



**HAL**  
open science

# Nouvelles stratégies synthétiques vers la formation d'hétérodinucléotides incorporant un motif difluorophosphinothioate.

Matthieu Roignant

## ► To cite this version:

Matthieu Roignant. Nouvelles stratégies synthétiques vers la formation d'hétérodinucléotides incorporant un motif difluorophosphinothioate.. Chimie organique. Normandie Université, 2019. Français. ⟨NNT: 2019NORMR090⟩. ⟨tel-03494312⟩

**HAL Id: tel-03494312**

**<https://theses.hal.science/tel-03494312v1>**

Submitted on 19 Dec 2021

**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.



HAL Authorization



Normandie Université

## THÈSE

**Pour obtenir le diplôme de doctorat**

**Spécialité Chimie**

**Préparée au sein de l'Université de Rouen Normandie**

# **Nouvelles stratégies synthétiques vers la formation d'hétérodinucléotides incorporant un motif difluorophosphinothioate**

**Présentée et soutenue par  
Mathieu ROIGNANT**

**Thèse soutenue publiquement le 18 décembre 2019  
devant le jury composé de**

Pr. Jacques Lebreton	Professeur Université de Nantes	Président
Pr. Isabelle Gillaizeau	Professeure Université d'Orléans	Rapportrice
Dr. Nicolas Blanchard	Directeur de recherche CNRS Laboratoire Innovation Moléculaire et Applications Mulhouse	Rapporteur
Dr. Angela Marinetti	Directrice de recherche CNRS Institut de Chimie des Substances Naturelles	Examinatrice
Pr. Serge Piettre	Professeur Université de Rouen Normandie	Directeur de thèse
Dr. Julien Brioche	Maître de conférences Université de Rouen Normandie	Encadrant

**Thèse dirigée par Serge PIETTRE, laboratoire COBRA**



*« Les difficultés ne visent pas à nuire mais à instruire »*

*Brian Tracy*



## Remerciements

Je tiens à remercier Jacques Lebreton, Professeur à l'université de Nantes, Isabelle Gillaizeau, Professeure à l'université d'Orléans, Nicolas Blanchard, Directeur de recherche CNRS au LIMA à Mulhouse et Angela Marinetti, Directrice de recherche CNRS de l'ICSN à Gif-sur-Yvette, pour m'avoir fait l'honneur d'évaluer ce travail.

J'adresse toute ma gratitude au Professeur Serge Piettre pour m'avoir accueilli au sein de son laboratoire, pour sa confiance et pour ses précieux conseils. Je remercie également le Docteur Julien Brioché pour son soutien, son aide et ses conseils avisés.

Cette thèse n'aurait pas eu la même saveur sans les personnes formidables côtoyées tout au long de cette aventure. Tout d'abord, je voudrais remercier Morgane, Chouquette, Florian, Wiwi et Kévin pour tous les merveilleux moments passés ensemble. Je n'oublie pas de remercier Mégane, Madeleine, Maxime, Steven, Alexis, Solène, Batoul, Dorra, Margot et Bambi pour votre bonne humeur quotidienne.

Un grand merci au plateau technique, en particulier Françoise, Tia<sup>2</sup> et Emilie pour leur gentillesse et leur aide précieuse facilitant grandement les analyses et les purifications.

Je suis sincèrement reconnaissant au Pr. Rebecca Deprez-Poulain de l'université de Lille, de m'avoir fait confiance dès la licence et de m'avoir donné goût à la recherche. Je remercierai également le Dr. Cyril Ronco de l'université de Sophia-Antipolis et le Dr. Ronan Gealageas de m'avoir épaulé au cours de ces stages et donné l'envie de poursuivre dans cette voie.

Une petite pensée pour mon cher Kévin. Malgré la distance, nous avons su rester proches et je souhaiterais juste te dire merci. Merci pour ta gentillesse, ton soutien, ta bonne humeur. Merci pour tout !

Si j'en suis là aujourd'hui, c'est grâce au soutien indéfectible de ma famille, mes parents, ma sœur et Marie. Il n'existe pas de mots pour vous remercier, ni pour traduire tout l'amour que je vous porte.



## Sommaire

Remerciements .....	5
Sommaire .....	7
Abréviations .....	13
Préambule .....	17
<b>Introduction .....</b>	<b>19</b>
<b>I. L'acide phosphorique 11 .....</b>	<b>24</b>
I.1. Exemples de molécules phosphatées essentielles au développement de la vie.....	24
I.1.a. Les phospholipides .....	24
I.1.b. Les nucléotides.....	25
I.1.b.i. L'adénosine triphosphate .....	25
I.1.b.ii. L'ADN et l'ARN .....	26
I.2. Analogues de phosphate .....	27
I.2.a. Dérivés de monoester et d'anhydride de phosphate .....	28
I.2.b. Analogues de diester de phosphate .....	29
I.2.b.i. Analogues de phosphate dépourvus de liaisons phosphore-carbone.....	30
I.2.b.ii. Analogues caractérisés par une liaison phosphore-carbone .....	31
I.2.b.iii. Analogues caractérisés par deux liaisons phosphore-carbone.....	32
I.2.b.iv. Analogues caractérisés par trois liaisons phosphore-carbone .....	33
I.2.b.v. Les dérivés cycliques .....	34
<b>II. Applications des oligonucléotides en chimie médicinale .....</b>	<b>35</b>
II.1. Les aptamères .....	35
II.2. Applications intracellulaires des ONs .....	35
II.2.a. ARN antisens .....	36
II.2.b. ARN interférent.....	37
II.2.c. Différences essentielles entre approches antisens et d'interférence .....	39
<b>III. Développement d'oligonucléotides antisens à visées thérapeutiques.....</b>	<b>40</b>
III.1. Synthèse et méthodes d'évaluation de la stabilité des ONMs.....	40
III.1.a. Synthèse automatisée d'oligonucléotides .....	40
III.1.b. Stabilité des duplex.....	42
III.1.c. Stabilité métabolique .....	43
III.2. Modifications des oligonucléotides.....	44
III.2.a. Généralités .....	44
III.2.b. Modifications de la base nucléique .....	46
III.2.b.i. Modification structurelle .....	46
III.2.b.ii. Introduction de substituant.....	47

III.2.b.iii. Remplacement de la base nucléique .....	48
III.2.c. Modifications du furanose .....	49
III.2.c.i. Modification en C2' .....	49
III.2.c.ii. Utilisation de furanose bloqué.....	50
III.2.c.iii. Remplacement du furanose par un analogue .....	51
III.2.d. Modification du pont phosphodiester naturel .....	51
III.2.d.i. Les isostères du pont phosphate naturel .....	52
III.2.d.ii. Phosphorothioates .....	53
III.2.d.iii. Phosphonates et phosphinates .....	54
<b>IV. Des phosphonates aux difluorophosphonates .....</b>	<b>58</b>
IV.1. Analogie structurelle .....	58
IV.2. Résistance du motif CF <sub>2</sub> à l'hydrolyse .....	59
IV.3. Apicophilicité.....	59
IV.4. Etude conformationnelle .....	60
<b>V. Synthèse du motif <i>gem</i>-difluorophosphate 150a et <i>gem</i>-difluorophosphinothioate 150b ..</b>	<b>61</b>
V.1. Prérequis .....	61
V.2. Développement synthétique menant au motif difluorophosphate .....	61
<b>Chapitre 1 .....</b>	<b>73</b>
<b>Rétrosynthèses.....</b>	<b>73</b>
<b>I. Voie A.....</b>	<b>76</b>
<b>II. Voie B.....</b>	<b>78</b>
<b>III. Voie C.....</b>	<b>79</b>
<b>IV. Voie D .....</b>	<b>80</b>
<b>V. Voie E.....</b>	<b>82</b>
<b>VI. Conclusions .....</b>	<b>83</b>
<b>Chapitre 2 .....</b>	<b>85</b>
<b>Synthèses des fragments-cibles à partir d'un substrat commun : l'allofuranose 221 .....</b>	<b>85</b>
<b>I. Synthèse de l'oléfine 185 .....</b>	<b>89</b>
I.1. Synthèse du fragment oléfinique <b>185a</b> à partir d'un nucléoside .....	89
I.1.a. Rétrosynthèse .....	89
I.1.b. Protection du diol C2'-C3' .....	90
I.1.c. Oxydation .....	91
I.1.d. Oléfination .....	92

I.1.e. Protection de l'uridine .....	92
I.2. Synthèse du fragment oléfinique <b>185b</b> à partir de l'allofuranose.....	93
I.2.a. Rétrosynthèse .....	93
I.2.b. Synthèse.....	94
<b>II. Synthèse du fragment iodé 186 .....</b>	<b>96</b>
II.1. Rétrosynthèse.....	96
II.2. Synthèse .....	96
II.3. Etude de synthèses plus convergentes vers <b>186</b> .....	97
II.3.a. Stratégie n° 1 .....	98
II.3.b. Stratégie n° 2.....	99
II.3.c. Insertion des autres bases nucléiques .....	100
II.4. Protection de la thymine .....	101
II.4.a. A partir du fragment <b>186a</b> .....	101
II.4.b. Stratégie hydrozirconation-iodation.....	102
II.4.c. Stratégie hydroboration-oxydation-iodation .....	103
<b>III. Conclusions .....</b>	<b>104</b>
<b>Chapitre 3 .....</b>	<b>107</b>
<b><i>Formation de la seconde liaison phosphore-carbone et du motif difluorophosphinothioate .....</i></b>	<b><i>107</i></b>
<b>I. Etude de la voie A .....</b>	<b>111</b>
<b>II. Etude de la voie B .....</b>	<b>113</b>
II.1. Formation de la seconde liaison P-C par addition radicalaire .....	113
II.2. Formation de la seconde liaison P-C par voie ionique : réaction de Michaelis-Becker .....	114
<b>III. Etude de la voie C : application sur un substrat modèle .....</b>	<b>117</b>
III.1. Etude sur un substrat modèle : rétrosynthèse .....	117
III.2. Formation de la première liaison P-C .....	118
III.3. Etude de la réactivité H-P=O versus H-P=S pour la formation de la seconde liaison P-C.....	118
<b>IV. Conclusions .....</b>	<b>122</b>
<b>Chapitre 4 .....</b>	<b>125</b>
<b><i>Insertion d'une seconde base nucléique .....</i></b>	<b><i>125</i></b>
<b>I. Introduction d'une seconde base nucléique.....</b>	<b>129</b>
I.1. A partir du furanose <b>197d</b> incorporant une thymine protégée.....	129
I.2. A partir de l'intermédiaire mixte thymine-furanose <b>197a</b> .....	131
<b>II. Phénomène de trans-N-glycosylation .....</b>	<b>136</b>

II.1. Etat de l'art .....	136
II.2. Solutions envisagées .....	142
II.2.a. Addition contrôlée .....	142
II.2.b. Ordre d'insertion.....	144
II.2.b.i. Utilisation d'un intermédiaire cytosine .....	144
II.2.b.ii. Utilisation d'un intermédiaire adénine.....	152
<b>III. Conclusions .....</b>	<b>153</b>
<b>Conclusions et perspectives.....</b>	<b>155</b>
<b>I. Conclusions .....</b>	<b>157</b>
<b>II. Perspectives .....</b>	<b>161</b>
II.1. Affranchissement du problème de trans- <i>N</i> -glycosylation.....	161
II.1.a. Protection du <i>H</i> -phosphinothioate <b>165d</b> .....	161
II.1.b. Insertion de la base avant la formation du motif difluorophosphate.....	162
II.1.c. Blocage de l'oxygène en position C2' .....	162
II.2. Fin de la synthèse .....	163
<b>Experimental part .....</b>	<b>167</b>
<b>I. General procedure.....</b>	<b>169</b>
<b>II. Characterizations chapter 2.....</b>	<b>170</b>
<b>III. Characterizations chapter 3.....</b>	<b>193</b>
<b>IV. Characterizations chapter 4.....</b>	<b>208</b>
<b>Annexes .....</b>	<b>223</b>
Structure Macugen.....	225
Structure Vitravène <sup>TM</sup> .....	226
Structure Kynamro <sup>®</sup> .....	227
Structure Spinraza <sup>®</sup> .....	228
Structure Eteplirsén.....	229
Structure Tegsedi .....	230
Exemples d'attribution des signaux des spectres RMN du composé mixte adénine-furanose 197c .....	231
Exemples d'attribution des signaux des spectres RMN du dinucléotide T <sup>-PMB</sup> -U 183a .....	237





## Abréviations

°C	: degrés Celsius
1,2-DCE	: 1,2-dichloroéthane
9-BBN	: 9-borabicyclo[3.3.1]nonane
A	: adénine
Ac	: acétate
AcOH	: acide acétique
ADN	: acide désoxyribonucléique
ADP	: adénosine diphosphate
AMPc	: adénosine monophosphate cyclique
ARN	: acide ribonucléique
ARNm	: acide ribonucléique messenger
ATCC	: acide trichloroisocyanurique
ATP	: adénosine triphosphate
B	: base nucléique
Bn	: benzyle
BSA	: bis(triméthylsilyl)acétamide
BuLi ou <i>n</i> -BuLi	: <i>n</i> -butyllithium
Bz	: benzoyle
C	: cytosine
c	: concentration
CAN	: nitrate de cérium et d'ammonium
CCM	: chromatographie sur couche mince
CSA	: acide camphorsulfonique
DBU	: 1,8-diazabicyclo[5.4.0]undéc-7-ène
DCC	: dicyclohexylcarbodiimide
DIPEA	: diisopropyléthylamine
DLP	: peroxyde de dilauroyle
DMAC	: diméthylacétamide
DMAP	: 4-diméthylaminopyridine
DMF	: diméthylformamide
DMSO	: diméthylsulfoxyde
DMT	: diméthoxytrityle
dsDNA	: double stranded deoxyribonucleic acid
E	: énergie
EDCI.HCl	: 1-éthyl-3-(3-diméthylaminopropyl)carbodiimide
EI	: ionisation électronique (Electron ionization)
éq/eq	: équivalent
ESI	: ionisation par électronébuliseur (Electrospray ionization)
Et	: éthyle
EtOH	: éthanol
FDA	: food and drug administration
FTIR	: spectroscopie infrarouge à transformée de Fourier (Fourier transformation infrared spectroscopy)

## Abréviations

g	: gramme
G	: guanine
GP	: groupement protecteur
h	: heure
HCV	: virus de l'hépatite C
HF	: fluorure d'hydrogène
HMPT	: tris(diméthylamino)phosphine
HOBt	: hydroxybenzotriazole
HOMO	: highest occupied molecular orbital
HRMS	: spectrométrie de masse haute résolution (High-resolution mass spectrometry)
Hz	: Hertz
IBX	: acide 2-iodobenzoïque
<i>J</i>	: constante de couplage
KHMDS	: bis(triméthylsilyl)amide de potassium
KOH	: hydroxyde de potassium
LA	: Lewis acid (acide de Lewis)
LNA	: locked nucleic acid
LUMO	: lowest unoccupied molecular orbital
M	: molaire
<i>m</i> -CPBA	: acide <i>mé</i> ta-chloroperbenzoïque
MeCN	: acétonitrile
MeOH	: méthanol
mg	: milligramme
MHz	: mégahertz
min	: minute
mL	: millilitre
μL	: microlitre
mm	: millimètre
μm	: micromètre
mmol	: millimole
MMT	: monométhoxytrityle
M.p. ou T <sub>m</sub>	: température de fusion (Melting point)
MOE	: méthoxyéthyle
MSNT	: 1-(2-mésitylènesulfonyl)-3-nitro-1 <i>H</i> -1,2,4-triazole
NaOH	: hydroxyde de sodium
nm	: nanomètre
ON	: oligonucléotide
ONM	: oligonucléotide modifié
P-C	: phosphore-carbone
<i>p</i> -ClBz	: <i>para</i> -chlorobenzoyl
pH	: potentiel hydrogène
Ph	: phényle
P <sub>i</sub>	: phosphore inorganique
PMB	: <i>para</i> -méthoxybenzyle
P-O	: phosphore-oxygène

## Abréviations

ppm	: partie par million
py	: pyridine
Rdt	: rendement
<i>R<sub>f</sub></i>	: rapport frontal
RISC	: RNA-induced silencing complex
RMN	: résonance magnétique nucléaire
SN <sub>2</sub>	: substitution nucléophile de type 2
SOMO	: single occupied molecular orbital
ssDNA	: single stranded deoxyribonucleic acid
T	: thymine
t.a.	: température ambiante
TBAF	: fluorure de tétra- <i>n</i> -butylammonium
TBDMS	: <i>tert</i> -butyldiméthylsilyle
TBDPS	: <i>tert</i> -butyldiphénylsilyle
TBPP	: peroxyphosphate de <i>tert</i> -butyle
TBTU	: 2-(1 <i>H</i> -benzotriazole-1-yl)-1,1,3,3-tétraméthylammonium tétrafluoroborate
TEBAC	: chlorure de benzyltriéthylammonium
Temp.	: température
TEMPO	: (2,2,6,6-tétraméthylpipéridin-1-yl)oxyle
THF	: tétrahydrofurane
TLC	: Thin layer chromatography (chromatographie sur couche mince)
TMS	: triméthylsilyle
TMSOTf	: triflate de triméthylsilyle
Tps	: temps
U	: uracile
UPLC-MS	: ultra performance liquid chromatography – mass spectroscopy
UV	: ultraviolet
VIH	: virus de l'immunodéficience humaine
v/v	: volume/volume



## Préambule

Ce projet de thèse a été réalisé au sein du laboratoire COBRA, dans l'équipe « synthèse de biomolécules fluorées ». Depuis de nombreuses années, le thème de la chimie des phosphono(thio)- et phosphino(thio)-nucléosides fluorés est étudié par l'équipe du Pr. Serge Piettre. Le motif difluorophosphate a ainsi été ciblé dans l'optique de développer un nouvel analogue du pont phosphodiester naturel présent dans les chaînes d'acides nucléiques.

Grâce aux travaux menés précédemment, une première stratégie synthétique a été mise au point conduisant à la formation d'un homodinucléotide (thymine-thymine) comportant les motifs difluorophosphate et difluorophosphinothioate. Dans la continuité de ces travaux, ce projet de thèse a été fondé sur le développement d'une seconde stratégie menant à un hétérodinucléotide (ayant deux bases nucléiques différentes) caractérisé par la présence des motifs difluorophosphate et difluorophosphinothioate.

Le manuscrit se compose d'une introduction centrée sur le rôle et le développement d'analogues de l'acide phosphorique, le motif difluorophosphate et les précédents travaux du laboratoire. Par la suite, le premier chapitre présentera les différentes rétrosynthèses envisagées menant à un hétérodinucléotide. Dans un second temps, les stratégies synthétiques et les problèmes rencontrés lors de la formation du dinucléotide-cible seront abordés dans les chapitres 2, 3 et 4. Enfin, la conclusion de ces travaux et les perspectives envisagées seront dressées.



# Introduction



# Sommaire

<b>I. L'acide phosphorique 11</b> .....	<b>24</b>
<b>I.1. Exemples de molécules phosphatées essentielles au développement de la vie</b> .....	<b>24</b>
I.1.a. Les phospholipides .....	24
I.1.b. Les nucléotides .....	25
I.1.b.i. L'adénosine triphosphate .....	25
I.1.b.ii. L'ADN et l'ARN.....	26
<b>I.2. Analogues de phosphate</b> .....	<b>27</b>
I.2.a. Dérivés de monoester et d'anhydride de phosphate .....	28
I.2.b. Analogues de diester de phosphate .....	29
I.2.b.i. Analogues de phosphate dépourvus de liaisons phosphore-carbone .....	30
I.2.b.ii. Analogues caractérisés par une liaison phosphore-carbone .....	31
I.2.b.iii. Analogues caractérisés par deux liaisons phosphore-carbone .....	32
I.2.b.iv. Analogues caractérisés par trois liaisons phosphore-carbone.....	33
I.2.b.v. Les dérivés cycliques .....	34
<b>II. Applications des oligonucléotides en chimie médicinale</b> .....	<b>35</b>
<b>II.1. Les aptamères</b> .....	<b>35</b>
<b>II.2. Applications intracellulaires des ONs</b> .....	<b>35</b>
II.2.a. ARN antisens.....	36
II.2.b. ARN interférent .....	37
II.2.c. Différences essentielles entre approches antisens et d'interférence .....	39
<b>III. Développement d'oligonucléotides antisens à visées thérapeutiques</b> .....	<b>40</b>
<b>III.1. Synthèse et méthodes d'évaluation de la stabilité des ONMs</b> .....	<b>40</b>
III.1.a. Synthèse automatisée d'oligonucléotides.....	40
III.1.b. Stabilité des duplex .....	42
III.1.c. Stabilité métabolique.....	43
<b>III.2. Modifications des oligonucléotides</b> .....	<b>44</b>
III.2.a. Généralités .....	44
III.2.b. Modifications de la base nucléique .....	46
III.2.b.i. Modification structurelle .....	46
III.2.b.ii. Introduction de substituant .....	47
III.2.b.iii. Remplacement de la base nucléique .....	48
III.2.c. Modifications du furanose.....	49
III.2.c.i. Modification en C2' .....	49

III.2.c.ii. Utilisation de furanose bloqué .....	50
III.2.c.iii. Remplacement du furanose par un analogue .....	51
III.2.d. Modification du pont phosphodiester naturel .....	51
III.2.d.i. Les isostères du pont phosphate naturel .....	52
III.2.d.ii. Phosphorothioates .....	53
III.2.d.iii. Phosphonates et phosphinates .....	54
<b>IV. Des phosphonates aux difluorophosphonates .....</b>	<b>58</b>
IV.1. Analogie structurelle.....	58
IV.2. Résistance du motif CF <sub>2</sub> à l'hydrolyse.....	59
IV.3. Apicophilicité.....	59
IV.4. Etude conformationnelle.....	60
<b>V. Synthèse du motif gem-difluorophosphate 150a et gem-difluorophosphinothioate 150b.....</b>	<b>61</b>
V.1. Prérequis.....	61
V.2. Développement synthétique menant au motif difluorophosphate .....	61

De nombreuses fonctions chimiques sont indispensables au développement de la vie. Dans le corps humain (Figure 1), la fonction amide **1** permet l'assemblage d'acides aminés pour former les protéines. Ces biomolécules assurent des fonctions variées telles que, par exemple, la catalyse de réactions biochimiques et la défense contre les agressions bactériologiques. Dans le monde vivant, la fonction ester **2** est également importante puisque celle-ci permet le stockage des acides gras ayant notamment un rôle métabolique et structurel.<sup>1</sup> En effet, les acides gras se combinent aux fonctions alcool du glycérol **5** pour former les triglycérides **3**, qui sont une source d'énergie importante pour le corps humain. Les acides gras servent également à la synthèse d'autres lipides comme les phospholipides, dont l'utilité sera discutée ultérieurement. La fonction alcool **4** est, quant à elle, présente dans les lipides (glycérol **5**, cholestérol **6**, etc...) mais également dans des biomolécules essentielles à tout être vivant, les glucides **7**.<sup>2</sup> Celle-ci est également impliquée dans la formation de biomolécules plus complexes. Afin de former le lien phosphate **9** des acides nucléiques **10**, les fonctions alcool des nucléosides **8** s'associent à une molécule essentielle à la vie, l'acide phosphorique **11**, qui fera l'objet d'une étude bibliographique plus approfondie dans la partie suivante.

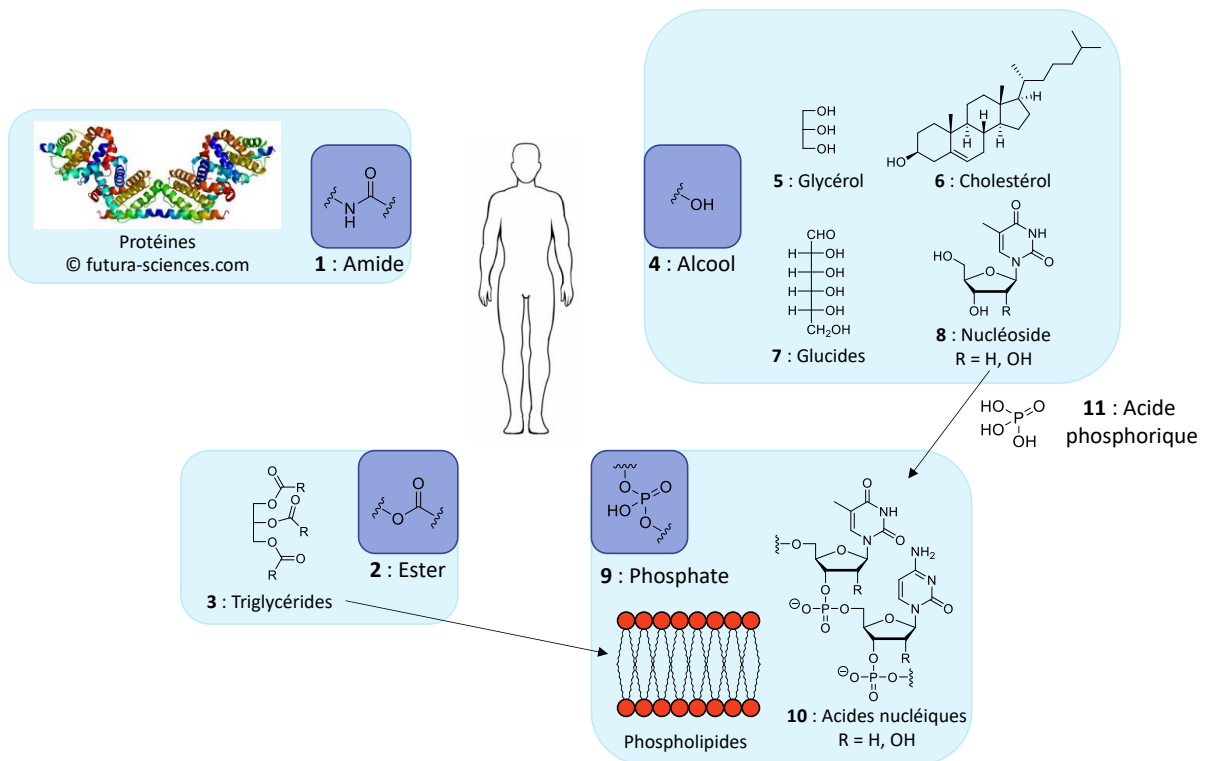


Figure 1 : Fonctions essentielles pour le fonctionnement du corps humain

<sup>1</sup> Moghadasian, M. H. and Shahidi, F. *International Encyclopedia of Public Health (Second Edition)* **2017**, 114-122.

<sup>2</sup> Blanco, A. and Blanco, G. *Medical Biochemistry* **2017**, 73-97.

## I. L'acide phosphorique **11**

Dans le corps humain, l'acide phosphorique **11** se trouve sous sa forme phosphate, monoester de phosphate **12** ou diester de phosphate **9**, entre-autres, et est à l'origine de nombreuses biomolécules essentielles au développement de la vie (Figure 2).<sup>3</sup>

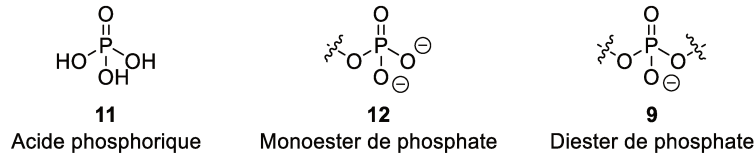


Figure 2 : Esters de phosphate

### I.1. Exemples de molécules phosphatées essentielles au développement de la vie

#### I.1.a. Les phospholipides

Les phospholipides sont des composants essentiels de la bicouche lipidique des membranes cellulaires. Ceux-ci sont composés de deux acides gras liés à un **glycérol phosphorylé** en position 3 et peuvent être nommés selon le substituant ( $R_3$ ) présent sur le phosphate (Figure 3). Par exemple, si ce substituant est une sérine, ce phospholipide sera une phosphatidylsérine **14**, alors que si  $R_3$  est une choline, celui-ci se nommera phosphatidylcholine **15**.

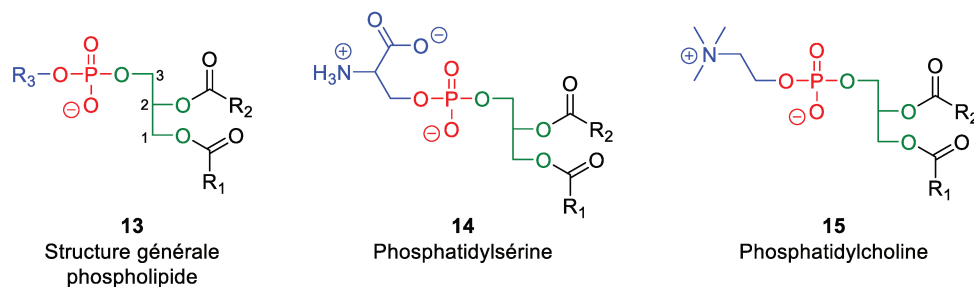


Figure 3 : Structures de phospholipides

Les phospholipides sont composés d'une partie hydrophobe (chaînes alkyles des acides gras) et d'une extrémité hydrophile (groupement phosphate). En milieu aqueux, les phospholipides s'agencent de manière à orienter les queues hydrophobes d'un côté et les têtes hydrophiles de l'autre, ce qui forme une monocouche. Dans le monde vivant, deux monocouches s'associent entre-elles pour former la bicouche lipidique qui constitue les membranes cellulaires (Figure 4).

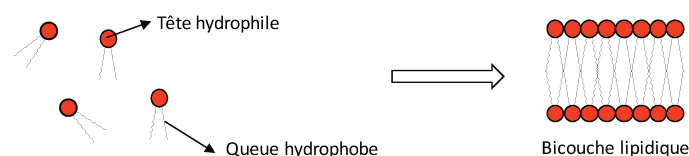


Figure 4 : Formation de la bicouche lipidique

<sup>3</sup> Westheimer, F. *Science* **1987**, 235, 1173-1178.

### I.1.b. Les nucléotides

Un nucléoside est composé d'une **base nucléique** et d'un **pentose**, tandis qu'un nucléotide est la combinaison d'un nucléoside avec un à trois groupes **phosphates** (Figure 5). Les nucléotides ont un rôle crucial dans l'organisme, étant notamment impliqués dans la production énergétique du corps, ainsi que dans la signalisation cellulaire et dans l'information génétique.

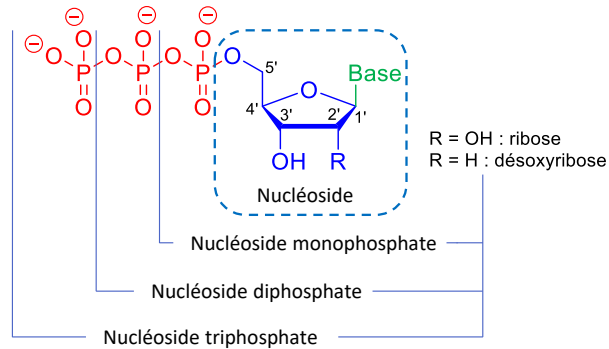


Figure 5 : Nomenclature

#### I.1.b.i. L'adénosine triphosphate

L'adénosine triphosphate (ATP) **16** est un nucléotide triphosphate portant une adénine comme base nucléique (Figure 6). L'hydrolyse de l'ATP libère une grande quantité d'énergie et génère de l'adénosine diphosphate **18** (ADP) et un phosphate inorganique  $P_i$  **17**. Cette énergie est ensuite utilisée pour la locomotion, pour la réalisation des réactions chimiques du métabolisme ou bien pour la division cellulaire. Par la suite, l'ADP se recombine avec un phosphate grâce à l'énergie dégagée par le catabolisme et reforme ainsi une molécule d'ATP.

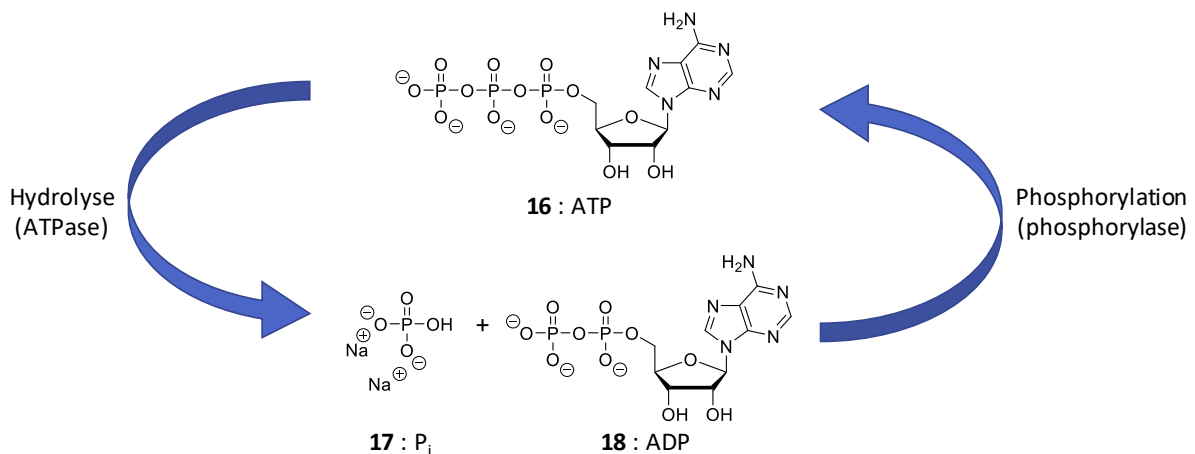


Figure 6 : Cycle de l'ATP

L'ATP **16** a également un rôle de précurseur dans des voies de signalisations cellulaires. En effet, celui-ci est le substrat de l'adénylate cyclase, qui produit l'adénosine monophosphate

cyclique **19** (AMPc) (Schéma 1). L'AMPc est un second messenger impliqué, par exemple, dans l'activation d'enzymes permettant la phosphorylation de protéines spécifiques.<sup>4</sup>

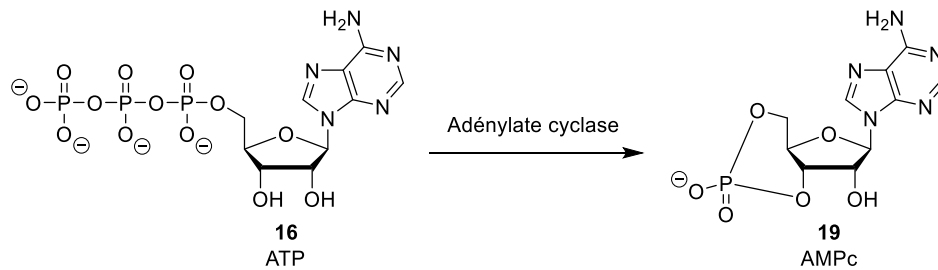


Schéma 1 : Action de l'adénylate cyclase sur l'ATP

L'ATP est également impliquée dans la synthèse de polynucléotides cruciaux pour tout être vivant, à savoir l'ADN et l'ARN.

### 1.1.b.ii. L'ADN et l'ARN

Lorsque des unités nucléosidiques sont reliées entre-elles par des ponts phosphodiester, celles-ci forment l'acide désoxyribonucléique **20** (ADN) et l'acide ribonucléique **21** (ARN) (Figure 7). Les unités nucléosidiques de l'ADN sont composées de 2'-désoxy-D-ribose lié par une liaison osidique à des bases nucléiques (thymine **22**, cytosine **23**, adénine **26** et guanine **27**). Tandis que dans l'ARN, le pentose est le D-ribose lié à des bases nucléiques identiques à l'ADN sauf pour la thymine qui est remplacée par l'uracile **24**. La numérotation de chaque atome des bases pyrimidiques et puriques est représentée sur les noyaux pyrimidine **25** et purine **28**.

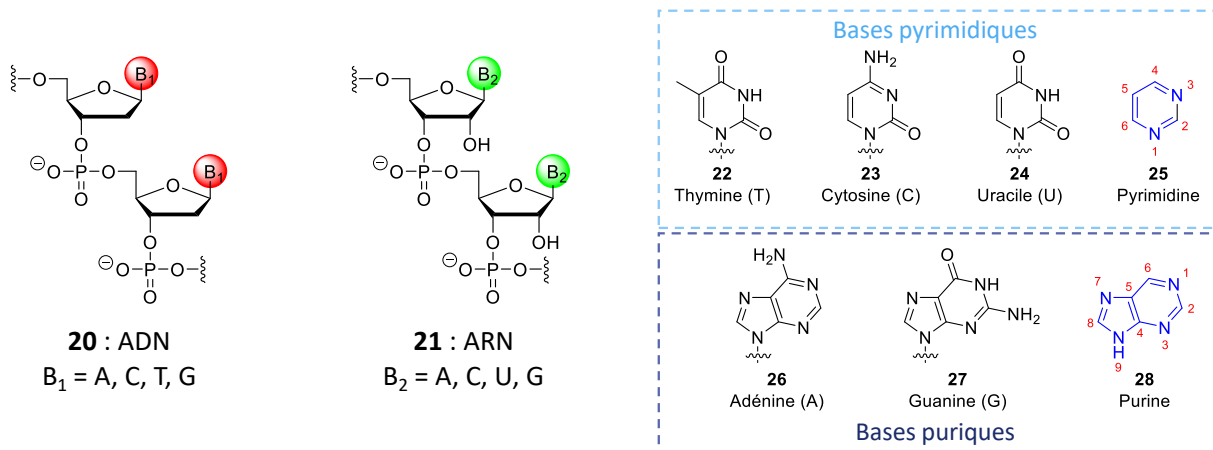


Figure 7 : Structure de l'ADN et l'ARN

L'ADN contient l'information génétique de chaque individu et adopte une structure en double hélice, formée de deux chaînes complémentaires de nucléotides liées entre-elles par des liaisons hydrogène selon l'hybridation de Watson et Crick.<sup>5</sup> L'appariement des bases pyrimidiques avec les bases puriques

<sup>4</sup> Kuszak, A. J. and Sunahara, R. K. *Handbook of Cell Signaling (Second Edition)* **2010**, 1389-1397.

<sup>5</sup> Watson, J. D. and Crick, F. H. C. *Nature* **1953**, *171*, 737-738.

par liaisons hydrogène entre deux brins d'ADN engendrent la formation de la double hélice. La thymine et l'adénine s'apparient au moyen de deux liaisons hydrogène tandis que la cytosine et la guanine s'apparient grâce à trois liaisons hydrogène (Figure 8). L'information génétique portée par une séquence d'ADN est transcrite, dans le sens de l'extrémité 5' vers l'extrémité 3', en chaîne complémentaire d'ARN messager qui est ensuite traduite en protéines à l'aide de ribosomes.

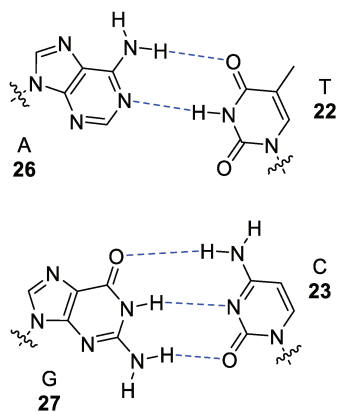


Figure 8 : Appariement des bases nucléiques

La fonction phosphate est donc très présente dans l'organisme et est impliquée dans de nombreux mécanismes biologiques. Vu l'importance de cette fonction, les chimistes médicaux s'y sont intéressés dans le but de développer de nouveaux médicaments et outils biologiques. Cependant, la liaison phosphore-oxygène est facilement hydrolysable en milieu biologique sous l'action d'enzymes. Afin d'augmenter la stabilité métabolique de cette fonction, de nombreux analogues<sup>6</sup> et isostères<sup>7</sup> du groupement phosphate ont donc été développés. Dans la partie suivante, des exemples d'analogues de phosphate seront présentés tandis que les isostères seront abordés ultérieurement.

## 1.2. Analogues de phosphate

La synthèse de nombreux analogues de phosphate est décrite dans la littérature depuis parfois plus d'un siècle. L'impact de l'incorporation de ces analogues de phosphate sur l'activité et la stabilité des biomolécules a été étudié intensivement au cours des dernières décennies. Dans cette partie, il ne s'agit pas d'établir une liste exhaustive de tous les analogues de phosphates utilisés, mais de présenter la diversité de ces modifications dans des applications médicales. Après avoir présenté quelques dérivés de monoester et d'anhydride de phosphate, les modifications évoquées se concentreront sur les analogues de diester de phosphate.

<sup>6</sup> Analogue : Fonction possédant une structure commune. Dans notre cas, chaque analogue sera centré sur un atome de phosphore.

<sup>7</sup> Isostère : Fonction mimant la fonction initiale tout en ayant une structure différente.

I.2.a. Dérivés de monoester et d'anhydride de phosphate

Il existe de nombreux analogues de monoester de phosphate.<sup>8</sup> De Clercq a développé des analogues non-hydrolysables de phosphate, en remplaçant l'oxygène estérifié par un motif CH<sub>2</sub> (phosphonate **29**, Figure 9), dans le but d'évaluer leur activité antivirale.<sup>9</sup> Ainsi, les composés à base d'adénine **30** et de cytosine **31** ont présentés de bonnes activités antivirales contre le virus de l'Herpès et le cytomégalovirus. Par la suite, cette étude a été complétée par les travaux de Kim où des phosphonates comportant des guanines ont été synthétisés (**32-35**) afin d'inhiber le virus de l'Herpès.<sup>10</sup> Néanmoins, aucune des structures **32-35** ne s'est révélée être active sur le virus.

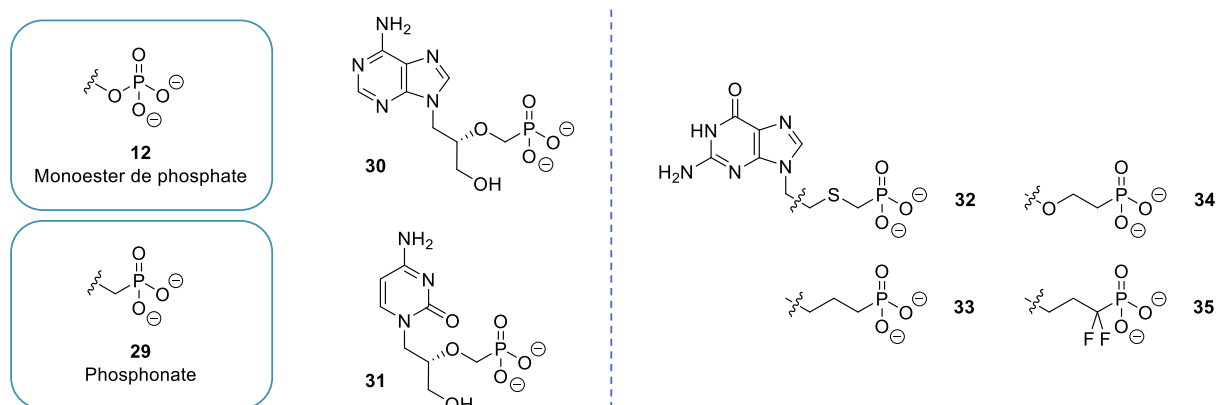


Figure 9 : Analogues de monoester de phosphate

Les études antivirales de De Clercq ont abouti au développement d'un médicament, le ténofovir **36**. En effet, cette molécule a montré une bonne activité sur le virus de l'immunodéficience humaine (VIH).<sup>11</sup> Le VIH est un rétrovirus agissant sur le système immunitaire en prenant le contrôle des globules blancs, coordonnant la réponse immunitaire lorsqu'un virus est détecté par l'organisme (lymphocytes CD4), en y intégrant son matériel génétique. L'ARN viral est transcrit en ADN grâce à la transcriptase inverse (enzyme virale) puis cet ADN est intégré au génome du lymphocyte CD4. Par la suite, ce nouveau génome est transcrit en ARN puis traduit en protéines par la cellule. La protéase, autre enzyme virale, intervient pour terminer la fabrication des protéines et permet la prolifération du virus dans l'organisme. A l'heure actuelle, les traitements utilisés sont des trithérapies composées de trois substances antivirales dont deux inhibiteurs nucléosidiques/nucléotidiques de la transcriptase inverse. C'est à ce titre que le ténofovir **36** est utilisé ainsi que l'emtricitabine **37** et l'AZT **38** (zidovudine), premier antiviral utilisé dans le traitement du VIH (Figure 10).

<sup>8</sup> Elliott, T. S.; Slowey, A.; Ye, Y. and Conway, S. J. *Med. Chem. Commun.* **2012**, 3, 735-751.

<sup>9</sup> Clercq, E. D.; Sakuma, T.; Baba, M.; Pauwels, R.; Balzarini, J.; Rosenberg, I. and Holy, A. *Antiviral Res.* **1987**, 8, 261-272.

<sup>10</sup> Kim, C. U.; Luh, B. Y.; Misco, P. F.; Bronson, J. J.; Hitchcock, M. J. M.; Ghazzouli, I. and Martin, J. C. *J. Med. Chem.* **1990**, 33, 1207-1213.

<sup>11</sup> Balzarini, J.; Holy, A.; Jindrich, J.; Naesens, L.; Snoeck, R.; Schols, D. and De Clercq, E. *Antimicrob. Agents Chemother.* **1993**, 37, 332-338.

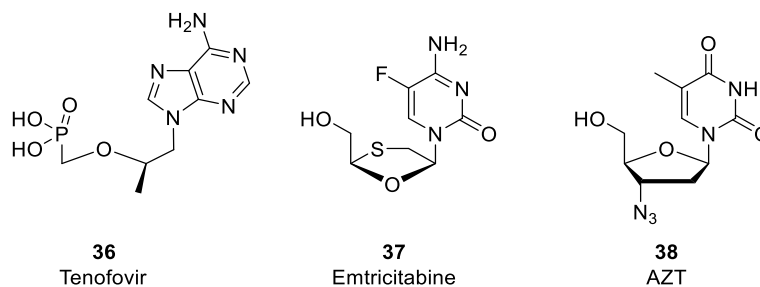


Figure 10 : Inhibiteurs nucléosidiques/nucléotidiques de la transcriptase inverse du VIH

Dans l'optique d'étudier les fonctions de l'ATP dans l'organisme, la synthèse d'analogues d'anhydride de phosphate **39** non-hydrolysables a été réalisée par plusieurs équipes. Ainsi, l'oxygène lié aux deux atomes de phosphore a été remplacé soit par un motif  $\text{CH}_2$ <sup>12</sup> (**40**), par un azote<sup>13</sup> (**41**) ou par des motifs fluorés ( $\text{CHF}$  **42** et  $\text{CF}_2$  **43**)<sup>14</sup> (Figure 11). Chacune de ces modifications a entraîné une augmentation de la stabilité métabolique des analogues synthétisés.

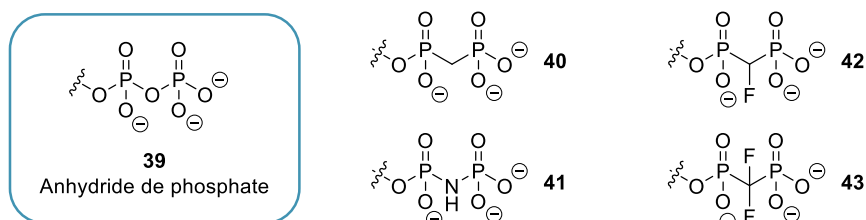


Figure 11 : Analogues d'anhydride de phosphate

### I.2.b. Analogues de diester de phosphate

Depuis plusieurs décennies, un grand nombre d'analogues de diesters de phosphate ont été synthétisés. Certains de ces dérivés sont présentés ci-dessous et ont été classés en cinq catégories (Figure 12) : les dérivés de phosphate **44**, les phosphonates et phosphonothioates **45**, les phosphinates et phosphinothioates **46**, les oxydes de phosphine **47** et les dérivés cycliques **48**. Chaque catégorie est illustrée ci-après par des exemples issus de la littérature.

<sup>12</sup> Myers, T. C.; Nakamura, K. and Flesher, J. W. *J. Am. Chem. Soc.* **1963**, *85*, 3292-3295.

<sup>13</sup> Yount, R. G.; Babcock, D.; Ballantyne, W. and Ojala, D. *Biochemistry* **1971**, *10*, 2484-2489.

<sup>14</sup> Blackburn, G. M.; Kent, D. E. and Kolkman, F. J. *Chem. Soc., Perkin Trans. 1* **1984**, 1119-1125.

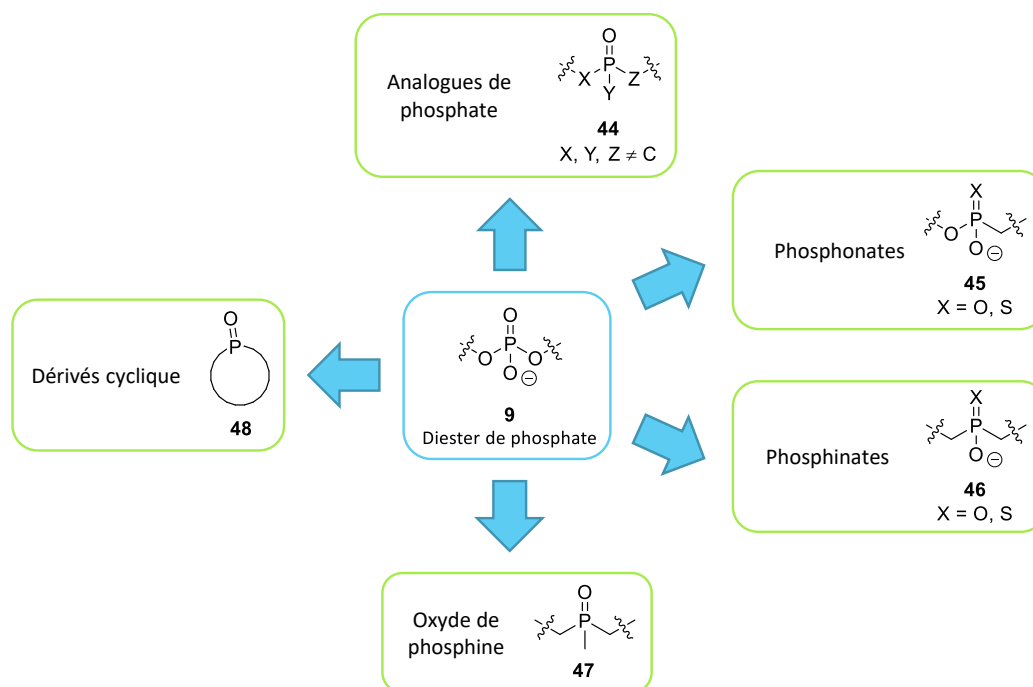


Figure 12 : Analogues de diester de phosphate

### 1.2.b.i. Analogues de phosphate dépourvus de liaisons phosphore-carbone

Dans cette catégorie sont regroupés tous les analogues au sein desquels un ou plusieurs atome(s) d'oxygène lié(s) au phosphore est (sont) substitué(s) par un (des) hétéroatome(s) ; des exemples sont présentés ci-dessous (Figure 13).

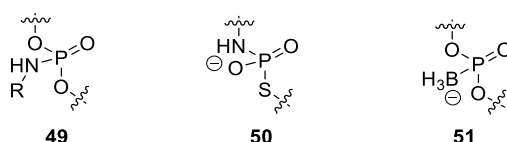


Figure 13 : Analogues de phosphate non-carboné

Le motif phosphoramidate **49** est trouvé dans l'un des médicaments utilisé pour traiter le virus de l'hépatite C, le sofosbuvir **52** (Figure 14). L'hépatite C est une inflammation du foie, plus ou moins détectable, provoquée par un virus.<sup>15</sup> Ce dernier pénètre dans les cellules hépatiques afin d'y libérer son ARN viral. Celui-ci sera traduit en une polyprotéine composée d'un ensemble d'unités protéiques, dont NS3A/4A, NS5A et NS5B, qui sont à l'origine de la réplication du virus et donc de la prolifération de ce dernier. Les traitements actuels sont des antiviraux dits « à action directe » ciblant les différentes unités protéiques. Le traitement le plus utilisé est constitué d'un inhibiteur du complexe NS5A (velpatasvir **53**) et d'un inhibiteur de la polymérase NS5B (sofosbuvir **52**), permettant une guérison virale efficace.<sup>16</sup>

<sup>15</sup> Pawlotsky, J.-M. *Med. Sci.* **2002**, *18*, 303-314.

<sup>16</sup> Gane, E. J.; Stedman, C. A.; Hyland, R. H.; Ding, X.; Svarovskaia, E.; Symonds, W. T.; Hinds, R. G. and Berrey, M. M. *N. Engl. J. Med.* **2013**, *368*, 34-44.

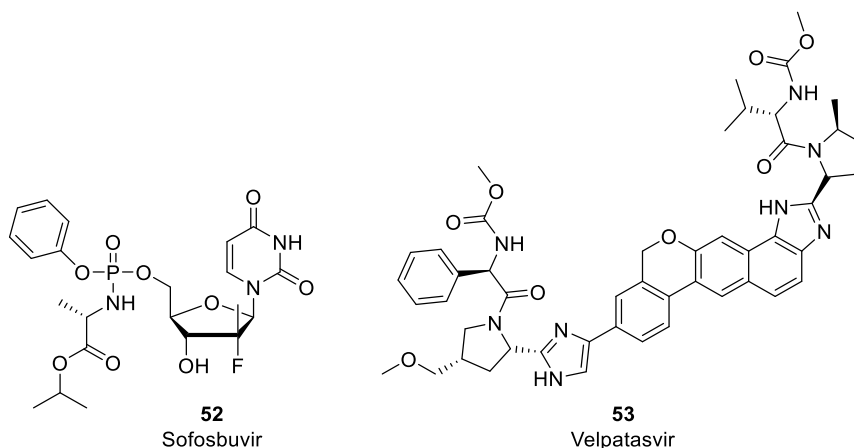


Figure 14 : Inhibiteurs du VHC

Le phosphorothioamidate **50** et le boranophosphate **51** ont été développés respectivement par Conway<sup>17</sup> et par Sood,<sup>18</sup> afin d'augmenter la stabilité du pont phosphodiester liant deux unités nucléosidiques. Les stabilités des deux groupements fonctionnels ont été testées à différents pH. Une dégradation du lien **50** a été observée à pH acide. Le pont **51** a présenté de meilleurs résultats puisqu'une résistance a été observée en conditions acide et basique.

#### *1.2.b.ii. Analogues caractérisés par une liaison phosphore-carbone*

Les phosphonates sont des dérivés possédant une liaison phosphore-carbone et trois liaisons phosphore-oxygène. Deux possibilités structurales existent, le méthylphosphonate **54** et le phosphonate **55** (Figure 15). D'autres analogues comportant toujours cette liaison P-C, mais dans lesquels un ou plusieurs atome(s) d'oxygène a (ont) été remplacé(s) par d'autres hétéroatomes (S, N principalement) ont également été décrit (par exemple : phosphonamide **56**,<sup>19</sup> phosphonothioamide **57**,<sup>20</sup> phosphonothioate **58**,<sup>21</sup> phosphonodithioate **59**,<sup>22</sup> phosphonotrithioate **60**<sup>22</sup>).

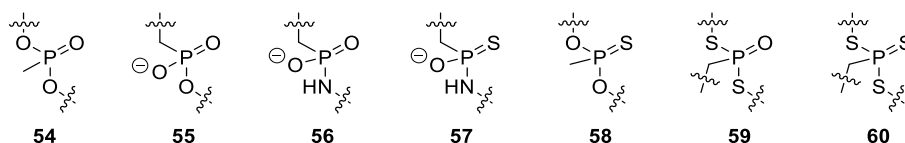


Figure 15 : Les phosphonates

Les analogues méthylphosphonates<sup>23</sup> **54** et méthylphosphonothioates **58** ont été développés dans le but d'augmenter la stabilité métabolique du pont phosphodiester naturel. Ces deux analogues

<sup>17</sup> Conway, L. P.; Mikkola, S.; O'Donoghue, A. C. and Hodgson, D. R. W. *Org. Biomol. Chem.* **2016**, *14*, 7361-7367.

<sup>18</sup> Sood, A.; Shaw, B. R. and Spielvogel, B. F. *J. Am. Chem. Soc.* **1990**, *112*, 9000-9001.

<sup>19</sup> Yoshino, K.; Goto, K.; Yoshiizumi, K.; Morita, T. and Tsukamoto, G. *J. Med. Chem.* **1990**, *33*, 2192-2196.

<sup>20</sup> Rushing, S. D. and Hammer, R. P. *J. Am. Chem. Soc.* **2001**, *123*, 4861-4862.

<sup>21</sup> Padmapriya, A. A. and Agrawal, S. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 761-764.

<sup>22</sup> Kim, S.; Mimikakis, J. L. and Roundhill, D. M. *Phosphorus, Sulfur Silicon Relat. Elem.* **1992**, *68*, 119-128.

<sup>23</sup> Miller, P. S.; Yano, J.; Yano, E.; Carroll, C.; Jayaraman, K. and Ts'o, P. O. P. *Biochemistry* **1979**, *18*, 5134-5143.

ont ainsi présenté une résistance accrue vis-à-vis de l'action des nucléases. Une introduction du lien **58** dans une chaîne polynucléotide a également été réalisée. Une déstabilisation du duplex résultant avec une chaîne d'ARN complémentaire a été observée.<sup>24</sup> Les phosphonates **55** seront développés plus en détail postérieurement (Introduction, III.2.d.iii).

Le développement d'inhibiteurs calciques a conduit à la synthèse du fostédil **61** présentant une bonne activité inhibitrice (Figure 16). Cependant, des effets d'anti-hypertension et anti-arythmique ont été observés.<sup>25</sup> Ainsi, Yoshino a développé des analogues comportant le motif phosphonamide **56** afin de pallier les effets secondaires du fostédil **61**.<sup>19</sup> Quelques exemples sont présentés ci-dessous (**62a-c**). L'amidophosphonate **62a** a montré une meilleure activité inhibitrice que le fostédil **61** et l'inhibiteur calcique de référence (diltiazem **63**). Cependant, aucun autre développement n'a été réalisé sur celui-ci signifiant potentiellement que les effets secondaires n'ont pu être éliminés.

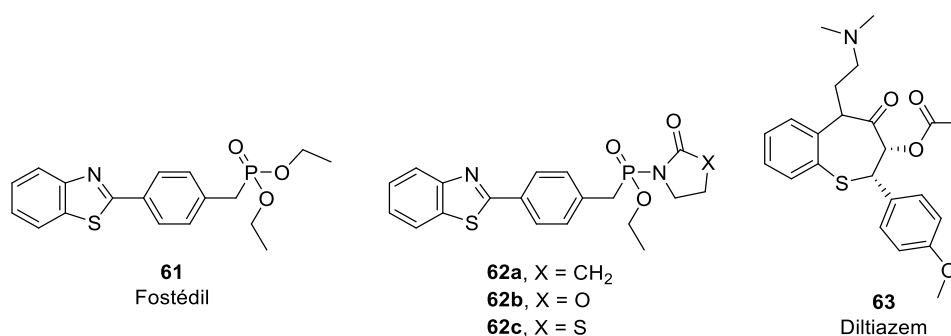


Figure 16 : Travaux de Yoshino

### 1.2.b.iii. Analogues caractérisés par deux liaisons phosphore-carbone

La fonction parentale est définie par la présence de deux liaisons phosphore-carbone et deux liaisons phosphore-oxygène. Cette catégorie est notamment illustrée par le motif phosphinate **64** mais également par d'autres analogues où les atomes d'oxygène ont été remplacés par un ou deux hétéroatome(s) (S, N) (par exemple : phosphinothioates **65**,<sup>26</sup> phosphinodithioates **66**<sup>27</sup> et phosphinamides **67**<sup>28</sup>) (Figure 17).

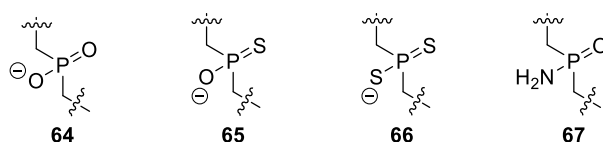


Figure 17 : Les phosphinates

<sup>24</sup> La notion de stabilité du duplex sera abordée plus en détail dans la partie III.1.b. Dans les grandes lignes, il s'agit d'étudier la force d'appariement entre un brin d'ARN ou d'ADN et une chaîne polynucléotidique modifiée.

<sup>25</sup> Pelc, L. R.; Gross, G. J. and Warltier, D. C. *Gen. Pharmac.* **1985**, *16*, 585-590.

<sup>26</sup> Sekine, M.; Matsuzaki, J.; Satoh, M. and Hata, T. *J. Org. Chem.* **1982**, *47*, 571-573.

<sup>27</sup> Navech, J.; Revel, M. and Kraemer, R. *Tetrahedron Lett.* **1985**, *26*, 207-210.

<sup>28</sup> Harger, M. J. P. and Westlake, S. *Tetrahedron* **1982**, *38*, 1511-1515.

A titre d'exemple, le motif phosphinate **64** a été étudié dans des composés développés par Acher afin de trouver de nouveaux agonistes de récepteurs métabotropiques du glutamate.<sup>29</sup> Cette recherche a permis récemment de développer les composés **68** à **70** ayant un potentiel thérapeutique prometteur dans le cadre de la maladie de Parkinson, l'anxiété, la schizophrénie et la douleur (Figure 18).<sup>30</sup>

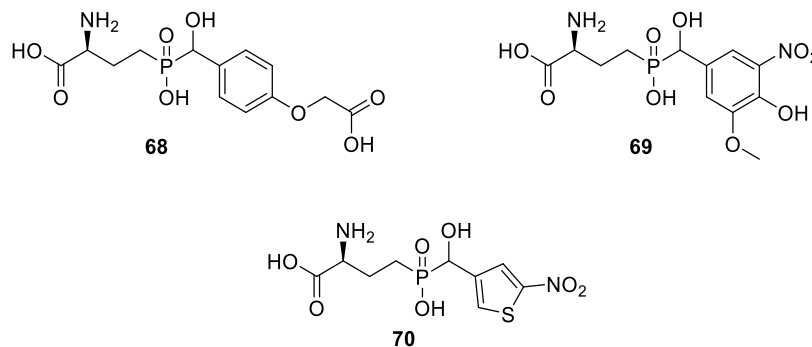


Figure 18 : Phosphinates développés par Acher

#### 1.2.b.iv. Analogues caractérisés par trois liaisons phosphore-carbone

Ces composés sont caractérisés par le remplacement des trois liaisons simple phosphore-oxygène par trois liaisons phosphore-carbone. Les dérivés d'oxyde de phosphine **47** ont notamment été introduits par Collingwood afin de développer un nouveau type de liaison entre deux nucléosides et ainsi d'augmenter la stabilité du nouveau pont formé (Figure 19).<sup>31</sup> L'introduction du lien oxyde de phosphine **47**, dans une chaîne nucléosidique, a permis de stabiliser métaboliquement la chaîne, mais a généré une déstabilisation du duplex formé.

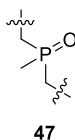


Figure 19 : Oxyde de phosphine

<sup>29</sup> Selvam, C.; Oueslati, N.; Lemasson, I. A.; Brabet, I.; Rigault, D.; Courtiol, T.; Cesarini, S.; Triballeau, N.; Bertrand, H.-O.; Goudet, C.; Pin, J.-P. and Acher, F. C. *J. Med. Chem.* **2010**, *53*, 2797-2813.

<sup>30</sup> Selvam, C.; Lemasson, I. A.; Brabet, I.; Oueslati, N.; Karaman, B.; Cabaye, A.; Tora, A. S.; Commare, B.; Courtiol, T.; Cesarini, S.; McCort-Tranchepain, I.; Rigault, D.; Mony, L.; Bessiron, T.; McLean, H.; Leroux, F. R.; Colobert, F.; Daniel, H.; Goupil-Lamy, A.; Bertrand, H.-O.; Goudet, C.; Pin, J.-P. and Acher, F. C. *J. Med. Chem.* **2018**, *61*, 1969-1989.

<sup>31</sup> Collingwood, S. P. and Taylor, R. J. *Synlett* **1998**, *09*, 283-285.

## 1.2.b.v. Les dérivés cycliques

Ce type de modification a été étudié dans le but de développer un analogue stable du pont phosphodiester naturel. Par exemple, les travaux de Dupouy<sup>32</sup> (**71**) et Maturano<sup>33</sup> (**72**) ont été réalisés afin de lier des unités nucléosidiques entre-elles et de mimer ainsi la conformation structurelle des nucléotides dans un brin d'ADN ou d'ARN. Dans le cas du phosphate cyclique à cinq atomes **71**, seule la synthèse a été décrite et aucune information sur la stabilité de celui-ci n'a été rapportée. En revanche, la synthèse du motif cyclique à 6 atomes **72** et son incorporation dans une chaîne nucléotidique ont été réalisées. La présence de ce motif **72** a permis d'augmenter à la fois la stabilité métabolique et celle du duplex formé avec un brin complémentaire d'acide nucléique.<sup>34</sup>

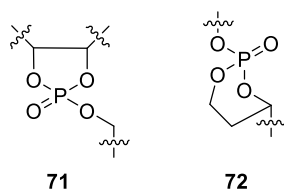


Figure 20 : Les dérivés cycliques

Les modifications du motif phosphate présentées ont été utilisées en chimie médicinale notamment dans le développement de médicaments antiviraux (ténofovir et sofosbuvir) et dans l'étude de nouveaux liens internucléosidiques. Ces liens ont aussi été introduits dans des chaînes polynucléotidiques (oligonucléotides) dont les applications médicales ont connu un véritable essor au cours des dernières années.

<sup>32</sup> Dupouy, C.; Lavedan, P. and Escudier, J.-M. *Eur. J. Org. Chem.* **2008**, 2008, 1285-1294.

<sup>33</sup> Maturano, M.; Catana, D.-A.; Lavedan, P.; Tarrat, N.; Saffon, N.; Payrastra, C. and Escudier, J.-M. *Eur. J. Org. Chem.* **2012**, 2012, 721-730.

<sup>34</sup> Catana, D.-A.; Renard, B.-L.; Maturano, M.; Payrastra, C.; Tarrat, N. and Escudier, J.-M. *J. Nucleic Acids* **2012**, 2012, 17.

## II. Applications des oligonucléotides en chimie médicinale

Un oligonucléotide (ON) est une chaîne de nucléotides de huit à plusieurs dizaines d'unités. La longueur d'un oligonucléotide est notée par le suffixe « mère » précédé du préfixe correspondant au nombre d'unités nucléotidiques. Par extension, le chiffre lui-même est parfois utilisé pour décrire l'oligonucléotide : ainsi un oligonucléotide comprenant huit unités nucléotidiques est un octamère ou un 8-mère. Certains d'entre eux ont été développés et modifiés à des fins thérapeutiques. Depuis la première synthèse d'un oligonucléotide rapportée par Khorana en 1972,<sup>35</sup> de nombreuses applications impliquant des ONs ont été développées telles que les aptamères, les oligonucléotides leurres, les antigènes, les ribozymes, l'ARN antisens et l'ARN interférent.

### II.1. Les aptamères

Tout d'abord, les aptamères sont des oligonucléotides possédant une grande affinité et spécificité pour des cibles moléculaires variées (acides aminés, protéines,...), et sélectionnés à partir d'une banque aléatoire de séquences.<sup>36</sup> Ceux-ci ont un large spectre d'applications et peuvent être utilisés comme outils de détection, de régulation ou thérapeutique.<sup>37</sup> En 2004, un aptamère thérapeutique a été développé contre la dégénérescence maculaire liée à l'âge et approuvé par la Food and Drug Administration (FDA), à savoir le Macugen<sup>38</sup> (28-mère).<sup>39</sup>

Les aptamères ont essentiellement des applications extracellulaires mais d'autres oligonucléotides présentent des activités intracellulaires.

### II.2. Applications intracellulaires des ONs

Premièrement, les oligonucléotides leurres (1) sont des courtes séquences d'ADN double brin ciblant, par exemple, les protéines responsables de la transcription (Figure 21). L'activité de cette protéine est donc inhibée et les gènes ciblés ne sont pas transcrits.<sup>40</sup>

En 1957, Felsenberg a montré l'existence de la formation d'une triple hélice dans laquelle deux chaînes polyuridine sont hybridées avec une chaîne polyadénine.<sup>41</sup> A la suite de cette découverte, le développement d'oligonucléotide dit « antigène » (2) a vu le jour. Ces chaînes polynucléotidiques se

---

<sup>35</sup> Khorana, H. G.; Agarwal, K. L.; Büchi, H.; Caruthers, M. H.; Gupta, N. K.; Klbppe, K.; Kumar, A.; Ohtsuka, E.; RajBhandary, U. L.; van de Sande, J. H.; Sgaramella, V.; Tebao, T.; Weber, H. and Yamada, T. *J. Mol. Biol.* **1972**, *72*, 209-217.

<sup>36</sup> Blind, M. and Blank, M. *Mol. Ther. Nucleic Acids* **2015**, *4*, e223.

<sup>37</sup> Chauveau, F.; Pestourie, C. and Tavitian, B. *Pathol. Biol. (Paris)* **2006**, *54*, 251-258.

<sup>38</sup> Structure disponible en annexes

<sup>39</sup> Ng, E. W. M.; Shima, D. T.; Calias, P.; Cunningham, E. T.; Guyer, D. R. and Adamis, A. P. *Nat. Rev. Drug Discovery* **2006**, *5*, 123-132.

<sup>40</sup> Mann, M. J. and Dzau, V. J. *J. Clin. Invest.* **2000**, *106*, 1071-1075.

<sup>41</sup> (a) Felsenfeld, G.; Davies, D. R. and Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023-2024. (b) Rich, A. *Nature* **1958**, *181*, 521-525.

fixent sur la double hélice de l'ADN par des liaisons hydrogène de type Hoogsteen<sup>42</sup> et inhibent la transcription de l'ADN en ARNm et, par conséquent, la production de la protéine correspondante.<sup>43</sup> Cependant, la complexité du transport d'un oligonucléotide à l'intérieur du noyau de la cellule reste un frein au développement de médicaments antigènes.

Une dernière approche consiste à introduire, dans le cytoplasme des cellules, des ribozymes (3) qui sont des oligonucléotides catalysant le clivage d'ARNm spécifique. Cette activité catalytique a été mise en évidence par Cech<sup>44</sup> et Altman<sup>45</sup>, ce qui leur a valu le prix Nobel de chimie en 1989.

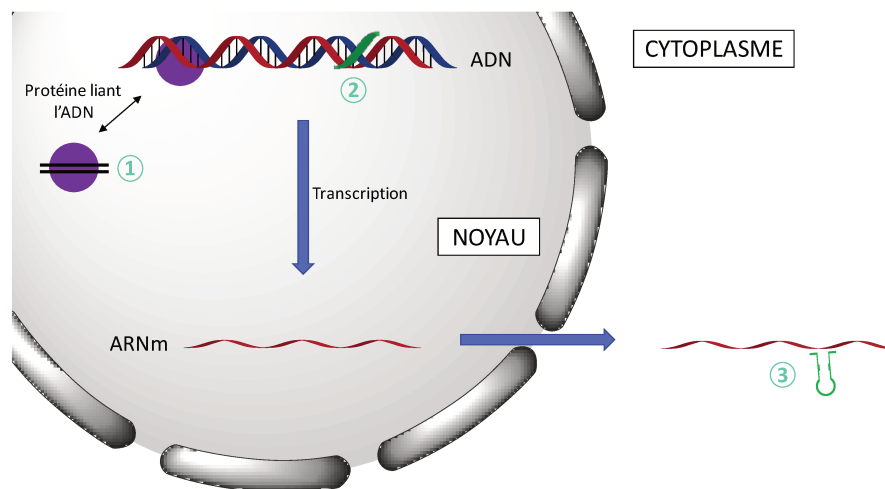


Figure 21 : Applications des oligonucléotides dans la cellule  
(1) Oligonucléotides leurres, (2) Approche antigène, (3) Ribozymes

Enfin, deux applications thérapeutiques des oligonucléotides ont connu un essor considérable au cours des dernières décennies : l'ARN antisens et l'ARN interférent.

### II.2.a. ARN antisens

En 1978, Zamecnik et Stephenson ont inventé le concept d'antisens en utilisant un tridécamère de type déoxyribonucléotide complémentaire de la partie terminale de la séquence du virus du sarcome de Rous afin d'inhiber sa répllication virale.<sup>46</sup> Dans une approche thérapeutique classique ciblant les enzymes, l'activité de la protéine, ayant un rôle primordial dans la pathologie que l'on souhaite traiter, est inhibée par l'action d'une petite molécule organique (principe actif) dans son site catalytique (Figure 22). Pour certaines maladies, les protéines membranaires (récepteurs) sont ciblées par l'action d'une substance permettant d'activer (agoniste) ou d'inactiver (antagoniste) une fonction

<sup>42</sup> Hoogsteen, K. *Acta Cryst.* **1963**, *16*, 907-916.

<sup>43</sup> Hélène, C.; Garestier, T.; Giovannangeli, C. and Sun, J.-S. *Gene Ther. Mol. Biol.* **1997**, *1*, 467-474.

<sup>44</sup> Grabowski, P. J.; Zaug, A. J. and Cech, T. R. *Cell* **1981**, *23*, 467-476.

<sup>45</sup> Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N. and Altman, S. *Cell* **1983**, *35*, 849-857.

<sup>46</sup> Zamecnik, P. C. and Stephenson, M. L. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 280-284.

cellulaire particulière. En comparaison, la stratégie antisens permet de moduler la production de protéine en appariant un oligonucléotide thérapeutique à sa séquence complémentaire d'ARNm de manière à empêcher la traduction de ce dernier. Sous l'action de la RNase H, la partie ARNm du duplex ON-ARNm formé est dégradée.

A ce jour, cette approche thérapeutique a conduit à la mise sur le marché de cinq oligonucléotides<sup>47</sup> et une trentaine de candidats sont en phase de tests cliniques.<sup>48</sup> En 1998, le Fomivirsen (Vitravene™, 21-mère) a été le premier médicament antisens approuvé par la FDA. Celui-ci est utilisé dans le traitement de la rétinite à cytomégalovirus chez les patients immunodéprimés. En 2013 a été approuvé, par la FDA, le Mipomersen (Kynamro®, 20-mère) dans le traitement d'hypercholestérolémie. Plus récemment, en 2016, deux autres médicaments ont été validés par la FDA, le Nusinersen (Spinraza®) et l'Eteplirsén (31-mère). L'Eteplirsén est prescrit dans le traitement de certaines mutations causant la dystrophie musculaire de Duchenne. Le Nusinersen (Spinraza®) est quant à lui utilisé dans le traitement de l'amyotrophie spinale. Le dernier oligonucléotide antisens thérapeutique, l'Inotersen (Tegsedi, 20-mère), a été approuvé en 2018, dans le traitement de la polyneuropathie de stade 1 ou de stade 2 chez les patients adultes atteints d'amylose à transthyrétine héréditaire.

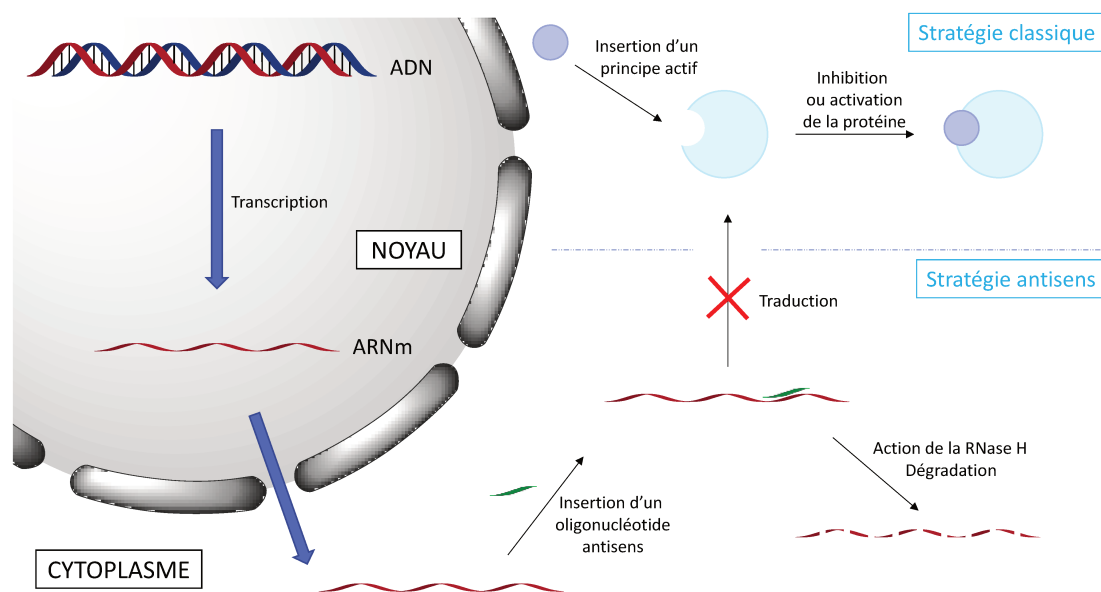


Figure 22 : Stratégie classique versus stratégie antisens

### II.2.b. ARN interférent

En 1990, Jorgensen fait une découverte fortuite en voulant renforcer la couleur des pétunias. En effet, l'introduction d'un transgène de chalcone synthase codant un pigment de la fleur génère des

<sup>47</sup> Structures disponibles en annexes

<sup>48</sup> Crooke, S. T.; Witztum, J. L.; Bennett, C. F. and Baker, B. F. *Cell Metab.* **2018**, *27*, 714-739.

fleurs partiellement ou totalement blanches. Il en déduit que le vecteur introduit a inhibé l'expression du gène naturel (Figure 23).<sup>49</sup>

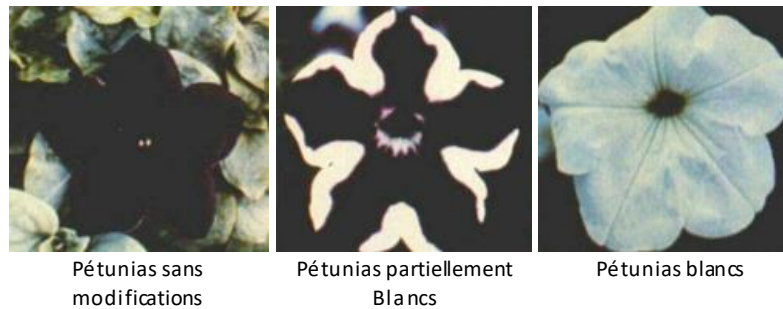


Figure 23 : Décoloration des pétunias ©1990 by American Society of Plant Biologists

A la fin des années 90, Fire et Mello ont montré qu'en introduisant de l'ARN double brin dans les cellules de nématode l'expression de protéine pouvait être réduite, découvrant par là le processus d'interférence.<sup>50</sup> Lorsqu'un ARN double brin est introduit dans la cellule, ce dernier, sous l'action de l'enzyme Dicer, est clivé en petits ARN interférent double brin de 22 nucléotides (Figure 24). Ces fragments sont reconnus par le complexe protéique RISC (RNA-Induced Silencing Complex). La protéine Argonaute-2 (Ago2), faisant partie du complexe protéique RISC, va alors cliver le brin dit passager entraînant l'activation du complexe. Le brin restant va servir de guide pour rechercher les ARNm complémentaire afin de les dégrader et empêcher la production de la protéine correspondante.

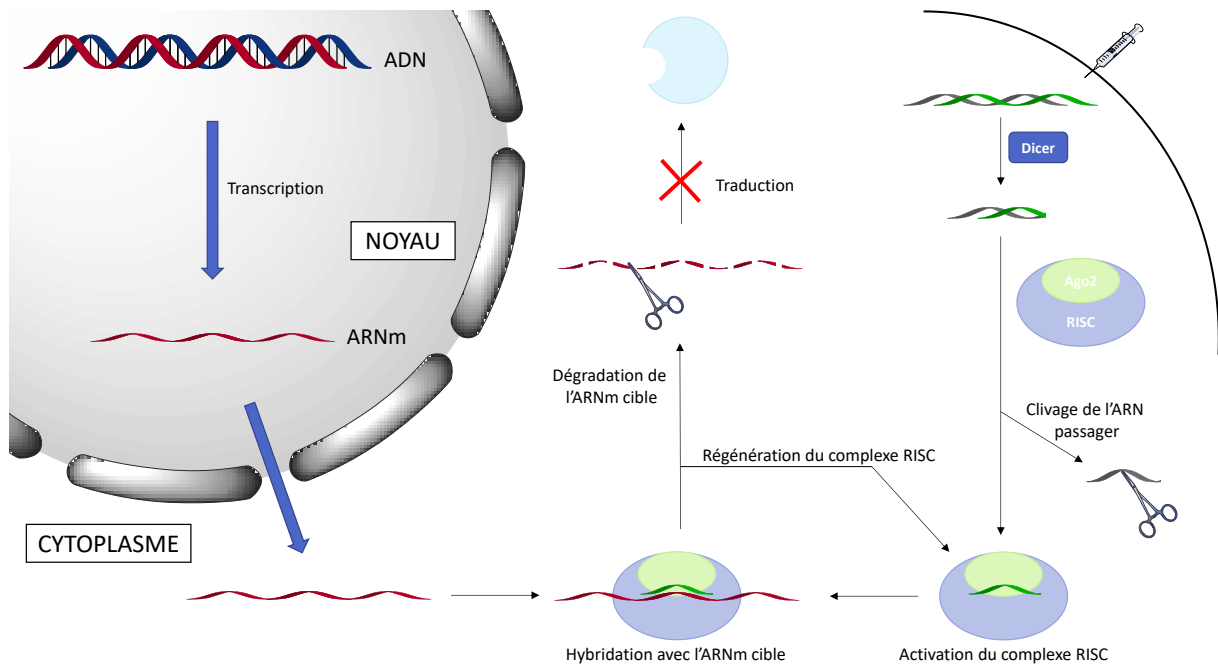


Figure 24 : ARN interférent

<sup>49</sup> Napoli, C.; Lemieux, C. and Jorgensen, R. *Plant Cell* **1990**, 2, 279-289.

<sup>50</sup> Fire, A.; Xu, S.; Montgomery, M. K.; Kostas, S. A.; Driver, S. E. and Mello, C. C. *Nature* **1998**, 391, 806.

A l'heure actuelle, l'ARN interférent s'est davantage imposé comme un outil permettant d'étudier la fonction d'un gène.<sup>51</sup>

Depuis 2018, un médicament interférent est présent sur le marché, le Onpatro® qui est utilisé pour le traitement de l'amylose héréditaire à transthyréline chez les adultes atteints de polyneuropathie de stade 1 ou 2 (Figure 25). De plus, une vingtaine de composés sont en phases d'essais cliniques.<sup>48</sup>

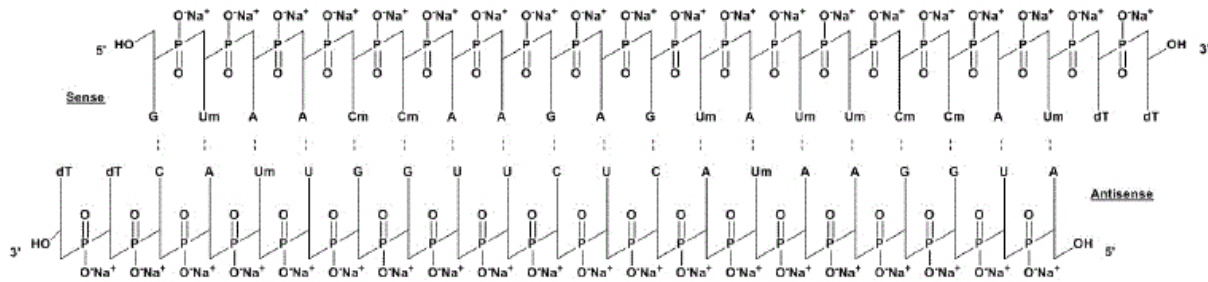


Figure 25 : Structure de l'Onpatro® © The New England Journal of Medicine  
(nature du pont phosphoré : phosphate)

### II.2.c. Différences essentielles entre approches antisens et d'interférence

La différence majeure entre les deux approches repose sur l'utilisation d'un oligonucléotide simple brin pour la stratégie antisens et d'un oligonucléotide double brin pour l'approche interférence. Les problèmes de passage membranaire sont donc accrus dans le cas de la stratégie interférente expliquant ainsi le fait qu'il n'y ait qu'un seul médicament interférent sur le marché à l'heure actuelle.

La seconde différence concerne l'enzyme spécifique utilisée pour dégrader le brin d'ARNm-cible. Dans le cas de l'approche antisens, la RNase H est activée pour couper le brin d'ARN. Dans la stratégie d'interférence, la chaîne nucléique est clivée par l'action du complexe protéique RISC.

De manière générale et quelle que soit l'approche utilisée, des modifications doivent être apportées aux oligonucléotides pour permettre leur utilisation à des fins thérapeutiques.

<sup>51</sup> Davidson, B. L. and McCray Jr, P. B. *Nature Reviews Genetics* **2011**, 12, 329.

### III. Développement d'oligonucléotides antisens à visées thérapeutiques

Pour être utilisés comme médicaments, les ONs monoplexes antisens doivent respecter certaines conditions. Tout d'abord, le duplex formé (ON-ARN) doit être suffisamment stable dans des conditions physiologiques. De plus, l'interaction entre l'oligonucléotide et le brin d'ARN naturel doit être spécifique et sélective. Enfin, l'oligonucléotide modifié (ONM) doit être résistant métaboliquement et être capable de traverser les membranes cellulaires.<sup>52</sup> Par la suite, les méthodes de détermination de la stabilité d'un duplex seront abordées mais, avant cela, la synthèse automatisée de ces ONMs sera présentée.

#### III.1. Synthèse et méthodes d'évaluation de la stabilité des ONMs

##### III.1.a. Synthèse automatisée d'oligonucléotides

Comme pour la synthèse peptique, la synthèse oligonucléotidique est réalisée sur support solide impliquant cycles de protections, déprotections, couplages et lavages. A l'heure actuelle, la méthode la plus utilisée est la synthèse par la méthode phosphoramidite (Schéma 2). Celle-ci s'effectue dans le sens inverse de la biosynthèse des ADN, c'est-à-dire dans le sens 3' vers 5'.

La synthèse débute par le clivage du groupement diméthoxytrityle (DMT) du nucléoside **73**, fixé à la résine, en présence d'acide trichloroacétique (étape 1). Le nucléoside suivant **76** est ajouté sous la forme d'un phosphoramidite protégé. Après activation par protonation du groupement diisopropylamino à l'aide de tétrazole **75**, le couplage entre **76** et l'hydroxyle en C5' de **74** est réalisé (étape 2). La seconde liaison phosphore-oxygène est ainsi formée et le triester de phosphite **77** est obtenu. Du fait de son instabilité, le triester de phosphite est ensuite oxydé en phosphotriester **78** en présence d'iode, d'eau et de pyridine (étape 3). Les extrémités C5' des pendants nucléosidiques **74** n'ayant pas réagi doivent être bloquées pour éviter des réactions secondaires. Ceci est réalisé (étape 4) par la formation d'acétate **79** (anhydride acétique, *N*-méthylimidazole). Si la séquence de l'oligonucléotide (*n*+1 cycles) est celle souhaitée, le cycle est arrêté, le composé est libéré de la résine et les groupements protecteurs clivés pour mener à **80** (étape 5). Le groupement DMT est coupé en présence d'acide trichloroacétique tandis que les chaînes 2-cyanoéthyle des phosphates sont éliminées en présence d'ammoniaque. Sinon, le cycle est répété.

Cette synthèse automatisée a été illustrée avec une séquence de monomères ADN mais cette stratégie peut également être utilisée pour former des oligonucléotides ARN ou mixte ARN-ADN.

---

<sup>52</sup> Uhlmann, E. and Peyman, A. *Chem. Rev.* **1990**, *90*, 543-584.

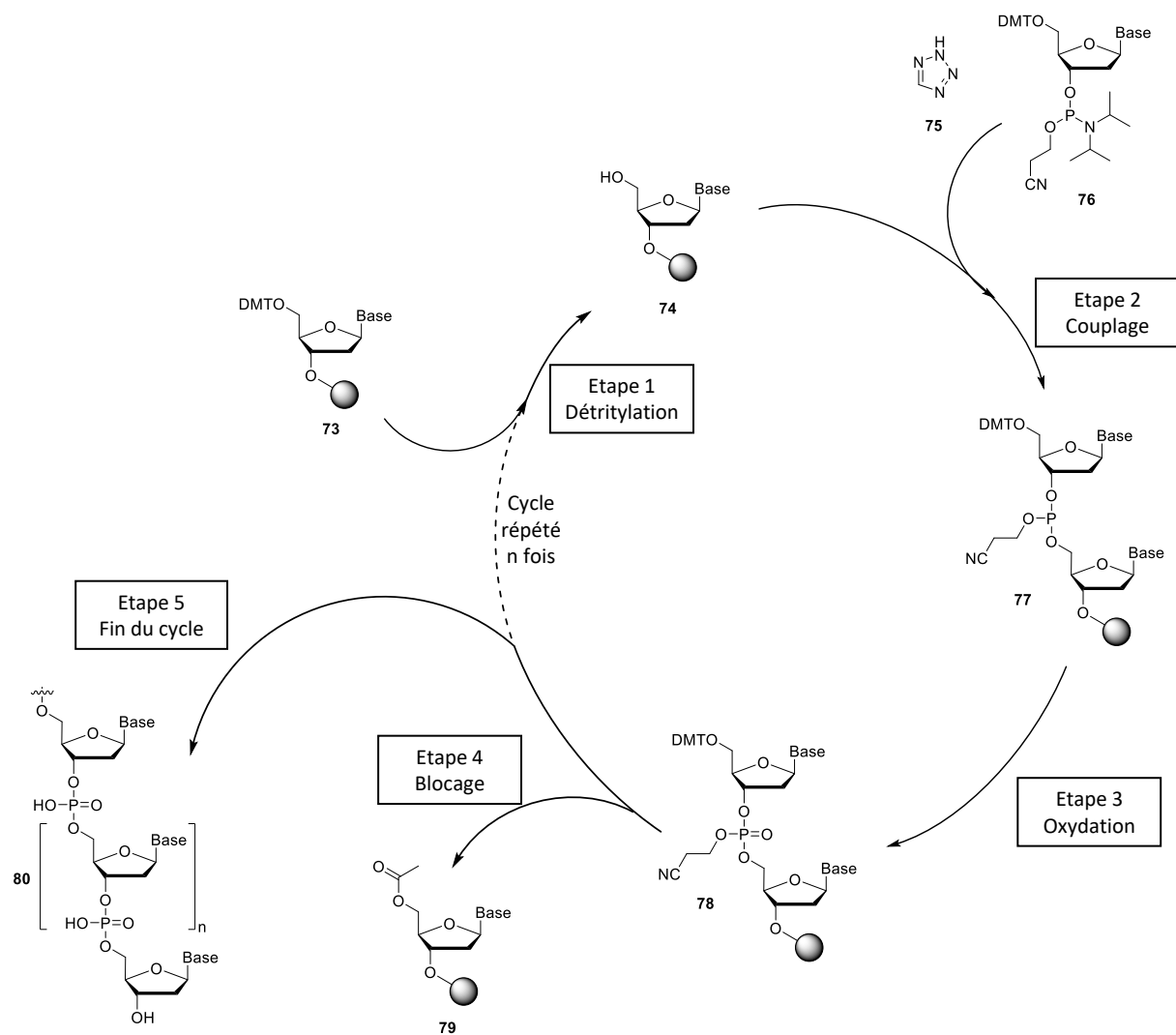


Schéma 2 : Synthèse automatisée d'oligonucléotides

Grâce à cette synthèse automatisée, les rendements moyens obtenus à chaque cycle sont de l'ordre de 98-99%. Plus le nombre de cycle augmente, plus le rendement diminue (Tableau 1).

Tableau 1 : Rendements globaux de la synthèse automatisée en fonction du nombre de cycles

Rendement par cycle		90%	95%	97%	98.5%	99.5%
Nombre de cycles						
10		35%	60%	74%	86%	95%
20		12%	36%	54%	74%	91%
50		0.5%	8%	22%	47%	78%
100		<0.01%	0.6%	5%	22%	61%

De ce fait, plusieurs centaines de milligrammes de nucléosides sont requis afin d'avoir une quantité suffisante d'oligonucléotide final pour réaliser les études physicochimiques et de stabilité métabolique.

III.1.b. Stabilité des duplex

La stabilité d'un duplex est mesurée par sa température de fusion (melting temperature  $T_M$ ), définie comme étant la température à laquelle 50% d'un duplex en solution est dissocié en deux simples brins. Cette température est dépendante de la concentration en oligonucléotides dans le milieu mais également des propriétés du solvant. La méthode expérimentale la plus utilisée pour déterminer  $T_M$  est d'observer le changement de l'absorbance du duplex à 260 nm en fonction de la température (Figure 26). La mesure s'effectue dans une cellule thermostatée, dans un spectromètre UV-visible en continu, en augmentant lentement la température.  $T_M$  est lue comme étant la valeur au milieu de la courbe entre les plateaux ADN double brin (dsDNA) et ADN simple brin (ssDNA).

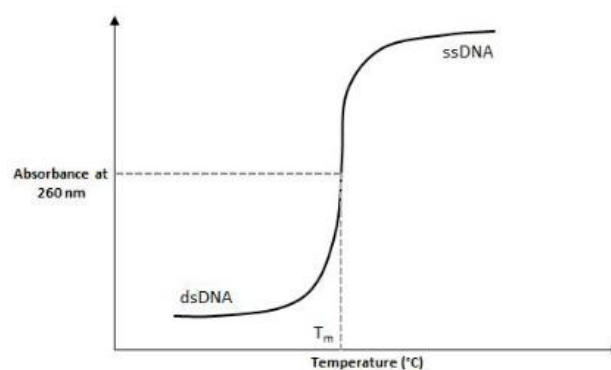


Figure 26 : Mesure de  $T_M$  par absorbance © Merck

La valeur de  $T_M$  peut également être déterminée par calcul en s'appuyant sur deux méthodes. Pour les oligonucléotides de moins de 15 bases nucléiques, la formule de Marmur Doty est utilisée. Celle-ci repose directement sur le nombre d'adénine, de thymine, de cytosine et de guanine.<sup>53</sup> Une seconde formule basée sur les paramètres thermodynamiques de l'oligonucléotide est utilisée dans le cas des séquences de 15 à 120 bases (méthode des voisins proche).<sup>54</sup>

Généralement, sous conditions physiologiques, 12 paires de bases sont requises pour avoir une bonne hybridation entre l'oligonucléotide et sa cible. De plus statistiquement, une séquence de 17 nucléotides différents n'apparaît qu'une seule fois dans le génome humain.<sup>55</sup> Donc, pour être sélectif d'un gène, un oligonucléotide doit inclure entre 15 et 20 nucléotides.

<sup>53</sup> Marmur, J. and Doty, P. J. *Mol. Biol.* **1962**, 5, 109-118.

<sup>54</sup> Freier, S. M.; Kierzek, R.; Jaeger, J. A.; Sugimoto, N.; Caruthers, M. H.; Neilson, T. and Turner, D. H. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, 83, 9373-9377.

<sup>55</sup> Thuong, N. T. and Hélène, C. *Angew. Chem. Int. Ed.* **1993**, 32, 666-690.

### III.1.c. Stabilité métabolique

De nombreuses enzymes peuvent cibler les nucléotides et les oligonucléotides tels que les phosphatases, les enzymes de restriction ou les nucléases. Tout d'abord, les phosphatases sont responsables de l'hydrolyse d'un groupement phosphate sur une molécule. Les enzymes de restriction sont des protéines coupant un fragment d'ADN à une séquence nucléotidique spécifique. Enfin, les nucléases agissent sur l'hydrolyse de la liaison phosphodiester, présente dans les brins d'ARN et d'ADN, et se partagent en deux catégories, les endonucléases et les exonucléases. Les endonucléases sont des enzymes capables de couper en milieu de chaîne alors que les exonucléases hydrolysent exclusivement aux extrémités de la chaîne. A titre d'exemple, l'enzyme Dicer intervenant dans la stratégie d'ARN interférent est une endonucléase venant cliver des fragments double brins d'ARN.<sup>56</sup> La RNase H, intervenant elle dans la stratégie antisens, est une exonucléase qui libère les extrémités 3'-OH et 5'-phosphate dans un double brin hybride ADN-ARN.<sup>57</sup>

Un oligonucléotide est généralement incubé, à 37 °C, dans un tampon contenant des phosphodiesterases afin de déterminer sa stabilité métabolique. Des échantillons sont ensuite prélevés à des temps prédéfinis puis analysés par électrophorèse sur un gel d'agarose afin de visualiser la dégradation des oligonucléotides.

Par exemple, récemment, l'équipe de Brown a étudié la stabilité d'oligonucléotides modifiés comprenant un lien amide en lieu et place du phosphodiester naturel et une structure de type nucléoside bloquée (Locked Nucleic Acid, LNA, Introduction, III.2.c.ii).<sup>58</sup> Pour déterminer leur stabilité, les oligonucléotides modifiés ont été incubés dans un tampon phosphodiesterase 1 et dans un tampon de sérum de veau fœtal. Après migration par électrophorèse, des tâches noires plus ou moins intenses ont été observées (Figure 27). La disparition des tâches intenses, correspondant aux ONs de départ, résulte de la dégradation de ces derniers. Dans le cas des oligonucléotides de contrôle **ON1/ON8**, une dégradation totale a été observée dans les deux tampons. L'ajout d'un LNA (**ON10/ON14**) n'a pas permis d'augmenter la stabilité des oligonucléotides résultants. En revanche, en ajoutant un second LNA, une faible dégradation des oligonucléotides **ON11** et **ON15** a été obtenue.

---

<sup>56</sup> Bernstein, E.; Caudy, A. A.; Hammond, S. M. and Hannon, G. J. *Nature* **2001**, *409*, 363-366.

<sup>57</sup> Schultz, S. J. and Champoux, J. J. *Virus Res.* **2008**, *134*, 86-103.

<sup>58</sup> Thorpe, C.; Epple, S.; Woods, B.; El-Sagheer, A. H. and Brown, T. *Org. Biomol. Chem.* **2019**, *17*, 5341-5348.

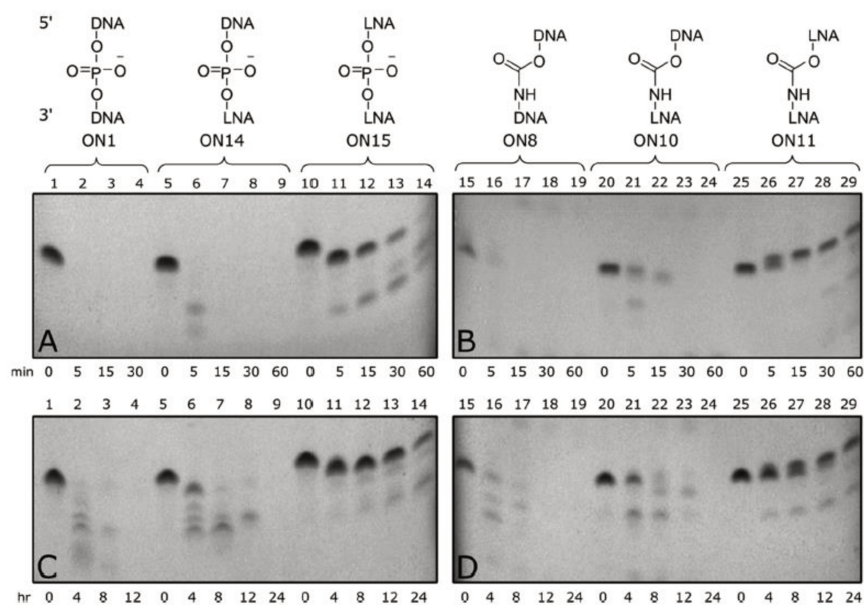


Figure 27 : Résultats de l'étude de Brown  
 A et B : Résultats provenant du tampon phosphodiesterase 1  
 C et D : Résultats provenant du tampon sérum de veau fœtal  
 © The Royal Society Of Chemistry 2019

Ainsi, en mettant de côté l'aspect transport de l'ONM à sa cible, un oligonucléotide modifié sera potentiellement efficace si le duplex ONM-ARN/ADN formé est plus stable que le duplex naturel et si celui-ci est résistant envers les nucléases. Pour répondre à ces critères, de nombreuses modifications ont été apportées sur les oligonucléotides.

### III.2. Modifications des oligonucléotides

Avant de présenter les différentes modifications réalisées sur les ONs, quelques généralités sur la configuration et la conformation des duplex seront présentées.

#### III.2.a. Généralités

Tout d'abord, les nucléosides existent sous la forme de deux anomères,  $\alpha$  et  $\beta$  (exemples de la thymidine, Figure 28). Cependant, dans les chaînes d'ARN et d'ADN, seul l'anomère  $\beta$  est présent.<sup>59</sup>

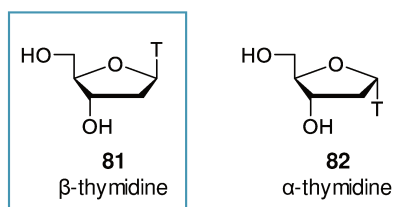


Figure 28 : Anomères  $\alpha$  et  $\beta$  de la thymidine

<sup>59</sup> Séquin, U. *Experientia* **1973**, 29, 1059-1062.

Altona a étudié les différentes conformations du cycle furanique et les a représentées sous la forme d'un cercle, dit de pseudorotation, en fonction des valeurs des angles internes adoptés par le furanose (Figure 29).<sup>60</sup>

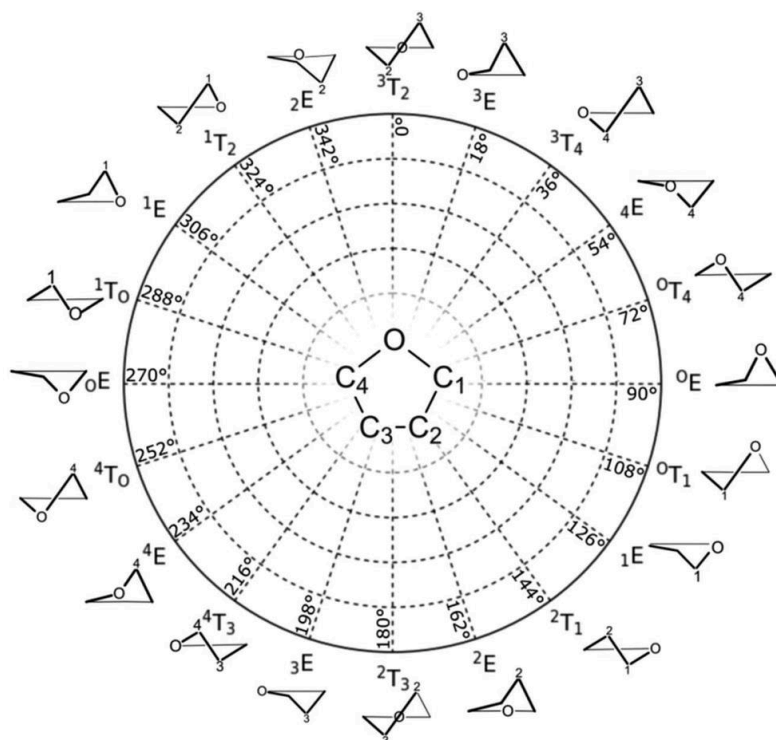


Figure 29 : Cercle de pseudorotation du furanose © JACS 1972

T : Conformation Twist, E : Conformation Enveloppe

Le chiffre en haut à gauche (ex : <sup>3</sup>E) correspond à la position se trouvant sous la conformation endo.

Le chiffre en bas à droite ou à gauche (ex : <sup>2</sup>T<sub>1</sub> ou <sup>3</sup>E) coïncide avec la position étant sous la conformation exo.

Deux conformations sont essentiellement observées dans les chaînes nucléiques, la conformation C2'-endo **83** est présente dans les chaînes d'ADN (Figure 30).<sup>61</sup> Dans les chaînes ARN, les furanoses adoptent une conformation C3'-endo **84**. Il est à noter que les conformations C2'-endo et C3'-endo peuvent également être nommé respectivement conformation « sud » et conformation « nord ».

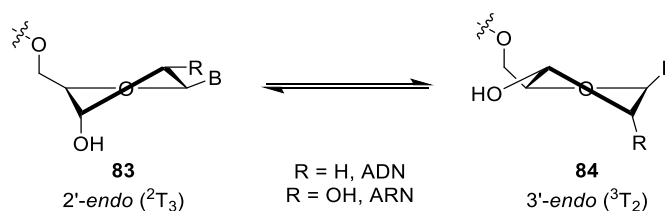


Figure 30 : Conformation du furanose

Dans la stratégie antisens, les furanoses d'un ON thérapeutique doivent majoritairement avoir une conformation de type 3'-endo proche de la conformation présente dans la chaîne d'ARN

<sup>60</sup> Altona, C. and Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205-8212.

<sup>61</sup> (a) Neidle, S. *Principles of Nucleic Acid Structure* **2008**, 20-37. (b) Neidle, S. *Principles of Nucleic Acid Structure* **2008**, 204-248.

complémentaire cible pour former un duplex stable. Afin de satisfaire toutes ces exigences, de nombreuses modifications ont été étudiées notamment au niveau de la **base nucléique**, du **furanose** ou du **pont phosphodiester** (Figure 31). Chacune de ces modifications peut affecter la reconnaissance moléculaire, la structure ou la fonction du nucléotide.

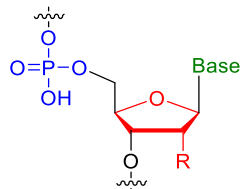


Figure 31 : Modifications possibles sur un nucléotide

Les modifications sur la base nucléique et le furanose seront brièvement présentées puis les changements effectués au niveau du lien phosphodiester seront abordés plus en détail.

### III.2.b. Modifications de la base nucléique

Les bases nucléiques peuvent être modifiées afin d'augmenter l'affinité de l'oligonucléotide avec le brin complémentaire. Quelques exemples sont présentés ci-dessous.

#### III.2.b.i. *Modification structurelle*

L'oxygène en C2 ou C4 de la thymine peut être remplacé par un soufre, plus volumineux, plus polarisable et moins bon accepteur de liaisons hydrogène (Figure 32).

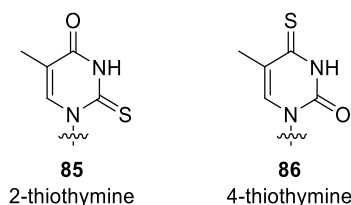


Figure 32 : Modification structurelle de la thymine

L'introduction d'une 2-thiothymine **85**, dans un oligonucléotide, permet de former des duplex stables thermiquement avec l'ARN complémentaire. En comparaison avec une chaîne polythymine non-modifiée, l'introduction de 2-thiothymine a permis d'augmenter la  $T_M$  de +3 °C pour un 8-mère et de +16 °C pour un 15-mère. Une déstabilisation du duplex a été observée lors de l'introduction d'une 4-thiothymine **86** (-16 °C pour un 8-mère).<sup>62</sup> Plus récemment, Ostergaard a montré que l'introduction de 2-thiothymine **85**, à des positions spécifiques dans un oligonucléotide, permet de moduler la sélectivité allélique.<sup>63</sup>

<sup>62</sup> Ishikawa, T.; Yoneda, F.; Tanaka, K. and Fuji, K. *Bioorg. Med. Chem. Lett.* **1991**, *1*, 523-526.

<sup>63</sup> Østergaard, M. E.; Southwell, A. L.; Kordasiewicz, H.; Watt, A. T.; Skotte, N. H.; Doty, C. N.; Vaid, K.; Villanueva, E. B.; Swayze, E. E.; Bennett, C. F.; Hayden, M. R. and Seth, P. P. *Nucleic Acids Res.* **2013**, *41*, 9634-9650.

## III.2.b.ii. Introduction de substituant

La position C5 des bases pyrimidiques est la plus simple à modifier d'un point de vue synthétique (Figure 33).

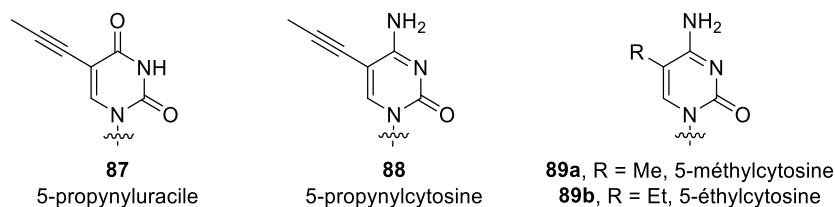


Figure 33 : Introduction de groupement sur la position 5 des bases pyrimidiques

Par exemple, un groupement propyne peut y être inséré. Dans une étude de Froehler, l'introduction de la 5-propynyluracile **87** dans un  $\beta$ -oligonucléotide a permis de stabiliser le duplex ONM-ARN (entrée 2, +8.5 °C par rapport à la référence) (Tableau 2). Le même résultat a été obtenu avec l'introduction de la 5-propynylcytosine **88** (entrée 3, +7.0 °C).<sup>64</sup> Cependant, ces composés ont présenté une faible résistance aux nucléases. Une étude similaire a été réalisée par Morvan sur un  $\alpha$ -oligonucléotide (c'est-à-dire une chaîne comportant des  $\alpha$ -nucléotides non naturels).<sup>65</sup> En comparaison avec un  $\beta$ -oligonucléotide (entrée 4), l'utilisation d' $\alpha$ -oligonucléotide a entraîné une déstabilisation du complexe (entrée 5, -4.8 °C en version ADN et -2.7 °C en version ARN) mais une résistance aux nucléases a été observée.<sup>66</sup> La synthèse d'un  $\alpha$ -oligonucléotide comportant la 5-propynyluracile **87** a permis de stabiliser des duplex en version ONM-ADN (+9.9 °C) et en version ONM-ARN (+11.5 °C) (entrée 6).

Tableau 2 : Résultats des études comportant un motif 5-propynyl

	Entrée	Oligonucléotide	Version ADN	Version ARN
			$T_M$ (°C)	$T_M$ (°C)
Froehler	1	$\beta$ -phosphodiester	/	65.5
	2	5-propynyluracile $\beta$ -phosphodiester	/	74.0
	3	5-propynylcytosine $\beta$ -phosphodiester	/	73.0
Morvan	4	$\beta$ -phosphodiester	46.4	45.9
	5	$\alpha$ -phosphodiester	41.8	43.2
	6	5-propynyluracile $\alpha$ -phosphodiester	51.7	54.7

<sup>64</sup> Froehler, B. C.; Wadwani, S.; Terhorst, T. J. and Gerrard, S. R. *Tetrahedron Lett.* **1992**, 33, 5307-5310.

<sup>65</sup> Morvan, F.; Porumb, H.; Degols, G.; Lefebvre, I.; Pompon, A.; Sproat, B. S.; Rayner, B.; Malvy, C.; Lebleu, B. and Imbach, J. L. *J. Med. Chem.* **1993**, 36, 280-287.

<sup>66</sup> Morvan, F.; Zeidler, J. and Rayner, B. *Tetrahedron* **1998**, 54, 71-82.

Sági a réalisé une étude sur l'introduction d'un groupement en C5 des bases pyrimidiques (Tableau 3).<sup>67</sup> Le motif 5-méthylcytosine **89a** a été introduit dans un oligonucléotide ADN. Une augmentation de la stabilité de cet ON a été observée en comparaison avec une chaîne cytosine non-modifiée (+5.9 °C) (entrée 2). La substitution par un groupement plus large (éthylcytosine **89b**) a entraîné une diminution de l'affinité avec le brin complémentaire (-3.1 °C) (entrée 3).

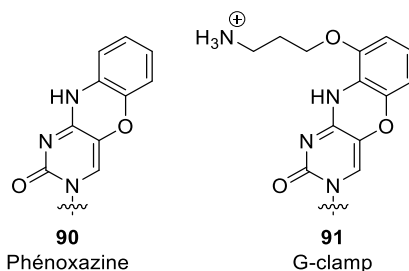
Tableau 3 : Résultats de l'étude de Sági

Entrée	Oligonucléotide	$T_M$ (°C)	$\Delta T_M^{68}$ (°C)
1	Cytosine	84.3	/
2	5-méthylcytosine	90.2	+5.9
3	5-éthylcytosine	81.2	-3.1

### III.2.b.iii. Remplacement de la base nucléique

Les bases nucléiques peuvent être remplacées par des analogues, comme par exemple, la phénoxazine **90** (Tableau 4). Matteucci a réalisé l'introduction de cet analogue de la cytosine dans un oligonucléotide d'une longueur de 10 unités.<sup>69</sup> En comparaison avec leur oligonucléotide référence comportant la 5-méthylcytosine **89a**, l'insertion d'une phénoxazine **90** a permis l'augmentation de la stabilité du duplex (entrée 2, +6.5 °C). L'ajout d'une amine protonée (**91**) a permis de stabiliser davantage le duplex (entrée 3, +18.0 °C) grâce à la formation de quatre liaisons hydrogène entre le G-clamp **91** et la guanine complémentaire (Figure 34).

Tableau 4 : Résultats des travaux de Matteucci



Entrée	X	$T_M$ (°C)	$\Delta T_M$ (°C)
1	5-méthylcytosine <b>89a</b>	50.5	/
2	Phénoxazine <b>90</b>	57.0	+6.5
3	G-clamp <b>91</b>	68.5	+18.0

Oligonucléotide test : 5'-TCTCXCTCTC

<sup>67</sup> Sági, J.; Szemző, A.; Ötvös, L.; Vorlíčková, M. and Kypr, J. *Int. J. Biol. Macromol.* **1991**, *13*, 329-336.

<sup>68</sup>  $\Delta T_M$ : Différence entre  $T_M$  de l'oligonucléotide modifié et  $T_M$  de l'oligonucléotide non-modifié

<sup>69</sup> Lin, K.-Y. and Matteucci, M. D. *J. Am. Chem. Soc.* **1998**, *120*, 8531-8532.

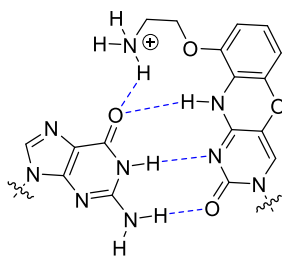


Figure 34 : Liaisons hydrogènes entre G-clamp et guanine

### III.2.c. Modifications du furanose

Les principales modifications apportées au furanose se font en introduisant un groupement en C2', en bloquant sa conformation ou bien en remplaçant ce dernier par un cycle analogue.

#### III.2.c.i. *Modification en C2'*

L'une des modifications les plus étudiées, sur le furanose, est l'insertion d'un substituant en C2' (Figure 35).<sup>70</sup>

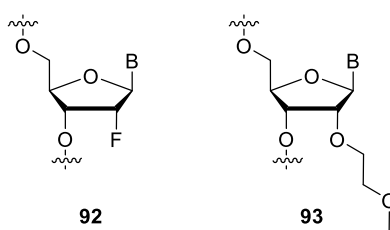


Figure 35 : Modification du substituant en C2' du furanose

Par exemple, l'introduction d'un fluor (**92**) a permis d'améliorer la stabilité envers les nucléases et d'augmenter l'affinité et la sélectivité avec le brin d'ARN-cible. En effet, lors de l'utilisation d'un 2'-déoxy-2'-fluoro nucléoside, Kawasaki a obtenu des valeurs de température de fusion supérieures à celle de l'oligonucléotide non-modifié (de +7.9 °C à +25.2 °C).<sup>71</sup> Cette modification est présente dans un des oligonucléotides approuvés par la FDA, le Macugen.

Le groupement MOE **93** (MéthOxyEthyl) est l'une des modifications la plus utilisée en stratégie antisens ; celle-ci est présente dans le Kynamro™, le Spinraza® et le Tegsedi. Comme pour l'introduction d'un fluor, ce groupement a permis d'augmenter la stabilité vis-à-vis des nucléases mais également d'améliorer l'affinité de l'oligonucléotide avec le brin d'ARN complémentaire en favorisant la conformation C3'-endo.<sup>72</sup>

<sup>70</sup> Egli, M.; Minasov, G.; Tereshko, V.; Pallan, P. S.; Teplova, M.; Inamati, G. B.; Lesnik, E. A.; Owens, S. R.; Ross, B. S.; Prakash, T. P. and Manoharan, M. *Biochemistry* **2005**, *44*, 9045-9057.

<sup>71</sup> Kawasaki, A. M.; Casper, M. D.; Freier, S. M.; Lesnik, E. A.; Zounes, M. C.; Cummins, L. L.; Gonzalez, C. and Cook, P. D. *J. Med. Chem.* **1993**, *36*, 831-841.

<sup>72</sup> Teplova, M.; Minasov, G.; Tereshko, V.; Inamati, G. B.; Cook, P. D.; Manoharan, M. and Egli, M. *Nat. Struct. Mol. Biol.* **1999**, *6*, 535.

## III.2.c.ii. Utilisation de furanose bloqué

La conformation C3'-*endo* du cycle furanosique peut notamment être bloquée par la création d'un pont entre les positions C2' et C4' du cycle (Figure 36).

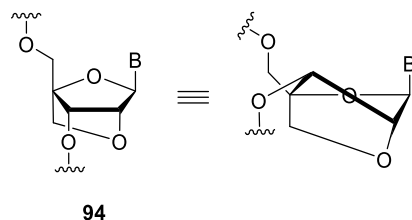


Figure 36 : Insertion d'un pont bloquant la conformation du furanose

Ce type de nucléoside bloqué est communément appelé « Locked Nucleic Acid » **94** (LNA). Un LNA a pour fonction de fixer la conformation du furanose sous la forme 3'-*endo*. De ce fait, l'appariement avec les brins d'ARN complémentaire est facilité et permet la formation de duplex stables.<sup>73</sup> Ainsi, lors de l'incorporation d'un LNA dans un oligonucléotide, une augmentation drastique de l'affinité avec le brin complémentaire d'ADN ou d'ARN a été observée. En effet, chaque insertion d'un LNA a permis d'augmenter la température de fusion du duplex d'environ +5 °C.<sup>74</sup> De plus, comme vu précédemment (III.1.c), l'ajout d'un LNA a entraîné une augmentation de la stabilité des composés vis-à-vis des nucléases. Au vu de leur potentiel thérapeutique prometteur, le développement de LNA a connu un véritable essor et une grande variété de modification (**95-99**) ont vu le jour (Figure 37).<sup>75</sup> A l'heure actuelle, quelques oligonucléotides incorporant des unités LNA sont en phase d'essais cliniques.<sup>48</sup>

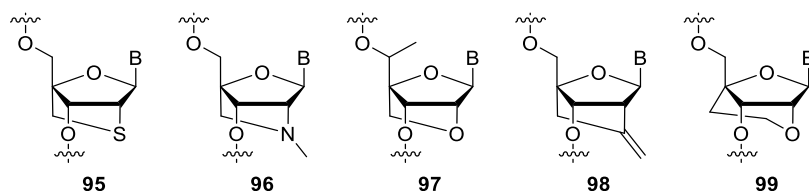


Figure 37 : Exemples de structures LNA

<sup>73</sup> Obika, S.; Nanbu, D.; Hari, Y.; Morio, K.-i.; In, Y.; Ishida, T. and Imanishi, T. *Tetrahedron Lett.* **1997**, *38*, 8735-8738.

<sup>74</sup> Obika, S.; Nanbu, D.; Hari, Y.; Andoh, J.-i.; Morio, K.-i.; Doi, T. and Imanishi, T. *Tetrahedron Lett.* **1998**, *39*, 5401-5404.

<sup>75</sup> Wan, W. B. and Seth, P. P. *J. Med. Chem.* **2016**, *59*, 9645-9667.

## III.2.c.iii. Remplacement du furanose par un analogue

Pour finir, l'augmentation de la taille du cycle (hexose **100** et oxépane **101**) ainsi que son remplacement par un analogue (morpholino **102**) ont également été étudiés (Figure 38).

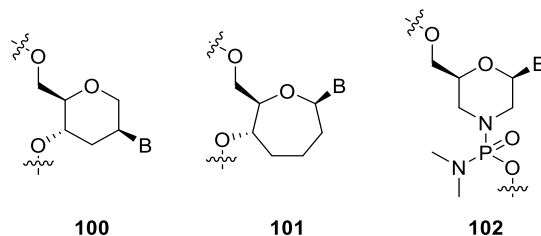


Figure 38 : Analogues du furanose

Le motif hexose **100**, dont la base nucléique n'a pas été insérée en position anomérique, a été introduit par Van Aerschot dans un oligonucléotide 6-mère et un 12-mère.<sup>76</sup> Une augmentation drastique de la stabilité du duplex et de la stabilité métabolique a été observée. En effet, une température de fusion  $T_M$  du duplex de +34 °C a été obtenue dans le cas du 6-mère et de +36 °C pour le 12-mère. Cette modification a donc permis la formation d'un duplex très stable. De plus, une résistance envers les phosphodiesterases du venin de serpent a été observée.

L'ajout d'un carbone supplémentaire pour former l'oxépane **101** a été réalisé par Damha.<sup>77</sup> Malgré une résistance enzymatique observée dans un sérum de veau foetal, l'introduction de ce motif dans un oligonucléotide a entraîné une déstabilisation du duplex. Une diminution de la valeur de  $T_M$  de -20 °C a été observée pour un duplex oxépane-ARN.

Le motif morpholinophosphorodiamidate **102** a montré une bonne résistance envers des nucléases, des protéases, des estérases et d'autres enzymes biologiques.<sup>78</sup> Le potentiel prometteur du motif morpholino a conduit au développement d'un oligonucléotide antisens contre la dystrophie musculaire de Duchenne, approuvé récemment par la FDA, à savoir l'Eteplirsén.

III.2.d. Modification du pont phosphodiester naturel

Etudiée depuis le début des années 90, la modification du pont phosphodiester a permis de mettre en évidence un certain nombre d'analogues. Certaines modifications ont déjà été présentées précédemment (I.2.b). Ici, quelques isostères du pont phosphodiester seront présentés avant de nous intéresser plus particulièrement aux analogues se rapprochant du domaine de recherche du laboratoire, à savoir les phosphorothioates, les phosphonates et les phosphinates.

<sup>76</sup> Aerschot Van, A.; Verheggen, I.; Hendrix, C. and Herdewijn, P. *Angew. Chem. Int. Ed.* **1995**, *34*, 1338-1339.

<sup>77</sup> Sabatino, D. and Damha, M. J. *J. Am. Chem. Soc.* **2007**, *129*, 8259-8270.

<sup>78</sup> Nan, Y. and Zhang, Y.-J. *Front. Microbiol.* **2018**, *9*, 750-750.

## III.2.d.i. Les isostères du pont phosphate naturel

De nombreux isostères du pont phosphodiester naturel ont été étudiés au cours des dernières années. Des dérivés guanidinium **103**,<sup>79</sup> sulfone **104**,<sup>80</sup> amide **105**,<sup>81</sup> sulfonamide **106**,<sup>82</sup> éther **107**,<sup>83</sup> silyle **108**<sup>84</sup> et triazole **109**<sup>85</sup> ont ainsi été développés (Figure 39).

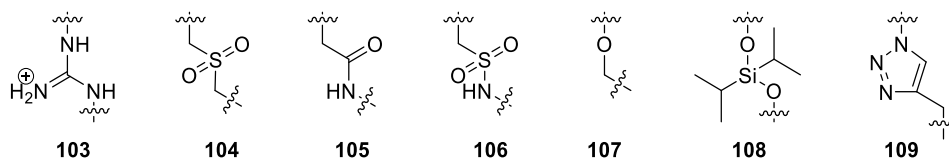
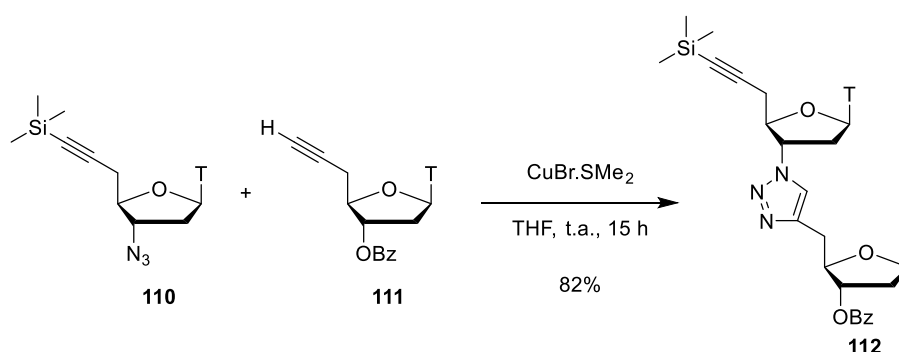


Figure 39 : Isostères du phosphate naturel

En guise d'exemple, le groupe d'Isobe a développé un lien triazole **109** afin de lier deux unités nucléosidiques (Schéma 3). Assemblé par réaction de cycloaddition 1,3-dipolaire de Huisgen entre deux nucléosides, un comportant une fonction azoture **110** et le second une fonction alcyne **111**, ce lien a montré d'excellentes propriétés. En effet, la formation d'une double hélice stable avec un brin d'ADN naturel a été observée. Une augmentation drastique de la température de fusion (+41.1 °C) a été obtenue avec la présence du lien triazole.

Schéma 3 : Formation d'un dinucléotide comportant le motif triazole **109**

Comme second exemple, le pont amide **105** a été présenté comme étant un isostère du pont phosphodiester naturel. En effet, De Mesmaeker a montré que l'introduction d'un lien amide, en lieu et place du phosphate naturel, a permis d'augmenter la stabilité thermique d'un duplex oligodéoxynucléotide-ARN (+0.5 °C). Résultant de la liaison de deux nucléosides **113** et **114** (Schéma 4), le pont amide a également montré une résistance envers les 3'-exonucléases.

<sup>79</sup> Park, M. and Bruice, T. C. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3982-3986.

<sup>80</sup> Huang, Z.; Schneider, K. C. and Benner, S. A. *J. Org. Chem.* **1991**, *56*, 3869-3882.

<sup>81</sup> De Mesmaeker, A.; Waldner, A.; Lebreton, J.; Hoffmann, P.; Fritsch, V.; Wolf, R. M. and Freier, S. M. *Angew. Chem. Int. Ed.* **1994**, *33*, 226-229.

<sup>82</sup> Korotkovs, V.; Reichenbach, L. F.; Pescheteau, C.; Burley, G. A. and Liskamp, R. M. J. *J. Org. Chem.* **2019**, *84*, 10635-10648.

<sup>83</sup> Sengupta, J. and Bhattacharjya, A. *J. Org. Chem.* **2008**, *73*, 6860-6863.

<sup>84</sup> Cormier, J. F. and Ogilvie, K. K. *Nucleic Acids Res.* **1988**, *16*, 4583-4594.

<sup>85</sup> Isobe, H.; Fujino, T.; Yamazaki, N.; Guillot-Nieckowski, M. and Nakamura, E. *Org. Lett.* **2008**, *10*, 3729-3732.

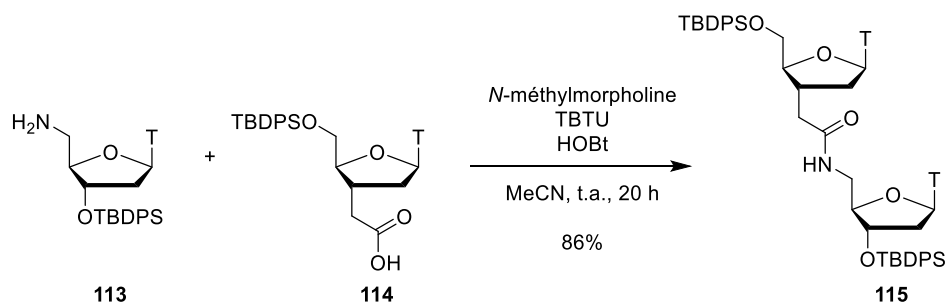


Schéma 4 : Synthèse d'un dinucléotide comportant le motif amide 105

## III.2.d.ii. Phosphorothioates

Le lien phosphorothioate **116** est la modification la plus utilisée dans les oligonucléotides thérapeutiques. En effet, celle-ci est présente dans quatre des cinq médicaments antisens présents sur le marché (Vitravene™, Mipomersen, Spinraza et Tegsedi).

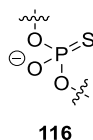


Figure 40 : Lien phosphorothioate

Bien qu'une augmentation de la stabilité métabolique soit obtenue en présence du lien phosphorothioate, une réduction de l'affinité pour le brin complémentaire a été observée.<sup>86</sup> Néanmoins, cet analogue active la RNase H (à l'origine de la dégradation des ARNm-cibles) et offre des avantages pharmacocinétiques, il permet notamment l'entrée des ONMs dans certaines cellules grâce à des interactions avec des protéines de surface.<sup>87</sup> Il est à noter que l'insertion de ce lien entraîne la formation de diastéréoisomères, liée à la présence d'un centre phosphore stéréogénique, expliquant potentiellement la déstabilisation du duplex. La synthèse du motif phosphorothioate, se faisant de manière non-sélective, conduit à la formation de nombreux stéréoisomères. Par exemple, le Spinraza, contenant ce motif, est un mélange de 131072 stéréoisomères. Récemment, Baran a mis au point une synthèse permettant de contrôler la stéréochimie du lien phosphorothioate **116**.<sup>88</sup> A partir d'un réactif dérivé du limonène dont la configuration du phosphore a été bloquée sous la forme (*R*) pour **118** et (*S*) pour **122**, une première étape a été réalisée afin d'obtenir sélectivement les intermédiaires **119** et **123** dont les configurations sont inversées, avec d'excellents rendements (Schéma 5). Une étape de couplage avec un nucléoside **120** a permis d'obtenir un dinucléotide avec une seconde inversion de configuration. Ainsi à partir du réactif (*R*) **118**, un dinucléotide **121** de configuration (*R*) a été obtenu.

<sup>86</sup> Freier, S. M. and Altmann, K. H. *Nucleic Acids Res.* **1997**, *25*, 4429-4443.

<sup>87</sup> Fritz, E. *Antisense Nucleic Acid Drug Dev.* **2000**, *10*, 117-121.

<sup>88</sup> Knouse, K. W.; deGruyter, J. N.; Schmidt, M. A.; Zheng, B.; Vantourout, J. C.; Kingston, C.; Mercer, S. E.; McDonald, I. M.; Olson, R. E.; Zhu, Y.; Hang, C.; Zhu, J.; Yuan, C.; Wang, Q.; Park, P.; Eastgate, M. D. and Baran, P. S. *Science* **2018**, *361*, 1234-1238.

Le réactif (S) **122** a, lui, conduit au dinucléotide **124** de configuration (S). Cette synthèse a été réalisée avec différentes combinaisons de nucléosides avec des rendements compris entre 61% et 91%. Cette méthode pourrait être appliquée dans le futur pour la synthèse de nouveaux oligonucléotides thérapeutiques comportant le motif phosphorothioate tout en contrôlant sa stéréochimie. Cette stratégie pourrait ainsi permettre l'étude des différentes configurations d'un oligonucléotide afin de déterminer la forme la plus active sur l'ARN-cible.

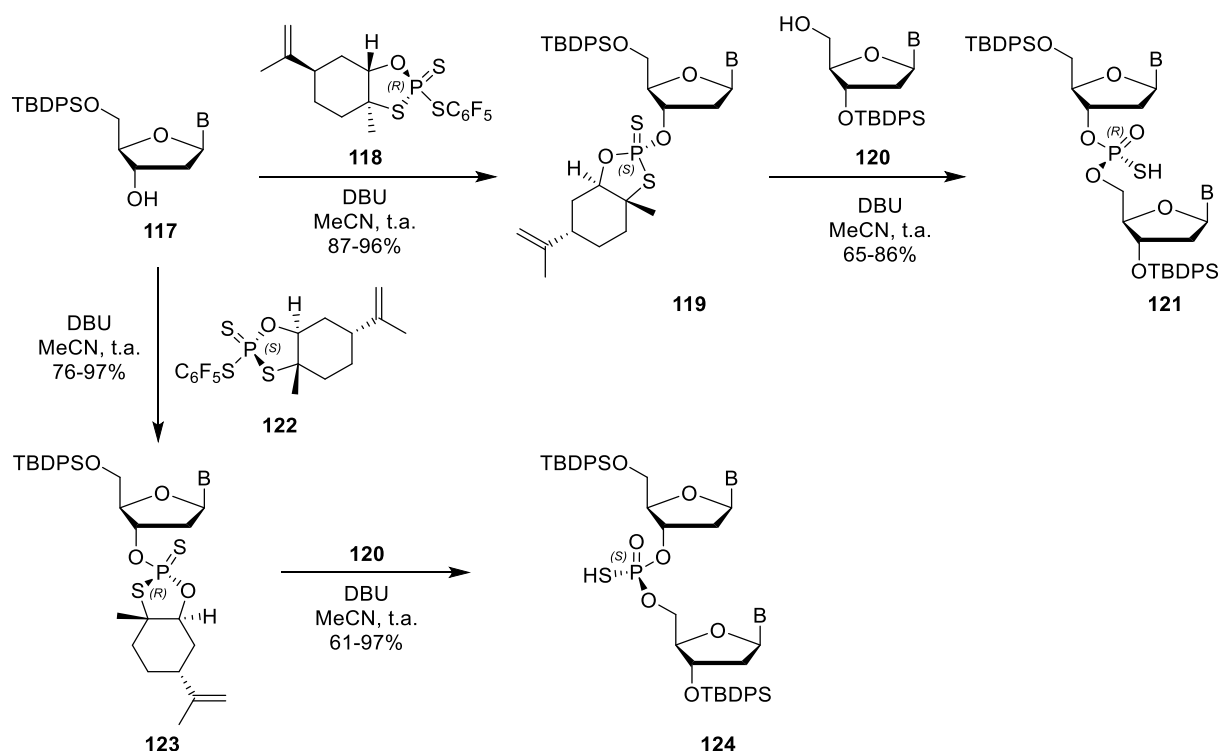


Schéma 5 : Travaux de Baran

### III.2.d.iii. Phosphonates et phosphinates

Dans cette partie seront donc abordés les analogues comportant une ou deux liaison(s) phosphore-carbone. Quatre types de modification sont alors possibles (Figure 41). N'ayant pas été introduit dans un oligonucléotide, le phosphinate **126** ne sera pas abordé.

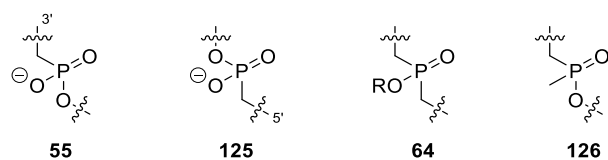


Figure 41 : Analogues possédant une ou deux liaison(s) phosphore-carbone

La synthèse du 3'-méthylène phosphonate **55** a été rapportée par An.<sup>89</sup> La liaison phosphore-carbone a été obtenue par réaction d'Arbuzov (Schéma 6). Une substitution nucléophile

<sup>89</sup> An, H.; Wang, T.; Maier, M. A.; Manoharan, M.; Ross, B. S. and Cook, P. D. *J. Org. Chem.* **2001**, *66*, 2789-2801.

du bis(triméthylsilyl)phosphonite sur l'iodonucléoside **127** a été effectuée afin de former le *H*-phosphinate **128** avec des rendements moyens. Par la suite, celui-ci a été transformé en phosphoramidite **130** pour être introduit dans un oligonucléotide (**131**) par synthèse automatisée. Le 3-hydroxypropionitrile a été condensé sur **128**, à l'aide de DCC, pour former le cyanoéthyl *H*-phosphinate **129**. Puis, celui-ci a été activé sous la forme chlorophosphonate afin de réaliser l'addition de la diisopropylamine et de former **130**.

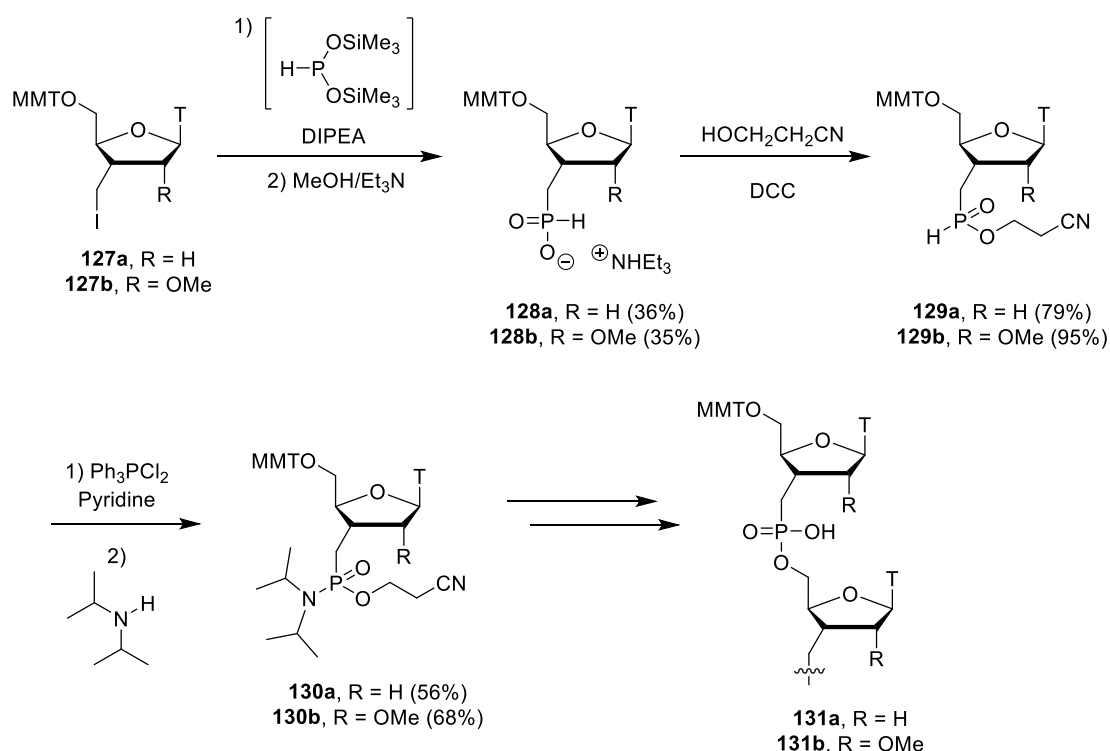


Schéma 6 : Synthèse du motif 3'-méthylène phosphonate 55

Dans la suite de l'étude, les propriétés biophysiques d'oligonucléotides modifiés comportant le lien phosphonate **55** couplé à des modifications en C2' ont été analysées. Les duplex oligonucléotide-ARN résultants ont présentés une complète résistance aux phosphodiesterases du venin de serpent. En comparaison avec un brin ADN phosphorothioate, une meilleure hybridation a été observée avec une augmentation de la température de fusion de +3.6 °C par modification.

Le pont 5'-méthylène phosphonate **125** a été introduit par Caruthers en 1993.<sup>90</sup> Hutter a réalisé une étude afin de comparer le lien 5'-méthylène phosphonate **125** au phosphinate **64** et au bis(méthylène)sulfone **104**.<sup>91</sup> L'introduction du diéthyle méthylphosphonate a été réalisée selon une méthode développée par Tanaka (Schéma 7).<sup>92</sup> Cette dernière a consistée en l'ouverture du cycle oxétane **132** par le carbanion dérivé du diéthyl phosphonite en présence d'un acide de Lewis. Après

<sup>90</sup> Böhringer, M. P.; Graff, D. and Caruthers, M. H. *Tetrahedron Lett.* **1993**, *34*, 2723-2726.

<sup>91</sup> Hutter, D.; Blaettler, M. O. and Benner, S. A. *Helv. Chim. Acta* **2002**, *85*, 2777-2806.

<sup>92</sup> Tanaka, H.; Fukui, M.; Haraguchi, K.; Masaki, M. and Miyasaka, T. *Tetrahedron Lett.* **1989**, *30*, 2567-2570.

inversion de la stéréochimie du groupement hydroxyle en position C3' par réaction de Mitsunobu, une monodéprotection du phosphonate **134** a été réalisée à l'aide de bromure de lithium. Le nucléotide résultant **135** a été condensé avec l'hydroxyle en C3' d'un nucléoside protégé **136** pour former le dinucléotide **137**.

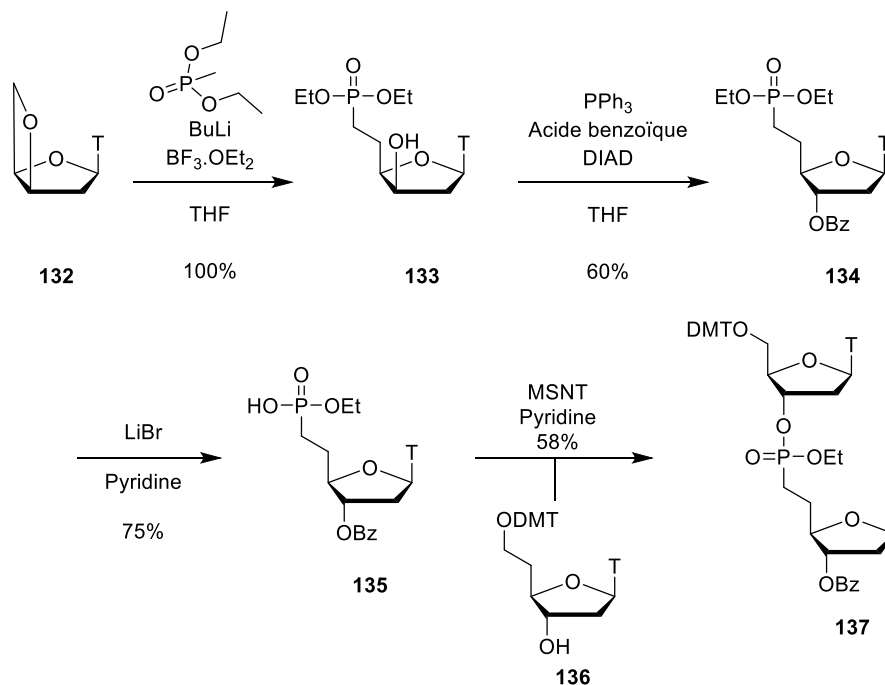


Schéma 7 : Synthèse du motif 5'-méthylène phosphonate **125**

Par la suite, ce dinucléotide **137** a été introduit dans différents oligonucléotides ADN par synthèse automatisée. Une déstabilisation des duplex a été observée dans chaque cas. Concernant la stabilité métabolique, une résistance a été observée dans le cas de la phosphodiesterase de bovin, qui clive spécifiquement la liaison P-O en position C5'.<sup>93</sup> Au contraire, une dégradation totale a été obtenue dans le cas de l'utilisation de la phosphodiesterase du venin de serpent spécifique à l'hydrolyse de la liaison P-O en position C3'.

Les nucléotides modifiés contenant le lien phosphinate **64** ont quant à eux été développés par Collingwood.<sup>94</sup> En suivant la méthode de Tanaka, l'action du carbanion du phosphinate généré à l'aide de BuLi a permis l'ouverture du cycle oxétane **132** en présence d'un acide de Lewis (Schéma 8). Après inversion de la stéréochimie de l'hydroxyle en C3' par réaction de Mitsunobu et déprotection du phosphinate du nucléotide **138**, une addition nucléophile de **139** sur l'aldéhyde **140** en présence de DBU a été réalisée afin de former le dinucléotide **141**. Le composé **142** a été obtenu par

<sup>93</sup> Szabó, T.; Kers, A. and Stawinski, J. *Nucleic Acids Res.* **1995**, 23, 893-900.

<sup>94</sup> Collingwood, S. P. and Baxter, A. D. *Synlett* **1995**, 1995, 703-705.

désoxygénation de Barton-McCombie entraînant le remplacement de l'hydroxyle en alpha du phosphore par un hydrogène.

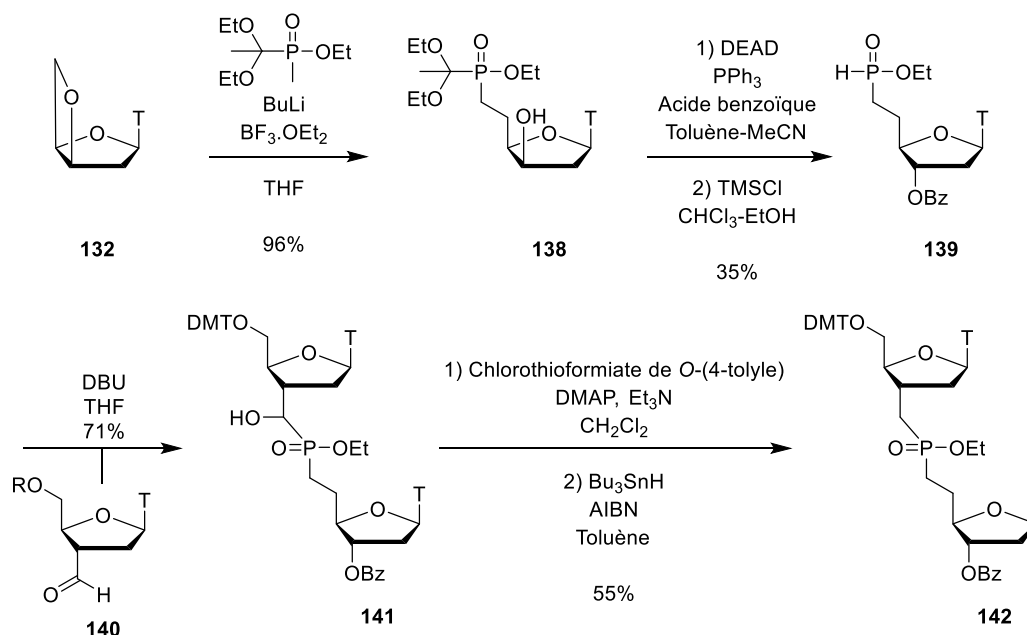


Schéma 8 : Synthèse du dinucléotide 142

Malgré une résistance observée envers les 3'-exonucléases, la stabilité du duplex a été réduite de -2.2 °C par modification en série ARN et -5.5 °C pour la série ADN. Cette déstabilisation pourrait être expliquée par l'absence d'un substituant électro-négatif en position C3' favorisant la conformation 3'-endo.<sup>61</sup>

Le remplacement d'un ou deux atome(s) d'oxygène sur le motif phosphodiester naturel a permis le développement d'analogues résistants aux nucléases. Le motif 3'-méthylène phosphonate **55** a montré les meilleurs résultats, en comparaison avec le 5'-méthylène phosphonate **125** et le phosphinate **64**, en présentant une résistance métabolique et en permettant la stabilisation du duplex avec l'ARN complémentaire.

La modification du pont phosphodiester naturel est un enjeu majeur pour améliorer la stabilité, l'affinité et la sélectivité d'un oligonucléotide thérapeutique. De ce fait, le laboratoire travaille au développement de nouveaux analogues du phosphate naturel caractérisés par la présence d'un motif CF<sub>2</sub> en lieu et place d'un oxygène estérifié.

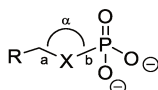
#### IV. Des phosphonates aux difluorophosphonates

Depuis quelques années, le motif  $\text{CF}_2$  s'est imposé comme étant un bon mime de l'oxygène estérifié du groupement phosphate naturel. Dans cette partie seront présentés les aspects électroniques et structuraux faisant de ce motif un isostère de l'oxygène.

##### IV.1. Analogie structurelle

Chambers et O'Hagan ont montré l'existence d'une analogie structurelle entre le phosphate **12** et le motif **144** par étude cristallographique (Tableau 5).<sup>95</sup> L'introduction d'un motif  $\text{CH}_2$  (**143**, entrée 2) ou d'un motif CHF (**145**, entrée 3) à la place de l'oxygène estérifié a entraîné une forte diminution de l'angle en comparaison avec le phosphate naturel **12** ( $-6.6^\circ$  pour  $\text{CH}_2$ ,  $-5.4^\circ$  pour CHF). Dans le cas de  $\text{CF}_2$  **144**, l'angle  $\alpha$  a une valeur légèrement inférieure à celle de **12** ( $-1.9^\circ$ , entrée 4) mais celui-ci a été compensé par une augmentation globale des longueurs de liaison a et b. D'un point de vue structurel, le groupement  $\text{CF}_2$  peut être considéré comme un bon mime de l'oxygène du phosphate naturel.

Tableau 5 : Données géométriques et  $\text{pK}_a$  du phosphate et de ces analogues



Entrée	X	Composé	$\alpha$ ( $^\circ$ )	a ( $\text{\AA}$ )	b ( $\text{\AA}$ )	$\text{pK}_a_2$
1	O	<b>12</b>	118.7	1.43	1.59	6.4
2	$\text{CH}_2$	<b>143</b>	112.1	1.807	1.51	7.6
3	CHF	<b>145</b>	113.3	1.82	1.50	6.5
4	$\text{CF}_2$	<b>144</b>	116.5	1.85	1.496	5.4

Chambers et O'Hagan ont également réalisé une étude comparative des valeurs de  $\text{pK}_a$  de deuxième déprotonation du phosphate **12** et des phosphonates **143-145**. L'insertion d'un motif  $\text{CH}_2$  **143** a diminué l'acidité de la molécule ( $\text{pK}_a_2 = 7.6$ , entrée 2) en comparaison avec le phosphate naturel **12** ( $\text{pK}_a_2 = 6.4$ , entrée 1). Tandis que l'incorporation d'un motif  $\text{CF}_2$  **144** en lieu et place de l'oxygène estérifié a augmenté le caractère acide ( $\text{pK}_a_2 = 5.4$ , entrée 4) à cause de l'effet inductif électroattracteur des atomes de fluor.<sup>96</sup> L'introduction d'un seul atome de fluor **145** a permis de moduler l'acidité de la molécule afin d'obtenir une valeur proche du phosphate naturel ( $\text{pK}_a_2 = 6.5$ , entrée 3). Cependant, la présence d'un seul fluor génère une difficulté synthétique supplémentaire puisqu'un nouveau centre stéréogène est créé. Notons que le motif CHF **145** a été

<sup>95</sup> (a) Chambers, R. D.; O'Hagan, D.; Lamont, R. B. and Jaina, S. C. *J. Chem. Soc., Chem. Commun.* **1990**, 1053-1054. (b) Nieschalk, J.; Batsanov, A. S.; O'Hagan, D. and Howard, J. *Tetrahedron* **1996**, 52, 165-176. (c) Okaya, Y. *Acta Cryst.* **1966**, 20, 712-715.

<sup>96</sup> Blackburn, G. M. and Kent, D. E. *J. Chem. Soc., Chem. Commun.* **1981**, 511-513.

inséré dans un dinucléotide par Baxter et Collingwood ; cependant, aucune introduction de ce dinucléotide au sein d'un oligonucléotide n'a été réalisée.<sup>97</sup>

#### IV.2. Résistance du motif CF<sub>2</sub> à l'hydrolyse

En série phosphate, pour éviter l'hydrolyse de la liaison phosphore-oxygène par les enzymes (Introduction, I.2.a), Myers a remplacé l'oxygène estérifié du phosphate **12** par un motif CH<sub>2</sub> **143** dans son étude (Figure 42).<sup>12</sup> Une résistance aux nucléases a ainsi été observée. Par la suite, Blackburn a introduit un groupement difluorométhylène à la place de l'oxygène estérifié.<sup>14</sup> Dans le développement d'inhibiteurs de la transcriptase inverse du VIH, la présence du motif **144** a permis d'obtenir des composés possédant une activité inhibitrice et une résistance accrue face aux enzymes déphosphorylantes.<sup>98</sup> Le motif CF<sub>2</sub> est donc un bon isostère de l'oxygène estérifié du phosphate.

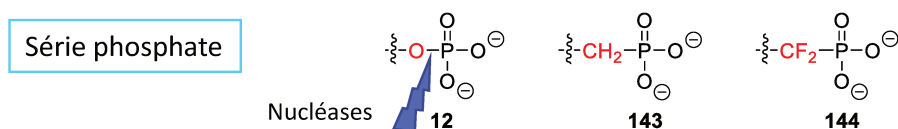


Figure 42 : Modifications en série phosphate

#### IV.3. Apicophilicité

Dans un composé pentacoordiné à géométrie bipyramidale trigonale, l'apicophilicité est définie comme la préférence des substituants à occuper les positions apicales.<sup>99</sup> Ainsi, lors de l'addition d'un nucléophile sur le phosphate naturel **12**, les substituants vont se positionner préférentiellement en position apicale, en fonction de leur nature, pour former le composé pentacoordiné **146** (Schéma 9). Dans son étude théorique, Thatcher a montré que le groupement difluorométhylène **147** avait tendance à également se placer en position apicale.<sup>100</sup> Ainsi, le remplacement d'un oxygène par un motif CF<sub>2</sub> n'aurait pas d'influence significative sur la géométrie des analogues de phosphate.

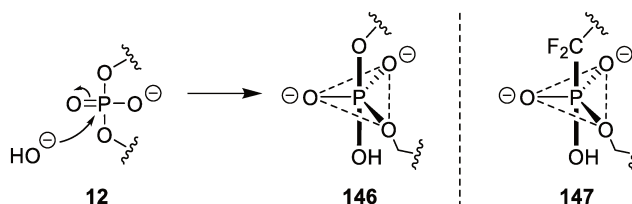


Schéma 9 : Influence du motif CF<sub>2</sub> sur l'apicophilie d'un phosphore pentacoordiné

<sup>97</sup> Baxter, A. D.; Baylis, E. K.; Collingwood, S. P.; Taylor, R. J.; De Mesmaeker, A. and Schmit, C. **1994**, EP0614907A1.

<sup>98</sup> J. Hamilton, C.; M. Roberts, S. and Shipitsin, A. *Chem. Comm.* **1998**, 0, 1087-1088.

<sup>99</sup> Trippett, S. *Phosphorus and Sulfur and the Related Elements* **1976**, 1, 89-98.

<sup>100</sup> Thatcher, G. R. J. and Campbell, A. S. *J. Org. Chem.* **1993**, 58, 2272-2281.

IV.4. Etude conformationnelle

Au sein du laboratoire, une étude combinant analyses RMN et calculs théoriques a été réalisée afin d'étudier l'influence du motif CF<sub>2</sub> en position C3' sur la conformation d'un ribofuranose. Cette étude a montré que l'insertion du motif CF<sub>2</sub> en position C3' déplace l'équilibre conformationnel vers la conformation C3'-*endo* qui est celle privilégiée dans les duplex ARN/ARN (Figure 43).<sup>101</sup>

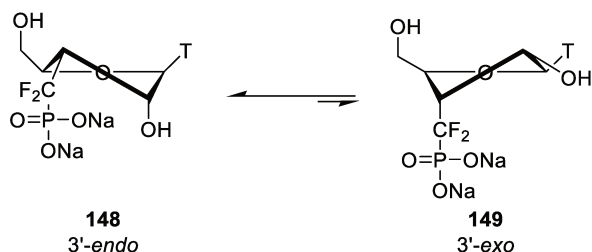


Figure 43 : Conformation du furanose

Dans ces travaux réalisés sur des phosphonates, le motif CF<sub>2</sub> s'est avéré être un bon isostère de l'oxygène. Cependant, aucune étude ne rapporte l'introduction du lien difluorophosphate<sup>102</sup> dans un oligonucléotide afin d'étudier ses propriétés physicochimiques. Ceci peut être expliqué par la présence d'une liaison phosphore-oxygène pouvant être clivée par les nucléases. Ainsi, le laboratoire a transposé ces résultats à la synthèse de phosphinates afin d'étudier des analogues potentiellement non-hydrolysables (**150a** et **150b**) (Figure 44).

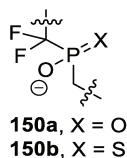


Figure 44 : Gem-difluorophosphate **150a** et gem-difluorophosphinothioate **150b**

<sup>101</sup> Fressigné, C.; Piettre, S.; Condamine, E.; Altona, C. and Gautier, A. *Tetrahedron* **2005**, *61*, 4769-4777.

<sup>102</sup> Pour plus d'informations sur les synthèses et applications des fluorophosphonates, voir : Romanenko, V. D. and Kukhar, V. P. *Chem. Rev.* **2006**, *106*, 3868-3935.

## V. Synthèse du motif *gem*-difluorophosphinate **150a** et *gem*-difluorophosphinothioate **150b**

### V.1. Prérequis

En série phosphinate, Collingwood a réalisé la synthèse d'oligonucléotides comportant un motif phosphinate **64**. Ces composés ont montré une résistance vis-à-vis des nucléases. Cependant, une déstabilisation des duplex oligonucléotide-ARN a été observée puisque le motif CH<sub>2</sub> ne semble pas mimer les propriétés stéréoelectroniques d'un atome d'oxygène (Figure 45).<sup>94</sup>

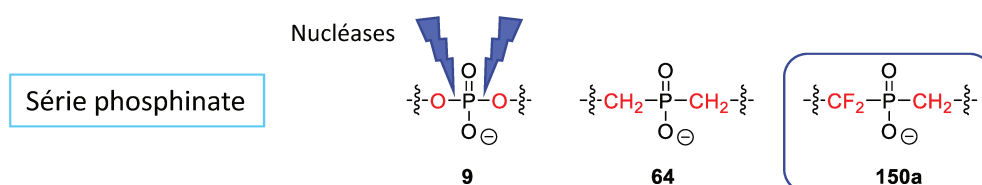


Figure 45 : Modifications en série phosphinate

En combinant les résultats précédents, le laboratoire a posé l'hypothèse que le pont difluorométhylphosphinate **150a** serait un bon bioisostère du pont phosphate naturel et a donc travaillé sur la synthèse de ce motif. Depuis plusieurs années, des méthodologies de synthèse menant au difluorométhylphosphinate ont été mises au point au laboratoire avec pour objectif d'introduire un tel motif dans un dinucléotide en lieu et place d'un pont phosphate naturel (Schéma 10). L'objectif final de ces travaux est d'obtenir le dinucléotide **151**, avec des groupements protecteurs spécifiques (DMT et TBDMS) et sous la forme phosphoramidite, afin que celui-ci soit introduit dans un oligonucléotide par synthèse automatisée dans le but d'étudier ses propriétés physicochimiques (stabilité du duplex, stabilité métabolique).

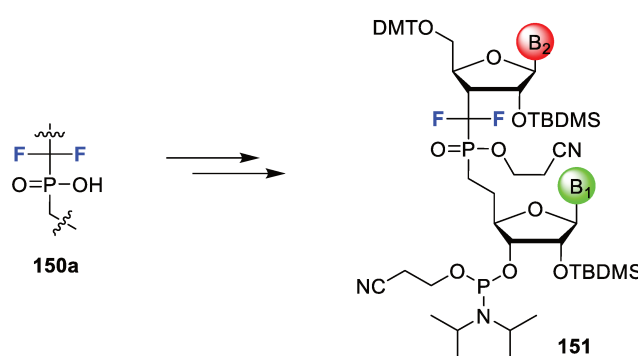


Schéma 10 : Objectif final

### V.2. Développement synthétique menant au motif difluorophosphinate

Le laboratoire travaille donc sur la formation et l'introduction du motif difluorophosphinate **150a** dans un dinucléotide. Il est à noter que celui-ci peut avoir une utilité en stratégie antisens mais également en interférence, ou encore dans le développement de médicaments antiviraux.

La formation du lien difluorophosphinate a d'abord été étudiée sur un substrat modèle.<sup>103</sup> La méthodologie repose sur la construction successive de deux liaisons phosphore-carbone à partir du sel de sodium de l'acide hypophosphoreux **153**. Ainsi, la première liaison phosphore-carbone a été obtenue par addition radicalaire du sel de sodium de l'acide hypophosphoreux **153** sur la difluorooléfine **152**, en présence d'un initiateur radicalaire, le peroxy-pivalate de *tert*-butyle (TBPP). Cette méthode a généré l'acide *H*-phosphinique **154** avec un excellent rendement de 85% (Schéma 11). Ensuite, deux méthodes ont été étudiées afin de former la seconde liaison phosphore-carbone. La labilité de la liaison phosphore-hydrogène a offert la possibilité de générer une seconde fois un radical centré sur l'atome de phosphore et de l'additionner sur un alcène. Ainsi, en présence de TBPP, le composé **154** a été additionné au méthylène cyclohexane **155** pour mener au composé **156**, isolé sous forme d'ester méthylique avec un rendement de 53%. La seconde méthode a impliqué l'utilisation d'un intermédiaire P<sup>III</sup>. L'emploi d'un excès de chlorure de triméthylsilyle a permis de former quantitativement le phosphonite bissilylé **157** correspondant. Puis, une addition 1,4 (réaction d'Abramov) a été réalisée sur la but-3-èn-2-one pour former le composé **158**, isolé une nouvelle fois sous forme d'ester méthylique par action de diazométhane, avec un rendement de 57%. Ces deux stratégies de synthèse complémentaires ont donc permis de former la seconde liaison P-C avec des rendements similaires.

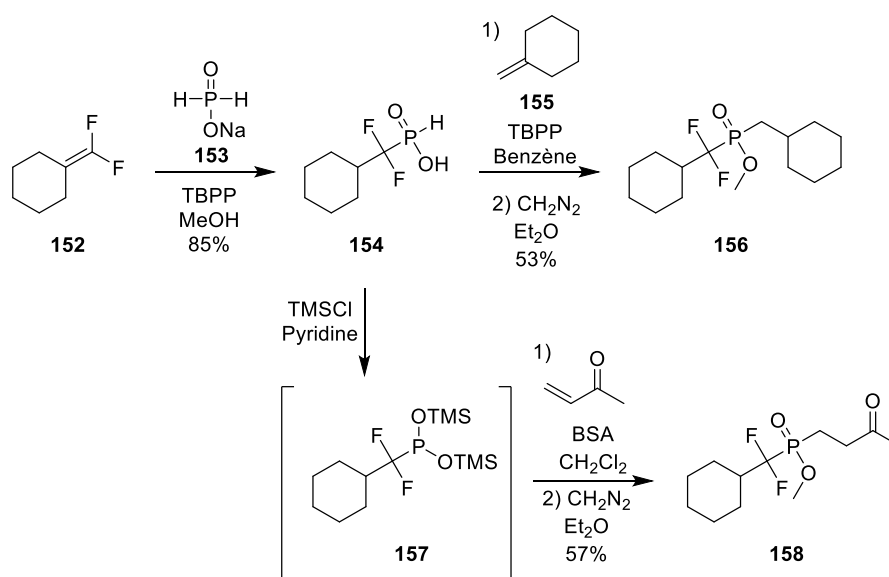


Schéma 11 : Formation du lien difluorophosphinate

Par la suite, cette stratégie de synthèse a été appliquée aux furanoses. Dans un premier temps, la synthèse du furanose *gem*-difluorométhylénique **163** a été réalisée en plusieurs étapes à partir du D-xylose **159** (Schéma 12). Une protection sélective du diol en C1' et C2' sous forme d'acétonide a été réalisée puis, l'alcool terminal en C5' a été protégé à l'aide de chlorure de *para*-chlorobenzoyle avec

<sup>103</sup> Gautier, A.; Garipova, G.; Salcedo, C.; Balieu, S. and Piettre, S. R. *Angew. Chem. Int. Ed.* **2004**, *43*, 5963-5967.

un excellent rendement global de 77%. L'alcool en C3' a été oxydé en cétone **162** à l'aide de TEMPO et d'acide trichloroisocyanurique (ATCC) avec un rendement de 84%.<sup>104</sup> Sur cette cétone, une réaction d'oléfination a été effectuée à l'aide d'HMPT et de dibromodifluorométhane,<sup>105</sup> afin d'obtenir la difluorooléfine **163** avec un rendement global de 44% sur quatre étapes, à partir du D-xylose **159**.

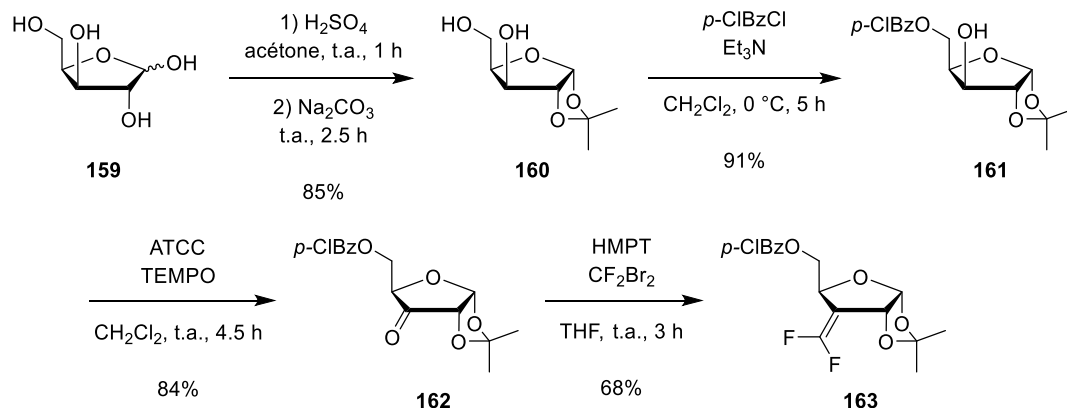


Schéma 12 : Synthèse de la difluorooléfine **163**

L'encombrement stérique du groupement isopropylidène en position 1,2 du furanose **163** a un rôle primordial lors de l'addition radicalaire du sel de l'acide hypophosphoreux. En effet, celui-ci permet de bloquer la face  $\alpha$ , d'induire le transfert d'un atome d'hydrogène en position C3' sur la face  $\beta$  et ainsi de former le stéréoisomère **165a** désiré.<sup>106</sup> La première liaison phosphore-carbone a ainsi été obtenue par une réaction radicalaire entre la difluorooléfine **163** et le sel de sodium de l'acide hypophosphoreux, à l'aide de TBPP, suivie d'une hydrolyse en présence de  $\text{KHSO}_4$ , avec un rendement de 64% (Schéma 13).

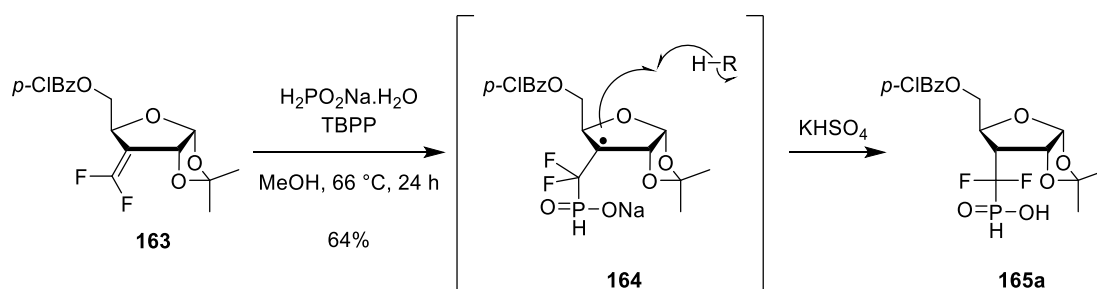


Schéma 13 : Formation de la première liaison P-C

Lors de l'addition radicalaire de l'acide *H*-phosphinique **165a** avec l'alcène **166** en présence de peroxyde de dilauryle (DLP), un mélange inexploitable a été obtenu et la présence du composé-cible

<sup>104</sup> Nomura, M.; Sato, T.; Washinosu, M.; Tanaka, M.; Asao, T.; Shuto, S. and Matsuda, A. *Tetrahedron* **2002**, *58*, 1279-1288.

<sup>105</sup> (a) Houlton, J. S.; Motherwell, W. B.; Ross, B. C.; Tozer, M. J.; Williams, D. J. and Slawin, A. M. Z. *Tetrahedron* **1993**, *49*, 8087-8106. (b) Serafinowski, P. J. and Barnes, C. L. *Tetrahedron* **1996**, *52*, 7929-7938. (c) Flynn, R. M.; Burton, D. J. and Wiemers, D. M. *J. Fluorine Chem.* **2008**, *129*, 583-589.

<sup>106</sup> Lopin, C.; Gautier, A.; Gouhier, G. and Piettre, S. R. *J. Am. Chem. Soc.* **2002**, *124*, 14668-14675.

**168a** n'a pas été observée (Schéma 14).<sup>107</sup> Ce résultat peut être dû à la forte acidité du substrat **165a** pouvant notamment entraîner l'hydrolyse de l'acétonide et donc la formation de sous-produit. Afin de diminuer l'acidité du substrat de départ, le sel de triéthylamine **165b** a été utilisé et a permis d'obtenir **168b** avec des rendements compris entre 35 et 49% mais également le sous-produit **169b** et le phosphonate **170b**. La formation de **169b** peut être expliquée par l'association d'un radical centré sur l'adduit carboné **167b** et du radical généré lors du clivage homolytique de DLP.

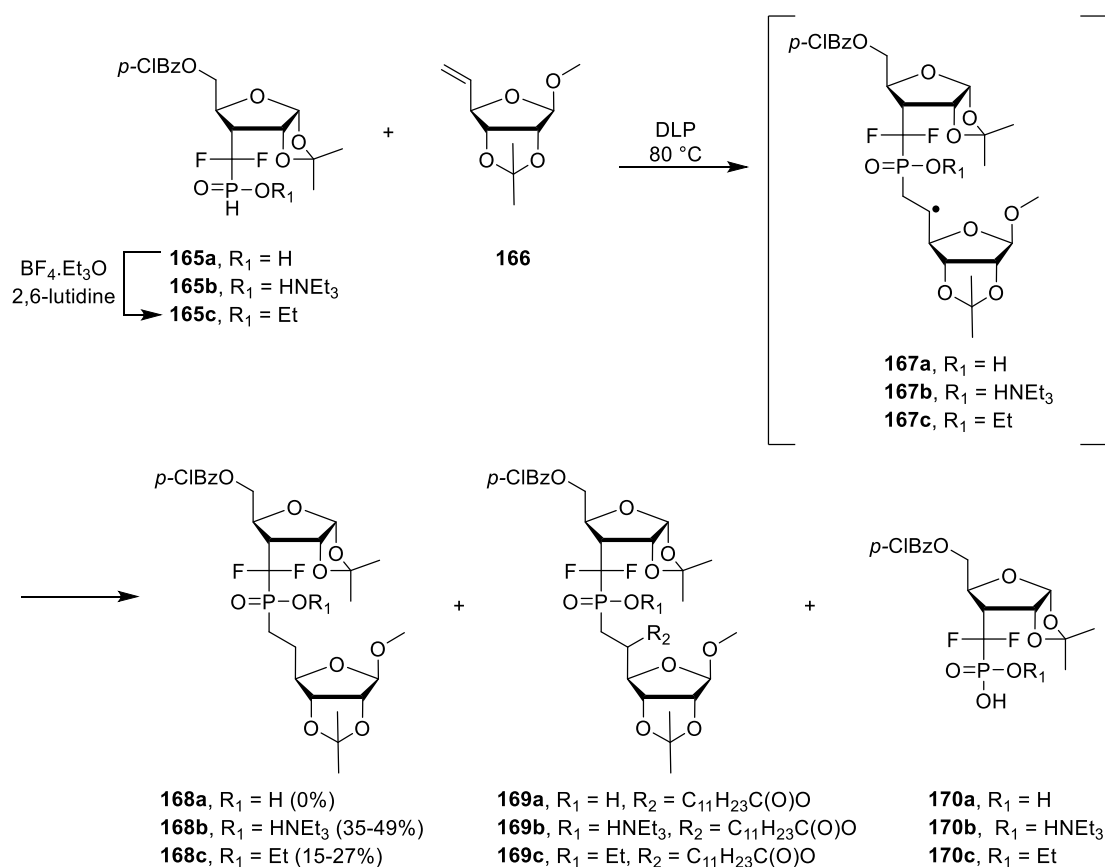
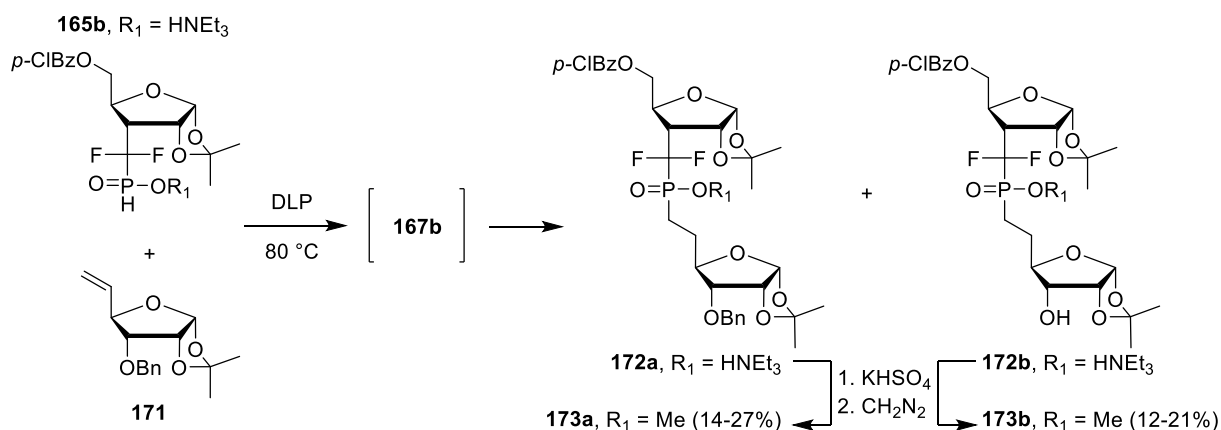


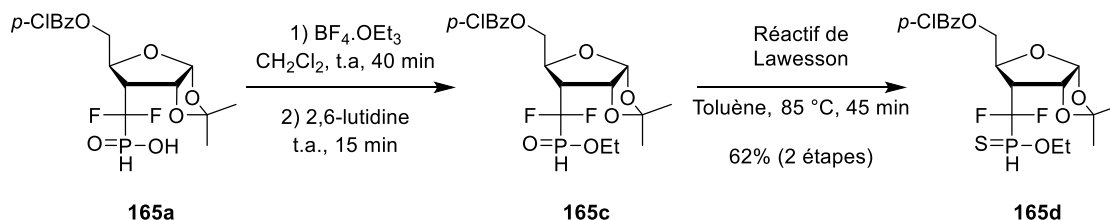
Schéma 14 : Etude de l'addition radicalaire de **165** sur l'oléfine **166**

Un résultat similaire a été observé lors de l'addition radicalaire de **165b** sur le furanose **171** (Schéma 15). En effet, après acidification ( $\text{KHSO}_4$ ) et méthylation à l'aide de diazométhane, les composés **173a** et **173b** ont été isolés avec des rendements respectifs compris entre 14-27% et 12-21%. La formation de **173b** a résulté d'une débenzylation de l'alcool en C3' du furanose du bas provenant d'un transfert intramoléculaire [1,5] d'hydrogène par le radical initial **167b** suivi d'une fragmentation de la liaison benzylique générant **173b** (après acidification et estérification à l'aide de diazométhane).

<sup>107</sup> Zhang, J.; Lambert, E.; Xu, Z.-F.; Brioché, J.; Remy, P. and Piettre, S. R. *J. Org. Chem.* **2019**, *84*, 5245-5260.

Schéma 15 : Addition radicalaire de **165b** sur **171**

Dans le but de réaliser la réaction radicalaire avec un substrat neutre, l'acide *H*-phosphinique **165a** a été estérifié à l'aide du sel de Meerwein en *O*-éthyl-*H*-phosphinate **165c** (Schéma 14). Celui-ci a directement été engagé dans l'étape d'addition radicalaire du fait de sa faible stabilité (facilement oxydable en acide correspondant (**170c**)). Malgré une réactivité accrue, le composé-cible **168c** n'a été obtenu qu'avec des faibles rendements (15-27%) en mélange avec les sous-produits **169c** et **170c**. La liaison P-H relativement forte des composés **165a-c** est à l'origine de la formation de sous-produits et des faibles rendements. Ainsi, en s'appuyant sur des travaux de Piettre,<sup>108</sup> la synthèse des *O*-éthyl-*H*-phosphinothioates analogues a été envisagée afin d'augmenter la labilité de la liaison P-H. En effet, l'énergie de dissociation de la liaison P-H des composés P=S est plus faible que celle des composés P=O, rendant l'utilisation des *H*-phosphinothioates susceptibles d'éviter ces réactions secondaires.<sup>109</sup> Ainsi, une synthèse du motif difluoro-*H*-phosphinothioate, jusqu'alors non-décrite dans la littérature, a été mise au point. L'acide **165a** a été estérifié à l'aide du sel de Meerwein (Schéma 16). Puis, une thiolation subséquente de l'ester **165c** a été réalisée à l'aide du réactif de Lawesson. Le difluoro-*H*-phosphinothioate **165d** a donc été obtenu en sept étapes, avec un rendement global de 17% à partir du D-xylose.

Schéma 16 : Synthèse du *H*-phosphinothioate **165d**

<sup>108</sup> Piettre, S. R. *Tetrahedron Lett.* **1996**, *37*, 4707-4710.

<sup>109</sup> Jessop, C. M.; Parsons, A. F.; Routledge, A. and Irvine, D. J. *Eur. J. Org. Chem.* **2006**, *2006*, 1547-1554.

La réaction radicalaire entre le *H*-phosphinate **165c** et l'alcène **166** a conduit au difuranose **168c** avec un faible rendement (27%) (Schéma 17). Par contre, la même réaction d'addition radicalaire, menée avec le *H*-phosphinothioate **165d**, a généré le phosphinothioate analogue **168d** avec un rendement de 71%.

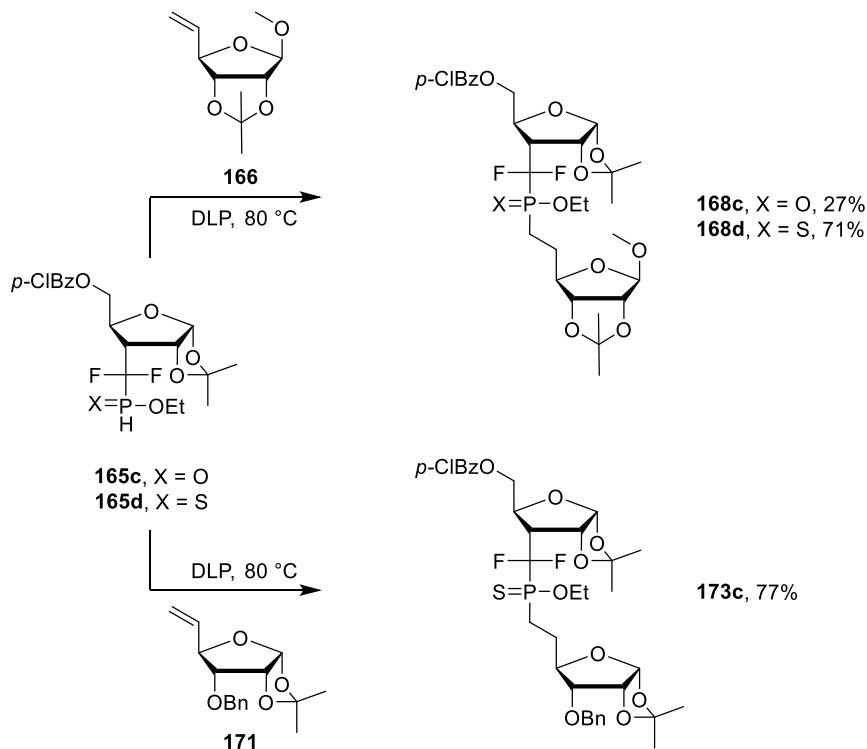


Schéma 17 : Formation de la seconde liaison phosphore-carbone

Le bon déroulement de la chaîne radicalaire résulte de trois étapes distinctes : initiation, propagation, terminaison (Schéma 18). L'étape d'initiation va entraîner la formation du radical phosphoré **174**. Ce dernier va réagir sur l'oléfine **166** pour permettre la formation du radical-adduit **167**. Celui-ci va arracher l'hydrogène de **165** pour former le composé-cible **168**, générer un nouveau radical phosphoré **174** et permettre, ainsi, la propagation de la chaîne radicalaire.

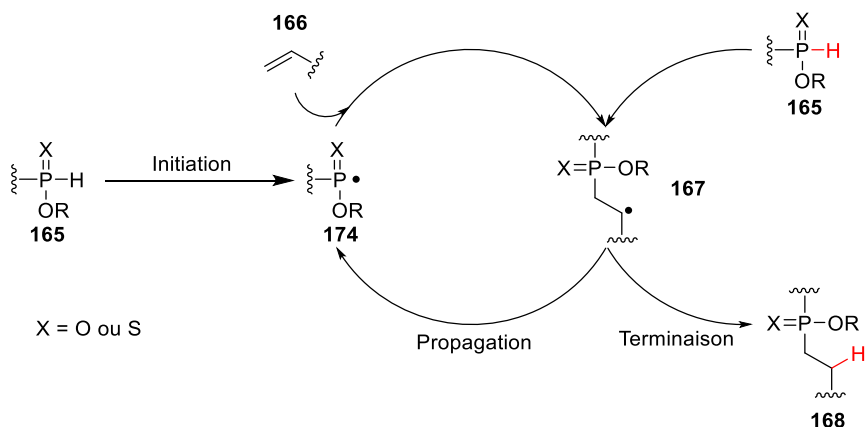


Schéma 18 : Déroulement de la chaîne radicalaire pour former **168**

En partant du principe que pour une même température, chaque étape est thermodynamiquement favorable dans le cas du P=O ou du P=S, la différence de réactivité se ferait donc au niveau des énergies d'activation menant aux états de transition des différentes étapes élémentaires. Expérimentalement, seuls des sous-produits provenant du radical-adduit P=O **167c** ont été observés. Par conséquent, on peut en déduire que c'est l'énergie d'activation nécessaire à l'étape du transfert d'hydrogène entre **167c** et **165** qui est en cause (A, Figure 46). Dans le cas du phosphinate **167c**, cette dernière  $E_a^{(P=O)}$  serait plus grande que l'énergie d'activation  $E_a^{\text{sous-produit}}$  responsable de la formation des sous-produits et que celle  $E_a^{(P=S)}$  du phosphinothioate **167d** permettant la formation de **168d**. Une explication pourrait être apportée à l'aide des orbitales frontières (B, Figure 46). La haute énergie d'activation du P=O **167c** résulterait d'une différence énergétique  $\Delta E^{(P=O)}$  importante entre les orbitales SOMO de **167c** et HOMO de **165c**. Cette différence se traduirait en un état de transition déstabilisé ne favorisant pas le bon déroulement de la chaîne et expliquant la formation des sous-produits. Dans le cas du *H*-phosphinothioate **165d**, une diminution du  $\Delta E^{(P=S)}$  entre la SOMO de **167d** et la HOMO de **165d** générerait une stabilisation supérieure de l'état de transition résultant en une énergie d'activation plus faible pour l'étape d'arrachage de l'hydrogène de **165d**. L'augmentation de vitesse de cette étape (Arrhénius)<sup>110</sup> a pour conséquence de rendre la réaction plus sélective et de supprimer les processus secondaires observés dans le cas de **165c**.

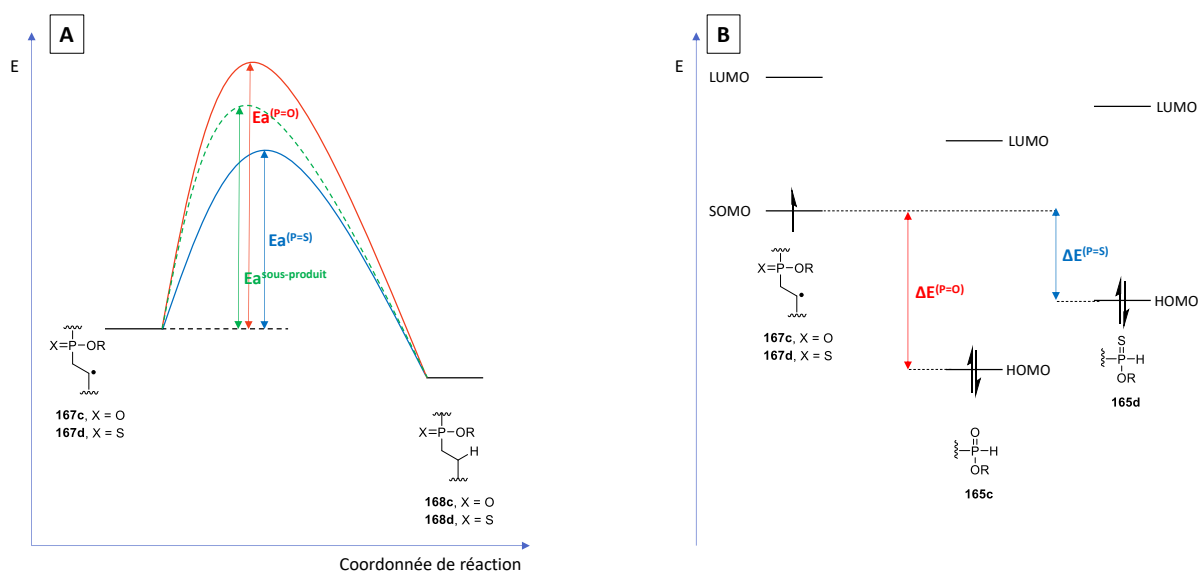


Figure 46 : (A) Comparaison des énergies d'activation nécessaire à la formation de **168**  
(B) Différence de réactivité entre **165c** et **165d** expliquée à l'aide des orbitales frontières

<sup>110</sup> Arrhenius, S. *Selected Readings in Chemical Kinetics* **1967**, 31-35.

Cette synthèse a donc été appliquée au furanose-cible **171** permettant de synthétiser le difuranose **173c** avec un très bon rendement (Schéma 17). Il est à noter que l'utilisation du *H*-phosphinothioate **165d** à la place du *H*-phosphinate **165c** a également permis d'éliminer le phénomène de débenzylation, pour des raisons analogues.

La réaction de Michaelis-Becker a également été envisagée afin de former la seconde liaison phosphore-carbone par une substitution nucléophile impliquant l'anion *H*-phosphinothioate **165d** et un partenaire électrophile. Cette réaction avait déjà été étudiée sur un substrat modèle permettant une optimisation des conditions.<sup>111</sup> En présence d'hydroxyde de potassium et d'un catalyseur de transfert de phase, le chlorure de benzyltriéthylammonium (TEBAC), la substitution nucléophile de l'anion phosphinate **175** sur l'iodure d'alkyle **176** a été réalisée. (Schéma 19). Notons que ces travaux précédents avaient déjà montré que l'utilisation d'un partenaire électrophile alkyle iodé était nécessaire pour réaliser efficacement la réaction de Michaelis-Becker. En effet, ceux-ci se sont révélés plus réactifs que les composés bromés et chlorés analogues et de meilleurs rendements ont été obtenus. Comme pour la formation de la seconde liaison P-C par addition radicalaire, une différence de réactivité avait également été observée entre le P=O **175a** et le P=S **175b**. La seconde liaison phosphore-carbone **177** avait aussi pu être créée avec un excellent rendement de 89 % à partir du *H*-phosphinothioate **175b** contre 13% à partir du *H*-phosphinate **175a**.

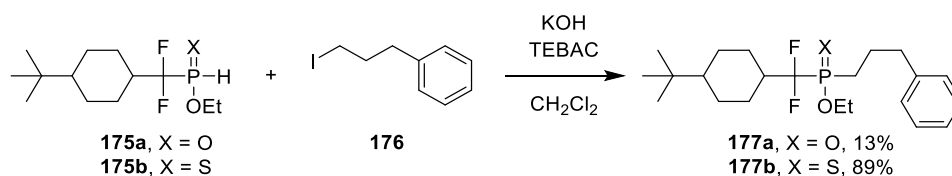


Schéma 19 : Formation de la seconde liaison phosphore-carbone par voie ionique

Ces conditions ont par la suite été appliquées à la formation du difuranose **173c** (Schéma 20).<sup>111</sup> Celui-ci a été obtenu par substitution nucléophile de l'iode dans **178** par l'anion du phosphinothioate **165d** avec un excellent rendement de 86%.

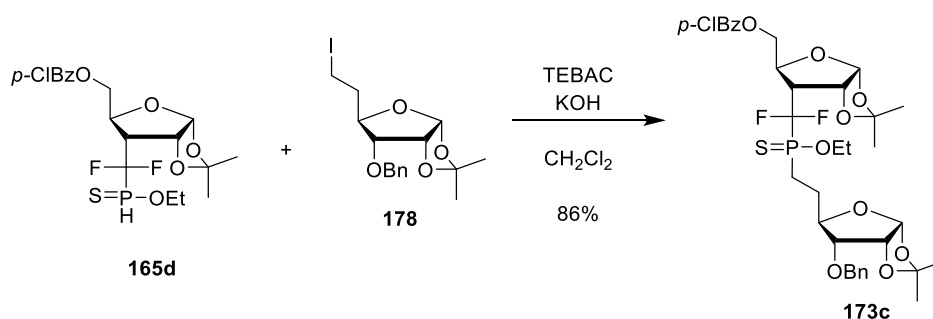
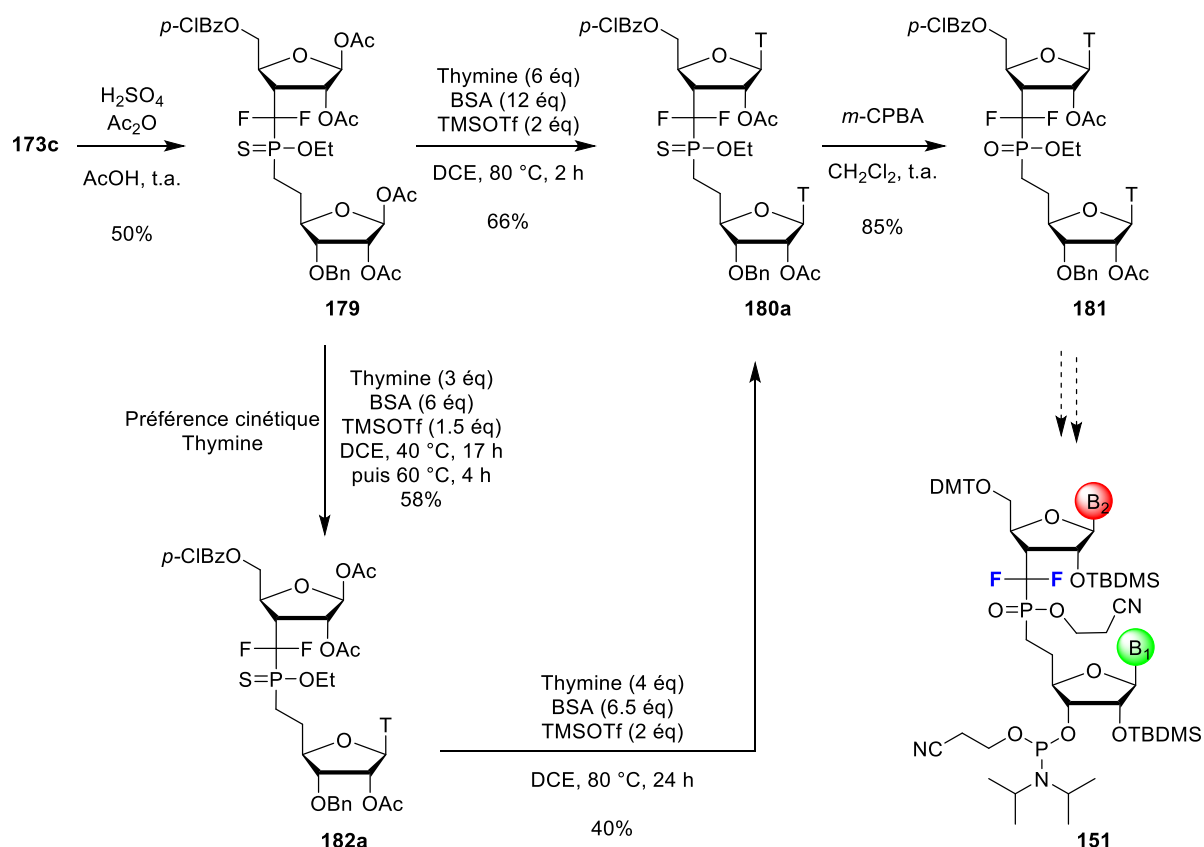


Schéma 20 : Synthèse du difuranose **173c** par voie ionique

<sup>111</sup> Thèse Dr. Emilie Lambert, Université de Rouen, *Préparation de phosphinates difluorés, isostères des diesters de phosphate. Application à la synthèse de dinucléotides modifiés*. Soutenue le 17 mai 2013.

A partir de l'intermédiaire **173c**, l'acétolyse des deux acétonides et la protection subséquente des alcools C1'-C2' résultants sous forme d'acétate ont été réalisées en milieu acide pour obtenir le tétraacétate **179** (Schéma 21). Deux thymine ont ensuite été introduites selon les conditions de Vorbrüggen afin de former le dinucléotide **180a**. Le motif final **181** a été obtenu par oxydation du P=S en P=O à l'aide d'acide *m*-chloroperbenzoïque (*m*-CPBA). Une préférence cinétique sur le furanose du bas a été observée lors de l'introduction d'une seule thymine conduisant à la formation d'un intermédiaire nucléotide-furanose **182a** avec un rendement de 58% (Dr. Jun Zhang, travaux non publiés du laboratoire). Cette préférence cinétique peut être expliquée soit par un encombrement stérique d'origine conformationnelle, soit par une stabilisation de l'oxocarbénium formé sur le furanose du haut par assistance anchimérique du soufre, ralentissant ainsi l'introduction de la base nucléique sur ce furanose.<sup>112</sup> Pour finir, une seconde thymine a ensuite été introduite pour former le même dinucléotide **180a** avec un rendement de 40%.

Schéma 21 : Synthèse menant au motif difluorophosphinate **150a**

<sup>112</sup> Yokomatsu, T.; Sada, T.; Shimizu, T. and Shibuya, S. *Tetrahedron Lett.* **1998**, *39*, 6299-6302.

Cette première stratégie valide la synthèse d'un dinucléotide modifié caractérisé par un pont difluorophosphinyle et par deux bases nucléiques identiques. De plus, l'obtention de l'intermédiaire **182a** permet d'entrevoir l'introduction d'une seconde base nucléique différente et donc l'accès à des hétérodinucléotides.<sup>113</sup> Cependant, cette stratégie manque d'efficacité (rendement global de 5% en onze étapes à partir du D-xylose) puisque quelques étapes de synthèse doivent encore être réalisées afin d'obtenir le dinucléotide-cible **151**. Par ailleurs, quelques centaines de milligrammes de **151** doivent être produit avant d'être introduit dans un oligonucléotide, la formation du duplex, incorporant ce dernier, permettant *in fine* de réaliser les études physicochimiques sur celui-ci et, ainsi, de déterminer l'apport du motif difluorophosphate à la stabilisation de ce duplex.

L'accès aux hétérodinucléotides, par l'intermédiaire de la synthèse de première génération, est fondé sur la préparation d'un intermédiaire mixte de type nucléotide-furanose **182** reposant sur un phénomène de préférence cinétique non-garanti avec les bases nucléiques naturelles ou modifiées.

L'objectif de cette thèse est donc de mettre au point une synthèse de seconde génération afin d'améliorer l'efficacité synthétique mais également de permettre un accès simple aux homodinucléotides<sup>114</sup> **180** et hétérodinucléotides **183** (Figure 47).

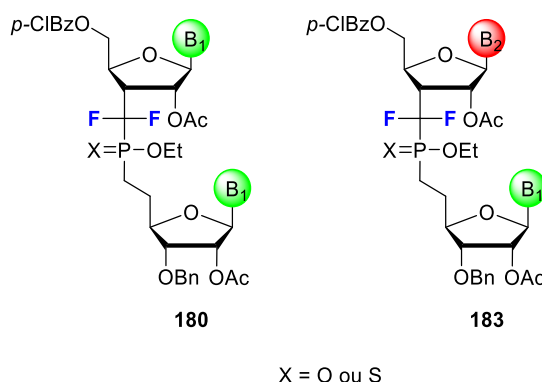


Figure 47 : Dinucléotides-cibles

<sup>113</sup> Hétérodinucléotide : dinucléotide possédant deux bases nucléiques différentes.

<sup>114</sup> Homodinucléotide : dinucléotide possédant deux bases nucléiques identiques.

En nous limitant aux bases nucléiques présentes dans l'ADN, seize combinaisons de dinucléotides sont possibles (Tableau 6).

Tableau 6 : Combinaison de dinucléotides possibles

Nucléobases	Thymine	Cytosine	Adénine	Guanine
<b>Thymine</b>	TT	TC	TA	TG
<b>Cytosine</b>	CT	CC	CA	CG
<b>Adénine</b>	AT	AC	AA	AG
<b>Guanine</b>	GT	GC	GA	GG



# Chapitre 1

*Rétrosynthèses*



Dans ce chapitre seront présentées les analyses rétrosynthétiques envisagées menant à la formation d'un hétérodinucléotide incorporant le motif difluorophosphinate.

Chaque voie rétrosynthétique présentée repose sur les savoir-faire du laboratoire, c'est-à-dire (i) la synthèse du motif difluorophosphin(othio)ate par réaction séquentielle du sel de sodium de l'acide hypophosphoreux avec, d'abord, un difluoroalcène (formation de la liaison  $CF_2-P$ ), et, ensuite, soit un alcène, soit un dérivé iodé (formation de la liaison  $P-CH_2$ ) ; et (ii) en exploitant l'efficacité synthétique supérieure des dérivés  $P=S$  par rapport aux analogues  $P=O$ . Cinq stratégies différentes susceptibles de mener aussi bien aux homodinucléotides qu'aux hétérodinucléotides ont ainsi été mises en évidence (Figure 48). Ces stratégies focalisent sur la formation de la seconde liaison phosphore-carbone menant à la formation du pont phosphinothioate. Une première (voie A) prévoit l'installation séquentielle des deux bases après la formation du lien difluorophosphinothioate. Deux autres (voies B et C) reposent sur l'introduction de la seconde base après la formation du pont. Et les deux dernières (voies D et E) envisagent la construction du pont à partir de deux fragments contenant déjà les bases nucléiques.

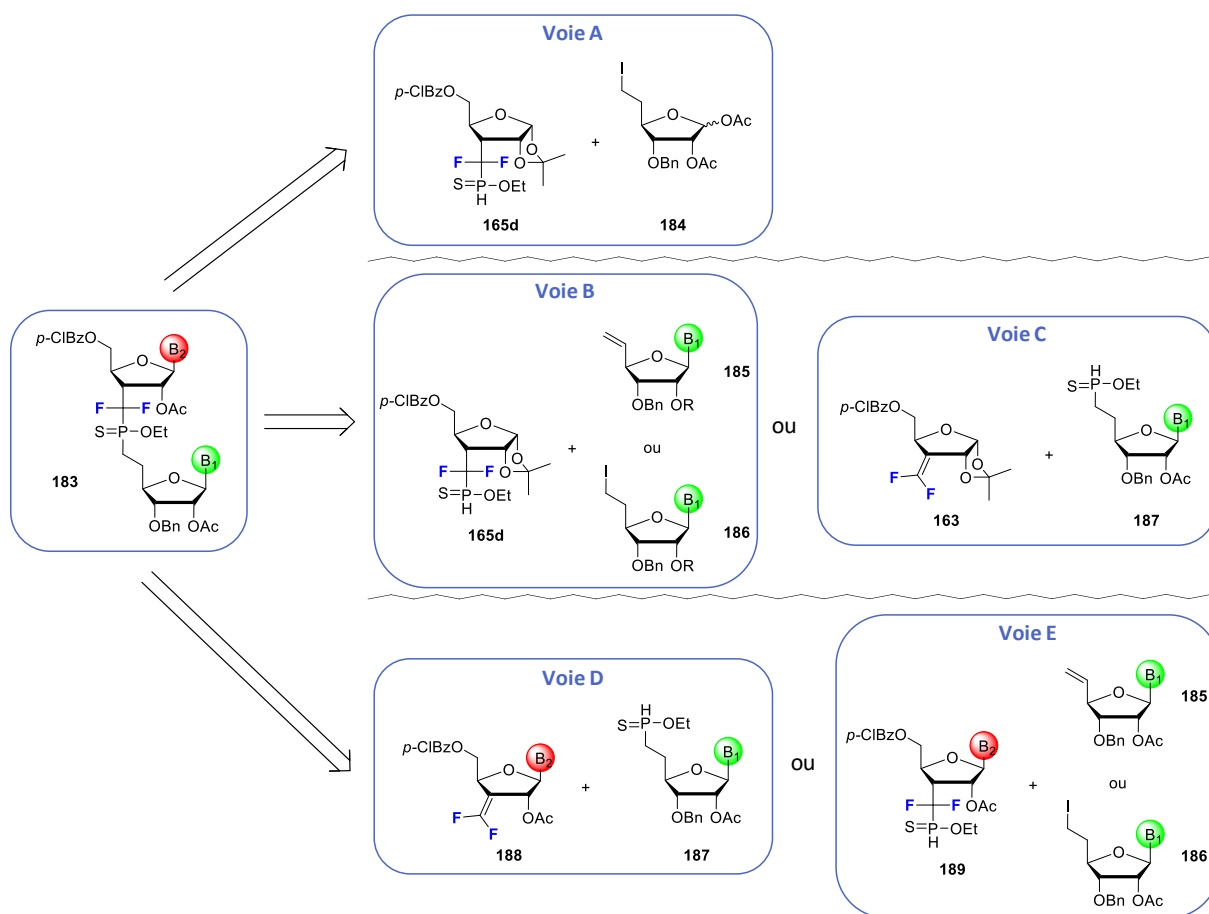


Figure 48 : Nouvelles stratégies envisagées

Par la suite, chaque voie envisagée sera présentée plus en détail et les avantages et inconvénients de chacune d'elles seront discutés.

## I. Voie A

Cette stratégie est parallèle aux travaux de Lebreton dont le but est de synthétiser des hétérodinucléotides comportant un pont amide. Cette approche est fondée sur l'introduction successive des bases nucléiques après la construction du lien amide grâce à des protections différentes du diol C1'-C2' de chaque furanose.<sup>115</sup> L'intermédiaire **192** a été formé avec un rendement de 65%, à partir de deux furanoses, l'un protégé sous la forme acétonide **190** et l'autre sous la forme bisacétate **191** (Schéma 22). Ainsi, une thymine a pu être insérée sur le furanose bisacétate selon les conditions de Vorbrüggen avec un rendement de 89%. Après acétylolyse de l'acétonide **193** en milieu acide et protection subséquente des alcools C1'-C2' résultants sous forme d'acétates, un uracile a été inséré sur le second furanose avec un rendement de 83% sur deux étapes. Cette stratégie a permis d'obtenir trois dinucléotides, deux homodinucléotides (T-T et U-U) et un hétérodinucléotide **195** (T-U).

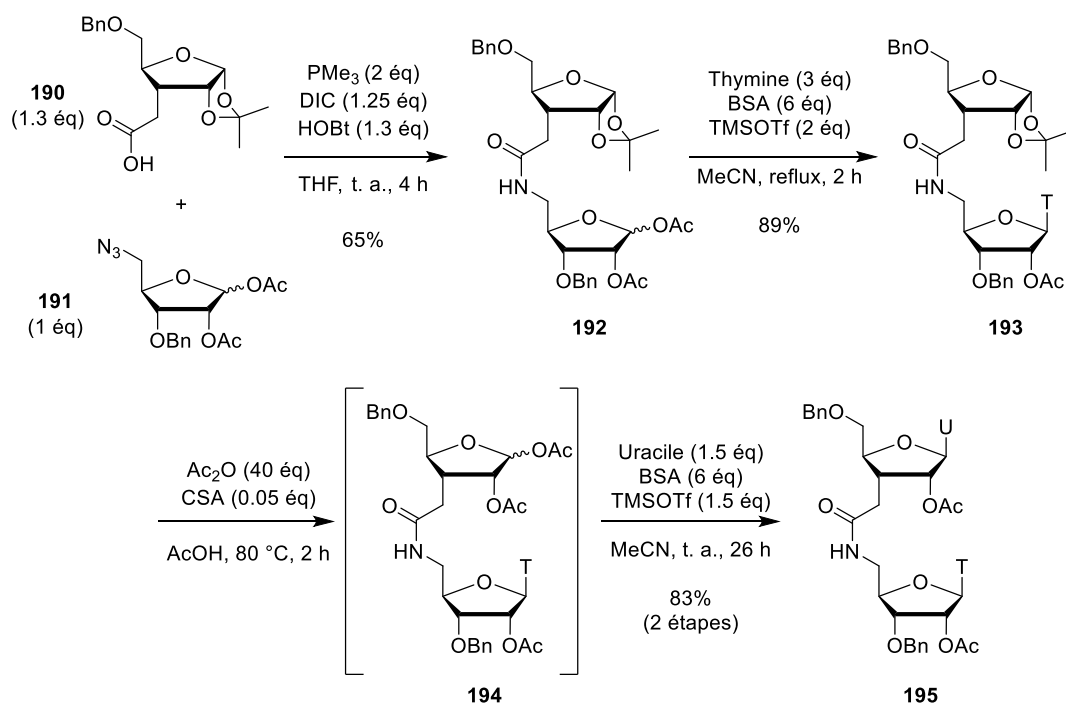


Schéma 22 : Stratégie de Lebreton

Dans notre cas, l'hétérodinucléotide-cible **183**, par une disconnection au niveau de la position anomérique portant la base nucléique  $B_2$ , serait obtenu à partir d'un synthon pouvant correspondre au triacétate **182** (Schéma 23). Celui-ci proviendrait d'une modification du groupement protecteur du diol C1'-C2' du furanose du haut et similairement à **183**, une coupure de la position anomérique portant, cette fois, la nucléobase  $B_1$  pourrait être réalisée. Ainsi, le triacétate **182** serait formé à partir du substrat bisacétate-acétonide **196**. En réalisant la disconnection de la liaison P- $\text{CH}_2$ , l'étape-clé de

<sup>115</sup> Arzel, L.; Dubreuil, D.; Dénès, F.; Silvestre, V.; Mathé-Allainmat, M. and Lebreton, J. *J. Org. Chem.* **2016**, *81*, 10742-10758.

cette rétrosynthèse, entraînant la formation de **196**, résulterait de la combinaison de deux synthons dont les équivalents synthétiques pourraient être le *H*-phosphinothioate **165d** et l'iodure d'alkyle **184**.

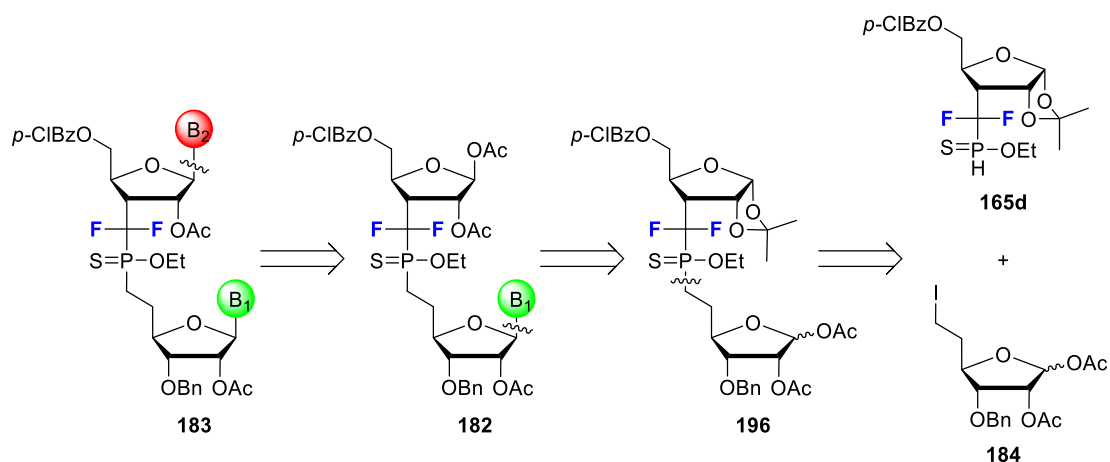


Schéma 23 : Rétrosynthèse voie A

Le difluoro-*H*-phosphinothioate **165d**, déjà synthétisé au laboratoire, pourrait ainsi être associé par voie ionique au fragment bisacétate **184**, dont la synthèse reste à mettre au point, pour différencier nos deux furanoses. Cette stratégie s'est montrée efficace pour la synthèse d'hétérodinucléotides avec un pont amide. Elle pourrait donc également l'être en présence d'un pont difluorophosphinothioate et générerait des composés analogues au composé mixte nucléotide-furanose **182** obtenue par le Dr. Jun Zhang au laboratoire (Introduction, voir Schéma 21, page 69).

D'emblée, il faut noter qu'à la fois l'approche développée ci-dessus et celle du laboratoire (Schéma 20 et Schéma 21) se rejoignent par l'introduction tardive des bases nucléiques dans la synthèse : chaque approche possède donc un intermédiaire de type bisfuranoside **179** et **196**, commun aux seize différents homodinucléotides et hétérodinucléotides ; cet intermédiaire est proche des produits finaux.

## II. Voie B

Le début de cette rétrosynthèse est similaire à celle proposée pour la voie A jusqu'à la formation du triacétate **182**. Dans ce cas, ce dernier serait obtenu à partir d'un changement du groupement protecteur du diol C1'-C2' et pourrait provenir du substrat nucléotide-acétonide **197**. Une disconnection de la liaison P-CH<sub>2</sub> entraînerait la formation de **197** à partir de synthons provenant du *H*-phosphinothioate **165d** et d'un substrat, où une base nucléique serait déjà en place, de type oléfinique **185** ou iodé **186**.

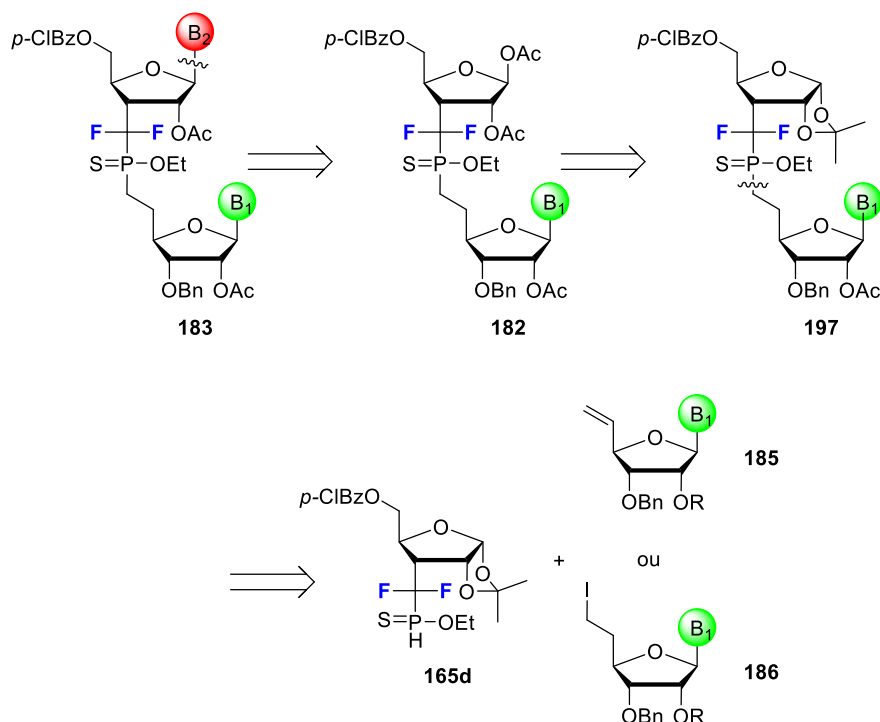


Schéma 24 : Rétrosynthèse voie B

Fondée sur les acquis du laboratoire, cette stratégie permettrait l'assemblage du *H*-phosphinothioate **165d** à un fragment comportant une base nucléique par formation de la liaison P-CH<sub>2</sub> (Schéma 24). La seconde liaison phosphore-carbone pourrait être obtenue en suivant des conditions développées au laboratoire, soit par addition radicalaire avec l'oléfine **185**, soit par réaction avec l'iodure d'alkyle **186**. Les synthèses des huit nucléosides **185** et **186** sont donc à développer.

Si la voie A et B se rejoignent au niveau du fragment **182**, il faut cependant noter ici que la voie B nécessite de synthétiser indépendamment quatre nucléosides de type **185** ou **186**, porteur chacun de la base désirée. Cette approche implique donc la préparation de cinq substrats distincts avant couplage.

III. Voie C

A la différence des deux précédentes rétrosynthèses, une disconnection au niveau de la liaison P-CF<sub>2</sub> générerait des synthons pouvant se traduire en difluorooléfine **163** et *H*-phosphino(thio)ate **187** (Schéma 25). A noter que l'addition radicalaire de **187** sur **163** générerait le radical-adduit **198**.

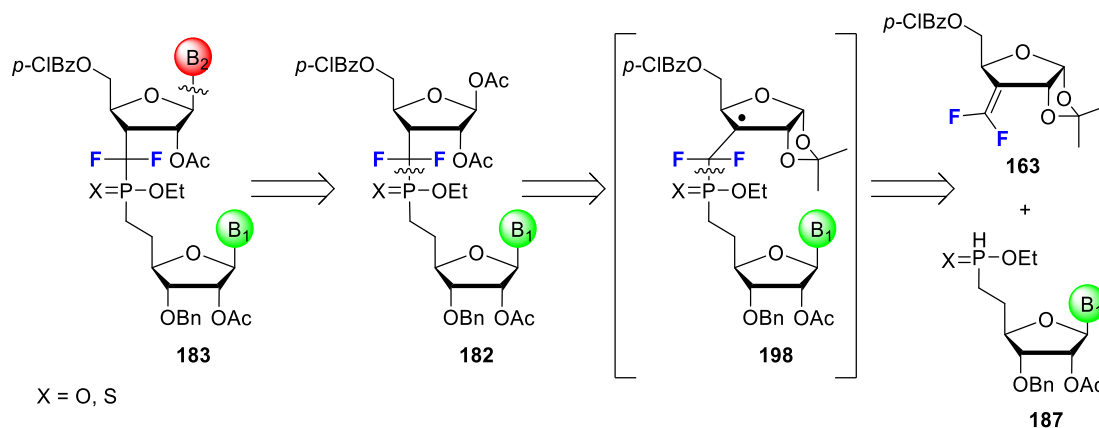


Schéma 25 : Rétrosynthèse voie C

La formation du pont difluorophosphinothioate est ici envisagée via une approche non-documentée dans la littérature. En revanche, il est à noter que la synthèse d'un motif difluorophosphinate a été rapportée par Piettre entre un *H*-phosphinate et une difluorooléfine modèle.<sup>116</sup> Cette stratégie repose donc sur la formation d'une liaison P-CF<sub>2</sub> en vue d'accéder au motif difluorophosphin(othio)ate via une addition radicalaire entre le *H*-phosphin(othio)ate **187** et la difluorooléfine **163** (synthèse déjà réalisée au laboratoire). Cette stratégie permettrait également de réaliser une étude sur la différence de réactivité entre H-P=O et H-P=S ainsi que sur la stéréosélectivité observée en position C3' lors de l'addition radicalaire de **187** sur **163**. Les travaux développés au laboratoire indiquent qu'une hydrogénation du radical-adduit **198** par la face β devrait être préférée (Schéma 13, page 63).

Ainsi dans cette approche, la synthèse de cinq substrats distincts est nécessaire avant de réaliser l'étape de couplage, comme pour la voie B.

<sup>116</sup> Piettre, S. R. *Tetrahedron Lett.* **1996**, 37, 2233-2236.

IV. Voie D

En comparaison avec les premières approches, l'hétérodinucléotide-cible **183** serait obtenu par une disconnection au niveau de la liaison P-CF<sub>2</sub>. Cette dernière générerait donc deux synthons portant des bases nucléiques et pouvant correspondre à la difluorooléfine **188** et le *H*-phosphino(thio)ate **187** (Schéma 26).

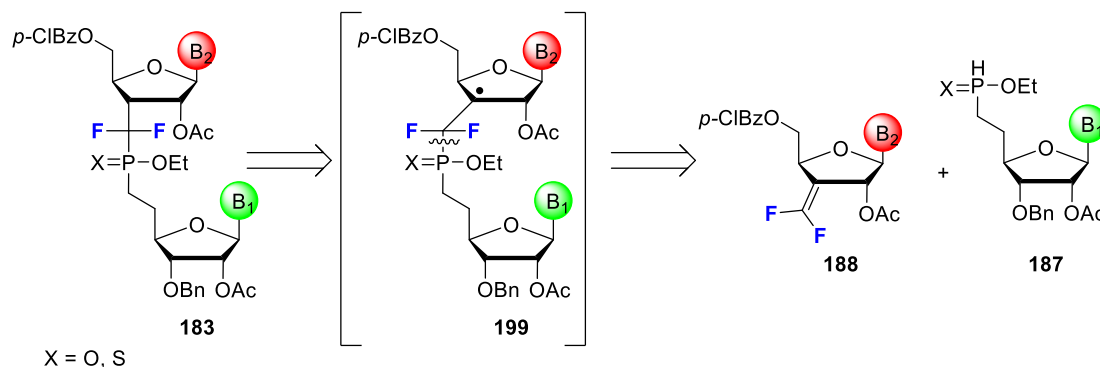


Schéma 26 : Rétrosynthèse voie D

Cette approche nécessite le développement d'une méthodologie de synthèse afin d'accéder au motif difluorophosphin(othio)ate par addition radicalaire de **187** sur **188** formant un radical-adduit **199**. Cette stratégie est similaire à celle présentée dans la voie C et nécessite la préparation de huit substrats différents avant couplage et formation du motif difluorophosphi(othio)ate.

De plus, lorsque l'on compare la voie C à la voie D, une interrogation subsiste au sujet de la stéréochimie du centre stéréogène en C3' résultant de l'addition radicalaire de **187** sur **188**. En effet, d'après une étude menée par Matsuda, lors de la désoxygénation en position C3' du nucléoside **200** par réaction de Barton-McCombie, un radical tertiaire centré en C3' est formé (**201**, Schéma 27).<sup>117</sup> Le transfert d'hydrogène entre le radical-adduit **201** et Bu<sub>3</sub>SnH génère le composé **202**, obtenu sous la forme d'un mélange d'épimères en C3' inséparables (ratio 4/1). Ce transfert prend place majoritairement par la face alpha du furanose dans le cas de Matsuda. Cette voie D impliquerait également la formation d'un radical centré en C3' et l'obtention d'un épimère analogue est fort probable. Dans notre cas, la configuration absolue en C3' souhaitée est analogue à celle de l'épimère minoritaire **202b**. Ces résultats indiquent clairement l'effet stéréodirecteur délétère que risque de jouer la base nucléique présente dans **188**.

<sup>117</sup> Hattori, H.; Nozawa, E.; Iino, T.; Yoshimura, Y.; Shuto, S.; Shimamoto, Y.; Nomura, M.; Fukushima, M.; Tanaka, M.; Sasaki, T. and Matsuda, A. *J. Med. Chem.* **1998**, *41*, 2892-2902.

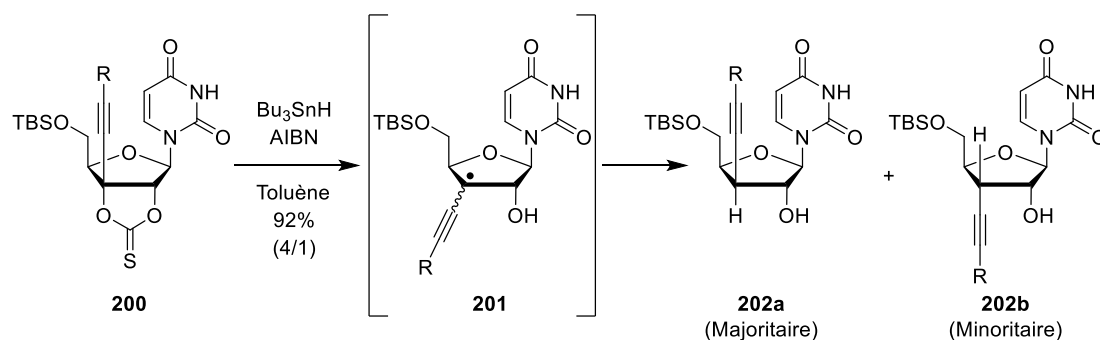


Schéma 27 : Résultats de l'étude menée par Matsuda

En revanche, une réaction similaire a été effectuée à partir d'un furanose possédant un acétonide sur les positions C1' et C2'. Dans ce cas, un seul épimère en C3' a été obtenu via le radical tertiaire centré C3' **204**. En effet, dans les travaux de Redlich, une désoxygénation de Barton-McCombie a été réalisée sur **203** afin de réduire la position C3' (Schéma 28).<sup>118</sup> Le composé-cible **205** a été obtenu sous l'unique configuration allofuranose. Les additions radicalaires réalisées au laboratoire sur le difluoroalcène **163** exploitent déjà l'effet stéréodirecteur d'un tel acétonide (voir Schéma 13, page 63).

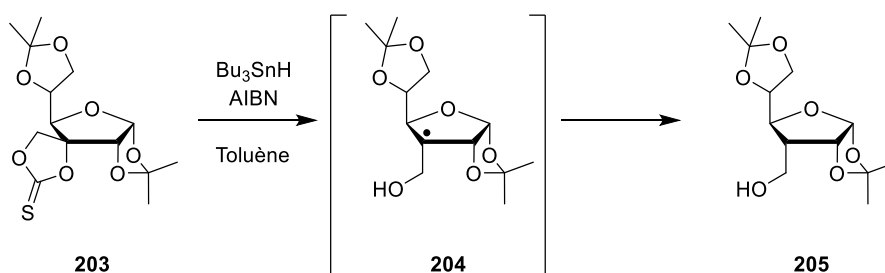


Schéma 28 : Travaux de Redlich

Cette quatrième approche, qui repose sur la construction de la seconde liaison P-C par une voie radicalaire, montre la nécessité de la présence de l'acétonide afin d'induire la stéréochimie désirée en C3'. Ceci rend cette approche D non viable.

<sup>118</sup> Redlich, H.; Sudau, W. and Paulsen, H. *Tetrahedron* **1985**, *41*, 4253-4260.

## V. Voie E

Afin d'éviter le problème de sélectivité en C3' de la voie D, la voie E a été envisagée (Schéma 29). La formation de l'hétérodinucléotide **183** résulterait, cette fois-ci, d'une disconnection au niveau de la liaison P-CH<sub>2</sub> et donc de l'association de synthons correspondant au *H*-phosphinothioate **189** et à un second fragment incorporant une base nucléique **185** ou **186**.

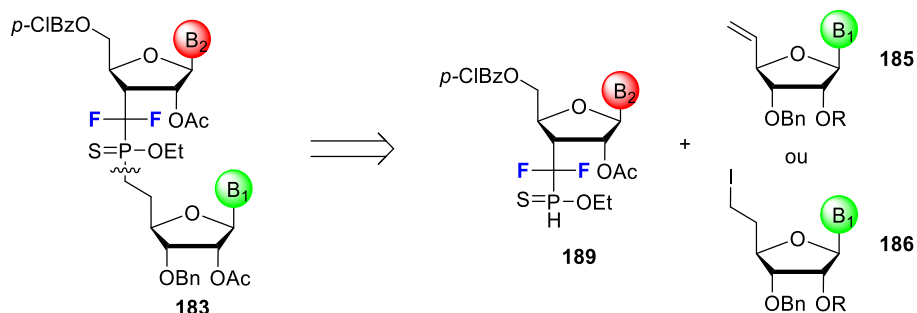
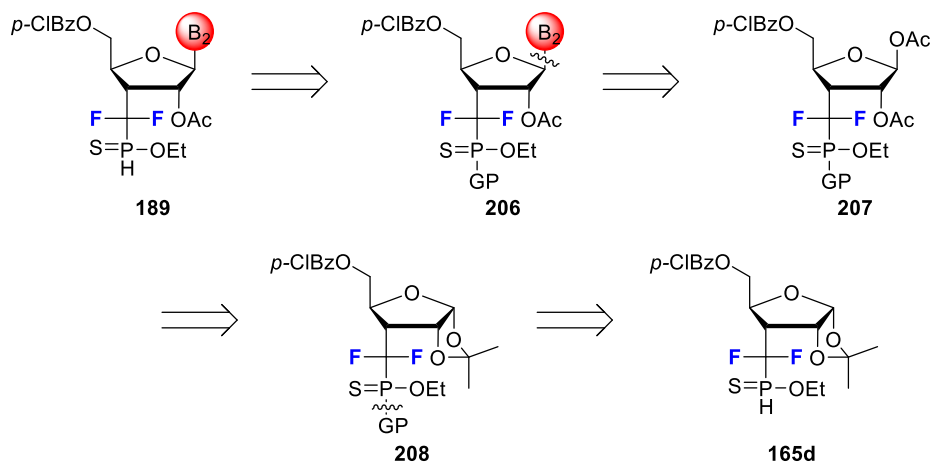


Schéma 29 : Rétrosynthèse voie E

La clé de cette stratégie repose sur la préparation de **189** dont la configuration en C3' a été contrôlée. Ainsi, en suivant des protocoles mis au point au laboratoire, la seconde liaison P-CH<sub>2</sub> pourrait être obtenue soit par addition radicalaire sur l'oléfine **185**, soit par un processus de type S<sub>N</sub>2 avec le nucléoside iodé **186**.

Le *H*-phosphinothioate **189** serait obtenu par le clivage d'un groupement protecteur (GP) du phosphinothioate **206** (Schéma 30). Une disconnection au niveau de la liaison osidique entre le furanose et la base nucléique B<sub>2</sub> générerait un synthon furanosique pouvant se traduire en bisacétate **207**. Ce dernier résulterait d'une modification du groupement protecteur du diol C1'-C2' de l'acétonide **208**. Enfin, une disconnection de la liaison P-GP permettrait la formation d'un synthon correspondant au *H*-phosphinothioate **165d**, dont la synthèse a déjà été réalisée au laboratoire.

Schéma 30 : Rétrosynthèse du *H*-phosphinothioate **189**

Le développement d'un groupement protecteur du *H*-phosphinothioate **189** résistant aux conditions acides de déprotection du diol C1'-C2' et d'introduction de la base devrait ainsi être réalisé puisqu'à l'heure actuelle, seules des protections stables en milieu basique sont présentes dans la littérature. Cette stratégie repose sur un travail de méthodologie de protection-déprotection et sur la préparation de douze substrats différents. Au vu du travail conséquent nécessaire à son développement, cette stratégie a été mise en attente.

## VI. Conclusions

D'un point de vue stratégique, les voies D et E ont été mises de côté. Celles-ci nécessitent notamment un développement synthétique important, avec respectivement la synthèse de huit et douze substrats différents avant l'étape de couplage.

Les travaux expérimentaux de cette thèse sont donc focalisés sur les stratégies A, B et C. Elles ont l'avantage de reposer sur des stratégies de construction de liaison P-C et des fragments (**163** et **165d**) déjà développés au laboratoire. Les autres substrats, dont les synthèses seront présentées dans le chapitre 2, peuvent quant à eux être préparés à partir d'une molécule commune, à savoir l'allofuranose. Les chapitres 3 et 4 décriront le développement synthétique de ces trois stratégies.







## Sommaire

<b>I. Synthèse de l'oléfine 185</b> .....	<b>89</b>
<b>I.1. Synthèse du fragment oléfinique 185a à partir d'un nucléoside</b> .....	<b>89</b>
I.1.a. Rétrosynthèse.....	89
I.1.b. Protection du diol C2'-C3' .....	90
I.1.c. Oxydation.....	91
I.1.d. Oléfination.....	92
I.1.e. Protection de l'uridine.....	92
<b>I.2. Synthèse du fragment oléfinique 185b à partir de l'allofuranose</b> .....	<b>93</b>
I.2.a. Rétrosynthèse.....	93
I.2.b. Synthèse .....	94
<b>II. Synthèse du fragment iodé 186</b> .....	<b>96</b>
<b>II.1. Rétrosynthèse</b> .....	<b>96</b>
<b>II.2. Synthèse</b> .....	<b>96</b>
<b>II.3. Etude de synthèses plus convergentes vers 186</b> .....	<b>97</b>
II.3.a. Stratégie n° 1 .....	98
II.3.b. Stratégie n° 2 .....	99
II.3.c. Insertion des autres bases nucléiques .....	100
<b>II.4. Protection de la thymine</b> .....	<b>101</b>
II.4.a. A partir du fragment <b>186a</b> .....	101
II.4.b. Stratégie hydrozirconation-iodation .....	102
II.4.c. Stratégie hydroboration-oxydation-iodation.....	103
<b>III. Conclusions</b> .....	<b>104</b>



Dans ce chapitre seront abordées les différentes voies de synthèse des fragments **184-186** (Figure 49). Ceux-ci seront utilisés dans les stratégies de synthèse A et B (Chapitre 3, I et II). Dans le cas de la stratégie B, la thymine et l'uracile ont été choisis dans un premier temps puisque ces nucléobases ont déjà été utilisées avec succès dans une stratégie de synthèse d'un hétérodinucléotide dont l'introduction des bases se fait après la construction du pont internucléosidique (Chapitre 1, I). L'approche synthétique C ayant été étudiée sur un substrat modèle (Chapitre 3, III.1), le fragment **187** n'a pas été synthétisé.

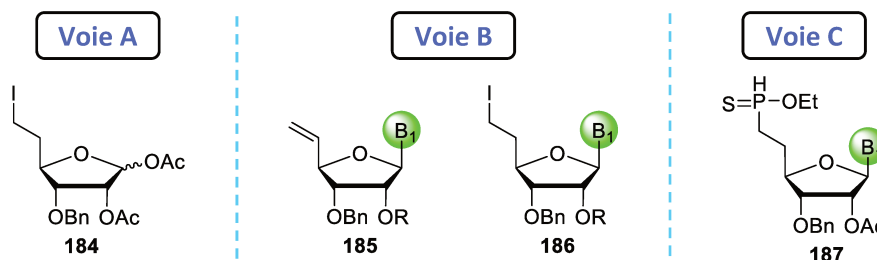


Figure 49 : Fragments synthétisés à partir de l'allofuranose

### I. Synthèse de l'oléfine **185**

Afin d'obtenir un nucléoside oléfinique **185**, deux voies de synthèses ont été envisagées. La première se ferait à partir d'un nucléoside et la seconde, d'un allofuranose.

#### I.1. Synthèse du fragment oléfinique **185a** à partir d'un nucléoside

##### I.1.a. Rétrosynthèse

Le nucléotide-cible **185a** pourrait être obtenu par le clivage du groupement benzylidène présent dans **212** suivi de la protection de l'alcool en C2' (Schéma 31).<sup>119</sup> L'alcène est en principe accessible par réaction d'oléfination, à partir de l'aldéhyde **211** lui-même obtenu par oxydation de l'alcool **210**. Ce dernier résulterait de la protection du diol C2'-C3' de l'uridine **209**. Le principal avantage de cette stratégie est d'avoir la possibilité d'introduire un groupement protecteur de notre choix sur l'alcool en position C2' après le clivage sélectif du benzylidène. D'un point de vue pratique, cette synthèse se ferait à partir de nucléosides commerciaux.

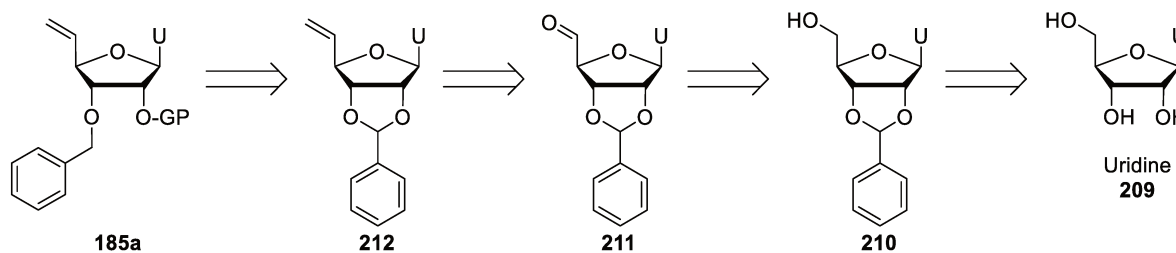


Schéma 31 : Rétrosynthèse du nucléoside **185a**

<sup>119</sup> Salinas, J. C.; Migawa, M. T.; Merner, B. L. and Hanessian, S. *J. Org. Chem.* **2014**, *79*, 11651-11660.

La séquence oxydation-oléfination a été réalisée auparavant sur des nucléosides ; quelques exemples sont présents dans la littérature dont les travaux de Nielsen<sup>120</sup> et Gravier-Pelletier<sup>121</sup> (Schéma 32). Le premier a réalisé une oxydation à l'aide du périodinate de Dess-Martin sur le dérivé d'uridine **213** afin d'obtenir l'aldéhyde correspondant **214**. Ensuite, l'oléfine souhaitée **215** a été formée par réaction de Wittig sur **214**. Gravier-Pelletier, quant à elle, a réalisé une oxydation du dérivé d'uridine **216** à l'aide d'acide 2-iodobenzoïque (IBX) afin de former l'aldéhyde **217** qui a été introduit directement dans une réaction d'oléfination, également selon des conditions de Wittig. L'oléfine **218** a été obtenue avec un rendement global de 61%. Vu ces exemples, une application de la séquence oxydation-oléfination sur notre dérivé uridine **209** s'avère donc envisageable.

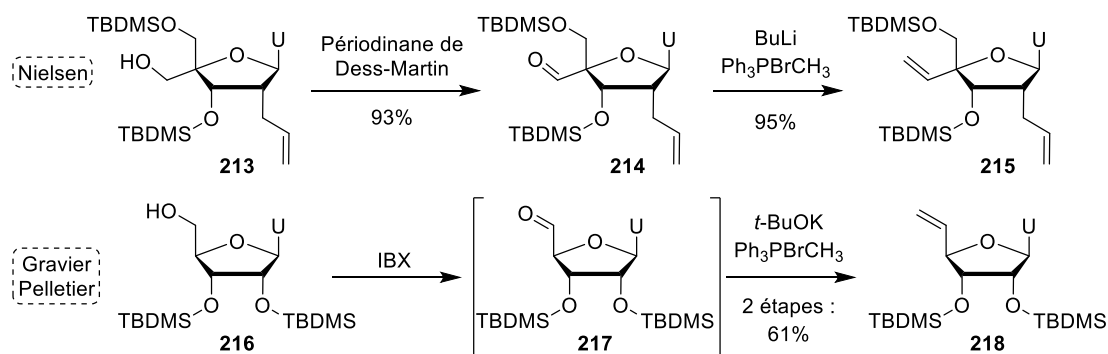


Schéma 32 : Travaux de Nielsen et Gravier-Pelletier

### I.1.b. Protection du diol C2'-C3'

En suivant un protocole de la littérature,<sup>122</sup> le diol C2'-C3' de l'uridine **209** a été protégé sous forme de benzylidène avec un rendement de 80% et a été obtenu sous la forme d'un mélange de diastéréoisomères (ratio 6/1). La structure du diastéréoisomère majoritaire a été déterminée par NOESY (Schéma 33). La présence du carbone asymétrique du benzylidène ne devrait pas avoir d'influence sur la stratégie oxydation-oléfination envisagée. De plus, lors du clivage sélectif du benzylidène, celui-ci disparaîtra.

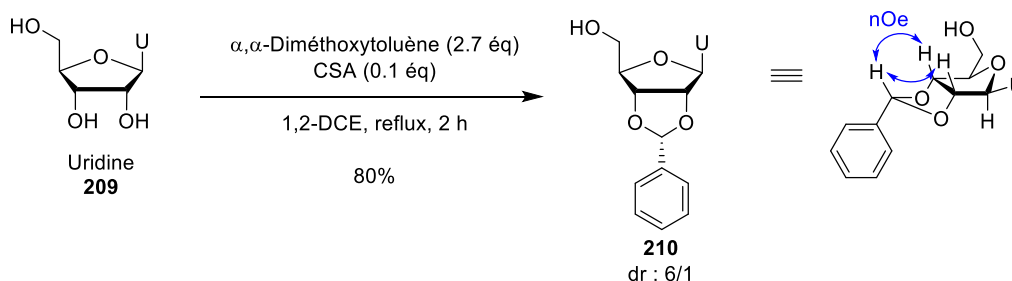


Schéma 33 : Protection du diol C2'-C3'

<sup>120</sup> Albæk, N.; Petersen, M. and Nielsen, P. *J. Org. Chem.* **2006**, *71*, 7731-7740.

<sup>121</sup> Fer, M. J.; Doan, P.; Prangé, T.; Calvet-Vitale, S. and Gravier-Pelletier, C. *J. Org. Chem.* **2014**, *79*, 7758-7765.

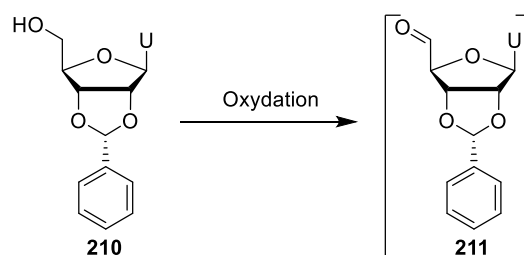
<sup>122</sup> a) Pour le protocole, voir : Johnson, D. C. and Widlanski, T. S. *Org. Lett.* **2004**, *6*, 4643-4646. b) Pour la caractérisation de **210**, voir : Li, C. and Vasella, A. *Helv. Chim. Acta* **1993**, *76*, 211-221.

I.1.c. Oxydation

De nombreuses méthodes ont été utilisées pour oxyder l'alcool en position C5' des furanoses comme, par exemple, l'utilisation d'IBX<sup>123</sup> ou du réactif de Sarett<sup>124</sup> mais également les conditions de Swern,<sup>125</sup> Parikh-Doering,<sup>126</sup> Pfitzner-Moffatt<sup>127</sup> ou Dess Martin.<sup>119</sup> Ces quatre dernières ont été testées sur l'alcool **210** pour obtenir l'aldéhyde **211** et les résultats sont présentés dans le Tableau 7.

Dans un premier temps, l'utilisation des conditions de Swern (entrée 1) ou encore du réactif de Burgess dans le DMSO<sup>128</sup> (entrée 2) n'a pas permis de convertir l'alcool **210** en aldéhyde **211**. Les conditions de Parikh-Doering (entrée 3), Dess-Martin (entrée 4) et Pfitzner-Moffatt (entrée 5) ont conduit à une conversion totale de l'alcool de départ. Pour ces trois réactions, plusieurs tentatives de purification (sur silice, alumine et florisil) de l'aldéhyde **211** ont été entreprises. Dans chaque cas, une dégradation totale, pouvant notamment résulter de l'hydratation de **211** sur support (silice, alumine ou florisil), a été observée.

Tableau 7 : Résultats des différentes oxydations testées



Entrée	Conditions		Temps (h)	Conversion de 210 en 211
1	Swern	(COCl) <sub>2</sub> (1.1 éq), Et <sub>3</sub> N (5 éq), DMSO (2 éq), CH <sub>2</sub> Cl <sub>2</sub>	2.5	0
2	Burgess	C <sub>8</sub> H <sub>18</sub> N <sub>2</sub> O <sub>4</sub> S (1.3 éq), DMSO	72	0
3	Parikh-Doering	DIPEA (5 éq), SO <sub>3</sub> -Py (3 éq), DMSO (10 éq), CH <sub>2</sub> Cl <sub>2</sub>	3	100
4	Dess Martin	Périodinane de Dess Martin (1.3 éq), DMSO, CH <sub>2</sub> Cl <sub>2</sub>	22	100
5	Pfitzner-Moffatt	DCC (3 éq), acide dichloroacétique (1 éq), DMSO	24	100
6	Pfitzner-Moffatt	EDCI.HCl (3 éq), acide dichloroacétique (1 éq), DMSO	20	100

Par conséquent, l'étude de l'étape d'oléfination a dû être envisagée à partir du mélange brut réactionnel résultant de l'étape d'oxydation. Ce dernier devant être dépourvu au maximum de

<sup>123</sup> More, J. D. and Finney, N. S. *Org. Lett.* **2002**, *4*, 3001-3003.

<sup>124</sup> Meurillon, M.; Gallier, F.; Peyrottes, S. and Périgaud, C. *Tetrahedron* **2009**, *65*, 6039-6046.

<sup>125</sup> Orban, J. and Reid, B. R. *J. Labelled Compd. Radiopharm.* **1988**, *18*, 195-198.

<sup>126</sup> Mahadevegowda, S. H. and Khan, F. A. *Tetrahedron Lett.* **2014**, *55*, 4400-4403.

<sup>127</sup> Rozners, E.; Katkevica, D.; Bizdena, E. and Strömberg, R. *J. Am. Chem. Soc.* **2003**, *125*, 12125-12136.

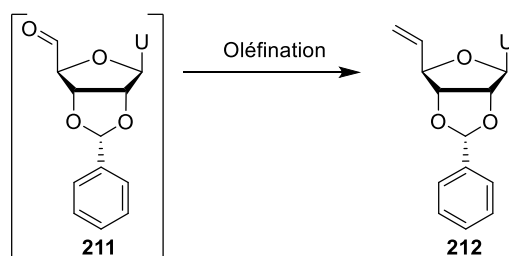
<sup>128</sup> Sultane, P. R. and Bielawski, C. W. *J. Org. Chem.* **2017**, *82*, 1046-1052.

sous-produits, l'oxydation suivant un protocole Pfitzner-Moffatt a été privilégiée. Le dicyclohexylcarbodiimide (DCC), initialement utilisé, a été remplacé par l'EDCI.HCl (entrée 6) afin d'éliminer les sous-produits de réaction par simple lavage aqueux. De ce fait, la présence de sous-produits a été limitée dans le mélange brut réactionnel des conditions de Pfitzner-Moffatt au contraire des mélanges bruts des réactions de Parikh-Doering et de Dess-Martin, dans lesquels la présence de sous-produits non-séparables a été observée. De plus, une conversion totale de l'alcool **209** a permis d'envisager l'utilisation du mélange brut réactionnel directement dans l'étape d'oléfination.

#### I.1.d. Oléfination

Le mélange brut réactionnel de l'aldéhyde **211** a donc été engagé dans une réaction d'oléfination afin d'obtenir l'alcène **212**. Pour cela, un nouveau travail d'optimisation a été entrepris et les résultats sont présentés dans le Tableau 8. Aucune des conditions testées (Wittig<sup>126</sup> (entrée 1), Tebbe<sup>129</sup> (entrée 2) ou Petasis<sup>130</sup> (entrée 3)) n'a permis l'obtention de l'oléfine **212**. En effet, dans le cas des conditions de Wittig, seul l'aldéhyde de départ a été observé dans le spectre de RMN du <sup>1</sup>H du mélange brut réactionnel après traitement. Les conditions de Tebbe et Petasis ont, quant à elles, donné lieu à la formation de mélanges complexes dans lesquels ni **211**, ni **212** n'ont été observés par spectrométrie de RMN du <sup>1</sup>H des mélanges bruts réactionnels.

Tableau 8 : Résultats des différentes réactions d'oléfination testées



Entrée	Conditions	Temps	<b>212</b>
<b>1</b>	Wittig Ph <sub>3</sub> PCH <sub>3</sub> Br (3 éq), <i>n</i> -BuLi (3 éq), THF	4 h	/
<b>2</b>	Tebbe Cp <sub>2</sub> TiCH <sub>2</sub> AlCl(CH <sub>3</sub> ) <sub>2</sub> , Pyridine, THF	4.5 h	/
<b>3</b>	Petasis Cp <sub>2</sub> Ti(CH <sub>3</sub> ) <sub>2</sub> (3 éq), Toluène	18 h	/

#### I.1.e. Protection de l'uridine

La fonction urée de l'uridine étant potentiellement à l'origine des problèmes rencontrés dans notre cas, une protection de l'azote N<sup>3</sup> a donc été envisagée (Schéma 34). Dans un premier temps, l'uridine a été protégée par un groupement *para*-méthoxybenzyle (PMB), à l'aide de DBU, avec un

<sup>129</sup> Tebbe, F. N.; Parshall, G. W. and Reddy, G. S. *J. Am. Chem. Soc.* **1978**, *100*, 3611-3613.

<sup>130</sup> Petasis, N. A. and Bzowej, E. I. *J. Am. Chem. Soc.* **1990**, *112*, 6392-6394.

rendement de 87%. Par la suite, les conditions de Pfitzner-Moffatt et de Swern ont été testées afin de former le composé **220**. En ce qui concerne les conditions de Pfitzner-Moffatt, après 48 h de réaction, seul le réactif de départ **219** a été observé dans le spectre de RMN du  $^1\text{H}$  du mélange brut réactionnel. En revanche, dans le spectre de RMN du  $^1\text{H}$  du mélange brut réactionnel, de nombreux signaux dans la zone des aldéhydes (aux alentours de 9-10 ppm) ont été observés lors de l'utilisation de la réaction de Swern. Ces signaux sont synonymes à la fois de la présence de l'aldéhyde **220** mais également de sous-produits non-séparables par purification du fait de l'instabilité potentielle de **220**. L'étape d'oléfination n'a donc pas été réalisée à cause de la présence de ces sous-produits.

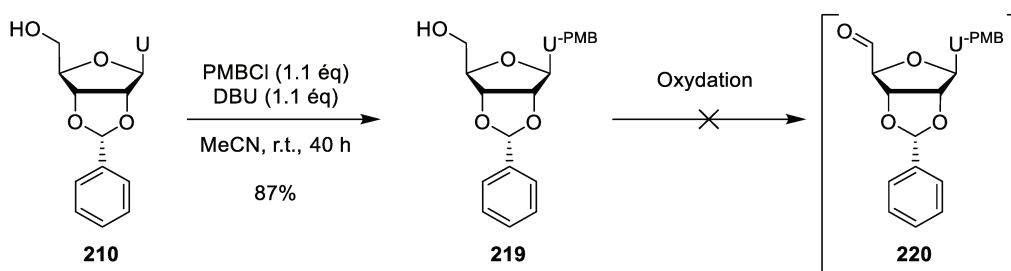


Schéma 34 : Essai d'oxydation sur l'uridine protégée

Le composé-cible **185a** n'ayant pas pu être obtenu à partir de l'uridine, la seconde voie de synthèse a été envisagée.

## I.2. Synthèse du fragment oléfinique **185b** à partir de l'allofuranose

### I.2.a. Rétrosynthèse

Une disconnection au niveau de la position anomérique et de la thymine générerait un synthon furanosique pouvant correspondre au bisacétate **223a** (Schéma 35). Ce dernier résulterait d'une modification de groupement protecteur sur l'acétonide **171**. L'interconversion de la liaison C=C présente dans **171** en diol vicinal **222**, qui peut lui-même être relié à l'analogue protégé **221** de l'allofuranose, clôturerait cette rétrosynthèse. Le composé **221** est disponible commercialement. L'avantage de cette stratégie, par rapport à celle présentée précédemment, est qu'elle permet l'introduction de bases nucléiques différentes lors de la dernière étape. Cependant, celle-ci s'avère plus longue d'une étape et nécessite la présence d'un groupement acétate en C2' afin d'introduire stéréosélectivement la base nucléique sur la position anomérique du furanose. Ce dernier devra être remplacé par un groupement protecteur de type éther silylé (TBDMS) en vue de l'introduction du dinucléotide final dans un oligonucléotide par synthèse automatisée rallongeant ainsi la séquence réactionnelle globale de deux étapes.

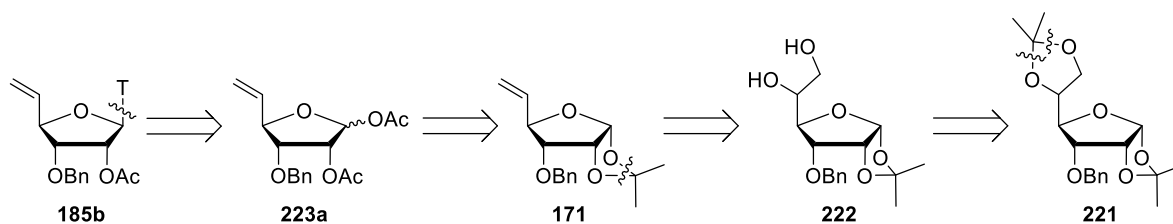


Schéma 35 : Rétrosynthèse du nucléotide **185b**

### I.2.b. Synthèse

Dans un premier temps, l'acétone C5'-C6' de l'allofuranose commercial **221** a été hydrolysé sélectivement en milieu acide (Schéma 36).<sup>131</sup> Le diol **222** formé a été transformé en alcène **171** dans un même ballon après une étape d'iodation de l'alcool primaire suivie d'une étape d'oléfination en présence d'un excès de triphénylphosphine.<sup>132</sup> Il est à noter que l'oléfine **171** a pu être isolée avec un rendement de 87% en deux étapes à partir de **221** contre 36% en huit étapes à partir du D-xylose **159**, lors des précédents travaux réalisés au laboratoire par les Drs. Jun Zhang et Emilie Lambert.

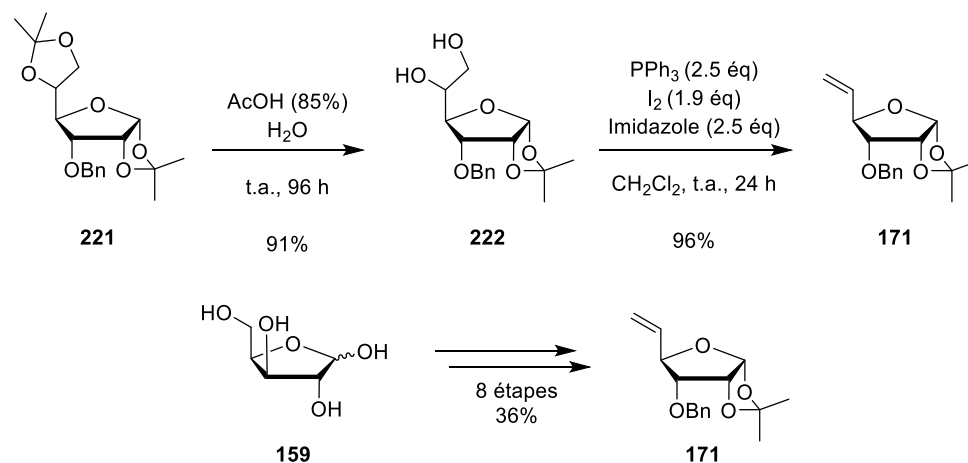


Schéma 36 : Synthèse de l'oléfine **171** à partir des composés **221** ou **159**

Par la suite, l'acétone **171** a été clivé sous conditions acides, puis les alcools C1'-C2' résultants ont été protégés à l'aide de groupements acétyles (Schéma 37). Il est à noter que lors de l'utilisation d'acide sulfurique en quantité catalytique, le produit désiré **223a** a été obtenu en mélange (ratio 2.3/1) avec le sous-produit **223b** avec un rendement de 79%. Ce dernier n'est généralement pas formé dans ces conditions mais il a tout de même été précédemment rapporté dans une étude de Mereyala.<sup>133</sup> Ces deux produits se sont avérés inséparables par chromatographie sur gel de silice. L'augmentation du temps de réaction, l'ajout d'acide sulfurique ou le chauffage du mélange réactionnel n'ont pas permis de supprimer ce sous-produit. Cependant, la formation de **223b** a pu être évitée en réalisant

<sup>131</sup> Chapron, C.; Glen, R.; La Colla, M.; Mayes, B. A.; McCarville, J. F.; Moore, S.; Moussa, A.; Sarkar, R.; Seifer, M.; Serra, I. and Stewart, A. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 2699-2702.

<sup>132</sup> Garegg, P. J. *Pure and Applied Chemistry* **1984**, *56*, 845.

<sup>133</sup> Mereyala, H. B. and Pola, P. *Synth. Commun.* **2003**, *33*, 2547-2552.

séquentiellement l'hydrolyse de l'acétonide en milieu acide, puis la protection du mélange brut réactionnel contenant le diol **224** sous forme d'acétates dans des conditions réactionnelles classiques. Ainsi, le bisacétate **223a** a pu être isolé avec un rendement de 71% en deux étapes à partir de **171**. Finalement, une troisième méthode utilisant l'acide camphorsulfonique (CSA) en quantité catalytique a été testée.<sup>115</sup> Dans ce cas, le produit désiré a pu être obtenu en une seule étape, avec un excellent rendement de 92%. Enfin, la thymine a été introduite selon les conditions de Vorbrüggen (BSA, TMSOTf) afin d'obtenir le nucléoside **185b** avec un bon rendement.<sup>134</sup> Celui-ci a été isolé sous forme d'un mélange d'anomères  $\alpha/\beta$  (ratio de 6/94).

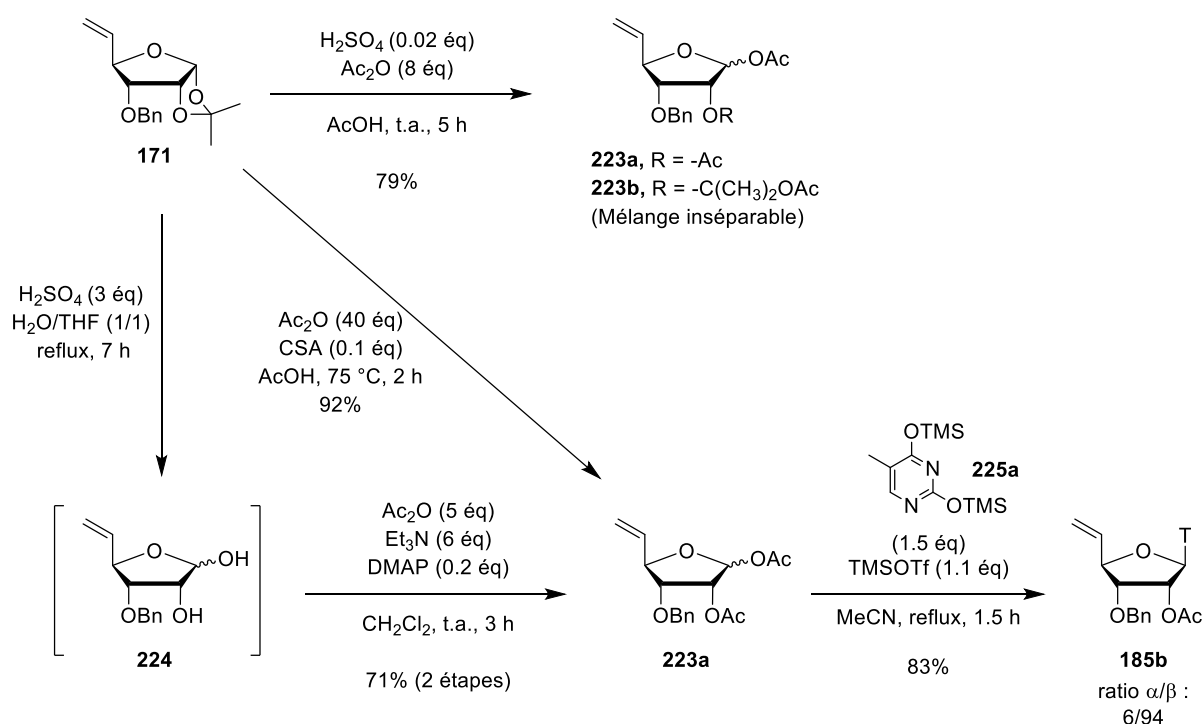


Schéma 37 : Introduction de la thymine

En résumé, le nucléoside-cible **185b** a été synthétisé, à l'échelle de plusieurs grammes, en quatre étapes à partir de l'allofuranose commercial **221** et avec un rendement global de 67%.

<sup>134</sup> Vorbrüggen, H.; Krolikiewicz, K. and Bennua, B. *Chem. Ber.* **1981**, *114*, 1234-1255.

## II. Synthèse du fragment iodé **186**

### II.1. Rétrosynthèse

Le nucléoside iodé **186a** peut donc être relié au nucléoside **229** par une séquence interconversion I/OH-rétroprotection (Schéma 38). Une disconnection de la thymine permet d'envisager un synthon correspondant au bisacétate **228**. Ce dernier peut être relié au fragment **227** par une séquence rétroprotection-rétrodéprotection. Enfin, une rétroprotection de l'éther silylé **227**, une rétrooxydation et une rétrohydroboration permettent d'envisager sa formation à partir d'un synthon correspondant à l'oléfine **171**. Celle-ci constitue donc le point de départ de la synthèse.

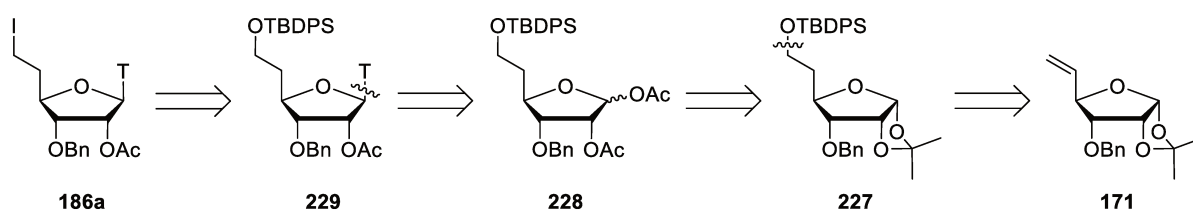


Schéma 38 : Rétrosynthèse du nucléoside iodé **186a**

### II.2. Synthèse

A partir de l'oléfine **171**, précédemment synthétisée (Chapitre 2, 1.2.b), une séquence hydroboration-oxydation a été réalisée à l'aide de 9-BBN suivie d'un traitement au peroxyde d'hydrogène en milieu basique (Schéma 39). Ainsi, l'alcool **226** a pu être isolé quantitativement. La protection de l'alcool **226** sous forme d'éther silylé TBDPS a été réalisée afin de résister aux conditions acides utilisées lors de la déprotection-protection du diol C1'-C2' pour former le bisacétate **228**.

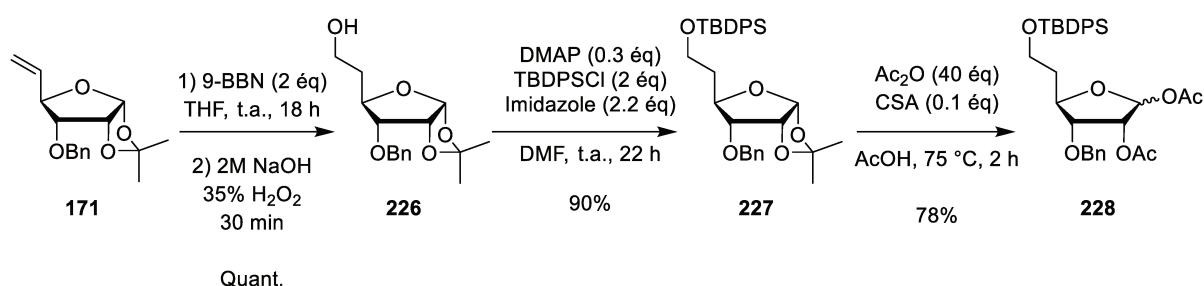


Schéma 39 : Synthèse du fragment **228**

La thymine a ensuite été introduite selon les conditions de Vorbrüggen, avec un rendement de 87%. Le nucléoside **229** a été isolé sous la forme d'un mélange d'anomères  $\alpha/\beta$  présentant un ratio de 6/94 (Schéma 40). L'alcool **230** a ensuite été obtenu avec un excellent rendement après clivage du groupement TBDPS par un mélange fluorure de triéthylammonium/acide fluorhydrique. L'étape d'iodation, à l'aide de triphénylphosphine et d'iode, termine la synthèse et permet l'obtention du composé **186a** (huit étapes avec un rendement global de 43%, à partir de l'allofuranose **221**).

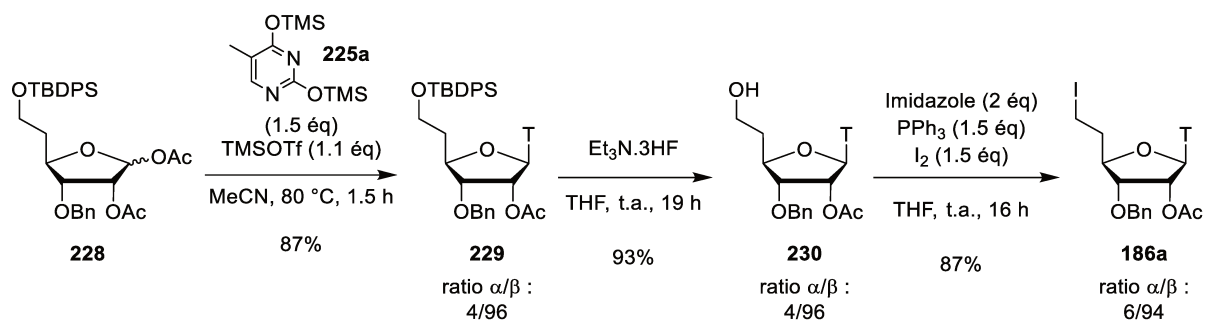


Schéma 40 : Synthèse du fragment iodé **186a**

### II.3. Etude de synthèses plus convergentes vers **186**

Afin de développer une synthèse permettant l'introduction de la base nucléique en dernière étape, de nouvelles stratégies synthétiques ont été étudiées. Une disconnection de la thymine générerait un synthon pouvant correspondre au bisacétate **184**. Deux stratégies rétrosynthétiques peuvent être envisagées en jouant sur l'ordre de la séquence réactionnelle. Dans une première, le bisacétate **184** résulterait d'une séquence interconversion I/OH - rétrodéprotection de l'éther silylé **228**. Dans la seconde, une rétroprotection, une rétrodéprotection et une interconversion I/OH permettraient la formation de **184** à partir d'un synthon pouvant correspondre à l'alcool **226**.

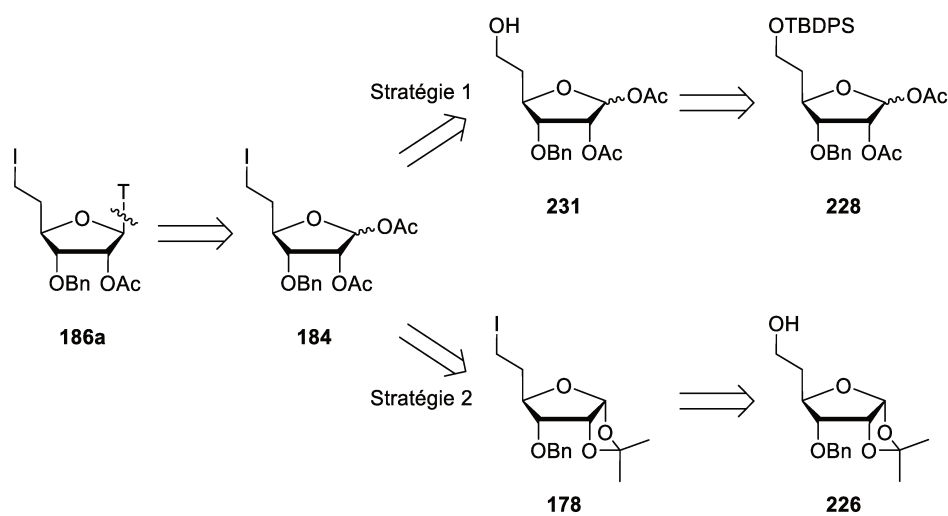
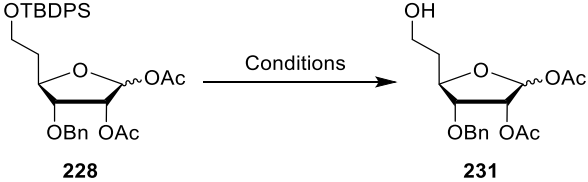


Schéma 41 : Rétrosynthèse menant au nucléoside **186a**

II.3.a. Stratégie n° 1

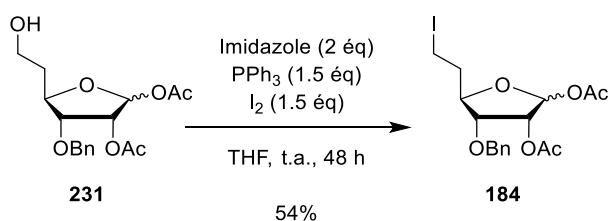
Afin d'étudier la stratégie de synthèse n°1 vers **184** à partir du fragment **228**, obtenu précédemment (Chapitre 2, II.2), des tests de déprotection de l'alcool en C6' ont été réalisés dans un premier temps et les résultats sont présentés dans le Tableau 9. Lors de l'utilisation de TBAF<sup>135</sup> (entrée 1), l'alcool désiré a été observé dans le mélange brut réactionnel. Néanmoins, le produit-cible **231** n'a pu être isolé pur du fait de la présence de sous-produits non-séparables par chromatographie sur gel de silice. L'utilisation d'acide fluorhydrique (HF), en présence de pyridine<sup>136</sup> (entrée 2) ou de triéthylamine<sup>137</sup> (entrée 3), a permis d'obtenir l'alcool **231**, avec des rendements de 50% et 68% respectivement.

Tableau 9 : Résultats des conditions de déprotection de l'alcool en C6'



Entrée	Conditions	Temps	Rendement
1	TBAF (4 éq), THF, t.a.	18 h	X
2	HF.pyridine (12 éq), THF/pyridine, t.a.	18 h	50%
3	Et <sub>3</sub> N.3HF (12 éq), THF, t.a.	18 h	68%

Dans un second temps, une étape d'iodation réalisée sur l'alcool **231** a permis d'obtenir le fragment-cible **184**, avec un rendement moyen de 54% (Schéma 42), soit un rendement global de 22% en sept étapes à partir de l'allofuranose protégé **221**.

Schéma 42 : Synthèse du fragment **184**

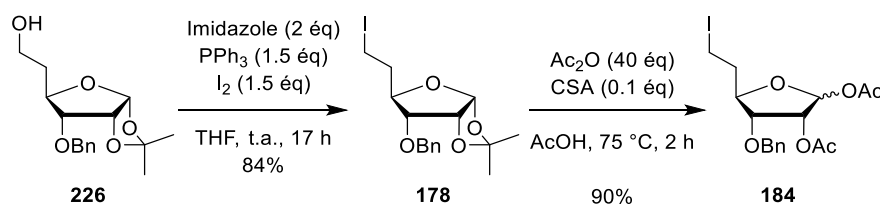
<sup>135</sup> Wang, L.; Feng, S.; An, L.; Gu, G. and Guo, Z. *J. Org. Chem.* **2015**, *80*, 10060-10075.

<sup>136</sup> Wang, S.; Sun, J.; Zhang, Q.; Cao, X.; Zhao, Y.; Tang, G. and Yu, B. *Angew. Chem. Int. Ed.* **2018**, *57*, 2884-2888.

<sup>137</sup> Prakash, T. P.; Siwkowski, A.; Allerson, C. R.; Migawa, M. T.; Lee, S.; Gaus, H. J.; Black, C.; Seth, P. P.; Swayze, E. E. and Bhat, B. *J. Med. Chem.* **2010**, *53*, 1636-1650.

II.3.b. Stratégie n° 2

En raison du faible rendement obtenu sur les étapes de clivage de l'éther silylé TBDPS et d'iodation, la seconde approche vers **184** à partir de l'alcool **226** synthétisé précédemment (Chapitre 2, II.2) a été entreprise. Dans un premier temps, l'introduction de l'iode en C6' a été réalisée à partir de l'alcool **226** à l'aide de triphénylphosphine et d'iode (Schéma 43). Ainsi, l'iodure d'alkyle **178** a pu être isolé avec un excellent rendement de 84%. Dans un second temps, le bisacétate **184** a été synthétisé via une séquence déprotection-protection du diol C1'-C2' en présence d'anhydride acétique et une quantité catalytique de CSA. Le furanose-cible **184** a donc été obtenu en cinq étapes avec un rendement global de 66% à partir de l'allofuranose protégé **221** contre 22% lors de la mise en œuvre de la stratégie n°1.

Schéma 43 : Optimisation de la synthèse du fragment **184**

Afin de valider notre hypothèse d'introduire la nucléobase lors de la dernière étape, la thymine a été introduite dans les conditions de Vorbrüggen permettant l'accès au nucléoside **186a** sous la forme d'un mélange  $\alpha/\beta$  (ratio 4/96) avec un rendement de 80% (Schéma 44). En comparaison avec la stratégie développée précédemment (Chapitre 2, II.2), le nucléoside **186a** a ainsi été obtenu en six étapes avec un rendement global de 53% à partir de **221**, au lieu de huit étapes et un rendement global de 43% en partant de l'allofuranose **221**.

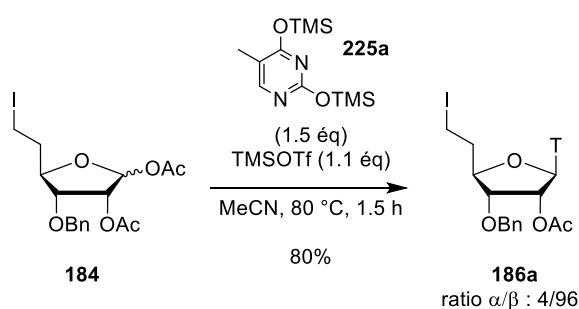
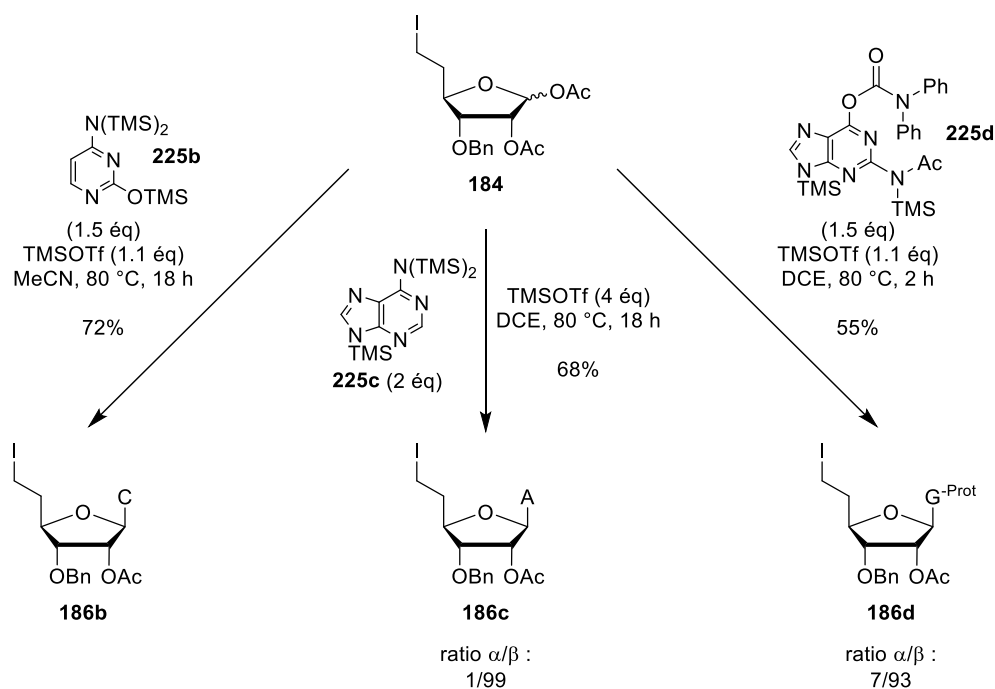


Schéma 44 : Introduction de la thymine

## II.3.c. Insertion des autres bases nucléiques

La mise au point d'une synthèse efficace du fragment iodé **184** a permis d'envisager la préparation des différents nucléosides. Ainsi, les trois autres bases nucléiques présentes dans l'ADN – cytosine, adénine et guanine – ont pu être introduites.

Tout d'abord, la cytosine silylée **225b** a été introduite sur la position anomérique du furanose **184** dans des conditions similaires à l'introduction de la thymine. Le nucléoside **186b** désiré a été obtenu avec un bon rendement de 72% (Schéma 45). Il est à noter que, dans ce cas, le nucléoside cytosine **186b** a été obtenu uniquement sous la forme de l'anomère  $\beta$ .

Schéma 45 : Introduction des bases nucléiques sur le fragment iodé **184**

L'introduction de l'adénine **225c**, au préalable silylée à l'aide de chlorure de triméthylsilyle (TMSCl), a permis la formation du nucléoside **186c** avec un rendement de 68% et sous la forme d'un mélange  $\alpha/\beta$  (1/99).<sup>138</sup>

Dans le cas de la guanine, cette dernière a au préalable été protégée suivant le protocole de Zou, afin d'éviter un mélange de régioisomères N<sup>7</sup>/N<sup>9</sup> (Figure 50) et de favoriser la proportion d'isomère N<sup>9</sup> (présente dans les nucléosides et nucléotides naturels).<sup>139</sup> Dans un premier temps, l'amine libre a été

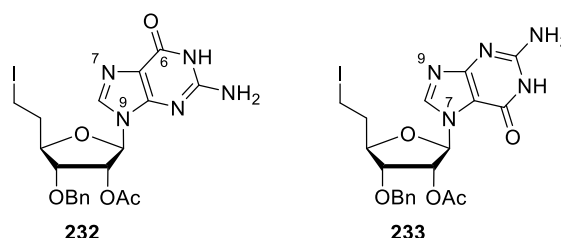


Figure 50 : Régioisomères de la guanine

<sup>138</sup> Fujino, T.; Kogashi, K.; Okada, K.; Mattarella, M.; Suzuki, T.; Yasumoto, K.; Sogawa, K. and Isobe, H. *Chem. Asian J.* **2015**, *10*, 2683-2688.

<sup>139</sup> Zou, R. and Robins, M. J. *Can. J. Chem.* **1987**, *65*, 1436-1437.

protégée sous forme d'acétate à l'aide d'anhydride acétique, à 160 °C, dans le diméthylacétamide (DMAc) (Schéma 46). Puis l'oxygène en position 6 de la guanine a été protégé à l'aide d'un groupement diphenylcarbamoyle. Ainsi la guanine bis-protégée **235** a pu être introduite sur le fragment **184** avec un rendement de 55% et isolé comme seul isomère N<sup>9</sup> et en mélange  $\alpha/\beta$  avec un ratio 7/93.

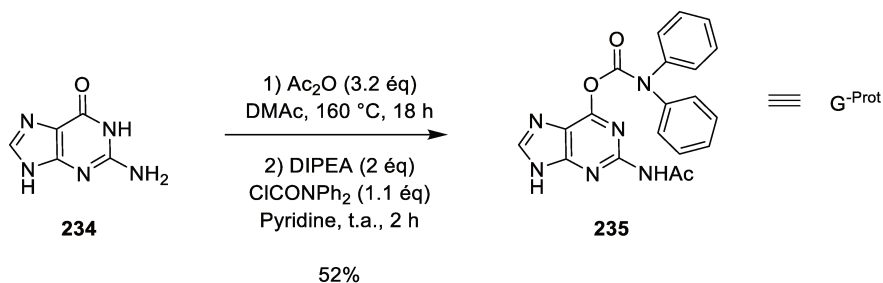
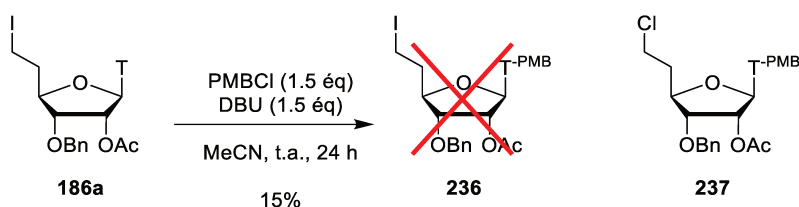


Schéma 46 : Protection de la guanine

#### II.4. Protection de la thymine

##### II.4.a. A partir du fragment **186a**

Anticipant d'éventuels problèmes lors de l'étape de formation de la liaison P-CH<sub>2</sub> par une réaction de Michaelis-Becker réalisée sous conditions basiques, la protection de l'azote N<sup>3</sup> de la thymine par un groupement PMB à partir de PMBCl a été envisagée. Cependant, le produit désiré **236** n'a pas été observé et le chlorure d'alkyle **237** résultant d'une substitution de l'iode par le chlore a été isolé avec un rendement de 15% (Schéma 47). En effet, une modification de la valeur du déplacement chimique du carbone porteur de l'halogène a été observée due au remplacement de l'iode par un chlore (Pour CH<sub>2</sub>-I : 0.54 ppm, pour CH<sub>2</sub>-Cl : 40.8 ppm). La structure **237** a été confirmée par spectrométrie de masse haute résolution. Le faible rendement peut être en partie expliqué par une réaction de déshydrohalogénéation compétitive à partir du substrat **186a** ou du produit **237**, respectivement, sous l'action de la base. En effet, la présence de signaux oléfiniques a été observée par spectrométrie de RMN du <sup>1</sup>H du le mélange brut réactionnel.



II.4.b. Stratégie hydrozirconation-iodation

A partir du fragment nucléosidique **185b**, dont la synthèse a été décrite précédemment (Chapitre 2, I.2.b), la thymine a été protégée à l'aide du *para*-méthoxybenzyle, en présence de DBU, avec un rendement de 88% (Schéma 48).

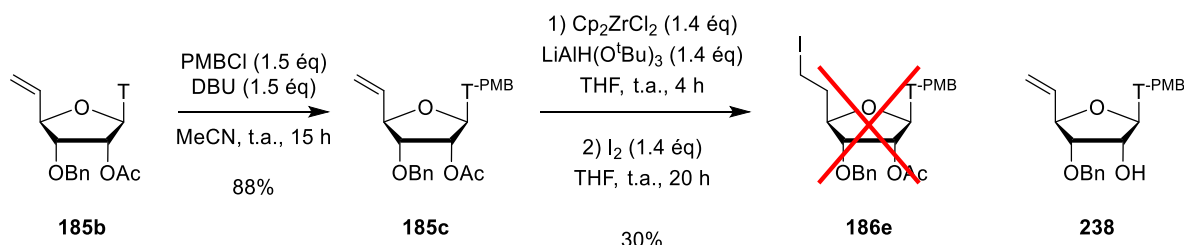


Schéma 48 : Protection de la thymine et essai de la séquence hydrozirconation-iodation

Lors des travaux précédents du laboratoire, l'alcène **171** a été converti en iodure d'alkyle **178** à l'aide d'une séquence d'hydrozirconation-iodation (Schéma 49).<sup>111</sup> Cette séquence a ainsi été testée en vue d'obtenir l'alkyle iodé **186e**.

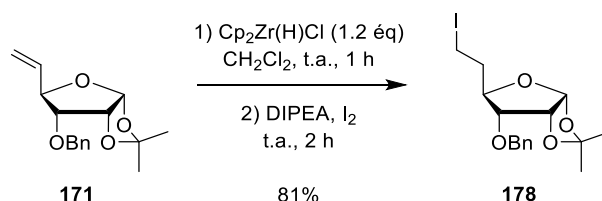


Schéma 49 : Séquence hydrozirconation-iodation sur l'oléfine 171

Pour cela, l'oléfine **185b** a été soumise au réactif de Schwartz commercial suivi par l'addition d'une solution d'iode afin d'intercepter l'intermédiaire organométallique résultant de l'étape d'hydrozirconation (Schéma 48). Cependant, aucune réaction n'a été observée. L'oléfine **185c** a donc été mise en présence d'un réactif de Schwartz généré *in situ* à l'aide du dichlorure de zirconocène et du tri-*tert*-butyloxy-hydrure d'aluminium et de lithium.<sup>140</sup> Néanmoins, même en augmentant la quantité de réactif à trois équivalents, le composé-cible **186e** n'a pas été observé et seul l'alcool **238** provenant d'une déacétylation de la position C2' a été isolé avec un rendement de 30%.

Suite à ces résultats, le groupement protecteur ester en position C2' a été remplacé par un éther silylé TBDMS (Schéma 50). Pour ce faire, l'alcool en C2' a été déprotégé, de manière quantitative, à l'aide de carbonate de potassium dans le méthanol. Puis, l'alcool libre **238** a été protégé à l'aide de chlorure de *tert*-butyldiméthylsilyle pour donner l'oléfine **239**, avec un rendement de 75%. La séquence hydrozirconation-iodation, à l'aide du réactif de Schwartz commercial ou généré *in situ*, a alors été

<sup>140</sup> Zhao, Y. and Snieckus, V. *Org. Lett.* **2014**, *16*, 390-393.

testée sur l'oléfine **239** mais sans succès. Dans ce cas, seule l'oléfine de départ a été retrouvée dans le mélange brut réactionnel.

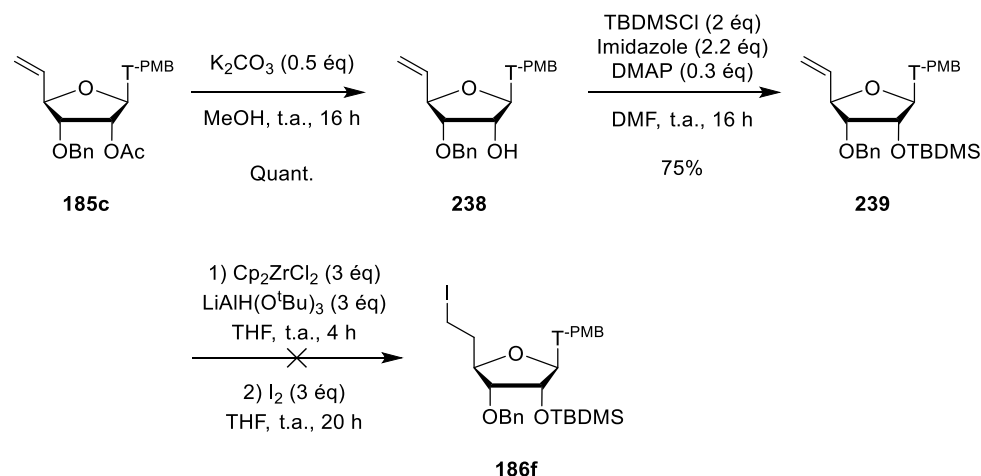


Schéma 50 : Modification du groupement en C2' pour essai de la séquence hydrozirconation-iodation

Vu les résultats non-concluants de la séquence hydrozirconation-iodation, une séquence hydroboration-oxydation-iodation a été envisagée.

#### II.4.c. Séquence hydroboration-oxydation-iodation

Les conditions basiques du traitement de l'étape d'oxydation réalisée après l'hydroboration de l'oléfine (NaOH) pouvant entraîner la déprotection du groupement acétyle en C2', la séquence hydroboration-oxydation a été réalisée à partir de l'oléfine **239** possédant un groupement OTBDMS en position C2' (Schéma 51). L'alcool **240** a été obtenu par l'action du 9-BBN sur l'oléfine **239** suivi d'un traitement à l'eau oxygénée en présence d'hydroxyde de sodium. Enfin, la conversion de l'alcool C6' en iode, à l'aide de triphénylphosphine et d'iode, a permis d'obtenir le nucléoside **186f** possédant une thymine protégée avec un rendement de 87%. Globalement, ce fragment a été obtenu en neuf étapes à partir de l'allofuranose protégé **221** avec un rendement de 31%.

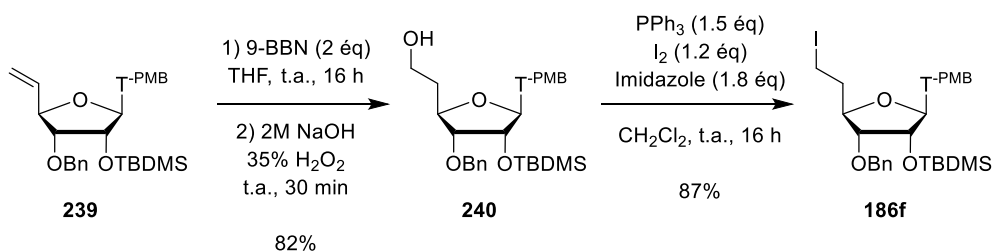


Schéma 51 : Séquence d'hydroboration-oxydation-iodation

### III. Conclusions

Dans le but d'explorer les stratégies synthétiques envisagées (A et B), la synthèse des différents fragments-cibles a été réalisée à partir d'un substrat commun, l'allofuranose **221** (Figure 51).

Ainsi, les fragments oléfiniques **185b** et **185c**, obtenus respectivement en quatre et cinq étapes, pourront être utilisés dans la formation de la seconde liaison phosphore-carbone par voie radicalaire. Les nucléosides iodés **186a-d**, **186f** et **184** seront utilisés pour réaliser la seconde liaison P-CH<sub>2</sub> par une réaction de Michaelis-Becker. Il est à noter que le furanose **184** a pu être utilisé pour introduire les quatre bases nucléiques présentes dans l'ADN avec des rendements compris entre 55% et 80% (**186a-d**) améliorant ainsi l'efficacité de la synthèse des fragments nucléosidiques.

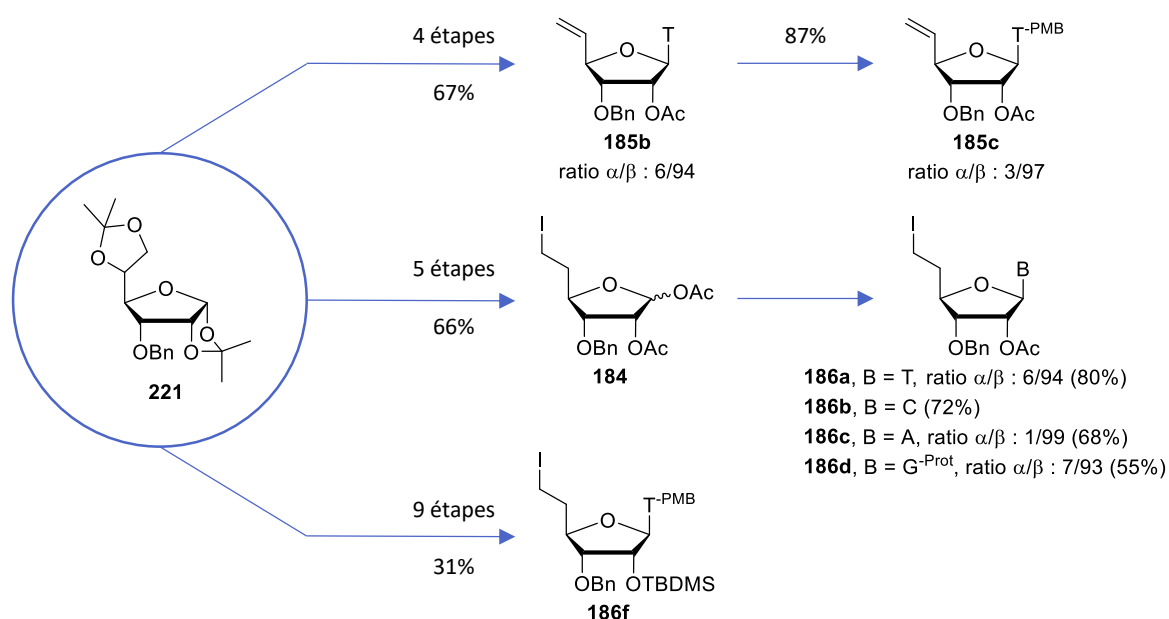


Figure 51 : Récapitulatif des fragments synthétisés

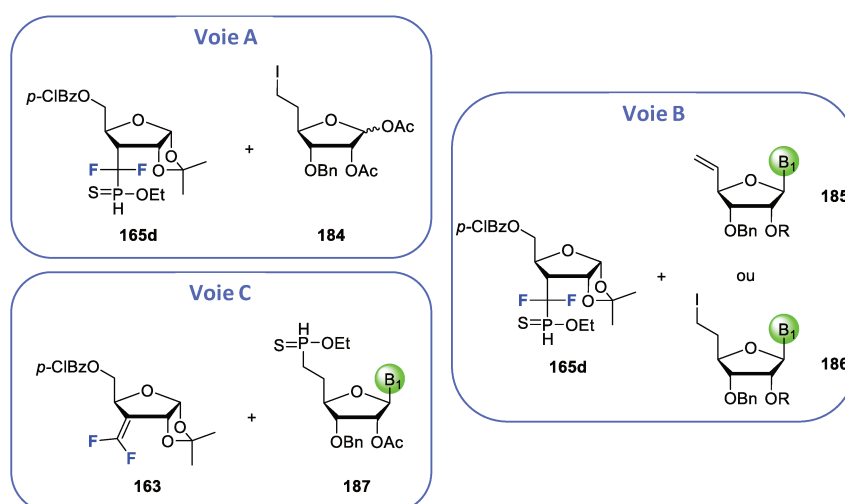
La synthèse des fragments-cibles a été réalisée avec succès ; ceux-ci ont été utilisés dans les différentes voies de synthèse envisagées pour former une seconde liaison P-CH<sub>2</sub>, soit par réaction radicalaire, soit par réaction ionique. Ces synthèses seront développées dans le chapitre suivant.





# Chapitre 3

*Formation de la seconde liaison phosphore-carbone et du motif difluorophosphinothioate*





## Sommaire

<b><i>I. Etude de la voie A</i></b> .....	<b>111</b>
<b><i>II. Etude de la voie B</i></b> .....	<b>113</b>
<b>II.1. Formation de la seconde liaison P-C par addition radicalaire</b> .....	<b>113</b>
<b>II.2. Formation de la seconde liaison P-C par voie ionique : réaction de Michaelis-Becker</b> .....	<b>114</b>
<b><i>III. Etude de la voie C : application sur un substrat modèle</i></b> .....	<b>117</b>
<b>III.1. Etude sur un substrat modèle : rétrosynthèse</b> .....	<b>117</b>
<b>III.2. Formation de la première liaison P-C</b> .....	<b>118</b>
<b>III.3. Etude de la réactivité H-P=O versus H-P=S pour la formation de la seconde liaison P-C</b> .....	<b>118</b>
<b><i>IV. Conclusions</i></b> .....	<b>122</b>



Dans ce chapitre, les différentes voies proposées (Chapitre 1) conduisant à la formation d'une seconde liaison P-C seront abordées. Les voies A et B ont été développées à partir des fragments nucléosidiques précédemment synthétisés tandis que la voie C a été étudiée sur un substrat modèle.

### I. Etude de la voie A

En s'appuyant sur des travaux de l'équipe du Pr. Lebreton,<sup>115</sup> présentés dans le chapitre 1 (I), le dinucléotide **183** pourrait être obtenu par introduction successive des bases nucléiques sur l'intermédiaire **196** présentant des protections différentes des diols C1'-C2' sur les deux furanoses (Schéma 52). Ainsi la seconde liaison phosphore-carbone serait obtenue à partir des fragments *H*-phosphinothioate **165d** et le bisacétate **184**.

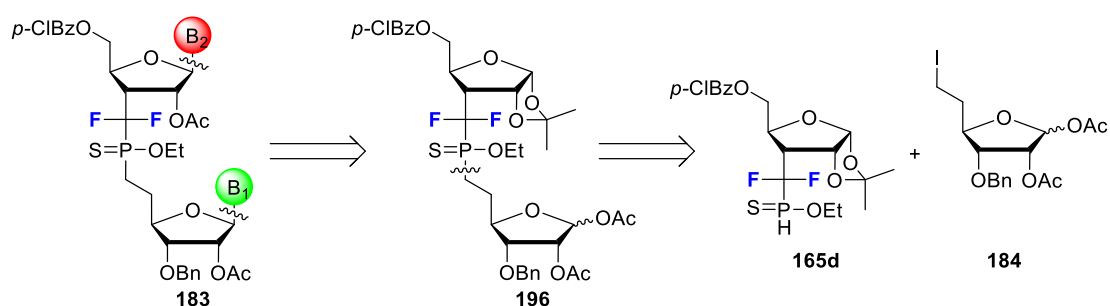


Schéma 52 : Stratégie A

Le *H*-phosphinothioate **165d** a été synthétisé plusieurs fois, selon le protocole élaboré au laboratoire et décrit précédemment (Introduction, V.2), en 7 étapes avec un rendement global de 12%. Une modification des conditions de purification a permis la séparation des deux diastéréoisomères et leur caractérisation totale (Experimental part, III.1).

En suivant les conditions mises au point au laboratoire par Dr. Emilie Lambert,<sup>111</sup> un essai de réaction de Michaelis-Becker entre **165d** et **184** a été réalisé (Schéma 53). Après un jour de réaction, une faible conversion du *H*-phosphinothioate **165d** en intermédiaire acétonide-bisacétate **196** a été observée en spectrométrie de RMN du <sup>19</sup>F. Dans le spectre de RMN du <sup>1</sup>H du mélange brut réactionnel, un ratio 1/1 a été observé entre l'iodure d'alkyle **184** et le produit attendu **196**. Il est à noter que le *H*-phosphinothioate **165d** a également été observé, dans le spectre de RMN du <sup>31</sup>P découplé du proton (ratio **165d**/**196** : 1/1.5). Une purification en phase normale n'a pas permis de séparer **184** et **196**. Une purification en phase inverse a donc été réalisée pour éliminer l'iodure d'alkyle **184** n'ayant pas réagi du produit **196**, qui n'a cependant été isolé qu'avec un faible rendement de 7%. Ceci peut être le résultat de la dégradation compétitive de **184** (hydrolyse de la fonction ester ?) et/ou **196**.

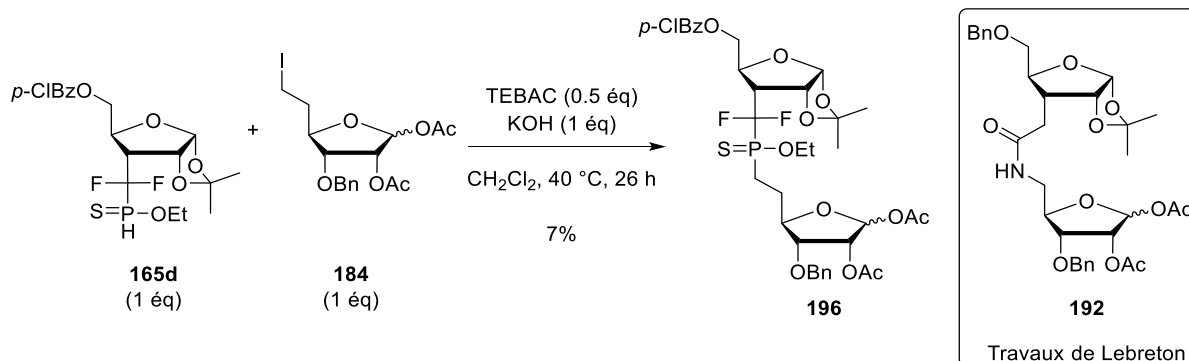


Schéma 53 : Réaction de Michaelis-Becker entre **165d** et **184**

Un second essai a été réalisé dans le dichloroéthane à 60 °C. Dans ce cas, après 48 h de réaction, dans le mélange brut réactionnel, la formation du produit **196** n'a pas été observée en spectrométrie de RMN du  $^1\text{H}$ , du  $^{19}\text{F}$  et du  $^{31}\text{P}$ . Après purification en phase normale, un mélange comportant le *H*-phosphinothioate **165d** et un sous-produit non-séparable présentant des signaux dans les spectres RMN du  $^{19}\text{F}$  découplé du  $^1\text{H}$  et du  $^{31}\text{P}$  découplé du  $^1\text{H}$  a été obtenu (ratio 1.3/1). Il est à noter que le iodonucléoside **184** est également présent dans ce mélange (ratio **165d/184** : 1/2 dans le spectre de RMN du  $^1\text{H}$ ).

Les déplacements chimiques de ce sous-produit, qui se trouve sous la forme d'un mélange de diastéréoisomères (ratio A/B : 2/1), sont les suivants :

Diastéréoisomère A : RMN  $^{19}\text{F}$   $\{^1\text{H}\}$  :  $\delta$  -99.8 (dd, 1F,  $J = 291.0$  Hz, 84.4 Hz) et -110.5 (dd, 1F,  $J = 291.0$  Hz, 87.9 Hz) ; RMN  $^{31}\text{P}$   $\{^1\text{H}\}$  :  $\delta$  84.8 (dd,  $J = 87.9$  Hz, 84.4 Hz).

Diastéréoisomère B : RMN  $^{19}\text{F}$   $\{^1\text{H}\}$  :  $\delta$  -98.6 (dd, 1F,  $J = 291.4$  Hz, 111.2 Hz) et -112.2 (dd, 1F,  $J = 291.4$  Hz, 81.1 Hz) ; RMN  $^{31}\text{P}$   $\{^1\text{H}\}$  :  $\delta$  85.8 (dd,  $J = 111.2$  Hz, 81.1 Hz).

En spectrométrie de RMN du  $^{31}\text{P}$ , le *H*-phosphinothioate **165d** a un déplacement chimique d'environ 60 ppm. D'après l'expérience précédente et les résultats du laboratoire,<sup>111</sup> la formation du motif difluorophosphinothioate engendre des déplacements chimiques autour de 94-90 ppm. Ainsi, une dégradation du composé **196** pourrait être à l'origine de ce sous-produit. D'après les travaux du Dr. Emilie Lambert,<sup>111</sup> il est à noter que les déplacements chimiques d'un motif S=P-OH, résultant d'une hydrolyse de S=P-OEt, se situent à environ 70 ppm et que ceux de O=P-OEt se trouvent dans la zone des 40 ppm. A l'heure actuelle, sa structure n'a pas été élucidée.

Les résultats escomptés n'ayant pas été obtenus grâce à cette voie A, d'autres stratégies ont donc été envisagées.

## II. Etude de la voie B

Comme présenté dans le chapitre 1 (II), l'hétérodinucléotide **183** proviendrait de l'intermédiaire nucléotide-furanose **197** (Schéma 54). Par disconnection au niveau de la liaison P-CH<sub>2</sub>, ce dernier pourrait être obtenu à partir de synthons issus du *H*-phosphinothioate **165d** et de fragments incorporant une base nucléique de type oléfinique **185** ou iodé **186**.

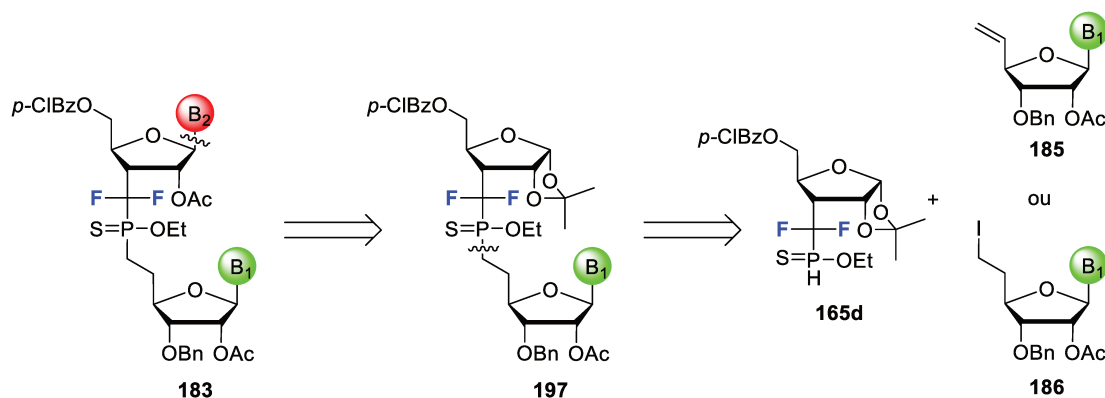


Schéma 54 : Stratégie B

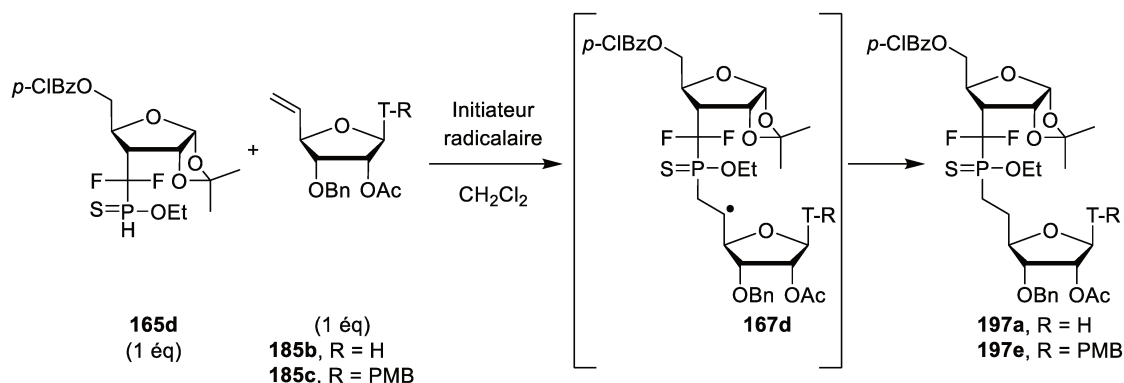
Fondée sur la synthèse de première génération développée au laboratoire, la formation de la liaison P-CH<sub>2</sub> pourrait être étudiée soit par addition radicalaire entre le *H*-phosphinothioate **165d** et l'oléfine **185**, soit par réaction de Michaelis-Becker entre **165d** et l'iodonucléoside **186**.

### II.1. Formation de la seconde liaison P-C par addition radicalaire

Des tests de formation de la seconde liaison P-C par voie radicalaire entre le *H*-phosphinothioate **165d** et le nucléoside non-protégé **185b** ou protégé **185c** ont été réalisés. Les résultats de cette étude sont présentés dans le Tableau 10. Deux initiateurs radicalaires ont été testés, à savoir le DLP et le TBPP. Dans un premier temps, l'addition radicalaire a été testée sur l'oléfine comportant la thymine non-protégée **185b** (entrée 1 et 2). En ce qui concerne l'entrée 1, des traces de produit ont été détectées mais une quantité importante d'oléfine **185b** a également été observée dans le mélange brut réactionnel en spectrométrie de RMN du <sup>1</sup>H. L'utilisation de TBPP a conduit à un résultat similaire, des signaux correspondants à l'oléfine **185b**, au *H*-phosphinothioate **165d** et au produit attendu **197a** ayant été observés dans le mélange brut réactionnel en spectrométrie de RMN du <sup>1</sup>H, du <sup>19</sup>F et du <sup>31</sup>P (entrée 2). Deux autres essais ont été réalisés à partir d'une oléfine comportant une thymine PMB-protégée **185c** (entrée 3 et 4). Les résultats obtenus sont identiques à ceux des deux premières entrées.

Finalement, aucune différence de réactivité n'a été observée lors de l'utilisation d'un alcène porteur d'une thymine protégée ou non. Dans chaque cas, seules des traces de l'intermédiaire souhaité **197a** ou **197e** ont été observés en TLC-MS (Thin-Layer Chromatography – Mass Spectroscopy) ainsi qu'en

spectrométrie de RMN du  $^{19}\text{F}$  et du  $^{31}\text{P}$ , en concordance avec les valeurs des déplacements chimiques observées par Dr. Jun Zhang (entre -100 ppm et -110 ppm pour l'analyse RMN  $^{19}\text{F}$  et aux alentours de 90 ppm pour l'analyse RMN  $^{31}\text{P}$ ). Cependant, ce composé n'a pu être isolé du fait de sa faible quantité dans le milieu.

Tableau 10 : Tests d'addition radicalaire entre **165d** et **185**

Entrée	R	Initiateur radicalaire	Eq	Temp. (°C)	Tps (h)	197a ou e
1	H	DLP	0.2 + 0.4	85	5	Traces
2	H	TBPP	0.3	65	24	Traces
3	PMB	DLP	0.2 + 0.4	85	5	Traces
4	PMB	TBPP	0.3 + 0.3	65	6	Traces

La présence de la base nucléique, sur le substrat oléfinique **187**, pourrait avoir un effet délétère sur cette réaction d'addition radicalaire. Cet effet pourrait être expliqué par l'encombrement stérique généré par la thymine qui empêcherait le radical-adduit **167d** d'induire l'étape de propagation en venant arracher l'hydrogène de **165d** pour former le composé-cible **197**. Ainsi, des réactions secondaires pourraient prendre place comme, par exemple, l'attaque du radical-induit **167d** sur la base nucléique dont des exemples ont été rapportés dans la littérature.<sup>141</sup>

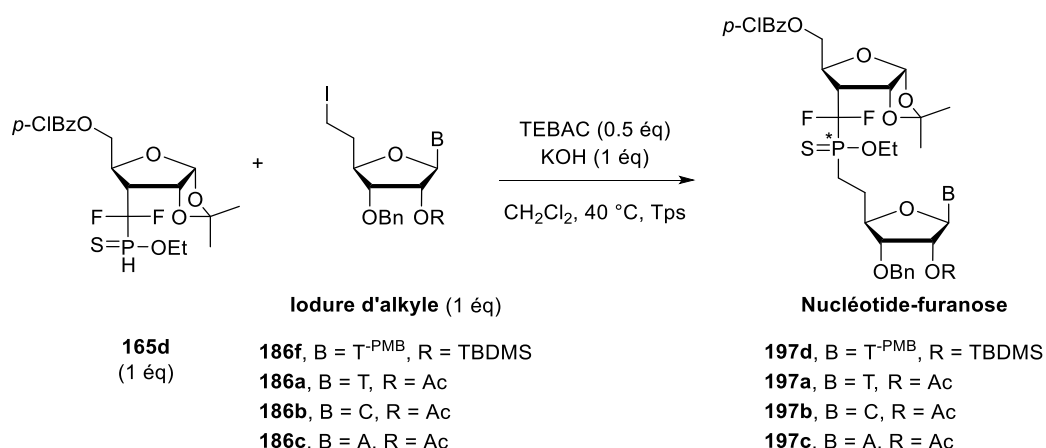
## II.2. Formation de la seconde liaison P-C par voie ionique : réaction de Michaelis-Becker

Aucun résultat probant n'ayant été obtenu par addition radicalaire, une stratégie ionique a logiquement été envisagée. La réaction de Michaelis-Becker a été testée sur un nucléoside protégé et sur un autre non-protégé (Tableau 11).

Dans un premier temps, la formation de la seconde liaison P-C a été testée entre le *H*-phosphinothioate **165d** et le nucléoside thymine protégé **186f** (entrée 1), en suivant les conditions mise au point précédemment au laboratoire. En présence d'un médiateur de transfert de phase (TEBAC) et d'une

<sup>141</sup> (a) Muller, E.; Gasparutto, D.; Jaquinod, M.; Romieu, A. and Cadet, J. *Tetrahedron* **2000**, *56*, 8689-8701.  
 (b) Boussicault, F.; Kaloudis, P.; Caminal, C.; Mulazzani, Q. G. and Chatgililoglu, C. *J. Am. Chem. Soc.* **2008**, *130*, 8377-8385.

base (KOH), l'anion phosphinothioate généré réalise alors une substitution nucléophile de type  $S_N2$  sur l'iodure d'alkyle **186f** pour former le difluorophosphinothioate **197d** souhaité. Ce dernier a été isolé avec un rendement moyen de 63% et sous forme d'un mélange diastéréoisomérique au niveau du phosphore (ratio 1.4/1, déterminé par spectrométrie de RMN du  $^{31}P$ ). Il est à noter que de l'iodure d'alkyle **186f** non-consommé a pu être isolé après purification par chromatographie sur gel de silice à hauteur de 21%. Les substrats étant en mélange équimolaire, ce résultat peut être expliqué soit par la présence d'impuretés dans le *H*-phosphinothioate **165d**, soit par la dégradation partielle de celui-ci dans les conditions réactionnelles.

Tableau 11 : Réaction de Michaelis-Becker entre le *H*-phosphinothioate **165d** et un alkyle iodé **186**

Entrée	N°	Base	R	Tps (h) <sup>(a)</sup>	Nucléotide-furanose	
					Rdt (%) <sup>(b)</sup>	Ratio <sup>(d)</sup>
1	<b>186f</b>	T- <sup>PMB</sup>	TBDMS	3	<b>197d</b> , 63% (21%) <sup>(c)</sup>	1.7/1
2	<b>186a</b>	T	Ac	3	<b>197a</b> , 47% (16%) <sup>(c)</sup>	1.7/1
3	<b>186b</b>	C	Ac	20	<b>197b</b> , 55%	2.4/1
4	<b>186c</b>	A	Ac	40	<b>197c</b> , 52% (13%) <sup>(c)</sup>	2/1

<sup>(a)</sup> Suivi de réaction par CCM jusqu'à la consommation totale de **165d**. <sup>(b)</sup> Rendements isolés. <sup>(c)</sup> Rendements isolés des substrats iodures d'alkyle non-consommés. <sup>(d)</sup> Dû à la chiralité du phosphore, ratios déterminés à partir des mélanges bruts réactionnels par spectrométrie de RMN du  $^{31}P$ .

Dans un second temps, la réaction de Michaelis-Becker a été réalisée entre le *H*-phosphinothioate **165d** et le dérivé nucléosidique non-protégé **186a** (entrée 2). En suivant les conditions décrites précédemment, le difluorophosphinothioate **197a** a été obtenu avec un rendement de 47% et sous la forme d'un mélange diastéréoisomérique 1.7/1 (ratio P\* déterminé par analyse de RMN du  $^{31}P$ ). Dans ce cas, l'iodure d'alkyle **186a** a aussi été récupéré en fin de réaction à hauteur de 16%. Il est à noter qu'un essai, dans lequel le *H*-phosphinothioate **165d** a été mis en excès (1.2 éq), a été réalisé afin de limiter la quantité d'iodure d'alkyle en fin de réaction. Cependant, l'iodure d'alkyle a de nouveau été

récupéré après purification. Les autres réactions de Michaelis-Becker ont donc été effectuées en mélange équimolaire.

Les essais de réactions de Michaelis-Becker avec le nucléoside thymine non-protégée **186a** étant concluants, la préparation d'autres dérivés mixtes nucléotide-furanose a été entreprise à partir du difluorophosphinothioate **165d** et des iodures d'alkyles **186b** et **186c** portant respectivement la cytosine ou l'adénine.

En suivant les conditions décrites précédemment mais avec un temps de réaction un peu plus long, le couplage du fragment **165d** et du nucléoside cytosine **186b** a conduit à la formation du difluorophosphinothioate **197b** (entrée 3). Dans ce cas, l'iodure d'alkyle **186b** non-consommé a pu être séparé par plusieurs purifications successives en phase normale ; cependant, un rendement isolé de ce dernier n'a pu être déterminé. Ici, le composé **197b** a été obtenu sous la forme d'un mélange diastéréoisomérique 2.4/1 (ratio P\* déterminé par spectrométrie de RMN du  $^{31}\text{P}$  du mélange brut réactionnel), avec un rendement de 55%. La formation de la seconde liaison P-C entre le *H*-phosphinothioate **165d** et le nucléoside adénine **186c** a aussi été réalisée selon les conditions optimisées (Entrée 4). La purification par chromatographie en phase normale n'a pas permis de séparer l'iodure d'alkyle **186c** n'ayant pas réagi et le composé désiré **197c**. Par conséquent, une purification en phase inverse a été entreprise. Le composé **197c** a finalement été isolé avec un rendement de 52% sous la forme d'un mélange diastéréoisomérique 2/1 (ratio P\* déterminé par spectrométrie de RMN du  $^{31}\text{P}$  du mélange brut réactionnel) et le dérivé iodé **186c** a été récupéré à hauteur de 14%.

Il est à noter que le *H*-phosphinothioate **165d** a été introduit dans la réaction de Michaelis-Becker sous la forme d'un mélange diastéréoisomérique avec un ratio 1/1 (déterminé par spectrométrie de RMN du  $^{31}\text{P}$ ). Les différences notables des ratios obtenus après la formation de la seconde liaison P-C ne sont pour l'instant pas expliquées.

Malgré les résultats peu concluants de la construction de la seconde liaison phosphore-carbone par voie radicalaire, la voie ionique a permis l'obtention de quatre intermédiaires avancés **197a** (T), **197b** (C), **197c** (A) et **197d** (T- $\text{P}^{\text{MB}}$ ), incluant pour l'instant une seule des deux bases nucléiques requises pour les dinucléotides-cibles modifiés.

### III. Etude de la voie C : application sur un substrat modèle

Pour rappel, le dinucléotide **183** pourrait être obtenu à partir de l'intermédiaire **197** qui résulterait de l'addition du *H*-phosphinothioate **187** sur la difluorooléfine **163** (Schéma 55). Cette stratégie est innovante puisqu'à ce jour aucune réaction d'addition radicalaire d'un *H*-phosphinothioate **187** sur une difluorooléfine **163** n'est décrite dans la littérature.

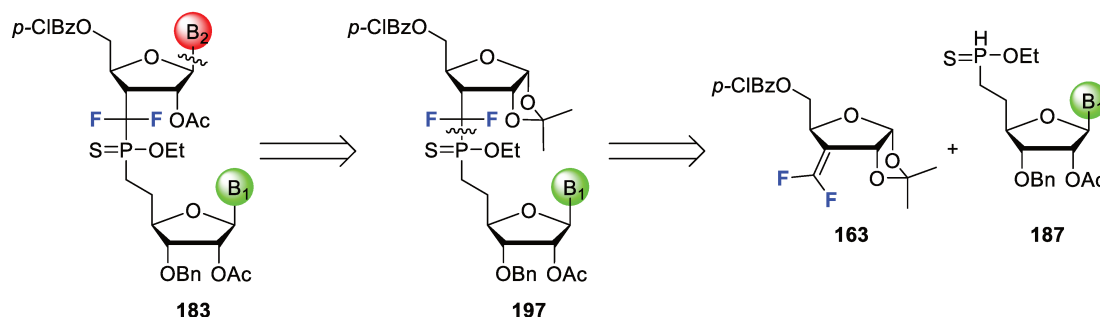
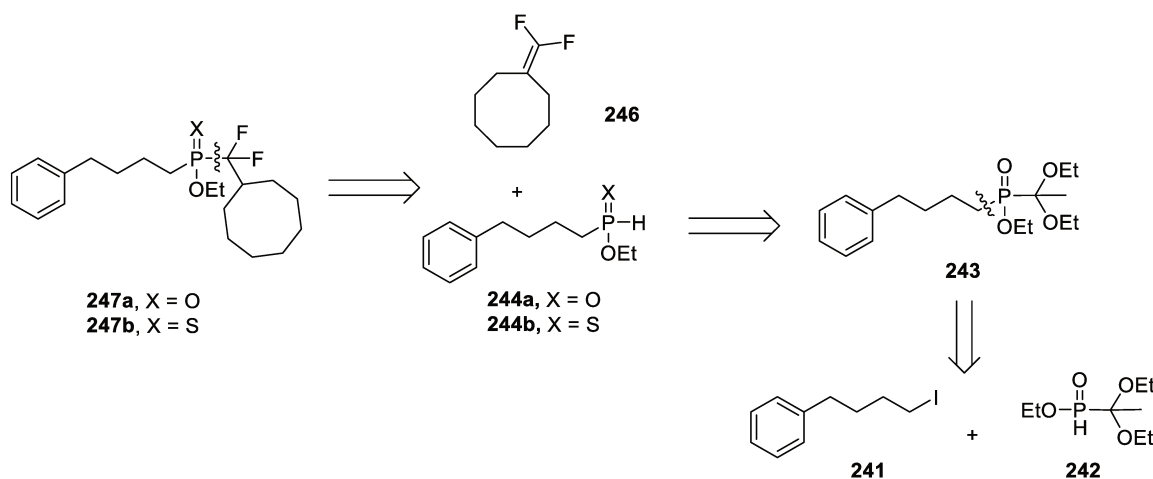


Schéma 55 : Stratégie C

Afin de vérifier la faisabilité de cette stratégie de synthèse, une étude préliminaire a été réalisée sur un substrat modèle.

#### III.1. Etude sur un substrat modèle : rétrosynthèse

La rétrosynthèse du substrat modèle **247** est présentée dans le Schéma 56. En disconnectant au niveau de la liaison P-CF<sub>2</sub>, le substrat modèle **247** pourrait être obtenu à partir de la difluorooléfine **246** et le *H*-phosphinate **244a** (H-P=O) ou le *H*-phosphinothioate **244b** (H-P=S). Cette étape serait l'occasion d'étudier la différence de réactivité entre H-P=O et H-P=S déjà observée lors des travaux précédents du laboratoire (Introduction, V.2). Le composé **244b** serait préparé via une étape de thiolation en présence du réactif de Lawesson sur **244a**. Quant au *H*-phosphinate **244a**, celui-ci serait obtenu par déprotection du phosphinate **243**, qui résulterait de l'association entre l'iodure d'alkyle **241** et le *H*-phosphinate **242**.

Schéma 56 : Rétrosynthèse du substrat modèle **247**

## III.2. Formation de la première liaison P-C

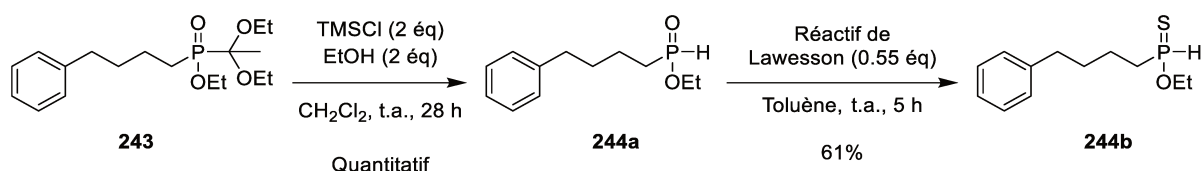
Dans un premier temps, la formation de la première liaison phosphore-carbone par substitution  $S_N2$  a été étudiée à partir du 1-iodo-4-phénylbutane commercial **241** et du *H*-phosphinate **242** synthétisé suivant un protocole de la littérature.<sup>142</sup> Un travail d'optimisation a été réalisé afin d'obtenir le composé **243** (Tableau 12). Tout d'abord, en suivant une procédure expérimentale,<sup>143</sup> le *H*-phosphinate **242** a été mis en large excès en présence de KHMDS (entrée 1). La formation du composé **243** a été observée par analyse de RMN du  $^{31}\text{P}$  du mélange réactionnel brut. Cependant, ce dernier n'a pas pu être isolé pur du fait de sa coélution avec le *H*-phosphinate **242** durant la purification par chromatographie sur gel de silice. Par conséquent, l'iodure d'alkyle **241** a été mis en excès, ce qui a permis d'isoler le composé **243** avec un rendement moyen de 36% (entrée 2). Un dernier test a été réalisé (entrée 3) dans lequel tous les réactifs ont été mis en quantité équimolaire et après un temps de réaction allongé à 22 h, le phosphinate **243** a été isolé avec un bon rendement de 76%.

Tableau 12 : Optimisation des conditions pour former **243**

Entrée	241	242	KHMDS	Temp. (°C)	Tps (h)	Rdt
1	1 éq	2.5 éq	2.5 éq	-78 à t.a.	6.5 h	/
2	1.5 éq	1.05 éq	1 éq	-78 à -20	5.7 h	36%
3	1.05 éq	1 éq	1 éq	-78 à t.a.	22 h	76%

## III.3. Etude de la réactivité H-P=O versus H-P=S pour la formation de la seconde liaison P-C

Afin d'effectuer la formation de la seconde liaison P-C par réaction  $S_N2$ , le phosphinate **243** a été déprotégé à l'aide de TMSCl et d'éthanol avec un rendement quantitatif (Schéma 57).<sup>143</sup> Le *H*-phosphinate résultant **244a** a aussi été transformé en *H*-phosphinothioate **244b** à l'aide du réactif de Lawesson avec un rendement de 61%, en vue d'étudier une éventuelle différence de réactivité H-P=O/H-P=S.

Schéma 57 : Synthèse du *H*-phosphinothioate **244b**<sup>142</sup> Fougère, C.; Guénin, E.; Hardouin, J. and Lecouvey, M. *Eur. J. Org. Chem.* **2009**, 2009, 6048-6054.<sup>143</sup> Fairhurst, R. A.; Collingwood, S. P. and Lambert, D. *Synlett* **2001**, 2001, 0473-0476.

La difluorooléfine modèle **246** a été synthétisée à partir de la cyclooctanone commerciale **245**, par réaction d'oléfination en présence de dibromodifluorométhane et de HMPT (Schéma 58).<sup>144</sup> Le faible rendement de cette réaction peut être expliqué par le fait qu'il s'agisse d'un composé très volatil. Une perte significative de produit est observée lors de l'évaporation des solvants après traitement et purification.

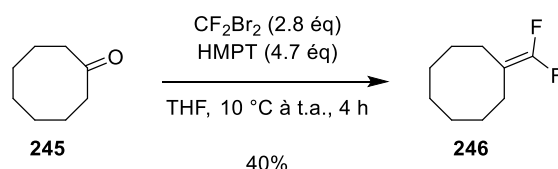


Schéma 58 : Synthèse de la difluorooléfine modèle **246**

Grâce aux fragments **244a** et **244b**, la réaction d'addition radicalaire avec l'oléfine modèle a pu être étudiée et les résultats sont présentés dans le Tableau 13. Lors de la réaction entre le *H*-phosphinate **244a** et la difluorooléfine **246**, avec l'utilisation de DLP comme initiateur radicalaire, le composé-cible n'a pas été observé. Seuls les réactifs de départ (**244a** et **246**) et un sous-produit **248a** (<sup>31</sup>P δ : 25 ppm ; ratio **244a/248a** : 2/1 déterminé par spectrométrie de RMN du <sup>31</sup>P) provenant de la dégradation (oxydation ?) de **244a** ont été observés en spectrométrie de RMN (entrée 1). Piettre a obtenu le motif difluorophosphinate par réaction entre un *H*-phosphinate et une difluorooléfine-modèle dans des conditions similaires.<sup>108</sup> Ainsi, dans notre cas, la formation de la molécule-cible **247a** aurait dû être détectée dans les spectres de RMN du milieu réactionnel. Néanmoins, seule la dégradation de **244a** a été observée.

En revanche, lors de la réalisation de la réaction radicalaire entre le *H*-phosphinothioate **244b** et la difluorooléfine **246**, le composé-cible **247b** a été obtenu en mélange (rendement de 74%) avec un sous-produit non-séparable **248b** (<sup>31</sup>P δ : 104 ppm) dans un ratio 1/1 (déterminé par spectrométrie de RMN du <sup>31</sup>P) (entrée 2). Dans ce cas, le *H*-phosphinothioate **244b** a totalement été transformé en **247b** et **248b**. Comme dans les travaux précédents du laboratoire (Introduction, V.2), une différence de réactivité a été observée entre H-P=O **244a** et H-P=S **244b**.

En dépit de la présence de ce sous-produit (**248b**), un essai de formation de la seconde liaison phosphore-carbone a été réalisé sur la difluorooléfine-cible **163** en utilisant des conditions identiques mais dans ce cas, seuls le sous-produit **248b** et la difluorooléfine **163** ont été observés en spectrométrie de RMN du <sup>31</sup>P et du <sup>19</sup>F, respectivement (entrée 3). En utilisant l'initiateur radicalaire TBPP à la place de DLP, un résultat similaire a été obtenu (entrée 4).

<sup>144</sup> Naae, D. G. and Burton, D. J. *Synth. Commun.* **1973**, *3*, 197-200.

Afin de déterminer l'origine de ce sous-produit, une réaction de contrôle a été réalisée entre le *H*-phosphinothioate **244b** et les deux initiateurs radicalaires (entrée 5 et 6). Dans les deux cas, *H*-phosphinothioate **244b** a totalement été converti en un sous-produit **248b** identique à celui formé précédemment. A l'heure actuelle, les structures de **248a** et **248b** n'ont pas été déterminées. Cependant, il est à noter que la disparition du signal du P-H en spectrométrie de RMN du  $^1\text{H}$  a été observée.

Tableau 13 : Essais de formation de la seconde liaison phosphore-carbone

$\text{244a, X = O}$   
 $\text{244b, X = S}$

$\text{247a, X = O}$   
 $\text{247b, X = S}$

$\text{248a, X = O}$   
 $\text{248b, X = S}$

Entrée	X	Difluorooléfine	Initiateur radicalaire	Tps (h)	Temp. (°C)	247 <sup>(c)</sup>	248 <sup>(c)</sup>	Substrat 244 <sup>(c)</sup>
1	O	<b>246</b>	DLP <sup>(a)</sup>	4	80	/	1	2
2	S	<b>246</b>	DLP <sup>(a)</sup>	5	80	1	1	/
3	S	<b>163</b>	DLP <sup>(a)</sup>	5	80	/	1	/
4	S	<b>163</b>	TBPP <sup>(b)</sup>	24	66	/	1	/
5	S	/	DLP <sup>(a)</sup>	4	80	/	1	/
6	S	/	TBPP <sup>(b)</sup>	24	66	/	1	/

<sup>(a)</sup> Ajout de 0.2 éq toutes les 2 h, au total utilisation de 0.6 éq de DLP. <sup>(b)</sup> 0.3 éq de TBPP. <sup>(c)</sup> Ratio déterminé par spectrométrie de RMN du  $^{31}\text{P}$  découplé.

Des conditions microondes ont été développées par Parsons afin de former une liaison phosphore-carbone entre un *H*-phosphinothioate **249** et une oléfine **250** avec un rendement de 93%.<sup>109</sup> Ces conditions ont donc été testées mais dans notre cas seuls les substrats de départ ont été observés après traitement (Schéma 59).

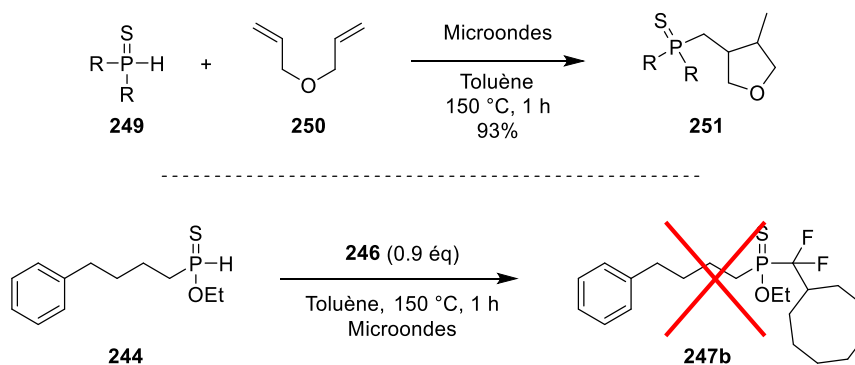


Schéma 59 : Essai de formation de la seconde liaison P-C à l'aide de microondes

Comme dans les travaux précédents réalisés au laboratoire, une différence de réactivité entre le *H*-phosphinate **244a** et le *H*-phosphinothioate **244b** a pu être observée. Le composé P=S s'est révélé plus réactif que son homologue P=O. Cependant en raison de la quantité non-négligeable de sous-produit formé et de la non-réactivité du *H*-phosphinothioate **244b** sur la difluorooléfine-cible **163**, la stratégie C a été abandonnée.

#### IV. Conclusions

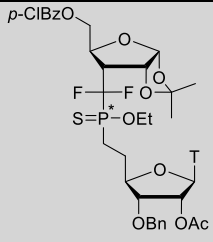
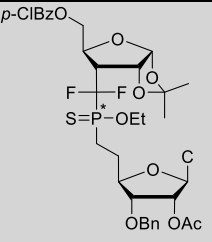
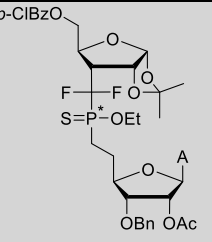
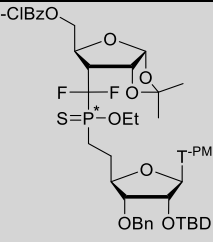
Plusieurs stratégies de construction de la seconde liaison phosphore-carbone ont été étudiées à partir des substrats synthétisés précédemment (Chapitre 2).

Dans un premier temps, la voie C a été réalisée sur un substrat modèle. Cependant celle-ci a été abandonnée à la suite de la formation d'un sous-produit en quantité non-négligeable lors de l'étude de la réactivité d'un *H*-phosphinate ou *H*-phosphinothioate modèle sur une difluorooléfine modèle. De plus, aucune réaction entre ce même *H*-phosphinothioate modèle et la difluorooléfine furanosique cible n'a été observée. Néanmoins, cette étude a tout de même permis d'observer une différence de réactivité entre H-P=O et H-P=S.

Ensuite, la voie A a permis la formation de l'intermédiaire **196** où une base nucléique pourrait être introduite par la suite ; cependant, celui-ci a été obtenu avec un faible rendement de 7%.

La formation de la seconde liaison phosphore-carbone par voie radicalaire (Voie B) a conduit à la formation de traces du composé souhaité, tandis que la voie ionique, par réaction de Michaelis-Becker, a permis d'obtenir quatre intermédiaires mixtes de type nucléotide-furanose, avec des rendements compris entre 47% et 63% (Tableau 14). La suite de la synthèse vers un hétérodinucléotide à partir des différents intermédiaires obtenus sera discutée dans le chapitre suivant.

Tableau 14 : Récapitulatifs des intermédiaires synthétisés par la voie B

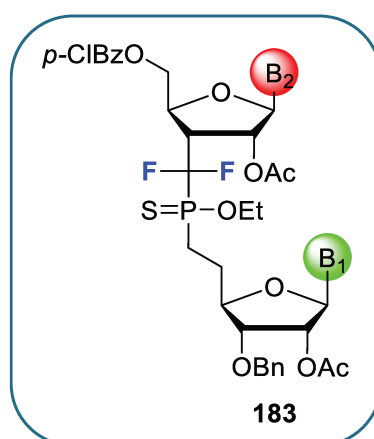
Nucléosides	Thymine	Cytosine	Adénine	Thymine <sup>PMB</sup>
<b>Produits</b>	 <p style="text-align: center;"><b>197a</b></p>	 <p style="text-align: center;"><b>197b</b></p>	 <p style="text-align: center;"><b>197c</b></p>	 <p style="text-align: center;"><b>197d</b></p>
<b>Rendements</b>	47%	55%	52%	63%
<b>P* ratio</b>	1.7/1	2.4/1	2/1	1.7/1





# Chapitre 4

*Insertion d'une seconde base nucléique*





## Sommaire

<b><i>I. Introduction d'une seconde base nucléique</i></b> .....	<b>129</b>
<b>I.1. A partir du furanose 197d incorporant une thymine protégée</b> .....	<b>129</b>
<b>I.2. A partir de l'intermédiaire mixte thymine-furanose 197a</b> .....	<b>131</b>
<b><i>II. Phénomène de trans-N-glycosylation</i></b> .....	<b>136</b>
<b>II.1. Etat de l'art</b> .....	<b>136</b>
<b>II.2. Solutions envisagées</b> .....	<b>142</b>
II.2.a. Addition contrôlée .....	142
II.2.b. Ordre d'insertion .....	144
II.2.b.i. Utilisation d'un intermédiaire cytosine .....	144
II.2.b.ii. Utilisation d'un intermédiaire adénine .....	152
<b><i>III. Conclusions</i></b> .....	<b>153</b>



## I. Introduction d'une seconde base nucléique

L'introduction d'une seconde base nucléique sur le furanose du haut a été étudiée afin d'accéder aux hétérodinucléotides **183** (Schéma 60). Une disconnection au niveau de la position anomérique portant la base nucléique B<sub>2</sub> de **183** générerait un synthon de type nucléotide-furanose dont un équivalent synthétique est représenté par l'intermédiaire **182**. Comme présenté précédemment, celui-ci peut être relié au composé **197** par une modification du groupement protecteur du diol C1'-C2'.

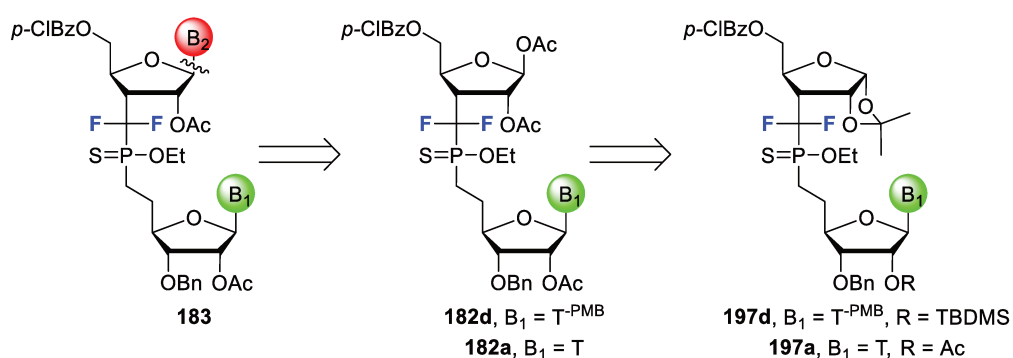


Schéma 60 : Rétrosynthèse de l'hétérodinucléotide **183**

Cette approche repose donc sur l'obtention de triacétates intermédiaires **182** provenant d'une séquence acétylase de l'acétonide-protection du diol résultant en C1'-C2' à partir des intermédiaires de synthèse **197** obtenus précédemment (Chapitre 3, II.2).

### I.1. A partir du furanose **197d** incorporant une thymine protégée

Dans un premier temps, le remplacement du groupement silylé TBDMS par un groupement acétyle a été réalisé afin de résister aux conditions acides de déprotection-protection du diol C1'-C2' et d'insertion de la base nucléique (Schéma 61). Pour ce faire, l'éther silylé a été clivé à l'aide d'une solution d'acide fluorhydrique dans la pyridine avec un rendement de 82%. L'alcool résultant **252** a ensuite été protégé sous la forme d'acétate avec un rendement analogue de 83%.

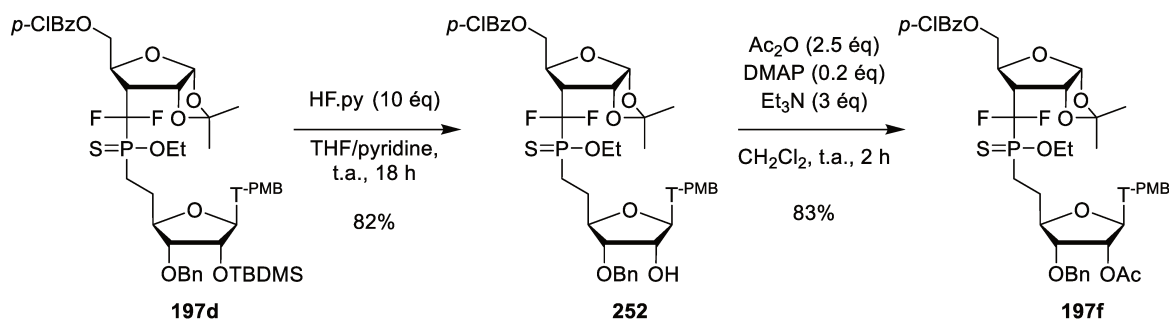


Schéma 61 : Modification du groupement protecteur en C2'

En suivant des conditions développées au laboratoire, une séquence clivage de l'acétonide-protection des alcools C1'-C2' résultants sous forme d'acétates a été réalisée sur le composé **197f** (Schéma 62).

Le triacétate **182d** a ensuite été placé en présence d'uracile silylé **225e** sous les conditions classiques de Vorbrüggen permettant le greffage de la seconde base nucléique avec, cependant, un rendement très moyen. Il est à noter la présence de sous-produits dans le mélange brut réactionnel, détecté par spectrométrie de RMN du  $^{19}\text{F}$  découplé du proton (ratio **183a**/sous-produit : 1.7/1).<sup>145</sup> Néanmoins, ceux-ci n'ont pas pu être isolés lors de la purification. Un premier hétérodinucléotide thymine<sup>PMB</sup>-uracile **183a**, noté T<sup>PMB</sup>-U,<sup>146</sup> a ainsi été synthétisé en quatre étapes à partir de l'intermédiaire mixte nucléotide-furanose **197f** avec un rendement global de 14%.

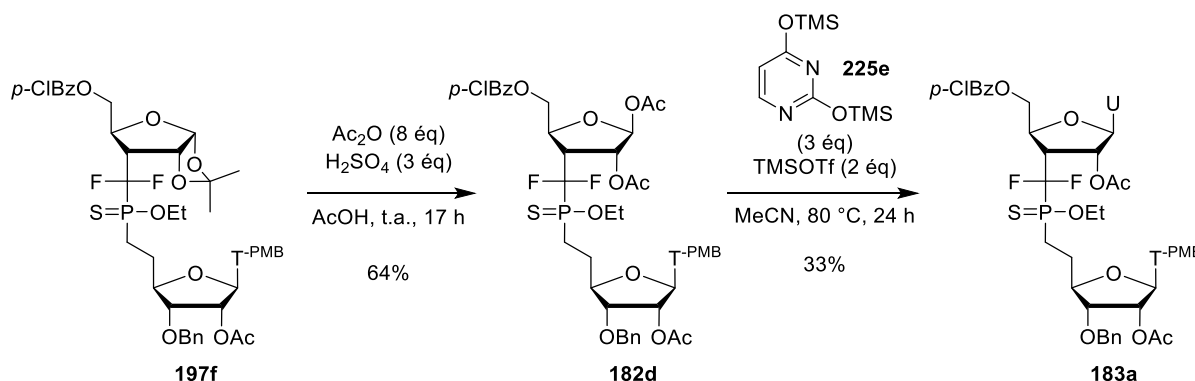


Schéma 62 : Introduction d'une seconde base nucléique

En suivant un protocole de la littérature,<sup>147</sup> un essai de déprotection du groupement PMB, présent sur la thymine, a été réalisé à l'aide de nitrate de cérium et d'ammonium à température ambiante, puis à 40 °C (Schéma 63). Après deux jours, aucune réaction n'a eu lieu et seul l'hétérodinucléotide **183a** de départ a été observé dans le mélange brut réactionnel en spectrométrie de RMN du  $^1\text{H}$ , du  $^{19}\text{F}$  et du  $^{31}\text{P}$ .

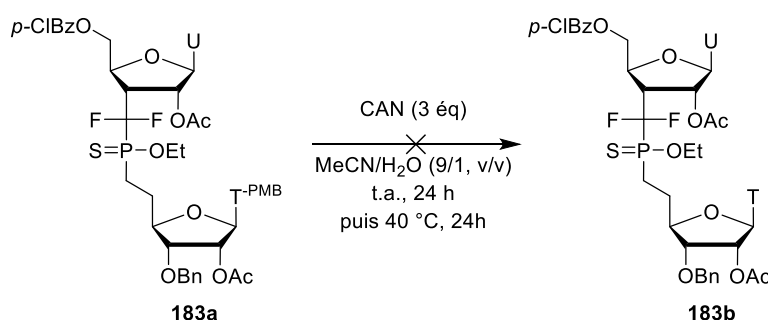


Schéma 63 : Tentative de déprotection de la thymine

La synthèse d'un premier hétérodinucléotide mixte T<sup>PMB</sup>-U **183a** a donc été réalisée avec succès. Cependant, le groupement protecteur n'a pu être retiré à ce jour. D'autres méthodes de

<sup>145</sup> Déplacements chimiques des sous-produits :

$^{19}\text{F}$  { $^1\text{H}$ } :  $\delta$  -100.8 (dd,  $J$  = 290.2 Hz, 82.7 Hz), -100.9 (dd,  $J$  = 292.2 Hz, 108.2 Hz), -111.7 (dd,  $J$  = 290.2 Hz, 102.9 Hz), -114.7 (dd,  $J$  = 292.2 Hz, 80.8 Hz)

<sup>146</sup> Note : la notation dinucléotide B<sub>1</sub>-B<sub>2</sub> signifie que la première base citée (B<sub>1</sub>) est celle du furanose du bas tandis que la seconde (B<sub>2</sub>) est celle du furanose du haut.

<sup>147</sup> Nawale, G. N.; Gore, K. R.; Höbartner, C. and Pradeepkumar, P. I. *Chem. Comm.* **2012**, 48, 9619-9621.

déprotection devraient être essayées notamment par utilisation d' $\text{AlCl}_3$ <sup>148</sup> ou de DDQ.<sup>149</sup> Cependant par manque de substrat, celles-ci n'ont pas pu être testées. En parallèle à cette étude, l'introduction d'une seconde base nucléique a été explorée à partir du fragment mixte thymine non-protégé-furanose **197a**.

### I.2. A partir de l'intermédiaire mixte thymine-furanose **197a**

Comme précédemment, une modification des groupements protecteurs en position C1' et C2' a été réalisée en milieu acide conduisant à la formation de l'intermédiaire triacétate **182a**, isolé avec un rendement de 44%, malgré une conversion totale observée en spectrométrie de RMN du  $^{19}\text{F}$  (Schéma 64). Une dégradation sur la colonne de silice peut être à l'origine de ce rendement moyen. Afin de comparer cette stratégie avec celle du laboratoire (présentée en Introduction, V.2), une seconde thymine a été introduite avec succès et l'homodinucléotide T-T **180a** a été isolé avec un rendement de 46%.

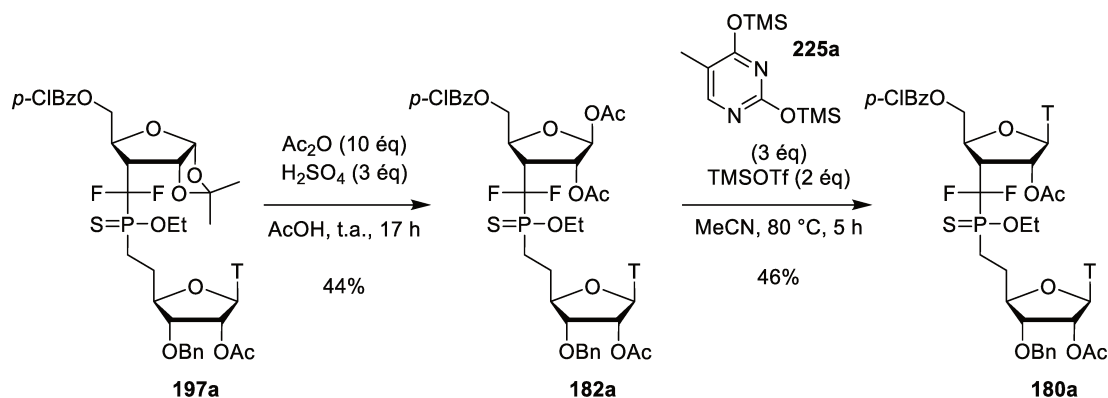


Schéma 64 : Insertion d'une seconde thymine sur l'intermédiaire thymine non protégée **197a**

La stratégie développée précédemment au laboratoire par les Drs. Jun Zhang et Emilie Lambert avait permis de synthétiser l'homodinucléotide T-T **180a**, par addition séquentielle des deux bases pyrimidiques, en quatre étapes avec un rendement global de 8% (rendements en bleu sur le Schéma 65). L'étape de Michaelis-Becker a été réalisée avec deux furanoses (**165a** et **178**) portant chacun un groupement acétonide protecteur du diol C1'-C2'. Après conversion du bisacétate en intermédiaire tétraacétate **179**, une addition contrôlée d'une seule thymine a permis d'obtenir l'intermédiaire mixte thymine-furanose **182a** avec un rendement de 58%. Cet intermédiaire étant commun aux deux stratégies, nous pouvons ainsi comparer l'efficacité synthétique des deux approches. Dans le cas de la nouvelle stratégie présentée, **182a** a été obtenu en deux étapes à partir de **186a** après une réaction de Michaelis-Becker et la conversion de l'acétonide **197a** résultant en

<sup>148</sup> Takahiko, A.; Yasuhiro, T.; Masahiro, K.; Hiroyuki, N. and Shoichiro, O. *Bull. Chem. Soc. Jpn.* **1991**, *64*, 2266-2269.

<sup>149</sup> Horita, K.; Yoshioka, T.; Tanaka, T.; Oikawa, Y. and Yonemitsu, O. *Tetrahedron* **1986**, *42*, 3021-3028.

triacétate **182a**. Pour terminer cette synthèse, une seconde thymine a été introduite afin de former l'homodinucléotide T-T **180a**. Grâce à la nouvelle stratégie, **180a** a donc été obtenu en trois étapes à partir de **165d** avec un rendement global de 12% (rendements en rouge sur le Schéma 65). Dans le cas de la synthèse d'homodinucléotide, la stratégie à privilégier reste celle mise au point au laboratoire précédemment où le fragment tétraacétate **179** permet l'introduction directe de deux bases nucléiques identiques (66% de rendement pour l'introduction de deux thymines). La nouvelle stratégie permet, elle, d'accéder efficacement à des hétérodinucléotides. Celle-ci présente donc un rendement légèrement meilleur que la synthèse de première génération, et permet un gain relatif d'une étape chimique.

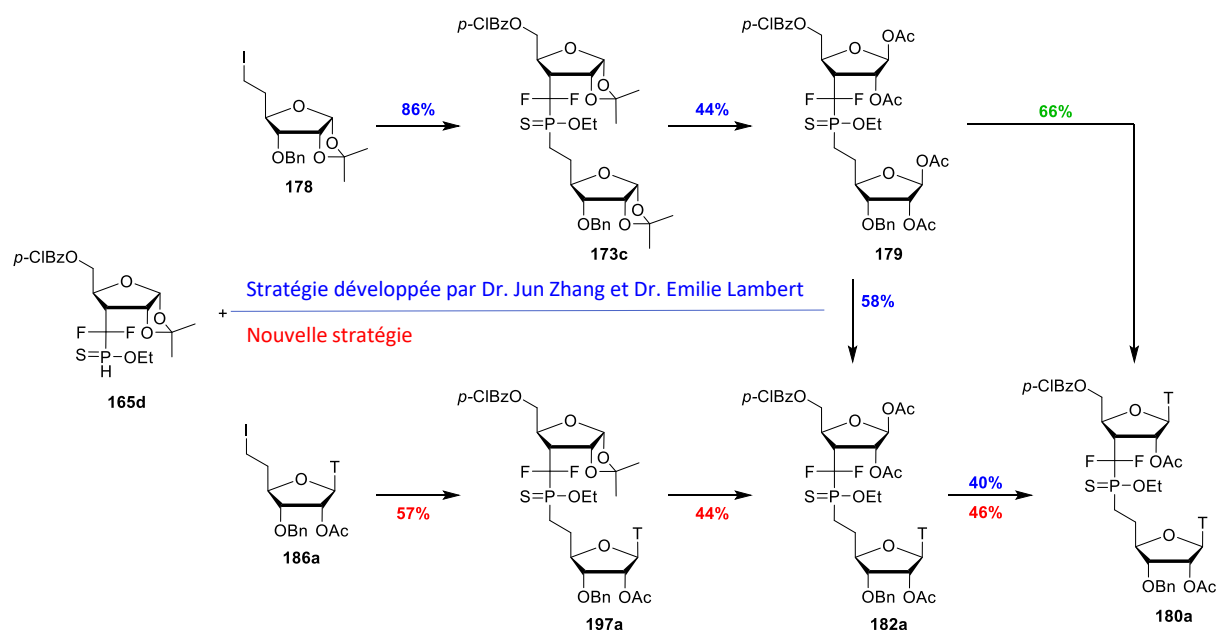


Schéma 65 : Comparaison des voies développées menant à un dinucléotide thymine-thymine **180a**

Par la suite, l'introduction de bases nucléiques différentes, sur l'intermédiaire **182a**, a été étudiée. Tout d'abord, la synthèse de l'hétérodinucléotide T-U a été envisagée afin de comparer cette stratégie à celle ayant mené à l'hétérodinucléotide T<sup>PMB</sup>-U. Ainsi, l'introduction de l'uracile a été réalisée dans les conditions classiques de Vorbrüggen à partir de **182a** (Schéma 66). Un mélange complexe, difficilement exploitable en spectrométrie de RMN du <sup>1</sup>H, <sup>19</sup>F et <sup>31</sup>P, a été obtenu. Cependant, il est à noter la présence d'un signal à -78.4 ppm dans le spectre de RMN du <sup>19</sup>F découplé du proton du mélange brut réactionnel, témoignant d'une possible dégradation de **182a**. Deux dinucléotides ont tout de même été isolés après purification par chromatographie en phase inverse. L'hétérodinucléotide thymine-uracile **183b** attendu a été détecté à l'état de traces. Quant au second dinucléotide, isolé à hauteur de 5%, une disparition des signaux caractéristiques de la thymine a été observée dans le spectre de RMN du <sup>1</sup>H. L'hypothèse d'une substitution de la thymine par l'uracile ayant mené à la formation d'un homodinucléotide U-U a été avancée. Cette dernière a été validée par la présence de

signaux caractéristiques de deux uraciles sur le spectre proton, et corroborée par l'analyse par spectrométrie de masse haute résolution. Un échange de base entre la thymine et l'uracile a ainsi été mis en évidence. La faible quantité de produit obtenu n'a pas permis de réaliser l'ensemble des analyses par spectrométrie de RMN ( $^{13}\text{C}$ , COSY, HSQC, HMBC). Un second essai a été réalisé à température ambiante pendant 24 h. Dans ce cas, aucune introduction d'uracile n'a eu lieu et seul le substrat de départ **182a** a été observé dans les différentes analyses RMN.

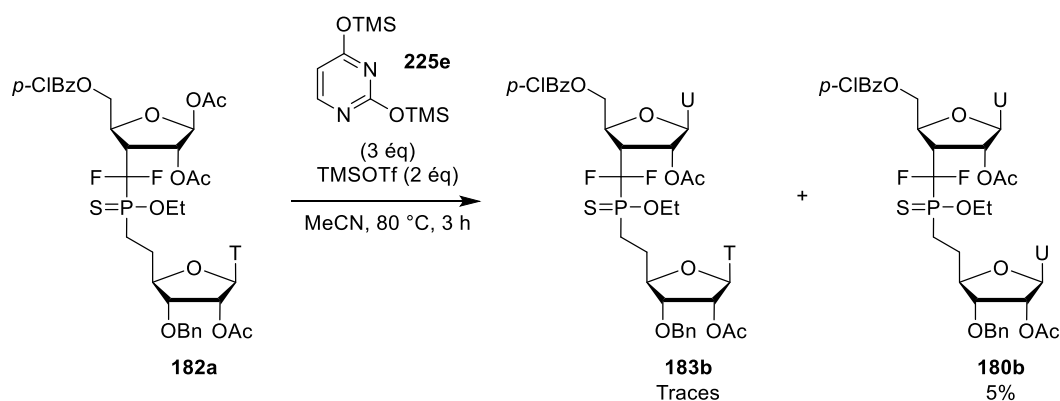


Schéma 66 : Insertion de l'uracile sur l'intermédiaire **182a**

Afin de vérifier si l'échange de base observé précédemment n'est pas spécifiquement dû à l'uracile, l'introduction d'une autre base nucléique (la cytosine) a alors été étudiée sur l'intermédiaire **182a**. Néanmoins, un phénomène similaire a été observé (Schéma 67).

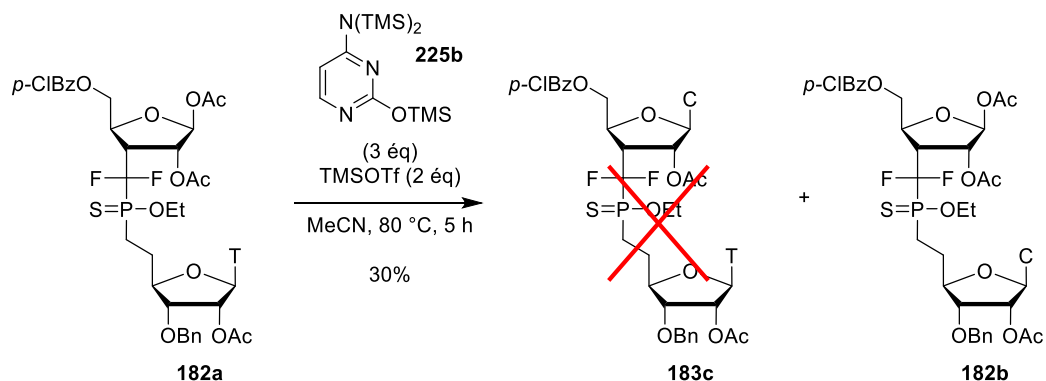


Schéma 67 : Insertion d'une cytosine sur l'intermédiaire **127**

Dans le mélange brut réactionnel, un seul produit (sous la forme de deux diastéréoisomères dû au phosphore chiral) a été observé par spectrométrie de RMN du  $^{19}\text{F}$ . Ceux-ci ont été isolés avec un rendement de 30%. La comparaison entre les spectres de RMN du  $^1\text{H}$  de **182a**, du mélange brut réactionnel et du produit purifié a clairement montré la disparition des signaux caractéristiques de la thymine (NH, CH et méthyle) (Figure 52). De plus, l'apparition du signal de l'amine de la cytosine est à noter dans les spectres de RMN du  $^1\text{H}$  du mélange brut réactionnel et après purification. Une analyse par spectrométrie de masse haute résolution a confirmé la formation du composé **182b**. A ce stade, le remplacement de la thymine par une cytosine a été établi. Cependant, une interrogation subsiste

concernant la position anomérique occupée par la cytosine. La faible quantité obtenue après purification n'a pas permis de réaliser les expériences de RMN complémentaires pour déterminer la position exacte de la cytosine. Néanmoins, par la suite (voir II.2.b.i), le substrat **182b** a été synthétisé par une autre voie à partir d'un fragment porteur au départ de la cytosine sur le furanose du bas (**197b**). Une comparaison des spectres de RMN du  $^1\text{H}$ , du  $^{19}\text{F}$  découplé du  $^1\text{H}$  et du  $^{31}\text{P}$  découplé du  $^1\text{H}$  a été réalisée et a permis de valider la structure de **182b** (Figure 53, Figure 54 et Figure 55).

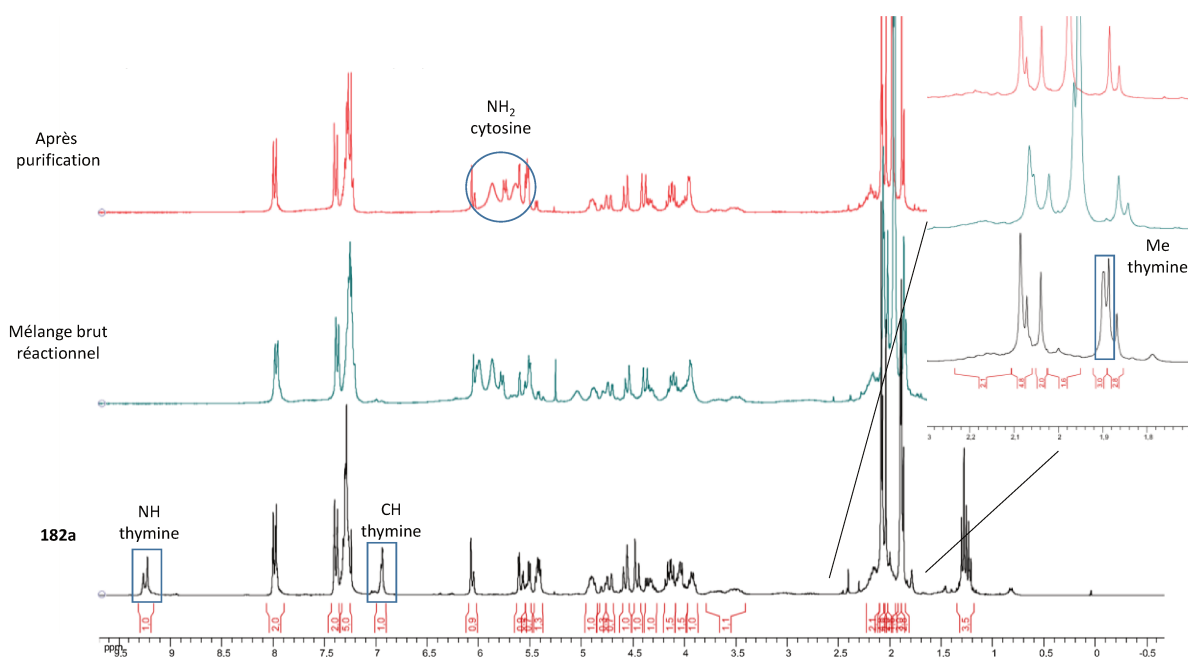


Figure 52 : Comparaison des spectres RMN du  $^1\text{H}$  lors de l'étape d'introduction d'une cytosine sur **182a**

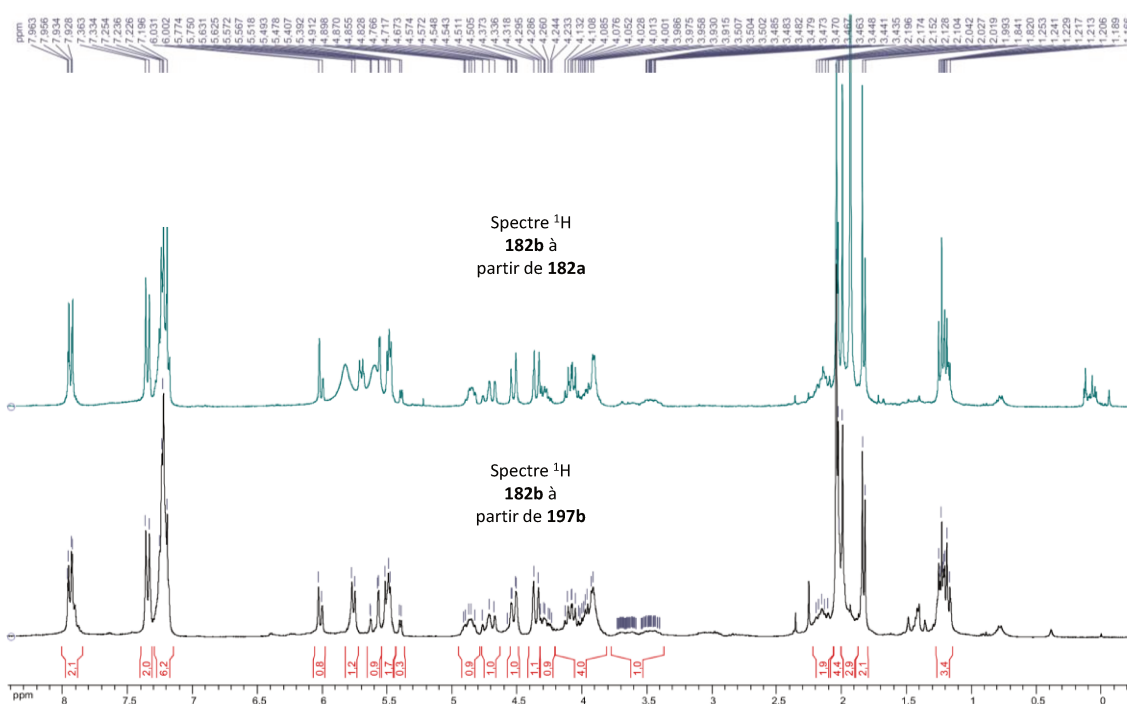


Figure 53 : Comparatif des spectres RMN du  $^1\text{H}$  de **182b**

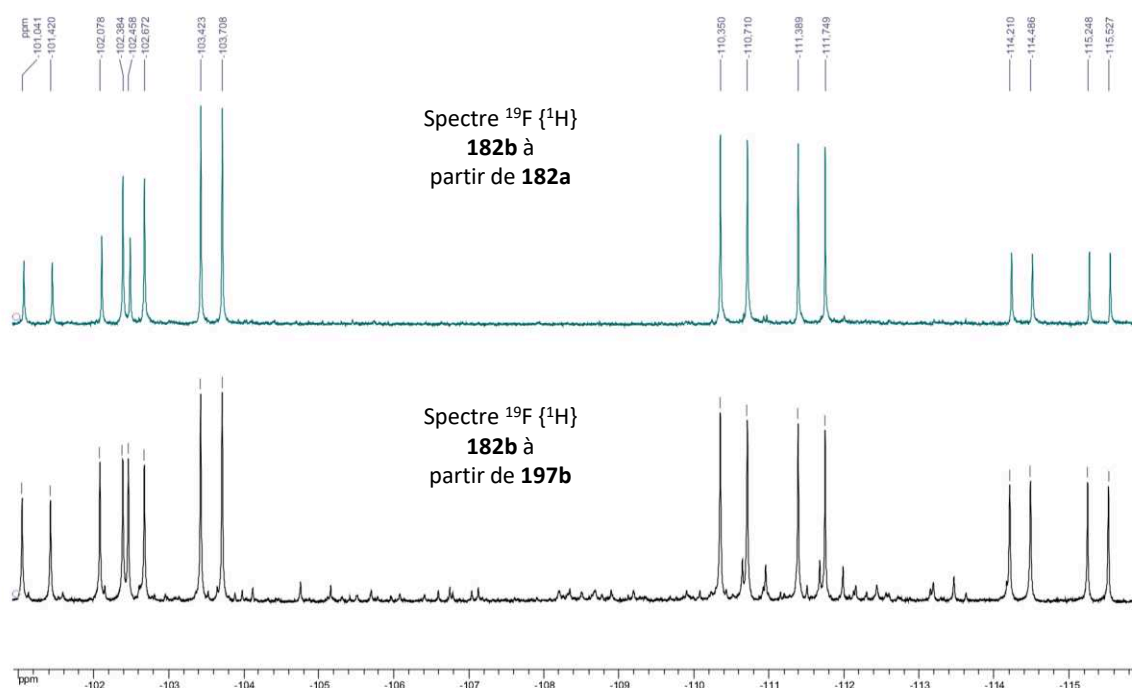


Figure 54 : Comparatif des spectres RMN du  $^{19}\text{F} \{^1\text{H}\}$  de **182b**

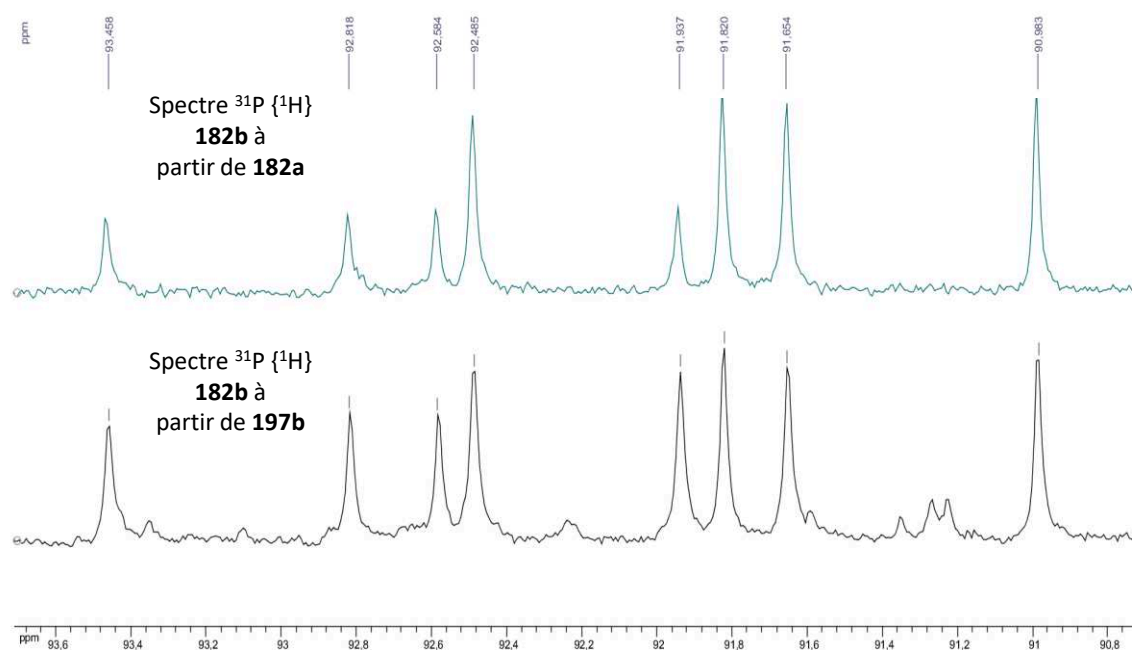
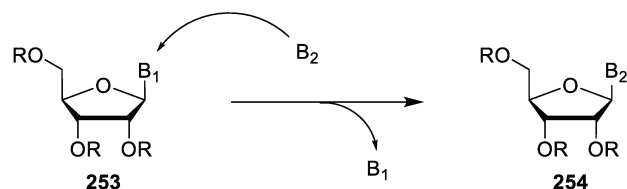


Figure 55 : Comparatif des spectres RMN du  $^{31}\text{P} \{^1\text{H}\}$  de **182b**

La substitution d'une base nucléique par une seconde, différente, constitue le phénomène de trans-*N*-glycosylation. Un tel processus n'est que peu évoqué dans la littérature récente, lors de transformations organochimiques, et constitue donc un problème synthétique supplémentaire qui se greffe sur notre approche. Avant de définir les solutions mises en œuvre pour éviter ou contourner le phénomène de trans-*N*-glycosylation, celui-ci est présenté ci-après au travers d'études précédentes décrites dans la littérature.

## II. Phénomène de trans-*N*-glycosylation

Dans le vivant, par utilisation d'enzymes (purine nucléoside phosphorylase, glycosyltransférase par exemple), chaque base nucléique peut être remplacée par une autre (Schéma 68).<sup>150</sup> Dans notre cas, nous allons nous intéresser plus spécifiquement à la substitution de nucléobase par voie chimique permettant notamment d'accéder à de nouveaux nucléosides à partir d'un substrat analogue.

Schéma 68 : Trans-*N*-glycosylation

### II.1. Etat de l'art

La trans-*N*-glycosylation par voie chimique a été développée en laboratoire à la fin des années 1960 par Shimizu et Miyaki.<sup>151</sup> C'est un processus dans lequel la partie furanose d'un nucléoside est transférée d'une base hétérocyclique à une autre permettant ainsi la formation d'un nouveau nucléoside. Dans les travaux de Miyaki, le transfert de la partie furanose du nucléoside **255**, portant une cytosine silylée, sur l'adénine **26** a été réalisé en présence stœchiométrique de HgBr<sub>2</sub> comme acide de Lewis. Dans ces conditions, l'adénosine **256**, résultant d'une réaction de trans-*N*-glycosylation, a été isolée avec un rendement de 44% (Schéma 69).<sup>152</sup> Dans cette étude, un dérivé de base purique, la théophylline **258**, a également substitué l'uracile du nucléoside **257**.

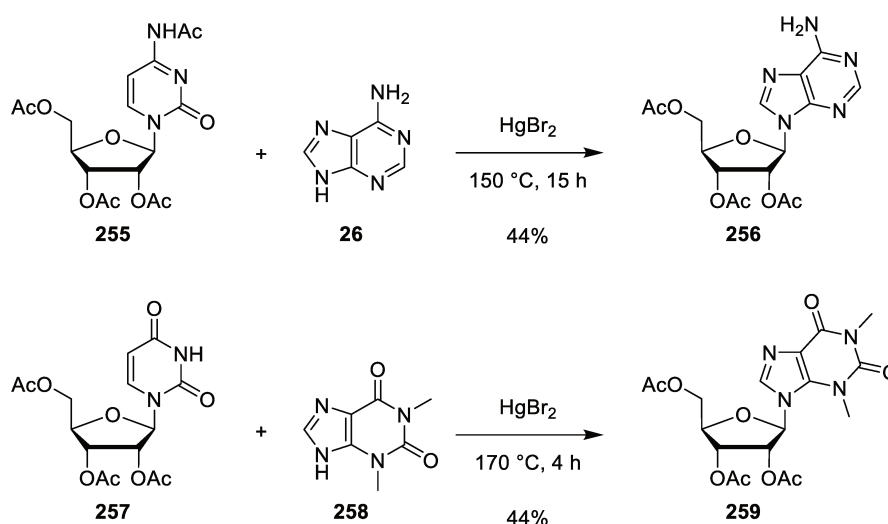


Schéma 69 : Travaux de Miyaki

<sup>150</sup> Mikhailopulo, I. A. *Curr. Org. Chem.* **2007**, *11*, 317-335.

<sup>151</sup> Shimizu, B. and Miyaki, M. *Tetrahedron Lett.* **1968**, *9*, 855-859.

<sup>152</sup> Miyaki, M.; Saito, A. and Shimizu, B. *Chem. Pharm. Bull.* **1970**, *18*, 2459-2468.

Plus récemment, dans une étude visant à synthétiser des nucléosides 2'-OME, une *trans*-*N*-glycosylation du furanose **260**, portant une cytosine acétylée, sur une adénine benzoylée **261** a été rapportée par Beigelman en utilisant les conditions classiques de Vorbrüggen (Schéma 70).<sup>153</sup>

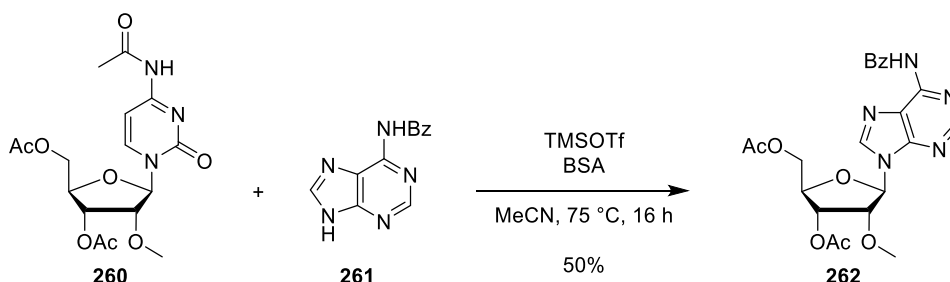


Schéma 70 : Travaux de Beigelman

De nombreux travaux utilisent ce processus afin de substituer principalement la thymine par une autre base nucléique. En effet, cette approche permet l'accès à des nucléosides modifiés portant des bases différentes à partir d'un même nucléoside. En guise d'exemple, Sawamoto a réalisé la substitution de la thymine d'un nucléoside LNA **263** par les trois autres bases nucléiques de l'ADN dans les conditions de Vorbrüggen (Schéma 71).<sup>154</sup> Ainsi les nucléosides adénine benzoylée **264**, 5-méthylcytosine **265** et guanine diprotégée **266** ont été obtenus respectivement avec des rendements de 83%, 78% et 71%.

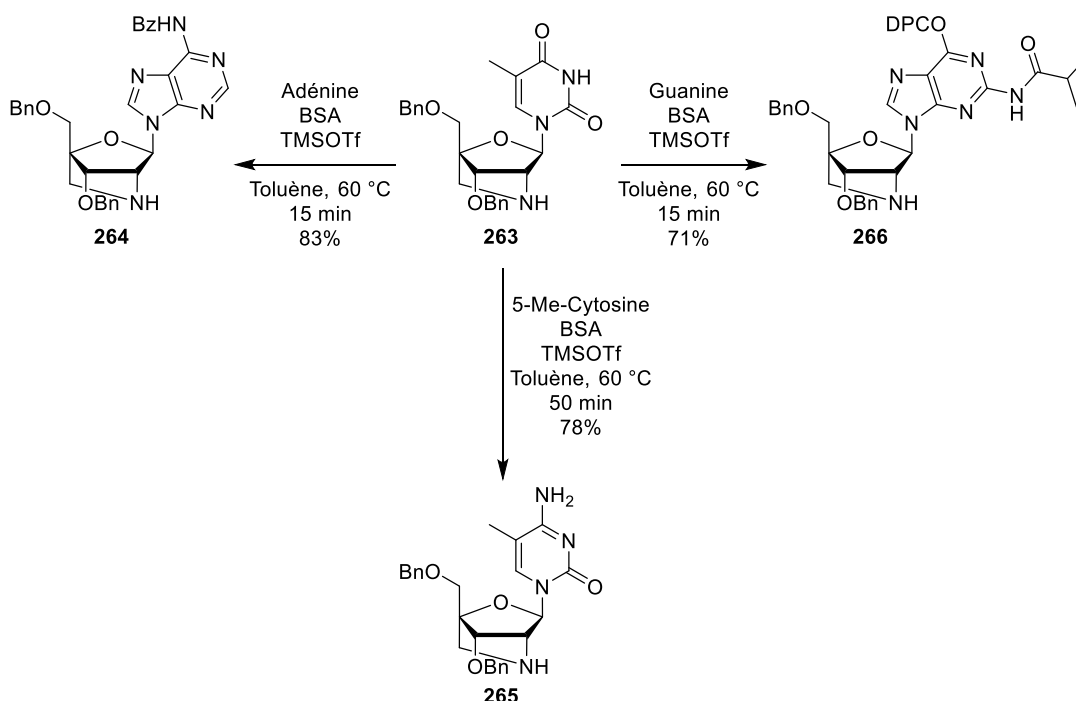


Schéma 71 : Travaux de Sawamoto

<sup>153</sup> Beigelman, L.; Haeberli, P.; Sweedler, D. and Karpeisky, A. *Tetrahedron* **2000**, *56*, 1047-1056.

<sup>154</sup> Sawamoto, H.; Arai, Y.; Yamakoshi, S.; Obika, S. and Kawanishi, E. *Org. Lett.* **2018**, *20*, 1928-1931.

La trans-*N*-glycosylation a été développée originellement comme la substitution d'une base pyrimidique par une base purique. Cependant, cette méthode a aussi été observée dans le sens inverse comme le montre les travaux de Suhadolnik portant sur le développement d'une méthode efficace afin d'obtenir un 3'-déoxynucléoside.<sup>155</sup> A partir de la 3-déoxyadénine bisbenzoylée **267**, une réaction de trans-*N*-glycosylation a été réalisée en présence de SnCl<sub>4</sub> (5 éq) et d'uracile silylé **225e** (Schéma 72). Ainsi, la 3'-déoxyuridine **268** a été obtenue avec un rendement de 61%.

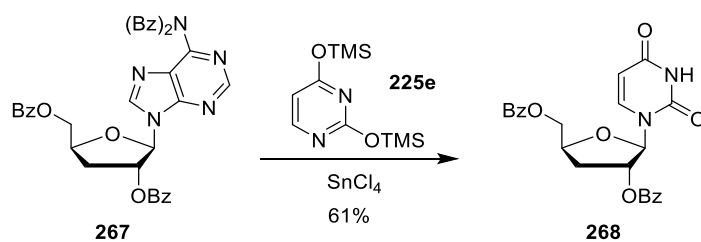


Schéma 72 : Travaux de Suhadolnik

Ces différentes études montrent qu'il est possible de moduler les caractères nucléofuge et nucléophile des différentes bases nucléiques afin que celles-ci soient substituées par une autre. En effet, la nucléophilie des nucléobases peut être augmentée par modification chimique : par exemple, la silylation de l'uracile pour mener à **225e**. L'insertion de groupement(s) électroattracteur(s) comme, par exemple, benzoyle ou acétate, permet d'augmenter le caractère nucléofuge de la base nucléique du nucléoside. Ainsi, en théorie, n'importe quelle base nucléique doit être capable d'en substituer une autre.

Les travaux de Boryski ont permis de mettre en évidence que la présence d'un cation oxonium est nécessaire afin d'initier la réaction d'échange de la base nucléique. De plus, une rétention de configuration est observée lors de la présence d'un groupement acétyle en C2'.<sup>156</sup> En se basant sur des résultats récents publiés par Boryski, un mécanisme plausible de la réaction de trans-*N*-glycosylation peut être proposé.<sup>157</sup> Après activation de la base en milieu acide (acide de Lewis), deux mécanismes d'élimination de cette dernière sont possibles (Schéma 73). La première élimination de B<sub>1</sub> (en bleu) conduit à la formation de l'oxocarbenium **271**. L'addition d'un nucléophile (B<sub>2</sub>) sur ce carbonyle activé permet la formation d'un nouveau nucléoside **273**. Dans ce cas, l'addition pourrait prendre place à la fois sur la face  $\alpha$  et sur la face  $\beta$ , et un mélange d'anomères devrait être obtenu. L'oxonium **271** peut également être stabilisé par attaque nucléophile du groupement acétyle en C2' pour mener à l'intermédiaire 1,2-acyloxonium **272**. Ce dernier peut également être formé par un mécanisme de type S<sub>N</sub>2 (en rouge), le groupement acétate en C2' substituant la base nucléique activée par l'acide de Lewis.

<sup>155</sup> Suhadolnik, R. J. and Uematsu, T. *Carbohydr. Res.* **1978**, *61*, 545-548.

<sup>156</sup> Boryski, J. *Nucleosides and Nucleotides* **1996**, *15*, 771-791.

<sup>157</sup> Boryski, J. *Curr. Org. Chem.* **2008**, *12*, 309-325.

Une seconde substitution nucléophile par une nucléobase  $B_2$  permet la formation du nucléoside **273** avec cette fois-ci une rétention globale de configuration. L'équilibre est déplacé vers la formation du nucléoside thermodynamiquement le plus stable.

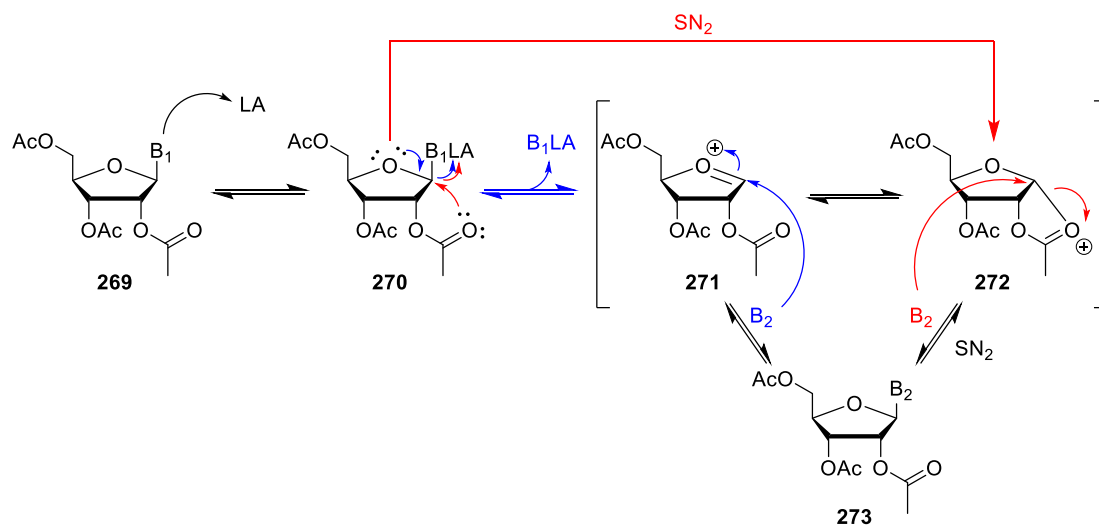


Schéma 73 : Mécanisme proposé de la trans-N-glycosylation

Le Schéma 74 explique ce phénomène sur des dinucléotides. Si l'on considère que la base nucléique  $B_2$  ne présente pas un caractère nucléophile suffisant pour substituer  $B_1$  (cas n° 1), l'oxonium cyclique **275a** ou l'acyloxonium **276a** est formé à partir de l'intermédiaire **274B<sub>1</sub>**. Ainsi, l'introduction d'une seconde base nucléique entraîne la formation de l'hétérodinucléotide **277B<sub>1</sub>-B<sub>2</sub>**, dans ce cas, le phénomène de trans-N-glycosylation n'a pas lieu. Dans le cas où  $B_2$  présenterait un caractère nucléophile suffisant pour substituer  $B_1$ , le cas n° 2 résulterait du processus de trans-N-glycosylation. L'oxonium cyclique **278a** ou **278b** est obtenu à la suite de l'élimination de  $B_1$  en présence d'un acide de Lewis. La nucléobase  $B_2$  prend donc place sur le furanose du bas pour donner **274B<sub>2</sub>**. Dans le milieu réactionnel se trouve donc **274B<sub>2</sub>**, l'acide de Lewis et la base nucléique  $B_1$  éliminée. L'intermédiaire cationique **275b** ou **276b** peut donc se former et la base nucléique  $B_1$  peut venir s'additionner pour donner l'hétérodinucléotide **277B<sub>2</sub>-B<sub>1</sub>**. En revanche, si la base nucléique  $B_2$  ne substitue que partiellement  $B_1$  (cas n° 3), alors un mélange de quatre espèces (**274B<sub>1</sub>**, **274B<sub>2</sub>**,  $B_1$  et  $B_2$ ) est présent dans le milieu réactionnel. La présence des deux bases nucléiques  $B_1$  et  $B_2$  dans le milieu réactionnel peut ainsi engendrer la formation de deux hétérodinucléotides (**277B<sub>1</sub>-B<sub>2</sub>** et **277B<sub>2</sub>-B<sub>1</sub>**) et deux homodinucléotides (**279B<sub>1</sub>-B<sub>1</sub>** et **279B<sub>2</sub>-B<sub>2</sub>**). Ainsi, lors de l'addition d'une seconde base nucléique sur un substrat triacétate de type **274**, le phénomène de trans-N-glycosylation pourrait être responsable de la formation de deux homodinucléotides et deux hétérodinucléotides.

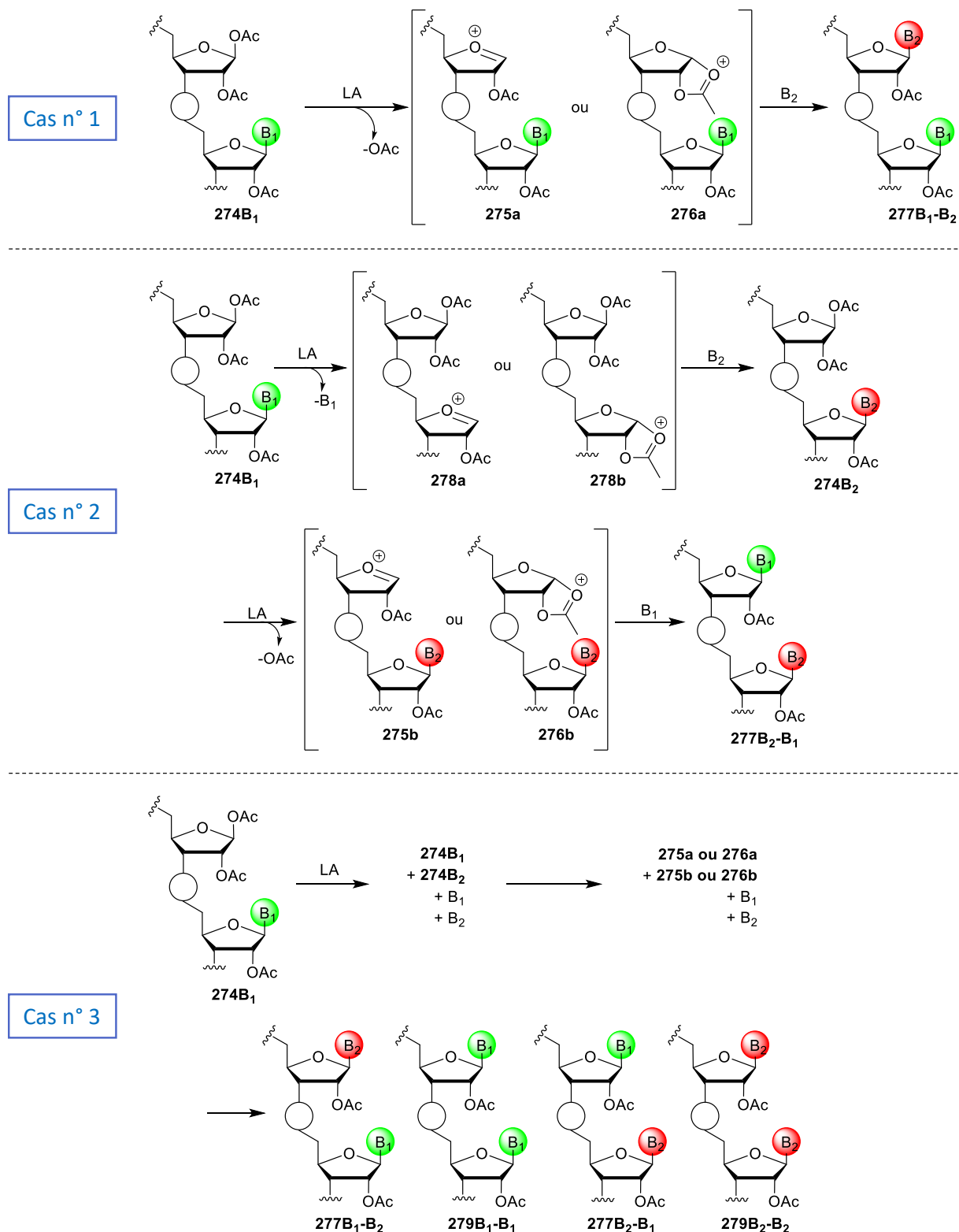


Schéma 74 : Présentation générique du processus de trans-N-glycosylation dans le cas des dinucléotides  
Le lien internucléosidique a été représenté sur la forme d'un cercle et les positions C3' et C5' n'ont pas été définies

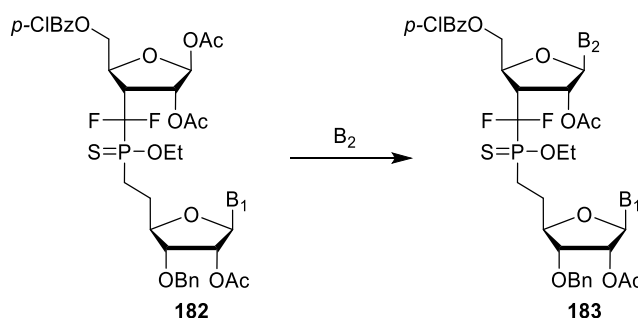
- Cas n° 1 : B<sub>2</sub> (1 équivalent) ne substitue pas B<sub>1</sub>
- Cas n° 2 : B<sub>2</sub> (1 équivalent) substitue totalement B<sub>1</sub>
- Cas n° 3 : B<sub>2</sub> (1 équivalent) substitue partiellement B<sub>1</sub>

Ces différents cas ont été observés lors de nos expériences. Au cours de l'introduction d'un uracile sur un substrat thymine<sup>-PMB</sup>-furanose, seule l'introduction de l'uracile sur le furanose du haut a été

observée ce qui correspond au cas n° 1. Le cas n° 2 a été obtenu lors de l'insertion d'une cytosine sur un substrat thymine-furanose. Le phénomène de trans-*N*-glycosylation a eu lieu et un substrat cytosine-furanose a été isolé. Dans ce cas, il est à noter que la thymine présente dans le milieu n'a pas réagi sur la position anomérique du furanose du haut pour donner un hétérodinucléotide C-T. Pour finir, l'introduction d'un uracile sur le même substrat thymine-furanose a entraîné l'obtention d'un mélange complexe défini comme étant la combinaison d'un hétérodinucléotide T-U et d'un homodinucléotide U-U. Cependant, le phénomène de trans-*N*-glycosylation remet en question l'hétérodinucléotide T-U défini comme tel dans un premier temps. En effet, ce dernier peut au regard du phénomène décrit ci-dessus, être soit un T-U, soit un U-T. Dans ce cas, on peut donc noter que le caractère nucléophile de l'uracile semble donc logiquement relativement proche de la thymine et entraîne la formation de plusieurs composés comme illustré par le cas n° 3.

Les travaux présents dans la littérature n'ont été réalisés que sur des nucléosides mais en se basant sur ces études et, vu les résultats de nos expériences, ces données sont potentiellement extrapolables aux dinucléotides. Ainsi, un tableau a été établi afin de déterminer les combinaisons possibles lors d'une stratégie d'introduction successive de bases nucléiques après la formation du lien internucléosidique (Tableau 15). La première lettre de chaque combinaison correspond à la base nucléique portée par le furanose du bas et la seconde au furanose du haut.

Tableau 15 : Combinaisons possibles de dinucléotide B<sub>1</sub>-B<sub>2</sub>



B <sub>1</sub> nucléobase en place	B <sub>2</sub> nucléobase à introduire				
	Thymine	Cytosine	Uracile	Adénine	Guanine
Thymine	TT	TC	TU	TA	TG
Cytosine	CT	CC	CU	CA	CG
Uracile	UT	UC	UU	UA	UG
Adénine	AT	AC	AU	AA	AG
Guanine	GT	GC	GU	GA	GG

En se basant sur nos expériences et la littérature,<sup>154</sup> dans les conditions de réactions utilisées (base silylée + TMSOTf), l'introduction d'une seconde base nucléique différente semble compromise

sur notre substrat thymine-furanose **182a** puisque la thymine peut être substituée par les trois autres bases présentes dans l'ADN. Ce phénomène pourrait également être rencontré lors de l'insertion d'une nucléobase sur les substrats cytosine-furanose **197b** et adénine-furanose **197c**, précurseurs des triacétates **182b** et **182c**. Afin d'éviter ce problème, deux solutions ont été envisagées.

## II.2. Solutions envisagées

### II.2.a. Addition contrôlée

Afin d'utiliser le processus de trans-*N*-glycosylation à notre avantage et en s'inspirant des travaux de Sawamoto (Chapitre 4, II.1), l'introduction contrôlée d'une seconde base nucléique (cytosine) sur un homodinucléotide T-T **180a** a été envisagée. Cependant, il est difficile d'anticiper quelle thymine sera substituée par la cytosine puisque chaque furanose possède un -OAc en position C2' permettant de stabiliser le cation oxocarbénium. De plus, le phosphinothioate peut également participer à la stabilisation de l'oxonium formé comme l'a montré une étude de Yokomatsu sur des phosphonothioates analogues.<sup>112</sup> Deux possibilités de mécanisme sont donc à envisager. La première est la formation du cation **280** sur le furanose du haut engendrant ainsi l'insertion de la cytosine sur celui-ci et la formation de **183c**. La seconde correspond à la formation du cation **281** sur le furanose du bas, stabilisé par l'assistance anchimérique du soufre générant un cycle à 7 chaînons, puis à l'insertion de la cytosine sur ce dernier conduisant à **183d**, avec l'obtention potentiel d'un mélange  $\alpha/\beta$  non-contrôlable (Schéma 75).

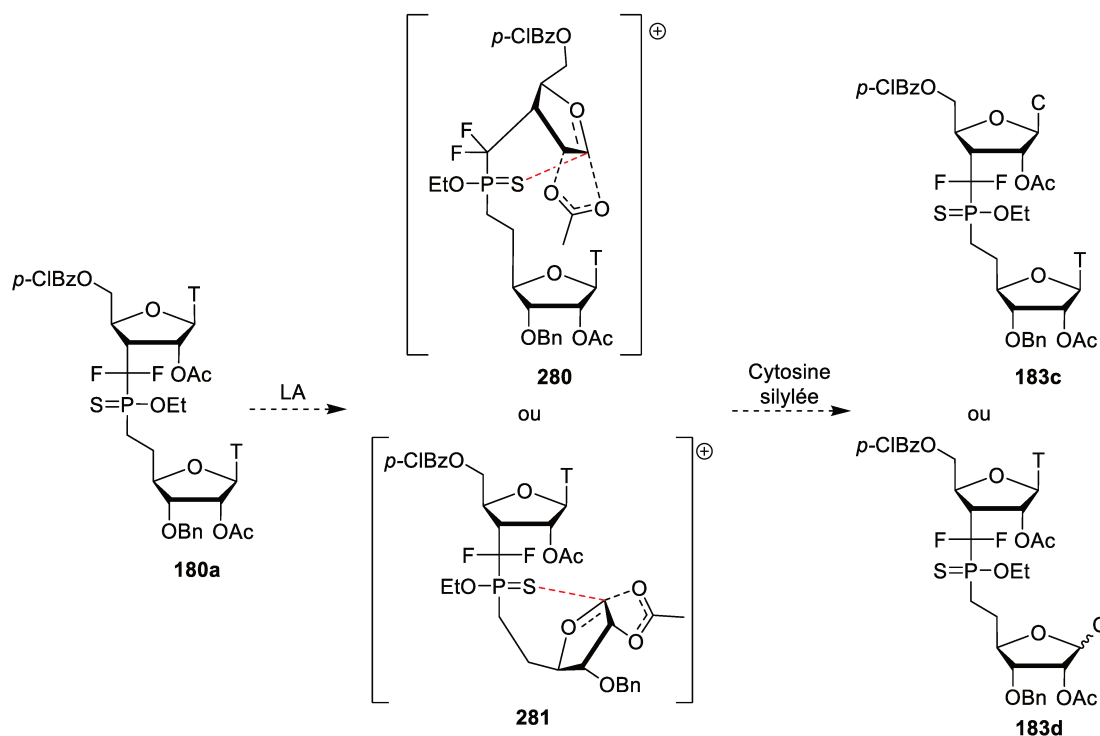
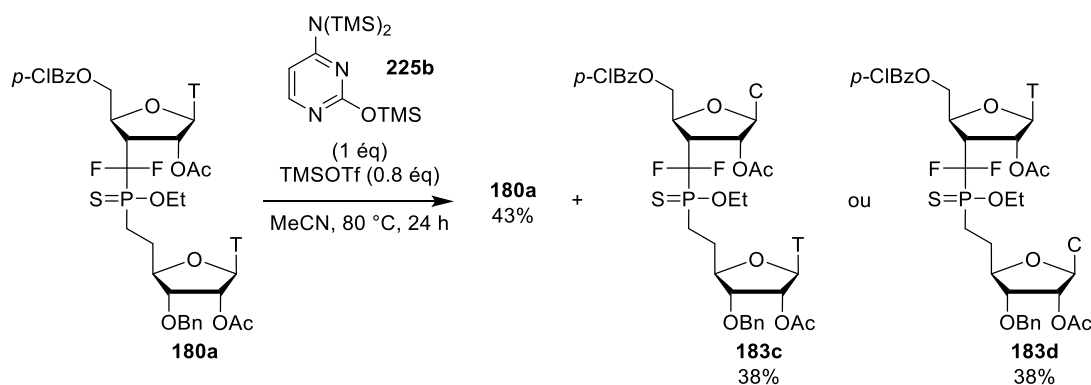


Schéma 75 : Mécanismes possibles pour l'addition contrôlée de la cytosine  
(note : les liaisons en pointillé noir peuvent en réalité être des liaisons  $\sigma$  ou  $\pi$ )

Ainsi, le dinucléotide **180a** a été mis en présence d'un équivalent de cytosine silylée **225b** et de TMSOTf pendant 24 h, à 80 °C (Schéma 76). Après purification, deux dinucléotides ont été séparés. Tout d'abord, le premier, correspondant à l'homodinucléotide T-T de départ **180a**, a été isolé à hauteur de 43%. Le second, correspondant à un dinucléotide porteur d'une thymine et d'une cytosine, a été isolé avec un rendement de 38% et confirme que le processus de trans-*N*-glycosylation a pris place dans cette réaction. Le problème suivant est de déterminer sur quel furanose ce processus s'est déroulé, et donc d'établir la structure du dinucléotide forme T-C (**183c**) ou C-T (**183d**).

Schéma 76 : Résolution cinétique à partir du dinucléotide T-T **180a**

La spectrométrie de masse haute résolution confirme la nature hétérodinucléotidique du produit. Cependant, la réaction ayant été réalisée sur petite échelle, sa structure exacte n'a pas pu être déterminée, à ce stade, par les études par spectrométrie de RMN. L'hétérodinucléotide C-T **183d** ayant été synthétisé par la suite et sa structure ayant été déterminée de manière inambigüe (voir II.2.b.i), une comparaison des spectres de RMN a pu être réalisée afin de déterminer la position de la thymine. Seule l'analyse RMN du <sup>19</sup>F découplé du <sup>1</sup>H a permis une comparaison claire avec **183d** (Figure 56). Ainsi, les signaux montrent que les deux composés seraient différents et donc que le composé **183** formé lors de la réaction de trans-*N*-glycosylation sur l'homodinucléotide T-T **180a** serait l'hétérodinucléotide T-C **183c**.

Une préférence cinétique pour le furanose du bas avait été observée lors de l'introduction d'une seule thymine sur un composé tétraacétate lors des travaux du Dr. Jun Zhang (Introduction, V.2). Dans notre cas, la cytosine a été introduite sur le furanose du haut. La différence entre ces deux expériences est la présence d'une thymine en position anomérique sur le furanose du haut à la place du -OAc. Ainsi, dans les conditions réactionnelles, la thymine activée par un acide de Lewis serait un meilleur nucléofuge que -OAc. Ceci expliquerait le fait qu'une base nucléique ait été introduite sur le furanose du haut dans notre cas et pas lors de l'expérience du Dr. Jun Zhang.

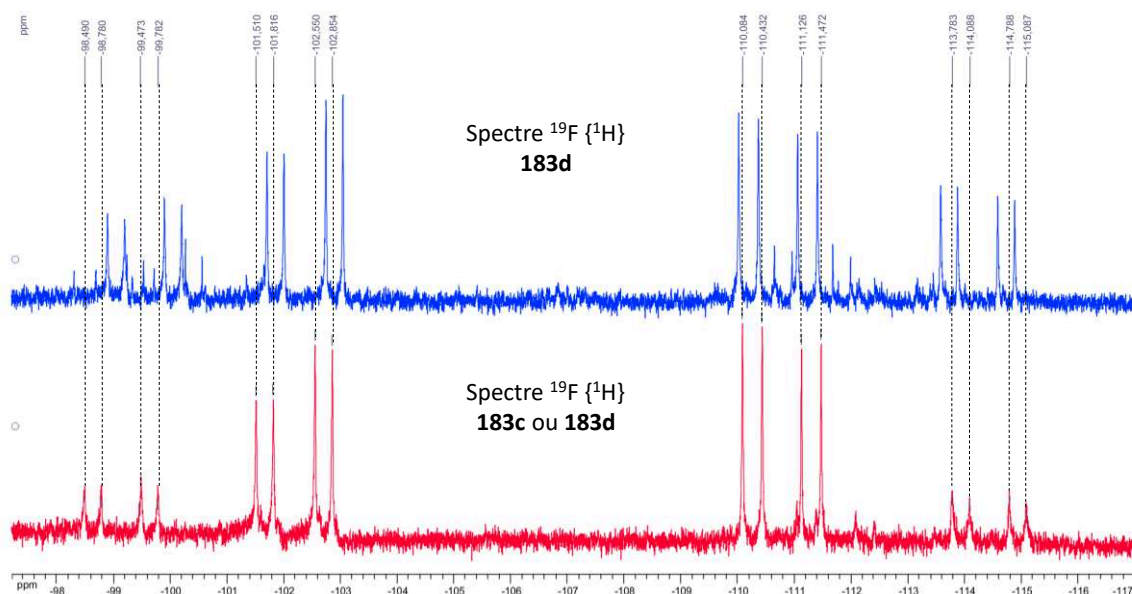


Figure 56 : Comparaison entre le spectre  $^{19}\text{F}$   $\{^1\text{H}\}$  lors de l'insertion de la cytosine sur T-T **180a** et C-T **183d**

Une solution alternative susceptible de contourner le phénomène de trans-*N*-glycosylation a également été explorée.

### II.2.b. Ordre d'insertion

Cette stratégie est basée sur l'application des données de la littérature sur des nucléosides. Grâce à ces résultats, une tentative d'établir un ordre d'introduction des bases nucléiques a été définie.

#### II.2.b.i. *Utilisation d'un intermédiaire cytosine*

En s'appuyant sur les travaux de Sawamoto, la thymine peut être substituée par les autres bases nucléiques dans des conditions similaires (base silylée, TMSOTf). Ainsi, un intermédiaire nucléotide-furanose **182** portant une autre base que la thymine pourrait être utilisée afin d'introduire une seconde base nucléique différente (Schéma 77).

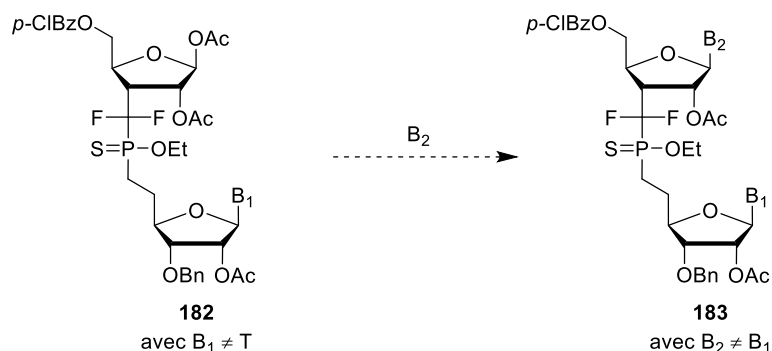


Schéma 77 : Stratégie ordre d'insertion

Dans un premier temps, une stratégie menant à un hétérodinucléotide cytosine-thymine **183d** a été envisagée à partir de l'intermédiaire mixte cytosine-furanose **197b** (Schéma 78). Tout d'abord, l'acétonide a été clivé en milieu acide puis les alcools C1'-C2' résultants ont été protégés sous forme d'acétates. L'intermédiaire triacétate **182b** a été obtenu avec un faible rendement de 25%. Lors de cette étape, un sous-produit a notamment été détecté en spectrométrie de RMN du  $^{19}\text{F}$  découplé du proton dont les déplacements chimiques sont les suivants : -96.0 ppm (dd, 1F,  $J = 288.6$  Hz, 72.6 Hz), -111.2 ppm (dd, 1F,  $J = 288.6$  Hz, 85.5 Hz). De plus, une analyse HRMS de la fraction contenant ce dernier après purification a été réalisé afin de déterminer sa masse moléculaire (HRMS (ESI-) : 922.1880). Une différence de 37 est ainsi obtenue avec **182b**. Celle-ci ne correspond ni à l'acétylation de la base (ajout d'un acétate : +43), ni à la substitution de cette dernière par un -OAc (MW tétraacétate : 834). A l'heure actuelle, la structure de ce sous-produit n'a pu être déterminée.

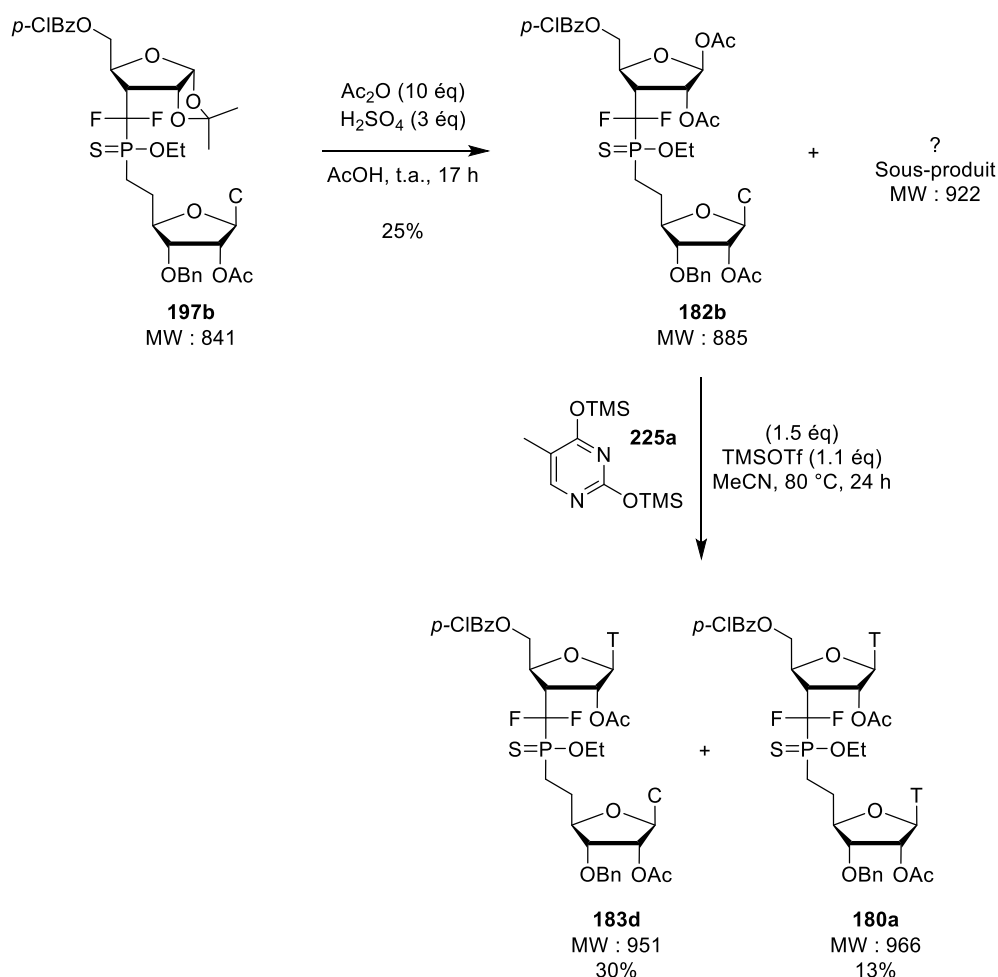


Schéma 78 : Stratégie de synthèse du dinucléotide cytosine-thymine **183d**

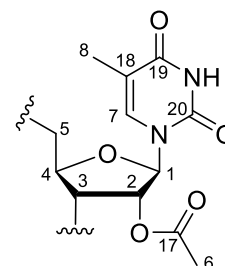
Il a été montré, lors d'un suivi réactionnel par RMN du  $^{19}\text{F}$  découplé du proton, que le sous-produit se forme essentiellement lors du traitement basique effectué (solution aqueuse de bicarbonate de saturée). D'autres traitements ont ainsi été réalisés :

- Evaporation du mélange brut réactionnel puis coévaporation au toluène pour éliminer les dernières traces d'acide acétique.
- Utilisation de  $\text{NaHCO}_3$  solide, filtration puis coévaporation au toluène.
- Lavage à l'eau distillée.
- Lavage avec un tampon à  $\text{pH} = 7.33$ .

Cependant, aucun de ces traitements n'a permis d'éliminer ce sous-produit. L'introduction d'une thymine a donc été réalisée sur l'intermédiaire cytosine **182b**. Une diminution du nombre d'équivalent de la thymine silylée a été effectuée (1.5 éq au lieu de 3 éq) afin de potentiellement limiter le phénomène de trans-*N*-glycosylation. Ainsi, l'hétérodinucléotide C-T **183d** a été obtenu avec un rendement de 30%. Il est à noter que dans ce cas, l'absence d'un groupement électroattracteur sur la cytosine n'a pas empêché la réaction de trans-*N*-glycosylation et le dinucléotide T-T **180a** a également été isolé à hauteur de 13%. Un rendement de 44% a été obtenu lors de l'introduction d'une seconde thymine à partir de **197a** pour former l'homodinucléotide T-T où le processus de trans-*N*-glycosylation n'est pas observable. Dans le cas de l'introduction de la thymine sur l'intermédiaire cytosine **182b** et en prenant en compte le possible phénomène de trans-*N*-glycosylation observé, un rendement global en hétérodinucléotides quasi similaire a donc été obtenu. Dans ce cas, la réaction de trans-*N*-glycosylation semble avoir été limitée mais sans être supprimée pour autant.

#### Détermination de la position anomérique de la thymine et de la cytosine :

Il est à noter que cette étude par spectrométrie de RMN 1D et 2D a été réalisée sur un mélange de diastéréoisomères (ratio 2/1, dû à la chiralité du phosphore) (Figure 58, Figure 59, Figure 60, Figure 61, Figure 62 et Figure 63). A l'aide des signaux caractéristiques de la thymine, le proton anomérique H1 a pu être déterminé. Corrélant avec un méthyle (H8) et deux carbonyles (C19 et C20) en analyse de RMN COSY et HMBC respectivement, le signal à 6.90 ppm sur le spectre de RMN du  $^1\text{H}$  correspond à H7. L'analyse HSQC permet d'attribuer son signal (138.0 ppm) sur le spectre de RMN du  $^{13}\text{C}$ . H7 corrèle également avec un signal à 94.0 ppm en HMBC et à 5.15 ppm en NOESY. L'analyse HMBC révèle que ces deux signaux proviennent de la même position. Etant donné que les signaux de la thymine ont déjà été attribués, cette position est donc définie comme le proton anomérique H1. En analyse HMBC, ce dernier révèle trois tâches de corrélation à 75.4 ppm, à 180.0 ppm (C7) et à 149.7 ppm (C20, vu également par H7). A l'aide de l'HSQC, le CH possédant le signal à 75.4 ppm, sur le spectre RMN du  $^{13}\text{C}$ ,



peut être placé sur le spectre RMN du  $^1\text{H}$  (5.84 ppm). Ce signal corrèle en analyse HMBC avec un signal à 170.2 ppm correspondant à la zone des carbonyles. Ainsi, ce CH est défini comme étant la position 2 du furanose. Grâce à cette dernière et l'analyse COSY, la position du proton H3 peut être définie à 4.00 ppm. Ce dernier corrèle également, en analyse COSY, avec le signal à 4.82-4.77 ppm pouvant correspondre à H4. Le spectre RMN du  $^1\text{H}$  révèle la présence de deux atomes d'hydrogène sous ce massif. Cette analyse est corroborée par l'analyse HSQC qui montre la présence d'un CH (C4), ayant une correspondance sous le pic de chloroforme deutéré ( $\approx 77$  ppm), et d'un  $\text{CH}_2$ , réparti sous ce massif et sur un second à 4.60-3.92 ppm et correspondant au signal à 64.4 ppm sur le spectre  $^{13}\text{C}$ . L'analyse HMBC révèle une corrélation entre H4-H5 et trois signaux. Le premier représente un massif à 44 ppm sur le spectre  $^{13}\text{C}$  et ne montre pas de correspondance en analyse HSQC. Le deuxième est la corrélation de H5 avec le signal de C4 à 77 ppm (également observée avec le massif à 4.60-3.92 ppm comportant le second H5). La troisième tâche de corrélation se situe, elle, à 166 ppm signifiant la présence d'un carbonyle dans le voisinage de H4 ou H5. Seul le H5 du furanose du haut présente un carbonyle proche (C14). Cela permet d'affirmer que la thymine a été introduite sur ce furanose et, ainsi, que la cytosine se situe sur le furanose du bas. En se basant sur les analyses des différentes molécules synthétisées, le massif à 44 ppm sur le spectre  $^{13}\text{C}$  peut être attribué à C3. Son absence de corrélation en analyse HSQC peut être expliquée par la multiplicité engendrée par la présence du  $\text{CF}_2$  entraînant l'augmentation du temps de relaxation de cette position.

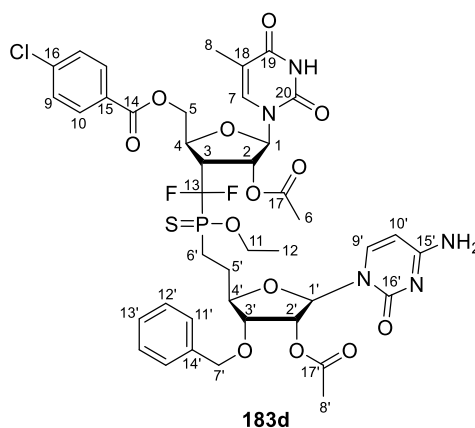


Figure 57 : Hétérodinucléotide cytosine-thymine synthétisé

Chapitre 4 - Insertion d'une seconde base nucléique

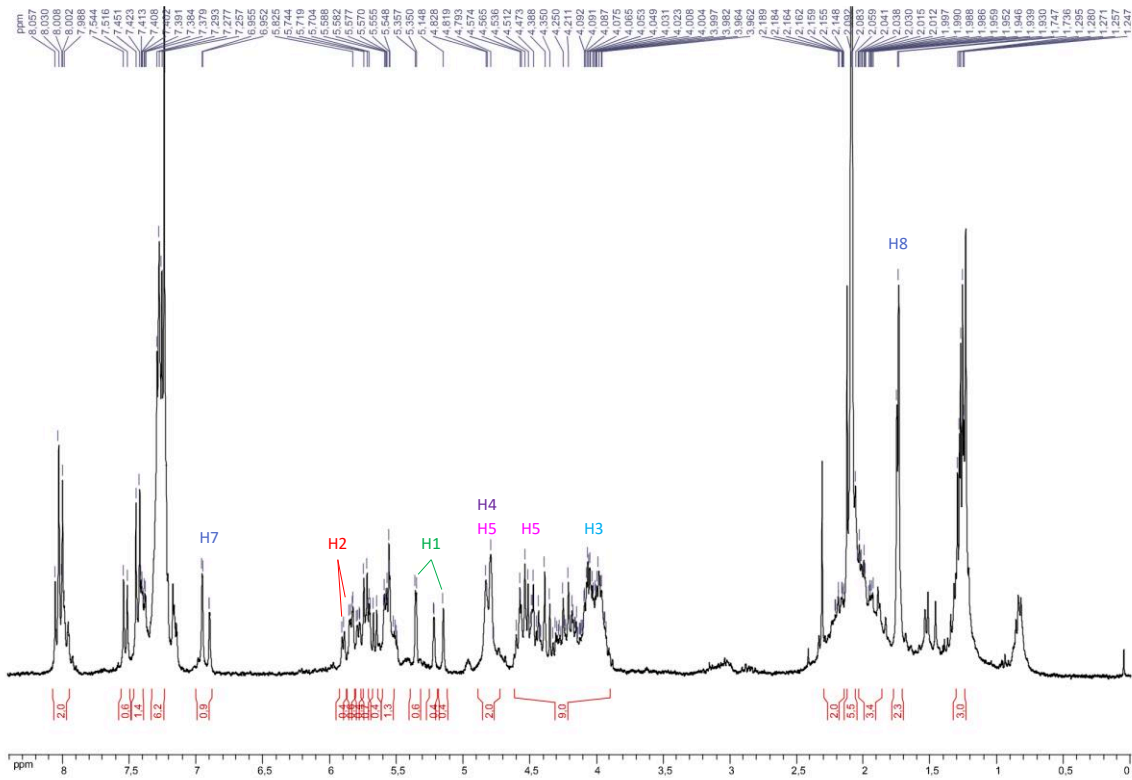


Figure 58 : Spectre RMN du  $^1\text{H}$  de **183d**

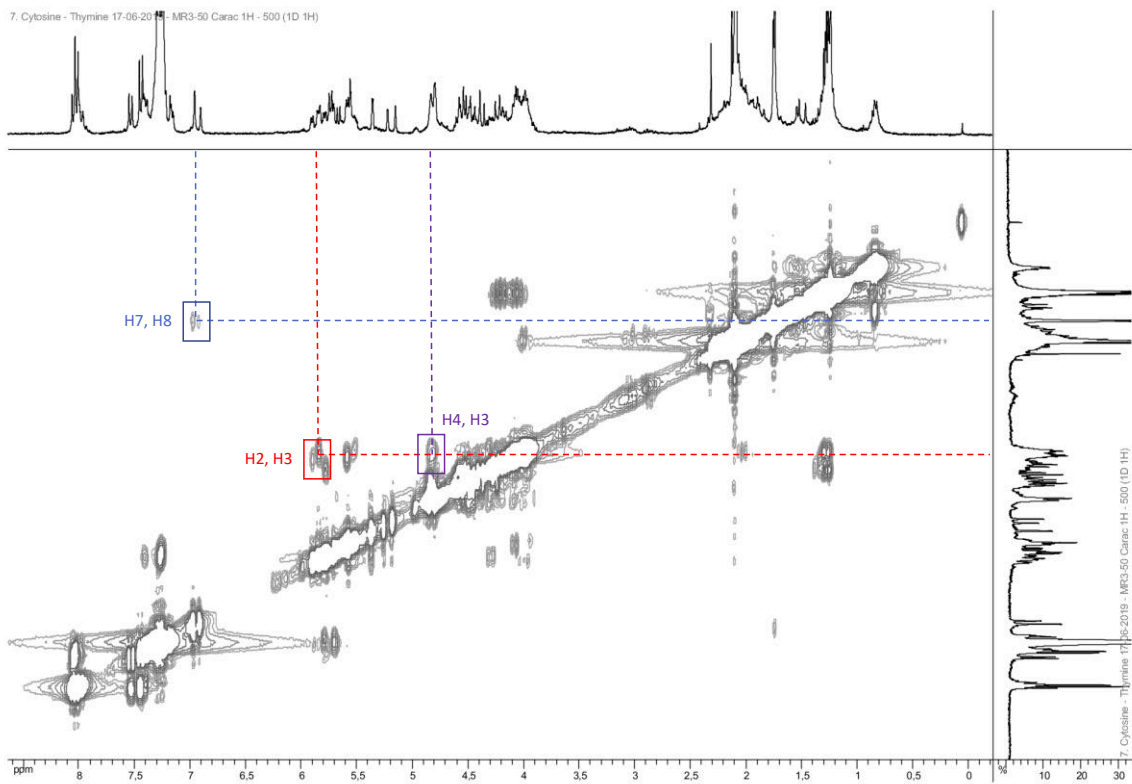


Figure 59 : Spectre RMN COSY de **183d**



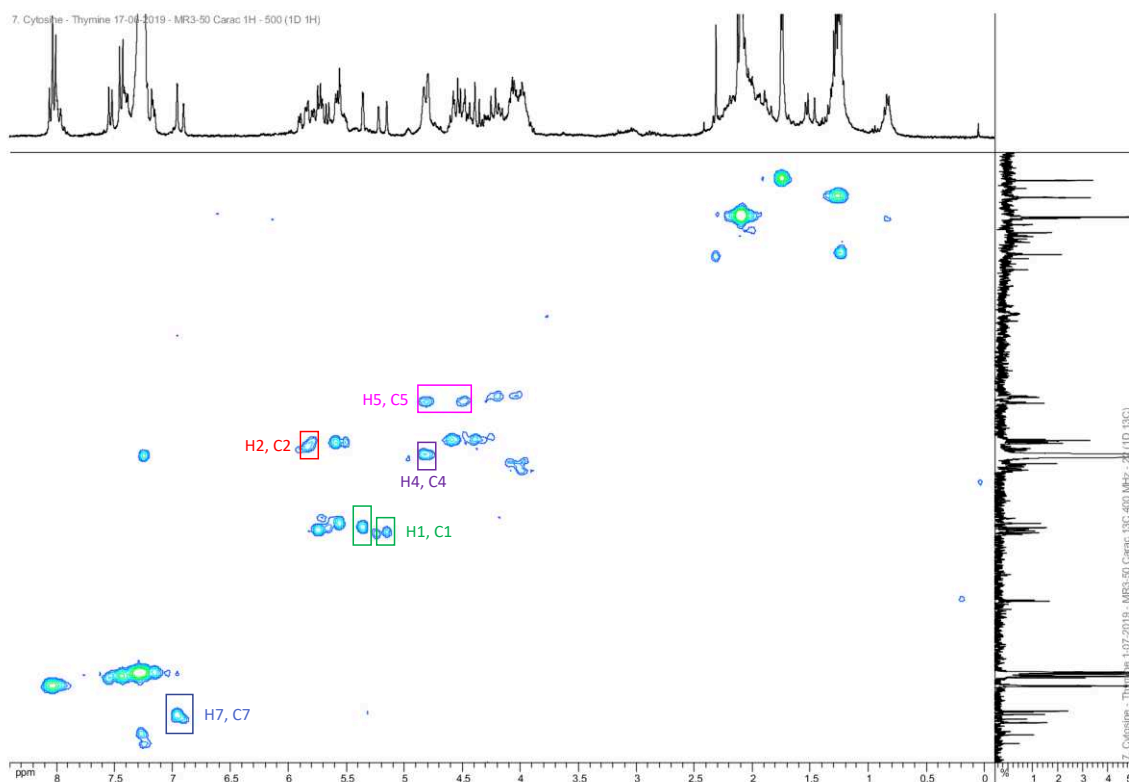


Figure 62 : Analyse HSQC de 183d

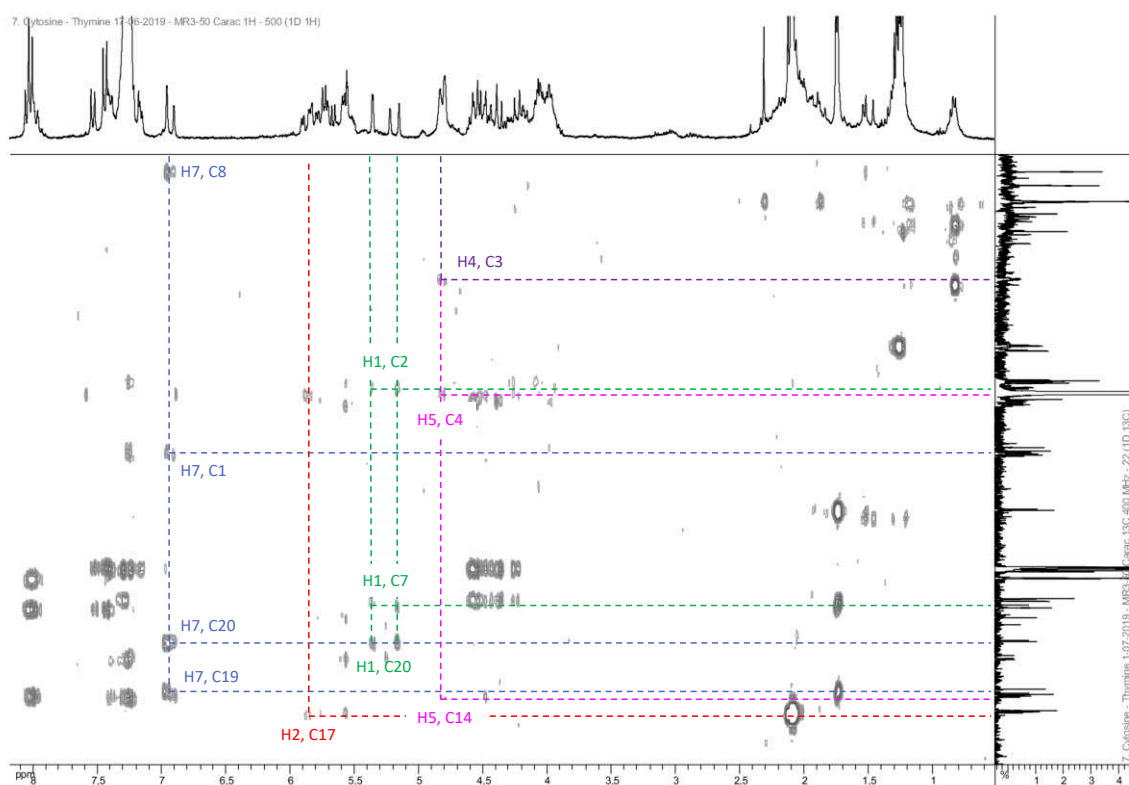


Figure 63 : Analyse HMBC de 183d

Un essai d'introduction de la guanine protégée **225f** a été réalisée sur l'intermédiaire **182b** (Schéma 79). Deux fractions ont été isolées après chromatographie en phase normale. Dans ces fractions, des traces du substrat **182b** (ayant une masse de 885) et de l'hétérodinucléotide **183e** ou **183f** (ayant une masse de 1213) ont été détectées en UPLC-MS (Ultra Performance Liquid Chromatography - Mass Spectrometry). La présence de l'homodinucléotide **180c** (ayant une masse de 1490) n'a pas pu être observée en analyse UPLC-MS à cause de la limite du spectromètre de masse à 1250. La détermination de la présence de **180c** n'a pu être réalisée à l'aide du spectromètre de masse de haute résolution à cause d'un problème technique. Il est cependant à noter que l'intermédiaire mixte guanine protégée-furanose **182d** (ayant une masse de 1162) n'a pas été détecté en UPLC-MS. A l'heure actuelle, cet essai a permis d'observer le produit **183e** ou **183f** à l'état de traces. Cependant, d'après les analyses RMN, un sous-produit (**180c** ?) est également présent.

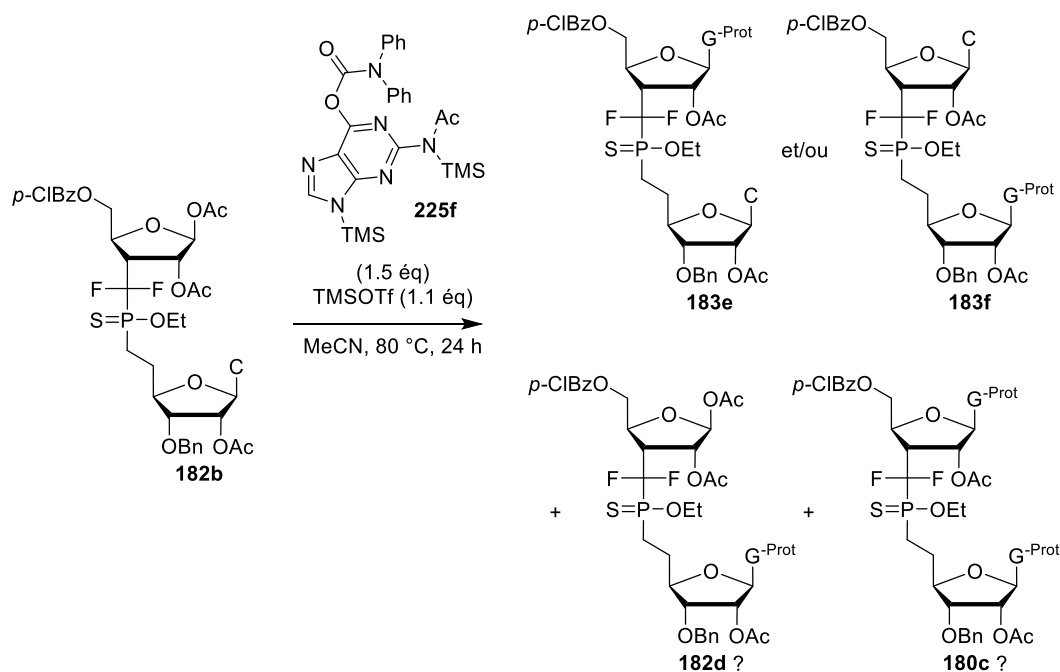
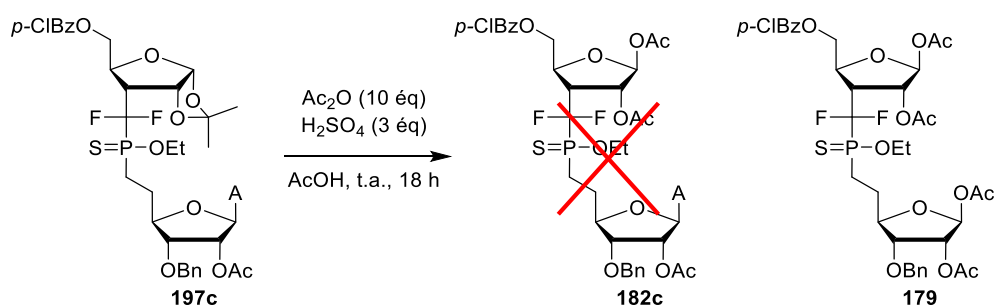


Schéma 79 : Introduction de la guanine protégée **225f** sur l'intermédiaire nucléotide-bisacétate **182b**

## II.2.b.ii. Utilisation d'un intermédiaire adénine

Une cible hétérodinucléotidique supplémentaire a été envisagée à partir d'un intermédiaire **197c** portant une adénine (Schéma 80). Lors de l'étape d'acétylolyse de l'acétonide, la perte de l'adénine a été observée par spectrométrie de RMN du  $^1\text{H}$  dans le mélange brut réactionnel. Les signaux observés en spectrométrie de RMN du  $^{19}\text{F}$  correspondent à ceux de l'intermédiaire tétraacétate **179** déjà synthétisé au laboratoire. Le composé adénine-furanose **197c** a donc totalement été converti en tétraacétate **179**. Cette expérience montre néanmoins que, sous certaines conditions, l'adénine peut être un bon nucléofuge et favoriser donc le phénomène de trans-*N*-glycosylation. D'autres conditions de clivage de l'acétonide pourraient être testées comme, par exemple, une diminution de la quantité d'acide sulfurique (inférieure à 3 éq) ou l'utilisation catalytique d'acide camphorsulfonique à la place de l'acide sulfurique. Cependant, par manque de temps, ces essais n'ont pas été entrepris.

Schéma 80 : Voie synthétique développée à partir de l'intermédiaire **197c**

### III. Conclusions

Lors du développement de la synthèse d'hétérodinucléotides, de nombreux problèmes ont été rencontrés, parmi lesquels le phénomène de *trans-N*-glycosylation. Les expériences décrites ci-avant ont permis de mettre ce problème en lumière, qui peut être rencontré dans toutes les approches synthétiques mettant en jeu l'insertion des bases nucléiques après la construction du lien internucléosidique. Des solutions ont été testées dans le but d'éviter ce phénomène. Ainsi, grâce à une stratégie basée sur l'ordre d'introduction des bases nucléiques, un hétérodinucléotide C-T **183d** a pu être synthétisé (Figure 64). D'autres voies de synthèse ont permis l'obtention d'un autre hétérodinucléotide T<sup>PMB</sup>-U **183a** et de deux homodinucléotides T-T **180a** et U-U **180b**.

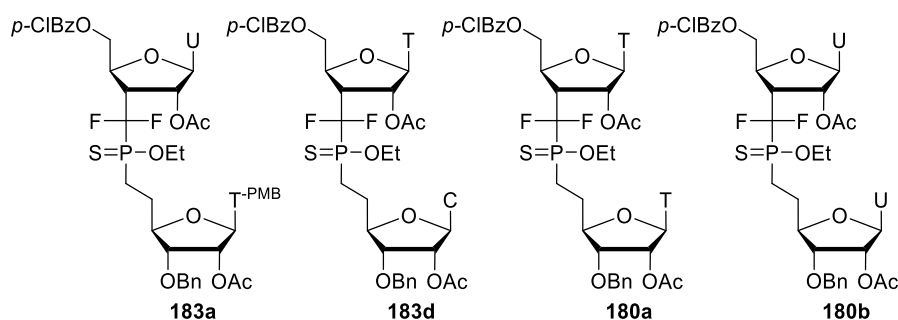


Figure 64 : Dinucléotides synthétisés

En suivant la méthode développée par les Drs. Jun Zhang et Emilie Lambert, les cinq homodinucléotides peuvent être obtenus. Dans le cas de la synthèse d'hétérodinucléotides, en utilisant les conditions bases silylées-TMSOTf, l'obtention de dinucléotide avec la thymine inchangée sur le furanose du bas semble compromise à cause du phénomène de *trans-N*-glycosylation. Cependant, il est à noter que lors de l'utilisation de la thymine PMB-protégée, ce processus n'a pas été observé et l'introduction d'un uracile a été réalisée avec succès. Le groupement PMB pourrait ralentir l'activation de la base nucléique par un acide de Lewis et donc éviter le départ de cette dernière. Si des conditions de déprotections efficaces du PMB peuvent être mises au point, cette stratégie pourrait être utilisée afin de s'affranchir des réactions de *trans-N*-glycosylation, au moins dans ce cas. Une autre solution visant à monosilyler la thymine portée par le furanose du bas pourrait potentiellement réduire le caractère nucléofuge de la base et limiter le processus de *trans-N*-glycosylation. Lors de l'utilisation d'un substrat ayant la cytosine déjà en place, l'hétérodinucléotide C-T a été obtenu majoritairement. Néanmoins, le phénomène de *trans-N*-glycosylation a tout de même été observé. D'autres essais de déprotection de l'acétonide doivent être entrepris avant de conclure sur la faisabilité d'hétérodinucléotides à partir du substrat adénine-furanose **197c**.



## **Conclusions et perspectives**



## I. Conclusions

L'objectif principal de ces travaux était de développer une stratégie de seconde génération conduisant à la formation d'hétérodinucléotides caractérisés par un motif difluorophosphinothioate en lieu et place du pont phosphodiester naturel. Pour cela, trois stratégies basées sur la construction successive de deux liaisons phosphore-carbone ont été étudiées (Figure 65).

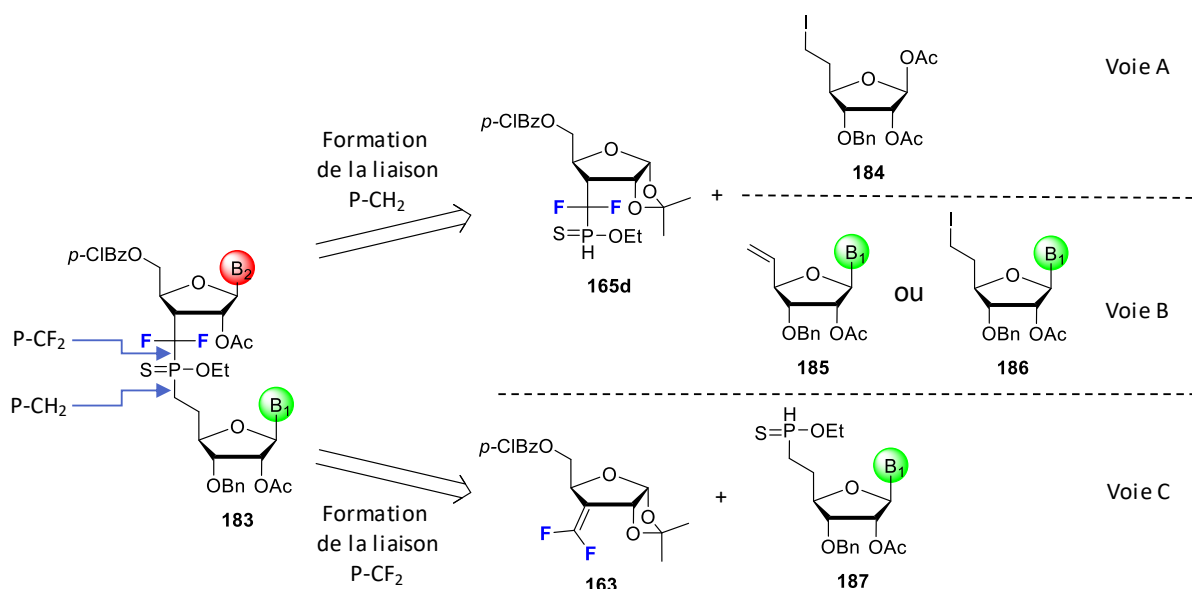


Figure 65 : Stratégies étudiées

Dans un premier temps, plusieurs fragments ont été synthétisés à partir d'un substrat commun, à savoir l'allofuranose **221** (Schéma 81). Ceux-ci ont été utilisés dans les stratégies A et B afin d'étudier la construction de la seconde liaison phosphore-carbone. Dans cette étude, le substrat **184** a permis l'accès à quatre nucléosides (**186a-d**) possédant des bases nucléiques différentes.

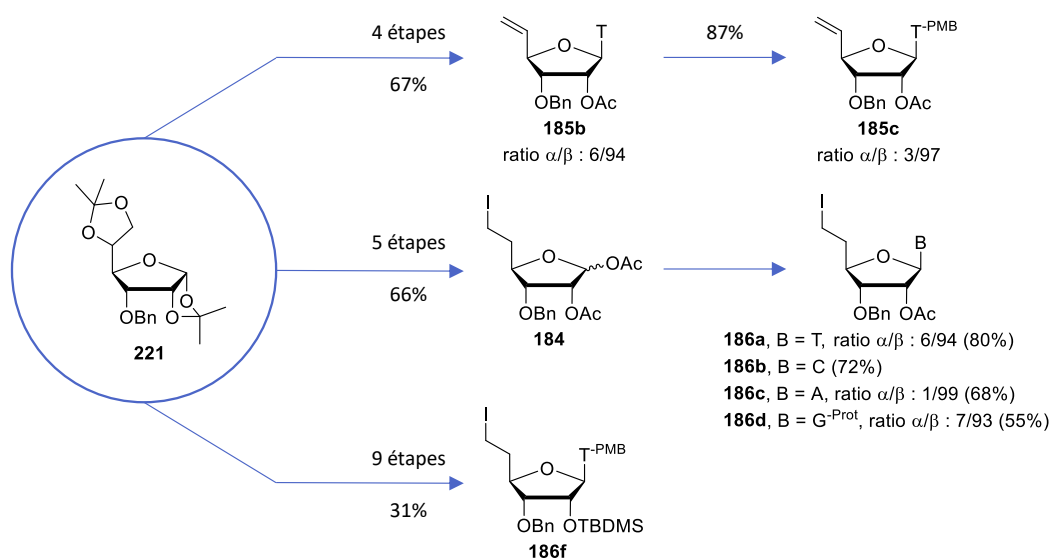


Schéma 81 : Fragments obtenus à partir de l'allofuranose **221**

Une première stratégie (voie A) visant à introduire une première base nucléique après la formation du lien difluorophosphinothioate (formation de la liaison P-CH<sub>2</sub>) et reposant sur la différenciation des furanoses a été étudiée. Néanmoins, le manque d'efficacité lors des essais de couplage entre le *H*-phosphinothioate **165d** et le fragment **184** nous a conduit à envisager la deuxième stratégie. Au contraire des travaux de Lebreton, cette voie synthétique A ne s'avère pas efficace dans notre cas pour l'instant.

Une seconde stratégie (voie C), réalisée sur un substrat modèle, menant à la formation d'une seconde liaison phosphore-carbone (P-CF<sub>2</sub>) par addition radicalaire d'un *H*-phosphinothioate **244b** sur une difluorooléfine modèle **246** a été étudiée. Lors du développement de cette voie synthétique, une étude comparative de la réactivité entre H-P=O **244a** et H-P=S **244b** a pu être réalisée (Schéma 82). La formation de la liaison P-CF<sub>2</sub> a été observée seulement avec le composé **244b**. Cependant, la formation d'un sous-produit en quantité non-négligeable et la non-réactivité du *H*-phosphinothioate **244b** sur la difluorooléfine d'intérêt nous ont conduit à abandonner cette stratégie.

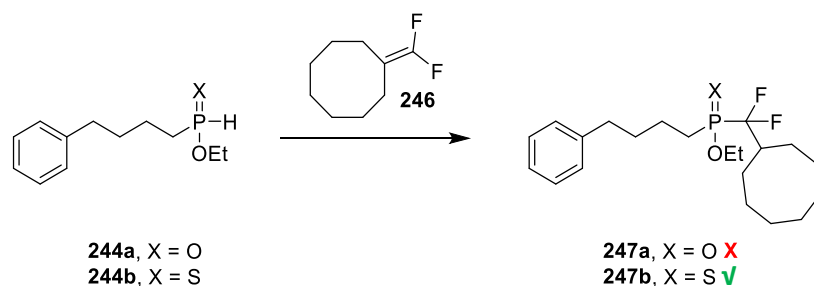


Schéma 82 : Différence de réactivité entre H-P=O et H-P=S

Ces résultats décevants nous ont amenés à considérer une troisième stratégie (voie B) reposant sur la formation de la seconde liaison phosphore-carbone (P-CH<sub>2</sub>) par addition radicalaire ou ionique (Schéma 83). Lors de l'addition radicalaire du *H*-phosphinothioate **165d** sur les oléfines **185b** et **185c**, seules des traces des intermédiaires nucléotide-furanose **197d** et **197a** ont été observées dans les mélanges bruts réactionnels par spectrométrie de RMN du <sup>19</sup>F découplé du proton. Une réaction de Michaelis-Becker pour former la seconde liaison P-CH<sub>2</sub> a, par contre, permis de synthétiser quatre intermédiaires de type nucléotide-furanose (**197a-d**) avec des rendements compris entre 47% et 63%.

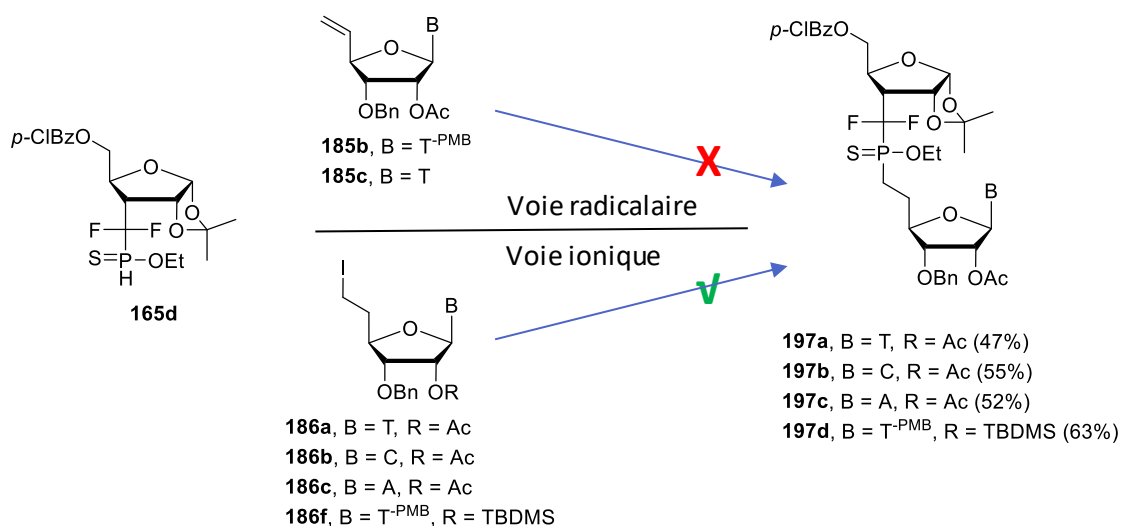


Schéma 83 : Stratégie radicalaire versus stratégie ionique

Avant l'étude de l'introduction d'une seconde base nucléique différente, un homodinucléotide T-T a été synthétisé à partir de l'intermédiaire **197a** afin de comparer cette stratégie aux travaux précédents du laboratoire. La synthèse de première génération (développée précédemment au laboratoire) reste la plus adaptée à la formation d'homodinucléotides. Notre stratégie de seconde génération conduit, quant à elle, à l'obtention d'intermédiaire nucléotide-furanose de manière plus efficace ; ces derniers constituent la clé permettant l'accès aux hétérodinucléotides.

Par la suite, l'introduction d'une seconde base nucléique différente a été réalisée grâce à notre stratégie de seconde génération. Dans un premier temps, un hétérodinucléotide T-<sup>PMB</sup>-U **183a** a été obtenu avec succès. Dans un second temps, une tentative d'introduction d'un uracile sur un substrat comportant une thymine non-protégée a été entreprise. Néanmoins, une substitution de la thymine par l'uracile a été observée. Un problème identique a été rencontré lors de l'addition d'une cytosine sur le même substrat thymine non-protégée. Ainsi, ces difficultés synthétiques ont mis en évidence le phénomène de trans-*N*-glycosylation dans notre stratégie d'introduction successive des bases nucléiques après formation du lien internucléosidique. Deux solutions possibles ont été mises en œuvre pour tenter de s'affranchir de ce phénomène.

La première solution a consisté à réaliser une substitution contrôlée d'une base par une autre. Cependant, la structure du dinucléotide obtenu n'a pu être déterminée avec certitude. Une seconde méthode a donc été mise en place en essayant de définir un ordre préférentiel d'introduction des nucléobases. Grâce à cette solution, un hétérodinucléotide supplémentaire (C-T **183d**) a été synthétisé (Figure 66). Il est également à noter qu'un autre hétérodinucléotide C-G<sup>Prot</sup> ou G<sup>Prot</sup>-C a été détecté à l'état de traces.

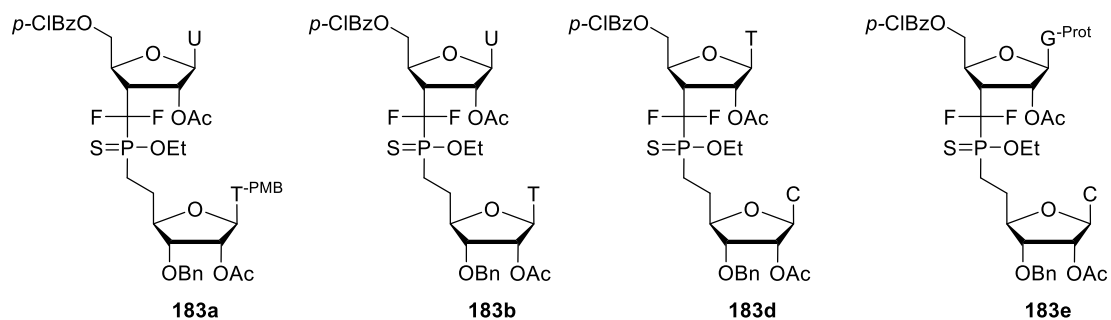


Figure 66 : Récapitulatif des hétérodimères synthétisés

Le principal objectif de ces travaux était de mettre au point une voie de synthèse efficace menant à des homodinucléotides ou hétérodinucléotides. Malgré des difficultés rencontrées, la voie B a permis d'atteindre en partie cet objectif. De nombreux problèmes restent cependant à résoudre pour transformer ces premiers résultats positifs en une stratégie de synthèse efficace d'hétérodinucléotides.

## II. Perspectives

D'autres solutions permettant d'éviter le problème de trans-*N*-glycosylation seront évoquées avant de présenter la fin de synthèse menant au dinucléotide-cible.

### II.1. Affranchissement du problème de trans-*N*-glycosylation

#### II.1.a. Protection du *H*-phosphinothioate **165d**

Une première solution, déjà présentée comme étant la voie E dans le chapitre 1 (V), consisterait à protéger le *H*-phosphinothioate **165d** afin d'introduire une base nucléique sur ce fragment. Un groupement protecteur résistant aux conditions acides de l'hydrolyse de l'acétonide et d'introduction de la base nucléique doit être développé. Lors de la synthèse automatisée, un groupement 2-cyanoéthyle **282**, résistant aux conditions acides, est utilisé pour protéger le phosphate (Schéma 84). Celui-ci est libéré en milieu basique pour former le phosphodiester **284** et l'acrylonitrile **283**, par un processus de type rétro-Michael.

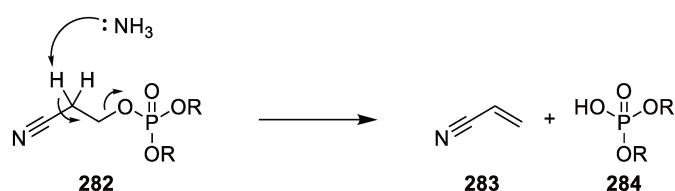


Schéma 84 : Hydrolyse du 2-cyanoéthyle

Par analogie, un groupement similaire pourrait être utilisé afin de protéger le *H*-phosphinothioate **165d** (Schéma 85). Le composé **208** pourrait être obtenu par réaction de Michaelis-Becker avec le 3-iodopropanenitrile. Par la suite, la séquence déprotection-protection du diol C1'-C2', suivie de l'introduction de la base nucléique seraient réalisées selon des conditions mises au point au laboratoire. Seules les conditions de la dernière étape d'hydrolyse du 2-cyanoéthyle devraient être déterminées.

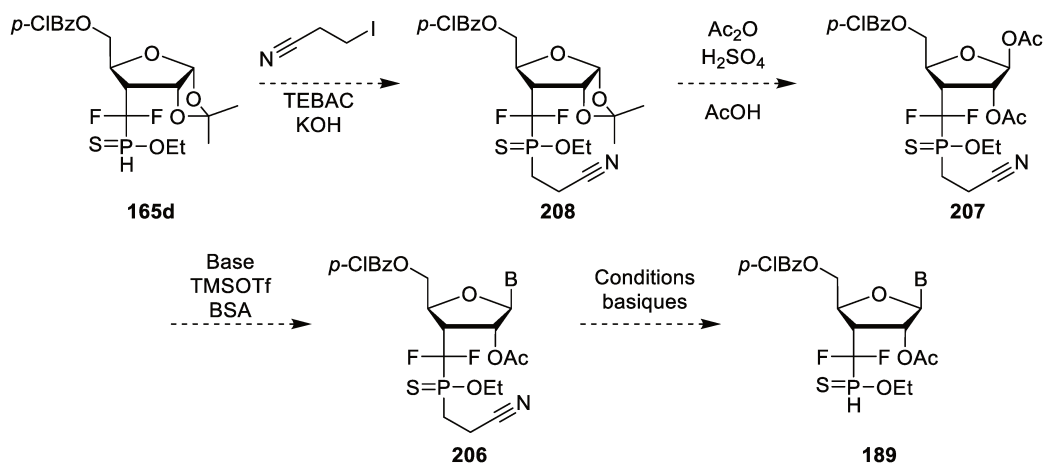


Schéma 85 : Stratégie de protection du *H*-phosphinothioate **165d**

### II.1.b. Insertion de la base avant la formation du motif difluorophosphinate

La principale solution pour éliminer les réactions de trans-*N*-glycosylation est d'insérer les bases nucléiques avant de former le lien internucléosidique. En s'inspirant de travaux de la littérature portant sur le développement de bioisostères de l'acide carboxylique,<sup>158</sup> le motif *H*-difluorophosphinate pourrait être obtenu à partir d'un lien difluorophosphonate. A partir du phosphonate de diester **285**, dont la synthèse a déjà été réalisée au laboratoire,<sup>106</sup> une mono-déprotection d'un des groupements éthyle d'un oxygène sp<sup>3</sup> pourrait être effectuée en présence de pyridine et de NaI (Schéma 86). Selon les conditions de Vilsmeier, le fragment **286** pourrait être converti en phosphonochloridate **287**, en présence de chlorure d'oxalyle. A partir de cet intermédiaire **287**, le *H*-phosphinate **288** pourrait être formé par réduction du chlore à l'aide de LiAlH(O<sup>t</sup>Bu)<sub>3</sub>.<sup>158</sup>

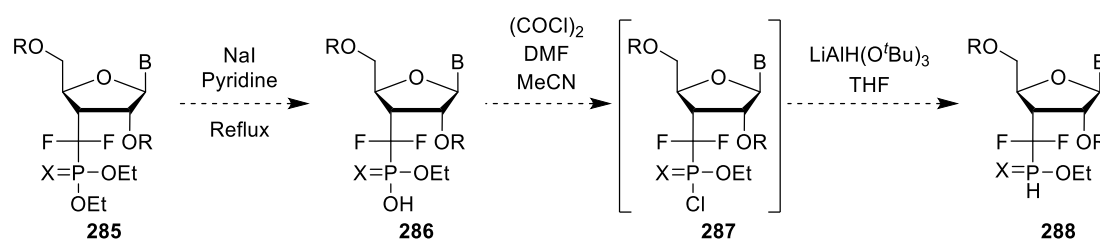


Schéma 86 : Formation du *H*-phosphinate à partir du phosphonate

### II.1.c. Blocage de l'oxygène en position C2'

Une dernière méthode serait de bloquer l'oxygène en position C2' afin d'éviter que celui-ci ne vienne stabiliser l'intermédiaire cation oxocarbenium responsable en partie du phénomène de trans-*N*-glycosylation. A la fin de son étude, Sawamoto a montré que lorsqu'un groupement acétate a été introduit sur l'azote pontant en C2' **289**, aucune réaction de trans-*N*-glycosylation n'a été observée (Schéma 87).<sup>154</sup> Le doublet non-liant de l'azote joue potentiellement un rôle important dans la stabilisation de l'oxocarbenium. Ainsi, par analogie, si l'oxygène en C2' est bloqué, les réactions de trans-*N*-glycosylation pourraient être évitées.

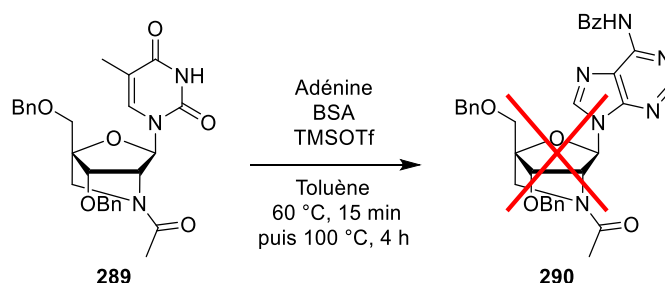


Schéma 87 : Blocage de la position C2'

<sup>158</sup> Pyun, H.-J.; Clarke, M. O.; Cho, A.; Casarez, A.; Ji, M.; Fardis, M.; Pastor, R.; Sheng, X. C. and Kim, C. U. *Tetrahedron Lett.* **2012**, 53, 2360-2363.

Pour bloquer l'oxygène sur cette position C2', un groupement volumineux tel que le TBDPS (**182e**) pourrait y être introduit afin d'éviter la stabilisation du cation. (Figure 67). Ces solutions pourraient favoriser la formation d'un hétérodinucléotide B<sub>1</sub>-B<sub>2</sub> (**183f**).

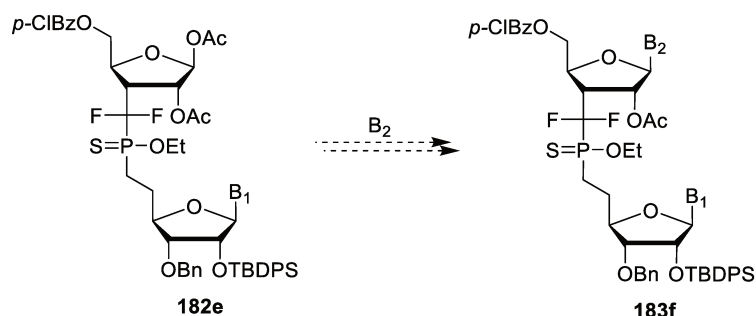


Figure 67 : Proposition de groupement protecteur de l'oxygène en C2'

## II.2. Fin de la synthèse

Afin d'être introduit dans un oligonucléotide par synthèse automatisée, une modification des groupements protecteurs doit être effectuée. Tout d'abord, les groupements protecteurs de **183** devront être clivés afin d'obtenir les alcools libres **291** (Schéma 88). Les alcools résultants pourront ainsi être protégés sous la forme d'éther silylé **292a**. Par la suite, le phosphinothioate **292a** pourra, ou non, être converti en phosphinate **292b** à l'aide de *m*-CPBA.<sup>107</sup> La fin de la synthèse consistera à cliver sélectivement l'alcool primaire en position C5' du furanose du haut, à réaliser une hydrogénolyse du groupement benzyloxy et à déséthylater afin de générer le motif acide difluorophosphin(othio)ique, qu'il conviendra de protéger sous forme d'ester 2-cyanoéthyle.

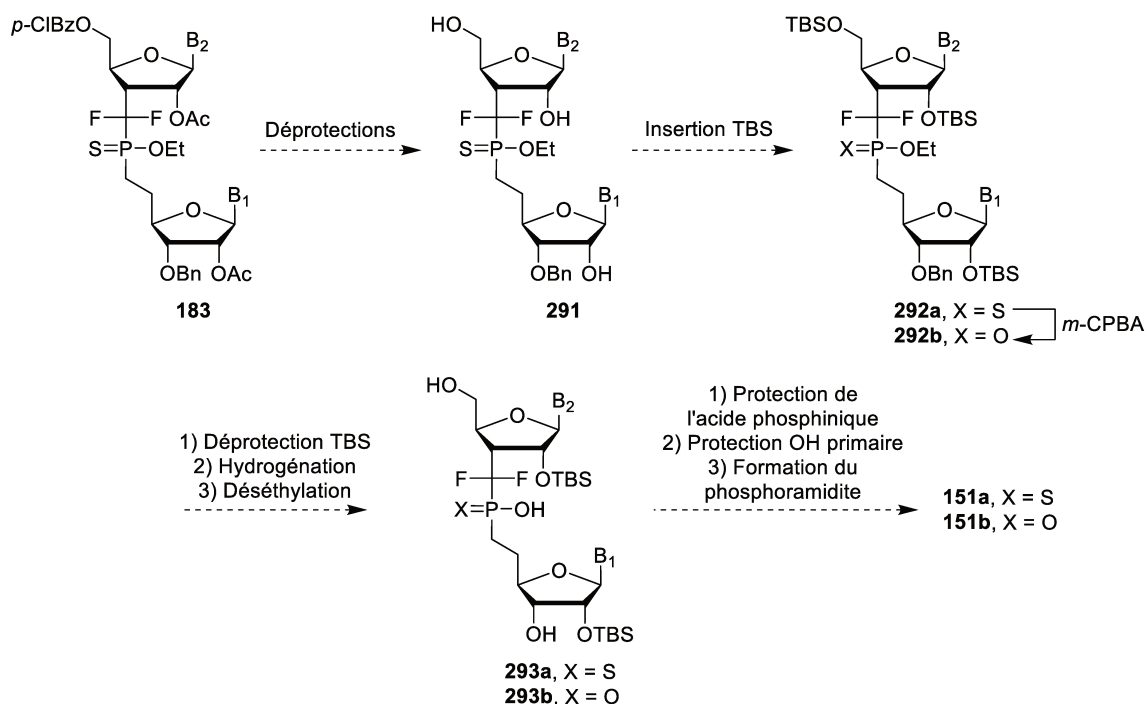


Schéma 88 : Fin de la synthèse

Pour finir, un groupement diméthoxytrityle et la formation du phosphoramidite devront être effectués afin de pouvoir introduire le dinucléotide dans un oligonucléotide par synthèse automatisée et de pouvoir étudier ainsi les propriétés physicochimiques des motifs difluorophosphate et difluorophosphinothioate.

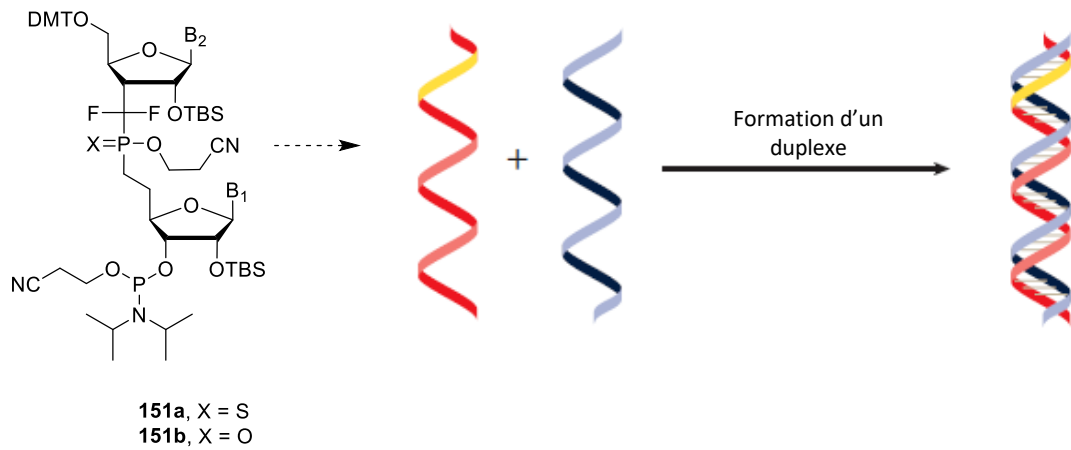


Figure 68 : Insertion dans un oligonucléotide





**Experimental part**



## I. General procedure

All reactions were performed, under an atmosphere of Argon, using dried glassware and magnetic stirring. Chemicals and solvents were purchased from Acros, Aldrich, Carbosynth and VWR and used without further purification. All reaction solvents were used anhydrous. 1,2-dichloroethane and acetonitrile were dried by distillation over calcium hydride. Tetrahydrofuran was distilled over sodium/benzophenone. Dichloromethane and toluene were dried with solvent purification system. Methanol was used in HPLC quality. Dimethylformamide and dimethylsulfoxide were purchased dried. Column chromatography purifications were carried out on silica gel (60-200  $\mu\text{m}$ ) from VWR chemicals. Thin-layer chromatography (TLC) was performed on Merck TLC silica gel 60 F<sub>254</sub> aluminium sheets. Compounds were visualized by illumination with a UV lamp (254 nm) or by TLC stain (KMnO<sub>4</sub> or Ceric Ammonium Molybdate). Purifications in reverse phase were realized on Puriflash 4250 from Interchim<sup>®</sup>, with UV detector. Two methods were used:

- Method A: Column Nucleodur PFP 250\*21.2 mm\_5  $\mu\text{m}$ ; Acetonitrile/Milli-Q<sup>®</sup> water, gradient from 50/50 to 85/15 in 25 min; UV detection at 230 nm and 254 nm.
- Method B: Hypersil Gold C18 250\*30 mm\_5  $\mu\text{m}$ ; Acetonitrile/0.1% formic acid, gradient from 50/50 to 100/0 in 90 min; UV detection at 230 nm and 254 nm.
- Method C: Synchronis C18 250x21.2 mm\_5  $\mu\text{m}$ ; Acetonitrile/0.1% formic acid, gradient from 45/55 to 100/0 in 70 min; UV detection at 230 nm and 254 nm.

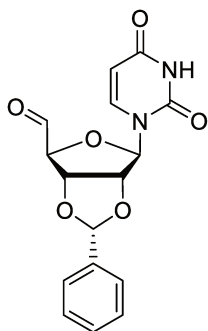
<sup>1</sup>H NMR (300 MHz), <sup>19</sup>F NMR (282 MHz), <sup>31</sup>P NMR (121 MHz) and <sup>13</sup>C NMR (75 MHz or 100 MHz) spectra were realized on Bruker AVIII-300 MHz and Bruker AVIII HD-400 MHz spectrometers. The solvent was chloroform deuterated (CDCl<sub>3</sub>) with a calibration at 7.24 ppm for <sup>1</sup>H spectra and 77.23 ppm for <sup>13</sup>C spectra. Chemical shifts ( $\delta$ ) are expressed in parts per million (ppm) and coupling constant ( $J$ ) in Hertz (Hz). Proton decoupled are noted {<sup>1</sup>H}. The following abbreviations were used to designate multiplicities: s (singlet), d (doublet), dd (doublet of doublet), t (triplet), dt (doublet of triplet), m (multiplet) and br (broad signal). Additional 2D NMR experiments (COSY, HSQC, HMBC, NOESY) were performed to determined correct structure.

IR spectra were recorded on a Perkin-Elmer 16PC IRTF spectrometer with diamond ATR accessory. High Resolution Mass Spectra were recorded on a JEOL AccuTOF™ 4G mass spectrometer (for EI mode) and a Waters LCT Premier spectrometer (for ESI ionization). Optical rotations were performed on a Perkin-Elmer 341 polarimeter using the sodium D line (589 nm) and the values are given in unit deg.mL.dm<sup>-1</sup>.g<sup>-1</sup>. Elemental analysis was realized on Thermo EAGER 300 Flash 2000 (ThermoFisher scientific). Melting points were measured on Stuart SMP3.

## II. Characterizations chapter 2

### II.1. 6-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-2-phenyltetrahydrofuro[3,4-d][1,3]dioxole-4-carbaldehyde **211**

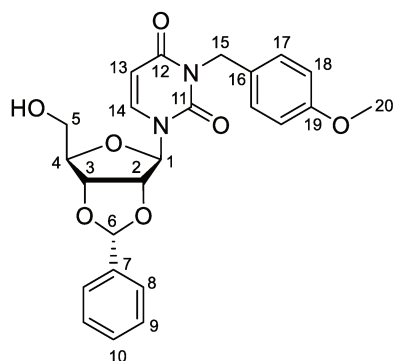
Protected uridine **210** (500.0 mg, 1.51 mmol, 1 eq) was dissolved in DMSO (3.75 mL). EDCI.HCl (862.7 mg, 4.53 mmol, 3 eq) and dichloroacetic acid (125  $\mu$ L, 1.51 mmol, 1 eq) were added and the reaction mixture was stirred for 24 h at room temperature. After this time, water (7.0 mL) and ethyl acetate (7.0 mL) were added in the mixture under vigorous stirred. The aqueous layer was extracted twice with AcOEt. Then, the organic layer was washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure to afford the crude product **211** as a white solid. **211** was engaged without further purification in the next step.



Chemical Formula: C<sub>16</sub>H<sub>14</sub>N<sub>2</sub>O<sub>6</sub>  
Exact Mass: 330.09

### II.2. 1-((2R,3aR,4R,6R,6aR)-6-(hydroxymethyl)-2-phenyltetrahydrofuro[3,4-d][1,3]dioxol-4-yl)-3-(4-methoxybenzyl)pyrimidine-2,4(1H,3H)-dione **219**

To a solution of **210** (200 mg, 0.60 mmol, 1 eq) in acetonitrile (3.0 mL) were added *para*-methoxybenzoyl (90  $\mu$ L, 0.66 mmol, 1.1 eq) and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) (98  $\mu$ L, 0.66 mmol, 1.1 eq). The reaction mixture was stirred at room temperature for 40 h. After this time, the reaction mixture was diluted with ethyl acetate and washed with a saturated aqueous solution of sodium bicarbonate. The aqueous layer was extracted twice with ethyl acetate. The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography on silica gel (dichloromethane/methanol, gradient from 100/0 to 98/2) to afforded **219** as an 8.5:1.5 mixture of diastereoisomers and as a white solid (235 mg, 87%).



Chemical Formula: C<sub>24</sub>H<sub>24</sub>N<sub>2</sub>O<sub>7</sub>  
Exact Mass: 452.16

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.51-7.39 (m, 7H, H<sup>Ar</sup>, H<sup>17</sup>), 7.25 (d, 1H, *J* = 7.7 Hz, H<sup>14</sup>), 6.87\* (d, 0.4H, *J* = 8.7 Hz, H<sup>18</sup>), 6.82 (d, 1.6H, *J* = 8.7 Hz, H<sup>18</sup>), 6.07\* (s, 0.2H, H<sup>6</sup>), 5.99 (s, 0.8H, H<sup>6</sup>), 5.79\* (d, 0.2H, *J* = 7.8 Hz, H<sup>13</sup>), 5.74 (d, 0.8H, *J* = 7.7 Hz, H<sup>13</sup>), 5.61 (d, 1H, *J* = 2.7 Hz, H<sup>1</sup>), 5.18 (dd, 1H, AB syst, *J* = 6.8 Hz, 2.7 Hz, H<sup>2</sup>), 5.09 (dd, 1H, AB syst, *J* = 6.8 Hz, 3.0 Hz, H<sup>4</sup>), 5.04 (d, 1H, AB syst, *J* = 13.8 Hz, H<sup>15</sup>), 4.99 (d, 1H, AB syst, *J* = 13.8 Hz, H<sup>15</sup>), 4.44 (m, 0.8H, H<sup>3</sup>), 4.31\* (m, 0.2H, H<sup>3</sup>), 3.93\* (br s, 0.2H, H<sup>5</sup>), 3.90 (br m, 0.8H, H<sup>5</sup>), 3.83 (dd, 0.8H, *J* = 6.9 Hz, 3.0 Hz, H<sup>5</sup>), 3.79-3.78\* (br m, 0.8H, H<sup>20</sup>), 3.76 (s, 2.4H, H<sup>5</sup>, H<sup>20</sup>), 2.97-2.93 (m, 0.8H, OH), 2.76-2.74\* (m, 0.2H, OH).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ 162.5 (C<sup>12</sup>), 159.3 (C<sup>19</sup>), 151.2 (C<sup>11</sup>), 140.9 (C<sup>14</sup>), 135.9 (C<sup>7</sup>), 131.0 (3C, C<sup>Ar</sup>, C<sup>16</sup>), 130.1 (C<sup>10</sup>), 128.7 (2C, C<sup>17</sup>), 126.8 (2C, C<sup>Ar</sup>), 113.9 (2C, C<sup>18</sup>), 107.9 (C<sup>6</sup>), 102.5 (C<sup>13</sup>), 97.0 (C<sup>1</sup>), 86.9 (C<sup>3</sup>), 84.5 (C<sup>2</sup>), 81.5 (C<sup>4</sup>), 62.9 (C<sup>5</sup>), 55.4 (C<sup>20</sup>), 43.8 (C<sup>15</sup>).

HRMS calculated for C<sub>24</sub>H<sub>25</sub>N<sub>2</sub>O<sub>7</sub> [M+H]<sup>+</sup>: 453.1656. Found 453.1662.

FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 3070, 2945, 1661, 1511, 1454, 1242, 1093, 806, 698.

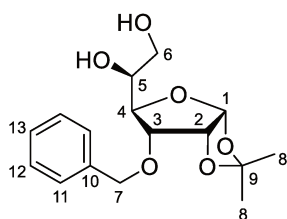
M.p. = 60-61 °C.

R<sub>f</sub> (dichloromethane/methanol, 98/2, v/v) = 0.52.

II.3. (S)-1-((3aR,5R,6R,6aR)-6-(benzyloxy)-2,2-dimethyltetrahydrofuro[2,3-d][1,3] dioxol-5-yl)ethane-1,2-diol **222**

Following a protocol of the literature,<sup>159</sup> 3-*O*-benzyl-1,2:5,6-di-*O*-isopropylidene-α-D-allofuranose (5.00 g, 14.2 mmol, 1 eq) was dissolved in acetic acid (20.0 mL) and water (5.0 mL). The mixture was stirred at room temperature for 96 h. The solution was poured into an ice-cold solution of sodium hydroxide (15.0 g in 75.0 mL of water) with vigorous stirring and addition of ice to keep the temperature below 20 °C. The mixture was extracted with ethyl acetate and the combined organics were washed with brine, dried over MgSO<sub>4</sub> and concentrated under reduced pressure to afford the expected product as a colorless oil (4.05 g, 91%). The diol **222** was used in the next step without further purification.

<sup>159</sup> Yoshikawa, M.; Okaichi, Y.; Cheon Cha, B. and Kitagawa, I. *Tetrahedron* **1990**, *46*, 7459-7470.



Chemical Formula: C<sub>16</sub>H<sub>22</sub>O<sub>6</sub>  
Exact Mass: 310.14

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.36-7.30 (m, 5H, H<sup>A</sup>), 5.74 (d, 1H, *J* = 3.9 Hz, H<sup>1</sup>), 4.77 (d, 1H, AB syst, *J* = 11.1 Hz, H<sup>7</sup>), 4.59 (dd, apparent t, 1H, *J* = 3.9 Hz, H<sup>2</sup>), 4.54 (d, 1H, AB syst, *J* = 11.1 Hz, H<sup>7</sup>), 4.08 (dd, 1H, *J* = 9.0 Hz, 3.3 Hz, H<sup>4</sup>), 3.99 (q, 1H, *J* = 8.7 Hz, 4.5 Hz, H<sup>5</sup>), 3.91 (dd, 1H, *J* = 9.0 Hz, 4.5 Hz, H<sup>3</sup>), 3.69-3.64 (m, 2H, H<sup>6</sup>), 2.54 (d, 1H, *J* = 3.9 Hz, OH), 2.47 (dd, 1H, *J* = 7.5 Hz, 5.7 Hz, OH), 1.57 (s, 3H, H<sup>8</sup>), 1.34 (s, 3H, H<sup>8</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ 136.9 (C<sup>10</sup>), 128.8 (2C, C<sup>12</sup>), 128.5 (3C, C<sup>11</sup>, C<sup>13</sup>), 113.4 (C<sup>9</sup>), 104.4 (C<sup>1</sup>), 79.3 (C<sup>4</sup>), 77.5 (C<sup>2</sup>), 77.0 (C<sup>3</sup>), 72.4 (C<sup>7</sup>), 71.0 (C<sup>5</sup>), 63.3 (C<sup>6</sup>), 27.0 (C<sup>8</sup>), 26.8 (C<sup>8</sup>).

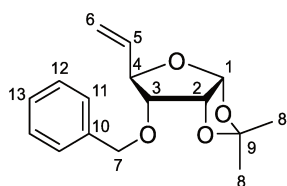
FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 3418, 2986, 2934, 1682, 1497, 1454, 1372, 1310, 1215, 1166, 1120, 1079, 1016, 870, 738, 698, 660, 611, 507.

[α]<sub>D</sub><sup>23</sup> + 60.4 (c 1.70, CHCl<sub>3</sub>).

R<sub>f</sub> (dichloromethane/methanol, 95/5, v/v) = 0.48.

#### II.4. (3aR,5R,6R,6aR)-6-(benzyloxy)-2,2-dimethyl-5-vinyltetrahydro[2,3-d][1,3] dioxole **171**

To a 0 °C solution of diol **222** (3.9 g, 12.6 mmol, 1 eq) in dichloromethane (110.0 mL) were added (under vigorous stirring) triphenylphosphine (8.26 g, 31.5 mmol, 2.5 eq), imidazole (2.14 g, 31.5 mmol, 2.5 eq) and iodine (6.08 g, 23.9 mmol, 1.9 eq). The solution was stirred at room temperature for 24 h. After this time, the mixture was washed with a saturated aqueous solution of Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub>. The organic layer was washed with water, brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under vacuum. The crude residue was purified by flash chromatography on silica gel (petroleum ether/ethyl acetate, gradient from 9/1 to 8/2) to afford olefin **171** as colorless oil (3.32 g, 96%).



Chemical Formula: C<sub>16</sub>H<sub>20</sub>O<sub>4</sub>  
Exact Mass: 276.14

**According to literature data:**<sup>160</sup>

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.37-7.26 (m, 5H, H<sup>A</sup>), 5.80 (ddd, 1H, *J* = 17.1 Hz, 10.4 Hz, 6.7 Hz, H<sup>5</sup>), 5.72 (d, 1H, *J* = 3.7 Hz, H<sup>1</sup>), 5.44 (ddd, apparent dt, 1H, *J* = 17.1 Hz, 1.4 Hz, H<sup>6</sup>), 5.26 (ddd, 1H, *J* = 10.5 Hz, 1.4 Hz, 1.0 Hz, H<sup>6</sup>), 4.74 (d, 1H, AB syst, *J* = 12.2 Hz, H<sup>7</sup>), 4.60 (d, 1H, AB syst, *J* = 12.2 Hz, H<sup>7</sup>), 4.55 (dd, apparent t, 1H, *J* = 4.0 Hz, H<sup>2</sup>), 4.47 (dd, 1H, *J* = 8.9 Hz, 6.6 Hz, H<sup>4</sup>), 3.50 (dd, 1H, *J* = 8.9 Hz, 4.3 Hz, H<sup>3</sup>), 1.61 (s, 3H, H<sup>8</sup>), 1.35 (s, 3H, H<sup>8</sup>).

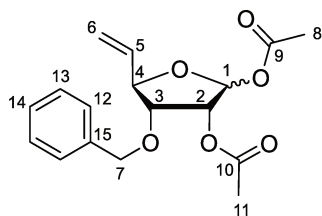
<sup>160</sup> Krohn, K.; Flörke, U.; Keine, J. and Terstiege, I. *J. Carbohydr. Chem.* **1998**, *17*, 171-195.

$^{13}\text{C}$   $\{^1\text{H}\}$  NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  137.7 ( $\text{C}^{10}$ ), 135.0 ( $\text{C}^5$ ), 128.5 (2C,  $\text{C}^{12}$ ), 128.1 ( $\text{C}^{13}$ ), 128.0 (2C,  $\text{C}^{11}$ ), 118.9 ( $\text{C}^6$ ), 113.0 ( $\text{C}^9$ ), 103.9 ( $\text{C}^1$ ), 81.9 ( $\text{C}^3$ ), 79.2 ( $\text{C}^4$ ), 77.7 ( $\text{C}^2$ ), 72.3 ( $\text{C}^7$ ), 26.9 ( $\text{C}^8$ ), 26.6 ( $\text{C}^8$ ).

Rf (petroleum ether/ethyl acetate, 9/1, v/v) = 0.34.

II.5. (2R,3R,4R,5R)-4-(benzyloxy)-5-vinyltetrahydrofuran-2,3-diyl diacetate **223a**

According to literature method,<sup>115</sup> to a solution of **171** (100 mg, 364.7  $\mu\text{mol}$ , 1 eq) in acetic acid (20.0 mL) were added anhydride acetic (1.3 mL, 14.6 mmol, 40 eq) and camphorsulfonic acid (4 mg, 18.2  $\mu\text{mol}$ , 0.05 eq). The mixture was stirred at 75 °C for 1 h and a second portion of camphorsulfonic acid was then added. The reaction mixture was stirred for additional 1 h. After this time, the reaction mixture was diluted with dichloromethane, washed with water and a saturated aqueous solution of sodium bicarbonate, dried over  $\text{Na}_2\text{SO}_4$  and concentrated under reduced pressure. The crude residue, as a mixture 4:1 of anomers, determined by  $^1\text{H}$  NMR, was purified by flash chromatography on silica gel (pentane/ethyl acetate, 8/2) to afford the expected product **223a** as a colorless oil (105 mg, 92%).



Chemical Formula:  $\text{C}_{17}\text{H}_{20}\text{O}_6$   
Exact Mass: 320.13

**Characterizations realized on mixture 9:1 of pure anomers (determined by  $^1\text{H}$  NMR):**

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ) (\*refers to the minor anomer when unambiguous distinction is possible):  $\delta$  7.28-7.18 (m, 5H,  $\text{H}^{\text{Ar}}$ ), 6.32\* (d, 0.1H,  $J = 4.5$  Hz,  $\text{H}^1$ ), 6.07 (s, 0.9H,  $\text{H}^1$ ), 5.75 (ddd, 0.9H,  $J = 17.1$  Hz, 10.2 Hz, 6.8 Hz,  $\text{H}^5$ ), 5.66\* (ddd, 0.1H,  $J = 17.1$  Hz, 10.2 Hz, 6.8 Hz,  $\text{H}^5$ ), 5.28 (ddd, apparent td, 0.9H,  $J = 17.1$  Hz, 1.3 Hz,  $\text{H}^6$ ), 5.27\* (dm, 0.1H,  $J = 17.1$  Hz,  $\text{H}^6$ ), 5.22 (d, 1H,  $J = 4.4$  Hz,  $\text{H}^2$ ), 5.15 (ddd, apparent td, 0.9H,  $J = 10.2$  Hz, 1.3 Hz,  $\text{H}^6$ ), 5.15-5.10\* (m, 0.1H,  $\text{H}^6$ ), 4.55 (d, 1H, AB syst,  $J = 11.5$  Hz,  $\text{H}^7$ ), 4.48-4.38 (m, 1H,  $\text{H}^4$ ), 4.48 (d, 1H, AB syst,  $J = 11.5$  Hz,  $\text{H}^7$ ), 3.93 (dd, 0.9H,  $J = 7.8$  Hz, 4.4 Hz,  $\text{H}^3$ ), 3.79\* (dd, 0.1H,  $J = 6.2$  Hz, 5.6 Hz,  $\text{H}^3$ ), 2.05 (s, 2.5H,  $\text{H}^{11}$ ), 2.03\* (s, 0.5H,  $\text{H}^{11}$ ), 1.97 (s, 3H,  $\text{H}^8$ ).

$^{13}\text{C}$   $\{^1\text{H}\}$  NMR (75 MHz,  $\text{CDCl}_3$ ) (\*refers to the minor anomer when unambiguous distinction is possible):  $\delta$  170.0\* ( $\text{C}^{10}$ ), 169.9 ( $\text{C}^{10}$ ), 169.1 ( $\text{C}^9$ ), 137.6\* ( $\text{C}^{15}$ ), 137.4 ( $\text{C}^{15}$ ), 136.2 ( $\text{C}^5$ ), 134.7\* ( $\text{C}^5$ ), 128.6 (2C,  $\text{C}^{13}$ ), 128.1 ( $\text{C}^{14}$ ), 127.9 (2C,  $\text{C}^{12}$ ), 118.3 ( $\text{C}^6$ ), 98.5 ( $\text{C}^1$ ), 94.4\* ( $\text{C}^1$ ), 83.8\* ( $\text{C}^4$ ), 83.0 ( $\text{C}^4$ ), 80.8 ( $\text{C}^3$ ), 79.1\* ( $\text{C}^3$ ), 73.8 ( $\text{C}^2$ ), 73.30 ( $\text{C}^7$ ), 73.28\* ( $\text{C}^7$ ), 70.7\* ( $\text{C}^2$ ), 21.2 ( $\text{C}^{11}$ ), 20.8 ( $\text{C}^8$ ), 20.7\* ( $\text{C}^8$ ).

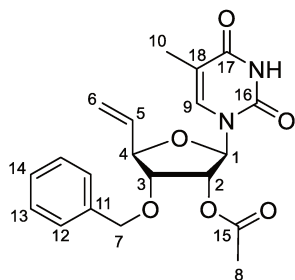
HRMS (TOF ESI+) calculated for  $\text{C}_{17}\text{H}_{24}\text{NO}_6$   $[\text{M}+\text{NH}_4]^+$ : 338.1604. Found 338.1605.

FTIR ( $\nu_{\text{max}}$   $\text{cm}^{-1}$ ) 2989, 1742, 1455, 1370, 1212, 1120, 1095, 1062, 1004, 951, 738, 698.

Rf (petroleum ether/ethyl acetate, 4/1, v/v) = 0.25 and 0.37.

II.6. (3R,4R,5R)-4-(benzyloxy)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-5-vinyltetrahydrofuran-3-yl acetate **185b**

According to Vorbrüggen conditions,<sup>134</sup> to a solution of diacetate **223a** (4.9 g, 15.3 mmol, 1 eq) in acetonitrile (70.0 mL) were added thymine (2.9 g, 22.9 mmol, 1.5 eq) and *N,O*-bis(trimethylsilyl)acetamide (11.2 mL, 45.9 mmol, 3 eq) at room temperature. The reaction mixture was stirred at 50 °C for 5 minutes. After cooling at room temperature, trimethylsilyl trifluoromethanesulfonate (3.0 mL, 16.8 mmol, 1.1 eq) was added and the solution was heated at reflux for 1.5 hours. After this time, the mixture was quenched with a saturated aqueous solution of sodium bicarbonate and extracted with ethyl acetate. The organic layer was washed with water, brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under vacuum. The crude residue was purified by flash chromatography on silica gel (petroleum ether/ethyl acetate, gradient from 6/4 to 5/5) to afford the nucleoside **185b** as white solid (4.73 g, 80%).<sup>161</sup>



Chemical Formula: C<sub>20</sub>H<sub>22</sub>N<sub>2</sub>O<sub>6</sub>  
Exact Mass: 386.15

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 8.30 (s, 1H, NH), 7.35-7.25 (m, 5H, H<sup>A</sup>), 7.01 (d, 1H, *J* = 1.1 Hz, H<sup>9</sup>), 5.88 (ddd, 1H, *J* = 17.0 Hz, 10.4 Hz, 6.5 Hz, H<sup>5</sup>), 5.80 (d, 1H, *J* = 3.4 Hz, H<sup>1</sup>), 5.42 (ddd, apparent dt, 1H, *J* = 17.2 Hz, 1.2 Hz, H<sup>6</sup>), 5.36 (dd, 1H, *J* = 5.7 Hz, 3.4 Hz, H<sup>2</sup>), 5.31 (ddd, apparent dt, 1H, *J* = 10.5 Hz, 1.1 Hz, H<sup>6</sup>), 4.59 (d, 1H, AB syst, *J* = 11.2 Hz, H<sup>7</sup>), 4.50 (d, 1H, AB syst, *J* = 11.2 Hz, H<sup>7</sup>), 4.29 (dd, apparent t, 1H, *J* = 6.7 Hz, H<sup>4</sup>), 4.03 (dd, 1H, *J* = 7.1 Hz, 5.7 Hz, H<sup>3</sup>), 2.10 (s, 3H, H<sup>8</sup>), 1.90 (d, 3H, *J* = 1.1 Hz, H<sup>10</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ 170.3 (C<sup>15</sup>), 163.9 (C<sup>17</sup>), 150.2 (C<sup>16</sup>), 137.3 (C<sup>11</sup>), 136.5 (C<sup>9</sup>), 134.4 (C<sup>5</sup>), 128.7 (2C, C<sup>13</sup>), 128.3 (C<sup>14</sup>), 128.2 (2C, C<sup>12</sup>), 119.4 (C<sup>6</sup>), 111.6 (C<sup>18</sup>), 90.3 (C<sup>1</sup>), 82.8 (C<sup>4</sup>), 79.7 (C<sup>3</sup>), 73.9 (C<sup>2</sup>), 73.8 (C<sup>7</sup>), 21.0 (C<sup>8</sup>), 12.8 (C<sup>10</sup>).

HRMS (TOF ESI+) calculated for C<sub>20</sub>H<sub>23</sub>N<sub>2</sub>O<sub>6</sub> [M+H]<sup>+</sup>: 387.1556. Found 387.1550.

FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 3070, 2945, 1739, 1716, 1666, 1470, 1373, 1233, 1217, 1050.

[α]<sub>D</sub><sup>22</sup> + 63.1 (c 1.30, CHCl<sub>3</sub>).

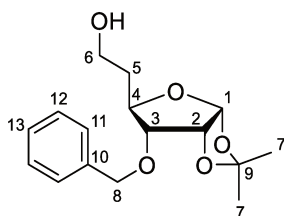
M.p. = 46-55 °C.

R<sub>f</sub> (petroleum ether/ethyl acetate, 3/2, v/v) = 0.21.

<sup>161</sup> Mixture of anomers α/β (ratio 6/94).

II.7. 2-((3aR,5R,6R,6aR)-6-(benzyloxy)-2,2-dimethyltetrahydrofuro[2,3-d][1,3]dioxol-5-yl)ethanol **226**

According to literature protocol,<sup>162</sup> to a 0° C solution of olefin **171** (4.05 g, 14.7 mmol, 1 eq) in tetrahydrofuran (80.0 mL) was added 9-borabicyclo[3.3.1]nonane (0.5M in THF, 59.0 mL, 29.4 mmol, 2 eq). The reaction mixture was allowed to warm up at room temperature and stirred for 18 h. After this time, the reaction mixture was cooled in an ice-water bath and treated carefully with an aqueous solution of sodium hydroxide (220.0 mL, 2M in H<sub>2</sub>O), followed by hydrogen peroxide (20.0 mL, 35% in H<sub>2</sub>O). The cold bath was removed and the mixture was stirred for 20 min. The aqueous layer was extracted with ethyl acetate (3 times). The organic extracts were washed with water, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography on silica gel (petroleum ether/ethyl acetate, gradient from 7/3 to 5/5) to afford the alcohol **226** as a colorless oil (4.3 g, 99%).



Chemical Formula: C<sub>16</sub>H<sub>22</sub>O<sub>5</sub>  
Exact Mass: 294.15

**According to literature data:**<sup>163</sup>

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.35-7.29 (m, 5H, H<sup>Ar</sup>), 5.67 (d, 1H, *J* = 3.8 Hz, H<sup>1</sup>), 4.73 (d, 1H, AB syst, *J* = 11.7 Hz, H<sup>8</sup>), 4.57 (dd, apparent t, 1H, *J* = 4.2 Hz, H<sup>2</sup>), 4.52 (d, 1H, AB syst, *J* = 11.7 Hz, H<sup>8</sup>), 4.07 (ddd, 1H, *J* = 8.9 Hz, 7.4 Hz, 5.0 Hz, H<sup>4</sup>), 3.72-3.68 (m, 2H, H<sup>6</sup>), 3.43 (dd, 1H, *J* = 8.9 Hz, 4.2 Hz, H<sup>3</sup>), 2.41 (s, 1H, -OH), 1.96-1.86 (m, 1H, H<sup>5</sup>), 1.81-1.70 (m, 1H, H<sup>5</sup>), 1.57 (s, 3H, H<sup>7</sup>), 1.33 (s, 3H, H<sup>7</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ 137.2 (C<sup>10</sup>), 128.7 (2C, C<sup>12</sup>), 128.4 (C<sup>13</sup>), 128.3 (2C, C<sup>11</sup>), 113.1 (C<sup>9</sup>), 104.2 (C<sup>1</sup>), 81.8 (C<sup>3</sup>), 77.5 (C<sup>4</sup>), 77.0 (C<sup>2</sup>), 72.3 (C<sup>8</sup>), 60.5 (C<sup>6</sup>), 35.0 (C<sup>5</sup>), 26.8 (C<sup>7</sup>), 26.7 (C<sup>7</sup>).

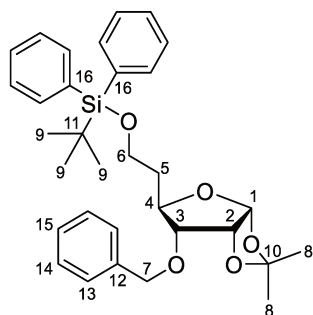
**R<sub>f</sub>** (petroleum ether/ethyl acetate, 1/1, v/v) = 0.36.

<sup>162</sup> Das, S. N.; Chowdhury, A.; Tripathi, N.; Jana, P. K. and Mandal, S. B. *J. Org. Chem.* **2015**, *80*, 1136-1148.

<sup>163</sup> Rauter, A. P.; Figueiredo, J. A.; Ismael, M. I. and Justino, J. J. *Carbohydr. Chem.* **2004**, *23*, 513-528.

II.8. 2-((3aR,5R,6R,6aR)-6-(benzyloxy)-2,2-dimethyltetrahydrofuro[2,3-d][1,3]dioxol-5-yl)ethoxy)(tert-butyl)diphenylsilane **227**

To a solution of **226** (167 mg, 0.568 mmol, 1 eq) and Imidazole (85 mg, 1.25 mmol, 2.2 eq) in *N,N*-dimethylformamide (10.0 mL) were added *tert*-butyldiphenylsilyl chloride (300  $\mu$ L, 1.13 mmol, 2 eq) and 4-dimethylaminopyridine (21 mg, 0.170 mmol, 0.3 eq). The resulting mixture was stirred for 40 h at room temperature. The mixture was quenched with water and extracted with twice with dichloromethane. The organic phase was washed with brine, dried over  $\text{Na}_2\text{SO}_4$  and concentrated under reduced pressure. The crude residue was purified by flash chromatography on silica gel (pentane/ethyl acetate, gradient from 96/4 to 90/10) to afford the expected product **227** as a colorless oil (273 mg, 90%).



Chemical Formula:  $\text{C}_{32}\text{H}_{40}\text{O}_5\text{Si}$   
Exact Mass: 532.26

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.68-7.64 (m, 4H,  $\text{H}^{\text{Ar}}$ ), 7.40-7.30 (m, 11H,  $\text{H}^{\text{Ar}}$ ), 5.68 (d, 1H,  $J = 3.8$  Hz,  $\text{H}^1$ ), 4.75 (d, 1H, AB syst,  $J = 12.0$  Hz,  $\text{H}^7$ ), 4.52 (dd, t apparent, 1H,  $J = 4.2$  Hz,  $\text{H}^2$ ), 4.52 (d, 1H, AB syst,  $J = 12.0$  Hz,  $\text{H}^7$ ), 4.20 (ddd, td apparent, 1H,  $J = 8.8$  Hz, 3.6 Hz,  $\text{H}^4$ ), 3.80-3.75 (m, 2H,  $\text{H}^6$ ), 3.41 (dd, 1H,  $J = 8.9$  Hz, 4.2 Hz,  $\text{H}^3$ ), 2.06-1.95 (m, 1H,  $\text{H}^5$ ), 1.74-1.63 (m, 1H,  $\text{H}^5$ ), 1.56 (s, 3H,  $\text{H}^8$ ), 1.33 (s, 3H,  $\text{H}^8$ ), 1.03 (s, 9H,  $\text{H}^9$ ).

$^{13}\text{C}$  { $^1\text{H}$ } NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  137.9 ( $\text{C}^{12}$ ), 135.81 (2C,  $\text{C}^{\text{Ar}}$ ), 135.79 (2C,  $\text{C}^{\text{Ar}}$ ), 134.2 ( $\text{C}^{16}$ ), 134.0 ( $\text{C}^{16}$ ), 129.7 (2C,  $\text{C}^{\text{Ar}}$ ), 128.6 (2C,  $\text{C}^{15}$ ,  $\text{C}^{\text{Ar}}$ ), 128.1 (3C,  $\text{C}^{\text{Ar}}$ ), 127.8 (4C,  $\text{C}^{\text{Ar}}$ ), 112.9 ( $\text{C}^{10}$ ), 104.2 ( $\text{C}^1$ ), 82.2 ( $\text{C}^3$ ), 77.5 ( $\text{C}^2$ ), 74.9 ( $\text{C}^4$ ), 72.3 ( $\text{C}^7$ ), 60.8 ( $\text{C}^6$ ), 35.6 ( $\text{C}^5$ ), 27.1 (3C,  $\text{C}^9$ ), 26.9 ( $\text{C}^8$ ), 26.8 ( $\text{C}^8$ ), 19.4 ( $\text{C}^{11}$ ).

HRMS (TOF ESI+) calculated for  $\text{C}_{32}\text{H}_{44}\text{NSiO}_5$  [ $\text{M} + \text{NH}_4$ ] $^+$ : 550.2889. Found 550.2991.

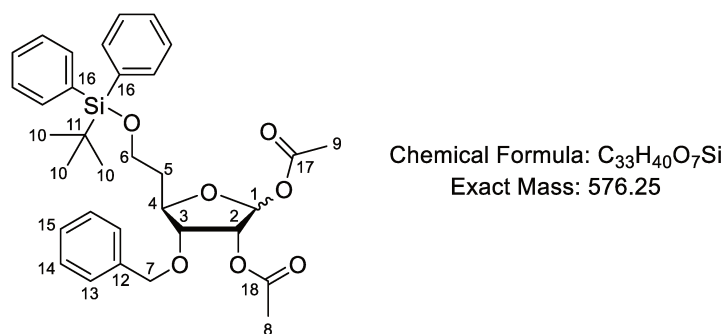
FTIR ( $\nu_{\text{max}}$   $\text{cm}^{-1}$ ) 2930, 2857, 1589, 1472, 1455, 1427, 1372, 1248, 1105, 1019, 872, 822, 736, 699.

$[\alpha]_{\text{D}}^{20} + 42.0$  (c 1.00,  $\text{CHCl}_3$ ).

R<sub>f</sub> (pentane/diethyl ether, 9/1, v/v) = 0.29.

II.9. (3R,4R,5R)-4-(benzyloxy)-5-(2-((*tert*-butyldiphenylsilyl)oxy)ethyl) tetrahydrofuran-2,3-diyl diacetate **228**

According to literature method,<sup>115</sup> to a solution of **227** (250 mg, 0.470 mmol, 1 eq) in acetic acid (25.0 mL) were added anhydride acetic (1.8 mL, 18.8 mmol, 40 eq) and camphorsulfonic acid (5 mg, 235  $\mu$ mol, 0.05 eq). The mixture was stirred at 75 °C for 1 h and a second portion of camphorsulfonic acid was then added. The reaction mixture was stirred for additional 1 h. After this time, the reaction mixture was diluted with dichloromethane, washed with water and a saturated aqueous solution of sodium bicarbonate, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue, as a mixture 4:1 of anomers (determined by <sup>1</sup>H NMR), was purified by flash chromatography on silica gel (pentane/ethyl acetate, gradient from 95/5 to 80/20) to afford the expected product **228** as a colorless oil (210 mg, 78%).



**Characterizations realized on mixture 4:1 of pure anomers (determined by <sup>1</sup>H NMR):**

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) (\*refers to the minor diastereoisomer when unambiguous distinction is possible)  $\delta$  7.61-7.57 (m, 4H, H<sup>Ar</sup>), 7.35-7.18 (m, 11H, H<sup>Ar</sup>), 6.28\* (d, 0.2H, *J* = 3.9 Hz, H<sup>1</sup>), 6.03 (s, 0.8H, H<sup>1</sup>), 5.20 (d, 0.8H, *J* = 4.5 Hz, H<sup>2</sup>), 5.07\* (dd, 0.2H, *J* = 6.5 Hz, 4.7 Hz, H<sup>2</sup>), 4.53\* (d, 0.2H, AB syst, *J* = 11.8 Hz, H<sup>7</sup>), 4.51 (d, 0.8H, AB syst, *J* = 11.3 Hz, H<sup>7</sup>), 4.42\* (d, 0.2H, AB syst, *J* = 11.8 Hz, H<sup>7</sup>), 4.37-4.34 (m, 0.2H, H<sup>4</sup>), 4.34 (d, 0.8H, AB syst, *J* = 11.3 Hz, H<sup>7</sup>), 4.16 (ddd, apparent td, 0.8H, *J* = 7.7 Hz, 4.8 Hz, H<sup>4</sup>), 3.90 (dd, 0.8H, *J* = 7.6 Hz, 4.7 Hz, H<sup>3</sup>), 3.80\* (dd, 0.2H, *J* = 6.5 Hz, 5.2 Hz, H<sup>3</sup>), 3.73-3.65 (m, 2H, H<sup>6</sup>), 2.04 (s, 3H, H<sup>8</sup>), 1.92 (s, 3H, H<sup>9</sup>), 1.90-1.64 (m, 2H, H<sup>5</sup>), 0.982 (s, 7H, H<sup>10</sup>), 0.977 (s, 2H, H<sup>10</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) (\*refers to the minor diastereoisomer when unambiguous distinction is possible)  $\delta$  170.22\* (C<sup>17</sup> or 18), 170.18\* (C<sup>17</sup> or 18), 170.1 (C<sup>18</sup>), 169.3 (C<sup>17</sup>), 137.9\* (C<sup>12</sup>), 137.5 (C<sup>12</sup>), 135.7 (4C, C<sup>Ar</sup>), 134.0 (C<sup>16</sup>), 133.94 (C<sup>16</sup>), 133.87\* (C<sup>16</sup>), 133.7\* (C<sup>16</sup>), 129.8 (2C, C<sup>Ar</sup>), 128.7 (2C, C<sup>13</sup>), 128.2 (C<sup>15</sup>), 128.1 (2C, C<sup>Ar</sup>), 127.9 (4C, C<sup>Ar</sup>), 98.8 (C<sup>1</sup>), 94.5\* (C<sup>1</sup>), 80.7 (C<sup>3</sup>), 80.5\* (C<sup>4</sup>), 79.3 (C<sup>4</sup>), 78.7\* (C<sup>3</sup>), 74.1 (C<sup>2</sup>), 73.3 (C<sup>7</sup>), 73.1\* (C<sup>7</sup>), 71.1\* (C<sup>2</sup>), 60.7 (C<sup>6</sup>), 60.3\* (C<sup>6</sup>), 37.6 (C<sup>5</sup>), 36.4\* (C<sup>5</sup>), 27.1 (3C, C<sup>10</sup>), 21.4\* (C<sup>9</sup>), 21.3 (C<sup>9</sup>), 21.0 (C<sup>8</sup>), 20.8\* (C<sup>8</sup>), 19.4 (C<sup>11</sup>).

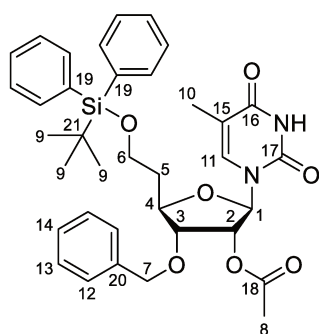
HRMS (TOF ESI+) calculated for C<sub>33</sub>H<sub>44</sub>NO<sub>7</sub>Si [M+NH<sub>4</sub>]<sup>+</sup>: 594.2887. Found 594.2899.

FTIR ( $\nu_{\max}$  cm<sup>-1</sup>) 2931, 2858, 1744, 1428, 1369, 1215, 1103, 1007, 955, 732, 699.

R<sub>f</sub> (pentane/ethyl acetate, 9/1, v/v) = 0.43 and 0.26.

II.10. (2R,3R,4R,5R)-4-(benzyloxy)-5-(2-((tert-butyl)diphenylsilyl)oxy)ethyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-3-yl acetate **229**

To a solution of diacetate **228**, mixture 8:2 of anomers, (210 mg, 0.364 mmol, 1 eq) in acetonitrile (3.0 mL) were added thymine (69 mg, 0.546 mmol, 1.5 eq) and *N,O*-bis(trimethylsilyl)acetamide (270  $\mu$ L, 1.09 mmol, 3 eq) at room temperature. The reaction mixture was stirred at 50 °C for 5 minutes. After cooling at room temperature, trimethylsilyl trifluoromethanesulfonate (72  $\mu$ L, 0.40 mmol, 1.1 eq) was added and the solution was heated at reflux for 1.5 h. After this time, the mixture was quenched with a saturated aqueous solution of sodium bicarbonate and extracted with ethyl acetate. The organic layer was washed with water, brine, dried over  $\text{Na}_2\text{SO}_4$  and concentrated under vacuum. The crude residue was purified by flash chromatography on silica gel (cyclohexane/ethyl acetate, 3/2) to afford the nucleoside **229** as white solid (203 mg, 87%).<sup>164</sup>



Chemical Formula:  $\text{C}_{36}\text{H}_{42}\text{N}_2\text{O}_7\text{Si}$   
Exact Mass: 642.28

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.85 (s, 1H, NH), 7.66-7.62 (m, 4H,  $\text{H}^{\text{Ar}}$ ), 7.41-7.25 (m, 11H,  $\text{H}^{\text{Ar}}$ ), 6.93 (d, 1H,  $J = 1.1$  Hz,  $\text{H}^{11}$ ), 5.74 (d, 1H,  $J = 3.9$  Hz,  $\text{H}^1$ ), 5.30 (dd, 1H,  $J = 5.9$  Hz, 4.0 Hz,  $\text{H}^2$ ), 4.55 (d, 1H, AB syst,  $J = 11.3$  Hz,  $\text{H}^7$ ), 4.44 (d, 1H, AB syst,  $J = 11.3$  Hz,  $\text{H}^7$ ), 4.16-4.10 (m, 1H,  $\text{H}^4$ ), 3.43 (dd, apparent t, 1H,  $J = 6.4$  Hz,  $\text{H}^3$ ), 3.80-3.76 (m, 2H,  $\text{H}^6$ ), 2.09 (s, 3H,  $\text{H}^8$ ), 2.02-1.91 (m, 1H,  $\text{H}^5$ ), 1.88 (d, 3H,  $J = 1.1$  Hz,  $\text{H}^{10}$ ), 1.87-1.76 (m, 1H,  $\text{H}^5$ ), 1.04 (s, 9H,  $\text{H}^9$ ).

$^{13}\text{C}$  { $^1\text{H}$ } NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  170.4 ( $\text{C}^{18}$ ), 163.7 ( $\text{C}^{16}$ ), 150.1 ( $\text{C}^{17}$ ), 137.5 ( $\text{C}^{20}$ ), 136.4 ( $\text{C}^{11}$ ), 135.8 (4C,  $\text{C}^{\text{Ar}}$ ), 133.84 ( $\text{C}^{19}$ ), 133.81 ( $\text{C}^{19}$ ), 129.9 (2C,  $\text{C}^{\text{Ar}}$ ), 128.7 (2C,  $\text{C}^{12}$ ), 128.3 ( $\text{C}^{14}$ ), 128.1 (2C,  $\text{C}^{\text{Ar}}$ ), 127.9 (4C,  $\text{C}^{\text{Ar}}$ ), 111.5 ( $\text{C}^{15}$ ), 89.8 ( $\text{C}^1$ ), 79.7 ( $\text{C}^3$ ), 79.0 ( $\text{C}^4$ ), 73.9 ( $\text{C}^2$ ), 73.5 ( $\text{C}^7$ ), 60.6 ( $\text{C}^6$ ), 36.3 ( $\text{C}^5$ ), 27.1 (3C,  $\text{C}^9$ ), 21.0 ( $\text{C}^8$ ), 19.4 ( $\text{C}^{21}$ ), 12.8 ( $\text{C}^{10}$ ).

HRMS (TOF ESI-) calculated for  $\text{C}_{36}\text{H}_{43}\text{N}_2\text{O}_7\text{Si}$  [M-H] $^-$ : 641.2683. Found 641.3687.

FTIR ( $\nu_{\text{max}}$   $\text{cm}^{-1}$ ) 2931, 1689, 1371, 1228, 1082, 822, 737, 700, 613.

$[\alpha]_{\text{D}}^{21} + 24.7$  (c 1.00,  $\text{CHCl}_3$ ).

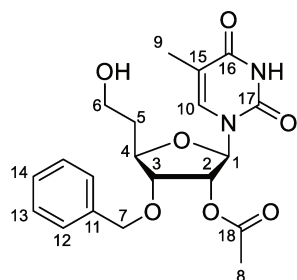
M.p. = 51-60 °C.

Rf (petroleum ether/ethyl acetate, 3/2, v/v) = 0.39.

<sup>164</sup> Mixture of anomers  $\alpha/\beta$  (ratio 6/94).

II.11. (2R,3R,4R,5R)-4-(benzyloxy)-5-(2-hydroxyethyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-3-yl acetate **230**

To a 0 °C solution of **229** (750 mg, 1.17 mmol, 1 eq) in tetrahydrofuran (8.0 mL) was added dropwise triethylamine trihydrofluoride (2.4 mL, 14.9 mmol, 12 eq). The mixture was stirred at room temperature for 17 h. The solvents were removed under reduced pressure. The crude residue was purified by flash chromatography on silica gel (dichloromethane/methanol, gradient from 100/0 to 98/2) to afford the alcohol **230** as a white solid (420 mg, 89%).<sup>165</sup>



Chemical Formula: C<sub>20</sub>H<sub>24</sub>N<sub>2</sub>O<sub>7</sub>  
Exact Mass: 404.16

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 9.35 (s, 1H, NH), 7.33-7.28 (m, 5H, H<sup>Ar</sup>), 6.99 (d, 1H, *J* = 1.1 Hz, H<sup>10</sup>), 5.60 (d, 1H, *J* = 3.0 Hz, H<sup>1</sup>), 5.42 (dd, 1H, *J* = 5.6 Hz, 3.0 Hz, H<sup>2</sup>), 4.58 (d, 1H, AB syst, *J* = 11.2 Hz, H<sup>7</sup>), 4.49 (d, 1H, AB syst, *J* = 11.2 Hz, H<sup>7</sup>), 4.17-4.11 (m, 1H, H<sup>3</sup>), 4.10-4.05 (m, 1H, H<sup>4</sup>), 3.75 (br m, 2H, H<sup>6</sup>), 2.17 (br s 1H, OH), 2.10 (s, 3H, H<sup>8</sup>), 1.95-1.86 (m, 2H, H<sup>5</sup>), 1.90 (d, 3H, *J* = 1.1 Hz, H<sup>9</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ 170.0 (C<sup>18</sup>), 163.9 (C<sup>16</sup>), 150.3 (C<sup>17</sup>), 137.4 (C<sup>10</sup>), 137.3 (C<sup>11</sup>), 128.7 (2C, C<sup>13</sup>), 128.4 (3C, C<sup>12</sup>, C<sup>14</sup>), 111.7 (C<sup>15</sup>), 91.8 (C<sup>1</sup>), 80.4 (C<sup>4</sup>), 79.5 (C<sup>3</sup>), 73.84 (C<sup>7</sup>), 73.77 (C<sup>2</sup>), 59.8 (C<sup>6</sup>), 35.2 (C<sup>5</sup>), 21.0 (C<sup>8</sup>), 19.4 (C<sup>9</sup>).

HRMS (TOF ESI-) calculated for C<sub>20</sub>H<sub>23</sub>N<sub>2</sub>O<sub>7</sub> [M-H]<sup>-</sup>: 403.1505. Found 403.1508.

FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 3658, 2981, 2899, 1682, 1374, 1228, 1048, 698, 426.

[α]<sub>D</sub><sup>22</sup> + 35.3 (c 1.00, CHCl<sub>3</sub>).

M.p. = 46-52 °C.

R<sub>f</sub> (dichloromethane/methanol, 95/5, v/v) = 0.36.

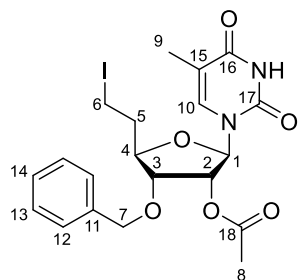
II.12. (2R,3R,4R,5R)-4-(benzyloxy)-5-(2-iodoethyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-3-yl acetate **186a**

**Method A:** To a solution of alcohol **230** (378 mg, 0.935 mmol, 1 eq) in tetrahydrofuran (6.0 mL, 0.05 M) were added imidazole (127 mg, 1.87 mmol, 2 eq), iodine (355 mg, 1.40 mmol, 1.5 eq) and triphenylphosphine (367 mg, 1.40 mmol, 1.5 eq). The reaction mixture was stirred at room temperature for 17 h. After this time, the reaction mixture was quenched with a saturated aqueous solution of Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub> and extracted with ethyl acetate. The combined organic extracts 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

<sup>165</sup> Mixture of anomers α/β (ratio 6/94).

flash chromatography on silica gel (cyclohexane/ethyl acetate, 6/4) to afford the alkyl iodine **186a** as a white solid (420 mg, 87%).<sup>166</sup>

**Method B:** According to protocol literature,<sup>134</sup> to a solution of diacetate **184**, as a 4:1 mixture of anomers, (100 mg, 0.223 mmol, 1 eq) in acetonitrile (1.8 mL) were added thymine (42 mg, 0.334 mmol, 1.5 eq) and *N,O*-bis(trimethylsilyl)acetamide (163  $\mu$ L, 0.669 mmol, 3 eq) at room temperature. The reaction mixture was stirred at 50 °C for 5 minutes. After cooling at room temperature, trimethylsilyl trifluoromethanesulfonate (44  $\mu$ L, 0.245 mmol, 1.1 eq) was added and the solution was heated at 80 °C for 1.5 h. After this time, the mixture was quenched with a saturated aqueous solution of sodium bicarbonate and extracted with ethyl acetate. The organic layer was washed with water, brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under vacuum. The crude residue was purified by flash chromatography on silica gel (cyclohexane/ethyl acetate, gradient from 7/3 to 6/4) to afford the nucleoside **186a** as white solid (91 mg, 80%).



Chemical Formula: C<sub>20</sub>H<sub>23</sub>IN<sub>2</sub>O<sub>6</sub>  
Exact Mass: 514.06

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)  $\delta$  9.14 (s, 1H, NH), 7.34-7.28 (m, 5H, H<sup>A</sup>), 6.93 (d, 1H, *J* = 1.1 Hz, H<sup>10</sup>), 5.54 (d, 1H, *J* = 3.1 Hz, H<sup>1</sup>), 5.42 (dd, 1H, *J* = 5.9 Hz, 3.1 Hz, H<sup>2</sup>), 4.57 (d, 1H, AB syst, *J* = 11.6 Hz, H<sup>7</sup>), 4.46 (d, 1H, AB syst, *J* = 11.6 Hz, H<sup>7</sup>), 4.08 (dd, 1H, *J* = 7.4 Hz, 5.9 Hz, H<sup>3</sup>), 3.98 (ddd, dt apparent, 1H, *J* = 7.9 Hz, 3.6 Hz, H<sup>4</sup>), 3.24-3.09 (m, 2H, H<sup>6</sup>), 2.21-2.04 (m, 2H, H<sup>5</sup>), 2.10 (s, 3H, H<sup>8</sup>), 1.90 (d, 3H, *J* = 1.1 Hz, H<sup>9</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>)  $\delta$  170.4 (C<sup>18</sup>), 163.9 (C<sup>16</sup>), 150.1 (C<sup>17</sup>), 137.7 (C<sup>10</sup>), 137.3 (C<sup>11</sup>), 128.8 (2C, C<sup>13</sup>), 128.8 (C<sup>14</sup>), 128.5 (2C, C<sup>12</sup>), 111.6 (C<sup>15</sup>), 92.1 (C<sup>1</sup>), 81.6 (C<sup>4</sup>), 78.9 (C<sup>3</sup>), 73.9 (C<sup>2</sup>), 73.8 (C<sup>7</sup>), 37.1 (C<sup>5</sup>), 21.0 (C<sup>8</sup>), 12.7 (C<sup>9</sup>), 0.54 (C<sup>6</sup>).

HRMS (TOF ESI-) calculated for C<sub>20</sub>H<sub>22</sub>N<sub>2</sub>O<sub>6</sub>I [M-H]<sup>-</sup>: 513.0523. Found 513.0508.

FTIR ( $\nu_{\max}$  cm<sup>-1</sup>) 3190, 3032, 2926, 1682, 1454, 1370, 1226, 1091, 898, 737, 697, 580.

[ $\alpha$ ]<sub>D</sub><sup>21</sup> + 52.0 (c 1.00, CHCl<sub>3</sub>).

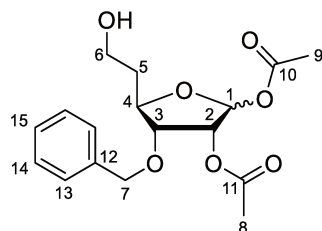
M.p. = 46-50 °C.

R<sub>f</sub> (petroleum ether/ethyl acetate, 1/1, v/v) = 0.31.

<sup>166</sup> Mixture of anomers  $\alpha/\beta$  (ratio 6/94).

II.13. (3R,4R,5R)-4-(benzyloxy)-5-(2-hydroxyethyl)tetrahydrofuran-2,3-diyl diacetate **231**

To a 0 °C solution of **228**, as a 4:1 mixture of anomers, (100 mg, 0.173 mmol, 1 eq) in tetrahydrofuran (1.2 mL) was added dropwise triethylamine trihydrofluoride (354  $\mu$ L, 2.07 mmol, 12 eq). The solution was stirred at room temperature for 43 h. The solvents were removed under reduced pressure. The crude residue, as a mixture 4:1 of anomers, determined by  $^1\text{H}$  NMR, was purified by flash chromatography on silica gel (cyclohexane/ethyl acetate, gradient from 6/4 to 5/5) to afford alcohol **231** as a colorless oil (110 mg, 50%)



Chemical Formula:  $\text{C}_{17}\text{H}_{22}\text{O}_7$   
Exact Mass: 338.14

**Characterizations realized on mixture 5:1 of pure anomers (determined by  $^1\text{H}$  NMR):**

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  (\*refers to the minor diastereoisomer when unambiguous distinction is possible) 7.33-7.26 (m, 5H,  $\text{H}^{\text{Ar}}$ ), 6.31\* (d, 0.2H,  $J = 4.5$  Hz,  $\text{H}^1$ ), 6.07 (s, 0.8H,  $\text{H}^1$ ), 5.28 (d, 0.8H,  $J = 4.2$  Hz,  $\text{H}^2$ ), 5.19\* (dd, 0.2H,  $J = 6.2$  Hz, 4.6 Hz,  $\text{H}^2$ ), 4.61 (d, 0.8H, AB syst,  $J = 11.2$  Hz,  $\text{H}^7$ ), 4.60\* (d, 0.2H, AB syst,  $J = 11.2$  Hz,  $\text{H}^7$ ), 4.46\* (d, 0.2H, AB syst,  $J = 11.2$  Hz,  $\text{H}^7$ ), 4.43 (d, 0.8H, AB syst,  $J = 11.2$  Hz,  $\text{H}^7$ ), 4.30-4.24\* (m, 0.2H,  $\text{H}^4$ ), 4.20-4.14 (m, 0.8H,  $\text{H}^4$ ), 3.99 (dd, 0.8H,  $J = 8.0$  Hz, 4.3 Hz,  $\text{H}^3$ ), 3.82\* (dd, t apparent, 0.2H,  $J = 6.1$  Hz,  $\text{H}^3$ ), 3.74-3.64 (m, 2H,  $\text{H}^6$ ), 2.20 (br s, 1H, OH), 2.09 (s, 2.4H,  $\text{H}^8$ ), 2.08\* (s, 0.6H,  $\text{H}^8$ ), 2.07\* (s, 0.6H,  $\text{H}^9$ ), 2.03 (s, 2.4H,  $\text{H}^9$ ), 1.96-1.72 (m, 2H,  $\text{H}^5$ ).

$^{13}\text{C}$  { $^1\text{H}$ } NMR (75 MHz,  $\text{CDCl}_3$ ) (\*refers to the minor diastereoisomer when unambiguous distinction is possible)  $\delta$  170.1\* ( $\text{C}^{11}$ ), 170.0\* ( $\text{C}^{11}$ ), 169.9\* ( $\text{C}^{10}$ ), 169.2 ( $\text{C}^{10}$ ), 137.5\* ( $\text{C}^{12}$ ), 137.1 ( $\text{C}^{12}$ ), 128.7 (2C,  $\text{C}^{14}$ ), 128.6\* (2C,  $\text{C}^{14}$ ), 128.3 ( $\text{C}^{15}$ ), 128.2 (2C,  $\text{C}^{13}$ ), 128.0\* (2C,  $\text{C}^{13}$ ), 98.6 ( $\text{C}^1$ ), 94.5\* ( $\text{C}^1$ ), 81.4\* ( $\text{C}^4$ ), 80.6 (2C,  $\text{C}^3$ ,  $\text{C}^4$ ), 79.0\* ( $\text{C}^3$ ), 74.43 ( $\text{C}^7$ ), 73.41 ( $\text{C}^2$ ), 73.2\* ( $\text{C}^7$ ), 70.6\* ( $\text{C}^2$ ), 60.1 ( $\text{C}^6$ ), 59.8\* ( $\text{C}^6$ ), 36.9 ( $\text{C}^5$ ), 35.9\* ( $\text{C}^5$ ), 21.2 ( $\text{C}^9$ ), 20.8 ( $\text{C}^8$ ), 20.7\* ( $\text{C}^8$ ).

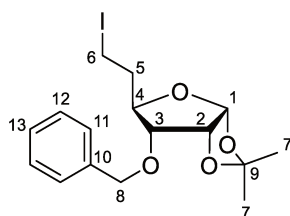
HRMS (TOF ESI+) calculated for  $\text{C}_{17}\text{H}_{26}\text{NO}_7$  [ $\text{M}+\text{NH}_4$ ] $^+$ : 356.1709. Found 356.1709.

FTIR ( $\nu_{\text{max}}$   $\text{cm}^{-1}$ ) 3473, 2941, 1741, 1370, 1215, 1097, 1010, 954, 749, 699.

R<sub>f</sub> (petroleum ether/ethyl acetate, 3/2, v/v) = 0.23 and 0.15.

II.14. (3aR,5R,6R,6aR)-6-(benzyloxy)-5-(2-iodoethyl)-2,2-dimethyltetrahydrofuro[2,3-d][1,3]dioxole **178**

Iodine (97.5 mg, 0.39 mmol, 1.2 eq) and imidazole (39 mg, 0.58 mmol, 1.8 eq) were successively added to a stirred solution of triphenylphosphine (126 mg, 0.48 mmol, 1.5 eq) in dichloromethane (3.5 mL) at room temperature. The resulting solution was stirred for 10 min before addition of a solution of alcohol **226** (93.0 mg, 0.32 mmol, 1 eq) in dichloromethane (1.0 mL). The reaction mixture was stirred 6 h at room temperature and then was quenched with addition of a saturated aqueous solution of sodium thiosulfate. The aqueous layer was extracted with dichloromethane, washed with brine, dried over MgSO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography on silica gel (petroleum ether/ethyl acetate, 9/1) to afford iodoalkyl **178** as a colorless oil (66 mg, 51%).



Chemical Formula: C<sub>16</sub>H<sub>21</sub>IO<sub>4</sub>  
Exact Mass: 404.05

**According to literature data:**<sup>107</sup>

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.39-7.29 (m, 5H, H<sup>A</sup>), 5.68 (d, 1H, *J* = 3.8 Hz, H<sup>1</sup>), 4.77 (d, 1H, AB syst, *J* = 12.0 Hz, H<sup>8</sup>), 4.54 (dd, apparent t, 1H, *J* = 4.0 Hz, H<sup>2</sup>), 4.52 (d, 1H, AB syst, *J* = 12.0 Hz, H<sup>8</sup>), 4.04 (ddd, apparent td, 1H, *J* = 8.7 Hz, 3.4 Hz, H<sup>4</sup>), 3.41 (dd, 1H, *J* = 9.0 Hz, 4.3 Hz, H<sup>3</sup>), 3.25-3.10 (m, 2H, H<sup>6</sup>), 2.29-2.17 (m, 1H, H<sup>5</sup>), 2.00-1.88 (m, 1H, H<sup>5</sup>), 1.58 (s, 3H, H<sup>7</sup>), 1.34 (s, 3H, H<sup>7</sup>).

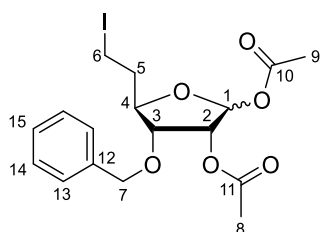
<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ 137.5 (C<sup>10</sup>), 128.7 (2C, C<sup>12</sup>), 128.3 (C<sup>13</sup>), 128.1 (2C, C<sup>11</sup>), 113.1 (C<sup>9</sup>), 104.0 (C<sup>1</sup>), 81.2 (C<sup>3</sup>), 78.0 (C<sup>4</sup>), 77.2 (C<sup>2</sup>, overlap with solvent signal and deduced from DEPT experiment), 72.2 (C<sup>8</sup>), 37.1 (C<sup>5</sup>), 26.9 (C<sup>7</sup>), 26.8 (C<sup>7</sup>), 0.7 (C<sup>6</sup>).

**R<sub>f</sub>** (petroleum ether/ethyl acetate, 9/1, v/v) = 0.20.

II.15. (3R,4R,5R)-4-(benzyloxy)-5-(2-iodoethyl)tetrahydrofuran-2,3-diyl diacetate **184**

**Method A:** To a solution of **231** (46 mg, 0.136 mmol, 1 eq) in tetrahydrofuran (3.0 mL, 0.05 M) were added imidazole (19 mg, 0.272 mmol, 2 eq), iodine (51 mg, 0.204 mmol, 1.5 eq) and triphenylphosphine (54 mg, 0.204 mmol, 1.5 eq). The reaction mixture was stirred at room temperature for 48 h. After this time, the reaction mixture was quenched with a saturated aqueous solution of sodium thiosulfate and extracted with ethyl acetate. The combined organic extracts were washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue, as a mixture 4:1 of anomers (determined by <sup>1</sup>H NMR), was purified by flash chromatography on silica gel (cyclohexane/ethyl acetate, 6/4) to afford the alkyl iodine **184** as a white solid (33 mg, 54%).

**Method B:** According to literature method,<sup>115</sup> to a solution of **178** (2.5 g, 2.62 mmol, 1 eq) in acetic acid (100.0 mL) were added anhydride acetic (23.4 mL, 247.6 mmol, 40 eq) and camphorsulfonic acid (72 mg, 0.310 mmol, 0.05 eq). The mixture was stirred at 75 °C for 1 h and a second portion of camphorsulfonic acid was then added. The reaction mixture was stirred for additional 1 h. After this time, the reaction mixture was diluted with dichloromethane, washed with water and a saturated aqueous solution of sodium bicarbonate, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue, as a mixture 4:1 of anomers, determined by <sup>1</sup>H NMR, was purified by flash chromatography on silica gel (cyclohexane/diethyl ether, gradient from 8/2 to 7/3) to afford the expected product **184** as a colorless oil (2.5 g, 90%).



Chemical Formula: C<sub>17</sub>H<sub>21</sub>IO<sub>6</sub>  
Exact Mass: 448.04

**Characterizations realized on mixture 4:1 of pure anomers (determined by <sup>1</sup>H NMR):**

**<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)** (\*refers to the minor diastereoisomer when unambiguous distinction is possible) δ 7.34-7.27 (m, 5H, H<sup>Ar</sup>), 6.30\* (d, 0.2H, *J* = 4.5 Hz, H<sup>1</sup>), 6.09 (s, 0.8H, H<sup>1</sup>), 5.26 (d, 0.8H, *J* = 4.5 Hz, H<sup>2</sup>), 5.13\* (dd, 0.2H, *J* = 6.3 Hz, 4.5 Hz, H<sup>2</sup>), 4.61\* (d, 0.2H, AB syst, *J* = 11.4 Hz, H<sup>7</sup>), 4.60 (d, 0.8H, AB syst, *J* = 11.4 Hz, H<sup>7</sup>), 4.46\* (d, 0.2H, AB syst, *J* = 11.4 Hz, H<sup>7</sup>), 4.42 (d, 0.8H, AB syst, *J* = 11.4 Hz, H<sup>7</sup>), 4.22-4.16\* (m, 0.2H, H<sup>4</sup>), 4.07 (ddd, td apparent, 0.8H, *J* = 8.1 Hz, 3.9 Hz, H<sup>4</sup>), 3.92 (dd, 0.8H, *J* = 7.8 Hz, 4.5 Hz, H<sup>3</sup>), 3.75\* (dd, 0.2H, *J* = 6.3 Hz, 5.7 Hz, H<sup>3</sup>), 3.21-3.06 (m, 2H, H<sup>6</sup>), 2.22-2.07 (m, 1H, H<sup>5</sup>), 2.10 (s, 2.4H, H<sup>8</sup>), 2.09\* (s, 0.6H, H<sup>8</sup>), 2.03 (s, 3H, H<sup>9</sup>), 1.96-1.72 (m, 1H, H<sup>5</sup>).

**<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>)** (\*refers to the minor diastereoisomer when unambiguous distinction is possible) δ 169.9 (C<sup>11</sup>), 169.0 (C<sup>10</sup>), 137.2 (C<sup>12</sup>), 128.7 (2C, C<sup>14</sup>), 128.3 (C<sup>15</sup>), 128.2\* (2C, C<sup>14</sup>), 128.1 (2C, C<sup>13</sup>), 128.0\* (2C, C<sup>13</sup>), 98.6 (C<sup>1</sup>), 94.3\* (C<sup>1</sup>), 82.7\* (C<sup>4</sup>), 81.7 (C<sup>4</sup>), 79.8 (C<sup>3</sup>), 73.6 (C<sup>2</sup>), 73.3 (C<sup>7</sup>), 73.2\* (C<sup>7</sup>), 70.8\* (C<sup>2</sup>), 38.9 (C<sup>5</sup>), 37.6\* (C<sup>5</sup>), 21.3 (C<sup>9</sup>), 20.9 (C<sup>8</sup>), 20.7\* (C<sup>8</sup>), 0.2 (C<sup>6</sup>).

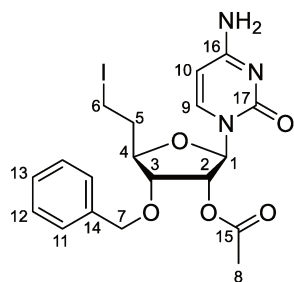
**HRMS (TOF ESI+)** calculated for C<sub>17</sub>H<sub>25</sub>NO<sub>6</sub> [M+NH<sub>4</sub>]<sup>+</sup>: 466.0727. Found 466.0738.

**FTIR (ν<sub>max</sub> cm<sup>-1</sup>)** 2928, 1742, 1369, 1213, 1179, 1123, 1098, 1034, 1009, 951, 738, 698.

**R<sub>f</sub>** (cyclohexane/diethyl ether, 3/2, v/v) = 0.37 and 0.23.

II.16. (2R,3R,4R,5R)-2-(4-amino-2-oxopyrimidin-1(2H)-yl)-4-(benzyloxy)-5-(2-iodoethyl) tetrahydrofuran-3-yl acetate **186b**

According to protocol literature,<sup>134</sup> a solution of cytosine (186 mg, 1.67 mmol, 1.5 eq) and *N,O*-bis(trimethylsilyl)acetamide (1.1 mL, 4.44 mmol, 4 eq) in acetonitrile (15.0 mL) was heated at 80 °C for 10 min. After cooling down to room temperature, a solution of diacetate **184**, as a mixture 4:1 of anomers, (500 mg, 1.11 mmol, 1 eq) in acetonitrile (17.0 mL) and trimethylsilyl trifluoromethanesulfonate (222  $\mu$ L, 1.22 mmol, 1.1 eq) were added and the mixture reaction was stirred at 80 °C for 18 h. After this time, the mixture was quenched with a saturated aqueous solution of sodium bicarbonate and extracted with ethyl acetate. The organic layer was washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under vacuum. The crude residue was purified by flash chromatography on silica gel (dichloromethane/methanol, 95/5) to afford the nucleoside **186b** as yellow solid (600 mg, 72%).



Chemical Formula: C<sub>19</sub>H<sub>22</sub>I N<sub>3</sub>O<sub>5</sub>  
Exact Mass: 499.06

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)  $\delta$  7.34-7.26 (m, 6H, H<sup>9</sup>, H<sup>A</sup>), 5.87 (d, 1H, *J* = 7.4 Hz, H<sup>10</sup>), 5.64 (d, 1H, *J* = 1.9 Hz, H<sup>1</sup>), 5.56 (dd, 1H, *J* = 4.8 Hz, 1.9 Hz, H<sup>2</sup>), 4.58 (d, 1H, AB syst, *J* = 11.4 Hz, H<sup>7</sup>), 4.41 (d, 1H, AB syst, *J* = 11.4 Hz, H<sup>7</sup>), 4.06-3.96 (m, 2H, H<sup>3</sup>, H<sup>4</sup>), 3.27-3.14 (m, 2H, H<sup>6</sup>), 2.25-2.14 (m, 2H, H<sup>5</sup>), 2.11 (s, 3H, H<sup>8</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>)  $\delta$  170.2 (C<sup>15</sup>), 166.4 (C<sup>16</sup>), 155.7 (C<sup>17</sup>), 141.8 (C<sup>9</sup>), 137.4 (C<sup>14</sup>), 128.7 (2C, C<sup>12</sup>), 128.32 (C<sup>13</sup>), 128.27 (2C, C<sup>11</sup>), 95.8 (C<sup>10</sup>), 92.5 (C<sup>1</sup>), 81.2 (C<sup>4</sup>), 79.3 (C<sup>3</sup>), 74.2 (C<sup>2</sup>), 73.6 (C<sup>7</sup>), 37.3 (C<sup>5</sup>), 21.1 (C<sup>8</sup>), 0.7 (C<sup>6</sup>).

HRMS (TOF ESI-) calculated for C<sub>19</sub>H<sub>21</sub>N<sub>3</sub>O<sub>5</sub> [M-H]<sup>-</sup>: 498.0526. Found 498.0534.

FTIR ( $\nu_{\max}$  cm<sup>-1</sup>) 3330, 3023, 1742, 1641, 1482, 1372, 1230, 1096, 742, 698, 594.

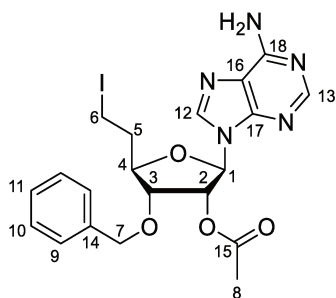
[ $\alpha$ ]<sub>D</sub><sup>22</sup> + 79.5 (c 1.00, CHCl<sub>3</sub>).

M.p. = 87-88 °C.

R<sub>f</sub> (dichloromethane/methanol, 95/5, v/v) = 0.23.

II.17. (2R,3R,4R,5R)-2-(6-amino-9H-purin-9-yl)-4-(benzyloxy)-5-(2-iodoethyl)tetrahydrofuran-3-yl acetate **186c**

According to the literature,<sup>138</sup> a mixture of adenine (754 mg, 5.58 mmol, 2 eq), trimethylsilyl chloride (3.5 mL, 27.9 mmol, 10 eq) and hexamethyldisilazane (14.0 mL) was heated at reflux temperature for 21 h. After removal volatile materials under reduced pressure, the remaining gum was mixed with diacetate **184**, as a 4:1 mixture of anomers, (1.25 g, 2.79 mmol, 1 eq), trimethylsilyl trifluoromethanesulfonate (2.0 mL, 11.2 mmol, 4 eq) and 1,2-dichloroethane (27.0 mL), and the mixture was stirred at room temperature for 17 h. After this time, the reaction mixture was quenched with a saturated aqueous solution of sodium bicarbonate and extracted with dichloromethane (3 times). The combined organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude material was purified by flash chromatography on silica gel (ethyl acetate/cyclohexane, gradient from 9/1 to 10/0) to afford the alkyl iodine **186c** as a yellow solid (976 mg, 68%).<sup>167</sup>



Chemical Formula: C<sub>20</sub>H<sub>22</sub>N<sub>5</sub>O<sub>4</sub>  
Exact Mass: 523.07

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 8.26 (s, 1H, H<sup>13</sup>), 7.79 (s, 1H, H<sup>12</sup>), 7.38-7.26 (m, 5H, H<sup>Ar</sup>), 6.34 (s, 2H, NH<sub>2</sub>), 5.97 (d, 1H, *J* = 3.3 Hz, H<sup>1</sup>), 5.89 (dd, 1H, *J* = 5.3 Hz, 3.3 Hz, H<sup>2</sup>), 4.62 (d, 1H, AB syst, *J* = 11.5 Hz, H<sup>7</sup>), 4.54 (dd, 1H, *J* = 6.7 Hz, 5.3 Hz, H<sup>3</sup>), 4.46 (d, 1H, AB syst, *J* = 11.5 Hz, H<sup>7</sup>), 4.15 (ddd, td apparent, 1H, *J* = 7.1 Hz, 4.9 Hz, H<sup>4</sup>), 3.15-3.07 (m, 2H, H<sup>6</sup>), 2.22-2.14 (m, 2H, H<sup>5</sup>), 2.10 (s, 3H, H<sup>8</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ 170.1 (C<sup>15</sup>), 156.0 (C<sup>16</sup>), 153.2 (C<sup>13</sup>), 149.4 (C<sup>17</sup>), 140.0 (C<sup>12</sup>), 137.3 (C<sup>14</sup>), 128.7 (2C, C<sup>10</sup>), 128.4 (3C, C<sup>9</sup>, C<sup>11</sup>), 120.5 (C<sup>18</sup>), 88.0 (C<sup>1</sup>), 81.9 (C<sup>4</sup>), 78.7 (C<sup>3</sup>), 74.0 (C<sup>2</sup>), 73.4 (C<sup>7</sup>), 37.0 (C<sup>5</sup>), 20.9 (C<sup>8</sup>), 0.9 (C<sup>6</sup>).

HRMS (TOF ESI+) calculated for C<sub>20</sub>H<sub>23</sub>N<sub>5</sub>O<sub>4</sub>I [M+H]<sup>+</sup>: 524.0795. Found 524.0792.

FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 3314, 3122, 1743, 1641, 1227, 1097, 698, 643.

[α]<sub>D</sub><sup>22</sup> + 25.9 (c 1.00, CHCl<sub>3</sub>).

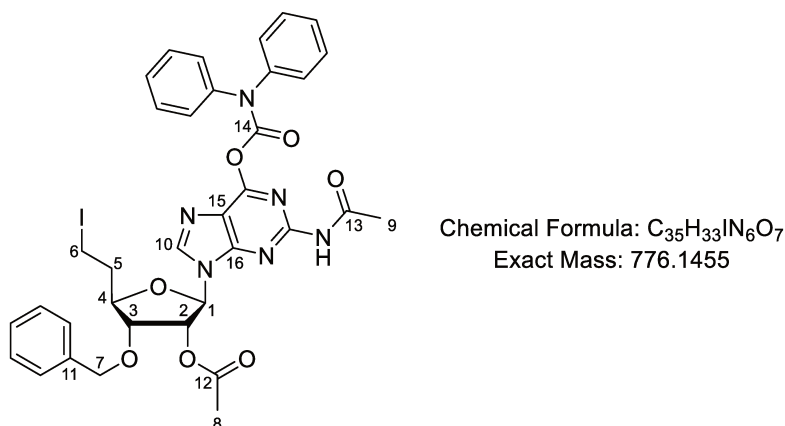
M.p. = 74-75 °C.

R<sub>f</sub> (ethyl acetate) = 0.30.

<sup>167</sup> Mixture of anomers α/β (ratio 1/99).

II.18. (2R,3R,4R,5R)-2-(2-acetamido-6-((diphenylcarbamoyl)oxy)-9H-purin-9-yl)-4-(benzyloxy)-5-(2-iodoethyl)tetrahydrofuran-3-yl acetate **186d**

Guanine was protected following a procedure described in the literature.<sup>139</sup> A solution of protected guanine (130 mg, 0.335 mmol, 1.5 eq) and *N,O*-bis(trimethylsilyl)acetamide (330  $\mu$ L, 1.34 mmol, 4 eq) in 1,2-dichloroethane (3.0 mL) was heated at 80 °C for 10 min. After cooling down to room temperature, a solution of diacetate **184** (100 mg, 0.223 mmol, 1 eq) in 1,2-dichloroethane (1.0 mL) and trimethylsilyl trifluoromethanesulfonate (44  $\mu$ L, 0.245 mmol, 1.1 eq) were added and the mixture reaction was stirred at 80 °C for 2 h. After this time, the mixture was quenched with a saturated aqueous solution of sodium bicarbonate and extracted with dichloromethane. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under vacuum. The crude residue was purified by flash chromatography on silica gel (cyclohexane/ethyl acetate, 6/4). Further purification by precipitation (dichloromethane/pentane) gave the nucleoside **186d** as white solid (95 mg, 55%).<sup>168</sup>



<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)  $\delta$  8.36 (s, 1H, NH), 7.99 (s, 1H, H<sup>10</sup>), 7.52-7.30 (m, 15H, H<sup>Ar</sup>), 5.99 (d, 1H, *J* = 3.0 Hz, H<sup>1</sup>), 5.85 (dd, 1H, *J* = 5.4 Hz, 3.0 Hz, H<sup>2</sup>), 4.93 (br m, 1H, H<sup>3</sup>), 4.77 (d, 1H, AB syst, *J* = 11.4 Hz, H<sup>7</sup>), 4.72 (d, 1H, AB syst, *J* = 11.4 Hz, H<sup>7</sup>), 4.25 (m, 1H, H<sup>4</sup>), 3.36-3.18 (m, 2H, H<sup>6</sup>), 2.42 (s, 3H, H<sup>9</sup>), 2.39-2.33 (m, 2H, H<sup>5</sup>), 2.18 (s, 3H, H<sup>8</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>)  $\delta$  170.1 (3C, C<sup>12</sup>, C<sup>13</sup>, C<sup>14</sup>), 156.4 (C<sup>16</sup>), 154.2, 152.0, 150.4, 143.5 (C<sup>10</sup>), 141.8, 137.6 (C<sup>11</sup>), 129.4 (C<sup>Ar</sup>), 128.7 (C<sup>Ar</sup>), 128.3 (C<sup>Ar</sup>), 121.9 (C<sup>15</sup>), 88.5 (C<sup>1</sup>), 82.3 (C<sup>4</sup>), 78.7 (C<sup>3</sup>), 74.2 (C<sup>2</sup>), 73.5 (C<sup>7</sup>), 36.8 (C<sup>5</sup>), 25.3 (C<sup>9</sup>), 20.9 (C<sup>8</sup>), 1.3 (C<sup>6</sup>).

HRMS calculated for C<sub>35</sub>H<sub>34</sub>N<sub>6</sub>O<sub>7</sub>I [M+H]<sup>+</sup>: 777.1534. Found 777.1533.

FTIR ( $\nu_{\max}$  cm<sup>-1</sup>) 1738, 1621, 1586, 1491, 1452, 1408, 1369, 1278, 1209, 1182, 1058, 981, 693.

[ $\alpha$ ]<sub>D</sub><sup>25</sup> + 19.1 (c 1.00, CHCl<sub>3</sub>).

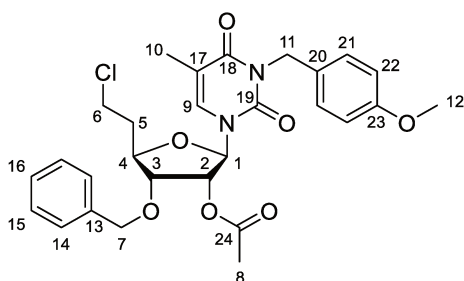
M.p. = 92-93 °C.

R<sub>f</sub> (cyclohexane/ethyl acetate, 1/1, v/v) = 0.42.

<sup>168</sup> Mixture of anomers  $\alpha/\beta$  (ratio 7/93).

II.19. (2R,3R,4R,5R)-4-(benzyloxy)-5-(2-chloroethyl)-2-(3-(4-methoxybenzyl))-5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-3-yl acetate **237**

To a solution of **186a** (50 mg, 0.124 mmol, 1 eq) in acetonitrile (1.0 mL) were added *para*-methoxybenzoyl (25  $\mu$ L, 0.186 mmol, 1.5 eq) and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) (26  $\mu$ L, 0.186 mmol, 1.5 eq). The reaction mixture was stirred at room temperature for 24 h. After this time, the reaction mixture was diluted with ethyl acetate and washed with a saturated aqueous solution of sodium bicarbonate. The aqueous layer was extracted twice with ethyl acetate. The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography on silica gel (cyclohexane/ethyl acetate, 7/3) to afford **237** (10 mg, 15%).



Chemical Formula: C<sub>28</sub>H<sub>31</sub>ClN<sub>2</sub>O<sub>7</sub>  
Exact Mass: 542.18

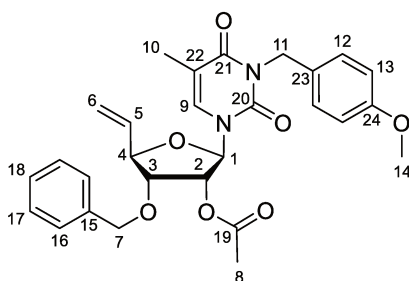
<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)  $\delta$  7.42 (d, 2H,  $J$  = 9.0 Hz, H<sup>21</sup>), 7.37-7.29 (m, 5H, H<sup>A</sup>), 6.89 (d, 1H,  $J$  = 0.9 Hz, H<sup>9</sup>), 6.82 (d, 2H,  $J$  = 9.0 Hz, H<sup>22</sup>), 5.64 (d, 1H,  $J$  = 3.0 Hz, H<sup>1</sup>), 5.41 (dd, 1H,  $J$  = 5.1 Hz, 3.0 Hz, H<sup>2</sup>), 5.05 (d, 1H, AB syst,  $J$  = 13.5 Hz, H<sup>11</sup>), 4.99 (d, 1H, AB syst,  $J$  = 13.5 Hz, H<sup>11</sup>), 4.60 (d, 1H, AB syst,  $J$  = 11.4 Hz, H<sup>7</sup>), 4.48 (d, 1H, AB syst,  $J$  = 11.4 Hz, H<sup>7</sup>), 4.09-4.02 (m, 2H, H<sup>3</sup>, H<sup>4</sup>), 3.75 (s, 3H, H<sup>12</sup>), 3.62-3.54 (m, 2H, H<sup>6</sup>), 2.16-2.05 (m, 2H, H<sup>5</sup>), 2.09 (s, 3H, H<sup>8</sup>), 1.90 (d, 3H,  $J$  = 0.9 Hz, H<sup>10</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>)  $\delta$  170.3 (C<sup>24</sup>), 163.4 (C<sup>18</sup>), 159.3 (C<sup>19</sup>), 150.8 (C<sup>20</sup>), 137.3 (C<sup>13</sup>), 135.2 (C<sup>9</sup>), 131.0 (2C, C<sup>21</sup>), 129.1 (C<sup>23</sup>), 128.8 (2C, C<sup>15</sup>), 128.4 (C<sup>16</sup>), 128.3 (2C, C<sup>14</sup>), 113.9 (2C, C<sup>22</sup>), 111.0 (C<sup>17</sup>), 92.2 (C<sup>1</sup>), 79.3 (C<sup>3</sup> or <sup>4</sup>), 78.8 (C<sup>3</sup> or <sup>4</sup>), 73.9 (C<sup>2</sup>), 73.8 (C<sup>7</sup>), 55.4 (C<sup>12</sup>), 44.1 (C<sup>11</sup>), 40.9 (C<sup>6</sup>), 36.0 (C<sup>5</sup>), 21.0 (C<sup>8</sup>), 13.5 (C<sup>10</sup>).

HRMS (TOF ESI+) calculated for C<sub>28</sub>H<sub>32</sub>N<sub>2</sub>O<sub>7</sub>Cl [M+H]<sup>+</sup>: 543.1898. Found 543.1893.

II.20. (2R,3R,4R,5R)-4-(benzyloxy)-2-(3-(4-methoxybenzyl))-5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)-5-vinyltetrahydrofuran-3-yl acetate **185c**

To a solution of nucleoside **185b** (4.66 g, 12.1 mmol, 1 eq) in acetonitrile (100.0 mL) were added 4-methoxybenzyl chloride (2.47 mL, 18.2 mmol, 1.5 eq) and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) (2.72 mL, 18.2 mmol, 1.5 eq). The reaction mixture was stirred at room temperature for 17 h. After this time, the reaction mixture was diluted with ethyl acetate and washed with a saturated aqueous solution of sodium bicarbonate. The aqueous layer was extracted twice with ethyl acetate. The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under reduced pressure. The crude residue was purified by flash chromatography on silica gel (dichloromethane) to afford the desired protected nucleoside **185c** as a gummy solid (5.32 g, 87%).<sup>169</sup>



Chemical Formula: C<sub>28</sub>H<sub>30</sub>N<sub>2</sub>O<sub>7</sub>  
Exact Mass: 506.21

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.43 (d, 2H, *J* = 8.7 Hz, H<sup>12</sup>), 7.34-7.27 (m, 5H, H<sup>Ar</sup>), 7.00 (d, 1H, *J* = 1.1 Hz, H<sup>9</sup>), 6.81 (d, 2H, *J* = 8.7 Hz, H<sup>13</sup>), 5.89 (ddd, 1H, *J* = 17.1 Hz, 10.5 Hz, 6.5 Hz, H<sup>5</sup>), 5.87 (d, 1H, *J* = 2.9 Hz, H<sup>1</sup>), 5.42 (ddd, apparent dt, 1H, *J* = 17.1 Hz, 1.1 Hz, H<sup>6</sup>), 5.37 (dd, 1H, *J* = 5.5 Hz, 3.0 Hz, H<sup>2</sup>), 5.31 (ddd, apparent dt, 1H, *J* = 10.5 Hz, 1.1 Hz, H<sup>6</sup>), 5.07 (d, 1H, AB syst, *J* = 13.5 Hz, H<sup>11</sup>), 4.99 (d, 1H, AB syst, *J* = 13.5 Hz, H<sup>11</sup>), 4.60 (d, 1H, AB syst, *J* = 11.3 Hz, H<sup>7</sup>), 4.48 (d, 1H, AB syst, *J* = 11.3 Hz, H<sup>7</sup>), 4.41 (dd, 1H, *J* = 7.4 Hz, 6.6 Hz, H<sup>4</sup>), 3.98 (dd, 1H, *J* = 7.4 Hz, 5.7 Hz, H<sup>3</sup>), 3.76 (s, 3H, H<sup>14</sup>), 2.10 (s, 3H, H<sup>8</sup>), 1.90 (d, 3H, *J* = 1.1 Hz, H<sup>10</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ 170.1 (C<sup>19</sup>), 163.4 (C<sup>21</sup>), 159.3 (C<sup>24</sup>), 150.9 (C<sup>20</sup>), 137.3 (C<sup>15</sup>), 134.4 (C<sup>5</sup>), 134.1 (C<sup>9</sup>), 131.0 (2C, C<sup>12</sup>), 129.2 (C<sup>23</sup>), 128.7 (2C, C<sup>17</sup>), 128.3 (C<sup>18</sup>), 128.2 (2C, C<sup>16</sup>), 119.3 (C<sup>6</sup>), 113.9 (2C, C<sup>13</sup>), 110.9 (C<sup>22</sup>), 90.6 (C<sup>1</sup>), 82.6 (C<sup>4</sup>), 79.9 (C<sup>3</sup>), 74.0 (C<sup>2</sup>), 73.8 (C<sup>7</sup>), 55.5 (C<sup>14</sup>), 44.2 (C<sup>11</sup>), 21.0 (C<sup>8</sup>), 13.6 (C<sup>10</sup>).

HRMS (TOF ESI+) calculated for C<sub>28</sub>H<sub>34</sub>N<sub>3</sub>O<sub>7</sub> [M+ NH<sub>4</sub>]<sup>+</sup>: 524.2397. Found 524.2397.

FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 3039, 2925, 2838, 1612, 1745, 1704, 1667, 1643, 1512, 1460.

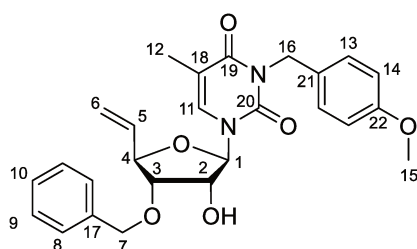
[α]<sub>D</sub><sup>21</sup> + 49.1 (c 0.90, CHCl<sub>3</sub>).

R<sub>f</sub> (dichloromethane) = 0.10.

<sup>169</sup> Mixture of anomers α/β (ratio 3/97).

II.21. 1-((2R,3R,4S,5R)-4-(benzyloxy)-3-hydroxy-5-vinyltetrahydrofuran-2-yl)-3-(4-methoxybenzyl)-5-methylpyrimidine-2,4(1H,3H)-dione **238**

The protected nucleoside **185c** (3.0 g, 5.92 mmol, 1 eq) was dissolved in methanol (15.0 mL) and potassium carbonate (410 mg, 2.96 mmol, 0.5 eq) was added. The mixture was stirred at room temperature for 18 h. After this time, the mixture was neutralized with a 1M aqueous solution of hydrochloric acid. Methanol was eliminated under reduced pressure. The aqueous phase was extracted with dichloromethane. The organic phase was washed with a saturated aqueous solution of sodium bicarbonate, dried over MgSO<sub>4</sub> and concentrated under reduced pressure to afford the expected product as a gummy solid (2.7 g, 98%) which was used without further purification.<sup>170</sup>



Chemical Formula: C<sub>26</sub>H<sub>28</sub>N<sub>2</sub>O<sub>6</sub>  
Exact Mass: 464.19

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.43 (d, 2H, *J* = 8.8 Hz, H<sup>14</sup>), 7.37-7.31 (m, 5H, H<sup>A'</sup>), 7.04 (d, 1H, *J* = 1.2 Hz, H<sup>11</sup>), 6.80 (d, 2H, *J* = 8.8 Hz, H<sup>13</sup>), 5.89 (ddd, 1H, *J* = 17.1 Hz, 10.5 Hz, 6.6 Hz, H<sup>5</sup>), 5.73 (d, 1H, *J* = 3.7 Hz, H<sup>1</sup>), 5.39 (ddd, apparent dt, 1H, *J* = 17.1 Hz, 1.2 Hz, H<sup>6</sup>), 5.29 (ddd, apparent dt, 1H, *J* = 10.5 Hz, 1.2 Hz, H<sup>6</sup>), 5.07 (d, 1H, AB syst, *J* = 13.5 Hz, H<sup>16</sup>), 4.99 (d, 1H, AB syst, *J* = 13.5 Hz, H<sup>16</sup>), 4.67 (d, 1H, AB syst, *J* = 11.5 Hz, H<sup>7</sup>), 4.62 (d, 1H, AB syst, *J* = 11.5 Hz, H<sup>7</sup>), 4.47 (dd, apparent br t, 1H, *J* = 6.2 Hz, H<sup>4</sup>), 4.21 (br dd, 1H, *J* = 9.2 Hz, 5.2 Hz, H<sup>2</sup>), 3.96 (t, 1H, *J* = 5.7 Hz, H<sup>3</sup>), 3.76 (s, 3H, H<sup>15</sup>), 3.02 (d, 1H, *J* = 5.0 Hz, OH), 1.90 (d, 3H, *J* = 1.2 Hz, H<sup>12</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ 163.4 (C<sup>19</sup>), 159.2 (C<sup>22</sup>), 151.2 (C<sup>20</sup>), 136.9 (C<sup>17</sup>), 134.9 (C<sup>11</sup>), 134.4 (C<sup>5</sup>), 130.9 (2C, C<sup>13</sup>), 129.2 (C<sup>21</sup>), 128.8 (2C, C<sup>9</sup>), 128.2 (C<sup>10</sup>), 128.0 (2C, C<sup>8</sup>), 118.9 (C<sup>6</sup>), 113.8 (2C, C<sup>14</sup>), 110.5 (C<sup>18</sup>), 92.7 (C<sup>1</sup>), 82.7 (C<sup>4</sup>), 80.7 (C<sup>3</sup>), 73.4 (C<sup>2</sup>), 73.2 (C<sup>7</sup>), 55.4 (C<sup>15</sup>), 44.1 (C<sup>16</sup>), 13.5 (C<sup>12</sup>).

HRMS calculated for C<sub>26</sub>H<sub>32</sub>N<sub>3</sub>O<sub>6</sub> [M+ NH<sub>4</sub>]<sup>+</sup>: 482.2286. Found 482.2261.

FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 3428, 2930, 1700, 1666, 1637, 1511, 1462, 1244, 1030.

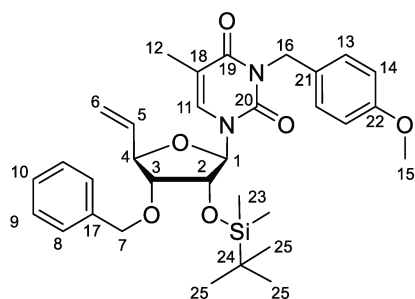
[α]<sub>D</sub><sup>24</sup> + 11.6 (c 1.21, CHCl<sub>3</sub>).

R<sub>f</sub> (pentane/diethyl ether, 7/3, v/v) = 0.27.

<sup>170</sup> Mixture of anomers α/β (ratio 4/96).

II.22. 1-((2R,3R,4R,5R)-4-(benzyloxy)-3-((tert-butyldimethylsilyl)oxy)-5-vinyltetrahydrofuran-2-yl)-3-(4-methoxybenzyl)-5-methylpyrimidine-2,4(1H,3H)-dione  
**239**

To a solution of alcohol **238** (1.52 g, 3.27 mmol, 1 eq) and Imidazole (490 mg, 7.19 mmol, 2.2 eq) in *N,N*-dimethylformamide (50.0 mL) were added *tert*-butyldimethylsilyl chloride (986 mg, 6.54 mmol, 2 eq) and 4-dimethylaminopyridine (120 mg, 0.981 mmol, 0.3 eq). The resulting mixture was stirred for 18 h at room temperature. After this time, the reaction mixture was quenched with water and extracted twice with dichloromethane. The organic layer was washed with brine, dried over MgSO<sub>4</sub> and concentrated under reduced pressure. The crude residue was purified by flash chromatography on silica gel (dichloromethane) to afford the expected product **239** as a colorless oil (1.56 g, 83%).



Chemical Formula: C<sub>32</sub>H<sub>42</sub>N<sub>2</sub>O<sub>6</sub>Si  
 Exact Mass: 578.28

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.46 (d, 2H, *J* = 8.8 Hz, H<sup>13</sup>), 7.32-7.26 (m, 5H, H<sup>A</sup>), 7.08 (d, 1H, *J* = 1.2 Hz, H<sup>11</sup>), 6.80 (d, 2H, *J* = 8.8 Hz, H<sup>14</sup>), 5.94 (ddd, 1H, *J* = 17.2 Hz, 10.5 Hz, 6.1 Hz, H<sup>5</sup>), 5.83 (d, 1H, *J* = 3.2 Hz, H<sup>1</sup>), 5.43 (ddd, apparent dt, 1H, *J* = 17.2 Hz, 1.3 Hz, H<sup>6</sup>), 5.30 (ddd, apparent dt, 1H, *J* = 10.5 Hz, 1.3 Hz, H<sup>6</sup>), 5.04 (s, 2H, H<sup>16</sup>), 4.67 (d, 1H, AB syst, *J* = 11.7 Hz, H<sup>7</sup>), 4.61 (br t, 1H, *J* = 6.5 Hz, H<sup>4</sup>), 4.48 (d, 1H, AB syst, *J* = 11.7 Hz, H<sup>7</sup>), 4.26 (dd, 1H, *J* = 4.3 Hz, 3.2 Hz, H<sup>2</sup>), 3.76 (s, 3H, H<sup>15</sup>), 3.59 (dd, 1H, *J* = 6.5 Hz, 4.6 Hz, H<sup>3</sup>), 1.89 (d, 3H, *J* = 1.1 Hz, H<sup>12</sup>), 0.86 (s, 9H, H<sup>25</sup>), 0.06 (s, 3H, H<sup>23</sup>), 0.05 (s, 3H, H<sup>23</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ 163.5 (C<sup>19</sup>), 159.3 (C<sup>22</sup>), 150.9 (C<sup>20</sup>), 137.7 (C<sup>17</sup>), 135.3 (C<sup>5</sup>), 133.6 (C<sup>11</sup>), 131.2 (2C, C<sup>13</sup>), 129.4 (C<sup>21</sup>), 128.6 (2C, C<sup>9</sup>), 128.0 (C<sup>10</sup>), 127.8 (2C, C<sup>8</sup>), 118.6 (C<sup>6</sup>), 113.8 (2C, C<sup>14</sup>), 110.2 (C<sup>18</sup>), 92.0 (C<sup>1</sup>), 82.3 (C<sup>4</sup>), 80.9 (C<sup>3</sup>), 74.5 (C<sup>2</sup>), 72.7 (C<sup>7</sup>), 55.4 (C<sup>15</sup>), 44.0 (C<sup>16</sup>), 25.9 (3C, C<sup>25</sup>), 18.3 (C<sup>24</sup>), 13.6 (C<sup>12</sup>), -4.5 (C<sup>23</sup>), -4.8 (C<sup>23</sup>).

HRMS (TOF ESI+) calculated for C<sub>32</sub>H<sub>46</sub>N<sub>3</sub>O<sub>6</sub>Si [M+ NH<sub>4</sub>]<sup>+</sup>: 596.3150. Found 596.3153.

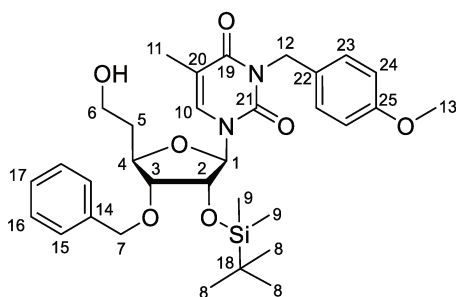
FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 2952, 2925, 2851, 2253, 1699, 1662, 1640, 1613, 1512, 1461.

[α]<sub>D</sub><sup>21</sup> + 66.8 (c 0.79, CHCl<sub>3</sub>).

R<sub>f</sub> (pentane/diethyl ether, 7/3, v/v) = 0.54.

II.23. 1-((2R,3R,4R,5R)-4-(benzyloxy)-3-((tert-butyldimethylsilyl)oxy)-5-(2-hydroxyethyl)tetrahydrofuran-2-yl)-3-(4-methoxybenzyl)-5-methylpyrimidine-2,4(1H,3H)-dione **240**

To a 0 °C solution of olefin **239** (2.32 g, 4.02 mmol, 1 eq) in tetrahydrofuran (100.0 mL) was added a solution of 9-borabicyclo[3.3.1]nonane (0.5M in THF, 16.1 mL, 8.04 mmol, 2 eq). The reaction mixture was allowed to warm up at room temperature and stirred for 18 h. The reaction mixture was cooled in an ice-water bath, quenched carefully with an aqueous solution of sodium hydroxide (60.3 mL, 2M in H<sub>2</sub>O), followed by hydrogen peroxide (5.2 mL, 35% in H<sub>2</sub>O), warmed up to 20 °C and stirred for 30 min. The aqueous layer was extracted three times with ethyl acetate. The organic extracts were washed with water, dried over Na<sub>2</sub>SO<sub>4</sub> and evaporated under reduced pressure. The crude residue was purified by flash chromatography on silica gel (pentane/ethyl acetate, gradient from 7/3 to 6/4) to afford alcohol **240** as a gummy solid (1.96 g, 82%).



Chemical Formula: C<sub>32</sub>H<sub>44</sub>N<sub>2</sub>O<sub>7</sub>Si  
Exact Mass: 596.29

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.46 (d, 2H, *J* = 8.8 Hz, H<sup>23</sup>), 7.35-7.26 (m, 5H, H<sup>Ar</sup>), 7.02 (br q, 1H, *J* = 2.2 Hz, 1.1 Hz, H<sup>10</sup>), 6.80 (d, 2H, *J* = 8.8 Hz, H<sup>24</sup>), 5.71 (d, 1H, *J* = 3.0 Hz, H<sup>1</sup>), 5.03 (s, 2H, H<sup>12</sup>), 4.70 (d, 1H, AB syst, *J* = 11.3 Hz, H<sup>7</sup>), 4.42 (d, 1H, AB syst, *J* = 11.3 Hz, H<sup>7</sup>), 4.35 (dd, 1H, *J* = 4.7 Hz, 3.0 Hz, H<sup>2</sup>), 4.26 (ddd, apparent br dd, 1H, *J* = 12.9 Hz, 7.0 Hz, H<sup>4</sup>), 3.81-3.72 (m, 2H, H<sup>6</sup>), 3.75 (br s, 3H, H<sup>13</sup>), 3.60 (dd, 1H, *J* = 7.0 Hz, 4.7 Hz, H<sup>3</sup>), 2.10 (t, 1H, *J* = 5.9 Hz, OH), 1.96-1.93 (m, 2H, H<sup>5</sup>), 1.90 (d, 3H, *J* = 1.1 Hz, H<sup>11</sup>), 0.87 (s, 9H, H<sup>8</sup>), 0.07 (s, 3H, H<sup>9</sup>), 0.06 (s, 3H, H<sup>9</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ 163.4 (C<sup>19</sup>), 159.3 (C<sup>25</sup>), 150.9 (C<sup>21</sup>), 137.4 (C<sup>14</sup>), 133.9 (C<sup>10</sup>), 131.2 (2C, C<sup>23</sup>), 129.4 (C<sup>22</sup>), 128.7 (2C, C<sup>16</sup>), 128.2 (C<sup>17</sup>), 127.9 (2C, C<sup>15</sup>), 113.8 (2C, C<sup>24</sup>), 110.4 (C<sup>20</sup>), 92.9 (C<sup>1</sup>), 81.0 (C<sup>3</sup>), 80.8 (C<sup>4</sup>), 73.8 (C<sup>2</sup>), 72.7 (C<sup>7</sup>), 60.4 (C<sup>6</sup>), 55.5 (C<sup>13</sup>), 44.0 (C<sup>12</sup>), 36.2 (C<sup>5</sup>), 25.9 (3C, C<sup>8</sup>), 18.2 (C<sup>18</sup>), 13.6 (C<sup>11</sup>), -4.3 (C<sup>9</sup>), -4.9 (C<sup>9</sup>).

HRMS (TOF ESI+) calculated for C<sub>32</sub>H<sub>45</sub>N<sub>2</sub>O<sub>7</sub>Si [M+H]<sup>+</sup>: 597.2991. Found 597.3003.

FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 3457, 2982, 2856, 1701, 1666, 1638, 1512, 1462, 1246, 1056, 832, 775.

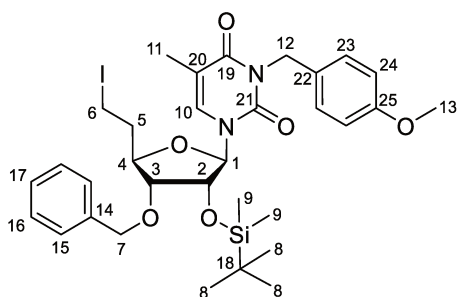
[α]<sub>D</sub><sup>24</sup> + 56.9 (c 1.26, CHCl<sub>3</sub>).

R<sub>f</sub><sub>product</sub> (petroleum ether/ethyl acetate, 1/1, v/v) = 0.42.

R<sub>f</sub><sub>intermediate</sub> (petroleum ether/ethyl acetate, 1/1, v/v) = 0.26.

II.24. 1-((2R,3R,4R,5R)-4-(benzyloxy)-3-((tert-butyl dimethylsilyl)oxy)-5-(2-iodoethyl) tetrahydrofuran-2-yl)-3-(4-methoxybenzyl)-5-methylpyrimidine-2,4(1H,3H)-dione **186f**

To a solution of alcohol **240** (1.9 g, 3.19 mmol, 1 eq) in tetrahydrofuran (70.0 mL, 0.05 M) were added imidazole (434 mg, 6.37 mmol, 2 eq), iodine (1.21 g, 4.78 mmol, 1.5 eq) and triphenylphosphine (1.25 g, 4.78 mmol, 1.5 eq). The reaction mixture was stirred at room temperature for 17 h. After this time, the reaction mixture was quenched with a saturated aqueous solution of sodium thiosulfate and extracted with ethyl acetate. The combined organic extracts 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 on silica gel (pentane/diethyl ether, gradient from 9/1 to 6/4) to afford the alkyl iodide **186f** as a white solid (1.97 g, 87%).



Chemical Formula: C<sub>32</sub>H<sub>43</sub>IN<sub>2</sub>O<sub>6</sub>Si  
Exact Mass: 706.19

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.45 (d, 2H, *J* = 8.7 Hz, H<sup>23</sup>), 7.34-7.28 (m, 5H, H<sup>Ar</sup>), 6.91 (br q, 1H, *J* = 2.2 Hz, 1.1 Hz, H<sup>10</sup>), 6.82-6.79 (d, 2H, *J* = 8.7 Hz, H<sup>24</sup>), 5.65 (d, 1H, *J* = 4.2 Hz, H<sup>1</sup>), 5.03 (s, 2H, H<sup>12</sup>), 4.71 (d, 1H, AB syst, *J* = 11.9 Hz, H<sup>7</sup>), 4.46 (d, 1H, AB syst, *J* = 11.9 Hz, H<sup>7</sup>), 4.43 (dd, apparent t, 1H, *J* = 4.5 Hz, H<sup>2</sup>), 4.15 (ddd, apparent dt, 1H, *J* = 8.4 Hz, 5.2 Hz, H<sup>4</sup>), 3.76 (s, 3H, H<sup>13</sup>), 3.58 (dd, apparent t, 1H, *J* = 5.2 Hz, H<sup>3</sup>), 3.27-3.10 (m, 2H, H<sup>6</sup>), 2.20-2.11 (m, 2H, H<sup>5</sup>), 1.91 (d, 3H, *J* = 1.1 Hz, H<sup>11</sup>), 0.85 (s, 9H, H<sup>8</sup>), 0.04 (s, 3H, H<sup>9</sup>), -0.01 (s, 3H, H<sup>9</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ 163.4 (C<sup>19</sup>), 159.3 (C<sup>25</sup>), 150.8 (C<sup>21</sup>), 137.6 (C<sup>14</sup>), 134.7 (C<sup>10</sup>), 131.2 (2C, C<sup>23</sup>), 129.3 (C<sup>22</sup>), 128.7 (2C, C<sup>16</sup>), 128.2 (C<sup>17</sup>), 127.9 (2C, C<sup>15</sup>), 113.8 (2C, C<sup>24</sup>), 110.5 (C<sup>20</sup>), 93.4 (C<sup>1</sup>), 81.8 (C<sup>4</sup>), 80.4 (C<sup>3</sup>), 73.6 (C<sup>2</sup>), 72.6 (C<sup>7</sup>), 55.5 (C<sup>13</sup>), 44.0 (C<sup>12</sup>), 38.0 (C<sup>5</sup>), 25.9 (3C, C<sup>8</sup>), 18.2 (C<sup>18</sup>), 13.6 (C<sup>11</sup>), 0.5 (C<sup>6</sup>), -4.5 (C<sup>9</sup>), -4.8 (C<sup>9</sup>).

HRMS (TOF ESI+) calculated for C<sub>32</sub>H<sub>44</sub>IN<sub>2</sub>O<sub>6</sub>Si [M+H]<sup>+</sup>: 707.2008. Found 707.2005.

FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 3073, 2932, 2858, 1745, 1704, 1643, 1667, 1612, 1512, 1460.

[α]<sub>D</sub><sup>21</sup> + 33.8 (c 1.35, CHCl<sub>3</sub>).

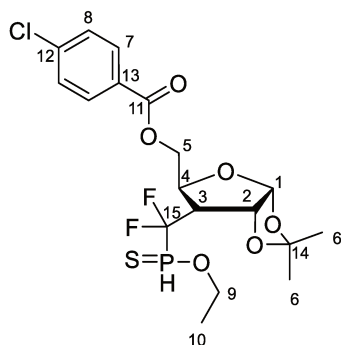
M.p. = 40-42 °C.

R<sub>f</sub><sub>product</sub> (petroleum ether/ethyl acetate, 9/1, v/v) = 0.19.

## III. Characterizations chapter 3

III.1. 5-(4-chlorobenzoyl)-3-desoxy-3-(ethyl difluoromethyl-H-phosphinothioate)-1,2-O-isopropylidene- $\alpha$ -D-ribofuranose **165d**<sup>107</sup>

To a solution of crude phosphinate **165c** in toluene (9.0 mL) was added Lawesson reagent (388 mg, 0.96 mmol, 0.55 eq). The reaction mixture was stirred at 85 °C for 45 min. After this time, the reaction was concentrated. The crude residue was purified by flash chromatography on silica gel (pentane/ethyl acetate, gradient from 9/1 to 8/2). Further purification was realized by flash chromatography on silica gel (pentane/diethyl ether, gradient from 9/1 to 8/2) to afford *H*-phosphinothioate **165d** as a colorless oil (298 mg, 36%) and as a 1:1 mixture of diastereoisomers (P center), determined by <sup>31</sup>P NMR from crude residue. For analytical purposes, a pure fraction of both diastereoisomers were isolated and characterized separately.



Chemical Formula: C<sub>18</sub>H<sub>22</sub>ClF<sub>2</sub>O<sub>6</sub>PS  
Exact Mass: 470.05

**Diastereoisomer A:**

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)  $\delta$  7.95 (d, 2H,  $J$  = 8.6 Hz, H<sup>7</sup>), 7.88 (dd, 1H,  $J$  = 588.4 Hz, 11.5 Hz, PH), 7.38 (d, 2H,  $J$  = 8.6 Hz, H<sup>8</sup>), 5.90 (d, 1H,  $J$  = 3.9 Hz, H<sup>1</sup>), 5.07 (dd, apparent t, 1H,  $J$  = 4.7 Hz, H<sup>2</sup>), 4.78-4.72 (m, 2H, H<sup>4</sup>, H<sup>5</sup>), 4.36-4.25 (m, 2H, H<sup>5</sup>, H<sup>9</sup>), 4.23-4.13 (m, 1H, H<sup>9</sup>), 2.99-2.83 (m, 1H, H<sup>3</sup>), 1.51 (s, 3H, H<sup>6</sup>), 1.34 (t, 3H,  $J$  = 7.0 Hz, H<sup>10</sup>), 1.30 (s, 3H, H<sup>6</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>)  $\delta$  165.3 (C<sup>11</sup>), 139.8 (C<sup>13</sup>), 131.3 (2C, C<sup>7</sup>), 128.9 (2C, C<sup>8</sup>), 128.3 (C<sup>12</sup>), 121.3 (ddd, <sup>1</sup>J<sub>C-F</sub> = 281.9 Hz, <sup>1</sup>J<sub>C-F</sub> = 265.1 Hz, <sup>1</sup>J<sub>C-P</sub> = 107.5 Hz, C<sup>15</sup>), 113.3 (C<sup>14</sup>), 105.4 (C<sup>1</sup>), 79.0 (d, <sup>2</sup>J<sub>C-F</sub> = 9.3 Hz, C<sup>2</sup>), 74.7 (d, <sup>2</sup>J<sub>C-F</sub> = 6.4 Hz, C<sup>4</sup>), 64.6 (C<sup>5</sup>), 63.2 (d, <sup>2</sup>J<sub>C-P</sub> = 7.1 Hz, C<sup>9</sup>), 48.8 (dt, <sup>2</sup>J<sub>C-P</sub> = 22.1 Hz, <sup>2</sup>J<sub>C-F</sub> = 17.3 Hz, C<sup>3</sup>), 26.7 (C<sup>6</sup>), 26.5 (C<sup>6</sup>), 16.3 (d, <sup>2</sup>J<sub>C-P</sub> = 6.6 Hz, C<sup>10</sup>).

<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)  $\delta$  -106.34 (dddd, apparent ddt, 1F, <sup>2</sup>J<sub>F-F</sub> = 286.4 Hz, <sup>2</sup>J<sub>F-P</sub> = 90.5 Hz, <sup>3</sup>J<sub>F-H</sub> = 11.7 Hz), -108.34 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 286.4 Hz, <sup>2</sup>J<sub>F-P</sub> = 91.5 Hz, <sup>3</sup>J<sub>F-H</sub> = 20.5 Hz).

<sup>31</sup>P NMR (121 MHz, CDCl<sub>3</sub>)  $\delta$  61.1 (dddt, apparent dtt, <sup>1</sup>J<sub>P-H</sub> = 588.4 Hz, <sup>2</sup>J<sub>P-F</sub> = 91.2 Hz, <sup>2</sup>J<sub>P-C</sub> = 10.9 Hz).

HRMS (TOF ESI-) calculated for C<sub>18</sub>H<sub>21</sub>F<sub>2</sub>O<sub>6</sub>PSCI [M-H]<sup>-</sup>: 469.0631. Found 469.0626.

FTIR ( $\nu_{\max}$  cm<sup>-1</sup>) 2988, 1723, 1595, 1489, 1455, 1402, 1384, 1269.5, 1217.1, 1170, 1091.3, 1063, 1026, 1014, 757, 662, 528, 467, 464, 414.

$[\alpha]_{\text{D}}^{21} + 35.3$  (c 1.00, CHCl<sub>3</sub>).

R<sub>f</sub>A (pentane/ethyl acetate, 85/15, v/v) = 0.71.

**Diastereoisomer B:**

**<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)** δ 7.95 (d, 2H, *J* = 8.6 Hz, H<sup>7</sup>), 7.85 (ddd, 1H, *J* = 595.9 Hz, 9.6 Hz, 1.4 Hz, PH), 7.38 (d, 2H, *J* = 8.6 Hz, H<sup>8</sup>), 5.88 (d, 1H, *J* = 3.8 Hz, H<sup>1</sup>), 5.07 (m, 1H, H<sup>4</sup>), 4.86 (dd, apparent t, 1H, *J* = 4.1 Hz, H<sup>2</sup>), 4.74 (br dd, 1H, AB syst, *J* = 12.6 Hz, 1.4 Hz, H<sup>5</sup>), 4.40 (dd, 1H, AB syst, *J* = 12.6 Hz, 4.5 Hz, H<sup>5</sup>), 4.36-4.26 (m, 1H, H<sup>9</sup>), 4.21-4.11 (m, 1H, H<sup>9</sup>), 2.95-2.77 (m, 1H, H<sup>3</sup>), 1.60 (s, 3H, H<sup>6</sup>), 1.33 (t, 3H, *J* = 7.1 Hz, H<sup>10</sup>), 1.31 (s, 3H, H<sup>6</sup>).

**<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>)** δ 165.4 (C<sup>11</sup>), 139.8 (C<sup>13</sup>), 131.3 (2C, C<sup>7</sup>), 128.9 (2C, C<sup>8</sup>), 128.5 (C<sup>12</sup>), 120.4 (ddd, apparent td, <sup>1</sup>*J*<sub>C-F</sub> = 269.5 Hz, <sup>1</sup>*J*<sub>C-P</sub> = 107.5 Hz, C<sup>15</sup>), 113.5 (C<sup>14</sup>), 104.8 (C<sup>1</sup>), 79.0 (dd, <sup>2</sup>*J*<sub>C-P</sub> = 5.8 Hz, 3.3 Hz, C<sup>2</sup>), 74.7 (d, <sup>2</sup>*J*<sub>C-F</sub> = 4.2 Hz, C<sup>4</sup>), 63.7 (C<sup>5</sup>), 63.0 (d, <sup>2</sup>*J*<sub>C-P</sub> = 7.1 Hz, C<sup>9</sup>), 50.7 (dt, <sup>2</sup>*J*<sub>C-P</sub> = 21.8 Hz, <sup>2</sup>*J*<sub>C-F</sub> = 13.6 Hz, C<sup>3</sup>), 26.6 (C<sup>6</sup>), 26.4 (C<sup>6</sup>), 16.3 (d, <sup>2</sup>*J*<sub>C-P</sub> = 6.4 Hz, C<sup>10</sup>).

**<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)** δ -97.6 (ddd, 1F, <sup>2</sup>*J*<sub>F-F</sub> = 303.8 Hz, <sup>2</sup>*J*<sub>F-P</sub> = 99.7 Hz, <sup>3</sup>*J*<sub>F-H</sub> = 9.8 Hz), -109.7 (dddd, 1F, <sup>2</sup>*J*<sub>F-F</sub> = 303.8 Hz, <sup>2</sup>*J*<sub>F-P</sub> = 94.7 Hz, <sup>3</sup>*J*<sub>F-H</sub> = 18.8 Hz, 9.6 Hz).

**<sup>31</sup>P NMR (121 MHz, CDCl<sub>3</sub>)** δ 62.9 (dm, <sup>1</sup>*J*<sub>P-H</sub> = 595.9 Hz).

**HRMS (TOF ESI-)** calculated for C<sub>18</sub>H<sub>21</sub>F<sub>2</sub>O<sub>7</sub>PSCI [M-H+O]<sup>-</sup>: 485.0402. Found 485.0406.

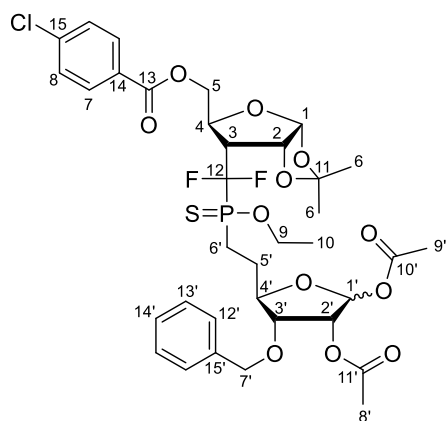
**FTIR (ν<sub>max</sub> cm<sup>-1</sup>)** 2987, 1722, 1594, 1489, 1455, 1402, 1384, 1270, 1216, 1170, 1091, 1036, 1012, 926, 905, 873, 849, 755, 659, 528, 478.

**[α]<sub>D</sub><sup>21</sup>** + 56.2 (c 1.00, CHCl<sub>3</sub>).

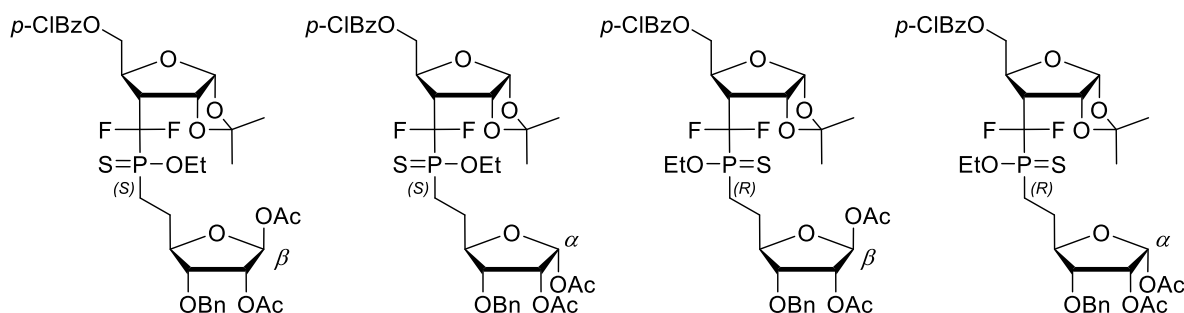
**R<sub>f</sub>** (pentane/ethyl acetate, 85/15, v/v) = 0.50.

III.2. (3R,4R,5R)-4-(benzyloxy)-5-(2-(((3aR,5S,6R,6aR)-5-(((4-chlorobenzoyl)oxy)methyl)-2,2-dimethyltetrahydrofuro[2,3-d][1,3]dioxol-6-yl)difluoromethyl)(ethoxy)phosphorothioyl)ethyl)tetrahydrofuran-2,3-diyl diacetate **196**

To a suspension of potassium hydroxide (30 mg, 0.542 mmol, 1 eq) and benzyltriethylammonium chloride (61 mg, 0.271 mmol, 0.5 eq) in dichloromethane (5.0 mL) was added dropwise, at room temperature, a solution of *H*-phosphinothioate **165d**, ratio 1:1 (A:B) determined by <sup>31</sup>P NMR, (300 mg, 0.542 mmol, 1 eq) in dichloromethane (5.0 mL), followed by a solution of alkyl iodide **184** (243 mg, 0.542 mmol, 1 eq) in dichloromethane (7.5 mL). The reaction mixture was stirred at 40 °C for 26 h. After cooling down to room temperature, the reaction solution was quenched with a saturated aqueous solution of ammonium chloride. The separated aqueous layer was extracted with dichloromethane. 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 on silica gel (pentane/diethyl ether, gradient from 8/2 to 1/1) followed by a purification on Interchim Puriflash® (method A) to afford the expected product as a white solid (A+B: 20 mg, C+D: 11 mg, 7%). Three diastereoisomers were characterized.



Chemical Formula:  $C_{35}H_{42}ClF_2O_{12}PS$   
Exact Mass: 790.18



**Diastereoisomers A+B (ratio A:B in  $^{31}P \{^1H\}$  - 1:1):<sup>171</sup>**

$^1H$  NMR (300 MHz,  $CDCl_3$ )  $\delta$  7.97 (2d, 2H,  $J = 8.7$  Hz,  $H^7$ ), 7.40 (2d, 2H,  $J = 8.7$  Hz,  $H^8$ ), 7.35-7.26 (m, 5H,  $H^{Ar}$ ), 6.30 (d, 0.5H,  $J = 4.5$  Hz,  $H^{1'}$ ), 6.09 (s, 0.5H,  $H^{1'}$ ), 5.86 (d, 0.5H,  $J = 3.6$  Hz,  $H^1$ ), 5.85 (d, 0.5H,  $J = 3.9$  Hz,  $H^1$ ), 5.28 (d, 0.5H,  $J = 4.2$  Hz,  $H^2$ ), 5.15 (dd, 0.5H,  $J = 6.3$  Hz, 4.5 Hz,  $H^2$ ), 4.97 (dd, apparent t, 0.9H,  $J = 3.9$  Hz,  $H^2$ ), 4.89 (dd, apparent t, 0.1H,  $J = 4.2$  Hz,  $H^2$ ), 4.82-4.68 (m, 2H,  $H^4$ ,  $H^5$ ), 4.62 (d, 0.5H, AB syst,  $J = 11.7$  Hz,  $H^{7'}$ ), 4.60 (d, 0.5H, AB syst,  $J = 11.4$  Hz,  $H^{7'}$ ), 4.48 (d, 0.5H, AB syst,  $J = 11.7$  Hz,  $H^{7'}$ ), 4.43 (d, 0.5H, AB syst,  $J = 11.4$  Hz,  $H^{7'}$ ), 4.35-4.01 (m, 4H,  $H^{4'}$ ,  $H^5$ ,  $H^9$ ), 3.93 (dd, 0.5H,  $J = 7.8$  Hz, 4.5 Hz,  $H^{3'}$ ), 3.77 (dd, 0.5H,  $J = 6.3$  Hz, 6.0 Hz,  $H^{3'}$ ), 3.28-2.99 (m, 1H,  $H^3$ ), 2.36-2.14 (m, 2H,  $H^6$ ), 2.11-2.10 (3s, 3H,  $H^8$ ,  $H^9$ ), 2.09 (s, 1.5H,  $H^8$ ,  $H^9$ ), 2.05 (s, 1.1H,  $H^9$ ), 2.04 (s, 0.4H,  $H^9$ ), 2.00-1.80 (m, 2H,  $H^5$ ), 1.55 (s, 1.7H,  $H^6$ ), 1.49 (s, 1.3H,  $H^6$ ), 1.33-1.25 (m, 6H,  $H^6$ ,  $H^{10}$ ).

$^{13}C \{^1H\}$  NMR (75 MHz,  $CDCl_3$ )  $\delta$  " $C^{12}$ , unobserved due to multiplicity" 170.1 ( $C^{11}$ ), 170.0 ( $C^{11}$ ), 169.9 ( $C^{10}$ ), 169.2 ( $C^{10}$ ), 165.5 ( $C^{13}$ ), 164.4 ( $C^{13}$ ), 139.8 ( $C^{14}$ ), 137.5 ( $C^{15}$ ), 137.2 ( $C^{15}$ ), 131.4 (2C,  $C^7$ ), 129.0 (2C,  $C^8$ ), 128.8 (2C,  $C^{13'}$ ), 128.5 ( $C^{15}$ ), 128.4 ( $C^{14'}$ ), 128.3 ( $C^{14'}$ ), 128.2 (2C,  $C^{12'}$ ), 128.0 (2C,  $C^{12'}$ ), 113.4 ( $C^{11}$ ), 105.3 ( $C^1$ ), 98.6 ( $C^1$ ), 94.4 ( $C^1$ ), 82.5 (d,  $^3J_{C-P} = 17.3$  Hz,  $C^4$ ), 81.4 (d,  $^3J_{C-P} = 18.0$  Hz,  $C^4$ ), 80.3 ( $C^3$ ), 79.1 (d,  $^3J_{C-P} = 10.4$  Hz,  $C^2$ ), 78.8 (d,  $^3J_{C-P} = 9.6$  Hz,  $C^2$ ), 78.4 ( $C^3$ ), 75.0 (br dd,  $^3J_{C-P} = 5.8$  Hz, 3.8 Hz,  $C^4$ ), 73.8 ( $C^2$ ), 73.5 ( $C^7$ ), 73.3 ( $C^7$ ), 71.0 ( $C^2$ ), 64.9 (d,  $^4J_{C-P} = 3.3$  Hz,  $C^5$ ), 64.7 (d,  $^4J_{C-P} = 3.8$  Hz,  $C^5$ ), 63.7 (d,  $^2J_{C-P} = 6.5$  Hz,  $C^9$ ), 63.7 (d,  $^2J_{C-P} = 7.4$  Hz,  $C^9$ ), 48.6-48.1 (m,  $C^3$ ), 46.9-46.6 (m,  $C^3$ ), 28.0 (d,  $^1J_{C-P} = 75.0$  Hz,  $C^6$ ), 26.9 (d,  $^1J_{C-P} = 72.4$  Hz,  $C^6$ ), 27.0 ( $C^6$ ), 26.8 ( $C^6$ ), 26.7 ( $C^6$ ), 26.6 ( $C^6$ ), 25.4 ( $C^5$ ), 21.32 ( $C^9$ ), 21.28 ( $C^9$ ), 20.9 ( $C^8$ ), 20.8 ( $C^8$ ), 16.6 (d,  $^3J_{C-P} = 6.2$  Hz,  $C^{10}$ ), 16.5 (d,  $^3J_{C-P} = 6.6$  Hz,  $C^{10}$ ).

<sup>171</sup> An impurity was observed with diastereoisomers A+B in  $^{19}F \{^1H\}$  and  $^{31}P \{^1H\}$  NMR (ratio 7:6:1, A:B:impurity), with chemical shifts:

$^{19}F \{^1H\}$  NMR:  $\delta$  -95.3 (dd, 1F,  $^2J_{F-F} = 281.7$  Hz,  $^2J_{F-P} = 66.3$  Hz), -101.7 (dd, 1F,  $^2J_{F-F} = 281.7$  Hz,  $^2J_{F-P} = 93.3$  Hz).

$^{31}P \{^1H\}$  NMR:  $\delta$  147.5 (dd,  $^2J_{P-F} = 93.3$  Hz, 66.3 Hz).

**<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)** *Diastereoisomer A*:  $\delta$  -100.2 (ddd, 1F,  $^2J_{F-F}$  = 289.3 Hz,  $^2J_{F-P}$  = 81.2 Hz,  $^3J_{F-H}$  = 8.7 Hz), -111.4 (ddd, 1F,  $^2J_{F-F}$  = 293.3 Hz,  $^2J_{F-P}$  = 86.5 Hz,  $^3J_{F-H}$  = 24.0 Hz).  
*Diastereoisomer B*:  $\delta$  -99.2 (ddd, 1F,  $^2J_{F-F}$  = 290.2 Hz,  $^2J_{F-P}$  = 108.5 Hz,  $^3J_{F-H}$  = 6.2 Hz), -113.1 (ddd, 1F,  $^2J_{F-F}$  = 290.2 Hz,  $^2J_{F-P}$  = 78.1 Hz,  $^3J_{F-H}$  = 25.9 Hz).

**<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>)** *Diastereoisomer A*:  $\delta$  91.0 (dd,  $^2J_{P-F}$  = 86.5 Hz,  $^2J_{P-F}$  = 81.2 Hz).  
*Diastereoisomer B*:  $\delta$  91.6 (dd,  $^2J_{P-F}$  = 108.5 Hz,  $^2J_{P-F}$  = 78.1 Hz).

**HRMS** calculated for C<sub>35</sub>H<sub>42</sub>F<sub>2</sub>NaO<sub>12</sub>PS<sup>35</sup>Cl [M+Na]<sup>+</sup>: 813.1689. Found 813.1696. Calculated for C<sub>35</sub>H<sub>42</sub>F<sub>2</sub>NaO<sub>12</sub>PS<sup>37</sup>Cl [M+Na]<sup>+</sup>: 815.1615. Found 815.1658.

**FTIR ( $\nu_{\max}$  cm<sup>-1</sup>)** 2990, 2930, 1724, 1594, 1372, 1216, 1170, 1091, 1012, 758, 699, 628, 457.

**Rf<sub>A+B</sub>** (pentane/diethyl ether, 1/1, v/v) = 0.44.

**Diastereoisomers C+D (ratio C:D in <sup>31</sup>P {<sup>1</sup>H} - 9:1):**

**<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)** (\*refers to the diastereoisomer B when unambiguous distinction is possible)  
 $\delta$  7.97 (d, 2H,  $J$  = 8.7 Hz, H<sup>7</sup>), 7.40 (d, 2H,  $J$  = 8.7 Hz, H<sup>8</sup>), 7.35-7.28 (m, 5H, H<sup>A</sup>), 6.09 (s, 1H, H<sup>1</sup>), 5.83 (d, 1H,  $J$  = 3.6 Hz, H<sup>1</sup>), 5.29 (d, 0.9H,  $J$  = 4.5 Hz, H<sup>2</sup>), 5.28\* (d, 0.1H,  $J$  = 5.4 Hz, H<sup>2</sup>), 5.00\* (dd, apparent t, 0.1H,  $J$  = 3.9 Hz, H<sup>2</sup>), 4.94 (dd, apparent t, 0.9H,  $J$  = 3.9 Hz, H<sup>2</sup>), 4.79 (d, 1H,  $J$  = 12.3 Hz, H<sup>5</sup>), 4.70 (br dd, 1H,  $J$  = 11.1 Hz, 4.8 Hz, H<sup>4</sup>), 4.62 (d, 1H, AB syst,  $J$  = 11.4 Hz, H<sup>7</sup>), 4.44 (d, 1H, AB syst,  $J$  = 11.4 Hz, H<sup>7</sup>), 4.33-4.21 (m, 2H, H<sup>5</sup>, H<sup>9</sup>), 4.10-4.03 (m, 2H, H<sup>4</sup>, H<sup>9</sup>), 3.94 (dd, 1H,  $J$  = 8.7 Hz, 4.5 Hz, H<sup>3</sup>), 3.36-2.99 (m, 1H, H<sup>3</sup>), 2.40-2.18 (m, 2H, H<sup>6</sup>), 2.11 (s, 3H, H<sup>8</sup>), 2.05 (s, 2.7H, H<sup>9</sup>), 2.04\* (s, 0.3H, H<sup>9</sup>), 2.04-1.97 (m, 2H, H<sup>5</sup>), 1.55 (s, 1H, H<sup>6</sup>), 1.48 (s, 2H, H<sup>6</sup>), 1.33-1.25 (m, 6H, H<sup>6</sup>, H<sup>10</sup>).

**<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>)**  $\delta$  "C<sup>12</sup>, unobserved due to multiplicity" 170.0 (C<sup>11</sup>), 169.1 (C<sup>10</sup>), 165.5 (C<sup>13</sup>), 139.8 (C<sup>14</sup>), 137.3 (C<sup>15</sup>), 131.4 (2C, C<sup>7</sup>), 128.9 (2C, C<sup>8</sup>), 128.8 (2C, C<sup>13</sup>), 128.5 (C<sup>15</sup>), 128.4 (C<sup>14</sup>), 128.1 (2C, C<sup>12</sup>), 113.5 (C<sup>11</sup>), 105.3 (C<sup>1</sup>), 98.6 (C<sup>1</sup>), 81.6 (d,  $^3J_{C-P}$  = 17.3 Hz, C<sup>4</sup>), 80.4 (C<sup>3</sup>), 79.1 (d,  $^3J_{C-P}$  = 10.1 Hz, C<sup>2</sup>), 75.1 (C<sup>4</sup>), 73.9 (C<sup>2</sup>), 73.4 (C<sup>7</sup>), 64.9 (d,  $^4J_{C-P}$  = 3.2 Hz, C<sup>5</sup>), 63.7 (d,  $^2J_{C-P}$  = 6.6 Hz, C<sup>9</sup>), 48.3-48.0 (m, C<sup>3</sup>), 28.4 (d,  $^1J_{C-P}$  = 75.5 Hz, C<sup>6</sup>), 26.9 (C<sup>6</sup>), 26.5 (C<sup>6</sup>), 26.4 (d,  $^2J_{C-P}$  = 2.5 Hz, C<sup>5</sup>), 21.4 (C<sup>9</sup>), 21.0 (C<sup>8</sup>), 16.5 (d,  $^3J_{C-P}$  = 6.0 Hz, C<sup>10</sup>).

**<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)** *Diastereoisomer C*:  $\delta$  -100.3 (ddd, 1F,  $^2J_{F-F}$  = 289.9 Hz,  $^2J_{F-P}$  = 81.2 Hz,  $^3J_{F-H}$  = 8.8 Hz), -113.1 (ddd, 1F,  $^2J_{F-F}$  = 289.9 Hz,  $^2J_{F-P}$  = 87.1 Hz,  $^3J_{F-H}$  = 24.0 Hz).  
*Diastereoisomer D*:  $\delta$  -102.7 (ddd, 1F,  $^2J_{F-F}$  = 291.3 Hz,  $^2J_{F-P}$  = 97.0 Hz,  $^3J_{F-H}$  = 12.7 Hz), -107.8 (ddd, 1F,  $^2J_{F-F}$  = 291.3 Hz,  $^2J_{F-P}$  = 121.0 Hz,  $^3J_{F-H}$  = 17.2 Hz).

**<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>)** *Diastereoisomer C*:  $\delta$  91.4 (dd,  $^2J_{P-F}$  = 87.1 Hz,  $^2J_{P-F}$  = 81.2 Hz).  
*Diastereoisomer D*:  $\delta$  95.5 (dd,  $^2J_{P-F}$  = 121.0 Hz,  $^2J_{P-F}$  = 97.0 Hz).

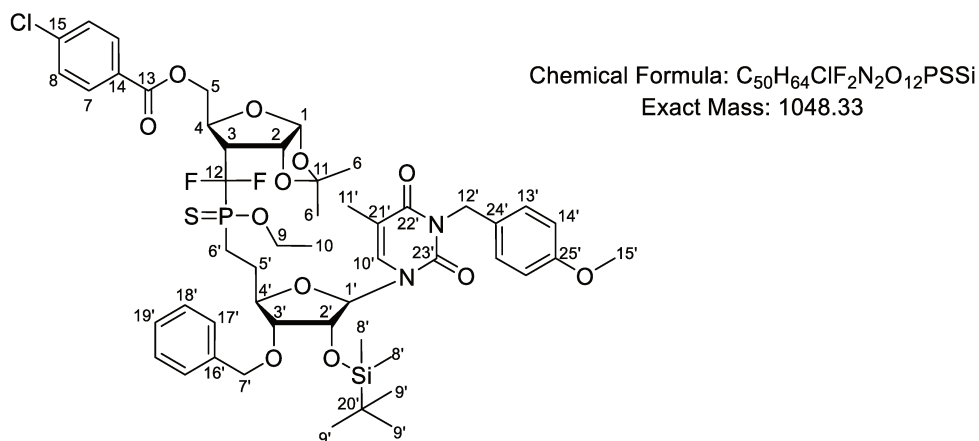
**HRMS** calculated for C<sub>35</sub>H<sub>42</sub>F<sub>2</sub>NaO<sub>12</sub>PS<sup>35</sup>Cl [M+Na]<sup>+</sup>: 813.1689. Found 813.1681. Calculated for C<sub>35</sub>H<sub>42</sub>F<sub>2</sub>NaO<sub>12</sub>PS<sup>37</sup>Cl [M+Na]<sup>+</sup>: 815.1659. Found 815.1678.

**FTIR ( $\nu_{\max}$  cm<sup>-1</sup>)** 2987, 1744, 1722, 1594, 1455, 1372, 1272, 1216, 1171, 1093, 1013, 955, 879, 778, 758, 734, 699, 684, 630, 526.

**Rf<sub>C</sub>** (pentane/diethyl ether, 1/1, v/v) = 0.51.

III.3. ((3aR,5S,6R,6aR)-6-(((2-((2R,3R,4R,5R)-3-(benzyloxy)-4-((tert-butyl dimethylsilyl)oxy)-5-(3-(4-methoxybenzyl)-5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl) tetrahydrofuran-2-yl)ethyl)(ethoxy)phosphorothioyl)difluoromethyl)-2,2-dimethyltetrahydrofuro[2,3-d][1,3]dioxol-5-yl) methyl 4-chlorobenzoate **197d**

To a suspension of potassium hydroxide (32 mg, 0.577 mmol, 1 eq) and benzyltriethylammonium chloride (66 mg, 0.289 mmol, 0.5 eq) in dichloromethane (5.0 mL) was added dropwise, at room temperature, a solution of *H*-phosphinothioate **165d** (1:1 ratio (A:B) determined by  $^{31}\text{P}$  NMR) (271 mg, 0.577 mmol, 1 eq) in dichloromethane (5.0 mL), followed by a solution of alkyl iodide **186f** (407 mg, 0.577 mmol, 1 eq) in dichloromethane (7.5 mL). The reaction mixture was stirred at 40 °C for 3 h. After cooling down to room temperature, the reaction solution was quenched with a saturated aqueous solution of ammonium chloride. The separated aqueous layer was extracted with dichloromethane. The combined organic layers were washed with brine, dried over  $\text{Na}_2\text{SO}_4$  and concentrated under reduced pressure. The crude residue, as a 1.7:1 (A:B) mixture of diastereoisomers (P center), determined by  $^{31}\text{P}$  NMR, was purified by flash chromatography on silica gel (pentane/ethyl acetate, gradient from 95/5 to 80/20) to afford the expected product as a white solid (415 mg, 69%). For analytical purposes, a pure fraction of diastereoisomer A was isolated and characterized separately.



**Diastereoisomer A:**

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.97 (d, 2H,  $J = 8.7$  Hz,  $\text{H}^7$ ), 7.46 (d, 2H,  $J = 8.7$  Hz,  $\text{H}^{13'}$ ), 7.40 (d, 2H,  $J = 8.7$  Hz,  $\text{H}^8$ ), 7.33-7.27 (m, 5H,  $\text{H}^{\text{Ar}}$ ), 7.01 (d, 1H,  $J = 1.1$  Hz,  $\text{H}^{10'}$ ), 6.80 (d, 2H,  $J = 8.7$  Hz,  $\text{H}^{14'}$ ), 5.80 (d, 1H,  $J = 3.7$  Hz,  $\text{H}^1$ ), 5.73 (d, 1H,  $J = 3.4$  Hz,  $\text{H}^{1'}$ ), 5.03 (s, 2H,  $\text{H}^{12'}$ ), 4.94 (dd, apparent t, 1H,  $J = 4.0$  Hz,  $\text{H}^2$ ), 4.78 (br d, 1H,  $J = 12.5$  Hz,  $\text{H}^5$ ), 4.72-4.68 (m, 1H,  $\text{H}^4$ ), 4.70 (d, 1H, AB syst,  $J = 11.7$  Hz,  $\text{H}^7$ ), 4.44 (d, 1H, AB syst,  $J = 11.7$  Hz,  $\text{H}^7$ ), 4.35-4.22 (m, 3H,  $\text{H}^{2'}$ ,  $\text{H}^5$ ,  $\text{H}^9$ ), 4.14-4.03 (m, 2H,  $\text{H}^{4'}$ ,  $\text{H}^9$ ), 3.75 (s, 3H,  $\text{H}^{15'}$ ), 3.55 (dd, 1H,  $J = 6.0$  Hz, 4.7 Hz,  $\text{H}^3$ ), 3.16-3.00 (m, 1H,  $\text{H}^3$ ), 2.54-2.28 (m, 2H,  $\text{H}^6$ ), 2.09-1.98 (m, 2H,  $\text{H}^5$ ), 1.93 (d, 3H,  $J = 0.9$  Hz,  $\text{H}^{11'}$ ), 1.44 (s, 3H,  $\text{H}^6$ ), 1.28 (t, 3H,  $J = 7.0$  Hz,  $\text{H}^{10}$ ), 1.13 (s, 3H,  $\text{H}^6$ ), 0.86 (s, 9H,  $\text{H}^9$ ), 0.06 (s, 3H,  $\text{H}^8$ ), 0.04 (s, 3H,  $\text{H}^8$ ).

$^{13}\text{C}$   $\{^1\text{H}\}$  NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  " $\text{C}^{12}$ , unobserved due to multiplicity" 165.5 ( $\text{C}^{13}$ ), 163.4 ( $\text{C}^{22}$ ), 159.3 ( $\text{C}^{25'}$ ), 150.9 ( $\text{C}^{23'}$ ), 139.8 ( $\text{C}^{14}$ ), 137.5 ( $\text{C}^{16'}$ ), 133.6 ( $\text{C}^{10'}$ ), 131.3 (2C,  $\text{C}^7$  or 8), 131.2 (2C,  $\text{C}^7$  or 8),

129.3 (C<sup>24</sup>), 129.0 (2C, C<sup>13</sup>), 128.7 (2C, C<sup>18</sup>), 128.5 (C<sup>15</sup>), 128.2 (C<sup>19</sup>), 127.8 (2C, C<sup>17</sup>), 113.8 (2C, C<sup>14</sup>), 113.4 (C<sup>11</sup>), 110.5 (C<sup>21</sup>), 105.2 (C<sup>1</sup>), 92.3 (C<sup>1</sup>), 81.3 (C<sup>3</sup> or 4<sup>1</sup>), 81.0 (C<sup>3</sup> or 4<sup>1</sup>), 79.1 (d, <sup>3</sup>J<sub>C-P</sub> = 9.5 Hz, C<sup>2</sup>), 75.0 (t, <sup>2</sup>J<sub>C-F</sub> = 5.6 Hz, C<sup>4</sup>), 74.3 (C<sup>2</sup>), 72.7 (C<sup>7</sup>), 64.8 (br s, C<sup>5</sup>), 63.8 (d, <sup>2</sup>J<sub>C-P</sub> = 8.5 Hz, C<sup>9</sup>), 55.4 (C<sup>15</sup>), 48.3-47.6 (m, C<sup>3</sup>), 44.0 (C<sup>12</sup>), 28.4 (d, <sup>1</sup>J<sub>C-P</sub> = 75.4 Hz, C<sup>6</sup>), 26.8 (C<sup>6</sup>), 26.3 (C<sup>6</sup>), 25.8 (4C, C<sup>5</sup>, C<sup>9</sup>), 18.2 (C<sup>20</sup>), 16.4 (d, <sup>3</sup>J<sub>C-P</sub> = 6.3 Hz, C<sup>10</sup>), 13.6 (C<sup>11</sup>), -4.4 (C<sup>8</sup>), -4.9 (C<sup>8</sup>).

<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>) δ -100.2 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 290.2 Hz, <sup>2</sup>J<sub>F-P</sub> = 81.5 Hz, <sup>3</sup>J<sub>F-H</sub> = 9.5 Hz), -111.0 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 290.2 Hz, <sup>2</sup>J<sub>F-P</sub> = 88.0 Hz, <sup>3</sup>J<sub>F-H</sub> = 23.6 Hz).

<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>) δ 91.4 (dd, <sup>2</sup>J<sub>P-F</sub> = 88.0 Hz, <sup>2</sup>J<sub>P-F</sub> = 81.5 Hz).

HRMS (TOF ESI+) calculated for C<sub>50</sub>H<sub>64</sub>F<sub>2</sub>N<sub>2</sub>O<sub>12</sub>NaSiPSCI [M+Na]<sup>+</sup>: 1071.3241. Found 1071.3235.

FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 2959, 2929, 2858, 1725, 1667, 1512, 1461, 1374.

[α]<sub>D</sub><sup>22</sup> + 66.2 (c 1.60, CHCl<sub>3</sub>).

M.p. = 51-53 °C.

Rf<sub>A</sub> (petroleum ether/ethyl acetate, 4/1, v/v) = 0.36.

**Diastereoisomer B (determined by subtraction of diastereoisomer A from a 1:1 mixture A:B, determined by <sup>31</sup>P NMR):**

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.96 (d, 2H, J = 8.7 Hz, H<sup>7</sup>), 7.46 (d, 2H, J = 8.7 Hz, H<sup>13</sup>), 7.39 (d, 2H, J = 8.7 Hz, H<sup>8</sup>), 7.33-7.27 (m, 5H, H<sup>A</sup>), 7.00 (d, 1H, J = 1.1 Hz, H<sup>10</sup>), 6.80 (d, 2H, J = 8.7 Hz, H<sup>14</sup>), 5.86 (d, 1H, J = 3.6 Hz, H<sup>1</sup>), 5.72 (d, 1H, J = 3.7 Hz, H<sup>1</sup>), 5.03 (s, 2H, H<sup>12</sup>), 4.96 (dd, apparent br t, 1H, J = 3.6 Hz, H<sup>2</sup>), 4.82-4.72 (m, 2H, H<sup>4</sup>, H<sup>5</sup>), 4.69 (d, 1H, AB syst, J = 11.6 Hz, H<sup>7</sup>), 4.43 (d, 1H, AB syst, J = 11.6 Hz, H<sup>7</sup>), 4.36-4.22 (m, 3H, H<sup>2</sup>, H<sup>5</sup>, H<sup>9</sup>), 4.12-4.02 (m, 2H, H<sup>4</sup>, H<sup>9</sup>), 3.75 (s, 3H, H<sup>15</sup>), 3.57-3.53 (m, 1H, H<sup>3</sup>), 3.23-3.00 (m, 1H, H<sup>3</sup>), 2.51-2.23 (m, 2H, H<sup>6</sup>), 2.06-2.00 (m, 2H, H<sup>5</sup>), 1.92 (d, 3H, J = 1.1 Hz, H<sup>11</sup>), 1.54 (s, 3H, H<sup>6</sup>), 1.31 (s, 3H, H<sup>6</sup>), 1.30 (t, 3H, J = 6.9 Hz, H<sup>10</sup>), 0.85 (s, 9H, H<sup>9</sup>), 0.05 (s, 3H, H<sup>8</sup>), 0.02 (s, 3H, H<sup>8</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ "C<sup>12</sup>, unobserved due to multiplicity" 165.4 (C<sup>13</sup>), 163.4 (C<sup>22</sup>), 159.3 (C<sup>25</sup>), 150.9 (C<sup>23</sup>), 139.8 (C<sup>14</sup>), 137.5 (C<sup>16</sup>), 133.9 (C<sup>10</sup>), 131.3 (2C, C<sup>7</sup> or 8), 131.2 (2C, C<sup>7</sup> or 8), 129.3 (C<sup>24</sup>), 129.0 (2C, C<sup>13</sup>), 128.7 (2C, C<sup>18</sup>), 128.5 (C<sup>15</sup>), 128.1 (C<sup>19</sup>), 127.8 (2C, C<sup>17</sup>), 113.8 (2C, C<sup>14</sup>), 113.5 (C<sup>11</sup>), 110.5 (C<sup>21</sup>), 105.3 (C<sup>1</sup>), 92.5 (C<sup>1</sup>), 81.3 (C<sup>4</sup>), 80.8 (C<sup>3</sup>), 78.8 (d, <sup>3</sup>J<sub>C-P</sub> = 10.5 Hz, C<sup>2</sup>), 75.0 (C<sup>4</sup>), 74.0 (C<sup>2</sup>), 72.7 (C<sup>7</sup>), 64.8 (C<sup>5</sup>), 63.2 (d, <sup>2</sup>J<sub>C-P</sub> = 6.6 Hz, C<sup>9</sup>), 55.4 (C<sup>15</sup>), 48.3-47.6 (m, C<sup>3</sup>), 44.0 (C<sup>12</sup>), 28.9 (d, J = 71.0 Hz, C<sup>6</sup>), 27.0 (C<sup>6</sup>), 26.8 (C<sup>6</sup>), 26.5 (C<sup>5</sup>), 25.0 (3C, C<sup>9</sup>), 18.2 (C<sup>20</sup>), 16.6 (d, <sup>3</sup>J<sub>C-P</sub> = 6.3 Hz, C<sup>10</sup>), 13.5 (C<sup>11</sup>), -4.4 (C<sup>8</sup>), -4.9 (C<sup>8</sup>).

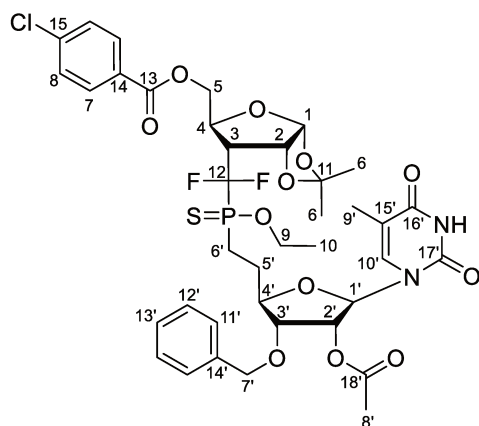
<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>) δ -99.0 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 291.1 Hz, <sup>2</sup>J<sub>F-P</sub> = 109.6 Hz, <sup>3</sup>J<sub>F-H</sub> = 7.0 Hz), -112.9 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 291.1 Hz, <sup>2</sup>J<sub>F-P</sub> = 79.5 Hz, <sup>3</sup>J<sub>F-H</sub> = 26.5 Hz).

<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>) δ 91.9 (dd, <sup>2</sup>J<sub>P-F</sub> = 109.6 Hz, <sup>2</sup>J<sub>P-F</sub> = 79.5 Hz).

Rf<sub>B</sub> (petroleum ether/ethyl acetate, 4/1, v/v) = 0.22.

III.4. (((3aR,5S,6R,6aR)-6-(((2-((2R,3R,4R,5R)-4-acetoxy-3-(benzyloxy)-5-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-2-yl)ethyl)(ethoxy)phosphorothioyl)difluoromethyl)-2,2-dimethyltetrahydrofuro[2,3-d][1,3]dioxol-5-yl)methyl 4-chlorobenzoate **197a**

To a suspension of potassium hydroxide (35 mg, 0.615 mmol, 1 eq) and benzyltriethylammonium chloride (70 mg, 0.308 mmol, 0.5 eq) in dichloromethane (5.0 mL) was added dropwise, at room temperature, a solution of *H*-phosphinothioate **165d** (1:1 ratio (A:B) determined by  $^{31}\text{P}$  NMR) (289 mg, 0.615 mmol, 1 eq) in dichloromethane (6.0 mL), followed by a solution of alkyl iodide **186a** (316 mg, 0.615 mmol, 1 eq) in dichloromethane (8.0 mL). The reaction mixture was stirred at 40 °C for 3 h. After cooling down to room temperature, the reaction solution was quenched with a saturated aqueous solution of ammonium chloride. The separated aqueous layer was extracted with dichloromethane. The combined organic layers were washed with brine, dried over  $\text{Na}_2\text{SO}_4$  and concentrated under reduced pressure. The crude residue, as a 1.7:1 (A:B) mixture of diastereoisomers (determined by  $^{31}\text{P}$  NMR), was purified by flash chromatography on silica gel (pentane/ethyl acetate, gradient from 8/2 to 5/5) followed by a purification on Interchim Puriflash® (method A) to afford the expected product as a white solid (250 mg, 47%). For analytical purposes, a pure fraction of diastereoisomer A and diastereoisomer B were isolated and characterized separately.



Chemical Formula:  $\text{C}_{38}\text{H}_{44}\text{ClF}_2\text{N}_2\text{O}_{12}\text{PS}$   
Exact Mass: 856.20

**Diastereoisomer A:**

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  9.26 (s, 1H, NH), 7.96 (d, 2H,  $J = 8.7$  Hz,  $\text{H}^7$ ), 7.39 (d, 2H,  $J = 8.7$  Hz,  $\text{H}^8$ ), 7.33-7.27 (m, 5H,  $\text{H}^{\text{Ar}}$ ), 6.98 (d, 1H,  $J = 1.0$  Hz,  $\text{H}^{10'}$ ), 5.83 (d, 1H,  $J = 3.8$  Hz,  $\text{H}^1$ ), 5.70 (d, 1H,  $J = 3.5$  Hz,  $\text{H}^{1'}$ ), 5.39 (dd, 1H,  $J = 5.3$  Hz, 3.6 Hz,  $\text{H}^2$ ), 4.96 (dd, apparent t, 1H,  $J = 4.0$  Hz,  $\text{H}^2$ ), 4.78 (d, 1H,  $J = 12.5$  Hz,  $\text{H}^5$ ), 4.73-4.68 (br dd, 1H,  $J = 11.1$  Hz, 4.8 Hz,  $\text{H}^4$ ), 4.59 (d, 1H, AB syst,  $J = 11.3$  Hz,  $\text{H}^7$ ), 4.47 (d, 1H, AB syst,  $J = 11.3$  Hz,  $\text{H}^7$ ), 4.34-4.21 (m, 2H,  $\text{H}^5$ ,  $\text{H}^9$ ), 4.11-3.92 (m, 3H,  $\text{H}^9$ ,  $\text{H}^3$ ,  $\text{H}^4$ ), 3.15-2.99 (m, 1H,  $\text{H}^3$ ), 2.42-2.26 (m, 2H,  $\text{H}^6$ ), 2.09 (s, 3H,  $\text{H}^8$ ), 2.09-2.00 (m, 2H,  $\text{H}^5$ ), 1.92 (d, 3H,  $J = 1.0$  Hz,  $\text{H}^9$ ), 1.46 (s, 3H,  $\text{H}^6$ ), 1.27 (t, 3H,  $J = 7.0$  Hz,  $\text{H}^{10}$ ), 1.21 (s, 3H,  $\text{H}^6$ ).

$^{13}\text{C}$  { $^1\text{H}$ } NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  " $\text{C}^{12}$ , unobserved due to multiplicity" 170.2 ( $\text{C}^{18'}$ ), 165.5 ( $\text{C}^{13}$ ), 163.9 ( $\text{C}^{16'}$ ), 150.1 ( $\text{C}^{17'}$ ), 139.8 ( $\text{C}^{14}$ ), 137.2 ( $\text{C}^{14'}$ ), 136.6 ( $\text{C}^{10'}$ ), 131.3 (2C,  $\text{C}^7$ ), 128.9 (2C,  $\text{C}^8$ ), 128.7 (2C,  $\text{C}^{12'}$ ), 128.43 ( $\text{C}^{15}$ ), 128.40 ( $\text{C}^{13'}$ ), 128.2 (2C,  $\text{C}^{11'}$ ), 113.3 ( $\text{C}^{11}$ ), 111.7 ( $\text{C}^{15'}$ ), 105.2 ( $\text{C}^1$ ), 90.6 ( $\text{C}^{1'}$ ),

81.3 (d,  $^3J_{C-P} = 16.5$  Hz, C<sup>4'</sup>), 79.4 (C<sup>3'</sup>), 79.2 (d,  $^3J_{C-P} = 9.8$  Hz, C<sup>2</sup>), 75.0 (dd, apparent t,  $^3J_{C-P} = 5.2$  Hz, C<sup>4</sup>), 73.9 (C<sup>2'</sup>), 73.7 (C<sup>7'</sup>), 64.8 (d,  $^4J_{C-P} = 1.9$  Hz, C<sup>5</sup>), 63.8 (br d,  $^2J_{C-P} = 6.0$  Hz, C<sup>9</sup>), 48.5-47.8 (m, C<sup>3</sup>), 28.0 (d,  $^1J_{C-P} = 76.5$  Hz, C<sup>6'</sup>), 26.8 (C<sup>6</sup>), 26.5 (C<sup>6</sup>), 25.2 (d,  $^2J_{C-P} = 3.0$  Hz, C<sup>5</sup>), 20.9 (C<sup>8</sup>), 16.4 (d,  $^3J_{C-P} = 6.3$  Hz, C<sup>10</sup>), 12.7 (C<sup>9'</sup>).

**<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)**  $\delta$  -100.1 (ddd, 1F,  $^2J_{F-F} = 290.7$  Hz,  $^2J_{F-P} = 81.6$  Hz,  $^3J_{F-H} = 9.0$  Hz), -111.1 (ddd, 1F,  $^2J_{F-F} = 290.7$  Hz,  $^2J_{F-P} = 88.1$  Hz,  $^3J_{F-H} = 23.4$  Hz).

**<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>)**  $\delta$  91.2 (dd,  $^2J_{P-F} = 88.1$  Hz,  $^2J_{P-F} = 81.6$  Hz).

**HRMS (TOF ESI+)** calculated for C<sub>38</sub>H<sub>44</sub>F<sub>2</sub>N<sub>2</sub>O<sub>12</sub>NaPS<sup>35</sup>Cl [M+Na]<sup>+</sup>: 879.1907. Found 879.1927. Calculated for C<sub>38</sub>H<sub>44</sub>F<sub>2</sub>N<sub>2</sub>O<sub>12</sub>NaPS<sup>37</sup>Cl [M+Na]<sup>+</sup>: 881.1877. Found 881.1909.

**FTIR ( $\nu_{\max}$  cm<sup>-1</sup>)** 2992, 1691, 1372, 1271, 1226, 1091, 1013, 877, 758, 699, 581, 489, 409.

**[ $\alpha$ ]<sub>D</sub><sup>22</sup> + 66.2** (c 1.00, CHCl<sub>3</sub>).

**M.p.** = 91-92 °C.

**Rf<sub>A</sub>** (petroleum ether/ethyl acetate, 1/1, v/v) = 0.31.

#### **Diastereoisomer B:<sup>172</sup>**

**<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)**  $\delta$  9.15 (s, 1H, NH), 7.96 (d, 2H,  $J = 8.5$  Hz, H<sup>7</sup>), 7.39 (d, 2H,  $J = 8.5$  Hz, H<sup>8</sup>), 7.33-7.26 (m, 5H, H<sup>Ar</sup>), 6.96 (d, 1H,  $J = 0.9$  Hz, H<sup>10'</sup>), 5.86 (d, 1H,  $J = 3.2$  Hz, H<sup>1</sup>), 5.62 (d, 1H,  $J = 2.8$  Hz, H<sup>1</sup>), 5.40 (dd, 1H,  $J = 5.2$  Hz, 3.1 Hz, H<sup>2'</sup>), 4.96 (dd, apparent t, 1H,  $J = 3.7$  Hz, H<sup>2</sup>), 4.81-4.74 (m, 2H, H<sup>4</sup>, H<sup>5</sup>), 4.57 (d, 1H, AB syst,  $J = 11.1$  Hz, H<sup>7'</sup>), 4.45 (d, 1H, AB syst,  $J = 11.1$  Hz, H<sup>7'</sup>), 4.32 (dd, 1H,  $J = 12.6$  Hz, 4.6 Hz, H<sup>5</sup>), 4.26-4.20 (m, 1H, H<sup>9</sup>), 4.13-3.99 (m, 2H, H<sup>9</sup>, H<sup>3</sup>), 3.96-3.90 (m, 1H, H<sup>4'</sup>), 3.24-3.08 (m, 1H, H<sup>3</sup>), 2.32-2.14 (m, 3H, H<sup>5'</sup>, H<sup>6'</sup>), 2.08 (s, 3H, H<sup>8</sup>), 1.95-1.81 (m, 1H, H<sup>5</sup>), 1.90 (d, 3H,  $J = 0.9$  Hz, H<sup>9</sup>), 1.53 (s, 3H, H<sup>6</sup>), 1.30 (s, 3H, H<sup>6</sup>), 1.29 (t, 3H,  $J = 7.0$  Hz, H<sup>10</sup>).

**<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>)**  $\delta$  "C<sup>12</sup>, unobserved due to multiplicity" 170.3 (C<sup>18'</sup>), 165.4 (C<sup>13</sup>), 163.7 (C<sup>16'</sup>), 150.0 (C<sup>17'</sup>), 139.8 (C<sup>14</sup>), 137.2 (C<sup>10'</sup>), 137.1 (C<sup>14'</sup>), 131.3 (2C, C<sup>7</sup>), 129.0 (2C, C<sup>8</sup>), 128.8 (2C, C<sup>11'</sup>), 128.5 (2C, C<sup>15</sup>, C<sup>13'</sup>), 128.4 (2C, C<sup>12'</sup>), 113.4 (C<sup>11</sup>), 111.7 (C<sup>15</sup>), 105.3 (C<sup>1</sup>), 91.3 (C<sup>1</sup>), 81.3 (d,  $^3J_{C-P} = 17.6$  Hz, C<sup>4'</sup>), 79.3 (C<sup>3'</sup>), 78.9 (d,  $^3J_{C-P} = 10.0$  Hz, C<sup>2</sup>), 75.2-75.0 (br m, C<sup>4</sup>), 73.9 (C<sup>2'</sup>), 73.8 (C<sup>7'</sup>), 64.8 (d,  $^4J_{C-P} = 3.6$  Hz, C<sup>5</sup>), 63.2 (d,  $^2J_{C-P} = 7.4$  Hz, C<sup>9</sup>), 47.0-46.3 (m, C<sup>3</sup>), 27.0 (C<sup>6</sup>), 26.8 (d,  $^1J_{C-P} = 72.3$  Hz, C<sup>6'</sup>), 26.7 (C<sup>6</sup>), 25.4 (br s, C<sup>5'</sup>), 20.9 (C<sup>8</sup>), 16.4 (d,  $^3J_{C-P} = 6.3$  Hz, C<sup>10</sup>), 12.7 (C<sup>9'</sup>).

**<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)**  $\delta$  -99.2 (ddd, 1F,  $^2J_{F-F} = 290.5$  Hz,  $^2J_{F-P} = 108.9$  Hz,  $^3J_{F-H} = 6.2$  Hz), -113.0 (ddd, 1F,  $^2J_{F-F} = 290.5$  Hz,  $^2J_{F-P} = 78.3$  Hz,  $^3J_{F-H} = 26.5$  Hz).

**<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>)**  $\delta$  91.4 (dd,  $^2J_{P-F} = 108.9$  Hz,  $^2J_{P-F} = 78.3$  Hz).

**FTIR ( $\nu_{\max}$  cm<sup>-1</sup>)** 2981, 1692, 1372, 1224, 1091, 1014, 757, 408.

**[ $\alpha$ ]<sub>D</sub><sup>22</sup> + 42.8** (c 1.00, CHCl<sub>3</sub>).

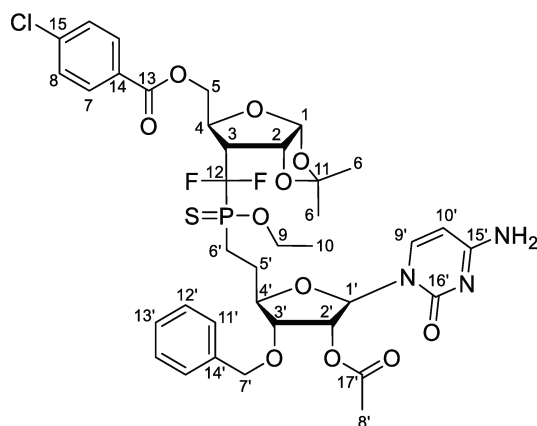
**M.p.** = 91-92 °C.

**Rf<sub>B</sub>** (petroleum ether/ethyl acetate, 1/1, v/v) = 0.26.

<sup>172</sup> 2 impurities were present with B: first is anomer  $\alpha$  of thymine and second had a chemistry shift of -147.8 in <sup>31</sup>P {<sup>1</sup>H} NMR (dd,  $J = 91.6$  Hz, 65.6 Hz).

III.5. ((3aR,5S,6R,6aR)-6-(((2-((2R,3R,4R,5R)-4-acetoxy-5-(4-amino-2-oxopyrimidin-1(2H)-yl)-3-(benzyloxy)tetrahydrofuran-2-yl)ethyl)(ethoxy)phosphorothioyl)difluoromethyl)-2,2-dimethyltetrahydrofuro[2,3-d][1,3]dioxol-5-yl)methyl 4-chlorobenzoate **197b**

To a suspension of potassium hydroxide (37 mg, 0.659 mmol, 1 eq) and benzyltriethylammonium chloride (75 mg, 0.659 mmol, 0.5 eq) in dichloromethane (5.0 mL) was added dropwise, at room temperature, a solution of *H*-phosphinothioate **165d** (1:1 ratio (A:B) determined by  $^{31}\text{P}$  NMR) (310 mg, 0.659 mmol, 1 eq) in dichloromethane (6.0 mL), followed by a solution of alkyl iodide **186b** (330 mg, 0.659 mmol, 1 eq) in dichloromethane (8.0 mL). The reaction mixture was stirred at 40 °C for 20 h. After cooling down to room temperature, the reaction solution was quenched with a saturated aqueous solution of ammonium chloride. The separated aqueous layer was extracted with dichloromethane. The combined organic layers were washed with brine, dried over  $\text{Na}_2\text{SO}_4$  and concentrated under reduced pressure. The crude residue, as a 2.1:1 (A:B) mixture of diastereoisomers (determined by  $^{31}\text{P}$  NMR), was purified by flash chromatography on silica gel (dichloromethane/methanol, gradient from 97/3 to 96/4) to afford the expected product as a white solid (302 mg, 55%). For analytical purposes, a pure fraction of diastereoisomer A was isolated and characterized separately. A fraction of diastereoisomer B was isolated in inseparable mixture of four products (ratio 1:4:5:20, impurities:B).



Chemical Formula:  $\text{C}_{37}\text{H}_{43}\text{ClF}_2\text{N}_3\text{O}_{11}\text{PS}$   
Exact Mass: 841.20

**Diastereoisomer A:**<sup>173</sup>

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.87 (d, 2H,  $J = 8.4$  Hz,  $\text{H}^7$ ), 7.29 (d, 2H,  $J = 8.4$  Hz,  $\text{H}^8$ ), 7.23-7.14 (m, 6H,  $\text{H}^{\text{Ar}}$ ,  $\text{H}^9$ ), 5.86 (d, 1H,  $J = 7.5$  Hz,  $\text{H}^{10'}$ ), 5.74 (d, 1H,  $J = 3.6$  Hz,  $\text{H}^1$ ), 5.58 (d, 1H,  $J = 1.2$  Hz,  $\text{H}^{1'}$ ), 5.41 (dd, 1H,  $J = 4.2$  Hz, 1.8 Hz,  $\text{H}^{2'}$ ), 4.87 (t, 1H,  $J = 3.6$  Hz,  $\text{H}^2$ ), 4.68 (br d, 1H,  $J = 12.3$  Hz,  $\text{H}^5$ ), 4.61 (dd, 1H,  $J = 10.5$  Hz, 4.5 Hz,  $\text{H}^4$ ), 4.49 (d, 1H, AB syst,  $J = 11.1$  Hz,  $\text{H}^7$ ), 4.33 (d, 1H, AB syst,  $J = 11.1$  Hz,  $\text{H}^7$ ), 4.22 (dd, 1H,  $J = 12.6$  Hz, 4.8 Hz,  $\text{H}^5$ ), 4.17-4.11 (m, 1H,  $\text{H}^9$ ), 4.01-3.82 (m, 3H,  $\text{H}^9$ ,  $\text{H}^{3'}$ ,  $\text{H}^4$ ), 3.05-2.89 (m, 1H,  $\text{H}^3$ ), 2.34-2.21 (m, 2H,  $\text{H}^6$ ), 2.00 (s, 3H,  $\text{H}^8$ ), 2.00-1.92 (m, 2H,  $\text{H}^{5'}$ ), 1.36 (br s, 3H,  $\text{H}^6$ ), 1.17 (t, 3H,  $J = 7.2$  Hz,  $\text{H}^{10}$ ), 1.11 (br s, 3H,  $\text{H}^6$ ).

<sup>173</sup> An impurity was present with A in  $^{31}\text{P}$   $\{^1\text{H}\}$  NMR (ratio 1:10, impurity:A), with chemical shifts at 58.7 (dd,  $J = 89.5$  Hz, 52.2 Hz) and 58.0 (dd,  $J = 89.5$  Hz, 49.1 Hz).

**<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>)**  $\delta$  "C<sup>12</sup>, unobserved due to multiplicity" 170.0 (C<sup>17</sup>), 165.5 (C<sup>13</sup>), 164.9 (C<sup>15'</sup>), 154.3 (C<sup>16'</sup>), 141.8 (C<sup>9'</sup>), 139.8 (C<sup>14</sup>), 137.3 (C<sup>14'</sup>), 131.3 (2C, C<sup>7</sup>), 128.9 (2C, C<sup>8</sup>), 128.7 (2C, C<sup>12'</sup>), 128.41 (C<sup>15</sup>), 128.35 (C<sup>13'</sup>), 128.2 (2C, C<sup>11'</sup>), 113.3 (C<sup>11</sup>), 105.2 (C<sup>1</sup>), 95.9 (C<sup>10'</sup>), 91.8 (C<sup>1'</sup>), 81.2 (d, <sup>3</sup>J<sub>C-P</sub> = 16.7 Hz, C<sup>4'</sup>), 79.6 (C<sup>3'</sup>), 79.1 (d, <sup>3</sup>J<sub>C-P</sub> = 9.9 Hz, C<sup>2</sup>), 75.0 (C<sup>4</sup>), 74.3 (C<sup>2'</sup>), 73.7 (C<sup>7'</sup>), 64.9 (br s, C<sup>5</sup>), 63.8 (d, <sup>2</sup>J<sub>C-P</sub> = 6.3 Hz, C<sup>9</sup>), 48.6-47.8 (m, C<sup>3</sup>), 28.2 (br d, <sup>1</sup>J<sub>C-P</sub> = 75.5 Hz, C<sup>6'</sup>), 26.8 (C<sup>6</sup>), 26.6 (C<sup>6</sup>), 25.1 (br s, C<sup>5'</sup>), 21.0 (C<sup>8'</sup>), 16.4 (d, <sup>3</sup>J<sub>C-P</sub> = 6.3 Hz, C<sup>10</sup>).

**<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)**  $\delta$  -99.9 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 290.5 Hz, <sup>2</sup>J<sub>F-P</sub> = 81.8 Hz, <sup>3</sup>J<sub>F-H</sub> = 8.7 Hz), -111.2 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 290.5 Hz, <sup>2</sup>J<sub>F-P</sub> = 87.7 Hz, <sup>3</sup>J<sub>F-H</sub> = 24.0 Hz).

**<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>)**  $\delta$  91.1 (dd, <sup>2</sup>J<sub>P-F</sub> = 81.8 Hz, <sup>2</sup>J<sub>P-F</sub> = 87.7 Hz).

**HRMS (TOF ESI+)** calculated for C<sub>37</sub>H<sub>44</sub>F<sub>2</sub>N<sub>3</sub>O<sub>11</sub>PS<sup>35</sup>Cl [M+H]<sup>+</sup>: 842.2091. Found 842.2106. Calculated for C<sub>37</sub>H<sub>44</sub>F<sub>2</sub>N<sub>3</sub>O<sub>11</sub>PS<sup>37</sup>Cl [M+H]<sup>+</sup>: 844.2061. Found 844.2091.

**FTIR (ν<sub>max</sub> cm<sup>-1</sup>)** 2980, 1722, 1646, 1488, 1373, 1273, 1226, 1169, 1090, 1013, 757, 729, 628.

**[α]<sub>D</sub><sup>21</sup>** + 74.6 (c 1.00, CHCl<sub>3</sub>).

**M.p.** = 88-89 °C.

**R<sub>f</sub>** (dichloromethane/methanol, 95/5, v/v) = 0.30.

**Diastereoisomer B:**<sup>174</sup>

**<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)**  $\delta$  7.86 (d, 2H, *J* = 8.4 Hz, H<sup>7</sup>), 7.29 (d, 2H, *J* = 8.4 Hz, H<sup>8</sup>), 7.24-7.15 (m, 6H, H<sup>Ar</sup>, H<sup>9</sup>), 5.81-5.71 (m, 2H, H<sup>1</sup>, H<sup>10</sup>), 5.51 (br s, 1H, H<sup>1'</sup>), 5.42-5.41 (m, 1H, H<sup>2</sup>), 4.87 (t, 1H, *J* = 3.6 Hz, H<sup>2</sup>), 4.70 (d, 1H, *J* = 11.7 Hz, H<sup>5</sup>), 4.65-4.64 (m, 1H, H<sup>4</sup>), 4.47 (d, 1H, AB syst, *J* = 11.1 Hz, H<sup>7</sup>), 4.32 (d, 1H, AB syst, *J* = 11.1 Hz, H<sup>7'</sup>), 4.28-4.10 (m, 2H, H<sup>5</sup>, H<sup>9</sup>), 4.04-3.72 (m, 3H, H<sup>9</sup>, H<sup>3'</sup>, H<sup>4</sup>), 3.11-2.98 (m, 1H, H<sup>3</sup>), 2.92-2.66 (m, 1H, H<sup>6</sup>), 2.29-2.06 (m, 2H, H<sup>5</sup>, H<sup>6</sup>), 1.99 (s, 3H, H<sup>8</sup>), 1.93-1.77 (m, 1H, H<sup>5</sup>), 1.44 (s, 3H, H<sup>6</sup>), 1.21 (s, 3H, H<sup>6</sup>), 1.19 (t, 3H, *J* = 6.9 Hz, H<sup>10</sup>).

**<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>)**  $\delta$  "C<sup>12</sup>, unobserved due to multiplicity" 170.0 (C<sup>17</sup>), 165.4 (C<sup>13</sup> or <sup>15'</sup>), 165.3 (C<sup>13</sup> or <sup>15'</sup>), 154.4 (C<sup>16'</sup>), 142.3 (C<sup>9'</sup>), 139.8 (C<sup>14</sup>), 137.2 (C<sup>14'</sup>), 131.3 (2C, C<sup>7</sup>), 128.9 (2C, C<sup>8</sup>), 128.7 (2C, C<sup>12'</sup>), 128.41 (C<sup>15</sup>), 128.3 (2C, C<sup>11'</sup>), 128.2 (C<sup>13'</sup>), 113.4 (C<sup>11</sup>), 105.2 (C<sup>1</sup>), 95.7 (C<sup>10'</sup>), 92.4 (C<sup>1'</sup>), 81.1 (d, <sup>3</sup>J<sub>C-P</sub> = 17.2 Hz, C<sup>4'</sup>), 79.6 (C<sup>3'</sup>), 78.8 (d, <sup>3</sup>J<sub>C-P</sub> = 9.6 Hz, C<sup>2</sup>), 75.0 (br s, C<sup>4</sup>), 74.3 (C<sup>2'</sup>), 73.7 (C<sup>7'</sup>), 64.7 (br s, C<sup>5</sup>), 63.2 (d, <sup>2</sup>J<sub>C-P</sub> = 7.1 Hz, C<sup>9</sup>), 47.1-46.4 (m, C<sup>3</sup>), 26.9 (d, <sup>1</sup>J<sub>C-P</sub> = 75.0 Hz, C<sup>6'</sup>), 26.9 (C<sup>6</sup>), 26.7 (C<sup>6</sup>), 25.4 (C<sup>5'</sup>), 21.0 (C<sup>8'</sup>), 16.4 (d, <sup>3</sup>J<sub>C-P</sub> = 6.3 Hz, C<sup>10</sup>).

**<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)**  $\delta$  -99.0 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 290.2 Hz, <sup>2</sup>J<sub>F-P</sub> = 108.0 Hz, <sup>3</sup>J<sub>F-H</sub> = 5.6 Hz), -112.8 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 290.2 Hz, <sup>2</sup>J<sub>F-P</sub> = 78.4 Hz, <sup>3</sup>J<sub>F-H</sub> = 25.9 Hz).

**<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>)**  $\delta$  91.4 (dd, <sup>2</sup>J<sub>P-F</sub> = 108.0 Hz, <sup>2</sup>J<sub>P-F</sub> = 78.4 Hz).

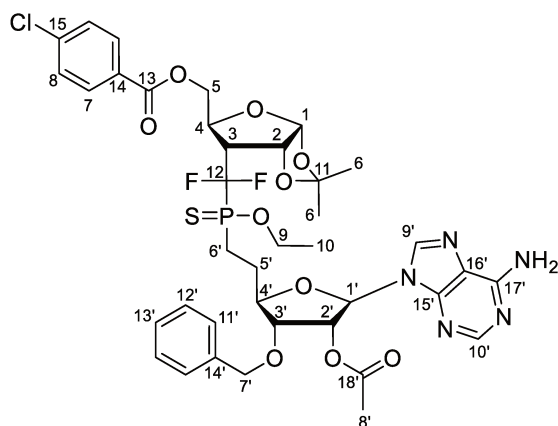
**FTIR (ν<sub>max</sub> cm<sup>-1</sup>)** 2985, 1723, 1646, 1488, 1373, 1228, 1091, 1013, 758, 698, 596.

**R<sub>f</sub>** (dichloromethane/methanol, 95/5, v/v) = 0.30.

<sup>174</sup> Three impurities were present with B in <sup>31</sup>P {<sup>1</sup>H} NMR (ratio 1/2, impurities/B), with chemical shifts at 148.3 (dd, *J* = 93.2 Hz, 63.3 Hz), 144.4 (dd, *J* = 84.5 Hz, 66.3 Hz), 40.4 (dd, *J* = 101.3 Hz, 52.2 Hz) and 39.4 (dd, *J* = 101.3 Hz, 46.9 Hz).

III.6. ((3aR,5S,6R,6aR)-6-(((2-((2R,3R,4R,5R)-4-acetoxy-5-(6-amino-9H-purin-9-yl)-3-(benzyloxy)tetrahydrofuran-2-yl)ethyl)(ethoxy)phosphorothioyl)difluoromethyl)-2,2-dimethyltetrahydrofuro[2,3-d][1,3]dioxol-5-yl)methyl 4-chlorobenzoate **197c**

To a suspension of potassium hydroxide (27 mg, 0.478 mmol, 1 eq) and benzyltriethylammonium chloride (54 mg, 0.239 mmol, 0.5 eq) in dichloromethane (5.0 mL) was added dropwise, at room temperature, a solution of *H*-phosphinothioate **165d** (1:1 ratio (A:B) determined by  $^{31}\text{P}$  NMR) (225 mg, 0.478 mmol, 1 eq) in dichloromethane (7.0 mL), followed by a solution of alkyl iodide **186c** (250 mg, 0.478 mmol, 1 eq) in dichloromethane (8.0 mL). The reaction mixture was stirred at 40 °C for 40 h. After cooling down to room temperature, the reaction solution was quenched with a saturated aqueous solution of ammonium chloride. The separated aqueous layer was extracted with dichloromethane. The combined organic layers were washed with brine, dried over  $\text{Na}_2\text{SO}_4$  and concentrated under reduced pressure. The crude residue, as a 2.4:1 (A:B) mixture of diastereoisomers (determined by  $^{31}\text{P}$  NMR), was purified by flash chromatography on silica gel (dichloromethane/methanol, gradient from 99/1 to 98/2) followed by a purification on Interchim Puriflash® (method C) to afford the expected product as a white solid (212 mg, 52%). For analytical purposes, a pure fraction of diastereoisomer A and diastereoisomer B were isolated and characterized separately.



Chemical Formula:  $\text{C}_{38}\text{H}_{43}\text{ClF}_2\text{N}_5\text{O}_{10}\text{PS}$   
Exact Mass: 865.21

**Diastereoisomer A:**

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.30 (s, 1H,  $\text{H}^{10'}$ ), 7.95 (d, 2H,  $J = 8.7$  Hz,  $\text{H}^7$ ), 7.82 (s, 1H,  $\text{H}^9$ ), 7.37 (d, 2H,  $J = 8.7$  Hz,  $\text{H}^8$ ), 7.35-7.29 (m, 5H,  $\text{H}^{\text{Ar}}$ ), 6.06 (br s, 2H,  $\text{NH}_2$ ), 6.01 (d, 1H,  $J = 3.6$  Hz,  $\text{H}^1$ ), 5.90 (dd, 1H,  $J = 5.1$  Hz, 3.9 Hz,  $\text{H}^2$ ), 5.77 (d, 1H,  $J = 3.6$  Hz,  $\text{H}^1$ ), 4.87 (dd, apparent t, 1H,  $J = 3.9$  Hz,  $\text{H}^2$ ), 4.76 (d, 1H,  $J = 12.0$  Hz,  $\text{H}^5$ ), 4.67 (br dd, 1H,  $J = 10.8$  Hz, 4.5 Hz,  $\text{H}^4$ ), 4.63 (d, 1H, AB syst,  $J = 11.4$  Hz,  $\text{H}^7$ ), 4.54 (d, 1H, AB syst,  $J = 11.4$  Hz,  $\text{H}^7$ ), 4.45 (dd, apparent t, 1H,  $J = 6.0$  Hz,  $\text{H}^3$ ), 4.32-4.18 (m, 2H,  $\text{H}^5$ ,  $\text{H}^9$ ), 4.15-4.09 (m, 1H,  $\text{H}^4$ ), 4.06-3.99 (m, 1H,  $\text{H}^9$ ), 3.13-2.97 (m, 1H,  $\text{H}^3$ ), 2.39-2.26 (m, 2H,  $\text{H}^6$ ), 2.23-2.13 (m, 2H,  $\text{H}^5$ ), 2.09 (s, 3H,  $\text{H}^8$ ), 1.41 (s, 3H,  $\text{H}^6$ ), 1.24 (t, 3H,  $J = 6.9$  Hz,  $\text{H}^{10}$ ), 1.01 (s, 3H,  $\text{H}^6$ ).

$^{13}\text{C}$   $\{^1\text{H}\}$  NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  " $\text{C}^{12}$ , unobserved due to multiplicity" 170.0 ( $\text{C}^{18}$ ), 165.4 ( $\text{C}^{13}$ ), 155.9 ( $\text{C}^{16}$ ), 153.4 ( $\text{C}^{10}$ ), 149.6 ( $\text{C}^{15}$ ), 139.7 (2C,  $\text{C}^{14}$ ,  $\text{C}^9$ ), 137.2 ( $\text{C}^{14}$ ), 131.3 (2C,  $\text{C}^7$ ), 128.9 (2C,  $\text{C}^8$ ),

128.7 (2C, C<sup>12'</sup>), 128.4 (2C, C<sup>15</sup>, C<sup>13'</sup>), 128.3 (2C, C<sup>11'</sup>), 120.5 (C<sup>17'</sup>), 113.2 (C<sup>11</sup>), 105.2 (C<sup>1</sup>), 87.8 (C<sup>1'</sup>), 81.9 (d, <sup>3</sup>J<sub>C-P</sub> = 16.7 Hz, C<sup>4'</sup>), 79.3 (C<sup>3'</sup>), 79.0 (d, <sup>3</sup>J<sub>C-P</sub> = 9.8 Hz, C<sup>2</sup>), 75.0 (dd, apparent t, <sup>3</sup>J<sub>C-P</sub> = 4.7 Hz, C<sup>4</sup>), 74.1 (C<sup>2'</sup>), 73.5 (C<sup>7'</sup>), 64.8 (d, <sup>4</sup>J<sub>C-P</sub> = 2.2 Hz, C<sup>5</sup>), 63.6 (d, <sup>2</sup>J<sub>C-P</sub> = 6.0 Hz, C<sup>9</sup>), 48.5-47.7 (m, C<sup>3</sup>), 28.0 (d, <sup>1</sup>J<sub>C-P</sub> = 75.2 Hz, C<sup>6'</sup>), 26.7 (C<sup>6</sup>), 26.2 (C<sup>6</sup>), 25.3 (br s, C<sup>5'</sup>), 20.9 (C<sup>8'</sup>), 16.4 (d, <sup>3</sup>J<sub>C-P</sub> = 6.0 Hz, C<sup>10</sup>).

<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>) δ -100.2 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 289.6 Hz, <sup>2</sup>J<sub>F-P</sub> = 80.9 Hz, <sup>3</sup>J<sub>F-H</sub> = 8.5 Hz), -111.4 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 289.6 Hz, <sup>2</sup>J<sub>F-P</sub> = 87.1 Hz, <sup>3</sup>J<sub>F-H</sub> = 24.0 Hz).

<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>) δ 91.2 (dd, <sup>2</sup>J<sub>P-F</sub> = 87.1 Hz, <sup>2</sup>J<sub>P-F</sub> = 80.9 Hz).

HRMS calculated for C<sub>38</sub>H<sub>44</sub>F<sub>2</sub>N<sub>5</sub>O<sub>10</sub>PS<sup>35</sup>Cl [M+H]<sup>+</sup>: 866.2203. Found 866.2216. Calculated for C<sub>38</sub>H<sub>44</sub>F<sub>2</sub>N<sub>5</sub>O<sub>10</sub>PS<sup>37</sup>Cl [M+H]<sup>+</sup>: 868.2199. Found 868.2215.

FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 2984, 1723, 1634, 1594, 1372, 1226, 1171, 1091, 1013, 757, 699, 631, 529, 422.

[α]<sub>D</sub><sup>21</sup> + 24.1 (c 1.00, CHCl<sub>3</sub>).

M.p. = 82-83 °C.

Rf<sub>A</sub> (ethyl acetate/cyclohexane, 4/1, v/v) = 0.35.

#### Diastereoisomer B:

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 8.26 (s, 1H, H<sup>10'</sup>), 7.95 (d, 2H, J = 8.7 Hz, H<sup>7</sup>), 7.80 (s, 1H, H<sup>9</sup>), 7.37 (d, 2H, J = 8.7 Hz, H<sup>8</sup>), 7.33-7.29 (m, 5H, H<sup>Ar</sup>), 6.04 (br s, 2H, NH<sub>2</sub>), 5.96 (d, 1H, J = 2.7 Hz, H<sup>1</sup>), 5.85 (dd, 1H, J = 5.4 Hz, 3.0 Hz, H<sup>2</sup>), 5.83 (d, 1H, J = 3.6 Hz, H<sup>1</sup>), 4.92 (dd, apparent t, 1H, J = 3.6 Hz, H<sup>2</sup>), 4.78 (d, 1H, J = 12.0 Hz, H<sup>5</sup>), 4.72 (br d, 1H, J = 4.2 Hz, H<sup>4</sup>), 4.63 (d, 1H, AB syst, J = 11.4 Hz, H<sup>7</sup>), 4.57 (dd, 1H, J = 6.9 Hz, 5.7 Hz, H<sup>3</sup>), 4.52 (d, 1H, AB syst, J = 11.4 Hz, H<sup>7</sup>), 4.32-4.26 (m, 1H, H<sup>5</sup>), 4.24-4.17 (m, 1H, H<sup>9</sup>), 4.16-4.09 (m, 1H, H<sup>4'</sup>), 4.06-3.99 (m, 1H, H<sup>9</sup>), 3.20-3.08 (m, 1H, H<sup>3</sup>), 2.28-2.10 (m, 3H, H<sup>5'</sup>, H<sup>6'</sup>), 2.10 (s, 3H, H<sup>8</sup>), 1.97-1.87 (m, 1H, H<sup>5'</sup>), 1.52 (s, 3H, H<sup>6</sup>), 1.27 (s, 3H, H<sup>6</sup>), 1.24 (t, 3H, J = 6.9 Hz, H<sup>10</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ "C<sup>12</sup>, unobserved due to multiplicity" 170.1 (C<sup>18'</sup>), 165.4 (C<sup>13</sup>), 155.7 (C<sup>16'</sup>), 153.0 (C<sup>10'</sup>), 149.4 (C<sup>15'</sup>), 140.0 (C<sup>9'</sup>), 139.7 (C<sup>14</sup>), 137.2 (C<sup>14'</sup>), 131.3 (2C, C<sup>7</sup>), 128.9 (2C, C<sup>8</sup>), 128.8 (2C, C<sup>12'</sup>), 128.51 (2C, C<sup>11'</sup>), 128.47 (2C, C<sup>15</sup>, C<sup>13'</sup>), 120.5 (C<sup>17'</sup>), 113.4 (C<sup>11</sup>), 105.2 (C<sup>1</sup>), 88.2 (C<sup>1'</sup>), 82.0 (d, <sup>3</sup>J<sub>C-P</sub> = 17.5 Hz, C<sup>4'</sup>), 78.8 (C<sup>2</sup> or <sup>3'</sup>), 78.7 (C<sup>2</sup> or <sup>3'</sup>), 75.0 (dd, <sup>3</sup>J<sub>C-P</sub> = 5.8 Hz, 2.5 Hz, C<sup>4</sup>), 72.2 (C<sup>2</sup>), 73.6 (C<sup>7'</sup>), 64.7 (d, <sup>4</sup>J<sub>C-P</sub> = 3.3 Hz, C<sup>5</sup>), 63.0 (d, <sup>2</sup>J<sub>C-P</sub> = 7.4 Hz, C<sup>9</sup>), 47.0-46.3 (m, C<sup>3</sup>), 27.0 (C<sup>6</sup>), 26.6 (C<sup>6</sup>), 26.4 (d, J = 73.2 Hz, C<sup>6'</sup>), 25.1 (br s, C<sup>5'</sup>), 20.9 (C<sup>8'</sup>), 16.5 (d, <sup>3</sup>J<sub>C-P</sub> = 6.3 Hz, C<sup>10</sup>).

<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>) δ -99.2 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 290.5 Hz, <sup>2</sup>J<sub>F-P</sub> = 108.3 Hz, <sup>3</sup>J<sub>F-H</sub> = 6.2 Hz), -113.1 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 290.5 Hz, <sup>2</sup>J<sub>F-P</sub> = 78.1 Hz, <sup>3</sup>J<sub>F-H</sub> = 25.9 Hz).

<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>) δ 91.6 (dd, <sup>2</sup>J<sub>P-F</sub> = 108.3 Hz, <sup>2</sup>J<sub>P-F</sub> = 78.1 Hz).

HRMS calculated for C<sub>38</sub>H<sub>44</sub>F<sub>2</sub>N<sub>5</sub>O<sub>10</sub>PS<sup>35</sup>Cl [M+H]<sup>+</sup>: 866.2203. Found 866.2217. Calculated for C<sub>38</sub>H<sub>44</sub>F<sub>2</sub>N<sub>5</sub>O<sub>10</sub>PS<sup>37</sup>Cl [M+H]<sup>+</sup>: 868.2174. Found 868.2215.

FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 2990, 1723, 1636, 1594, 1373, 1226, 1091, 1013, 798, 757, 699, 648, 528.

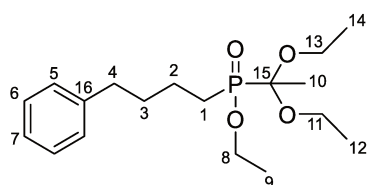
[α]<sub>D</sub><sup>21</sup> + 36.5 (c 1.00, CHCl<sub>3</sub>).

M.p. = 72-73 °C.

Rf<sub>B</sub> (ethyl acetate/cyclohexane, 4/1, v/v) = 0.25.

III.7. Ethyl (1,1-diethoxyethyl)(4-phenylbutyl)phosphinate **243**

According to a literature protocol,<sup>143</sup> to a -78 °C solution of ethyl (1,1-diethoxyethyl)phosphinate **242** (200 mg, 0.95 mmol, 1 eq) in tetrahydrofuran (7.4 mL, 0.56 M) was added a solution of KHMDS (950  $\mu$ L, 0.95 mmol, 1 eq, 1M in tetrahydrofuran) over 15 min. After stirring for 1 h, (4-iodobutyl)benzene **241** (260 mg, 0.99 mmol, 1.0 eq) was added and the resulting mixture was stirred for 24 h and allowed to warm to room temperature. After this time, the mixture was quenched with a saturated aqueous solution of ammonium chloride and extracted with ethyl acetate. The separated organic layer was washed with brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated reduced pressure. The crude residue was purified by flash chromatography on silica gel (petroleum ether 100% then ethyl acetate 100%) to afford the expected product **243** as a colorless oil (186 mg, 57%).



Chemical Formula: C<sub>18</sub>H<sub>31</sub>O<sub>4</sub>P  
Exact Mass: 342.20

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)  $\delta$  7.16-7.00 (m, 5H, H<sup>Ar</sup>), 4.11-3.93 (m, 2H, H<sup>8</sup>), 3.65-3.42 (m, 4H, H<sup>11</sup>, H<sup>13</sup>), 2.48 (t, 2H, *J* = 7.2 Hz, H<sup>4</sup>), 1.68-1.51 (m, 6H, H<sup>1</sup>, H<sup>2</sup>, H<sup>3</sup>), 1.32 (d, 3H, *J* = 11.0 Hz, H<sup>10</sup>), 1.14 (t, 3H, *J* = 7.0 Hz, H<sup>9</sup>), 1.05 (t, 6H, *J* = 7.0 Hz, H<sup>12</sup>, H<sup>14</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>)  $\delta$  142.3 (C<sup>16</sup>), 128.7 (2C, C<sup>5</sup> or <sup>6</sup>), 128.5 (2C, C<sup>5</sup> or <sup>6</sup>), 126.0 (C<sup>7</sup>), 101.4 (d, <sup>1</sup>J<sub>C-P</sub> = 137.4 Hz, C<sup>15</sup>), 61.6 (d, <sup>2</sup>J<sub>C-P</sub> = 4.7 Hz, C<sup>8</sup>), 58.3 (d, <sup>3</sup>J<sub>C-P</sub> = 4.7 Hz, C<sup>11</sup> or <sup>13</sup>), 57.8 (d, <sup>3</sup>J<sub>C-P</sub> = 7.0 Hz, C<sup>11</sup> or <sup>13</sup>), 35.8 (C<sup>4</sup>), 33.0 (d, <sup>2</sup>J<sub>C-P</sub> = 14.1 Hz, C<sup>2</sup>), 25.7 (d, <sup>1</sup>J<sub>C-P</sub> = 87.0 Hz, C<sup>1</sup>), 21.0 (d, <sup>3</sup>J<sub>C-P</sub> = 5.4 Hz, C<sup>3</sup>), 20.7 (d, <sup>2</sup>J<sub>C-P</sub> = 12.2 Hz, C<sup>10</sup>), 16.9 (d, <sup>3</sup>J<sub>C-P</sub> = 5.9 Hz, C<sup>9</sup>), 15.8 (C<sup>12</sup> or <sup>14</sup>), 15.5 (C<sup>12</sup> or <sup>14</sup>).

<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>)  $\delta$  49.0.

HRMS (TOF ESI+) calculated for C<sub>18</sub>H<sub>31</sub>O<sub>4</sub>PNa [M+Na]<sup>+</sup>: 365.1852. Found 365.1866.

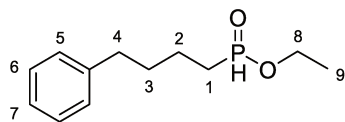
FTIR ( $\nu_{\max}$  cm<sup>-1</sup>) 3435, 2930, 1701, 1667, 1637, 1511, 1454, 1243, 1039, 743.

R<sub>f</sub> (ethyl acetate) = 0.46.

III.8. Ethyl (4-phenylbutyl)phosphinate **244a**

To a 0 °C solution of phosphinate **243** (1.91 g, 5.6 mmol, 1 eq) in ethanol (651  $\mu$ L, 11.2 mmol, 2 eq) and dichloromethane (20.0 mL) was added trimethylsilyl chloride (1.42 mL, 11.2 mmol, 2 eq). The reaction mixture was allowed to room temperature and stirred for 30 h (after 28 h, addition of 0.5 eq of trimethylsilyl chloride and ethanol). The reaction was quenched with a saturated solution of sodium bicarbonate and extracted with three times with dichloromethane, dried over MgSO<sub>4</sub> and concentrated under reduced pressure.

The crude was engaged in the next step without further purification.



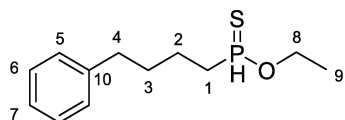
Chemical Formula: C<sub>12</sub>H<sub>19</sub>O<sub>2</sub>P  
Exact Mass: 226.11

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.25-7.09 (m, 5H, H<sup>Ar</sup>), 7.01 (br d, 1H, *J* = 526.1 Hz, PH), 4.15-3.95 (m, 2H, H<sup>8</sup>), 2.58 (t, 2H, *J* = 7.1 Hz, H<sup>4</sup>), 1.78-1.55 (m, 6H, H<sup>1</sup>, H<sup>2</sup>, H<sup>3</sup>), 1.29 (t, 3H, *J* = 7.0 Hz, H<sup>9</sup>).

<sup>31</sup>P NMR (121 MHz, CDCl<sub>3</sub>) δ 38.5 (dm, *J*<sub>P-H</sub> = 526.1 Hz).

### III.9. O-ethyl (4-phenylbutyl)phosphinothioate 244b

Phosphinate **244a** (42 mg, 0.186 mmol, 1 eq) was dissolved in toluene (1.3 mL) and Lawesson reagent (41 mg, 0.102 mmol, 0.55 eq) was added under Argon. The reaction mixture was stirred at room temperature for 5 h. After this time, toluene was removed under reduced pressure. The crude product was purified by flash chromatography on silica gel (cyclohexane/ethyl acetate, 95/5) to afford the expected product as a colorless oil (34 mg, 75%).



Chemical Formula: C<sub>12</sub>H<sub>19</sub>OPS  
Exact Mass: 242.09

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.62 (dt, 1H, *J* = 502.0 Hz, 2.9 Hz, PH), 7.24-7.18 (m, 2H, H<sup>Ar</sup>), 7.13-7.08 (m, 3H, H<sup>Ar</sup>), 4.15-4.01 (m, 1H, H<sup>8</sup>), 4.00-3.86 (m, 1H, H<sup>8</sup>), 2.57 (br t, 2H, *J* = 7.0 Hz, H<sup>4</sup>), 2.04-1.90 (m, 2H, H<sup>1</sup>), 1.71-1.57 (m, 4H, H<sup>2</sup>, H<sup>3</sup>), 1.27 (t, 3H, *J* = 7.0 Hz, H<sup>9</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ 141.9 (C<sup>10</sup>), 128.54 (2C, C<sup>5</sup> or <sup>6</sup>), 128.52 (2C, C<sup>5</sup> or <sup>6</sup>), 126.1 (C<sup>7</sup>), 62.2 (d, <sup>2</sup>*J*<sub>C-P</sub> = 6.9 Hz, C<sup>8</sup>), 35.6 (C<sup>4</sup>), 34.2 (d, <sup>1</sup>*J*<sub>C-P</sub> = 73.9 Hz, C<sup>1</sup>), 32.1 (d, <sup>2</sup>*J*<sub>C-P</sub> = 16.4 Hz, C<sup>2</sup>), 21.8 (d, <sup>3</sup>*J*<sub>C-P</sub> = 2.8 Hz, C<sup>3</sup>), 16.0 (d, <sup>3</sup>*J*<sub>C-P</sub> = 7.0 Hz, C<sup>9</sup>).

<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>) δ 75.2 (dm, *J*<sub>P-H</sub> = 502.0 Hz).

HRMS (TOF ESI+) calculated for C<sub>12</sub>H<sub>20</sub>O<sub>2</sub>P [M-S+O+H]<sup>+</sup>: 227.1201. Found 227.1203.

Elemental analysis calculated for C<sub>12</sub>H<sub>19</sub>OPS: C, 59.48; H, 7.90; S, 13.23. Found: C, 58.40; H, 7.99; S, 14.14.

FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 2930, 2320, 1739, 1496, 1453, 1020, 923, 743, 697, 622.

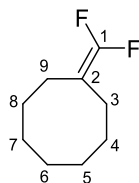
R<sub>f</sub> (petroleum ether/ethyl acetate, 95/5, v/v) = 0.25.

### III.10. (difluoromethylene)cyclooctane 246

To a -10 °C solution of cyclooctanone **245** (2.0 g, 15.8 mmol, 1 eq) in tetrahydrofuran (25.0 mL) was added dropwise dibromodifluoromethane (4.0 mL, 44.2 mmol, 2.8 eq) followed by hexamethylphosphorous triamide (13.5 mL, 74.3 mmol, 4.7 eq). *NB: the temperature should be kept below 5 °C during the addition of hexamethylphosphorous triamide.* After stirring at room temperature for 4 h, the solution was cooled at 0 °C and water was added to the reaction mixture. The aqueous layer was extracted twice with pentane. The combined organic layers were dried over Na<sub>2</sub>SO<sub>4</sub>, filtered and concentrated under vacuum at room temperature. The crude residue was purified by flash

chromatography on silica gel (pentane) to give *gem*-difluoroolefin **246** as a colorless oil (1.11 g, 44%).

**NB: 246 is highly volatile, evaporation should be done at room temperature.**



Chemical Formula: C<sub>9</sub>H<sub>14</sub>F<sub>2</sub>

Exact Mass: 160.11

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 2.02-1.97 (m, 4H, H<sup>3</sup>, H<sup>9</sup>), 1.51-1.49 (m, 4H, H<sup>4</sup>, H<sup>8</sup>), 1.41-1.38 (m, 6H, H<sup>5</sup>, H<sup>6</sup>, H<sup>7</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ 152.3 (t, <sup>1</sup>J<sub>C-F</sub> = 282.5 Hz, C<sup>1</sup>), 89.8 (t, <sup>2</sup>J<sub>C-F</sub> = 16.5 Hz, C<sup>2</sup>), 27.0 (2C, C<sup>5</sup>, C<sup>7</sup>), 26.4 (C<sup>6</sup>), 25.9 (2C, C<sup>3</sup>, C<sup>9</sup>), 25.8 (2C, C<sup>4</sup>, C<sup>8</sup>).

<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>) δ -95.8.

HRMS (IE) calculated for C<sub>9</sub>H<sub>14</sub>F<sub>2</sub>: 160.10636. Found 160.10699.

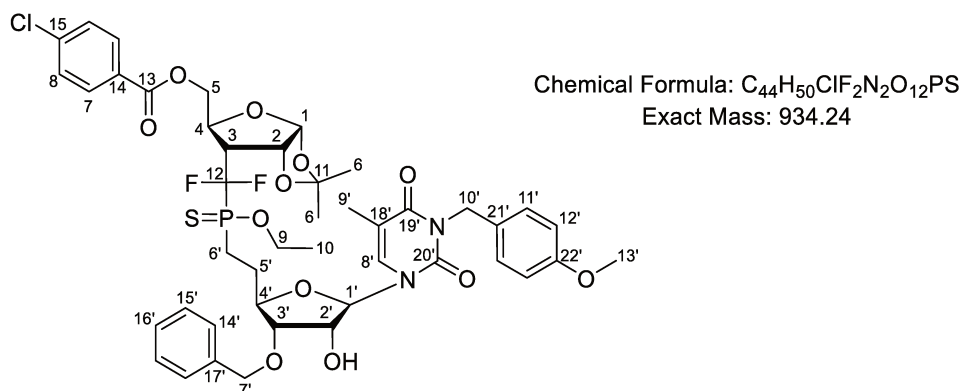
FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 2924, 2857, 1741, 1468, 1449, 1280, 1251, 1206, 1133, 1000.

R<sub>f</sub> (pentane) = 1.

## IV. Characterizations chapter 4

IV.1. (((3aR,5S,6R,6aR)-6-(((2-((2R,3S,4R,5R)-3-(benzyloxy)-4-hydroxy-5-(3-(4-methoxybenzyl)-5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-2-yl)ethyl)(ethoxy)phosphorothioyl)difluoromethyl)-2,2-dimethyltetrahydrofuro[2,3-d][1,3]dioxol-5-yl)methyl 4-chlorobenzoate **252**

To a 0 °C solution of **197d** 1.4:1 (ratio (A:B) determined by  $^{31}\text{P}$  NMR) (22 mg, 0.021 mmol, 1 eq) in mixture of tetrahydrofuran and pyridine (200  $\mu\text{L}$ /200  $\mu\text{L}$ ) was added dropwise hydrogen fluoride pyridine (20  $\mu\text{L}$ , 0.222 mmol, 10 eq). The solution was stirred at room temperature for 18 h. After this time, a saturated aqueous solution of sodium bicarbonate was added slowly to the mixture. The resulting mixture was extracted twice with ethyl acetate. The combined organic layers were washed with brine, dried over  $\text{Na}_2\text{SO}_4$  and concentrated under reduced pressure. The crude residue, as a 1.4:1 (A:B) mixture of diastereoisomers (determined by  $^{31}\text{P}$  NMR), was purified by flash chromatography on silica gel (pentane/diethyl ether, gradient from 6/4 to 3/7) to afford alcohol **252** as a white solid (16 mg, 82%). For analytical purposes, a pure fraction of diastereoisomer A was isolated and characterized separately.

**Diastereoisomer A:**

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.97 (d, 2H,  $J = 8.7$  Hz,  $\text{H}^7$ ), 7.45-7.34 (m, 9H,  $\text{H}^{11'}$ ,  $\text{H}^8$ ,  $\text{H}^{\text{Ar}}$ ), 6.98 (d, 1H,  $J = 1.2$  Hz,  $\text{H}^{8'}$ ), 6.81 (d, 2H,  $J = 8.7$  Hz,  $\text{H}^{12'}$ ), 5.80 (d, 1H,  $J = 3.7$  Hz,  $\text{H}^1$ ), 5.63 (d, 1H,  $J = 3.9$  Hz,  $\text{H}^{1'}$ ), 5.06 (d, 1H, AB syst,  $J = 13.5$  Hz,  $\text{H}^{10'}$ ), 4.97 (d, 1H, AB syst,  $J = 13.5$  Hz,  $\text{H}^{10'}$ ), 4.92 (dd, apparent t, 1H,  $J = 4.0$  Hz,  $\text{H}^2$ ), 4.78 (br d, 1H,  $J = 12.4$  Hz,  $\text{H}^5$ ), 4.72-4.68 (m, 1H,  $\text{H}^4$ ), 4.64 (s, 2H,  $\text{H}^7$ ), 4.34-4.22 (m, 3H,  $\text{H}^{2'}$ ,  $\text{H}^5$ ,  $\text{H}^9$ ), 4.14-4.06 (m, 1H,  $\text{H}^9$ ), 4.03-3.98 (m, 1H,  $\text{H}^4$ ), 3.95-3.91 (m, 1H,  $\text{H}^{3'}$ ), 3.75 (s, 3H,  $\text{H}^{13'}$ ), 3.07 (dddd, apparent dtd, 1H,  $J_{\text{H-F}} = 23.6$  Hz, 9.7 Hz,  $J = 4.6$  Hz,  $\text{H}^3$ ), 2.92 (d, 1H,  $J = 5.4$  Hz, OH), 2.42-2.28 (m, 2H,  $\text{H}^6$ ), 2.13-1.96 (m, 2H,  $\text{H}^5$ ), 1.92 (d, 3H,  $J = 1.0$  Hz,  $\text{H}^9$ ), 1.45 (s, 3H,  $\text{H}^6$ ), 1.28 (t, 3H,  $J = 6.9$  Hz,  $\text{H}^{10}$ ), 1.10 (s, 3H,  $\text{H}^6$ ).

$^{13}\text{C}$   $\{^1\text{H}\}$  NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  " $\text{C}^{12}$ , unobserved due to multiplicity" 165.5 ( $\text{C}^{13}$ ), 163.3 ( $\text{C}^{19}$ ), 159.3 ( $\text{C}^{22}$ ), 151.1 ( $\text{C}^{20}$ ), 139.8 ( $\text{C}^{14}$ ), 136.7 ( $\text{C}^{17}$ ), 134.7 ( $\text{C}^8$ ), 131.3 (2C,  $\text{C}^7$ ), 131.2 (2C,  $\text{C}^8$ ), 129.2 ( $\text{C}^{21}$ ), 129.00 (2C,  $\text{C}^{11'}$  or  $\text{C}^{15}$ ), 128.97 (2C,  $\text{C}^{11'}$  or  $\text{C}^{15}$ ), 128.8 ( $\text{C}^{16}$ ), 128.5 ( $\text{C}^{15}$ ), 128.3 (2C,  $\text{C}^{14}$ ), 113.9 (2C,  $\text{C}^{12}$ ), 113.3 ( $\text{C}^{11}$ ), 110.8 ( $\text{C}^{18}$ ), 105.2 ( $\text{C}^1$ ), 93.4 ( $\text{C}^1$ ), 81.1 (d,  $^3J_{\text{C-P}} = 17.5$  Hz,  $\text{C}^4$ ), 80.5 ( $\text{C}^3$ ), 79.1 (d,  $^3J_{\text{C-P}} = 10.2$  Hz,  $\text{C}^2$ ), 75.0 (t,  $^3J_{\text{C-P}} = 5.8$  Hz,  $\text{C}^4$ ), 73.4 ( $\text{C}^7$ ), 73.2 ( $\text{C}^2$ ), 64.8 (d,  $^4J_{\text{C-P}} = 3.1$  Hz,  $\text{C}^5$ ),

63.9 (dd,  $^2J_{C-P} = 7.5$  Hz, 2.4 Hz, C<sup>9</sup>), 55.4 (C<sup>13'</sup>), 48.3 (C<sup>3</sup>), 44.1 (C<sup>10'</sup>), 27.5 (d,  $^1J_{C-P} = 75.8$  Hz, C<sup>6'</sup>), 26.8 (C<sup>6</sup>), 26.3 (C<sup>6</sup>), 25.6 (d,  $^1J_{C-P} = 2.9$  Hz, C<sup>5'</sup>), 16.5 (d,  $^3J_{C-P} = 6.4$  Hz, C<sup>10</sup>), 13.5 (C<sup>9'</sup>).

**<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)**  $\delta$  -100.2 (ddd, 1F,  $^2J_{F-F} = 291.0$  Hz,  $^2J_{F-P} = 81.6$  Hz,  $^3J_{F-H} = 9.3$  Hz), -111.2 (ddd, 1F,  $^2J_{F-F} = 291.0$  Hz,  $^2J_{F-P} = 88.1$  Hz,  $^3J_{F-H} = 24.3$  Hz).

**<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>)**  $\delta$  91.3 (dd,  $^2J_{P-F} = 88.1$  Hz,  $^2J_{P-F} = 81.6$  Hz).

**HRMS (TOF ESI+)** calculated for C<sub>44</sub>H<sub>54</sub>F<sub>2</sub>N<sub>3</sub>O<sub>12</sub>PSCI [M+NH<sub>4</sub>]<sup>+</sup>: 952.2817. Found 952.2822.

**FTIR ( $\nu_{\max}$  cm<sup>-1</sup>)** 2918, 1644, 1512, 1245, 1014, 777, 758.

**[ $\alpha$ ]<sub>D</sub><sup>22</sup>** 27.7 (c 1.60, CHCl<sub>3</sub>).

**M.p.** = 71-72 °C.

**R<sub>f</sub>A** (pentane/diethyl ether, 3/7, v/v) = 0.42.

**Diastereoisomer B (determined by subtraction of diastereoisomer A from a 1.2:1 mixture A:B, determined by <sup>31</sup>P NMR):**

**<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)**  $\delta$  7.97 (d, 2H,  $J = 8.7$  Hz, H<sup>7</sup>), 7.45-7.31 (m, 9H, H<sup>11'</sup>, H<sup>8</sup>, H<sup>Aa</sup>), 6.97 (d, 1H,  $J = 1.1$  Hz, H<sup>8'</sup>), 6.81 (d, 2H,  $J = 8.7$  Hz, H<sup>12'</sup>), 5.87 (d, 1H,  $J = 3.6$  Hz, H<sup>1</sup>), 5.57 (d, 1H,  $J = 3.6$  Hz, H<sup>1'</sup>), 5.06 (d, 1H, AB syst,  $J = 13.4$  Hz, H<sup>10'</sup>), 4.97 (d, 1H, AB syst,  $J = 13.4$  Hz, H<sup>10</sup>), 4.95 (dd, apparent t, 1H,  $J = 4.2$  Hz, H<sup>2</sup>), 4.80 (d, 1H,  $J = 11.4$  Hz, H<sup>5</sup>), 4.72-4.67 (m, 1H, H<sup>4</sup>), 4.62 (s, 2H, H<sup>7'</sup>), 4.36-4.21 (m, 3H, H<sup>2'</sup>, H<sup>5</sup>, H<sup>9</sup>), 4.14-3.91 (m, 3H, H<sup>3'</sup>, H<sup>4'</sup>, H<sup>9</sup>), 3.75 (s, 3H, H<sup>13'</sup>), 3.22-2.95 (m, 1H, H<sup>3</sup>), 2.95 (d, 1H,  $J = 5.4$  Hz, OH), 2.41-2.28 (m, 2H, H<sup>6'</sup>), 2.12-1.97 (m, 2H, H<sup>5'</sup>), 1.91 (d, 3H,  $J = 1.1$  Hz, H<sup>9'</sup>), 1.54 (s, 3H, H<sup>6</sup>), 1.44 (s, 3H, H<sup>6</sup>), 1.30 (t, 3H,  $J = 6.6$  Hz, H<sup>10</sup>).

**<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>)**  $\delta$  "C<sup>12</sup>, unobserved due to multiplicity" 165.3 (C<sup>13</sup>), 163.3 (C<sup>19'</sup>), 159.2 (C<sup>22'</sup>), 151.0 (C<sup>20'</sup>), 139.7 (C<sup>14</sup>), 136.6 (C<sup>17'</sup>), 134.9 (C<sup>8'</sup>), 131.2 (2C, C<sup>7</sup>), 130.9 (2C, C<sup>8</sup>), 129.0 (C<sup>21'</sup>), 128.85 (2C, C<sup>11'</sup> or 15'), 128.82 (2C, C<sup>11'</sup> or 15'), 128.6 (C<sup>16'</sup>), 128.3 (C<sup>15</sup>), 128.2 (2C, C<sup>14'</sup>), 113.7 (2C, C<sup>12'</sup>), 113.3 (C<sup>11</sup>), 110.6 (C<sup>18'</sup>), 105.0 (C<sup>1</sup>), 93.8 (C<sup>1'</sup>), 81.1 (d,  $^3J_{C-P} = 17.1$  Hz, C<sup>4'</sup>), 80.3 (C<sup>3'</sup>), 78.7 (d,  $^3J_{C-P} = 10.8$  Hz, C<sup>2</sup>), 74.9 (t,  $^3J_{C-P} = 5.3$  Hz, C<sup>4</sup>), 73.3 (C<sup>7'</sup>), 72.9 (C<sup>2'</sup>), 64.6 (d,  $^4J_{C-P} = 2.8$  Hz, C<sup>5</sup>), 63.0 (d,  $^2J_{C-P} = 7.0$  Hz, C<sup>9</sup>), 55.3 (C<sup>13'</sup>), 48.4-47.7 (m, C<sup>3</sup>), 43.9 (C<sup>10'</sup>), 29.7 (C<sup>6'</sup>), 26.9 (C<sup>6</sup>), 26.5 (C<sup>6</sup>), 25.8 (d,  $^1J_{C-P} = 3.4$  Hz, C<sup>5'</sup>), 16.5 (d,  $^3J_{C-P} = 6.1$  Hz, C<sup>10</sup>), 13.2 (C<sup>9'</sup>).

**<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)**  $\delta$  -99.1 (ddd, 1F,  $^2J_{F-F} = 290.5$  Hz,  $^2J_{F-P} = 109.6$  Hz,  $^3J_{F-H} = 6.8$  Hz), -113.0 (ddd, 1F,  $^2J_{F-F} = 290.5$  Hz,  $^2J_{F-P} = 78.8$  Hz,  $^3J_{F-H} = 25.9$  Hz).

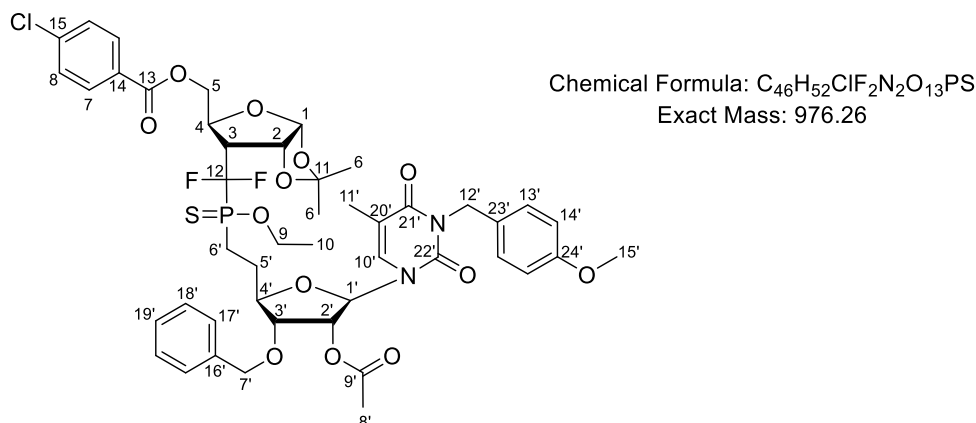
**<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>)**  $\delta$  91.6 (dd,  $^2J_{P-F} = 109.6$  Hz,  $^2J_{P-F} = 78.8$  Hz).

**R<sub>f</sub>B** (pentane/diethyl ether, 3/7, v/v) = 0.30.

IV.2. ((3aR,5S,6R,6aR)-6-(((2-((2R,3R,4R,5R)-4-acetoxy-3-(benzyloxy)-5-(3-(4-methoxybenzyl)-5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-2-yl)ethyl)(ethoxy)phosphorothioyl)difluoromethyl)-2,2-dimethyltetrahydrofuro[2,3-d][1,3]dioxol-5-yl)methyl 4-chlorobenzoate **197f**

To a solution of **252** (1.6:1 ratio (A:B) determined by <sup>31</sup>P NMR) (138 mg, 147.7  $\mu$ mol, 1 eq), triethylamine (62  $\mu$ L, 443.1  $\mu$ mol, 3 eq) and 4-dimethylaminopyridine (4 mg, 29.5  $\mu$ mol, 0.2 eq) in dichloromethane (8.0 mL) was added dropwise anhydride acetic (35  $\mu$ L, 369.2  $\mu$ mol, 2.5 eq). After addition, the mixture was stirred at room temperature for 2 h. After this time, the reaction mixture was quenched with addition of water. The separated organic phase was washed with water, a

saturated aqueous solution of sodium bicarbonate and brine, dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated under reduced pressure. The crude residue, as a 1.1:1 (A:B) mixture of diastereoisomers (determined by <sup>31</sup>P NMR), was purified by flash chromatography on silica gel (pentane/diethyl ether, gradient from 6/4 to 3/7) to afford the expected product as a white solid (120 mg, 83%). For analytical purposes, a pure fraction of diastereoisomer A was isolated and characterized separately.



#### Diastereoisomer A:

<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>) δ 7.97 (d, 2H, *J* = 8.7 Hz, H<sup>7</sup>), 7.44 (d, 2H, *J* = 8.7 Hz, H<sup>13'</sup>), 7.39 (d, 2H, *J* = 8.7 Hz, H<sup>8</sup>), 7.34-7.29 (m, 5H, H<sup>A'</sup>), 6.94 (d, 1H, *J* = 1.1 Hz, H<sup>10'</sup>), 6.82 (d, 2H, *J* = 8.7 Hz, H<sup>14'</sup>), 5.79 (d, 1H, *J* = 3.6 Hz, H<sup>1</sup>), 5.74 (d, 1H, *J* = 3.3 Hz, H<sup>1'</sup>), 5.38 (dd, 1H, *J* = 4.9 Hz, 3.4 Hz, H<sup>2</sup>), 5.06 (d, 1H, AB syst, *J* = 13.3 Hz, H<sup>12'</sup>), 4.97 (d, 1H, AB syst, *J* = 13.3 Hz, H<sup>12'</sup>), 4.92 (dd, apparent t, 1H, *J* = 3.8 Hz, H<sup>2</sup>), 4.78 (br d, 1H, *J* = 12.0 Hz, H<sup>5</sup>), 4.72-4.67 (br dd, 1H, *J* = 11.1 Hz, 4.8 Hz, H<sup>4</sup>), 4.60 (d, 1H, AB syst, *J* = 11.2 Hz, H<sup>7</sup>), 4.48 (d, 1H, AB syst, *J* = 11.2 Hz, H<sup>7</sup>), 4.33-4.21 (m, 2H, H<sup>5</sup>, H<sup>9</sup>), 4.14-3.92 (m, 3H, H<sup>3'</sup>, H<sup>4'</sup>, H<sup>9</sup>), 3.75 (s, 3H, H<sup>15'</sup>), 3.15-2.99 (m, 1H, H<sup>3</sup>), 2.45-2.26 (m, 2H, H<sup>6</sup>), 2.11-1.98 (m, 2H, H<sup>5</sup>), 2.09 (s, 3H, H<sup>8</sup>), 1.92 (d, 3H, *J* = 1.1 Hz, H<sup>11'</sup>), 1.43 (s, 3H, H<sup>6</sup>), 1.28 (t, 3H, *J* = 7.1 Hz, H<sup>10</sup>), 1.10 (s, 3H, H<sup>6</sup>).

<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>) δ "C<sup>12</sup>, unobserved due to multiplicity" 170.2 (C<sup>9'</sup>), 165.5 (C<sup>13</sup>), 163.3 (C<sup>21'</sup>), 159.3 (C<sup>24'</sup>), 150.8 (C<sup>22'</sup>), 139.8 (C<sup>14</sup>), 137.3 (C<sup>16'</sup>), 134.7 (C<sup>10'</sup>), 131.4 (2C, C<sup>7</sup>), 131.1 (2C, C<sup>8</sup>), 129.1 (C<sup>23'</sup>), 129.0 (2C, C<sup>13'</sup>), 128.8 (2C, C<sup>18'</sup>), 128.5 (C<sup>15</sup>), 128.4 (C<sup>19'</sup>), 128.2 (2C, C<sup>17'</sup>), 113.9 (2C, C<sup>14'</sup>), 113.3 (C<sup>11</sup>), 111.1 (C<sup>20'</sup>), 105.2 (C<sup>1</sup>), 91.1 (C<sup>1</sup>), 81.3 (d, <sup>3</sup>J<sub>C-P</sub> = 16.5 Hz, C<sup>4</sup>), 79.6 (C<sup>3</sup>), 79.1 (d, <sup>3</sup>J<sub>C-P</sub> = 10.0 Hz, C<sup>2</sup>), 75.1 (t, <sup>3</sup>J<sub>C-P</sub> = 5.5 Hz, C<sup>4</sup>), 74.1 (C<sup>2</sup>), 73.8 (C<sup>7</sup>), 64.8 (d, <sup>4</sup>J<sub>C-P</sub> = 3.2 Hz, C<sup>5</sup>), 63.8 (dd, <sup>2</sup>J<sub>C-P</sub> = 6.7 Hz, 1.7 Hz, C<sup>9</sup>), 55.4 (C<sup>15'</sup>), 48.3-48.0 (m, C<sup>3</sup>), 44.2 (C<sup>12'</sup>), 28.1 (d, <sup>1</sup>J<sub>C-P</sub> = 75.3 Hz, C<sup>6</sup>), 26.8 (C<sup>6</sup>), 26.3 (C<sup>6</sup>), 25.2 (br s, C<sup>5'</sup>), 21.0 (C<sup>8'</sup>), 16.5 (d, <sup>3</sup>J<sub>C-P</sub> = 6.4 Hz, C<sup>10</sup>), 13.5 (C<sup>11'</sup>).

<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>) δ -100.2 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 291.0 Hz, <sup>2</sup>J<sub>F-P</sub> = 81.3 Hz, <sup>3</sup>J<sub>F-H</sub> = 9.5 Hz), -111.3 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 291.0 Hz, <sup>2</sup>J<sub>F-P</sub> = 88.0 Hz, <sup>3</sup>J<sub>F-H</sub> = 23.6 Hz).

<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>) δ 91.2 (dd, <sup>2</sup>J<sub>P-F</sub> = 88.0 Hz, <sup>2</sup>J<sub>P-F</sub> = 81.3 Hz).

HRMS (TOF ESI+) calculated for C<sub>46</sub>H<sub>52</sub>F<sub>2</sub>N<sub>2</sub>O<sub>13</sub>NaPSCI [M+Na]<sup>+</sup>: 999.2477. Found 999.2482.

FTIR (ν<sub>max</sub> cm<sup>-1</sup>) 2924, 1650, 1512, 1373, 1229, 1092, 1014, 777, 759.

[α]<sub>D</sub><sup>22</sup> + 27.7 (c 1.00, CHCl<sub>3</sub>).

M.p. = 45-60 °C.

R<sub>f</sub>A (pentane/diethyl ether, 2/3, v/v) = 0.36.

**Diastereoisomer B (determined by subtraction of diastereoisomer A from a 1.3:1 mixture A:B, determined by  $^{31}\text{P}$  NMR):**

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  7.96 (d, 2H,  $J = 8.7$  Hz,  $\text{H}^7$ ), 7.43 (d, 2H,  $J = 8.7$  Hz,  $\text{H}^{13'}$ ), 7.39 (d, 2H,  $J = 8.7$  Hz,  $\text{H}^8$ ), 7.34-7.26 (m, 5H,  $\text{H}^{\text{Ar}}$ ), 6.93 (d, 1H,  $J = 1.1$  Hz,  $\text{H}^{10'}$ ), 6.81 (d, 2H,  $J = 8.7$  Hz,  $\text{H}^{14'}$ ), 5.86 (d, 1H,  $J = 3.6$  Hz,  $\text{H}^1$ ), 5.68 (d, 1H,  $J = 3.1$  Hz,  $\text{H}^1$ ), 5.40-5.38 (m, 1H,  $\text{H}^2$ ), 5.06 (d, 1H, AB syst,  $J = 13.5$  Hz,  $\text{H}^{12'}$ ), 4.99-4.94 (m, 1H,  $\text{H}^{12'}$ ), 4.92 (dd, apparent t, 1H,  $J = 4.0$  Hz,  $\text{H}^2$ ), 4.83-4.74 (m, 2H,  $\text{H}^4$ ,  $\text{H}^5$ ), 4.58 (d, 1H, AB syst,  $J = 11.1$  Hz,  $\text{H}^7$ ), 4.46 (d, 1H, AB syst,  $J = 11.1$  Hz,  $\text{H}^7$ ), 4.35-4.20 (m, 2H,  $\text{H}^5$ ,  $\text{H}^9$ ), 4.14-3.92 (m, 3H,  $\text{H}^3$ ,  $\text{H}^4$ ,  $\text{H}^9$ ), 3.74 (s, 3H,  $\text{H}^{15'}$ ), 3.22-2.99 (m, 1H,  $\text{H}^3$ ), 2.43-2.22 (m, 2H,  $\text{H}^6$ ), 2.17-2.02 (m, 2H,  $\text{H}^5$ ), 2.08 (s, 3H,  $\text{H}^8$ ), 1.91 (d, 3H,  $J = 1.1$  Hz,  $\text{H}^{11'}$ ), 1.54 (s, 3H,  $\text{H}^6$ ), 1.30 (s, 3H,  $\text{H}^6$ ), 1.29 (t, 3H,  $J = 7.1$  Hz,  $\text{H}^{10}$ ).

$^{13}\text{C}$  { $^1\text{H}$ } NMR (75 MHz,  $\text{CDCl}_3$ )  $\delta$  " $\text{C}^{12}$ , unobserved due to multiplicity" 170.2 ( $\text{C}^9$ ), 165.4 ( $\text{C}^{13}$ ), 163.3 ( $\text{C}^{21}$ ), 159.3 ( $\text{C}^{24}$ ), 150.8 ( $\text{C}^{22}$ ), 139.8 ( $\text{C}^{14}$ ), 137.2 ( $\text{C}^{16'}$ ), 134.8 ( $\text{C}^{10'}$ ), 131.3 (2C,  $\text{C}^7$ ), 131.0 (2C,  $\text{C}^8$ ), 129.1 ( $\text{C}^{23}$ ), 128.9 (2C,  $\text{C}^{13'}$ ), 128.7 (2C,  $\text{C}^{18'}$ ), 128.4 ( $\text{C}^{15}$ ), 128.3 ( $\text{C}^{19'}$ ), 128.2 (2C,  $\text{C}^{17'}$ ), 113.9 (2C,  $\text{C}^{14'}$ ), 113.4 ( $\text{C}^{11}$ ), 111.0 ( $\text{C}^{20'}$ ), 105.2 ( $\text{C}^1$ ), 91.7 ( $\text{C}^1$ ), 81.1 (d,  $^3J_{\text{C-P}} = 17.7$  Hz,  $\text{C}^4$ ), 79.4 ( $\text{C}^3$ ), 78.8 (d,  $^3J_{\text{C-P}} = 9.9$  Hz,  $\text{C}^2$ ), 75.1 ( $\text{C}^4$ ), 74.0 ( $\text{C}^2$ ), 73.8 ( $\text{C}^7$ ), 64.7 ( $\text{C}^5$ ), 63.1 (d,  $^2J_{\text{C-P}} = 7.2$  Hz,  $\text{C}^9$ ), 55.4 ( $\text{C}^{15'}$ ), 48.5-47.7 (m,  $\text{C}^3$ ), 44.1 ( $\text{C}^{12'}$ ), 27.3 ( $\text{C}^6$ ), 27.0 ( $\text{C}^6$ ), 26.6 ( $\text{C}^6$ ), 25.5 ( $\text{C}^5$ ), 21.0 ( $\text{C}^8$ ), 16.6 (d,  $^3J_{\text{C-P}} = 6.4$  Hz,  $\text{C}^{10}$ ), 13.4 ( $\text{C}^{11'}$ ).

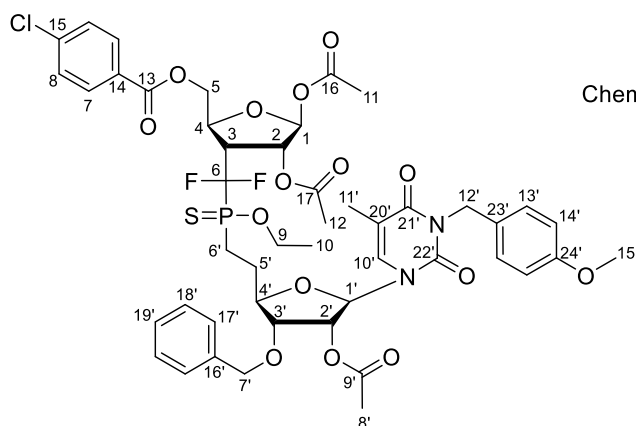
$^{19}\text{F}$  NMR (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -99.2 (ddd, 1F,  $^2J_{\text{F-F}} = 291.0$  Hz,  $^2J_{\text{F-P}} = 108.6$  Hz,  $^3J_{\text{F-H}} = 6.2$  Hz), -113.0 (ddd, 1F,  $^2J_{\text{F-F}} = 291.0$  Hz,  $^2J_{\text{F-P}} = 78.8$  Hz,  $^3J_{\text{F-H}} = 26.6$  Hz).

$^{31}\text{P}$  { $^1\text{H}$ } NMR (121 MHz,  $\text{CDCl}_3$ )  $\delta$  91.6 (dd,  $^2J_{\text{P-F}} = 108.6$  Hz,  $^2J_{\text{P-F}} = 78.8$  Hz).

$\text{Rf}_B$  (pentane/diethyl ether, 4/6, v/v) = 0.26.

IV.3. (2S,3R,4R,5S)-4-(((2-((2R,3R,4R,5R)-4-acetoxy-3-(benzyloxy)-5-(3-(4-methoxybenzyl)-5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-2-yl)ethyl)(ethoxy)phosphorothioyl)difluoromethyl)-5-(((4-chlorobenzoyl)oxy)methyl)tetrahydrofuran-2,3-diyl diacetate **182d**

To a solution of **197f** (1.4:1 ratio (A:B) determined by  $^{31}\text{P}$  NMR) (12 mg, 12.3  $\mu\text{mol}$ , 1 eq) in acetic acid (180  $\mu\text{L}$ ) were added anhydride acetic (10  $\mu\text{L}$ , 98.4  $\mu\text{mol}$ , 8 eq) and sulfuric acid (2  $\mu\text{L}$ , 36.9  $\mu\text{mol}$ , 3 eq). After addition, the mixture was stirred at room temperature for 17 h. After this time, the reaction mixture was poured into ice-water (1.1 mL) and stirred for 1.5 h. The aqueous phase was extracted twice with ethyl acetate. The separated organic phase was washed with water, a saturated aqueous solution of sodium bicarbonate and brine, dried over  $\text{MgSO}_4$  and concentrated under reduced pressure. The crude residue, as a 1.4:1 (A:B) mixture of diastereoisomers (determined by  $^{31}\text{P}$  NMR), was purified by flash chromatography on silica gel (pentane/diethyl ether, gradient from 7/3 to 6/4) to afford the expected product as a white solid (8 mg, 64%).



Chemical Formula: C<sub>47</sub>H<sub>52</sub>ClF<sub>2</sub>N<sub>2</sub>O<sub>15</sub>PS  
Exact Mass: 1020.25

**<sup>1</sup>H NMR (300 MHz, CDCl<sub>3</sub>)** (\*refers to the diastereoisomer B when unambiguous distinction is possible)  $\delta$  8.00\* (d, 0.8H,  $J$  = 8.7 Hz, H<sup>7</sup>), 7.99 (d, 1.2H,  $J$  = 8.7 Hz, H<sup>7</sup>), 7.44-7.39 (m, 4H, H<sup>8</sup>, H<sup>13'</sup>), 7.33-7.28 (m, 5H, H<sup>A'</sup>), 6.91\* (d, 0.4H,  $J$  = 1.1 Hz, H<sup>10'</sup>), 6.90 (d, 0.6H,  $J$  = 1.1 Hz, H<sup>10'</sup>), 6.81\* (d, 0.8H,  $J$  = 8.7 Hz, H<sup>14'</sup>), 6.80 (d, 1.2H,  $J$  = 8.7 Hz, H<sup>14'</sup>), 6.08 (s, 0.6H, H<sup>1</sup>), 6.06\* (d, 0.4H,  $J$  = 1.3 Hz, H<sup>1</sup>), 5.69 (d, 0.6H,  $J$  = 3.0 Hz, H<sup>1</sup>), 5.66\* (d, 0.4H,  $J$  = 3.0 Hz, H<sup>1</sup>), 5.52 (d, 0.6H,  $J$  = 4.4 Hz, H<sup>2</sup>), 5.45\* (d, 0.4H,  $J$  = 4.2 Hz, H<sup>2</sup>), 5.42-5.38 (m, 1H, H<sup>2</sup>), 5.06 (d, 1H, AB syst,  $J$  = 13.4 Hz, H<sup>12'</sup>), 4.97 (d, 1H, AB syst,  $J$  = 13.4 Hz, H<sup>12'</sup>), 4.92-4.90 (m, 1H, H<sup>4</sup>), 4.80\* (br d, 0.4H,  $J$  = 11.9 Hz, H<sup>5</sup>), 4.73 (br d, 0.6H,  $J$  = 12.3 Hz, H<sup>5</sup>), 4.60 (d, 0.6H, AB syst,  $J$  = 11.0 Hz, H<sup>7</sup>), 4.59\* (d, 0.4H, AB syst,  $J$  = 11.0 Hz, H<sup>7</sup>), 4.47 (d, 1H, AB syst,  $J$  = 11.2 Hz, H<sup>7</sup>), 4.34-4.29 (m, 1H, H<sup>5</sup>), 4.19-4.09 (m, 2H, H<sup>9</sup>), 4.05-3.92 (m, 2H, H<sup>3'</sup>, H<sup>4'</sup>), 3.75 (s, 3H, H<sup>15'</sup>), 3.51-3.44 (m, 1H, H<sup>3</sup>), 2.26-2.13 (m, 2H, H<sup>6'</sup>), 2.11-2.06 (m, 2H, H<sup>5'</sup>), 2.09 (s, 1.8H, H<sup>11, 12 or 8'</sup>), 2.08\* (s, 1.2H, H<sup>11, 12 or 8'</sup>), 2.03 (s, 3H, H<sup>11, 12 or 8'</sup>), 1.91 (d, 3H,  $J$  = 0.8 Hz, H<sup>11'</sup>), 1.90 (s, 1.8H, H<sup>11, 12 or 8'</sup>), 1.88\* (s, 1.2H, H<sup>11, 12 or 8'</sup>), 1.28 (t, 1.8H,  $J$  = 7.0 Hz, H<sup>10</sup>), 1.23\* (t, 1.2H,  $J$  = 7.0 Hz, H<sup>10</sup>).

**<sup>13</sup>C {<sup>1</sup>H} NMR (75 MHz, CDCl<sub>3</sub>)** "C<sup>3</sup>, C<sup>6</sup>, unobserved due to multiplicity" (\*refers to the diastereoisomer B when unambiguous distinction is possible)  $\delta$  170.2 (C<sup>16, 17 or 9'</sup>), 169.7 (C<sup>16, 17 or 9'</sup>), 169.0\* (C<sup>16, 17 or 9'</sup>), 168.8 (C<sup>16, 17 or 9'</sup>), 165.2 (C<sup>13</sup>), 163.3 (C<sup>21</sup>), 159.3 (C<sup>24'</sup>), 150.9 (C<sup>22'</sup>), 150.8\* (C<sup>22'</sup>), 140.04 (C<sup>14</sup>), 139.99\* (C<sup>14</sup>), 137.2 (C<sup>16'</sup>), 134.9\* (C<sup>10'</sup>), 134.6 (C<sup>10'</sup>), 131.4\* (2C, C<sup>7</sup>), 131.3 (2C, C<sup>7</sup>), 131.1\* (2C, C<sup>13'</sup>), 131.0 (2C, C<sup>8</sup>), 129.1 (C<sup>23'</sup>), 129.0 (2C, C<sup>13'</sup>), 128.8 (2C, C<sup>18'</sup>), 128.5 (C<sup>19'</sup>), 128.39 (C<sup>15</sup>), 128.36\* (2C, C<sup>17'</sup>), 128.3 (2C, C<sup>17'</sup>), 113.9 (2C, C<sup>14'</sup>), 111.1 (C<sup>20'</sup>), 98.7\* (C<sup>1</sup>), 98.6 (C<sup>1</sup>), 92.0\* (C<sup>1</sup>), 91.5 (C<sup>1</sup>), 81.1 (br s, C<sup>4</sup>), 80.9\* (br s, C<sup>4</sup>), 79.5 (C<sup>3</sup>), 79.4\* (C<sup>3</sup>), 77.4 (C<sup>4</sup>, overlap with solvent signal and deduced from DEPT experiment), 75.8 (d,  $J$  = 6.5 Hz, C<sup>2</sup>), 75.0\* (d,  $J$  = 8.3 Hz, C<sup>2</sup>), 74.0 (C<sup>2</sup>), 73.9 (C<sup>7</sup>), 65.9 (br s, C<sup>5</sup>), 63.3 (d,  $^2J_{C-P}$  = 6.6 Hz, C<sup>9</sup>), 55.4 (C<sup>15'</sup>), 44.2 (C<sup>12'</sup>), 26.95 (d,  $^1J_{C-P}$  = 71.3 Hz, C<sup>6'</sup>), 26.87\* (d,  $^1J_{C-P}$  = 69.1 Hz, C<sup>6'</sup>), 26.0\* (br s, C<sup>5'</sup>), 25.3 (br s, C<sup>5'</sup>), 21.1 (C<sup>11, 12 or 8'</sup>), 21.0 (2C, C<sup>11, 12 or 8'</sup>), 16.6-16.3 (m, C<sup>10</sup>), 13.5 (C<sup>11'</sup>).

**HRMS (TOF ESI+)** calculated for C<sub>47</sub>H<sub>52</sub>F<sub>2</sub>N<sub>2</sub>O<sub>15</sub>NaPS<sup>35</sup>Cl [M+Na]<sup>+</sup>: 1043.2380. Found 1043.2380. Calculated for C<sub>47</sub>H<sub>52</sub>F<sub>2</sub>N<sub>2</sub>O<sub>15</sub>NaPS<sup>37</sup>Cl [M+Na]<sup>+</sup>: 1045.2351. Found 1045.2382.

**FTIR** ( $\nu_{\max}$  cm<sup>-1</sup>) 3350, 2976, 2877, 1746, 1707, 1669, 1645, 1454, 1373, 1271, 1230, 1089, 1043, 1014, 777.

#### Diastereoisomer A:

**<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)**  $\delta$  -102.9 (ddd, 1F,  $^2J_{F-F}$  = 292.5 Hz,  $^2J_{F-P}$  = 80.7 Hz,  $^3J_{F-H}$  = 10.7 Hz), -111.3 (ddd, 1F,  $^2J_{F-F}$  = 292.5 Hz,  $^2J_{F-P}$  = 102.1 Hz,  $^3J_{F-H}$  = 20.9 Hz).

**<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>)**  $\delta$  91.6 (dd,  $^2J_{P-F}$  = 102.1 Hz,  $^2J_{P-F}$  = 80.7 Hz).

**R<sub>f</sub>A** (pentane/diethyl ether, 3/7, v/v) = 0.38.

**Diastereoisomer B:**

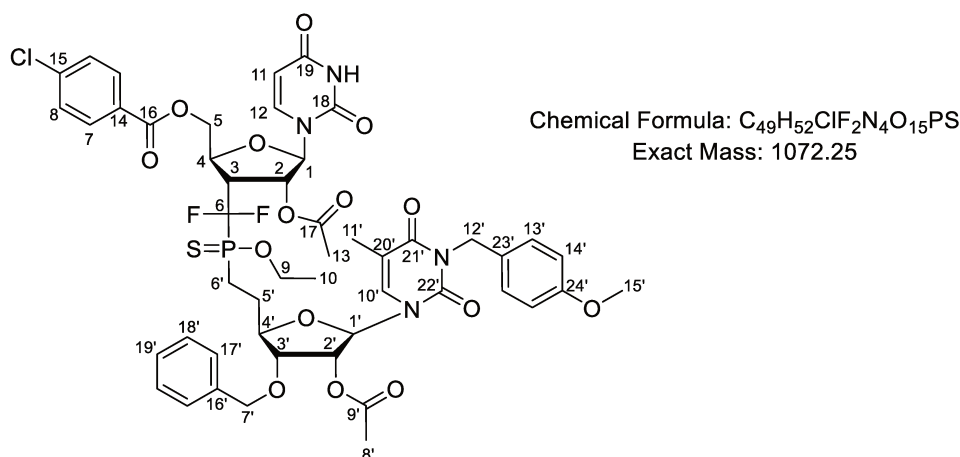
$^{19}\text{F}$  NMR (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -101.6 (ddd, 1F,  $^2J_{\text{F-F}} = 292.5$  Hz,  $^2J_{\text{F-P}} = 108.0$  Hz,  $^3J_{\text{F-H}} = 5.4$  Hz), -115.1 (ddd, 1F,  $^2J_{\text{F-F}} = 292.5$  Hz,  $^2J_{\text{F-P}} = 77.6$  Hz,  $^3J_{\text{F-H}} = 24.8$  Hz).

$^{31}\text{P}$  { $^1\text{H}$ } NMR (121 MHz,  $\text{CDCl}_3$ )  $\delta$  92.7 (dd,  $^2J_{\text{P-F}} = 108.0$  Hz,  $^2J_{\text{P-F}} = 77.6$  Hz).

$R_f_B$  (pentane/diethyl ether, 3/7, v/v) = 0.23.

IV.4. (((2S,3R,4R,5R)-4-acetoxy-3-(((2-((2R,3R,4R,5R)-4-acetoxy-3-(benzyloxy)-5-(3-(4-methoxybenzyl))-5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-2-yl)ethyl)(ethoxy)phosphorothioyl)difluoromethyl)-5-(2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl) tetrahydrofuran-2-yl)methyl 4-chlorobenzoate **183a**

According to a protocol literature,<sup>134</sup> to a solution of uracil (13 mg, 114.6  $\mu\text{mol}$ , 3 eq) in acetonitrile (170  $\mu\text{L}$ ) was added *N,O*-bis(trimethylsilyl)acetamide (61  $\mu\text{L}$ , 248.3  $\mu\text{mol}$ , 6.5 eq). The reaction mixture was stirred at 80 °C for 30 minutes. After cooling at room temperature, a solution of **182d** (2:1 ratio (A:B) determined by  $^{31}\text{P}$  NMR) (39 mg, 38.2  $\mu\text{mol}$ , 3 eq) in acetonitrile (500  $\mu\text{L}$ ) and trimethylsilyl trifluoromethanesulfonate (14  $\mu\text{L}$ , 76.4  $\mu\text{mol}$ , 2 eq) were added to the reaction mixture. After addition, the reaction mixture was stirred at 80 °C for 24 h. After this time, the reaction mixture was diluted with dichloromethane and washed with a saturated aqueous solution of sodium bicarbonate and brine, dried over  $\text{MgSO}_4$  and concentrated under reduced pressure. The crude residue, as a 1.4:1 (A:B) mixture of diastereoisomers (determined by  $^{31}\text{P}$  NMR), was purified by flash chromatography on silica gel (pentane/ethyl acetate, gradient from 6/4 to 4/6) to afford the expected product as a white solid (14 mg, 33%).



$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ) (\*refers to the diastereoisomer B when unambiguous distinction is possible)  $\delta$  8.46\* (br d, 0.4H,  $J = 1.7$  Hz, NH), 8.42 (br d, 0.6H,  $J = 1.7$  Hz, NH), 7.99-7.90 (m, 2H,  $\text{H}^7$ ), 7.38-7.35 (m, 4H,  $\text{H}^8$ ,  $\text{H}^{13'}$ ), 7.26-7.24 (m, 5H,  $\text{H}^{\text{Ar}}$ ), 7.16\* (d, 0.4H,  $J = 8.2$  Hz,  $\text{H}^{12}$ ), 7.09 (d, 0.6H,  $J = 8.2$  Hz,  $\text{H}^{12}$ ), 6.87 (br m, 1H,  $\text{H}^{10'}$ ), 6.80 (d, 2H,  $J = 8.5$  Hz,  $\text{H}^{14'}$ ), 5.74 (dd, 0.6H,  $J = 6.4$  Hz, 1.7 Hz,  $\text{H}^2$ ), 5.69-5.67\* (m, 0.4H,  $\text{H}^2$ ), 5.65 (dd, t apparent, 1H,  $J = 2.9$  Hz,  $\text{H}^1$ ), 5.57 (dd, 0.6H,  $J = 8.1$  Hz, 2.0 Hz,  $\text{H}^{11}$ ), 5.53\* (dd, 0.4H,  $J = 8.2$  Hz, 2.2 Hz,  $\text{H}^{11}$ ), 5.38-5.25 (m, 2H,  $\text{H}^1$ ,  $\text{H}^2$ ), 5.01 (d, 1H, AB syst,  $J = 13.3$  Hz,  $\text{H}^{12'}$ ), 4.92 (d, 1H, AB syst,  $J = 13.3$  Hz,  $\text{H}^{12'}$ ), 4.83-4.71 (m, 2H,  $\text{H}^5$ ,  $\text{H}^4$ ), 4.56-4.39 (m, 3H,  $\text{H}^5$ ,  $\text{H}^7$ ),

4.14-4.02 (m, 2H, H<sup>9</sup>), 3.99-3.85 (m, 3H, H<sup>3</sup>, H<sup>3'</sup>, H<sup>4</sup>), 3.70 (s, 3H, H<sup>15'</sup>), 2.15-2.11 (m, 2H, H<sup>6</sup>), 2.07-2.01 (m, 6H, H<sup>13</sup>, H<sup>8'</sup>), 1.98-1.92 (m, 2H, H<sup>5'</sup>), 1.85 (s, 3H, H<sup>11'</sup>), 1.25-1.16 (m, 3H, H<sup>10</sup>).

**<sup>13</sup>C {<sup>1</sup>H} NMR (100 MHz, CDCl<sub>3</sub>)** “C<sup>6</sup>, unobserved due to multiplicity” (\*refers to the diastereoisomer B when unambiguous distinction is possible) δ 170.6 (C<sup>17</sup> or <sup>9'</sup>), 170.4\* (C<sup>17</sup> or <sup>9'</sup>), 170.3 (C<sup>17</sup> or <sup>9'</sup>), 170.1\* (C<sup>17</sup> or <sup>9'</sup>), 165.4 (C<sup>16</sup>), 165.3\* (C<sup>16</sup>), 163.33 (C<sup>21'</sup>), 163.32\* (C<sup>21'</sup>), 162.8 (C<sup>19</sup>), 162.7\* (C<sup>19</sup>), 159.3 (C<sup>24'</sup>), 150.88\* (C<sup>22'</sup>), 150.85 (C<sup>22'</sup>), 149.73\* (C<sup>18</sup>), 149.67 (C<sup>18</sup>), 142.7 (C<sup>12</sup>), 141.7\* (C<sup>12</sup>), 140.2\* (C<sup>14</sup>), 140.1 (C<sup>14</sup>), 137.2 (C<sup>16'</sup>), 134.8\* (C<sup>10'</sup>), 134.6 (C<sup>10'</sup>), 131.6\* (2C, C<sup>7</sup>), 131.4 (2C, C<sup>7</sup>), 131.0 (2C, C<sup>8</sup>), 129.2\* (C<sup>13'</sup>), 129.1 (3C, C<sup>13'</sup>, C<sup>23'</sup>), 128.8 (2C, C<sup>18'</sup>), 128.5 (C<sup>19'</sup>), 128.3 (2C, C<sup>17'</sup>), 128.2 (C<sup>15</sup>), 113.9 (2C, C<sup>14'</sup>), 111.1 (C<sup>20'</sup>), 103.1\* (C<sup>11</sup>), 103.0 (C<sup>11</sup>), 95.3 (C<sup>1</sup>), 94.2\* (C<sup>1</sup>), 91.5\* (C<sup>1</sup>), 91.4 (C<sup>1</sup>), 81.2\* (d, <sup>3</sup>J<sub>C-P</sub> = 17.7 Hz, C<sup>4</sup>), 81.0 (d, <sup>3</sup>J<sub>C-P</sub> = 17.1 Hz, C<sup>4</sup>), 79.5 (C<sup>3</sup>), 79.3\* (C<sup>3</sup>), 77.4 (C<sup>4</sup>, overlap with solvent signal and deduced from DEPT experiment), 75.8 (d, <sup>3</sup>J<sub>C-P</sub> = 7.5 Hz, C<sup>2</sup>), 75.5\* (d, <sup>3</sup>J<sub>C-P</sub> = 7.5 Hz, C<sup>2</sup>), 74.0 (C<sup>2</sup>), 73.9 (C<sup>7</sup>), 73.8\* (C<sup>7</sup>), 64.8 (C<sup>5</sup>), 63.5 (d, <sup>1</sup>J<sub>C-P</sub> = 6.0 Hz, C<sup>9</sup>), 55.4 (C<sup>15</sup>), 44.2 (C<sup>12</sup>), 43.8-42.9 (m, C<sup>3</sup>), 27.2\* (d, <sup>1</sup>J<sub>C-P</sub> = 72.6 Hz, C<sup>6</sup>), 27.0 (d, <sup>1</sup>J<sub>C-P</sub> = 74.5 Hz, C<sup>6</sup>), 25.7\* (br s, C<sup>5</sup>), 25.2 (br s, C<sup>5</sup>), 21.2\* (2C, C<sup>13</sup>, C<sup>8'</sup>), 21.0 (2C, C<sup>13</sup>, C<sup>8'</sup>), 16.5 (m, C<sup>10</sup>), 13.5 (C<sup>11'</sup>).

**HRMS (TOF ESI-)** calculated for C<sub>49</sub>H<sub>51</sub>F<sub>2</sub>N<sub>4</sub>O<sub>15</sub>PSCI [M-H]<sup>-</sup>: 1071.2471. Found 1071.2466.

**FTIR (ν<sub>max</sub> cm<sup>-1</sup>)** 3356, 3262, 2922, 2849, 1626, 1337, 1157, 1092, 843, 697, 530, 411.

**Diastereoisomer A:**

**<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)** δ -102.1 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 291.0 Hz, <sup>2</sup>J<sub>F-P</sub> = 80.1 Hz, <sup>3</sup>J<sub>F-H</sub> = 9.3 Hz), -110.6 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 291.0 Hz, <sup>2</sup>J<sub>F-P</sub> = 101.0 Hz, <sup>3</sup>J<sub>F-H</sub> = 24.3 Hz).

**<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>)** δ 91.4 (dd, <sup>2</sup>J<sub>P-F</sub> = 101.0 Hz, <sup>2</sup>J<sub>P-F</sub> = 80.1 Hz).

**Rf<sub>A</sub>** (pentane/ethyl acetate, 1/1, v/v) = 0.22.

**Diastereoisomer B:**

**<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)** δ -102.0 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 290.5 Hz, <sup>2</sup>J<sub>F-P</sub> = 102.1 Hz, <sup>3</sup>J<sub>F-H</sub> = 9.6 Hz), -113.4 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 290.5 Hz, <sup>2</sup>J<sub>F-P</sub> = 79.2 Hz, <sup>3</sup>J<sub>F-H</sub> = 24.0 Hz).

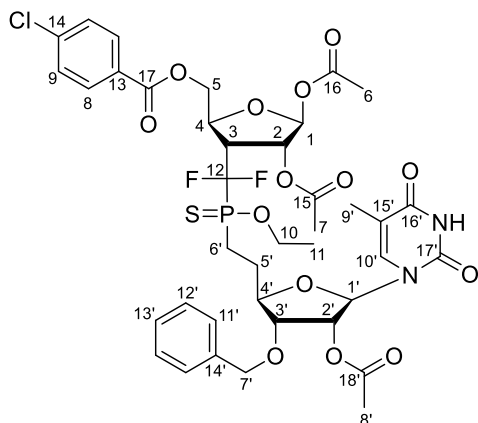
**<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>)** δ 91.5 (dd, <sup>2</sup>J<sub>P-F</sub> = 102.1 Hz, <sup>2</sup>J<sub>P-F</sub> = 79.2 Hz).

**Rf<sub>B</sub>** (pentane/ethyl acetate, 1/1, v/v) = 0.17.

IV.5. (2S,3R,4R,5S)-4-(((2-((2R,3R,4R,5R)-4-acetoxy-3-(benzyloxy)-5-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-2-yl)ethyl)(ethoxy)phosphorothioyl) difluoromethyl)-5-(((4-chlorobenzoyl)oxy)methyl)tetrahydrofuran-2,3-diyl diacetate  
**182a**

To a solution of **197a** (1.9:1 ratio (A:B) determined by <sup>31</sup>P NMR) (155 mg, 0.181 mmol, 1 eq) in acetic acid (2.5 mL) were added anhydride acetic (170 μL, 1.81 mmol, 10 eq) and sulfuric acid (30 μL, 0.543 mmol, 3 eq). After addition, the mixture was stirred at room temperature for 18 h. After this time, the reaction mixture was poured into ice-water (12.5 mL) and stirred for 1.5 h. The aqueous phase was extracted twice with ethyl acetate. The separated organic phase was washed with water, a saturated aqueous solution of sodium bicarbonate and brine, dried over MgSO<sub>4</sub> and concentrated under reduced pressure. The crude residue, as a 1.5:1 (A:B) mixture of diastereoisomers (determined

by  $^{31}\text{P}$  NMR), was purified by flash chromatography on silica gel (dichloromethane/methanol, gradient from 100/0 to 98/2) to afford the expected product as a white solid (72 mg, 44%).



Chemical Formula:  $\text{C}_{39}\text{H}_{44}\text{ClF}_2\text{N}_2\text{O}_{14}\text{PS}$   
Exact Mass: 900.19

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ) (\*refers to the diastereoisomer B when unambiguous distinction is possible)  $\delta$  9.26\* (s, 0.4H, NH), 9.22 (s, 0.6H, NH), 7.99\* (d, 0.8H,  $J = 8.6$  Hz,  $\text{H}^8$ ), 7.98 (d, 1.2H,  $J = 8.6$  Hz,  $\text{H}^8$ ), 7.394\* (d, 0.8H,  $J = 8.6$  Hz,  $\text{H}^9$ ), 7.390 (d, 1.2H,  $J = 8.6$  Hz,  $\text{H}^9$ ), 7.32-7.26 (m, 5H,  $\text{H}^{\text{Ar}}$ ), 6.95\* (d, 0.4H,  $J = 1.0$  Hz,  $\text{H}^{10'}$ ), 6.94\* (d, 0.6H,  $J = 1.1$  Hz,  $\text{H}^{10'}$ ), 6.07 (br s, 0.6H,  $\text{H}^1$ ), 6.05\* (d, 0.4H,  $J = 1.2$  Hz,  $\text{H}^1$ ), 5.61 (d, 0.6H,  $J = 3.2$  Hz,  $\text{H}^1$ ), 5.57 (d, 0.4H,  $J = 3.0$  Hz,  $\text{H}^1$ ), 5.54\* (d, 0.4H,  $J = 3.0$  Hz,  $\text{H}^2$ ), 5.51 (d, 0.6H,  $J = 4.5$  Hz,  $\text{H}^2$ ), 5.45-5.40 (m, 1H,  $\text{H}^2$ ), 4.94-4.87 (m, 1H,  $\text{H}^4$ ), 4.79\* (br d, 0.4H,  $J = 13.4$  Hz,  $\text{H}^5$ ), 4.73 (br dd, 0.6H,  $J = 12.4$  Hz, 1.1 Hz,  $\text{H}^5$ ), 4.58 (d, 0.6H, AB syst,  $J = 11.2$  Hz,  $\text{H}^7$ ), 4.57\* (d, 0.4H, AB syst,  $J = 11.2$  Hz,  $\text{H}^7$ ), 4.460 (d, 0.6H, AB syst,  $J = 11.2$  Hz,  $\text{H}^7$ ), 4.459\* (d, 0.4H, AB syst,  $J = 11.2$  Hz,  $\text{H}^7$ ), 4.35 (dd, 0.6H,  $J = 12.2$  Hz, 5.2 Hz,  $\text{H}^5$ ), 4.31\* (dd, 0.4H,  $J = 12.2$  Hz, 5.2 Hz,  $\text{H}^5$ ), 4.19-4.09 (m, 2H,  $\text{H}^{10}$ ), 4.08-4.01 (m, 1H,  $\text{H}^3$ ), 3.99-3.89 (m, 1H,  $\text{H}^4$ ), 3.76-3.42 (m, 1H,  $\text{H}^3$ ), 2.20-2.13 (m, 2H,  $\text{H}^6$ ), 2.09 (s, 1.8H,  $\text{H}^{6,7\text{ or }8}$ ), 2.07\* (s, 1.2H,  $\text{H}^{6,7\text{ or }8}$ ), 2.03 (s, 3H,  $\text{H}^{6,7\text{ or }8}$ ), 2.02-1.97 (m, 2H,  $\text{H}^5$ ), 1.90 (d, 3H,  $J = 0.8$  Hz,  $\text{H}^9$ ), 1.89 (s, 1.8H,  $\text{H}^{6,7\text{ or }8}$ ), 1.87\* (s, 1.2H,  $\text{H}^{6,7\text{ or }8}$ ), 1.28 (t, 1.8H,  $J = 7.0$  Hz,  $\text{H}^{11}$ ), 1.23\* (t, 1.2H,  $J = 7.0$  Hz,  $\text{H}^{11}$ ).

$^{13}\text{C}$   $\{^1\text{H}\}$  NMR (75 MHz,  $\text{CDCl}_3$ ) " $\text{C}^{12}$ , unobserved due to multiplicity" (\*refers to the diastereoisomer B when unambiguous distinction is possible)  $\delta$  170.3 ( $\text{C}^{18}$ ), 169.7 ( $\text{C}^{15}$ ), 168.9\* ( $\text{C}^{16}$ ), 168.8 ( $\text{C}^{16}$ ), 165.2 ( $\text{C}^{17}$ ), 163.9\* ( $\text{C}^{16}$ ), 163.9 ( $\text{C}^{16}$ ), 150.2 ( $\text{C}^{17}$ ), 150.1\* ( $\text{C}^{17}$ ), 140.0 ( $\text{C}^{13}$ ), 139.9\* ( $\text{C}^{13}$ ), 137.4\* ( $\text{C}^{10}$ ), 137.2 ( $\text{C}^{14}$ ), 137.0 ( $\text{C}^{10}$ ), 131.31\* (2C,  $\text{C}^9$ ), 131.26 (2C,  $\text{C}^9$ ), 129.0 (2C,  $\text{C}^8$ ), 128.7 (2C,  $\text{C}^{11}$ ), 128.5 ( $\text{C}^{13}$ ), 128.4 ( $\text{C}^{14}$ ), 128.3 (2C,  $\text{C}^{12}$ ), 111.7 ( $\text{C}^{15}$ ), 111.6\* ( $\text{C}^{15}$ ), 98.6 ( $\text{C}^1$ ), 91.7\* ( $\text{C}^1$ ), 91.3 ( $\text{C}^1$ ), 81.0 (d,  $^3J_{\text{C-P}} = 16.7$  Hz,  $\text{C}^4$ ), 79.3 ( $\text{C}^3$ ), 79.1\* ( $\text{C}^3$ ), 77.9 (d,  $^3J_{\text{C-P}} = 3.1$  Hz,  $\text{C}^4$ ), 75.7 (d,  $^3J_{\text{C-P}} = 7.0$  Hz,  $\text{C}^2$ ), 75.0\* (d,  $^3J_{\text{C-P}} = 8.6$  Hz,  $\text{C}^2$ ), 73.9 ( $\text{C}^2$ ), 73.8 ( $\text{C}^7$ ), 65.9\* (br s,  $\text{C}^5$ ), 65.8 (br s,  $\text{C}^5$ ), 63.2 (d,  $^2J_{\text{C-P}} = 6.5$  Hz,  $\text{C}^{10}$ ), 43.6-42.9 (m,  $\text{C}^3$ ), 26.8 (d,  $^1J_{\text{C-P}} = 73.1$  Hz,  $\text{C}^6$ ), 26.7\* (d,  $^1J_{\text{C-P}} = 70.7$  Hz,  $\text{C}^6$ ), 25.8\* (br s,  $\text{C}^5$ ), 25.2 (br s,  $\text{C}^5$ ), 21.1 ( $\text{C}^{6,7\text{ or }8}$ ), 21.0 ( $\text{C}^{6,7\text{ or }8}$ ), 20.9 ( $\text{C}^{6,7\text{ or }8}$ ), 16.5 (d,  $^3J_{\text{C-P}} = 6.2$  Hz,  $\text{C}^{11}$ ), 12.6 ( $\text{C}^9$ ).

HRMS (TOF ESI+) calculated for  $\text{C}_{39}\text{H}_{44}\text{F}_2\text{N}_2\text{O}_{14}\text{NaPS}^{35}\text{Cl}$   $[\text{M}+\text{Na}]^+$ : 923.1805. Found 923.1823. Calculated for  $\text{C}_{39}\text{H}_{44}\text{F}_2\text{N}_2\text{O}_{14}\text{NaPS}^{37}\text{Cl}$   $[\text{M}+\text{Na}]^+$ : 925.1776. Found 925.1806.

FTIR ( $\nu_{\text{max}}$   $\text{cm}^{-1}$ ) 2930, 1746, 1692, 1454, 1371, 1270, 1217, 1090, 1014, 956, 758.

#### Diastereoisomer A:

$^{19}\text{F}$  NMR (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -103.0 (ddd, 1F,  $^2J_{\text{F-F}} = 292.6$  Hz,  $^2J_{\text{F-P}} = 80.7$  Hz,  $^3J_{\text{F-H}} = 10.7$  Hz), -111.3 (ddd, 1F,  $^2J_{\text{F-F}} = 292.6$  Hz,  $^2J_{\text{F-P}} = 100.8$  Hz,  $^3J_{\text{F-H}} = 20.9$  Hz).

$^{31}\text{P}$   $\{^1\text{H}\}$  NMR (121 MHz,  $\text{CDCl}_3$ )  $\delta$  91.6 (dd,  $^2J_{\text{P-F}} = 100.8$  Hz,  $^2J_{\text{P-F}} = 80.7$  Hz).

$R_f$  (dichloromethane/methanol, 95/5, v/v) = 0.38.

**Diastereoisomer B:**

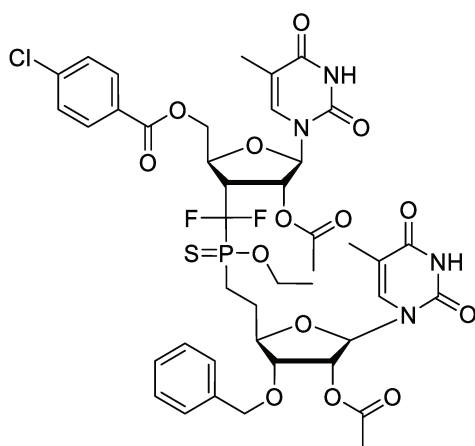
$^{19}\text{F}$  NMR (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -101.7 (ddd, 1F,  $^2J_{\text{F-F}} = 293.0$  Hz,  $^2J_{\text{F-P}} = 107.2$  Hz,  $^3J_{\text{F-H}} = 5.9$  Hz), -115.1 (ddd, 1F,  $^2J_{\text{F-F}} = 293.0$  Hz,  $^2J_{\text{F-P}} = 78.5$  Hz,  $^3J_{\text{F-H}} = 25.5$  Hz).

$^{31}\text{P}$  { $^1\text{H}$ } NMR (121 MHz,  $\text{CDCl}_3$ )  $\delta$  92.5 (dd,  $^2J_{\text{P-F}} = 107.2$  Hz,  $^2J_{\text{P-F}} = 78.5$  Hz).

$R_f$  (dichloromethane/methanol, 95/5, v/v) = 0.34.

IV.6. ((2S,3R,4R,5R)-4-acetoxy-3-(((2R,3R,4R,5R)-4-acetoxy-3-(benzyloxy)-5-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-2-yl)ethyl)(ethoxy)phosphorothioyl)difluoromethyl)-5-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-2-yl) methyl 4-chlorobenzoate **180a**

According to a protocol literature,<sup>134</sup> to a solution of thymine (11 mg, 84.6  $\mu\text{mol}$ , 1 eq) in acetonitrile (125  $\mu\text{L}$ ) was added *N,O*-bis(trimethylsilyl)acetamide (45  $\mu\text{L}$ , 183.3  $\mu\text{mol}$ , 6.5 eq). The reaction mixture was stirred at 80 °C for 5 minutes. After cooling at room temperature, a solution of **182a** (2:1 ratio (A:B) determined by  $^{31}\text{P}$  NMR) (25 mg, 28.2  $\mu\text{mol}$ , 1 eq) in acetonitrile (400  $\mu\text{L}$ ) and trimethylsilyl trifluoromethanesulfonate (10  $\mu\text{L}$ , 56.4  $\mu\text{mol}$ , 2 eq) were added to the reaction mixture. After addition, the mixture was stirred at 80 °C for 24 h. After this time, the reaction mixture was diluted with dichloromethane and washed with a saturated aqueous solution of sodium bicarbonate and brine, dried over  $\text{MgSO}_4$  and concentrated under reduced pressure. The crude residue, as a 2.4:1 (A:B) mixture of diastereoisomers (determined by  $^{31}\text{P}$  NMR), and was purified by flash chromatography on silica gel (dichloromethane/methanol, gradient from 99/1 to 98/2) to afford the expected product as a white solid (12 mg, 46%).



Chemical Formula:  $\text{C}_{42}\text{H}_{46}\text{ClF}_2\text{N}_4\text{O}_{14}\text{PS}$   
Exact Mass: 966.21

**According to literature data:<sup>107</sup>**

**HRMS (TOF ESI+)** calculated for  $\text{C}_{42}\text{H}_{46}\text{F}_2\text{N}_4\text{O}_{14}\text{PSNa}^{35}\text{Cl}$   $[\text{M}+\text{Na}]^+$ : 989.2023. Found 989.2028. Calculated for  $\text{C}_{42}\text{H}_{46}\text{F}_2\text{N}_4\text{O}_{14}\text{PSNa}^{37}\text{Cl}$   $[\text{M}+\text{Na}]^+$ : 991.1994. Found 991.2026.

**Diastereoisomer A:**

$^{19}\text{F}$  NMR (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -102.9 (dd, 1F,  $^2J_{\text{F-F}} = 291.7$  Hz,  $^2J_{\text{F-P}} = 82.5$  Hz), -110.5 (dd, 1F,  $^2J_{\text{F-F}} = 291.7$  Hz,  $^2J_{\text{F-P}} = 99.5$  Hz).

$^{31}\text{P}$  { $^1\text{H}$ } NMR (121 MHz,  $\text{CDCl}_3$ )  $\delta$  91.7 (dd,  $^2J_{\text{P-F}} = 99.5$  Hz,  $^2J_{\text{P-F}} = 82.5$  Hz).

$R_{\text{fA}}$  (dichloromethane/methanol, 95/5, v/v) = 0.38.

**Diastereoisomer B:**

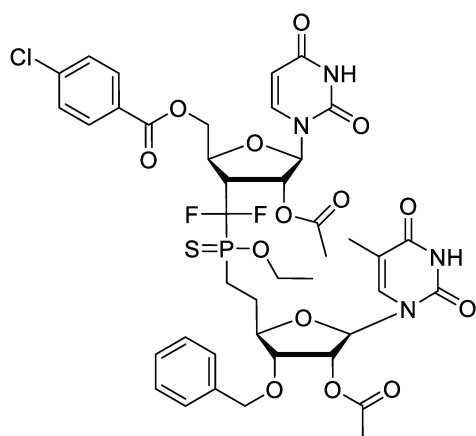
$^{19}\text{F}$  NMR (282 MHz,  $\text{CDCl}_3$ ) -99.7 (dd, 1F,  $^2J_{\text{F-F}} = 290.0$  Hz,  $^2J_{\text{F-P}} = 92.0$  Hz), -113.0 (dd, 1F,  $^2J_{\text{F-F}} = 290.0$  Hz,  $^2J_{\text{F-P}} = 81.2$  Hz).

$^{31}\text{P}$  { $^1\text{H}$ } NMR (121 MHz,  $\text{CDCl}_3$ )  $\delta$  91.0 (dd,  $^2J_{\text{P-F}} = 92.0$  Hz,  $^2J_{\text{P-F}} = 81.2$  Hz).

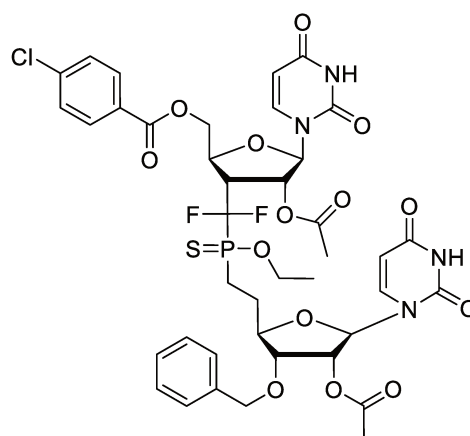
$R_{\text{fB}}$  (dichloromethane/methanol, 95/5, v/v) = 0.34.

IV.7. T-U 183b and U-U 180b

According to a protocol literature,<sup>134</sup> to a solution of uracil (27 mg, 0.240 mmol, 1 eq) in acetonitrile (400  $\mu\text{L}$ ) was added *N,O*-bis(trimethylsilyl)acetamide (130  $\mu\text{L}$ , 0.520 mmol, 6.5 eq). The reaction mixture was stirred at 60 °C for 5 minutes. After cooling at room temperature, a solution of **182a** (2:1 ratio (A:B) determined by  $^{31}\text{P}$  NMR) (72 mg, 0.080 mmol, 1 eq) in acetonitrile (1.2 mL) and trimethylsilyl trifluoromethanesulfonate (30  $\mu\text{L}$ , 0.160 mmol, 2 eq) were added to the reaction mixture. After addition, the mixture was stirred at 80 °C for 24 h. After this time, the reaction mixture was diluted with dichloromethane and washed with a saturated aqueous solution of sodium bicarbonate and brine, dried over  $\text{Na}_2\text{SO}_4$  and concentrated under reduced pressure. The crude residue was purified by flash chromatography on silica gel (dichloromethane/methanol, gradient from 99/1 to 98/2) followed by a purification on Interchim Puriflash® (method C) to afford the dinucleotide T-U **183b** as a white solid (3.5 mg, impurities present in  $^1\text{H}$  NMR) and the dinucleotide U-U **180b** as a white solid (4.2 mg, 5%).



Chemical Formula:  $\text{C}_{41}\text{H}_{44}\text{ClF}_2\text{N}_4\text{O}_{14}\text{PS}$   
Exact Mass: 952.20



Chemical Formula:  $\text{C}_{40}\text{H}_{42}\text{ClF}_2\text{N}_4\text{O}_{14}\text{PS}$   
Exact Mass: 938.18

**T-U (not pure in  $^1\text{H}$  NMR,  $^{19}\text{F}$  and  $^{31}\text{P}$  NMR were only analysed)**

**Diastereoisomer A:**

$^{19}\text{F}$  NMR (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -103.0 (ddd, 1F,  $^2J_{\text{F-F}} = 289.9$  Hz,  $^2J_{\text{F-P}} = 84.9$  Hz,  $^3J_{\text{F-H}} = 11.8$  Hz), -110.5 (ddd, 1F,  $^2J_{\text{F-F}} = 289.9$  Hz,  $^2J_{\text{F-P}} = 98.1$  Hz,  $^3J_{\text{F-H}} = 22.6$  Hz).

$^{31}\text{P}$   $\{^1\text{H}\}$  NMR (121 MHz,  $\text{CDCl}_3$ )  $\delta$  91.9 (dd,  $^2J_{\text{P-F}} = 98.1$  Hz,  $^2J_{\text{P-F}} = 84.9$  Hz).

HRMS (TOF ESI+) calculated for  $\text{C}_{41}\text{H}_{45}\text{N}_4\text{O}_{14}\text{F}_2\text{PSCl}$   $[\text{M}+\text{H}]^+$ : 953.2047. Found 953.2051.

**Diastereoisomer B:**

$^{19}\text{F}$   $\{^1\text{H}\}$  NMR (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -98.5 (dd, 1F,  $^2J_{\text{F-F}} = 289.6$  Hz,  $^2J_{\text{F-P}} = 88.3$  Hz), -111.1 (dd, 1F,  $^2J_{\text{F-F}} = 289.6$  Hz,  $^2J_{\text{F-P}} = 80.6$  Hz).

$^{31}\text{P}$   $\{^1\text{H}\}$  NMR (121 MHz,  $\text{CDCl}_3$ )  $\delta$  90.6 (dd,  $^2J_{\text{P-F}} = 88.3$  Hz,  $^2J_{\text{P-F}} = 80.6$  Hz).

**U-U (Not enough product to realized  $^{13}\text{C}$  NMR, HSQC and HMBC)**

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ )  $\delta$  8.69 (s, 1H), 8.53 (s, 1H), 8.00 (d, 2H,  $J = 8.4$  Hz), 7.41 (d, 2H,  $J = 8.4$  Hz), 7.35-7.27 (m, 5H), 7.18 (d, 1H,  $J = 8.1$  Hz), 7.14 (d, 1H,  $J = 8.1$  Hz), 5.80 (dd, 1H,  $J = 6.3$  Hz, 1.8 Hz), 5.72 (d, 1H,  $J = 8.1$  Hz), 5.62 (d, 1H,  $J = 8.4$  Hz), 5.55 (d, 1H,  $J = 3.0$  Hz), 5.43 (dd, 1H,  $J = 5.7$  Hz, 3.3 Hz), 5.35 (d, 1H,  $J = 1.8$  Hz), 4.87-4.79 (m, 2H), 4.58 (d, 1H, AB syst,  $J = 11.1$  Hz), 4.57-4.52 (m, 1H), 4.46 (d, 1H, AB syst,  $J = 11.1$  Hz), 4.24-3.86 (m, 5H), 2.21 (m, 2H), 2.10 (d, 6H,  $J = 1.2$  Hz), 2.07-1.95 (m, 2H), 1.27 (t, 3H,  $J = 6.9$  Hz).

HRMS (TOF ESI+) calculated for  $\text{C}_{40}\text{H}_{43}\text{N}_4\text{O}_{14}\text{F}_2\text{PS}^{35}\text{Cl}$   $[\text{M}+\text{H}]^+$ : 939.1891. Found 939.1882.

**Diastereoisomer A:**

$^{19}\text{F}$  NMR (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -102.7 (ddd, 1F,  $^2J_{\text{F-F}} = 292.3$  Hz,  $^2J_{\text{F-P}} = 83.2$  Hz,  $^3J_{\text{F-H}} = 10.2$  Hz), -110.5 (ddd, 1F,  $^2J_{\text{F-F}} = 292.3$  Hz,  $^2J_{\text{F-P}} = 97.6$  Hz,  $^3J_{\text{F-H}} = 22.8$  Hz).

$^{31}\text{P}$   $\{^1\text{H}\}$  NMR (121 MHz,  $\text{CDCl}_3$ )  $\delta$  91.8 (dd,  $^2J_{\text{P-F}} = 97.6$  Hz,  $^2J_{\text{P-F}} = 83.2$  Hz).

**Diastereoisomer B:**

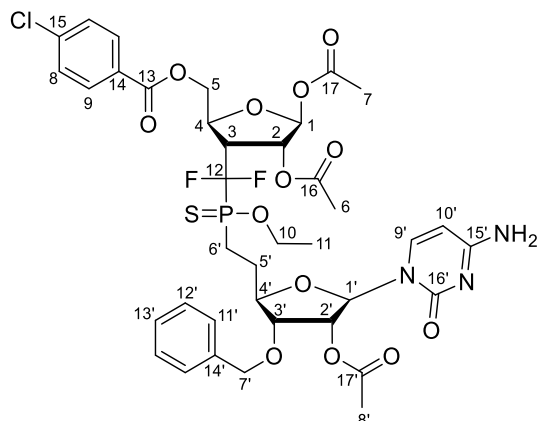
$^{19}\text{F}$  NMR (282 MHz,  $\text{CDCl}_3$ )  $\delta$  -97.7 (ddm, 1F,  $^2J_{\text{F-F}} = 289.3$  Hz,  $^2J_{\text{F-P}} = 86.3$  Hz), -113.6 (ddd, 1F,  $^2J_{\text{F-F}} = 289.3$  Hz,  $^2J_{\text{F-P}} = 81.5$  Hz,  $^3J_{\text{F-H}} = 27.6$  Hz).

$^{31}\text{P}$   $\{^1\text{H}\}$  NMR (121 MHz,  $\text{CDCl}_3$ )  $\delta$  90.3 (dd,  $^2J_{\text{P-F}} = 86.3$  Hz,  $^2J_{\text{P-F}} = 81.5$  Hz).

IV.8. (2S,3R,4R,5S)-4-(((2-((2R,3R,4R,5R)-4-acetoxy-5-(4-amino-2-oxopyrimidin-1(2H)-yl)-3-(benzyloxy)tetrahydrofuran-2-yl)ethyl)(ethoxy)phosphorothioyl)difluoromethyl)-5-(((4-chlorobenzoyl)oxy)methyl)tetrahydrofuran-2,3-diyl diacetate **182b**

To a solution of **197b** (192 mg, 0.229 mmol, 1 eq) in acetic acid (3.0 mL) were added anhydride acetic (216  $\mu\text{L}$ , 2.29 mmol, 10 eq) and sulfuric acid (36  $\mu\text{L}$ , 0.687 mmol, 3 eq). After addition, the mixture was stirred at room temperature for 16 h. After this time, the reaction mixture was quenched with water. The aqueous phase was extracted twice with ethyl acetate. The separated organic phase was washed with a saturated aqueous solution of sodium bicarbonate and brine, dried over  $\text{Na}_2\text{SO}_4$  and concentrated under reduced pressure. The crude residue was purified by flash chromatography on

silica gel (dichloromethane/methanol, 98/2) to afford the expected product as a mixture of diastereoisomers (ratio was impossible to determine by NMR) and as a white solid (46 mg, 23%).



Chemical Formula:  $C_{38}H_{43}ClF_2N_3O_{13}PS$   
Exact Mass: 885.19

**$^1H$  NMR (300 MHz,  $CDCl_3$ )** (\*refers to the diastereoisomer B when unambiguous distinction is possible)  $\delta$  7.99\* (d, 0.8H,  $J = 8.7$  Hz,  $H^9$ ), 7.98 (d, 1.2H,  $J = 8.7$  Hz,  $H^9$ ), 7.39 (d, 2H,  $J = 8.7$  Hz,  $H^8$ ), 7.29-7.23 (m, 6H,  $H^{9'}$ ,  $H^{A'}$ ), 6.07 (br s, 0.6H,  $H^1$ ), 6.04\* (br s, 0.4H,  $H^1$ ), 5.83 (d, 1H,  $J = 7.5$  Hz,  $H^{10'}$ ), 5.62 (d, 0.6H,  $J = 1.5$  Hz,  $H^1$ ), 5.57\* (d, 0.4H,  $J = 1.8$  Hz,  $H^1$ ), 5.56-5.50 (m, 1.6H,  $H^2$ ,  $H^2$ ), 5.44\* (d, 0.4H,  $J = 4.5$  Hz,  $H^2$ ), 4.94-4.87 (m, 1H,  $H^4$ ), 4.79\* (br d, 0.4H,  $J = 12.0$  Hz,  $H^5$ ), 4.73 (br d, 0.6H,  $J = 10.8$  Hz,  $H^5$ ), 4.57 (br d, 0.6H, AB syst,  $J = 11.1$  Hz,  $H^7$ ), 4.56\* (br d, 0.4H, AB syst,  $J = 11.1$  Hz,  $H^7$ ), 4.39 (d, 1H, AB syst,  $J = 11.1$  Hz,  $H^7$ ), 4.36-4.30 (m, 1H,  $H^5$ ), 4.17-4.07 (m, 2H,  $H^{10}$ ), 4.05-3.95 (m, 2H,  $H^{3'}$ ,  $H^4$ ), 3.75-3.47 (m, 1H,  $H^3$ ), 2.23-2.14 (m, 2H,  $H^6$ ), 2.08 (s, 3H,  $H^6$  or  $8'$ ), 2.07\* (s, 1.2H,  $H^6$  or  $8'$ ), 2.03 (s, 1.8H,  $H^6$  or  $8'$ ), 2.02-1.97 (m, 2H,  $H^5$ ), 1.88 (s, 1.8H,  $H^7$ ), 1.86\* (s, 1.2H,  $H^7$ ), 1.27 (d, 1.8H,  $J = 7.2$  Hz,  $H^{11}$ ), 1.23 (d, 1.2H,  $J = 6.9$  Hz,  $H^{11}$ ).

**$^{13}C$  { $^1H$ } NMR (75 MHz,  $CDCl_3$ )** " $C^{12}$ , unobserved due to multiplicity" (\*refers to the diastereoisomer B when unambiguous distinction is possible)  $\delta$  170.2\* ( $C^{16}$ ), 170.1 ( $C^{16}$ ), 169.8 ( $C^{17}$ ), 169.7\* ( $C^{17}$ ), 169.0\* ( $C^{17}$ ), 168.9 ( $C^{17}$ ), 166.1 ( $C^{15}$ ), 165.2 ( $C^{13}$ ), 155.3 ( $C^{16'}$ ), 142.5\* ( $C^9$ ), 142.0 ( $C^9$ ), 140.0 ( $C^{14}$ ), 139.9\* ( $C^{14}$ ), 137.3 ( $C^{14}$ ), 131.33\* (2C,  $C^9$ ), 131.28 (2C,  $C^9$ ), 129.0 (2C,  $C^8$ ), 128.7 (2C,  $C^{12}$ ), 128.4 (4C,  $C^{15}$ ,  $C^{11'}$ ,  $C^{13'}$ ), 98.6 ( $C^1$ ), 95.4 ( $C^{10'}$ ), 92.9\* ( $C^1$ ), 92.4 ( $C^1$ ), 80.8 (d,  $^3J_{C-P} = 17.0$  Hz,  $C^4$ ), 79.7 ( $C^3$ ), 79.4\* ( $C^3$ ), 77.9 ( $C^4$ , overlap with solvent signal and deduced from DEPT experiment), 75.7 (d,  $^3J_{C-P} = 6.0$  Hz,  $C^2$ ), 75.0\* (d,  $^3J_{C-P} = 8.0$  Hz,  $C^2$ ), 74.3 ( $C^2$ ), 73.7 ( $C^7$ ), 65.9\* ( $C^5$ ), 65.8 ( $C^5$ ), 63.3-63.0 (m,  $C^{10}$ ), 44.1-42.0 (m,  $C^3$ ), 26.9 (d,  $^1J_{C-P} = 71.9$  Hz,  $C^6$ ), 26.7\* (d,  $^1J_{C-P} = 68.9$  Hz,  $C^6$ ), 25.7\* (br s,  $C^5$ ), 25.3 (br s,  $C^5$ ), 21.3-20.8\* (m, 3C,  $C^6$ ,  $C^7$ ,  $C^8$ ), 16.5 (d,  $^3J_{C-P} = 5.5$  Hz,  $C^{11}$ ).

**HRMS** calculated for  $C_{38}H_{44}F_2N_3O_{13}PS^{35}Cl$  [ $M+H$ ] $^+$ : 886.1989. Found 886.1987. Calculated for  $C_{38}H_{44}F_2N_3O_{13}PS^{37}Cl$  [ $M+H$ ] $^+$ : 888.1960. Found 888.1982.

**FTIR** ( $\nu_{max}$   $cm^{-1}$ ) 2930, 1745, 1650, 1489, 1371, 1273, 1230, 1091, 1014, 758.

**R<sub>f</sub>** (dichloromethane/methanol, 95/5, v/v) = 0.63.

#### Diastereoisomer A:

**$^{19}F$  NMR (282 MHz,  $CDCl_3$ )**  $\delta$  -103.1 (ddd, 1F,  $^2J_{F-F} = 292.4$  Hz,  $^2J_{F-P} = 81.2$  Hz,  $^3J_{F-H} = 11.3$  Hz), -111.0 (ddd, 1F,  $^2J_{F-F} = 292.4$  Hz,  $^2J_{F-P} = 101.0$  Hz,  $^3J_{F-H} = 19.7$  Hz).

**$^{31}P$  { $^1H$ } NMR (121 MHz,  $CDCl_3$ )**  $\delta$  91.7 (dd,  $^2J_{P-F} = 101.0$  Hz,  $^2J_{P-F} = 81.2$  Hz).

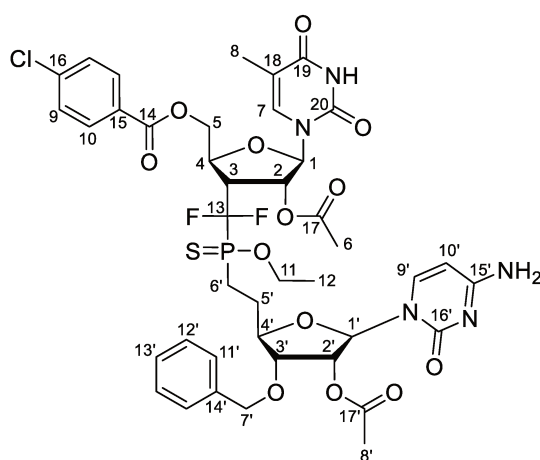
#### Diastereoisomer B:

**$^{19}F$  NMR (282 MHz,  $CDCl_3$ )**  $\delta$  -101.7 (ddd, 1F,  $^2J_{F-F} = 292.7$  Hz,  $^2J_{F-P} = 106.6$  Hz,  $^3J_{F-H} = 6.2$  Hz), -114.9 (ddd, 1F,  $^2J_{F-F} = 292.7$  Hz,  $^2J_{F-P} = 78.7$  Hz,  $^3J_{F-H} = 25.1$  Hz).

$^{31}\text{P}$  { $^1\text{H}$ } NMR (121 MHz,  $\text{CDCl}_3$ )  $\delta$  92.7 (dd,  $^2J_{\text{P-F}} = 106.6$  Hz,  $^2J_{\text{P-F}} = 78.7$  Hz).

IV.9. ((2S,3R,4R,5R)-4-acetoxy-3-(((2-((2R,3R,4R,5R)-4-acetoxy-5-(4-amino-2-oxopyrimidin-1(2H)-yl)-3-(benzyloxy)tetrahydrofuran-2-yl)ethyl)(ethoxy)phosphorothioyl) difluoromethyl)-5-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)tetrahydrofuran-2-yl)methyl 4-chlorobenzoate **183d**

According to a protocol literature,<sup>134</sup> to a solution of thymine (8.1 mg, 64.4  $\mu\text{mol}$ , 1 eq) in acetonitrile (250  $\mu\text{L}$ ) was added *N,O*-bis(trimethylsilyl)acetamide (31  $\mu\text{L}$ , 128.7  $\mu\text{mol}$ , 3 eq). The reaction mixture was stirred at 60 °C for 5 minutes. After cooling at room temperature, a solution of **182b**, ratio 1.3:1 (A:B) determined by  $^{31}\text{P}$  NMR, (38 mg, 42.9  $\mu\text{mol}$ , 1 eq) in acetonitrile (650  $\mu\text{L}$ ) and trimethylsilyl trifluoromethanesulfonate (8.5  $\mu\text{L}$ , 47.2  $\mu\text{mol}$ , 2 eq) were added to the reaction mixture. After addition, the mixture was stirred at 80 °C for 24 h. After this time, the reaction mixture was diluted with dichloromethane and washed with a saturated aqueous solution of sodium bicarbonate and brine, dried over  $\text{Na}_2\text{SO}_4$  and concentrated under reduced pressure. The crude residue was purified by flash chromatography on silica gel (dichloromethane/methanol, gradient from 100/0 to 95/5) to afford the dinucleotide C-T **183d** as a 2:1 (A:B) mixture of diastereoisomers (P center), determined by  $^{31}\text{P}$  NMR from crude residue, and as a white solid (12 mg, 30%).



Chemical Formula:  $\text{C}_{41}\text{H}_{45}\text{ClF}_2\text{N}_5\text{O}_{13}\text{PS}$   
Exact Mass: 951.21

$^1\text{H}$  NMR (300 MHz,  $\text{CDCl}_3$ ) (\*refers to the diastereoisomer B when unambiguous distinction is possible)  $\delta$  8.06-8.00 (2d, 2H,  $J = 8.4$  Hz,  $\text{H}^{10}$ ), 7.53\* (d, 0.8H,  $J = 8.4$  Hz,  $\text{H}^9$ ), 7.44 (d, 1.2H,  $J = 8.4$  Hz,  $\text{H}^9$ ), 7.29-7.23 (m, 6H,  $\text{H}^{9'}$ ,  $\text{H}^{\text{Ar}}$ ), 6.95 (d, 0.6H,  $J = 0.9$  Hz,  $\text{H}^7$ ), 6.90\* (d, 0.4H,  $J = 0.9$  Hz,  $\text{H}^7$ ), 5.90\* (d, 0.4H,  $J = 4.8$  Hz,  $\text{H}^2$ ), 5.84 (dd, 0.6H,  $J = 6.9$  Hz, 2.7 Hz,  $\text{H}^2$ ), 5.78\* (dd, 0.4H,  $J = 5.7$  Hz, 1.5 Hz,  $\text{H}^2$ ), 5.73 (d, 0.6H,  $J = 7.5$  Hz,  $\text{H}^{10'}$ ), 5.66 (d, 0.4H,  $J = 7.5$  Hz,  $\text{H}^{10'}$ ), 5.60-5.49 (m, 1.2H,  $\text{H}^{1'}$ ,  $\text{H}^2$ ), 5.35\* (d, 0.6H,  $J = 2.1$  Hz,  $\text{H}^{1'}$ ), 5.22 (d, 0.4H,  $J = 0.9$  Hz,  $\text{H}^{1'}$ ), 5.15 (s, 0.4H,  $\text{H}^{1'}$ ), 4.82-4.77 (m, 2H,  $\text{H}^4$ ,  $\text{H}^5$ ), 4.60-3.92 (m, 8H,  $\text{H}^3$ ,  $\text{H}^5$ ,  $\text{H}^{11}$ ,  $\text{H}^{3'}$ ,  $\text{H}^4$ ,  $\text{H}^7$ ), 2.21-2.15 (m, 2H,  $\text{H}^6$ ), 2.09 (s, 3.6H,  $\text{H}^6$ ,  $\text{H}^8$ ), 2.08 (s, 2.4H,  $\text{H}^6$ ,  $\text{H}^8$ ), 2.05-1.93 (m, 3H,  $\text{H}^8$ ,  $\text{H}^{5'}$ ), 1.75\* (s, 1.2H,  $\text{H}^8$ ), 1.74 (s, 0.8H,  $\text{H}^8$ ), 1.27\* (t, 1.2H,  $J = 7.2$  Hz,  $\text{H}^{12}$ ), 1.26\* (d, 1.8H,  $J = 6.9$  Hz,  $\text{H}^{12}$ ).

$^{13}\text{C}$  { $^1\text{H}$ } NMR (100 MHz,  $\text{CDCl}_3$ ) " $^{13}\text{C}$ , unobserved due to multiplicity" (\*refers to the diastereoisomer B when unambiguous distinction is possible)  $\delta$  170.7\* ( $\text{C}^{17}$  or  $17'$ ), 170.4 ( $\text{C}^{17}$  or  $17'$ ), 170.3\* ( $\text{C}^{17}$  or  $17'$ ), 170.1 ( $\text{C}^{17}$  or  $17'$ ), 166.1\* ( $\text{C}^{15}$ ), 165.8 ( $\text{C}^{15}$ ), 165.5\* ( $\text{C}^{14}$ ), 165.3 ( $\text{C}^{14}$ ), 165.0 ( $\text{C}^{19}$ ), 163.8 ( $\text{C}^{19}$ ), 155.1\* ( $\text{C}^{16}$ ), 154.8 ( $\text{C}^{16}$ ), 149.7 ( $\text{C}^{20}$ ), 149.4\* ( $\text{C}^{20}$ ), 144.8 ( $\text{C}^9$ ), 142.8 ( $\text{C}^9$ ), 140.0 ( $\text{C}^{15}$ ), 139.9\* ( $\text{C}^{15}$ ), 139.1\* ( $\text{C}^7$ ),

138.0 (C<sup>7</sup>), 137.5\* (C<sup>14'</sup>), 137.2 (C<sup>14'</sup>), 131.4\* (2C, C<sup>10</sup>), 131.2 (2C, C<sup>10</sup>), 129.3\* (2C, C<sup>12'</sup>), 129.0 (2C, C<sup>12'</sup>), 128.8 (C<sup>16</sup>), 128.6 (2C, C<sup>11'</sup>), 128.5\* (2C, C<sup>11'</sup>), 128.2\* (2C, C<sup>9</sup>), 128.1 (2C, C<sup>9</sup>), 128.04\* (C<sup>13'</sup>), 127.96 (C<sup>13'</sup>), 111.3 (C<sup>18</sup>), 111.0\* (C<sup>18</sup>), 95.5\* (C<sup>1'</sup>), 95.2 (C<sup>1</sup>), 94.8\* (C<sup>10'</sup>), 94.4\* (C<sup>10'</sup>), 94.0 (C<sup>1</sup>), 93.0 (C<sup>1</sup>), 80.5 (d, <sup>3</sup>J<sub>C-P</sub> = 17.2 Hz, C<sup>4'</sup>), 79.7\* (C<sup>3'</sup>), 79.3\* (d, <sup>3</sup>J<sub>C-P</sub> = 17.2 Hz, C<sup>4'</sup>), 79.0 (C<sup>3'</sup>), 77.5 (C<sup>4</sup>), 75.7 (d, <sup>3</sup>J<sub>C-P</sub> = 8.0 Hz, C<sup>2</sup>), 75.4 (d, <sup>3</sup>J<sub>C-P</sub> = 6.9 Hz, C<sup>2</sup>), 74.1 (C<sup>2'</sup>), 73.8\* (C<sup>2'</sup>), 73.5 (C<sup>7</sup>), 73.2\* (C<sup>7</sup>), 64.7 (C<sup>5</sup>), 64.4\* (C<sup>5</sup>), 63.3 (d, <sup>2</sup>J<sub>C-P</sub> = 6.9 Hz, C<sup>11</sup>), 63.0\* (d, <sup>2</sup>J<sub>C-P</sub> = 7.7 Hz, C<sup>11</sup>), 43.9-43.5 (m, C<sup>3</sup>), 26.4 (d, <sup>1</sup>J<sub>C-P</sub> = 74.0 Hz, C<sup>6'</sup>), 24.5 (C<sup>5'</sup>), 21.2\* (C<sup>6</sup> or 8'), 21.0 (C<sup>6</sup> or 8'), 20.94\* (C<sup>6</sup> or 8'), 20.88 (C<sup>6</sup> or 8'), 16.33\* (C<sup>12</sup>), 16.28 (C<sup>12</sup>), 12.3 (C<sup>8</sup>), 12.2\* (C<sup>8</sup>).

**HRMS (TOF ESI+)** calculated for C<sub>41</sub>H<sub>46</sub>F<sub>2</sub>N<sub>5</sub>O<sub>13</sub>PSCI [M+H]<sup>+</sup>: 952.2207. Found 952.2211.

**FTIR** (ν<sub>max</sub> cm<sup>-1</sup>) 2920, 1723, 1372, 1222, 1013, 758, 402.

**Rf** (dichloromethane/methanol, 95/5, v/v) = 0.30.

**Diastereoisomer A:**

**<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)** δ -102.4 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 292.4 Hz, <sup>2</sup>J<sub>F-P</sub> = 84.3 Hz, <sup>3</sup>J<sub>F-H</sub> = 10.7 Hz), -110.7 (ddd, 1F, <sup>2</sup>J<sub>F-F</sub> = 292.4 Hz, <sup>2</sup>J<sub>F-P</sub> = 98.4 Hz, <sup>3</sup>J<sub>F-H</sub> = 24.0 Hz).

**<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>)** δ 92.4 (dd, <sup>2</sup>J<sub>P-F</sub> = 98.4 Hz, <sup>2</sup>J<sub>P-F</sub> = 84.3 Hz).

**Diastereoisomer B:**

**<sup>19</sup>F NMR (282 MHz, CDCl<sub>3</sub>)** δ -99.6 (ddm, 1F, <sup>2</sup>J<sub>F-F</sub> = 283.0 Hz, <sup>2</sup>J<sub>F-P</sub> = 84.0 Hz), -112.6 (ddm, 1F, <sup>2</sup>J<sub>F-F</sub> = 283.0 Hz, <sup>2</sup>J<sub>F-P</sub> = 83.7 Hz).

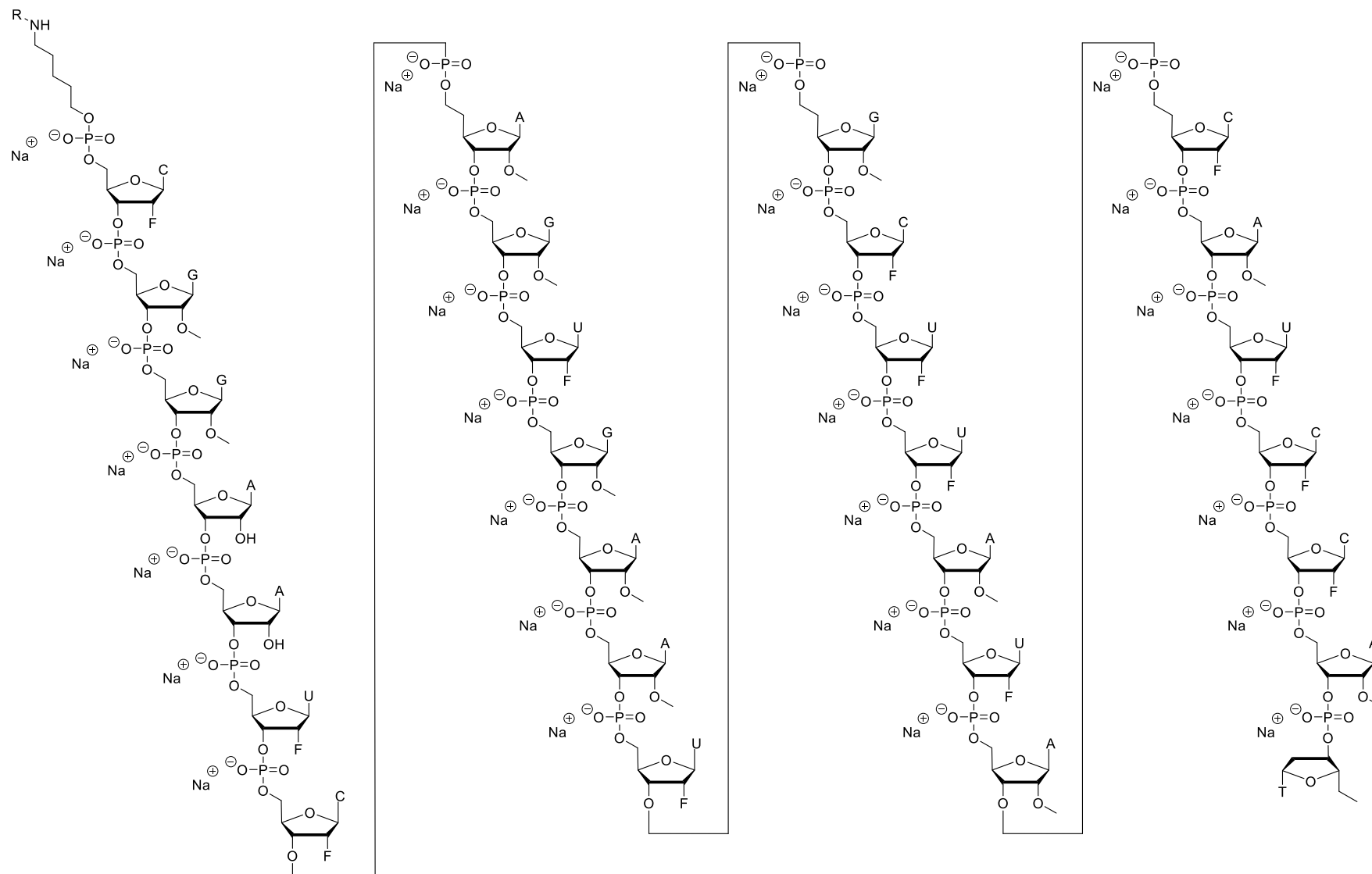
**<sup>31</sup>P {<sup>1</sup>H} NMR (121 MHz, CDCl<sub>3</sub>)** δ 90.3 (dd, <sup>2</sup>J<sub>P-F</sub> = 84.0 Hz, <sup>2</sup>J<sub>P-F</sub> = 83.7 Hz).



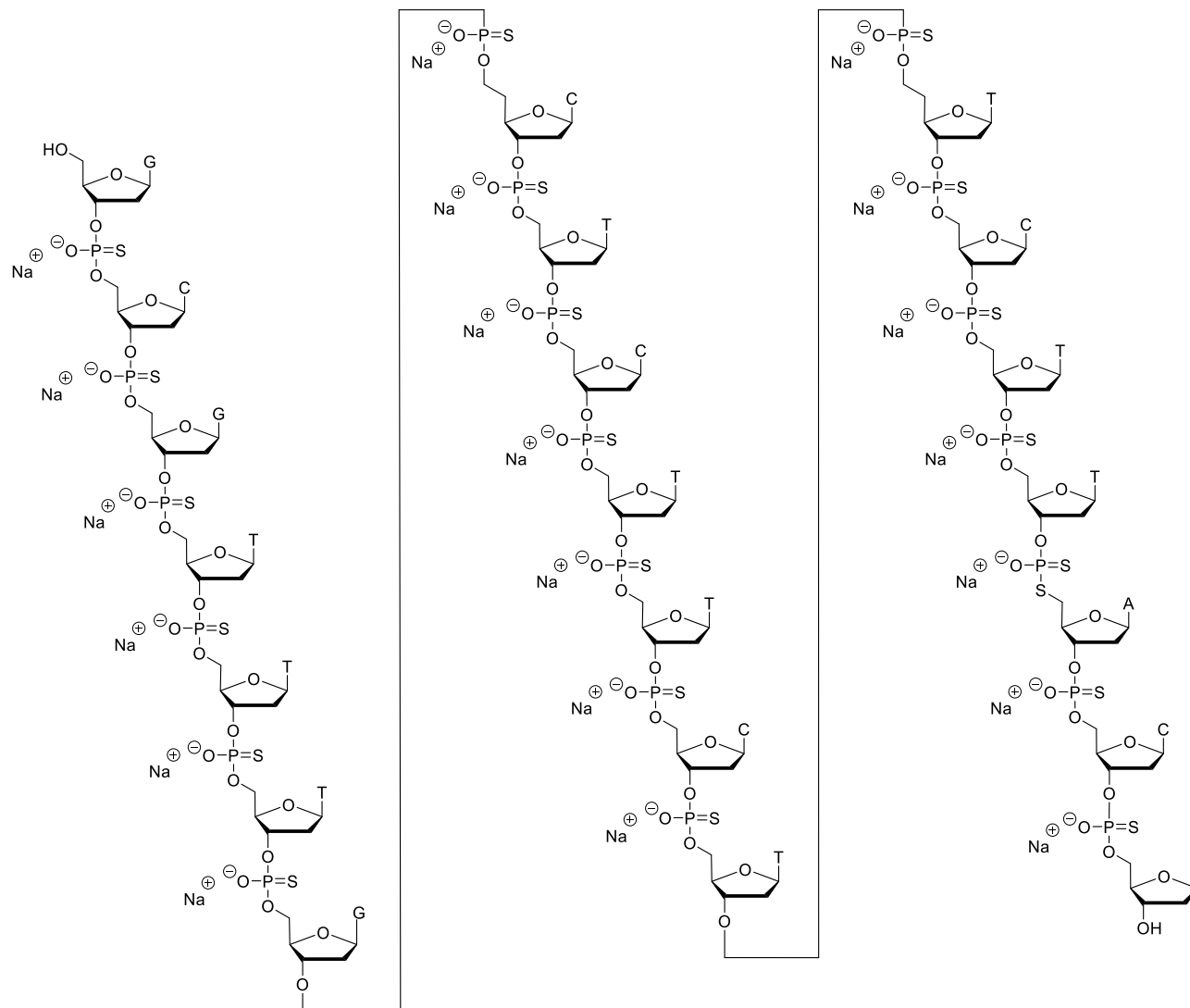
# **Annexes**



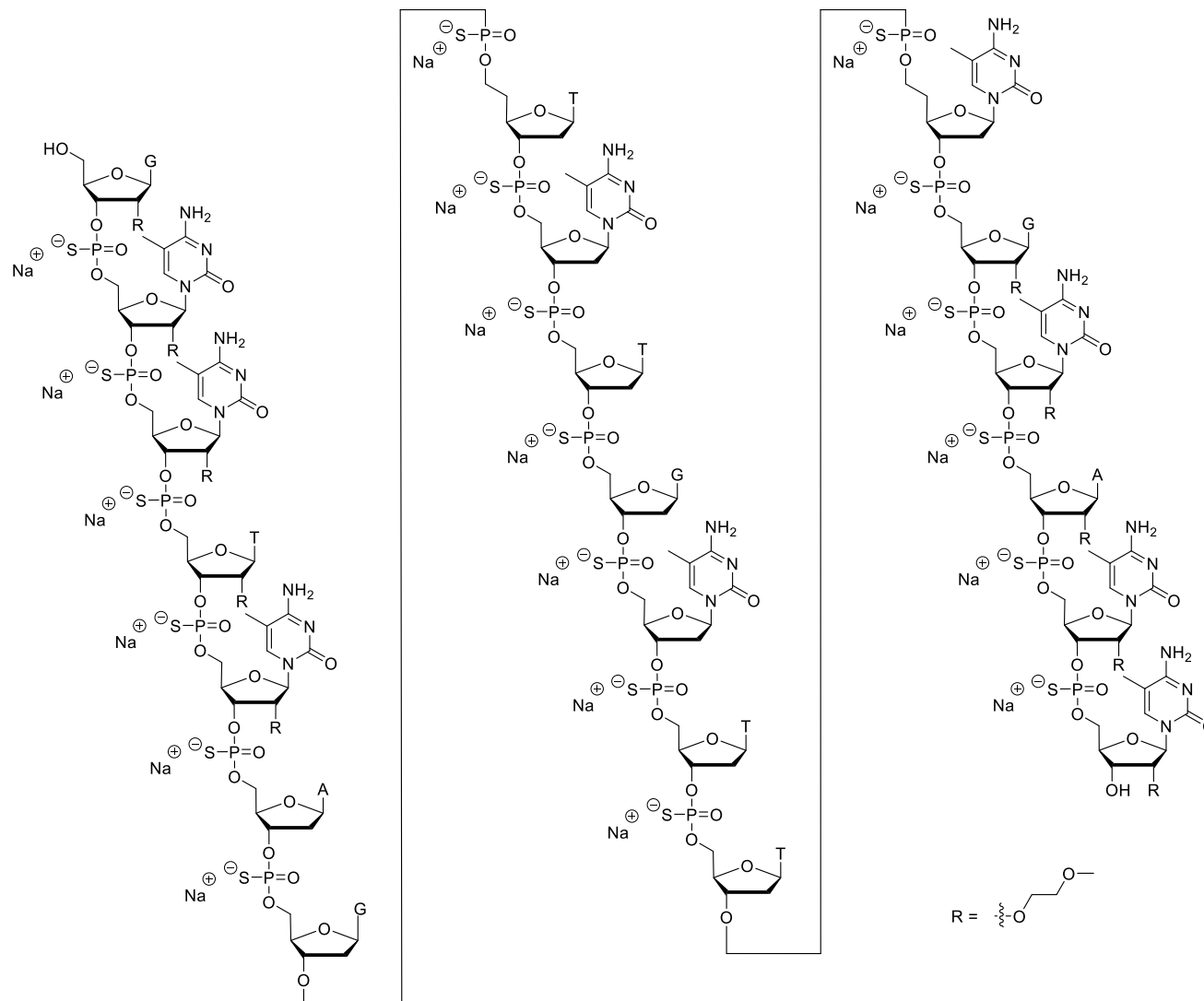
## Structure Macugen



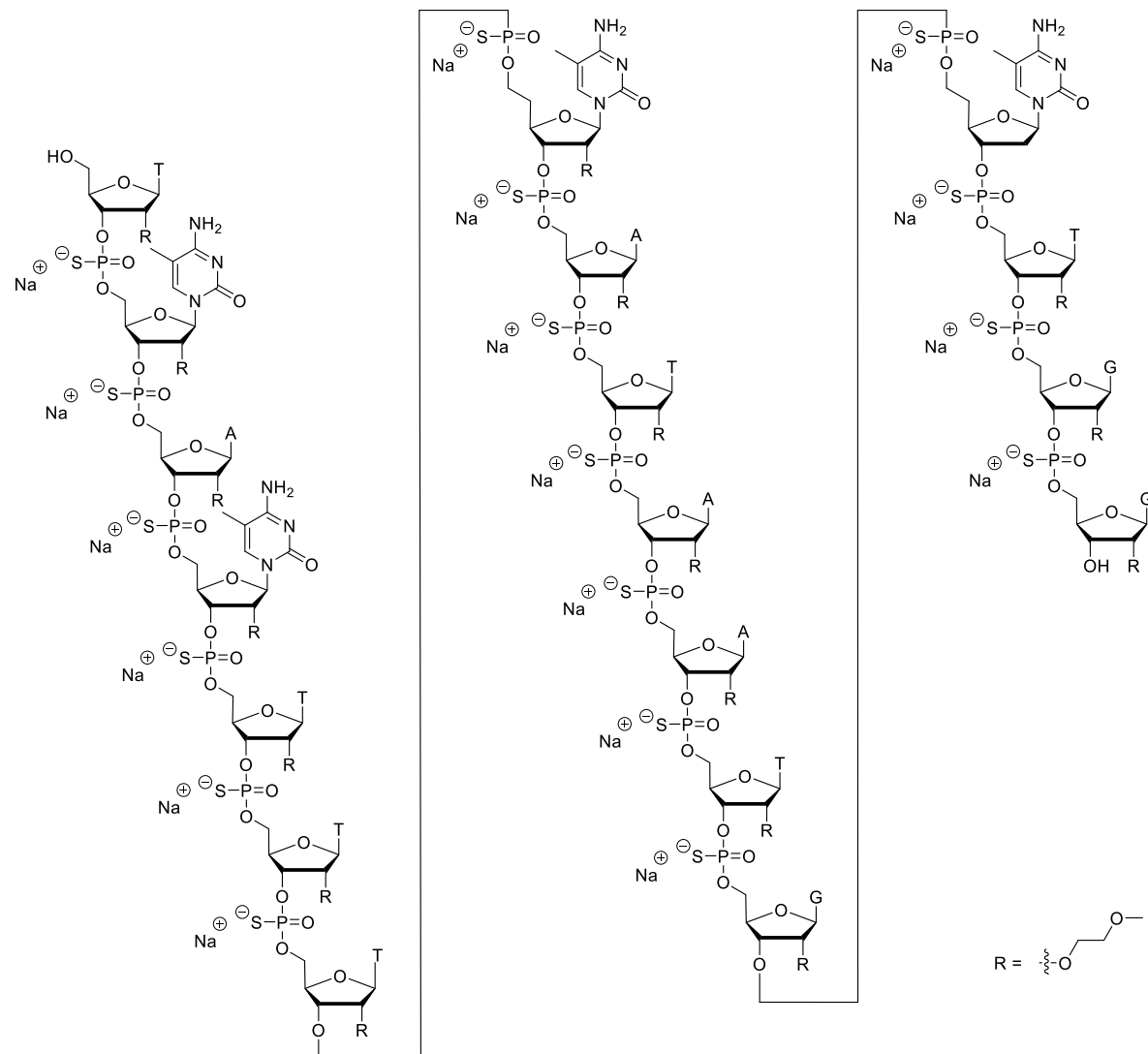
### Structure Vitravène™



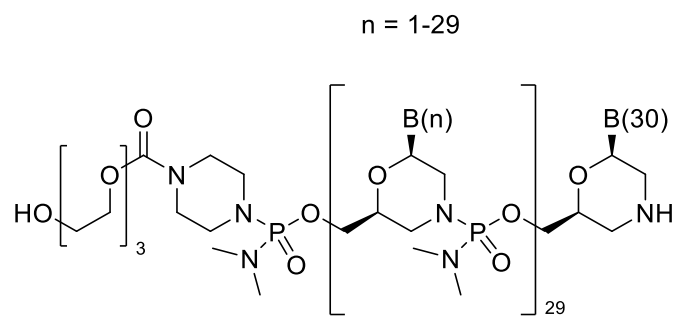
### Structure Kynamro®



### Structure Spinraza®

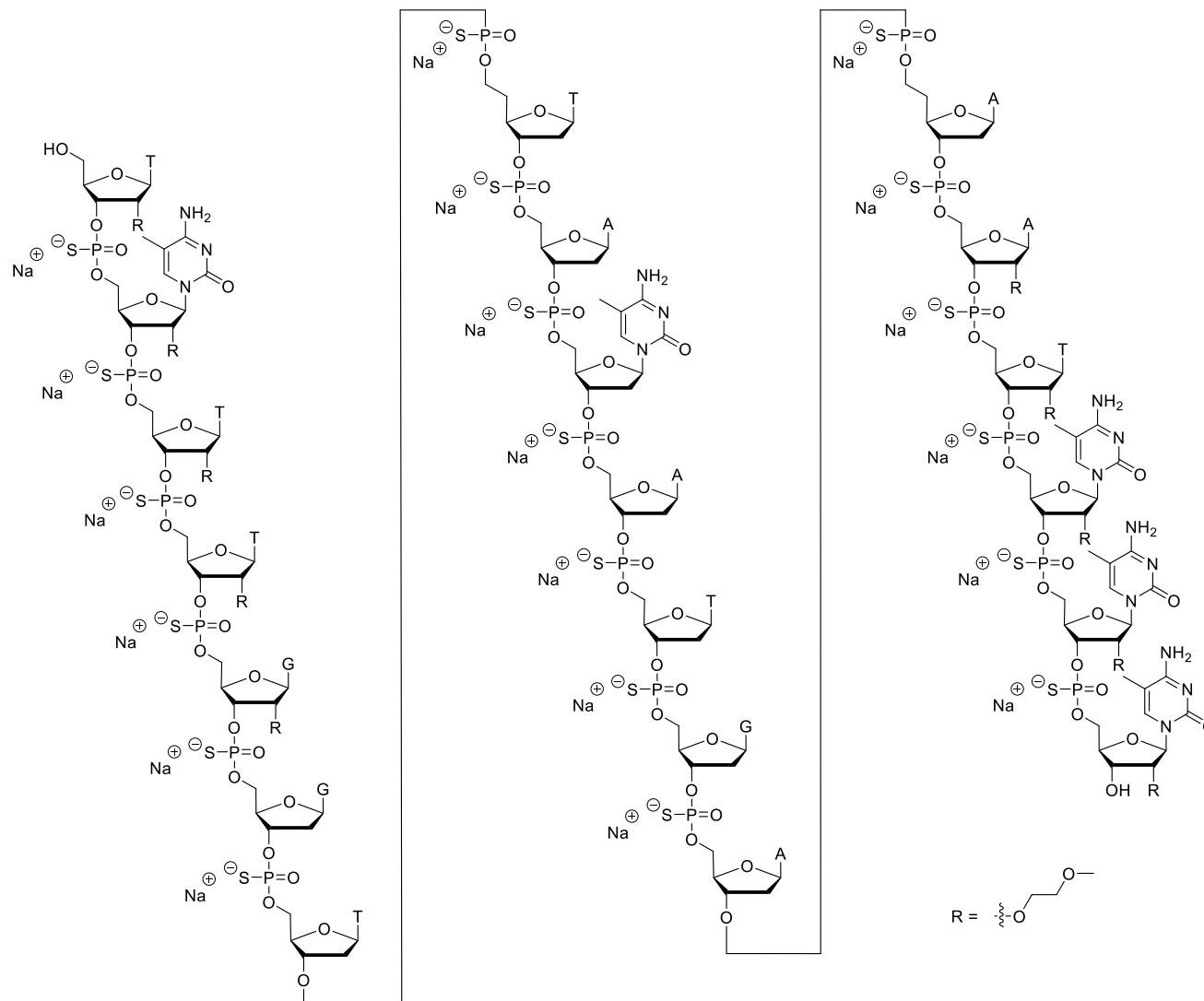


## Structure Eteplirsén



B(1-30) :  
 C-T-C-C-A-A-C-A-T-C-A-A-G-G-A  
 -A-G-A-T-G-G-C-A-T-T-T-C-T-A-G

### Structure Tegsedi



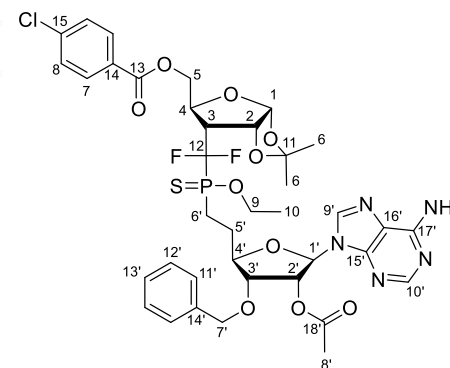
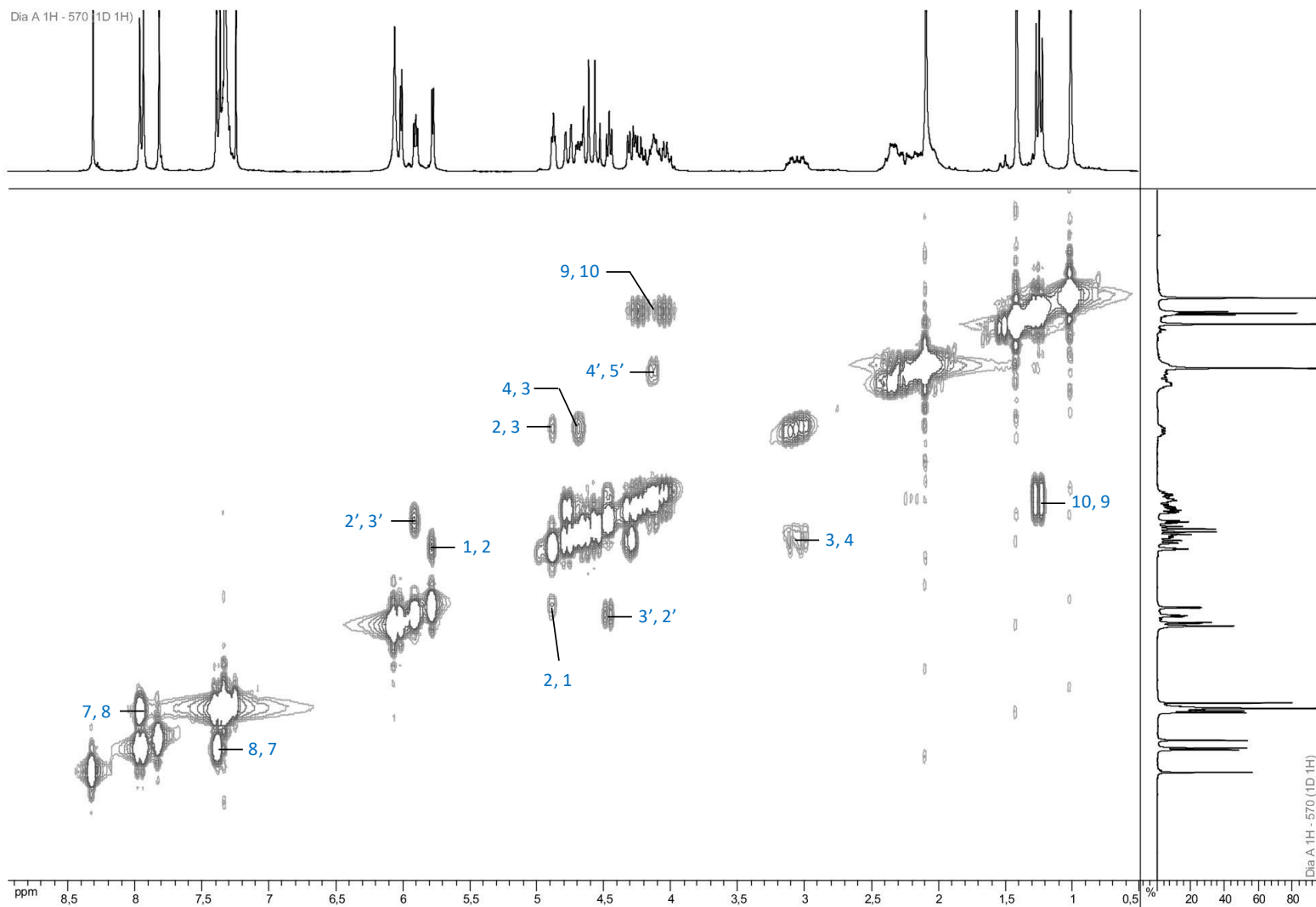
## Exemples d'attribution des signaux des spectres RMN du composé mixte adénine-furanose 197c

Dans le cas de ce composé, les spectres utilisés ne correspondent qu'au seul diastéréoisomère A.

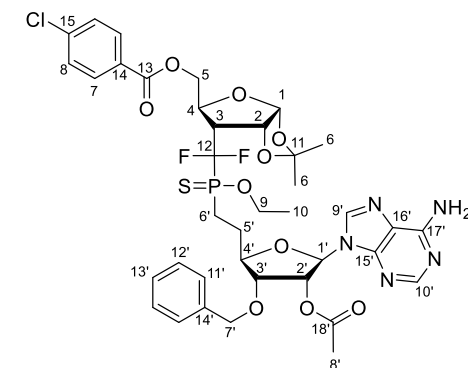
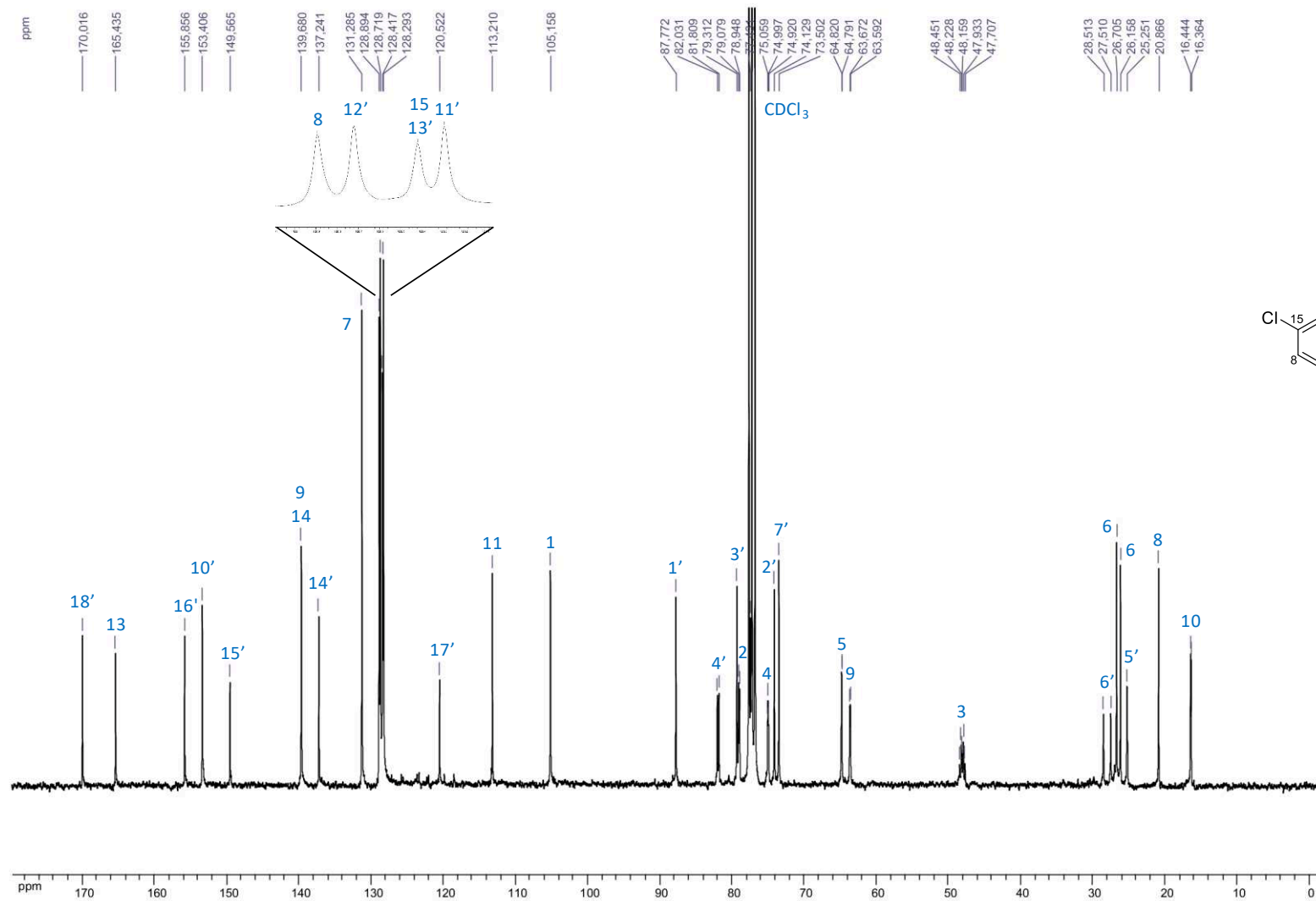
C12 n'est pas observé dans les différents spectres à cause de ses couplages avec les atomes de fluor et de phosphore. Dans le spectre RMN  $^1\text{H}$ , le triplet (1.24 ppm) correspond à H10. La corrélation du proton H10 observée en COSY permet de déterminer la position de H9 (4.32-4.18 ppm et 4.06-3.99 ppm). Les signaux de H6 (1.01 ppm et 1.41 ppm) corrélient avec le signal de C11 (113.2 ppm) en HMBC. C11 permet de déterminer la position de H1 (5.77 ppm) en HMBC. A partir de H1 et de l'expérience COSY, les protons H2 (4.87 ppm), H3 (3.13-2.97 ppm) et H4 (4.67 ppm) peuvent être placés sur le spectre  $^1\text{H}$  et donc sur le spectre  $^{13}\text{C}$  à partir de la HSQC. Le singulet à 2.09 ppm sur le spectre  $^1\text{H}$  est celui de H8'. En HMBC, ce signal permet d'établir la position du carbonyle C18' (170.0 ppm) et ce dernier détermine la position de H2' (5.90 ppm). En COSY, H2' aide au placement de H3' (4.45 ppm) qui est confirmé en HMBC car ce dernier corréle avec H7' (4.63 ppm et 4.54 ppm). Par ailleurs, H7' permet d'élucider les déplacements de C11' (128.3 ppm) et C14' (137.2 ppm). En HMBC, H2' corréle avec deux autres signaux qui correspondent à C4' (81.9 ppm) et C1' (105.2 ppm). L'analyse HSQC permet donc de déterminer les protons H4' (4.13 ppm) et H1' (5.77 ppm). La corrélation de H1' (5.77 ppm), en HMBC, avec les signaux de l'adénine C9' (139.7 ppm) et C15' (149.6 ppm) permet de confirmer le déplacement chimique de H1' sur le spectre  $^1\text{H}$ . Le signal à 6.06 ppm est celui de  $\text{NH}_2$  de l'adénine car aucune correspondance avec un carbone n'est observée en HSQC. En HMBC,  $\text{NH}_2$  permet donc de positionner C17' (120.5 ppm). Deux massifs sont présents entre 2.39 ppm et 2.13 ppm, ils correspondent au H5' et H6'. Le massif de H5' (2.23-2.13 ppm) est validé par la corrélation avec H4' (4.15-4.09 ppm) en COSY et donc par déduction, le massif de H6' correspond au second (2.39-2.26 ppm). Concernant l'adénine, deux signaux caractéristiques sont observés à 8.30 ppm et 7.82 ppm. Le second a déjà été défini comme appartenant à H9' donc le premier correspond à H10'. H10' permet ainsi de déterminer la valeur du signal de C16' (155.9 ppm) en HMBC. Pour finir avec le furanose du bas, C12' est placé à 128.7 ppm et C13' à 128.4 ppm. L'intensité du signal de C12' est supérieure à C13' car il correspond à deux carbones. A ce stade, il ne reste donc que les signaux H5, H7, H8, C13, C14 et C15 à attribuer. Le C13, étant un carbonyle, est placé à 165.4 ppm sur le spectre  $^{13}\text{C}$ . Grâce à ce dernier, les positions de H5 (4.32-4.18 ppm) et de H7 (7.95 ppm) peuvent être déterminées sur le spectre  $^1\text{H}$ . En COSY, H7 corréle avec un signal à 7.37 ppm correspondant à H8. Enfin en HMBC, H7 permet de déterminer la position de C14 (139.7 ppm) sur le spectre  $^{13}\text{C}$  et H8 la position de C15 (128.4 ppm).



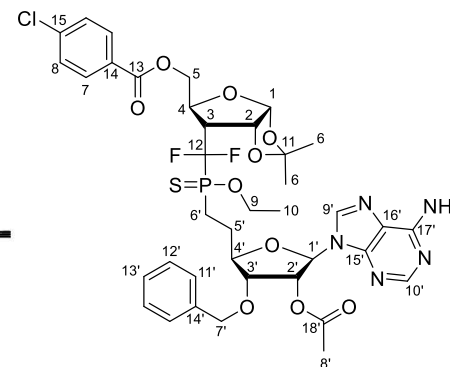
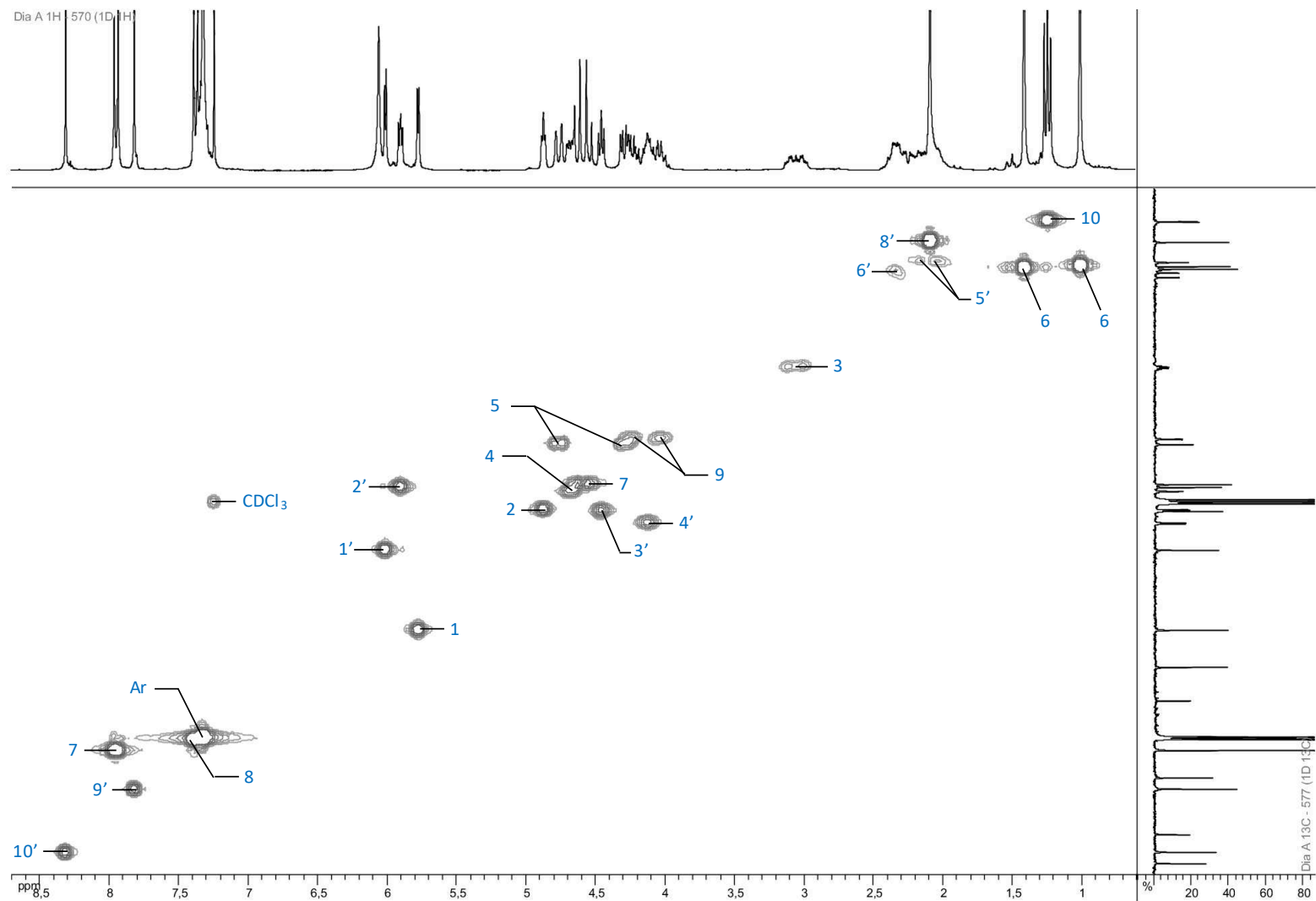
Spectre COSY de **197c**



Spectre <sup>13</sup>C de 197c



Spectre HSQC de **197c**





## Exemples d'attribution des signaux des spectres RMN du dinucléotide

### T<sup>PMB</sup>-U 183a

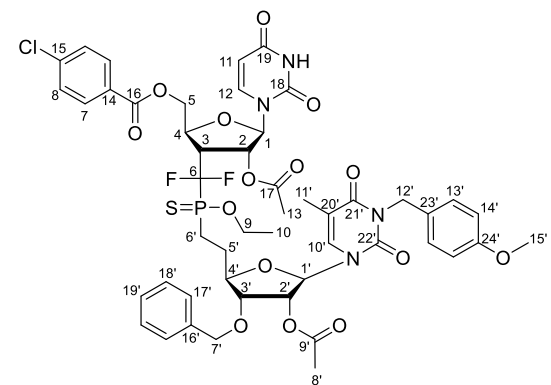
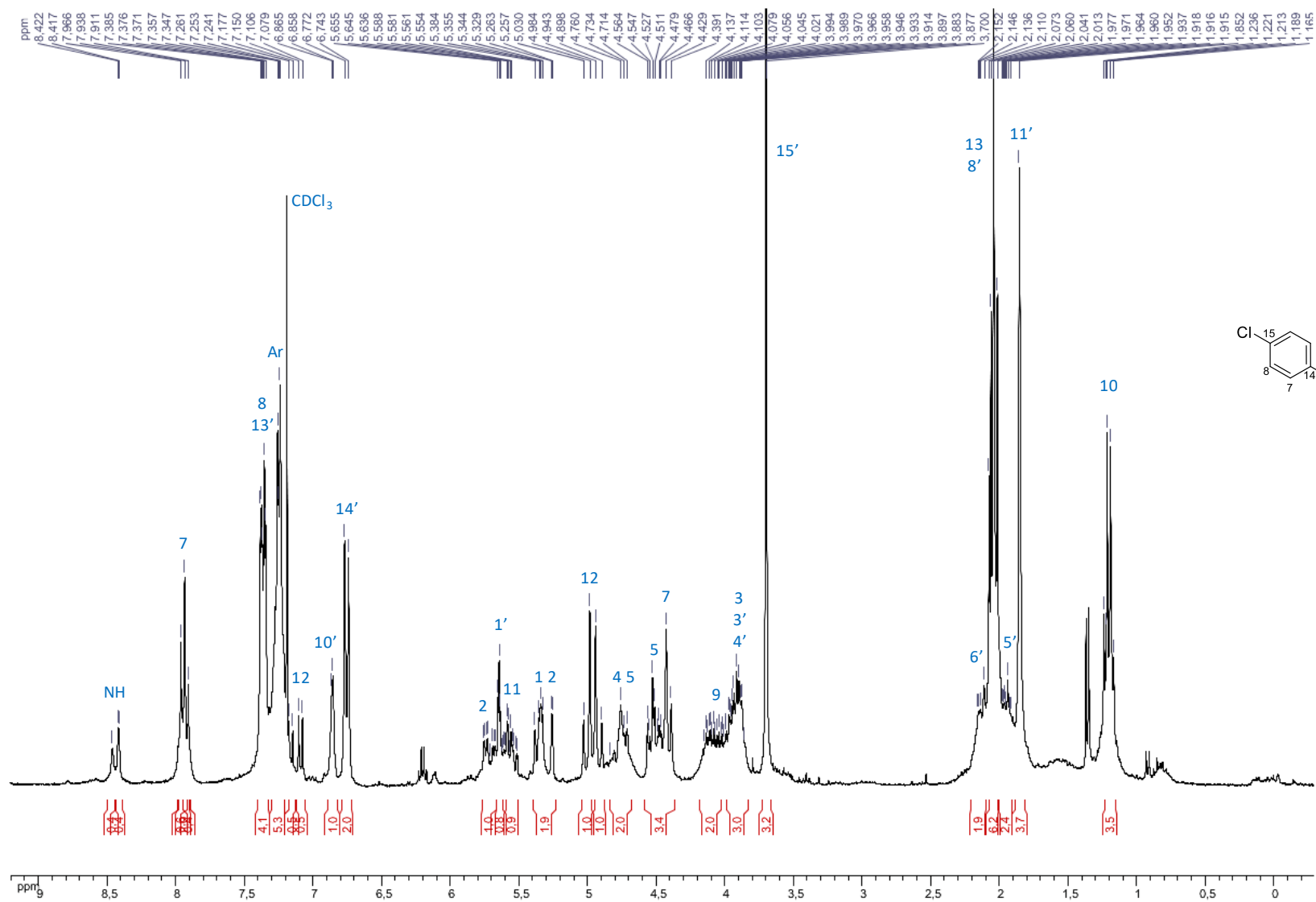
Dans le cas, les spectres analysés correspondent à un mélange des deux diastéréoisomères A et B ratio A/B : 1.1/1, déterminé à partir du spectre RMN de <sup>31</sup>P découplé du proton après purification).

C6 n'est pas observé dans les différents spectres à cause de ses couplages avec les fluors et le phosphore. Le premier massif (1.25-1.16 ppm) observé sur le spectre <sup>1</sup>H et étant assimilé à H10 permet de déterminer les signaux correspondant à H9 en COSY (4.14-4.02 ppm). Cette analyse peut être poursuivie grâce aux signaux caractéristiques des deux bases nucléiques. A 6.87 ppm se trouve le signal typique de H10' de la thymine. Ce déplacement chimique est confirmé en COSY par corrélation avec H11' à 1.85 ppm. A partir de celui-ci et de l'HMBC, les positions de C1' (91.4 ppm), C11' (13.5 ppm), C21' (163.3 ppm) et C22' (150.8 ppm) peuvent être déterminées. H11' permet de positionner le C20' à 111.1 ppm. En HMBC, C21' permet de déterminer le signal de H12' (5.01 ppm). Ce dernier aide au positionnement de C13' (129.1 ppm) et C23' (129.1 ppm). Pour finir avec les signaux de la thymine, le singulet à 3.70 ppm est le pic de H15'. En HMBC, ce dernier permet de positionner C24' (159.3 ppm) et par la même occasion H14' (6.80 ppm). A 7.09 ppm, le doublet observé est typique au H12 de l'uridine et permet de déterminer la position de H1 (5.38-5.25 ppm) en NOESY, celle de H11 en COSY (5.57 ppm) et en HMBC, celles de C18 (149.7 ppm) et C19 (162.7 ppm). A noter que le NH de l'uridine est positionné à 8.42 ppm sur le spectre <sup>1</sup>H. Les deux groupements acétates sont à la même position sur les spectres <sup>1</sup>H et <sup>13</sup>C. Les deux méthyles (H13 et H8') se trouvent à 2.07-2.01 ppm en <sup>1</sup>H et 21.0 ppm en <sup>13</sup>C tandis que les deux carbonyles (C17 et C9') sont aux alentours de 170.0 ppm en <sup>13</sup>C.

Les protons H1 et H1' ne permettent pas de déterminer les positions respectives de H2 et H2'. A partir d'ici, l'attribution des derniers signaux se fera à partir d'hypothèse. Par analogie avec les autres molécules synthétisées, le H7 se situe à 7.99-7.90 ppm sur le spectre <sup>1</sup>H. Celui-ci permet de placer le H8 (7.38-7.35 ppm) à l'aide du COSY et C14 (140.2 ppm) et C16 (165.4 ppm) en HMBC. De plus, H8 permet de déterminer la position de C15 (128.2 ppm). Un signal à 4.56-4.39 ppm, en HMBC, corrèle avec des signaux aromatiques. Vu que H12' est déjà placé, ce signal correspond donc à H7'. Toujours en HMBC, ce dernier permet le placement des signaux de C3' (79.3 ppm), de C16' (137.2 ppm) et de C17' (128.3 ppm). Le C3' permet de déterminer H2' (5.38-5.25 ppm) en HMBC et en COSY. Deux signaux sont encore à attribués dans le spectre <sup>13</sup>C, dans la partie aromatique, 128.8 ppm et 128.5 ppm et il reste deux carbones aromatiques disponibles C18' et C19'. L'intensité du pic à 128.8 ppm est supérieure à celle de 128.5 ppm, par déduction, C18' se situe à 128.8 ppm et C19' à 128.5 ppm. A ce stade, il reste donc à placer H2, H3, H4, H5, H4', H5' et H6'. Par analogie avec le spectre <sup>13</sup>C des composés similaires synthétisés, H3, H5' et H6' peuvent être déduit. En effet, sur le spectre <sup>13</sup>C, C3 a

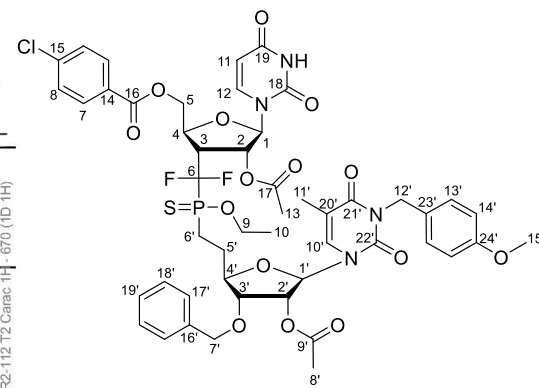
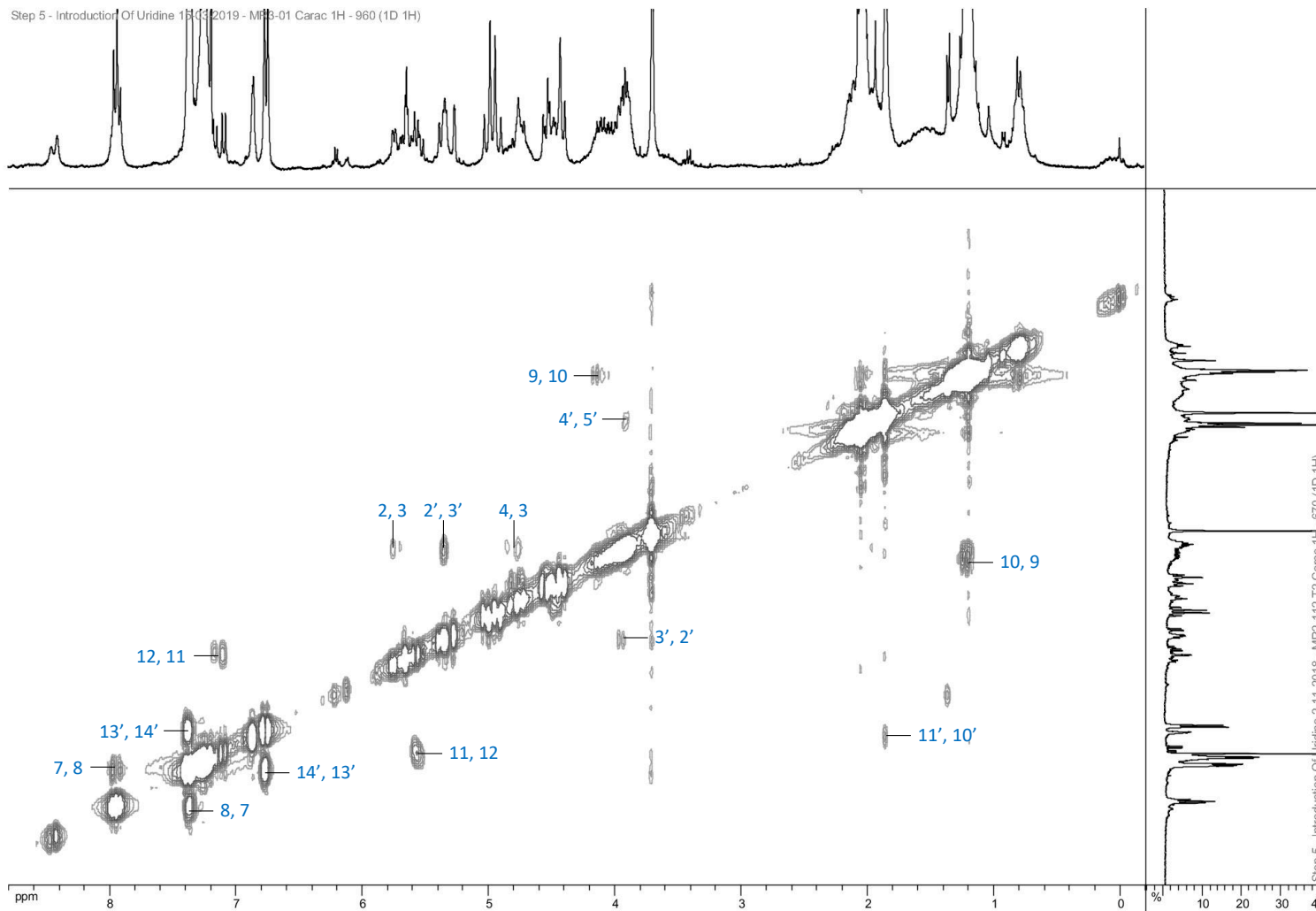
un signal sous la forme d'un multiplet à 43.8-42.9 ppm. A partir de cette valeur, sa position sur le spectre  $^1\text{H}$  peut être déduite par HSQC (3.99-3.85 ppm). Ce dernier corrèle avec le signal correspondant à H2 (5.74 ppm) en HMBC. De plus, une seconde corrélation est observée en COSY à 4.83-4.71 ppm. Par élimination, ce massif correspond à H4. H5' et H6' se trouvent dans les massifs à 2.15-2.11 ppm et à 1.98-1.92 ppm. L'un des deux corrèle avec un signal à 3.99-3.85 ppm. Par déduction, H4' correspond donc à 3.99-3.85 ppm, H5' à 1.98-1.92 ppm et H6' à 2.15-2.11 ppm. Le signal restant correspond donc à H5 réparti sur les massifs 4.83-4.71 ppm et 4.56-4.39 ppm.

Spectre <sup>1</sup>H de 183a



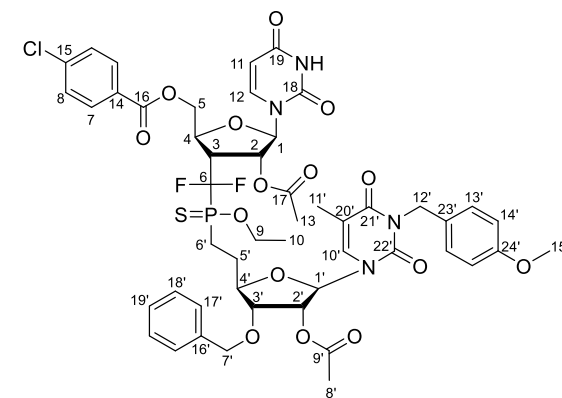
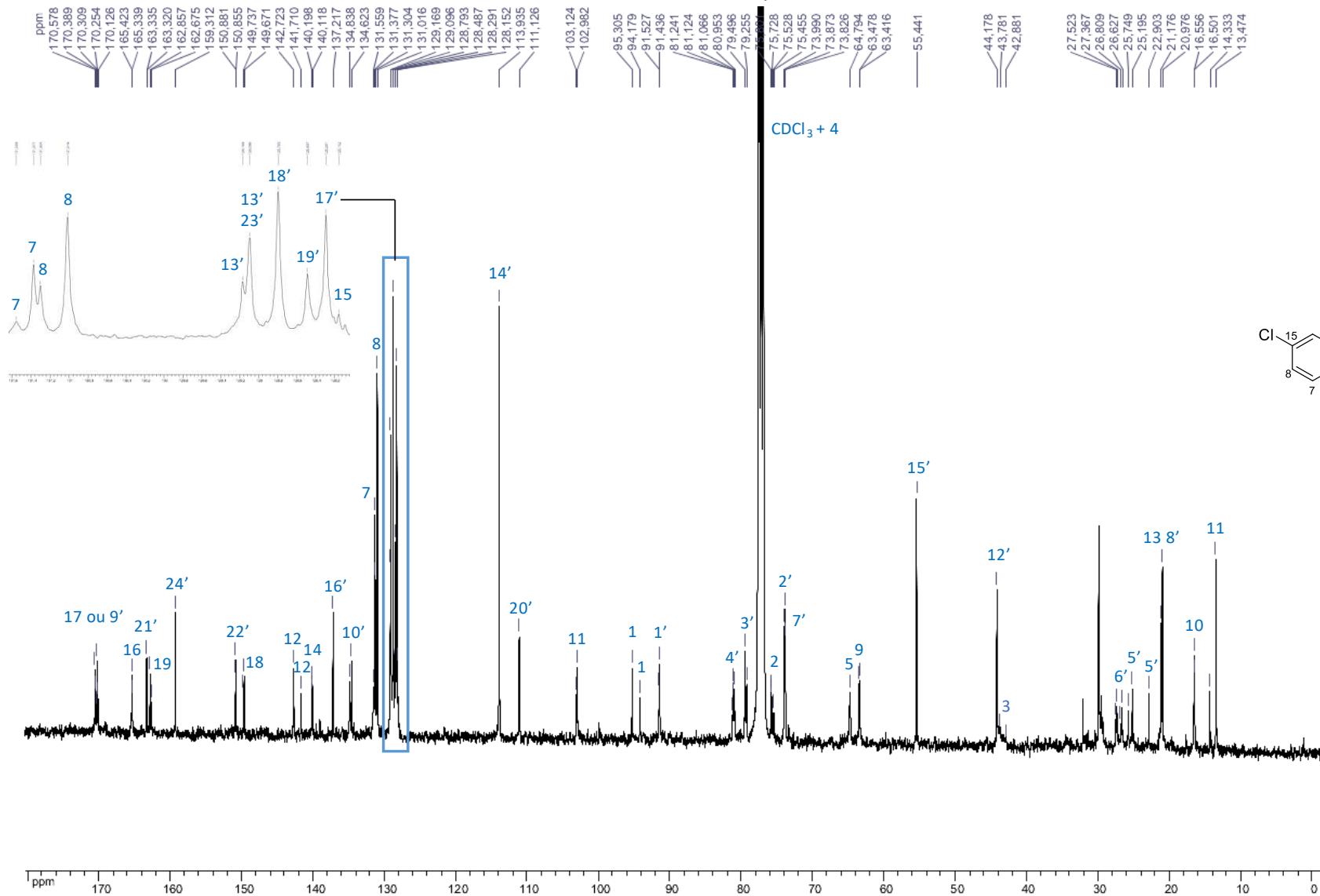
Spectre COSY de 183a

Step 5 - Introduction Of Uridine 183-03-2019 - MR3-01 Carac 1H - 960 (1D 1H)



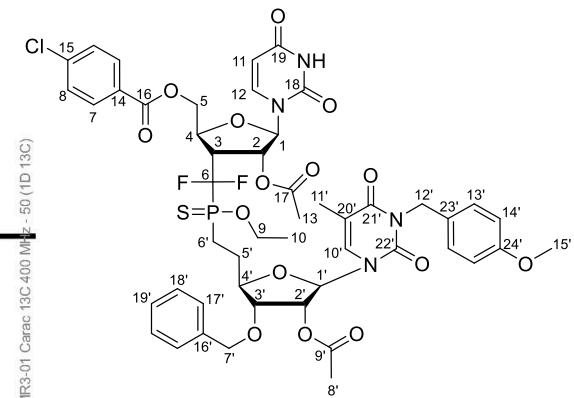
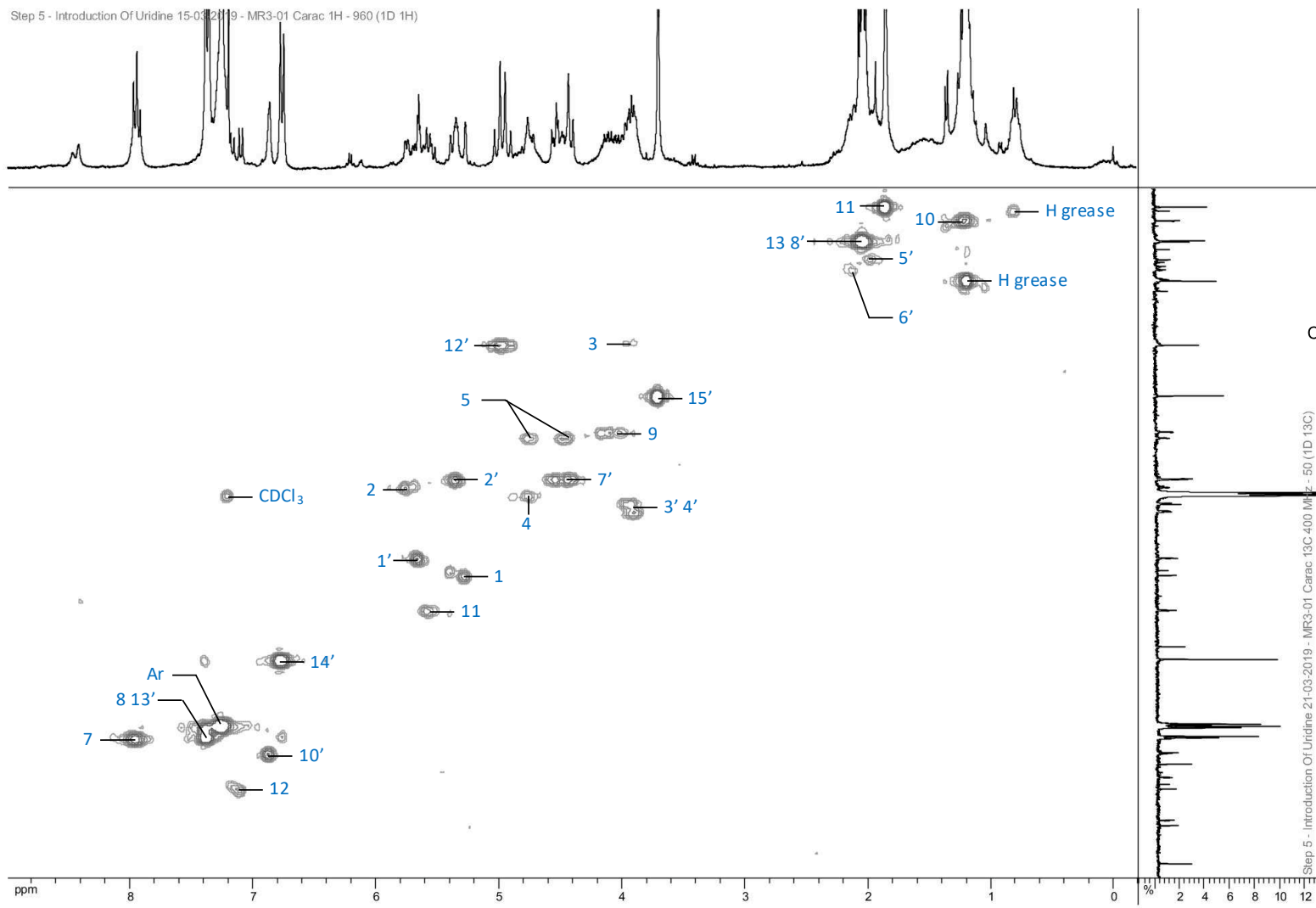
Step 5 - Introduction Of Uridine 2-11-2018 - MR2-112 T2 Carac 1H - 670 (1D 1H)

Spectre <sup>13</sup>C de 183a



Spectre HSQC de 183a

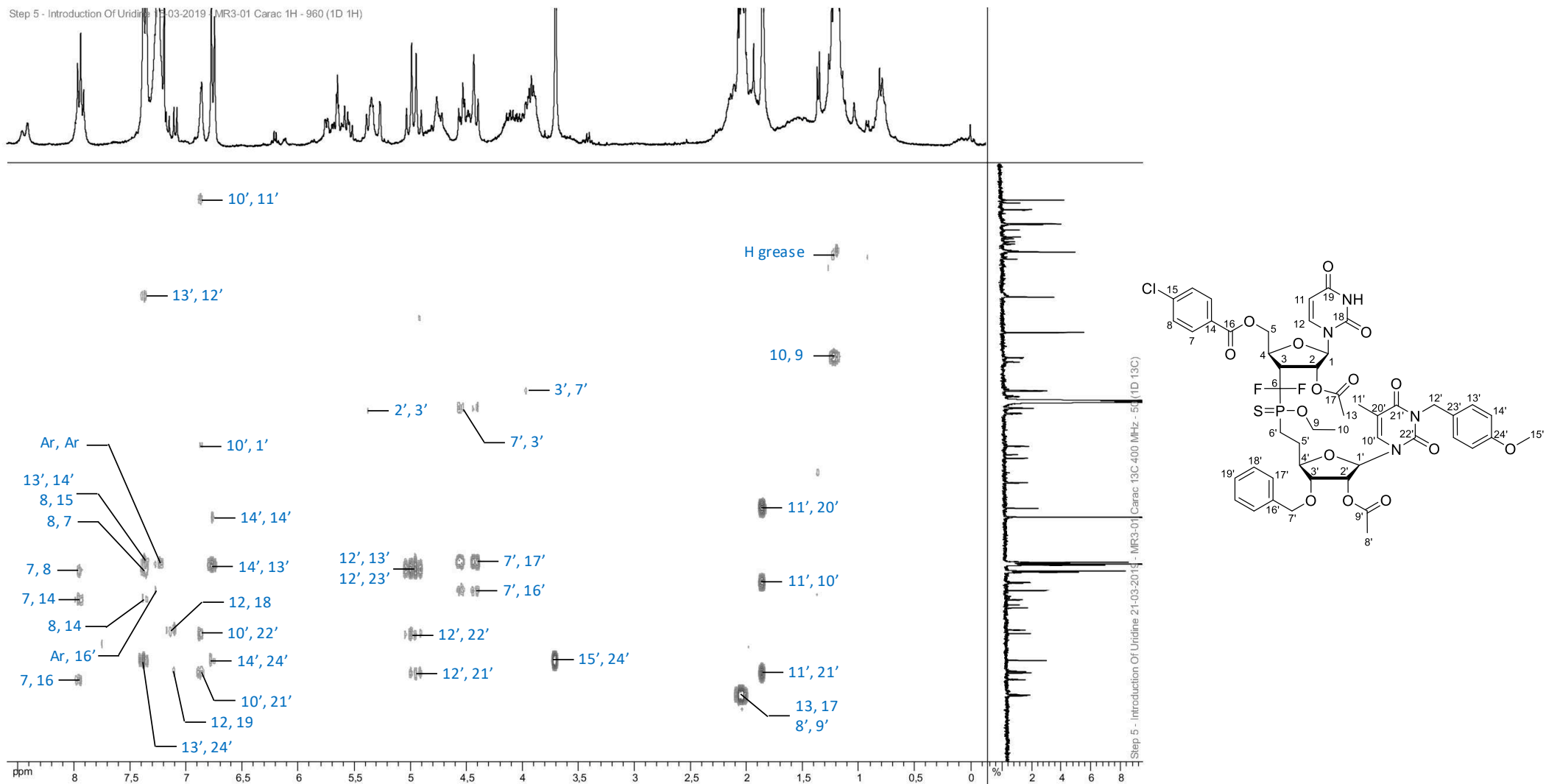
Step 5 - Introduction Of Uridine 15-03-2019 - MR3-01 Carac 1H - 960 (1D 1H)



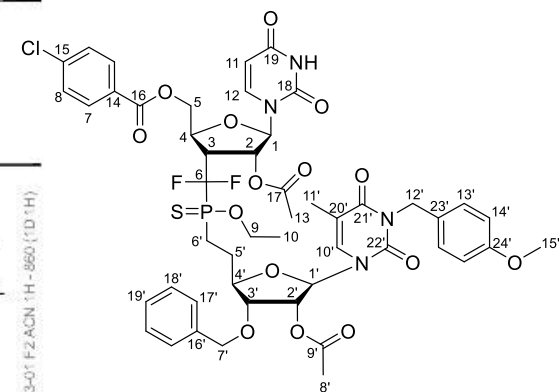
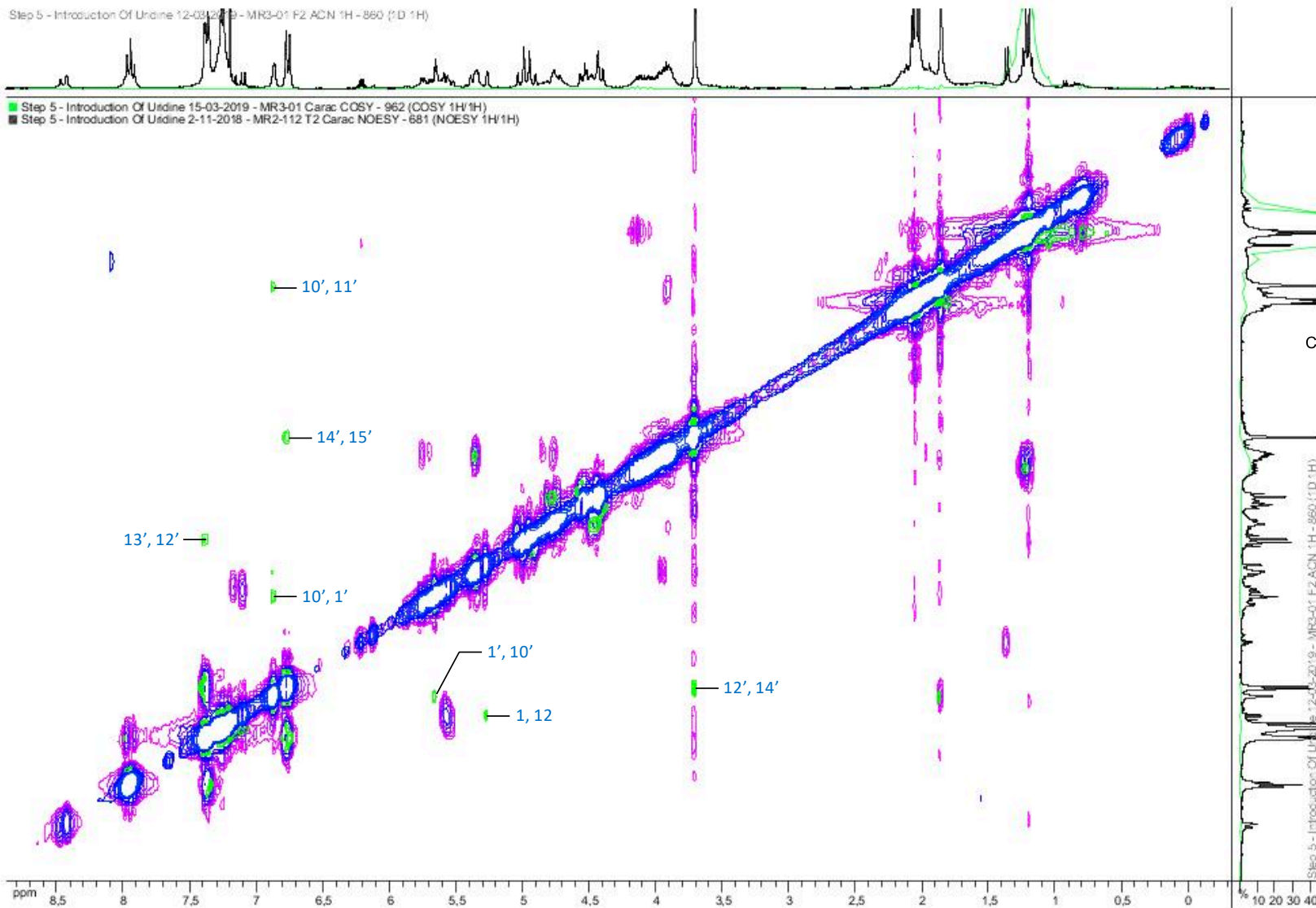
Step 5 - Introduction Of Uridine 21-03-2019 - MR3-01 Carac 13C 400 MHz - 50 (1D 13C)

Spectre HMBC de 183a

Step 5 - Introduction Of Uridine 21-03-2019 - MR3-01 Carac 1H - 960 (1D 1H)



Spectre NOESY de 183a





Les oligonucléotides (ONs) naturels sont une classe majeure de biomolécules ayant un potentiel thérapeutique prometteur. Cependant, le développement d'ONs thérapeutiques rencontre deux problèmes majeurs : un manque de stabilité envers les nucléases et un manque de sélectivité envers une cible thérapeutique. Pour pallier ces problèmes, les chercheurs académiques et industriels ont travaillé sur le développement d'ONs modifiés, incorporant notamment de nouveaux analogues du pont phosphodiester naturel. Dans ce contexte, la synthèse de nouveaux analogues du lien naturel a été étudiée au laboratoire, à savoir les motifs difluorophosphinate (**R-CF<sub>2</sub>-P(O)(OEt)CH<sub>2</sub>-R**) et difluorophosphinothioate (**R-CF<sub>2</sub>-P(S)(OEt)CH<sub>2</sub>-R**). Ainsi, grâce aux travaux menés précédemment, une synthèse de première génération a été mise au point conduisant à la formation d'un homodinucléotide comportant ces motifs. Le fragment-clé **furanose-CF<sub>2</sub>-P(S)(OEt)CH<sub>2</sub>-furanose**, permettant l'introduction de deux bases nucléiques identiques, est obtenu par une construction successive des deux liaisons P-C. Une oxydation du motif **R-CF<sub>2</sub>-P(S)(OEt)CH<sub>2</sub>-R** permet l'obtention du lien **R-CF<sub>2</sub>-P(O)(OEt)CH<sub>2</sub>-R**. Dans la continuité de ces travaux, ce projet de thèse est focalisé sur la mise au point d'une synthèse de seconde génération menant à des hétérodinucléotides caractérisés par la présence du motif difluorophosphin(othio)ate et deux nucléobases différentes. A partir d'un substrat commun, l'allofuranose, la synthèse des différents nucléosides intermédiaires a été réalisée. Dans un premier temps, ceux-ci ont ensuite été utilisés afin de former un nouveau fragment-clé **furanose-CF<sub>2</sub>-P(S)(OEt)CH<sub>2</sub>-nucléoside** à partir duquel une seconde base nucléique différente a été incorporée. Cependant, au cours de cette étude, un phénomène de substitution d'une base nucléique par une autre a pu être mis en évidence. En dépit de ce processus dit de trans-*N*-glycosylation, six dinucléotides (T<sup>PMB</sup>-U, T-U, U-U, T-T, C-T, C-G<sup>Prot</sup>) ont pu être préparés.

**Mots clés :** difluorophosphinothioate, difluorophosphinate, nucléosides, hétérodinucléotide, réaction de Michaelis-Becker, trans-*N*-glycosylation.

Oligonucleotides (ONs) represent a major class of bioactive molecules with a great potential in medicinal chemistry. However, the development of therapeutic ONs is facing two major problems: a lack of stability toward nucleases and a lack of selectivity toward a specific therapeutic target. To overcome those problems, academic and industrial chemists are working on the development of modified ONs, incorporating in particular new analogues of the natural phosphodiester bridge. In this context, the synthesis of new analogues was studied in the laboratory, namely the difluorophosphinate (**R-CF<sub>2</sub>-P(O)(OEt)CH<sub>2</sub>-R**) and difluorophosphinothioate (**R-CF<sub>2</sub>-P(S)(OEt)CH<sub>2</sub>-R**) units. Thanks to previous studies, a first-generation synthesis was developed towards a homodinucléotide incorporating a difluorophosphinate moiety. A key fragment **furanose-CF<sub>2</sub>-P(S)(OEt)CH<sub>2</sub>-furanose**, allowing the introduction of two identical nucleic bases, was obtained by the sequential construction of the two P-C bonds. Oxidation of **R-CF<sub>2</sub>-P(S)(OEt)CH<sub>2</sub>-R** gave access to the **R-CF<sub>2</sub>-P(O)(OEt)CH<sub>2</sub>-R** unit. This thesis describes the development of a second-generation synthesis towards heterodinucléotides incorporating a difluorophosphin(othio)ate as internucleosidic bridge and two different nucleic bases. Starting from a common substrate, allofuranose, the synthesis of key fragments was worked out. These intermediates were used to build a difluorophosphinothioate internucleosidic bridge. Thus, a new key fragment **furanose-CF<sub>2</sub>-P(S)(OEt)CH<sub>2</sub>-nucleoside** was obtained and used to introduce a second, identical or different, nucleic base. However, a substitution process involving the replacement of a nucleic base with another one was highlighted. Despite this trans-*N*-glycosylation process, six dinucleotides (T<sup>PMB</sup>-U, T-U, U-U, T-T, C-T, C-G<sup>Prot</sup>) were prepared.

**Keywords:** difluorophosphinothioate, difluorophosphinate, nucleosides, heterodinucléotide, Michaelis-Becker reaction, trans-*N*-glycosylation.