la conception du plasmide pFL02-neoApra. Nous avons pour cela utilisé les techniques classiques de biologie moléculaire[284] que nous ne ferons que survoler pour illustrer notre stratégie de clonage.

### 1) Stratégie de clonage du plasmide pFL02-neoApra

La stratégie de clonage est résumée dans la figure III-10. Pour obtenir le fragment neo5-neo7 de l'opéron de biosynthèse de la néomycine, nous avons extrait l'ADN total d'une culture de la souche *S.fradiae* sauvage commerciale (référence ATCC 10745). Nous avons mis au point des conditions de PCR permettant d'amplifier ce fragment que nous avons ensuite purifié par extraction sur gel (voir partie expérimentale pour les modes opératoires et les amorces utilisées). Du fait du pourcentage important en bases G et C du génome de *S.fradiae*, ces conditions ont nécessité un certain travail d'optimisation. Le fragment purifié a été phosphorylé puis inséré par ligation dans un plasmide de clonage pUC19, précédemment digéré par l'enzyme de restriction *SmaI* (coupure à bouts francs) et déphosphorylé pour éviter la ligation du plasmide sur lui-même. L'insert du plasmide obtenu a ensuite été modifié. Une digestion par l'enzyme de restriction *AfeI* (coupure à bouts francs), qui possède deux sites dans la séquence de neo6 séparés de 725 paires de bases, a permis d'obtenir deux fragments d'ADN : un fragment de 725 paires de base correspondant à une large partie du gène neo6 et un fragment plus grand correspondant au reste du plasmide. Ce dernier fragment a été purifié par extraction sur gel puis circularisé par ligation pour donner un deuxième plasmide contenant un site *AfeI* au milieu d'une région homologue au fragment neo5-neo7 dans laquelle l'essentiel du gène neo6 a été supprimé.

Ce deuxième plasmide a de nouveau été digéré par *AfeI* de façon à y insérer par ligation une cassette de résistance à l'apramycine issue de la digestion du plasmide pOSV226 (plasmide fourni par JL Pernodet) par *EcoRV* (coupure à bouts francs) (voir partie expérimentale pour la séquence). On obtient alors un troisième plasmide pUC19 possédant la cassette Apra flanquée de deux régions homologues au fragment neo5-neo7.

Ce troisième plasmide a été digéré par les enzymes de restriction *NdeI* et *XbaI* pour en séparer l'insert qui a été purifié par extraction sur gel puis inséré par ligation dans le plasmide pFL02. Ce dernier résulte d'une modification du plasmide pFL01 fourni par Jean-Luc Pernodet consistant en l'introduction d'un site de clonage multiple (« multi cloning site ») comprenant les sites de restriction *AflII, BglII, BsrGI* et *NdeI*. La digestion du plasmide pFL02 par les enzymes de restriction *NdeI* et *SpeI (*compatible *XbaI*) a permis d'insérer par ligation le fragment neo5-Apra-neo7. Cette ligation est très efficace car on a deux fragments à bouts cohésifs. On obtient alors le plasmide pFL02-neoApra.

A chaque étape du clonage, la séquence des plasmides a été vérifiée par carte de restriction et par PCR avec des amorces spécifiques. Les plasmides ont été amplifiés après transformation de bactéries *E.coli* DH5α puis purifiés par mini- ou midi-prep (kit NucleoSpin, Macherey-Nagel). Le séquençage du plasmide final, pFL02-neoApra, a révélé l'existence d'un certain nombre de mutations

ponctuelles dans les gènes neo5 et neo7, qui n'ont cependant pas été considérées gênantes pour la recombinaison avec le génome de *S.fradiae* sauvage.

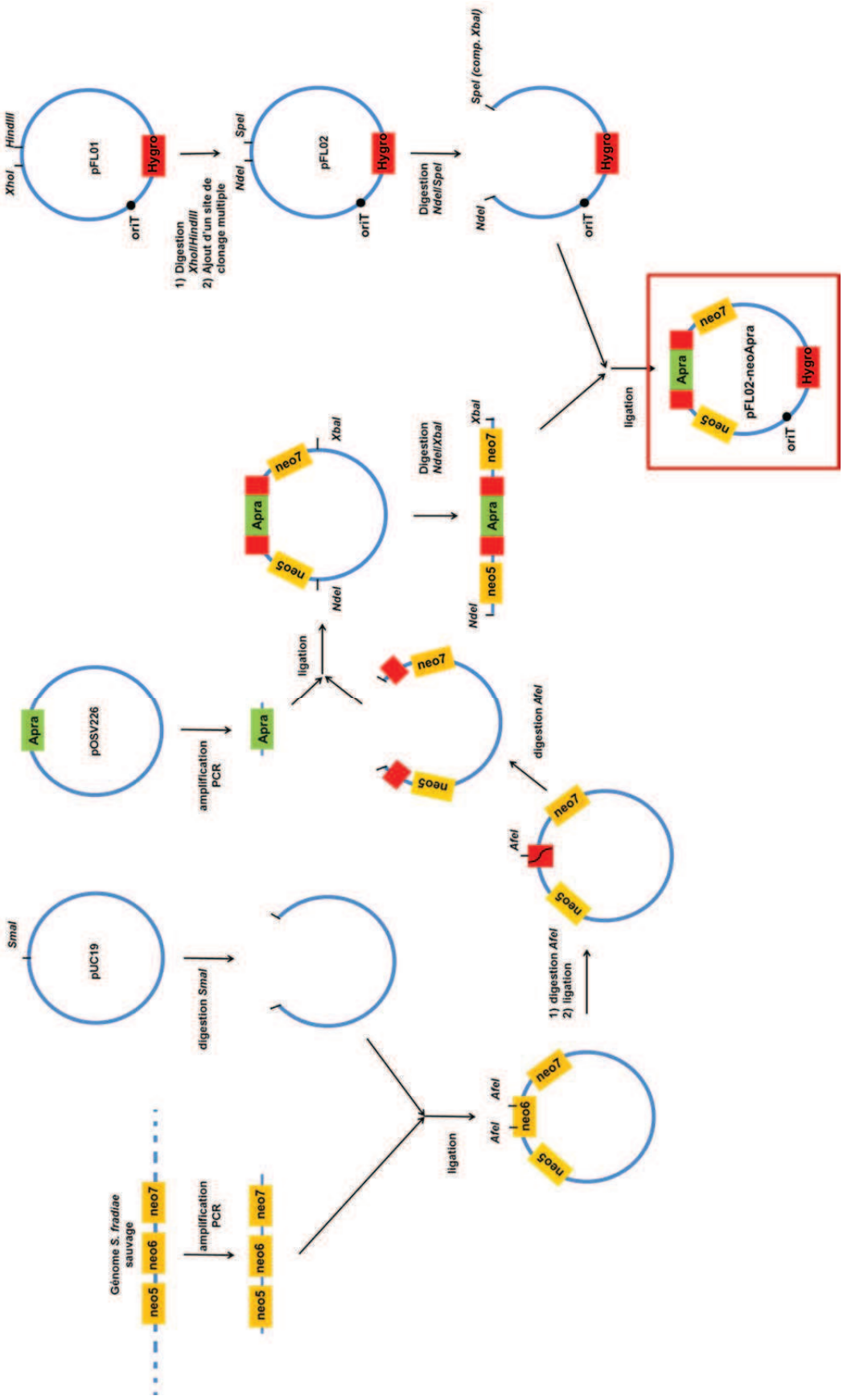



Figure III-80 Stratégie de clonage du plasmide pFL02-neoApra

### 2) Conjugaison interspécifique *E.coli* S17.1 / *S.fradiae* : suppression du gène neo6 par double recombinaison homologue

Une fois le plasmide pFL02-neoApra obtenu, nous sommes passés à l'expérience de double recombinaison homologue proprement dite selon le protocole décrit dans le livre de référence « Practical Streptomyces Genetics » écrit par Hopwood et ses collaborateurs[253] (voir partie expérimentale). Son principe est simple : on transforme la souche d'*E.coli* S17.1 avec le plasmide pFL02-neoApra. Cette souche est capable d'échanger ce plasmide par conjugaison avec *S.fradiae*. Par recombinaison homologue, le plasmide peut s'intégrer au génome de *S.fradiae*. Les bactéries issues d'un événement de simple recombinaison vont être résistantes aux deux antibiotiques et auront toujours une copie du gène neo6. Par contre, un événement de double recombinaison conduit au remplacement des gènes neo5-neo6-neo7 par le fragment neo5-Apra-neo7 porté par le plasmide pFL02-neoApra avec perte de la résistance à l'hygromycine (voir figure III-5).

Les bactéries sont sélectionnées pour leur résistance à l'apramycine et leur sensibilité à l'hygromycine. Ce criblage phénotypique a été fait pour environ une quarantaine de clones résistants à l'apramycine et nous a permis d'isoler une dizaine de clones issus d'un événement de double recombinaison. Nous avons pu vérifier par PCR la présence de la cassette de résistance à l'apramycine et l'absence de la cassette de résistance à l'hygromycine (voir figure III-11 et partie expérimentale). Nous avons ensuite travaillé sur un seul des clones identifiés (le clone n°8), que l'on appellera désormais souche Δneo6. Nous avons néanmoins fait des stocks de spores pour deux autres clones issus d'un événement de double recombinaison.

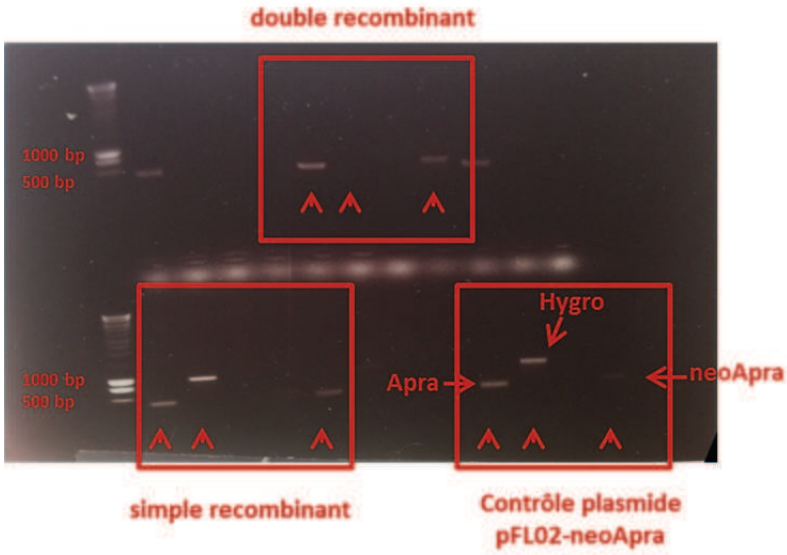



**Figure III-81 Vérification par PCR de l'événement de double recombinaison**
Chaque cadre rouge correspond à l'analyse PCR d'un mutant ou du plasmide controle. Dans le cas d'une double recombinaison, on a une bande caractéristique de la cassette de résistance à l'apramycine (Apra) (≈ 500 paires de bases) mais pas celle de l'hygromycine (Hygro) (≈ 800 paires de bases). On peut également voir le fragment correspondant à la liaison neo5/apra (≈ 500 paires de bases). Dans le cas d'une simple recombinaison ou du plasmide contrôle, les trois bandes sont présentes.

L'étape suivante a été la vérification de l'absence de production de néomycine par la souche Δneo6. Pour ce faire, nous avons cultivé la souche Δneo6 dans des conditions où la souche sauvage produit de la néomycine (milieu solide SFM pendant 7 jours à 30°C), puis recouvert la boite de pétri d'un top SNA ensemencé avec une bactérie sensible à la néomycine, *Micrococcus luteus* (voir partie expérimentale). La même expérience a été faite en parallèle avec la souche sauvage de *S.fradiae* et une souche obtenue après recombinaison simple du plasmide pFL02-neoApra. Les résultats de cette expérience sont présentés dans la figure III-12. Dans le cas de la souche sauvage, on voit un halo plus clair autour de la colonie qui correspond à une zone où la croissance de *M.luteus* est inhibée : il y a bien production de néomycine (contrôle positif). Par contre, dans le cas de la souche Δneo6, il n'y a pas de halo d'inhibition et donc, comme attendu, pas de production de néomycine : le mutant Δneo6 a bien le phénotype désiré. Enfin, le clone issu d'une simple recombinaison du plasmide semble produire encore un peu de néomycine car on peut distinguer un petit halo d'inhibition. Ce clone possède en effet encore une copie du gène neo6 mais le plasmide pFL02-neoApra est intégré dans le fragment neo5-neo7 ce qui induit sans doute un effet polaire qui bouleverse l'expression des gènes en aval et donc la biosynthèse de la DOS. Ces résultats ont plus tard été confirmés par HPLC-MS (voir partie suivante).

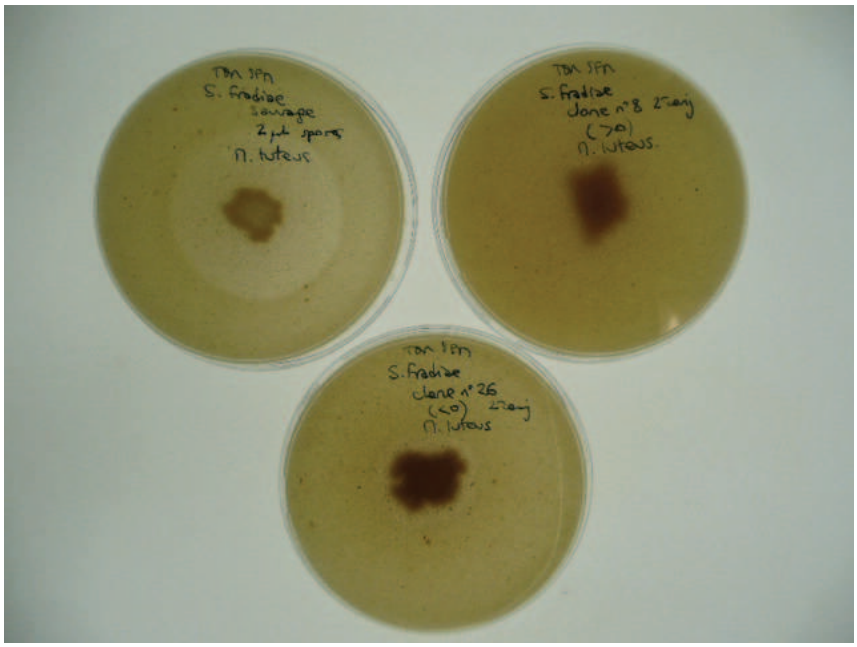


**Figure III-82 Antibiogramme comparé de la souche Δneo6 (en haut, à droite) avec la souche *S.fradiae* sauvage (en haut, à gauche) et la souche issue d'une simple recombinaison du plasmide pFL02-neoApra (en bas).**

Dans le but d'ajouter un contrôle supplémentaire, nous avons souhaité effectuer la complémentation génétique de la souche Δneo6 en réinsérant dans le génome de cette souche le gène neo6 sous le contrôle du promoteur ermE*, puissant promoteur constitutif des Streptomyces. Le protocole suivi a été identique à celui de recombinaison précédemment décrit, mettant cette fois en jeu la souche Δneo6 et le plasmide pOSV557 fourni par Jean-Luc Pernodet dans lequel nous avions inséré le gène neo6 (voir partie expérimentale). Ce plasmide, inspiré du vecteur intégratif pSET152, possède le site d'attachement du bactériophage phi C31 qui permet son intégration dans le chromosome de certaines souches de *Streptomyces*[317]. Cet essai de complémentation génétique n'a cependant pas été concluant. Plusieurs possibilités peuvent être proposées pour expliquer cet échec. L'effet polaire important de la délétion de neo6 pourrait empêcher l'expression correcte des gènes de biosynthèse en aval de la séquence codante[5]. Pour tester cette hypothèse, il serait intéressant d'utiliser le système de cassettes de résistance excisables développé récemment dans l'équipe de Jean-Luc Pernodet[318]. Ce système détourne en effet l'élément intégratif pSAM2 de *Streptomyces ambofaciens* pour exciser, en phase avec dans le cadre de lecture choisi, la cassette de résistance

[5] Cette explication serait en accord avec notre observation d'un fort effet polaire chez la souche de *S.fradiae* issue d'une simple recombinaison du plasmide pFL02-neoApra.

introduite lors de l'inactivation d'un gène par double recombinaison homologue. Il permet donc d'éliminer les éventuels effets polaires dus à un décalage du cadre de lecture.

Le plasmide de complémentation pourrait d'autre part ne pas être intégré dans le génome de *S.fradiae* ou bien l'expression du gène neo6 pourrait ne pas s'effectuer correctement. Si tel était le cas, on pourrait envisager d'utiliser d'autres systèmes d'insertion chromosomique spécifiques d'autres régions du génome et/ou mettre le gène neo6 sous le contrôle d'autres promoteurs de *Streptomyces*, voire sous le contrôle de son promoteur naturel. Enfin, il reste possible que les mutations ponctuelles introduites par double recombinaison dans les gènes neo5 et neo7 soient délétères.

Par manque de temps, nous n'avons pas poursuivi ces investigations et nous sommes concentrés sur l'utilisation de la souche mutante Δneo6, incapable de synthétiser de la DOS, dans le cadre d'une stratégie de mutasynthèse. Dans cette optique l'étape suivante consistait à mettre en place une méthode de purification et d'analyse de molécules de la famille des aminoglycosides avant de pouvoir passer à la préparation des mutasynthons et à la mutasynthèse proprement dite.

## *B - Purification et quantification de la production d'aminoglycosides*

Comme nous l'avons vu en introduction, la stratégie de mutasynthèse repose de façon cruciale sur notre capacité à détecter, identifier et quantifier la production des éventuels produits non-naturels formés dans le milieu de culture bactérien. Dans ce but, nous avons mis en place une stratégie d'analyse des molécules de la famille des aminoglycosides reposant sur une purification par échangeuse d'ion en SPE (pour « Solid Phase Extraction » en anglais) suivie d'une analyse par HPLC-MS/MS.

### 1) Purification par échangeuse d'ions (SPE)

Les aminoglycosides sont des molécules extrêmement polycationiques à pH physiologique en raison de leurs nombreuses fonctions amines. Cette propriété peut être exploitée avantageusement dans leur purification par colonne échangeuse d'ions à partir d'un milieu de culture[319].

Nous avons développé un protocole s'appuyant sur l'utilisation de colonnes SPE de type Strata-X-CW (Phenomenex) (voir partie expérimentale). Ces colonnes sont composées de billes de polymère greffées par des fonctions acide carboxylique (voir figure III-13). Elles sont faiblement polyanioniques

à pH physiologique et adsorbent donc sélectivement les molécules fortement polycationiques telles que les aminoglycosides.

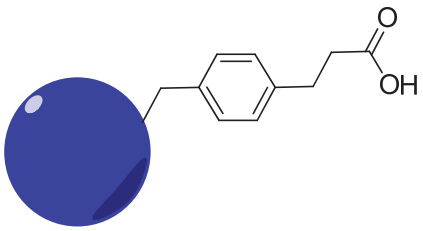



**Figure III-83 Greffage des billes de polymère des colonnes SPE Strata X-CW**

Le protocole de purification que nous avons développé comporte différentes phases de lavage et une élution à l'ammoniaque. Il permet une purification des échantillons suffisante pour permettre leur injection dans notre système HPLC.

### 2) Analyse des échantillons par HPLC-MS/MS

L'analyse des échantillons en HPLC-MS/MS a été effectuée en collaboration avec Anne Regazzetti au laboratoire C-TAC. Elle repose sur l'utilisation d'une colonne C18 en présence d'acide pentafluoropropionique, qui conférent à la colonne des propriétés d'échangeuse d'ions (voir partie expérimentale).

Pour tester notre méthode d'analyse, nous avons tout d'abord effectué des essais sur des échantillons commerciaux purs de néomycine d'abord par introduction directe (infusion) dans le spectromètre de masse (figure III-14), puis après passage par la colonne HPLC (figure III-15). On retrouve bien le spectre de masse attendu pour la néomycine, qui a un temps de rétention $t_R$ d'environ 9.5 min ($m/z \approx 615$). Une analyse en $MS^n$ a permis d'identifier le schéma de fragmentation de la néomycine. En particulier, en fragmentant le composé correspondant à la masse de 455, il est possible de faire apparaitre un fragment d'une masse de 163 correspondant à la DOS (voir partie expérimentale).

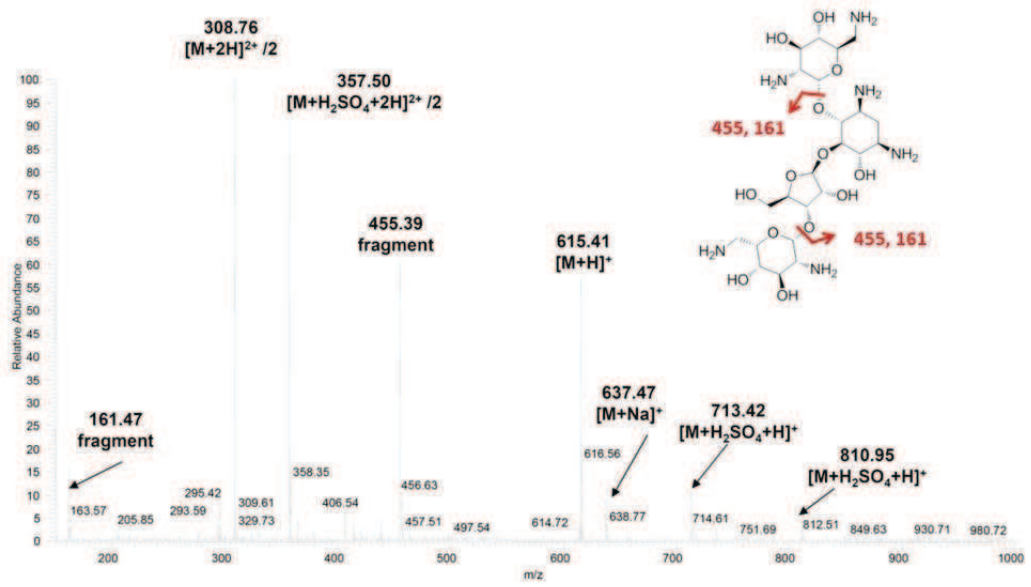



**Figure III-84 Spectre de masse du sulfate de néomycine obtenu par infusion**

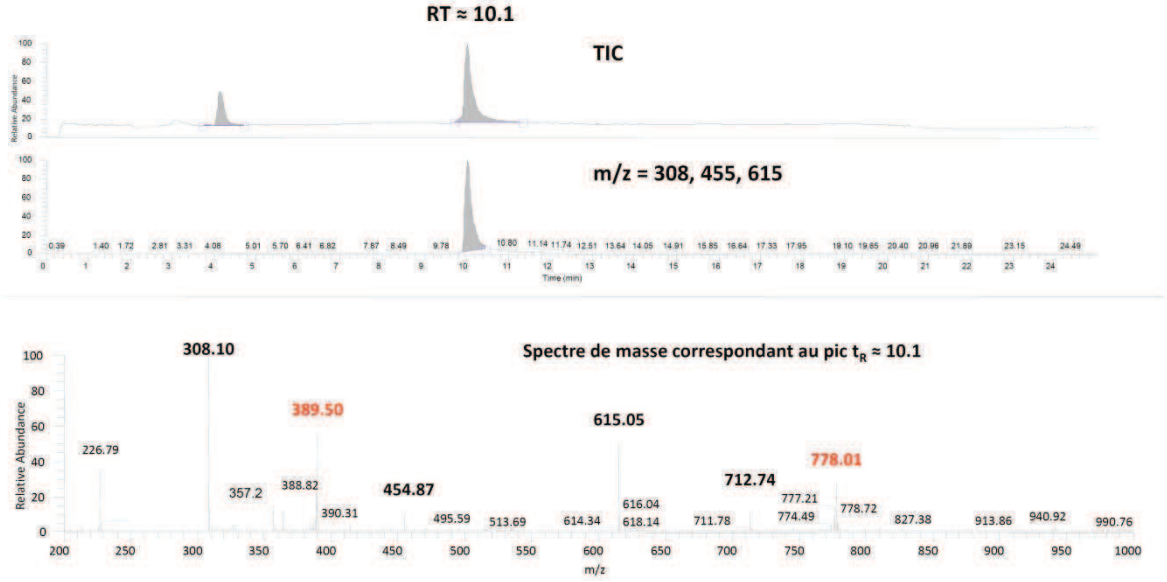



**Figure III-85 Caractérisation en HPLC-MS de la néomycine purifiée à partir d'un milieu de culture TSB (chromatogramme reconstitué à m/z = 615, 455, 308) et spectre de masse associé au pic à $t_R$ = 10.1)**

Nous nous sommes ensuite intéressés aux possibles biais expérimentaux liés à la purification sur colonne SPE. Nous avons pu vérifier que cette purification ne modifiait pas l'intégrité de la néomycine (même spectre de masse) et nous avons pu estimer la perte de produit liée à la purification en comparant la variation de l'aire sous la courbe en fonction de la concentration en

néomycine avec ou sans purification par SPE (voir figure III-16). Comme dans toute purification, il y a une perte de produit mais celle-ci reste néanmoins compatible avec la détection par HPLC-MS.

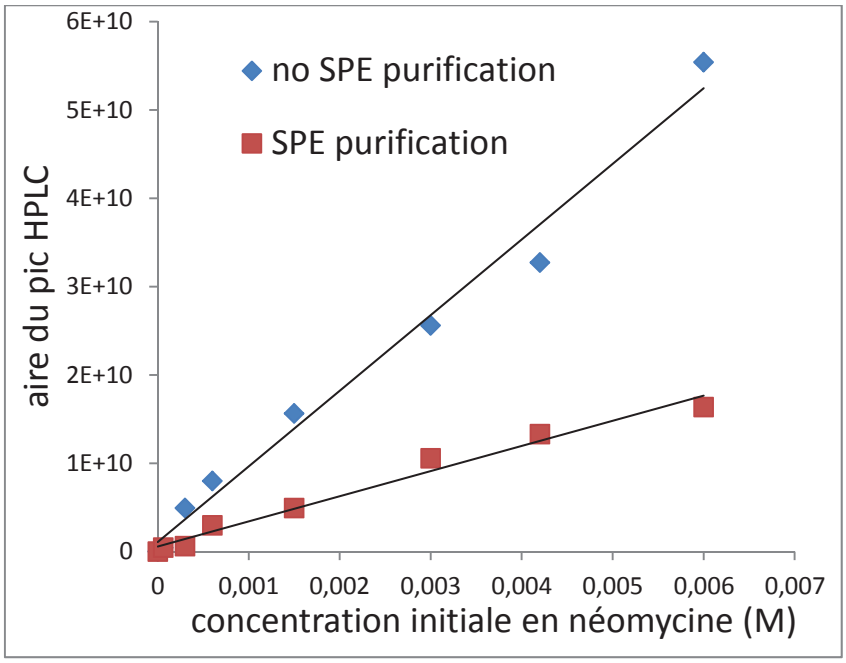



**Figure III-86 Evaluation de la perte de néomycine après purification SPE**

L'étape suivante a été de vérifier que l'on était capable de détecter la néomycine présente dans un mélange complexe comme un milieu de culture. Nous avons calibré notre méthode d'analyse en testant différentes concentrations de néomycine dans un milieu de culture, le TSB, de façon à réaliser une droite d'étalonnage. Chaque échantillon a été purifié sur SPE puis analysé en HPLC-MS. L'aire du pic de néomycine obtenu est alors reliée à la concentration initiale en néomycine du milieu de culture. Comme on peut le voir dans la figure III-17, la limite de détection de notre méthode d'analyse se situe aux alentours de 10 µg/mL (c'est-à-dire environ 16 µM) et la réponse est quasi-linéaire jusqu'à environ 500 µg/mL (environ 0.8 mM). L'ordre de grandeur des concentrations détectées semble correct au vu des concentrations en antibiotiques habituellement produites par les souches du genre *Streptomyces* (de l'ordre de la centaine de microgrammes par millilitre). Lors d'une expérience de mutasynthèse, on peut par ailleurs espérer être dans cette gamme de concentration si l'on a des mutasynthons suffisamment bien reconnus par la voie de biosynthèse modifiée.

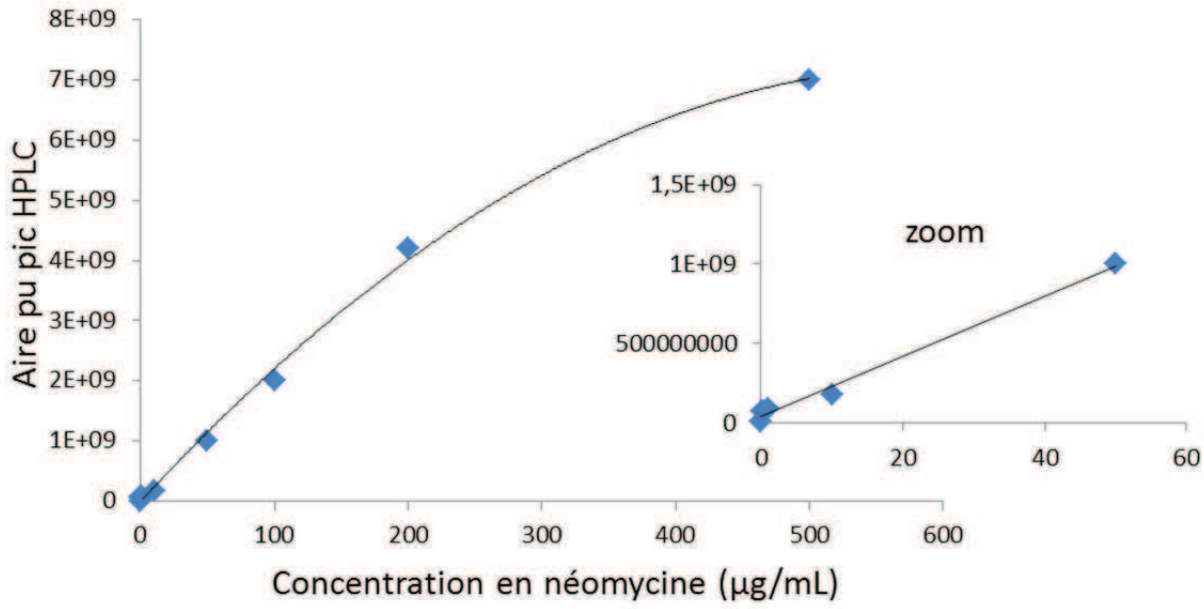



**Figure III-87 Etalonnage de la mesure par HPLC-MS de la concentration en néomycine d'un milieu de culture TSB après purification par SPE**

## *C - Optimisation de la production de néomycine par la souche sauvage : définition des conditions de mutasynthèse*

### 1) Criblage de différents milieux de culture

Une mesure de la concentration en néomycine d'un milieu de culture étant disponible, nous avons cherché à optimiser les conditions de culture de la souche *S.fradiae* sauvage de façon à obtenir une production maximale de néomycine. Le présupposé de cette démarche étant que ces conditions sont aussi les conditions optimales pour les expériences de mutasynthèse (ce qui resterait néanmoins à vérifier).

Nous nous sommes fondés sur les conditions classiques de culture des *Streptomyces* ainsi que sur la description de milieux de culture connus pour susciter la production d'antibiotiques (culture dans un fernbach à picots de 250 mL pendant 7 jours à 30 °C avec une agitation à 220 rpm, voir partie expérimentale). Nous nous sommes en particulier intéressés au milieu de culture minimal décrit dans les années 80 par Berwick[320] pour la production de néomycine chez S.fradiae dans l'idée qu'il serait plus simple d'y purifier d'éventuels produits issus de la mutasynthèse.

Nous avons testé 7 milieux de culture différents et avons constaté que le milieu TSB était le plus adapté à la production de néomycine, avec des concentrations finales de l'ordre de 480 µg/mL (voir figure III-18). Dans nos mains, le milieu minimal précédemment décrit ne conduit qu'à des

productions très modestes loin des rendements annoncés. Dans la suite de notre étude, toutes les expériences sont effectuées dans du milieu TSB.

Dans le but de permettre une éventuelle purification préparative d'un produit obtenu par mutasynthèse, nous avons adapté notre protocole à l'utilisation d'un fermenteur de 4L. De façon satisfaisante, la fermentation de *S.fradiae* sauvage dans du TSB à cette échelle conduit à des titres en néomycine importants (de l'ordre de 600 µg/mL) (voir partie expérimentale).

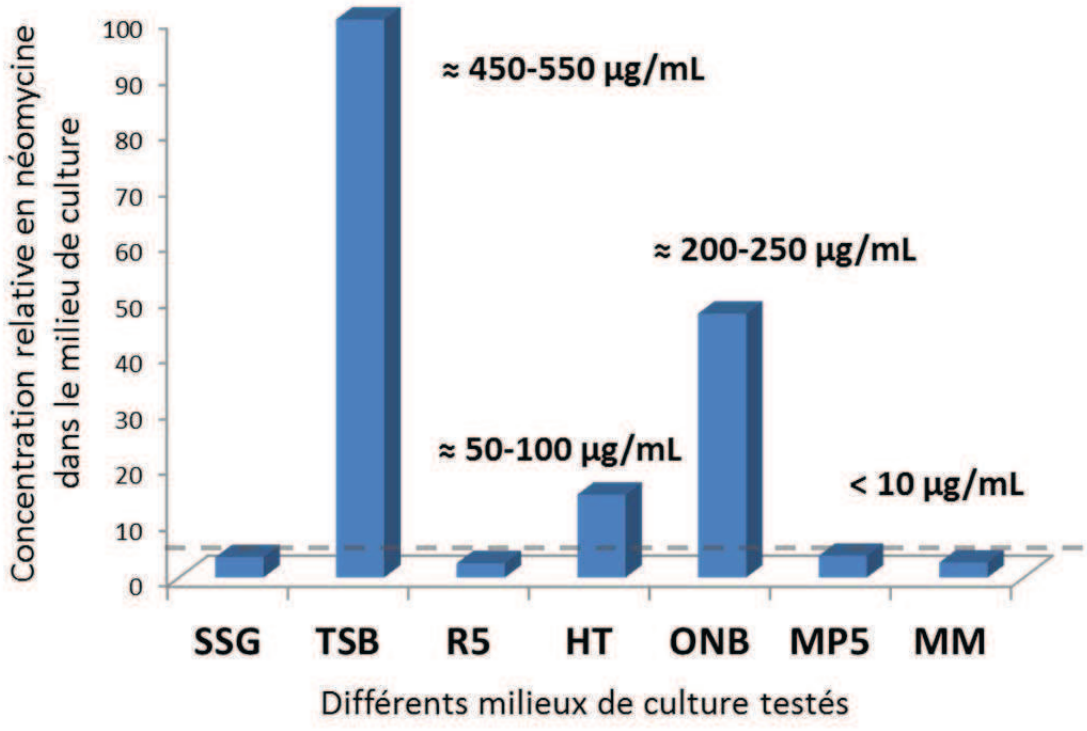



**Figure III-88 Comparaison de la production en antibiotique de la souche *S.fradiae* sauvage dans différents milieux**

### 2) Cinétique de production de la néomycine

La cinétique de production de la néomycine était importante à connaitre pour planifier au mieux l'ajout des mutasynthons. Nous nous sommes donc intéressés à cette cinétique dans la souche de *S.fradiae* sauvage. Dans la figure III-19, on voit que la production de néomycine commence entre le deuxième et le troisième jour et tend vers un plateau aux alentours du sixième jour. L'augmentation du titre en néomycine pendant le septième jour est sans doute due à la concentration progressive du

milieu de culture par évaporation. Pour éviter ce phénomène, on peut travailler dans un incubateur INFORS où le taux d'humidité est contrôlé.

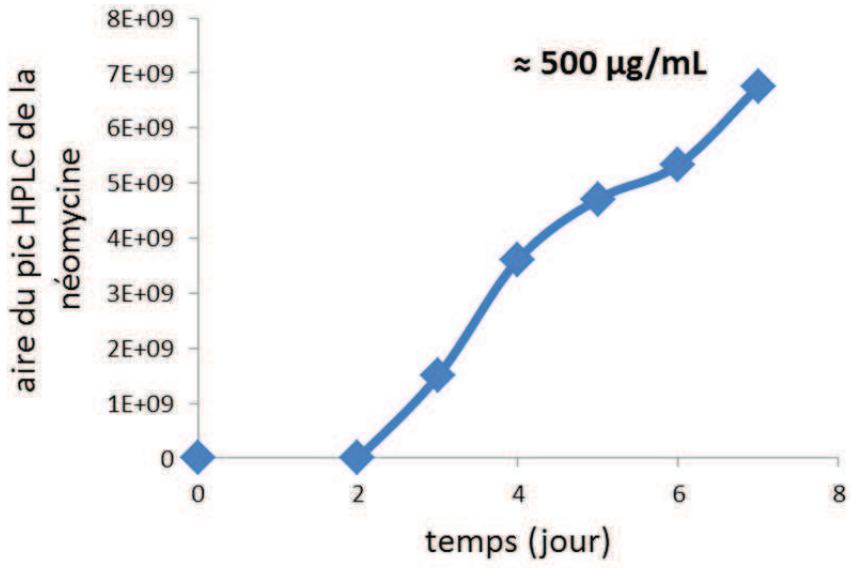



**Figure III-89 Cinétique de production de la néomycine par *S.fradiae* dans le TSB. La concentration finale en néomycine avoisine les 500 µg/mL.**

Au vu de cette cinétique, à l'inverse de ce qui a été fait dans les premières expériences de mutasynthèse[250], on peut imaginer qu'il est inutile d'ajouter les mutasynthons avant le deuxième jour de culture permettant ainsi de potentiellement éviter leur éventuelle dégradation ou leur consommation par des processus cellulaires autres que la voie de biosynthèse modifiée. Alternativement, on peut envisager d'ajouter le mutasynthon par petites quantités au fur et à mesure de la fermentation voire même de façon continue. Ce mode opératoire permet de s'assurer qu'il n'y a pas d'accumulation de certains des métabolites intermédiaires de la voie de biosynthèse qui sont potentiellement toxiques pour la bactérie. Il permet aussi de diminuer les chances de dégradation du mutasynthon utilisé. Ce protocole a été employé par l'équipe de Kirschning pour la mutasynthèse de nouveaux analogues de geldanamycine à partir d'une souche de *S.hygroscopus* mutée[321]. Nous reviendrons dans la partie III-II-E-1 sur les différentes possibilités d'introduction des mutasynthons dans le milieu de culture que nous avons testées (voir partie expérimentale).

## *D - Synthèse des mutasynthons*

L'étape suivante du projet était la synthèse d'une petite librairie d'analogues d'aminoglycosides pouvant servir de mutasynthons. Les premiers mutasynthons vers lesquels nous nous sommes tournés sont des métabolites naturels de la voie de biosynthèse de la néomycine dont la synthèse est rapidement décrite dans la partie III-II-D-1. Ces mutasynthons sont utilisés comme contrôles positifs du bon fonctionnement de la voie de biosynthèse modifiée, la complémentation chimique de la souche Δneo6 avec ces substrats naturels devant théoriquement restaurer la production de néomycine.

Nous avons dans un second temps synthétisé des métabolites non-naturels de la voie de biosynthèse de la néomycine en nous concentrant sur des analogues de la DOS ou des analogues pseudodisaccharidiques de la paromamine. La transformation de ces mutasynthons par la souche Δneo6 conduirait à des analogues de néomycine. Leur synthèse est décrite dans la partie III-II-D-2. On peut trouver dans la partie expérimentale l'intégralité des protocoles de synthèse de ces produits ainsi que leur caractérisation complète.

### 1) Métabolites « naturels »

Les métabolites naturels de la biosynthèse de la néomycine peuvent être obtenus par dégradation sélective de néomycine ou de paromamine selon des protocoles décrits. Ces produits étant très hydrosolubles, leur purification est habituellement effectuée par précipitation sélective, recristallisation ou échange d'ion.

#### a) Synthèse de la désoxystreptamine

La désoxystreptamine (DOS) peut être obtenue par dégradation de la néomycine dans du HBr concentré porté à reflux pendant trois jours. Nous avons adapté le protocole décrit par l'équipe de Puglisi[241] pour produire de la DOS à l'échelle de la dizaine de grammes (voir figure III-20). La difficulté de cette synthèse réside dans la purification de la DOS à partir du brut réactionnel. Nous avons adopté une stratégie reposant sur un traitement au charbon actif, des filtrations successives (les produits de dégradation polymérisent) et l'emploi d'une phase échangeuse d'ion de type Amberlyst A-15.

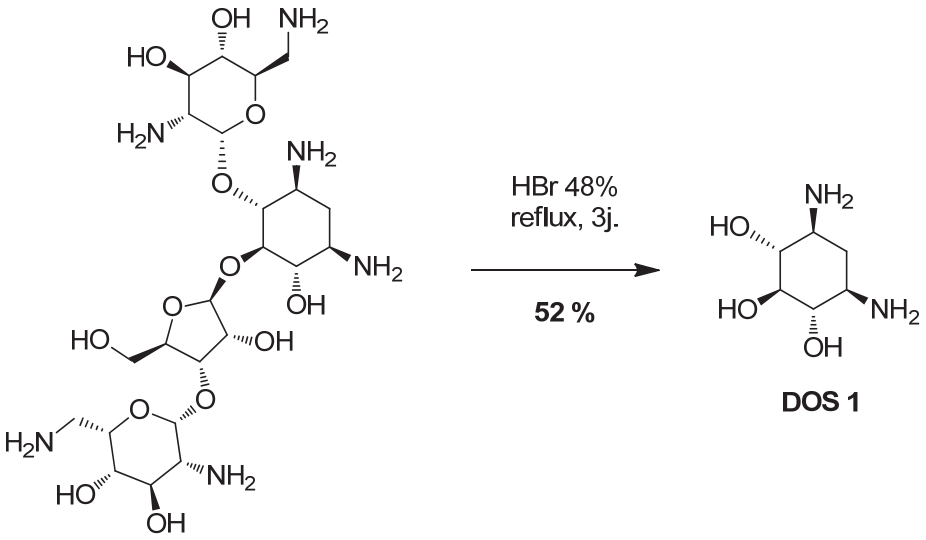



Figure III-90 Synthèse de la DOS par dégradation de la néomycine

### b) Synthèse de la paromamine et de la néamine

La paromamine est obtenue à partir de la dégradation de la paromomycine en milieu acide dans le méthanol. Nous nous sommes inspirés du protocole publié par Spencer[278] en l'adaptant à la production de paromamine à l'échelle de la dizaine de grammes et en optimisant son rendement. On récupère le chlorhydrate de paromamine par simple précipitation.

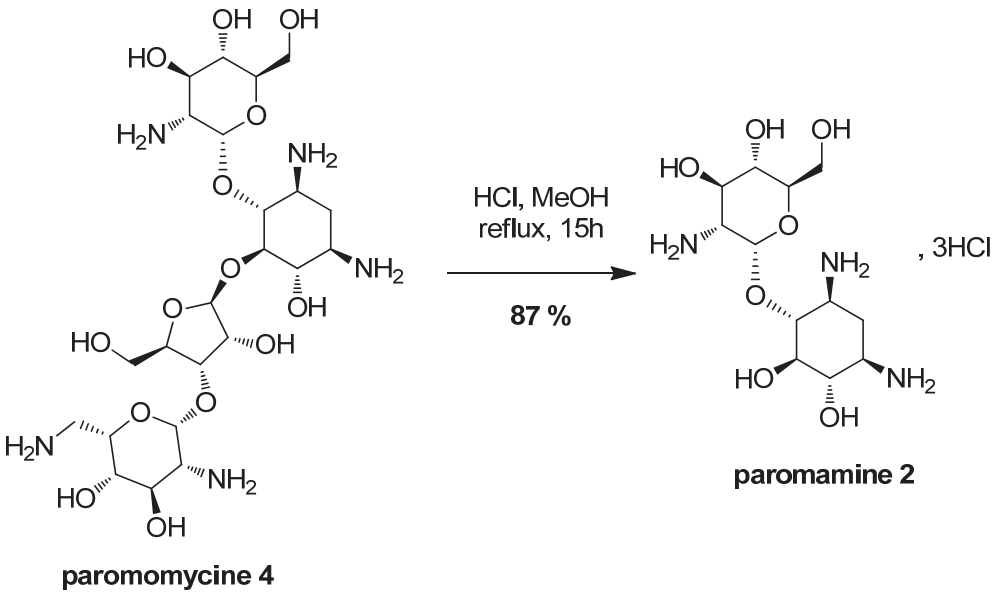



Figure III-91 Synthèse de la paromamine par dégradation sélective de la paromomycine

La néamine est obtenue selon un protocole similaire à la paromamine en dégradant la néomycine selon le protocole décrit par Wong[322].

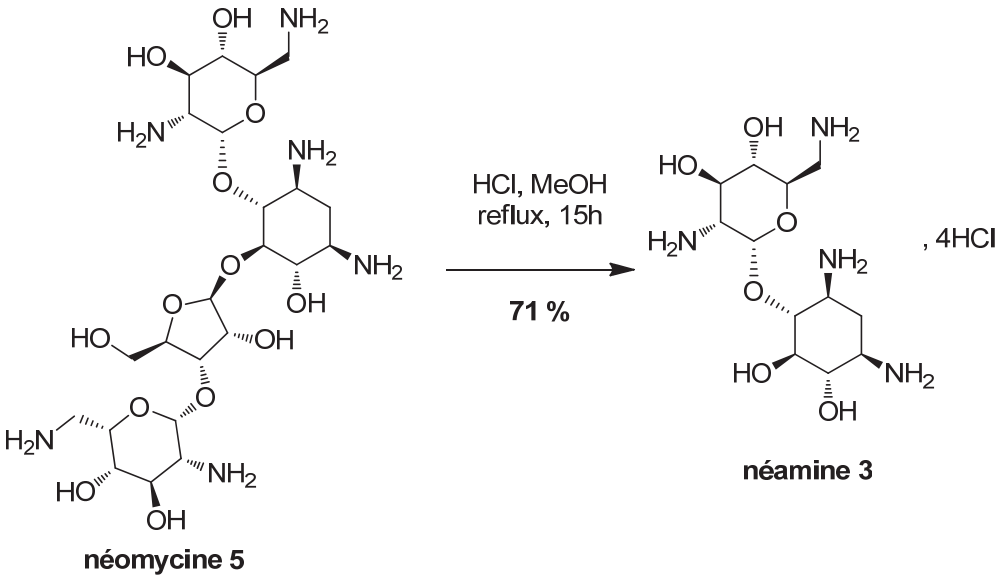



Figure III-92 Synthèse de la néamine par dégradation sélective de la néomycine

### 2) Métabolites « non-naturels »

#### a) Synthèse de la streptidine et de la streptamine

La streptidine et la streptamine ont fait partie des premières molécules à être testées dans les études initiales de mutasynthèse d'analogues d'aminoglycosides de Rinehart[251,250]. Du fait de sa similitude structurale avec la DOS, la streptamine a longtemps constitué le seul exemple de mutasynthon exogène pris en charge par la voie de biosynthèse modifiée pour donner un analogue de néomycine, l'hybrimycine A, possédant un alcool supplémentaire en position 2. Il nous a semblé intéressant de revenir à ces exemples initiaux pour pouvoir comparer notre souche Δneo6 avec la souche initiale isolée par Rinehart, la souche DSM 41550.

Ces deux analogues de la DOS peuvent être obtenus très simplement par dégradation de la streptomycine. Nous avons suivi les protocoles initialement décrits par Folkers[323,324] et adaptés par l'équipe de Puglisi[241]. La synthèse de la streptidine n'a pas posé de problème particulier et elle a pu être purifiée par recristalisation à l'échelle de plusieurs grammes. La purification de la streptamine

est plus délicate et peut se faire par traitement par une phase échangeuse d'ion selon le même protocole que la DOS.

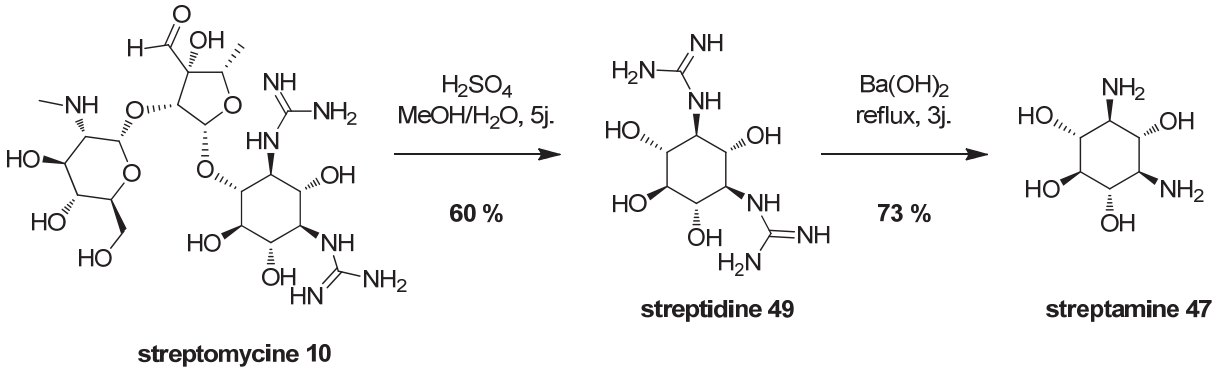



**Figure III-93 Synthèse de la streptidine et de la streptamine par dégradation sélective de la streptomycine**

### b) Synthèse du *cis*-1,3-diaminocyclopent-4,5-diol C5

Des études préliminaires au laboratoire ont démontré que le *cis*-1,3-diaminocyclopent-4,5-diol pouvait constituer un mime convenable de la DOS vis-à-vis de partenaires nucléiques[62,325] ou protéiques[130]. Nous avons donc voulu tester si ce cycle à 5 pouvait être reconnu par la voie de biosynthèse de la néomycine dans la souche mutante Δneo6.

La première étape de la synthèse met en jeu une réaction de Diels-Alder entre le cyclopentadiène fraichement distillé à partir de son dimère et le diazocarboxylate de dibenzyle selon un protocole décrit dans la littérature[326]. L'étape suivante repose sur la dihydroxylation à l'osmate de potassium de l'hydrazine bicyclique **75** obtenue[327,328]. Le contrôle de la stéréosélectivité de la réaction (dihydroxylation sur la face *exo* de l'hydrazine exclusivement) est assuré par l'encombrement stérique causé par un des deux groupements carbamate qui vient pointer vers la face *endo*. Le composé dihydroxylé **77** obtenu peut alors être ouvert par hydrogénolyse catalysée au palladium sur charbon pour donner le *cis*-1,3-diaminocyclopent-4,5-diol **78** avec un très bon rendement de 82 % sur 3 étapes à partir du cyclopentadiène à l'échelle de plusieurs grammes (voir figure III-24). Dans la suite du manuscrit, ce composé sera nommé **C5**.

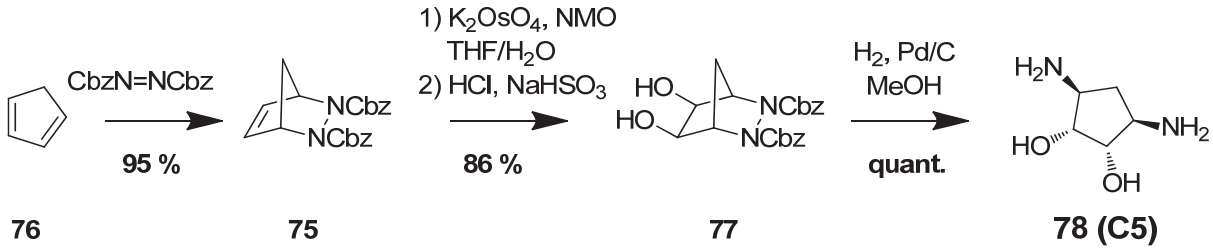



**Figure III-94 Synthèse du *cis*-1,3-diaminocyclopent-4,5-diol 79, aussi appelé C5**

### c) Synthèse de la paromamine avec un azoture en position 6' (paro$N_3$)

Comme nous l'avons vu dans la partie introductive, les travaux initiaux de Sato suggèrent qu'il est possible d'utiliser des analogues de paromamine en tant que mutasynthons[305]. Nos propres travaux suggèrent par ailleurs que la voie de biosynthèse de la néomycine a une certaine tolérance pour les substituants en position 6' de la paromamine (voir ultérieurement dans la partie III-II-E-1).

Dans cette optique, nous avons voulu développer un analogue de paromamine modifié en position 6' par une fonction azoture. A travers la réaction de cycloaddition de Huisgen (chimie « click »), la fonction azoture offre en effet de nombreuses potentialités de fonctionalisation ultérieure des éventuels produits formés par mutasynthèse. On peut de plus envisager de développer des réactions bioorthogonales extrêmement efficaces et compatibles avec des systèmes biologiques complexes[329]. Il existe ainsi dans la littérature des exemples où de tels composés sont utilisés *in vivo* pour marquer spécifiquement une cible biologique d'intérêt[330]. L'intérêt de développer un analogue d'aminoglycoside possédant une fonction azoture serait donc très grand, à la fois pour pouvoir synthétiser rapidement d'autres analogues par chimie click mais aussi pour étudier les cibles biologiques des aminoglycosides. Dans la suite du manuscrit, l'analogue de paromamine modifié en position 6' par une fonction azoture sera nommé **paro$N_3$**.

La synthèse de **paro$N_3$** commence par la synthèse de la paromamine **2** à partir de la paromomycine **4** selon le protocole précédemment évoqué. Après avoir protégé les amines par des groupements *tert*-butoxycarbonyle (Boc), on peut sélectivement faire réagir le chlorure de tosyle sur la fonction alcool primaire du composé **79**, plus réactive que les fonctions alcools secondaires plus encombrées. On obtient alors le composé **80** dont on peut substituer le tosylate qui est un bon groupe partant par l'azoture de sodium. La dernière étape consiste à déprotéger le produit azidé **81** obtenu pour donner **paro$N_3$** avec un rendement global 26.5 % en 5 étapes à partir de la paromomycine **4**.

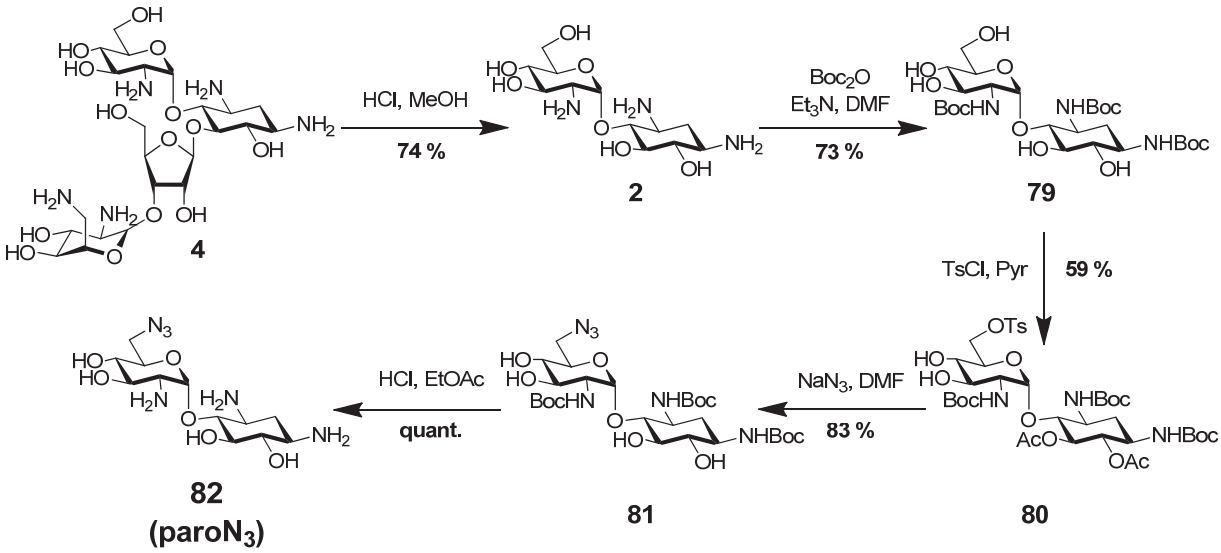
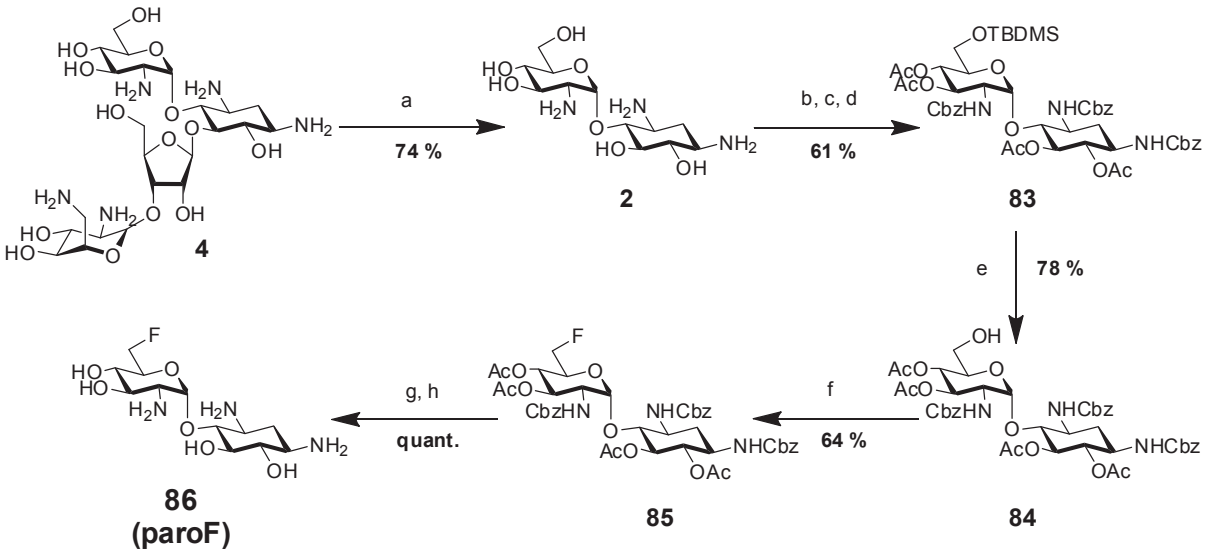



**Figure III-95 Synthèse de la paromamine avec un azoture en position 6' (paroN$_3$)**

### d) Synthèse de la paromamine fluorée en position 6' (paroF)

La synthèse de la paromamine fluorée en position 6', que nous appellerons **paroF**, sera détaillée ultérieurement dans le chapitre sur l'utilisation de sondes fluorées dans l'étude des ARN (partie IV III B 4 a). Elle met en jeu une série de protections et déprotections sélectives permettant de fluorer spécifiquement la fonction alcool en 6' (voir figure III-26) et conduit à la paromamine fluorée en position 6' en 8 étapes à partir de la néomycine avec un bon rendement global de 22.5 %.




a) HCl, MeOH, reflux, 24h; b) CbzCl, $Na_2CO_3$, toluene/eau/acetone, -5°C jusqu'à T.A., 4h;
c) TBDMSCl, DMAP, $Et_3N$, DMF, -5°C, 1h; d) $Ac_2O$, DMAP, pyridine, T.A., 1 d.; e) HF-pyridine, pyridine, THF, T.A., 4h;
f) DAST, dichloroethane sec, -5°C jusqu'à T.A., 5 h; g) NaOMe, MeOH, THF, dioxane, T.A., 3h;
h) $H_2$, Pd/C, dioxane, MeOH, eau, T.A., 2j.

**Figure III-96 Synthèse de la paromamine fluorée en position 6' (paroF)**

## *E - Mutasynthèse*

### 1) Complémentation chimique avec les métabolites naturels de la voie de biosynthèse de la néomycine

#### a) Vérification du phénotype des mutants DSM 41550 et Δneo6 par HPLC-MS

Avant de commencer les expériences de complémentation chimique, nous avons tout d'abord confirmé par HPLC-MS les résultats du criblage phénotypique ayant conduit à la sélection des souches mutantes DSM 41550 et Δneo6. Comme on peut le voir sur la figure III-27, les deux souches de *S.fradiae* mutantes ne produisent plus de néomycine lorsqu'elles sont cultivées dans le TSB sans ajout de DOS exogène (le pic résiduel que l'on peut voir pour les souches mutées correspond à la limite de détection de notre méthode d'analyse et représente du bruit de fond. Son spectre de masse ne correspond pas à la néomycine).

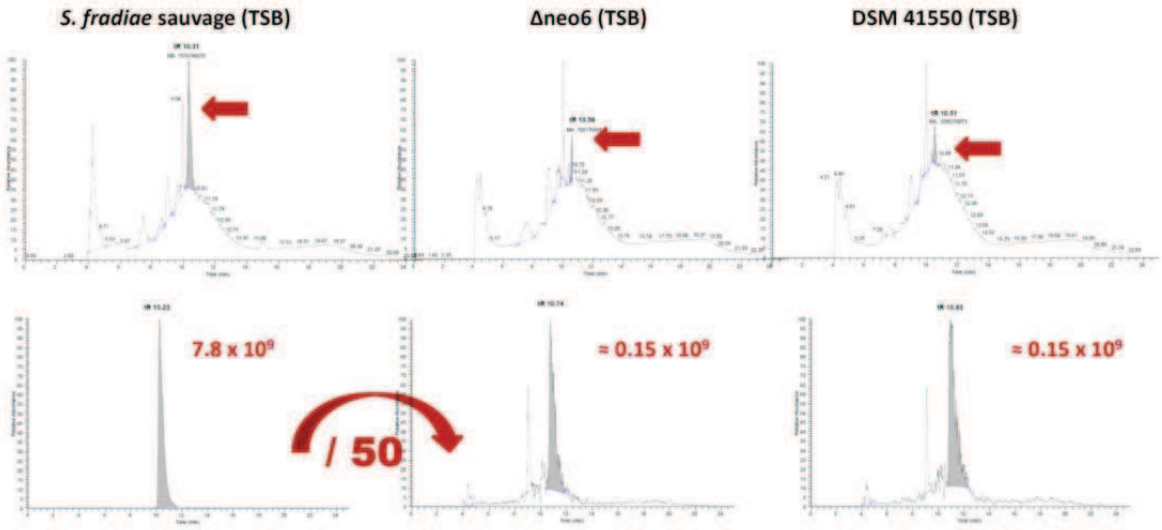



**Figure III-97 Comparaison de la production de néomycine de la souche sauvage de S.fradiae avec les souches mutantes Δneo6 et DSM41550.**

En haut, chromatogrammes HPLC-MS non filtrés, en bas, chromatogrammes correspondant à la masse de la néomycine (m/z = 615). Les aires des pics HPLC sont notées en rouge. La souche sauvage produit de l'ordre de 400 à 500 µg/mL de néomycine. Les souches mutantes ne produisent pas de concentration mesurable de néomycine (pic HPLC dans le bruit de fond, production inférieure au seuil de détection <10 µg/mL).

Par la suite, chacune des séries d'expériences de mutasynthèse est effectuée en parallèle d'une expérience de production de néomycine par la souche sauvage dans le TSB. Cette expérience sert à la fois de calibration et de contrôle positif, permettant de corriger les éventuelles variations stochastiques de la production d'aminoglycosides liées au milieu ou aux conditions d'incubation.

Cette vérification faite, nous sommes passés aux expériences de mutasynthèse avec des métabolites naturels de la voie de biosynthèse de la néomycine, que l'on peut retrouver dans la figure III-28.

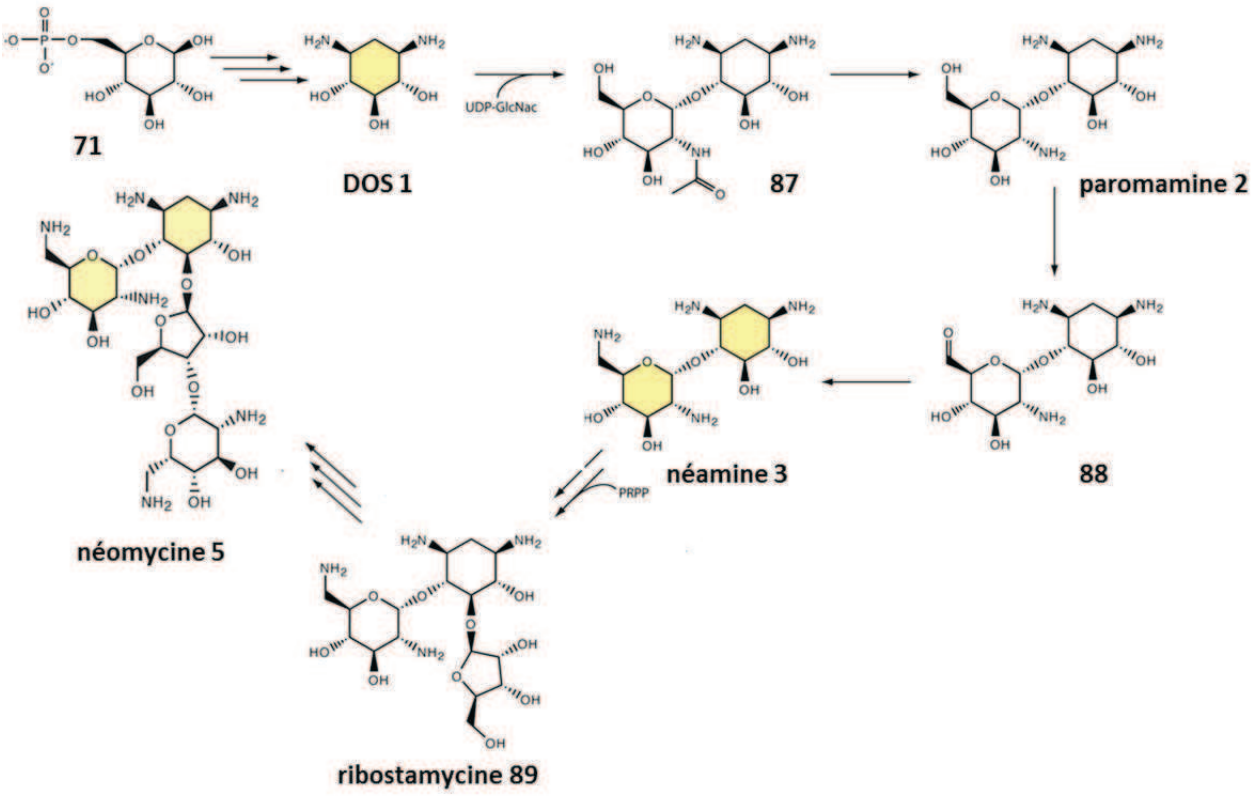



**Figure III-98 La voie de biosynthèse de la néomycine chez *S.fradiae* et les différents métabolites naturels**

### b) Complémentation avec la désoxystreptamine (DOS) et étude de l'influence de la concentration en DOS sur la régulation de la biosynthèse chez *S.fradiae*

La première expérience visait à vérifier qu'il était effectivement possible de restaurer la production de néomycine en ajoutant de la DOS dans le milieu de culture des souches mutées DSM 41550 et Δneo6. Nous avons testé différentes concentrations en DOS et différents protocoles d'ajout au milieu de culture : en une fois, au début de la culture ou après quelques jours, ou bien par additions successives au fur et à mesure de la culture (voir partie expérimentale et partie III-II-C-2). N'ayant pas constaté de différences significatives entre les différents protocoles, nous avons adopté un protocole standard qui, sauf indication contraire, consiste à ajouter le mutasynthon en une fois en début de culture de façon à atteindre une concentration de 2mM. La figure III-29 présente de façon synthétique les résultats de ces expériences.

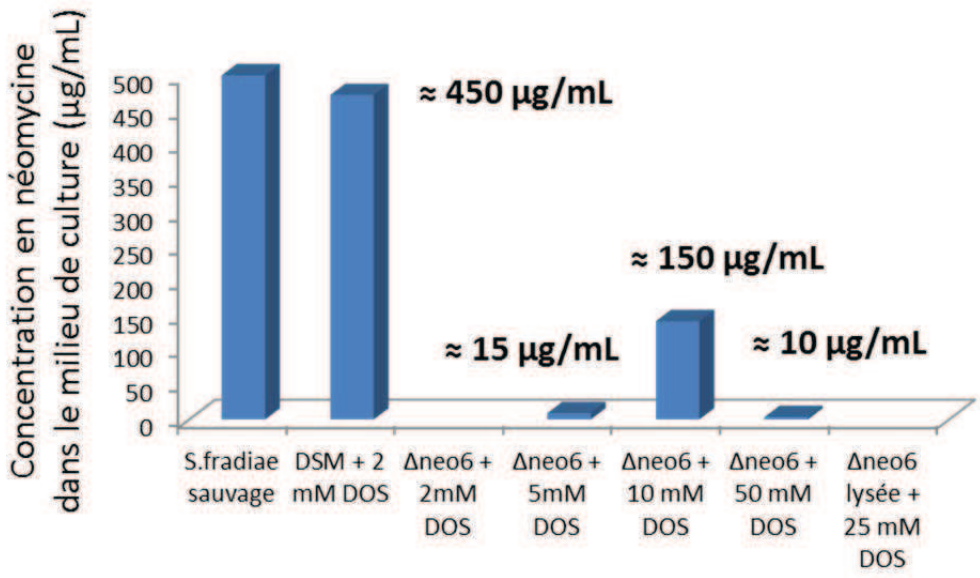



**Figure III-99 Complémentation chimique de DSM 41550 et Δneo6 par la DOS**

Conformément à ce qui avait été décrit par Rinehart[250], on retrouve de très bons niveaux de production de néomycine lorsqu'on complémente la souche DSM 41550 avec 2 mM de DOS. Par contre, dans les mêmes conditions, le mutant Δneo6 ne conduit pas à une production de néomycine détectable et ce n'est qu'à des concentrations en DOS plus importantes que l'on restaure la production. La concentration optimale de DOS pour la production de néomycine semble être aux alentours de 10 mM, une concentration plus importante conduisant à l'abolition de la production de néomycine[6]. Suspectant qu'un mécanisme de rétrocontrôle négatif était la raison de cette abolition, nous avons analysé l'influence de la concentration en DOS sur la production de néomycine par la souche sauvage. Les résultats de ces expériences, présentés dans la figure III-30, montrent qu'à une concentration de 50 mM en DOS, on ne détecte pas de production de néomycine par la souche sauvage. A 40 mM, on peut évaluer la concentration finale en néomycine à 50 µg/mL, bien loin des 500 µg/mL habituellement obtenus dans les conditions de cultures classiques.

[6] Les concentrations en DOS utilisées sont inférieures au seuil de toxicité conduisant à la mort des bactéries (biomasse récupérée en fin de fermentation semblable à une expérience contrôle). L'abolition de la production de néomycine n'est donc pas due à la mort des bactéries.

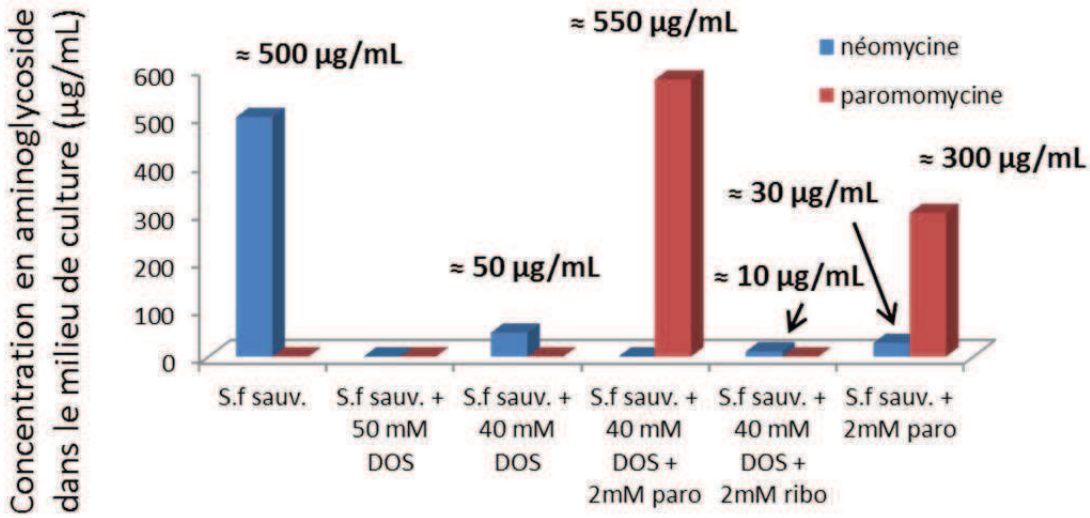



**Figure III-100 Influence de la concentration de divers métabolites sur la production d'aminoglycoside par la souche *S.fradiae* sauvage[7]**

On peut trouver deux explications à ces observations. La première est qu'il y a bien un rétrocontrôle négatif de la concentration de DOS sur la production de néomycine : une forte concentration de DOS conduisant à l'inhibition d'une ou plusieurs étapes de la voie de biosynthèse de la néomycine ou à la diminution de l'expression des gènes codant pour les enzymes correspondantes. Cette explication offre l'avantage de faire sens d'un point de vue biologique. L'autre explication possible est qu'à de telles concentrations en DOS, on sature les colonnes échangeuses d'ion servant à purifier les aminoglycosides à partir du milieu de culture bactérien : la néomycine et la DOS étant en compétition pour la fixation sur les colonnes SPE, un excès de DOS empêche la fixation de la néomycine et donc sa détection en HPLC-MS.

Pour répondre à cette question, en présence d'une forte concentration en DOS, nous avons cherché à restaurer chez la souche de *S.fradiae* sauvage la production de néomycine en ajoutant dans le milieu un métabolite plus avancé que le DOS dans la voie de biosynthèse afin de dépasser

[7] NB : les concentrations en paromomycine ne sont que des extrapolations à partir de la courbe de calibration de la néomycine en HPLC-MS. Pour être rigoureux, il faudrait refaire une calibration de l'aire des pics HPLC-MS avec différentes solutions de paromomycine dans le TSB purifiées par SPE. On notera au passage qu'il en est de même avec toute molécule non naturelle éventuellement formée par mutasynthèse : la HPLC-MS ne permet pas d'avoir accès à la concentration absolue sans calibration.

l'éventuelle étape limitante. L'ajout de 2 mM de ribostamycine ne modifie pas significativement le niveau de production de la néomycine. Par contre l'ajout de 2mM de paromamine conduit à une restauration de la production d'aminoglycoside mais l'aminoglycoside produit est un analogue de la néomycine où l'amine en position 6' est remplacée par un alcool ($t_R$ =9.65 contre 10.1 pour la néomycine, m/z = 616 contre 615 pour la néomycine) (voir figure III-31). Ce produit semble donc être issu de la transformation directe de la paromamine en paromomycine, qui résulte d'un contournement de l'étape d'amination réductrice qui conduit normalement de la paromamine à la néamine.

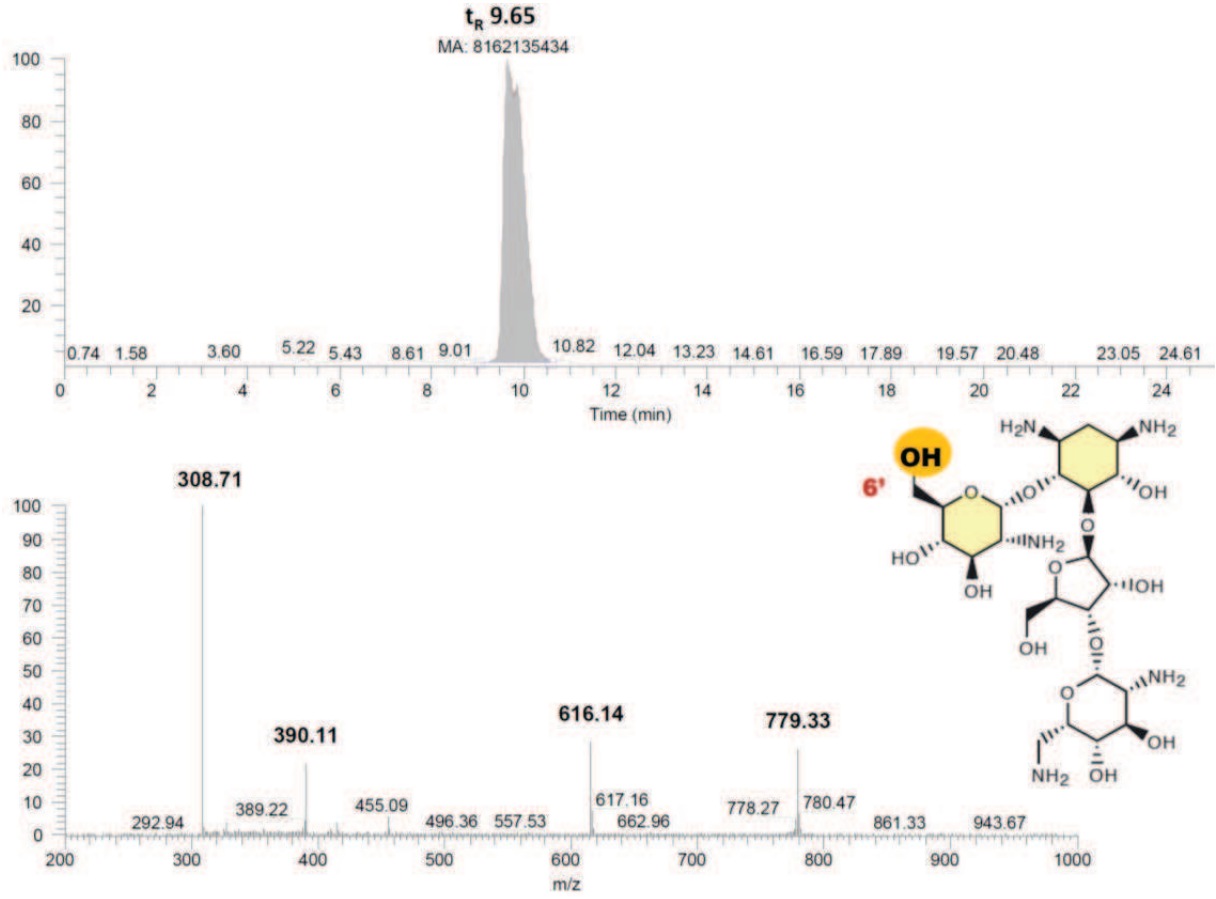



**Figure III-101 Spectre HPLC-MS et structure chimique de la paromomycine.**
La paromomycine est un analogue de néomycine hydroxylé en position 6' produit par complémentation de *S.fradiae* par la paromamine (2mM) en présence de 40 mM de DOS

Au terme de cette étude, et même si d'autres expériences de vérification seraient sans doute nécessaires, l'hypothèse de la compétition entre DOS et néomycine sur les colonnes SPE est écartée et il semble bien que l'on soit en présence d'un mécanisme de rétrocontrôle négatif de la production

de néomycine par la DOS. Il serait extrêmement intéressant de comprendre ce mécanisme de régulation qui rejoint le domaine de recherche actuellement en plein développement de la régulation de la production des métabolites secondaires chez les *Streptomyces*[331]. D'un point de vue pratique, cela signifie qu'il faut sans doute trouver un compromis pour chacun des mutasynthons testés entre une concentration suffisante pour conduire à une prise en charge efficace par la voie de biosynthèse mais n'induisant pas d'inhibition de cette voie de biosynthèse. Ce compromis est probablement différent pour les deux mutants DSM 41550 et Δneo6, ce qui complique encore l'optimisation du niveau de production des molécules éventuellement obtenues par mutasynthèse.

Un autre enseignement de ces expériences est que la voie de biosynthèse de la néomycine semble être très flexible à l'instar de ce que les travaux de Sonhg ont montré avec la voie de biosynthèse de la kanamycine[311]. En particulier, on peut envisager d'utiliser des mutasynthons modifiés en position 6' pour obtenir des analogues de néomycine eux-mêmes modifiés en position 6'. Dans ce but, nous avons développé deux analogues de paromamine, **paro$N_3$** et **paroF**, dont la synthèse est décrite précédemment et qui feront l'objet de la partie III-II-E-2 b.

La différence de comportement entre les deux souches mutantes, DSM 41550 et Δneo6, est assez complexe à analyser en raison de la présence potentielle d'un grand nombre de mutations inconnues chez DSM 41550 (souche non séquencée). On peut par exemple envisager qu'une mutation dans un transporteur membranaire facilite l'absorption cytoplasmique de la DOS exogène chez DSM 41550. Pour tester ce paramètre nous avons conduit une expérience de complémentation à la DOS sur un lysat cellulaire de la souche Δneo6 (lyse par sonication) (voir figure III-29). Cette expérience n'a pas restauré la production de néomycine possiblement du fait de l'altération de l'intégrité de la voie de biosynthèse. Il faudrait donc envisager d'utiliser des lyses plus douces telles la lyse chimique ou la lyse par la presse de French.

Il existe par ailleurs une différence essentielle entre les deux mutants qui pourrait être à l'origine de leur différence de comportement. La souche DSM41550 a été obtenue par mutations ponctuelles aléatoires provoquées par un milieu mutagène[251]. Dans la souche Δneo6, nous avons uniquement supprimé la quasi-totalité du gène neo6. Si, comme le suggèrent nos essais de surexpression de différentes enzymes clés de la biosynthèse (chapitre II), la néomycine est produite par un complexe biosynthétique comprenant plusieurs partenaires enzymatiques, la délétion de neo6 peut déstabiliser ce complexe et conduire à une perte d'efficacité globale de la voie de biosynthèse. Cette délétion pourrait ainsi influencer des étapes enzymatiques ultérieures où la protéine Neo6 n'a pas de rôle direct. A l'inverse, dans la souche DSM 41550, aucun des partenaires protéiques potentiels n'a

été complètement supprimé et donc, à moins d'avoir des mutations ponctuelles empêchant l'interaction entre deux partenaires, l'éventuel complexe biosynthétique ne serait pas modifié dans son ensemble. Des observations récentes effectuées dans le groupe de Spencer à Cambridge viennent étayer cette hypothèse : en créant un mutant de *S.fradiae* où le gène responsable de la glycosylation du DOS pour conduire à la paromamine (neo8) avait été supprimé, ils se sont aperçus, à travers des expériences de complémentation chimique, que l'efficacité et la sélectivité de l'ensemble des étapes biosynthétiques ultérieures semblaient modifiées[275].

Pour mieux comprendre la différence d'efficacité dans les voies de biosynthèse modifiées des souches Δneo6 et DSM 41550, nous avons un temps songé à réintroduire chez le mutant Δneo6 le gène neo6 modifié de façon à rendre l'enzyme Neo6 inactive. L'analyse de Neo6 par alignement de séquence (logiciel Clustal$N_2$) avec les enzymes de voies de biosynthèse similaires comme BtrR (butirosine), TbmB (tobramycine), KanB (kanamycine), GntA (gentamicine) permet en effet de détecter un résidu extrêmement conservé, la lysine catalytique K202, responsable de la liaison avec le cofacteur PLP (pyridoxal phosphate). Muter ce résidu permettrait d'obtenir une enzyme Neo6 inactivée sans toutefois modifier ses éventuelles interactions avec des partenaires protéiques. Nous avons cependant renoncé à ce projet par manque de temps.

En conclusion, on peut restaurer la production de néomycine par complémentation chimique avec la DOS dans les deux souches de *S.fradiae* mutées DSM 41550 et Δneo6 mais ces deux souches n'ont pas le même comportement. La suite de l'analyse était d'utiliser ces souches dans une stratégie de mutasynthèse avec d'autres métabolites naturels ou avec des mutasynthons éxogènes et d'analyser l'influence de leur différence de comportement avec ces nouveaux substrats.

### c) Complémentation avec la néamine et la ribostamycine

Lorsque les deux souches DSM41550 et Δneo6 sont complémentées avec 2 mM de néamine ou de ribostamycine, on restaure la production de néomycine (voir figure III-32). Ces résultats sont en contradiction avec les observations que Rinehart avait faites sur le mutant DSM 41550[250] mais vont dans le sens des résultats de l'équipe de Sato au cours de leur étude sur des mutants DOS- de *S.ribosidificus* and *S.kanamyceticus*[305]. Nos observations prouvent en effet qu'il est possible d'utiliser des mutasynthons pseudo-disaccharidiques de type néamine.

Une différence de concentration finale en néomycine est observée entre les deux souches mutées, DSM 41550 étant systématiquement celle qui produit le plus de néomycine comme c'était le cas lors de la complémentation avec la DOS.

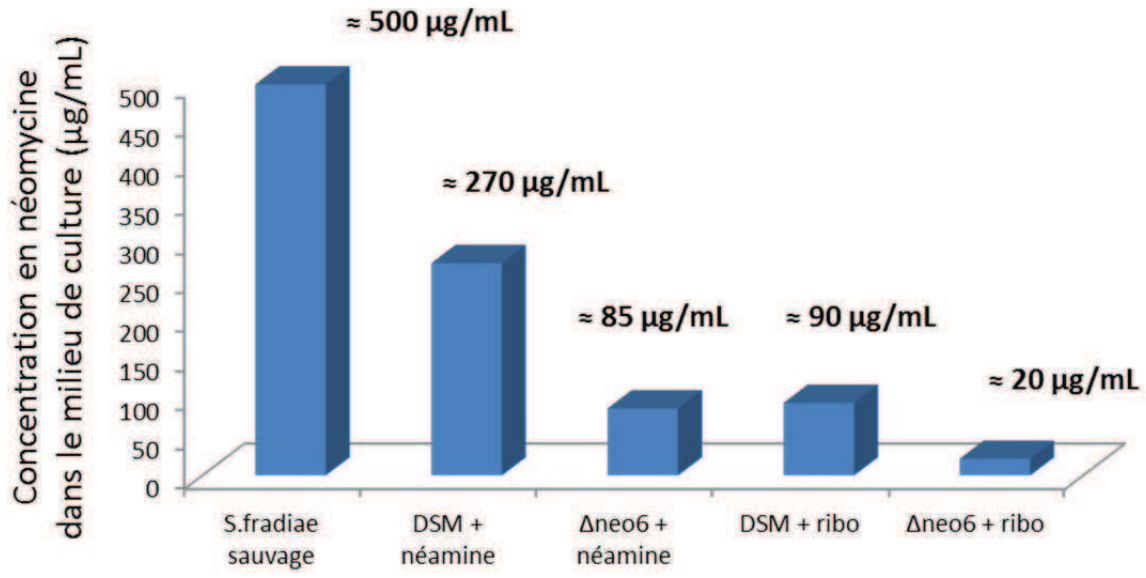



**Figure III-102 Complémentation de DSM 41550 et Δneo6 par la néamine et la ribostamycine**

### d) Complémentation avec la paromamine

Comme ce qui avait été observé pour la souche sauvage (qui produit naturellement dans nos conditions exclusivement de la néomycine), la présence de paromamine conduit à la production de paromomycine (voir figure III-33). Là encore les deux souches n'ont pas le même comportement : la souche DSM 41550 produit un mélange de néomycine (environ 140 µg/mL) et de paromomycine (environ 35 µg/mL) tandis que la souche Δneo6 produit exclusivement de la paromomycine (environ 60 µg/mL) (voir figure III-34). De façon intéressante, la souche Δneo6 conduit à une production plus importante de paromomycine que la souche DSM 41550, ce qui laisse possiblement entrevoir un intérêt en mutasynthèse.

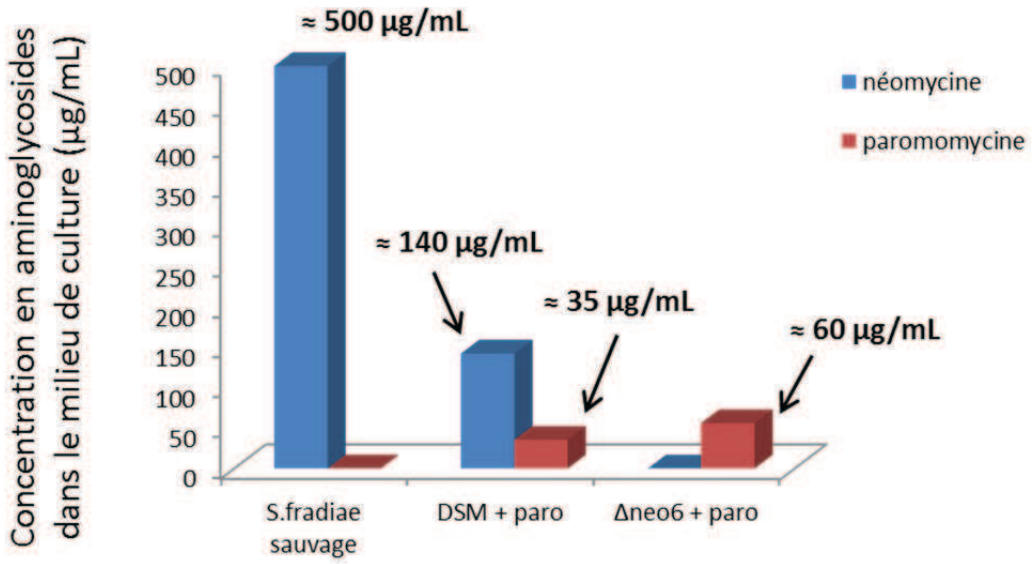



**Figure III-103 Complémentation de DSM 41550 et Δneo6 par la paromamine**

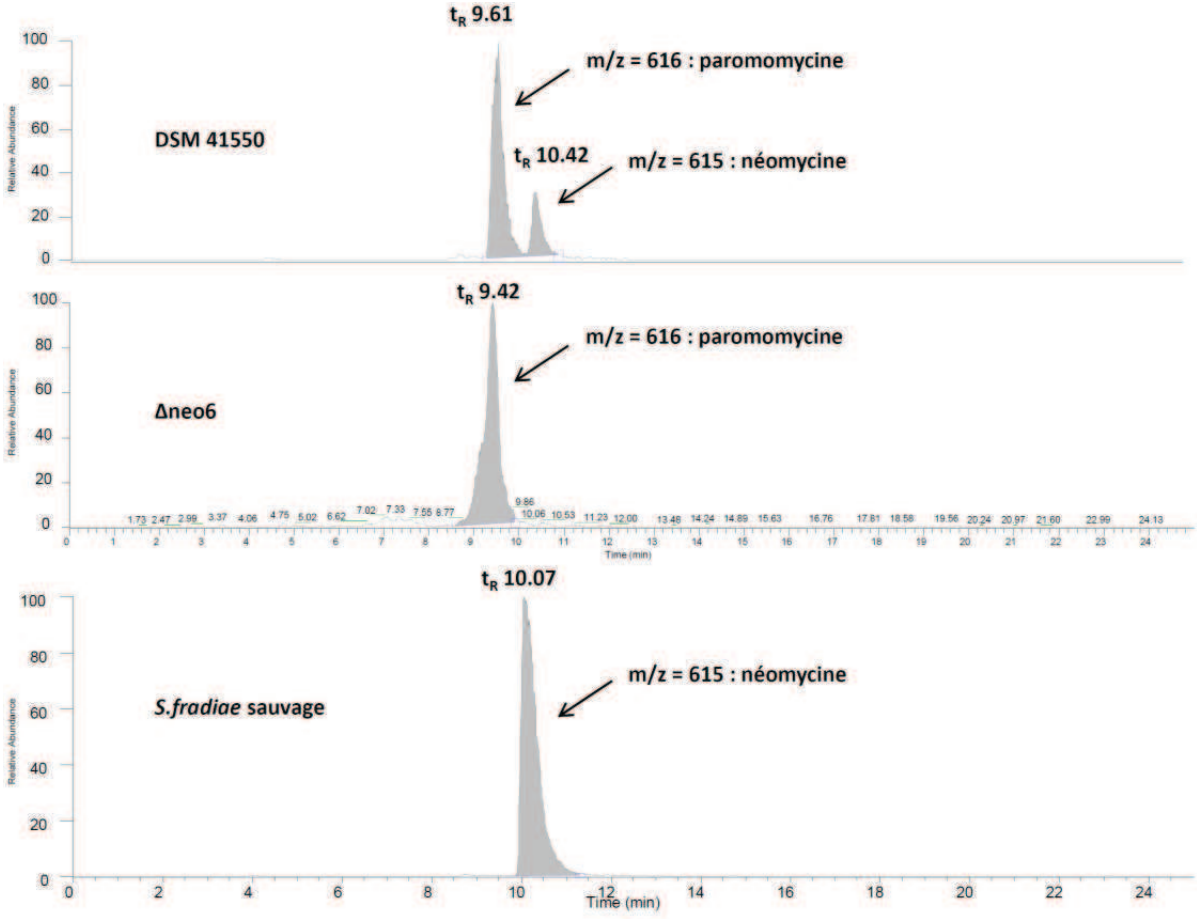



**Figure III-104 Spectres HPLC-MS des expériences de complémentation de DSM 41550 et Δneo6 avec la paromamine**

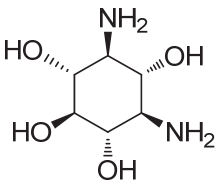
### 2) Complémentation avec des mutasynthons non naturels

Encouragés par les résultats des expériences de complémentation chimique avec des métabolites naturels de la voie de biosynthèse, nous nous sommes tournés vers des mutasynthons exogènes pouvant servir de mimes de ces métabolites. Il n'existe malheureusement que très peu de molécules commerciales avec les propriétés désirées. Nous nous sommes donc concentrés tout d'abord sur une librairie restreinte d'analogues potentiellement intéressants. La synthèse de ces analogues a été décrite précédemment (partie III-II-D) à l'exception du composé **90**, qui nous a été fourni par Aurélie Blond, doctorante dans notre équipe de recherche. Enfin, nous avons également utilisé comme mutasynthons deux antibiotiques disponibles commercialement, la kanamycine **6** et la gentamicine **45**. Afin de détecter les produits éventuellement formés, nous avons mis au point un algorithme de criblage à partir des spectres HPLC-MS, repérant les ions de masse identique aux intermédiaires biosynthétiques attendus pour chaque mutasynthon.

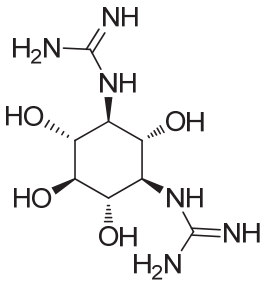
#### a) Complémentation avec des analogues de DOS

Nous nous sommes tout d'abord tournés vers deux mutasynthons qui avaient été testés dans les premières expériences de Rinehart[250], la streptamine **47** et la streptidine **76** (voir figure III-35). En complémentant la souche DSM 41550 avec de la streptamine, Rinehart a décrit l'obtention d'« hybrimicine », un analogue de néomycine résultant du remplacement du fragment DOS par la streptamine. Comme nous l'avons vu dans la partie introductive, ce composé a représenté le premier produit non naturel obtenu par mutasynthèse dont on a une caractérisation fiable (MS, RMN $^{13}C$, …). En ce qui concerne la streptidine, si dans ces premières études sur le mutant DSM 41550 elle ne conduit pas à la formation de nouveau produit par mutasynthèse, certaines études ultérieures sur d'autres mutants démontrent le contraire.

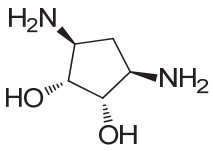
De façon surprenante, nous n'avons jamais réussi à reproduire ces résultats malgré nos essais répétés. Dans nos mains, ces deux molécules ne sont transformées ni par la souche DSM 41550, ni pour la souche Δneo6. Une explication possible est que nous ne sommes pas exactement dans les conditions de mutasynthèse des expériences précédemment décrites. En particulier, il n'a pas été possible de retrouver la composition exacte des milieux de culture utilisés ni le protocole précis d'introduction des mutasynthons. Ces résultats restent néanmoins surprenants.



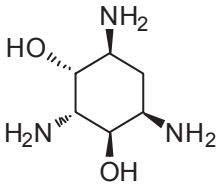

**streptamine 47**




**streptidine 49**




*cis*-1,3-diaminocyclopent-4,5-diol (**C5**)




*cis*-1,3-*trans*-5-triaminocyclohex-4,6-diol (**90**)

**Figure III-105 Les analogues de DOS testés comme mutasynthons**

Nous ensuite utilisé la molécule **C5** (*cis*-1,3-diaminocyclopent-4,5-diol) qui possède en commun avec la DOS le motif diamine en cis. Comme nous l'avons évoqué précédemment, des travaux antérieurs du laboratoire avaient démontré que cette molécule pouvait constituer un mime convenable de la DOS vis-à-vis de partenaires nucléiques[62,325] ou protéiques[130]. Nous avons effectué diverses expériences de mutasynthèse sur les deux souches, DSM 41550 et Δneo6, en faisant varier la concentration du mutasynthon (2.5, 5, 10 et 25 mM) et son protocole d'introduction dans le milieu de culture. Ces expériences ont donné des résultats identiques pour les deux souches de façon très reproductible et ne nous ont pas permis d'identifier la production d'un analogue de néomycine. Par contre, de façon intéressante, nous avons détecté par HPLC-MS la présence d'un métabolite secondaire potentiel (m/z = 293) qui pourrait être un analogue de néamine résultant de l'incorporation du mutasynthon **C5** (composé **91**, voir figure III-36).

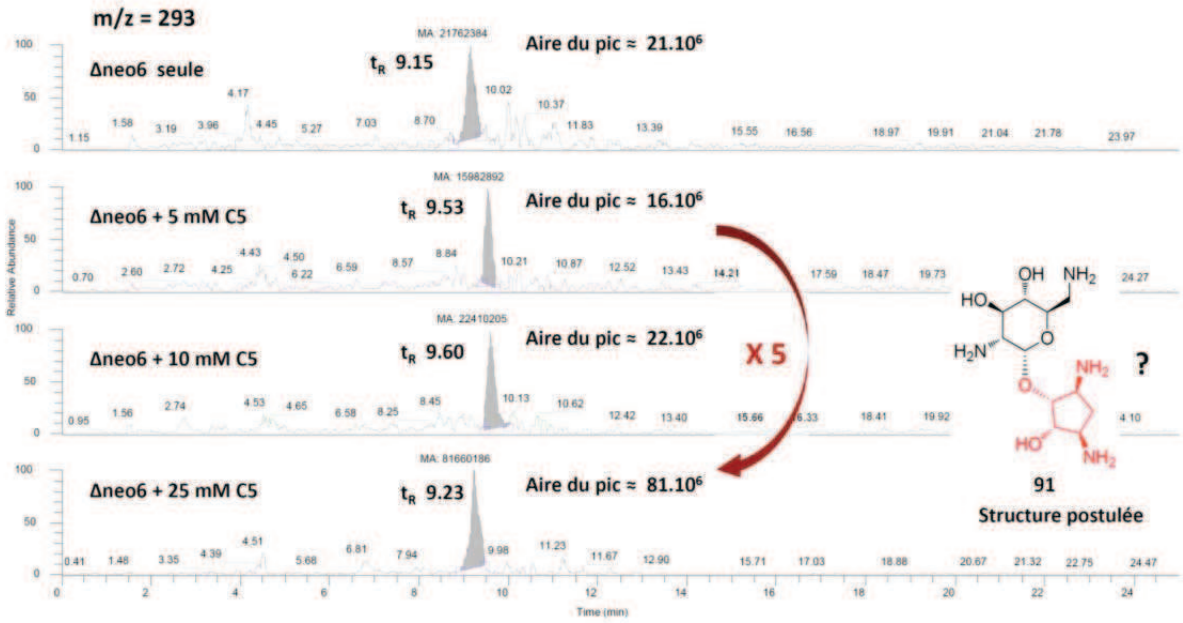



**Figure III-106 Complémentation de Δneo6 avec le mutasynthon C5**
(chromatogramme reconstitué à m/z =293)

Comme on peut le voir sur la figure III-36, l'identification de ce composé reste sujette à caution. En effet, dans l'expérience de contrôle négatif que nous avons effectuée (Δneo6 seule), on trouve aussi un pic pour un composé à m/z = 293. Cependant, la fragmentation de ce pic n'est pas du tout la même dans le contrôle négatif, où on a que du bruit de fond, et dans les expériences de complémentation avec **C5**, où on observe un pic bien défini correspondant à une déshydratation (m/z – 18) (voir figure III-37). Cette fragmentation est donc compatible avec la structure postulée de la molécule mais ne permet malheureusement pas de retrouver un fragment qui serait caractéristique du mutasynthon C5 (m/z = 133, voir partie expérimentale). Ce pic peut néanmoins être fragmenté de façon successive pour donner un produit à m/z = 245 puis un produit à m/z = 219, que nous n'avons pas pu attribuer.

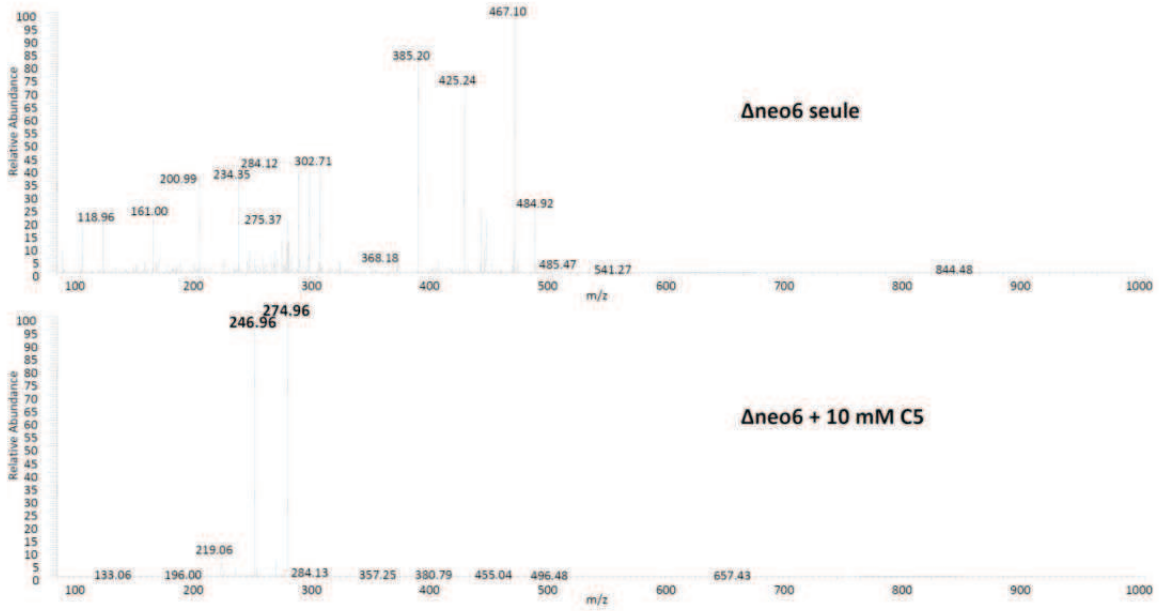



**Figure III-107 Fragmentation de l'ion à m/z = 293 pour Δneo6 seule et Δneo6 + 10 mM C5.**

Lors de l'analyse de ces expériences, trois détails intrigants ont attiré notre attention. Tout d'abord, on peut noter que le temps de rétention de ce produit ($t_R$ = 9.53) est compatible avec celui de son cousin chimique la néamine ($t_R$ = 9.25, voir partie expérimentale). Ensuite, l'aire du pic à m/z = 293 observé lors des expériences de mutasynthèse semble être directement liée à la quantité de mutasynthon introduite dans le milieu : une multiplication par 5 de la concentration du mutasynthon conduit à une multiplication par 5 de l'aire du pic à m/z = 293 en HPLC-MS (voir figure III-36). Il y a donc une réponse dose-dépendante. Enfin, l'apparition progressive de ce pic à m/z = 293 s'accompagne de l'apparition d'un autre pic moins marqué à $t_R \approx 4.7$ et correspondant à un produit de m/z = 294. De façon très intéressante, ce pic semble être absent du témoin négatif et pourrait correspondre à un analogue de paromamine ($t_R$ = 6.62) et donc à un précurseur biosynthétique de l'analogue de néamine. C'est un argument de plus en faveur de l'existence d'analogues non naturels d'aminoglycoside résultant de la transformation de la molécule **C5** par les souches mutées DSM 41550 et Δneo6. Il faut cependant rester très prudent car on se situe aux limites de détection de notre appareil de HPLC-MS.

Dans le but d'éclaircir cette question, nous avons cherché à effectuer des expériences de mutasynthèse à plus grande échelle de façon à être au dessus de la limite de détection des éventuels produits formés. On peut trouver dans la partie expérimentale les conditions de culture que nous avons développées et qui permettent d'effectuer des expériences de mutasynthèse dans un

fermenteur de 4L. Malheureusement, nos travaux ont été arrêtés par des problèmes techniques sur notre appareil HPLC (pannes répétées).

Par ailleurs, comme malheureusement les aminoglycosides ne sont pas visibles en UV, nous nous sommes également penchés sur des méthodes nous permettant de dérivatiser ces produits directement dans le milieu de culture, de façon à pouvoir les détecter en UV et à les purifier par HPLC semi-préparative. Pour ce faire, nous nous sommes inspirés de méthodes de dérivatisation des aminoglycosides par des groupes protecteurs Fmoc[332] (fluorenylmethyloxycarbonyl) ou DNB[272,280] (2,4-dinitrophényl) directement à partir de milieux complexes. L'inconvénient de ces méthodes est que les produits obtenus sont très peu solubles ce qui complique leur purification. Le développement de ces méthodes est actuellement en cours au laboratoire.

Les expériences de mutasynthèse avec les autres mutasynthons ont, elles aussi, souffert des problèmes techniques de notre appareil HPLC et ne sont en conséquence que des études préliminaires. Nous avons cependant pu effectuer une expérience préliminaire de mutasynthèse avec le mutasynthon **90** qui semble donner des résultats encourageants : on peut voir, à la limite du seuil de détection, l'apparition d'un produit absent du contrôle négatif, pouvant correspondre à un analogue de néomycine résultant de l'incorporation du mutasynthon (figure III-38). Il serait intéressant d'effectuer d'autres expériences pour confirmer cette observation et éventuellement isoler le produit formé.

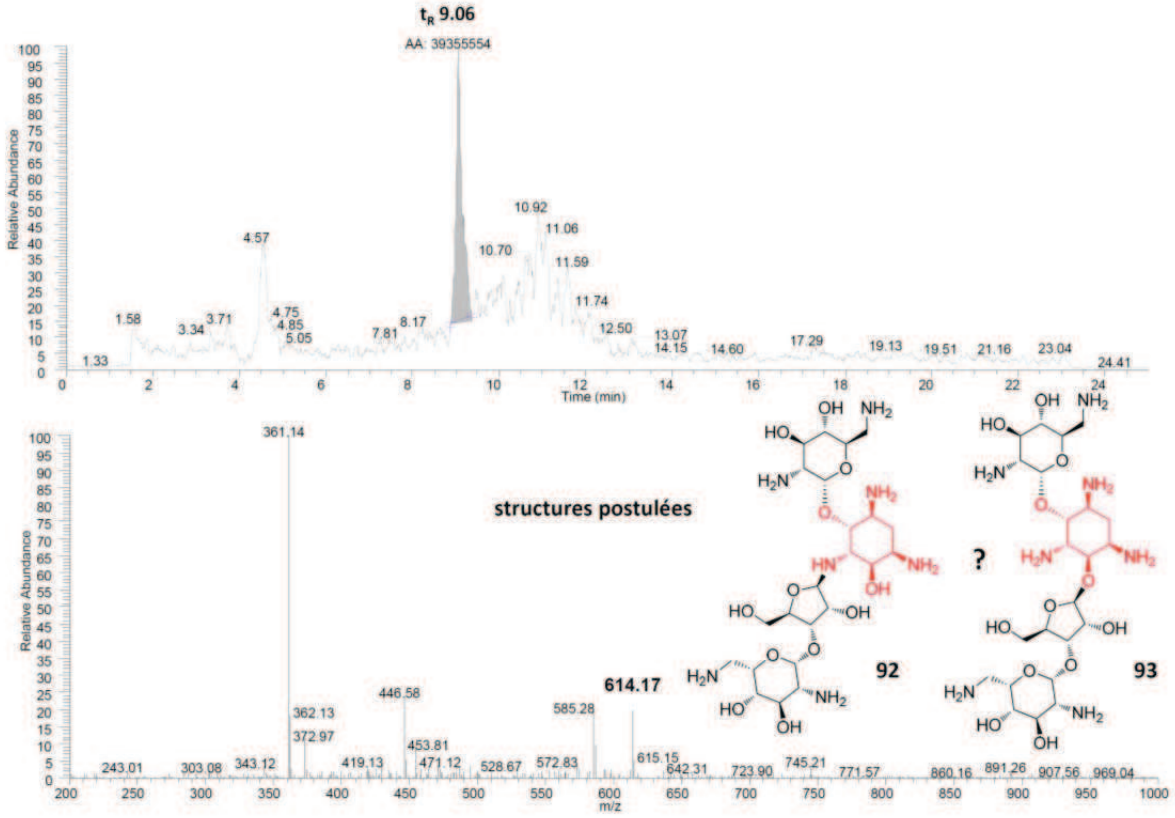
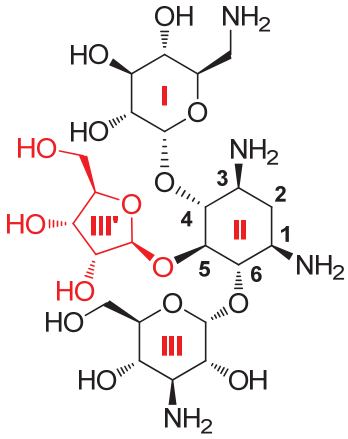



**Figure III-108 Complémentation de Δneo6 avec le mutasynthon 90**
(spectre de masse extrait à $t_R$ = 9.06 min)

### b) Complémentation avec des analogues de métabolites plus complexes

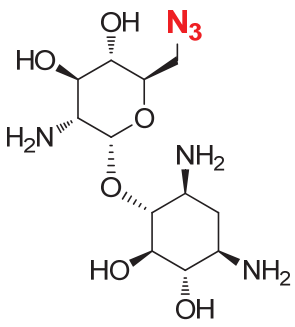
L'équipe d'Asensio a récemment décrit des hybrides ribostamycine/kanamycine très encombrés et peu susceptibles d'être modifiés par des enzymes de résistance[222]. Dans ces analogues d'aminoglycosides, les trois alcools de la DOS sont substitués par des sucres (voir figure III-39).



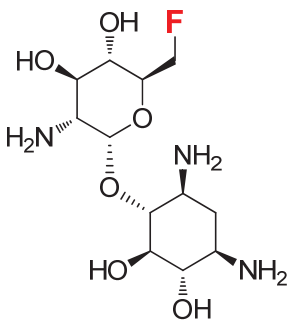

**24**

**Figure III-109 Exemple d'hybride kanamycine/ribostamycine développé par l'équipe d'Asensio**[222]

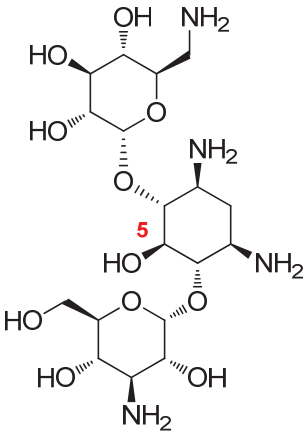
Dans le but d'obtenir de telles molécules par mutasynthèse, nous avons cherché à utiliser la kanamycine et la gentamicine comme mutasynthons (voir figure III-40). Dans ces molécules en effet, l'alcool en position 5 de la DOS est libre et peut donc potentiellement être modifié par les enzymes de la voie de biosynthèse de la néomycine, qui est fonctionalisée en position 4 et 5 de la DOS. Malheureusement, les concentrations que nous avons utilisées (2mM) se sont avérées létales pour les souches mutantes DSM 41550 et Δneo6. Il serait intéressant de recommencer cette expérience en diminuant la concentration de ces mutasynthons ou en rendant les souches résistantes à ces antibiotiques (transformation par un plasmide codant pour une enzyme de résistance).



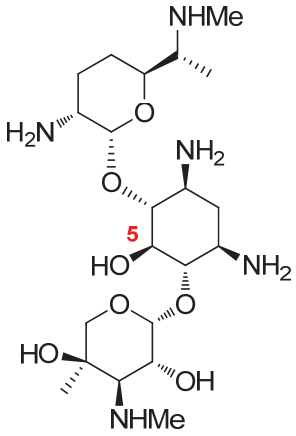
**paro$N_3$ 82**



**paroF 86**



**kanamycine 6**



**gentamicine 45**

**Figure III-110 Analogues de paromamine synthétisés pouvant servir de mutasynthons**

Les expériences de mutasynthèse sur les composés paroF et paro$N_3$ n'ont malheureusement pas pu être analysées en raison de problèmes techniques sur l'appareil de HPLC-MS et du manque de temps pour trouver une solution alternative.

# III - Conclusion et perspectives

Au terme de ce chapitre, on peut tirer le bilan des progrès accomplis dans la mise en place d'une stratégie de mutasynthèse de nouveaux analogues d'aminoglycosides. La conception du mutant Δneo6 et d'une méthodologie d'analyse des aminoglycosides par HPLC-MS ont constitué des avancées significatives vers ce but. L'approche de la mutasynthèse a pu être validée par complémentation des souches de *S.fradiae* mutantes Δneo6 et DSM 41550 par des intermédiaires de biosynthèse, ce qui a permis de restaurer la production de néomycine. Au cours de ces études, nous avons mis à jour des mécanismes intéressants de régulation et de divergence de la voie de biosynthèse (production de néomycine et/ou de paromomycine en fonction des conditions). Nous avons par ailleurs développé une chimie permettant la synthèse d'analogues de métabolites naturels pouvant servir de mutasynthons. Certains des mutasynthons ont pu être testés même si nous avons malheureusement été assez vite limités par la sensibilité de notre méthode de détection.

L'insuffisante sensibilité de notre méthode de détection, associé à l'insuffisante purification des différents analytes par SPE, se sont avérés être des problèmes plus importants qu'initialement prévus puisque les pistes intéressantes que nous avons pu découvrir n'ont pas pu être confirmées de façon certaine, se situant à la limite de détection. A ce problème de sensibilité se sont ajoutés de nombreux problèmes techniques liés à la vétusté de notre appareillage qui nous ont empêchés de mener à terme l'analyse des différentes expériences de mutasynthèse que nous avions effectuées.

Si on se réfère à des travaux de mutasynthèse similaires au nôtre, comme par exemple les travaux de l'équipe de Kirschning sur les analogues d'ansamitocine[264] ou de geldanamycine[263,321], on constate que les composés formés par mutasynthèse sont produits à des concentrations très faibles (de l'ordre du µg/mL voire moins, c'est-à-dire plus de 100 fois moins que la concentration du produit naturel obtenu par fermentation de la souche sauvage, respectivement *Streptomyces hygroscopicus* et *Actinosynnema pretiosum*). Il est donc clair que ces études nécessitent des méthodes très pointues de détection et d'analyse. La plupart d'entre elles utilisent d'ailleurs une analyse par UPLC-HRMS (pour « Ultra Performance Liquid Chromatography » couplée à un spectromètre haute résolution). Cette détection est très sensible et permet de plus d'identifier relativement simplement les éventuels produits formés. L'accès à une telle technologie permettrait sans doute d'accélérer considérablement nos travaux.

Une autre des difficultés à laquelle nous avons été confrontés était l'analyse des spectres HPLC-MS des milieux de culture purifiés. Malgré la stratégie de criblage mise en place, il nous a bien

souvent été très difficile d'analyser les changements de ces spectres dans les différentes conditions de culture : l'analyse par l'œil humain (plein d'espoir) apporte potentiellement un biais non négligeable. De plus, dans notre méthode actuelle on ne trouve que ce que l'on cherche : si le mutasynthon utilisé conduit à un produit inattendu, il peut très bien passer inaperçu s'il n'est pas produit à des concentrations qui rendent visible son pic en HPLC-MS. Par ailleurs, sauf cas très favorable où il y a production massive de métabolites secondaires (comme par exemple quand on complémente DSM 41550 avec son métabolite naturel, la DOS), on ne sait pas *a priori* si les changements observés dans les spectres HPLC-MS sont statistiquement significatifs. Pour dépasser cette difficulté, on pourrait envisager d'utiliser des outils d'analyse statistique systématique des produits trouvés dans le milieu de culture. On peut notamment songer à adapter les techniques d'analyse en composantes principales (PCA en anglais), très utilisées en économie ou en informatique, à l'analyse des variations de la composition moléculaire d'un milieu de culture. On pourrait ainsi détecter de façon semi-automatique les éventuels produits formés par mutasynthèse. Cette méthode d'analyse a été employée avec succès par l'équipe de Rolf Müller en Allemagne pour l'analyse du métabolome de diverses souches bactériennes. On peut citer par exemple ses travaux novateurs d'analyse systématique du métabolome de *Myxococcus xanthus* ayant permis l'identification de la Myxoprincomide[333].

En conclusion, ce projet de recherche, initié au début de ma thèse, offre des perspectives très intéressantes. Les études que nous avons menées ont permis de soulever des questions intéressantes à la fois sur la régulation de la production de néomycine chez *S.fradiae* mais aussi sur la flexibilité de sa voie de biosynthèse. Par ailleurs, certaines pistes obtenues lors de la complémentation des souches DSM 41550 et Δneo6 avec les divers mutasynthons que nous avons synthétisés sont encourageantes et justifient la mise en place d'autres tests sur une plus grande librairie de mutasynthons. Enfin, les derniers développements méthodologiques de l'étude de la production de métabolites secondaires chez les *Streptomyces* devraient permettre de lever les difficultés rencontrées dans l'analyse des métabolites formés.

# IV. Developpement de sondes fluorées pour l'étude des ARN en RMN $^{19}F$

Dans la partie introductive, nous avons pu voir le rôle croissant accordé aux ARN dans de nombreux phénomènes biologiques clés. En dépit de cette importance, leur potentiel thérapeutique reste relativement inexploité, en particulier en comparaison avec les cibles protéiques ou ADN. Ceci est lié à notre compréhension parcellaire de la structure et de la dynamique des ARN ainsi que de leur interaction avec d'éventuels partenaires moléculaires. La conception de ligands affins et spécifiques de ces biomolécules polyanioniques et très flexibles est en conséquence rendue assez complexe.

Nous avons cherché à développer un outil supplémentaire pour l'étude des ARN qui puisse donner accès à leur caractère dynamique. En tirant parti de nos études sur les analogues d'aminoglycosides, nous avons développé des sondes fluorées permettant d'étudier les ARN en spectroscopie RMN $^{19}F$.

Après avoir rapidement rappelé l'intérêt de la RMN, et plus particulièrement de la RMN du fluor, dans l'étude des ARN, nous résumerons les travaux préliminaires à l'origine de notre projet. Ceci nous amènera à nous pencher plus spécifiquement sur les sondes fluorées développées au laboratoire. Nous étudierons leur emploi dans la caractérisation de la structure et de la dynamique de changement conformationnel de certains ARN avant de nous intéresser à leur utilisation dans la caractérisation de l'affinité de certains ligands vis-à-vis de leur cible nucléique.

# I - La RMN du fluor dans l'étude des ARN

## *A - L'intérêt de la RMN : donner une information dynamique à l'échelle atomique*

Comme nous l'avons vu dans la partie introductive, les ARN sont des biomolécules très flexibles capables de former des structures tridimensionnelles complexes. Ils sont bien souvent en équilibre entre plusieurs conformations et leur activité biologique est dépendante de leur capacité à bouleverser cet équilibre suivant leur état fonctionnel[81]. Cet aspect dynamique, que l'on retrouve par exemple dans les phénomènes d'adaptation conformationnelle d'un ARN lors de la fixation de son ligand, est crucial pour comprendre le fonctionnement biologique des ARN et pour concevoir de nouveaux ligands capables d'interférer spécifiquement avec ceux-ci.

Parmi les méthodes d'analyse disponibles, la résonance magnétique nucléaire (RMN) joue un rôle à part en raison de sa capacité à fournir une information dynamique à l'échelle atomique[334]. A travers un simple jeu de paramètres (le déplacement chimique (δ), les couplages entre spin, les temps de relaxation, …) extrêmement sensibles à l'environnement électronique et chimique des atomes étudiés, la RMN permet d'avoir accès à quantité d'informations sur la structure des biomolécules d'intérêt, les caractéristiques de leurs dynamiques moléculaires ou leur mode d'interaction avec différents partenaires.

A condition de s'intéresser à des objets de taille raisonnable (moins de 150 nucléotides environ)[335], la RMN permet ainsi d'étudier à la fois la structure des ARN mais aussi leurs changements conformationnels[336]. On trouve dans la littérature de nombreuses structures complexes d'ARN résolues par RMN et confirmées par d'autres techniques d'analyse telles que la cristallographie. Les propriétés intrinsèques de la RMN pouvant permettre des mesures rapides, elle est de plus très utilisée dans l'étude de phénomènes dynamiques comme les cinétiques de repliement ou de modification chimique des ARN (par exemple, modification post-transcriptionelle).

Par ailleurs, un des grands avantages de la RMN est de permettre la caractérisation précise de l'interaction des ARN avec d'éventuels ligands, même si ceux-ci n'ont que des affinités modestes (Kd de l'ordre du millimolaire). Pour ce faire, on peut par exemple procéder à des expériences dites d'« empreintes », qui consistent à enregistrer le spectre RMN 2D homo- ou hétéronucléaire d'un ARN en présence et en absence de son ligand. La fixation du ligand induit la modification d'un certain nombre de signaux correspondant aux nucléotides qui interagissent avec lui, permettant de visualiser l'empreinte de fixation du ligand sur l'ARN. Comme nous le verrons ultérieurement, cette technique a été utilisée au laboratoire pour déterminer les sites de fixation d'analogues

d'aminoglycosides sur de l'ARNt$^{Lys}_3$ enrichi en $^{15}$N à l'aide d'expériences de RMN hétéronucléaires $^{1}$H/$^{15}$N de type TROSY[62]. Nous avons également utilisé cette approche pour étudier les sites de fixations de nos sondes fluorées sur un analogue du site A du ribosome, l'ARN 16S23, à l'aide d'expériences de RMN 2D homonucléaire $^{1}$H/$^{1}$H (TOCSY sur les protons H5-H6 des nucléobases et NOESY). Pour étudier l'interaction d'un ligand avec une cible ARN bien caractérisée, il est également possible de se contenter d'expériences de RMN $^{1}$H 1D en observant les variations de déplacement chimiques des protons imino de l'ARN (entre 10 ppm et 15 ppm) lors de l'ajout du ligand (voir figure IV-1).

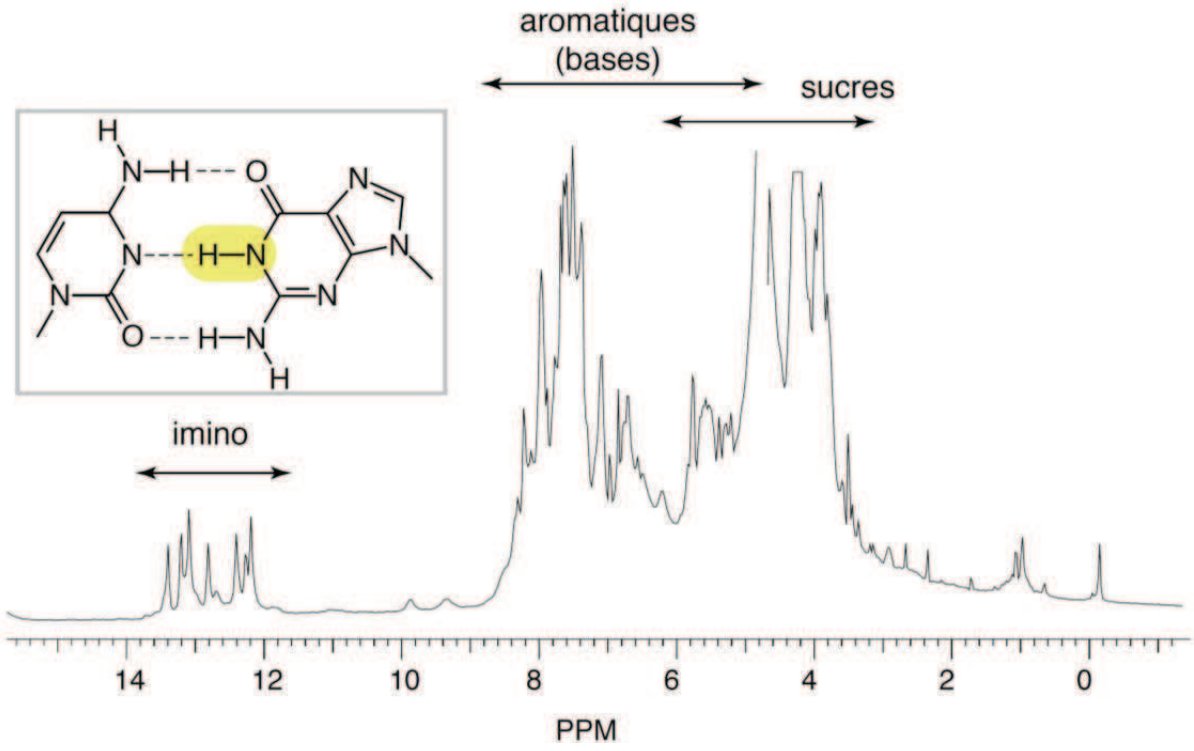



**Figure IV-111 Exemple de spectre RMN $^{1}$H d'un ARN**
Les protons imino correspondent aux NH des bases G ou U appariées.

Nous avons utilisé cette approche pour étudier l'interaction d'un certain nombre d'analogues d'aminoglycosides avec l'ARN 16S23. Enfin, il est également possible de mettre en place une stratégie d'observation du ligand lors de l'interaction avec sa cible (expériences de type STD (Saturation Transfert Difference) ou de WaterLOGSY par exemple), ce que j'ai eu l'opportunité de faire pour identifier des analogues d'aminoglycosides lors de mon Master 2[130].

En conclusion, en dépit de sa sensibilité modeste en comparaison avec d'autres techniques d'analyse, la richesse des informations fournies par la RMN en fait une méthode de choix dans

l'étude structurale dynamique de l'activité biologique des ARN et dans la conception de nouveaux ligands. Les expériences développées dans la littérature ont toutefois comme inconvénient de souvent nécessiter de grandes quantités de biomolécules marquées (par exemple ARN enrichis en $^{15}N$/ $^{13}C$), ce qui peut être cher et techniquement difficile. Par ailleurs, les spectres obtenus peuvent être très complexes même pour des ARN relativement courts (nombreux recouvrements de pics, analyse difficile des différents conformères, …) et leur analyse nécessite donc souvent une véritable expertise scientifique.

Pour dépasser en partie ces limitations, la RMN $^{19}F$ a été récemment exploitée pour l'étude des ARN. Cette méthode s'est révélée à la fois simple d'utilisation et extrêmement utile.

## *B - La RMN $^{19}F$ comme outil d'étude des biomolécules*

### 1) Les avantages de la RMN $^{19}F$

La RMN du fluor peut être exploitée de façon très efficace dans l'analyse de systèmes biologiques complexes et ceci pour 4 raisons principales[337] :

- Le fluor joue un rôle mineur en biologie et est par conséquent quasiment absent des systèmes étudiés ce qui diminue le risque de signaux parasites non introduits par l'expérimentateur[338].
- L'isotope $^{19}F$, de spin nucléaire ½ et d'abondance relative naturelle de 100 %, est très sensible dans les expériences de RMN en raison de son fort rapport gyromagnétique. On peut évaluer la sensibilité théorique du noyau fluor à 0.83 fois celle du noyau proton, soit bien plus que ce qui peut être atteint avec les autres noyaux usuels de la RMN comme l'azote $^{15}N$ ou le carbone $^{13}C$. Ces derniers ont de plus l'inconvénient d'avoir une abondance naturelle faible. En pratique, la sensibilité du noyau fluor est même encore plus grande que celle du proton en raison de l'absence de certains problèmes expérimentaux usuels de la RMN $^{1}H$ comme la déformation de la ligne de base liée à la suppression des signaux du solvant par exemple[337].
- Le fluor possède une gamme de déplacements chimiques très large allant jusqu'à plus de 80 GHz pour les composés organiques (plus de 300 ppm sur un spectromètre 282 MHz) soit presque 25 fois plus que pour le proton. Ceci diminue d'autant le risque de recouvrement des signaux et facilite l'attribution des spectres. Par ailleurs, le déplacement chimique du fluor est extrêmement sensible à son environnement chimique et électronique, ce qui en fait une bonne sonde RMN. Ainsi, une simple modification du pH ou de la concentration en sels

d'un échantillon RMN va directement se répercuter sur le déplacement chimique du fluor. De la même façon, l'interaction d'une petite molécule fluorée avec une macromolécule bouleverse complétement le spectre fluor de la molécule. En plus de modifier le déplacement chimique, cette interaction conduit souvent à un élargissement du signal du fluor dû à une diminution drastique du temps de relaxation $T_2$.

- Enfin, les mesures en RMN $^{19}F$ sont robustes et peu sujettes à des interférences. En particulier, il est possible d'utiliser une large variété de tampons, détergents ou petites molécules sans nécessairement devoir utiliser des solvants deutérés comme c'est le cas en RMN $^{1}H$ par exemple. Les conditions d'étude des biomolécules sont donc moins contraintes en RMN $^{19}F$. Elles peuvent donc plus facilement être optimisées pour la stabilité ou la bioactivité de la cible d'intérêt.

Les avantages de la RMN $^{19}F$ ont été pour le moment majoritairement exploités pour étudier la structure et la dynamique de protéines modifiées par fluoration[339]. Il est également possible d'étudier l'interaction de petites molécules fluorées avec des partenaires protéiques. On peut notamment citer les travaux de l'équipe de Claudio Dalvit qui développe depuis une dizaine d'années de nouvelles techniques de criblage en RMN $^{19}F$. En exploitant l'élargissement du signal du fluor d'une petite molécule fluorée lors de son interaction avec une macromolécule, il a été possible de développer un criblage moyen débit capable d'identifier, parmi un mélange de ligands potentiels, les molécules qui se liaient à une cible protéique particulière[340]. Une fois que l'on a identifié un ligand fluoré approprié (affinité modérée et un seul site d'interaction), il est possible de s'en servir dans un second temps comme reporter pour cribler des librairies de ligands non fluorés en effectuant des expériences de compétition[337,341,342]. Les ligands non fluorés qui interagissent spécifiquement avec la cible biologique d'intérêt vont pouvoir chasser la molécule fluorée reporter et induire un changement de son signal en RMN $^{19}F$ (le pic de la sonde (S) passe d'un signal élargi dû à l'interaction avec la protéine à un signal fin, voir figure IV-2). Cette technique, nommée FAXS (pour « fluorine chemical shift anisotropy and exchange screening ») semble être employée avec succès dans l'industrie pharmaceutique dans des stratégies de synthèse d'inhibiteurs par fragment. Une preuve de concept de cette stratégie, le criblage d'une bibliothèque de ligands de la trypsine bovine, a récemment été publiée [343].

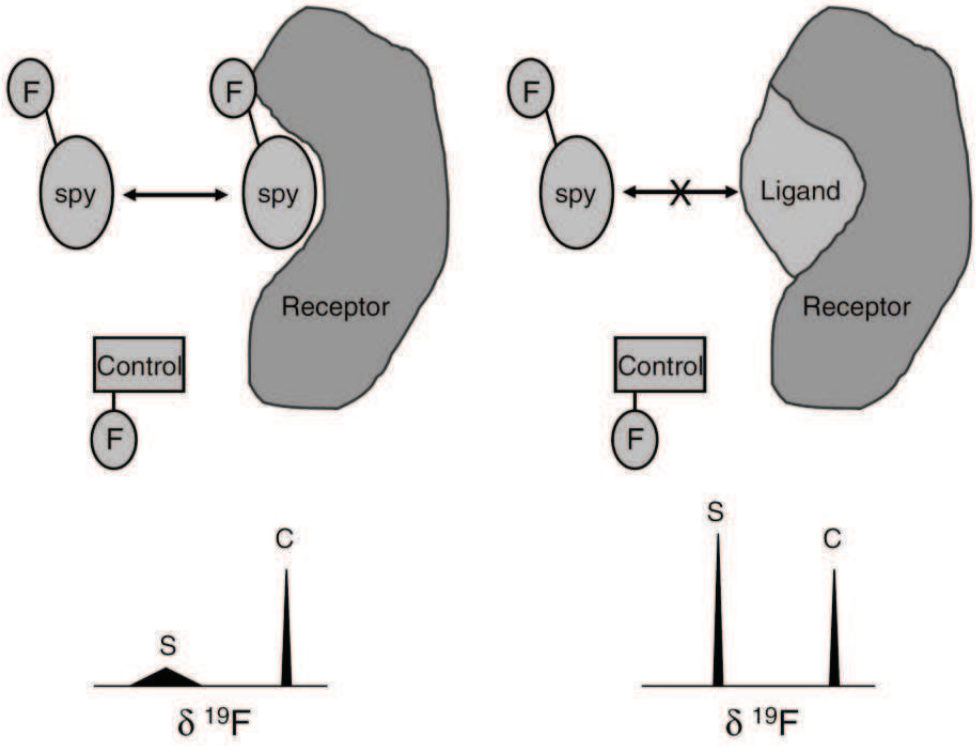
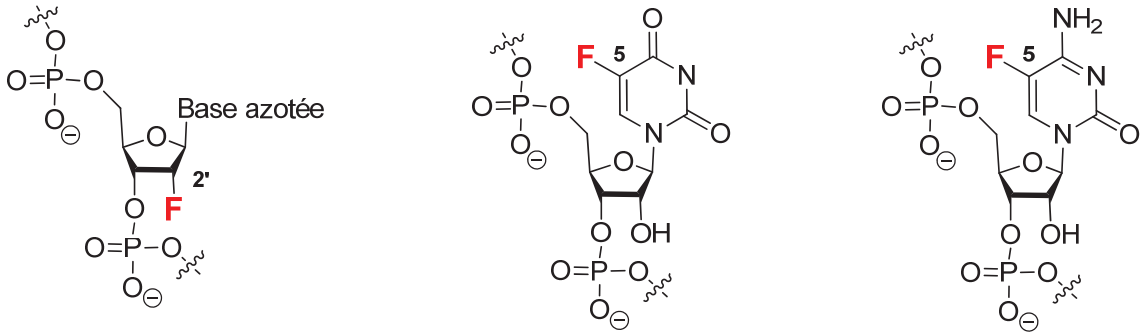



**Figure IV-112 Principe du criblage par RMN $^{19}$F en utilisant une molécule rapportrice fluorée (méthode FAXS)[337].**

### 2) Application de la RMN $^{19}$F à l'étude des ARN

#### a) Stratégie de marquage de la cible

Si la RMN du fluor a pour le moment essentiellement été appliquée à l'étude des protéines, on trouve néanmoins quelques exemples où elle a été exploitée avec succès dans l'étude des ARN. Une stratégie populaire consiste à fluorer la cible biologique avec des approches hémi-synthétiques ou chemoenzymatiques, puis à suivre les modifications du signal du fluor par RMN[344]. Ces méthodes de fluoration sont cependant souvent assez peu efficaces et manquent de spécificité ce qui limite considérablement leur portée. Une alternative consiste à synthétiser entièrement l'ARN cible en phase solide en intégrant spécifiquement dans certaines positions choisies des nucléotides non-naturels fluorés. On peut par exemple utiliser des riboses fluorés en position 2'[345] (composé **94**) ou bien des pyrimidines fluorées en position 5[346] (voir figure IV-3) (composés **95** et **96**). Cette stratégie de synthèse totale n'est cependant raisonnable que si l'ARN visé n'est pas trop grand et ne comporte pas trop de modifications post-transcriptionelles.

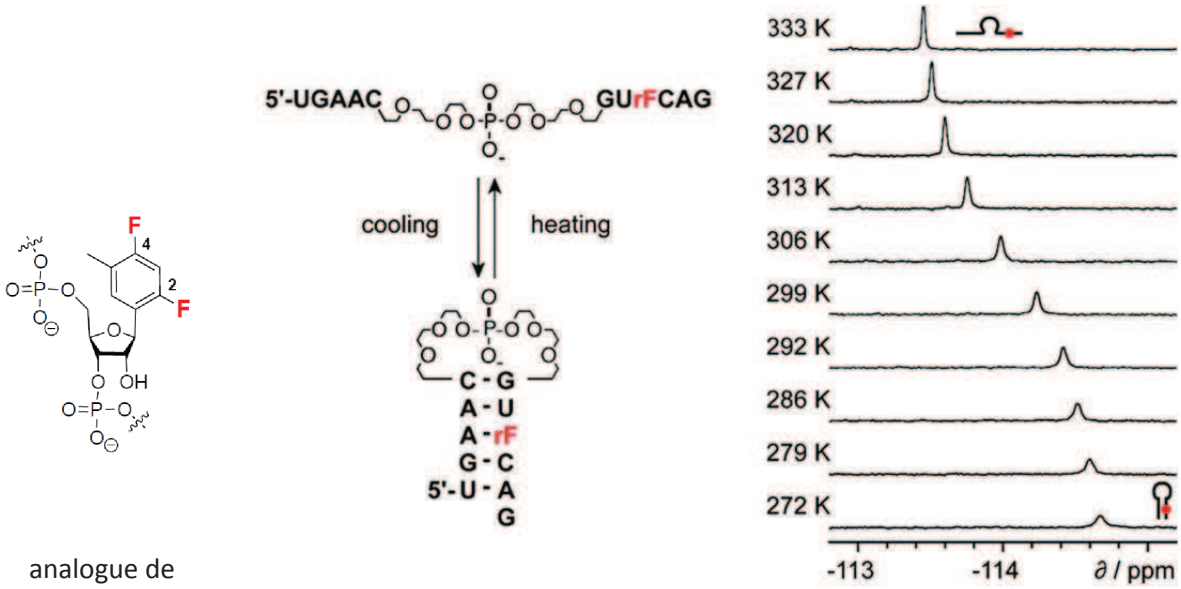



94 95 96

**Figure IV-113 Exemple de nucléotides fluorés utilisés dans les stratégies de marquage d'ARN**

Ces stratégies de marquage des nucléobases des ARN ont permis d'explorer diverses facettes des propriétés biologiques des ARN. Il a par exemple été possible d'étudier des phénomènes de repliement de certains ARN comme le ribozyme Hammerhead[347]. La RMN $^{19}F$ a également permis l'étude des cinétiques de transition entre différents type de structures secondaires[348], ou encore les mécanismes d'hybridation entre oligonucléotides[349].

Le groupe de Micura a récemment décrit une stratégie originale d'insertion de nucléotides fluorés à des positions spécifiques de séquences d'ARN par synthèse en phase solide. Ces nucléotides ont été par exemple utilisés dans la caractérisation des réarrangements conformationnels d'ARN bistables[345]. Cette même approche a été utilisé en intégrant l'analogue de nucléotide fluoré **97** à une tige boucle, rendant possible l'étude de sa dénaturation en fonction de la température en suivant les variations de déplacement chimique du signal du fluor[350] (figure IV-4). De façon satisfaisante, la température de fusion (Tm) de l'ARN calculé à partir de cette méthode est parfaitement en accord avec celle qui peut être calculée par des méthodes classiques utilisant la spectroscopie UV ou la microcalorimètrie. Ceci démontre la pertinence des approches utilisant la RMN.




analogue de nucléotide fluoré **97**

**Figure IV-114 Suivi de la dénaturation d'une pseudo tige boucle d'ARN par RMN $^{19}F$** [350]
La variation du déplacement chimique du fluor du nucléotide modifié (à droite) peut être simplement corrélée à la transition de structure secondaire de l'ARN. A gauche, structure du nucléotide fluoré **97** utilisé dans cette étude.

Ce même groupe a plus tard mis en place une stratégie astucieuse de détection de ligands d'ARN s'appuyant là encore sur l'utilisation d'ARN marqués comportant un ou plusieurs nucléotides fluorés (voir figure IV-5). En fluorant spécifiquement un nucléotide proche du site de liaison, il est possible de détecter la fixation d'un ligand potentiel en observant les perturbations du signal du nucléotide en RMN $^{19}F$ du fluor[351]. En guise de contrôle, on peut intégrer un deuxième nucléotide fluoré dans une zone où aucune interaction n'est attendue et où on ne s'attend donc pas à avoir de modification du signal du fluor.

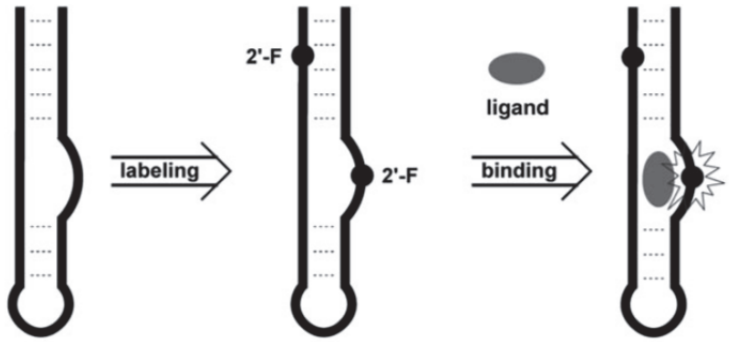



**Figure IV-115 Principe général de l'utilisation d'ARN fluorés dans le criblage de nouveaux ligands**[351]

La fixation d'un ligand induit un bouleversement du signal fluor du nucléotide proche du site de fixation. Le signal fluor du nucléotide contrôle qui est éloigné du site de liaison n'est quant à lui pas modifié.

Comme preuve de concept, les auteurs ont appliqué cette stratégie à l'interaction d'un aptamère avec son ligand. Ils ont modifié un aptamère à la tobramycine en intégrant au niveau de sa boucle d'interaction un nucléotide comportant une base 2'-fluoroadenosine. Un autre nucléotide fluoré, la 2'-fluorocytidine, a été placé au niveau de la tige comme référence interne. En présence d'une concentration croissante de tobramycine, le signal fluoré correspondant à la 2'-fluoroadénosine diminue tandis qu'un nouveau signal, plus déblindé, apparaît. Ceci traduit un échange lent entre le ligand et la cible et la coexistence d'une forme liée et d'une forme libre de l'ARN. Si à l'inverse on était en échange rapide, on verrait un signal moyenné entre forme libre et liée. Le signal fluoré de la 2'-fluorocytidine reste quant à lui inchangé, ce qui démontre bien la spécificité de l'interaction de la tobramycine avec la boucle de l'aptamère. (Figure IV-6).

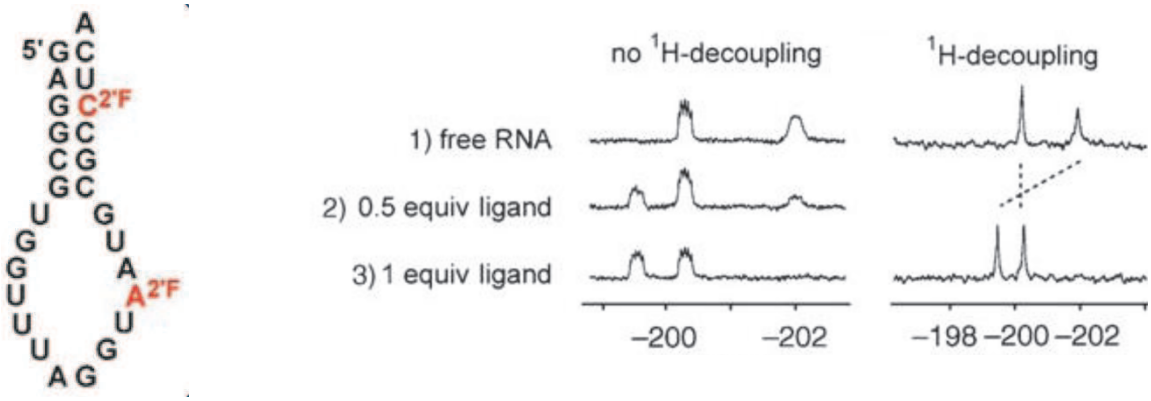



**Figure IV-116 Séquence de l'ARN utilisé (à gauche) et spectre RMN du fluor en présence ou en absence de ligand (à droite)** (adapté de[351]).

### b) Les inconvénients du marquage de la cible

On l'aura compris, fluorer directement un ARN pour ensuite le suivre en RMN $^{19}F$ représente une stratégie très intéressante d'étude des ARN notamment en raison de la simplicité d'analyse des résultats. Cette stratégie présente néanmoins un certain nombre d'inconvénients. Tout d'abord, elle implique de disposer de méthodes de marquage efficaces et spécifiques des cibles visées. Dans l'état actuel des connaissances, il n'est pas possible de marquer des ARN trop grands et/ou disposant de nombreuses bases modifiées. On est ainsi contraint d'étudier de petits ARN que l'on est capable de synthétiser en phase solide, c'est-à-dire inférieurs à 100 voire 150 nucléotides en fonction des technologies disponibles. Enfin, cette stratégie ne constitue en aucun cas une méthode générale d'analyse des ARN puisqu'il faut à chaque fois adapter cette méthode à partir de connaissances sur la structure et les différents sites de fixation de la cible étudiée.

Par ailleurs, dès lors que l'on modifie une biomolécule, on prend le risque d'interférer avec son activité biologique. S'il semble que les fluorations de nucléobases soient assez bien tolérées par les ARN utilisés dans les exemples précédents, ne modifiant que très légèrement leur structure et leur interactions avec des ligands, on ne peut cependant pas préjuger du caractère général d'une telle assertion. La substitution d'un proton par un élément fortement électronégatif est rarement sans conséquence sur les propriétés physicochimiques (polarité, pK, conformation…) d'une molécule.

Pour ces raisons, il nous semblait intéressant de développer des sondes fluorées externes utilisables en RMN $^{19}F$, capables de se lier à une grande variété de cibles biologiques et ne nécessitant aucune modification de celles-ci.

### c) Utilisation d'une sonde fluorée externe

De façon surprenante, alors que l'utilisation de sondes fluorées externes a été bien étudiée dans le cas des protéines (voir les travaux de Dalvit exposés précédemment), il n'y a pas dans la littérature d'exemple de cette stratégie pour étudier les ARN par RMN $^{19}F$ (voir le principe général exposé dans la figure IV-7). Notre objectif a donc été de développer de telles sondes en nous appuyant sur l'architecture des aminoglycosides, connus pour être de bons ligands universels d'ARN. Ces sondes permettraient en effet d'étudier par RMN $^{19}F$ la structure et les interactions des ARN sans toutefois nécessiter leur fluoration de façon covalente.

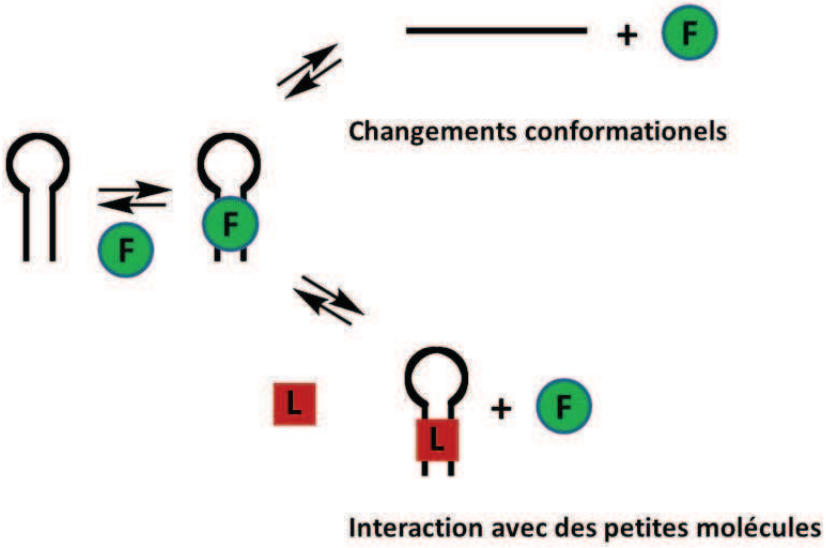



**Figure IV-117 Principe de l'utilisation d'une sonde fluorée externe (en vert) dans l'étude des ARN**

# II - Résultats antérieurs du laboratoire

## *A - Contexte général : synthèse de ligands d'ARN par fragment*

Ce projet s'inscrit dans le contexte plus général de la synthèse de ligands d'ARN par fragment. La partie qui suit expose les travaux préliminaires effectués au laboratoire qui ont permis la mise en place de mon projet de recherche.

Initialement, mon équipe s'est particulièrement intéressée à la conception de ligands venant interagir avec l'ARNt$^{Lys}_{3}$ en raison de son importance dans les premières étapes du cycle de réplication du VIH (figure IV-8) (voir le chapitre I, partie I-C-4 pour plus de détails).

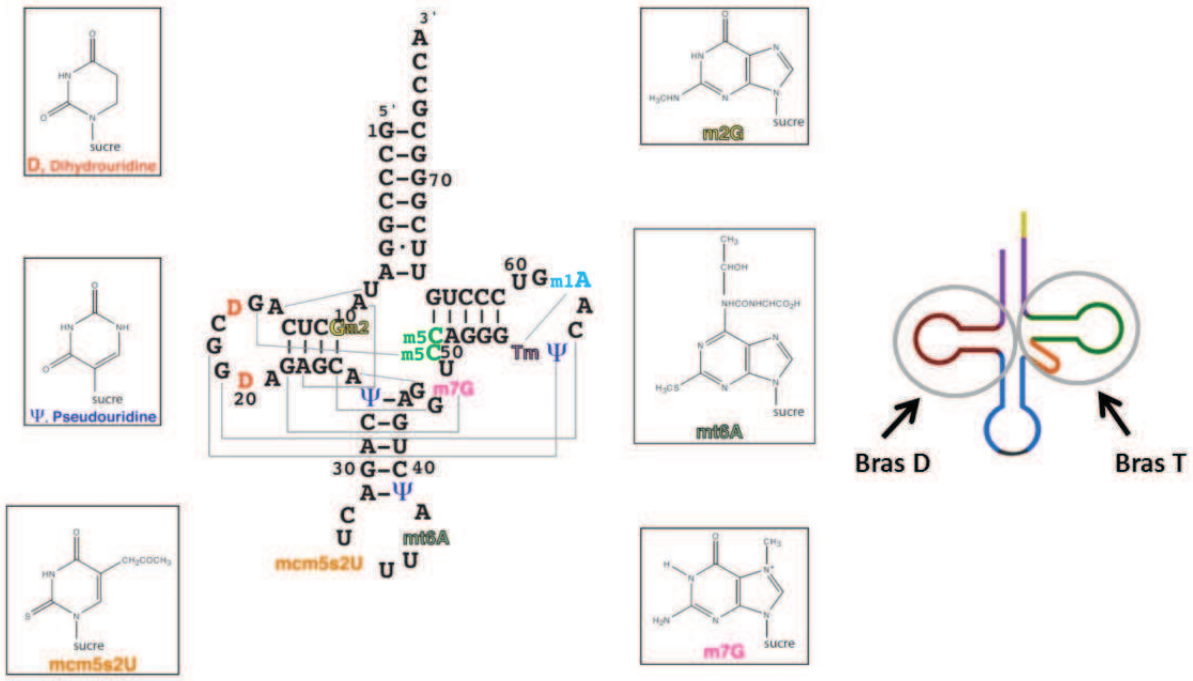



**Figure IV-118 L'ARNt$^{Lys}_{3}$, une cible particulièrement intéressante dans la conception de ligands d'ARN.** Les bases non canoniques sont en couleur dans la séquence et leur structure chimique est représentée dans les encarts (figure adaptée de la thèse de Pierre Barraud)

Pour ce faire, une méthodologie de criblage de fragments à l'aide d'une sonde RMN à flux continu a été développée [352]. Cette stratégie permet l'automatisation du criblage d'un grand nombre de molécules, injectées dans le spectromètre à travers un système de tubulures HPLC couplé à un robot de pipetage programmable. Il est alors possible d'effectuer une variété d'expériences RMN pour étudier l'interaction potentielle des petites molécules avec l'ARN d'intérêt. Dans le cas de l'ARNt$^{Lys}_{3}$, l'étude des variations de déplacements chimiques des protons imino (situés entre 10 et 15

ppm) par RMN $^{1}$H 1D en présence de petites molécules a permis de détecter des ligands potentiels. Dans un second temps, il a été possible d'évaluer leur affinité par titration en RMN 1D. Des expériences d'empreintes en RMN $^{15}$N/$^{1}$H sur de l'ARNt$^{Lys}_{3}$ enrichi en $^{15}$N, dont les pics avaient été préalablement attribués[353,354], ont ensuite permis d'étudier précisément leur site de fixation. Certains fragments ont pu être combinés entre eux pour donner des ligands de plus haute affinité, spécifique du bras T de l'ARNt$^{Lys}_{3}$[62,325].

Au cours de ces études, il a été démontré que le fragment *cis*-1,3-diaminocyclopentanol **98**, qui sera dans la suite du manuscrit appelé **DACP**, constituait un ligand convenable de l'ARNt$^{Lys}_{3}$ disposant d'une affinité modeste de l'ordre du millimolaire. Ce fragment, synthétisé au laboratoire par désymétrisation d'hydrazines bycicliques[314], constitue en effet un mime simplifié du cycle central **DOS** de la plupart des aminoglycosides, connus pour être des « ligands universels » d'ARN (voir figure IV-9). J'ai d'ailleurs pu exploiter cette analogie structurale entre **DACP** et **DOS** au cours de mon M2 pour concevoir des mimes d'aminoglycosides capables d'inhiber une enzyme de résistance inactivant la quasi-totalité des aminoglycosides connus[130].

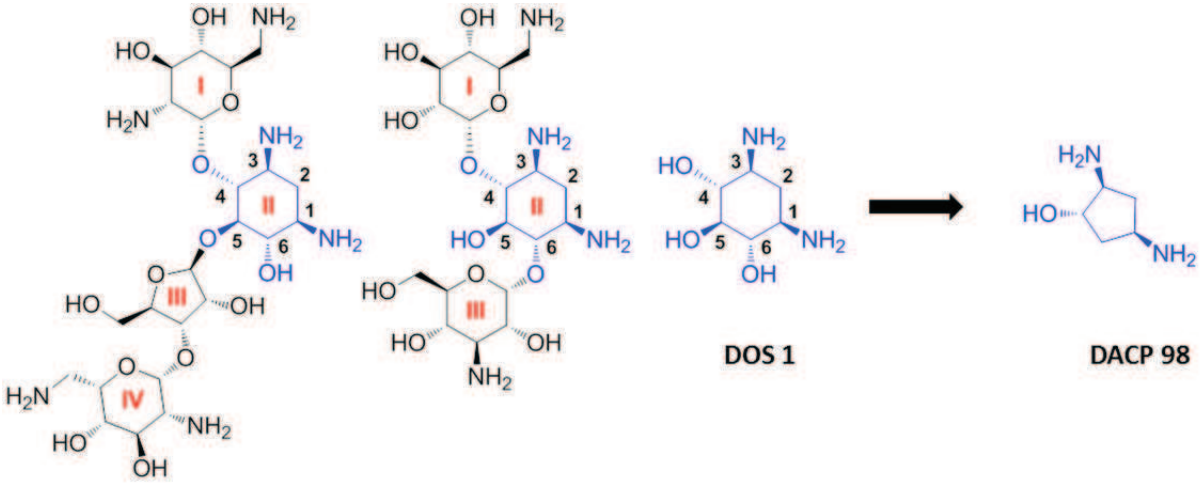
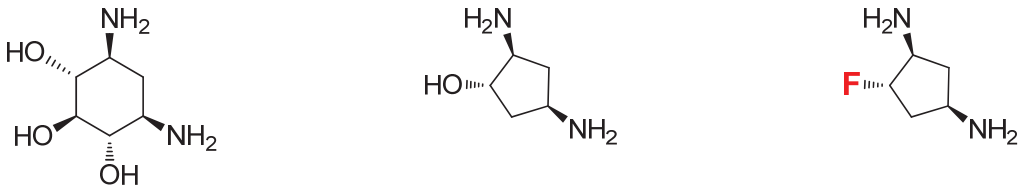



**Figure IV-119 Le cycle diaminocyclopentanol DACP, un mime structural de DOS**

L'intérêt de ce mime simplifié d'aminoglycosides a conduit le laboratoire à chercher à augmenter la diversité moléculaire de ce fragment diaminocyclopentanol, par exemple en introduisant un atome de fluor. Cette diversification a fait l'objet de la thèse de Morgane Pasco et des travaux du docteur Roba Moumné lors de son séjour post-doctoral au laboratoire. Un certain

nombre d'analogues du DACP ont pu être synthétisés[316] et leurs propriétés de reconnaissance vis-à-vis de diverses cibles ARN ont été testées[355]. Dans la suite de notre exposé, nous nous concentrerons plus spécifiquement sur un analogue fluoré particulier du DACP, le *cis*-1,3-diamino-4-fluorocyclopentane **99** (nommé **DACPF** dans la suite du manuscrit) (voir figure IV-10). Les propriétés de reconnaissance de cette molécule fluorée la rendent en effet particulièrement intéressante dans l'étude des ARN par RMN $^{19}F$[355].




**DOS** **DACP (98)** **DACPF (99)**

**Figure IV-120 Le DACPF, analogue fluoré de la DOS**

## *B - Le DACPF comme sonde dans l'étude des ARN par RMN $^{19}F$*

Les travaux présentés ici sont essentiellement le fait de Morgane Pasco et de Roba Moumné. Ils ont fait l'objet d'une publication dans JACS en 2010[355] et d'une publication dans JOC en 2011[316].

### 1) Synthèse du DACPF

La synthèse du DACPF en mélange racémique est très simple et s'effectue en 4 étapes à partir du cyclopentadiène **76** avec un très bon rendement global de 63 % (figure IV-11). Elle s'appuie sur une méthodologie de désymétrisation des hydrazines bicycliques développée depuis quelques années au laboratoire[314]. La première étape de la synthèse met en jeu une réaction de Diels-Alder entre le cyclopentadiène fraichement distillé à partir de son dimère et le diazocarboxylate de dibenzyle selon un protocole décrit dans la littérature[326]. L'hydrazine bicyclique **75** qui résulte de cette cycloaddition peut être obtenue à l'échelle de 25 g avec de très bons rendements par simple précipitation dans le cyclohexane.

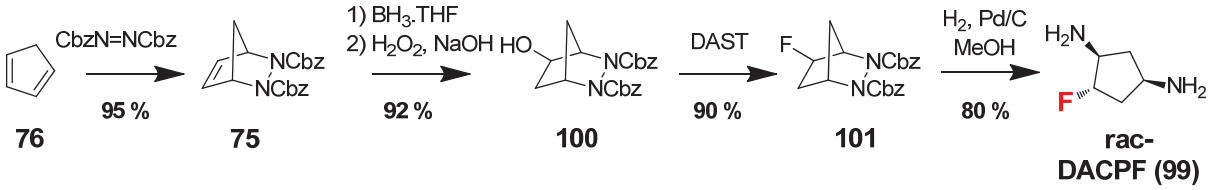



Figure IV-121 Synthèse du DACPF en mélange racémique

L'étape suivante consiste à hydroxyler l'hydrazine bicyclique à l'aide d'une réaction d'hydroboration suivie d'un traitement à l'eau oxygénée pour donner le produit **100**. Le contrôle de la stéréosélectivité de l'attaque du borane sur la face *exo* de l'hydrazine est assuré par l'encombrement stérique causé par un des deux groupements carbamate qui vient pointer vers la face *endo,* rendant impossible toute attaque sur cette face. On peut noter à ce stade qu'une hydroboration assymétrique catalysée au rhodium ou à l'iridium a été développée au laboratoire de façon à obtenir un alcool énantioenrichi[314]. La fluoration de l'hydrazine **100** est effectué par du DAST (pour « diethylaminosulfur trifluoride ») et passe par un intermédiaire aziridinium résultant de la participation du doublet non-liant de l'azote (voir le mécanisme dans la figure IV-12). Le composé intermédiaire formé est un composé méso et conduit donc à la racémisation des produits.

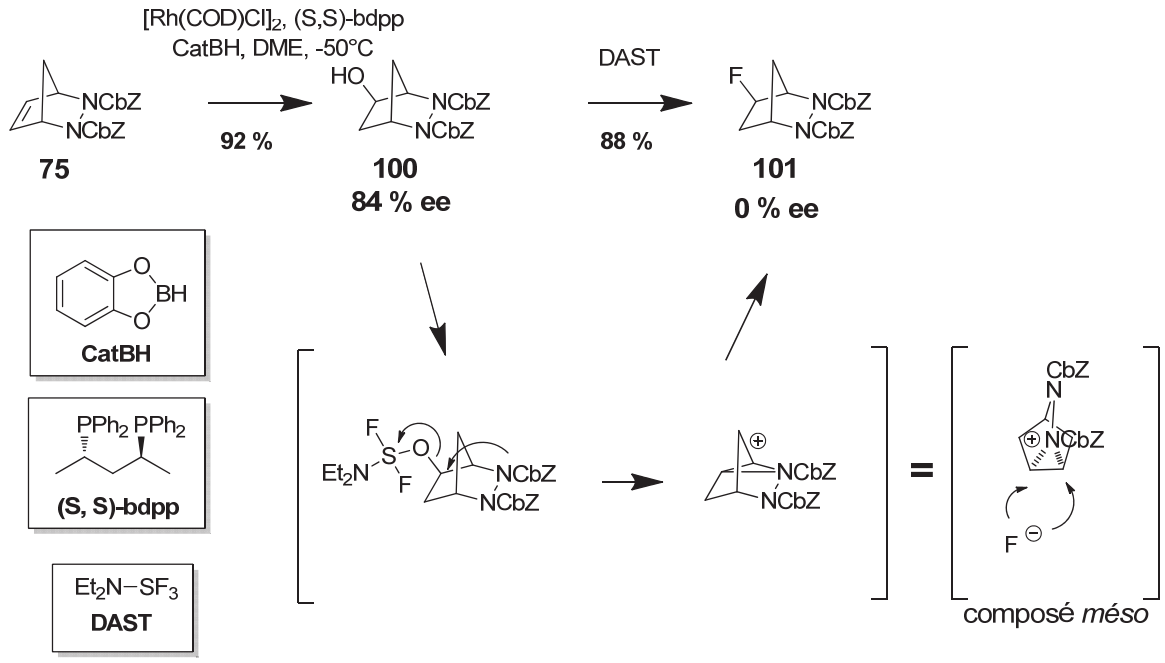
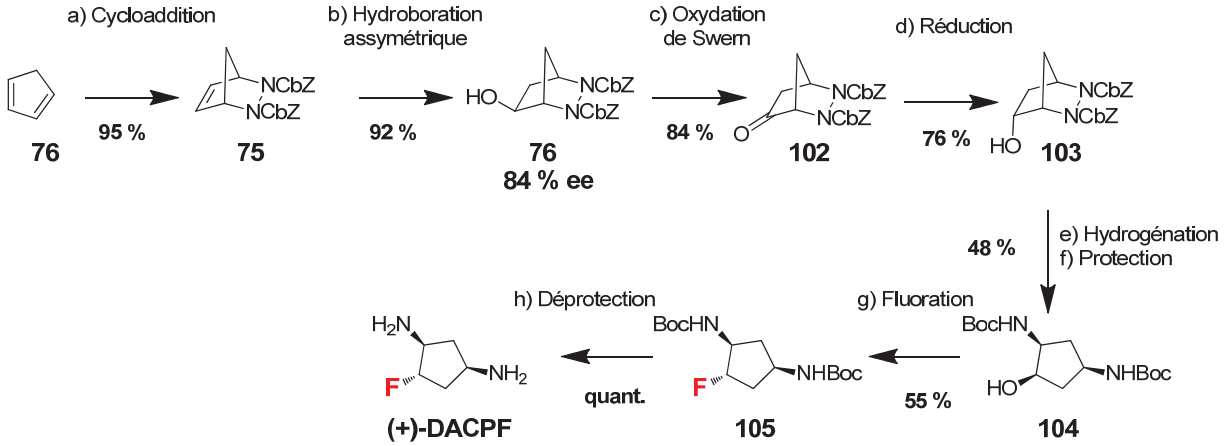



Figure IV-122 Mécanisme de la fluoration de l'hydrazine bicyclique hydroxylée : exemple de la racémisation d'un alcool énantioenrichi issu d'une hydroboration énantiosélective

Enfin, le traitement de l'hydrazine fluorée **101** par du palladium sous atmosphère d'hydrogène permet d'ouvrir le cycle de façon à obtenir le **DACPF** sous forme racémique (voir la partie expérimentale pour plus de détails).

Alternativement, la synthèse de chacun des deux énantiomères du **DACPF** a aussi été développée au laboratoire[355]. Elle passe par l'hydroboration énantiosélective de l'hydrazine en adaptant le ligand chiral en fonction de l'énantiomère voulu. Comme exposé plus tôt, la fluoration ne peut se faire directement sur le bicycle car il y a racémisation. La réaction d'hydroxylation est donc suivie d'une séquence d'oxydation selon la méthode de Swern/réduction au borane pour inverser la configuration de l'alcool, puis d'une ouverture du cycle par hydrogénation catalysée au palladium. Le 1,3-diaminocyclopentanol obtenu est alors protégé par des groupements Boc dans les conditions de Schotten-Baumann, puis fluoré au DAST ce qui s'effectue avec une inversion de configuration. Après déprotection des groupements Boc par traitement au HCl (gaz), on obtient un des deux énantiomères du DACPF en fonction du ligand utilisé lors de la réaction d'hydroboration. L'emploi du (R, R)-bdpp en tant que ligand chiral dans la réaction d'hydroboration conduit *in fine* à l'obtention du (+)-(1R,3S,4S)-cis-1,3-diamine-4-fluorocyclopentane, aussi noté (+)-DACPF (voir figure IV-13).




a) CbZNH=NHCbZ; b)$[Rh(COD)Cl]_2$, (R,R)-bdpp, CatBH, DME, -50°C puis $H_2O_2$, NaOH; c) $(COCl)_2$, $Et_3N$, DMSO; d) $BH_3$.THF
e) $H_2$, Pd/C, MeOH; f) $Boc_2O$, $K_2CO_3$, $H_2O$/THF; g) DAST, DCM, 0°C; h) HCl(g), AcOEt

**Figure IV-123 Synthèse énantiosélective du (+)-DACPF**

### 2) Interaction avec les ARN : le cas de l'ARNt$^{Lys}_3$

L'affinité du DACPF pour l'ARNt$^{Lys}_3$ a pu être évaluée à 0.9 mM à l'aide d'une expérience de titration en RMN $^{1}$H en regardant la variation des protons iminos de l'ARN au fur et à mesure de l'augmentation de la concentration en DACPF.

Les sites mis en jeu lors de l'interaction entre l'ARNt$^{Lys}_3$ et le DACPF ont pu être déterminés par la méthodologie dites des « empreintes » qui consiste à superposer des spectres RMN 2D hétéronucléaires $^{1}$H/$^{15}$N (TROSY) de l'ARN seul ou en présence de son ligand. Les bases dont les signaux RMN sont perturbés correspondent au site de fixation du ligand. Il a été ainsi constaté que le DACPF se liait majoritairement au bras T de l'ARNt$^{Lys}_3$, avec toutefois un site secondaire d'interaction sur le bras D (voir figure IV-14). (Dans la majorité des expériences RMN présentées, les conditions sont les suivantes : tampon phosphate $KHPO_4$ 10mM, pH 6.5, 50 mM KCl, 293K, la concentration de l'ARN et du ligand éventuel étant indiquée (pour plus de précisions, voir la partie expérimentale ou les articles cités).

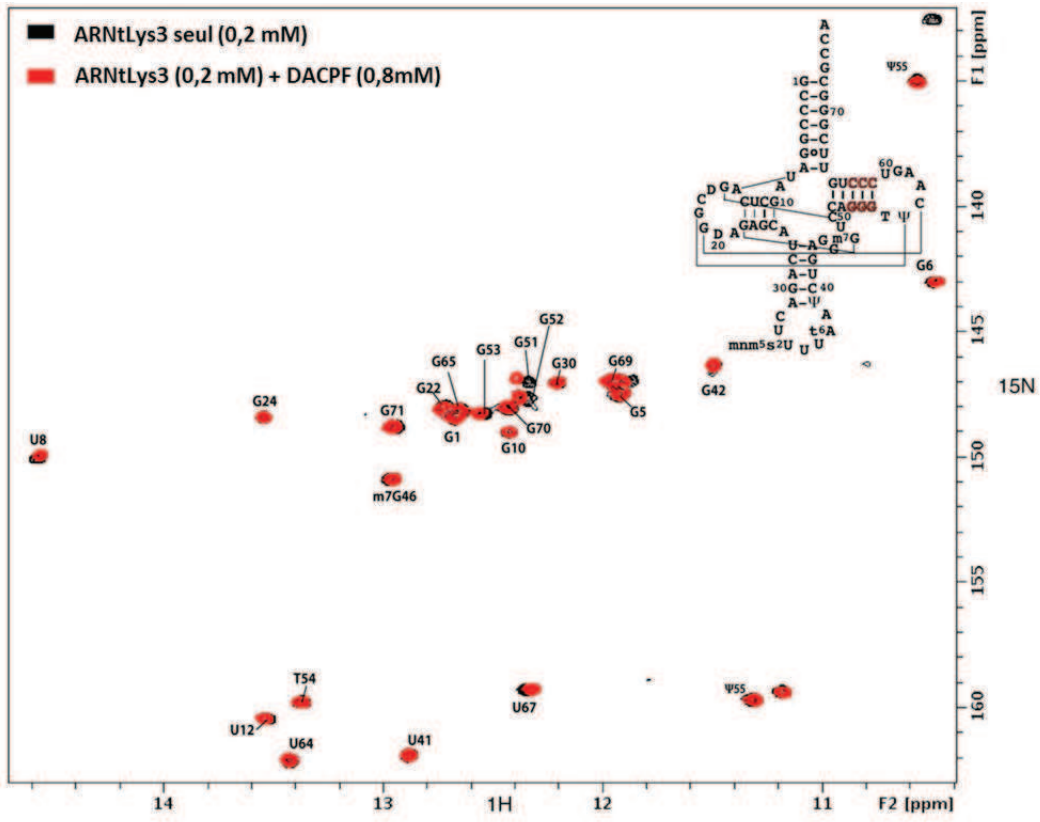



**Figure IV-124 Empreinte TROSY de l'ARNt$^{Lys}_3$ seul (en noir) et de l'ARNt$^{Lys}_3$ en présence du DACPF (en rouge).** Les bases surlignées en rouge correspondent au site d'interaction de DACPF/ ARNt$^{Lys}_3$.

De façon très intéressante, lorsque l'on suit l'interaction du DACPF avec l'ARNt$^{Lys}_3$ par RMN $^{19}$F, on constate que le singulet correspondant au signal du fluor du DACPF en mélange racémique seul se dédouble et est décalé vers les champs faibles (déblindage du signal) (voir la figure IV-15). Lorsque le DACPF fluoré est mis en présence du seul bras T de l'ARNt$^{Lys}_3$, ce dédoublement caractéristique est préservé tandis que si le DACPF est mis en présence du bras D de l'ARNt$^{Lys}_3$, il n'y a plus de dédoublement. Ceci est en accord avec les données d'empreintes 2D suggérant que le bras T constitue le site d'interaction majoritaire de ce dernier, les interactions avec le bras D ne constituant que des interactions non spécifiques de faible affinité.

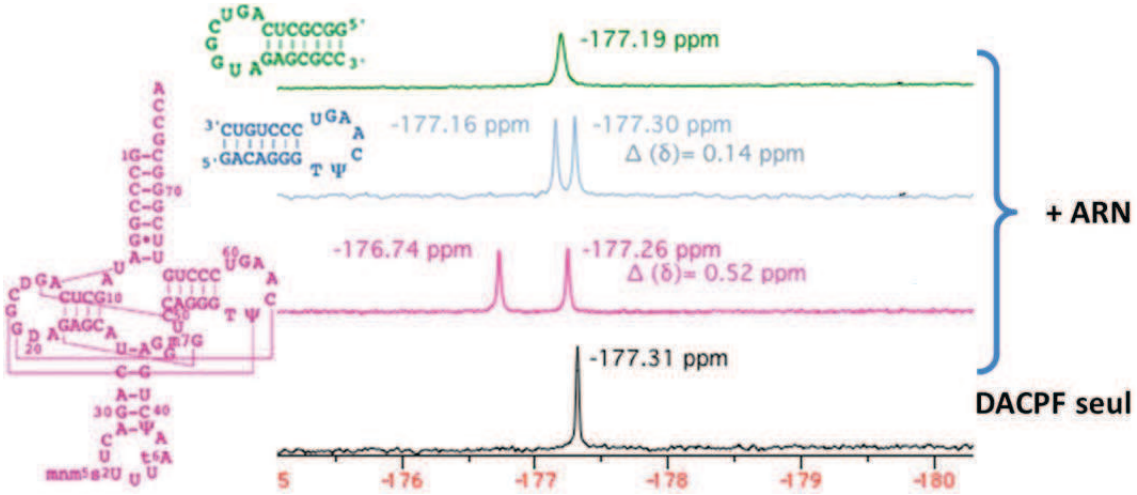



**Figure IV-125 Suivi en RMN $^{19}$F de l'interaction du DACPF avec l'ARNt$^{Lys}_3$**
En noir, spectre fluor du DACPF seul puis spectre résultant de l'interaction du DACPF avec l'ARNt$^{Lys}_3$ entier (en rose), avec le bras T (en bleu) ou avec le bras D (en vert) de l'ARNt$^{Lys}_3$
(DACPF (1 mM), ARN (62.5 µM) (16 : 1), tampon phosphate 10mM, 50 mM KCl, pH 6.5)

Ces observations suggèrent la formation de deux complexes diastéréomériques entre chacun des deux énantiomères du DACPF et l'ARNt$^{Lys}_3$. Il y a donc reconnaissance supramoléculaire chirale de la molécule fluorée par l'ARN, avec, pour chaque énantiomère, un environnement électronique et physicochimique différent pour l'atome de fluor résultant en un déplacement chimique différent. Comme la cible et le ligand sont en échange rapide par rapport au temps caractéristique de la mesure RMN, chacun des deux signaux observés correspond au signal moyenné en forme libre et liée des deux énantiomères.

Pour vérifier cette explication sur l'origine de ce dédoublement, il a été décidé de suivre par RMN $^{19}F$ l'interaction de l'$ARNt^{Lys}_3$ avec chacun des deux énantiomères du DACPF séparément. Cette expérience, présentée dans la figure IV-16, a été rendue possible par la mise au point d'une synthèse énantiosélective de chacun des deux énantiomères. Comme attendu, en mélange avec l'$ARNt^{Lys}_3$, chaque énantiomère donne bien un seul signal du fluor avec un déplacement chimique différent, qui correspond à un des deux pics du doublet observé avec la molécule racémique. En guise de contrôle, il a été vérifié par RMN que les deux énantiomères avaient la même affinité pour l'$ARNt^{Lys}_3$ (de l'ordre du micromolaire) ainsi que le même site d'interaction sur le bras T. Ceci permet d'affirmer que les différences entre déplacement chimique correspondent bien à la formation de deux complexes diastéréomériques et non à une interaction différente pour chacun des deux énantiomères. Un certain nombre d'autres analogues fluorés du DACP ont par ailleurs été testés. Ils ont le même comportement que le DACPF en RMN $^{19}F$ lorsqu'ils sont mis présence de l'$ARNt^{Lys}_3$. Le dédoublement observé est cependant moins important ce qui diminue leur intérêt dans la conception de sondes en RMN $^{19}F$. On peut en conclure que le dédoublement du fluor observé lors de l'interaction de ces molécules fluorées avec un ARN est très fortement dépendant de la géométrie de l'interaction entre ces deux partenaires et plus particulièrement de la position de l'atome de fluor dans les deux complexes diastéréotopiques formés.

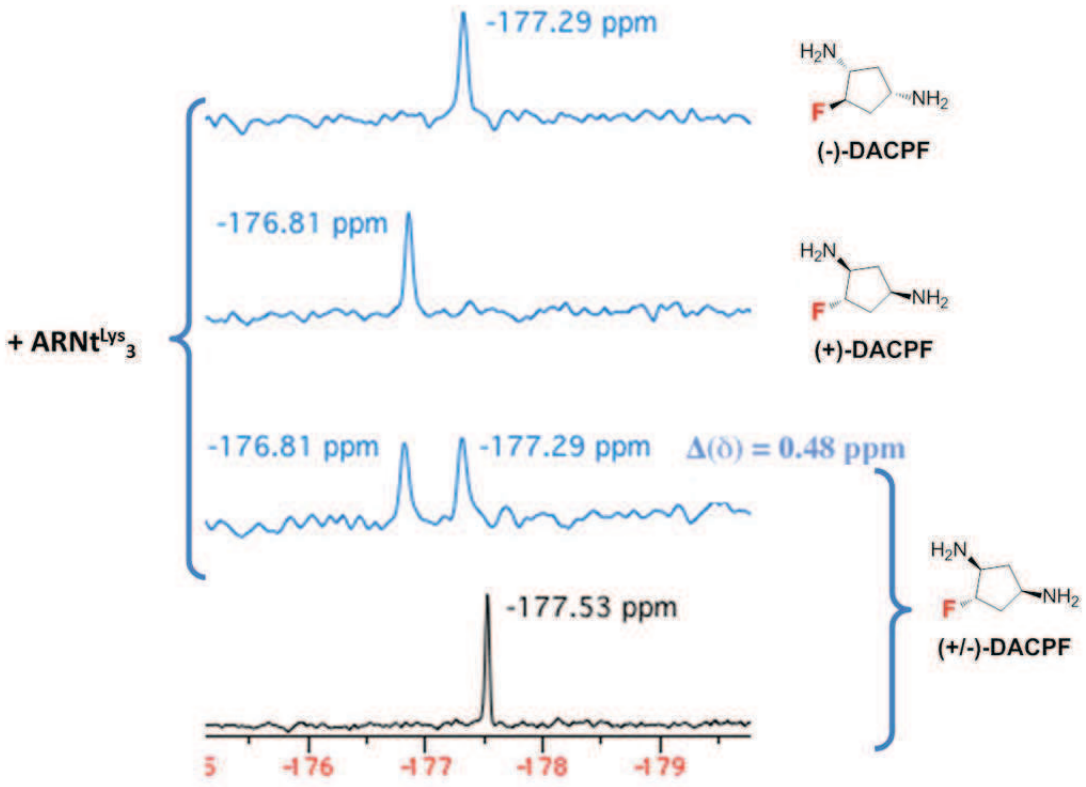



**Figure IV-126 Suivi par RMN $^{19}$F de l'interaction de chacun des deux énantiomères du DACPF avec l'ARNt$^{Lys}_3$.**

(DACPF (1 mM), ARNt$^{Lys}_3$ (62.5 µM) (16 : 1), tampon phosphate 10mM, pH 6.5)

L'influence de la stœchiométrie ARN/DACPF sur le dédoublement du signal du fluor a également été étudiée (figure IV-17). Comme attendu, plus on augmente le ratio ARNt$^{Lys}_3$/DACPF, plus on déplace l'équilibre de l'interaction vers la forme liée du DACPF et plus on accentue le dédoublement du signal du fluor et son déblindage. Ces résultats confirment le fait que l'on est en échange rapide : le signal RMN observé correspond à la moyenne des signaux des formes libres et liées du DACPF, pondérée de leur proportion respective.

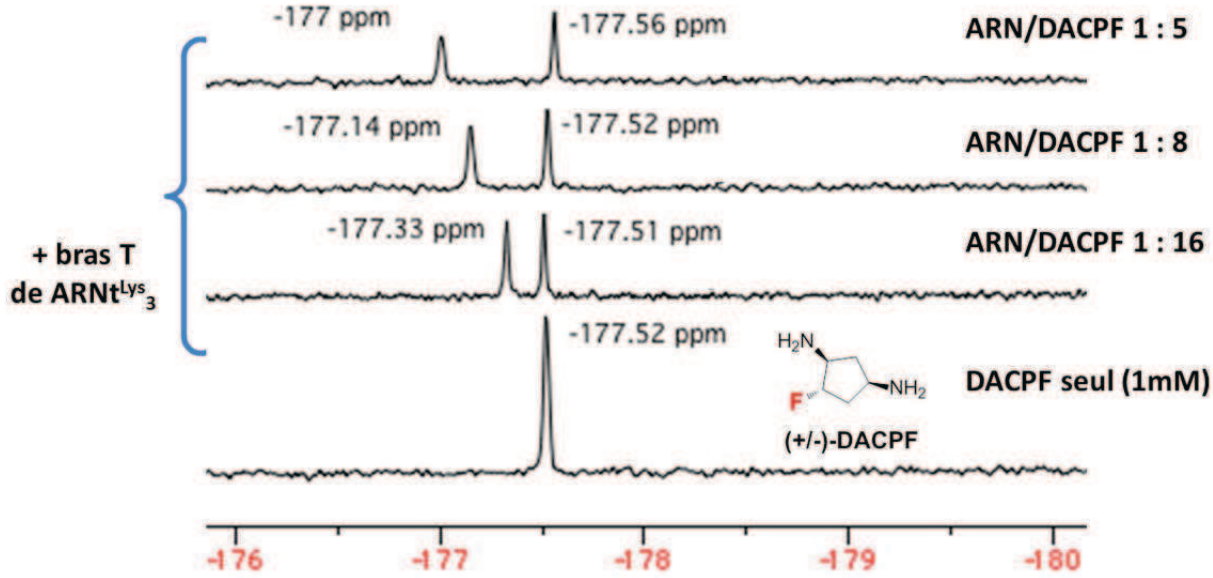



**Figure IV-127 Influence de la stœchiométrie ARN/DACPF sur le dédoublement du signal du fluor**

(DACPF (1 mM), bras T ARNt$^{Lys}_3$ (62.5 µM à 200 µM) (16 : 1), tampon phosphate 10mM, pH 6.5)

### 3) Le DACPF comme sonde topologique des ARN

Le dédoublement du signal du fluor du DACPF racémique en présence d'ARN en fait potentiellement une très bonne sonde de la structure des ARN en RMN $^{19}$F. En effet, ce dédoublement est extrêmement dépendant de la structure de l'ARN avec lequel le DACPF interagit. Pour tester cette possibilité, le docteur Roba Moumné a suivi en RMN $^{19}$F l'interaction du DACPF avec un certain nombre d'ARN tels que l'ARNt$_m^{Met}$, l'ARNt$_f^{Met}$ ou l'ARNt$^{Phe}$, qui présentent des structures secondaires en feuille de trèfle similaires à celle de l'ARNt$^{Lys}_3$ mais des structures tertiaires légèrement différentes. En dépit de leur similarité structurale, pour chaque ARN, on obtient un dédoublement différent du signal du DACPF (figure IV-18).

Ces expériences permettent de souligner la sensibilité du dédoublement du signal du fluor du DACPF vis-à-vis de la structure de l'ARN étudié. Le bouleversement du signal du fluor par rapport à la sonde libre constitue donc en quelque sorte une signature topologique de l'ARN mis en contact avec la sonde. De façon très intéressante, le DACPF dispose d'une affinité modeste (Kd de l'ordre du millimolaire) pour une grande variété d'ARN différents. Ceci permet de potentiellement étudier de nombreuses cibles biologiques différentes et aussi de s'assurer que l'interaction du DACPF avec ses cibles ne modifie pas trop la structure ou les propriétés physicochimiques de celles-ci. On peut ainsi envisager d'utiliser le DACPF comme sonde structurale dans l'étude des ARN par RMN $^{19}$F.

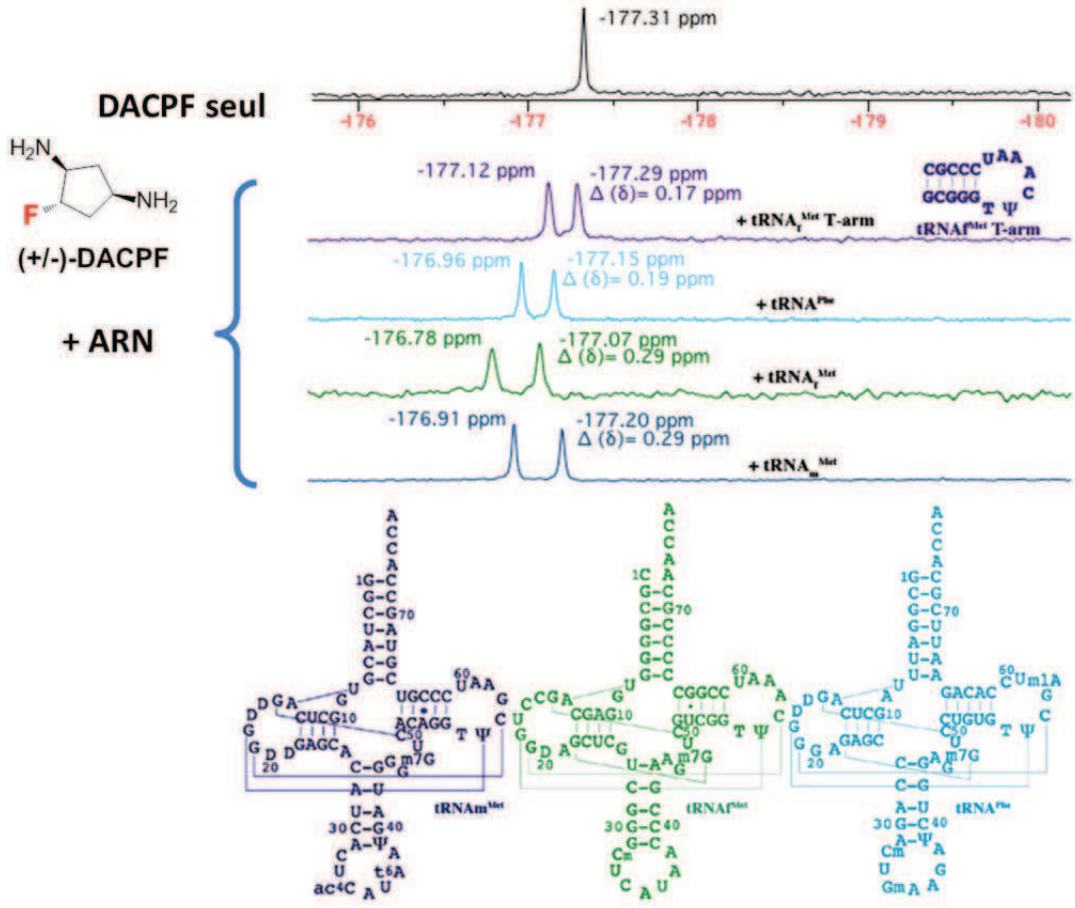



**Figure IV-128 Interaction du DACPF avec divers ARN : le dédoublement du signal du fluor constitue une signature topologique de la structure des ARN testés**
(DACPF (1 mM), ARN (62.5 µM) (16 : 1), tampon phosphate 10mM, 50 mM KCl, pH 6.5)

### 4) Suivi de changements conformationnels: exemple de la dénaturation d'une tige boucle

Ayant démontré que l'on pouvait utiliser le DACPF comme sonde de structure des ARN, nous avons voulu évaluer son utilisation pour suivre des phénomènes dynamiques de changement structurels des ARN. Pour ce faire, nous avons choisi le cas simple de la dénaturation thermique progressive de la tige boucle correspondant au bras T de l'ARNt$^{Lys}_3$.

Lorsque la température de l'échantillon RMN augmente, le déplacement chimique de chacun des énantiomères varie et se déplace vers les champs faibles (déblindage). Par ailleurs, l'écart entre les deux signaux se réduit progressivement. Une coalescence complète est atteinte à 350 K, température à laquelle l'ARN est totalement déstructuré comme l'attestent les expériences UV réalisées sur le même échantillon (voir figure IV-19).

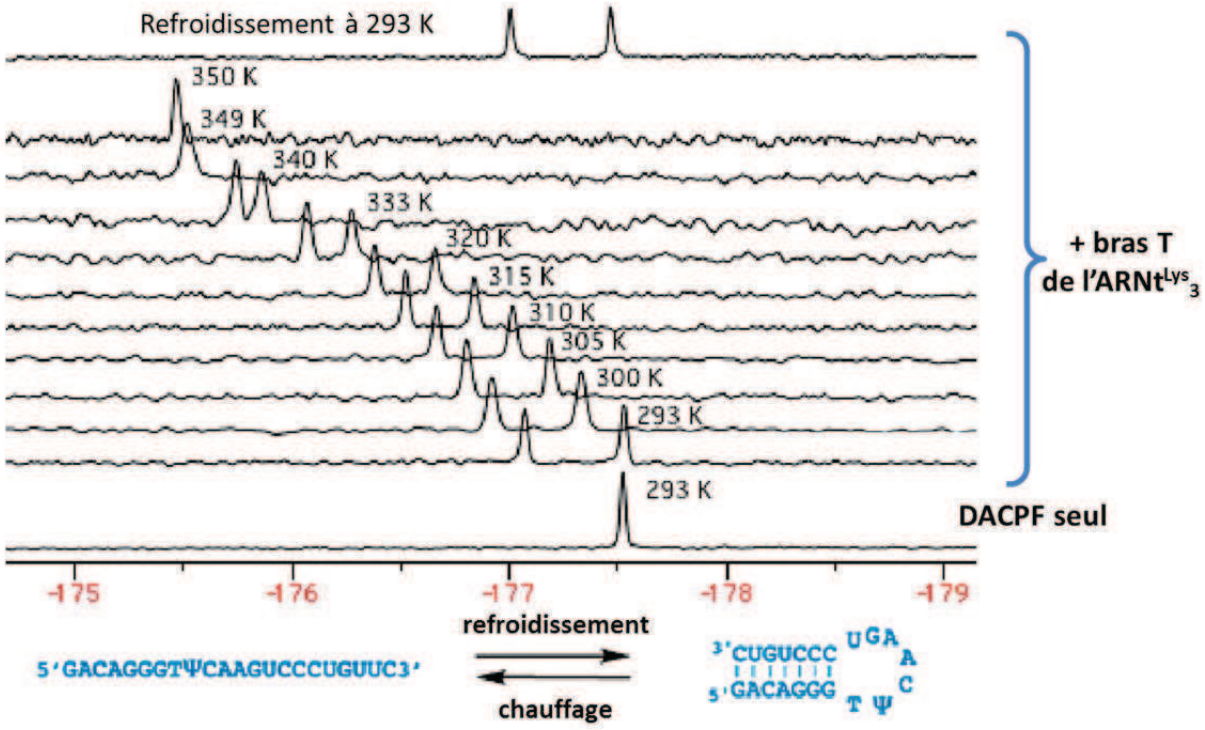



**Figure IV-129 Suivi de la dénaturation du bras T de l'ARNt$^{Lys}_3$ en RMN $^{19}$F**
(DACPF (1 mM), ARN (167 µM) (6 : 1), tampon phosphate 10mM, pH 6.5, température variable)

Lorsque l'on refroidit l'échantillon RMN, on retrouve le dédoublement caractéristique de l'interaction du DACPF sous forme racémique avec le bras T de l'ARNt$^{Lys}_3$, ce qui signe le repliement de l'ARN dans sa structure en tige-boucle originale.

Cette expérience constitue une preuve de concept de l'intérêt des sondes fluorées externes pour suivre en RMN $^{19}$F le changement conformationnel d'un ARN. De façon très intéressante, à l'inverse de la méthodologie de marquage des ARN développée par l'équipe de Micura, cette stratégie ne nécessite pas de modification de la cible ARN. Par ailleurs, cette méthode peut se faire sur un simple spectromètre 300 MHz (valeur du $^{1}$H) avec des expériences de seulement quelques minutes en raison de la sensibilité de la RMN $^{19}$F. Ceci est à comparer avec les expériences de suivi de changement conformationnel par RMN $^{1}$H qui nécessitent souvent des spectromètres à haut champ et des temps d'accumulation relativement longs (dans notre cas, nous avons utilisé un spectromètre 600 MHz équipé d'une cryosonde).

# III - Résultats personnels

Au cours de mon travail de thèse, je me suis appuyé sur ces travaux pour étendre l'utilisation de sondes fluorées à l'étude en RMN $^{19}F$ d'interactions ARN/petites molécules. En m'intéressant au cas particulier d'un aptamère à la néomycine, j'ai développé une méthode d'étude des changements conformationnels induits lors de l'interaction d'un ARN avec son ligand.

En m'appuyant sur les travaux préliminaires initiés par Roba Moumné, j'ai par ailleurs exploité les propriétés du DACPF pour caractériser précisément l'affinité de petites molécules pour une cible nucléique particulière mimant le site de décodage du ribosome, l'ARN 16S23. Ceci m'a amené à concevoir une nouvelle sonde fluorée plus sensible que le DACPF, la paromamine fluorée, directement inspirée de la structure des aminoglycosides. L'ensemble de ces résultats a fait l'objet d'une publication dans Angew. Chem. Int. Ed. en 2012 dont on peut trouver une version en annexe[356].

## *A - Suivi de changements conformationnels: capture conformationnelle d'un aptamère par son ligand*

Dans l'expérience de dénaturation du bras T de l'ARNt$^{Lys}_3$ précédemment évoquée, les changements de température n'influent pas uniquement sur le repliement de l'ARN mais aussi sur le Kd du DACPF et sur les déplacements chimiques du fluor qui sont tous deux fonction de la température. Il est par conséquent délicat de faire une analyse quantitative de la dénaturation de l'ARN car cela nécessiterait de déconvoluer ces trois phénomènes.

Pour dépasser cette limitation, nous nous sommes appliqués à mettre en place le suivi d'un changement conformationnel d'ARN ne nécessitant pas de changement de température. Nous avons pour cela choisi d'étudier le phénomène de capture conformationnelle d'un aptamère par son ligand en nous intéressant plus particulièrement au cas d'un aptamère à la néomycine récemment décrit par l'équipe de Suess[357]. Cet aptamère, petite séquence de 27 nucléotides capable de reconnaitre la néomycine avec une affinité très forte (Kd de l'ordre du nanomolaire), a été obtenu en utilisant une stratégie d'évolution dirigée de type SELEX (voir partie I-I-C-2). Sa structure et les implications de sa dynamique conformationnelle en terme d'activité biologique ont été étudiées en détail dans les équipes de Suess et Wöhnert[358,36]. Ceci en fait une bonne cible pour établir une preuve de concept de l'intérêt d'utiliser des sondes fluorées pour étudier les changements conformationnels des ARN en RMN $^{19}F$.

Cet aptamère est particulièrement intéressant car il est en équilibre entre plusieurs conformations lorsqu'il est seul en solution. Sa liaison avec la néomycine se traduit par un bouleversement du spectre RMN $^{1}$H des protons imino de l'aptamère. On parle de capture conformationnelle de l'aptamère par son ligand car l'équilibre entre les différentes conformations de l'aptamère est complétement déplacé vers une conformation « liée » au départ très minoritaire (voir figure IV-20). Il y a ainsi changement drastique d'une conformation « ON » libre à une conformation « OFF » où l'aptamère, lié à la néomycine, adopte une conformation rigide qui empêche l'expression d'éventuels gènes situés en aval à l'instar du fonctionnement des riboswitches naturels.

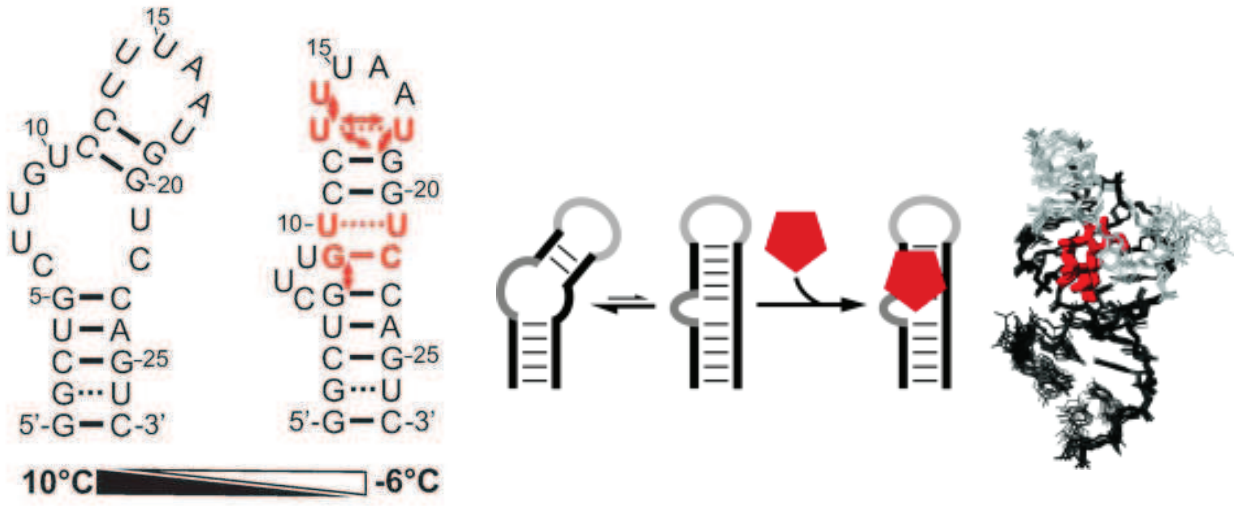
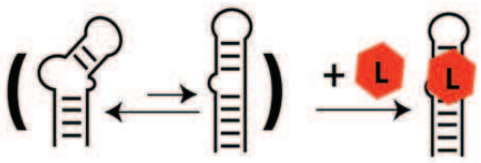



**Figure IV-130 Structure de l'aptamère à la néomycine étudié**[358]
En solution, l'aptamère est en équilibre entre plusieurs conformations en fonction de la température. La liaison de la néomycine (en rouge) déplace cet équilibre vers une conformation « liée » plus rigide.

Afin de pouvoir tester l'utilisation du DACPF dans le suivi de ces changements conformationnels en RMN $^{19}$F, il était important de bien caractériser cette transition en RMN $^{1}$H pour pouvoir disposer d'un contrôle fiable. La figure IV-21 permet de constater que la liaison de la néomycine à l'aptamère induit bien un bouleversement des protons imino de ce dernier, ce qui traduit un changement drastique de conformation (voir partie expérimentale).



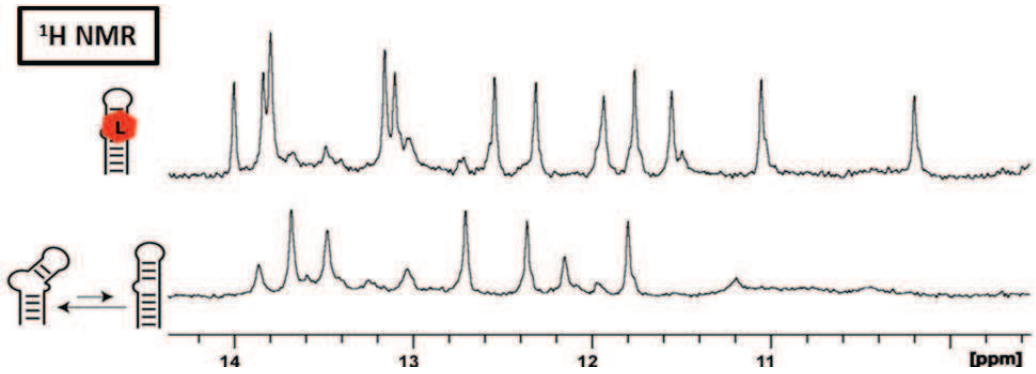



**Figure IV-131 Suivi en RMN $^{1}$H des variations des signaux des protons imino de l'aptamère lors de la fixation de la néomycine (en rouge)**

(ARN (200 µM), néomycine (200 µM (1 : 1), tampon phosphate 10 mM, 50 mM KCl, pH 6.5, 293K)

Lorsqu'on étudie l'interaction de l'aptamère avec le DACPF, on constate que l'équilibre conformationel de l'aptamère ne semble pas modifié comme l'attestent en RMN $^{1}$H les signaux des protons iminos qui restent semblables à ceux de l'aptamère seul. Cette propriété du DACPF est particulièrement importante pour son emploi en tant que sonde en RMN $^{19}$F car on s'assure alors que la sonde n'introduit qu'une perturbation négligeable du système biologique étudié. En suivant le décalage progressif du signal du proton imino du nucléotide $G_{25}$ avec l'augmentation de la concentration en DACPF, nous avons effectué une titration en RMN $^{1}$H et évalué l'affinité du DACPF à 0.8 mM (+/- 0.2 mM) (voir partie expérimentale).

De façon très intéressante, on constate en RMN $^{19}$F le dédoublement du signal du DACPF, ce qui constitue une signature de son interaction avec l'ARN que l'on va pouvoir exploiter pour suivre les changements conformationnels de ce dernier (voir figure IV-22).

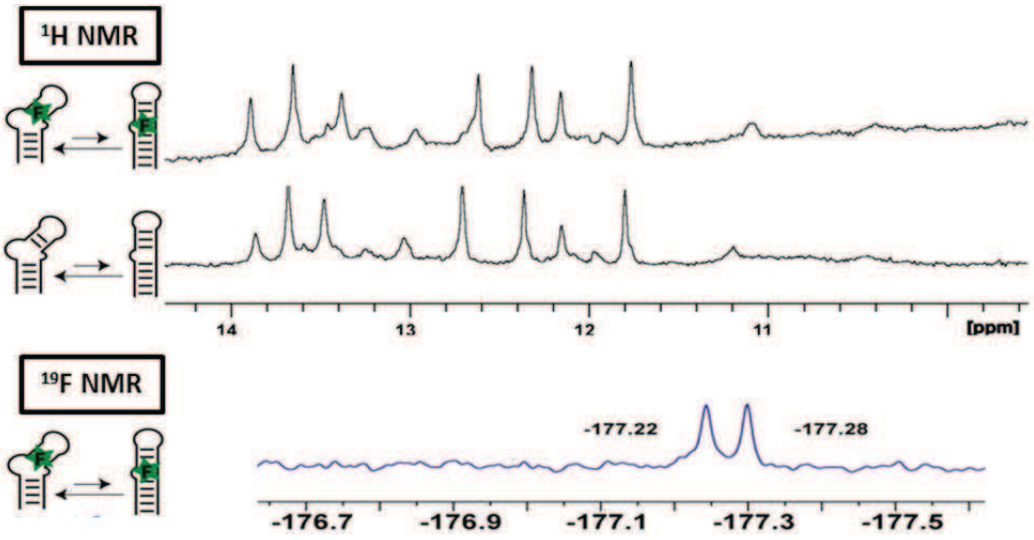



**Figure IV-132 Suivi par RMN $^{1}H$ et $^{19}F$ de l'interaction du DACPF (en vert) avec l'aptamère à la néomycine (en noir)**

(ARN (200 µM), DACPF (200 µM (1 : 1), tampon phosphate 10 mM, 50 mM KCl, pH 6.5, 293K)

L'étape suivante était de voir si le DACPF permettait de suivre les changements conformationnels induits par la fixation de la néomycine sur l'aptamère. Nous avons donc effectué une expérience de compétition entre le DACPF et la néomycine dont la représentation schématique et le suivi en RMN $^{1}H$ et $^{19}F$ sont présentés dans la figure IV-23.

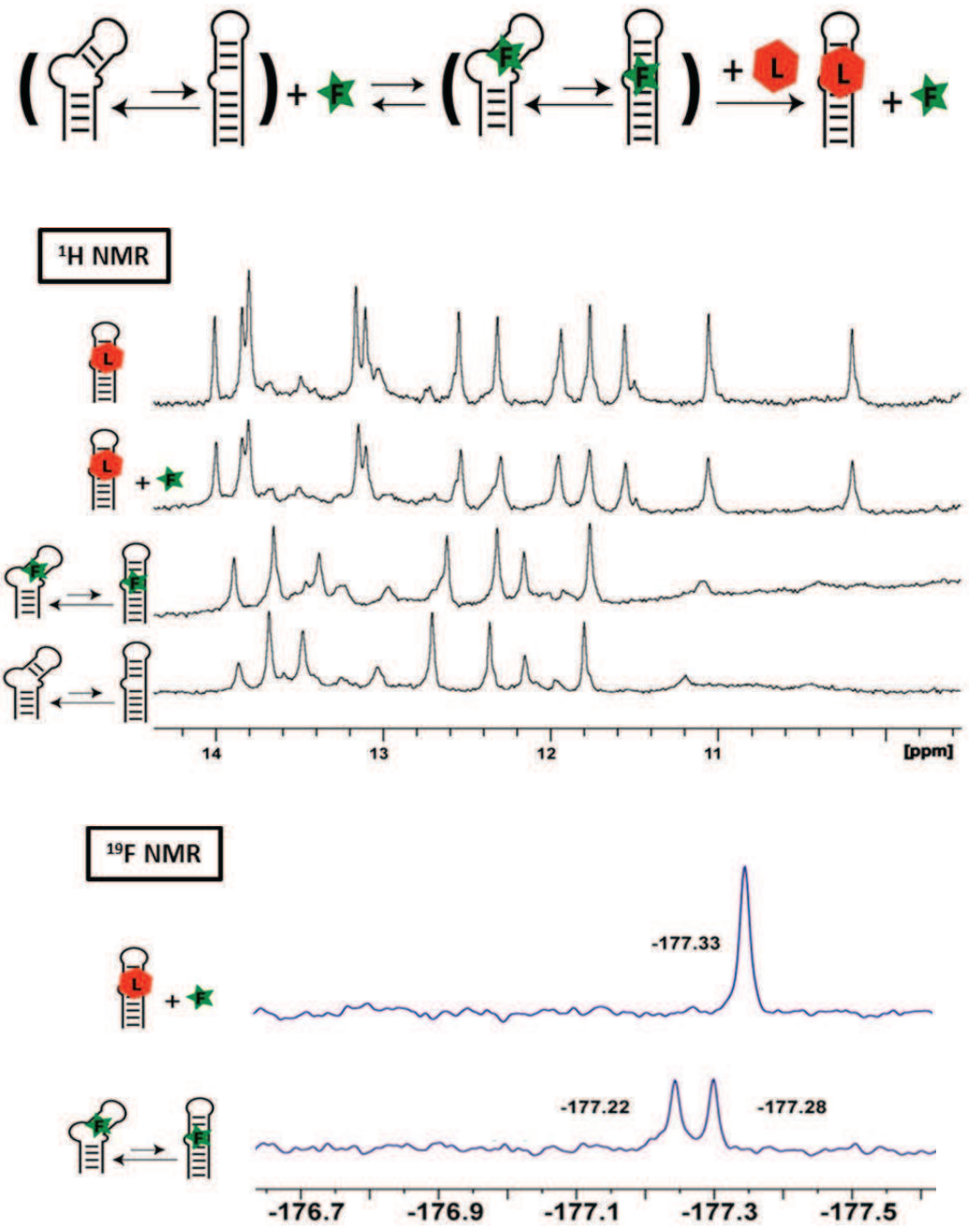



**Figure IV-133 Suivi en RMN $^{1}$H et $^{19}$F de la capture conformationnelle (en noir) de l'aptamère par la néomycine (en rouge) en présence de DACPF (en vert)**

(ARN (200 µM), DACPF (200 µM (1 : 1), tampon phosphate 10 mM, 50 mM KCl, pH 6.5, 293K puis ajout de néomycine (300 µM (1 : 1.5)

Lors de l'addition de la néomycine, on peut constater un bouleversement des signaux des protons imino en RMN $^{1}$H identique à celui qui est observé dans l'expérience sans DACPF (figure IV-21). En RMN $^{19}$F, on peut voir une coalescence du signal du DACPF pour donner le signal du DACPF

libre, ce que l'on peut interpréter comme résultant du déplacement total du DACPF par la néomycine qui induit un changement conformationnel de l'aptamère.

Cette expérience démontre que l'on peut utiliser le déplacement compétitif d'un petit ligand fluoré pour suivre de façon très simple en RMN $^{19}F$ la capture conformationnelle d'un aptamère par son ligand. Cette approche présente l'avantage d'être extrêmement simple à mettre en œuvre car elle ne requiert aucune attribution complexe des spectres RMN de l'ARN. Par ailleurs, contrairement aux méthodologies classiques de la RMN, cette approche n'est pas limitée par la taille de l'ARN étudié (un ARN trop grand conduit par exemple souvent à un spectre RMN $^{1}H$ difficilement interprétable en raison des recouvrements de pics et de l'augmentation du temps de relaxation $T_2$), ni par son marquage ou ses éventuelles modifications post-transcriptionnelles.

Les perspectives offertes par cette expérience sont multiples. On peut en effet envisager d'utiliser cet aptamère en combinaison avec le DACPF pour développer une méthodologie de criblage d'analogue d'aminoglycosides (la durée d'une expérience de déplacement, réalisée sur un 300 MHz, est d'eniron 3 min). La simplicité des spectres du fluor permet même d'envisager une automatisation de leur analyse. Il est alors possible d'imaginer une stratégie de criblage RMN à débit élevé combinant test en flux continu de mélanges de ligands potentiels, analyse des résultats assistée par ordinateur et préparation des échantillons par des automates robotisés.

## *B - Evaluation de l'affinité de ligands vis-à-vis d'une cible ARN : exemple de l'ARN 16S23*

### 1) La cible : l'ARN 16S23, analogue du site de décodage du ribosome bactérien

Ayant démontré l'intérêt d'utiliser des sondes fluorées en RMN $^{19}F$ pour caractériser la topologie des ARN mais aussi leur dynamique conformationnelle, nous avons souhaité évaluer leur utilisation comme outils dans l'étude de l'interaction des ARN avec de petites molécules. Nous avons pour cela choisi d'étudier une cible nucléique très bien caractérisée, l'ARN 16S23 (voir figure IV-24).

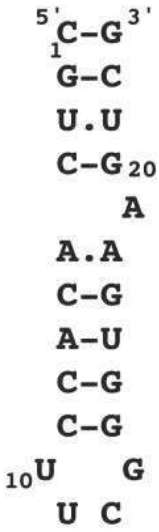
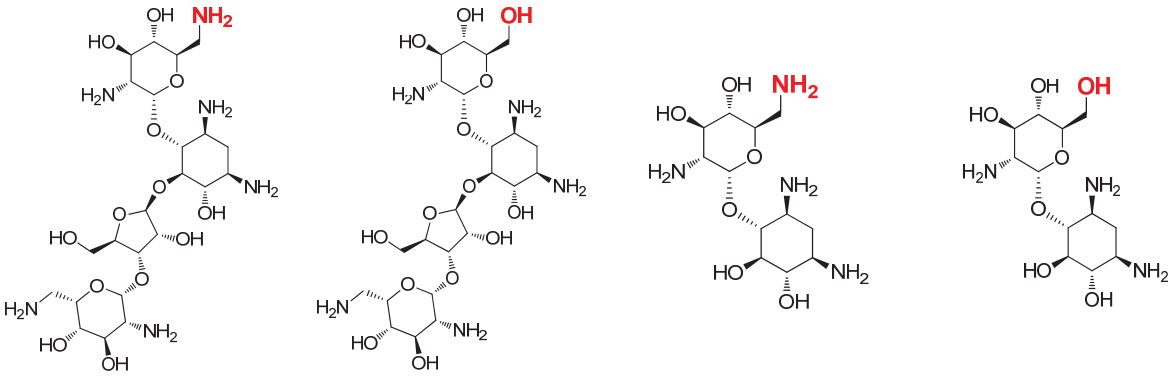



**Figure IV-134 Structure secondaire de l'ARN 16S23 utilisé, mime du site de décodage du ribosome bactérien.**
Les tirets représentent les paires de bases canoniques et les points les paires de bases non-canoniques.

Comme il avait été montré par l'équipe de Puglisi au milieu des années 90[111,112], ce petit ARN de 23 nucléotides constitue en effet un mime du site de décodage du ribosome bactérien, cible d'un certain nombre d'antibiotiques comme les aminoglycosides par exemple[8]. On dispose donc de nombreux ligands spécifiques de cette cible ARN, dont l'affinité a été déterminée par diverses méthodes comme la spectrométrie de masse[116], la RMN $^{1}$H[111,44], la résonnance plasmonique de surface (SPR)[117], la microcalorimètrie[119,98] ou bien encore des expériences quantitatives d'empreintes protégeant la cible vis-à-vis d'une modification chimique (« quantitative chemical footprinting »)[111,44]. S'il est difficile de comparer la valeur absolue de ces affinités d'une étude à l'autre en raison de la grande diversité des conditions expérimentales utilisées (tampon, force ionique, température…), on peut néanmoins retenir un classement global des ligands par ordre d'affinité. Ainsi, on peut ainsi retenir que la néomycine est plus affine que la paromomycine, elle-même plus affine que la

[8] Pour être plus précis, l'équipe de Puglisi a travaillé avec une version légèrement plus longue de cet ARN, comportant 27 nucléotides et non 23. Nous avons démontré par des calculs de Kd en RMN $^{1}$H que l'ARN 16S23 avait les mêmes propriétés de reconnaissance que le mime original de 27 nucléotides et nous avons donc choisi de travailler avec le mime le plus court pour des raisons de coût.

néamine. La paromamine est quant à elle encore moins affine que la néamine vis-à-vis de l'ARN 16S23 (voir figure IV-25).



**néomycine 5** **paromomycine 4** **néamine 3** **paromamine 2**

**Figure IV-135 Quelques ligands aminoglycosidiques de l'ARN 16S23 classés par affinité.**
A gauche, la néomycine est la molécule la plus affine, à droite, la paromamine est la moins affine.

### 2) Principe de la mesure d'affinité : compétition avec la sonde fluorée

Le principe de l'expérience, représenté de façon schématique sur la figure IV-26, est simple : il s'agit d'effectuer une compétition entre la sonde fluorée d'affinité connue pour la cible ARN et un ligand d'affinité inconnue. En suivant par RMN $^{19}F$ le signal du fluor de la sonde au fur et à mesure de la titration, nous pouvons déduire la proportion de sonde libre en fonction de la quantité de ligand compétiteur ajouté et donc l'affinité du ligand compétiteur. En début de titration (situation A de la figure IV-26), on retrouve en RMN $^{19}F$ le doublet caractéristique de la fixation de la sonde sur l'ARN. En fin de titration en revanche, l'intégralité de la sonde a été chassée de sa cible nucléique et on retrouve en RMN $^{19}F$ le singulet correspondant à la sonde fluorée libre en solution (situation B de la figure IV-26).

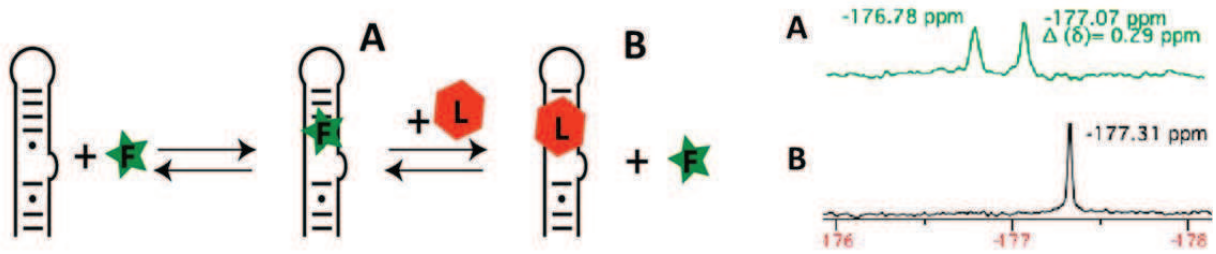



**Figure IV-136 Principe de la mesure par RMN $^{19}$F de l'affinité d'un ligand (en rouge) pour une cible ARN (en noir) par compétition avec une sonde fluorée (en vert)**

Pour calibrer la méthode, nous nous sommes tout d'abord penchés sur des ligands aminoglycosidiques dont l'affinité pour l'ARN 16S23 avait déjà été évaluée par d'autres méthodes expérimentales.

### 3) Utilisation du DACPF comme ligand compétitif

#### a) Etude de l'interaction du DACPF avec l'ARN 16S23

##### *i. Calcul de l'affinité du DACPF pour l'ARN 16S23 par titration en RMN 1D*

Pour interpréter les expériences d'interaction compétitive d'un ligand avec le DACPF, et notamment pour en tirer l'affinité du ligand pour l'ARN 16S23, il importe d'avoir accès en premier lieu à l'affinité du DACPF pour le 16S23. Pour ce faire, nous avons mis en place deux expériences complémentaires de titration par RMN : suivi des signaux de la cible ARN en RMN $^{1}$H et suivi du signal fluor de la sonde en RMN $^{19}$F.

Lors de l'ajout progressif du DACPF sur l'ARN 16S23, on peut suivre en RMN $^{1}$H le changement du déplacement chimique des protons imino de l'ARN (observation de la cible). En traçant la variation de déplacement chimique des protons imino correspondant aux bases $U_{16}$ et $G_{17}$ en fonction de la concentration en DACPF, on peut déduire l'affinité du DACPF pour l'ARN 16S23 à l'aide d'un modèle non-linéaire d'interaction à un site (programme MC-fit développé au laboratoire[359]) (voir figure IV-27). L'analyse des deux courbes obtenues nous donne un Kd de l'ordre de 2mM pour le DACPF envers l'ARN 16S23.

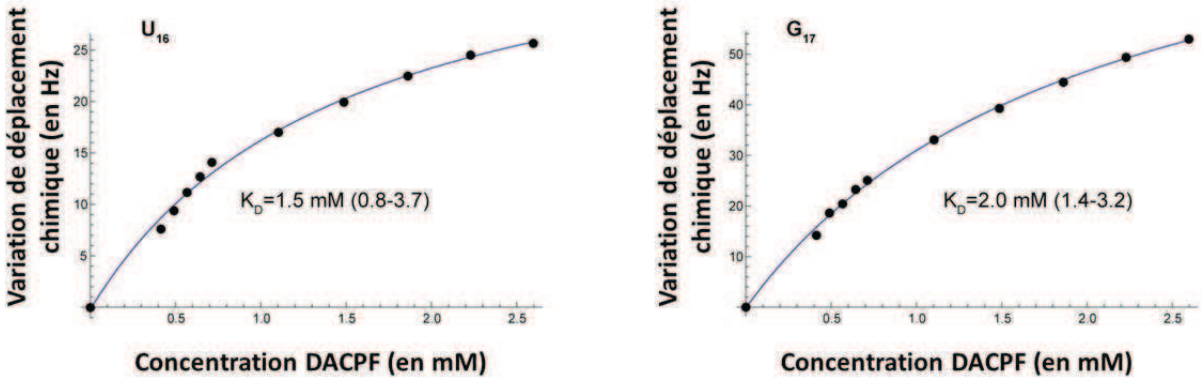



**Figure IV-137 Suivi en RMN $^{1}H$ de la variation des protons imino des nucléotides $U_{16}$ et $G_{17}$ de l'ARN 16S23 en fonction de la concentration en DACPF**

(ARN (300 µM), tampon phosphate 10 mM, 50 mM KCl, pH 6.5, 293K, concentration variable de DACPF)

On peut également observer l'interaction de l'ARN 16S23 avec le DACPF en se plaçant du côté du ligand en suivant en RMN $^{19}F$ les variations du déplacement chimique du DACPF au fur et à mesure de l'augmentation de sa concentration dans l'échantillon RMN. On obtient là encore une courbe de titration, dont on peut extraire, après le même traitement mathématique que précédemment, l'affinité du DACPF envers l'ARN 16S23 (voir figure IV-28).

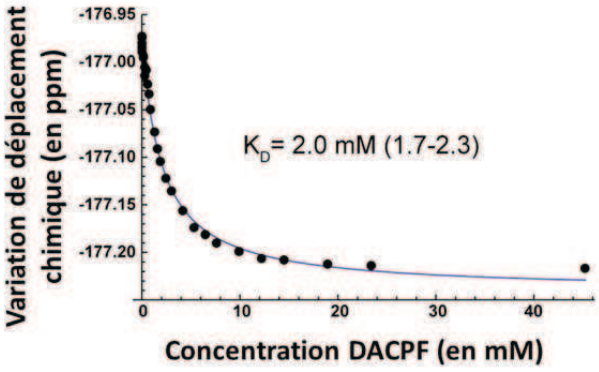


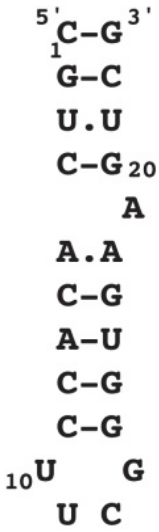

**Figure IV-138 Suivi en RMN $^{19}F$ de la variation du signal du fluor du DACPF en fonction de sa concentration**

(ARN (300 µM), tampon phosphate 10 mM, 50 mM KCl, pH 6.5, 293K, concentration variable de DACPF)

Nous avons été heureux de constater que les deux méthodes de titration donnaient sensiblement la même valeur de Kd pour le DACPF, de l'ordre de 2 mM. Cette affinité est cohérente avec ce qui est publié dans la littérature pour la désoxystreptamine (DOS) qui a un Kd de l'ordre du millimolaire pour l'ARN 16S23[241]. Ces études RMN permettent par ailleurs d'affirmer que l'on est bien en échange rapide par rapport au temps caractéristiques des mesures RMN, à l'instar de ce que l'on observe pour la DOS. Cela signifie que les signaux observés en RMN correspondent à la moyenne des signaux de la forme libre et de la forme liée, pondérée de leur abondance respective. Des ligands d'affinité plus élevée pour le 16S23, comme la néomycine ou la paromomycine, sont eux en échange lent.

### *ii. Etude de l'interaction DACPF/ARN 16S23 en RMN 2D : détermination des sites d'interaction*

Pour affiner notre compréhension de l'interaction du DACPF avec l'ARN 16S23, nous avons décidé d'étudier le complexe formé par RMN 2D $^{1}H/^{1}H$ à l'image de ce qu'avait fait l'équipe de Puglisi pour différents aminoglycosides comme par exemple la paromomycine[44,111,112], la gentamicine[104] ou la desoxystreptamine (DOS)[241]. La comparaison du mode d'interaction du DACPF avec l'ARN 16S23 par rapport à celui de la DOS est particulièrement instructive car ces deux composés sont assez proches structurellement, possédant tous les deux le motif *cis*-1,3-diamino essentiel à la reconnaissance du site A de l'ARN ribosomique 16S. Par ailleurs, dans la majorité des études structurales conduites sur l'interaction des aminoglycosides avec l'ARNr 16S, la DOS joue un rôle de première importance à travers l'interaction par liaison hydrogène de ses 2 amines en position 1 et 3 avec des paires de bases G-U ou G-G du grand sillon de l'ARN 16S. Je me suis donc attelé, avec l'aide de Valery Larue et de Carine Tisné, à l'étude de l'ARN 16S23 en présence de DACPF. Pour ce faire, nous nous sommes appuyés sur une série d'expériences RMN TOCSY et NOESY en faisant varier les temps de mélange (par exemple, nous avons utilisés trois temps de mélange différents pour les NOESY : 300 ms, 400 ms, 600 ms ; voir la partie expérimentale pour plus de détails). Afin d'avoir un spectre de référence, la quasi-totalité des signaux RMN des protons H-5,H-6,H-8,H-1' et H-2'de l'ARN 16S23 seul ont pu être attribués en accord avec ce qui avait été précédemment publié[44,241] (voir figure IV-29) .




| ARN 16S entier | ARN 16S23 | H8 | H6 | H5 | H2 | H1' | H2' |
|---|---|---|---|---|---|---|---|
| C 1404 | C1 | | 7,657 | 5,281 | | 5,566 | 4,760 |
| G 1405 | G2 | 7,529 | | | | 5,700 | - |
| U 1406 | U3 | | 7,521 | 5,371 | | 5,810 | 4,430 |
| C1407 | C4 | | 8,002 | 5,816 | | 5,705 | - |
| A 1408 | A5 | 8,141 | | | 7.834 | 5,987 | 4,72 / 4,48 |
| C 1409 | C6 | | 7,521 | 5,373 | | 5,230 | - |
| A 1410 | A7 | 8,172 | | | 8,074 | 5,869 | 4,310 |
| C 1411 | C8 | | 7,611 | 5,200 | | 5,422 | - |
| C1412 | C9 | | 7,649 | 5,430 | | 5,509 | 4,680 |
| | U10 | | 7,791 | 5,771 | | 5,621 | 4.81 |
| | U11 | | 8,043 | 5,890 | | 6,132 | 4,700 |
| | C12 | | 7,718 | 6,132 | | 5,982 | 4,830 |
| | G13 | 7,894 | | | | 6,006 | 4,480 |
| G 1488 | G14 | 8,338 | | | | 5,847 | 4,620 |
| G 1489 | G15 | 7,263 | | | | 5,835 | 4,570 |
| U 1490 | U16 | | 7,665 | 5,162 | | 5,563 | 4,770 |
| G 1491 | G17 | 7,714 | | | | 5,700 | - |
| A 1492 | A18 | 8,185 | | | 7,900 | 5,960 | 4,72 / 4,48 |
| A 1493 | A19 | 8,362 | | | 8,164 | 6,066 | 4,790 |
| G 1494 | G20 | 7,448 | | | | 5,594 | - |
| U 1495 | U21 | | 7,603 | 5,130 | | 5,482 | 4,680 |
| C 1496 | C22 | | 8,005 | 5,757 | | 5,762 | - |
| G 1497 | G23 | - | | | | - | - |

**Figure IV-139 Tableau d'attribution des signaux RMN des protons H-8/H-6/H-5/H-1'/H-2' de l'ARN 16S23.**

Les correspondances avec la numérotation de l'ARN 16S total sont indiquées à gauche du tableau. (conditions : ARN (1.5 mM), tampon phosphate $KHPO_4$ 100 % $D_2O$ 10 mM, pH 6.5, 50 mM KCl, 278K, temps de mélange 400 ms)

Une fois cette attribution faite, il a été possible de suivre en NOESY et TOCSY le déplacement des pics de l'ARN en fonction de la concentration en DACPF. Nous avons ainsi testé 4 concentrations de DACPF différentes (0 mM (spectre de référence), 2.25 mM, 4.5 mM, 9 mM) à deux températures différentes (278K et 293K) en faisant varier les temps de mélange des différentes expériences RMN (voir partie expérimentale). A chaque fois, l'ensemble de l'attribution des pics a pu être menée à bien. Un exemple des expériences que nous avons menées est présenté dans la figure IV-30. La superposition de spectres TOCSY permet de suivre les déplacements chimiques des pics de la région H5-H6 de l'ARN 16S23 en absence ou en présence de DACPF. Sur cet exemple, on voit bien que certaines bases ont leurs pics RMN modifiés en présence de DACPF : il s'agit des sites d'interaction préférentiels.

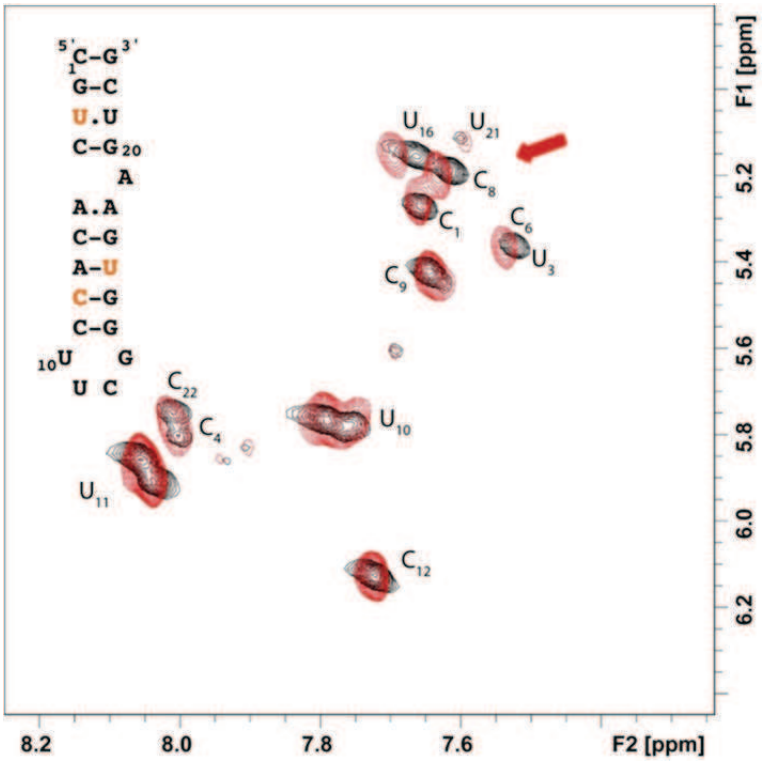



**Figure IV-140 Superposition du spectre TOCSY de l'ARN 16S23 seul (1.5 mM) (en noir) et de l'ARN en présence de 3 équivalents de DACPF (4.5 mM) (en rouge).**

Les bases dont le déplacement chimique est perturbé sont indiquées en rouge sur la structure secondaire de l'ARN.

(conditions : ARN (1.5 mM), tampon phosphate $KHPO_4$ 100 % $D_2O$ 10 mM, pH 6.5, 50 mM KCl, 278K, temps de mélange 45 ms)

En dépit de ce large éventail de conditions expérimentales, nous n'avons pas été en mesure de mettre en évidence en NOESY des pics d'interaction intermoléculaire entre le DACPF et l'ARN 16S23 contrairement à ce qui avait été observé pour la DOS ou d'autres aminoglycosides de meilleure affinité[241]. L'ensemble de ces données nous a cependant permis d'effectuer une analyse de perturbation des déplacements chimiques des pics des différents nucléotides selon la méthode CSP (pour « Chemical Shift Perturbation ») de façon à obtenir une empreinte de la fixation du ligand sur l'ARN. Nous avons pu en déduire que le DACPF avait 2 sites d'interactions équivalents sur l'ARN 16S23 semblables à ceux de la DOS (représentés sur la figure IV-31).

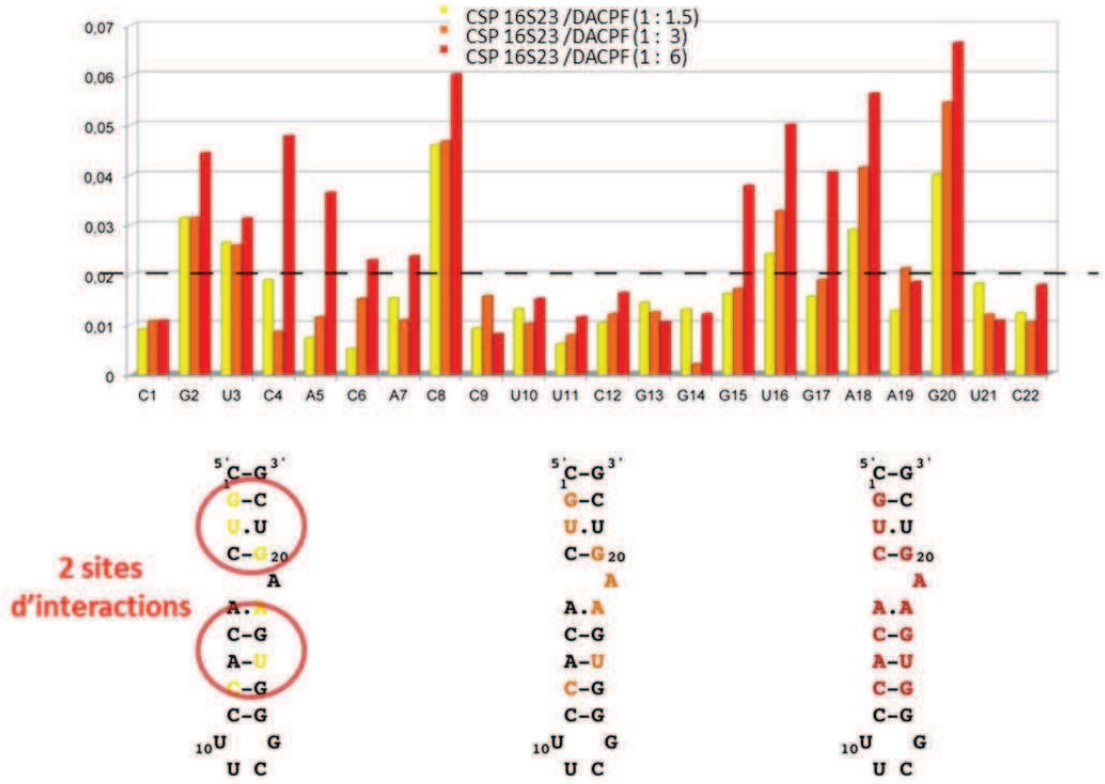



**Figure IV-141 Etude des perturbations des déplacements chimiques des différents nucléotides de l'ARN 16S23 (1.5 mM) en fonction de la concentration en DACPF (2.25 mM (jaune), 4.5 mM (orange) et 9 mM (rouge))**

En conclusion, ces expériences nous ont permis une caractérisation précise de l'interaction de la sonde fluorée DACPF avec sa cible l'ARN 16S23. Le DACPF a une affinité apparente de 2mM pour l'ARN 16S23 et possède deux sites d'interaction. Nous avons ensuite envisagé d'utiliser cette sonde pour caractériser l'affinité de ligands envers l'ARN 16S23 à travers des expériences de compétition.

### b) Résultats des compétitions avec le DACPF

En suivant le principe exposé précédemment (partie IV-III-B-2), nous avons effectué des expériences de compétition entre le DACPF et des ligands connus de l'ARN 16S23 d'affinités variées. L'intégralité de ces expériences étant présentée dans la partie expérimentale, nous nous concentrerons ici sur quelques-unes des expériences les plus significatives.

La figure IV-32 représente une expérience de compétition représentative de l'ensemble des expériences effectuées. Il s'agit des spectres RMN $^{19}F$ obtenus lors de l'augmentation graduelle de la

concentration en néamine dans un échantillon contenant du DACPF et de l'ARN 16S23 (voir partie expérimentale pour plus de détails).

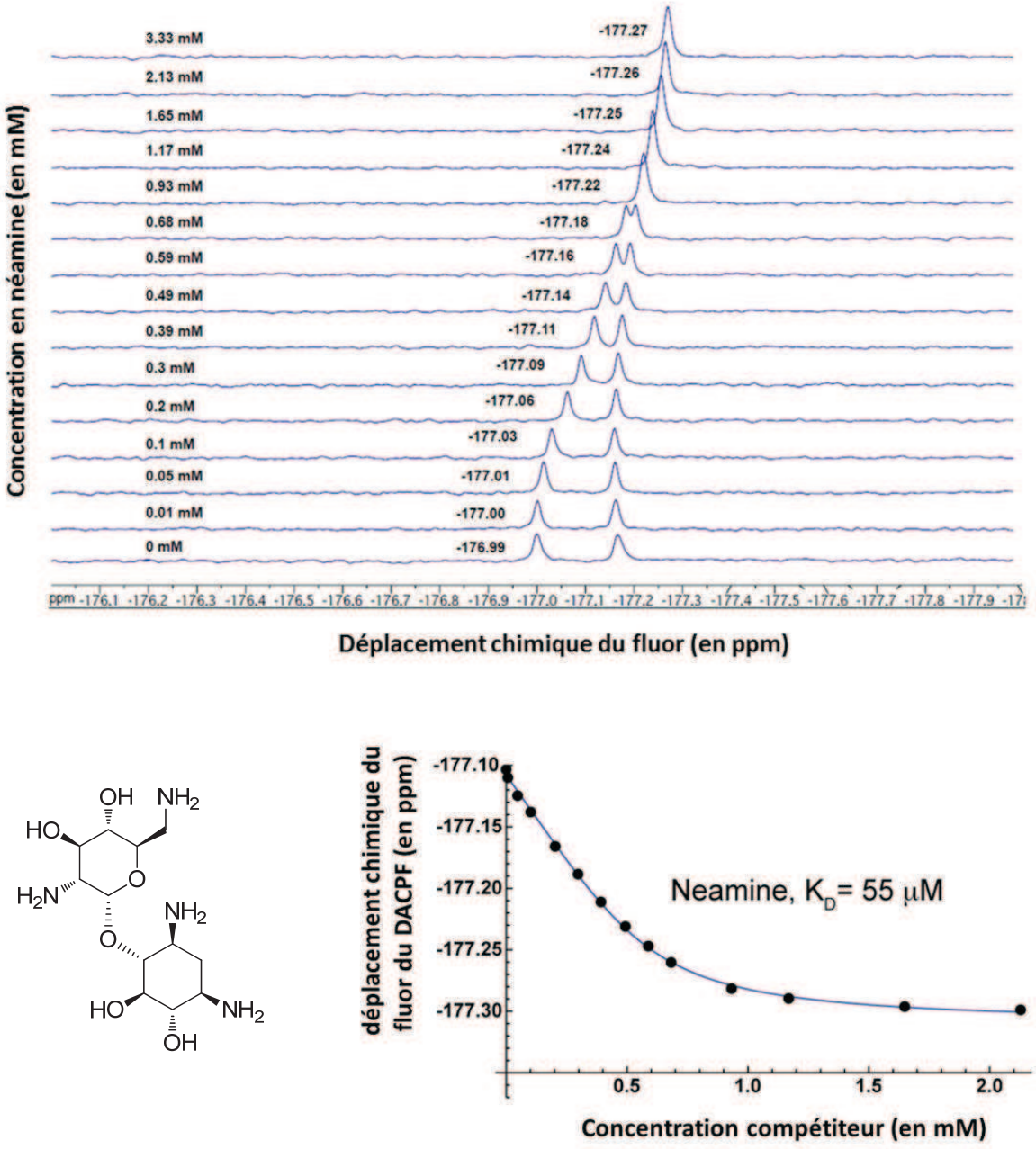



**Figure IV-142 Exemple de la fixation compétitive de la néamine sur le 16S23 en présence de DACPF.** On peut extraire l'affinité de la néamine pour l'ARN 16S23 à partir de la variation en RMN $^{19}F$ du signal de la sonde en fonction de la concentration en néamine.

Au début de l'expérience, il n'y a pas de néamine et on observe donc en RMN $^{19}F$ le doublet caractéristique de l'interaction de la sonde fluorée DACPF avec l'ARN 16S23. Plus on ajoute de néamine, plus le signal de fluor va être modifié pour tendre vers une coalescence du doublet observé en un singulet correspondant au signal de la sonde libre. Ce point constitue la fin de l'expérience de compétition, le ligand compétiteur (ici, la néamine) ayant complètement déplacé la sonde fluorée de l'ARN. En suivant le signal le plus déblindé en RMN $^{19}F$, qui correspond à un des deux énantiomères de la sonde, on peut tracer une courbe de variation du déplacement chimique du fluor en fonction de la concentration en ligand compétiteur. Cette courbe peut être exploitée pour déduire l'affinité du ligand compétiteur à partir de l'affinité de la sonde. On a en effet un système d'équation relativement simple à 5 équations et 5 inconnues qui peut être résolu analytiquement[342,360]. Nous avons utilisé le logiciel Mathematica pour extraire de ces courbes l'affinité du ligand compétiteur mais aussi, en guise de contrôle, le nombre de ses sites d'interaction (voir partie expérimentale pour les détails de l'équation et du modèle développé). A ce stade, il est nécessaire de préciser que la méthode que nous cherchions à développer vise avant tout à établir un classement qualitatif des différents ligands par ordre d'affinité et à obtenir un ordre de grandeur de leur affinité et non pas une valeur précise.

En suivant cette méthodologie, nous avons effectué des expériences de compétition avec 9 ligands différents. Dans cette partie, nous nous concentrerons sur la néomycine, la paromomycine, la néamine et la paromamine mais dont on peut trouver l'intégralité des spectres RMN dans la partie expérimentale[9]. Les expériences de compétition avec la néomycine, la paromamine et la néamine ont été effectuées en duplicata ou en triplicata de façon à tester leur reproductibilité. Ces expériences se sont révélées relativement reproductibles, avec des variations maximales de l'ordre de 50 % sur les valeurs des Kd trouvées. L'essentiel des variations est due à une différence dans le déplacement chimique initial de la sonde fluorée, qui est extrêmement sensible à la qualité de l'échantillon ARN et du tampon. Travailler avec strictement le même tampon de départ (même lot d'ARN, même lot de tampon, etc..) permet de réduire grandement ces variations et donc d'évaluer de façon plus précise les Kd respectifs des différents ligands. En suivant ces précautions, la reproductibilité des expériences est alors très bonne et parfaitement compatible avec notre intention de classer qualitativement les ligands suivant leur affinité.

[9] La ribostamycine, la gentamicine, la kanamycine, la streptomycine et la DOS sont les cinq autres ligands étudiés dont l'analyse n'a cependant pas été poussée jusqu'au bout.

Logiquement, au cours d'une expérience de compétition, plus le ligand compétiteur est affin pour l'ARN 16S23, moins il faut en ajouter pour chasser le DACPF. Ceci est illustré dans la figure IV-33 où l'on perçoit très clairement que la décroissance du signal du fluor la plus rapide est obtenue lors de la compétition avec la néomycine, ligand le plus affin des trois ligands testés.

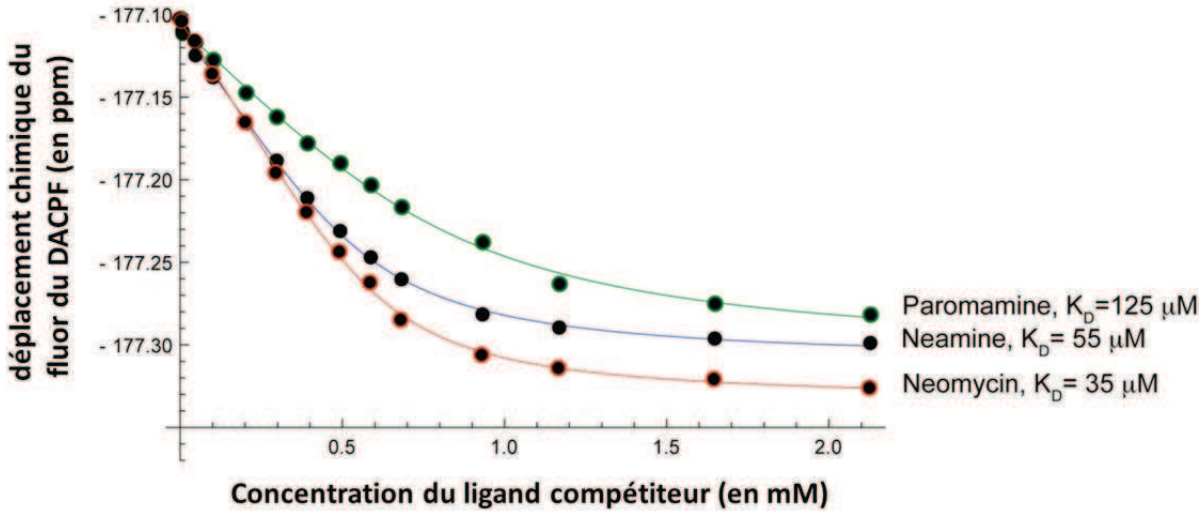



**Figure IV-143 Comparaison de la paromamine, de la néamine et de la néomycine.**

Cette expérience montre que l'on est capable, à l'aide d'une simple expérience de compétition avec une sonde fluorée, de classer qualitativement des ligands en fonction de leur affinité pour la cible ARN. On peut même évaluer leur affinité respective à 125 µM pour la paromamine, 55 µM pour la néamine et 35 µM pour la néomycine, classement qui est en accord avec les données précédemment publiées[117,44]. L'affinité de la paromomycine est évaluée à environ 36 µM soit à peu près la même valeur que pour la néomycine.

En guise de contrôle, nous avons effectué des expériences de compétition pour la néomycine et la paromomycine sur un ARN légèrement plus long que l'ARN 16S23, l'ARN 16S27. C'est cet ARN de 27 nucléotides, possédant 2 paires G-C de plus que l'ARN 16S23 à l'extrémité de la tige-boucle, qui avait été initialement employé par l'équipe de Puglisi comme mime du site A de l'ARN 16S[111](voir partie expérimentale). Nous avons pu vérifier que la néomycine et la paromomycine avaient environ la même affinité pour les deux ARN, ce qui valide à nouveau notre choix de travailler sur l'ARN 16S23, plus court et donc moins cher à faire synthétiser en phase solide.

| Ligand compétiteur | ARN utilisé | Kd (µM) | Kd publiés (µM) (SPR[117]/RMN[44]) (conditions sans sels) |
|---|---|---|---|
| **Néomycine** | 16S23<br><br><br><br>16S27 | 35<br>55<br>47<br>65 | 0.02/0.5 |
| **Paromomycine** | 16S23<br><br><br>16S27 | 36<br>52<br>73 | 0.2/0.5 |
| **Néamine** | 16S23 | 55<br>67 | 8/100 |
| **Paromamine** | 16S23 | 125<br>207 | >100/- |

**Tableau IV-1 Tableau récapitulatif des résultats des expériences de compétition avec le DACPF**

### c) Interprétation critique des résultats

De façon très satisfaisante, l'emploi du DACPF comme sonde reportrice permet de suivre très simplement en RMN $^{19}$F les interactions de l'ARN 16S23 avec différents ligands. Les expériences de titration que nous avons effectuées ont permis d'obtenir un ordre de grandeur cohérent des différents ligands étudiés ainsi qu'un classement qualitatif des ligands par affinité en accord avec les données publiées dans la littérature.

Notre méthode a cependant deux inconvénients qui en limitent l'usage. Tout d'abord, il est difficile de parvenir à chasser complétement la sonde fluorée de l'ARN lorsqu'on étudie des ligands d'affinité proches de celle du DACPF comme la DOS. On est alors contraint d'augmenter considérablement la concentration en ligand compétiteur, ce qui a tendance à faire précipiter l'ARN

et donc à fausser les mesures de déplacements chimiques (voir partie expérimentale). Le DACPF ne permet par conséquent d'étudier que des ligands dont le Kd pour l'ARN 16S23 est inférieur ou égal au millimolaire. Il faut par ailleurs que les ligands soient suffisamment solubles dans l'eau pour pouvoir en préparer des solutions concentrées à ajouter petit à petit dans l'échantillon RMN sans trop le diluer. En pratique, ces limitations ne sont pas vraiment pas gênantes.

Le deuxième inconvénient de l'utilisation du DACPF pour les expériences de compétition est qu'il n'est pas possible de classer les bons ligands (Kd < 40 µM) entre eux. En particulier, les affinités trouvées pour la néomycine et la paromomycine sont similaires alors que la néomycine est décrite dans la littérature comme étant plus affine(voir par exemple la référence[117]). On s'attendrait par ailleurs à une différence plus marquée entre l'affinité de la néamine et celle de la néomycine. Ceci peut s'expliquer par une trop grande différence entre l'affinité de la sonde pour l'ARN (environ 2 mM) et celle de ces ligands (de l'ordre du µM). La sonde n'est alors pas assez sensible pour permettre une mesure précise d'affinités aussi élevées. Diverses études théoriques suggèrent d'ailleurs une limite de sensibilité aux alentours de 2 ordres de grandeur autour de l'affinité de la sonde[361,362]. Dans notre cas, les ligands d'affinité supérieure à environ 40 µM ne peuvent ainsi pas être classés avec précision.

Pour dépasser cette limitation, nous avons souhaité développer une sonde d'affinité plus élevée pour l'ARN 16S23, qui permettrait de classer plus précisément les ligands de haute affinité. Pour concevoir une telle sonde, nous nous sommes directement inspirés de la structure des aminoglycosides et plus particulièrement de la paromamine dont nous avons conçu un analogue fluoré. La synthèse de la paromamine fluorée (nommée paroF dans la suite du manuscrit) et son utilisation dans des expériences de compétition sont détaillées dans la partie suivante.

### 4) Développement d'une sonde fluorée plus sensible :la paromamine fluorée

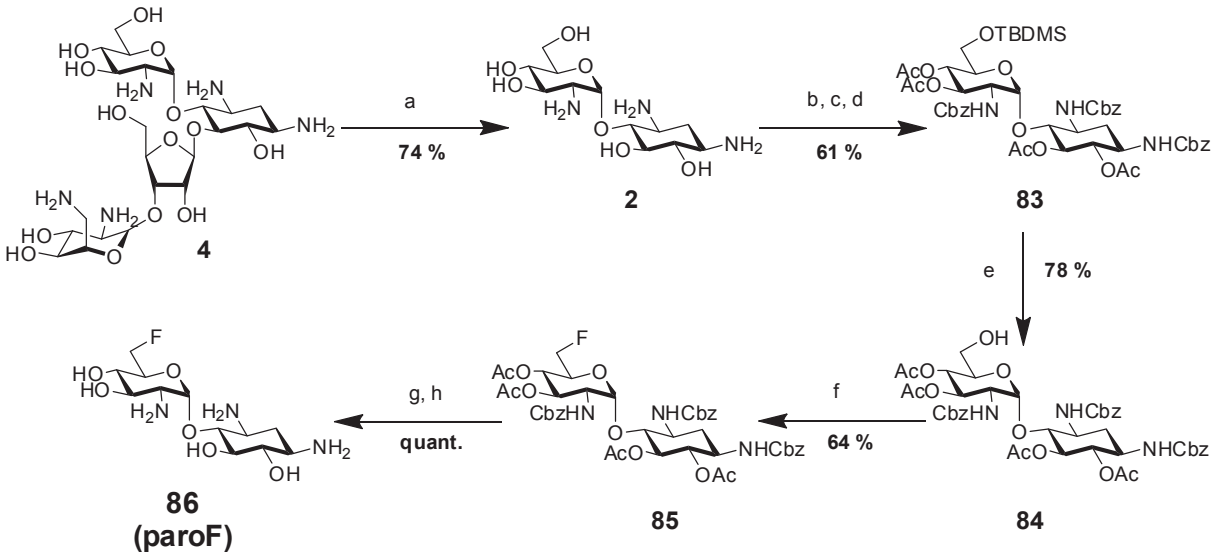
#### a) Synthèse de la paromamine fluorée (paroF)

L'ensemble des études structurales des aminoglycosides de la famille de la néomycine en interaction avec l'ARN 16S montraient la proximité de la position 6' du cycle I des aminoglycosides avec le grand sillon de l'ARN[88]. Pour concevoir une sonde fluorée de cet ARN, il nous a donc semblé intéressant d'essayer de fluorer sélectivement la position 6' de la paromamine (voir figure IV-34).

Figure IV-144 La sonde paroF (à gauche), un analogue fluoré de la paromamine (à droite)

La stratégie de synthèse que nous avons mise en place est relativement simple (figure IV-35). Elle consiste tout d'abord à obtenir la paromamine **2** à partir de la dégradation de la paromomycine **4** en milieu acide dans le méthanol. Nous nous sommes inspirés du protocole publié par Spencer[278] en l'adaptant à la production de paromamine à l'échelle de la dizaine de grammes. L'étape suivante consiste à protéger les amines par des groupements carbamates de benzyl selon le protocole décrit par Tsumura[363] et repris par Wong[322]. A ce stade, le produit obtenu est très peu soluble dans une grande variété de solvants ce qui complique sa purification. A l'aide d'un suivi de la réaction par HPLC, nous avons optimisé les conditions réactionnelles de façon à obtenir un produit assez propre que nous avons purifié par précipitation et triturations successives (voir partie expérimentale). Ce produit peut alors être protégé sélectivement en position 6' par réaction du TBDMSCl sur l'alcool primaire dans la DMF. Pour éviter la protection des alcools secondaires malgré le caractère très encombrant du groupement silylé, la réaction est conduite en condition cinétique (1h à -5°C, excès de TBDMSCl) et est suivie par HPLC. Lorsque la HPLC indique une conversion totale du produit de départ, la réaction est stoppée par ajout de méthanol. Le produit obtenu n'étant toujours pas très soluble dans la plupart des solvants usuels, le brut réactionnel est directement engagé dans une réaction de peracétylation par traitement à l'anhydride acétique dans la pyridine en présence de DMAP pour donner le produit **83**. Cette réaction s'est avérée plus difficile que prévu en raison de la faible réactivité de l'alcool en position 5 du cycle DOS. Certaines stratégies de synthèse (par exemple celle développée par Wong[322]) utilisent d'ailleurs l'encombrement stérique de cet alcool pour sélectivement peracetyler tous les autres alcools de la molécule avant de travailler sur l'alcool en position 5 laissé libre. Là encore, la HPLC s'est révélée extrêmement utile à l'optimisation des conditions réactionnelles de la peracétylation. Le produit **83**, entièrement protégé, est soluble dans le dichlorométhane et peut être facilement purifié par flash chromatographie sur gel de silice.

L'étape suivante de déprotection du groupement protecteur silylé pour donner le composé **84** a tout d'abord été tentée à l'aide de fluorure de tétrabutylammonium (TBAF) mais ces conditions classiques conduisent à un mélange de produits résultant de la migration des groupements acétate sur l'alcool primaire déprotégé. La déprotection est donc effectuée au HF.Pyridine dans les conditions douces et moins basiques développées par Trost[364], puis adaptées par Carreira[365] qui était confronté au même problème de migration de groupements acétates. La suite de la synthèse n'a pas posé de problème majeur. La fluoration du composé **84** est effectuée en présence de trifluorure de diethylaminosulfure (DAST) selon des conditions classiques[366] dans le dichloroethane anhydre à 0°C pour donner le composé **85**. Les alcools secondaires sont ensuite déprotégés au méthanolate de sodium puis le produit obtenu est hydrogénolysé dans un mélange méthanol/dioxanne/eau en présence de palladium pour donner la paromamine fluorée **paroF** en 8 étapes avec un rendement global de 22.5 %. La paromamine fluorée obtenue est enantiomériquement pure car d'une part le substrat initial, la paromomycine, est un produit naturel énantiomériquement pur et d'autre part la séquence réactionnelle proposée ne conduit pas à une racémisation ou une perte d'excès énantiomérique (voir partie expérimentale).




a) HCl, MeOH, reflux, 24h; b) CbzCl, $Na_2CO_3$, toluene/eau/acetone, -5°C jusqu'à T.A., 4h;
c) TBDMSCl, DMAP, $Et_3N$, DMF, -5°C, 1h; d) $Ac_2O$, DMAP, pyridine, T.A., 1 d.; e) HF-pyridine, pyridine, THF, T.A., 4h;
f) DAST, dichloroethane sec, -5°C jusqu'à T.A., 5 h; g) NaOMe, MeOH, THF, dioxane, T.A., 3h;
h) $H_2$, Pd/C, dioxane, MeOH, eau, T.A., 2j.

**Figure IV-145 Stratégie de synthèse de la paromamine fluorée paroF**

### b) Etude de l'interaction de paroF avec l'ARN 16S23 et optimisation des conditions expérimentales pour les expériences de compétition.

Notre approche a été la même que celle employée plus tôt pour le DACPF. Pour effectuer des compétitions avec la paroF et en tirer l'affinité des ligands compétiteurs pour l'ARN 16S23, il importe tout d'abord d'évaluer l'affinité de la sonde fluorée pour l'ARN. Nous avons donc procédé à des titrations en RMN $^{1}H$ à différentes concentrations en sels selon le même protocole que précédemment avec le DACPF. La variation des protons iminos des bases $U_{16}$ et $G_{17}$ de l'ARN 16S23 au fur et à mesure de l'augmentation de la concentration en paroF permet d'avoir accès au Kd de la sonde fluorée pour l'ARN (figure IV-36, voir partie expérimentale pour plus de détails).

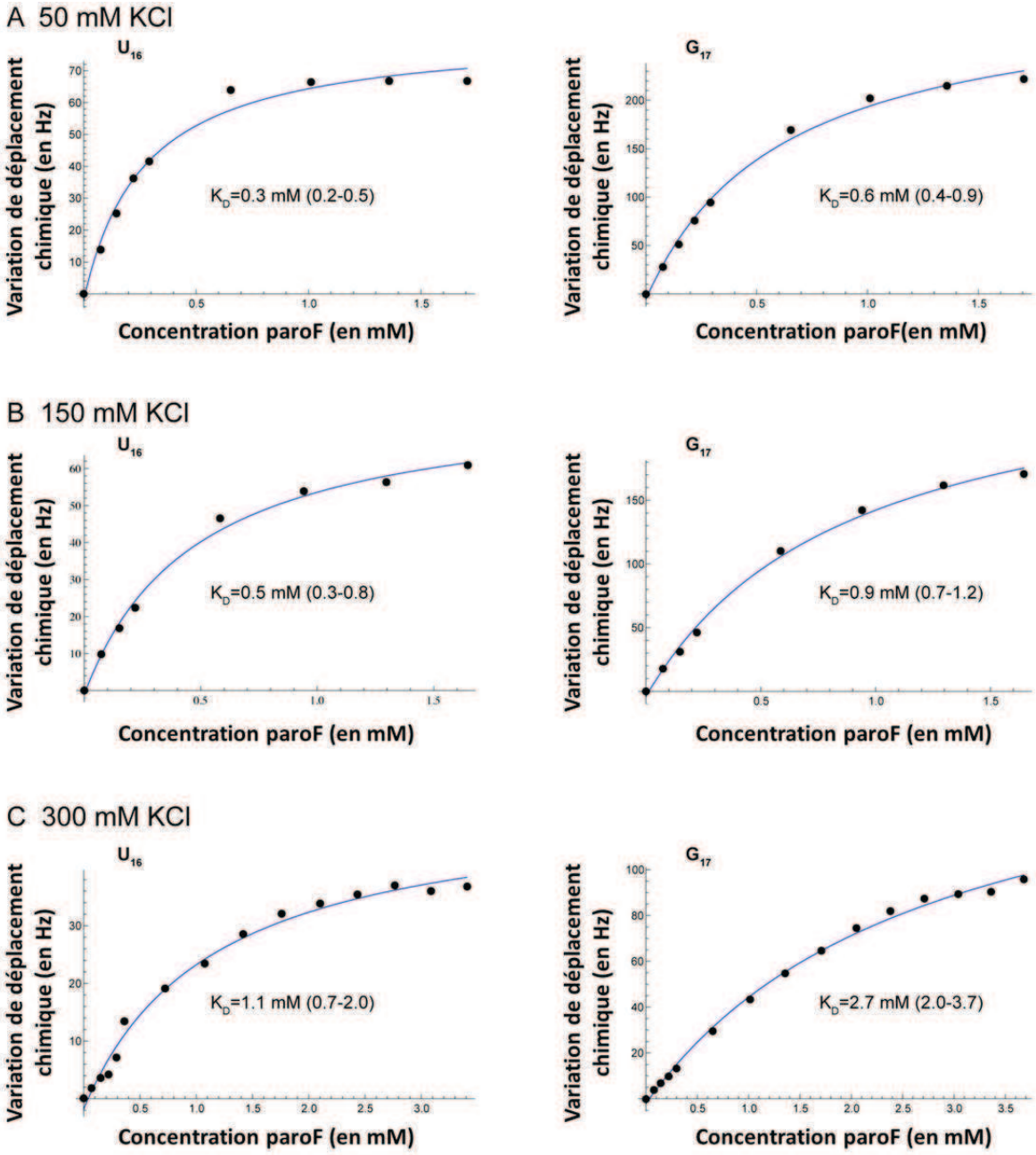



**Figure IV-146 Mesure de l'affinité de la paromamine fluorée pour l'ARN 16S23 par RMN $^{1}$H via la variation du déplacement chimique des protons imino des bases U16 et G17 lors d'addition croissante de paromamine fluorée.**

Comme prévu, en se plaçant dans les mêmes conditions expérimentales que pour les compétions avec le DACPF ([KCl] = 50 mM), l'affinité de la paroF calculée par RMN $^{1}$H est un peu plus de 6 fois plus forte que le DACPF (Kd de 300 µM contre 2 mM). La sonde paroF est donc plus affine que le DACPF pour l'ARN 16S23, ce qui laisse espérer une plus grande sensibilité. Malheureusement, en RMN $^{19}$F, cette plus grande affinité se traduit par un élargissement du signal du fluor lors de

l'interaction de la sonde avec l'ARN qui rend impossible son observation. Ceci est lié à la meilleure affinité de la sonde, qui passe alors en mode d'échange intermédiaire par rapport au temps caractéristique de la RMN. Il est néanmoins possible de retrouver un signal RMN fin en augmentant la concentration en sels de l'échantillon ([KCl] augmenté à 300 mM) (voir figure IV-37). Cette augmentation a deux conséquences intéressantes. Tout d'abord, augmenter la concentration en sels revient à diminuer l'affinité apparente de la sonde pour sa cible ARN. En effet, comme pour la majorité des aminoglycosides, une large part de l'affinité de cette molécule provient des interactions électrostatiques de ses amines chargées positivement avec les phosphates de l'ARN chargés négativement à pH physiologique. En conséquence, une multiplication de la concentration en sels par 6 conduit à une augmentation du Kd par un facteur 3 environ, de 300 µM à 1.1 mM (cf titration en RMN $^{1}$H). On repasse donc dans des valeurs d'affinité compatibles avec un échange rapide et donc un signal fin en RMN $^{19}$F. Par ailleurs, si l'on augmente la concentration en sels, on écrante davantage les interactions électrostatiques qui constituent souvent une large partie des interactions non spécifiques. Ces conditions expérimentales permettent ainsi de se concentrer sur les ligands affins et spécifiques de la cible nucléique d'intérêt (voir la publication de Pilch[98] qui étudie la dépendance du Kd de certains aminoglycosides en fonction de la concentration en sels).

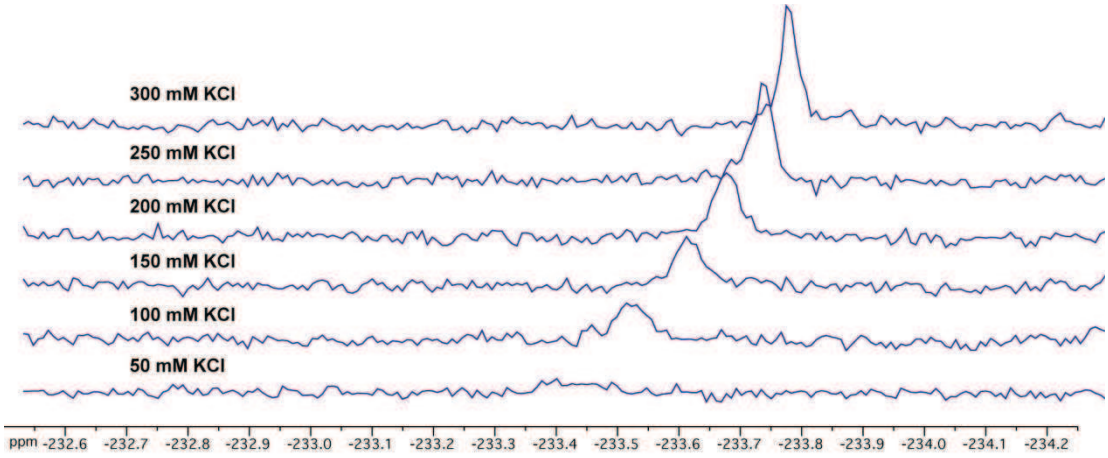



**Figure IV-147 Effet de la concentration en sels sur la finesse du pic du fluor de paroF interagissant avec l'ARN 16S23**

### c) Résultats de compétitions avec la paroF

Une fois les conditions expérimentales optimisées pour les expériences de compétition avec la paroF, plus sensible que le DACPF, nous nous sommes de nouveau penchés sur l'étude des 3 ligands

de forte affinité pour l'ARN 16S23, la néomycine, la paromomycine et la néamine, que nous n'avions pas réussi à classer de façon satisfaisante avec le DACPF.

Les résultats des expériences sont présentés dans la figure IV-38. La sonde paroF permet de classer ces trois ligands par ordre d'affinité. Cette fois, et contrairement à ce que l'on observe avec le DACPF, il y a une différence nette entre le comportement de la néomycine et celui de la paromomycine. La néamine est quant à elle très clairement le ligand le moins affin. On peut évaluer l'affinité de ces trois molécules à, respectivement, 250 µM, 90 µM et 60 µM. De façon intéressante, l'emploi d'une sonde fluorée plus affine et l'augmentation de la concentration en sels permet donc d'augmenter les différences d'affinité apparente entre bons ligands et donc de les classer plus précisément (voir le travail de Pilch qui étudie la variation de l'affinité de certains aminoglycosides en fonction de la concentration en sels[98]). Les inconvénients d'une telle approche sont qu'une sonde plus affine peut potentiellement modifier les propriétés de sa cible. Ce n'est pas ce qui est observé avec l'ARN 16S23 à une concentration de 300 mM en KCl. Par contre, l'interaction de paroF avec l'aptamère à la néomycine décrit précédemment à une concentration de 50 mM en KCl conduit effectivement à une modification de la structure de l'aptamère (voir partie perspectives IV-V-D). D'autre part, les fortes concentrations en sels utilisées pour classer les ligands très affins entre eux sont relativement éloignées des conditions physiologiques classiques ce qui peut conduire à un certain nombre de biais expérimentaux.

On peut remarquer par ailleurs que les spectres fluor ne comportent pas de doublet comme avec le DACPF mais un singulet qui se décale vers les champs forts (blindage) au fur et à mesure de la titration, jusqu'à atteindre le signal de la sonde libre. Ceci est dû au fait que la sonde paroF est enantiomériquement pure.

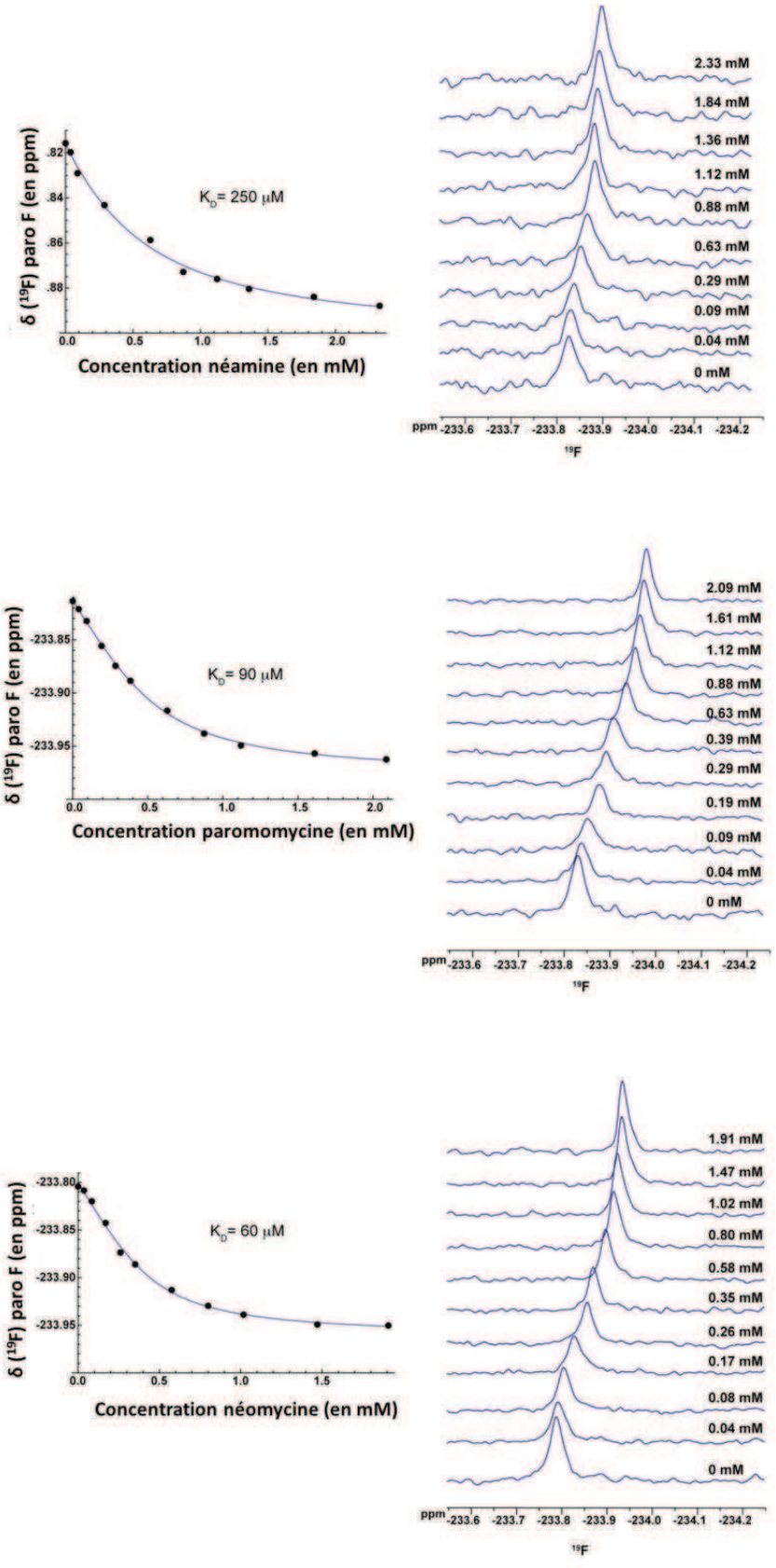



**Figure IV-148 Compétition avec la paroF : les ligands de forte affinité peuvent être classés entre eux.**

En guise de contrôle, nous avons voulu vérifier si on trouvait les mêmes ordres de grandeurs pour l'affinité de ces 3 ligands en effectuant directement une titration en RMN $^{1}H$. Le suivi des protons imino de la base $U_{19}$ de l'ARN 16S23 permet de déterminer une affinité de l'ordre de 450 µM pour la néamine dans les conditions expérimentales de la compétition avec la paroF (Tampon phosphate 10 mM, 300 mM KCl, pH 6.5, 293K). Ce résultat est en accord convenable avec les 250 µM trouvés en RMN $^{19}F$. Par contre, il n'a pas été possible de mesurer l'affinité de la néomycine et de la paromomycine en suivant ce protocole en RMN $^{1}H$ car ces molécules ont une trop bonne affinité pour l'ARN et sont donc en échange lent. Ceci illustre un des autres avantages de l'utilisation de sondes fluorées pour mesurer de façon indirecte l'affinité de ligands très affins.

# IV - Conclusion

Dans ce chapitre, nous avons exposé l'intérêt d'utiliser des analogues fluorés de fragments aminoglycosidiques comme sondes pour étudier les ARN en RMN $^{19}$F. Nous en avons illustré certains des avantages par rapport aux autres méthodes expérimentales : simplicité d'utilisation, sensibilité de la RMN $^{19}$F, possibilité d'étudier une large variété de ligands et d'ARNs non marqués et de taille illimitée…

Nous avons montré qu'un analogue fluoré de la DOS, le DACPF, permettait d'obtenir, à travers son interaction avec diverses cibles ribonucléiques, une signature topologique de celles-ci *via* le dédoublement de son signal du fluor en RMN. Cette propriété a été exploitée pour suivre de façon très simple en RMN $^{19}$F des phénomènes complexes de changement conformationnel d'ARN lié, soit à un événement de dénaturation/renaturation, soit à la capture conformationelle d'un aptamère par son ligand.

Cette même sonde fluorée a aussi été mise à profit pour caractériser l'interaction de petites molécules avec un analogue du site de décodage de l'ARN 16S, l'ARN 16S23. Nous avons développé une méthodologie combinant des expériences de fixation compétitive suivie en RMN $^{19}$F et modélisation des courbes obtenues pour évaluer l'affinité des différents ligands et les classer de façon qualitative. Dans le but d'évaluer de façon plus précise l'affinité de très bons ligands de l'ARN 16S23, nous avons développé une sonde plus sensible que le DACPF, la paromamine fluorée paroF. Cette sonde nous a permis de caractériser l'interaction de ligands de très forte affinité pour l'ARN 16S23 comme la néomycine et la paromomycine. Nous disposons ainsi de deux sondes fluorées pour la mise en place de criblage de nouveaux ligands d'ARN: le DACPF, utile dans les premières phases de criblage pour identifier les « hits » par exemple lors d'une approche de synthèse par fragment, et la paroF, sonde plus sensible, à utiliser lors d'études de type structure/affinité visant à améliorer un fragment déjà connu. Nous travaillons par ailleurs à développer d'autres sondes fluorées avec des propriétés différentes (voir perspectives et partie expérimentale).

De façon très intéressante, ces sondes fluorées, directement inspirées des aminoglycosides qui sont souvent considérés comme des « ligands universels d'ARN », semblent se fixer à de nombreuses cibles ARN ce qui rend possible leur utilisation pour l'étude de systèmes biologiques très variés. Leur simplicité d'utilisation, leur généralité potentielle ainsi que la bioorthogonalité de l'atome du fluor et sa sensibilité en RMN $^{19}$F feront peut-être de ces sondes des outils de première importance dans l'étude des ARN et le développement de nouveaux ligands.

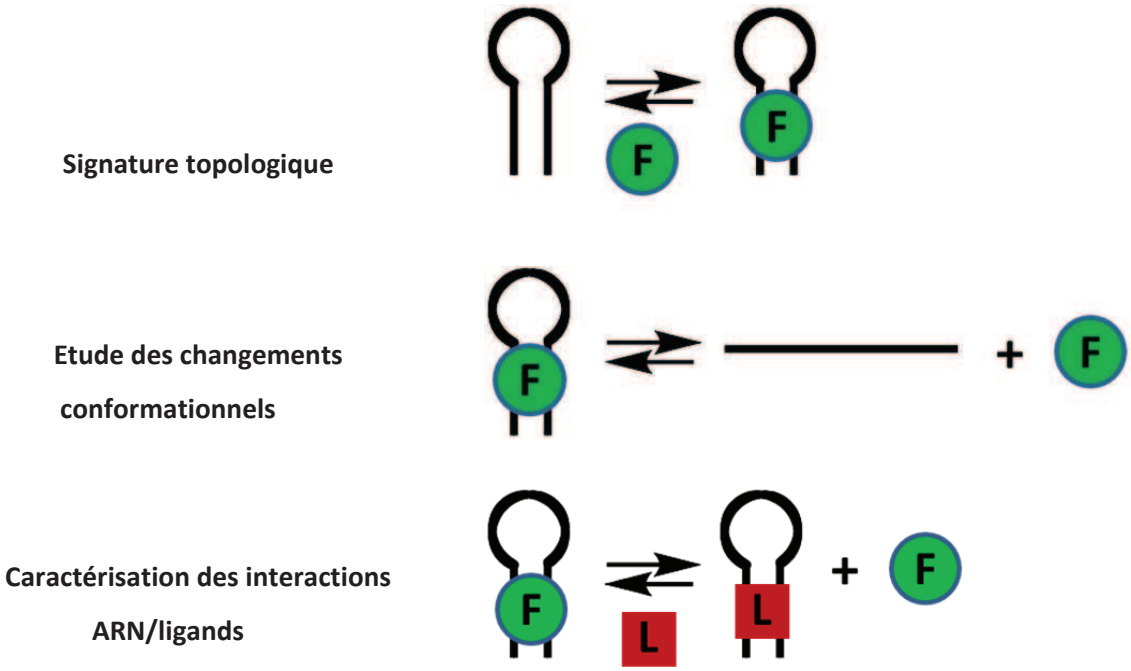



Figure IV-149 L'utilisation de sondes fluorées dans l'étude des ARN par RMN $^{19}$F offre de nombreuses possibilités

# V - Perspectives

## *A - Criblage en RMN $^{19}$F de nouveaux ligands d'ARN*

Comme nous l'avons évoqué dans la conclusion, les sondes DACPF et paroF que nous avons développées au laboratoire sont des outils potentiellement très utiles dans une approche de synthèse de ligand par fragment. Ces outils vont donc tout naturellement rejoindre l'arsenal dont nous disposons au laboratoire pour tester et évaluer l'affinité des différents fragments que nous développons pour se lier spécifiquement à diverses structures ribonucléiques. En particulier, de nombreux ligands de l'ARNt$^{Lys}_3$ ont été développés au laboratoire[62,325,367] et les méthodologies développées en RMN $^{19}$F vont être utilisées pour plus précisément caractériser leur interaction avec leur cible ARN.

## *B - Etude en RMN $^{19}$F de la dynamique de systèmes biologiques*

Les propriétés du DACPF et de la paroF permettent de suivre « en temps réel » des événements de changements de conformation d'ARN variés. A condition que les temps de mesure soient inférieurs aux temps caractéristiques des phénomènes étudiés, on peut ainsi imaginer les utiliser pour caractériser des cinétiques de système biologiques d'intérêt.

En guise d'exemple, nous sommes en train de développer au laboratoire le suivi par RMN $^{19}$F de l'initiation de la transcription inverse du VIH par hybridation entre le site de liaison PBS (Primer Binding Site) de l'ARN viral et l'ARNt$^{Lys}_3$ assistée par la nucléocapside. Utiliser les sondes fluorées pour mettre en évidence ce phénomène représenterait en effet une approche alternative aux approches de RMN $^{1}$H actuellement développées au laboratoire[59,60,61].

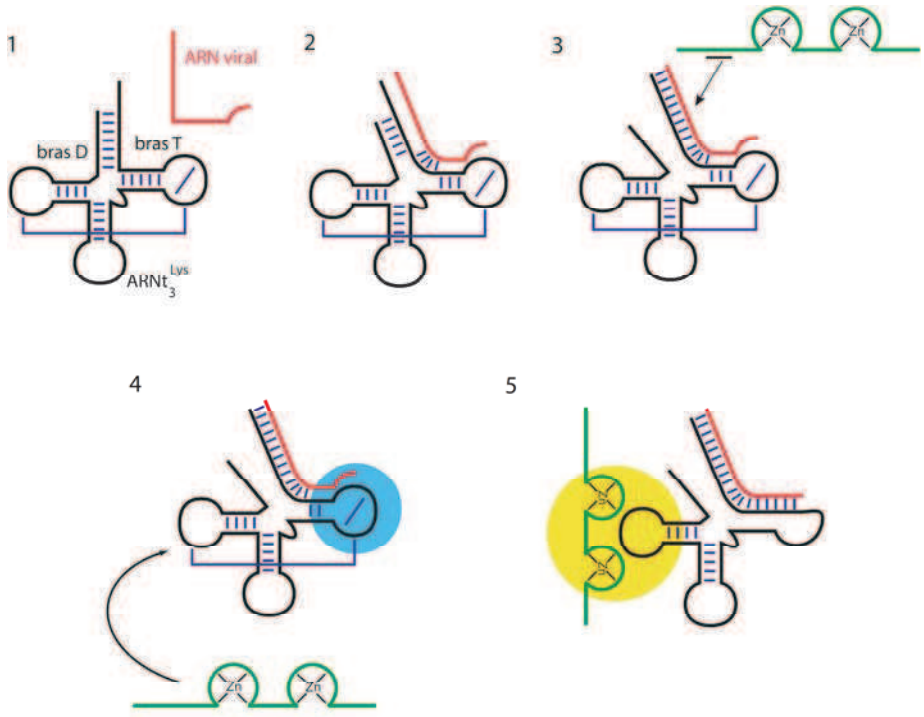



**Figure IV-150 L'initiation de la transcription inverse du VIH par hybridation entre le PBS de l'ARN viral (en rouge) et l' ARNt$^{Lys}_3$ (en noir), assistée par la nucléocapside (en vert)**[61]

Nous avons par ailleurs initié des contacts préliminaires avec le docteur T. Bizebard de l'IBPC qui étudie le fonctionnement d'hélicases d'ARN issues d'extrémophiles psychrophiles, c'est-à-dire adaptés à des environnements très froids[368]. Son équipe est particulièrement intéressée dans l'étude des changements conformationnels de doubles brins d'ARN en présence de ces enzymes et aux grandeurs cinétiques ou thermodynamiques associées (énergie d'activation, entropie, enthalpie, …) La méthodologie qu'ils utilisent pour étudier ces changements conformationnels, le transfert de fluorescence (FRET), a l'inconvénient de nécessiter un double marquage de la cible de l'hélicase avec un groupe émetteur sur un brin de l'ARN et un groupe récepteur sur l'autre brin (voir figure IV-41). Il n'est pas exclu *a priori* que ces modifications changent le comportement du système biologique d'intérêt. Le suivi de ce genre de phénomène en RMN $^{19}$F à l'aide des sondes fluorées que nous avons développées au laboratoire permettrait de dépasser cette limitation. La mise en place de cette expérience semble simple à mettre en œuvre, toutes les conditions expérimentales nécessaires au bon fonctionnement des hélicases étant compatibles avec la RMN $^{19}$F, et permettrait de comparer les données obtenues avec les deux méthodes. On peut envisager dans un second temps de travailler sur des ARN plus complexes.

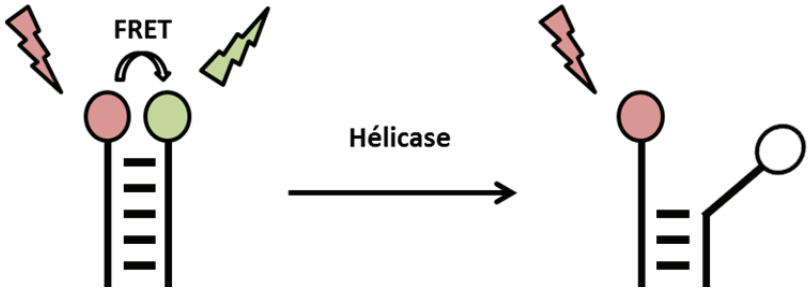



**Figure IV-151 Méthodologie de suivi de l'activité d'une hélicase à ARN par FRET**[368]
L'utilisation d'une sonde fluorée permettrait d'obtenir le même type d'information en s'affranchissant de la nécessité de marquer l'ARN.

## *C - Paromamine fluorée et aptamère, vers un système artificiel de régulation traductionnelle ?*

Au cours du développement de la paromamine fluorée paroF, nous avons eu l'opportunité de rencontrer et d'initier une collaboration avec le docteur Beatrix Suess, une des chercheurs à l'origine de la conception de l'aptamère à la néomycine[358] utilisé pour le suivi de changements conformationels d'ARN en RMN $^{19}$F dans la partie IV III A. Un des objectif de son équipe de recherche est de développer des systèmes de régulation artificiels basés sur l'utilisation d'aptamères en tant que riboswitches[357369]. En particulier, l'aptamère à la néomycine que nous avons utilisé a été intégré dans un système de régulation traductionnel artificiel chez la levure *S. cerevisiae*[36] (voir figure IV-42). La présence de néomycine permet de diminuer l'efficacité de la traduction d'un gène reporter, ici la protéine fluorescente GFP.

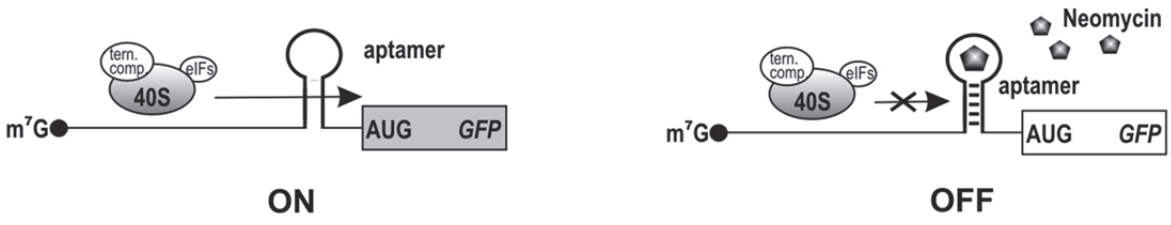



**Figure IV-152 Utilisation d'un aptamère à la néomycine en tant que régulateur traductionnel.**
En présence de néomycine (à droite), l'aptamère adopte une conformation rigide qui inhibe l'initiation de la traduction et donc l'expression du gène reporter GFP[36].

Dans le but d'améliorer cette inhibition, le groupe de Suess est constamment à la recherche de nouveaux aptamères plus efficaces mais aussi de nouvelles molécules effectrices capables d'être

reconnues spécifiquement par ces aptamères. En particulier, l'utilisation de la néomycine comme molécule effectrice présente un certain nombre d'inconvénients liés à sa non-spécificité et à sa toxicité (interaction avec le ribosome notamment) et à la difficulté de cette molécule à passer les membranes cytoplasmiques en raison de son grand nombre d'amines chargées au pH physiologique.

La conception de sondes fluorées comme outil d'étude des ARN par RMN $^{19}$F nous amène à développer une variété de molécules reconnaissant diverses cibles ARN. Une des caractéristiques essentielles d'une sonde est de ne pas trop perturber le système biologique étudié. Ainsi, l'affinité modeste du DACPF pour l'aptamère à la néomycine nous a permis de vérifier que son interaction ne modifiait pas l'équilibre conformationnel de l'ARN qui a donc pu être étudié en RMN $^{19}$F avec le miminum de biais expérimentaux. La paromamine fluorée quant à elle s'est révélée être un trop bon ligand de l'aptamère pour constituer une sonde convenable. Lors de la fixation de la paromamine fluorée sur l'aptamère, on peut en effet voir en RMN $^{1}$H la modification des signaux des protons imino de l'ARN qui résulte d'un réarrangement structural de l'ARN (voir figure IV-43).

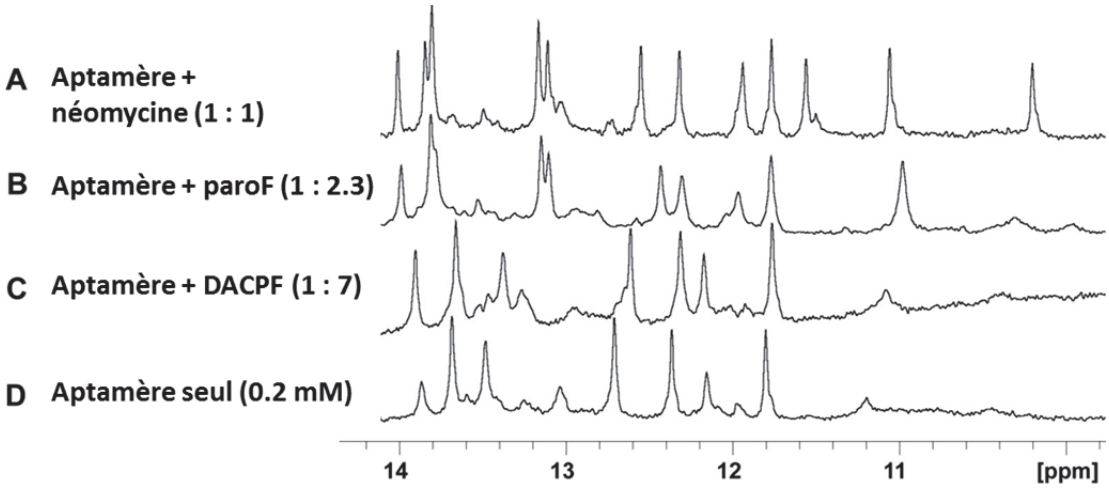


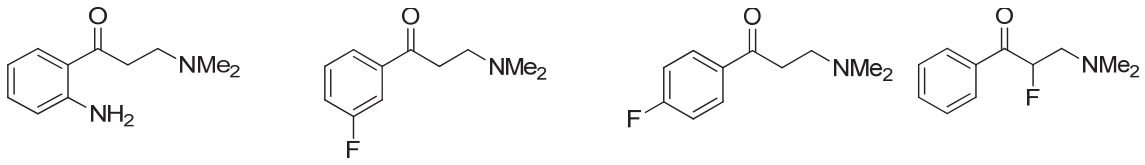

**Figure IV-153 Suivi des protons de l'aptamère à la néomycine en RMN $^{1}$H.**
Alors que l'interaction avec le DACPF n'induit pas de modification des signaux RMN, l'interaction avec paroF se traduit par une modification du spectre semblable à celui causé par la présence de néomycine.

Si ce réarrangement de l'aptamère induit par l'interaction avec la paromamine fluorée condamne l'utilisation de cette dernière comme sonde structurale, il constitue une piste intéressante pour l'emploi de cette molécule comme molécule effectrice dans un système de régulation basé sur l'aptamère à la néomycine à l'instar de ce qui est présenté dans la figure IV-42. Utiliser la paromamine fluorée plutôt que la néomycine présenterait plusieurs avantages : tout d'abord, la paromamine fluorée est un beaucoup moins bon ligand du ribosome que la néomycine, ce qui

diminue d'autant sa toxicité potentielle. D'autre part, la paromamine fluorée est beaucoup moins chargée que la néomycine (3 amines contre 6), ce qui aurait tendance à faciliter son passage dans le cytoplasme. Pour tester cette possibilité, un premier lot de paromamine fluorée a été envoyé à l'équipe de Beatrix Suess et les tests d'inhibition de la traduction de la GFP chez *S. cerevisiae* selon un système précédemment publié[36] sont en cours.

## *D - Développement de nouvelles sondes fluorées : analogues fluorés de la kynuramine*

Encouragés par les succès du développement des sondes fluorées DACPF et paroF, nous avons souhaité développer d'autres sondes fluorées permettant l'étude des ARN en RMN $^{19}F$. En particulier, il nous semblait très intéressant de développer des sondes permettant l'étude simultanée de différentes régions d'un même ARN. En guise d'exemple, nous sommes revenus à l'une des cibles ARN les plus étudiées au laboratoire, l'ARNt$^{Lys}_3$. Le criblage initial de fragments se liant à cette cible avait permis d'obtenir des ligands spécifiques du bras T de l'ARNt$^{Lys}_3$, comme le DACP qui a conduit au développement du DACPF, mais aussi des ligands spécifiques du bras D, à l'instar de la kynuramine dont la structure est présentée dans la figure IV-44[62].




**kynuramine 106** **107** **108** **109**

**Figure IV-154 La kynuramine, ligand spécifique du bras D de l'ARNt$^{Lys}_3$ et les analogues fluorés pouvant servir de sonde en RMN $^{19}F$.**

Nous avons donc cherché à développer des sondes fluorées s'inspirant de la structure de la kynuramine pour pouvoir, au cours d'une même expérience en RMN $^{19}F$, suivre les interactions sur le bras D de l' $ARNt^{Lys}_3$ à l'aide des nouvelles sondes fluorées développées, tout en suivant les interactions sur le bras T de l'$ARNt^{Lys}_3$ à l'aide du DACPF.

La synthèse des deux premières sondes présentées, les analogues de kynuramine fluorés **107** et **108**, ne présente pas de difficulté notoire et a pu être effectué par une réaction de Mannich à partir de 3- ou 4-fluoro-acétophénone, de paraformaldéhyde et de chlorhydrate de diméthylamine (voir partie expérimentale). Leur test en tant que sondes spécifiques du bras D de l'$ARNt^{Lys}_3$ est en cours au laboratoire. Nous avons tout d'abord pu vérifier en RMN $^{1}H$ que ces molécules étaient bien des ligands du bras D de l'$ARNt^{Lys}_3$ avec des affinités de l'ordre du millimolaire. Les études préliminaires en RMN $^{19}F$ laissent penser que ces sondes sont difficiles à déplacer, même avec de très bons ligands. Ceci laisse présager de l'existence de sites d'interactions non spécifiques de faible affinité qui vont peut-être nous contraindre à modifier la structure de ces sondes.

La synthèse de la dernière sonde fluorée, le composé **109**, s'appuie sur la formation d'un éther d'enol silylé sur lequel on fait agir une source de fluor « électrophile » comme le Select-Fluor. Elle n'a pas pu être menée à son terme en raison de l'instabilité du composé final, qui se dégrade très probablement selon un mécanisme de type « rétro-Mannich » (voir partie expérimentale).

En conclusion, les analogues fluorés de kynuramine semblent être des composés prometteurs pour le développement de sondes en RMN $^{19}F$. Nos premières expériences donnent des résultats encourageants qui justifient un effort d'amélioration de leurs propriétés.

# Conclusion Générale

Le projet de thèse qui m'a été confié s'inscrivait dans la thématique générale de la conception de nouveaux ligands d'ARN basés sur l'architecture des aminoglycosides. Ce projet très multidisciplinaire s'est appuyé sur trois approches complémentaires qui sont rappelées dans la figure V-1.

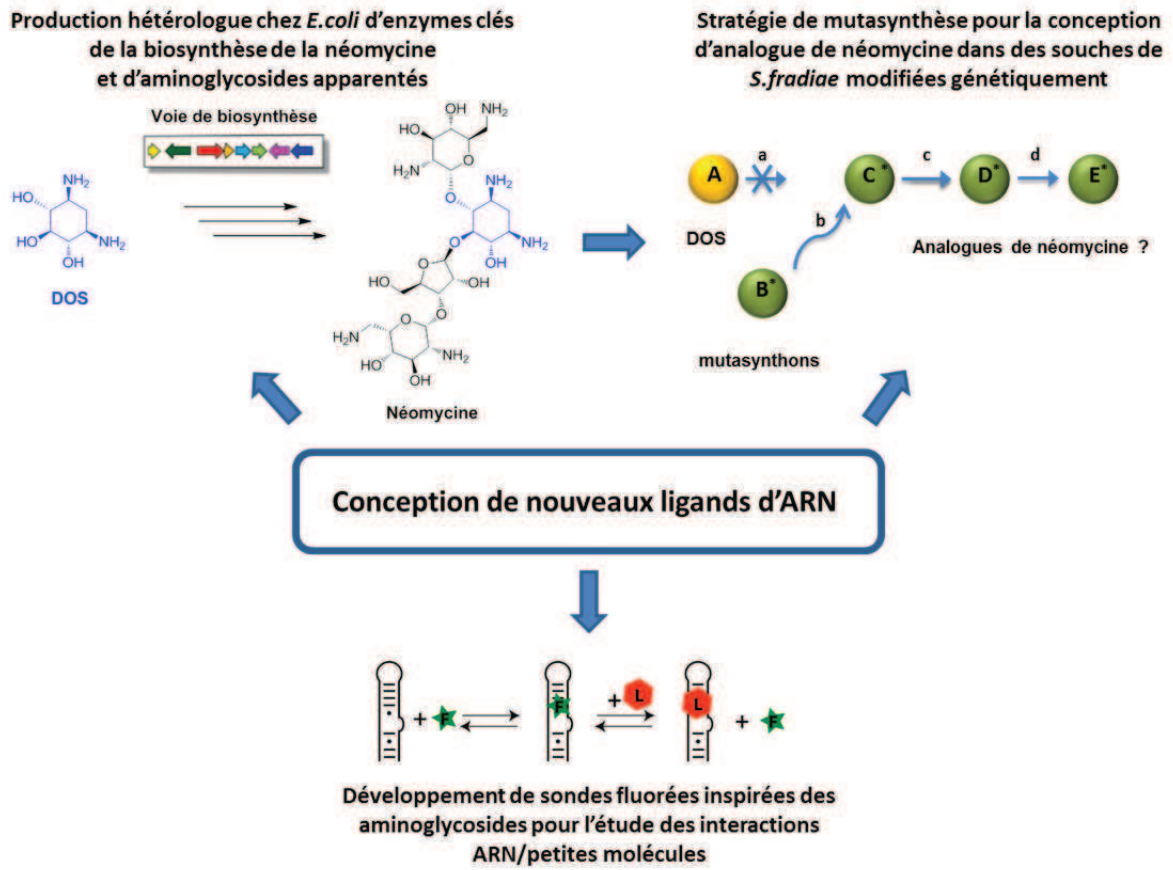



**Figure V-1 Représentation schématique du projet de thèse**

La première approche mise en place visait à mieux comprendre le fonctionnement de la voie de biosynthèse de la néomycine chez *S.fradiae* et de voies de biosynthèse d'aminoglycosides apparentées. Nous voulions exprimer de façon hétérologue chez *E.coli* certaines des enzymes clés de ces biosynthèses pour permettre leur étude biochimique et structurale. Malgré la très large palette expérimentale testée, la majorité de ces enzymes n'ont pas pu être exprimées de façon soluble en quantité suffisamment importante. Une hypothèse pouvant expliquer ces difficultés était l'interdépendance des différentes enzymes de biosynthèse, assemblées en un assemblage supramoléculaire complexe. Nous avons donc développé certains des outils nécessaires à la

confirmation expérimentale de cette hypothèse et avons obtenu des résultats préliminaires qui semblent la confirmer.

Nous avons ensuite décidé de nous concentrer sur l'ingénierie génétique de la voie de biosynthèse de la néomycine directement chez *S.fradiae* afin de mettre en place une stratégie de mutasynthèse pour produire des analogues de néomycine. Nous avons obtenu des résultats significatifs dans la mise en place de cette méthodologie, développant à la fois une souche délétée du gène Neo6 (Δneo6) codant une enzyme initiale dans la voie de biosynthèse, des protocoles de complémentation puis d'analyse des produits éventuellement formés par mutasynthèse ainsi qu'une chimie originale permettant la synthèse des différents analogues requis. Notre approche a été validée par la complémentation de la souche Δneo6 par des métabolites naturels de la biosynthèse de la néomycine et par sa comparaison avec une souche commercialement disponible, la souche DSM 41550. Les pistes obtenues lors des expériences de mutasynthèse avec des analogues non naturels (mutasynthons) sont très prometteuses et en cours d'analyse au laboratoire. Au cours de nos travaux, nous avons par ailleurs mis en évidence d'intéressants phénomènes de régulation et de divergence de la voie de biosynthèse de la néomycine chez *S.fradiae* pour l'étude desquels nous avons proposé des pistes d'étude.

Dans le but de faciliter le développement de nouveaux ligands d'ARN, nous avons par ailleurs développé des sondes fluorées s'inspirant de la structure des aminoglycosides, le DACPF et la paroF. Ces sondes se sont révélées être des outils extrêmement utiles pour étudier en RMN $^{19}F$ la structure et la dynamique des ARN ainsi que leurs interactions avec divers ligands. Les perspectives ouvertes par ces travaux sont multiples. En particulier, on peut exploiter ces sondes pour mettre en place un criblage moyen-débit en RMN $^{19}F$ de nouveaux ligands d'ARN. Nous envisageons aussi d'utiliser ces sondes pour suivre en RMN $^{19}F$ la dynamique de systèmes biologiques complexes comme l'initiation de la transcription inverse du VIH.

# V. Partie expérimentale

# I - Partie expérimentale du chapitre III : Mutasynthèse

## *A - Production de spores de S.fradiae*

Les bactéries du genre Streptomyces ont un cycle de développement particulier au regard des bactéries majoritairement utilisées dans les laboratoires. Sur milieu solide, il commence par une phase dite de croissance végétative où la germination d'une spore donne naissance à un mycélium primaire ramifié se développant en surface et à l'intérieur du substrat nutritif. Cette phase est suivie d'une phase de croissance aérienne où des hyphes aériens s'érigent à partir du mycélium primaire puis sporulent. Ces différentes phases sont présentées dans la figure S-1.

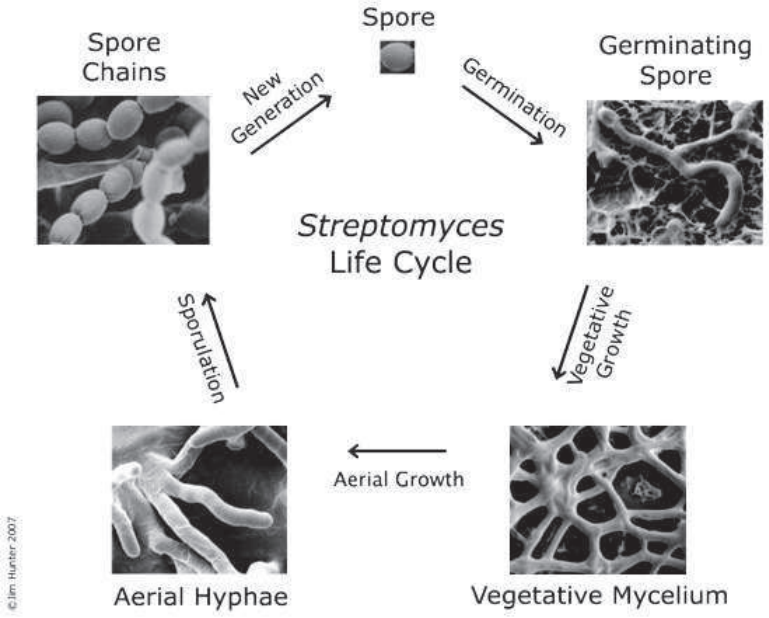



**Figure S-155 Les différentes phases du cycle de développement de *S.fradiae***
(référence : http://openwetware.org/wiki/Streptomyces)

Les souches de *S.fradiae* sont systématiquement gardées sous forme de spores à -20°C ou à -80°C dans une solution de glycérol stérile à 20 % final. L'ensemble des manipulations de ces stocks doit se faire en environnement parfaitement stérile (hotte à flux laminaire, instruments stérilisés, solutions stériles,…).

Pour préparer un stock de spores à partir d'une bactérie isolée, prélever un peu de mycélium (les hyphes aériens contiennent des spores) avec un cure-dent, l'écraser entre deux cure-dents puis l'étaler sur une boite de SFM, milieu de culture solide qui favorise la sporulation. Si on est en possession d'un stock de spores, on peut se contenter d'étaler environ $10^8$ spores par boite de SFM (habituellement environ 10 µL du stock). Mettre à pousser à 30°C pendant 7 jours. A ce stade, la boite de pétri est recouverte de mycélium blanc et de spores (non visibles à l'œil nu). On peut alors récolter les spores selon le protocole suivant :

1) Sur chaque boite SFM, ajouter 3 mL d'eau stérile puis gratter le mycélium avec une anse de platine.

2) Récupérer le mélange avec une pipette stérile et filtrer sur coton stérile de façon à éliminer le mycélium en laissant passer les spores

3) Répéter l'opération plusieurs fois

4) Centrifuger le filtrat. Le culot est constitué de spores et peut être repris dans une solution de glycérol stérile, transféré dans un contenant approprié puis gardé à -20°C.

Pour évaluer la concentration d'une solution stock de spores, on peut procéder à des étalements en duplicatat ou triplicatat de dilutions successives de la solution stock sur boites de SFM. Le nombre de colonies se formant sur chaque boite permet d'en déduire la concentration initiale du stock. Habituellement, les stocks de spores de *S.fradiae* ont une concentration variant de $10^9$ à $10^{10}$ spores/mL.

Préparation du milieu SFM

(parfois aussi appelé MS dans la littérature pour « Mannitol soya flour medium »)

- 20 g mannitol
- 20 g farine de soja précuite (Biocoop)
- 20 g agar
- $H_2O$ q.s.p. 1000 mL

## *B - Extraction d'ADN total de Streptomyces : technique Fast-prep*

Cette technique d'extraction utilise un homogénéisateur de type Fast-prep.

1. Ensemencer 5mL de milieu TSB avec environ 10µL de spores. Incuber à 30°C pendant 2 à 3 jours.
2. Transférer la culture dans un tube Fast-prep et centrifuger pendant 5 min à 13000 g Travailler dans la glace à partir de cette étape
3. Eliminer le surnageant et ajouter 400mg de billes de verre, 400µL d'eau stérile et 400µL de phénol/chloroforme (phénol : chloroforme : alcool isoamylique [25 :24 :1])
4. Faire 3 runs au fast-prep : vitesse 4 pendant 30 sec avec une pause d'une minute dans la glace à chaque fois
5. Centrifuger 10 min à 13000g
6. Récupérer le surnageant et y ajouter 400µL de phénol/chloroforme
7. Centrifuger 10 min à 13000g
8. Récupérer le surnageant
9. Précipiter l'ADN avec un volume d'isopropanol (congeler à -20°C favorise la précipitation)
10. Centrifuger 10 min à 13000g et éliminer le surnageant
11. Laver le culot à l'éthanol 70% et sécher le culot au speed-vac

12. Resuspendre le culot dans 50 à 200µL de tampon TE (Tris 10 mM, pH 8, EDTA 1mM) + RNase (10µg de RNase/mL de TE)
13. Incuber 1 h à 37°C
14. Stocker l'ADN extrait à 4°C jusqu'à utilisation (peut se garder pendant plusieurs mois).

## *C - Conditions PCR*

**1) Amplification du fragment neo5-neo7 à partir de l'ADN génomique de *S.fradiae* sauvage**

On utilise une polymèrase Phusion (NEB) et les amorces suivantes :

neo5-7F : ACAACTCGCACCCGAACATCCAACAACC

neo5-7R : GGACGTGCACCAGGTCGGGCCTCCAG

Le fragment neo5-neo7 (4483 paires de bases) est amplifié selon le programme suivant :

5min dénaturation initiale à 97°C

puis 30 cycles de :

- 30 sec à 97 °C (dénaturation)
- 45 sec à 62 °C (hybridation)
- 10 min à 72 °C (élongation)

Le cycle se termine par 10 min à 72°C (élongation finale)

**2) Amplification du gène neo6 à partir de l'ADN génomique de *S.fradiae* sauvage**

On utilise une polymérase Accuprime G-C rich (Invitrogen) et les amorces :

neo6F : ACACCCACTGCAGGCCCTCAGGAAAGGAAG size 1324

neo6R : GGTGCTCGAGGGCGGTCAAGTGGCCAGGTC

Le fragment d'ADN contenant le gène neo6 (1324 paires de bases) est amplifié selon le programme suivant :

5min dénaturation initiale à 97°C

puis 26 cycles de :

- 30 sec à 97 °C (dénaturation)
- 45 sec à 60 °C (hybridation)
- 2min30 à 72 °C (élongation)

Le cycle se termine par 10 min à 72°C (élongation finale)

**3) PCR de vérification des clones lors de la double recombinaison**

On utilise une polymérase Taq (Qiagen) et les amorces suivantes :

- Contrôle de la présence de la cassette Apra
  acc2F : CGGATCTCGGCCCAGTTGAC
  acc2R : CGTTCGATCCTGCCACGTCG
  Taille du fragment attendu : 500 paires de bases

- Contrôle de la présence de la cassette Hygro
  AG-hyg1 : TCGGATGATTCCTACGCGAG
  Ώ HYG2 : TCCTCGAACACCTCGAAGTC
  Taille du fragment attendu : 872 paires de bases

- Contrôle de la liaison neo5-7 avec Apra
  neo5-7scrR : GTCTGCATGATGCTTCCCGGTCCT
  apraF : ACTGAGGAGGTCTACCGGAAGCAGA
  Taille du fragment attendu : 500 paires de bases

Pour l'ensemble de ces fragments, on utilise le même programme d'amplification :

5min dénaturation initiale à 95°C

puis 30 cycles de :

- 30 sec à 95 °C (dénaturation)
- 30 sec à 55 °C (hybridation)
- 2min à 72 °C (élongation)

Le cycle se termine par 10 min à 72°C (élongation finale)

## *D - Conjugaison Streptomyces/E.coli : délétion de neo6 par double recombinaison homologue*

Ce mode opératoire est adapté d'un protocole qui m'a été transmis par Florence Lorieux (équipe de Jean-Luc Pernodet). Il a été initialement décrit dans le livre de référence « Practical Streptomyces Genetics » de Hopwood et ses collaborateurs[253]. En voici les différentes étapes :

1. Préparer un stock de cellules électrocompétentes d'*E.coli* S17.1 (souche contenant la machinerie nécessaire à la conjugaison avec des bactéries du genre *Streptomyces*[370,371]) ou souche équivalente comme *E.coli* ET12567 (pUZ8002) (souche ne méthylant pas le plasmide échangé ce qui permet d'éviter la dégradation de celui-ci par des enzyme de restriction methyl-spécifiques lors de la conjugaison dans certaines souches de *Streptomyces*[372])
2. Transformer les bactéries *E.coli* par le plasmide à conjuguer (électroporation). Eventuellement, étalement sur boite LB agar avec sélection antibiotique (ici, hygromycine 50 µg/mL, apramycine 50 µg/mL) et isolement d'un clone pour lancer la préculture.

3. Préculture dans le milieu de culture LB sur la nuit à 37°C en présence d'antibiotiques (résistance portée par le plasmide de conjugaison et éventuellement par la souche bactérienne elle-même) (ici, hygromycine 50 µg/mL, apramycine 50 µg/mL).
4. Diluer la préculture au 1/100$^{ème}$ dans 50 mL de LB + antibiotiques (hygromycine et apramycine 50 µg/mL), et laisser pousser à 37°C jusqu'à obtenir une densité optique (DO) de 0.4 à 0.6 (à 600 nm).
5. Centrifuger les bactéries et laver le culot obtenu avec 50 mL de LB. Recommencer l'opération une à deux fois puis resuspendre le culot bactérien dans 5 mL de LB.
6. Pendant ce temps, pour chaque expérience de conjugaison, ajouter environ $10^8$ spores de *Streptomyces* dans 500 µl de milieu de culture 2TY. Mettre à chauffer pendant 10 min à 50°C puis refroidir à 0°C et maintenir sur la glace.
7. Ajouter aux 500 µl de spores, 500 µl des *E.coli* préparées précédemment. Mélanger et centrifuger rapidement (20 sec). Enlever le surnageant et reprendre dans la goutte de liquide restant.
8. Étaler sur boites HT agar contenant 10 mM de $MgCl_2$ (bien séchées pendant 1 heure sous la hotte aspirante) et incuber à 30°C pendant 16 à 20 h.
9. Avec un étaloir ou une anse de platine et un peu de LB, racler les bactéries *E.coli* sur trouvant en surface des boites de pétri. (Les *streptomyces* sont légèrement enfouies dans l'agar). Enlever le liquide avec une pipette. Répéter l'opération 2 à 3 fois puis couler un top SNA de 3ml, contenant 40µg/mL d'acide nalidixique (pour tuer les *E.coli* restantes) et la sélection antibiotique (apramycine 50 µg/mL). Mettre à pousser à 30° pour 5 à 7 jours.
10. Repiquer les exconjugants potentiels sur milieu sélectifs (ici, SFM ou HT avec apramycine 50 µg/mL et/ou hygromycine 50 µg/mL) contenant 25 µg/ml d'acide nalidixique. Les ex-conjugants sont sélectionnés pour leur résistance à l'apramycine (50 µg/mL) et leur sensibilité à l'hygromycine (50 µg/mL). Les bactéries résistantes à l'apramycine et sensibles à l'hygromycine ont subi un événement de double recombinaison du plasmide échangé par conjugaison. Le gène cible (dans notre cas, neo6) est alors remplacé par une cassette de résistance à l'apramycine.

Milieux utilisés:

- **LB**

  10 g tryptone
  5 g extrait de levure
  10 g NaCl
  $H_2O$ q.s.p. 1000 mL
  (alternativement, on peut acheter du LB déjà préparé en poudre (réf. Sigma Aldrich L3022))

- **HT**

  1 g extrait de levure
  1 g extrait de boeuf
  10 g dextrine

2 g tryptone
1mL CoCl2, 7 H2O (20 g/l)
20 g agar
$H_2O$ q.s.p. 1000 mL
pH ajusté à 7,3 (NaOH 1M ~5ml)

- **2TY**
  16 g tryptone
  10 g extrait de levure
  5 g NaCl
  $H_2O$ q.s.p. 1000 mL

- **SNA**
  8 g Bacto Nutrient Broth
  5 g agar
  $H_2O$ q.s.p. 1000 mL

- **SFM**
  20 g mannitol
  20 g farine de soja précuite (référence Biocoop)
  20 g agar
  $H_2O$ q.s.p. 1000 mL

## *E - Test phénotypique de production de néomycine sur milieu solide*

Les souches à tester sont repiquées sur une boite de pétri contenant du milieu SFM puis incubées pendant 7 jours à 30°C (condition de production de néomycine de la souche sauvage). Après 7 jours, on recouvre les boites d'un top SNA ensemencé avec une bactérie reportrice, qui peut être *M.luteus* ou *E.coli* DH5α, toutes deux sensibles à la néomycine (4 mL SNA pour environ 400 µL d'une culture bactérienne à DO 0.6. Maintenir le SNA à 45°C pour qu'il reste liquide avant l'ensemencement). On peut également utiliser une bactérie transformée avec un plasmide conférant une résistance à la néomycine (par exemple, le plasmide pBK-CMV) pour tester la spécificité de l'inhibition de croissance. Après incubation à 30 °C pour une nuit supplémentaire, l'éventuelle production de néomycine par les souches de S.fradiae se lit à la présence de halo d'inhibition de la souche reportrice.

## *F - Production de néomycine par fermentation de S.fradiae : Optimisation des conditions de cultures pour la mutasynthèse*

### 1) Production de néomycine à petite échelle (fernbach)

Nous avons testé différents milieux de culture pour la production de néomycine par la souche sauvage *S.fradiae*. Les conditions de culture donnant les titres les plus importants en néomycine

sont : 30°C, 220 rpm, environ 7 jours de culture dans un fernbach à picots contenant du milieu de culture TSB. Ensemencer 50 mL de TSB avec environ 10 µL d'une solution stock de spores à $10^{10}$ spores/mL.

**2) Production de néomycine à grande échelle (fermenteur)**

Nous avons également testé la production de *S.fradiae* en fermenteur de 4L. Les titres en néomycine obtenus sont également très bons (de l'ordre de 600 µg/mL).

Le protocole est le suivant : préculture pendant 72 h à 30°C de 4 x 80 mL de TSB inoculés avec 20 µL de spores à $5.10^9$ spores/mL puis inoculation de 3.5 L de TSB dans le fermenteur stérile avec l'ensemble de ces précultures (320 mL). On incube sous agitation (200-500 rpm) pendant 96 h en maintenant une température de 30°C, une pression partielle en oxygène et un pH stable. On ajoute de l'antimousse régulièrement pour éviter que le fermenteur déborde (référence antimousse Rhône-Poulenc épuisée mais qu'on peut remplacer par l'antimousse A6207 de Sigma Aldrich). En fin de fermentation, on centrifuge le milieu de culture pendant 1h à 8000 g et on récupère le surnageant qui contient la néomycine. Les titres en néomycine obtenus à cette échelle sont légèrement supérieurs à ceux que l'on peut obtenir en fernbach à picots (de l'ordre de 600 µg/mL).

**3) Expériences de mutasynthèse**

Les expériences de mutasynthèse sont effectuées dans les mêmes conditions que la production de néomycine par fermentation de la souche sauvage de *S.fradiae* : Dans un fernbach à picots de 250mL, incuber de 50 mL de TSB ensemencés avec 10 µL d'un stock de spores de la souche d'intérêt (concentration d'environ $10^9$ à $10^{10}$ spores/mL) pendant 7 jours à 30°C sous une agitation de 220 rpm. Les mutasynthons sont ajoutés à une concentration finale variable (habituellement 2 mM) soit au début de la culture, soit par batchs successifs au fur et à mesure de la culture. Idéalement, on prépare une solution concentrée de mutasynthon dans l'eau que l'on stérilise par filtration à travers des filtres 0.2 µm. En fin de culture, on centrifuge les milieux de culture pendant 20 min à 8000 g pour récupérer le surnageant qui contient la néomycine ou les éventuels analogues formés.

**4) Différents milieux testés**

i. **TSB** (référence sigma 22092)

17.0 g digestion enzymatique de caséine (digestion pancréatique)
3.0 g digestion enzymatique de soja (digestion à la papaïne)
5.0 g NaCl
2.5 g phosphate de potasium $K_2HPO_4$
2.5 g dextrose
pH final : 7.3 ± 0.2 à 25°C

$H_2O$ q. s. p. 1000ml

Alternativement, simplement ajouter de l'eau au mélange commercial

ii. **ONB** (reference oxoid CM0001)

1.0 g extrait de viande de bœuf
2.0 g extrait de levure
5.0 g peptone
5.0 g chlorure de sodium
pH final : 7,4 ± 0,2 à 25°C
$H_2O$ q. s. p. 1000ml
Alternativement, simplement ajouter de l'eau au mélange commercial

iii. **MP5** (milieu décrit par JL pernodet pour la production de congocidine[373])

7 g extrait de levure
5 g NaCl
1 g $NaNO_3$
36 ml glycérol
20.92 g MOPS (acide 3-(N-morpholino)propanesulfonique) q. s. p. 0.1M final
$H_2O$ q. s. p. 1000ml
Ajuster à pH = 7.5 avec NaOH

**iv. Milieu minimal décrit par Berwick pour la production de néomycine**[320]

0.5 g $CaCl_2$, $6H_2O$
20 g glucose
5 g glycine
0.5 g $KH_2PO_4$
10 g $MgSO_4$, $7H_2O$
5 g NaCl
10 g citrate de sodium
$H_2O$ q.s.p 1000 mL

v. **R5** (fourni par Jean-Luc Pernodet)

103 g Sucrose
0.25 g $K_2SO_4$
10.12 g $MgCl_2$, $6H_2O$
10 g glucose
0.1 g Difco casaminoacids
2ml solution d'éléments à l'état de trace
5 g extrait de levure
5.73 g tampon TES
$H_2O$ q.s.p 1000 mL
Au moment de l'utilisation, complémenter avec :
5ml $KH_2PO_4$ (0.5%)
2ml $CaCl_2$ $2H_2O$ (5M)
7.5ml L-proline (20%)
3.5ml NaOH (1N)

**Solution d'éléments à l'état de trace (solution d'Okanashi)**[374]

40 mg $ZnCl_2$
200 mg $Fe_3Cl_3$, 6 $H_2O$
10 mg $CuCl_2$, 2 $H_2O$
10 mg $MnCl_2$, 4 $H_2O$
10 mg $Na_2B_4O_7$, 10 $H_2O$
10 mg $(NH_4)_6$ $Mo_7O_{24}$, 4 $H_2O$
$H_2O$ q.s.p 1000 mL

vi. **HT**

1 g extrait de levure
1 g extrait de bœuf
10 g dextrine
2 g tryptone
1mL CoCl2, 7 H2O (20 g/l)
20 g agar
$H_2O$ q.s.p. 1000 mL
pH ajusté à 7,3 (NaOH 1M ~5ml)

vii. **SSG**

10 g amidon soluble
10 g glucose
10 g glycerol
2.5 g farine de mais
5 g peptone
2 g extrait de levure
1 g NaCl
1 g $CaCO_3$
$H_2O$ q.s.p. 1000 mL
pH ajusté à 7,2

## *G - Purification d'aminoglycosides par colonne échangeuse d'ions (SPE)*

1) Prélever 10 mL de milieu de culture issu d'une fermentation de *S.fradiae*. Centrifuger à vitesse maximale (8000 g pour les falcons 50 mL) pendant 20 min. Récupérer le surnageant et le diluer avec 40 mL de tampon phosphate $KHPO_4$ 50 mM, pH 6.5.
2) Filtrer la solution obtenue à 0.45 µm. On peut à ce stade congeler les échantillons à -20 °C pour les analyser ultérieurement.
3) Préparer les colonnes échangeuses d'ions (cartouche d'extraction en phase solide SPE) Strata X-CW (Phenomenex) en les rinçant avec 25 mL de méthanol puis 25 mL d'eau. Pour accélérer le passage des différentes solutions, on travaille en aspiration avec une trompe à eau (pression de l'ordre de 30 mmHg) à l'aide d'un extracteur sous vide.
4) Charger les échantillons sur les colonnes SPE.
5) Rincer les colonnes avec 50 mL de tampon phosphate puis avec 50 mL de méthanol. Bien sécher les colonnes.
6) Eluer les colonnes avec 2 x 5 mL d'ammoniac aqueux $NH_4OH$ à 1M. On peut ici couper l'aspiration et laisser couler goutte à goutte pour obtenir un meilleur décrochage des analogues d'aminoglycosides de la colonne. Alternativement, il est possible d'utiliser un gradient d'ammoniac pour décrocher progressivement les molécules accrochées sur la colonne en recueillant les différentes fractions. Cette méthode conduit à une purification plus importante des produits.
7) Lyophiliser les échantillons. On peut garder les échantillons lyophilisés à -20 °C jusqu'à analyse.

## *H - Analyse des analogues d'aminoglycosides purifiés par HPLC-MS/MS*

### 1) Chromatographie liquide haute performance (HPLC)

Après lyophilisation, les échantillons purifiés par SPE sont repris dans 500µl d'$H_2O$ et 25µl sont injectés. Ils sont conservés dans le passeur à 4°C pendant la durée de l'expérience.

La séparation des analytes est réalisée sur un système HPLC Surveyor® (ThermoScientific, Courtaboeuf, France) comprenant une pompe HPLC et un passeur d'échantillon permettant de thermostater les échantillons et la colonne HPLC.

La phase stationnaire est constituée par une silice greffée octadodécyle (EC 250/2 Nucleodur Pyramide, 5µm, 250mm x 2.1mm, Macherey-Nagel, Hoerdt, Fr). La colonne est maintenue à 20°C. Un gradient d'élution binaire est composé de A - $H_2O$ 0.1% *acide pentafluoropropionique (v/v) et B-acétonitrile 0.1% acide pentafluoropropionique. Le débit est fixé à* 200µL.$min^{-1}$.

| temps | % A | %B |
|---|---|---|
| **0** | 90 | 10 |
| **15** | 40 | 60 |
| **17** | 40 | 60 |
| **20** | 90 | 10 |
| **25** | 90 | 10 |

**2) Spectrométrie de masse (MS)**

La détection est effectuée par spectrométrie de masse à l'aide d'un appareil LCQ Advantage® (ThermoScientific, Courtaboeuf, France) utilisant comme technique d'ionisation l'électronébulisation et comme analyseur un piège à ions. Le logiciel Xcalibur® est utilisé pour l'acquisition et le traitement des données.

L'ionisation est effectuée en mode positif, l'optimisation des paramètres de détection est réalisée en injectant simultanément et en continu dans la source d'ionisation :

- une solution de néomycine à la concentration de $1.10^{-6}$ $mol.L^{-1}$ à un débit de $5\mu l.min^{-1}$ à l'aide d'une seringue de 500µl SGE (infusion).

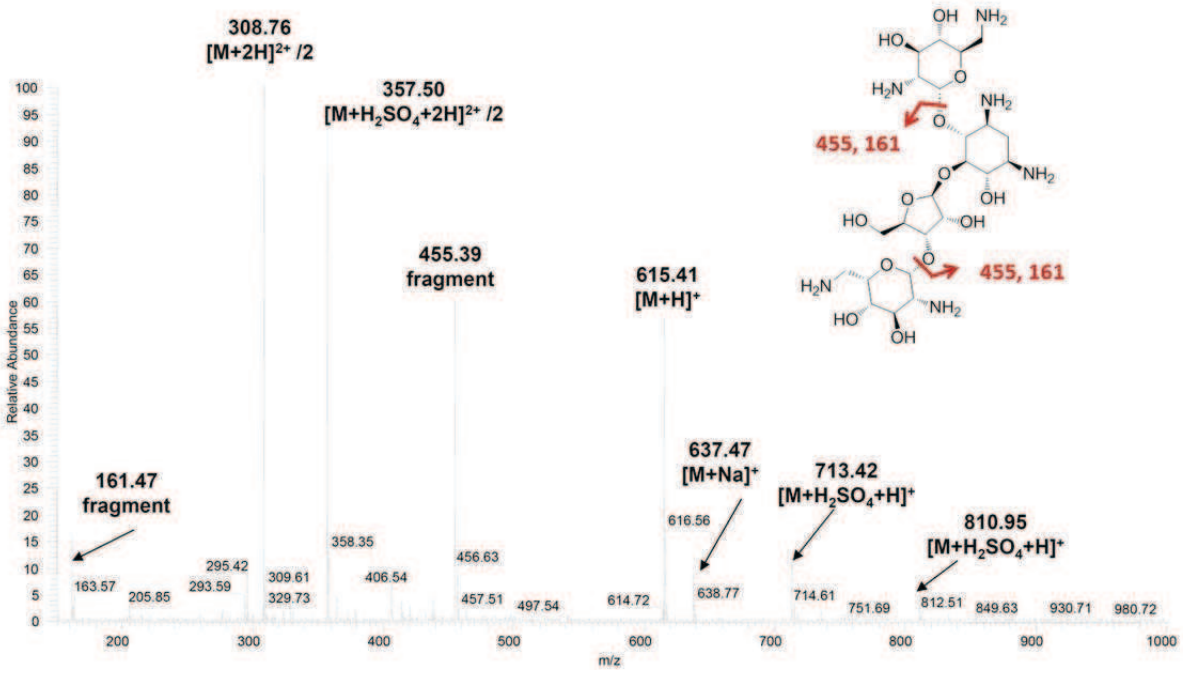
- et la phase mobile à un débit de $200\mu l.min^{-1}$.

Les paramètres retenus après optimisation sont les suivants : différence de potentiel appliquée au capillaire de transfert 13.0V, différence de potentiel appliquée à l'électronébulisation 4.0kV ; température du capillaire de transfert 20°C ; débit du gaz fourreau 60u.a. ; débit du gaz auxiliaire 20u.a. .

L'acquisition des données est réalisée en mode "full scan" sur un intervalle de 100,00 -1000,00 (*m/z*). Trois microscans sont nécessaires pour produire un scan et la durée maximale d'accumulation des ions dans l'analyseur est fixée à 50 ms. Les profils d'élution sont obtenus en sélectionnant le rapport *m/z* de l'ion moléculaire considéré avec un intervalle de +/- *m/z*=0,5 et les temps de rétention correspondant à chaque pic repérés sont donnés en min.

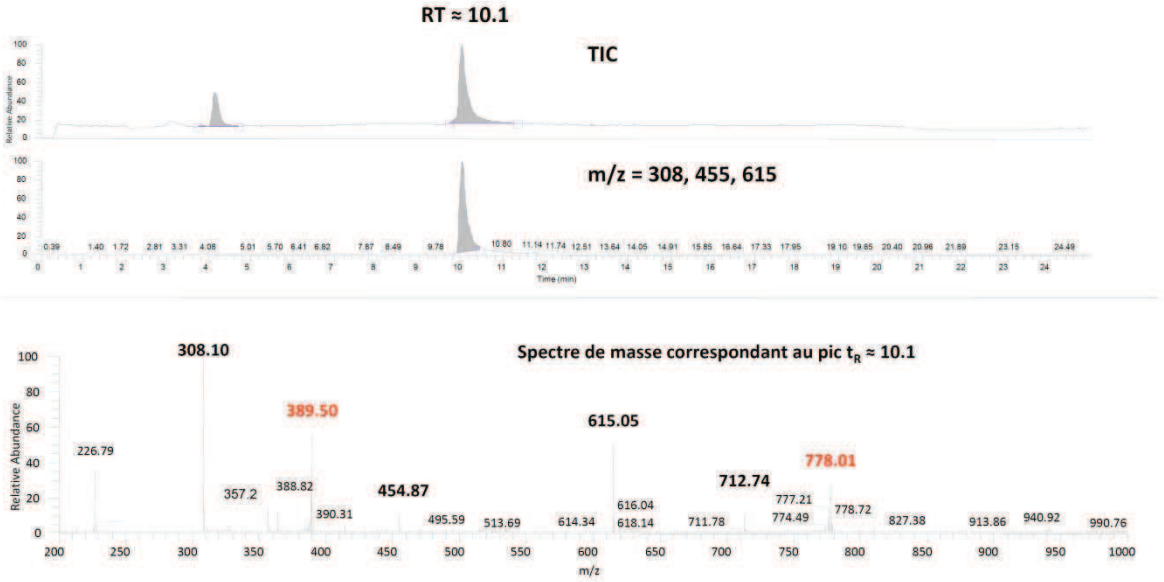
## 3) Spectres de références

### a) Spectre de masse du sulfate de néomycine (infusion)



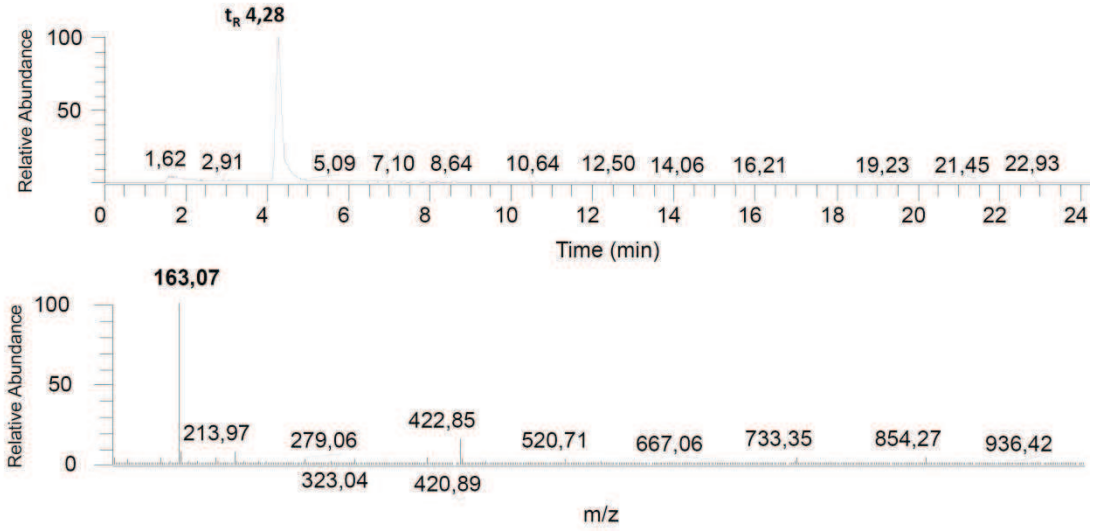

En fractionnant le fragment à 455, on peut obtenir un fragment correspondant à la néamine à 323. En fragmentant davantage, on obtient un fragment de 162 correspondant à la DOS.

### b) Spectre HPLC-MS de la néomycine après purification sur SPE

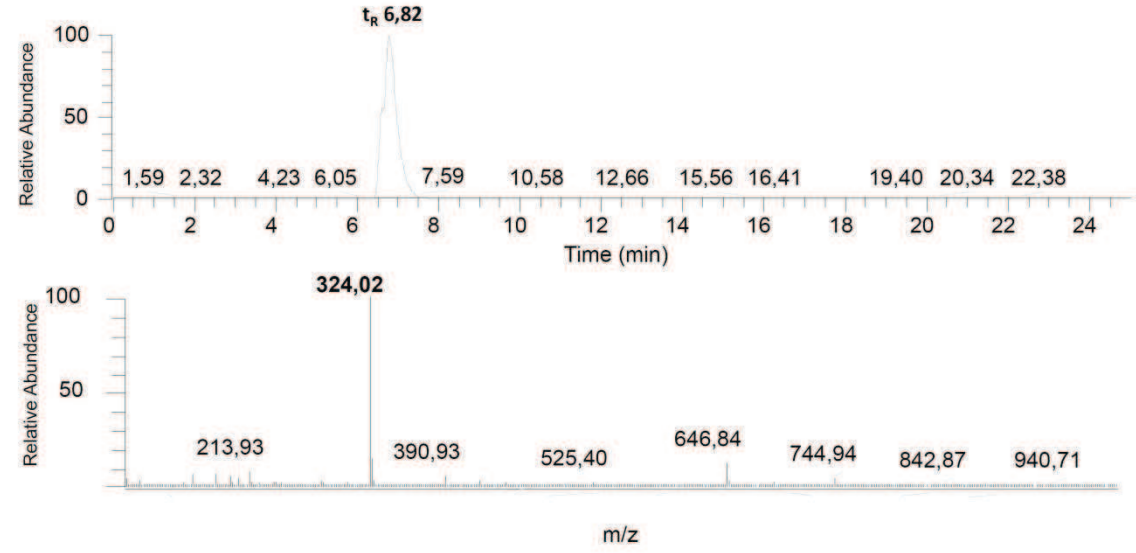


Le pic de néomycine sort à environ 10.1 min. Son spectre de masse est identique à celui obtenu par infusion à l'exception des deux pics marqués en rouge qui résultent probablement de la formation d'un adduit avec l'*acide pentafluoropropionique utilisé dans les solvants de HPLC. En optimisant les conditions d'ionisation, il est toutefois possible de supprimer ces pics parasites.*

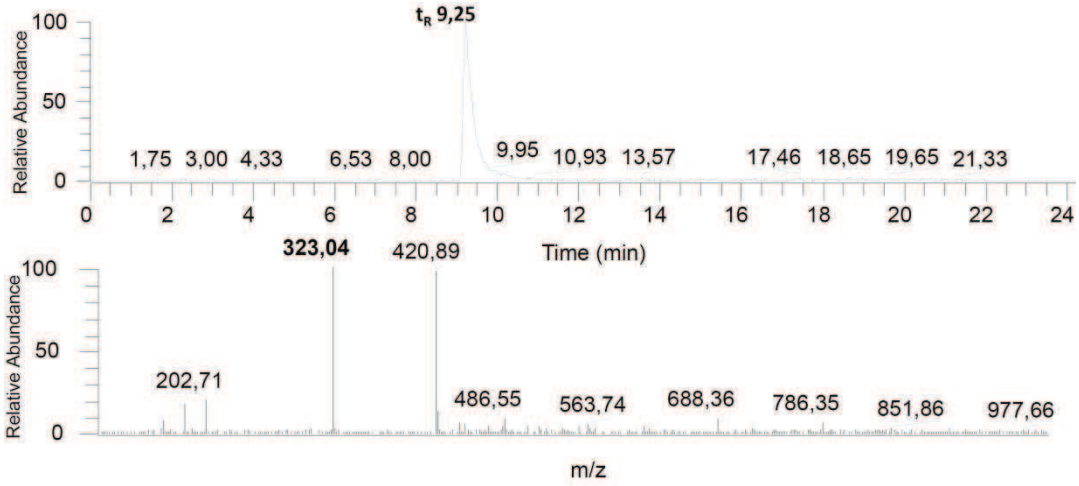
### c) Spectre HPLC-MS de la DOS



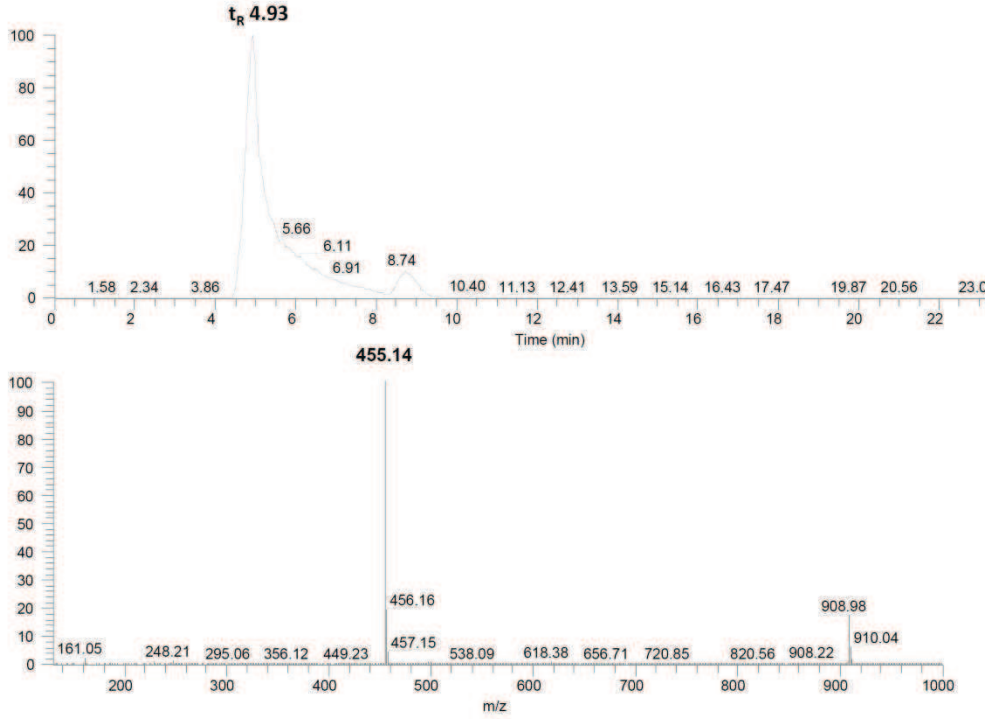

### d) Spectre HPLC-MS de la paromamine



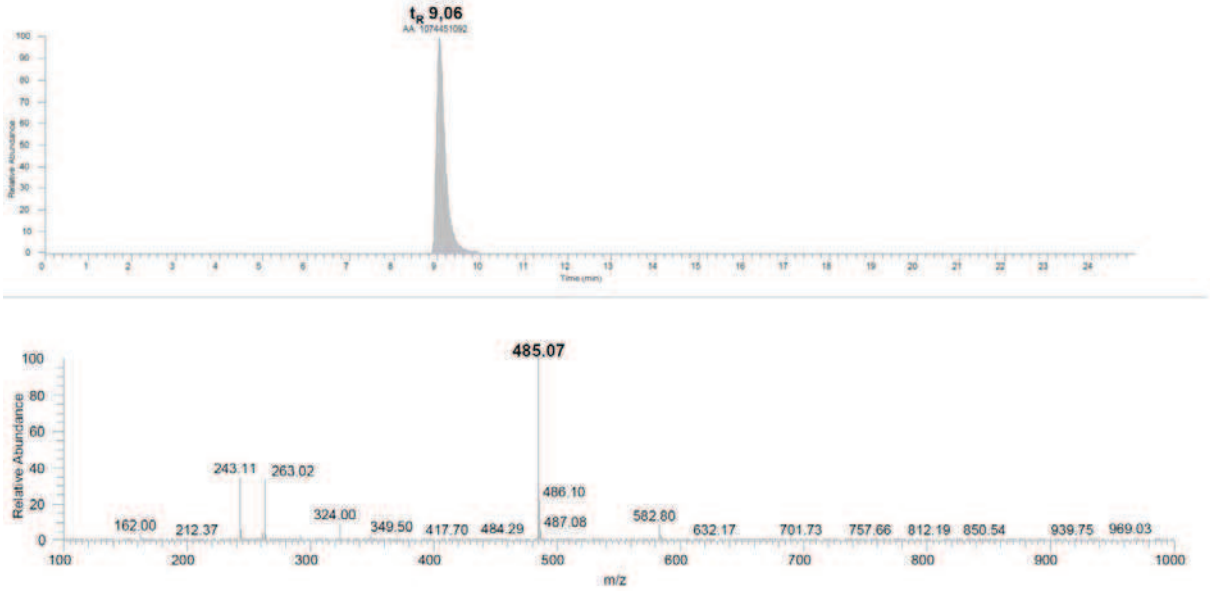
**e) Spectre HPLC-MS de la néamine**



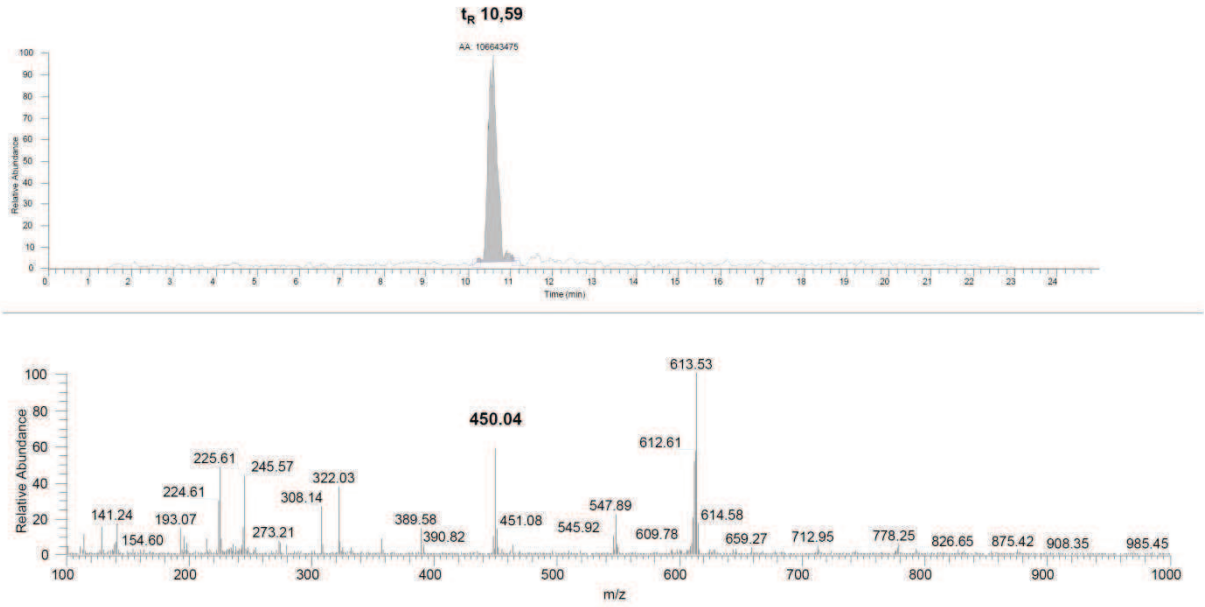

**f) Spectre HPLC-MS de la ribostamycine**



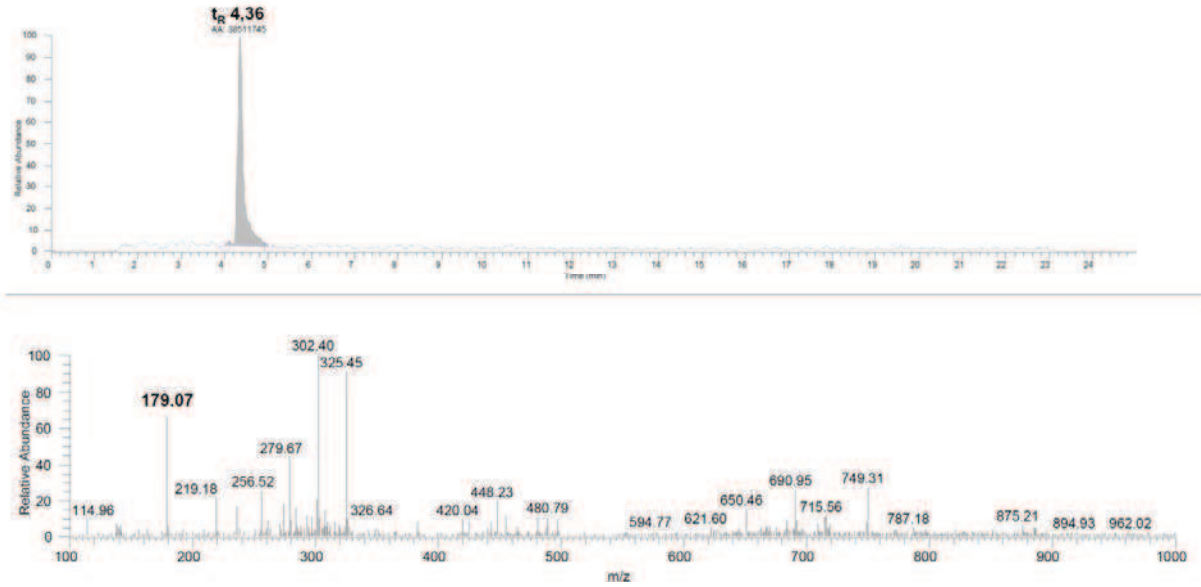
### g) Spectre HPLC-MS de la kanamycine



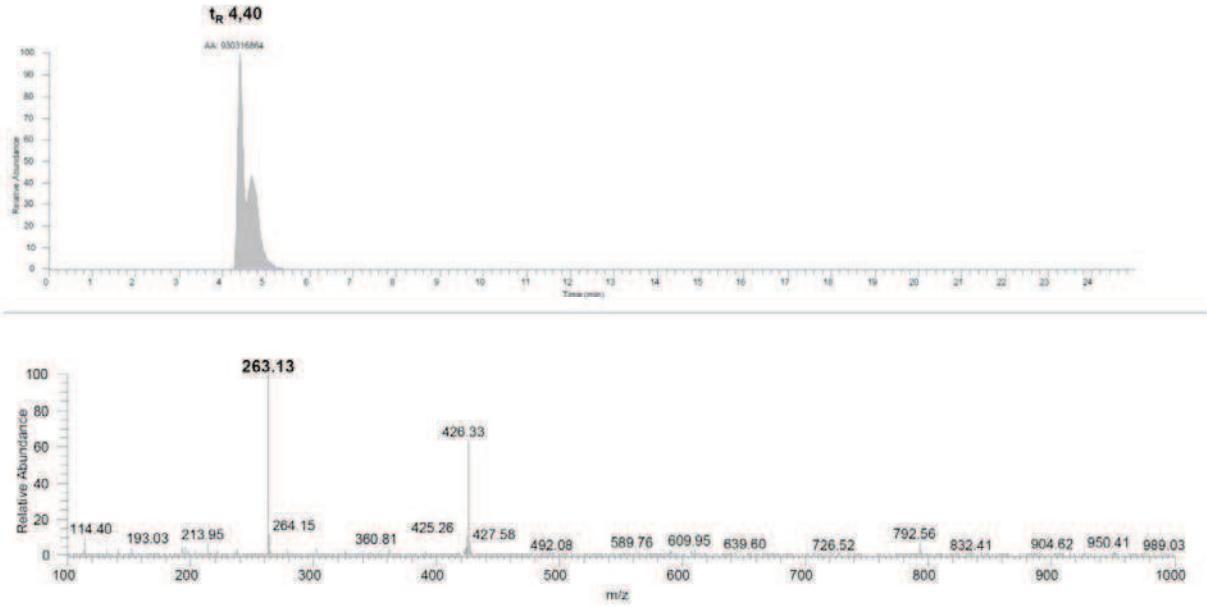

### h) Spectre HPLC-MS de la gentamicine



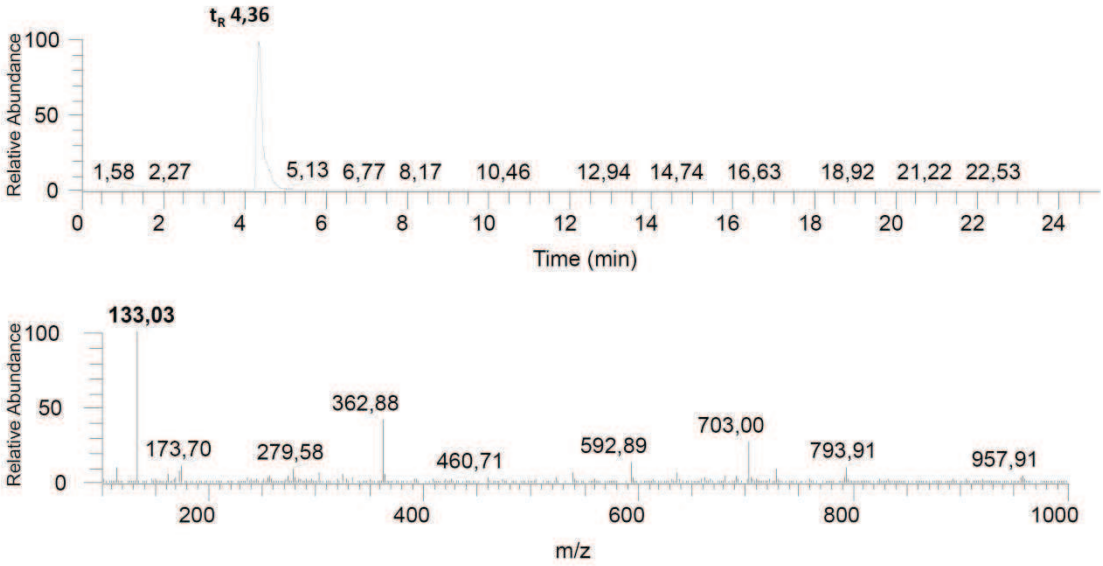
**i) Spectre HPLC-MS de la streptamine**




**j) Streptre HPLC-MS de la streptidine**



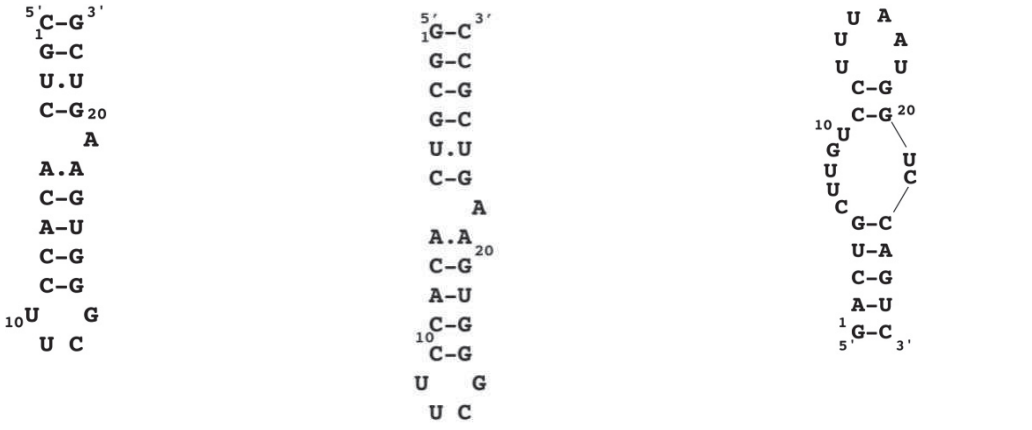
**k) Spectre HPLC-MS du cis-1,3-diamino-cyclopentan-4,5-diol C5**



# II - Partie expérimentale du chapitre IV : $^{19}$F RMN

## *A - Préparation des ARN*

L'ARN 16S23, l'ARN 16S27 et l'aptamère à la néomycine (figure S-2) ont été acheté chez ThermoFisher Scientific, sous forme lyophilisée et protégée en 2' par des groupements ACE (bis(acetoxyethoxy)-methyl) (figure S-3).



**16S23** **16S27** **Aptamère à la néomycine**

**Figure S-156 L'ARN 16S23 et l'aptamère à la néomycine**

Les échantillons ont été déprotégés à l'aide du protocole suivant (figure S-3) :

1. Centrifugation rapide des échantillons d'ARN lyophilisés (10 sec sur une centrifugeuse de paillasse)
2. Dilution des échantillons dans 400 µL de la solution de déprotection fournie (acide acétique 100 mM ajusté à pH 3,8 par ajout de TEMED (Tétraméthyléthylènediamine)) et pipetage jusqu'à dilution complète du culot d'ARN
3. Vortex pendant 10 secondes
4. Centrifugation rapide
5. Incubation de la solution pendant 30 minutes à 60°C
6. Refroidissement rapide des tubes dans un bain de glace à 0°C pendant 2 min

7. Dilution de l'échantillon dans un grand volume (environ 10 mL) du tampon voulu pour les analyses ultérieures. Dans la plupart des expériences RMN, nous avons utilisé du tampon phosphate (PBS) (10 mM, 50 mM KCl, pH 6,5)
8. Concentration des échantillons par dialyse ou centrifugation à travers des systèmes de centriprep.
9. Congélation des solutions jusqu'à utilisation ultérieure

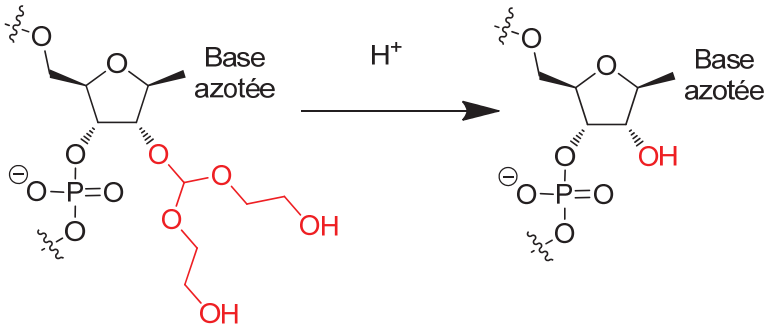



**Figure S-157 déprotection des groupements ACE**

L'ARNt$^{Lys}_3$ a été fourni par Marjorie Catala, ingénieure d'étude dans mon équipe au sein de l'UMR 8015 (figure S-4). Il a été obtenu par surexpression chez *E.coli* et purification par échangeuse d'ions et tamis moléculaire selon le protocole développé au laboratoire par Carine Tisné[353]. Cet ARN de transfert recombinant possède 7 des 13 modifications post-transcriptionnelles normalement trouvées chez les eucaryotes. Il a été montré qu'il était actif dans le processus de rétrotranscription du VIH et donc qu'il constituait un modèle valable pour son étude.

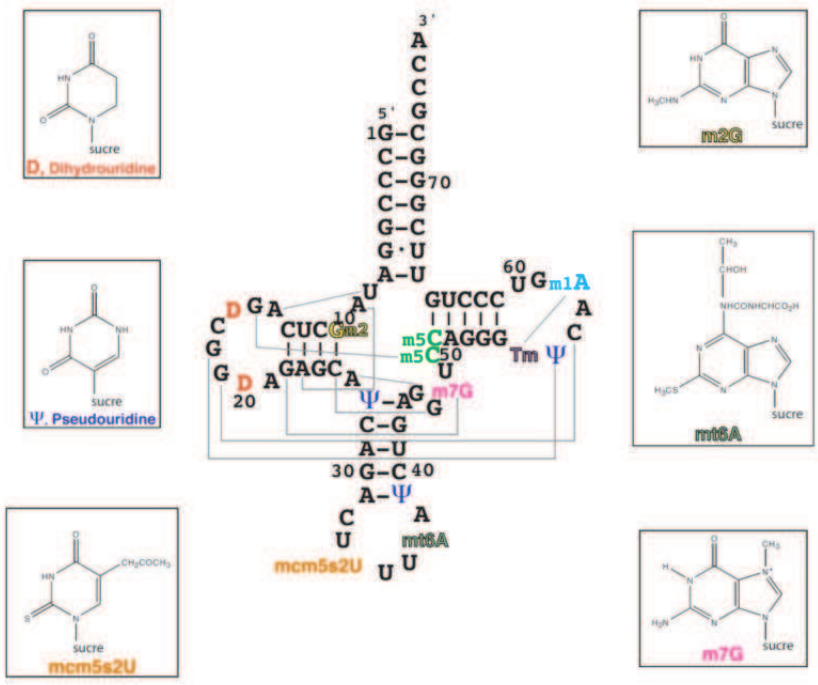



**Figure S-158 L' ARNt$^{Lys}_3$ (figure adaptée de la thèse de Pierre Barraud)**

## *B - Généralités sur les expériences RMN*

Les expériences de RMN du proton ont été acquises sur un spectromètre RMN Bruker Avance DRX 600 équipé d'une cryosonde TXI (600 MHz). Les expériences du RMN du fluor ont été acquises sur un Bruker Avance 300 équipé d'une sonde QNP (282 MHz) et sont découplées du proton (séquence Waltz-16 avec $\gamma B_1$=1 kHz). Les déplacements chimiques du fluor sont calibrés par rapport au trifluoroéthanol (TFE) (0,1% v/v final est systématiquement ajouté aux échantillons). Pour les expériences de titration utilisant un mélange racémique de la sonde fluorée **DACPF**, on suit le signal $^{19}$F correspondant à l'énantiomère le plus déblindé.

## *C - Caractérisation des interactions entre les sondes fluorées et leurs cibles ribonucléiques*

### 1) Mesure en RMN $^{1}$H et RMN $^{19}$F de l'affinité de la sonde fluorée DACPF vis-à-vis de l'ARN 16S23

Pour déterminer l'affinité de la sonde fluorée DACPF vis-à-vis de l'ARN 16S23, nous avons mis en place deux expériences de titration par RMN complémentaires : suivi des signaux de la cible ARN en $^{1}$H RMN et suivi du signal fluor de la sonde en $^{19}$F RMN.

Pour les expériences en RMN du proton, le spectre de l'ARN 16S23 dans un tampon phosphate (0.2 mM dans 10 mM tampon KPO4, pH 6.5 et 50 mM KCl) a été enregistré à 293K en utilisant une séquence Watergate pour la suppression d'eau[375]. Les signaux des protons imino des bases $U_{16}$ et $G_{17}$ de l'ARN 16S23 sont suivis tout au long de l'addition progressive de sonde fluorée (11 points, de 0 µM à 2.6 mM). Leurs variations ont permis de déduire le Kd de la sonde fluorée pour l'ARN en utilisant un modèle non-linéaire à un site d'interaction (programme MC-fit[359]). Les valeurs des Kd sont indiquées avec un intervalle de confiance de 95 %.

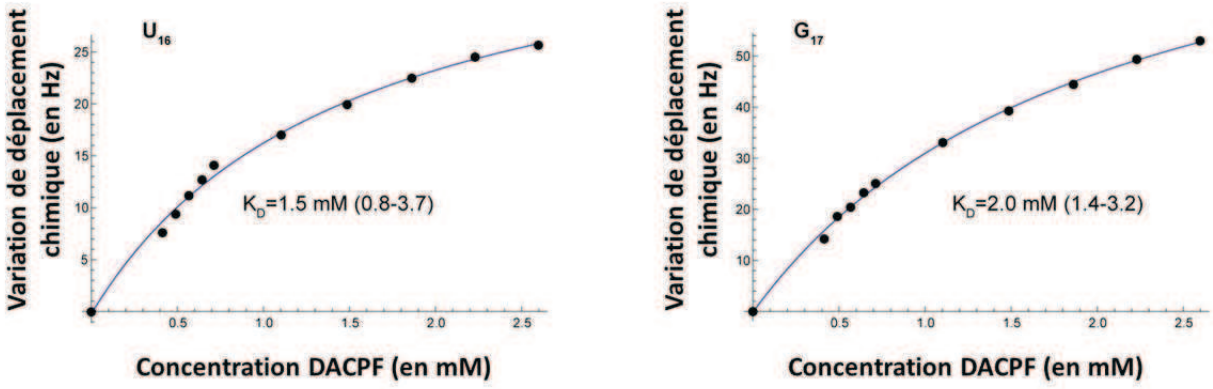



**Figure S-159 Mesure de l'affinité du DACPF pour l'ARN 16S23 par RMN $^{1}$H via la variation du déplacement chimique des protons imino des bases $U_{16}$ et $G_{17}$ lors de l'augmentation de la concentration en DACPF.**

Pour les expériences en RMN du fluor, l'ARN 16S23 (0.3 mM) a été mis dans un tampon phosphate (10 mM tampon KPO4, pH 6.5 et 50 mM KCl) puis mis en présence de concentrations croissantes de sonde fluorée (29 points, de 0 µM à 50 mM). Les variations du signal correspondant à l'énantiomère le plus déblindé ont été traités comme précédemment.

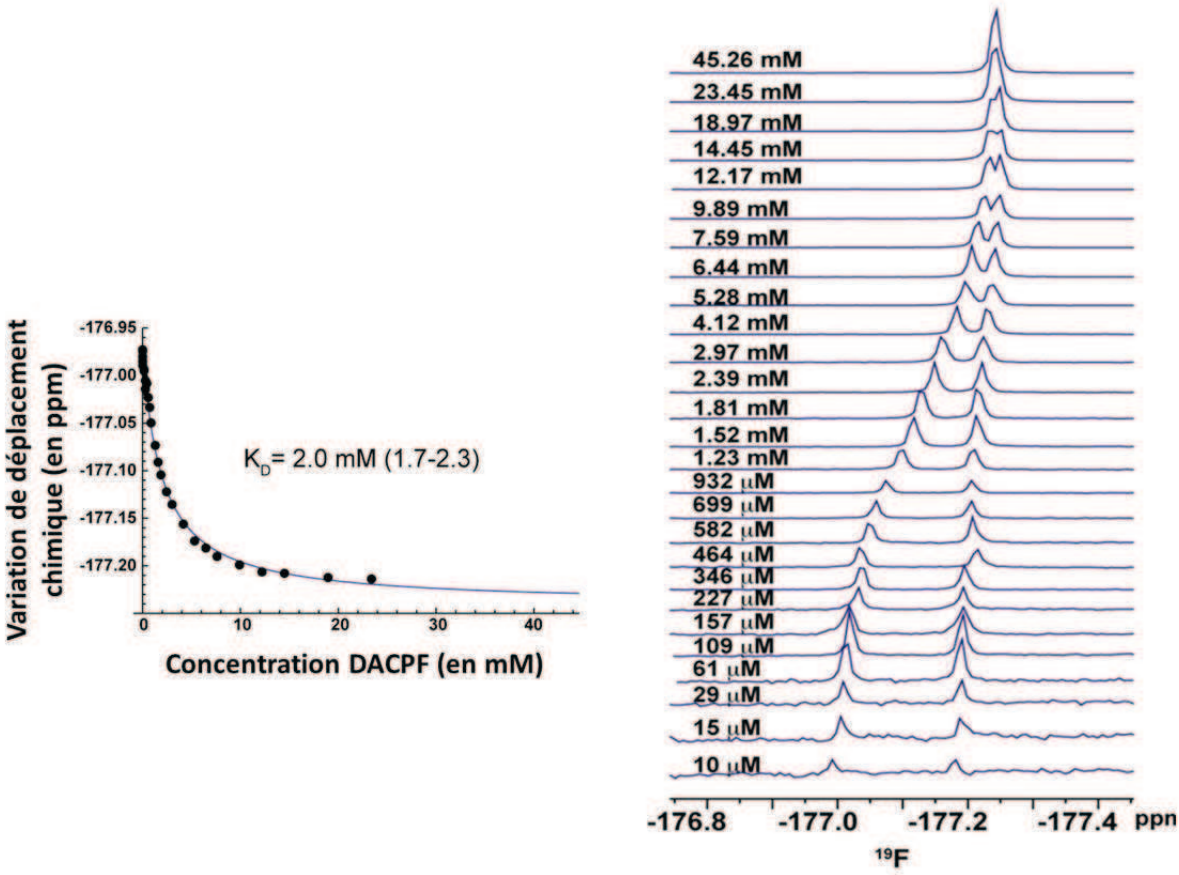



**Figure S-160 Mesure de l'affinité du DACPF pour l'ARN 16S23 par RMN $^{19}$F via la variation du déplacement chimique du fluor du DACPF lors de l'expérience de titration.**

## 2) Mesure en RMN $^{1}$H de l'affinité de la sonde fluorée DACPF vis-à-vis de l'aptamère à la néomycine

L'affinité du DACPF vis-à-vis de l'aptamère à la néomycine a été déterminée selon la même méthode que précédemment avec l'ARN 16S23. En RMN $^{1}$H, le spectre de l'ARN dans un tampon phosphate (0.2 mM dans 10 mM tampon KPO4, pH 6.5 et 50 mM KCl) a été enregistré à 293K en utilisant une séquence Watergate pour la suppression d'eau[375]. Le signal du proton imino de la base $G_{25}$ de l'aptamère est suivi tout au long de l'addition progressive de DACPF (11 points, de 0 µM à 3.3 mM). Leurs variations ont permis de déduire le Kd de la sonde fluorée pour l'ARN en utilisant un modèle non-linéaire à un site d'interaction (programme MC-fit[359]). Les valeurs des Kd sont indiquées avec un intervalle de confiance de 95 %.

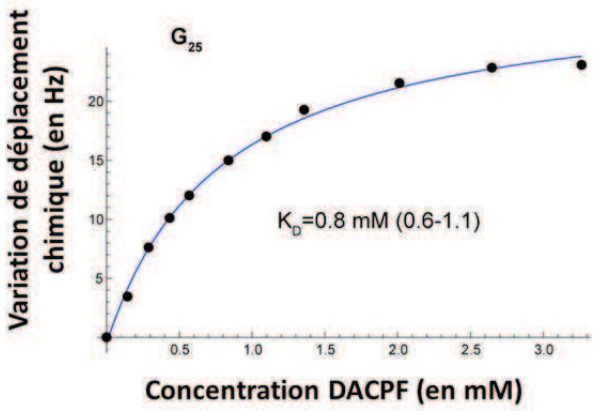



**Figure S-161 Mesure de l'affinité du DACPF pour l'aptamère à la néomycine par RMN $^{1}$H via la variation du déplacement chimique du proton imino de la bases $G_{25}$ lors de l'augmentation de la concentration en DACPF.**

### 3) Mesure en RMN $^{1}$H de l'affinité de la sonde fluorée ParoF vis-à-vis de l'ARN 16S23

L'affinité de la paromamine fluorée vis-à-vis de l'ARN 16S23 a été déterminée selon la même méthode que pour le DACPF en RMN $^{1}$H avec une séquence de suppression d'eau de type Watergate[375] (ARN (0.3 mM), 10 mM tampon KPO4, pH 6.5, 293K). La variation des protons imino des bases $U_{16}$ et $G_{17}$ au fur et à mesure de l'augmentation de la concentration de la paromamine fluorée (15 points, de 0 µM à 4.05 mM) permet la détermination du Kd de la paromamine fluorée pour l'ARN en utilisant le logiciel MC-fit[359]. Cette expérience a été effectuée à 3 concentrations différentes en sels (50 mM, 150 mM et 300 mM KCl).

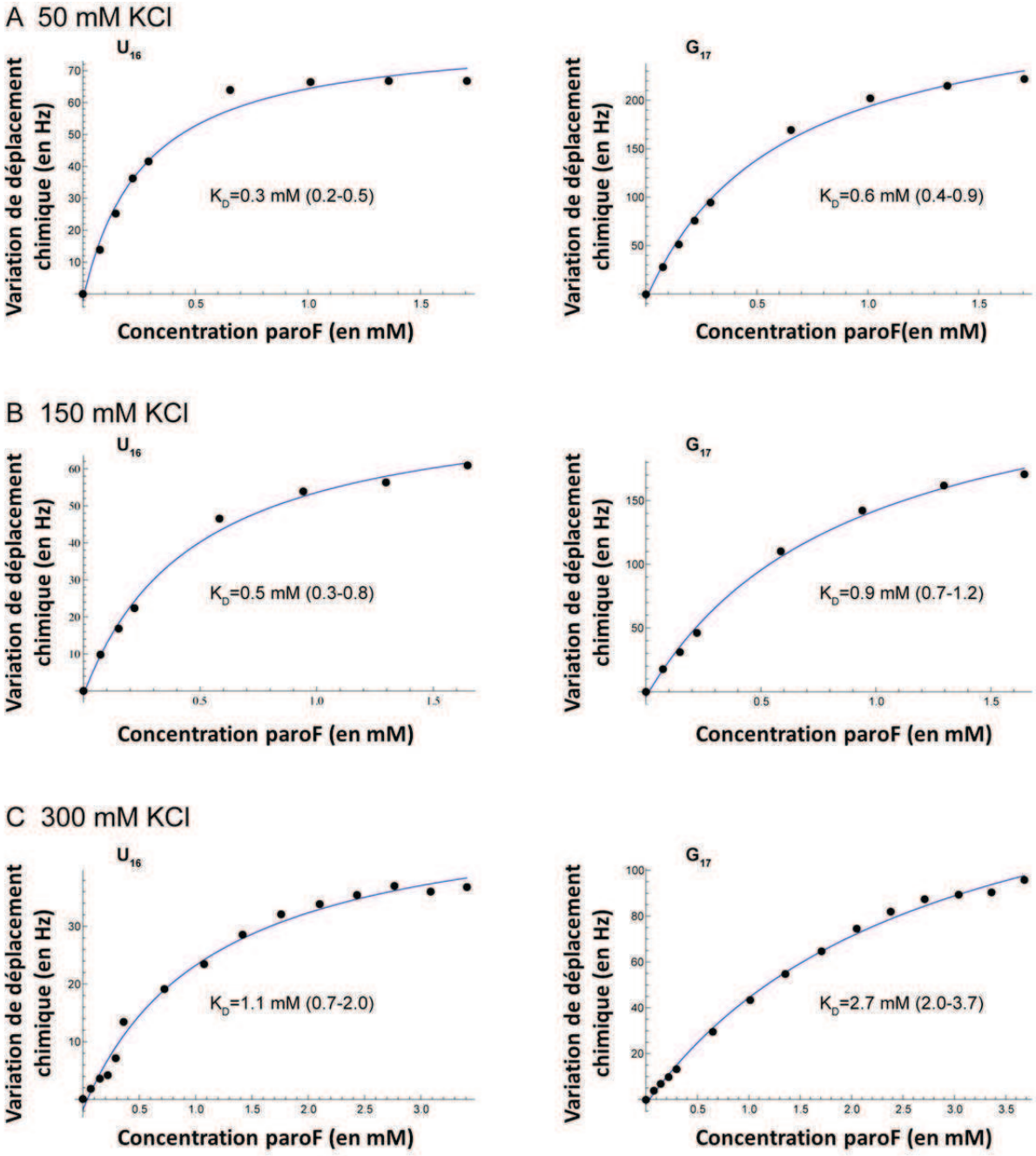



**Figure S-162 Mesure de l'affinité de la paromamine fluorée pour l'ARN 16S23 par RMN $^{1}$H via la variation du déplacement chimique des protons imino des bases $U_{16}$ et $G_{17}$ lors de l'augmentation de la concentration en paromamine fluorée.**

### 4) Analyse de l'interaction DACPF/ARN 16S23 par NOESY et TOCSY

L'ARN 16S23 a été préparé dans un tampon entièrement deutéré $D_2O$ (10 mM $KPO_4$ pH 6.5, 50 mM KCl) à une concentration de 1.5 mM. Une série d'expérience NMR TOCSY (temps de mélange 45 ms) et NOESY (temps de mélange 300, 400 et 600 ms) ont été enregistré à deux températures différentes, 293 K et 278 K, avec ou sans la sonde fluorée DACPF préalablement solubilisée dans $D_2O$. Nous avons testé plusieurs concentrations de sonde fluorée ( 2.25 mM, 4.5 mM et 9 mM).

Les signaux RMN des protons H-8/H-6/H-5/H-1' de l'ARN 16S23 seul ont tout d'abord été entièrement réattribués et sont en accord avec la littérature[241]. Pour chaque nouvelle expérience, l'intégralité des signaux sont attribués de façon à pouvoir suivre leurs variations, à l'exception des signaux de la guanidine $G_{23}$ qui n'ont pas pu être attribués. Pour chaque nucléotide de l'ARN 16S23, la perturbation globale de déplacement chimique (CSP pour « overall chemical shift perturbation ») est calculée selon la formule suivante :

$$CSP = \sqrt{(\Delta\delta_{H8-H6})^2 + (\Delta\delta_{H5})^2 + (\Delta\delta_{H1'})^2}$$

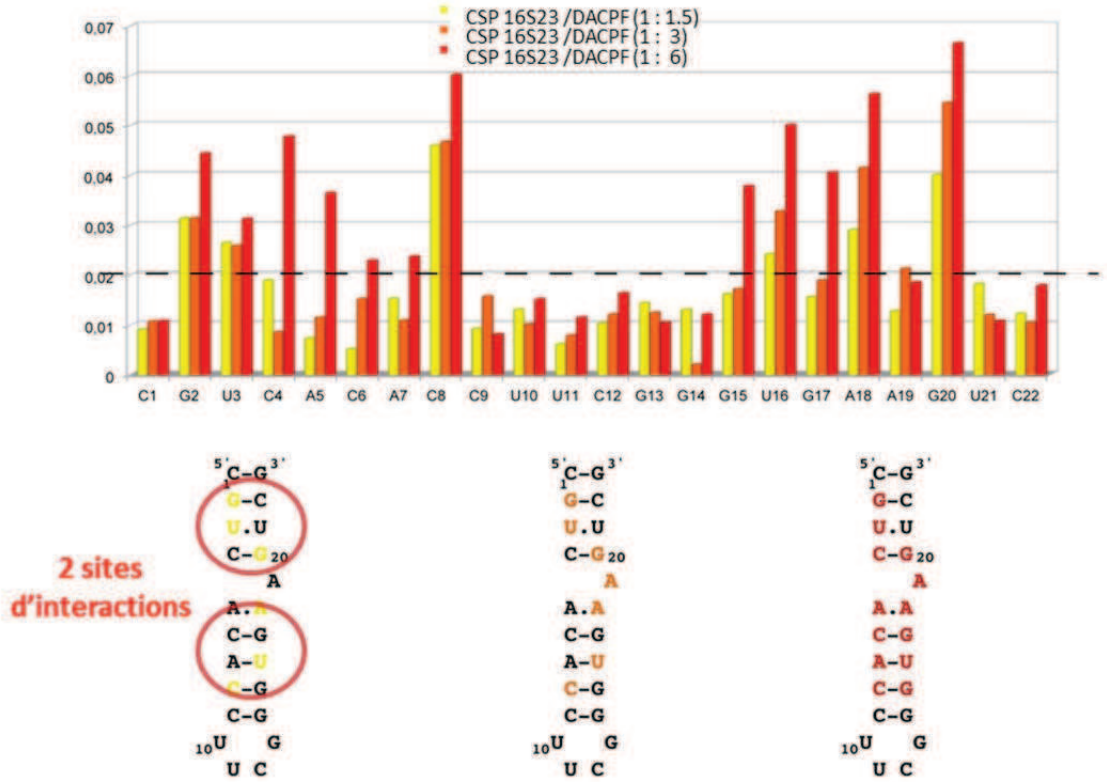



**Figure S-163 Analyse CSP des sites d'interaction du DACPF Avec l'ARN 16S23**

## *D - Suivi de la capture conformationnelle de l'aptamère par la néomycine*

L'aptamère à la néomycine (0.3 mM) dans un tampon phosphate ($KHPO_4$ 10mM, 50 mM KCl, 0.1 % TFE v/v, pH 6.5) est mis en présence de DACPF (0.7 mM final, rapport 1 : 2.3) à 293K. On ajoute alors de la néomycine (0.45 mM final, rapport 1 : 1.5).

## *E - Expériences de compétition sonde fluorée/ligand/ARN 16S23 ou 16S27*

### 1) Conditions expérimentales des expériences de compétition

Pour le DACPF, les conditions sont les suivantes : ARN (0.3 mM), DACPF (0.7 mM, rapport 1 : 2.3), TFE 1% v/v, tampon phosphate 10 mM, pH 6.5, 50 mM KCl, 293K. On ajoute des concentrations croissantes de ligands compétiteurs jusqu'à coalescence du doublet du signal du fluor du DACPF en RMN $^{19}F$. On arrête la titration lorsque le signal observé en RMN $^{19}F$ correspond au signal de la sonde libre (δ∼-177.3 ppm). Dans certains cas, principalement pour les titrations de ligands de faible affinité, l'apparition d'un précipité blanc à de fortes concentrations en ligand (précipitation du complexe ARN/ligand ?) contraint d'arrêter la titration prématurément. Un spectre peut être accumulé en 128 scans (environ 3 min d'accumulation sur le spectromètre RMN 300 MHz utilisé).

Pour la paroF, on utilise les conditions suivantes : ARN (0.3 mM), paroF (0.7 mM), TFE 1% v/v, tampon phosphate 10 mM, pH 6.5, 300 mM KCl, 293K. On ajoute des concentrations croissantes de ligand compétiteur jusqu'à obtenir un signal de la sonde paroF correspondant à la sonde libre (δ∼-233.9 ppm). Un spectre peut être accumulé en 512 scans (environ 12 min d'accumulation sur le spectromètre RMN 300 MHz utilisé).

### 2) Fit des expériences de compétition

La fraction de la quantité de sonde fluorée liée à la cible ARN, $\left[{}^{19}F_Probe_{Bound}\right]/\left[{}^{19}F_Probe_{Total}\right]$, en présence d'un ligand compétiteur est donnée par l'équation (1) suivante[342360] :

$$\frac{\left[{}^{19}F_Probe_{Bound}\right]}{\left[{}^{19}F_Probe_{Total}\right]} = \frac{2\sqrt{(a^2-3b)}\,\cos(\theta/3) - a}{3K_D + 2\sqrt{(a^2-3b)}\,\cos(\theta/3) - a} \qquad (1)$$

où : $\theta = ar\ \cos\left[\frac{-2a^2+9ab-27c}{2\sqrt{(a^2-3b)^3}}\right]$, avec :

$$a = K_D(^{19}F_probe) + K_D(ligand) + \left[^{19}F_Probe_{Total}\right] + \ [RNA_ligand] - [RNA_{total}]$$

$$b = \{[I] - [RNA_{total}]\}K_D + \left\{\left[^{19}F_Probe_{Total}\right] - [RNA_{total}]\right\}K_{D_ligand} + K_D(^{19}F_probe)K_D(ligand)$$

$$c = -K_D(^{19}F_probe)K_D(ligand)[RNA_{total}]$$

$K_D(^{19}F_probe)$ et $K_D(ligand)$ sont, respectivement, les constantes de dissociations de la sonde fluorée et du ligand vis-à-vis de l'ARN. $\left[^{19}F_Probe_{Total}\right]$ et $[RNA_ligand]$ sont les concentrations totales de la sonde fluorée et du ligand. Enfin, $[RNA_{total}]$ est la concentration totale des sites d'interaction entre la sonde et l'ARN. Cette concentration est un multiple entier de la concentration totale d'ARN (par exemple, cette concentration correspond à deux fois la concentration totale en ARN si il existe deux sites d'interaction).

Cette fraction est accessible expérimentalement car elle est directement liée au déplacement chimique du signal du fluor de la sonde, noté $\delta\left[^{19}F_Probe_{Total}\right]$. Etant dans des conditions d'échange rapide, ce signal correspond en effet à la moyenne du signal de la sonde libre ($\delta\left[^{19}F_Probe_{Free}\right]$) et de la sonde liée ($\delta\left[^{19}F_Probe_{Bound}\right]$), pondérée de leur concentration respective (équation (2)).

$$\delta\left[^{19}F_Probe_{Total}\right] = \frac{\delta\left[^{19}F_Probe_{Free}\right] * \left[^{19}F_{Probe_{Free}}\right] + \delta\left[^{19}F_{Probe_{Bound}}\right] * \left[^{19}F_Probe_{Bound}\right]}{\left[^{19}F_Probe_{Total}\right]} \quad (2)$$

L'équation (1) est rentrée dans le logiciel Mathematica (Wolfram) et la solution réelle de cette équation est calculée à partir de la variation des déplacements chimiques de la sonde lors de l'ajout de ligand compétiteur ($\delta\left[^{19}F_Probe_{Total}\right]$). Les constantes fournies pour résoudre le modèle sont : $K_D$ ($^{19}$F_probe), $\left[^{19}F_Probe_{Total}\right]$.

Dans les expériences de titrations utilisant la sonde DACPF, un Kd de 2 mM a été choisi (tampon phosphate 10 mM tampon KPO4, pH 6.5 et 50 mM KCl). Pour les expériences de titrations utilisant la paromamine fluorée, un Kd de 1.1 mM a été choisi (tampon phosphate 10 mM tampon KPO4, pH 6.5 et 300 mM KCl).

$[RNA_{total}]$ est un des résultats du modèle et peut donc être utilisé pour déterminer le caractère vraisemblable des résultats trouvés.

Voici un exemple de l'interface développée dans Mathematica :

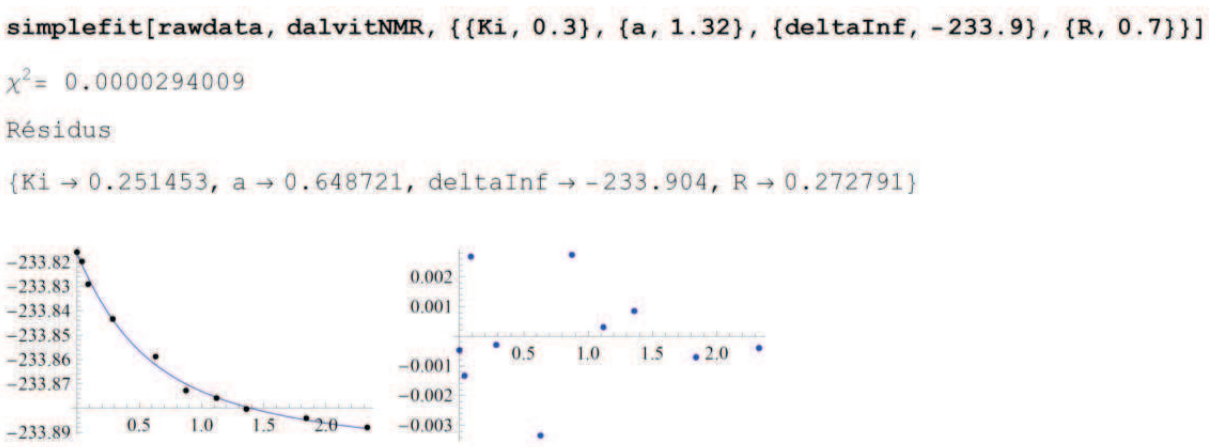



**Figure S-164 Interface de l'algorithme développé dans Mathematica**

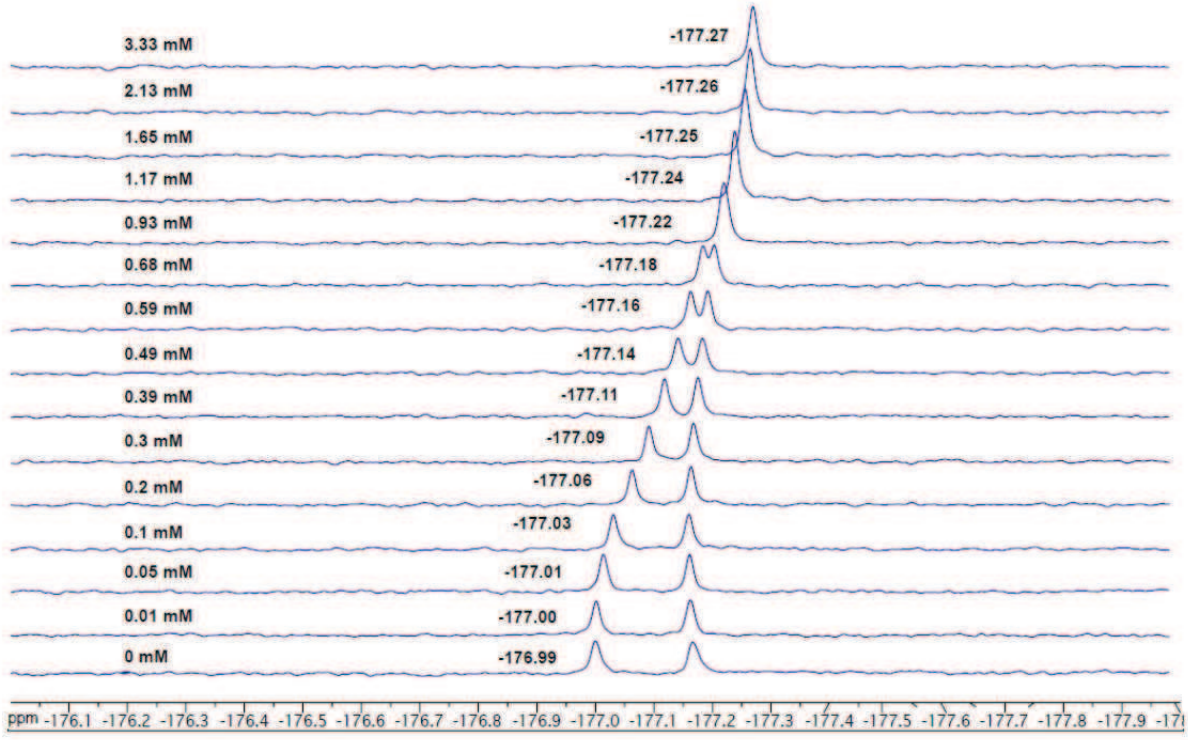
Le modèle (nommé simplefit) utilise trois données de départ : 2 constantes, la concentration du ligand fluorée (notée F), son affinité pour l'ARN (notée Keq) et une matrice, la variation du signal du fluor en fonction de la concentration en ligand compétiteur (notée rawdata). Il donne 4 valeurs utilisables en sortie.

Ki est la valeur que l'on cherche initialement à obtenir grâce au modèle. Elle correspond au Kd trouvé pour le ligand compétiteur (ici la néamine). $\chi^2$ représente la qualité de fit des données, plus la valeur est faible, plus le modèle utilisé est en adéquation avec les données expérimentales. Les résidus sont présentés à droite. Ils correspondent aux écarts des données expérimentales avec le modèle. Ceci permet de voir s'il y a un profil de déviation par rapport au modèle, qui serait lié à une erreur systématique. Enfin, deltainf et R sont des données déduites par le modèle qui correspondent directement à des données expérimentales. Elles peuvent donc servir de contrôle interne au modèle. Deltainf correspond au déplacement chimique du fluor obtenu en fin d'expérience de compétition quand l'intégralité de la sonde fluorée a été déplacée. R correspond à la concentration en ARN

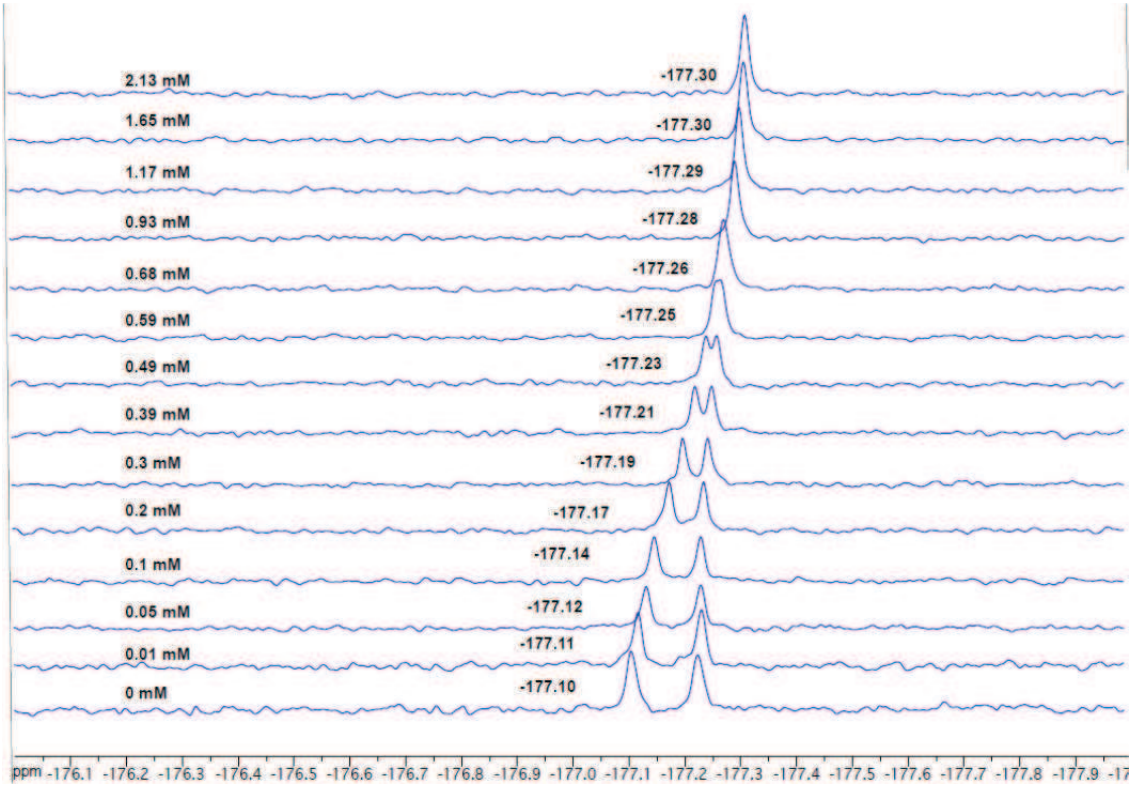
apparente. Si la sonde fluorée possède plusieurs sites d'interaction sur l'ARN et si le ligand compétiteur déplace l'ensemble des molécules de sondes fluorées lors de son interaction, R est un multiple de la concentration initiale en ARN dans l'échantillon correspondant au nombre de sites de la sonde fluorée sur l'ARN.

## 3) Spectres RMN des différentes expériences de compétition

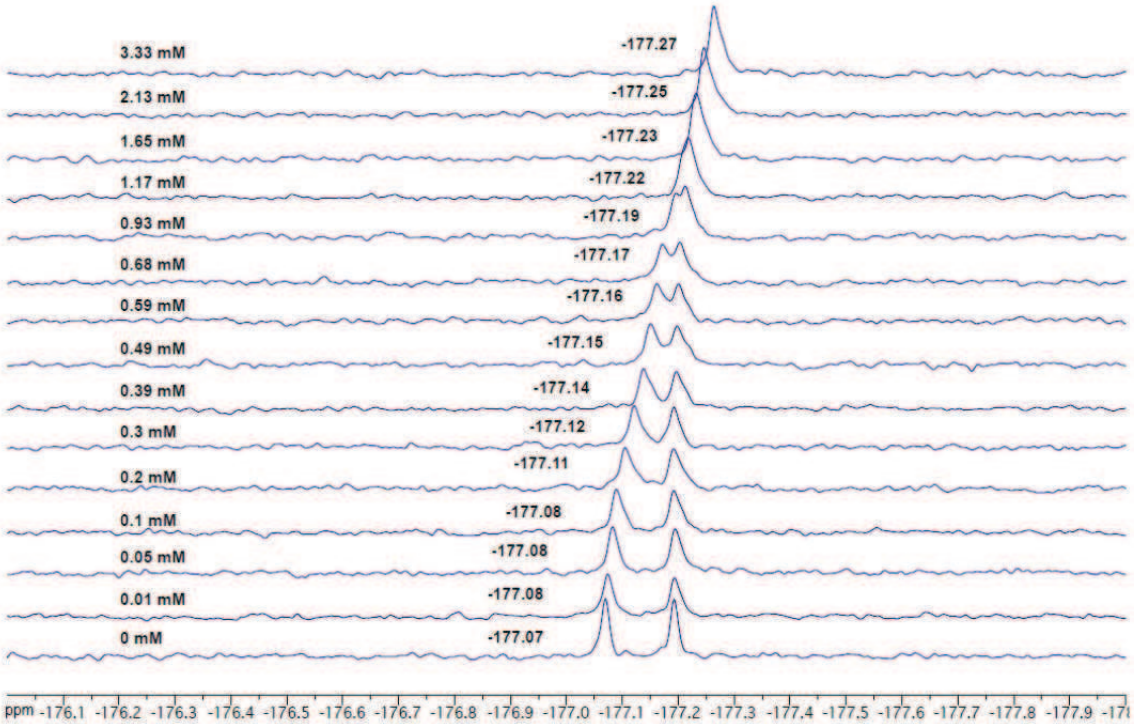
### a) DACPF/néamine/16S23



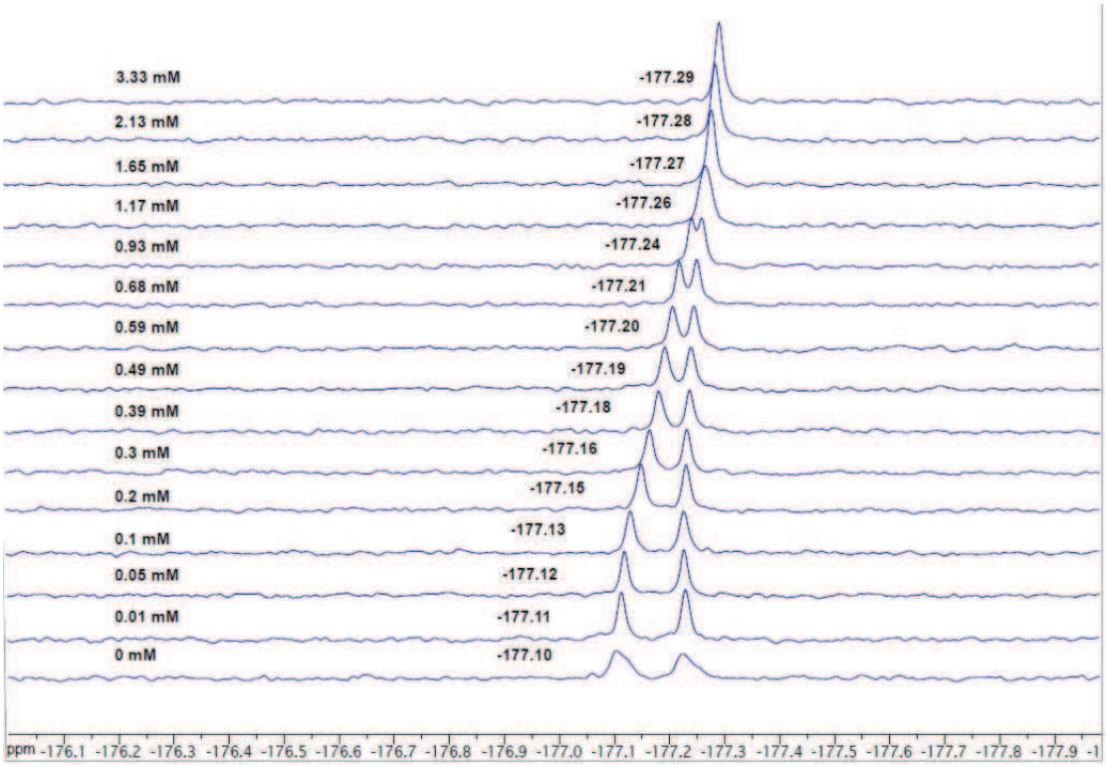
**b) DACPF/néamine/16S23 n°2**



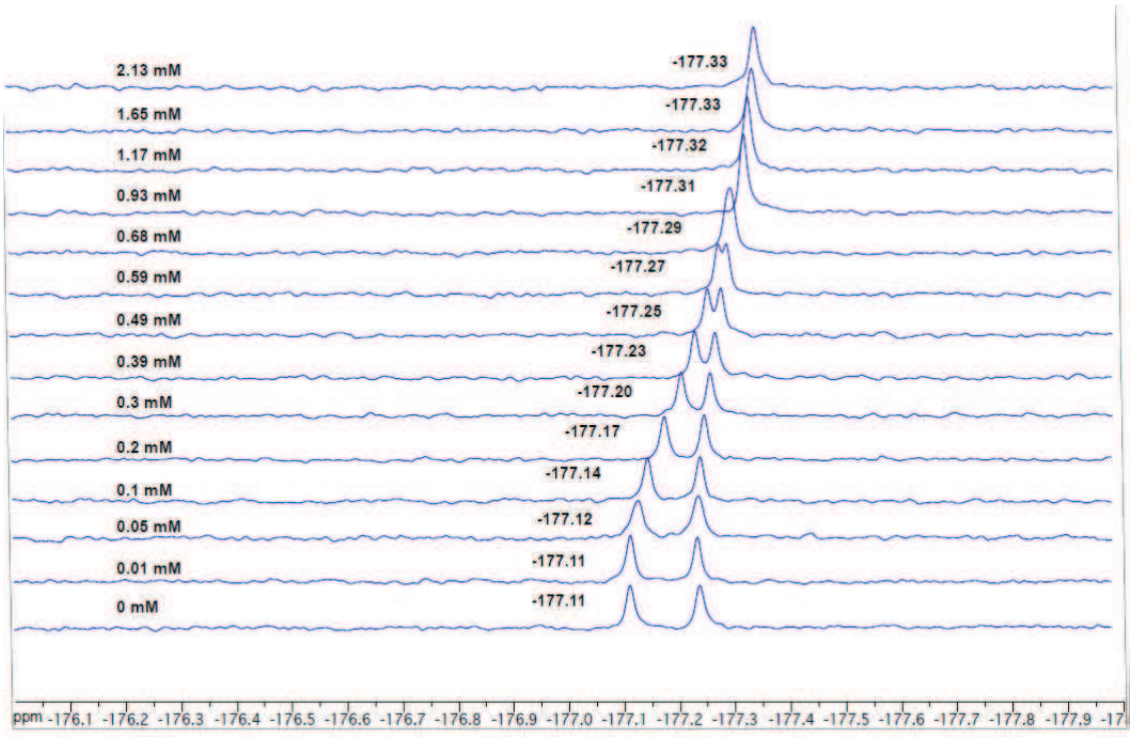

**c) DACPF/paromamine/16S23**



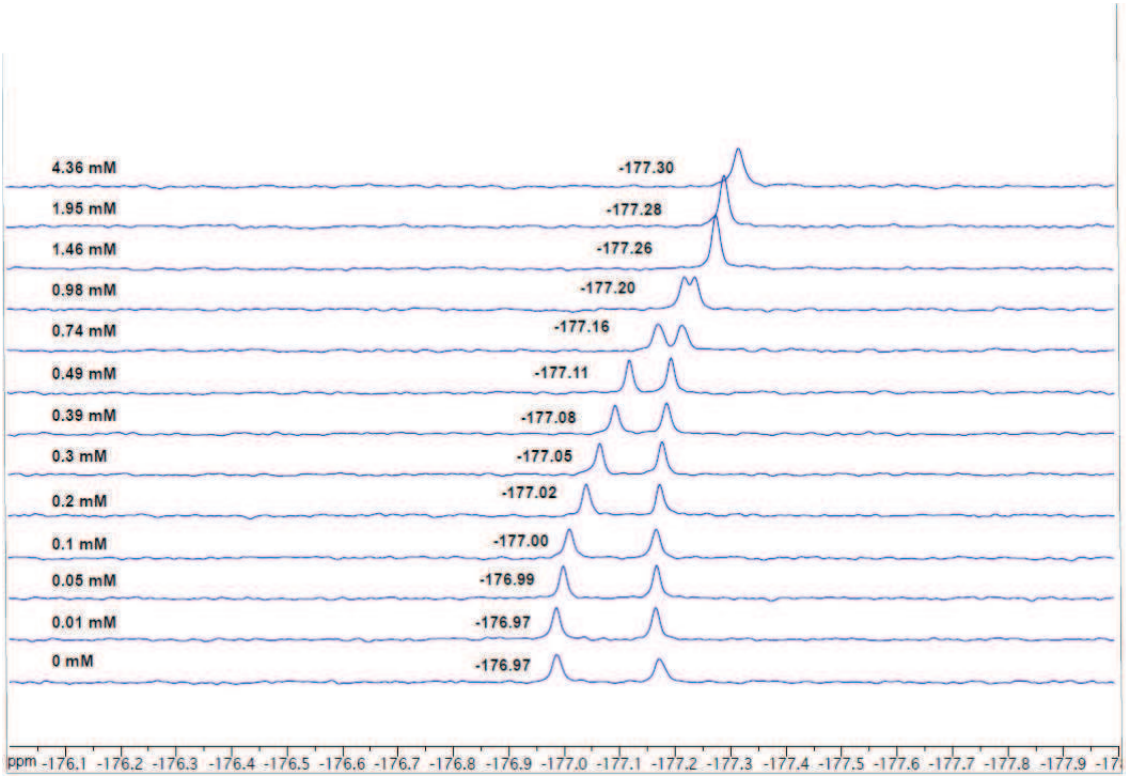
**d) DACPF/paromamine/16S23 n°2**



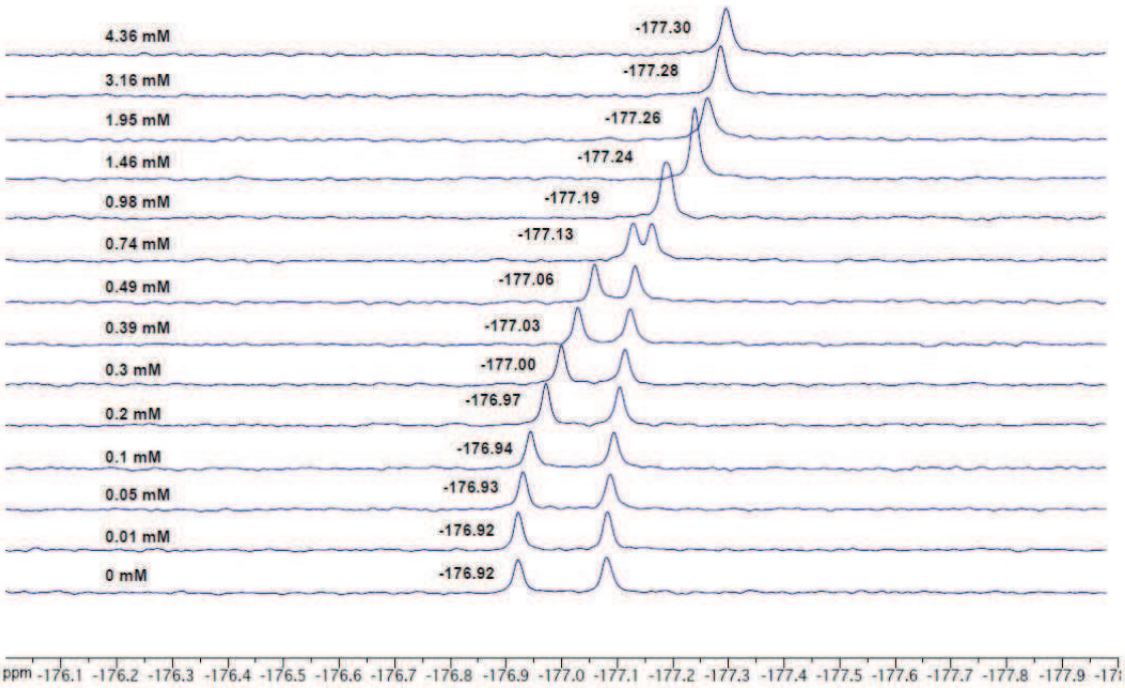

**e) DACPF/néomycine/16S23**



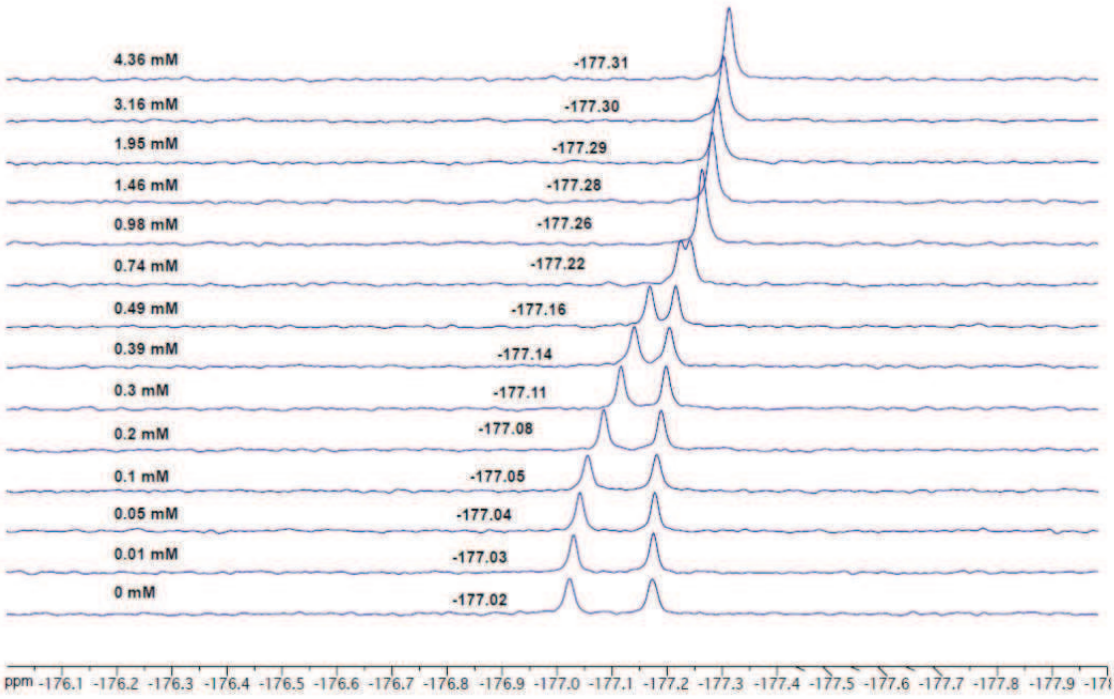
**f) DACPF/néomycine/16S23 n°2**



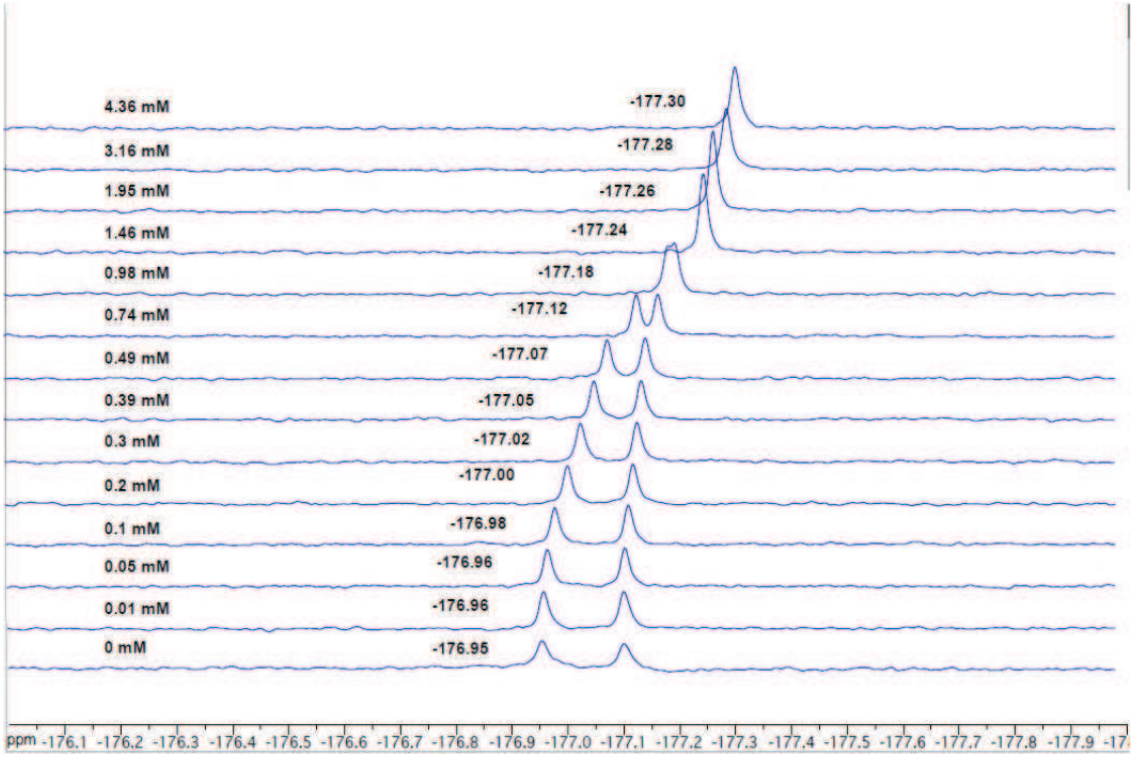

**g) DACPF/néomycine/16S27**



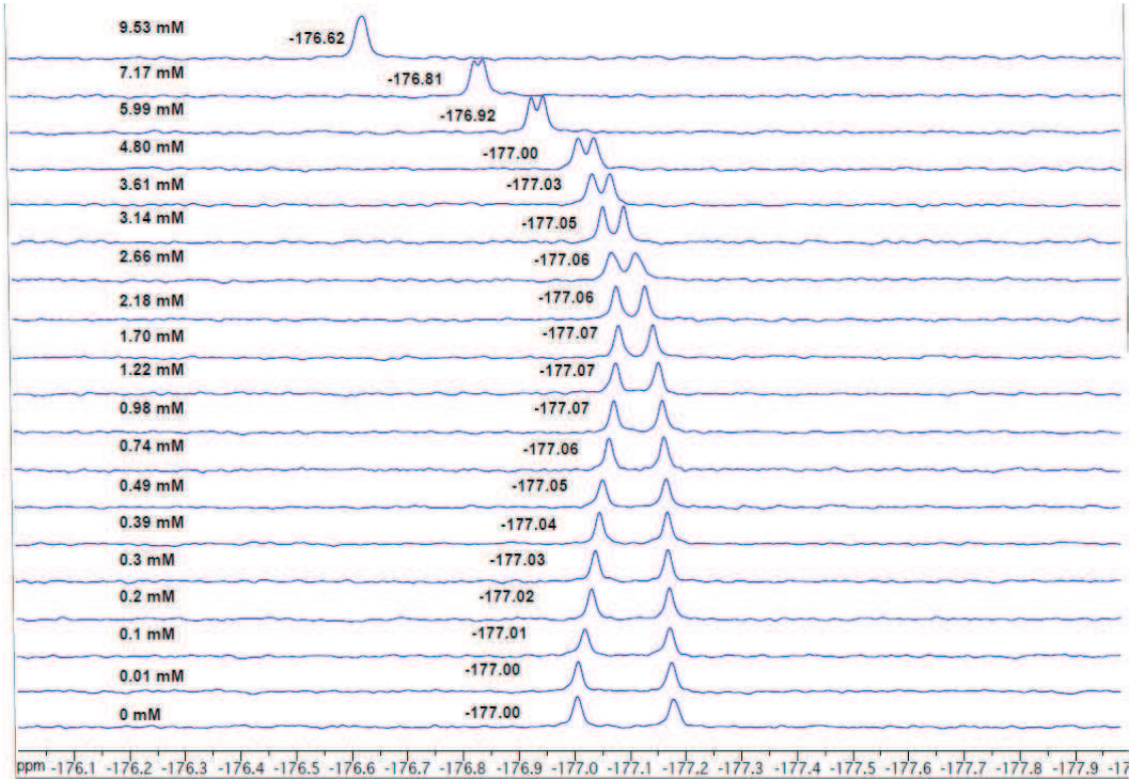
**h) DACPF/paromomycine/16S23**



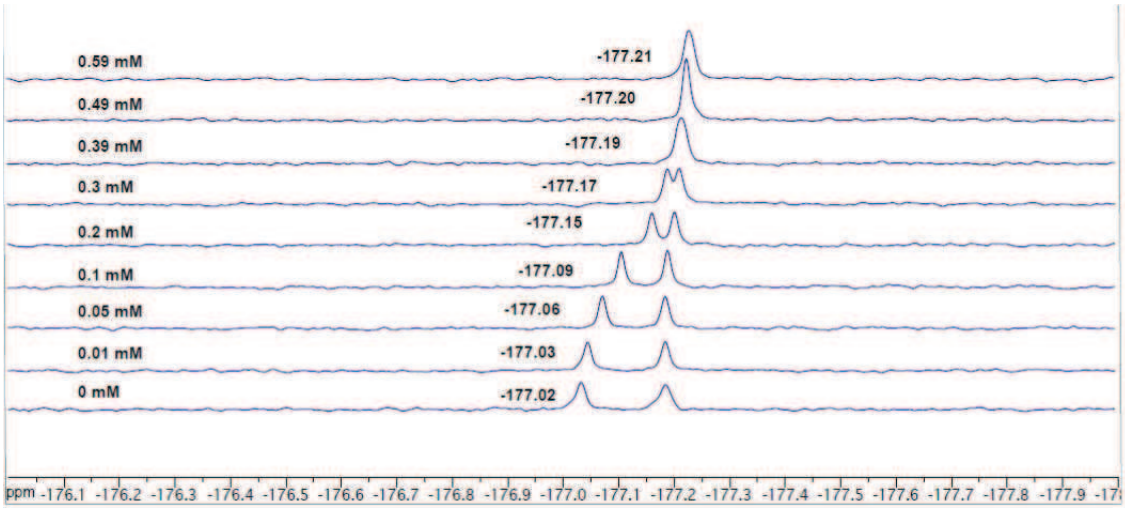

**i) DACPF/paromomycine/16S27**



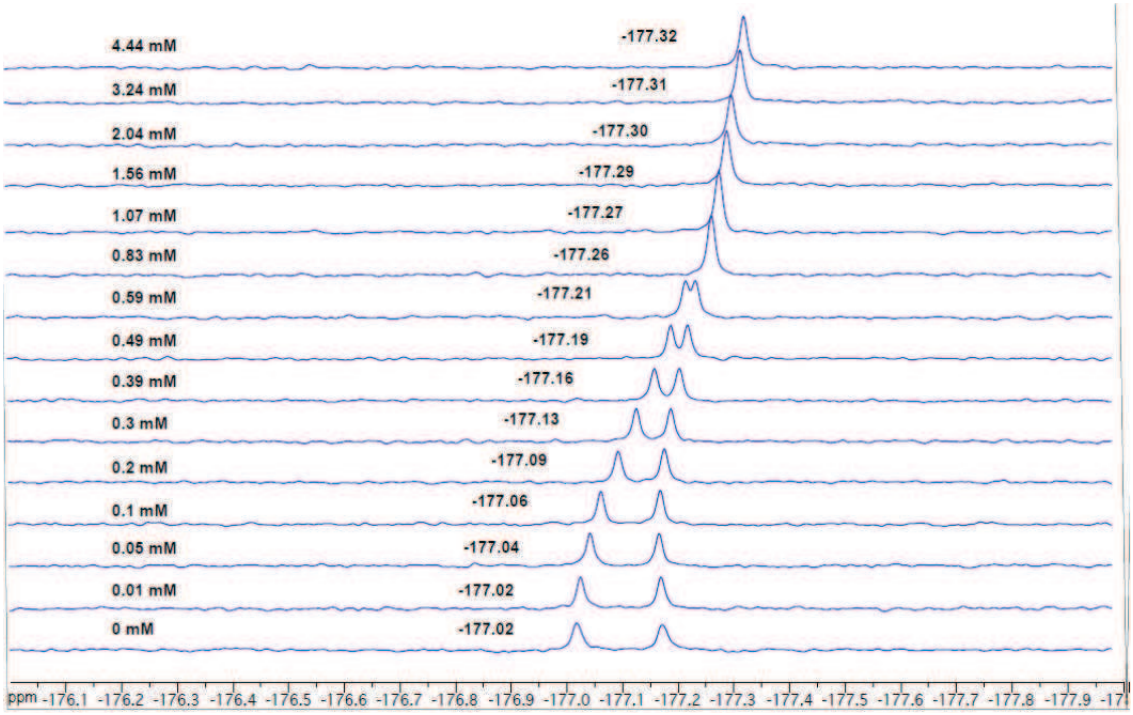
**j) DACPF/DOS/16S23**



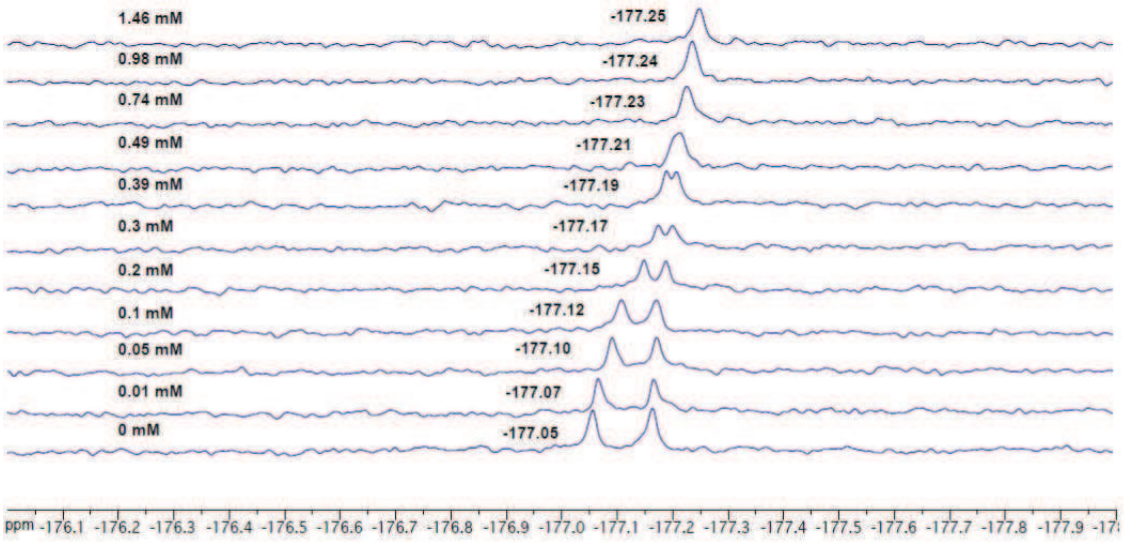

**k) DACPF/Kanamycine/16S23**



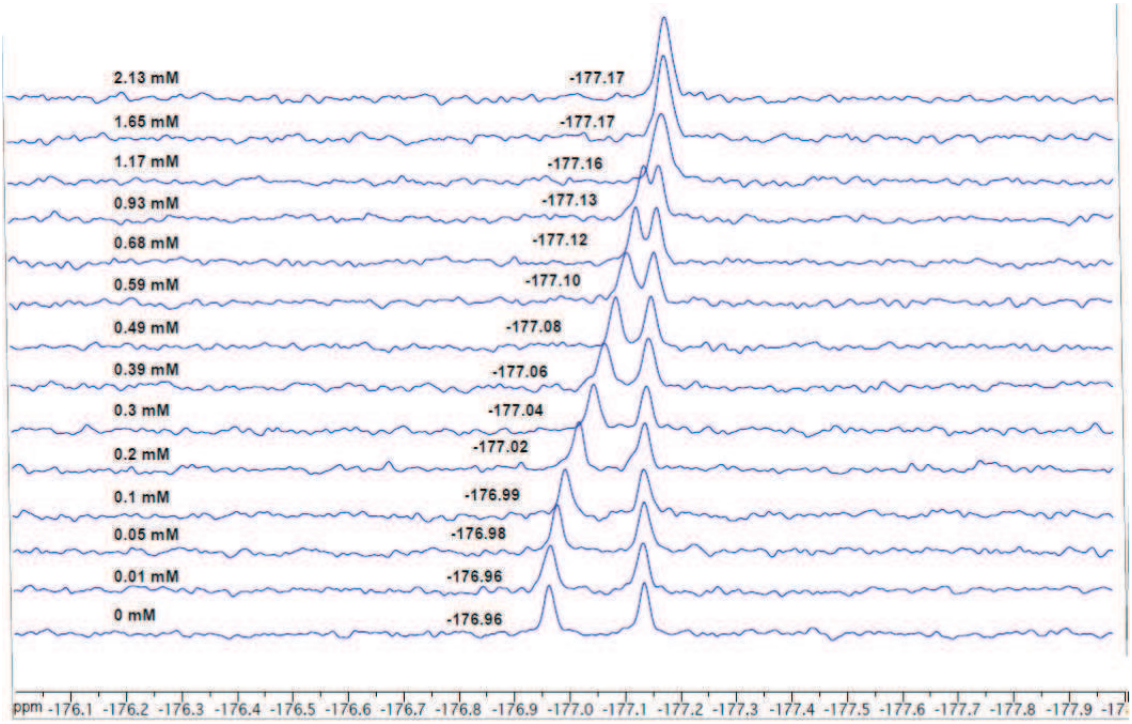
l) DACPF/Gentamicine/16S23



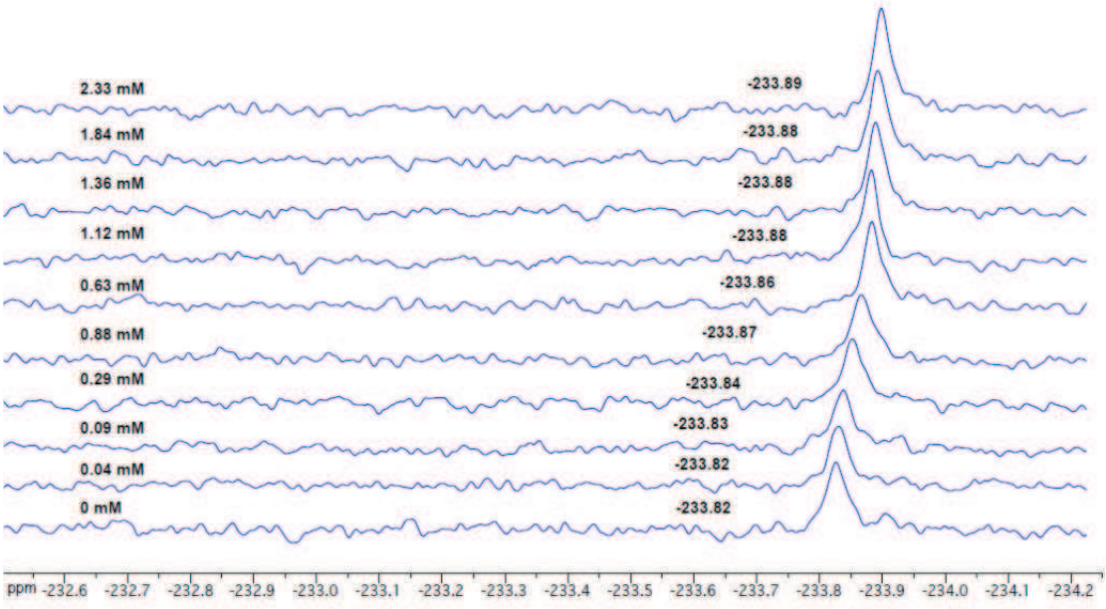

m) DACPF/Streptomycine/16S23



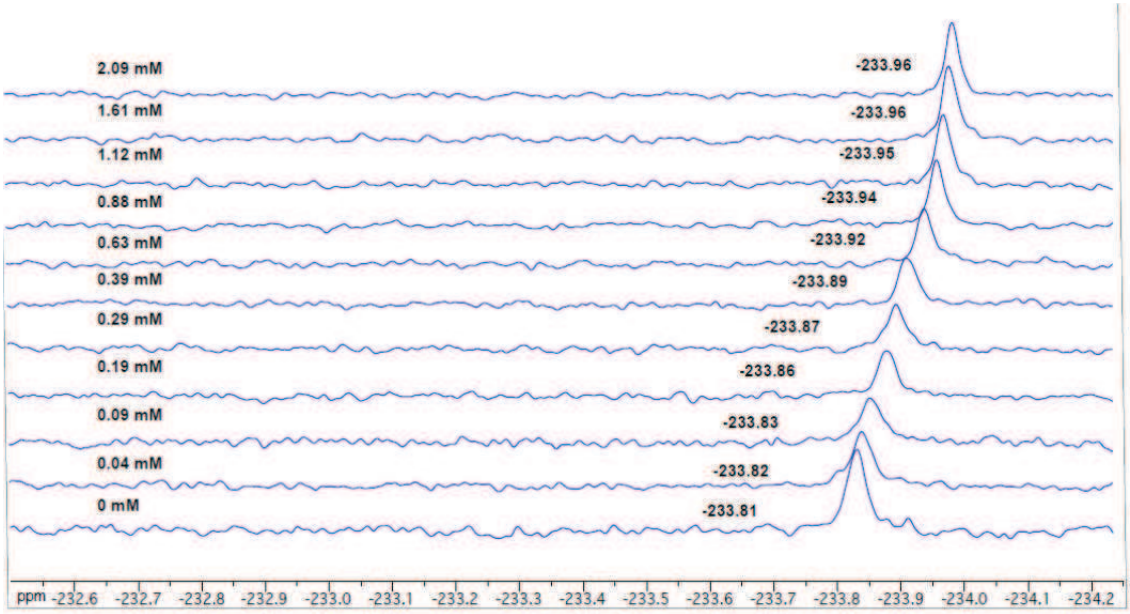
**n) DACPF/Ribostamycine/16S23**



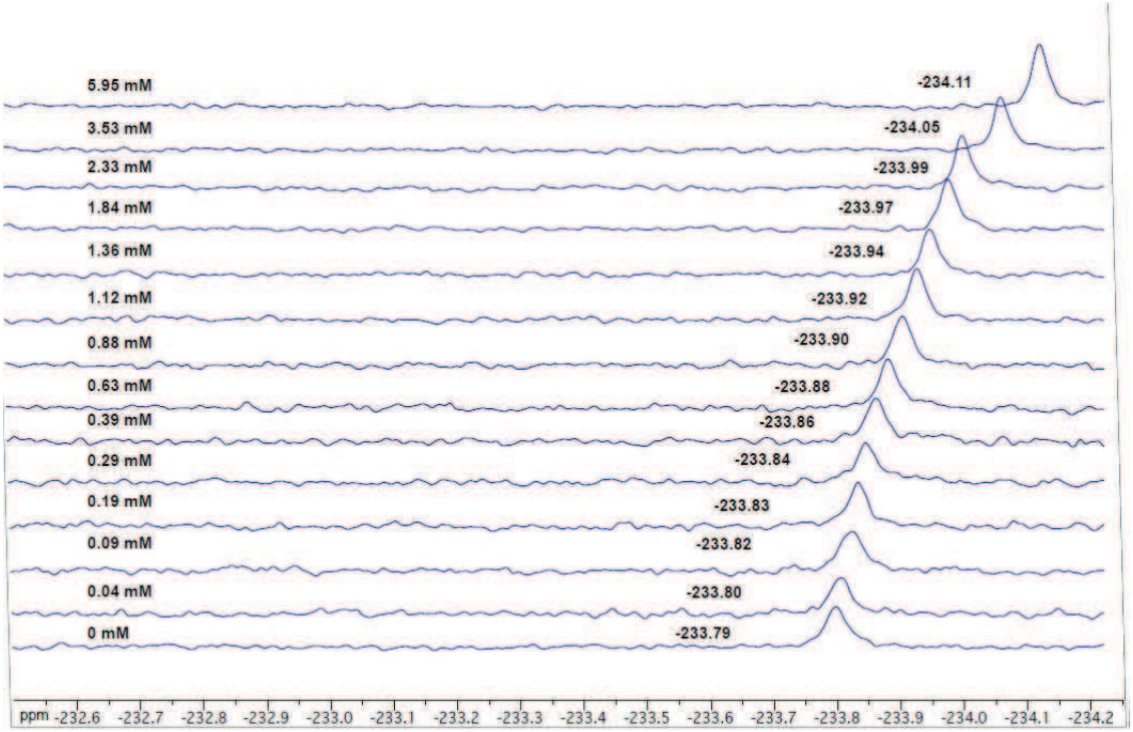

**o) ParoF/néamine/16S23**



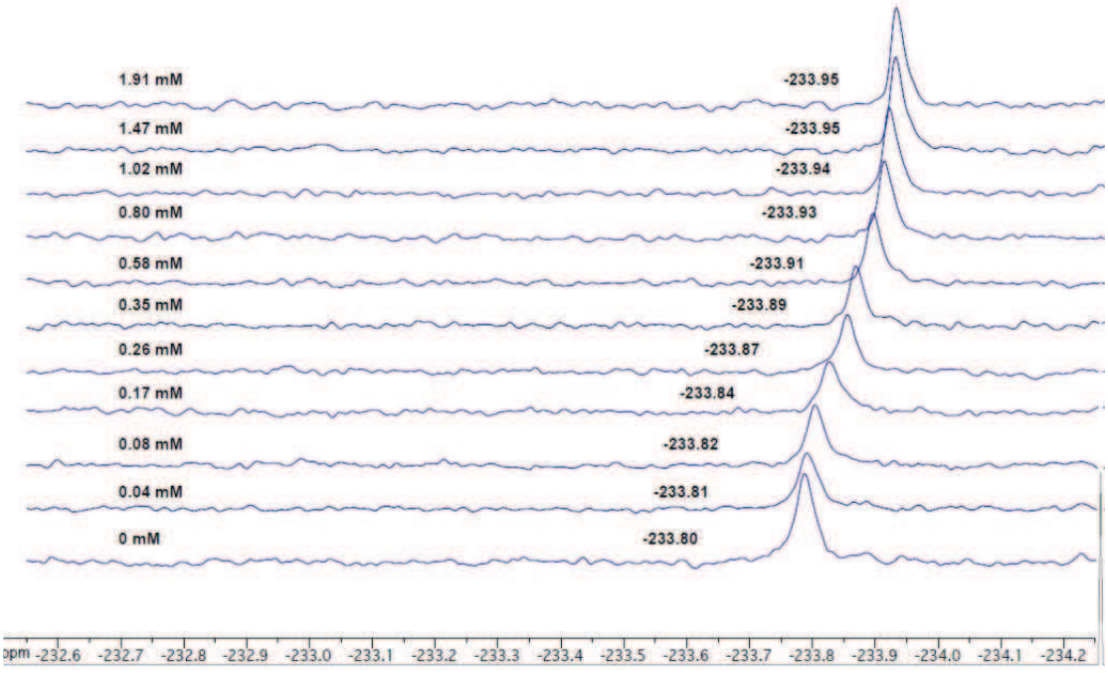
**p) ParoF/paromomycine/16S23**




**q) ParoF/ribostamycine/16S23**



**r) ParoF/néomycine/16S23**



# III - Partie expérimentale : Synthèse organique

## *A - Généralités*

Les solvants utilisés lors des réactions nécessitant des conditions anhydres sont soit achetés anhydres (Pyridine, dimethylformamide, DME) soit distillés sur sodium-benzophénone (THF).

Les chromatographies sur couche mince sont réalisées sur plaque de silicagel de 0,2 mm d'épaisseur sur support d'aluminium (Merck 60 F254), puis observées à la lumière UV et révélées par une solution d'acide phosphomolybdique, de permanganate de potassium ou de ninhydrine.

Les purifications sur silice sont réalisées selon des techniques de chromatographie éclair, sous moyenne pression, avec de la silice SDS 60 A C.C 35-70 µm.

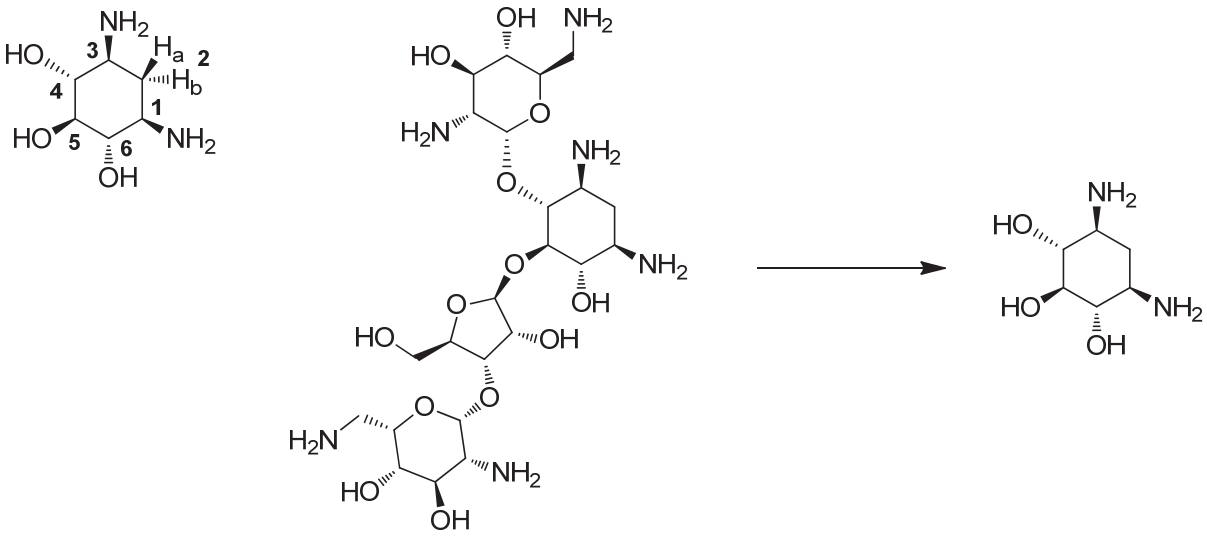
Les spectres RMN des produits synthétisés sont enregistrés sur un appareil Bruker Avance 300 (300 MHz en proton, 75 MHz en carbone 13) et sur un appareil Bruker Avance 400 (400 MHz en proton, 100 MHz en carbone 13). Les déplacements chimiques (δ) sont exprimés en ppm par rapport au tetraméthylsilane (TMS) ou par rapport au résidu de solvant non deutéré. Les constantes de couplage sont exprimées en Hertz (Hz). Les abréviations suivantes sont utilisées pour la multiplicité des signaux : s (singulet), d (doublet), t (triplet), q (quadruplet), m (multiplet ou massif). Sauf indication contraire, les spectres ont été enregistrés à 293 K. Certains spectres ont été enregistrés à des températures plus élevés en raison de la propriété de certaines molécules à être en mélange complexe de rotamères à 293 K, conduisant à des signaux très élargis (noté b dans les descriptions RMN). L'ensemble des structures ont été attribuées à l'aide d'expériences de DEPT, COSY, HSQC et HMBC.

Les spectres de masse (MS) par ionisation chimique sont réalisés sur un appareil ZQ 2000 Waters muni d'une source d'électrospray. Les spectres de masse haute résolution (HRMS) ont été obtenus à l'aide d'un spectromètre ZABSpec TOF de Micromass possédant une géométrie EBE TOF du Centre Régional de Mesures Physiques de l'Ouest.

Les analyses HPLC sont réalisées sur une colonne C18 Stability 100/5µM (Cluzeau). La méthode d'analyse développée, identique pour tous les produits, est la suivante : gradient 20/80 → 90/10 acétonitrile/eau en 20 minutes puis palier de 20 min à 90/10 acétonitrile/eau. Le débit utilisé est de 1mL/min et la détection UV est effectuée à 190 nm.

## *B - Synthèse d'aminoglycosides naturels par dégradation*

### 1) Desoxystreptamine DOS 1



**Formule brute** : $C_6H_{14}N_2O_3$

**M** : 162 g.mol$^{-1}$ (324 g.mol$^{-1}$ sous forme de bromhydrate)

**Mode opératoire :**

Dans un ballon de 2L, on dissout 60 g de sulfate de néomycine (66 mmol ; 1 eq.) dans 300 mL de solution aqueuse d'acide bromhydrique à 48 %. La solution est portée à reflux et agitée pendant 3 jours. Après avoir refroidi la solution, on tire à sec au rotavapor de façon à obtenir une épaisse huile noire. On reprend cette huile dans 100 mL d'eau et on tire de nouveau à sec. On répète l'opération une fois. Après avoir repris le résidu restant dans 100 mL d'eau, on filtre sur célite. Le filtrat est récupéré, traité au charbon actif puis de nouveau filtré sur célite. On obtient alors une solution jaune clair très acide (pH = 1) à laquelle on ajoute des pastilles de soude NaOH de façon à arriver à un pH basique (pH ≈ 9-10) en prenant garde au dégagement de chaleur dû à la réaction acido-basique. Le changement de pH se traduit par un changement de la couleur de la solution, qui devient alors verte sombre et légèrement trouble (précipité de NaBr).

En parallèle, on prépare la phase Amberlyst A-15. Après avoir fait gonfler la phase dans l'eau pendant quelques dizaines de minutes (40 g de phase pour 60 g de néomycine de départ), on rince abondamment à l'eau jusqu'à obtenir un surnageant incolore de pH neutre. On lave ensuite la phase avec une solution d'ammoniaque ($NH_4OH$ 4M) jusqu'à l'obtention d'un lavis incolore lui aussi. Après avoir de nouveau rincé abondamment la phase à l'eau, on la régénère en l'incubant pendant 1 heure avec 100 mL d'une solution aqueuse d'acide chlorhydrique (HCl 6M) puis en la rinçant de nouveau abondamment à l'eau.

La phase Amberlyst A-15 peut alors être ajoutée au mélange réactionnel pré-purifié, ce qui conduit à un mélange de pH légèrement acide. Après une dizaine de minutes d'agitation douce, le mélange est filtré et la phase est lavée plusieurs fois à l'eau. La désoxystreptamine est éluée avec 2 x 50 mL d'une solution d'ammoniaque ($NH_4OH$ 4M). L'éluat est évaporé puis longuement séché à la pompe pour donner 5,5 g de désoxystreptamine sur forme de base (34 mmol, 52 %). La désoxystreptamine a pu être recristallisée à partir d'un mélange eau/ éthanol (9 : 1) pour former des aiguilles blanches (environ 50 mg de DOS pour 2 mL).

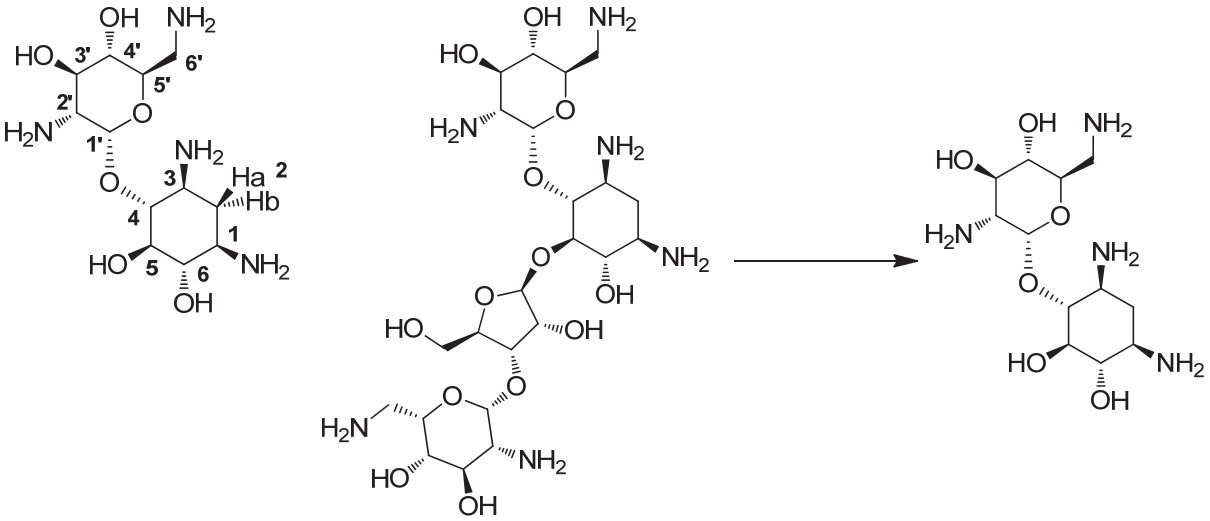
**Solide brun** (5,5 g, 52 %)

**RMN $^{1}$H** (300 MHz, $D_2O$) δ 3,15 (t, 1H, J = 8,7 Hz, H-5) ; 3,02 (bt, 2H, J = 9,0 Hz, H-4, H-6) ; 2,61 (td, 2H, J = 11 Hz, 3,9 Hz, H-1, H-3) ; 1,87 (dt, 1H, J = 12,4 Hz, 3,9 Hz, H-2a) ; 1,08 (q, J = 12,4 Hz, 1H, H-2b) ;

**RMN $^{13}$C** (75 MHz, $D_2O$) δ 77,3 (C-4, C-6) ; 75,6 (C-5) ; 50,4 (C-1, C-3) ; 35,3 (C-2) ;

**SM** (ES) : 163 ($[M+H]^+$).

**2) Néamine 3**



**Formule brute** : $C_{12}H_{26}N_4O_6$

**M** : 322 $g.mol^{-1}$ (468 $g.mol^{-1}$ sous forme de chlorhydrate)

**Mode opératoire :**

Le sulfate de néomycine (2g, 2.2 mmol) est dissous dans une solution d'acide chlorhydrique méthanolique (0.9 N, 80 mL). La solution est agitée à reflux pendant 15h pour donner une solution jaune clair. La réaction est alors refroidie à température ambiante puis concentrée de moitié avant d'être placée à 0 °C. Le solide blanc qui précipite est récupéré par centrifugation (3000 g, 4 min) pour donner 720 mg de chlorhydrate de néamine (720 mg, 1.6 mmol, 71 %).

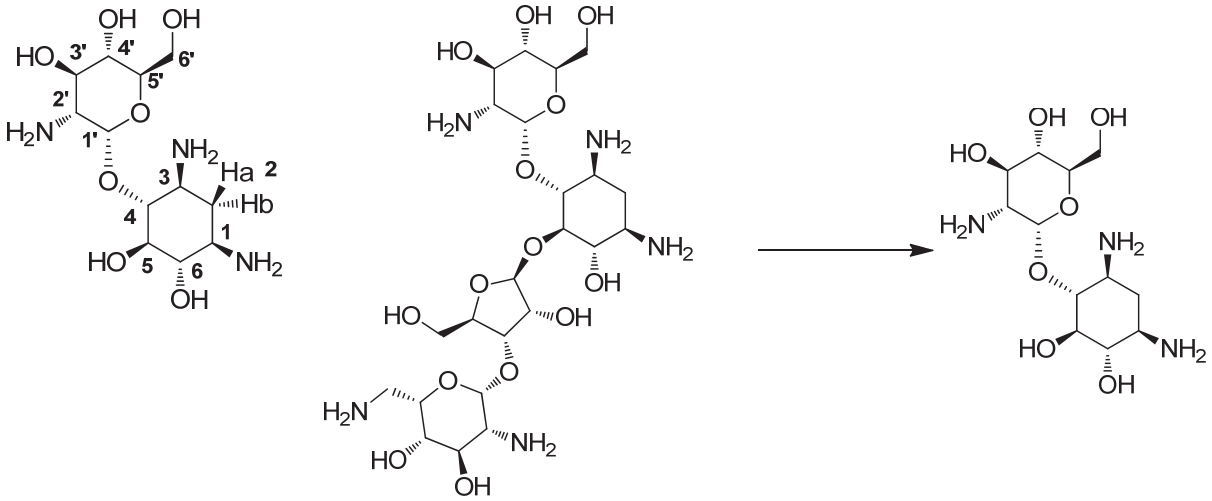
**Solide blanc** (720 mg, 71 %)

**RMN $^{1}$H** (300 MHz, $D_2O$) δ 5.79 (d, 4 Hz, 1H, H-1'), 3.94-3.81 (m, 3H, H-3', H-5', H-5), 3.62-3.52 (m, 1H, H-4), 3.51-3.28 (m, 5H, H-6, H-2', H-4', H-6'), 3.27-3.11 (m, 2H, H-1, H-3), 2.38 (dt, J = 12.5 Hz, 4.3 Hz, 1H, H-2a), 1.77 (q, J = 12.5 Hz, 1H, H-2b ) ;

**RMN $^{13}$C** (75 MHz, $D_2O$) δ. 96.0 (C-1'), 77.56 (C-4), 75.1 (C-5), 72.4 (C-6), 70.6 (C-5'), 69.1 (C-4'), 68.1 (C-3'), 53.4 (C-6'), 49.6 (C-2'), 48.3 (C-1), 40.1 (C-3), 28.2 (C-2) ;

**SM** (ESI+) : 323 ($[M+H]^+$).

### 3) Paromamine 2



**Formule brute** : $C_{12}H_{25}N_3O_7$

**M** : 322 g.mol$^{-1}$ (433 g.mol$^{-1}$ sous forme de chlorhydrate)

**Mode opératoire :**

Le sulfate de paromomycine (20g, 28.02 mmol) est dissous dans une solution d'acide chlorhydrique méthanolique (0.9 N, 330 mL) et agité à reflux pendant 15h jusqu'à apparition d'un précipité blanc. Après avoir laissé refroidir la solution à température ambiante, le solide blanc est récupéré par centrifugation (3000 g, 4min) pour donner 6.7 g de paromamine sous forme de chlorhydrate (6.7g, 15.5 mmol, 55%). Le surnageant est alors concentré sous pression réduite jusqu'à un volume de 50 mL, auquel on ajoute de nouveau 100 mL HCl méthanolique (1.25 N). La solution est portée au reflux pendant 12 h supplémentaires et donne, après centrifugation, encore 3.8 g de chlorhydrate de paromamine (rendement total : 10.5 g, 24.36 mmol, 87%).

**Solide blanc** (10.5 g, 87%)

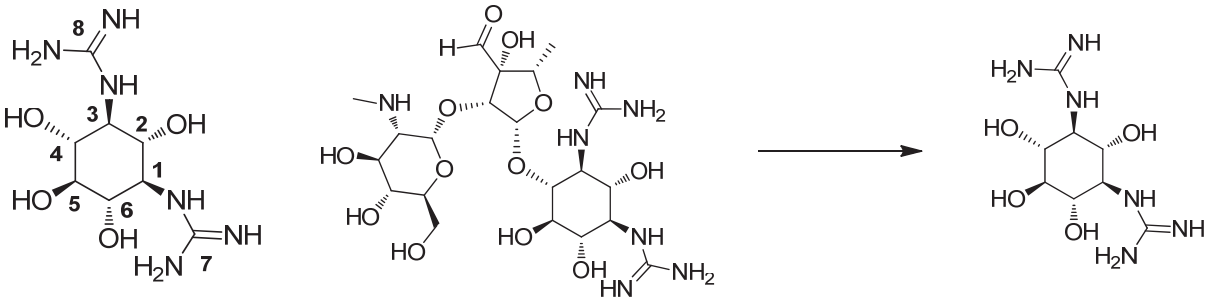
**RMN $^{1}$H** (300 MHz, $D_2O$) δ 5.53 (d, 1H, J = 4 Hz, H-1'), 3.86 – 3.66 (m, 4H, H-3', H-5', H-5, H-4), 3.66 – 3.50 (m, 2H, H-6'), 3.50 – 3.39 (m, 2H, H-2', H-4'), 3.38 – 3.22 (m, 3H, H-1, H-3, H-6), 2.37 (dt, J = 12.7, 4.3 Hz, 1H, H-2a), 1.72 (q, 1H, J = 12.6 Hz, H-2b);

**RMN $^{13}$C** (75 MHz, $D_2O$) δ 97.0 (C-1'), 80.3 (C-4), 74.6 (C-5), 73.5 (C-6), 72.3 (C-5'), 69.3 (C-4'), 68.9 (C-3'), 60.3 (C-6'), 53.9 (C-2'), 49.6 (C-1), 48.7 (C-3), 28.3 (C-2);

**MS** (ESI+): 324 ($[M+H]^+$);

**HRMS** for $C_{12}H_{25}N_3O_7$: ($[M+H]^+$): calculé: 324.17708, trouvé: 324.1770 (0 ppm); ($[M+Na]^+$): calculé: 346.15902, trouvé: 346.1592 (0 ppm).

**4) Streptidine 49**



**Formule brute** : $C_8H_{18}N_6O_4$

**M** : 262 $g.mol^{-1}$ (456 $g.mol^{-1}$ sous forme de sels de sulfate)

**Mode opératoire :**

La streptomycine (10,7 g ; 7.35 mmol; 1 eq.) est dissoute dans 100 mL d'un mélange MeOH/eau (8 : 2). On ajoute 3 mL d'acide sulfurique fumant et on agite le mélange obtenu à température ambiante pendant 5 jours. La streptomycine se solubilise en environ 2 h puis il y apparition lente d'un précipité blanc. Après 5 jours, on filtre, on lave le filtrat avec 50 mL de MeOH/eau (4 : 1) et on le sèche longuement à la pompe pour obtenir 3,2 g de solide blanc. Le produit peut être recristallisé dans l'acide sulfurique 2% dans l'eau pour donner des aiguilles blanches (2 g ; 4.4 mmol ; 60 %).

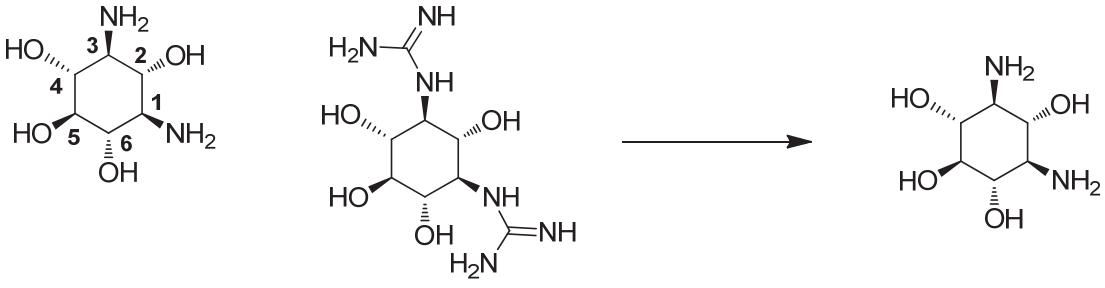
**Cristaux blancs (aiguilles)** (2 g, 60 %)

**RMN $^{1}$H** (400 MHz, $D_2O$) δ 7.45 (d, J = 6.5 Hz, 2H, NH-1, NH-3), 7.14 – 6.70 (bs, 6H, NH), 5.73 (bs, 1H), 5.30 (m, 3H, OH), 3.19 (m, 3H, H-1, H-3, H-6), 3.11 (m, 3H, H-2, H-4, H-5) ;

**RMN $^{13}$C** (100 MHz, DMSO) δ 158.4 (C-8, C-7), 76.4 (C-5), 72.1 (C-4, C-6), 71.3 (C-2), 59.4 (C-1, C-3) ;

**SM** (ES) : 263 ($[M+H]^+$).

**5) Streptamine 47**



**Formule brute** : $C_6H_{14}N_2O_4$

**M** : 178 $g.mol^{-1}$

**Mode opératoire :**

La streptidine recristalisée (2g ; 4,4 mmol ; 1eq.) est dissoute dans 80 mL d'une solution aqueuse saturée d'hydroxyde de baryum $Ba(OH)_2$. Le mélange est porté à reflux et agité pendant 3 jours. Après avoir refroidi le milieu réactionnel, on le filtre puis on ajoute doucement quelques gouttes d'acide sulfurique (≈ 500 µL) de façon à repasser en milieu acide. Il y a alors formation d'un précipité blanc de sulfate de baryum $BaSO_4$ que l'on filtre. La solution résultante est alors évaporée pour donner, après séchage à la pompe, 1,2 g de streptamine sous forme de sels de sulfate (3.2 mmol, 73 %). Si nécessaire, ses sels peuvent être repurifiés par échange d'ions à l'aide d'une phase Amberlyst A-15 selon le même protocole que celui employé pour la désoxystreptamine.

**Solide brun** (1.2 g, 73 %)

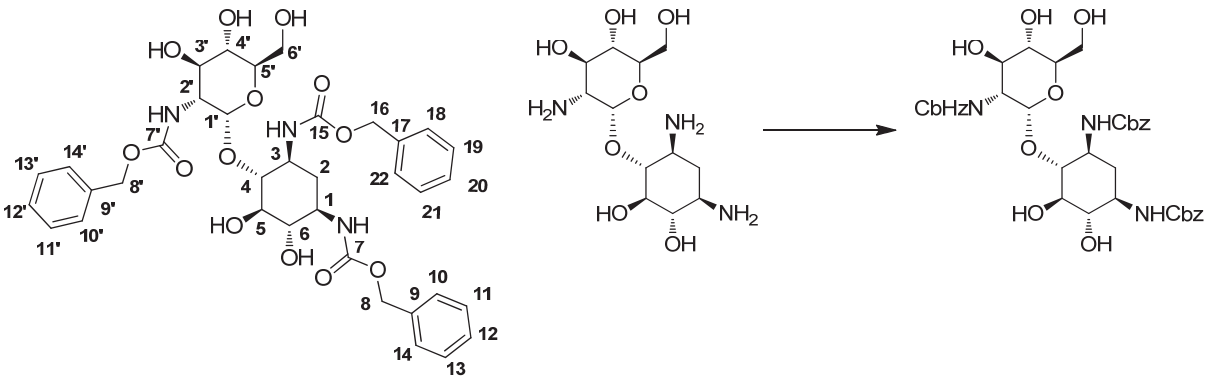
**RMN $^{1}$H** (400 MHz, $D_2O$) δ 3.70 (bt, J = 10.7 Hz, 1H, H-2), 3.40 (m, 2H, H-4, H-6), 3.32 (m, 1H, H-5), 3.06 (m, 2H, H-1, H-3) ;

**RMN $^{13}$C** (100 MHz, $D_2O$) δ 74.4 (C-5), 69.3, 66.1 (C-2), 56.3 (C-1, C-3) ;

**SM** (ES) : 179 ($[M+H]^+$).

## *C - Synthèse de la paromamine avec un fluor en 6' paroF*

### 1) Composé 110



**Formule brute** : $C_{36}H_{43}N_3O_{13}$

**M** : 726 g.mol$^{-1}$

**Mode opératoire :**

On dissout de la paromamine **2** (5 g, 11.57 mmol, 1 eq) et du carbonate de sodium (15g, 141.52 mmol, 12.2 eq) dans 50 mL d'eau. On agite vigoureusement jusqu'à dissolution totale des produits, puis on ajoute doucement de l'acétone (150 mL) à température ambiante. Le mélange est alors refroidi à -5°C à l'aide d'un bain eau/glace/NaCl et a tendance à devenir hétérogène (formation d'un solide blanc pâteux). A ce stade, il est important de s'assurer qu'il y a encore une bonne agitation. Le mélange réactionnel hétérogène est agité vigoureusement au cours de l'addition goute à goute sur une période de 20 min de CbzCl en solution dans le toluène (30 % en masse, 8.4 g CbzCl dans 19.6 g de toluène, 49.24 mmol, 4.25 eq.). Après une demi-heure, le mélange réactionnel est laissé remonter progressivement à température ambiante et agité pendant 4 heures supplémentaires. Le mélange réactionnel est alors gardé à 4 °C pendant la nuit sans agitation pour promouvoir la précipitation d'un solide blanc. Ce solide est récupéré par filtration, lavé avec une solution d'éther (20 mL) dans l'eau (400 mL), trituré en plusieurs fois dans 300 mL de HCl 1N jusqu'à disparition totale de la mousse résultant de la consommation du carbonate de sodium excédentaire, puis lavé à nouveau à l'eau (200 mL). Le solide résultant est séché longuement à la pompe puis trituré dans du dioxane chaud (120 °C) (7.5g du solide pour 150 mL de dioxane) pour donner le composé **110** sous forme de solide blanc (6.18 g, 8.52 mmol, 74 %). Le dioxane issu de la trituration est évaporé et le solide qui en résulte est de nouveau trituré dans le dioxane pour donner 900 mg de produit supplémentaire (rendement total : 7.08 g, 9.76 mmol, 84 %).

**Solide blanc** (7.08 g, 84 %)

**HPLC** : $t_R$ 13.9 min

**RMN $^{1}$H** (400 MHz, DMSO) δ 7.42-7.25 (m, 15H), 7.17 (d, 1H, J = 8 Hz), 7.11 (d, 1H, J = 8 Hz), 6.85 (d, 1H, J = 8 Hz), 5.20-4.93 (m, 8H), 4.93-4.85 (m, 2H), 4.81 (m, 1H), 4.25 (m, 1H), 3.57 (m, 4H), 3.44 (m, 2H), 3.30 (m, 4H), 3.11 (m, 1H), 1.78 (bd, 1H, J = 12.6 Hz), 1.31 (qd, 1H, J = 12.6 Hz);

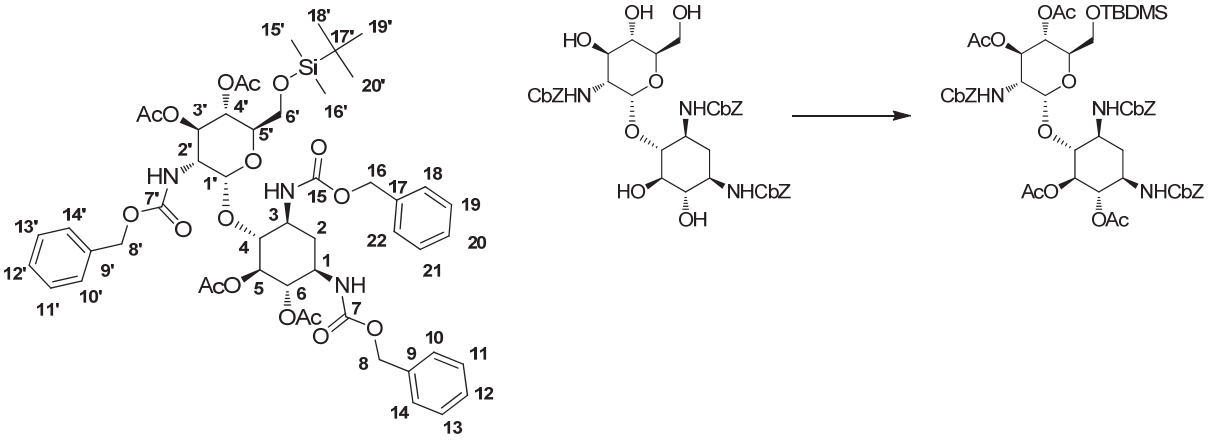
**RMN $^{13}$C** (75 MHz, DMSO) δ (300 MHz, DMSO) δ 156.7, 156.2, 156.1, 137.8, 137.7, 137.5, 129.2-127.9, 99.4, 82.5, 77.0, 74.5, 73.3, 71.6, 70.4, 65.8, 65.6, 65.6, 60.6, 56.7, 51.6, 50.7, 35.4;

**MS** (ESI+): 748 ($[M+Na]^+$)

**HRMS** for $C_{36}H_{43}N_3O_{13}$: ($[M+Na]^+$): calculé: 748.26881, trouvé : 748.2689 (0 ppm); ($[M+K]^+$): calculé: 764.24275, trouvé: 764.2431 (0 ppm).

| Composé 110-DMSO | C | δ ($^{13}$C) | δ ($^{1}$H) | m, nH, J |
|---|---|---|---|---|
| **1'** | CH | 99.4 | 5.06 | m, 1H |
| **2'** | CH | 56.7 | 3.42 | m, 1H |
| **3'** | CH | 71.6 | 3.45 | m, 1H |
| **4'** | CH | 70.4 | 3.29 | m, 1H |
| **5'** | CH | 73.3 | 3.55 | m, 1H |
| **6'** | $CH_2$ | 60.6 | 3.57 | m, 2H |
| **NH (2')** | - | - | 6.85 | d, 1H, J = 8 Hz |
| **1** | CH | 51.6 | 3.27 | m, 1H |
| **2** | $CH_2$ | 35.4 | 1.31<br>1.78 | qd, 1H, J = 12.6 Hz<br>bd, 1H, J = 12.6 Hz |
| **3** | CH | 50.7 | 3.41 | m, 1H |
| **4** | CH | 82.5 | 3.31 | m, 1H |
| **5** | CH | 77.1 | 3.29 | m, 1H |
| **6** | CH | 74.5 | 3.10 | m, 1H |
| **NH (1)** | - | - | 7.11 | d, 1H, J = 8 Hz |
| **NH (3)** | - | - | 7.17 | d, 1H, J = 8 Hz |
| **7, 7', 15** | Cq | 156.7, 156.2, 156.1 | - | |
| **8, 8', 16** | $CH_2$ | 65.8, 65.6, 65.6 | 4.85-5.12 | m, 6H |
| **9, 9', 17** | Cq | 137.8, 137.7, 137.5 | - | |
| **10, 11, 12, 13, 14<br>10', 11', 12', 13', 14'<br>18, 19, 20, 21, 22** | $CH_{aro}$ | 127.9-129.2 | 7.25-7.42 | m, 15H |

## 2) Composé 83



**Formule brute** : $C_{50}H_{65}N_3O_{17}Si$

**M** : 1008 $g.mol^{-1}$

**Mode opératoire :**

Sous atmosphere d'argon, le composé **110** (5.74 g, 7.91 mmol, 1 eq.) et la 4-(dimethylamino)pyridine (DMAP ; 967 mg, 7.91 mmol, 1eq.) sont dissous dans la DMF (80 mL). La solution est refroidie à -5°C dans un bain eau/glace/NaCl, puis la triéthylamine ($Et_3N$; 2.86 mL, 20.58 mmol, 2.6 eq.) et le TBDMSCl (2.38g, 15.83 mmol, 2 eq.) sont ajoutés. La réaction est agitée à -5°C et son avancement est suivi par CCM et par HPLC (voir méthode générale dans les généralités). Après une heure, la HPLC montre une disparition du produit de départ ($t_R$ 13.9 min) et l'apparition d'un nouveau produit ($t_R$ 21.7 min) correspondant au composé **110** dont l'alcool primaire en 6' est protégé par un groupement TBDMS. La réaction est alors stoppée par ajout de 50 mL de MeOH et les solvants sont évaporés sous pression réduite. Le brut réactionnel est séché à la pompe et directement engagé dans l'étape suivante sans purification supplémentaire. Le brut réactionnel est ainsi dissous dans la pyridine (128 mL, 1.58 mol, 200 eq.) avec de la DMAP (3.87 g, 31.66 mmol, 4 eq.) et de l'acide acétique (37.3 mL, 396 mmol, 50 eq.). La solution est agitée pendant 24 heures à température ambiante jusqu'à disparition complète du produit de départ ($t_R$ 21.7 min) et apparition d'un composé majeur ($t_R$ 26.7 min). La réaction est alors refroidie à 0°C et arrêtée par ajout de MeOH (80mL). La solution est évaporée puis reprise dans 100 mL de DCM et 100 mL d'eau. La phase aqueuse est extraite avec du DCM (3 x 100 mL) et les phases oragniques sont rassemblées, lavées avec de la saumure, séchées sur $MgSO_4$, puis évaporées pour donner 12.4 g d'une huile orangée. Ce produit brut est purifié par chromatographie éclair sur gel de silice (éluant : DCM / $Et_2O$) pour donner 5.7 g du composé **83** sous forme de poudre blanche (rendement total sur les deux étapes : 5.7 g, 9.76 mmol, 72 %).

**Solide blanc** (5.7 g, 72 %)

**CCM** R*f* =0.37 ($CH_2Cl_2$/$Et_2O$, 85:15)

**HPLC**: $t_R$ 26.7 min

**RMN $^1$H** (400 MHz, $CDCl_3$, 293 K) δ 7.45-7.30 (m, 15H), 5.67 (d, 0.9H, J = 7.8 Hz), 5.27 (m, 1H), 5.25-5.00 (m, 9H), 4.93 (m, 2H), 4.85 (bt, 1H, J = 10.0 Hz), 4.19 (bt, 1H, J = 7.9 Hz), 4.06 ( td, 1H, J = 10.1, 3.0 Hz), 3.91 (m, 1H), 3.84 (m, 1H), 3.76 (bt, 1H, J = 9.6 Hz), 3.60-3.50 (m, 2H), 2.49 (bd, 1H, J = 13.0 Hz), 1.93 (s, 3H), 1.90 (s, 3H), 1.87 (s, 3H), 1.83 (s, 3H), 1.47 (m, 1H), 0.90 (s, 9H), 0.04 (s, 3H), 0.02 (s, 3H);

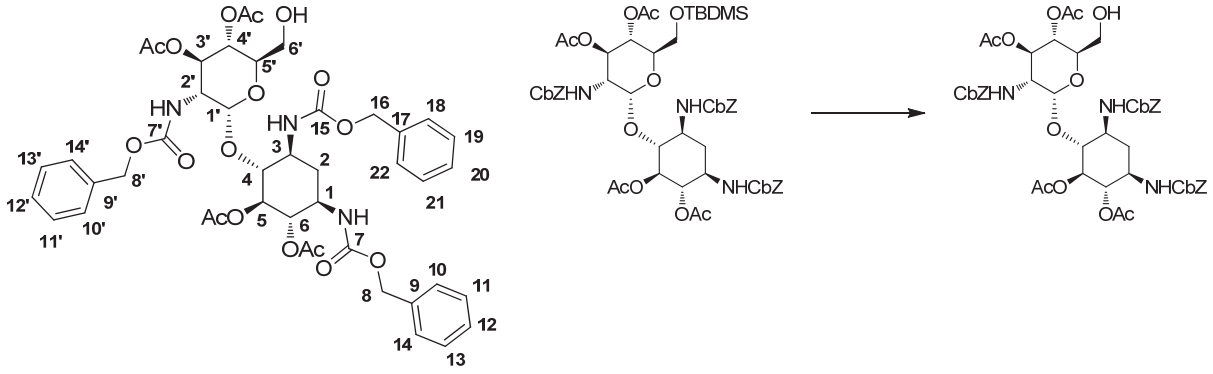
**RMN $^{13}$C** (75 MHz, $CDCl_3$, 293 K) δ 171.1, 170.6, 169.5, 169.4, 156.0, 155.6, 155.5, 136.3-136.2, 128.5-127.6, 98.3, 77.8, 74.2, 73.7, 71.3, 70.1, 68.7, 67.0-66.9, 62.3, 53.4, 50.2, 49.6, 35.0, 25.9, 20.5-20.4, 18.4, -5.4, -5.6;

**MS** (ESI+): 1030 ($[M+Na]^+$)

**HRMS** for $C_{50}H_{65}N_3O_{17}NaSi$: ($[M+Na]^+$): calculé: 1030.39755, trouvé: 1030.3977 (0 ppm); ($[M+K]^+$): calculé: 1030.37148, trouvé: 1030.3722 (1 ppm).

| Composé 83– $CDCl_3$ | C | δ ($^{13}C$) | δ ($^{1}H$) | m, nH, J |
|---|---|---|---|---|
| **1'** | CH | 98.3 | 5.07 | m, 1H |
| **2'** | CH | 53.4 | 4.06 | td, 1H, J = 10.1, 3.0 Hz |
| **3'** | CH | 71.3 | 5.10 | m, 1H |
| **4'** | CH | 68.7 | 4.93 | m, 1H |
| **5'** | CH | 70.1 | 4.19 | bt, 1H, J = 7.9 Hz |
| **6'** | $CH_2$ | 62.3 | 3.50 – 3.60 | Massif, 2H |
| **NH (2')** | - | - | ? | |
| **1** | CH | 49.6 | 3.91 | m, 1H |
| **2** | $CH_2$ | 35.0 | 1.47<br>2.49 | m, 1H<br>bd, 1H, J = 13.0 Hz |
| **3** | CH | 50.2 | 3.84 | m, 1H |
| **4** | CH | 77.8 | 3.76 | bt, 1H, J = 9.6 Hz |
| **5** | CH | 74.2 | 5.27 | bt, 1H, J = 9 Hz |
| **6** | CH | 73.7 | 4.85 | bt, 1H, J = 10 Hz |
| **NH (1)** | - | - | 5.67 | d, 0.9H, J = 7.8Hz |
| **NH (3)** | - | - | 4.92 | |
| **7, 7', 15** | C=O | 155.5, 155.6, 156.0 | - | |
| **8, 8', 16** | $CH_2$ | 66.9 - 67.0 | 5.00 - 5.25 | m, 6H |
| **9, 9', 17** | Cq | 136.2 - 136.3 | - | |
| **10, 11, 12, 13, 14<br>10', 11', 12', 13', 14'<br>18, 19, 20, 21, 22** | $CH_{aro}$ | 127.6 -128.5 | 7.30 - 7.45 | m, 15H |
| **Groupements acétyl** | $CH_3$ | 20.4 – 20.5 | 1.83, 1.87, 1.90, 1.93 | 4 s, 3H |
| **Groupements acétyl** | C=O | 169.4, 169.5, 170.6, 171.1 | - | |
| **15',16'** | CH3 | -5.42, -5.61 | 0.04 ; 0.02 | 2s, 2 x 3H |
| **17'** | C | 18.4 | - | |
| **18', 19', 20'** | CH3 | 25.9 | 0.9 | s, 9H |

### 3) Composé 84



**Formule brute** : $C_{44}H_{51}N_3O_{17}$

**M** : 894 g.mol$^{-1}$

**Mode opératoire :**

Le composé **83** (437 mg, 0.44 mmol, 1eq.) est dissous dans 8 mL de THF anhydre sous atmosphère d'argon. Une solution stock de HF·pyridine dans un mélange pyridine/THF anhydre (rapport volumique 1:2:2, HF·pyridine/pyridine/THF) est préparée et 2.45 mL de cette solution sont ajoutés au mélange réactionnel (ce qui correspond à environ 19.5 mmol de HF, soit 45 eq.). La réaction est suivie par CCM et HPLC. Au bout de 4 heures d'agitation à température ambiante, la réaction est diluée avec de l'EtOAc (50 mL) et stoppée par ajout de $NaHCO_3$(10 % dans eau, 40 mL). La phase aqueuse est extraite à l'EtOAc (3 x 50 mL), les phases organiques sont rassemblées, lavées au $NaHCO_3$(10 % dans eau, 40 mL) et à la saumure (40 mL), séchées au $MgSO_4$, puis évaporées pour donner 430 mg de produit brut. Ce produit peut directement être utilisé dans l'étape suivante sans nécessiter de purification supplémentaire. Alternativement, il peut être purifié par chromatographie éclair sur gel de silice (éluant : EtOAc/cyclohexane, 75/25) pour donner le composé **84** sous forme de poudre blanche (303 mg, 0.34 mmol, 78 %).

**Solide blanc** (303 mg, 78 %)

**CCM** R*f* =0.36 (EtOAc/Cyclohexane, 75:25)

**HPLC** characterization: $t_R$ 19.9 min

**RMN $^1$H** (400 MHz, $CDCl_3$, 323 K) δ 7.43-7.32 (m, 15H), 5.52 (bs, 1H), 5.24-4.99 (m, 6H), 5.23 (m, 1H), 5.15 (m, 1H), 5.10 (m, 1H), 4.94 (m, 2H), 4,09 (m, 1H), 4,01 (bdd, 1H, J = 6.0, 10.0 Hz), 3.89 (m, 1H), 3.86 (m, 1H), 3.79 (m, 1H), 3.62 (bs, 1H), 3.45 (bs, 1H), 2.42 (bd, 1H, J = 13.0 Hz), 2.00 (s, 3H), 1.93 (s, 3H), 1.89 (s, 3H), 1.85 (s, 3H), 1.68 (m, 1H);

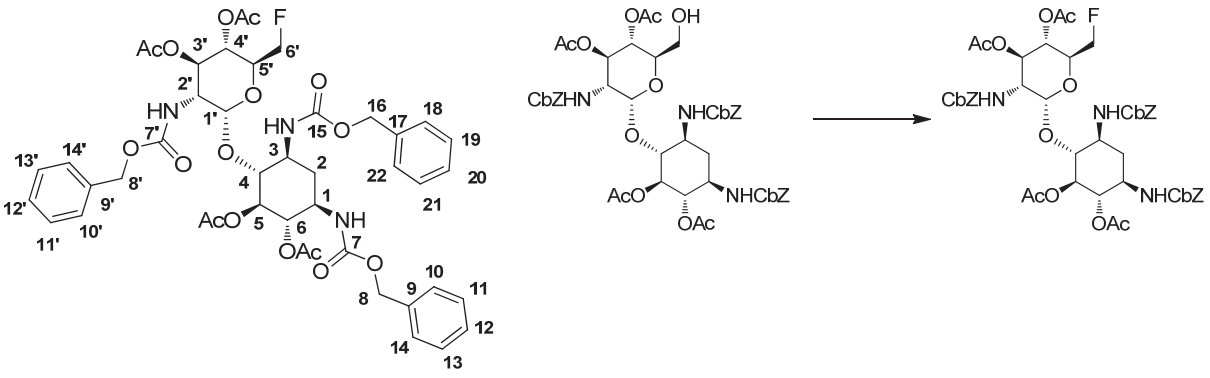
**RMN $^{13}$C** (75 MHz, $CDCl_3$, 323 K) δ 170.8, 170.4, 169.9, 169.6, 156.0-155.6, 136.5-136.0, 128.0-128.6, 99.2, 79.4, 74.5, 73.7, 71.0, 71.0, 69.1, 67.3-66.9, 61.5, 53.9, 50.8, 49.7, 34.7, 20.4-20.2;

**MS** (ESI+): 916 ($[M+Na]^+$)

**HRMS** for $C_{44}H_{51}N_3O_{17}$ : ($[M+Na]^+$): calculé: 916.31107, trouvé: 916.3112 (0 ppm); ($[M+K]^+$): calculé: 932.28501, trouvé: 932.2868 (2 ppm).

| Composé 84 – $CDCl_3$ | C | δ ($^{13}C$) | δ ($^{1}H$) | m, nH, J |
|---|---|---|---|---|
| **1'** | CH | 99.2 | 5.10 | m, 1H |
| **2'** | CH | 53.9 | 4.08 | m, 1H |
| **3'** | CH | 71.0 | 5.15 | m, 1H |
| **4'** | CH | 69.1 | 4.93 | m, 1H |
| **5'** | CH | 71.0 | 4.01 | bdd, 1H, J = 6.0, 10.0 Hz |
| **6'** | $CH_2$ | 61.5 | 3.45, 3.62 | 2 x bs, 1H et 1H |
| **NH (2')** | - | - | ? | |
| **1** | CH | 49.7 | 3.88 | m, 1H |
| **2** | $CH_2$ | 34.7 | 1.59<br>2.43 | m, 1H<br>bd, 1H, J = 13.0 Hz |
| **3** | CH | 50.8 | 3.84 | m, 1H |
| **4** | CH | 79.4 | 3.83 | m, 1H |
| **5** | CH | 74.5 | 5.23 | m, 1H |
| **6** | CH | 73.7 | 4.92 | m, 1H |
| **NH (1)** | - | - | 5.49 | bs, 1H |
| **NH (3)** | - | - | - | |
| **7, 7', 15** | C=O | 155.6 - 156.0 | - | |
| **8, 8', 16** | $CH_2$ | 66.9 - 67.3 | 4.99 - 5.24 | m, 6H |
| **9, 9', 17** | Cq | 136.0 - 136.5 | - | |
| **10, 11, 12, 13, 14<br>10', 11', 12', 13', 14'<br>18, 19, 20, 21, 22** | $CH_{aro}$ | 128.0 -128.6 | 7.32 - 7.43 | m, 15H |
| **Groupements acétyl** | $CH_3$ | 20.2 - 20.4 | 1.85, 1.89,<br>1.94, 2.00 | 4 s, 3H |
| **Groupements acétyl** | C=O | 169.6, 169.9, 170.4, 170.8 | - | |

**4) Composé 85**



**Formule brute** : $C_{44}H_{50}FN_3O_{16}$

**M** : 896 g.mol$^{-1}$

**Mode opératoire :**

Dans de la verrerie sèche sous atmosphère d'argon, on dissout le composé **84** (295 mg, 0.33 mmol, 1 eq.) dans 3 mL de DCE anhydre. La solution est refroidie à -5 °C puis le DAST est lentement ajouté (87 µL, 0.66 mmol, 2 eq.). Après une heure d'agitation à -5 °C, on laisse la solution progressivement remonter à température ambiante et on agite pour 4 heures supplémentaires. A ce stade, la HPLC indique la consommation totale du produit de départ ($t_R$ 19.9 min) et l'apparition d'un nouveau composé majoritaire ($t_R$ 21.7 min). La réaction est alors stoppée par ajout de 4 mL de MeOH et le solvant est évaporé sous pression réduite. Le brut réactionnel est purifié par chromatographie éclair sur gel de silice (éluant : DCM/EtOAc) pour donner le composé **85** sous forme de poudre blanche (188 mg, 0.22 mmol, 64 %).

**Solide blanc** (188 mg, 64 %)

**CCM** R*f* =0.22 ($CH_2Cl_2$/ $Et_2O$, 75:25), 0.44 ($CH_2Cl_2$/EtOAc, 75:25)

**HPLC** : $t_R$ 21.7 min

**RMN $^1$H** (300 MHz, DMSO) δ 7.51 (d, 1H, J = 8.6 Hz), 7.44-7.27 (m, 15H), 6.63 (d, 1H, J = 8.9 Hz), 5.19 (t, 1H, J = 9.4 Hz), 5.15-5.02 (m, 4H), 5.02-4.89 (m, 5H), 4.72 (t, 1H, J = 10.2 Hz), 4.55 (dd, 1H, J = 10.6, 49.4 Hz), 4.27 (dd, 1H, J = 10.9, 47.2 Hz), 3.97 (dd, 1H, J = 9.0, 30.7 Hz), 3.86-3.66 (m, 4H), 1.97-1.76 (m, 13H), 1.70 (q, 1H, J = 12.2 Hz);

**RMN $^{13}$C** (375 MHz, DMSO) δ 170.0, 169.9, 169.5, 169.4, 156.3, 155.9, 155.7, 137.6, 137.2, 137.1, 128.8-127.9, 97.6, 80.9 (d, J = 171 Hz), 77.8, 75.0, 74.5, 70.9, 68.2 (d, J = 18 Hz), 67.5 (d, J = 6 Hz), 66.1, 65.9, 65.5, 53.5, 49.8, 48.6, 34.1, 20.7-20.6;

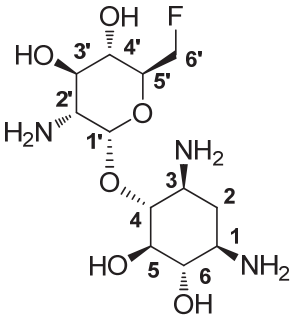
**RMN $^{19}$F** (282 MHz, DMSO) δ -236.03 (td, J = 47.2 Hz, 30.1 Hz);

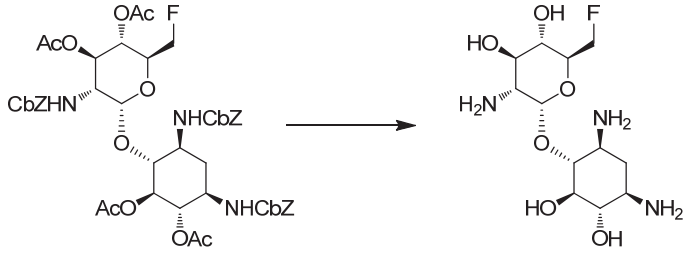
**MS** (ESI+): 918 ([M+Na]$^+$)

**HRMS** for $C_{44}H_{50}N_3O_{13}F$: ([M+Na]$^+$): calculé: 918.30673, trouvé: 918.3066 (0 ppm); ([M+K]$^+$): calculé: 934.28067, trouvé: 934.2792 (2 ppm);

| Composé 85 – DMSO | C | δ ($^{13}$C) | δ ($^{1}$H) | m, nH, J |
|---|---|---|---|---|
| **1'** | CH | 97.9 | 5.12 | m, 1H |
| **2'** | CH | 53.9 | 3.80 | m, 1H |
| **3'** | CH | 71.4 | 4.97 | m, 1H |
| **4'** | CH | 67.8 (d, 6 Hz) | 4.94 | m, 1H |
| **5'** | CH | 68.5 (d, 18 Hz) | 3.97 | dd, 1H, J = 9.0, 30.7 Hz |
| **6'** | $CH_2$ | 80.9 (d, 171 Hz) | 4.55, 4.27 | dd, 1H, 10.6,49.4<br>dd, 1H, 10.9,47.2 |
| **NH (2')** | - | - | 6.56 | d, 1H, J = 8.9 Hz |
| **1** | CH | 49.0 | 3.76 | m, 1H |
| **2** | $CH_2$ | 34.2 | 1.70<br>1.82 | qd, 1H, J = 12.2 Hz<br>m, 1H |
| **3** | CH | 50.2 | 3.75 | m, 1H |
| **4** | CH | 78.3 | 3.71 | m, 1H |
| **5** | CH | 75.3 | 5.19 | t, 1H, J = 9.4 Hz |
| **6** | CH | 74.6 | 4.72 | t, 1H, J = 10.2 Hz |
| **NH (1)** | - | - | 7.49 | d, 1H, J = 8.6 Hz |
| **NH (3)** | - | - | 7.38 | Cosy |
| **7, 7', 15** | C=O | 155.9, 156.0, 156.3 | - | |
| **8, 8', 16** | $CH_2$ | 65.7, 66.0, 66.3 | 4.89 - 5.14 | m, 6H |
| **9, 9', 17** | Cq | 137.2, 137.4, 137.7 | - | |
| **10, 11, 12, 13, 14<br>10', 11', 12', 13', 14'<br>18, 19, 20, 21, 22** | $CH_{aro}$ | 127.9 -128.9 | 7.28 - 7.41 | m, 15H |
| **Groupements acétate** | $CH_3$ | 20.6 - 20.8 | 1.80, 1.87,<br>1.90 | 1 s 6H, 2 s 3H |
| **Groupements acétate** | C=O | 169.4, 169.5, 169.9, 170.1 | - | |

### 5) 6'-désoxy-6'-fluoro-paromamine (paroF)





**Formule brute** : $C_{12}H_{24}FN_3O_6$

**M** : 325 g.mol$^{-1}$

**Mode opératoire :**

Le composé **85** (176 mg, 197 µmol, 1eq.) est solubilisé dans 8 mL d'un mélange MeOH/THF/dioxane (2:4:2) et traité avec une quantité catalytique de NaOMe dans le MeOH (environ 100 µmol, 0.5 eq. de MeONa). Au bout de 3 heures d'agitation à température ambiante, la CCM et la HPLC indiquent une complète consommation du produit de départ ($t_R$ 21.7 min) et l'apparition d'un produit unique ($t_R$ 15.8 min), correspondant au produit déacétylé. Le mélange réactionnel est alors évaporé sous pression réduite pour donner une huile jaune. Le résidu est repris dans un mélange dioxane/MeOH/$H_2O$ (6:3:3, 12 mL) auquel on ajoute une solution aqueuse d'HCl (1.2N, 1 mL, environ 6 eq.) et 70 mg de Pd/C 10%. Le mélange est agité sous atmosphère d'hydrogène pendant 2 jours à température ambiante puis filtré sur célite, évaporé, repris dans l'eau et enfin lyophilisé pour donner la paromamine fluorée en position 6' **paroF** sous forme de sels de chlorhydrate (solide beige floconneux, 85 mg, 197 µmol, rendement quantitatif).

**Solide beige** (85 mg, quant.)

**$^1$H NMR** (400 MHz, $D_2O$) δ 5.45 (d, 1H, J = 3.8 Hz), 4.56 (dd, 2H, J = 47.1, 3.2 Hz), 3.93 (ddt, 1H, J = 25.1, 10.1, 3.2 Hz), 3.72 (dd, 1H, J = 9.3, 10.5 Hz), 3.55-3.42 (m, 3H), 3.39 (m, 1H), 3.18 (dd, 1H, J = 10.7, 3.9 Hz), 3.15 – 3.03 (m, 2H), 2.19 (dt, 1H, J = 12.6, 4.3 Hz), 1.50 (q, 1H, J = 12.6 Hz);

**$^{13}$C NMR** (300 MHz, $D_2O$) δ 96.9, 82.3 (d, J = 168 Hz), 79.7, 74.7, 72.3, 72.1 (d, J = 17.4 Hz), 68.8, 68.2 (d, J = 7 Hz), 53.7, 49.5, 48.5, 28.2;

**$^{19}$F NMR** (282 MHz, $D_2O$) δ -233.7 (td, J = 47.1 Hz, 25.1 Hz);

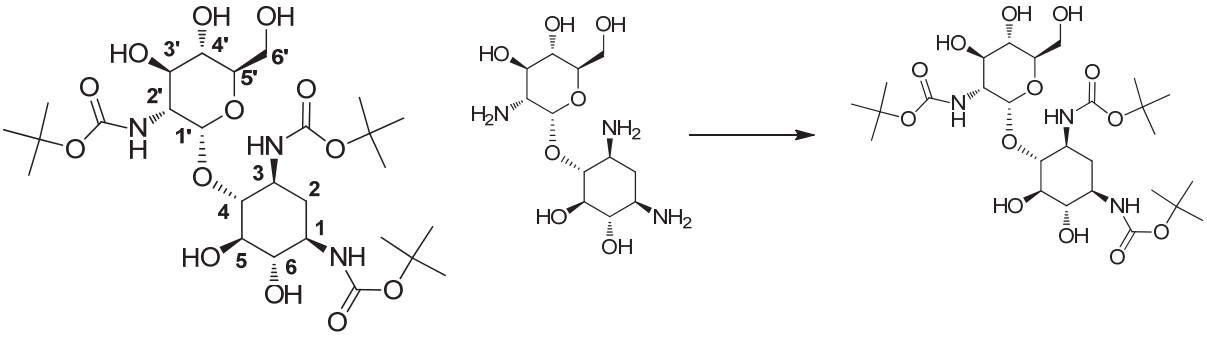
**MS** (ESI+): 326 ([M+Na]$^+$)

**HRMS** for $C_{12}H_{25}N_3O_6F$ : ([M+Na]$^+$): calculé: 326.17274, trouvé: 326.1727 (0 ppm); ([M+K]$^+$): calculé: 348.15468, trouvé: 348.1566 (0 ppm);.

| **paroF – $D_2O$** | **C** | **δ ($^{13}C$)** | **δ ($^{1}H$)** | **m, nH, J** |
|---|---|---|---|---|
| **1'** | CH | 97.7 | 5.42 | d, 1H, 3.8 Hz |
| **2'** | CH | 54.2 | 3.15 | dd, 1H, J = 10.7, 3.9 Hz |
| **3'** | CH | 70.1 | 3.72 | dd, 1H, J = 9.3, 10.5 Hz |
| **4'** | CH | 68.3 (d, 7 Hz) | 3.43 | m, 1H |
| **5'** | CH | 71.8 (d, 17.4 Hz) | 3.90 | ddt, 1H, J = 25.1, 10.1, 3.2 Hz |
| **6'** | $CH_2$ | 82.1 (d, 168 Hz) | 4.59 | dd, 2H, J = 47.1, 3.2 Hz |
| **NH (2')** | - | - | - | |
| **1** | CH | 50.2 | 3.09 | m, 1H |
| **2** | $CH_2$ | 30.8 | 1.47<br>2.16 | qd, 1H, J = 12.6 Hz<br>dt, 1H, J = 12.6, 4.3 |
| **3** | CH | 48.6 | 3.04 | m, 1H |
| **4** | CH | 83.3 | 3.44 | m, 1H |
| **5** | CH | 75.1 | 3.47 | m, 1H |
| **6** | CH | 73.1 | 3.37 | m, 1H |
| **NH (1)** | - | - | - | |
| **NH (3)** | - | - | - | |

## *D - Synthèse de la paromamine avec un azoture en 6' paroN3*

### 1) Composé 79



**Formule brute** : $C_{27}H_{49}N_3O_{13}$

**M** : 624 g.mol$^{-1}$

**Mode opératoire :**

On dissout la paromamine **2** dans 18 mL d'eau puis on ajoute de la DMF (50 mL), de la triethylamine (1.9 mL, 13.9 mmol, 6 eq.). Le milieu réactionnel est agité à température ambiante et le di-tert-butyl dicarbonate $(Boc)_2O$ est ajouté par petites portions successives (2.53 g, 11.6 mmol, 5 eq.). L'ensemble est ensuite agité pendant 8 heures, jusqu'à l'apparition en HPLC d'un composé majoritaire ($t_R$ = 11.1 min). Le mélange réactionnel est alors évaporé sous pression réduite pour donner un solide blanc. Ce solide est trituré dans le cyclohexane (2 x 20 mL) puis dans l'eau (2 x 10 mL) pour donner 1.18 g de paromamine protégée **79** sous forme de solide blanc (1.18 g, 1.9 mmol, 82 %).

**Solide brun** (1.2 g, 73 %)

**HPLC**: $t_R$ 11.2 min

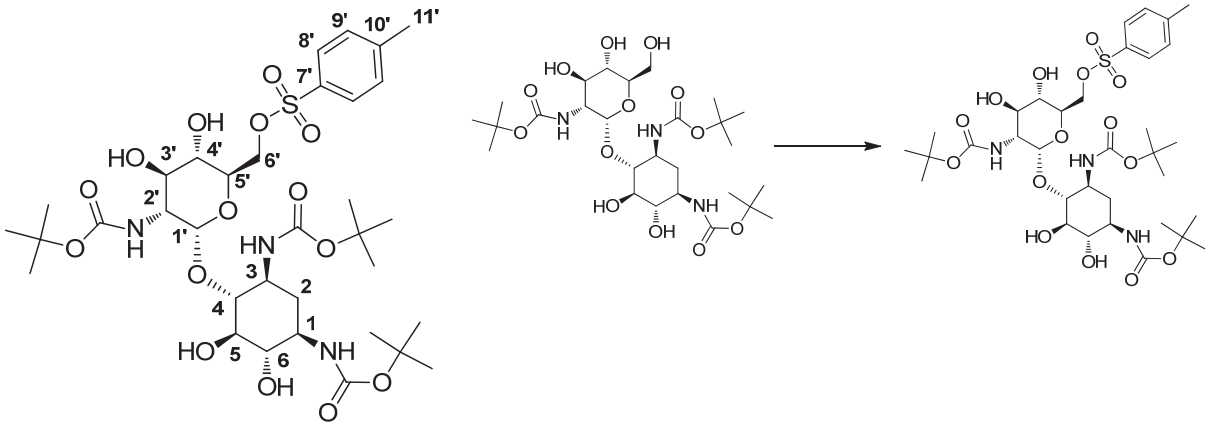
**$^1$H NMR** (400 MHz, DMSO) δ 6.66 (d, J = 6.8 Hz, 1H, NH), 6.49 (m, 1H, NH), 6.33 (d, J = 6.8 Hz, 1H, NH), 4.96 (m, 1H, H-1'), 4.85 (d, J = 4.8 Hz, OH-4'), 4.76 (d, J = 4.5 Hz, 1H, OH-6), 4.69 (m, J = 4.3 Hz, 1H, OH-3'), 4.62 (m, 1H, OH-5), 4.22 (m, 1H, OH-6'), 3.70 – 3.47 (m, 3H, H-5', H-6'), 3.42 – 3.16 (m, 6H, H-3, H 4, H-5, H-2', H-3', H-4'), 3.13 – 2.98 (m, 2H, H-1, H-6), 1.75 (bd, J = 10.4 Hz, 1H, H-2a), 1.38 (2 x s, 18H, Me Boc), 1.36 (s, 9H, Me Boc), 1.17 (m, 1H, H-2b);

**$^{13}$C NMR** (100 MHz, DMSO) δ 156.2 (CO Boc), 155.6 (CO Boc), 155.3 (CO Boc), 99.5 (C-1'), 82.3(C-4), 78.0 (2 x C Boc), 77.8 (C Boc), 77.4 (C-5), 74.1 (C-6), 72.9 (C-5'), 71.4 (C-3'), 70.3 (C-4'), 60.6 (C-6'), 56.0 (C-2'), 51.1 (C-1), 50.0 (C-3), 35.7 (C-2),28.6 (Me Boc), 28.5 (Me Boc), 27.2 (Me Boc);

**MS** (ESI+): 647 ($[M+Na]^+$)

**HRMS** for $C_{27}H_{49}N_3O_{13}$ : ($[M+Na]^+$): calculé: 646.31631, trouvé: 646.3159 (0 ppm); ($[M+K]^+$): calculé: 662.29025, trouvé: 662.2898 (0 ppm);

**2) Composé 80**



**Formule brute** : $C_{34}H_{55}N_3O_{15}S$

**M** : 778 g.mol$^{-1}$

**Mode opératoire :**

Le composé **79** (1.17 g, 1.87 mmol, 1 eq.) est dissous dans la pyridine anhydre (20 mL) sous atmosphère d'argon dans un ballon sec. La solution est refroidie à 0°C et on y ajoute goutte à goutte une solution de TsCl (390 mg, 2.06 mmol, 1.1 eq.) dans la pyridine (10 mL). Au bout d'une heure d'agitation à 0°C, on laisse la solution remonter progressivement à température ambiante et on agite pendant 3 heures supplémentaires. A ce stade, la HPLC indique une conversion totale du produit de départ ($t_R$ 11.2 min) et l'apparition d'un composé majoritaire ($t_R$ 17.1 min). La réaction est alors stoppée par ajout de 20 mL de MeOH et le solvant est évaporé sous pression réduite. Le brut réactionnel est purifié par chromatographie éclair sur gel de silice (éluant : DCM/MeOH, 96/4) pour donner le composé **80** sous forme de poudre brune (853 mg, 1.10 mmol, 59 % ).

**Solide brun** (853 mg, 59 %)

**CCM** R*f* =0.37 ($CH_2Cl_2$/ MeOH, 96:4)

**HPLC**: $t_R$ 17.2 min

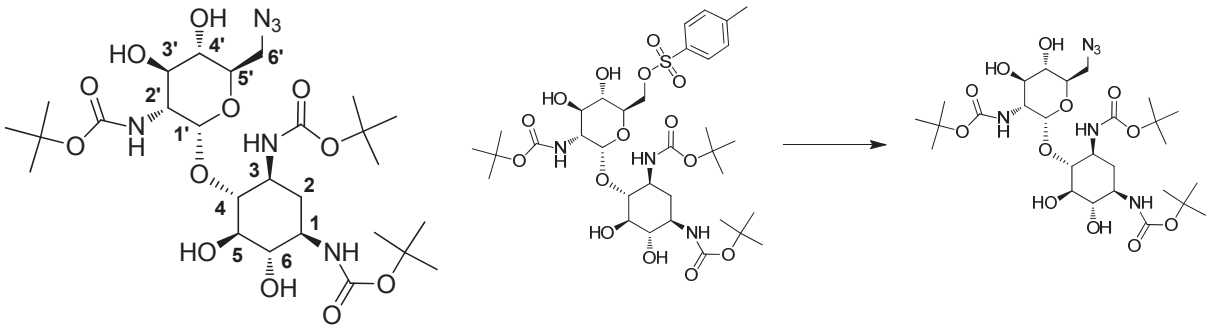
**$^1$H NMR** (400 MHz, DMSO) δ 7.77 (d, 2H, H-8'), 7.47 (d, 2H, H-9'), 6.66 (d, J = 7.7 Hz, 1H, NH), 6.60 (d, J = 7.7 Hz, 1H, NH), 6.34 (d, J = 7.1 Hz, 1H, NH), 5.17 (d, J = 4.4 Hz, OH-4'), 4.91 (m, 1H, H-1'), 4.85 – 4.69 (m, 2H, OH-6, OH-3'), 4.50 (m, 1H, OH-5), 4.23 (m, 2H, H-6'), 3.70 (bd, J = 9.3 Hz, 1H, H-5'), 3.35 – 3.13 (m, 7H, H-2', H-3', H-4', H-1, H-3, H-4, H-5), 3.12-2.97 (m, 1H, H-6), 2.42 (m, 3H, H-11'), 1.70 (bd, J = 12.9 Hz, 1H, H-2a), 1.38 (s, 18H, 2 x Me Boc), 1.28 (s, 9H, Me Boc), 1.16 (m, 1H, H-2b);

**$^{13}$C NMR** (100 MHz, DMSO) δ 156.2 (CO Boc), 155.7 (CO Boc), 155.2 (CO Boc), 145.3 (C-7'), 132.9 (C-10'), 130.6 (C-9'), 128.0 (C-8'), 99.8 (C-1'), 82.6(C-4), 78.2(2 x C Boc), 77.9(C Boc), 77.6(C-5), 74.2(C-6), 71.4 (C-3'), 69.8 (C-5'), 69.7 (C-4'), 69.6 C-6'), 55.9 (C-2'), 51.2 (C-1), 49.8 (C-3), 35.7 (C-2), 28.7 (2 x Me Boc), 28.5 (Me Boc), 21.6 (C-11');

**MS** (ESI+): 801 ([M+Na]$^+$)

**HRMS** for $C_{34}H_{55}N_3O_{15}S$ : ($[M+H]^+$): calculé: 778.34267, trouvé: 778.3413 (0 ppm); ($[M+Na]^+$): calculé: 800.32516, trouvé: 800.3249 (0 ppm); ($[M+K]^+$): calculé: 816.29855, trouvé: 816.2973 (2 ppm);

### 3) Composé 81



**Formule brute** : $C_{27}H_{48}N_6O_{12}$

**M** : 649 g.mol$^{-1}$

**Mode opératoire :**

Le composé **80** (229 mg, 0.295 mmol, 1 eq.) et l'azoture de sodium $NaN_3$ (192 mg, 2.95 mmol, 10 eq.) sont dissous dans la DMF (6 mL). Le mélange réactionnel est agité à 100 °C pendant 5 heures puis à température ambiante pendant la nuit. A ce stade, la HPLC indique une consommation totale du produit de départ ($t_R$ 17.1 min) et l'apparition d'un nouveau produit majoritaire ($t_R$ 14.6 min). Le mélange réactionnel est alors centrifugé (2 min, 3000 g) de façon à se débarrasser de l'azoture de sodium excédentaire non soluble restant, puis évaporé pour donner un solide blanc. (Pour éviter le risque d'explosion lié à la présence d'azoture de sodium, il vaut mieux filtrer la solution quand on travaille sur grosses quantités.) Ce solide est trituré dans l'eau (5 mL) puis longuement séché à la pompe pour donner 158 mg de composé **81** sous forme de solide blanc (128 mg, 0.243 mmol, 83 %).

**Solide blanc** (158 mg, 83 %)

**HPLC**: $t_R$ 14.6 min

**IR**: bande caractéristique de la fonction azoture à 2070 cm$^{-1}$

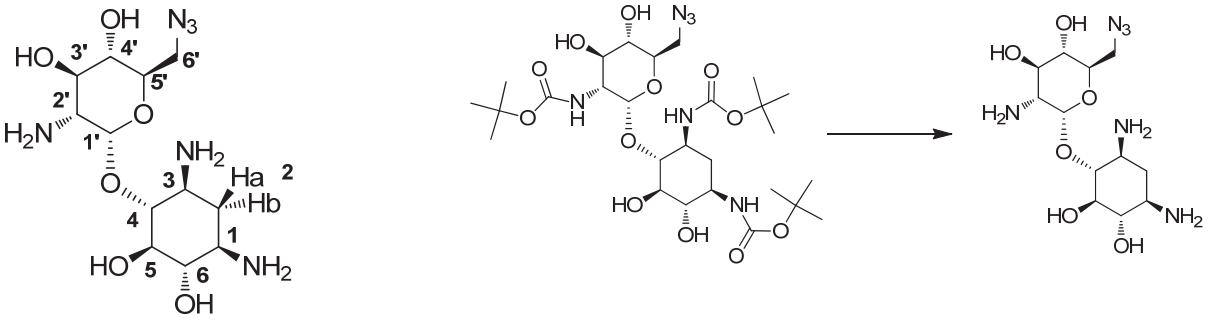
**$^1$H NMR** (400 MHz, DMSO) δ 6.76 (d, 1H, J = 7.7 Hz, NH-1), 6.72 (d, 1H, J = 7.2 Hz, NH-3), 6.45 (d, 1H, J = 6.2 Hz, NH-2), 5.29 (m, 1H, OH-4'), 5.12 (m, 1H, H-1'), 4.89 (m, 2H, OH-3', OH-6), 4.68 (m, 1H, OH-5), 3.80 (bd, J = 9.95 Hz, 1H, H-5'), 3.71 (d, J = 13 Hz, 1H, H-6'), 3.5 – 3.37 (m, 4H, H-2', H-3', H-4', H-6'), 3.37 – 3.17 (m, 4H, H-1, H-3, H-4, H-5), 3.12 (m, 1H, H-6), 1.8 (bd, J = 12.2 Hz, 1H, H-2a), 1.44 (m, 27 H, Me Boc), 1.30 (m, 1H, H-2b) ;

**$^{13}$C NMR** (100 MHz, DMSO) δ 156.2 (CO Boc), 155.6 (CO Boc), 155.3 (CO Boc), 99.5 (C-1'), 82.4(C-4), 78.0 (2 x C Boc), 77.8 (C Boc), 77.4 (C-5), 74.1 (C-6), 71.3 (C-5'), 71.2 (C-3'), 70.7 (C-4'), 55.9 (C-2'), 51.1 (C-1), 50.9 (C-6'), 49.9 (C-3), 35.7 (C-2),28.6 (Me Boc), 28.5 (Me Boc), 27.2 (Me Boc);

**MS** (ESI+): 671 ([M+Na]$^+$)

**HRMS** for $C_{27}H_{48}N_6O_{12}$ : ([M+Na]$^+$): calculé: 671.32224, trouvé: 671.3222 (0 ppm); ([M+K]$^+$): calculé: 687.29618, trouvé: 687.2955 (1 ppm);.

**4) 6'-désoxy-6'-azido-paromamine (paro$N_3$).**



**Formule brute** : $C_{12}H_{24}N_6O_6$

**M** : 348 g.mol$^{-1}$ (458 g.mol$^{-1}$ pour le chlorhydrate)

**Mode opératoire :**

Le composé **80** (80 mg, 0.123 mmol, 1 eq.) est dissous/suspendu dans de l'EtOAc anhydre (10 mL). On fait buller du HCl (gaz) dans le mélange réactionnel pendant environ une minute. Le composé **paro$N_3$** précipite sous forme de chlorhydrate et est isolé par filtration. Le composé est repris dans l'eau (2 mL) puis lyophilisé pour donner un solide blanc floconneux (56 mg, 0.123 mmol, rendement quantitatif).

**Solide blanc** (158 mg, quant.)

**IR**: bande caractéristique de la fonction azoture à 2070 cm$^{-1}$

**$^1$H NMR** (300 MHz, $D_2O$) δ 5.54 (d, 1H, J = 4.1 Hz, H-1'), 3.82 – 3.68 (m, 4H, H-6, H-3', H-5', H-6'), 3.58 – 3.39 (m, 5H, H-1, H-4, H-5, H-4', H-6'), 3.39 – 3.30 (m, 1H, H-2'), 3.22 (m, 1H, H-3), 2.38 (dt, 1H, J = 12.7, 4.8 Hz, H-2a), 1.75 (q, 1H, J = 12.8 Hz, H-2b);

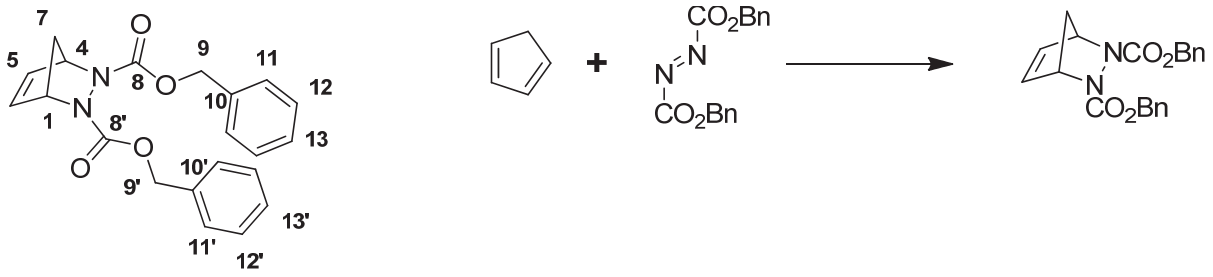
**$^{13}$C NMR** (75 MHz, $D_2O$) δ 97.1 (C-1'), 80.1 (C-4), 74.7 (C-5), 72.3 (C-6), 72.0 (C-5'), 69.9 (C-4'), 68.7 (C-3'), 53.8 (C-2'), 50.8 (C-6'), 49.5 (C-1), 48.5 (C-3), 28.2 (C-2);

**MS** (ESI+): 371 ([M+Na]$^+$)

**HRMS** for $C_{12}H_{24}N_6O_6$ : ([M+H]$^+$): calculé: 349.18356, trouvé: 349.1834 (0 ppm).

## *E - Synthèse du 1,3-diamino-cyclopentan-4,5-diol C5*

### 1) 2,3-diazabicyclo[2.2.1]hept-5-ène-2,3-dicarboxylate de dibenzyle (75)




**Formule brute** : $C_{21}H_{20}N_2O_4$

**M** : 364 g.mol$^{-1}$

**Mode opératoire :**

25g de diazocarboxylate de dibenzyle (84 mmol) et 50 mL de dichlorométhane sont introduits dans un ballon de 250 mL. Le milieu est refroidi à 0 °C, puis 16 mL de cyclopentadiène (193 mmol) fraîchement distillé sont ajoutés goutte à goutte jusqu'à décoloration complète de la solution. La température est laissée remonter progressivement à T.A. Après deux heures d'agitation, le solvant est évaporé. L'huile résultante est agitée une nuit dans 100 mL de cyclohexane, puis le solide obtenu est filtré et séché à la pompe, pour donner le composé **75** (30.5 g, 84 mmol) de façon quantitative

**Solide blanc** (30.5 g, quant.)

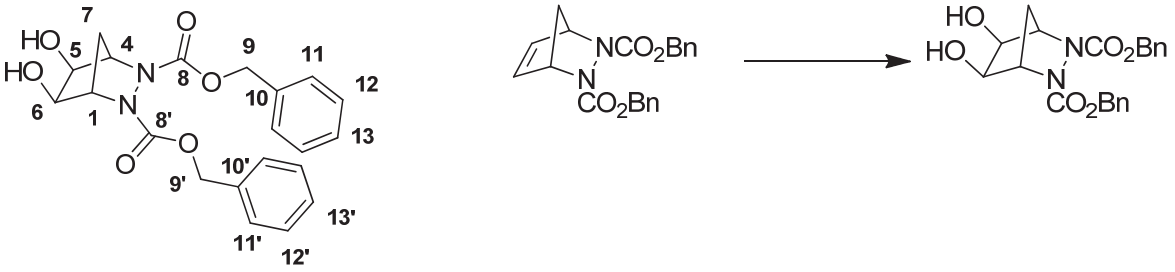
**CCM** R*f* =0.12 (Cyclohexane/ EtOAc, 90:10)

**RMN $^{1}$H** (300 MHz, CDCl3) δ 7.25-7.40 (m, 10H, H-11, H-12, H-13, H-11', H-12' et H-13'), 6.44 (s, 2H, H-5 et H-6), 5.15 (m, 6H, H1, H-4, H-9 et H-9'), 1.70 (s, 2H, H-7).

**RMN $^{13}$C** (75 MHz, CDCl3) δ 158.8 (C-8 et C-8'), 134.9 et 138.0 (C-5 et C-5'), 136.0 (C-10 et C-10'), 128.6, 128.2, 128.0 (C-11, C-12, C-13, C-11', C-12', C-13'), 68.2 (C-9 et C-9'), 65.6 (C-4 et C-1), 48.2 (C-7).

**SM** (ESI+) **:** 387 ([M+Na]$^{+}$)et 403 ([M+K]$^{+}$).

**2) 5,6-Dihydroxy-2,3-diaza-bicyclo[2.2.1]heptane-2,3-dicarboxylate de dibenzyle (77)**



**Formule brute** : $C_{21}H_{22}N_2O_6$

**M** : 398 $g.mol^{-1}$

**Mode opératoire :**

10 g du composé **75** (27,4 mmol) sont introduits dans un ballon de 2 L et dissous dans un mélange THF/eau (290 mL : 30 mL). 4 g de NMO (34,1 mmol) et 80 mg d'osmate de potassium (0.22 mmol) sont ajoutés au milieu qui prend une coloration brune-verte. Après 24 h d'agitation à T.A., de l'acide chlorhydrique 6 N (400 mL) et une solution aqueuse à 15 % de NaHSO3 (200 mL) sont ajoutés, puis l'ensemble est agité 2 h à T.A. 200 mL d'acétate d'éthyle sont ensuite ajoutés et les phases sont séparées. La phase aqueuse est réextraite par de l'acétate d'éthyle (4 x 250 mL). Les phases organiques rassemblées sont séchées sur MgSO4, filtrées et évaporées. Le produit brut est purifié par chromatographie éclair sur gel de silice (éluant : Cy/AcOEt 60 : 40) pour donner le composé **77** (8.7g, 12.6 mmol) avec un rendement de 86 %.

**Huile incolore** (8.7 g, 86 %)

**CCM** R*f* =0.24 (Cyclohexane/ EtOAc, 60:40)

**RMN $^{1}$H** (300 MHz, CDCl3, 50 °C) δ 7.26-7.32 (m, 10H, H-11, H-12, H-13, H-11', H-12' et

H-13'), 5.16 (m, 4H, H-9 et H-9'), 4.38 (sl, 2H, H-1 et H-4), 3.98 (sl, 2H, H-5 et H-6), 3.74 (sl, 2H, OH), 2.04 (d, 1H, *J* = 11.0 Hz, H-7), 1.63 (d, 1H, *J* = 11.3 Hz, H-7) ;

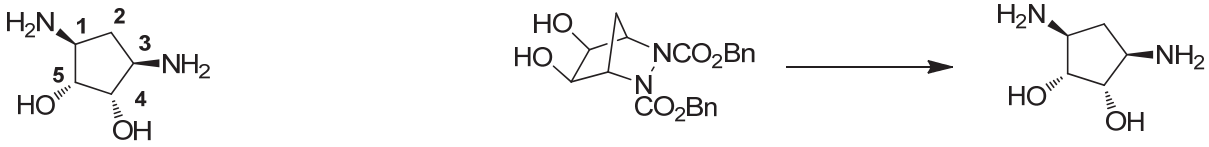
**RMN $^{13}$C** (75 MHz, CDCl3) δ 157.2 (C-8 et C-8'), 135.8 (C-10 et C-10'), 128.6, 128.3, 127.9

(C-11, C-12, C-13, C-11', C-12' et C-13'), 70.5 (C-5 et C-6), 68.4 ( C-9 et C-9'), 63.3 (C-1 et C-4),

31.2 (C-7) ;

**SM** (ES) **:** 421 ($[M+Na]^+$)et 437 ($[M+K]^+$).

### 3) 1,3-diamino-cyclopentan-4,5-diol (C5)



**Formule brute** : $C_5H_{12}N_2O_2$

**M** : 132 $g.mol^{-1}$ (205 $g.mol^{-1}$ sous forme de chlorhydrate)

**Mode opératoire :**

Le diol bicyclique **77** (4.3 g, 10.8 mmol) est agité dans 80 mL de méthanol en présence de Pd/C (10% en masse, 2.5 g, ≈ 20%mol) sous atmosphère d'hydrogène pendant trois jours. Le milieu réactionnel est filtré sur célite, lavé avec 100 mL de HCl 6M, puis évaporé au rotavapor puis séché 2 jours à la pompe pour donner 4.6 g du produit **C5** sous forme de chlorhydrate avec un rendement quantitatif.

**Solide jaune** (4.6 g, quant.)

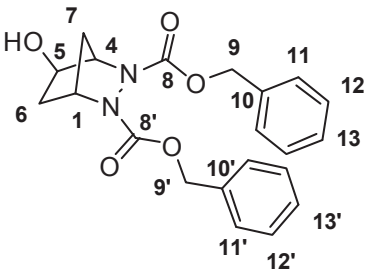
**RMN $^{1}$H** (300 MHz, MeOD-d4) δ 3.66 (m, 2H, H-4 et H5), 3.09 (ddm, J = 8.8, 8.0 Hz, 2H, H-1, H-3), 2.36 (dt, J = 13.5 et 8.0 Hz, 1H, H-2), 1.06 (dt, J = 13.5 et 8.8 Hz, 1H, H-2) ;

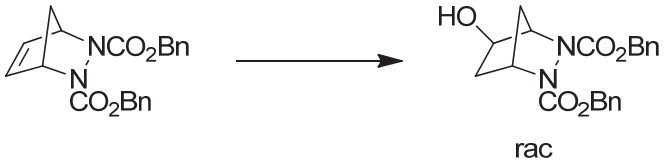
**RMN $^{13}$C** (75 MHz, MeOD-d4) δ 79.8 (C-4 et C-5), 57.0 (C-1 et C-3), 39.0 (C-2) ;

**MS** (ESI+) **:** 133 ($[M+H]^+$)et 155 ($[M+Na]^+$).

## *F - Synthèse du 1,3-diamino-4-fluoro-cyclopentane DACPF*

### 1) 5-Hydroxy-2,3-diaza-bicyclo[2.2.1]heptane-2,3-dicarboxylate de dibenzyle (100)




**Formule brute** : $C_{21}H_{22}N_2O_5$

**M** : 382 $g.mol^{-1}$

**Mode opératoire :**

L'hydrazine **75** (5 g, 13.7 mmol) est introduite dans un ballon de 250 mL avec 40 mL de THF, puis le milieu est refroidi à 0 °C. Du borane (1 M dans le THF, 16.5 mL) est additionné et l'ensemble est agité pendant 4 h à 0 °C. Du peroxyde d'hydrogène (30 % dans l'eau, 30 mL) et du NaOH (3 M dans l'eau, 75 mL) sont ensuite ajoutés doucement et le mélange biphasique est agité pendant 24 h en laissant la température remonter progressivement à température ambiante. Les deux phases sont alors séparées et la phase aqueuse est extraite plusieurs fois avec du dichlorométhane. Les phases organiques sont rassemblées, lavées avec une solution aqueuse saturée de NaCl (5 mL), séchées sur MgSO4, filtrées et évaporées. Le brut est purifié par chromatographie éclair sur gel de silice (éluant : Cy/AcOEt 50 : 50) pour donner **100** (4.6 g, 12.1 mmol) avec un rendement de 88 %.

**Huile incolore** (4.6 g, 88 %)

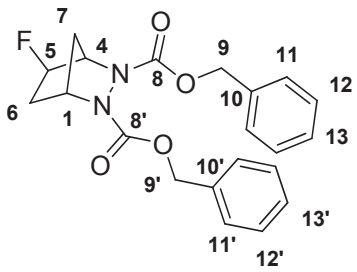
**CCM** R*f* =0.40 (cyclohexane/ EtOAc, 1:1)

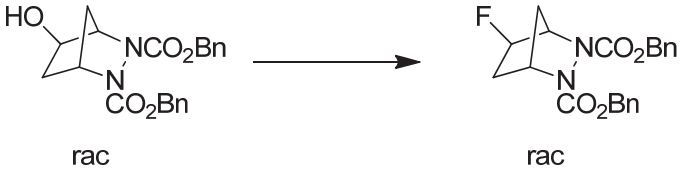
**RMN $^{1}$H** (300 MHz, CDCl3, 323 K) δ 7.32 (m, 10H, H-11, H-12, H-13, H-11', H-12', H-13'), 5.17 (m, 4H, H-9 et H-9'), 4.61 (sl, 1H, H-1), 4.39 (sl, 1H, H-4), 4.14 (sl, 1H, H-5), 2.16 (ddl, *J* = 13.8 et 6.6 Hz, 1H, H-6), 1.99 (dl, *J* = 10.8 Hz, 1H, H-7), 1.65 (dl, *J* = 10.8 Hz, 1H, H-7), 1.50 (dl, *J* = 13.8 Hz, 1H, H-6) ;

**RMN $^{13}$C** (75 MHz, CDCl3, 323 K) δ 157.9 et 156.9 (C-8 et C-8'), 136.1 et 136.0 (C-10 et C-10'), 128.5, 128.3, 128.1 et 127.9 (C-11, C-12, C-13), 70.2 (C-5), 68.1 et 68.0 (C-9 et C-9'), 64.5 (C-4), 59.7 (C-1), 39.2 (C-6), 33.9 (C-7) ;

**MS** (ESI+) **:** 405 ($[M+Na]^+$) et 421 ($[M+K]^+$).

**2) 5-Fluoro-2,3-diaza-bicyclo[2.2.1]heptane-2,3-dicarboxylate de dibenzyle 101**




**Formule brute** : $C_{21}H_{21}FN_2O_4$

**M** : 384 g.mol$^{-1}$

**Mode opératoire :**

Une solution de composé **100** (900 mg, 2.3 mmol, 1 eq.) dans le DCM (16 mL) est refroidie à 0°C avant d'ajouter le DAST (460 µL, 4.7 mmol, 2.05 eq.). Le mélange réactionnel est agité à température ambiante pendant 1 heure. On stoppe alors la réaction par l'addition d'eau (10 mL) et extraction avec du DCM (3 X 10 mL). Les phases organiques sont regroupées, lavées avec de la saumure, séchées sur $MgSO_4$, puis évaporées. Le brut réactionnel est purifié par chromatographie éclair sur gel de silice (éluant cyclohexane/EtOAc, 8:2) pour donner le produit **101** pur sous forme d'huile incolore (812 mg, 90%).

**Huile incolore** (812 mg, 90%)

**CCM** R*f* = 0.16(Cyclohexane/ EtOAc, 9:1)

**RMN $^{1}$H** (300 MHz, CDCl3, 318 K) δ 7.26 (10H, m, H-11, H-12, H-13, H-11', H-12', H-13'); 5.15 (2H, d, *J* = 12 Hz, H-9 ou H-9'); 5.1 (2H, d, *J* = 12 Hz, H-9 ou H-9'); 4.79 (1H, d, *J* = 39 Hz, H-5); 4.60 (1H, bs, H-4); 4.57 (1H, bs, H-1); 2.10 (1H, t, *J* = 14 Hz, H-6); 1.87 (1H, d, *J* = 11 Hz, H-7); 1.76 (1H, d, *J* = 14 Hz, H-6); 1.67 (1H, d, *J* = 11 Hz, H-7);

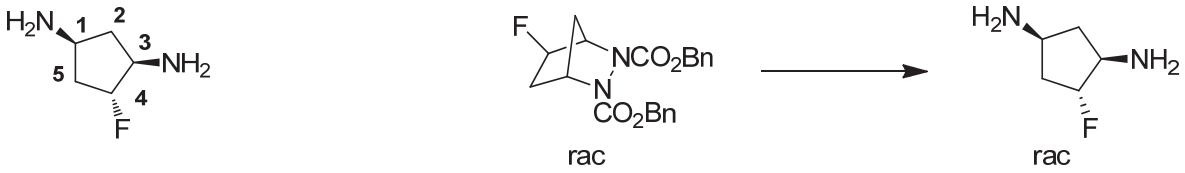
**RMN $^{13}$C** (75 MHz, CDCl3, 318 K) δ157.2 (s, C-8 ou C-8'), 156.9 (s, C-8 ou C-8'), 135.8 (s, C-10, C-10'); 128.5 (s, C-12, C-12'); 128.3 (s, C-11, C-11'); 127.9 (s, C-13, C-13'); 89 (d, *J* = 190 Hz, C-5); 68.1 (d, *J* = 8.9 Hz, C-9, C-9'); 61.9 (d, *J* = 28 Hz, C-4); 59.1 (s, C-1); 37.4 (d, *J* = 17 Hz, C-6); 34.1 (s, C-7);

**RMN $^{19}$F** (282 MHz, CDCl3, 318 K, découplage du proton) δ -170.6 (bs);

**MS** (ESI+) **:** 407($[M+Na]^+$)

**HRMS** pour $C_{21}H_{21}FN_2O_4$: ($[M+Na]^+$): calculé: 407.13830, trouvé: 407.1378 (0 ppm);

### 3) 1,3-diamino-4-fluoro-cyclopentane (DACPF)




**Formule brute** : $C_5H_{11}FN_2$

**M** : 118 g.mol$^{-1}$ (191 g.mol$^{-1}$ sous forme de chlorhydrate)

**Mode opératoire :**

Le composé **101** (727 mg; 1.9 mmol) est agité dans 5 mL de méthanol en présence de Pd/C (10% en masse, 604 mg ≈ 30%mol) sous atmosphère d'hydrogène pendant 24 h. Le milieu réactionnel est filtré sur célite, lavé avec 100 mL MeOH/DCM 1 :1, puis évaporé. Le produit est alors solubilisé dans l'EtOAc puis traité au $HCl_{(g)}$ pour donner le produit **DACPF** qui précipite sous forme de chlorhydrate (289mg, 80 % rendement).

**Poudre blanche** (289 mg, 80 %)

**RMN $^{1}$H** (300 MHz, $D_2O$) δ 5.35 (1H, ddt, *J* = 52 Hz, 7.9 Hz, 3.9 Hz, H-4), 3.86-4.04 (2H, m, H-1, H-3), 2.81 (1H, dt, *J* = 13.6 Hz, 6.6 Hz, H-2), 2.44-2.59 (1H, m, H-5), 2.30 (1H, ddt, *J* = 25 Hz, 17 Hz, 7.8 Hz, H-5), 1.88 (1H, dt, *J* = 13 Hz, 10.5 Hz, H-2) ;

**RMN $^{13}$C** (75 MHz, $D_2O$) δ 93.7 (d, *J* = 183 Hz, C-4), 54.9 (d, *J* = 25 Hz, C-3), 46.9 (s, C-1), 34.9 (d, *J* = 23 Hz, C-5), 32.7 (d, *J* = 3 Hz, C-2) ;

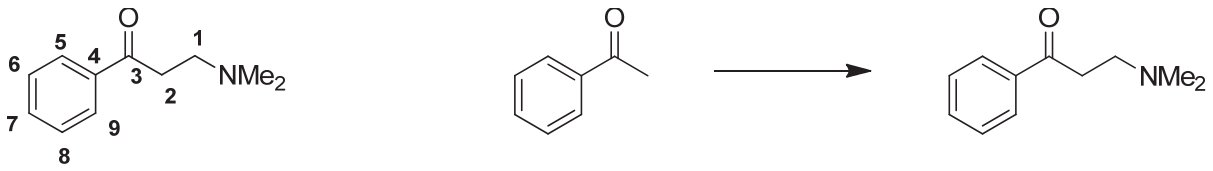
**RMN $^{19}$F** (282 MHz, $D_2O$, découplage du proton) δ -177.34 (bs);

**MS** (ESI+) **:** 119 ($[M+H]^+$).

**HRMS** pour $C_5H_{11}FN_2$: ($[M+Na]^+$): calculé: 141.0804, trouvé: 141.0795 (1 ppm);

## *G - Synthèse de β aminocétones par la réaction de Mannich*

### 1) 3-dimethylamino-1-phenyl-propane-1-one (111)




**Formule brute** : $C_{11}H_{15}NO$

**M** : 177 $g.mol^{-1}$

**Mode opératoire :**

Le paraformaldéhyde (1,43 g, 47,6 mmol, 1 éq.), le chlorhydrate de diméthylamine (11,64 g, 0,14 mol ; 3 éq.), l'acétophénone (13 mL; 0,11 mol ; 2,3 éq.), et 0,2 mL d'acide chlorhydrique 35% sont dissous dans 10 mL d'éthanol. Le mélange est agité au reflux pendant 5 heures. A l'arrêt de la réaction, de l'acétone froide (50 mL) est versée dans le mélange encore chaud, et des cristaux apparaissent. Après quelques heures à 0 °C, les cristaux sont filtrés et rincés à l'acétone froide (2x20 mL), puis séchés. Ils sont ensuite solubilisés dans 20 mL d'eau et sont extraits par 2x35 mL d'acétate d'éthyle. La phase aqueuse est ensuite basifiée par du carbonate de potassium en poudre (pH = 10), puis extraite par 5x35 mL d'acétate d'éthyle. Les phases organiques sont réunies, séchées sur MgSO4, filtrées puis concentrées sous vide pour donner 5,75 g du composé **111** sous la forme d'une huile jaune (68 %).

**Huile jaune** (6,51 g, 77%)

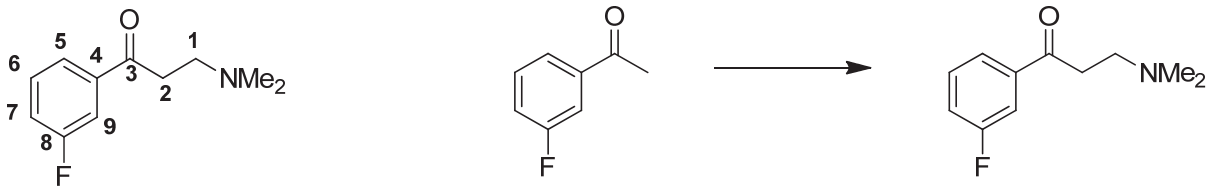
**RMN $^{1}$H** (400 MHz, CDCl3) δ 7.86 (dd, 2H, J = 7.1 Hz, 1.2 Hz, H-5, H-9), 7.44 (t, 1H, J =7,4 Hz, 1,2 Hz, H-7), 7.35 (dd, 2H, J = 7.4 Hz, 7.1 Hz, H-6, H-8), 3.05 (t, 2H, J = 7.3 Hz, H-2), 2.66 (t, 2H, J = 7.3 Hz, H-1), 2.19 (s, 6H, NMe2) ;

**RMN $^{13}$C** (100 MHz, CDCl3) δ 199.0 (C-3), 137.0 (C-4), 133.0 (C-7), 128.6 (C-5, C-9),

128.1 (C-6, C-8), 36.9 (C-2), 54.4 (C-1), 45.5 (2 Me) ;

**SM** (ES) : 178 [MH+].

**2) 3-(dimethylamino)-1-(3-fluorophenyl)propan-1-one (107)**




**Formule brute** : $C_{11}H_{14}FNO$

**M** : 195 $g.mol^{-1}$

**Mode opératoire :**

Le paraformaldéhyde (100 mg, 3,33 mmol, 1 éq.), le chlorhydrate de diméthylamine (815 mg, 10 mmol ; 3 éq.), la 3-fluoro-acétophénone (940 µL; 7,65 mmol ; 2,3 éq.), et 50 µL d'acide chlorhydrique 35% sont dissous dans 1,5 mL d'éthanol. Le mélange est agité au reflux pendant 5 heures. A l'arrêt de la réaction, de l'acétone froide (8 mL) est versée dans le mélange encore chaud, et des cristaux apparaissent. Après quelques heures à 0 °C, les cristaux sont filtrés et rincés à l'acétone froide (2x10 mL), puis séchés. Ils sont ensuite solubilisés dans 3 mL d'eau et sont extraits par 2x5 mL d'acétate d'éthyle. La phase aqueuse est ensuite basifiée par du carbonate de potassium en poudre (pH = 10), puis extraite par 5x5 mL d'acétate d'éthyle. Les phases organiques sont réunies, séchées sur MgSO4, filtrées puis concentrées sous vide pour donner 428 mg du composé **107** sous la forme d'une huile jaune (66 %).

**Huile jaune** (428 mg, 66%)

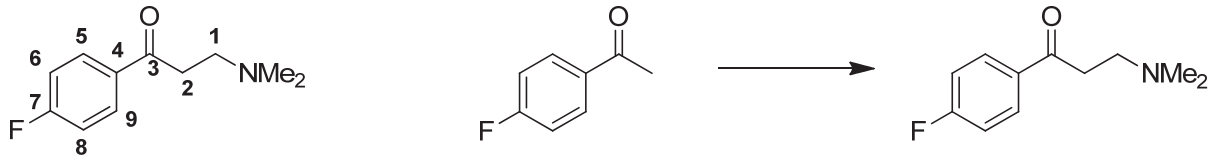
**RMN $^{1}$H** (300 MHz, CDCl3) δ 7.72 (m, 1H, H-5), 7.61 (bdt, 1H, J = 9.2 Hz, 2.4 Hz, H-9), 7.42 (m, 1H, H-7), 7.23 (btd, J =9.7 Hz, 2.8 Hz, 1H, H-6), 3.11 (t, 2H, J = 7.5 Hz, H- 2), 2.72 (t, 2H, J = 7.5 Hz, H-1), 2.26 (s, 6H, NMe2).

**RMN $^{13}$C** (75 MHz, CDCl3) δ 199.8 (d, J = 2.4 Hz, C-3), 162.9 (d, J = 248.8 Hz, C-8), 138.9 (d, J = 6.3 Hz, C-4), 130.3 (d, J = 9 Hz, C-6), 123.8 (d, J = 3.4 Hz, C-5), 120.1 (d, J = 21.9 Hz, C- 7), 114.8 (d, J = 21 Hz, C-9), 54.1 (C-1), 45.5 (2 Me), 37.0 (C-2);

**RMN $^{19}$F** (282MHz, CDCl3, découplage du proton, 318 K) δ -111.79

**SM** (ES) : 196 [MH+].

### 3) 3-(dimethylamino)-1-(4-fluorophenyl)propan-1-one (108)




**Formule brute** : $C_{11}H_{14}FNO$

**M** : 195 $g.mol^{-1}$

**Mode opératoire :**

Le paraformaldéhyde (100 mg, 3,33 mmol, 1 éq.), le chlorhydrate de diméthylamine (815 mg, 10 mmol ; 3 éq.), la 4-fluoro-acétophénone (925 µL; 7,65 mmol ; 2,3 éq.), et 50 µL d'acide chlorhydrique 35% sont dissous dans 1,5 mL d'éthanol. Le mélange est agité au reflux pendant 5 heures. A l'arrêt de la réaction, de l'acétone froide (8 mL) est versée dans le mélange encore chaud, et des cristaux apparaissent. Après quelques heures à 0 °C, les cristaux sont filtrés et rincés à l'acétone froide (2x10 mL), puis séchés. Ils sont ensuite solubilisés dans 3 mL d'eau et sont extraits par 2x5 mL d'acétate d'éthyle. La phase aqueuse est ensuite basifiée par du carbonate de potassium en poudre (pH = 10), puis extraite par 5x5 mL d'acétate d'éthyle. Les phases organiques sont réunies, séchées sur MgSO4, filtrées puis concentrées sous vide pour donner 196 mg du composé 108 sous la forme d'une huile jaune (30 %).

**Huile jaune** (196 mg, 30%)

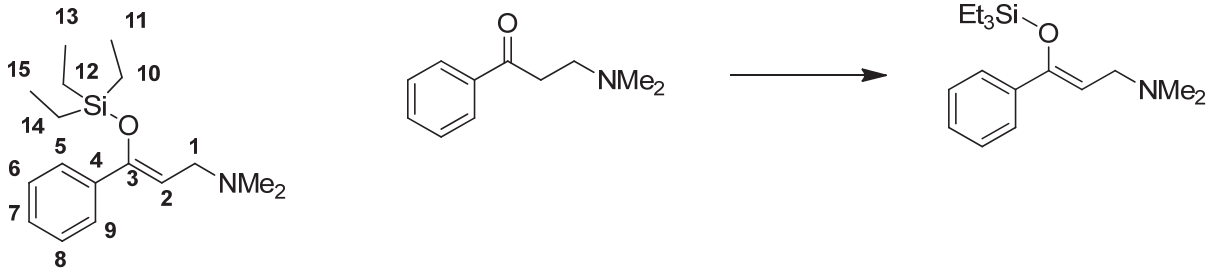
**RMN $^{1}$H** (300 MHz, CDCl3) δ 8.01 (m, 2H, H-5, H-9), 7.44 (bt, 1H, J = 8.5Hz, H-6, H-8), 3.05 (t, 2H, J = 7,3 Hz, H- 2), 2.66 (t, 2H, J = 7,3 Hz, H-1), 2.19 (s, 6H, NMe2) ;

**RMN $^{13}$C** (75 MHz, CDCl3) δ 197.5 (C-3), 165.7 (d, J = 254.5 Hz,C-7), 133.3 (d, J = 3.5 Hz, C-4), 130.7 (d, J = 8.6 Hz,C-5, C-9), 115.7 (d, J = 21.4 Hz, C-6, C-8), 54.3 (C-1), 45.5 (2 Me), 36.8 (C-2) ;

**RMN $^{19}$F** (282MHz, CDCl3, découplage du proton, 318 K) δ -111.79

**SM** (ES) : 196 [MH+].

**4) N,N-dimethyl-3-phenyl-3-((triethylsilyl)oxy)prop-2-en-1-amine (112)**




**Formule brute** : $C_{17}H_{29}NOSi$

**M** : 291 $g.mol^{-1}$

**Mode opératoire :**

On travaille sous atmosphère d'argon avec une verrerie séchée sous vide au heat gun. La kynuramine **111** est séchée au pentoxyde de phosphore $P_2O_5$ pendant 2 jours dans un dessiccateur puis dissoute dans un bicol avec 90 mL de THF fraichement distillé sur sodium/benzophénone (2,6 g; 14,6 mmol ; 1eq.). On ajoute à cette solution le TESCl (2,96 mL; 17,5 mmol, 1,2 eq.) et on agite pendant quelques minutes. Dans un tricol, on dilue 14,6 mL de diisopropylamidure de lithium (LDA)(solution à 1M dans le THF, 14,6 mmol ; 1 eq.) avec 45 mL de THF fraichement distillé et on refroidit le mélange à -78°C. Après avoir agité la solution pendant quelques minutes, on transfère goutte à goutte la solution du bicol (composé **111** et TESCl) dans le tricol. On agite pendant 2 heures à -78°C puis pendant 2 heures à température ambiante avant de tirer à sec. Le résidu est alors trituré dans le pentane (3x75mL) pour obtenir une huile jaune pâle (4,3 g, rendement quant.). Le brut réactionnel est très propre (quelques traces de TESCl seulement) et ne nécessite pas nécessairement de purification.

Il peut néanmoins être purifié par flash chromatographie sur colonne de silice (DCM/MeOH/$Et_3N$, 90/9/1) pour obtenir le composé **112** avec un modeste rendement de 12% dû à la déprotection de l'éther d'énol silylé pour redonner le composé de départ.

**Huile incolore** (500mg, 12%)[302]

**RMN $^1$H** (300 MHz, CDCl3) δ 7.86 (dd, 2H, J = 7.1 Hz, 1.2 Hz, H-5, H-9), 7.44 (t, 1H, J =

7.4 Hz, 1.2 Hz, H-7), 7.35 (dd, 2H, J = 7.4 Hz, 7.1 Hz, H-6, H-8), 3.05 (t, 2H, J = 7.3 Hz, H-2), 2.66 (t, 2H, J = 7.3 Hz, H-1), 2.19 (s, 6H, NMe2) ;

**RMN $^{13}$C** (75 MHz, CDCl3) δ 151.9 (C-3), 139.2 (C-4), 128.0 (C-6, C-7), 127.9 (C-8),

125.8 (C-5, C-9), 107.4 (C-2), 55.2 (C-1), 36.9 (C-2); 45.2 (2 Me), 6.7 (C-10, C-12, C-14), 5.3 (C-11, C-13, C-15) ;

# VI. Bibliographie

(1) Rozenski, J.; Crain, P. F.; McCloskey, J. a The RNA Modification Database: 1999 update. *Nucleic acids research* **1999**, *27*, 196–7.

(2) Dunin-Horkawicz, S.; Czerwoniec, A.; Gajda, M. J.; Feder, M.; Grosjean, H.; Bujnicki, J. M. MODOMICS: a database of RNA modification pathways. *Nucleic acids research* **2006**, *34*, D145–9.

(3) Carell, T.; Brandmayr, C.; Hienzsch, A.; Müller, M.; Pearson, D.; Reiter, V.; Thoma, I.; Thumbs, P.; Wagner, M. Structure and function of noncanonical nucleobases. *Angewandte Chemie (International ed. in English)* **2012**, *51*, 7110–31.

(4) Tinoco, I.; Bustamante, C. How RNA folds. *Journal of molecular biology* **1999**.

(5) Shi, H.; Moore, P. B. The crystal structure of yeast phenylalanine tRNA at 1.93 A resolution: a classic structure revisited. *RNA (New York, N.Y.)* **2000**, *6*, 1091–105.

(6) Leontis, N. B.; Westhof, E. Geometric nomenclature and classification of RNA base pairs. *RNA (New York, N.Y.)* **2001**, *7*, 499–512.

(7) Batey, R.; Rambo, R.; Doudna, J. Tertiary Motifs in RNA Structure and Folding. *Angewandte Chemie (International ed. in English)* **1999**, *38*, 2326–2343.

(8) Cruz, J. A.; Westhof, E. The dynamic landscapes of RNA architecture. *Cell* **2009**, *136*, 604–9.

(9) Fürtig, B.; Wenter, P.; Pitsch, S.; Schwalbe, H. Probing mechanism and transition state of RNA refolding. *ACS chemical biology* **2010**, *5*, 753–65.

(10) Rosenberg, J. M.; Seeman, N. C.; Kim, J. J.; Suddath, F. L.; Nicholas, H. B.; Rich, A. Double helix at atomic resolution. *Nature* **1973**, *243*, 150–4.

(11) Leontis, N. B.; Lescoute, A.; Westhof, E. The building blocks and motifs of RNA architecture. *Current opinion in structural biology* **2006**, *16*, 279–87.

(12) Vicens, Q.; Westhof, E. RNA as a drug target: the case of aminoglycosides. *Chembiochem : a European journal of chemical biology* **2003**, *4*, 1018–23.

(13) JACOB, F.; MONOD, J. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of molecular biology* **1961**, *3*, 318–56.

(14) Thomas, J. R.; Hergenrother, P. J. Targeting RNA with small molecules. *Chemical reviews* **2008**, *108*, 1171–224.

(15) Avery, O. T.; Macleod, C. M.; McCarty, M. Studies on the chemical nature of the substance inducing transformation of the Pneumococcal types : induction of transformation by a

desoxyribonucleic acid isolated from Pneumococcus type III. *The Journal of experimental medicine* **1944**, *79*, 137–58.

(16) Watson, J. D.; Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **1953**, *171*, 737–8.

(17) Crick, F. Central dogma of molecular biology. *Nature* **1970**, *227*, 561–563.

(18) Crick, F. The biological replication of macromolecules. *Symp.Soc.Exp.Biol.* **1958**, *XII*, 138.

(19) Ramakrishnan, V. Ribosome structure and the mechanism of translation. *Cell* **2002**, *108*, 557–72.

(20) Mattick, J. S. RNA regulation: a new genetics? *Nature reviews. Genetics* **2004**, *5*, 316–23.

(21) Nóbrega, M. A.; Zhu, Y.; Plajzer-Frick, I.; Afzal, V.; Rubin, E. M. Megabase deletions of gene deserts result in viable mice. *Nature* **2004**, *431*, 988–93.

(22) Biémont, C.; Vieira, C. Genetics: junk DNA as an evolutionary force. *Nature* **2006**, *443*, 521–4.

(23) Winkler, W. C.; Breaker, R. R. Genetic control by metabolite-binding riboswitches. *Chembiochem : a European journal of chemical biology* **2003**, *4*, 1024–32.

(24) Mandal, M.; Boese, B.; Barrick, J. E.; Winkler, W. C.; Breaker, R. R. Riboswitches control fundamental biochemical pathways in Bacillus subtilis and other bacteria. *Cell* **2003**, *113*, 577–86.

(25) Chabanon, H.; Mickleburgh, I.; Hesketh, J. Zipcodes and postage stamps: mRNA localisation signals and their trans-acting binding proteins. *Briefings in functional genomics & proteomics* **2004**, *3*, 240–56.

(26) Hesketh, J. 3'-Untranslated regions are important in mRNA localization and translation: lessons from selenium and metallothionein. *Biochemical Society transactions* **2004**, *32*, 990–3.

(27) Soukup, J. K.; Soukup, G. a Riboswitches exert genetic control through metabolite-induced conformational change. *Current opinion in structural biology* **2004**, *14*, 344–9.

(28) Ellington, A. D.; Szostak, J. W. In vitro selection of RNA molecules that bind specific ligands. *Nature* **1990**, *346*, 818–22.

(29) Zimmermann, G. R.; Jenison, R. D.; Wick, C. L.; Simorre, J. P.; Pardi, A. Interlocking structural motifs mediate molecular discrimination by a theophylline-binding RNA. *Nature structural biology* **1997**, *4*, 644–9.

(30) Sudarsan, N.; Barrick, J. E.; Breaker, R. R. Metabolite-binding RNA domains are present in the genes of eukaryotes. *RNA (New York, N.Y.)* **2003**, *9*, 644–7.

(31) Mironov, A. S.; Gusarov, I.; Rafikov, R.; Lopez, L. E.; Shatalin, K.; Kreneva, R. A.; Perumov, D. A.; Nudler, E. Sensing small molecules by nascent RNA: a mechanism to control transcription in bacteria. *Cell* **2002**, *111*, 747–56.

(32) Serganov, A.; Huang, L.; Patel, D. J. Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. *Nature* **2009**, *458*, 233–7.

(33) McDaniel, B. A. M.; Grundy, F. J.; Artsimovitch, I.; Henkin, T. M. Transcription termination control of the S box system: direct measurement of S-adenosylmethionine by the leader RNA. *Proceedings of the National Academy of Sciences of the United States of America* **2003**, *100*, 3083–8.

(34) Lee, E. R.; Blount, K. F.; Breaker, R. R. Roseoflavin is a natural antibacterial compound that binds to FMN riboswitches and regulates gene expression. *RNA biology* **2009**, *6*, 1–8.

(35) Montange, R. K.; Batey, R. T. Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature* **2006**, *441*, 1172–5.

(36) Weigand, J. E.; Schmidtke, S. R.; Will, T. J.; Duchardt-Ferner, E.; Hammann, C.; Wöhnert, J.; Suess, B. Mechanistic insights into an engineered riboswitch: a switching element which confers riboswitch activity. *Nucleic acids research* **2011**, *39*, 3363–72.

(37) Werstuck, G.; Green, M. R. Controlling gene expression in living cells through small molecule-RNA interactions. *Science (New York, N.Y.)* **1998**, *282*, 296–8.

(38) Weigand, J. E.; Suess, B. Aptamers and riboswitches: perspectives in biotechnology. *Applied microbiology and biotechnology* **2009**, *85*, 229–36.

(39) Schluenzen, F.; Tocilj, A.; Zarivach, R.; Harms, J.; Gluehmann, M.; Janell, D.; Bashan, A.; Bartels, H.; Agmon, I.; Franceschi, F.; Yonath, A. Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell* **2000**, *102*, 615–23.

(40) Ban, N.; Nissen, P.; Hansen, J.; Moore, P. B.; Steitz, T. A. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. *Science (New York, N.Y.)* **2000**, *289*, 905–20.

(41) Wimberly, B. T.; Brodersen, D. E.; Clemons, W. M.; Morgan-Warren, R. J.; Carter, A. P.; Vonrhein, C.; Hartsch, T.; Ramakrishnan, V. Structure of the 30S ribosomal subunit. *Nature* **2000**, *407*, 327–39.

(42) Clardy, J.; Walsh, C. Lessons from natural molecules. *Nature* **2004**, *432*, 829–37.

(43) Newman, D. J.; Cragg, G. M. Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010. *Journal of natural products* **2012**, *75*, 311–35.

(44) Fourmy, D.; Recht, M. I.; Puglisi, J. D. Binding of neomycin-class aminoglycoside antibiotics to the A-site of 16 S rRNA. *Journal of molecular biology* **1998**, *277*, 347–62.

(45) Fourmy, D.; Yoshizawa, S.; Puglisi, J. D. Paromomycin binding induces a local conformational change in the A-site of 16 S rRNA. *Journal of molecular biology* **1998**, *277*, 333–45.

(46) Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* **1983**, *35*, 849–57.

(47) Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **1982**, *31*, 147–57.

(48) Reiter, N. J.; Osterman, A.; Torres-Larios, A.; Swinger, K. K.; Pan, T.; Mondragón, A. Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature* **2010**, *468*, 784–9.

(49) Masquida, B.; Westhof, E. RNase P: at last, the key finds its lock. *RNA (New York, N.Y.)* **2011**, *17*, 1615–8.

(50) Winkler, W. C.; Nahvi, A.; Roth, A.; Collins, J. a; Breaker, R. R. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **2004**, *428*, 281–6.

(51) Cech, T. R. RNA finds a simpler way. *Nature* **2004**, *428*, 263–4.

(52) Gilbert, W. Origin of life: The RNA world. *Nature* **1986**, *319*, 618–618.

(53) Cech, T. R. The RNA Worlds in Context. *Cold Spring Harbor perspectives in biology* **2011**.

(54) Giegé, R.; Frugier, M.; Rudinger, J. tRNA mimics. *Current opinion in structural biology* **1998**, *8*, 286–93.

(55) Barré-Sinoussi, F.; Chermann, J. C.; Rey, F.; Nugeyre, M. T.; Chamaret, S.; Gruest, J.; Dauguet, C.; Axler-Blin, C.; Vézinet-Brun, F.; Rouzioux, C.; Rozenbaum, W.; Montagnier, L. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science (New York, N.Y.)* **1983**, *220*, 868–71.

(56) Temin, H. M.; Mizutani, S. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* **1970**, *226*, 1211–3.

(57) Baltimore, D. RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* **1970**, *226*, 1209–11.

(58) Tisné, C. Structural bases of the annealing of primer tRNA(3Lys) to the HIV-1 viral RNA. *Current HIV research* **2005**, *3*, 147–56.

(59) Tisné, C.; Roques, B. P.; Dardel, F. The annealing mechanism of HIV-1 reverse transcription primer onto the viral genome. *The Journal of biological chemistry* **2004**, *279*, 3588–95.

(60) Barraud, P.; Gaudin, C.; Dardel, F.; Tisné, C. New insights into the formation of HIV-1 reverse transcription initiation complex. *Biochimie* **2007**, *89*, 1204–10.

(61) Sleiman, D.; Goldschmidt, V.; Barraud, P.; Marquet, R.; Paillart, J.-C.; Tisné, C. Initiation of HIV-1 reverse transcription and functional role of nucleocapsid-mediated tRNA/viral genome interactions. *Virus research* **2012**.

(62) Chung, F.; Tisné, C.; Lecourt, T.; Dardel, F.; Micouin, L. NMR-guided fragment-based approach for the design of tRNA(Lys3) ligands. *Angewandte Chemie (International ed. in English)* **2007**, *46*, 4489–91.

(63) Moumné, R.; Catala, M.; Larue, V.; Micouin, L.; Tisné, C. Fragment-based design of small RNA binders: Promising developments and contribution of NMR. *Biochimie* **2012**, *94*, 1607–19.

(64) Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E.; Mello, C. C. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **1998**, *391*, 806–11.

(65) Mello, C. C.; Conte, D. Revealing the world of RNA interference. *Nature* **2004**, *431*, 338–42.

(66) Sharp, P. A. The centrality of RNA. *Cell* **2009**, *136*, 577–80.

(67) Carthew, R. W.; Sontheimer, E. J. Origins and Mechanisms of miRNAs and siRNAs. *Cell* **2009**, *136*, 642–55.

(68) Friedman, R. C.; Farh, K. K.-H.; Burge, C. B.; Bartel, D. P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome research* **2009**, *19*, 92–105.

(69) Esteller, M. Non-coding RNAs in human disease. *Nature reviews. Genetics* **2011**, *12*, 861–74.

(70) Burnett, J. C.; Rossi, J. J. RNA-Based Therapeutics: Current Progress and Future Prospects. *Chemistry & biology* **2012**, *19*, 60–71.

(71) Cooper, T. a; Wan, L.; Dreyfuss, G. RNA and disease. *Cell* **2009**, *136*, 777–93.

(72) Hopkins, A. L.; Groom, C. R. The druggable genome. *Nature reviews. Drug discovery* **2002**, *1*, 727–30.

(73) Overington, J. P.; Al-Lazikani, B.; Hopkins, A. L. How many drug targets are there? *Nature reviews. Drug discovery* **2006**, *5*, 993–6.

(74) Zaman, G. J. R.; Michiels, P. J. a; van Boeckel, C. a a Targeting RNA: new opportunities to address drugless targets. *Drug discovery today* **2003**, *8*, 297–306.

(75) Hoffmann, T.; Metternich, R. The Future of Medicinal Chemistry. *Angewandte Chemie International Edition* **2012**, *51*, 2–4.

(76) Lipinski, C. a; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews* **2001**, *46*, 3–26.

(77) Foloppe, N.; Matassova, N.; Aboul-Ela, F. Towards the discovery of drug-like RNA ligands? *Drug discovery today* **2006**, *11*, 1019–27.

(78) Aboul-ela, F. Strategies for the design of RNA-binding small molecules. *Future medicinal chemistry* **2010**, *2*, 93–119.

(79) Fulle, S.; Gohlke, H. Molecular recognition of RNA: challenges for modelling interactions and plasticity. *Journal of molecular recognition : JMR* **2010**, *23*, 220–31.

(80) Vicens, Q.; Westhof, E. Molecular recognition of aminoglycoside antibiotics by ribosomal RNA and resistance enzymes: an analysis of x-ray crystal structures. *Biopolymers* **2003**, *70*, 42–57.

(81) Hermann, T. Rational ligand design for RNA: the role of static structure and conformational flexibility in target recognition. *Biochimie* **2002**, *84*, 869–75.

(82) Dethoff, E. A.; Chugh, J.; Mustoe, A. M.; Al-Hashimi, H. M. Functional complexity and regulation through RNA dynamics. *Nature* **2012**, *482*, 322–30.

(83) Leulliot, N.; Varani, G. Current topics in RNA-protein recognition: control of specificity and biological function through induced fit and conformational capture. *Biochemistry* **2001**, *40*, 7947–56.

(84) Hammond, M. C. RNA folding: a tale of two riboswitches. *Nature chemical biology* **2011**, *7*, 342–3.

(85) Walter, F.; Pütz, J.; Giegé, R.; Westhof, E. Binding of tobramycin leads to conformational changes in yeast tRNA(Asp) and inhibition of aminoacylation. *The EMBO journal* **2002**, *21*, 760–8.

(86) Guan, L.; Disney, M. D. Recent advances in developing small molecules targeting RNA. *ACS chemical biology* **2012**, *7*, 73–86.

(87) Dervan, P. B.; Edelson, B. S. Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Current opinion in structural biology* **2003**, *13*, 284–99.

(88) Magnet, S.; Blanchard, J. S. Molecular insights into aminoglycoside action and resistance. *Chemical reviews* **2005**, *105*, 477–98.

(89) Tor, Y. The ribosomal A-site as an inspiration for the design of RNA binders. *Biochimie* **2006**, *88*, 1045–51.

(90) Vakulenko, S. B.; Mobashery, S. Versatility of aminoglycosides and prospects for their future. *Clinical microbiology reviews* **2003**, *16*, 430–50.

(91) Schatz, A.; Bugle, E.; Waksman, S. A. Streptomycin, a Substance Exhibiting Antibiotic Activity Against Gram-Positive and Gram-Negative Bacteria.∗†. *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)* **1944**, *55*, 66–69.

(92) Walsh, C. Antibiotics: actions, origins, resistance. *ASM Press* **2003**.

(93) Fernandes, P. Antibacterial discovery and development--the failure of success? *Nature biotechnology* **2006**, *24*, 1497–503.

(94) Coates, A.; Hu, Y.; Bax, R.; Page, C. The future challenges facing the development of new antimicrobial drugs. *Nature reviews. Drug discovery* **2002**, *1*, 895–910.

(95) EMEA The bacterial challenge: time to react. **2009**.

(96) Boucher, H. W.; Talbot, G. H.; Bradley, J. S.; Edwards, J. E.; Gilbert, D.; Rice, L. B.; Scheld, M.; Spellberg, B.; Bartlett, J. Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **2009**, *48*, 1–12.

(97) Houghton, J. L.; Green, K. D.; Chen, W.; Garneau-Tsodikova, S. The future of aminoglycosides: the end or renaissance? *Chembiochem : a European journal of chemical biology* **2010**, *11*, 880–902.

(98) Kaul, M.; Barbieri, C. M.; Pilch, D. S. Defining the basis for the specificity of aminoglycoside-rRNA recognition: a comparative study of drug binding to the A sites of Escherichia coli and human rRNA. *Journal of molecular biology* **2005**, *346*, 119–34.

(99) Wang, H.; Tor, Y. RNA–Aminoglycoside Interactions: Design, Synthesis, and Binding of "Amino-aminoglycosides" to RNA. *Angewandte Chemie International Edition* **1998**, *37*, 109–111.

(100) Hermann, T.; Westhof, E. Docking of cationic antibiotics to negatively charged pockets in RNA folds. *Journal of medicinal chemistry* **1999**, *42*, 1250–61.

(101) Wang, H.; Tor, Y. Electrostatic Interactions in RNA Aminoglycosides Binding. *Journal of the American Chemical Society* **1997**, *119*, 8734–8735.

(102) Hendrix, M.; Priestley, E. S.; Joyce, G. F.; Wong, C. H. Direct observation of aminoglycoside-RNA interactions by surface plasmon resonance. *Journal of the American Chemical Society* **1997**, *119*, 3641–8.

(103) Verhelst, S. H. L.; Michiels, P. J. A.; van der Marel, G. A.; van Boeckel, C. A. A.; van Boom, J. H. Surface plasmon resonance evaluation of various aminoglycoside-RNA hairpin interactions reveals low degree of selectivity. *Chembiochem : a European journal of chemical biology* **2004**, *5*, 937–42.

(104) Yoshizawa, S.; Fourmy, D.; Puglisi, J. D. Structural origins of gentamicin antibiotic action. *The EMBO journal* **1998**, *17*, 6437–48.

(105) Carter, A. P.; Clemons, W. M.; Brodersen, D. E.; Morgan-Warren, R. J.; Wimberly, B. T.; Ramakrishnan, V. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. *Nature* **2000**, *407*, 340–8.

(106) François, B.; Russell, R. J. M.; Murray, J. B.; Aboul-ela, F.; Masquida, B.; Vicens, Q.; Westhof, E. Crystal structures of complexes between aminoglycosides and decoding A site oligonucleotides: role of the number of rings and positive charges in the specific binding leading to miscoding. *Nucleic acids research* **2005**, *33*, 5677–90.

(107) Han, Q.; Zhao, Q.; Fish, S.; Simonsen, K. B.; Vourloumis, D.; Froelich, J. M.; Wall, D.; Hermann, T. Molecular recognition by glycoside pseudo base pairs and triples in an apramycin-RNA complex. *Angewandte Chemie (International ed. in English)* **2005**, *44*, 2694–700.

(108) Moitessier, N.; Westhof, E.; Hanessian, S. Docking of aminoglycosides to hydrated and flexible RNA. *Journal of medicinal chemistry* **2006**, *49*, 1023–33.

(109) Davies, J.; Gilbert, W.; Gorini, L. Streptomycin, suppression, and the code. *Proceedings of the National Academy of Sciences of the United States of America* **1964**, *51*, 883–90.

(110) Moazed, D.; Noller, H. F. Interaction of antibiotics with functional sites in 16S ribosomal RNA. *Nature* **1987**, *327*, 389–94.

(111) Recht, M. I.; Fourmy, D.; Blanchard, S. C.; Dahlquist, K. D.; Puglisi, J. D. RNA sequence determinants for aminoglycoside binding to an A-site rRNA model oligonucleotide. *Journal of molecular biology* **1996**, *262*, 421–36.

(112) Fourmy, D.; Recht, M. I.; Blanchard, S. C.; Puglisi, J. D. Structure of the A site of Escherichia coli 16S ribosomal RNA complexed with an aminoglycoside antibiotic. *Science (New York, N.Y.)* **1996**, *274*, 1367–71.

(113) Yoshizawa, S.; Fourmy, D.; Puglisi, J. D. Recognition of the codon-anticodon helix by ribosomal RNA. *Science (New York, N.Y.)* **1999**, *285*, 1722–5.

(114) Vicens, Q.; Westhof, E. Crystal structure of paromomycin docked into the eubacterial ribosomal decoding A site. *Structure (London, England : 1993)* **2001**, *9*, 647–58.

(115) Vicens, Q.; Westhof, E. Crystal structure of a complex between the aminoglycoside tobramycin and an oligonucleotide containing the ribosomal decoding a site. *Chemistry & biology* **2002**, *9*, 747–55.

(116) Griffey, R. H.; Hofstadler, S. A.; Sannes-Lowery, K. A.; Ecker, D. J.; Crooke, S. T. Determinants of aminoglycoside-binding specificity for rRNA by using mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 10129–33.

(117) Wong, C. H.; Hendrix, M.; Priestley, E. S.; Greenberg, W. A. Specificity of aminoglycoside antibiotics for the A-site of the decoding region of ribosomal RNA. *Chemistry & biology* **1998**, *5*, 397–406.

(118) Barbieri, C. M.; Srinivasan, A. R.; Pilch, D. S. Deciphering the origins of observed heat capacity changes for aminoglycoside binding to prokaryotic and eukaryotic ribosomal RNA a-sites: a calorimetric, computational, and osmotic stress study. *Journal of the American Chemical Society* **2004**, *126*, 14380–8.

(119) Kaul, M.; Barbieri, C. M.; Pilch, D. S. Fluorescence-based approach for detecting and characterizing antibiotic-induced conformational changes in ribosomal RNA: comparing aminoglycoside binding to prokaryotic and eukaryotic ribosomal RNA sequences. *Journal of the American Chemical Society* **2004**, *126*, 3447–53.

(120) Shandrick, S.; Zhao, Q.; Han, Q.; Ayida, B. K.; Takahashi, M.; Winters, G. C.; Simonsen, K. B.; Vourloumis, D.; Hermann, T. Monitoring molecular recognition of the ribosomal decoding site. *Angewandte Chemie (International ed. in English)* **2004**, *43*, 3177–82.

(121) Kramer, E. B.; Farabaugh, P. J. The frequency of translational misreading errors in E. coli is largely determined by tRNA competition. *RNA (New York, N.Y.)* **2007**, *13*, 87–96.

(122) Ogle, J. M.; Brodersen, D. E.; Clemons, W. M.; Tarry, M. J.; Carter, A. P.; Ramakrishnan, V. Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science (New York, N.Y.)* **2001**, *292*, 897–902.

(123) Mingeot-Leclercq, M. P.; Tulkens, P. M. Aminoglycosides: nephrotoxicity. *Antimicrobial agents and chemotherapy* **1999**, *43*, 1003–12.

(124) Moestrup, S. K.; Cui, S.; Vorum, H.; Bregengård, C.; Bjørn, S. E.; Norris, K.; Gliemann, J.; Christensen, E. I. Evidence that epithelial glycoprotein 330/megalin mediates uptake of polybasic drugs. *The Journal of clinical investigation* **1995**, *96*, 1404–13.

(125) Wright, G. D. Mechanisms of resistance to antibiotics. *Current opinion in chemical biology* **2003**, *7*, 563–9.

(126) Yoneyama, H.; Katsumata, R. Antibiotic resistance in bacteria and its future for novel antibiotic development. *Bioscience, biotechnology, and biochemistry* **2006**, *70*, 1060–75.

(127) Cundliffe, E. How antibiotic-producing organisms avoid suicide. *Annual review of microbiology* **1989**, *43*, 207–33.

(128) Böttger, E. C.; Springer, B.; Prammananan, T.; Kidan, Y.; Sander, P. Structural basis for selectivity and toxicity of ribosomal antibiotics. *EMBO reports* **2001**, *2*, 318–23.

(129) Casin, I.; Hanau-Berçot, B.; Podglajen, I.; Vahaboglu, H.; Collatz, E. Salmonella enterica serovar Typhimurium bla(PER-1)-carrying plasmid pSTI1 encodes an extended-spectrum aminoglycoside 6'-N-acetyltransferase of type Ib. *Antimicrobial agents and chemotherapy* **2003**, *47*, 697–703.

(130) Lombès, T.; Bégis, G.; Maurice, F.; Turcaud, S.; Lecourt, T.; Dardel, F.; Micouin, L. NMR-guided fragment-based approach for the design of AAC(6')-Ib ligands. *Chembiochem : a European journal of chemical biology* **2008**, *9*, 1368–71.

(131) Dietz, H. C. New therapeutic approaches to mendelian disorders. *The New England journal of medicine* **2010**, *363*, 852–63.

(132) Rowe, S. M.; Clancy, J. P. Pharmaceuticals targeting nonsense mutations in genetic diseases: progress in development. *BioDrugs : clinical immunotherapeutics, biopharmaceuticals and gene therapy* **2009**, *23*, 165–74.

(133) Gorini, L.; Kataja, E. Phenotypic repair by Streptomycin of defective genotypes in E.coli. *Proceedings of the National Academy of Sciences of the United States of America* **1964**, *51*, 487–93.

(134) Burke, J. F.; Mogg, A. E. Suppression of a nonsense mutation in mammalian cells in vivo by the aminoglycoside antibiotics G-418 and paromomycin. *Nucleic acids research* **1985**, *13*, 6265–72.

(135) Bedwell, D. M.; Kaenjak, A.; Benos, D. J.; Bebok, Z.; Bubien, J. K.; Hong, J.; Tousson, A.; Clancy, J. P.; Sorscher, E. J. Suppression of a CFTR premature stop mutation in a bronchial epithelial cell line. *Nature medicine* **1997**, *3*, 1280–4.

(136) Wilschanski, M.; Yahav, Y.; Yaacov, Y.; Blau, H.; Bentur, L.; Rivlin, J.; Aviram, M.; Bdolah-Abram, T.; Bebok, Z.; Shushi, L.; Kerem, B.; Kerem, E. Gentamicin-induced correction of CFTR function in patients with cystic fibrosis and CFTR stop mutations. *The New England journal of medicine* **2003**, *349*, 1433–41.

(137) Nudelman, I.; Rebibo-Sabbah, A.; Cherniavsky, M.; Belakhov, V.; Hainrichson, M.; Chen, F.; Schacht, J.; Pilch, D. S.; Ben-Yosef, T.; Baasov, T. Development of novel aminoglycoside (NB54) with reduced toxicity and enhanced suppression of disease-causing premature stop mutations. *Journal of medicinal chemistry* **2009**, *52*, 2836–45.

(138) Brook, J. D.; McCurrach, M. E.; Harley, H. G.; Buckler, A. J.; Church, D.; Aburatani, H.; Hunter, K.; Stanton, V. P.; Thirion, J. P.; Hudson, T. Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. *Cell* **1992**, *68*, 799–808.

(139) Warf, M. B.; Berglund, J. A. MBNL binds similar RNA structures in the CUG repeats of myotonic dystrophy and its pre-mRNA substrate cardiac troponin T. *RNA (New York, N.Y.)* **2007**, *13*, 2238–51.

(140) Cardani, R.; Mancinelli, E.; Rotondo, G.; Sansone, V.; Meola, G. Muscleblind-like protein 1 nuclear sequestration is a molecular pathology marker of DM1 and DM2. *European journal of histochemistry : EJH* **2006**, *50*, 177–82.

(141) Lee, M. M.; Childs-Disney, J. L.; Pushechnikov, A.; French, J. M.; Sobczak, K.; Thornton, C. A.; Disney, M. D. Controlling the specificity of modularly assembled small molecules for RNA via ligand module spacing: targeting the RNAs that cause myotonic muscular dystrophy. *Journal of the American Chemical Society* **2009**, *131*, 17464–72.

(142) Lee, M. M.; Pushechnikov, A.; Disney, M. D. Rational and modular design of potent ligands targeting the RNA that causes myotonic dystrophy 2. *ACS chemical biology* **2009**, *4*, 345–55.

(143) Lapidot, A.; Berchanski, A.; Borkow, G. Insight into the mechanisms of aminoglycoside derivatives interaction with HIV-1 entry steps and viral gene transcription. *The FEBS journal* **2008**, *275*, 5236–57.

(144) Brady, J.; Kashanchi, F. Tat gets the "green" light on transcription initiation. *Retrovirology* **2005**, *2*, 69.

(145) Mei, H. Y.; Mack, D. P.; Galan, A. A.; Halim, N. S.; Heldsinger, A.; Loo, J. A.; Moreland, D. W.; Sannes-Lowery, K. A.; Sharmeen, L.; Truong, H. N.; Czarnik, A. W. Discovery of selective, small-molecule inhibitors of RNA complexes--I. The Tat protein/TAR RNA complexes required for HIV-1 transcription. *Bioorganic & medicinal chemistry* **1997**, *5*, 1173–84.

(146) Blount, K. F.; Tor, Y. A tale of two targets: differential RNA selectivity of nucleobase-aminoglycoside conjugates. *Chembiochem : a European journal of chemical biology* **2006**, *7*, 1612–21.

(147) Blount, K. F.; Zhao, F.; Hermann, T.; Tor, Y. Conformational constraint as a means for understanding RNA-aminoglycoside specificity. *Journal of the American Chemical Society* **2005**, *127*, 9818–29.

(148) Frankel, A. D.; Young, J. A. HIV-1: fifteen proteins and an RNA. *Annual review of biochemistry* **1998**, *67*, 1–25.

(149) Zapp, M. L.; Stern, S.; Green, M. R. Small molecules that selectively block RNA binding of HIV-1 Rev protein inhibit Rev function and viral production. *Cell* **1993**, *74*, 969–78.

(150) Lacourciere, K. A.; Stivers, J. T.; Marino, J. P. Mechanism of neomycin and Rev peptide binding to the Rev responsive element of HIV-1 as determined by fluorescence and NMR spectroscopy. *Biochemistry* **2000**, *39*, 5630–41.

(151) Cho, J.; Rando, R. R. Specificity in the binding of aminoglycosides to HIV-RRE RNA. *Biochemistry* **1999**, *38*, 8548–54.

(152) Dibrov, S. M.; Johnston-Cox, H.; Weng, Y.-H.; Hermann, T. Functional architecture of HCV IRES domain II stabilized by divalent metal ions in the crystal and in solution. *Angewandte Chemie (International ed. in English)* **2007**, *46*, 226–9.

(153) Spahn, C. M.; Kieft, J. S.; Grassucci, R. A.; Penczek, P. A.; Zhou, K.; Doudna, J. A.; Frank, J. Hepatitis C virus IRES RNA-induced changes in the conformation of the 40s ribosomal subunit. *Science (New York, N.Y.)* **2001**, *291*, 1959–62.

(154) Seth, P. P.; Miyaji, A.; Jefferson, E. A.; Sannes-Lowery, K. A.; Osgood, S. A.; Propp, S. S.; Ranken, R.; Massire, C.; Sampath, R.; Ecker, D. J.; Swayze, E. E.; Griffey, R. H. SAR by MS: discovery of a new class of RNA-binding small molecules for the hepatitis C virus: internal ribosome entry site IIA subdomain. *Journal of medicinal chemistry* **2005**, *48*, 7099–102.

(155) Parsons, J.; Castaldi, M. P.; Dutta, S.; Dibrov, S. M.; Wyles, D. L.; Hermann, T. Conformational inhibition of the hepatitis C virus internal ribosome entry site RNA. *Nature chemical biology* **2009**, *5*, 823–5.

(156) Carnevali, M.; Parsons, J.; Wyles, D. L.; Hermann, T. A modular approach to synthetic RNA binders of the hepatitis C virus internal ribosome entry site. *Chembiochem : a European journal of chemical biology* **2010**, *11*, 1364–7.

(157) Stage, T. K.; Hertel, K. J.; Uhlenbeck, O. C. Inhibition of the hammerhead ribozyme by neomycin. *RNA (New York, N.Y.)* **1995**, *1*, 95–101.

(158) Mikkelsen, N. E.; Brännvall, M.; Virtanen, A.; Kirsebom, L. A. Inhibition of RNase P RNA cleavage by aminoglycosides. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 6155–60.

(159) Waksman, S. A.; Lechevalier, H. A. Neomycin, a New Antibiotic Active against Streptomycin-Resistant Bacteria, including Tuberculosis Organisms. *Science (New York, N.Y.)* **1949**, *109*, 305–7.

(160) UMEZAWA, H. Kanamycin: its discovery. *Annals of the New York Academy of Sciences* **1958**, *76*, 20–6.

(161) Cragg, G. M.; Grothaus, P. G.; Newman, D. J. Impact of natural products on developing new anti-cancer agents. *Chemical reviews* **2009**, *109*, 3012–43.

(162) Koehn, F. E.; Carter, G. T. The evolving role of natural products in drug discovery. *Nature reviews. Drug discovery* **2005**, *4*, 206–20.

(163) Wilkinson, B.; Micklefield, J. Mining and engineering natural-product biosynthetic pathways. *Nature chemical biology* **2007**, *3*, 379–86.

(164) Baltz, R. H. Renaissance in antibacterial discovery from actinomycetes. *Current opinion in pharmacology* **2008**, *8*, 557–63.

(165) Fenical, W.; Jensen, P. R. Developing a new resource for drug discovery: marine actinomycete bacteria. *Nature chemical biology* **2006**, *2*, 666–73.

(166) Baltz, R. H. Antimicrobials from Actinomycetes: Back to the Future. *Microbe* **2007**, *2*, 125–131.

(167) Genilloud, O.; González, I.; Salazar, O.; Martín, J.; Tormo, J. R.; Vicente, F. Current approaches to exploit actinomycetes as a source of novel natural products. *Journal of industrial microbiology & biotechnology* **2011**, *38*, 375–89.

(168) Lane, D. J.; Pace, B.; Olsen, G. J.; Stahl, D. A.; Sogin, M. L.; Pace, N. R. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proceedings of the National Academy of Sciences of the United States of America* **1985**, *82*, 6955–9.

(169) Handelsman, J.; Rondon, M. R.; Brady, S. F.; Clardy, J.; Goodman, R. M. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry & biology* **1998**, *5*, R245–9.

(170) Curtis, T. P.; Sloan, W. T.; Scannell, J. W. Estimating prokaryotic diversity and its limits. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 10494–9.

(171) Schloss, P. D.; Handelsman, J. Status of the microbial census. *Microbiology and molecular biology reviews : MMBR* **2004**, *68*, 686–91.

(172) Pace, N. R. A Molecular View of Microbial Diversity and the Biosphere. *Science* **1997**, *276*, 734–740.

(173) Amann, R. I.; Ludwig, W.; Schleifer, K. H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiological reviews* **1995**, *59*, 143–69.

(174) Kaeberlein, T.; Lewis, K.; Epstein, S. S. Isolating “uncultivable” microorganisms in pure culture in a simulated natural environment. *Science (New York, N.Y.)* **2002**, *296*, 1127–9.

(175) Watve, M. G.; Tickoo, R.; Jog, M. M.; Bhole, B. D. How many antibiotics are produced by the genus Streptomyces? *Archives of microbiology* **2001**, *176*, 386–90.

(176) Bode, H. B.; Müller, R. The impact of bacterial genomics on natural product research. *Angewandte Chemie (International ed. in English)* **2005**, *44*, 6828–46.

(177) Bentley, S. D.; Chater, K. F.; Cerdeño-Tárraga, A.-M.; Challis, G. L.; Thomson, N. R.; James, K. D.; Harris, D. E.; Quail, M. A.; Kieser, H.; Harper, D.; Bateman, A.; Brown, S.; Chandra, G.; Chen, C. W.; Collins, M.; Cronin, A.; Fraser, A.; Goble, A.; Hidalgo, J.; Hornsby, T.; Howarth, S.; Huang, C.-H.; Kieser, T.; Larke, L.; Murphy, L.; Oliver, K.; O’Neil, S.; Rabbinowitsch, E.; Rajandream, M.-

A.; Rutherford, K.; Rutter, S.; Seeger, K.; Saunders, D.; Sharp, S.; Squares, R.; Squares, S.; Taylor, K.; Warren, T.; Wietzorrek, A.; Woodward, J.; Barrell, B. G.; Parkhill, J.; Hopwood, D. A. Complete genome sequence of the model actinomycete Streptomyces coelicolor A3(2). *Nature* **2002**, *417*, 141–7.

(178) Ikeda, H.; Ishikawa, J.; Hanamoto, A.; Shinose, M.; Kikuchi, H.; Shiba, T.; Sakaki, Y.; Hattori, M.; Omura, S. Complete genome sequence and comparative analysis of the industrial microorganism Streptomyces avermitilis. *Nature biotechnology* **2003**, *21*, 526–31.

(179) Challis, G. L. Mining microbial genomes for new natural products and biosynthetic pathways. *Microbiology (Reading, England)* **2008**, *154*, 1555–69.

(180) Oliynyk, M.; Samborskyy, M.; Lester, J. B.; Mironenko, T.; Scott, N.; Dickens, S.; Haydock, S. F.; Leadlay, P. F. Complete genome sequence of the erythromycin-producing bacterium Saccharopolyspora erythraea NRRL23338. *Nature Biotechnology* **2007**, *25*, 447–453.

(181) Fischbach, M. A.; Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. *Chemical reviews* **2006**, *106*, 3468–96.

(182) Medema, M. H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M. a; Weber, T.; Takano, E.; Breitling, R. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic acids research* **2011**, *39*, 339–46.

(183) Hertweck, C. Hidden biosynthetic treasures brought to light. *Nature chemical biology* **2009**, *5*, 450–2.

(184) Cane, D. E.; Ikeda, H. Exploration and mining of the bacterial terpenome. *Accounts of chemical research* **2012**, *45*, 463–72.

(185) Lautru, S.; Deeth, R. J.; Bailey, L. M.; Challis, G. L. Discovery of a new peptide natural product by Streptomyces coelicolor genome mining. *Nature chemical biology* **2005**, *1*, 265–9.

(186) Gross, H.; Stockwell, V. O.; Henkels, M. D.; Nowak-Thompson, B.; Loper, J. E.; Gerwick, W. H. The genomisotopic approach: a systematic method to isolate products of orphan biosynthetic gene clusters. *Chemistry & biology* **2007**, *14*, 53–63.

(187) Ishida, K.; Lincke, T.; Behnken, S.; Hertweck, C. Induced biosynthesis of cryptic polyketide metabolites in a Burkholderia thailandensis quorum sensing mutant. *Journal of the American Chemical Society* **2010**, *132*, 13966–8.

(188) Laureti, L.; Song, L.; Huang, S.; Corre, C.; Leblond, P.; Challis, G. L.; Aigle, B. Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in Streptomyces ambofaciens. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108*, 6258–63.

(189) Piel, J. Approaches to capturing and designing biologically active small molecules produced by uncultured microbes. *Annual review of microbiology* **2011**, *65*, 431–53.

(190) Rusch, D. B.; Halpern, A. L.; Sutton, G.; Heidelberg, K. B.; Williamson, S.; Yooseph, S.; Wu, D.; Eisen, J. A.; Hoffman, J. M.; Remington, K.; Beeson, K.; Tran, B.; Smith, H.; Baden-Tillson, H.; Stewart, C.; Thorpe, J.; Freeman, J.; Andrews-Pfannkoch, C.; Venter, J. E.; Li, K.; Kravitz, S.; Heidelberg, J. F.; Utterback, T.; Rogers, Y.-H.; Falcón, L. I.; Souza, V.; Bonilla-Rosso, G.; Eguiarte, L. E.; Karl, D. M.; Sathyendranath, S.; Platt, T.; Bermingham, E.; Gallardo, V.; Tamayo-Castillo, G.; Ferrari, M. R.; Strausberg, R. L.; Nealson, K.; Friedman, R.; Frazier, M.; Venter, J. C. The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS biology* **2007**, *5*, e77.

(191) Eisen, J. A. Environmental shotgun sequencing: its potential and challenges for studying the hidden world of microbes. *PLoS biology* **2007**, *5*, e82.

(192) Fisch, K. M.; Gurgui, C.; Heycke, N.; van der Sar, S. A.; Anderson, S. A.; Webb, V. L.; Taudien, S.; Platzer, M.; Rubio, B. K.; Robinson, S. J.; Crews, P.; Piel, J. Polyketide assembly lines of uncultivated sponge symbionts from structure-based gene targeting. *Nature chemical biology* **2009**, *5*, 494–501.

(193) Kirschning, A.; Taft, F.; Knobloch, T. Total synthesis approaches to natural product derivatives based on the combination of chemical synthesis and metabolic engineering. *Organic & biomolecular chemistry* **2007**, *5*, 3245–59.

(194) Planson, A.-G.; Carbonell, P.; Grigoras, I.; Faulon, J.-L. A retrosynthetic biology approach to therapeutics: from conception to delivery. *Current opinion in biotechnology* **2012**.

(195) Weissman, K. J.; Leadlay, P. F. Combinatorial biosynthesis of reduced polyketides. *Nature reviews. Microbiology* **2005**, *3*, 925–36.

(196) Wong, F. T.; Khosla, C. Combinatorial biosynthesis of polyketides--a perspective. *Current opinion in chemical biology* **2012**, *16*, 117–23.

(197) Llewellyn, N. M.; Spencer, J. B. Biological chemistry: enzymes line up for assembly. *Nature* **2007**, *448*, 755–6.

(198) Ro, D.-K.; Paradise, E. M.; Ouellet, M.; Fisher, K. J.; Newman, K. L.; Ndungu, J. M.; Ho, K. a; Eachus, R. a; Ham, T. S.; Kirby, J.; Chang, M. C. Y.; Withers, S. T.; Shiba, Y.; Sarpong, R.; Keasling, J. D. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **2006**, *440*, 940–3.

(199) Chang, M. C. Y.; Eachus, R. a; Trieu, W.; Ro, D.-K.; Keasling, J. D. Engineering Escherichia coli for production of functionalized terpenoids using plant P450s. *Nature chemical biology* **2007**, *3*, 274–7.

(200) Trefzer, A.; Blanco, G.; Remsing, L.; Künzel, E.; Rix, U.; Lipata, F.; Braña, A. F.; Méndez, C.; Rohr, J.; Bechthold, A.; Salas, J. a Rationally designed glycosylated premithramycins: hybrid aromatic polyketides using genes from three different biosynthetic pathways. *Journal of the American Chemical Society* **2002**, *124*, 6056–62.

(201) Salas, A. P.; Zhu, L.; Sánchez, C.; Braña, A. F.; Rohr, J.; Méndez, C.; Salas, J. A. Deciphering the late steps in the biosynthesis of the anti-tumour indolocarbazole staurosporine: sugar donor substrate flexibility of the StaG glycosyltransferase. *Molecular microbiology* **2005**, *58*, 17–27.

(202) Salas, J. A.; Méndez, C. Engineering the glycosylation of natural products in actinomycetes. *Trends in microbiology* **2007**, *15*, 219–32.

(203) Usui, T.; Umezawa, S. Total synthesis of neomycin B. *The Journal of antibiotics* **1987**, *40*, 1464–7.

(204) Greenberg, W. A.; Priestley, E. S.; Sears, P. S.; Alper, P. B.; Rosenbohm, C.; Hendrix, M.; Hung, S.-C.; Wong, C.-H. Design and Synthesis of New Aminoglycoside Antibiotics Containing Neamine as an Optimal Core Structure: Correlation of Antibiotic Activity with in Vitro Inhibition of Translation. *Journal of the American Chemical Society* **1999**, *121*, 6527–6541.

(205) Umezawa, S.; Tsuchiya, T.; Jikihara, T.; Umezawa, H. Synthesis of 3’.4'-dideoxyneamine active aginst kanamycin-resistant E. coli and P. aeruginosa. *The Journal of antibiotics* **1971**, *24*, 711–2.

(206) Hanessian, S.; Butterworth, R. F.; Nakagawa, T. Aminoglycoside antibiotics: chemical transformation of paromamine into 3’-epiparomamine. *Carbohydrate research* **1973**, *26*, 261–3.

(207) Hanessian, S.; Massé, R.; Capmeau, M. L. Aminoglycoside antibiotics: synthesis of 5’'-amino-5''-deoxyneomycin and 5''-amino-5''-deoxyparomomycin. *The Journal of antibiotics* **1977**, *30*, 893–6.

(208) Hanessian, S.; Vatele, J. M. Aminoglycoside antibiotics 4'-deoxyneomycin and 4'-deoxyparomamine. *The Journal of antibiotics* **1980**, *33*, 675–8.

(209) Pathak, R.; Böttger, E. C.; Vasella, A. Design and Synthesis of Aminoglycoside Antibiotics to Selectively Target 16S Ribosomal RNA Position 1408. *Helvetica Chimica Acta* **2005**, *88*, 2967–2985.

(210) Li, J.; Wang, J.; Czyryca, P. G.; Chang, H.; Orsak, T. W.; Evanson, R.; Chang, C.-W. T. Application of glycodiversification: expedient synthesis and antibacterial evaluation of a library of kanamycin B analogues. *Organic letters* **2004**, *6*, 1381–4.

(211) Baker, T. J.; Luedtke, N. W.; Tor, Y.; Goodman, M. Synthesis and Anti-HIV Activity of Guanidinoglycosides. *The Journal of Organic Chemistry* **2000**, *65*, 9054–9058.

(212) Luedtke, N. W.; Carmichael, P.; Tor, Y. Cellular uptake of aminoglycosides, guanidinoglycosides, and poly-arginine. *Journal of the American Chemical Society* **2003**, *125*, 12374–5.

(213) Kondo, S.; Iinuma, K.; Yamamoto, H.; Maeda, K.; Umezawa, H. Letter: Syntheses of 1-n-(S)-4-amino-2-hydroxybutyryl)-kanamycin B and -3', 4'-dideoxykanamycin B active against kanamycin-resistant bacteria. *The Journal of antibiotics* **1973**, *26*, 412–5.

(214) Kondo, S.; Iinuma, K.; Yamamoto, H.; Ikeda, Y.; Maeda, K. Letter: Synthesis of (S)-4-amino-2-hydroxybutyryl derivatives of 3',4'-dideoxykanamycin B and their antibacterial activities. *The Journal of antibiotics* **1973**, *26*, 705–7.

(215) Haddad, J.; Kotra, L. P.; Llano-Sotelo, B.; Kim, C.; Azucena, E. F.; Liu, M.; Vakulenko, S. B.; Chow, C. S.; Mobashery, S. Design of novel antibiotics that bind to the ribosomal acyltransfer site. *Journal of the American Chemical Society* **2002**, *124*, 3229–37.

(216) Haddad, J.; Vakulenko, S.; Mobashery, S. An Antibiotic Cloaked by Its Own Resistance Enzyme. *Journal of the American Chemical Society* **1999**, *121*, 11922–11923.

(217) François, B.; Szychowski, J.; Adhikari, S. S.; Pachamuthu, K.; Swayze, E. E.; Griffey, R. H.; Migawa, M. T.; Westhof, E.; Hanessian, S. Antibacterial aminoglycosides with a modified mode of binding to the ribosomal-RNA decoding site. *Angewandte Chemie (International ed. in English)* **2004**, *43*, 6735–8.

(218) Hanessian, S.; Tremblay, M.; Swayze, E. E. Tobramycin analogues with C-5 aminoalkyl ether chains intended to mimic rings III and IV of paromomycin. *Tetrahedron* **2003**, *59*, 983–993.

(219) Stathakis, C. I.; Mavridis, I.; Kythreoti, G.; Papakyriakou, A.; Katsoulis, I. A.; Cottin, T.; Anastasopoulou, P.; Vourloumis, D. Second generation analogs of rigid 6,7-spiro scaffolds targeting the bacterial ribosome. *Bioorganic & medicinal chemistry letters* **2010**, *20*, 7488–92.

(220) Katsoulis, I. A.; Kythreoti, G.; Papakyriakou, A.; Koltsida, K.; Anastasopoulou, P.; Stathakis, C. I.; Mavridis, I.; Cottin, T.; Saridakis, E.; Vourloumis, D. Synthesis of 5,6-spiroethers and evaluation of their affinities for the bacterial A site. *Chembiochem : a European journal of chemical biology* **2011**, *12*, 1188–92.

(221) Fridman, M.; Belakhov, V.; Yaron, S.; Baasov, T. A new class of branched aminoglycosides: pseudo-pentasaccharide derivatives of neomycin B. *Organic letters* **2003**, *5*, 3575–8.

(222) Revuelta, J.; Vacas, T.; Corzana, F.; Gonzalez, C.; Bastida, A.; Asensio, J. L. Structure-based design of highly crowded ribostamycin/kanamycin hybrids as a new family of antibiotics. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2010**, *16*, 2986–91.

(223) Zhao, F.; Zhao, Q.; Blount, K. F.; Han, Q.; Tor, Y.; Hermann, T. Molecular recognition of RNA by neomycin and a restricted neomycin derivative. *Angewandte Chemie (International ed. in English)* **2005**, *44*, 5329–34.

(224) Bastida, A.; Hidalgo, A.; Chiara, J. L.; Torrado, M.; Corzana, F.; Pérez-Cañadillas, J. M.; Groves, P.; Garcia-Junceda, E.; Gonzalez, C.; Jimenez-Barbero, J.; Asensio, J. L. Exploring the use of conformationally locked aminoglycosides as a new strategy to overcome bacterial resistance. *Journal of the American Chemical Society* **2006**, *128*, 100–16.

(225) Kirk, S. R.; Luedtke, N. W.; Tor, Y. Neomycin−Acridine Conjugate: A Potent Inhibitor of Rev-RRE Binding. *Journal of the American Chemical Society* **2000**, *122*, 980–981.

(226) Luedtke, N. W.; Liu, Q.; Tor, Y. RNA-ligand interactions: affinity and specificity of aminoglycoside dimers and acridine conjugates to the HIV-1 Rev response element. *Biochemistry* **2003**, *42*, 11391–403.

(227) Kaiser, M.; Sainlos, M.; Lehn, J.-M.; Bombard, S.; Teulade-Fichou, M.-P. Aminoglycoside-quinacridine conjugates: towards recognition of the P6.1 element of telomerase RNA. *Chembiochem : a European journal of chemical biology* **2006**, *7*, 321–9.

(228) Pokrovskaya, V.; Belakhov, V.; Hainrichson, M.; Yaron, S.; Baasov, T. Design, synthesis, and evaluation of novel fluoroquinolone-aminoglycoside hybrid antibiotics. *Journal of medicinal chemistry* **2009**, *52*, 2243–54.

(229) Lee, J.; Kwon, M.; Lee, K. H.; Jeong, S.; Hyun, S.; Shin, K. J.; Yu, J. An approach to enhance specificity against RNA targets using heteroconjugates of aminoglycosides and chloramphenicol (or linezolid). *Journal of the American Chemical Society* **2004**, *126*, 1956–7.

(230) Riguet, E.; Tripathi, S.; Chaubey, B.; Désiré, J.; Pandey, V. N.; Décout, J.-L. A peptide nucleic acid-neamine conjugate that targets and cleaves HIV-1 TAR RNA inhibits viral replication. *Journal of medicinal chemistry* **2004**, *47*, 4806–9.

(231) Hyun, S.; Lee, K. H.; Yu, J. A strategy for the design of selective RNA binding agents. Preparation and RRE RNA binding affinities of a neomycin-peptide nucleic acid heteroconjugate library. *Bioorganic & medicinal chemistry letters* **2006**, *16*, 4757–9.

(232) Erlanson, D. A.; McDowell, R. S.; O'Brien, T. Fragment-based drug discovery. *Journal of medicinal chemistry* **2004**, *47*, 3463–82.

(233) Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. Fragment-based lead discovery. *Nature reviews. Drug discovery* **2004**, *3*, 660–72.

(234) Sucheck, S. J.; Wong, A. L.; Koeller, K. M.; Boehr, D. D.; Draker, K.; Sears, P.; Wright, G. D.; Wong, C.-H. Design of Bifunctional Antibiotics that Target Bacterial rRNA and Inhibit Resistance-Causing Enzymes. *Journal of the American Chemical Society* **2000**, *122*, 5230–5231.

(235) Agnelli, F.; Sucheck, S. J.; Marby, K. A.; Rabuka, D.; Yao, S.-L.; Sears, P. S.; Liang, F.-S.; Wong, C.-H. Dimeric aminoglycosides as antibiotics. *Angewandte Chemie (International ed. in English)* **2004**, *43*, 1562–6.

(236) Riguet, E.; Désiré, J.; Boden, O.; Ludwig, V.; Göbel, M.; Bailly, C.; Décout, J.-L. Neamine dimers targeting the HIV-1 TAR RNA. *Bioorganic & medicinal chemistry letters* **2005**, *15*, 4651–5.

(237) Tok, J. B.; Dunn, L. J.; Des Jean, R. C. Binding of dimeric aminoglycosides to the HIV-1 rev responsive element (RRE) RNA construct. *Bioorganic & medicinal chemistry letters* **2001**, *11*, 1127–31.

(238) Wang, H.; Tor, Y. Dimeric aminoglycosides: Design, synthesis and RNA binding. *Bioorganic & Medicinal Chemistry Letters* **1997**, *7*, 1951–1956.

(239) Kumar, S.; Arya, D. P. Recognition of HIV TAR RNA by triazole linked neomycin dimers. *Bioorganic & medicinal chemistry letters* **2011**, *21*, 4788–92.

(240) Liu, X.; Thomas, J. R.; Hergenrother, P. J. Deoxystreptamine dimers bind to RNA hairpin loops. *Journal of the American Chemical Society* **2004**, *126*, 9196–7.

(241) Yoshizawa, S.; Fourmy, D.; Eason, R. G.; Puglisi, J. D. Sequence-specific recognition of the major groove of RNA by deoxystreptamine. *Biochemistry* **2002**, *41*, 6263–70.

(242) Thomas, J. R.; Liu, X.; Hergenrother, P. J. Size-specific ligands for RNA hairpin loops. *Journal of the American Chemical Society* **2005**, *127*, 12434–5.

(243) Zhou, Y.; Gregor, V. E.; Sun, Z.; Ayida, B. K.; Winters, G. C.; Murphy, D.; Simonsen, K. B.; Vourloumis, D.; Fish, S.; Froelich, J. M.; Wall, D.; Hermann, T. Structure-guided discovery of novel aminoglycoside mimetics as antibacterial translation inhibitors. *Antimicrobial agents and chemotherapy* **2005**, *49*, 4942–9.

(244) Jöge, T.; Jesberger, M.; Bröker, P.; Kirschning, A. Synthetic access to spacer-linked 3,6-diamino-2,3,6-trideoxy-alpha-D-glucopyranosides--potential aminoglycoside mimics for the inhibition of the HIV-1 TAR-RNA/Tat-peptide complex. *Carbohydrate research* **2007**, *342*, 1704–14.

(245) Westermann, B.; Dorner, S. Synthesis of multivalent aminoglycoside mimics via the Ugi multicomponent reaction. *Chemical communications (Cambridge, England)* **2005**, 2116–8.

(246) Green, K. D.; Chen, W.; Houghton, J. L.; Fridman, M.; Garneau-Tsodikova, S. Exploring the substrate promiscuity of drug-modifying enzymes for the chemoenzymatic generation of N-acylated aminoglycosides. *Chembiochem : a European journal of chemical biology* **2010**, *11*, 119–26.

(247) Llewellyn, N. M.; Spencer, J. B. Chemoenzymatic acylation of aminoglycoside antibiotics. *Chemical communications (Cambridge, England)* **2008**, 3786–8.

(248) Kirschning, A.; Hahn, F. Merging Chemical Synthesis and Biosynthesis: A New Chapter in the Total Synthesis of Natural Products and Natural Product Libraries. *Angewandte Chemie (International ed. in English)* **2012**, *3*, 4012–4022.

(249) Weist, S.; Süssmuth, R. D. Mutational biosynthesis--a tool for the generation of structural diversity in the biosynthesis of antibiotics. *Applied microbiology and biotechnology* **2005**, *68*, 141–50.

(250) Rinehart, K. L. Mutasynthesis of new antibiotics. *Pure and Appl. Chem.* **1977**, *49*, 1361–1384.

(251) Shier, W. T.; Rinehart, K. L.; Gottlieb, D. Preparation of four new antibiotics from a mutant of Streptomyces fradiae. *Proceedings of the National Academy of Sciences of the United States of America* **1969**, *63*, 198–204.

(252) Keasling, J. D. Synthetic biology for synthetic chemistry. *ACS chemical biology* **2008**, *3*, 64–76.

(253) Kieser, T.; Bibb, M. J. .; Buttner, M. J. .; Chater, K. F. .; Hopwood, D. A. . Practical Streptomyces Genetics. *The John Innes Foundation* **2000**.

(254) Raynal, A.; Karray, F.; Tuphile, K.; Darbon-Rongère, E.; Pernodet, J.-L. Excisable cassettes: new tools for functional analysis of Streptomyces genomes. *Applied and environmental microbiology* **2006**, *72*, 4839–44.

(255) Fischbach, M. A.; Lai, J. R.; Roche, E. D.; Walsh, C. T.; Liu, D. R. Directed evolution can rapidly improve the activity of chimeric assembly-line enzymes. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 11951–6.

(256) Reetz, M. T.; Carballeira, J. D. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nature protocols* **2007**, *2*, 891–903.

(257) Park, S.-H.; Park, H.-Y.; Sohng, J. K.; Lee, H. C.; Liou, K.; Yoon, Y. J.; Kim, B.-G. Expanding substrate specificity of GT-B fold glycosyltransferase via domain swapping and high-throughput screening. *Biotechnology and bioengineering* **2009**, *102*, 988–94.

(258) Zhang, H.; Boghigian, B. A.; Armando, J.; Pfeifer, B. A. Methods and options for the heterologous production of complex natural products. *Natural product reports* **2011**, *28*, 125–51.

(259) Dutton, C. J.; Gibson, S. P.; Goudie, A. C.; Holdom, K. S.; Pacey, M. S.; Ruddock, J. C.; Bu'Lock, J. D.; Richards, M. K. Novel avermectins produced by mutational biosynthesis. *The Journal of antibiotics* **1991**, *44*, 357–65.

(260) Lowden, P. a S.; Böhm, G. a; Metcalfe, S.; Staunton, J.; Leadlay, P. F. New rapamycin derivatives by precursor-directed biosynthesis. *Chembiochem : a European journal of chemical biology* **2004**, *5*, 535–8.

(261) Gregory, M. A.; Petkovic, H.; Lill, R. E.; Moss, S. J.; Wilkinson, B.; Gaisser, S.; Leadlay, P. F.; Sheridan, R. M. Mutasynthesis of rapamycin analogues through the manipulation of a gene governing starter unit biosynthesis. *Angewandte Chemie (International ed. in English)* **2005**, *44*, 4757–60.

(262) Weissman, K. J. Mutasynthesis - uniting chemistry and genetics for drug discovery. *Trends in biotechnology* **2007**, *25*, 139–42.

(263) Eichner, S.; Floss, H. G.; Sasse, F.; Kirschning, A. New, highly active nonbenzoquinone geldanamycin derivatives by using mutasynthesis. *Chembiochem : a European journal of chemical biology* **2009**, *10*, 1801–5.

(264) Knobloch, T.; Harmrolfs, K.; Taft, F.; Thomaszewski, B.; Sasse, F.; Kirschning, A. Mutational biosynthesis of ansamitocin antibiotics: a diversity-oriented approach to exploit biosynthetic flexibility. *Chembiochem : a European journal of chemical biology* **2011**, *12*, 540–7.

(265) Ankenbauer, R. G.; Staley, A. L.; Rinehart, K. L.; Cox, C. D. Mutasynthesis of siderophore analogues by Pseudomonas aeruginosa. *Proceedings of the National Academy of Sciences of the United States of America* **1991**, *88*, 1878–82.

(266) Hojati, Z.; Milne, C.; Harvey, B.; Gordon, L.; Borg, M.; Flett, F.; Wilkinson, B.; Sidebottom, P. J.; Rudd, B. A. M.; Hayes, M. A.; Smith, C. P.; Micklefield, J. Structure, biosynthetic origin, and engineered biosynthesis of calcium-dependent antibiotics from Streptomyces coelicolor. *Chemistry & biology* **2002**, *9*, 1175–87.

(267) Weist, S.; Bister, B.; Puk, O.; Bischoff, D.; Pelzer, S.; Nicholson, G. J.; Wohlleben, W.; Jung, G.; Süssmuth, R. D. Fluorobalhimycin--a new chapter in glycopeptide antibiotic research. *Angewandte Chemie (International ed. in English)* **2002**, *41*, 3383–5.

(268) Weist, S.; Kittel, C.; Bischoff, D.; Bister, B.; Pfeifer, V.; Nicholson, G. J.; Wohlleben, W.; Süssmuth, R. D. Mutasynthesis of glycopeptide antibiotics: variations of vancomycin's AB-ring amino acid 3,5-dihydroxyphenylglycine. *Journal of the American Chemical Society* **2004**, *126*, 5942–3.

(269) Kalaitzis, J. A.; Izumikawa, M.; Xiang, L.; Hertweck, C.; Moore, B. S. Mutasynthesis of enterocin and wailupemycin analogues. *Journal of the American Chemical Society* **2003**, *125*, 9290–1.

(270) Shen, B. Accessing natural products by combinatorial biosynthesis. *Science's STKE : signal transduction knowledge environment* **2004**, *2004*, pe14.

(271) Llewellyn, N. M.; Spencer, J. B. Biosynthesis of 2-deoxystreptamine-containing aminoglycoside antibiotics. *Natural product reports* **2006**, *23*, 864–74.

(272) Kudo, F.; Yamamoto, Y.; Yokoyama, K.; Eguchi, T.; Kakinuma, K. Biosynthesis of 2-deoxystreptamine by three crucial enzymes in Streptomyces fradiae NBRC 12773. *The Journal of antibiotics* **2005**, *58*, 766–74.

(273) Huang, F.; Haydock, S. F.; Mironenko, T.; Spiteller, D.; Li, Y.; Spencer, J. B. The neomycin biosynthetic gene cluster of Streptomyces fradiae NCIMB 8233: characterisation of an aminotransferase involved in the formation of 2-deoxystreptamine. *Organic & biomolecular chemistry* **2005**, *3*, 1410–8.

(274) Kudo, F.; Kawabe, K.; Kuriki, H.; Eguchi, T.; Kakinuma, K. A new family of glucose-1-phosphate/glucosamine-1-phosphate nucleotidylyltransferase in the biosynthetic pathways for antibiotics. *Journal of the American Chemical Society* **2005**, *127*, 1711–8.

(275) Fan, Q.; Huang, F.; Leadlay, P. F.; Spencer, J. B. The neomycin biosynthetic gene cluster of Streptomyces fradiae NCIMB 8233: genetic and biochemical evidence for the roles of two glycosyltransferases and a deacetylase. *Organic & biomolecular chemistry* **2008**, *6*, 3306–14.

(276) Kudo, F.; Eguchi, T. Biosynthetic enzymes for the aminoglycosides butirosin and neomycin. *Methods in enzymology* **2009**, *459*, 493–519.

(277) Wehmeier, U. F.; Piepersberg, W. Enzymology of aminoglycoside biosynthesis-deduction from gene clusters. *Methods in enzymology* **2009**, *459*, 459–91.

(278) Truman, A. W.; Huang, F.; Llewellyn, N. M.; Spencer, J. B. Characterization of the enzyme BtrD from Bacillus circulans and revision of its functional assignment in the biosynthesis of butirosin. *Angewandte Chemie (International ed. in English)* **2007**, *46*, 1462–4.

(279) Huang, F.; Spiteller, D.; Koorbanally, N. A.; Li, Y.; Llewellyn, N. M.; Spencer, J. B. Elaboration of neosamine rings in the biosynthesis of neomycin and butirosin. *Chembiochem : a European journal of chemical biology* **2007**, *8*, 283–8.

(280) Yokoyama, K.; Yamamoto, Y.; Kudo, F.; Eguchi, T. Involvement of two distinct N-acetylglucosaminyltransferases and a dual-function deacetylase in neomycin biosynthesis. *Chembiochem : a European journal of chemical biology* **2008**, *9*, 865–9.

(281) Kurumbang, N. P.; Liou, K.; Sohng, J. K. Biosynthesis of paromamine derivatives in engineered Escherichia coli by heterologous expression. *Journal of applied microbiology* **2010**, *108*, 1780–8.

(282) Stano, N. M.; Patel, S. S. T7 lysozyme represses T7 RNA polymerase transcription by destabilizing the open complex during initiation. *The Journal of biological chemistry* **2004**, *279*, 16136–43.

(283) Miroux, B.; Walker, J. E. Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *Journal of molecular biology* **1996**, *260*, 289–98.

(284) Sambrook, J.; Russell, D. W. Molecular Cloning: A laboratory manual. *CSHL Press* **2001**.

(285) Betancor, L.; Fernández, M.-J.; Weissman, K. J.; Leadlay, P. F. Improved catalytic activity of a purified multienzyme from a modular polyketide synthase after coexpression with

Streptomyces chaperonins in Escherichia coli. *Chembiochem : a European journal of chemical biology* **2008**, *9*, 2962–6.

(286) Krupa, J. C.; Shaya, D.; Chi, L.; Linhardt, R. J.; Cygler, M.; Withers, S. G.; Mort, J. S. Quantitative continuous assay for hyaluronan synthase. *Analytical biochemistry* **2007**, *361*, 218–25.

(287) An, S.; Kumar, R.; Sheets, E. D.; Benkovic, S. J. Reversible compartmentalization of de novo purine biosynthetic complexes in living cells. *Science (New York, N.Y.)* **2008**, *320*, 103–6.

(288) Weissman, K. J.; Müller, R. Protein-protein interactions in multienzyme megasynthetases. *Chembiochem : a European journal of chemical biology* **2008**, *9*, 826–48.

(289) Park, J. W.; Hong, J. S. J.; Parajuli, N.; Jung, W. S.; Park, S. R.; Lim, S.-K.; Sohng, J. K.; Yoon, Y. J. Genetic dissection of the biosynthetic route to gentamicin A2 by heterologous expression of its minimal gene set. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 8399–404.

(290) Kurumbang, N. P.; Liou, K.; Sohng, J. K. Biosynthesis of ribostamycin derivatives by reconstitution and heterologous expression of required gene sets. *Applied biochemistry and biotechnology* **2011**, *163*, 373–82.

(291) Kurumbang, N. P.; Park, J. W.; Yoon, Y. J.; Liou, K.; Sohng, J. K. Heterologous production of ribostamycin derivatives in engineered Escherichia coli. *Research in microbiology* **2010**, *161*, 526–33.

(292) Welch, M.; Govindarajan, S.; Ness, J. E.; Villalobos, A.; Gurney, A.; Minshull, J.; Gustafsson, C. Design parameters to control synthetic gene expression in Escherichia coli. *PloS one* **2009**, *4*, e7002.

(293) Davis, G. D.; Elisee, C.; Newham, D. M.; Harrison, R. G. New fusion protein systems designed to give soluble expression in Escherichia coli. *Biotechnology and bioengineering* **1999**, *65*, 382–8.

(294) Arnvig, K. B.; Pennell, S.; Gopal, B.; Colston, M. J. A high-affinity interaction between NusA and the rrn nut site in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 8325–30.

(295) Niiranen, L.; Espelid, S.; Karlsen, C. R.; Mustonen, M.; Paulsen, S. M.; Heikinheimo, P.; Willassen, N. P. Comparative expression study to increase the solubility of cold adapted Vibrio proteins in Escherichia coli. *Protein expression and purification* **2007**, *52*, 210–8.

(296) Maurice, F.; Broutin, I.; Podglajen, I.; Benas, P.; Collatz, E.; Dardel, F. Enzyme structural plasticity and the emergence of broad-spectrum antibiotic resistance. *EMBO reports* **2008**, *9*, 344–9.

(297) Smialowski, P.; Martin-Galiano, A. J.; Mikolajka, A.; Girschick, T.; Holak, T. A.; Frishman, D. Protein solubility: sequence based prediction and experimental verification. *Bioinformatics (Oxford, England)* **2007**, *23*, 2536–42.

(298) Horcajada, C.; Guinovart, J. J.; Fita, I.; Ferrer, J. C. Crystal structure of an archaeal glycogen synthase: insights into oligomerization and substrate binding of eukaryotic glycogen synthases. *The Journal of biological chemistry* **2006**, *281*, 2923–31.

(299) Vetting, M. W.; Frantom, P. A.; Blanchard, J. S. Structural and enzymatic analysis of MshA from Corynebacterium glutamicum: substrate-assisted catalysis. *The Journal of biological chemistry* **2008**, *283*, 15834–44.

(300) Rinehart, K. L.; Stroshane, R. M. Biosynthesis of aminocyclitol antibiotics. *The Journal of antibiotics* **1976**, *29*, 319–53.

(301) Reinhart, K. L.; Malik, J. M.; Nystrom, R. S.; Stroshane, R. M.; Truitt, S. T.; Taniguchi, M.; Rolls, J. P.; Haak, W. J.; Ruff, B. A. Biosynthetic incorporation of (1-13C)glucosamine and (6-13C)glucose into neomycin. *Journal of the American Chemical Society* **1974**, *96*, 2263–5.

(302) Shier, W. T.; Rinehart, K. L. Preparation of two new aminoglycoside antibiotics. *The Journal of antibiotics* **1970**, *23*, 51–3.

(303) Shier, W. T.; Ogawa, S.; Hichens, M.; Rinehart, K. L. Chemistry and biochemistry of the neomycins. XVII. Bioconversion of aminocyclitols to aminocyclitol antibiotics. *The Journal of antibiotics* **1973**, *26*, 551–61.

(304) Cleophax, J.; Gero, S. D.; Leboul, J.; Akhtar, M.; Barnett, J. E.; Pearce, C. J. A chiral synthesis of D-(+)-2,6-dideoxystreptamine and its microbial incorporation into novel antibiotics. *Journal of the American Chemical Society* **1976**, *98*, 7110–2.

(305) Kojima, M.; Sato, A. Letter: Microbial semi-synthesis of aminoglycosidic antibiotics by mutants of S. ribosidificus and S. kanamyceticus. *The Journal of antibiotics* **1973**, *26*, 784–6.

(306) Shier, W. T.; Schaefer, P. C.; Gottlieb, D.; Rinehart, K. L. Use of mutants in the study of aminocyclitol antibiotic biosynthesis and the preparation of the hybrimycin C complex. *Biochemistry* **1974**, *13*, 5073–8.

(307) Taylor, H. D.; Schmitz, H. Antibiotics derived from a mutant of Bacillus circulans. *The Journal of antibiotics* **1976**, *29*, 532–5.

(308) Siebenberg, S.; Bapat, P. M.; Lantz, A. E.; Gust, B.; Heide, L. Reducing the variability of antibiotic production in Streptomyces by cultivation in 24-square deepwell plates. *Journal of bioscience and bioengineering* **2010**, *109*, 230–4.

(309) Daum, S. J.; Rosi, D.; Goss, W. Production of antibiotics by biotransformation of 2,4,6/3,5-pentahydroxycyclohexanone and 2,4/3,5-tetrahydroxycyclohexanone by a deoxystreptamine-negative mutant of Micromonospora purpurea. *Journal of the American Chemical Society* **1977**, *99*, 283–4.

(310) Rosi, D.; Goss, W. A.; Daum, S. J. Mutational biosynthesis by idiotrophs of Micromonospora purpurea. I. Conversion of aminocyclitols to new aminoglycoside antibiotics. *The Journal of antibiotics* **1977**, *30*, 88–97.

(311) Park, J. W.; Park, S. R.; Nepal, K. K.; Han, A. R.; Ban, Y. H.; Yoo, Y. J.; Kim, E. J.; Kim, E. M.; Kim, D.; Sohng, J. K.; Yoon, Y. J. Discovery of parallel pathways of kanamycin biosynthesis allows antibiotic manipulation. *Nature chemical biology* **2011**, *7*, 843–52.

(312) Juguet, M.; Lautru, S.; Francou, F.-X.; Nezbedová, S.; Leblond, P.; Gondry, M.; Pernodet, J.-L. An iterative nonribosomal peptide synthetase assembles the pyrrole-amide antibiotic congocidine in Streptomyces ambofaciens. *Chemistry & biology* **2009**, *16*, 421–31.

(313) Nguyen, H. C.; Karray, F.; Lautru, S.; Gagnat, J.; Lebrihi, A.; Huynh, T. D. H.; Pernodet, J.-L. Glycosylation steps during spiramycin biosynthesis in Streptomyces ambofaciens: involvement of three glycosyltransferases and their interplay with two auxiliary proteins. *Antimicrobial agents and chemotherapy* **2010**, *54*, 2830–9.

(314) Pérez Luna, A.; Ceschi, M.-A.; Bonin, M.; Micouin, L.; Husson, H.-P.; Gougeon, S.; Estenne-Bouhtou, G.; Marabout, B.; Sevrin, M.; George, P. Enantioselective desymmetrization of meso bicyclic hydrazines: a novel approach to the asymmetric synthesis of polysubstituted amino cyclopentanic cores. *The Journal of organic chemistry* **2002**, *67*, 3522–4.

(315) Bournaud, C.; Chung, F.; Pérez Luna, A.; Pasco, M.; Errasti, G.; Lecourt, T.; Micouin, L. Stereoselective Transformations of meso Bicyclic Hydrazines: Versatile Access to Functionalized Aminocyclopentanes. *Synthesis* **2009**, 869–887.

(316) Pasco, M.; Moumné, R.; Lecourt, T.; Micouin, L. Stereoselective synthesis of fluorinated 1,3-cis-diaminocyclopentanes. *The Journal of organic chemistry* **2011**, *76*, 5137–42.

(317) Bierman, M.; Logan, R.; O'Brien, K.; Seno, E. T.; Rao, R. N.; Schoner, B. E. Plasmid cloning vectors for the conjugal transfer of DNA from Escherichia coli to Streptomyces spp. *Gene* **1992**, *116*, 43–9.

(318) Raynal, A.; Karray, F.; Tuphile, K.; Darbon-Rongère, E.; Pernodet, J.-L. Excisable Cassettes: New Tools for Functional Analysis of Streptomyces Genomes. *Applied and Environmental Microbiology* **2006**, *72*, 4839–4844.

(319) Fedeniuk, R. W.; Shand, P. J. Theory and methodology of antibiotic extraction from biomatrices. *Journal of chromatography. A* **1998**, *812*, 3–15.

(320) Berwick, P. G. Beta-lactam and aminoglycoside production from streptomycetes. *The Journal of applied bacteriology* **1988**, *64*, 9–15.

(321) Eichner, S.; Eichner, T.; Floss, H. G.; Fohrer, J.; Hofer, E.; Sasse, F.; Zeilinger, C.; Kirschning, A. Broad substrate specificity of the amide synthase in S. hygroscopicus--new 20-membered macrolactones derived from geldanamycin. *Journal of the American Chemical Society* **2012**, *134*, 1673–9.

(322) Park, W. K. C.; Auer, M.; Jaksche, H.; Wong, C.-H. Rapid Combinatorial Synthesis of Aminoglycoside Antibiotic Mimetics: Use of a Polyethylene Glycol-Linked Amine and a Neamine-Derived Aldehyde in Multiple Component Condensation as a Strategy for the Discovery of New Inhibitors of the HIV RNA Rev Responsiv. *Journal of the American Chemical Society* **1996**, *118*, 10150–10155.

(323) Peck, R. L.; Graber, R. P.; Walti, A.; Peel, E. W.; Hoffhine, J. C. E.; Folkers, K. Streptomyces antibiotics; hydrolytic cleavage of streptomycin to streptidine. *Science* **1946**, 29.

(324) Peck, R. L.; Hoffhine, J. C. E.; Peel, E. W.; Graber, R. P.; Holly, F. W.; Mozingo, R.; Folkers, K. . *Science* **1946**, 776.

(325) Chung, F.; Tisné, C.; Lecourt, T.; Seijo, B.; Dardel, F.; Micouin, L. Design of tRNA(Lys)3 ligands: fragment evolution and linker selection guided by NMR spectroscopy. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2009**, *15*, 7109–16.

(326) Stout, D. M.; Takaya, T.; Meyers, A. I. Unequivocal synthesis of N-substituted 1,4-dihydropyridines. *The Journal of Organic Chemistry* **1975**, *40*, 563–569.

(327) Mellor, J. M.; Smith, N. M. Reductive cleavage of the nitrogen–nitrogen bond in hydrazine derivatives. *J. Chem. Soc., Perkin Trans. 1* **1984**, 2927–2931.

(328) Grabowski, S.; Armbruster, J.; Prinzbach, H. Biocatalysis in the chiral recognition of meso-diamides — an efficient route from cyclic olefinic hydrocarbons to optically pure diamino-polyols. *Tetrahedron Letters* **1997**, *38*, 5485–5488.

(329) Jewett, J. C.; Bertozzi, C. R. Cu-free click cycloaddition reactions in chemical biology. *Chemical Society reviews* **2010**, *39*, 1272–9.

(330) Dumont, A.; Malleron, A.; Awwad, M.; Dukan, S.; Vauzeilles, B. Click-mediated labeling of bacterial membranes through metabolic modification of the lipopolysaccharide inner core. *Angewandte Chemie (International ed. in English)* **2012**, *51*, 3143–6.

(331) van Wezel, G. P.; McDowall, K. J. The regulation of the secondary metabolism of Streptomyces: new links and experimental advances. *Natural product reports* **2011**, *28*, 1311–33.

(332) Stead, D. A.; Richards, R. M. Sensitive fluorimetric determination of gentamicin sulfate in biological matrices using solid-phase extraction, pre-column derivatization with 9-fluorenylmethyl chloroformate and reversed-phase high-performance liquid chromatography. *Journal of chromatography. B, Biomedical applications* **1996**, *675*, 295–302.

(333) Cortina, N. S.; Krug, D.; Plaza, A.; Revermann, O.; Müller, R. Myxoprincomide: a natural product from Myxococcus xanthus discovered by comprehensive analysis of the secondary metabolome. *Angewandte Chemie (International ed. in English)* **2012**, *51*, 811–6.

(334) Pellecchia, M.; Bertini, I.; Cowburn, D.; Dalvit, C.; Giralt, E.; Jahnke, W.; James, T. L.; Homans, S. W.; Kessler, H.; Luchinat, C.; Meyer, B.; Oschkinat, H.; Peng, J.; Schwalbe, H.; Siegal, G. Perspectives on NMR in drug discovery: a technique comes of age. *Nature reviews. Drug discovery* **2008**, *7*, 738–45.

(335) Allain, F. H.; Varani, G. How accurately and precisely can RNA structure be determined by NMR? *Journal of molecular biology* **1997**, *267*, 338–51.

(336) Latham, M. P.; Brown, D. J.; McCallum, S. A.; Pardi, A. NMR methods for studying the structure and dynamics of RNA. *Chembiochem : a European journal of chemical biology* **2005**, *6*, 1492–505.

(337) Dalvit, C. Ligand- and substrate-based 19F NMR screening: Principles and applications to drug discovery. *Progress in Nuclear Magnetic Resonance Spectroscopy* **2007**, *51*, 243–271.

(338) Cobb, S. L.; Murphy, C. D. 19F NMR applications in chemical biology. *Journal of Fluorine Chemistry* **2009**, *130*, 132–143.

(339) Danielson, M. A.; Falke, J. J. Use of 19F NMR to Probe Protein Structure and Conformational Changes. *Annual Review of Biophysics and Biomolecular Structure* **1996**, *Vol. 25*, 163–195.

(340) Dalvit, C.; Gossert, A. D.; Coutant, J.; Piotto, M. Rapid acquisition of 1H and 19F NMR experiments for direct and competition ligand-based screening. *Magnetic resonance in chemistry : MRC* **2011**, *49*, 199–202.

(341) Dalvit, C.; Fagerness, P. E.; Hadden, D. T. A.; Sarver, R. W.; Stockman, B. J. Fluorine-NMR experiments for high-throughput screening: theoretical aspects, practical considerations, and range of applicability. *Journal of the American Chemical Society* **2003**, *125*, 7696–703.

(342) Dalvit, C.; Mongelli, N.; Papeo, G.; Giordano, P.; Veronesi, M.; Moskau, D.; Kümmerle, R. Sensitivity improvement in 19F NMR-based screening experiments: theoretical considerations and experimental applications. *Journal of the American Chemical Society* **2005**, *127*, 13380–5.

(343) Vulpetti, A.; Hommel, U.; Landrum, G.; Lewis, R.; Dalvit, C. Design and NMR-based screening of LEF, a library of chemical fragments with different local environment of fluorine. *Journal of the American Chemical Society* **2009**, *131*, 12949–59.

(344) Hennig, M.; Scott, L. G.; Sperling, E.; Bermel, W.; Williamson, J. R. Synthesis of 5-fluoropyrimidine nucleotides as sensitive NMR probes of RNA structure. *Journal of the American Chemical Society* **2007**, *129*, 14911–21.

(345) Kreutz, C.; Kählig, H.; Konrat, R.; Micura, R. Ribose 2'-F labeling: a simple tool for the characterization of RNA secondary structure equilibria by 19F NMR spectroscopy. *Journal of the American Chemical Society* **2005**, *127*, 11558–9.

(346) Puffer, B.; Kreutz, C.; Rieder, U.; Ebert, M.-O.; Konrat, R.; Micura, R. 5-Fluoro pyrimidines: labels to probe DNA and RNA secondary structures by 1D 19F NMR spectroscopy. *Nucleic acids research* **2009**, *37*, 7728–40.

(347) Hammann, C.; Norman, D. G.; Lilley, D. M. Dissection of the ion-induced folding of the hammerhead ribozyme using 19F NMR. *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98*, 5503–8.

(348) Arnold, J. R.; Fisher, J. Structural equilibria in RNA as revealed by 19F NMR. *Journal of biomolecular structure & dynamics* **2000**, *17*, 843–56.

(349) Trempe, J. F.; Wilds, C. J.; Denisov, A. Y.; Pon, R. T.; Damha, M. J.; Gehring, K. NMR solution structure of an oligonucleotide hairpin with a 2'F-ANA/RNA stem: implications for RNase H

specificity toward DNA/RNA hybrid duplexes. *Journal of the American Chemical Society* **2001**, *123*, 4896–903.

(350) Graber, D.; Moroder, H.; Micura, R. 19F NMR spectroscopy for the analysis of RNA secondary structure populations. *Journal of the American Chemical Society* **2008**, *130*, 17230–1.

(351) Kreutz, C.; Kählig, H.; Konrat, R.; Micura, R. A general approach for the identification of site-specific RNA binders by 19F NMR spectroscopy: proof of concept. *Angewandte Chemie (International ed. in English)* **2006**, *45*, 3450–3.

(352) Tisné, C.; Guillière, F.; Dardel, F. NMR-based identification of peptides that specifically recognize the d-arm of tRNA. *Biochimie* **2005**, *87*, 885–8.

(353) Tisné, C.; Rigourd, M.; Marquet, R.; Ehresmann, C.; Dardel, F. NMR and biochemical characterization of recombinant human tRNA(Lys)3 expressed in Escherichia coli: identification of posttranscriptional nucleotide modifications required for efficient initiation of HIV-1 reverse transcription. *RNA (New York, N.Y.)* **2000**, *6*, 1403–12.

(354) Tisné, C.; Roques, B. P.; Dardel, F. Heteronuclear NMR studies of the interaction of tRNA(Lys)3 with HIV-1 nucleocapsid protein. *Journal of molecular biology* **2001**, *306*, 443–54.

(355) Moumné, R.; Pasco, M.; Prost, E.; Lecourt, T.; Micouin, L.; Tisné, C. Fluorinated diaminocyclopentanes as chiral sensitive NMR probes of RNA structure. *Journal of the American Chemical Society* **2010**, *132*, 13111–3.

(356) Lombès, T.; Moumné, R.; Larue, V.; Prost, E.; Catala, M.; Lecourt, T.; Dardel, F.; Micouin, L.; Tisné, C. Investigation of RNA-Ligand Interactions by (19) F NMR Spectroscopy Using Fluorinated Probes. *Angewandte Chemie (International ed. in English)* **2012**.

(357) Weigand, J. E.; Sanchez, M.; Gunnesch, E.-B.; Zeiher, S.; Schroeder, R.; Suess, B. Screening for engineered neomycin riboswitches that control translation initiation. *RNA (New York, N.Y.)* **2008**, *14*, 89–97.

(358) Duchardt-Ferner, E.; Weigand, J. E.; Ohlenschläger, O.; Schmidtke, S. R.; Suess, B.; Wöhnert, J. Highly modular structure and ligand binding by conformational capture in a minimalistic riboswitch. *Angewandte Chemie (International ed. in English)* **2010**, *49*, 6216–9.

(359) Dardel, F. MC-Fit: using Monte-Carlo methods to get accurate confidence limits on enzyme parameters. *Computer applications in the biosciences : CABIOS* **1994**, *10*, 273–5.

(360) Wang, Z. X. An exact mathematical expression for describing competitive binding of two different ligands to a protein molecule. *FEBS letters* **1995**, *360*, 111–4.

(361) Zhang, X.; Sänger, A.; Hemmig, R.; Jahnke, W. Ranking of high-affinity ligands by NMR spectroscopy. *Angewandte Chemie (International ed. in English)* **2009**, *48*, 6691–4.

(362) Thordarson, P. Determining association constants from titration experiments in supramolecular chemistry. *Chemical Society reviews* **2011**, *40*, 1305–23.

(363) Umezawa, S.; Koto, S.; Tatsuta, K.; Tsumura, T. Studies of aminosugars. XXI. The total synthesis of kanamycin C. *Bulletin of the Chemical Society of Japan* **1969**, *42*, 529–33.

(364) Trost, B. M.; Caldwell, C. G.; Murayama, E.; Heissler, D. Sulfur-substituted dienes and the silylene protecting group in synthesis. Deoxypillaromycinone. *The Journal of Organic Chemistry* **1983**, *48*, 3252–3265.

(365) Carreira, E. M.; Du Bois, J. (+)-Zaragozic Acid C: Synthesis and Related Studies. *Journal of the American Chemical Society* **1995**, *117*, 8106–8125.

(366) Singh, R. P.; Shreeve, J. M. Recent Advances in Nucleophilic Fluorination Reactions of Organic Compounds Using Deoxofluor and DAST. *Synthesis* **2002**, 2561–2578.

(367) Moumné, R.; Larue, V.; Seijo, B.; Lecourt, T.; Micouin, L.; Tisné, C. Tether influence on the binding properties of tRNALys3 ligands designed by a fragment-based approach. *Organic & biomolecular chemistry* **2010**, *8*, 1154–9.

(368) Cartier, G.; Lorieux, F.; Allemand, F.; Dreyfus, M.; Bizebard, T. Cold adaptation in DEAD-box proteins. *Biochemistry* **2010**, *49*, 2636–46.

(369) Suess, B.; Entian, K.-D.; Kötter, P.; Weigand, J. E. Aptamer-regulated expression of essential genes in yeast. *Methods in molecular biology (Clifton, N.J.)* **2012**, *824*, 381–91.

(370) Simon, R.; Priefer, U.; Pühler, A. A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Bio/Technology* **1983**, *1*, 784–791.

(371) McFarlane, G. J. B.; Machray, G. C.; Stewart, W. D. P. A simplfied method for conjugal gene transfer into the filamentous cyanobacterium Anabaena sp. ATCC 27893. *Journal of Microbiological Methods* **1987**, *6*, 301–305.

(372) Paget, M. S.; Chamberlin, L.; Atrih, A.; Foster, S. J.; Buttner, M. J. Evidence that the extracytoplasmic function sigma factor sigmaE is required for normal cell wall structure in Streptomyces coelicolor A3(2). *Journal of bacteriology* **1999**, *181*, 204–11.

(373) Pernodet, J. L.; Alegre, M. T.; Blondelet-Rouault, M. H.; Guérineau, M. Resistance to spiramycin in Streptomyces ambofaciens, the producer organism, involves at least two different mechanisms. *Journal of general microbiology* **1993**, *139*, 1003–11.

(374) Okanishi, M.; Suzuki, K.; Umezawa, H. Formation and reversion of Streptomycete protoplasts: cultural condition and morphological study. *Journal of general microbiology* **1974**, *80*, 389–400.

(375) Piotto, M.; Saudek, V.; Sklenár, V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *Journal of biomolecular NMR* **1992**, *2*, 661–5.

(376) Bonnet, J.; Subsoontorn, P.; Endy, D. Rewritable digital data storage in live cells via engineered control of recombination directionality. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 8884–9.

(377) Dubey, G. P.; Ben-Yehuda, S. Intercellular nanotubes mediate bacterial communication. *Cell* **2011**, *144*, 590–600.

(378) Toor, N.; Keating, K. S.; Taylor, S. D.; Pyle, A. M. Crystal structure of a self-spliced group II intron. *Science (New York, N.Y.)* **2008**, *320*, 77–82.

(379) Lautru, S.; Song, L.; Demange, L.; Lombès, T.; Galons, H.; Challis, G. L.; Pernodet, J.-L. A Sweet Origin for the Key Congocidine Precursor 4-Acetamidopyrrole-2-carboxylate. *Angewandte Chemie (International ed. in English)* **2012**.

# VII. Annexes

# I - Résumé de la thèse

Les ARN constituent des cibles thérapeutiques extrêmement intéressantes bien qu'encore assez peu exploitées. En effet, les obstacles pour la conception de ligands spécifiques de ces cibles non traditionnelles, polyanioniques et très flexibles, sont encore loin d'être levés. Les aminoglycosides, utilisés depuis longtemps pour leurs propriétés antibiotiques, sont souvent décrits comme des « ligands universels » d'ARN. Leur structure constitue donc une architecture favorable pour l'élaboration de nouveaux ligands spécifiques des ARN.

Le but de cette thèse a été de développer une méthode systémique originale combinant chimie organique et microbiologie pour synthétiser de nouvelles molécules de structure analogue aux aminoglycosides, se fixant de façon spécifique sur des cibles ARN. Ce travail repose sur la compréhension récente des voies de biosynthèse des aminoglycosides permettant leur ingénierie rationnelle selon une stratégie de mutasynthèse. Cette approche expérimentale s'appuie sur la conception de mimes de métabolites naturels pouvant être transformés par des bactéries génétiquement modifiées. Le développement de méthodologies novatrices en ingénierie métabolique, synthèse organique et chimie analytique nous a permis de concevoir des analogues d'aminoglycosides fluorés qui se sont avérées être d'excellentes sondes dans l'étude des ARN par RMN du fluor.

**Mots-clés :** Aminoglycosides, ARN, ligands, biosynthèse, mutasynthèse, RMN $^{19}F$

# II - Production scientifique

## *A - Communications orales*

- **Séminaire de l'Institut de Génétique et Microbiologie** (IGM) – Université Paris Sud – Orsay/France *(Fev. **2010**)*
  « Chemical and metabolic engineering approaches to the synthesis of novel aminoglycosides »
- **Séminaire de l'Institut Médicament Toxicolologie Chimie Environnement** (IMTCE) – Université Paris Descartes – Paris/France *(Nov. **2010**)*
  Présentation du projet IGEM de l'équipe de Paris 2010 « Every Bacteria counts ! »
- **Réunion RMN Ile de France** – Orsay/France (*Fev. **2011***)
  « $^{19}$F NMR as a tool for RNA analysis »
- **Séminaire de l'Institut de Génétique et Microbiologie** (IGM) – Université Paris Sud – Orsay/France *(Avril **2011**)*
  Présentation du projet IGEM de l'équipe de Paris 2010 « Every Bacteria counts ! »
- **Journée d'étude « Biologie Synthétique » de la société de biologie** – Paris/France (*Sept. **2011***)
  « Introduction à la compétition IGEM et présentation des travaux de l'équipe de Paris »
- **Congrès annuel de la Société Française de Biochimie et de Biologie Moléculaire** (SFBBM) - Ax les Thermes/France *(Oct. **2011**)*
  « Metabolic engineering and chemical strategies for the synthesis of novel aminoglycosides analogues »
- **Séminaire de l'Institut Médicament Toxicolologie Chimie Environnement** (IMTCE) – Université Paris Descartes – Paris/France *(Mars **2012**)*
  Présentation du projet IGEM de l'équipe de Paris 2011 « Tube or not Tube ? Towards harnessing bacterial nanotubes by and for synthetic biology »
- **Journée de chimie organique et chimie biologique de la montagne Sainte-Geneviève** – Paris/France *(Mai **2012**)*
  « Investigation of ligand-RNA interaction by $^{19}$F NMR using fluorinated probes »

## *B - Posters*

- **Paris Interdisciplinary PhD Symposium "Numbers in living systems"** - Paris/France *(Dec. **2009**)*
- **Paris Interdisciplinary PhD Symposium "Shapes in living systems" –** Paris/France (*Jun*. ***2011***)
- **European Symposium on Organic Chemistry** (ESOC) – Crete/Greece *(Jul. **2011**)*
  « Stereoselective synthesis of fluorinated diaminocyclopentanes as chiral sensitive NMR probes of RNA structure »
- **European Science Foundation (ESF) Conference " Synthetic Biology of Antibiotic Production" –** Sant Feliu de Guixols/Spain *(Oct. **2011**)*
  « Metabolic engineering and chemical strategies for the synthesis of novel aminoglycosides analogues »

## *C - Organisation de congrès*

- **Rencontres de Chimie Organique (RCO) -** Paris/France *(avril **2009**)*
  Membre du comité d'organisation
- **Paris Interdisciplinary PhD Symposium "Numbers in living systems"** - Paris/France *(dec. **2009**)*
  Co-fondateur et membre du comité d'organisation
- **Paris Interdisciplinary PhD Symposium "Shapes in living systems" –** Paris/France (*Juin **2011***)
  Co-fondateur et membre du comité d'organisation

## *D - Enseignements*

- **Chargé de TD** « Biologie Synthétique » (32h, 2009-2010)
  niveau M1, Master AIV, Université Paris Descartes
- **Encadrement du stage d'Aleksandra Nivina**
  (stage volontaire de 4$^{ème}$ année de pharmacie, 2 mois)
- **Encadrement de Paola Frigols-Arroyo** (stage ERASMUS, 6 mois)
- **Encadrant de l'équipe IGEM Paris 2010** (8 mois)
- **Encadrant de l'équipe IGEM Paris 2011** (8 mois)

## *E - Publications*

### 1) Liste des publications

1. “NMR-guided fragment-based approach for the design of AAC(6')-Ib ligands.”
   **Lombès T**, Bégis G, Maurice F, Turcaud S, Lecourt T, Dardel F, Micouin L.
   *Chembiochem*, **2008**, Jun 16;9(9):1368-71
   DOI: 10.1002/cbic.200700677

2. “2,3-DIAZABICYCLO[2.2.1]HEPT-5-ENE-2,3-DICARBOXYLIC ACID, 2,3-DIETHYL ESTER”
   **Lombès T**, Prasad B, Gilbertson S, Micouin L.
   *e-EROS,* **2012,** Mar 15
   DOI: 10.1002/047084289X.rn01437

3. “A Sweet Origin for the Key Congocidine Precursor 4-Acetamidopyrrole-2-carboxylate”
   Lautru S, Song L, Demange L, **Lombès T**, Galons H, Challis GL, Pernodet JL.
   *Angew. Chem. Int. Ed*., **2012**, Jul 23;51(30):7454-8
   DOI: 10.1002/anie.201201445

4. “Investigation of RNA-ligand interaction by $^{19}$F NMR spectroscopy using fluorinated probes.”
   **Lombès T**, Moumné R, Larue V, Prost E, Catala M, Lecourt T, Dardel F, Micouin L, Tisné C.
   *Angew. Chem. Int. Ed*., **2012**, Aug 21
   [accepté, publié sur internet, DOI: 10.1002/anie.201204083]

## 2) Articles en pdf

### a) Publication 1

CHEMBIOCHEM

DOI: 10.1002/cbic.200700677

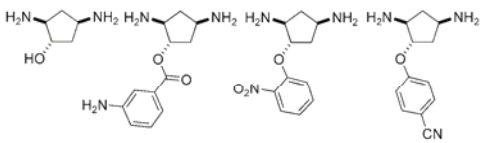
# NMR-Guided Fragment-Based Approach for the Design of AAC(6′)-Ib Ligands

Thomas Lombès,[a] Guillaume Bégis,[a] Frédérique Maurice,[b] Serge Turcaud,[a] Thomas Lecourt,[a] Frédéric Dardel,*[b] and Laurent Micouin*[a]

Aminoglycosides are potent, broad-spectrum antibiotics of major clinical importance.[1] Unfortunately, the emergence of resistance is increasingly restricting their use as antibacterials.[2] Among the common mechanisms of resistance, enzymatic modification of aminoglycosides by N-acetylation at the 6′-position is one of the most widespread. Two functional classes of N-acetyltransferases (AAC(6′)) have been described: AAC(6′)-I, which confers resistance to amikacin but not to gentamicin, and AAC(6′)-II, with the reciprocal selectivity. Among these two classes, the aminoglycoside 6′-*N*-acetyltransferase type Ii (AAC(6′)-Ii) has been intensively investigated, with the crystallographic determination of its structure and the design of several bisubstrate inhibitors.[3] Moreover, the first inhibitor of this enzyme active in cells has recently been reported, providing the proof of concept for the pharmaceutical potential of such compounds.[4] The structure of AAC(6′)-Iy has also been reported.[5]

As part of a general program on the fragment-based design of aminoglycoside mimics,[6] we have been interested in the study of another AAC(6′)-I subclass: AAC(6′)-Ib. Unlike in the cases of AAC(6′)-Ii and Iy, which are chromosomally encoded (and hence confined to a single bacterial species), the gene for AAC(6′)-Ib is carried by mobile elements such as the integrons, and is present in over 70% of AAC(6′)-producing Gram-negative clinical isolates.[7] Furthermore, some isoforms of this class have recently evolved in clinical isolates to provide resistance to both amikacin and gentamicin[8] or to some fluoroquinolines.[9] The low sequence identity (<20%) of ACC(6′)-Ib with the well-characterized ACC(6′)-Ii and ACC(6′)-Iy prompted us to evaluate a fragment-based approach for the design of ligands of this enzyme, rather than a rational design based on the structural data available for ACC(6′)-Ii and ACC(6′)-Iy enzymes. Our aim was to design ligands that were structurally distinct from aminoglycosides, which would be less likely to be deactivated by other resistance enzymes.[10] Here we report an NMR-guided design of micromolar ligands for this enzyme, as well as their conversion into a bisubstrate inhibitor.

In preliminary studies, we used NMR detection techniques to identify compound **1** (Scheme 1) as a primary fragment for the design of ligands.[11] This compound is readily available from cyclopentadiene on a large scale in a racemic or enantioenriched form.[12] Both saturation transfer difference (STD)[13] and reverse n.O.e. pumping experiments[14] were used to verify its ability to mimic the central desoxystreptamine core of kanamycin binding to the enzyme (Figure 1). Furthermore, the specificity of this binding was confirmed by a substantial decrease in the STD signal of compound **1** when the experiment was conducted in the presence of kanamycin, showing competitive binding of this fragment to the active site of the enzyme (for details see the Supporting Information).




**Scheme 1.** Dissociation constants of selected ligands determined by fluorescence spectroscopy experiments.

Compound **1** was then coupled through ether or ester linkages with several substituted aromatic moieties capable of mimicking ring 1 of aminosides, leading to racemic neamine analogues (Scheme 2).

All the compounds were then evaluated by reverse n.O.e. pumping experiments for their ability to bind AAC(6′)-Ib.

It was expected that the common aminocyclopentanic moiety would bind in most cases, leading to intermolecular n.O.e. transfers, whereas efficient n.O.e. transfers from the aromatic appendage would only occur with the best ligands (Figure 2). This NMR screening provided a qualitative estimation of binding affinity for all the compounds, which can be classified into three groups: no signals from the aromatic components were observed with compounds **5** and **13**, whereas weak n.O.e transfer onto the aromatic components of compounds **14**, **15**, **8**, **18**, **12** and **11** could be detected, and strong n.O.e. transfers were observed with compounds **7**, **4**, **6**, **16** and **9**. Although this "ligand-observe" set of experiments provided a qualitative affinity score for the enzyme, we could not at this stage exclude the possibility that the binding might be occur-

[a] T. Lombès, Dr. G. Bégis, S. Turcaud, Dr. T. Lecourt, Dr. L. Micouin
Chimie Thérapeutique, UMR 8638
Université Paris Descartes, CNRS
4 avenue de l'Observatoire, 75006 Paris (France)
Fax: (+33) 1-43-29-14-03
E-mail: laurent.micouin@univ-paris5.fr

[b] Dr. F. Maurice, Dr. F. Dardel
Cristallographie et RMN Biologiques
Université Paris Descartes, CNRS
4 avenue de l'Observatoire, 75006 Paris (France)
Fax: (+33) 1-53-73-99-25
E-mail: frederic.dardel@univ-paris5.fr

Supporting information for this article is available on the WWW under http://www.chembiochem.org or from the author.


© 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

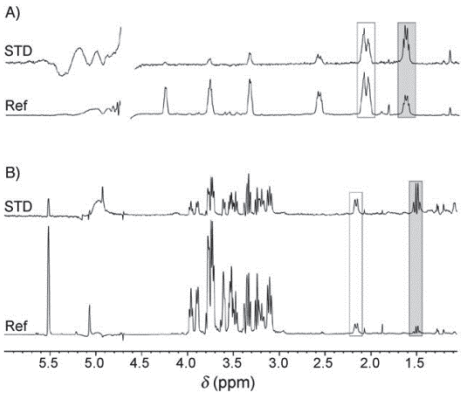



**Figure 1.** STD analysis of the binding of kanamycin (bottom) and diaminocyclopentanol **1** (top; the water signal has been removed for clarity) to AAC(6′)-1b. The boxed peaks correspond to the cyclic methylenic protons located between the two amino groups. In each case, the shaded box corresponds to the proton *syn* to the amines and shows the strongest STD signal.

ring out of the enzyme active site. Since the size of AAC(6′)-Ib (43 kD, 2×196 amino acids) precluded the direct investigation of foot-printing by "target-observe" 2D NMR techniques (such as HSQC or TROSY experiments), the binding of representative ligands was then investigated by fluorescence spectroscopy. We indeed observed that binding of compounds **1**, **15**, **4** and **9** induced strong quenching of the fluorescence of tryptophan residues (50–70%). As AAC(6′)-I enzymes share conserved tryptophan residues within their active sites, on which the 2-DOS moiety has been observed to stack, this fluorescence quench might be indirect evidence that these compounds are indeed interacting with tryptophans located in the AAC(6′)-Ib active site. Furthermore, it allowed the calculation of their dissociation constants by titration. Thus, $K_d$ values were established to be 41± 3 μM for compound **1**, 57±6 μM for compound **15**, 23±2 μM for compound **4**, and 20±1.2 μM for compound **9** (Scheme 1).

This correlates very well with the NMR results, as the additional interactions observed with **4** and **9** result in a twofold improvement of affinity, whereas the small loss observed with **15** is correlated with very weak intermolecular n.O.e. transfer.

The final validation of this NMR-guided ligand optimization was performed by preparing a bisubstrate ligand from compound **9**. For this study, we decided to resynthesise this compound in enantioenriched form, since the coupling of a racemic material with the CoA moiety would deliver diastereomers (Scheme 3). Compound **22** was then selectively reduced, deprotected and coupled with CoA by the one-pot procedure recently reported by Auclair and co-workers (Scheme 3).

The dissociation constant of compound **24** was then evaluated by fluorescence titration as 950±190 nM, showing an improved affinity relative to the starting optimized fragment **9**. Standard AAC(6′)-Ib inhibition essays with kanamycin as a substrate showed that compound **24** is an inhibitor, with a $K_i$ value of about 500±300 nM, whereas neither CoA nor compound **9** showed any significant inhibitory properties. No synergistic activity could be observed when **24** was tested in combination with kanamycin against a *E. coli* strain harbouring the AAC(6′)-Ib encoding gene, as would be expected with a compound bearing an unmodified CoA moiety.[15]

In conclusion, we have shown in this work that "ligand-observe" NMR techniques such as reverse n.O.e. pumping can be used to drive the detection and optimisation of ligands for AAC(6′)-Ib, one of the most clinically important resistance enzymes to aminoglycosides. Interestingly, this fragment-based approach has been conducted without any need for NMR

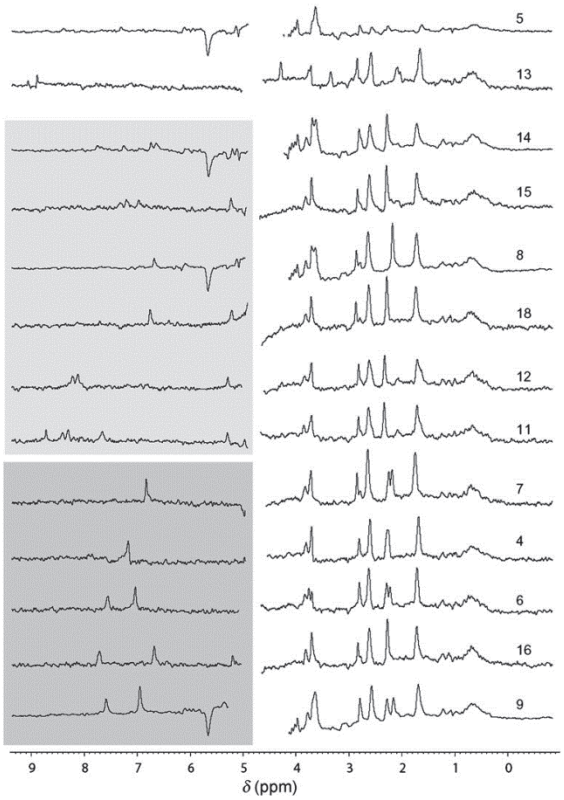
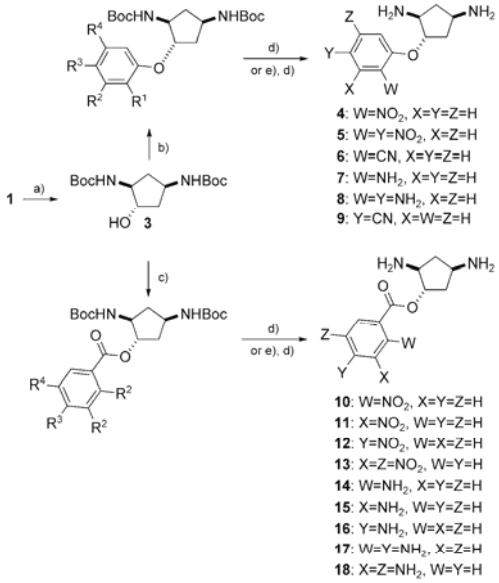
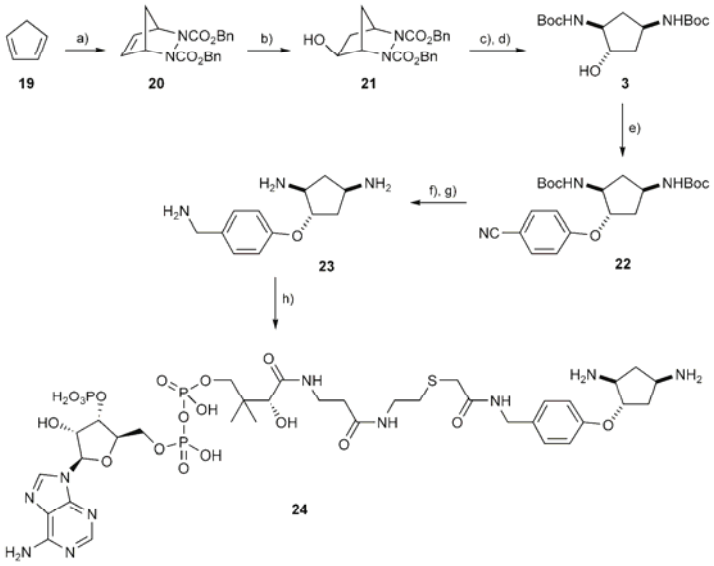



**Figure 2.** Qualitative binding analysis by reverse-n.O.e pumping experiments of compounds **4**–**18**. Spectra are ordered by increasing affinities. The water signal (4.7 ppm) has been removed for clarity.

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 




**Scheme 2.** a) $Boc_2O$, NaOH, dioxane/$H_2O$, 84%; b) KHMDS, THF, ArF, 66–88%; c) DCC, cat. DMAP, $CH_2Cl_2$, ArCOOH, 47–66%; d) HCl (gas), AcOEt, quantitative; e) $H_2$, 10% Pd/C, AcOEt or THF, 43–93%.




**Scheme 3.** a) DBAD, $CH_2Cl_2$, quantitative; b) CatBH, 1% $[Rh(COD)Cl]_2$, 2% (*R,R*)-BDPP, DME, −50 °C then $H_2O_2$, 77%, *ee* = 83%; c) $H_2$, $PtO_2$, AcOH, quantitive; d) $Boc_2O$, NaOH 88% from **21**; e) *p*-fluorocyanobenzene, KHMDS, 4 °C to RT, 68%. f) $H_2$, 10% Pd/C, MeOH, 18 h; g) $HCl_g$, AcOEt, 10 min., 73% from **22**; h) Auclair's procedure (ref. [4]), preparative HPLC purification, 40%.

structural assignment of the target, with the validation of the binding site for the ligands by fluorescence spectroscopy and the preparation of a bisubstrate inhibitor. Despite the modest inhibition constant of compound **24** in relation to bisubstrate inhibitors described for AAC(6′)-Ii, this result shows that ligands and inhibitors of AAC(6′)-Ib can be constructed from non-aminoglycoside-like fragments, and that NMR can be a powerful tool for qualitatively estimation of their binding affinities to the enzyme. The design of more drug-like, nonbisubstrate inhibitors, as well as cocrystallisation studies, are underway.

## Acknowledgements

*F.M. was supported by a studentship from the Ministère de la Recherche. Dr. A. Lombès is acknowledged for help in UV experiments. Dr. A. Regazzetti is acknowledged for MS experiments. This work was supported by ANRS, CNRS and the Ministère de la Recherche.*

**Keywords:** aminoglycosides · antibiotics · coenzyme A · fragment-based synthesis · NMR spectroscopy

[1] S. B. Vakulenko, S. Mobashery, *Clin. Microbiol. Rev.* **2003**, *16*, 430–450.
[2] S. Magnet, J. S. Blanchard, *Chem. Rev.* **2005**, *105*, 477–498.
[3] a) L. E. Wybenga-Groot, K.-A. Draker, G. D. Wright, A. M. Berghuis, *Structure* **1999**, *7*, 497–507; b) D. L. Burk, N. Ghuman, L. E. Wybenga-Groot, A. M. Berghuis, *Protein Sci.* **2003**, *12*, 426–437; c) K.-A. Draker, D. B. Northrop, G. D. Wright, *Biochemistry* **2003**, *42*, 6565–6574; d) K.-A. Draker, G. D. Wright, *Biochemistry* **2004**, *43*, 446–454; e) J. W. Williams, D. B. Northrop, *J. Antibiot.* **1979**, *32*, 1147–1154; f) F. Gao, X. Yan, O. M. Baettig, A. M. Berghuis, K. Auclair, *Angew. Chem.* **2005**, *117*, 7019–7022; *Angew. Chem. Int. Ed.* **2005**, *44*, 6859–6862.
[4] F. Gao, X. Yan, T. Shakya, O. M. Baettig, S. Ait-Mohand-Brunet, A. M. Berghuis, G. D. Wright, K. Auclair, *J. Med. Chem.* **2006**, *49*, 5273–5281.
[5] M. W. Vetting, S. Magnet, E. Nieves, S. L. Roderick, J. S. Blanchard, *Chem. Biol.* **2004**, *11*, 565–573.
[6] F. Chung, C. Tisné, T. Lecourt, F. Dardel, L. Micouin, *Angew. Chem.* **2007**, *119*, 4573–4575; *Angew. Chem. Int. Ed.* **2007**, *46*, 4489–4491.
[7] A. C. Fluit, F. J. Schmitz, *Clin. Microbiol. Infect.* **2004**, *10*, 272–288.
[8] I. Casin, B. Hanau-Bercot, I. Podglajen, H. Vahaboglu, E. Collatz, *Antimicrob. Agents Chemother.* **2003**, *47*, 697–703.
[9] A. Robicsek, J. Strahilevitz, G. A. Jacoby, M. Macielag, D. Abbanat, C. H. Park, K. Bush, D. C. Hooper, *Nat. Med.* **2006**, *12*, 83–88.
[10] The potential to escape bacterial resistance by enzymatic inactivation through the design of conformationally restricted aminoglycosides has recently been proposed

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 

in a series of very elegant works: a) J. L. Asensio, A. Hidalgo, A. Bastida, M. Torrado, F. Corzana, J. L. Chiara, E. García-Junceda, J. Canada, J. Jimènez-Barbero, *J. Am. Chem. Soc.* **2005**, *127*, 8278–8279; b) A. Bastida, A. Hidalgo, J. L. Chiara, M. Torrado, F. Corzana, J. M. Pérez-Canadillas, P. Groves, E. García-Junceda, C. Gonzalez, J. Jimènez-Barbero, J. L. Asensio, *J. Am. Chem. Soc.* **2006**, *128*, 100–116; c) F. Corzana, I. Cuesta, A. Bastida, A. Hidalgo, M. Latorre, C. González, A. García-Junceda, J. Jiménez-Barbero, J. L. Asensio, *Chem. Eur. J.* **2005**, *11*, 5102–5113.

[11] F. Maurice, G. Begis, L. Micouin, F. Dardel, *C. R. Chim.* **2006**, *9*, 413–419.

[12] a) A. Pérez Luna, M.-A. Ceschi, M. Bonin, L. Micouin, H.-P. Husson, S. Gougeon, G. Estenne Bouthou, B. Marabout, M. Sevrin, P. George, *J. Org. Chem.* **2002**, *67*, 3522–3524; b) A. Pérez Luna, M. Bonin, L. Micouin, H.-P. Husson, *J. Am. Chem. Soc.* **2002**, *124*, 12098–12099.

[13] M. Mayer, B. Meyer, *J. Am. Chem. Soc.* **2001**, *123*, 6108–6117.

[14] A. Chen, M. J. Shapiro, *J. Am. Chem. Soc.* **2000**, *122*, 414–415.

[15] Compounds bearing a CoA moiety are probably unable to penetrate the bacterial envelope, as observed in ref. [3e] and discussed in ref. [2].

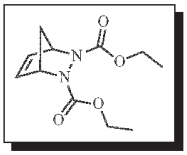
Received: November 9, 2007
Published online on May 7, 2008

 © 2008 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 

## b) Publication 2



# 2,3-Diazabicyclo[2.2.1]hept-5-ene-2,3-dicarboxylic Acid, 2,3-Diethyl Ester[1]



*[14011-60-0]* $C_{11}H_{16}N_2O_4$ (MW 240.26)
InChI = 1S/C11H16N2O4/c1-3-16-10(14)12-8-5-6-9(7-8)13(12)11(15)17-4-2/h5-6,8-9H,3-4,7H2,1-2H3
InChIKey = XMAKWQGDHULQOS-UHFFFAOYSA-N

(building block for numerous stereoselective transformations)

*Alternate Names:* 2,3-Diazabicyclo[2.2.1]hept-5-ene-2,3-dicarboxylic acid, diethyl ester(6CI,7CI,8CI,9CI); diethyl 2,3-diazabicyclo[2.2.1]hept-5-ene-2,3-dicarboxylate;Diethyl 1,2,3,6-tetrahydro-3,6-methano-1,2-pyridazinedicarboxylate; NSC 44163.

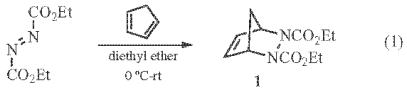
*Physical Data:* mp 38–39 °C; bp 131 °C (1 mm); d 1.272 ± 0.06 g $cm^{-3}$.

*Solubility:* alcohol, ether, and most organic solvents.

*Form Supplied in:* clear fluid oil, solidifies on long standing, rare chemical.

*Analysis of Reagent Purity:* $^1$H NMR.

*Preparative Method:* To a freshly distilled cyclopentadiene (1.5–2 equiv) solution in diethyl ether at 0 °C was added the diethyl azodicarboxylate (1 equiv).[2] The reaction was allowed to stir for 6–12 h at room temperature until full consumption of the diethyl azodicarboxylate. The solvent was evaporated under reduced pressure and the crude clear oil was purified by silica-gel flash column chromatography with EtOAc/hexanes as eluent (20–40% gradient) to afford the 2,3-diazabicyclo[2.2.1]hept-5-ene-2,3-dicarboxylic acid, 2,3-diethyl ester as a clear viscous liquid in 95% yield (eq 1).



(1)

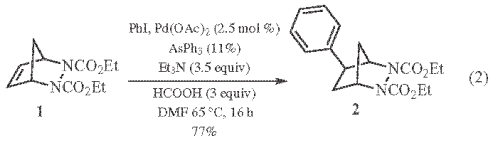
The corresponding dibenzyl ester or ditertbutyl ester can be prepared in almost quantitative yield in a similar manner.[3]

*Handling, Storage, and Precautions:* Very stable. Can be stored in the refrigerator for several months. Stable at room temperature, turns into pale yellow viscous oil on long standing.

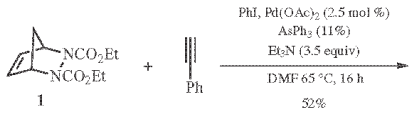
---

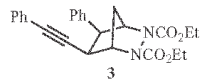
**Electrophilic Additions to C=C Double Bond Without Ring Cleavage.**

***Hydroarylation.*** The reaction of bicyclic hydrazine (**1**) with aryl or styryl halides in the presence of an *in situ*-generated palladium catalyst, stabilized by triphenylarsine, affords the *exo*-configured hydroarylation product (**2**) (eq 2).[4] Using electron-deficient aromatic groups tends to favor fragmentation of the bicycle. Replacing the acid by phenylacetylene in the hydroarylation reaction leads to a sequential arylation–alkynylation reaction (eq 3).



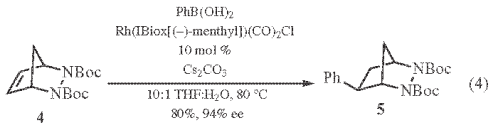
(2)



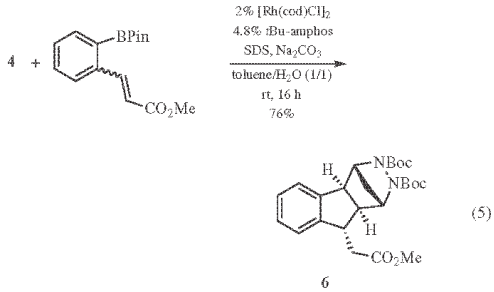
(3)

Enantioselective hydroarylation of compound (**4**)[5] has been developed using rhodium catalyst and several arylboronic acids. The use of a *N*-heterocyclic carbene (NHC) complex Rh(IBiox[(−)-menthyl])$(CO)_2$Cl gives excellent results (eq 4). In some cases, similar selectivities can be achieved using [{Rh(cod)OH}$_2$] as a precatalyst and (*R,S*)-*t*-Bu-josiphos as a ligand.[6]



(4)

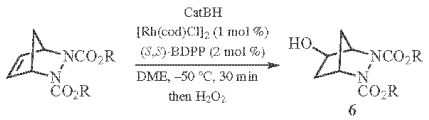
The combination of an electron-rich, sterically bulky ligand and a functionalized boronic ester enables the formation of a polycyclic compound in a fully distereoselective arylative cyclization reaction (eq 5).[7]



(5)

***Hydroboration.*** The olefin can be subjected to a fully diastereoselective hydroboration reaction.[8] Enantioselective hydroboration catalyzed by $[Rh(COD)Cl]_2$ was recently developed (eq 6).[9] The enantioselectivity of the reaction can be completely

reversed when switching from rhodium to iridium using the same chiral ligand.[10] Compounds (**8**) are excellent precursors for hydroxylated[11] or fluorinated diaminocyclopentanes.[12]

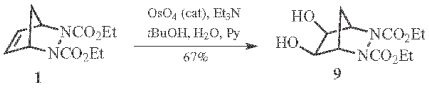



(6)

1 R = Et
7 R = Bn

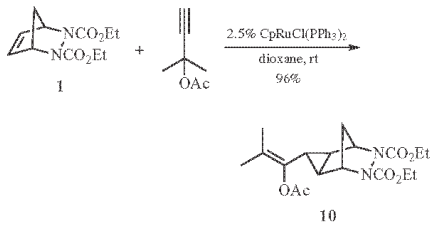
**8a** R = Et, 49%, 83% ee
**8b** R = Bn, 90%, 84% ee

***Dihydroxylation.*** The double bond of bicyclic hydrazine displays excellent reactivity toward electrophiles and can thus be smoothly dihydroxylated using a catalytic amount of osmium tetroxide (eq 7).[13] The exclusive *exo*-dihydroxylated product in this reaction is an excellent precursor for the preparation of hydroxylated diaminocyclopentanes after cleavage of the N–N bond.




(7)

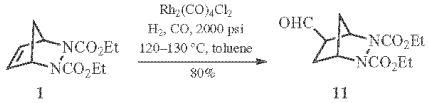
***Cyclopropanation.*** A ruthenium-catalyzed cyclopropanation of diazabicyclo alkene (**1**) with tertiary propargylic acetates affording 1,2,3-trisubstituted vinyl cyclopropane (**10**) as a single diastereomer was reported (eq 8).[14]




(8)

The cyclopropanation of several bicyclic hydrazines has also been reported using rhodium[15] or platinum[16] catalyst.

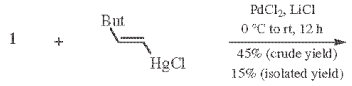
***Hydroformylation.*** The rhodium-catalyzed racemic hydroformylation of bicyclic hydrazine (**1**) to form the *exo*-aldehyde (**11**) proceeds in good yield at elevated temperatures (eq 9).[17] The hydroformylated product is sensitive and easily undergoes epimerization to the *endo*-aldehyde or fragmentation of the bicycle.

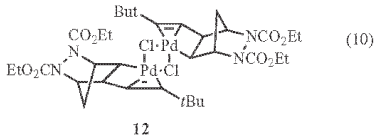



(9)

This reaction can also be performed in an enantioselective way, using $Rh(CO)_2(acac)$ as a precatalyst and various chiral diphosphines, leading to excellent conversions but only moderate *ee* values.[18]

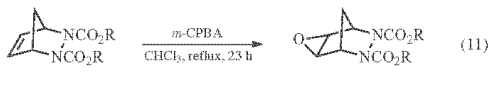
***Vinylpalladation.*** Bicyclic hydrazine (**1**) reacts with vinylmercuric chloride and $Li_2PdCl_4$ to form stable, isolable adduct (**12**) in which the vinyl and PdCl groups have been added in a *cis*-, *exo*-fashion across the C=C double bond (eq 10).[19] One example of Ni-catalyzed hydrovinylation of bicyclic hydrazine has also been reported.[20]




(10)

***Epoxidation.*** The diastereoselective epoxidation of the reactive C=C double bond of bicyclic hydrazines (**1**) or (**7**) with *m*-CPBA has been reported (eq 11).[13] These epoxides can generate transient aziridinium intermediates under various acidic conditions, which can be trapped by several nucleophiles in a regioselective way.[21]

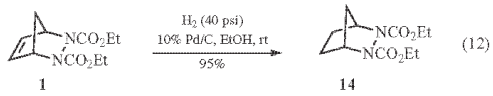



(11)

1 R = Et
7 R = Bn

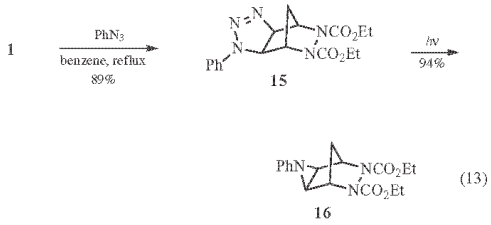
**13a** R = Et, 94%
**13b** R = Bn, 63%

***Hydrogenation.*** Although most of the attention has been focused on several electrophilic additions as well as metal-catalyzed reactions to the C=C double bond of the bicyclic hydrazine (**1**), its catalytic hydrogenation in the presence of 10% palladium on charcoal in absolute ethanol has been reported to give (**14**) in 95% yield (eq 12).[2]



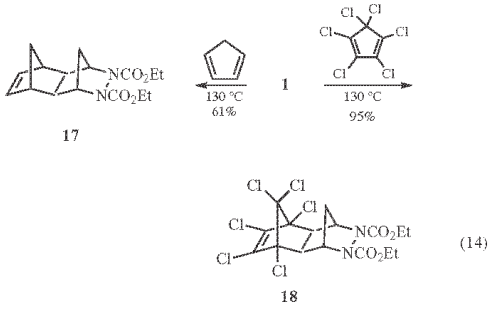

(12)

***Cycloaddition Reactions.*** Several [2 + 2],[22] [3 + 2],[23] or [4 + 2][24] cycloaddition reactions have been reported on bicyclic hydrazines. For instance, the cycloaddition of hydrazine (**1**) with various alkyl, aryl, and sulfonylazides lead to the formation of triazolines in high yields that can be photolyzed to the corresponding aziridines(eq 13).[25]



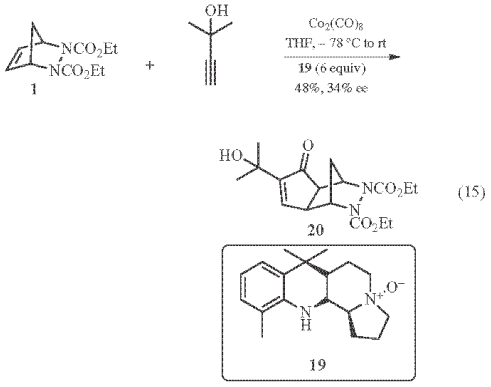

(13)

The [4+2] cycloaddition of hydrazine (**1**) with cyclopentadieneproceeds in moderate yield, however, with hexachlorocyclopentadiene at 130°C the adduct (**18**) is obtained in 95% yield as a stable, colorless, and crystalline solid (eq 14).[24]




(14)

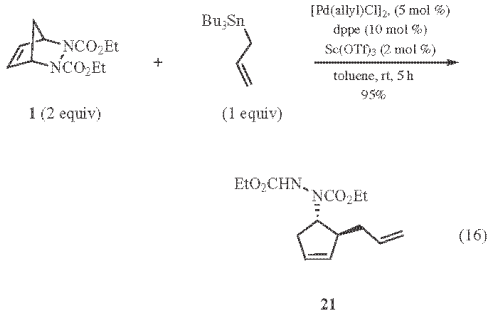
In addition to these cycloadditions, Pauson–Khand reaction[26] of (**1**) with an alkyne and carbon monoxide to afford a cyclopentenone adduct has also been reported. An enantioselective reaction has been evaluated on (**1**), with chiral amine *N*-oxides with modest enantioselectivities observed (eq 15).[27]




(15)

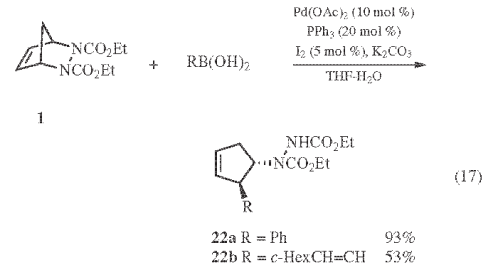
***Synthetic Transformations with Ring Cleavage.*** When the reaction with electrophiles is not fast enough, the metallated bicyclic hydrazines can undergo a favorable $\beta$-fragmentation that can be exploited to get access to functionalized hydrazinocyclopentenes of great synthetic interest.

***Palladium-Catalyzed Ring-Opening Reactions.*** The palladium-catalyzed ring opening of (**1**) with various organometallic reagents can be exploited in allylation reactions. The palladium-catalyzed ring opening of bicyclic hydrazine (**1**) using allyltributyltin or allyltrimethylsilane delivers 1,2-disubstituted cyclopentenes in a diastereoselective fashion (eq 16).[28,29] This reaction is catalyzed by Lewis acids and takes place faster when conducted in ionic liquids. Similar reactivity was reported with vinyl-, phenyl-, furyl-, and thienylstannanes as nucleophiles.[30] The ring opening of (**1**) using azidostannane is also possible under similar conditions.[31]



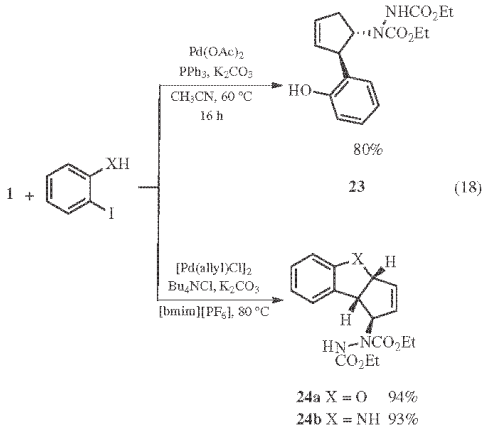

(16)

An efficient related synthesis of functionalized cyclopentenes was performed by treating the bicyclic hydrazine (**1**) with *in situ*-generated allyl- and benzylindium or gallium species in the presence of palladium catalyst, a ligand, and a Lewis acid.[32] Similarly, bicyclic hydrazine (**1**) undergoes a facile iodine-assisted stereoselective ring opening on reaction with aryl- or alkenylboronic acids affording *trans*-3,4-disubstituted hydrazinocyclopentenes in good to excellent yields (eq 17).[33]

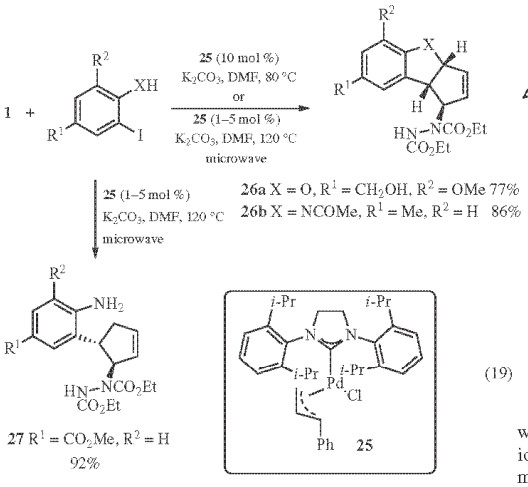



(17)

A palladium-catalyzed tandem ring opening–ring closing reaction of bicyclic hydrazines with *ortho*-functionalized aryl iodides was reported to afford cyclopentannulated heterocycles. The reaction proceeds with high chemo- and stereoselectivity in ionic liquids. Arylated hydrazinocyclopentenes are formed in acetonitrile with triphenylphosphine (eq 18).[34]

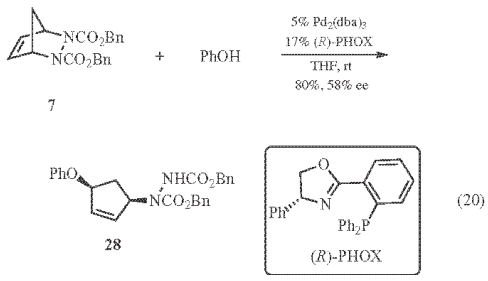



Interestingly, when triphenylphosphine is replaced with a *N*-heterocyclic carbene ligand, the reaction of bicyclic hydrazines with polysubstituted 2-iodophenols affords cyclopentannulated products as single diastereomers in high yields.[35] Under similar conditions, *N*-(2-iodophenyl)amides also react with (**1**) to give cyclopentannulated heterocycles as single diastereomers in good yields. However, with electron-deficient *N*-(2-iodophenyl)amides, only arylated hydrazinocyclopentenes with loss of the amine acetyl moiety are observed (eq 19).[35]

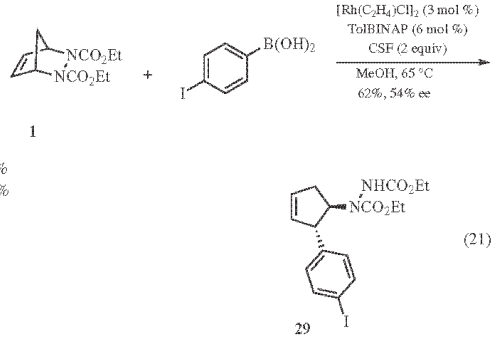



This palladium–NHC-catalyzed cyclopentannulation of (**1**) with numerous *ortho*-functionalized aryl iodides can be performed under conventional heating as well as under microwave irradiation without employing ionic liquids as solvents.

Palladium-catalyzed ring opening can also lead to *cis*-1,4-disusbstituted cyclopentenes. Enantioselectivities up to 58% can be obtained in the presence of P, N ligands (eq 20).[36]

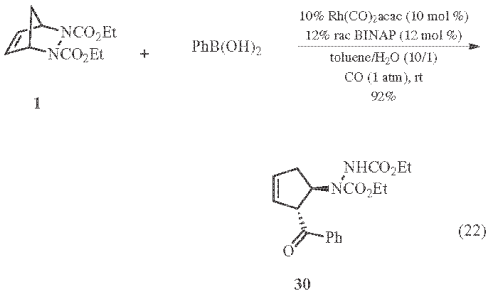


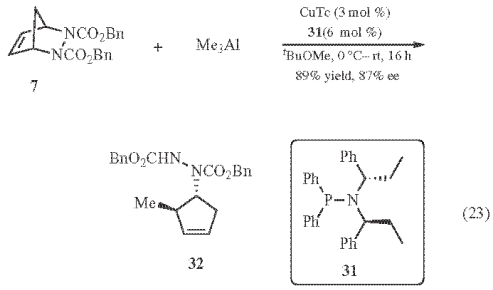

***Rhodium-Catalyzed Ring-Opening Reactions.*** Although palladium catalyzes the ring opening of bicyclic hydrazines with arylboronic acids, this fragmentation leads to racemic *trans*-1,2-disubstituted cyclopentenes. The use of rhodium and chiral phosphine ligand as a catalytic system for asymmetric arylation of (**1**) to give regioselective and diastereoselective access to arylated hydrazinocyclopentenes in an enantioenriched form has been reported. Enantioselectivities up to 89% were obtained for this transformation using $[Rh(C_2H_4)Cl]_2$ and TolBINAP as a bidentate ligand and arylboronic acids as arylating agents (eq 21).[37] A major improvement of this reaction has been recently described, with the use of *t*Bu-Josiphos ligand and [Rh(cod)OH]$_2$ as precatalyst.[6] In addition to aromatic boronic acids, alkynylboronic esters have also been used in rhodium-catalyzed desymmetrization of bicyclic hydrazines.[38]



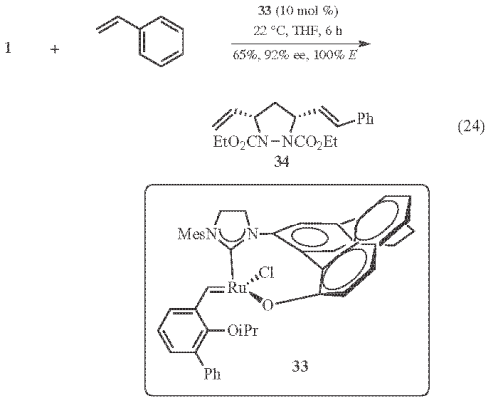

Densely functionalized *trans*-3-acyl,2-hydrazinocyclopentenes were obtained stereoselectively upon reaction with acyl anion equivalents generated in situ from organoboron and carbon monoxide in the presence of rhodium catalyst (eq 22).[39]

***Copper-Catalyzed Ring-Opening Reactions.*** Regioselective and antistereoselective asymmetric ring opening of diazabicyclic alkenes with hard alkyl metals and copper phosphinamide catalysts has been reported. The copper phosphinamide-catalyzed reaction of (**7**) with trialkylaluminum reagents is highly efficient and permits the desymmetrization of bicyclic hydrazines with high enantiomeric excesses (eq 23).[40]

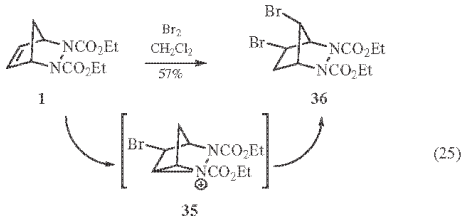


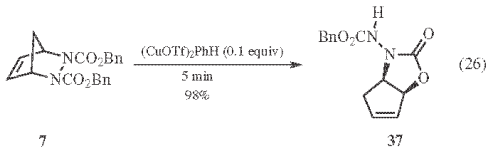

***Ruthenium-Catalyzed Ring-Opening Metathesis/Cross-Metathesis.*** Ring-opening metathesis oligomerization of (**1**) is promoted by Grubbs catalyst. In addition, the ring-opening cross-metathesis has been shown to deliver *E*/*Z* (2.5:1) mixture of cross-coupled products in the presence of 4-nitrostyrene with the same catalyst.[41] On the contrary, the use of catalyst (**33**) enables the enantioselective ring-opening metathesis/cross-metathesis of (**1**) with enantioselectivities ranging from 68% to 92% (eq 24).[42]




**Rearrangements.** Bicyclic hydrazines can undergo interesting skeletal rearrangements during electrophilic additions to the C=C double bond. Bromination of (**1**) is reported to trigger such a skeletal rearrangement via a transient secondary cation that is stabilized by one nitrogen, leading to an aziridinium (**35**) that can be attacked in a regioselective manner to afford compound (**36**) (eq 25).[43]




Moreover, bicyclic hydrazines are also prone to rearrange under Bronsted or Lewis acid catalysis to diastereoselectively form interesting [5.5] bicyclic hydrazines (eq 26). These building blocks can be further functionalized, for instance, in the same pot by a copper-catalyzed *N*-arylation.[44] They are also interesting starting material for the fast synthesis of amino cyclitols.[45]




**Related Reagents.** Dimethyl 2,3-diazabicyclo[2.2.1]-hept-5-ene-2,3-bis(carboperoxoate); diisopropyl 2,3-diazabicyclo[2.2.1]hept-5-ene-2,3-dicarboxylate; di-*tert*-butyl 2,3-diazabicyclo[2.2.1]hept-5-ene-2,3-dicarboxylate; dibenzyl 2,3-diazabicyclo[2.2.1]hept-5-ene-2,3-bis(carboperoxoate); 2,3-diazabicyclo[2.2.1]hept-5-ene-*N*,*N*′-phthalazide; 1,1′-(2,3-diazabicyclo[2.2.1]hept-5-ene-2,3-diyl)bis(3-tosylpropan-1-one); 1,1′-(2,3-diazabicyclo[2.2.1]hept-5-ene-2,3-diyl)bis(3,3,3-trichloropropan-1-one).


1. (a) Bournaud, C.; Chung, F.; Pérez Luna, A. P.; Pasco, M.; Errasti, G.; Lecourt, T.; Micouin, L., *Synthesis* **2009**, *6*, 869. (b) Sajisha, V. S.; Anas, S.; John, J.; Radhakrishnan, K. V., *Synlett* **2009**, *18*, 2885.
2. Gassman, P. G.; Mansfield, K. T., *Org. Synth.* **1969**, *49*, 1.
3. Helmchen, G.; Hoffmann, R. W.; Mulzer, E., Eds., *Methoden der Organischen Chemie (Houble-Weyl)*; Thieme: Stuttgart, **1995**, Vol. *21*, p. 5451.
4. (a) Storsberg, J.; Nandakumar, M. V.; Sankaranarayanan, S.; Kaufmann, D. E., *Adv. Synth. Catal.* **2001**, *343*, 177. (b) Yao, M. L.; Adiwidjaja, G.; Kaufmann, D. E., *Angew. Chem., Int. Ed.* **2002**, *41*, 3375.
5. Bexrud, J.; Lautens, M., *Org. Lett.* **2010**, *12*, 3160.
6. Menard, F.; Lautens, M., *Angew. Chem., Int. Ed.* **2008**, *47*, 2085.
7. Lautens, M.; Mancuso, J., *J. Org. Chem.* **2004**, *69*, 3478.
8. Allred, E. L.; Anderson, C. L.; Smith, R. L., *Tetrahedron Lett.* **1966**, *9*, 951.
9. Pérez Luna, A.; Ceschi, M.-A., Bonin, M.; Micouin, L.; Husson, H.-P., Gougeon, S.; Estenne-Bouhtou, G.; Marabout, B.; Servin, M.; George, P., *J. Org. Chem.* **2002**, *67*, 3522.

10. Pérez Luna, A.; Bonin, M.; Micouin, L.; Husson, H.-P., *J. Am. Chem. Soc.* **2002**, *124*, 12098.
11. (a) Chung, F.; Tisné, C., Lecourt, T.; Dardel, F.; Micouin, L., *Angew. Chem., Int. Ed.* **2007**, *46*, 4489. (b) Lombès, T.; Bégis, G.; Maurice, F.; Turcaud, S.; Lecourt, T.; Dardel, F.; Micouin, L., *ChemBioChem* **2008**, *9*, 1368. (c) Chung, F.; Tisné, C.; Lecourt, T.; Seijo, B.; Dardel, F.; Micouin, L., *Chem. Eur. J.* **2009**, *15*, 7109.
12. (a) Moumné, R., Pasco, M.; Prost, E.; Lecourt, T.; Micouin, L.; Tisné, C., *J. Am. Chem. Soc.* **2010**, *132*, 13111. (b) Pasco, M.; Moumné, R., Lecourt, T.; Micouin, L., *J. Org. Chem.* **2011**, *76*, 5237.
13. (a) Mellor, J. M.; Smith, N. M., *J. Chem. Soc., Perkin Trans. 1*. **1984**, 2927. (b) Grabowski, S.; Armbruster, J.; Prinzbach, H., *Tetrahedron Lett.* **1997**, *38*, 5485.
14. Tenaglia, A.; Marc, S., *J. Org. Chem.* **2006**, *71*, 3569.
15. Tseng, N.-W.; Mancuso, J.; Lautens, M., *J. Am. Chem. Soc.* **2006**, *128*, 5338.
16. Bigeault, J.; Giordano, L.; de Riggi, I.; Gimbert, Y.; Buono, G., *Org. Lett.* **2007**, *9*, 3567.
17. Wilson, R. M.; Schnapp, K. A.; Merwin, R. K.; Ranganathan, R.; Moats, D. L.; Conrad, T. T., *J. Org. Chem.* **1986**, *51*, 4028.
18. (a) Huang, J.; Bunel, E.; Allgeier, A.; Tedrow, J.; Storz, T.; Preston, J.; Correl, T.; Manley, D.; Soukup, T.; Jensen, R.; Syed, R.; Moniz, G.; Larsen, R.; Martinelli, M.; Reider, P. J., *Tetrahedron Lett.* **2005**, *46*, 7831. (b) Bournaud, C.; Lecourt, T.; Micouin, L.; Méliet, C.; Agbossou-Niedercorn, F., *Eur. J. Org. Chem.* **2008**, 2298.
19. (a) Larock, R. C.; Hershberger, S. S.; Takagi, K.; Mitchell, M. A., *J. Org. Chem.* **1986**, *51*, 2450. (b) Larock, R. C.; Takagi, K.; Hershberger, S. S.; Mitchell, M. A., *Tetrahedron Lett.* **1981**, *22*, 5231.
20. Liu, W.; RajanBabu, T. V., *J. Org. Chem.* **2010**, *75*, 7636.
21. Bournaud, C.; Bonin, M.; Micouin, L., *Org. Lett.* **2006**, *8*, 3041.
22. (a) Warrener, R. N.; Bina, S.; Butler, D. N.; Pitt, I. G.; Russel, R. A., *Tetrahedron Lett.* **1990**, *31*, 7073. (b) Warrener, R. N.; Pitt, I. G.; Nunn, E. E.; *Tetrahedron Lett.* **1994**, *35*, 621.
23. (a) Huisgen, R.; Moebius, L.; Mueller, G.; Stangl, H.; Szeimies, G.; Vernon, J. M., *Chem. Ber.* **1965**, *98*, 3992. (c) Wade, P. A.; Hinney, H. R.; *J. Am. Chem. Soc.* **1979**, *101*, 1319.
24. (a) Kuderna, J. G.; Sims, J. W.; Wikstrom, J. F.; Soloway, S. B., *J. Am. Chem. Soc.* **1959**, *81*, 382. (b) Hünig, S.; Kraft, P.; Klarner, F.-G., Artschwanger-Perl, U.; Peters, K.; Von Schnering, H. G., *Leibigs Ann.* **1995**, 351.
25. Stout, D. M.; Takaya, T.; Meyers, A. I., *J. Org. Chem.* **1975**, *40*, 563.
26. Pauson, P. L.; Khand, I. U., *Ann. N. Y. Acad. Sci.* **1977**, *295*, 2.
27. Derdau, V.; Laschat, S., *J. Organomet. Chem.* **2002**, *642*, 131.
28. Radhakrishnan, K. V.; Sajisha, V. S.; Anas, S.; Krishnan, K. S., *Synlett* **2005**, *15*, 2273.
29. Sajisha, V. S.; Mohanlal, S.; Anas, S.; Radhakrishnan, K. V., *Tetrahedron* **2006**, *62*, 3997.
30. Sajisha, V. S.; Radhakrishnan, K. V., *Adv. Synth. Catal.* **2006**, *348*, 924.
31. Anas, S.; Sajisha, V. S.; John, J.; Joseph, N.; George, S. C., Radhakrishnan, K. V., *Tetrahedron* **2008**, *64*, 9689.
32. (a) John, J.; Anas, S.; Sajisha, V. S.; Viji, S.; Radhakrishnan, K. V., *Tetrahedron Lett.* **2007**, *48*, 7225. (b) John, J.; Adarsh, B.; Radhakrishnan, K. V., *Tetrahedron* **2010**, *66*, 1383.
33. (a) John, J.; Sajisha, V. S.; Mohanlal, S.; Radhakrishnan, K. V., *Chem. Commun.* **2006**, 3510. (b) Anas, S.; John, J.; Sajisha, V. S.; John, J.; Rajan, R.; Suresh, E.; Radhakrishnan, K. V., *Org. Biomol. Chem.* **2007**, *5*, 4010.
34. John, J.; Indu, U.; Suresh, E.; Radhakrishnan, K. V.; *J. Am. Chem. Soc.* **2009**, *131*, 5042.
35. Bhanu Prasad, B. A.; Buechele, A. E.; Gilbertson, S. R., *Org. Lett.* **2010**, *12*, 5422.
36. Pérez Luna, A.; Cesario, M.; Bonin, M.; Micouin, L., *Org. Lett.* **2003**, *5*, 4771.
37. (a) Bertolini, F.; Macchia, F.; Pineschi, M., *Tetrahedron Lett.* **2006**, *47*, 9173. (b) Panteleev, J.; Menard, F.; Lautens, M., *Adv. Synth. Catal.* **2008**, *350*, 2893.
38. Crotti, S.; Bertolini, F.; Macchia, F.; Pineschi, M., *Chem. Commun.* **2008**, 3127.
39. Menard, F.; Weise, C. F.; Lautens, M., *Org. Lett.* **2007**, *9*, 5365.
40. (a) Pineschi, M.; Moro, F. D.; Crotti, P.; Macchia, F., *Org. Lett.* **2005**, *7*, 3605. (b) Palais, L.; Bournaud, C.; Micouin, L.; Alexakis, A., *Chem. Eur. J.* **2010**, *16*, 2567.
41. Ellis, J. M.; King, S. B., *Tetrahedron Lett.* **2002**, *43*, 5833.
42. (a) Van Velduizen, J. J.; Gillingham, D. G.; Garber, S. B.; Kataoka, O.; Hoveyda, A. H., *J. Am. Chem. Soc.* **2003**, *125*, 12502. (b) Hoveyda, A. H.; Lombardi, P. J.; O'Brien, R. V.; Zhugralin, A. R., *J. Am. Chem. Soc.* **2009**, *131*, 8378.
43. (a) Raasch, M. S., *J. Org. Chem.* **1975**, *40*, 161.
44. Martins, A.; Lemouzy, S.; Lautens, M., *Org. Lett.* **2009**, *11*, 181.
45. Bournaud, C.; Robic, D.; Bonin, M.; Micouin, L., *J. Org. Chem.* **2005**, *70*, 331.


B. A. Bhanu Prasad & Scott R. Gilbertson
*University of Houston, Houston, TX, USA*

Thomas Lombès & Laurent Micouin
*UMR 8638, CNRS, Paris Descartes University, Paris, France*

## c) Publication 3

Angewandte Chemie

Biosynthetic Pathway

DOI: 10.1002/ange.201201445

# A Sweet Origin for the Key Congocidine Precursor 4-Acetamidopyrrole-2-carboxylate**

*Sylvie Lautru,* Lijiang Song, Luc Demange, Thomas Lombès, Hervé Galons, Gregory L. Challis, and Jean-Luc Pernodet*

Pyrrole groups are found in nature in two primary metabolites (heme and tryptophan) and in a large variety of secondary metabolites (e.g. pyoluteorin, clorobiocin, congocidine, and prodigiosins). The diversity of these metabolites is mirrored in the multiple biosynthetic pathways leading to pyrrole groups. Six different biosynthetic pathways have been characterized to date,[1,2] involving various primary metabolite precursors, such as amino acids (glycine, proline, serine, threonine, and tryptophan),[1–10] dicarboxylic acids (malonate, oxaloacetate, and succinate),[1,2,7,8] or *N*-(5′-phosphoribosyl)-anthranilate.[1]

Pyrrolamides are a family of natural products, synthesized by *Streptomyces* and related actinobacteria, that all contain one or more pyrrole-2-carboxamide moieties in their structure. Most pyrrolamides, such as the well-characterized congocidine (**1**, also called netropsin; Figure 1) and distamycin, bind noncovalently in the DNA minor groove with some sequence specificity,[11] but pyrronamycin B has been suggested to bind DNA covalently.[12] This capacity to bind DNA confers on them a variety of biological activities, such as antiviral, antibacterial, antitumor, and anthelmintic activities, but also renders them too toxic for clinical use. Nonetheless, because molecules binding to DNA at specific positions can manipulate the expression of genes involved in various diseases (such as cancer), congocidine and distamycin have prompted the synthesis of many structurally related molecules that bind in the minor groove of DNA at various defined sequences.[13]

The nature and biosynthetic origin of the pyrrolamide pyrrole precursor are unknown. A retrobiosynthetic analysis of pyrrolamide structures suggests 4-aminopyrrole-2-carboxylate as the potential pyrrole precursor common to all pyrrolamides. However, this remains to be established and no biosynthetic pathway has been reported for the synthesis of this molecule. We report herein that 4-acetamidopyrrole-2-carboxylate (**10**) is the true precursor of congocidine and propose for this compound a biosynthetic pathway starting from *N*-acetylglucosamine-1-phosphate (**2**), involving carbohydrate metabolizing enzymes, and differing entirely from known pyrrole biosynthetic pathways.

We recently reported the identification, analysis, and heterologous expression of the first pyrrolamide gene cluster; the *cgc* gene cluster directs congocidine biosynthesis in *Streptomyces ambofaciens*.[14] Sequence analyses of the proteins encoded by the *cgc* cluster led us to propose a pathway for congocidine biosynthesis involving a noncanonical nonribosomal peptide synthetase (NRPS) and three putative precursors: guanidinoacetate, 3-aminopropionamidine, and 4-aminopyrrole-2-carboxylate. However, the biosynthetic origins of these precursors could not easily be inferred from the analysis of the *cgc* gene cluster. In particular, we could not identify any gene encoding homologues of known pyrrole biosynthetic enzymes, suggesting that the pyrrole groups in pyrrolamides could be synthesized through an entirely new pathway.

The *cgc18* gene encodes an NRPS module that has been proposed to recognize 4-aminopyrrole-2-carboxylate and catalyze its ATP-dependent activation as the corresponding peptidyl carrier protein (PCP)-bound thioester.[14] To investigate whether 4-aminopyrrole-2-carboxylate is indeed a precursor of congocidine, we used HPLC to compare the profile of metabolites in the culture supernatants of the "wild type" SPM110 *Streptomyces ambofaciens* strain and the CGCA004 mutant strain, in which *cgc18* has been disrupted.[14] A new compound accumulated in the culture supernatant of CGCA004 (retention time = 12.3 min, Figure 2a), which we expected to be 4-aminopyrrole-2-carboxylate. This metabolite peak was not detected in the CGCA004 chromatogram of our previous report[14] and on average, it was detected in approximately half of the cultures of strains deleted in *cgc*

[*] Dr. S. Lautru, Dr. J.-L. Pernodet
Institut de Génétique et Microbiologie
Université Paris-Sud, CNRS, UMR8621, 91405 Orsay (France)
E-mail: sylvie.lautru@igmors.u-psud.fr
Dr. L. Song, Prof. G. L. Challis
Department of Chemistry, University of Warwick,
Coventry CV4 7AL (UK)
Dr. L. Demange, Prof. H. Galons
Laboratoire de Chimie et Biochimie Pharmacologiques et Toxicologiques UMR 8601 CNRS
Université Paris Descartes, Sorbonne Paris Cité, UFR Biomédicale des Saints Pères
45 rue des Saints Pères, 75006 Paris (France)
T. Lombès
UMR 8638 CNRS-Paris Descartes, Faculté des Sciences Pharmaceutiques et Biologiques
4 av. de l'Observatoire, 75006 Paris (France)

[**] We would like to thank Dr. Christophe Corre for helpful discussions, Dr. Laurent Micouin for his help with the synthesis of the 4-acetamidopyrrole-2-methanol, and Dr. Cyrille Kouklovsky, Géraldine Le Goff, and Dr. Robert Thaï for their help with mass spectrometry. This work was supported in part by the European Union through the Integrated Project ActinoGEN (CT-2004-0005224) and the reintegration grant BOPA (CT-2005-029154), by the Pôle de Recherche et d'Enseignement Supérieur UniverSud Paris and by the Bettencourt Schueller Foundation.

Supporting information for this article (experimental details) is available on the WWW under http://dx.doi.org/10.1002/anie.201201445.



© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

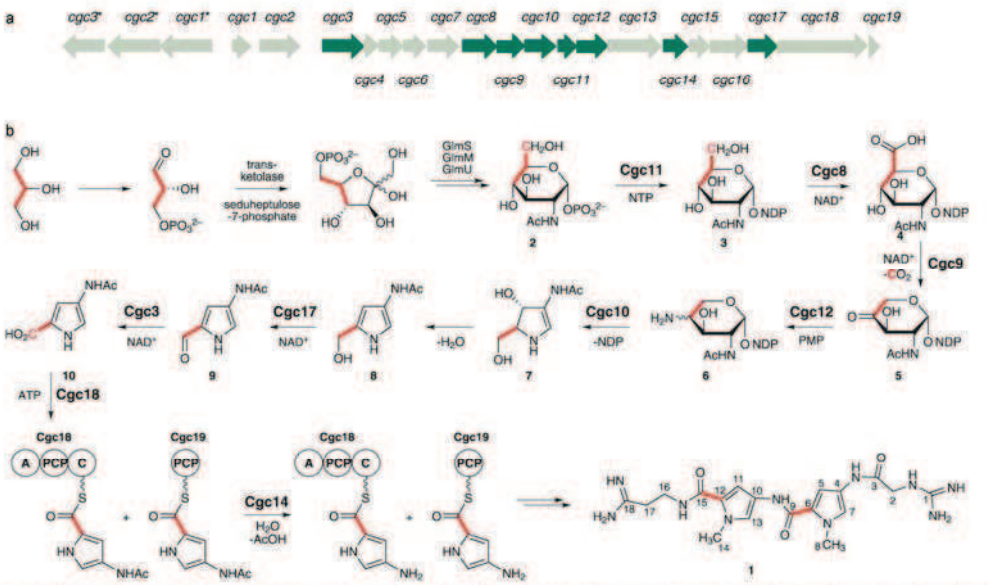



***Figure 1.*** Congocidine gene cluster (a) and proposed pathway for biosynthesis of 4-acetamidopyrrole-2-carboxylate (**10**) from fructose-6-phosphate by way of an intermediate **2** (b). [U-$^{13}$C]Glycerol is proposed to be converted into [U-$^{13}$C]glyceraldehyde-3-phosphate, which undergoes transaldolase-catalyzed condensation with seduheptulose-7-phosphate with loss of erythrose-4-phosphate to yield fructose-6-phosphate bearing $^{13}$C labels at C-4, C-5, and C-6. The positions of the $^{13}$C labels are highlighted in red. The primary metabolic enzymes GlmSMU catalyze conversion of fructose-6-phosphate into **2**. Genes involved in **10** biosynthesis are indicated in green.

genes not involved in the biosynthesis of **10**. It is likely that this variability is related to the variability of secondary metabolite production often seen in *Streptomyces*.[15] LC-MS analysis of the CGCA004 culture supernatant showed that this metabolite produces an *m/z* 168.9 ion rather than the expected *m/z* 127.1 ion (Supporting Information, Figure S1 a). MS/MS analyses suggested that the accumulated metabolite was **10** (Supporting Information, Figure S1 a). This hypothesis was confirmed by the preparation of a standard of **10**, whose retention time (Figure 2 b), UV/Vis spectrum, and MS/MS fragmentation pattern (Supporting Information, Figure S1 b) were identical to those recorded for the metabolite accumulated in the CGCA004 culture supernatant.

We next confirmed that **10** is an intermediate in congocidine biosynthesis and investigated the pathway for its assembly. Among the 22 genes constituting the *cgc* gene cluster, 13 could not be assigned a function in regulation of gene expression, congocidine resistance, or assembly of the putative precursors, guanidinoacetate, 3-aminopropionamidine, and **10**.[14] These 13 genes encode proteins with similarity to enzymes of known function (with the exception of *cgc7*) and are likely to be involved in the biosynthesis of congocidine precursors. Six of the genes encode proteins that are similar to enzymes involved in carbohydrate biosynthesis or transfer (Cgc8, putative NDP-hexose dehydrogenase; Cgc9, putative NDP-hexose epimerase/dehydrogenase; Cgc10, putative glycosyltransferase; Cgc11, putative hexose nucleotidyltransferase; Cgc12, putative DegT/DnrJ/EryC1/StrS aminotransferase; Cgc13, putative glycoside hydrolase). The presence of these genes in the *cgc* cluster is intriguing because congocidine does not contain a carbohydrate moiety. Yet, we have previously shown that at least one of these genes, *cgc10*, is involved in congocidine biosynthesis.[14] To determine whether the other five genes are also involved in congocidine biosynthesis, we constructed in-frame deletions of each within a bacterial artificial chromosome (BAC; pCGC002) containing the entire *cgc* cluster. The resulting constructs were heterologously expressed in *S. lividans* TK23. None of the *S. lividans* strains with the genetically modified BACs produced congocidine, which would elute at 15.1 min, (Figure 2 dand Supporting Information, Figure S3, left panels) except the CGCL047 strain (*Δcgc13*), which produced residual congocidine (15% of the "wild type" strain CGCL006). This indicates that the genes encoding putative carbohydrate biosynthesis/transfer enzymes are involved in congocidine biosynthesis and, therefore, are likely to participate in biosynthesis of a congocidine precursor.

We examined which of the carbon atoms of congocidine are carbohydrate-derived using incorporation experiments with labeled precursors. Cultures of *S. ambofaciens* SPM110 were pulse-fed with uniformly $^{13}$C-labeled glycerol ([U-$^{13}$C]glycerol; 4 × 100 μmol) and congocidine was purified by cation exchange chromatography. The DEPTQ-135 NMR spectrum (Supporting Information, Figure S4) of the isolated congocidine showed coupled doublets flanking the natural abundance signals for C-2/C-3, C-6/C-9, and C-12/C-15,

 © 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 

These are not the final page numbers!

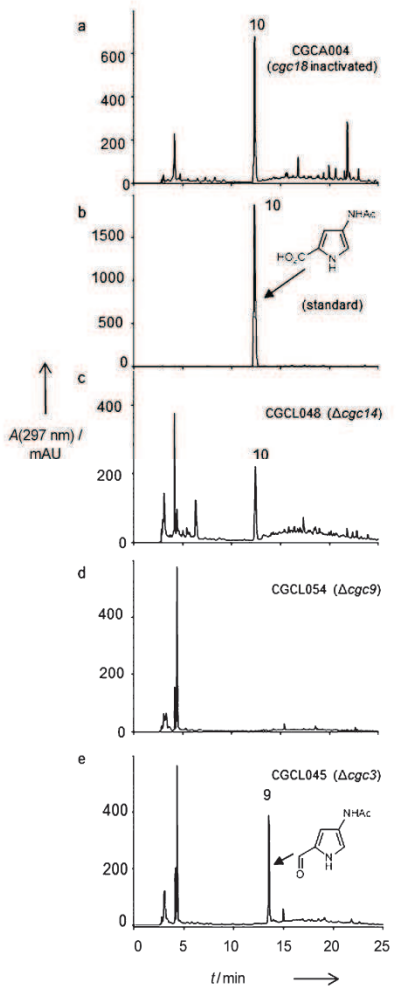



***Figure 2.*** LC analysis of culture supernatents of mutants with deletions in congocidine biosynthetic genes. Chromatograms from a) CGCA004 (disruption of *cgc18*); b) synthetic **10**; c) CGCL048 (*Δcgc14*); d) CGCL054 (*Δcgc9*), and e) CGCL045 (*Δcgc3*).

showing that a two-carbon fragment of glycerol was incorporated intact into C-2/C-3, C-6/C-9, and C-12/C-15 (see Figure 1 for carbon atom numbering). The guanidinoacetate precursor of congocidine has been suggested to originate from glycine by way of transguanylation.[16] This is consistent with the incorporation pattern seen for C-2/C-3 (Supporting Information, Figure S5). More significant, however, is the incorporation of two carbons from glycerol into the same position (C-2 and the carboxamide group) in both 4-aminopyrrole-2-carboxamide units of congocidine. This incorporation pattern is consistent with derivation of the 4-aminopyrrole-2-carboxamide units from fructose-6-phosphate by way of a **2** intermediate (Figure 1).[17] To verify this hypothesis, we next attempted to feed cultures of *S. ambofaciens* with [$^{15}$N,$^{13}$C-1] *N*-acetylglucosamine. Unfortunately, the addition of *N*-acetylglucosamine strongly suppresses congocidine production (presumably because of DasR-mediated regulation).[18] Thus, we were not able to obtain enough congocidine to determine whether *N*-acetylglucosamine was incorporated.[19]

Taken together these data suggest that the putative carbohydrate-processing enzymes encoded within the *cgc* cluster are responsible for the assembly of **10** from **2**. To further investigate this hypothesis and examine whether **10** is an intermediate in congocidine biosynthesis, we fed chemically synthesized **10** to cultures of the *cgc11*, *cgc8*, *cgc9*, *cgc12*, and *cgc10* deletion mutants. LC-MS analyses showed that production of congocidine was restored in all of the mutants (Supporting Information, Figure S3, central and right panels), indicating that **10** is a congocidine precursor and that *cgc11*, *cgc8*, *cgc9*, *cgc12*, and *cgc10* are all involved in its biosynthesis. When the *cgc13* deletion mutant was fed with **10**, congocidine production was increased by a factor of 2.5 (Supporting Information, Figure S3h). This result suggests that the reduced congocidine production in the *cgc13* deletion mutant is due to impaired synthesis of **10**.

To determine whether other *cgc* genes are involved in **10** biosynthesis, the six *cgc* genes with no assigned function (*cgc3*, *cgc4*, *cgc5*, *cgc6*, *cgc7*, and *cgc17*) were individually deleted in frame within the pCGC002 BAC and the resulting BACs heterologously expressed in *S. lividans* TK23. The *cgc4* deletion mutant produced residual congocidine, whereas congocidine production was abolished in the other deletion mutants (Supporting Information, Figure S3 f, g, left panels; Figure S6, left panels). The *cgc4*, *cgc5*, *cgc6*, and *cgc7* deletion mutants accumulated **10** (confirmed by MS analyses, Supporting Information, Figure S6, right panels), indicating that these genes are not involved in **10** biosynthesis. The *cgc3* and *cgc17* deletion mutants, however, did not produce **10**. The involvement of Cgc3 (a putative aldehyde dehydrogenase)[14] and Cgc17 (a putative alcohol dehydrogenase)[14] in **10** biosynthesis was confirmed by feeding **10** to the mutant strains, which restored congocidine production (Supporting Information, Figure S3 f, g, central and right panels).

The *cgc3* deletion mutant accumulated a new metabolite in the culture medium that was not found in the other mutants (Supporting Information, Figure S3g). MS and NMR spectroscopic analyses of the purified metabolite (Supporting Information, Figure S7) showed that it was 4-acetamidopyrrole-2-carboxaldehyde (**9**), which is likely to be the substrate of Cgc3 (Figure 3).

The *cgc17* gene encodes a putative alcohol dehydrogenase possibly involved in the oxidation of 4-acetamidopyrrole-2-methanol (**8**) to yield **9** (Figure 1). To verify this hypothesis, cultures of the *cgc11*, *cgc8*, *cgc9*, *cgc12 cgc10*, *cgc17*, and *cgc3* deletion mutants were fed with chemically synthesized **8** (Supporting Information, methods and Figure S8). Production of congocidine was restored for the first five mutant strains (Supporting Information, Figure S9). In contrast, feeding of **8** to the *cgc17* and *cgc3* deletion mutants did not


© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

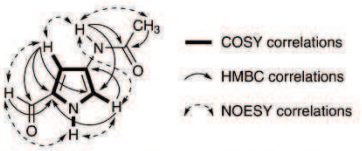



*Figure 3.* Summary of COSY, HMBC, and NOESY correlations recorded for **9** purified from the culture supernatant of the *cgc3* mutant.

restore congocidine production. These results are consistent with **8** being the substrate of Cgc17. Accumulation of **8** in the *cgc17* deletion mutant could not be detected using standard LC-MS analysis because this compound is unstable and poorly retained on reverse-phase HPLC columns.

Based on our data, we suggest the following pathway for **10** biosynthesis from **2** (Figure 1). Cgc11 catalyzes nucleotide triphosphate (NTP)-dependent conversion of **2** into the nucleotide diphosphate (NDP) derivative **3**, which undergoes Cgc8-catalyzed oxidation of the C-6 hydroxy group to a carboxy group to yield NDP-2-acetamido-2-deoxyglucopyranuronic acid (**4**). Cgc9 catalyzes the decarboxylation of **4** to afford NDP-threo-2-acetamido-2-deoxy-pentopyran-4-ulose (**5**). Reductive amination of the C-4 keto group of **5** is catalyzed by the putative pyridoxal-5-phosphate-dependent aminotransferase Cgc12. The next step involves intramolecular nucleophilic attack of the C-4 amino group on C-1 to displace the NDP and afford 4-acetamido-2,3-dihydropyrrole-2-methanol (**7**). This conversion of a carbohydrate into a pyrrole molecule is proposed to be catalyzed by Cgc10, an enzyme that resembles a glycosyltransferase. Loss of water from **7** to form **8** could then occur spontaneously (the formation of the resulting pyrrole aromatic ring is expected to be thermodynamically favorable). Finally, oxidation of the hydroxy group of **8** to the aldehyde **9** and subsequently the carboxylic acid **10** is catalyzed by Cgc17 and Cgc3, respectively.

The initial steps of our proposed pathway are catalyzed by enzymes (Cgc11, Cgc8, Cgc9, and Cgc12) that are homologous to enzymes catalyzing similar reactions in the biosynthesis of aminodeoxypentose sugars found in enediyne natural products such as maduropeptin, calicheamicin, and AT2433 (Supporting Information, Figure S10).[20,21] Carbohydrates incorporated in these natural products can be activated using UTP or TTP. By analogy, it appears likely that UTP or TTP is the NTP used to activate **2**. UDP-*N*-acetylglucosamine, derived from **2**, is an intermediate in the biosynthesis of bacterial peptidoglycan.[22,23] As the *cgc11* deletion mutant does not produce any congocidine, it is unlikely that UDP-*N*-acetylglucosamine is an intermediate in congocidine biosynthesis. Indeed, activation of **2** by a nucleotide triphosphate other than UTP (most likely TTP) could constitute the first committed step of **10** biosynthesis.

Our data show that Cgc13 is not essential for congocidine biosynthesis. Nonetheless the reduction of congocidine production in the *cgc13* deletion mutant, and the partial restoration of the production when the mutant is fed with **10**, indicates that Cgc13 may participate in the supply of **10**. The similarity of Cgc13 with glycoside hydrolases suggests that Cgc13 could be involved in providing **2**.

Our proposed pathway results in the biosynthesis of **10**. Yet, before condensation with the other congocidine precursors can occur, this molecule must be de-acetylated to yield 4-aminopyrrole-2-carboxylate. Among the proteins encoded by the *cgc* gene cluster, we identified Cgc14, a predicted amidohydrolase, as the enzyme likely to catalyze the de-acetylation. Therefore, we constructed a *cgc14* in-frame deletion mutant (CGCL048). This mutant did not produce congocidine but accumulated **10** (Figure 2c), consistent with the proposed role for Cgc14. However a mutant strain impaired in congocidine assembly (CGCA004) also accumulates **10** and not 4-aminopyrrole-2-carboxylate. These observations suggest that **10**, rather than 4-aminopyrrole-2-carboxylate, may be the substrate of the Cgc18 adenylation domain and that de-acetylation occurs after **10** has been loaded onto the Cgc18 and Cgc19 PCPs (Figure 1). Consistent with this hypothesis, feeding of the *cgc14* deletion mutant with commercially available 4-aminopyrrole-2-carboxylate did not restore congocidine production (Supporting Information, Figure S11). *N*-acetylation is a well characterized mechanism of arylamine detoxification and maintaining the 4-aminopyrrole-2-carboxylate in its *N*-acetylated form when free in solution would avoid the cytotoxic effects associated with arylamines.[24]

In conclusion, we have shown that **10** is the true intermediate in the biosynthesis of congocidine. Interestingly, the biosynthesis of **10** described herein is unprecedented in that it involves enzymes and precursors from carbohydrate metabolism. In particular, it bears no resemblance to the biosynthesis of other pyrrole-2-carboxylates incorporated into secondary metabolites such as clorobiocin,[6] coumermycin,[4,6] pyoluteorin,[3,25] and undecylprodiginine.[3,7,8] Although this pathway is likely to be involved in the biosynthesis of other members of the pyrrolamide family of metabolites, it is not currently known whether it is also utilized in the biosynthesis of other pyrrole-containing natural products. However, database searches identified a putative **10** biosynthetic gene cluster in the genome of *Micromonospora carbonacea var. Africana* ATCC 39149 (MCAG_05565 and MCAG_05568 to MCAG_05575). Moreover, it is interesting to note that this pathway, with the reductive amination reaction replaced by a ketoreduction, would be perfectly adapted to the biosynthesis of the furan precursors of proximicins, analogues of congocidine recently isolated from marine actinomycetes.[26]

Received: February 21, 2012
Revised: April 11, 2012
Published online: ■■ ■■, ■■■■

**Keywords:** biosynthesis · congocidine · natural products · pyrrolamide · pyrrole

[1] C. T. Walsh, S. Garneau-Tsodikova, A. R. Howard-Jones, *Nat. Prod. Rep.* **2006**, *23*, 517 – 531.
[2] S. Siebenberg, N. Burkard, A. Knuplesch, B. Gust, S. Grond, L. Heide, *ChemBioChem* **2011**, *12*, 2677 – 2685.

 © 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 

These are not the final page numbers!

[3] M. G. Thomas, M. D. Burkart, C. T. Walsh, *Chem. Biol.* **2002**, *9*, 171–184.
[4] H. Xu, Z.-X. Wang, J. Schmidt, L. Heide, S.-M. Li, *Mol. Genet. Genomics* **2002**, *268*, 387–396.
[5] S. Maharjan, N. Aryal, S. Bhattarai, D. Koju, J. Lamichhane, J. K. Sohng, *Appl. Microbiol. Biotechnol.* **2012**, *93*, 687–696.
[6] S. Garneau, P. C. Dorrestein, N. L. Kelleher, C. T. Walsh, *Biochemistry* **2005**, *44*, 2770–2780.
[7] S. Mo, P. K. Sydor, C. Corre, M. M. Alhamadsheh, A. E. Stanley, S. W. Haynes, L. Song, K. A. Reynolds, G. L. Challis, *Chem. Biol.* **2008**, *15*, 137–148.
[8] A. E. Stanley, L. J. Walton, M. Kourdi Zerikly, C. Corre, G. L. Challis, *Chem. Commun.* **2006**, 3981–3983.
[9] T. Nishizawa, S. Grüschow, D.-H. E. Jayamaha, C. Nishizawa-Harada, D. H. Sherman, *J. Am. Chem. Soc.* **2006**, *128*, 724–725.
[10] S. Kirner, P. E. Hammer, D. S. Hill, A. Altmann, I. Fischer, L. J. Weislo, M. Lanahan, K. H. van Pée, J. M. Ligon, *J. Bacteriol.* **1998**, *180*, 1939–1943.
[11] S. Neidle, *Nat. Prod. Rep.* **2001**, *18*, 291–309.
[12] A. Asai, Y. Sakai, H. Ogawa, Y. Yamashita, S. Kakita, K. Ochiai, T. Ashizawa, A. Mihara, T. Mizukami, H. Nakano, *J. Antibiot.* **2000**, *53*, 66–69.
[13] P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 2215–2235.
[14] M. Juguet, S. Lautru, F.-X. Francou, S. Nezbedová, P. Leblond, M. Gondry, J.-L. Pernodet, *Chem. Biol.* **2009**, *16*, 421–431.
[15] S. Siebenberg, P. M. Bapat, A. E. Lantz, B. Gust, L. Heide, *J. Biosci. Bioeng.* **2010**, *109*, 230–234.
[16] M. E. Wildfeuer, The biosynthesis of netropsin, PhD thesis, University of Delaware (USA), **1963**.
[17] Fructose-6-phosphate can also be formed from fructose-bisphosphate aldolase-mediated condensation of glyceraldehyde-3-phosphate and dihydroxyacetone phosphate with subsequent dephosphorylation. In this case, we would expect all six carbon atoms of fructose-6-phosphate to be labeled through incorporation of two molecules of [U-$^{13}$C]glycerol. However, fructose-bisphosphate aldolase does not appear to be expressed/active in *Streptomyces* species cultivated in rich growth media (see C. Corre, G. L. Challis, *ChemBioChem* **2005**, *6*, 2166–2170 and G. L. Challis, K. F. Chater, *Chem. Commun.* **2001**, 935–936). Thus, only incorporation of a single molecule of [U-$^{13}$C]glycerol into fructose-6-phosphate by way of the pentose phosphate pathway is observed.
[18] M. A. Świątek, E. Tenconi, S. Rigali, G. P. van Wezel, *J. Bacteriol.* **2012**, *194*, 1136–1144.
[19] We tried to feed [U-$^{13}$C]glucose to a mutant (CGCL022) that accumulates **10**. Some of the **10** was purified and analyzed by HRMS (see Figure S12). As expected, $^{13}$C-5 **10** was produced. However, in addition to the $^{13}$C-5 labeled compound and to the unlabeled compound, $^{13}$C-1, $^{13}$C-2, $^{13}$C-3, $^{13}$C-4 and $^{13}$C-7 labeled molecules were also observed. This nonspecific incorporation pattern can be expected with labeled glucose feeding experiments, as many primary metabolites that could be incorporated into the **2** precursor of **10** are derived from glucose.
[20] Q. Gao, C. Zhang, S. Blanchard, J. S. Thorson, *Chem. Biol.* **2006**, *13*, 733–743.
[21] S. G. van Lanen, T.-J. Oh, W. Liu, E. Wendt-Pienkowski, B. Shen, *J. Am. Chem. Soc.* **2007**, *129*, 13082–13094.
[22] D. Mengin-Lecreulx, J. van Heijenoort, *J. Bacteriol.* **1994**, *176*, 5788–5795.
[23] A. M. Gehring, W. J. Lees, D. J. Mindiola, C. T. Walsh, E. D. Brown, *Biochemistry* **1996**, *35*, 579–585.
[24] M. Payton, A. Mushtaq, T. W. Yu, L. J. Wu, J. Sinclair, E. Sim, *Microbiology* **2001**, *147*, 1137–1147.
[25] B. Nowak-Thompson, N. Chaney, J. S. Wing, S. J. Gould, J. E. Loper, *J. Bacteriol.* **1999**, *181*, 2166–2174.
[26] K. Schneider, et al., *Angew. Chem.* **2008**, *120*, 3302–3305; *Angew. Chem. Int. Ed.* **2008**, *47*, 3258–3261.



© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

Angewandte International Edition Chemie

RNA Structures

DOI: 10.1002/anie.201204083

# Investigation of RNA–Ligand Interactions by $^{19}$F NMR Spectroscopy Using Fluorinated Probes**

*Thomas Lombès, Roba Moumné, Valéry Larue, Elise Prost, Marjorie Catala, Thomas Lecourt, Frédéric Dardel, Laurent Micouin,* and Carine Tisné**

Ribonucleic acid (RNA) is now recognized as playing a key role in many biological functions, and is emerging as an important new drug target.[1] However, its therapeutic potential is still underexploited.[2] Indeed, the limited understanding of the interactions between small molecules and RNA still hampers rational drug development of RNA-targeting molecules. Among the different methods available to investigate binding between small molecules and RNA,[3] NMR spectroscopy is particularly attractive as it can deliver information on molecular interactions at the atomic level, including conformational rearrangements that can occur before or upon binding.[4] The dynamic nature of this interaction is particularly important in RNA-regulated pathways.[5]

Many NMR spectroscopy techniques have been developed to visualize dynamic RNA–ligand interactions, most of them based on the observation of either the target or ligand $^1$H nuclei.[6] However, some difficulties can occur when studying larger strands of RNA, as the number of detectable signals will increase. Introducing a specific label is one way to overcome this problem. An elegant method based on $^{19}$F NMR spectroscopy was proposed some years ago by Micura and co-workers. Introduction of a fluorine atom at a specific position of RNA allows local monitoring of binding events at this site.[7] One technical difficulty with this approach is the need to chemically modify the RNA, which can be difficult for large RNAs or for RNAs with modified nucleotides. Furthermore, this modification can affect RNA–ligand interactions.

We have recently shown that $^{19}$F NMR spectroscopy can be used to monitor the binding of racemic fluorinated molecules to various RNAs and that chiral recognition can be used to monitor the local conformation of the binding site.[8] The use of ligand-based binding-competition NMR screening using fluorinated ligands has been described by Dalvit for the investigation of protein–ligand interactions (FAXS technique),[9] but has not been applied to the study small molecules interacting with RNA. We report herein that competitive binding of fluorinated probes can be used to detect and quantify the interaction between unlabeled RNA and non-fluorinated ligands and to monitor dynamic RNA folding events (Figure 1).

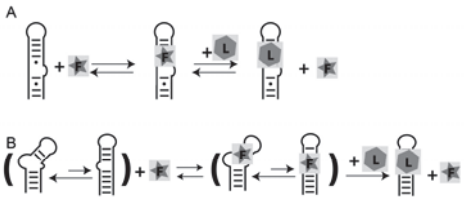



***Figure 1.*** Principle of the displacement experiments. A) Competition for binding to a structured RNA between a fluorinated probe (F, star) and an RNA ligand (L, hexagon). B) Competition for binding to a bistable RNA. In both cases, the fluorinated reporter is monitored by $^{19}$F NMR.

The binding of aminoglycosides with 16S23 RNA, a 23 nucleotide hairpin that mimics the decoding A-site of 16S ribosomal RNA, was first investigated as a case study. Although absolute $K_D$ values are difficult to compare since experimental conditions significantly vary from one study to another, the $K_D$ values of the most studied ligands of 16S A-site RNA can be ordered as follows: neomycin < paromomycin < neamine < paromamine, neomycin being the strongest binder and paromamine the weakest.[10] Fluorinated compound **1** being an analogue of desoxystreptamine (DOS), the core moiety of most aminoglycosides, we first checked that it bound to the 16S23 RNA.[11] Its binding was monitored using both $^{19}$F and $^1$H NMR spectroscopy (Figure 2). A $K_D$ of 2 mM was determined by NMR titration. As reported for DOS,[12] compound **1** interacts with two equivalent binding sites of 16S23 RNA in a fast-exchange regime.[13]

We then conducted ligand-based binding competition using compound **1** as a fluorinated spy probe and known binders of 16S23 RNA (neamine **2**, paromamine **3**, neomycin **4**) as competitors. Progressive displacement of the fluorinated reporter could be observed by increasing the concentration of the competitor (Figure 3), the $K_D$ value of paromamine, neamine, and neomycin for 16S23 could be estimated in the

[*] T. Lombès, Dr. R. Moumné, E. Prost, Dr. T. Lecourt, Dr. L. Micouin
UMR 8638 University Paris Descartes, CNRS, Faculté de Pharmacie
4 av. de l'Observatoire, 75006 Paris (France)
E-mail: laurent.micouin@parisdescartes.fr
Homepage: http://www.ssmip.cnrs.fr/eq3_micouin.html
Dr. V. Larue, M. Catala, Prof. F. Dardel, Dr. C. Tisné
UMR 8015 University Paris Descartes, CNRS, Faculté de Pharmacie
4 av. de l'Observatoire, 75006 Paris (France)
E-mail: carine.tisne@parisdescartes.fr
Homepage: http://lcrbw.pharmacie.univ-paris5.fr/spip.php?rubrique10

[**] Financial support from CNRS Bettencourt Schueller Foundation (PhD grant to T.L.), ANRS (grant to R.M.) and ANR (research project PCV TriggeRNA) is acknowledged.

Supporting information for this article (experimental details) is available on the WWW under http://dx.doi.org/10.1002/anie.201204083.



© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

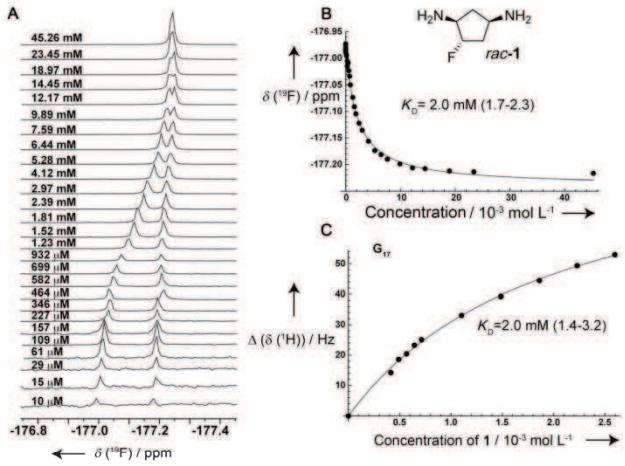
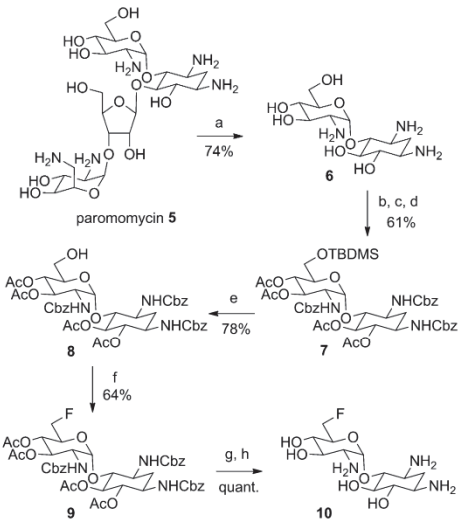



***Figure 2.*** Measurement of the $K_D$ between compound **1** and 16S23 RNA (0.3 mM) by $^{19}F$ NMR spectroscopy in $KPO_4$ buffer pH 6.5 containing 50 mM KCl at 293 K. A) $^{19}F$ chemical shift for increasing concentrations of probe. B) $K_D$ determination based on the $^{19}F$ NMR chemical-shift variation (observation of the probe). C) $K_D$ determination based on $^1H$ NMR chemical-shift variation of the 16S23 $G_{17}$ base (observation of the target).

range of 120 μM, 50 μM, and 35 μM, respectively. The ranking of the $K_D$ values of these molecules is thus in agreement with published data.[10] However, we would have expected a greater difference in the $K_D$ values of neomycin and neamine. The inability of compound **1** to accurately measure this difference might be from too great a difference in dissociation constants between reporter **1** and neomycin.[14]

For the purpose of ranking higher affinity ligands, we prepared a novel fluorinated reporter that could bind in a more selective and efficient manner. Compound **10** was therefore prepared from paromomycin **5** (Scheme 1). This compound showed a sixfold lower dissociation constant for the 16S23 RNA than compound **1**, as determined by $^1H$ NMR titration ($K_D = 300$ μM).[15] However, its improved $K_D$ led to a significant broadening of its $^{19}F$ NMR signal in the presence of 16S23 RNA, as a result of an intermediate exchange regime. Fortunately, a sharp signal could be recovered by increasing the ionic strength of the buffer to 300 mM KCl.[15] Working at a high salt concentration has two interesting consequences: 1) it decreases the $K_D$ of the reporter for the RNA target, and the interaction is back to a fast regime exchange on NMR timescale; 2) the electrostatic part of the binding interaction is lowered, and thus the promiscuous binding that often occurs with an RNA target is limited.[16] Using this reporter **10** under these conditions, we estimated a $K_D$ of 60 μM for neomycin, 90 μM for paromomycine and 250 μM for neamine (Figure 4). This ranking is in agreement with ones determined by other techniques.[10] Interestingly, raising the salt concentration increases $K_D$ values of 16S23 ligands, and tends to increase the differences between good ligands, thus allowing a more precise ranking.[17] Thus, neamime and neomycin, initially ranked very close to each other using reporter **1**, were better differentiated using a tighter 16S23 RNA binder like compound **10** at a high-salt concentration.[7] Thus, two fluorinated probes that present a broad spectrum of interaction with RNA targets can be used to design specific ligands towards an RNA target. In a fragment-based approach, compound **1** is a better suited molecule to perform competition experiments whereas, in optimization assays, compound **10** will be more efficient.

We then investigated the use of fluorinated compounds to monitor RNA conformational changes upon ligand binding (principle from Figure 1B). As a case study, we investigated the interaction of neomycin with a known neomycin aptamer.[18] This aptamer is particularly interesting since it adopts an unbound "on" conformation and a bound "off" conformation in the presence of its ligand.[18] This conformational switch can be visualized by $^1H$ NMR, (Figure 5A,D). Ligand binding induces extensive structural changes, as indicated by the appearance of new imino protons upon addition of neomycin (Figure 5A).[19] We first analyzed the interaction between the aptamer and com-




***Scheme 1.*** Reaction conditions: a) HCl, MeOH, heated to reflux, overnight; b) benzylchloroformate (CbzCl), $Na_2CO_3$, toluene/water/acetone, −5 °C to RT, 4 h; c) *tert*-butyldimethylsilyl chloride (TBDMSCl), 4-dimethylaminopyridine (DMAP), $Et_3N$, dimethylformamide (DMF), −5 °C, 1 h; d) $Ac_2O$, DMAP, pyridine, RT, 1 day; e) HF/pyridine, tetrahydrofuran (THF), RT, 4 h; f) diethylaminosulfur trifluoride (DAST), dry dichloroethane, −5 °C to RT, 5 h; g) NaOMe, MeOH/THF/dioxane, RT, 3 h; h) $H_2$, Pd/C, dioxane/MeOH/water, RT, 2 days.

 © 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 

These are not the final page numbers!

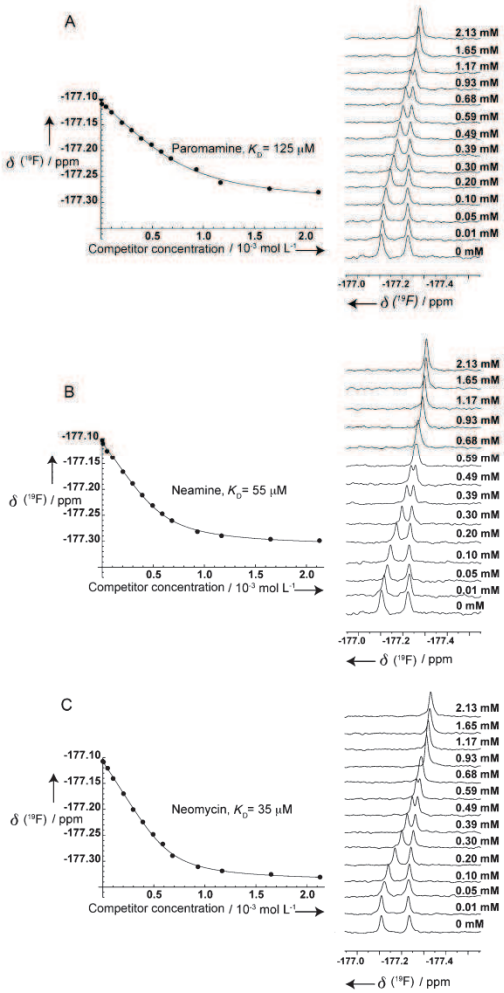



*Figure 3.* Competition experiments using compound **1** as the fluorinated probe. Right: $^{19}$F NMR spectra of compound **1** (0.7 mM) mixed with 16S23 RNA (0.3 mM) in $KPO_4$ buffer pH 6.5 containing 50 mM KCl at 293 K with increasing concentrations of A) paromamine, B) neamine, and C) neomycin.

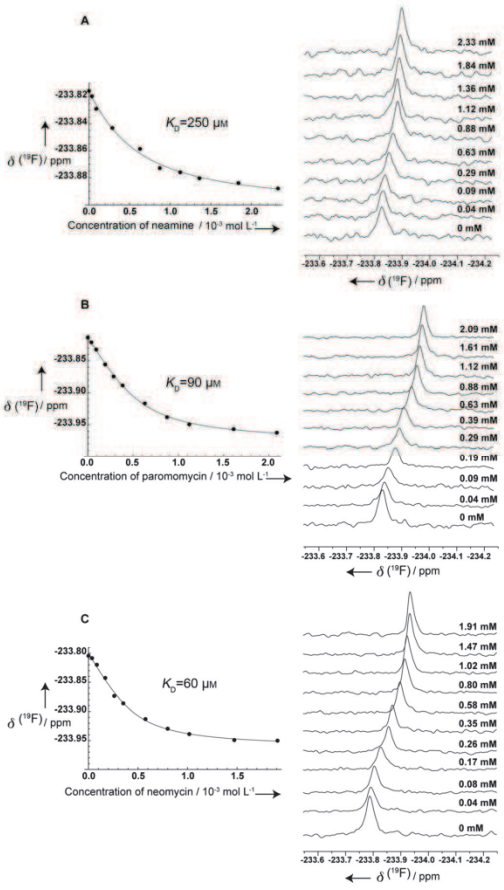



*Figure 4.* Competition experiments with **10** as the fluorinated probe: $^{19}$F NMR spectra of compound **10** (0.7 mM) mixed with 16S23 RNA (0.3 mM) in $KPO_4$ buffer pH 6.5, containing 300 mM KCL, at 293 K, with increasing concentrations of A) neamine, B) paromomycin, and C) neomycin.

pounds **1** and **10** (Figure 5). Both probes bind to the aptamer but with different outcomes (Figure 5B,C). Compound **1** binding does not alter the conformational equilibrium of the RNA target (Figure 5C). Only chemical-shift variations are observed upon addition of compound **1** leading to a $K_D$ value in the millimolar range. This result supports the idea that this external probe will induce only very limited bias in the study of the interaction between ligands and the RNA target. On the contrary, the higher affinity of compound **10** for the aptamer is illustrated by a complete disappearance of

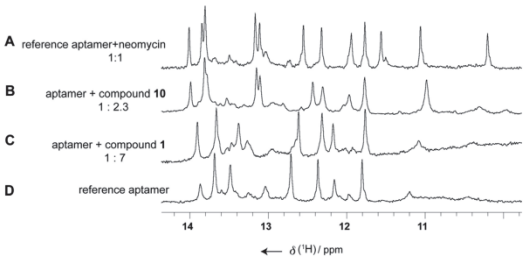



*Figure 5.* $^1$H NMR spectra of neomycin aptamer (0.2 mM) in $KPO_4$ buffer pH 6.5 containing 50 mM KCl, at 293 K with A) one equivalent of neomycin, B) compound **10** (0.46 mM), C) compound **1** (1.4 mM), and D) alone.


© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

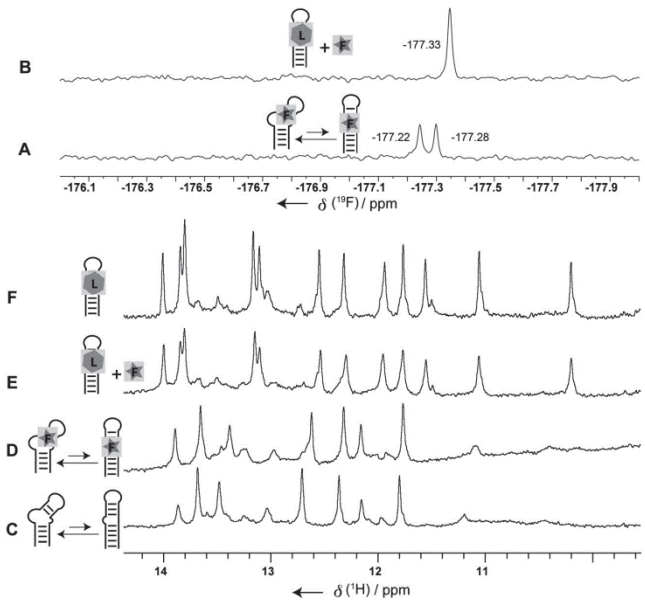



***Figure 6.*** Dynamic folding of neomycin aptamer monitored by NMR spectroscopy using the fluorinated reporter **1**. $^{19}$F NMR spectra of neomycin aptamer (0.3 mM) in a phosphate buffer (10 mM, pH 6.5) with 50 mM KCl and A) compound **1** (0.7 mM), B) compound **1** (0.7 mM) and neomycin (0.45 mM). $^{1}$H NMR spectra of neomycin aptamer (0.3 mM) in a phosphate buffer (10 mM, pH 6.5) with 50 mM KCl, C) alone, D) compound **1** (0.7 mM), E) compound **1** (0.7 mM) and neomycin (0.45 mM), or F) neomycin (0.3 mM).

its $^{19}$F NMR signal upon binding, as a result of an intermediate exchange regime, and the folding of the aptamer into a nearly bound conformation (Figure 5B). These experiments clearly show that a reporter with high affinity for the RNA target should be avoided for monitoring conformationally labile RNA such as riboswitches, since it can dramatically alter their native conformational equilibrium.

Interestingly, the $^{19}$F NMR signal of compound **1** is detectable upon binding to the neomycin aptamer (Figure 6A) and shows signal splitting owing to the binding of each compound **1** enantiomer to the aptamer.[8] A complete $^{19}$F NMR signal coalescence to the $^{19}$F signal of the free compound **1** was observed upon addition of neomycin to the mixture, which shows that the fluorinated probe was fully displaced (Figure 6B). Moreover, $^{1}$H NMR spectroscopy revealed a profound change in the aptamer structure that now shows a neomycin-bound conformation (Figure 6E). These experiments demonstrate that the competitive displacement of a small fluorinated binder can be used to easily monitor the conformational capture of a bi-stable RNA by its ligand.

In conclusion, we have shown that small fluorinated reporters are useful for monitoring the binding of unlabeled RNA with unlabeled ligands. The observation of $^{19}$F NMR chemical-shift variations can be used to qualitatively rank different binders for a structured RNA, such as the 16S23 RNA. Chemical modification of the reporter and tuning of experimental conditions enable the observation of ligands with different binding strengths. Fluorinated reporters can also be used to monitor the conformational capture of a bi-stable RNA upon binding. In this case, a reporter with a low affinity to the target behaves as a neutral partner and does not influence the equilibrium between the bound and unbound states, in contrast to a higher-affinity binder. The absence of any target or ligand labeling step, the 100% natural abundance of the $^{19}$F nucleus, the bio-orthogonality of the carbon–fluorine bond, as well as the high $^{19}$F NMR sensitivity make this method very attractive from a practical standpoint.[20] We believe that $^{19}$F NMR spectroscopy competitive screening will be a simple and efficient method for the discovery of new RNA binders.

Received: May 25, 2012
Published online: ■■ ■■, ■■■■

**Keywords:** aminoglycosides · $^{19}$F NMR spectroscopy · reporter ligands · riboswitches · RNA

[1] a) P. A. Sharp, *Cell* **2009**, *136*, 577–580; b) T. A. Cooper, L. Wan, G. Dreyfuss, *Cell* **2009**, *136*, 777–793; c) M. Esteller, *Nat. Rev. Genet.* **2011**, *12*, 861–874.
[2] L. Guan, M. D. Disney, *ACS Chem. Biol.* **2012**, *7*, 73–86.
[3] J. R. Thomas, P. J. Hergenrother, *Chem. Rev.* **2008**, *108*, 1171–1224.
[4] a) M. Pellecchia, I. Bertini, D. Cowburn, C. Dalvit, E. Giralt, W. Jahnke, T. L. James, S. W. Homas, H. Kessler, C. Luchinat, B. Meyer, H. Oschkinat, J. Peng, H. Schwalbe, G. Siegal, *Nat. Rev. Drug Discovery* **2008**, *7*, 738–745; b) C. Dalvit, *Drug Discovery Today* **2009**, *14*, 1051–1057.
[5] a) J. Almeida Cruz, E. Westhof, *Cell* **2009**, *136*, 604–609; b) J. Parsons, M. P. Castaldi, S. Dutta, S. M. Dibrov, D. L. Wyles, T. Hermann, *Nat. Chem. Biol.* **2009**, *5*, 823–825; c) U. Rieder, C. Kreutz, R. Micura, *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 10804–10809; d) K. E. Deigan, A. R. Ferré-d'Amaré, *Acc. Chem. Res.* **2011**, *44*, 1329–1338.
[6] a) J. Rinnenthal, J. Buck, J. Ferner, A. Wacker, B. Fürtig, H. Schwalbe, *Acc. Chem. Res.* **2011**, *44*, 1292–1301; b) L. Chen, E. Cressina, F. J. Leeper, A. G. Smith, C. Abell, *ACS Chem. Biol.* **2010**, *5*, 355–358.
[7] C. Kreutz, H. Kählig, R. Konrat, R. Micura, *Angew. Chem.* **2006**, *118*, 3528–3532; *Angew. Chem. Int. Ed.* **2006**, *45*, 3450–3453.
[8] R. Moumné, M. Pasco, E. Prost, T. Lecourt, L. Micouin, C. Tisné, *J. Am. Chem. Soc.* **2010**, *132*, 13111–13113.
[9] C. Dalvit, *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51*, 243–271.
[10] a) C.-H. Wong, M. Hendrix, E. S. Priestley, W. A. Greenberg, *Chem. Biol.* **1998**, *5*, 397; b) M. Kaul, C. M. Barbieri, D. S. Pilch, *J. Mol. Biol.* **2005**, *346*, 119–134.
[11] The corresponding non-fluorinated diaminocyclopentanol is a ligand of 16S23 RNA: L. Micouin, F. P. Dardel, C. Tisne-

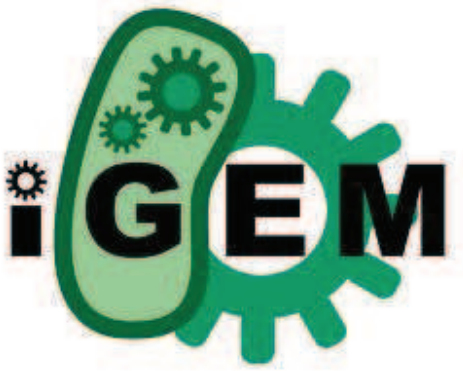
 © 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 

Vicrobeck, F. Maurice, M. Bonin, A. Perez Luna, C. J. Bournaud, G. Bégis, WO2006024784, **2006**.
[12] S. Yoshizawa, D. Fourmy, R. G. Eason, J. D. Puglishi, *Biochemistry* **2002**, *41*, 6263 – 6270.
[13] The binding sites were localized using NOESY experiments. See Supporting Information for details.
[14] X. Zhang, A. Sänger, R. Hemmig, W. Jahnke, *Angew. Chem.* **2009**, *121*, 6819; *Angew. Chem. Int. Ed.* **2009**, *48*, 6691.
[15] See Supporting Information.
[16] An alternative solution with probes undergoing intermediate exchange can be used for the monitoring of the recovery of the signal intensity upon adding of a competing molecule, as recently proposed: A. Vulpetti, U. Hommel, G. Landrum, R. Lewis, C. Dalvit, *J. Am. Chem. Soc.* **2009**, *131*, 12949 – 12959.
[17] The $K_D$ measurements of high affinity binders (neomycin, paromomycin) using imino proton chemical shifts were not possible because of a slow exchange regime, even at 300 mM KCl.
[18] J. E. Weigand, M. Sanchez, E. B. Gunnesch, S. Zeiher, R. Schroeder, B. Suess, *RNA* **2008**, *14*, 89 – 97.
[19] E. Duchardt-Ferner, J. E. Weigand, O. Ohlenschläger, S. R. Schmidtke, B. Suess, J. Wöhnert, *Angew. Chem.* **2010**, *122*, 6352 – 6355; *Angew. Chem. Int. Ed.* **2010**, *49*, 6216 – 6219.
[20] All the $^{19}F$ NMR spectroscopy experiments have been conducted on a 300 MHz (H value) spectrometer with a standard QNP probe. As **1** has been shown to bind to several structurally different RNAs, we believe that it should be an excellent tool to investigate many other RNAs.

 © 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim 

## *F - Compétition IGEM*

Au cours de ma thèse, j'ai eu l'opportunité de participer par deux fois, en 2010 et 2011, à la compétition IGEM (pour « International Genetically Engineered Machine » competition) en tant qu'encadrant de l'équipe parisienne.




Cette compétition internationale de biologie synthétique a vu le jour au MIT à Boston il y a 8 ans. En 2011, elle a regroupé plus de 3000 participants et 170 équipes venant des 5 continents. Son intérêt est qu'elle regroupe des étudiants du premier cycle universitaire venant d'horizons scientifiques très variés autour d'un projet commun : construire, en un été, un système biologique artificiel à partir de composants génétiques standardisés, les « biobricks ». Chaque année, les équipes rivalisent d'ingéniosité pour créer des constructions toujours plus complexes et ludiques à la fois (biosenseurs, calcul amorphe, bioremédiation, impression d'image sur un tapis bactérien…). Pour en savoir plus : http://igem.org/Main_Page

Les deux années que j'ai passée en tant qu'encadrant de l'équipe IGEM de Paris ont été pour moi une aventure unique extrêmement enrichissante, tant sur le plan scientifique qu'humain. J'ai voulu exposer rapidement dans ce manuscrit les deux projets que nous avons menés à bien. Ils ont été récompensés, en 2010, par le prix du meilleur projet de recherche fondamentale et en 2011, par le prix de la meilleure présentation (Europe).

Si vous êtes intéressés, je vous invite fortement à aller voir les sites web de ces deux projets :

http://2010.igem.org/Team:Paris_Liliane_Bettencourt

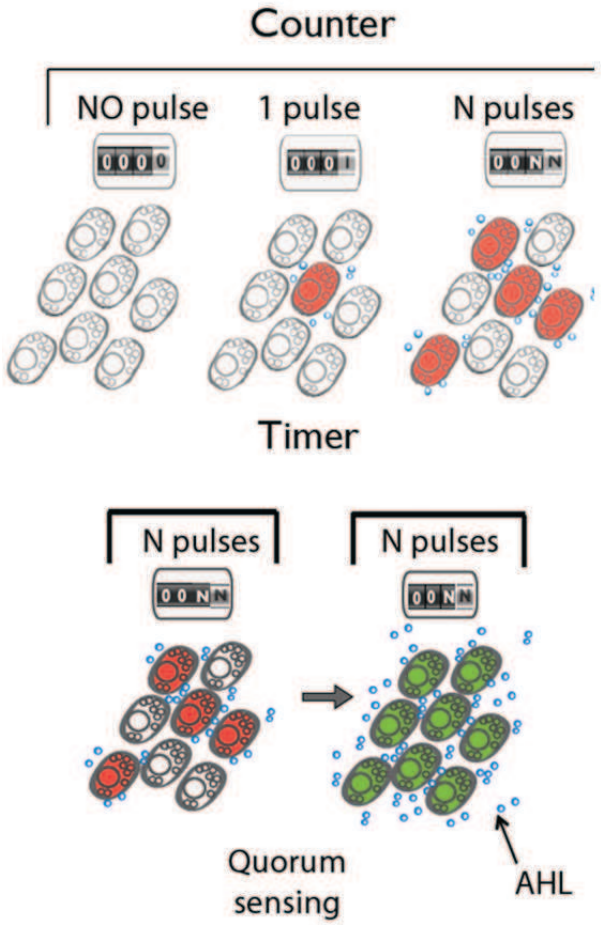
http://2011.igem.org/Team:Paris_Bettencourt

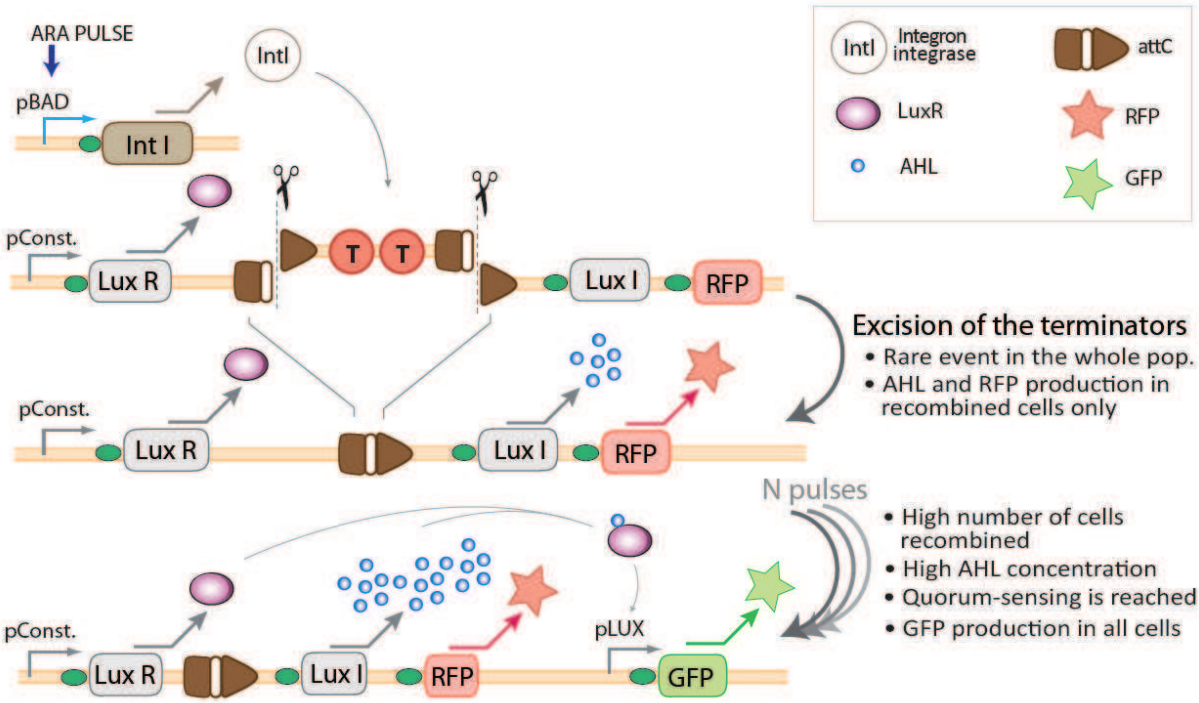
## 1) Projet IGEM Paris 2010 « Every Bacteria Counts »

En 2010, le projet de l'équipe consistait à développer des bactéries permettant de compter l'occurrence d'un événement (pulse d'arabinose) puis de déclencher une réponse à partir d'un seuil donné. Pour ce faire, nous avons développé deux stratégies différentes, une s'appuyant sur une réponse à l'échelle d'une population bactérienne (projet Pop-Counter), l'autre se situant à l'échelle de l'individu (projet Memo-Cell).

**Projet Pop-Counter**

Le projet Pop-Counter reposait l'induction à chaque événement de l'expression de l'intégrase IntI. Celle-ci va alors avoir une certaine probabilité d'exciser une partie de notre construction génétique, ce qui va conduire à un changement phénotypique mesurable. L'observation du phénotype de l'ensemble de la population bactérienne permet de remonter au nombre d'événement qui se sont produits. En détournant le système de quorum sensing, il est par ailleurs possible de déclencher un changement global de phénotype dans la population bactérienne à partir d'un certain nombre d'événements.

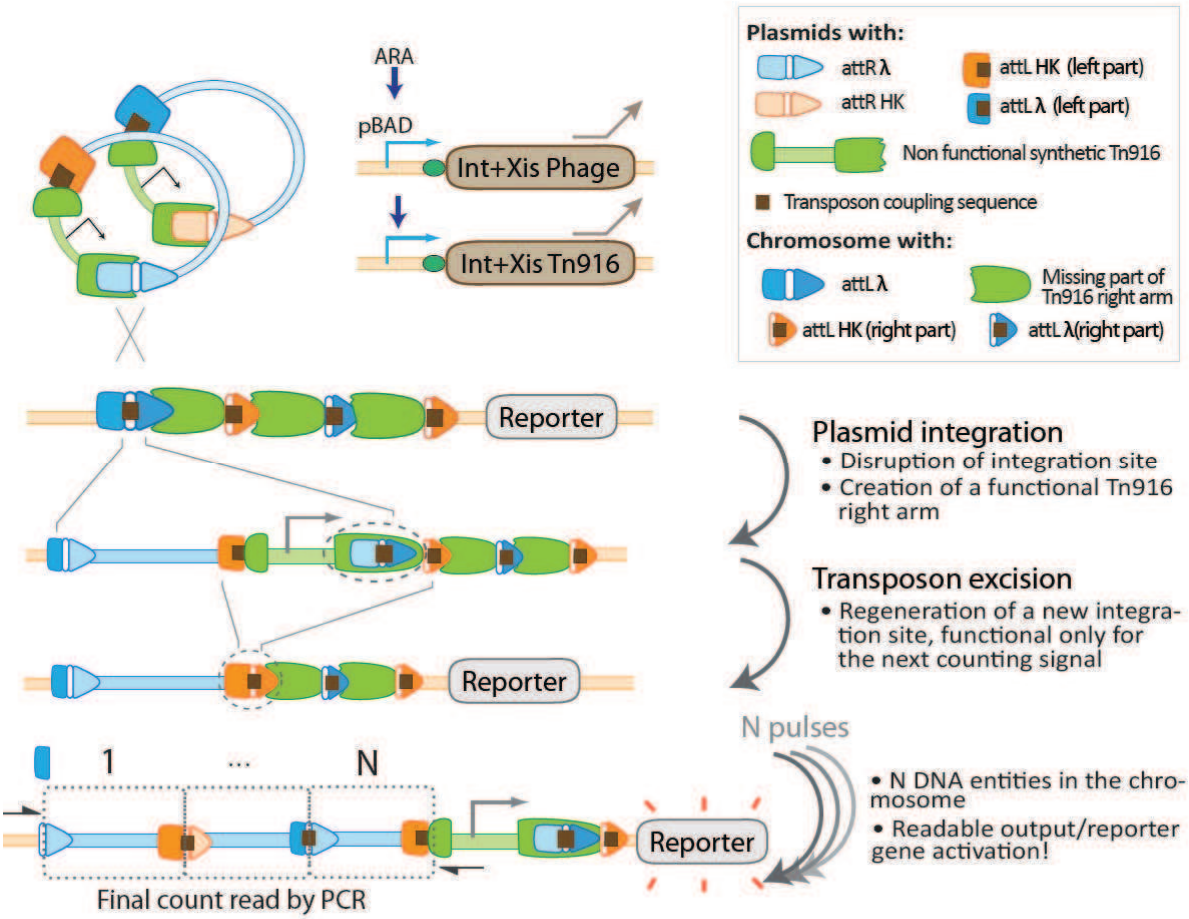





**Principe de fonctionnement du système Pop-Counter**

## Projet Memo-Cell

L'objectif du projet Memo-Cell était d'intégrer dans le chromosome bactérien un fragment d'ADN spécifique à chaque fois qu'un événement était détecté. Ce projet repose sur le détournement des systèmes d'intégration des phages Lambda et HK022 et sur l'ingénierie du système de recombinaison du transposon Tn916. Ce projet a inspiré des travaux qui ont récemment été publiés dans PNAS[376].

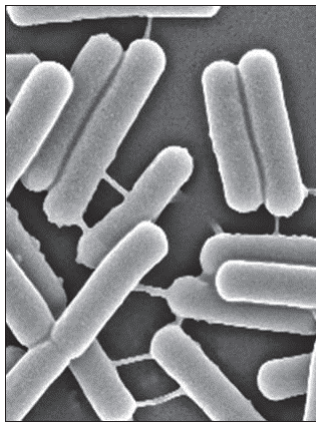



**Principe de fonctionement du système Memo-Cell**

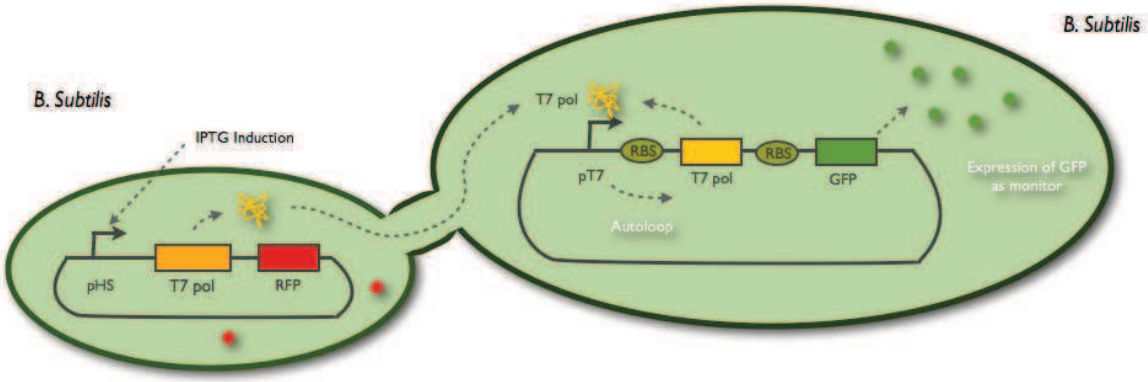
Pour plus d'informations : http://2010.igem.org/Team:Paris_Liliane_Bettencourt

### 2) Projet IGEM Paris 2011 « Tube or not Tube ? Towards harnessing bacterial nanotubes by and for synthetic biology »

En 2011, l'équipe de Ben-Yehuda publie dans Cell un article qui provoque beaucoup de débats dans la communauté scientifique[377]. Cet article décrit la découverture d'un supposé nouveau mode de communication entre bactéries *B.subtillus* : les nanotubes. Le projet de l'équipe IGEM Paris 2011 a été de développer des outils de biologie synthétique pour essayer de mettre en évidence l'existence de ces structures particulières et pour en caractériser le mode de fonctionnement. Nous avons ainsi développé une série d'émetteurs et de récepteurs permettant d'amplifier un signal moléculaire passant à travers ces nanotubes.



**Principe de la détection d'un nanotube**




**Exemple d'un système testé reposant sur l'autoamplification de la polymérase T7**

Pour plus d'informations : http://2011.igem.org/Team:Paris_Bettencourt

# III - Liste des abréviations

| | |
|---|---|
| µL | microlitre |
| 16S | ARNr de la petite sous-unité du ribosome bactérien |
| 30S | petite sous-unité du ribosome bactérien |
| 50S | grande sous-unité du ribosome bactérien |
| Ac | acétyl |
| ACE | bis(acetoxyethoxy)-methyl |
| ADN | Acide desoxyribonucléique |
| Apra | cassete de résistance à l'apramycine |
| aq. | aqueux |
| Ar | aromatique |
| ARN | Acide ribonucléique |
| ARNm | ARN messager |
| ARNnc | ARN non codant |
| ARNr | ARN ribosomique |
| ARNt | Acide ribonucléique de transfert |
| ARNt | ARN de transfert |
| ARNtLys3 | ARNt lysine isoaccepteur 3 |
| ARNv | ARN viral |
| Bn | Benzyle |
| Boc | tert-butyloxycarbonyle |
| C5 | 1,3-cis-diaminocyclopentan-4-ol |
| CatBH | catécholborane |
| Cbz | benzyloxycarbonyle |
| CCM | chromatographie sur couche mince |
| CSP | Chemical Shift Perturbation |
| d | doublet |
| Da | Dalton |
| DACP | cis-1,3-diaminocyclopentan-4-ol |
| DACPF | cis-1,3-diamino-4-fluorocyclopentane |
| DAST | diméthylaminosulfur trifluoride |
| DCE | dichloroéthane |
| DCM | dichlorométhane |
| DMAP | 4-diméthylaminopyridine |
| DMF | diméthylformamide |
| DMSO | diméthylsulfoxyde |
| DNB | dinitrophenyl |
| DO | densité optique |
| DOS | 2-désoxystreptamine |
| DTT | dithiothréitol |
| EBI | European Bioinformatics Institute |
| $EC_{50}$ | Concentration efficace médiane |
| EDTA | acide éthylène-diamine tétra-acétique |
| ee | excès énantiomérique |
| éq. | équivalent |
| ES | Electrospray |

| | |
|---|---|
| EtOAc | acétate d'éthyle |
| FAXS | fluorine chemical shift anisotropy and exchange screening |
| Fmoc | fluorenylmethyloxycarbonyl |
| g | gramme |
| GFP | Green Fluorescent Protein |
| GST | glutathion-S-transférase |
| h | heure |
| HPLC | chromatographie liquide haute pression |
| HRMS | spectroscopie de masse haute résolution |
| HSQC | Heteronuclear Single Quantum Coherence |
| HTS | Criblage haut débit ou "High Throughput Screening" |
| hyg ou Hygro | cassette de résistance à l'hygromycine |
| Hz | Hertz |
| IPTG | isopropyl β-D-1-thiogalactopyranoside |
| IR | Infrarouge |
| IRES | Internal Ribosome Entry Site |
| Ka | constante d'acidité |
| Kd | constante de dissociation |
| Ki | constante d'inhibition |
| L-LDH | lactate déshydrogénase |
| m | multiplet ou massif |
| M | masse molaire |
| MALDI | Désorption-ionisation laser assistée par matrice |
| Me | Méthyle |
| mL | millilitre |
| mM | millimolaire |
| mol. | mole |
| MS | spectroscopie de masse |
| NAD | nicotinamide adénine dinucléotide |
| Ni-NTA | colonne de nickel |
| NMO | oxyde de N-méthylmorpholine |
| nOe | nuclear Overhauser effect |
| NOESY | Nuclear Overhauser Effect SpectroscopY |
| NRPS | Synthase de peptide non ribosomique |
| Nu | nucléophile |
| paroF | paromamine fluorée en position 6' |
| paroN3 | paromamine azidée en position 6' |
| PBS | Primer Binding Site |
| PCA | analyse en composantes principales |
| PCR | Polymerase Chain Reaction |
| PDB | Precursor-Directed Biosynthesis |
| PDB | Protein Data Bank |
| PEG | polyéthylène glycol |
| Ph | Phényle |
| pH | potentiel hydrogène |
| PKS | Synthase de polycétide |
| PLP | pyridoxal phosphate |
| PMF | Peptide mass fingerprinting |

| | |
|---|---|
| PMSF | flurorure de phenylméthanesulfonyle |
| ppm | partie par million |
| PTC | premature codon terminaison |
| Pyr. Kin. | pyruvate kinase |
| q | quadruplet |
| q.s.p. | quantité suffisante pour |
| rac. | racémique |
| rdt | Rendement |
| Rf | rapport frontal |
| RMN | résonance magnétique nucléaire |
| RNase | Ribonucléase |
| rpm | rotation par minute |
| RRE | Rev Response Element |
| RT | Transcriptase Inverse |
| s | singulet |
| SDS | sodium dodecyl sulfate |
| sec | seconde |
| SELEX | Systematic Evolution of Ligands by Exponential Enrichment |
| siARN | petits ARN interférents |
| SPE | Solid Phase Extraction |
| SPR | résonnance plasmonique de surface |
| STD | Saturation Transfert Difference |
| t | triplet |
| T.A. | température ambiante |
| tampon TE | tampon Tris EDTA |
| TAR | “trans-activating response” element |
| TBAF | tetra-n-butylammonium fluoride |
| TBDMSCl | tert-Butyldimethylsilyl chloride |
| TFE | trifluoroéthanol |
| THF | tétrahydrofurane |
| TMS | tetraméthylsilane |
| TOCSY | TOtal Correlation SpectroscopY |
| $t_R$ | temps de rétention |
| Tris | Tris(hydroxyméthyl)aminométhane |
| TROSY | Transverse Relaxation Optimized SpectroscopY |
| Ts | tosyle ou para-toluènesulfonyle |
| tsr | thiostrepton |
| UDP | uridine diphosphate |
| UDP-GlucNAc | UDP-N-acetylglucosamine |
| UV | Ultra Violet |
| VIH | Virus de l'Immunodéficience Humaine |
| $\alpha_D$ | pouvoir rotatoire |

# IV - Liste des figures

Figure I-1 Structure primaire des ARN ..... 20

Figure I-2 Modification post transcriptionelle des ARN : synthèse de la pseudouridine..... 21

Figure I-3 Niveaux d'organisation de la structure des ARN ..... 21

Figure I-4 Exemples d'appariement des bases de l'ARN..... 22

Figure I-5 Comparaison structurale d'une hélice d'ARN avec une hélice d'ADN..... 23

Figure I-6 Variété des motifs structuraux secondaires de l'ARN ..... 24

Figure I-7 Représentation schématique de la transcription et de la traduction..... 26

Figure I-8 Le ratio séquences génomiques non-codantes/séquences codantes ..... 27

Figure I-9 Exemple de fonctionnement d'un riboswitch..... 28

Figure I-10 Le principe de la méthode SELEX ..... 29

Figure I-11 Structure du ribosome procaryote. ..... 31

Figure I-12 Rétrocontrole négatif de la glucosamine 6 phosphate sur la transcription de l'enzyme GlmS à travers un ribozyme [51] ..... 33

Figure I-13 L'initiation de la transcription inverse du VIH ..... 35

Figure I-14 Le principe de la capture conformationelle d'un ARN par son ligand. ..... 40

Figure I-15 Une colonie de *Streptomyces*(Streptomyces sp. Acta 1362) ..... 43

Figure I-16 Structure de différents aminoglycosides ..... 46

Figure I-17 L'ARN 16S d'*E.coli*. ..... 50

Figure I-18 Mécanisme d'action biologique des aminoglycosides..... 52

Figure I-19 Interaction d'aminoglycosides avec le site A de l'ARN 16S ..... 53

Figure I-20 Les principaux mécanismes de résistance aux aminoglycosides ..... 54

Figure I-21 Inactivation des aminoglycosides par l'enzyme de résistance AAC(6')-Ib11..... 55

Figure I-22 Principe général de la translecture induite par les aminoglycosides. ..... 57

Figure I-23 Principe du traitement de la dystrophie myotonique avec un ligand d'ARN ..... 59

Figure I-24 Un aperçu de la biodiversité des actinomycètes ..... 64

Figure I-25 Principe de la métagénomique ..... 68

Figure I-26 Hydrolyse sélective de la néomycine pour obtenir la DOS ou la néamine. ..... 72

Figure I-27 Stratégie de glycodiversification pour obtenir des analogues de kanamycine. ..... 73

Figure I-28 Un exemple de guanidinoglycoside, la kanamycine guanidinylée..... 74

Figure I-29 Structures de l'amikacine et l'arbekacine ..... 75

Figure I-30 Analogues de néamine développés par Mobashery..... 76

Figure I-31 Mécanisme de régénération d'un analogue de kanamycine..... 77

Figure I-32 Analogues de paromomycine développés par Hanessian ..... 78

Figure I-33 Analogues d'aminoglycosides de type spiro-ether ..... 78

Figure I-34 Hybrides d'aminoglycosides obtenus par ajout de cycle glycosidique ..... 79

Figure I-35 Etude structurale de la conformation de la néomycine en complexe avec le site A du ribosome ou l'ARN viral TAR et design d'un analogue contraint ..... 80

Figure I-36 Hybrides aminoglycosides/ base nucléique ..... 82

Figure I-37 Hybrides aminoglycosides/acridine ..... 83

Figure I-38 Hybride neomycine /cipofloxacine ..... 84

Figure I-39 Conjugué néamine-PNA développé par Décout ..... 85

Figure I-40 Stratégie de synthèse de ligand par fragment ..... 86

Figure I-41 Exemples de dimères de néamine synthétisés par Wong ..... 87

Figure I-42 Deux dimères de deoxystreptamine ..... 88

Figure I-43 Synthèse par fragment d'inhibiteur de l'enzyme de résistance AAC(6')-Ib11.). ..... 89

Figure I-44 Dimères de mimes simplifiés d'aminoglycosides ..... 90

Figure I-45 Exemple d'analogue d'aminoglycoside obtenu par réaction multicomposant ..... 90

Figure I-46 Approche chemoenzymatique pour la synthèse d'analogues d'aminoglycosides ..... 91

Figure I-47 Principe des stratégies de PDB et de mutasynthèse ..... 93

Figure I-48 Mutasynthèse de la doramectine ..... 96

Figure I-49 Sites de modification de la balhimycine permis par mutasynthèse ..... 97

Figure I-50 Stratégie de mutasynthèse pour la conception d'analogues de la Wailupemycine.... 98

Figure I-51 Approches pour le développement d'analogues d'aminoglycosides ..... 99

Figure I-52 Représentation schématique du projet de thèse ..... 100

Figure II-1 Structure de la DOS et de différents aminoglycosides ..... 105

Figure II-2 Voie de biosynthèse de la néamine à partir du glucose-6-phosphate ..... 108

Figure II-3 Amplification par PCR des gènes neo8, neo11, neo16, neo18. ..... 113

Figure II-4 Test d'expression de la protéine Neo8 dans la souche BL21 (DE3).. ..... 114

Figure II-5 Test d'expression de la protéine Neo11 dans la souche BL21 (DE3) ..... 114

Figure II-6 Purification par colonne d'affinité de la protéine Neo8 sous forme dénaturée ..... 116

Figure II-7 Test enzymatique de l'activité glycosyltransférase Neo8..... 118

Figure II-8 Purification des anticorps polyclonaux anti-Neo8. ..... 120

Figure II-9 Western blot utilisant les anticorps anti-Neo8 purifiés. ..... 121

Figure II-10 Western blot des fractions d'une chromatographie par exclusion de taille sur Superdex 200 d'un lysat cellulaire de *S.fradiae* en conditions natives ..... 122

Figure II-11 Protocole de purification de la protéine chimère GST-GenM1 ..... 125

Figure II-12 Digestion de la protéine GST-GenM1 par la thrombine ..... 126

Figure II-13 Expression et purification de la protéine NusA-GenM1 ..... 127

Figure II-14 Essais de digestion de la protéine NusA-GenM1. ..... 128

Figure II-15 Surexpression de la protéine chimère AAC(6')-Ib-GenM1 ..... 129

Figure II-16 Modèle structural de la protéine GenM1* ..... 131

Figure II-17 Surexpression et purification de la protéine GenM1*..... 133

Figure II-18 Surexpression et purification de la déactetylase BtrD..... 134

Figure III-1 Premiers essais de mutasynthèse sur le mutant de *S.fradiae* isolée par Rinehart.... 143

Figure III-2 Exemples d'analogues d'aminoglycosides obtenus par mutasynthèse..... 146

Figure III-3 L'opéron de biosynthèse de la néomycine chez *S.fradiae*..... 148

Figure III-4 Synthèse de la DOS à partir de glucose-6-phosphate..... 149

Figure III-5 Stratégie de délétion d'un gène par double recombinaison ..... 150

Figure III-6 Synthèse d'analogues de DOS par désymétrisation d'hydrazine bicycliques..... 151

Figure III-7 Synthèse d'analogues de paromamine modifiés en position 6' ..... 151

Figure III-8 Mutasynthèse d'analogues de néomycine : les différentes étapes de l'étude ..... 152

Figure III-9 Principe de la suppression de neo6 par double recombinaison homologue ..... 153

Figure III-10 Stratégie de clonage du plasmide pFL02-neoApra ..... 156

Figure III-11 Vérification par PCR de l'événement de double recombinaison. ..... 158

Figure III-12 Antibiogramme comparé de la souche Δneo6 avec la souche *S.fradiae* sauvage).. 159

Figure III-13 Greffage des billes de polymère des colonnes SPE Strata X-CW ..... 161

Figure III-14 Spectre de masse du sulfate de néomycine obtenu par infusion..... 162

Figure III-15 Caractérisation en HPLC-MS de la néomycine purifiée à partir de TSB..... 162

Figure III-16 Evaluation de la perte de néomycine après purification SPE ..... 163

Figure III-17 Calibration par HPLC-MS de la concentration en néomycine dans le TSB..... 164

Figure III-18 Production en antibiotique de la souche *S.fradiae* dans différents milieux..... 165

Figure III-19 Cinétique de production de la néomycine par *S.fradiae* dans le TSB. ..... 166

Figure III-20 Synthèse de la DOS par dégradation de la néomycine..... 168

Figure III-21 Synthèse de la paromamine par dégradation sélective de la paromomycine ..... 168

Figure III-22 Synthèse de la néamine par dégradation sélective de la néomycine..... 169

Figure III-23 Synthèse de streptidine et de streptamine par dégradation de la streptomycine . 170

Figure III-24 Synthèse du *cis*-1,3-diaminocyclopent-4,5-diol 79 (C5) ..... 171

Figure III-25 Synthèse de la paromamine avec un azoture en position 6' (paro$N_3$) ..... 172

Figure III-26 Synthèse de la paromamine fluorée en position 6' (paroF) ..... 173

Figure III-27 Production de néomycine de S.fradiae et des souches Δneo6 et DSM41550..... 174

Figure III-28 Voie de biosynthèse de la néomycine chez *S.fradiae* ..... 175

Figure III-29 Complémentation chimique de DSM 41550 et Δneo6 par la DOS ..... 176

Figure III-30 Influence de la concentration de divers métabolites sur la production d'aminoglycoside par la souche *S.fradiae* sauvage ..... 177

Figure III-31 Spectre HPLC-MS et structure chimique de la paromomycine..... 178

Figure III-32 Complémentation de DSM 41550 et Δneo6 par la néamine et la ribostamycine ... 181

Figure III-33 Complémentation de DSM 41550 et Δneo6 par la paromamine ..... 182

Figure III-34 Spectres HPLC-MS des expériences de complémentation de DSM 41550 et Δneo6 avec la paromamine ..... 182

Figure III-35 Les analogues de DOS testés comme mutasynthons ..... 184

Figure III-36 Complémentation de Δneo6 avec le mutasynthon C5 ..... 185

Figure III-37 Fragmentation comparée du pic à m/z = 293. ..... 186

Figure III-38 Complémentation de Δneo6 avec le mutasynthon 90 ..... 188

Figure III-39 Exemple d'hybride kanamycine/ribostamycine développé par Assensio ..... 189

Figure III-40 Analogues de paromamine synthétisés pouvant servir de mutasynthons..... 190

Figure IV-1 Exemple de spectre RMN $^{1}$H d'un ARN..... 196

Figure IV-2 Principe du criblage par RMN $^{19}$F en utilisant une molécule rapportrice fluorée. .... 199

Figure IV-3 Exemple de nucléotides fluorés utilisés dans les stratégies de marquage d'ARN .... 200

Figure IV-4 Suivi de la dénaturation d'une pseudo tige boucle d'ARN par RMN $^{19}$F ..... 201

Figure IV-5 Principe de l'utilisation d'ARN fluorés dans le criblage de nouveaux ligands ..... 202

Figure IV-6 Séquence de l'ARN utilisé par Micura et spectre RMN du fluor associé. ..... 202

Figure IV-7 Principe de l'utilisation d'une sonde fluorée externe dans l'étude des ARN ..... 204

Figure IV-8 L'ARNt$^{Lys}_{3}$..... 205

Figure IV-9 Le cycle diaminocyclopentanol DACP, un mime structural de DOS ..... 206

Figure IV-10 Le DACPF, analogue fluoré de la DOS..... 207

Figure IV-11 Synthèse du DACPF en mélange racémique..... 208

Figure IV-12 Mécanisme de la fluoration de l'hydrazine bicyclique hydroxylée ..... 208

Figure IV-13 Synthèse énantiosélective du (+)-DACPF..... 209

Figure IV-14 Empreinte TROSY de l'ARNt$^{Lys}_{3}$ seul et de l'ARNt$^{Lys}_{3}$ en présence du DACPF. ..... 210

Figure IV-15 Suivi en RMN $^{19}$F de l'interaction du DACPF avec l'ARNt$^{Lys}_{3}$ ..... 211

Figure IV-16 Suivi par RMN $^{19}F$ de l'interaction des énantiomères du DACPF avec l'ARNt$^{Lys}_3$..... 213

Figure IV-17 Influence de la stœchiométrie ARN/DACPF sur le dédoublement du signal..... 214

Figure IV-18 Interaction du DACPF avec divers ARN..... 215

Figure IV-19 Suivi de la dénaturation du bras T de l'ARNt$^{Lys}_3$ en RMN $^{19}F$ ..... 216

Figure IV-20 Structure de l'aptamère à la néomycine étudié..... 218

Figure IV-21 Suivi en RMN $^{1}H$ des variations des signaux des protons imino de l'aptamère lors de la fixation de la néomycine ..... 219

Figure IV-22 Suivi par RMN $^{1}H$ et $^{19}F$ de l'interaction du DACPF avec l'aptamère..... 220

Figure IV-23 Suivi en RMN $^{1}H$ et $^{19}F$ de la capture conformationnelle de l'aptamère par la néomycine en présence de DACPF..... 221

Figure IV-24 Structure secondaire de l'ARN 16S23..... 223

Figure IV-25 Quelques ligands aminoglycosidiques de l'ARN 16S23 classés par affinité. ..... 224

Figure IV-26 Principe de la mesure par RMN $^{19}F$ de l'affinité d'un ligand pour une cible ARN par compétition avec une sonde fluorée ..... 225

Figure IV-27 Suivi en RMN $^{1}H$ de la variation des protons imino des nucléotides $U_{16}$ et $G_{17}$ de l'ARN 16S23 en fonction de la concentration en DACPF..... 226

Figure IV-28 Suivi en RMN $^{19}F$ de la variation du signal du fluor du DACPF en fonction de sa concentration ..... 226

Figure IV-29 Tableau d'attribution des signaux RMN de l'ARN 16S23..... 228

Figure IV-30 Superposition du spectre TOCSY de l'ARN 16S23 seul ou avec le DACPF ..... 229

Figure IV-31 Etude des perturbations des déplacements chimiques des différents nucléotides de l'ARN 16S23 en fonction de la concentration en DACPF..... 230

Figure IV-32 Exemple de la fixation compétitive de la néamine sur le 16S23 en présence de DACPF ..... 231

Figure IV-33 Comparaison de la paromamine, de la néamine et de la néomycine. ..... 233

Figure IV-34 La sonde paroF, un analogue fluoré de la paromamine..... 236

Figure IV-35 Stratégie de synthèse de la paromamine fluorée paroF ..... 237

Figure IV-36 Mesure de l'affinité de la paromamine fluorée pour l'ARN 16S23 par RMN $^{1}H$ ..... 239

Figure IV-37 Effet de la concentration en sels sur la finesse du pic du fluor de paroF ..... 240

Figure IV-38 Compétition avec la sonde paroF ..... 242

Figure IV-39 L'utilisation de sondes fluorées dans l'étude des ARN ..... 245

Figure IV-40 L'initiation de la transcription inverse du VIH ..... 247

Figure IV-41 Méthodologie de suivi de l'activité d'une hélicase à ARN par FRET ..... 248

Figure IV-42 Utilisation d'un aptamère à la néomycine en tant que régulateur traductionnel. . 248

Figure IV-43 Suivi des protons de l'aptamère à la néomycine en RMN $^{1}$H.. ..... 249

Figure IV-44 La kynuramine et les analogues fluorés pouvant servir de sonde en RMN $^{19}$F. ..... 250

Figure S-1 Les différentes phases du cycle de développement de *S.fradiae* ..... 258

Figure S-2 L'ARN 16S23 et l'aptamère à la néomycine ..... 275

Figure S-3 déprotection des groupements ACE ..... 276

Figure S-4 L' ARNt$^{Lys}_{3}$ ..... 277

Figure S-5 Mesure de l'affinité du DACPF pour l'ARN 16S23 par RMN $^{1}$H. ..... 278

Figure S-6 Mesure de l'affinité du DACPF pour l'ARN 16S23 par RMN $^{19}$F ..... 279

Figure S-7 Mesure de l'affinité du DACPF pour l'aptamère à la néomycine par RMN $^{1}$H ..... 280

Figure S-8 Mesure de l'affinité de la paromamine fluorée pour l'ARN 16S23 par RMN $^{1}$H. ..... 281

Figure S-9 Analyse CSP des sites d'interaction du DACPF Avec l'ARN 16S23 ..... 282

Figure S-10 Interface de l'algorithme développé dans Mathematica ..... 285