



**HAL**  
open science

# Synthèse et étude de nouveaux ligands d'ARN pour le ciblage des systèmes toxine-antitoxine : vers des nouvelles thérapies antibiotiques

Céline Martin

► **To cite this version:**

Céline Martin. Synthèse et étude de nouveaux ligands d'ARN pour le ciblage des systèmes toxine-antitoxine : vers des nouvelles thérapies antibiotiques. Chimie thérapeutique. Université Côte d'Azur, 2021. Français. NNT : 2021COAZ4051 . tel-03767527

**HAL Id: tel-03767527**

**<https://theses.hal.science/tel-03767527v1>**

Submitted on 2 Sep 2022

**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.



$$\rho \left( \frac{\partial v}{\partial t} + v \cdot \nabla v \right) = -\nabla p + \nabla \cdot T + f$$

$$e^{i\pi} + 1 = 0$$

# THÈSE DE DOCTORAT

Synthèse et étude de nouveaux ligands  
d'ARN pour le ciblage des systèmes  
Toxine-Antitoxine : vers des nouvelles  
thérapies antibiotiques

**Céline MARTIN**

Institut de Chimie de Nice (ICN)

Devant le jury, composé de :

Présentée en vue de l'obtention  
du grade de docteur en Chimie  
de l'Université Côte d'Azur

Dirigée par : Dr. Maria DUCA  
Co-encadrée par : Dr. Audrey DI  
GIORGIO

Soutenue le : 20 Juillet 2021

**Dr. Muriel AMBLARD**, Directrice de Recherche CNRS,  
Université de Montpellier

**Dr. Fabien DARFEUILLE**, Directeur de Recherche  
INSERM,

Université de Bordeaux

**Dr. Audrey DI GIORGIO**, Maître de Conférences,  
Université Côte d'Azur

**Dr. Boris VAUZEILLES**, Directeur de Recherche CNRS,  
Université Paris-Saclay



# Synthèse et étude de nouveaux ligands d'ARN pour le ciblage des systèmes Toxine-Antitoxine : vers des nouvelles thérapies antibiotiques

## **Jury :**

### Rapporteurs

**Dr. Muriel AMBLARD**, Directrice de Recherche CNRS, Institut des Biomolécules Max Mousseron, Université de Montpellier

**Dr. Boris VAUZEILLES**, Directeur de Recherche CNRS, Institut de Chimie des Substances Naturelles, Université Paris-Saclay

### Examineurs

**Dr. Fabien DARFEUILLE**, Directeur de Recherche INSERM, Acides nucléiques : Régulations Naturelles et Artificielles, Université de Bordeaux

**Dr. Audrey DI GIORGIO**, Maître de Conférences, Institut de Chimie de Nice, Université Côte d'Azur

### Directrice de thèse

**Dr. Maria DUCA**, Directrice de Recherche CNRS, Institut de Chimie de Nice, Université Côte d'Azur



## SYNTHÈSE ET ÉTUDE DE NOUVEAUX LIGANDS D'ARN POUR LE CIBLAGE DES SYSTÈMES TOXINE-ANTITOXINE : VERS DES NOUVELLES THÉRAPIES ANTIBIOTIQUES

La découverte de nouveaux antibiotiques est un défi important pour la chimie médicinale actuelle. En raison de la multirésistance des bactéries aux traitements antibiotiques actuels, il devient urgent de découvrir de nouvelles cibles pour des stratégies antibiotiques novatrices. Les systèmes toxine-antitoxine (TA) chez les bactéries sont de petits éléments génétiques composés d'un gène de toxine et de son antitoxine qui jouent un rôle central dans la survie des cellules bactériennes dans des conditions de stress. Les récentes études sur la capacité des systèmes TA à réguler la mort cellulaire bactérienne en font un point de départ pour le développement d'agents antimicrobiens puissants et nouveaux. Dans ce manuscrit, nous nous concentrons sur les systèmes TA de type I, constitués d'un ARN antitoxine interagissant avec l'ARN messager (ARNm) codant pour la toxine chez la bactérie *Helicobacter pylori*. Ce système TA pourrait représenter une nouvelle cible antimicrobienne dans le sens où des molécules capables d'inhiber l'interaction entre l'ARN antitoxine et l'ARN messager codant pour la toxine, pourraient activer la synthèse de la toxine par traduction de l'ARNm et conduire à la mort des cellules bactériennes. Dans le but de valider ce système TA en tant que cible de choix, nous avons préparé des séries de molécules dites ligands d'ARN conçues pour interagir avec les ARN du système TA de type I. Ce travail décrit la conception et la synthèse des composés ainsi que leurs évaluations biochimiques ( $K_D$ , inhibition du complexe) et biologiques (activité cellulaire antibactérienne). La validation de cette stratégie pourrait permettre de développer de nouveaux antibiotiques et de s'attaquer à un problème thérapeutique majeur actuel : la multirésistance des bactéries aux traitements antibiotiques connus.

**Mots clés :** Systèmes Toxine-antitoxine, Réaction d'UGI, Ligands d'ARN, Analogues de nucléobases, Aminoglycosides

## SYNTHESIS AND STUDY OF NEW RNA LIGANDS FOR THE TARGETING OF TOXIN-ANTITOXIN SYSTEMS: TOWARDS NEW ANTIBIOTIC THERAPIES

The discovery of new antibiotics is an important challenge in current medicinal chemistry. Due to the increasing multidrug resistance to current antibacterial therapies, it is urgent to discover new targets for innovative antibiotic strategies. Bacterial toxin-antitoxin (TA) systems are small genetic elements composed of a toxin gene and its cognate antitoxin that play a pivotal role in the survival of bacterial cells under stress conditions. Recent studies on the ability of TA systems to regulate bacterial cell death make them a starting point for the development of potent and new antimicrobial agents. In this manuscript, we focus on type I TA systems that consist of an antitoxin RNA interacting with the toxin-encoding messenger RNA (mRNA) in *Helicobacter pylori* bacteria. This TA system could represent a new target for molecules able to inhibit the complex formation between toxin and antitoxin RNAs could activate the toxin synthesis by mRNA translation leading thus the death of the bacterial cells. To validate this TA system as a target of choice, we prepared several series of RNA ligands designed to interact with the TA type I RNAs. This work describes their design and synthesis as well as their biochemical ( $K_D$ , complex inhibition) and biological (antibacterial cellular activity) evaluations. Validation of this promising strategy could help to develop new antibiotics to overcome a current major therapeutic issue: multidrug resistance of bacteria to known antibiotic treatments.

**Keywords:** Toxin-Antitoxin systems, UGI reaction, RNA ligands, Aminoglycosides, Nucleobase analogues



## Remerciements

Je souhaite d'abord remercier les Dr. Muriel Amblard et Dr. Boris Vauzeilles d'avoir accepté de juger ce travail en tant que rapporteurs.

Un grand merci au Dr. Fabien Darfeuille pour toutes ses discussions scientifiques sur le passionnant sujet qu'est le système toxine-antitoxine pendant ces trois années. Merci de m'avoir accueillie au sein de son laboratoire pendant une semaine de découverte qui a été très enrichissante pour une chimiste. Je remercie aussi son équipe pour sa gentillesse et ses conseils durant cette semaine et également l'évaluation des molécules.

J'exprime toute ma gratitude à ma directrice de thèse, Maria, de m'avoir donné l'occasion de découvrir et de travailler dans l'univers des ligands d'ARN à travers multiples projets. Je me souviens encore il y a quatre ans, de la fameuse question : pourquoi pas une thèse ? Quelques mois ont suffi pour que je me lance et je ne regrette pas. Merci, de m'avoir fait confiance pour réaliser ce projet et de m'avoir toujours encouragée et poussée pour en arriver là où j'en suis. Certes pas toujours facile mais je n'ai pas de mots pour décrire à quel point je suis reconnaissante.

Un grand merci à Audrey, pour toutes tes discussions scientifiques, pour ta rigueur, pour ta disponibilité, pour ton écoute. On n'oubliera pas la fameuse danse des Boc qui nous a bien fait réfléchir. Merci pour ton regard sans faille dans l'analyse des RMN et tant d'autres détails scientifiques.

Merci à toutes les deux, pour vos qualités humaines, votre soutien pour tous les aspects de la vie quotidienne, et notamment pour cette grande ligne droite qu'a été la rédaction.

Merci à Nadia, pour ton énergie débordante, tous ces moments de partage, tes conseils, ton écoute toujours au bon moment, mais aussi tes bisous, tes pépettes, tes rires que j'espère recroiser encore lors d'une rando en été comme en hiver dans nos montagnes. Stéphane, merci pour les nombreux échanges que nous avons pu avoir, tes conseils techniques pour le labo, ton avis précieux sur l'anisotropie mais également autour des moments conviviaux avec tes gâteaux et les repas d'équipe.

L'équipe TNA, c'est aussi plein de belles rencontres avec pleins de bons moments de partage, de rigolades, de folie, d'énervements (les racks), de chants (Dance monkey) ou de goûters (16h précis) qui ont rythmé le quotidien du laboratoire mais également les apéros (jamais un lundi 17), les raclettes, les randos et le ski (la piste rouge). Pleins de souvenirs mais surtout un grand merci à Maurinne, Sandra, Mélanie, Eleonora, Marc et Sylvain d'avoir partagé votre joie et votre bonne humeur ces derniers mois sans oublier évidemment Carla et Carlo! Je tiens particulièrement à remercier Marc, ça a été un plaisir de partager la paillasse avec toi, avec ces nombreux échanges scientifiques, techniques ou juste de la vie de tous les jours, fière de voir plein de détails qui ne se perdent pas. Merci de m'avoir également



supportée tous les jours mais sache que le Paic doit rester à gauche sur le lavabo ! Et merci Momo pour ton aide précieuse et ton soutien de ce dernier mois mais aussi pour tous ces moments à côté de toi au bureau et au labo. Une pensée également pour les stagiaires ou les anciens qui ont partagé beaucoup de bons moments avec nous, Bénédicte, Naila, Paula, Alessandro, Anaïs, Benjamin.

Une thèse, ce sont aussi des liens forts avec un binôme de rêve, Chloé, une complémentarité parfaite qui nous a permis d'affronter ce travail avec un soutien mutuel. Jamais en compétition, mais toujours là pour apprendre l'une de l'autre. Ce fût un plaisir de pouvoir partager tant de moments, d'apprécier ta façon de penser ou de travailler (tes petites tâches CCM), ta fascination pour les cristaux (d'éléphant), mais aussi nos excursions en congrès et surtout nos moments de pétage de câbles. Je ne saurai oublier Lou (un mot merci), déjà quatre ans, le temps passe vite, ta joie de vie, tes moments de défouloirs, tes petites attentions, tes cheveux, nos soirées pizzas et tant de choses. Merci mes Spies, de m'avoir supportée avec mes humeurs, mes défauts (oups le chocolat), les moments café mais aussi de m'avoir soutenue même à distance avec nos petits appels. Les Totaly Spies, c'est aussi Guigui où j'ai pu compter sur toi. Merci à vous trois, pour ces trois belles années de rire et de partage sans fin. La thèse c'est trois années, mais l'amitié qu'on a créée continuera encore bien plus longtemps !

J'ai également une pensée pour toutes les personnes que j'ai pu rencontrer au sien de l'ICN, de tous les étages, merci pour ces moments de convivialité qui ont rythmé ces trois années. Merci notamment à Hortense, Sarah, Maïlys, Paula, Steve pour votre gentillesse, vos points de vue extérieurs et votre soutien de toutes heures !

Une pensée spéciale pour mes amis d'ailleurs : Mes Neuneus, Flora, Aurel, Béné, Maëva, Gwen, Adrien, Caro, Anaïs, Camille, Justine, Laura, les Montpelliérains pour tous les bons moments passés en dehors du laboratoire, l'aide et le soutien moral tout au long de la thèse.

Un immense merci évidemment à ma famille. Merci à mes parents de m'avoir soutenue et accompagnée durant toutes ces années d'études. Merci à vous, mon frerot, ma sister, ma grumelle, mon Lulu (dindon) pour votre soutien et vos appels durant ces derniers mois qui m'ont permis de me vider la tête. Merci à GP-GM pour ces petites attentions durant les deux semaines de rédactions mais aussi à mes oncles.

À Rayane, je ne saurais te remercier pour tout ce que tu as fait pour moi, même à distance que ce soit depuis Montpellier ou depuis Houston. Tu as su me soutenir, m'aider face à mes émotions, mes doutes, m'encourager et particulièrement pendant toute cette période de rédaction, hâte de commencer une nouvelle aventure à tes côtés.

---

# Table des matières

---

<b>Liste des Abréviations.....</b>	<b>11</b>
<b>Introduction générale.....</b>	<b>13</b>
<b>Chapitre I - Bibliographie.....</b>	<b>17</b>
<b>I.1 Les bactéries et la cellule procaryote.....</b>	<b>19</b>
I.1.1 Définition d'une bactérie .....	19
I.1.2 L'enveloppe bactérienne .....	19
I.1.2.1 La membrane cytoplasmique .....	19
I.1.2.2 Paroi bactérienne .....	20
I.1.3 Contenu de la bactérie .....	22
I.1.4 Les bactéries pathogènes et maladies infectieuses .....	24
<b>I.2 Les antibiotiques et leurs mécanismes d'action .....</b>	<b>27</b>
I.2.1 Ciblage de la synthèse des enveloppes de la bactérie .....	28
I.2.2 Ciblage de la synthèse des acides nucléiques.....	29
I.2.3 Ciblage de la voie de biosynthèse du folate .....	30
I.2.4 Ciblage de l'intégrité de la membrane bactérienne.....	31
I.2.5 Ciblage de l'ARN ribosomal et inhibition de la synthèse des protéines.....	32
I.2.5.1 Les macrolides : inhibiteurs de l'élongation de la chaîne peptidyle.....	33
I.2.5.2 Les tétracyclines : inhibiteurs de la livraison des ARNt.....	34
I.2.5.3 Les aminoglycosides : inhibiteurs de l'étape de translocation .....	34
I.2.6 Vers de nouvelles cibles ? .....	36
<b>I.3 Phénomènes de résistance aux antibiotiques .....</b>	<b>38</b>
I.3.1 La réduction de la concentration d'antimicrobien au niveau de la cible .....	39
I.3.2 La modification ou l'inactivation du médicament .....	40
I.3.2.1 Inactivation de l'antibiotique par hydrolyse .....	40
I.3.2.2 Inactivation de l'antibiotique par transfert d'un groupement chimique .....	41
I.3.3 La modification de la cible .....	41
I.3.4 L'acquisition d'un processus cellulaire pour détourner l'action du médicament .....	42
I.3.5 Résistance, tolérance et persistance.....	42
<b>I.4 Systèmes Toxine-Antitoxine (TA).....</b>	<b>44</b>
I.4.1 Généralité sur les systèmes TA .....	44
I.4.1.1 Classification des systèmes TA .....	44
I.4.1.2 Mécanisme d'action des toxines dans les systèmes TA.....	46
I.4.1.3 Activités biologiques des systèmes TA.....	47

I.4.2	Comment cibler ces systèmes TA : une nouvelle classe d'antimicrobiens ? .....	49
I.4.2.1	Activation du système TA de type I .....	49
I.4.2.2	Activation du système TA de type II.....	49
I.4.2.2.1	L'activation directe.....	50
I.4.2.2.2	L'activation indirecte .....	50
I.4.3	Le système TA de type I <i>AapA/IsoA</i> chez <i>Helicobacter pylori</i> .....	51
I.4.3.1	Système TA de type I.....	52
I.4.3.2	Système <i>AapA/IsoA</i> chez <i>Helicobacter pylori</i> .....	54
I.4.3.2.1	<i>Helicobacter pylori</i> .....	54
I.4.3.2.2	Système TA <i>AapA/IsoA</i> .....	54
<b>I.5</b>	<b>Les ligands d'ARN .....</b>	<b>59</b>
I.5.1	ARN d'intérêt thérapeutique .....	59
I.5.2	Stratégies de ciblage d'ARN : oligonucléotides et petites molécules .....	60
I.5.2.1	Approche par des oligonucléotides .....	60
I.5.2.2	Approche par des petites molécules : des stratégies multiples.....	61
I.5.3	Ciblage d'ARN bactériens.....	64
I.5.4	Ciblage d'ARN viraux.....	69
I.5.5	Ciblage d'ARN eucaryotes.....	74
<b>Chapitre II - Synthèse et évaluation biologique de nouveaux ligands d'ARN .....</b>		<b>81</b>
<b>II.1</b>	<b>Introduction : description du projet .....</b>	<b>83</b>
II.1.1	Conception de ligands d'ARN inhibiteurs de la formation de complexes boucle-boucle.....	85
II.1.2	Choix des motifs .....	86
II.1.2.1	Motifs hétérocycliques.....	86
II.1.2.2	Motifs aminoglycosides et analogues.....	87
II.1.2.3	Motifs acides aminés.....	88
II.1.2.4	Motifs isocyanures .....	89
II.1.3	Préparation de trois séries de nouveaux ligands d'ARN multifonctionnels .....	90
<b>II.2</b>	<b>Préparation des composants pour la réaction de Ugi-4CR .....</b>	<b>92</b>
II.2.1	Réactions multicomposants.....	92
II.2.1.1	Réaction de Ugi.....	93
II.2.1.2	Diversité structurale.....	94
II.2.1.3	Conditions de réaction de la U-4CR .....	95
II.2.1.4	Réaction de U-4CR sur support solide .....	96
II.2.2	Fonctionnalisation du support solide polymérique .....	96
II.2.3	Préparation des motifs ayant la fonction amine.....	102
II.2.3.1	Préparation des nucléobases naturelles alkylées.....	102
II.2.3.2	Synthèse des analogues naphtyridines .....	106
II.2.3.2.1	Synthèse de l'analogue naphtyridine par la voie A.....	108
II.2.3.2.2	Synthèse de l'analogue naphtyridine par la voie B.....	111
II.2.3.3	Préparation de la nucléobase S et de son dérivé .....	112
II.2.3.4	Sélection d'autres motifs comportant la fonction amine .....	113
II.2.4	Préparation des motifs ayant la fonction aldéhyde .....	115

II.2.4.1	Préparation de l'aldéhyde 2-déoxystreptamine .....	115
II.2.4.1.1	Préparation du motif 2-déoxystreptamine .....	115
II.2.4.1.2	Synthèse de l'aldéhyde 2-déoxystreptamine.....	117
II.2.4.2	Préparation de l'aldéhyde glucosamine.....	120
II.2.5	Récapitulatif des composants synthétisés.....	122
<b>II.3</b>	<b>Préparation des ligands d'ARN sur support solide : Séries A et B .....</b>	<b>124</b>
II.3.1	Synthèses des molécules obtenues dans la série A .....	124
II.3.1.1	Etude préliminaire .....	124
II.3.1.2	Formation des composés 2,5-dicétopipérazines .....	127
II.3.1.3	Synthèse des composés de la série A .....	129
II.3.2	Synthèses des composés de la série B.....	132
II.3.3	Conclusion de la synthèse sur support solide.....	136
<b>II.4</b>	<b>Préparation des ligands d'ARN en solution : Série C.....</b>	<b>138</b>
II.4.1	Synthèses de la série C-1 : série naphtalène .....	138
II.4.2	Synthèses de la série C-2 : série 2-déoxystreptamine .....	145
II.4.3	Synthèses de la série C-3 : ensemble de sous-familles .....	147
II.4.4	Conclusion des composés obtenus par une réaction de Ugi en solution .....	150
<b>II.5</b>	<b>Evaluations biochimiques .....</b>	<b>151</b>
II.5.1	Évaluation de l'affinité des composés sur <i>AapA1</i> .....	151
II.5.1.1	Évaluation de l'affinité des composés issus de la synthèse sur phase solide : Séries A et B .....	153
II.5.1.2	Évaluation de l'affinité des composés issus de la synthèse en solution : Série C .....	155
II.5.1.2.1	Évaluation de l'affinité des composés de la série C-1.....	155
II.5.1.2.2	Évaluation de l'affinité des composés de la série C-2.....	159
II.5.1.2.3	Évaluation de l'affinité des composés de la série C-3.....	160
II.5.1.3	Évaluation de la sélectivité sur les meilleurs ligands identifiés sur l'ARN <i>AapA1</i> .....	162
II.5.2	Mise au point d'un test <i>in vitro</i> pour évaluer la capacité d'inhibition de l'interaction boucle-boucle.....	163
II.5.2.1	Test d'inhibition utilisant l'anisotropie de fluorescence.....	165
II.5.2.1.1	Anisotropie de fluorescence (AF).....	165
II.5.2.1.2	Mise au point du test .....	166
II.5.2.2	Test d'inhibition utilisant un couple FRET .....	171
II.5.2.3	Expérience de retardement sur gel d'électrophorèse.....	173
II.5.3	Évaluation de l'activité <i>in vivo</i> sur le système <i>AapA1/IsoA1</i> – « système rapporteur » .....	176
II.5.4	Évaluation de l'activité antimicrobienne <i>in vivo</i> .....	179
II.5.5	Conclusion de l'évaluation biochimique.....	181
<b>II.6</b>	<b>Conclusion et perspectives.....</b>	<b>183</b>
<b>Chapitre III - Synthèse de nouveaux ligands d'ARN ciblant l'ARN TAR du VIH.....</b>		<b>189</b>
<b>III.1</b>	<b>Introduction .....</b>	<b>191</b>

III.1.1 Description du projet.....	191
III.1.2 Publication “Unveiling RNA - Binding Properties of Verapamil and Preparation of New Derivatives as Inhibitors of HIV-1 Tat-TAR Interaction” .....	194
<b>III.2 Conclusion.....</b>	<b>205</b>
<b><i>Conclusion générale</i> .....</b>	<b>207</b>
<b><i>Annexes</i> .....</b>	<b>211</b>
<b><i>Partie expérimentale</i>.....</b>	<b>221</b>
<b><i>Références bibliographiques</i> .....</b>	<b>329</b>

---

# Liste des Abréviations

---

<b>2,5-DKP</b> : 2,5-dicétopipérazine	<b>DMAP</b> : 4-Diméthylaminopyridine
<b>2DCS</b> : Criblage combinatoire deux dimensions	<b>Dia</b> : Diastéréoisomère
<b>2-DOS</b> : 2-désoxystreptamine	<b>DMF</b> : <i>N,N</i> -Diméthylformamide
<b>A</b> : Adénine	<b>Dnases</b> : Désoxyribonucléases
<b>AA</b> : Acide aminé	<b>ECDC</b> : Centre européen de prévention et de contrôle des maladies
<b>AapA</b> : Antisens-associated peptide family A	<b>EF-G</b> : Facteur d'élongation
<b>ACCs</b> : <i>N</i> -acétyltransférases	<b>EF-Tu</b> : Facteur d'élongation thermo instable
<b>ADN</b> : Acide DésoxyriboNucléique	<b>EGFP</b> : Enhanced Green Fluorescent Protein
<b>AF</b> : Anisotropie de fluorescence	<b>éq.</b> ou <b>équiv.</b> : équivalent
<b>AG</b> : Aminoglycoside	<b>FAM</b> : 5'6-Carboxyfluorescéine
<b>Ala</b> : Alanine	<b>FcS</b> : Criblage ciblé
<b>ALIS</b> : Système Automatisé d'Identification de Ligands	<b>FDA</b> : US Food and Drug Administration
<b>AMES</b> : Enzymes Modificateurs d'Aminoglycosides	<b>FID</b> : Déplacement d'Indicateur Fluorescent
<b>AMO</b> : AntiMicroARN Oligonucleoide	<b>fMet-ARNt</b> : ARN de transfert initiateur
<b>ANTs</b> : <i>O</i> -adényltransférases	<b>FMN</b> : Flavine MonoNucléotide
<b>APHs</b> : <i>O</i> -phosphotransférases	<b>FRET</b> : Transfert d'Energie entre molécules Fluorescentes
<b>ARN</b> : Acide RiboNucléique	<b>G</b> : Guanine
<b>ARNnc</b> : ARN non codants	<b>Gly</b> : Glycine
<b>ARNr</b> : ARN ribosomique	<b>h et H</b> : Hélices
<b>ARNt</b> : ARN de transfert	<b>His</b> : Histidine
<b>ATP</b> : Adénosine Triphosphate	<b>HMBC</b> : Heteronuclear Multiple Bond Correlation
<b>AUG</b> : Codon de démarrage	<b>HPLC</b> : Chromatographie en phase Liquide à Haute Performance
<b>Boc</b> : <i>tert</i> -butoxycarbonyle	<b>HRMS</b> : Spectrométrie de Masse à Haute Résolution
<b>C</b> : Cytosine	<b>HTS</b> : Criblage Haut débit
<b>CDK9</b> : Kinase 9 Dépendante de la Cycline	<b>IC<sub>50</sub></b> : Concentration Inhibitrice médiane
<b>CLIP</b> : Chemical Cross-Linking and Isolation by Pull Down	<b>IsoA</b> : Inhibitor of small ORF family A
<b>CMI</b> : Concentration Minimale d'Inhibition	<b>ITC</b> : Titration Calorimétrique Isotherme
<b>CN</b> : Chimiothèque Nationale	<b>K<sub>D</sub></b> : Constante de dissociation
<b>DCE</b> : 1,2-Dichloroéthane	
<b>DHF A</b> : Acide dihydrofolique	
<b>DIS</b> : Site d'Initiation de Dimérisation	

**LC-MS** : Chromatographie Liquide-Spectrométrie de Masse

**LNA** : Acide Nucléique bloqué (Locked Nucleic Acid)

**Leu** : Leucine

**LPS** : Lipopolysaccharides

**Lys** : Lysine

**M** : Molaire

**MCR** : Réaction Multicomposant

**MD** : Dynamique Moléculaire

**miARN** : microARN

**miR** : microARN

**n** : nano

**NAG** ou **GlcNac** : *N*-acétylglucosamine

**NAM** ou **MurNAc** : Acide *N*-acétylmuramique

**nd** : Non déterminé

**Néa** : Néamine

**Néo** : Néomycine

**NPET** : Nascent Peptide Exit Tunnel = tunnel de sortie du polypeptide naissant

**nt** : Nucléotides

**OMS** : Organisation Mondiale de la Santé

**PABA** : Acide 4-Aminobenzoïque

**pH** : Potentiel hydrogène

**Phe** : Phénylalanine

**PLP** : Protéine de Liaison aux Pénicillines

**PNA** : Acide Nucléique Peptidique

**PPI** : Interactions Protéine-Protéine

**Pre-** : Précurseur

**Pri-** : Primaire

**PSK** : Post-Segregational Killing

**PTC** : Centre peptidyl transférase

**P-TEFb** : Facteur d'élongation

**Py.** : Pyridine

**quant.** : Quantitatif

**RBS** : Site de fixation du ribosome

**RMN** : Résonance Magnétique Nucléaire

**RNase III** : RiboNucléase III

**RRE** : Rev-Response-Element

**S** : Unité de Svedberg

**SARM** : *S. aureus* résistant à la méticilline

**SD** : Séquence de Shine-Dalgarno

**siARN** : Petits ARN interférents

**Site A** : Site Accepteur

**Site E** : Site de sortie

**Site P** : Site Peptidyle

**SMA** : Maladie d'Amyotrophie Spinale

**SMM** : Small Molecule Microarray

**SPR** : Surface Plasmon Resonance

**SPS** : Support Solide Polymérique

**SRAS-CoV** : Coronavirus du Syndrome Respiratoire Sévère

**StARTS** : Relations Structure-Activité par le Séquençage

**T** : Thymine

**t.a.** : Température ambiante

**TA** : Toxine-Antitoxine

**TAR** : Réponse de Trans-Activation

**THF** : Tétrahydrofurane

**THF A** : Acide tétrahydrofolique

**Tm** : Température de fusion

**tr** : Temps de rétention

**U-4CR** : Réaction de Ugi à 4 composants

**UDC** : Ugi-Déprotection-Cyclisation

**uS** : Unité de protéines

**UTR** : Région non-Transcrite

**VHC** : Virus de l'hépatite C

**VIH** : Virus de l'Immunodéficience Humaine

---

# Introduction générale

---





La découverte de nouveaux agents antimicrobiens est aujourd'hui l'un des plus grands défis de la chimie médicinale en raison du nombre toujours croissant de bactéries multirésistantes. En 2014, il a été estimé que 700 000 personnes dans le monde mouraient actuellement à cause d'agents infectieux multirésistants et que ces derniers pourraient être responsables en 2050 d'environ 10 millions de morts par an, soit plus que le cancer d'après un rapport de l'économiste Jim O'Neill<sup>1</sup>. Le nombre de projets de recherche consacrés à la découverte de nouveaux antibiotiques diminue, notamment dans le milieu industriel, en raison d'un certain nombre de difficultés importantes rencontrées dans la recherche d'antimicrobiens efficaces. En outre, la difficulté d'identifier de nouvelles structures chimiques possédant des propriétés antibiotiques parallèlement au manque de nouvelles cibles par rapport à celles déjà connues pour induire l'apparition de résistances, limitent fortement la découverte de nouvelles générations d'antibiotiques.

Dans ce contexte, nous proposons d'innover en recherchant de nouvelles molécules et, plus important encore, des cibles originales. Ainsi, ce projet s'inscrit dans **la validation d'une nouvelle cible par la conception de nouvelles molécules ligands d'ARN agissant sur les systèmes Toxine-Antitoxine (TA)**. Dans ces systèmes, propres aux bactéries, la toxine est une protéine toxique pour la bactérie qui est produite uniquement en cas de stress tel qu'une variation de pH, de température ou en présence d'un traitement antibiotique. Au cours de ce travail de thèse, nous nous sommes intéressés au ciblage des systèmes TA de type I dans lesquels une interaction ARN/ARN entre l'ARNm codant pour la toxine et un ARN non codant appelé antitoxine est mise en jeu afin d'inhiber la production de la toxine. Nous avons ainsi synthétisé une nouvelle collection de ligands d'ARN dans le but d'inhiber les interactions ARN-ARN et d'activer artificiellement la production de la toxine pour induire la mort cellulaire bactérienne. La cible choisie au cours de ce travail est le **système *AapA/IsoA* de *Helicobacter pylori***.

La recherche de petites molécules capables de se lier à des ARN d'intérêt thérapeutique est un domaine en plein essor depuis plusieurs années. Une des cibles les plus étudiées est l'ARN ribosomique (ARNr) chez la bactérie, qui est d'ailleurs la cible de plusieurs classes d'antibiotiques sur le marché. Outre l'avancement dans la conception de ligands d'ARN, de nombreux progrès ont été également établis dans la compréhension des ARN et de leurs caractéristiques biologiques. Ces progrès permettent de mettre en avant de nouvelles cibles intéressantes et pertinentes telles que les systèmes TA qui doivent être étudiées avec l'utilisation de petites molécules.

Dans ce contexte, le développement de ce projet de thèse s'inscrit non seulement dans la validation d'une nouvelle cible antimicrobienne mais aussi dans la découverte de nouvelles molécules antibiotiques. Ainsi, ce manuscrit est organisé en trois parties :

Le **chapitre I** est consacré à une introduction générale pouvant être divisée en trois parties. Dans la première partie, il est présenté les différentes classes d'antibiotiques existantes ainsi que les différents mécanismes de résistance associés. Ensuite, les systèmes Toxine-Antitoxine sont décrits pour amener à la cible qui fait l'objet de ce travail qu'est le système TA *AapA/IsoA* présent chez la bactérie *Helicobacter pylori*. Enfin, à travers différents exemples ciblant les ARN bactériens, les ARN viraux et les ARN eucaryotes nous illustrerons les différents outils existants pour cibler les ARN avec de petites molécules.

Le **chapitre II** porte sur le développement de nouveaux ligands d'ARN à travers une réaction multicomposants : la réaction de Ugi. Cette librairie composée de trois séries a été ensuite évaluée par des tests biochimiques *in vitro* mis au point au laboratoire permettant de mesurer leur capacité à inhiber le complexe boucle-boucle du système TA. Des tests *in vivo* sur les bactéries ont aussi été réalisés sur l'ensemble des composés. Les synthèses effectuées et les résultats des évaluations biochimique et biologique seront présentés dans ce chapitre.

Le **chapitre III** décrit la synthèse de nouveaux ligands d'ARN ciblant **l'ARN TAR du VIH**. Cet ARN est une cible prometteuse pour le développement de nouvelles thérapies anti-VIH, mais aussi comme modèle idéal pour valider la découverte de ligands originaux d'ARN. Dans ce chapitre, nous décrivons le criblage d'une bibliothèque de composés du laboratoire et l'identification du composé vérapamil en tant que ligand d'ARN pouvant interagir avec l'ARN TAR. L'étude des relations structure-activité qui nous a permis d'obtenir un composé interagissant plus efficacement avec l'ARN TAR, sera détaillée dans ce chapitre.

Une conclusion générale nous permettra enfin de résumer l'ensemble des résultats obtenus et décrire les perspectives que ce travail particulièrement original et prometteur ouvre pour le futur.

---

# **Chapitre I - Bibliographie**

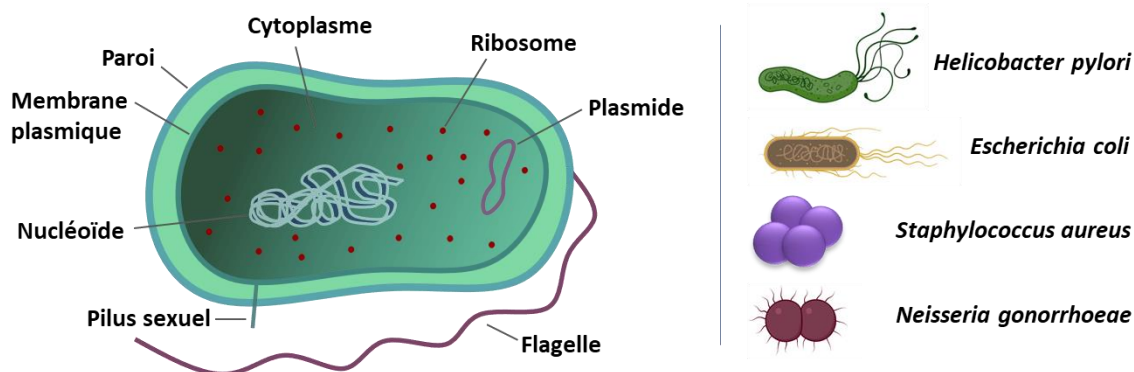
---



## I.1 Les bactéries et la cellule procaryote

### I.1.1 Définition d'une bactérie

Une bactérie est un micro-organisme unicellulaire de petite taille de l'ordre du micron (0,5 à 10-15  $\mu\text{m}$ ) faisant partie de la famille des procaryotes avec les archés<sup>2</sup>. La structure de la bactérie est relativement simple et ne possède pas de membrane nucléaire entourant le matériel génétique qui est entièrement contenu dans le cytoplasme (*figure 1*).



**Figure 1 :** (A) Représentation simple de la composition d'une bactérie. (B) Représentation non exhaustive de différentes morphologies des bactéries les plus connues : morphologie de forme sphérique « cocci » représentant les *Staphylococcus aureus* et *Neisseria gonorrhoeae*, de forme en bâtonnets « bacille » pour *Escherichia coli*, et de forme de « spirale » pour *Helicobacter pylori*.

La nomenclature scientifique de ces bactéries est constituée de deux noms latins : le premier nom correspond au genre et le deuxième correspond à l'espèce (ex : *Staphylococcus aureus*). D'autres classifications existent selon le caractère génotypique ou phénotypique qui est pris en compte. La classification à Gram classe les bactéries en deux : Gram positif ou Gram négatif. Cette dénomination provient de la différence de propriété liée à la paroi cellulaire en particulier de l'épaisseur de celle-ci. Cette différenciation a été mise en évidence par Christian Gram en 1884 par un test de coloration de Gram au cristal de violet.

Afin de détailler les différences structurales entre ces deux groupes de bactéries, les différentes composantes de la bactérie seront brièvement détaillées.

### I.1.2 L'enveloppe bactérienne

#### I.1.2.1 La membrane cytoplasmique

D'une épaisseur de 8 nm environ, la membrane cytoplasmique qui entoure la surface interne de la bactérie, est constituée d'une bicouche phospholipidique. Celle-ci comporte une tête hydrophile de triester de glycérols orientée vers l'extérieur de la membrane, et une région hydrophobe de chaînes d'acides gras hydrocarbonées constituant le cœur de cette membrane.

Des protéines associées à cette barrière sont présentes et vont permettre le contrôle des flux d'échanges avec le milieu extérieur par différents mécanismes et selon la nature du substrat, ce qui leur confère une perméabilité sélective. En effet, l'entrée ou la sortie peut s'effectuer par une diffusion libre ou par des protéines telles que des perméases et des transporteurs (passifs et actifs). D'autres protéines sont également présentes et sont impliquées dans la production d'énergie (ATPase) ou dans la synthèse de peptidoglycanes (Protéine de Liaison aux Pénicillines ; PLP).

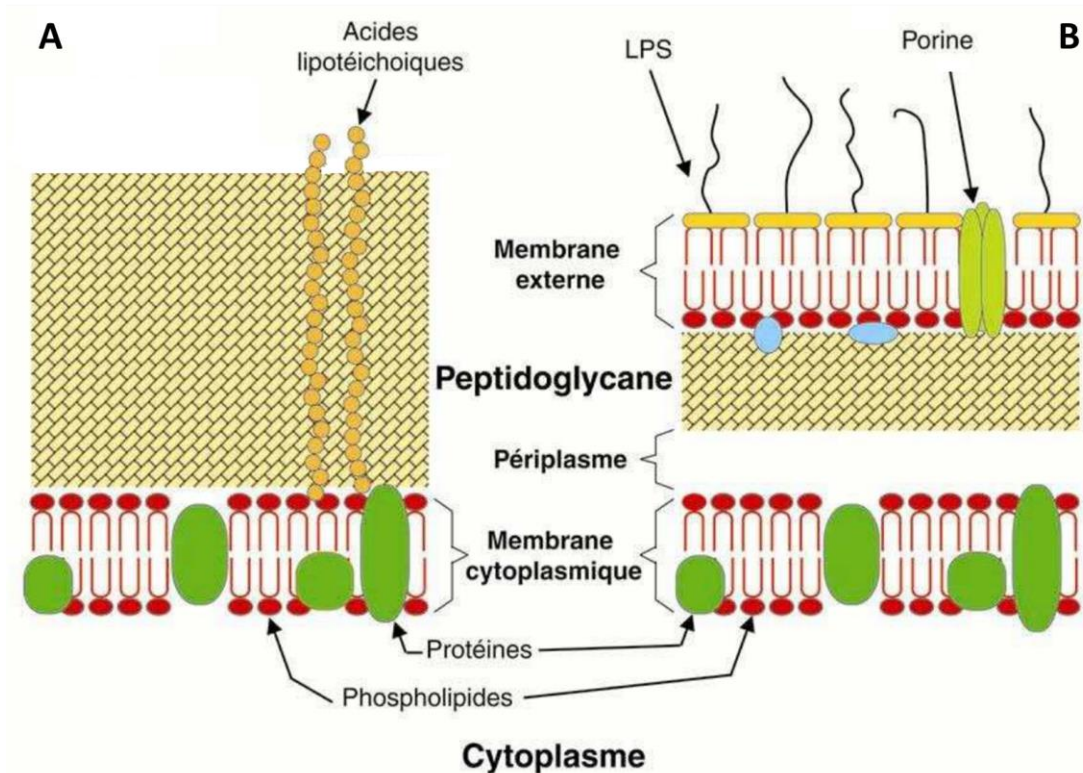
### I.1.2.2 Paroi bactérienne

La paroi bactérienne est composée d'une épaisseur de peptidoglycane. Cet hétéropolymère comporte une alternance de deux polysaccharides différents : le *N*-acétylglucosamine (NAG) et l'acide *N*-acétylmuramique (NAM) qui sont reliés par une liaison osidique. Sur la partie muramique, une structure térapeptidique est formée de quatre acides aminés différents (L-alanine, D-alanine, acide D-glutamique, L-lysine ou acide diaminopimélique) d'alternance L et D.

Pour former une structure tridimensionnelle en forme de feuillets, une réticulation de ce polysaccharide est nécessaire par deux connexions distinctes : soit une liaison directe entre deux chaînes térapeptidiques soit une liaison interpeptidique par un pentapeptide (présente exclusivement chez les bactéries à Gram positif).

L'ensemble de la biosynthèse du peptidoglycane nécessite plusieurs enzymes spécifiques qui sont regroupées sous le nom de PLP (Protéines Liant les Pénicillines) comme les transglycosylases et les transpeptidases.

La paroi bactérienne a pour rôle le maintien de la pression osmotique (jusqu'à 20 bars) et confère à la bactérie une certaine rigidité ce qui donne une morphologie spécifique. L'épaisseur de la paroi de peptidoglycanes va ainsi caractériser le type de la bactérie selon si elle est à Gram négatif (épaisseur fine) ou à Gram positif (paroi épaisse) (*figure 2*). Cependant, cette différence ne s'établit pas seulement sur l'épaisseur de cette paroi.



**Figure 2 :** Comparaison de l'organisation des parois cellulaires chez la bactérie à Gram positif (A) et la bactérie à Gram négatif (B)<sup>3</sup>. LPS = lipopolysaccharides, lipoprotéines (en bleu clair).

### Paroi bactérienne chez la bactérie à Gram positif :

Chez les bactéries à Gram positif, afin de relier la paroi très réticulée de peptidoglycane à la membrane cytoplasmique, des glycérides sont présents et confèrent à la surface une charge négative :

- Les acides téichoïques qui sont des polymères glycérophosphates ou polyalcools ancrés dans la paroi et qui dépassent du côté externe de celle-ci.
- Les acides lipotéichoïques qui possèdent un diglycéride en plus et traversent intégralement la paroi pour se fixer sur la membrane cytoplasmique.

### Paroi bactérienne chez la bactérie à Gram négatif :

La paroi des bactéries à Gram négatif est recouverte d'une membrane externe complexe. Sa structure comporte une double couche phospholipidique qui incorpore différentes protéines dont des porines. Ces dernières sont diverses mais spécifiques pour permettre le transport de nombreux substrats notamment des éléments nutritifs.

Du côté externe de la membrane, une couche formée de lipopolysaccharides (LPS) est présente. Ces LPS sont constitués de lipides A (dits toxiques) qui sont dans la partie interne de la membrane, une partie polysaccharide (antigène O) présente hors de la membrane et une partie intermédiaire (Core ou cœur) constituée d'un polysaccharide reliant les parties du lipide A et de l'antigène O.



Cette deuxième membrane octroie aux bactéries à Gram négatif des propriétés d'imperméabilité renforcée grâce à la sélectivité du transport par les porines et à une barrière dense d'oligosaccharides qui lui donne un rôle de tamis moléculaire.

### I.1.3 Contenu de la bactérie

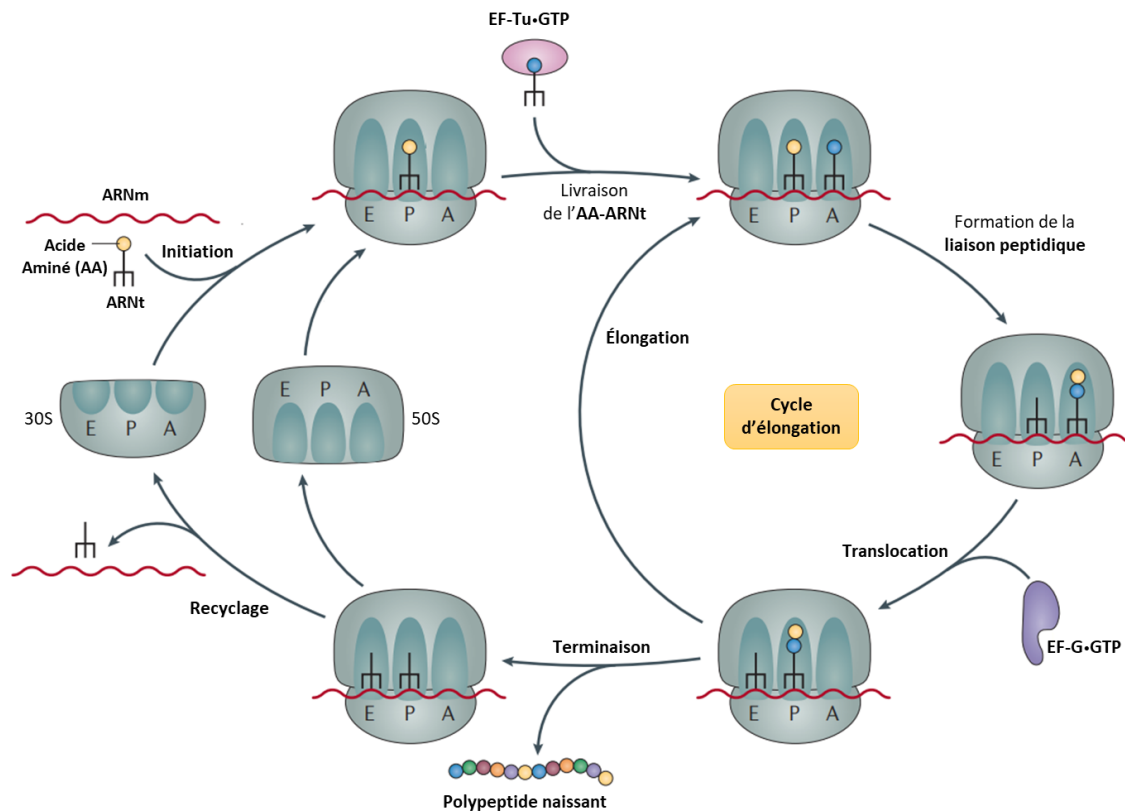
Dans le cytoplasme de la bactérie réside le matériel génétique et divers éléments protéiques.

Le matériel génétique d'une bactérie est regroupé dans le nucléoïde qui inclut un ADN chromosomique circulaire surenroulé et différents plasmides. Ces plasmides sont des gènes supplémentaires non essentiels pour le fonctionnement d'une bactérie mais qui permettent la production de composés possédant des propriétés supplémentaires telles que les bactériocines mais aussi des vecteurs impliqués dans la résistance aux antibiotiques.

Au cours de l'étape de transcription, l'ADN qui est en double brin unique et compacté, doit être déroulé en une structure simple brin pour permettre l'appariement des bases complémentaires. Dans ce processus, différentes enzymes tels que les topoisomérases (ADN gyrase et la topoisomérase IV) qui jouent un rôle essentiel. Ce sont des modulateurs de la topologie de l'ADN.

Un autre élément important et très présent dans ce cytoplasme est le ribosome. Cette machinerie macromoléculaire permet la polymérisation des acides aminés formant les différents polypeptides lors de la traduction<sup>4</sup>. Cet ensemble composé majoritairement de protéines forme un complexe dit 70S. La partie 30S contient l'ARN ribosomique (ARNr) 16S, et la grande partie 50S contient les ARNr 5S et 23S. Trois cavités se distinguent et permettent le processus de la traduction : le site E (site de sortie : Exit), site A (site Accepteur) et le site P (Peptidyle).

Grâce à tous ces éléments, le déroulement de la traduction s'opère en quatre étapes : l'initiation, l'élongation, la terminaison et le recyclage (*schéma 1*) :



*Schéma 1 : Mécanisme du fonctionnement du ribosome bactérien lors de la traduction<sup>5</sup>.*

L'étape d'initiation comporte la présence d'un codon de démarrage (AUG) sur l'ARNm en correspondance avec un ARN de transfert initiateur (fMet-ARNt) qui vient se placer dans le site P.

L'étape d'élongation comporte l'arrivée d'un nouvel ARN de transfert (ARNt) portant un acide aminé et l'anti-codon correspondant. Le ribosome place l'ARNt dans le site A grâce à la complémentarité de séquence entre l'anti-codon de l'ARNt et le codon de l'ARNm. Il s'ensuit alors la formation de la liaison peptidique entre les acides aminés des sites P et A.

Afin de réaliser d'autres ligations, une importante étape de « translocation » est nécessaire. Les ARN sont déplacés dans les sites E et P, ce qui va permettre l'arrivée d'un nouvel ARNt pour continuer l'élongation jusqu'à l'arrivée d'un codon stop. Ce codon stop engendrera une hydrolyse du polypeptide naissant.

Lors de l'étape de terminaison, le peptide qui vient d'être hydrolysé, va être entraîné dans un tunnel (NPET = Nascent Peptide Exit Tunnel) situé dans la partie 50S. Le polypeptide va alors se replier avant d'être libéré sous forme de protéine dans le cytoplasme.

Enfin, l'étape de recyclage permet la réutilisation des entités par séparation des deux sous-unités.

Enfin, parallèlement aux composants que nous venons de décrire, d'autres éléments variables sont également présents selon la nature de la bactérie tels que la capsule pour l'attachement de la bactérie à sa surface, le glycocalyx pour la formation du biofilm, les flagelles pour la mobilité ou les pili

permettant le processus de conjugaison entre bactéries. Ces différents éléments ainsi que d'autres, confèrent à la bactérie des propriétés permettant de survivre dans des conditions environnementales spécifiques.

L'ensemble de ces micro-organismes sont présents en harmonie avec l'environnement et également chez les êtres humains où ils sont bénéfiques pour l'organisme. Néanmoins, certaines de ces bactéries provoquent des déséquilibres dans l'organisme et peuvent être à l'origine de nombreuses maladies. Ce sont alors des bactéries pathogènes.

#### **I.1.4 Les bactéries pathogènes et maladies infectieuses**

Les bactéries pathogènes sont des bactéries capables d'induire une pathologie chez le sujet infecté. L'Organisation mondiale de la Santé (OMS) a lancé en 2017 un cri d'alerte sur le fait que « **les "super-bactéries" résistantes aux antibiotiques mettent en échec la médecine moderne et menacent de nous renvoyer un siècle en arrière** ». Ainsi, l'une des dix plus grandes menaces pour la santé publique, est la résistance aux antimicrobiens<sup>6</sup>.

Ces préoccupations ont incité l'OMS à lancer un plan d'action mondial sur la résistance aux antimicrobiens en 2015 que ce soit pour le secteur de la santé humaine mais également de la santé animale et de l'agriculture. Afin de guider le développement de nouveaux antimicrobiens, l'OMS a établi une liste prioritaire classant les bactéries résistantes aux antibiotiques au niveau mondial<sup>7</sup>. Cette liste est un outil qui promeut et guide la recherche, la découverte et le développement de nouveaux antibiotiques. Cette liste est organisée en trois niveaux : critique, élevé et moyen (*tableau 1*) et regroupe les différents pathogènes en fonction de l'espèce et du type de résistance que ces bactéries ont développée.

**Tableau 1 :** Classements des 12 bactéries les plus menaçantes au niveau mondial avec leurs familles d'antibiotiques résistantes : Priorité important, Priorité haute et Priorité moyenne :

	Bactéries	Conséquences	Résistance
CRITIQUE	<i>Acinetobacter baumannii</i>	Infection lors de blessures	Carbapénèmes
	<i>Pseudomonas aeruginosa</i>	Infections	Carbapénèmes
	<i>Enterobacteriaceae*</i>	Infections stomacales	Carbapénèmes, 3 <sup>rd</sup> génération de céphalosporines
	<i>Enterococcus faecium</i>	Méningite néonatale	Vancomycine
HAUTE	<i>Staphylococcus aureus</i>	Infections de la peau	Méticilline, intermédiaire ou complète à la vancomycine
	<i>Helicobacter pylori</i>	Problèmes gastriques, ulcères	Clarithromycine
	<i>Campylobacter spp.</i>	Gastroentérites,	Fluoroquinolones
	<i>Salmonellae</i>	Salmonelle	Fluoroquinolones
	<i>Neisseria gonorrhoeae</i>	Gonorrhée	Céphalosporines et fluoroquinolones
MEDIUM	<i>Streptococcus pneumoniae</i>	Pneumonie, méningite, Bactériémie fébrile	Pénicillines
	<i>Haemophilus influenzae</i>	Méningite, épiglottites	Ampicilline
	<i>Shigella spp.</i>	Dysenterie	Fluoroquinolones

\* Enterobacteriaceae inclue : *Klebsiella pneumoniae*, *Escherichia coli*, *Enterobacter spp.*, *Serratia spp.*, *Proteus spp.*, et *Providencia spp*, *Morganella spp.*

La tuberculose causée par le *Mycobacterium tuberculosis* est un cas particulier et un programme de recherche spécifique est dédié à la lutte contre cette maladie.

Le milieu hospitalier représente un environnement favorable pour la propagation d'agents pathogènes et les infections qui se développent dans ce contexte sont appelées infections nosocomiales. Les bactéries regroupées sous le nom « **ESKAPE** » qui est l'acronyme de *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* et *Enterobacter spp.*, sont présentes dans la plupart des infections nosocomiales pouvant être mortelles. Ces bactéries résistent de plus en plus aux meilleurs antibiotiques présents sur le marché comme la famille des carbapénèmes et les céphalosporines de 3<sup>ème</sup> génération.

Afin d'apporter une réponse au problème de multirésistance, des organisations régionales et/ou nationales ont été créées et ont été chargées de la surveillance de différents isolats résistants à ces antibiotiques. Au niveau Européen, le Centre Européen de Prévention et de Contrôle des maladies « ECDC » (European Center for Disease Prevention and Control) regroupe les données sur de nombreuses maladies (tuberculoses, covid-19) et sur les agents pathogènes bactériens<sup>8</sup>.

Dans son plan d'action, l'OMS s'est également engagée à promouvoir la bonne gestion de l'utilisation d'antibiotiques permettant de cibler ces agents pathogènes. Afin, de visualiser facilement l'importance du choix approprié de ces antibiotiques, un classement de ces derniers a été réalisé : la classification

de AWaRe pour Access, Watch et Reserve<sup>9</sup>. Cette classification permet de prioriser l'utilisation d'antibiotiques ne présentant pour l'instant que peu de résistance et qui sont actifs contre de nombreux pathogènes (classe Access). En revanche, l'utilisation d'antibiotiques efficaces sur des bactéries multirésistantes doit être limitée et réservée en dernier recours (classe Reserve).

Cette classification des antibiotiques rassemble les différentes classes existantes qui seront détaillées dans la partie suivante suivie par les différents mécanismes de résistance rapportés à ce jour.

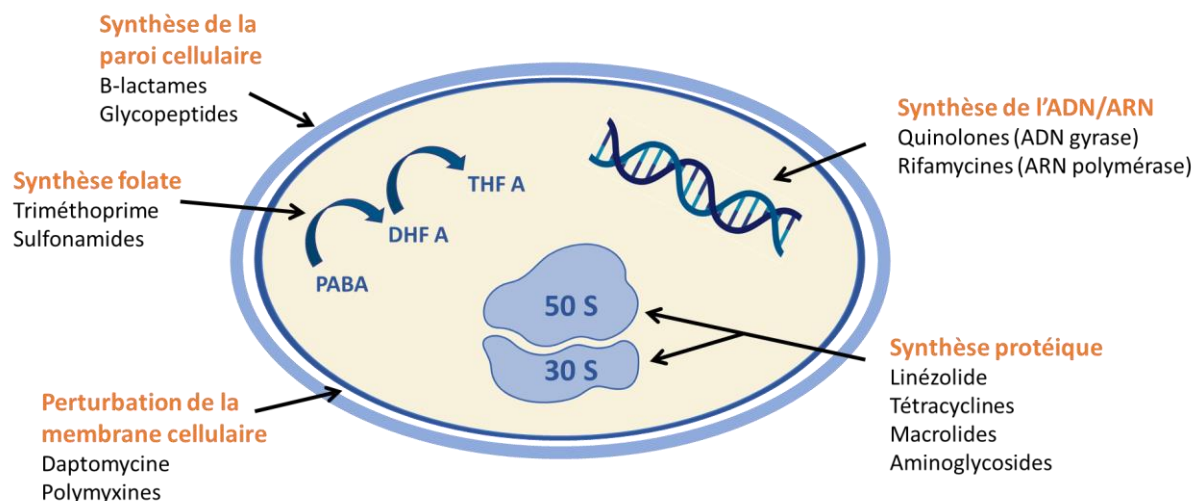
## I.2 Les antibiotiques et leurs mécanismes d'action

Le mot antibiotique vient du grec *anti* : « contre », et *bios* : « la vie ». C'est une substance capable de tuer les bactéries (bactéricide) ou d'inhiber leur croissance (bactériostatique). Les antibiotiques sont présents naturellement dans certains micro-organismes. Ils permettent de contrer ou de tuer d'autres micro-organismes qui pourraient les menacer dans leur biotope<sup>10,11</sup>.

En 1945, Fleming, Chain et Florey ont reçu le prix Nobel de médecine pour la découverte de la pénicilline<sup>10</sup>. C'est en 1928, qu'Alexander Fleming a accidentellement mis en avant l'action de la pénicilline (de structure  $\beta$ -lactame) qui provient du champignon de genre *Penicillium* et qui a un effet inhibiteur sur la croissance des *staphylocoques*. De nombreuses études sur de nouveaux composés provenant de différentes souches de micro-organismes ont suivi, mais de composés synthétiques ont aussi été développés notamment par Domagk pour les sulfamides (Protosil®). Ensuite, en 1943, l'introduction du dépistage de la souche *streptomycète* par Selman Waksman a permis la découverte de la streptomycine qui provient de la *Streptomyces griseus*<sup>11</sup>. La streptomycine est le premier antibiotique actif contre la tuberculose et le premier antibiotique aminoglycoside. Cette découverte a été récompensée par un prix Nobel en 1952.

Par la suite, de nombreuses découvertes ont permis le développement d'une multitude de classes d'antibiotiques notamment pendant la période de l'âge d'or (1940-1990). Cette période peut être découpée en deux parties dont la première correspond à la découverte de la plupart des familles d'antibiotiques existantes qui seront décrites ci-après. Puis, à partir de 1960, une phase d'optimisation en chimie médicinale de ces antibiotiques a conduit à différentes molécules afin de faire face aux vagues successives de bactéries pathogènes résistantes à ces antibiotiques mais également à la toxicité de certains d'entre eux.

À des fins thérapeutiques, ces antibiotiques vont cibler des voies métaboliques essentielles au bon fonctionnement de la bactérie notamment pour sa croissance et/ou pour sa survie. L'ensemble de ces antibiotiques peuvent être regroupés selon leur mode d'action. Les cibles principalement exploitées en thérapie sont : la synthèse des enveloppes bactériennes, la synthèse des acides nucléiques, la voie de biosynthèse du folate, l'intégrité de la membrane bactérienne et la synthèse des protéines (*figures 3*).

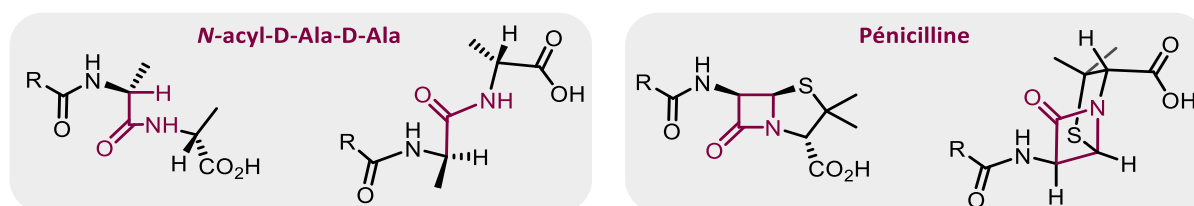


**Figure 3 :** Les différentes cibles thérapeutiques des molécules antibiotiques avec des exemples de familles d'antibiotiques<sup>12</sup>.

Ci-dessous, les principales caractéristiques de ces antibiotiques seront illustrées pour expliquer ces différentes cibles thérapeutiques.

### 1.2.1 Ciblage de la synthèse des enveloppes de la bactérie

L'assemblage de la paroi bactérienne constituée de peptidoglycanes implique des enzymes dites Protéines Liant les Pénicillines (PLP). Ces PLP, présentes uniquement chez la bactérie, sont la cible d'une classe d'antibiotiques les **bêta-lactamines**<sup>13,14</sup> telles que les pénicillines (*figure 4*). Ces molécules possèdent une structure comportant un cycle  $\beta$ -lactame qui rentre en compétition avec les substrats naturels (un dipeptide d'alanine : D-Ala-Ala des mucopeptides).

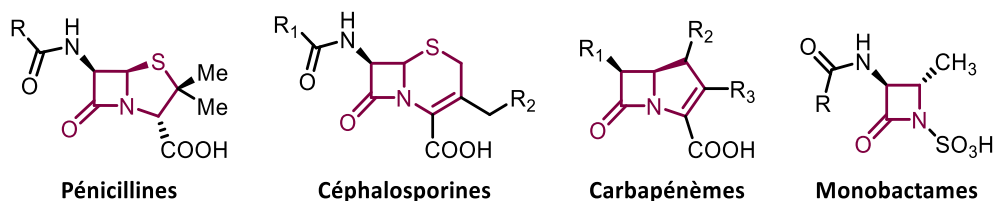


**Figure 4 :** Homologie de la structure du dipeptide N-acyl-D-Ala-Ala avec la structure de la pénicilline. La pénicilline comportant un noyau  $\beta$ -lactame en violet.

Cette homologie structurale permet de cibler le site catalytique formant un complexe stable acyl-enzyme par ouverture du cycle lactame. Cela engendre une perturbation du maintien de la paroi bactérienne favorisant l'autolyse de la membrane et provoque un effet létal lors de la phase de croissance de la bactérie.

Cette famille d'antibiotiques est divisée en quatre sous-familles selon le motif fusionné au noyau  $\beta$ -lactame : les pénicillines comportant un cycle à cinq chaînons avec un atome de soufre, les céphalosporines comportant un cycle à six chaînons avec un atome de soufre intégré, les

carbapénèmes portant un cycle de type cyclopentène ou 2,3-dihydrothiophène et les monobactames ne procédant pas de cycle fusionné (*figure 5*).



*Figure 5 : Représentation de différentes structures des noyaux de la famille des  $\beta$ -lactamines.*

Ces modifications de cycles sont le résultat d'un travail d'optimisation qui a été effectué pour pallier l'apparition de résistances qui sont causées par des  $\beta$ -lactamases qui dégradent les pénicillines et les céphalosporines.

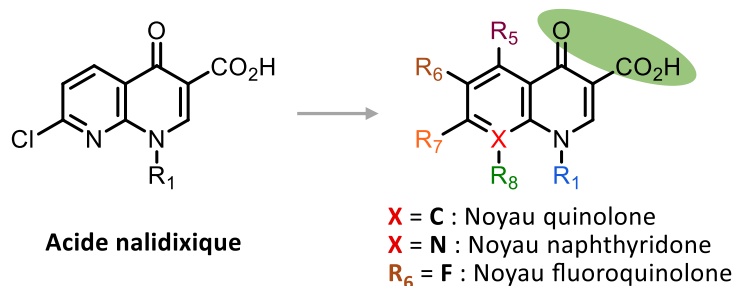
D'autres antibiotiques ciblent également la formation de cette membrane. La famille des **glycopeptides**<sup>15</sup>, tels que la Vancomycine, se fixe à une sous unité d'un précurseur du peptidoglycane (lipide II) contenant une terminaison D-Alanine-D-Alanine par des interactions hydrogènes. Ces interactions perturbent la suite du processus de l'assemblage du peptidoglycane. Enfin, la **fosfomycine**, constituée d'un époxyde, va agir au niveau intracytoplasmique en inhibant le processus de formation de l'acide *N*-acétyl-muramique (NAM).

## 1.2.2 Ciblage de la synthèse des acides nucléiques

La réplication du génome bactérien et sa transcription ont aussi été exploitées en tant que cibles thérapeutiques pour le développement d'antibiotiques. Les **quinolones**<sup>16</sup> agissent en tant qu'inhibiteurs sur les protéines qui ont pour rôle la modulation de la topologie de l'ADN telles que l'ADN gyrase et la topoisomérase IV. Ces antimicrobiens forment un complexe ternaire médicament-enzyme-ADN. Ce complexe entraîne un changement conformationnel de l'enzyme provoquant une instabilité de l'ADN et engendrant l'arrêt du processus de réplication ou de transcription. Cet arrêt provoque soit une inhibition de la réplication de l'ADN, soit une activation de la réponse au stress de l'ADN bactérien.

Dans cette famille, on retrouve l'acide nalidixique qui possède un noyau 1,8-naphtyridine. À la suite de l'identification de cet inhibiteur, différentes modifications ont été apportées afin d'en optimiser l'activité thérapeutique (*figure 6*).





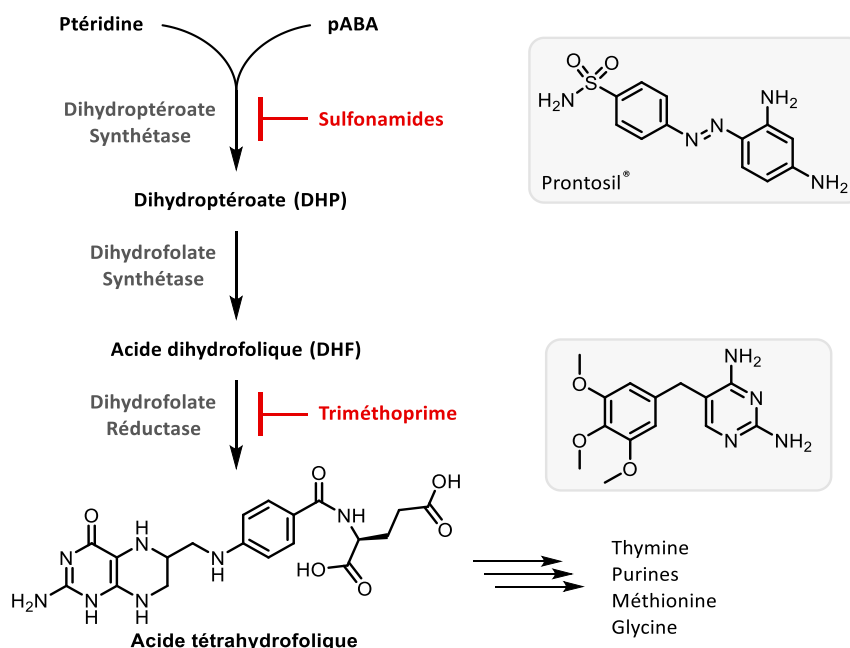
**Figure 6** : Structure générale de la famille des quinolones comportant une fonction céto-carbonyle. Il y a 6 positions majeures de modifications : R<sub>1</sub>, R<sub>5</sub>, R<sub>6</sub>, R<sub>7</sub>, R<sub>8</sub>. X = C définit la famille des quinolones, X = N définit la famille des naphthyridones et X = C avec R<sub>6</sub> = F définit la famille des fluoroquinolones.

Parmi ces modifications, on trouve l'insertion de l'atome de fluor en position R<sub>6</sub> qui apporte un noyau de type **fluoroquinolone**. Toutes ces modifications ont donné naissance à plusieurs générations de quinolones qui ont permis d'améliorer les propriétés pharmacocinétiques et toxicologiques, ainsi que d'élargir le spectre d'activité antimicrobienne.

Deux autres classes d'antibiotiques qui ciblent la formation d'acides nucléiques sont connues à ce jour. Tout d'abord la classe des **rifamycines** qui ciblent l'ARN polymérase ce qui a pour conséquence l'arrêt de la synthèse des ARN messagers. Ensuite la classe des composés **nitro-imidazoles**, dont la métronidazole, qui agit spécifiquement sur des bactéries anaérobies provoquant des coupures sur l'ADN.

### I.2.3 Ciblage de la voie de biosynthèse du folate

La voie de biosynthèse de l'acide tétrahydrofolique est essentielle pour la synthèse des nucléotides et de certains acides aminés. La famille des sulfamides identifiée notamment grâce aux travaux de Domagk dans les années 30, a permis de développer de nouveaux agents antibactériens. Le silfamidochrysoïde (Prontosil®)<sup>10</sup> et d'autres sulfamides sont associés au triméthoprime donnant un effet synergique. L'action de ces molécules intervient à différentes étapes dans l'inhibition de la voie de biosynthèse de l'acide tétrahydrofolique<sup>17</sup> (schéma 2).

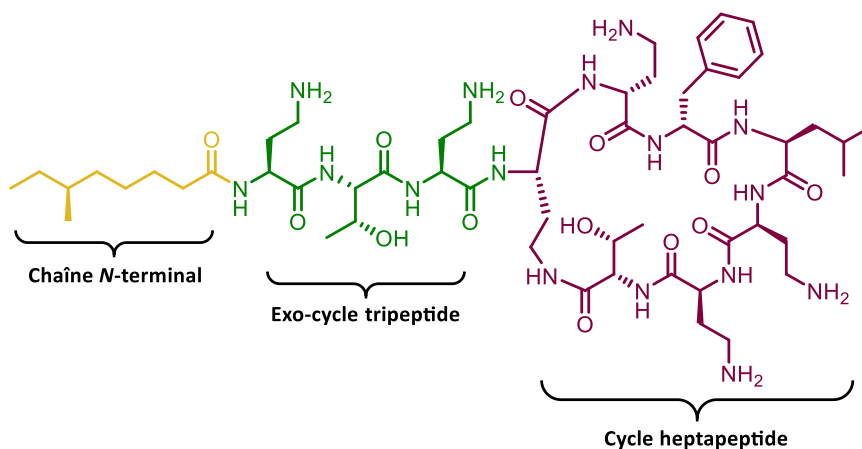


*Schéma 2 : Mécanisme d'action des antibiotiques de type sulfonamides et du triméthoprime sur une partie de la voie de biosynthèse de l'acide folique<sup>17</sup>. pABA = Acide 4-aminobenzoïque.*

Ces deux compétiteurs induisent une inhibition de la synthèse de l'ADN, de l'ARN et des protéines et attribuent ainsi des propriétés bactéricides à ces antibiotiques.

### 1.2.4 Ciblage de l'intégrité de la membrane bactérienne

La membrane bactérienne est essentielle dans le maintien de l'intégrité de la bactérie notamment pour le contrôle des échanges. La famille des **polymyxines**<sup>18</sup> comprend des antibiotiques qui vont venir interagir avec la membrane bactérienne chez les bactéries à Gram négatif. Les polymyxines sont des lipopeptides cycliques polybasiques qui proviennent de bactérie du sol : *Paenibacillus polymyxa* (figure 7).



*Figure 7 : Structure de l'antibiotique polymyxine B qui peut être segmenté en différents domaines selon les interactions possibles avec les lipides A.*

Les colistines font partie de cette famille où l'acide aminé D-phénylalanine (D-Phe) est remplacé par une D-leucine (D-Leu) par rapport aux polymyxines B. Leur mode d'action est conféré par leur caractère amphiphile qui leur permet de pénétrer dans la cellule bactérienne à Gram négatif. Elles vont cibler spécifiquement les lipopolysaccharides (lipide A) en s'insérant parmi ces derniers à l'aide d'un ensemble d'interactions électrostatiques et hydrophobes. Le complexe stable polymyxine-LPS va désorganiser la membrane ce qui va engendrer la perméabilité de celle-ci.

La **daptomycine**<sup>19</sup> est un **lipopeptide** parmi les derniers mis sur le marché en 2003. Cet antibiotique cible également la membrane des bactéries à Gram positif, mais son mécanisme d'action exact reste encore en discussion.

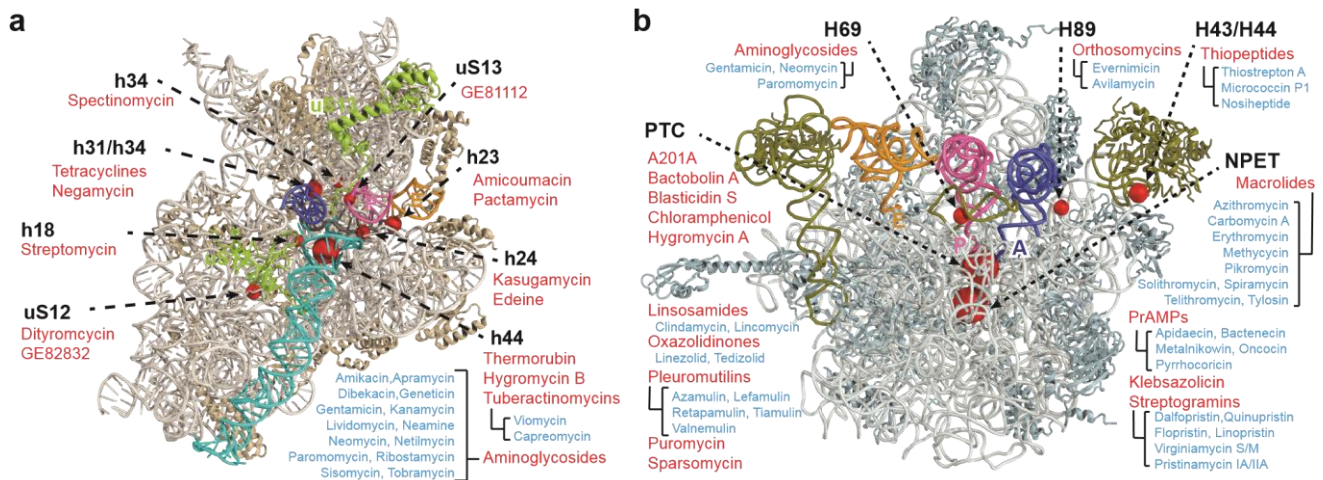
## I.2.5 Ciblage de l'ARN ribosomal et inhibition de la synthèse des protéines

La synthèse de protéines implique l'action du ribosome qui est une plateforme complexe et très dynamique (cf. **chap.1.1.3**)<sup>4</sup>. Les quatre étapes clés qui constituent le processus de la traduction (l'initiation, l'élongation, la terminaison et le recyclage) sont la cible de nombreux antibiotiques. Les actions de ces antibiotiques sont résumées dans ce tableau non exhaustif (**tableau 2**)<sup>5</sup>.

**Tableau 2 : Mécanismes d'action de différentes familles d'antibiotiques inhibiteurs de la synthèse des protéines : PTC = peptidyl transferase center :**

Famille	Composés antibiotiques	Mécanismes d'inhibition	Cible
Aminoglycoside	Streptomycine	Translocation	30S
	Paromomycine, néomycine	Translocation	30S
	Puromycine	PTC	50S
Chloramphénicol	Chloramphénicol	PTC	50S
Lincosamide	Lincomycine	PTC	50S
Macrolide	Erythromycine, azithromycine	Élongation	50S
Oxazolidinone	Linézolide	PTC	50S
Pleuromutiline	Tiamulin, Lefamulin	PTC	50S
Stréptogramine	Dalfopristine, quinupristine	PTC	50S
Tétracycline	Doxycycline, tigecycline, Omadacycline	Livraison de l'ARNt	30S
Thiopéptides	Thiostrepton	Facteur d'accroche	50S
Autres	Blasticidine	PTC et terminaison	50S
	Edéine, Kasugamycine	Initiation	30S
	Acide Fusidique	Élongation et recyclage	EF-G
	Kirromycine,	Élongation	EF-Tu

Ces antibiotiques vont agir par diverses interactions sur différentes hélices (h et H) que constituent les ARN ribosomiques sur les deux sous-unités 30S ou 50S du ribosome (*figure 8*).

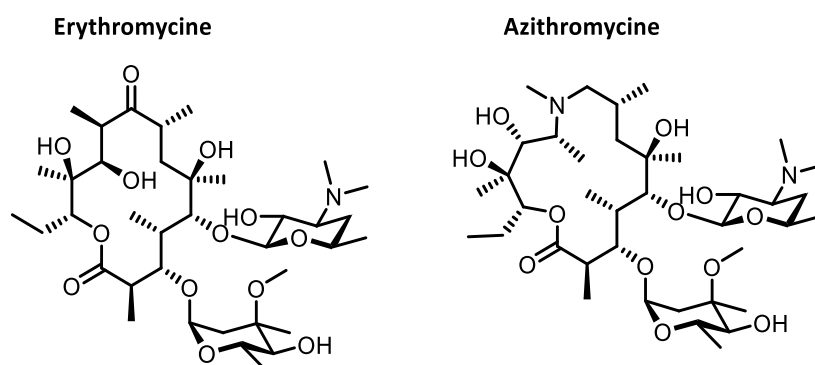


**Figure 8 :** Sites d'interactions pour une grande majorité d'antibiotiques ciblant la sous-unité 30S (a) et 50S (b) du ribosome bactérien<sup>20</sup>. En rouge le nom des familles des antibiotiques et en bleu les noms des antibiotiques. Les sphères rouges représentent les sites d'interactions. H et h : numéro des hélices des ARNr, uS : unité de protéines, PTC : Peptidyl Transferase Center, NPET : Nascent Peptide Exit Tunnel.

Trois classes de ces antibiotiques sont décrites ci-dessous : la famille des **macrolides**, des **tétracyclines** et la grande famille des **aminoglycosides**.

### 1.2.5.1 Les macrolides : inhibiteurs de l'élongation de la chaîne peptidyle

Les macrolides sont des structures cycliques possédant un noyau lactonique avec une taille du cycle variable entre 12 à 16 atomes<sup>21</sup>. Ce cycle est décoré avec des aminoglycosides ou des glycosides neutres à diverses positions (*figure 9*).



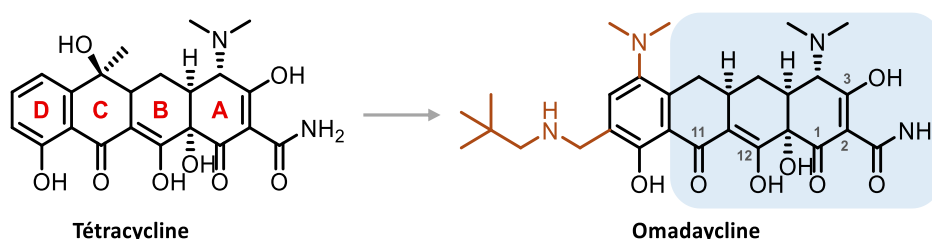
**Figure 9 :** Représentation de quelques antibiotiques de la famille des macrolides : Erythromycine (14 atomes) et Azithromycine (15 atomes).

Ces structures vont venir se positionner dans le tunnel de sortie adjacent du site PTC pour interagir avec une partie de l'ARN ribosomique 23S dans la sous-unité (50S). La progression de la chaîne polypeptidique naissante va ainsi être perturbée. Cependant, le mécanisme d'action est probablement plus compliqué et ne se limite pas uniquement à empêcher la sortie du peptide formé. En effet, les

macrolides agissent probablement sur la traduction en reconnaissant sélectivement certaines séquences de l'ARNm.

### I.2.5.2 Les tétracyclines : inhibiteurs de la livraison des ARNt

Les composés de la famille des tétracyclines<sup>22</sup> possèdent un noyau tétracyclique désigné A, B, C et D (*figure 10*). Les composés sont des compétiteurs qui vont inhiber la livraison des aminoacyl-ARNt au site accepteur du site A dans la sous-unité 30S du ribosome. Cette compétition réversible attribue une activité bactériostatique à ces antibiotiques.



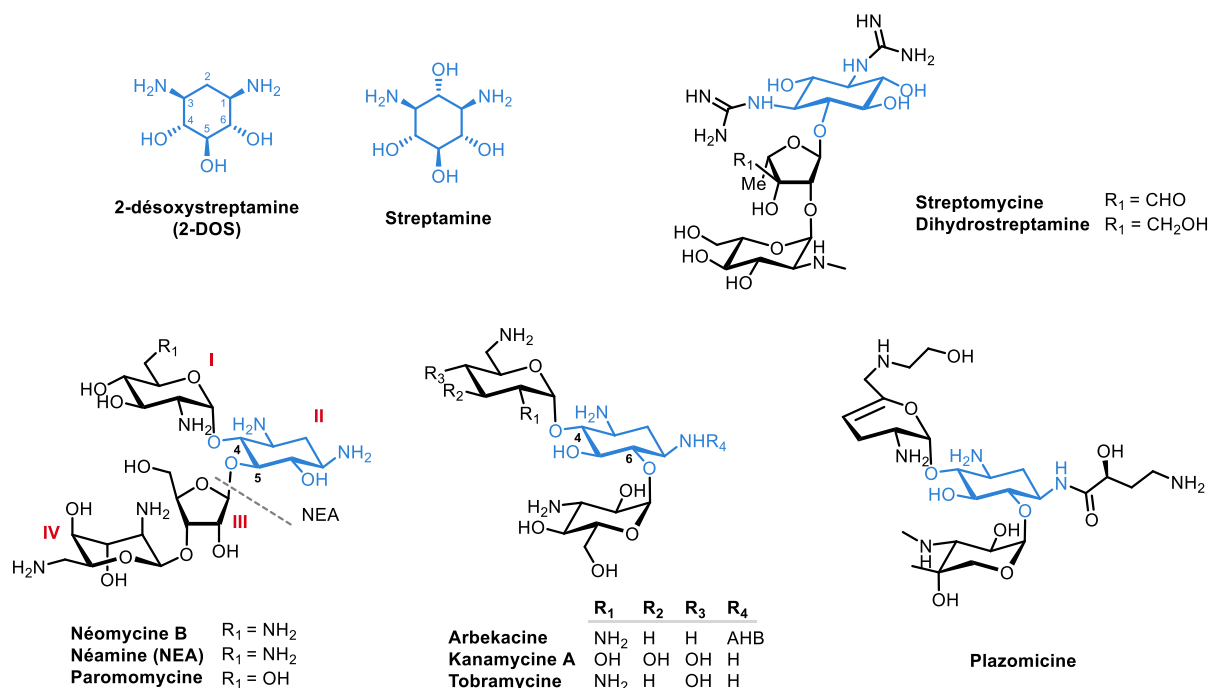
*Figure 10* : Représentation des structures de la tétracycline et de l'omadacycline<sup>23</sup>. Diverses modifications peuvent être apportées en positions 7 et 9 (marron). Des modifications dans la partie bleue induit une baisse de l'activité antibiotique : la stéréochimie reste importante pour l'activité.

La structure minimale considérée pour avoir cette activité antibiotique est la 6-désoxy-6-déméthyltétracycline. Dans ce motif, certains groupements fonctionnels sont indispensables pour l'activité (motif bleu) et d'autres entraînent l'amélioration des propriétés pharmacocinétiques (motif marron). Une des propriétés importantes est l'effet chélatant qu'apportent les tétracyclines grâce à différentes fonctions chimiques.

### I.2.5.3 Les aminoglycosides : inhibiteurs de l'étape de translocation

En 1943, le dépistage des streptomycètes initié par Selman Waksman a permis la découverte de la streptomycine qui provient de *Streptomyces griseus* et est active contre la tuberculose<sup>24</sup>. À partir de l'étude de différentes souches produisant un grand nombre de composés, une multitude d'aminoglycosides ont pu être découverts. Les aminoglycosides (AG) sont donc pour la plupart des produits naturels mais certains sont semi-synthétiques tels que la plazomicine approuvée en 2018<sup>25</sup> (*figure 11*).

Ces composés antimicrobiens sont des aminosucres constitués d'un noyau central qui peut être soit de type « aminocyclitol » (2-désoxystreptamine : 2-DOS) tel que la néomycine et la paromomycine, soit de type « streptamine » comme la streptomycine et la dihydrostreptamine. La famille d'AG contenant un noyau 2-DOS peut être classée également selon la substitution des glycosides sur ce noyau : 4,5-disubstitués ou 4,6-disubstitués (*figure 11*).



**Figure 11 :** Représentation non exhaustive des différents sous-groupes de la famille des aminoglycosides d'origine naturelle ou semi-synthétique. En bleu la structure du noyau 2-désoxystreptamine ou du noyau streptamine. AHB = (S)-4-amino-2-hydroxybutyryle (chaîne latérale).

Les AG agissent sur la traduction et induisent, dans la majorité des cas, une mauvaise lecture de l'ARN messager (ARNm) en inhibant l'étape de translocation.

Lors de la traduction, le ribosome assure l'appariement entre l'anticodon de l'ARNt et le codon de l'ARNm grâce à deux nucléotides hautement conservés présents sur la sous-unité de l'ARNr 16S : A1492 et A1493 (site A de la sous-unité 30S). Ces deux nucléotides permettent un réarrangement de l'hélice h44 (ARNr 16S) qui induit la reconnaissance codon/anticodon dans le site A. Ce réarrangement entraînant un basculement de ces deux nucléotides vers l'extérieur doit être compensé par des interactions supplémentaires qui sont stabilisées si l'interaction codon/anticodon est correcte. Cette stabilisation permet ainsi une activation de la traduction de l'ARNm par la suite.

Dans ce contexte, certains AG stabilisent ce réarrangement en se fixant dans la boucle interne de l'hélice h44 comme la paromomycine (figure 12).

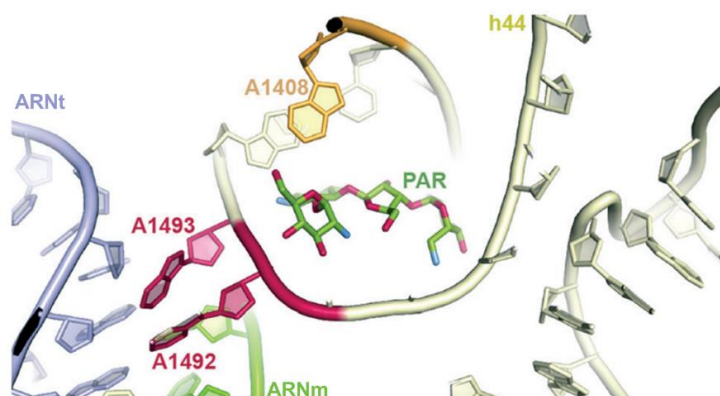


Figure 12 : Structure du complexe paromomycine-ribosome dans le site A de l'ARNr 16S<sup>24</sup>.

Ainsi, ces AG vont induire une mauvaise lecture de l'ARNm par stabilisation d'un ARNt erroné et qui ne correspond pas au codon de l'ARNm provoquant une mauvaise traduction des protéines. Les protéines défectueuses libérées vont s'accumuler dans la membrane cellulaire ce qui engendre une mauvaise perméabilité et favorise ensuite l'augmentation de la concentration d'AG au niveau intracellulaire. L'ensemble de ces phénomènes confère aux AG des propriétés bactéricides.

## I.2.6 Vers de nouvelles cibles ?

Ces dernières décennies, la famille des **oxazolidinones** est la seule nouvelle famille d'antibiotiques à avoir été découverte et développée notamment avec le **linézolide** en 2000. Cependant, cette famille est seulement active envers les bactéries à Gram positif (figure 13).

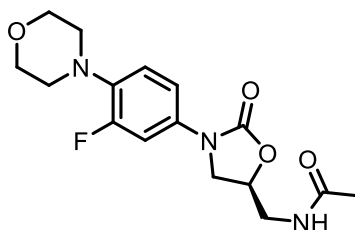


Figure 13 : Structure du composé linézolide.

Cet antibiotique inhibe la synthèse protéique de la bactérie en ciblant le site PTC du ribosome bactérien par des interactions avec l'ARN ribosomique (décrit précédemment). L'action du linézolide est bactériostatique.

Actuellement, pour le développement de nouveaux agents antimicrobiens, deux problèmes majeurs doivent être surmontés : la difficulté de perméabilité des membranes bactériennes et la toxicité des composés. En outre, le nombre de cibles spécifiques disponibles reste limité. Tous ces problèmes ont entraîné une forte diminution des efforts déployés pour la découverte de nouveaux antibiotiques, tant dans l'industrie que dans le monde universitaire.

Les quelques améliorations de certaines de ces classes d'antibiotiques ont été développées pour les bactérie à Gram positif<sup>26</sup> et les bactéries à Gram négatif<sup>27</sup> mais sans valider de nouvelles cibles thérapeutiques.

Ces améliorations constantes sont en très grande partie réalisées pour lutter contre les bactéries résistantes à ces antibiotiques. En effet, malgré un nombre important d'agents antimicrobiens parmi lesquels il est possible de choisir pour lutter contre une infection, l'apparition de résistances, parfois de façon très rapide après l'introduction de certains antibiotiques en thérapie, reste un problème de santé publique majeur.



### I.3 Phénomènes de résistance aux antibiotiques

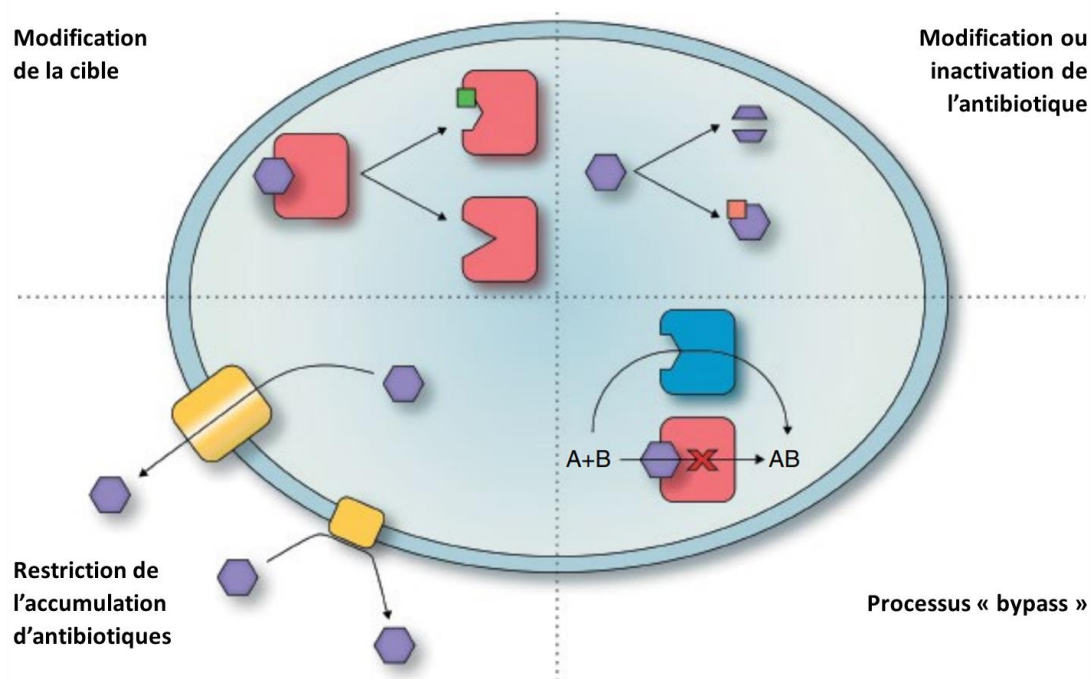
Les antibiotiques font partie des médicaments les plus fréquemment prescrits aujourd'hui, la résistance microbienne, due aux pressions de l'évolution et à une mauvaise utilisation, menace dangereusement le maintien de leur efficacité. Étant donné qu'au cours des dernières décennies, les infections dues à des bactéries résistantes aux médicaments n'ont cessé d'augmenter, la multirésistance est actuellement un problème de santé majeur qui menace de plus en plus la population.

Le phénomène de résistance peut être inné rendant la population bactérienne insensible à un mode d'action d'un antibiotique. A l'inverse, la résistance peut être acquise résultant de mécanismes liés à l'ADN par des transferts de gènes ou résultant d'une mutation.

En général, l'utilisation du terme de résistance permet de décrire la capacité héréditaire des micro-organismes à croître en présence de concentrations élevées d'un antibiotique indépendamment de la durée du traitement<sup>28</sup>.

Ces mécanismes de résistance peuvent être regroupés en quatre catégories (*figure 14*)<sup>29,30,31,32</sup>:

- La réduction de la concentration de médicament au niveau de la cible,
- La modification ou l'inactivation du médicament,
- La modification de la cible,
- L'acquisition d'un processus cellulaire pour détourner le médicament.



**Figure 14 :** Représentation des différents mécanismes impliqués dans le phénomène de résistance aux antibiotiques<sup>32</sup>.

Dans cette partie, ces quatre catégories seront décrites avec des exemples spécifiques pour chacune d'elles.

### I.3.1 La réduction de la concentration d'antimicrobien au niveau de la cible

Le premier phénomène que les bactéries peuvent utiliser pour limiter l'effet d'un antimicrobien est la réduction de la concentration intracellulaire du médicament ce qui affecte son efficacité d'action. Cette concentration peut être modulée par la bactérie selon deux mécanismes : la réduction de la perméabilité membranaire ou l'augmentation de l'efflux<sup>31</sup>.

- **La réduction de la perméabilité<sup>32</sup> :**

La notion de perméabilité est intrinsèquement plus prononcée chez la bactérie à Gram négatif par la présence d'une paroi épaisse. Cependant, des mécanismes de transport sont présents tels que des porines qui sont des protéines peu sélectives et larges. Les bactéries peuvent utiliser deux mécanismes afin de réduire la perméabilité dans l'ensemble de la population. Elles peuvent employer un mécanisme quantitatif basé sur la diminution du nombre de porines ou elles peuvent utiliser un mécanisme qualitatif basé sur la réduction de la largeur du passage formé par ces protéines (mutation dans les gènes). Ce type de résistance a été particulièrement observé par exemple avec la famille de  $\beta$ -lactamines.

- **L'augmentation de l'efflux<sup>31</sup> :**

Les pompes à efflux présentes naturellement chez les bactéries, permettent d'expulser des déchets métaboliques mais également des déchets toxiques tels que des métaux lourds ou encore des antibiotiques. Ce processus peut être qualifié de détoxification. Néanmoins, elles sont également responsables de manière significative de la multirésistance vis-à-vis des antibiotiques que ce soit de façon naturelle ou acquise.

Les pompes à efflux sont des protéines transmembranaires actives nécessitant une source d'énergie lors du transport. Ces protéines fonctionnent par gradient de protons ou d'ions  $\text{Na}^+$  ou par l'hydrolyse de l'ATP. Ces transporteurs sont souvent étroitement liés selon le substrat reconnu, la famille de substrat ou un éventail large de substrats. Ces transporteurs sont regroupés en cinq familles :

- **RND** (Resistance Nodulation Division)
- **MFS** (Major Facilitator Superfamily)
- **SMR** (Small Multidrug Resistance)
- **MATE** (Multidrug and Toxic compound Extrusion)
- **ABC** (ATP-Binding Cassette)

Certaines de ces pompes sont spécifiques selon l'espèce de la bactérie. Les pompes RND concernent en majorité les bactéries à Gram négatif, tandis que les pompes de type MSF sont retrouvées chez les bactéries à Gram positif. Ces différents systèmes d'efflux semblent concerner plus particulièrement la famille des fluoroquinolones et des macrolides.

Afin de remédier à ce mécanisme d'efflux, des recherches sur des inhibiteurs de pompes à efflux sont en cours. Ces inhibiteurs sont conçus pour être des inhibiteurs compétitifs ou non-compétitifs<sup>33</sup>. Une des molécules les plus connues pour être un inhibiteur de canaux est le composé vérapamil qui présente un effet inhibiteur de l'efflux chez *Mycobacterium tuberculosis*<sup>34</sup>.

### 1.3.2 La modification ou l'inactivation du médicament

Outre le fait d'empêcher les antibiotiques d'atteindre leur cible, les bactéries peuvent induire des modifications au niveau de la structure de ces molécules ou en inactiver certaines. Dans les deux cas de figure, cela va engendrer la perte de l'activité antimicrobienne. Les modifications sont effectuées par différentes enzymes qui agissent soit par hydrolyse, soit par transfert d'une fonction chimique sur l'antibiotique.

#### 1.3.2.1 Inactivation de l'antibiotique par hydrolyse

Un des mécanismes de résistance majeurs contre la famille des  $\beta$ -lactames est l'action d'enzymes catalysant l'hydrolyse du cycle  $\beta$ -lactame caractéristique de cette famille. Ces enzymes sous le nom de  $\beta$ -lactamase provoquent la formation de  $\beta$ -acides aminés inertes induisant l'inhibition de l'effet de ces antibiotiques sur les PLPs.

Ces  $\beta$ -lactamases sont nombreuses (classées sous 4 classes : A, B, C et D) et peuvent être définies selon leur mécanisme catalytique : les métallo- $\beta$ -lactamases dont le mécanisme d'hydrolyse est médié par des ions de zinc et les protéases à sérine. L'efficacité de ces enzymes est rendue possible par le fait qu'elles nécessitent seulement une molécule d'eau comme co-substrat et sont présentes à l'extérieur du cytoplasme. Afin de contrer en partie ce mécanisme de résistance des inhibiteurs de  $\beta$ -lactamase à sérine ont été développés (figure 15)<sup>35</sup>.

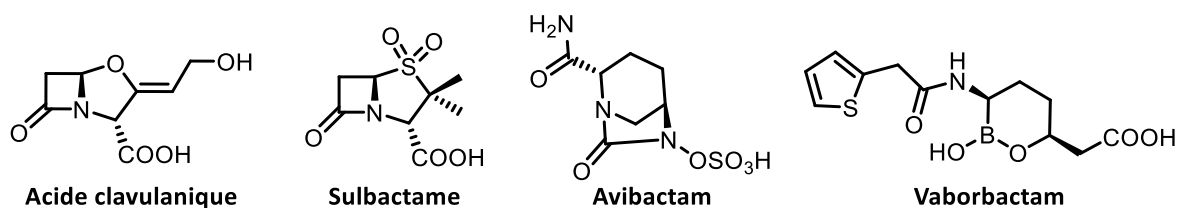


Figure 15 : Représentation des structures de différents inhibiteurs de  $\beta$ -lactamase.

Le premier de ces inhibiteurs de  $\beta$ -lactamase est l'acide clavulanique, une molécule d'origine naturelle. D'autres inhibiteurs possédant une structure qui mime le noyau de  $\beta$ -lactame possèdent des propriétés de faux substrat et induisent la dégradation de l'enzyme. Ces inhibiteurs sont utilisés en association avec un antibiotique de la famille des  $\beta$ -lactamines.

L'inactivation de l'antibiotique par l'hydrolyse est également retrouvée chez la famille des macrolides. Les liaisons esters présentes dans la structure de cette famille peuvent être hydrolysées par des thioestérases induisant la dégradation de l'antibiotique.

### I.3.2.2 Inactivation de l'antibiotique par transfert d'un groupement chimique

Certaines modifications induites par des transférases à l'aide de co-substrats peuvent affecter l'affinité de la molécule envers sa cible<sup>30,31,32</sup>. Ces modifications sont réalisées par la fonctionnalisation à l'aide d'acyles, de phosphates, de nucléotides, de riboses et de thiols.

Un des exemples les plus connus de résistance comportant ces modifications est la transformation des aminoglycosides<sup>24</sup>. Les trois enzymes principales opérant pour la modification de ces aminoglycosides sur des groupements hydroxyles et amines sont : les *N*-acétyltransférases (AACs), les *O*-phosphotransférases (APHs) et les *O*-adényltransférases (ANTs).

D'autres antibiotiques sont également modifiés par ces types d'enzymes. Il peut être cité l'exemple du chloramphénicol qui est modifié par des acétyltransférases, ou la rifampicine qui est modifié par des *O*-ribosylations, ou la fosfomycine qui est modifiée par des groupements glutathions.

### I.3.3 La modification de la cible

Parallèlement à la réduction de la concentration intracellulaire et à l'inactivation de l'antibiotique, des modifications au niveau de la cible peuvent aussi avoir lieu afin d'augmenter sa résistance aux antibiotiques<sup>30,31,32</sup>. Ces modifications doivent empêcher l'interaction du médicament sur la cible tout en gardant l'activité de la cible elle-même. Elles peuvent être induites directement lors de la traduction ou induites par des enzymes au niveau post-transcriptionnel.

Pour la famille des fluoroquinolones qui inhibe l'ADN gyrase et des topoisomérases de type IV, leur action antibiotique peut subir une baisse à la suite d'une mutation au niveau du site de liaison des fluoroquinolones empêchant ainsi leur action.

Pour les familles des macrolides, des lincosamides et des stréptogramines, une méthyltransférase va induire la mono- ou di- méthylation spécifiquement au niveau des adénines comme l'A2503 dans l'ARNr 23S ce qui va provoquer une gêne stérique favorisant une résistance. Cette méthylation est

également observée au niveau du site de liaison de certains aminoglycosides sur la guanine G1405 ou de l'adénine A1408 de l'ARNr 16S.

Outres ces modifications, un mécanisme de protection de la cible peut également affecter l'efficacité des antibiotiques. Le plus connu induit les résistances aux tétracyclines. Les bactéries résistantes à cette famille d'antibiotiques vont à l'aide de protéines du ribosome, déloger ces tétracyclines en évitant leur fixation à leur cible.

### **I.3.4 L'acquisition d'un processus cellulaire pour détourner l'action du médicament**

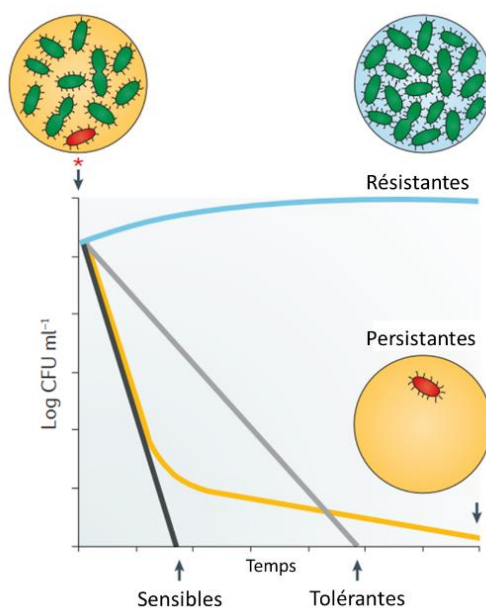
Certains phénomènes de résistance peuvent être classés dans la catégorie « bypass »<sup>32</sup>. Dans ce contexte, la bactérie acquiert des mécanismes qui vont lui permettre de contourner l'action du médicament et de réaliser sa fonction sans qu'elle soit perturbée.

Un des exemples les plus importants illustrant ce phénomène est l'acquisition d'une enzyme (PLP2a) qui est une alternative aux enzymes PLPs. En effet, l'acquisition de cette enzyme permet d'effectuer des liaisons croisées dans la synthèse du peptidoglycane. Ces PLP2a sont réfractaires aux actions des antibiotiques qui ciblent généralement ces enzymes PLPs par une acylation inefficace de la sérine active du site catalytique.

### **I.3.5 Résistance, tolérance et persistance**

Ces différents phénomènes de résistance illustrés ci-dessus résultent d'une adaptation génétique de la population bactérienne qui se traduit par une forte augmentation de la valeur de la CMI (Concentration Minimale d'Inhibition). Ainsi, les bactéries qui développent ces mécanismes d'adaptation permettent à l'ensemble de la population de croître normalement même en présence d'une forte concentration d'antibiotiques. Cette dernière peut par ailleurs engendrer une toxicité élevée pour l'hôte.

D'autres phénomènes présents chez la bactérie impactent également l'action d'antimicrobiens et sont basés sur la tolérance et la persistance<sup>28,36</sup> (*figure 16*). Ces deux phénomènes n'impactent pas la valeur de la CMI en comparaison du phénomène de la résistance.



**Figure 16** : Représentation de la cinétique de la mise à mort par l'utilisation d'antibiotiques (astérisque rouge) selon le type de cellules : résistantes (bleu), sensibles (gris foncé), tolérantes (gris clair) et persistantes (orange). CFU = Unités formant des colonies<sup>37</sup>.

Les cellules dites **tolérantes** sont des cellules qui, avant l'exposition aux antibiotiques, n'ont pas une croissance normale et ne possèdent pas de facteur de résistance. La tolérance est ainsi induite par un facteur extérieur tel que l'environnement qui engendre une croissance lente voire un état de dormance. La tolérance permet aux bactéries de survivre pendant la durée du traitement sur l'ensemble d'une population.

En ce qui concerne les bactéries **persistantes**, plusieurs phénomènes sont possibles et toujours à l'étude. Certaines bactéries au sein d'une population deviennent persistantes grâce à l'acquisition d'un état de non-division qui rend l'action des antibiotiques inefficace. Ces cellules persistantes sont ensuite capables, lors de l'arrêt du traitement antibiotique, de régénérer une population saine à croissance normale.

Plusieurs mécanismes peuvent induire le phénomène de persistance et un de ces mécanismes est basé sur l'activation des systèmes Toxine-Antitoxine (TA) qui représentent la cible biologique de ce travail de thèse. Ces systèmes TA seront décrits dans la section suivante<sup>36</sup>.

## I.4 Systèmes Toxine-Antitoxine (TA)

Les systèmes TA sont composés de deux éléments génétiques codant l'un pour une toxine stable capable d'induire un ralentissement de la croissance bactérienne (effet bactériostatique) voire la mort de la bactérie (effet bactéricide) et l'autre pour une antitoxine labile. L'antitoxine a pour rôle d'inhiber l'activité ou l'expression de la toxine. En cas de stress tel qu'une variation de la température, du pH ou encore un traitement antibiotique, les bactéries activent les systèmes TA en inhibant la production de l'antitoxine de manière à produire la toxine nocive pour la bactérie elle-même.

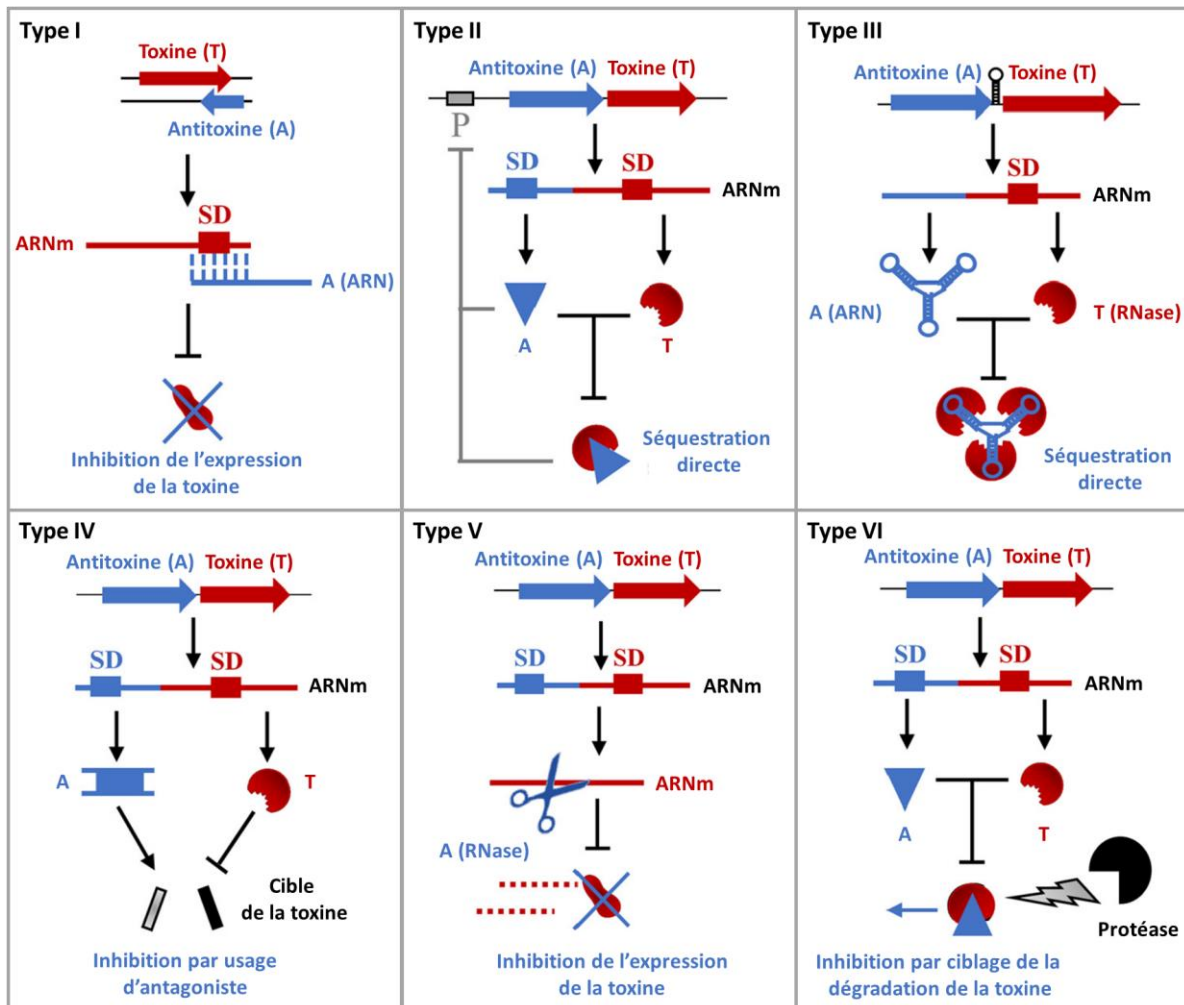
Ces systèmes permettent donc d'arrêter la croissance bactérienne pour mieux résister à un traitement antibiotique ou encore d'induire la mort des bactéries sensibles pour ne laisser survivre que les bactéries résistantes. Il s'agit donc d'un phénomène possible de persistance que les bactéries ont mis en place pour assurer la survie de la population en conditions défavorables<sup>36</sup>.

Les systèmes TA ont été initialement identifiés dans les plasmides bactériens chez *Escherichia coli* en 1983<sup>38</sup>. Après 30 ans de recherches sur les systèmes TA, différents types de systèmes ont été découverts, tant sur les plasmides que sur les chromosomes. L'action de leurs toxines et leur rôle biologique ont aussi fait l'objet de nombreux travaux et seront détaillés dans la section suivante.

### I.4.1 Généralité sur les systèmes TA

#### I.4.1.1 Classification des systèmes TA

Les toxines de tous les systèmes TA sont des protéines capables de cibler un processus cellulaire spécifique de la bactérie, alors que les antitoxines peuvent être soit une protéine soit un petit ARN non-codant. Les systèmes TA ont ainsi été classés en **6 types** en fonction de la nature de ces deux éléments dont le mécanisme de la neutralisation de la toxine par l'antitoxine <sup>39</sup> (*figure 17*). D'autres systèmes peuvent également être définis comme de nouveaux types de système TA (type VII et VIII)<sup>40</sup>.



**Figure 17 :** Vue d'ensemble des différents systèmes Toxine-Antitoxine de type I à VI. L'antitoxine est représentée en bleu et la toxine en rouge<sup>41</sup>. A = Antitoxine, T = Toxine, SD = Séquence de Shine-Dalgarno.

Dans le système de **type I**, les antitoxines sont des petits ARN non-codants qui régulent l'expression de la toxine en inhibant la traduction de l'ARNm codant pour la toxine. Dans le système de **type II**, l'antitoxine est une protéine qui va séquestrer la toxine en formant un complexe protéine-protéine.

Dans le système de **type III**, l'antitoxine est également un petit ARN non-codant et la toxine possède une activité endoribonucléase (endoRNase).

Dans toutes les autres classes de systèmes TA, les antitoxines sont des protéines. Dans le système de **type IV**, l'antitoxine va venir protéger la cible de la toxine en se liant dans un site adjacent (ex *FtsZ*). Pour les systèmes TA de **type V**, l'antitoxine a une activité d'endoribonucléase ce qui va provoquer la coupure de l'ARNm de la toxine. La toxine est un peptide hydrophobe qui va provoquer la formation de pores dans la membrane (formation de cellules fantômes, ex : *GhoST*). Enfin, pour le TA de **type VI**, la toxine va interférer avec la réplication de l'ADN en se fixant à l'ADN polymérase III (ex : *SocAB*). Et l'antitoxine est une protéine « adaptive » protéolytique qui inhibe la toxicité en favorisant sa dégradation par une protéase (ClpXP).

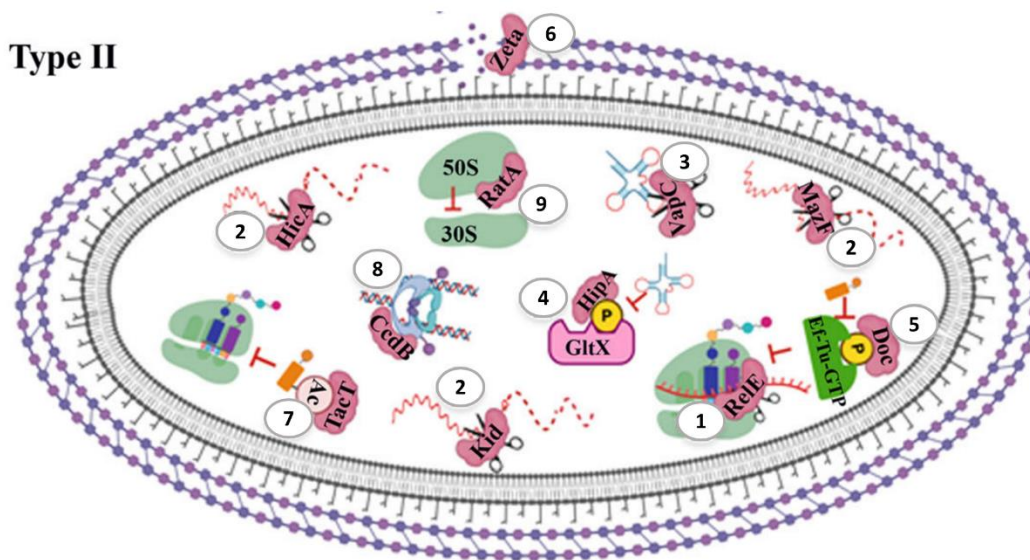


Les systèmes du type II, sont les plus étudiés et décrits, ce qui permet d'attribuer un grand nombre d'actions pour les toxines de cette famille.

### I.4.1.2 Mécanisme d'action des toxines dans les systèmes TA

Lorsque les toxines sont exprimées et ne sont pas inhibées par l'antitoxine, leurs activités au niveau moléculaire sont larges et variées. De nombreuses toxines ont été décrites et sont capables de cliver, dégrader ou modifier leurs cibles cellulaires<sup>40</sup>. Ces mécanismes peuvent obstruer la physiologie globale de la bactérie même lorsque la toxine est faiblement exprimée.

De nombreux exemples d'activités peuvent être représentés (*figure 18*), grâce à l'étude approfondie du système TA de type II.



**Figure 18 :** Représentation des différentes activités des toxines des systèmes Toxine-Antitoxine de type II<sup>40</sup>.  
(Voir le texte pour la description).

Les toxines de système de type II, les plus abondantes, sont des nucléases. Ces enzymes regroupent les désoxyribonucléase (DNases) (ex : *RaIR*), les endonucléases à ARNm dépendants du ribosome (**1** : *RelE*), les endonucléases à ARNm indépendants du ribosome (**2** : *MazF*, *Kid* et *HicA*) et des enzymes qui vont cliver des ARNt et des ARNr (**3** : *VapC*).

D'autres toxines vont modifier leurs cibles cellulaires au niveau post-traductionnel. Ces toxines sont des kinases qui peuvent venir phosphoryler les ARNt (**4** : *HipA*) ou le facteur d'élongation (**5** : *Doc*) mais aussi phosphoryler des entités qui vont inhiber la synthèse de la paroi cellulaire (**6** : *Zeta*). D'autres enzymes vont permettre de réaliser des acétylations notamment au niveau de l'ARNt chargé, ce qui mène à l'inhibition de la traduction (**7** : *TacT*). Par ailleurs, certaines toxines vont agir comme des empoisonneurs de la gyrase ce qui conduit à la fragmentation de l'ADN (**8** : *CcdB*). Enfin, la toxine *Rata* (**9**) est une protéine qui inhibe la traduction en prévenant l'assemblage des sous-unités du ribosome.

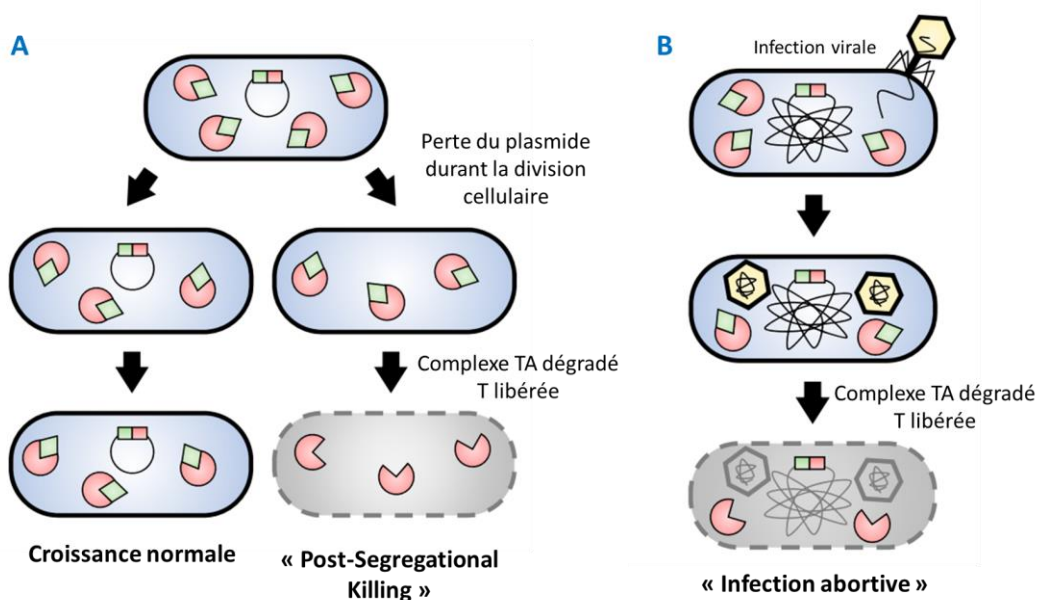
Pour les autres systèmes TA, de nombreuses toxines vont également cliver les ARNm libres (type III). Pour les systèmes TA de type I et V, certaines toxines vont perturber la membrane interne de la bactérie, tandis que pour les TA de type IV et VI, les toxines vont agir sur la division cellulaire ou inhiber la réplication de l'ADN.

Ainsi, l'ensemble de ces toxines intervient dans différents processus comme la traduction, la réplication ou encore la formation de la paroi bactérienne. Ces actions sont essentielles pour la prolifération et la survie bactériennes.

### I.4.1.3 Activités biologiques des systèmes TA

Les systèmes TA retrouvés dans les plasmides et les chromosomes sont impliqués dans de nombreux processus cellulaires par l'activité variée de la toxine conférant un effet bactériostatique (arrêt de la croissance) ou bactéricide (mort de la bactérie).

Ces systèmes TA sont impliqués dans l'adaptation bactérienne à un nouvel environnement allant par exemple de la mise en dormance pour la préservation des ressources énergétiques, de l'état de persistance<sup>37</sup>, de la formation de biofilm, de la virulence, de la mort cellulaire (« Post-Segregational Killing ») à la protection anti-phage<sup>39</sup> (infection abortive). Ces deux derniers rôles sont décrits ci-après (figure 19).



**Figure 19 :** Représentation des modèles « Post-Segregational Killing » (A) et de « l'infection abortive » (B). Les gènes TA (rectangles) et les protéines sont représentés en rouge (toxines) et en vert (antitoxines)<sup>42</sup>.

- **Le rôle de la mort programmée « Post-Segregational Killing (PSK) » :**

Les premiers systèmes TA découverts ont été caractérisés comme des « éléments de dépendance » qui participent à l'entretien des plasmides après la division cellulaire par un processus appelé PSK :

« Post Segregational Killing »<sup>38,43</sup> (*figure 19, A*). Ces systèmes TA sont principalement présents sur des plasmides.

Ce mécanisme repose sur la stabilité différentielle de la toxine et de l'antitoxine. Lors de la croissance cellulaire, les bactéries-filles héritent d'une copie des plasmides contenant les systèmes TA. Lorsqu'elles reçoivent ces plasmides, le ratio Toxine-Antitoxine permet un développement normal de ces cellules filles. Cependant, lorsqu'une bactérie-fille n'hérite pas de ce plasmide, dans le cytoplasme, il y a toujours la présence de ce complexe Toxine-Antitoxine. Cette antitoxine qui est labile, va être vite dégradée par exemple par une protéase et non régénérée<sup>42</sup>. Ainsi, la toxine stable est libérée et va entraîner la mort de la cellule-fille.

- **Protection anti-phage et « infection abortive » :**

Certains systèmes TA, notamment de type III, sont impliqués dans la protection contre les bactériophages par le biais d'un processus connu sous le nom « d'infection abortive »<sup>42</sup> (*figure 19, B*). Ce mécanisme de protection est un mécanisme inné qui va contrer la propagation des phages en provoquant le suicide des cellules infectées dans le but d'induire la survie uniquement des cellules non infectées. Ainsi, cette stratégie permet d'empêcher la réplication et la propagation de ces phages. Cependant, le mécanisme déclenchant les modules TA de type III reste encore inconnu<sup>39</sup>.

En ce qui concerne la formation de cellules persistantes, l'implication du système TA reste mal connue et n'a pas été encore complètement élucidée. Cependant, de nombreuses études ont montré l'implication de ces systèmes dans le phénomène de persistance<sup>36</sup>. L'activation de la persistance serait dirigée par divers stimuli environnementaux<sup>44</sup> comprenant une variation de la température, un traitement antibiotique, une augmentation du pH, un stress oxydatif ou encore des dommages à l'ADN. Ces différents stimuli, notamment dans le système de type II, activent des voies de signalisation en réponse au stress tel que la production de l'alarmone (p)ppGpp ou de la réponse SOS (LexA, RecA) responsable de la dégradation de l'antitoxine par exemple<sup>44,45</sup>.

Cependant, les mécanismes permettant la régulation et la sortie de l'état de persistance lorsque les conditions extérieures sont favorables restent à être élucidés<sup>44,46,37</sup>. Une des pistes étudiées pour les systèmes de type II comprend le mécanisme de la « coopérativité conditionnelle », qui passe par un dynamisme de la concentration des toxines dans la cellule par exemple.

## I.4.2 Comment cibler ces systèmes TA : une nouvelle classe d'antimicrobiens ?

Comme expliqué dans cette section, les systèmes TA contribuent à la résistance multiple aux antibiotiques puisqu'ils sont impliqués très fortement dans le phénomène de persistance. De plus en plus de preuves démontrent que la persistance contribue également à la pathogénicité. Dans ce contexte, les systèmes TA pourraient représenter une cible thérapeutique intéressante, prometteuse et originale. Le ciblage des systèmes TA est une stratégie particulièrement attrayante puisque ces systèmes TA n'ont pas d'homologues chez les humains. Et une molécule capable d'activer artificiellement ces systèmes TA ne serait pas toxique pour la bactérie en tant que telle puisque la molécule toxique reste la toxine. Le fait que la molécule toxique soit propre à la bactérie laisse espérer pas ou peu de développement de résistances puisqu'elle serait reconnue comme propre au microorganisme.

De nombreuses stratégies ont été étudiées ces dernières années notamment pour le ciblage des systèmes TA de type I et II<sup>47,48</sup>.

### I.4.2.1 Activation du système TA de type I

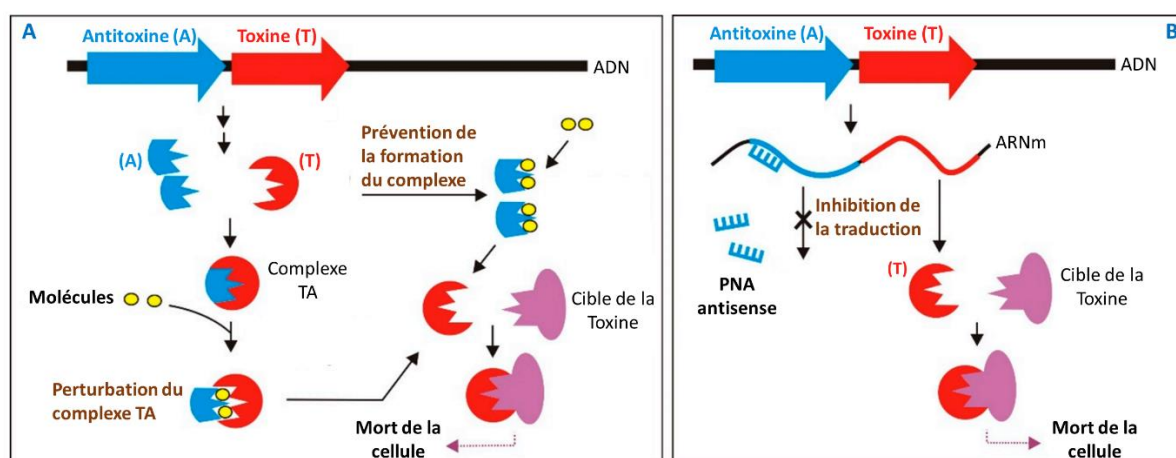
Les systèmes TA de type I possèdent principalement des propriétés bactéricides, c'est-à-dire que la toxine induit la mort bactérienne. Il s'agit donc de cibles particulièrement prometteuses pour le développement de nouveaux antibiotiques. Néanmoins, la seule stratégie figurant dans la littérature concernant les systèmes de type I, est l'optimisation directe d'une toxine. En effet, de nombreuses toxines de ce système ciblent la paroi cellulaire induisant une toxicité chez la bactérie. Ainsi, il paraît intéressant de convertir celles-ci en une drogue active. Felden *et al.* ont étudié le système TA *sprA1/sprA1<sub>AS</sub>* chez *Staphylococcus aureus*<sup>49,50</sup>. Dans ce système, le peptide cyclique PepA1 endommage la membrane bactérienne par la formation de pores. Cependant, PepA1 est également toxique pour les érythrocytes. Après, une optimisation de ce peptide, un heptapeptide a été obtenu et a induit une perméabilité de la membrane et a ainsi conduit à la mort bactérienne. Ce potentiel antimicrobien est actif contre *S. aureus* résistant à la *méticilline* (SARM) mais également contre d'autres bactéries gram-positives et gram-négatives.

### I.4.2.2 Activation du système TA de type II

Comme mentionné précédemment, les systèmes TA de type II sont les plus étudiés à ce jour et plusieurs façons de les cibler ont été décrits. De manière simple, des molécules pourraient être utilisées pour perturber ou inhiber le complexe Toxine-Antitoxine (activation directe) ou activer d'autres partenaires capables de prévenir la formation de ce complexe (activation indirecte).

## I.4.2.2.1 L'activation directe

Plusieurs approches ont été proposées pour le ciblage direct des TA de type II. Une des propositions pour l'activation directe est d'agir au niveau du site de reconnaissance de ce complexe Toxine-Antitoxine pour induire la libération de l'activité de la toxine en utilisant des inhibiteurs d'interactions protéine-protéine (PPI) <sup>47,51</sup> (figure 20, A). Des octapeptides ont ainsi été évalués pour leur capacité à inhiber le complexe *PemK/PemI* où la toxine engendre une forte toxicité dans les cellules des bactéries *Bacillus*. Ces octapeptides ont été conçus en mimant le domaine C-terminal de l'antitoxine *PemI* impliqué dans la reconnaissance avec la toxine. Ces études ont montré la faisabilité de l'approche mais, cependant, une baisse de l'activité de la toxine a été observée.



**Figure 20 :** Représentation des stratégies antibactériennes proposées sur l'activation directe des toxines des systèmes TA<sup>48</sup>. **A)** Perturbation du complexe protéique et antitoxine et/ou prévention de la formation de complexes protéiques. **B)** Inhibition de la traduction de l'antitoxine.

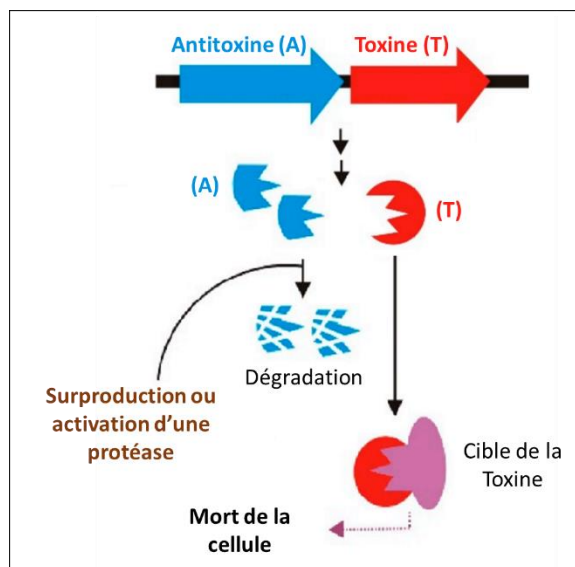
Une autre stratégie est de passer par l'inhibition de la traduction de l'antitoxine<sup>48</sup> (figure 20, B). Les exemples qui illustrent cette stratégie passent par l'utilisation des ARN antisens. Ces derniers entraînent l'inhibition de la traduction par complémentarité avec la partie de l'ARNm codant l'antitoxine. Les acides nucléiques peptidiques (PNA) antisens ont été étudiés dans ce contexte. Cette stratégie a été appliquée dans les systèmes TA *mazEF* et *hipBA* qui entraînent une diminution de la croissance cellulaire chez *Escherichia coli*.

## I.4.2.2.2 L'activation indirecte

L'activation indirecte du système TA implique des facteurs cellulaires qui sont interconnectés avec ces systèmes tels que les protéases dégradant l'antitoxine.

Dans les systèmes de type II, l'activation de la toxine nécessite la protéolyse de l'antitoxine par des protéases bactériennes comme Lon et ClpP qui sont les plus connues. Des études ont montré que les molécules acyldepsipeptides (ADEP) activent spécifiquement la protéase bactérienne ClpP<sup>48</sup> (figure 21). Cette activation incontrôlée pour la bactérie induit l'inhibition de la division cellulaire

bactérienne et entraîne la mort cellulaire. Celle-ci serait également due à l'activation de certaines toxines répondant à l'augmentation de cette protéase.



**Figure 21** : Exemple d'une des stratégies basées sur l'activation indirecte des toxines des systèmes TA : activation des protéases Lon ou ClpP qui dégradent l'antitoxine<sup>48</sup>.

D'autres approches ont également été envisagées impliquant d'autres facteurs jouant un rôle dans la reconnaissance de ces systèmes TA<sup>48</sup>.

L'ensemble de ces recherches montre un potentiel important pour l'activation des systèmes TA en tant que nouvelle cible thérapeutique. Dans ce contexte, nous proposons dans ce travail, la validation d'un système TA de type I en tant que nouvelle cible antimicrobienne. Nous avons choisi d'étudier un système TA de type I chez *Helicobacter pylori* découvert très récemment et décrit dans le paragraphe suivant.

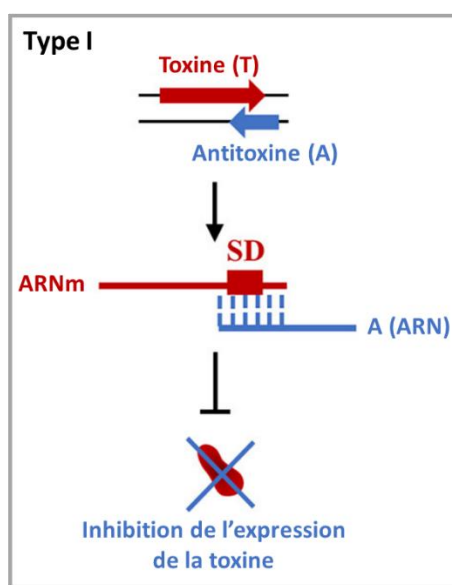
### I.4.3 Le système TA de type I *AapA/IsoA* chez *Helicobacter pylori*

Ce système TA de type I, appelé ***AapA/IsoA*** (Antisens-associated peptide family A / Inhibitor of small ORF family A) est présent dans les chromosomes chez la bactérie *Helicobacter pylori*. Ce nouveau système TA a été découvert et étudié par l'équipe du Dr. Fabien Darfeuille au laboratoire : Acides nucléiques : Régulations Naturelles et Artificielles (ARNA) de l'Université de Bordeaux avec laquelle nous collaborons pour le développement de ce projet.

Dans cette partie sera développé de façon générale le système TA de type I dans un premier temps, puis les caractéristiques du système *AapA/IsoA* impliquant les différents travaux de l'équipe du Dr. Darfeuille pour amener à la nouvelle stratégie de ciblage des systèmes TA de type I.

### I.4.3.1 Système TA de type I

Comme mentionné précédemment, les systèmes TA de type I sont constitués d'un ARN non-codant qui représente l'antitoxine et d'un ARN messager codant pour la toxine (*figure 22*). Dans ces systèmes, les antitoxines ARN reconnaissent spécifiquement par complémentarité des bases l'ARNm de la toxine via une interaction ARN-ARN. La formation de ce complexe inhibe la traduction de l'ARNm et bloque ainsi la production de la toxine.



*Figure 22* : Représentation du système TA de type I. SD = Séquence de Shine-Dalgarno.

Quand la bactérie se trouve en conditions de stress environnemental, tel que des variations de température, la présence de stress oxydant, le manque de nutriment ou encore le traitement antibiotique, les bactéries activent la production des toxines.

- **La toxine**

Dans le système TA de **Type I**, la majorité des toxines ont des caractéristiques structurales conservées. Ce sont des peptides hydrophobes de structure alpha-hélicoïdale (20-65 acides-aminés) possédant un domaine transmembranaire qui leur permet d'interagir avec la membrane cytoplasmique de la bactérie<sup>52</sup>.

Ces toxines peuvent être classées en deux sous-groupes selon leur mode d'action et leur localisation. D'abord, les toxines associées à la membrane agissent par formation de pores et /ou par condensation nucléoïde. Certains de ces peptides formant des pores vont être responsables de la perturbation du potentiel de la membrane cytoplasmique. Cela va engendrer une baisse de la production de l'ATP et de la force motrice des protons. Il sera alors observé soit un arrêt de la croissance, soit une lyse cellulaire. Les toxines possédant une partie de leur squelette dans le cytoplasme vont quant à elles, induire une distribution aberrante d'ADN induisant une condensation nucléoïde.

Contrairement à ces toxines, il a été également décrit que certaines toxines (telles que *SymE* et *RalR*) ne comportant pas de parties transmembranaires vont alors cliver des acides nucléiques. Ces toxines agissent comme des endoribonucléases pour induire des coupures dans les ARNm ou l'ADN.

#### ▪ L'antitoxine

En ce qui concerne l'antitoxine, elle va intervenir sur la régulation de l'expression de la toxine. Cette régulation se produit soit par la dégradation de l'ARNm de la toxine, soit par l'inhibition de sa traduction soit en utilisant ces deux mécanismes<sup>53</sup>. Ceci dépend du site de l'interaction de l'antitoxine vis-à-vis de l'ARNm mais également de la structure secondaire de celui-ci.

En effet, l'interaction entre les deux ARNs se fait par une complémentarité de séquence variant entre 18 nucléotides et 75 nucléotides au niveau des sites spécifiques à la traduction de l'ARNm qui code la toxine. Les antitoxines qui vont réprimer la traduction auront des complémentarités avec l'ARNm au niveau de différents modules nécessaires à la traduction comme le site de fixation du ribosome (RBS : ribosome binding site)<sup>54</sup>.

En ce qui concerne la dégradation de l'ARNm, il semblerait que l'interaction de l'antitoxine se fasse au niveau de l'extrémité 3' ce qui pourrait stimuler la dégradation du duplex étendu ainsi formé. Cette dégradation permettrait un renouvellement rapide de l'ARNm de la toxine.

#### ▪ Rôles biologiques

Le premier système TA découvert est le système *hok/sok* présent dans les plasmides chez *Escherichia coli*<sup>43</sup>. Ce système est impliqué dans le maintien de plasmide par le processus « post-segregational killing ». Depuis, de nombreux systèmes TA de type I ont été identifiés dans les plasmides mais aussi dans les chromosomes.

Pour les systèmes TA de type I situés dans les chromosomes, des études ont montré leurs implications dans le mécanisme de persistance<sup>55</sup>. Pour beaucoup d'entre eux, leurs rôles biologiques ainsi que leur mode d'activation restent encore à élucider.

Cependant, certains systèmes TA de type I chromosomiques impliquent des toxines qui, une fois exprimées, sont capables d'induire la mort des cellules bactériennes. C'est le cas pour le système *AapA/IsoA* découvert chez *Helicobacter pylori*.



### 1.4.3.2 Système AapA/IsoA chez *Helicobacter pylori*

#### 1.4.3.2.1 *Helicobacter pylori*

*Helicobacter pylori* est une bactérie en forme de spirale à Gram négatif comportant 4 à 6 flagelles lui permettant une mobilité. C'est un micro-organisme majoritairement dominant du microbiote de l'estomac humain et qui infecte plus de la moitié de la population. La colonisation est généralement acquise dans l'enfance avant 10 ans et en l'absence d'antibiotique persiste à vie<sup>56,57</sup>.

*H. pylori* est une bactérie productrice d'uréase ce qui lui confère une adaptation au pH acide permettant de coloniser les cellules épithéliales de l'estomac. D'autres propriétés lui permettent aussi de s'adapter aux conditions de stress environnemental, ce qui explique sa virulence.

Cette colonisation de la muqueuse gastrique entraîne un environnement inflammatoire par des protéines cytotoxiques ce qui peut provoquer des lésions importantes, allant de l'inflammation aiguë ou chronique, aux ulcères gastroduodénaux et au cancer gastrique.

L'implication d'*H. pylori* de façon directe ou indirecte dans le développement du cancer gastrique a été largement démontrée et ce cancer correspond au 3<sup>ème</sup> cancer le plus mortel dans le monde<sup>58</sup>. Le traitement de ces cancers passe d'abord par l'éradication de cette bactérie. Une trithérapie incluant l'utilisation d'antibiotiques et d'inhibiteurs de pompe à protons est utilisée. Cependant, la résistance de *H. pylori* aux antibiotiques, notamment à la clarithromycine, est de plus en plus fréquente ce qui met à mal le schéma thérapeutique pour lutter contre cette bactérie et donc certaines des maladies qu'elle cause<sup>7</sup>. Il a été récemment démontré que *H. pylori* contient plusieurs copies d'une nouvelle famille de systèmes TA appelée *AapA/IsoA* dans lesquels *AapA* est la toxine (protéine) et *IsoA* l'antitoxine (ARN).

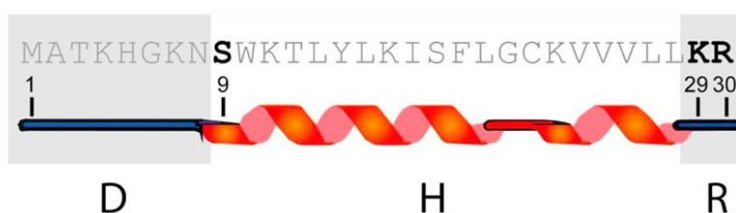
#### 1.4.3.2.2 Système TA *AapA/IsoA*

L'étude complète du transcriptome principal de *H. pylori* (souche 26695) a montré une très grande diversité génétique et notamment la présence de petits ARN non-codants dont la nouvelle famille de système TA de type I appelé *AapA/IsoA*<sup>59</sup>.

Ce système TA représente des petits modules d'expressions se trouvant à 6 endroits spécifiques du génome chromosomique de la bactérie avec une grande homologie structurale entre elles. Les modules *AapA1/IsoA1* et *AapA3/IsoA3* sont les plus exprimés et ainsi les plus étudiés.

#### I.4.3.2.2.1 Structure des toxines *AapA*

Chacun de ces petits ARNm *AapA* code pour une petite protéine de 23 à 30 acides aminés. Des études ont permis de montrer que la surexpression de *AapA1* induit la mort cellulaire de la bactérie sans provoquer de lyse cellulaire<sup>59</sup>. Ainsi, ces toxines semblent agir sur l'intégrité de membrane de la bactérie. Le peptide *AapA1* qui a été particulièrement étudié, correspond à une petite protéine hydrophobe qui contient un domaine transmembranaire alpha-hélicoïdal. A l'aide d'études de caractérisations biochimiques et de techniques biophysiques, la structure du peptide *AapA1* a ainsi été déterminée avec la spécificité de chacun de ses domaines (*figure 23*)<sup>60</sup>.



**Figure 23** : Structure du peptide *AapA1* comportant trois domaines distincts<sup>60</sup>. D = Désordonné, H = Hélice, R = Chargée positivement.

La toxine est composée de trois domaines :

Un domaine C-terminal où la présence d'acides aminés de charge positive est déterminante pour l'affinité à la membrane qui est de l'ordre du nanomolaire.

Une partie centrale comporte la séquence minimale pour obtenir une toxicité et forme une hélice-alpha qui lui confère une localisation transmembranaire.

Un domaine N-terminal déplié qui est désordonné et chargé positivement. Il ne contribue pas à la toxicité mais est important pour la reconnaissance des premiers nucléotides de l'antitoxine *IsoA1*.

Dans son ensemble, ce petit peptide possède une forte affinité pour la membrane. En pénétrant dans celle-ci, le peptide conduit à une désorganisation lipidique entraînant une diminution de la bicouche de la membrane à proximité du peptide. Cependant le mécanisme d'action affectant l'intégrité de la membrane et la toxicité reste encore mal compris.

Il a été suggéré dans une autre étude que l'activation de cette toxine est causée par un stress oxydatif par dégradation des antitoxines<sup>61</sup>. Ce stress enclencherait une transformation morphologique rapide des bactéries en forme de spirale en cellules coccoïdes rondes faisant référence à des cellules dormantes.

## I.4.3.2.2 Mécanisme de la traduction et de la répression de la toxine

La régulation de la traduction d'une toxine *AapA* se fait par l'interaction d'un petit ARN non-codant spécifique<sup>41</sup>. Cette régulation a été étudiée pour le système *AapA1/IsoA1* et peut être schématisée en plusieurs étapes allant de la traduction de la toxine à sa répression<sup>59</sup> (figure 24).

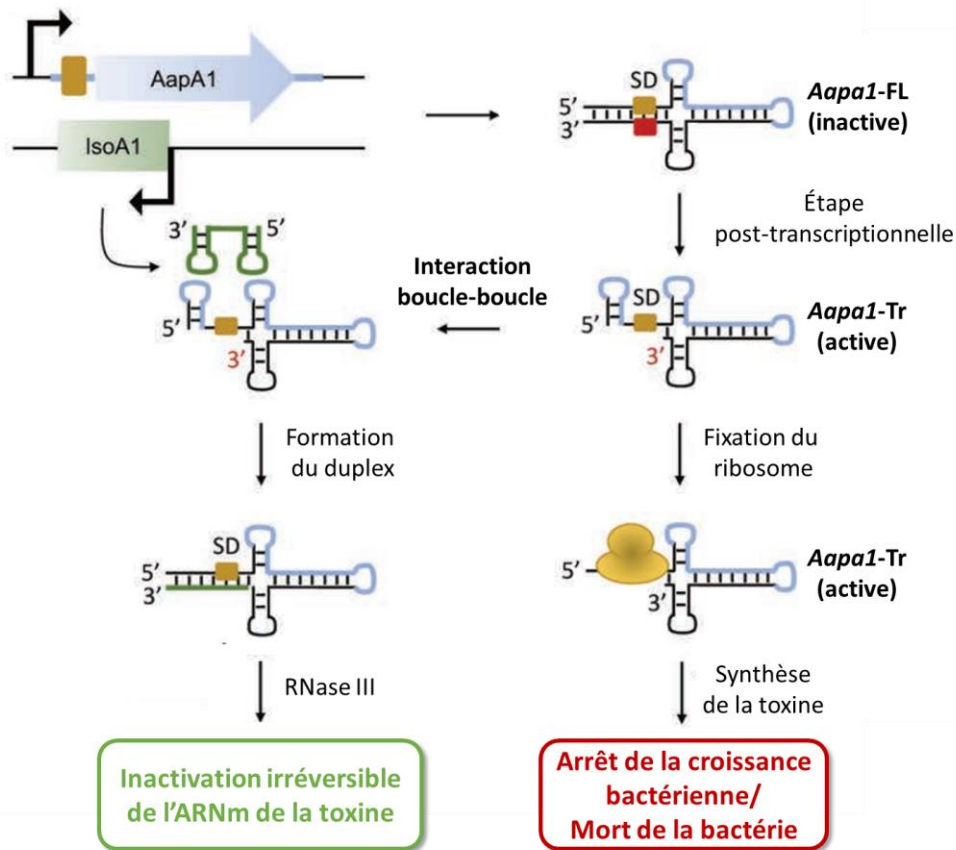


Figure 24 : Mécanisme de régulation de la toxine *AapA1* dans le système Toxine-Antitoxine *AapA1/IsoA1*<sup>62</sup>.

**Étape 1 :** D'abord, la transcription du gène codant *AapA1* génère un ARN stable et inactif (*AapA1-FL*). En effet, l'initiation de la traduction d'ARNm dépend de l'interaction d'une partie du ribosome sur une séquence en amont du codon de départ dite la séquence Shine-Dalgarno (SD)<sup>63</sup>.

**Étape 2 :** Pour rendre actif *AapA1-FL*, une étape post-transcriptionnelle est nécessaire dans laquelle le site de liaison du ribosome (*AapA1-Tr*) est disponible. A cette étape, en l'absence de l'antitoxine (*IsoA1*), la traduction de la toxine est possible provoquant l'arrêt de la croissance et la mort de la bactérie.

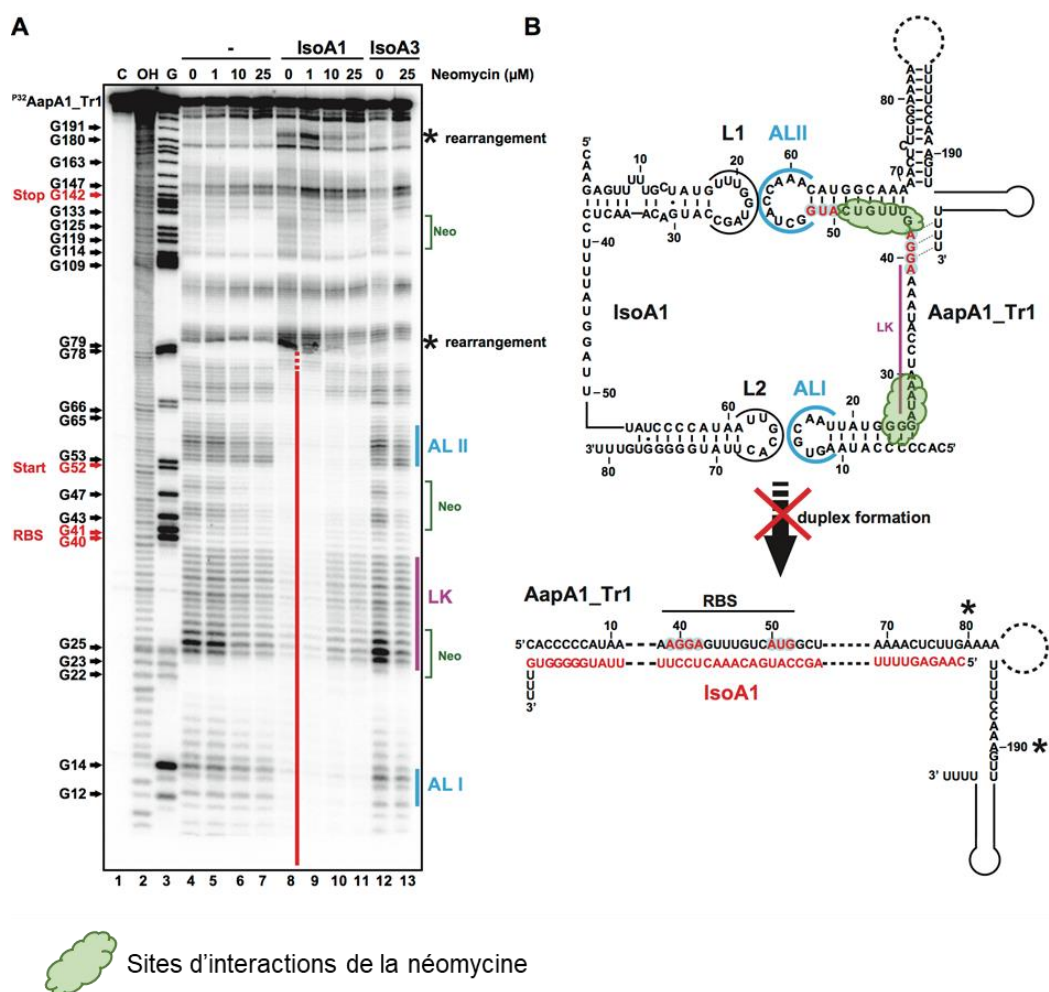
**Étape 3 :** Cependant, une fois l'ARNm activé et en présence d'*IsoA1*, leur interaction conduit à la formation de deux structures en tige-boucle *via* la formation d'interaction de type « kissing complex » ou d'interaction boucle-boucle.

**Étape 4 :** Cette interaction boucle-boucle va favoriser la formation d'un duplex étendu grâce à la complémentarité de l'*IsoA1* avec de la région de l'UTR 5' de l'ARNm de la toxine. Ce duplex étendu sera rapidement dégradé par une RNase III permettant un renouvellement rapide de l'ARNm *AapA1*.

Ainsi, l'ARN antitoxine *IsoA1* induit l'inhibition et la dégradation de la toxine *AapA1*. Ce mécanisme a été également validé pour le système *AapA3/IsoA3*<sup>64</sup>.

1.4.3.2.2.3 Validation de l'interaction boucle-boucle : formation du « kissing complex ».

Lors de l'étape 4 mentionnée précédemment, une interaction boucle-boucle est formée. Afin de confirmer la présence de ce complexe des expériences d'empreintes *in vitro* ont été réalisées en absence et en présence d'un antibiotique : la néomycine<sup>59</sup> (figure 25).



**Figure 25 :** A) Digestion au plomb de l'ARN *AapA1* (4) radiomarqué en absence et en présence de concentrations croissantes de néomycine (1, 10 et 25 μM). Les régions protégées par la néomycine sont indiquées en vert. En présence de *IsoA1*, le duplex étendu est formé (8) et en présence de néomycine celui-ci est partiellement inhibé (9, 10 et 11). B) Structure secondaire prévue pour la formation du « kissing complex ».

La néomycine fait partie de la famille des aminoglycosides largement étudiée et est illustrée dans la figure 11 (cf. page 35). Les aminoglycosides en général, et la néomycine en particulier, ont déjà été décrits pour leur capacité à interférer avec la formation de complexes boucle-boucle, notamment dans

la formation du complexe DIS (Site d'Initiation de Dimérisation) présent dans le génome de l'ARN du HIV-1 (cf. **Chap.I.4.**). La néomycine interagit avec le complexe en se fixant au niveau des boucles, stabilise le « kissing loop » et empêche ainsi la formation d'un duplex étendu indispensable à la prolifération virale.

Sur la base de ces résultats, la néomycine a été choisie comme composé modèle afin de confirmer la formation du complexe *AapA1/IsoA1 in vitro* ainsi que la possibilité d'inhiber la formation du complexe étendu. Des expériences de digestion de l'ARN en présence d'acétate de plomb ont été réalisées et une protection du clivage a été observée sur les sites de fixation de la néomycine. Ce composé prévient la formation de duplex étendu.

Ainsi, cette étude démontre, que l'étape limitante pour la formation du duplex étendu entre *AapA-Tr1* et *IsoA1*, passe par la formation d'un complexe boucle-boucle. Cette expérience montre également que des **ligands d'ARN**, tels que la néomycine, sont en mesure d'inhiber ces interactions dans ce système TA. Dans ce travail de thèse en collaboration avec l'équipe du Dr. Fabien Darfeuille au laboratoire ARNA, nous avons décidé d'explorer d'un point de vue fondamental la validation du système *AapA/IsoA* en tant que cible antimicrobienne potentielle et de concevoir des composés ligands d'ARN qui seraient capables de se fixer sur l'ARN et inhiber la formation du complexe boucle-boucle ou encore de le stabiliser.

Nous avons choisi ce système puisque comme mentionné précédemment *H. pylori* est une bactérie particulièrement agressive, très répandue et résistante à certains traitements. Par ailleurs, le système *AapA/IsoA* est capable d'induire la mort irréversible des cellules bactériennes et représente ainsi une cible de choix. Enfin, nos collaborateurs sont des experts dans l'étude de ce système et de la bactérie qui représente un modèle idéal afin de valider notre approche. **Les ligands d'ARN** synthétisés nous permettront d'étudier la faisabilité de cette approche antimicrobienne mais pourraient aussi représenter des nouvelles molécules pour une application en thérapie antibiotique.

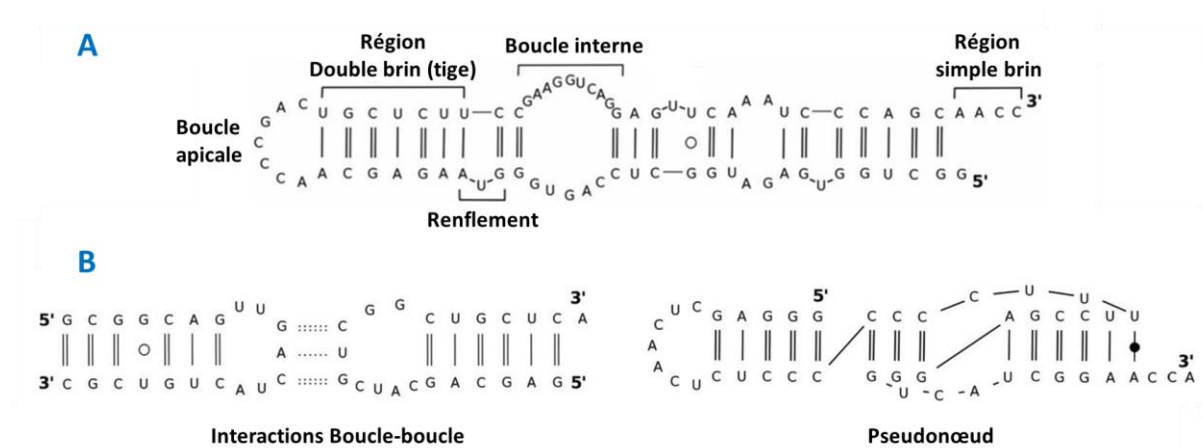
## I.5 Les ligands d'ARN

### I.5.1 ARN d'intérêt thérapeutique

Au cours des dernières décennies, le rôle de l'ARN dans un grand nombre de processus biologiques tels que la transcription, la traduction mais aussi la régulation de l'expression des gènes, a été largement mis en évidence. Grâce à une meilleure compréhension de l'ensemble de ces processus, le domaine de la biologie de l'ARN émerge comme un champ essentiel pour la compréhension d'un grand nombre de phénomènes et la pandémie actuelle causée par le SARS-CoV2, un virus possédant un génome ARN, souligne encore plus l'intérêt d'étudier le fonctionnement de cette macromolécule ainsi que son ciblage dans un but thérapeutique.

L'ARN est majoritairement connu pour ses rôles dans la transcription et la traduction incluant la famille des ARN messagers (ARNm), les ARN de transfert (ARNt) et les ARN ribosomiques (ARNr). Ces deux derniers ont été précédemment cités comme cibles de plusieurs classes d'antibiotiques. Néanmoins, d'autres ARN appartenant à la famille des ARN non-codants (ARNnc) possèdent des fonctions essentielles de régulation. Chez l'humain, les ARNnc représentent plus de 70% de son génome<sup>65</sup>.

Les ARNnc sont constitués de structures très dynamiques et présentent des structures secondaires et tertiaires variables et complexes. La structure secondaire est constituée de régions de doubles brins et de régions simple brin qui représentent des motifs uniques comme des renflements ainsi que des boucles apicales ou internes (*figure 26*).



**Figure 26 :** A) Représentation des différentes structures secondaires de l'ARN. B) Exemples de structure tertiaire : interaction boucle-boucle et pseudonœud<sup>66</sup>.

La formation de ces structures secondaires peut entraîner des interactions dans l'espace et ainsi conduire à des structures tertiaires comme des pseudonœuds très complexes, des interactions boucle-boucle ou encore la formation de structure G-quadruplex. Ces topologies distinctes,

tridimensionnelles, peuvent ainsi offrir des sites de reconnaissance pour des peptides (ligands naturels de l'ARN) ou des petites molécules.

Les divers rôles biologiques attribués à l'ARN sont présents dans les processus de l'expression ou de la réplication chez les bactéries et les virus. Pour le génome humain, les ARNnc régulent un bon nombre de processus cellulaires et revêtent donc des fonctions physiologiques essentielles. Cependant, certains se retrouvent à être dérégulés (surexprimés ou sous-exprimés) et ont donc été directement liés au développement de nombreuses maladies, telles que les cancers ou des maladies neurodégénératives.

Récemment, l'amélioration de la connaissance du domaine des ARN a été mise en avant par de nouvelles innovations notamment avec de nouveaux vaccins contre le coronavirus du syndrome respiratoire aigu sévère (SARS) utilisant un ARNm codant des protéines de surface, ou avec le risdiplam qui cible l'épissage de l'ARNm dans la maladie d'amyotrophie spinale de type 1 (SMA)<sup>67</sup>.

Le fait que plusieurs molécules soient d'ores et déjà sur le marché en tant que médicament, telles qu'un grand nombre d'antibiotiques ciblant le ribosome bactérien ou encore le risdiplam entré sur le marché en 2020, montre que l'étude de nouvelles approches thérapeutiques basées sur le ciblage de l'ARN représente une stratégie particulièrement prometteuse. Le potentiel thérapeutique repose actuellement sur deux approches comprenant **l'utilisation d'oligonucléotides** ou **l'utilisation de petites molécules**.

## I.5.2 Stratégies de ciblage d'ARN : oligonucléotides et petites molécules

### I.5.2.1 Approche par des oligonucléotides

L'approche basée sur le ciblage des ARN par des oligonucléotides repose sur l'appariement des nucléobases via des interactions Watson–Crick spécifiques avec un ARN cible dans le but de dégrader ou d'inhiber le fonctionnement de celui-ci. Différentes classes d'oligonucléotides sont regroupées selon leurs caractéristiques. Les aptamères agissent mécaniquement sur des cibles protéiques ou d'ARN, tandis que les petits ARN interférents ou les oligonucléotides antisens, interfèrent directement avec certains groupes d'ARN. Ces oligonucléotides peuvent être capables de recruter des enzymes pour induire le clivage de l'ARNm cible<sup>68</sup>.

En dépit de leur promesse, les oligonucléotides non modifiés sont rapidement dégradés par des nucléases cellulaires *in vivo*. Pour améliorer leur stabilité métabolique ainsi que leur efficacité de pénétration cellulaire, différentes modifications ont pu être apportées telles que les modifications du pyranose en position 2', le verrouillage des acides nucléiques (« locked nucleic acids » ou LNA), la modification du lien phosphodiester ou le changement du squelette. Grâce aux différentes

modifications, l'approche basée sur l'utilisation d'oligonucléotides a amené à plusieurs succès thérapeutiques et certains de ces outils sont utilisés en clinique<sup>69</sup> (*tableau 3*).

**Tableau 3** : Quelques exemples non exhaustifs des différentes classes d'oligonucléotides en phase clinique ou actuellement approuvés :

Nom	Type	Maladie ciblée	Phase Clinique
<b>Patisiran</b>	siARN	Amyloïdose héréditaire à transthyrétine (hATTR)	Approuvé
<b>Nusinersen</b>	Antisense (AMO)	SMN2 intron 7 spinal muscular atrophy	Approuvé
<b>Miravirsen</b>	AntagomiR (AMO spécifique aux anti-microARN)	miR-122 hépatite C (VHC)	Phase clinique II
<b>Pegaptanib</b>	Aptamer	Dégénérescence maculaire liée à l'âge néovasculaire (DMLA)	Approuvé

Parmi ces oligonucléotides, certains doivent former un complexe avec la cible ARN pour exercer une activité biologique ce qui limite leur efficacité uniquement aux régions non structurées et donc accessibles. De plus, du fait de leur poids moléculaire élevé et de leur charge négative, la délivrance et la biodistribution nécessitent souvent des injections intraveineuses ou sous-cutanées, mais aussi intra-oculaires ou dans la moelle épinière. Les paramètres pharmacocinétiques et pharmacodynamiques continuent donc de limiter leur utilisation thérapeutique même si des progrès dans ce domaine sont en cours.

Une alternative à ces différents inconvénients notamment pour **la reconnaissance de domaines hautement structurés**, peut être le développement de petites molécules ciblant ces potentiels sites de liaison.

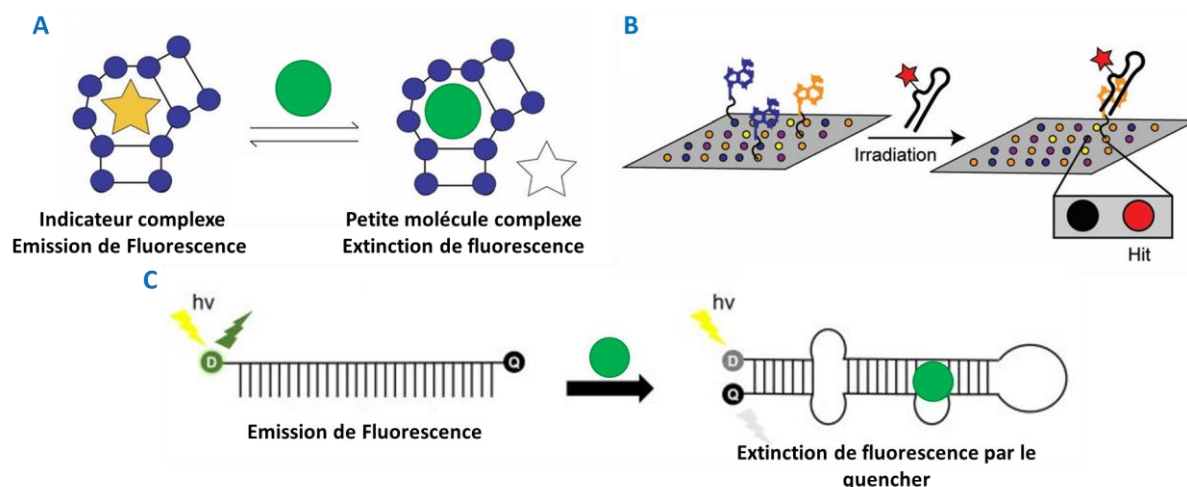
### 1.5.2.2 Approche par des petites molécules : des stratégies multiples

Une grande partie des molécules visant l'ARN a été découverte par criblages de chimiothèques pour découvrir des squelettes chimiques capables d'interagir avec l'ARN. Différentes librairies académiques ou industrielles et de tailles variables ont été étudiées<sup>70</sup>. Différentes approches de criblage peuvent être utilisées : criblage ciblé (FcS), criblage à haut débit (HTS pour « High-Throughput Scening ») et HTS suivi de l'optimisation des leads (HTS-LO). Ces collections peuvent contenir des composés en solution ou retenus sur une surface<sup>71</sup>.

Les méthodes le plus souvent utilisées lors des criblages reposent sur la mesure de l'affinité de liaisons relatives par des tests cellulaires *in vitro* basés sur la sensibilité de fluorescence, d'anisotropie ou de luminescence ou encore par des tests *in silico*. Différentes méthodologies peuvent être utilisées telles que l'intégration de marqueur de fluorescence directement sur l'ARN cible, l'utilisation d'un couple de



transfert d'énergie de résonance de Förster (FRET) (DNA beacon) ou encore le dosage par un test de déplacement d'indicateur de fluorescence (FID)<sup>72</sup> (figure 27).



**Figure 27 :** Représentation de certains tests basés sur la fluorescence reportés dans la littérature. **A)** Principe du test de dosage par FID : l'affinité de la liaison est mesurée par une petite molécule d'intérêt « déplaçant » une molécule marquée indicatrice de liaison à l'ARN, ce qui provoque un changement de fluorescence de l'indicateur. **B)** Principe du microréseau à petites molécules (SMM) : irradiation des ARN labelisés par un fluorophore détectant la liaison de la molécule à l'ARN par une variation de fluorescence. **C)** Principe du test utilisant un système FRET : Extinction de la fluorescence par rapprochement du quencher (Q) et du fluorophore (D = donneur).

Afin d'étudier le mode d'interaction des ligands identifiés après criblage, des analyses complémentaires sont également utilisées<sup>66</sup>. Les méthodes physico-chimiques primaires permettant l'étude d'un état statique comprennent l'étude de la structure de l'ARN par rayon X ou par Cryo-EM. La dynamique des complexes peut être étudiée par RMN en solution. Des paramètres thermodynamiques ( $T_m$ , ITC) et cinétiques (SPR : résonance plasmonique de surface) peuvent aussi être déterminés par des études biophysiques. Cependant ces techniques ont des limites qui comprennent notamment le temps et la quantité d'échantillon nécessaire.

La compréhension fondamentale de ce qui entraîne la sélectivité au niveau des petites molécules reste néanmoins un défi. Ces dernières années, différentes équipes essaient de développer d'autres technologies afin d'en comprendre un peu plus sur le potentiel caractère sélectif des petites molécules afin de déterminer des motifs de reconnaissance de l'ARN.

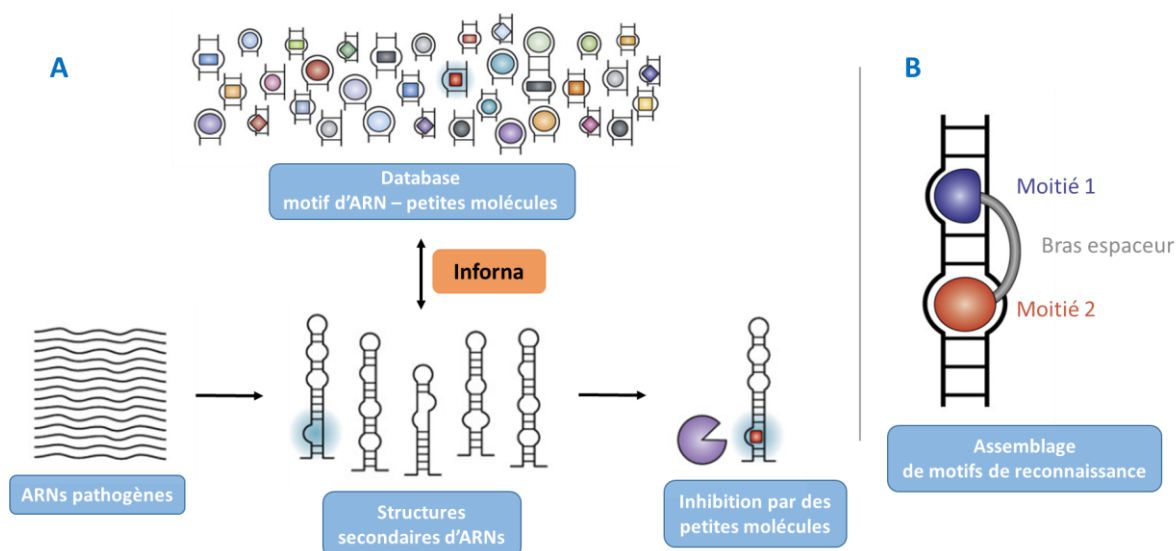
L'équipe de Nickbarg (Merck & Co.) utilise un Système Automatisé d'Identification de Ligands (ALIS) pour effectuer un criblage haut-débit de molécules en utilisant de la chromatographie liquide-spectrométrie de masse (LC-MS) sur un ensemble varié de 42 ARN<sup>73</sup>. Cette méthode mène ensuite à une analyse par composantes principales comparant les ligands obtenus à des ligands formant des interactions protéine/petites-molécules. **L'abondance d'hétérocycles** contenant une amine

aromatique ou des motifs de type amidine est ressortie comme une des caractéristiques de sélectivité envers les cibles ARN.

Hargrove et al. ont analysé les ligands d'ARN connus dont les aminoglycosides par des méthodes chemo-informatiques et par des analyses par composantes principales, afin d'en ressortir des propriétés physico-chimiques, structurales et spatiales<sup>74,75,76</sup>. La reconnaissance des ligands d'ARN vis-à-vis de leur cible semblerait privilégier des interactions de type **liaisons hydrogènes** et **d'empilements des bases**, tandis que la reconnaissance des molécules ciblant les protéines est plus propice à des interactions hydrophobes.

Ce même groupe a également réalisé la mise en place d'une base de données des ligands connus pour cibler l'ARN appelée R-BIND<sup>77</sup>. L'étude des structures contenues dans cette base de données a permis de faire ressortir différentes caractéristiques propres aux ligands d'ARN par rapport aux ligands protéiques telles que : **un nombre plus élevé d'atomes d'azote, un nombre plus faible d'atomes de carbone hybridés  $sp_3$  et de stéréocentres.**

De manière intéressante, l'équipe du professeur Disney a développé une plateforme informatique appelée **Inforna** (figure 28, A). Cet outil comprend un criblage combinatoire bidimensionnel (2DCS)<sup>78</sup> qui permet le criblage d'un grand nombre de motifs structuraux secondaires d'ARN et de larges chimiothèques de composés. Puis, l'ensemble des données obtenues après criblage est organisé par une approche appelée « Structure Activity Relationships Through Sequencing » (StARTS)<sup>79</sup>. Enfin, un score est attribué aux meilleures combinaisons ligand/structure d'ARN afin d'appliquer la stratégie à une cible spécifique. La combinaison de ces trois étapes a donné naissance à Inforna<sup>80</sup>.



**Figure 28 :** A) Illustration de l'identification de sonde chimique par l'outil Inforna pour un ARN pathogène cible. B) Conceptions de ligands multivalents liant deux structures chimiques reconnaissant un motif spécifique de la structure secondaire de l'ARN cible<sup>81</sup>.

Grâce à cette méthodologie et à partir d'un ARN d'intérêt, par la comparaison de structures secondaires présentes dans celui-ci, il est possible de concevoir des sondes chimiques sélectives. Cette approche a été utilisée dans plusieurs études notamment pour le ciblage des miRNA<sup>82</sup>. Cet outil a permis également de concevoir des « **ligands multivalents** » permettant de rendre le ligand plus sélectif pour l'ARN cible (*figure 28, B*).

D'autres équipes, notamment le groupe Pr. Zimmerman, utilisent l'approche de composés multivalents pour cibler des ARN comportant des répétitions de trinucleotides, ce qui a permis d'améliorer significativement la puissance du ligand monovalent<sup>83</sup>.

En définitive, l'ensemble de ces techniques dont le criblage virtuel, le criblage à haut débit ou encore la modélisation moléculaire, ont amené à la conception et/ou l'identification de petites molécules pouvant se lier à des cibles ARN de manière efficace et spécifique. Cet engouement pour aider à concevoir des ligands sélectifs est une source d'inspiration et tous les travaux futurs dans ce domaine permettront d'améliorer la conception de nouveaux ligands.

Dans la littérature depuis ces vingt dernières années, un nombre croissant d'études sur des petites molécules ligands d'ARN est reporté. Les ligands d'ARN connus peuvent cibler des ARN contenus dans **des bactéries, des virus et des eucaryotes**. Ci-après, un certain nombre d'exemples caractéristiques de ces différentes cibles ainsi que la contribution des travaux amenés par notre équipe et qui ont permis de sélectionner des éléments clés dans la conception du projet de ce manuscrit seront décrits.

### 1.5.3 Ciblage d'ARN bactériens

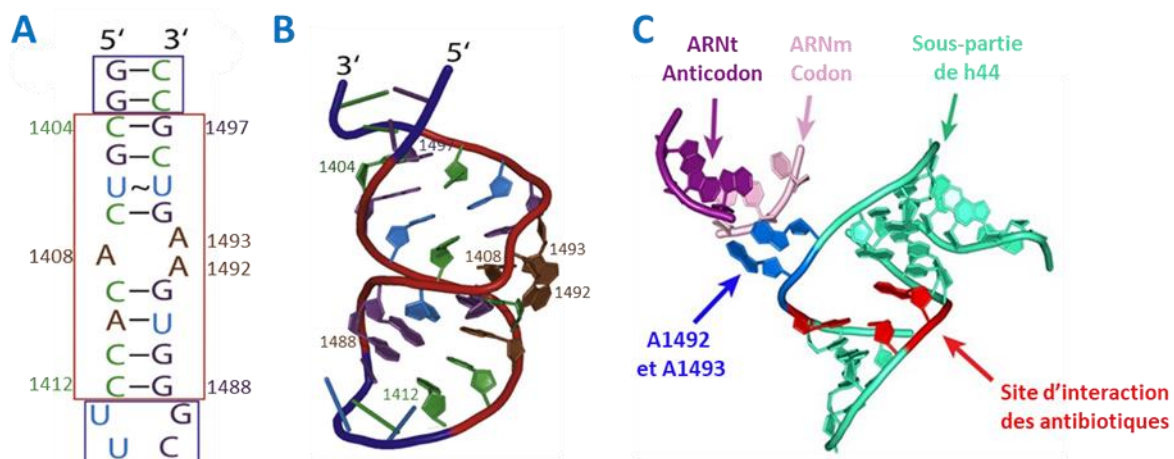
Depuis de très nombreuses années, l'ARN ribosomique (ARNr) procaryote est la cible d'un grand nombre d'antimicrobiens naturels, semi-synthétiques ou synthétiques (par exemple, les aminosides, les tétracyclines, les macrolides, les oxazolidinones et les phénicolés) afin d'inhiber la synthèse protéique et ainsi la prolifération bactérienne (cf. **Chap. 1.2.5**). Les ARNr sont un composant majoritaire de la cellule bactérienne (90% de tout l'ARN bactérien) ce qui en fait une cible privilégiée.

La compréhension des mécanismes moléculaires d'action des antibiotiques agissant sur le ribosome bactérien a été un moteur majeur pour la découverte et le développement de ces antibiotiques ainsi qu'une source d'inspiration pour le développement de molécules ciblant l'ARN. Des études cristallographiques sur le ribosome procaryote menées par Ramakrishnan, Steitz et Yonath ont permis de comprendre le fonctionnement de cette machine macromoléculaire et l'importance des ARNr (5S, 16S et 23S) qui la constitue. Cette découverte leur a valu le prix Nobel de Chimie en 2009<sup>84</sup>.

D'autres études cristallographiques à l'aide des rayons X et de la spectroscopie RMN en solution ont pu révéler le fonctionnement de nombreux antibiotiques aminoglycosides, notamment ceux comportant un noyau 2-déoxystreptamine dans le site de décodage dans la partie 30S<sup>85,86,87</sup>.

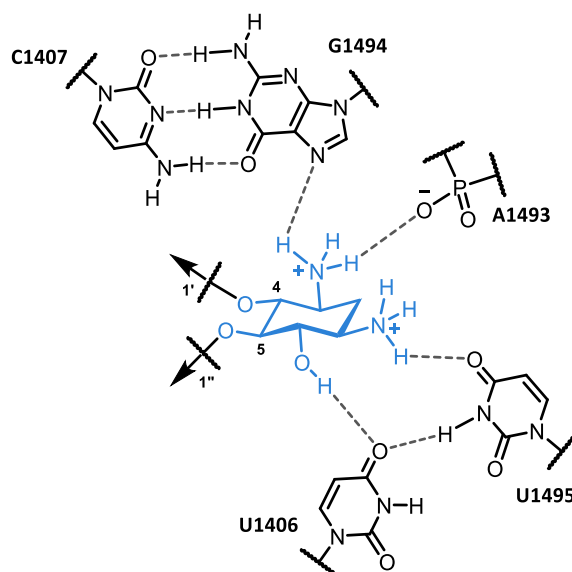
Pour rappel, le site A possède deux conformations en équilibre fonctionnant comme un interrupteur. La position « fermée » est inactive, tandis que la position « ouverte » oriente les deux adénines 1492 et 1493 vers l'extérieur ce qui permet la reconnaissance des bases vers le complexe codon-anticodon. Les interactions des antibiotiques ciblant ce site A vont alors stabiliser la conformation « active » ce qui induit une lecture aberrante de l'ARNm lors de l'étape de décodage (*figure 29*).

Le **site A** se trouvant dans l'hélice 44 (h44), peut être représenté par une structure en tige-boucle comprenant une région double brin et un renflement interne comportant des bases non appariées (*figure 29*).



**Figure 29 :** A) Structure secondaire du site A chez *E. Coli*. La boîte rouge représente les nucléotides présents dans le site d'interaction<sup>88</sup> alors que les autres résidus sont ajoutés lors des études *in vitro*. B) Structure tridimensionnelle du site A dans la position « fermée ». C) Structure tridimensionnelle du site A dans la position « active »<sup>89</sup>.

Des études structurales de la classe des aminoglycosides de type 2-déoxystreptamine notamment pour la néomycine et la paromomycine ont permis de souligner l'importance des cycles I et II de ces aminoglycosides (cf. *figure 11*, page 35) dans des interactions clé se trouvant dans la poche du site A<sup>87</sup> (*figure 30*). En revanche, les anneaux III et IV contribuent généralement moins aux interactions spécifiques avec la cible.

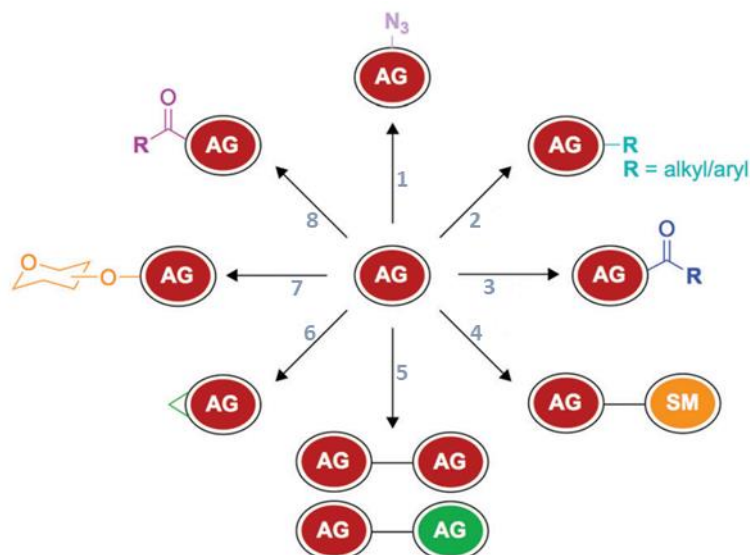


**Figure 30 :** Illustration de l'interaction du cycle II (2-déoxystreptamine) de la néomycine B dans la poche de liaison du site A montrant l'importance du cycle 2-déoxystreptamine pour l'interaction<sup>87</sup>.

Ces interactions sont de type électrostatique et sont dues aux charges cationiques portées par les groupements amines qui sont protonés dans les conditions physiologiques.

Néanmoins, ces aminoglycosides sont la cible de plusieurs phénomènes de résistance spécifique induite par différentes enzymes comme détaillé à la section **Chap.I.1.3**. Outre les phénomènes de résistance, les aminoglycosides ont une toxicité élevée en raison d'une mauvaise métabolisation et des dommages cellulaires observés induisant une néphrotoxicité et une ototoxicité importante. Cependant, les aminoglycosides restent une classe d'antibiotiques très utilisée dans les traitements du fait du large spectre d'activité.

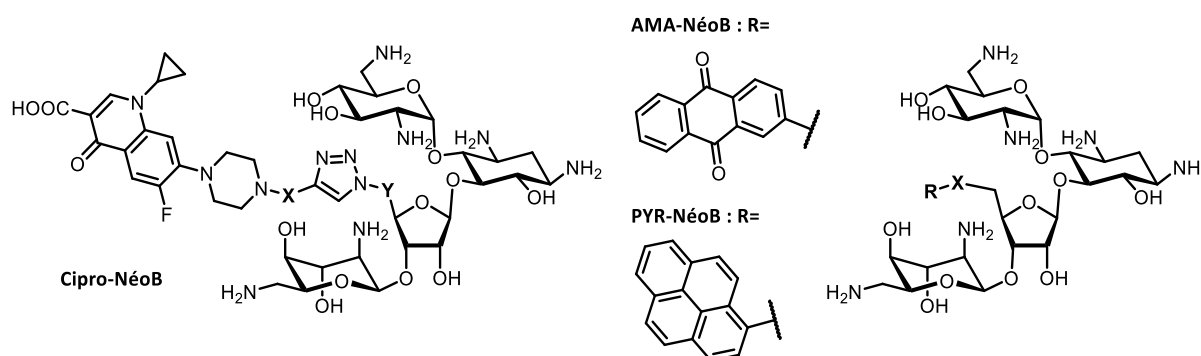
Dans ce contexte, et à l'aide de ces études structurales, de nombreux groupes de recherches ont étudié les aminoglycosides pour leur modification chimique ou leur conjugaison à d'autres motifs afin de tenter de pallier ces différents problèmes (*figure 31*).



**Figure 31 :** Vue générale sur les différentes modifications réalisées sur les composés aminoglycosides (AG) : (1) modification des noyaux AG, (2) introduction de divers substituants alkyles/aryles et (3) de groupement acyle, (4) développement d'hétéroconjugués, (5) développement d'AG homo- et d'hétérodimères, (6) formation d'AG avec une conformation restreinte, (7) création de nouveaux AG par ajout de glycoside supplémentaire, (8) développement de nouveaux AG par des réactions chemo-enzymatiques.

Ces stratégies ont permis de diminuer la toxicité et de réduire l'action de certaines enzymes notamment lors du développement de différents analogues semi-synthétiques de la kanamycine et de la gentamycine tels que la plazomicine approuvée en 2018<sup>25</sup>.

Pour les analogues de la néomycine, de nombreuses modifications ont été apportées dont la conjugaison de divers motifs sur la position 5'' du cycle III à l'aide de différents bras espaceurs. Quelques exemples sont présentés ci-après (figure 32).



**Figure 32 :** Structures des conjugués néomycine décrites par l'équipe de Baasov et l'équipe de Arya. X et Y représentant différents bras espaceurs.

En 2009, l'équipe du Pr. Baasov a synthétisé un composé hybride conjugué entre la ciprofloxacine et la néomycine<sup>90</sup> (figure 32, Cipro-NéoB). Ces deux antibiotiques sont reliés par un bras de type triazole comportant une longueur de chaîne et des fonctions variables de part et d'autre. Comparé à la néomycine, cet assemblage a permis une amélioration de l'inhibition de la croissance bactérienne (les

bactéries à Gram négatif et le *Staphylococcus aureus* résistant à la méticilline (SARM) à Gram positif), mais reste inférieur en comparaison de l'activité de la ciprofloxacine.

Les conjugués d'antraquinone (*figure 32, AMA-NéoB*) ont été synthétisés en 2017 par l'équipe d'Arya avec des bras espaceurs de type thio-urée<sup>91</sup>. Dans cette étude, l'affinité envers le site A a été mesurée par des tests de déplacement avec la néomycine comportant une sonde fluorescente (F-Néo) *via* la sensibilité de la fluorescence. Outre ce test, des évaluations de l'inhibition de la croissance (CMI) ont été réalisées sur des bactéries résistantes à la néomycine contre les AMEs. Ainsi, ces conjugués AMA-NéoB ont démontré une affinité pour le site A comparable à la néomycine, mais leurs propriétés antibactériennes et leur capacité à agir comme substrats contre ces AMEs sont variables selon le bras espaceur. D'autres conjugués ont été élaborés par la même équipe et plus récemment, des conjugués néomycine-pyrène<sup>92</sup> (*figure 32, PYR-NéoB*). Même si l'affinité a été améliorée, l'activité inhibitrice n'a pas été suffisante contre certaines souches résistantes.

Malgré des stratégies diverses et considérables, aucune de ces modifications n'a abouti à de nouveaux candidats médicaments. Cependant, tous ces efforts de synthèse ont permis, par ailleurs, de montrer la capacité des aminoglycosides à interagir avec l'ARN et qu'il est possible de les détourner de leur cible principale pour leur permettre d'interagir avec d'autres cibles ARN<sup>93</sup>.

Une autre façon de cibler les ARN bactériens est de faire appel aux **riboswitchs** décrits pour la première fois en 2002 et considérés comme des cibles antibactériennes potentielles par le fait qu'ils sont principalement retrouvés chez les bactéries<sup>94,95</sup>. Ces domaines résident dans la région 5' non traduite (UTR) d'ARN messagers et régulent l'expression du gène induite par un ligand spécifique endogène. Ces ARN sont impliqués dans la biosynthèse et le transport de métabolites essentiels comme des coenzymes ou des nucléobases et leurs dérivés.

La régulation de ces ARN est possible par la présence d'une poche de liaison sélective pour une petite molécule dans la partie aptamère. La liaison du ligand engendre un changement conformationnel qui active ou réprime l'expression de l'ARNm d'où le nom d'interrupteurs génétiques (ribo-commutation = riboswitch). Par ailleurs, les ARNm sont la cible de composés naturels couvrant une large gamme de molécules comprenant les ions, la S-adénosylméthionine, les acides aminés, la glucosamine-6-phosphate, le pyrophosphate de thiamine ou encore le mononucléotide de flavine (FMN).

Le mécanisme de rétroaction négative permet de moduler le métabolisme de la bactérie et fait des riboswitch une cible attrayante. Le riboswitch à FMN (Flavine MonoNucléotide) a particulièrement été étudié et est impliqué dans la régulation de l'expression de gènes dans la biosynthèse et le transport des riboflavines qui sont essentielles à la croissance des bactéries. En 2015, Howe *et al.* ont effectué

un criblage d'environ 57 000 molécules par la méthode ALIS. Le composé **ribocil** est ressorti comme pouvant être un ligand synthétique distinct par rapport au ligand naturel<sup>96</sup> (figure 33).

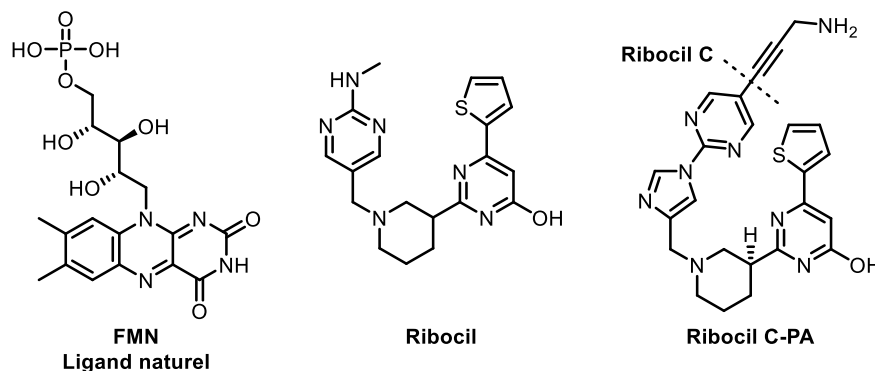


Figure 33 : Structure chimique du ligand naturel pour le riboswitch à Flavine MonoNucléotide.

Diverses caractérisations biophysiques, telles que la cristallographie, ont amené à la découverte d'autres dérivés capables de cibler cet ARN avec une affinité similaire de l'ordre du nanomolaire<sup>97</sup>. Une modification remarquable a été rapportée par Hergenrother et *al.* qui ont synthétisé le **ribocil C-PA** après une étude de relations structure-activité. Cette modification a permis une augmentation de la concentration intracellulaire du ribocil C-PA chez *E. coli* entraînant une amélioration de l'activité antimicrobienne (CMI de >64 à 4 µg/mL)<sup>98</sup>.

Malgré l'observation de résistances fréquentes lors du ciblage des riboswitchs, dans une perspective plus large, ces travaux ont montré la capacité de pouvoir cibler des ARN par d'autres molécules que les aminoglycosides et la possibilité d'en améliorer l'affinité ainsi que les propriétés pharmacocinétiques.

### I.5.4 Ciblage d'ARN viraux

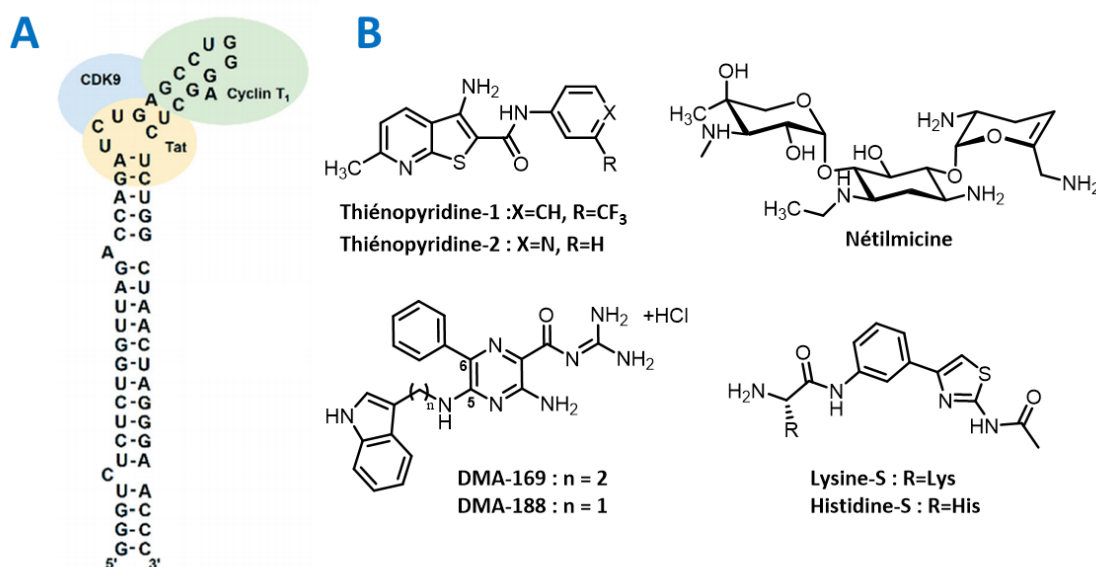
Les ARN non-codants présents chez les virus représentent de nombreuses opportunités de ciblage par des petites molécules, notamment ceux qui se trouvent dans des régions hautement conservées du génome et des transcrits. Actuellement, les ARNnc les plus étudiés sont retrouvés dans le virus de l'immunodéficience humaine (VIH), dans le virus de l'hépatite C (VHC), dans le coronavirus du syndrome respiratoire sévère (SRAS-CoV) et dans le virus de la grippe A. Les ligands ciblant ces virus se lient à différentes architectures des ARNnc affectant la stabilité structurale ou le changement conformationnel permettant de bloquer des processus essentiels à la réplication virale<sup>99,100,101</sup>.

Dans cette partie, il sera essentiellement détaillé deux ARN cibles du virus de l'immunodéficience humaine que sont les éléments **TAR** et **DIS**. En effet, le génome du VIH plusieurs ARN régulateurs essentiels dans plusieurs processus tels que la régulation transcriptionnelle, la transcription inverse, la traduction des protéines virales ou encore le transport nucléocytoplasmique<sup>102,103</sup>.



L'ARN **TAR** pour « Trans-Activation Response element » est essentiel dans la réplication efficace du VIH. Ce fragment d'ARN de 59 nucléotides est situé dans les régions 3'-UTR de chaque ARN viral et présente une structure tige-boucle avec un renflement (*figure 34, A*). L'ARN TAR interagit avec la protéine virale Tat et avec le facteur d'élongation P-TEFb lui-même constitué de la cycline T1 et de la kinase 9 dépendante de la cycline (CDK9), contrôlant ainsi la transcription. La spécificité de l'interaction de Tat est médiée par une région riche en arginine qui reconnaît le site du renflement. Ce dernier est propice à l'interaction de différents petits ligands capables de bloquer l'interaction TAR-Tat afin d'inhiber la réplication virale.

L'ARN TAR représente donc non seulement une cible antivirale prometteuse mais aussi un modèle pour l'étude de ligands d'ARN puisqu'il a été largement étudié. Par exemple, l'affinité de nombreux aminoglycosides a été étudiée envers cet ARN<sup>93</sup>. Récemment, plusieurs nouveaux composés ont été découverts ciblant cet élément régulateur (*figure 34, B*).



**Figure 34 :** A) Représentation de la structure secondaire de l'ARN TAR possédant le site de liaison du peptide Tat. B) Structures chimiques de molécules se liant à l'ARN TAR.

Parmi les ligands découverts récemment, l'équipe du Pr. Schneekloth a effectué un criblage de 20000 molécules et déterminé la sélectivité de ces molécules par un test utilisant un ARN TAR conjugué à un fluorophore<sup>104</sup>. A la suite de ce criblage, la molécule **thiénopyridine-1** est ressortie avec une affinité de 2,4  $\mu\text{M}$  ( $K_D$ ) pouvant ainsi être définie comme ligand d'ARN. L'activité antivirale envers le VIH a été mesurée dans les lymphoblastes T (CEM-SS) induisant un effet cytopathique  $CE_{50}$  de 28  $\mu\text{M}$  sans toxicité même à des concentrations allant jusqu'à 1 mM.

Des modifications de ce squelette (**thiénopyridine-2**) ont permis d'augmenter la solubilité permettant de réaliser des analyses structurales par RMN en solution mettant en avant un mode de liaison dans

une région inférieure du renflement, probablement dans le petit sillon mineur et ce site diffère de celui des autres classes de molécules étudiées pour cibler cet ARN <sup>105</sup>.

Une autre approche permettant la découverte de nouveaux chémotypes a été envisagée par un ensemble d'études dynamiques de l'ARN TAR obtenues par simulation de dynamique moléculaire (MD) ensuite affinées par spectroscopie RMN<sup>106</sup>. L'équipe du Pr. d'Al-Hashimi, qui a développé cette méthode, a étudié une large chimiothèque de composés et a identifié six nouveaux ligands d'ARN dont le **nétilmicine** et le **diméthylamiloride** (DMA) avec une affinité de 1,35  $\mu\text{M}$  et 122  $\mu\text{M}$ <sup>107</sup>. Le nétilmicine a montré une sélectivité pour TAR en présence d'un grand excès d'ARNt et est capable d'inhiber l'activation médiée par la protéine Tat de 81% dans les lignées cellulaires infectées par l'isolat VIH-1 NL4-3.

Par ailleurs, le composé DMA ayant une affinité faible envers la cible TAR, l'équipe d'Hargrove a décidé de générer une nouvelle série d'analogues en modulant diverses positions (C(5) et C(6)) donnant le composé **DMA-169**<sup>108</sup> (*figure 34, B*). Un test de déplacement de la protéine Tat a permis par ailleurs de mesurer la capacité d'inhibition de ce composé *in vitro* de 4,4  $\mu\text{M}$ . La réduction de la flexibilité par le nombre de carbone  $sp^3$  en position (5) a donné le composé **DMA-188** qui a conduit à une augmentation de 2 à 3 fois pour l'affinité de l'ARN TAR ainsi qu'à une sélectivité relativement améliorée<sup>109</sup>.

Au travers de cette étude, l'équipe du Pr. Hargrove a comparé la sélectivité sur différents ARN du VIH. Un de ces analogues a pu être mis en avant pour sa sélectivité envers l'ARN ESSV. Cette ARN a été identifiée comme une cible critique pour l'épissage de l'ARNm viral et la production virale<sup>110</sup>.

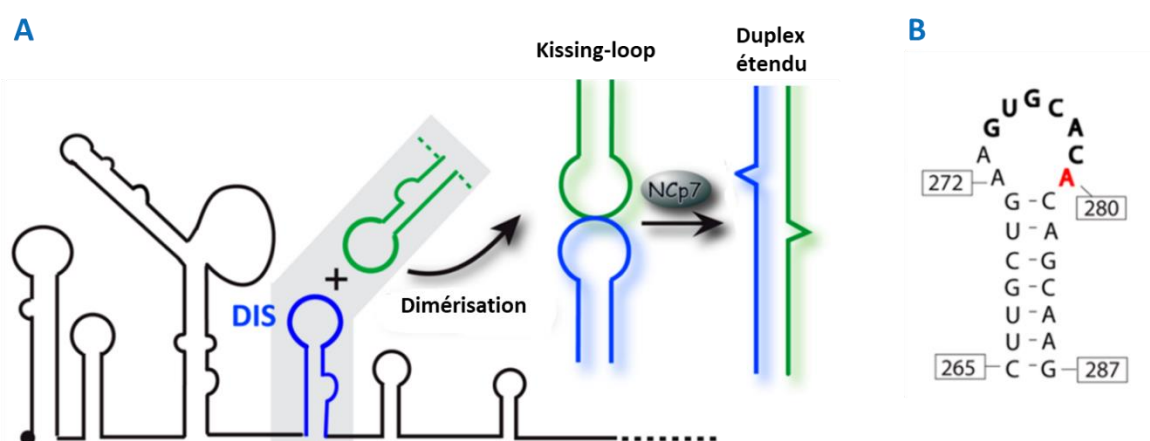
Notre équipe a également contribué à la conception de nouveaux squelettes permettant de cibler l'ARN TAR. Au niveau de la jonction de la tige et du renflement de la séquence TAR, on distingue une paire de bases AU en position 22 et 44 (*figure 34, B*). Afin de cibler cette paire de nucléobases, des conjugués comportant une nucléobase artificielle précédemment découverte par l'équipe du Dr Benhida, ont permis de former un triplex S-AU à l'aide d'interactions de Hoogsteen<sup>111</sup>.

Dans une première série de conjugués<sup>112</sup>, la nucléobase a été assemblée avec un motif 2-désoxyribose comprenant des esters d'acides aminés sur les positions 3' et 5'. Malheureusement, cette classe de composés n'a donné aucune activité dans les analyses cellulaires du VIH en raison d'une dégradation rapide des liaisons esters labiles. Suite à cela, un ensemble de 14 conjugués de mono- ou di- acides aminés avec la nucléobase S a été étudié<sup>113</sup>. L'affinité de ces molécules a été testée par la mesure de la fluorescence par titration avec l'ARN TAR comportant un fluorophore en position terminale. Seul le conjugué **histidine-S** a montré une affinité de 17,5  $\mu\text{M}$  et une sélectivité en présence d'ARNt et d'ADN en excès. Le conjugué **lysine-S** a quant à lui donné une liaison plus affine envers la cible ( $K_D = 7,5 \mu\text{M}$ ),

mais a perdu un peu de sa sélectivité en présence d'ARNt et d'ADN. A l'aide d'études thermodynamiques, il a pu être montré la contribution spécifique des liaisons hydrogènes de la nucléobase S et que l'ajout à des positions optimales d'acides aminés a permis d'obtenir des interactions spécifiques additionnelles. Enfin, des tests d'activité antivirale de ces conjugués ont été effectués dans des cellules humaines infectées par le VIH, ce qui a conduit à un  $IC_{50}$  de 0,6 et 0,4  $\mu M$  pour le composé lysine-S et le composé histidine-S respectivement.

D'autres approches ont également été étudiées pour cette cible notamment avec l'aide de peptides cycliques par l'équipe du Pr. Varani<sup>114</sup>. Ces différents exemples montrent que l'ARN TAR est un modèle pertinent pour pouvoir découvrir et développer de nouveaux chénotypes de ligand d'ARN. De plus, face à l'émergence possible de variants résistants du VIH, il est indispensable de continuer l'étude de ces différents modèles ainsi que le développement de nouvelles molécules antivirales.

Une autre cible attrayante pour cibler le VIH est le processus de **la dimérisation du génome rétroviral** qui intervient dans plusieurs étapes clés du cycle de vie du rétrovirus comme l'étape d'encapsidation. Cette dimérisation implique une structure tige-boucle appelée **site d'initiation de dimérisation (DIS)**. L'étape de dimérisation découle d'une interaction ARN boucle-boucle qui a lieu grâce aux séquences auto-complémentaires de la boucle formant des paires de bases intermoléculaires entre deux structures DIS. Cette reconnaissance médiée par une séquence de 6 nucléotides va entraîner une conversion en un duplex étendu par l'extension des interactions de Watson-Crick dans la région apicale de l'ARN DIS<sup>115</sup> (figure 35).



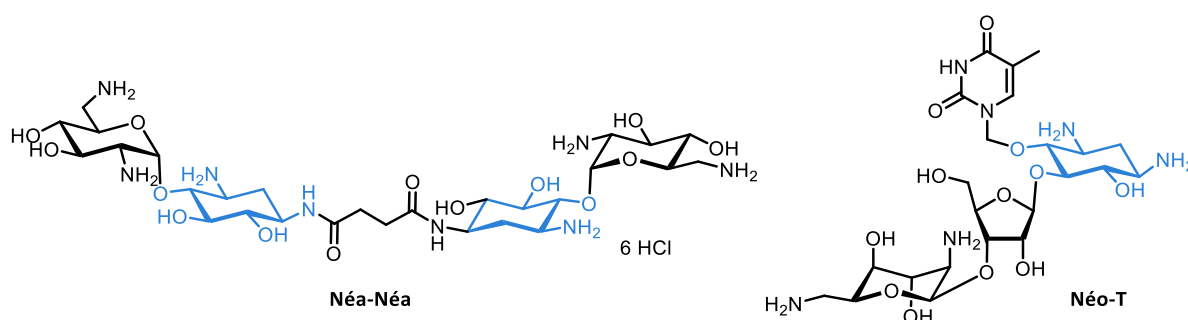
**Figure 35 :** A) Représentation de l'interaction entre deux ARN DIS lors de la dimérisation. Formation d'un complexe intermédiaire de boucle-boucle d'embrassement « kissing-loop ». B) Séquence de la tige-boucle de l'ARN DIS : en noir foncé la séquence de reconnaissance permettant l'interaction boucle-boucle<sup>116</sup>.

De nombreuses études ont été réalisées *in vitro* afin d'évaluer le mécanisme de formation du complexe « kissing-loop » et de la formation du duplex étendu. Ces études se sont basées sur la compréhension de la séquence de la tige et de la boucle dans la reconnaissance boucle-boucle<sup>117</sup>, mais comprennent

également des études thermodynamiques de ce complexe<sup>118</sup> et des études sur la protéine de nucléocapside (NCp7) qui est impliquée dans la formation du duplex étendu<sup>119</sup>.

Des structures cristallines et des analyses RMN en solution ont révélé de fortes similitudes structurales avec le site ribosomal A 16S (site A)<sup>115</sup>. Plusieurs aminoglycosides ont été étudiés pour leur capacité à interférer avec le processus de dimérisation par interaction avec ce complexe DIS (deux unités par complexe). La paromomycine, la néomycine et la lividomycine ont induit une stabilisation du « kissing-loop » suffisante pour empêcher la conversion en un duplex étendu<sup>120,121</sup>. Ces aminosides, 4,5-disubstitués de type 2-DOS, possèdent une affinité d'environ 50 fois supérieure pour ce complexe par rapport à leur cible principale qu'est le site ribosomal A bactérien. Cette amélioration de l'affinité s'explique par une forte implication du cycle II des aminoglycosides (noyau 2-DOS). L'analyse d'empreintes de la séquence DIS du VIH-1 révèle une protection par ces aminoglycosides dans les cellules et dans les virions.

A partir de ces analyses, la dimérisation de la néamine par un bras espaceur de longueur appropriée (4-6 Å) représente une base rationnelle dans la conception d'un nouveau ligand pour ce complexe boucle-boucle<sup>122</sup>. Ce dimère néamine **Néa-Néa** relié par un bras espaceur entre les amines  $N_1$  a été synthétisé par l'équipe de Dr. Marquet et du Dr. Dumas<sup>123</sup> (*figure 36*).



**Figure 36** : Représentation de la structure chimique du dimère-néamine (**Néa-Néa**) et du conjugué thymine-néomycine (**Néo-T**).

Les études biochimiques de ce dimère de néamine ont révélé la protection du complexe boucle-boucle lors du clivage sélectif induit par le plomb II et la stabilisation du complexe boucle-boucle à 37°C affectant la formation du duplex étendu.

Les équipes de Dr. Dumas et du Dr. Van Delft ont aussi élaboré un dérivé de la néomycine comportant une thymine par remplacement du cycle I (**Néo-T**) afin de réaliser une liaison de Watson-Crick avec l'adénine 280 située dans le complexe boucle-boucle<sup>116</sup>. Des analyses *in vitro* par ITC (Isothermal Titration Calorimetry) ont montré que le composé Néo-T possédait une meilleure spécificité envers le complexe boucle-boucle de l'ARN DIS, mais l'affinité mesurée restait plus faible que la néomycine par comparaison. Pour ces deux composés, l'activité potentielle antivirale ou antibactérienne n'a pas été abordée.

D'autres composés, tels que les composés naphtyridines ont également été étudiés dans le même contexte, c'est-à-dire dans le but de stabiliser un état conformationnel d'un ARN tel que le ribozyme en tête de marteau (HHR) par l'équipe du Dr. Nakatani<sup>124</sup>.

A travers ces exemples, l'études des interactions boucle-boucle sont sources d'inspiration et montrent la faisabilité des approches de criblage de ces complexes par des dérivés d'aminoglycosides ou autres petites molécules menant à une stabilisation de structure d'ARN. Ainsi, cette approche pourrait être proposée pour cibler d'autres systèmes de reconnaissance passant par des interactions boucle-boucle, notamment pour le ciblage de **systèmes Toxine-Antitoxine** qui fait l'objet de ce travail de thèse.

### I.5.5 Ciblage d'ARN eucaryotes

Les ARNnc eucaryotes jouent un rôle important dans presque tous les processus cellulaires. Certains ARN régulent la transcription, la traduction et l'épissage alternatif, et d'autres agissent sur des voies de signalisation. Cependant lorsque ces ARN sont dérégulés (sous-exprimés ou surexprimés) les processus dans lesquels ils interviennent ne se déroulent plus correctement et ces dérégulations ont été liées à un grand nombre de pathologies humaines, allant des cancers aux maladies neurodégénératives et neuromusculaires. De nombreuses molécules ont été étudiées pour chacune de ces maladies via le ciblage de structures particulières de différents ARN impliqués<sup>125,126,127</sup> (*tableau 4*).

*Tableau 4 : Différentes structures d'ARN ciblées par des petites molécules dans diverses maladies :*

Structure de l'ARN	Localisation	Maladie associée
Répétitions trinuéotidiques	Introns, 5'- et 3'-UTR	Maladie d'Huntington, le syndrome du X fragile, Dystrophie myotonique ...
miARN oncogènes	-	Cancers variés
G-quadruplexes	5'-UTRs, sites d'épissage	Maladie d'Alzheimer, cancer du sein, de la prostate, pancréatique
Sites d'épissage	Jonctions Exon-intron	Maladie de Parkinson, Amyotrophie spinale proximale

Un des exemples les plus importants et récents de petites molécules ligands d'ARN interférant avec des ARN eucaryotes est le risdiplam mis au point par Roche<sup>128</sup> avec son homologue branaplam<sup>129</sup> (phase clinique) (*figure 37*). Ces molécules agissent contre l'amyotrophie spinale proximale qui est le résultat d'une perte fonctionnelle du gène de survie du motoneurone 1 (SMN1) entraînant une perte musculaire et une mort précoce<sup>66</sup>. Les humains possèdent un gène paralogue appelé SMN2 qui est non fonctionnel. Cependant, le risdiplam va induire l'inclusion de l'exon 7 lors de l'épissage du pré-ARNm de la protéine SMN2 qui devient fonctionnelle. Cet épissage restauré va ainsi contrebalancer l'absence de la protéine SMN1.

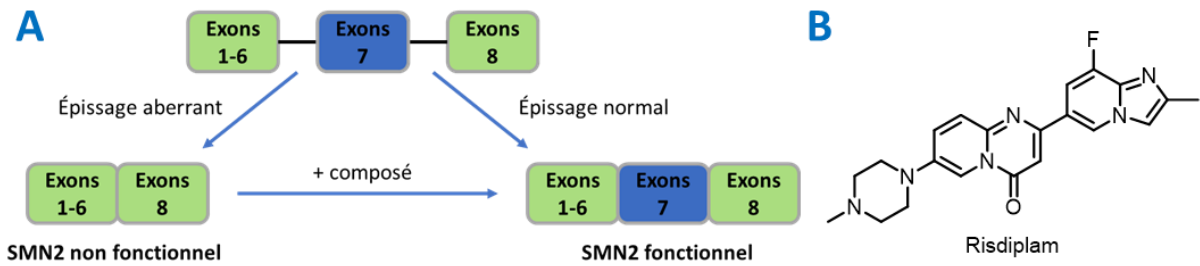


Figure 37 : A) Représentation de l'épissage de la protéine SMN2<sup>130</sup>. B) Structure chimique du risdiplam.

Un autre progrès majeur dans le domaine des ligands d'ARN est la construction de composés chimères appelés RIBOTAC qui ciblent la biogénèse des miARN et qui ont été mis au point par l'équipe du Pr. Disney<sup>131</sup> (figure 38).

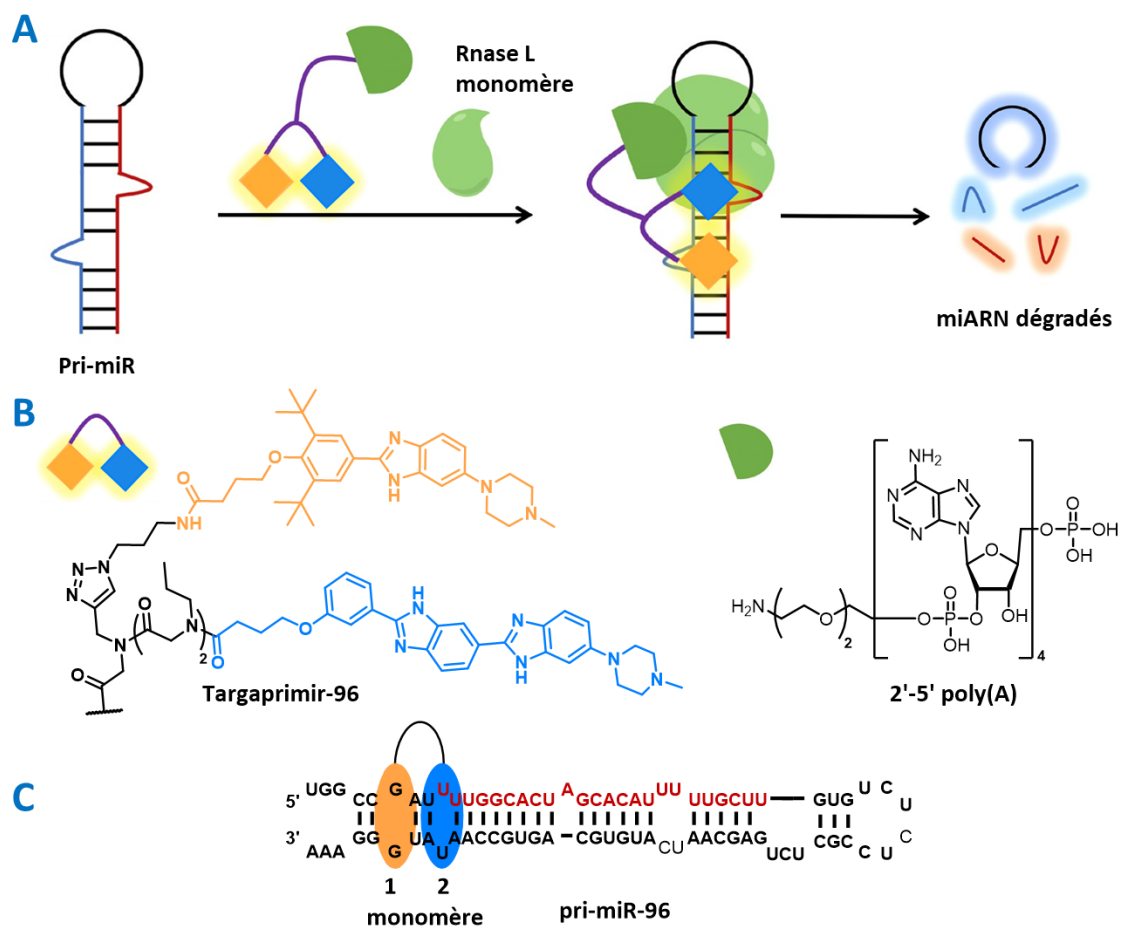


Figure 38 : A) Principe du RIBOTAC faisant appel à une RNase dégradant les miARN ; B) Structures chimiques du Targaprimir-96 et de l'oligonucléotide 2'-5' poly(A) ; C) Sites d'interactions du Targaprimir-96 sur le pri-miR-96.

Cette technologie a été mise au point pour l'étude du miR-96 qui est oncogène et surexprimé dans plusieurs cancers et inhibe l'expression de la protéine FOXO1 notamment dans les cancers du sein triple négatif (TNBC). La biogénèse des miARN implique la présence de précurseurs (pri-miARN et pre-miARN) qui sont clivés par deux enzymes (Drosha et Dicer) et qui vont permettre la formation du

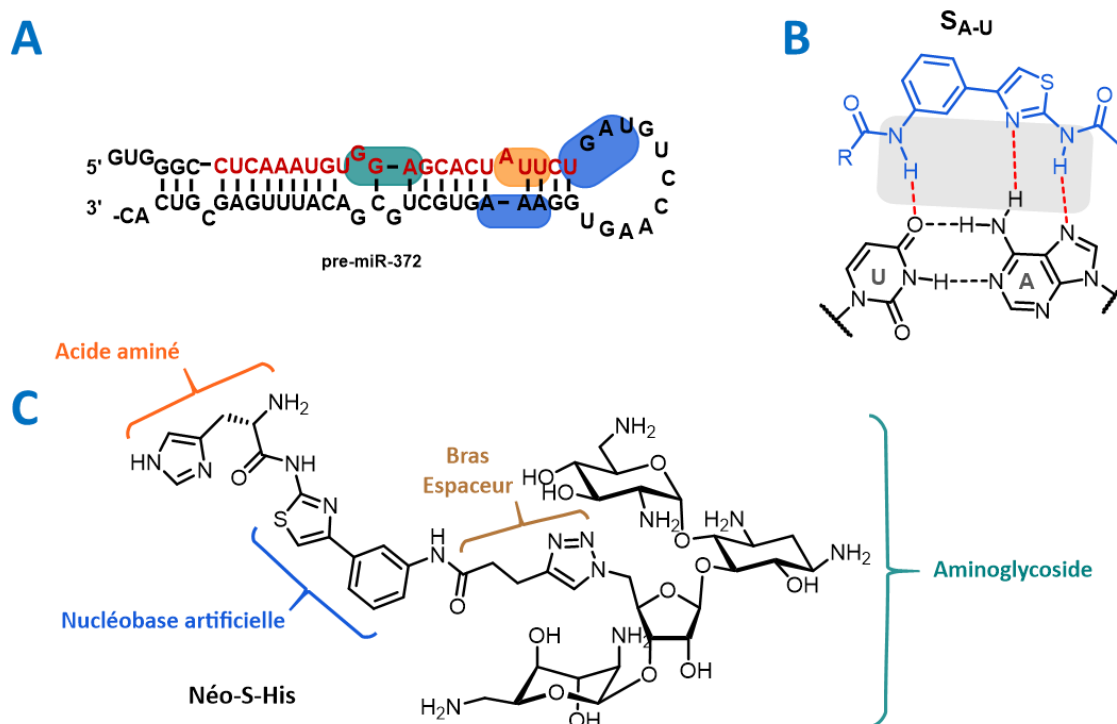
microARN mature. Il est donc intéressant de venir inhiber la biogenèse de miARN surexprimés par liaison sur un des deux précurseurs afin d'éviter leur maturation en miARN.

Grâce à l'outil Inforna décrit précédemment (cf. **Chap.I.5.2.2**), un ligand monomère 1 (*figure 37, B, C*) a été identifié pour cibler un site de liaison spécifique du pri-miR-96 qui possède une structure en épingle à cheveux. Après optimisation, un deuxième monomère 2 intégrant une partie bis-benzimidazole a été ajoutée pour cibler un deuxième site d'interaction adjacent à ce premier site de liaison<sup>132</sup>. Ce composé bivalent **targaprimir-96** se lie avec une affinité de l'ordre du nanomolaire au pri-miR-96. Des tests intracellulaires ont été effectués pour déterminer l'activité inhibitrice de l'enzyme Drosha qui est de l'ordre de 50 nM. Ce composé a été évalué sur la biogénèse du miR-96, donnant une inhibition de plus de 400 fois supérieure à celle du composé monomère dans les cellules cancéreuses du sein. Des études *in vivo* chez des modèles de cancers du sein triple négatif montrent une inhibition de la croissance tumorale.

Dans le but d'ajouter une activité de dégradation de l'ARN pour ce précurseur du miR-96, de la même manière qu'il a été effectué précédemment dans le développement des PROTAC, Disney et son équipe ont ajouté au targaprimiR-96 un module (2'-5' poly(A) oligonucléotide) capable de recruter une **RNase L endogène**. Cette dernière induit le clivage et la dégradation de l'ARN cible. Le composé tagaprimir-96 possède ainsi une activité inhibitrice de la biogénèse du miR-96 largement meilleure d'au moins 5 fois.

Dans la littérature, d'autres miARN ont été ciblés en utilisant diverses approches<sup>133,134,135</sup>. Notre équipe a également contribué à l'étude de différents ligands ciblant la biogénèse des miARN en particulier le miR-372. Ce microARN oncogène est impliqué dans le développement et la progression de plusieurs cancers tels que le cancer de l'œsophage, de la thyroïde ou encore de l'adénocarcinome gastrique. Ce miR-372 lorsqu'il est surexprimé, va inhiber la biogénèse de la cible protéique LATS2 qui est une protéine suppresseur de tumeur. Afin d'inhiber ce microARN, une approche par conjugaison a été utilisée afin d'élaborer différents ligands d'ARN multifonctionnels (*figure 39*). Cette approche permet d'assembler différents motifs qui sont connus pour être affins pour l'ARN et/ou sélectifs sur un motif défini.

Dans le but de mieux comprendre les motifs essentiels dans la reconnaissance d'ARN et l'activité sur le pré-miR-372, des conjugués d'aminoglycosides ont été étudiés contenant par exemple la néomycine. La fonctionnalisation de cet aminoglycoside a été effectuée en 5' par une fonction azoture afin d'effectuer une chimie « click » avec des alcynes comportant des nucléobases naturelles ou artificielles<sup>136</sup>.



**Figure 39 :** A) Interactions du conjugué Néo-S-His sur le pre-miR-372 ; B) Interaction de Hoogsteen pour la nucléobase S ; C) Structure chimique du conjugué Néo-S-His pour cibler le pre-miR-372 par l'équipe Duca.

**Le choix des aminoglycosides** est basé sur leur forte affinité avec les motifs d'ARN en particulier avec les structures tige-boucle et leur capacité d'inhiber le processus de clivage par Dicer sur différents pre-microARN<sup>137</sup>. En effet, ces composés sont capables de former des liaisons hydrogènes sélectives avec les bases non-appariées des boucles et des renflements mais aussi des liaisons électrostatiques fortes avec le squelette phosphate grâce à la présence d'amines primaires sur les aminosucre qui constituent les aminoglycosides.

**Le choix des nucléobases naturelles et artificielles** a été basé sur la possibilité de former des liaisons hydrogènes spécifiques avec les bases non-appariées pour les nucléobases naturelles ou des bases appariées pour les nucléobases artificielles. Ces dernières sont en effet capables de reconnaître spécifiquement des paires AU ou GC comme décrit précédemment<sup>138</sup>.

Lors d'une première étude, le conjugué qui a montré les meilleurs résultats en termes d'affinité et capacité d'inhibition était constitué de la base artificielle « S » utilisée précédemment dans le ciblage de l'ARN TAR du VIH<sup>113</sup>. Le conjugué **Néo-S** (figure 39) a montré des résultats prometteurs dans l'inhibition du clivage du pre-miR-372 *in vitro* mais également lors de tests évaluant la prolifération des cellules d'adénocarcinome gastrique.

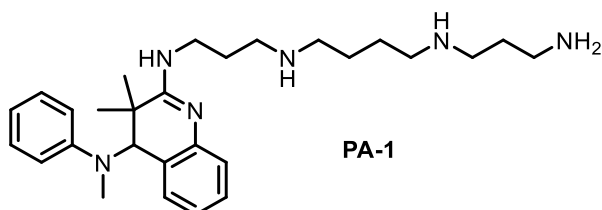
Dans le but d'améliorer la sélectivité et l'affinité par une approche de chimie médicinale, diverses modifications ont été apportées. Dans un premier temps, des modifications pour évaluer l'importance de la longueur et **du type du bras espaceur** ont été faites<sup>139</sup>. En parallèle, dans la même étude, la



modification de l'aminoglycoside a été effectuée en le remplaçant par des parties de la néomycine telles que la 2-désoxystreptamine ou la néamine. Ces modifications n'ont pas donné d'amélioration significative pour l'activité et une longueur étendue du bras espaceur s'est avérée délétère. En revanche, la modification de la nucléobase artificielle et notamment l'augmentation de l'encombrement par ajout d'un noyau aromatique supplémentaire a mené à une forte amélioration de l'activité. À partir de ces modifications, et dans un deuxième temps, la partie *N*-terminale de la nucléobase « S » a été substituée par divers acides aminés<sup>140</sup>. **Les acides aminés** apportent une reconnaissance supplémentaire et sont connus pour interagir avec l'ARN par des interactions électrostatiques et par des interactions supplémentaires conférées par leur chaîne latérale. Ainsi, trois acides aminés basiques ont été choisis tels que la lysine, l'arginine et l'histidine.

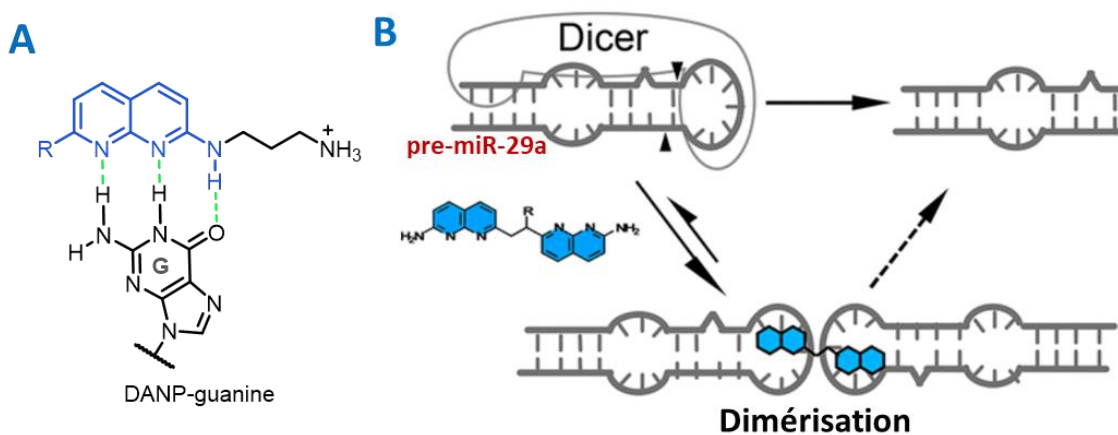
De cette étude, le composé **Néo-S-His** a montré une affinité 16,4 nM envers le pre-miR-372 (*figure 39, C*). L'ajout du groupement histidine a aussi permis d'améliorer la sélectivité grâce à la formation d'interactions supplémentaires au niveau de la jonction de la boucle apicale. L'activité inhibitrice *in vitro* du clivage de Dicer a pu ainsi être améliorée pour atteindre 0.79  $\mu$ M alors que le composé parent Néo-S présentait une valeur d'IC<sub>50</sub> de 2.42  $\mu$ M. Des études d'empreintes enzymatique et de docking moléculaire ont permis de déterminer les interactions de ce composé sur le pre-miR-372 se situant dans le site de liaison de l'enzyme Dicer (*figure 39, A*). Enfin, ce composé conserve son activité dans la réduction du miR-372 au niveau intracellulaire.

D'autres composés ont également été évalués pour le ciblage du pre-miR-372 dans notre équipe. Un criblage de 640 molécules constituant la Chimiothèque Essentielle et représentant la diversité moléculaire de la Chimiothèque Nationale (70000 composés) a permis d'identifier 3 composés capables de se lier au pre-miR-372<sup>141</sup>. Le plus prometteur est le composé **spermidine-amidine (PA-1)** qui possède une bonne sélectivité pour le pré-miR-372 en présence d'ARNt et d'ADN (*figure 40*). De plus, ce composé inhibe *in vitro* la biogénèse du miR-372 (IC<sub>50</sub> = 1,06  $\mu$ M) et a une activité dans les cellules d'adénocarcinome gastrique. Une étude du miRnome des cellules traitées a permis de confirmer l'inhibition d'un petit nombre de miARN oncogènes dans ces cellules tumorales et les miARN ciblés partageaient non seulement un site de liaison similaire mais aussi la cible protéique LATS2.



**Figure 40** : Structure chimique du composé spermidine-amidine évaluée lors d'un criblage dans l'équipe du Dr. Duca.

Enfin, l'équipe du professeur Nakatani a étudié le ciblage des miARN, notamment du pre-miR-29a, avec des composés comportant le motif naphthyridine. Ce motif est capable d'interagir avec la guanine ou cytosine par trois liaisons hydrogènes (*figure 41, A*). En 2016, le composé DANPH<sup>+</sup> a été étudié pour interagir avec un renflement mono-nucléotidique comportant une cytosine et a ainsi permis d'inhiber l'action de Dicer dans la biogénèse de ce miARN<sup>142</sup>. Afin d'augmenter l'affinité de ces composés, des études rapportent l'utilisation de dimères ou tétramères de différents analogues de DANP<sup>143</sup> dont la formation de complexes entre deux tiges-boucles<sup>144</sup>. Plus récemment, en 2021, un dimère de naphthyridine a été conçu pour se lier à la boucle en épingle à cheveux du pre-miR-29a et induire sa dimérisation permettant l'inhibition de l'action de l'enzyme Dicer<sup>145</sup> (*figure 41, B*).



**Figure 41 :** *A) Structure et interactions du complexe DANP-guanine. B) Représentation de l'inhibition de l'action de Dicer sur le pre-miR-29a.*

À travers ces différents exemples de ligands d'ARN qui ciblent les ARNnc des eucaryotes, nous avons illustré que différentes stratégies ont été développées et ont mené à la conception de ligands affins et sélectifs pour leur cible. L'approche utilisant l'outil Inforna montre que les renflements ou boucles internes sont des cibles qui peuvent être privilégiées par certaines structures chimiques. La conception de ligands multifonctionnels permet d'augmenter considérablement la spécificité pour une cible donnée, permet de concevoir des molécules de façon rationnelle et ouvre la voie à une optimisation successive. Dans le cas particulier des interactions boucle-boucle, il a été montré qu'il est possible de stabiliser la formation du complexe « kissing loop » ou encore inhiber la formation du complexe. Au cours de ce projet, nous souhaitons inhiber l'interaction boucle-boucle des systèmes Toxine-Antitoxine et ces exemples sont particulièrement importants dans ce contexte.



---

# **Chapitre II - Synthèse et évaluation biologique de nouveaux ligands d'ARN**

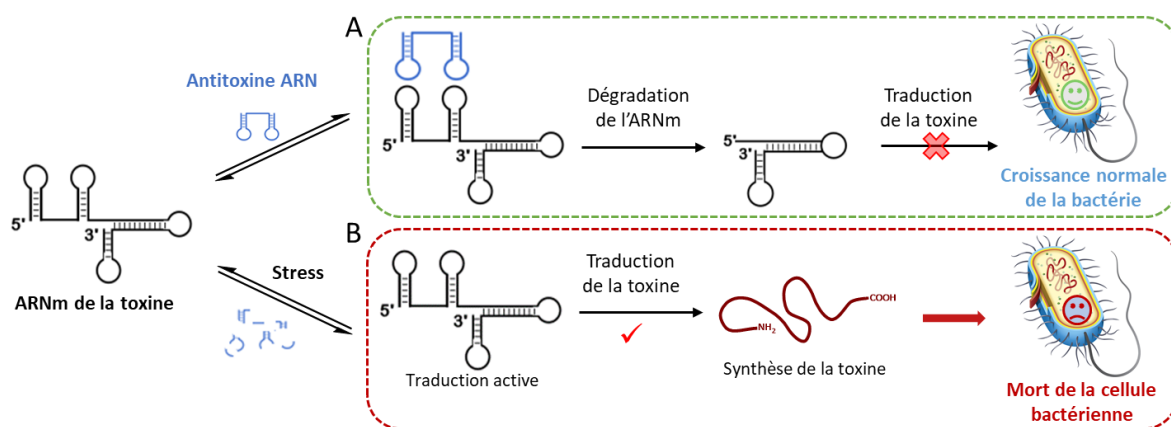
---



## II.1 Introduction : description du projet

Comme introduit précédemment, les bactéries possèdent un étonnant système qu'est le système Toxine-Antitoxine. Même si de nombreuses questions subsistent en ce qui concerne l'activation, le rôle biologique et le mécanisme potentiellement réversible de certains de ces systèmes, il est fort intéressant de mener de nouvelles études pouvant améliorer leur compréhension et pouvant valider leur utilisation en tant que cible antimicrobienne.

Chez les bactéries comportant un système TA de type I (composé de deux ARNs), la traduction de la toxine est contrôlée en permanence par l'antitoxine qui est un ARN non codant. Ce contrôle va induire une dégradation de l'ARNm codant la toxine par une RNase, et par conséquent celle-ci ne sera pas traduite. Ainsi, lorsque cet équilibre s'opère, la bactérie se développe normalement (*figure 42, A*).



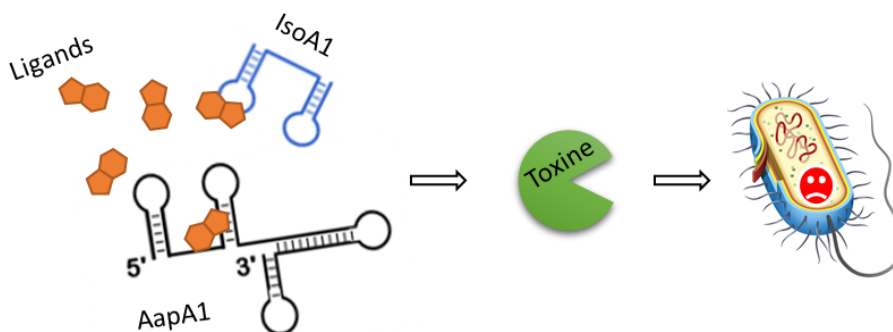
**Figure 42 :** Représentation simplifiée des systèmes TA de type I. **A)** équilibre des interactions des ARNs toxine-antitoxine permettant l'inhibition du peptide toxine et la croissance cellulaire bactérienne. **B)** Conditions de stress permettant la traduction du peptide toxine entraînant l'arrêt de la croissance ou la mort de la cellule bactérienne.

Pendant, lorsque la bactérie est soumise à des conditions non favorables pour une croissance normale, telles que l'augmentation de la température ou du pH, la présence d'antibiotiques, le stress oxydatif ou les dommages à l'ADN, l'antitoxine est alors dégradée et ne permet plus d'interagir avec l'ARNm de la toxine. Dans cette situation, la traduction de la toxine est activée et remplit sa fonction biologique soit en inhibant la croissance bactérienne soit en provoquant la mort de la cellule bactérienne (*figure 42, B*).

Dans ce dernier cas de figure, différentes équipes se sont posées la question : « **Est-il possible d'activer ces systèmes TA de manière artificielle pour provoquer la traduction de la toxine ?** ». Ces réflexions ont mené à différentes recherches ciblant le système TA de Type II (inhibition d'interactions protéine-protéine) mais aucune pour le type I<sup>48</sup>.

Afin d'étendre ce champ d'action, nous nous sommes intéressés aux systèmes de type I notamment le système *AapA/IsoA* présent chez la bactérie *Helicobacter pylori*. Dans ce système TA, la toxine codée par *AapA1* est un peptide qui induit la mort de la bactérie. La possibilité de pouvoir activer artificiellement la traduction de la toxine pourrait s'inscrire dans **une nouvelle approche antibiotique** après validation de cette nouvelle cible. Dans ce but, nous proposons de développer des ligands d'ARN capables d'inhiber l'interaction ARN-ARN impliquée dans les systèmes TA de type I et plus particulièrement dans le système *AapA/IsoA*.

La stratégie proposée dans ce manuscrit provient d'une étude préliminaire menée par l'équipe du Dr. Darfeuille<sup>59</sup>. En effet, l'aminoglycoside néomycine, ligand d'ARN efficace mais très peu sélectif, a montré la capacité d'inhiber l'interaction entre l'ARNm codant pour la toxine et celui de l'antitoxine dans le système *AapA/IsoA*. Ce premier résultat nous a permis de concevoir un nouveau projet portant sur la conception de ligands d'ARN inhibiteurs d'interactions boucle-boucle comme celles formées par *AapA* et *IsoA* afin d'étudier la possibilité d'actionner le mécanisme de la traduction de la toxine. Dans ce contexte, nous nous sommes intéressés au cours de **ces travaux de thèse à l'élaboration de nouveaux ligands d'ARN comme potentiels inhibiteurs de la formation d'interactions boucle-boucle**. De fait, la conception de ces ligands est fondée sur la reconnaissance des structures se situant au niveau des boucles ou à la jonction des tige-boucles de ces ARNs. Ceci, dans le but d'inhiber la formation de l'interaction boucle-boucle ce qui entrainera la traduction de la toxine *AapA* et conduira donc à la mort de la cellule bactérienne (*figure 43*).



**Figure 43 :** Représentation schématique de la stratégie permettant l'activation de la traduction de la toxine par l'inhibition de l'interaction entre l'ARN de l'antitoxine et l'ARNm de la toxine.

Comme mentionné précédemment, le système *AapA/IsoA* est présent dans le génome d'*H. pylori* en plusieurs copies. Il y a donc plusieurs avantages dans le ciblage de ce système : (i) le système *AapA/IsoA* induit la mort de la bactérie, (ii) l'antibiotique capable d'activer la production de la toxine ne serait pas toxique et la molécule toxique pour la bactérie resterait la toxine produite par la bactérie elle-même et (iii) puisque présent à plusieurs sites génomiques, même si une mutation avait lieu sur un de ces sites, il est difficilement envisageable que la bactérie arrive à muter tous les sites du système

*AapA/IsoA*. L'ensemble de ces paramètres en font une cible de choix qui laisse présager peu ou pas de développement de résistance si des ligands efficaces seront identifiés.

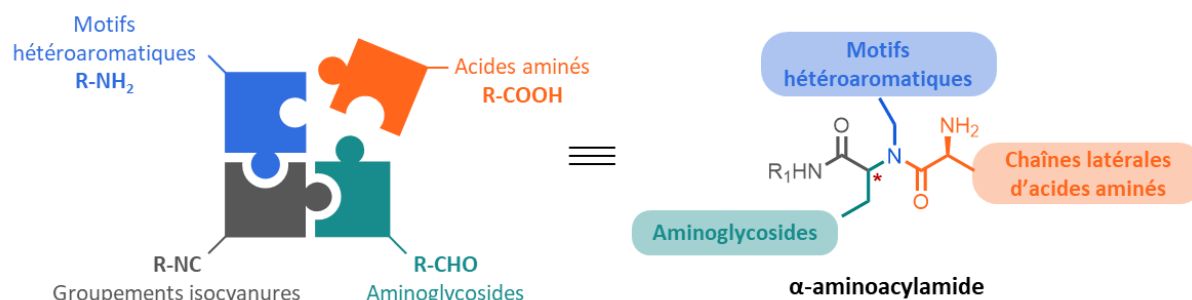
### II.1.1 Conception de ligands d'ARN inhibiteurs de la formation de complexes boucle-boucle

Grâce à l'expertise de notre équipe, nous avons décidé d'utiliser **une approche basée sur la conception de ligands multifonctionnels** pour explorer l'inhibition des complexes ARN-ARN et ainsi constituer une bibliothèque de ligands d'ARN. Les ligands multifonctionnels sont conçus pour contenir plusieurs motifs d'interaction avec l'ARN et leur utilisation par le passé a permis d'obtenir un bon équilibre entre l'affinité et la spécificité grâce à la combinaison de motifs capables de former des interactions fortes (électrostatiques ou d'empilement) ainsi que des liaisons plus spécifiques telles que des interactions de Van der Waals ou des liaisons hydrogène.

Pour la préparation de cette nouvelle bibliothèque de composés, nous avons utilisé une réaction multicomposants (MCR) qui permet d'assembler divers motifs, donc divers fragments, en une seule et même réaction<sup>146</sup>. De nombreuses MCR existent et permettent d'obtenir des structures complexes avec des ramifications variées notamment à la suite de réactions supplémentaires « post-MCR ». Dans le but de préparer des chimiothèques de conjugués multifonctionnels, nous avons choisi d'utiliser la **réaction de Ugi**, découverte par Ivar Karl Ugi en 1959<sup>147</sup>. Cette réaction permet d'obtenir un squelette pseudo-dipeptidique (ou  $\alpha$ -aminoacylamide) rappelant les peptides. Ces derniers sont connus pour interagir avec l'ARN et sont résistants dans les milieux biologiques. Cette réaction, qui est facile à mettre en œuvre, permet également d'explorer de nouveaux composés non linéaires qui sont peu présents dans les ligands d'ARN connus.

Cette réaction de Ugi a été réalisée par la condensation de quatre fonctions chimiques ce qui correspond parfaitement à notre stratégie de conception de ligands multifonctionnels. Afin d'élaborer une diversification rapide et facile de notre bibliothèque, chacune de ces fonctions comporte une famille de fragments spécifiques et connus pour interagir avec l'ARN (*figure 44*) : l'**acide carboxylique** (fragment orange) comporte des acides aminés, l'**amine** (fragment bleu) comporte des composés hétéroaromatiques, l'**aldéhyde** (fragment vert) comporte des motifs d'aminoglycosides ou analogues et l'**isocyanure** permet de réaliser la réaction soit en support solide soit en solution (fragment gris). Par ailleurs, le substituant du composé isocyanure pourra lui aussi être modulé par la suite.





**Figure 44 :** Conceptions des ligands d'ARN par une approche par fragments en utilisant une réaction de Ugi pour former des motifs  $\alpha$ -aminoacylamides soit sur support solide polymérique ( $R_1 = \text{résine}$ ) soit en solution ( $R_1 = t\text{Bu}$ ).

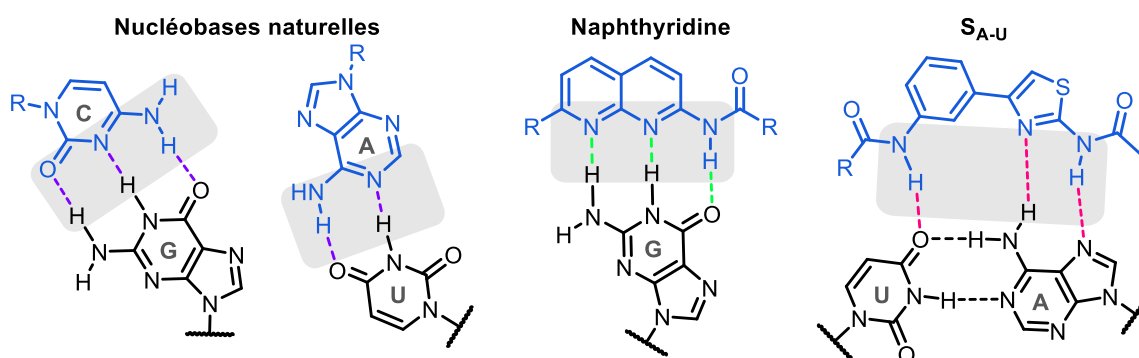
Ci-après, il sera détaillé le choix de ces différents motifs pour chacun des composants de la réaction de Ugi.

## II.1.2 Choix des motifs

### II.1.2.1 Motifs hétérocycliques

Pour former des interactions spécifiques soit avec les nucléobases non appariées au niveau des boucles apicales ou renflements, soit avec les paires de bases au niveau des tiges des ARNs, différents motifs hétérocycliques spécifiques ont été choisis.

Dans le cas des motifs ciblant les nucléobases non appariées, nous avons utilisé les quatre nucléobases naturelles (A, C, G et T) qui peuvent interagir spécifiquement avec leur nucléobase complémentaire dans les boucles apicales. Les interactions de ces nucléobases sont rendues possibles par les liaisons hydrogènes de type Watson-Crick (*figure 45*).



**Figure 45 :** Représentation des composés hétérocycliques choisis et leurs interactions sélectives avec les nucléobases correspondantes.

D'autres motifs sont également connus et sont capables d'interagir spécifiquement avec des nucléobases naturelles non appariées tels que le dérivé 1,8-naphtyridine. Comme nous l'avons décrit dans le chapitre I, ce motif est très étudié par l'équipe du Pr. Nakatani. En 2014, un ligand composé de quatre motifs naphtyridines a montré une stabilisation de deux ARN structurés en tige-boucle par

l'interaction de ces motifs avec des résidus guanines dans les boucles apicales<sup>144</sup>. Ces études ont ainsi montré la capacité de ce motif naphtyridine à interagir avec la nucléobase guanine à l'aide de trois liaisons hydrogènes (*figure 45*).

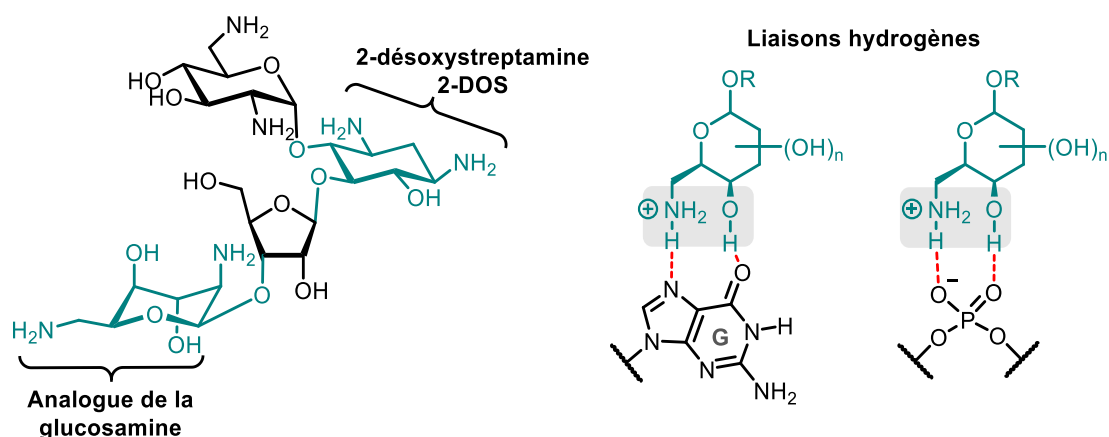
Enfin, des motifs hétérocycliques de type nucléobase artificielle ont été choisis pour interagir avec des paires de bases d'ARN au niveau des tiges et ainsi former des triplets avec celui-ci. Ces motifs vont spécifiquement former des interactions de type Hoogsteen avec une paire de bases de l'ARN. De nombreux motifs formant des triplets d'ARN sont connus<sup>138</sup>, nous nous sommes focalisés sur le motif S<sub>A-U</sub> précédemment utilisé et étudié dans divers travaux de notre équipe<sup>111</sup>. Ce motif forme trois interactions spécifiques avec la paire de bases A-U (*figure 45*).

Pour réaliser la réaction de Ugi, un bras espaceur contenant la fonction amine est assemblé sur chaque motif sélectionné précédemment. Les différentes synthèses de ces motifs seront décrites dans la suite de ce chapitre. Au-delà de ces composés, nous avons également choisi d'utiliser certains composés commerciaux hétérocycliques tels que des dérivés comportant des noyaux benzothiazoles ou imidazoles fréquemment employés au sein de ligands d'ARN très affins et sélectifs<sup>81</sup>.

Enfin, ces motifs confèrent aux molécules nouvellement synthétisées un caractère hydrophobe et conduisent à une meilleure affinité et une meilleure sélectivité en permettant la formation d'interactions d'empilement supplémentaires avec l'ARN.

### II.1.2.2 Motifs aminoglycosides et analogues

Dans l'étude préliminaire, la néomycine a montré une protection de l'ARNm de la toxine *AapA* et notamment *AapA1* dans le système TA étudié<sup>59</sup>. La néomycine B fait partie de la famille des aminoglycosides et est très étudiée en tant que ligand d'ARN. Ce composé est connu pour se positionner au niveau des régions dans lesquelles la structure en double hélice d'ARN est affectée par la présence d'une boucle ou d'un renflement, c'est-à-dire au niveau des jonctions simple-brin/double-brin. De plus, la néomycine a montré une capacité à interférer avec les interactions boucle-boucle dans des études menées sur l'ARN DIS du VIH<sup>121</sup>. La forte affinité que ces aminoglycosides présentent pour les ARNs s'explique par des interactions électrostatiques intermoléculaires entre les groupes ammonium chargés positivement et les groupements phosphates qui constituent le squelette de l'ARN. Néanmoins, il est à noter que les aminoglycosides peuvent aussi former des liaisons hydrogènes avec les bases non appariées<sup>148</sup>. De ce fait, le choix d'utiliser la néomycine B constitue un bon point de départ pour réaliser la synthèse de nouveaux ligands d'ARN pour ce projet (*figure 46*).



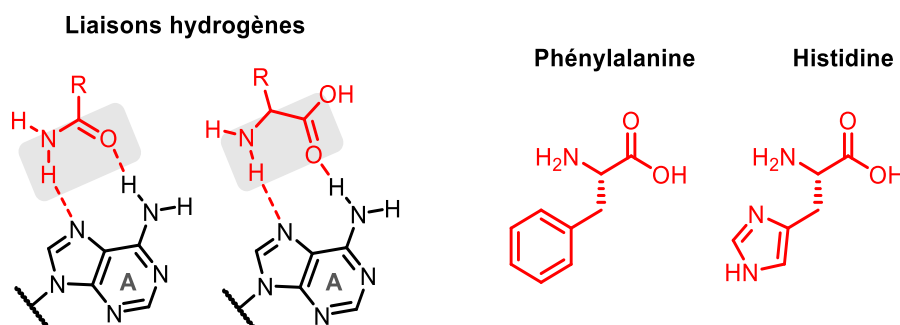
**Figure 46 :** Représentation des motifs choisis de l'aminoglycoside néomycine B et leurs interactions avec les différents constituants de l'ARN.

Cependant, le haut poids moléculaire ainsi que les phénomènes de résistance et de toxicité de la néomycine limitent son utilisation clinique<sup>24</sup>. Afin de profiter des propriétés intéressantes de cet aminoglycoside en tant que ligand d'ARN mais d'avoir des caractéristiques plus favorables, nous avons donc sélectionné des sous-parties de son squelette. Le motif de la **2-déoxystreptamine (2-DOS)**, noyau central de ces aminoglycosides, est souvent considéré comme un motif essentiel dans les interactions avec l'ARN du site A ou de l'ARN DIS (cf. *Chap. I.5.4*). Il est également mis en avant pour être affiné envers les boucles des structures ARN en épingles à cheveux<sup>149</sup> C'est la raison pour laquelle nous l'avons choisi pour la suite de nos recherches. Un autre motif a été également sélectionné, la **glucosamine**, qui constitue un analogue de l'aminosucrose du cycle IV de la néomycine.

En raison des interactions électrostatiques mais aussi des liaisons hydrogènes spécifiques que ces composés peuvent former avec des paires de bases non appariées, ces motifs peuvent apporter une forte affinité pour la cible. Par ailleurs, ce type de composé devrait améliorer la solubilité des composés synthétisés dans les milieux aqueux. Dans le but de pouvoir les utiliser comme fonction aldéhyde, ces deux composés seront synthétisés avec un bras espaceur contenant cette fonction aldéhyde.

### II.1.2.3 Motifs acides aminés

Les acides aminés sont des constituants des ligands naturels de l'ARN que sont les peptides (par exemple le peptide Tat qui interagit avec l'ARN TAR). Des interactions de type liaison hydrogène peuvent être observées notamment avec les bases non appariées, mais aussi avec le squelette chargé de l'ARN (*figure 47*).



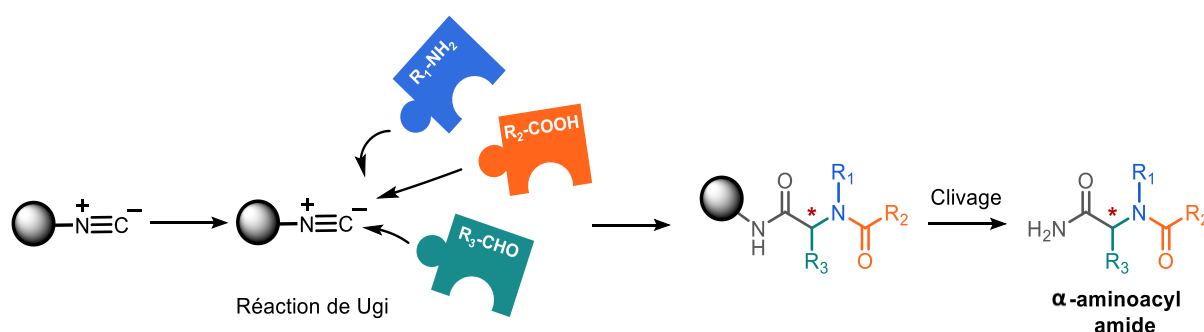
**Figure 47 :** Représentation des interactions des acides aminés avec les nucléobases telles que l'adénine. Exemples d'acides aminés choisis.

Ces acides aminés comportent une fonction acide carboxylique qui permet une utilisation directe sans modification pour la réaction de Ugi. Une protection de l'amine alpha sera par conséquent utilisée pour permettre leur utilisation.

Différents acides aminés ont été choisis pour la synthèse de cette bibliothèque par rapport à leur chaîne latérale qui confère des interactions supplémentaires spécifiques. Des acides aminés aromatiques, tels que la phénylalanine et le tryptophane, permettent d'obtenir des interactions hydrophobes et d'empilement. Les acides aminés basiques tels que la lysine et l'histidine peuvent apporter des interactions électrostatiques supplémentaires.

### II.1.2.4 Motifs isocyanures

Dans un premier temps, nous avons souhaité mettre au point la réaction de Ugi pour la préparation des nouveaux ligands d'ARN qui seront présentés dans ce chapitre sur phase solide polymérique (SPS) (*schéma 3*).



**Schéma 3 :** Représentation de la réaction de Ugi sur support solide polymérique employée dans ce manuscrit.

Le recours à un quatrième composant s'est porté sur la possibilité de pouvoir utiliser cette fonction comme réactif sur un support solide<sup>150</sup>. Une réaction à trois composants comparables à la réaction de Ugi telle que la réaction de Passerini aurait pu suffire, mais cette dernière aurait nécessité la fonctionnalisation d'un des fragments avec une fonction isocyanure et aurait conduit à un squelette portant une fonction ester, peu stable en milieu biologique. C'est pourquoi, l'application de la réaction

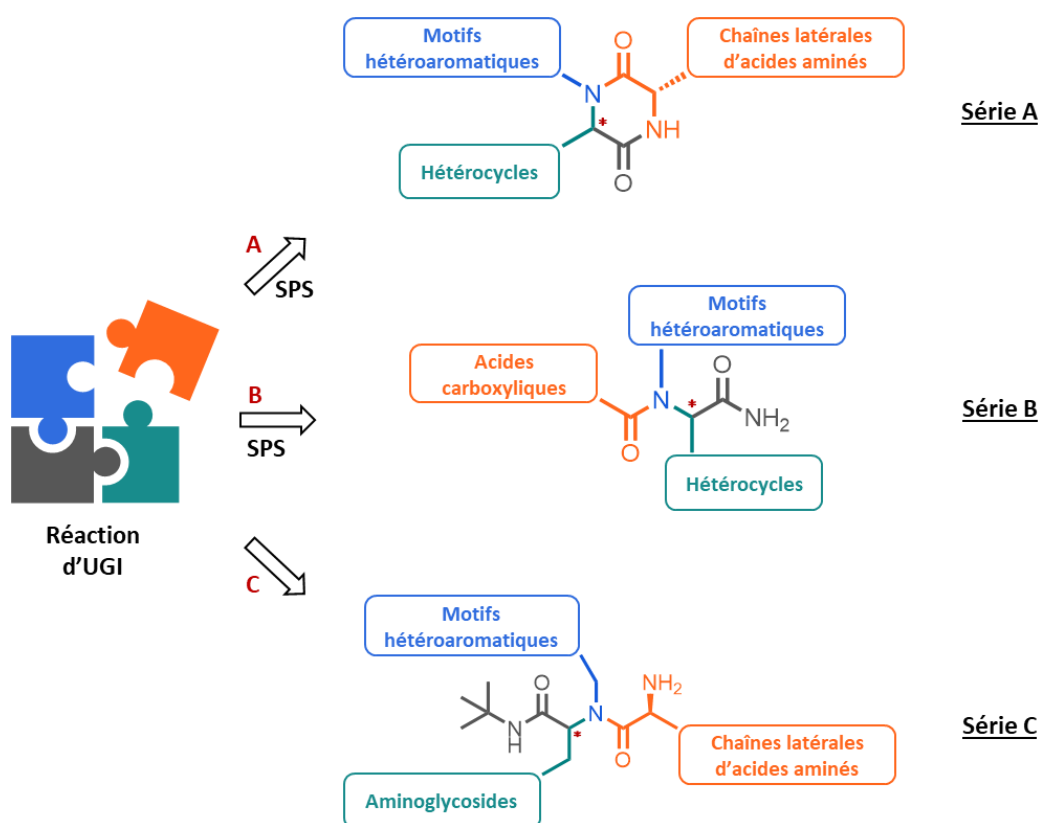
d'Ugi par l'utilisation de la phase solide rend notre méthodologie encore plus favorable pour la préparation d'une large collection de composés. Cette approche nous offre également d'autres avantages. Tout d'abord, elle permet de lancer en parallèle de nombreuses réactions permettant un accès rapide aux composés souhaités. Cela est possible par la facilité de préparer une grande quantité de support solide fonctionnalisé avec la fonction isocyanure. De plus, l'utilisation de support solide permet théoriquement de fournir des composés relativement purs sans traitement post réactionnel et de s'affranchir des étapes de purification permettant un gain de temps considérable. Un autre avantage est que l'étape de clivage de la résine peut être réalisée dans des conditions acides ce qui conduit à une déprotection des groupements protecteurs acido-labiles concomitante avec le clivage.

Enfin, dans l'optique d'une future optimisation d'un produit sortant, l'introduction d'un quatrième fragment remplaçant la fonction isocyanure du support solide pourra permettre d'augmenter la sélectivité ou l'affinité du composé.

### II.1.3 Préparation de trois séries de nouveaux ligands d'ARN multifonctionnels

Dans ce **chapitre II**, il sera brièvement présenté dans une première partie la réaction de Ugi. Ensuite, la préparation de l'ensemble des composants de cette réaction sera décrite. Il sera question d'abord de la fonctionnalisation du support solide pour l'obtention de la fonction isocyanure. Dans le but d'utiliser des groupements fonctionnels de type *tert*-butoxycarbonate (Boc), le choix de la résine s'est porté sur la résine Rink amide MBHA. Puis, il sera détaillé les différentes synthèses des motifs comportant la fonction amine (10 synthons) et la fonction aldéhyde (2 synthons).

Dans une deuxième partie, une description des travaux permettant l'obtention de trois séries de composés sera décrite à la suite d'une réaction de Ugi et selon les motifs utilisés : La **série A et B** sont réalisées en phase solide et la **série C** est réalisée par des réactions en phase liquide (*figure 48*).



**Figure 48** : Représentation des différentes séries obtenues par une réaction de Ugi au cours du travail de thèse. Les séries A et B ont été obtenues sur support solide polymérique (SPS). La série C a été obtenue par réaction en phase liquide.

Enfin, dans une dernière partie, il sera abordé les différents tests biochimiques que j'ai pu mettre en place au sein du laboratoire permettant d'évaluer l'affinité et la sélectivité des composés synthétisés pour les ARNs cibles, mais aussi la capacité de ces composés à interférer avec l'interaction Toxine-Antitoxine. Des tests biologiques développés et réalisés par l'équipe Dr. Darfeuille sur le système TA de *Helicobacter pylori* compléteront cette partie. Au cours du travail de thèse, j'ai eu l'opportunité de travailler au sein du laboratoire du Dr. Darfeuille afin d'effectuer moi-même une partie des tests biologiques. Enfin, une évaluation de l'activité antimicrobienne sur plusieurs souches bactériennes nous a permis d'identifier des nouveaux composés potentiellement antibiotiques.

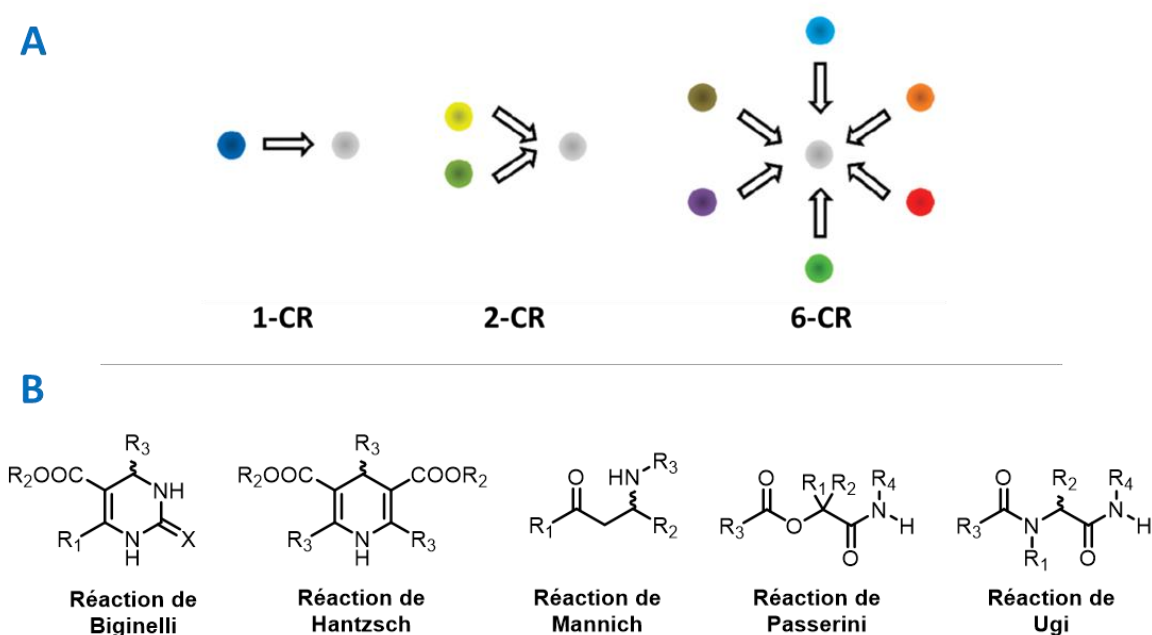
## II.2 Préparation des composants pour la réaction de Ugi-4CR

Dans ce travail de thèse, la préparation de nouveaux ligands d'ARN a été réalisée *via* une réaction multicomposants (MCR) qui repose sur une autre approche que la synthèse dite « classique ». La MCR que nous avons choisie est la réaction de Ugi qui dans un premier temps sera brièvement introduite. Ensuite, nous nous intéresserons aux synthèses des différents composants choisis dans ce projet.

### II.2.1 Réactions multicomposants

L'importance des petites molécules organiques dans la recherche en chimie médicinale est incontestée. De manière générale, faire appel à une synthèse multi-étape pour l'obtention d'une diversité structurale peut être long et fastidieux<sup>151</sup>.

Pour pallier cette problématique, les réactions multicomposants ont été découvertes à la fin du 19<sup>ème</sup> siècle et largement développées par la suite permettant d'accéder à une large diversité structurale et sont notamment utilisées en chimie combinatoire<sup>146</sup>. Les réactions multicomposants sont définies comme des réactions où plus de deux réactifs de départ réagissent ensemble de manière ordonnée formant une molécule cible (*figure 49, A*).



**Figure 49 :** **A)** Comparaison entre une réaction classique à un (1-CR) ou deux composants (2-CR) et une réaction multicomposant à six composants (6-CR)<sup>152</sup>. **B)** Exemples de structures des composés obtenus à la suite d'une réaction multicomposants.

De nombreuses MCR ont été décrites depuis 1850, date de la publication de la réaction de Strecker qui est considérée comme la première MCR<sup>146</sup>. D'autres réactions ont suivi dont les plus connues sont Hantzsch (1882), Biginelli (1891), Mannich (1912), Passerini (1921) et Ugi<sup>153</sup> (1959) (*figure 49, B*).

D'autres structures plus complexes peuvent être également obtenues par des successions de MCR mais aussi par des réactions post-MCR<sup>151</sup>. Grâce à ces nombreuses réactions, divers produits ont été étudiés pour différentes applications en biologie<sup>152</sup>.

L'approche de synthèse par MCR offre en théorie plusieurs avantages qui sont considérés en faveur de la chimie verte. Le principe d'économie d'atomes est mis en avant par le fait que la majorité des atomes initiaux sont incorporés dans la structure finale. Ensuite, ces réactions génèrent souvent très peu ou pas de produits secondaires et nécessitent peu de purifications. Par ailleurs, le produit est souvent isolé avec un rendement élevé. Tous ces avantages permettent d'avoir un gain de temps considérable et des réactions peu onéreuses comparées à une synthèse multi-étape classique.

L'ensemble de ces avantages et la simplicité de mise en œuvre, font de cette approche par MCR un outil adapté pour l'accès à une large série de molécules. Dans le but d'obtenir une diversité rapide d'une structure donnée, la réaction de Ugi a ainsi été choisie.

### II.2.1.1 Réaction de Ugi

La réaction de Ugi<sup>153</sup> développée en 1959 par Ivar Karl Ugi est une réaction à quatre composants (U-4CR) qui sont principalement une amine, un aldéhyde (ou une cétone), un acide carboxylique et un isocyanure. Cette réaction est une des plus importantes réactions de la classe des réactions multicomposants à base d'isocyanure (IMCR)<sup>154</sup> et permet l'accès à des structures peptidomimétiques.

L'isocyanure qui peut à la fois se comporter comme un nucléophile et comme un électrophile définit cette importante classe d'IMCR. Le mécanisme admis dans la littérature pour la réaction de Ugi se divise en plusieurs étapes réversibles et une étape finale irréversible selon deux voies <sup>155</sup> (schéma 4).

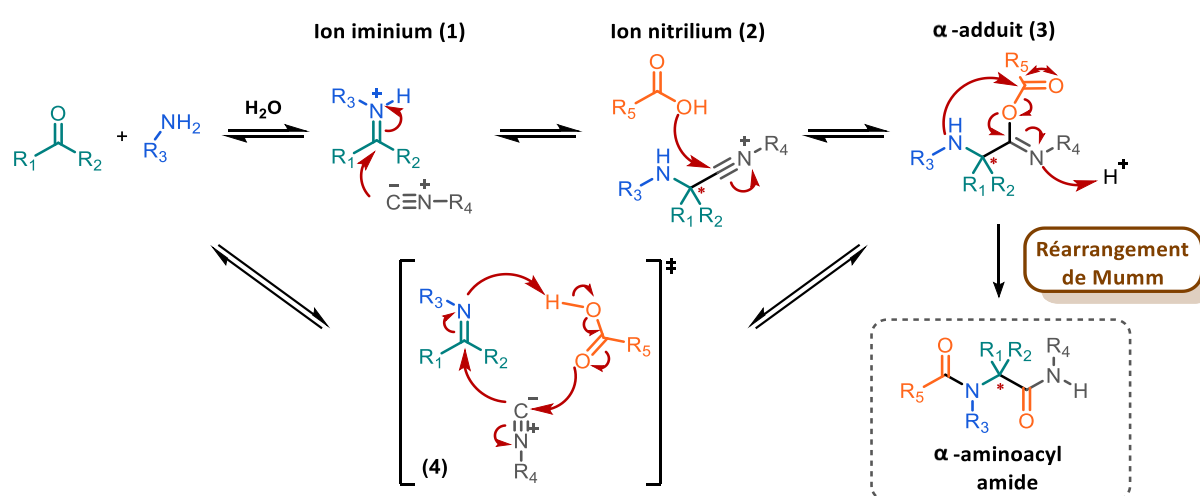


Schéma 4 : Mécanisme de la réaction de Ugi admis dans la littérature.

L'étape initiale est la formation d'une imine (ou iminium) **1** *in situ* à partir de l'amine et du carbonylé. La deuxième étape est une attaque nucléophile de l'isocyanure sur cet intermédiaire pour conduire à



la formation de l'ion nitrilium **2** fortement réactif. Ensuite, l'acide carboxylique réagit sur le centre électrophile de l'isocyanure du composé **2** pour former un  $\alpha$ -adduit activé **3**. Cette  $\alpha$ -adduit peut également être décrit par le passage d'un état transitionnel **4** à trois composantes. Enfin, la dernière étape irréversible est un réarrangement intramoléculaire de Mumm, dans lequel l'amine va attaquer le carbonyle de l' $\alpha$ -adduit pour former le produit bis-amide ou  $\alpha$ -aminoacylamide qui possède un centre stéréogénique. Ainsi, la force motrice de cette réaction est l'oxydation du carbone de l'isocyanure en un amide et donc la formation de deux nouvelles liaisons peptidiques.

### II.2.1.2 Diversité structurale

La réactivité particulière de l'isocyanure permet l'utilisation de différentes catégories d'acide donnant une diversité du squelette de la réaction de Ugi. Par exemple, l'acide carbonique, l'acide thiocarboxylique,  $\text{HN}_3$ , l'eau,  $\text{H}_2\text{S}$ ,  $\text{HNCO}$ ,  $\text{HNCS}$ , et le phénol (réaction Ugi-Smiles) permettent d'accéder à des produits de type  $\alpha$ -acylaminothiocarbonamides, tétrazoles,  $\alpha$ -aminoamides,  $\alpha$ -aminothioamides, hydantoïnes, thiohydantoïnes et  $\alpha$ -aminoarylamides<sup>146</sup> (schéma 5).

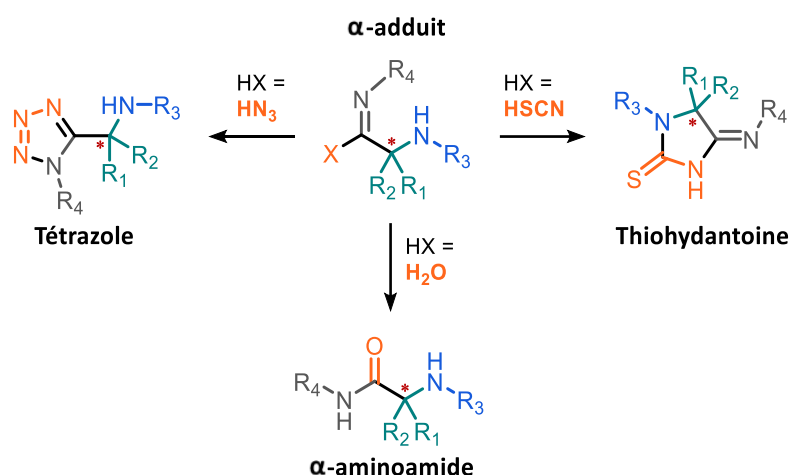


Schéma 5 : Exemples de squelettes formés par une réaction Ugi-type selon l'acide utilisé.

La particularité des structures des composants utilisées lors de la réaction de Ugi permet d'obtenir une diversification structurale post réactionnelle importante notamment pour accéder à des hétérocycles. La procédure UDC ou Ugi-Déprotection-Cyclisation donne accès à la famille des pipérazines<sup>156</sup> ou des  $\gamma$ -lactames<sup>157</sup> par exemple. Les groupements fonctionnels introduits peuvent aussi conduire à des réactions phares telles que Ugi-Heck, Ugi-Diels-Alder ou encore Ugi-Click<sup>154</sup>. La versatilité de la réaction en fait donc une méthode de choix pour la préparation de larges collections de composés.

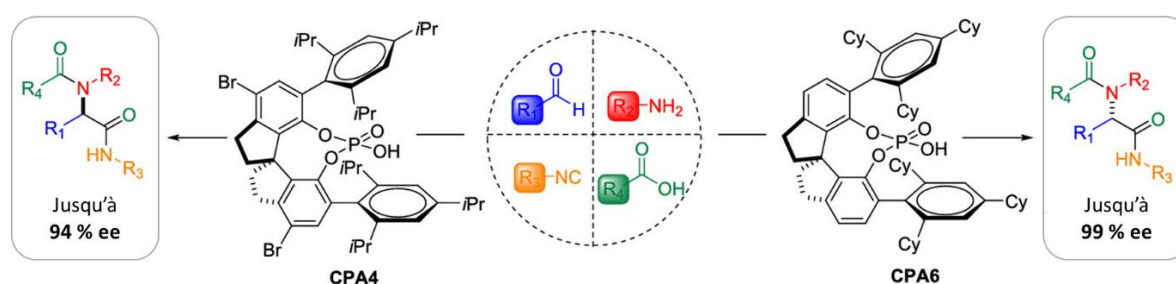
### II.2.1.3 Conditions de réaction de la U-4CR

Comme le montrent les exemples cités précédemment, la réaction de Ugi a été très étudiée et par conséquent, différents paramètres sont importants pour réaliser cette réaction<sup>158</sup>.

Tout d'abord, un paramètre important est le choix du solvant qui peut être contrasté dans la littérature. Les solvants polaires et protiques tels que le méthanol, l'éthanol ou le 2,2,2-trifluoroéthanol sont les plus communément utilisés pour la réaction de U-4CR. Cette préférence s'explique par la formation de liaison hydrogène avec les intermédiaires réactionnels en particulier pour la stabilisation de l'ion iminium. D'autres solvants aprotiques polaires (THF, DMF ou dichlorométhane) peuvent cependant donner de bons résultats dans certains cas, en particulier lorsqu'ils permettent d'améliorer la solubilité de certains réactifs. De même, des mélanges de solvants peuvent être utilisés tels que le CH<sub>2</sub>Cl<sub>2</sub>/MeOH ou le THF/MeOH.

D'un point de vue cinétique, la réaction est réalisée dans des conditions de haute concentration allant de 0,5 à 2 M. L'ordre d'ajout des différents réactifs est quant à lui peu important même s'il est recommandé de préformer le composé carbonylé avec l'amine avant l'ajout des deux autres réactifs. Afin d'augmenter l'efficacité de la formation de l'imine, des acides de Lewis ou de Bronsted peuvent faciliter l'attaque nucléophile de cette amine.

Enfin, comme mentionné dans la description du mécanisme réactionnel de la U-4CR, le produit obtenu est un mélange d'énantiomères du fait de la formation d'un centre chiral. Pour pallier ce manque de stéréocontrôle qui est une des limitations de cette réaction, diverses méthodes comme l'utilisation d'amines inductrices de chiralité<sup>154</sup> ont émergé. Plus récemment, Zhang et *al.*<sup>159</sup> ont décrit une réaction de Ugi énantiosélective avec l'utilisation de dérivés d'acides phosphoriques asymétriques comme catalyseurs (*schéma 6*).



*Schéma 6* : Réaction énantiosélective de Ugi en présence d'un catalyseur chiral d'acide phosphorique (CPA)<sup>160</sup>.

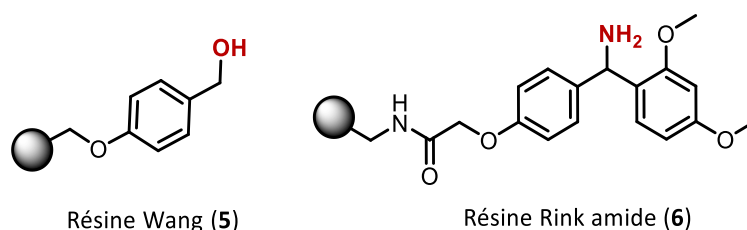
Malgré le manque de stéréospécificité, la réaction de Ugi peut se réaliser dans divers solvants et des conditions douces et conduire à la préparation d'un seul énantiomère si la réaction est effectuée en présence d'une amine chirale ou d'un catalyseur chiral. Les avantages de l'utilisation de cette réaction restent donc très nombreux.

### II.2.1.4 Réaction de U-4CR sur support solide

Pour le développement de ce projet de thèse, nous avons choisi d'utiliser dans un premier temps la réaction de Ugi sur support solide polymérique (SPS). Ainsi, les conditions décrites précédemment peuvent être appliquées mais avec une légère adaptation. La composition du support solide qui est majoritairement en polystyrène, ne permet pas l'utilisation du méthanol seul comme solvant. En effet, ce solvant polaire protique a un effet rétractant de la résine et donc empêche une utilisation maximale de la surface réactionnelle. Afin d'envisager un gonflage idéal de la résine, un mélange de solvants peut être utilisé comme les mélanges  $\text{CH}_2\text{Cl}_2/\text{MeOH}$  ou  $\text{THF}/\text{MeOH}$ .

Un autre paramètre à adapter est le nombre d'équivalents des différents réactifs. La synthèse sur SPS donne un avantage dans les étapes de traitement par rapport à la synthèse en solution. En effet, un excès de réactifs est facilement éliminé sur SPS par simple lavage sans compromettre la pureté du produit obtenu. Ainsi, pour la réaction de Ugi, un excès de 2 éq. à 5 éq. peut être utilisé par rapport au support solide.

De nombreuses IMCR ont été rapportées dans la littérature par réaction sur SPS<sup>161</sup> et en particulier les réactions de Passerini, de Ugi et de Groebke-Blackburn-Binaymé. Différents supports solides ont été développés depuis l'introduction du support solide par Merrifield en 1963<sup>162</sup> et sont facilement accessibles. Ainsi, un choix important de fonctions chimiques et de spécificité du site d'ancrage existe. Selon le site d'ancrage, des conditions basiques (comme lors de l'utilisation de la résine Wang **5**) ou des conditions acides (comme lors de l'utilisation de la résine Rink amide **6**) sont requises (*figure 50*).



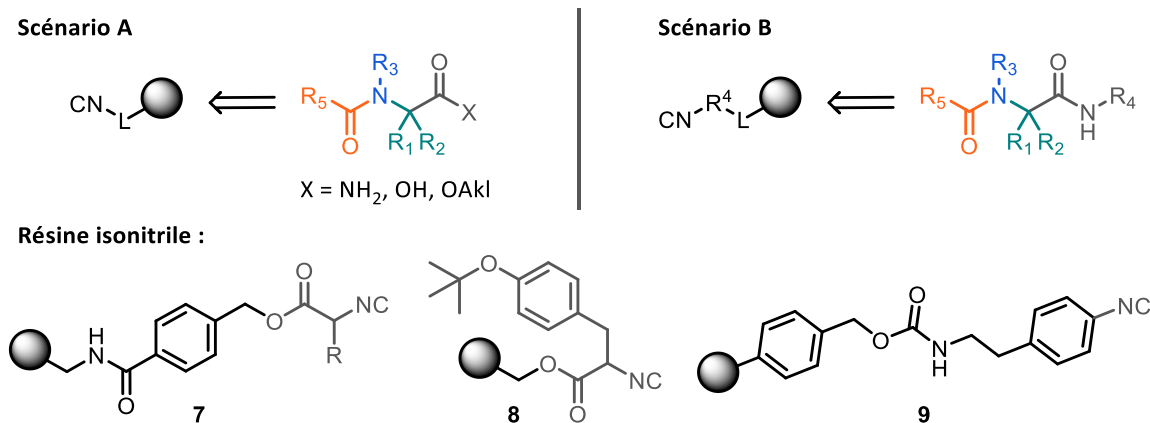
**Figure 50 :** Exemple de support utilisé en synthèse sur support solide polymérique.

Dans le cadre de notre projet, nous avons décidé d'utiliser un support solide fonctionnalisé par une fonction isocyanure. Cependant, ce type de support n'est pas commercial et doit donc être préparé en amont.

### II.2.2 Fonctionnalisation du support solide polymérique

Pour fonctionnaliser un support solide polymérique deux approches sont possibles et permettent d'accéder à deux scénarios suite à une réaction de Ugi<sup>161</sup> (*schéma 51*). Dans le scénario A, le groupement fonctionnel est attaché directement au support solide. Ce support permet d'accéder,

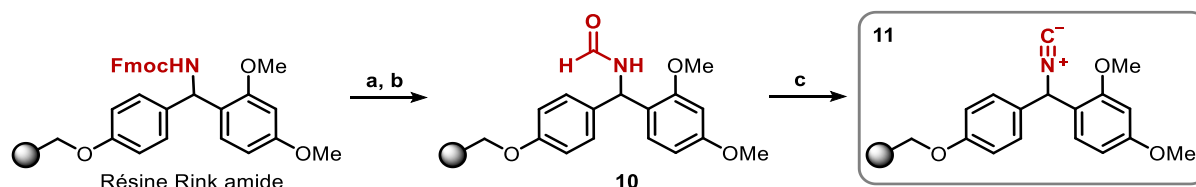
selon les conditions de clivage, à une fonction de type amide, acide carboxylique ou ester. Dans le scénario B, un bras espaceur est inséré entre la fonction isocyanure et le support solide tel que les résines **7**<sup>163</sup>, **8**<sup>164</sup> et **9**<sup>165</sup>. Dans la majorité des cas, le motif du bras espaceur est présent sur la structure de la molécule obtenue à la suite de l'étape de clivage du support solide.



**Figure 51 :** Deux scénarios (A et B) possibles selon le type de fonctionnalisation de la résine avec une fonction isocyanure<sup>161</sup>. Exemples illustrés dans la littérature : la résine **7** et **8** sont sensibles en milieu basique et la résine **9** est dite « safety catch linker » et nécessite une étape supplémentaire pour induire le clivage du support. R = groupements divers introduits par l'aldéhyde (R<sub>1</sub> et R<sub>2</sub>), par l'amine (R<sub>3</sub>), par l'isocyanure (R<sub>4</sub>) et par l'acide carboxylique (R<sub>5</sub>) après une réaction de Ugi.

Dans le cadre de ce projet, le « scénario A » a été envisagé avec une résine **Rink amide** ou une résine **Rink amide MBHA**, où l'étape de clivage de la résine est réalisée dans des conditions acides ce qui conduit à une déprotection des groupements protecteurs acido-labiles concomitante avec ce clivage. Par l'utilisation de cette résine, des composés de type amidique seront obtenus à la suite à la réaction de Ugi<sup>150</sup>.

En 2001, Chen et *al.* ont décrit pour la première fois la préparation de la résine Rink-isocyanure pour la synthèse de 3-acylamido imidazo[1,2-a]pyridines<sup>166</sup>. Nous avons donc choisi d'utiliser cette procédure que nous avons réalisée en trois étapes à partir de différentes résines Rink Amide commerciales (**Schéma 7**).



**Schéma 7 :** Description de la fonctionnalisation de la résine Rink amide en résine Rink-isocyanure **11** en trois étapes. Conditions : **a)** pipéridine/DMF (2/8), t.a., 30-40 min ; **b)** HCOOH, DIC, py., CH<sub>2</sub>Cl<sub>2</sub>, t.a., 2h30 (3 cycles) **c)** POCl<sub>3</sub>, DIEA, CH<sub>2</sub>Cl<sub>2</sub>, t.a., 2h30 (3 cycles).

La première étape est le clivage du groupement protecteur 9-fluorénylméthyle carbamate (Fmoc) dans de la pipéridine/DMF pour libérer la fonction amine. Cette amine est ensuite engagée dans une étape de formylation donnant la résine **10** à l'aide de l'acide formique suite à une activation avec le *N,N'*-

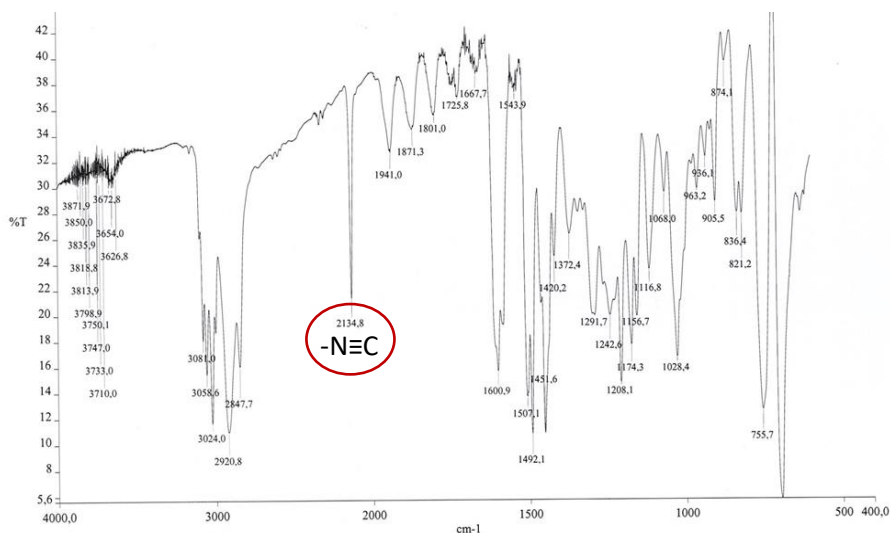
diisopropylcarbodiimide (DIC) en présence de pyridine dans le  $\text{CH}_2\text{Cl}_2$ . Enfin, un traitement par du  $\text{POCl}_3$  en présence de *N,N*-diisopropyléthylamine (DIEA) dans le  $\text{CH}_2\text{Cl}_2$  permet une déshydratation du groupement formamide donnant la résine rink-isocyanure **11**.

Nous avons donc appliqué cette procédure à la résine Rink amide et la résine Rink amide MBHA. Cette dernière possède une accroche plus longue et nécessite aussi des conditions plus drastiques lors du clivage du support solide par rapport à la résine Rink amide. Cette fonctionnalisation a également été réalisée sur la résine NovaGel™ Rink amide qui a l'avantage de pouvoir admettre le méthanol comme solvant sans phénomène de rétraction de celle-ci.

Afin de surveiller la progression des réactions et l'obtention de la fonction isocyanure, différentes analyses qualitatives et quantitatives ont été réalisées.

- **Analyses qualitatives :**

Pour les deux premières étapes, l'avancement de la réaction a été suivi par un test de Kaiser qui est couramment utilisé pour le suivi de couplage peptidique sur SPS. Ensuite pour l'étape c, des analyses par spectroscopie infrarouge ont été nécessaires afin de visualiser l'apparition de la bande caractéristique de la fonction isocyanure à  $2135\text{ cm}^{-1}$  (figure 52).



**Figure 52 :** Analyse par spectroscopie Infrarouge de la résine Rink amide isocyanure (KBr) :  $\nu = 2135\text{ cm}^{-1}$  ( $\text{C}\equiv\text{N}$  isocyanure).

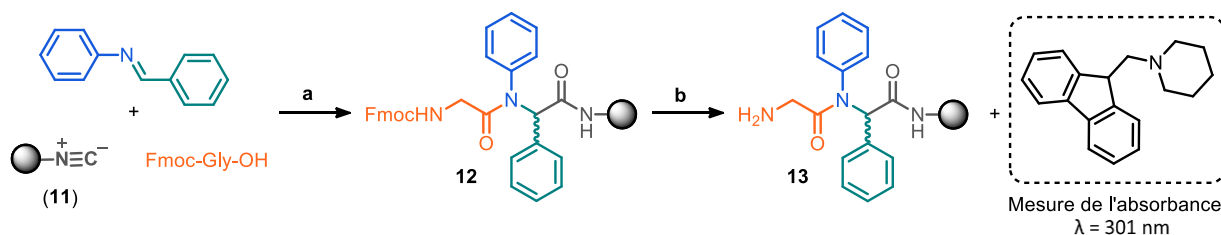
La caractérisation par IR a aussi été réalisée pour les deux intermédiaires (**étapes 1 et 2**) préparés sur la résine Rink amide. L'apparition et la disparition des bandes caractéristiques du formamide ont pu être observées ( $\nu = 1690\text{ cm}^{-1}$ ,  $\text{C}=\text{O}$ ) (annexe 3). Pour la résine Rink amide MBHA, le suivi de ces 2 intermédiaires n'a pas été possible par IR en raison de la structure du bras espaceur contenant déjà des liaisons amides et donc seule l'analyse finale a été réalisée.

Cependant, ces analyses n'ont pas permis d'obtenir des informations sur le taux exact de la fonctionnalisation ainsi des analyses quantitatives ont été envisagées.

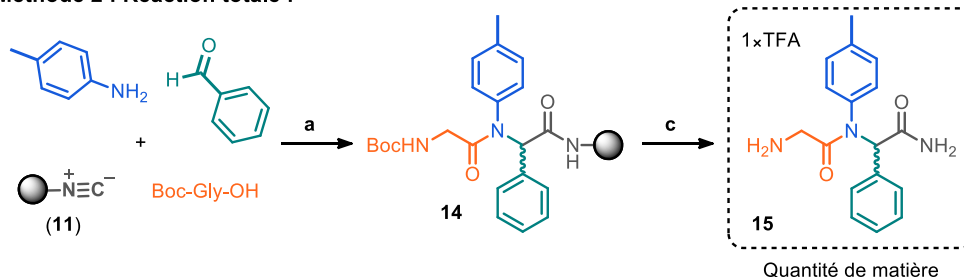
- **Analyses quantitatives :**

Deux approches pour quantifier la fonctionnalisation de la résine ont été employées. Dans un premier temps, la procédure décrite par Henkel et *al.*<sup>167</sup> a été appliquée. Dans cette procédure, la quantification est obtenue par l'utilisation d'un dosage de l'adduit fulvène (dibenzofulvène) sur lequel s'est ajouté la pipéridine lors du clivage du groupement Fmoc (*Schéma 8*). Puis, ce test a été adapté pour obtenir un test témoin *via* la formation du produit **15** du support solide.

**Méthode 1 : Dosage par l'adduit fulvène :**



**Méthode 2 : Réaction totale :**



**Schéma 8 :** Représentation des méthodes pour déterminer quantitativement le taux de fonctionnalisation de la résine. **Méthode 1)** Dosage par l'adduit dibenzofulvène-pipéridine à la longueur d'onde  $\lambda = 301 \text{ nm}$ . **Méthode 2)** Quantité obtenue du produit **15** à la suite au clivage de la résine. Conditions : **a)**  $\text{CH}_2\text{Cl}_2/\text{MeOH}$ , t.a., 5j ; **b)** pipéridine/DMF (2/8), 15 min, t.a. ; **c)** TFA,  $\text{CH}_2\text{Cl}_2$ , t.a., 1h30.

Le dosage de l'adduit fulvène a été réalisé à la suite d'une réaction de Ugi impliquant un excès des différents réactifs (5 éq.) par rapport à la résine afin de favoriser une conversion totale. L'imine commerciale alanine *N*-benzylidèneaniline a été mise à réagir avec la glycine protégée par le groupement Fmoc (*N* $\alpha$ -Fmoc-Gly-OH) et la résine fonctionnalisée à étudier. La réaction a été laissée sous agitation dans un mélange  $\text{CH}_2\text{Cl}_2/\text{MeOH}$  pendant 5 jours. Après traitement et séchage de cette résine-Ugi (**12**), une analyse par IR a permis de visualiser la disparition de la bande de l'isocyanure montrant une réaction totale. Ensuite, une quantité précise de la résine-Ugi (**12**) a été mise dans un volume donné d'une solution de pipéridine (20%) dans le DMF puis analysée par spectroscopie d'absorbance après dilution ( $\lambda = 301 \text{ nm}$  de l'adduit dibenzofulvène-pipéridine). L'absorbance obtenue a permis de remonter au rendement sur l'ensemble des 5 étapes. Cette expérience a été répétée 4 fois pour obtenir une valeur moyenne.

En parallèle sur le même lot de résine fonctionnalisée, la réaction témoin a été effectuée. Cette réaction de Ugi a été réalisée avec du benzaldéhyde (2 éq.), de la *p*-toluidine (2 éq.) en présence de la glycine protégée avec un groupement Boc (*N* $\alpha$ -Boc-Gly). Après 6 jours de réaction, le produit résine-Ugi (**14**) a été clivé et la quantité obtenue du produit **15** a permis de déterminer le rendement sur les 5 étapes. La pureté du produit obtenu a été confirmée par une analyse par HPLC à 214 nm et est supérieure à 95%.

Ces deux approches ont été comparées sur un même lot de résine fonctionnalisée (*tableau 5*).

**Tableau 5** : Comparaison des résultats obtenus avec le test par dosage de l'adduit fulvène (*méthode 1*) et avec le test témoin (*méthode 2*) pour la résine Rink amide MBHA :

Méthodes	Solvants	Taux de substitution (mmol/g)		Rendement (%)
		Théorique	Expérimental	
<b>1</b>	CH <sub>2</sub> Cl <sub>2</sub> /MeOH	0,79	0,33	42 % <sup>b</sup>
<b>2</b>	CH <sub>2</sub> Cl <sub>2</sub> /MeOH	0,79	0,43	54 %

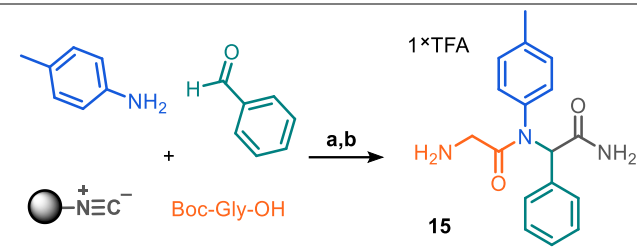
<sup>a</sup> nombre d'équivalents par rapport au taux de substitution de la résine de départ non fonctionnalisée. <sup>b</sup> moyenne effectuée sur 4 expériences.

En considérant que les étapes de la réaction de Ugi et de déprotection du groupement Fmoc ou Boc sont totales, la méthode 1 utilisant le dosage de l'adduit fulvène, nous a donné un taux de fonctionnalisation de 42 %. Pour la méthode 2, le produit **15** a été obtenu après 5 étapes avec un rendement de 54 % ce qui correspond au taux de fonctionnalisation. La différence entre les deux méthodes pourrait être due aux différentes erreurs de mesure notamment pour la méthode 1 ou encore d'une déprotection incomplète du groupement Fmoc.

Ces taux de fonctionnalisation sont encourageants et nous permettent par la suite d'utiliser cette réaction témoin pour chaque lot de résine fonctionnalisée afin d'ajuster le rendement final des réactions de Ugi.

Par cette réaction témoin, nous avons aussi déterminé les conditions de la réaction de Ugi sur support solide qui seront utilisées par la suite. Ces différentes conditions étudiées sont résumées dans le *tableau 6*.

Tableau 6 : Optimisation des conditions de la réaction de Ugi sur support solide :



Résine	Entrée	Solvants	Équiv. <sup>a</sup>	Taux de substitution (mmol/g)		Rendement (%)
				Théorique	Expérimental	
Rink amide MBHA	1	CH <sub>2</sub> Cl <sub>2</sub> /MeOH	2	0,79	0,43	54 %
	2a	CH <sub>2</sub> Cl <sub>2</sub> /MeOH	2	0,79	0,19	24 %
	2b	CH <sub>2</sub> Cl <sub>2</sub> /MeOH	1	0,79	0,17	22 %
	2c	THF/MeOH	2	0,79	0,16	21 %
	3	CH <sub>2</sub> Cl <sub>2</sub> /MeOH	2	0,52	0,20	38 %
Rink amide NovaGel	4a	CH <sub>2</sub> Cl <sub>2</sub> /MeOH	2	0,66	0,17	26 %
	4b	THF/MeOH	2	0,66	0,079	12 %
Rink amide	5	CH <sub>2</sub> Cl <sub>2</sub> /MeOH	5	0,45	0,19	43 %

Conditions : **a**) t.a., 5j ; **b**) TFA, CH<sub>2</sub>Cl<sub>2</sub>, t.a., 1h30. <sup>a</sup> nombre d'équivalents par rapport au taux de substitution de la résine de départ non fonctionnalisée.

À partir des résultats de la méthode 2 décrits précédemment (**entrée 1**), le nombre d'équivalents (**entrée 2b**) et le mélange de solvants (**entrée 2c**) ont été évalués à partir d'un nouveau lot de résine fonctionnalisée. En comparaison avec les conditions initiales (**entrée 2a**), il n'a pas été observé de différence significative sur le rendement total sur les 5 étapes. Cependant, un manque de répétabilité a été observé entre les trois lots fonctionnalisés pour cette même résine (**entrées 1, 2a et 3**).

Différentes résines ont également été étudiées pour comparer le taux de fonctionnalisation. La résine Rink amide NovaGel™ (**entrée 4**), qui se prête à l'utilisation du méthanol comme solvant, ne nous a pas apporté un meilleur taux de fonctionnalisation. Enfin, pour la résine Rink amide (**entrée 5**), nous avons obtenu un rendement satisfaisant de 43 % avec un nombre d'équivalents supérieur (5 éq.), afin d'obtenir une réaction totale.

Ainsi, les différentes analyses qualitatives et quantitatives nous ont permis de mettre au point la fonctionnalisation des résines ainsi que les conditions privilégiées pour la réaction de Ugi. En conclusion, deux types de résines fonctionnalisées par un groupement isocyanure seront utilisées : la résine Rink amide et la résine Rink amide MBHA toutes deux sensibles en milieu acide.



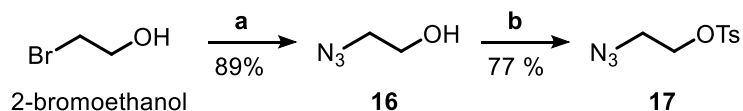
Afin de réaliser les futures réactions de Ugi, les synthèses nécessaires des deux autres composants de la réaction de Ugi qui comportent soit la fonction amine soit la fonction aldéhyde, seront décrites. Pour rappel, les motifs comportant la fonction acide carboxylique sont des acides aminés commerciaux.

### II.2.3 Préparation des motifs ayant la fonction amine

Les synthèses des motifs des nucléobases modifiées (A, C, G et T), des analogues du motif naphthyridine et de la nucléobase S seront détaillées ci-après.

#### II.2.3.1 Préparation des nucléobases naturelles alkylées

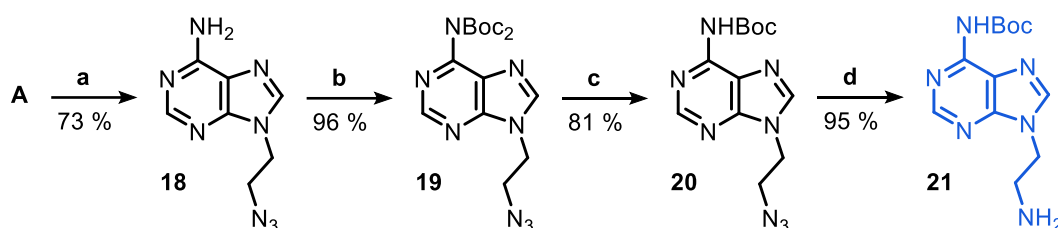
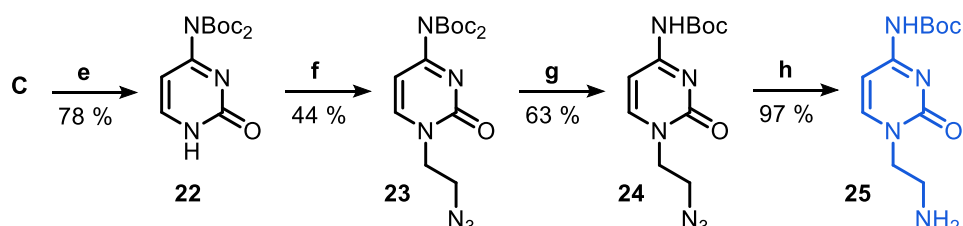
La stratégie utilisée pour la synthèse des nucléobases modifiées a été dirigée par le choix des groupements protecteurs. Afin d'éliminer l'ensemble des groupements protecteurs lors du clivage de la résine Rink amide, qui est une résine acido-labile, le groupement *tert*-butoxycarbonyle (Boc) a été choisi. Ce groupement a donc été utilisé pour protéger les amines exo-cycliques de ces nucléobases. Pour introduire, au niveau des nucléobases, un bras espaceur porteur de la fonction amine primaire impliquée dans la réaction de Ugi, nous avons choisi d'alkyler les nucléobases avec un bras activé par un tosylate et contenant un groupement azoture (2-azidoéthyl-*p*-toluènesulfonate) **17** qui est couramment utilisé au sein de notre laboratoire. La fonction azoture permet de régénérer l'amine primaire par hydrogénation catalytique, elle est donc compatible avec le groupement Boc. La préparation du bras N<sub>3</sub>-OTs est décrite dans le *schéma 9*.



*Schéma 9* : Synthèse du composé 2-azidoéthyl-*p*-toluènesulfonate **17**. Conditions : **a**) NaN<sub>3</sub>, H<sub>2</sub>O, 80°C, 16h ; **b**) *p*Ts-Cl, Et<sub>3</sub>N, CH<sub>2</sub>Cl<sub>2</sub>, t.a., 16h.

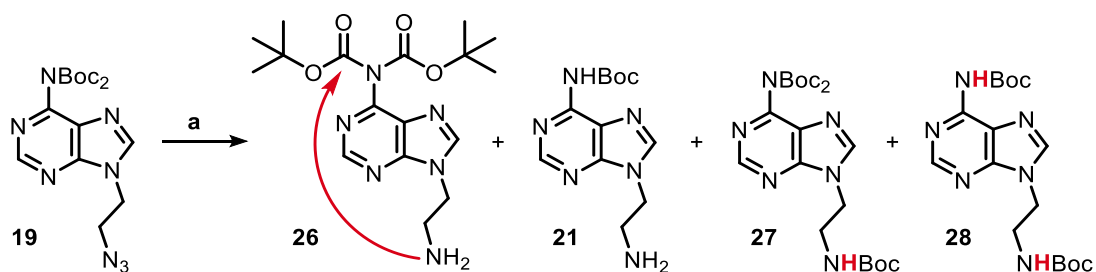
La synthèse du composé **17** a été réalisée en deux étapes à partir du 2-bromoéthanol commercial. Une substitution nucléophile de l'atome de brome par l'azoture en milieu aqueux donne lieu dans un premier temps au composé **16** avec un rendement de 89 %. La fonction hydroxyle du composé **16** est ensuite transformée en tosylate avec un rendement de 77 %. Le caractère nucléofuge de ce groupement permet d'envisager les différentes alkylations sur les nucléobases souhaitées.

Au travers de cette stratégie, les nucléobases modifiées adénine (A), cytosine (C) et guanine (G) ont été synthétisées en suivant les protocoles décrits dans la littérature<sup>168,169,170</sup>. La synthèse des nucléobases modifiées de l'adénine et de la cytosine suivent la même stratégie de synthèse (*schéma 10*).

Voie de synthèse de l'adénine **21**:Voie de synthèse de la cytosine **25** :

**Schéma 10** : Synthèses des nucléobases modifiées de l'adénine **21** et de la cytosine **25**. Conditions : **a**) 2-azidoéthyl-*p*-toluènesulfonate,  $K_2CO_3$ , DMF, 60°C, 16h ; **b**)  $Boc_2O$ , DMAP, THF, t.a., 16h ; **c**)  $K_2CO_3$ , MeOH, 56°C, 2h ; **d**)  $H_2$ , Pd/C, EtOH, t.a., 16h ; **e**) (i)  $Boc_2O$ , DMAP, THF, t.a., 3h, (ii)  $NaHCO_3$ , MeOH, reflux, 2h ; **f**) 2-azidoéthyl-*p*-toluènesulfonate, NaH, DMF, 60°C, 16h ; **g**)  $K_2CO_3$ , MeOH, 56°C, 1h ; **h**)  $H_2$ , Pd/C,  $CH_2Cl_2$ , t.a., 16h.

La synthèse du composé **21** a été réalisée en quatre étapes. L'alkylation régiosélective en position  $N_9$  de l'adénine commerciale a été effectuée en présence de  $K_2CO_3$  et du dérivé tosylate **17**. Le composé alkylé **18** a été obtenu avec un rendement de 73 %. La présence d'un composé *N*-alkylé sur l'amine exo-cyclique a également été observée lors de cette étape. Ensuite, une protection de l'amine exo-cyclique a été réalisée en présence de DMAP donnant le composé di-Boc **19** avec un rendement de 96 %. À cette étape de la synthèse, nous avons converti le composé di-protégé **19** en composé mono-protégé **20**. Cette suppression sélective d'un groupement Boc est nécessaire avant l'étape d'hydrogénation catalytique qui suit. En effet, lorsque nous avons réalisé l'étape d'hydrogénation catalytique directement sur l'intermédiaire **19**, un mélange de différents produits (**21**, **26**, **27** et **28**) a été obtenu et caractérisé donnant un rendement faible du composé stable souhaité **21** (*schéma 11*).



**Schéma 11** : Structures des sous-produits obtenus lors de l'hydrogénation catalytique du composé di-Boc **19** : **a**)  $H_2$ , Pd/C, EtOH, t.a., 16h.

Dans un premier temps, le produit issu de l'hydrogénation a donné le produit di-Boc **26** qui a été isolé et identifié. Cependant, très vite, il a été observé une migration d'un groupement *tert*-butoxycarbonyl du composé **26** sur l'amine primaire du bras espaceur, donnant un mélange de produits tri-Boc **27** et

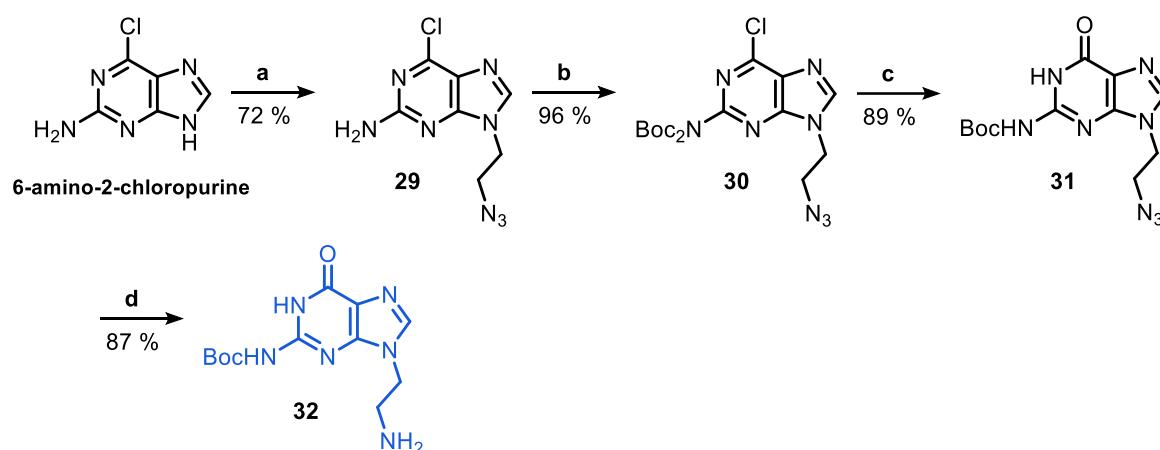
di-Boc **28**. Cette migration intra-ou inter molécules provient d'une instabilité du deuxième groupement Boc de l'amine exo-cyclique. La caractérisation de ces sous-produits **27** et **28** a été réalisée par RMN  $^1\text{H}$  en comparant les déplacements des hydrogènes NH (*annexe 2*).

Afin d'éviter ce problème majeur retrouvé dans l'ensemble de nos synthèses possédant une amine exo-cyclique protégée par deux groupement Boc, une étape supplémentaire a donc été introduite avant l'hydrogénation catalytique.

Dans la littérature, une suppression sélective du groupement Boc pour des amines primaires est réalisée par le biais d'acides de Lewis. L'interaction bivalente de l'acide avec les deux carbonyles des groupements Boc entraîne cette déprotection<sup>171</sup>. Darnbrough et *al.* ont décrit une élimination sélective d'un des groupements Boc de composés amine-aryl dans des conditions basiques<sup>172</sup>. La réaction est effectuée dans le méthanol en utilisant un excès de  $\text{K}_2\text{CO}_3$  à reflux. Par cette dernière étude, nous avons appliqué ces mêmes conditions en utilisant un excès de  $\text{K}_2\text{CO}_3$  dans le méthanol à reflux qui nous ont permis de cliver sélectivement un Boc de **19** pour obtenir le composé **20** avec un rendement de 81 %. Ensuite, l'étape finale d'hydrogénation catalytique permettant d'obtenir l'amine souhaitée a été réalisée dans des conditions standard et a permis d'obtenir le produit souhaité **21** avec un rendement de 95 % après purification.

Similaire à la synthèse de l'adénine, la préparation de l'analogue contenant la cytosine modifiée **25** a été effectuée en quatre étapes (cf. *schéma 10*). Dans le cas de cette nucléobase, une protection de l'amine exo-cyclique est nécessaire avant de procéder à l'alkylation régiosélective en  $N_1$ . Dans nos conditions de protection, la présence d'un intermédiaire tri-Boc avec un groupement sur la position  $N_1$ , a été observée. Pour libérer cette position, le mélange réactionnel a été porté à reflux dans le méthanol pendant 1h en présence de  $\text{NaHCO}_3$  donnant le composé protégé **22** pour un rendement global de 78 %. Ce dernier a ensuite été activé dans le DMF par de l'hydrure de sodium avant l'ajout du réactif **17**. Le composé **23** résultant de cette alkylation sélective a été alors obtenu avec un rendement de 44 %. La présence du composé mono-Boc **24** (8-10 %) a également été observée lors de cette étape. Une élimination sélective d'un des groupements Boc du composé **23** a ensuite été effectuée donnant le composé mono-Boc **24** avec un rendement de 63 %, mais également une faible proportion de produit sans groupement protecteur. Enfin, l'hydrogénation catalytique a permis d'obtenir la cytosine modifiée **25** avec un rendement de 97 %.

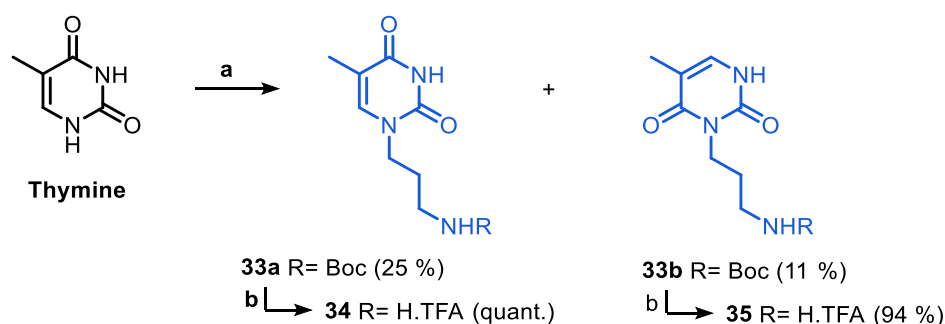
En ce qui concerne la nucléobase guanine modifiée, une autre stratégie a été utilisée en partant de la 6-amino-2-chloropurine commerciale qui est un analogue de la guanine fréquemment utilisé<sup>168</sup> (*schéma 12*).



**Schéma 12** : Synthèse de la guanine modifiée **32**. Conditions : **a**) 2-azidoéthyl-*p*-toluènesulfonate,  $K_2CO_3$ , DMF,  $60^\circ C$ , 16h ; **b**)  $Boc_2O$ , DMAP, THF, t.a., 16h ; **c**) 3-hydroxypropionitrile, NaH, THF,  $-78^\circ C$  à t.a., 16h ; **d**)  $H_2$ , Pd/C, EtOH, t.a., 16h.

L'utilisation de la 2-chloro-6-amino-purine permet d'augmenter la régiosélectivité de l'alkylation en position  $N_9$  par l'effet de l'atome du chlore en position 2<sup>173</sup>. Contrairement au résultat obtenu avec la guanine commerciale (30 % de rendement), le composé **29** a été obtenu avec un rendement de 72 % en mettant la 2-chloro-6-aminopurine en présence du réactif 2-azidoéthyl-*p*-toluènesulfonate **17** et de la base  $K_2CO_3$ . Le composé **29** a ensuite été doublement protégé et le composé **30** isolé avec un rendement de 96 %. Le remplacement l'atome de chlore du composé **30** par un atome d'oxygène a été réalisé au moyen du 3-hydroxypropionitrile (4.5 éq.) en présence d'hydrure de sodium (5 éq.). Cette conversion passe par une première substitution nucléophile aromatique suivie d'une  $\beta$ -élimination de celui-ci. Le composé **31** a été obtenu avec un rendement de 89 %. Lors de cette étape, un des groupements Boc a été clivé et le composé mono-Boc souhaité avant l'hydrogénation isolé. Cette dernière étape d'hydrogénation forme un précipité dans l'éthanol permettant d'obtenir le produit **32** sans purification.

Enfin, la quatrième nucléobase modifiée est la nucléobase thymine qui a été préalablement synthétisée au sein du laboratoire. Pour obtenir ce fragment, une synthèse en deux étapes a été réalisée à partir de la thymine commerciale (*schéma 13*).



**Schéma 13** : Synthèse de la thymine modifiée en position  $N_1$  (**34**) et en position  $N_3$  (**35**). Conditions : **a**) 1-(*N*-tert-butoxycarbonyl)amino-3-bromopropane,  $K_2CO_3$ , DMF,  $65^\circ C$ , 7h ; **b**) TFA,  $CH_2Cl_2$ , t.a., 16h.

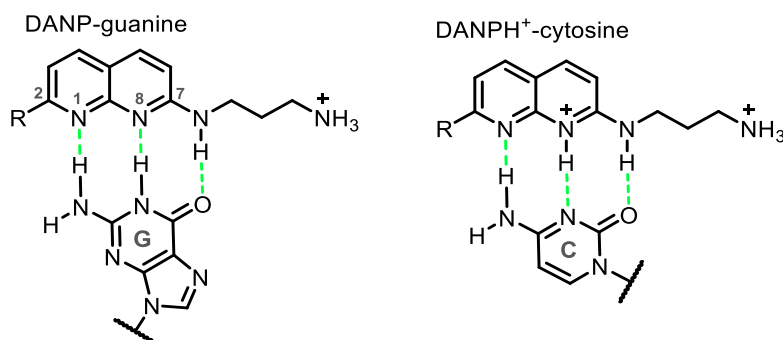
Dans une première étape, une alkylation non sélective a été réalisée en présence du réactif 1-(*N*-*tert*-butoxycarbonyl)amino-3-bromopropane et de la base  $K_2CO_3$ . Par cette alkylation, deux composés ont été obtenus dont le composé **32a** alkylé en position  $N_1$  avec un rendement de 25 % et le composé **32b** alkylé en position  $N_3$  avec un rendement de 11 %. Ce dernier a néanmoins été utilisé lors des réactions de Ugi effectuées successivement car il peut aussi former des liaisons hydrogènes avec l'ARN, même si ces interactions sont à priori moins spécifiques que celles de la thymine alkylée en position  $N_1$ . Enfin, un clivage du groupement Boc a été réalisée séparément pour accéder aux composés **34** alkylé en position  $N_1$  et **35** alkylé en position  $N_3$  avec des rendements quantitatifs et de 94 % respectivement.

En résumé, dans cette partie, quatre bases naturelles ont été obtenues comportant la fonction amine sur le bras espaceur se trouvant soit en position  $N_9$  pour les dérivés de l'**adénine** (**21**, 4 étapes, 54 %) et de la **guanine** (**32**, 4 étapes, 54 %) soit en position  $N_1$  pour le dérivé de la **cytosine** (**25**, 4 étapes, 21 %) et de la **thymine** (**34**, 2 étapes 25 %) ainsi qu'en position  $N_3$  pour la **thymine** (**35**, 2 étapes, 11 %).

### II.2.3.2 Synthèse des analogues naphtyridines

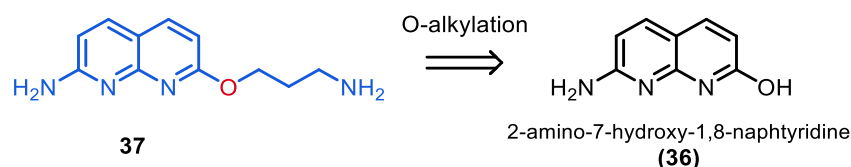
Dans cette partie, la synthèse des analogues naphtyridines sera abordée. Le noyau 1,8-naphtyridine est un composé hétérocyclique de la famille des diazanaphtalènes comportant deux cycles à six chaînons accolés avec deux atomes d'azote en positions 1 et 8. Les molécules possédant cette structure sont nombreuses et ont été beaucoup étudiées pour différentes activités biologiques<sup>174</sup>. Ce motif est présent notamment dans la famille des antibiotiques ciblant l'ADN gyrase bactérien tels que l'acide nalidixique<sup>16</sup>. Ce motif a également été employé dans des études supramoléculaires ou intramoléculaires mettant en avant la possibilité que les atomes d'azote agissent en tant que donneurs ou accepteurs de liaison hydrogène<sup>175,176,177</sup>.

Les analogues 1,8-naphtyridine ont particulièrement été étudiés par l'équipe du Pr. Nakatani à travers divers travaux ciblant dans un premier temps l'ADN pour la détection de polymorphismes nucléotidiques simples<sup>178</sup>. Par la suite, ce type de composé a été étudié sur une structure d'ARN en tige-boucle contenant un renflement et il a été montré l'interaction du composé DANP avec les nucléobases guanine et cytosine par trois liaisons hydrogènes<sup>179</sup> (*figure 53*).



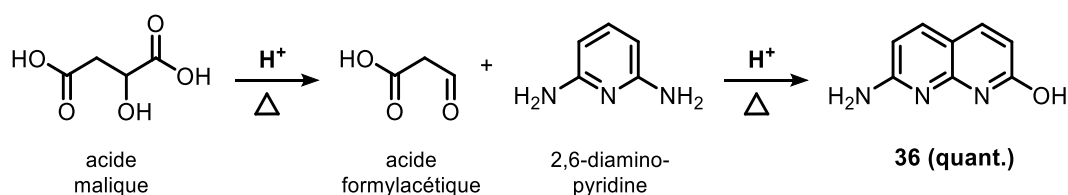
**Figure 53 :** Structure de la DANP et des liaisons hydrogènes dans les complexes DANP-guanine et DANPH<sup>+</sup>-cytosine dans une solution à pH = 7.0.

Au regard de ces études, nous avons choisi d'utiliser ce support 1,8-naphtyridine pour avoir la possibilité de former des interactions hydrogènes avec les guanines non appariées présentes notamment dans les boucles apicales ou dans les renflements des ARNs étudiés. Afin d'insérer le bras espaceur permettant l'introduction de la fonction de l'amine primaire nécessaire pour la réaction de Ugi, il a été envisagé une *O*-alkylation permettant un accès rapide vers l'analogue **37** à partir de motif 2-amino-7-hydroxy-1,8-naphtyridine (**36**) (schéma 14).



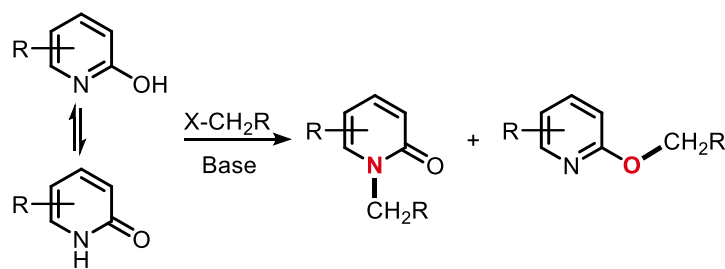
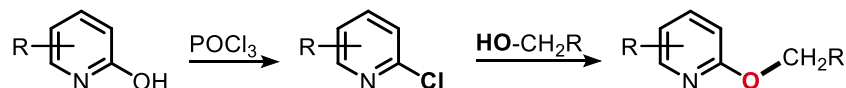
**Schéma 14 :** Représentation de l'analogue 1,8-naphtyridine 36 obtenu par une *O*-alkylation.

Afin d'accéder à cet analogue, le motif 2-amino-7-hydroxy-1,8-naphtyridine (**36**) a d'abord été synthétisé en utilisant la procédure décrite par Newkome et *al.* au travers d'une réaction « one pot »<sup>180</sup> (schéma 15).



**Schéma 15 :** Synthèse du motif 2-amino-7-hydroxy-1,8-naphtyridine **36**. Conditions : H<sub>2</sub>SO<sub>4</sub>, 110°C, 3h, 90 % à quant.

Le traitement de l'acide malique dans l'acide sulfurique à 110°C a permis d'obtenir l'acide formylacétique par décarboxylation. Le traitement de cet acide avec de la 2,6-diaminopyridine dans une réaction de type « Knorr », a conduit au composé **36** souhaité après neutralisation avec une solution ammoniacale (25 %). Le produit **36** a été isolé par précipitation donnant des rendements allant de 90 % à quantitatif. Une fois le composé **36** obtenu, deux stratégies de synthèse se distinguent pour accéder au composé **37** (schéma 16).

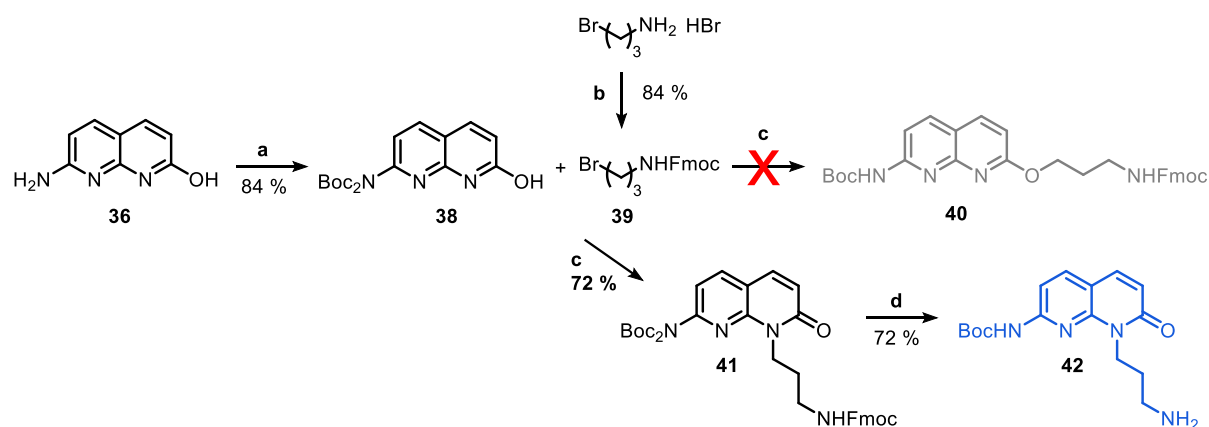
**Voie A : N-alkylation VS O-alkylation****Voie B : via une S<sub>N</sub>Ar**

**Schéma 16 :** Description des voies de synthèse A et B afin d'obtenir le composé O-alkylé. S<sub>N</sub>Ar = Substitution nucléophile aromatique.

Le motif central du 2-amino-7-hydroxy-1,8-naphtyridine ressemble fortement au motif pyridone avec un cycle accolé. Les pyridones sont connues pour réagir avec un dérivé halogéné afin d'obtenir selon diverses conditions les composés N- ou O-alkylés (*schéma 16, Voie A*). La **Voie A** permet d'obtenir les composés O-alkylés en partant directement du motif (**36**). La **Voie B** nécessite une halogénéation préalable pour ensuite réaliser une substitution nucléophile aromatique avec un composé hydroxyle et permettant d'obtenir exclusivement le composé O-alkylé. Ces deux voies de synthèses ont été exploitées et sont décrites ci-dessous.

### II.2.3.2.1 Synthèse de l'analogue naphthyridine par la voie A

La régiosélectivité de l'alkylation sur le motif pyridone dépend dans un premier temps de la base utilisée mais également du réactif halogéné, du substrat pyridone et du solvant utilisé. Rashad et *al.*<sup>181</sup> ont décrit une procédure permettant une O-alkylation d'un dérivé pyridone dans des conditions utilisant K<sub>2</sub>CO<sub>3</sub> dans l'acétone. Avec un motif relativement proche du nôtre, Leonard et *al.*<sup>182</sup> ont utilisé des conditions similaires pour une O-alkylation du composé de départ 2-hydroxy-4-méthyl-7-amino-1,8-naphtyridine. Ces conditions ont été choisies pour la voie de synthèse A décrite ci-dessous pour tenter de réaliser la O-alkylation sur le composé 2-amino-7-hydroxy-1,8-naphtyridine **36** (*schéma 17*). Dans cette voie de synthèse, une étape de protection de l'amine exo-cyclique a été réalisée en amont, afin d'obtenir la protection pour l'étape de la réaction de Ugi. L'utilisation de la DMAP a été nécessaire pour obtenir le composé **38** di-protégé avec un rendement de 84 %.



**Schéma 17** : Synthèse de l'analogue naphtyridine N-alkylé 42. Conditions : **a)**  $\text{Boc}_2\text{O}$ , DMAP, THF, t.a., 16h ; **b)** Fmoc-OSu,  $\text{NaHCO}_3$ ,  $\text{H}_2\text{O}$ /dioxane, t.a., 16h ; **c)**  $\text{K}_2\text{CO}_3$ , acétone,  $56^\circ\text{C}$ , 16h ; **d)** pipéridine/DMF (2/8), t.a., 3h.

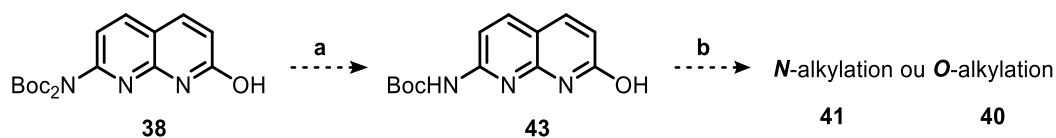
Une fois protégé, le composé **38** a été engagé dans la réaction d'alkylation avec le réactif halogéné **39**. Ce réactif **39** a été préalablement préparé à partir du bromhydrate de 3-bromopropylamine qui a été protégé par le groupement 9-fluorénylméthyle carbamate (Fmoc) orthogonal au groupement Boc avec un rendement de 84 %. Ensuite, l'alkylation du composé **38** a été réalisée par l'ajout de la base  $\text{K}_2\text{CO}_3$  dans l'acétone pour amorcer la réaction avant l'ajout du composé **39**. Contrairement à ce que nous avons envisagé à partir de la littérature, cette réaction n'a pas été en faveur de la O-alkylation (**40**) mais a conduit au composé N-alkylé **41**. À la suite de ce résultat nous avons décidé de poursuivre la préparation du composé **42** et de l'étudier ce composé en tant que fragment de ligands d'ARN. Ce composé pourra être comparé avec le composé O-alkylé par la suite.

Ainsi, la dernière étape de cette synthèse a consisté à cliver le groupement Fmoc par la pipéridine dans le DMF du composé **41**. Dans un premier temps, nous avons réalisé la synthèse sur une petite échelle et nous avons isolé le composé **42** comportant un seul Boc avec un rendement de 72 %. Une faible proportion de produit résultant d'une migration du groupement Boc sur l'amine primaire a néanmoins été identifiée (confirmé par RMN  $^1\text{H}$  et COSY).

Cependant, il faut noter que lorsque cette synthèse a été réalisée à l'échelle du gramme, l'étape d'alkylation a conduit à un clivage partiel du groupement Fmoc post alkylation. Pour pallier ce problème, le mélange réactionnel après alkylation a été directement engagé pour le clivage du groupement Fmoc. Pour ces réactions à l'échelle du gramme, un rendement du composé **42** de 55 % a été obtenu sur les deux étapes.

Ce rendement moyen sur deux étapes peut s'expliquer en partie par l'obtention d'une migration de Boc qui a été plus importante lors de la synthèse sur une plus grande échelle. Afin d'éviter la migration du groupement Boc, il faudra convertir le composé di-Boc **38** en composé mono-Boc **43** (schéma 18).

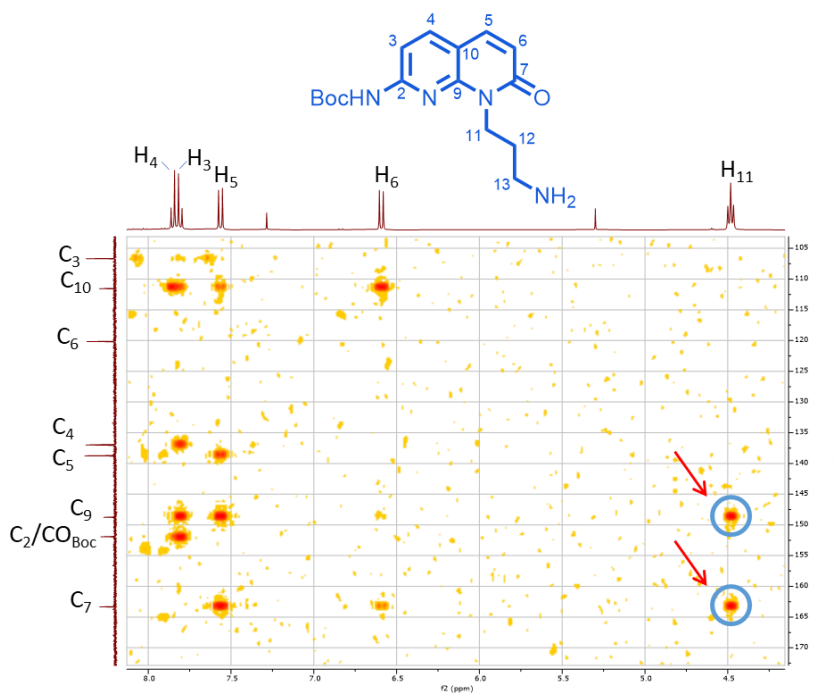




**Schéma 18** : Conversion du composé di-Boc **38** en un composé mono-Boc **43** : **a)**  $\text{K}_2\text{CO}_3$ , MeOH,  $60^\circ\text{C}$ , 4h ; **b)**  $\text{K}_2\text{CO}_3$ , acétone,  $56^\circ\text{C}$ , 16h.

L'étape d'alkylation (étape b) qui suivra, pourra être réalisée dans les mêmes conditions que précédemment et il sera ainsi intéressant d'étudier le résultat obtenu (*N*-alkylé ou *O*-alkylé).

Par ailleurs, la régiosélectivité de la réaction d'alkylation a été confirmée par RMN  $^{13}\text{C}$  avec le déplacement chimique du carbone  $\text{CH}_2\text{-N}$  dans la région de 40 ppm qui correspond à une *N*-alkylation tandis que pour un composé résultant d'une *O*-alkylation ce carbone sort avec un déplacement chimique aux alentours de 60 ppm. Une analyse supplémentaire par HMBC nous a confirmé les corrélations attendues pour les protons  $\text{CH}_2\text{-N}$  (H-11) avec les carbones C-9 et C-7 prouvant que les conditions réactionnelles utilisées favorisent une attaque nucléophile par l'atome de l'azote N-8 (figure 54).

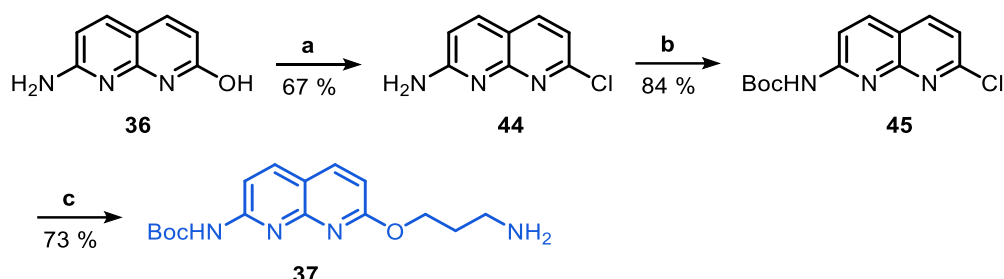


**Figure 54** : Spectre HMBC du composé *N*-alkylé **42** réalisé dans le chloroforme-*d*.

Au vu des résultats obtenus et face à cette régiosélectivité, nous avons ensuite décidé d'opter pour la **voie B** afin d'obtenir le produit *O*-alkylé souhaité.

## II.2.3.2.2 Synthèse de l'analogue naphtyridine par la voie B

La stratégie de la **voie B** (cf. *schéma 16*) repose principalement sur l'intermédiaire 2-amino-7-chloro-[1,8]-naphtyridine **44** sur lequel un bras hydroxyle est introduit via une substitution nucléophile aromatique (*schéma 19*).



*Schéma 19* : Synthèse de l'analogue naphtyridine O-alkylé **37**. Conditions : **a)** POCl<sub>3</sub>, toluène, reflux, 72h ; **b)** Boc<sub>2</sub>O, THF, t.a., 16h ; **c)** 3-amino-1-propanol, NaH, THF, 60°C, 16h.

La chloration du 2-amino-7-hydroxy-[1,8]-naphtyridine (**36**) a été réalisée en suivant les procédures déjà décrites dans la littérature<sup>183,184</sup>. La réaction implique l'utilisation du réactif POCl<sub>3</sub> comme agent de chloration et comme solvant. Ainsi, après 72h de réaction et après neutralisation par une solution ammoniacale (25 %), le produit **44** a été isolé avec un rendement de 67 % sans purification. L'atome de chlore situé en position 7 permet d'envisager par la suite une substitution nucléophile aromatique. Avant d'effectuer cette substitution nucléophile, l'amine exo-cyclique du composé **44** a été protégée par un groupement Boc. Lors de cette protection, seule la mono-protection a été observée (**45**, 80 %) sans utilisation de la DMAP. Cette différence de réactivité par rapport au composé **38** (cf. *schéma 17*) pourrait s'expliquer par la présence de l'atome de chlore qui diminue la réactivité de l'amine exocyclique plus fortement que l'atome d'oxygène. Sur le composé **44**, une S<sub>N</sub>Ar a été réalisée avec le 3-amino-1-propanol sans avoir recours à une protection de l'amine primaire comme décrit par Tarrago et *al.*<sup>185</sup>. En effet, l'utilisation de l'hydruide de sodium permet de former l'alcoolate qui lui est plus nucléophile que l'amine primaire. Dans ces conditions, le composé O-alkylé **37** a pu être directement obtenu avec un rendement de 73 % dans le THF.

Pour confirmer la O-alkylation, une analyse par HMBC a aussi été réalisée et nous avons observé la corrélation attendue entre les protons CH<sub>2</sub>-N (H-11) et le carbone C-7, démontrant bien la O-alkylation (*figure 55*).

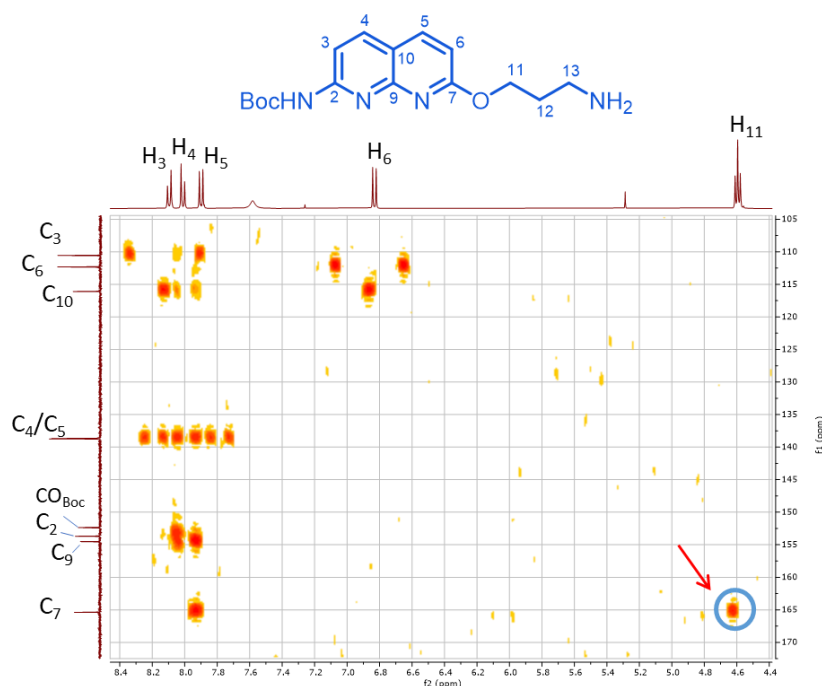


Figure 55 : Spectre HMBC du composé O-alkylé **37** réalisé dans le chloroforme-d.

Pour conclure, les deux voies de synthèse présentées nous ont permis de moduler la position du bras espaceur portant la fonction amine pour obtenir deux composés distincts : le composé N-alkylé **42** (4 étapes, 44 %) et le composé O-alkylé **37** (4 étapes, 41 %), qui pourraient avoir des affinités différentes comme fragments de ligands d'ARN.

### II.2.3.3 Préparation de la nucléobase S et de son dérivé

Le motif de la **nucléobase S** permet de réaliser un triplex avec les paires de bases A-U<sup>111</sup> (figure 56). Pour la synthèse de ce dérivé S<sub>A-U</sub>, nous avons ajouté via un couplage peptidique l'acide aminé glycine agissant comme bras espaceur portant la fonction amine souhaitée pour la réaction de Ugi selon la méthode décrite par Joly et al.<sup>113</sup> (figure 56).

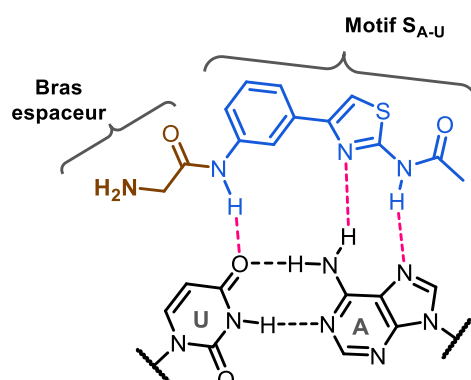
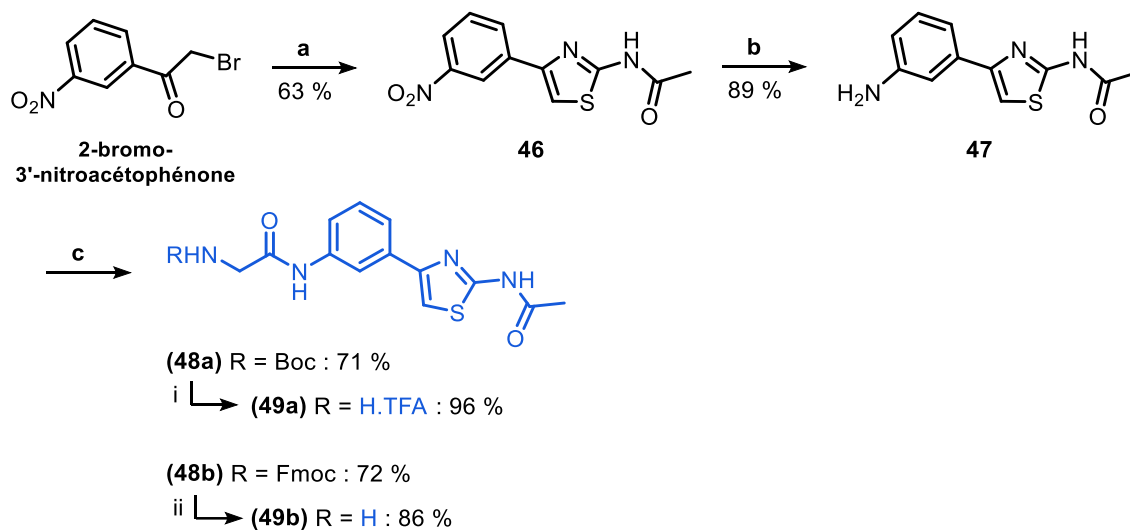


Figure 56 : Présentation de la structure du dérivé de la nucléobase S.

Les deux premières étapes de la synthèse permettent la préparation du motif S<sub>A-U</sub> (**46**) (schéma 20). Une cyclo-condensation de type Hantzsch a été réalisée entre le composé 2-bromo-3'-

nitroacétophénone et la *N*-acétylthiourée permettant de former le cycle aminothiazole **46**. Une simple filtration de la réaction a conduit au composé **46** avec un rendement de 63 %.



**Schéma 20** : Synthèse des dérivés de la nucléobase artificielle « S » **49a** et **49b**. Conditions : **a**) *N*-acétylthiourée, EtOH, 80°C, 1h ; **b**) Pd/C, NaBH<sub>4</sub>, CH<sub>2</sub>Cl<sub>2</sub>/MeOH, t.a., 16h ; **c**) RHN-Gly-COOH, EDC, HOSu, CH<sub>2</sub>Cl<sub>2</sub>/DMF, t.a., 16h ; **i**) TFA, CH<sub>2</sub>Cl<sub>2</sub>, t.a., 4h ou **ii**) pipéridine/DMF, t.a., 4h.

Ensuite, une étape de réduction du groupement nitro en groupement amine a été effectuée à l'aide du palladium sur charbon et du borohydrure de sodium dans un mélange CH<sub>2</sub>Cl<sub>2</sub>/MeOH. Le motif S<sub>A-U</sub> **47** a été alors obtenu après purification avec un rendement de 89 %.

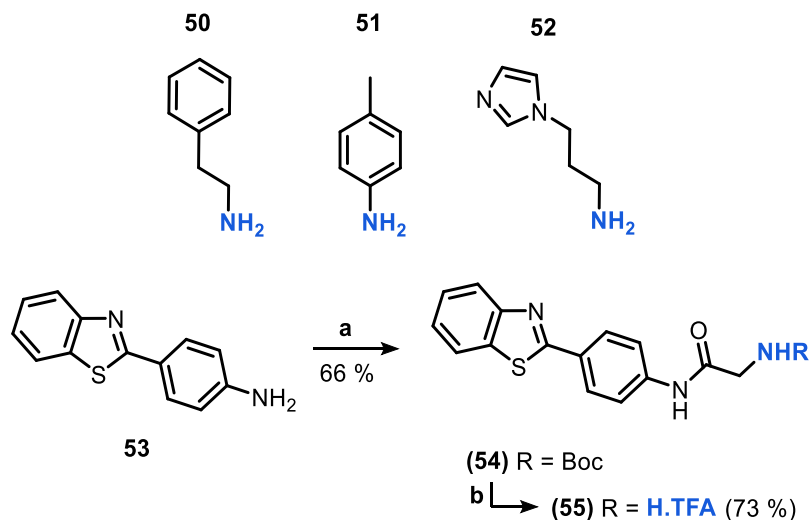
L'amine exo-cyclique du composé **47** a ensuite été engagée dans un couplage peptidique avec la glycine protégée soit avec un groupement Boc soit avec un groupement Fmoc. L'activation de l'acide carboxylique en ester activé a été réalisée avec les agents de couplages 1-éthyl-3-(3-diméthylaminopropyl)carbodiimide (EDC) et *N*-hydroxysuccinimide (HOSu). Le composé **48a** a été obtenu après purification avec un rendement de 71 %, alors que le composé **48b** a été isolé par simple filtration pour un rendement 72 %. La dernière étape est la suppression des groupements protecteurs. Le composé **48a** a été traité par l'acide trifluoroacétique dans le CH<sub>2</sub>Cl<sub>2</sub> donnant le composé **49a** sous forme protonée et avec un rendement de 96 %. Le composé **48b** a été déprotégé en présence de pipéridine dans le DMF ce qui a permis d'obtenir le composé **49b** sous forme neutre avec un rendement de 86 %.

Par cette stratégie de synthèse, les composés **47**, **49a** et **49b** comportant une fonction amine ont été obtenus et peuvent être utilisés dans la réaction de Ugi.

### II.2.3.4 Sélection d'autres motifs comportant la fonction amine

Pour pouvoir enrichir la nouvelle collection de ligands d'ARN, nous avons sélectionné d'autres motifs hétéroaromatiques commerciaux (**50**, **51** et **52**) permettant un accès rapide à cette bibliothèque

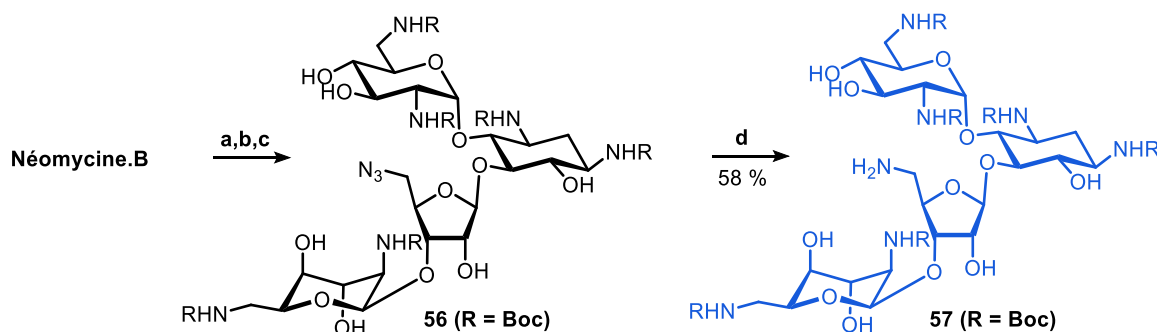
(schéma 21). Ces différents motifs ont été choisis pour des potentielles interactions d'empilement et/ou des liaisons hydrogènes envers l'ARN, ainsi que leur caractère hydrophobe.



**Schéma 21 :** A) Structures des fragments commerciaux sélectionnés. B) Synthèse du composé 55. Conditions : a) BocHN-Gly-COOH, HBTU, Et<sub>3</sub>N, CH<sub>2</sub>Cl<sub>2</sub>, t.a., 16h ; b) TFA, CH<sub>2</sub>Cl<sub>2</sub>, t.a., 16h.

Pour le composé 55, à partir de la même stratégie que l'analogue S<sub>A-U</sub>, il a été introduit un bras espaceur comportant la fonction amine à partir du composé commercial 53. Ainsi, après un couplage peptidique entre le 2-(4-aminophényl)benzothiazole (53) et la glycine N-protégée par un groupement Boc, l'intermédiaire 54 a été obtenu avec un rendement de 66 %. Ensuite, le groupement Boc a été éliminé pour obtenir le composé 55 sous forme protonée avec un rendement de 73 %.

Enfin, nous avons choisi d'utiliser la néomycine comportant la fonction amine sur la position 5'' du cycle III (schéma 22). La néomycine pourra apporter des interactions électrostatiques intermoléculaires entre les groupes chargés ammoniums et les groupements phosphates qui constituent le squelette de l'ARN. De plus, les motifs aminosucre dont elle est constituée peuvent interagir de manière plus sélective via des liaisons hydrogènes avec les bases non appariées de la cible.



**Schéma 22 :** Synthèse de la néomycine modifiée 57. Conditions : a) Boc<sub>2</sub>O, Et<sub>2</sub>N, MeOH/H<sub>2</sub>O, t.a., 16h, 66 % ; b) chlorure de 2,4,6-triisopropylbenzènesulfonyle, py., t.a., 16h, 60 % ; c) NaN<sub>3</sub>, DMF, 100°C, 16h, 60 % ; d) H<sub>2</sub>, Pd/C, EtOH, t.a., 16h.

Suivant le protocole décrit par Wang et *al.*<sup>186</sup>, le composé **57** a pu être synthétisé à partir de la néomycine B trisulfate dont les amines ont été protégées par le groupement Boc. Ensuite, une activation de l'hydroxyle primaire en un groupement partant a permis d'introduire une fonction azoture *via* une substitution nucléophile donnant le composé **56**. Enfin, une hydrogénation catalytique conduit au composé **57** avec un rendement de 58 %. Ainsi, le dérivé néomycine modifié (**57**) a été obtenu en 4 étapes pour un rendement global de 14 %.

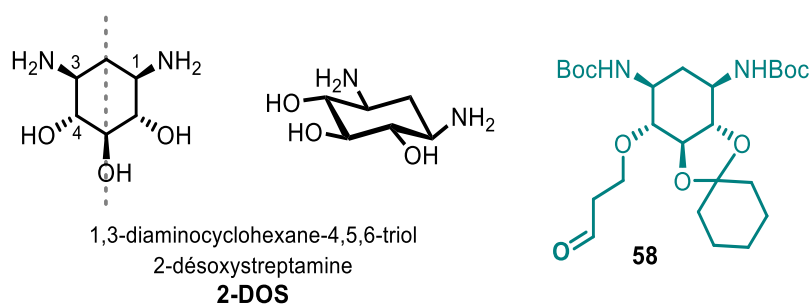
## II.2.4 Préparation des motifs ayant la fonction aldéhyde

Les synthèses des analogues de la 2-déoxystreptamine et de la glucosamine sont détaillées ci-dessous. Ces analogues comportent un bras espaceur avec une fonction aldéhyde permettant la réaction de Ugi future.

### II.2.4.1 Préparation de l'aldéhyde 2-déoxystreptamine

#### II.2.4.1.1 Préparation du motif 2-déoxystreptamine

Caractéristique d'une partie des aminoglycosides, le motif 2-déoxystreptamine (2-DOS) est un composé « méso » en raison de la présence de cinq stéréocentres adjacents les uns aux autres formant un plan de symétrie (*figure 57*).

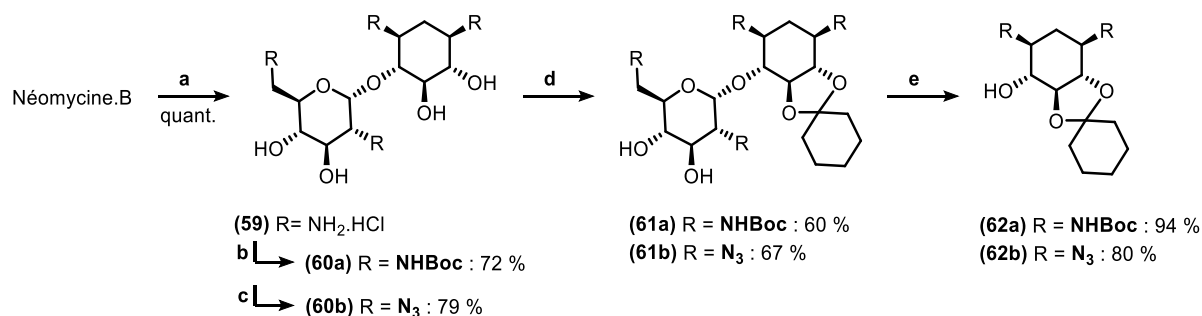


**Figure 57** : Structure chimique du motif 2-déoxystreptamine (2-DOS) et de l'analogue 2-DOS modifiée **58**.

L'orientation des groupements périphériques de ce motif est essentielle pour l'établissement des interactions avec l'ARN notamment l'ARN du site A bactérien<sup>187</sup> ou l'ARN DIS<sup>121</sup>. Pour cela, nous avons décidé d'insérer sur la position 4 un bras espaceur contenant la fonction aldéhyde sur un squelette protégé pour obtenir le composé **58**. Dans cette stratégie, nous avons fait le choix d'utiliser des groupements protecteurs acido-labiles. Ces derniers peuvent ainsi être éliminés en même temps que ceux des autres fragments à la suite de la réaction de Ugi. Ainsi, dans un premier temps, il a fallu synthétiser le squelette 2-DOS permettant la synthèse du composé **58**.

L'accès au squelette 2-déoxystreptamine a été largement étudié dans la littérature. La 2-DOS peut en effet être obtenue à partir de motifs variés ou directement par une acidolyse de la néomycine B<sup>188</sup> En

2009, Klemm et *al.*<sup>189</sup> ont décrit la synthèse du motif 2-DOS protégé **62** à partir de l'hydrolyse de la néomycine suivie de différentes étapes de protection permettant une modification unique sur la position 4 libre (*schéma 23*). Dans cette procédure, les fonctions amines sont protégées soit par un groupement Boc (**62a**), soit en une fonction azoture (**62b**), et dans les deux cas, les fonctions hydroxyles sont protégées par un groupement acétal cyclique.



**Schéma 23** : Synthèses des composés intermédiaires **62a** et **62b**. Conditions : **a)** HCl conc., MeOH, 65°C, 16h ; **b)** Boc<sub>2</sub>O, Et<sub>3</sub>N, dioxane/H<sub>2</sub>O, 60°C, 16h ; **c)** TfN<sub>3</sub>, CuSO<sub>4</sub>/Et<sub>3</sub>N, CH<sub>2</sub>Cl<sub>2</sub>/H<sub>2</sub>O/MeOH (1:2:3 v/v/v), t.a., 16h ; **d)** 2,2-diméthoxycyclohexane, *p*-TsOH, 50 mbar, 50°C, 2h30 ; **e)** (i) NaIO<sub>4</sub>, MeOH, t.a., 16h, (ii) *n*-butylamine, MeOH, t.a., 16h.

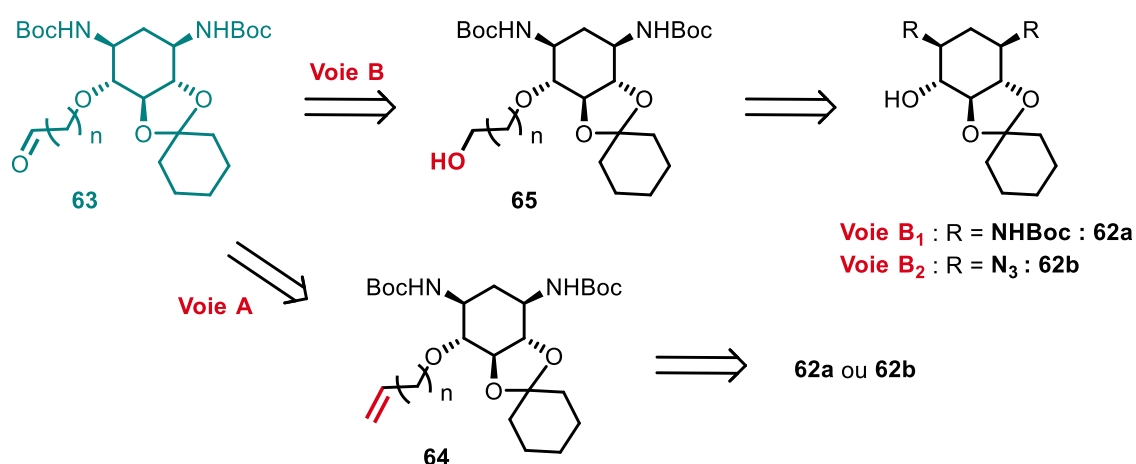
Durant cette thèse, les deux composés **62a** et **62b** ont été synthétisés. La première étape de cette synthèse repose sur l'hydrolyse de la néomycine B trisulfate en milieu acide à reflux donnant le composé néamine **59** par une filtration à froid de manière quantitative. Ensuite, les fonctions amines de la néamine **59** ont été protégées donnant accès aux composés comportant soit des groupements Boc **60a** avec un rendement de 72 %, soit des fonctions azoture **60b** avec un rendement de 79 %. Cette dernière protection a été réalisée par un « diazotransfert » avec de l'azoture de trifluorométhanesulfonyle (TfN<sub>3</sub>) préparé *in situ*<sup>190</sup> en présence d'une quantité catalytique de sulfate de cuivre (II).

En gardant à l'esprit que la structure 2-déoxystreptamine est un composé méso, il est indispensable d'orienter la synthèse permettant l'accès à un seul stéréoisomère notamment pour l'accès à la position 4 du cycle de la 2-DOS. Afin de contrôler l'accès à un seul stéréoisomère, une protection sélective des hydroxyles en position 5- et 6- a été effectuée par le 1,1-diméthoxycyclohexane en présence d'acide *p*-toluènesulfonique. Cette réaction a été menée sous pression réduite (55 mbar) à 50°C permettant d'obtenir un mélange de produits mono-acétal ou di-acétal. Le produit di-protégé a facilement été converti en un composé mono-acétal **61** par ajout de méthanol<sup>191</sup>. Le rendement global de cette réaction est de 60 % et 67 % pour les composés respectifs **61a** et **61b**. Enfin, l'accès au motif 2-DOS s'est fait en deux étapes. En effet, un traitement des composés **61a** et **61b** par du périodate de sodium (NaIO<sub>4</sub>) a permis une coupure oxydative des 1,2-diols<sup>192</sup>. Le dialdéhyde qui en résulte a ensuite été mis en présence de *n*-butylamine afin de procéder à une β-élimination. Les composés **62a** et **62b** ont ainsi été obtenus avec des rendements respectifs de 94 % et 80 %.

En conclusion, la préparation des composés énantio-purs **62a** (protection Boc) et **62b** (protection par un azoture) comportant un hydroxyle libre en position 4 a été effectuée en 4 étapes avec des rendements globaux de 40 % et 42 % respectivement.

### II.2.4.1.2 Synthèse de l'aldéhyde 2-désoxystreptamine

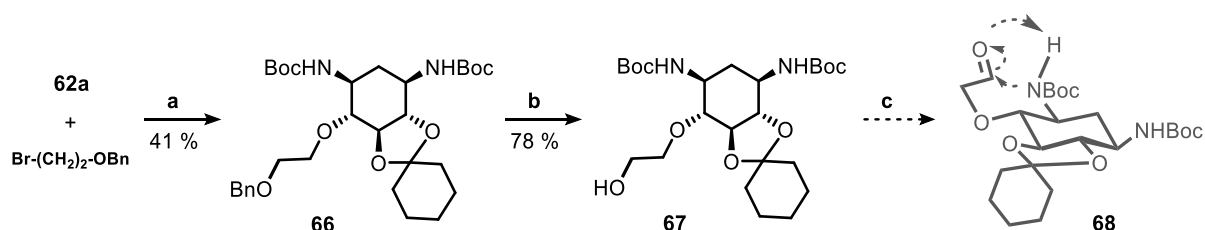
Pour obtenir l'aldéhyde **63**, différentes voies sont accessibles et sont illustrées dans le schéma **24**. Dans la **voie A**, l'obtention de l'aldéhyde peut être envisagée au travers d'une ozonolyse<sup>193</sup> de l'intermédiaire **64**, alors que dans le cas de la **voie B**, la stratégie repose sur une oxydation de l'alcool **65**. Nous avons opté pour la **voie B** qui ne nécessite pas d'appareillage spécifique d'ozonolyse.



*Schéma 24 : Analyse rétro-synthétique pour accéder au composé **65** comportant la fonction aldéhyde.*

L'incorporation des chaînes en position 4 fait appel à la réaction de Williamson entre les composés **62a/62b** et un dérivé halogéné. Ici, notre choix s'est porté sur l'utilisation d'un alcool protégé par un groupement benzyle pouvant être supprimé à la suite d'une hydrogénation catalytique.

Dans une étude préliminaire réalisée en partie par un stagiaire de master 2, O. BORODIN, nous avons décidé dans un premier temps, d'alkyler l'intermédiaire **62a** avec le réactif Br-(CH<sub>2</sub>)<sub>2</sub>-OBn (schéma **25**).



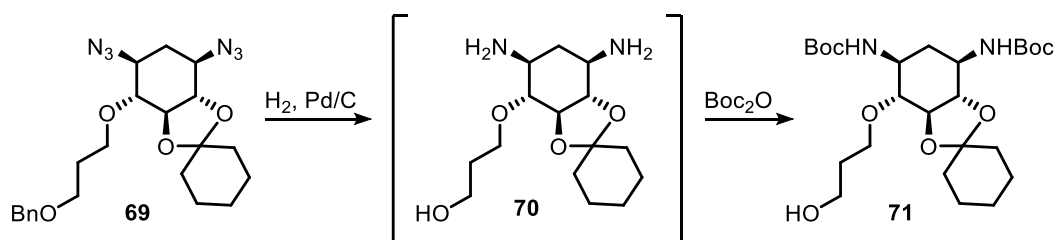
*Schéma 25 : Synthèse de l'analogue 2-DOS modifiée **68**. Conditions : a) benzyl 3-bromopropyl éther, NaH, TBAI, toluène, t.a., 16h ; b) H<sub>2</sub>, Pd/C, EtOH, t.a., 16h ; c) IBX, DMSO, t.a.*

La première étape est la substitution nucléophile d'ordre 2 entre le réactif bromé et l'alcoolate du composé **62a** réalisée en présence de l'hydrure de sodium (NaH) dans le toluène<sup>189,139</sup>. En utilisant ces conditions, nous avons observé une dégradation du produit, ainsi qu'une élimination d'un des groupement Boc des amines réduisant les rendements obtenus du composé **66** qui n'ont pas dépassé



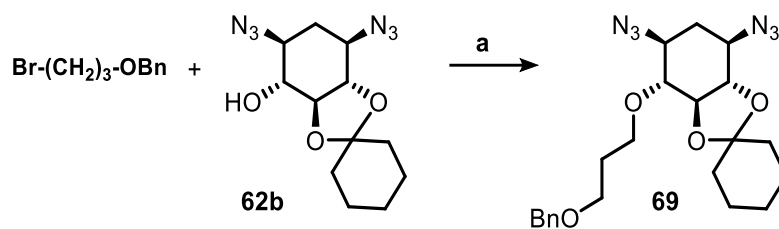
les 41 %. Une fois obtenu, le composé **66** a été engagé dans une étape d'hydrogénation pour fournir l'hydroxyle libre pour un rendement de 78 %. Enfin, une étape d'oxydation pour former l'aldéhyde **67** a été réalisée par un agent oxydant doux qui est l'acide 2-iodoxybenzoïque (IBX). Lors de cette étape, un sous-produit récurrent est apparu, provenant probablement d'une attaque de type 6-exo-trig entre l'amine et le carbonyle de l'aldéhyde conduisant à des réactions secondaires et ne permettant pas d'obtenir convenablement l'aldéhyde souhaité **68**.

Ainsi, face à ces deux problèmes lors des étapes d'alkylation et d'hydrogénation catalytique, une deuxième stratégie a été employée pour accéder à l'aldéhyde souhaité. Premièrement, afin d'éviter la possible cyclisation avec l'amine du squelette 2-DOS et le bras alkylé, un carbone supplémentaire a été introduit pour éviter la formation d'un cycle à 6. Puis, afin de se prémunir contre les réactions parasites dues à la réactivité des amines protégées par un groupement Boc, nous avons introduit une fonction azoture à la place des fonctions amines. Cette stratégie reste assez intéressante car la diamine désirée (**71**) peut être obtenue lors de l'élimination du benzyle (**69**) par hydrogénation catalytique donnant l'intermédiaire **70** (schéma 26). Puis, la diamine **70** peut être également protégée par un groupement Boc par la suite et fournir le composé **71**.



*Schéma 26 : Représentation de l'étape clé de la stratégie de la Voie B<sub>2</sub> pour passer des protections azotures en groupement Boc.*

Ainsi, dans la nouvelle stratégie établie, le composé **69** a d'abord été synthétisé. Au travers de la réaction de Williamson sur le composé **62b**, une optimisation des conditions réactionnelles a été nécessaire. Les différentes conditions testées sont détaillées dans le *tableau 7* ci-dessous.

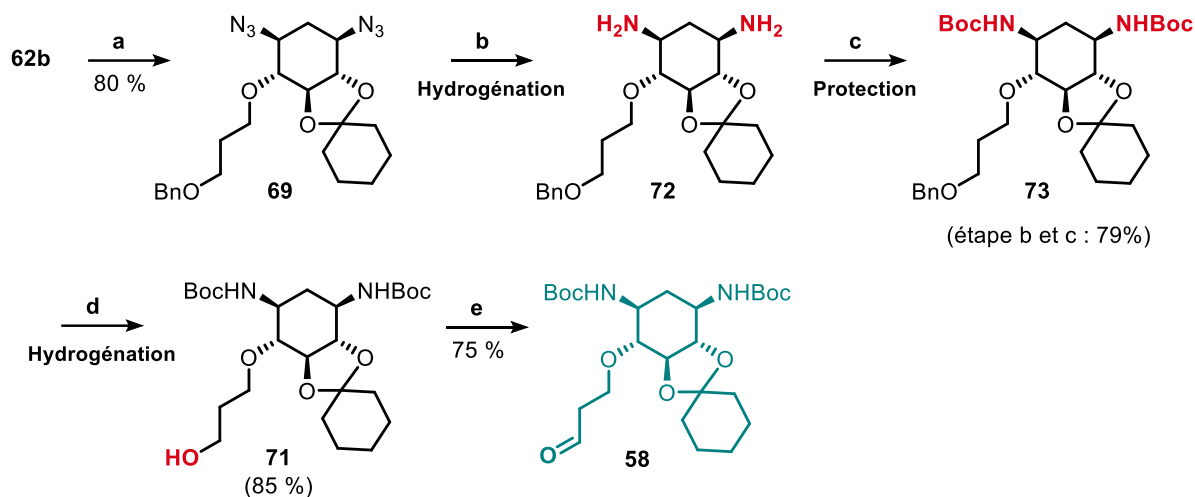
Tableau 7 : Optimisation des conditions de la réaction d'alkylation pour obtenir le composé **69**.

Entrée	TBAI (Nb d'éq.)	NaH (Nb d'éq.)	R-Br (Nb d'éq.)	Solvant	<b>62b</b> <sup>a</sup>	<b>69</b>
<b>1</b>	0.2	2.5	2	THF	30 %	45 %
<b>2</b>	0.2	4	2	THF	nd	40 %
<b>3</b>	1	4	1.5	THF	40 %	21 %
<b>4</b>	1	4	1.8	DMF	9 %	65 %
<b>5</b>	1.2	4	1.2	DMF	38 %	45 %
<b>6</b>	1.2	4	3	DMF	nd	<b>80 %</b>

Conditions : **a**) t.a., 16h ; **nd** : non déterminé. <sup>a</sup> produit obtenu après purification.

Dans un premier temps (**entrée 1**), les conditions décrites par Klemm et *al.*<sup>189</sup> ont été mises en œuvre. Ces conditions reposent sur l'utilisation de 2.5 équivalents de NaH et 0.2 équivalents de TBAI dans le THF. Un rendement de 45 % a été obtenu avec une faible dégradation du produit de départ qui a pu être récupéré à la suite de la purification (30 %). L'augmentation du nombre d'équivalents du réactif TBAI ou de NaH (**entrées 2 et 3**) n'a pas permis d'améliorer le rendement de manière significative. En revanche, le remplacement du THF par du DMF (**entrée 4**) a permis d'obtenir un rendement de 65 %. Enfin, l'utilisation d'un excès du réactif halogéné (**entrée 6**) entraîne une augmentation notable du rendement (80 %). Ce rendement a été doublé en comparaison avec celui obtenu lors de l'autre stratégie de synthèse (cf. *schéma 25*) utilisant le composé de départ **62a**.

Une fois obtenu, le composé **69** a été engagé dans une étape d'hydrogénation pour fournir de façon inattendue le produit **72** dans lequel seules les fonctions azotures ont été réduites (*schéma 27*).



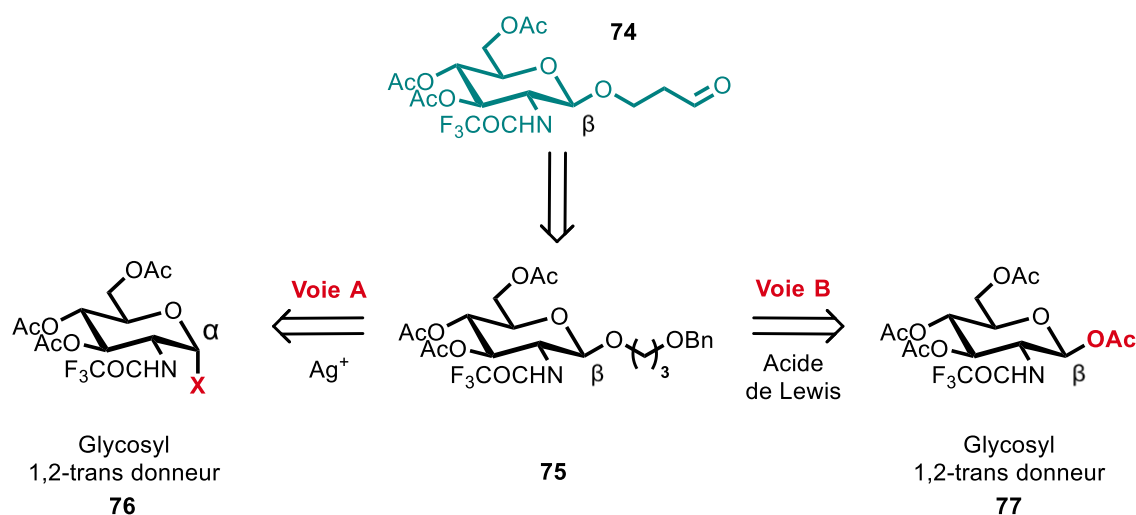
**Schéma 27** : Synthèse de l'analogue 2-DOS modifiée **58**. Conditions : **a**) benzyl 3-bromopropyl éther, NaH, TBAI, DMF, t.a., 16h ; **b**) H<sub>2</sub>, Pd/C, EtOH, t.a., 16h, **c**) Boc<sub>2</sub>O, Et<sub>3</sub>N, dioxane, t.a., 16h ; **d**) H<sub>2</sub>, Pd/C, EtOH, t.a., 16h ; **e**) IBX, DMSO, t.a., 6h.

Le composé **72**, a été directement engagé dans l'étape de protection des amines par les groupements Boc (**étape c**). Le deuxième intermédiaire **73** a été alors obtenu avec un rendement global de 79 % après purification pour les étapes b et c. Afin de réduire le groupement *O*-benzyle, une deuxième hydrogénation (**étape d**) a été réessayée sur l'intermédiaire **73**, donnant le composé souhaité **71** avec un rendement de 85 %. Cet enchaînement des étapes b, c et d a pu également être réalisé sans effectuer l'étape de purification de l'intermédiaire **73** en réalisant un simple traitement aqueux suivi d'une étape d'hydrogénation. Ainsi, le composé **71** a été obtenu avec un rendement de 65 % sur 3 étapes.

Enfin, pour obtenir le composé **58**, une étape d'oxydation sur le composé **71** a été réalisée par un agent oxydant doux, l'acide 2-iodoxybenzoïque (IBX) avec un rendement de 75 %. Ainsi, l'aldéhyde 2-DOS (**58**) a pu être obtenu en 8 étapes avec un rendement global de 14 %.

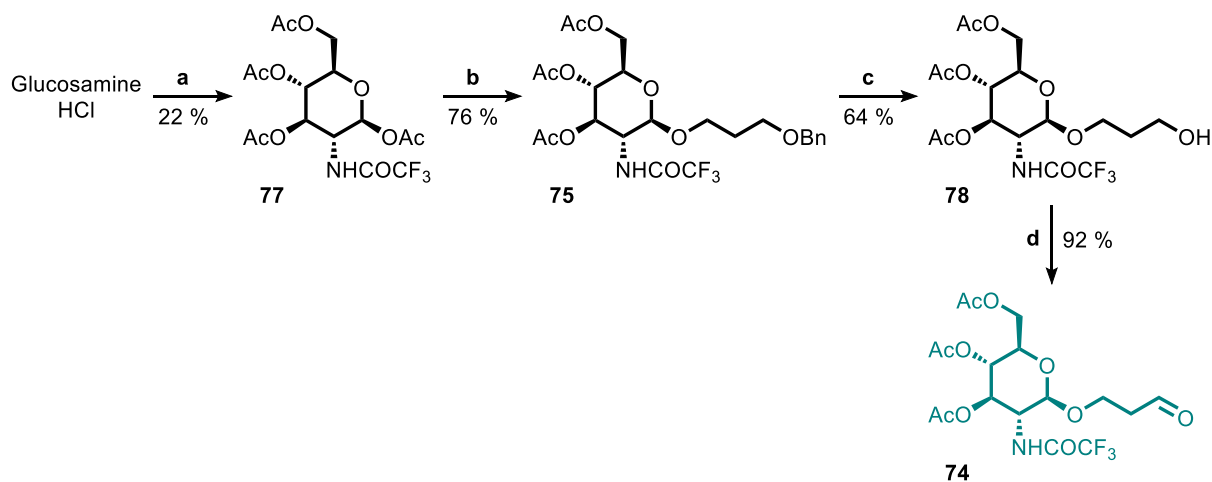
#### II.2.4.2 Préparation de l'aldéhyde glucosamine

Le motif glucosamine représente le cycle IV de la néomycine et interagit avec l'ARN via des interactions électrostatiques. Cet aminoglycoside a déjà fait l'objet d'études où différentes modifications sur le lien glycosidique ont été réalisées pour être évaluées dans le contexte du ciblage d'ARN<sup>194</sup>. Westermann et *al.* ont réalisé une réaction de Ugi avec quatre monomères de glucosamines modifiés<sup>195</sup>. Sur cette base, nous avons décidé de synthétiser l'anomère β (**74**) (*schéma 28*). Des groupements acétyles et trifluoroacétamide ont été utilisés pour protéger les différentes fonctions de la glucosamine, ces protections pouvant être éliminées par un traitement basique après la réaction de Ugi. Par analogie avec le composé 2-DOS modifié (**58**), nous avons opté pour une stratégie similaire pour introduire la fonction aldéhyde avec une longueur de bras espaceur identique (**75**).



**Schéma 28** : Analyse rétro-synthétique pour obtenir la glucosamine modifiée **74**. X = Halogène.

Le composé **75** implique une étape de glycosylation<sup>196</sup>. La **voie A** fait appel à une réaction de Koenings-Knorr (S<sub>N</sub>2) impliquant un composé halogénure **76** et un sel d'argent comme promoteur. Dans le cas de la **voie B**, le composé **75** est exclusivement issu de l'anomère β du composé **77** en présence d'acide de Lewis. Notre choix de synthèse s'est porté sur la voie B car l'obtention du réactif de départ **77** en deux étapes successives, est plus rapide que celle du composé **76** (schéma 29).



**Schéma 29** : Synthèse de l'analogue glucosamine modifiée **74**. Conditions : **a**) i) trifluoroacétate d'éthyle, Et<sub>3</sub>N, MeOH, t.a., 16h, ii) Ac<sub>2</sub>O, py., t.a., 16h ; **b**) 3-benzyloxy-1-propanol, BF<sub>3</sub>.OEt<sub>2</sub>, CH<sub>2</sub>Cl<sub>2</sub>, t.a., 16h ; **c**) H<sub>2</sub>, Pd/C, EtOH, t.a., 16h ; **d**) IBX, DMSO, t.a., 6h.

Tout d'abord, la protection des fonctions de la glucosamine commerciale a été réalisée en deux étapes. Le trifluoroacétate d'éthyle puis l'anhydride acétique (Ac<sub>2</sub>O) ont respectivement permis la protection de la fonction amine et des fonctions hydroxyles, conduisant à un mélange α/β. Après séparation, l'anomère β **77** a été isolé minoritairement avec un rendement de 22 %. D'autres stratégies permettent d'améliorer la sélectivité de l'anomère β mais nécessitent des étapes supplémentaires<sup>197</sup>. Ensuite, sur l'anomère β **77**, une réaction de glycosylation avec le 3-benzyloxy-1-propanol a été

réalisée en présence de l'acide de Lewis diéthyléthérate de trifluorure de bore ( $\text{BF}_3 \cdot \text{OEt}_2$ )<sup>198</sup>. Cette étape implique la formation d'un ion oxazolidinium par la suite d'une assistance anchimérique de l'acétamide. Cet ion intermédiaire va favoriser une 1,2-trans glycosylation et donc l'anomère  $\beta$  **75** a été obtenu avec un rendement de 76 %. Ensuite, le groupement *O*-Benzyle a été éliminé par une hydrogénation catalytique donnant le composé **78** avec un rendement de 64 %. Enfin, un traitement de l'alcool **78** par de l'IBX a permis de le transformer en aldéhyde **74** avec un rendement de 92 %. Ainsi, le synthon glucosamine comportant la fonction aldéhyde (**74**) a été obtenu avec un rendement global de 11 % sur 4 étapes.

### II.2.5 Récapitulatif des composants synthétisés

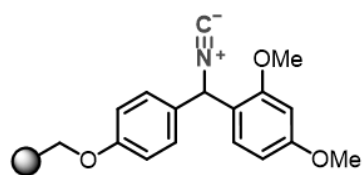
La réaction de Ugi fait réagir quatre fonctions distinctes qui sont l'aldéhyde, l'amine, l'acide carboxylique et l'isocyanure. Dans cette partie, nous avons décrit les synthèses des différentes fragments choisies préalablement et résumées dans la figure ci-après (*figure 58*).

Dans un premier temps, nous avons réalisé la fonctionnalisation de la résine pour obtenir la fonction isocyanure. Des caractérisations qualitatives par IR et quantitatives par dosage ou par réactivité, ont pu être réalisées pour différents lots de résine fonctionnalisée. Les taux de fonctionnalisation ont été déterminés et se situent entre 22 % à 54 % pour la résine Rink amide MBHA et de 43 % pour la résine Rink amide. Grâce à ce test quantitatif, nous avons pu aussi mettre en place les conditions de la réaction de Ugi sur support solide et de l'étape de clivage du support qui seront appliquées dans la partie suivante.

Parallèlement, l'ensemble des synthons comportant une fonction aldéhyde (2 motifs) et une fonction amine (10 motifs) par des synthèses variées et multi-étapes a été obtenu avec succès pour des synthèses allant de 2 étapes à 8 étapes. Lors de ces différentes synthèses, j'ai pu encadrer un stagiaire durant son master 1 et 2 qui a préparé plusieurs synthons notamment sur grandes quantités dont les synthons de l'amine naphtyridine *N*-alkylé (**42**) et de l'aldéhyde 2-DOS (**58**).

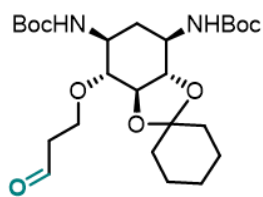
À partir de ces différents fragments, nous avons réalisé différentes séries *via* la réaction de Ugi. Dans la partie suivante, nous détaillerons les trois séries de composés (cf. *figure 48*, page **91**) dont les deux premières (**séries A et B**) ont été obtenues grâce à une méthodologie de synthèse sur support solide fonctionnalisé comme décrit précédemment, et ensuite la **série C** qui a été réalisée grâce à la réaction de Ugi effectuée en solution.

## Support solide

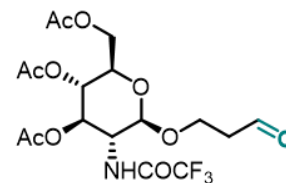


3 étapes  
Rink Amide ou MBHA

## Aldéhydes



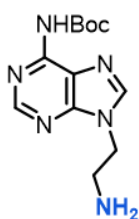
58  
(8 étapes - 14 %)



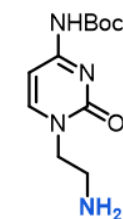
74  
(4 étapes - 11 %)

## Amines

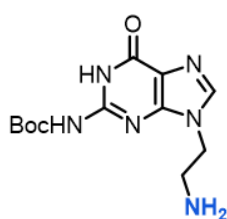
21  
4 étapes - 54 %



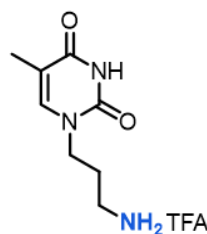
25  
4 étapes - 21 %



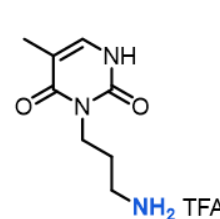
32  
4 étapes - 54 %



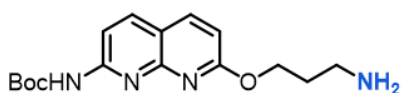
34  
2 étapes - 25 %



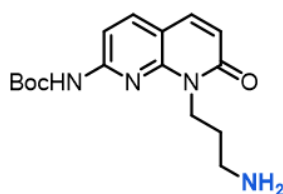
35  
2 étapes - 11 %



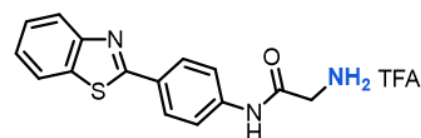
37  
4 étapes - 41 %



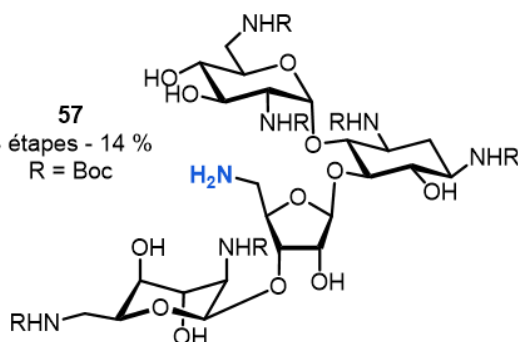
42  
4 étapes - 44 %



55  
2 étapes - 48 %



57  
4 étapes - 14 %  
R = Boc



R = H : 45  
2 étapes - 56 %  
R = Gly-NH<sub>2</sub> : 49a/b  
4 étapes - 35 %

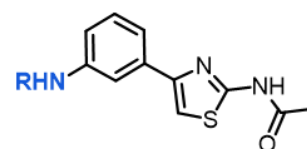
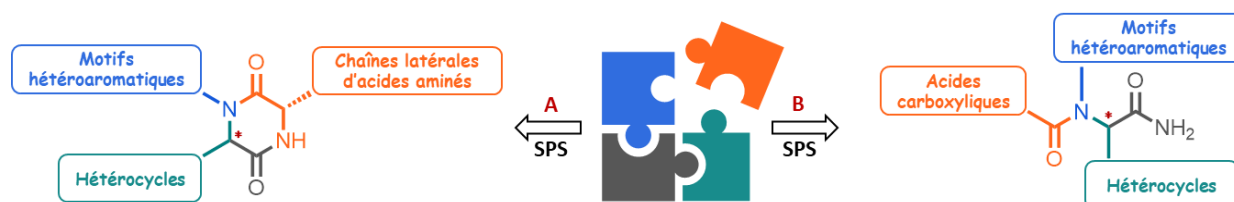


Figure 58 : Représentation des composants isocyanures, aldéhydes et amines synthétisées.

## II.3 Préparation des ligands d'ARN sur support solide : Séries A et B

Pour explorer l'inhibition des complexes ARN-ARN, nous avons dans un premier temps réalisé la préparation d'une première série de ligands avec l'utilisation d'un support solide polymérique. Dans le but d'effectuer les synthèses par une réaction de Ugi, nous avons précédemment fonctionnalisé le support solide avec la fonction isocyanure (cf. *Chap. II.2.2*). A partir de ce support solide, nous avons pu explorer deux séries de structures différentes qui sont les séries A et B décrites dans cette partie (*figure 59*).



**Figure 59 :** Représentation des deux séries obtenues sur synthèse sur support solide. Série A : Structure 2,5-dicétopipérazine. Série B : Structure linéaire =  $\alpha$ -aminoacylamide.

### II.3.1 Synthèses des molécules obtenues dans la série A

#### II.3.1.1 Etude préliminaire

Afin de quantifier la fonctionnalisation de la résine, nous avons utilisé précédemment (cf. *Chap. II.2.2*) un test modèle qui est une réaction de Ugi avec des fragments simples pour obtenir un composé linéaire (**15**) de type amidique<sup>150</sup>. Ce composé a été obtenu avec une pureté supérieure à 95% et un rendement global sur 5 étapes entre 24 % et 54 % par rapport au taux de fonctionnalisation de la résine Rink amide MBHA commercial (*tableau 8, entrée 1*). Ces résultats nous ont permis dans un premier temps de vérifier si les conditions choisies précédemment sont optimales pour d'autres motifs proches de ce composé témoin. Pour réaliser cette étude préliminaire, nous avons fait varier les différents fragments à partir d'un même lot de résine fonctionnalisée (support  $S_1$ ) (*tableau 8, entrées 1, 2 et 3*).

Tableau 8 : Résultats des réactions dérivées du test témoin pour obtenir des composés Ugi-linéaires.

Support **S**<sub>1</sub> = Rink amide MBHA, **S**<sub>2</sub> = Rink amide

Entrée	Composé	AA : R <sub>1</sub>	Amine : R <sub>2</sub>	Aldéhyde : R <sub>3</sub>	Support	Rendement <sup>a</sup>	Pureté <sup>b</sup>
1	15	Gly			S <sub>1</sub>	24 %	100 %
2	79	Gly			S <sub>1</sub>	18 %	95 %
3	80	Lys			S <sub>1</sub>	13 %	85 %
4	81	Lys			S <sub>2</sub> <sup>c</sup>	13 %	95 %
5	82	Phe			S <sub>2</sub> <sup>c</sup>	37 %	91 %

**Conditions :** **1)** Résine-isocyanure (1 éq.), aldéhyde (2 éq.), amine (2éq) et acide aminé (AA) (2 éq), CH<sub>2</sub>Cl<sub>2</sub>/MeOH, t.a, 6 jours, **2)** pour la résine S<sub>1</sub> : 95% TFA dans le CH<sub>2</sub>Cl<sub>2</sub>, pour la résine S<sub>2</sub> :15 % TFA dans le CH<sub>2</sub>Cl<sub>2</sub>, t.a., 1h30. <sup>a</sup> Rendement calculé par rapport au « loading » de la résine de départ. <sup>b</sup> Pureté mesurée par HPLC à UV = 214 nm. <sup>c</sup> Rendement maximal déterminé à 43% sur les cinq étapes.

En comparaison avec la réaction témoin, l'aldéhyde avec un motif indole (**entrée 2**) a donné le composé **79**. Ce produit a été obtenu avec un résultat comparable au composé **15** avec un rendement de 18 % (pureté 95 %). En revanche, lorsque la lysine a été introduite à la place de la glycine (**entrée 3**), le rendement du composé **80** (13 %) a diminué. Cela pourrait être dû à une plus faible réactivité de cet acide dans la réaction de Ugi. Néanmoins, une pureté de 85 % a été obtenue. Ces mélanges (**15**, **79** et **80**) ont été caractérisés sans purification par RMN <sup>1</sup>H et par HRMS.

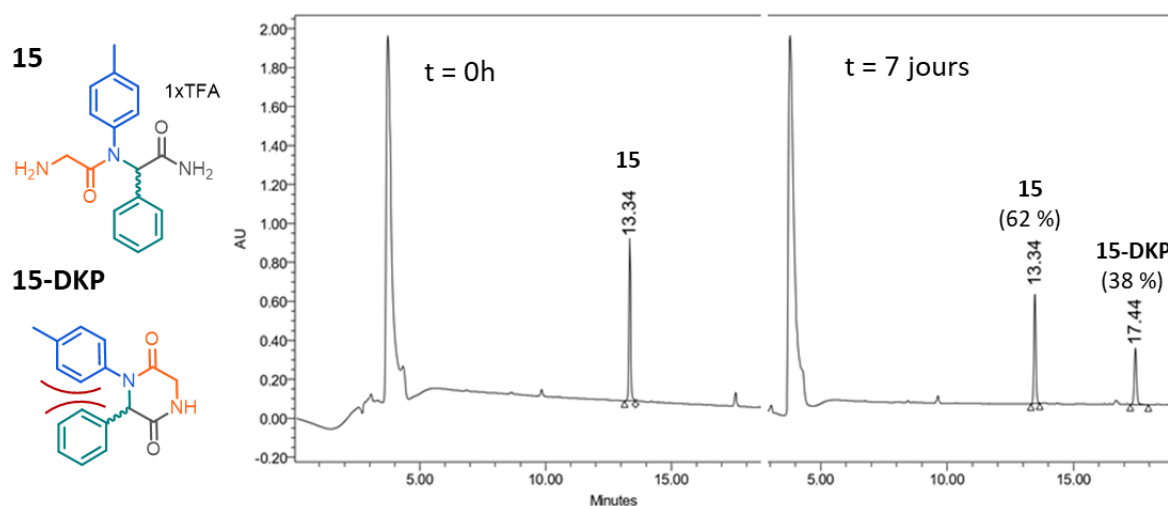
En parallèle, nous avons comparé les rendements de la réaction de Ugi avec un autre support solide fonctionnalisé préalablement, qui est la résine Rink amide (S<sub>2</sub>) dont le taux de fonctionnalisation est de 43 %. Ce support nécessite des conditions plus douces pour l'étape de clivage de la résine (15 % de TFA dans du CH<sub>2</sub>Cl<sub>2</sub>). Avec cette résine-isocyanure, il a été préparé les composés **81** (**entrée 4**) et **82** (**entrée 5**). Comme observé précédemment, l'utilisation de la lysine (**81**) a conduit à un faible rendement (13 % et une pureté de 95 %) par comparaison avec le composé phénylalanine **82** (37 % de rendement et une pureté de 91 %). Ces deux derniers composés (**81** et **82**) ont été caractérisés sans purification par RMN <sup>1</sup>H et confirmés par HRMS.

Ainsi, au travers de ces premiers résultats, nous avons remarqué qu'il existe une différence notable de réactivité de certains acides aminés comme la lysine. Le changement de support solide quant à lui n'impacte pas réellement le rendement ou la pureté des produits obtenus.



Les produits décrits dans le *tableau 8* ont tous été obtenus sous forme linéaire. Cependant dans la littérature, de nombreux exemples ont été reportés dans lesquels des acides aminés sont employés dans une réaction de Ugi et engendrent une cyclisation après déprotection de l'amine alpha conduisant à des réactions post-Ugi décrites ci-après et donnant lieu à des composés 2,5-dicétopipérazines (2,5-DKP)<sup>199</sup> (*figure 60*). Dans ce contexte, nous avons étudié la stabilité des composés du *tableau 8*, afin de pouvoir évaluer leur activité biologique par la suite.

Le but de cette étude a donc été de suivre la formation possible de la cyclisation des produits **15**, **79** à **82** pour former des produits 2,5-DKP en milieu tamponné (à pH = 7.2). Un suivi cinétique par analyse HPLC a été réalisé à différents intervalles : t = 0 min, t = 24h et t = 7 jours. Comme attendu, nous avons pu observer une cyclisation spontanée avec une cinétique de formation du composé cyclique plus ou moins rapide mais non totale au bout de 7 jours : exemple pour le composé **15** (*figure 60*).



**Figure 60** : Etude de la stabilité du composé **15** dans le tampon TRIS à pH = 7.2. Le composé linéaire a un temps de rétention de 13.34 min et le composé cyclisé a un temps de rétention de 17.44 min.

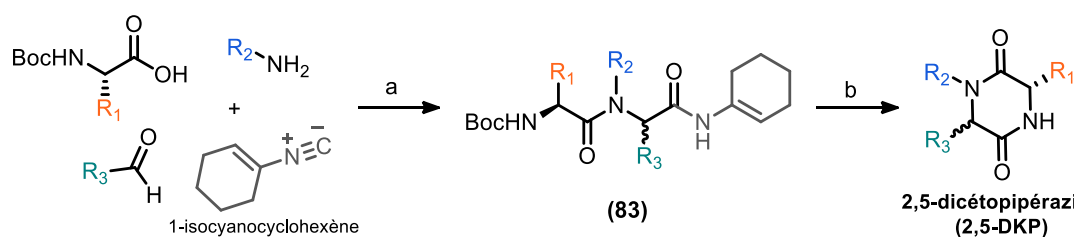
Cette cyclisation privilégiée par une formation d'un cycle à 6 n'a cependant pas été observée durant l'étape de clivage du support solide en milieu acide. Cette cinétique de cyclisation très lente (38 % en 7 jours pour le composé **15**) peut s'expliquer en partie par un encombrement stérique prononcé avec la présence de deux groupements aromatiques sur les positions des fonctions aldéhyde et amine. Ainsi, les 5 composés linéaires **15**, **79** à **82** ayant une pureté comprise entre 85 % et 100 % et au vu de leur lente cinétique de cyclisation en solution seront étudiés pour leur affinité envers l'ARN.

En conclusion, cette étude nous a permis d'identifier les fragments pouvant conduire à des produits de Ugi linéaires et ceux pouvant induire la formation de 2,5-DKP cycliques. Ces deux séries, linéaire et cyclique, seront préparées et évaluées en tant que ligands d'ARN sur nos ARN cibles (cf. *Chap. II-5*).

### II.3.1.2 Formation des composés 2,5-dicétopipérazines

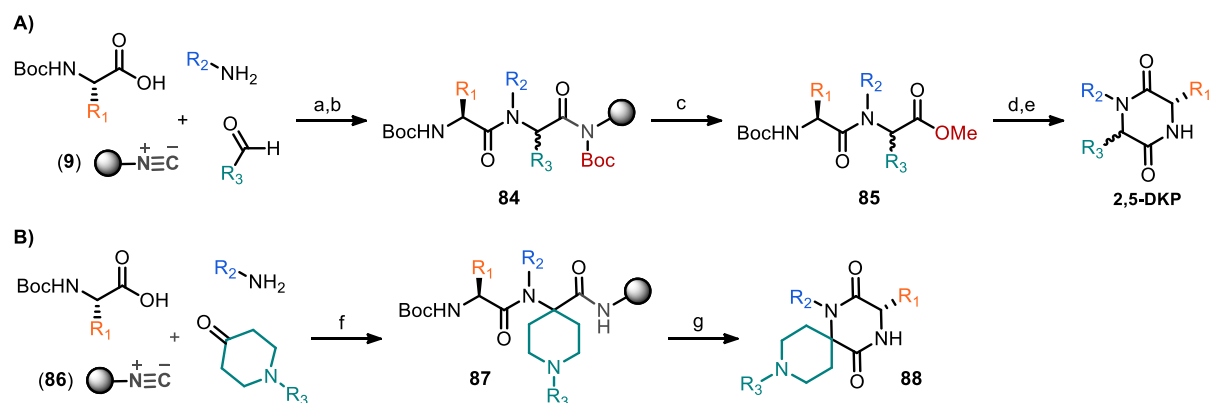
Dans la littérature, différentes méthodologies sont rapportées pour la formation des produits 2,5-DKP à la suite d'une procédure dite Ugi-Déprotection-Cyclisation (UDC)<sup>199,161</sup>. Nous avons fait le choix d'utiliser une synthèse par activation de l'amide provenant du groupement isocyanure pour induire la cyclisation.

Quand les réactions de Ugi ont lieu en solution, il est possible de sélectionner des isocyanures particuliers qui permettent d'obtenir après la réaction de Ugi, un groupement amide terminal pouvant être converti en un groupement labile pour permettre la cyclisation. Différents isocyanures de ce type ont été rapportés dans la littérature<sup>199</sup>, parmi lesquels le 1-isocyanocyclohexène décrit par Armstrong en 1996<sup>200</sup>. Cet isocyanure a ensuite été utilisé par Hulme et *al.*<sup>201</sup> pour des réactions de Ugi avec des acides aminés protégés par un groupement Boc (*schéma 30*). Le clivage de ce dernier en milieu acide permet d'activer la partie cyclohexène de l'intermédiaire **83** en un ion *N*-acyliminium labile qui induit la cyclisation pour obtenir le composé 2,5-dicétopipérazine.



*Schéma 30* : Synthèse de composés 2,5-dicétopipérazines à partir d'une activation de l'amide en solution.  
 Conditions : **a)** MeOH, t.a. ; **b)** 10% TFA, DCE.

Dans le cadre des réactions de Ugi en phase solide, les fonctions isocyanures sont portées par le support solide. Deux méthodes différentes ont été décrites pour la cyclisation permettant la formation de composé 2,5-DKP : soit une cyclisation après clivage de la résine (*schéma 31, A*), soit une cyclisation induite lors du clivage de la résine (*schéma 31, B*)<sup>199</sup>.



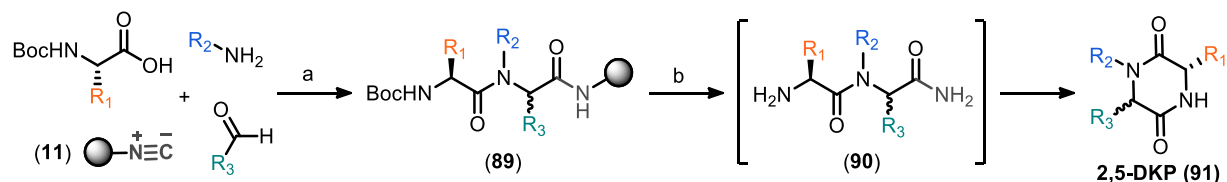
**Schéma 31** : Synthèse de composés 2,5-dicétopipérazines à partir d'une activation de l'amide provenant d'une résine-isocyanure. **A)** Cyclisation après le clivage de la résine. **B)** Cyclisation pendant le clivage de la résine. Conditions : a) MeOH/THF ; b)  $\text{Boc}_2\text{O}$ ,  $\text{Et}_3\text{N}$ , DMAP,  $\text{CH}_2\text{Cl}_2$  ; c) NaOMe, MeOH/THF ; d) 10 % AcCl, MeOH, e) 5 %  $\text{Et}_2\text{NH}$ , DCE ou Dowex basique (1x8-50) ; f) MeOH/THF (1:1), 65°C ; g) 1) TFA,  $\text{CH}_2\text{Cl}_2$  t.a., 2) AcOH, toluène, 90°C.

Dans le premier cas (*schéma 31, A*), une résine-isocyanure **9** (cf. *schéma 51*) nommée « résine safety-catch » est employée avec un acide aminé protégé par un groupement Boc<sup>165</sup>. Une fois le produit Ugi formé, une étape de protection de l'amide par un groupement Boc est nécessaire pour fournir l'intermédiaire **84**. Cette étape caractéristique du support solide permet son clivage en milieu basique pour former une fonction ester intermédiaire **85** qui va permettre la cyclisation. Cette dernière est ensuite réalisée dans des conditions légèrement basiques pour donner les composés 2,5-DKP. Cette méthodologie a également été employée sur d'autres résines-isocyanures dans le but d'obtenir des composés cyclisés post-Ugi par des divers nucléophiles<sup>202,203</sup>.

En ce qui concerne la cyclisation pendant le clivage du support solide (*schéma 31, B*), Habashita et *al.*<sup>204</sup> ont mis au point la synthèse de spiro-2,5-DKP à partir d'une résine méthylène-isocyanure **86** (préparée à partir de la résine « aminomethylated »). La réaction initiale d'Ugi a exigé le chauffage à 65 °C pour réaliser la conversion de la pipéridone vers le produit **87**. Une élimination du groupement Boc suivie du clivage du support avec de l'acide acétique fournit le produit spiro-2,5-DKP (**88**). Une élévation de la température à 90°C est nécessaire pour procéder à cette cyclisation. Dans cette même étude, la résine Rink amide-isocyanure a également été utilisée. Cependant, l'élimination du groupement Boc a été réalisée avec du TMSOTf et de la 2,6-lutidine suivie par une cyclisation dans l'AcOH/toluène à 90°C. Enfin, Chen et *al.*<sup>166</sup> ont également utilisé la résine Rink amide-isocyanure avec un acide aminé protégé par un groupement Fmoc. Cette protection rajoute une étape d'élimination du groupement Fmoc supplémentaire pour ensuite procéder à l'étape de clivage induisant la cyclisation par de l'AcOH dans le DCE. Cette cyclisation a également nécessité un chauffage à 60°C pour avoir une conversion totale.

Au travers de ces exemples, nous avons pu remarquer que l'ensemble des procédures décrites utilise des températures élevées pour convertir les produits linéaires en produits cyclisés. Dans ce manuscrit,

nous envisageons un schéma réactionnel utilisant des groupements protecteurs acido-labiles de type Boc sur les différents acides aminés. Ainsi, l'étape de clivage de la résine-Ugi (**89**) réalisée en milieu acide devrait entraîner une cyclisation des composés linéaires (**90**) en composés 2,5-DKP (**91**) directement sans étape supplémentaire (schéma **32**).



**Schéma 32** : Description du schéma réactionnel permettant d'obtenir des composés 2,5-dicétopipérazines (**91**) en deux étapes à partir d'une résine-isocyanure. Conditions : **a)** CH<sub>2</sub>Cl<sub>2</sub>/MeOH, t.a., 4-6 jours ; **b)** TFA, CH<sub>2</sub>Cl<sub>2</sub>, t.a., 1h30.

La synthèse des différents composés préparés selon cette méthodologie (**Série A**) sera ainsi décrite dans la section suivante.

### II.3.1.3 Synthèse des composés de la série A

Dans ce projet de thèse, l'ensemble des fragments portant la fonction amine présente un bras espaceur entre cette fonction et le composé. Afin d'étudier la capacité de ces composés à former un produit de type 2,5-dicétopipérazine, nous avons réalisé deux tests dans lesquels l'amine aromatique utilisée pour les composés **A1** à **A2** précédemment a été remplacée par une amine *N*-Boc-1,4-butanediamine avec la lysine (tableau **9**).

**Tableau 9** : Résultats des tests de cyclisation en présence de la *N*-Boc-1,3-propanediamine pour former des produits 2,5-DKP :

Entrée	Composé	Aldéhyde : R <sub>1</sub>	Support	Rendement <sup>a</sup>	Pureté <sup>b</sup>
<b>1</b>	A1		S <sub>1</sub> (54 %)	17 %	70 %
<b>2</b>	A2		S <sub>1</sub> (24 %)	11 %	52 %

**Conditions** : Support S<sub>1</sub> : Rink amide MBHA ; **1**) Résine-isocyanure (1 éq.), aldéhyde (2 éq.), amine (2 éq) et acide aminé (2 éq), CH<sub>2</sub>Cl<sub>2</sub>/MeOH, t.a, 6 jours, **2**) 95% TFA dans le CH<sub>2</sub>Cl<sub>2</sub>. <sup>a</sup> Rendement calculé par rapport au « loading » de la résine de départ. <sup>b</sup> Pureté mesurée par HPLC à UV = 214 nm.

Le mélange **A1** (entrée **1**), a été obtenu avec un faible rendement de 17 % compte tenu de la fonctionnalisation de la résine (54 %). Lors de cette réaction, une faible proportion de produit linéaire

a été observée en HPLC ce qui conduit à une pureté de 70 % du composé DKP. Après un deuxième traitement par trituration, le produit **A1** a été obtenu avec une pureté supérieure à 95 % et a pu être caractérisé par RMN  $^1\text{H}$  et confirmé par HRMS. En ce qui concerne le mélange **A2 (entrée 2)**, le produit 2,5-DKP a également été isolé, mais une plus grande proportion de produit linéaire a été obtenue après clivage de la résine entraînant une baisse de la pureté à 52 %. Une purification par HPLC semi-préparative a donc été nécessaire pour isoler le composé **A2** et ce dernier a été caractérisé par RMN  $^1\text{H}$  et LCMS.

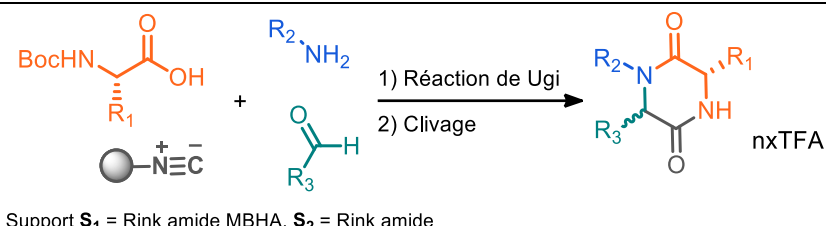
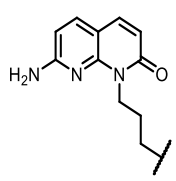









Il semble donc que la présence d'une chaîne alkyle soit favorable pour la cyclisation, probablement car une telle chaîne n'induit pas de gêne stérique lors de la cyclisation, contrairement à ce que nous avons précédemment observé avec les anilines. Les deux fragments naphtyridine *N*-alkylé (**42**) et adénine modifiée (**21**) préalablement synthétisés, ont aussi été utilisés pour former des composés 2,5-DKP *tableau 10* (page suivante).

L'amine naphtyridine *N*-alkylée (**entrées 1 à 7**) a été employée dans la réaction en présence de trois acides aminés (phénylalanine, lysine et tryptophane) en tant qu'acides et des motifs hétéroaromatiques ou aromatiques en tant qu'aldéhydes. L'ensemble de ces motifs conduit à la préparation de 7 composés et pour chacune de ces réactions, il a fallu réaliser une étape de cyclisation supplémentaire après le clivage de la résine, car la proportion en produit 2,5-DKP observée par HPLC par rapport au produit linéaire était trop faible. Par exemple, après clivage de la résine, le résultat du produit formé dans l'**entrée 1** a été obtenu sous forme linéaire avec une pureté de 80 %. Afin d'obtenir une cyclisation totale, ces composés ont donc été placés en solution dans le méthanol pour une durée de deux à cinq jours, sauf pour le produit **A8 (entrée 6)** pour lequel une température de 50°C a été requise pendant 2 jours pour obtenir un taux de cyclisation total.

Globalement, les rendements de cette série sont relativement faibles (entre 7 % et 36 %) par rapport au taux de fonctionnalisation des différents supports utilisés, le meilleur rendement ayant été obtenu avec le mélange **A8** (groupements tryptophane et naphthalène, **entrée 6**) qui a conduit à un composé 2,5-DKP avec une pureté de 82 %. Les combinaisons de groupements phénylalanine/indole (**entrée 1**) et lysine/indole (**entrée 5**) semblent en revanche peu adaptés pour l'obtention de produits de Ugi sur support solide (rendements 8 % et 7 % respectivement).

En ce qui concerne le pourcentage en produit cyclisé (pureté), l'acide aminé tryptophane (**entrées 6 et 7**) a donné des résultats plus élevés (de l'ordre des 80 %) par rapport aux composés obtenus avec la phénylalanine et la lysine (de l'ordre de 60 %).

**Tableau 10 :** Résultats obtenus par l'utilisation du motif naphtyridine N-alkylé (**42**) et de l'adénine modifiée (**21**) pour les composés 2,5-DKP de la série A :

								
Entrée	Composé	Amine : R <sub>2</sub>	AA : R <sub>1</sub>	Aldéhyde : R <sub>3</sub>	Support (Rendement maximal)	Rendement <sup>a</sup>	Pureté <sup>b</sup>	
1	A3		Phe		S <sub>1</sub> (26%)	8 %	Linéaire (80 %)	
2	A4				S <sub>1</sub> (26%)	13 %	54 %	
3	A5				S <sub>1</sub> (26%)	13 %	60 %	
4	A6				S <sub>2</sub> (43%)	24 %	63 %	
5	A7			Lys		S <sub>2</sub> (43%)	7 %	50 %
6	A8			Trp		S <sub>2</sub> (43%)	36 %	82 %
7	A9				S <sub>1</sub> (nd)	17 %	81 %	
8	A10			Trp		S <sub>2</sub> (43%)	39 %	80 %
9	A11			Phe		S <sub>1</sub> (nd)	47 %	68 %

**Conditions :** **1)** Résine-isocyanure (1 éq.), aldéhyde (2 éq.), (Boc)-amine (2éq) et acide aminé (2 éq), CH<sub>2</sub>Cl<sub>2</sub>/MeOH, t.a, 6 jours, **2)** Résine S<sub>1</sub> : 95% TFA dans le CH<sub>2</sub>Cl<sub>2</sub>, Résine S<sub>2</sub> :15 % TFA dans le CH<sub>2</sub>Cl<sub>2</sub>, t.a., 1h30. <sup>a</sup> Rendement calculé par rapport au « loading » de la résine de départ. <sup>b</sup> Pureté mesurée par HPLC à 214 nm. <sup>c</sup> Rendement maximal déterminé à 43% sur les cinq étapes.

Pour l'évaluation de ces composés en tant que ligands d'ARN qui est envisagée dans la suite du projet, les mélanges de composés de faible pureté (**A3**, **A5** et **A6**) ont été purifiés par HPLC semi-préparative et caractérisés par RMN <sup>1</sup>H et leur structure a été confirmée par HRMS. En revanche, nous avons décidé d'utiliser les mélanges des produits obtenus avec une pureté supérieure ou égale à 80 % (composés **A8** et **A9**) sans purification supplémentaire. Une analyse par HRMS a confirmé la structure des composés souhaités. La purification ultérieure de ces composés pourra être envisagée si leur activité biologique résultant est particulièrement intéressante.

Pour les composés contenant l'adénine modifiée (**A10-11**, entrées **8** et **9**), nous pouvons observer que les rendements sont légèrement meilleurs et comparables au rendement du produit **A8**. Il semble donc que ce motif soit favorable à une réaction de Ugi sur support solide et ce indépendamment de la combinaison acide aminé/aldéhyde. Pour ce qui est de la conversion en produit cyclique, le composé

**A10** a nécessité un chauffage à 50°C pour atteindre le taux de cyclisation maximum donnant une pureté du produit à 80 %, alors qu'avec le composé cyclisé **A11**, la conversion totale n'a pas nécessité de chauffage et a donné une pureté de 68 %. Une purification a donc été réalisée sur ce dernier avant son évaluation biologique.

Pour conclure sur la préparation de la série A contenant des produits 2,5-DKP préparés sur phase solide, nous avons pu mettre au point une méthodologie de synthèse dans laquelle l'amine alpha des acides aminés est directement impliquée dans la réaction de cyclisation. L'élévation de la température a été nécessaire quand l'encombrement stérique était trop important (composés **A8** et **A10**). Grâce à cette méthodologie, nous avons pu obtenir **9 composés** de la famille 2,5-dicétopipérazine avec des rendements allant de 8 à 39% sur les cinq étapes et des puretés allant de 50 à 80%, ce qui a rendu une étape de purification par HPLC semi-préparative nécessaire. L'ensemble de ces composés sera étudié pour leur activité biologique sur les systèmes Toxine-Antitoxine qui font l'objet de ce travail de thèse. Sur ces 9 composés obtenus, 5 composés seront étudiés sous forme purifiés et 4 composés seront étudiés sous forme de mélange avec une pureté supérieur à 80 % (*annexe 3*).

Dans la conception de notre projet, des composés linéaires ont aussi été préparés pour comparaison. Cependant, les acides aminés choisis ne permettant pas de favoriser une structure linéaire, plusieurs stratégies ont dû être utilisées pour obtenir une nouvelle série de molécules linéaires : la **série B**.

### II.3.2 Synthèses des composés de la série B

Pour obtenir des composés linéaires par une réaction de Ugi, plusieurs possibilités peuvent être envisagées en fonction de l'acide carboxylique employé. Dans la littérature, Portlock *et al.*,<sup>150</sup> ont décrit la formation d'une librairie de composés  $\alpha$ -aminoacylamides à partir d'acides aminés dans lesquels l'amine était encombrée (amine tertiaire ne permettant pas la formation de produit 2,5-DKP). Dans cette étude, une résine Rink amide a été utilisée et a permis d'obtenir les composés résultants comportant une fonction amide après l'étape de clivage de la résine. Une autre possibilité est l'utilisation d'acides carboxyliques ne possédant pas d'amine en position alpha mais mimant la chaîne latérale des acides aminés. Nous avons choisi cette deuxième possibilité pour réaliser la série B afin d'obtenir des composés linéaires pseudo-dipeptides (*tableau 11* et *12*).

Pour le début de cette série, nous avons utilisé les composés naphtyridine *N*-alkylé (**42**), adénine modifiée (**21**) et cytosine modifiée (**25**) en tant que composants portant la fonction amine. Pour les acides carboxyliques, nous avons d'abord sélectionné l'acide phénylacétique mimant la phénylalanine.

Tableau 11 : Résumé des synthèses utilisant l'acide phénylacétique pour la série B sur support solide :

Entrée	Composé	Amine : R <sub>2</sub>	Rendement <sup>a</sup>	Pureté <sup>d</sup>
1	B1a		29 % <sup>b</sup>	56 %
2	B2a		26 % <sup>c</sup>	88 %
3	B3a		9 % <sup>b</sup>	42 %

Support S<sub>1</sub> = Rink amide MBHA,

**Conditions :** **1)** Résine-isocyanure (1 éq.), aldéhyde (2 éq.), (Boc)-amine (2éq.) et acide aminé (2 éq.), CH<sub>2</sub>Cl<sub>2</sub>/MeOH, t.a, 6 jours, **2)** 95% TFA dans le CH<sub>2</sub>Cl<sub>2</sub>, t.a., 1h30. <sup>a</sup> Rendement calculé par rapport au « loading » de la résine de départ. <sup>b</sup> Taux de fonctionnalisation de la résine non déterminé. <sup>c</sup> Taux de fonctionnalisation de 38 %. <sup>d</sup> Pureté mesurée par HPLC à 214 nm.

En utilisant les conditions de réactions de Ugi précédentes et la résine Rink amide MBHA, nous avons d'abord synthétisé le composé **B1a** (entrée 1). Après clivage de la résine, nous avons obtenu un rendement de 29 %. Cependant, lors de l'étape du clivage de support solide, une hydrolyse de la fonction amide a été observée. En effet, l'analyse par HPLC nous a montré un mélange de plusieurs produits qui ont pu être identifiés (*figure 61*). Le produit souhaité (**B1a**) étant néanmoins majoritaire (53 %, tr = 14,88 min).

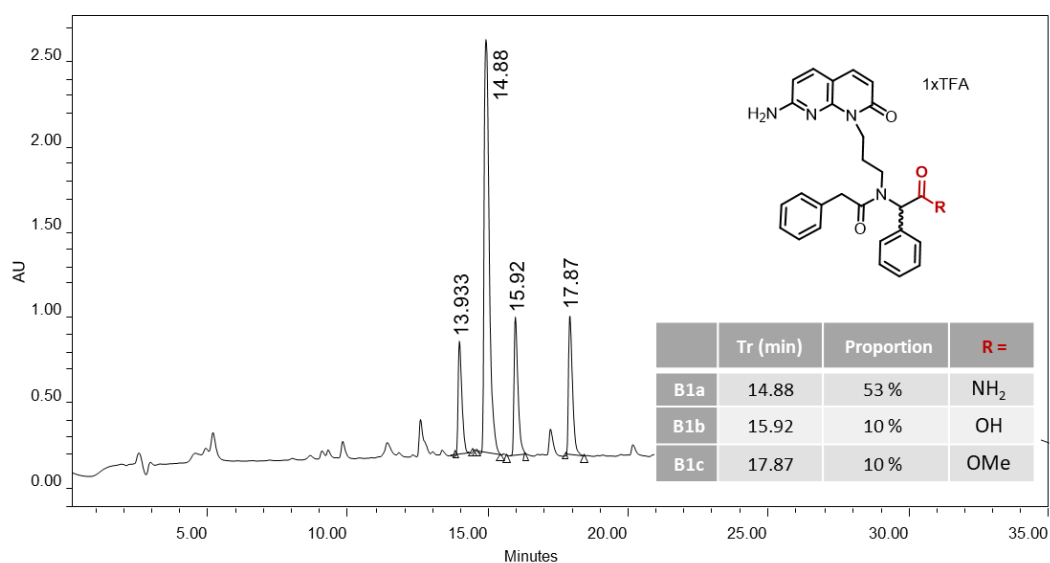
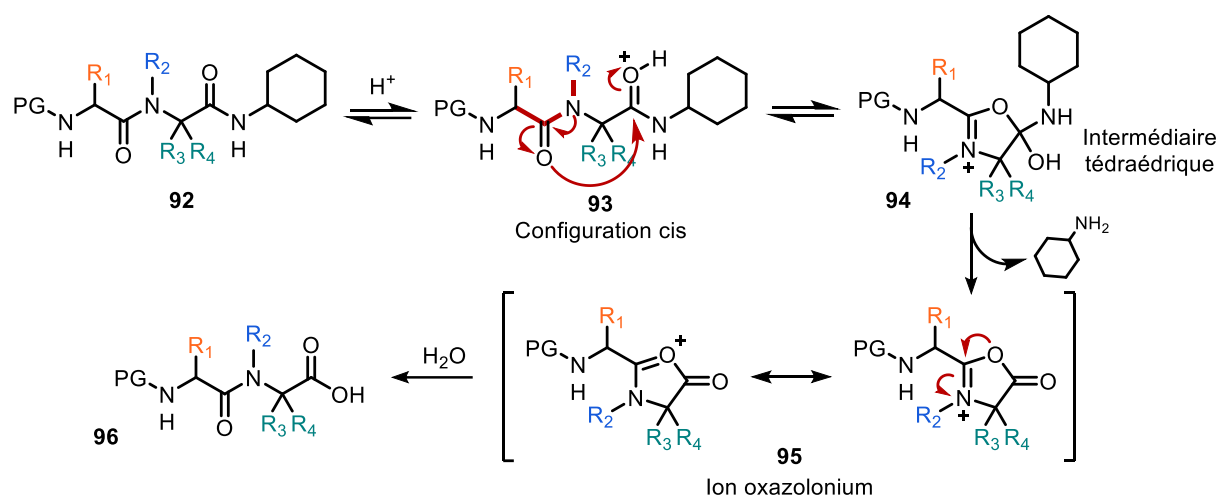


Figure 61 : Profil HPLC du mélange B1.



Ce type d'hydrolyse a précédemment été rapportée dans la littérature pour la formation de composés  $\alpha,\alpha$ -dialkylglycines<sup>205,206</sup>. Plus récemment, en 2021, Ramos-Tomillero et *al.*<sup>207</sup> ont rapporté l'hydrolyse de composés de type  $\alpha,\alpha$ -dialkylglycine *N*-alkylé (**92**) via une réaction de Ugi (*schéma 33*). Un intermédiaire oxazolonium (**95**) est d'abord obtenu à l'aide d'une configuration *cis* par rapport au carboxamide de l'intermédiaire **93** avant d'être hydrolysé par une molécule d'eau donnant l'acide correspondant **96**.



*Schéma 33* : Acidolyse des composés  $\alpha,\alpha$ -dialkylglycine *N*-alkylés **92** pour donner le composé **96** via l'intermédiaire d'ions oxazolonium **95**.

Cette hydrolyse est plus ou moins importante selon le motif présent en position *N*-acylée et elle est favorisée par des motifs aromatiques riches en électrons<sup>208</sup>.

Dans notre cas, c'est la présence d'eau lors de l'étape de clivage qui a conduit à l'hydrolyse et à l'acide correspondant (**B1b**) dans une proportion importante de 14 % pour le mélange **B1** (*figure 61*, *tr* = 15,92 min). Nous avons également mis en évidence le dérivé ester méthylique (**B1c**) dans la même proportion de 14 % (*figure 61*, *tr* = 17,87 min), la présence de méthanol provenant de l'étape de rinçage de la résine à la suite du clivage du support. En remplaçant le méthanol par l'acétonitrile en tant que solvant, ce sous-produit a pu être évité mais l'acide carboxylique correspondant s'est révélé être toujours présent. Enfin, le dernier sous-produit (**B1d**) a également été identifié (proportion de 12%, *figure 61*, *tr* = 13.93 min) et correspond au produit de la réaction de Ugi avec l'acide benzoïque comme motif acide, la présence de ce dernier provenant d'une bouteille commerciale de benzaldéhyde contenant cette impureté. Ainsi, le mélange **B1** est composé d'un ensemble de 4 produits **B1a** (53 %), **B1b** (10 %), **B1c** (10 %) et **B1d** (12 %) donnant un rendement global de 85 %.

Néanmoins, ces sous-produits n'ont pas été observés lorsque nous avons utilisé l'adénine modifiée (**21**) et la cytosine modifiée (**25**) pour former les composés linéaires **B2a** (*tableau 11*, *entrée 2*) et **B3a** (*tableau 11*, *entrée 3*). Cette différence pourrait être due à l'encombrement qui est plus important avec une chaîne alkyle plus courte pour ces deux amines et ne favoriserait pas la formation de

l'intermédiaire tétraédrique (cf. *schéma 33*). Le composé **B2a** a été obtenu avec un rendement de 26 % et une pureté de 88 %. En ce qui concerne le composé **B3a**, un manque de solubilité de la cytosine modifiée, même après un ajout de DMF, a conduit à un rendement plus faible (9 %) et à une pureté de 42 %. Pour améliorer ce rendement, il serait intéressant dans le futur d'essayer cette réaction à une température plus élevée permettant une solubilité totale du mélange réactionnel.

Afin, d'évaluer si d'autres acides carboxyliques entraînent une hydrolyse, nous avons réalisé la synthèse des composés **B4** à **B7** à partir du naphtyridine *N*-alkylé (**42**) (*tableau 12*).

**Tableau 12** : Résumé des synthèses utilisant le motif naphtyridine *N*-alkylé (**42**) avec divers acides carboxyliques pour la série B sur support solide :

Entrée	Composé	Amine : R <sub>2</sub>	Acide carboxylique : R <sub>1</sub>	Aldéhyde : R <sub>3</sub>	Rendement <sup>a</sup>	Pureté <sup>c</sup>
1	B4a				32 % <sup>b</sup>	69 %
2	B5a				17 % <sup>b</sup>	52 %
3	B6a				11 % <sup>b</sup>	70 %
4	B7a				- % <sup>b</sup>	- %

**Conditions** : 1) Résine-isocyanure (1 éq.), aldéhyde (2 éq.), amine (2 éq.) et acide aminé (2 éq.), CH<sub>2</sub>Cl<sub>2</sub>/MeOH, t.a, 6 jours, 2) 95% TFA dans le CH<sub>2</sub>Cl<sub>2</sub>, t.a., 1h30. <sup>a</sup> Rendement calculé par rapport au « loading » de la résine de départ. <sup>b</sup> Taux de fonctionnalisation de la résine non déterminé. <sup>c</sup> Pureté mesurée par HPLC à 214 nm.

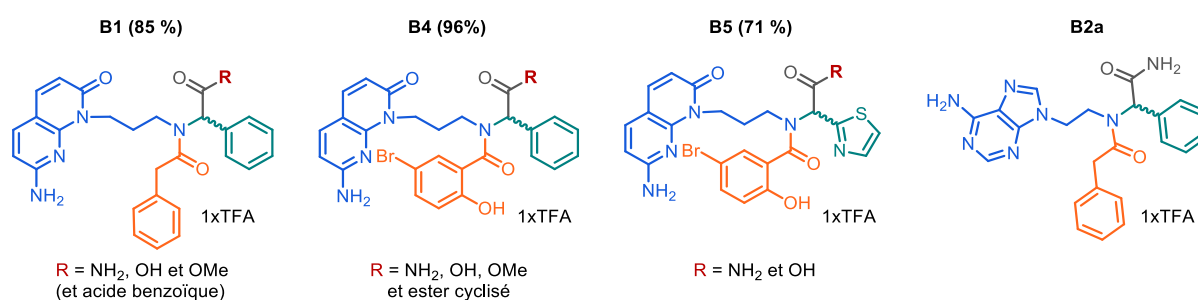
Pour le composé **B4a** comportant l'acide 5-bromosalicylique, une proportion de 69 % a été identifiée pour un rendement de 32 %. Nous avons également obtenu les mêmes sous-produits résultant de l'hydrolyse comme observés avec le mélange **B1** (*tableau 11*). L'acide carboxylique (**B4b**) et l'ester méthylique (**B4c**) ont été obtenus dans les proportions de 7 % et de 9 % respectivement. Un autre sous-produit a également été observé. En effet, l'acide 5-bromosalicylique contient un groupement hydroxyle qui peut se comporter comme un nucléophile lors de l'hydrolyse. Un ester cyclisé (**B4d**) résultant de cette attaque nucléophile a été identifié dans une proportion de 11 %. Ainsi, le mélange **B4** est composé d'un ensemble de 4 produits **B4a** (69 %), **B4b** (7 %), **B4c** (9 %) et **B4d** (11 %) donnant une proportion globale de 96 %.

Ensuite avec ce même acide (**entrée 2**), nous avons utilisé le 2-thiazolecarboxaldehyde pour former le mélange **B5a** de deux composés avec un rendement de 17 %. Le composé linéaire souhaité (**B5a**) a été obtenu dans une proportion de 52 %. L'acide carboxylique (**B5b**) a aussi été observé à hauteur de 19 % donnant une proportion globale de 71 %.

Pour le composé **B6a** (**entrée 3**) comportant un imidazole qui pourrait mimer l'acide aminé histidine, l'hydrolyse n'a pas été observée. Ainsi, le composé **B6a** a été obtenu avec un rendement de 11 % et une pureté de 71 %.

Enfin, l'utilisation de l'acide 1*H*-benzimidazole-6-carboxylique (**entrée 4**) ne nous a pas permis d'accéder au composé **B5a** en raison de la faible solubilité de l'acide de départ. Pour remédier à ce problème, nous pourrions avoir recours soit à une élévation de la température, soit à une protection de l'amine du benzimidazole permettant une meilleure solubilité.

En conclusion sur cette série B, nous avons pu obtenir des composés linéaires avec différents fragments amines synthétisés précédemment. Les rendements ont été globalement faibles en particulier en raison de la formation de plusieurs produits secondaires provenant de l'hydrolyse de la fonction amide, observée avec le motif naphtyridine *N*-alkylé. Afin d'étudier ces composés en tant que ligands d'ARN, nous avons fait le choix de ne pas les purifier et d'étudier l'affinité des mélanges (*figure 62*). Ainsi les mélanges **B1**, **B4** et **B5** en quantité suffisante (rendement > 15 %) seront employés en tant que tels lors de l'évaluation biologique. En revanche, le composé **B2a** a été purifié et caractérisé par RMN du proton.



**Figure 62 :** Composés issus de série B pour être étudiés en tant que ligands d'ARN.

### II.3.3 Conclusion de la synthèse sur support solide

A partir de la synthèse sur support solide en utilisant une réaction de Ugi, nous avons exploré la préparation de deux séries de molécules de structure 2,5-dicétopipérazine pour la **série A** et de structure linéaire pseudo-dipeptide pour la **série B** (*figure 63*).

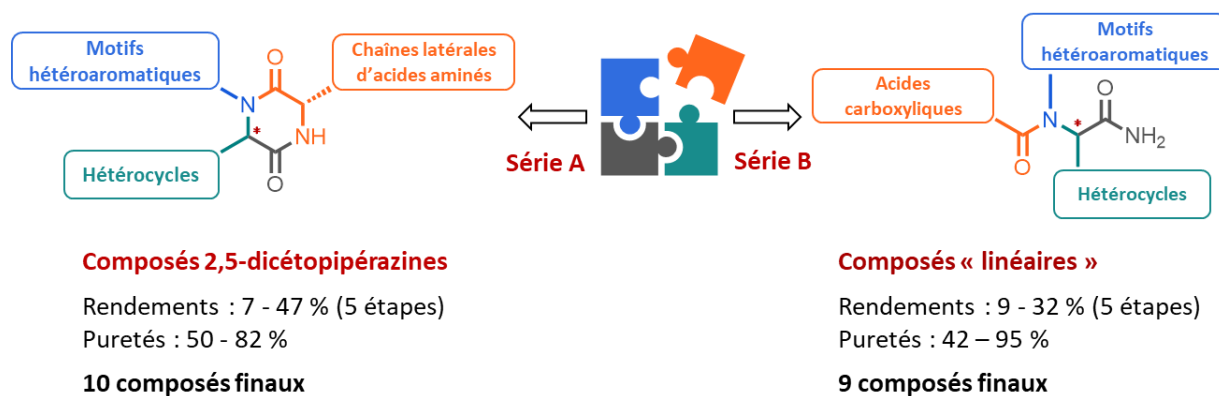


Figure 63 : Bilan des molécules obtenues par la synthèse sur support solide.

Le choix de l'acide aminé dans la conception du projet nous a permis d'obtenir ces différentes structures. Dans l'étude préliminaire, **5 composés (15, 79 à 82)** de puretés comprises entre 85 % à 100 % ont présenté une stabilité suffisante sous cette forme linéaire (pas de cyclisation observée) et seront étudiés lors de l'évaluation biologique (*annexe 3*).

Lorsque nous avons utilisé différents fragments aminés tels que le fragment naphtyridine *N*-alkylé (**42**) et l'adénine modifiée (**21**), l'encombrement stérique étant faible, des composés cycliques de type 2,5-DKP ont pu être obtenus. La formation de composés 2,5-DKP a été réalisée directement lors de l'étape de clivage de la résine en milieu acide. Le rendement de la réaction de cyclisation a pu être amélioré grâce au chauffage pour certains produits. Au total, nous avons obtenu **6 composés (A2, A3, A5, A6, A11a et A11b)** de pureté >95 % après purification) et **4 mélanges (A1, A8, A9 et A10)** de pureté de 81 % à 95 % qui seront étudiés en tant que ligands d'ARN (*annexe 3*).

Pour préparer les composés linéaires de la série B, nous avons utilisé des acides carboxyliques ne comportant pas d'amine en position alpha, cette dernière étant nécessaire pour la cyclisation. Néanmoins, nous avons observé dans certains cas, plusieurs sous-produits issus d'une hydrolyse de la fonction amide. Ainsi, pour réaliser les études biologiques, le mélange des produits linéaires (**3 mélanges, B1, B4 et B5**) a été utilisé (*annexe 3*). Une purification ultérieure sera réalisée pour les mélanges démontrant une affinité élevée vis-à-vis de l'ARN et des résultats d'activité prometteurs. Nous avons également purifié le composé **B2a** par HPLC semi-préparative.

En conclusion, nous avons pu obtenir **19 produits linéaires ou cyclisés** en appliquant la réaction de Ugi sur support solide. Puisque dans certains cas, le rendement de la réaction n'était pas satisfaisant ou la pureté des produits obtenus pas suffisante et donc nécessitant une étape de purification ultérieure (HPLC semi-préparative), nous avons fait le choix de changer de méthodologie pour une synthèse en phase liquide. Cette dernière nécessitera une étape de purification avant déprotection mais pourrait être une méthodologie plus adaptée selon les motifs utilisés dans la réaction, en particulier pour les motifs réalisés en plusieurs étapes tels que la 2-désoxystreptamine.

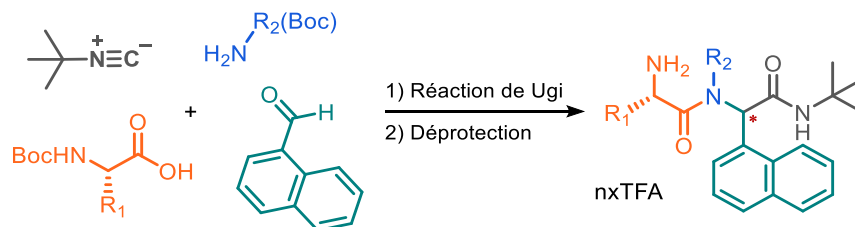
## II.4 Préparation des ligands d'ARN en solution : Série C

Dans cette partie, les différents composés issus de la réaction de Ugi réalisée en solution avec l'isocyanure *tert*-butyle (disponible commercialement) seront décrits. L'ensemble des fragments synthétisés dans la partie *Chap. II.2* sera utilisé dans un premier temps dans une série dite « naphtalène » (**Série C-1**) à partir du 1-naphthaldéhyde. Dans un deuxième temps, la série avec l'aldéhyde 2-désoxystreptamine sera présentée formant la **Série C-2**. Enfin, une troisième partie présentera les autres synthèses réalisées au cours de ce projet (**Série C-3**).

### II.4.1 Synthèses de la série C-1 : série naphtalène

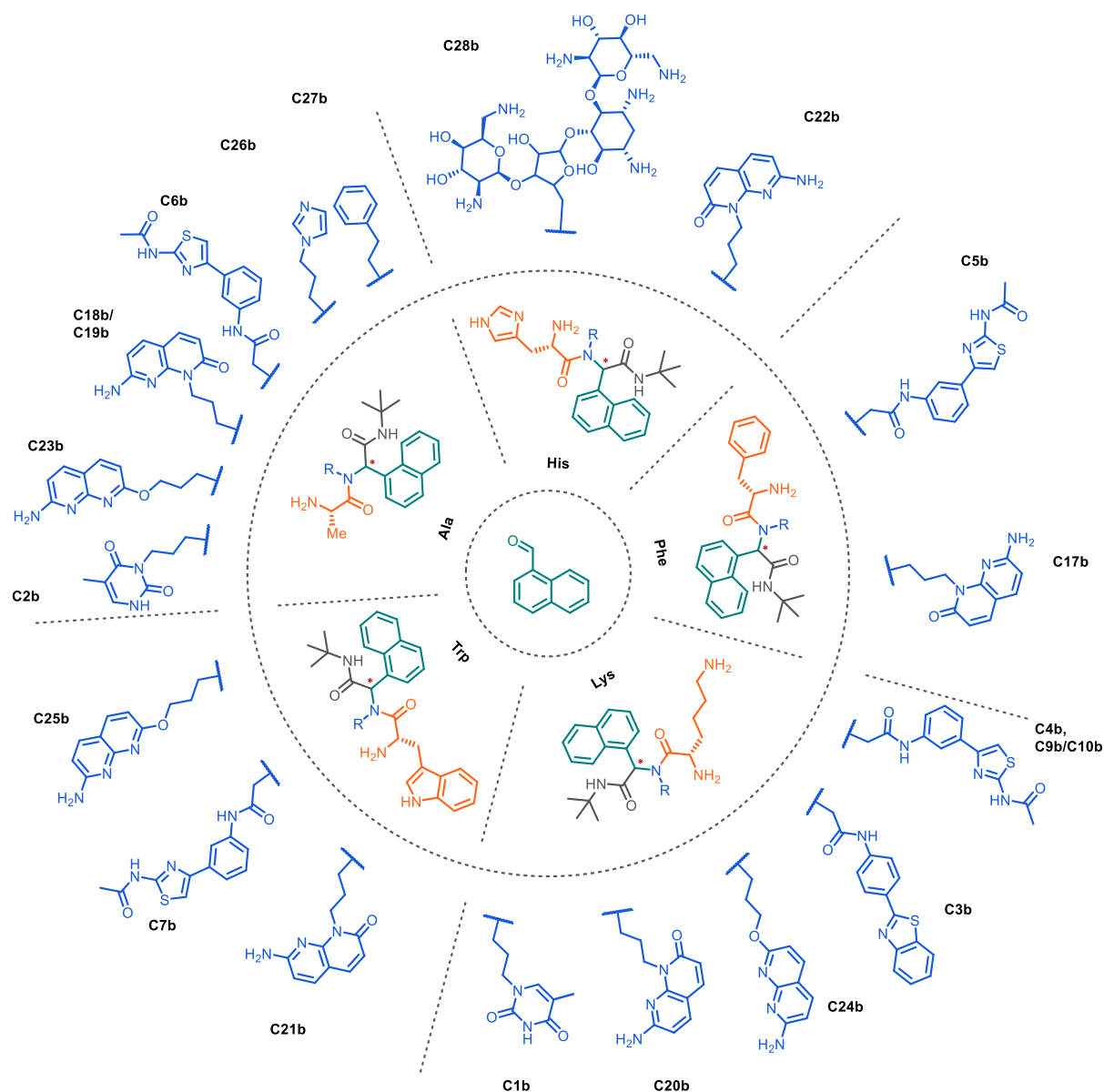
Dans cette première série, notre choix s'est porté sur le motif 1-naphthaldéhyde qui devrait conférer des propriétés intercalantes aux ligands d'ARN synthétisés. Pour réaliser les réactions d'Ugi, les cinq acides aminés choisis préalablement : l'alanine (Ala), la phénylalanine (Phe), le tryptophane (Trp), l'histidine (His) et la lysine (Lys) ont été utilisés. Enfin, les différents fragments comportant la fonction amine, synthétisés et décrits précédemment, ont permis de réaliser plusieurs sous-séries.

Pour préparer cette large série, chaque réaction a été réalisée selon une procédure générale<sup>158</sup> en deux étapes (*schéma 34*).



**Schéma 34** : Description de la synthèse permettant d'obtenir les composés de la série C-1. Conditions : **1)** MeOH, t.a., 24h à 72h ; **2)** TFA, CH<sub>2</sub>Cl<sub>2</sub>, t.a., 4h.

Les réactions de Ugi réalisées en solution sont effectuées dans le méthanol en utilisant un équivalent pour chaque réactif. Le suivi par chromatographie sur couche mince nous a permis d'adapter le temps nécessaire à la réaction pouvant varier de 24h à 72h. Une purification du produit Ugi formé (désigné par la lettre **XXa**) a été nécessaire donnant un mélange de 2 diastéréoisomères difficilement séparables. Cependant, certains de ces diastéréoisomères ont pu être séparés. Ensuite, une étape d'élimination des groupements protecteurs acido-labiles a été réalisée. Cette étape a été effectuée avec de l'acide trifluoroacétique dans du CH<sub>2</sub>Cl<sub>2</sub>. Un mélange de diastéréoisomères sous forme de sels de TFA a été obtenu (désigné par la lettre **XXb**). Cette procédure nous a permis d'élaborer un ensemble de 18 molécules présentées dans la *figure 64* de pureté supérieure à 95 %.



**Figure 64 :** Représentation des molécules de la *série C-1*. Chaque composé est obtenu à partir du 1-naphthalaldéhyde en faisant varier l'acide aminé puis le fragment portant l'amine. Les composés obtenus sont sous forme de sels de TFA, non représentés ici.

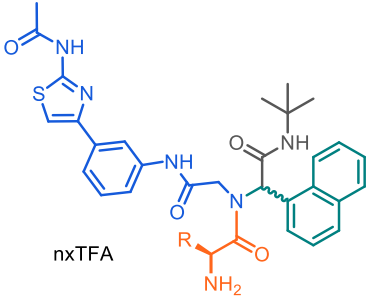
Cette partie s'inscrit dans la continuité des travaux initiés par le stagiaire de master 2, O. BORODIN qui a réalisé la synthèse (étape de la réaction de Ugi et l'étape de déprotection) des composés thymines **C1b** et **C2b** ainsi que le composé **C3b** et **C4b** (structures *figure 64*). Ces molécules ont été obtenues à partir d'amines protonées sous forme de sels de TFA. Pour réaliser la réaction de Ugi avec ces différentes amines, un ajout d'hydroxyde de potassium (1 éq.) a été effectué<sup>158</sup>. Les rendements de l'étape de Ugi pour ces quatre composés (**C1a**, **C2a**, **C3a** et **C4b**) ont ainsi été obtenus entre 33 % et 51 % (*tableau 13*). Pour l'étape de déprotection, les molécules **C1b**, **C2b**, **C3b** et **C4b** ont été obtenues après filtration dans des rendement satisfaisants (*tableau 13*).

**Tableau 13** : Rendements obtenus pour les composés **C1b**, **C2b**, **C3b** et **C4b** :

Composé	Rendements	
	Réaction d'Ugi (a)	Étape de déprotection (b)
<b>C1b</b>	36 %	85 %
<b>C2b</b>	33 %	91 %
<b>C3b</b>	46 %	92 %
<b>C4b</b>	51 %	60 %

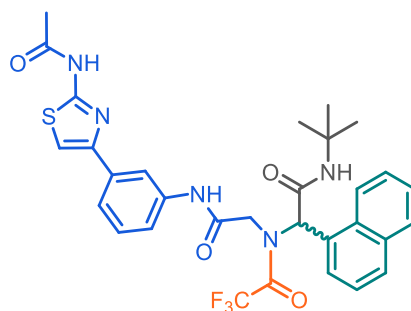
À la suite de ces travaux, une série comportant la nucléobase artificielle « S » a été synthétisée à la suite du composé **C4b** préparé précédemment. Différents acides aminés ont été choisis pour conduire aux composés **C5b** (Phe), **C6b** (Ala), **C7b** (Trp) et **C8b** (His) et en resynthétisant le composé **C4b** également donnant les diastéréoisomères séparés **C9a** et **C10a** (*tableau 14*). Afin de comprendre les faibles rendements de la réaction de Ugi, obtenus par O. BORODIN, à partir d'amines protonées sous forme de sels de TFA, nous avons utilisé la même amine (**49b**) qui est donc sous forme protonée (sel de TFA) pour les composés **C5a** et **C6a**.

**Tableau 14** : Rendements de la réaction de Ugi obtenus pour les composés de la famille de la nucléobase artificielle « S » :

	R = AA	Composé	Rendement
		Phe	<b>C5a</b>
	Ala	<b>C6a</b>	36 % <sup>a</sup>
	Lys	<b>C4a</b> (dia 1 + 2)	51 % <sup>a</sup>
		<b>C9a</b> (dia 1)	37 % <sup>b</sup>
		<b>C10a</b> (dia 2)	37 % <sup>b</sup>
	Trp	<b>C7a</b>	73 % <sup>b</sup>
	His	<b>C8a</b>	83 % <sup>b</sup>

<sup>a</sup> réaction réalisé avec l'amine sous forme de sel de TFA, <sup>b</sup> réaction réalisée avec l'amine sous forme neutre.

Lors de la réaction de Ugi des composés **C5a** et **C6a**, il a pu être identifié un même sous-produit (**C11**) (*figure 65*). Ce sous-produit provient de la réaction de l'acide trifluoroacétique qui correspond au contre-ion de l'amine. Cet acide rentrant en compétition avec l'acide aminé utilisé engendre une diminution des rendements pour les composés **C5a** (50 %) et **C6a** (36 %).



**Figure 65 :** Structure du sous-produit **C11** obtenu dans la synthèse utilisant la nucléobase « S ».

Pour s'affranchir de ce problème, la même nucléobase « S » sous forme amine a été synthétisée mais avec une amine sous forme neutre **49b** (cf. *Chap.II.2.3*), permettant ainsi d'éviter l'ajout d'une base lors de la réaction. Les intermédiaires protégés **C7a** et **C8a** ont ainsi été obtenus avec des rendements de 73 % et 83 % sous forme d'un mélange de 2 diastéréoisomères (*tableau 14*). En revanche, les diastéréoisomères protégés **C9a** et **C10a** ont pu être séparés avec des rendements de 37 % chacun.

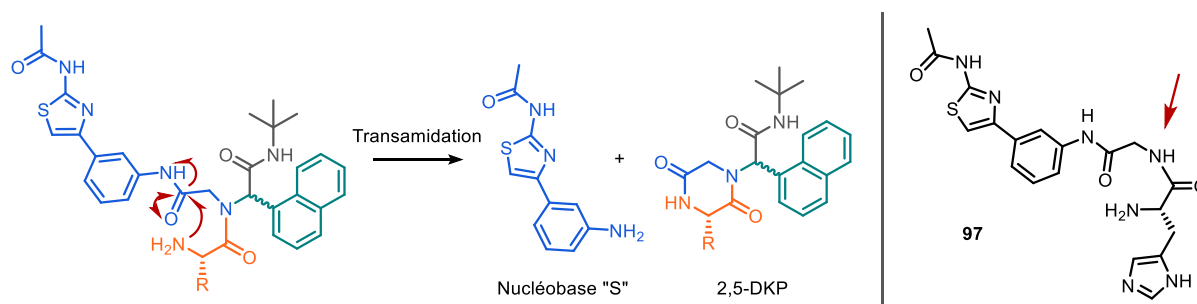
Après avoir réalisé la caractérisation de ces différents intermédiaires, l'étape suivante consiste en la déprotection de ces composés en milieu acide. Ainsi, les composés **C6b**, **C9b** et **C10b** ont été obtenus avec des rendements supérieurs à 70 % (*tableau 15*). Cependant, pour les composés **C5b** et **C7b**, une étape de purification par HPLC semi-préparative a été nécessaire pour avoir des produits purs engendrant de faibles rendements entre 34 % et 18 %.

**Tableau 15 :** Rendements de l'étape de déprotection obtenus pour les composés de la famille de la nucléobase artificielle « S » :

 nxTFA	R = AA	Composé	Rendement
		Phe	<b>C5b</b>
	Ala	<b>C6b</b>	85 %
	Lys	<b>C4b</b>	60 %
		<b>C9b</b>	71 %
		<b>C10b</b>	77 %
	Trp	<b>C7b</b>	18 %
	His	<b>C8b</b>	- %

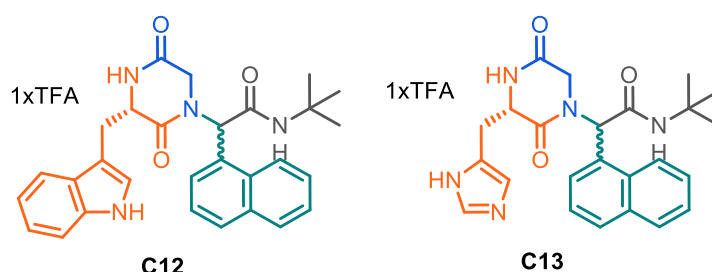
Après analyse des sous-produits obtenus lors de ces purifications pour les composés **C5b** et **C7b**, une réaction secondaire a été identifiée et elle se déroule une fois la déprotection de l'amine effectuée (*schéma 35*). Cette réaction est une transamidation intramoléculaire formant un sous-produit de structure 2,5-dicétopipérazine et libère ainsi la nucléobase « S ». Cette transamidation est probablement induite par la formation d'un cycle à 6 favorisée par la présence d'un groupement *N*-alkylé et la présence d'un groupement amine aromatique lui conférant un caractère nucléofuge. En effet, cette transamidation n'avait pas été observée pour des produits étudiés par Joly *et al.*, utilisant la même nucléobase « S » avec l'histidine par exemple (**97**)<sup>113</sup>.





**Schéma 35 :** Représentation de la transamidation obtenue dans la synthèse de la famille utilisant la nucléobase « S » après l'étape de déprotection. Structures du composé **97** étudié par Joly et al.

Ainsi, lors de la purification du composé **C7b**, le sous-produit 2,5-DKP **C12** a été obtenu avec un rendement de 30 % (figure 66). Pour la déprotection du composé histidine **C8a**, la réaction de transamidation a été observée conduisant directement aux composés 2,5-DKP **C13**. Le produit souhaité (**C8b**) n'a donc pas pu être isolé.



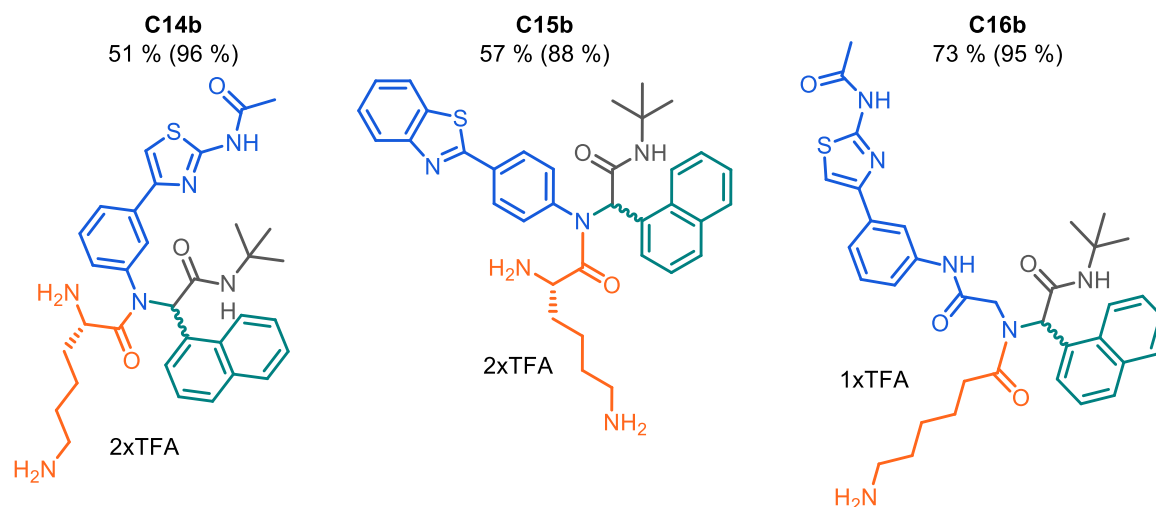
**Figure 66 :** Représentation des sous-produits 2,5-DKP **C12** et **C13**.

Ces différents résultats nous ont mené à réaliser une étude de stabilité de l'ensemble des composés de cette série (**C5b**, **C6b**, **C7b**, **C9b** et **C10b**). Ainsi, la réaction de transamidation a été observée pour chaque composé avec des cinétiques différentes pour chaque acide aminé. En effet, le composé Ala (**C6b**) présente une cinétique de dégradation plus lente que le composé avec le Trp (**C7b**) par exemple. De manière générale, par leur faible stabilité, ces 5 composés n'ont pas été étudiés pour leur activité biologique.

Outre le motif de la nucléobase « S », cette réaction de transamidation a aussi été observée lors de l'étude de stabilité du composé **C3b** (cf. tableau 13) comportant le motif 2-(4-aminophényl)benzothiazole. Par conséquent, ce composé ne sera également pas étudié.

Afin de pouvoir utiliser ces motifs pour obtenir des ligands d'ARN nous avons choisi d'utiliser directement l'amine aromatique dans la réaction de Ugi. Ainsi, le composé **C14a** comportant la nucléobase « S » et le composé **C15a** comportant le motif 2-(4-aminophényl)benzothiazole ont été synthétisés (figure 67). Un chauffage du milieu réactionnel à 50°C a été cependant nécessaire pour favoriser la solubilité des amines de départ permettant d'obtenir des rendements de 51 % et de 57 % pour les composés **C14a** et **C15a**. Pour l'étape de déprotection, grâce au fait que la réaction secondaire

ne peut plus se réaliser, des rendements de 96 % et de 88 % ont été obtenus pour les composés **C14b** et **C15b**.

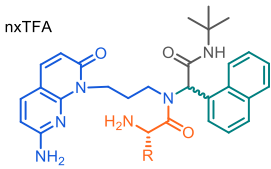
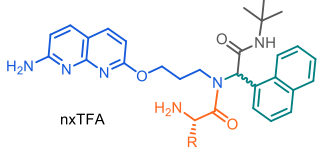


**Figure 67 :** Structures des composés **C14b**, **C15b** et **C16b**. Les pourcentages représentent le rendement de la réaction de Ugi et le rendement de l'étape de déprotection figure entre parenthèses.

En parallèle, le composé **C16b** comportant la glycine comme bras espaceur a été synthétisé mais en utilisant l'acide 6-(Boc-amino)hexanoïque qui ne possède pas d'amine en position alpha. Cet acide donne après la réaction de Ugi un mélange d'énantiomères avec un rendement de 73 % (**C16a**) et ensuite un rendement de 95 % pour l'étape de déprotection (**C16b**). Ainsi, les composés **C14b** et **C16b** seront étudiés et comparés entre eux, afin d'évaluer l'influence du motif glycine et de la présence de l'amine alpha sur l'affinité envers les ARN et sur l'activité biologique.

Nous nous sommes ensuite intéressés à l'utilisation des amines comportant les motifs naphtyridines. Dans ce contexte, nous avons synthétisé 8 composés (*tableau 16*). Les composés issus du motif *N*-alkylé (première partie du tableau, **42**) pourront ainsi être comparés avec ceux issus du motif *O*-alkylé (deuxième partie du tableau, **37**) lors de l'étude d'affinité. Ces composés ont été obtenus avec des bons à très bons rendements et la réaction de Ugi effectuée en phase liquide s'est révélée particulièrement avantageuse et aisée. Ainsi, les rendements de la réaction de Ugi sont compris entre 40 % à 90 % et l'étape de déprotection entre 69 % à 93 %.

Tableau 16 : Rendements obtenus pour les composés de la famille issue des motifs naphtyridines :

Structure	R = AA	Composé	Rendements	
			Réaction d'Ugi (a)	Étape de déprotection (b)
	Phe	<b>C17</b>	67 %	93 %
	Ala	<b>C18</b>	31 %	85 %
		<b>C19</b>	36 %	74 %
	Lys	<b>C20</b>	59 %	88 %
	Trp	<b>C21</b>	73 %	81 %
	His	<b>C22</b>	63 %	83 %
	Ala	<b>C23</b>	75 %	88 %
	Lys	<b>C24</b>	41 %	69 %
	Trp	<b>C25</b>	90 %	82 %

Enfin, pour compléter cette série permettant d'étendre la collection de ligands, les composés **C26b** et **C27b**, ainsi que le composé contenant la néomycine modifiée **C28b** ont été synthétisés (*figure 68*).

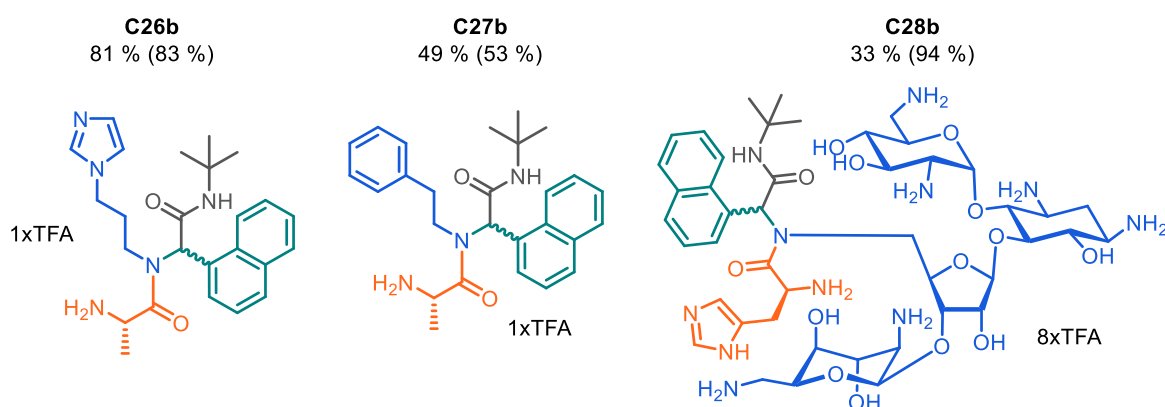


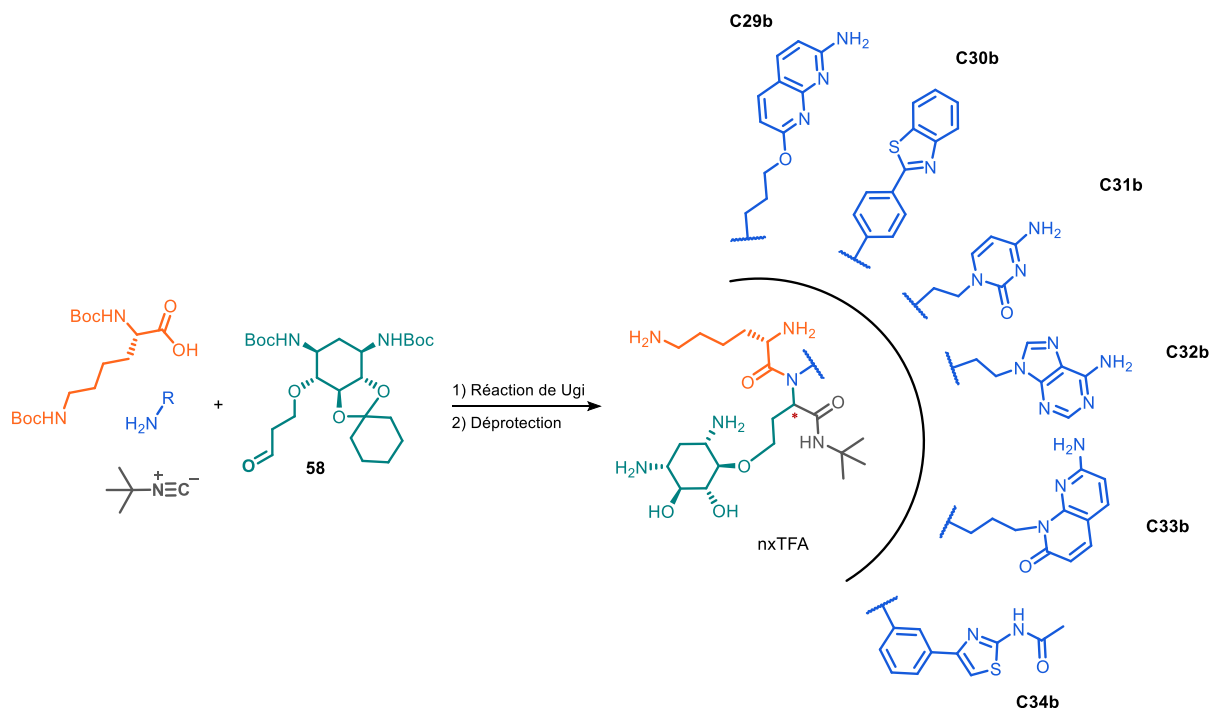
Figure 68 : Structures des composés **C26b**, **C27b** et **C28b**. Les pourcentages représentent le rendement de la réaction de Ugi et le rendement de l'étape de déprotection figure entre parenthèses.

En conclusion de cette série C-1, nous avons synthétisé une collection de 26 molécules dont 3 issues de sous-produits. Ces composés ont été caractérisés par RMN du  $^1\text{H}$  et du  $^{13}\text{C}$ . La présence de diastéréoisomères (mélanges racémiques) ainsi que de conformères (cis et trans autour de la liaison amide) ont rendu l'analyse et l'attribution des signaux compliquées engendrant une difficulté à déterminer la pureté des produits obtenus. Pour cela, des analyses HPLC ont été réalisées afin de vérifier cette pureté.

Les molécules de cette série ont été étudiées pour déterminer leur affinité envers les ARN cibles sélectionnés. Cependant, en raison de leur instabilité, l'affinité des composés **C3b**, **C5b**, **C6b**, **C7b**, **C9b** et **C10b** n'a pas pu être évaluée.

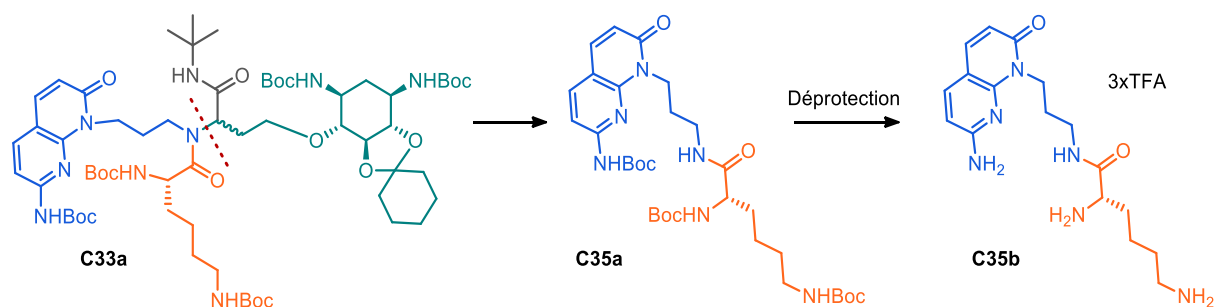
## II.4.2 Synthèses de la série C-2 : série 2-désoxystreptamine

À la suite de la synthèse de la 2-désoxystreptamine modifiée avec une fonction aldéhyde (**58**), une série appelée « 2-DOS » a été préparée et est décrite dans cette partie. Afin, de faire varier les fragments possédant la fonction amine, la lysine a été utilisée comme acide aminé (*schéma 36*).



*Schéma 36* : Description de la synthèse permettant d'obtenir les composés de la série C-2. Conditions : **1)** MeOH, t.a., 24h à 72h ; **2)** TFA, CH<sub>2</sub>Cl<sub>2</sub>, t.a., 4h.

Dans cette série, la réaction de Ugi n'a pas toujours permis d'obtenir des produits purs. Plusieurs purifications particulièrement difficiles ont été nécessaires faisant diminuer les rendements de la première étape, notamment pour les intermédiaires des composés **C29a**, **C31a**, **C33a** (*tableau 17, page suivante*). De plus, certains de ces composés se sont dégradés lors de la purification ou lorsque le temps de réaction de Ugi était trop long (5 jours). Cette dégradation conduit à une structure ayant une liaison peptidique entre la lysine et l'amine utilisée comme pour le composé **C35b** déterminé lors de la synthèse du composé **C33a** (*schéma 37*). Ce sous-produit a aussi été observé pour les composés **C29a** et **C34a**.



**Schéma 37** : Formation du sous-produit **C35b** obtenu lors de la purification du composé **C33a**. Le composé **C35b** est isolé après déprotection. Conditions de la déprotection : TFA, CH<sub>2</sub>Cl<sub>2</sub>, t.a., 4h.

À la suite de ces purifications, les rendements obtenus de la réaction de Ugi sont compris entre 23 % et 76% (*tableau 17*).

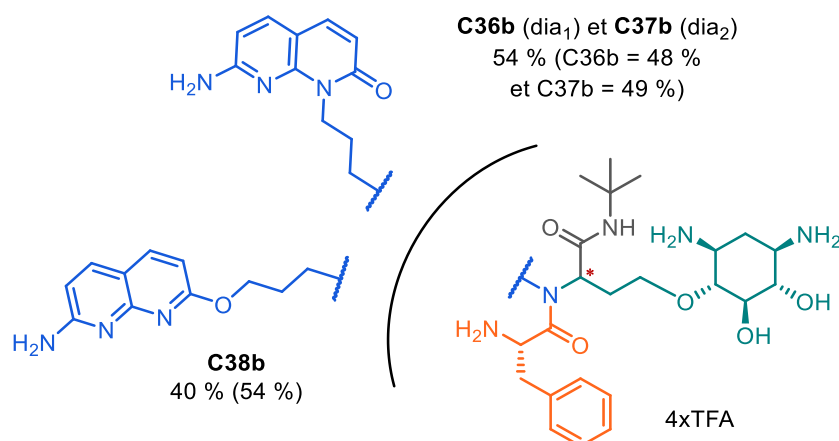
**Tableau 17** : Rendements obtenus pour les composés de la famille « 2-DOS » :

Composé	Rendements	
	Réaction d'Ugi (a)	Étape de déprotection (b)
<b>C29</b>	40 %	97 %
<b>C30</b>	76 %	97 %
<b>C31</b>	38 %	81 %
<b>C32</b>	60 %	54 %
<b>C33</b>	23 %	52 %
<b>C34</b>	65 %	47 %

Lors de la synthèse du composé **C33b**, la séparation du sous-produit **C35b** a été possible seulement lors de la purification après l'étape de déprotection. Cette déprotection en milieu acide a permis d'éliminer les groupement Boc et le groupement acétal. Ainsi, le composé **C33b** (52 % de rendement) a pu être séparé du sous-produit **C35b**. Ce fragment **C35b** a été évalué en tant que ligand d'ARN pour être comparé à la molécule **C33b**.

Pour l'étape de déprotection des autres composés, les rendements sont compris entre 47 % et 97 % avec une purification par HPLC semi-préparative pour les composés **C31b**, **C32b** et **C34b**. L'étude de stabilité (à pH =7.2) de ces différents produits n'a montré cependant aucune dégradation.

Enfin, dans cette série, deux molécules ont été synthétisées avec la phénylalanine. Il s'agit des composés **C38b** et des diastéréoisomères **C36b** et **C37b** qui ont pu être séparés et qui ont également nécessité une purification à la suite de l'étape de déprotection (*figure 69*).



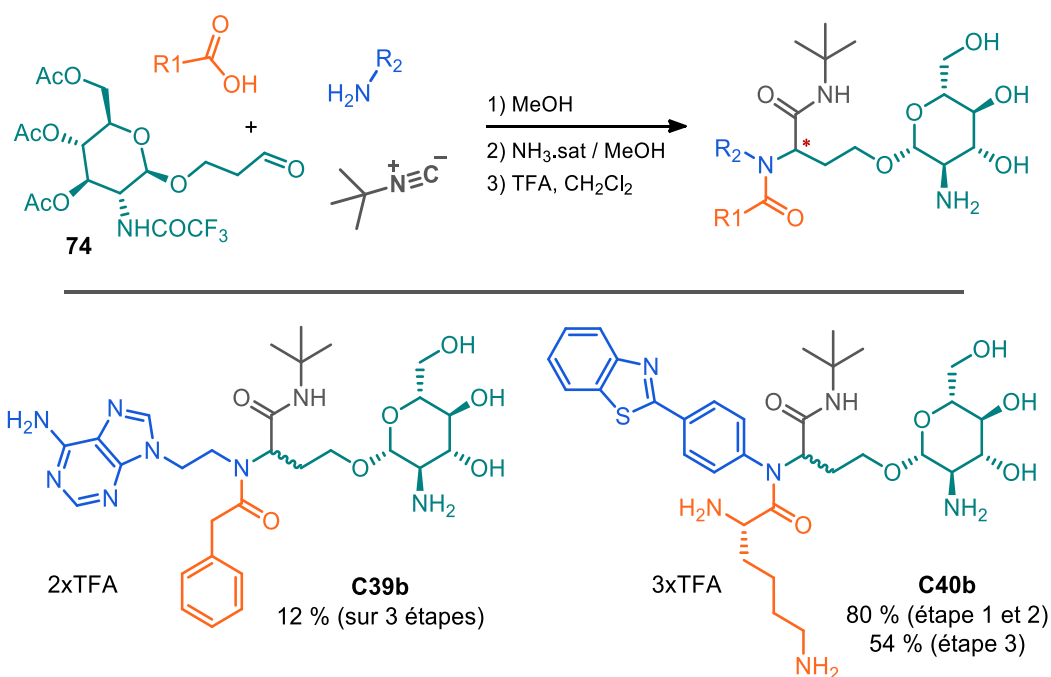
**Figure 69** : Structures des composés **C36b**, **C37b** et **C38b**. Les pourcentages représentent le rendement de la réaction de Ugi et le rendement de l'étape de déprotection figure entre parenthèses.

En conclusion sur cette série C-2 « 2-DOS », 9 composés dont deux diastéréoisomères séparés ont été synthétisés. Le sous-produit **C35b** a également été isolé. Par la suite, l'affinité de ces **10 composés** envers les ARNs cibles a été étudiée.

### II.4.3 Synthèses de la série C-3 : ensemble de sous-familles

Dans cette dernière série C-3, plusieurs motifs différents ont été utilisés et les composés obtenus peuvent être regroupés par sous-famille.

Dans une première sous-famille, des composés contenant la glucosamine modifiée **74** ont été synthétisés **C39b** et **C40b** (schéma **38**). Pour préparer ces deux composés, une étape post-Ugi a dû être ajoutée afin d'éliminer les groupements acétyle et le groupement trifluoroacétamide présents sur le précurseur **74**. L'élimination de ces groupements a été réalisée en présence d'une solution d'ammoniaque saturée dans le méthanol. Dans une dernière étape, l'élimination des groupements Boc a été réalisée avec de l'acide trifluoroacétique dans le CH<sub>2</sub>Cl<sub>2</sub> pour fournir les composés **C39b** et **C40b**.

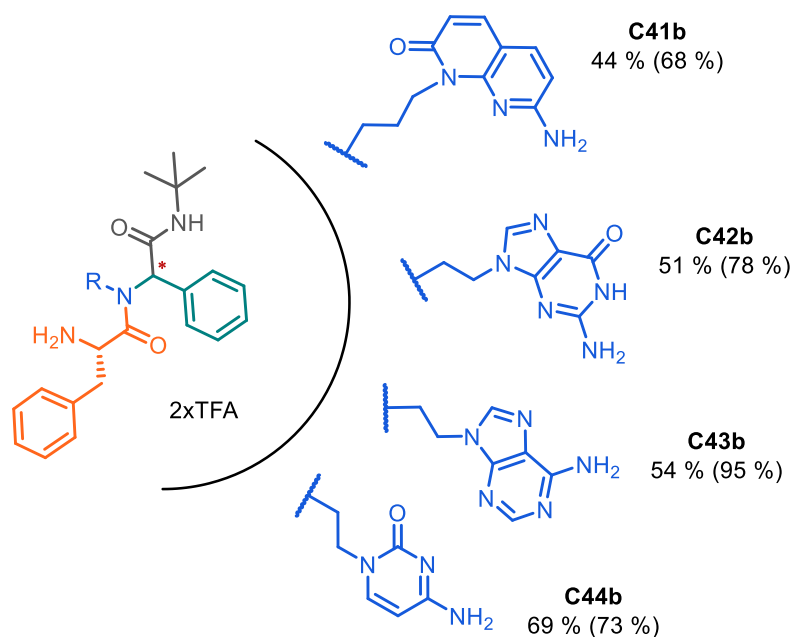


*Schéma 38 : Description de la synthèse permettant d'obtenir les composés C39b et C40b.*

En ce qui concerne le composé **C39b**, les trois étapes ont été réalisées à la suite sans purifications intermédiaires. Une purification finale a donc été réalisée après déprotection des groupements protecteurs pour donner le composé souhaité avec un rendement total de 12 % sur les trois étapes. Lors de cette réaction de Ugi, un sous-produit a été identifié, issu d'une coupure de la liaison glycosidique. Cette hydrolyse a fortement diminué le rendement global de la réaction de Ugi.

Le deuxième composé (**C40b**) de cette sous-famille a été obtenu en trois étapes. Le composé **C40a** a été isolé après purification avec un rendement de 80 % sur deux étapes (réaction de Ugi et déprotection en milieu basique). Dans cette réaction de Ugi, aucun sous-produit issu de l'hydrolyse n'a été observé. Enfin, l'étape de déprotection finale (étape 3) a été réalisée et a permis d'obtenir le composé **C40b** avec un rendement de 54 % après purification.

Dans une autre sous-famille (*figure 70*), une réaction de Ugi entre le benzaldéhyde, la phénylalanine et quatre fragments amines préalablement synthétisés a été réalisée.



**Figure 70 :** Structures des composés **C41b**, **C42b**, **C43b** et **C44b**. Les pourcentages représentent le rendement de la réaction de Ugi et le rendement de l'étape de déprotection figure entre parenthèses.

L'ensemble de ces 4 composés a été obtenu avec des rendements satisfaisants compris entre 44 % et 69 % pour l'étape de la réaction de Ugi et entre 68 % et 95 % pour l'étape de déprotection.

Enfin dans une dernière sous-famille, une série de composés comportant l'aldéhyde thiazole et le fragment naphtyridine *N*-alkylé ainsi que différents acides aminés a été synthétisée (*tableau 18*). Cette série devrait permettre d'établir une comparaison avec les molécules de la série naphthalène lors de l'évaluation biologique.

**Tableau 18 :** Rendements obtenus pour les composés de la famille thiazole :

R = AA	Composé	Rendements	
		Réaction d'Ugi (a)	Étape de déprotection (b)
Phe	<b>C45</b>	78 %	86 %
Ala	<b>C46</b>	82 %	60 %
Lys	<b>C47</b>	69 %	81 %
Trp	<b>C48</b>	83 %	87 %
His	<b>C49</b>	61 %	95 %
Adénine-Imidazole-Lys	<b>C50</b>	54 %	72 %

Ces composés ont été obtenus sans difficultés avec des rendements très satisfaisants allant de 61 % à 83 % pour l'étape de la réaction de Ugi et allant de 60 % à 95 % pour l'étape de déprotection. Enfin, la dernière molécule de ce tableau est le composé comportant les fragments adénine/imidazole/Lys (**C50b**). Ce composé a été obtenu avec des rendements de 54 % pour la réaction de Ugi et de 72 % pour l'étape de déprotection.



#### II.4.4 Conclusion des composés obtenus par une réaction de Ugi en solution

En conclusion de cette partie, une librairie de **48 molécules** constituées de deux diastéréoisomères (en proportion racémique) a été conçue dont 4 sous forme de diastéréoisomères séparés. Cette librairie a été synthétisée pour interagir avec la cible ARN et inhiber la formation du complexe boucle-boucle du système Toxine-Antitoxine *AapA1/IsoA1*. L'ensemble des structures sont résumées dans l'*annexe 3*. Dans une première série regroupant un ensemble de **26 molécules**, le motif naphthalène a été utilisé donnant la **série C-1**. Pour la **série C-2** comportant le motif 2-désoxystreptamine, **10 molécules** ont été obtenues. Enfin, dans une série plus variée, la série **C-3**, **12 molécules** ont été préparées. Lors de ces synthèses, j'ai pu avoir l'opportunité d'encadrer un stagiaire de Master 2, M. PANOSSETTI, qui a participé à la synthèse de 4 molécules finales.

Lors de ces différentes synthèses, plusieurs sous-produits ont été identifiés notamment pour les séries C-1 et C-2. En effet, certaines de ces molécules présentent une instabilité notamment celles avec le motif de la nucléobase « S » et le motif 2-(4-aminophényl)benzothiazole comportant une glycine comme bras espaceur. Ainsi, ces 6 composés (**C3b**, **C5b**, **C6b**, **C7b**, **C9b** et **C10b**) n'étaient pas suffisamment stables pour pouvoir effectuer leur évaluation biologique.

Afin d'étudier l'affinité de ces composés pour le complexe boucle-boucle décrit dans ce projet, plusieurs tests biochimiques ont été réalisés pour déterminer leur sélectivité, leur capacité à inhiber la formation du complexe et ainsi d'activer le système Toxine-Antitoxine *in vitro* et dans les bactéries. L'ensemble de ces études sont décrites dans la section suivante.

## II.5 Evaluations biochimiques

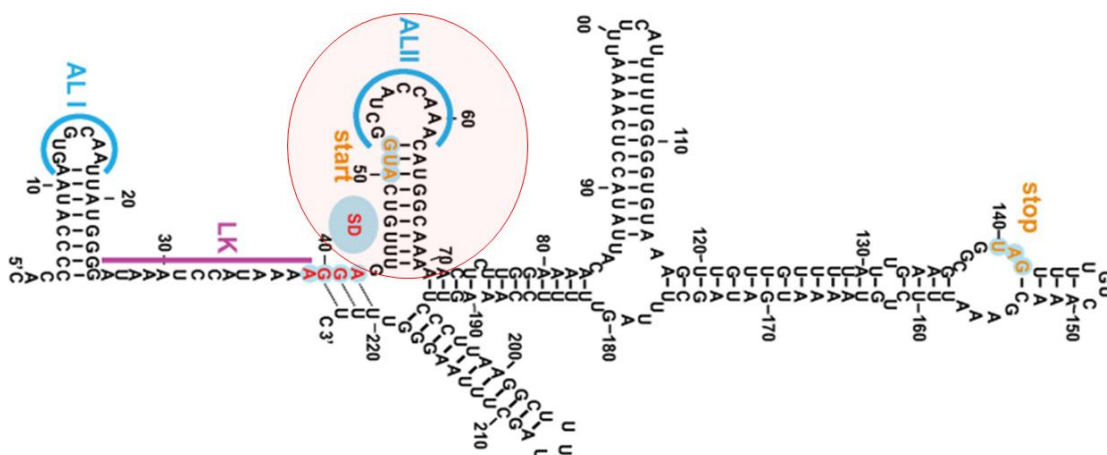
Comme mentionné auparavant, le but de ce projet de recherche est d'identifier des nouveaux ligands d'ARN susceptibles d'inhiber la formation du complexe Toxine-Antitoxine de type I constitué d'une interaction ARN/ARN de type boucle-boucle. La formation de ce que l'on appelle « kissing complex » conduit par la suite à un duplex étendu ce qui inhibe l'expression de la toxine en condition de croissance normale pour la bactérie. Les ligands synthétisés au cours du travail de thèse ont été conçus afin d'inhiber l'interaction boucle-boucle et induire la synthèse de la toxine létale pour la bactérie.

Dans ce contexte, nous nous sommes focalisés sur le système *AapA/IsoA* présent chez la bactérie *Helicobacter pylori* et notamment sur le système de *AapA1/IsoA1* (5 autres systèmes *AapA/IsoA* sont en effet présents dans le génome de la bactérie). Une fois les composés synthétisés, nous avons donc choisi d'étudier *in vitro* leur capacité (i) à se fixer sur les séquences cibles de manière affine et sélective et (ii) à inhiber la formation du duplex étendu. L'ensemble des tests effectués dans ce cadre sont décrits dans la suite de ce chapitre et sont basés sur l'utilisation de fragments d'ARN fluorescents.

Les composés synthétisés ont aussi été évalués *in vivo* pour leur capacité à inhiber la formation du duplex étendu à travers un test cellulaire basé sur l'utilisation d'un système rapporteur développé par l'équipe du Dr. Darfeuille. Lors de ce projet, j'ai eu l'opportunité de participer à l'évaluation de certains composés grâce à un stage effectué au Laboratoire ARNA à Bordeaux. Enfin, une évaluation biologique des composés sur différentes souches bactériennes a été réalisée par un prestataire et sera présentée à la fin de ce chapitre.

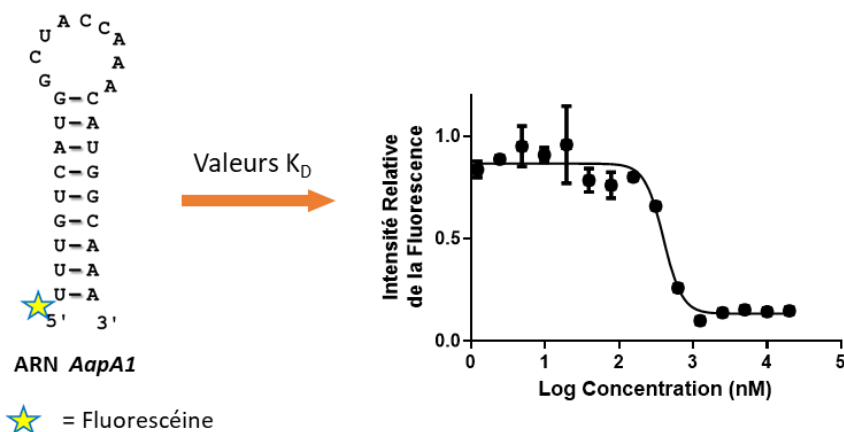
### II.5.1 Évaluation de l'affinité des composés sur *AapA1*

Le système Toxine-Antitoxine *AapA1/IsoA1* étudié comporte deux ARN. Le complexe boucle-boucle est induit par deux tige-boucles présentes sur chacun des deux brins d'ARN dont la boucle ALI et ALII pour l'ARN *AapA1* (figure 71). Dans un premier temps, nous avons décidé d'étudier la capacité des molécules synthétisées à interagir avec ALII (entouré en rouge). Nous avons alors effectué des études en utilisant une séquence marquée par un fluorophore ce qui nous a permis de mesurer des constantes de dissociation ( $K_D$ ).



**Figure 71 :** Structures primaire et secondaire de l'ARN AapA1. Les boucles ALI et ALII interviennent dans la reconnaissance avec l'ARN IsoA1 formant des interactions boucle-boucles d'abord et un duplex étendu ensuite. La partie entourée en rouge représente la séquence étudiée dans ce travail.

Dans un premier temps, pour réaliser ces tests biochimiques et ainsi quantifier l'affinité de ces molécules, la séquence sélectionnée (boucle ALII de AapA1) a été choisie. Cette séquence de 27 nucléobases a été préalablement marquée en position 5' par une fluorescéine (FAM) (figure 72).



**Figure 72 :** Représentation de la séquence d'AapA1 étudié pour évaluer l'affinité des composés synthétisés.

La mesure de la fluorescence en présence de concentrations croissantes d'un ligand permet de mesurer les valeurs de  $K_D$  et ainsi quantifier l'interaction de la molécule étudiée avec l'ARN<sup>209</sup>. En effet, la liaison d'une molécule à l'ARN induit un changement de conformation qui influence l'environnement de la sonde elle-même (même si le site de liaison est éloigné). Ce changement de conformation entraîne l'extinction de la fluorescence du fluorophore. Ainsi, la variation de fluorescence en fonction de la concentration du composé étudié peut être mesurée. Les courbes obtenues sont ensuite exploitées par une régression linéaire permettant d'obtenir la valeur de  $K_D$ .


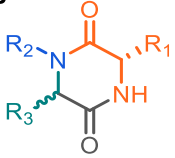
Les mesures de  $K_D$  ont été réalisées sur toutes les séries synthétisées dans ce manuscrit et sont décrites dans cette partie. Afin de vérifier la répétabilité des résultats et d'obtenir une valeur moyenne pour chaque composé, les mesures ont été effectuées au cours de trois manipulations indépendantes, chacune effectuée en duplicata.

Enfin, la sélectivité des composés présentant les meilleures  $K_D$  envers l'ARN *AapA1* a été évaluée. Pour cela, les constantes d'affinité ont été mesurées en présence d'autres acides nucléiques présents en grande quantité dans les cellules : l'ARNt et l'ADN. Ainsi, l'affinité des composés a été mesurée en présence d'un excès (100 équiv.) d'ARNt ( $K'_D$ ) puis un excès (100 équiv.) d'ADN ( $K''_D$ ). Les valeurs obtenues ont été comparées aux valeurs de constantes de dissociation ( $K_D$ ) du composé en présence de l'ARN *AapA1* afin d'en déduire la sélectivité.

### II.5.1.1 Évaluation de l'affinité des composés issus de la synthèse sur phase solide : Séries A et B

Dans ce projet, deux séries A et B ont été obtenues qui diffèrent selon l'acide carboxylique utilisé dans la réaction de Ugi, mais également selon le groupement amine  $R_2$  employé (*pour les structures voir l'annexe 3*). Ces deux séries ont conduit à des composés linéaires et des composés cycliques sous forme de 2,5-DKP. Ces composés sont rassemblés dans le *tableau 19*. Pour rappel, les numéros indiqués font référence soit à des composés ayant été purifiés soit à des composés présents sous forme de mélanges de plusieurs produits tels que les mélanges **B1**, **B4** et **B5**. Les constantes d'affinité ont été évaluées et sont décrites dans le *tableau 19*.

**Tableau 19 :** Valeurs des constantes de dissociation ( $K_D$ ) des composés des *séries A et B* issue de la synthèse sur support solide pour l'ARN *AapA1*.

Linéaire					2,5-DKP					
			$R_4$ = Acides aminés ou acides carboxyliques						$R_1$ = Chaînes latérales d'acides aminés	
Entrée	N°	Structure	Pureté <sup>a</sup>	$K_D$ ( $\mu$ M) <sup>b</sup>	Entrée	N°	Structure	Pureté <sup>a</sup>	$K_D$ ( $\mu$ M) <sup>b</sup>	
1	<b>15</b>	Linéaire	100 %	<b>22,6</b>	10	<b>A1</b>	DKP	95 %	<b>19,1</b>	
2	<b>79</b>	Linéaire	95 %	> 1mM	11	<b>A2</b>	DKP	100 %	nd	
3	<b>80</b>	Linéaire	85 %	> 1mM	12	<b>A3</b>	DKP	100 %	> 50	
4	<b>81</b>	Linéaire	95 %	<b>5,76</b>	13	<b>A5</b>	DKP	100 %	> 40	
5	<b>82</b>	Linéaire	91 %	>85	14	<b>A6</b>	DKP	100 %	nd	
6	<b>B1</b>	Linéaire	86 %	> 1mM	15	<b>A8</b>	DKP	82 %	<b>13,3</b>	
7	<b>B4</b>	Linéaire	96 %	>275	16	<b>A9</b>	DKP	81 %	> 1mM	
8	<b>B5</b>	Linéaire	77 %	> 1mM	17	<b>A10</b>	DKP	80 %	> 1mM	
9	<b>B2a</b>	Linéaire	100 %	nd	18	<b>A11a</b>	DKP	100 %	> 1mM	
					19	<b>A11b</b>	DKP	100 %	> 1mM	

<sup>a</sup> Pureté mesurée par HPLC à UV = 214 nm. <sup>b</sup> Ces manipulations n'ont été effectuées qu'une seule fois et dans ce cas l'erreur de la mesure est  $\pm 10\%$ . **nd** : Non déterminé.

Dans la première partie du tableau (**entrées 1 à 5**), les composés sont issus des tests préliminaires et comportent des groupements aromatiques et des fonctions amines protonables. En revanche, les composés linéaires de la série B (**entrées 6 à 9**) ne contiennent que des motifs aromatiques.

Parmi les 9 composés, seuls deux montrent une certaine affinité : le composé **15** (**entrée 1**) et le composé **81** (**entrée 4**) avec un  $K_D$  respectivement de 22,6  $\mu\text{M}$  et de 5,76  $\mu\text{M}$ . Ces premiers résultats semblent indiquer que la présence de fonctions amines non aromatiques soit nécessaire pour obtenir une affinité. Le meilleur composé de cette série linéaire, c'est-à-dire le composé **81**, contient d'ailleurs un motif lysine avec deux groupements amines (chaîne latérale et amine alpha) et un motif naphthalène. Le fait que cette combinaison de motifs lysine/naphthalène engendre une affinité est un résultat intéressant puisque ces motifs sont repris dans certains composés de la série C-1.

Pour les autres composés de forme linéaire ne comportant pas d'amines non aromatiques (**entrées 6 à 9**), de faibles réponses sont obtenues pour certains d'entre eux et d'autres ne présentent aucune affinité jusqu'à une concentration de 1 mM. Ces résultats pourraient s'expliquer par leur faible solubilité lors des tests réalisés. De plus, ces composés possèdent peu de motifs pouvant réaliser des interactions efficaces avec la cible.

En ce qui concerne les structures cycliques (2,5-DKP), on note aussi des différences importantes d'affinité. Deux composés présentent une affinité prometteuse. Il s'agit du composé **A1** (**entrée 10**) qui contient 2 motifs lysines et un motif tryptophane avec un  $K_D$  de 19,1  $\mu\text{M}$  et le composé **A8** (**entrée 15**) qui contient aussi un motif tryptophane associé à deux motifs aromatiques : le naphtyridine *N*-alkylé et le naphthalène. Cette combinaison tryptophane/naphthalène conduit à une amélioration de l'affinité avec une valeur de  $K_D$  de 13,3  $\mu\text{M}$ .

Le motif naphtyridine *N*-alkylé se retrouve aussi dans les composés des **entrées 12 à 17**, certains d'entre eux présentent d'ailleurs une petite affinité comme **A3** et **A5**. Par ailleurs, lorsque le motif naphtyridine est remplacé par une adénine (composé **A10**, **A11a** (dia1) et **A11b** (dia2)), l'affinité mesurée est extrêmement faible voire absente. La comparaison entre le composé **A10** (Adénine) comportant les motifs tryptophane/naphthalène couplés à l'adénine et le composé **A8** (naphtyridine *N*-alkylé) comportant les mêmes motifs couplés au motif naphtyridine montre que le naphtyridine est beaucoup plus favorable pour l'affinité que l'adénine.

Globalement, l'affinité des composés appartenant aux séries A et B n'est pas très satisfaisante excepté pour les composés **15**, **81**, **A1** et **A8** qui possèdent des valeurs de  $K_D$  du même ordre de grandeur c'est-à-dire dans le bas micromolaire. Il faut noter que le meilleur composé **A8** ( $K_D$  de 13,3  $\mu\text{M}$ ) est un mélange avec une pureté de 80%. Son activité pourrait être meilleure une fois le composé purifié. Afin de confirmer cette hypothèse, il serait intéressant dans un deuxième temps d'évaluer le composé **A8**

avec une pureté supérieure ou égale à 95 %. Enfin, ce composé comporte le motif naphtyridine *N*-alkylé et il serait intéressant de pouvoir comparer ce même composé comportant le motif *O*-alkylé plus favorable aux interactions. En effet, la position différente du substituant de ces deux motifs doit probablement entraîner des liaisons différentes sur l'ARN et donc une affinité également différente.

### II.5.1.2 Évaluation de l'affinité des composés issus de la synthèse en solution : Série C

Comme décrit précédemment, différentes séries ont été préparées en employant la réaction de Ugi en solution en présence du *tert*-butyl isocyanure (**Série C-1, C-2 et C-3**). Pour chacune de ces séries, la constante de dissociation a été étudiée sur l'ARN *AapA1* et les résultats obtenus sont décrits dans la suite de cette partie.

#### II.5.1.2.1 Évaluation de l'affinité des composés de la série C-1

Dans le cas des composés de la **série C-1**, le motif central choisi est l'aldéhyde représenté par le 1-naphtaldéhyde. Ce motif apporte des propriétés intercalantes aux molécules synthétisées. L'affinité a été mesurée pour les **19 molécules** de cette série.

Le **tableau 20** illustre les valeurs de  $K_D$  obtenues lorsque le naphthalène a été couplé à l'amine thymine modifiée (**34 et 35**), 1-(3-aminopropyl)imidazole, 2-phényléthylamine et néomycine modifiée (**57**). Le couplage du naphthalène avec la thymine et la lysine (**C1b**) conduit à une bonne affinité avec un  $K_D$  de 16,9  $\mu$ M. Lorsque la lysine est remplacée par l'alanine (**C2b**) l'affinité diminue ( $K_D > 20\mu$ M) ce qui suggère que la chaîne latérale de l'acide aminé est importante pour l'interaction. Lorsque le naphthalène est couplé au motif propyl-1*H*-imidazole et à l'alanine (**C26b**) le  $K_D$  monte à 39,4  $\mu$ M indiquant que le motif portant l'amine est aussi fortement lié à l'affinité et que l'ensemble des motifs agissent de manière coopérative pour interagir avec l'ARN cible.

**Tableau 20 :** Valeurs des constantes de dissociation ( $K_D$ ) des composés **C1b**, **C2b**, **C26b**, **C27b** et **C28b** pour l'ARN *AapA1* :

Composé	Structure : Amine/Aldéhyde/AA	$K_D$ ( $\mu$ M)
<b>C1b</b>	Thymine( <i>N</i> <sub>1</sub> ) -Naphtalène-Lys	16,9 <sup>a</sup>
<b>C2b</b>	Thymine( <i>N</i> <sub>3</sub> ) -Naphtalène-Ala	>20 <sup>a</sup>
<b>C26b</b>	Propyl-1 <i>H</i> -Imidazole -Naphtalène-Ala	39,4 <sup>a</sup>
<b>C27b</b>	Ethylbenzène-Naphtalène-Ala	1,25 <sup>a</sup>
<b>C28b</b>	Néomycine-Naphtalène-His	<b>0,0435 ± 0,021</b>
	Néomycine B	2,00 <sup>a</sup>

<sup>a</sup> Ces manipulations n'ont été effectuées qu'une seule fois et dans ce cas l'erreur de la mesure est  $\pm 10\%$ .

La comparaison de **C2b** et **C26b** avec un composé portant un motif éthylbenzène et l'alanine couplés au naphthalène (**C27b**) est en revanche surprenante puisque ce dernier composé présente un  $K_D$  de 1,25  $\mu\text{M}$  qui est le meilleur observé. Ceci suggère que non seulement les motifs doivent être choisis judicieusement pour former les interactions les plus favorables, mais aussi que leur disposition tridimensionnelle et l'encombrement du composé influent fortement sur l'affinité.

L'analogie contenant la néomycine couplée au naphthalène et à l'histidine (**C28b**) montre quant à lui un  $K_D$  de 42,5 nM deux ordres de grandeur plus petit par rapport au meilleur composé **C27b**. Non seulement l'ajout de néomycine, connue pour sa forte affinité envers plusieurs motifs d'ARN, a amélioré l'affinité des conjugués, mais sa conjugaison avec les motifs naphthalène et histidine a permis d'augmenter l'affinité de la néomycine seule qui est de l'ordre du micromolaire ( $K_D = 2,00 \mu\text{M}$ ).

Le *tableau 21* illustre les valeurs de  $K_D$  mesurées pour les molécules comportant la nucléobase S comme fragment amine, ainsi que les sous-produits obtenus lors de la synthèse de cette sous-famille. Comme mentionné précédemment, les composés **C3b**, **C5b**, **C6b**, **C7b**, **C9b** et **C10b** n'étaient pas suffisamment stables pour effectuer les études concernant leur activité biologique.

**Tableau 21 :** Valeurs des constantes de dissociation ( $K_D$ ) des composés issus de la synthèse avec la nucléobase artificielle « S » pour l'ARN AapA1 :

Composé	Structure : Amine/Aldéhyde/AA	$K_D$
<b>C11</b>	S(Gly)-Naphthalène-TFA	>1 mM <sup>a</sup>
<b>C12</b>	2,5-DKP : Gly -Naphthalène-Trp	>300 $\mu\text{M}$ <sup>a</sup>
<b>C13</b>	2,5-DKP : Gly -Naphthalène-His	36,6 $\mu\text{M}$ <sup>a</sup>
<b>C14b</b>	S-Naphthalène-Lys	<b>120 <math>\pm</math> 1,4 nM</b>
<b>C16b</b>	S(Gly)-Naphthalène- $\epsilon$ -Ahx	517 $\pm$ 145 nM
<b>C15b</b>	2-(4-aminophényl)benzothiazole-Naphthalène-Lys	<b>152 <math>\pm</math> 83 nM</b>

<sup>a</sup> Ces manipulations n'ont été effectuées qu'une seule fois et dans ce cas l'erreur de la mesure est  $\pm 10\%$ .

Ces résultats montrent que les sous-produits de structure 2,5-DKP obtenus qui ne comportent pas le motif « S » (composés **C12** et **C13**) présentent une faible affinité pour l'ARN cible. De la même manière, le sous-produit **C11** dans lequel le TFA a pris le rôle d'acide carboxylique au cours de la réaction de Ugi présente une très faible affinité. En revanche, les composés **C14b** et **C16b** ont montré une affinité très similaire avec des valeurs de  $K_D$  de 120 nM et de 517 nM. Ces deux composés ne comportent que très peu de différences structurales puisque le composé **C16b** comprend un bras de liaison de type glycine sur la nucléobase et l'amine  $\alpha$  de l'acide aminé lysine est absente. Une légère perte d'affinité pour ce composé comparé au composé **C14b** suggère que l'amine  $\alpha$  joue un rôle dans l'interaction avec la cible alors que le bras de liaison additionnel n'est pas nécessaire pour l'affinité.

Le composé **C15b** comporte le motif 2-(4-aminophényl)benzothiazole et a été synthétisé pour pouvoir comparer son affinité et activité avec le composé **C14b**. **C15b** possède une valeur de  $K_D$  de 152 nM comparable à celle du composé **C14b** ce qui illustre que le motif phényl/thiazole ou une substitution en position 4 du phényle par un motif benzothiazole sont favorables pour l'affinité (figure 73). Il est à noter que le composé **81** issu de la synthèse sur support solide (phényle/naphtalène/Lys) et ne comportant pas de motif phényl/thiazole ou phényl/benzothiazole présente une valeur de  $K_D$  de 5,76  $\mu$ M. Ainsi, la structure centrale de ces trois composés est suffisante pour interagir avec cet ARN mais le motif hétéroaromatique améliore de plus de 50 fois l'affinité.

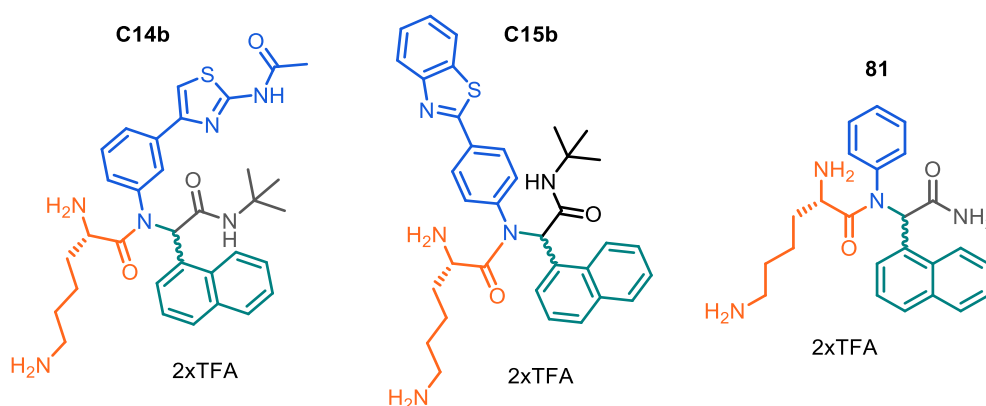


Figure 73 : Représentation des structures des composés **C14b**, **C15b** et **81**.

En conclusion, la nucléobase S et le motif naphthalène conjugués à des acides aminés basiques sont particulièrement favorables pour l'affinité et permettent d'atteindre des  $K_D$  de l'ordre du nanomolaire.

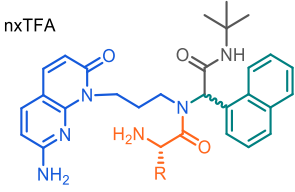
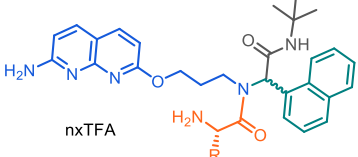
Une partie des composés appartenant à la série C-1 inclut des conjugués comportant le motif naphtyridine *N*- ou *O*-alkylés en présence de différents acides aminés. Les constantes d'affinités sont représentées dans le *tableau 22*.

L'analyse des résultats obtenus montre que l'acide aminé influe fortement sur l'affinité. En effet, les composés portant une lysine (**C20b** et **C24b**) possèdent la meilleure affinité pour *AapA1* qui est de 0,62  $\mu$ M et de 0,29  $\mu$ M. Comme nous pouvons le prévoir, le motif naphtyridine *O*-alkylé est donc plus favorable pour l'interaction par rapport au motif *N*-alkylé. Ensuite les composés possédant les valeurs de  $K_D$  plus favorables suivent l'ordre des acides aminés Trp>His>Ala>>Phe lorsque le motif naphtyridine *N*-alkylé est utilisé. Par ailleurs, la phénylalanine ne possédant pas d'hétéroatomes sur son cycle aromatique en comparaison avec le Trp et l'His, fait diminuer drastiquement l'affinité. Cet ordre d'affinité est également retrouvé pour les 3 composés contenant le motif naphtyridine *O*-alkylé même si dans ce cas les valeurs de  $K_D$  sont très similaires et en dessous du micromolaire. Enfin, il est intéressant de noter que la comparaison des deux diastéréoisomères des composés **C18b** et **C19b** montre une différence d'affinité en fonction de la stéréochimie. Ceci suggère que non seulement les



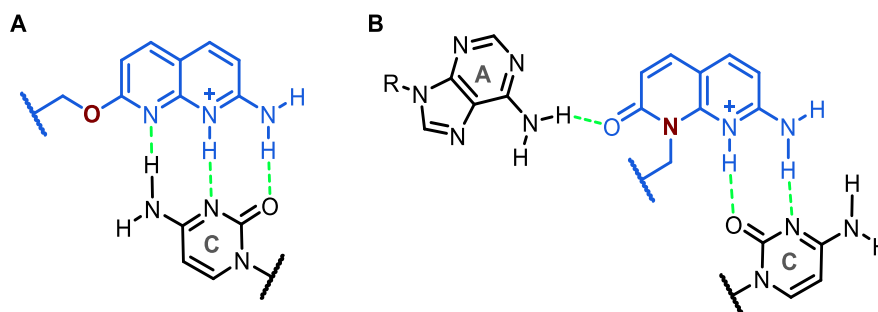
substituants présents sur le composé sont importants, mais que leur distribution dans l'espace affecte aussi l'interaction avec la cible.

**Tableau 22 :** Valeurs des constantes de dissociation ( $K_D$ ) des composés issus de la synthèse avec les motifs naphtyridines *N*- et *O*-alkylés pour l'ARN *AapA1* :

Structure	Composé	R = AA	$K_D$ ( $\mu\text{M}$ )
	<b>C17b</b>	Phe	>70 <sup>a</sup>
	<b>C18b</b>	Ala	4,06 ± 0,20
	<b>C19b</b>		15,3 <sup>a</sup>
	<b>C20b</b>	Lys	<b>0,62 ± 0,21</b>
	<b>C21b</b>	Trp	1,26 ± 0,08
	<b>C22b</b>	His	1,88 ± 0,18
	<b>C23b</b>	Ala	0,83 ± 0,20
	<b>C24b</b>	Lys	<b>0,29 ± 0,11</b>
	<b>C25b</b>	Trp	0,43 ± 0,05

<sup>a</sup> Ces manipulations n'ont été effectuées qu'une seule fois et dans ce cas l'erreur de la mesure est ±10%.

Une comparaison entre la position *N*- et *O*-alkylée, montre que les composés ayant le motif *O*-alkylé forment des interactions plus fortes avec *AapA1*. Effectivement, le motif *O*-alkylé devrait permettre d'obtenir 3 liaisons sur les motifs cytosines ou guanines présents dans la boucle de l'ARN *AapA1* (figure 74). Ces trois interactions induisent une amélioration de l'affinité qui est de 2 à 4 fois plus élevée qu'avec le motif *N*-alkylé. En effet, le motif *N*-alkylé peut quant à lui interagir avec deux liaisons hydrogènes avec un résidu cytosine par exemple ce qui expliquerait la plus faible affinité. Cependant, ce motif pourrait interagir avec une deuxième nucléobase différente par la fonction carbonyle du motif expliquant en partie l'affinité qui est néanmoins observée.



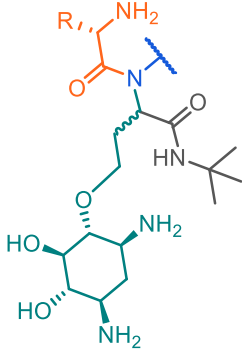
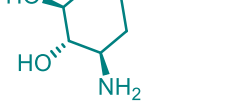
**Figure 74 :** A) Représentation des interactions potentielles du fragments naphtyridine *O*-alkylé avec la cytosine non appariée. B) Représentation des interactions potentielles du fragment naphtyridine *N*-alkylé avec la cytosine non appariée et une autre nucléobase (ici l'adénine).

En conclusion de cette série C-1, nous avons pu mettre en évidence **8 composés** possédant une affinité pour la cible ARN submicromolaire. Une étude plus approfondie de la sélectivité de ces ligands sera donc effectuée par la suite.

### II.5.1.2.2 Évaluation de l'affinité des composés de la série C-2

La série C-2 possède comme noyau central le motif 2-déoxystreptamine (2-DOS). Dans une première partie, ce motif est conjugué avec la lysine et différentes amines. Les valeurs de  $K_D$  obtenues sont illustrées dans le *tableau 23*. Le motif 2-DOS est connu pour interagir au travers de liaisons hydrogènes avec l'ARN et joue un rôle très important dans l'interaction des aminoglycosides avec la cible ARN. L'utilisation de la 2-DOS conjuguée nous a permis d'obtenir des affinités de l'ordre du nanomolaire.

**Tableau 23** : Valeurs des constantes de dissociation ( $K_D$ ) des composés issus de la synthèse avec l'aldéhyde 2-déoxystreptamine pour l'ARN AapA1 :

	Composé	Amine	$K_D$ (nM)
 Lys	<b>C29b</b>	Naphtyridine O-alkyle	<b>119 ± 33</b>
	<b>C30b</b>	2-(4-aminophényl)benzothiazole	<b>101 ± 42</b>
	<b>C31b</b>	Cytosine	1,84 ± 0,50 μM
	<b>C32b</b>	Adénine	850 ± 87
	<b>C33b</b>	Naphtyridine N-alkyle	239 ± 160
 Phe	<b>C34b</b>	Nucléobase « S »	217 ± 71
	<b>C36b</b>	Naphtyridine N-alkyle	175 ± 96
	<b>C37b</b>	Naphtyridine N-alkyle	222 ± 57
	<b>C38b</b>	Naphtyridine O-alkyle	<b>67,3 ± 21,9</b>
	<b>C35b</b>	Lys- Naphtyridine N-alkyle	837 <sup>a</sup>

<sup>a</sup> Ces manipulations n'ont été effectuées qu'une seule fois et dans ce cas l'erreur de la mesure est ±10%.

Tout d'abord, la comparaison des conjugués **C29b**, **C30b**, **C33b** et **C34b** avec leurs analogues dans lesquels la 2-DOS est remplacée par le naphthalène (série C1, composés **C24b**, **C15b**, **C20b**, **C14b**) a montré une légère amélioration de l'affinité. Ces quatre conjugués possèdent ainsi des valeurs de  $K_D$  de 119, 101, 239 et 217 nM.

Les nucléobases naturelles cytosine et adénine interagissent avec leur nucléobase naturelle complémentaire par deux liaisons de Watson-Crick. Les composés comportant ces motifs **C31b** (C) et **C32b** (A), possèdent ainsi de très bonnes affinités avec des valeurs de  $K_D$  submicromolaires même si moins favorables que les composés contenant le motif naphtyridine. Les composés comportant le motif naphtyridine O-alkylé et le benzothiazole substitué tels que les composés **C29b** et **C30b**, possèdent les meilleures valeurs de  $K_D$  qui sont respectivement de 119 nM et 101 nM.

La comparaison des composés comportant les motifs naphtyridines *N*- (**C33b**) et *O*-alkylés (**C29b**) présentent des  $K_D$  de 239 nM et 119 nM ce qui confirme les observations précédentes indiquant que le motif *O*-alkylé est légèrement plus favorable pour l'affinité (**C29b**). Par ailleurs, l'ajout du motif 2-DOS sur le sous-produit **C35b** obtenu lors de la synthèse du composé **C33b**, confère une affinité de 3.5 fois supérieure au composé **C33b**.

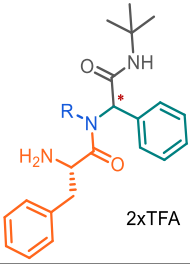
Enfin, sur ces deux motifs naphtyridines, le remplacement de l'acide aminé lysine avec la phénylalanine a été évalué. Pour le motif *N*-alkylé (composés **C36b** et **C37b**), le changement de cette acide aminé n'a pas influencé la valeur de la constante de dissociation. De plus, la différence entre les deux diastéréoisomères n'a pas influencé significativement la valeur de l'affinité entre eux. En revanche pour le composé *O*-alkylé (composé **C33b**) avec la phénylalanine, une amélioration de l'affinité de presque de 2 fois a été constatée avec une valeur de 67,3 nM par rapport au composé **C29b**.

Comme pour les composés de la série C-1, des études ultérieures de sélectivité seront effectuées.

### II.5.1.2.3 Évaluation de l'affinité des composés de la série C-3

Dans la série de conjugués C-3, différentes sous-familles ont été synthétisées. Dans une première partie, nous avons évalué l'effet du motif amine en présence de l'acide aminé phénylalanine et du benzaldéhyde (*tableau 24*). Comme évoqué dans les différents exemples précédents, le motif naphtyridine *N*-alkylé (**C41b**) a apporté une meilleure affinité que les bases naturelles G (**C42b**), A (**C43b**), et C (**C44b**).

**Tableau 24** : Valeurs des constantes de dissociation ( $K_D$ ) des composés **C41b**, **C42b**, **C43b** et **C44b** pour l'ARN *AapA1* :



Composé	Amine	$K_D$ ( $\mu\text{M}$ )
<b>C41b</b>	Naphtyridine <i>N</i> -alkyle	18,3
<b>C42b</b>	Guanine	>1 mM
<b>C43b</b>	Adénine	>46
<b>C44b</b>	Cytosine	>1 mM

<sup>a</sup> Ces manipulations n'ont été effectuées qu'une seule fois et dans ce cas l'erreur de la mesure est  $\pm 10\%$ .

Fait intéressant, l'analogue comportant un motif naphthalène (**C17b**, > 70  $\mu\text{M}$ ) n'avait pas montré d'affinité envers l'ARN *AapA1* alors que le remplacement du motif naphthalène par un motif phényle, a nettement amélioré l'affinité pour le composé **C41b** (18,3  $\mu\text{M}$ ).

Ensuite, la série naphtyridine *N*-alkylé/naphthalène/AA (série C-1, *tableau 22*) peut être également comparée avec la série comportant les motifs *N*-alkylé/thiazole/AA (*tableau 25*). Les valeurs de constantes de dissociation de cette dernière sous-famille sont résumées dans le *tableau 25* suivant.

**Tableau 25** : Valeurs des constantes de dissociation ( $K_D$ ) des composés de la sous-famille naphtyridine N-alkylé/thiazole/AA pour l'ARN AapA1 :

Composé	R = AA	$K_D$ ( $\mu\text{M}$ )
<b>C45b</b>	Phe	>60 <sup>a</sup>
<b>C46b</b>	Ala	37,4 <sup>a</sup>
<b>C47b</b>	Lys	<b>2,34 ± 0,44</b>
<b>C48b</b>	Trp	>280 <sup>a</sup>
<b>C49b</b>	His	19,5 <sup>a</sup>
Adénine-Imidazole-Lys	<b>C50b</b>	Lys 46 <sup>a</sup>

<sup>a</sup> Ces manipulations n'ont été effectuées qu'une seule fois et dans ce cas l'erreur de la mesure est  $\pm 10\%$ .

L'ordre de l'affinité des molécules avec les différents acides aminés reste identique à l'exception du composé tryptophane qui a induit dans cette série une diminution d'affinité (**C48b**): Lys>>His>Ala>Phe>>Trp. La lysine apporte une meilleure affinité mais cette dernière reste 4 fois plus faible que celle du composé comportant le motif naphthalène **C20b** (0,62  $\mu\text{M}$ ). Ces résultats montrent une différence d'interaction entre le motif naphthalène qui a des propriétés intercalantes et le motif thiazole. Enfin, la dernière molécule comportant la nucléobase Adénine et le fragment imidazole (**C50b**, 46  $\mu\text{M}$ ) confirme la différence observée avec les nucléobases naturelles moins affines envers l'ARN comme décrit précédemment.

Enfin dans cette série, le motif glucosamine a été utilisé en tant qu'aldéhyde. Néanmoins, le composé Adénine-Glucosamine-Phénylacétique (**C39b**) ne permet pas d'obtenir une molécule avec une affinité satisfaisante (*tableau 26*). En revanche, pour la structure 2-(4-aminophényl)benzothiazole/Glucosamine/Lys (**C40b**) la valeur de  $K_D$  est de 314 nM et donc ce composé possède une affinité intéressante. Ainsi, les composés avec le motif 2-(4-aminophényl)benzothiazole (**C40b**, **C15b** et **C30b**) sont tous affins et ont des valeurs de constantes d'affinité dans le bas nanomolaire envers l'ARN AapA1.

**Tableau 26** : Valeurs des constantes de dissociation ( $K_D$ ) des composés **C39b** et des trois composés ayant le motif 2-(4-aminophényl)benzothiazole pour l'ARN AapA1 :

Composé	Structure : Amine/Aldéhyde/AA	$K_D$
<b>C39b</b>	Adénine-Glucosamine-Phénylacétique	>1 mM
<b>C40b</b>	2-(4-aminophényl)benzothiazole-Glucosamine-Lys	<b>314 ± 64 nM</b>
<b>C15b</b>	2-(4-aminophényl)benzothiazole-Naphtalène-Lys	152 ± 83,0 nM
<b>C30b</b>	2-(4-aminophényl)benzothiazole-DOS-Lys	101 ± 42 nM

<sup>a</sup> Ces manipulations n'ont été effectuées qu'une seule fois et dans ce cas l'erreur de la mesure est  $\pm 10\%$ .

Ainsi, les études sur ce composé **C40b** seront poursuivies.

Parallèlement à l'affinité, la sélectivité reste un paramètre important afin de comprendre le mécanisme d'interaction des composés et leur applicabilité lors d'études cellulaires sur les bactéries. Nous avons donc tout d'abord choisi d'étudier la sélectivité des composés présentant une affinité intéressante pour la cible.

### II.5.1.3 Évaluation de la sélectivité sur les meilleurs ligands identifiés sur l'ARN *AapA1*.

La sélectivité des 20 meilleurs composés ( $K_D < 2 \mu\text{M}$ ) vis-à-vis de différents acides nucléiques a été ensuite évaluée. Dans une première série d'expériences, la constante d'affinité ( $K'_D$ ) en présence d'un excès de 100 équivalents d'ARNt a été mesurée (*tableau 27*). Un rapport  $K'_D/K_D$  de 1 indique une parfaite sélectivité, alors que des ratios supérieurs suggèrent un manque de sélectivité. Ce manque de sélectivité peut s'expliquer par la capacité des composés à se fixer sur d'autres acides nucléiques.

**Tableau 27** : Valeurs des constantes de dissociation ( $K_D$ ) pour l'ARN *AapA1* et les valeurs en présence d'un excès de compétiteur d'ARNt ( $K'_D$ ) et d'ADN ( $K''_D$ ).

Série	Composé	$K_D$ (nM) <sup>a</sup>	$K'_D$ (nM) <sup>b</sup>	$K'_D/K_D$	$K''_D$ (nM) <sup>b</sup>	$K''_D/K_D$
C-1	<b>C14b</b>	120 ± 1,4	240 ± 105	<b>2,0</b>	583 ± 17	<b>4,9</b>
	<b>C15b</b>	152 ± 84	92,6 ± 31,8	0,6	400 <sup>c</sup>	2,6
	<b>C16b</b>	517 ± 150	401 ± 51	0,8	1199 ± 21	2,3
	<b>C20b</b>	622 ± 211	513 ± 67	0,8	1499 ± 550	2,4
	<b>C21b</b>	1261 ± 78	660 ± 102	0,5	3237 ± 1684	2,6
	<b>C22b</b>	1882 ± 180	1641 ± 605	0,9	5800 <sup>c</sup>	<b>3,1</b>
	<b>C23b</b>	834 ± 204	780 ± 154	0,9	4022 ± 13	<b>4,8</b>
	<b>C24b</b>	287 ± 113	220 ± 132	0,8	738 ± 412	2,6
	<b>C24b</b>	428 ± 46	362 ± 57	0,8	1167 ± 67	2,7
	<b>C28b</b>	42,5 ± 21,0	20,0 ± 0,5	0,5	110 ± 2,5	2,6
C-2	<b>C29b</b>	119 ± 33	83,4 ± 11,8	0,7	319 ± 114	2,7
	<b>C30b</b>	101 ± 42	85,5 ± 2,8	0,9	248 ± 75	2,5
	<b>C31b</b>	1835 ± 500	712 ± 111	0,4	2063 ± 1729	<b>1,1</b>
	<b>C32b</b>	850 ± 87	551 ± 90	0,7	1873 ± 525	2,2
	<b>C33b</b>	239 ± 160	210 ± 14	0,9	697 ± 89	2,9
	<b>C34b</b>	217 ± 71	136 ± 24	0,6	1076 ± 805	<b>5,0</b>
	<b>C36b</b>	175 ± 96	214 ± 13	1,2	620 ± 53	<b>3,6</b>
	<b>C37b</b>	222 ± 57	131 ± 65	0,6	558 ± 11	2,5
<b>C38b</b>	67,3 ± 21,9	139 ± 34	<b>2,1</b>	245 <sup>c</sup>	<b>3,6</b>	
C-3	<b>C40b</b>	314 ± 64	267 ± 135	0,9	1235 ± 645	<b>3,9</b>

Manipulations ont été effectuées : <sup>(a)</sup> 3 fois ; <sup>(b)</sup> 2 fois ; <sup>(c)</sup> 1 fois.

Ainsi, l'étude des valeurs des constantes d'affinité ( $K'_D$ ), nous montrent une sélectivité intéressante envers l'ARN *AapA1*. Une très faible diminution d'activité (facteur 2) est observée pour les molécules **C14b** et **C38b** qui présentent les meilleures affinités pour *AapA1*.

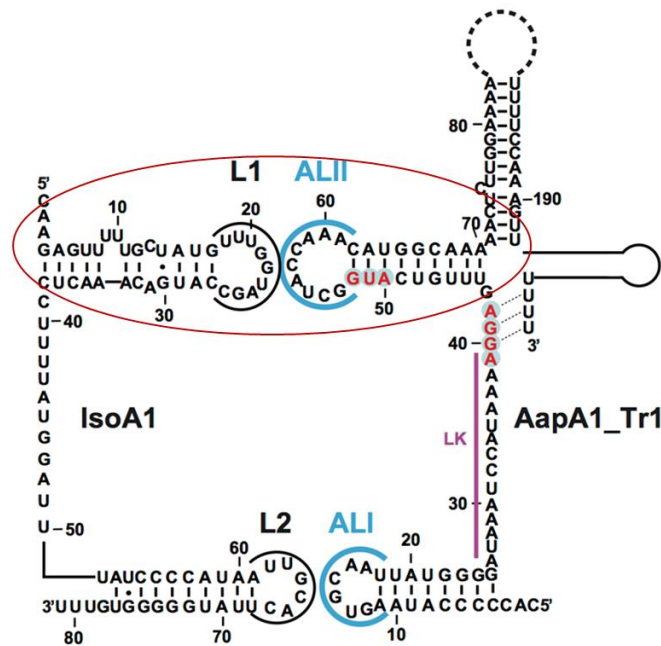
La deuxième série d'expériences de compétition est réalisée en présence d'un excès de 100 équivalents d'un duplex d'ADN donnant les valeurs de constantes de dissociation «  $K''_D$  ». De même le ratio  $K''_D/K_D$  égale à 1 indique une parfaite sélectivité. Ainsi, les ratios  $K''_D/K_D$  observés sont globalement supérieurs et sont comprises entre 2 et 3 majoritairement. Néanmoins, la molécule **144b** ayant une plus faible affinité par rapport aux autres composés ( $K_D = 1,85 \mu\text{M}$ ), possède une meilleure sélectivité que ce soit en présence d'ARNt et d'ADN. A l'opposé, il peut être distingué un ensemble de 7 composés montrant une perte d'affinité importante jusqu'à un ratio  $K''_D/K_D$  de 5.

Les composés **C14b** et **C34b** contenant la nucléobase artificielle S montrent une perte de sélectivité de 5 fois qui peut s'expliquer par une ressemblance avec la structure de l'ADN et l'interaction de cette nucléobase formant un triplex avec les paires de bases AT également. Ainsi l'ajout des motifs naphthalène ou 2-DOS n'ont pas permis d'améliorer la sélectivité. Néanmoins, le composé **C16b** qui comporte la nucléobase « S » mais avec la glycine en bras espaceur permet de garder la sélectivité envers un excès d'ADN ( $K''_D/K_D$  de 2,3).

En conclusion, ces premiers tests de sélectivité nous ont permis de confirmer que la plupart des composés possède une excellente sélectivité en présence d'ARNt et d'ADN puisque seulement trois composés sur les vingt composés évalués ont montré une différence significative notamment en présence d'ADN.

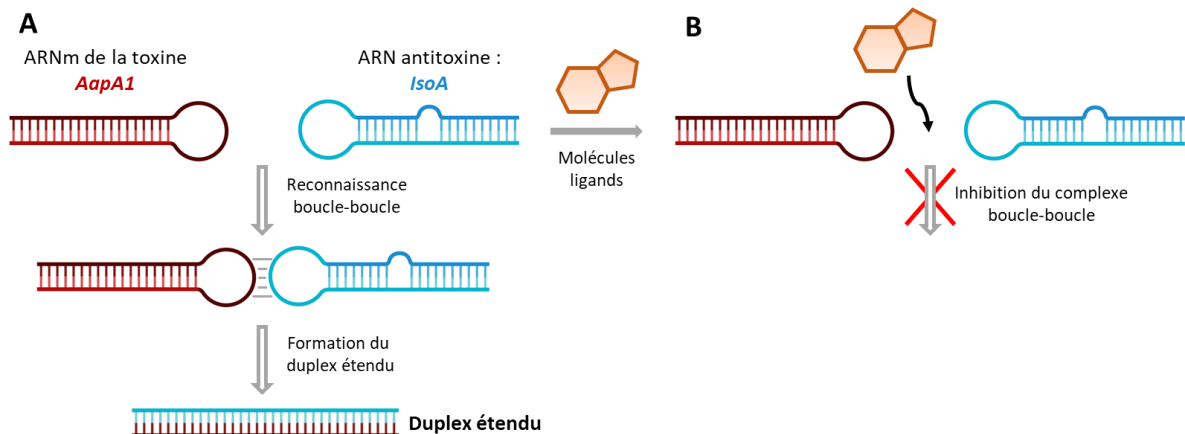
### II.5.2 Mise au point d'un test *in vitro* pour évaluer la capacité d'inhibition de l'interaction boucle-boucle

Pour évaluer la capacité des meilleurs composés à interférer avec l'interaction ARN-ARN Toxine-Antitoxine, un test biochimique *in vitro* a été mis au point au cours du travail de thèse. Lors du test d'affinité, une partie de la séquence de l'ARNm *AapA1* a été étudiée. Cette partie correspond à une des deux tige-boucles impliquées dans la reconnaissance boucle-boucle. Cette interaction permet la formation du duplex étendu entre les ARN *AapA1* et *IsoA1*<sup>59</sup>. Ainsi, pour réaliser ce test, nous avons choisi la séquence complémentaire de la tige boucle (boucle L1) précédemment étudiée se trouvant sur l'ARN *IsoA* (figure 75).



**Figure 75 :** Représentation de l'interaction boucle-boucle entre les ARN AapA1 et IsoA. Dans ce projet, les boucles L1 (pour l'ARNnc IsoA1) et ALII (pour l'ARNm AapA1) ont été sélectionnées.

Pour réaliser l'étude de l'interaction entre ces deux tige-boucles, plusieurs tests complémentaires ont été explorés selon le modèle suivant (figure 76). En absence d'inhibiteur les deux séquences initialement structurées en tige-boucle pourront interagir en formant le « kissing complex » qui va à son tour évoluer vers un duplex étendu (A). En présence d'un inhibiteur la formation du « kissing complex » ne pourra pas avoir lieu (B).



**Figure 76 :** Représentation du test d'inhibition étudié. **A)** Cas de figure lorsque les deux ARN (AapA1 et IsoA1) sont assemblés pour former un duplex étendu. **B)** Cas de figure où une potentielle molécule a un effet d'inhibition du complexe boucle-boucle.

Au cours de la mise au point du test, nous avons choisi d'abord d'étudier la formation du duplex étendu à partir des deux tiges-boucles et ensuite de l'appliquer à l'évaluation de l'activité d'inhibition des composés.

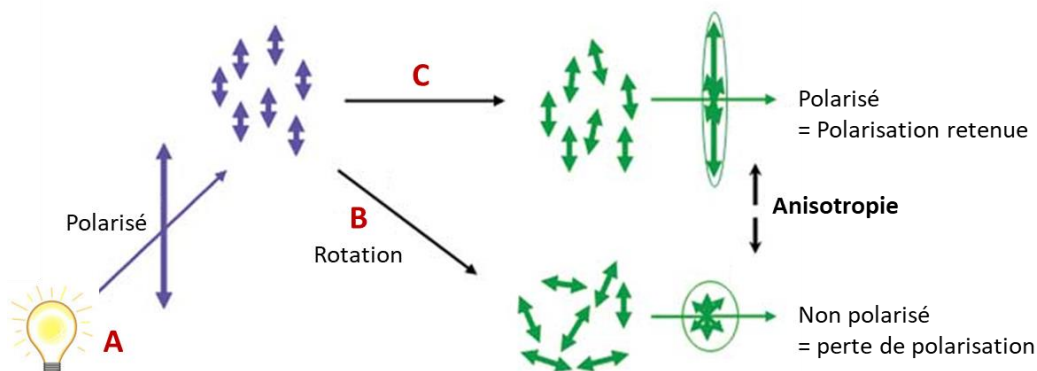
Dans le premier test mis en place, la séquence de l'ARNm *AapA1* utilisée pour l'étude des constantes de dissociation et marquée avec le fluorophore FAM a été utilisée. Grâce à ce marquage fluorescent, nous avons exploré la possibilité de suivre l'inhibition de formation du complexe boucle-boucle par mesure de l'anisotropie de fluorescence qui est décrite dans cette partie. Par ailleurs, nous avons aussi choisi d'explorer une méthodologie FRET (« Förster resonance energy transfer ») dans laquelle nous avons utilisé une construction comportant un couple de fluorophore/quencher (couple FRET). Ces tests seront décrits dans une deuxième partie. Enfin, dans une dernière partie, nous décrirons les études effectuées afin de visualiser la formation des deux tige-boucle et du duplex étendu par électrophorèse qui pourra ensuite être appliquée à l'étude de l'activité d'inhibition pour certains composés.

### II.5.2.1 Test d'inhibition utilisant l'anisotropie de fluorescence

Afin d'étudier l'inhibition de la formation du complexe boucle-boucle, nous avons utilisé une approche selon laquelle les valeurs de l'anisotropie de fluorescence sont mesurées.

#### II.5.2.1.1 Anisotropie de fluorescence (AF)

L'anisotropie est une mesure de la dépolariation de fluorescence. Les principaux événements influençant la valeur de l'anisotropie de fluorescence (AF) sont les mouvements rotationnels du chromophore<sup>210</sup>. Pour déterminer cette rotation, les mesures de AF utilisent la lumière polarisée plane d'un échantillon possédant un fluorophore (*figure 77, A*). Dans un modèle simplifié, si la molécule liée au fluorophore est petite et n'est pas liée à sa cible (par exemple une protéine), sa rotation moléculaire est rapide et entraîne une dépolariation de la fluorescence (*figure 77, B*). Or, si la molécule se lie à la cible de taille plus élevée, la molécule aura une rotation plus faible dirigée par la rotation de la cible, alors la polarisation est conservée (*figure 77, C*). Ce changement est mesuré à l'aide d'un polarimètre.



**Figure 77 :** Représentation schématique d'anisotropie de fluorescence. **A)** Excitation sélective du fluorophore aligné sur les dipôles. **B)** La diffusion rotationnelle entraîne une dépolariation - anisotropie inférieure. **C)** Les fluorophores liés à une macromolécules diffusent ou tournent plus lentement. La polarisation est conservée - anisotropie est plus élevée.



L'étude de l'anisotropie de fluorescence est donc une technique basée sur l'une des caractéristiques de l'émission de la fluorescence qui est la différence de polarisation<sup>210</sup>. L'anisotropie de fluorescence (AF) ou polarisation de fluorescence (PF) est définie en fonction des composantes parallèle ( $I_{||}$ ) et orthogonale ( $I_{\perp}$ ) de l'intensité totale d'émission de la fluorescence. L'étude du degré de polarisation d'un fluorophore est inversement proportionnelle à sa rotation moléculaire donnant l'équation :

$$r = \frac{I_{||} - GI_{\perp}}{I_{||} + 2GI_{\perp}}$$

G étant un facteur de correction de sensibilité pour les deux modes de détection pour les filtres, les polariseurs et les monochromateurs de l'appareil.

La valeur d'AF est corrélée à la vitesse de diffusion rotationnelle de la molécule étudiée. Cette technique est sensible aux propriétés intrinsèques des molécules, telles que le volume et la forme moléculaire, mais également à des facteurs environnementaux (la température et la viscosité de la solution). Un autre paramètre important à prendre en compte est la durée de vie de fluorescence du fluorophore.

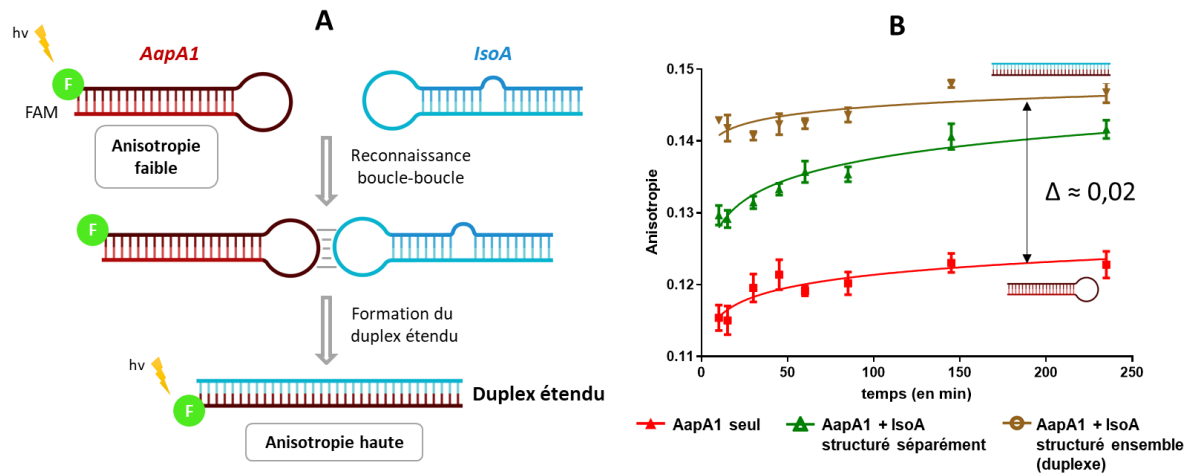
L'anisotropie de fluorescence est avantageuse dans l'étude des interactions moléculaires telles que les événements de liaison protéine-protéine, protéine-ADN et protéine-ligand<sup>210</sup>. L'utilisation de l'AF a également été décrite dans des études ARN-ligand. A titre d'exemple, l'étude de l'interaction de la néomycine avec l'ARN RRE (Rev-Response-Element) du VIH-1 a été réalisé via des expériences utilisant l'anisotropie avec la sonde FAM attachée à la néomycine<sup>211</sup>.

Pour résumer, une molécule fluorescente ayant un taux de rotation plus élevé présente généralement des valeurs d'anisotropie plus faibles. *A contrario*, les molécules fluorescentes avec des taux de rotation plus faibles donnent des valeurs d'anisotropie plus élevées. Dans ce contexte, nous avons exploité cette différence de rotation dans le cas d'une structure de tige-boucle et d'un duplex étendu.

#### II.5.2.1.2 Mise au point du test

Pour réaliser un test d'inhibition en profitant des mesures d'anisotropie, la sonde fluorescente (FAM) permettant la mesure de l'AF est attachée sur l'ARNm *AapA1*. Lorsque cet ARN est de structure tige-boucle, la rotation devrait être plus rapide que lorsque le duplex étendu entre *AapA1* et *IsoA1* est formé (*figure 78, A*). Ainsi, une différence d'anisotropie devrait être mesurable entre les deux états (tige-boucles et duplex étendu). Pour vérifier et mesurer cet écart de valeur d'anisotropie, un premier test a été réalisé (*figure 78, B*). Dans ce test, nous avons mesuré la valeur d'anisotropie pour le duplex étendu qui est préformé en structurant ensemble les ARN *AapA1* et *IsoA1* (courbe marron, *figure 78, B*). De même, la valeur d'anisotropie a été mesurée pour l'ARN *AapA1* de structure tige-boucle (courbe

rouge, *figure 78, B*). Les valeurs d'anisotropie ont été mesurées à intervalles de temps réguliers en incubant les échantillons à 37°C pour une concentration d'ARN de 50 nM. Nous avons ainsi pu déterminer un delta se rapprochant de 0,02 entre la valeur de l'ARN *AapA1* et du duplex étendu. Cette valeur correspond aux valeurs reportées précédemment dans la littérature pour des interactions similaires.



**Figure 78 :** **A)** Représentation du test de mesure de l'anisotropie de fluorescence. **B)** Résultat obtenu comparant les valeurs d'anisotropie de fluorescence de *AapA1* structuré en tige-boucle (rouge), du duplex *AapA1/IsoA1* préformé (marron) et des structures tige-boucles *AapA1* et *IsoA1* incubées à 37°C (vert). Conditions : Tampon = HEPES à 20 mM, pH 7.4, 20 mM NaCl, 140 mM KCl, et 3 mM MgCl<sub>2</sub>; [ARN] = 50 nM ; incubation à 37°C.

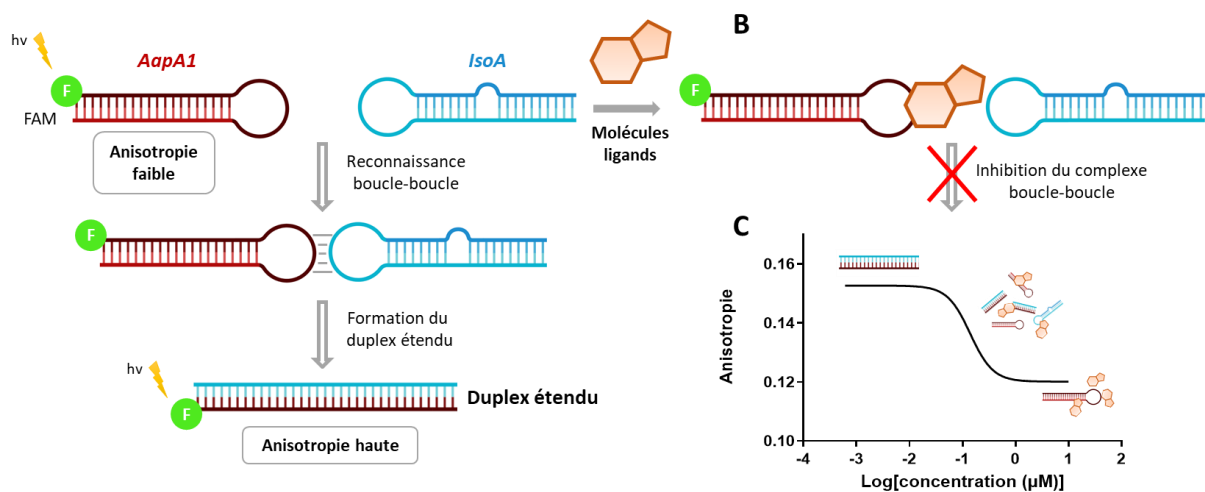
Dans cette même expérience, nous avons évalué le comportement des deux tiges-boucles des ARN *AapA1* et *IsoA1* qui ont été structurées séparément puis mises à interagir ensemble comme représenté dans la *figure 78, A*. A temps défini, les mesures d'anisotropie ont été relevées à 37°C donnant une courbe qui tend vers la courbe du duplex étendu (courbe verte, *figure 78, B*). L'interprétation de cette courbe montre que la formation du duplex étendu est rapide. En effet, à  $t = 15$  min, la valeur de l'anisotropie correspond à une valeur plus haute que la courbe de l'ARN *AapA1* de structure tige-boucle. Ceci démontre que la reconnaissance entre les deux ARN est rapide pour ensuite former le duplex étendu.

A l'aide de ce test, différents paramètres ont été étudiés pour déterminer les conditions optimales qui permettent de différencier les deux structures et ainsi d'évaluer l'activité des composés. Ainsi, la nature du tampon (HEPES ou TRIS), la température de lecture (4°C, 25°C et 37°C), la concentration de sels (NaCl, KCl et MgCl<sub>2</sub>), la température d'incubation (25°C et 37°C) ont été évaluées. Cela a permis de sélectionner le tampon HEPES à 20 mM, avec une valeur de pH 7.4, une concentration de 20 mM de NaCl, 140 nM de KCl, et 3 mM MgCl<sub>2</sub>, une incubation à 37°C pendant 2h. Ces conditions correspondent aux conditions qui sont les plus proches des conditions physiologiques. Par ailleurs, l'évolution de ces complexes présentant une cinétique rapide, nous avons choisi d'incuber les

complexes à 37°C mais d'effectuer les lectures à 4°C après 20 minutes d'incubation à cette température afin de stabiliser la structure au moment de la lecture. Ce dernier paramètre est très important puisqu'il permet d'avoir des valeurs constantes.

Dans un deuxième temps, nous avons aussi étudié la concentration d'ARN à utiliser lors de ces tests. Le but de ce test d'inhibition est en effet de pouvoir cribler un grand nombre de composés en utilisant un minimum d'ARN. De ce fait, nous avons réalisé les expériences à la même concentration que celle des tests d'affinités qui est de 5 nM. Par ailleurs, la concentration d'ARN n'est pas prise en compte au niveau de l'équation d'anisotropie. Ainsi, la variation de la concentration ne devrait pas influencer la valeur d'anisotropie. En utilisant les mêmes conditions que le graphique précédent (à une concentration de 50 nM) (*figure 78, B*) mais à une concentration de 5 nM, le delta d'anisotropie entre l'ARN *AapA1* et le duplex étendu n'a pas varié. Il sera donc possible d'utiliser une concentration moindre d'ARN pour étudier l'activité d'inhibition des composés.

Au travers de ces résultats encourageants, l'inhibition de ce complexe en présence de 7 de nos meilleures ligands d'*AapA1* a été explorée (*figure 79*).



**Figure 79 :** **A)** Représentation schématique de test basé sur la mesure par anisotropie de fluorescence. **B)** Cas de figure d'une molécule inhibitrice du complexe boucle-boucle. **C)** Représentation du résultat des valeurs d'anisotropie obtenues dans le cas d'inhibition du complexe boucle-boucle. F = Fluorescéine = FAM.

Deux cas de figure sont ainsi envisageables : (i) le composé n'est pas capable d'inhiber la formation du complexe boucle-boucle, dans ce cas la valeur de l'anisotropie devrait rester constante avec une valeur correspondant au duplex étendu (la valeur haute d'anisotropie) et (ii) le ligand est capable de perturber la formation du complexe étendu (*figure 79, B*), en variant la concentration du ligand, les valeurs d'anisotropie doivent former une courbe sinusoidale avec deux plateaux distincts (*figure 79, C*). Dans ce cas de figure, pour une concentration faible du composé, la formation du duplex étendu est privilégiée et par conséquent une valeur d'anisotropie haute est obtenue (plateau haut). Cependant, lorsque la concentration augmente et permet l'inhibition de la formation du duplex

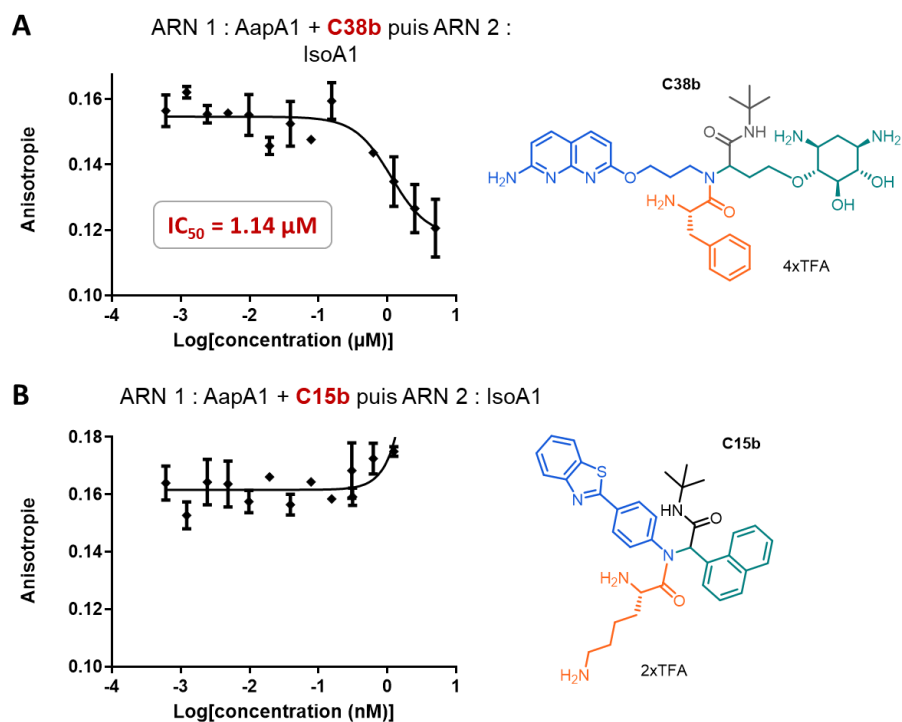
étendu, un plateau de valeur d'anisotropie plus faible sera observé. Cette valeur correspondant ainsi à l'anisotropie de l'ARNm *AapA1* de structure tige-boucle en présence de molécules.

Pour réaliser ce test de criblage préliminaire, des concentrations croissantes des molécules synthétisées ont été incubées avec *AapA1* à 4°C pendant une nuit pour favoriser leur interaction avec l'ARN, puis *IsoA1* est ajouté. Pour étudier l'influence de l'ordre d'ajout des ARN, deux expériences indépendantes ont été réalisées. Dans ces expériences, les molécules sont d'abord mises en présence soit de l'ARN *AapA1* soit de l'ARN *IsoA*, ensuite le deuxième ARN est ajouté. Pour chacune de ces expériences, les valeurs d'anisotropie ont été mesurées après incubation à 37°C pendant 2h. Un contrôle du delta d'anisotropie entre le duplex étendu préformé et de l'ARNm *AapA1* a aussi été vérifié.

Nous avons ainsi testé pour l'instant 7 composés : **C14b**, **C15b**, **C20b**, **C25b**, **C28b**, **C34b** et **C38b**. Les conjugués **C14b**, **C15b**, **C20b** et **C25b** appartiennent à la série C-1 et ont montré des résultats particulièrement prometteurs pour leur affinité et sélectivité envers la cible ARN. Le composé **C28b** fait aussi partie de cette série mais c'est le seul composé contenant la néomycine. Enfin, les conjugués **C34b** et **C38b** appartiennent à la série C-2 et ont aussi montré d'excellentes affinités et sélectivités pour la cible. Ci-dessous, les deux types de profils qui ont été obtenus pour l'expérience où l'ARN *AapA1* a été ajouté avant l'ARNnc *IsoA1* (*figure 80*).

Les gammes de concentrations des molécules employées dans ces tests ont été choisies sur la base des concentrations utilisées pour la mesure des  $K_D$ . La figure 80 illustre le résultat obtenu pour le conjugué **C38b** (*figure 80, A*), et la courbe montre bien l'inhibition de la formation du complexe boucle-boucle. Une valeur d' $IC_{50}$  de 1,35  $\mu\text{M}$  a ainsi été mesurée. Cependant, la gamme de concentration nécessite d'être adaptée afin de confirmer le résultat obtenu et d'obtenir une meilleure visualisation des deux plateaux souhaités.

De manière similaire, le conjugué **C20b** présente une valeur de  $IC_{50}$  de 10.4  $\mu\text{M}$  et le composé **C28b** de 253  $\mu\text{M}$ . Pour les conjugués **C14b**, **C25b** et **C34b**, la gamme de concentration devra également être adaptée puisque même si un début d'inhibition est observé, la gamme utilisée n'est pas adaptée à une mesure précise de l' $IC_{50}$ .



**Figure 80 :** Résultats obtenus pour les molécules **C38b** (A) et **C15b** (B) pour le test d'inhibition du complexe boucle-boucle AapA1 et IsoA1. Conditions : incubation à 37°C pendant 2h ; Tampon = HEPES à 20 mM, pH 7.4, 20 mM NaCl, 140 mM KCl, et 3 mM MgCl<sub>2</sub> ; [ARN] = 5 nM.

En revanche, Le conjugué **C15b** (figure 80, B) montre un plateau de valeur haute d'anisotropie correspondant à la formation du duplex étendu. A plus haute concentration, l'anisotropie semble augmenter ce qui pourrait indiquer une stabilisation du « kissing loop », c'est-à-dire d'un état intermédiaire entre la tige-boucle et le duplex étendu, mais ce résultat reste à confirmer. En effet, l'inhibition du complexe boucle-boucle n'a pas été observée pour la gamme de concentration de la molécule utilisée. Cette gamme de concentration devra donc être adaptée pour les prochaines expériences afin de vérifier le résultat obtenu.

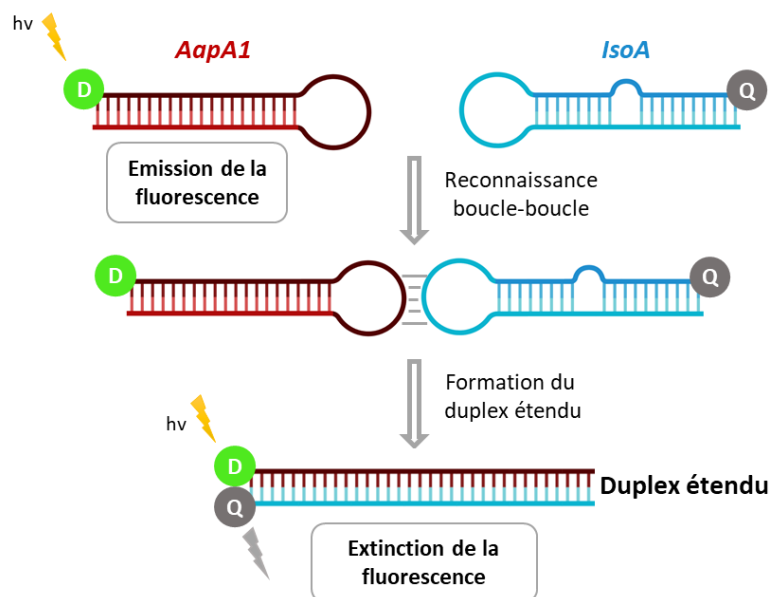
Dans le cas de l'ordre de l'ajout des ARN, celui-ci ne semble pas avoir une influence sur les résultats. La valeur de l' $\text{IC}_{50}$  obtenue pour la molécule **C38b** est de 1,36  $\mu\text{M}$ . Dans le futur, l'ensemble des gammes de concentration doivent être ajustées pour confirmer les résultats obtenus. Néanmoins ces résultats sont particulièrement prometteurs puisque très peu d'exemples de ligands d'ARN capables d'inhiber des interactions boucle-boucle ont été reportés dans la littérature et les composés illustrés dans cette partie enrichissent donc fortement ce domaine de recherche.

En conclusion de cette partie, nous avons pu mettre en place un test de criblage permettant d'évaluer la capacité des molécules à inhiber la formation du complexe boucle-boucle et de quantifier l'inhibition par une mesure des valeurs de  $\text{IC}_{50}$ . Lors d'un criblage de 7 molécules, la molécules **C38b** a montré un effet d'inhibition sur le complexe AapA1/IsoA1 étudié avec un  $\text{IC}_{50}$  de 1,14  $\mu\text{M}$ . Ce résultat est issu d'une étude préliminaire et est encourageant. Comme mentionné, très peu d'études sur la

formation et l'inhibition d'interactions boucle-boucle existent dans la littérature, et donc ce type de test est particulièrement original. Dans ce contexte, nous avons décidé d'entamer l'étude d'autres tests afin de caractériser non seulement la formation des complexes mais aussi la capacité d'inhibition des composés synthétisés. Ces autres tests sont encore en cours de réalisation et pourraient confirmer ces résultats dans le but d'avoir un test fiable.

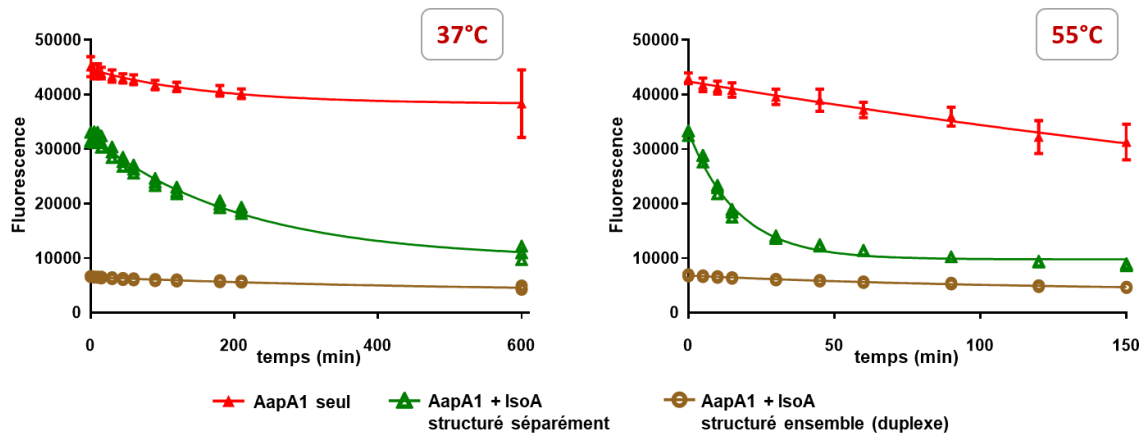
### II.5.2.2 Test d'inhibition utilisant un couple FRET

Dans une deuxième partie, pour réaliser le test d'inhibition, nous avons exploré une autre construction de marquage des ARN utilisant le principe de FRET (*figure 81*). Dans cette étude, l'ARNm *AapA1* est marqué en position 5' avec le fluorophore FAM qui émet de la fluorescence lorsqu'il est sous forme de tige-boucle, tandis que sur l'ARN *IsoA1*, en position en 3', un marquage par un quencher EDQ est réalisé. Lorsque les deux ARN sont structurés ensemble pour former le duplex étendu, le quencher est positionné côte à côte du fluorophore FAM. Au travers de ce rapprochement, l'énergie de fluorescence est transférée au quencher qui éteint la fluorescence du fluorophore FAM (principe de FRET).



**Figure 81 :** Représentation schématique de test d'inhibition utilisant la technique par FRET. D = donneur = FAM, Q = quencher = EDQ.

Ce test nous a permis de confirmer plusieurs informations retrouvées dans le test d'inhibition décrit précédemment avec l'anisotropie de fluorescence. Par cette construction, nous avons réalisé une étude de la formation du duplex étendu qui est présenté ci-dessous (*figure 82*). Pour réaliser ces tests, une concentration de 50 nM en ARN a également été utilisée. Les mesures de fluorescence ont été réalisées en fonction du temps à 37°C (*figure 82, A*) et à 55°C (*figure 82, B*).



**Figure 82 :** Courbes de la cinétique de la formation du duplex étendu en fonction du temps et selon la température : A) 37°C et B) 55°C. Conditions : HEPES à 20 mM, pH 7.4, 20 mM NaCl, 140 mM KCl, et 3 mM  $MgCl_2$ ;  $[ARN] = 50$  nM.

Comme nous en avons fait l'hypothèse, l'ARN *AapA1* marqué à la fluorescence donne les valeurs de fluorescence les plus hautes (courbe rouge, *figure 82*). Lorsque *AapA1* et *IsoA1* sont structurés ensemble pour former le duplex étendu, nous observons une excellente extinction de fluorescence (courbe marron, *figure 82*). Lorsque les deux ARN structurés séparément sont incubés à 37°C ou à 55°C, la formation du duplex étendu peut être observée grâce à la variation de la fluorescence qui se rapproche au fur et à mesure du temps de celle du duplex (courbe verte, *figure 82*). Cette extinction de fluorescence indique la formation du duplex étendu avec une cinétique plus rapide pour une incubation à 55°C qu'à 37°C. Il est également observé qu'à  $t = 0$  la fluorescence est d'ores et déjà plus faible que celle de l'ARN *AapA1* seul ce qui montre que l'interaction des deux ARN est très rapide pour ensuite former le duplex étendu après ajout du deuxième ARN.

Par cette expérience, l'influence du tampon ainsi que la concentration en magnésium de ces tampons ont été étudiés à 37°C et à 55°C. Les meilleures conditions déterminées sont les mêmes que celles du test d'anisotropie (tampon HEPES à 20 mM, pH 7.4, 20 mM NaCl, 140 mM KCl, et 3 mM  $MgCl_2$ ).

Dans l'intention de pouvoir utiliser un système FRET pour réaliser un test de criblage comme pour le test d'anisotropie ci-avant, une construction avec le système fluorophore/quencher sur un seul des deux ARN est aussi envisagée. Cette construction pourrait nous permettre d'étudier la capacité d'inhibition des molécules grâce à une extinction de fluorescence lors de l'inhibition.

En conclusion, nous avons pu mettre au point un deuxième test de criblage et étudier en détail la cinétique de la formation du duplex étendu. Les résultats obtenus suggèrent qu'un temps minimum de 2h d'incubation à 37°C est nécessaire pour étudier efficacement la formation du complexe étendu et l'effet des molécules. Néanmoins, lorsque nous avons essayé de tester des composés, nous avons observé une extinction de fluorescence due à l'interaction du composé avec l'ARN, ce qui est la base du test que nous avons utilisé pour mesurer les  $K_D$  et décrit précédemment. D'autres essais seront

nécessaires pour vérifier si un tel type de test peut donc être utilisé pour évaluer l'activité d'inhibition des composés.

Dans ce contexte, nous avons souhaité tester une autre construction dans laquelle un des ARN est marqué en position 5' par le fluorophore Cy3 donneur alors que l'autre est marqué en position 3' par le fluorophore Cy5 accepteur. Lorsque les deux fluorophores sont proches (formation du duplex étendu) l'efficacité du FRET (« FRET efficiency » ou E-FRET) est haute alors que lorsqu'ils sont éloignés (inhibition de la formation du duplex) la valeur sera basse. Le E-FRET est calculé comme le rapport entre l'intensité de fluorescence du fluorophore accepteur (Cy5) et la somme des intensités de fluorescence pour les deux fluorophores Cy3 et Cy5. Le calcul d'un rapport de fluorescence pourrait nous affranchir du problème d'extinction de fluorescence en présence des composés observé pour le test FRET précédent. L'évaluation des composés en utilisant ce test est en cours actuellement.

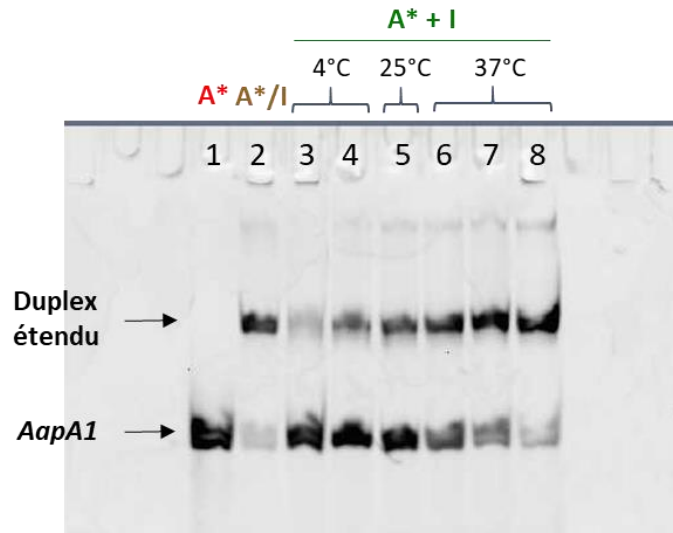
Afin de compléter ces études, nous avons aussi souhaité visualiser la formation du « kissing complex » et du duplex étendu par retardement sur gel d'électrophorèse et ces études seront décrites dans le paragraphe suivant.

### II.5.2.3 Expérience de retardement sur gel d'électrophorèse

Pour mieux caractériser les structures formées par les deux ARN et aussi confirmer si une molécule est capable d'inhiber un complexe boucle-boucle, nous avons décidé de réaliser d'autres expériences complémentaires dont des études de retardement sur gel d'électrophorèse. Ces expériences ont été notamment utilisées pour l'étude d'aminoglycosides sur l'ARN DIS<sup>121,212</sup>. La technique par électrophorèse permet de faire migrer les ARN selon leur poids moléculaires et leur structure. Ainsi, une différence entre une structure tige boucle (*AapA1*) et une structure en duplex étendu (*AapA1/IsoA1*) pourront être distinguées par électrophorèse.

Ainsi, dans une première expérience, nous avons étudié la cinétique de la formation du duplex étendu selon la température et le temps (*figure 83*).

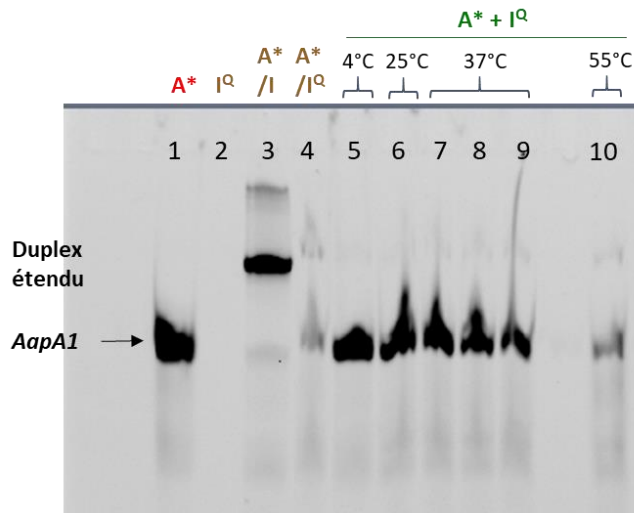




**Figure 83 :** Image du gel d'électrophorèse (20% acrylamide/bis-acrylamide 19 :1, TBE 1x) : la **ligne 1** représente AapA1-FAM seul de structure tige-boucle (bande AapA1) ; la **ligne 2** représente le duplex étendu AapA1/IsoA1 préformé (bande duplex étendu) ; ligne 3 à 8 représentent l'interaction entre les tige-boucles AapA1-FAM et IsoA1 à différentes températures d'incubation (4°C, 25°C ou 37°C) et pour une durée déterminée : **ligne 3** = 0 min, 4°C. **ligne 4** = 2h, 4°C. **ligne 5** = 2h, 25°C. **ligne 6** = 15 min, 37°C. **ligne 7** = 30 min, 37°C et **ligne 8** = 2h, 37°C. Migration à 4°C.

La **ligne 1 et 2** représentent les bandes témoins de AapA1 de structure tige boucle (bande AapA1) et du duplex étendu AapA1/IsoA1 préformé (bande duplex étendu). Ensuite, pour avoir un témoin des deux ARN ajoutés ensemble sans formation de duplex étendu, les ARN ont été incubés à 4°C avant d'être rassemblés et analysés (**ligne 3**). Ainsi, la bande majoritaire correspond à la bande de AapA1 seul montrant que le duplex étendu n'a pas encore été formé. Ensuite, pour la **ligne 4, 5 et 8**, les ARN ont été incubés ensemble à 4°C, 25°C et 37°C pendant 2h. Ces résultats montrent une faible formation du duplex étendu à des température basses de 4°C et 25°C. A contrario, à 37°C, la formation du duplex est totale. En effet, une étude cinétique de la formation du duplex étendu peut être visualisée en observant les **lignes 6** (t = 15 min), **7** (t = 30 min) et **8** (t = 2h) dans lesquelles les deux ARN incubés à 37°C ont été déposés. La formation rapide du complexe boucle-boucle duplex étendu corrèle donc bien avec les études précédentes.

Afin de vérifier si la bande formée lors de l'ajout des deux ARN ensemble correspond bien au duplex étendu ou à une possible formation du complexe boucle-boucle intermédiaire (« kissing loop »), une deuxième expérience a été réalisée. Pour cela, l'ARN IsoA1 marqué avec le quencher EDQ est incubé avec AapA1-5'-FAM ce qui induit l'extinction de la fluorescence correspondante à la bande du duplex étendu (**figure 84**).



**Figure 84** : Image du gel d'électrophorèse (20% acrylamide/bis-acrylamide 19 :1, TBE 1x): la **ligne 1** représente AapA1-FAM seul de structure tige-boucle (bande AapA1) ; la **ligne 2** représente l'ARN IsoA1 marqué avec le quencher EDQ ; la **ligne 3** et **4** représentent le duplex étendu AapA1/IsoA1 préformé soit avec l'ARN IsoA non marqué (**ligne 3**, bande duplex) soit avec l'ARN IsoA-EDQ (**ligne 4**, bande du duplex étendu non visible) ; **ligne 5** à **10** représentent l'interaction entre les tige-boucles AapA1-FAM et IsoA1-EDQ à différentes températures d'incubation (4°C, 25°, 37°C ou 55°C) et pour une durée déterminée : **ligne 5** = 2h, 4°C. **ligne 6** = 2h, 25°C. **ligne 7** = 15 min, 37°C. **ligne 8** = 30 min, 37°C. **ligne 9** = 2h, 37°C et **ligne 10** = 2h, 55°C. Migration à 4°C.

Par cette expérience, pour la **ligne 4**, une extinction de la fluorescence du duplex étendu préformé utilisant le couple FAM/EDQ est observée par comparaison du duplex étendu comportant seulement le marqueur FAM (**ligne 3**). Ainsi, dans les mêmes conditions que précédemment, la formation du duplex étendu a été réalisée à 4°C (**ligne 5**) et 25°C (**ligne 6**) après incubation pendant 2h et à 37°C pour différents temps (15 min = **ligne 7**, 30 min = **ligne 8** et 2h = **ligne 9**). Une 4<sup>ème</sup> température a été rajoutée qui est une incubation à 55°C (**ligne 10**). Pour ces différentes conditions d'incubation, nous pouvons visualiser l'absence de la bande du duplex étendu. Ainsi, il est bien question de la formation du duplex étendu et non du complexe boucle-boucle. L'absence de bande pour les conditions de la ligne 10, confirme la conversion totale en duplex étendu à 55°C pendant 2h.

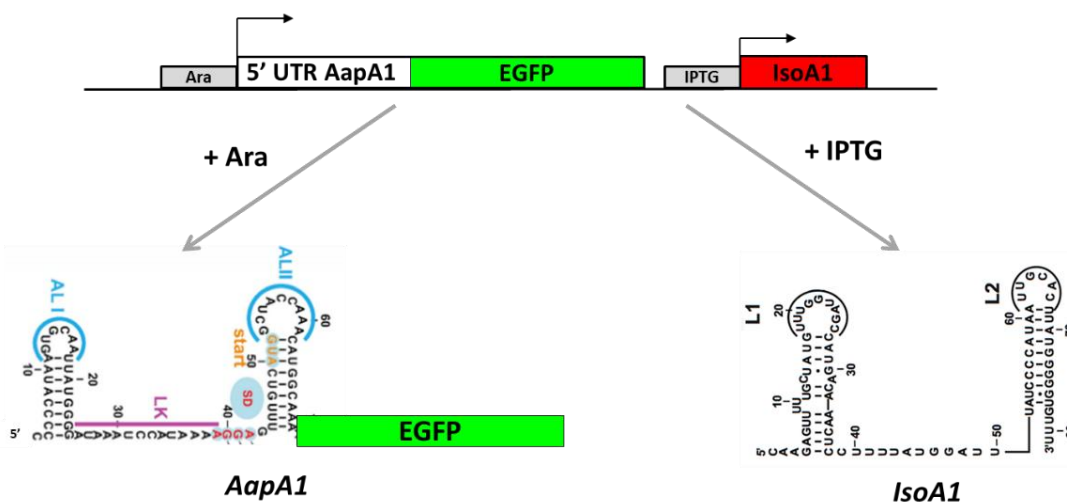
En conclusion de ces deux expériences, nous pouvons confirmer la formation du duplex étendu lorsque les deux ARN sont mis ensemble. Par ailleurs, ce test pourra être utilisé pour confirmer si une potentielle molécule montre un effet inhibiteur du complexe boucle-boucle déterminé par le test d'inhibition précédemment décrit.

Dans le même principe que développé dans cette partie, l'équipe du Dr Darfeuille a également élaboré un test *in vivo* basé sur l'utilisation d'un système rapporteur et portant sur les séquences entières des deux ARN étudiés. Ce test est effectué *in vivo* dans la bactérie et il est décrit dans la section suivante.

### II.5.3 Evaluation de l'activité *in vivo* sur le système *AapA1/IsoA1* – « système rapporteur »

Pour évaluer l'inhibition du système TA *AapA1/IsoA1*, l'équipe du Dr. Darfeuille a développé un test permettant de quantifier la traduction de la toxine *AapA1* par une méthode utilisant une protéine fluorescente : la EGFP (Enhanced Green Fluorescent Protein).

Ce test a été réalisé au sein de la bactérie *Escherichia coli* qui ne possède pas le système TA *AapA1/IsoA1* propre à *H. pylori* et qui n'est donc pas sensible à la toxine *AapA1*. Pour pouvoir étudier la séquence du système TA *AapA1/IsoA1*, un plasmide a été conçu et est présenté ci-dessous (*figure 85*). Cette construction permet de transcrire indépendamment l'ARNm *AapA1* et l'ARNnc *IsoA1* dans la bactérie après transfection. Pour quantifier la traduction de la toxine *AapA1* une sonde GFP a été ajoutée à la séquence de l'ARNm de la toxine. Lorsque la séquence de l'ARN *AapA1/GFP* est traduite par induction à l'arabinose (Ara), la toxine et l'EGFP seront exprimées. Ainsi, la traduction de la toxine pourra être suivie par mesure de la fluorescence de la protéine GFP. Concernant la séquence de l'ARNnc *IsoA1*, l'induction de *IsoA1* se fait par l'action de l'IPTG (isopropyl  $\beta$ -D-1-thiogalactopyranoside).



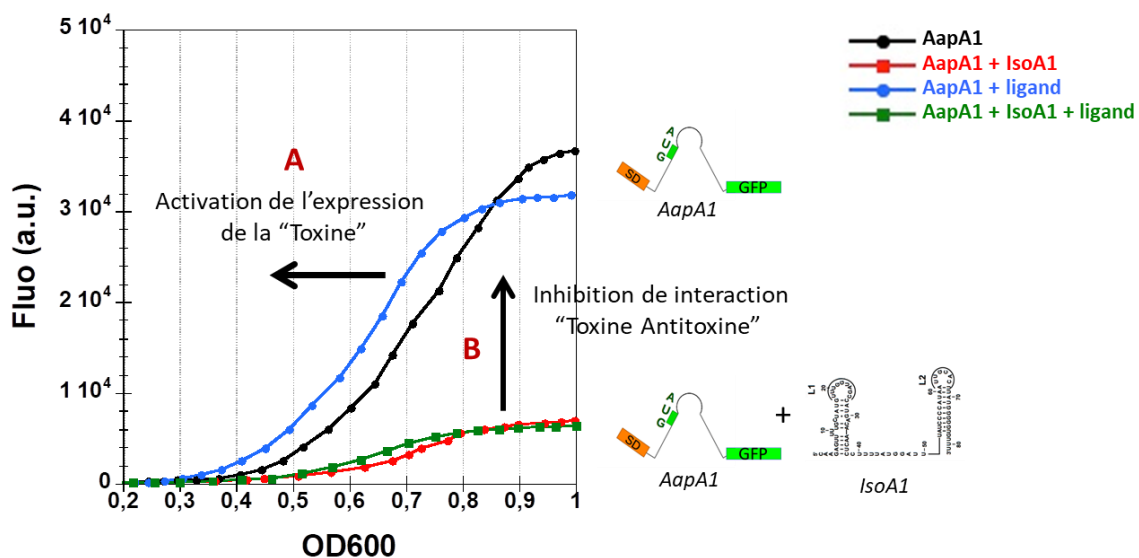
**Figure 85 :** Construction du plasmide pour réaliser le test du système rapporteur. Une protéine fluorescence (EGFP) est insérée à la séquence de l'ARNm *AapA1*. Une induction à l'arabinose (Ara) permet la transcription de l'ARNm. Une deuxième cassette comporte la séquence de l'ARNnc *IsoA1* qui est induit par l'action de l'IPTG (isopropyl  $\beta$ -D-1-thiogalactopyranoside).

Outre la mesure de la fluorescence par la protéine GFP, une mesure optique à OD<sub>600nm</sub> est également réalisée et permet de suivre la croissance bactérienne. Grâce à cette mesure, la viabilité bactérienne en présence des composés pourra être déterminée.

Ainsi grâce à cette construction, il est possible de réaliser deux expériences en absence ou présence de molécules (*figure 86*). Pour chacune des expériences, les mesures de l'OD<sub>600nm</sub> et de la fluorescence

sont réalisées en fonction du temps. Ainsi, les expériences peuvent être tracées par la mesure de l'OD<sub>600nm</sub> en fonction de la valeur de fluorescence.

Pour la première expérience (*figure 86, A*), une induction à l'Ara permet de visualiser la viabilité de la bactérie par une mesure de l'OD<sub>600nm</sub> et d'obtenir une courbe témoin de l'expression de la toxine en regardant l'expression de la protéine GFP (fluorescence élevée) (*figure 86*, courbe noire). Lorsqu'une molécule est ajoutée, son influence sur l'expression de la toxine et sur la viabilité de la bactérie pourra être évaluée (*figure 86*, courbe bleue). Dans le cas représenté sur ce graphique, la molécule **C26b** n'a pas un effet sur la croissance bactérienne et donc n'est pas toxique pour la bactérie. De même l'expression de la traduction de la toxine n'a pas été influencée.



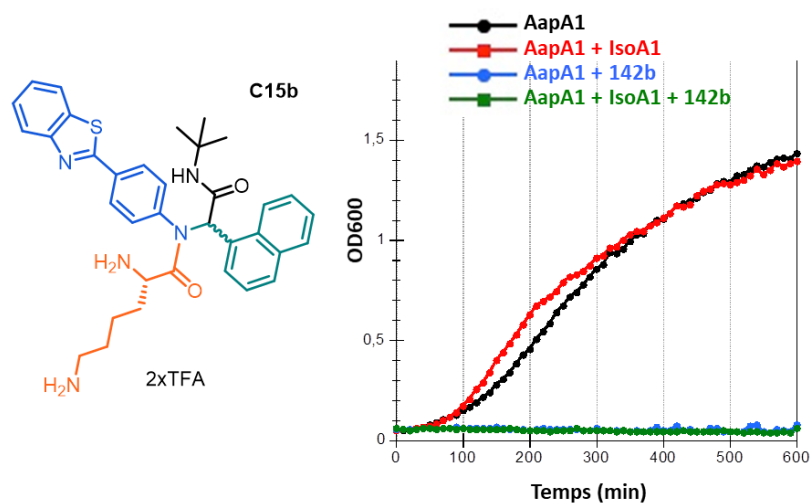
**Figure 86 :** Représentation d'un graphique type du test du système rapporteur en mesurant l'OD<sub>600nm</sub> en fonction de la fluorescence (protéine GFP). **A)** Induction par l'Ara permet l'expression de la toxine AapA1 (courbe noire). En présence du conjugué **C26b** mais en absence d'IsoA1 ni la croissance ni l'expression de la toxine sont affectées (courbe bleue). **B)** Induction par l'Ara et IPTG permettant l'inhibition de la toxine par la formation du complexe Toxine-Antitoxine (courbe rouge). En présence du conjugué **C26b** on n'observe pas de changement e qui indique que le composé n'est pas inhibiteur (courbe verte).

Dans un deuxième test (*figure 86, B*), une induction en présence de l'Ara et de l'IPTG permet d'obtenir les conditions de la formation du complexe Toxine-Antitoxine (*figure 86*, courbe rouge). En absence de molécules, une inhibition de la traduction de la toxine par l'antitoxine est ainsi possible. Dans ce cas, la protéine GFP ne sera pas traduite et par conséquent une absence de fluorescence sera observée sans compromettre la croissance de la bactérie.

Lorsqu'une molécule est ajoutée deux scénarios sont possibles. Si la molécule n'a pas d'effet sur l'inhibition de système TA, la protéine GFP ne sera pas traduite et donc la courbe sera identique à la courbe témoin (*figure 86*, courbe verte). En revanche, si une molécule engendre l'inhibition du système Toxine/Antitoxine, l'expression de la toxine sera possible. Ainsi, une augmentation de la

fluorescence par la protéine GFP sera observée se rapprochant de la courbe témoin *Aapa1* (courbe noire). Ce dernier cas représente le modèle idéal et le but du projet décrit dans ce manuscrit.

Une partie des molécules synthétisées au cours de ce projet a été testée à 50  $\mu\text{M}$  en utilisant ce test. Au cours de ces tests, j'ai pu réaliser moi-même la première étape (*figure 86, A*) de l'étude permettant de regarder l'effet des molécules sur la croissance bactérienne grâce à un séjour dans le laboratoire du Dr. Fabien Darfeuille. Par ailleurs, l'effet de la molécule sur la traduction de la toxine *AapA1* a aussi été analysé. Lors de ces tests, seules les molécules **C28b** (Néomycine/naphtalène/His) et **C15b** (*figure 87*) ont montré une toxicité envers la bactérie *E. coli*. Sur le graphique ci-dessous, il est observé une absence totale de valeur d'OD<sub>600nm</sub> à 50  $\mu\text{M}$  de la molécule **C15b** montrant une toxicité intéressante. Malheureusement, la toxicité ne permet pas d'évaluer l'action de la molécule sur le système Toxine-Antitoxine étudié. En effet, la toxicité est forcément due à un autre mécanisme puisque la toxine *IsoA1* n'est pas toxique pour *E. coli* et la mort de la bactérie empêche l'évaluation de l'effet sur le système *AapA1/IsoA1*. Néanmoins, un tel résultat reste très intéressant pour deux raisons. Tout d'abord, le conjugué **C15b** représente une nouvelle molécule possédant une activité antimicrobienne. Deuxièmement, la toxicité envers *E. coli* n'implique pas que cette molécule ne possède pas d'effet sur le système TA étudié.



**Figure 87** : Résultats obtenus du test du système rapporteur du composé **C15b**.

Pour la deuxième partie du test donnant les informations sur l'effet d'inhibition du système Toxine-Antitoxine, les molécules ont également été testées à 50  $\mu\text{M}$ . A ce jour, aucune molécule n'a montré d'inhibition pour le système étudié alors que certaines ont montré une activité lors des tests d'inhibition *in vitro*. Ceci pourrait être dû à plusieurs raisons.

Tout d'abord, même si certains conjugués possèdent une excellente affinité pour la cible et une bonne capacité d'inhibition du complexe TA dans le système simplifié *in vitro*, il est possible qu'*in vivo* ces

activités ne soient pas suffisantes pour induire l'expression de la toxine. Par ailleurs, la sélectivité est un autre paramètre à étudier. Nous avons obtenu des résultats prometteurs en présence de larges excès d'autres acides nucléiques compétiteurs, mais il est possible que certaines molécules puissent interagir avec d'autres ARN intracellulaires, même sans induire d'effet biologique. De plus, toute la difficulté de ces tests est que la cible considérée est complètement nouvelle et nous ne disposons pas de contrôle positif, c'est-à-dire une petite molécule capable d'activer artificiellement le système TA exprimé par le système rapporteur. Enfin, ce test étant réalisé directement chez une bactérie, il est possible que des problèmes de pénétration cellulaire soient rencontrés.

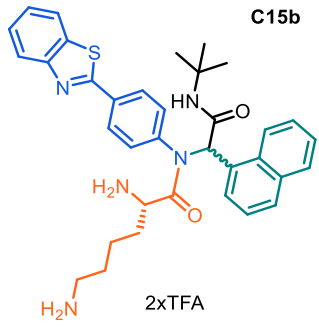
#### II.5.4 Evaluation de l'activité antimicrobienne *in vivo*

La capacité des molécules synthétisées à inhiber la croissance bactérienne a été évaluée sur différentes souches bactériennes dont certaines font partie du groupe ESKAPE. Ce groupe rassemble les six bactéries pathogènes couramment associées à la résistance aux antimicrobiens et qui appartiennent à la liste des pathogènes prioritaires définis par l'OMS (cf. *Chap.I.1*)

Par le biais d'un prestataire, la plateforme BAC-SCREEN de Marseille, un criblage de l'ensemble des molécules synthétisées a été réalisé par le Dr. Véronique SINOUE en évaluant tout d'abord le pourcentage d'inhibition de la croissance bactérienne à une concentration de 10 µM. Pour ce faire, l'activité *in vivo* est estimée par lecture de la DO<sub>600nm</sub> après incubation en présence/absence des composés. Le pourcentage d'inhibition de la croissance bactérienne est calculé selon la formule :

$$\% \text{ Inhibition} = 100 - \frac{DO_{\text{avec composé}} \times 100}{DO_{\text{sans composé}}}$$

L'activité de l'ensemble des composés a été évaluée sur les souches de bactérie à Gram positif : *S. aureus* (CIP 76.25) et de bactéries de Gram négatif : *K. pneumoniae* (CIP 82.91), *E. coli* (CIP 54.81), *P. aeruginosa* (CIP 100720), *A. baumannii* (ATCC 19606), *Enterobacter spp* (ATCC 13048) à une concentration de 10 µM. De ce criblage, seule la molécule **C15b** a montré des activités intéressantes et ces activités sont résumées dans le *tableau 28*. Pour ce composé, la CMI ou concentration minimale inhibitrice a été déterminée. La CMI est la plus petite concentration d'antibiotique qui inhibe toute croissance visible d'une souche bactérienne après 18-24 heures de culture à 37°C. Cette valeur caractérise l'effet bactériostatique d'un antibiotique et permet de déterminer si une souche est sensible à l'antibiotique testé.

**Tableau 28 :** Activités *in vitro* du composé **C15b** sur 6 souches bactériennes. nd = non déterminé.


Souches	Bactérie à Gram	% Inhibition (à 10 µM)	CMI (µM)
<i>S. aureus</i> (CIP 76.25)	G+	<b>95,4</b>	10
<i>K. pneumoniae</i> (CIP 82.91)	G-	51,8	>40
<i>E. coli</i> (CIP 54.81)	G-	<b>99,6</b>	40
<i>P. aeruginosa</i> (CIP 100720)	G-	3,2	nd
<i>A. baumannii</i> (ATCC 19606)	G-	29,8	nd
<i>Enterobacter spp</i> (ATCC 13048)	G-	8,4	nd

Pour les 6 souches évaluées, le composé **C15b** inhibe la croissance bactérienne à plus de 50% pour trois d'entre elles, qui est de 51,8 % pour *K. pneumoniae*, de 95,4 % pour *S. aureus* et de 99,6 % pour *E. coli*. Ces différents résultats révèlent que la molécule peut cibler des bactéries à Gram positif comme à Gram négatif. Pour ces trois souches, la CMI a été déterminée et une sensibilité intéressante est retrouvée pour la souche de *S. aureus* avec une CMI de 10 µM.

Par ce criblage, il est intéressant de noter que les molécules se rapprochant de la structure de **C15b** ont eu des activités inhibitrices de la croissance bactérienne inférieures à 36 % que ce soit pour le changement du fragment amine 2-(4-aminophényl)benzothiazole par un motif naphtyridine *N*- ou *O*-alkylé (**C20b** et **C24b**) ou par la nucléobase artificiel « S » (**C14b**) soit par le changement du fragment aldéhyde par le motif 2-DOS (**C30b**) ou glucosamine (**C40b**).

Pour approfondir la sensibilité du composé **C15b** envers *S. aureus*, l'activité pour différents isolats de la souche *S. aureus* a été évaluée dont certaines sont résistantes à la néomycine (SA010) ou résistante à la méticilline (SA003, 09 et 12) (**tableau 29**). Pour comparaison, la sensibilité de ces différentes souches à la néomycine a été indiquée dans le tableau.

**Tableau 29 :** Activité *in vitro* du composé **C15b** pour différents isolats cliniques de *S. aureus* :

Isolat clinique de <i>S. aureus</i>	% Inhibition (à 10 µM)	CMI (µM)	Sensibilité de la néomycine
<i>S. aureus</i> (CIP 76.25)	95,4	10	2 µg/mL
SA003 (Méticilline R)	nd	20	8 µg/mL
SA009 (Méticilline R)	65,8	<b>20 (16,4 µg/ml)</b>	>128 µg/mL
SA010	nd	10	4 µg/mL
SA012 (Méticilline R)	nd	10	8 µg/mL
SA016	nd	20	4 µg/mL
SA020	nd	10-20	4 µg/mL

L'ensemble des résultats montre une efficacité du même ordre de grandeur quel que soit le niveau de sensibilité à la néomycine des isolats/souches testés ou des isolats résistant à la méticilline. Ces

résultats encourageants ne permettent pas en revanche de montrer le mécanisme d'action. Néanmoins, il est important de noter que le composé **C15b** est un des meilleurs ligands d'ARN identifié lors de l'étude des constantes de dissociation et a montré une forte activité antibiotique lors de l'étude *in vivo* pour l'inhibition du système Toxine-Antitoxine. Il sera donc particulièrement intéressant d'étudier plus en détail le mécanisme d'action en évaluant notamment sa capacité à inhiber la synthèse des protéines chez la bactérie de la même manière que plusieurs classes d'antibiotiques ligands d'ARN.

### II.5.5 Conclusion de l'évaluation biochimique

En conclusion de cette partie, nous avons entrepris l'évaluation biochimique de la collection de 61 molécules synthétisées dont 6 composés ont été évalués sous forme de diastéréoisomères séparés. Cette évaluation biochimique s'est effectuée sur le système TA de type I *AapA1/IsoA1* présent chez la bactérie *Helicobacter pylori*.

Dans une première partie, nous avons évalué l'affinité *in vitro* de ces molécules sur une partie de la séquence de l'ARN *AapA1* qui correspond à une tige-boucle impliquée dans la reconnaissance boucle-boucle du système TA *AapA1/IsoA1*. Les résultats sont les suivants :

- 13 composés  $K_D > 1$  mM.
- 23 composés  $2 \mu\text{M} < K_D < 1$  mM.
- **22 composés  $K_D < 2 \mu\text{M}$ .**

Sur les meilleurs composés ( $K_D < 2 \mu\text{M}$ ), la sélectivité a été évaluée en présence d'ARNt et d'ADN en excès (100 éq.). A travers cette évaluation, la plupart des composés possèdent une excellente sélectivité en présence d'ARNt et d'ADN.

Dans un deuxième temps, nous avons étudié l'interaction de cette tige-boucle *AapA1* avec la tige-boucle correspondante sur l'ARN *IsoA1*. Par cette construction, nous avons mis au point un test *in vitro* pour évaluer la capacité d'inhibition du complexe boucle-boucle formé par les deux ARN induisant ensuite la formation d'un duplex étendu. Grâce à trois études complémentaires, nous avons pu évaluer la cinétique de la formation du duplex étendu (i) par une mesure par anisotropie de fluorescence (ii) par une mesure de fluorescence d'un couple FRET (iii) par des expériences par gel d'électrophorèse. Ces trois études nous ont permis de définir des conditions optimales pour évaluer par la suite l'activité de nos molécules sur l'inhibition de ce système. Les conditions choisies sont : tampon HEPES à 20 mM, pH 7.4, 20 mM NaCl, 140 mM KCl, et 3 mM  $\text{MgCl}_2$  et un temps d'incubation de 2h à une température de 37°C. Au moyen du test d'inhibition utilisant la mesure par l'anisotropie de fluorescence, nous avons commencé par étudier 7 des meilleurs composés affins envers l'ARN *AapA1*. Lors de ce premier test de criblage permettant d'évaluer la capacité des molécules à inhiber la formation du complexe boucle-



boucle et de quantifier l'inhibition par une mesure des valeurs de  $IC_{50}$ , la molécule **C38b** a montré un effet d'inhibition sur le complexe *AapA1/IsoA1* avec un  $IC_{50}$  de 1,14  $\mu$ M. Ce résultat étant issu d'une étude préliminaire sur 7 composés seulement, il nous semble encourageant. D'autres composés, tels que **C14b**, **C25b** et **C34b** semblent inhiber la formation du complexe TA, mais ces résultats ont besoin d'être confirmés car la gamme de concentration utilisée n'était pas adaptée à une évaluation précise de l'inhibition. Il est à noter que le composé **C15b** a montré un effet différent des autres composés testés puisqu'il induit une augmentation d'anisotropie qui suggère la formation ou la stabilisation d'une structure différente de la tige-boucle ou du duplex étendu. Des tests supplémentaires d'optimisation sont en cours ainsi que le criblage de l'ensemble des molécules synthétisées.

Dans un troisième temps, une partie des molécules a été évaluée pour son activité d'inhibition par un test basé sur l'utilisation d'un système rapporteur chez la bactérie *E. coli* et impliquant les séquences entières des deux ARN étudiés contrairement au test précédent qui utilise deux fragments en tige-boucle d'environ 27 nucléotides.

Enfin, dans une dernière partie, nous avons évalué l'activité antimicrobienne *in vivo* sur 6 souches bactériennes pour l'ensemble des molécules synthétisées. De manière intéressante, la molécule **C15b** qui est un des meilleurs ligands d'ARN et un des composés ayant donné un effet lors de l'étude d'anisotropie, ressort avec des activités intéressantes non seulement sur des souches sensibles mais aussi résistantes aux antibiotiques actuellement disponibles. Par ailleurs, son effet antibiotique a été confirmé sur les bactéries utilisées lors du test utilisant le système rapporteur. Il sera donc très important d'étudier en détail le mécanisme d'action de ce composé.

Ces résultats préliminaires d'évaluation des composés synthétisés sont très prometteurs même si l'activation de la toxine chez la bactérie n'a pas encore été obtenue. En effet, nous avons pu identifier des composés possédant une excellente affinité pour l'ARN, une très bonne sélectivité et la capacité d'inhiber la formation du complexe Toxine-Antitoxine *in vitro*. Le composé **C15b** a par ailleurs montré une très bonne activité antibiotique sur plusieurs souches bactériennes.

## II.6 Conclusion et perspectives

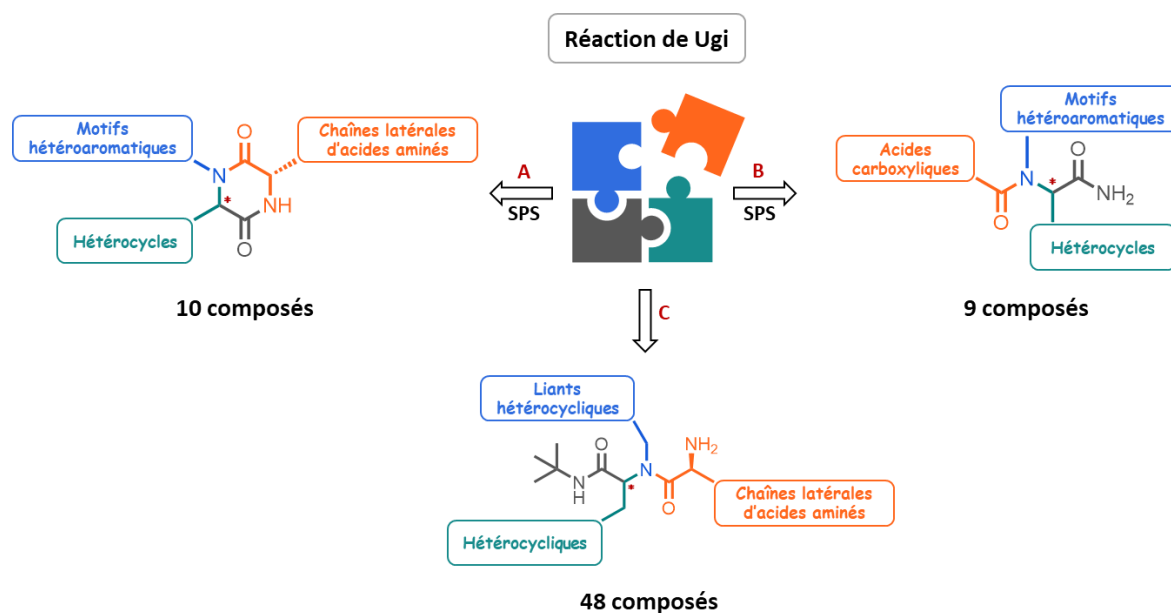
L'objectif principal de ce travail de thèse était la synthèse d'une librairie de ligands d'ARN conçus pour interagir avec un système Toxine-Antitoxine de type I : *AapA1/IsoA1* chez la bactérie *Helicobacter pylori*. A partir de cette librairie, le but était d'identifier des ligands d'ARN capables d'activer artificiellement le système TA pour induire la production de la toxine.

Le projet s'est articulé selon **trois axes majeurs** (i) la préparation des composants de la réaction de Ugi (ii) la synthèse de trois séries de molécules et (iii) l'évaluation biochimique sur le complexe boucle-boucle *in vitro* et chez la bactérie.

La préparation de la nouvelle collection de ligands d'ARN a porté sur la synthèse de composés issus de la réaction de Ugi qui est une réaction multicomposants permettant d'obtenir des composés de squelette pseudo-dipeptidique (ou  $\alpha$ -aminoacylamide). Cette réaction a été réalisée sur support solide et en solution.

Pour réaliser l'ensemble de ces séries par réaction de Ugi, la préparation des 3 fragments de type amine, aldéhyde et isocyanure a été nécessaire (cf. *figure 59* ou *annexe 4*) alors que nous avons utilisé des acides carboxyliques commerciaux (acides aminés). Pour la synthèse sur support solide polymérique, la fonctionnalisation de la résine comportant la fonction isocyanure a été réalisée avec succès sur la résine Rink amide MBHA et la résine Rink amide. En parallèle, les motifs choisis pour interagir spécifiquement avec l'ARN ont été synthétisés menant à **10 composants** portant la fonction amine sur des motifs hétéroaromatiques ainsi que **2 composants** portant la fonction aldéhyde sur des motifs aminoglycosides et analogues. Ces synthèses comportaient de 2 à 8 étapes et les produits souhaités ont été obtenus avec succès. Lors de ce travail, j'ai eu l'opportunité d'encadrer deux stages de master (M1 et M2) avec comme objectif la préparation de quelques-uns de ces synthons sur plus grandes quantités.

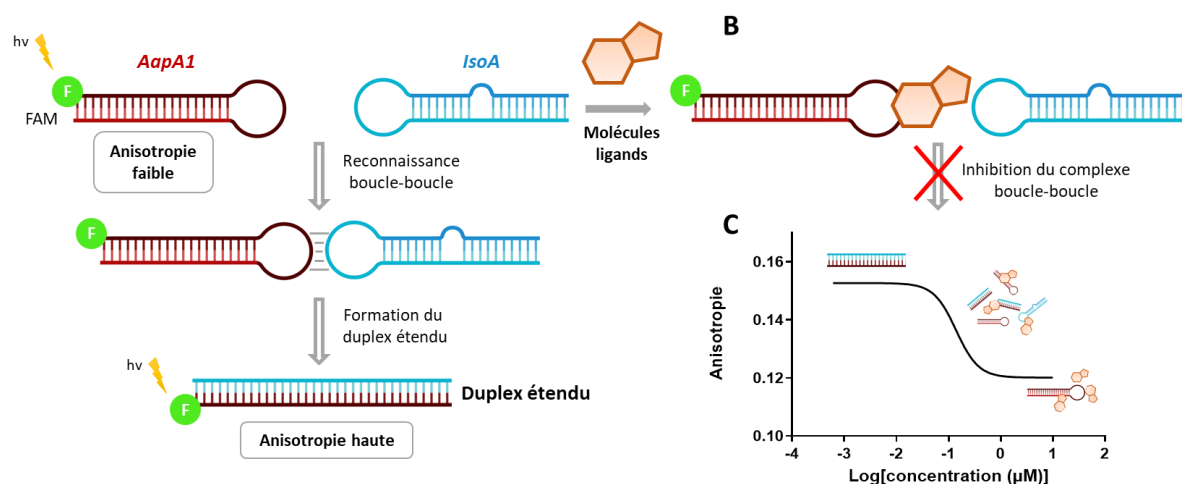
Le deuxième axe du projet a été ensuite consacré à la synthèse d'**une librairie de 67 composés** sous forme de mélanges de deux diastéréoisomères ou d'énantiomères (59 composés) mais également de diastéréoisomères séparés (8 composés) (*figure 88* et *annexe 3*). La réaction de Ugi sur support solide nous a permis d'explorer des composés de structure 2,5-dicétopipérazine (**Série A**) mais aussi de structure linéaire (**Série B**). Concernant, la **série C**, réalisée en solution au moyen de l'isocyanure *tert*-butyle, un ensemble de **48 molécules** a été préparé : le motif naphthalène a été utilisé donnant la **série C-1 (26 molécules)**, le motif 2-désoxystreptamine a permis d'obtenir la **série C-2 (10 molécules)** et une série plus large constitue la **série C-3 (12 molécules)**.



**Figure 88 :** Bilan des trois séries réalisées au cours de ce manuscrit en utilisant la réaction multicomposants de Ugi. Les séries A et B ont été préparées sur support solide polymérique (SPS). La série C est issue de réactions en solution.

A partir de cette bibliothèque, plusieurs études biochimiques et biologiques ont été réalisées. L'affinité des composés a été mesurée dans un premier temps sur une séquence de l'ARN *AapA1* qui est l'une des tige-boucles impliquées dans la reconnaissance boucle-boucle du complexe Toxine-Antitoxine *AapA1/IsoA1*. Ces études d'affinité ainsi que les études de sélectivité (présence d'excès d'ARNt ou d'ADN) ont permis de mettre en avant 20 composés ayant une affinité ( $K_D$ ) inférieure à 2  $\mu\text{M}$ . Parmi ces composés, la plupart possède une excellente sélectivité en présence d'ARNt et d'ADN. Ainsi, l'ensemble de ces résultats montre qu'une **approche basée sur la conception de ligands multifonctionnels** élaborés avec des motifs spécifiquement choisis pour leur affinité envers l'ARN, permet d'obtenir des ligands très affins envers un ARN, dans notre cas, de structure tige-boucle.

Pour évaluer la capacité des meilleurs composés à interférer avec l'interaction ARN-ARN Toxine-Antitoxine, un test d'inhibition a été mis en place au sein du laboratoire. Pour cela, deux séquences de tige-boucle constituant le système TA *AapA1/IsoA1* ont été choisies (*figure 89*) et étudiées. Différents tests biochimiques *in vitro* ont été explorés (i) un test utilisant de l'anisotropie de fluorescence (*figure 89*) (ii) un test utilisant un couple de fluorophore/quencher (FRET) et (iii) des expériences de retardement sur gel d'électrophorèse. Ces trois études nous ont permis d'étudier la formation du complexe boucle-boucle en suivant l'apparition du duplex étendu entre *AapA1/IsoA1* sur les séquences sélectionnées.



**Figure 89 :** Représentation schématique de test basé sur la mesure par anisotropie de fluorescence. B) Cas de figure d'une molécule inhibitrice du complexe boucle-boucle. C) Représentation du résultat des valeurs d'anisotropie obtenues dans le cas d'inhibition du complexe boucle-boucle. F=Fluorescéine = FAM.

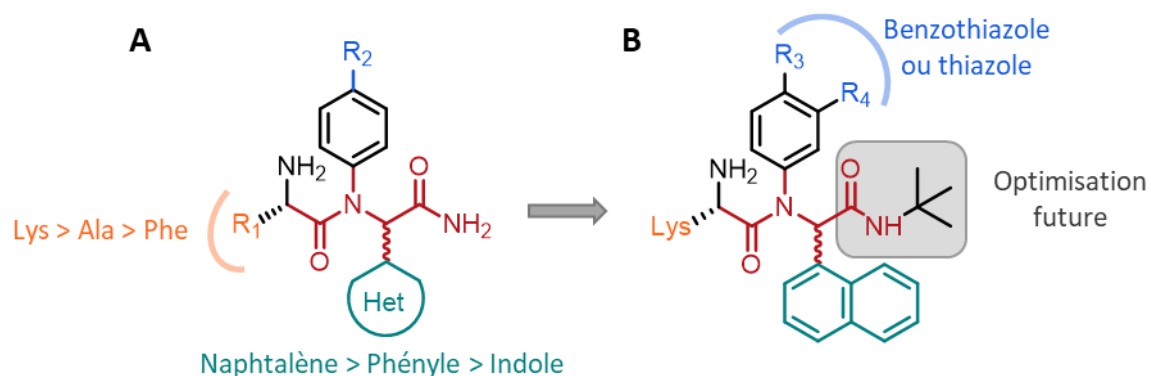
Le test d'inhibition de l'interaction boucle-boucle utilisant la mesure d'anisotropie de fluorescence, nous a permis de commencer **la mise au point d'un test de criblage** répondant aux objectifs de ce projet (*figure 89, A*). Ce test innovant par le fait que peu d'études ont été réalisées sur des interactions boucle-boucles est encourageant. Actuellement, 7 composés (**C14b**, **C15b**, **C20b**, **C25b**, **C28b**, **C3b8** et **C34b**) ont été étudiés et leur étude doit être poursuivie notamment en adaptant la gamme de concentration. Un des résultats encourageants a été obtenu pour le composé **C38b** avec une valeur d'inhibition ( $IC_{50}$ ) de 1,35  $\mu\text{M}$ . Ce composé est également un des meilleurs composés en termes d'affinité avec l'ARN *AapA1*. Ces résultats préliminaires nécessitent d'être donc confirmés et ce premier criblage devra être étendu à l'ensemble de la librairie. Par la suite, les résultats pourront être confirmés par des expériences de retardement sur gel d'électrophorèse qui ont d'ores et déjà été mis au point.

Parallèlement, l'équipe du Dr Darfeuille a mis en place un test de criblage sur *E. coli* basé sur un système rapporteur contenant le système TA *AapA1/IsoA1*. Pour l'instant, aucune des molécules n'a montré un effet d'inhibition mais le composé **C15b** a montré une excellente activité antibiotique.

Le criblage de l'ensemble des composés sur plusieurs souches bactériennes appartenant pour certains au groupe ESKAPE, a montré des activités intéressantes sur trois d'entre elles (*K. pneumoniae*, *S. aureus* et *E. coli*.) et le composé **C15b** se confirme en tant que composé antibiotique. Nous envisageons maintenant l'étude du mécanisme d'action intracellulaire.

Grâce aux études effectuées, nous avons d'ores et déjà pu dégager des relations structure-activité (*figures 90 et 91*). La combinaison de motifs simples (phényle, naphthalène ou indole) tels que ceux utilisés dans l'étude préliminaire pour la synthèse sur phase solide (**15**, **79**, **80**, **81** et **82**) a donné lieu à des composés présentant des résultats d'affinité prometteurs, en particulier pour deux d'entre eux

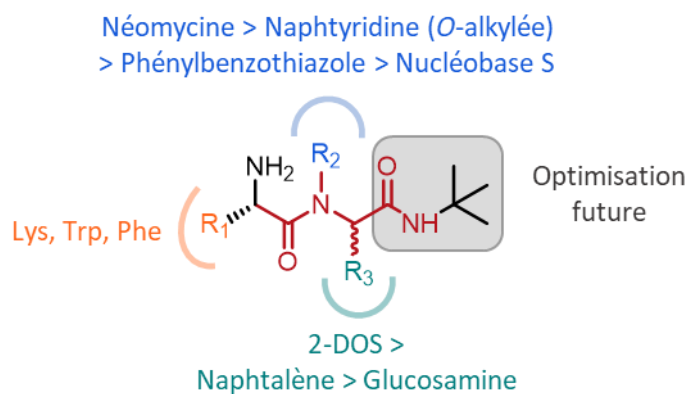
(15 et 81) indiquant que les produits de Ugi pouvaient être un excellent point de départ pour la préparation de ligands d'ARN (*figure 90, A*). Certains des composés 2,5-DKP de la série A ont montré une certaine affinité qui reste modeste, dans le bas micromolaire. Les composés de la Série B n'ont pas montré d'activité mais la présence de substituants peu fonctionnalisés peut peut-être expliquer ce manque d'activité.



**Figure 90 :** Relations structure-activité entre les composés issus de la phase solide (A) et de la série C-1 (B).

En revanche, les conjugués issus de la synthèse en solution (série C-1) et constitués de la combinaison du motif naphthalène et lysine avec les motifs phénylthiazole (**C14b**) et phénylbenzothiazole (**C15b**), nous a permis d'obtenir des affinités de l'ordre du nanomolaire (*figure 90, B*) de même que les composés avec le motif naphtyridine et un acide aminé basique tel que lysine ou tryptophane (**C20b**, **C24b** et **C25b**) dans la série C-1. On peut noter aussi que l'utilisation de la néomycine (**C28b**) a fortement amélioré l'affinité ( $K_D$  de 42 nM).

De manière intéressante, dans la série C-2, le motif 2-DOS a été conjugué aux meilleurs motifs hétéroaromatiques identifiés dans la série C-1 (naphtyridine, benzothiazole et thiazole) et à la lysine ou la phénylalanine, ce qui nous a permis de nous rapprocher de l'affinité du composé **C28b** notamment dans le cas du composé **C38b** qui présente un  $K_D$  de 67 nM. La glucosamine employée dans la série C-3 en revanche a donné des résultats intéressants mais des affinités qui restent de l'ordre du micromolaire dans la plupart des cas. Il semble donc important d'avoir un motif de type aminoglycoside ou 2-désoxystreptamine, un hétérocycle aromatique et un acide aminé basique ou aromatique, afin d'obtenir les meilleurs résultats d'affinité (*figure 91*).



**Figure 91** : Relation structure-activité de la série C.

Un des meilleurs ligands identifiés à ce jour, **C38b**, est aussi le meilleur inhibiteur du complexe TA *AapA1/IsoA1 in vitro*. Ceci suggère qu'une optimisation à partir de ce composé pourrait nous mener à des inhibiteurs efficaces chez la bactérie également. Parmi les approches d'optimisation que nous envisageons, la modification du motif isocyanure n'a pas encore été étudiée pour l'instant mais elle est facilement accessible. L'introduction dans cette position d'un nouveau motif d'interaction pour l'ARN cible pourrait nous mener à l'activité biologique souhaitée. Parallèlement, l'étude sur un effet possible de l'orientation des motifs induit par les deux diastéréoisomères est envisagée.



---

# **Chapitre III - Synthèse de nouveaux ligands d'ARN ciblant l'ARN TAR du VIH**

---





## III.1 Introduction

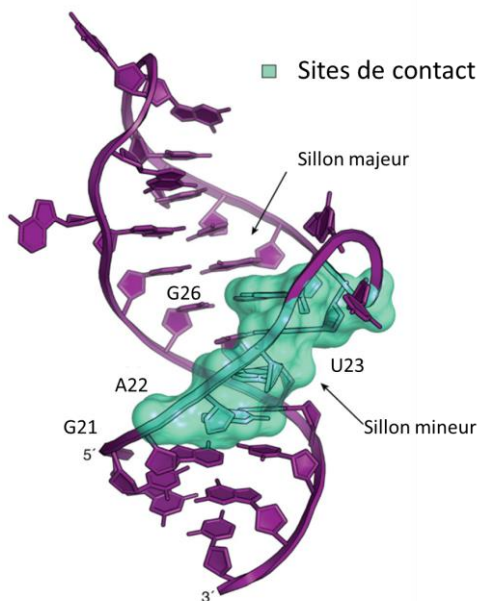
Parallèlement au projet principal de ce manuscrit, j'ai eu l'opportunité de participer à plusieurs projets dont la synthèse de nouveaux ligands d'ARN ciblant l'ARN TAR du VIH qui a fait l'objet d'une publication intitulée « **Unveiling RNA - Binding Properties of Verapamil and Preparation of New Derivatives as Inhibitors of HIV-1 Tat-TAR Interaction** » dans le journal **ChemPlusChem** en 2020 avec Serena De Piccoli, Marc Gaysinski, Cécile Becquart, Stéphane Azoulay, Audrey Di Giorgio et Maria Duca.

Une description du projet sera détaillée dans un premier temps puis les résultats obtenus seront décrits après la publication.

### III.1.1 Description du projet

Comme mentionné dans le **Chapitre I**, l'ARN du VIH-1, notamment l'ARN TAR, est une cible prometteuse pour le développement de nouvelles thérapies anti-VIH, mais aussi comme modèle idéal pour valider la découverte de ligands originaux d'ARN.

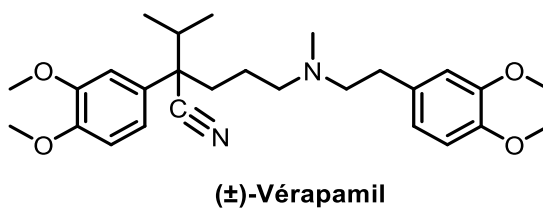
Pour rappel, l'ARN **TAR** pour « Trans-Activation Response element » est essentiel dans la réplication efficace du VIH. Ce fragment d'ARN de 59 nucléotides est situé dans les régions 3'-UTR de chaque ARN viral et présente une structure en tige-boucle avec un renflement à trois nucléotides (U<sub>23</sub>, C<sub>24</sub> et U<sub>25</sub>). L'ARN TAR interagit avec la protéine virale Tat et avec le facteur d'élongation P-TEFb lui-même constitué de la cycline T1 et de la kinase 9 dépendante de la cycline (CDK9), formant ainsi un complexe qui contrôle la transcription virale. La spécificité de l'interaction de Tat est médiée par une région riche en arginine qui reconnaît la région autour du renflement de TAR. Cette interaction Tat-TAR est essentielle à la réplication virale. Ainsi, des ligands d'ARN capables d'empêcher la formation de cette interaction pourraient inhiber la réplication virale<sup>213</sup> (*figure 92*).



**Figure 92 :** Représentation du modèle utilisé lors des études sur l'ARN TAR du VIH-1. Les sites d'interactions souvent rencontrés sont représentés par la surface verte<sup>213</sup>.

Ces dernières décennies, de nombreux ligands de TAR ont été identifiés, essentiellement par criblage, dont certains récemment comme par exemple la molécule thiénopyridine<sup>105</sup> (cf. *figure 34*, page 70) découverte par un criblage de 20000 molécules ou la molécule diméthylamiloride<sup>108</sup> (DMA-169) (cf. *figure 34*, page 70) issue d'un criblage d'une large chimiothèque suivi d'une optimisation. Dans ce contexte, notre équipe a voulu tester une bibliothèque de composés présents au laboratoire (188 molécules) dans le but d'identifier de nouveaux motifs capables de se lier à des ARN et en particulier l'ARN TAR. Ces motifs, même s'ils présentent une faible affinité, pourront être associés à d'autres composés connus pour interagir avec des ARN, de façon à obtenir des ligands affins et sélectifs. Nous avons déjà utilisé cette approche multimodal e pour découvrir des molécules actives biologiquement<sup>141</sup>.

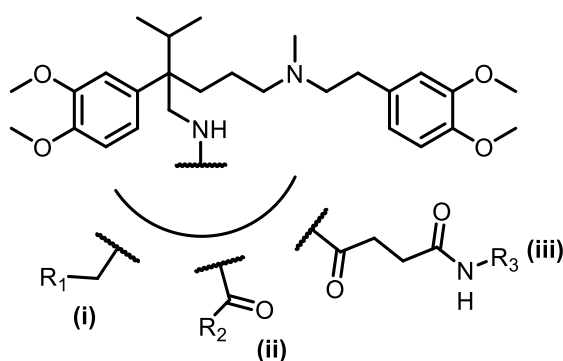
Le criblage des 188 molécules sur un fragment d'ARN TAR de 27 nucléotides marqué par un groupement fluorescent, a permis d'identifier plusieurs composés dont un de structure originale pour un ligand d'ARN, le vérapamil (*figure 93*), actuellement utilisé en clinique comme antiarythmique en tant que bloqueur des canaux calciques.



**Figure 93 :** Structure du vérapamil.

Nous avons donc décidé de partir de ce composé commercial possédant une structure originale pour un ligand d'ARN et de le conjuguer avec d'autres motifs identifiés lors de ce criblage initial (à savoir la phéothiazine, le naphthalène, la quinolone, le benzimidazole, l'indole, le thiazole, l'acide 3,5dibromosalicylique et le 4-diméthylaminobenzaldehyde) ainsi qu'avec des motifs connus pour la liaison ARN (tels que le benzothiazole, le phénythiazole, l'imidazole et des acides aminés : lysine, histidine, phénylalanine et arginine) ; le but étant de découvrir des ligands originaux, affins et sélectifs.

C'est dans ce cadre que j'ai participé à l'élaboration de ces conjugués dont la préparation a nécessité trois voies de fonctionnalisation différentes selon le motif introduit sur le vérapamil (*figure 94*). Le couplage des motifs sur le vérapamil a nécessité dans un premier temps la réduction de la fonction nitrile en amine primaire qui sert de point d'encrage pour la condensation des autres motifs. Cette dernière a ainsi été couplée (i) avec des motifs aldéhydiques via des aminations réductrices, (ii) avec des acides carboxyliques en utilisant des réactifs de synthèse peptidique et (iii) avec des motifs portant une fonction amine via l'anhydride succinique. Ces trois méthodes de couplage sont détaillées dans l'article ci-après.



**Figure 94 :** Modifications apportées au composé vérapamil par trois fonctionnalisations différentes : (i) amination réductive ; (ii) couplage peptidique et (iii) insertion d'un bras espaceur pour réaliser un couplage peptidique.  $R_1$  = motifs issus d'aldéhydes ;  $R_2$  = motifs issus d'acides carboxyliques ;  $R_3$  = motifs issus d'amines.

Une fois l'ensemble des conjugués synthétisés, plusieurs tests biochimiques ont été réalisés afin d'évaluer ces composés. Dans un premier temps, l'affinité ( $K_D$ ) pour l'ARN TAR a été mesurée puis la sélectivité des composés vis-à-vis d'autres constituants nucléiques largement présents dans les cellules (ARNt et ADN) a été évaluée. L'évaluation de l'inhibition de l'interaction Tat/TAR ( $IC_{50}$ ) a ensuite été réalisée sur les ligands les plus affins et les plus sélectifs. Enfin, une dernière étude a été menée afin d'étudier plus précisément le mécanisme d'interaction moléculaire du meilleur ligand par rapport au composé parent non modifié (Vérapamil-réduit). Cette étude a été réalisée par de la spectroscopie de RMN.

### **III.1.2 Publication “Unveiling RNA - Binding Properties of Verapamil and Preparation of New Derivatives as Inhibitors of HIV-1 Tat-TAR Interaction”**

ChemPlusChem 2020, 85, 207–216. <https://doi.org/10.1002/cplu.201900650>



by an arginine-rich region that recognizes the bulged region of TAR.<sup>[8]</sup> The inhibition of Tat/TAR interaction should thus induce the inhibition of viral replication and a various RNA binders able to block this interaction have been reported.<sup>[6]</sup> TAR RNA thus represents particularly relevant model system to study ligand binding to RNA as well as a promising target for the discovery of compounds with anti-HIV activity.<sup>[9]</sup>

During recent years, various small-molecule RNA binders able to target biologically relevant RNAs have been identified.<sup>[1a,c,10]</sup> One major example is represented by RNA-interacting compounds acting as pre-mRNA splicing modifiers of SMN2 (survival motor neuron 2 protein). Some of these ligands recently found clinical application in the treatment of Spinal Muscular Atrophy (SMA).<sup>[11]</sup> Very specific RNA binders targeting the production of oncogenic microRNAs in cancer cells and *in vivo* have also been reported.<sup>[12]</sup> In this case, an approach called InfoRNA was developed to identify lead binders using a statistical prediction of compounds affinity and selectivity based on a previous screening of a large library of compounds against a library of RNA motifs.<sup>[13]</sup> This allowed for the discovery of compounds bearing an exceptional specificity for miR-96 precursor or for miR-210 precursor both in cells and in animals models.<sup>[10a,12c]</sup>

We also discovered various RNA ligands upon combination of different RNA binding domains carefully selected to bring both affinity and selectivity for the target.<sup>[14]</sup> For example, we conjugated artificial nucleobases, that are able to interact selectively with RNA base pairs, with basic amino acids that can bring affinity for an efficient targeting of TAR RNA *in vitro* and in infected cells.<sup>[15]</sup> A similar approach was used for the preparation of C-nucleosides targeting prokaryotic A-site RNA.<sup>[16]</sup> We also conjugated known antibiotics, such as aminoglycosides, with artificial nucleobases and amino acids leading to multimodal RNA ligands able to selectively target oncogenic miRNAs *in vitro* and in cancer cells overexpressing these miRNAs.<sup>[14a]</sup> Concomitantly to the focused design of new RNA ligands, we performed screening of larger libraries of compounds against oncogenic miRNAs.<sup>[17]</sup> This led to the identification of an array of chemical scaffolds bearing favorable properties in terms of binding and selectivity *in vitro* but also inside cancer cells.

Here, we describe the screening of an in-house library of commercial compounds that allowed for the identification of verapamil, a phenylalkylamine calcium channel blocker currently in clinical use as an antiarrhythmic compound (Figure 1B), as a potential scaffold for the preparation of new RNA ligands. As mentioned above, the conjugation of RNA ligands with other RNA binding domains led to promising and selective RNA binders. We thus decided to prepare a reduced form of verapamil and to conjugate it with fragments identified as RNA binders during the screening of the library and with known RNA-binding moieties led to the preparation of a series of verapamil analogs and to the improvement of TAR RNA affinity and selectivity. Noteworthy, the conjugation of verapamil to an indole fragment through reductive amination led to a compound (**2h**) which is able to efficiently inhibit Tat/TAR interaction *in vitro* with 18.8  $\mu\text{M}$  IC<sub>50</sub>. This represents a very

promising inhibition activity compared to small-molecule RNA ligands already reported in the literature.<sup>[9,18]</sup> The detailed study of the molecular mechanism of interaction of **2h** with TAR RNA led to important insights about the chemical features that are relevant for binding and for the design of future efficient TAR binders and potential antiviral compounds. Verapamil has never been considered for RNA binding and this work represents the first example of verapamil chemical modification applied to RNA targeting thus opening new perspectives in the RNA targeting field.

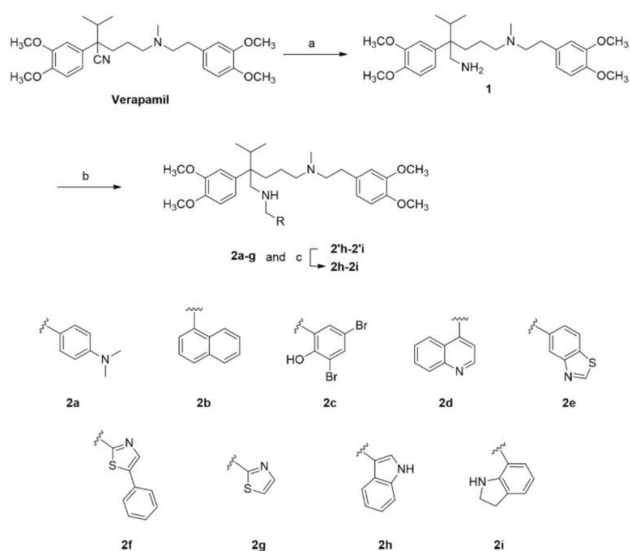
## Results and Discussion

### Library screening

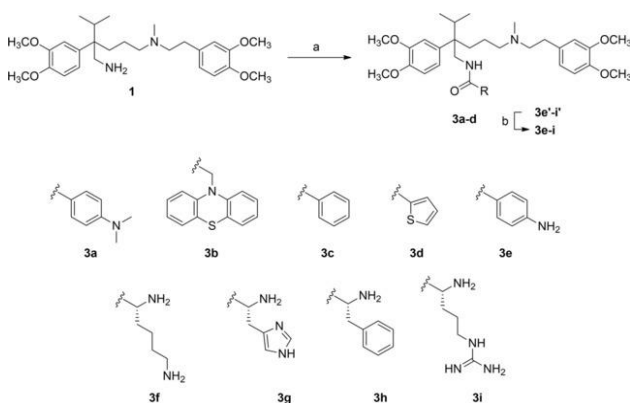
The library of compounds and fragments chosen for this study belongs to an in-house collection of commercial compounds and contains 188 molecules (see Table S1 in the Supporting Information for the complete list) whose 90% can be considered as fragments (MW < 300, ClogP, H-bond donors and acceptors and n° of rotatable bonds < 3).<sup>[19]</sup> The remaining 10% are molecules bearing higher molecular weight but that could present an interest as scaffolds for the preparation of new RNA binders. We also decided to add to this library some antibiotics that are efficient and non-selective RNA binders, such as aminoglycosides and tetracyclines that could function as positive controls. The aim of the initial screening was to identify compounds bearing even modest affinity for TAR RNA in order to later improve their binding and biological activity upon chemical modification as usually performed in a fragment-based approach. Hits growing (addition of substituents) and linking (combination with other hits) could lead indeed to relevant improvements of affinity and of selectivity allowing for the discovery of new scaffolds and for the design of RNA ligands with potential antiviral activity.

First, we screened the library measuring the affinity toward a 27-mer TAR fragment labeled in 5' position with Alexa<sup>488</sup>. This kind of fluorescence-based assay relies on the fact that binding of an efficient ligand to the labeled target induces a change in fluorescence intensity that is dependent on the concentration of ligand and it has been largely validated in the literature.<sup>[20]</sup> We screened fragments at 10  $\mu\text{M}$  and 100  $\mu\text{M}$  leading to the identification of 9 compounds showing affinity (at least 20% fluorescence variation) at both concentrations (guanine, thio-guanine, phenothiazine, neomycin, oxytetracycline, 1,6-dibromo-2-hydroxynaphthalene-3-carboxylic acid, cinchonidine, verapamil and apramycin, Figure S1 in the Supporting Information) and 9 compounds showing some affinity only at 100  $\mu\text{M}$  concentration (2-aminobenzimidazole, 5-aminomethylindole, 2-aminothiazole, 3,5-dibromosalicylic acid, 3,5-dichlorosalicylic acid, 3-acetylindole, 4-dimethylamino benzaldehyde, 2,6-diaminopurine and puromycin, Figure S2). Among the strongest binders, neomycin, oxytetracycline, apramycin and puromycin are known RNA binders able to bind prokaryotic ribosomal RNA and impair protein synthesis in bacteria. Phenothiazine, cinchonidine and naphthalene derivative are known intercalating

agents and have been previously studied for RNA binding.<sup>[21]</sup> Purines, such as guanine and thioguanine have also been studied as RNA binding motifs.<sup>[14c,22]</sup> The remaining compound, verapamil, revealed to have interesting affinity for TAR and this pharmacophore has never been explored for RNA binding. It thus seemed an appropriate starting point for the development of new and original RNA ligands. Indeed, verapamil is a calcium channel blocker currently used clinically for the treatment of arrhythmia and other cardiovascular pathologies.<sup>[23]</sup> We thus decided to take advantage of verapamil to synthesize conjugates with other scaffolds identified during the initial screening (namely phenothiazine, naphthalene, quinolone, benzimidazole, indole, thiazole, 3,5-dibromo-salicylic acid and 4-dimethylaminobenzaldehyde) as well as with known RNA binding domains (such as benzothiazole, phenylthiazole, imidazole and amino acids lysine, histidine, phenylalanine and arginine).



**Scheme 1.** Synthetic procedure for the preparation of reduced verapamil **1** as well as of reductive amination products **2a–i**. Reagents: a) LiAlH<sub>4</sub>, THF, 60 °C, overnight; b) aldehyde, NaBH<sub>3</sub>CN, CH<sub>2</sub>Cl<sub>2</sub>, r.t., overnight; c) TFA, CH<sub>2</sub>Cl<sub>2</sub>, r.t., overnight.



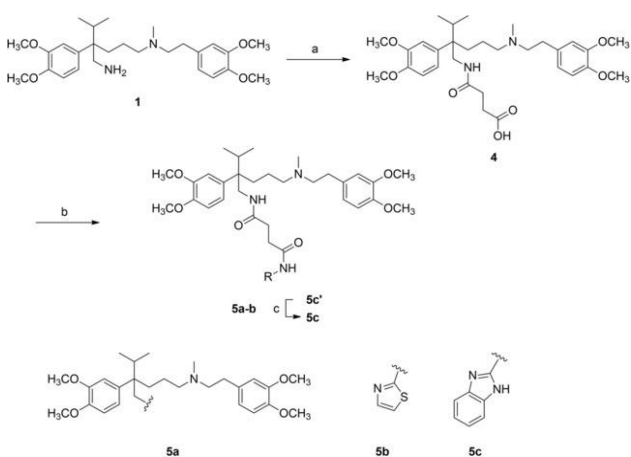
**Scheme 2.** Synthesis of amide-linked conjugates **3a–i**. Reagents: a) carboxylic acid, HBTU, DIPEA, CH<sub>2</sub>Cl<sub>2</sub>, r.t., overnight; b) TFA, CH<sub>2</sub>Cl<sub>2</sub>, r.t., overnight.

## Synthesis of a new series of verapamil analogs

In order to conjugate verapamil to other RNA-binding motifs, we first chose to reduce verapamil nitrile group using LiAlH<sub>4</sub> in THF to obtain desired amine-containing compound **1** in 96% yield as previously reported in the literature (Scheme 1).<sup>[23]</sup> Starting from this compound, we chose three different functionalization routes: (i) reductive amination upon reaction with aldehydes (Scheme 1), (ii) amide formation upon reaction with carboxylic acids (Scheme 2) and (iii) introduction of a carboxyl group on verapamil using succinic anhydride followed by amide formation upon reaction with amines (Scheme 3).

Regarding the reductive amination reaction, we chose to react **1** with aldehyde-containing compounds that demonstrated to have some affinity during fragment screening (Scheme 1): 4-dimethylaminobenzaldehyde (as in **2a**), 1-naphthaldehyde (as in **2b**), 3,5-dibromo-2-hydroxybenzaldehyde (as in **2c**), 4-quinolinecarboxaldehyde (as in **2d**), *N*-Boc-3-formyl-1*H*-indole-1-carboxylate (as in **2'h**). Beside these five aldehyde compounds, we also chose four different heterocyclic moieties containing an aldehyde group and known as interesting moieties for RNA targeting:<sup>[10c]</sup> benzothiazole (as in **2e**), phenylthiazole (as in **2f**), thiazole (as in **2g**) and *N*-Boc-indoline (as in **2'i**). Reductive amination in the presence of these aldehydes was performed using classical conditions (NaBH<sub>3</sub>CN in CH<sub>2</sub>Cl<sub>2</sub>/MeOH 50:50) and led to the formation of compounds **2a–g** and **2'h–2'i** in 52–68% yields. Compounds **2'h** and **2'i** containing a Boc protecting group were further deprotected in the presence of TFA in CH<sub>2</sub>Cl<sub>2</sub> leading to **2h** and **2i** in 94% and 78% yields, respectively.

For the preparation of analogs bearing an amide linkage between verapamil and the modifying moiety, we chose two different strategies. The first one is the direct coupling of carboxylic acids on the amino group of compound **1** (Scheme 2). In this case, we also took advantage of chemical moieties identified during the screening process by coupling 4-dimethylaminobenzoic acid (as in **3a**), 10*H*-phenothiazine-10-



**Scheme 3.** Synthetic procedure for the synthesis of compound **4** bearing a succinamic acid linker and of amide-linked derivatives **5a–c**. Reagents: a) succinic anhydride, CH<sub>2</sub>Cl<sub>2</sub>, r.t., overnight; b) amine, HOBT, EDC, Et<sub>3</sub>N, CH<sub>2</sub>Cl<sub>2</sub>, r.t., overnight; c) TFA, CH<sub>2</sub>Cl<sub>2</sub>, r.t., overnight.



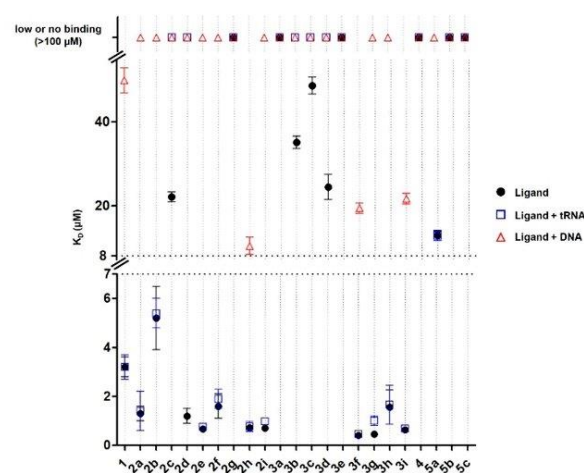
carboxylic acid (as in **3b**), benzoic acid (as in **3c**), 2-thiophenecarboxylic acid (as in **3d**) and 4-*N*-tert-butoxycarbonyl-aminobenzoic acid (as in **3'e**). We also conjugated **1** with *N* $\alpha$ ,*N* $\epsilon$ -di-Boc-L-lysine (as in **3'f**), *N* $\alpha$ ,*N* $\epsilon$ -di-Boc-L-histidine (as in **3'g**), *N* $\alpha$ -Boc-L-phenylalanine (as in **3'h**) and *N* $\alpha$ -Boc-*N* $\gamma$ -(4-methoxy-2,3,6-trimethylbenzenesulfonyl)-L-arginine (as in **3'i**). Amino acids are indeed RNA binding agents as they constitute natural RNA ligands, i.e. peptides. Peptide coupling between compound **1** and these carboxylic acids in the presence of HOBt, EDC and Et<sub>3</sub>N in CH<sub>2</sub>Cl<sub>2</sub> led to the formation of compounds **3a–d** and **3'e–3'i** in 56–86% yields. Compounds **3'e–3'h** containing a Boc group were deprotected in the presence of TFA in CH<sub>2</sub>Cl<sub>2</sub> leading to desired products **3e–3h** in 79–95% yields. Compound **3'i** containing a Mtr and a Boc protecting groups was first deprotected on the *N* $\alpha$  in the presence of TFA in CH<sub>2</sub>Cl<sub>2</sub> then treated with a solution of phenol/TFA 5% w/w to remove the Mtr group of *N* $\gamma$ . These two steps led to **3i** in 95% yield.

Finally, compound **1** was modified with the introduction of a succinamic acid side chain that was introduced upon reaction of **1** with succinic anhydride in CH<sub>2</sub>Cl<sub>2</sub> leading to compound **4** in 95% yield (Scheme 3). **4** was then coupled with compound **1** to form a verapamil dimer (as in compound **5a**) as well as with 2-aminothiazole (as in compound **5b**) in the presence of HOBt, EDC, Et<sub>3</sub>N in CH<sub>2</sub>Cl<sub>2</sub> leading to desired compounds **5a** and **5b** and **5'c** in 27% and 25% yield, respectively. Compound **4** was also coupled with *N*-Boc-2-aminobenzimidazole (as in compound **5'c**) in the presence of HBTU and DIPEA in CH<sub>2</sub>Cl<sub>2</sub> leading to a mixture of **5'c** and **5c**. Addition of TFA in CH<sub>2</sub>Cl<sub>2</sub> led to desired compound **5c** in 25% yield over two steps.

All synthesized compounds were characterized by NMR, HRMS and ultimately purity was verified by HPLC before their biological evaluation.

### Evaluation of binding affinity and selectivity of the newly synthesized verapamil analogs

All synthesized compounds **2a–i**, **3a–i** and **5a–c** were evaluated as TAR RNA binders using the same fluorescence-based assay that was employed for the initial screening but performing this assay over a range of ligands concentrations. Compounds **1** (reduced verapamil) and **4** (verapamil conjugated to the succinamic acid linker) as well as all moieties that were employed for conjugates synthesis were used as controls. The measurement of fluorescence variation as a function of compounds concentration led to the calculation of dissociation constants ( $K_D$ ) as reported in Figure 2. First of all, while verapamil showed a  $K_D > 100 \mu\text{M}$  that could not be measured precisely, compound **1** bears a  $K_D$  of  $3.2 \mu\text{M}$  demonstrating a strong improvement of the affinity for the target once the nitrile group is reduced. Reductive amination products **2a–2i** were the most promising TAR binders with  $K_D$  in the low micromolar range from  $0.66 \mu\text{M}$  for **2e**,  $0.71$  and  $0.72 \mu\text{M}$  for **2i** and **2h**, respectively, until  $22.1 \mu\text{M}$  for **2c**. Amide formation products **3a–3i** were globally less strong binders except **3f**, **3g** and **3i** that showed a submicromolar  $K_D$  ( $0.41 \mu\text{M}$ ,  $0.45 \mu\text{M}$  and



**Figure 2.**  $K_D$  values (●) of compounds **2a–i**, **3a–i** and **5a–c** against HIV1 TAR RNA 27-mer fragment labeled at the 5'-end with Alexa<sup>488</sup>. Selectivity for TAR in the presence of 100 eq. of tRNA (■) and of DNA (▲) is reported.  $K_D$  values were determined from duplicates performed over three independent experiments using the change in fluorescence intensity. Error bars represent the standard deviation determined from three independent experiments.

$0.63 \mu\text{M}$ , respectively. Finally, amide formation products **5a–5c** showed weak affinity with  $12.9 \mu\text{M}$   $K_D$  for **5a** and no measurable binding for **5b** and **5c**. These results demonstrated that benzothiazole, indole and indoline as in **2e**, **2h**, **2i** as well as basic amino acids lysine, histidine and arginine as in **3f**, **3g** and **3i** are more favorable for binding once conjugated to verapamil with  $K_D$  values in the submicromolar range. 4-dimethylaminophenyl, naphthalene, quinoline and phenylthiazole moieties as in **2a**, **2b**, **2d**, **2f** and **3h** keep affinity for the target with  $K_D$  in the low micromolar range but globally loose affinity compared to the first group of ligands. Compounds **2g**, **3b**, **3c** and **3d** bearing thiazole, phenothiazine, phenyl and thiophene substituents, respectively, showed only low affinity with  $K_D$  over  $10 \mu\text{M}$ . Finally, compounds **2g**, **3a**, **3e**, **5b** and **5c** bearing thiazole (conjugated *via* amide formation as in **2g** and **5b**), 4-dimethylaminophenyl (conjugated *via* amide formation), 4-aminophenyl and benzimidazole showed no affinity for the target. It is important to note that all moieties tested without conjugation with verapamil showed very low affinity for the target and a  $K_D$  value could not be measured in the tested range of concentrations.

Affinity measurements alone are not representative of the potential of an RNA ligand, since a lack of selectivity could hamper the biological activity. In order to appreciate the selectivity of each compound we decided to measure  $K_D$  values in the presence of a large excess of competitors such as tRNA and DNA (100 eq. of each competitor). As illustrated in Figure 2, these experiments allowed us to discriminate the most selective compounds. While **1**, **2a–b**, **2e–f**, **2h–i**, **3f–i** and **5a** retained affinity in the presence of tRNA, only compounds **1**, **2h**, **3f** and **3i** retained TAR RNA affinity in the presence of DNA. This suggested that these latter were much more favorable for future studies keeping in mind that **2h**, **3f** and **3i** are also the strongest binders.

Based on these results, we analyzed the thermodynamic features of the interactions of the most promising compounds **2h**, **3f** and **3i** in comparison with parent compound **1** as well as compound **2a** both being good but non-selective ligands and the results are as summarized in Table 1.<sup>[24]</sup> Free energies of Gibbs ( $\Delta G^\circ$ ) were first calculated from the dissociation constants ( $\Delta G^\circ = -RT \ln K_D$ ) for all compounds and found to be very similar (from  $-30.0$  to  $-36.5$  kJ/mol). To complete the thermodynamic binding profiles, the enthalpic ( $\Delta H^\circ$ ) and the entropic ( $-T\Delta S^\circ$ ) energy contributions were determined after the determination of  $\Delta G^\circ_T$  at several temperatures (278–308 K). In all cases, the interaction is driven by the enthalpy suggesting the formation of specific interactions, such as hydrogen bonds, with the RNA target. In compounds **2h** and **3f**, the entropic contribution is close to zero while for compounds **2a** and **3i** its relevance is more important suggesting that electrostatic interactions and desolvation probably contribute to binding of these latter compounds. Furthermore, the total energy represented by the  $\Delta G^\circ$  value is composed of two components: (i) the first one reflects the non-electrostatic interactions (nonionic hydrophobic effects driven by entropy, specific interactions, including H-bonds, van der Waals interactions, and  $\pi$ -stacking) that contribute to the total free energy ( $\Delta G^\circ_{\text{nel}}$  value) and ii) the pure electrostatic contribution given by ionic interactions occurring between two groups of opposite charge ( $\Delta G^\circ_{\text{el}}$  value) that can be highly dependent on the salt concentration. As suggested by the enthalpy values,  $\Delta G^\circ_{\text{nel}}$  is over 90% for all compounds with **2h** showing 99% of non-electrostatic interactions formed. Altogether, binding studies demonstrated that various compounds are good RNA binders, but only few conjugates are selective for TAR and that compound **2h**, conjugate between verapamil and indole using reductive amination, is the best and most selective ligand.

### Inhibition of TAT/Tar interaction

Once identified the strongest and most selective binders, namely compound **2h** followed by **1**, **3f** and **3i**, we wondered if these ligands could be able to inhibit Tat/TAR interaction *in vitro* and we compared their efficacy with compound **2a**. To measure the ability of the synthesized compounds to displace a Tat fragment from a preformed TAR/Tat complex we employed a FRET assay, developed from a previously described procedure,<sup>[25]</sup> using a fluoresceinated Tat peptide fragment

(amino acids 48–57) and a Dabcyl-labeled TAR fragment (nucleotides 18–44). In the absence of ligand, association of these two components results in an efficient quenching of the dyes. Upon addition of increasing amounts of compounds **1**, **2a**, **2h**, **3f** and **3i**, the fluorescence intensity increases in almost all cases, thus confirming that these ligands can displace the Tat fragment from a preformed Tat/TAR complex.  $IC_{50}$  values for each compound are given in Table 1. These experiments showed that reduced verapamil (**1**) and compound **2a** were not efficient inhibitors of Tat/TAR interaction with  $IC_{50}$  of 1.87 mM for **1** and no inhibition for **2a** and that all other ligands were able to inhibit the formation of the complex with micromolar  $IC_{50}$ . Compounds **3f** and **3i** showed an  $IC_{50}$  of 302  $\mu\text{M}$  and 404  $\mu\text{M}$ , respectively, further confirming the potential of basic amino acids lysine and arginine for the preparation of RNA binders as previously reported. The best TAR ligand **2h** showed an  $IC_{50}$  of 18.8  $\mu\text{M}$  thus demonstrating to be not only the strongest and most selective ligand but also the most efficient inhibitor of Tat/TAR interaction. This results further underlines how inhibition studies are essential for the identification of the most promising binders since dissociation constants alone do not allow the discrimination between equally efficient ligands. All *in vitro* studies converged to the demonstration that reduced verapamil conjugated to an indole *via* reductive amination, as in compound **2h**, led to the best binding features and *in vitro* activity thus suggesting that the mode of binding and the site of interaction of **2h** is particularly favorable for inhibition and deserved to be studied more in detail.

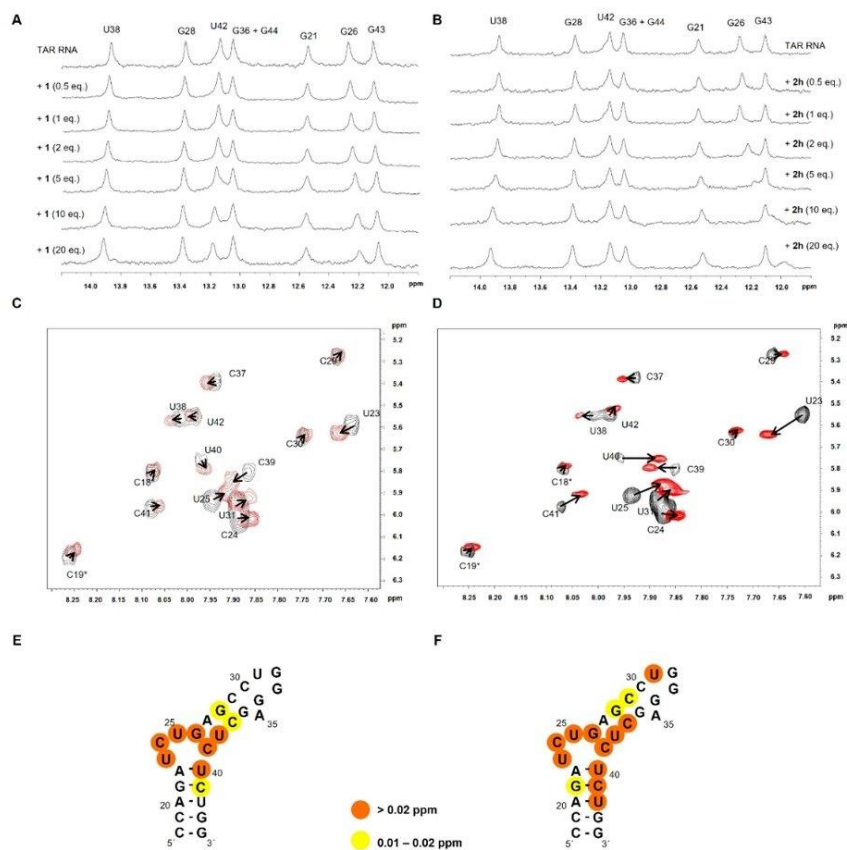
### Evaluation of the molecular mechanism of interaction

In order to study more in detail the molecular mechanism of interaction of compound **2h** with TAR, we decided to identify which residues were affected upon binding in comparison to non-modified parent compound **1**. Ligand-bound and free states of HIV-1 TAR RNA were examined using NMR spectroscopy. Therefore, evidence of how the ligand interacts with TAR can be provided by following the changes in chemical shifts of TAR as a ligand is titrated into solution. Upon titration of **1** into TAR, the most significant chemical shifts changes were observed for residues in the bulge area of TAR (Figure 3), confirming that verapamil specifically interacts with this bulge. The most affected residues are located between U23 and G26 as well as between U38 and U40. Titration of **2h** leads to similar

**Table 1.** Thermodynamic parameters for ligands/TAR RNA interaction and  $IC_{50}$  for the inhibition of Tat/TAR interaction.

ID	$\Delta G^\circ$ [kJmol <sup>-1</sup> ]	$\Delta H^\circ$ <sup>[a]</sup> [kJmol <sup>-1</sup> ]	$T\Delta S^\circ$ <sup>[a]</sup> [kJmol <sup>-1</sup> ]	$\Delta G^\circ_{\text{nel}}$ <sup>[b]</sup> [kJmol <sup>-1</sup> ]	$IC_{50}$ [ $\mu\text{M}$ ]
<b>1</b>	$-32.1 \pm 2.7$	$-50.5 \pm 4.7$	$-18.4 \pm 1.2$	$-28.8$ (90%)	$1870 \pm 12$
<b>2a</b>	$-32.6 \pm 1.8$	$-16.1 \pm 1.2$	$16.5 \pm 1.5$	not determined	no inhibition
<b>2h</b>	$-34.8 \pm 3.1$	$-33.3 \pm 1.6$	$1.5 \pm 1.0$	$-34.4$ (99%)	$18.8 \pm 1.1$
<b>3f</b>	$-36.5 \pm 3.1$	$-35.1 \pm 2.9$	$1.4 \pm 1.0$	$-32.8$ (90%)	$302 \pm 5$
<b>3i</b>	$-36.0 \pm 3.2$	$-24.2 \pm 2.1$	$11.8 \pm 1.2$	$-35.0$ (97%)	$404 \pm 7$

[a] Determined by temperature effect experiments by using the equation  $\Delta G^\circ_T = \Delta H^\circ_{\text{Tr}} + \Delta C_p(T - \text{Tr}) - T\Delta S^\circ_{\text{Tr}} - T\Delta C_p \ln(T/\text{Tr})$ . See the Supporting Information for definitions and further details. [b] Determined by salt effect experiments by using the equation  $\log(K_D) = \log(K_{\text{nel}}) - ZY \log[\text{KCl}]$ . See the Supporting Information for definitions and further details. The percentage of non-electrostatic interactions ( $\Delta G^\circ_{\text{nel}}/\Delta G^\circ$ ) is given in parentheses.

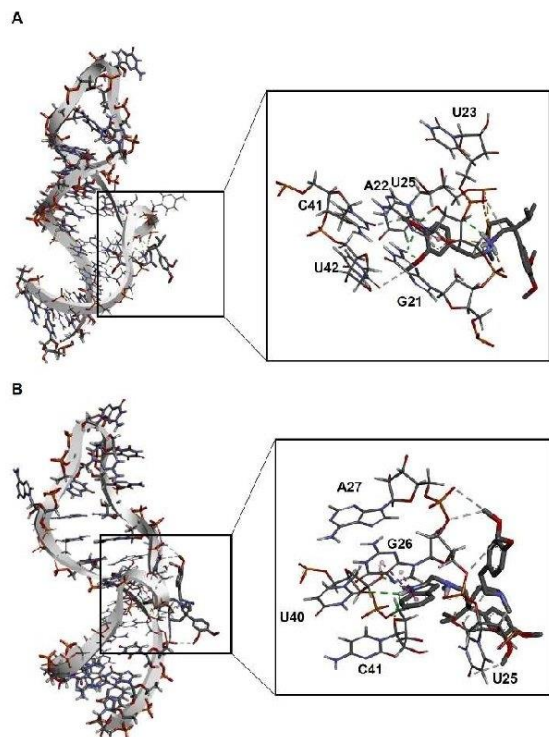


**Figure 3.** A) and B) Stacked plot of 1D NMR spectra of the imino region of 50  $\mu\text{M}$  TAR RNA with increasing concentration of **1** (A) and **2h** (B). The spectra were collected at 286 K in a  $\text{H}_2\text{O}/\text{D}_2\text{O}$  (90/10) buffer (20 mM phosphate and 50 mM NaCl, pH 7.4). C) and D) 2D-TOCSY spectra showing pyrimidine H5–H6 cross-peaks for TAR. Black: free TAR (100  $\mu\text{M}$ ); red: **1** (C) or **2h** (D)/TAR complex at a ratio 5:1. Arrows indicate chemical shift changes on ligand binding. The spectra were acquired at 308 K in a  $\text{D}_2\text{O}$  buffer (50 mM NaCl, 20 mM phosphate, pH 7.4); E) and F) 27-mer TAR RNA primary and secondary structure illustrating the nucleotides that are the most affected by the presence of **1** (E) or **2h** (F).

chemical-shift changes but two important differences were observed: (i) the interaction with TAR is extended to more residues since the strongest variations are observed from U23 to G26 and between C37 and U42 and (ii) G26 chemical shift is particularly modified by **2h** with 0.3 ppm variation *versus* 0.05 ppm for **1**. This demonstrates that the binding site of **1** and **2h** remains the same before and after conjugation but that further interactions are formed with the target by the additional moiety thus reinforcing the interaction and rendering it more efficient in terms of selectivity and inhibition activity. Interestingly, other compounds of the series that showed good affinity do not show the same level of inhibition. As an example, compound **2a**, containing a dimethylaminophenyl substituent conjugated upon reductive amination, bears a  $K_D$  of 1.3  $\mu\text{M}$  but no inhibition of Tat/TAR interaction. Indeed, the NMR study of its site of interaction (Figure S7) confirmed that this kind of chemical modification induces a different binding site compared to **1** and **2h**. In fact, G26 is not affected by **2a** while residues from U38 to U42 are only slightly shifted.

Confirmation of these results came from molecular docking that we performed using AutoDock 4. We employed a flexible ligand and TAR receptor to calculate independent docking

events utilizing a Lamarckian genetic algorithm and selected the 10 docking poses bearing the lowest energy in the presence of compounds **1** and **2h**. As illustrated in Figure 4, docking suggested that the interaction of both compounds occurs at the level of the bulge site. The obtained results suggested the formation of H-bonds and electrostatic interactions with phosphates and bases located in the region between G21 and U25 as well as between C41 and U42 in the case of compound **1** (Figure 4A). Molecular docking is also consistent with the formation of more specific hydrophobic interactions between the indole moiety and G26 as well as various H-bonds with U26–A27 and U40–C41 in the case of **2h** (Figure 4B). The comparison of these results with the molecular docking of compound **2a** (Figure S8) suggested that this latter interacts more specifically with the stem region with residues G17 and C19 as well as with residues U38, C39, U40 and U42 as suggested by NMR studies. **2a** thus seems to interact in a very different manner compared to **1** and **2h** further demonstrating that the site of interaction is extremely important for an efficient inhibition of RNA interaction with its biological partners. i.e. TAR interaction with Tat peptide. Indole moiety forms additional hydrophobic interactions with residue G26 as



**Figure 4.** Docking of **1** (A) and **2h** (B) with the 27-mer TAR hairpin loop performed by using autodock 4, in which the grid boxes were fixed on the entire RNA sequence.

well as hydrogen bonds with U40 and C41 thus increasing the strength and the number of interactions and improving the selectivity and the inhibition activity. Efficient and selective binding at a particular site of the target is thus essential to have good inhibition activity as demonstrated by the comparison of **2h** with **2a**.

## Conclusion

The screening of a library of commercially available compounds for their ability to bind HIV-1 TAR RNA led to the discovery of a small number of hits and in particular to the identification of verapamil, a phenylalkylamine calcium channel blocker currently in clinical use as an antiarrhythmic compound, as a promising scaffold for the preparation of new RNA binders. Verapamil has never been studied as an RNA ligand and we chose it for the synthesis of a new series of RNA ligands obtained upon conjugation of verapamil to other hits identified during screening or other known RNA-binding motifs. This led to the synthesis of 21 new compounds. Some of these compounds demonstrated to be very good TAR binders and ligands **2h**, **3f** and **3i** also showed selectivity in the presence of a large excess of both tRNA and DNA. **2h**, where indole fragment was coupled to verapamil, was the most efficient inhibitor of Tat/TAR interaction with an  $IC_{50}$  of 18.8  $\mu$ M and the detailed study of its molecular mechanism of interaction

demonstrated that this compound binds to the bulge site of TAR efficiently and selectively thus being a promising hit for future optimization.

Drug discovery in the field of RNA targeting is an interesting challenge for medicinal chemistry with great potential for future therapies. Library screening showed to be an efficient approach for the discovery of new RNA ligands bearing biological activity. Here, we thus described a screening that led to the discovery of verapamil, a drug currently employed in the treatment of some cardiovascular diseases, as an efficient TAR RNA binding agent. The chemical modification of this marketed drug showed that the biological activity can be greatly improved thus leading to strong and selective binders of HIV-1 TAR RNA. Even if the most active and selective compound **2h** needs further optimization before antiviral studies in infected cells, this work demonstrated that the study of existing drugs as scaffolds for the preparation of efficient RNA ligands could be an interesting strategy toward original ligands in the field of RNA targeting by small molecules.

## Experimental section

### Chemistry

**Materials:** Reagents and solvents were purchased from Merck and Carlo Erba Reagents and used without further purification. All reactions involving air- or moisture-sensitive reagents or intermediates were performed under an argon atmosphere. Flash column chromatography was carried out on silica gel columns (Interchim Puriflash silica HP 15  $\mu$ m) on a Puriflash XS420 system (Interchim). Analytical thin-layer chromatography (TLC) was conducted on Sigma Aldrich precoated silica gel and compounds were visualized by irradiation (254 nm) and/or by staining with ninhydrin and phosphomolybdic acid. HPLC was performed using a Water Alliance 2695 pump coupled with Waters 996 photodiode array detector and a Waters XSELECT™ CSH™ Fluoro-Phenyl (2.5  $\mu$ m, 4.6  $\times$  150 mm ColumnXP for analytical HPLC and 10  $\times$  150 mm for semipreparative HPLC). All HPLC analyses were run at room temperature using a gradient of CH<sub>3</sub>CN containing 0.1% TFA (eluent B) and water containing 0.1% TFA (eluent A) from 5% to 100% of B in 30 min at a flow rate of 1 mL/min for the analytical column and 3.5 mL/min for the semipreparative column. <sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded on a Bruker AC 200 MHz or a Bruker AC 400 MHz spectrometer. Chemical shifts are reported in parts per million (ppm,  $\delta$ ) referenced to the residual <sup>1</sup>H resonance of the solvent (CDCl<sub>3</sub>,  $\delta$  7.26; CD<sub>3</sub>OD  $\delta$  3.31; DMSO-*d*<sub>6</sub>  $\delta$  2.50). Splitting patterns are designated as follows: s (singlet), d (doublet), t (triplet), m (multiplet), br (broad). Coupling constants (J values) are listed in hertz (Hz). Low resolution mass spectra (MS) were obtained with a Bruker Daltonics Esquire 3000+ electrospray spectrometer equipped with API ionization source. High resolution mass spectra (HRMS) were obtained with a LTQ Orbitrap hybrid mass spectrometer with an electrospray ionization probe (ThermoScientific, San Jose, CA) by direct infusion from a pump syringe to confirm correct molar mass and high purity of compounds.

**General procedure for reductive amination (General procedure A):** To a solution of 1.3 eq. of the appropriate aldehyde (0.15 mmol) in CH<sub>2</sub>Cl<sub>2</sub> or CH<sub>2</sub>Cl<sub>2</sub>/MeOH (1:1) (4 mL) is added 1 eq. of compound **1** (50 mg, 0.11 mmol). When necessary the pH is adjusted to pH = 5–6 by addition of acetic acid. Then NaBH<sub>3</sub>CN (14 mg, 0.22 mmol, 2 eq.) is added and the reaction mixture is stirred at room

temperature overnight. The solvent is then removed under reduced pressure.  $\text{CH}_2\text{Cl}_2$  is added to the crude product and the organic phase is washed with a solution of  $\text{NaHCO}_3$  (10%), water and brine. After drying over  $\text{MgSO}_4$  and filtration, the solvent is evaporated under reduced pressure and the residue is purified by flash chromatography on a silica gel column using a gradient of  $\text{CH}_2\text{Cl}_2/\text{MeOH}$  9:1 as the eluent.

**(±)-5-[(3,4-Dimethoxyphenethyl)methylamino]-2-(3,4-dimethoxyphenyl)-2-isopropylpentyl-N1-(tert-butyl-1H-indolecarboxylate-3-methyl)amine (2' h)**

Compound 2' h was prepared following general procedure A in the presence of *tert*-butyl 3-formyl-1H-indole-1-carboxylate (36 mg). Compound 2' h was obtained as a yellow solid: 44 mg (59%);  $R_f = 0.25$  ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  95:5);  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 8.09$  (d,  $J = 8.2$  Hz, 1H), 7.59 (d,  $J = 7.7$  Hz, 1H), 7.56 (s, 1H), 7.31–7.15 (m, 2H), 6.86–6.76 (m, 4H), 6.74 (d,  $J = 1.9$  Hz, 1H), 6.66 (dd,  $J = 8.1$ , 1.9 Hz, 1H), 3.94 (d,  $J = 13.7$  Hz, 1H), 3.89 (d,  $J = 14.1$  Hz, 1H), 3.80–3.74 (m, 9H), 3.68 (s, 3H), 3.07–2.97 (m, 2H), 2.71–2.58 (m, 4H), 2.48 (t,  $J = 7.6$  Hz, 2H), 2.26 (s, 3H), 1.99–1.91 (m, 1H), 1.86–1.77 (m, 2H), 1.65 (s, 9H), 1.46–1.26 (m, 2H), 0.76, 0.71 (2d,  $J = 6.8$  Hz, 6H);  $^{13}\text{C}$  NMR (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 151.0, 150.5, 149.8, 149.0, 148.8, 137.1, 136.7, 133.6, 131.6, 125.5, 125.1, 123.7, 122.0, 121.7, 120.6, 120.4, 116.2, 113.7, 113.5, 113.2, 112.3, 84.9, 59.9, 58.8, 56.6, 56.5, 56.4, 56.4, 53.7, 47.8, 45.6, 42.3, 35.3, 33.0, 32.2, 28.4, 22.1, 18.5, 18.4$ ; MS (ESI):  $m/z$  688.27 (M+H)<sup>+</sup> (theoretical 688.42).

**(±)-5-[(3,4-Dimethoxyphenethyl)methylamino]-2-(3,4-dimethoxyphenyl)-2-isopropylpentyl-N1-(1H-indole-3-methyl)amine (2 h)**

Compound 2 h was dissolved in  $\text{CH}_2\text{Cl}_2$  (3 mL) and TFA (195  $\mu\text{L}$ ; 2.55 mmol, 40 eq.) was added. The reaction mixture was stirred at 35 °C for 3 hours. The solvent was removed under reduce pressure. The crude product was then triturated with  $\text{Et}_2\text{O}$  and after filtration, 5 b was obtained as a brown solid: 35 mg (94%);  $R_f = 0.3$  ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  93:7);  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 7.79$  (d,  $J = 7.8$  Hz, 1H, h), 7.55 (s, 1H), 7.47 (d,  $J = 8.0$  Hz, 1H), 7.25–7.13 (m, 2H), 6.92–6.85 (m, 2H), 6.82–6.72 (m, 2H), 6.55–6.46 (m, 2H), 4.56 (d,  $J = 14.0$  Hz, 1H), 4.49 (d,  $J = 13.9$  Hz, 1H), 3.84–3.73 (m, 9H), 3.63–3.51 (m, 5H), 3.32–3.05 (m, 4H), 3.04–2.87 (m, 2H), 2.79 (s, 3H), 2.08–1.73 (m, 3H), 1.61–1.43 (m, 2H), 0.71, 0.55 (2d,  $J = 6.7$  Hz, 3H);  $^{13}\text{C}$  NMR (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 150.8, 150.2, 149.8, 149.8, 138.0, 131.6, 130.1, 129.4, 128.8, 123.6, 122.3, 121.6, 121.5, 119.0, 113.7, 113.4, 113.2, 112.8, 112.6, 104.89, 58.4, 57.2, 56.5, 56.5, 56.3, 50.0, 46.9, 44.0, 40.5, 34.3, 31.0, 30.3, 19.8, 17.9, 17.7$ ; HRMS (ESI):  $m/z$  588.37988 (M+H)<sup>+</sup> ( $\text{C}_{36}\text{H}_{50}\text{N}_3\text{O}_4$  requires 588.37958).

## NMR experiments

### Materials and equipment

NMR experiments were conducted as described in a previous work and detailed below.<sup>261</sup> High resolution NMR experiments were recorder on a BRUKER AVANCE Ultra shield DRX 500 spectrometer operating at 500.13 MHz for  $^1\text{H}$ , equipped with a temperature control unit (BCU 6.0, BVT 3000), and an inverse probe head (5 mm PHTXI 1H-13 C/15 N Z-GRD). Proton chemical shift was referenced internally by setting the carrier frequency on water at the center of the spectrum (4.71 ppm at 13 °C and 4.70 ppm at 35 °C). Chemical shifts ( $\delta$ ) are expressed in parts per million (ppm). All NMR experiments were carried out using standard pulse sequences

supplied by the spectrometer manufacturer (BRUKER). 1D and 2D spectra were processed using TOPSPIN 2.1 NMR Software (BRUKER).

For the preparation of all NMR samples, a folding of TAR RNA (100  $\mu\text{M}$ ) was performed in the appropriate buffer (vide infra) as described in material and methods section of the paper. After refolding, the NMR sample (alone or with the appropriate amount of ligand) is incorporated into a Shigemi NMR tube.

$^1\text{H}$  NMR imino proton spectra were recorded in a  $\text{H}_2\text{O}/\text{D}_2\text{O}$  (90/10) buffer containing 20 mM phosphate and 50 mM NaCl at 286 °K (13 °C) by using a WATERGATE 3-9-19 water suppression. Each proton NMR spectrum was acquired using 10.964 KHz Spectral Width (SW), 64 K complex data point, acquisition time (aq) of 2.98 s, relaxation delay (D1) of 1 s, number of scan (ns) between 1000 and 2000, number of dummy scan (ds) 4 and a 90° flip angle pulse width. Water suppression was achieved using WATERGATE pulse sequence. Gradient pulse were sine shape (SINE.100), 1.5 ms long (P16) with 100  $\mu\text{s}$  gradient recovery delay (D16) and strengths set to 8.44 Gauss.cm<sup>-1</sup> (20%). A 45.6  $\mu\text{s}$  delay (D19) was used for binomial water suppression. Prior to Fourier transformation, the fids were multiplied by an exponential line broadening function of 3 Hz.

gs-TOCSY Phase sensitive (States – TPPI mode) experiments using MLEV 17 pulse sequence for spin lock were recorded in a D20 buffer (50 mM NaCl, 20 mM phosphate, pH 7.4) at 308 K (35 °C) by using a WATERGATE 3-9-19 water suppression. Each TOCSY 2D NMR spectrum was acquired with a spectral width of 5 KHz in both dimension, 2 K complex data point in F2, 256 t1 increments (between 32 and 64 scans by increment) in F1, 0.20 s for aq and D1 of 2 s. MLEV 17 pulse sequence for spin lock was set to 60 ms. Water suppression was achieved using WATERGATE pulse sequence. Gradient pulse were sine shape (SINE.100), 1.5 ms long (P16) with 100  $\mu\text{s}$  gradient recovery delay (D16) and strengths set to 8.44 Gauss.cm<sup>-1</sup> (20%). A 100  $\mu\text{s}$  delay (D19) was used for binomial water suppression. Prior to Fourier transformation a QSINE window function (SSB=2) was applied in both dimension and the data were zero filled and linear predicted (NC=32) to 1 K data points in F1.

**Molecular Modeling and Docking:** The experimental structure of RNA TAR complexed with ligand rbt158 (pdb code 1UUI) was used as input for our calculations. As described in our previous work,<sup>271</sup> for docking with AutoDock, polar hydrogen atoms, Kollman united charges and solvent parameters were applied to the RNA using pmol2q script ([http://www.sourcefiles.org/Scientific/Biology/Proteins/pmol2q\\_2.3.0.tar.gz](http://www.sourcefiles.org/Scientific/Biology/Proteins/pmol2q_2.3.0.tar.gz)). This script converts the .pdb file format to of the RNA template to the .pdbqt file format that is compatible with AutoDock program version 4 (<http://autodock.scripps.edu/>). The compounds .pdbqt files compatible with Autodock program version 4 were prepared from the .pdb files obtained from <http://ligand-expo.rcsb.org/> (NMY).

RNA-ligand molecular dockings were conducted using AutoDock program version 4. The rotational bonds of the ligand were treated as flexible, whereas the receptor was kept rigid. Grid box was fixed in order to include the entire RNA sequence. RNA-ligand interactions were analyzed and visualized using Discovery Studio Visualizer version 4.1 (<http://accelrys.com/products/discovery-studio/>).

## Biochemistry

Unless otherwise stated, all reagents and solvents were of analytical grade and from Sigma (St Louis, U.S.A.). HEPES and all inorganic salts for buffers were purchased from Merck (molecular biology grade). RNA and DNA oligonucleotides were purchased from IBA GmbH and used without further purification. Oligonucleotide

sequences: TAR (5'-Alexa488-CCAGGUCUGAGCCUGGGAGCUCC-CUGG-3'), DNA (5'-CGTTTTTATTTTGC-3' and 5'-GCAAAAA-TAAAAACG-3'). A mixture of pre- and mature yeast tRNAs (containing over 30 different species) was purchased from Sigma (type X-SA). Stocks of tRNAmix can be quantified in its native form (without base hydrolysis) using an extinction coefficient of 9640 cm<sup>-1</sup> M<sup>-1</sup> per base.

**Buffers:** All buffers were filtered through 0.22 μm Millipore filters (GP ExpressPLUS membrane). A small aliquot (100 mL) was first filtered and then discarded to avoid any contaminants that might be leached from the filter. The solutions to be used in the fluorescence experiments were prepared by diluting the concentrated stocks in Milli-Q water and filtered again as described above. All standard fluorescence measurements were performed in buffer A (20 mM HEPES (pH 7.4 at 25 °C), 20 mM NaCl, 140 mM KCl and 3 mM MgCl<sub>2</sub>). FRET experiments were performed in buffer B (50 mM tris buffer (pH 7.4 at 25 °C), 20 mM KCl and 0.005% Tween 20). For competitive experiments in the presence of a tRNA, a mixture of pre- and mature yeast tRNAs (containing over 30 different species from baker's yeast (*S. cerevisiae*, Sigma, type X-SA)) was added to buffer A to obtain a 100-fold nucleotide excess regarding TAR RNA.

**Inhibition of Tat/TAR interaction:** As previously described,<sup>[28]</sup> Ligand solutions and RNA (40 nM working solution) were prepared as described above in buffer B (50 mM tris buffer (pH 7.4 at 25 °C), 20 mM KCl and 0.005% Tween 20). Labeled Tat peptide was prepared at 40 nM in buffer B and mixed to an equal volume of TAR RNA for 20 min at room temperature to form the Tat/TAR complex before adding the ligand. The appropriate ligand solution (30 μL) was then added to a well of a non-treated black 384-well plate, in triplicate, and 30 μL of the Tat/TAR solution was added. This subsequent dilution lowered the final Tat/TAR concentration to 10 nM. Fluorescence was measured as described above after 30 min of incubation at room temperature.

**Binding studies:** Ligand solutions were prepared as serial dilutions in buffer A at a concentration four times higher than the desired final concentration to allow for the subsequent dilution during the addition of the RNA solution. Binding experiments were performed in 384-well plates (Greiner bio-one) in a final volume of 60 μL using a 5070 EpMotion automated pipetting system (Eppendorf). Each experiment was performed in duplicate. 10 nM TAR beacon was used in each well. Each ligand was added in 15 dilutions (from 0.030 nM to 0.5 μM) and the fluorescence increase measured after 4 hours. The fluorescence was measured on a GeniosPro (Tecan) with an excitation filter of 485 ± 10 nm and an emission filter of 535 ± 15 nm. Each point was measured 10 times with a 500 μs integration time and averaged. Binding data were analyzed using Graphpad Prism 5 software. Unless otherwise stated, binding profiles were well modeled using a simple model assuming the one to one stoichiometry. Refolding of the RNA was performed using a thermocycler (ThermoStat Plus Eppendorf) as follow: the RNA, diluted in 1 mL of buffer A at a concentration of 200 nM, was first denatured by heating to 90 °C for 2 min, then cooled to 4 °C for 10 min followed by incubation at 20 °C for 15 min. After refolding, the RNA was diluted to a working concentration of 10 nM through addition of the appropriate amount of buffer A. The tube was mixed and 50 μL of the RNA solution was added to each well containing ligand. This subsequent dilution lowered the final RNA concentration to 5 nM. To study the temperature dependence, the plates were incubated after overnight equilibrium at different temperature ranging from 5 °C to 35 °C.

**Data analysis:** Binding data (K<sub>D</sub> and FRET experiments) were analyzed using Prism 5 (GraphPad Software) by nonlinear regression following Equation (1):

$$Y = \text{Bottom} + (\text{Top} - \text{Bottom}) / (1 + 10^{[(\log K_{50} - X) \times \text{HillSlope}]}) \quad (1)$$

K<sub>D</sub> values were converted to ΔG° values as ΔG° = -RT ln K<sub>D</sub>. Binding data were analyzed using the nonlinear least squares numerical solver-based binding data global analysis program GraphPad, in which the calculated binding surface is obtained using a numerical constrained optimization chemical equilibrium solver. Unless otherwise stated, binding profiles were well modeled using a simple model assuming the one to one stoichiometry. A higher initial fluorescence value is observed in the presence of tRNA, which is consistent with the modification of the polarity of the solvent and a small fluorescence of the tRNA mixture.

For thermodynamic analysis, ΔG° values were plotted versus T for the three-parameter fit. Nonlinear regression in Prism 4 (GraphPad Software) was used to fit Equation (2) to the data:

$$\Delta G^{\circ}_T = \Delta H^{\circ}_{Tr} + \Delta C_p (T - Tr) - T \Delta S^{\circ}_{Tr} - T \Delta C_p \ln (T/Tr) \quad (2)$$

where Tr is a constant reference temperature (in our study Tr = 293.15 K), and the three fit parameters are ΔH°<sub>Tr</sub>, the change in enthalpy upon binding at Tr; ΔS°<sub>Tr</sub>, the change in entropy upon binding at Tr; and ΔC<sub>p</sub>, the change in heat capacity. Starting values for the three parameters did not affect the final values. ΔC<sub>p</sub> was assumed to be independent of temperature; inclusion of a ΔC<sub>p</sub>/ΔT term in the analysis did not improve the quality of the fits and gave larger standard errors for the returned parameters.

ΔH°<sub>T</sub> and ΔS°<sub>T</sub> were calculated by using Equation (3):

$$\Delta H^{\circ}_T = \Delta H^{\circ}_{Tr} + \Delta C_p (T - Tr) \quad \text{and} \quad \Delta S^{\circ}_T = \Delta S^{\circ}_{Tr} + \Delta C_p \ln (T/Tr) \quad (3)$$

Salt dependence of K<sub>D</sub> was analyzed by the Equation (4):

$$\log(K_D) = \log(K_{\text{nel}}) - Z\psi \log [KCl] \quad (4)$$

where K<sub>nel</sub> is the dissociation constant at the standard state in 1 M KCl, Z is the number of ions displaced from the nucleic acid (essentially the number of intermolecular ion pairs) and ψ is the fractional probability of a counterion being thermodynamically associated with each phosphate of the RNA number of cations. K<sub>nel</sub> and Zψ were treated as fitting parameters.

## Acknowledgments

We thank J. M. Guignon (Plateforme Bernard Rossi, CEA TIRO, Nice, France) for HRMS analyses. A PhD fellowship (to C.M.) was supported by grant from the Agence Nationale de la Recherche (ANR-17-CE18-0009).

**Keywords:** inhibitors · molecular docking · RNA · structure-activity relationships · therapeutics

- [1] a) A. Di Giorgio, M. Duca, *MedChemComm* **2019**, *10*, 1242–1255; b) M. D. Disney, *J. Am. Chem. Soc.* **2019**, *141*, 6776–6790; c) A. Donlic, A. E. Hargrove, *Wiley Interdiscip. Rev.: RNA* **2018**, *9*, e1477; d) K. D. Warner, C. E. Hajdin, K. M. Weeks, *Nat. Rev. Drug Discovery* **2018**, *17*, 547–558.
- [2] M. Matsui, D. R. Corey, *Nat. Rev. Drug Discovery* **2017**, *16*, 167–179.
- [3] E. Valeur, P. Jimonet, *J. Med. Chem.* **2018**, *61*, 9004–9029.
- [4] D. N. Wilson, *Nat. Rev. Microbiol.* **2014**, *12*, 35–48.

- [5] T. Hermann, *Wiley Interdiscip. Rev.: RNA* **2016**, *7*, 726–743.
- [6] A. Blond, E. Ennifar, C. Tisne, L. Micouin, *ChemMedChem* **2014**, *9*, 1982–1996.
- [7] M. Stevens, E. De Clercq, J. Balzarini, *Med. Res. Rev.* **2006**, *26*, 595–625.
- [8] K. S. Long, D. M. Crothers, *Biochemistry* **1995**, *34*, 8885–8895.
- [9] a) S. Kumar, P. Kellish, W. E. Robinson, Jr., D. Wang, D. H. Appella, D. P. Arya, *Biochemistry* **2012**, *51*, 2331–2347; b) S. Kumar, N. Ranjan, P. Kellish, C. Gong, D. Watkins, D. P. Arya, *Org. Biomol. Chem.* **2016**, *14*, 2052–2056; c) N. N. Patwardhan, L. R. Ganser, G. J. Kapral, C. S. Eubanks, J. Lee, B. Sathyamoorthy, H. M. Al-Hashimi, A. E. Hargrove, *MedChemComm* **2017**, *8*, 1022–1036; d) J. Sztuba-Solinska, S. R. Shenoy, P. Gareiss, L. R. Krumpke, S. F. Le Grice, B. R. O’Keefe, J. S. Schneekloth, Jr., *J. Am. Chem. Soc.* **2014**, *136*, 8402–8410; e) M. Zeiger, S. Stark, E. Kalden, B. Ackemann, J. Ferner, U. Scheffer, F. Shoja-Bazargani, V. Erdel, H. Schwalbe, M. W. Gobel, *Bioorg. Med. Chem. Lett.* **2014**, *24*, 5576–5580.
- [10] a) M. D. Disney, A. M. Winkelsas, S. P. Velagapudi, M. Southern, M. Fallahi, J. L. Childs-Disney, *ACS Chem. Biol.* **2016**, *11*, 1720–1728; b) C. S. Eubanks, B. Zhao, N. N. Patwardhan, R. D. Thompson, Q. Zhang, A. E. Hargrove, *J. Am. Chem. Soc.* **2019**, *141*, 5692–5698; c) B. S. Morgan, J. E. Forte, R. N. Culver, Y. Zhang, A. E. Hargrove, *Angew. Chem. Int. Ed.* **2017**, *56*, 13498–13502; d) B. S. Morgan, J. E. Forte, A. E. Hargrove, *Nucleic Acids Res.* **2018**, *46*, 8025–8037.
- [11] a) J. Palacino, S. E. Swalley, C. Song, A. K. Cheung, L. Shu, X. Zhang, M. Van Hoosier, Y. Shin, D. N. Chin, C. G. Keller, M. Beibel, N. A. Renaud, T. M. Smith, M. Salcius, X. Shi, M. Hild, R. Servais, M. Jain, L. Deng, C. Bullock, M. McLellan, S. Schuierer, L. Murphy, M. J. Blommers, C. Blaustein, F. Berenshteyn, A. Lacoste, J. R. Thomas, G. Roma, G. A. Michaud, B. S. Tseng, J. A. Porter, V. E. Myer, J. A. Tallarico, L. G. Hamann, D. Curtis, M. C. Fishman, W. F. Dietrich, N. A. Dales, R. Sivasankaran, *Nat. Chem. Biol.* **2015**, *11*, 511–517; b) H. Ratni, M. Ebeling, J. Baird, S. Bendels, J. Bylund, K. S. Chen, N. Denk, Z. Feng, L. Green, M. Guerard, P. Jablonski, B. Jacobsen, O. Khwaja, H. Kletzl, C. P. Ko, S. Kustermann, A. Marquet, F. Metzger, B. Mueller, N. A. Naryshkin, S. V. Paushkin, E. Pinard, A. Poirier, M. Reutlinger, M. Weetall, A. Zeller, X. Zhao, L. Mueller, *J. Med. Chem.* **2018**, *61*, 6501–6517; c) M. Sivaramakrishnan, K. D. McCarthy, S. Campagne, S. Huber, S. Meier, A. Augustin, T. Heckel, H. Meistermann, M. N. Hug, P. Birrer, A. Moursy, S. Khawaja, R. Schmucki, N. Berntenis, N. Giroud, S. Golling, M. Tzouros, B. Banfai, G. Duran-Pacheco, J. Lamerz, Y. Hsiu Liu, T. Luebbbers, H. Ratni, M. Ebeling, A. Clery, S. Paushkin, A. R. Krainer, F. H. Allain, F. Metzger, *Nat. Commun.* **2017**, *8*, 1476.
- [12] a) C. L. Haga, S. P. Velagapudi, J. R. Strivelli, W. Y. Yang, M. D. Disney, D. G. Phinney, *ACS Chem. Biol.* **2015**, *10*, 2267–2276; b) S. P. Velagapudi, M. D. Cameron, C. L. Haga, L. H. Rosenberg, M. Lafitte, D. R. Duckett, D. G. Phinney, M. D. Disney, *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 5898–5903; c) S. P. Velagapudi, S. M. Gallo, M. D. Disney, *Nat. Chem. Biol.* **2014**, *10*, 291–297.
- [13] M. D. Disney, A. J. Angelbello, *Acc. Chem. Res.* **2016**, *49*, 2698–2704.
- [14] a) D. D. Vo, C. Becquart, T. P. A. Tran, A. Di Giorgio, F. Darfeuille, C. Staedel, M. Duca, *Org. Biomol. Chem.* **2018**, *16*, 6262–6274; b) D. D. Vo, M. Duca, *Methods Mol. Biol.* **2017**, *1517*, 137–154; c) D. D. Vo, C. Staedel, L. Zehnacker, R. Benhida, F. Darfeuille, M. Duca, *ACS Chem. Biol.* **2014**, *9*, 711–721; d) D. D. Vo, T. P. Tran, C. Staedel, R. Benhida, F. Darfeuille, A. Di Giorgio, M. Duca, *Chem. Eur. J.* **2016**, *22*, 5350–5362.
- [15] J. P. Joly, G. Mata, P. Eldin, L. Briant, F. Fontaine-Vive, M. Duca, R. Benhida, *Chem. Eur. J.* **2014**, *20*, 2071–2079.
- [16] J. P. Joly, M. Gaysinski, L. Zara, M. Duca, R. Benhida, *Chem. Commun.* **2019**, *55*, 10432–10435.
- [17] a) C. Becquart, M. Le Roch, S. Azoulay, P. Uriac, A. Di Giorgio, M. Duca, *ACS Omega* **2018**, *3*, 16500–16508; b) C. Staedel, T. P. A. Tran, J. Giraud, F. Darfeuille, A. Di Giorgio, N. J. Tourasse, F. Salin, P. Uriac, M. Duca, *Sci. Rep.* **2018**, *8*, 1667.
- [18] a) N. W. Luedtke, Q. Liu, Y. Tor, *Biochemistry* **2003**, *42*, 11391–11403; b) A. C. Stelzer, A. T. Frank, J. D. Kraz, M. D. Swanson, M. J. Gonzalez-Hernandez, J. Lee, I. Andricioaei, D. M. Markovitz, H. M. Al-Hashimi, *Nat. Chem. Biol.* **2011**, *7*, 553–559; c) J. R. Thomas, P. J. Hergenrother, *Chem. Rev.* **2008**, *108*, 1171–1224.
- [19] D. A. Erlanson, B. J. Davis, W. Jahnke, *Cell Chem. Biol.* **2019**, *26*, 9–15.
- [20] J. R. Thomas, X. Liu, P. J. Hergenrother, *J. Am. Chem. Soc.* **2005**, *127*, 12434–12435.
- [21] M. Mayer, P. T. Lang, S. Gerber, P. B. Madrid, I. G. Pinto, R. K. Guy, T. L. James, *Chem. Biol.* **2006**, *13*, 993–1000.
- [22] K. F. Blount, Y. Tor, *ChemBioChem* **2006**, *7*, 1612–1621.
- [23] a) B. Bogner, S. Ahmed, M. L. Kuppusamy, K. Selvendiran, M. Khan, J. Jeko, O. H. Hankovszky, T. Kalai, P. Kuppusamy, K. Hideg, *Bioorg. Med. Chem.* **2010**, *18*, 2954–2963; b) L. J. Theodore, W. L. Nelson, R. H. Zobrist, K. M. Giacomini, J. C. Giacomini, *J. Med. Chem.* **1986**, *29*, 1789–1792.
- [24] a) G. Klebe, *Nat. Rev. Drug Discovery* **2015**, *14*, 95–110; b) A. Umuhire Juru, N. N. Patwardhan, A. E. Hargrove, *ACS Chem. Biol.* **2019**, *14*, 824–838.
- [25] A. I. Murchie, B. Davis, C. Isel, M. Afshar, M. J. Drysdale, J. Bower, A. J. Potter, I. D. Starkey, T. M. Swarbrick, S. Mirza, C. D. Prescott, P. Vaglio, F. Aboul-ela, J. Karn, *J. Mol. Biol.* **2004**, *336*, 625–638.
- [26] L. Pascale, A. L. Gonzalez, A. Di Giorgio, M. Gaysinski, J. Teixido Closa, R. E. Tejedor, S. Azoulay, N. Patino, *J. Biomol. Struct. Dyn.* **2016**, *34*, 2327–2338.
- [27] T. P. Tran, D. D. Vo, A. Di Giorgio, M. Duca, *Bioorg. Med. Chem.* **2015**, *23*, 5334–5344.
- [28] L. Pascale, S. Azoulay, A. Di Giorgio, L. Zenacker, M. Gaysinski, P. Clayette, N. Patino, *Nucleic Acids Res.* **2013**, *41*, 5851–5863.

---

Manuscript received: October 29, 2019  
 Revised manuscript received: December 9, 2019  
 Accepted manuscript online: December 10, 2019

## III.2 Conclusion

La découverte de médicaments dans le domaine du ciblage d'ARN est un défi grandissant pour la chimie médicinale avec un grand potentiel pour les thérapies futures. Afin d'identifier les interactions possibles entre ligand et l'ARN, le criblage de bibliothèque s'avère une approche efficace et essentielle pour la découverte de nouveau chémotype capable de former de nouveaux réseaux de liaison. L'ARN TAR du VIH-1 représente une cible pertinente pour pouvoir découvrir et développer de nouveaux chémotypes de ligands d'ARN. De plus, face à l'émergence possible de variants VIH résistants aux médicaments actuels, il est indispensable de continuer l'étude de ce modèle de façon à développer de nouvelles molécules antivirales.

Dans ce contexte, un criblage nous a permis d'identifier une nouvelle structure qu'est le vérapamil, qui n'a jamais été considéré comme ligand d'ARN. Afin d'optimiser ce composé, une série de 21 conjugués de vérapamil a été préparée en couplant sur le vérapamil d'autres motifs interagissant avec TAR. L'étude de l'affinité de ces conjugués a démontré que l'approche « multimodale » permettait d'améliorer nettement l'affinité d'un facteur 10 pour atteindre des constantes submicromolaires ( $K_D$ ). Après l'étude de la sélectivité réalisée en présence d'excès d'ARNt et d'ADN, trois ligands prometteurs ont gardé leur affinité dont le composé **2h** (figure 95) qui présente un motif indolique.

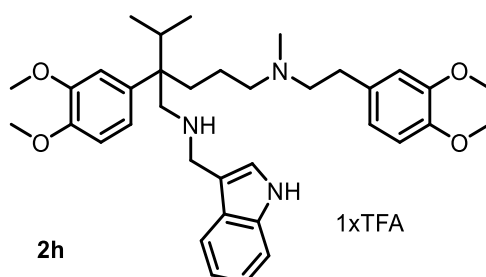


Figure 95 : Structure du composé vérapamil **2h**.

Parmi ces trois conjugués, le composé **2h** s'est montré le plus efficace pour inhiber l'interaction Tat/TAR avec un  $IC_{50}$  de 18,8  $\mu$ M. L'étude détaillée de son réseau d'interactions avec TAR a mis en évidence que le groupement indole permettait d'établir des liaisons hydrogènes et des interactions hydrophobes spécifiques avec le renflement de TAR permettant au composé de se lier efficacement et sélectivement. Néanmoins, ce composé doit encore être optimisé avant d'être étudié pour son activité antivirale dans les cellules infectées.

Ainsi, par cette étude, il a été montré que l'approche par criblage reste encore très prometteuse dans la découverte de nouveaux ligands d'ARN. Des molécules déjà connues pour leur activité biologique sont une source d'inspiration considérable et peuvent être utilisées en tant que ligands d'ARN par suite d'une optimisation par la chimie médicinale.





---

# Conclusion générale

---



L'ARN a longtemps été considéré comme un simple intermédiaire dans la transmission et l'expression de l'information génétique au cours desquelles l'ADN est transcrit en ARN qui est ensuite traduit en protéines. Depuis plusieurs années, le rôle de l'ARN en tant qu'outil essentiel et cible thérapeutique a été largement étudié. L'année 2020 a marqué un tournant pour le rôle de l'ARN comme cible thérapeutique par (i) la mise sur le marché des vaccins à ARN pour lutter contre le SARS-CoV2, (ii) le prix Nobel de Chimie 2020 attribué à Emmanuelle Charpentier et Jennifer Doudna pour le système CRISPR-Cas9 où l'ARN joue un rôle essentiel pour la reconnaissance de l'ADN ciblé par la protéine Cas9 et (iii) la mise sur le marché d'un nouveau médicament, le Risdiplam, qui modifie l'épissage alternatif de l'ARNm codant pour la protéine SMN2, pour la maladie de l'atrophie spinale musculaire. Ces exemples démontrent l'implication importante de l'ARN dans de nombreux processus biologiques chez les eucaryotes mais qui est tout aussi importante dans les cellules procaryotes notamment chez les bactéries.

Dans le cadre de ce travail, nous avons choisi d'étudier un processus d'interaction ARN-ARN présent chez la bactérie dans les systèmes Toxine-Antitoxine de type I dont le système *AapA/IsoA* chez *Helicobacter pylori*. Ce système TA est constitué d'un ARNm codant pour une toxine et d'un ARN non codant permettant d'inhiber la traduction de l'ARNm par des interactions boucle-boucle. Pour activer artificiellement la production de la toxine et induire la mort cellulaire bactérienne, nous avons conçu, synthétisé et étudié une librairie de composés capable d'interférer avec la reconnaissance boucle-boucle de ce système TA. Ce système a donc été choisi dans le contexte du développement et de la validation de nouvelles cibles antimicrobiennes mais aussi pour la découverte de nouvelles molécules antibiotiques.

Cette collection a été conçue par une approche de ligands multifonctionnels avec des fragments connus pour interagir avec l'ARN. Une collection de 67 composés a ainsi été synthétisée en utilisant la réaction multicomposants de Ugi à travers deux méthodologies différentes : soit par l'utilisation d'un support solide permettant d'obtenir des composés cycliques de type 2,5-dicétopipérazine ainsi que des linéaires, soit en solution par l'utilisation d'un isocyanure commercial (*tert*-butyle isocyanure).

Pour évaluer la capacité des composés synthétisés à inhiber l'interaction boucle-boucle, différents tests *in vitro* ont été développés. La cinétique de passage de la structure boucle-boucle du système TA *AapA1/IsoA1* en un duplex étendu stable a d'abord été étudiée par différentes constructions. A partir de ces résultats, nous avons mis en place un test d'inhibition utilisant des mesures d'anisotropie. Les premiers tests d'inhibition réalisés ont montré des résultats prometteurs pour le composé **C38b** qui fait partie des 20 molécules les plus affines pour l'ARN *AapA1*. A travers ce test, nous avons pu mettre en avant la capacité de certains ligands d'ARN d'inhiber la reconnaissance ARN-ARN, cette dernière étant très peu étudiée actuellement dans la littérature. Même si pour l'instant, les molécules ne sont

pas capables d'interférer dans un modèle plus complexe réalisé directement chez une bactérie, cette approche originale vers la découverte de nouveaux antibiotiques reste ambitieuse et stimulante et l'optimisation du composé **C38b** est d'ores et déjà envisagée. Cette approche, si validée, pourrait être ensuite appliquée à d'autres pathogènes majeurs tels que *Staphylococcus aureus* et *Mycobacterium tuberculosis*.

Enfin, l'activité d'inhibition de la croissance bactérienne sur 6 souches de bactéries pathogènes a permis de découvrir un nouveau composé (**C15b**) qui a des activités d'inhibition intéressantes chez *K. pneumoniae*, *S. aureus* et *E. coli*. Il est donc prévu d'évaluer le mécanisme d'action de ce composé qui pourra ensuite être optimisé afin d'en améliorer l'activité antimicrobienne et les propriétés pharmacodynamiques/pharmacocinétiques vers une future application thérapeutique.

Parallèlement à ce travail, nous avons effectué la conception d'une nouvelle série de ligands d'ARN sur l'ARN TAR du HIV-1. Dans ce contexte, nous avons tout d'abord effectué un criblage d'une collection de composés propre au laboratoire. Ceci a permis d'identifier un nouveau chémotype capable de former des liaisons avec l'ARN TAR, le composé vérapamil. L'optimisation de ce composé à travers la conception de 21 conjugués a permis de mettre en avant le composé **2h**. Ce dernier a montré non seulement une excellente affinité et sélectivité pour TAR, mais aussi la capacité d'inhiber la formation du complexe Tat/TAR avec un  $IC_{50}$  de l'ordre du micromolaire. Ainsi, l'approche par criblage reste encore très prometteuse pour la découverte de nouveaux ligands d'ARN. Ce travail a été publié dans le journal ChemPlusChem en 2020.

L'ensemble de ce travail a été présenté à plusieurs congrès jeunes chercheurs et internationaux et m'a permis d'être récompensée par le prix du meilleur poster aux Rencontres Internationales de Chimie Thérapeutique (RICT) en 2019. Par ailleurs, j'ai eu l'opportunité d'encadrer un stage M1 et un stage M2 et de participer à d'autres projets de l'équipe. Plusieurs publications sont ainsi en cours de rédaction et la voie est ouverte pour la découverte de nouveaux ligands d'ARN efficaces et vers leur application thérapeutique.

---

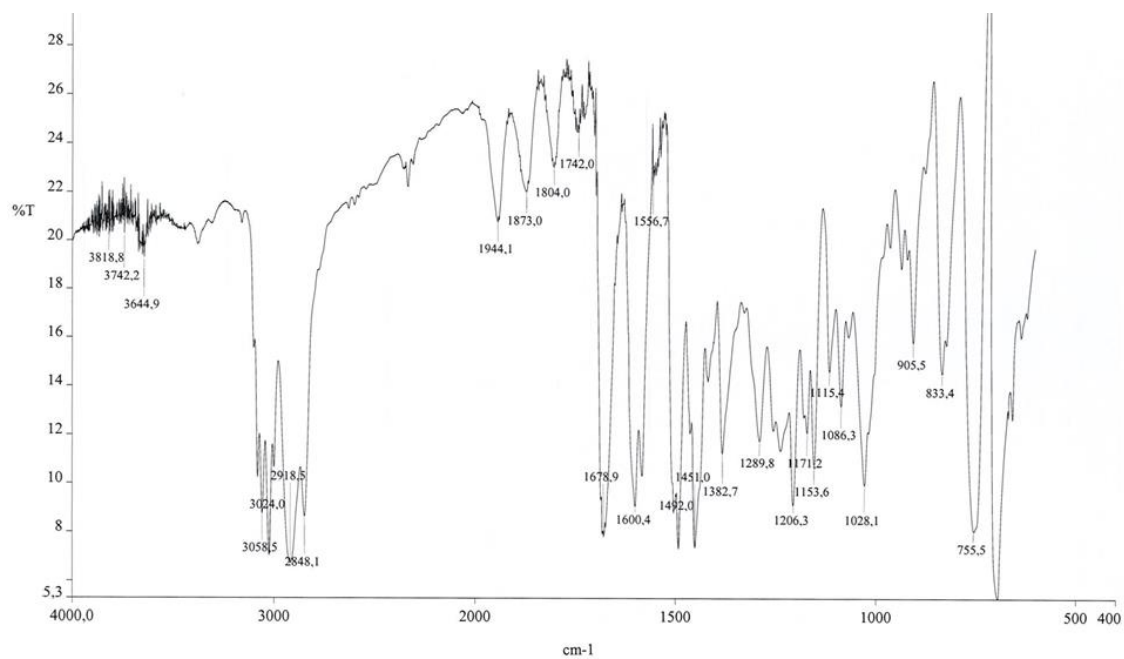
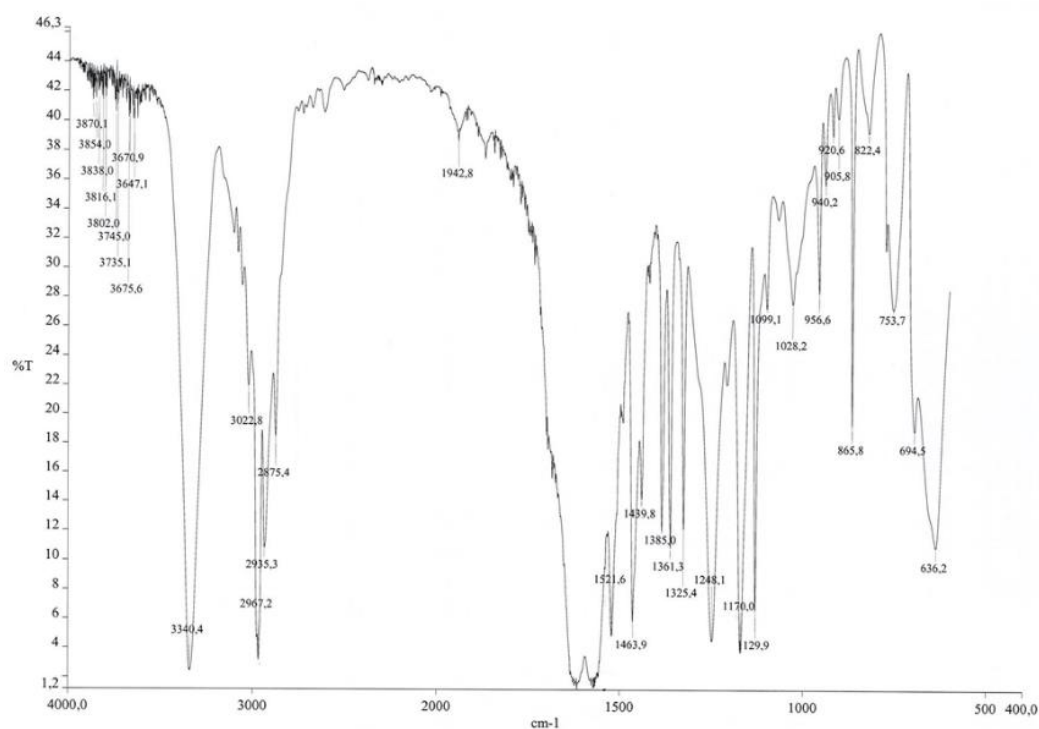
# Annexes

---

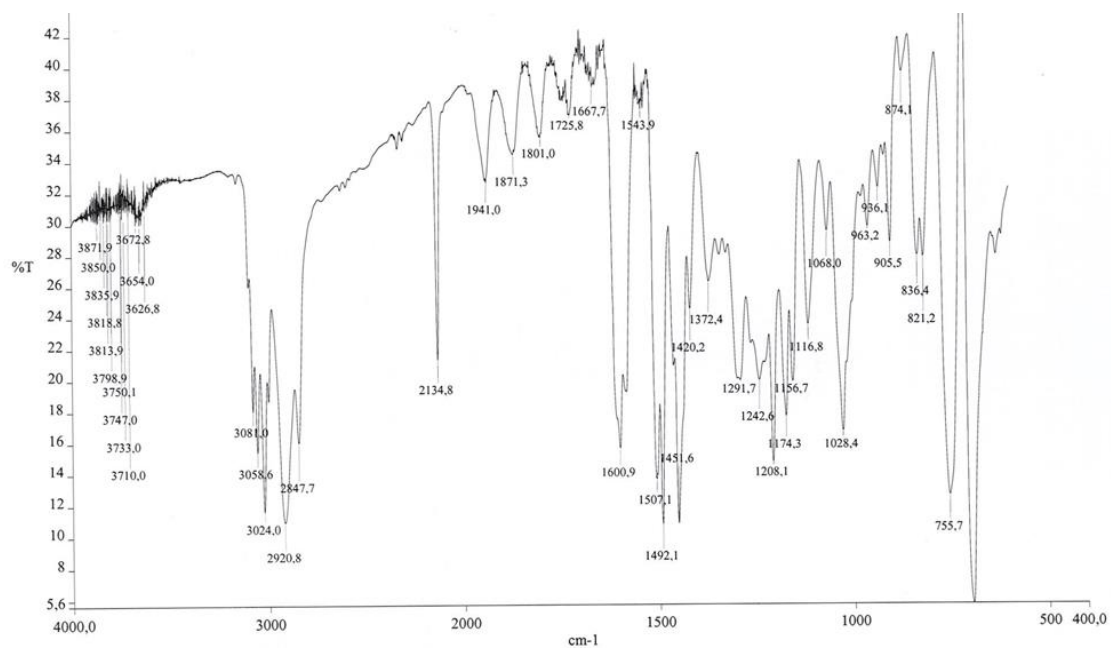


**Annexe 1 :** Caractérisation par spectroscopie infrarouge (KBr) de la fonctionnalisation de la résine

Rink amide :

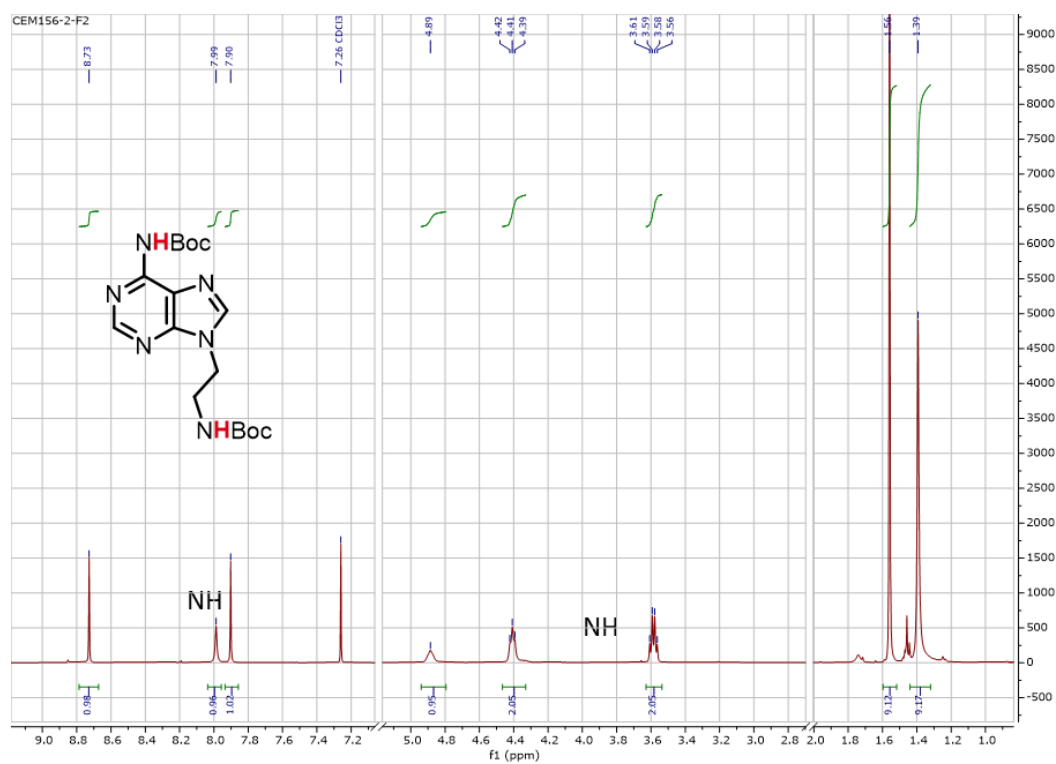
**Étape 1 :** Résine Rink amide déprotégée : résine-NH<sub>2</sub>**Étape 2 :** Résine Rink amide *N*-formyl :  $\nu = 1690 \text{ cm}^{-1}$  (C=O amide),  $\nu = 3340,4 \text{ cm}^{-1}$  (N-H amide) :



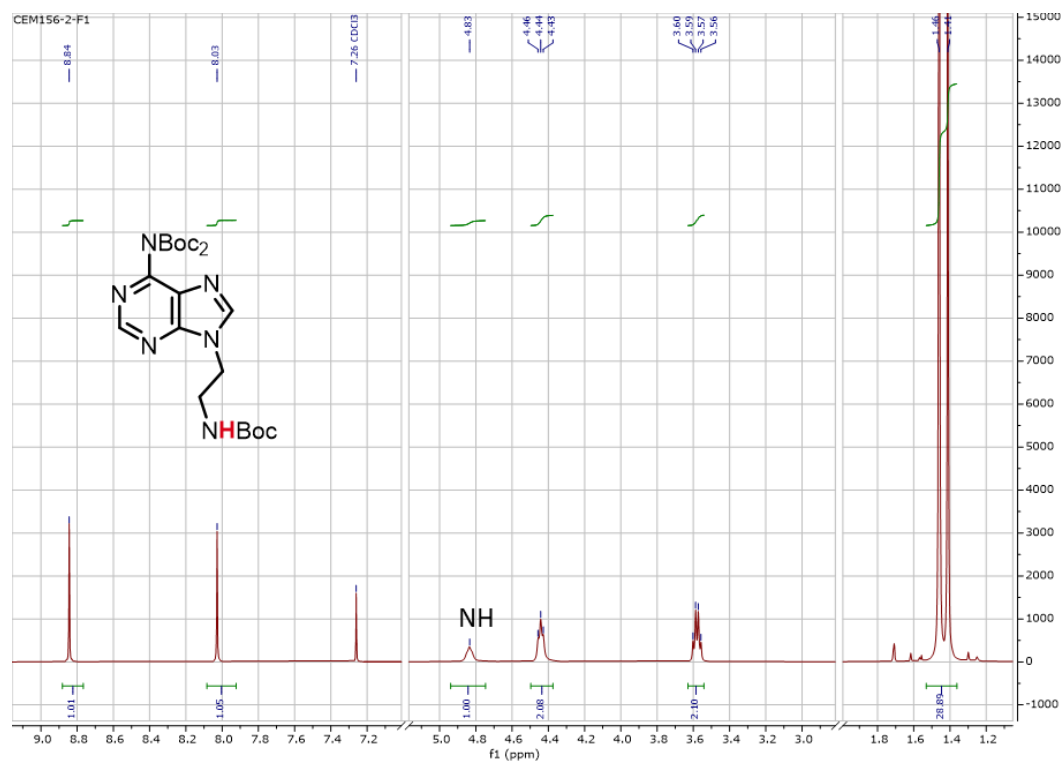
**Étape 3 : Résine Rink amide isocyanure :  $\nu = 2134,8 \text{ cm}^{-1}$  (C≡N isocyanure) :**

**Annexe 2 : RMN  $^1\text{H}$  des produits secondaires obtenus par migration du groupement protecteur Boc :**

**A) produit Adénine di-Boc (28) :**

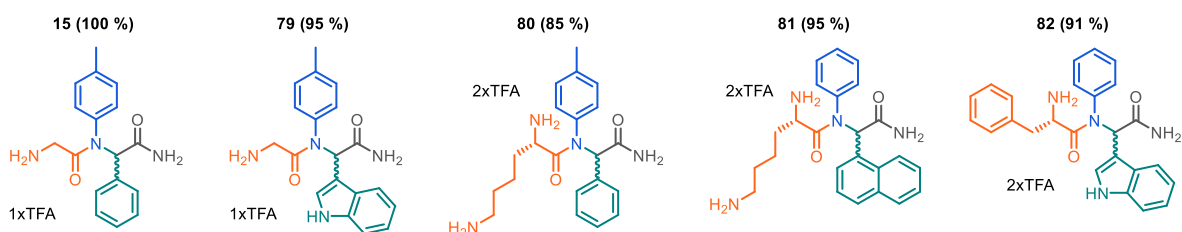


**B) Produit Adénine tri-Boc (27).**

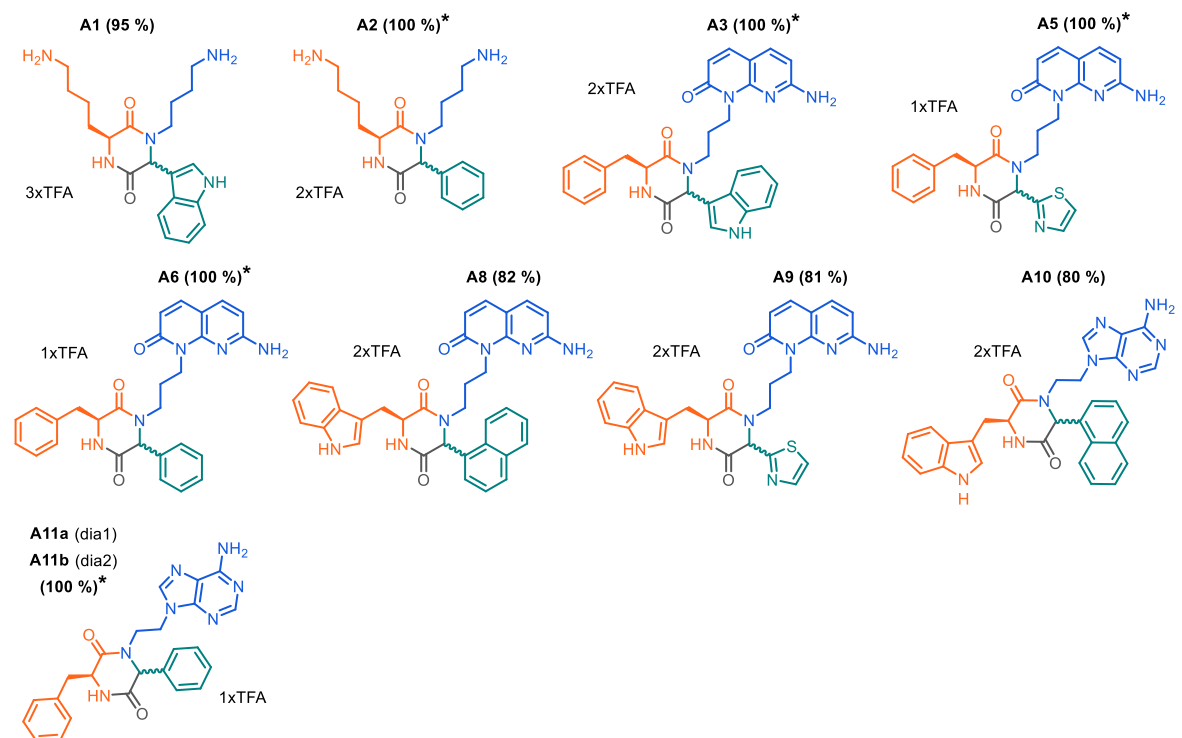


## Annexe 3 : Structures des composés des Séries A et B puis des Séries C-1, C-2 et C-3 :

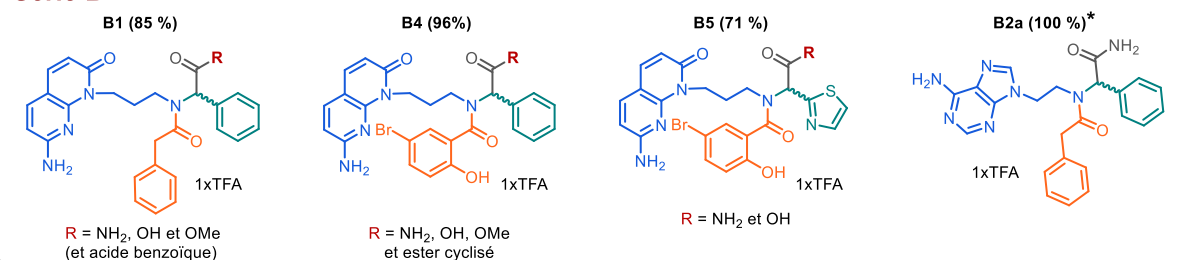
## Etude préliminaire



## Série A

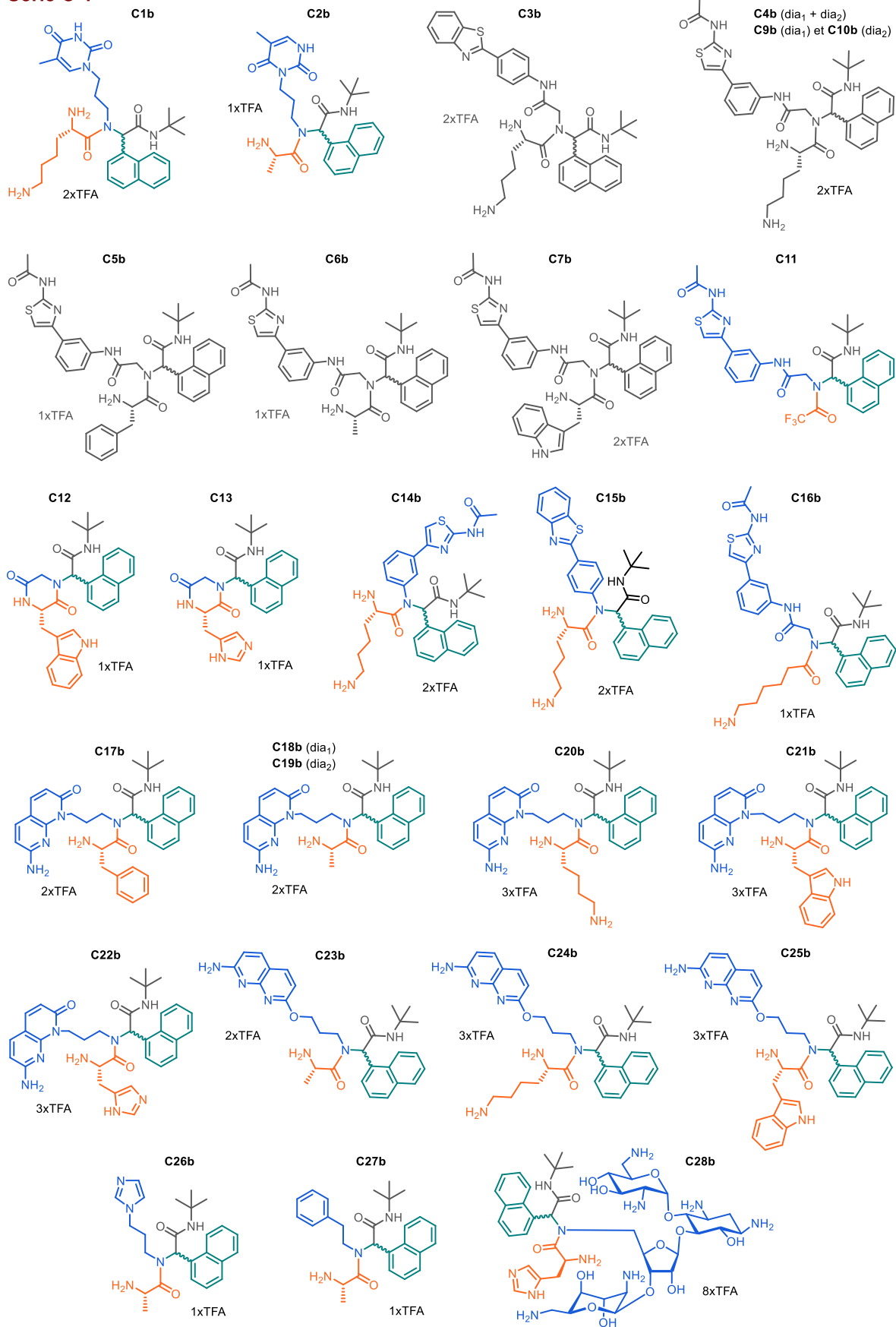


## Série B

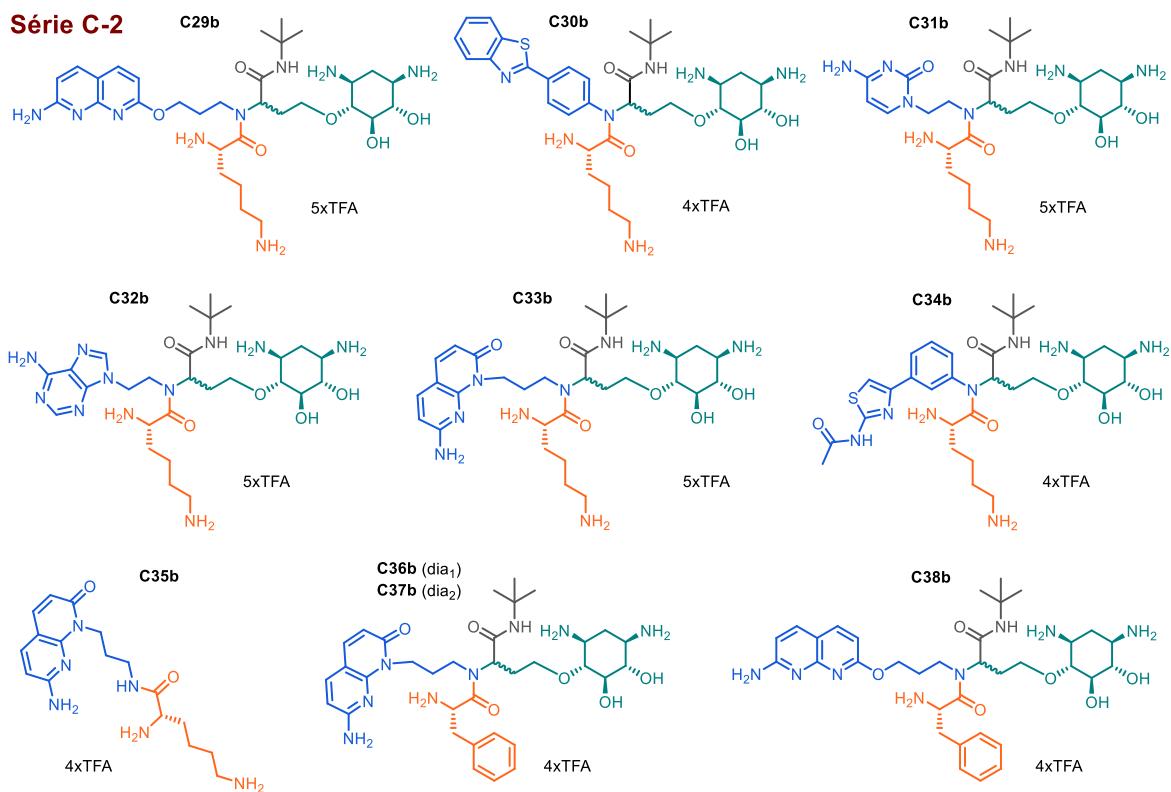


(\* purifiés par HPLC semi-préparative)

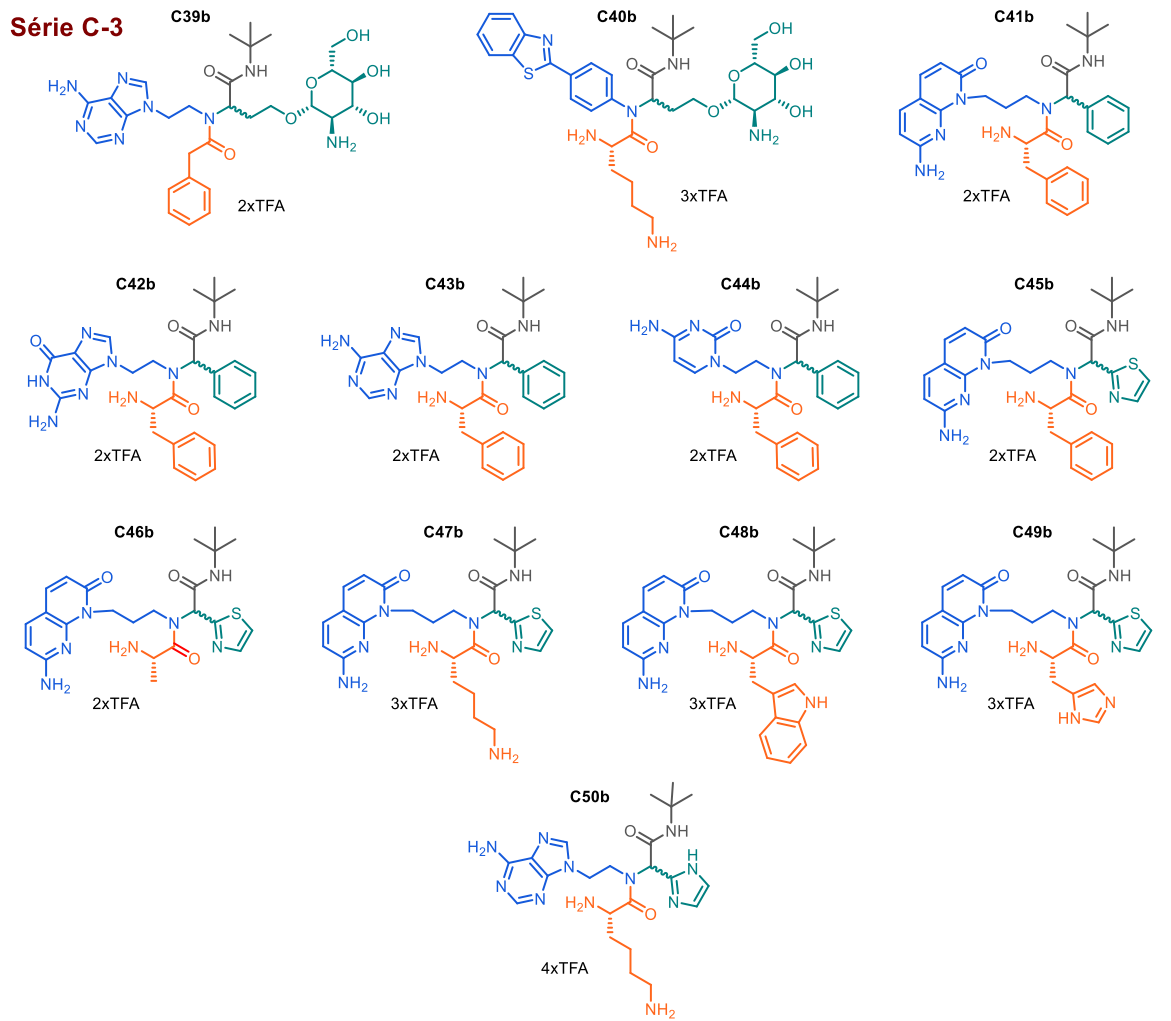
## Série C-1



## Série C-2

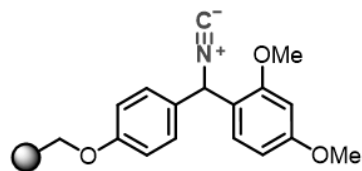


## Série C-3

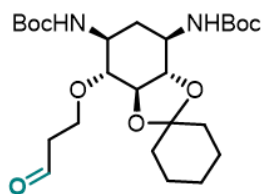
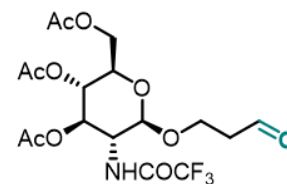


## Annexe 4 : Représentation des composants isocyanures, aldéhydes et amines synthétisées.

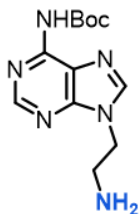
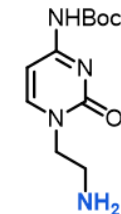
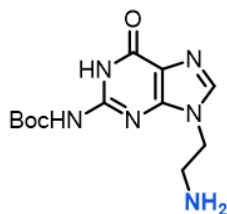
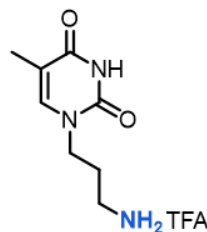
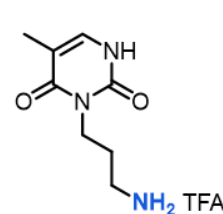
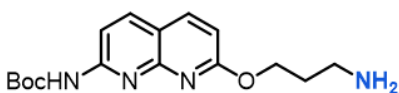
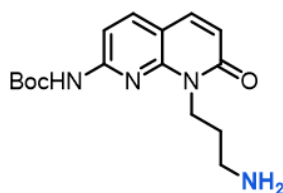
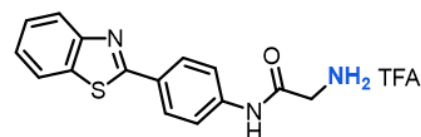
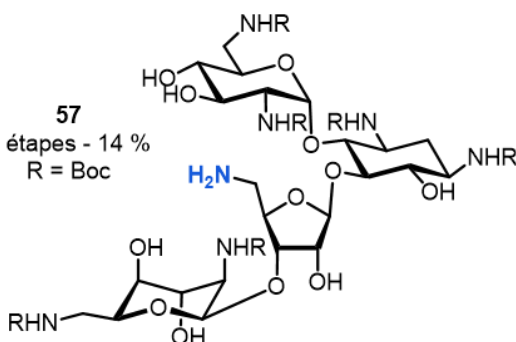
## Support solide

3 étapes  
Rink Amide ou MBHA

## Aldéhydes

58  
(8 étapes - 14 %)74  
(4 étapes - 11 %)

## Amines

21  
4 étapes - 54 %25  
4 étapes - 21 %32  
4 étapes - 54 %34  
2 étapes - 25 %35  
2 étapes - 11 %37  
4 étapes - 41 %42  
4 étapes - 44 %55  
2 étapes - 48 %57  
4 étapes - 14 %  
R = BocR = H : 45  
2 étapes - 56 %  
R = Gly-NH<sub>2</sub> : 49a/b  
4 étapes - 35 %



---

# Partie expérimentale

---





## A) Chemistry

### General procedures and Instrumentation:

Reagents and solvents were purchased from Aldrich, Alfa Aesar, Fluorochem, Carlo Erba, Acros, Novabiochem® and used without further purification. All reactions involving air- or moisture-sensitive reagents or intermediates were performed under an argon atmosphere. Flash column chromatography was carried out on silica gel columns (Interchim Puriflash silica HP 15 to 50µm) Puriflash® XS420 system (Interchim). Analytical thin-layer chromatography (TLC) was conducted on Sigma Aldrich precoated silica gel and compounds were visualized by irradiation (254 nm) and/or by staining with ninhydrin or vanilin stains.

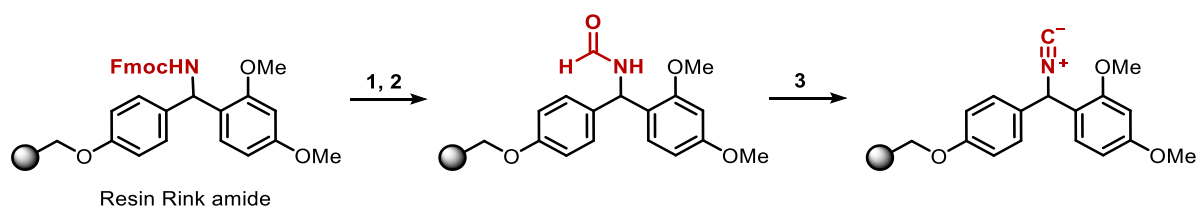
**HPLC analyses** were performed using a Water Alliance 2695 pump coupled to a Water 996 photodiode array detector and a two different column were used. **Method A:** Phenomenex SynergiTM FusionRP 80A reverse-phase C18 column (250 x 4.6 mm, 5 µm) for analytical HPLC analysis and reverse-phase C18 column (250 x 10 mm, 5 µm) for semipreparative HPLC. **Method B:** X-Select CSHTM fluorophenyl reverse-phase C18 column (50 x 4.6 mm, 5 µm) for analytical HPLC analysis and reverse-phase C18 column (250 x 10 mm, 5 µm) for semipreparative HPLC. All HPLC analyses were run at room temperature using a gradient of CH<sub>3</sub>CN containing 0.1% TFA (eluent B) and water containing 0.1% TFA (eluent A) from 5% to 100% of B in 30 min at a flow rate of 1 mL/min for the analytical column and 3.5 mL/min for the semipreparative column.

<sup>1</sup>H and <sup>13</sup>C NMR spectra were recorded on Bruker AC 200 and 400 MHz spectrometers. Chemical shifts are reported in parts per million (ppm, δ) referenced to the residual <sup>1</sup>H resonance of the solvent (D<sub>2</sub>O, δ 4.79; CDCl<sub>3</sub>, δ 7.26; CD<sub>3</sub>OD, δ 3.31 and DMSO-*d*<sub>6</sub>, δ 2.50) and referenced to the residual <sup>13</sup>C resonance of the solvent (CDCl<sub>3</sub>, δ 77.2; CD<sub>3</sub>OD, δ 49.0 and DMSO-*d*<sub>6</sub>, δ 39.52). Splitting patterns are designed as follows: s (singlet), d (doublet), t (triplet), q (quadruplet) p (pentuplet), m (multiplet), and br (broad). Coupling constant (*J*) are listed in hertz (Hz).

Low resolution mass spectra (MS) were obtained with a Bruker Daltonics Esquire 3000+ electrospray spectrometer equipped with API ionization source. High-resolution mass spectrometry (HRMS) was carried out on a LTQ Orbitrap hybrid mass spectrometer with an electrospray ionization probe (ThermoScientific, San Jose, CA) by direct infusion from a pump syringe, to confirm the correct molar mass and high purity of the compounds.

## Experimental part – Chapter II – Functionalization of resin

### Functionalization of resin:



The solid-phase synthesis for the functionalization of the resin was carried on an orbital shaker, at room temperature, in tubes with frits. Before each reaction, the resin was swollen in  $\text{CH}_2\text{Cl}_2$  (10 mL/g of resin) overnight. Each synthesis was carried out in 500 mg scale using Fmoc-protected Rink amide MBHA resin 100-200 mesh (loading 0.36, 0.52 or 0.79 g/mol) or Fmoc-protected Rink amide resin 100-200 mesh (loading 0.45 g/mol) or in 250 mg scale using Fmoc-protected Rink amide NovaGel™ (loading 0.66 g/mol). The steps 1 and 2 were followed by a Kaiser test.

#### Step 1: Fmoc cleavage:

Fmoc cleavage of *N*-Fmoc Rink resin (500 mg) was carried out using a solution of piperidine-DMF (20/80 v/v, 10 mL/g of resin). The resin was stirred 30 min at room temperature, then drained and washed several times with DMF (3×4 mL), MeOH (2×4 mL) and  $\text{CH}_2\text{Cl}_2$  (3×4 mL). The reaction was checked using a Kaiser test (if not positive, step 1 was repeated).

#### Step 2: Coupling step:

Coupling reaction was carried out starting from resins swollen in  $\text{CH}_2\text{Cl}_2$  (4 mL) after step 1. Pyridine (1 eq.) and  $\text{HCO}_2\text{H}$  (5 eq.) were added, followed by DIC (5 eq.). The suspension was stirred for 2h30 at room temperature until completion (usually 3 cycles were necessary, as judged by a Kaiser test). The resin was washed several times with  $\text{CH}_2\text{Cl}_2$  (3×4 mL), MeOH (2×4 mL) and  $\text{CH}_2\text{Cl}_2$  (3×4 mL).

#### Step 3: Dehydration step:

*N*-formylated Rink resin was suspended in dry  $\text{CH}_2\text{Cl}_2$  (4 mL). After the addition of anhydrous *N,N*-diisopropylamine (15 eq.), anhydrous  $\text{POCl}_3$  (5 eq.) was added dropwise. The suspension was stirred for 2h30 at room temperature. The cycle was repeated twice. The resin was drained and washed several times with  $\text{CH}_2\text{Cl}_2$  (3×3 mL), MeOH (2×3 mL),  $\text{CH}_2\text{Cl}_2$  (3×3 mL), MeOH (2×3 mL) and  $\text{Et}_2\text{O}$  (2×3 mL). Drying under reduced pressure afforded the Rink-isocyanure resin as a pale-yellow resin. The resin was stored under argon and the presence of the isocyanide group was confirmed by IR spectroscopy.

Rink Amide MBHA isocyanide: IR (KBr) :  $\nu = 2135$  ( $\text{C}\equiv\text{N}$ )

Rink Amide isocyanide: IR (KBr) :  $\nu = 2135$  ( $\text{C}\equiv\text{N}$ )

**Estimation of isocyanide resin substitution rate:****Method 1:**

A mixture of *N*-benzylideneaniline (5 eq.) and Fmoc-Gly-OH (5eq.) was poured into a reaction tube containing the appropriate swollen resin-isocyanide (1 eq.) suspended in dry CH<sub>2</sub>Cl<sub>2</sub>/MeOH (0.4 mL, v/v). The mixture was stirred for 4-6 days at room temperature, then drained and washed several times with DMF (3×3 mL), MeOH (3×3 mL) and CH<sub>2</sub>Cl<sub>2</sub> (3×3 mL). Drying under reduced pressure afforded the Ugi-resin as a pale-yellow resin. Approximately 10 mg of resin were transferred to a 10 mL flask and 1 mL of a 20% piperazine solution in DMF (v/v) was added. The suspension was agitated for 15 minutes and diluted with DMF up to the 10 mL. A volume of 4 mL of the filtrate was transferred to a 25 mL flask and diluted with DMF (21 mL). The filtrate was analyzed for dibenzofulvene adduct by UV-Vis. spectrometer. For each time point, absorbance spectrum was recorded and after subtraction of blank (spectrum for deprotection solution itself), absorbance at 301 nm (*A*<sub>t</sub>) was noted. The substitution rate was calculated by inserting the absorption at 301 nm and the amount of resin tested in the equation below:

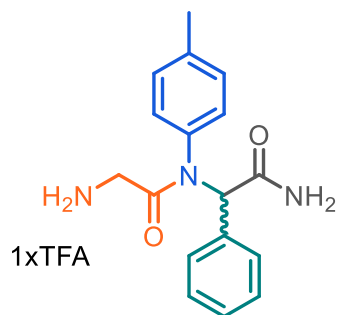
$$\text{loading (mmol.g}^{-1}\text{)} = \frac{A_t \times D \times V_0}{\epsilon \times l \times m}$$

Where: **A**<sub>t</sub>= Absorbance of sample (λ= 301 nm), **D**= Dilution factor, **V**<sub>0</sub>= Sample volume (ml), **ε**= Molar absorption coefficient at 301 nm: ε<sub>301 nm</sub> = 7800 ml.mmol<sup>-1</sup>.cm<sup>-1</sup>, **l**= optical path length of the cell in cm and **m** = sample weight of the resin (mg).

**Method 2:**

The appropriate benzaldehyde (2 eq.) and *p*-toluidine (2 eq.) were dissolved in dry MeOH (0.2 mL) and stirred for 1h at room temperature. Then, a solution of Boc-Gly-OH (2 eq.) in dry MeOH (0.2 mL) was added to the mixture. The mixture was poured into a reaction tube containing the appropriate swollen resin-isocyanide (1 eq.) suspended in dry CH<sub>2</sub>Cl<sub>2</sub> (0.4 mL). The mixture was stirred for 4-6 days at room temperature, then drained and washed several times with DMF (3×3 mL), MeOH (3×3 mL) and CH<sub>2</sub>Cl<sub>2</sub> (3×3 mL). The product was removed from the resin upon cleavage with an appropriated stock solution of 15% trifluoroacetic acid (TFA) (Rink amide) or 95% TFA (Rink amide MBHA) in CH<sub>2</sub>Cl<sub>2</sub> for 1h30. The remaining cleavage mixture of TFA was removed under reduced pressure by co-evaporating twice with toluene, then the crude mixture was triturated in Et<sub>2</sub>O and filtrated.

Example of yield (five steps), purity, and some characterizations of mixture of enantiomers obtained using method 2:

***p*-Toluidine-(R/S)-Phenyl-Gly-CONH<sub>2</sub> – (15)**

**m<sub>Resin</sub>** = 250 mg : Rink Amide MBHA (loading =0.79 mmol/g).

**m** = 19.2 mg

**Yield** = 24%

**Purity** = 100%

**HPLC**: *rt*<sub>dia1</sub> = 13.35 min (method A).

**HRMS (ESI)**: *m/z* 298.15512 [M+H]<sup>+</sup> (C<sub>17</sub>H<sub>20</sub>N<sub>3</sub>O<sub>2</sub> requires 298.15500).

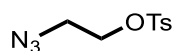
<sup>1</sup>H NMR (400 MHz, MeOD)  $\delta$  = 7.79 – 6.31 (m, 9H, 9CH<sub>Ar</sub>), 6.16 (s, 1H, CH), 3.52 – 3.38 (m, 2H, CH<sub>2</sub>), 2.24 (s, 3H, CH<sub>3</sub>).

<sup>13</sup>C NMR (101 MHz, MeOD)  $\delta$  = 174.4 (C=O), 167.6 (C=O), 140.4, 135.9, 134.9 (3C<sub>IV,Ar</sub>), 131.8, 131.6, 130.9, 129.6, 129.3, 123.5 (9CH<sub>Ar</sub>), 66.2 (CH), 42.1 (CH<sub>2</sub>), 21.0 (CH<sub>3</sub>).

## Experimental part – Chapter II - Synthesis of amine fragments

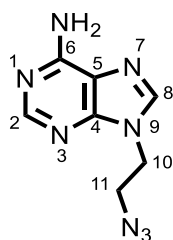
### 1) Synthesis of Boc-Adenine-NH<sub>2</sub> (21):

#### 2-azidoethyl-*p*-toluènesulfonate (17):



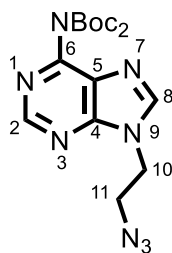
Compound 2-azidoethyl-*p*-toluènesulfonate **17** was obtained according to previously published procedure<sup>214</sup> in 69 % yield using compound commercially available 2-bromoethanol as starting material.

#### N<sup>9</sup>-(2-azidoethyl)-adenine (18):



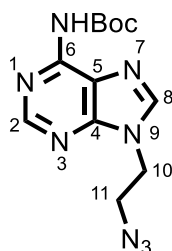
To a solution of adenine (200 mg, 1.48 mmol, 1 eq.) in anhydrous DMF (10 mL) was added NaH (35.5 mg, 1.48 mmol, 1 eq.). The reaction was stirred at room temperature for 30 min. before addition of 2-azidoethyl-*p*-toluènesulfonate (392 mg, 1.60 mmol, 1.1 eq.). The reaction mixture was stirred overnight at 60°C, then concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 96:4, v/v) to afford compound (**18**) (220 mg, 73%) as a white solid.

**R<sub>f</sub>** = 0.28 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 96:4); **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>) δ = 8.16, 8.15 (2s, 1H, H<sub>2</sub> + H<sub>8</sub>), 7.24 (s, 2H, NH<sub>2</sub>), 4.34 (t, *J* = 5.7 Hz, 2H, H<sub>10</sub>), 3.81 (t, *J* = 5.7 Hz, 2H, H<sub>11</sub>); **<sup>13</sup>C NMR** (101 MHz, DMSO-*d*<sub>6</sub>) δ = 156.0 (C<sub>6</sub>), 152.6 (C<sub>2</sub>), 149.6 (C<sub>4</sub>), 140.9 (C<sub>8</sub>), 118.7 (C<sub>5</sub>), 49.7 (C<sub>11</sub>), 42.5 (C<sub>10</sub>); **MS (ESI)**: *m/z* 204.93 [M+H]<sup>+</sup> (theoretical *m/z* 205.09).

***N*<sup>6</sup>,*N*<sup>6</sup>-Bis(*tert*-butoxycarbonyl)-*N*<sup>9</sup>-(2-azidoethyl)-adenine (**19**)**

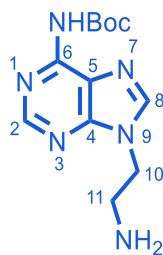
To a suspension of compound **18** (200 mg, 0.98 mmol, 1 eq.) and DMAP (12.0 mg, 0.098 mmol, 0.1 eq.) in THF (9 mL) at 0°C, Boc<sub>2</sub>O (500 mg, 2.27 mmol, 2.3 eq.) was added. The reaction mixture was stirred overnight at room temperature, then was concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 95:5, v/v) to afford compound **19** (380 mg, 96%) as a white solid.

*R*<sub>f</sub> = 0.58 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 95:5); <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ = 8.86 (s, 1H, H<sub>2</sub>), 8.67 (s, 1H, H<sub>8</sub>), 4.50 (t, *J* = 5.6 Hz, 2H, H<sub>10</sub>), 3.89 (t, *J* = 5.6 Hz, 2H, H<sub>11</sub>), 1.36 (s, 18H, 6CH<sub>3</sub>.Boc); <sup>13</sup>C NMR (101 MHz, DMSO-*d*<sub>6</sub>) δ = 153.3 (C<sub>4</sub>), 151.6 (C<sub>2</sub>), 150.0 (2C<sub>O</sub>Boc), 149.1 (C<sub>6</sub>), 146.9 (C<sub>8</sub>), 127.8 (C<sub>5</sub>), 83.3 (2C<sub>N</sub>.Boc), 49.6 (C<sub>11</sub>), 43.1 (C<sub>10</sub>), 27.3 (6CH<sub>3</sub>.Boc); **MS (ESI):** *m/z* 404.73 [M+H]<sup>+</sup> (theoretical *m/z* 405.20).

***N*<sup>6</sup>-(*tert*-butoxycarbonyl)-*N*<sup>9</sup>-(2-azidoethyl)-adenine (**20**)**

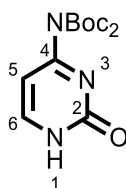
To a solution of compound **19** (3.68 g, 9.10 mmol, 1 eq.) in MeOH (160 mL) was added K<sub>2</sub>CO<sub>3</sub> (2.51 g, 18.2 mmol, 2.5 eq.). The reaction mixture was stirred 2h at reflux, then concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 95:5, v/v) to afford compound **20** (2.24 g, 81%) as a white solid.

*R*<sub>f</sub> = 0.35 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 95:5); <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ = 8.60 (s, 1H, H<sub>2</sub>), 8.34 (s, 1H, H<sub>8</sub>), 4.36 (t, *J* = 5.7 Hz, 2H, H<sub>10</sub>), 3.13 (t, *J* = 5.7 Hz, 2H, H<sub>11</sub>), 1.60 (s, 9H, 3CH<sub>3</sub>.Boc); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD) δ = 153.2 (C<sub>2</sub>), 152.8, 152.4 (C<sub>4</sub> + C<sub>6</sub>), 151.3 (CO<sub>Boc</sub>), 145.3 (C<sub>8</sub>), 122.9 (C<sub>5</sub>), 82.8 (C<sub>N</sub>.Boc), 51.3 (C<sub>11</sub>), 44.4 (C<sub>10</sub>), 28.4 (3CH<sub>3</sub>.Boc); **MS (ESI):** *m/z* 304.73 [M+H]<sup>+</sup> (theoretical *m/z* 305.15).

**Boc-Adenine-NH<sub>2</sub> = N<sup>6</sup>-(tert-butoxycarbonyl)-N<sup>9</sup>-(2-aminoethyl)-adenine (21)**

To a solution of compound **20** (2.33 g, 7.65 mmol, 1 eq.) in EtOH (40 mL) was added 10% Pd/C (200 mg). The system was degassed three times (vacuum/N<sub>2</sub>), then stirred under H<sub>2</sub> (1 atm) for 12h. The mixture was filtered through a pad of Celite<sup>®</sup> which was washed with EtOH. After solvent removal, the crude product was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH [10% NH<sub>4</sub>OH] gradient 100:0 to 9:1, v/v) to afford compound **21** (2.23 g, 95%) as a white solid.

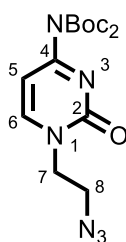
**R<sub>f</sub>** = 0.29 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH [10% NH<sub>4</sub>OH] 92:8); **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD) δ = 8.56 (s, 1H, H<sub>2</sub>), 8.29 (s, 1H, H<sub>8</sub>), 4.35 (t, *J* = 6.1 Hz, 2H, H<sub>10</sub>), 3.12 (t, *J* = 6.1 Hz, 2H, H<sub>11</sub>), 1.59 (s, 9H, 3CH<sub>3</sub>.Boc); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD) δ = 153.0 (C<sub>2</sub>), 152.9 (C<sub>4</sub>), 152.5 (C<sub>6</sub>), 151.2 (CO<sub>Boc</sub>), 145.4 (C<sub>8</sub>), 123.0 (C<sub>5</sub>), 82.7 (C<sub>IV.Boc</sub>), 47.45 (C<sub>10</sub>), 42.0 (C<sub>11</sub>), 28.4 (3CH<sub>3</sub>.Boc); **MS (ESI)**: *m/z* 278.87 [M+H]<sup>+</sup> (theoretical *m/z* 279.16).

**2) Synthesis of Boc-Cytosine-NH<sub>2</sub> (25):****N<sup>4</sup>,N<sup>4</sup>-Bis(tert-butoxycarbonyl)-cytosine (22):**

To a suspension of cytosine (1.00 g, 9.00 mmol, 1 eq.) and DMAP (110 mg, 0.90 mmol, 0.1 eq.) in THF (40 mL) at 0°C, Boc<sub>2</sub>O (6.87 g, 31.50 mmol, 3.5eq.) was added. The reaction mixture was stirred at room temperature until the starting solid was entirely solubilized due to alkylation, then it was concentrated under reduced pressure. The fully protected compound was dissolved in MeOH (40 mL) and a saturated NaHCO<sub>3</sub> aq. solution (20 mL) was added. The reaction mixture was stirred 2h at reflux to free the N1 position. After solvent removal, the aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub>. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 95:5, v/v) to afford compound **22** (2.18 g, 78%) as a white solid.

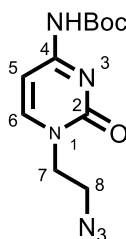
**R<sub>f</sub>** = 0.65 (EtOAc 100%); **<sup>1</sup>H NMR** (400 MHz, CDCl<sub>3</sub>) δ = 7.78 (d, *J* = 7.1 Hz, 1H, H<sub>6</sub>), 7.09 (d, *J* = 7.1 Hz, 1H, H<sub>5</sub>), 1.54 (s, 18H, 6CH<sub>3</sub>.Boc); **<sup>13</sup>C NMR** (101 MHz, CDCl<sub>3</sub>) δ = 163.8 (C<sub>4</sub>), 158.6 (C<sub>2</sub>), 149.6 (2CO<sub>Boc</sub>), 146.1 (C<sub>6</sub>), 96.9 (C<sub>5</sub>), 85.0 (C<sub>IV.Boc</sub>), 27.8 (6CH<sub>3</sub>.Boc); **MS (ESI)**: *m/z* 312.6 [M+H]<sup>+</sup> (theoretical *m/z* 312.16).



***N*<sup>1</sup>-(2-azidoethyl)-*N*<sup>4</sup>,*N*<sup>4</sup>-Bis(*tert*-butoxycarbonyl)-cytosine (**23**):**

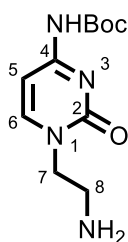
To a solution of compound **22** (1.63 g, 5.23 mmol, 1 eq.) in anhydrous DMF (40 mL) was added NaH (151 mg, 6.28 mmol, 1.2 eq.) The reaction was stirred at room temperature for 30 min. before addition of 2-azidoethyl-*p*-toluenesulfonate (2.00 g, 8.30 mmol, 1.6 eq.). The reaction mixture was stirred overnight at 60°C, then concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 98:2, v/v) to afford compound (**23**) (875 mg, 44%).

*R*<sub>f</sub> = 0.27 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 98:2); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ = 7.53 (d, *J* = 7.3 Hz, 1H, H<sub>6</sub>), 7.05 (d, *J* = 7.3 Hz, 1H, H<sub>5</sub>), 3.97 – 3.90 (m, 2H, H<sub>8</sub>), 3.77 – 3.70 (m, 2H, H<sub>7</sub>), 1.54 (s, 18H, 6CH<sub>3</sub>.Boc); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ = 162.9 (C<sub>4</sub>), 154.8 (C<sub>2</sub>), 149.6 (2CO<sub>Boc</sub>), 148.6 (C<sub>6</sub>), 96.3 (C<sub>5</sub>), 85.1 (2C<sub>IV.Boc</sub>), 50.5 (C<sub>8</sub>), 49.2 (C<sub>7</sub>), 27.8 (6CH<sub>3</sub>.Boc).

***N*<sup>1</sup>-(2-azidoethyl)-*N*<sup>4</sup>-(*tert*-butoxycarbonyl)-cytosine (**24**):**

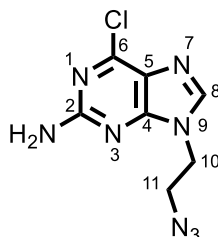
To a solution of compound **23** (450 mg, 1.18 mmol, 1 eq.) in MeOH (10 mL) were added K<sub>2</sub>CO<sub>3</sub> (164 mg, 1.18 mmol, 1 eq.). The reaction mixture was stirred 1h at reflux, then concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 96:4, v/v) to afford compound **24** (208 mg, 63%) as a white solid.

*R*<sub>f</sub> = 0.12 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 98:2); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ = 7.55 (d, *J* = 7.3 Hz, 1H, H<sub>6</sub>), 7.42 (s, NH), 7.20 (d, *J* = 7.3 Hz, 1H, H<sub>5</sub>), 3.99 – 3.93 (m, 2H, H<sub>8</sub>), 3.77 – 3.70 (m, 2H, H<sub>7</sub>), 1.51 (s, 9H, 3CH<sub>3</sub>.Boc); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ = 163.1 (C<sub>4</sub>), 155.7 (C<sub>2</sub>), 151.3 (CO<sub>Boc</sub>), 149.0 (C<sub>6</sub>), 94.9 (C<sub>5</sub>), 83.0 (C<sub>IV.Boc</sub>), 50.4 (C<sub>8</sub>), 49.4 (C<sub>7</sub>), 28.1 (3CH<sub>3</sub>.Boc); **MS (ESI)**: *m/z* 281.6 [M+H]<sup>+</sup> (theoretical *m/z* 281.14).

**Boc-Cytosine-NH<sub>2</sub> : N<sup>1</sup>-(2-aminoethyl)-N<sup>4</sup>-(*tert*-butoxycarbonyl)-cytosine (25):**

To a solution of compound **24** (197 mg, 0.70 mmol, 1 eq.) in CH<sub>2</sub>Cl<sub>2</sub> (2 mL) was added 10% Pd/C (10 mg). The system was degassed three times (vacuum/N<sub>2</sub>), then stirred under H<sub>2</sub> (1 atm.) for 3h. The resulting precipitate was diluted with EtOH. Then, the mixture was filtered through a pad of Celite® and washed with EtOH. After solvent removal, the crude product was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH [10% NH<sub>4</sub>OH] gradient 100:0 to 9:1, v/v) to afford compound **25** (173 g, 97%) as a white solid.

R<sub>f</sub> = 0.22 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH [10% NH<sub>4</sub>OH] 91:9); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ = 7.68 (d, *J* = 7.2 Hz, 1H, H<sub>6</sub>), 7.13 (br d, *J* = 7.2 Hz, 1H, H<sub>5</sub>), 3.91 (t, *J* = 5.9 Hz, 2H, H<sub>7</sub>), 3.08 (t, *J* = 5.9 Hz, 2H, H<sub>8</sub>), 1.49 (s, 9H, 3CH<sub>3</sub>.Boc); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ = 162.8 (C<sub>4</sub>), 156.2 (C<sub>2</sub>), 151.5 (CO<sub>Boc</sub>), 149.2 (C<sub>6</sub>), 94.6 (C<sub>5</sub>), 82.7 (C<sub>IV.Boc</sub>), 53.5 (C<sub>7</sub>), 40.3 (C<sub>8</sub>), 28.2 (3CH<sub>3</sub>.Boc); MS (ESI): *m/z* 255.7 [M+H]<sup>+</sup> (theoretical *m/z* 255.15).

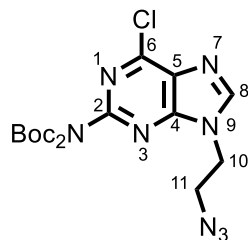
**3) Synthesis of Boc-Guanine-NH<sub>2</sub> (32):****2-Amino-6-chloro-N<sup>9</sup>-(2-azidoethyl)purine (29):**

To a suspension of 6-amino-2-chloropurine (300 mg, 1.77 mmol, 1 eq.) in anhydrous DMF (8 mL) was added K<sub>2</sub>CO<sub>3</sub> (366 mg, 2.65 mmol, 1.5 eq.). The reaction was stirred at 85°C for 30 min. before addition of 2-azidoethyl-*p*-toluenesulfonate (470 mg, 1.95 mmol, 1.1 eq.). The mixture was stirred overnight at room temperature then concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 98:2, v/v) to afford compound **29** (303 mg, 72%) as a white solid.

R<sub>f</sub> = 0.30 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ = 8.15 (s, 1H, H<sub>8</sub>), 6.96 (s, NH<sub>2</sub>), 4.24 (t, *J* = 5.7 Hz, 2H, H<sub>10</sub>), 3.79 (t, *J* = 5.7 Hz, 2H, H<sub>11</sub>); <sup>13</sup>C NMR (101 MHz, DMSO-*d*<sub>6</sub>) δ = 159.8 (C<sub>2</sub>), 154.2 (C<sub>4</sub>),

149.4 (C<sub>6</sub>), 143.2 (C<sub>8</sub>), 123.3 (C<sub>5</sub>), 49.3 (C<sub>11</sub>), 42.5 (C<sub>10</sub>); **MS (ESI):**  $m/z$  239.4 [M+H]<sup>+</sup> (theoretical  $m/z$  239.06).

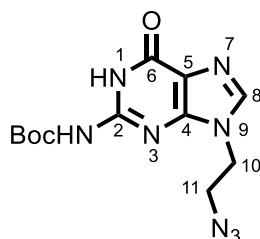
**N,N'-Bis(tert-butyl)-(9-(2-azidoethyl)-6-chloro-9H-purin-2-yl)carbamate (30):**



To a suspension of compound **29** (200 mg, 0.838 mmol, 1 eq.) and DMAP (18.4 mg, 0.151 mmol, 0.2 eq.) in anhydrous THF (7 mL) at 0°C, Boc<sub>2</sub>O (548 mg, 2.51 mmol, 3 eq.) was added. The reaction mixture was stirred overnight at room temperature then concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v) to afford compound **30** (95 mg, 57%) as a white solid.

**R<sub>f</sub>** = 0.25 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 99:1); **<sup>1</sup>H NMR** (400 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  = 8.81 (s, 1H, H<sub>8</sub>), 4.52 – 4.45 (m, 2H, H<sub>10</sub>), 3.91 – 3.84 (m, 2H, H<sub>11</sub>), 1.38 (s, 18H, 6CH<sub>3</sub>.Boc); **<sup>13</sup>C NMR** (101 MHz, DMSO-*d*<sub>6</sub>)  $\delta$  = 153.0 (C<sub>4</sub>), 150.8 (C<sub>2</sub>), 150.1 (2CO<sub>Boc</sub>), 149.1, 148.8 (C<sub>6</sub> + C<sub>8</sub>), 129.3 (C<sub>5</sub>), 83.3 (2C<sub>IV.Boc</sub>), 49.6 (C<sub>11</sub>), 43.4 (C<sub>10</sub>), 27.3 (6CH<sub>3</sub>.Boc); **MS (ESI):**  $m/z$  439.4 [M+H]<sup>+</sup> (theoretical  $m/z$  439.16).

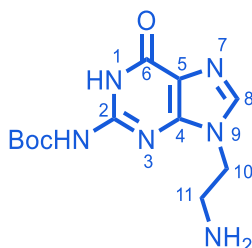
**N-tert-butyl-(9-(2-azidoethyl)-6-oxo-6,9-dihydro-1H-purin-2-yl)carbamate (31):**



A solution of 3-hydroxypropionitrile (109 mg, 1.54 mol, 4.5 eq.) in anhydrous THF (10 mL) was cooled to -70°C, and NaH (69 mg, 1.70 mmol, 5 eq.) was added and stirred for 30 min. Compound **30** (150 mg, 0.34 mmol, 1 eq.) in THF (2 mL) was added in one portion via syringe and the reaction was allowed to warm to room temperature. The reaction was stirred for 3h, then a saturated solution of NH<sub>4</sub>Cl was added and the solvent was evaporated. Water was added, and the aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub>. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and then concentrated under reduced pressure. The crude residue was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 96:4, v/v) to afford compound **31** (97 mg, 89%) as a white solid.

$R_f = 0.38$  ( $\text{CH}_2\text{Cl}_2$ -MeOH 95:5);  $^1\text{H NMR}$  (400 MHz,  $\text{DMSO}-d_6$ )  $\delta = 11.35$  (br, NH), 7.97 (s, 1H,  $\text{H}_8$ ), 4.22 (t,  $J = 5.7$  Hz, 2H,  $\text{H}_{10}$ ), 3.79 (t,  $J = 5.7$  Hz, 2H,  $\text{H}_{11}$ ), 1.49 (s, 9H,  $3\text{CH}_3$ .*Boc*);  $^{13}\text{C NMR}$  (101 MHz,  $\text{DMSO}-d_6$ )  $\delta = 155.1$ , 153.9 ( $\text{C}_6 + \text{CO}_{\text{Boc}}$ ), 149.3, 147.7 ( $\text{C}_4 + \text{C}_2$ ), 139.5 ( $\text{C}_8$ ), 119.5 ( $\text{C}_5$ ), 82.5 ( $\text{C}_{\text{IV.Boc}}$ ), 49.6 ( $\text{C}_{11}$ ), 42.6 ( $\text{C}_{10}$ ), 27.8 ( $3\text{CH}_3$ .*Boc*); **MS (ESI):**  $m/z$  321.5  $[\text{M}+\text{H}]^+$  (theoretical  $m/z$  321.14).

**Boc-Guanine-NH<sub>2</sub> - *N*-*tert*-butyl-(9-(2-aminoethyl)-6-oxo-6,9-dihydro-1*H*-purin-2-yl)carbamate (32):**

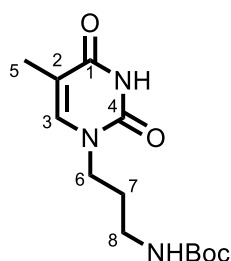


To a solution of compound **31** (800 mg, 2.50 mmol, 1 eq) in  $\text{CH}_2\text{Cl}_2$  (50 mL) was added 10% of Pd/C (80 mg). The system was degassed three times (vacuum/ $\text{N}_2$ ), then stirred under  $\text{H}_2$  (1 atm.) for 16h. The resulting precipitate was diluted with EtOH. Then, the mixture was filtered through a pad of Celite<sup>®</sup> which was washed with  $\text{CH}_2\text{Cl}_2$  and EtOH. After solvent removal, the crude was precipitated in  $\text{CH}_2\text{Cl}_2$  to afford compound **32** (736 mg, 88%) as a white solid.

$R_f = 0.20$  ( $\text{CH}_2\text{Cl}_2$ -MeOH [10%  $\text{NH}_4\text{OH}$ ] 91:9);  $^1\text{H NMR}$  (400 MHz,  $\text{DMSO}-d_6$ )  $\delta = 7.89$  (s, 1H,  $\text{H}_8$ ), 3.99 (t,  $J = 6.2$  Hz, 2H,  $\text{H}_{10}$ ), 2.90 (t,  $J = 6.2$  Hz, 2H,  $\text{H}_{11}$ ), 1.48 (s, 9H,  $3\text{CH}_3$ .*Boc*);  $^{13}\text{C NMR}$  (101 MHz,  $\text{DMSO}-d_6$ )  $\delta = 155.4$ , 154.2 ( $\text{C}_6 + \text{CO}_{\text{Boc}}$ ), 149.2 ( $\text{C}_4$ ), 147.8 ( $\text{C}_2$ ), 139.7 ( $\text{C}_8$ ), 119.4 ( $\text{C}_5$ ), 82.0 ( $\text{C}_{\text{IV.Boc}}$ ), 46.2 ( $\text{C}_{10}$ ), 41.17 ( $\text{C}_{11}$ ), 27.8 ( $3\text{CH}_3$ .*Boc*); **MS (ESI):**  $m/z$  295.6  $[\text{M}+\text{H}]^+$  (theoretical  $m/z$  295.15).

**4) Synthesis of *N*<sup>1</sup>- and *N*<sup>3</sup>- Thymine-NH<sub>2</sub> (34 and 35):**

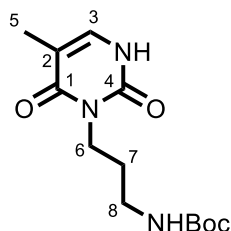
***N*<sup>1</sup>-(3-(*N*<sup>1</sup>-*tert*-butoxycarbonyl)aminopropyl)-thymine (33a)**



To a solution of thymine (780 mg, 6.20 mmol, 1.2 eq.) in DMF (37 mL) was added  $\text{K}_2\text{CO}_3$  (932 mg, 6.72 mmol, 1.3 eq.), then 3-(Boc-amino)propyl bromide (1.23 g, 5.17 mmol, 1 eq.) was added under argon and the mixture was stirred at 60°C overnight. The mixture was concentrated under vacuum and the residual crude sample was dissolved in  $\text{CH}_2\text{Cl}_2$ . The mixture was filtered off, then the filtrate was concentrated under reduced pressure. The crude residue was purified by flash chromatography ( $\text{CH}_2\text{Cl}_2$ /Acetone gradient 100:0 to 80:20, v/v) to provide **33a** (370 mg, 25%) as a white powder.

$R_f = 0.58$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 9:1);  $^1\text{H NMR}$  (200 MHz, CDCl<sub>3</sub>)  $\delta = 9.51$  (s, 1H, NH), 7.05 (s, 1H, H<sub>3</sub>), 5.10 (br t, 1H, NH), 3.76 (t,  $J = 6.5$  Hz, 2H, H<sub>6</sub>), 3.14 (dd,  $J = 11.6, 5.6$  Hz, 2H, H<sub>8</sub>), 1.91 (s, 3H, H<sub>5</sub>), 1.93 – 1.68 (m, 2H, H<sub>7</sub>), 1.42 (s, 9H, 3CH<sub>3</sub>.Boc);  $^{13}\text{C NMR}$  (50 MHz, CDCl<sub>3</sub>)  $\delta = 164.6$  (C<sub>1</sub>), 156.2 (CO<sub>Boc</sub>), 151.5 (C<sub>4</sub>), 140.6 (C<sub>3</sub>), 111.2 (C<sub>2</sub>), 79.6 (2C<sub>IV.Boc</sub>), 45.8 (C<sub>6</sub>), 37.1 (C<sub>8</sub>), 29.6 (3CH<sub>3</sub>.Boc), 28.5 (C<sub>7</sub>), 12.5 (C<sub>5</sub>); **MS (ESI)**:  $m/z$  589.1 [2M+Na]<sup>+</sup> (theoretical  $m/z$  589.3).

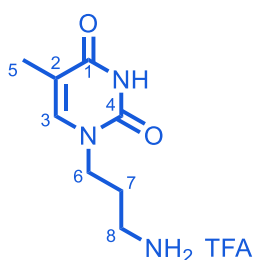
### ***N*<sup>3</sup>-(3-(*N*'-tert-butoxycarbonyl)aminopropyl)-thymine (33b)**



Compound **33b** was isolated together with compound **33a** during the same purification step. Compound **33b** (160 mg, 11%) was obtained as a white powder.

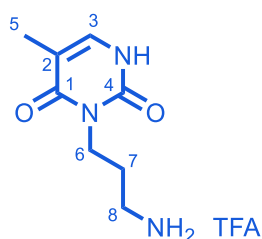
$R_f = 0.58$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 9:1);  $^1\text{H NMR}$  (200 MHz, CDCl<sub>3</sub>)  $\delta = 10.18$  (s, 1H, NH), 7.07 (d,  $J = 5.2$  Hz, 1H, H<sub>3</sub>), 5.31 (br t, 1H, NH), 4.01 (t,  $J = 6.3$  Hz, 2H, H<sub>6</sub>), 3.09 (dd,  $J = 11.0, 5.5$  Hz, 2H, H<sub>8</sub>), 1.92 (s, 3H, H<sub>5</sub>), 1.99 – 1.32 (m, 2H, H<sub>7</sub>), 1.43 (s, 9H, 3CH<sub>3</sub>.Boc);  $^{13}\text{C NMR}$  (50 MHz, CDCl<sub>3</sub>)  $\delta = 164.6$  (C<sub>1</sub>), 156.9 (CO<sub>Boc</sub>), 153.3 (C<sub>4</sub>), 134.9 (C<sub>3</sub>), 110.1 (C<sub>2</sub>), 79.2 (2C<sub>IV.Boc</sub>), 38.2 (C<sub>6</sub>), 37.3 (C<sub>8</sub>), 28.6 (3CH<sub>3</sub>.Boc), 28.1 (C<sub>7</sub>), 13.1 (C<sub>5</sub>); **MS (ESI)**:  $m/z$  283.9 [M+H]<sup>+</sup> (theoretical  $m/z$  284.2).

### ***N*<sup>1</sup>-Thymine-NH<sub>2</sub> = *N*<sup>1</sup>-(3-aminopropyl)thymine (34)**



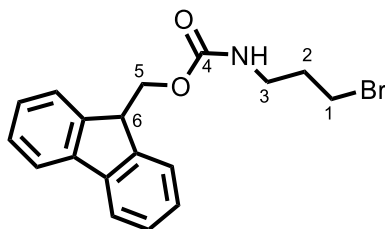
To a solution of protected compound **33a** (74.0 mg, 0.26 mmol) in CH<sub>2</sub>Cl<sub>2</sub> was added 50 equiv. of TFA. The reaction mixture was stirred at room temperature for 4h. The solvent and the remaining TFA were then removed under reduced pressure by co-evaporating twice with toluene, then the crude mixture was triturated in Et<sub>2</sub>O. After filtration, compound **34** (78.0 mg, 100%) as a white solid was obtained.

$R_f = 0.12$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 8:2);  $^1\text{H NMR}$  (200 MHz, CD<sub>3</sub>OD)  $\delta = 7.46$  (s, 1H, H<sub>3</sub>), 3.84 (t,  $J = 6.6$  Hz, 2H, H<sub>6</sub>), 2.99 (t,  $J = 7.3$  Hz, 2H, H<sub>8</sub>), 2.19 – 1.83 (m, 2H, H<sub>7</sub>), 1.87 (s, 3H, H<sub>5</sub>);  $^{13}\text{C NMR}$  (50 MHz, CD<sub>3</sub>OD)  $\delta = 166.8$  (C<sub>1</sub>), 153.4 (C<sub>4</sub>), 142.7 (C<sub>3</sub>), 111.8 (C<sub>2</sub>), 46.1 (C<sub>6</sub>), 37.8 (C<sub>8</sub>), 28.1 (C<sub>7</sub>), 12.2 (C<sub>5</sub>); **MS (ESI)**:  $m/z$  184.1 [M+H]<sup>+</sup> (theoretical  $m/z$  184.1).

**$N^1$ -Thymine- $NH_2 = N^3$ -(3-aminopropyl)thymine (35) :**

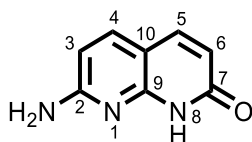
To a solution of protected compound **33b** (172 mg, 0.61 mmol) in  $CH_2Cl_2$  was added 50 equiv. of TFA. The reaction mixture was stirred at room temperature for 4h. The solvent and the remaining TFA were then removed under reduced pressure by co-evaporating twice with toluene, then the crude mixture was triturated in  $Et_2O$ . After filtration, compound **35** (170 mg, 94%) as a white solid was obtained.

$R_f = 0.33$  ( $CH_2Cl_2$ -MeOH 8:2);  $^1H$  NMR (200 MHz,  $CD_3OD$ )  $\delta = 7.28$  (s, 1H,  $H_3$ ), 4.03 (t,  $J = 6.3$  Hz, 2H,  $H_6$ ), 2.96 (t,  $J = 6.9$  Hz, 2H,  $H_8$ ), 2.13 – 1.86 (m, 2H,  $H_7$ ), 1.87 (s, 3H,  $H_5$ );  $^{13}C$  NMR (50 MHz,  $CD_3OD$ )  $\delta = 166.6$  ( $C_1$ ), 153.6 ( $C_4$ ), 137.8 ( $C_3$ ), 109.9 ( $C_2$ ), 38.3 ( $C_6 + C_8$ ), 26.9 ( $C_7$ ), 12.8 ( $C_5$ ); **MS (ESI):**  $m/z$  184.1 [ $M+H$ ] $^+$  (theoretical  $m/z$  184.1).

**5) Synthesis of Boc-Naphthyridine( $N$ -alkyl)- $NH_2$  (42)** **$N$ -Fmoc-3-bromopropylamine (39):**

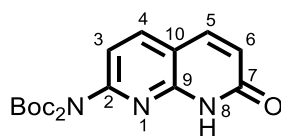
To a solution of 3-bromopropylamine hydrobromide (3.00 g, 13.7 mmol, 1 eq.) in water/dioxane (40 mL/70 mL) were added  $NaHCO_3$  (2.30 g, 27.4 mmol, 2 eq.) and 9-fluorenylmethyl  $N$ -succinimidyl carbonate (5.08 g, 15.1 mmol, 1.1 eq.). The reaction mixture was stirred overnight at room temperature, then concentrated under reduced pressure. The aqueous layer was extracted with  $EtOAc$ . The combined organic layers were washed with brine, dried over  $Na_2SO_4$ , and concentrated under reduced pressure. The crude residue was purified by flash chromatography (cHex/ $EtOAc$  gradient 95:5 to 85:15, v/v) to afford compound **39** (4.12 g, 84 %) as a white solid.

$R_f = 0.43$  (cHex- $EtOAc$  8:2);  $^1H$  NMR (400 MHz,  $CDCl_3$ )  $\delta = 7.77$  (d,  $J = 7.5$  Hz, 2H,  $2CH_{Fmoc}$ ), 7.59 (d,  $J = 7.5$  Hz, 2H,  $2CH_{Fmoc}$ ), 7.41 (td,  $J = 7.5, 1.2$  Hz, 2H,  $2CH_{Fmoc}$ ), 7.33 (td,  $J = 7.5, 1.2$  Hz, 2H,  $2CH_{Fmoc}$ ), 4.91 (br s, NH), 4.44 (d,  $J = 6.7$  Hz, 2H,  $H_5$ ), 4.21 (t,  $J = 6.8$  Hz, 1H,  $H_6$ ), 3.41 (t,  $J = 6.3$  Hz, 2H,  $H_1$ ), 3.34 (dt,  $J = 6.3$  Hz, 2H,  $H_3$ ), 2.06 (p,  $J = 6.3$  Hz, 2H,  $H_2$ );  $^{13}C$  NMR (101 MHz,  $CDCl_3$ )  $\delta = 156.6$  ( $C_4$ ), 144.0, 141.5 ( $C_{IV.Fmoc}$ ), 127.8, 127.2, 125.1, 120.1 ( $8CH_{Fmoc}$ ), 66.7 ( $C_6$ ), 47.4 ( $C_5$ ), 39.5 ( $C_3$ ), 32.5 ( $C_2$ ), 30.8 ( $C_1$ ); **MS (ESI):**  $m/z$  359.73 [ $M+H$ ] $^+$  (theoretical  $m/z$  360.06).

**2-Amino-7-hydroxy-[1,8]-naphthyridine (36):**

Malic acid (1.35 g, 10.1 mmol, 1.1 eq.) and 2,6-diaminopyridine (1.00 g, 9.16 mmol, 1 eq.) were put together in a round flask and at 0°C and concentrated sulfuric acid (5 mL) was added dropwise. The reaction mixture was stirred at 110 °C for 3h. After cooling, the reaction solution was neutralized with NH<sub>4</sub>OH at 0°C until pH 8. The mixture was filtered to afford compound **36** (1.48 g, quant.) as a brown solid.

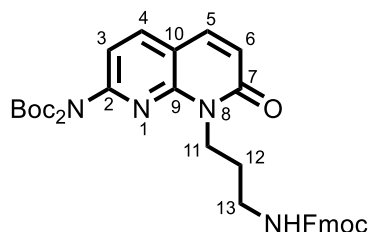
<sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ = 11.94 (s, 1H, NH), 7.73 – 7.66 (m, 2H, H<sub>4</sub> + H<sub>5</sub>), 6.99 (s, 2H, NH<sub>2</sub>), 6.35 (d, *J* = 8.5 Hz, 1H, H<sub>6</sub>), 6.12 (d, *J* = 9.2 Hz, 1H, H<sub>3</sub>); <sup>13</sup>C NMR (50 MHz, DMSO-*d*<sub>6</sub>) δ = 163.7, 160.6 (C<sub>2</sub> + C<sub>7</sub>), 150.4 (C<sub>9</sub>), 139.7, 137.3 (C<sub>4</sub> + C<sub>5</sub>), 114.8 (C<sub>10</sub>), 105.2, 104.9 (C<sub>3</sub> + C<sub>6</sub>); MS (ESI): *m/z* 162.20 [M+H]<sup>+</sup> (theoretical *m/z* 162.07).

**N<sup>2</sup>,N<sup>2</sup>-Bis(*tert*-butyl)-(7-oxo-7,8-dihydro[1,8]naphthyridin-2-yl)carbamate (38):**

To a suspension of compound **36** (3.00 g, 18.6 mmol, 1 eq.) in THF (120 mL) at 0°C, DMAP (1.00 g, 8.20 mmol, 0.4 eq.) and Boc<sub>2</sub>O (15.02 g, 68.8 mmol, 3.7 eq.) were added. The reaction mixture was stirred overnight at room temperature, then concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and then concentrated under reduced pressure. The crude residue was purified by flash chromatography (cHex/EtOAc gradient 90:10 to 40:60, v/v) to afford compound **38** (5.64 g, 84%) as a brown solid.

R<sub>f</sub> = 0.29 (cHex-EtOAc 4:6); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ = 9.78 (s, NH), 7.87 (d, *J* = 8.3 Hz, 1H, H<sub>4</sub>), 7.68 (d, *J* = 9.5 Hz, 1H, H<sub>5</sub>), 7.30 (d, *J* = 8.3 Hz, 1H, H<sub>3</sub>), 6.65 (d, *J* = 9.5 Hz, 1H, H<sub>6</sub>), 1.49 (s, 18H, 6CH<sub>3</sub>.Boc); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ = 163.4 (C<sub>7</sub>), 152.7 (C<sub>2</sub>), 150.9 (2C=O<sub>Boc</sub>), 148.3 (C<sub>2</sub>), 138.6 (C<sub>5</sub>), 137.8 (C<sub>4</sub>), 122.7 (C<sub>6</sub>), 114.8 (C<sub>3</sub>), 112.6 (C<sub>10</sub>), 84.0 (2C<sub>IV</sub>.Boc), 28.0 (6CH<sub>3</sub>.Boc); MS (ESI): *m/z* 361.93 [M+H]<sup>+</sup> (theoretical *m/z* 362.12).

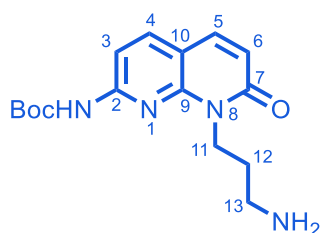
***N*<sup>2</sup>-*N*<sup>2</sup>-(*tert*-butyl)-(8-(*N*'-Fmoc-3-aminopropyl)-7-oxo-7,8-dihydro[1,8]naphthyridin-2-yl)carbamate (41):**



To a solution of compound **38** (200 mg, 0.553 mmol, 1 eq.) in acetone (8 ml), K<sub>2</sub>CO<sub>3</sub> (305 mg, 2.21 mmol, 4 eq.) was added and the reaction was stirred for 30 min before to adding compound **39** (399 mg, 1.11 mmol, 2 eq.). The reaction mixture was stirred overnight at reflux, then the solvent was concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and then concentrated under reduced pressure. The crude residue was purified by flash chromatography (cHex/EtOAc gradient 100:0 to 50:50, v/v) to afford compound **41** (254 mg, 72%) as a pale-yellow solid.

*R*<sub>f</sub> = 0.34 (cHex-EtOAc 5:5); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ = 7.89 (d, *J* = 8.3 Hz, 1H, H<sub>4</sub>), 7.75 (d, *J* = 7.5 Hz, 2H, 2CH<sub>Fmoc</sub>), 7.69 – 7.54 (m, 3H, 2CH<sub>Fmoc</sub> + H<sub>6</sub>), 7.42 – 7.24 (m, 5H, 4CH<sub>Fmoc</sub> + H<sub>3</sub>), 6.74 (d, *J* = 9.4 Hz, 1H, H<sub>7</sub>), 6.13 (t, *J* = 6.3 Hz, NH), 4.54 (t, *J* = 6.0 Hz, 2H, H<sub>11</sub>), 4.34 (d, *J* = 7.4 Hz, 2H, CH<sub>2.Fmoc</sub>), 4.25 (t, *J* = 7.3 Hz, 1H, CH<sub>Fmoc</sub>), 3.16 – 3.07 (m, 2H, H<sub>13</sub>), 2.04 – 1.92 (m, 2H), 1.52 (s, 18H, 6CH<sub>3.Boc</sub>); **MS (ESI):** *m/z* 640.67 [M+H]<sup>+</sup> (theoretical *m/z* 641.30).

**Boc-Naphthyridine(*N*-alkyl)-NH<sub>2</sub>: *N*<sup>2</sup>-(*tert*-butyl)-(8-(3-aminopropyl)-7-oxo-7,8-dihydro[1,8]naphthyridin-2-yl)carbamate (42):**



To a solution of compound **41** (393 mg, 0.610 mmol, 1 eq.) in DMF (1.6 ml) was added piperidine (400 μL, 4.05 mmol, 6.6 eq.). The reaction mixture was stirred at room temperature for 3h, then the solvent was concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub>. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and then concentrated under reduced pressure. The crude residue was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH [10% NH<sub>4</sub>OH] gradient 100:0 to 92:8, v/v) to afford compound **42** (140 mg, 72%) as a pale-yellow solid.

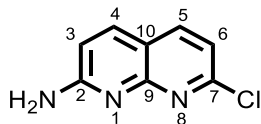
*R*<sub>f</sub> = 0.40 (CH<sub>2</sub>Cl<sub>2</sub>/MeOH [10% NH<sub>4</sub>OH] 90/10); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) δ = 7.85 (d, *J* = 8.5 Hz, 1H, H<sub>4</sub>), 7.81 (d, *J* = 8.5 Hz, 1H, H<sub>3</sub>), 7.56 (d, *J* = 9.4 Hz, 1H, H<sub>5</sub>), 7.40 (s, NH), 6.60 (d, *J* = 9.4 Hz, 1H, H<sub>6</sub>), 4.48 (t, *J*



= 6.9 Hz, 2H, H<sub>11</sub>), 2.73 (t, *J* = 6.6 Hz, 2H, H<sub>13</sub>), 1.91 (p, *J* = 6.8 Hz, 2H, H<sub>12</sub>), 1.55 (s, 9H, 3CH<sub>3</sub>.Boc); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>) δ = 163.4 (C<sub>7</sub>), 151.9, 151.91 (C<sub>2</sub> + CO<sub>Boc</sub>), 148.8 (C<sub>9</sub>), 138.8 (C<sub>4</sub>), 137.0 (C<sub>5</sub>), 120.2 (C<sub>6</sub>), 111.6 (C<sub>10</sub>), 106.7 (C<sub>3</sub>), 81.8 (C<sub>IV.Boc</sub>), 39.3 (C<sub>13</sub>), 38.5 (C<sub>11</sub>), 31.7 (C<sub>12</sub>), 28.3 (3CH<sub>3</sub>.Boc); **MS (ESI):** *m/z* 319.00 [M+H]<sup>+</sup> (theoretical *m/z* 319.18).

## 6) Synthesis of Boc-Naphthyridine(O-alkyl)-NH<sub>2</sub> (37)

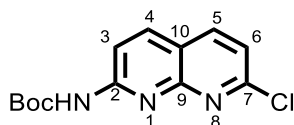
### 2-Amino-7-chloro-[1,8]-naphthyridine (44):



2-Amino-7-hydroxy-1,8-naphthyridine (**36**) (1.00 g, 0.621 mmol, 1 eq.) was suspended in anhydrous toluene (6 mL) and 6 mL of POCl<sub>3</sub> were added. The reaction mixture was stirred at reflux for 3 days. After cooling, the reaction solution was neutralized with NH<sub>4</sub>OH at 0°C, and was filtered. The aqueous phase was extracted three times with CH<sub>2</sub>Cl<sub>2</sub>. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and then concentrated under reduced pressure to give the crude product **44** as a yellow powder (884 mg, 80%).

*R<sub>f</sub>* = 0.39 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 95:5); <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ = 8.08 (d, *J* = 8.2 Hz, 1H, H<sub>5</sub>), 7.96 (d, *J* = 8.8 Hz, 1H, H<sub>4</sub>), 7.18 (d, *J* = 8.2 Hz, 1H, H<sub>6</sub>), 7.04 (s, 2H, NH<sub>2</sub>), 6.82 (d, *J* = 8.8 Hz, 1H, H<sub>3</sub>); <sup>13</sup>C NMR (101 MHz, DMSO-*d*<sub>6</sub>) δ = 161.5 (C<sub>2</sub>), 156.5 (C<sub>7</sub>), 151.0 (C<sub>9</sub>), 139.6 (C<sub>5</sub>), 137.4 (C<sub>4</sub>), 117.2 (C<sub>6</sub>), 115.7 (C<sub>10</sub>), 113.5 (C<sub>3</sub>); **MS (ESI):** *m/z* 180.4 [M+H]<sup>+</sup> (theoretical *m/z* 180.03).

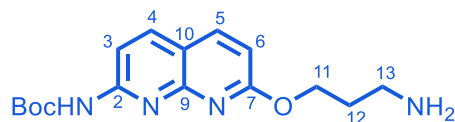
### N<sup>2</sup>-(*tert*-butyl)-(7-chloro[1,8]naphthyridin-2-yl)carbamate (45):



To a suspension of compound **44** (2.74 g, 16.92 mmol, 1 eq.) in anhydrous THF (140 mL) at 0°C, Boc<sub>2</sub>O (18.0 g, 84.58 mmol, 5 eq.) was added. The reaction mixture was stirred at reflux for 3 days then concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub>. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and then concentrated under reduced pressure. The crude residue was purified by flash chromatography (cHex/EtOAc gradient 100:0 to 50:50, v/v) to afford compound **45** (3.78 g, 80%) as a pale-yellow solid.

*R<sub>f</sub>* = 0.55 (cHex-EtOAc 6:4); <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ = 10.51 (s, 1H, NH), 8.43 (d, *J* = 9.0 Hz, 1H, H<sub>4</sub>), 8.41 (d, *J* = 8.4 Hz, 1H, H<sub>5</sub>), 8.19 (d, *J* = 9.0 Hz, 1H, H<sub>3</sub>), 7.54 (d, *J* = 8.4 Hz, 1H, H<sub>6</sub>), 1.50 (s, 9H, 3CH<sub>3</sub>.Boc); <sup>13</sup>C NMR (101 MHz, DMSO-*d*<sub>6</sub>) δ = 155.6, 154.3, 152.7, 152.5 (C<sub>2</sub> + C<sub>7</sub> + C<sub>9</sub> + CO<sub>Boc</sub>), 140.4 (C<sub>5</sub>), 139.3 (C<sub>4</sub>), 121.1 (C<sub>6</sub>), 118.4 (C<sub>10</sub>), 114.3 (C<sub>3</sub>), 80.4 (C<sub>IV.Boc</sub>), 27.9 (3CH<sub>3</sub>.Boc); **MS (ESI):** *m/z* 279.73 [M+H]<sup>+</sup> (theoretical *m/z* 280.08).

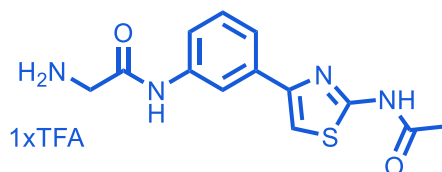
**Boc-Naphthyridine(O-alkyl)-NH<sub>2</sub>: N<sup>2</sup>-(*tert*-butyl)-(7-(3-aminopropoxy)-[1,8]naphthyridin-2-yl)carbamate (37):**



To a solution of 3-amino-1-propanol (1.64 mL, 21.5 mmol, 3 eq.) in THF (75 mL) was added NaH (1.00 g, 43.0 mmol, 6 eq.) and the reaction was stirred for 30 min before to adding compound **45** (2.00 g, 7.17 mmol, 1 eq.). The reaction mixture was stirred overnight at 60°C, then concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub>. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and then concentrated under reduced pressure. The crude residue was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH [10% NH<sub>4</sub>OH] gradient 100:0 to 90:10, v/v) to afford compound **37** (1.67 g, 73%) as a pale-yellow solid.

$R_f$  = 0.18 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH [10% NH<sub>4</sub>OH] 93:7); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  = 8.10 (d,  $J$  = 8.7 Hz, 1H, H<sub>3</sub>), 8.01 (d,  $J$  = 8.7 Hz, 1H, H<sub>4</sub>), 7.90 (d,  $J$  = 8.7 Hz, 1H, H<sub>5</sub>), 7.58 (s, 1H, NH), 6.83 (d,  $J$  = 8.7 Hz, 1H, H<sub>6</sub>), 4.59 (t,  $J$  = 6.4 Hz, 2H, H<sub>11</sub>), 2.90 (t,  $J$  = 6.9 Hz, 2H, H<sub>13</sub>), 2.04 – 1.89 (m, 2H, H<sub>12</sub>), 1.53 (s, 9H, 3CH<sub>3</sub>.Boc); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta$  = 165.4 (C<sub>7</sub>), 154.5 (C<sub>9</sub>), 153.7 (C<sub>2</sub>), 152.4 (CO<sub>Boc</sub>), 138.8, 138.7 (C<sub>4</sub> + C<sub>5</sub>), 116.1 (C<sub>10</sub>), 112.3 (C<sub>6</sub>), 110.6 (C<sub>3</sub>), 81.6 (C<sub>IV</sub>.Boc), 64.6 (C<sub>11</sub>), 39.4 (C<sub>13</sub>), 33.1 (C<sub>12</sub>), 28.4 (3CH<sub>3</sub>.Boc); **MS (ESI):**  $m/z$  = 319.00 [M+H]<sup>+</sup> (theoretical  $m/z$  319.10).

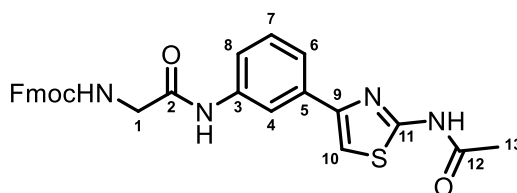
**7) Synthesis of S-Gly-NH<sub>2</sub> (TFA salt) (49a):**



Compound S-Gly-NH<sub>2</sub> (TFA salt) **49a** was obtained according to previously published procedure<sup>113</sup> in 38% yield using compound commercially available 2-bromo-3'-nitroacetophenone as starting material.

**8) Synthesis of S-Gly-NH<sub>2</sub> (49b):**

**N $\alpha$ -Fluorenylmethoxycarbonyl-[N-(2-N-acetylamino-4-(3-aminophenyl)-thiazole)]-tert-butoxyglycinamide (48b):**

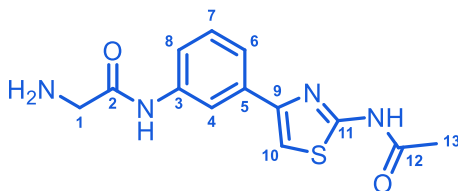


To a solution of **Fmoc-Gly-OH** (765 mg, 2.58 mmol, 1eq.) in CH<sub>2</sub>Cl<sub>2</sub>/DMF (4 mL/0.5 mL) was added 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (690 mg, 3.60 mmol, 1.4 eq.) and *N*-

hydroxysuccinimide (414 mg, 3.60 mmol, 1.4 eq.) and the mixture was stirred 30 min at room temperature. The compound **47**<sup>113</sup> (600 mg, 2.58 mmol, 1 eq.) was added and the reaction mixture was stirred overnight at room temperature, then the solvent was concentrated under reduced pressure. After solvent removal, the crude was precipitated in CH<sub>2</sub>Cl<sub>2</sub> to afford compound **48b** (926 mg, 70%) as a white solid.

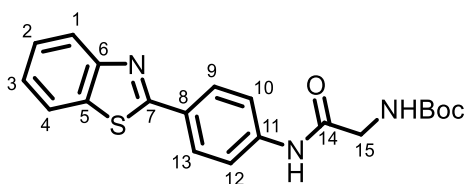
<sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ = 10.04 (s, NH), 8.28 (s, 1H, H<sub>4</sub>), 7.90, 7.73 (2d, *J* = 7.4 Hz, 4H, 4CH<sub>Fmoc</sub>), 7.66 – 7.54 (m, 1H, H<sub>6</sub>), 7.51 (s, 1H, H<sub>10</sub>), 7.46 – 7.30 (m, 6H, 4CH<sub>Fmoc</sub> + H<sub>7-8</sub>), 4.32 (d, *J* = 6.8 Hz, 2H, CH<sub>2Fmoc</sub>), 4.28 – 4.18 (m, 1H, CH<sub>Fmoc</sub>), 3.82 (d, *J* = 6.1 Hz, 2H, H<sub>1</sub>), 2.15 (s, 3H, H<sub>13</sub>); <sup>13</sup>C NMR (101 MHz, DMSO-*d*<sub>6</sub>) δ = 168.8, 168.6 (C<sub>2</sub> + C<sub>12</sub>), 158.0 (C<sub>11</sub>), 148.7 (C<sub>9</sub>), 142.6 (2C<sub>IV.Fmoc</sub>), 139.5 (C<sub>IV.Fmoc</sub>), 139.2 (C<sub>3</sub>), 137.5 (C<sub>IV.Fmoc</sub>), 134.9 (C<sub>5</sub>), 129.2 (C<sub>7</sub>), 129.0, 127.4, 121.4 (6CH<sub>Fmoc</sub>), 120.7 (C<sub>6</sub>), 120.1 (2CH<sub>Fmoc</sub>), 118.7 (C<sub>8</sub>), 117.0 (C<sub>4</sub>), 108.1 (C<sub>10</sub>), 66.1 (CH<sub>2.Fmoc</sub>), 45.3 (CH<sub>Fmoc</sub>), 44.3 (C<sub>1</sub>), 22.5 (C<sub>13</sub>); **MS (ESI):** *m/z* 513.5 [M+H]<sup>+</sup> (theoretical *m/z* 513.16).

**S-Gly-NH<sub>2</sub> = *N*-(2-*N*-acetylamino-4-(3-aminophenyl)-thiazole)]-*tert*-butoxyglycinamide (49b) :**



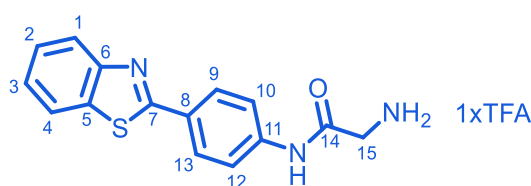
To a solution of compound **48b** (800 mg, 1.56 mmol, 1 eq.) in DMF (4 ml) was added piperidine (1.00 mL, 10.1 mmol, 6.5 eq.). The reaction mixture was stirred overnight at room temperature, then the solvent was concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub>. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and then concentrated under reduced pressure. The crude residue was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH [10% NH<sub>4</sub>OH] gradient 100:0 to 93:7, v/v) to afford compound **49b** (389 mg, 86%) as a white solid.

*R<sub>f</sub>* = 0.44 (CH<sub>2</sub>Cl<sub>2</sub>/MeOH [10% NH<sub>4</sub>OH] 93/7); <sup>1</sup>H NMR (400 MHz, DMSO-*d*<sub>6</sub>) δ = 8.27 (s, 1H, H<sub>4</sub>), 7.60 – 7.55 (m, 1H, H<sub>6</sub>), 7.52 (s, 1H, H<sub>10</sub>), 7.53 – 7.47 (m, 1H, H<sub>8</sub>), 7.35 (t, *J* = 7.8 Hz, 1H, H<sub>7</sub>), H<sub>1</sub> (signal under the D<sub>2</sub>O peak), 2.16 (s, 1H, H<sub>13</sub>); <sup>13</sup>C NMR (101 MHz, DMSO-*d*<sub>6</sub>) δ = 170.8 (C<sub>2</sub>), 168.7 (C<sub>12</sub>), 158.0 (C<sub>11</sub>), 148.6 (C<sub>9</sub>), 139.1 (C<sub>3</sub>), 134.9 (C<sub>5</sub>), 129.1 (C<sub>7</sub>), 120.7 (C<sub>6</sub>), 118.5 (C<sub>8</sub>), 116.6 (C<sub>4</sub>), 108.0 (C<sub>10</sub>), 44.7 (C<sub>1</sub>), 22.5 (C<sub>13</sub>); **MS (ESI):** *m/z* 291.4 [M+H]<sup>+</sup> (theoretical *m/z* 291.09).

**9) Synthesis of 2-Amino-N-[4-(benzothiazol-2-yl)phenyl]acetamide (TFA salt) (55)****2-(N-tert-butoxycarbonyl)amino-N'-[4-(benzothiazol-2-yl)phenyl]acetamide (54):**

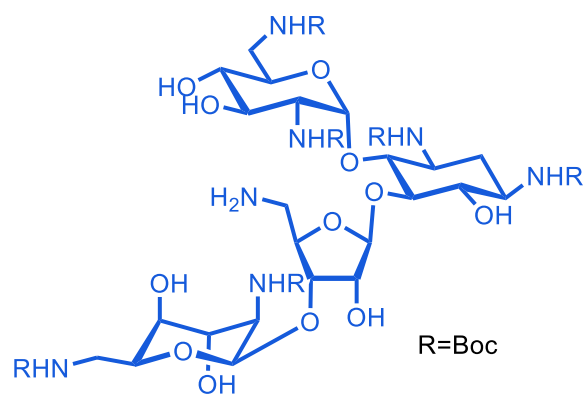
To a solution of Boc-Gly-OH (354mg, 2.02 mmol, 1 eq.) in  $\text{CH}_2\text{Cl}_2$  (5.5 mL) was added triethylamine (560  $\mu\text{L}$ , 4.04 mmol, 2 eq.) and HBTU (918 mg, 2.42 mmol, 1.2 eq.) and the mixture was stirred 30 min at room temperature. The compound 2-(4-aminophenyl)benzothiazole (548 mg, 2.42 mmol, 1.2 eq.) was added and the reaction mixture was stirred overnight at room temperature, then the solvent was concentrated under reduced pressure. After solvent removal, the crude was precipitated in  $\text{CH}_2\text{Cl}_2$  and washed with EtOH ( $3 \times 2$  mL) to afford compound **54** (513 mg, 66%) as a white solid.

$R_f = 0.40$  (cHex-EtOAc 5:5);  $^1\text{H NMR}$  (200 MHz,  $\text{DMSO}-d_6$ )  $\delta = 10.30$  (s, 1H, NH), 8.20 – 7.92 (m, 4H,  $H_1 + H_4 + H_9 + H_{13}$ ), 7.88 – 7.72 (m, 2H,  $H_{10} + H_{12}$ ), 7.60 – 7.38 (m, 2H,  $H_2 + H_3$ ), 7.14 (t,  $J = 4.6$  Hz, 1H, NH), 3.77 (d,  $J = 4.6$  Hz, 2H,  $H_{15}$ ), 1.40 (s, 9H,  $3\text{CH}_3.\text{Boc}$ );  $^{13}\text{C NMR}$  (50 MHz,  $\text{DMSO}-d_6$ )  $\delta = 168.8$  ( $C_7$ ), 167.0 ( $C_{14}$ ), 156.0 ( $\text{CO}_{\text{Boc}}$ ), 153.7 ( $C_6$ ), 141.8 ( $C_{11}$ ), 134.3 ( $C_5$ ), 128.1 ( $C_9 + C_{13}$ ), 127.5 ( $C_8$ ), 126.6 ( $C_2$ ), 125.2 ( $C_3$ ), 122.6 ( $C_4$ ), 122.2 ( $C_1$ ), 119.3 ( $C_{10} + C_{12}$ ), 78.1 ( $C_{\text{IV.Boc}}$ ), 43.9 ( $C_{15}$ ), 28.2 ( $3\text{CH}_3.\text{Boc}$ ); **MS (ESI):**  $m/z = 384.1$  [ $\text{M}+\text{H}$ ] $^+$  (theoretical  $m/z$  384.1).

**2-Amino-N-[4-(benzothiazol-2-yl)phenyl]acetamide (55)**

To a solution of protected compound **54** (319 mg, 0.83 mmol) in  $\text{CH}_2\text{Cl}_2$  was added 50 equiv. of TFA. The reaction mixture was stirred at room temperature for 4h. The solvent and the remaining TFA were then removed under reduced pressure by co-evaporating twice with toluene, then the crude mixture was triturated in  $\text{Et}_2\text{O}$ . After filtration, compound **55** (241 mg, 73%) as a white solid was obtained.

$R_f = 0.47$  ( $\text{CH}_2\text{Cl}_2$ -MeOH 80:20);  $^1\text{H NMR}$  (200 MHz,  $\text{DMSO}-d_6$ )  $\delta = 10.91$  (s, 1H, NH), 8.27 (s br, 2H,  $H_{17}$ ), 8.17 – 7.98 (m, 4H,  $H_1 + H_4 + H_9 + H_{11}$ ), 7.81 (d,  $J = 8.7$  Hz, 2H,  $H_{10} + H_{12}$ ), 7.59 – 7.38 (m, 2H,  $H_2 + H_3$ ), 3.87 (s, 2H,  $H_{15}$ );  $^{13}\text{C NMR}$  (50 MHz,  $\text{DMSO}-d_6$ )  $\delta = 166.8$  ( $C_7$ ), 165.4 ( $C_{14}$ ), 153.6 ( $C_6$ ), 141.0 ( $C_{11}$ ), 134.4 ( $C_5$ ), 128.3 ( $C_9 + C_{13}$ ), 128.2 ( $C_8$ ), 126.7 ( $C_2$ ), 125.4 ( $C_3$ ), 122.7 ( $C_4$ ), 122.3 ( $C_1$ ), 119.5 ( $C_{10} + C_{12}$ ), 41.2 ( $C_{15}$ ); **MS (ESI):**  $m/z = 284.1$  [ $\text{M}+\text{H}$ ] $^+$  (theoretical  $m/z$  284.1).

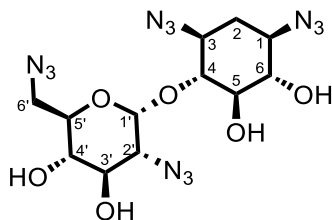
**10) Synthesis of Boc-Neomycin-NH<sub>2</sub> (57):**

Compound Boc-Neo-NH<sub>2</sub> **57** was obtained according to previously published procedure<sup>186</sup> in 14% yield using compound commercially available neomycin trisulfate salt hydrate as starting material.

## Experimental part – Chapter II - Synthesis of aldehyde fragments

### 1) Synthesis of Boc-DOS<sub>(Acetal)</sub>-CHO (58):

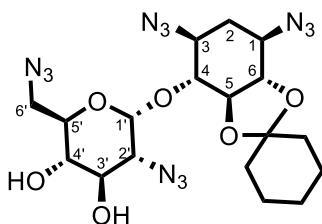
#### 1,3,2',6'-Tetraazidoneamine (60b):



To a solution of NaN<sub>3</sub> (1.67 g, 25.6 mmol, 18.2 eq.) in H<sub>2</sub>O/CH<sub>2</sub>Cl<sub>2</sub> (15 mL, 1:1 v/v) at 0 °C was added Tf<sub>2</sub>O (2.40 mL, 14.2 mmol, 10 eq.). The reaction mixture was stirred at room temperature for 3h. After, a saturated NaHCO<sub>3</sub> solution (10 mL) was added. Then, the aqueous layer was separated, and extracted 3 times with CH<sub>2</sub>Cl<sub>2</sub> (7 mL). The organic layers were combined to afford 28 mL of TfN<sub>3</sub> solution. Then, to a solution of neamine 4.HCl<sup>189</sup> (660 mg, 1.41 mmol, 1eq) in H<sub>2</sub>O (15 mL) was added the TfN<sub>3</sub> solution, MeOH (40 mL), and a solution of CuSO<sub>4</sub> (25 mg) in a mixture of MeOH/H<sub>2</sub>O/Et<sub>3</sub>N (5 mL, 3:3:4 v/v/v). The reaction mixture was stirred overnight at room temperature, then solid NaHCO<sub>3</sub> (1 g) was added portion wise, and the organic solvents were evaporated. The remaining aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography (cHex/EtOAc gradient 80:20 to 40:60, v/v) to afford compound **60b** (476 mg, 79%) as a white solid.

$R_f$  = 0.39 (cHex-EtOAc 4:6); <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 5.65 (d,  $J$  = 3.7 Hz, 1H, H<sub>1'</sub>), 4.18 (ddd,  $J$  = 9.8, 5.4, 2.4 Hz, 1H, H<sub>5'</sub>), 3.87 (dd,  $J$  = 10.5, 8.9 Hz, 1H, H<sub>3'</sub>), 3.56 – 3.22 (m, 7H, 5CH + H<sub>6'</sub>), 3.12 (dd,  $J$  = 10.5, 3.7 Hz, 1H, H<sub>2'</sub>), 2.25 (dt,  $J$  = 13.0, 4.1 Hz, 1H, H<sub>2,eq</sub>), 1.49 – 1.35 (m, 1H, H<sub>2,ax</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 99.7 (C<sub>1'</sub>), 81.1 (C<sub>4</sub>), 77.9 (CH), 77.9 (CH), 73.2 (C<sub>5'</sub>), 72.7 (CH), 72.2 (C<sub>3'</sub>), 64.7 (C<sub>2'</sub>), 61.8, 60.9 (C<sub>1</sub> + C<sub>3</sub>), 52.6 (C<sub>6'</sub>), 33.2 (C<sub>2</sub>);

#### 1,3,2',6'-Tetraazido-5,6-O-cyclohexylideneneamine (61b):



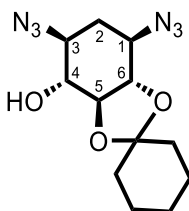
To a solution of compound **60b** (2.10 g, 4.92 mmol, 1 eq.) in dry DMF (8 mL) were added *p*-TsOH monohydrate (45.0 mg, 0.0481 mmol, 0.04 eq.) and 1,1-dimethoxycyclohexane (2.99 mL, 19.7 mmol, 4 eq.). The reaction mixture was heated at 50°C under 55 mbar in a rotary evaporator and after 1h, *p*-TsOH monohydrate (0.04 eq.) was added again. This solution was heated for 1h30 in a rotary

evaporator (50 °C, 50 mbar) and triethylamine (1 mL) was added. Then, water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography (cHex/EtOAc gradient 100:0 to 40:60, v/v) to afford the desired compound **61b** (1.67 g, 67%) separated from the 1,3,2',6'-tetraazido-3',4',5,6-di-*O*-cyclohexylideneamine (570 mg, 20%) as white solids.

Finally, the 1,3,2',6'-tetraazido-3',4',5,6-di-*O*-cyclohexylideneamine (570 mg, 0.972 mmol, 1 eq.) [*R<sub>f</sub>* = 0.80 (cHex-EtOAc 5:5)] was fully converted into the desired product **61b** by the following procedure: dry dimethylformamide (8 mL), *p*TsOH monohydrate (35 mg, 0.184 mmol, 0.18 eq.) and dry methanol (150 μL) were added to 1,3,2',6'-tetraazido-3',4',5,6-di-*O*-cyclohexylideneamine (570 mg, 0.972 mmol, 1 eq.). This solution was heated for 30 min. in a rotary evaporator (50 °C, 50 mbar) and every 30 min. *p*-TsOH monohydrate (0.18 eq.) and dry methanol (150 μL) were added until TLC showed completion of the reaction. Then, the reaction mixture was quenched with triethylamine (1 mL). Water was added and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography (cHex/EtOAc gradient 100:0 to 40:60, v/v) to afford the desired compound **61b** (274 mg, 56%).

*R<sub>f</sub>* = 0.44 (cHex-EtOAc 6:4); <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ = 5.51 (d, *J* = 3.7 Hz, 1H, H<sub>1'</sub>), 4.03 (ddd, *J* = 10.0, 5.9, 2.4 Hz, 1H, H<sub>5'</sub>), 3.87 (m, H<sub>3'</sub> + H<sub>4'</sub>), 3.80 (ddd, *J* = 11.4, 10.0, 4.7 Hz, 1H, H<sub>1'</sub>), 3.69 – 3.42 (m, 4H, 3CH + H<sub>6'</sub>), 3.37 (dd, *J* = 10.0, 8.7 Hz, 1H, H<sub>4'</sub>), 3.16 (dd, *J* = 10.5, 3.7 Hz, 1H, H<sub>2'</sub>), 2.32 (dt, *J* = 13.3, 4.7 Hz, 1H, H<sub>2,eq</sub>), 1.78 – 1.58 (m, 8H, 4CH<sub>2,Acetal</sub>), 1.53 – 1.38 (m, 3H, CH<sub>2,Acetal</sub> and H<sub>2,ax</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD) δ = 114.3 (C<sub>IV,Acetal</sub>), 98.0 (C<sub>1'</sub>), 80.8, 80.8 (C<sub>5-6</sub>), 78.3 (C<sub>4</sub>), 73.4 (C<sub>5'</sub>), 72.7 (C<sub>4'</sub>), 71.9 (C<sub>3'</sub>), 64.3 (C<sub>2'</sub>), 62.4 (C<sub>3</sub>), 58.7 (C<sub>I</sub>), 52.6 (C<sub>6'</sub>), 37.3, 37.2 (2CH<sub>2,Acetal</sub>), 34.8 (C<sub>2</sub>), 26.1, 24.9, 24.8 (3CH<sub>2,Acetal</sub>);

### 1,3-Diazido-5,6-*O*-cyclohexylidene-2-deoxystreptamine (**62b**):

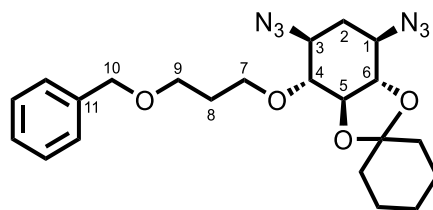


A solution of compound **61b** (1.94 g, 3.83 mmol, 1 eq.) in methanol (120 mL) was added NaIO<sub>4</sub> (6.15 g, 28.7 mmol, 7.5 eq.) at 0 °C. Then, the reaction mixture was stirred overnight at room temperature (TLC of the mixture showed an intermediate at *R<sub>f</sub>* = 0.75 (cHex-EtOAc 5:5)). The precipitate was filtered through a pad of celite and washed with MeOH. After solvent removal, the crude was diluted with EtOAc and neutralized with a saturated solution of NaHCO<sub>3</sub>. After layers were separated, the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and then concentrated under reduced pressure. The crude residue was then dissolved in

methanol (200 mL), *n*-butylamine (600  $\mu$ L, 11.5 mol, 3 eq.) was added dropwise and the reaction mixture was stirred overnight at room temperature. After solvent removal, water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and then concentrated under reduced pressure. The crude residue was purified by flash chromatography (cHex/EtOAc gradient 100:0 to 80:20, v/v) to afford compound **62b** (910 mg, 81%) as a yellow oil.

$R_f$  = 0.42 (cHex-EtOAc 7:3); <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 3.81 – 3.68 (m, 1H, H<sub>1</sub>), 3.68 – 3.58 (m, 1H, H<sub>4</sub>), 3.48 – 3.38 (m, 3H, H<sub>3</sub> + H<sub>5</sub> + H<sub>6</sub>), 2.21 (dt,  $J$  = 13.3, 4.9 Hz, 1H, H<sub>2,eq</sub>), 1.75 – 1.56 (m, 8H, 4CH<sub>2,Acetal</sub>), 1.51 – 1.40 (m, 2H, CH<sub>2,Acetal</sub>), 1.34 (dt,  $J$  = 13.3, 11.8 Hz, 1H, H<sub>2,ax</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 113.7 (C<sub>IV,Acetal</sub>), 80.8, 80.6 (C<sub>5-6</sub>), 75.3 (C<sub>4</sub>), 64.3 (C<sub>3</sub>), 59.0 (C<sub>1</sub>), 37.4, 37.3 (2CH<sub>2,Acetal</sub>), 34.8 (C<sub>2</sub>), 26.1, 24.8, 24.7 (3CH<sub>2,Acetal</sub>);

### 1,3-Diazido-5,6-*O*-cyclohexylidene-4-*O*-(3-(benzyloxy)propoxy)-2-deoxystreptamine (**69**):

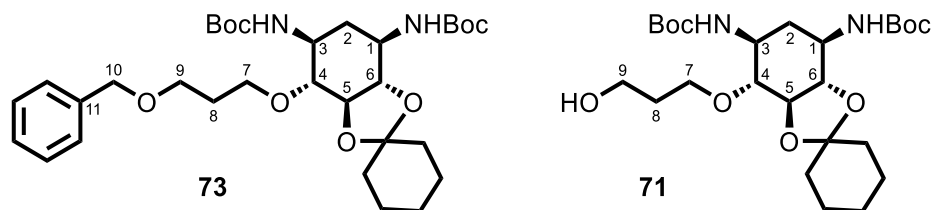


To a solution of compound **62b** (278 mg, 0.950 mmol, 1 eq.) in DMF (3 mL) was added NaH (91.1 mg, 3.78 mmol, 4 eq.) at 0°C. After stirring 1h, TBAI (418 mg, 1.14 mmol, 1.2 eq.) and benzyl 3-bromopropyl ether (0.5 mL, 2.83 mmol, 3 eq.) were added and the reaction was stirred overnight at room temperature. Water was added and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography (cHex/EtOAc gradient 100:0 to 75:25, v/v) to afford compound **69** (335 mg, 80%) as a white solid.

$R_f$  = 0.60 (cHex-EtOAc 8:2); <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 7.36 – 7.24 (m, 5H, 5CH<sub>Ar</sub>), 4.50 (s, 2H, H<sub>10</sub>), 3.99 (dt,  $J$  = 9.7, 6.1 Hz, 1H, H<sub>7</sub>), 3.78 – 3.38 (m, 8H, H<sub>1'</sub> + H<sub>3'-6'</sub> + H<sub>7</sub> + H<sub>9</sub>), 2.14 (dt,  $J$  = 13.5, 5.0 Hz, 1H, H<sub>2,eq</sub>), 1.95 – 1.84 (m, 2H, H<sub>8</sub>), 1.74 – 1.55 (m, 8H, 4CH<sub>2,Acetal</sub>), 1.46 – 1.38 (m, 2H, CH<sub>2,Acetal</sub>), 1.30 (dt,  $J$  = 13.5, 11.7 Hz, 1H, H<sub>2,ax</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 139.9 (C<sub>11</sub>), 129.4, 128.8, 128.6 (5CH<sub>Ar</sub>), 113.7 (C<sub>IV,Acetal</sub>), 83.2 (C<sub>4</sub>), 81.0, 80.8 (C<sub>5-6</sub>), 74.0 (C<sub>10</sub>), 69.2, 68.2 (C<sub>7</sub> + C<sub>9</sub>), 62.6 (C<sub>3</sub>), 58.8 (C<sub>1</sub>), 37.4, 37.2 (2CH<sub>2,Acetal</sub>), 34.9 (C<sub>2</sub>), 31.5 (C<sub>8</sub>), 26.1, 24.9, 24.9 (3CH<sub>2,Acetal</sub>); **MS (ESI):**  $m/z$  443.3 [M+H]<sup>+</sup> (theoretical  $m/z$  443.24).



**1,3-Bis-*N*-(*tert*-butoxycarbonyl)-5,6-*O*-cyclohexylidene-4-*O*-(3-hydroxypropoxy)-2-deoxystreptamine (71):**



To a solution of compound **69** (216 mg, 0.490 mmol, 1 eq.) in EtOH (3 mL) was added 10% Pd/C (20 mg). The system was degassed three times (vacuum/N<sub>2</sub>), then stirred under H<sub>2</sub> (1 atm) for 12h. The mixture was filtered through a pad of celite, and which was washed with EtOH. TLC of the mixture showed an intermediate at  $R_f = 0.11$  (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 9:1). After solvent removal, the crude was dissolved in dioxane (2 mL) and at 0°C, TEA (134 μL, 0.980 mmol, 2 eq.) and Boc<sub>2</sub>O (319 mg, 1.46 mmol, 3 eq.) were added. The reaction mixture was stirred overnight at room temperature and the reaction mixture was concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and then concentrated under reduced pressure to afford the intermediate **73**. [ $R_f = 0.32$  (cHex-EtOAc 7:3)]. The crude residue was diluted with EtOH (2 mL), and 10% Pd/C (15 mg) was added. The system was degassed three times (vacuum/N<sub>2</sub>), then stirred under H<sub>2</sub> (1 atm) for 3h. The mixture was filtered through a pad of celite which was washed with EtOH and CH<sub>2</sub>Cl<sub>2</sub>. After solvent removal, the crude was purified by flash chromatography (cHex/EtOAc gradient 80:20 to 40:60, v/v) to afford compound **71** (65%) as a white solid. [ $R_f = 0.47$  (cHex-EtOAc 4:6)].

**1,3-Bis-*N*-(*tert*-butoxycarbonyl)-5,6-*O*-cyclohexylidene-4-*O*-(3-(benzyloxy)propoxy)-2-deoxystreptamine (73):**

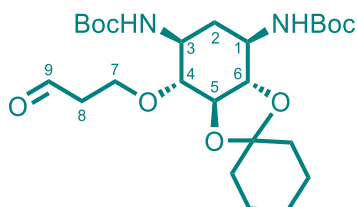
$R_f = 0.11$  (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 9:1); <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 7.38 - 7.22$  (m, 5H, 5CH<sub>Ar</sub>), 4.50 (s, 2H, H<sub>10</sub>), 3.91, 3.70 – 3.57 (dt and m,  $J = 9.7, 6.0$  Hz, 2H, H<sub>7</sub>), 3.70 – 3.57 (m, 3H, H<sub>1</sub> + H<sub>9</sub>), 3.54 – 3.44 (m, 1H, H<sub>3</sub>), 3.44 – 3.28 (m, 3H, H<sub>4-6</sub>), 2.08 (dt,  $J = 13.1, 4.7$  Hz, 1H, H<sub>2,eq</sub>), 1.84 (p,  $J = 6.0$  Hz, 2H, H<sub>8</sub>), 1.71 – 1.54 (m, 8H, 4CH<sub>2,Acetal</sub>), 1.52 – 1.24 (m, 21H, CH<sub>2,Acetal</sub> + H<sub>2,ax</sub> + 6CH<sub>3,Boc</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 158.0, 157.7$  (2CO<sub>Boc</sub>), 139.9 (C<sub>11</sub>), 129.3, 128.8, 128.6 (5CH<sub>Ar</sub>), 112.9 (C<sub>IV,Acetal</sub>), 82.3, 82.0, 80.2, 80.0, 79.8 (C<sub>4-6</sub> + 2C<sub>IV,Boc</sub>), 74.0 (C<sub>10</sub>), 69.1, 68.5 (C<sub>7</sub> + C<sub>9</sub>), 52.8 (C<sub>3</sub>), 49.8 (C<sub>1</sub>), 37.5 (2CH<sub>2,Acetal</sub>), 37.3 (C<sub>2</sub>), 31.5 (C<sub>8</sub>), 28.8, 28.7 (6CH<sub>3,Boc</sub>), 26.2, 24.9, 24.8 (3CH<sub>2,Acetal</sub>); **MS (ESI):**  $m/z$  613.3 [M+Na]<sup>+</sup> (theoretical  $m/z$  613.35).

**1,3-Bis-*N*-(*tert*-butoxycarbonyl)-5,6-*O*-cyclohexylidene-4-*O*-(3-hydroxypropoxy)-2-deoxystreptamine (71):**

$R_f = 0.47$  (cHex-EtOAc 4:6); <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 3.91, 3.73 - 3.58$  (dt and m,  $J = 9.7, 5.9$  Hz, 2H, H<sub>7</sub>), 3.73 – 3.58 (m, 3H, H<sub>1</sub> + H<sub>9</sub>), 3.56 – 3.44 (m, 1H, H<sub>3</sub>), 3.44 – 3.26 (m, 3H, H<sub>4-6</sub>), 2.09 (dt,  $J = 13.2,$

4.7 Hz, 1H, H<sub>2.eq</sub>), 1.83 – 1.71 (m, 2H, H<sub>8</sub>), 1.72 – 1.54 (m, 8H, 4CH<sub>2.Acetal</sub>), 1.54 – 1.25 (m, 21H, CH<sub>2.Acetal</sub> + 6CH<sub>3.Boc</sub> + H<sub>2.ax</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 158.0, 157.7 (2CO<sub>Boc</sub>), 112.8 (C<sub>IV.Acetal</sub>), 82.2, 81.9, 80.2, 80.0 (C<sub>4-6</sub> + 2C<sub>IV.Boc</sub>), 69.2 (C<sub>7</sub>), 60.3 (C<sub>9</sub>), 52.6 (C<sub>3</sub>), 49.8 (C<sub>1</sub>), 37.5 (2CH<sub>2.Acetal</sub>), 37.2 (C<sub>2</sub>), 34.0 (C<sub>8</sub>), 28.7, 28.7 (6CH<sub>3.Boc</sub>), 26.1, 24.8, 24.8 (3CH<sub>2.Acetal</sub>); **MS (ESI):**  $m/z$  523.3 [M+Na]<sup>+</sup> (theoretical  $m/z$  523.30).

**Boc-DOS<sub>(Acetal)</sub>-CHO = 1,3-Bis-*N*-(*tert*-butoxycarbonyl)-5,6-*O*-cyclohexylidene-4-*O*-(3-oxopropoxy)-2-deoxystreptamine (58):**

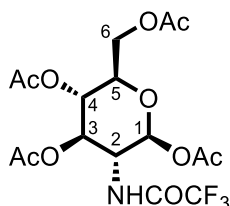


IBX (97.0 mg, 0.345 mmol, 1.1 eq.) was dissolved in DMSO (2 mL) and compound **71** (157 mg, 0.314 mmol, 1 eq.) in DMSO (1 mL) was added. The reaction mixture was stirred overnight at room temperature, then was filtered through a pad of Celite<sup>®</sup> which was washed with EtOAc. Water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and then concentrated under reduced pressure. The crude residue was purified by flash chromatography (cHex/EtOAc gradient 90:10 to 70:30, v/v) to afford compound **58** (117 mg, 75%) as a white solid.

$R_f$  = 0.30 (cHex-EtOAc 7:3); <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  = 9.76 (t,  $J$  = 1.7 Hz, 1H, H<sub>9</sub>), 4.71 (s, 2NH), 4.12, 3.89 (dt and ddd,  $J$  = 10.2, 5.6 Hz and  $J$  = 10.2, 7.7, 4.8 Hz, 2H, H<sub>7</sub>), 3.72 – 3.26 (m, 5H, H<sub>1</sub> + H<sub>3-6</sub>), 2.74 – 2.51 (m, 2H, H<sub>8</sub>), 2.49 – 2.39 (m, 1H, H<sub>2.eq</sub>), 1.67 – 1.21 (m, 29H, 5CH<sub>2.Acetal</sub> + 6CH<sub>3.Boc</sub> + H<sub>2.ax</sub>); <sup>13</sup>C NMR (101 MHz, CDCl<sub>3</sub>)  $\delta$  = 201.7 (C<sub>9</sub>), 155.5, 155.3 (2CO<sub>Boc</sub>), 112.4 (C<sub>IV.Acetal</sub>), 80.7, 79.8, 78.5 (C<sub>4-6</sub> + 2C<sub>IV.Boc</sub>), 64.9 (C<sub>7</sub>), 52.0, 49.2 (C<sub>1</sub> + C<sub>3</sub>), 44.2 (C<sub>8</sub>), 42.1 (C<sub>2</sub>), 36.6, 36.2 (2CH<sub>2.Acetal</sub>), 28.4, 28.4 (6CH<sub>3.Boc</sub>), 25.1, 23.9, 23.8 (3CH<sub>2.Acetal</sub>); **MS (ESI):** *do not respond*.

**2) Synthesis of Glucosamine<sub>(Acetyl)</sub>-CHO (74):**

**1,3,4,6-Tetra-*O*-acetyl-2-deoxy-2-trifluoroacetamido- $\beta$ -D-glucopyranose (77):**



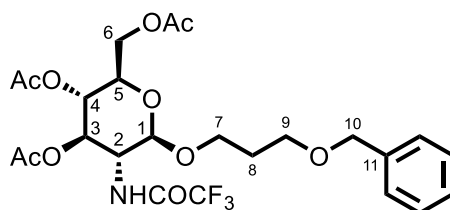
To a solution of D-Glucosamine hydrochloride (3.00 g, 13.9 mmol, 1 eq.) in MeOH (15 mL), triethylamine (1.94 mL, 13.9 mmol, 1 eq.) and ethyl trifluoroacetate (4.97 mL, 13.9 mmol, 1 eq.) were added. The reaction mixture was stirred overnight at room temperature, then concentrated under

reduced pressure. The crude was used as such without purification for the next step.  $R_f = 0.51$  ( $\text{CH}_2\text{Cl}_2$ -MeOH 8:2).

The crude was taken in anhydrous pyridine (40 mL) and acetic anhydride (10.5 mL, 111 mmol, 8 eq.) was added to the solution. The reaction mixture was stirred overnight at room temperature, then concentrated under reduced pressure. Saturated solution of  $\text{NaHCO}_3$  was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over  $\text{Na}_2\text{SO}_4$  and then concentrated under reduced pressure. The crude residue was purified by flash chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  gradient 100:0 to 98:2, v/v) to afford compound  $\beta$  **77** (1.46 g, 80%) as a white solid.

$R_{f,\alpha} = 0.42$  (cHex-EtOAc 5:5);  $R_{f,\beta} = 0.48$  (cHex-EtOAc 5:5);  $^1\text{H NMR}$  (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 5.85$  (d,  $J = 8.8$  Hz, 1H,  $\text{H}_1$ ), 5.35 (dd,  $J = 10.1, 9.3$  Hz, 1H,  $\text{H}_3$ ), 5.11 – 5.02 (t,  $J = 10.1$  Hz, 1H,  $\text{H}_4$ ), 4.30 (dd,  $J = 12.4, 4.6$  Hz, 1H,  $\text{H}_{6a}$ ), 4.12 (m, 2H,  $\text{H}_2 + \text{H}_{6b}$ ), 3.95 (ddd,  $J = 10.1, 4.6, 2.3$  Hz, 1H,  $\text{H}_5$ ), 2.08, 2.06, 2.02, 1.98 (4s, 12H,  $4\text{CH}_{3,\text{Ac}}$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 172.3, 171.6, 171.2, 170.4$  ( $4\text{CO}_{\text{Ac}}$ ), 159.4, 159.1 ( $\text{CO}_{\text{TFA}}$ ), 118.3, 115.8 ( $\text{CF}_3$ ), 92.7 ( $\text{C}_1$ ), 73.9 ( $\text{C}_5$ ), 73.2 ( $\text{C}_3$ ), 69.5 ( $\text{C}_4$ ), 62.8 ( $\text{C}_6$ ), 54.6 ( $\text{C}_2$ ), 20.6, 20.5, 20.4, 20.3 ( $4\text{CH}_{3,\text{Ac}}$ ); **MS (ESI)**:  $m/z$  441.87 [ $\text{M-H}$ ] $^-$  (theoretical  $m/z$  442.10).

### $\beta$ -D-Glucopyranoside,3-((benzyloxy)propoxy)-2-deoxy-2[(trifluoroacetyl)amino]-3,4,6-triacetate (**75**)

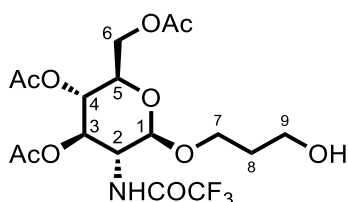


To a solution of compound **77** (1.30 g, 2.93 mmol, 1 eq.) and 3-benzyloxy-1-propanol (697  $\mu\text{L}$ , 4.40 mmol, 1.5 eq.) in anhydrous  $\text{CH}_2\text{Cl}_2$  (15 mL), boron trifluoride diethyl etherate (1.1 mL, 8.80 mmol, 3 eq.) was added under argon. The reaction mixture was stirred overnight at room temperature, then was concentrated under reduced pressure. Water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over  $\text{Na}_2\text{SO}_4$  and then concentrated under reduced pressure. The crude residue was purified by flash chromatography (cHex/EtOAc gradient 100:0 to 75:25, v/v) to afford compound **75** (983 mg, 76%) as a white solid.

$R_{f,\beta} = 0.66$  (cHex-EtOAc 5:5);  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta = 7.36 - 7.23$  (m, 5H,  $5\text{CH}_{\text{AR}}$ ), 5.28 (dd,  $J = 10.6, 9.2$  Hz, 1H,  $\text{H}_3$ ), 5.01 (dd,  $J = 10.1, 9.2$  Hz, 1H,  $\text{H}_4$ ), 4.70 (d,  $J = 8.4$  Hz, 1H,  $\text{H}_1$ ), 4.29 (dd,  $J = 12.3, 4.6$  Hz, 1H,  $\text{H}_{6a}$ ), 4.13 (dd,  $J = 12.3, 2.4$  Hz, 1H,  $\text{H}_{6b}$ ), 3.98 – 3.91, 3.67 – 3.61 (2m, 2H,  $\text{H}_7$ ), 3.90 (dd,  $J = 10.6, 8.4$  Hz, 1H,  $\text{H}_2$ ), 3.80 (ddd,  $J = 10.1, 4.6, 2.4$  Hz, 1H,  $\text{H}_5$ ), 3.54 (t,  $J = 6.3$  Hz, 1H,  $\text{H}_9$ ), 2.04, 2.01, 1.97 (3s, 9H,  $3\text{CH}_{3,\text{Ac}}$ ), 1.85 (p,  $J = 6.3$  Hz, 1H);  $^{13}\text{C NMR}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta = 172.3, 171.6, 171.2$  ( $3\text{CO}_{\text{Ac}}$ ), 159.7, 159.3, 158.9, 158.6 ( $\text{CO}_{\text{TFA}}$ ), 139.8 ( $\text{C}_{11}$ ), 129.4, 128.8, 128.6 ( $5\text{CH}_{\text{AR}}$ ), 121.6, 118.8, 115.9, 113.1 ( $\text{CF}_3$ ), 101.4

(C<sub>1</sub>), 74.0 (C<sub>10</sub>), 73.4 (C<sub>3</sub>), 73.0 (C<sub>5</sub>), 70.0 (C<sub>4</sub>), 67.9, 67.9 (C<sub>7</sub> + C<sub>9</sub>), 63.1 (C<sub>6</sub>), 55.9 (C<sub>2</sub>), 31.0 (C<sub>8</sub>), 20.6, 20.6, 20.4 (3C<sub>H<sub>3</sub>Ac</sub>); **MS (ESI)**: *m/z* 548.3 [M-H]<sup>-</sup> (theoretical *m/z* 548.17).

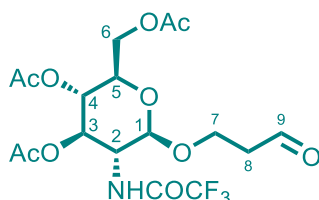
**β-D-Glucopyranoside,3-hydroxypropoxy-2-deoxy-2[(trifluoroacetyl)amino]-3,4,6-triacetate (78):**



To a solution of compound **75** (1.71 g, 3.12 mmol, 1 eq.) in EtOH (14 mL) was added 10% Pd/C (150 mg). The system was degassed three times (vacuum/N<sub>2</sub>), then stirred under H<sub>2</sub> (1 atm) for 4h. The mixture was filtered through a pad of Celite<sup>®</sup> which was washed with EtOH. After solvent removal, the crude was purified by flash chromatography (cHex/EtOAc gradient 100:0 to 75:25, v/v) to afford compound **78** (911 mg, 64%) as a white solid.

**R<sub>f,α</sub>** = 0.40 (cHex-EtOAc 4:6); **R<sub>f,β</sub>** = 0.23 (cHex-EtOAc 4:6); **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD) δ = 5.28 (dd, *J* = 10.4, 9.7 Hz, 1H, H<sub>3</sub>), 5.01 (t, *J* = 9.7 Hz, 1H, H<sub>4</sub>), 4.70 (d, *J* = 8.4 Hz, 1H, H<sub>1</sub>), 4.30 (dd, *J* = 12.3, 4.6 Hz, 1H, H<sub>6a</sub>), 4.14 (dd, *J* = 12.3, 2.3 Hz, 1H, H<sub>6b</sub>), 3.99 – 3.86 (m, 2H, H<sub>2</sub>), 3.99 – 3.86, 3.70 – 3.53 (2m, 2H, H<sub>7</sub>), 3.82 (ddd, *J* = 9.7, 4.6, 2.3 Hz, 1H, H<sub>5</sub>), 3.70 – 3.53 (m, 3H, H<sub>7b</sub> + H<sub>9</sub>), 2.06, 2.01, 1.97 (3s, 9H, 3C<sub>H<sub>3</sub>Ac</sub>), 1.83 – 1.71 (m, 2H, H<sub>8</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD) δ = 172.4, 171.6, 171.2 (3CO<sub>Ac</sub>), 159.7, 159.3, 159.0, 158.6 (CO<sub>TFA</sub>), 121.6, 118.8, 115.9, 113.1 (CF<sub>3</sub>), 101.4 (C<sub>1</sub>), 73.5 (C<sub>3</sub>), 73.0 (C<sub>5</sub>), 70.0 (C<sub>4</sub>), 68.0 (C<sub>7</sub>), 63.2 (C<sub>6</sub>), 59.5 (C<sub>9</sub>), 55.9 (C<sub>2</sub>), 33.5 (C<sub>8</sub>), 20.6, 20.6, 20.4 (3C<sub>H<sub>3</sub>Ac</sub>); **MS (ESI)**: *m/z* 457.93 [M-H]<sup>-</sup> (theoretical *m/z* 458.13).

**β-D-Glucopyranoside,3-oxopropoxy 2-deoxy-2[(trifluoroacetyl)amino]-, 3,4,6-triacetate (74):**



IBX (1.46 g, 2.35 mmol, 1.2 eq.) was dissolved in DMSO (10 mL) and the compound **78** (900 mg, 1.95 mmol, 1 eq.) in DMSO (2 mL) was added. The reaction mixture was stirred at room temperature for 6h, then was filtered through a pad of celite which was washed with EtOAc. Water was added, and the aqueous layer was extracted with EtOAc. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and then concentrated under reduced pressure. The crude residue was purified by flash

chromatography (cHex/EtOAc gradient 100:0 to 50:50, v/v) to afford compound **74** (820 mg, 92%) as a white solid.

$R_{f,B} = 0.29$  (cHex-EtOAc 5:5);  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta = 9.77$  (s, 1H,  $\text{H}_9$ ), 7.06 (d,  $J = 9.1$  Hz, NH), 5.28 (dd,  $J = 10.6, 9.7$  Hz, 1H,  $\text{H}_3$ ), 5.11 (t,  $J = 9.7$  Hz, 1H,  $\text{H}_4$ ), 4.69 (d,  $J = 8.4$  Hz, 1H,  $\text{H}_1$ ), 4.28 (dd,  $J = 12.3, 4.7$  Hz, 2H,  $\text{H}_{6a}$ ), 4.18 (dd,  $J = 12.3, 2.3$  Hz, 1H,  $\text{H}_{6b}$ ), 4.16 – 4.09 (m,  $\text{H}_{7a}$ ), 4.09 – 4.00 (m, 1H,  $\text{H}_2$ ), 3.88 (ddd,  $J = 10.4, 8.6, 4.4$  Hz, 1H,  $\text{H}_{7b}$ ), 3.75 (ddd,  $J = 9.7, 4.7, 2.3$  Hz, 1H,  $\text{H}_5$ ), 2.89 – 2.79 (1m, 2H,  $\text{H}_{8a}$ ), 2.67 (dt,  $J = 9.5, 4.4$  Hz, 1H,  $\text{H}_{8b}$ ), 2.12, 2.06, 2.05 (3s, 9H,  $3\text{CH}_{3,Ac}$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{CDCl}_3$ )  $\delta = 200.7$  ( $\text{C}_9$ ), 171.3, 170.9, 169.5 ( $3\text{CO}_{Ac}$ ), 158.1, 157.7, 157.4, 157.0 ( $\text{CO}_{TFA}$ ), 120.0, 117.2, 114.3, 111.4 ( $\text{CF}_3$ ), 100.9 ( $\text{C}_1$ ), 72.1, 72.0 ( $\text{C}_3 + \text{C}_5$ ), 68.5 ( $\text{C}_4$ ), 63.6 ( $\text{C}_7$ ), 62.0 ( $\text{C}_6$ ), 54.6 ( $\text{C}_2$ ), 43.8 ( $\text{C}_8$ ), 20.8, 20.7, 20.5 ( $3\text{CH}_{3,Ac}$ ).

## Experimental part – Chapter II - Ugi Reaction: Series A and B

### General procedure:

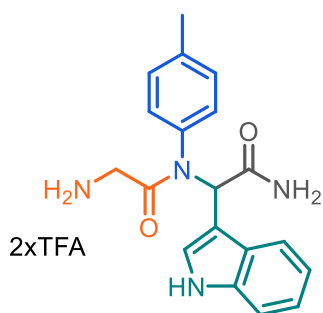
The solid-phase synthesis for the Ugi reactions was carried on an orbital shaker at room temperature, in tubes with frits. Before each reaction, the resin was swollen in CH<sub>2</sub>Cl<sub>2</sub> during 4h. The number of equivalents was calculated in relation to the initial loading of the commercial resin.

The appropriate aldehyde (2 eq.) and amine (2 eq.) were dissolved in dry MeOH (0.2 mL) and stirred for 1h at room temperature. Then, a solution of carboxylic acid (2 eq.) in dry MeOH (0.2 mL) was added to the mixture. The mixture was poured into a reaction tube containing the appropriate swollen resin-isocyanide (1 eq.) suspended in dry CH<sub>2</sub>Cl<sub>2</sub> (0.6 mL). The mixture was left stirring for 4-6 days at room temperature, then drained and washed several times with DMF (3×3 mL), MeOH (3×3 mL) and CH<sub>2</sub>Cl<sub>2</sub> (3×3 mL). The product was removed from the resin upon cleavage with a stock solution of 15% trifluoroacetic acid (TFA) (Rink amide) or 95% TFA (Rink amide MBHA) in CH<sub>2</sub>Cl<sub>2</sub> for 1h30. The remaining cleavage mixture of TFA was removed under reduced pressure by co-evaporating twice with toluene, then the crude mixture was triturated in Et<sub>2</sub>O and filtrated.

For 2,5-diketopiperazine compounds, the cyclization step was carried out in solution in MeOH for 2 to 5 days at room temperature or 50 ° C for 2 days.

Yields, purities, and some characterizations of **mixtures of diastereomers or enantiomers** obtained using different reagents and resins are summarized in **table** (NMR data as a mixture of 2 diastereomers). **For convenience reasons, the name of the molecules is written based on their fragments.**

### *p*-Toluidine-(R/S)-Indole-Gly- CONH<sub>2</sub> (**79**)



**m**<sub>Resin</sub> = 250 mg: Rink Amide MBHA (loading =0.79 mmol/g).

**m** = 20.0 mg

**Yield** = 18%

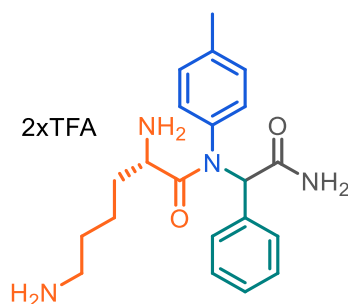
**Purity** = 95%

**HPLC**: rt = 13.35 min (method A).

**HRMS (ESI)**: *m/z* 337.16620 [M+H]<sup>+</sup> (C<sub>19</sub>H<sub>21</sub>N<sub>4</sub>O<sub>2</sub> requires 337.16590).

<sup>1</sup>H NMR (400 MHz, MeOD) δ = 7.73 – 6.68 (m, 4H, 4CH<sub>Ar</sub>), 7.44 (d, *J* = 8.0 Hz, 1H, CH<sub>Ar</sub>), 7.27 (d, *J* = 8.0 Hz, 1H, CH<sub>Ar</sub>), 7.11 – 7.05, 7.03 – 6.95 (2m, 2H, 2CH<sub>Ar</sub>), 6.93 (s, 1H, CH<sub>Ar</sub>), 6.60 (s, 1H, CH), 3.44 (s, 2H, CH<sub>2</sub>), 2.19 (s, 3H, CH<sub>3</sub>).

<sup>13</sup>C NMR (101 MHz, MeOD) δ = 174.9 (C=O), 167.4 (C=O), 140.2, 137.5, 136.1 (3C<sub>IV,Ar</sub>), 131.6, 131.0, 130.6 (4CH<sub>Ar</sub>), 128.2 (C<sub>IV,Ar</sub>), 127.8, 122.9, 120.5, 119.1, 112.5 (5CH<sub>Ar</sub>), 108.6 (C<sub>IV,Ar</sub>), 58.3 (CH), 42.1 (CH<sub>2</sub>), 21.0 (CH<sub>3</sub>).

***p*-Toluidine-(R/S)-Phenyl-L-Lys- CONH<sub>2</sub> (80)**

$m_{\text{Resin}} = 250 \text{ mg}$ : Rink Amide MBHA (loading =0.79 mmol/g).

$m = 14.7 \text{ mg}$

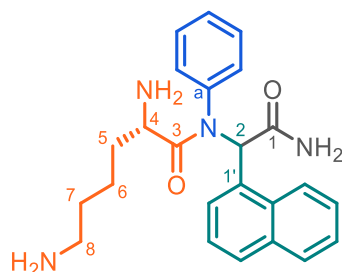
**Yield** = 13%

**Purity** = 85%

**HPLC**:  $rt_{\text{dia}1+2} = 11.50 \text{ min}$  (method A).

**HRMS (ESI)**:  $m/z$  369.22897  $[M+H]^+$  ( $C_{21}H_{29}N_4O_2$  requires 369.22850).

**<sup>1</sup>H NMR** (400 MHz, MeOD)  $\delta = 7.75 - 6.42$  (m, 9H, 9CH<sub>Ar</sub>), 6.30, 5.96 (2s, 1H, CH), 3.95 - 3.51 (m, 1H, CH), 3.02 - 2.74 (m, 2H, CH<sub>2</sub>), 2.26, 2.23 (2s, 3H, CH<sub>3</sub>), 1.90 - 1.12 (m, 6H, 3CH<sub>2</sub>).

**Phenyl-(R/S)-Naphthalene-L-Gly- CONH<sub>2</sub> (81)**

$m_{\text{Resin}} = 300 \text{ mg}$ : Rink Amide (loading =0.45 mmol/g).

$m = 9.80 \text{ mg}$

**Yield** = 13%

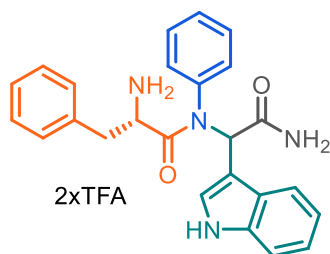
**Purity** = 95%

**HPLC**:  $rt_{\text{dia}1+2} = 15.55 \text{ min}$  (method A).

**HRMS (ESI)**:  $m/z$  405.22836  $[M+H]^+$  ( $C_{24}H_{29}N_4O_2$  requires 405.22850).

**<sup>1</sup>H NMR** (400 MHz, MeOD)  $\delta = 8.12 - 7.00$ , 6.88 - 6.70, 6.63 - 6.50, 6.14 - 5.99 (4m, 13H, 12CH<sub>Ar</sub> + H<sub>2</sub>), 4.18 - 4.08, 4.01 - 3.92 (2m, 1H, H<sub>4</sub>), 2.85 (2t,  $J = 6.8 \text{ Hz}$ , 2H, H<sub>8</sub>), 1.89 - 1.62 (m, 2H, H<sub>5</sub>), 1.55 - 1.12 (m, 4H, H<sub>6</sub> + H<sub>7</sub>).

**<sup>13</sup>C NMR** (101 MHz, MeOD)  $\delta = 174.1$ , 170.4, 169.9 ( $C_1 + C_3$ ), 136.1, 136.0, 133.3, 133.0, 131.7, 131.5 ( $3C_{IV,Ar}$ ), 130.1, 129.9, 129.7, 129.4, 129.4, 129.3, 129.2, 129.0, 129.0, 128.8, 128.7, 128.3, 127.6, 127.4, 126.5, 126.4, 125.2, 125.0, 122.9, 122.7, 122.3 ( $C_{IV,Ar} + 12CH_{Ar}$ ), 61.2, 60.9 ( $C_2$ ), 51.5, 51.3 ( $C_4$ ), 38.8, 38.8 ( $C_8$ ), 29.8, 29.3 ( $C_5$ ), 25.9, 25.8 ( $C_7$ ), 21.2, 20.7 ( $C_6$ ).

**Phenyl-(R/S)-Indole-L-Phe- CONH<sub>2</sub> (82)**

$m_{\text{Resin}} = 300 \text{ mg}$ : Rink Amide (loading =0.45 mmol/g).

$m = 32.0 \text{ mg}$

**Yield** = 37%

**Purity** = 91%

**HPLC**:  $rt_{\text{dia}1} = 15.74 \text{ min}$  and  $rt_{\text{dia}2} = 15.96 \text{ min}$  (method A).

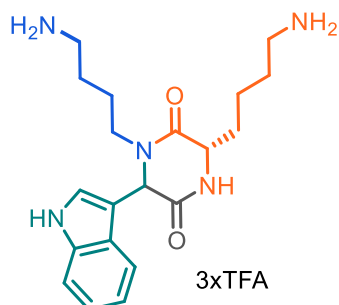
**HRMS (ESI)**:  $m/z$  413.19736  $[M+H]^+$  ( $C_{25}H_{25}N_4O_2$  requires 413.19720).

**<sup>1</sup>H NMR** (400 MHz, MeOD)  $\delta = 7.75 - 6.71$  (m, 15H, 15CH<sub>Ar</sub>), 6.59, 6.54 (2s, 1H, CH), 3.93, 3.85 (dd and t,  $J_{\text{dia}1} = 9.0, 4.4 \text{ Hz}$  and  $J_{\text{dia}2} = 7.2 \text{ Hz}$ , 1H, CH), 3.31 - 3.08, 2.92 - 2.79 (2m, 2H, CH<sub>2</sub>).

**<sup>13</sup>C NMR** (101 MHz, MeOD)  $\delta = 175.0$ , 174.7 (C=O), 169.7, 169.5 (C=O), 139.6, 138.8, 137.5, 137.5, 135.5, 135.5, 134.1, 131.9 ( $4C_{IV,Ar}$ ), 131.6, 131.1, 130.5, 130.5, 130.2, 130.1, 130.1, 130.0, 129.9, 129.9, 129.7,

128.8, 128.7, 128.3, 128.1, 127.9, 127.9, 125.9, 123.1, 122.9, 121.4, 120.5, 120.5, 119.3, 119.2, 112.5 (15 $\underline{\text{C}}_{\text{Ar}}$ ), 108.6, 108.5 ( $\text{C}_{\text{IV,Ar}}$ ), 58.8, 58.4 ( $\underline{\text{C}}_{\text{H}}$ ), 54.6, 54.1 ( $\underline{\text{C}}_{\text{H}}$ ), 38.5, 37.9 ( $\underline{\text{C}}_{\text{H}_2}$ ).

#### 4-Aminobutyl-(R/S)-Indole-L-Lys-2,5-DKP (A1)



$m_{\text{Resin}} = 250 \text{ mg}$  : Rink Amide MBHA (loading = 0.79 mmol/g).

$m = 23.7 \text{ mg}$

**Yield** = 17%

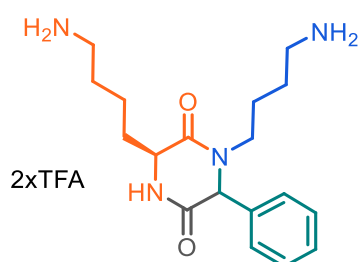
**Purity** = 95%

**HPLC**:  $rt_{\text{dia1}} = 8.49 \text{ min}$  and  $rt_{\text{dia2}} = 8.85 \text{ min}$  (method A).

**HRMS (ESI)**:  $m/z$  372.23975 [ $\text{M}+\text{H}$ ] $^+$  ( $\text{C}_{19}\text{H}_{28}\text{N}_5\text{O}_2$  requires 372.23940).

$^1\text{H NMR}$  (400 MHz, MeOD)  $\delta = 7.72, 7.57$  (2d,  $J = 7.9 \text{ Hz}$ , 1H,  $\underline{\text{C}}_{\text{HAr}}$ ), 7.48 – 7.36, 7.18 – 7.02 (2m, 3H, 3 $\underline{\text{C}}_{\text{HAr}}$ ), 7.33, 7.20 (2s, 1H,  $\underline{\text{C}}_{\text{HAr}}$ ), 5.45, 5.38 (2s, 1H,  $\underline{\text{C}}_{\text{H}}$ ), 4.38, 4.10 – 4.00 (t and m,  $J = 4.4 \text{ Hz}$ , 1H,  $\underline{\text{C}}_{\text{H}}$ ), 3.95 – 3.53, 3.16 – 2.76 (2m, 6H, 3 $\underline{\text{C}}_{\text{H}_2}$ ), 2.13 – 1.26 (m, 10H, 5 $\underline{\text{C}}_{\text{H}_2}$ ).

#### 4-Aminobutyl-(R/S)-Phenyl-L-Lys-2,5-DKP (A2)



$m_{\text{Resin}} = 250 \text{ mg}$ : Rink Amide MBHA (loading = 0.79 mmol/g).

$m = 11.3 \text{ mg}$

**Yield** = 11 %

**Purity** = 52 %

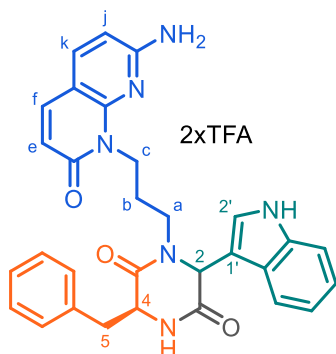
**HPLC**:  $rt_{\text{dia1+2}} = 5.57 \text{ min}$  (method B).

**MS (ESI)**:  $m/z$  333.20 [ $\text{M}+\text{H}$ ] $^+$  (theoretical  $m/z$  332.22).

Purified by semipreparative HPLC (method B):

$^1\text{H NMR}$  (400 MHz, MeOD)  $\delta$  7.51 – 7.33 (m, 5H, 5 $\underline{\text{C}}_{\text{HAr}}$ ), 5.20, 5.11 (2s, 1H,  $\underline{\text{C}}_{\text{H}}$ ), 4.35 – 3.28, 4.06 – 3.84 (2m, 1H,  $\underline{\text{C}}_{\text{H}}$ ), 4.00 – 3.89, 3.83 – 3.70, 3.00 – 2.77 (3m, 2H,  $\underline{\text{C}}_{\text{H}_2}$ ), 3.00 – 2.77 (m, 4H, 2 $\underline{\text{C}}_{\text{H}_2}$ ), 2.09 – 1.85, 1.81 – 1.32 (m, 10H, 5 $\underline{\text{C}}_{\text{H}_2}$ ).

#### Naphthyridine(N-alkyl)-(R/S)-Indole-L-Phe-2,5-DKP (A3)



$m_{\text{Resin}} = 300 \text{ mg}$ : Rink Amide MBHA (loading = 0.52 mmol/g).

$m = 11.2 \text{ mg}$

**Yield** = nd %

**Purity** = nd %

**HPLC**:  $rt_{\text{dia1}} = 17.70 \text{ min}$  and  $rt_{\text{dia2}} = 18.08 \text{ min}$  (method A).

**HRMS (ESI)**:  $m/z$  521.22957 [ $\text{M}+\text{H}$ ] $^+$  ( $\text{C}_{30}\text{H}_{29}\text{N}_6\text{O}_3$  requires 521.22957).

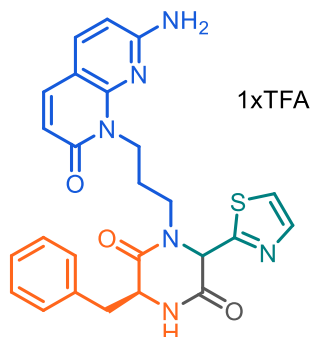
Purified by semipreparative HPLC (method B):

$^1\text{H NMR}$  (400 MHz, MeOD)  $\delta = 7.71 – 7.57$  (m, 2H,  $H_f + H_k$ ), 7.44 – 7.22, 7.17 – 6.84 (2m, 9H, 9 $\underline{\text{C}}_{\text{HAr}}$ ), 6.49, 6.47 (2d,  $J = 8.2 \text{ Hz}$ , 1H,  $H_j$ ), 6.28, 6.25 (2d,  $J = 9.2 \text{ Hz}$ , 1H,  $H_e$ ), 6.01 (s, 1H,  $H_2$ ), 5.69, 5.30 (2s, 1H,  $H_2$ ), 4.75



– 4.70, 4.52 – 4.46 (2m, 1H,  $H_d$ ), 4.44 – 4.26, 2.94 – 2.83 (2m, 2H,  $H_c$ ), 4.06 – 3.88, 3.63 – 3.51, 2.68 – 2.58 (3m, 2H,  $H_a$ ), 3.47 – 3.27, 3.12 (m and dd,  $J = 13.7, 4.3$  Hz, 2H,  $H_5$ ), 1.99 – 1.52 (m, 2H,  $H_b$ ).

### Naphthyridine(*N*-alkyl)-(R/S)-Thiazol-L-Phe-2,5-DKP (A5)



$m_{\text{Resin}} = 250$  mg: Rink Amide MBHA (loading = 0.79 mmol/g).

$m = 14.0$  mg

**Yield** = 13 %

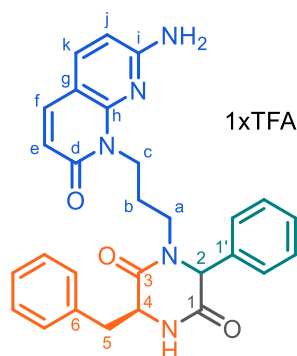
**Purity** = 60 %

**HPLC**:  $rt_{\text{dia1}} = 17.00$  min and  $rt_{\text{dia2}} = 17.84$  min (method A).

**HRMS (ESI)**:  $m/z$  521.22957  $[M+H]^+$  ( $C_{30}H_{29}N_6O_3$  requires 521.22957).

Purified by semipreparative HPLC (method B): Purity = 84 %.

### Naphthyridine(*N*-alkyl)-(R/S)-Phenyl-L-Phe-2,5-DKP (A6)



$m_{\text{Resin}} = 300$  mg : Rink Amide (loading = 0.45 mmol/g)

$m = 12.9$  mg

**Yield** = 24%

**Purity** = 63%

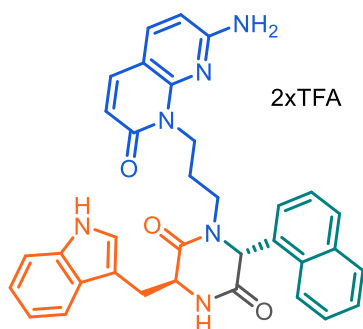
**HPLC**:  $rt_{\text{dia1+2}} = 18.59$  min (method A).

**HRMS (ESI)**:  $m/z$  482.21924  $[M+H]^+$  ( $C_{28}H_{28}N_5O_3$  requires 482.21867).

Purified by semipreparative HPLC (method B):

**$^1\text{H NMR}$**  (400 MHz, MeOD)  $\delta = 7.71 - 7.59$  (m, 2H,  $H_f + H_k$ ), 7.40 – 7.19, 7.16 – 7.08, 6.94 – 6.88 (3m, 8H,  $8\text{CH}_{Ar}$ ), 7.03 (t,  $J = 7.6$  Hz, 1H,  $\text{CH}_{Ar}$ ), 6.50, 6.48 (2d,  $J = 8.3$  Hz, 1H,  $H_j$ ), 6.39 (d,  $J = 7.3$  Hz, 1H,  $\text{CH}_{Ar}$ ), 6.29, 6.28 (2d,  $J = 9.2$  Hz, 1H,  $H_e$ ), 5.32, 5.25 (2s, 1H,  $H_2$ ), 4.72 – 4.67, 4.55 – 4.48 (2m, 1H,  $H_4$ ), 4.32, 3.94 – 3.77, 3.51 – 3.41 (t and 2m,  $J = 7.3$  Hz, 2H,  $H_c$ ), 4.10 – 3.95, 3.94 – 3.77, 2.78 – 2.66, 2.55 – 2.45 (4m, 2H,  $H_a$ ), 3.53 – 3.32, 3.13 – 3.03 (2m, 2H,  $H_5$ ), 1.98 – 1.84, 1.75 – 1.51 (2m, 2H,  $H_b$ ).

**$^{13}\text{C NMR}$**  (101 MHz, MeOD)  $\delta = 168.8, 168.2$  ( $C_1$ ), 167.4, 167.1 ( $C_3$ ), 165.7, 161.9 ( $C_i$ ), 150.9, 150.8 ( $C_h$ ), 140.3 ( $C_j$ ), 139.3 ( $C_k$ ), 137.9, 137.1, 137.0, 136.4 ( $C_{1'} + C_6$ ), 131.8, 131.6, 130.1, 129.9, 129.9, 129.7, 129.5, 129.3, 128.7, 128.5, 128.5, 128.0 ( $10\text{CH}_{Ar}$ ), 114.8, 114.7 ( $C_e$ ), 108.5, 108.4 ( $C_g$ ), 106.7, 106.6 ( $C_e$ ), 64.0, 62.9 ( $C_2$ ), 57.7, 57.1 ( $C_4$ ), 43.2, 41.6 ( $C_a$ ), 40.3 ( $C_5$ ), 39.2, 39.0 ( $C_c$ ), 25.4, 24.3 ( $C_b$ ).

**Naphthyridine(*N*-alkyl)-(R/S)-Naphthalene-L-Trp-2,5-DKP (A8)**

$m_{\text{Resin}}$  = 300 mg: Rink Amide (loading =0.45 mmol/g).

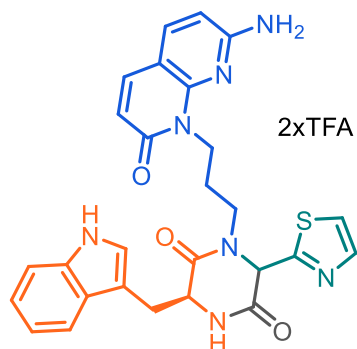
$m$  = 39.0 mg

**Yield** = 36%

**Purity** = 82%

**HPLC:**  $rt_{\text{dia1}}$  = 19.42 min and  $rt_{\text{dia2}}$  = 19.54 min (method A).

**HRMS (ESI):**  $m/z$  571.24567  $[M+H]^+$  ( $C_{34}H_{31}N_6O_3$  requires 571.24522).

**Naphthyridine(*N*-alkyl)-(R/S)-Thiazol-L-Trp-2,5-DKP (A9)**

$m_{\text{Resin}}$  = 300 mg: Rink Amide MBHA (loading =0.52 mmol/g).

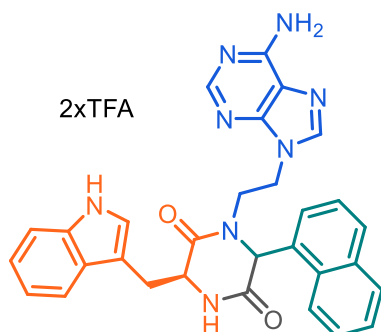
$m$  = 16.8 mg

**Yield** = 17%

**Purity** = 81%

**HPLC:**  $rt_{\text{dia1}}$  = 16.31 min and  $rt_{\text{dia2}}$  = 17.30 min (method A).

**HRMS (ESI):**  $m/z$  528.18182  $[M+H]^+$  ( $C_{27}H_{26}N_7O_3S$  requires 528.18123).

**Adenine-(R/S)-Naphthalene-L-Trp-2,5-DKP (A10)**

$m_{\text{Resin}}$  = 300 mg: Rink Amide (loading =0.45 mmol/g).

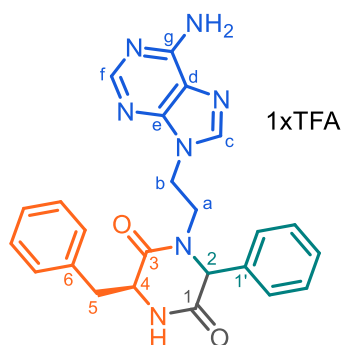
$m$  = 28.1 mg

**Yield** = 39%

**Purity** = 80%

**HPLC:**  $rt_{\text{dia1}}$  = 13.76 min and  $rt_{\text{dia2}}$  = 14.39 min (method A).

**HRMS (ESI):**  $m/z$  531.22546  $[M+H]^+$  ( $C_{30}H_{27}N_8O_2$  requires 531.22515).

**Adenine-(R/S)-Phenyl-L-Trp-2,5-DKP (A11a and A11b)**

$m_{\text{Resin}}$  = 208 mg: Rink Amide MBHA (loading =0.52 mmol/g).

$m$  = 28.4 mg

**Yield** = 47%

**Purity** = 68%

**HPLC:**  $rt_{\text{dia1}}$  = 13.10 min and  $rt_{\text{dia2}}$  = 13.45 min (method A).

**HRMS (ESI):**  $m/z$  531.22546  $[M+H]^+$  ( $C_{30}H_{27}N_8O_2$  requires 531.22515).

Purified by semipreparative HPLC (method B):

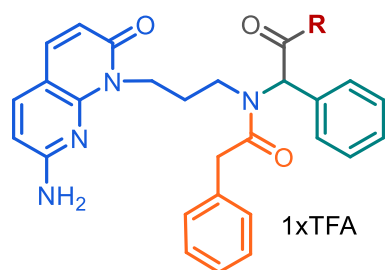
**A11a** (dia1):

$^1\text{H NMR}$  (400 MHz,  $\text{DMSO-}d_6$ )  $\delta$  = 8.29 (s, NH), 8.09 (s, 1H,  $H_f$ ), 7.33 (s, 1H,  $H_c$ ), 7.40 – 7.18, 7.18 – 7.10 (2m, 10H,  $10\text{CH}_{Ar}$ ), 4.58 (t,  $J$  = 4.3 Hz, 1H,  $H_d$ ), 4.34 (s, 1H,  $H_2$ ), 4.23, 3.97 (2dt,  $J$  = 13.3, 5.8 Hz, 2H,  $H_b$ ), 4.11, 2.79 (2dt,  $J$  = 12.8, 5.8 Hz, 2H,  $H_a$ ), 3.14, 2.91 (dd,  $J$  = 13.7, 4.3 Hz, 1H,  $H_5$ ).

$^{13}\text{C NMR}$  (101 MHz,  $\text{DMSO-}d_6$ )  $\delta$  = 166.0 ( $C_1$ ), 164.8 ( $C_3$ ), 156.0 ( $C_g$ ), 152.5 ( $C_f$ ), 149.5 ( $C_e$ ), 140.2 ( $C_c$ ), 136.9, 136.2 ( $2\text{C}_{IV,Ar}$ ), 130.2, 128.9, 128.4, 128.1, 127.1, 126.7 ( $10\text{CH}_{Ar}$ ), 118.7 ( $C_d$ ), 64.1 ( $C_2$ ), 54.2 ( $C_4$ ), 43.8 ( $C_a$ ), 40.6 ( $C_b$ ), 37.6 ( $C_5$ ).

**A11a** (dia2): purity= 60 %.

### Naphthyridine(*N*-alkyl)-(R/S)-Phenyl-Phenylacetic-C(O)R (**B1**)



R =  $\text{NH}_2$ , OH et OMe  
(and benzoic acid)

$m_{\text{Resin}}$  = 280mg: Rink Amide MBHA (loading =0.52 mmol/g).

$m$  = 23.7 mg

Yield = 28%

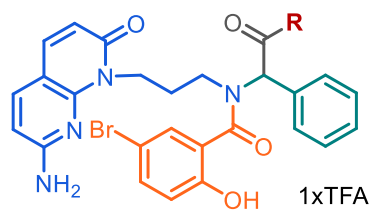
**(B1a)** R =  $\text{NH}_2$  (53%),  $r_t$  = 17.53 min (A), **HRMS** (ESI):  $m/z$  470.21921  $[\text{M}+\text{H}]^+$  ( $\text{C}_{27}\text{H}_{28}\text{N}_5\text{O}_3$  requires 470.21867).

**(B1b)** R = OH (12%),  $r_t$  = 19.10 min (A).

**(B1c)** R = OMe (11%),  $r_t$  = 21.72 min (A), **HRMS** (ESI):  $m/z$  485.21856  $[\text{M}+\text{H}]^+$  ( $\text{C}_{28}\text{H}_{29}\text{N}_4\text{O}_4$  requires 485.21833).

**(B1d)** R = by-product benzoic acid (10%),  $r_t$  = 16.40 min (A), **HRMS** (ESI):  $m/z$  456.20335  $[\text{M}+\text{H}]^+$  ( $\text{C}_{26}\text{H}_{26}\text{N}_5\text{O}_3$  requires 456.20302).

### Naphthyridine(*N*-alkyl)-(R/S)-Phenyl-(5-bromo-2-hydroxybenzoic)-C(O)R (**B4**)



R =  $\text{NH}_2$ , OH, OMe  
and cyclized by-product

$m_{\text{Resin}}$  = 250mg: Rink Amide MBHA (loading =0.52 mmol/g).

$m$  = 27.8 mg

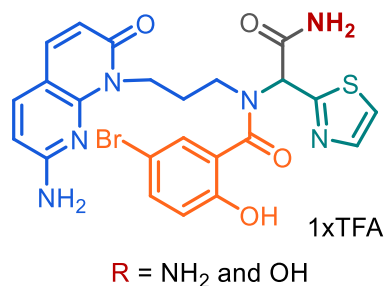
Yield = 32%

**(B4a)** R =  $\text{NH}_2$  (69%),  $r_t$  = 17.13 min (A), **HRMS** (ESI):  $m/z$  550.10870  $[\text{M}+\text{H}]^+$  ( $\text{C}_{26}\text{H}_{25}\text{N}_5\text{O}_4\text{Br}$  requires 550.10844).

**(B4b)** R = OH (7%),  $r_t$  = 17.86 min (A).

**(B4c)** R = OMe (9%),  $r_t$  = 19.23 min (A).

**(B4d)** R = cyclized by-product (11%),  $r_t$  = 20.09 min (A).

**Naphthyridine(*N*-alkyl)-(R/S)-Thiazol-(5-bromo-2-hydroxybenzoic)- CONH<sub>2</sub> (B5)**

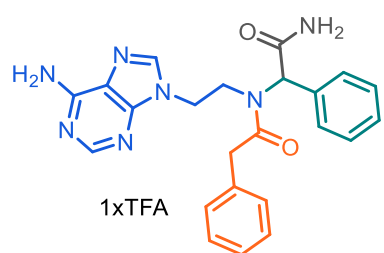
$m_{\text{Resin}}$  = 250mg: Rink Amide MBHA (loading =0.52 mmol/g).

$m$  = 14.9 mg

Yield = 17%

(B5a) R = NH<sub>2</sub> (53%), rt = 15.82 min (A), HRMS (ESI):  $m/z$  557.06122 [M+H]<sup>+</sup> (C<sub>23</sub>H<sub>22</sub>N<sub>6</sub>O<sub>4</sub>BrS requires 557.06011).

(B5b) R = OH (18%), rt = 16.76 min (A).

**Indole-(R/S)-Phenyl-Phenylacetic-C(O)R (B2a)**

$m_{\text{Resin}}$  = 200 mg: Rink Amide MBHA (loading =0.52 mmol/g).

$m$  = 14.7 mg

Yield = 26%

Purity = 88%

HPLC: rt = 8.06 min (method B).

HRMS (ESI):  $m/z$  430.19879 [M+H]<sup>+</sup> (C<sub>23</sub>H<sub>24</sub>N<sub>7</sub>O<sub>2</sub> requires 430.19860).

Purified by semipreparative HPLC (method B): NMR data as a mixture of 2 conformers:

<sup>1</sup>H NMR (400 MHz, MeOD)  $\delta$  = 8.33, 8.20 (2s, 1H, CH), 7.98, 7.91 (2s, 1H, CH), 7.49 – 7.22, 7.22 – 7.16, 7.01 – 6.94 (3m, 10H, 10CH<sub>Ar</sub>), 6.06, 5.77 (2s, 1H, CH), 4.40 – 4.19, 4.07 – 3.59 (2m, 6H, 3CH<sub>2</sub>).

<sup>13</sup>C NMR (101 MHz, MeOD)  $\delta$  = 175.3, 174.7, 173.7 (C=O), 152.0, 150.2 (2C<sub>IV,Ar</sub>), 146.7, 145.3, 145.3, 145.0 (2CH), 135.9, 135.7 (2C<sub>IV,Ar</sub>), 131.3, 130.5, 130.3, 130.2, 130.1, 130.0, 129.8, 129.6, 129.6, 128.4 (10CH<sub>Ar</sub>), 119.7 (C<sub>IV,Ar</sub>), 65.5, 64.2 (CH), 46.2, 45.3, 44.7, 43.9, 42.3, 41.1 (3CH<sub>2</sub>).

## Experimental part – Chapter II - Ugi Reaction: Serie C

### General procedures :

#### General procedure A - Synthesis of UGI products :

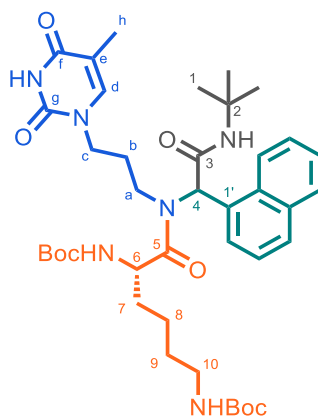
A mixture of amine (1 eq.) and aldehyde (1 eq.) in MeOH or CH<sub>2</sub>Cl<sub>2</sub>/MeOH was stirred at room temperature for 2h. In cases when the amine was taken as a salt, a solution of KOH (1 eq.) in MeOH (0.1 mL) was added as well. Carboxylic acid (1 eq.) and isocyanide (1 eq.) were added, and the reaction mixture was stirred at room temperature for 24 to 72 hours as monitored by TLC. The reaction mixture was concentrated under reduced pressure. A saturated NaHCO<sub>3</sub> aq. was added, and the aqueous layer was extracted with CH<sub>2</sub>Cl<sub>2</sub>. The combined organic layers were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub>, and concentrated under reduced pressure. The crude residue was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH or cHex/EtOAc) to afford a mixture of diastereomers.

#### General procedure B - Boc and acetal groups removal:

To a solution of protected compounds in CH<sub>2</sub>Cl<sub>2</sub> were added 50 equiv. of TFA. The reaction mixture was stirred at room temperature for 4h. The solvent and the remaining TFA were then removed under reduced pressure by co-evaporating twice with toluene, then the crude mixture was triturated in Et<sub>2</sub>O. After filtration, a **mixture of diastereomers (TFA salt)** was obtained.

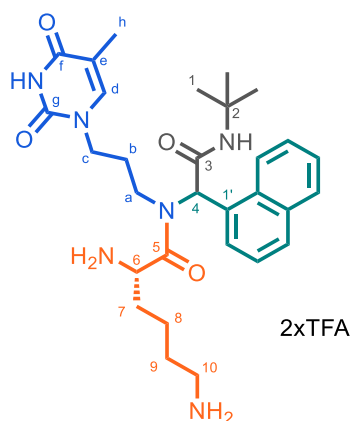
**For convenience reasons, the name of the molecules is written based on their fragments.**

## Serie C-1

***N*<sup>1</sup>-Thymine-(*R/S*)-Naphthalene-(Boc)-L-Lys(Boc)-*tert*-butylacetamide (**C1a**):**

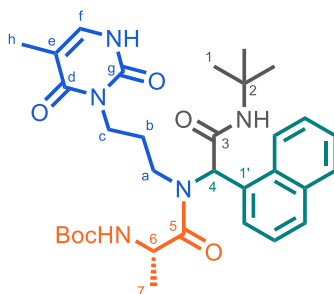
Compound **C1a** was prepared following general A procedure using the TFA salt of *N*<sup>1</sup>-thymine-NH<sub>2</sub> (**34**) (53.5 mg, 0.18 mmol) as the amine, **1-naphthaldehyde** (24.7 μL, 0.18 mmol) as aldehyde, **Boc-L-Lys(Boc)-OH** (62.3 mg, 0.18 mmol) as the acid and **tert-butyl isocyanide** (20.4 μL, 0.18 mmol) in MeOH (1.0 mL). After purification by flash chromatography (cHex/EtOAc gradient 60:40 to 10:90, v/v), a mixture of diastereomers **C1a** was obtained as a white solid (45.0 mg, 33%).

$R_{f, dia1+2} = 0.36$  (cHex-EtOAc 20:80); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.41 - 8.28$ ,  $8.01 - 7.77$ ,  $7.75 - 7.34$  (3m, 7H, 7CH<sub>Ar</sub>), 6.95, 6.77, 6.65 (3s, 1H, H<sub>d</sub>), 6.81, 6.63, 6.60 (3s, 1H, H<sub>4</sub>), 4.83 - 4.72, 4.52 - 4.20 (2m, 1H, H<sub>6</sub>), 3.60 - 3.12 (m, 4H, H<sub>a</sub> + H<sub>c</sub>), 3.11 - 2.92 (m, 2H, H<sub>10</sub>), 1.88 - 1.59 (m, 5H, H<sub>h</sub> + H<sub>7</sub>), 1.59 - 1.27 (m, 31H, 6CH<sub>3, Boc</sub> + H<sub>1</sub> + H<sub>8-9</sub>), 1.88 - 1.59, 1.04 - 0.64, 0.57 - 0.39 (3m, 2H, H<sub>b</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 176.2$ , 175.8 (C<sub>5</sub>), 171.8, 171.7, 171.4 (C<sub>3</sub>), 166.8, 166.8 (C<sub>f</sub>), 158.5, 158.4, 157.8, 157.7 (2CO<sub>Boc</sub>), 152.5, 152.4, 152.3 (C<sub>g</sub>), 142.7, 142.5, 142.5 (C<sub>d</sub>), 135.1, 135.1, 135.0, 134.0, 133.8, 133.7, 133.0, 132.8, 132.4 (3C<sub>IV, Ar</sub>), 131.1, 130.7, 130.6, 130.2, 130.0, 129.9, 128.5, 128.5, 128.3, 128.2, 128.1, 127.5, 127.4, 127.4, 126.3, 126.1, 125.3, 124.7, 123.9 (7CH<sub>Ar</sub>), 111.1, 111.0, 110.8 (C<sub>e</sub>), 80.8, 80.5, 80.4, 79.8 (2C<sub>IV, Boc</sub>), 62.3, 60.4, 59.4 (C<sub>4</sub>), 53.0, 52.6, 52.4, 52.3 (C<sub>2</sub> + C<sub>6</sub>), 46.8, 46.4, 45.9, 44.0, 43.8 (C<sub>a</sub> + C<sub>c</sub>), 41.1 (C<sub>10</sub>), 33.9, 32.9 (C<sub>7</sub>), 30.9, 30.7, 30.6 (C<sub>b</sub> + C<sub>9</sub>), 28.8, 28.8 (6CH<sub>3, Boc</sub> + C<sub>1</sub>), 24.6, 24.3, 24.2 (C<sub>8</sub>), 12.3, 12.3 (C<sub>h</sub>); **MS (ESI):**  $m/z$  751.5 [M+H]<sup>+</sup> (theoretical  $m/z$  751.44).

***N*<sup>1</sup>-Thymine-(R/S)-Naphthalene-L-Lys-*tert*-butylacetamide (C1b):**

General procedure B was employed for the deprotection of compound **C1a** (22.0 mg, 0.029 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C1b** was isolated as a white powder (20.7 mg, 91%).

**HPLC:**  $rt_{dia1}$  = 13.01 min and  $rt_{dia2}$  = 13.19 min (method A); NMR data as a mixture of 2 diastereomers; **<sup>1</sup>H NMR** (400 MHz, D<sub>2</sub>O)  $\delta$  = 7.99 – 7.89 (m, 1H, CH<sub>Ar</sub>), 7.89 – 7.82 (m, 1H, CH<sub>Ar</sub>), 7.81 – 7.72, 7.69 – 7.54 (m, 3H, 3CH<sub>Ar</sub>), 7.44 – 7.32 (m, 2H, 2CH<sub>Ar</sub>), 6.74, 6.59 (2s, 1H, H<sub>d</sub>), 6.55, 6.31 (2s, 1H, H<sub>4</sub>), 4.54 – 4.44 (2m, 1H, H<sub>6</sub>), 3.43 – 3.34, 3.33 – 3.12, 2.70 – 2.58 (3m, 4H, H<sub>a</sub> + H<sub>c</sub>), 3.08 – 2.98 (m, 2H, H<sub>10</sub>), 2.09 – 1.95 (m, 2H, H<sub>7</sub>), 1.82 – 1.67 (m, 2H, H<sub>9</sub>), 1.69, 1.65 (2s, 3H, H<sub>h</sub>), 1.68 – 1.45 (m, 2H, H<sub>8</sub>), 1.34, 1.29 (2s, 9H, H<sub>i</sub>), 1.83 – 1.47, 1.39 – 1.22, 0.77 - 0.60 (2m, 2H, H<sub>b</sub>); **<sup>13</sup>C NMR** (101 MHz, D<sub>2</sub>O)  $\delta$  = 170.8, 170.8, 170.3, 170.0 (C<sub>3</sub> + C<sub>5</sub>), 166.6, 166.5 (C<sub>f</sub>), 151.2, 151.1 (C<sub>g</sub>), 142.3, 142.2 (C<sub>d</sub>), 133.2, 133.1, 131.7, 131.4 (2C<sub>IV,Ar</sub>), 130.1, 130.0, 129.3, 129.1, 128.7, 128.5 (C<sub>IV,Ar</sub> + 2CH<sub>Ar</sub>), 128.2, 127.8, 127.8, 127.5, 126.8, 126.8, 125.2, 125.1, 122.5, 121.7 (5CH<sub>Ar</sub>), 110.6 (C<sub>e</sub>), 61.1, 59.9 (C<sub>4</sub>), 51.9, 51.6 (C<sub>2</sub>), 51.0, 50.9 (C<sub>6</sub>), 45.3, 45.1, 42.6, 42.5 (C<sub>a</sub> + C<sub>c</sub>), 39.0, 39.0 (C<sub>10</sub>), 30.8, 30.5 (C<sub>7</sub>), 29.4, 29.2 (C<sub>b</sub>), 27.6, 27.6 (C<sub>1</sub>), 26.5 (C<sub>9</sub>), 21.6, 20.8 (C<sub>8</sub>), 11.2, 11.2 (C<sub>h</sub>); **HRMS (ESI):**  $m/z$  551.33453 [M+H]<sup>+</sup> (C<sub>30</sub>H<sub>43</sub>N<sub>6</sub>O<sub>4</sub> requires 551.33403).

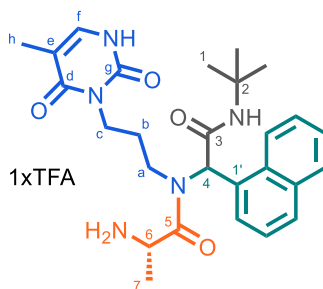
***N*<sup>3</sup>-Thymine-(R/S)-Naphthalene-(Boc)-L-Ala-*tert*-butylacetamide (C2a):**

Compound **C2a** was prepared following general A procedure using the TFA salt of *N*<sup>3</sup>-thymine-NH<sub>2</sub> (**35**) (54.0 mg, 0.18 mmol) as the amine, **1-naphthaldehyde** (24.7  $\mu$ L, 0.18 mmol) as aldehyde, **Boc-Ala-OH** (34.1 mg, 0.18 mmol) as the acid and *tert*-butyl isocyanide (20.4  $\mu$ L, 0.18 mmol) in MeOH (0.6 mL).

After purification by flash chromatography (cHex/EtOAc gradient 65:35 to 15:85, v/v), a mixture of diastereomers **C2a** was obtained as a white solid (49.0 mg, 46%).

$R_{f, dia1+2} = 0.47$  (cHex-EtOAc 10:90); NMR data as a mixture of 2 diastereomers and conformers;  $^1\text{H NMR}$  (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 8.34 - 8.28$ ,  $7.99 - 7.77$ ,  $7.76 - 7.63$ ,  $7.60 - 7.29$  (4m, 7H,  $7\text{CH}_{Ar}$ ),  $7.14$ ,  $7.12$ ,  $7.09$ ,  $7.08$  (3d and br d,  $J_{dia1} = 1.0$  Hz and  $J_{dia2} = 0.7$  Hz, 1H,  $H_f$ ),  $6.81$ ,  $6.73$ ,  $6.66$ ,  $6.59$  (4s, 1H,  $H_4$ ),  $4.90 - 4.74$ ,  $4.54 - 4.32$  (2m, 1H,  $H_6$ ),  $3.63 - 2.96$  (m, 4H,  $H_a + H_c$ ),  $1.89 - 1.57$ ,  $0.79 - 0.40$  (2m, 2H,  $H_b$ ),  $1.79$ ,  $1.79$ ,  $1.78$ ,  $1.74$  (4s, 3H,  $H_h$ ),  $1.56 - 1.22$  (m, 21H,  $3\text{CH}_{3.Boc} + H_1 + H_7$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 176.9$ ,  $176.7$ ,  $176.4$  ( $C_5$ ),  $171.8$ ,  $171.4$  ( $C_3$ ),  $166.1$ ,  $165.9$  ( $C_d$ ),  $157.6$ ,  $157.4$  ( $\text{CO}_{Boc}$ ),  $153.1$ ,  $153.0$  ( $C_g$ ),  $137.4$ ,  $137.2$  ( $C_f$ ),  $135.1$ ,  $134.0$ ,  $133.9$ ,  $132.9$ ,  $132.8$  ( $3\text{C}_{IV.Ar}$ ),  $131.1$ ,  $130.7$ ,  $130.6$ ,  $130.2$ ,  $129.8$ ,  $128.6$ ,  $128.4$ ,  $128.3$ ,  $128.1$ ,  $128.0$ ,  $127.3$ ,  $127.3$ ,  $127.2$ ,  $126.1$ ,  $125.9$ ,  $125.3$ ,  $124.6$ ,  $123.7$  ( $7\text{CH}_{Ar}$ ),  $109.8$ ,  $109.7$  ( $C_e$ ),  $80.7$ ,  $80.4$ ,  $80.3$  ( $\text{C}_{IV.Boc}$ ),  $62.7$ ,  $60.4$ ,  $59.7$  ( $C_4$ ),  $52.8$ ,  $52.3$ ,  $52.2$  ( $C_2$ ),  $C_6$  (signal under the  $\text{CD}_3\text{OD}$  peak),  $44.9$  ( $C_a$ ),  $38.9$ ,  $38.5$  ( $C_c$ ),  $29.8$  ( $C_b$ ),  $28.9$ ,  $28.8$ ,  $28.8$ ,  $28.8$  ( $3\text{CH}_{3.Boc} + C_1$ ),  $18.6$ ,  $18.3$ ,  $18.2$  ( $C_7$ ),  $12.9$  ( $C_h$ ); **MS (ESI)**:  $m/z$  594.3  $[\text{M}+\text{H}]^+$  (theoretical  $m/z$  594.33).

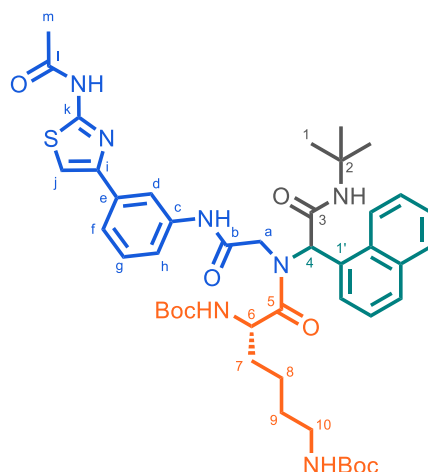
### ***N*<sup>3</sup>-Thymine-(R/S)-Naphthalene-L-Ala-*tert*-butylacetamide (**C2b**):**



General procedure B was employed for the deprotection of compound **C2a** (33.0 mg, 0.056 mmol) in 1 mL of  $\text{CH}_2\text{Cl}_2$  and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C2b** was isolated as a white solid (30.0 mg, 89%).

**HPLC**:  $rt_{dia1} = 14.79$  min and  $rt_{dia2} = 15.09$  min (method A); NMR data as a mixture of 2 diastereomers;  $^1\text{H NMR}$  (400 MHz,  $\text{D}_2\text{O}$ )  $\delta = 7.98 - 7.89$ ,  $7.86 - 7.78$  (m, 2H,  $2\text{CH}_{Ar}$ ),  $7.76 - 7.67$ ,  $7.64 - 7.51$  (2m, 3H,  $3\text{CH}_{Ar}$ ),  $7.39 - 7.28$  (m, 2H,  $2\text{CH}_{Ar}$ ),  $7.18$ ,  $7.15$  (2d,  $J_{dia1} = 1.0$  Hz and  $J_{dia2} = 0.7$  Hz, 1H,  $H_f$ ),  $6.49$ ,  $6.25$  (2s, 1H,  $H_4$ ),  $4.57$ ,  $4.53$  (2q,  $J = 6.8$  Hz, 1H,  $H_6$ ),  $3.49 - 3.28$ ,  $3.24 - 3.03$  (2m, 2H,  $H_c$ ),  $3.24 - 3.03$ ,  $2.67 - 2.51$  (2m, 2H,  $H_a$ ),  $1.79 - 1.46$  (m, 6H,  $H_h + H_7$ ),  $1.79 - 1.50$ ,  $1.37 - 1.09$ ,  $0.64 - 0.45$  (3m, 2H,  $H_b$ ),  $1.32$ ,  $1.28$  (2s, 9H,  $H_1$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{D}_2\text{O}$ )  $\delta = 171.8$ ,  $171.4$ ,  $170.9$ ,  $170.2$  ( $C_3 + C_5$ ),  $165.7$ ,  $165.6$  ( $C_d$ ),  $151.9$ ,  $151.9$  ( $C_g$ ),  $137.0$  ( $C_f$ ),  $133.0$ ,  $132.9$ ,  $131.7$ ,  $131.5$  ( $2\text{C}_{IV.Ar}$ ),  $130.7$ ,  $130.6$ ,  $129.2$ ,  $129.1$  ( $2\text{CH}_{Ar}$ ),  $128.7$ ,  $128.5$  ( $\text{C}_{IV.Ar}$ ),  $128.1$ ,  $127.9$ ,  $127.7$ ,  $127.3$ ,  $126.7$ ,  $126.6$ ,  $124.9$ ,  $124.8$ ,  $122.4$ ,  $121.7$  ( $5\text{CH}_{Ar}$ ),  $109.4$ ,  $109.3$  ( $C_e$ ),  $61.0$ ,  $59.7$  ( $C_4$ ),  $51.9$ ,  $51.7$  ( $C_2$ ),  $47.5$ ,  $47.5$  ( $C_6$ ),  $43.5$ ,  $43.4$  ( $C_a$ ),  $37.5$ ,  $37.3$  ( $C_c$ ),  $28.8$ ,  $28.46$  ( $C_b$ ),  $27.6$ ,  $27.6$  ( $C_1$ ),  $16.7$ ,  $16.5$  ( $C_7$ ),  $12.0$ ,  $12.0$  ( $C_h$ ); **HRMS (ESI)**:  $m/z$  494.27649  $[\text{M}+\text{H}]^+$  ( $\text{C}_{27}\text{H}_{36}\text{N}_5\text{O}_4$  requires 494.27618).



**S-Gly-(R/S)-Naphthalene-(Boc)-(L)-Lys(Boc)-tert-butylacetamide (C9a and C10a):**

Compounds **C9b** and **C10a** were prepared following general procedure A using **S-Gly-NH<sub>2</sub> (49b)** (43.5 mg, 0.15 mmol) as the amine, **1-naphthaldehyde** (20.4  $\mu$ L, 0.15 mmol) as aldehyde, **Boc-L-Lys(Boc)-OH** (51.9 mg, 0.15 mmol) as the acid and **tert-butyl isocyanide** (17.0  $\mu$ L, 0.15 mmol) in MeOH (0.3 mL). The reaction mixture was stirring at 50°C. After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), the diastereomers **C9a (dia1)** (47.9 mg, 37%) and **C10a (dia2)** (47.0 mg, 37%) were separated as white solids.

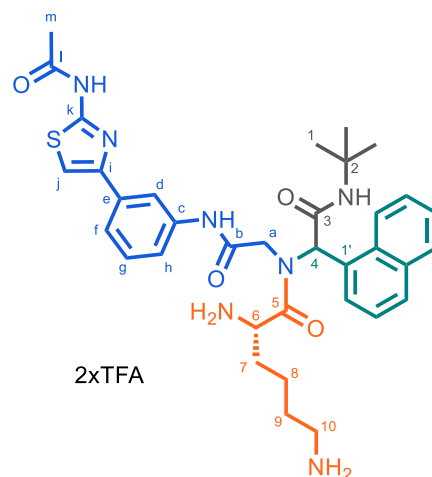
$R_{f, \text{dia1}} = 0.57$  and  $R_{f, \text{dia2}} = 0.45$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3);

**C9a – Dia1:** NMR data as a single conformer; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.13 - 7.72$  (m, 4H, 3CH<sub>Ar</sub> + H<sub>d</sub>), 7.72 – 7.40 (m, 4H, 3CH<sub>Ar</sub> + H<sub>f</sub>), 7.40 – 7.16 (m, 4H, CH<sub>Ar</sub> + H<sub>g-h</sub> + H<sub>j</sub>), 6.69 (s, 1H, H<sub>4</sub>), 4.48, 3.77 – 3.52 (d and m,  $J = 18.5$  Hz, 2H, H<sub>a</sub>), 4.27 (dd,  $J = 4.3, 8.8$  Hz, 1H, H<sub>6</sub>), 2.96 – 2.73 (m, 2H, H<sub>10</sub>), 2.21 (s, 3H, H<sub>m</sub>), 1.86 – 1.59 (m, 2H, H<sub>7</sub>), 1.57 – 1.23 (m, 31H, 6CH<sub>3, Boc</sub> + H<sub>1</sub> + H<sub>8-9</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta = 177.0$  (C<sub>5</sub>), 172.6 (C<sub>3</sub>), 170.9 (C<sub>i</sub>), 170.1 (C<sub>b</sub>), 159.4, 158.4, 158.2 (2CO<sub>Boc</sub> + C<sub>k</sub>), 150.8 (C<sub>i</sub>), 139.4, 136.5 (C<sub>c</sub> + C<sub>e</sub>), 135.4, 133.8, 131.8 (3C<sub>IV, Ar</sub>), 131.2, 130.0, 129.9, 128.9, 128.6, 127.4, 126.4, 124.3 (7CH<sub>Ar</sub> + C<sub>g</sub>), 123.2 (C<sub>f</sub>), 120.5 (C<sub>h</sub>), 118.6 (C<sub>d</sub>), 108.8 (C<sub>j</sub>), 80.7, 79.8 (2C<sub>IV, Boc</sub>), 61.9 (C<sub>4</sub>), 54.8, 53.2, 52.8, 52.7 (C<sub>2</sub> + C<sub>6</sub>), 50.3 (C<sub>a</sub>), 40.8 (C<sub>10</sub>), 32.0 (C<sub>7</sub>), 30.6 (C<sub>9</sub>), 28.9, 28.8 (6CH<sub>3, Boc</sub> + C<sub>1</sub>), 24.2 (C<sub>8</sub>), 22.6 (C<sub>m</sub>); **MS (ESI):**  $m/z$  858.07 [M+H]<sup>+</sup> (theoretical  $m/z$  858.42).

**C10a – Dia2:** NMR data as a mixture of 2 conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.40 - 8.25, 8.02 - 7.83, 7.82 - 7.73$  (3m, 3H, 3CH<sub>Ar</sub>), 8.06, 7.89 (2s, 1H, H<sub>d</sub>), 7.65 – 7.43 (m, 5H, 4CH<sub>Ar</sub> + H<sub>f</sub>), 7.43 – 7.21 (m, 3H, H<sub>g-h</sub> + H<sub>j</sub>), 6.91, 6.58 (2s, 1H, H<sub>4</sub>), 4.73, 4.52 (2dd,  $J_{\text{conf.1}} = 4.1, 9.5$  Hz and  $J_{\text{conf.2}} = 4.6, 8.7$  Hz, 1H, H<sub>6</sub>), 4.10, 3.84 – 3.64, 3.25 (2d and m,  $J_{\text{conf.1}} = 18.2$  Hz and  $J_{\text{conf.2}} = 16.5$  Hz, 2H, H<sub>a</sub>), 3.11 – 2.92 (m, 2H, H<sub>10</sub>), 2.22, 2.21 (2s, 3H, H<sub>m</sub>), 1.90 – 1.64 (m, 2H, H<sub>7</sub>), 1.64 – 1.22 (m, 31H, 6CH<sub>3, Boc</sub> + H<sub>1</sub> + H<sub>8-9</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta = 177.3, 176.2$  (C<sub>5</sub>), 173.1, 172.3 (C<sub>3</sub>), 170.8 (C<sub>i</sub>), 170.3, 169.1 (C<sub>b</sub>), 159.4, 159.3, 158.5, 158.4, 157.5 (2CO<sub>Boc</sub> + C<sub>k</sub>), 153.4, 150.9 (C<sub>i</sub>), 139.8, 139.6, 136.6, 136.4 (C<sub>c</sub> + C<sub>e</sub>), 135.5, 135.4, 133.7, 133.5, 132.1, 131.7 (3C<sub>IV, Ar</sub>), 131.5, 131.2, 130.1, 130.0, 129.8, 128.9, 128.7, 128.4, 127.7, 127.5,

126.4, 126.4, 124.8, 124.1 ( $7\text{CH}_{Ar} + C_g$ ), 123.1, 123.0 ( $C_f$ ), 121.0, 120.2 ( $C_h$ ), 119.1, 118.4 ( $C_d$ ), 108.8 ( $C_j$ ), 80.9, 80.5, 79.8 ( $2C_{IV.Boc}$ ), 63.6, 62.0 ( $C_4$ ), 53.3, 52.9, 52.7, 52.6 ( $C_2 + C_6$ ), 49.9 ( $C_a$ ), 41.0, 41.0 ( $C_{10}$ ), 33.4, 32.5 ( $C_7$ ), 31.0, 30.6 ( $C_9$ ), 29.0, 28.9, 28.8, 28.8, 28.7 ( $6\text{CH}_{3.Boc} + C_1$ ), 24.6, 24.0 ( $C_8$ ), 22.6 ( $C_m$ ); **MS (ESI)**:  $m/z$  858.00  $[\text{M}+\text{H}]^+$  (theoretical  $m/z$  858.42).

**S-Gly-(R/S)-Naphthalene-L-Lys-tert-butylacetamide (C9b and C10b):**



**Dia1 (C9b)**: General procedure B was employed for the deprotection of compound **C9a** (30.0 mg, 0.035 mmol) in 1 mL of  $\text{CH}_2\text{Cl}_2$  and in the presence of 50 equiv. of TFA. After filtration, a diastereomer **C9b** was isolated as a white powder (24.0 mg, 77%).

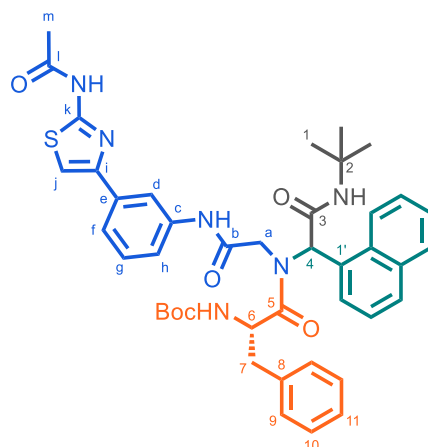
**HPLC**:  $\text{rt}_{\text{dia1}} = 12.34$  min (method B);  $^1\text{H NMR}$  (400 MHz,  $\text{D}_2\text{O}$ )  $\delta = 7.80$  (d,  $J = 8.6$  Hz, 1H,  $\text{CH}_{Ar}$ ), 7.64 – 7.24 (m, 5H,  $5\text{CH}_{Ar}$ ), 7.24 – 7.12 (m, 2H,  $\text{CH}_{Ar} + \text{H}_f$ ), 7.11 – 7.00 (m, 1H,  $\text{H}_g$ ), 6.92 – 6.79 (m, 2H,  $\text{H}_j + \text{H}_4$ ), 6.68 (br, 1H,  $\text{H}_d$ ), 6.56 (d,  $J = 7.3$  Hz, 1H,  $\text{H}_h$ ), 4.64 – 4.50 (m, 1H,  $\text{H}_6$ ), 4.34 – 4.12 (m, 2H,  $\text{H}_a$ ), 3.02 (t,  $J = 7.5$  Hz, 2H,  $\text{H}_{10}$ ), 2.17 (s, 3H,  $\text{H}_m$ ), 2.10 – 1.93 (m, 2H,  $\text{H}_7$ ), 1.81 – 1.68 (m, 2H,  $\text{H}_9$ ), 1.66 – 1.44 (m, 2H,  $\text{H}_8$ ), 1.33 (s, 9H,  $\text{H}_1$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{D}_2\text{O}$ )  $\delta = 171.4$ , 171.0, 170.1 ( $C_l + C_3 + C_5$ ), 166.5 ( $C_b$ ), 158.7 ( $C_k$ ), 147.3 ( $C_i$ ), 136.0, 133.4 ( $C_c + C_e$ ), 133.3, 131.7 ( $2C_{IV.Ar}$ ), 130.7 ( $\text{CH}_{Ar}$ ), 129.1, 128.9, 128.6, 127.2, 126.4, 125.3 ( $C_{IV.Ar} + 5\text{CH}_{Ar} + C_g$ ), 122.9 ( $\text{CH}_{Ar}$ ), 122.6 ( $C_f$ ), 120.4 ( $C_h$ ), 118.2 ( $C_d$ ), 109.2 ( $C_j$ ), 58.4 ( $C_4$ ), 51.8 ( $C_2$ ), 51.3 ( $C_6$ ), 48.6 ( $C_a$ ), 39.0 ( $C_{10}$ ), 29.8 ( $C_7$ ), 27.6 ( $C_1$ ), 26.5 ( $C_9$ ), 22.1 ( $C_m$ ), 20.9 ( $C_8$ ); **MS (ESI)**:  $m/z$  658.13  $[\text{M}+\text{H}]^+$  (theoretical  $m/z$  658.32).

**Dia2 (C10b)**: General procedure B was employed for the deprotection of compound **C10a** (30.0 mg, 0.035 mmol) in 1 mL of  $\text{CH}_2\text{Cl}_2$  and in the presence of 50 equiv. of TFA. After filtration, a diastereomer **C10b** was isolated as a white powder (21.9 mg, 71%).

**HPLC**:  $\text{rt}_{\text{dia2}} = 12.68$  min (method B);  $^1\text{H NMR}$  (400 MHz,  $\text{D}_2\text{O}$ )  $\delta = 7.89$  – 7.72 (m, 1H,  $\text{CH}_{Ar}$ ), 7.72 – 7.24 (m, 6H,  $6\text{CH}_{Ar}$ ), 7.23 – 6.99 (m, 2H,  $\text{H}_f + \text{H}_g$ ), 6.98 – 6.85 (m, 1H,  $\text{H}_j$ ), 6.80 – 6.61 (m, 3H,  $\text{H}_d + \text{H}_h + \text{H}_4$ ), 4.53 (br s, 1H,  $\text{H}_6$ ), 4.46 – 4.10 (m, 2H,  $\text{H}_a$ ), 2.96 (t,  $J = 7.5$  Hz, 2H,  $\text{H}_{10}$ ), 2.30 – 1.91 (m, 5H,  $\text{H}_m + \text{H}_7$ ), 1.78 – 1.38 (m, 4H,  $\text{H}_{8-9}$ ), 1.30 (s, 9H,  $\text{H}_1$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{D}_2\text{O}$ )  $\delta = 171.7$ , 171.4, 170.8 ( $C_l + C_3 + C_5$ ), 166.3 ( $C_b$ ), 158.8 ( $C_k$ ), 147.4 ( $C_i$ ), 136.4, 133.4, 133.3, 131.7 ( $2C_{IV.Ar} + C_c + C_e$ ), 130.7 ( $\text{CH}_{Ar}$ ), 129.2, 129.0, 128.8,

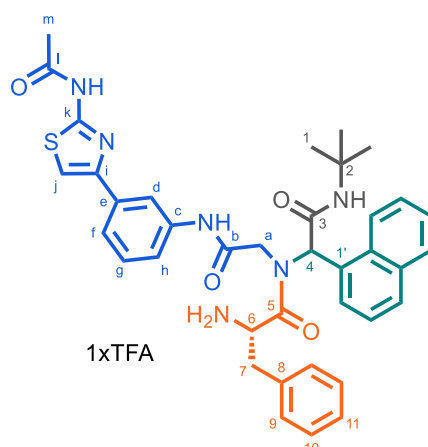
127.3, 126.5, 125.2 ( $C_{IV.Ar} + 5\underline{C}_{HAr} + C_g$ ), 122.7, 122.6 ( $\underline{C}_{HAr} + C_f$ ), 120.4 ( $C_h$ ), 118.3 ( $C_d$ ), 109.3 ( $C_j$ ), 59.3 ( $C_4$ ), 51.9 ( $C_2$ ), 51.3 ( $C_6$ ), 49.3 ( $C_a$ ), 38.9 ( $C_{10}$ ), 30.3 ( $C_7$ ), 27.6 ( $C_1$ ), 26.3 ( $C_9$ ), 22.1 ( $C_m$ ), 21.3 ( $C_8$ ); **MS (ESI):**  $m/z$  658.13  $[M+H]^+$  (theoretical  $m/z$  658.32).

**S-Gly-(R/S)-Naphthalene-(Boc)-L-Phe-*tert*-butylacetamide (C5a):**



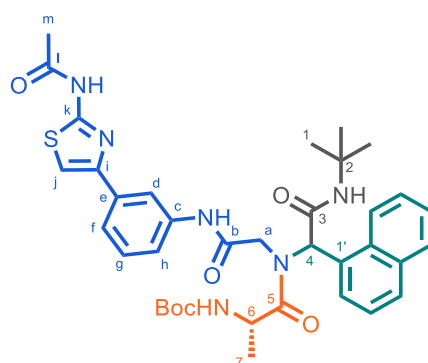
Compound **C5a** was prepared following general procedure A procedure using the TFA salt of **S-Gly-NH<sub>2</sub> (49a)** (29.1 mg, 0.10 mmol) as the amine, **1-naphthaldehyde** (13.6  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-Phe-OH** (26.5 mg, 0.10 mmol) as the acid and **tert-butyl isocyanide** (11.3  $\mu$ L, 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography ( $CH_2Cl_2$ /MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C5a** was obtained as a yellow solid (39.1 mg, 51%).

$R_{f,dia1+2} = 0.36$  ( $CH_2Cl_2$ -MeOH 96:4); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz,  $CD_3OD$ )  $\delta = 8.32 - 7.74$  (m, 4H,  $3\underline{C}_{HAr} + H_d$ ),  $7.74 - 6.96$  (m, 13H,  $9\underline{C}_{HAr} + H_{f-h} + H_j$ ), 6.89, 6.67, 6.41 (3s, 1H,  $H_4$ ), 5.19 - 5.07, 4.91 - 4.74, 4.61 - 4.44 (3m, 1H,  $H_6$ ), 3.97 - 3.70, 3.67 - 3.34 (2m, 2H,  $H_a$ ), 3.38 - 2.81 (m, 2H,  $H_7$ ), 2.20, 2.19 (2s, 3H,  $H_m$ ), 1.68 - 0.94 (m, 18H,  $3\underline{C}_{H_3.Boc} + H_1$ ); **<sup>13</sup>C NMR** (101 MHz,  $CD_3OD$ )  $\delta = 176.4, 176.2, 175.4$  ( $C_5$ ), 172.3 ( $C_3$ ), 170.9 ( $C_i$ ), 170.3, 169.1 ( $C_b$ ), 164.8, 159.4, 159.3, 158.1, 157.9, 157.0 ( $CO_{Boc} + C_k$ ), 150.8 ( $C_i$ ), 139.9, 139.7, 139.5, 138.4, 138.3, 138.2 ( $C_c + C_e$ ), 136.6, 136.5, 136.4, 135.5, 135.4, 135.3, 133.7, 133.5, 131.8, 131.6, 131.3, 131.2 ( $4C_{IV.Ar} + \underline{C}_{HAr}$ ), 130.8, 130.6, 130.4, 130.4, 130.1, 130.0, 129.9, 129.8, 129.6, 129.4, 129.4, 128.9, 128.7, 128.6, 127.9, 127.7, 127.6, 127.5, 126.6, 126.5, 126.3 ( $10\underline{C}_{HAr} + C_g$ ), 124.5, 124.1 ( $\underline{C}_{HAr}$ ), 123.3, 123.1 ( $C_f$ ), 120.9, 120.6, 120.4 ( $C_h$ ), 119.0, 118.7, 118.5 ( $C_d$ ), 108.8, 108.7 ( $C_j$ ), 81.0, 80.8, 80.5 ( $C_{IV.Boc}$ ), 62.3 ( $C_4$ ), 54.8, 54.8, 54.2, 53.4, 53.2, 52.7, 52.6 ( $C_2 + C_6$ ),  $C_a$  (signal under the  $CD_3OD$  peak), 40.3, 38.5, 38.4 ( $C_7$ ), 28.9, 28.9, 28.7, 28.7, 28.6, 28.1 ( $3\underline{C}_{H_3.Boc} + C_1$ ), 22.60 ( $C_m$ ); **MS (ESI):**  $m/z$  777.00  $[M+H]^+$  (theoretical  $m/z$  777.34).

**S-Gly-(R/S)-Naphthalene-L-Phe-*tert*-butylacetamide (C5b):**

General procedure B was employed for the deprotection of compound **C5a** (30.3 mg, 0.039 mmol) in 1 mL of  $\text{CH}_2\text{Cl}_2$  and in the presence of 50 equiv. of TFA. After filtration, the crude mixture was purified by semipreparative HPLC (method B) leading to the mixture of diastereomers **C5b** as a white powder (10.6 mg, 34%).

**HPLC:**  $\text{rt}_{\text{dia}1} = 15.92$  min and  $\text{rt}_{\text{dia}2} = 16.23$  min (method B); **NMR** data as a mixture of 2 diastereomers;  $^1\text{H NMR}$  (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 8.03 - 7.40$ ,  $7.39 - 7.10$  (2m, 17H,  $17\text{CH}_{\text{Ar}}$ ),  $6.98$ ,  $6.79$  (2s, 1H,  $\text{H}_4$ ),  $4.65$ ,  $4.59 - 4.49$  (t and m,  $J = 6.4$  Hz, 1H,  $\text{H}_6$ ),  $4.32$ ,  $4.09 - 3.73$  (d and m,  $J = 17.8$  Hz, 2H,  $\text{H}_a$ ),  $3.53 - 3.08$  (m, 2H,  $\text{H}_7$ ),  $2.22$ ,  $2.22$  (2s, 3H,  $\text{H}_m$ ),  $1.44$ ,  $1.40$  (2s, 9H,  $\text{H}_1$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 172.0$ ,  $171.0$ ,  $170.9$  ( $\text{C}_3 + \text{C}_5$ ),  $168.7$ ,  $168.7$ ,  $168.3$ ,  $168.3$  ( $\text{C}_b + \text{C}_l$ ),  $159.5$  ( $\text{C}_k$ ),  $150.7$  ( $\text{C}_i$ ),  $139.5$ ,  $138.2$  ( $\text{C}_c + \text{C}_e$ ),  $136.5$ ,  $135.6$ ,  $134.9$ ,  $133.5$ ,  $131.5$ ,  $131.5$ ,  $131.4$  ( $4\text{C}_{\text{IV.Ar}} + \text{CH}_{\text{Ar}}$ ),  $131.1$ ,  $131.0$ ,  $130.9$ ,  $130.6$ ,  $130.2$ ,  $130.1$ ,  $130.0$ ,  $129.8$ ,  $129.7$ ,  $129.3$ ,  $128.9$ ,  $128.3$ ,  $128.2$ ,  $128.0$ ,  $127.5$ ,  $127.1$ ,  $126.3$ ,  $126.1$ ,  $125.0$  ( $10\text{CH}_{\text{Ar}} + \text{C}_g$ ),  $124.8$ ,  $124.5$  ( $\text{CH}_{\text{Ar}}$ ),  $123.3$ ,  $123.2$  ( $\text{C}_f$ ),  $120.8$ ,  $120.6$  ( $\text{C}_h$ ),  $118.7$ ,  $118.6$  ( $\text{C}_d$ ),  $110.4$  ( $\text{C}_j$ ),  $57.9$ ,  $57.6$  ( $\text{C}_4$ ),  $54.2$ ,  $52.7$ ,  $52.5$  ( $\text{C}_2 + \text{C}_6$ ),  $49.78$  ( $\text{C}_a$ ),  $38.8$ ,  $37.7$  ( $\text{C}_7$ ),  $28.8$ ,  $28.8$  ( $\text{C}_1$ ),  $22.7$  ( $\text{C}_m$ ); **MS (ESI):**  $m/z$  677.00 [ $\text{M}+\text{H}$ ] $^+$  (theoretical  $m/z$  677.29).

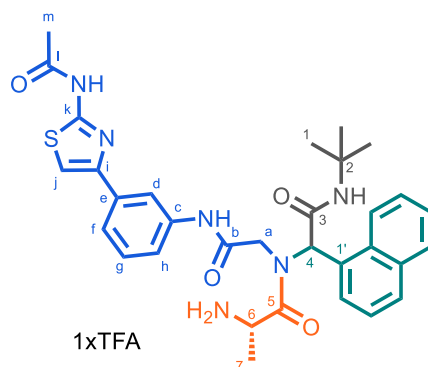
**S-Gly-(R/S)-Naphthalene-(Boc)-L-Ala-*tert*-butylacetamide (C6a):**

Compound **C6a** was prepared following general procedure A procedure using the TFA salt of **S-Gly-NH<sub>2</sub>** (**49a**) (43.5 mg, 0.15 mmol) as the amine, **1-naphthaldehyde** (20.4  $\mu\text{L}$ , 0.15 mmol) as aldehyde, **Boc-L-**

**Ala-OH** (28.4 mg, 0.15 mmol) as the acid and **tert-butyl isocyanide** (17.0  $\mu$ l, 0.15 mmol) in MeOH (0.3 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 98:2, v/v), a mixture of diastereomers **C6a** was obtained as a white solid (48.5 mg, 46%).

$R_{f, \text{dia}1+2} = 0.29$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 98:2); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.48, 8.13 - 7.69$  (d and m,  $J = 8.6$  Hz, 4H, 3CH<sub>Ar</sub> + H<sub>d</sub>), 7.64 – 7.34 (m, 5H, 4CH<sub>Ar</sub> + H<sub>f</sub>), 7.64 – 7.34, 7.35 – 7.12 (2m, 1H, H<sub>h</sub>), 7.35 – 7.12 (m, 2H, H<sub>g</sub> + H<sub>j</sub>), 6.97, 6.75, 6.58 (3s, 1H, H<sub>4</sub>), 4.85 – 4.76, 4.68 – 4.57, 4.41 – 4.30 (3m, 1H, H<sub>6</sub>), 4.50, 4.10, 3.86, 3.20 (4d,  $J_{\text{dia}1} = 18.6$  Hz and  $J_{\text{dia}2} = 16.6$  Hz, 2H, H<sub>a</sub>), 2.19, 2.18 (s, 3H, H<sub>m</sub>), 1.54 – 1.29 (m, 21H, 3CH<sub>3, Boc</sub> + H<sub>1</sub> + H<sub>7</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 178.4, 177.7$  (C<sub>5</sub>), 173.2, 172.4 (C<sub>3</sub>), 170.8, 170.3, 169.1 (C<sub>l</sub> + C<sub>b</sub>), 159.3, 158.0 (CO<sub>Boc</sub> + C<sub>k</sub>), 150.9, 150.8 (C<sub>i</sub>), 139.8, 139.6, 136.5, 136.3 (C<sub>c</sub> + C<sub>e</sub>), 135.5, 135.3, 133.8, 133.6, 131.9, 131.8, 131.5, 131.2 (3C<sub>IV, Ar</sub> + CH<sub>Ar</sub>), 130.0, 129.9, 129.7, 128.9, 128.6, 128.5, 127.7, 127.3, 126.4, 126.3, 125.1, 124.4, 124.0 (6CH<sub>Ar</sub> + C<sub>g</sub>), 123.2, 123.0 (C<sub>f</sub>), 120.9, 120.6, 120.3 (C<sub>h</sub>), 119.0, 118.7, 118.5 (C<sub>d</sub>), 108.8 (C<sub>j</sub>), 80.8, 80.7, 80.5 (C<sub>IV, Boc</sub>), 63.7, 62.0 (C<sub>4</sub>), 53.3, 53.2, 52.7, 52.6 (C<sub>2</sub>), 50.4, 49.9, 49.7, 49.3, 49.3, 49.1 (C<sub>a</sub> + C<sub>6</sub>), 28.9, 28.8, 28.8, 28.7 (3CH<sub>3, Boc</sub> + C<sub>1</sub>), 22.6 (C<sub>m</sub>), 18.6, 17.8, 17.2 (C<sub>7</sub>); **MS (ESI)**:  $m/z$  701.00 [M+H]<sup>+</sup> (theoretical  $m/z$  701.31).

### S-Gly-(R/S)-Naphthalene-L-Ala-tert-butylacetamide (**C6b**):

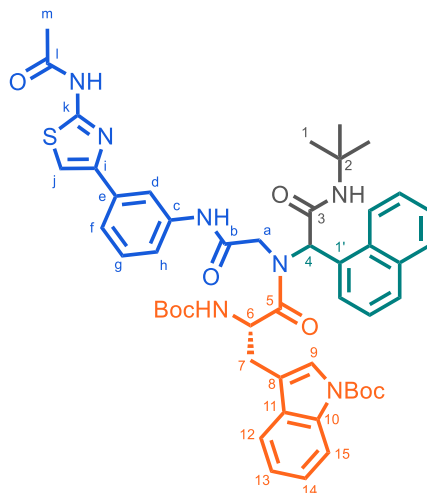


General procedure B was employed for the deprotection of compound **C6b** (30.0 mg, 0.043 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C6b** was isolated as a yellow-pale powder (28.2 mg, 92%).

**HPLC**:  $rt_{\text{dia}1} = 14.15$  min and  $rt_{\text{dia}2} = 14.65$  min (method B); NMR data as a mixture of 2 diastereomers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.00 - 7.71, 7.69 - 7.35$  (2m, 10H, 7CH<sub>Ar</sub> + H<sub>d</sub> + H<sub>f-g</sub>), 7.32 – 7.09, 6.92 – 6.81 (2m, 1H, H<sub>h</sub>), 7.22, 7.14 (2s, 1H, H<sub>j</sub>), 6.87, 6.78 (2s, 1H, H<sub>4</sub>), 4.50, 4.36 (2q,  $J = 6.8$  Hz, 1H, H<sub>6</sub>), 4.31, 4.03 – 3.96 (d and m,  $J = 17.9$  Hz, 2H, H<sub>a</sub>), 2.22, 2.22 (2s, 3H, H<sub>m</sub>), 1.58, 1.57 (d,  $J = 6.8$  Hz, 3H, H<sub>7</sub>), 1.42, 1.41 (2s, 9H, H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 173.1, 172.6$  (C<sub>5</sub>), 171.9, 171.8 (C<sub>3</sub>), 170.9 (C<sub>l</sub>), 168.3 (C<sub>b</sub>), 159.4 (C<sub>k</sub>), 150.8, 150.7 (C<sub>i</sub>), 139.5, 138.8, 136.5, 136.3 (C<sub>c</sub> + C<sub>e</sub>), 135.5, 135.3, 133.9, 133.5 (2C<sub>IV, Ar</sub>), 131.8, 131.5, 131.4 (C<sub>IV, Ar</sub> + CH<sub>Ar</sub>), 130.1, 129.9, 129.8, 129.7, 129.3, 129.1, 128.3, 128.1, 127.5, 127.4, 126.4, 126.2, 124.8, 124.1 (6CH<sub>Ar</sub> + C<sub>g</sub>), 123.3, 123.2 (C<sub>f</sub>), 120.7, 120.6 (C<sub>h</sub>), 118.7, 118.6 (C<sub>d</sub>), 108.8,

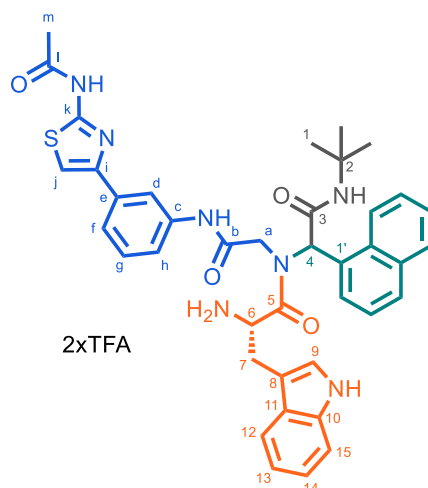
108.8 (C<sub>j</sub>), 61.1 (C<sub>4</sub>), 52.7, 52.6 (C<sub>2</sub>), 49.5, 49.9 (C<sub>a</sub>), 49.1, 48.9 (C<sub>6</sub>), 28.8, 28.7 (C<sub>1</sub>), 22.6 (C<sub>m</sub>), 17.4, 16.9 (C<sub>7</sub>); **MS (ESI)**:  $m/z$  600.87 [M+H]<sup>+</sup> (theoretical  $m/z$  601.26).

**S-Gly-(R/S)-Naphthalene-(Boc)-L-Trp(Boc)-tert-butylacetamide (C7a):**



Compound **C6a** was prepared following general procedure A using **S-Gly-NH<sub>2</sub> (49b)** (29.1 mg, 0.10 mmol) as the amine, **1-naphthaldehyde** (13.6  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-Trp(Boc)-OH** (40.5 mg, 0.10 mmol) as the acid and **tert-butyl isocyanide** (11.3  $\mu$ L, 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 96:4, v/v), a mixture of diastereomers **C7a** was obtained as a white solid (66.8 mg, 73%).

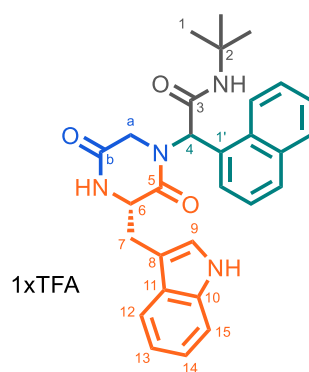
**R<sub>f,dia1+2</sub>** = 0.25 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 96:4); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.25 – 6.88 (m, 17H, 17CH<sub>Ar</sub>), 6.87, 6.61, 6.41 (3s, 1H, H<sub>4</sub>), 5.23 – 5.15, 5.01 – 4.84, 4.72 – 4.62 (3m, 1H, H<sub>6</sub>), 4.12 – 3.80, 3.52 – 3.38 (d and 2m,  $J$  = 18.7 Hz, 2H, H<sub>a</sub>), 3.37 – 3.01 (m, 2H, H<sub>7</sub>), 2.19 (s, 3H, H<sub>m</sub>), 1.74 – 1.07, 0.94 – 0.78 (2m, 21H, 6CH<sub>3,Boc</sub> + H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 176.0, 175.3 (C<sub>5</sub>), 172.9, 172.8, 172.4 (C<sub>3</sub>), 170.9 (C<sub>i</sub>), 170.1 (C<sub>b</sub>), 159.4, 159.2, 158.7, 157.2 (2CO<sub>Boc</sub> + C<sub>k</sub>), 150.7 (C<sub>i</sub>), 139.7, 139.5, 136.8, 136.5, 136.4, 136.3, 135.6, 135.4, 135.3, 133.7 (3C<sub>IV,Ar</sub> + C<sub>c</sub> + C<sub>e</sub>), 131.9, 131.8, 131.6, 131.5, 131.3, 131.2 (2C<sub>IV,Ar</sub> + CH<sub>Ar</sub>), 130.1, 130.0, 129.8, 128.8, 128.7, 128.5, 127.7, 127.5, 127.3, 126.5, 126.3, 125.6, 125.5, 125.2, 124.0, 123.8, 123.7, 123.6, 123.2, 123.2, 123.1, 120.8, 120.4, 120.3, 120.3, 120.3, 120.2, 119.9 (9CH<sub>Ar</sub> + C<sub>f-h</sub>), 118.9, 118.4, 117.8, 117.4, 117.2, 116.1, 116.1, 114.6, 111.8 (C<sub>d</sub> + C<sub>8</sub> + C<sub>12</sub> + C<sub>15</sub>), 108.8, 108.8 (C<sub>j</sub>), 84.8, 84.7, 84.5, 81.1, 80.8, 80.6 (2C<sub>IV,Boc</sub>), 63.7, 62.8, 62.3 (C<sub>4</sub>), 53.2, 52.8, 52.7, 52.5, 52.0 (C<sub>2</sub> + C<sub>6</sub>), 49.9, 49.9, 49.3, 49.2 (C<sub>a</sub>), 29.6, 29.0, 28.9, 28.8, 28.7, 28.7, 28.5, 28.4, 27.8 (6CH<sub>3,Boc</sub> + C<sub>1</sub> + C<sub>7</sub>), 22.6 (C<sub>m</sub>); **MS (ESI)**:  $m/z$  914.07 [M+H]<sup>+</sup> (theoretical  $m/z$  914,39).

**S-Gly-(R/S)-Naphthalene-L-Trp-*tert*-butylacetamide (C7b):**

General procedure B was employed for the deprotection of compound **C7a** (38.7 mg, 0.042 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, the crude mixture was purified by semipreparative HPLC (method B) leading to the mixture of diastereomers **C7b** as a white powder (7.1 mg, 18%) and the cyclic product **C12** as white powder (7.2 mg, 30%).

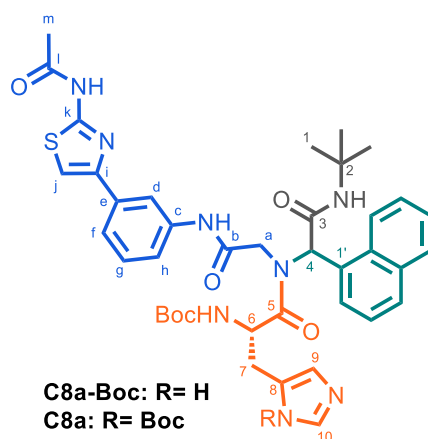
**C7b**: HPLC:  $rt_{dia1}$  = 16.39 min and  $rt_{dia2}$  = 16.57 min (method B); NMR data as a mixture of 2 diastereomers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.30 – 8.20, 8.03 – 6.86 (m, 16H, 17CH<sub>Ar</sub>), 6.66 (s, 1H, H<sub>4</sub>), 4.68 – 4.61, 4.57 (m and t,  $J$  = 7.4 Hz, 1H, H<sub>6</sub>), 4.27, 3.94 – 3.70 (d and m,  $J$  = 18.0 Hz, 2H, H<sub>a</sub>), 3.62 – 3.52, 3.42 – 3.23 (2m, 2H, H<sub>7</sub>), 2.21, 2.21 (2s, 3H, H<sub>m</sub>), 1.45, 1.40 (2s, 9H, H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 172.4, 172.4, 172.0, 171.9, 170.9, 168.4 (C<sub>b</sub> + C<sub>l</sub> + C<sub>3</sub> + C<sub>5</sub>), 159.5, 159.5 (C<sub>k</sub>), 151.7, 150.8 (C<sub>i</sub>), 139.5, 138.3, 136.6, 135.4, 134.2, 133.8, 133.5, 131.5, 131.4, 131.0, 130.0, 130.0, 129.8, 129.7, 129.3, 128.9, 128.5, 128.3, 128.2, 128.2, 128.1, 127.6, 127.5, 127.4, 126.3, 126.2, 126.2, 125.9, 125.8, 124.1, 123.4, 123.3, 123.1, 122.9, 120.7, 120.6, 120.4, 119.7, 119.3, 119.0, 118.7, 118.6 (7C<sub>IV,Ar</sub> + 15CH<sub>Ar</sub>), 112.8, 112.8, 108.9, 108.8, 107.8, 107.3 (C<sub>j</sub> + C<sub>8</sub> + C<sub>15</sub>), 61.8, 61.4 (C<sub>4</sub>), 53.2, 53.1, 52.8, 52.6 (C<sub>2</sub> + C<sub>6</sub>), 49.0 (C<sub>a</sub>) 28.9, 28.8 (C<sub>1</sub> + C<sub>7</sub>), 22.6 (C<sub>m</sub>); MS (ESI):  $m/z$  716,07 [M+H]<sup>+</sup> (theoretical  $m/z$  716,30).

**C12**: 2-((S)-3-((1*H*-indol-3-yl)methyl)-2,5-dioxopiperazin-1-yl)-*N*-(*tert*-butyl)-2-(naphthalen-1-yl)acetamide:



**HPLC:**  $rt_{dia1}$  = 17.08 min and  $rt_{dia2}$  = 17.69 min (method B); NMR data as a mixture of 2 diastereomers;  **$^1H$  NMR** (400 MHz,  $CD_3OD$ )  $\delta$  = 8.08 – 7.80, 7.57 – 7.39 (2m, 6H,  $6CH_{Ar}$ ), 7.64, 7.58 (2d,  $J$  = 7.9 Hz, 1H,  $H_{12}$ ), 7.37 – 7.32 (2m, 1H,  $H_{15}$ ), 7.25 – 6.97 (m, 4H,  $CH_{Ar}$  +  $H_9$  +  $H_{13-14}$ ), 6.87, 6.46 (2s, 1H,  $H_4$ ), 4.47, 4.35 (2t,  $J_{dia1}$  = 4.5 Hz and  $J_{dia2}$  = 5.6 Hz, 1H,  $H_6$ ), 3.84, 3.75, 2.93, 2.41 (4d,  $J_{dia1}$  = 17.6 Hz and  $J_{dia2}$  = 17.0 Hz, 2H,  $H_{\alpha,dia2}$ ), 3.52 – 3.16 (m, 2H,  $H_7$ ), 1.40, 1.36 (2s, 9H,  $H_1$ );  **$^{13}C$  NMR** (101 MHz,  $CD_3OD$ )  $\delta$  = 171.4, 170.9 ( $C_5$ ), 169.5, 168.8, 167.9, 167.3 ( $C_b + C_3$ ), 138.1, 137.9 ( $C_{10}$ ), 135.5, 135.2, 133.3, 133.1 ( $2C_{IV,Ar}$ ), 131.3, 131.01, 131.0, 130.9, 130.1, 129.9, 128.9, 128.7, 128.6, 128.2, 127.5, 127.3, 126.1, 126.1 ( $C_{IV,Ar}$  +  $6CH_{Ar}$  +  $C_{11}$ ), 125.4, 125.4 ( $C_9$ ), 123.8, 123.6 ( $CH_{Ar}$ ), 122.5, 122.4, 120.2, 120.1, 119.8, 119.3 ( $C_{12-14}$ ), 112.4, 112.4 ( $C_{15}$ ), 109.3, 109.2 ( $C_8$ ), 59.2, 58.6 ( $C_4$ ), 58.1, 57.7 ( $C_6$ ), 52.5, 52.4 ( $C_2$ ), 48.3, 47.5 ( $C_a$ ), 30.9, 30.5 ( $C_7$ ), 28.9, 28.8 ( $C_1$ ); **HRMS (ESI):**  $m/z$  483.23911 [ $M+H$ ]<sup>+</sup> ( $C_{29}H_{31}N_4O_3$  requires 483.23907).

### S-Gly-(R/S)-Naphthalene-(Boc)-L-His(Boc)-tert-butylacetamide (**C8a**):



Compound **C8a** was prepared following general procedure A using **S-Gly-NH<sub>2</sub> (49b)** (29.1 mg, 0.10 mmol) as the amine, **1-naphthaldehyde** (13.6  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-His(Boc)-OH** (53.7 mg, 0.10 mmol) as the acid and **tert-butyl isocyanide** (11.3  $\mu$ L, 0.10 mmol) in MeOH (0.2 mL). The reaction mixture was stirred at 50°C. After purification by flash chromatography ( $CH_2Cl_2/MeOH$  gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C8a** (39.0 mg, 45%) and a mixture of diastereomers **C8a-Boc** (29.0 mg, 38%) were separated as white solids.

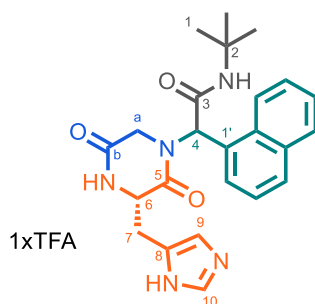
**(C8a)  $R_{f,dia1+2}$**  = 0.36 ( $CH_2Cl_2$ -MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers;  **$^1H$  NMR** (400 MHz,  $CD_3OD$ )  $\delta$  = 8.60 – 7.76 (m, 5H,  $3CH_{Ar}$  +  $H_d$  +  $H_{10}$ ), 7.75 – 7.09 (m, 9H,  $4CH_{Ar}$  +  $H_{f-h}$  +  $H_j$  +  $H_9$ ), 6.93, 6.66, 6.50 (3s, 1H,  $H_4$ ), 5.24 – 5.09, 5.01 – 4.66 (2m, 1H,  $H_6$ ), 4.60 – 4.44, 4.32 – 4.18, 3.94 – 3.40 (3m, 2H,  $H_a$ ), 3.23 – 2.80 (m, 2H,  $H_7$ ), 2.20, 2.20 (2s, 3H,  $H_m$ ), 1.65 – 1.11 (m, 27H,  $6CH_{3,Boc}$  +  $H_1$ );  **$^{13}C$  NMR** (101 MHz,  $CD_3OD$ )  $\delta$  = 175.8, 175.6, 175.1 ( $C_5$ ), 172.8, 172.4, 172.2, 172.2 ( $C_3$ ), 170.8 ( $C_i$ ), 170.1, 169.7, 169.2 ( $C_b$ ), 159.3, 158.1, 157.8, 157.01 ( $CO_{Boc} + C_k$ ), 150.9, 148.2, 148.1, 148.0 ( $CO_{Boc} + C_i$ ), 140.1, 140.1, 139.8, 139.7, 139.4, 139.3, 138.2, 136.5, 136.3, 135.5, 135.3, 133.8, 133.5 ( $4C_{IV,Ar}$  +  $C_{10}$ ), 131.9, 131.8, 131.5, 131.4, 131.3, 131.2, 130.1, 130.0, 129.9, 129.7, 128.9, 128.7, 128.5, 128.4, 127.7, 127.4, 127.4, 126.4, 126.4 ( $2C_{IV,Ar}$  +  $7CH_{Ar}$ ), 124.6, 124.2 ( $CH_{Ar}$ ), 123.9, 123.1, 123.0 ( $C_f$ ), 120.9, 120.61,



120.4 ( $C_h$ ), 119.0, 118.8, 118.6 ( $C_d$ ), 116.5, 116.3, 116.1 ( $C_g$ ), 108.8, 108.7 ( $C_j$ ), 87.0, 86.9, 81.0, 80.8, 80.5 ( $2C_{IV.Boc}$ ), 63.4, 62.2 ( $C_4$ ), 53.2, 52.7, 52.6, 52.3, 51.8 ( $C_2 + C_6$ ), 50.1, 49.9 ( $C_a$ ), 32.4, 31.1, 31.0 ( $C_7$ ), 29.0, 29.0, 28.9, 28.8, 28.7, 28.6, 28.0, 28.0 ( $6\text{CH}_{3.Boc} + C_1$ ), 22.6 ( $C_m$ ); **MS (ESI)**:  $m/z$  866.93 [ $M+H$ ]<sup>+</sup> (theoretical  $m/z$  867.39).

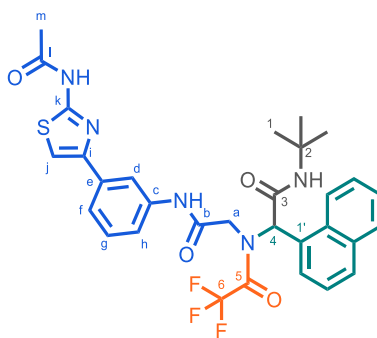
**(C8a-Boc)** NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.11 – 6.78 (m, 14H, 14CH<sub>Ar</sub>), 6.69, 6.51 (2s, 1H, H<sub>4</sub>), 5.16 – 4.71, 4.67 – 4.53 (2m, 1H, H<sub>6</sub>), 4.50 – 4.04, 3.84 – 3.40 (2m, 2H, H<sub>a</sub>), 3.27 – 2.84 (m, 2H, H<sub>7</sub>), 2.22 (s, 3H, H<sub>m</sub>), 1.57 – 1.00 (m, 18H, 3CH<sub>3.Boc</sub> + H<sub>1</sub>); **MS (ESI)**:  $m/z$  767.07 [ $M+H$ ]<sup>+</sup> (theoretical  $m/z$  767.33).

**2-((S)-3-((1H-indol-3-yl)methyl)-2,5-dioxopiperazin-1-yl)-N-(tert-butyl)-2-(naphthalen-1-yl)acetamide (C13):**



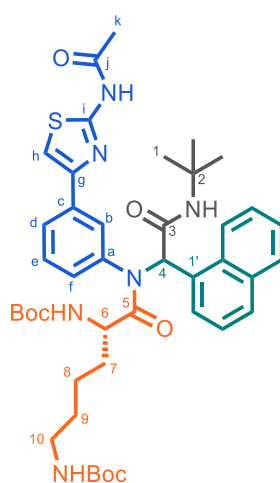
General procedure B was employed for the deprotection of compound **C8a** (61.7 mg, 0.071 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, the crude mixture was purified by semipreparative HPLC (method B) leading to the cyclic product **C13** as white powder (30.7 mg, 79%).

**HPLC**:  $rt_{dia1}$  = 10.22 min and  $rt_{dia2}$  = 10.48 min (method B); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.87 (s, 1H, H<sub>10</sub>), 7.99 – 7.92 (m, 2H, 2CH<sub>Ar</sub>), 7.86 – 7.80, 7.63 – 7.47 (2m, 5H, 5CH<sub>Ar</sub>), 7.40, 7.32 (2s, 1H, H<sub>9</sub>), 6.81, 6.80 (2s, 1H, H<sub>4</sub>), 4.52 (dd,  $J$  = 9.1, 4.8 Hz, 1H, H<sub>6</sub>), 4.19, 3.96, 2.94, 2.93 (4d,  $J$  = 17.0 Hz, 2H, H<sub>a</sub>), 3.51 – 3.41, 3.39 – 3.23 (2m, 2H, H<sub>7</sub>), 1.39, 1.37 (2s, 9H, H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 171.1, 171.1 ( $C_5$ ), 168.4, 168.0, 167.2, 167.2 ( $C_b + C_3$ ), 135.5, 135.4 ( $C_{IV.Ar}$ ), 135.2, 135.2 ( $C_{10}$ ), 133.3, 133.1, 131.3, 131.2, 131.1, 131.0 (3C<sub>IV.Ar</sub>), 130.3, 130.2, 130.0, 129.6, 129.0, 128.9, 128.5, 128.3, 127.6, 127.6, 126.3, 126.2, 123.7, 123.4 (7CH<sub>Ar</sub>), 119.2, 118.9 ( $C_9$ ), 58.9, 58.3 ( $C_4$ ), 55.6, 55.5 ( $C_6$ ), 52.6, 52.5 ( $C_2$ ), 48.3, 47.0 ( $C_a$ ), 29.8, 29.1 ( $C_7$ ), 28.9, 28.8 ( $C_1$ ); **HRMS (ESI)**:  $m/z$  434.21902 [ $M+H$ ]<sup>+</sup> ( $C_{24}H_{28}N_5O_3$  requires 434.21867).

**S-Gly-(R/S)-Naphthalene-Trifluoroacetic-*tert*-butylacetamide (C11):**

Compound **C11** was a by-product produced by the general procedure **A** using salt TFA of **S-NH<sub>2</sub> (49a)** (17.0 mg, 0.073 mmol) as the amine, **1-naphthaldehyde** (9.9  $\mu$ L, 0.073 mmol) as aldehyde, **Boc-L-Phe-OH** (26.5 mg, 0.10 mmol) as the acid and ***tert*-butyl isocyanide** (11.3  $\mu$ L, 0.10 mmol) in MeOH (0.2 mL). The reaction mixture was stirred at room temperature. After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 98:2, v/v), a mixture of enantiomers **C11** was obtained as a white solid (17.2 mg, 28%).

$R_f$  = 0.36 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); **HPLC**:  $t_r$  = 21.52 min (method B); NMR data as a mixture of 2 conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.00 – 7.79 (m, 4H, 3CH<sub>Ar</sub> + H<sub>d</sub>), 7.71 – 7.44 (m, 5H, 4CH<sub>Ar</sub> + H<sub>f</sub>), 7.33 – 7.13 (m, 3H, H<sub>g-h</sub> + H<sub>j</sub>), 6.59 (s, 1H, H<sub>4</sub>), 4.23, 3.89, 3.78, 3.55 (4d,  $J_{Conf.1}$  = 19.1 Hz and  $J_{Conf.2}$  = 16.2 Hz, 2H, H<sub>a</sub>), 2.21 (s, 3H, H<sub>m</sub>), 1.47, 1.44 (s, 9H, H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 171.7, 171.3, 170.9 (C<sub>1</sub> + C<sub>3</sub> + C<sub>5</sub>), 168.1 (C<sub>b</sub>), 159.4 (C<sub>k</sub>), 150.8 (C<sub>i</sub>), 139.4, 136.6 (C<sub>c</sub> + C<sub>e</sub>), 135.4, 133.8 (2C<sub>IV,Ar</sub>), 132.0, 130.2, 130.0, 129.6, 129.4, 128.9 (C<sub>IV,Ar</sub> + 4CH<sub>Ar</sub>), 127.8, 126.4, 123.8 (3CH<sub>Ar</sub>), 123.3 (C<sub>f</sub>), 120.5 (C<sub>h</sub>), 118.8 (C<sub>d</sub>), 108.9 (C<sub>j</sub>), 63.2 (C<sub>4</sub>), 53.0, 52.9 (C<sub>2</sub>), 50.5, 49.1 (C<sub>a</sub>), 28.8, 28.7 (C<sub>1</sub>), 22.6 (C<sub>m</sub>); **HRMS (ESI)**:  $m/z$  626.20502 [M+H]<sup>+</sup> (C<sub>31</sub>H<sub>31</sub>F<sub>3</sub>N<sub>5</sub>O<sub>4</sub>S requires 626.20434).

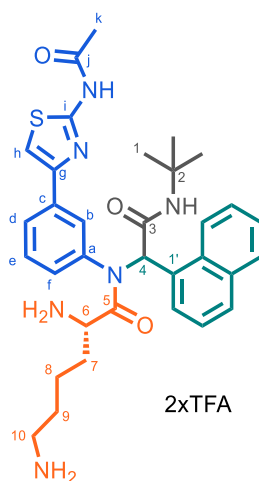
**S-(R/S)-Naphthalene-(Boc)-L-Lys(Boc)-*tert*-butylacetamide (C14a):**

Compound **C14a** was prepared following general procedure **A** using **S-NH<sub>2</sub> (45)** (17.0 mg, 0.073 mmol) as the amine, **1-naphthaldehyde** (9.90  $\mu$ L, 0.073 mmol) as aldehyde, **Boc-L-Lys(Boc)-OH** (25.3 mg,

0.073 mmol) as the acid and **tert-butyl isocyanide** (8.2  $\mu$ l, 0.073 mmol) in MeOH (0.2 mL). The reaction mixture was stirred at 50°C. After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 96:4, v/v), a mixture of diastereomers **C14a** was obtained as a white solid (30.0 mg, 51%).

**R<sub>f,dia1+2</sub>** = 0.25 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 96:4); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.55 – 8.34, 8.16 – 7.96, 7.94 – 6.80 (3m, 10H, 9CH<sub>Ar</sub> + H<sub>4</sub>), 6.73 – 5.86 (3m, 3H, CH<sub>Ar</sub>), 4.28 – 4.03 (m, 1H, H<sub>6</sub>), 2.95 – 2.69 (m, 2H, H<sub>10</sub>), 2.22, 2.17, 2.16 (3s, 3H, H<sub>k</sub>), 1.84 – 0.78 (m, 33H, 6CH<sub>3.Boc</sub> + H<sub>1</sub> + H<sub>7-9</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 175.3, 174.8 (C<sub>5</sub>), 171.7, 170.9, 170.8 (C<sub>j</sub> + C<sub>3</sub>), 159.2, 158.4, 158.0 (2CO<sub>Boc</sub> + C<sub>i</sub>), 150.3, 149.5 (C<sub>g</sub>), 139.8, 139.7, 139.7, 136.8, 136.1, 135.5, 135.2, 135.1, 134.9, 133.8, 133.7, 132.8, 132.2, 131.1 (5C<sub>IV.Ar</sub>), 130.3, 130.2, 130.1, 129.7, 129.5, 129.3, 129.0, 128.9, 128.2, 127.8, 127.1, 126.9, 126.5, 126.1, 126.1, 125.7, 125.1, 125.0, 124.3, 124.2 (11CH<sub>Ar</sub>), 109.6, 109.2, 108.8, 108.3 (C<sub>h</sub>), 80.3, 79.8 (2C<sub>IV.Boc</sub>), 62.4, 62.0 (C<sub>4</sub>), 53.4, 52.4, 52.3 (C<sub>2</sub> + C<sub>6</sub>), 40.9 (C<sub>10</sub>), 33.6, 32.1 (C<sub>7</sub>), 30.8, 30.1 (C<sub>9</sub>), 28.9, 28.8, 23.9 (6CH<sub>3.Boc</sub> + C<sub>1</sub>), 23.8, 23.8 (C<sub>8</sub>), 22.6, 22.6, 22.6 (C<sub>k</sub>); **MS (ESI)**: *m/z* 801.00 [M+H]<sup>+</sup> (theoretical *m/z* 801.40).

#### S-(R/S)-Naphthalene-L-Lys-*tert*-butylacetamide (**C14b**):

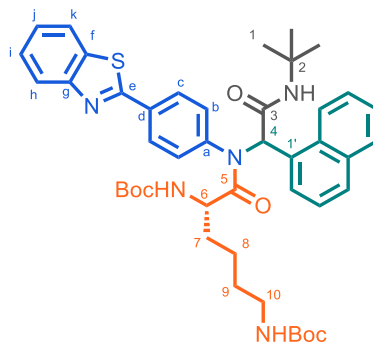


General procedure B was employed for the deprotection of compound **C14a** (24.4 mg, 0.031 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C14b** was isolated as a white powder (24.0 mg, 94%).

**HPLC**: *rt*<sub>dia1</sub> = 16.47 min and *rt*<sub>dia2</sub> = 16.91 min (method A); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.43, 8.18 – 8.03, 7.89 – 7.45 (d and 3m, *J* = 8.0 Hz, 7H, 7CH<sub>Ar</sub>), 7.43 – 7.08 (m, 2H, 2CH<sub>Ar</sub>), 7.43 – 7.08, 6.77, 6.71 (m and 2t, *J* = 7.8 Hz, 1H, H<sub>e</sub>), 7.43 – 7.08, 6.42, 5.96 (m and 2s, 1H, H<sub>h</sub>), 7.03, 6.96 (2d, *J*<sub>dia1</sub> = 8.9 Hz and *J*<sub>dia2</sub> = 4.9 Hz, 1H, H<sub>4</sub>), 6.50, 6.34, 5.91 (s and 2d, *J* = 7.8 Hz and *J* = 8.2 Hz, 1H, CH<sub>Ar</sub>), 4.08 – 4.02, 3.88 – 3.77 (2m, 1H, H<sub>6</sub>), 2.94 – 2.63 (m, 2H, H<sub>10</sub>), 2.23, 2.17 (2s, 3H, H<sub>k</sub>), 1.87 – 1.63 (m, 2H, H<sub>7</sub>), 1.63 – 1.18 (m, 13H, H<sub>1</sub> + H<sub>8-9</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 171.9, 171.7, 171.5, 171.4 (C<sub>5</sub>), 171.0, 170.9, 170.5, 169.9, 169.8 (C<sub>j</sub> + C<sub>3</sub>), 159.8, 159.5 (C<sub>i</sub>), 150.0, 149.7, 149.2, 149.1 (C<sub>g</sub>), 138.8, 138.7, 137.9, 137.3, 136.5, 136.2, 135.9, 135.2, 135.2, 134.9,

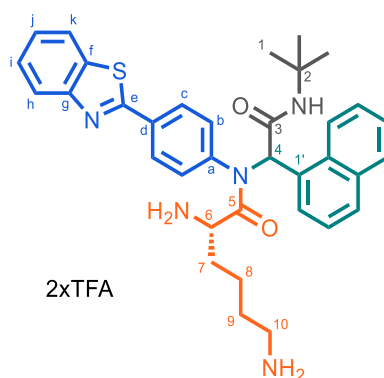
133.7, 133.7, 133.5 (4C<sub>IV,Ar</sub>), 132.1, 131.9, 131.6, 131.4, 131.3, 131.0, 130.7, 130.7, 130.5, 130.4, 130.4, 130.2, 129.9, 129.8, 129.7, 129.6, 129.3, 128.3, 128.1, 127.8, 127.8, 127.6, 127.5, 127.3, 127.3, 127.2, 127.0, 126.2, 126.1, 125.7 (C<sub>IV,Ar</sub> + 11C<sub>H,Ar</sub>), 124.7, 124.5, 124.2, 124.0 (C<sub>h</sub>), 62.9, 62.3 (C<sub>4</sub>), 52.8, 52.6, 52.5 (C<sub>2</sub> + C<sub>6</sub>), 40.1, 40.1 (C<sub>10</sub>), 31.6, 31.6, 30.7, 30.6 (C<sub>7</sub>), 28.9, 28.9, 28.8 (C<sub>1</sub>), 27.8, 27.8 (C<sub>9</sub>), 23.0, 22.6, 22.5, 22.0, 21.8 (C<sub>k</sub> + C<sub>8</sub>); **HRMS (ESI)**:  $m/z$  601.29529 [M+H]<sup>+</sup> (C<sub>33</sub>H<sub>41</sub>N<sub>6</sub>O<sub>3</sub>S requires 601.29554).

**4-(Benzothiazol-2-yl)phenyl-(R/S)-Naphthalene-(Boc)-L-Lys(Boc)-tert-butylacetamide (C15a):**



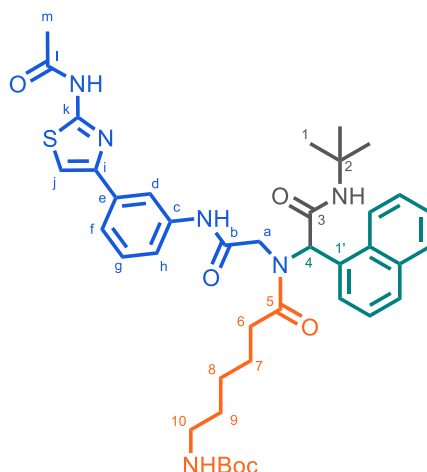
Compound **C15a** was prepared following general procedure A using **2-(4-aminophenyl)benzothiazole** (34.0 mg, 0.15 mmol) as the amine, **1-naphthaldehyde** (20.4  $\mu$ L, 0.15 mmol) as aldehyde, **Boc-L-Lys(Boc)-OH** (51.9 mg, 0.15 mmol) as the acid and **tert-butyl isocyanide** (17.0  $\mu$ L, 0.15 mmol) in MeOH (0.3 mL). The reaction mixture was stirred at 50°C. After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C15a** was obtained as a yellow-pale solid (68.1 mg, 58%).

$R_{f, \text{dia}1+2} = 0.44$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.25 - 7.87$ ,  $7.86 - 7.32$  (2m, 13H, 13C<sub>H,Ar</sub>),  $7.30 - 7.12$  (m, 2H, 2C<sub>H,Ar</sub>), 7.01, 6.98 (2s, 1H, H<sub>4</sub>), 6.89 - 6.80, 6.73 - 6.64, 6.43 - 6.31, 6.16 - 5.96 (4m, 2NH), 4.22 - 4.03 (m, 1H, H<sub>6</sub>), 2.96 - 2.75 (m, 2H, H<sub>10</sub>), 1.82 - 1.01 (m, 33H, 6C<sub>H<sub>3</sub>,Boc</sub> + H<sub>1</sub> + H<sub>7-9</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta = 175.2$ , 174.8 (C<sub>5</sub>), 171.6, 171.5 (C<sub>3</sub>), 168.5, 168.4 (C<sub>e</sub>), 158.4, 158.0, 157.4 (2CO<sub>Boc</sub>), 155.0, 154.9 (C<sub>g</sub>), 142.4 (C<sub>a</sub>), 136.2, 136.2 (C<sub>f</sub>), 135.2, 135.0, 134.2, 133.8, 133.7, 132.5, 132.4, 132.1 (3C<sub>IV,Ar</sub> + C<sub>d</sub>), 130.4, 130.2, 130.2, 129.8, 129.5, 128.2, 127.9, 127.7, 127.2, 127.1, 126.8, 126.1, 125.8, 124.9, 124.2, 123.9, 123.0 (15C<sub>H,Ar</sub>), 80.4, 79.7 (2C<sub>IV,Boc</sub>), 62.5, 61.9 (C<sub>4</sub>), 53.5, 53.3 (C<sub>6</sub>), 52.5, 52.3 (C<sub>2</sub>), 41.0 (C<sub>10</sub>), 33.4, 31.9 (C<sub>7</sub>), 30.3, 30.1 (C<sub>9</sub>), 28.9, 28.9, 28.8, 28.8, 28.7 (6C<sub>H<sub>3</sub>,Boc</sub> + H<sub>1</sub>), 23.9, 23.8 (C<sub>8</sub>); **MS (ESI)**:  $m/z$  794.07 [M+H]<sup>+</sup> (theoretical  $m/z$  794.39).

**4-(Benzothiazol-2-yl)phenyl-(R/S)-Naphthalene-L-Lys-tert-butylacetamide (C15b):**

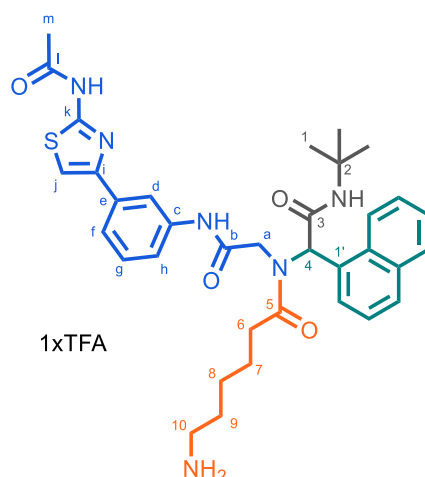
General procedure B was employed for the deprotection of compound **C15a** (53.2 mg, 0.067 mmol) in 1 mL of  $\text{CH}_2\text{Cl}_2$  and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C15b** was isolated as a yellow-pale powder (48.2 mg, 88%).

**HPLC:**  $rt_{\text{dia1}} = 18.44$  min and  $rt_{\text{dia2}} = 18.97$  min (method A); NMR data as a mixture of 2 diastereomers;  **$^1\text{H}$  NMR** (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 8.19 - 7.99$ ,  $7.98 - 7.84$ ,  $7.79 - 7.57$ ,  $7.54 - 7.30$ ,  $7.27 - 7.13$  (5m, 14H,  $14\text{CH}_{\text{Ar}}$ ),  $7.09$ ,  $7.02$  (2s, 1H,  $\text{H}_4$ ),  $6.79$ ,  $6.10$  (d and br s,  $J = 8.04$  Hz, 1H,  $\text{CH}_{\text{Ar}}$ ),  $4.13$ ,  $3.85$  (2t,  $J_{\text{dia1}} = 5.2$  Hz and  $J_{\text{dia2}} = 6.0$  Hz, 1H,  $\text{H}_6$ ),  $2.90$ ,  $2.81$  (2t,  $J_{\text{dia1}} = 7.3$  Hz and  $J_{\text{dia2}} = 7.5$  Hz, 2H,  $\text{H}_{10}$ ),  $1.89 - 1.63$  (m, 2H,  $\text{H}_7$ ),  $1.63 - 1.22$  (m, 4H,  $\text{H}_{8-9}$ ),  $1.43$ ,  $1.38$  (2s, 9H,  $\text{H}_1$ );  **$^{13}\text{C}$  NMR** (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 171.6$ ,  $171.3$  ( $\text{C}_3$ ),  $170.4$ ,  $169.8$  ( $\text{C}_5$ ),  $168.0$ ,  $167.9$  ( $\text{C}_e$ ),  $154.9$ ,  $154.8$  ( $\text{C}_g$ ),  $141.1$ ,  $141.1$  ( $\text{C}_a$ ),  $136.2$ ,  $136.1$  ( $\text{C}_f$ ),  $135.2$ ,  $134.9$ ,  $134.4$ ,  $133.6$ ,  $133.4$ ,  $132.8$ ,  $131.8$ ,  $131.5$ ,  $131.3$  ( $3\text{C}_{\text{IV,Ar}} + \text{C}_{\text{HAr}} + \text{C}_d$ ),  $130.8$ ,  $130.6$ ,  $130.3$ ,  $129.9$ ,  $129.8$ ,  $129.7$ ,  $129.3$ ,  $128.4$ ,  $128.3$ ,  $128.0$ ,  $127.9$ ,  $127.8$ ,  $127.4$ ,  $127.2$ ,  $127.0$ ,  $126.9$ ,  $126.1$ ,  $125.8$ ,  $124.5$ ,  $124.0$ ,  $123.9$ ,  $123.9$ ,  $123.0$ ,  $122.9$  ( $14\text{C}_{\text{HAr}}$ ),  $63.1$ ,  $62.3$  ( $\text{C}_4$ ),  $52.8$  ( $\text{C}_2$ ),  $52.6$ ,  $52.5$  ( $\text{C}_6$ ),  $40.1$  ( $\text{C}_{10}$ ),  $31.6$ ,  $30.5$  ( $\text{C}_7$ ),  $28.9$ ,  $28.8$  ( $\text{C}_1$ ),  $27.8$ ,  $27.8$  ( $\text{C}_9$ ),  $22.9$ ,  $21.8$  ( $\text{C}_8$ ); **HRMS (ESI):**  $m/z$  594.28992 [ $\text{M}+\text{H}$ ] $^+$  ( $\text{C}_{35}\text{H}_{40}\text{N}_5\text{O}_2\text{S}$  requires 594.28972).

**S-Gly-(R/S)-Naphthalene-(Boc)- $\epsilon$ -Ahx-*tert*-butylacetamide (C16a):**

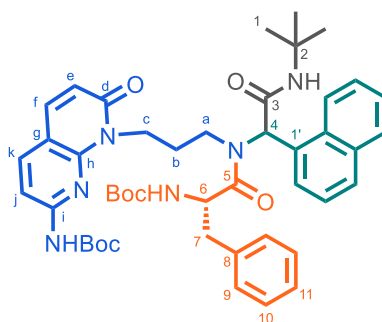
Compound **C16a** was prepared following general procedure A using **S-Gly-NH<sub>2</sub> (49b)** (40.6mg, 0.14 mmol) as the amine, **1-naphthaldehyde** (19.0  $\mu$ L, 0.14 mmol) as aldehyde, **Boc- $\epsilon$ -Ahx-OH** (32.4 mg, 0.14 mmol) as the acid and ***tert*-butyl isocyanide** (15.8  $\mu$ L, 0.14 mmol) in MeOH (0.2 mL). The reaction mixture was stirred at 50°C. After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of enantiomers **C16a** was obtained as a pale-yellow solid (66.8 mg, 73%).

$R_f$  = 0.30 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.05, 8.06 (2s, 1H, H<sub>d</sub>), 7.98 – 7.73 (m, 3H, 3CH<sub>Ar</sub>), 7.69 – 7.34 (m, 6H, 4CH<sub>Ar</sub> + H<sub>f</sub> + H<sub>h</sub>), 7.29 (t,  $J$  = 7.7 Hz, 1H, H<sub>g</sub>), 7.23, 7.21 (2s, 1H, H<sub>i</sub>), 6.66, 6.36 (2s, 1H, H<sub>d</sub>), 4.00 – 3.75, 3.45, 3.14 (m and d,  $J_{Conf.1}$  = 18.3 Hz and  $J_{Conf.2}$  = 16.0 Hz, 2H, H<sub>a</sub>), 3.08 – 2.90 (m, 2H, H<sub>10</sub>), 2.90 – 2.77, 2.65 – 2.45, 2.44 – 2.30 (2m, 2H, H<sub>6</sub>), 2.19, 2.18 (2s, 3H, H<sub>m</sub>), 1.86 – 1.62 (m, 2H, H<sub>7</sub>), 1.44 (m, 22H, 3CH<sub>3.Boc</sub> + H<sub>1</sub> + H<sub>8-9</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 177.0, 176.7 (C<sub>5</sub>), 173.0, 172.8 (C<sub>3</sub>), 170.8 (C<sub>l</sub>), 170.6, 170.0 (C<sub>b</sub>), 159.4, 158.5, 158.4 (CO<sub>Boc</sub> + C<sub>k</sub>), 150.8, 150.8 (C<sub>i</sub>), 139.9, 139.7, 136.6, 136.6 (C<sub>c</sub> + C<sub>e</sub>), 135.5, 135.4, 133.6, 133.2, 132.1, 131.7 (3C<sub>IV.Ar</sub>), 131.5, 131.2, 130.3, 130.1 (2CH<sub>Ar</sub> + C<sub>g</sub>), 130.1, 130.0, 128.8, 128.8, 128.6, 128.5, 127.7, 127.6, 126.5, 126.5, 123.9, 123.5 (5CH<sub>Ar</sub>), 123.2, 123.0 (C<sub>f</sub>), 120.4, 120.3 (C<sub>h</sub>), 118.6, 118.5 (C<sub>d</sub>), 108.9, 108.8 (C<sub>j</sub>), 79.8 (C<sub>IV.Boc</sub>), 64.2, 61.9 (C<sub>4</sub>), 52.8, 52.7 (C<sub>2</sub>), 50.1 (C<sub>a</sub>), 41.2 (C<sub>10</sub>), 34.6, 34.3 (C<sub>6</sub>), 30.8, 30.7 (C<sub>9</sub>), 28.9, 28.9, 28.8, (3CH<sub>3.Boc</sub> + C<sub>1</sub>), 27.6, 27.5 (C<sub>8</sub>), 25.8, 25.7 (C<sub>7</sub>), 22.6 (C<sub>m</sub>); **MS (ESI):**  $m/z$  743.00 [M+H]<sup>+</sup> (theoretical  $m/z$  743.36).

**S-Gly-(R/S)-Naphthalene- $\epsilon$ -Ahx-*tert*-butylacetamide (C16b):**

General procedure B was employed for the deprotection of compound **C16a** (43.3 mg, 0.058 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of enantiomers **C16b** was isolated as a pale-yellow powder (43.0 mg, 95%).

**HPLC:** rt = 20.19 min (method A); **NMR data** as a mixture of 2 conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.07 – 7.77 (m, 4H, 3CH<sub>Ar</sub> + H<sub>d</sub>), 7.67 – 7.46 (m, 5H, 4CH<sub>Ar</sub> + H<sub>f</sub>), 7.44 – 7.27 (m, 2H, H<sub>g-h</sub>), 7.23, 7.23 (2s, 1H, H<sub>j</sub>), 6.68, 6.37 (2s, 1H, H<sub>4</sub>), 3.98, 3.87, 3.53, 3.19 (4d,  $J_{Conf.1}$  = 18.3 Hz and  $J_{Conf.2}$  = 16.5 Hz, 1H, H<sub>a</sub>), 2.95 – 2.80 (m, 2H, H<sub>10</sub>), 2.68 – 2.51, 2.46 – 2.34 (2m, 2H, H<sub>6</sub>), 2.20, 2.20 (2s, 3H, H<sub>m</sub>), 1.88 – 1.57 (m, 4H, H<sub>7</sub> + H<sub>9</sub>), 1.58 – 1.23 (m, 2H, H<sub>8</sub>), 1.53, 1.45 (2s, 9H, H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 176.8, 176.5 (C<sub>5</sub>), 173.0, 172.8 (C<sub>3</sub>), 170.9 (C<sub>l</sub>), 170.7, 169.9 (C<sub>b</sub>), 159.5 (C<sub>k</sub>), 150.8, 150.7 (C<sub>i</sub>), 139.9, 139.6, 136.6 (C<sub>c</sub> + C<sub>e</sub>), 135.6, 135.4, 133.6, 133.2, 132.0, 131.7 (3C<sub>IV,Ar</sub>), 131.5, 131.3, 130.4, 130.1, 130.1 (2CH<sub>Ar</sub> + C<sub>g</sub>), 128.8, 128.5, 128.5, 127.7, 127.6, 126.6, 126.5, 124.0, 123.4 (5CH<sub>Ar</sub>), 123.3, 123.0 (C<sub>f</sub>), 120.5, 120.4 (C<sub>h</sub>), 118.6, 118.5 (C<sub>d</sub>), 108.9, 108.8 (C<sub>j</sub>), 64.2, 61.7 (C<sub>4</sub>), 53.2, 52.7 (C<sub>2</sub>), 50.1 (C<sub>a</sub>), 40.5 (C<sub>10</sub>), 34.2, 34.1 (C<sub>6</sub>), 28.9, 28.8 (C<sub>1</sub>), 28.4, 28.3 (C<sub>9</sub>), 27.1, 26.9 (C<sub>8</sub>), 25.4, 25.3 (C<sub>7</sub>), 22.6 (C<sub>m</sub>); **HRMS (ESI):**  $m/z$  643.30597 [M+H]<sup>+</sup> (C<sub>35</sub>H<sub>43</sub>N<sub>6</sub>O<sub>4</sub>S requires 643.30610).

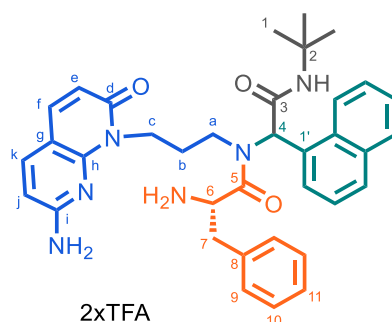
**(Boc)-Naphthyridine(N-alkyl)-(R/S)-Naphthalene-(Boc)-L-Phe-*tert*-butylacetamide (C17a):**

Compound **C17a** was prepared following general procedure A using **Boc-naphthyridine(N-alkyl)-NH<sub>2</sub>** (**42**) (31.8 mg, 0.10 mmol) as the amine, **1-naphthaldehyde** (13.6  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-**

**Phe-OH** (26.5 mg, 0.10 mmol) as the acid and **tert-butyl isocyanide** (11.3  $\mu$ l, 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C17a** was obtained as a white solid (53.5 mg, 67%).

$R_{f, \text{dia}1+2} = 0.38$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 96:4); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 7.98 - 7.81$  (m, 2H, H<sub>j</sub> + H<sub>k</sub>), 7.74 - 7.58 (m, 2H, CH<sub>Ar</sub> + H<sub>f</sub>), 8.13 - 8.08, 7.80 - 7.74, 7.56 - 7.11 (3m, 10H, 10CH<sub>Ar</sub>), 7.11 - 7.03, 7.00 - 6.84 (2m, 1H, CH<sub>Ar</sub>), 6.79, 6.72, 6.50, 6.48 (4s, 1H, H<sub>4</sub>), 6.44, 6.38, 6.32 (3d,  $J = 9.4$  Hz, 1H, H<sub>e</sub>), 5.12 - 5.05, 4.94 - 4.62 (2m, 1H, H<sub>6</sub>), 4.19 - 3.76 (m, 2H, H<sub>c</sub>), 3.67 - 3.29 (m, 2H, H<sub>a</sub>), 3.27 - 2.75 (m, 2H, H<sub>7</sub>), 1.96 - 1.67, 0.67 - 0.34 (2m, 1H, H<sub>b</sub>), 1.67 - 1.08 (m, 27H, 6CH<sub>3, Boc</sub> + H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 174.8, 173.9$  (C<sub>5</sub>), 171.7, 171.5 (C<sub>3</sub>), 165.1, 165.1, 164.7 (C<sub>d</sub>), 157.5, 154.2, 154.1 (2CO<sub>Boc</sub> + C<sub>i</sub>), 149.0, 148.9 (C<sub>h</sub>), 140.0, 139.9 (C<sub>k</sub>), 139.4, 139.2 (C<sub>f</sub>), 138.5, 138.4 (C<sub>8</sub>), 134.5, 134.5, 133.8, 133.6, 132.5, 132.4 (3C<sub>IV, Ar</sub>), 130.5, 130.5, 130.4, 130.1, 129.7, 129.7, 129.4, 129.4, 129.3, 128.2, 128.0, 127.8, 127.8, 127.6, 126.9, 126.8, 125.7, 125.4, 124.4, 123.8 (12CH<sub>Ar</sub>), 119.5, 119.4 (C<sub>e</sub>), 112.9, 112.7 (C<sub>g</sub>), 108.8, 108.6 (C<sub>j</sub>), 82.0, 82.0, 80.8, 80.7, 80.4 (2C<sub>IV, Boc</sub>), 62.1, 59.9, 59.3 (C<sub>4</sub>), 54.6, 54.2, 53.2 (C<sub>6</sub>), 52.5, 52.4, 52.2 (C<sub>2</sub>), 45.4, 45.0, 44.7 (C<sub>a</sub>), 40.1, 39.2, 39.0, 39.0 (C<sub>c</sub> + C<sub>7</sub>), 30.1 (C<sub>b</sub>), 28.9, 28.8, 28.7, 28.7, 28.7, 28.1 (6CH<sub>3, Boc</sub> + C<sub>1</sub>); **MS (ESI):**  $m/z$  804.93 [M+H]<sup>+</sup> (theoretical  $m/z$  805.43).

#### Naphthyridine(*N*-alkyl)-(*R/S*)-Naphthalene-*L*-Phe-*tert*-butylacetamide (**C17b**):



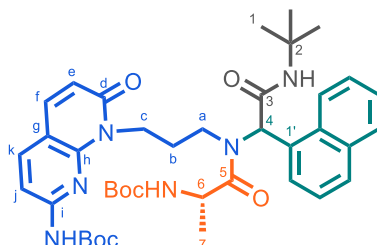
General procedure B was employed for the deprotection of compound **C17a** (39.7 mg, 0.049 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C17b** was isolated as a white powder (38.1 mg, 93%).

**HPLC:**  $rt_{\text{dia}1+2} = 19.81$  min (method A); NMR data as a mixture of 2 diastereomers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 7.91 - 7.73, 7.54 - 7.43$  (2m, 4H, 4CH<sub>Ar</sub>), 7.69 - 7.54 (m, 3H, CH<sub>Ar</sub> + H<sub>f</sub> + H<sub>k</sub>), 7.43 - 7.07 (m, 7H, 7CH<sub>Ar</sub>), 6.84, 6.47 (2s, 1H, H<sub>4</sub>), 6.47, 6.43 (2d,  $J = 8.5$  Hz, 1H, H<sub>j</sub>), 6.22, 6.18 (2d,  $J = 9.2$  Hz, 1H, H<sub>e</sub>), 4.48, 4.39 (2dd,  $J = 8.3, 6.4$  Hz and  $J = 7.7, 4.3$  Hz, 1H, H<sub>6</sub>), 4.13 - 3.78 (m, 2H, H<sub>c</sub>), 3.64 - 2.88 (m, 4H, H<sub>a</sub> + H<sub>7</sub>), 1.58 - 1.45, 1.44 - 1.11, 0.95 - 0.71, 0.65 - 0.45 (4m, 2H, H<sub>b</sub>), 1.35, 1.31 (2s, 9H, H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 172.3, 171.4$  (C<sub>5</sub>), 170.7, 169.9 (C<sub>3</sub>), 165.6, 165.5 (C<sub>d</sub>), 161.5, 161.3 (C<sub>i</sub>), 150.5, 150.4 (C<sub>h</sub>), 140.3, 140.2 (C<sub>f</sub>), 139.2, 139.2 (C<sub>k</sub>), 135.4, 134.9, 134.7, 134.7, 133.8, 133.5, 131.9, 131.3 (3C<sub>IV, Ar</sub> + C<sub>8</sub>), 130.9, 130.8, 130.6, 130.2, 130.1, 123.0, 128.9, 128.6, 128.5, 128.3, 128.0, 127.1, 125.7,



124.2, 123.9 ( $12\text{CH}_{Ar}$ ), 114.6 ( $C_e$ ), 108.6 ( $C_g$ ), 106.7, 106.6 ( $C_j$ ), 60.8, 59.3 ( $C_4$ ), 53.9, 53.2 ( $C_6$ ), 52.5, 52.4 ( $C_2$ ), 44.9, 44.8 ( $C_a$ ), 39.0, 38.8, 38.6, 38.1 ( $C_c + C_7$ ), 30.7, 30.2 ( $C_b$ ), 28.8, 28.7 ( $C_1$ ); **HRMS (ESI):**  $m/z$  605.32404  $[\text{M}+\text{H}]^+$  ( $\text{C}_{36}\text{H}_{41}\text{N}_6\text{O}_3$  requires 605.32347).

**(Boc)-Naphthyridine(*N*-alkyl)-(R/S)-Naphthalene-(Boc)-L-Ala-*tert*-butylacetamide (C18a and C19a):**



Compounds **C18a** and **C19a** were prepared following general procedure A using **Boc-naphthyridine(*N*-alkyl)-NH<sub>2</sub> (42)** (31.8 mg, 0.10 mmol) as the amine, **1-naphthaldehyde** (13.6  $\mu\text{L}$ , 0.10 mmol) as aldehyde, **Boc-L-Ala-OH** (18.9 mg, 0.10 mmol) as the acid and ***tert*-butyl isocyanide** (11.3  $\mu\text{L}$ , 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  gradient 100:0 to 96:4, v/v), the diastereomers **C18a (dia1)** (22.7 mg, 31%) and **C19a (dia2)** (26.3 mg, 36%) were separated as white solids.

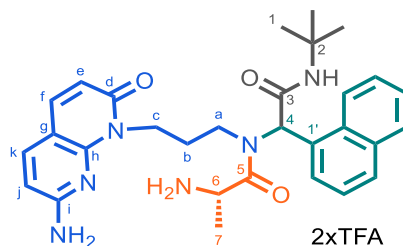
$R_{f,\text{dia1}} = 0.31$  and  $R_{f,\text{dia2}} = 0.29$  ( $\text{CH}_2\text{Cl}_2\text{-MeOH}$  96:4);

**C18a – Dia1:** NMR data as a mixture of 2 conformers; **<sup>1</sup>H NMR** (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 8.27, 7.64 - 7.56$  (d and m,  $J = 8.4$  Hz, 1H,  $\text{CH}_{Ar}$ ), 7.97, 7.92 (2d,  $J = 8.6$  Hz, 1H,  $H_k$ ), 7.87, 7.83 (2d,  $J = 8.6$  Hz, 1H,  $H_j$ ), 7.76, 7.73 – 7.65 (d and m,  $J = 9.4$  Hz, 1H,  $H_f$ ), 7.73 – 7.65, 7.64 – 7.56 (2m, 1H,  $\text{CH}_{Ar}$ ), 7.49 – 7.24 (m, 3H,  $3\text{CH}_{Ar}$ ), 7.20, 7.09 (2d,  $J = 7.1$  Hz, 1H,  $\text{CH}_{Ar}$ ), 6.95, 6.90 (2t,  $J = 7.7$  Hz, 1H,  $\text{CH}_{Ar}$ ), 6.64, 6.50 (2s, 1H,  $H_4$ ), 6.43, 6.37 (2d,  $J = 9.4$  Hz, 1H,  $H_e$ ), 4.90 – 4.78, 4.46 (m and q,  $J = 6.6$  Hz, 1H,  $H_6$ ), 4.16 – 3.80 (m, 2H,  $H_c$ ), 3.66 – 2.98 (m, 2H,  $H_a$ ), 1.83 – 1.18, 0.77 – 0.54 (2m, 2H,  $H_b$ ), 1.62, 1.60, 1.51, 1.47, 1.39, 1.32 (6s, 27H,  $6\text{CH}_{3,\text{Boc}} + H_1$ ), 1.33 – 1.26 (m, 3H,  $H_7$ ); **<sup>13</sup>C NMR** (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 176.8, 176.7$  ( $C_5$ ), 171.4, 171.3 ( $C_3$ ), 165.1 ( $C_d$ ), 157.8, 157.7, 154.3, 154.2, 154.2, 154.0 ( $2\text{CO}_{\text{Boc}} + C_i$ ), 149.2, 149.0 ( $C_h$ ), 140.1, 139.8 ( $C_k$ ), 139.5, 139.3 ( $C_f$ ), 134.6, 134.5, 133.7, 133.6, 132.6, 131.8 ( $3\text{C}_{\text{IV},Ar}$ ), 130.6, 130.1, 130.0, 129.6, 128.1, 127.9, 127.9, 127.0, 126.9, 125.6, 125.3, 125.2, 123.6 ( $7\text{CH}_{Ar}$ ), 119.5, 119.4 ( $C_e$ ), 112.8 ( $C_g$ ), 108.8, 108.7 ( $C_j$ ), 82.0, 81.9, 80.6, 80.4 ( $2\text{C}_{\text{IV},\text{Boc}}$ ), 62.3, 60.1 ( $C_4$ ), 52.8, 52.4 ( $C_2$ ), 48.2, 47.7 ( $C_6$ ), 45.3, 45.0 ( $C_a$ ), 39.4, 39.0 ( $C_c$ ), 30.2 ( $C_b$ ), 28.9, 28.8, 28.7, 28.7 ( $6\text{CH}_{3,\text{Boc}} + C_1$ ), 18.6, 18.3 ( $C_7$ ); **MS (ESI):**  $m/z$  728, 87  $[\text{M}+\text{H}]^+$  (theoretical  $m/z$  729.40).

**C19a – Dia1:** NMR data as a single conformer; **<sup>1</sup>H NMR** (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 7.95$  (d,  $J = 8.6$  Hz, 1H,  $H_k$ ), 7.86 (d,  $J = 8.6$  Hz, 1H,  $H_j$ ), 7.78 (d,  $J = 8.7$  Hz, 1H,  $\text{CH}_{Ar}$ ), 7.73 (d,  $J = 9.3$  Hz, 1H,  $H_f$ ), 7.65 – 7.57, 7.48 – 7.38 (m, 3H,  $3\text{CH}_{Ar}$ ), 7.34 (d,  $J = 8.1$  Hz, 1H,  $\text{CH}_{Ar}$ ), 7.14 (d,  $J = 7.0$  Hz, 1H,  $\text{CH}_{Ar}$ ), 6.85 (t,  $J = 7.7$  Hz, 1H,  $\text{CH}_{Ar}$ ), 6.69 (s, 1H,  $H_4$ ), 6.44 (d,  $J = 9.3$  Hz, 1H,  $H_e$ ), 4.57 – 4.47 (m, 1H,  $H_6$ ), 4.15 – 4.04, 3.93 – 3.81 (2m,

2H, H<sub>c</sub>), 3.48 – 3.33 (m, 2H, H<sub>a</sub>), 1.99 – 1.84, 0.69 – 0.47 (2m, 1H, H<sub>b</sub>), 1.61, 1.47, 1.27 (3s, 27H, 6CH<sub>3</sub>.Boc + H<sub>1</sub>), 1.38 (d, *J* = 6.9 Hz, 3H, H<sub>7</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD) δ = 176.3 (C<sub>5</sub>), 171.8 (C<sub>3</sub>), 165.2 (C<sub>d</sub>), 157.5, 154.2, 154.1 (2CO<sub>Boc</sub> + C<sub>i</sub>), 149.0 (C<sub>h</sub>), 139.9 (C<sub>k</sub>), 139.4 (C<sub>f</sub>), 134.4, 133.9, 132.5 (3C<sub>IV.Ar</sub>), 130.1, 129.6, 127.9, 127.8, 126.9, 125.3, 124.7 (7CH<sub>Ar</sub>), 119.4 (C<sub>e</sub>), 112.9 (C<sub>g</sub>), 108.8 (C<sub>j</sub>), 82.0, 80.3 (2C<sub>IV.Boc</sub>), 59.2 (C<sub>4</sub>), 52.3 (C<sub>2</sub>), 48.4 (C<sub>6</sub>), 44.9 (C<sub>a</sub>), 39.0 (C<sub>c</sub>), 30.1 (C<sub>b</sub>), 28.8, 28.8, 28.7 (6CH<sub>3</sub>.Boc + C<sub>1</sub>), 18.2 (C<sub>7</sub>); **MS (ESI):** *m/z* 728,93 [M+H]<sup>+</sup> (theoretical *m/z* 729.40).

### Naphthyridine(*N*-alkyl)-(R/S)-Naphthalene-L-Ala-*tert*-butylacetamide (**C18b** and **C19b**):



#### **C18b – Dia1**

General procedure B was employed for the deprotection of compound **C18a** (16.5 mg, 0.023 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a diastereomer **C18b** was isolated as a white powder (12.6 mg, 74%).

**HPLC:** *rt*<sub>dia1</sub> = 16.73 min (method A); NMR data as a mixture of 2 conformers; <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) δ = 7.80 – 7.62 (m, 3H, CH<sub>Ar</sub> + H<sub>f</sub> + H<sub>k</sub>), 7.59 – 7.44 (m, 3H, 3CH<sub>Ar</sub>), 7.33 (d, *J* = 7.2 Hz, 1H, CH<sub>Ar</sub>), 7.07 – 7.00 (m, 1H, CH<sub>Ar</sub>), 6.91 – 6.81 (m, 1H, CH<sub>Ar</sub>), 6.60, 6.53 (2d, *J* = 8.6 Hz, 1H, H<sub>j</sub>), 6.30, 6.28 (2s, 1H, H<sub>4</sub>), 6.24 – 6.15 (m, 1H, H<sub>e</sub>), 4.91, 4.60 (2q, *J* = 6.8 Hz, 1H, H<sub>6</sub>), 4.09 – 3.89, 3.80 – 3.65 (2m, 2H, H<sub>c</sub>), 3.25 – 3.13, 3.06 – 2.93, 2.89 – 2.78 (3m, 2H, H<sub>a</sub>), 1.82 – 1.66, 1.10 – 0.95 (2m, 1H, H<sub>b</sub>), 1.63, 1.53 (2d, *J* = 6.8 Hz, 3H, H<sub>7</sub>), 1.28 (s, 9H, H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O) δ = 172.0 (C<sub>5</sub>), 171.0 (C<sub>3</sub>), 164.6 (C<sub>d</sub>), 159.0 (C<sub>i</sub>), 147.1 (C<sub>h</sub>), 140.3 (C<sub>f</sub>), 139.4 (C<sub>k</sub>), 132.5, 131.2 (2C<sub>IV.Ar</sub>), 130.1, 129.2 (2CH<sub>Ar</sub>), 128.2 (C<sub>IV.Ar</sub>), 127.4, 126.8, 126.2, 124.3, 121.5 (5CH<sub>Ar</sub>), 113.7 (C<sub>e</sub>), 108.5 (C<sub>g</sub>), 107.0 (C<sub>j</sub>), 60.4 (C<sub>4</sub>), 51.8 (C<sub>2</sub>), 47.6 (C<sub>6</sub>), 43.8 (C<sub>a</sub>), 37.8 (C<sub>c</sub>), 28.9 (C<sub>b</sub>), 27.6 (C<sub>1</sub>), 17.0 (C<sub>7</sub>); **HRMS (ESI):** *m/z* 529.29266 [M+H]<sup>+</sup> (C<sub>30</sub>H<sub>37</sub>N<sub>6</sub>O<sub>3</sub> requires 529.29217).

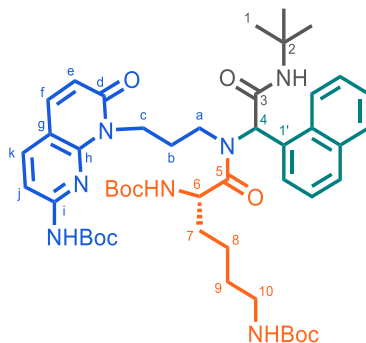
#### **C19b – Dia1**

General procedure B was employed for the deprotection of compound **C19a** (20.3 mg, 0.028 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a diastereomer **C19b** was isolated as a white powder (16.3 mg, 77%).

**HPLC:** *rt*<sub>dia2</sub> = 17.03 min (method A); NMR data as a single conformer; <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O) δ = 7.71 (d, *J* = 8.6 Hz, 1H, H<sub>k</sub>), 7.67 – 7.57 (m, 3H, 2CH<sub>Ar</sub> + H<sub>f</sub>), 7.56 – 7.54 (m, 2H, 2CH<sub>Ar</sub>), 7.26 (d, *J* = 7.7 Hz, 1H, CH<sub>Ar</sub>), 7.05 (d, *J* = 7.7 Hz, 1H, CH<sub>Ar</sub>), 6.82 (t, *J* = 7.7 Hz, 1H, CH<sub>Ar</sub>), 6.56 (d, *J* = 8.6 Hz, 1H, H<sub>j</sub>), 6.49 (s, 1H, H<sub>4</sub>), 6.21 (d, *J* = 9.2 Hz, 1H, H<sub>e</sub>), 4.47 (q, *J* = 6.7 Hz, 1H, H<sub>6</sub>), 3.99 – 3.88, 3.58 – 3.48 (2m, 2H, H<sub>c</sub>), 3.42 –

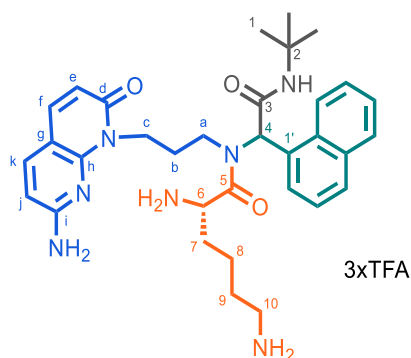
3.29, 3.17 – 3.06 (2m, 2H, H<sub>d</sub>), 1.69 – 1.52, 0.52 – 0.37 (2m, 2H, H<sub>b</sub>), 1.64 (d, *J* = 6.7 Hz, 3H, H<sub>7</sub>), 1.24 (s, 9H, H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O) δ = 171.2 (C<sub>5</sub>), 170.3 (C<sub>3</sub>), 164.4 (C<sub>d</sub>), 158.8 (C<sub>i</sub>), 146.7 (C<sub>h</sub>), 140.2 (C<sub>f</sub>), 139.5 (C<sub>k</sub>), 132.4, 131.4 (2C<sub>IV,Ar</sub>), 129.9, 129.0 (2CH<sub>Ar</sub>), 128.3 (C<sub>IV,Ar</sub>), 127.4, 127.2, 126.2, 124.4, 122.1 (5CH<sub>Ar</sub>), 113.8 (C<sub>e</sub>), 108.4 (C<sub>g</sub>), 107.0 (C<sub>j</sub>), 58.9 (C<sub>4</sub>), 51.6 (C<sub>2</sub>), 47.6 (C<sub>6</sub>), 43.5 (C<sub>a</sub>), 37.6 (C<sub>c</sub>), 29.1 (C<sub>b</sub>), 27.6 (C<sub>1</sub>), 16.4 (C<sub>7</sub>); **HRMS (ESI):** *m/z* 529.29248 [M+H]<sup>+</sup> (C<sub>30</sub>H<sub>37</sub>N<sub>6</sub>O<sub>3</sub> requires 529.29217).

**(Boc)-Naphthyridine(*N*-alkyl)-(R/S)-Naphthalene-(Boc)-L-Lys(Boc)-*tert*-butylacetamide**  
**(C20a):**



Compound **C20a** was prepared following general procedure A using **Boc-naphthyridine(*N*-alkyl)-NH<sub>2</sub> (42)** (31.8 mg, 0.10 mmol) as the amine, **1-naphthaldehyde** (13.6 μL, 0.10 mmol) as aldehyde, **Boc-Lys(Boc)-OH** (34.6 mg, 0.10 mmol) as the acid and ***tert*-butyl isocyanide** (11.3 μL, 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 96:4, v/v), a mixture of diastereomers **C20a** was obtained as a pale-white solid (52.1 mg, 59%).

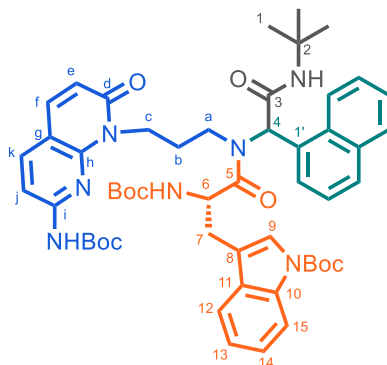
**R<sub>f,dia1</sub>** = 0.32 and **R<sub>f,dia2</sub>** = 0.35 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 96:4); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ = 7.99 – 7.81 (m, 2H, H<sub>j</sub> + H<sub>k</sub>), 8.20 – 8.13, 7.81 – 7.57 (2m, 3H, 2CH<sub>Ar</sub> + H<sub>f</sub>), 7.57 – 7.33 (m, 3H, 3CH<sub>Ar</sub>), 7.33 – 7.13 (m, 1H, CH<sub>Ar</sub>), 7.05 – 6.84 (m, 1H, CH<sub>Ar</sub>), 6.70, 6.65 (2s, 1H, H<sub>4</sub>), 6.49 – 6.32 (m, 1H, H<sub>e</sub>), 4.79 – 4.70, 4.50 – 4.22 (2m, 1H, H<sub>6</sub>), 4.17 – 3.93, 3.92 – 3.77 (2m, 2H, H<sub>c</sub>), 3.61 – 3.29 (m, 2H, H<sub>a</sub>), 3.13 – 2.95 (m, 2H, H<sub>10</sub>), 1.99 – 1.15, 0.72 – 0.47 (2m, 44H, 9CH<sub>3,Boc</sub> + H<sub>b</sub> + H<sub>1</sub> + H<sub>7-9</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD) δ = 175.7, 175.7, 175.6 (C<sub>5</sub>), 171.7, 171.5, 171.4 (C<sub>3</sub>), 165.2, 165.1, 165.0 (C<sub>d</sub>), 158.5, 158.2, 157.9, 157.8, 154.3, 154.2, 154.1 (3CO<sub>Boc</sub> + C<sub>i</sub>), 149.2, 149.0 (C<sub>h</sub>), 140.1, 139.9, 139.8, 139.5, 139.4, 139.3 (C<sub>f</sub> + C<sub>k</sub>), 134.7, 134.5, 133.8, 133.7, 132.8, 132.5 (3C<sub>IV,Ar</sub>), 130.6, 130.1, 130.0, 129.6, 128.1, 127.9, 127.8, 127.0, 126.9, 125.7, 125.3, 124.6, 123.8 (7CH<sub>Ar</sub>), 119.5, 119.4 (C<sub>e</sub>), 112.9, 112.8 (C<sub>g</sub>), 108.8, 108.7 (C<sub>j</sub>), 82.0, 82.0, 80.7, 80.5, 80.3, 79.8 (3C<sub>IV,Boc</sub>), 59.9, 59.2, 59.2 (C<sub>4</sub>), 53.0, 52.7, 52.6, 52.3, 52.3, 52.1 (C<sub>2</sub> + C<sub>6</sub>), 45.3, 44.9, 44.9 (C<sub>a</sub>), 41.1, 41.0 (C<sub>10</sub>), 39.4, 39.0, 39.0 (C<sub>c</sub>), 33.7, 33.1, 32.8 (C<sub>7</sub>), 30.9, 30.6, 30.5, 30.1 (C<sub>b</sub> + C<sub>9</sub>), 28.9, 28.9, 28.8, 28.7 (9CH<sub>3,Boc</sub> + C<sub>1</sub>), 24.5, 24.3, 24.2 (C<sub>8</sub>); **MS (ESI):** *m/z* 885.93 [M+H]<sup>+</sup> (theoretical *m/z* 886.51).

**Naphthyridine(*N*-alkyl)-(R/S)-Naphthalene-L-Lys-*tert*-butylacetamide (C20b):**

General procedure B was employed for the deprotection of compound **C20a** (35.2 mg, 0.040 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C20b** was isolated as a white powder (32.2 mg, 88%).

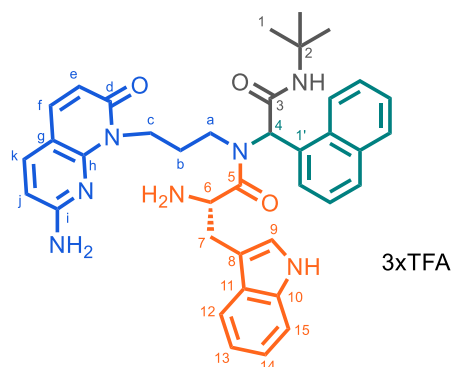
**HPLC:**  $rt_{dia1}$  = 14.92 min and  $rt_{dia2}$  = 15.15 min (method A); NMR data as a mixture of 2 diastereomers; **<sup>1</sup>H NMR** (400 MHz, D<sub>2</sub>O)  $\delta$  = 7.75 – 7.41 (m, 6H, 4CH<sub>Ar</sub> + H<sub>f</sub> + H<sub>k</sub>), 7.34, 7.26 (2d,  $J$  = 8.2 Hz, 1H, CH<sub>Ar</sub>), 7.07, 7.03 (2d,  $J$  = 7.2 Hz, 1H, CH<sub>Ar</sub>), 6.87 – 6.76 (m, 1H, CH<sub>Ar</sub>), 6.54, 6.49 (2d,  $J$  = 8.4 Hz, 1H, H<sub>j</sub>), 6.52, 6.29 (2s, 1H, H<sub>4</sub>), 6.23, 6.13 (2d,  $J$  = 9.3 Hz, 1H, H<sub>e</sub>), 4.53, 4.48 (2t,  $J$  = 5.9 Hz, 1H, H<sub>6</sub>), 4.01 – 3.82, 3.81 – 3.51 (2m, 2H, H<sub>c</sub>), 3.50 – 3.32, 3.28 – 2.96, 2.94 – 2.75 (3m, 2H, H<sub>o</sub>), 3.13 – 2.95 (m, 2H, H<sub>10</sub>), 2.17 – 1.88 (m, 2H, H<sub>7</sub>), 1.88 – 1.43 (m, 4H, H<sub>8-9</sub>), 1.88 – 1.43, 1.11 – 0.96, 0.57 – 0.41 (3m, 2H, H<sub>b</sub>), 1.29, 1.25 (2s, 9H, H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, D<sub>2</sub>O)  $\delta$  = 170.9, 170.9 (C<sub>5</sub>), 170.1, 170.0 (C<sub>3</sub>), 164.6, 164.6 (C<sub>d</sub>), 159.3, 159.2 (C<sub>i</sub>), 147.6, 147.1 (C<sub>h</sub>), 140.3, 140.2 (C<sub>f</sub>), 139.2, 139.1 (C<sub>k</sub>), 132.6, 132.4, 131.4, 131.1 (2C<sub>IV,Ar</sub>), 130.2, 129.9, 129.3, 129.0 (2CH<sub>Ar</sub>), 128.2, 128.1 (C<sub>IV,Ar</sub>), 127.2, 126.9, 126.2, 126.2, 124.4, 122.1, 121.5 (5CH<sub>Ar</sub>), 113.5 (C<sub>e</sub>), 108.6, 108.5 (C<sub>g</sub>), 107.1, 106.8 (C<sub>j</sub>), 60.3, 58.9 (C<sub>4</sub>), 51.8, 51.5 (C<sub>2</sub>), 51.0, 50.8 (C<sub>6</sub>), 43.9, 43.6 (C<sub>a</sub>), 39.0 (C<sub>10</sub>), 37.6, 37.5 (C<sub>c</sub>), 31.1, 30.6 (C<sub>7</sub>), 29.1, 29.1 (C<sub>b</sub>), 27.5, 27.5 (C<sub>1</sub>), 26.6, 26.5 (C<sub>9</sub>), 21.8, 21.0 (C<sub>8</sub>); **HRMS (ESI):**  $m/z$  586.35059 [M+H]<sup>+</sup> (C<sub>33</sub>H<sub>44</sub>N<sub>7</sub>O<sub>3</sub> requires 586.35001).

**(Boc)-Naphthyridine(*N*-alkyl)-(R/S)-Naphthalene-(Boc)-L-Trp(Boc)-*tert*-butylacetamide  
(C21a):**



Compound **C21a** was prepared following general procedure A using **Boc-naphthyridine(*N*-alkyl)-NH<sub>2</sub> (42)** (31.8 mg, 0.10 mmol) as the amine, **1-naphthaldehyde** (13.6  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-Trp(Boc)-OH** (40.5 mg, 0.10 mmol) as the acid and ***tert*-butyl isocyanide** (11.3  $\mu$ L, 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C21a** was obtained as a white solid (59.6 mg, 63%).

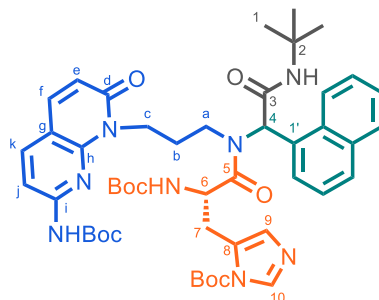
$R_{f,dia1}$  = 0.41 and  $R_{f,dia2}$  = 0.46 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 96:4); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.18 – 8.01 (m, 1H, H<sub>15</sub>), 7.94 – 7.54 (m, 4H, H<sub>f</sub> + H<sub>j-k</sub>), 7.93 – 6.75 (m, 11H, 11CH<sub>Ar</sub>), 6.79, 6.73, 6.44, 6.46 (4s, 1H, H<sub>4</sub>), 6.43–6.24 (m, 1H, H<sub>e</sub>), 5.19 – 5.08, 5.05 – 4.93, 5.92 – 4.71 (3m, 1H, H<sub>6</sub>), 4.30 – 3.95, 3.90 – 3.69 (2m, 2H, H<sub>c</sub>), 3.68 – 3.22 (m, 2H, H<sub>a</sub>), 3.52 – 3.22, 3.19 – 2.84 (2m, 2H, H<sub>7</sub>), 1.95 – 0.81, 0.74 – 0.50, 0.39 – 0.22 (3m, 2H, H<sub>b</sub>), 1.77 – 0.81 (m, 36H, 9CH<sub>3,Boc</sub> + H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 174.8, 174.5, 174.3, 173.3 (C<sub>5</sub>), 171.9, 171.7, 171.6 (C<sub>3</sub>), 165.2, 165.1, 164.7 (C<sub>d</sub>), 158.1, 157.6, 157.6, 157.6, 154.1, 154.1, 154.0 (2CO<sub>Boc</sub> + C<sub>i</sub>), 151.0, 149.0, 148.9 (CO<sub>Boc</sub> + C<sub>h</sub>), 140.0, 139.8 (C<sub>k</sub>), 139.3, 139.2 (C<sub>f</sub>), 136.8, 134.5, 134.4, 133.8, 133.5, 132.5, 132.3, 131.8 (5C<sub>IV,Ar</sub>), 130.4, 130.1, 129.7, 128.1, 128.0, 127.8, 127.6, 126.9, 126.8, 125.6, 125.4, 125.3, 124.4, 123.8, 123.6, 123.5 (10CH<sub>Ar</sub>), 120.3, 120.2 (C<sub>12</sub>), 119.4, 119.3 (C<sub>e</sub>), 117.8, 117.7 (C<sub>8</sub>), 116.1 (C<sub>15</sub>), 112.8 (C<sub>g</sub>), 108.7, 108.6 (C<sub>j</sub>), 84.7, 84.7, 82.1, 82.0, 81.9, 80.9, 80.4 (3C<sub>IV,Boc</sub>), 59.9, 59.4 (C<sub>4</sub>), 53.2, 52.8, 52.6, 52.5, 52.3, 51.8 (C<sub>2</sub> + C<sub>6</sub>), 45.2, 44.7 (C<sub>a</sub>), 39.1, 39.1 (C<sub>c</sub>), 30.8, 30.2, 29.9 (C<sub>b</sub>), 29.4, 28.9, 28.8, 28.7, 28.7, 28.5, 28.4, 28.1 (9CH<sub>3,Boc</sub> + C<sub>1</sub> + C<sub>7</sub>); **MS (ESI):**  $m/z$  944.07 [M+H]<sup>+</sup> (theoretical  $m/z$  944.49).

**Naphthyridine(*N*-alkyl)-(R/S)-Naphthalene-L-Trp-*tert*-butylacetamide (**C21b**):**

General procedure B was employed for the deprotection of compound **C21a** (40.7 mg, 0.043 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C21b** was isolated as a pale-white powder (34.3 mg, 81%).

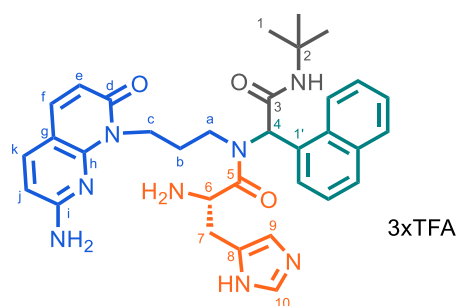
**HPLC:**  $rt_{dia1}$  = 19.30 min and  $rt_{dia2}$  = 19.55 min (method A); NMR data as a mixture of 2 diastereomers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 7.92 – 7.86, 7.80 – 7.71, 7.70 – 7.52 (3m, 3H, 3CH<sub>Ar</sub>), 7.64 – 7.52 (m, 2H, H<sub>f</sub> + H<sub>k</sub>), 7.51 – 7.33 (m, 4H, 4CH<sub>Ar</sub>), 7.28 (s, 1H, C<sub>9</sub>), 7.22 – 7.03, 7.01 – 6.95 (2m, 4H, 4CH<sub>Ar</sub>), 6.85, 6.31 (2s, 1H, H<sub>4</sub>), 6.44, 6.37 (2d,  $J$  = 8.5 Hz, 1H, H<sub>j</sub>), 6.19, 6.16 (2d,  $J$  = 9.2 Hz, 1H, H<sub>e</sub>), 4.61, 4.44 (2dd,  $J_{dia1}$  = 6.5, 8.4 Hz and  $J_{dia2}$  = 3.9, 9.2 Hz, 1H, H<sub>6</sub>), 4.14 – 4.04, 3.90 – 3.77, 3.74 – 3.60 (3m, 2H, H<sub>c</sub>), 3.74 – 3.60, 3.51 – 3.43, 3.40 – 3.18 (3m, 2H, H<sub>7</sub>), 3.60 – 3.52, 3.13 – 3.00, 2.89 – 2.76 (3m, 2H, H<sub>a</sub>), 1.55 – 1.41, 1.40 – 1.14, 0.95 – 0.80, 0.65 – 0.51 (4m, 2H, H<sub>b</sub>), 1.34, 1.30 (2s, 9H, H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 172.4, 171.4 (C<sub>5</sub>), 171.2, 170.6 (C<sub>3</sub>), 165.5 (C<sub>d</sub>), 161.5, 161.2 (C<sub>i</sub>), 150.4, 150.3 (C<sub>h</sub>), 140.2 (C<sub>f</sub>), 139.2, 139.0 (C<sub>k</sub>), 138.4, 138.2 (C<sub>10</sub>), 134.7, 134.5, 133.8, 133.4, 131.8, 131.1 (3C<sub>IV,Ar</sub>), 130.8, 130.8, 130.0, 129.9 (2CH<sub>Ar</sub>), 128.6, 128.4, 128.3, 128.0, 127.1, 127.1 (C<sub>IV,Ar</sub> + 3CH<sub>Ar</sub>), 126.1 (C<sub>9</sub>), 125.7, 125.4, 124.2, 123.5, 123.0, 120.6, 120.5 (4CH<sub>Ar</sub>), 119.0, 118.9 (C<sub>12</sub>), 114.6 (C<sub>e</sub>), 112.8 (C<sub>15</sub>), 108.6, 108.6 (C<sub>g</sub>), 108.0, 107.7 (C<sub>8</sub>), 106.6, 106.5 (C<sub>j</sub>), 61.2, 59.4 (C<sub>4</sub>), 53.5, 52.5, 52.4, 52.4 (C<sub>2</sub> + C<sub>6</sub>), 45.0, 44.8 (C<sub>a</sub>), 38.7, 38.6 (C<sub>c</sub>), 30.7, 30.4 (C<sub>b</sub>), 29.1, 28.8, 28.7, 28.7 (C<sub>1</sub> + C<sub>7</sub>); **HRMS (ESI):**  $m/z$  644.33478 [M+H]<sup>+</sup> (C<sub>38</sub>H<sub>42</sub>N<sub>7</sub>O<sub>3</sub> requires 644.33436).

**(Boc)-Naphthyridine(*N*-alkyl)-(*R/S*)-Naphthalene-(Boc)-L-His(Boc)-*tert*-butylacetamide  
(**C22a**):**



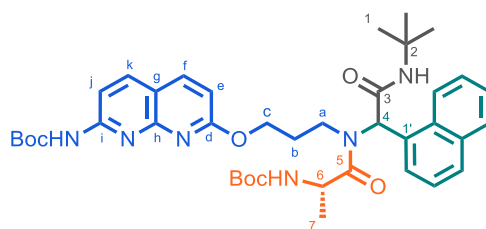
Compound **C22a** was prepared following general procedure A using **Boc-naphthyridine(*N*-alkyl)-NH<sub>2</sub> (42)** (31.8 mg, 0.10 mmol) as the amine, **1-naphthaldehyde** (13.6  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-His(Boc)-OH** (35.5 mg, 0.10 mmol) as the acid and ***tert*-butyl isocyanide** (11.3  $\mu$ L, 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C22a** was obtained as a white solid (51.0 mg, 63%).

$R_{f,dia1}$  = 0.32 and  $R_{f,dia2}$  = 0.35 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 96:4); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.21 – 8.02 (m, 1H, H<sub>10</sub>), 7.99 – 7.80 (m, 2H, H<sub>j</sub> + H<sub>k</sub>), 7.79 – 7.55 (m, 3H, 2CH<sub>Ar</sub> + H<sub>f</sub>), 7.55 – 7.10 (m, 5H, 4CH<sub>Ar</sub> + H<sub>9</sub>), 7.01 – 6.85 (m, 1H, CH<sub>Ar</sub>), 6.69, 6.54 (2s, 1H, H<sub>4</sub>), 6.40 (d,  $J$  = 9.1 Hz, 1H, H<sub>e</sub>), H<sub>6</sub> (signal under the D<sub>2</sub>O peak), 4.29 – 3.76 (m, 2H, H<sub>c</sub>), 3.71 – 3.26 (m, 2H, H<sub>a</sub>), 3.18 – 2.80 (m, 2H, H<sub>7</sub>), 1.97 – 1.08, 0.99 – 0.43 (2m, 38H, 9CH<sub>3,Boc</sub> + H<sub>b</sub> + H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 174.7, 174.3 (C<sub>5</sub>), 171.6, 171.5 (C<sub>3</sub>), 165.1, 165.0, 164.9 (C<sub>d</sub>), 157.5, 157.4, 154.2 (2CO<sub>Boc</sub> + C<sub>i</sub>), 149.2, 149.1, 148.2, 148.1 (CO<sub>Boc</sub> + C<sub>h</sub>), 140.2, 140.1, 139.9 (C<sub>k</sub>), 139.4, 139.3 (C<sub>f</sub>), 138.4, 138.2, 138.2 (C<sub>10</sub>), 134.5, 133.8, 133.6, 132.5, 132.4 (3C<sub>IV,Ar</sub>), 130.3, 130.2, 129.9, 129.7, 128.1, 127.9, 127.8, 127.7, 126.9, 125.6, 125.4, 124.5, 123.7 (7CH<sub>Ar</sub> + C<sub>8</sub>), 119.5, 119.4 (C<sub>e</sub>), 116.4, 116.1 (C<sub>9</sub>), 112.8, 112.8 (C<sub>g</sub>), 108.7 (C<sub>j</sub>), 87.0, 86.9, 81.9, 81.8, 80.6, 80.4 (3C<sub>IV,Boc</sub>), 60.2, 59.3 (C<sub>4</sub>), 52.7, 52.6 (C<sub>2</sub>), 52.2 (C<sub>6</sub>), 45.4, 45.0, 44.9 (C<sub>a</sub>), 39.4, 39.1 (C<sub>c</sub>), 32.5, 31.8, 31.5 (C<sub>7</sub>), 30.4, 30.0 (C<sub>b</sub>), 28.9, 28.8, 28.7, 28.6, 28.1 (9CH<sub>3,Boc</sub> + C<sub>1</sub>); **MS (ESI):**  $m/z$  894.93 [M+H]<sup>+</sup> (theoretical  $m/z$  895.47).

**Naphthyridine(*N*-alkyl)-(R/S)-Naphthalene-L-His-*tert*-butylacetamide (**C22b**):**

General procedure B was employed for the deprotection of compound **C22a** (36.4 mg, 0.041 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C22b** was isolated as a pale-white powder (31.6 mg, 83%).

**HPLC:**  $rt_{dia1}$  = 15.02 min and  $rt_{dia2}$  = 15.20 min (method A); **NMR** data as a mixture of 2 diastereomers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.94, 8.90 (2d,  $J$  = 0.9 Hz, 1H, H<sub>10</sub>), 7.87 – 7.70, 7.69 – 7.42 (2m, 8H, 5CH<sub>Ar</sub> + H<sub>f</sub> + H<sub>k</sub> + C<sub>9</sub>), 7.32 – 7.23 (m, 1H, CH<sub>Ar</sub>), 7.12 – 7.00 (m, 1H, CH<sub>Ar</sub>), 6.71, 6.54 (2s, 1H, H<sub>4</sub>), 6.47, 6.46 (2d,  $J$  = 8.5 Hz, 1H, H<sub>j</sub>), 6.20, 6.16 (2d,  $J$  = 9.2 Hz, 1H, H<sub>e</sub>), 4.78 – 4.72 (m, 1H, H<sub>6</sub>), 4.11 – 3.81 (m, 2H, H<sub>c</sub>), 3.52, 3.43 (2d,  $J_{dia1}$  = 5.6 and  $J_{dia2}$  = 7.3 Hz, 1H, H<sub>7</sub>), 3.38 – 3.17, 3.08 – 2.93 (2m, 2H, H<sub>a</sub>), 1.65 – 1.45, 0.96 – 0.70 (2m, 2H, H<sub>b</sub>), 1.33, 1.32 (2m, 9H, H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 172.3, 171.9 (C<sub>5</sub>), 169.9, 169.4 (C<sub>3</sub>), 165.6 (C<sub>d</sub>), 161.6, 161.5 (C<sub>i</sub>), 150.4, 150.3 (C<sub>n</sub>), 140.4, 140.4 (C<sub>j</sub>), 139.2, 139.1 (C<sub>k</sub>), 136.4, 136.1 (C<sub>10</sub>), 134.7, 134.6, 133.7, 133.4 (2C<sub>IV,Ar</sub>), 131.3, 131.0, 131.0, 131.0, 130.2, 130.1, 128.9, 128.5, 128.1, 128.1, 127.5, 127.2, 127.0 (C<sub>IV,Ar</sub> + 5CH<sub>Ar</sub> + C<sub>8</sub>), 125.8, 125.5, 123.9, 123.3, 120.6, 119.9 (2CH<sub>Ar</sub> + C<sub>9</sub>), 114.6, 114.5 (C<sub>e</sub>), 108.7 (C<sub>g</sub>), 106.8, 106.8 (C<sub>j</sub>), 60.8, 60.4 (C<sub>4</sub>), 52.8, 52.6 (C<sub>2</sub>), 51.4, 51.0 (C<sub>6</sub>), 45.0, 44.8 (C<sub>a</sub>), 38.8, 38.6 (C<sub>c</sub>), 30.8, 30.5 (C<sub>b</sub>), 28.8, 28.7 (C<sub>1</sub>), 28.1, 27.5 (C<sub>7</sub>); **HRMS (ESI):**  $m/z$  595.31451 [M+H]<sup>+</sup> (C<sub>33</sub>H<sub>39</sub>N<sub>8</sub>O<sub>3</sub> requires 595.31396).

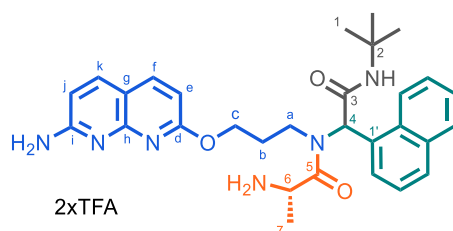
**(Boc)-Naphthyridine(O-alkyl)-(R/S)-Naphthalene-(Boc)-L-Ala-*tert*-butylacetamide (**C23a**):**

Compound **C23a** was prepared following general procedure A using **Boc-naphthyridine(O-alkyl)-NH<sub>2</sub> (37)** (31.8 mg, 0.10 mmol) as the amine, **1-naphthaldehyde** (13.6  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-Ala-OH** (18.9 mg, 0.10 mmol) as the acid and **tert-butyl isocyanide** (11.3  $\mu$ L, 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 98:2, v/v), a mixture of diastereomers **C23a** was obtained as a yellowish solid (54.7 mg, 75%).



$R_{f, \text{dia}1+2} = 0.22$  ( $\text{CH}_2\text{Cl}_2$ -MeOH 98:2); NMR data as a mixture of 2 diastereomers and conformers;  $^1\text{H NMR}$  (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 8.40 - 8.30$ ,  $7.91 - 7.84$ ,  $7.82 - 7.40$ ,  $7.39 - 7.25$  (4m, 7H,  $7\text{CH}_{\text{Ar}}$ ),  $8.15 - 8.08$ ,  $8.07 - 8.03$ ,  $8.02 - 7.92$  (3m, 3H,  $H_f + H_{j-k}$ ),  $6.88$ ,  $6.77$ ,  $6.62$  (3s, 1H,  $H_4$ ),  $6.72$ ,  $6.69$ ,  $6.55$  (3d,  $J = 8.7$  Hz, 1H,  $H_e$ ),  $4.97 - 4.87$ ,  $4.67 - 4.45$  (2m, 1H,  $H_6$ ),  $4.10 - 3.81$  (m, 2H,  $H_c$ ),  $3.74 - 3.58$ ,  $3.57 - 3.27$  (2m, 2H,  $H_a$ ),  $1.90 - 1.72$ ,  $1.66 - 1.26$ ,  $0.95 - 0.79$ ,  $0.65 - 0.44$  (4m, 2H,  $H_b$ ),  $1.68 - 1.20$  (m, 30H,  $6\text{CH}_{3, \text{Boc}} + H_1 + H_7$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 176.8$ ,  $176.7$ ,  $176.4$  ( $C_5$ ),  $171.9$ ,  $171.4$  ( $C_3$ ),  $166.0$ ,  $165.9$  ( $C_d$ ),  $157.8$ ,  $157.4$ ,  $157.3$ ,  $155.6$ ,  $155.5$  ( $2\text{CO}_{\text{Boc}}$ ),  $155.1$ ,  $155.1$ ,  $154.2$ ,  $154.2$  ( $C_{h+i}$ ),  $140.3$ ,  $140.3$  ( $C_f$ ),  $140.0$ ,  $140.0$ ,  $139.9$  ( $C_k$ ),  $135.0$ ,  $134.9$ ,  $134.0$ ,  $133.8$ ,  $133.1$ ,  $133.0$ ,  $132.3$  ( $3\text{C}_{\text{IV, Ar}}$ ),  $130.0$ ,  $130.4$ ,  $130.1$ ,  $129.7$ ,  $128.7$ ,  $128.5$ ,  $128.4$ ,  $128.1$ ,  $128.0$ ,  $127.9$ ,  $127.2$ ,  $127.2$ ,  $126.2$ ,  $126.0$ ,  $125.3$ ,  $124.9$ ,  $123.8$  ( $7\text{CH}_{\text{Ar}}$ ),  $117.2$ ,  $117.1$  ( $C_g$ ),  $113.1$ ,  $113.1$ ,  $113.0$ ,  $112.9$  ( $C_e$ ),  $112.0$ ,  $112.0$ ,  $111.8$  ( $C_j$ ),  $82.1$ ,  $82.1$ ,  $82.0$ ,  $80.7$ ,  $80.4$ ,  $80.2$  ( $2\text{C}_{\text{IV, Boc}}$ ),  $65.3$ ,  $64.7$ ,  $64.5$  ( $C_c$ ),  $62.5$ ,  $60.2$ ,  $59.2$  ( $C_4$ ),  $52.6$ ,  $52.5$ ,  $52.3$  ( $C_2$ ),  $48.2$ ,  $47.7$  ( $C_6$ ),  $44.5$ ,  $44.5$ ,  $44.2$  ( $C_a$ ),  $30.7$ ,  $30.5$ ,  $30.4$ ,  $30.2$  ( $C_b$ ),  $28.9$ ,  $28.8$ ,  $28.8$ ,  $28.6$  ( $6\text{CH}_{3, \text{Boc}} + C_1$ ),  $19.0$ ,  $18.4$ ,  $18.1$  ( $C_7$ ); **MS (ESI)**:  $m/z$  729.00  $[\text{M}+\text{H}]^+$  (theoretical  $m/z$  729.88).

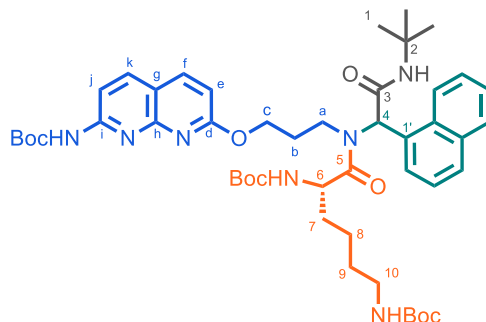
### Naphthyridine(*O*-alkyl)-(R/S)-Naphthalene-L-Ala-*tert*-butylacetamide (**C23b**):



General procedure B was employed for the deprotection of compound **C23a** (38.0 mg, 0.052 mmol) in 1 mL of  $\text{CH}_2\text{Cl}_2$  and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C23b** was isolated as a yellow-pale powder (34.8 mg, 88%).

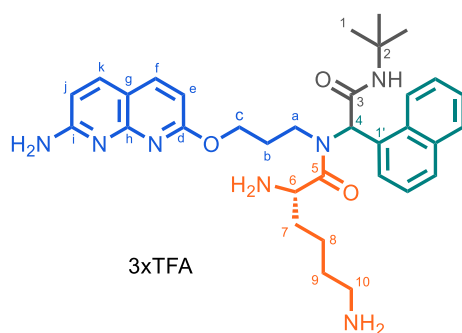
**HPLC**:  $rt_{\text{dia}1} = 15.33$  min and  $rt_{\text{dia}2} = 15.61$  min (method A); NMR data as a mixture of 2 diastereomers and conformers;  $^1\text{H NMR}$  (400 MHz,  $\text{D}_2\text{O}$ )  $\delta = 8.23$ ,  $8.18$  (2d,  $J = 9.2$  Hz, 1H,  $H_k$ ),  $7.97$ ,  $7.93$ ,  $7.92$ ,  $7.87$  (4d,  $J = 8.7$  Hz, 1H,  $H_f$ ),  $7.76 - 7.68$ ,  $7.63 - 7.52$ ,  $7.51 - 7.25$ ,  $7.23 - 7.15$  (4m, 6H,  $6\text{CH}_{\text{Ar}}$ ),  $7.10 - 6.91$  (m, 2H,  $\text{CH}_{\text{Ar}} + H_j$ ),  $6.57$ ,  $6.34$ ,  $6.21$ ,  $6.11$  (4s, 1H,  $H_4$ ),  $6.49$ ,  $6.46$ ,  $6.40$ ,  $6.31$  (4d,  $J = 8.7$  Hz, 1H,  $H_e$ ),  $4.95$ ,  $4.59$ ,  $4.49$  (3q,  $J = 6.8$  Hz, 1H,  $H_6$ ),  $4.29 - 4.07$ ,  $3.97 - 3.84$ ,  $3.83 - 3.63$  (3m, 2H,  $H_c$ ),  $3.56 - 3.34$ ,  $3.33 - 3.14$  (2m, 2H,  $H_a$ ),  $1.85 - 1.72$ ,  $1.52 - 1.35$ ,  $1.02 - 0.86$  (3m, 2H,  $H_b$ ),  $1.79$ ,  $1.67$ ,  $1.56$  (3d,  $J = 6.8$  Hz, 3H,  $H_7$ ),  $1.30$ ,  $1.30$ ,  $1.26$  (3m, 9H,  $H_1$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{D}_2\text{O}$ )  $\delta = 172.0$ ,  $171.4$  ( $C_3$ ),  $170.9$ ,  $170.4$  ( $C_5$ ),  $164.4$ ,  $164.1$  ( $C_d$ ),  $154.9$  ( $C_i$ ),  $145.2$ ,  $145.1$  ( $C_h$ ),  $143.3$ ,  $143.1$  ( $C_k$ ),  $139.8$ ,  $139.8$  ( $C_f$ ),  $132.6$ ,  $132.6$ ,  $131.7$ ,  $131.4$  ( $2\text{C}_{\text{IV, Ar}}$ ),  $129.6$ ,  $129.3$ ,  $128.8$ ,  $128.7$ ,  $128.5$ ,  $128.3$ ,  $128.1$ ,  $127.4$ ,  $127.4$  ( $\text{C}_{\text{IV, Ar}} + 4\text{CH}_{\text{Ar}}$ ),  $126.1$ ,  $126.0$ ,  $125.0$ ,  $124.9$ ,  $122.5$ ,  $121.9$  ( $3\text{CH}_{\text{Ar}}$ ),  $111.3$ ,  $111.3$  ( $C_g$ ),  $111.1$ ,  $111.0$  ( $C_e$ ),  $110.4$  ( $C_j$ ),  $64.1$ ,  $63.9$  ( $C_c$ ),  $60.7$ ,  $58.9$  ( $C_4$ ),  $52.0$ ,  $51.8$ ,  $51.7$  ( $C_2$ ),  $47.5$ ,  $47.2$  ( $C_6$ ),  $43.4$ ,  $42.8$  ( $C_a$ ),  $29.2$ ,  $28.4$  ( $C_b$ ),  $27.6$ ,  $27.6$ ,  $27.4$  ( $C_1$ ),  $16.9$ ,  $16.4$ ,  $15.8$  ( $C_7$ ); **HRMS (ESI)**:  $m/z$  529.29272  $[\text{M}+\text{H}]^+$  ( $\text{C}_{30}\text{H}_{37}\text{N}_6\text{O}_3$  requires 529.29217).

**(Boc)-Naphthyridine(O-alkyl)-(R/S)-Naphthalene-(Boc)-L-Lys(Boc)-tert-butylacetamide  
(C24a):**



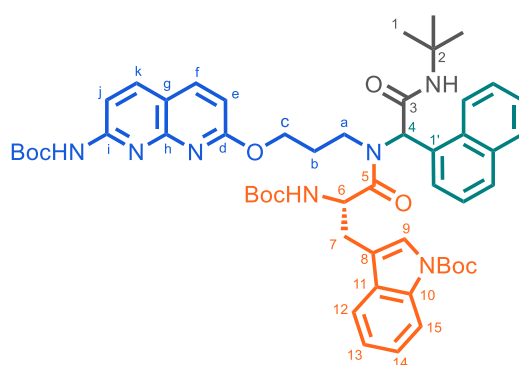
Compound **C24a** was prepared following general procedure A using **Boc-naphthyridine(O-alkyl)-NH<sub>2</sub> (37)** (31.8 mg, 0.10 mmol) as the amine, **1-naphthaldehyde** (13.6  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-Lys(Boc)-OH** (34.6 mg, 0.10 mmol) as the acid and **tert-butyl isocyanide** (11.3  $\mu$ L, 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C24a** was obtained as a white solid (36.6 mg, 41%).

$R_{f, dia1+2} = 0.20$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 98:2); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.20 – 8.13, 8.12 – 7.97 (2m, 3H, H<sub>f</sub> + H<sub>j-k</sub>), 8.33 – 8.26, 7.92 – 7.43, 7.42 – 7.29 (3m, 7H, 7CH<sub>Ar</sub>), 6.91, 6.77, 6.59 (3s, 1H, H<sub>4</sub>), 6.69, 6.75, 6.61 (2d,  $J$  = 8.7 Hz, 1H, H<sub>e</sub>), 4.63 – 4.46 (2m, 1H, H<sub>6</sub>), 4.12 – 3.85 (2m, 2H, H<sub>c</sub>), 3.79 – 3.63, 3.62 – 3.45 (2m, 2H, H<sub>a</sub>), 3.12 – 3.02, 3.01 – 2.92 (2m, 2H, H<sub>10</sub>), 1.91 – 1.80, 1.63 – 1.23, 0.90 – 0.84, 0.67 – 0.53 (4m, 44H, 9CH<sub>3, Boc</sub> + H<sub>b</sub> + H<sub>1</sub> + H<sub>7-9</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 175.9, 175.8 (C<sub>5</sub>), 171.9, 171.5 (C<sub>3</sub>), 166.0, 165.9 (C<sub>d</sub>), 158.4, 157.6, 155.5 (3CO<sub>Boc</sub>), 155.1 (C<sub>i</sub>), 154.2 (C<sub>h</sub>), 140.3, 140.0 (C<sub>k</sub> + C<sub>f</sub>), 135.0, 134.9, 134.0, 133.7, 133.1 (3C<sub>IV, Ar</sub>), 130.4, 130.0, 129.7, 128.5, 128.4, 127.9, 127.2, 127.1, 126.2, 126.0, 124.7, 123.9 (7CH<sub>Ar</sub>), 117.2, 117.1 (C<sub>g</sub>), 113.3, 113.2 (C<sub>e</sub>), 112.1, 112.0 (C<sub>j</sub>), 82.0, 80.7, 80.4, 80.2, 79.7 (3C<sub>IV, Boc</sub>), 65.3, 64.7, 64.5 (C<sub>c</sub>), 62.3, 60.0, 59.2 (C<sub>4</sub>), 52.9, 52.5, 52.4, 52.2 (C<sub>2</sub> + C<sub>6</sub>), 44.6, 44.4, 44.3 (C<sub>a</sub>), 41.1, 41.0 (C<sub>10</sub>), 33.9, 33.1 (C<sub>7</sub>), 30.8, 30.6, 30.2 (C<sub>b</sub> + C<sub>9</sub>), 28.9, 28.8, 28.8, 28.7, 28.5 (9CH<sub>3, Boc</sub> + C<sub>1</sub>), 24.3, 24.2 (C<sub>8</sub>); **MS (ESI):**  $m/z$  886.13 [M+H]<sup>+</sup> (theoretical  $m/z$  886.51).

**Naphthyridine(*O*-alkyl)-(R/S)-Naphthalene-L-Lys-*tert*-butylacetamide (**C24b**):**

General procedure B was employed for the deprotection of compound **C24a** (30.0 mg, 0.037 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C24b** was isolated as a yellow-pale powder (23.9 mg, 70%).

**HPLC:**  $rt_{dia1}$  = 14.11 min and  $rt_{dia2}$  = 14.63 min (method A); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta$  = 8.33, 8.28 (2d,  $J$  = 9.3 Hz, 1H, H<sub>k</sub>), 8.08, 8.06 (2d,  $J$  = 8.7 Hz, 1H, H<sub>f</sub>), 7.80 – 7.21, 7.15 – 6.98 (2m, 7H, 7CH<sub>Ar</sub>), 7.15 – 6.98 (m, 1H, H<sub>j</sub>), 6.65 – 6.59 (m, 1H, H<sub>e</sub>), 6.63, 6.40, 6.22 (3s, 1H, H<sub>4</sub>), 4.99, 4.57, 4.52 (3t,  $J$  = 6.2 Hz, 1H, H<sub>6</sub>), 4.37 – 4.22, 4.04 – 3.94, 3.88 – 3.79 (3m, 2H, H<sub>c</sub>), 3.59 – 3.49, 3.41 – 3.28 (2m, 2H, H<sub>a</sub>), 3.12 – 2.99 (m, 2H, H<sub>10</sub>), 2.17 – 2.00 (m, 2H, H<sub>7</sub>), 1.87 – 1.43 (m, 4H, H<sub>8-9</sub>), 1.33, 1.31, 1.28 (s, 9H, H<sub>1</sub>), 1.87 – 1.75, 1.75 – 1.43, 1.09 – 0.92 (m, 2H, H<sub>b</sub>); <sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O)  $\delta$  = 171.0, 170.9 (C<sub>3</sub>), 170.3, 170.2 (C<sub>5</sub>), 164.5, 164.1 (C<sub>d</sub>), 155.0 (C<sub>i</sub>), 145.3 (C<sub>h</sub>), 143.5, 143.4 (C<sub>k</sub>), 140.1 (C<sub>f</sub>), 132.8, 132.6, 131.7, 131.4 (2C<sub>IV,Ar</sub>), 129.6, 129.3, 128.7, 128.3, 128.1, 127.6, 127.4 (C<sub>IV,Ar</sub> + 4CH<sub>Ar</sub>), 126.2, 126.0, 125.1, 122.6, 121.9 (3CH<sub>Ar</sub>), 117.7, 114.8 (C<sub>g</sub>), 111.9, 111.5, 111.2, 111.0 (C<sub>e</sub>), 110.4 (C<sub>j</sub>), 64.1, 63.9 (C<sub>c</sub>), 60.6, 58.9 (C<sub>4</sub>), 52.0, 51.8, 51.6 (C<sub>2</sub>), 51.3, 51.1, 50.8 (C<sub>6</sub>), 43.7, 42.9 (C<sub>a</sub>), 39.0 (C<sub>10</sub>), 30.9, 30.5 (C<sub>7</sub>), 29.2, 28.5 (C<sub>b</sub>), 27.5, 27.5, 27.4 (C<sub>1</sub>), 26.6 (C<sub>9</sub>), 21.7, 21.0 (C<sub>8</sub>); **HRMS (ESI):**  $m/z$  586.35065 [M+H]<sup>+</sup> (C<sub>33</sub>H<sub>44</sub>N<sub>7</sub>O<sub>3</sub> requires 586.35001).

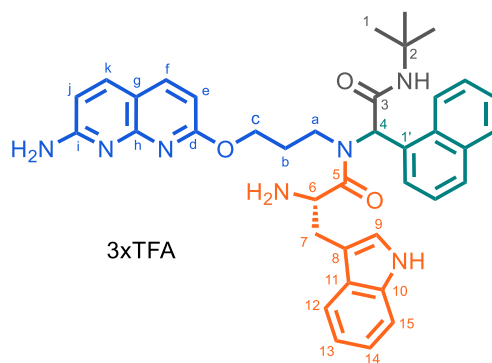
**(Boc)-Naphthyridine(*O*-alkyl)-(R/S)-Naphthalene-(Boc)-L-Trp(Boc)-*tert*-butylacetamide (**C25a**):**

Compound **C25a** was prepared following general procedure A using **Boc-naphthyridine(*O*-alkyl)-NH<sub>2</sub> (37)** (31.8 mg, 0.10 mmol) as the amine, **1-naphthaldehyde** (13.6  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-**

**Trp(Boc)-OH** (40.4 mg, 0.10 mmol) as the acid and **tert-butyl isocyanide** (11.3  $\mu$ l, 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography (cHex/EtOAc gradient 100:0 to 75:25, v/v), a mixture of diastereomers **C25a** was obtained as a white solid (84.5 mg, 90%).

$R_{f,dia1+2} = 0.26$  (cHex-EtOAc 70:30); NMR data as a mixture of 2 diastereomers and conformers;  $^1\text{H NMR}$  (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 8.19 - 7.98, 7.97 - 7.81, 7.80 - 7.59$  (3m, 7H,  $4\text{CH}_{Ar} + H_f + H_{j-k}$ ),  $7.58 - 6.95$  (m, 8H,  $8\text{CH}_{Ar}$ ),  $6.90, 6.62, 6.56$  (3s, 1H,  $H_4$ ),  $6.69, 6.57, 6.32$  (3d,  $J = 8.7$  Hz, 1H,  $H_e$ ),  $5.24 - 5.18, 5.10 - 4.99, 4.96 - 4.77$  (m, 1H,  $H_6$ ),  $4.19 - 4.12, 4.06 - 3.97, 3.93 - 3.83, 3.82 - 3.62$  (4m, 2H,  $H_c$ ),  $3.82 - 3.62, 3.63 - 3.36$  (2m, 2H,  $H_a$ ),  $3.35 - 3.26, 3.19 - 2.94$  (2m, 2H,  $H_7$ ),  $1.76 - 1.11, 1.10 - 0.84$  (m, 36H,  $9\text{CH}_{3.Boc} + H_1$ ),  $1.76 - 1.11, 0.68 - 0.51, 0.50 - 0.35$  (3m, 2H,  $H_b$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 175.1, 174.6, 173.9, 172.9, 171.9, 171.8$  ( $C_3 + C_5$ ),  $166.0, 165.9$  ( $C_d$ ),  $158.1, 157.3, 157.3, 155.4, 155.4, 155.0, 155.0, 154.2, 154.2, 150.9$  ( $3\text{CO}_{Boc}, C_h + C_i$ ),  $139.9, 139.9$  ( $C_f + C_k$ ),  $136.7, 134.9, 134.8, 134.0, 133.7, 133.1, 132.7, 131.9, 131.7$  ( $5\text{C}_{IV.Ar}$ ),  $130.5, 130.5, 130.4, 129.9, 129.8, 128.5, 128.4, 128.2, 128.0, 127.8, 127.2, 127.2, 127.0, 126.1, 126.0, 125.4, 125.4, 125.2, 124.6, 123.7, 123.7, 123.6$  ( $10\text{CH}_{Ar}$ ),  $120.3, 120.0, 117.7$  ( $\text{CH}_{Ar}$ ),  $117.7, 117.5, 117.2, 117.1$  ( $C_g + C_8$ ),  $116.2$  ( $C_{15}$ ),  $113.3, 113.2, 113.0$  ( $C_e$ ),  $111.9, 111.8$  ( $C_j$ ),  $83.4, 83.2, 80.6, 79.6, 79.2, 78.9$  ( $3\text{C}_{IV.Boc}$ ),  $65.3, 64.6, 64.4$  ( $C_c$ ),  $59.8, 59.5$  ( $C_4$ ),  $52.8, 52.6, 52.4, 52.3, 51.9$  ( $C_2 + C_6$ ),  $44.5, 44.1$  ( $C_a$ ),  $30.7, 30.4, 29.7$  ( $C_b$ ),  $28.9, 28.9, 28.8, 28.7, 28.6, 28.5, 28.5, 27.9$  ( $9\text{CH}_{3.Boc} + C_1 + C_7$ ); **MS (ESI)**:  $m/z$  944.00  $[\text{M}+\text{H}]^+$  (theoretical  $m/z$  944.49).

### Naphthyridine(O-alkyl)-(R/S)-Naphthalene-L-Trp-tert-butylacetamide (**C25b**):

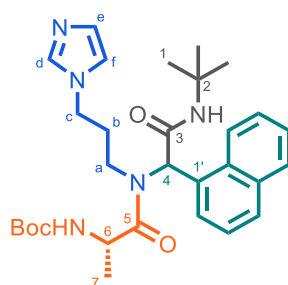


General procedure B was employed for the deprotection of compound **C25a** (50.0 mg, 0.053 mmol) in 1 mL of  $\text{CH}_2\text{Cl}_2$  and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C25b** was isolated as a brown-pale powder (42.5 mg, 82%).

**HPLC**:  $rt_{dia1} = 17.96$  min and  $rt_{dia2} = 18.14$  min (method A); NMR data as a mixture of 2 diastereomers;  $^1\text{H NMR}$  (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 8.23, 8.21$  (2d,  $J = 9.2$  Hz, 1H,  $H_k$ ),  $8.03, 7.98$  (2d,  $J = 8.7$  Hz, 1H,  $H_f$ ),  $8.00 - 7.95, 7.76 - 7.68, 7.63 - 7.48, 7.42 - 7.36, 7.33 - 7.26$  (9m, 5H,  $5\text{CH}_{Ar}$ ),  $7.19 - 7.12, 7.10 - 7.00, 6.98 - 6.90$  (3m, 4H,  $2\text{CH}_{Ar} + H_j + H_9$ ),  $6.56, 6.50$  (2d, 1H,  $H_e$ ),  $7.28, 6.50$  (2s, 1H,  $H_4$ ),  $4.67, 4.56$  (2dd,  $J_{dia1} = 9.2, 5.6$  Hz and  $J_{dia2} = 8.2, 5.0$  Hz, 1H,  $H_6$ ),  $4.06 - 3.99, 3.79 - 3.70$  (m, 2H,  $H_c$ ),  $3.78 - 3.33, 3.32 - 3.12,$

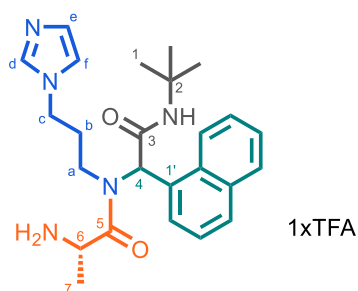
3.06 – 2.94 (3m, 4H, H<sub>a</sub> + H<sub>7</sub>), 1.38, 1.35 (2s, 9H, H<sub>1</sub>), 1.43 – 1.26, 1.29 – 1.27, 1.01 – 0.84, 0.82 – 0.67 (3m, 2H, H<sub>b</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD) δ = 172.5, 171.5, 171.5, 171.0 (C<sub>3</sub> + C<sub>5</sub>), 166.6, 166.5 (C<sub>d</sub>), 157.3, 157.2 (C<sub>h</sub>), 147.5, 147.5 (C<sub>i</sub>), 144.3, 144.3 (C<sub>k</sub>), 140.8 (C<sub>f</sub>), 138.3, 138.1 (C<sub>10</sub>), 134.9, 134.7, 133.9, 133.7, 132.3, 131.7 (3C<sub>IV,Ar</sub>), 130.5, 129.7, 129.6 (2C<sub>HA,r</sub>), 128.9, 128.7, 128.6, 128.3, 128.2, 127.4, 127.4 (C<sub>IV,Ar</sub> + 3C<sub>HA,r</sub>), 126.2, 126.1, 125.9, 125.6 (C<sub>HA,r</sub> + C<sub>9</sub>), 124.5, 123.9, 122.9, 122.9, 120.4, 118.8, 118.6 (4C<sub>HA,r</sub>), 112.8, 112.8 (C<sub>15</sub>), 112.6, 112.5 (C<sub>g</sub>), 112.3, 112.1 (C<sub>e</sub>), 111.6, 111.6 (C<sub>j</sub>), 108.0, 107.6 (C<sub>8</sub>), 65.4, 65.4 (C<sub>c</sub>), 61.0, 59.4 (C<sub>4</sub>), 53.2, 52.6, 52.4, 52.2 (C<sub>2</sub> + C<sub>6</sub>), 44.1, 43.9 (C<sub>a</sub>), 30.4, 30.1 (C<sub>b</sub>), 29.2, 28.8, 28.7, 28.5 (C<sub>1</sub> + C<sub>7</sub>); HRMS (ESI): *m/z* 644.33453 [M+H]<sup>+</sup> (C<sub>38</sub>H<sub>42</sub>N<sub>7</sub>O<sub>3</sub> requires 644.33436).

### Imidazole-(R/S)-Naphthalene-(Boc)-L-Ala-*tert*-butylacetamide (**C26a**):



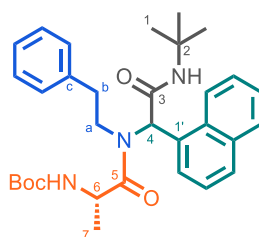
Compound **C26a** was prepared following general procedure A using **1-(3-aminopropyl)imidazole** (21.0 μL, 0.18 mmol) as the amine, **1-naphthaldehyde** (24.7 μL, 0.18 mmol) as aldehyde, **Boc-L-Ala-OH** (34.1 mg, 0.18 mmol) as the acid and **tert-butyl isocyanide** (20.4 μL, 0.18 mmol) in MeOH (0.7 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C26a** was obtained as a white solid (78.0 mg, 81%).

*R*<sub>f,dia1+2</sub> = 0.36 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ = 8.60 – 8.49, 8.09 – 7.81, 7.76 – 7.37 (3m, 7H, C<sub>HA,r</sub>), 7.25, 7.18, 7.07 (3s, 1H, H<sub>d</sub>), 6.90, 6.75, 6.69 (3s, 1H, H<sub>4</sub>), 6.84, 6.82, 6.74 (3br s, 1H, H<sub>e</sub>), 6.67, 6.58, 6.45 (3br s, 1H, H<sub>f</sub>), 4.92, 4.45, 4.29 (3q, *J* = 6.7 Hz, 1H, H<sub>6</sub>), 3.74 – 3.13 (m, 4H, H<sub>a</sub> + H<sub>c</sub>), 1.90 – 1.57, 1.41 – 1.11, 1.06 – 0.84, 0.55 – 0.37 (4m, 2H, H<sub>b</sub>), 1.51, 1.47, 1.39, 1.33 (4s, 18H, 3C<sub>H<sub>3</sub>,Boc</sub> + H<sub>1</sub>), 1.41 – 1.13 (m, 3H, H<sub>7</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD) δ = 177.3, 176.5, 176.3 (C<sub>5</sub>), 171.8, 171.8, 171.2 (C<sub>3</sub>), 157.9, 157.5, 157.4 (CO<sub>Boc</sub>), 138.0, 137.7 (C<sub>d</sub>), 135.1, 135.0, 134.0, 133.9, 133.7, 133.3, 133.0, 132.4 (3C<sub>IV,Ar</sub>), 131.0, 130.6, 130.5, 130.1, 129.9, 129.8, 128.9, 128.8, 128.6, 128.5, 128.5, 128.1, 128.2, 128.1, 128.0, 127.5, 127.4, 126.6, 126.4, 125.5, 125.0, 123.9 (7C<sub>HA,r</sub> + H<sub>e</sub>), 120.4, 120.0 (C<sub>f</sub>), 80.6, 80.5, 80.3 (C<sub>IV,Boc</sub>), 62.2, 60.1, 58.9 (C<sub>4</sub>), 52.5, 52.4, 52.3 (C<sub>2</sub>), 48.2, 47.9 (C<sub>6</sub>), 45.6, 45.1, 44.0, 43.8, 43.7 (C<sub>a</sub> + C<sub>c</sub>), 32.4, 31.7, 31.3 (C<sub>b</sub>), 28.9, 28.8, 28.8 (3C<sub>H<sub>3</sub>,Boc</sub> + C<sub>1</sub>), 19.0, 18.1, 17.8 (C<sub>7</sub>); MS (ESI): *m/z* 536.6 [M+H]<sup>+</sup> (theoretical *m/z* 536.32).

**Imidazole-(R/S)-Naphthalene-L-Ala-*tert*-butylacetamide (C26b):**

General procedure B was employed for the deprotection of compound **C26a** (60.0 mg, 0.112 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C26b** was isolated as a white powder (50.9 mg, 83%).

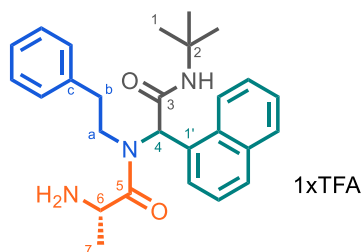
**HPLC:**  $rt_{dia1}$  = 12.93 min and  $rt_{dia2}$  = 13.01 min (method A); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, D<sub>2</sub>O)  $\delta$  = 8.05 – 7.85 (m, 3H, 2CH<sub>Ar</sub> + H<sub>d</sub>), 7.84 – 7.73, 7.70 – 7.56 (m, 3H, 3CH<sub>Ar</sub>), 7.50 – 7.37, 7.36 – 7.26 (2m, 2H, 2CH<sub>Ar</sub>), 7.20, 7.17, 7.14 (3d,  $J$  = 1.5 Hz, 1H, H<sub>f</sub>), 6.78, 6.76, 6.67, 6.64, (4br s, 1H, H<sub>e</sub>), 6.64, 6.40, 6.32, 6.18 (4s, 1H, H<sub>4</sub>), 4.95, 4.48, 4.36, 4.25, (4q,  $J$  = 6.8 Hz, 1H, H<sub>6</sub>), 3.91 – 3.53 (m, 2H, H<sub>c</sub>), 3.34 – 3.09, 3.06 – 2.90, 2.82 – 2.66 (2m, 2H, H<sub>a</sub>), 1.99 – 1.84, 1.78 – 1.47, 1.42 – 1.18, 0.74 – 0.43 (4m, 2H, H<sub>b</sub>), 1.59, 1.58, 1.52 (3d,  $J$  = 6.8 Hz, 3H, H<sub>7</sub>), 1.31, 1.27 (2s, 9H, H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, D<sub>2</sub>O)  $\delta$  = 172.0, 171.4, 171.1, 170.8, 170.3, 170.2 (C<sub>3</sub> + C<sub>5</sub>), 134.2, 134.1 (C<sub>d</sub>), 133.3, 133.2, 131.7, 131.5, 131.3 (2C<sub>IV,Ar</sub>), 130.6, 130.4, 130.3, 129.4, 129.2, 129.1, 129.1, 128.8 (C<sub>IV,Ar</sub> + 2CH<sub>Ar</sub>), 128.1, 128.0, 127.8, 127.4, 126.9, 125.5, 125.4, 122.5, 122.3, 121.9 (5CH<sub>Ar</sub>), 121.1, 121.0, 120.9, 120.7 (C<sub>e</sub> + C<sub>f</sub>), 60.8, 60.6, 59.0 (C<sub>4</sub>), 52.1, 51.8, 51.7 (C<sub>2</sub>), 47.6, 47.4 (C<sub>6</sub>), 46.1, 45.7, 45.5 (C<sub>c</sub>), 42.2, 42.2 (C<sub>a</sub>), 30.0, 29.9 (C<sub>b</sub>), 27.6, 27.6, 27.4 (C<sub>1</sub>), 16.7, 16.4, 15.9 (C<sub>7</sub>); **HRMS (ESI):**  $m/z$  436.27100 [M+H]<sup>+</sup> (C<sub>25</sub>H<sub>34</sub>N<sub>5</sub>O<sub>2</sub> requires 436.27100).

**Phenyl-(R/S)-Naphthalene-(Boc)-L-Ala-*tert*-butylacetamide (C27a):**

Compound **C27a** was prepared following general procedure A using **2-phenylethylamine** (22.3 mg, 0.184 mmol) as the amine, **1-naphthaldehyde** (25.0  $\mu$ L, 0.184 mmol) as aldehyde, **Boc-L-Ala-OH** (34.8 mg, 0.184 mmol) as the acid and ***tert*-butyl isocyanide** (20.8  $\mu$ L, 0.184 mmol) in MeOH (1.0 mL). After purification by flash chromatography (cHex/EtOAc gradient 85:15 to 65:35, v/v), a mixture of diastereomers **C27a** was obtained as a white solid (48.0 mg, 49%).

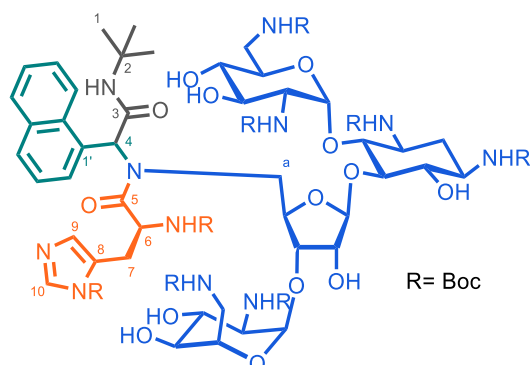
$R_{f, \text{dia}1+2} = 0.56$  (cHex-EtOAc 7:3); NMR data as a mixture of 2 diastereomers and conformers;  $^1\text{H NMR}$  (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 8.53 - 8.45$ ,  $8.12 - 7.79$ ,  $7.71 - 7.50$  (3m, 7H,  $7\text{CH}_{\text{Ar}}$ ),  $7.14 - 6.97$ ,  $6.62 - 4.40$  (2m, 5H,  $5\text{CH}_{\text{phenyl}}$ ),  $6.98$ ,  $6.95$ ,  $6.84$ ,  $6.72$  (4s, 1H,  $\text{H}_4$ ),  $5.02 - 4.92$ ,  $4.73 - 4.50$  (2m, 1H,  $\text{H}_6$ ),  $3.74 - 3.33$ ,  $3.22 - 3.11$  (m, 2H,  $\text{H}_a$ ),  $2.70 - 2.27$ ,  $1.64 - 1.10$  (2m, 2H,  $\text{H}_b$ ),  $1.64 - 1.10$  (m, 21H,  $3\text{CH}_{3, \text{Boc}} + \text{H}_1 + \text{H}_7$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 176.7$ ,  $176.5$ ,  $176.2$  ( $\text{C}_5$ ),  $172.0$ ,  $171.6$ ,  $171.5$  ( $\text{C}_3$ ),  $157.9$ ,  $157.6$ ,  $157.5$  ( $\text{CO}_{\text{Boc}}$ ),  $140.5$ ,  $139.6$ ,  $139.1$  ( $\text{C}_c$ ),  $135.3$ ,  $135.3$ ,  $135.2$ ,  $134.3$ ,  $134.1$ ,  $134.1$ ,  $133.5$ ,  $132.7$  ( $3\text{C}_{\text{IV, Ar}}$ ),  $131.1$ ,  $130.8$ ,  $130.7$ ,  $130.2$ ,  $129.9$ ,  $129.4$ ,  $129.4$ ,  $129.4$ ,  $129.3$ ,  $129.2$ ,  $129.1$ ,  $128.9$ ,  $128.7$ ,  $128.3$ ,  $128.2$ ,  $128.2$ ,  $127.5$ ,  $127.5$ ,  $127.4$ ,  $127.3$ ,  $126.9$ ,  $126.4$ ,  $126.2$ ,  $125.5$ ,  $125.0$ ,  $124.1$  ( $12\text{CH}_{\text{Ar}}$ ),  $80.7$ ,  $80.6$ ,  $80.4$  ( $\text{C}_{\text{IV, Boc}}$ ),  $62.5$ ,  $60.0$ ,  $59.2$  ( $\text{C}_4$ ),  $52.7$ ,  $52.6$ ,  $52.5$ ,  $52.3$  ( $\text{C}_2$ ),  $\text{C}_a + \text{C}_6$  (signal under the  $\text{CD}_3\text{OD}$  peak),  $47.8$  ( $\text{C}_6$ ),  $37.5$ ,  $37.1$ ,  $35.6$  ( $\text{C}_b$ ),  $28.9$ ,  $28.9$ ,  $28.8$ ,  $28.8$ ,  $28.8$ ,  $28.8$  ( $3\text{CH}_{3, \text{Boc}} + \text{C}_1$ ),  $19.1$ ,  $18.4$ ,  $18.2$  ( $\text{C}_7$ ); **MS (ESI)**:  $m/z$  532.2  $[\text{M}+\text{H}]^+$  (theoretical 532.32).

### Phenyl-(R/S)-Naphthalene-(L)-Ala-*tert*-butylacetamide (**C27b**):



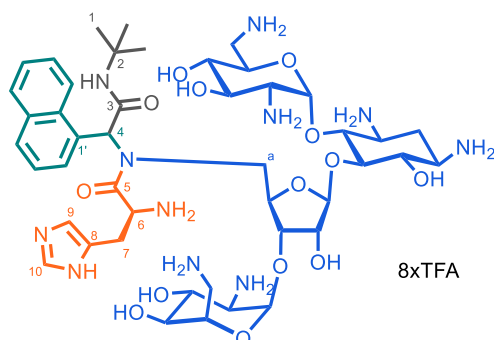
General procedure B was employed for the deprotection of compound **C27a** (22.7 mg, 0.043 mmol) in 1 mL of  $\text{CH}_2\text{Cl}_2$  and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C27b** was isolated as a white powder (12.2 mg, 53%).

**HPLC**:  $\text{rt}_{\text{dia}1} = 20.71\text{min}$  and  $\text{rt}_{\text{dia}2} = 21.22\text{ min}$  (method A); NMR data as a mixture of 2 diastereomers;  $^1\text{H NMR}$  (400 MHz,  $\text{D}_2\text{O}$ )  $\delta = 8.11 - 7.86$ ,  $7.76 - 7.54$  (2m, 7H,  $7\text{CH}_{\text{Ar}}$ ),  $7.17 - 7.06$  (m, 3H,  $3\text{CH}_{\text{phenyl}}$ ),  $6.83$ ,  $6.58$  (2s, 1H,  $\text{H}_4$ ),  $6.60 - 6.46$  (m, 2H,  $2\text{CH}_{\text{phenyl}}$ ),  $4.51$ ,  $4.18$  (2q,  $J = 6.8\text{ Hz}$ , 1H,  $\text{H}_6$ ),  $3.56 - 3.28$ ,  $3.21 - 3.07$  (2m, 2H,  $\text{H}_a$ ),  $2.60 - 2.47$ ,  $2.14 - 2.04$ ,  $2.01 - 1.89$ ,  $1.51 - 1.41$  (4m, 2H,  $\text{H}_b$ ),  $1.58$ ,  $1.55$  (2d,  $J = 6.9\text{ Hz}$ , 3H,  $\text{H}_7$ ),  $1.34$ ,  $1.31$  (2s, 9H,  $\text{H}_1$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{D}_2\text{O}$ )  $\delta = 171.9$ ,  $171.2$  ( $\text{C}_5$ ),  $170.8$ ,  $170.3$  ( $\text{C}_3$ ),  $137.6$ ,  $137.5$  ( $\text{C}_c$ ),  $133.5$ ,  $133.5$ ,  $132.2$ ,  $131.9$  ( $2\text{C}_{\text{IV, Ar}}$ ),  $130.4$ ,  $129.8$ ,  $129.6$ ,  $129.3$ ,  $129.2$  ( $\text{C}_{\text{IV, Ar}} + 2\text{CH}_{\text{Ar}}$ ),  $128.7$ ,  $128.6$ ,  $128.6$ ,  $128.2$ ,  $127.9$ ,  $127.9$ ,  $127.8$ ,  $126.8$ ,  $126.7$ ,  $126.7$ ,  $125.6$ ,  $122.7$ ,  $122.1$  ( $10\text{CH}_{\text{Ar}}$ ),  $60.4$ ,  $58.8$  ( $\text{C}_4$ ),  $51.9$ ,  $51.8$  ( $\text{C}_2$ ),  $47.7$ ,  $47.5$  ( $\text{C}_6$ ),  $47.2$ ,  $47.1$  ( $\text{C}_a$ ),  $35.5$ ,  $35.3$  ( $\text{C}_b$ ),  $27.7$ ,  $27.6$  ( $\text{C}_1$ ),  $16.7$ ,  $16.2$  ( $\text{C}_7$ ); **HRMS (ESI)**:  $m/z$  432.26480  $[\text{M}+\text{H}]^+$  ( $\text{C}_{27}\text{H}_{34}\text{N}_3\text{O}_2$  requires 432.26455).

**(Boc)-Neo-(R/S)-Naphthalene-(Boc)-L-His(Boc)-tert-butylacetamide (C28a):**

Compound **C28a** was prepared following general B procedure using **Boc-Neo-NH<sub>2</sub> (57)** (97.2 mg, 0.08 mmol) as the amine, **1-naphthaldehyde** (10.9  $\mu$ L, 0.08 mmol) as aldehyde, **Boc-L-His(Boc)-OH** (42.9 mg, 0.080 mmol) as the acid and **tert-butyl isocyanide** (9.0  $\mu$ L, 0.080 mmol) in MeOH (0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C28b** was obtained as a white solid (47.1 mg, 33%).

$R_f$ .<sub>dia1+2</sub> = 0.50 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.36 – 7.23 (m, 9H, 7CH<sub>Ar</sub> + H<sub>9-10</sub>), 6.65 (s, 1H, H<sub>4</sub>), 5.35 – 4.71 (m, 4H, 3CH + H<sub>6</sub>), 4.34 – 2.53 (m, 24H, 3CH<sub>2</sub> + 16CH + H<sub>7</sub>), 2.11 – 1.84, 1.72 – 0.74 (2m, 83H, 24CH<sub>3.Boc</sub> + CH<sub>2</sub> + H<sub>1</sub>); **MS (ESI)**:  $m/z$  calcd for C<sub>85</sub>H<sub>137</sub>N<sub>11</sub>O<sub>30</sub>: 895.98 [M+2H]<sup>2+</sup>; found: 895.87.

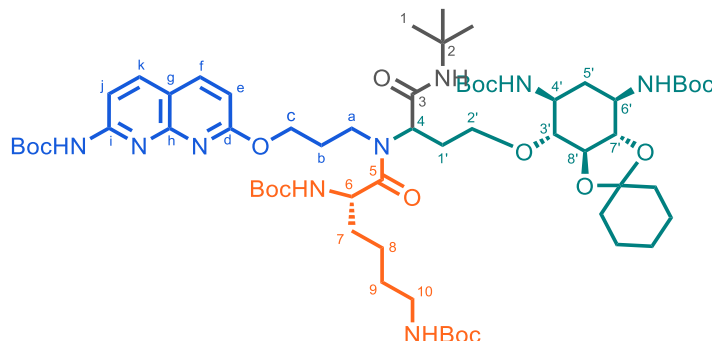
**Neo-(R/S)-Naphthalene-L-His-tert-butylacetamide (C28b):**

General procedure B was employed for the deprotection of compound **C28a** (30.1 mg, 0.017 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C28b** was isolated as a white powder (30.0 mg, 94%).

**HPLC**:  $rt_{dia1+2}$  = 10.20 min (method A); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, D<sub>2</sub>O)  $\delta$  = 8.77, 8.73, 8.57 (s, 1H, H<sub>10</sub>), 8.20 – 7.92, 7.83 – 7.47, 7.43 – 7.19 (3m, 8H, 7CH<sub>Ar</sub> + H<sub>9</sub>), 6.71, 6.38, 6.36 (3s, 1H, H<sub>4</sub>), 5.97 – 5.78, 5.22 – 5.03 (2m, 3H, 3CH), H<sub>4</sub> (signal in D<sub>2</sub>O peak), 4.37 – 2.82 (m, 24H, 4CH<sub>2</sub> + 16CH), 2.55 – 2.40, 2.03 – 1.78 (2m, 2H, CH<sub>2</sub>), 1.41 – 1.23 (m, 9H, H<sub>1</sub>); **HRMS (ESI)**:  $m/z$  990.52625 [M+H]<sup>+</sup> (C<sub>45</sub>H<sub>72</sub>N<sub>11</sub>O<sub>14</sub> requires 990.52547).

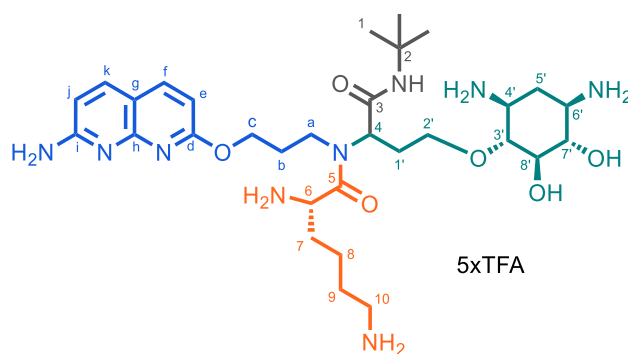


## Série C-2

**(Boc)-Naphthyridine(O-alkyl)-(R/S)-(Boc)-DOS<sub>(Acetal)</sub>-(Boc)-L-Lys(Boc)-tert-butylacetamide****(C29a):**

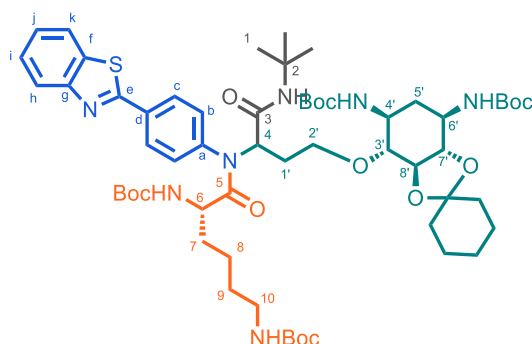
Compound **C29a** was prepared following general procedure A using **Boc-naphthyridine(O-alkyl)-NH<sub>2</sub>** (**37**) (47.7 mg, 0.15 mmol) as the amine, **Boc-DOS<sub>(Acetal)</sub>-CHO** (**58**) (74.7 mg, 0.15 mmol) as aldehyde, **Boc-L-Lys(Boc)-OH** (51.9 mg, 0.15 mmol) as the acid and **tert-butyl isocyanide** (17.0  $\mu$ l, 0.15 mmol) in MeOH (0.3 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C29a** was obtained as a white solid (73.7 mg, 40%).

$R_f$ .dia1+2 = 0.23 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.24 – 8.02 (m, 3H, H<sub>f</sub> + H<sub>j-k</sub>), 7.02 – 6.88 (m, 1H, H<sub>e</sub>), 5.22 – 5.14, 4.73 – 4.35 (2m, 1H, H<sub>4</sub>), 4.73 – 4.35 (m, 3H, H<sub>c</sub> + H<sub>6</sub>), 3.95 – 3.12 (m, 9H, H<sub>a</sub> + H<sub>2'-4'</sub> + H<sub>6'-8'</sub>), 3.23 – 2.92 (m, 2H, H<sub>10</sub>), 2.45 – 1.95, 1.80 – 1.06 (2m, 76H, 5CH<sub>2</sub>.Acetal + 15CH<sub>3</sub>.Boc + H<sub>b</sub> + H<sub>1'</sub> + H<sub>5'</sub> + H<sub>1</sub> + H<sub>7-9</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 176.9, 176.4 (C<sub>5</sub>), 172.0, 171.1 (C<sub>3</sub>), 166.5, 166.4 (C<sub>d</sub>), 158.5, 158.5, 158.2, 157.9, 157.7, 157.7, 155.8, 155.7, 155.5, 155.4, 154.3 (5CO<sub>Boc</sub> + C<sub>h-i</sub>), 140.7, 140.5, 140.2, 140.1 (C<sub>f</sub> + C<sub>k</sub>), 117.4, 117.2 (C<sub>g</sub>), 113.1, 112.9, 112.2, 112.1 (C<sub>IV.Acetal</sub> + C<sub>e</sub> + C<sub>j</sub>), 82.1, 82.1, 81.8, 81.3, 80.4, 80.2, 80.1, 79.8 (5C<sub>IV.Boc</sub> + C<sub>3'</sub> + C<sub>7'-8'</sub>), 68.7, 68.6 (C<sub>2'</sub>), 65.6, 65.0 (C<sub>c</sub>), 58.6 (C<sub>4</sub>), 52.8, 52.3, 52.2, 52.1 (C<sub>4'</sub> or C<sub>6'</sub> + C<sub>2</sub> + C<sub>6</sub>), C<sub>4'</sub> or C<sub>6'</sub> (signal under the CD<sub>3</sub>OD peak), 44.9, 42.6 (C<sub>a</sub>), 41.2, 41.0 (C<sub>10</sub>), 37.5, 37.4 (2CH<sub>2</sub>.Acetal), 33.0, 32.6, 30.7, 30.6, 29.2, 29.0, 28.9, 28.9, 28.9, 28.8, 28.8, 28.6 (15CH<sub>3</sub>.Boc + C<sub>b</sub> + C<sub>1'</sub> + C<sub>5'</sub> + C<sub>1</sub> + C<sub>7</sub> + C<sub>9</sub>), 26.1, 25.0, 24.9, 24.3 (3CH<sub>2</sub>.Acetal + C<sub>8</sub>); **MS (ESI):**  $m/z$  1228.87 [M+H]<sup>+</sup> (theoretical  $m/z$  1228.74).

**Naphthyridine(O-alkyl)-(R/S)-DOS-L-Lys-tert-butylacetamide (C29b):**

General procedure B was employed for the deprotection of compound **C29a** (42.7 mg, 0.035 mmol) in 1 mL of  $\text{CH}_2\text{Cl}_2$  and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C29b** was isolated as a white-pale powder (40.9 mg, 97%).

**HPLC:**  $\text{rt}_{\text{dia1}} = 8.50$  min and  $\text{rt}_{\text{dia2}} = 8.60$  min (method A); NMR data as a mixture of 2 diastereomers and conformers;  $^1\text{H NMR}$  (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 8.30 - 8.23$  (m, 1H,  $\text{H}_k$ ),  $8.20 - 8.11$  (m, 1H,  $\text{H}_f$ ),  $7.02 - 6.90$  (m, 2H,  $\text{H}_e + \text{H}_j$ ),  $5.19 - 5.12$ ,  $4.79 - 4.68$ ,  $4.64 - 4.37$  (3m, 1H,  $\text{H}_4$ ),  $4.79 - 4.68$ ,  $4.64 - 4.37$  (2m, 3H,  $\text{H}_c + \text{H}_6$ ),  $4.17 - 3.50$  (m, 4H,  $\text{H}_a + \text{H}_{2'}$ ),  $3.50 - 3.10$  (m, 5H,  $\text{H}_{3'-4'} + \text{H}_{6'-8'}$ ),  $3.03 - 2.88$  (m, 2H,  $\text{H}_{10}$ ),  $2.47 - 2.35$ ,  $2.01 - 1.64$  (2m, 2H,  $\text{H}_{5'}$ ),  $2.34 - 1.64$  (m, 8H,  $\text{H}_b + \text{H}_{1'} + \text{H}_{5'} + \text{H}_7 + \text{H}_9$ ),  $1.63 - 1.50$  (m, 2H,  $\text{H}_8$ ), 1.35, 1.33, 1.32 (3s, 9H,  $\text{H}_1$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 172.0$ , 171.5 ( $\text{C}_3 + \text{C}_5$ ), 167.6, 167.3 ( $\text{C}_d$ ), 157.5, 157.4 ( $\text{C}_i$ ), 148.0 ( $\text{C}_h$ ), 144.4, 144.3 ( $\text{C}_k$ ), 141.3, 141.3, 141.0 ( $\text{C}_j$ ), 112.7, 112.5, 112.4, 112.2, 112.1, 111.9, 111.6 ( $\text{C}_e + \text{C}_g + \text{C}_j$ ), 82.6, 82.2, 81.8 ( $\text{C}_{3'}$ ), 77.6, 77.5 ( $\text{C}_{7'}$ ), 74.5, 74.2 ( $\text{C}_{8'}$ ), 71.2, 69.9 ( $\text{C}_{2'}$ ), 67.0, 65.8 ( $\text{C}_c$ ), 58.8, 57.4 ( $\text{C}_4$ ), 52.6, 52.5, 52.3, 51.6, 51.1, 50.9 ( $\text{C}_{4'} + \text{C}_{6'} + \text{C}_2 + \text{C}_6$ ), 44.2, 43.6 ( $\text{C}_a$ ), 40.2 ( $\text{C}_{10}$ ), 32.1, 31.9, 31.3, 30.9 ( $\text{C}_b + \text{C}_{1'} + \text{C}_7$ ), 29.8, 29.7 ( $\text{C}_{5'}$ ), 28.8, 28.8, 28.7, 28.6, 28.3 ( $\text{C}_1 + \text{C}_9$ ), 22.8, 22.1 ( $\text{C}_8$ ); **HRMS** (ESI):  $m/z$  648.41852  $[\text{M}+\text{H}]^+$  ( $\text{C}_{31}\text{H}_{54}\text{N}_9\text{O}_6$  requires 648.41916).

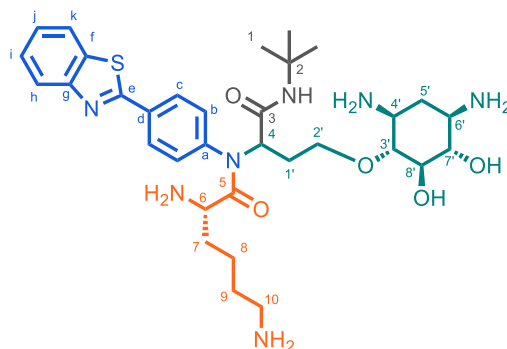
**4-(benzothiazol-2-yl)phenyl-(R/S)-(Boc)-DOS<sub>(Acetal)</sub>-(Boc)-L-Lys(Boc)-tert-butylacetamide (C30a):**

Compound **C30a** was prepared following general procedure A using **2-(4-aminophenyl)benzothiazole** (22.6 mg, 0.10 mmol) as the amine, **Boc-DOS<sub>(Acetal)</sub>-CHO (58)** (49.9 mg, 0.10 mmol) as aldehyde, **Boc-L-**

**Lys(Boc)-OH** (34.6 mg, 0.10 mmol) as the acid and **tert-butyl isocyanide** (11.3  $\mu$ l, 0.10 mmol) in MeOH/CH<sub>2</sub>Cl<sub>2</sub> (0.3 mL, 1/1, v/v). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C30a** was obtained as a yellow-pale solid (86.5 mg, 76%).

$R_{f, dia1} = 0.18$  and  $R_{f, dia2} = 0.20$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.26 - 8.17$  (m, 2H, H<sub>c</sub>), 8.09 – 8.00 (m, 2H, H<sub>h</sub> + H<sub>k</sub>), 7.87 – 7.52 (m, 2H, H<sub>b</sub>), 7.58 – 7.52 (m, 1H, H<sub>i</sub>), 7.50 – 7.43 (m, 1H, H<sub>j</sub>), 6.83 – 6.56, 6.48 – 6.32 (2m, 2NH), 5.12 – 4.92 (m, 1H, H<sub>4</sub>), 4.23 – 4.09, 4.04 – 3.94 (2m, 1H, H<sub>6</sub>), 3.90 – 3.73, 3.72 – 3.44 (2m, 2H, H<sub>2'</sub>), 3.72 – 3.44 (m, 2H, H<sub>4'</sub> + H<sub>6'</sub>), 3.44 – 3.23 (m, 3H, H<sub>3'</sub> + H<sub>7'-g'</sub>), 2.97 – 2.78 (m, 2H, H<sub>10</sub>), 2.20 – 1.80, 1.79 – 0.81 (2m, 65H, 5CH<sub>2, Acetal</sub> + 12CH<sub>3, Boc</sub> + H<sub>1'</sub> + H<sub>5'</sub> + H<sub>1</sub> + H<sub>7-9</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta = 175.6$  (C<sub>5</sub>), 172.1, 171.1 (C<sub>3</sub>), 168.3 (C<sub>e</sub>), 158.4, 158.0, 157.7 (4CO<sub>Boc</sub>), 155.2 (C<sub>g</sub>), 142.5 (C<sub>a</sub>), 136.4 (C<sub>f</sub>), 135.1, 134.9 (C<sub>d</sub>), 132.2 (2C<sub>b</sub>), 129.4 (2C<sub>c</sub>), 127.9 (C<sub>i</sub>), 127.0 (C<sub>j</sub>), 124.1, 123.1 (C<sub>h</sub> + C<sub>k</sub>), 113.0, 112.9 (C<sub>IV, Acetal</sub>), 82.5, 82.0, 80.4, 80.3, 79.7 (5C<sub>IV, Boc</sub> + C<sub>3'</sub> + C<sub>7'-g'</sub>), 68.8 (C<sub>2'</sub>), 59.7, 59.4 (C<sub>4</sub>), 53.4, 53.21, 52.8, 52.5 (C<sub>4'</sub> or C<sub>6'</sub> + C<sub>2</sub> + C<sub>6</sub>), C<sub>4'</sub> or C<sub>6'</sub> (signal under the CD<sub>3</sub>OD peak), 40.9 (C<sub>10</sub>), 37.5, 37.3 (2CH<sub>2, Acetal</sub>), 32.3, 32.1, 31.8, 31.6, 30.7, 30.7, 30.3, 30.1, 29.0, 29.0, 28.9, 28.8, 28.8, 28.7 (15CH<sub>3, Boc</sub> + C<sub>1'</sub> + C<sub>5'</sub> + C<sub>1</sub> + C<sub>7</sub> + C<sub>9</sub>), 26.1, 24.9, 24.1, 23.8 (3CH<sub>2, Acetal</sub> + C<sub>8</sub>); **MS (ESI)**:  $m/z$  1136.60 [M+H]<sup>+</sup> (theoretical  $m/z$  1136.63).

#### 4-(benzothiazol-2-yl)phenyl-(R/S)-DOS-L-Lys-tert-butylacetamide (**C30b**):

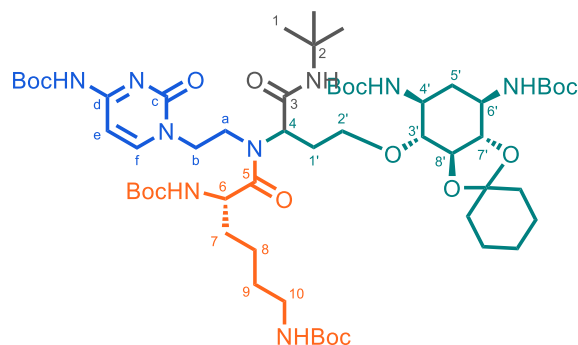


General procedure B was employed for the deprotection of compound **C30a** (59.1 mg, 0.052 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C30b** was isolated as a yellow-pale powder (56.1 mg, 97%).

**HPLC**:  $rt_{dia1} = 11.44$  min and  $rt_{dia2} = 11.80$  min (method A); NMR data of the mixture of 2 diastereomers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.25$  (d,  $J = 8.7$  Hz, 2H, H<sub>c</sub>), 8.06 (d,  $J = 7.4$  Hz, 1H, H<sub>h</sub>), 8.04 (d,  $J = 7.4$  Hz, 1H, H<sub>k</sub>), 7.88, 7.65 (br s and br d,  $J = 8.7$  Hz, 2H, H<sub>b</sub>), 7.57 (t,  $J = 7.4$  Hz, 1H, H<sub>i</sub>), 7.48 (t,  $J = 7.4$  Hz, 1H, H<sub>j</sub>), 5.33, 5.01 (t and dd,  $J_{dia1} = 7.5$  Hz and  $J_{dia2} = 8.3, 6.3$  Hz, 1H, H<sub>4</sub>), 4.20 – 4.08, 4.07 – 3.98, 3.85 – 3.74, 3.73 – 3.62 (4m, 2H, H<sub>2'</sub>), 4.08 – 3.98, 3.94 – 3.86 (2m, 1H, H<sub>6</sub>), 3.48 – 3.29 (m, 3H, H<sub>3'</sub> + H<sub>7'-g'</sub>), 3.29 – 3.07 (m, 2H, H<sub>4'</sub> + H<sub>6'</sub>), 2.92 – 2.77 (m, 2H, H<sub>10</sub>), 2.49 – 2.33, 1.98 – 1.64 (2m, 2H, H<sub>5'</sub>), 2.26 – 2.01, 1.97

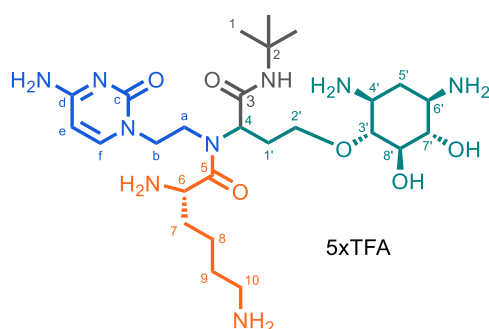
– 1.60 (2m, 2H, H<sub>1'</sub>), 1.97 – 1.60 (m, 2H, H<sub>7</sub>), 1.59 – 1.45 (m, 2H, H<sub>9</sub>), 1.44 – 1.19 (m, 2H, H<sub>8</sub>), 1.40, 1.34 (2s, 9H, H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD) δ = 171.5, 171.0 (C<sub>3</sub>), 170.6, 170.4 (C<sub>5</sub>), 167.9 (C<sub>e</sub>), 155.1 (C<sub>g</sub>), 141.4, 140.5 (C<sub>a</sub>), 136.4 (C<sub>f</sub>), 135.7 (C<sub>d</sub>), 132.3 (C<sub>b</sub>), 129.8 (C<sub>c</sub>), 128.0 (C<sub>i</sub>), 127.1 (C<sub>j</sub>), 124.2 (C<sub>h</sub>), 123.1 (C<sub>k</sub>), 82.6, 82.2 (C<sub>3'</sub>), 77.5, 77.4 (C<sub>7'</sub>), 74.5, 74.4 (C<sub>8'</sub>), 71.0, 70.2 (C<sub>2'</sub>), 61.1, 58.7 (C<sub>4</sub>), 52.7, 52.6, 52.5 (C<sub>2</sub> + C<sub>6</sub>), 51.5, 51.1, 50.8 (C<sub>4'</sub> + C<sub>6'</sub>), 40.1 (C<sub>10</sub>), 32.1, 31.3, 31.2, 30.8 (C<sub>1'</sub> + C<sub>7</sub>), 29.8 (C<sub>5'</sub>), 28.9, 28.8 (C<sub>1</sub>), 27.8, 27.7 (C<sub>9</sub>), 22.6, 22.2 (C<sub>8</sub>); HRMS (ESI): *m/z* 656.35919 [M+H]<sup>+</sup> (C<sub>33</sub>H<sub>50</sub>N<sub>7</sub>O<sub>5</sub>S requires 656.35886).

**(Boc)-Cytosine-(R/S)-(Boc)-DOS<sub>(Acetal)</sub>-(Boc)-L-Lys(Boc)-*tert*-butylacetamide (C31a):**



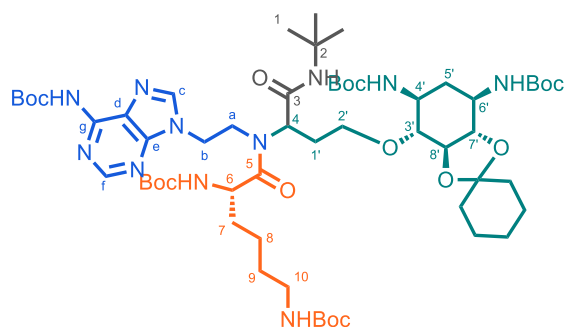
Compound **C31a** was prepared following general A procedure using **Boc-cytosine-NH<sub>2</sub> (25)** (38.1 mg, 0.15 mmol) as the amine, **Boc-DOS<sub>(Acetal)</sub>-CHO (58)** (74.7 mg, 0.15 mmol) as aldehyde, **Boc-Lys(Boc)-OH** (51.9 mg, 0.15 mmol) as the acid and ***tert*-butyl isocyanide** (17.0 μl, 0.15 mmol) in MeOH (0.3 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C31a** was obtained as a white solid (66.7 mg, 38%).

*R*<sub>f,dia1</sub> = 0.15 and *R*<sub>f,dia2</sub> = 0.17 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD) δ = 8.13 – 7.88, 7.85 – 7.73 (2m, 1H, H<sub>f</sub>), 7.35 – 7.17, 7.10 – 6.98 (2m, 1H, H<sub>e</sub>), 6.88 – 6.42 (m, 2NH), 5.22 – 5.06, 4.73 – 4.44 (2m, 2H, H<sub>4</sub> + H<sub>6</sub>), 4.43 – 3.22 (m, 11H, H<sub>a-b</sub> + H<sub>2'-4'</sub> + H<sub>6'-8'</sub>), 3.15 – 2.95 (m, 2H, H<sub>10</sub>), 2.43 – 1.97, 1.82 – 0.87 (2m, 74H, 5CH<sub>2</sub>.Acetal + 15CH<sub>3</sub>.Boc + H<sub>1'</sub> + H<sub>5'</sub> + H<sub>1</sub> + H<sub>7-9</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD) δ = 177.6, 175.7 (C<sub>5</sub>), 171.9, 170.7 (C<sub>3</sub>), 165.4, 165.2 (C<sub>d</sub>), 158.6, 158.5, 158.5, 158.2, 157.9, 157.8, 157.6 (4CO<sub>Boc</sub> + C<sub>c</sub>), 153.4, 150.8, 150.7, 150.5 (CO<sub>Boc</sub> + C<sub>f</sub>), 113.0, 112.9 (C<sub>IV</sub>.Acetal), 97.1, 96.8 (C<sub>e</sub>), 83.0, 83.0, 83.0, 82.1, 81.9, 81.8, 81.4, 81.3, 80.5, 80.1, 79.8, 79.7 (5C<sub>IV</sub>.Boc + C<sub>3'</sub> + C<sub>7'-8'</sub>), 68.8, 68.7 (C<sub>2'</sub>), 58.6, 58.4 (C<sub>4</sub>), 54.0, 52.9, 52.4, 52.4, 52.2 (C<sub>4'</sub> or C<sub>6'</sub> + C<sub>2</sub> + C<sub>6</sub>), C<sub>b</sub> + C<sub>4'</sub> or C<sub>6'</sub> (signal under the CD<sub>3</sub>OD peak), 44.2, 43.5 (C<sub>a</sub>), 41.1, 40.9, 40.3 (C<sub>10</sub>), 37.4 (2CH<sub>2</sub>.Acetal), 32.6, 31.7, 31.0, 30.7, 29.2, 29.1, 29.0, 28.9, 28.9, 28.8, 28.4 (15CH<sub>3</sub>.Boc + C<sub>1'</sub> + C<sub>5'</sub> + C<sub>1</sub> + C<sub>7</sub> + C<sub>9</sub>), 26.1, 25.1, 24.9, 24.7, 24.2, 24.0 (3CH<sub>2</sub>.Acetal + C<sub>8</sub>); MS (ESI): *m/z* 1164.60 [M+H]<sup>+</sup> (theoretical *m/z* 1164.71).

**Cytosine-(R/S)-DOS-L-Lys-*tert*-butylacetamide (C31b):**

General procedure B was employed for the deprotection of compound **C31a** (42.9 mg, 0.0369 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C31b** was isolated as a white powder (34.5 mg, 81%).

**HPLC:**  $rt_{dia1}$  = 6.33 min and  $rt_{dia2}$  = 6.43 min (method A); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.55 – 8.21, 7.75, 7.60, 7.53 (m and 3s, NH), 7.99, 7.91 (2d,  $J$  = 7.6 Hz, 1H, H<sub>f</sub>), 6.13 – 6.05 (m, 1H, H<sub>e</sub>), 5.20 – 5.08, 4.96 – 4.63, 4.62 – 4.35 (3m, 2H, H<sub>4</sub> + H<sub>6</sub>), 4.19 – 3.55 (m, 6H, H<sub>a+b</sub> + H<sub>2'</sub>), 3.55 – 3.10 (m, 5H, H<sub>3'-4'</sub> + H<sub>6'-8'</sub>), 3.07 – 2.88 (m, 2H, H<sub>10</sub>), 2.52 – 2.32, 1.98 – 1.40 (2m, 2H, H<sub>5'</sub>), 2.29 – 1.66 (m, 6H, H<sub>1'</sub> + H<sub>7</sub> + H<sub>9</sub>), 1.65 – 1.42 (m, 2H, H<sub>8</sub>), 1.35, 1.34, 1.33, 1.33 (4s, 9H, H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 172.7, 172.3, 171.8, 171.5, 170.5 (C<sub>3</sub> + C<sub>5</sub>), 163.3, 163.0, 162.0 (C<sub>d</sub>), 151.3, 151.0, 150.7, 150.2 (C<sub>c</sub> + C<sub>f</sub>), 95.2, 94.9, 94.7 (C<sub>e</sub>), 82.6, 82.0, 81.8 (C<sub>3'</sub>), 77.5, 77.4, 77.1 (C<sub>7'</sub>), 74.5, 74.2 (C<sub>8'</sub>), 70.7, 70.0, 69.8 (C<sub>2'</sub>), 58.1, 57.1 (C<sub>4</sub>), 52.9, 52.6, 52.4, 52.1, 51.7, 51.6, 51.0, 50.9, 50.5 (C<sub>4'</sub> + C<sub>6'</sub> + C<sub>2</sub> + C<sub>6</sub>), C<sub>b</sub> (signal under the CD<sub>3</sub>OD peak), 44.5, 43.5 (C<sub>a</sub>), 40.2 (C<sub>10</sub>), 32.3, 32.0, 31.5, 31.2, 31.0 (C<sub>1'</sub> + C<sub>7</sub>), 29.8, 29.7 (C<sub>5'</sub>), 28.9, 28.8, 28.7, 28.6, 28.3, 28.0 (C<sub>1</sub> + C<sub>9</sub>), 22.6, 22.1, 21.9 (C<sub>8</sub>); **HRMS** (ESI):  $m/z$  584.38806 [M+H]<sup>+</sup> (C<sub>26</sub>H<sub>50</sub>N<sub>9</sub>O<sub>6</sub> requires 584.38786).

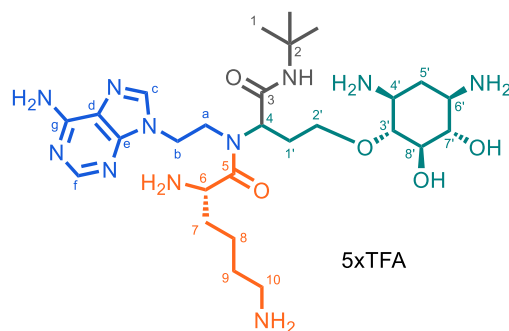
**(Boc)-Adenine-(R/S)-(Boc)-DOS<sub>(Acetal)</sub>-(Boc)-L-Lys(Boc)-*tert*-butylacetamide (C32a):**

Compound **C32a** was prepared following general procedure A using **Boc-adenine-NH<sub>2</sub> (21)** (41.7 mg, 0.15 mmol) as the amine, **Boc-DOS<sub>(Acetal)</sub>-CHO (58)** (74.7 mg, 0.15 mmol) as aldehyde, **Boc-L-Lys(Boc)-OH** (51.9 mg, 0.15 mmol) as the acid and ***tert*-butyl isocyanide** (17.0  $\mu$ l, 0.15 mmol) in MeOH (0.3 mL).

After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C32a** was obtained as a white solid (107 mg, 60%).

$R_{f, \text{dia1}} = 0.15$  and  $R_{f, \text{dia2}} = 0.17$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.63, 8.62, 8.58, 8.57$  (4s, 1H, H<sub>f</sub>), 8.44, 8.42, 8.36, 8.25 (4s, 1H, H<sub>c</sub>), 7.84 – 7.70, 7.04 – 6.90 (2m, NH), 5.22 – 5.06, 4.72 – 4.25 (2m, 1H, H<sub>4</sub>), 4.72 – 4.25 (m, 3H, H<sub>b</sub> + H<sub>6</sub>), 4.08 – 3.73, 3.72 – 3.21 (2m, 9H, H<sub>a</sub> + H<sub>2'-4'</sub> + H<sub>6'-8'</sub>), 3.14 – 2.91 (m, 2H, H<sub>10</sub>), 2.38 – 1.98, 1.81 – 1.09 (2m, 74H, 5CH<sub>2, Acetal</sub> + 15CH<sub>3, Boc</sub> + H<sub>1'</sub> + H<sub>5'</sub> + H<sub>1</sub> + H<sub>7-9</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 177.8, 171.9, 170.7$  (C<sub>3</sub> + C<sub>5</sub>), 158.5, 158.3, 157.7 (4C<sub>Boc</sub>), 153.4, 153.2, 153.1, 153.0, 152.9, 152.4, 152.3, 151.3, 151.2 (CO<sub>Boc</sub> + C<sub>e-g</sub>), 145.3, 145.2 (C<sub>c</sub>), 122.6 (C<sub>d</sub>), 113.0, 112.9 (C<sub>IV, Acetal</sub>), 82.7, 82.7, 81.9, 81.3, 80.5, 80.3, 80.2, 79.8, 79.8 (5C<sub>IV, Boc</sub> + C<sub>3'</sub> + C<sub>7'-8'</sub>), 68.6 (C<sub>2'</sub>), 58.6 (C<sub>4</sub>), 54.8, 52.9, 52.5, 52.4, 52.3, 52.2 (C<sub>4'</sub> or C<sub>6'</sub> + C<sub>2</sub> + C<sub>6</sub>), C<sub>4'</sub> or C<sub>6'</sub> (signal under the CD<sub>3</sub>OD peak), 44.7, 44.8, 43.6, 43.1 (C<sub>a-b</sub>), 41.1, 40.9 (C<sub>10</sub>), 37.7, 37.4, 37.4, 37.2 (2CH<sub>2, Acetal</sub>), 32.7, 30.9, 30.7, 29.2, 28.9, 28.9, 28.8, 28.8, 28.8, 28.5 (15CH<sub>3, Boc</sub> + C<sub>1'</sub> + C<sub>5'</sub> + C<sub>1</sub> + C<sub>7</sub> + C<sub>9</sub>), 26.1, 25.0, 24.8, 24.7 (3CH<sub>2, Acetal</sub> + C<sub>8</sub>); **MS (ESI):**  $m/z$  1188.80 [M+H]<sup>+</sup> (theoretical  $m/z$  1188.72).

#### Adenine-(R/S)-DOS-L-Lys-*tert*-butylacetamide (**C32b**):

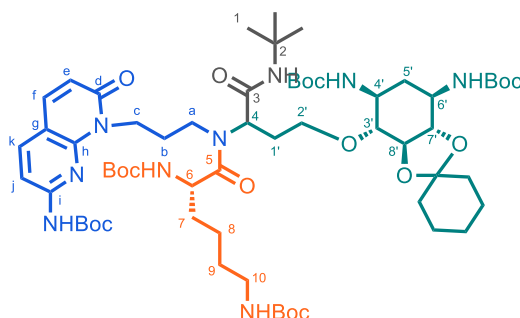


General procedure B was employed for the deprotection of compound **C32a** (51.2 mg, 0.043 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, the crude mixture was purified by semipreparative HPLC (method B) leading to a mixture of diastereomers **C32b** as a white-pale powder (27.2 mg, 54%).

**HPLC:**  $rt_{\text{dia1}} = 10.92$  min and  $rt_{\text{dia2}} = 11.10$  min (method B); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.46, 8.41, 8.40, 8.37$  (4s, 2H, H<sub>c</sub> + H<sub>f</sub>), 5.24 – 5.16, 4.75 – 4.45 (2m, 1H, H<sub>4</sub>), 5.01 – 4.74, 4.75 – 4.45 (2m, 1H, H<sub>6</sub>), 4.75 – 4.45 (m, 2H, H<sub>b</sub>) 4.22 – 3.64 (m, 4H, H<sub>a</sub> + H<sub>2'</sub>), 3.52 – 3.10 (m, 3H, H<sub>3'</sub> + H<sub>7'-8'</sub>), 3.35 – 3.10 (m, 2H, H<sub>4'</sub> + H<sub>6'</sub>), 3.04 – 2.89 (m, 2H, H<sub>10</sub>), 2.52 – 2.08, 2.07 – 1.65 (2m, 8H, H<sub>1'</sub> + H<sub>5'</sub> + H<sub>7</sub> + H<sub>9</sub>), 1.65 – 1.49, 1.48 – 1.27 (2m, 2H, H<sub>8</sub>), 1.37, 1.36, 1.36, 1.33 (4s, 9H, H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 172.6, 172.2, 171.6, 171.5, 171.3, 170.5$  (C<sub>3</sub> + C<sub>5</sub>), 159.2, 158.8 (C<sub>g</sub>), 152.6, 152.1, 150.8, 150.5 (C<sub>e</sub> + C<sub>f</sub>), 146.7, 145.7 (C<sub>c</sub>), 119.7, 119.4, 116.5, 116.2 (C<sub>d</sub>), 82.6, 82.1, 81.9 (C<sub>3'</sub>), 77.5, 77.3, 77.0 (C<sub>7'</sub>), 74.5, 74.2 (C<sub>8'</sub>), 70.8, 70.0 (C<sub>2'</sub>), 58.4, 58.1, 57.2 (C<sub>4</sub>), 52.9, 52.6, 52.5,

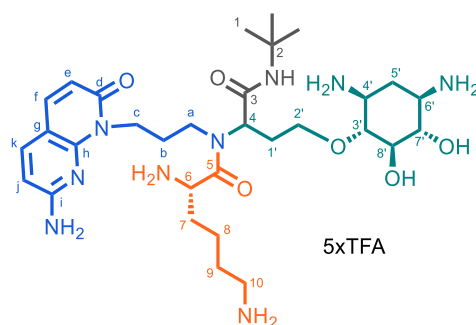
52.4, 52.3 ( $C_2 + C_6$ ), 51.7, 51.6, 51.3, 51.1, 50.9 ( $C_{4'} + C_{6'}$ ), 44.9, 44.7, 43.7, 43.7 ( $C_{a-b}$ ), 40.3, 40.2 ( $C_{10}$ ), 32.3, 32.1, 31.7, 31.2, 31.1 ( $C_{1'} + C_7$ ), 29.9, 29.8, 29.8 ( $C_{5'}$ ), 28.9, 28.8, 28.7 ( $C_1$ ), 28.5, 28.3 ( $C_9$ ), 22.8, 22.4, 22.3 ( $C_8$ ); HRMS (ESI):  $m/z$  608.39929 [ $M+H$ ] $^+$  ( $C_{27}H_{50}N_{11}O_5$  requires 608.39909).

**(Boc)-Naphthyridine(*N*-alkyl)-(R/S)-(Boc)-DOS<sub>(Acetal)</sub>-(Boc)-L-Lys(Boc)-*tert*-butylacetamide  
(**C33a**):**



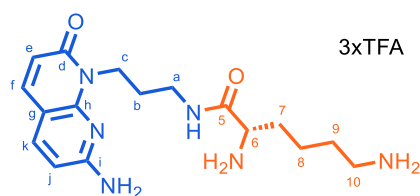
Compound **C33a** was prepared following general A procedure using **Boc-naphthyridine(*N*-alkyl)-NH<sub>2</sub> (42)** (31.8 mg, 0.10 mmol) as the amine, **Boc-DOS<sub>(Acetal)</sub>-CHO (58)** (49.9 mg, 0.10 mmol) as aldehyde, **Boc-L-Lys(Boc)-OH** (34.6 mg, 0.10 mmol) as the acid and ***tert*-butyl isocyanide** (11.3  $\mu$ l, 0.10 mmol) in MeOH/CH<sub>2</sub>Cl<sub>2</sub> (0.3 mL, 1/1, v/v). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C33a** was obtained as a yellow-pale solid (28.5 mg, 23%).

$R_{f, \text{dia}1} = 0.23$  and  $R_{f, \text{dia}2} = 0.26$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers;  $^1\text{H NMR}$  (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.06 - 7.97, 7.94 - 7.75$  (m, 3H,  $H_f + H_{f-k}$ ), 6.94, 6.84, 6.81 - 6.46 (br d, s and m, 2NH), 6.65 - 6.47 (m, 1H,  $H_e$ ), 5.16 - 5.04, 4.81 - 4.17, 4.73 - 4.35 (m, 3H,  $H_c + H_4 + H_6$ ), 4.00 - 3.17 (m, 9H,  $H_a + H_{2'-4'} + H_{6'-8'}$ ), 3.11 - 2.93 (m, 2H,  $H_{10}$ ), 2.43 - 1.86, 1.86 - 1.05 (2m, 76H,  $5\text{CH}_{2, \text{Acetal}} + 15\text{CH}_{3, \text{Boc}} + H_b + H_{1'} + H_{5'} + H_1 + H_{7-9}$ );  $^{13}\text{C NMR}$  (101 MHz, CD<sub>3</sub>OD)  $\delta = 176.7, 175.4$  ( $C_5$ ), 171.8, 171.0 ( $C_3$ ), 165.4, 165.3 ( $C_d$ ), 158.5, 158.2, 157.8, 157.7, 154.8, 154.7, 154.3, 154.2, 149.8, 149.7, 149.6 ( $5\text{CO}_{\text{Boc}} + C_{h-i}$ ), 140.4, 140.3, 139.8, 139.7, 139.6 ( $C_f + C_k$ ), 119.8, 119.8, 119.7 ( $C_e$ ), 113.0, 112.9, 112.8, 112.7 ( $C_{IV, \text{Acetal}} + C_g$ ), 108.8, 108.7 ( $C_j$ ), 82.2, 82.1, 82.0, 81.6, 81.3, 81.0, 80.6, 80.4, 80.3, 80.2, 80.1, 79.8 ( $5\text{C}_{IV, \text{Boc}} + C_{3'} + C_{7'-8'}$ ), 68.5 ( $C_{2'}$ ), 58.6, 56.5 ( $C_4$ ), 52.8, 52.4, 52.2, 52.1 ( $C_{4'} + C_{6'} + C_2 + C_6$ ), 44.0, 41.2, 41.2, 41.0, 40.1, 39.9, 39.7 ( $C_a + C_c + C_{10}$ ), 37.9, 37.4, 37.3 ( $2\text{CH}_{2, \text{Acetal}}$ ), 33.1, 32.8, 30.8, 30.7, 30.62, 30.6, 30.5, 29.1, 29.1, 28.1, 28.9, 28.8, 28.7, 28.7, 28.7, 28.6, 28.6, 27.9 ( $15\text{CH}_{3, \text{Boc}} + C_b + C_{1'} + C_{5'} + C_1 + C_7 + C_9$ ), 26.1, 25.0, 24.9, 24.9, 24.3, 23.7 ( $3\text{CH}_{2, \text{Acetal}} + C_8$ ).

**Naphthyridine(*N*-alkyl)-(R/S)-DOS-L-Lys-*tert*-butylacetamide (**C33b**):**

General procedure B was employed for the deprotection of compound **C33a** (23.0 mg, 0.019 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, the crude mixture was purified by semipreparative HPLC (method A) leading to a mixture of diastereomers **C33b** as a white-pale powder (12.0 mg, 52%) and the by-product **C35b** as white powder (7.5 mg).

**C33b**: HPLC:  $rt_{dia1+2} = 9.46$  min (method A); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 7.78 - 7.63$  (m, 2H, H<sub>f</sub> + H<sub>k</sub>), 6.54 – 6.44 (m, 1H, H<sub>j</sub>), 6.41 – 6.28 (m, 1H, H<sub>e</sub>), 5.14 – 5.05, 4.66 – 4.58 (2m, 1H, H<sub>4</sub>), 4.57 – 4.36 (m, 2H, H<sub>c</sub>), 4.57 – 4.36, 4.34 – 4.25 (2m, 1H, H<sub>6</sub>), 4.15 – 3.84, 3.78 – 3.63 (2m, 2H, H<sub>2'</sub>), 3.63 – 3.48 (m, 2H, H<sub>a</sub>), 3.48 – 3.29 (m, 3H, H<sub>3'</sub> + H<sub>7'-8'</sub>), 3.28 – 3.08 (m, 2H, H<sub>4'</sub> + H<sub>6'</sub>), 3.02 – 2.87 (m, 2H, H<sub>10</sub>), 2.48 – 2.34, 1.95 – 1.63 (2m, 2H, H<sub>5'</sub>), 2.31 – 1.95 (m, 2H, H<sub>1'</sub>), 1.94 – 1.63 (m, 6H, H<sub>b</sub> + H<sub>7</sub> + H<sub>9</sub>), 1.61 – 1.41 (m, 2H, H<sub>8</sub>), 1.32, 1.31, 1.29, 1.28 (4s, 9H, H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 171.9, 171.5, 171.4, 171.1$  (C<sub>3</sub> + C<sub>5</sub>), 165.8 (C<sub>d</sub>), 162.1 (C<sub>i</sub>), 150.9 (C<sub>h</sub>), 140.7, 140.6, 140.4, 139.4, 139.4, 139.3 (C<sub>k</sub> + C<sub>f</sub>), 114.8, 114.8 (C<sub>e</sub>), 108.6, 108.5 (C<sub>g</sub>), 106.8, 106.6 (C<sub>j</sub>), 82.5, 82.2, 81.7 (C<sub>3'</sub>), 77.6, 77.5 (C<sub>7</sub>), 74.4 (C<sub>8'</sub>), 71.08, 69.9 (C<sub>2'</sub>), 59.7, 57.0 (C<sub>4</sub>), 52.9, 52.5, 52.3 (C<sub>2</sub> + C<sub>6</sub>), 51.6, 51.1, 50.9 (C<sub>4'</sub> + C<sub>6'</sub>), 45.8, 44.1 (C<sub>a</sub>), 40.2 (C<sub>10</sub>), 39.4 (C<sub>c</sub>), 32.2, 31.9, 31.5, 31.4, 30.8, 30.5 (C<sub>b</sub> + C<sub>1'</sub> + C<sub>7</sub>), 29.8, 29.7 (C<sub>5'</sub>), 28.8, 28.8 (C<sub>1</sub>), 28.7, 28.1 (C<sub>9</sub>), 22.9, 22.7 (C<sub>8</sub>); HRMS (ESI):  $m/z$  648.41925 [M+H]<sup>+</sup> (C<sub>31</sub>H<sub>54</sub>N<sub>9</sub>O<sub>6</sub> requires 648.41916).

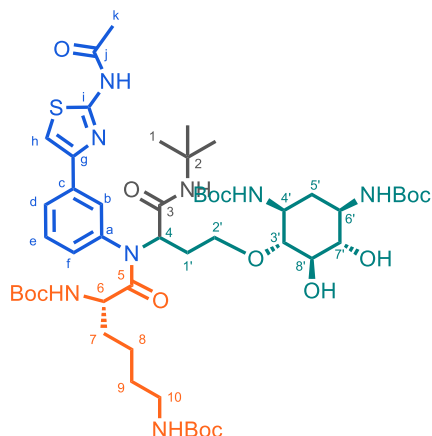
**C35b**: (S)-2,6-diamino-N-(3-(7-amino-2-oxo-1,8-naphthyridin-1(2H)-yl)propyl)hexanamide:

HPLC:  $rt = 8.52$  min (method A); <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 7.70$  (d,  $J = 9.2$  Hz, 1H, H<sub>f</sub>), 7.68 (d,  $J = 8.5$  Hz, 1H, H<sub>k</sub>), 6.49 (d,  $J = 8.5$  Hz, 1H, H<sub>j</sub>), 6.34 (d,  $J = 9.2$  Hz, 1H, H<sub>e</sub>), 4.50 (t,  $J = 7.4$  Hz, 2H, H<sub>c</sub>), 3.86 (t,  $J = 6.6$  Hz, 1H, H<sub>6</sub>), 3.50 – 3.37, 3.35 – 3.21 (2m, 2H, H<sub>a</sub>), 2.97 (t,  $J = 7.5$  Hz, 2H, H<sub>10</sub>), 2.06 – 1.82 (m, 4H, H<sub>b</sub> + H<sub>7</sub>), 1.82 – 1.66 (m, 2H, H<sub>9</sub>), 1.64 – 1.47 (m, 2H, H<sub>8</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 169.8$  (C<sub>5</sub>), 165.8 (C<sub>d</sub>), 162.1 (C<sub>i</sub>), 151.0 (C<sub>h</sub>), 140.5 (C<sub>f</sub>), 139.3 (C<sub>k</sub>), 114.7 (C<sub>e</sub>), 108.6 (C<sub>g</sub>), 106.7 (C<sub>j</sub>), 54.4 (C<sub>6</sub>), 40.3



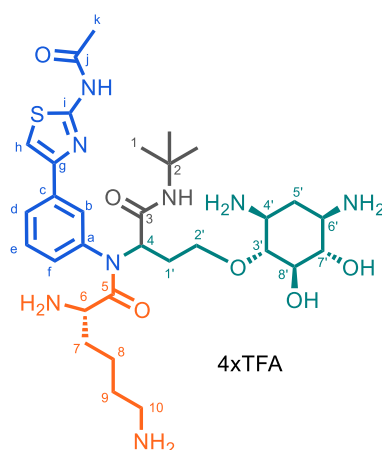
(C<sub>10</sub>), 39.8 (C<sub>c</sub>), 38.3 (C<sub>a</sub>), 32.1 (C<sub>7</sub>), 29.0 (C<sub>b</sub>), 28.3 (C<sub>9</sub>), 23.0 (C<sub>8</sub>); **HRMS** (ESI): *m/z* 347.21906 [M+H]<sup>+</sup> (C<sub>17</sub>H<sub>27</sub>N<sub>6</sub>O<sub>2</sub> requires 347.21900).

**S-(R/S)-(Boc)-DOS<sub>(Acetal)</sub>-(Boc)-L-Lys(Boc)-tert-butylacetamide (C34a):**



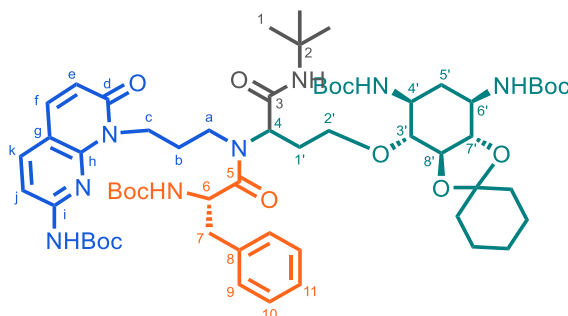
Compound **C34a** was prepared following general procedure A using **S-NH<sub>2</sub> (45)** (35.0 mg, 0.15 mmol) as the amine, **Boc-DOS<sub>(Acetal)</sub>-CHO (58)** (49.9 mg, 0.10 mmol) as aldehyde, **Boc-L-Lys(Boc)-OH** (34.6 mg, 0.10 mmol) as the acid and **tert-butyl isocyanide** (11.3  $\mu$ l, 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C34a** was obtained as a white-pale solid (74.0 mg, 65%).

**R<sub>f,dia1</sub>** = 0.13 and **R<sub>f,dia2</sub>** = 0.15 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.33 – 8.19, 8.06 – 7.77, 7.61 – 6.98 (3m, 5H, H<sub>b</sub> + H<sub>d-f</sub> + H<sub>h</sub>), 6.86 – 6.52 (m, NH), 5.02, 4.75 (2br s, 1H, H<sub>4</sub>), 4.79 – 4.69 (m, 1H, H<sub>6</sub>), 4.26 – 3.89, 3.88 – 3.55, 3.55 – 3.14 (3m, 7H, H<sub>2'-4'</sub> + H<sub>6'-8'</sub>), 2.98 – 2.68 (m, 2H, H<sub>10</sub>), 2.33 – 2.18 (m, 3H, H<sub>k</sub>), 2.33 – 2.18, 2.17 – 2.03, 2.02 – 1.76, 1.75 – 0.92 (4m, 65H, 5CH<sub>2,Acetal</sub> + 12CH<sub>3,Boc</sub> + H<sub>1'</sub> + H<sub>5'</sub> + H<sub>1</sub> + H<sub>7-9</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 175.9, 175.5 (C<sub>5</sub>), 170.9, 170.8 (C<sub>j</sub> + C<sub>3</sub>), 158.4, 158.0, 157.7 (4C<sub>Boc</sub> + C<sub>i</sub>), 151.6, 149.7, 140.3, 137.6 (C<sub>a</sub> + C<sub>c</sub>), 130.7, 130.3, 127.5, 117.3, 116.3, 114.2, 112.9 (C<sub>IV,Acetal</sub> + C<sub>b</sub> + C<sub>d-f</sub>), 109.9, 108.2 (C<sub>h</sub>), 81.9, 80.2, 79.8 (4C<sub>IV,Boc</sub> + C<sub>3'</sub> + C<sub>7'-8'</sub>), 68.9 (C<sub>2'</sub>), 59.5 (C<sub>4</sub>), 53.3, 53.1, 52.8, 52.5, 52.4 (C<sub>4'</sub> or C<sub>6'</sub> + C<sub>2</sub> + C<sub>6</sub>), C<sub>4'</sub> or C<sub>6'</sub> (signal under the CD<sub>3</sub>OD peak), 40.9, 40.8 (C<sub>10</sub>), 37.4, 37.3 (2CH<sub>2,Acetal</sub>), 33.2, 35.5, 31.7, 30.2, 30.0, 29.1, 29.0, 28.9, 28.8, 28.7 (12CH<sub>3,Boc</sub> + C<sub>1'</sub> + C<sub>5'</sub> + C<sub>1</sub> + C<sub>7</sub> + C<sub>9</sub>), 26.1, 24.9 (3CH<sub>2,Acetal</sub>), 24.1 (C<sub>8</sub>), 22.6 (C<sub>k</sub>); **MS (ESI):** *m/z* 1143.13 [M+H]<sup>+</sup> (theoretical *m/z* 1143.64).

**S-(R/S)-DOS-L-Lys-*tert*-butylacetamide (C34b):**

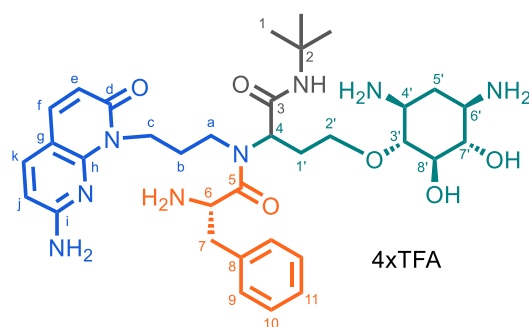
General procedure B was employed for the deprotection of compound **C34a** (49.0 mg, 0.043 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, the crude mixture was purified by semipreparative HPLC (method A) leading to a mixture of diastereomers **C34b** as a white-pale powder (22.5 mg, 47%).

**HPLC:**  $rt_{dia1}$  = 10.47 min and  $rt_{dia2}$  = 10.78 min (method A); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.48 – 8.18, 8.12 – 7.80 (2m, 2H, H<sub>b</sub> + H<sub>f</sub>), 7.67 – 7.60, 7.59 – 7.44, 7.43 – 7.22 (3m, 3H, H<sub>d</sub> + H<sub>e</sub> + H<sub>h</sub>), 5.33 – 5.24, 5.01 – 4.77 (2m, 1H, H<sub>4</sub>), 4.25 – 3.93, 3.80 – 3.61 (2m, 2H, H<sub>2'</sub>), 4.10 – 3.93, 3.92 – 3.79 (2m, 1H, H<sub>6</sub>), 3.45 – 3.27 (m, 3H, H<sub>3'</sub> + H<sub>7'-8'</sub>), 3.26 – 3.04 (m, 2H, H<sub>4'</sub> + H<sub>6'</sub>), 2.91 – 2.67 (m, 2H, H<sub>10</sub>), 2.43 – 2.30, 1.96 – 1.58 (2m, 2H, H<sub>5</sub>), 2.24 (s, 3H, H<sub>k</sub>), 2.28 – 2.00, 1.95 – 1.58 (2m, 2H, H<sub>1'</sub>), 1.95 – 1.58 (m, 2H, H<sub>7</sub>), 1.56 – 1.10 (m, 4H, H<sub>8-9</sub>), 1.38, 1.37, 1.32 (3s, 9H, H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 171.7, 171.0, 170.8, 170.5 (C<sub>j</sub> + C<sub>3</sub> + C<sub>5</sub>), 162.7, 160.0 (C<sub>i</sub>), 149.6, 149.4 (C<sub>g</sub>), 138.5, 137.8 (C<sub>a</sub> + C<sub>c</sub>), 131.3, 131.2, 130.6, 130.5 (C<sub>d-e</sub>), 128.7, 128.3, 128.2 (C<sub>b</sub> + C<sub>f</sub>), 110.6, 110.4, 110.2 (C<sub>h</sub>), 82.6, 82.1, 82.0 (C<sub>3'</sub>), 77.5, 77.4, 77.4 (C<sub>7'</sub>), 74.6, 74.5 (C<sub>8'</sub>), 71.1, 70.2 (C<sub>2'</sub>), 61.7, 58.7, 58.4 (C<sub>4</sub>), 52.7, 52.6, 52.5 (C<sub>2</sub> + C<sub>6</sub>), 51.5, 51.1, 51.1, 50.8 (C<sub>4'</sub> + C<sub>6'</sub>), 40.1 (C<sub>10</sub>), 32.1, 31.2, 30.9 (C<sub>1'</sub> + C<sub>7</sub>), 29.8 (C<sub>5'</sub>), 28.9, 28.8 (C<sub>1</sub>), 27.8, 27.7 (C<sub>9</sub>), 22.6 (C<sub>k</sub>), 22.4, 22.2 (C<sub>8</sub>); **HRMS** (ESI):  $m/z$  663.36572[M+H]<sup>+</sup> (C<sub>31</sub>H<sub>51</sub>N<sub>8</sub>O<sub>6</sub>S requires 663.36468).

**(Boc)-Naphthyridine(N-alkyl)-(R/S)-(Boc)-DOS<sub>(Acetal)</sub>-(Boc)-L-Phe-*tert*-butylacetamide****(C36a):**

Compound **C36a** was prepared following general procedure A using **Boc-naphthyridine(N-alkyl)-NH<sub>2</sub> (42)** (22,3 mg, 0.070 mmol) as the amine, **Boc-DOS<sub>(Acetal)</sub>-CHO (58)** (34.8 mg, 0.070 mmol) as aldehyde, **Boc-L-Phe-OH** (18,6 mg, 0.070 mmol) as the acid and ***tert*-butyl isocyanide** (7.10  $\mu$ l, 0.070 mmol) in MeOH (0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 96:4, v/v), a mixture of diastereomers **C36b** was obtained as a white solid (39.0 mg, 54%).

$R_{f,dia1+2} = 0.22$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 96:4); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.12 - 7.69$  (m, 3H, H<sub>f</sub> + H<sub>j</sub> + H<sub>k</sub>), 7.33 – 6.97 (m, 5H, H<sub>9-11</sub>), 6.65 – 6.45 (m, 1H, H<sub>e</sub>), 5.03 – 4.92, 4.91 – 4.74, 4.73 – 4.54 (3m, 2H, H<sub>4</sub> + H<sub>6</sub>), 4.53 – 4.24 (m, 2H, H<sub>c</sub>), 3.81 – 3.54, 3.53 – 3.18 (m, 9H, H<sub>a</sub> + H<sub>2'-4'</sub> + H<sub>6'-8'</sub>), 3.61 – 2.72 (m, 2H, H<sub>7</sub>), 2.32 – 1.68, 1.67 – 1.05 (m, 61H, 12CH<sub>3,Boc</sub> + 5CH<sub>2,Acetal</sub> + H<sub>b</sub> + H<sub>1'</sub> + H<sub>5'</sub> + H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta = 175.7, 175.3$  (C<sub>5</sub>), 171.8, 170.5 (C<sub>3</sub>), 165.2, 165.2 (C<sub>d</sub>), 157.9, 157.7, 157.5, 154.7, 154.6, 154.2, 149.7, 149.6 (4CO<sub>Boc</sub> + C<sub>h-i</sub>), 140.4, 140.3, 140.2, 140.2, 139.7, 139.6, 138.0 (C<sub>f</sub> + C<sub>k</sub> + C<sub>8</sub>), 131.1, 131.1, 130.4, 129.6, 129.5, 129.2, 127.8 (C<sub>9-11</sub>), 119.9, 119.7 (C<sub>e</sub>), 112.9, 112.8, 112.7 (C<sub>IV,Acetal</sub> + C<sub>g</sub>), 108.8, 108.7 (C<sub>j</sub>), 82.2, 82.2, 82.1, 81.6, 80.6, 80.6, 80.3, 80.2, 79.8 (4C<sub>IV,Boc</sub> + C<sub>3'</sub> + C<sub>7'-8'</sub>), 69.1, 68.5 (C<sub>2'</sub>), 58.6, 57.2 (C<sub>4</sub>), 54.0, 52.6, 52.4, 52.2 (C<sub>4'</sub> or C<sub>6'</sub> + C<sub>2</sub> + C<sub>6</sub>), C<sub>4'</sub> or C<sub>6'</sub> (signal under the CD<sub>3</sub>OD peak), 43.9, 42.9 (C<sub>c</sub>), 40.2, 40.0, 39.8, 39.6 (C<sub>a</sub> + C<sub>7</sub>), 37.4, 37.3 (2CH<sub>2,Acetal</sub>), 31.2, 30.4, 30.2, 29.1, 29.0, 28.9, 28.9, 28.8, 28.8, 28.7, 28.7, 28.7, 26.1, 26.1, 26.1 (12CH<sub>3,Boc</sub> + C<sub>b</sub> + C<sub>1'</sub> + C<sub>5'</sub> + C<sub>1</sub>), 24.9, 24.8, 24.7 (3CH<sub>2,Acetal</sub>); **MS (ESI):**  $m/z$  1147.27 [M+H]<sup>+</sup> (theoretical  $m/z$  1147.66).

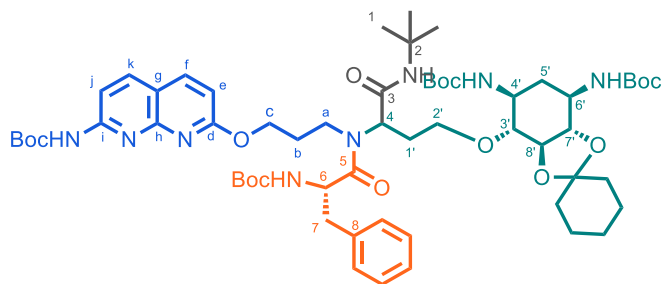
**Naphthyridine(*N*-alkyl)-(R/S)-DOS-L-Phe-*tert*-butylacetamide (**C36b** and **C37b**):**

General procedure B was employed for the deprotection of compound **C36a** (28.2 mg, 0.025 mmol) in 1 mL of  $\text{CH}_2\text{Cl}_2$  and in the presence of 50 equiv. of TFA. After filtration, the crude mixture was purified by flash chromatography on reversed phase ( $\text{H}_2\text{O}/\text{MeCN}$  gradient 95:5 to 50:50, v/v) leading to the diastereomer **C36b** ( $\text{dia}_1$ ) (13.3 mg, 48%) and **C37b** ( $\text{dia}_1:\text{dia}_2$ , ratio 3:7; 13.5 mg, 49%) which were separated and isolated as a white powder.

**C36b** – **Dia1**: **HPLC**:  $\text{rt}_{\text{dia1}} = 12.38$  min (method A); NMR data as a single conformer;  $^1\text{H NMR}$  (400 MHz,  $\text{D}_2\text{O}$ )  $\delta = 7.88$  (d,  $J = 9.2$  Hz, 1H,  $\text{H}_f$ ), 7.86 (d,  $J = 8.6$  Hz, 1H,  $\text{H}_k$ ), 7.48 – 7.00 (m, 5H,  $\text{H}_{9-11}$ ), 6.65 (d,  $J = 8.6$  Hz, 1H,  $\text{H}_j$ ), 6.50 (d,  $J = 9.2$  Hz, 1H,  $\text{H}_e$ ),  $\text{H}_4$  (signal under the  $\text{D}_2\text{O}$  peak), 4.45 – 4.15 (m, 3H,  $\text{H}_c + \text{H}_6$ ), 4.02 – 3.91, 3.73 – 3.63 (2m, 2H,  $\text{H}_{2'}$ ), 3.61 – 3.48 (m, 2H,  $\text{H}_{7'-8'}$ ), 3.48 – 3.34 (m, 1H,  $\text{H}_{3'}$ ), 3.33 – 3.20 (m, 2H,  $\text{H}_{4'} + \text{H}_{6'}$ ), 3.20 – 3.04, 2.83 – 2.71 (2m, 3H,  $\text{H}_a + \text{H}_7$ ), 2.53 – 2.41, 1.95 – 1.75 (2m, 2H,  $\text{H}_{5'}$ ), 2.28 – 2.15, 1.73 – 1.54 (2m, 2H,  $\text{H}_{1'}$ ), 1.95 – 1.75 (m, 2H,  $\text{H}_b$ ), 1.24 (s, 9H,  $\text{H}_1$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{D}_2\text{O}$ )  $\delta = 169.8$ , 169.6 ( $\text{C}_3 + \text{C}_5$ ), 165.0 ( $\text{C}_d$ ), 159.8 ( $\text{C}_i$ ), 148.4 ( $\text{C}_h$ ), 140.5, 139.5 ( $\text{C}_k + \text{C}_f$ ), 133.3 ( $\text{C}_8$ ), 129.1, 129.1, 128.0 ( $\text{C}_{9-11}$ ), 114.0 ( $\text{C}_e$ ), 108.4 ( $\text{C}_g$ ), 106.8 ( $\text{C}_j$ ), 80.6 ( $\text{C}_{3'}$ ), 75.4 ( $\text{C}_{7'}$ ), 72.5 ( $\text{C}_{8'}$ ), 70.0 ( $\text{C}_{2'}$ ), 56.8 ( $\text{C}_4$ ), 51.9 ( $\text{C}_2$ ), 51.6 ( $\text{C}_6$ ), 49.9, 49.2 ( $\text{C}_{4'} + \text{C}_{6'}$ ), 44.5 ( $\text{C}_a$ ), 38.2 ( $\text{C}_c$ ), 37.2 ( $\text{C}_7$ ), 29.4 ( $\text{C}_{1'}$ ), 28.1, 27.9 ( $\text{C}_b + \text{C}_{5'}$ ), 27.5 ( $\text{C}_1$ ); **HRMS** (ESI):  $m/z$  667.39288 [ $\text{M}+\text{H}$ ] $^+$  ( $\text{C}_{34}\text{H}_{51}\text{N}_8\text{O}_6$  requires 667.39261).

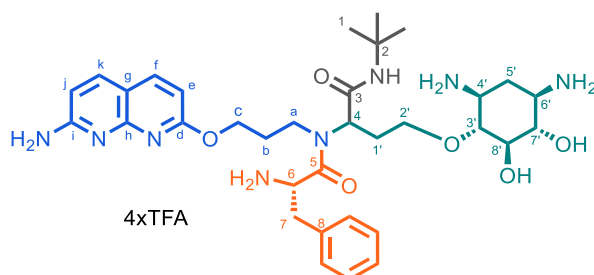
**C37b** – ( $\text{dia}_1:\text{dia}_2$ , ratio 3:7); **HPLC**:  $\text{rt}_{\text{dia2}} = 13.07$  min (method A); NMR data as a mixture of 2 diastereomers;  $^1\text{H NMR}$  (400 MHz,  $\text{D}_2\text{O}$ )  $\delta$  7.87 – 7.80 (m, 2H,  $\text{H}_f + \text{H}_k$ ), 7.43 – 6.95 (m, 5H,  $\text{H}_{9-11}$ ), 6.63, 6.62 (2d,  $J = 8.6$  Hz, 1H,  $\text{H}_j$ ), 6.48, 6.47 (2d,  $J = 9.2$  Hz, 1H,  $\text{H}_e$ ),  $\text{H}_4$  (signal under the  $\text{D}_2\text{O}$  peak), 4.40 – 4.23 (m, 1H,  $\text{H}_6$ ), 4.40 – 4.06 (m, 2H,  $\text{H}_c$ ), 3.99 – 3.81, 3.70 – 3.60 (2m, 2H,  $\text{H}_{2'}$ ), 3.59 – 3.46 (m, 2H,  $\text{H}_{7'-8'}$ ), 3.47 – 2.95, 2.84 – 2.69 (2m, 7H,  $\text{H}_a + \text{H}_{3'-4'} + \text{H}_{6'} + \text{H}_7$ ), 2.51 – 2.39, 1.98 – 1.73 (2m, 2H,  $\text{H}_{5'}$ ), 2.26 – 2.10, 1.98 – 1.73 (2m, 2H,  $\text{H}_{1'}$ ), 1.98 – 1.73, 1.70 – 1.50 (2m, 2H,  $\text{H}_b$ ), 1.24, 1.21 (2s, 9H,  $\text{H}_1$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{D}_2\text{O}$ )  $\delta = 169.7$ , 169.7, 169.6, 168.8 ( $\text{C}_3 + \text{C}_5$ ), 165.0, 164.7 ( $\text{C}_d$ ), 159.7 ( $\text{C}_i$ ), 148.3, 148.2 ( $\text{C}_h$ ), 140.4, 140.3, 139.4, 139.4 ( $\text{C}_k + \text{C}_f$ ), 133.4, 133.3 ( $\text{C}_8$ ), 129.1, 129.1, 127.8 ( $\text{C}_{9-11}$ ), 113.9 ( $\text{C}_e$ ), 108.3, 108.3 ( $\text{C}_g$ ), 106.8, 106.7 ( $\text{C}_j$ ), 80.5, 80.3 ( $\text{C}_{3'}$ ), 75.3 ( $\text{C}_{7'}$ ), 72.5 ( $\text{C}_{8'}$ ), 69.3 ( $\text{C}_{2'}$ ), 56.8, 55.6 ( $\text{C}_4$ ), 51.7, 51.6 ( $\text{C}_2 + \text{C}_6$ ), 49.9, 49.8, 49.2, 49.2 ( $\text{C}_{4'} + \text{C}_{6'}$ ), 42.3 ( $\text{C}_a$ ), 38.2, 37.9, 37.4 ( $\text{C}_c + \text{C}_7$ ), 29.8, 29.4 ( $\text{C}_{1'}$ ), 28.1, 27.9 ( $\text{C}_b + \text{C}_7$ ), 27.5, 27.4 ( $\text{C}_1$ ); **HRMS** (ESI):  $m/z$  667.39197 [ $\text{M}+\text{H}$ ] $^+$  ( $\text{C}_{34}\text{H}_{51}\text{N}_8\text{O}_6$  requires 667.39261).

**(Boc)-Naphthyridine(O-alkyl)-(R/S)-(Boc)-DOS<sub>(Acetal)</sub>-(Boc)-L-Phe-*tert*-butylacetamide**  
**(C38a):**



Compound **C38a** was prepared following general procedure A using **Boc-naphthyridine(O-alkyl)-NH<sub>2</sub> (37)** (31.8 mg, 0.10 mmol) as the amine, **Boc-DOS<sub>(Acetal)</sub>-CHO (58)** (49.9 mg, 0.10 mmol) as aldehyde, **Boc-L-Phe-OH** (34.5 mg, 0.13 mmol) as the acid and ***tert*-butyl isocyanide** (11.3  $\mu$ l, 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C38a** was obtained as a white solid (41.4 mg, 36%).

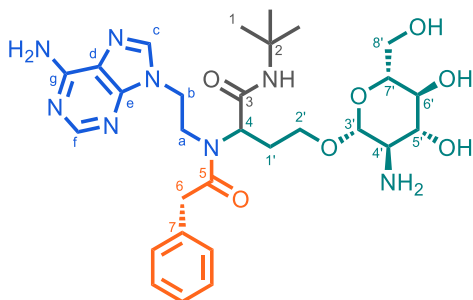
$R_{f, dia1+2} = 0.22$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 98:2); NMR data as a mixture of 2 diastereomers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.23 - 8.14$  (m, 1H, H<sub>k</sub>),  $8.16 - 8.07$  (m, 2H, H<sub>j</sub> + H<sub>f</sub>),  $7.37 - 7.16$  (m, 5H, 5CH<sub>Ar</sub>),  $6.94$  (d,  $J = 8.7$  Hz, 1H, H<sub>e</sub>),  $5.07 - 4.67$  (m, 2H, H<sub>4</sub> + H<sub>6</sub>),  $4.59 - 4.37$  (m, 2H, H<sub>c</sub>),  $3.88 - 3.17$  (m, 9H, H<sub>a</sub> + H<sub>2'-4'</sub> + H<sub>6'-8'</sub>),  $3.12 - 2.80$  (m, 2H, H<sub>7</sub>),  $2.31 - 2.04$ ,  $1.69 - 1.22$ ,  $1.03 - 0.82$  (3m, 61H, 5CH<sub>2, Acetal</sub> + 12CH<sub>3, Boc</sub> + H<sub>b</sub> + H<sub>1'</sub> + H<sub>5'</sub> + H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 175.6$ ,  $175.5$  (C<sub>5</sub>),  $171.9$ ,  $170.6$  (C<sub>3</sub>),  $166.5$ ,  $166.4$  (C<sub>d</sub>),  $157.9$ ,  $157.8$ ,  $157.7$ ,  $157.3$ ,  $155.8$ ,  $155.7$ ,  $155.4$ ,  $154.3$  (4CO<sub>Boc</sub>, C<sub>h+i</sub>),  $140.5$ ,  $140.5$  (C<sub>f</sub>),  $140.2$ ,  $140.1$  (C<sub>k</sub>),  $138.3$  (C<sub>8</sub>),  $131.1$ ,  $130.5$ ,  $129.6$ ,  $128.0$ ,  $127.9$  (5CH<sub>Ar</sub>),  $117.4$ ,  $117.2$  (C<sub>g</sub>),  $113.3$ ,  $113.1$ ,  $112.9$  (C<sub>IV, Acetal</sub> + C<sub>e</sub>),  $112.2$ ,  $112.0$  (C<sub>j</sub>),  $82.3$ ,  $82.1$ ,  $82.1$ ,  $81.8$ ,  $81.3$ ,  $80.7$ ,  $80.3$ ,  $80.1$  (4C<sub>IV, Boc</sub> + C<sub>3'</sub> + C<sub>7'-8'</sub>),  $69.3$ ,  $68.6$  (C<sub>2'</sub>),  $65.7$ ,  $64.9$  (C<sub>c</sub>),  $58.6$ ,  $58.0$  (C<sub>4</sub>),  $54.8$ ,  $54.3$ ,  $53.9$ ,  $52.8$ ,  $52.5$ ,  $52.3$ ,  $52.2$  (C<sub>4'</sub> or C<sub>6'</sub> + C<sub>2</sub> + C<sub>6</sub>), C<sub>4'</sub> or C<sub>6'</sub> (signal under the CD<sub>3</sub>OD peak),  $44.1$ ,  $42.6$  (C<sub>a</sub>),  $39.8$ ,  $38.9$  (C<sub>7</sub>),  $37.4$ ,  $37.3$  (2CH<sub>2, Acetal</sub>),  $31.4$ ,  $30.5$ ,  $29.1$ ,  $28.9$ ,  $28.8$ ,  $28.7$ ,  $28.5$  (12CH<sub>3, Boc</sub> + C<sub>b</sub> + C<sub>1'</sub> + C<sub>5'</sub> + C<sub>1</sub>),  $26.2$ ,  $24.9$  (3CH<sub>2, Acetal</sub>); **MS (ESI):**  $m/z$  1147.20 [M+H]<sup>+</sup> (theoretical  $m/z$  1147.66).

**Naphthyridine(O-alkyl)-(R/S)-DOS-L-Phe-*tert*-butylacetamide (C38b):**

General procedure B was employed for the deprotection of compound **C38a** (38.0 mg, 0.033 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C38b** was isolated as a white-pale powder (20.0 mg, 54%).

**HPLC:**  $rt_{dia1}$  = 11.77 min and  $rt_{dia2}$  = 12.38 min (method A); NMR data as a mixture of 2 diastereomers; **<sup>1</sup>H NMR** (400 MHz, D<sub>2</sub>O)  $\delta$  = 8.21, 8.20 (2d,  $J$  = 9.3 Hz, 1H, H<sub>k</sub>), 8.15, 8.15 (2d,  $J$  = 8.7 Hz, 1H, H<sub>f</sub>), 7.47 – 7.13 (m, 5H, 5CH<sub>Ar</sub>), 7.00 – 6.88 (m, 2H, H<sub>e</sub> + H<sub>j</sub>), 4.91 – 4.70 (m, 1H, H<sub>d</sub>), 4.63 – 4.53 (m, 1H, H<sub>6</sub>), 4.53 – 4.47, 4.46 – 4.28 (2m, 2H, H<sub>c</sub>), 4.03 – 3.95, 3.94 – 3.85, 3.82 – 3.64 (3m, 2H, H<sub>2'</sub>), 3.57 – 3.47 (m, 2H, H<sub>7'-8'</sub>), 3.47 – 3.06 (m, 7H, H<sub>a</sub> + H<sub>3'-4'</sub> + H<sub>6'</sub> + H<sub>7</sub>), 2.50 – 2.39, 1.88 – 1.68 (2m, 2H, H<sub>5'</sub>), 2.37 – 2.15, 2.09 – 1.68 (2m, 4H, H<sub>b</sub> + H<sub>1'</sub>), 1.35 – 1.19 (m, 9H, H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, D<sub>2</sub>O)  $\delta$  = 170.0, 169.7, 169.4, 169.2 (C<sub>3</sub> + C<sub>5</sub>), 165.4 (C<sub>d</sub>), 155.3 (C<sub>h</sub>), 146.0 (C<sub>i</sub>), 143.3 (C<sub>k</sub>), 140.4 (C<sub>f</sub>), 133.4, 133.3 (C<sub>8</sub>), 129.3, 129.2, 129.1, 128.1, 128.0 (5CH<sub>Ar</sub>), 111.5 (C<sub>g</sub>), 110.8, 110.8, 110.6 (C<sub>e</sub> + C<sub>j</sub>), 80.5, 80.4 (C<sub>3'</sub>), 75.3, 75.2 (C<sub>7'</sub>), 72.5 (C<sub>8'</sub>), 70.0, 69.4 (C<sub>2'</sub>), 64.4, 64.2 (C<sub>c</sub>), 57.2, 56.3 (C<sub>4</sub>), 52.1, 52.0, 51.7, 51.7 (C<sub>2</sub> + C<sub>6</sub>), 49.9, 49.2, 49.2 (C<sub>4'</sub> + C<sub>6'</sub>), 42.3 (C<sub>a</sub>), 37.2, 37.0 (C<sub>7</sub>), 29.8, 29.5 (C<sub>1'</sub>), 28.9, 28.5, 28.1 (C<sub>b</sub> + C<sub>5'</sub>), 27.5, 27.5 (C<sub>1</sub>); **HRMS (ESI):**  $m/z$  667.39288 [M+H]<sup>+</sup> (C<sub>34</sub>H<sub>51</sub>N<sub>8</sub>O<sub>6</sub> requires 667.39261).

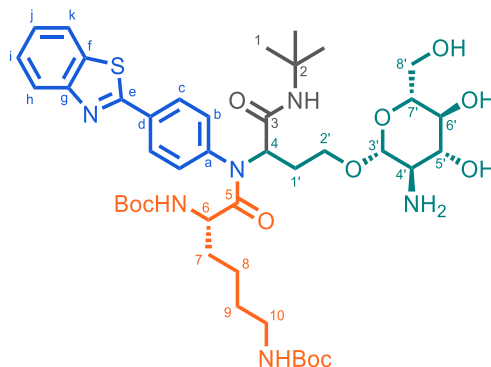
## Série C-3

Adenine-(R/S)-Glucosamine-Phenylacetic-*tert*-butylacetamide (**C39b**):

Compound **C39a** was prepared following general procedure A using **Boc-adenine-NH<sub>2</sub>** (**21**) (30.0 mg, 0.11 mmol) as the amine, **Glucosamine<sub>(Acetyl)</sub>-CHO** (**74**) (49.4 mg, 0.11 mmol) as aldehyde, **phenylacetic acid** (14.7 mg, 0.11 mmol) as the acid and ***tert*-butyl isocyanide** (12.2  $\mu$ l, 0.11 mmol) in MeOH (0.3 mL). After, the reaction mixture was concentrated under reduced pressure and a saturated methanolic ammonia solution (2 ml) was added. The resulting mixture was stirred 12h at room temperature. This step was repeated three times. The reaction mixture was then concentrated under reduced pressure. After general procedure B was employed for the deprotection in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, the crude mixture was purified by flash chromatography on reversed phase (H<sub>2</sub>O/MeCN gradient 95:5 to 50:50, v/v) leading to the diastereomers **C39b** as a white powder (11.3 mg, 12%).

**HPLC**:  $rt_{dia1+2}$  = 12.47 min (method A); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, D<sub>2</sub>O)  $\delta$  = 8.39, 8.34 (2s, 1H, H<sub>f</sub>), 8.23, 8.22, 8.16, 8.13 (4s, 1H, H<sub>c</sub>), 7.43 – 7.23, 7.16 – 7.08, 7.07 – 6.95 (3m, 5H, 5CH<sub>Ar</sub>), 4.93 – 4.86, 4.73 – 4.48 (2m, 4H, H<sub>3'</sub> + H<sub>4</sub> + H<sub>6</sub>), 4.13 – 3.56 (m, 9H, H<sub>a-</sub> + H<sub>2'</sub> + H<sub>5'</sub> + H<sub>8'</sub>), 3.50 – 3.37 (m, 2H, H<sub>6'</sub> + H<sub>7'</sub>), 3.08 – 2.95 (m, 1H, H<sub>4'</sub>), 2.40 – 2.00 (m, 2H, H<sub>1'</sub>), 1.32, 1.21, 1.19 (4s, 9H, H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 176.1, 170.2 (C<sub>3</sub> + C<sub>5</sub>), 149.5, 149.1, 145.0, 144.3, 144.0 (C<sub>c</sub> + C<sub>e-g</sub>), 134.3 (C<sub>7</sub>), 129.0, 128.9, 128.8, 128.8, 128.2, 128.0, 127.5, 127.3 (5CH<sub>Ar</sub>), 117.7 (C<sub>d</sub>), 98.9, 98.7, 98.6 (C<sub>1'</sub>), 76.1, 76.1 (C<sub>7'</sub>), 72.0 (C<sub>5'</sub>), 69.9, 69.7 (C<sub>6'</sub>), 65.7 (C<sub>2'</sub>), 60.5, 60.4 (C<sub>8'</sub>), 58.5, 58.2 (C<sub>4</sub>), 55.7 (C<sub>4'</sub>), 51.8, 51.6 (C<sub>2</sub>), 43.5, 42.9, 42.7, 40.5, 40.2 (C<sub>a</sub> + C<sub>b</sub> + C<sub>6</sub>), 29.3, 28.5 (C<sub>1'</sub>), 27.7, 27.6, 27.5 (C<sub>1</sub>).

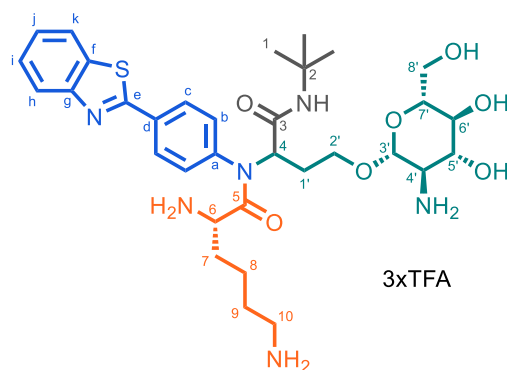
**4-(benzothiazol-2-yl)phenyl-(R/S)-Glucosamine<sub>(Acetyl)</sub>-(Boc)-L-Lys(Boc)-*tert*-butylacetamide (C40a):**



Compound **C40a** was prepared following general procedure A using **2-(4-aminophenyl)benzothiazole** (27.2 mg, 0.12 mmol) as the amine, **Glucosamine<sub>(Acetyl)</sub>-CHO (74)** (45.7 mg, 0.10 mmol) as aldehyde, **Boc-L-Lys(Boc)-OH** (34.6 mg, 0.10 mmol) as the acid and ***tert*-butyl isocyanide** (11.3  $\mu$ l, 0.10 mmol) in MeOH (0.3 mL). After, the reaction mixture was concentrated under reduced pressure and a saturated methanolic ammonia solution was added. The resulting mixture was stirred 12h at room temperature. This step was repeated three times. The reaction mixture was then concentrated under reduced pressure. The crude residue was purified by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 90:10, v/v), a mixture of diastereomers **C40a** was obtained as a yellow-pale solid (70.0 mg, 80%).

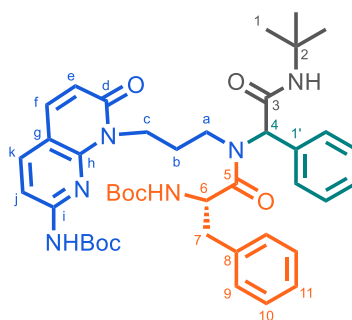
$R_{f,dia1+2} = 0.39$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 93:7); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.28 - 8.19$  (m, 2H, H<sub>c</sub>),  $8.10 - 8.02$  (m, 2H, H<sub>h</sub> + H<sub>k</sub>),  $7.86 - 7.53$  (m, 2H, H<sub>b</sub>),  $7.61 - 7.54$  (m, 1H, H<sub>i</sub>),  $7.52 - 7.45$  (m, 1H, H<sub>j</sub>),  $5.39, 5.11, 5.01 - 4.97$  (2t and m,  $J_{dia1} = 7.1$  Hz and  $J_{dia2} = 7.3$  Hz, 1H, H<sub>4</sub>),  $4.50, 4.47, 4.28, 4.21$  (4d,  $J_{conf.1+2} = 7.8$  Hz,  $J_{dia1} = 7.9$  Hz and  $J_{dia2} = 8.0$  Hz, 1H, H<sub>3'</sub>),  $4.18 - 4.08, 4.05 - 3.97$  (2m, 1H, H<sub>6</sub>),  $3.96 - 3.82, 3.81 - 3.62, 3.61 - 3.53$  (3m, 4H, H<sub>2'</sub> + H<sub>8'</sub>),  $3.52 - 3.22$  (m, 3H, H<sub>5'-7'</sub>),  $3.08 - 2.79$  (m, 2H, H<sub>10</sub>),  $2.69 - 2.56$  (m, 1H, H<sub>4'</sub>),  $1.92 - 1.74, 1.73 - 1.04$  (2m, 35H, 6CH<sub>3,Boc</sub> + H<sub>1'</sub> + H<sub>7-9</sub> + H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 175.8, 175.4$  (C<sub>5</sub>),  $171.8, 170.5$  (C<sub>3</sub>),  $168.2, 168.2$  (C<sub>e</sub>),  $158.4, 158.4, 157.9$  (2CO<sub>Boc</sub>),  $155.2$  (C<sub>g</sub>),  $142.4, 141.8$  (C<sub>a</sub>),  $136.4$  (C<sub>f</sub>),  $135.1, 134.9$  (C<sub>d</sub>),  $132.4$  (2C<sub>b</sub>),  $129.4$  (2C<sub>c</sub>),  $127.8$  (C<sub>i</sub>),  $126.9$  (C<sub>j</sub>),  $124.1, 123.0$  (C<sub>h</sub> + C<sub>k</sub>),  $104.6, 104.5, 103.6$  (C<sub>3'</sub>),  $80.4, 80.3, 79.7, 79.7$  (2C<sub>IV,Boc</sub>),  $78.2, 78.1, 77.6, 77.4, 76.9$  (C<sub>5'</sub> + C<sub>7'</sub>),  $71.7, 71.6$  (C<sub>6'</sub>),  $67.2, 67.1$  (C<sub>2'</sub>),  $62.7$  (C<sub>8'</sub>),  $59.4, 58.3, 58.3, 58.0, 57.0$  (C<sub>4'</sub> + C<sub>4</sub>),  $54.8, 53.4, 53.3, 52.4, 52.3$  (C<sub>2</sub> + C<sub>6</sub>),  $40.8$  (C<sub>10</sub>),  $32.1, 31.0, 30.7, 30.1, 30.0$  (C<sub>1'</sub> + C<sub>7</sub> + C<sub>9</sub>),  $29.0, 29.0, 28.9, 28.8, 28.8, 28.7$  (6CH<sub>3,Boc</sub> + H<sub>1</sub>),  $24.0, 23.7$  (C<sub>2</sub> + C<sub>8</sub>); **MS (ESI):**  $m/z$  873.27 [M+H]<sup>+</sup> (theoretical  $m/z$  873.44).



**4-(benzothiazol-2-yl)phenyl-(R/S)-Glucosamine-L-Lys-tert-butylacetamide (C40b):**

General procedure B was employed for the deprotection of compound **C40a** (54.4 mg, 0.062 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, the crude mixture was purified by flash chromatography on reversed phase (H<sub>2</sub>O/MeCN gradient 95:5 to 50:50, v/v) leading to the diastereomers **C40b** as a white powder (33.0 mg, 52%).

**HPLC:**  $rt_{dia1}$  = 11.89 min and  $rt_{dia2}$  = 12.82 min (method A); **NMR** data as a mixture of 2 diastereomers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.37 – 8.20 (m, 1H, H<sub>c</sub>), 8.11 – 8.03 (m, 2H, H<sub>h</sub> + H<sub>k</sub>), 8.03 – 7.92, 7.70 – 7.60 (2m, 2H, H<sub>b</sub>), 7.60 – 7.53 (m, 1H, H<sub>i</sub>), 7.49 (t,  $J$  = 7.6 Hz, 1H, H<sub>j</sub>), 5.19 – 5.10 (m, 1H, H<sub>4</sub>), 4.57, 4.48 (2d,  $J$  = 8.3 Hz, 1H, H<sub>3'</sub>), 4.10 – 4.01, 3.89 – 3.78 (2m, 1H, H<sub>6</sub>), 3.99 – 3.78, 3.75 – 3.59 (2m, 4H, H<sub>2'</sub> + H<sub>8'</sub>), 3.56, 3.49 (dd,  $J$  = 10.5, 8.2 Hz, 1H, H<sub>5'</sub>), 3.40 – 3.25 (m, 2H, H<sub>6'-7'</sub>), 2.97 – 2.87, 2.86 – 2.71 (m, 3H, H<sub>4'</sub> + H<sub>10</sub>), 2.18 – 1.89, 1.88 – 1.57 (2m, 4H, H<sub>1'</sub> + H<sub>7</sub>), 1.56 – 1.44 (m, 2H, H<sub>9</sub>), 1.43 – 1.11 (m, 2H, H<sub>8</sub>), 1.40, 1.34 (2s, 9H, H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 171.90, 170.6, 170.4, 170.2 (C<sub>3</sub> + C<sub>5</sub>), 167.9 (C<sub>e</sub>), 155.1 (C<sub>g</sub>), 142.0, 140.9 (C<sub>a</sub>), 136.4 (C<sub>f</sub>), 135.8, 135.6 (C<sub>d</sub>), 132.4 (2C<sub>b</sub>), 129.9 (2C<sub>c</sub>), 128.0 (C<sub>i</sub>), 127.2 (C<sub>j</sub>), 124.2, 123.1 (C<sub>h</sub> + C<sub>k</sub>), 100.4, 100.3 (C<sub>3'</sub>), 78.6, 78.4 (C<sub>7'</sub>), 73.9 (C<sub>5'</sub>), 71.9, 71.8 (C<sub>6'</sub>), 67.5, 67.4 (C<sub>2'</sub>), 62.4, 62.2 (C<sub>8'</sub>), 59.7, 59.5 (C<sub>4</sub>), 57.6, 57.5 (C<sub>4'</sub>), 52.7, 52.6, 52.5 (C<sub>2</sub> + C<sub>6</sub>), 40.1 (C<sub>10</sub>), 31.5, 31.4, 31.1, 30.8 (C<sub>1'</sub> + C<sub>7</sub>), 28.9, 28.8 (C<sub>1</sub>), 27.8, 27.7 (C<sub>9</sub>), 22.8, 22.2 (C<sub>8</sub>); **HRMS** (ESI):  $m/z$  673.33813 [M+H]<sup>+</sup> (C<sub>33</sub>H<sub>49</sub>N<sub>6</sub>O<sub>7</sub> requires 673.33779).

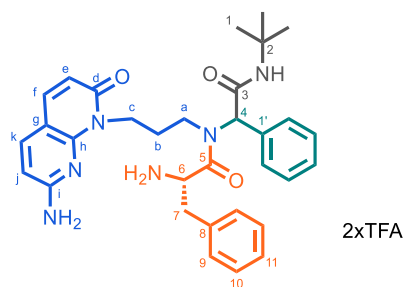
**(Boc)-Naphthyridine(N-alkyl)-(R/S)-Phenyl-(Boc)-(L)-Phe-tert-butylacetamide (C41a):**

Compound **C41a** was prepared following general procedure A using **Boc-naphthyridine(N-alkyl)-NH<sub>2</sub>** (**42**) (31.8 mg, 0.10 mmol) as the amine, **benzaldehyde** (10.2  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-Phe-**

**OH** (26.5 mg, 0.10 mmol) as the acid and **tert-butyl isocyanide** (11.3  $\mu$ l, 0.10 mmol) in MeOH/CH<sub>2</sub>Cl<sub>2</sub> (0.4 mL, 1/1, v/v). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C41a** was obtained as a white solid (55.8 mg, 82%).

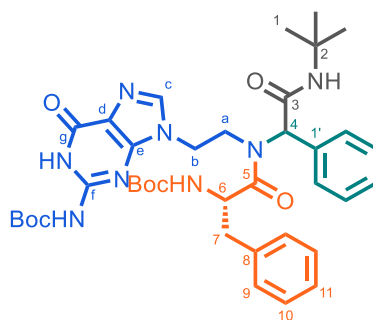
$R_{f, \text{dia}1+2} = 0.29$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.04 - 7.94$  (m, 1H, H<sub>k</sub>), 7.92 - 7.84 (m, 1H, H<sub>j</sub>), 7.82 - 7.72 (m, 1H, H<sub>f</sub>), 7.39 - 7.05, 7.02 - 6.91 (2m, 10H, 10CH<sub>Ar</sub>), 6.56 - 6.42 (m, 1H, H<sub>e</sub>), 5.94, 5.88, 5.76, 5.52 (4s, 1H, H<sub>4</sub>), 4.91 - 4.61 (m, 1H, H<sub>6</sub>), 4.35 - 3.98 (m, 2H, H<sub>c</sub>), 3.70 - 2.75 (m, 4H, H<sub>a</sub> + H<sub>7</sub>), 1.98 - 1.79, 1.67 - 1.01 (2m, 29H, 6CH<sub>3, Boc</sub> + H<sub>b</sub> + H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta = 174.7$  (C<sub>5</sub>), 171.1 (C<sub>3</sub>), 165.2 (C<sub>d</sub>), 160.3, 157.5, 154.5 (2CO<sub>Boc</sub> + C<sub>i</sub>), 149.4 (C<sub>h</sub>), 140.2 (C<sub>k</sub>), 139.5, 139.4 (C<sub>f</sub>), 138.4, 138.0 (C<sub>8</sub>), 136.5 (C<sub>1'</sub>), 130.8, 130.6, 130.3, 129.8, 129.5, 127.7, 127.6 (10CH<sub>Ar</sub>), 119.8, 119.7 (C<sub>e</sub>), 112.9, 112.8 (C<sub>g</sub>), 108.9, 108.7 (C<sub>j</sub>), 82.2, 82.1, 80.7, 80.5 (2C<sub>IV, Boc</sub>), 63.7, 63.3 (C<sub>4</sub>), 54.2, 54.0, 52.4, 52.2 (C<sub>2</sub> + C<sub>6</sub>), 44.9 (C<sub>a</sub>), 39.5, 39.3, 39.1 (C<sub>c</sub> + C<sub>7</sub>), 30.2, 29.8 (C<sub>b</sub>), 28.9, 28.8, 28.8, 28.6 (6CH<sub>3, Boc</sub> + C<sub>1</sub>); **MS (ESI)**:  $m/z$  754.93 [M+H]<sup>+</sup> (theoretical  $m/z$  755.41).

#### Naphthyridine(N-alkyl)-(R/S)-Phenyl-L-Phe-tert-butylacetamide (**C41b**):



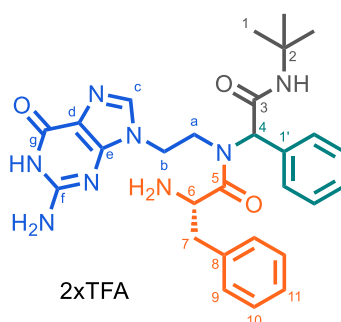
General procedure B was employed for the deprotection of compound **C41a** (23.5 mg, 0.031 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C41b** was isolated as a white powder (16.5 mg, 68%).

**HPLC**:  $rt_{\text{dia}1} = 17.31$  min and  $rt_{\text{dia}2} = 17.44$  min (method A); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta = 7.77 - 7.50$  (m, 2H, H<sub>f</sub> + H<sub>k</sub>), 7.43 - 7.11, 7.10 - 7.00, 6.89 - 6.83, 6.76 - 6.71 (4m, 10H, 10CH<sub>Ar</sub>), 6.48, 6.48 (2d,  $J = 8.5$  Hz, 1H, H<sub>j</sub>), 6.34 - 6.21 (m, 1H, H<sub>e</sub>), 5.93, 5.71, 5.61, 5.19 (4s, 1H, H<sub>4</sub>), 4.45 - 4.31 (m, 1H, H<sub>6</sub>), 4.26 - 4.06, 4.05 - 3.94, 3.90 - 3.78 (3m, 2H, H<sub>c</sub>), 3.41 - 3.00, 2.85 - 2.71 (2m, 2H, H<sub>a</sub>), 3.59 - 3.25, 3.00 - 2.88 (2m, 2H, H<sub>7</sub>), 2.11 - 1.47, 1.21 - 0.97 (2m, 2H, H<sub>b</sub>), 1.33, 1.31, 1.29 (3s, 9H, H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta = 171.6$ , 170.7, 170.5, 170.0 (C<sub>3</sub> + C<sub>5</sub>), 165.7, 165.6 (C<sub>d</sub>), 161.8, 161.6 (C<sub>i</sub>), 150.8, 150.7 (C<sub>h</sub>), 140.4, 140.2, 139.3 (C<sub>f</sub> + C<sub>k</sub>), 135.9, 135.4, 135.0 (C<sub>1'</sub> + C<sub>8</sub>), 131.1, 131.0, 131.0, 130.8, 130.7, 130.5, 130.5, 130.4, 130.4, 130.2, 130.1, 130.0, 129.9, 129.9, 129.7, 129.6, 129.3, 128.9 (10CH<sub>Ar</sub>), 114.9, 114.8 (C<sub>e</sub>), 108.6, 108.6 (C<sub>g</sub>), 106.7, 106.5 (C<sub>j</sub>), 64.6, 63.4 (C<sub>4</sub>), 53.9, 53.1 (C<sub>6</sub>), 52.4, 52.3 (C<sub>2</sub>), 44.9, 44.5 (C<sub>a</sub>), 39.0, 39.0, 38.7, 38.3 (C<sub>c</sub> + C<sub>7</sub>), 30.4, 29.7 (C<sub>b</sub>), 28.8, 28.7, 28.7 (C<sub>1</sub>); **HRMS (ESI)**:  $m/z$  555.30835 [M+H]<sup>+</sup> (C<sub>32</sub>H<sub>39</sub>N<sub>6</sub>O<sub>3</sub> requires 555.30782).

**(Boc)-Guanine-(R/S)-Phenyl-(Boc)-L-Phe-*tert*-butylacetamide (C42a):**

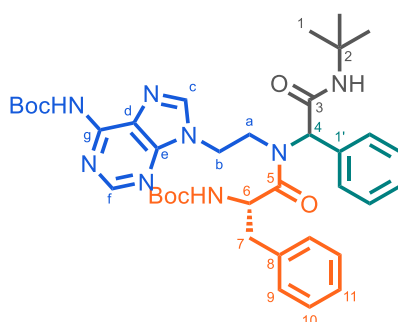
Compound **C42a** was prepared following general procedure A using **Boc-guanine-NH<sub>2</sub> (32)** (29.4 mg, 0.10 mmol) as the amine, **benzaldehyde** (10.2  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-Phe-OH** (26.5 mg, 0.10 mmol) as the acid and ***tert*-butyl isocyanide** (11.3  $\mu$ L, 0.10 mmol) in MeOH/DMF (0.4 mL, 1/1, v/v). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C42a** was obtained as a white solid (40.0 mg, 55%).

$R_{f, \text{dia}1+2} = 0.23$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  = 7.49, 7.39, 7.36, 7.07 (m, 1H, H<sub>c</sub>), 7.46 – 7.33, 7.32 – 7.05, 6.61 – 6.64 (3m, 10H, 10CH<sub>Ar</sub>), 6.05, 5.92, 5.84, 5.63 (4s, 1H, H<sub>4</sub>), 5.25 – 5.18, 5.13 – 5.07, 4.99 – 4.93, 4.89 – 4.81 (4m, 1H, H<sub>6</sub>), 4.37 – 4.20, 4.01 – 3.77, 3.70 – 3.29, 3.22 – 2.97 (4m, 4H, H<sub>a</sub> + H<sub>b</sub>), 3.22 – 2.77 (m, 2H, H<sub>7</sub>), 1.59, 1.58, 1.57 (3s, 9H, 3CH<sub>3, Boc</sub>), 1.45, 1.43, 1.42, 1.38, 1.35, 1.33, 1.30 (7s, 18H, 3CH<sub>3, Boc</sub> + H<sub>1</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 175.6, 175.1, 175.0, 174.5 (C<sub>5</sub>), 171.0, 170.5 (C<sub>3</sub>), 158.0, 157.7, 157.6, 157.5, 157.3, 157.2, 157.0 (CO<sub>Boc</sub> + C<sub>g</sub>), 155.6, 155.5, 155.5, 155.4 (CO<sub>Boc</sub>), 150.8, 150.7 (C<sub>e</sub>), 149.7, 149.6, 149.0, 148.9 (C<sub>f</sub>), 141.2, 140.9, 140.6, 140.5 (C<sub>c</sub>), 138.8, 138.4, 138.2, 138.1 (C<sub>8</sub>), 136.6, 136.3, 136.2 (C<sub>1'</sub>), 131.0, 130.9, 130.8, 130.8, 130.4, 130.1, 130.0, 130.0, 129.9, 129.7, 129.5, 129.5, 129.3, 129.2, 128.2, 128.0, 127.8, 127.6 (10CH<sub>Ar</sub>), 120.5, 120.4 (C<sub>d</sub>), 85.0, 84.9, 84.5, 81.0, 80.8, 80.7, 80.4 (2C<sub>IV, Boc</sub>), 65.0, 64.9, 64.3, 63.2 (C<sub>4</sub>), 54.2, 54.1, 53.6, 53.6 (C<sub>6</sub>), 52.6, 52.5, 52.3 (C<sub>2</sub>), 45.9, 45.4, 44.8, 44.4, 44.4, 43.2, 42.8, 41.5 (C<sub>a</sub> + C<sub>b</sub>), 40.9, 40.0, 38.9, 38.3 (C<sub>7</sub>), 28.9, 28.8, 28.8, 28.8, 28.7, 28.7, 28.5, 28.4, 28.3 (6CH<sub>3, Boc</sub> + C<sub>1</sub>); **MS (ESI):**  $m/z$  731.07 [M+H]<sup>+</sup> (theoretical  $m/z$  731.39).

**Guanine-(R/S)-Phenyl-L-Phe-*tert*-butylacetamide (C42b):**

General procedure B was employed for the deprotection of compound **C42a** (25.0 mg, 0.034 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C42b** was isolated as a white solid (24.5 mg, 95%).

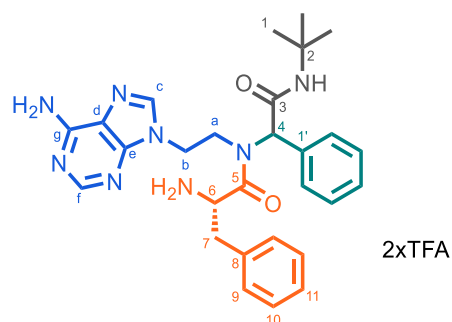
**HPLC:**  $rt_{dia1}$  = 14.15 min and  $rt_{dia2}$  = 14.60 min (method A); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, D<sub>2</sub>O)  $\delta$  = 8.49, 8.24, 8.22, 7.89 (4s, 1H, H<sub>c</sub>), 7.59 – 7.13, 6.68 – 6.58 (2m, 10H, CH<sub>Ar</sub>), 5.88, 5.76, 5.73, 5.29 (4s, 1H, H<sub>4</sub>), 5.10 – 5.02, 4.96 – 4.73, 4.57 – 4.29 (3m, 1H, H<sub>6</sub>), 4.57 – 4.29, 4.28 – 4.03, 4.00 – 3.58 (3m, 4H, H<sub>a</sub> + H<sub>b</sub>), 3.48 – 3.15, 3.11 – 2.97 (2m, 2H, H<sub>7</sub>), 1.30, 1.28, 1.28, 1.22 (4s, 9H, H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, D<sub>2</sub>O)  $\delta$  = 171.9, 171.6, 170.5, 170.5, 170.0, 169.5, 169.4, 169.2 (C<sub>g</sub> + C<sub>3</sub> + C<sub>5</sub>), 155.1, 154.9, 154.9 (C<sub>f</sub>), 149.4, 149.1, 148.8 (C<sub>e</sub>), 137.1, 136.7 (C<sub>c</sub>), 134.0, 133.5, 133.2, 132.4, 132.3 (C<sub>1'</sub> + C<sub>8</sub>), 129.9, 129.7, 129.7, 129.6, 129.4, 129.4, 129.3, 129.3, 129.2, 129.1, 128.9, 128.9, 128.8, 128.6, 128.4, 128.0 (10CH<sub>Ar</sub>), 108.2, 108.2 (C<sub>d</sub>), 64.7, 64.3, 63.6, 62.9 (C<sub>4</sub>), 54.2, 52.7, 52.4, 52.3, 52.1, 52.0, 51.8, 51.6 (C<sub>2</sub> + C<sub>6</sub>), 44.9, 44.2, 43.6, 43.0 (C<sub>a</sub> + C<sub>b</sub>), 37.4, 35.6, 35.2 (C<sub>7</sub>), 27.6, 27.5, 27.5, 27.3 (C<sub>1</sub>); **HRMS (ESI):**  $m/z$  531.28302 [M+H]<sup>+</sup> (C<sub>28</sub>H<sub>35</sub>N<sub>8</sub>O<sub>3</sub> requires 531.28266).

**(Boc)-Adenine-(R/S)-Phenyl-(Boc)-L-Phe-*tert*-butylacetamide (C43a):**

Compound **C43a** was prepared following general procedure A using **Boc-adenine-NH<sub>2</sub> (21)** (90.0 mg, 0.32 mmol) as the amine, **benzaldehyde** (32.7  $\mu$ L, 0.32 mmol) as aldehyde, **Boc-L-Phe-OH** (84.9 mg, 0.32 mmol) as the acid and ***tert*-butyl isocyanide** (36.2  $\mu$ L, 0.32 mmol) in MeOH/CH<sub>2</sub>Cl<sub>2</sub> (0.6 mL, 1/1, v/v). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:2, v/v), a mixture of diastereomers **C43a** was obtained as a pale-yellow solid (117.1 mg, 51%).

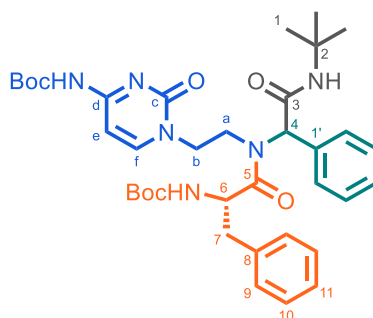
$R_{f,dia1+2} = 0.20$  ( $CH_2Cl_2$ -MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers;  $^1H$  NMR (400 MHz,  $CD_3OD$ )  $\delta = 8.53, 8.46, 8.38, 8.38$  (4s, 1H,  $H_f$ ), 8.02, 7.94, 7.89, 7.51 (4s, 1H,  $H_c$ ), 7.76 – 7.68, 7.47 – 7.15, 7.14 – 6.95, 6.69 – 6.61 (4m, 10H,  $10CH_{Ar}$ ), 6.05, 5.93, 5.82, 5.61 (4s, 1H,  $H_4$ ), 5.00 – 4.70 (m, 1H,  $H_6$ ), 4.48 – 4.22, 4.18 – 3.53, 3.45 – 3.33 (3m, 4H,  $H_a + H_b$ ), 3.19 – 2.77 (m, 2H,  $H_7$ ), 1.59, 1.59, 1.57, 1.56 (4s, 9H,  $3CH_{3.Boc}$ ), 1.43, 1.43, 1.38, 1.37, 1.34, 1.33, 1.32, 1.28 (8s, 18H,  $3CH_{3.Boc} + H_1$ );  $^{13}C$  NMR (101 MHz,  $CD_3OD$ )  $\delta = 175.9, 175.6, 175.1, 174.8$  ( $C_5$ ), 171.1, 170.9, 170.3 ( $C_3$ ), 158.0, 157.6, 157.1 ( $CO_{Boc}$ ), 153.0, 153.0, 152.6, 152.5, 152.4, 152.3, 152.2, 152.1, 151.0, 150.7 ( $CO_{Boc} + C_{e,f,g}$ ), 145.2, 145.0, 144.9, 144.7 ( $C_c$ ), 138.4, 138.3, 138.2, 137.9 ( $C_8$ ), 136.3, 136.2, 136.1, 136.0 ( $C_{1'}$ ), 130.9, 130.9, 130.8, 130.6, 130.4, 130.2, 123.0, 129.8, 129.6, 129.5, 129.5, 129.4, 129.2, 129.1, 128.1, 128.0, 127.8, 127.7 ( $10CH_{Ar}$ ), 122.8, 122.6 ( $C_d$ ), 82.7, 81.0, 80.7, 80.5 ( $2C_{IV.Boc}$ ), 65.0, 64.9, 64.5, 63.6 ( $C_4$ ), 54.4, 53.9, 53.6, 53.5 ( $C_6$ ), 52.6, 52.5, 52.4, 52.3 ( $C_2$ ), 45.8, 45.1, 44.3, 43.1, 42.9 ( $C_a + C_b$ ), 40.8, 40.4, 38.7, 38.3 ( $C_7$ ), 28.9, 28.8, 28.8, 28.7, 28.7, 28.4 ( $6CH_{3.Boc} + C_1$ ); **MS (ESI)**:  $m/z$  715.07  $[M+H]^+$  (theoretical  $m/z$  715.39).

### Adenine-(R/S)-Phenyl-L-Phe-*tert*-butylacetamide (**C43b**):



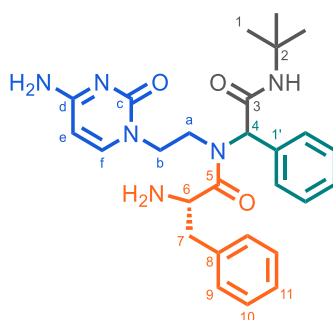
General procedure B was employed for the deprotection of compound **C43a** (70.0 mg, 0.098 mmol) in 1 mL of  $CH_2Cl_2$  and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C43b** was isolated as a white powder (27.9 mg, 78%).

**HPLC**:  $rt_{dia1} = 13.56$  min and  $rt_{dia2} = 13.92$  min (method A); NMR data as a mixture of 2 diastereomers and conformers;  $^1H$  NMR (400 MHz,  $CD_3OD$ )  $\delta = 8.30, 8.22, 8.20, 8.16$  (4s, 1H,  $H_f$ ), 7.98, 7.87, 7.66, 7.51 (4s, 1H,  $H_c$ ), 7.53 – 7.16, 6.85 – 6.79 (2m, 10H,  $10CH_{Ar}$ ), 6.12, 5.88, 5.80, 5.34 (4s, 1H,  $H_4$ ), 5.02 – 4.77, 4.42 – 4.25 (2m, 1H,  $H_6$ ), 4.42 – 4.14, 4.14 – 3.99, 3.99 – 3.83, 3.81 – 3.58, 3.40 – 3.17, 2.92 – 2.75 (5m, 4H,  $H_{a-b}$ ), 3.40 – 3.04 (m, 2H,  $H_7$ ), 1.35, 1.32 (2s, 9H,  $H_1$ );  $^{13}C$  NMR (101 MHz,  $CD_3OD$ )  $\delta = 171.9, 171.6, 171.2, 171.1, 170.9, 170.4, 170.3, 170.2$  ( $C_3 + C_5$ ), 155.1, 155.1, 153.8 ( $C_g$ ), 150.6, 150.3, 148.1, 148.0 ( $C_e + C_f$ ), 144.5, 144.4, 143.2, 143.0 ( $C_c$ ), 136.3, 136.3, 136.2, 135.4, 135.3, 135.1, 135.0, 134.8, 134.6 ( $C_{1'} + C_8$ ), 131.3, 131.2, 131.0, 130.9, 130.8, 130.7, 130.6, 130.5, 130.4, 130.4, 130.3, 130.2, 130.1, 130.1, 130.0, 129.8, 129.3, 129.2, 129.1, 129.0, 128.8 ( $10CH_{Ar}$ ), 120.0, 119.9, 119.6 ( $C_d$ ), 65.8, 65.6, 64.7, 63.1 ( $C_4$ ), 53.9, 53.8, 53.3, 53.1, 52.9, 52.9, 52.5, 52.4 ( $C_2 + C_6$ ), 47.6, 46.1, 45.4, 44.2, 44.2, 43.6, 43.3, 39.3 ( $C_a + C_b$ ), 39.1, 38.4, 37.5, 37.4 ( $C_7$ ), 28.8, 28.8, 28.6 ( $C_1$ ); **HRMS (ESI)**:  $m/z$  515.28809  $[M+H]^+$  ( $C_{28}H_{35}N_8O_2$  requires 515.28775).

**(Boc)-Cytosine-(R/S)-Phenyl-(Boc)-L-Phe-*tert*-butylacetamide (C44a):**

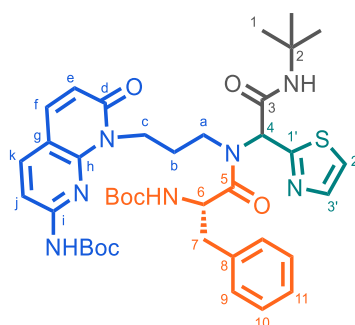
Compound **C44a** was prepared following general procedure A using **Boc-cytosine-NH<sub>2</sub> (25)** (60.0 mg, 0.24 mmol) as the amine, **benzaldehyde** (24.5  $\mu$ L, 0.24 mmol) as aldehyde, **Boc-L-Phe-OH** (63.7 mg, 0.24 mmol) as the acid and ***tert*-butyl isocyanide** (27.1  $\mu$ L, 0.24 mmol) in MeOH/CH<sub>2</sub>Cl<sub>2</sub> (0.4 mL, 1/1, v/v). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:2, v/v), a mixture of diastereomers **C44a** was obtained as a pale-yellow solid (111,7 mg, 69%).

$R_{f, \text{dia}1+2} = 0.20$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta = 7.77, 7.72, (2s, \text{NH}), 7.79 - 7.69, 7.50 - 7.16, 6.80 - 6.74 (3m, 11H, 10\text{CH}_{Ar} + H_f), 7.12 - 6.99 (m, 1H, H_e), 6.09, 6.02, 5.85, 5.67 (4s, 1H, H_4), 5.00 - 4.80 (m, 1H, H_6), 3.95 - 3.33, 3.25 - 3.06 (2m, 4H, H_a + H_b), 3.27 - 2.78 (m, 2H, H_7), 1.52, 1.50 (2s, 9H, 3\text{CH}_{3, \text{Boc}}), 1.44, 1.42, 1.41, 1.38, 1.36, 1.35, 1.34, 1.32 (8s, 18H, 3\text{CH}_{3, \text{Boc}} + H_1)$ ; **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta = 176.1, 175.7, 175.3, 174.8 (C_5), 170.9, 170.9, 170.8, 170.7, 170.6, 170.2, 170.1, 165.1, 164.8 (C_d + C_3), 158.2, 158.1, 157.9, 157.4, 157.1, 153.5, 153.4 (2\text{CO}_{\text{Boc}} + C_c), 150.5, 150.3, 150.0, 149.8 (C_f), 138.4, 138.2, 137.9, 137.8 (C_8), 136.8, 136.6, 136.5, 136.3 (C_{1'}), 130.9, 130.9, 130.8, 130.6, 130.4, 130.3, 130.1, 130.1, 130.0, 129.9, 129.8, 129.6, 129.6, 129.5, 129.4, 129.3, 128.0, 127.8, 127.7, 127.6 (10\text{CH}_{Ar}), 96.8, 96.5 (C_e), 82.9, 82.8, 81.0, 80.7, 80.4 (2\text{C}_{\text{IV, Boc}}), 65.1, 64.9, 64.8, 64.3, 63.2 (C_4), 54.1, 53.9, 53.4, 53.3 (C_6), 52.6, 52.5, 52.4, 52.3 (C_2), 51.0, 50.8, 50.1, 50.0 (C_b), 44.8, 44.5, 44.4 (C_d), 40.6, 40.2, 39.0, 38.6 (C_7), 28.8, 28.8, 28.7, 28.4, 28.3 (6\text{CH}_{3, \text{Boc}} + C_1)$ ; **MS (ESI):**  $m/z$  691.00 [M+H]<sup>+</sup> (theoretical  $m/z$  691.38).

**Cytosine-(R/S)-Phenyl-L-Phe-*tert*-butylacetamide (C44b):**

General procedure B was employed for the deprotection of compound **C44a** (47.9 mg, 0.069 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C44b** was isolated as a white powder (36.4 mg, 73%).

**HPLC:**  $rt_{dia1}$  = 13.68 min and  $rt_{dia2}$  = 13.84 min (method A); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 7.57, 7.07, 6.78, 7.54 – 7.46 (3d and m,  $J$  = 7.5 Hz, 1H,  $H_f$ ), 7.54 – 7.21, 7.03 – 6.97 (2m, 10H, 10CH<sub>Ar</sub>), 6.13, 5.88, 5.87, 5.42 (4s, 1H,  $H_4$ ), 5.95 – 5.88 (m, 1H,  $H_e$ ), 4.96 – 4.80, 4.40 – 4.34, 4.20 – 4.15 (3m, 1H,  $H_6$ ), 3.87 – 3.47, 3.41 – 3.00, 2.79 – 2.65, 2.53 – 2.42 (4m, 4H,  $H_a$  +  $H_b$ ), 3.28 – 3.01, 2.97– 2.85 (m, 2H,  $H_7$ ), 1.36, 1.36, 1.34, 1.33 (4s, 9H,  $H_1$ ); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 171.8, 171.6, 171.2, 171.2, 170.9, 170.4, 170.2, 170.1 ( $C_3$  +  $C_5$ ), 163.4, 163.4, 162.2, 162.2 ( $C_d$ ), 151.9, 150.9, 150.7, 150.0, 149.3, 149.1 ( $C_c$  +  $C_f$ ), 136.4, 136.4, 135.8, 135.7, 135.3, 135.2, 135.1, 134.9, 134.6 ( $C_{1'}$ + $C_8$ ), 131.4, 131.0, 130.9, 130.9, 130.7, 130.5, 130.5, 130.4, 130.3, 130.2, 130.1, 130.1, 129.3, 129.1, 128.9, 128.8 (10CH<sub>Ar</sub>), 95.5, 95.4, 94.7, 94.6 ( $C_e$ ), 65.6, 65.3, 64.4, 62.7 ( $C_4$ ), 55.5, 53.8, 53.5, 53.1, 53.0, 52.9, 52.8, 52.5, 52.4 ( $C_2$  +  $C_6$ ), 50.4, 50.3, 49.0 ( $C_b$ ), 46.3, 45.2, 43.8, 42.8 ( $C_a$ ), 39.2, 39.0, 38.3, 37.5 ( $C_7$ ), 28.8, 28.7, 28.6 ( $C_1$ ); **HRMS (ESI):**  $m/z$  491.27664 [M+H]<sup>+</sup> (C<sub>27</sub>H<sub>35</sub>N<sub>6</sub>O<sub>3</sub> requires 491.27652).

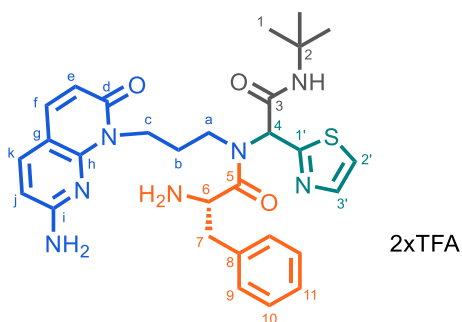
**(Boc)-Naphthyridine(N-alkyl)-(R/S)-Thiazole-(Boc)-L-Phe-*tert*-butylacetamide (C45a):**

Compound **C45a** was prepared following general A procedure using **Boc-naphthyridine(N-alkyl)-NH<sub>2</sub> (42)** (31.8 mg, 0.10 mmol) as the amine, **2-thiazolecarboxaldehyde** (9.4  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-Phe-OH** (26.5 mg, 0.10 mmol) as the acid and ***tert*-butyl isocyanide** (11.3  $\mu$ L, 0.10 mmol) in MeOH

(0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C45a** was obtained as a yellow-pale solid (59.0 mg, 78%).

$R_{f, \text{dia}1+2} = 0.38$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.03 - 7.95$  (m, 1H, H<sub>k</sub>), 7.92 - 7.84 (m, 1H, H<sub>j</sub>), 7.80, 7.79 (2d,  $J = 9.3$  Hz, 1H, H<sub>f</sub>), 7.71, 7.64, 7.61 (3d,  $J = 3,3$  Hz, 1H, H<sub>3'</sub>), 7.54, 7.49, 7.43, 7.41 (4d,  $J = 3,3$  Hz, 1H, H<sub>2'</sub>), 7.27 - 7.01 (m, 5H, 5CH<sub>Ar</sub>), 6.53, 6.51 (2d,  $J = 9.3$  Hz, 1H, H<sub>e</sub>), 4.75, 4.68 (2t,  $J = 7.2$  Hz, 1H, H<sub>6</sub>), 4.46 - 4.17 (m, 2H, H<sub>c</sub>), 3.79 - 3.40 (m, 2H, H<sub>a</sub>), 3.27 - 3.00, 2.94 - 2.72 (2m, 2H, H<sub>7</sub>), 2.09 - 1.74, 1.68 - 1.20 (2m, 2H, H<sub>b</sub>), 1.57, 1.55, 1.39, 1.38, 1.37, 1.33, 1.30, 1.30 (9s, 27H, 6CH<sub>3, Boc</sub> + H<sub>1</sub>), H<sub>4</sub> missing; <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 175.4, 175.0, 174.2$  (C<sub>5</sub>), 167.9 (C<sub>1</sub>), 167.0, 166.4 (C<sub>3</sub>), 165.1, 165.0 (C<sub>d</sub>), 157.5, 157.4, 154.6, 154.5, 154.2 (2CO<sub>Boc</sub> + C<sub>i</sub>), 149.5 (C<sub>h</sub>), 143.1, 142.8 (C<sub>3'</sub>), 140.2 (C<sub>k</sub>), 139.5 (C<sub>f</sub>), 138.2, 138.1 (C<sub>8</sub>), 130.5, 130.4, 129.4, 129.3, 127.6 (5CH<sub>Ar</sub>), 122.4, 122.2 (C<sub>2'</sub>), 119.7 (C<sub>e</sub>), 112.8, 112.8 (C<sub>g</sub>), 108.8, 108.8 (C<sub>j</sub>), 82.1, 80.8, 80.6 (2C<sub>IV, Boc</sub>), 53.8, 53.6, 52.9, 52.6, 52.5 (C<sub>2</sub> + C<sub>6</sub>), 46.8, 46.2 (C<sub>a</sub>), 39.5, 39.3, 39.2 (C<sub>c</sub> + C<sub>7</sub>), 29.7, 29.6 (C<sub>b</sub>), 28.8, 28.8, 28.8, 28.6 (6CH<sub>3, Boc</sub> + C<sub>1</sub>), C<sub>4</sub> missing; **MS (ESI):**  $m/z$  762.00 [M+H]<sup>+</sup> (theoretical  $m/z$  762.36).

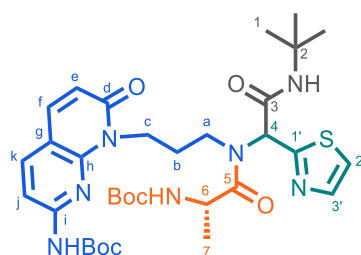
#### Naphthyridine(*N*-alkyl)-(R/S)-Thiazole-L-Phe-*tert*-butylacetamide (**C45b**):



General procedure B was employed for the deprotection of compound **C45a** (43.1 mg, 0.057mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C45b** was isolated as a yellow-pale powder (38.2 mg, 86%).

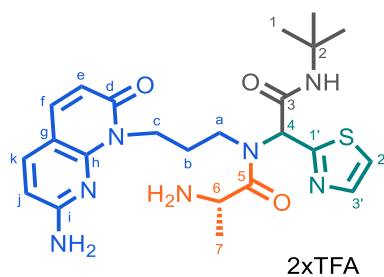
**HPLC:**  $rt_{\text{dia}1} = 16.01$  min and  $rt_{\text{dia}2} = 16.44$  min (method A); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 7.78, 7.74$  (2d,  $J = 3.3$  Hz, 1H, H<sub>3'</sub>), 7.72 - 7.61 (m, 2H, H<sub>f</sub> + H<sub>k</sub>), 7.56, 7.54, 7.51 (3d,  $J = 3.3$  Hz, 1H, H<sub>2'</sub>), 7.37 - 7.06 (m, 5H, 5CH<sub>Ar</sub>), 6.49 (d,  $J = 8.5$  Hz, 1H, H<sub>j</sub>), 6.33, 6.32 (2d,  $J = 9.2$  Hz, 1H, H<sub>e</sub>), 4.82 - 4.74, 4.57 - 4.10 (2m, 4H, H<sub>c</sub> + H<sub>6</sub>), 3.74 - 3.19 (m, 2H, H<sub>a</sub>), 3.33 - 2.83 (m, 2H, H<sub>7</sub>), 2.07 - 1.74, 1.68 - 1.26 (2m, 2H, H<sub>b</sub>), 1.34, 1.32, 1.31 (3s, 9H, H<sub>1</sub>), H<sub>4</sub> missing; <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 170.8, 170.3$  (C<sub>5</sub>), 168.0, 167.6 (C<sub>3</sub>), 166.1, 165.7, 165.6, 165.1 (C<sub>d</sub> + C<sub>1'</sub>), 161.9, 161.8 (C<sub>i</sub>), 150.9, 150.8 (C<sub>h</sub>), 143.5, 143.2 (C<sub>3'</sub>), 140.4, 140.4, 139.3 (C<sub>f</sub> + C<sub>k</sub>), 135.1, 135.0 (C<sub>8</sub>), 130.6, 130.5, 130.2, 130.1, 129.0, 128.8 (5CH<sub>Ar</sub>), 122.9, 122.8 (C<sub>2'</sub>), 114.9 (C<sub>e</sub>), 108.5 (C<sub>g</sub>), 106.7, 106.6 (C<sub>j</sub>), 53.5, 53.2, 52.8, 52.7 (C<sub>2</sub> + C<sub>6</sub>), 46.6, 45.6 (C<sub>a</sub>), 39.1, 38.9, 38.7, 38.3 (C<sub>c</sub> + C<sub>7</sub>), 29.6, 29.5 (C<sub>b</sub>), 28.7, 28.6 (C<sub>1</sub>), C<sub>4</sub> missing; **HRMS (ESI):**  $m/z$  562.26025 [M+H]<sup>+</sup> (C<sub>29</sub>H<sub>36</sub>N<sub>7</sub>O<sub>3</sub>S requires 562.25949).



**(Boc)-Naphthyridine(N-alkyl)-(R/S)-Thiazole-(Boc)-L-Ala-tert-butylacetamide (C46a):**

Compound **C46a** was prepared following general procedure A using **Boc-naphthyridine(N-alkyl)-NH<sub>2</sub> (42)** (31.8 mg, 0.10 mmol) as the amine, **2-thiazolecarboxaldehyde** (9.4  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-Ala-OH** (18.9 mg, 0.10 mmol) as the acid and **tert-butyl isocyanide** (11.3  $\mu$ L, 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C46a** was obtained as an orange solid (55.8 mg, 82%).

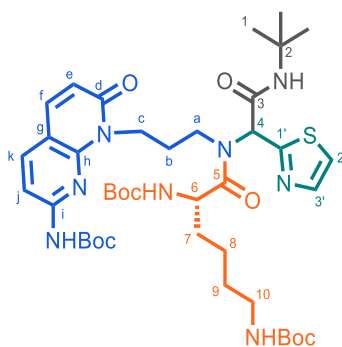
$R_{f,dia1}$  = 0.29 and  $R_{f,dia2}$  = 0.32 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.03 – 7.95 (m, 1H, H<sub>k</sub>), 7.91 – 7.83 (m, 1H, H<sub>j</sub>), 7.83 – 7.74 (m, 1H, H<sub>f</sub>), 7.72, 7.66, 7.60 (3d,  $J$  = 3.3 Hz, 1H, H<sub>3'</sub>), 7.55 – 7.40 (m, 1H, H<sub>2'</sub>), 7.01, 6.91, 6.76 (3d,  $J_{dia1}$  = 6.0 Hz and  $J_{dia2}$  = 6.9 Hz, NH), 6.59 – 6.45 (m, 1H, H<sub>e</sub>), 6.25, 6.06, 5.87 (3s, 1H, H<sub>4</sub>), 4.80 – 4.02 (m, 3H, H<sub>c</sub> + H<sub>6</sub>), 3.82 – 3.41, 3.12 – 3.00 (2m, 2H, H<sub>a</sub>), 2.30 – 1.50 (m, 2H, H<sub>b</sub>), 1.57, 1.56 (2s, 9H, 3CH<sub>3,Boc</sub>), 1.46 – 1.36, 1.35 – 1.22 (2m, 21H, 3CH<sub>3,Boc</sub> + H<sub>1</sub> + H<sub>7</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 176.8, 176.7, 175.9 (C<sub>5</sub>), 168.1, 167.9, 167.8 (C<sub>3</sub>), 166.6, 166.6, 165.2, 165.1, 164.5 (C<sub>d</sub> + C<sub>1'</sub>), 157.9, 157.6, 157.3, 154.6, 154.5, 154.2, 154.1 (2CO<sub>Boc</sub> + C<sub>i</sub>), 149.7, 149.6, 149.5 (C<sub>h</sub>), 143.0, 142.9, 142.7 (C<sub>3'</sub>), 140.2, 140.1 (C<sub>k</sub>), 139.6, 139.5 (C<sub>f</sub>), 122.7, 122.4, 122.3 (C<sub>2'</sub>), 119.7 (C<sub>e</sub>), 112.8, 112.7, 112.7 (C<sub>g</sub>), 108.9, 108.8, 108.7 (C<sub>j</sub>), 82.1, 82.0, 80.6, 80.4 (2C<sub>IV,Boc</sub>), 52.9, 52.7, 52.5 (C<sub>2</sub>), 48.2, 48.0, 47.7 (C<sub>6</sub>), 46.7, 46.4, 44.7 (C<sub>a</sub>), 39.7, 39.4, 39.3 (C<sub>c</sub>), 29.8, 29.8, 27.5 (C<sub>b</sub>), 28.9, 28.8, 28.8, 28.7, 28.7, 28.6 (6CH<sub>3,Boc</sub> + C<sub>1</sub>), 18.1, 17.8, 17.6 (C<sub>7</sub>), C<sub>4</sub> missing; **MS (ESI)**:  $m/z$  686.00 [M+H]<sup>+</sup> (theoretical  $m/z$  686.33).

**Naphthyridine(N-alkyl)-(R/S)-Thiazole-L-Ala-tert-butylacetamide (C46b):**

General procedure B was employed for the deprotection of compound **C46a** (41.1 mg, 0.056 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C46b** was isolated as a yellow-pale powder (25.3 mg, 60%).

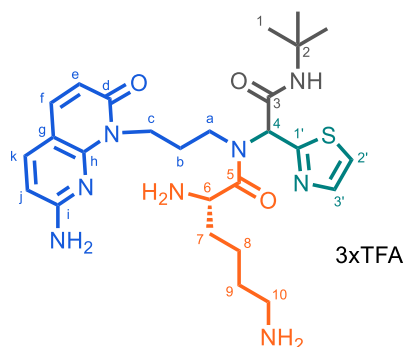
**HPLC:**  $rt_{dia1} = 13.66$  min and  $rt_{dia2} = 13.96$  min (method A); NMR data as a mixture of 2 diastereomers and conformers;  $^1\text{H NMR}$  (400 MHz,  $\text{D}_2\text{O}$ )  $\delta = 7.67 - 7.57$  (m, 2H,  $H_f + H_k$ ), 7.47, 7.42 (2d,  $J = 3.3$  Hz, 1H,  $H_{3'}$ ), 7.34, 7.29, 7.21 (3d,  $J = 3.3$  Hz, 1H,  $H_{2'}$ ), 6.52, 6.50 (2d,  $J = 8.5$  Hz, 1H,  $H_j$ ), 6.34, 6.29, 6.28 (3d,  $J = 9.1$  Hz, 1H,  $H_e$ ), 4.65 – 4.38 (m, 1H,  $H_6$ ), 4.25 – 3.80 (m, 2H,  $H_c$ ), 3.75 – 3.20, 3.01 – 2.86 (2m, 2H,  $H_a$ ), 1.99 – 1.77, 1.70 – 1.09 (2m, 2H,  $H_b$ ), 1.57, 1.55, 1.50, 1.42 (4d,  $J = 7.0$  Hz, 3H,  $H_7$ ), 1.22, 1.21, 1.20 (3s, 9H,  $H_1$ ),  $H_4$  missing;  $^{13}\text{C NMR}$  (101 MHz,  $\text{D}_2\text{O}$ )  $\delta = 171.4$ , 171.1, 170.6 ( $C_5$ ), 167.0, 166.7, 166.4 ( $C_3$ ), 164.6 ( $C_d$ ), 162.2 ( $C_i$ ), 159.4 ( $C_{1'}$ ), 147.6 ( $C_h$ ), 141.6 ( $C_{3'}$ ), 140.1 ( $C_f$ ), 139.2 ( $C_k$ ), 122.6 ( $C_{2'}$ ), 113.8 ( $C_e$ ), 108.1, 108.1 ( $C_g$ ), 106.7, 106.7 ( $C_j$ ), 53.3, 52.0, 51.9 ( $C_2$ ), 47.4, 47.2 ( $C_6$ ), 44.8, 44.1 ( $C_a$ ), 37.9, 37.8 ( $C_c$ ), 28.4, 28.3 ( $C_b$ ), 27.4, 27.3 ( $C_{10}$ ), 16.3, 16.2 ( $C_7$ ),  $C_4$  missing; **HRMS (ESI):**  $m/z$  486.22876  $[\text{M}+\text{H}]^+$  ( $\text{C}_{23}\text{H}_{32}\text{N}_7\text{O}_3\text{S}$  requires 486.22819).

**(Boc)-Naphthyridine(*N*-alkyl)-(R/S)-Thiazole-(Boc)-L-Lys(Boc)-*tert*-butylacetamide (**C47a**):**



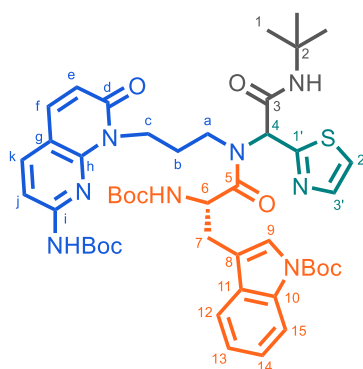
Compound **C47a** was prepared following general procedure A using **Boc-naphthyridine(*N*-alkyl)-NH<sub>2</sub> (42)** (31.8 mg, 0.10 mmol) as the amine, **2-thiazolecarboxaldehyde (42)** (9.4  $\mu\text{L}$ , 0.10 mmol) as aldehyde, **Boc-Lys(Boc)-OH** (34.6 mg, 0.10 mmol) as the acid and ***tert*-butyl isocyanide** (11.3  $\mu\text{L}$ , 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography ( $\text{CH}_2\text{Cl}_2/\text{MeOH}$  gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C47a** was obtained as a yellow solid (58.0 mg, 69%).

$R_{f,dia1+2} = 0.32$  ( $\text{CH}_2\text{Cl}_2$ -MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers;  $^1\text{H NMR}$  (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 8.03 - 7.94$  (m, 1H,  $H_k$ ), 7.93 – 7.73 (m, 2H,  $H_f + H_j$ ), 7.73, 7.70, 7.63 (3d,  $J = 3.3$  Hz, 1H,  $H_{3'}$ ), 7.54, 7.53, 7.47, 7.44 (4d,  $J = 3,3$  Hz, 1H,  $H_{2'}$ ), 6.97, 6.82, 6.72 (3d,  $J_{dia1} = 7.2$  Hz and  $J_{dia2} = 8.0$  Hz, NH), 6.57 – 6.45 (m, 1H,  $H_e$ ), 6.32, 6.20, 5.96, 5.89 (4s, 1H,  $H_4$ ), 4.67 – 4.14 (m, 3H,  $H_c + H_6$ ), 3.84 – 3.37 (m, 2H,  $H_a$ ), 3.18 – 2.90 (m, 2H,  $H_{10}$ ), 2.29 – 1.13 (m, 44H,  $9\text{CH}_{3,Boc} + H_b + H_1 + H_{7-9}$ ),  $H_4$  missing;  $^{13}\text{C NMR}$  (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 175.9$ , 175.2 ( $C_5$ ), 167.8, 166.7, 165.2 ( $C_d + C_{1'} + C_3$ ), 158.5, 158.3, 157.9, 157.6, 154.6, 154.5, 154.2, 154.1 ( $3\text{C}_{O,Boc} + C_i$ ), 149.6, 149.6 ( $C_h$ ), 143.0, 142.8 ( $C_{3'}$ ), 140.3, 140.2 ( $C_k$ ), 139.6, 139.5 ( $C_f$ ), 122.4, 122.3 ( $C_{2'}$ ), 119.8 ( $C_e$ ), 112.8, 112.8, 112.7 ( $C_g$ ), 108.8, 108.78, 108.7 ( $C_j$ ), 82.1, 82.0, 80.6, 80.4, 79.8 ( $3\text{C}_{IV,Boc}$ ), 62.5, 61.7, 61.7, 62.0 ( $C_4$ ), 52.9, 52.6, 52.5, 52.4, 52.4, 52.1 ( $C_2 + C_6$ ), 46.8, 44.7 ( $C_a$ ), 41.0, 40.9 ( $C_{10}$ ), 39.4, 39.4 ( $C_c$ ), 33.0, 32.7 ( $C_7$ ), 30.6, 30.0, 28.8, 28.8, 28.8, 28.7, 28.6, 28.6 ( $9\text{CH}_{3,Boc} + C_b + C_1 + C_9$ ), 24.1, 24.1, 23.9 ( $C_8$ ),  $C_4$  missing; **MS (ESI):**  $m/z$  843.07  $[\text{M}+\text{H}]^+$  (theoretical  $m/z$  843.44).

**Naphthyridine(*N*-alkyl)-(R/S)-Thiazole-L-Lys-*tert*-butylacetamide (**C47b**):**

General procedure B was employed for the deprotection of compound **C47a** (42.9 mg, 0.051 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C47b** was isolated as a yellow-pale powder (36.5 mg, 81%).

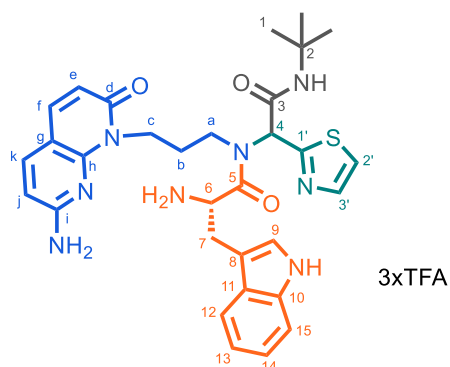
**HPLC**:  $t_{\text{dia}1+2}$  = 12.17 min (method A); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, D<sub>2</sub>O)  $\delta$  = 7.82 – 7.64 (m, 2H, H<sub>f</sub> + H<sub>k</sub>), 7.51, 7.49, 7.39 (3d,  $J$  = 3.3 Hz, 1H, H<sub>3'</sub>), 7.33, 7.30, 7.21 (3d,  $J$  = 3.3 Hz, 1H, H<sub>2'</sub>), 6.60, 6.58, 6.52 (3d,  $J$  = 8.2 Hz, 1H, H<sub>j</sub>), 6.37, 6.35 (2d,  $J$  = 9.4 Hz, 1H, H<sub>e</sub>), 4.65 – 4.59, 4.50 – 4.39 (2m, 1H, H<sub>6</sub>), 4.39 – 3.83 (m, 2H, H<sub>c</sub>), 3.75 – 3.37, 3.33 – 3.17 (m, 2H, H<sub>a</sub>), 3.07 – 2.88 (m, 2H, H<sub>10</sub>), 2.03 – 1.83 (m, 2H, H<sub>7</sub>), 1.82 – 1.58 (m, 2H, H<sub>9</sub>), 1.57 – 1.30 (m, 2H, H<sub>8</sub>), 2.03 – 1.57, 1.35 – 1.12 (2m, 2H, H<sub>b</sub>), 1.23, 1.22, 1.20 (3s, 9H, H<sub>1</sub>), H<sub>4</sub> missing; **<sup>13</sup>C NMR** (101 MHz, D<sub>2</sub>O)  $\delta$  = 170.4, 170.0 (C<sub>5</sub>), 166.9, 166.7 (C<sub>3</sub>), 164.6 (C<sub>d</sub>), 162.7, 161.9 (C<sub>1'</sub>), 159.1 (C<sub>i</sub>), 147.2 (C<sub>h</sub>), 141.6 (C<sub>3'</sub>), 140.2 (C<sub>f</sub>), 139.7 (C<sub>k</sub>), 122.7 (C<sub>2'</sub>), 114.0 (C<sub>e</sub>), 108.2 (C<sub>g</sub>), 106.9, 106.8 (C<sub>j</sub>), 52.0, 51.8 (C<sub>2</sub>), 50.7 (C<sub>6</sub>), 44.8, 44.0 (C<sub>a</sub>), 38.9 (C<sub>10</sub>), 38.0, 37.9 (C<sub>c</sub>), 30.4 (C<sub>7</sub>), 28.5, 28.4 (C<sub>b</sub>), 27.4 (C<sub>1</sub>), 26.5, 26.4 (C<sub>9</sub>), 21.2, 21.0 (C<sub>8</sub>), C<sub>4</sub> missing; **HRMS (ESI)**:  $m/z$  543.28650 [M+H]<sup>+</sup> (C<sub>26</sub>H<sub>39</sub>N<sub>8</sub>O<sub>3</sub>S requires 543.28603).

**(Boc)-Naphthyridine(*N*-alkyl)-(R/S)-Thiazole-(Boc)-L-Trp(Boc)-*tert*-butylacetamide (**C48a**):**

Compound **C48a** was prepared following general A procedure using **Boc-naphthyridine(*N*-alkyl)-NH<sub>2</sub> (**42**)** (31.8 mg, 0.10 mmol) as the amine, **2-thiazolecarboxaldehyde** (9.4  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-Trp(Boc)-OH** (40.5 mg, 0.10 mmol) as the acid and ***tert*-butyl isocyanide** (11.3  $\mu$ L, 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C48a** was obtained as a dark-yellow solid (75.0 mg, 83%).

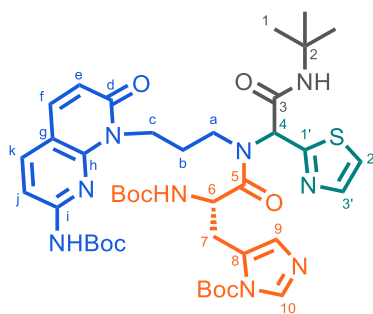
$R_{f, \text{dia}1+2} = 0.35$  ( $\text{CH}_2\text{Cl}_2$ -MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers;  $^1\text{H NMR}$  (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 8.22 - 7.44$  (m, 6H,  $H_f + H_{j-k} + H_{3'} + H_{12} + H_{15}$ ),  $7.44 - 7.26$  (m, 2H,  $H_{2'} + H_9$ ),  $7.25 - 6.97$  (m, 2H,  $H_{13-14}$ ),  $6.45$  (m, 1H,  $H_e$ ),  $6.23$ ,  $6.11$ ,  $5.94$ ,  $5.68$  (4s, 1H,  $H_4$ ),  $5.02 - 4.75$  (m, 1H,  $H_6$ ),  $4.44 - 4.12$  (m, 2H,  $H_c$ ),  $3.76 - 3.43$  (m, 2H,  $H_a$ ),  $3.36 - 2.85$  (m, 2H,  $H_7$ ),  $2.11 - 1.02$  (m, 38H,  $9\text{CH}_{3, \text{Boc}} + H_b + H_1$ );  $^{13}\text{C NMR}$  (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 175.3$ ,  $174.4$ ,  $173.5$  ( $C_5$ ),  $167.8$ ,  $167.7$ ,  $167.4$ ,  $166.6$ ,  $165.2$ ,  $165.1$ ,  $165.0$ ,  $164.9$  ( $C_d + C_{1'} + C_3$ ),  $157.9$ ,  $157.5$ ,  $157.4$ ,  $154.5$ ,  $154.4$ ,  $154.3$ ,  $154.1$ ,  $154.1$ ,  $154.0$  ( $2\text{CO}_{\text{Boc}} + C_i$ ),  $150.8$ ,  $149.6$ ,  $149.4$ ,  $149.4$  ( $\text{CO}_{\text{Boc}} + C_h$ ),  $143.1$ ,  $142.8$ ,  $142.7$  ( $C_{3'}$ ),  $140.1$ ,  $140.0$  ( $C_k$ ),  $139.4$  ( $C_f$ ),  $136.6$  ( $C_{10}$ ),  $131.7$ ,  $131.7$ ,  $131.6$  ( $C_{11}$ ),  $125.7$ ,  $125.3$ ,  $125.2$ ,  $123.6$  ( $C_9 + C_{13} + C_{14}$ ),  $122.2$ ,  $121.8$  ( $C_{2'}$ ),  $120.2$ ,  $120.0$ ,  $119.9$ ,  $119.7$ ,  $119.7$  ( $C_{12} + C_e$ ),  $117.3$ ,  $117.2$ ,  $116.8$  ( $C_8$ ),  $116.0$  ( $C_{15}$ ),  $112.7$ ,  $112.6$  ( $C_g$ ),  $108.7$ ,  $108.6$  ( $C_j$ ),  $84.6$ ,  $82.0$ ,  $80.9$ ,  $80.8$ ,  $80.7$  ( $3\text{C}_{\text{IV, Boc}}$ ),  $62.5$ ,  $61.4$ ,  $61.1$ ,  $62.2$  ( $C_4$ ),  $52.8$ ,  $52.6$ ,  $52.6$ ,  $52.5$ ,  $52.3$ ,  $52.1$ ,  $51.9$  ( $C_2 + C_6$ ),  $47.4$ ,  $46.3$ ,  $44.9$  ( $C_a$ ),  $39.8$ ,  $39.5$ ,  $39.3$  ( $C_c$ ),  $29.5$ ,  $29.0$ ,  $28.8$ ,  $28.6$ ,  $28.5$ ,  $27.5$  ( $9\text{CH}_{3, \text{Boc}} + C_b + C_1 + C_7$ ); **MS (ESI)**:  $m/z$  901.07  $[\text{M}+\text{H}]^+$  (theoretical  $m/z$  901.43).

### Naphthyridine(*N*-alkyl)-(*R/S*)-Thiazole-*L*-Trp-*tert*-butylacetamide (**C48b**):



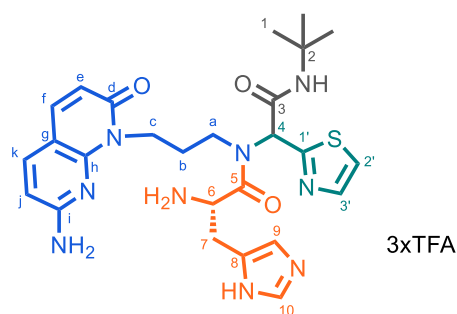
General procedure B was employed for the deprotection of compound **C48a** (53.9 mg, 0.060mmol) in 1 mL of  $\text{CH}_2\text{Cl}_2$  and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C48b** was isolated as a yellow-pale powder (48.8 mg, 87%).

**HPLC**:  $rt_{\text{dia}1} = 12.70$  min and  $rt_{\text{dia}2} = 13.08$  min (method A); NMR data as a mixture of 2 diastereomers and conformers;  $^1\text{H NMR}$  (400 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 7.71 - 7.55$  (m, 4H,  $H_f + H_k + H_{3'} + H_{12}$ ),  $7.50 - 7.41$  (m, 1H,  $H_{2'}$ ),  $7.42 - 7.31$  (m, 1H,  $H_{15}$ ),  $7.23 - 6.95$  (m, 3H,  $H_9 + H_{13} + H_{14}$ ),  $6.44$  (d,  $J = 8.5$  Hz, 1H,  $H_j$ ),  $6.29$ ,  $6.28$ , (2d,  $J = 9.2$  Hz, 1H,  $H_e$ ),  $4.80 - 4.72$ ,  $4.61 - 4.53$  (2m, 1H,  $H_6$ ),  $4.41 - 3.99$  (m, 2H,  $H_c$ ),  $3.64 - 3.09$  (m, 4H,  $H_a + H_7$ ),  $1.96 - 1.20$  (m, 2H,  $H_b$ ),  $1.34$ ,  $1.32$ ,  $1.31$ ,  $1.28$  (4s, 9H,  $H_1$ ),  $H_4$  missing;  $^{13}\text{C NMR}$  (101 MHz,  $\text{CD}_3\text{OD}$ )  $\delta = 171.4$ ,  $170.9$  ( $C_5$ ),  $168.0$ ,  $167.6$  ( $C_3$ ),  $166.2$ ,  $165.6$ ,  $165.6$ ,  $165.2$  ( $C_d + C_{1'}$ ),  $161.7$ ,  $161.7$  ( $C_i$ ),  $150.8$ ,  $150.7$  ( $C_h$ ),  $143.3$ ,  $143.2$  ( $C_{3'}$ ),  $140.4$ ,  $140.3$ ,  $139.3$ ,  $139.2$  ( $C_f + C_k$ ),  $138.3$ ,  $138.1$  ( $C_{10}$ ),  $128.3$ ,  $128.2$  ( $C_{11}$ ),  $125.9$ ,  $125.7$  ( $C_9$ ),  $123.1$ ,  $123.0$ ,  $122.9$ ,  $122.6$ ,  $120.7$ ,  $120.6$  ( $C_{2'} + C_{13} + C_{14}$ ),  $118.8$  ( $C_{12}$ ),  $114.9$  ( $C_e$ ),  $112.8$ ,  $112.7$  ( $C_{15}$ ),  $108.5$ ,  $107.6$ ,  $107.5$  ( $C_g + C_8$ ),  $106.6$ ,  $106.5$  ( $C_j$ ),  $52.7$ ,  $52.5$  ( $C_2 + C_6$ ),  $46.8$ ,  $45.6$  ( $C_a$ ),  $39.0$ ,  $38.9$  ( $C_c$ ),  $29.5$ ,  $29.5$  ( $C_b$ ),  $28.8$ ,  $28.7$ ,  $28.6$  ( $C_1 + C_7$ ),  $C_4$  missing; **HRMS (ESI)**:  $m/z$  601.27094  $[\text{M}+\text{H}]^+$  ( $\text{C}_{31}\text{H}_{37}\text{N}_8\text{O}_3\text{S}$  requires 601.27038).

**(Boc)-Naphthyridine(*N*-alkyl)-(R/S)-Thiazole-(Boc)-L-His(Boc)-*tert*-butylacetamide (**C49a**):**

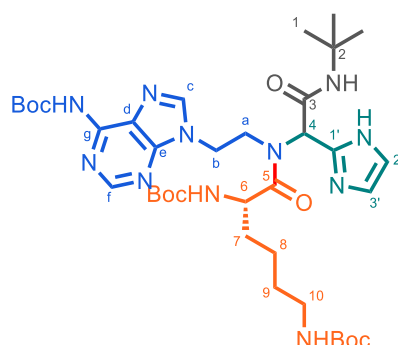
Compound **C49a** was prepared following general A procedure using **Boc-naphthyridine(*N*-alkyl)-NH<sub>2</sub> (42)** (31.8 mg, 0.10 mmol) as the amine, **2-thiazolecarboxaldehyde** (9.4  $\mu$ L, 0.10 mmol) as aldehyde, **Boc-L-His(Boc)-OH** (35.5 mg, 0.10 mmol) as the acid and ***tert*-butyl isocyanide** (11.3  $\mu$ L, 0.10 mmol) in MeOH (0.2 mL). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 97:3, v/v), a mixture of diastereomers **C49a** was obtained as a yellow solid (51.9 mg, 61%).

$R_f$ .<sub>dia1+2</sub> = 0.29 (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 97:3); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, CD<sub>3</sub>OD)  $\delta$  = 8.20 – 8.03 (m, 1H, H<sub>10</sub>), 8.02 – 7.94 (m, 1H, H<sub>k</sub>), 7.92 – 7.83 (m, 1H, H<sub>j</sub>), 7.83 – 7.74 (m, 1H, H<sub>f</sub>), 7.73 – 7.60 (m, 1H, H<sub>3'</sub>), 7.55 – 7.39 (m, 1H, H<sub>2'</sub>), 7.36 – 7.15 (m, 1H, H<sub>9</sub>), 6.56 – 6.47 (m, 1H, H<sub>e</sub>), 6.32, 6.19, 5.94 (3s, 1H, H<sub>4</sub>), 5.02 – 4.75 (m, 1H, H<sub>6</sub>), 4.55 – 4.18 (m, 2H, H<sub>c</sub>), 3.86 – 3.40 (m, 2H, H<sub>a</sub>), 3.29 – 3.16, 3.11 – 2.92, 2.91 – 2.76 (m, 2H, H<sub>7</sub>), 2.22 – 1.41 (m, 2H, H<sub>b</sub>), 1.64 – 1.15 (m, 36H, 9CH<sub>3.Boc</sub> + H<sub>1</sub>); **<sup>13</sup>C NMR** (101 MHz, CD<sub>3</sub>OD)  $\delta$  = 175.1, 174.9, 174.2 (C<sub>5</sub>), 168.0, 167.8, 166.8, 166.5, 165.1 (C<sub>d</sub> + C<sub>1'</sub> + C<sub>3</sub>), 157.9, 157.5, 157.3, 154.6, 154.6, 154.5, 154.2 (2CO<sub>Boc</sub> + C<sub>i</sub>), 149.7, 149.7, 148.1 (CO<sub>Boc</sub> + C<sub>h</sub>), 143.0, 142.9 (C<sub>3'</sub>), 140.1 (C<sub>k</sub>), 139.8, 139.7, 139.5, 139.4, 138.3 (C<sub>f</sub> + C<sub>8</sub> + C<sub>10</sub>), 122.5, 122.4, 122.2 (C<sub>2'</sub>), 119.8 (C<sub>e</sub>), 116.7, 116.4, 116.3 (C<sub>9</sub>), 112.8, 112.7 (C<sub>g</sub>), 108.7, 108.7 (C<sub>j</sub>), 87.0, 86.9, 82.0, 82.0, 80.7, 80.5 (3C<sub>IV.Boc</sub>), 61.8, 61.4 (C<sub>4</sub>), 52.9, 52.6, 52.5, 52.0, 51.8 (C<sub>2</sub> + C<sub>6</sub>), 46.8, 46.6 (C<sub>a</sub>), 39.8, 39.5, 39.4 (C<sub>c</sub>), 31.9, 31.6 (C<sub>7</sub>), 29.8 (C<sub>b</sub>), 28.8, 28.8, 28.7, 28.6, 28.6, 28.1, 28.0 (9CH<sub>3.Boc</sub> + C<sub>1</sub>); **MS (ESI):**  $m/z$  851.93 [M+H]<sup>+</sup> (theoretical  $m/z$  852.41).

**Naphthyridine(N-alkyl)-(R/S)-Thiazole-L-His-tert-butylacetamide (C49b):**

General procedure B was employed for the deprotection of compound **C49a** (34.6 mg, 0.041 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in the presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C49b** was isolated as a yellow-pale powder (38.2 mg, 95%).

**HPLC:**  $rt_{dia1}$  = 8.29 min and  $rt_{dia2}$  = 8.48 min (method B); NMR data as a mixture of 2 diastereomers and conformers; **<sup>1</sup>H NMR** (400 MHz, D<sub>2</sub>O)  $\delta$  = 8.72, 8.68, 8.65 (3d,  $J$  = 1,3 Hz, 1H, H<sub>10</sub>), 7.73, 7.72 (2d,  $J$  = 8.7 Hz, 1H, H<sub>k</sub>), 7.67, 7.66 (2d,  $J$  = 9.3 Hz, 1H, H<sub>f</sub>), 7.55, 7.49, 7.34 (3d,  $J$  = 3.3 Hz, 1H, H<sub>3'</sub>), 7.42 (s, 1H, H<sub>9</sub>), 7.32, 7.19 (2d,  $J$  = 3.3 Hz, 1H, H<sub>2'</sub>), 6.58, 6.56 (2d,  $J$  = 8.5 Hz, 1H, H<sub>j</sub>), 6.34, 6.33 (2d,  $J$  = 9.3 Hz, 1H, H<sub>e</sub>), H<sub>6</sub> (signal in D<sub>2</sub>O peak), 4.28 – 4.14, 4.13 – 4.02, 3.97 – 3.86 (3m, 2H, H<sub>c</sub>), 3.53 – 2.88 (m, 4H, H<sub>a</sub> + H<sub>7</sub>), 1.99 – 1.65, 1.60 – 1.45, 1.38 – 1.06 (m, 2H, H<sub>b</sub>), 1.20, 1.19, 1.18, 1.17 (4s, 9H, H<sub>1</sub>), H<sub>4</sub> missing; **<sup>13</sup>C NMR** (101 MHz, D<sub>2</sub>O)  $\delta$  = 168.9, 168.6 (C<sub>5</sub>), 167.0, 166.7 (C<sub>3</sub>), 164.4 (C<sub>d</sub>), 162.1, 161.3 (C<sub>1'</sub>), 158.7 (C<sub>i</sub>), 146.7 (C<sub>h</sub>), 141.8, 141.7 (C<sub>3'</sub>), 140.1, 140.0, 140.0 (C<sub>f</sub> + C<sub>k</sub>), 134.5 (C<sub>10</sub>), 125.1, 125.0 (C<sub>8</sub>), 122.9, 122.8 (C<sub>2'</sub>), 119.0, 119.0 (C<sub>9</sub>), 114.2 (C<sub>e</sub>), 108.1, 108.8 (C<sub>g</sub>), 106.9, 106.9 (C<sub>j</sub>), 52.1, 52.0 (C<sub>2</sub>), 49.4, 49.3 (C<sub>6</sub>), 44.4, 43.9 (C<sub>a</sub>), 37.9, 37.8 (C<sub>c</sub>), 28.4, 28.1 (C<sub>b</sub>), 27.4, 27.3 (C<sub>1</sub>), 26.0, 25.9 (C<sub>7</sub>), C<sub>4</sub> missing; **HRMS (ESI):**  $m/z$  552.25061 [M+H]<sup>+</sup> (C<sub>26</sub>H<sub>34</sub>N<sub>9</sub>O<sub>3</sub>S requires 552.24998).

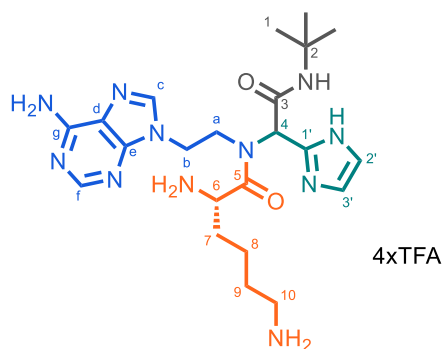
**(Boc)-Adenine-(R/S)-Imidazole-(Boc)-L-Lys(Boc)-tert-butylacetamide (C50a):**

Compound **C50a** was prepared following general procedure A using **Boc-adenine-NH<sub>2</sub> (21)** (27.8 mg, 0.10 mmol) as the amine, **2-imidazolecarboxaldehyde** (9.61 mg, 0.10 mmol) as aldehyde, **Boc-L-Lys(Boc)-OH** (34.6 mg, 0.10 mmol) as the acid and **tert-butyl isocyanide** (11.3  $\mu$ L, 0.10 mmol) in

MeOH/CH<sub>2</sub>Cl<sub>2</sub> (0.4 mL, 1/1, v/v). After purification by flash chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH gradient 100:0 to 95:5, v/v), a mixture of diastereomers **C50a** was obtained as a pale-yellow solid (42.7 mg, 54%).

$R_{f, \text{dia}1+2} = 0.27$  (CH<sub>2</sub>Cl<sub>2</sub>-MeOH 96:4); NMR data of one conformer of the mixture of diastereomers; <sup>1</sup>H NMR (400 MHz, CD<sub>3</sub>OD)  $\delta = 8.59, 8.57, 8.53, 8.50$  (4s, 1H, H<sub>f</sub>), 8.10, 8.07, 8.05 (3s, 1H, H<sub>c</sub>), 7.07, 7.03, 6.99 (3br s, 2H, H<sub>2'</sub> + H<sub>3'</sub>), 6.48 – 6.75 (br s, 1H, H<sub>4</sub>), 4.59 – 3.61 (m, 5H, H<sub>a-b</sub> + H<sub>6</sub>), 3.14 – 2.90 (m, 2H, H<sub>10</sub>), 1.86 – 1.80 (m, 42H, 6CH<sub>3, Boc</sub> + H<sub>1</sub> + H<sub>7-9</sub>); <sup>13</sup>C NMR (101 MHz, CD<sub>3</sub>OD)  $\delta = 176.2$  (C<sub>5</sub>), 168.1, 167.6, 167.4 (C<sub>3</sub>), 158.8, 158.5, 158.2, 157.9 (2CO<sub>Boc</sub>), 153.3, 153.2, 152.8, 152.7, 152.7, 152.3, 152.3, 151.2, 150.8 (CO<sub>Boc</sub> + C<sub>e</sub> + C<sub>f</sub> + C<sub>g</sub>), 145.4, 145.1 (C<sub>c</sub>), 143.7, 142.8 (C<sub>1'</sub>), 123.9, 123.0, 123.0, 122.9, 122.7 (C<sub>d</sub> + C<sub>2'</sub> + C<sub>3'</sub>), 82.7, 82.6, 81.2, 80.6, 80.4, 79.8 (3C<sub>IV, Boc</sub>), 54.8, 52.7, 52.7, 52.6, 52.5, 52.4 (C<sub>2</sub> + C<sub>6</sub>), 48.3, 46.1, 45.8, 44.0, 43.1, 42.9 (C<sub>a</sub> + C<sub>b</sub>), 40.9, 40.7 (C<sub>10</sub>), 32.7, 32.5, 32.2, 31.5, 30.6, 30.4 (C<sub>7</sub> + C<sub>9</sub>), 28.8, 28.8, 28.7, 28.5 (12CH<sub>3, Boc</sub> + C<sub>1</sub>), 24.0, 23.9, 23.8 (C<sub>8</sub>), C<sub>4</sub> missing; MS (ESI):  $m/z = 786.13$  [M+H]<sup>+</sup> (theoretical  $m/z$  786.46).

#### Adenine-(R/S)-Imidazole-L-Lys-*tert*-butylacetamide (**C50b**):



General procedure B was employed for the deprotection of compound **C50a** (30.8 mg, 0.039 mmol) in 1 mL of CH<sub>2</sub>Cl<sub>2</sub> and in presence of 50 equiv. of TFA. After filtration, a mixture of diastereomers **C50b** was isolated as a white powder (26.6 mg, 72%).

HPLC:  $rt_{\text{dia}1} = 7.48$  min and  $rt_{\text{dia}2} = 7.83$  min (method A); NMR data of one conformer of the mixture of diastereomers; <sup>1</sup>H NMR (400 MHz, D<sub>2</sub>O)  $\delta = 8.44, 8.42, 8.40, 8.39$  (4s, 1H, H<sub>f</sub>), 8.28, 8.27, 8.16 (3s, 1H, H<sub>c</sub>), 7.48, 7.43, 7.34 (3br s, 2H, H<sub>2'</sub> + H<sub>3'</sub>), 4.68 – 4.58 (m, 1H, H<sub>6</sub>), 4.58 – 4.20 (m, 2H, H<sub>b</sub>), 4.35 – 3.95 (m, 2H, H<sub>a</sub>), 3.07 – 2.87 (m, 2H, H<sub>10</sub>), 2.15 – 1.80 (m, 2H, H<sub>7</sub>), 1.79 – 1.36 (m, 4H, H<sub>8-9</sub>), 1.21, 1.19 (2s, 9H, H<sub>1</sub>), H<sub>4</sub> missing; <sup>13</sup>C NMR (101 MHz, D<sub>2</sub>O)  $\delta = 171.5, 170.9$  (C<sub>5</sub>), 163.4, 163.4 (C<sub>3</sub>), 149.9, 149.9, 148.6, 148.6 (C<sub>e</sub> + C<sub>g</sub>), 144.8, 144.7, 144.5, 144.3 (C<sub>c</sub> + C<sub>f</sub>), 138.8, 138.1 (C<sub>1'</sub>), 121.1, 120.5, 120.0 (C<sub>2'</sub> + C<sub>3'</sub>), 118.0 (C<sub>d</sub>), 52.7, 51.1, 51.0 (C<sub>2</sub> + C<sub>6</sub>), 49.3, 46.5, 43.1, 42.9 (C<sub>a</sub> + C<sub>b</sub>), 40.0, 38.9 (C<sub>10</sub>), 29.9, 29.9 (C<sub>7</sub>), 27.5, 27.4 (C<sub>1</sub>), 26.4, 26.3 (C<sub>9</sub>), 21.4, 21.0 (C<sub>8</sub>), C<sub>4</sub> missing; MS (ESI):  $m/z = 486.20$  [M+H]<sup>+</sup> (theoretical  $m/z$  486.30).

## B) Biochemistry

### Materials

All buffers and solutions employed in fluorescence-based assays were filtered through 0.22 µm Millipore filters (GP ExpressPLUS® membrane). Tris-HCl 20 mM, pH 7.2 containing 12 mM NaCl, 3 mM MgCl<sub>2</sub> and 1 mM DTT (Buffer A) was used for K<sub>D</sub> determination. HEPES 20 mM, pH 7.4, 20 mM NaCl, 140 mM KCl, et 3 mM MgCl<sub>2</sub> (Buffer B) was used for inhibition assays.

RNA and DNA oligonucleotides were purchased from Eurogentec (Liege, Belgium). For competition experiments in the presence of tRNA, a mixture of pre- and mature yeast tRNAs (from *Escherichia coli*, Sigma) was added to buffer A to obtain a 100-fold nucleotides excess regarding RNA *AapA1*. For competition experiments in the presence of a dsDNA, a 15-mer sequence (5'-CGTTTTTATTTTGC-3') and its complementary oligonucleotide, annealed beforehand, were added to buffer A to obtain a 100-fold nucleotide excess regarding RNA *AapA1*.

**Oligonucleotides are as follows:** *AapA1* (RNA messenger) and *IsoA1* (RNA antitoxin)

For K<sub>D</sub> determination: *AapA1*-FAM: 5'-FAM-UUU-GUC-AUG-GCU-ACC-AAA-CAU-GGC-AAA-3'

For inhibition assays:

*AapA1*-FAM: 5'-FAM-UUU-GUC-AUG-GCU-ACC-AAA-CAU-GGC-AAA-3'

*IsoA1*: 5'-UUU-GCC-AUG-UUU-GGU-AGC-CAU-GAC-AAA-3'

*IsoA1*-EDQ: 5'-UUU-GCC-AUG-UUU-GGU-AGC-CAU-GAC-AAA-EDQ-3'

### Procedure of Fluorescence-based assays

K<sub>D</sub> determinations were performed in 384-well black plates (Greiner bio-one) in a final volume of 60 µL, respectively, using a 5070 EpMotion automated pipetting system (Eppendorf). Each experiment was performed in duplicate and repeated three times. Refolding of the RNA was performed using a thermocycler (ThermoStatPlus Eppendorf) as follows: the RNA (5'-FAM-*AapA1*), was diluted in 1 mL of buffer A, was first denatured by heating to 90°C for 2 min and then cooled to 4°C for 10 min, followed by incubation at 25°C for 15 min. After refolding, the RNA was diluted to a working concentration of 10 nM through addition of the appropriate amount of buffer A.

During binding assays, 30 µL of RNA solution were added in each well containing 30 µL of each desired ligand concentration (final volume 60 µL). Each ligand was added in 15 dilutions (from 61 µM to 1 mM) and fluorescence was measured after incubating the plate at 4°C overnight. The fluorescence was measured at various temperature from 5°C to 35°C on a GeniosPro (Tecan).



### Procedure of Anisotropy assay:

Inhibition assays were performed in 384-well black plates (Greiner bio-one) in a final volume of 70  $\mu$ L using a 5070 EpMotion automated pipetting system (Eppendorf). Each experiment was performed in duplicate and repeated three times. Refolding of the RNA was performed using a thermocycler (ThermoStatPlus Eppendorf) as follows: the RNA 5'-FAM-*AapA1* or RNA *IsoA1*, was diluted in 1 mL of buffer B, was first denatured by heating to 90°C for 2 min and then cooled to 4°C for 10 min, followed by incubation at 25°C for 15 min. After refolding, the RNA was diluted to a working concentration of 11.7 nM for *AapA1* and 35 nM for *IsoA1* through addition of the appropriate amount of buffer B.

During inhibitions assays, 30  $\mu$ L of RNA *AapA1* solution was added in each well containing 30  $\mu$ L of each desired ligand concentration. Each ligand was added in 15 dilutions (from 61 pM to 1 mM). After incubating the plate at 4°C overnight, 10  $\mu$ L of RNA *IsoA1* solution was added in each well (final volume 70  $\mu$ L). The plate was incubated at 37°C for 2h then at 4°C for 20 min. The anisotropy was measured on a SpectraMax (Molecular Devices).

### Data analysis

Inhibition data were analyzed using GraphPad Prism 5 software using a nonlinear regression following the equation:

$$Y = Bottom + \frac{(Top - Bottom)}{(1 + 10^{(\log KD - X) \times hills\ slope})}$$

With bottom = Minimal value for Y; Top = Maximal value for Y;  $K_D = IC_{50}$  for the calculation of  $IC_{50}$ .

Binding profiles were well modeled using a simple model assuming a one-to-one stoichiometry.

### Procedure of electrophoretic mobility shift assay:

The RNA (0.2 nmol /100  $\mu$ L) 5'-FAM-*AapA*, *IsoA1* (or 3'-*IsoA1*-EDQ) and 5'-FAM-*AapA/IsoA1* (5'-FAM-*AapA/3'-IsoA1*-EDQ) were incubated separately for 2 min at 90°C, 10 min at 4°C and 15 min at 25°C in buffer B. Once mixed, they were incubated for 2h at 37°C, 25°C or 4°C adding RNA 5'-FAM-*AapA1* (10 $\mu$ L) and *IsoA1* (or 3'-*IsoA1*-EDQ) (10  $\mu$ L).

Conditions: 20% solution of acrylamide/bis-acrylamide 19:1 in TBE buffer 1x. Migration was performed at 4°C during 4h (9 mA).

The samples (RNA 20 pmol, V= 30 $\mu$ L) were loaded after addition of a sucrose solution (50%). Gels were scanned with a Versadoc (Biorad) at 480 nm.

**Procedure of GFP Reporter Assay:**

Calcium competent bacteria *E. coli* MG1655 are transformed by heat-shock at 42°C for 45 sec, with plasmids pBAD-5'UTR:AapA1(WT)/EGFP-IsoA1 or pBAD-5'UTR:AapA1(C49Amut)/EGFP-IsoA1. Expression of 5'UTR:AapA1/EGFP mRNA and IsoA1 sncRNA are independently induced by addition of arabinose and IPTG, respectively. For each plasmid, a bacterial culture (5 ml) is prepared in LB medium à 37°C, overnight, with an orbital shaking of 180 rpm. The culture is then diluted to an OD600 of 0.1, and split into samples (1 ml) to which are added arabinose (0.005%, W/V), or IPTG (1 mM), or both arabinose and IPTG. The molecule to be tested is next added (25 µL of a stock solution at 2 mM, 50 µM final concentration) to the 1 mL samples, which are dispensed (200 µL per well) in triplicate in a 96-well microplate. The microplate is placed in a FLUOstar Omega (BMG LABTECH) microplate reader prewarmed at 37°C, and both the OD600 and the fluorescence (Excitation wavelength: 485 nm; Emission wavelength: 520 nm) are recorded every 10 min over 10 hours with an orbital shaking of 600 rpm between recordings.



---

# Références bibliographiques

---



- (1) Home | AMR Review <https://amr-review.org/> (accessed 2021 -06 -02).
- (2) Nauciel, C.; Vildé, J.-L. *Bactériologie Médicale*, Masson.; 2005.
- (3) Morlot, C. Etude structurale de PBP3 et localisation des six “Penicillin-Binding Proteins” de *Streptococcus pneumoniae*: implication dans la croissance et la division bactérienne. Thèse doctoral, Université Joseph Fourier-Grenoble I, Grenoble, 2003.
- (4) Wilson, D. N. The A–Z of Bacterial Translation Inhibitors. *Crit. Rev. Biochem. Mol. Biol.* **2009**, *44* (6), 393–433. <https://doi.org/10.3109/10409230903307311>.
- (5) Wilson, D. N. Ribosome-Targeting Antibiotics and Mechanisms of Bacterial Resistance. *Nat. Rev. Microbiol.* **2014**, *12* (1), 35–48. <https://doi.org/10.1038/nrmicro3155>.
- (6) Dix ennemis que l’OMS devra affronter cette année <https://www.who.int/fr/news-room/spotlight/ten-threats-to-global-health-in-2019> (accessed 2021 -04 -08).
- (7) Tacconelli, E.; Carrara, E.; Savoldi, A.; Harbarth, S.; Mendelson, M.; Monnet, D. L.; Pulcini, C.; Kahlmeter, G.; Kluytmans, J.; Carmeli, Y.; Ouellette, M.; Outtersson, K.; Patel, J.; Cavalieri, M.; Cox, E. M.; Houchens, C. R.; Grayson, M. L.; Hansen, P.; Singh, N.; Theuretzbacher, U.; Magrini, N.; Aboderin, A. O.; Al-Abri, S. S.; Jalil, N. A.; Benzonana, N.; Bhattacharya, S.; Brink, A. J.; Burkert, F. R.; Cars, O.; Cornaglia, G.; Dyar, O. J.; Friedrich, A. W.; Gales, A. C.; Gandra, S.; Giske, C. G.; Goff, D. A.; Goossens, H.; Gottlieb, T.; Blanco, M. G.; Hryniewicz, W.; Kattula, D.; Jinks, T.; Kanj, S. S.; Kerr, L.; Kieny, M.-P.; Kim, Y. S.; Kozlov, R. S.; Labarca, J.; Laxminarayan, R.; Leder, K.; Leibovici, L.; Levy-Hara, G.; Littman, J.; Malhotra-Kumar, S.; Manchanda, V.; Moja, L.; Ndoye, B.; Pan, A.; Paterson, D. L.; Paul, M.; Qiu, H.; Ramon-Pardo, P.; Rodríguez-Baño, J.; Sanguinetti, M.; Sengupta, S.; Sharland, M.; Si-Mehand, M.; Silver, L. L.; Song, W.; Steinbakk, M.; Thomsen, J.; Thwaites, G. E.; Meer, J. W. van der; Kinh, N. V.; Vega, S.; Villegas, M. V.; Wechsler-Fördös, A.; Wertheim, H. F. L.; Wesangula, E.; Woodford, N.; Yilmaz, F. O.; Zorzet, A. Discovery, Research, and Development of New Antibiotics: The WHO Priority List of Antibiotic-Resistant Bacteria and Tuberculosis. *Lancet Infect. Dis.* **2018**, *18* (3), 318–327. [https://doi.org/10.1016/S1473-3099\(17\)30753-3](https://doi.org/10.1016/S1473-3099(17)30753-3).
- (8) Data from the ECDC Surveillance Atlas - Antimicrobial resistance <https://www.ecdc.europa.eu/en/antimicrobial-resistance/surveillance-and-disease-data/data-ecdc> (accessed 2021 -03 -23).
- (9) WHO Antibiotics Portal <https://aware.essentialmeds.org/list> (accessed 2021 -03 -23).
- (10) Bentley, R. Different Roads to Discovery; Prontosil (Hence Sulfa Drugs) and Penicillin (Hence  $\beta$ -Lactams). *J. Ind. Microbiol. Biotechnol.* **2009**, *36* (6), 775–786. <https://doi.org/10.1007/s10295-009-0553-8>.
- (11) Zaffiri, L.; Gardner, J.; Toledo-Pereyra, L. H. History of Antibiotics. From Salvarsan to Cephalosporins. *J. Invest. Surg.* **2012**, *25* (2), 67–77. <https://doi.org/10.3109/08941939.2012.664099>.
- (12) Wright, G. D. Q&A: Antibiotic Resistance: Where Does It Come from and What Can We Do about It? *BMC Biol.* **2010**, *8* (1), 123. <https://doi.org/10.1186/1741-7007-8-123>.
- (13) Wang, D. Y.; Abboud, M. I.; Markoulides, M. S.; Brem, J.; Schofield, C. J. The Road to Avibactam: The First Clinically Useful Non- $\beta$ -Lactam Working Somewhat like a  $\beta$ -Lactam. *Future Med. Chem.* **2016**, *8* (10), 1063–1084. <https://doi.org/10.4155/fmc-2016-0078>.

- (14) Lima, L. M.; Silva, B. N. M. da; Barbosa, G.; Barreiro, E. J.  $\beta$ -Lactam Antibiotics: An Overview from a Medicinal Chemistry Perspective. *Eur. J. Med. Chem.* **2020**, *208*, 112829. <https://doi.org/10.1016/j.ejmech.2020.112829>.
- (15) Zeng, D.; Debabov, D.; Hartsell, T. L.; Cano, R. J.; Adams, S.; Schuyler, J. A.; McMillan, R.; Pace, J. L. Approved Glycopeptide Antibacterial Drugs: Mechanism of Action and Resistance. *Cold Spring Harb. Perspect. Med.* **2016**, *6* (12). <https://doi.org/10.1101/cshperspect.a026989>.
- (16) Pham, T. D. M.; Ziora, Z. M.; Blaskovich, M. A. T. Quinolone Antibiotics. *Med Chem Comm* **2019**, *10* (10), 1719–1739. <https://doi.org/10.1039/C9MD00120D>.
- (17) Capasso, C.; Supuran, C. T. Sulfa and Trimethoprim-like Drugs – Antimetabolites Acting as Carbonic Anhydrase, Dihydropteroate Synthase and Dihydrofolate Reductase Inhibitors. *J. Enzyme Inhib. Med. Chem.* **2014**, *29* (3), 379–387. <https://doi.org/10.3109/14756366.2013.787422>.
- (18) Velkov, T.; Roberts, K. D. Discovery of Novel Polymyxin-Like Antibiotics. In *Polymyxin Antibiotics: From Laboratory Bench to Bedside*; Li, J., Nation, R. L., Kaye, K. S., Eds.; Advances in Experimental Medicine and Biology; Springer International Publishing: Cham, 2019; Vol. 1145, pp 343–362. [https://doi.org/10.1007/978-3-030-16373-0\\_20](https://doi.org/10.1007/978-3-030-16373-0_20).
- (19) Gray, D. A.; Wenzel, M. More Than a Pore: A Current Perspective on the In Vivo Mode of Action of the Lipopeptide Antibiotic Daptomycin. *Antibiotics* **2020**, *9* (1). <https://doi.org/10.3390/antibiotics9010017>.
- (20) Lin, J.; Zhou, D.; Steitz, T. A.; Polikanov, Y. S.; Gagnon, M. G. Ribosome-Targeting Antibiotics: Modes of Action, Mechanisms of Resistance, and Implications for Drug Design. *Annu. Rev. Biochem.* **2018**, *87* (1), 451–478. <https://doi.org/10.1146/annurev-biochem-062917-011942>.
- (21) Vázquez-Laslop, N.; Mankin, A. S. How Macrolide Antibiotics Work. *Trends Biochem. Sci.* **2018**, *43* (9), 668–684. <https://doi.org/10.1016/j.tibs.2018.06.011>.
- (22) Chopra, I.; Roberts, M. Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance. *Microbiol. Mol. Biol. Rev.* **2001**, *65* (2), 232–260. <https://doi.org/10.1128/MMBR.65.2.232-260.2001>.
- (23) Zhanel, G. G.; Esquivel, J.; Zelenitsky, S.; Lawrence, C. K.; Adam, H. J.; Golden, A.; Hink, R.; Berry, L.; Schweizer, F.; Zhanel, M. A.; Bay, D.; Lagacé-Wiens, P. R. S.; Walkty, A. J.; Lynch, J. P.; Karlowsky, J. A. Omadacycline: A Novel Oral and Intravenous Aminomethylcycline Antibiotic Agent. *Drugs* **2020**, *80* (3), 285–313. <https://doi.org/10.1007/s40265-020-01257-4>.
- (24) Garneau-Tsodikova, S.; Labby, K. J. Mechanisms of Resistance to Aminoglycoside Antibiotics: Overview and Perspectives. *MedChemComm* **2016**, *7* (1), 11–27. <https://doi.org/10.1039/C5MD00344J>.
- (25) Takahashi, Y.; Igarashi, M. Destination of Aminoglycoside Antibiotics in the ‘Post-Antibiotic Era.’ *J. Antibiot. (Tokyo)* **2018**, *71* (1), 4–14. <https://doi.org/10.1038/ja.2017.117>.
- (26) Koulenti; Xu; Mok; Song; Karageorgopoulos; Armaganidis; Lipman; SotiriosTsiodras. Novel Antibiotics for Multidrug-Resistant Gram-Positive Microorganisms. *Microorganisms* **2019**, *7* (8), 270. <https://doi.org/10.3390/microorganisms7080270>.
- (27) Koulenti, D.; Song, A.; Ellingboe, A.; Abdul-Aziz, M. H.; Harris, P.; Gavey, E.; Lipman, J. Infections by Multidrug-Resistant Gram-Negative Bacteria: What’s New in Our Arsenal and What’s in the Pipeline? *Int. J. Antimicrob. Agents* **2019**, *53* (3), 211–224. <https://doi.org/10.1016/j.ijantimicag.2018.10.011>.

- (28) Brauner, A.; Fridman, O.; Gefen, O.; Balaban, N. Q. Distinguishing between Resistance, Tolerance and Persistence to Antibiotic Treatment. *Nat. Rev. Microbiol.* **2016**, *14* (5), 320–330. <https://doi.org/10.1038/nrmicro.2016.34>.
- (29) Blair, J. M. A.; Webber, M. A.; Baylay, A. J.; Ogbolu, D. O.; Piddock, L. J. V. Molecular Mechanisms of Antibiotic Resistance. *Nat. Rev. Microbiol.* **2015**, *13* (1), 42–51. <https://doi.org/10.1038/nrmicro3380>.
- (30) C Reygaert, W.; Department of Biomedical Sciences, Oakland University William Beaumont School of Medicine, Rochester, MI, USA. An Overview of the Antimicrobial Resistance Mechanisms of Bacteria. *AIMS Microbiol.* **2018**, *4* (3), 482–501. <https://doi.org/10.3934/microbiol.2018.3.482>.
- (31) Marshall, R. L.; Bavro, V. N. Multidrug Resistance. In *Bacterial Resistance to Antibiotics – From Molecules to Man*; Bonev, B. B., Brown, N. M., Eds.; Wiley, 2019; pp 201–237. <https://doi.org/10.1002/9781119593522.ch9>.
- (32) Sharkey, L. K. R.; O’Neill, A. J. Molecular Mechanisms of Antibiotic Resistance – Part II. In *Bacterial Resistance to Antibiotics – From Molecules to Man*; Bonev, B. B., Brown, N. M., Eds.; Wiley, 2019; pp 27–50. <https://doi.org/10.1002/9781119593522.ch2>.
- (33) Green, A. T.; Moniruzzaman, M.; Cooper, C. J.; Walker, J. K.; Smith, J. C.; Parks, J. M.; Zgurskaya, H. I. Discovery of Multidrug Efflux Pump Inhibitors with a Novel Chemical Scaffold. *Biochim. Biophys. Acta BBA - Gen. Subj.* **2020**, *1864* (6), 129546. <https://doi.org/10.1016/j.bbagen.2020.129546>.
- (34) Gupta, S.; Cohen, K. A.; Winglee, K.; Maiga, M.; Diarra, B.; Bishai, W. R. Efflux Inhibition with Verapamil Potentiates Bedaquiline in Mycobacterium Tuberculosis. *Antimicrob. Agents Chemother.* **2014**, *58* (1), 574–576. <https://doi.org/10.1128/AAC.01462-13>.
- (35) Peilleron, L.; Cariou, K. Synthetic Approaches towards Avibactam and Other Diazabicyclooctane  $\beta$ -Lactamase Inhibitors. *Org. Biomol. Chem.* **2020**, *18* (5), 830–844. <https://doi.org/10.1039/C9OB02605C>.
- (36) Balaban, N. Q.; Helaine, S.; Lewis, K.; Ackermann, M.; Aldridge, B.; Andersson, D. I.; Brynildsen, M. P.; Bumann, D.; Camilli, A.; Collins, J. J.; Dehio, C.; Fortune, S.; Ghigo, J.-M.; Hardt, W.-D.; Harms, A.; Heinemann, M.; Hung, D. T.; Jenal, U.; Levin, B. R.; Michiels, J.; Storz, G.; Tan, M.-W.; Tenson, T.; Van Melderen, L.; Zinkernagel, A. Definitions and Guidelines for Research on Antibiotic Persistence. *Nat. Rev. Microbiol.* **2019**, *17* (7), 441–448. <https://doi.org/10.1038/s41579-019-0196-3>.
- (37) Hall, A. M.; Gollan, B.; Helaine, S. Toxin–Antitoxin Systems: Reversible Toxicity. *Curr. Opin. Microbiol.* **2017**, *36*, 102–110. <https://doi.org/10.1016/j.mib.2017.02.003>.
- (38) Ogura, T.; Hiraga, S. Mini-F Plasmid Genes That Couple Host Cell Division to Plasmid Proliferation. *Proc. Natl. Acad. Sci.* **1983**, *80* (15), 4784–4788. <https://doi.org/10.1073/pnas.80.15.4784>.
- (39) Harms, A.; Brodersen, D. E.; Mitarai, N.; Gerdes, K. Toxins, Targets, and Triggers: An Overview of Toxin-Antitoxin Biology. *Mol. Cell* **2018**, *70* (5), 768–784. <https://doi.org/10.1016/j.molcel.2018.01.003>.
- (40) Srivastava, A.; Pati, S.; Kaushik, H.; Singh, S.; Garg, L. C. Toxin-Antitoxin Systems and Their Medical Applications: Current Status and Future Perspective. *Appl. Microbiol. Biotechnol.* **2021**, *105* (5), 1803–1821. <https://doi.org/10.1007/s00253-021-11134-z>.



- (41) Riffaud, C.; Pinel-Marie, M.-L.; Felden, B. Cross-Regulations between Bacterial Toxin–Antitoxin Systems: Evidence of an Interconnected Regulatory Network? *Trends Microbiol.* **2020**, *28* (10), 851–866. <https://doi.org/10.1016/j.tim.2020.05.016>.
- (42) Fraikin, N.; Goormaghtigh, F.; Van Melderen, L. Type II Toxin-Antitoxin Systems: Evolution and Revolutions. *J. Bacteriol.* **2020**, *202* (7), e00763-19, /jb/202/7/JB.00763-19.atom. <https://doi.org/10.1128/JB.00763-19>.
- (43) Gerdes, K.; Rasmussen, P. B.; Molin, S. Unique Type of Plasmid Maintenance Function: Postsegregational Killing of Plasmid-Free Cells. *Proc. Natl. Acad. Sci.* **1986**, *83* (10), 3116–3120. <https://doi.org/10.1073/pnas.83.10.3116>.
- (44) Harms, A.; Maisonneuve, E.; Gerdes, K. Mechanisms of Bacterial Persistence during Stress and Antibiotic Exposure. *Science* **2016**, *354* (6318), aaf4268. <https://doi.org/10.1126/science.aaf4268>.
- (45) Gerdes, K.; Maisonneuve, E. Remarkable Functional Convergence: Alarmone PpGpp Mediates Persistence by Activating Type I and II Toxin-Antitoxins. *Mol. Cell* **2015**, *59* (1), 1–3. <https://doi.org/10.1016/j.molcel.2015.06.019>.
- (46) Page, R.; Peti, W. Toxin-Antitoxin Systems in Bacterial Growth Arrest and Persistence. *Nat. Chem. Biol.* **2016**, *12* (4), 208–214. <https://doi.org/10.1038/nchembio.2044>.
- (47) Chan, W. T.; Balsa, D.; Espinosa, M. One Cannot Rule Them All: Are Bacterial Toxins-Antitoxins Druggable? *FEMS Microbiol. Rev.* **2015**, *39* (4), 522–540. <https://doi.org/10.1093/femsre/fuv002>.
- (48) Równicki, M.; Lasek, R.; Trylska, J.; Bartosik, D. Targeting Type II Toxin–Antitoxin Systems as Antibacterial Strategies. *Toxins* **2020**, *12* (9), 568. <https://doi.org/10.3390/toxins12090568>.
- (49) Solecki, O.; Mosbah, A.; Baudy Floc’h, M.; Felden, B. Converting a Staphylococcus Aureus Toxin into Effective Cyclic Pseudopeptide Antibiotics. *Chem. Biol.* **2015**, *22* (3), 329–335. <https://doi.org/10.1016/j.chembiol.2014.12.016>.
- (50) Nicolas, I.; Bordeau, V.; Bondon, A.; Baudy-Floc’h, M.; Felden, B. Novel Antibiotics Effective against Gram-Positive and -Negative Multi-Resistant Bacteria with Limited Resistance. *PLOS Biol.* **2019**, *17* (7), e3000337. <https://doi.org/10.1371/journal.pbio.3000337>.
- (51) Williams, J. J.; Hergenrother, P. J. Artificial Activation of Toxin–Antitoxin Systems as an Antibacterial Strategy. *Trends Microbiol.* **2012**, *20* (6), 291–298. <https://doi.org/10.1016/j.tim.2012.02.005>.
- (52) Brielle, R.; Pinel-Marie, M.-L.; Felden, B. Linking Bacterial Type I Toxins with Their Actions. *Curr. Opin. Microbiol.* **2016**, *30*, 114–121. <https://doi.org/10.1016/j.mib.2016.01.009>.
- (53) Wen, J.; Fozo, E. SRNA Antitoxins: More than One Way to Repress a Toxin. *Toxins* **2014**, *6* (8), 2310–2335. <https://doi.org/10.3390/toxins6082310>.
- (54) Masachis, S.; Tourasse, N. J.; Lays, C.; Faucher, M.; Chabas, S.; Iost, I.; Darfeuille, F. A Genetic Selection Reveals Functional Metastable Structures Embedded in a Toxin-Encoding MRNA. *eLife* **2019**, *8*, e47549. <https://doi.org/10.7554/eLife.47549>.
- (55) Berghoff, B. A.; Wagner, E. G. H. RNA-Based Regulation in Type I Toxin–Antitoxin Systems and Its Implication for Bacterial Persistence. *Curr. Genet.* **2017**, *63* (6), 1011–1016. <https://doi.org/10.1007/s00294-017-0710-y>.
- (56) Cover, T. L.; Blaser, M. J. Helicobacter Pylori in Health and Disease. *Gastroenterology* **2009**, *136* (6), 1863–1873. <https://doi.org/10.1053/j.gastro.2009.01.073>.

- (57) Piscione, M.; Mazzone, M.; Di Marcantonio, M. C.; Muraro, R.; Mincione, G. Eradication of *Helicobacter Pylori* and Gastric Cancer: A Controversial Relationship. *Front. Microbiol.* **2021**, *12*, 630852. <https://doi.org/10.3389/fmicb.2021.630852>.
- (58) Cancer <https://www.who.int/fr/news-room/fact-sheets/detail/cancer> (accessed 2021 -05 -11).
- (59) Arnion, H.; Korkut, D. N.; Masachis Gelo, S.; Chabas, S.; Reignier, J.; Iost, I.; Darfeuille, F. Mechanistic Insights into Type I Toxin Antitoxin Systems in *Helicobacter Pylori*: The Importance of mRNA Folding in Controlling Toxin Expression. *Nucleic Acids Res.* **2017**, gkw1343. <https://doi.org/10.1093/nar/gkw1343>.
- (60) Korkut, D. N.; Alves, I. D.; Vogel, A.; Chabas, S.; Sharma, C. M.; Martinez, D.; Loquet, A.; Salgado, G. F.; Darfeuille, F. Structural Insights into the AapA1 Toxin of *Helicobacter Pylori*. *Biochim. Biophys. Acta BBA - Gen. Subj.* **2020**, *1864* (1), 129423. <https://doi.org/10.1016/j.bbagen.2019.129423>.
- (61) El Mortaji, L.; Tejada-Arranz, A.; Rifflet, A.; Boneca, I. G.; Pehau-Arnaudet, G.; Radicella, J. P.; Marsin, S.; De Reuse, H. A Peptide of a Type I Toxin–antitoxin System Induces *Helicobacter Pylori* Morphological Transformation from Spiral Shape to Coccoids. *Proc. Natl. Acad. Sci.* **2020**, *117* (49), 31398–31409. <https://doi.org/10.1073/pnas.2016195117>.
- (62) Gelo, S. M. When mRNA Folding Rules Gene Expression : Lessons from Type I Toxin-Antitoxin Systems. Thèse doctoral, Université de Bordeaux, Bordeaux, 2018.
- (63) Masachis, S.; Darfeuille, F. Type I Toxin-Antitoxin Systems: Regulating Toxin Expression via Shine-Dalgarno Sequence Sequestration and Small RNA Binding. **2018**, 18.
- (64) Masachis, S.; Tourasse, N. J.; Lays, C.; Faucher, M.; Chabas, S.; Iost, I.; Darfeuille, F. A Genetic Selection Reveals Functional Metastable Structures Embedded in a Toxin-Encoding mRNA. 35.
- (65) Warner, K. D.; Hajdin, C. E.; Weeks, K. M. Principles for Targeting RNA with Drug-like Small Molecules. *Nat. Rev. Drug Discov.* **2018**, *17* (8), 547–558. <https://doi.org/10.1038/nrd.2018.93>.
- (66) Falese, J. P.; Donlic, A.; Hargrove, A. E. Targeting RNA with Small Molecules: From Fundamental Principles towards the Clinic. *Chem. Soc. Rev.* **2021**, *50* (4), 2224–2243. <https://doi.org/10.1039/D0CS01261K>.
- (67) Commissioner, O. of the. FDA Approves Oral Treatment for Spinal Muscular Atrophy <https://www.fda.gov/news-events/press-announcements/fda-approves-oral-treatment-spinal-muscular-atrophy> (accessed 2021 -03 -27).
- (68) Angelbello, A. J.; Chen, J. L.; Childs-Disney, J. L.; Zhang, P.; Wang, Z.-F.; Disney, M. D. Using Genome Sequence to Enable the Design of Medicines and Chemical Probes. *Chem. Rev.* **2018**, *118* (4), 1599–1663. <https://doi.org/10.1021/acs.chemrev.7b00504>.
- (69) Yu, A.-M.; Choi, Y. H.; Tu, M.-J. RNA Drugs and RNA Targets for Small Molecules: Principles, Progress, and Challenges. *Pharmacol. Rev.* **2020**, *72* (4), 862–898. <https://doi.org/10.1124/pr.120.019554>.
- (70) Morgan, B. S.; Forte, J. E.; Hargrove, A. E. Insights into the Development of Chemical Probes for RNA. *Nucleic Acids Res.* **2018**, *46* (16), 8025–8037. <https://doi.org/10.1093/nar/gky718>.
- (71) Haniff, H. S.; Knerr, L.; Chen, J. L.; Disney, M. D.; Lightfoot, H. L. Target-Directed Approaches for Screening Small Molecules against RNA Targets. *SLAS Discov. Adv. Sci. Drug Discov.* **2020**, *25* (8), 869–894. <https://doi.org/10.1177/2472555220922802>.

- (72) Wicks, S. L.; Hargrove, A. E. Fluorescent Indicator Displacement Assays to Identify and Characterize Small Molecule Interactions with RNA. *Methods* **2019**, *167*, 3–14. <https://doi.org/10.1016/j.ymeth.2019.04.018>.
- (73) Rizvi, N. F.; Santa Maria, J. P.; Nahvi, A.; Klappenbach, J.; Klein, D. J.; Curran, P. J.; Richards, M. P.; Chamberlin, C.; Saradjian, P.; Burchard, J.; Aguilar, R.; Lee, J. T.; Dandliker, P. J.; Smith, G. F.; Kutchukian, P.; Nickbarg, E. B. Targeting RNA with Small Molecules: Identification of Selective, RNA-Binding Small Molecules Occupying Drug-Like Chemical Space. *SLAS Discov. Adv. Sci. Drug Discov.* **2020**, *25* (4), 384–396. <https://doi.org/10.1177/2472555219885373>.
- (74) Eubanks, C. S.; Forte, J. E.; Kapral, G. J.; Hargrove, A. E. Small Molecule-Based Pattern Recognition To Classify RNA Structure. *J. Am. Chem. Soc.* **2017**, *139* (1), 409–416. <https://doi.org/10.1021/jacs.6b11087>.
- (75) Eubanks, C. S.; Zhao, B.; Patwardhan, N. N.; Thompson, R. D.; Zhang, Q.; Hargrove, A. E. Visualizing RNA Conformational Changes via Pattern Recognition of RNA by Small Molecules. *J. Am. Chem. Soc.* **2019**, *141* (14), 5692–5698. <https://doi.org/10.1021/jacs.8b09665>.
- (76) Padroni, G.; Patwardhan, N. N.; Schapira, M.; Hargrove, A. E. Systematic Analysis of the Interactions Driving Small Molecule–RNA Recognition. *RSC Med. Chem.* **2020**, *11* (7), 802–813. <https://doi.org/10.1039/D0MD00167H>.
- (77) Morgan, B. S.; Forte, J. E.; Culver, R. N.; Zhang, Y.; Hargrove, A. E. Discovery of Key Physicochemical, Structural, and Spatial Properties of RNA-Targeted Bioactive Ligands. *Angew. Chem. Int. Ed.* **2017**, *56* (43), 13498–13502. <https://doi.org/10.1002/anie.201707641>.
- (78) Disney, M. D.; Labuda, L. P.; Paul, D. J.; Poplawski, S. G.; Pushechnikov, A.; Tran, T.; Velagapudi, S. P.; Wu, M.; Childs-Disney, J. L. Two-Dimensional Combinatorial Screening Identifies Specific Aminoglycoside–RNA Internal Loop Partners. *J. Am. Chem. Soc.* **2008**, *130* (33), 11185–11194. <https://doi.org/10.1021/ja803234t>.
- (79) Velagapudi, S. P.; Seedhouse, S. J.; Disney, M. D. Structure-Activity Relationships through Sequencing (StARTS) Defines Optimal and Suboptimal RNA Motif Targets for Small Molecules. *Angew. Chem. Int. Ed.* **2010**, *49* (22), 3816–3818. <https://doi.org/10.1002/anie.200907257>.
- (80) Disney, M. D.; Winkelsas, A. M.; Velagapudi, S. P.; Southern, M.; Fallahi, M.; Childs-Disney, J. L. Inforna 2.0: A Platform for the Sequence-Based Design of Small Molecules Targeting Structured RNAs. *ACS Chem. Biol.* **2016**, *11* (6), 1720–1728. <https://doi.org/10.1021/acscchembio.6b00001>.
- (81) Umuhire Juru, A.; Hargrove, A. E. Frameworks for Targeting RNA with Small Molecules. *J. Biol. Chem.* **2021**, *296*, 100191. <https://doi.org/10.1074/jbc.REV120.015203>.
- (82) Disney, M. D. Targeting RNA with Small Molecules To Capture Opportunities at the Intersection of Chemistry, Biology, and Medicine. *J Am Chem Soc* **2019**, *15*.
- (83) Luu, L. M.; Nguyen, L.; Peng, S.; Lee, J.; Lee, H. Y.; Wong, C.-H.; Hergenrother, P. J.; Chan, H. Y. E.; Zimmerman, S. C. A Potent Inhibitor of Protein Sequestration by Expanded Triplet (CUG) Repeats That Shows Phenotypic Improvements in a Drosophila Model of Myotonic Dystrophy. **2016**, *8*.
- (84) The Nobel Prize in Chemistry 2009 <https://www.nobelprize.org/prizes/chemistry/2009/summary/> (accessed 2021 -04 -10).
- (85) Fourmy, D.; Recht, M. I.; Puglisi, J. D. Binding of Neomycin-Class Aminoglycoside Antibiotics to the A-Site of 16 s rRNA. *J. Mol. Biol.* **1998**, *277* (2), 347–362. <https://doi.org/10.1006/jmbi.1997.1552>.

- (86) Blount, K. F.; Zhao, F.; Hermann, T.; Tor, Y. Conformational Constraint as a Means for Understanding RNA-Aminoglycoside Specificity. *J. Am. Chem. Soc.* **2005**, *127* (27), 9818–9829. <https://doi.org/10.1021/ja050918w>.
- (87) Zhao, F.; Zhao, Q.; Blount, K. F.; Han, Q.; Tor, Y.; Hermann, T. Molecular Recognition of RNA by Neomycin and a Restricted Neomycin Derivative. *Angew. Chem. Int. Ed.* **2005**, *44* (33), 5329–5334. <https://doi.org/10.1002/anie.200500903>.
- (88) Watkins, D.; Norris, F. A.; Kumar, S.; Arya, D. P. A Fluorescence-Based Screen for Ribosome Binding Antibiotics. *Anal. Biochem.* **2013**, *434* (2), 300–307. <https://doi.org/10.1016/j.ab.2012.12.003>.
- (89) O’Sullivan, M. E.; Perez, A.; Lin, R.; Sajjadi, A.; Ricci, A. J.; Cheng, A. G. Towards the Prevention of Aminoglycoside-Related Hearing Loss. *Front. Cell. Neurosci.* **2017**, *11*, 325. <https://doi.org/10.3389/fncel.2017.00325>.
- (90) Pokrovskaya, V.; Belakhov, V.; Hainrichson, M.; Yaron, S.; Baasov, T. Design, Synthesis, and Evaluation of Novel Fluoroquinolone–Aminoglycoside Hybrid Antibiotics. *J. Med. Chem.* **2009**, *52* (8), 2243–2254. <https://doi.org/10.1021/jm900028n>.
- (91) Degtyareva, N. N.; Gong, C.; Story, S.; Levinson, N. S.; Oyelere, A. K.; Green, K. D.; Garneau-Tsodikova, S.; Arya, D. P. Antimicrobial Activity, AME Resistance, and A-Site Binding Studies of Anthraquinone–Neomycin Conjugates. *ACS Infect. Dis.* **2017**, *3* (3), 206–215. <https://doi.org/10.1021/acsinfecdis.6b00176>.
- (92) Story, S.; Skriba, M. J.; Maiti, K.; Nihar Ranjan; Degtyareva, N. N.; Green, K. D.; Khodaverdian, V.; Oyelere, A. K.; Garneau-Tsodikova, S.; Arya, D. P. Synthesis, Antimicrobial Activity, Attenuation of Aminoglycoside Resistance in MRSA, and Ribosomal A-Site Binding of Pyrene-Neomycin Conjugates. *Eur. J. Med. Chem.* **2019**, *163*, 381–393. <https://doi.org/10.1016/j.ejmech.2018.11.022>.
- (93) Aradi, K.; Di Giorgio, A.; Duca, M. Aminoglycoside Conjugation for RNA Targeting: Antimicrobials and Beyond. *Chem. – Eur. J.* **2020**, *26* (54), 12273–12309. <https://doi.org/10.1002/chem.202002258>.
- (94) McCown, P. J.; Corbino, K. A.; Stav, S.; Sherlock, M. E.; Breaker, R. R. Riboswitch Diversity and Distribution. *RNA* **2017**, *23* (7), 995–1011. <https://doi.org/10.1261/rna.061234.117>.
- (95) Panchal, V.; Brenk, R. Riboswitches as Drug Targets for Antibiotics. *Antibiotics* **2021**, *10* (1), 45. <https://doi.org/10.3390/antibiotics10010045>.
- (96) Howe, J. A.; Wang, H.; Fischmann, T. O.; Balibar, C. J.; Xiao, L.; Galgoci, A. M.; Malinverni, J. C.; Mayhood, T.; Villafania, A.; Nahvi, A.; Murgolo, N.; Barbieri, C. M.; Mann, P. A.; Carr, D.; Xia, E.; Zuck, P.; Riley, D.; Painter, R. E.; Walker, S. S.; Sherborne, B.; de Jesus, R.; Pan, W.; Plotkin, M. A.; Wu, J.; Rindgen, D.; Cummings, J.; Garlisi, C. G.; Zhang, R.; Sheth, P. R.; Gill, C. J.; Tang, H.; Roemer, T. Selective Small-Molecule Inhibition of an RNA Structural Element. *Nature* **2015**, *526* (7575), 672–677. <https://doi.org/10.1038/nature15542>.
- (97) Rizvi, N. F.; Howe, J. A.; Nahvi, A.; Klein, D. J.; Fischmann, T. O.; Kim, H.-Y.; McCoy, M. A.; Walker, S. S.; Hruza, A.; Richards, M. P.; Chamberlin, C.; Saradjian, P.; Butko, M. T.; Mercado, G.; Burchard, J.; Strickland, C.; Dandliker, P. J.; Smith, G. F.; Nickbarg, E. B. Discovery of Selective RNA-Binding Small Molecules by Affinity-Selection Mass Spectrometry. *ACS Chem. Biol.* **2018**, *13* (3), 820–831. <https://doi.org/10.1021/acscchembio.7b01013>.

- (98) Motika, S. E.; Ulrich, R. J.; Geddes, E. J.; Lee, H. Y.; Lau, G. W.; Hergenrother, P. J. Gram-Negative Antibiotic Active Through Inhibition of an Essential Riboswitch. *J. Am. Chem. Soc.* **2020**, *142* (24), 10856–10862. <https://doi.org/10.1021/jacs.0c04427>.
- (99) Hermann, T. Small Molecules Targeting Viral RNA: Small Molecules Targeting Viral RNA. *Wiley Interdiscip. Rev. RNA* **2016**, *7* (6), 726–743. <https://doi.org/10.1002/wrna.1373>.
- (100) Hermann, T. Viral RNA Targets and Their Small Molecule Ligands. In *RNA Therapeutics*; Garner, A. L., Ed.; Topics in Medicinal Chemistry; Springer International Publishing: Cham, 2017; Vol. 27, pp 111–134. [https://doi.org/10.1007/7355\\_2016\\_20](https://doi.org/10.1007/7355_2016_20).
- (101) Di Giorgio, A.; Duca, M. Synthetic Small-Molecule RNA Ligands: Future Prospects as Therapeutic Agents. *MedChemComm* **2019**, *10* (8), 1242–1255. <https://doi.org/10.1039/C9MD00195F>.
- (102) Blond, A.; Ennifar, E.; Tisé, C.; Micouin, L. The Design of RNA Binders: Targeting the HIV Replication Cycle as a Case Study. *ChemMedChem* **2014**, *9* (9), 1982–1996. <https://doi.org/10.1002/cmdc.201402259>.
- (103) Le Grice, S. F. J. Targeting the HIV RNA Genome: High-Hanging Fruit Only Needs a Longer Ladder. In *The Future of HIV-1 Therapeutics*; Torbett, B. E., Goodsell, D. S., Richman, D. D., Eds.; Current Topics in Microbiology and Immunology; Springer International Publishing: Cham, 2015; Vol. 389, pp 147–169. [https://doi.org/10.1007/82\\_2015\\_434](https://doi.org/10.1007/82_2015_434).
- (104) Sztuba-Solinska, J.; Shenoy, S. R.; Gareiss, P.; Krumpke, L. R. H.; Le Grice, S. F. J.; O’Keefe, B. R.; Schneekloth, J. S. Identification of Biologically Active, HIV TAR RNA-Binding Small Molecules Using Small Molecule Microarrays. *J. Am. Chem. Soc.* **2014**, *136* (23), 8402–8410. <https://doi.org/10.1021/ja502754f>.
- (105) Abulwerdi, F. A.; Shortridge, M. D.; Sztuba-Solinska, J.; Wilson, R.; Le Grice, S. F. J.; Varani, G.; Schneekloth, J. S. Development of Small Molecules with a Noncanonical Binding Mode to HIV-1 Trans Activation Response (TAR) RNA. *J. Med. Chem.* **2016**, *59* (24), 11148–11160. <https://doi.org/10.1021/acs.jmedchem.6b01450>.
- (106) Frank, A. T.; Stelzer, A. C.; Al-Hashimi, H. M.; Andricioaei, I. Constructing RNA Dynamical Ensembles by Combining MD and Motionally Decoupled NMR RDCs: New Insights into RNA Dynamics and Adaptive Ligand Recognition. *Nucleic Acids Res.* **2009**, *37* (11), 3670–3679. <https://doi.org/10.1093/nar/gkp156>.
- (107) Stelzer, A. C.; Frank, A. T.; Kratz, J. D.; Swanson, M. D.; Gonzalez-Hernandez, M. J.; Lee, J.; Andricioaei, I.; Markovitz, D. M.; Al-Hashimi, H. M. Discovery of Selective Bioactive Small Molecules by Targeting an RNA Dynamic Ensemble. *Nat. Chem. Biol.* **2011**, *7* (8), 553–559. <https://doi.org/10.1038/nchembio.596>.
- (108) Patwardhan, N. N.; Ganser, L. R.; Kapral, G. J.; Eubanks, C. S.; Lee, J.; Sathyamoorthy, B.; Al-Hashimi, H. M.; Hargrove, A. E. Amiloride as a New RNA-Binding Scaffold with Activity against HIV-1 TAR. *MedChemComm* **2017**, *8* (5), 1022–1036. <https://doi.org/10.1039/C6MD00729E>.
- (109) Patwardhan, N. N.; Cai, Z.; Umuhire Juru, A.; Hargrove, A. E. Driving Factors in Amiloride Recognition of HIV RNA Targets. *Org. Biomol. Chem.* **2019**, *17* (42), 9313–9320. <https://doi.org/10.1039/C9OB01702J>.
- (110) Davila-Calderon, J.; Patwardhan, N. N.; Chiu, L.-Y.; Sugarman, A.; Cai, Z.; Penutmutchu, S. R.; Li, M.-L.; Brewer, G.; Hargrove, A. E.; Tolbert, B. S. IRES-Targeting Small Molecule Inhibits Enterovirus

- 71 Replication via Allosteric Stabilization of a Ternary Complex. *Nat. Commun.* **2020**, *11* (1), 4775. <https://doi.org/10.1038/s41467-020-18594-3>.
- (111) Guianvarc'h, D.; Benhida, R.; Fourrey, J.-L.; Maurisse, R.; Sun, J.-S. Incorporation of a Novel Nucleobase Allows Stable Oligonucleotide-Directed Triple Helix Formation at the Target Sequence Containing a Purine-pyrimidine Interruption. *Chem. Commun.* **2001**, No. 18, 1814–1815. <https://doi.org/10.1039/b103743a>.
- (112) Duca, M.; Malnuit, V.; Barbault, F.; Benhida, R. Design of Novel RNA Ligands That Bind Stem–Bulge HIV-1 TAR RNA. *Chem. Commun.* **2010**, *46* (33), 6162. <https://doi.org/10.1039/c0cc00645a>.
- (113) Joly, J.-P.; Mata, G.; Eldin, P.; Briant, L.; Fontaine-Vive, F.; Duca, M.; Benhida, R. Artificial Nucleobase-Amino Acid Conjugates: A New Class of TAR RNA Binding Agents. *Chem. - Eur. J.* **2014**, *20* (7), 2071–2079. <https://doi.org/10.1002/chem.201303664>.
- (114) Shortridge, M. D.; Wille, P. T.; Jones, A. N.; Davidson, A.; Bogdanovic, J.; Arts, E.; Karn, J.; Robinson, J. A.; Varani, G. An Ultra-High Affinity Ligand of HIV-1 TAR Reveals the RNA Structure Recognized by P-TEFb. *Nucleic Acids Res.* **2019**, *47* (3), 1523–1531. <https://doi.org/10.1093/nar/gky1197>.
- (115) Ennifar, E.; Paillart, J.-C.; Marquet, R.; Ehresmann, B.; Ehresmann, C.; Dumas, P.; Walter, P. HIV-1 RNA Dimerization Initiation Site Is Structurally Similar to the Ribosomal A Site and Binds Aminoglycoside Antibiotics. *J. Biol. Chem.* **2003**, *278* (4), 2723–2730. <https://doi.org/10.1074/jbc.M205726200>.
- (116) Ennifar, E.; Aslam, M. W.; Strasser, P.; Hoffmann, G.; Dumas, P.; van Delft, F. L. Structure-Guided Discovery of a Novel Aminoglycoside Conjugate Targeting HIV-1 RNA Viral Genome. *ACS Chem. Biol.* **2013**, *8* (11), 2509–2517. <https://doi.org/10.1021/cb400498n>.
- (117) Lodmell, J. S.; Ehresmann, C.; Ehresmann, B.; Marquet, R. Structure and Dimerization of HIV-1 Kissing Loop Aptamers 1 Edited by D. E. Draper. *J. Mol. Biol.* **2001**, *311* (3), 475–490. <https://doi.org/10.1006/jmbi.2001.4879>.
- (118) Weixlbaumer, A.; Werner, A.; Flamm, C.; Westhof, E. Determination of Thermodynamic Parameters for HIV DIS Type Loop–Loop Kissing Complexes. **8**.
- (119) Rist, M. J.; Marino, J. P. Mechanism of Nucleocapsid Protein Catalyzed Structural Isomerization of the Dimerization Initiation Site of HIV-1. **9**.
- (120) Ennifar, E. Targeting the Dimerization Initiation Site of HIV-1 RNA with Aminoglycosides: From Crystal to Cell. *Nucleic Acids Res.* **2006**, *34* (8), 2328–2339. <https://doi.org/10.1093/nar/gkl317>.
- (121) Bernacchi, S.; Freisz, S.; Maechling, C.; Spiess, B.; Marquet, R.; Dumas, P.; Ennifar, E. Aminoglycoside Binding to the HIV-1 RNA Dimerization Initiation Site: Thermodynamics and Effect on the Kissing-Loop to Duplex Conversion. *Nucleic Acids Res.* **2007**, *35* (21), 7128–7139. <https://doi.org/10.1093/nar/gkm856>.
- (122) Ennifar, E.; Paillart, J.; Bernacchi, S.; Walter, P.; Pale, P.; Decout, J.; Marquet, R.; Dumas, P. A Structure-Based Approach for Targeting the HIV-1 Genomic RNA Dimerization Initiation Site. *Biochimie* **2007**, *89* (10), 1195–1203. <https://doi.org/10.1016/j.biochi.2007.03.003>.
- (123) Bodlenner, A.; Alix, A.; Weibel, J.-M.; Pale, P.; Ennifar, E.; Paillart, J.-C.; Walter, P.; Marquet, R.; Dumas, P. Synthesis of a Neamine Dimer Targeting the Dimerization Initiation Site of HIV-1 RNA. *Org. Lett.* **2007**, *9* (22), 4415–4418. <https://doi.org/10.1021/ol701760k>.

- (124) Dohno, C.; Kimura, M.; Nakatani, K. Restoration of Ribozyme Tertiary Contact and Function by Using a Molecular Glue for RNA. *Angew Chem Int Ed* **2018**, *5*.
- (125) Meyer, S. M.; Williams, C. C.; Akahori, Y.; Tanaka, T.; Aikawa, H.; Tong, Y.; Childs-Disney, J. L.; Disney, M. D. Small Molecule Recognition of Disease-Relevant RNA Structures. *Chem. Soc. Rev.* **2020**, *49* (19), 7167–7199. <https://doi.org/10.1039/D0CS00560F>.
- (126) Donlic, A.; Hargrove, A. E. Targeting RNA in Mammalian Systems with Small Molecules. *Wiley Interdiscip. Rev. RNA* **2018**, *9* (4), e1477. <https://doi.org/10.1002/wrna.1477>.
- (127) Verma, A. K.; Khan, E.; Bhagwat, S. R.; Kumar, A. Exploring the Potential of Small Molecule-Based Therapeutic Approaches for Targeting Trinucleotide Repeat Disorders. *Mol. Neurobiol.* **2020**, *57* (1), 566–584. <https://doi.org/10.1007/s12035-019-01724-4>.
- (128) Ratni, H.; Ebeling, M.; Baird, J.; Bendels, S.; Bylund, J.; Chen, K. S.; Denk, N.; Feng, Z.; Green, L.; Guerard, M.; Jablonski, P.; Jacobsen, B.; Khwaja, O.; Kletzl, H.; Ko, C.-P.; Kustermann, S.; Marquet, A.; Metzger, F.; Mueller, B.; Naryshkin, N. A.; Paushkin, S. V.; Pinard, E.; Poirier, A.; Reutlinger, M.; Weetall, M.; Zeller, A.; Zhao, X.; Mueller, L. Discovery of Risdiplam, a Selective Survival of Motor Neuron-2 (SMN2) Gene Splicing Modifier for the Treatment of Spinal Muscular Atrophy (SMA). *J. Med. Chem.* **2018**, *61* (15), 6501–6517. <https://doi.org/10.1021/acs.jmedchem.8b00741>.
- (129) Campagne, S.; Boigner, S.; Rüdissler, S.; Moursy, A.; Gillioz, L.; Knörlein, A.; Hall, J.; Ratni, H.; Cléry, A.; Allain, F. H.-T. Structural Basis of a Small Molecule Targeting RNA for a Specific Splicing Correction. *Nat. Chem. Biol.* **2019**, *15* (12), 1191–1198. <https://doi.org/10.1038/s41589-019-0384-5>.
- (130) Connelly, C. M.; Moon, M. H.; Schneekloth, J. S. The Emerging Role of RNA as a Therapeutic Target for Small Molecules. *Cell Chem. Biol.* **2016**, *23* (9), 1077–1090. <https://doi.org/10.1016/j.chembiol.2016.05.021>.
- (131) Costales, M. G.; Matsumoto, Y.; Velagapudi, S. P.; Disney, M. D. Small Molecule Targeted Recruitment of a Nuclease to RNA. *J. Am. Chem. Soc.* **2018**, *140* (22), 6741–6744. <https://doi.org/10.1021/jacs.8b01233>.
- (132) Velagapudi, S. P.; Cameron, M. D.; Haga, C. L.; Rosenberg, L. H.; Lafitte, M.; Duckett, D. R.; Phinney, D. G.; Disney, M. D. Design of a Small Molecule against an Oncogenic Noncoding RNA. *Proc. Natl. Acad. Sci.* **2016**, *113* (21), 5898–5903. <https://doi.org/10.1073/pnas.1523975113>.
- (133) Di Giorgio, A.; Tran, T. P. A.; Duca, M. Small-Molecule Approaches toward the Targeting of Oncogenic MiRNAs: Roadmap for the Discovery of RNA Modulators. *Future Med. Chem.* **2016**, *8* (7), 803–816. <https://doi.org/10.4155/fmc-2016-0018>.
- (134) Lorenz, D. A.; Garner, A. L. Approaches for the Discovery of Small Molecule Ligands Targeting MicroRNAs. In *RNA Therapeutics*; Garner, A. L., Ed.; Topics in Medicinal Chemistry; Springer International Publishing: Cham, 2017; Vol. 27, pp 79–110. [https://doi.org/10.1007/7355\\_2017\\_3](https://doi.org/10.1007/7355_2017_3).
- (135) Verma et al. - 2020 - Exploring the Potential of Small Molecule-Based Th.Pdf.
- (136) Vo, D. D.; Stadel, C.; Zehnacker, L.; Benhida, R.; Darfeuille, F.; Duca, M. Targeting the Production of Oncogenic MicroRNAs with Multimodal Synthetic Small Molecules. *ACS Chem. Biol.* **2014**, *9* (3), 711–721. <https://doi.org/10.1021/cb400668h>.
- (137) Tran, T. P. A. Ribosome-Targeting Antibiotics as Inhibitors of Oncogenic MicroRNAs Biogenesis: Old Scaffolds for New Perspectives in RNA Targeting. *Bioorg Med Chem* **2015**, *11*.

- (138) Malnuit, V.; Duca, M.; Benhida, R. Targeting DNA Base Pair Mismatch with Artificial Nucleobases. Advances and Perspectives in Triple Helix Strategy. *Org Biomol Chem* **2011**, *9* (2), 326–336. <https://doi.org/10.1039/C0OB00418A>.
- (139) Vo, D. D.; Tran, T. P. A.; Staedel, C.; Benhida, R.; Darfeuille, F.; Giorgio, A. D.; Duca, M. Oncogenic MicroRNAs Biogenesis as a Drug Target: Structure–Activity Relationship Studies on New Aminoglycoside Conjugates. *Chem Eur J* **2016**, *22* (15), 5350–5362.
- (140) Org.Biomol.Chem.2018.Duca - Neo-B-AA - Mir372.Pdf.
- (141) Staedel, C.; Tran, T. P. A.; Giraud, J.; Darfeuille, F.; Di Giorgio, A.; Tourasse, N. J.; Salin, F.; Uriac, P.; Duca, M. Modulation of Oncogenic MiRNA Biogenesis Using Functionalized Polyamines. *Sci. Rep.* **2018**, *8* (1), 1667. <https://doi.org/10.1038/s41598-018-20053-5>.
- (142) Murata, A.; Otabe, T.; Zhang, J.; Nakatani, K. BzDANP, a Small-Molecule Modulator of Pre-MiR-29a Maturation by Dicer. *ACS Chem. Biol.* **2016**, *11* (10), 2790–2796. <https://doi.org/10.1021/acscchembio.6b00214>.
- (143) Murata, A.; Nakamori, M.; Nakatani, K. Modulating RNA Secondary and Tertiary Structures by Mismatch Binding Ligands. *Methods* **2019**, *167*, 78–91. <https://doi.org/10.1016/j.ymeth.2019.05.006>.
- (144) Hong, C.; Otabe, T.; Matsumoto, S.; Dohno, C.; Murata, A.; Hagihara, M.; Nakatani, K. Formation of a Ligand-Assisted Complex of Two RNA Hairpin Loops. *Chem. - Eur. J.* **2014**, *20* (18), 5282–5287. <https://doi.org/10.1002/chem.201304683>.
- (145) Murata, A.; Mori, Y.; Di, Y.; Sugai, A.; Das, B.; Takashima, Y.; Nakatani, K. Small Molecule-Induced Dimerization of Hairpin RNA Interfered with the Dicer Cleavage Reaction. *Biochemistry* **2021**, *60* (4), 245–249. <https://doi.org/10.1021/acs.biochem.0c00920>.
- (146) Dömling, A.; Ugi, I. Multicomponent Reactions with Isocyanides. *Angew Chem Int Ed* **2000**, *43*.
- (147) Versammlungsberichte. *Angew. Chem.* **1959**, *71* (11), 373–388. <https://doi.org/10.1002/ange.19590711110>.
- (148) Hermann, T.; Westhof, E. Aminoglycoside Binding to the Hammerhead Ribozyme: A General Model for the Interaction of Cationic Antibiotics with RNA. *10*.
- (149) Thomas, J. R.; Liu, X.; Hergenrother, P. J. Size-Specific Ligands for RNA Hairpin Loops. *J. Am. Chem. Soc.* **2005**, *127* (36), 12434–12435. <https://doi.org/10.1021/ja051685b>.
- (150) Portlock, D. E.; Naskar, D.; West, L.; Ostaszewski, R.; Chen, J. J. Solid-Phase Synthesis of Five-Dimensional Libraries via a Tandem Petasis–Ugi Multi-Component Condensation Reaction. *Tetrahedron Lett.* **2003**, *44* (27), 5121–5124. [https://doi.org/10.1016/S0040-4039\(03\)01119-5](https://doi.org/10.1016/S0040-4039(03)01119-5).
- (151) Ruijter, E.; Scheffelaar, R.; Orru, R. V. A. Multicomponent Reaction Design in the Quest for Molecular Complexity and Diversity. *Angew. Chem. Int. Ed.* **2011**, *50* (28), 6234–6246. <https://doi.org/10.1002/anie.201006515>.
- (152) Dömling, A.; Wang, W.; Wang, K. Chemistry and Biology Of Multicomponent Reactions. *Chem. Rev.* **2012**, *112* (6), 3083–3135. <https://doi.org/10.1021/cr100233r>.
- (153) Ugi, I.; Mayr, R.; Fetzer, U.; Steinbrückner, C. Versammlungsberichte. *Angew. Chem.* **1959**, *71* (11), 373–388. <https://doi.org/10.1002/ange.19590711110>.



- (154) Dömling, A. Recent Developments in Isocyanide Based Multicomponent Reactions in Applied Chemistry. *Chem. Rev.* **2006**, *106* (1), 17–89. <https://doi.org/10.1021/cr0505728>.
- (155) Alvim, H. G. O. What Do We Know about Multicomponent Reactions? Mechanisms and Trends for the Biginelli, Hantzsch, Mannich, Passerini and Ugi MCRs. *RSC Adv.* **2014**, *18*.
- (156) Dömling, A.; Huang, Y. Piperazine Scaffolds via Isocyanide-Based Multicomponent Reactions. *Synthesis* **2010**, *2010* (17), 2859–2883. <https://doi.org/10.1055/s-0030-1257906>.
- (157) Pharande, S. G. Synthesis of Lactams via Isocyanide-Based Multicomponent Reactions. *Synthesis* **2021**, *53* (03), 418–446. <https://doi.org/10.1055/s-0040-1706297>.
- (158) Marcaccini, S.; Torroba, T. The Use of the Ugi Four-Component Condensation. *Nat. Protoc.* **2007**, *2* (3), 632–639. <https://doi.org/10.1038/nprot.2007.71>.
- (159) Zhang, J.; Yu, P.; Li, S.-Y.; Sun, H.; Xiang, S.-H.; Wang, J. (Joelle); Houk, K. N.; Tan, B. Asymmetric Phosphoric Acid-Catalyzed Four-Component Ugi Reaction. *Science* **2018**, *361* (6407), eaas8707. <https://doi.org/10.1126/science.aas8707>.
- (160) Rocha, R. O.; Rodrigues, M. O.; Neto, B. A. D. Review on the Ugi Multicomponent Reaction Mechanism and the Use of Fluorescent Derivatives as Functional Chromophores. *ACS Omega* **2020**, *5* (2), 972–979. <https://doi.org/10.1021/acsomega.9b03684>.
- (161) Cankařová, N.; Krchňák, V. Isocyanide Multicomponent Reactions on Solid Phase: State of the Art and Future Application. *Int. J. Mol. Sci.* **2020**, *21* (23), 9160. <https://doi.org/10.3390/ijms21239160>.
- (162) Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85* (14), 2149–2154. <https://doi.org/10.1021/ja00897a025>.
- (163) Chen, J. J.; Golebiowski, A.; Klopfenstein, S. R.; McClenaghan, J.; Peng, S. X.; Portlock, D. E.; West, L. Multi-Component Synthesis of Imidazo[1,2-a] Annulated Heterocycles on  $\alpha$ -Isocyanato Resin Esters. *Synlett* **2001**, No. 8, 3.
- (164) Henkel, B.; Sax, M.; Dömling, A. New Method for the Preparation of Solid-Phase Bound Isocyanocarboxylic Acids and Ugi Reactions Therewith. *Tetrahedron Lett.* **2003**, *44* (37), 7015–7018. [https://doi.org/10.1016/S0040-4039\(03\)01793-3](https://doi.org/10.1016/S0040-4039(03)01793-3).
- (165) Hulme, C.; Peng, J.; Morton, G.; Salvino, J. M.; Herpin, T.; Labaudiniere, R. Novel Safety-Catch Linker and Its Application with a Ugi/De-BOC/Cyclization (UDC) Strategy to Access Carboxylic Acids, 1,4-Benzodiazepines, Diketopiperazines, Ketopiperazines and Dihydroquinoxalinones. *Tetrahedron Lett.* **1998**, *39* (40), 7227–7230. [https://doi.org/10.1016/S0040-4039\(98\)01593-7](https://doi.org/10.1016/S0040-4039(98)01593-7).
- (166) Chen, J. J.; Golebiowski, A.; McClenaghan, J.; Klopfenstein, S. R.; West, L. Universal Rink-Isonitrile Resin: Application for the Traceless Synthesis of 3-Acylamino Imidazo[1,2-a]Pyridines. *Tetrahedron Lett.* **2001**, *42* (12), 2269–2271. [https://doi.org/10.1016/S0040-4039\(01\)00159-9](https://doi.org/10.1016/S0040-4039(01)00159-9).
- (167) Henkel, B.; Weber, L. A Novel Four-Component Synthesis of *N*-Substituted Amino Acid Esters. *Synlett* **2002**, *2002* (11), 1877–1879. <https://doi.org/10.1055/s-2002-34864>.
- (168) Porcheddu, A.; Giacomelli, G.; Piredda, I.; Carta, M.; Nieddu, G. A Practical and Efficient Approach to PNA Monomers Compatible with Fmoc-Mediated Solid-Phase Synthesis Protocols. *Eur J Org Chem* **2008**, *12*.

- (169) Pothukanuri, S.; Pianowski, Z.; Winssinger, N. Expanding the Scope and Orthogonality of PNA Synthesis. *Eur. J. Org. Chem.* **2008**, 2008 (18), 3141–3148. <https://doi.org/10.1002/ejoc.200800141>.
- (170) Huang, Y.-C.; Cao, C.; Tan, X.-L.; Li, X.; Liu, L. Facile Solid-Phase Synthesis of PNA–Peptide Conjugates Using PNZ-Protected PNA Monomers. *Org Chem Front* **2014**, 1 (9), 1050–1054. <https://doi.org/10.1039/C4QO00217B>.
- (171) Stafford, J. A.; Brackeen, M. F.; Karanewsky, D. S.; Valvano, N. L. A Highly Selective Protocol for the Deprotection of BOC-Protected Amides and Carbamates. *Tetrahedron Lett.* **1993**, 34 (49), 7873–7876. [https://doi.org/10.1016/S0040-4039\(00\)61498-3](https://doi.org/10.1016/S0040-4039(00)61498-3).
- (172) Darnbrough, S.; Mervic, M.; Condon, S. M.; Burns, C. J. AN IMPROVED SYNTHESIS OF *N*-BOC PROTECTED ARYL AMINES. *Synth. Commun.* **2001**, 31 (21), 3273–3280. <https://doi.org/10.1081/SCC-100106036>.
- (173) Geen, G. R.; Grinter, T. J.; Kinsey, P. M.; Jarvest, R. L. The Effect of the C-6 Substituent on the Regioselectivity of *n*-Alkylation of 2-Aminopurines. *Tetrahedron* **1990**, 46 (19), 6903–6914. [https://doi.org/10.1016/S0040-4020\(01\)87878-9](https://doi.org/10.1016/S0040-4020(01)87878-9).
- (174) Litvinov, V. P. Chemistry and Biological Activities of 1,8-Naphthyridines. *Russ. Chem. Rev.* **2004**, 73 (7), 637–670. <https://doi.org/10.1070/RC2004v073n07ABEH000856>.
- (175) Ma, Y.; Kolotuchin, S. V.; Zimmerman, S. C. Supramolecular Polymer Chemistry: Self-Assembling Dendrimers Using the DDA·AAD (GC-like) Hydrogen Bonding Motif. *J. Am. Chem. Soc.* **2002**, 124 (46), 13757–13769. <https://doi.org/10.1021/ja0202006>.
- (176) Ghosh, K.; Sen, T.; Fröhlich, R. Hydrogen Bonded Assemblies of 1,8-Naphthyridine Derivatives: Discrete or Polymeric Structures in the Solid State. *J. Incl. Phenom. Macrocycl. Chem.* **2010**, 68 (1–2), 193–199. <https://doi.org/10.1007/s10847-010-9766-8>.
- (177) Ośmiałowski, B.; Kolehmainen, E.; Kalenius, E.; Behera, B.; Kauppinen, R.; Sievänen, E. Intermolecular Steric Hindrance in 7-Acylamino-[1H]-2-Oxo-1,8-Naphthyridines: NMR, ESI-MS, IR, and DFT Calculation Studies. *Struct. Chem.* **2011**, 22 (5), 1143–1151. <https://doi.org/10.1007/s11224-011-9808-x>.
- (178) Nakatani, K.; Sando, S.; Saito, I. Scanning of Guanine–Guanine Mismatches in DNA by Synthetic Ligands Using Surface Plasmon Resonance. *Nat. Biotechnol.* **2001**, 19 (1), 51–55. <https://doi.org/10.1038/83505>.
- (179) Zhang, J.; Takei, F.; Nakatani, K. Emission of Characteristic Fluorescence from the Ligand–Cytosine Complex in U<sub>A</sub>/ACU Bulged RNA Duplex. *Bioorg. Med. Chem.* **2007**, 15 (14), 4813–4817. <https://doi.org/10.1016/j.bmc.2007.05.007>.
- (180) Newkome, G. R.; Garbis, S. J.; Majestic, V. K.; Fronczek, F. R.; Chiari, G. Chemistry of Heterocyclic Compounds. 61. Synthesis and Conformational Studies of Macrocycles Possessing 1,8- or 1,5-Naphthyridino Subunits Connected by Carbon-Oxygen Bridges. *J. Org. Chem.* **1981**, 46 (5), 833–839. <https://doi.org/10.1021/jo00318a001>.
- (181) Rashad, A. E.; Shamroukh, A. H.; El-Hashash, M. A.; El-Faragy, A. F.; Yousif, N. M.; Salama, M. A.; Mostafa, A.; El-Shahat, M. Synthesis and Anti-Avian Influenza Virus (H5N1) Evaluation of Some Novel Nicotinonitriles and Their *N*-Acyclic Nucleosides: Synthesis and Anti-Avian Influenza Virus (H5N1) Evaluation of Some Novel Nicotinonitriles and Their *N*-Acyclic Nucleosides. *J. Heterocycl. Chem.* **2012**, 49 (5), 1130–1135. <https://doi.org/10.1002/jhet.966>.

- (182) Leonard, J. T.; Gangadhar, R.; Gnanasam, S. K.; Ramachandran, S.; Saravanan, M.; Sridhar, S. K. Synthesis and Pharmacological Activities of 1,8-Naphthyridine Derivatives. *Biol. Pharm. Bull.* **2002**, *25* (6), 798–802. <https://doi.org/10.1248/bpb.25.798>.
- (183) Tanaka, K.; Murakami, M.; Jeon, J.-H.; Chujo, Y. Enhancement of Affinity in Molecular Recognition Viahydrogen Bonds by POSS-Core Dendrimer and Its Application for Selective Complex Formation between Guanosine Triphosphate and 1,8-Naphthyridine Derivatives. *Org. Biomol. Chem.* **2012**, *10* (1), 90–95. <https://doi.org/10.1039/C1OB06630G>.
- (184) Goswami, S.; Mukherjee, R.; Mukherjee, R.; Jana, S.; Maity, A.; Adak, A. Simple and Efficient Synthesis of 2,7-Difunctionalized-1,8-Naphthyridines. *Molecules* **2005**, *10* (8), 929–936. <https://doi.org/10.3390/10080929>.
- (185) Tarrago, G.; Zidane, I.; Marzin, C.; Tep, A. Synthesis and Ionophore Properties of a Series of New Tetrapyrrolic Macrocycles. *Tetrahedron* **1988**, *44* (1), 91–100. [https://doi.org/10.1016/S0040-4020\(01\)85096-1](https://doi.org/10.1016/S0040-4020(01)85096-1).
- (186) Wang, H.; Tor, Y. RNA  $\pm$  Aminoglycoside Interactions: Design, Synthesis, and Binding of <sup>a</sup>Amino-Aminoglycosides<sup>o</sup> to RNA\*. *Angew. Chem. Int. Ed.* **1998**.
- (187) Yoshizawa, S.; Fourmy, D.; Eason, R. G.; Puglisi, J. D. Sequence-Specific Recognition of the Major Groove of RNA by Deoxystreptamine <sup>†</sup>. *Biochemistry* **2002**, *41* (20), 6263–6270. <https://doi.org/10.1021/bi0121609>.
- (188) Busscher, G. F.; Rutjes, F. P. J. T.; van Delft, F. L. 2-Deoxystreptamine: Central Scaffold of Aminoglycoside Antibiotics <sup>†</sup>. *Chem. Rev.* **2005**, *105* (3), 775–792. <https://doi.org/10.1021/cr0404085>.
- (189) Klemm, C. M.; Berthelmann, A.; Neubacher, S.; Arenz, C. Short and Efficient Synthesis of Alkyne-Modified Amino Glycoside Building Blocks. *Eur. J. Org. Chem.* **2009**, *17*, 2788–2794.
- (190) Alper, P. B.; Hung, S.-C.; Wong, C.-H. Metal Catalyzed Diazo Transfer for the Synthesis of Azides from Amines. *Tetrahedron Lett.* **1996**, *37* (34), 6029–6032. [https://doi.org/10.1016/0040-4039\(96\)01307-X](https://doi.org/10.1016/0040-4039(96)01307-X).
- (191) Canas-Rodriguez, A.; Galan Ruiz-Poveda, S.; Gomez-Sanchez, A.; Blanco Gonzalez, E. 5,6-O-Cyclohexylidene and 5,6-O-Cyclohexylidene-1,3-Di-N-Methoxycarbonyl Derivatives of 2-Deoxystreptamine. *Carbohydr. Res.* **1977**, *59* (1), 240–243. [https://doi.org/10.1016/S0008-6215\(00\)83309-1](https://doi.org/10.1016/S0008-6215(00)83309-1).
- (192) van den Broek, S. (Bas) A. M. W.; Gruijters, B. W. T.; Rutjes, F. P. J. T.; van Delft, F. L.; Blaauw, R. H. A Short and Scalable Route to Orthogonally O-Protected 2-Deoxystreptamine. *J. Org. Chem.* **2007**, *72* (9), 3577–3580. <https://doi.org/10.1021/jo062369y>.
- (193) 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 Responsive Element. *J. Am. Chem. Soc.* **1996**, *118* (42), 10150–10155. <https://doi.org/10.1021/ja9612817>.
- (194) Rosenbohm, C.; Vanden Berghe, D.; Vlietinck, A.; Wengel, J. Parallel Synthesis of a Small Library of Novel Aminoglycoside Analogs Based on 2-Amino-2-Deoxy-d-Glucose and d-Ribose Scaffolds. *Tetrahedron* **2001**, *57* (29), 6277–6287. [https://doi.org/10.1016/S0040-4020\(01\)00584-1](https://doi.org/10.1016/S0040-4020(01)00584-1).

- (195) Westermann, B.; Dörner, S. Synthesis of Multivalent Aminoglycoside Mimics via the Ugi Multicomponent Reaction. *Chem Commun* **2005**, No. 16, 2116–2118. <https://doi.org/10.1039/B501028D>.
- (196) Banoub, J.; Boullanger, P.; Lafont, D. Synthesis of Oligosaccharides of 2-Amino-2-Deoxy Sugars. *Chem. Rev.* **1992**, *92* (6), 1167–1195. <https://doi.org/10.1021/cr00014a002>.
- (197) Silva, D. J.; Wang, H.; Allanson, N. M.; Jain, R. K.; Sofia, M. J. Stereospecific Solution- and Solid-Phase Glycosylations. Synthesis of  $\beta$ -Linked Saccharides and Construction of Disaccharide Libraries Using Phenylsulfenyl 2-Deoxy-2-Trifluoroacetamido Glycopyranosides as Glycosyl Donors <sup>1</sup>. *J. Org. Chem.* **1999**, *64* (16), 5926–5929. <https://doi.org/10.1021/jo9903499>.
- (198) Lin, H.; Walsh, C. T. A Chemoenzymatic Approach to Glycopeptide Antibiotics. *J. Am. Chem. Soc.* **2004**, *126* (43), 13998–14003. <https://doi.org/10.1021/ja045147v>.
- (199) Borthwick, A. D. 2,5-Diketopiperazines: Synthesis, Reactions, Medicinal Chemistry, and Bioactive Natural Products. *Chem. Rev.* **2012**, *112* (7), 3641–3716. <https://doi.org/10.1021/cr200398y>.
- (200) Keating, T. A.; Armstrong, R. W. Postcondensation Modifications of Ugi Four-Component Condensation Products: 1-Isocyanocyclohexene as a Convertible Isocyanide. Mechanism of Conversion, Synthesis of Diverse Structures, and Demonstration of Resin Capture. *J. Am. Chem. Soc.* **1996**, *118* (11), 2574–2583. <https://doi.org/10.1021/ja953868b>.
- (201) Hulme, C.; Gore, V. “Multi-Component Reactions : Emerging Chemistry in Drug Discovery” “From Xylocain to Crixivan”; *Curr. Med. Chem.* **2003**, *10* (1), 51–80. <https://doi.org/10.2174/0929867033368600>.
- (202) Kennedy, A. L.; Fryer, A. M.; Josey, J. A. A New Resin-Bound Universal Isonitrile for the Ugi 4CC Reaction: Preparation and Applications to the Synthesis of 2,5-Diketopiperazines and 1,4-Benzodiazepine-2,5-Diones. *Org. Lett.* **2002**, *4* (7), 1167–1170. <https://doi.org/10.1021/ol0256015>.
- (203) Hulme, C.; Ma, L.; Cherrier, M.-P.; Romano, J. J.; Morton, G.; Duquenne, C.; Salvino, J.; Labaudiniere, R. Novel Applications of Convertible Isonitriles for the Synthesis of Mono and Bicyclic  $\gamma$ -Lactams via a UDC Strategy. *Tetrahedron Lett.* **2000**, *41* (12), 1883–1887. [https://doi.org/10.1016/S0040-4039\(00\)00052-6](https://doi.org/10.1016/S0040-4039(00)00052-6).
- (204) Habashita, H.; Kokubo, M.; Hamano, S.; Hamanaka, N.; Toda, M.; Shibayama, S.; Tada, H.; Sagawa, K.; Fukushima, D.; Maeda, K.; Mitsuya, H. Design, Synthesis, and Biological Evaluation of the Combinatorial Library with a New Spirodiketopiperazine Scaffold. Discovery of Novel Potent and Selective Low-Molecular-Weight CCR5 Antagonists. *J. Med. Chem.* **2006**, *49* (14), 4140–4152. <https://doi.org/10.1021/jm060051s>.
- (205) Aguiam, N. R.; Castro, V. I.; Ribeiro, A. I. F.; Fernandes, R. D. V.; Carvalho, C. M.; Costa, S. P. G.; Pereira-Lima, S. M. M. A.  $\alpha,\alpha$ -Dialkylglycines Obtained by Solid Phase Ugi Reaction Performed over Isocyanide Functionalized Resins. *Tetrahedron* **2013**, *69* (43), 9161–9165. <https://doi.org/10.1016/j.tet.2013.08.002>.
- (206) Jiang, W.-Q.; Costa, S. P. G.; Maia, H. L. S. Synthesis of N-Acyl-N, $\alpha,\alpha$ -Trialkyl and N-Acyl- $\alpha,\alpha$ -Dialkyl Glycines by Selective Cleavage of Ugi–Passerini Adducts. Qualitative Assessment of the Effect of Substituents on the Path and Yield of Reaction. *Org. Biomol. Chem.* **2003**, *1* (21), 3804–3810. <https://doi.org/10.1039/B307111C>.

- (207) Ramos-Tomillero, I.; Sánchez, M. K.; Rodríguez, H.; Albericio, F. Structure-Acid Lability Relationship of N-Alkylated  $\alpha,\alpha$ -Dialkylglycine Obtained via a Ugi Multicomponent Reaction. *Molecules* **2021**, *26* (1), 197. <https://doi.org/10.3390/molecules26010197>.
- (208) Samaritoni, J. G.; Copes, A. T.; Crews, D. K.; Glos, C.; Thompson, A. L.; Wilson, C.; O'Donnell, M. J.; Scott, W. L. Unexpected Hydrolytic Instability of N-Acylated Amino Acid Amides and Peptides. *J. Org. Chem.* **2014**, *79* (7), 3140–3151. <https://doi.org/10.1021/jo500273f>.
- (209) Llano-Sotelo, B.; Azucena, E. F.; Kotra, L. P.; Mobashery, S.; Chow, C. S. Aminoglycosides Modified by Resistance Enzymes Display Diminished Binding to the Bacterial Ribosomal Aminoacyl-TRNA Site. *Chem. Biol.* **2002**, *9* (4), 455–463. [https://doi.org/10.1016/S1074-5521\(02\)00125-4](https://doi.org/10.1016/S1074-5521(02)00125-4).
- (210) Lea, W. A.; Simeonov, A. Fluorescence Polarization Assays in Small Molecule Screening. *Expert Opin. Drug Discov.* **2011**, *6* (1), 17–32. <https://doi.org/10.1517/17460441.2011.537322>.
- (211) Luedtke, N. W.; Tor, Y. Fluorescence-Based Methods for Evaluating the RNA Affinity and Specificity of HIV-1 Rev-RRE Inhibitors. *Biopolymers* **2003**, *70* (1), 103–119. <https://doi.org/10.1002/bip.10428>.
- (212) Takahashi, K.-I.; Baba, S.; Chattopadhyay, P.; Koyanagi, Y.; Yamamoto, N.; Takaku, H.; Kawai, G. Structural Requirement for the Two-Step Dimerization of Human Immunodeficiency Virus Type 1 Genome. *RNA* **2000**, *6* (1), 96–102. <https://doi.org/10.1017/S1355838200991635>.
- (213) Chavali, S. S.; Bonn-Breach, R.; Wedekind, J. E. Face-Time with TAR: Portraits of an HIV-1 RNA with Diverse Modes of Effector Recognition Relevant for Drug Discovery. *J. Biol. Chem.* **2019**, *294* (24), 9326–9341. <https://doi.org/10.1074/jbc.REV119.006860>.
- (214) McConathy, J.; Zhou, D.; Shockley, S. E.; Jones, L. A.; Griffin, E. A.; Lee, H.; Adams, S. J.; Mach, R. H. Click Synthesis and Biologic Evaluation of ( *R* )- and ( *S* )-2-Amino-3-[1-(2-[  $^{18}$ F]Fluoroethyl)-1 *H* -[1,2,3]Triazol-4-yl]Propanoic Acid for Brain Tumor Imaging with Positron Emission Tomography. *Mol. Imaging* **2010**, *9* (6), 7290.2010.00025. <https://doi.org/10.2310/7290.2010.00025>.