



# Synthèse de sondes luminescentes pour la détection séquence spécifique d'ADN double brin

Guillaume Frémy

## ► To cite this version:

Guillaume Frémy. Synthèse de sondes luminescentes pour la détection séquence spécifique d'ADN double brin. Chimie organique. Université Grenoble Alpes, 2019. Français. NNT : 2019GREAV073 . tel-03134166

**HAL Id: tel-03134166**

**<https://theses.hal.science/tel-03134166>**

Submitted on 8 Feb 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.

## THÈSE

Pour obtenir le grade de

**DOCTEUR DE LA COMMUNAUTE UNIVERSITE  
GRENOBLE ALPES**

Spécialité : **Chimie Organique**

Arrêté ministériel : 25 mai 2016

Présentée par

**Guillaume FRÉMY**

Thèse dirigée par **Olivier SÉNÈQUE**  
et co-encadrée par **Muriel JOURDAN**

préparée au sein du **Laboratoire de Chimie et Biologie des Métaux**  
Dans l'**École Doctorale Chimie et Sciences du Vivant**

# Synthèse de sondes luminescentes pour la détection séquence- spécifique d'ADN double brin

Thèse soutenue publiquement le **9 décembre 2019**,  
devant le jury composé de :

**Madame Maryline BEYLER**

Maître de Conférences, Université de Bretagne Occidentale, Rapporteur

**Monsieur Vincent AUCAGNE**

Directeur de Recherche, CNRS-Orléans, Rapporteur

**Madame Frédérique LOISEAU**

Professeure des Universités, Université Grenoble Alpes, Présidente du jury

**Monsieur David MONCHAUD**

Directeur de Recherche, Université de Bourgogne, Examinateur

**Madame Muriel JOURDAN**

Maître de Conférences, Université Grenoble Alpes, Membre invitée

**Monsieur Jean-François CONSTANT**

Chargé de Recherche, Université Grenoble Alpes, Membre invité

**Monsieur Olivier SÉNÈQUE**

Directeur de Recherche, CNRS-Grenoble, Directeur de thèse





*A ma famille*

*« À force de persévérance et de courage,  
la petite fourmi finit par arriver au sommet de la montagne. »*



## Remerciements

Je tiens tout d'abord à remercier le jury de thèse composé du Dr Maryline Beyler, Dr Vincent Aucagne, Pr Frédérique Loiseau, et Dr David Monchaud, d'avoir accepté d'évaluer mon travail de thèse et pour la discussion enrichissante pendant la soutenance. Par ailleurs, un grand merci aux financements, le labex Arcane et l'EDCSV sans quoi rien n'aurait été possible.

J'ai effectué cette thèse au laboratoire de chimie et biologie des métaux dans l'équipe de physicochimie des métaux en biologie, dirigé par Geneviève Blondin. Merci à toi pour l'accueil dans ce qui est devenu mon second chez moi, Ô grande cheffe !

Un énorme merci à Olivier Sénèque, tu m'as accueilli et pris sous ton aile pour un stage de M2, qui s'est enchaîné par cette thèse dont voici l'aboutissement. Merci pour tout ce que tu m'as appris scientifiquement et personnellement, pour ta patience et ta compréhension, je n'oublierai jamais cette expérience. Merci pour tout le temps que tu m'as consacré pour les corrections de ce manuscrit, pour toutes communications orales et pour tout le reste.

En continuant sur les chefs, Muriel et Jean-François c'est votre tour. Merci à vous deux d'avoir pu être présents pour moi dès qu'il le fallait, d'avoir réalisés les manips de gels retard avec moi. Merci aussi pour les corrections, et d'avoir toujours répondu présents pour les réunions et répétitions.

Merci à toi Colette, pour les analyses MS et l'aide sur le Liberty1 qui nous a donné pas mal de fil à retordre. Jacques, à toi aussi merci pour les analyses MS, la résolution de structure et pour les diverses conversations. Christine, merci pour les analyses MALDI-TOF, au final on aura pas fait autant qu'espéré, mais je ne doute pas que mon successeur vienne te voir bientôt avec pleins d'analyses ! Merci à toi Pierre-Alain, pour avoir toujours été présent à la RMN et pour les conseils sur le matos de montagne. Je tiens aussi à remercier les collaborateurs que je n'ai pas eu le temps de connaître : Rémy Lartia pour la synthèse des brins d'ADN et Nathan McClenaghan pour la précieuse SiR-méthyl.

Didier et Christelle, merci de m'avoir suivi pendant ces trois années via les CSI, et merci d'avoir été présents à ma soutenance (merci la SNCF !).

Merci à tous les permanents de l'équipe ; Claude et Geneviève pour avoir gérées tout ce qui touche à l'administratif ; Amélie pour le prêt de ton bureau pendant la phase de rédaction ; Patrick pour les trucs et astuces, tes anecdotes et ta bonne humeur ; Martin, Ricardo, Jean-Marc, pour les discussions dans le couloir, au coin du labo, ou au petit déj le vendredi matin.

Merci à tous les non-permanents pour la superbe ambiance au laboratoire qui génère une atmosphère de travail agréable, mais aussi pour les différentes sorties bières, tacos... Elena, pour les petits cours sur la chimie de coordination et ta bonne humeur ; Céline pour m'avoir formé à la synthèse peptidique ; Coin-Coin pour la motivation que tu procures et pour tous ces bons moments passés au labo ; Jordan, mon comparse de master pour le soutien, allez on va pouvoir élever des chèvres dans le Vercors maintenant ! ; Patrick El présidente, pour ta bonne humeur et ton enthousiasme permanent ; Ji-Hyung, pour les fou-rires que tu procures avec ta maladresse ; Ikram, pour les pâtisseries. Merci à tous les stagiaires passés au laboratoire pour avoir contribué à l'ambiance : Margot, Kate, Clément, Clara, Oriane, Guillaume et Morgan.

Encore un et pas des moindres (non, non, je ne t'ai pas oublié !) : Thibault !!! Merci pour tout ce que tu as fait à l'époque où tu étais stagiaire et ce que tu as fait en revenant en thèse pour prendre la succession du projet, le plus beau cadeau que tu pouvais me faire. Mon regret est qu'on n'ait pas pu passer plus de temps ensemble au labo... Voilà je t'ai confié tous mes échantillons, et maintenant cette thèse qui t'es quand même un peu dédiée.

Merci aux gens du SyMMES, Yves, Florence, Johan, Harish, Coralie, Seydoux, Jérémy, Mickaël, pour ces moments partagés dans les couloirs.

Encore quelques derniers remerciements qui étaient inclassables, Laurent, je te remercie pour les discussions les fois où on s'est croisés et pour avoir apporté et testé la ligation de trois fragments au labo. Gwendoline, merci pour avoir géré ce qui relate de l'administratif avec le labex Arcane. Noëlle, merci pour les discussions et le soutien, ta bonne humeur permanente a égayé mes journées.

Merci à toute ma famille et mes amis pour leur soutien tout au long de ces trois années, pour leur compréhension sur la difficulté et le temps que prend cette expérience. Merci à mes grands-parents, passés ou présents. Merci à mes parents et mes frères de m'avoir toujours aidé pour que je puisse y arriver, pour leur réconfort et de m'avoir supporté malgré les périodes de stress que peut procurer ces 3 années.





## Table des matières

|                                                                          |    |
|--------------------------------------------------------------------------|----|
| Liste des abréviations .....                                             | 1  |
| Introduction .....                                                       | 11 |
| I Contexte .....                                                         | 11 |
| II Détection par fluorescence .....                                      | 11 |
| 1) Luminescence .....                                                    | 11 |
| 2) Transfert d'énergie résonant de Förster .....                         | 13 |
| III Cibler et détecter l'ADN.....                                        | 15 |
| 1) L'ADN .....                                                           | 15 |
| A Structure primaire .....                                               | 15 |
| B Structure secondaire .....                                             | 15 |
| 2) Sondes à ADN.....                                                     | 16 |
| A Sondes à ADN simple brin .....                                         | 17 |
| i Les molecular beacons .....                                            | 17 |
| ii Sondes FRET.....                                                      | 17 |
| B Sondes à ADN double brins .....                                        | 18 |
| i Oligonucléotides pour la formation de triplexes (TFO) .....            | 18 |
| ii Pyrole Imidazole Polyamides(PPI) .....                                | 19 |
| iii Sondes à base de polypeptides modifiés ou de protéines .....         | 19 |
| IV Les doigts de zinc .....                                              | 22 |
| 1) Description.....                                                      | 22 |
| 2) La protéine ZIF12 .....                                               | 24 |
| 3) Doigts de zinc, ADN et FRET : ZIF12 comme support de sonde FRET ..... | 25 |
| V Les lanthanides .....                                                  | 27 |
| 1) Propriétés fondamentales .....                                        | 27 |
| 2) Propriétés spectroscopiques .....                                     | 28 |
| A Emission des lanthanides .....                                         | 28 |
| B Temps de vie de luminescence.....                                      | 29 |
| C Utilisation pour des sondes luminescentes.....                         | 30 |

|                                                                                        |                                                                              |    |
|----------------------------------------------------------------------------------------|------------------------------------------------------------------------------|----|
| D                                                                                      | Effet « antenne » .....                                                      | 31 |
| 3)                                                                                     | Sondes luminescentes responsives à base de lanthanide .....                  | 32 |
| VI                                                                                     | Projet.....                                                                  | 34 |
| 1)                                                                                     | Sondes de première génération (FRET d'une antenne vers un lanthanide).....   | 34 |
| 2)                                                                                     | Sondes de seconde génération (FRET d'un lanthanide vers un fluorophore)..... | 36 |
| 3)                                                                                     | Cibler une séquence non palindromique .....                                  | 37 |
| Chapitre I : Sondes de première génération, FRET d'une antenne vers un lanthanide..... |                                                                              | 43 |
| I                                                                                      | Principe.....                                                                | 43 |
| II                                                                                     | Insertion de chromophores sur la protéine .....                              | 43 |
| 1)                                                                                     | Choix des chromophores.....                                                  | 43 |
| 2)                                                                                     | Introduction des chromophores sur un peptide .....                           | 45 |
| A                                                                                      | DOTA[Tb] .....                                                               | 45 |
| B                                                                                      | Carbostyryl-124 .....                                                        | 45 |
| 3)                                                                                     | Positions d'introduction des chromophores.....                               | 46 |
| III                                                                                    | Synthèse de la protéine .....                                                | 47 |
| 1)                                                                                     | Assemblage par Ligation Chimique Native (NCL) .....                          | 48 |
| A                                                                                      | Présentation .....                                                           | 48 |
| B                                                                                      | Formation du lien amide natif.....                                           | 48 |
| C                                                                                      | Cinétique .....                                                              | 49 |
| D                                                                                      | Introduction d'un thioester en C-terminal .....                              | 50 |
| 2)                                                                                     | Crypto-thioesters.....                                                       | 50 |
| IV                                                                                     | Stratégies de synthèse des sondes.....                                       | 52 |
| 1)                                                                                     | Assemblage en deux fragments .....                                           | 53 |
| 2)                                                                                     | Assemblage en trois fragments .....                                          | 54 |
| V                                                                                      | Synthèse de la sonde ZIF(Cs124 DOTA[Tb]).....                                | 56 |
| 1)                                                                                     | Fragment Ac-F1[Lys(DOTA[Tb])]-MPA .....                                      | 57 |
| 2)                                                                                     | Fragment F2[Cys(StBu) Glu(Cs124)]-SEA <sup>off</sup> .....                   | 59 |
| A                                                                                      | Aspartimides.....                                                            | 60 |
| B                                                                                      | Perte d'une masse de -121.....                                               | 62 |

|                                                                                      |                                                                                                     |    |
|--------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|----|
| 3)                                                                                   | Fragment F3- $NH_2$ .....                                                                           | 65 |
| 4)                                                                                   | Assemblage des différents fragments .....                                                           | 66 |
| 5)                                                                                   | Formation des doigts de zinc .....                                                                  | 67 |
| VI                                                                                   | Interaction de la sonde avec l'ADN .....                                                            | 68 |
| 1)                                                                                   | Mobilité électrophorétique.....                                                                     | 68 |
| A                                                                                    | Principe.....                                                                                       | 68 |
| B                                                                                    | Marquage .....                                                                                      | 69 |
| C                                                                                    | Mesures effectuées .....                                                                            | 69 |
| i                                                                                    | Protéine ZIF <sup>Zn</sup> (nue) .....                                                              | 70 |
| ii                                                                                   | Etude préliminaire sur la protéine ZIF <sup>Zn</sup> (Cs124   DOTA[Tb]).....                        | 70 |
| 2)                                                                                   | Luminescence .....                                                                                  | 71 |
| A                                                                                    | Caractérisation de la protéine ZIF <sup>Zn</sup> (Cs124   DOTA[Tb]) .....                           | 71 |
| B                                                                                    | Titration à l'ADN.....                                                                              | 72 |
| VII                                                                                  | Conclusion .....                                                                                    | 73 |
| Chapitre II : Sondes de seconde génération, FRET entre chromophores organiques ..... |                                                                                                     | 77 |
| I                                                                                    | Introduction .....                                                                                  | 77 |
| 1)                                                                                   | Principe .....                                                                                      | 77 |
| 2)                                                                                   | Stratégie de synthèse des sondes.....                                                               | 77 |
| 3)                                                                                   | Choix des fluorophores du système de FRET.....                                                      | 78 |
| II                                                                                   | Synthèse des protéines ZIF <sup>Zn</sup> (lissamine) et ZIF <sup>Zn</sup> (SiR-méthyl).....         | 81 |
| 1)                                                                                   | Synthèse des protéines ZIF(lissamine) et ZIF(SiR-méthyl) .....                                      | 81 |
| A                                                                                    | Fragment F1-MPA.....                                                                                | 81 |
| B                                                                                    | Fragments F2 .....                                                                                  | 82 |
| i                                                                                    | Peptide F2[Cys(StBu)   Lys(lissamine)]-SEA <sup>off</sup> .....                                     | 83 |
| ii                                                                                   | Peptide F2[Cys(StBu)   Lys(SiR-méthyl)]-SEA <sup>off</sup> .....                                    | 83 |
| C                                                                                    | Assemblages .....                                                                                   | 84 |
| 2)                                                                                   | Stabilité des protéines, conditions de stockage.....                                                | 85 |
| 3)                                                                                   | Méthylation au zinc.....                                                                            | 88 |
| III                                                                                  | Etude spectroscopique du système ZIF <sup>Zn</sup> (lissamine)/ZIF <sup>Zn</sup> (SiR-méthyl) ..... | 88 |
| A                                                                                    | Etude du mélange ZIF <sup>Zn</sup> (lissamine)/ZIF <sup>Zn</sup> (SiR-méthyl) .....                 | 88 |

|                                                          |                                                                                                                  |     |
|----------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|-----|
| B                                                        | Etude du mélange ZIF <sup>Zn</sup> (lissamine)/ZIF <sup>Zn</sup> (SiR-méthyl) avec la séquence d'ADN cible ..... | 90  |
| IV                                                       | Conclusion .....                                                                                                 | 92  |
| Chapitre III : Synthèse de complexes de lanthanides..... |                                                                                                                  | 97  |
| I                                                        | Introduction .....                                                                                               | 97  |
| 1)                                                       | Complexes bioconjuguables .....                                                                                  | 97  |
| 2)                                                       | Complexes d'euporium à absorption bi-photonique.....                                                             | 99  |
| 3)                                                       | Complexes de terbium à absorption bi-photonique .....                                                            | 101 |
| 4)                                                       | Conclusion .....                                                                                                 | 101 |
| II                                                       | Design des complexes de lanthanides, de leurs ligands et stratégie de synthèse .....                             | 102 |
| 1)                                                       | Design des complexes.....                                                                                        | 103 |
| A                                                        | Couplage par chimie click.....                                                                                   | 103 |
| B                                                        | Couplage lors de la synthèse peptidique .....                                                                    | 105 |
| C                                                        | Stratégies de couplage .....                                                                                     | 105 |
| 2)                                                       | Stratégie de synthèse.....                                                                                       | 106 |
| III                                                      | Synthèse des ligands .....                                                                                       | 109 |
| 1)                                                       | Synthèse du macrocycle DO3A <sup>(Me)3</sup> pic <sup>(Me)</sup> I.....                                          | 109 |
| A                                                        | Synthèse du composé 7.....                                                                                       | 109 |
| B                                                        | Synthèse du DO3A <sup>(Me)3</sup> .....                                                                          | 110 |
| C                                                        | Couplage final.....                                                                                              | 110 |
| 2)                                                       | Synthèse des substituants Ar <sub>i</sub> (i = 1, 2, 3).....                                                     | 111 |
| A                                                        | Ar <sub>1</sub> .....                                                                                            | 111 |
| B                                                        | Ar <sub>2</sub> .....                                                                                            | 112 |
| C                                                        | Ar <sub>3</sub> .....                                                                                            | 112 |
| 3)                                                       | Synthèse des ligands DO3ApicAr <sub>i</sub> .....                                                                | 113 |
| A                                                        | Couplages de Suzuki-Miyaura .....                                                                                | 113 |
| i                                                        | DO3ApicAr <sub>1</sub> .....                                                                                     | 113 |
| ii                                                       | DO3A <sup>(Me)3</sup> pic <sup>(Me)</sup> Ar <sub>1</sub> .....                                                  | 114 |
| B                                                        | Couplages de Sonogashira.....                                                                                    | 114 |
| i                                                        | DO3ApicAr <sub>2</sub> .....                                                                                     | 114 |
| ii                                                       | DO3A <sup>(Me)3</sup> pic <sup>(Me)</sup> Ar <sub>2</sub> .....                                                  | 115 |
| iii                                                      | DO3ApicAr <sub>3</sub> .....                                                                                     | 116 |

|                                                                                                      |                                                                                                                                               |     |
|------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 4)                                                                                                   | Remarque sur la caractérisation des ligands .....                                                                                             | 116 |
| 5)                                                                                                   | Conclusion .....                                                                                                                              | 117 |
| IV                                                                                                   | Complexes $[\text{Ln}(\text{DO3ApicAr}_i)]^-$ .....                                                                                           | 117 |
| 1)                                                                                                   | Synthèse .....                                                                                                                                | 117 |
| 2)                                                                                                   | Propriétés spectroscopiques des complexes.....                                                                                                | 120 |
| A                                                                                                    | Absorption et coefficients d'extinction molaire.....                                                                                          | 120 |
| B                                                                                                    | Luminescence .....                                                                                                                            | 120 |
| C                                                                                                    | Remarque sur la synthèse des complexes .....                                                                                                  | 124 |
| V                                                                                                    | Conclusion.....                                                                                                                               | 125 |
| Chapitre IV : Sondes de seconde génération, FRET d'un lanthanide vers un fluorophore organique ..... |                                                                                                                                               | 129 |
| I                                                                                                    | Introduction .....                                                                                                                            | 129 |
| II                                                                                                   | FRET du terbium vers un fluorophore .....                                                                                                     | 130 |
| 1)                                                                                                   | Synthèse de la protéine $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))$ .....                                                         | 130 |
| A                                                                                                    | Peptide $\text{F2}[\text{Cys}(\text{StBu}) \text{Lys}(\text{Tb}(\text{DO3ApicAr}_1))]\text{-SEA}^{\text{off}}$ .....                          | 131 |
| i                                                                                                    | Couplage du complexe $[\text{Tb}(\text{DO3ApicAr}_1)]\text{Na}$ .....                                                                         | 131 |
| ii                                                                                                   | Couplage du pro-ligand $\text{DO3A}^{(\text{Me})3}\text{pic}^{(\text{Me})}\text{Ar}_1$ .....                                                  | 131 |
| iii                                                                                                  | Hydrolyse.....                                                                                                                                | 132 |
| iv                                                                                                   | Métallation .....                                                                                                                             | 134 |
| B                                                                                                    | Assemblage.....                                                                                                                               | 135 |
| C                                                                                                    | Métallation au zinc .....                                                                                                                     | 136 |
| 2)                                                                                                   | Etude de la luminescence .....                                                                                                                | 137 |
| A                                                                                                    | Système FRET du terbium vers la lissamine .....                                                                                               | 137 |
| i                                                                                                    | Etude du mélange $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{lissamine})$ .....                       | 137 |
| ii                                                                                                   | Interaction du couple $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{lissamine})$ avec son ADN cible ... | 138 |
| iii                                                                                                  | Conclusion .....                                                                                                                              | 141 |
| B                                                                                                    | Système FRET du terbium vers la SiR-méthyl.....                                                                                               | 141 |
| 3)                                                                                                   | Conclusion .....                                                                                                                              | 143 |
| III                                                                                                  | FRET de l'euporium vers un fluorophore .....                                                                                                  | 144 |
| 1)                                                                                                   | Travaux préliminaires .....                                                                                                                   | 144 |
| i                                                                                                    | Ligation en conditions classiques .....                                                                                                       | 144 |

|     |                                                                                       |     |
|-----|---------------------------------------------------------------------------------------|-----|
| ii  | Changement des conditions de ligation .....                                           | 145 |
| iii | Conclusion .....                                                                      | 145 |
| 2)  | Mise en place d'une nouvelle voie de synthèse de protéines.....                       | 146 |
| A   | Utilisation de la résine HNB.....                                                     | 146 |
| i   | Principe .....                                                                        | 146 |
| ii  | Synthèse du groupement N-HNB-Cys(StBu).....                                           | 148 |
| iii | Réactivité du phénol.....                                                             | 148 |
| B   | Test des stratégies d'assemblages .....                                               | 150 |
| i   | Stratégie d'assemblage de C-vers-N.....                                               | 150 |
| ii  | Stratégie d'assemblage de N-vers-C.....                                               | 154 |
| iii | Conclusion sur les stratégie d'assemblages .....                                      | 158 |
| C   | Synthèse de la protéine ZIF(Eu(DO3ApicAr <sub>2</sub> )) .....                        | 159 |
| i   | Synthèse du peptide F2[Cys Lys(Eu(DO3ApicAr <sub>2</sub> ))]-HNB <sup>off</sup> ..... | 159 |
| ii  | Assemblage.....                                                                       | 164 |
| IV  | Conclusion .....                                                                      | 166 |
|     | Conclusion générale et perspectives.....                                              | 171 |
|     | Bibliographie .....                                                                   | 179 |
|     | Partie expérimentale .....                                                            | 197 |
|     | Annexes .....                                                                         | 260 |





## Liste des abréviations

### A

A : Adénine

Å : Angström ( $10^{-10}$  m)

a.a. : Acide aminé

AcOEt : Acétate d'éthyle

AcOH : Acide acétique

ADN : Acide Désoxyribonucléique

APS : Persulfate d'ammonium

Ar : Aromatique

ARN : Acide Ribonucléique

### B

B : Brilliance

BSA : Bovine serum albumine/ Albumine de sérum bovin

### C

C : Cytosine

CHCl<sub>3</sub> : Chloroforme

Cs : Carbostyryl

CT : Transfert de charge

### D

δ (ppm) : Déplacement chimique RMN

ΔM : Différence de masse

D : Donneur

Da : Dalton

DCM : Dichlorométhane

DEPT : Distortionless Enhanced Polarization Transfer

DIC : Diisopropylcarbodiimide

DIEA : Diisopropyléthylamine

DMF : Diméthylformamide

DMSO : Diméthylsulfoxyde

DO3A : Acide 1,4,7,10-tétraazacyclododécane-1,4,7-triacétique

DOTA : Acide 1,4,7,10-tétraazacyclododécane-1,4,7,10-tétraacétique

DTPA : Acide diéthylène triamine pentaacétique

DTT : Dithiothréitol

## **E**

$\epsilon$  : Coefficient d'extinction molaire

E : Efficacité

EDTA : Ethylènediaminetétraacétique

EMSA : Electrophoretic Mobility Shift Assay/ Décalage de mobilité électrophorétique

ESI : Electrospray Ionisation/ Ionisation par électronébuliseur

Et<sub>2</sub>O : Ether diéthylique

## **F**

FRET : Transfert d'énergie résonant de Förster

## **G**

G : Guanine

GFP : Green Fluorescent Protein/ Protéine fluorescente verte

Gn-HCl : Guanidine hydrochloride

GP : Groupement protecteur

GSH : Glutathion

## **H**

HEPES : Acide 4-(2-hydroxyéthyl)-1-pipérazine éthane sulfonique

HNB : N-(2-hydroxy-5-nitrobenzyl)

HPLC : High pressure liquid chromatography/ Chromatographie liquide haute pression

HRMS : High Resolution Mass Spectroscopy/ Spectroscopie de masse à haute résolution

## **J**

J(Hz) : Constante de couplage

## **L**

$\lambda$  : Longueur d'onde

LMCT : Transfert de charge ligand-métal

Ln : Lanthanide

## **M**

M : Masse

MeCN : Acétonitrile

MeOH : Méthanol

MesNa : 2-mercaptoéthanesulfonate sodium

MLCT : Transfert de charge métal-ligand

MPAA : Acide 4-mercaptophenylacétique

MPA : Acide mercaptopropionique

MS : Mass Spectroscopy/ Spectroscopie de masse

## **N**

NCL : Native chemical ligation/ Ligation chimique native

NHS : N-hydroxysuccinimide

## **P**

$\Phi$  : Rendement quantique

PAGE : Poly-acrylamide gel electrophoresis/ Electrophorèse sur gel de polyacrylamide

PBS : Tampon phosphate salin

PEG : Polyéthylène glycol

pH : Potentiel Hydrogène

PNA : Peptide nucleic acid/ Acide nucléique peptidique

PPI : Pyrole Imidazole Polyamides

PPh<sub>3</sub> : Triphénylphosphine

ppm : Partie par million

PS : Polystyrène

## **R**

R<sub>0</sub> : Rayon de Förster

RA : RinkAmide

RMN : Résonance Magnétique Nucléaire

r.t. : Room temperature/ Température ambiante

RX : Rayons X

## S

SEA : bis(2-sulfanyléthyl)amido

SPPS : Solid Phase Peptide Synthesis/ synthèse peptidique sur support solide

## T

T : Thymine

TACN : 1, 4, 7, Triazacyclononane

TALE : Transcription activator-like effector/ Effecteur TAL

TALEN : Transcription activator-like effector nucleases/ Nucléases effectrices de type activateur de transcription

TBAF : Fluorure de N-tétrabutylammonium

TBE : Tris, Borate, EDTA

TCEP : tris(2-carboxyéthyl)phosphine

TEA : Triéthylamine

TEMED : Tétraméthyléthylenediamine

TFA : Acide trifluoroacétique

THF : Tétrahydrofurane

TFO : Triplex Forming Oligonucleotide/ Oligonucléotide pour la formation de triplexes

TIS : Triisopropylsilane

Tris : Tris(hydroxyméthyl)aminométhane

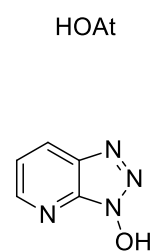
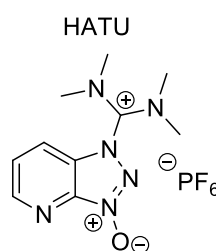
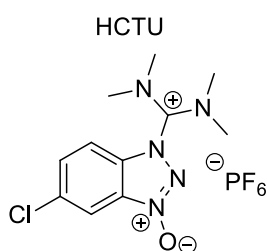
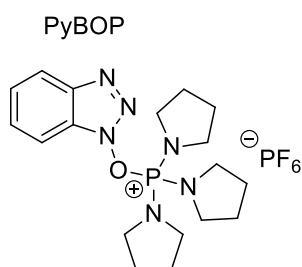
TTHA : Acide triéthylènetetraamine-N,N,N',N'',N''',N''''-hexaacétique

## U

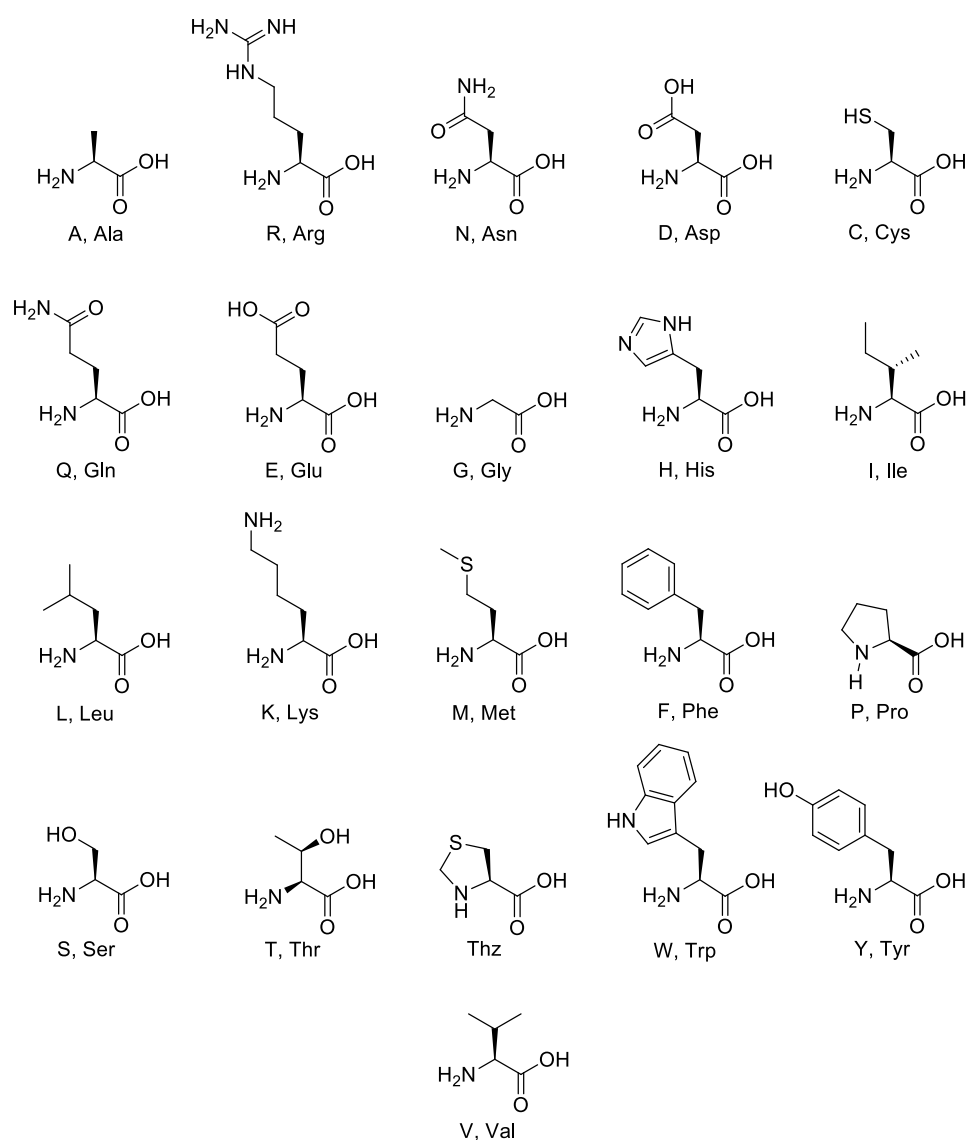
U : Uracile

UV : Ultraviolet

## Activateurs



## Abréviations des acides aminés



## Notation des peptides

Les peptides présentés dans ce manuscrit possèdent la notation suivante :

La partie N-terminale (notée Z dans l'exemple Figure 1) est d'abord indiquée en italique (si elle n'est pas composée d'un groupement NH<sub>2</sub> classique), séparée par un tiret du nom de la séquence (notée F dans l'exemple). Si un acide aminé particulier est présent dans le peptide, en portant par exemple une fonction, il est indiqué entre crochets. Finalement, la partie C-terminale (notée Y dans l'exemple) est indiquée en italique, séparée par un tiret.

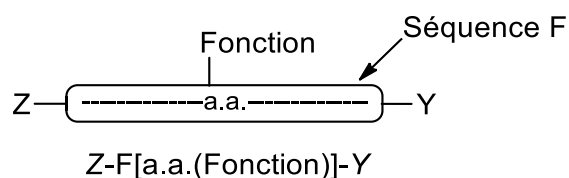


Figure 1: Exemple de notation d'un peptide comportant un acide aminé spécifique

Dans plusieurs cas, l'acide aminé présent en N-terminal devra être indiqué, pour cela l'acide aminé sera introduit entre les crochets et séparé par une barre de l'acide aminé comportant une fonction, comme dans l'exemple en Figure 2.

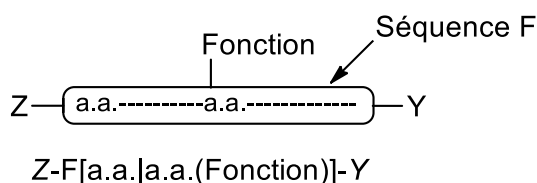


Figure 2: Exemple de notation d'un peptide avec un acide aminé en N-terminal devant être indiqué, et comportant un acide aminé spécifique

Si le peptide est toujours supporté sur une résine, cette dernière sera indiquée dans le dessin par une boule grise (Figure 3).

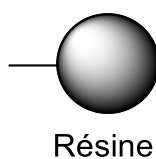


Figure 3: Représentation de la résine





# Introduction



# Introduction

## I Contexte

L'intérêt pour la détection de séquences spécifiques d'oligonucléotides en solution ou en milieu biologique (cellule, extrait cellulaire) s'est accrue au fil des dernières décennies. Il est crucial de pouvoir détecter sélectivement une séquence d'ADN par exemple pour le diagnostic médical, la sécurité alimentaire ou certaines recherches. Sur la base d'un calcul statistique, une séquence de 17 paires de bases devrait suffire à cibler sélectivement n'importe quel gène humain<sup>[1]</sup>. La luminescence est une des méthodes les plus simples à mettre en œuvre au niveau du laboratoire. Les méthodes actuelles de détection par luminescence reposent en général sur la détection d'un simple brin d'ADN par hybridation avec un ou deux brins complémentaires portant un fluorophore dont l'émission est modulée par la proximité d'un quencheur ou d'un autre fluorophore accepteur d'énergie. Le but de cette thèse est d'établir la preuve de concept de la détection d'un ADN double brin de séquence choisie en utilisant les doigts de zinc pour leurs propriétés de reconnaissance de l'ADN et des lanthanides pour leurs propriétés de luminescence particulières, très intéressantes pour la détection en milieu biologique.

## II Détection par fluorescence

### 1) Luminescence

La luminescence désigne l'émission d'un photon par une substance à partir d'un état électronique excité<sup>[2]</sup>. Lorsque l'état excité est un état singulet, on parle de fluorescence et lorsqu'il s'agit d'un état triplet, on parle de phosphorescence. Les substances luminescentes peuvent être des (macro)molécules contenant un squelette aromatique, des cations de type lanthanides, des complexes métalliques ou des nanoparticules. La luminescence fait le plus souvent suite à l'absorption d'un photon qui amène la substance dans un état singulet excité (Figure 4). Celui-ci se désactive de manière non-radiative vers son état singulet excité de plus basse énergie,  $S_1$ , à partir duquel le phénomène de fluorescence a lieu. Le temps de vie de fluorescence est de l'ordre de 10 nanosecondes pour des molécules organiques. Les substances dans leur état excité  $S_1$  peuvent également subir une conversion de spin pour passer dans leurs premier état excité triplet,  $T_1$ . Ce processus, appelé croisement intersystème, est théoriquement interdit et est donc plus lent que la fluorescence. La phosphorescence peut alors se produire

depuis l'état  $T_1$ . La transition  $T_1 \rightarrow S_0$  est également théoriquement interdite du fait du changement de spin et son temps de vie caractéristique va de la milliseconde à la seconde.

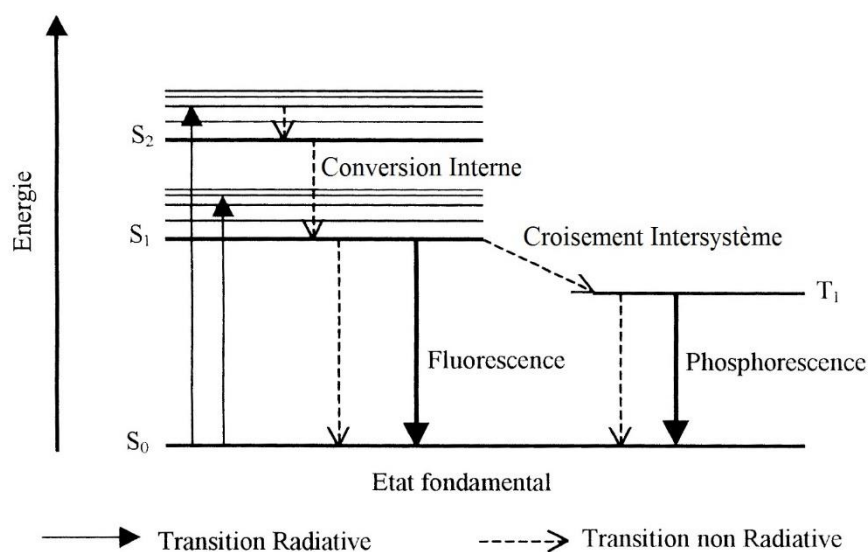


Figure 4 : Diagramme de Jablonski-Perrin

Un fluorophore est une molécule aromatique capable de fluorescer. Il se caractérise par les grandeurs suivantes :

- longueur d'onde d'excitation ( $\lambda_{ex}$ )
- longueur d'onde d'émission ( $\lambda_{em}$ )
- déplacement de Stokes, différence entre  $\lambda_{em}$  et  $\lambda_{ex}$
- rendement quantique ( $\Phi$ ), correspondant à la capacité d'émission de la lumière après absorption (défini par l'équation  $\Phi = (\text{nombre de photons émis}) / (\text{nombre de photons absorbés})$ )
- coefficient d'extinction molaire ( $\epsilon$ ), correspondant à l'efficacité d'absorption de la lumière (en  $M^{-1}cm^{-1}$ )
- brillance  $B$ , produit de  $\epsilon$  par  $\Phi$
- temps de vie de luminescence

Pour des applications en milieu biologique, il peut être intéressant de maximiser la brillance du fluorophore afin de le détecter plus facilement, en maximisant le rendement quantique et le coefficient d'extinction molaire. Un second facteur d'importance est le déplacement de Stokes, c'est-à-dire le décalage entre le spectre d'absorption et le spectre d'émission du fluorophore. Idéalement celui-ci doit être grand pour éviter le recouvrement de ces deux spectres et ainsi éviter les phénomènes de ré-absorption.

Dans l'optique de réaliser des mesures *in vivo*, une excitation et une émission aux hautes longueurs d'ondes est souhaitable. En effet, les milieux biologiques absorbent la lumière dans l'UV jusqu'au visible. La fenêtre de diagnostic idéale appelée fenêtre de transparence du milieu biologique est située dans le proche infrarouge, entre 650 et 1500nm<sup>[3]</sup> (Figure 5).

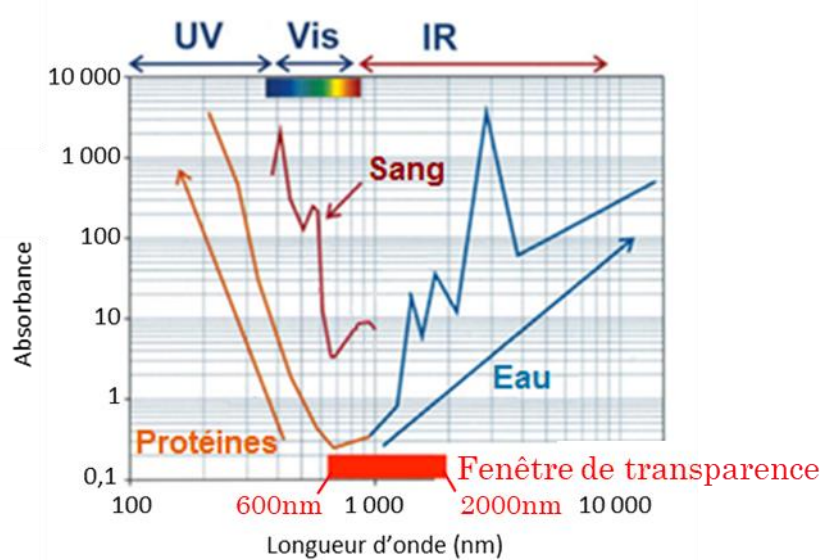


Figure 5: Absorption des milieux biologiques et fenêtre de diagnostic<sup>[3]</sup>

Les fluorophores sont depuis de nombreuses années utilisés pour concevoir des sondes<sup>[4],[5],[6]</sup> capables de révéler la présence d'un analyte (molécule, cation ...) dans un milieu, par un changement de leurs propriétés d'émission.

## 2) Transfert d'énergie résonant de Förster

Une des techniques les plus répandues pour la conception de sondes lumineuses est l'utilisation du transfert d'énergie résonant de Förster (FRET)<sup>[7]</sup>. Le FRET met en jeu un couple de fluorophores, l'un étant le donneur d'énergie et l'autre l'accepteur. Lorsque le donneur est excité (à une longueur d'onde notée  $\lambda_{ex}$ ), il peut se désexciter soit par voie radiative (en émettant des photons), soit par voie non radiative en transférant son énergie à l'accepteur qui se retrouve alors dans un état excité capable d'émettre de la lumière.

Le transfert d'énergie est dû à une interaction dipolaire entre les deux fluorophores. Son efficacité dépend de la distance entre le donneur et l'accepteur. On note le rayon de Förster  $R_0$ , la distance pour laquelle l'efficacité  $E$  du transfert est de 50% (Figure 6). Il dépend de la nature des chromophores utilisés, mais est en général compris entre 10 et 100 Å<sup>[8]</sup>. Pour que le transfert soit possible, il faut que le spectre d'émission du donneur recouvre au moins en partie le spectre d'excitation de l'accepteur (Figure 7). Plus le recouvrement est important, plus le transfert est

efficace. Le rendement quantique du donneur et le coefficient d'extinction molaire de l'accepteur sont deux paramètres à maximiser pour optimiser le transfert.

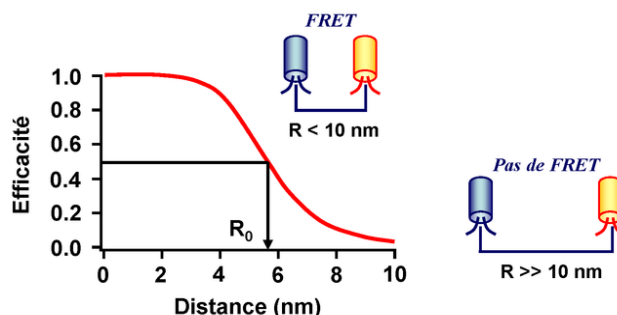


Figure 6: Evolution de l'efficacité d'un transfert donneur vers accepteur en fonction du rayon de Förster  $R_0$

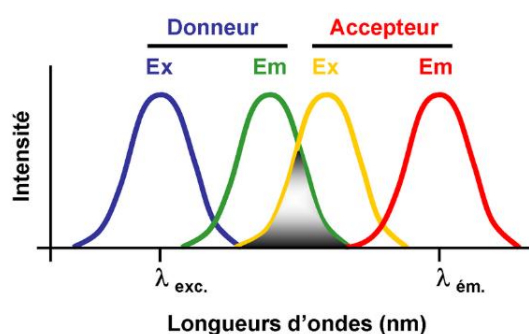


Figure 7: Spectres d'excitation (d'absorption) et d'émission du donneur et de l'accepteur

La dépendance en distance de ce transfert d'énergie est à la base du principe de détection de nombreuses sondes fluorescentes<sup>[9],[10]</sup>. Ces sondes doivent bien entendu intégrer deux fluorophores. Si l'analyte à détecter provoque le rapprochement des deux fluorophores, le transfert d'énergie devient plus efficace, l'émission du donneur diminue et celle de l'accepteur augmente (Figure 8). Les sondes FRET sont particulièrement intéressantes car elles permettent une détection ratiométrique, qui permet d'accéder à la concentration de l'analyte sans connaître la concentration de la sonde.

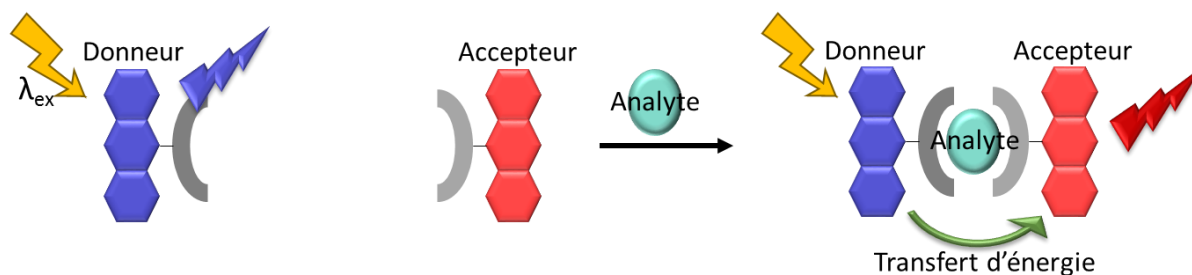


Figure 8: Principe de la modulation du transfert d'énergie par la présence de l'entité ciblée

Le choix des fluorophores pour la conception d'une sonde FRET dépend de nombreux paramètres pour adapter l'efficacité du transfert en présence de l'analyte ou du phénomène à

détecter (changement conformationnel par exemple) et au milieu dans lequel la détection doit être réalisée. C'est en utilisant ce principe que nous souhaitons détecter une séquence d'ADN.

### III Cibler et détecter l'ADN

#### 1) L'ADN

La molécule d'ADN (Acide DésoxyriboNucléique) est le support de l'information génétique de tout organisme vivant. Elle est située dans le cytoplasme des cellules procaryotes et dans le noyau des cellules eucaryotes. L'ADN est composé de deux structures : une structure primaire, et une autre secondaire.

##### A Structure primaire

La structure primaire est un polymère de nucléotides, appelé communément brin d'ADN. Ces nucléotides sont composés d'un acide ortho-phosphorique, lié à un sucre (le 2-désoxy-D-ribose), sur lequel est fixé à une base azotée aussi nommée base nucléique (Figure 9). Ces bases sont soit des bases puriques comme l'adénine (A) et la guanine (G), soit des bases pyrimidiques comme la cytosine (C) et la thymine (T). Un brin d'ADN est donc composé d'un squelette constitué par la répétition sucre-phosphate substitué par différentes bases azotées.

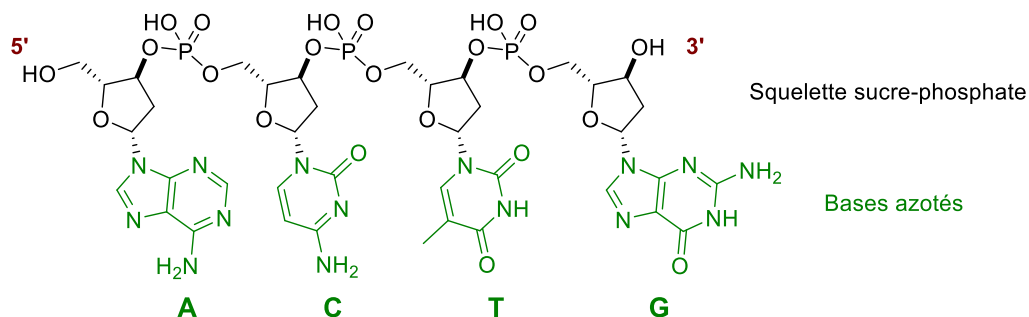


Figure 9: Structure primaire de l'ADN

La nomenclature d'un brin d'ADN comprend la notation des positions libres des sucres aux deux extrémités (5' à gauche et 3' à droite) ainsi que la représentation des bases nucléiques présentes par la première lettre de leur nom (soit pour l'exemple en Figure 9, 5'-ACTG-3').

##### B Structure secondaire

La structure secondaire de l'ADN retrouvée majoritairement dans les cellules a été décrite par J. Watson et F. Crick en 1953<sup>[11]</sup>. Elle a depuis été nommée ADN-B et est composée de deux brins, interagissant l'un avec l'autre par l'intermédiaire de leurs bases nucléiques. Ces dernières

s'associent spécifiquement par paires par des liaisons hydrogène, la thymine établit deux liaisons H avec l'adénine, et la guanine en établit trois avec la cytosine (Figure 10). Le duplex est également stabilisé par les interactions d'empilement de leurs cycles aromatiques<sup>[12]</sup>. L'appariement des deux brins d'ADN de manière antiparallèle donne une double hélice s'enroulant autour d'un axe, formant un petit et un grand sillon dans lesquels s'établissent des interactions spécifiques avec des protéines et divers ligands (Figure 11).

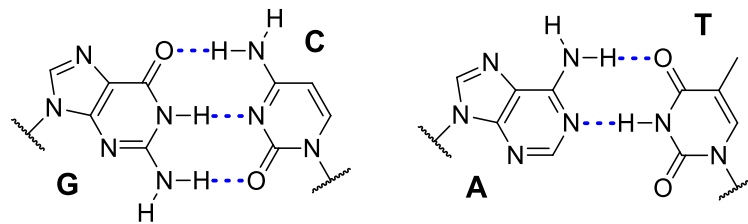


Figure 10: Liaisons hydrogène entre les bases nucléiques complémentaires

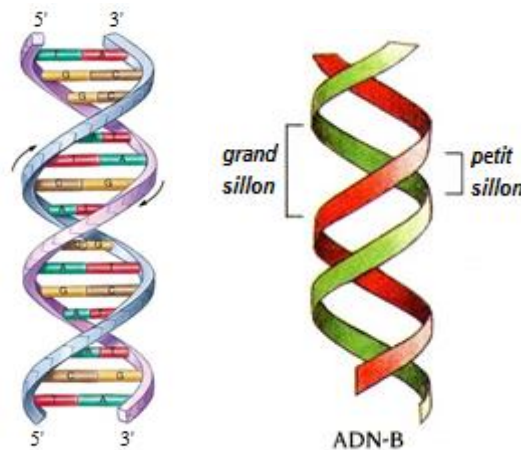


Figure 11: Appariement antiparallèle d'un double brin d'ADN

## 2) Sondes à ADN

Les principaux critères des sondes permettant de cibler spécifiquement des séquences d'ADN sont une interaction spécifique et stable dans le temps, une cinétique rapide, et une grande sensibilité. Parmi les différentes techniques permettant la détection d'ADN nous nous intéresserons uniquement à la détection par luminescence. En effet, les sondes luminescentes possèdent plusieurs avantages tels que leur faible coût, leur sensibilité, ainsi que la facilité et la rapidité de détection.

Parmi l'ensemble des sondes luminescentes décrites dans la littérature, on distingue deux grands types de sondes permettant de cibler l'ADN : les sondes ciblant un simple brin d'ADN <sup>[13],[14]</sup>; et celles ciblant de l'ADN sous forme de double brin <sup>[15]</sup>.

## A Sondes à ADN simple brin

Les sondes ciblant un simple brin d'ADN nécessitent de travailler dans des conditions dénaturantes<sup>[16]</sup> (par exemple en chauffant pour dissocier le duplex d'ADN).

### i Les molecular beacons

Un premier exemple de sonde est le molecular beacon<sup>[17]</sup>, qui est un oligonucléotide composé de la séquence complémentaire de celle ciblée comportant à ses extrémités quelques paires de bases complémentaires entre elles. Les extrémités sont fonctionnalisées par un couple fluorophore/quencher (Figure 12). En l'absence de la séquence d'ADN cible, la sonde est repliée. La séquence complémentaire est alors présente dans une boucle et les paires de bases complémentaires sont appariées (structure de type tige-boucle ou en épingle à cheveux). Le quencher est ainsi dans l'environnement du fluorophore et aucune luminescence n'est observée lorsque le fluorophore est excité. En présence de la séquence d'ADN cible, le brin complémentaire présent dans la sonde s'hybride avec sa séquence cible. Les extrémités ne sont plus appariées, éloignant le quencher du fluorophore, permettant une réponse lumineuse de ce dernier après excitation.

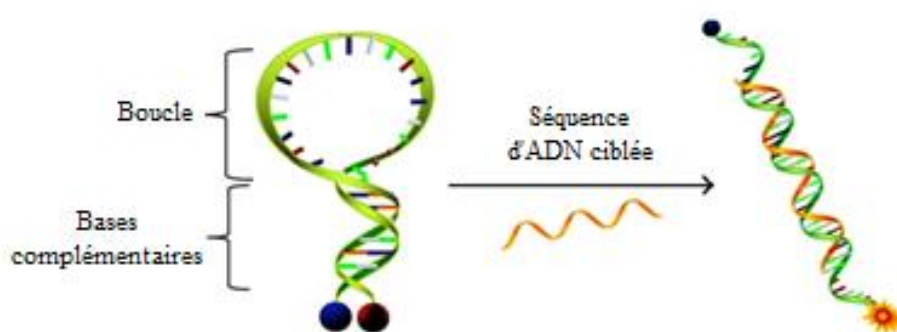


Figure 12: Principe d'un molecular beacon<sup>[18]</sup>

### ii Sondes FRET

Un autre exemple de sonde lumineuse est celle reposant sur un FRET entre deux fluorophores : un donneur, un autre accepteur<sup>[19]</sup>. Chacun est placé sur un oligonucléotide correspondant à une partie de la séquence d'ADN à cibler (Figure 13). En l'absence de la séquence d'ADN, le fluorophore donneur est excité, et son émission est observée. Lorsque les deux brins s'hybrident avec la séquence d'ADN cible, le fluorophore donneur est rapproché de l'accepteur. Ainsi, lorsque le donneur est excité, un transfert d'énergie a lieu vers l'accepteur et l'émission de ce dernier est observée.

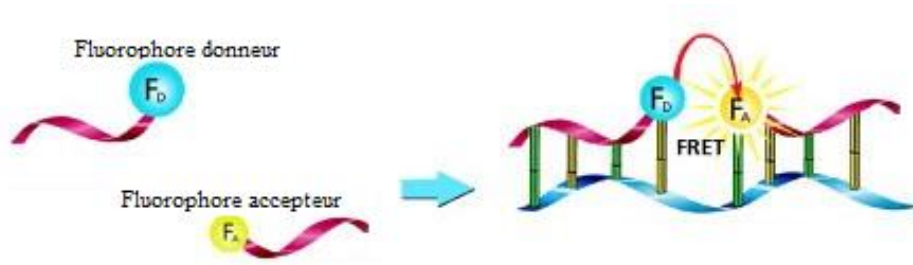


Figure 13: Principe de sondes FRET

La limite des systèmes présentés reste la nécessité de travailler dans des conditions dénaturantes empêchant alors leurs utilisations *in vivo*.

### B Sondes à ADN double brins

On dénombre trois grandes classes de sondes luminescentes permettant de cibler de l'ADN double brin et donc d'opérer en conditions non-dénaturantes :

- Les TFO : oligonucléotides formant des triplexes
- Les PPI : polyamides de pyroles et imidazoles
- Les sondes à base de polypeptides modifiés ou de protéines

#### i Oligonucléotides pour la formation de triplexes (TFO)

La première catégorie de sondes est basée sur la formation d'une triple hélice d'ADN entre le duplex d'ADN cible et un oligonucléotide TFO (triplex forming oligonucleotide)<sup>[20]</sup>. Les TFO sont capables de reconnaître de longues séquences d'ADN double brin (12-30 paires de bases), par interaction dans le grand sillon (Figure 14 gauche). Il est par exemple possible de concevoir une sonde en associant un fluorophore intercalant au TFO<sup>[21]</sup>. En solution, le fluorophore se désexcite préférentiellement par voie non radiative à cause des rotations autorisées. Lors de la formation du triplex le fluorophore s'intercale entre les paires de bases d'ADN, empêchant la rotation et favorisant la désexcitation par voie radiative (Figure 14 droite).

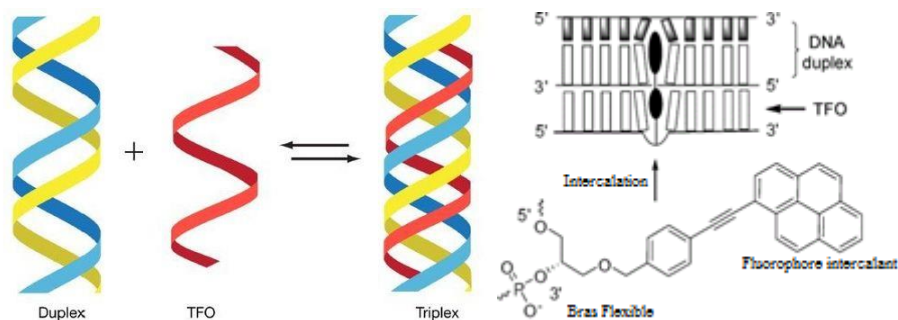


Figure 14: Principe des TFO ; à gauche interaction dans le grand sillon de l'oligonucléotide ; à droite principe d'un fluorophore intercalant<sup>[21]</sup>

L'interaction de la sonde pour la formation de la triple hélice requiert cependant la formation d'une liaison hydrogène supplémentaire, et donc un pH acide pour la protonation de la cytosine. Cette technique comporte plusieurs barrières, même si certaines approches ont été développées afin de les contourner<sup>[22]</sup>. On peut par exemple citer les contraintes d'une cinétique de formation très lente, l'instabilité de la triple hélice formée, les conditions de pH nécessaires, la dégradation possible de l'oligonucléotide.

### ii *Pyrole Imidazole Polyamides(PPI)*

Le second type de sonde est basé sur des polyamides. Ces polymères composés de pyrroles ou d'imidazole sont capables d'interagir avec une séquence définie d'ADN au sein du petit sillon<sup>[23],[24]</sup>. Une large gamme de fluorophores peuvent être introduits sur le polyamide<sup>[25]</sup>, qui lors de l'interaction avec l'ADN vont permettre la détection d'un signal fluorescent. Un exemple typique est la sonde décrite par P. Dervan<sup>[26]</sup>, avec l'introduction à une extrémité du polyamide d'un fluorophore intercalant (Figure 15). Ce type de sondes ne peut cependant pas discriminer les paires de bases A/T et T/A, posant un problème de sélectivité. Des améliorations ont été apportées ces deux dernières décennies, comme l'utilisation d'hydroxypyroles en plus du pyrole et de l'imidazole, pour permettre une lecture plus fidèle de l'ADN<sup>[27],[28]</sup>, mais au détriment de la stabilité du complexe.

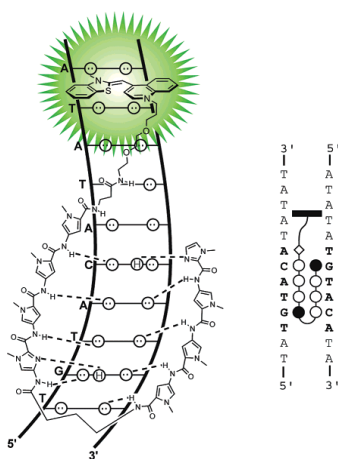


Figure 15: Principe des PPI avec la lecture de la séquence d'ADN par la chaîne, comportant un fluorophore intercalant représenté en vert<sup>[26]</sup>

### iii *Sondes à base de polypeptides modifiés ou de protéines*

Une autre catégorie de sondes est basée sur l'utilisation de petites protéines ou des séquences peptidiques porteuses de bases nucléiques, capables d'interagir avec une séquence spécifique d'ADN.

### (a) PNA Opener

L'interaction avec une séquence spécifique d'un double brin d'ADN est par exemple possible avec les PNA opener <sup>[29],[30]</sup> (PNA = peptide nucleic acid). Les PNA sont des peptides composés d'acides aminés dont les chaînes latérales sont des bases nucléiques (Figure 16). Un PNA complémentaire d'un ADN simple-brin peut s'hybrider avec celui-ci et peut même ouvrir un double brin d'ADN pour s'hybrider à son brin complémentaire. Le double brin d'ADN est déstabilisé par la formation d'un duplex PNA/ADN, ce dernier étant plus stable que le duplex ADN/ADN par suppression du phénomène de répulsion des groupements phosphates chargés négativement. La séquence ainsi libérée peut ensuite être mise en évidence par son interaction avec un molecular beacon, dont l'ouverture déclenche une réponse lumineuse (Figure 17).

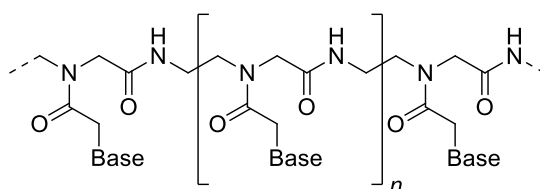


Figure 16: Structure d'un PNA

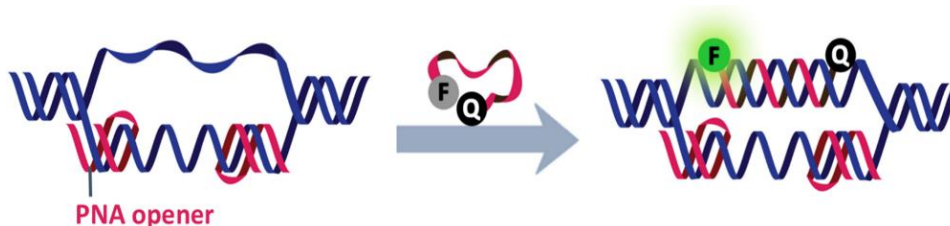


Figure 17: Principe d'un PNA Opener, avec interaction d'un molecular beacon<sup>[31]</sup>. F=Fluorophore ; Q=Quencher

Le principal désavantage de cette technique est une interaction des PNA avec l'ADN limitée à des séquences d'homo-purines ou d'homo-pyrimidines<sup>[32]</sup>.

### (b) TALE et doigts de zinc

Une dernière catégorie de sondes permettant de cibler l'ADN de manière séquence spécifique sont celles reposant sur l'utilisation de facteurs de transcription, tels que les Transcription Activating Like Effector (TALE) et les doigts de zinc.

Les TALE sont des protéines de 33-35 acides aminés capables de reconnaître une paire de base d'ADN. La fusion de TALE avec des protéines fluorescentes telles que la très utilisée GFP (Green Fluorescent Protein) a permis l'obtention de sondes fluorescentes<sup>[33]</sup>. Le principe de ces sondes reste similaire aux sondes FRET introduites en partie « Sondes FRET » (p.17), par dimérisation de deux protéines comportant un couple de fluorophores donneur/accepteur. Cet

exemple reste anecdotique, la majorité des TALE décrits étant couplés à des protéines de type nucléase actives sous forme de dimères<sup>[34]</sup>. Par exemple, la sonde TALEN (Transcription Activating Like Effector Nuclease) ainsi formée est capable de réaliser une coupure au niveau de la séquence d'ADN cible (Figure 18).

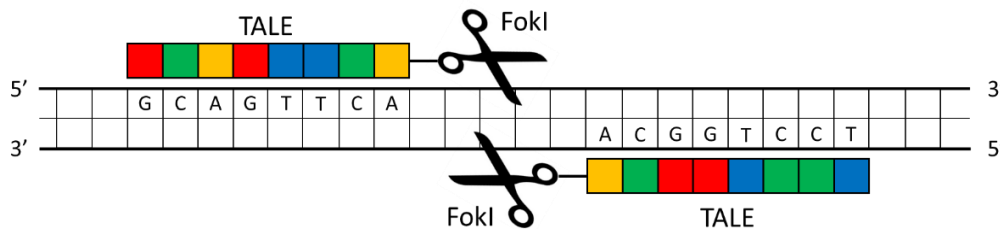


Figure 18: Principe d'un TALEN. Les TALE sont décorées par un système de nucléase complémentaire, qui lors de l'interaction avec la séquence d'ADN cible devient actif

Le même principe a été utilisé pour l'élaboration de nucléases artificielles avec des protéines à doigts de zinc<sup>[35]</sup>. Ces dernières sont fusionnées à un domaine de coupure qui se retrouve activé lors de leur dimérisation induite par l'interaction avec l'ADN (Figure 19). L'avantage de l'utilisation des doigts de zinc est qu'ils permettent de cibler trois paires de bases d'ADN avec une protéine de 30 acides aminés. Pour cette raison, nous nous intéressons à l'utilisation de doigts de zincs pour le développement de nos sondes lumineuses.

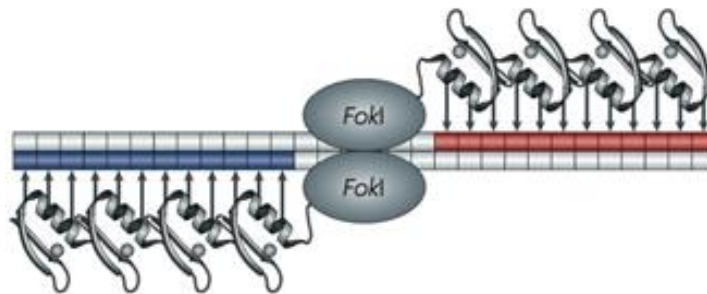


Figure 19: Principe des nucléases à doigt de zinc<sup>[35]</sup>. Les protéines à doigt de zinc permettent la lecture de l'ADN, qui en présence de la séquence cible d'ADN rend le domaine nucléase actif après dimérisation

Une sonde fluorescente basée sur un système similaire a été décrit par le groupe de D. Segal en 2005<sup>[36]</sup>. Deux protéines à doigts de zinc, l'une fusionnée avec le domaine N-terminal de la GFP et l'autre avec le domaine C-terminal dimérisent en présence de la séquence d'ADN double brin cible, ce qui permet de reconstituer une GFP active capable d'émettre sa fluorescence (Figure 20).

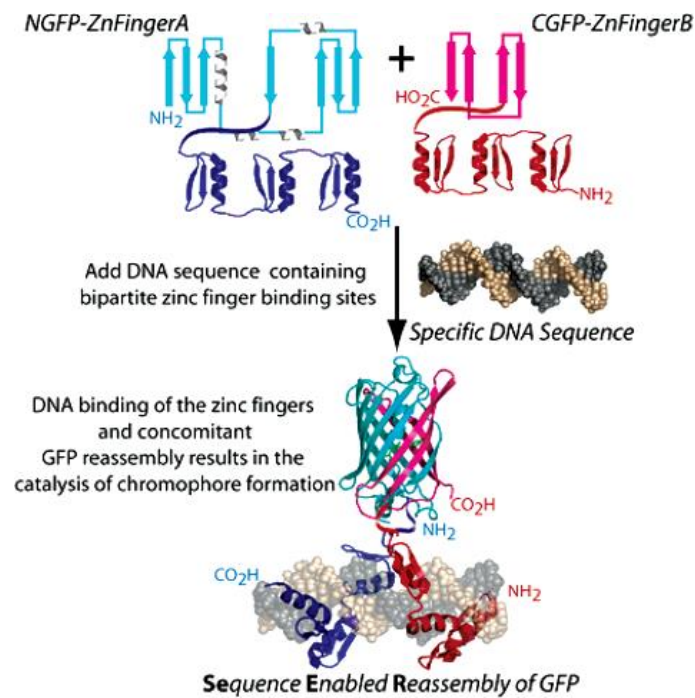


Figure 20 : Principe d'une sonde luminescente pour la détection d'ADN double brin basée sur des protéines à doigts de zinc<sup>[36]</sup>.

## IV Les doigts de zinc

### 1) Description

Dans ce projet, l'approche pour la détection d'une séquence d'ADN double brin est réalisée par l'utilisation de doigts de zinc. Les doigts de zinc possèdent des domaines de reconnaissance pouvant être n'importe quelle séquence d'ADN. Ce sont des petits domaines protéiques (20-30 acides aminés) dont la forme la plus répandue dans le vivant est du type  $\beta\beta\alpha$ <sup>[37],[38],[39]</sup>, c'est-à-dire adoptant une conformation d'un feuillet  $\beta$  à deux brins et d'une hélice  $\alpha$  (Figure 21). La séquence consensus de ces doigts de zinc est (Tyr/Phe)-Xaa-Cys-Xaa<sub>2,4</sub>-Cys-Xaa<sub>3</sub>-Phe-Xaa<sub>5</sub>-Leu-Xaa<sub>2</sub>-His-Xaa<sub>3,5</sub>-His, où Xaa peut correspondre à n'importe quel acide aminé (Figure 22). Cette séquence comporte donc 7 acides aminés indispensables pour la formation du doigt de zinc : 4 acides aminés coordonnant le zinc (2 cystéines et 2 histidines) et 3 acides aminés formant un cœur hydrophobe (tyrosine, phénylalanine, leucine). L'espacement entre ces acides aminés est essentiel pour le bon repliement de la protéine et la complexation du zinc. Les autres acides aminés peuvent par contre être changés à loisir sans affecter l'intégrité de la formation du doigt de zinc  $\beta\beta\alpha$ <sup>[40]</sup>.

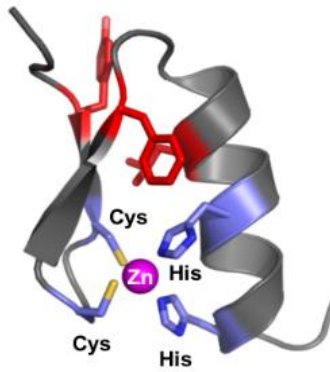


Figure 21: structure d'un doigt de zinc  $\beta\beta\alpha$ <sup>[41]</sup>

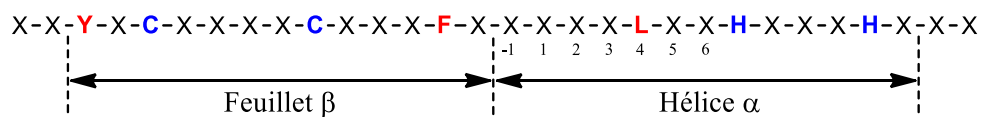


Figure 22: séquence d'un doigt de zinc  $\beta\beta\alpha$  (X peut être n'importe quel acide aminé)

De nombreux facteurs de transcription utilisent ces doigts de zinc pour « lire » l'ADN. Un doigt de zinc est capable d'interagir avec un triplet de paires de bases d'ADN. Dans le vivant, 3 à 6 motifs doigts de zinc sont en général enchainés, permettant ainsi la lecture de 9 à 18 paires de bases consécutives. Ce sont les acides aminés aux positions -1, 3 et 6 de l'hélice qui interagissent chacun avec une paire de bases, et l'acide aminé en position 2 est lui capable d'interagir avec la dernière paire de bases du triplet précédent<sup>[38],[42],[43]</sup> (Figure 23 gauche). Les acides aminés peuvent être changés pour cibler la séquence de trois paires de bases choisie sans affecter les propriétés de repliement du doigt de zinc. Un code de reconnaissance entre les acides aminés à ces positions et les paires de bases reconnues a été proposé (Figure 23 droite) mais ne fonctionne pas parfaitement. Des bases de données de doigts de zinc fonctionnant bien ensemble et des logiciels pour optimiser la séquence peptique en fonction de la séquence d'ADN à reconnaître sont disponibles sur divers sites web, comme par exemple « the Zinc Finger Consortium » (<http://www.zincfingers.org>), « ZiFiT Targeter » (<http://zifit.partners.org/ZiFiT/Introduction.aspx>) et « B1H screen of C2H2-ZF domains » (<http://zf.princeton.edu/b1h/index.html>). C'est en utilisant ce principe de modulation de la reconnaissance que les nucléases à doigt de zinc ont été développées<sup>[35]</sup>.

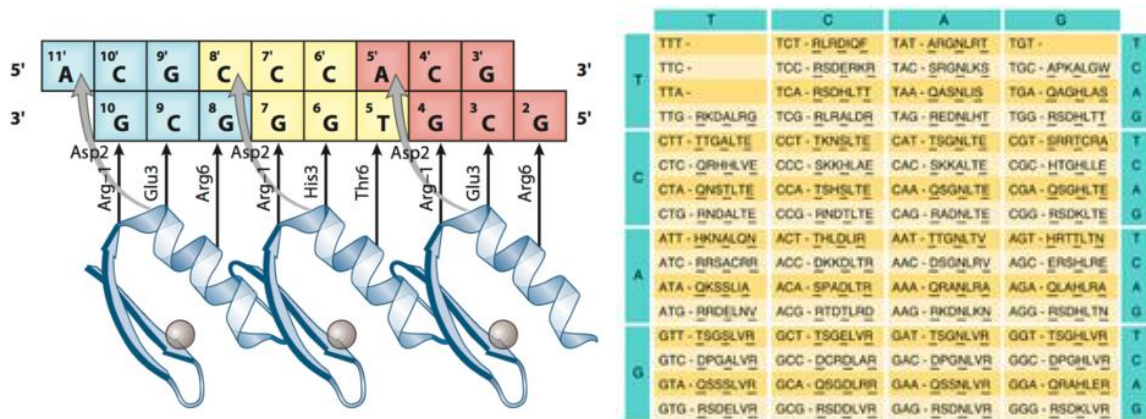


Figure 23: A gauche, interaction de trois doigts de zinc avec 9 paires de bases d'ADN (les acides aminés en positions -1, 2, 3, et 6 sont représentés)<sup>[38]</sup>. A droite code de reconnaissance des acides aminés avec la séquence de trois paires de bases choisie<sup>[44]</sup>.

## 2) La protéine ZIF12

Parmi les protéines à doigts de zinc décrites dans la littérature, une protéine de 72 acides aminés ZIF12, basée sur la protéine Zif268, décrite par B. Wang et C. Pabo<sup>[45],[46]</sup> semble très intéressante. Cette protéine est composée de deux doigts de zinc de 30 et 37 acides aminés (a.a) et d'une extension N-terminale de 15 a.a (Figure 24). Les deux doigts de zinc correspondent à la protéine Zif268, un facteur de transcription connu pour cibler les séquences 5'-TGG-3' et 5'-GCG-3'<sup>[37]</sup>. La partie N-terminale, elle, a été développée par B. Wang et C. Pabo afin de permettre la dimérisation de la protéine en présence d'une séquence palindromique d'ADN de séquence 5'-TGGGCGCGCCCA-3' (Figure 25). Cette extension interagit avec le doigt de zinc N-terminal de l'autre monomère (feuillet et hélice). Grâce à cette extension N-terminale, la protéine possède une affinité élevée en présence de la séquence d'ADN cible, avec une constante de dissociation de 15nM déterminée par électrophorèse (retard sur gel), contre 15μM pour la même protéine sans l'extension<sup>[45]</sup>.



Figure 24: Séquence de la protéine ZIF12 décrite par Wang et Pabo

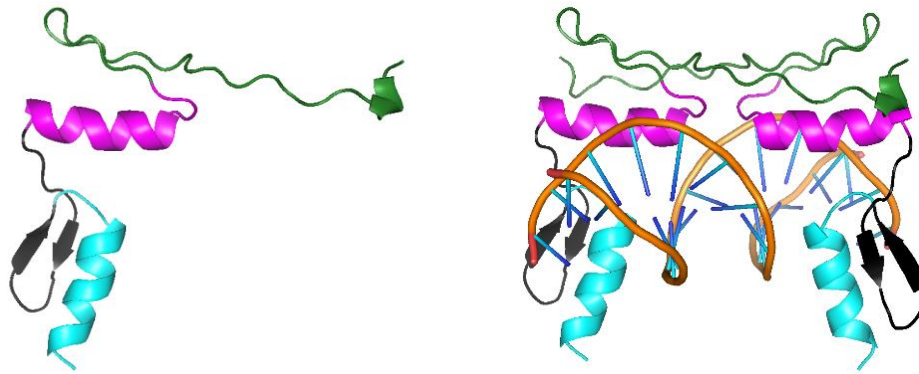


Figure 25: A gauche, protéine ZIF12. A droite, structure cristalline du dimère de ZIF12 autour de la séquence d'ADN 5'-TGGGCGC-GCCCA-3' (noté complexe ADN-(ZIF12)<sub>2</sub>)

La structure cristallographique de la structure du complexe ADN-(ZIF12)<sub>2</sub> obtenue par B. Wang et C. Pabo<sup>[46]</sup> montre deux caractéristiques intéressantes. Premièrement, l'interface de dimérisation entre l'extension N-terminale et le doigt de zinc de l'autre monomère implique l'hélice  $\alpha$  du doigt de zinc, mais pas les acides aminés interagissant avec l'ADN (Figure 26). Secondement, l'extension N-terminale n'est pas impliquée dans une quelconque interaction avec l'ADN. Ainsi, les acides aminés aux positions -1, 2, 3, 6 des doigts de zinc de cette protéine peuvent être modifiés sans affecter les propriétés de dimérisation, ouvrant alors la porte au ciblage de n'importe quelle séquence d'ADN.

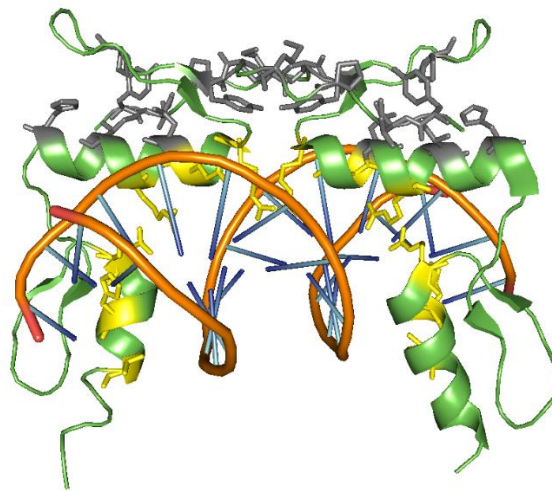


Figure 26: Structure du complexe ADN-(ZIF12)<sub>2</sub>, avec en gris les acides aminés de l'interaction entre l'extension N-terminale et l'hélice de l'autre monomère, et en jaune les acides aminés interagissant avec l'ADN

### 3) Doigts de zinc, ADN et FRET : ZIF12 comme support de sonde FRET

La protéine ZIF12 a été utilisée par l'équipe de H. Ueda en 2008<sup>[47]</sup> pour détecter la séquence palindromique 5'-TGGGCGCGCCCA-3' par un système de FRET. Pour cela, les protéines fluorescentes eCFP (Cyan) et eYFP (Jaune) ont été fusionnées à l'extrémité C-terminale de la

protéine ZIF12, pour donner un couple sonde ZIF12-CFP/ZIF12-YFP capable de se lier au palindrome (Figure 27). Ces protéines ont été obtenues par expression recombinante. Le ratio d'émission de fluorescence YFP/CFP augmente lorsque l'ADN cible est ajouté au couple de protéines sondes. Néanmoins, ce ratio varie peu, de 0,83 à 1,27, rendant la détection peu sensible, ce qui est un défaut récurrent des sondes FRET basées sur les protéines fluorescentes. Avec de tels systèmes, en milieu biologique, la détection devient peu sensible du fait de la fluorescence intrinsèque du milieu.

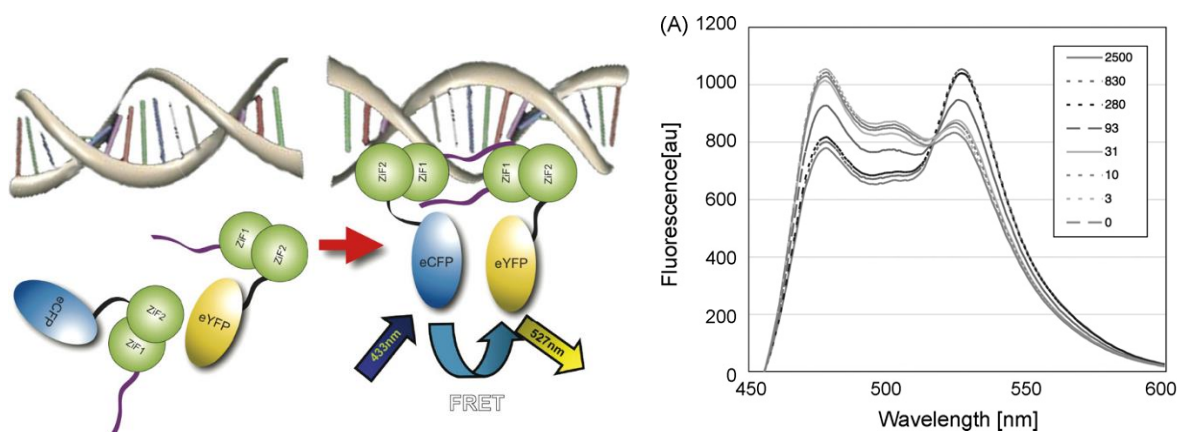


Figure 27: Sonde (ZIF12-CFP/ZIF12-YFP) décrite par l'équipe de H. Ueda<sup>[47]</sup> pour la détection du duplex de séquence 5'-TGGGCGCGCCCA-3'. A gauche : principe de fonctionnement. A droite : changement du spectre d'émission provoqué par l'ajout d'ADN (0 à 2500 nM) sur un mélange 5:1 de ZIF12-CFP (120μM) et ZIF12-YFP (600μM).

Le but de cette thèse était d'utiliser la protéine la protéine ZIF12 comme support pour la détection d'ADN double brin, à l'instar de ce qu'a proposé H. Ueda mais en implémentant un système FRET mettant en jeu des lanthanides(III) plutôt que des protéines fluorescentes afin d'améliorer les propriétés de luminescence et la détection. En effet, les ions lanthanide(III) ont des propriétés de luminescence très intéressantes pour la détection en milieu biologique, avec notamment de longs temps de vie de luminescence permettant d'éliminer la fluorescence intrinsèque du milieu lors des mesures.

## V Les lanthanides

Les lanthanides sont les 15 premiers éléments du bloc f de la classification périodique des éléments (Figure 28) allant du lanthane (Z=57) au lutécium (Z=71). Ils existent principalement sous le degré d'oxydation +III et sont notés  $\text{Ln}^{3+}$ . Leur configuration électronique est  $[\text{Xe}]4f^n$ , avec  $n=0$  pour le lanthane et  $n=14$  pour le lutécium.

|    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| H  |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    |    | He |
| Li | Be |    |    |    |    |    |    |    |    |    |    | B  | C  | N  | O  | F  | Ne |
| Na | Mg |    |    |    |    |    |    |    |    |    |    | Al | Si | P  | S  | Cl | Ar |
| K  | Ca | Sc | Ti | V  | Cr | Mn | Fe | Co | Ni | Cu | Zn | Ga | Ge | As | Se | Br | Kr |
| Rb | Sr | Y  | Zr | Nb | Mo | Tc | Ru | Rh | Pd | Ag | Cd | In | Sn | Sb | Te | I  | Xe |
| Cs | Ba | La | Hf | Ta | W  | Re | Os | Ir | Pt | Au | Hg | Tl | Pb | Bi | Po | At | Rn |
| Fr | Ra | Ac | Rf | Db | Sg | Bh | Hs | Mt | Ds | Rg | Cn |    |    |    |    |    |    |

|    |    |    |    |    |    |    |    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| Ce | Pr | Nd | Pm | Sm | Eu | Gd | Tb | Dy | Ho | Er | Tm | Yb | Lu |
| Th | Pa | U  | Np | Pu | Am | Cm | Bk | Cf | Es | Fm | Md | No | Lr |

Figure 28: Classification périodique des éléments; en violet les lanthanides

### 1) Propriétés fondamentales

Les orbitales 4f possèdent une faible extension radiale et sont blindées par les couches externes  $5s^2 5p^6$ , les rendant peu sensibles à leur environnement chimique et conférant une chimie de coordination quasi-identique à l'ensemble des lanthanides. Le remplissage progressif de la couche 4f entraîne une diminution du rayon ionique due à une charge partielle ressentie par les électrons  $5s^2 5p^6$  plus importante : c'est la contraction des lanthanides.

Les liaisons formées entre les lanthanides et leurs ligands sont purement électrostatiques et les ligands s'organisent autour du métal de manière à minimiser les interactions stériques répulsives. Les ions  $\text{Ln}^{3+}$  sont des acides durs suivant la classification de Pearson<sup>[48]</sup>. Ils interagissent préférentiellement avec des ligands durs tels que les atomes d'oxygène ou d'azote de manière à minimiser l'énergie électrostatique. Le nombre de coordination le plus fréquemment rencontré pour les ions  $\text{Ln}^{3+}$  est de 8 ou 9. Du fait de leur propriétés très similaires, un même ligand peut généralement complexer tous les lanthanides<sup>[49]</sup>.

On peut noter ici que la forte acidité des lanthanides au sens de Lewis favorise la formation d'hydroxocomplexes solubles  $\text{Ln}(\text{H}_2\text{O})_8(\text{OH})$  pour des pH allant de 3 à 6<sup>[50]</sup>. Les hydroxocomplexes de type  $\text{Ln}(\text{OH})_3$ , eux, sont insolubles et se forment facilement à pH neutre.

Pour une majeure partie des lanthanides, le remplissage des électrons de la couche 4f laisse des électrons célibataires et rend les ions lanthanides paramagnétiques, à l'exception du lanthane et du lutécium. Parmi l'ensemble des lanthanides, l'ion  $\text{Gd}^{3+}$  possède 7 électrons non appariés, lui conférant un moment magnétique élevé et un temps de relaxation électronique long par rapport aux autres éléments du bloc f<sup>[51]</sup>. Pour cette raison, il est très utilisé comme agent de contraste pour l'Imagerie par Résonance Magnétique (IRM)<sup>[52]</sup>.

## 2) Propriétés spectroscopiques

### A Emission des lanthanides

Tous les ions lanthanides sont luminescents, à l'exception du  $\text{La}^{3+}$  et du  $\text{Lu}^{3+}$ . L'émission des lanthanides est dû à trois types de transitions électroniques<sup>[53]</sup>:

- les transitions f-f, qui sont théoriquement interdites par la règle de Laporte,
- les transitions f-d, autorisées, et
- les transitions dues à des transferts de charge métal-ligand (MLCT) ou ligand-métal (LMCT), autorisées.

Les transitions f-d et de transfert de charge sont trop hautes en énergie pour être observées par spectrophotométrie UV-visible<sup>[53]</sup>. La règle de Laporte stipule que les transitions f-f, comme toute transition entre états de même parité dans un système centrosymétrique, sont interdites. Ces règles sont partiellement relaxées lorsque l'environnement du lanthanide est non centrosymétrique<sup>[54]</sup> et des transitions f-f peuvent être provoquées par l'absorption d'un photon. Néanmoins, les coefficients d'extinction molaire de ces transitions sont très faibles ( $< 10 \text{ M}^{-1} \text{cm}^{-1}$ ), rendant l'excitation directe des lanthanides particulièrement inefficace. Les lanthanides sont luminescents : le retour à l'état fondamental suivant une transition f-f peut se faire de manière radiative. Les raies d'émission sont fines et à longueur d'onde fixe, les orbitales  $5s^2 5p^6$  écrantant les orbitales f rendent ces dernières peu sensibles au champ de ligand. Ainsi, chaque lanthanide a un spectre d'émission caractéristique avec des raies quasi-monochromatiques à longueur d'onde fixe comme cela est représenté sur la Figure 29.

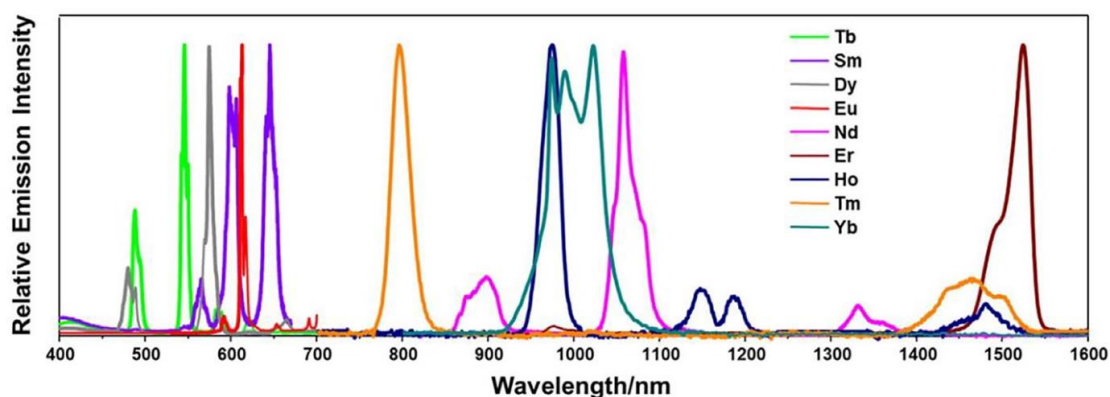


Figure 29: Spectres d'émission normalisés de complexes de lanthanides en solution<sup>[55]</sup>

### B Temps de vie de luminescence

Le retour à l'état fondamental du lanthanide excité étant un processus théoriquement interdit, le temps de vie de luminescence des lanthanides est souvent très long, de l'ordre de la microseconde pour ceux émettant dans le proche infrarouge et de la milliseconde pour ceux émettant dans le visible. Ceci permet de réaliser une détection de leur luminescence résolue en temps<sup>[56]</sup> (Figure 30). Ainsi, il suffit d'attendre quelques microsecondes après le flash d'excitation pour laisser s'éteindre la fluorescence du milieu biologique, puis de détecter l'émission du lanthanide seule. Ce type de détection permet d'améliorer considérablement le rapport signal/bruit, comme illustré sur la Figure 31<sup>[57]</sup>.

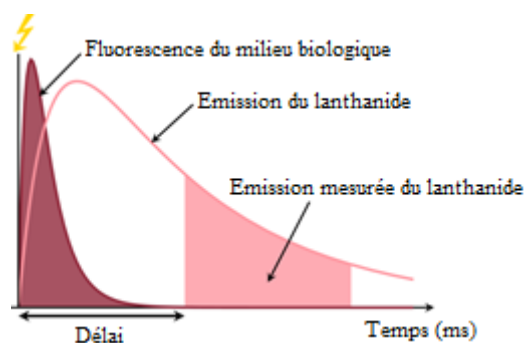


Figure 30: Principe de la détection résolue en temps. En violet, l'émission de fluorescence du milieu biologique. En rose l'émission du lanthanide. En attendant un délai après excitation, une mesure sélective de l'émission du lanthanide est possible.

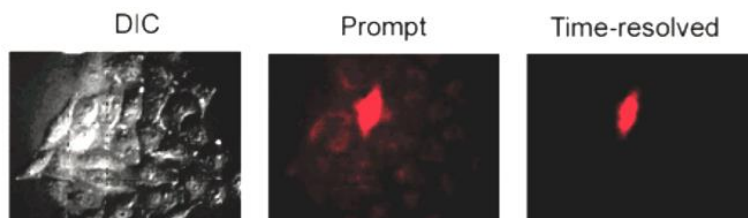


Figure 31 : Exemple de détection résolue en temps par microscopie<sup>[57]</sup>. A gauche, l'image montrée est un ensemble de cellules. Dans l'une d'elles seulement, un complexe d'euporium a été injecté. L'image au milieu est obtenue en détection « normale », c'est à dire sans délai après le flash : on voit une émission intense qui provient de la cellule contenant l'euporium mais aussi de la fluorescence des autres cellules. L'image à droite est enregistrée en détection résolue en temps. La fluorescence résiduelle des autres cellules n'est plus détectée.

Le temps de vie de luminescence des lanthanides est sensible à son environnement. En effet, les vibrations de liaisons O-H, N-H et C-H offrent des voies de désexcitations non radiatives au lanthanide ayant pour conséquence de diminuer son temps de vie<sup>[58]</sup>. Cette désexcitation est un phénomène à plusieurs phonons, la différence d'énergie du retour de l'état excité à l'état fondamental correspond à un multiple de l'énergie de vibration de la liaison. Par exemple, les vibreurs O-D possédant une énergie de vibration plus faible que les O-H rendent le phénomène moins probable et conduisent à une plus faible désactivation (Figure 32). Des équations phénoménologiques ont été établies et permettent de calculer le nombre de molécules d'eau présentes dans l'environnement du terbium et de l'euporium, à partir de leurs temps de vie de luminescence dans H<sub>2</sub>O et D<sub>2</sub>O<sup>[58],[59]</sup>.

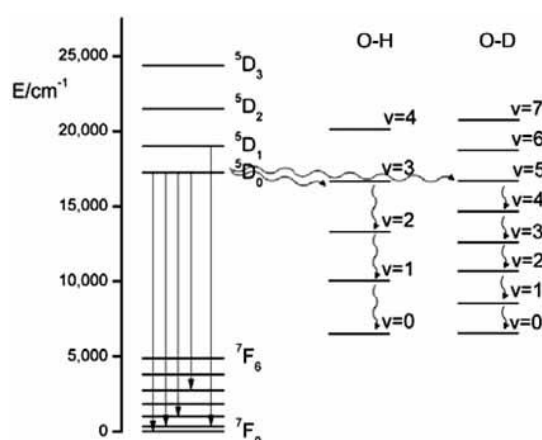


Figure 32: Exemple de voies de désexcitation par des oscillateurs O-H ou O-D de l'ion  $\text{Eu}^{3+}$  [60]

Le nombre de molécules d'eau présentes dans la sphère de coordination du lanthanide joue donc un rôle important et il est judicieux de pouvoir le minimiser en utilisant des ligands ne laissant pas de sites de coordination vacants.

### C Utilisation pour des sondes luminescentes

Les ions  $\text{Eu}^{3+}$  et  $\text{Tb}^{3+}$  sont les plus utilisés pour des applications de sondes luminescentes pour leurs longs temps de vie de luminescence, ainsi que leur émission dans les hautes longueurs d'ondes du visible (vert pour le terbium avec  $\lambda_{\text{em}} = 545 \text{ nm}$  et rouge pour l'euporium avec  $\lambda_{\text{em}} = 615 \text{ nm}$ ). Les ions  $\text{Dy}^{3+}$  ( $\lambda_{\text{em}} = 570 \text{ nm}$ ),  $\text{Sm}^{3+}$  ( $\lambda_{\text{em}} = 590 \text{ nm}$ ),  $\text{Nd}^{3+}$  ( $\lambda_{\text{em}} = 1065 \text{ nm}$ ),  $\text{Er}^{3+}$  ( $\lambda_{\text{em}} = 1510 \text{ nm}$ ),  $\text{Yb}^{3+}$  ( $\lambda_{\text{em}} = 980 \text{ nm}$ ) sont moins utilisés. Cependant ces trois derniers représentent un avantage pour la détection en milieu proche infra-rouge et sont de plus en plus présents<sup>[55]</sup>. Le diagramme partiel d'énergie entre les niveaux excités et les niveaux fondamentaux de ces ions sont présentés en Figure 33.

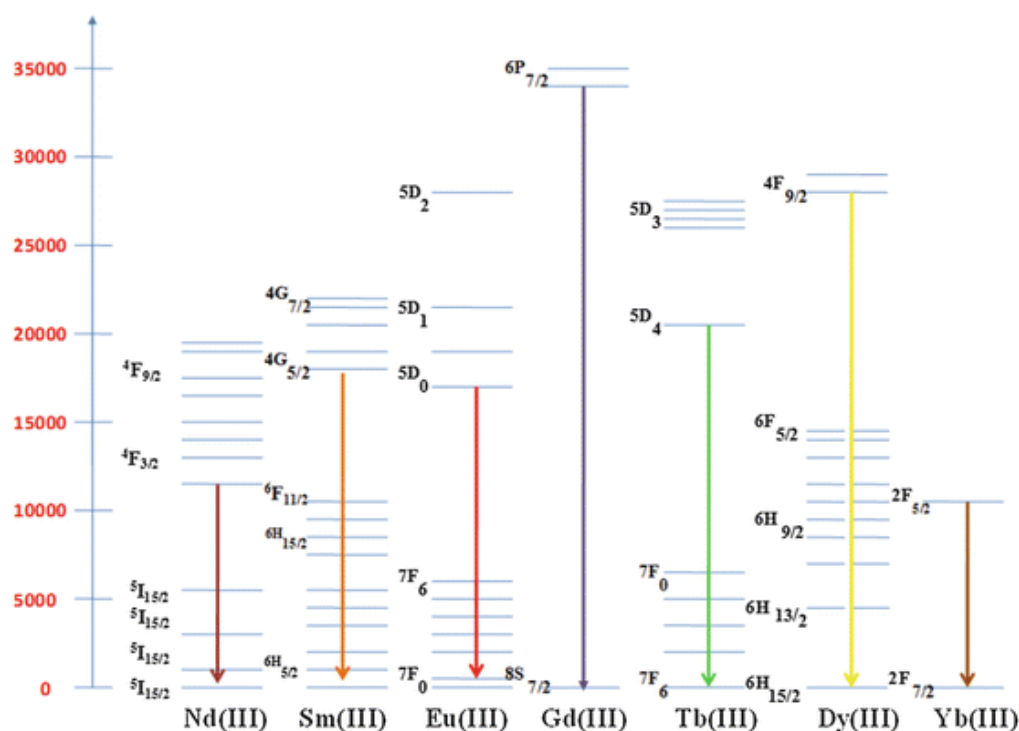


Figure 33: Diagramme partiel d'énergie pour les ions  $\text{Ln}^{3+}$  luminescents<sup>[61]</sup>

#### D Effet « antenne »

La voie la plus couramment utilisée pour exciter un lanthanide est par un transfert d'énergie provenant d'un chromophore organique (Figure 34). Cette stratégie appelée effet « antenne » découverte en 1942 par S. Weissman<sup>[62]</sup> permet d'utiliser un chromophore organique à haut coefficient d'extinction molaire, qui une fois irradié transmet son énergie au lanthanide. Le lanthanide peut ensuite se désexciter par voie radiative et non-radiative. L'effet « antenne » est efficace à courtes distances ( $R_0 < 15 \text{ \AA}$ )<sup>[63]</sup> et il est en général plus efficace pour des antennes coordonnées à l'ion  $\text{Ln}^{3+}$  (mécanisme de transfert d'énergie de type Dexter, à travers les orbitales) que pour des antennes distantes (mécanisme de type Förster par couplage dipolaire, à travers l'espace).

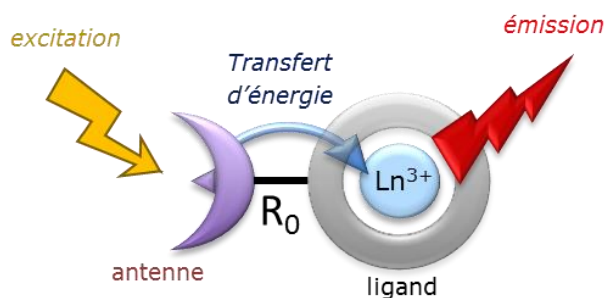


Figure 34: Principe de l'effet d'antenne, avec l'excitation du chromophore appelé antenne, suivi du transfert d'énergie au lanthanide et de son émission.

Lors de ce processus de sensibilisation, le chromophore absorbe un photon, puis relaxe pour se retrouver dans un état excité donneur (D) singulet, généralement  $S_1$  ou CT (transfert de charge), ou triplet ( $T_1$ ) (Figure 35).

L'état du donneur doit être adapté en énergie à l'état excité du lanthanide. Dans le cas d'un état donneur triplet  $T_1$ , des règles empiriques ont été développées pour le terbium et l'euprium<sup>[64],[65]</sup>. Pour le cas du terbium, le transfert est le plus efficace lorsque l'état triplet de l'antenne est situé entre 2 000 et 5 000  $\text{cm}^{-1}$  au-dessus de l'état accepteur  $^5D_4$  du terbium (20 400  $\text{cm}^{-1}$ ).<sup>[66]</sup> L'euprium possède plusieurs états excités proches en énergie pouvant être peuplés. Empiriquement, de bons transferts sont observés lorsque l'état triplet de l'antenne se situe au niveau de l'état  $^5D_2$  de l'euprium (21 500  $\text{cm}^{-1}$ ). Un transfert d'énergie « retour » du lanthanide vers l'état  $T_1$  de l'antenne (back transfer en anglais) est possible si ce dernier se trouve situé moins de 2 000  $\text{cm}^{-1}$  au-dessus de l'état excité émissif du lanthanide. Le transfert retour offre une voie de désexcitation supplémentaire au lanthanide, ce qui réduit son temps de vie.

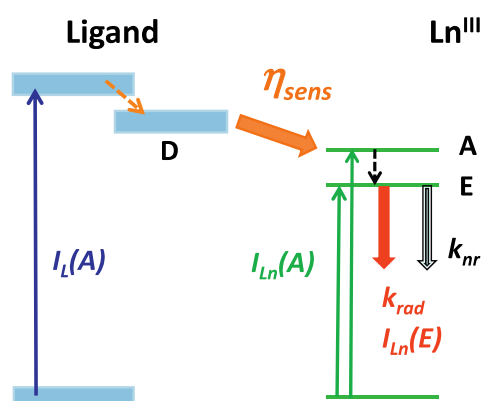


Figure 35: Diagramme de Jablonski-Perrin pour la sensibilisation d'un lanthanide ( $\text{Ln}^{\text{III}}$ ) par un ligand étant le chromophore

### 3) Sondes luminescentes responsives à base de lanthanide

Il existe plusieurs stratégies pour développer des sondes responsives, c'est à dire qui répondent à un analyte, à base de lanthanides. Il est possible de moduler l'émission d'un lanthanide lorsqu'un analyte :

- réduit ou augmente la distance antenne- $\text{Ln}^{3+}$  ( $R_0$  en Figure 34) en se liant à la sonde<sup>[67],[68],[69],[70]</sup>,
- change les propriétés photophysiques de l'antenne en se liant à celle-ci<sup>[71]</sup>,

- change la nature chimique de l'antenne par réaction chimique, altérant ainsi ses propriétés photophysiques<sup>[72],[73]</sup>,
- inhibe un transfert d'électron photoinduit qui désactive l'antenne<sup>[74],[75],[76]</sup>,
- perturbe les désexcitations non-radiatives du lanthanide en modifiant le nombre de molécules d'eau liée à l'ion  $\text{Ln}^{3+}$  <sup>[77],[78]</sup>,
- favorise une désexcitation non-radiative pour le lanthanide via un transfert d'énergie vers un chromophore quencheur ou un fluorophore<sup>[79],[80]</sup>.

Les systèmes FRET mettant en jeu un transfert d'énergie d'un lanthanide vers un fluorophore donnent des sondes très intéressantes pour la détection en milieu biologique avec une grande sensibilité<sup>[81],[82],[83],[84]</sup>. Le lanthanide, une fois sensibilisé par son antenne, transfère son énergie à un fluorophore, qui se désexcite en émettant de la lumière (Figure 36). Ce type de transfert lanthanide vers fluorophore possède des rayons de Förster élevés de l'ordre de 40 à 80 Å et permet de conserver l'avantage de la détection résolue en temps, la luminescence du fluorophore étant engendrée par le lanthanide excité.

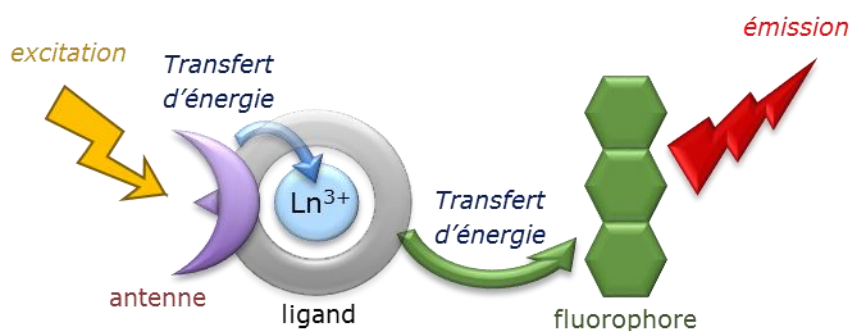


Figure 36: Principe du FRET d'un lanthanide vers un fluorophore

### VI Projet

Le projet vise à développer des sondes combinant les propriétés de reconnaissance d'ADN et de dimérisation de la protéine ZIF12 décrite par B. Wang et C. Pabo, et les propriétés de luminescence des lanthanides. Pour obtenir une réponse lumineuse sélective de la séquence d'ADN ciblée, deux types de FRET impliquant les lanthanides sont possibles :

- FRET d'une antenne vers un lanthanide (Figure 34)
- FRET d'un lanthanide vers un fluorophore (Figure 36)

#### 1) Sondes de première génération (FRET d'une antenne vers un lanthanide)

Le principe des sondes de première génération repose sur un FRET utilisant un transfert d'une antenne vers un lanthanide. Ce type de transfert n'étant efficace qu'à de courtes distances ( $< 15 \text{ \AA}$ ), l'introduction des chromophores à une faible distance l'un de l'autre sur le dimère formé est requise. Sur la structure RX du dimère (Figure 37), on observe que l'épingle du feuillet  $\beta$  du premier doigt de zinc (zone 2) d'un monomère est proche de l'extrémité de l'extension N-terminale de l'autre (zone 1). En introduisant une antenne et un lanthanide sur ces deux positions, le transfert d'énergie devrait être possible lors de la dimérisation. Ces positions permettent par ailleurs d'éviter la modification d'acides aminés importants pour la dimérisation.

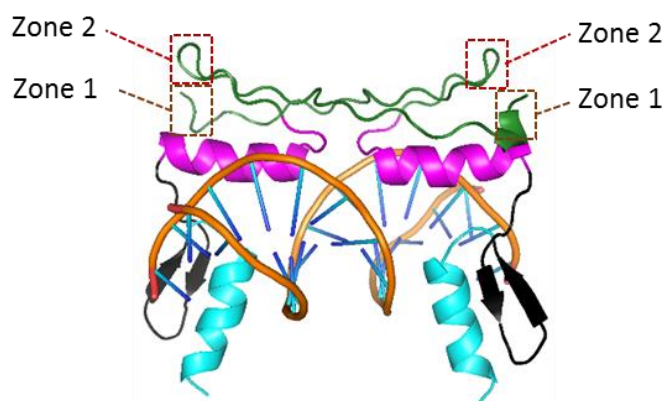


Figure 37: Positions possibles pour l'introduction des chromophores ne devant pas altérer la dimérisation

Avec cette stratégie il est possible d'envisager l'introduction du lanthanide et de son antenne sur un même monomère. Les zones 1 et 2 étant éloignées au sein d'un même monomère, le transfert d'énergie intramoléculaire de l'antenne au lanthanide devrait être peu efficace, alors que le transfert intermoléculaire ne pourra être efficace que lors de la formation du dimère. Cette approche permet de n'avoir à synthétiser qu'une protéine sonde pour la détection de la séquence d'ADN 5'-**TGGGCGCGCCCA**-3' (Figure 38).

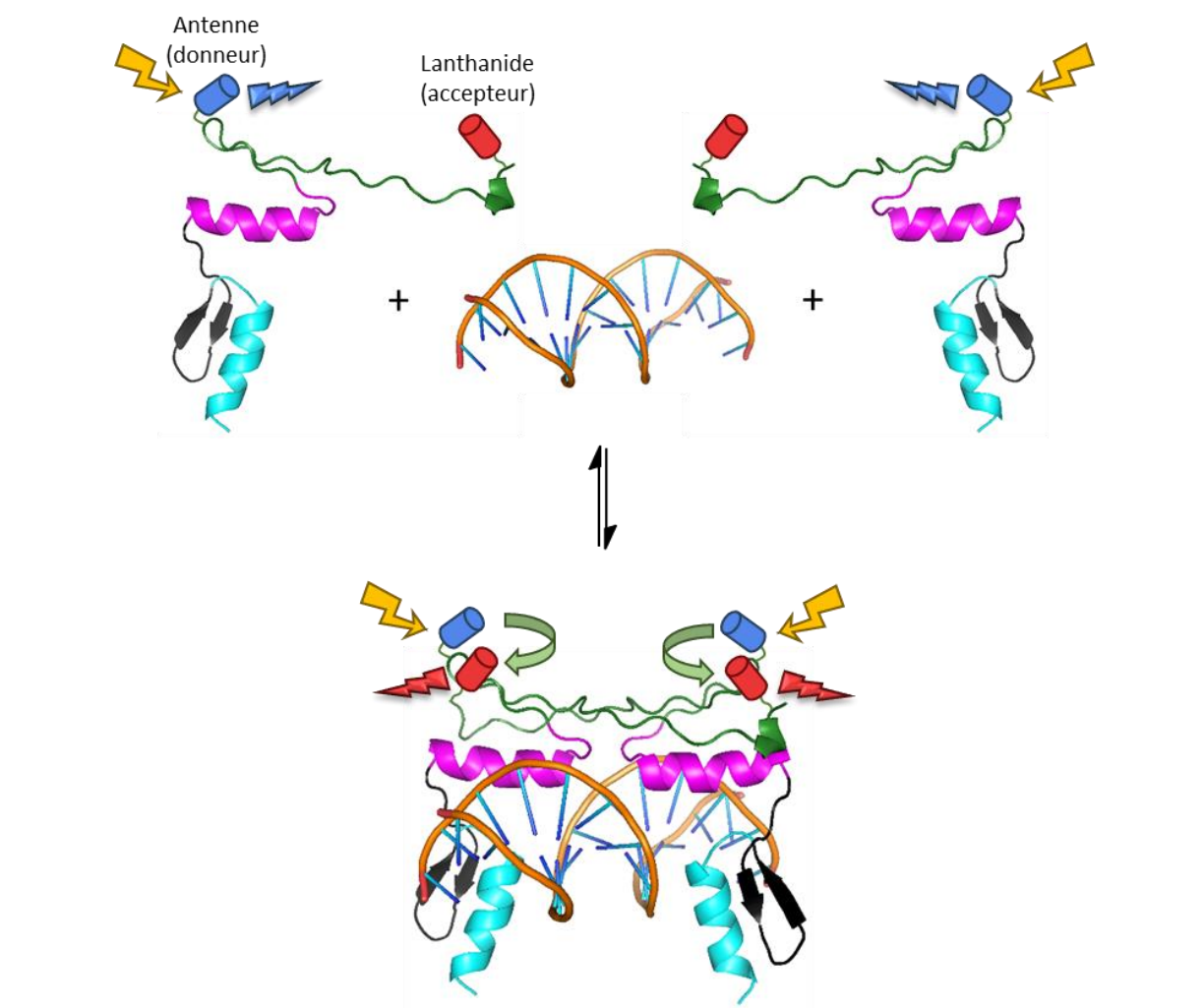


Figure 38: Principe de la sonde de première génération : en l'absence d'ADN, lorsque l'antenne de la sonde est excitée, son émission est observée ; en présence d'ADN, lorsque l'antenne est excitée, elle transfère son énergie au lanthanide fixé sur l'autre monomère. L'émission du lanthanide est alors mesurée.

Cette sonde permettra de mettre au point la synthèse par voie chimique de tels systèmes, mais aussi de montrer la possibilité de modifier certains acides aminés de la séquence initiale sans altérer la lecture de la séquence d'ADN, ni la dimérisation.

Si l'avantage de ce système est la synthèse d'une protéine unique servant de sonde, le rayon de Förster du FRET reste faible ( $<15\text{\AA}$ ). Le risque de ne pas rapprocher suffisamment l'antenne et le lanthanide lors de la formation du dimère est donc important. C'est pourquoi l'élaboration d'un autre système de FRET a été envisagé pour obtenir de meilleures réponses pour la suite du projet.

## 2) Sondes de seconde génération (FRET d'un lanthanide vers un fluorophore)

Comme dit plus tôt, le rayon de Förster d'un FRET lanthanide vers fluorophore est de l'ordre de 40-80Å. L'introduction d'un complexe de lanthanide (donneur FRET incorporant sa propre antenne) et d'un fluorophore (accepteur FRET) au sein d'un même monomère n'est cette fois plus possible car la taille de la molécule conduira inévitablement à un transfert efficace en l'absence d'ADN. L'introduction sélective d'un chromophore par monomère est donc nécessaire. En plaçant un complexe de lanthanide dans la zone 2 proposée dans la Figure 37 sur un monomère, puis le fluorophore sur une même position au sein de l'autre monomère, le transfert ne devrait être efficace que lors de la formation de l'hétéro-dimère (Figure 39).

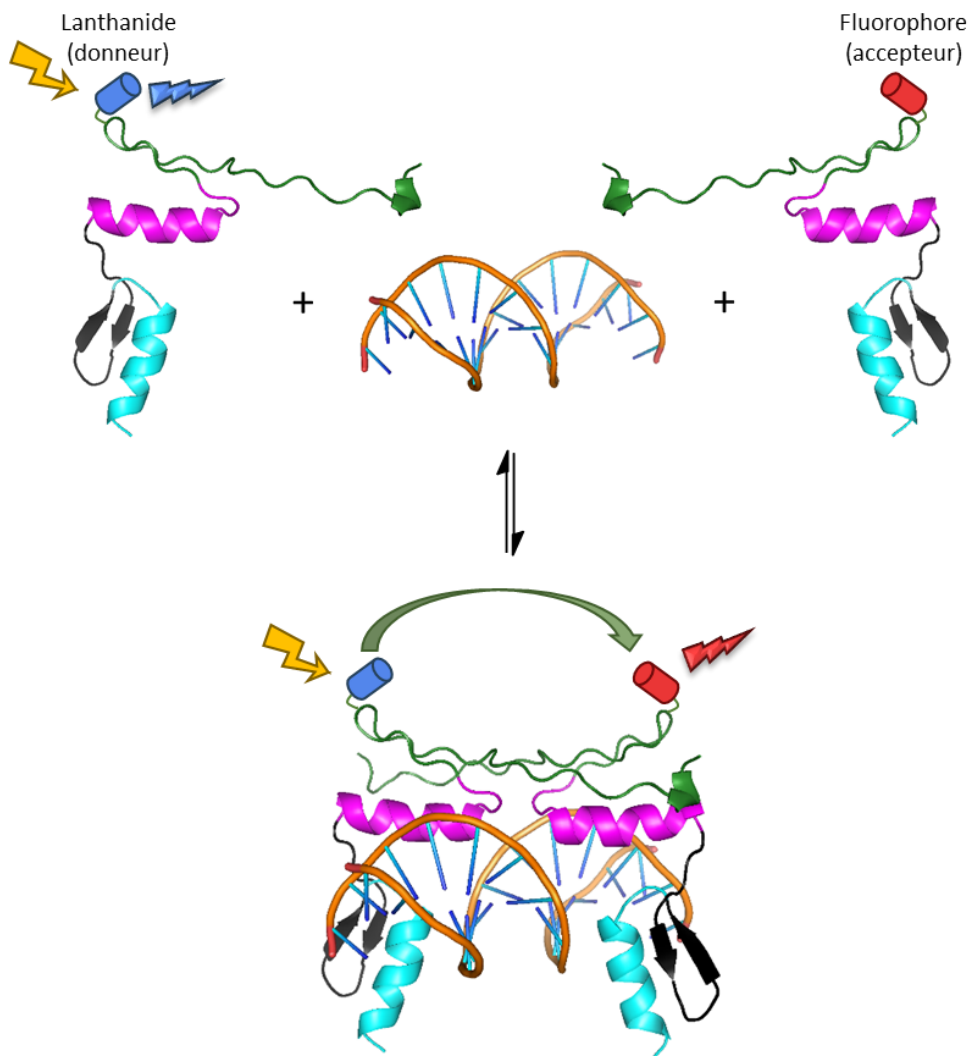


Figure 39: Principe de la sonde de seconde génération: en l'absence d'ADN, lorsque le complexe de lanthanide est excité, l'émission du lanthanide est observée ; en présence d'ADN, lorsque le complexe de lanthanide est excité, il transfère son énergie au fluorophore provenant de l'autre monomère. L'émission du fluorophore est alors mesurée.

Puisque la séquence d'ADN 5'-**TGGGCGCGCCCA**-3' ciblée est palindromique et que les deux monomères « donneur » et « accepteur » interagissent avec le même site de reconnaissance TGGGCG, la formation d'homo-dimères est possible (Figure 40). D'un point de vue statistique, la proportion d'hétéro-dimère formé permettant le transfert lanthanide vers fluorophore est de 50% ; le reste étant à 25% la formation d'un homo-dimère comportant le complexe de lanthanide, et 25% celui comportant le fluorophore.

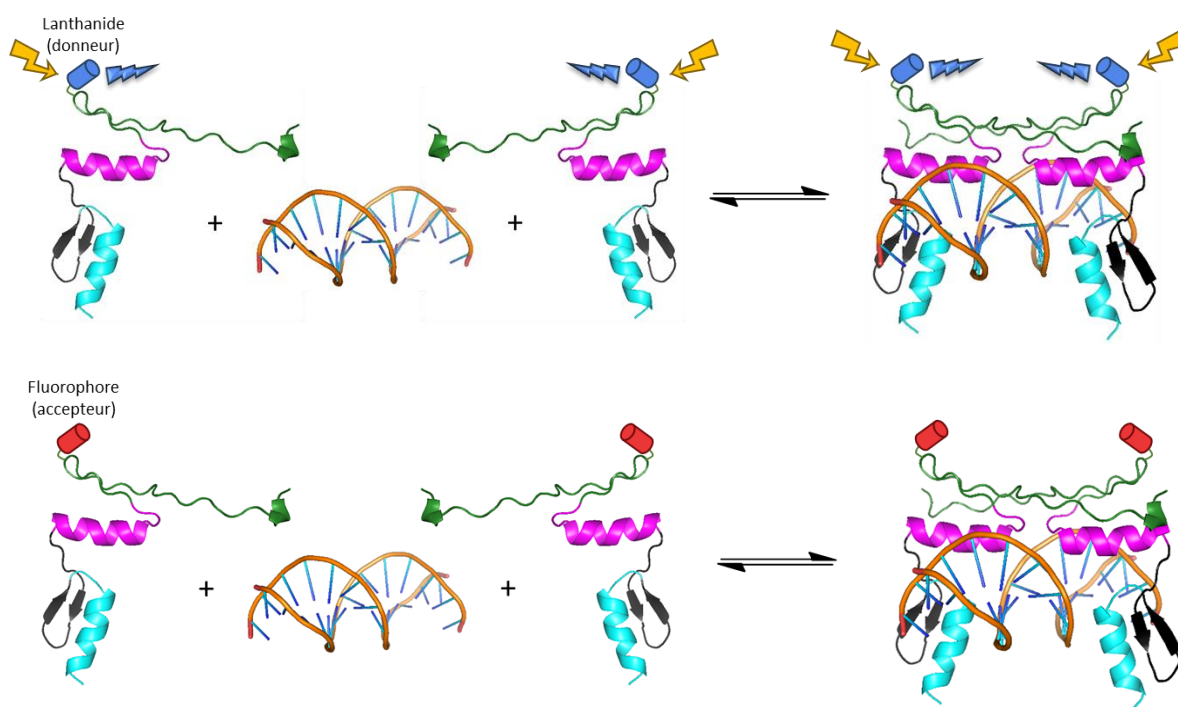


Figure 40: Formation d'homo-dimères possibles lors de l'interaction avec la séquence d'ADN cible

Pour éviter cette formation d'homo-dimère, il faut cibler une séquence non palindromique, ce qui représente l'objectif ultime du projet.

### 3) Cibler une séquence non palindromique

En synthétisant une sonde de seconde génération dont les acides aminés de reconnaissance d'ADN ont été modifiés pour cibler une autre séquence (par exemple 5'-**ACGCAC**-3'), la détection d'une séquence non palindromique d'ADN est possible. En effet, en mélangeant le monomère ciblant une séquence 5'-**TGGGCG**-3' comportant un complexe de lanthanide, avec un monomère comportant un fluorophore et ciblant une séquence 5'-**ACGCAC**-3', il se formera exclusivement un hétéro-dimère autour de la séquence d'ADN 5'-**TGGGCGACGCAC**-3', permettant un FRET du lanthanide vers le fluorophore (Figure 41).

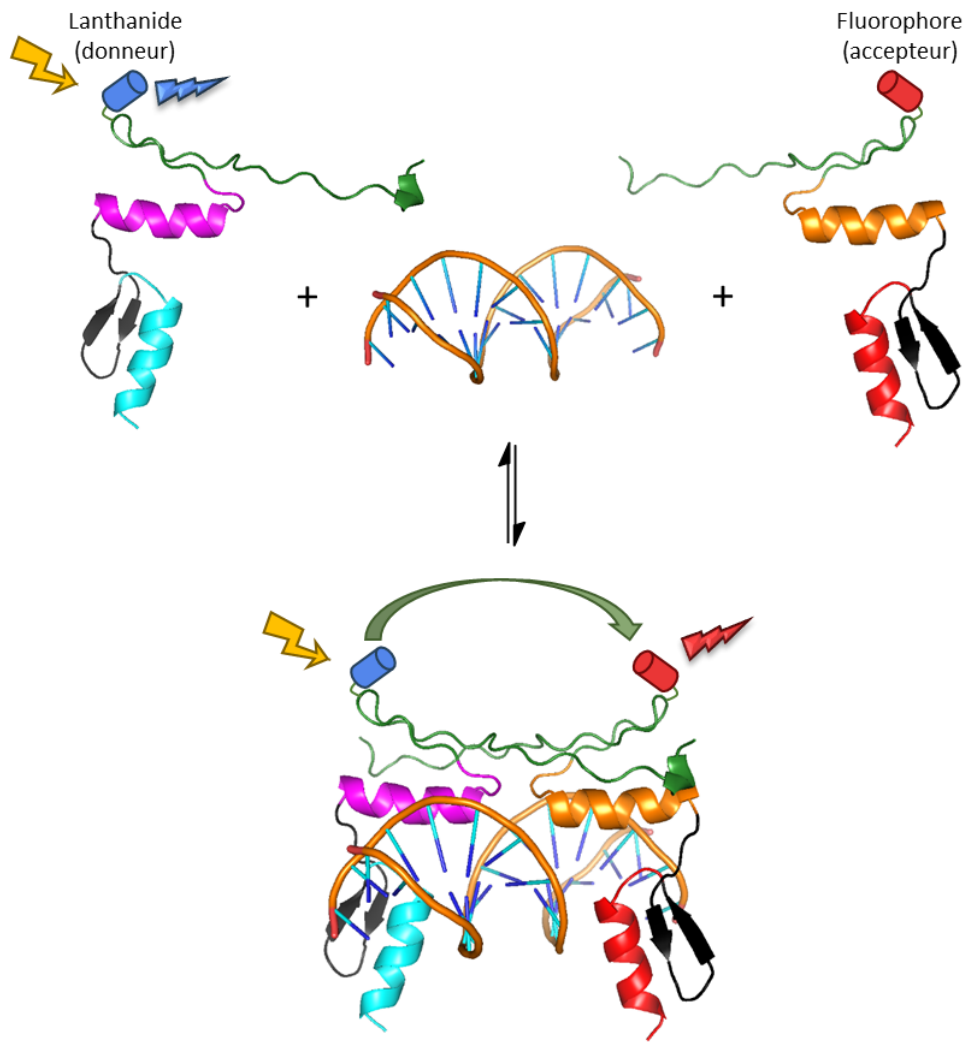


Figure 41: Formation d'un hétéro-dimère permettant de cibler une séquence non palindromique d'ADN

Ce travail de thèse se concentrera sur la preuve de concept démontrant la possibilité de décorer la protéine Zif12 avec des molécules lumineuses pour cibler sélectivement une séquence palindromique de 12 paires de bases par deux processus différents (FRET d'une antenne vers un lanthanide, puis FRET d'un lanthanide vers un fluorophore).





Chapitre I : Sondes de première  
génération, FRET d'une antenne vers  
un lanthanide



# Chapitre I : Sondes de première génération, FRET d'une antenne vers un lanthanide

## I Principe

Le principe des sondes de première génération repose sur un système de transfert d'énergie entre une antenne et un lanthanide, chaque monomère comportant en son sein l'antenne et le lanthanide (Figure 42). Cette approche permet de n'avoir qu'une seule protéine sonde pour la détection d'ADN, ce qui est un avantage en ce qui concerne la synthèse.

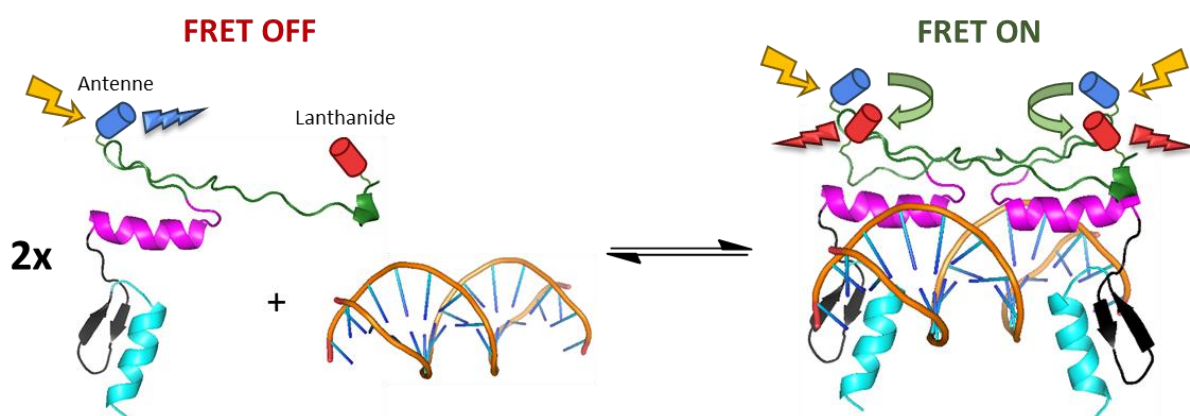


Figure 42: Principe d'une sonde de première génération

## II Insertion de chromophores sur la protéine

### 1) Choix des chromophores

Le système antenne/lanthanide à introduire sur la sonde doit répondre à deux critères :

- permettre un transfert d'énergie le plus efficace possible,
- être facile à introduire sur un peptide.

Le fait d'être commercial, ou facilement obtenu en quelques étapes de synthèse est un grand plus.

En 1995, le groupe R. Selvin a décrit qu'un composé de la famille des 2-quinolones (ou carbostyryl), noté Cs124 (Figure 43 en vert) est une excellente antenne pour la sensibilisation de la luminescence de l'euprimum et du terbium<sup>[85]</sup>. Récemment, le groupe de E. Borbas a montré que le Cs124 pouvait aussi

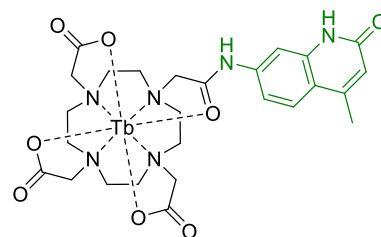


Figure 43: Complexe Cs124|DOTA[Tb] décrit par R. Selvin

sensibiliser l'ytterbium ou le neodyme<sup>[86]</sup>. Les complexes décrits par R. Selvin pour lesquels un bras d'un ligand DOTA (Figure 43), DTPA ou TTHA est amidé avec le Cs124, présentent des rendements quantiques élevés. Par exemple le ligand DOTA-Cs124, complexé à l'euprium possède un rendement quantique ( $\Phi$ ) de 41%, et de 73% avec le terbium<sup>[87]</sup>. Le couple Cs124/Tb a depuis été introduit sur divers molecular beacons<sup>[88]</sup>, et aptamères<sup>[89]</sup>, ainsi que dans un peptide Lanthanide-Binding-Tag, court peptide liant les lanthanides et pouvant servir de rapporteur luminescent pour une protéine, auquel il a pu être rattaché par ligation native<sup>[90]</sup>. Dans ce dernier cas, le Cs124 a été introduit lors de la SPPS via un acide aminé modifié (Figure 44) préalablement synthétisé en 3 étapes.

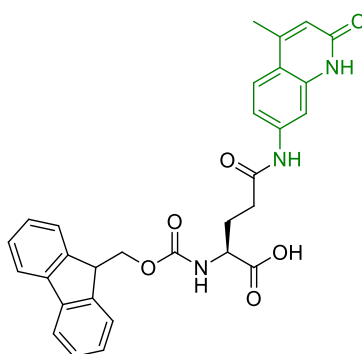


Figure 44: Acide aminé modifié utilisé par B.Imperiali<sup>[93]</sup>, permettant l'introduction du Cs124 lors d'une synthèse peptidique

Le couple Cs124/Tb<sup>3+</sup> a récemment été utilisé au laboratoire dans l'élaboration de sondes responsives au zinc<sup>[91]</sup>, ou à une séquence riche en bases AU d'ARN messenger<sup>[92]</sup>. En jouant sur la distance entre le Cs124 et un complexe de DOTA[Tb] introduits sur deux acides aminés différents d'un peptide, une modulation du transfert d'énergie a été rendue possible lorsque le peptide change de conformation en interagissant avec sa cible. Ces sondes s'allument en présence de leurs cibles mais présentent des rendements quantiques faibles (< 2%), du fait d'une distance trop importante entre le lanthanide et son antenne, contrairement aux complexes décrits par R. Selvin où le Cs124 est coordonné au lanthanide. L'ensemble de ces travaux montre la compatibilité des chromophores Cs124 et DOTA[Tb] avec la chimie de synthèse des peptides (assemblage du peptide par SPPS, ligation native<sup>[90],[92]</sup>), en faisant des candidats de choix pour leur utilisation sur la sonde de première génération.

## 2) Introduction des chromophores sur un peptide

### A DOTA[Tb]

Le complexe DOTA[Tb] est introduit sur une chaîne latérale d'un peptide via la conjugaison à une amine du pro-ligand DO3A-tris(*t*Bu)ester, DOTA dont trois des quatre fonctions acides ont été protégées en ester, commercial et peu coûteux (Figure 45). La conjugaison est facilement réalisable par réaction de l'acide carboxylique libre du ligand avec une fonction amine libre du peptide, conduisant à la formation d'un lien amide (Figure 46A). Une fois déprotégé au TFA, le ligand peut être métallé en phase aqueuse avec du terbium (Figure 46B).

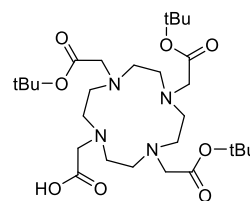


Figure 45: pro-ligand DO3A-tris(*t*Bu)ester

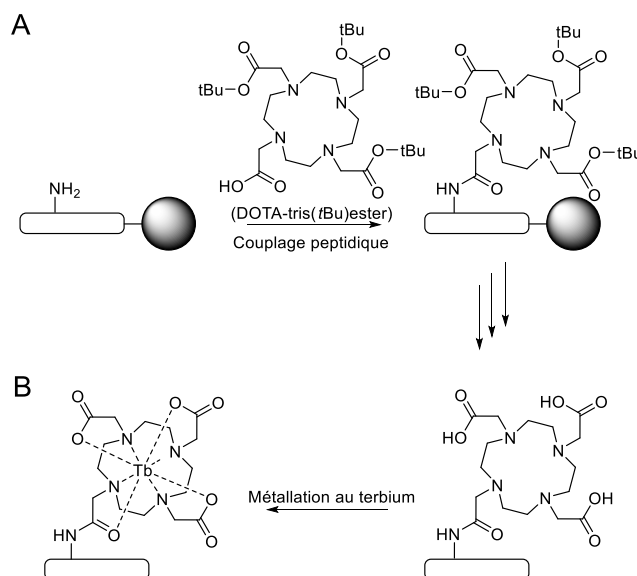


Figure 46 : A) Couplage du ligand DO3A-tris(*t*Bu)ester. B) Métallation par du terbium du ligand DOTA déprotégé

### B Carbostyryl-124

Le Cs124 est intégrable sur un peptide avec la méthode décrite par B. Imperiali<sup>[93]</sup> par la synthèse d'un acide aminé modifié (Figure 47). Le Cs124 est synthétisé en une étape par condensation de 1,3 diaminobenzène avec de l'éthylacétoacétate, puis l'amine du chromophore obtenu est couplée sur la chaîne latérale de l'acide glutamique pour obtenir l'acide aminé modifié Fmoc-Glu(Cs124)-OtBu, dont le groupement acido labile est éliminé dans le TFA. Il peut ensuite être utilisé comme un acide aminé classique lors de la synthèse peptidique.

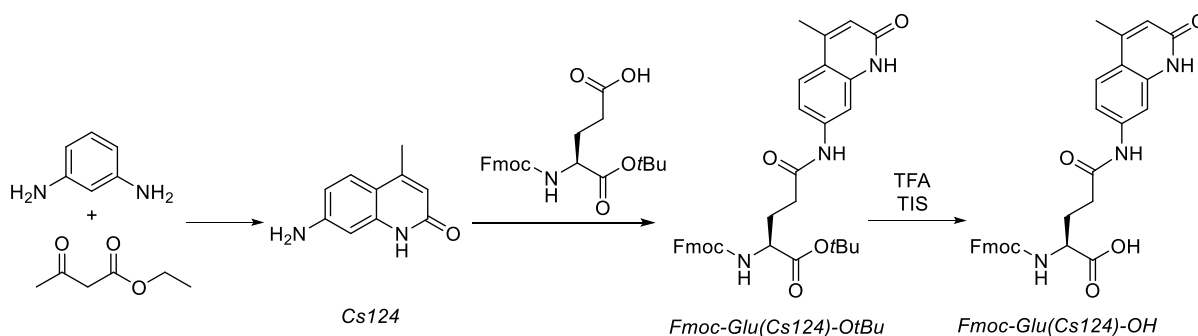


Figure 47: Synthèse de l'acide aminé modifié Fmoc-Glu(Cs124)-OH

### 3) Positions d'introduction des chromophores

Les zones globales proposées en partie « Sondes de première génération (FRET d'une antenne vers un lanthanide) » (p.34) pour l'introduction des chromophores laisse le choix sur plusieurs acides aminés pouvant être modifiés. Cependant une grande partie des acides aminés de la protéine possèdent un rôle essentiel au bon fonctionnement de la sonde et ne doivent pas être changés.

Sur la Figure 48 sont recensés les acides aminés dont les rôles sont :

- le repliement des doigts de zinc en un double feuillet  $\beta$  et d'une hélice  $\alpha$  sur le zinc (marqués par un triangle),
- l'interaction avec l'ADN (marqués par un carré),
- la formation du dimère (marqués par un point).

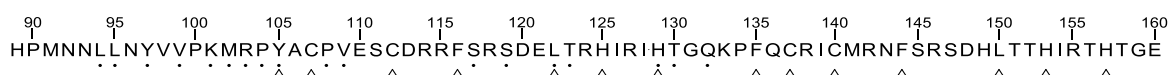


Figure 48: Recensement des acides aminés utiles pour la sonde

Sur la structure RX obtenue par B. Wang et C. Pabo, les 4 premiers acides aminés de la protéine ZIF (His<sup>89</sup> à Asn<sup>92</sup>) sont absents. La zone 1 proposée correspondant à l'extrémité de l'extension N-terminale susceptible de porter le complexe de lanthanide, comprend 8 acides aminés (His<sup>89</sup> à Asn<sup>96</sup> ; Figure 49). Les 4 premiers acides aminés sont éliminés de la sélection (His<sup>89</sup> à Asn<sup>92</sup>), puisque leur position par rapport à l'autre monomère n'est pas réellement connue. Parmi les acides aminés restants, deux jouent un rôle dans le processus de dimérisation (Leu<sup>94</sup> et Leu<sup>95</sup>), et l'Asn<sup>96</sup> est plus intégré dans l'interface de dimérisation que l'Asn<sup>93</sup>. Il semble donc judicieux d'introduire le chromophore en position n°93. De plus, cette position pointe davantage vers le second monomère que la position n°96.

L'antenne est, elle, introduite en zone 2 correspondant à l'épingle du feuillet  $\beta$  du premier doigt de zinc comportant 4 positions possibles (Val<sup>109</sup> à Cys<sup>112</sup> ; Figure 49). Dans cette zone, la Cys<sup>112</sup> intervient pour la coordination au zinc et ne peut être modifiée, de même que la Val<sup>109</sup> qui intervient dans la dimérisation. Seuls les acides aminés Glu<sup>110</sup> et Ser<sup>111</sup> peuvent être modifiés. Il a été choisi d'introduire l'antenne en position 110 en remplaçant l'acide glutamique, plus proche de l'autre monomère que la position 111.

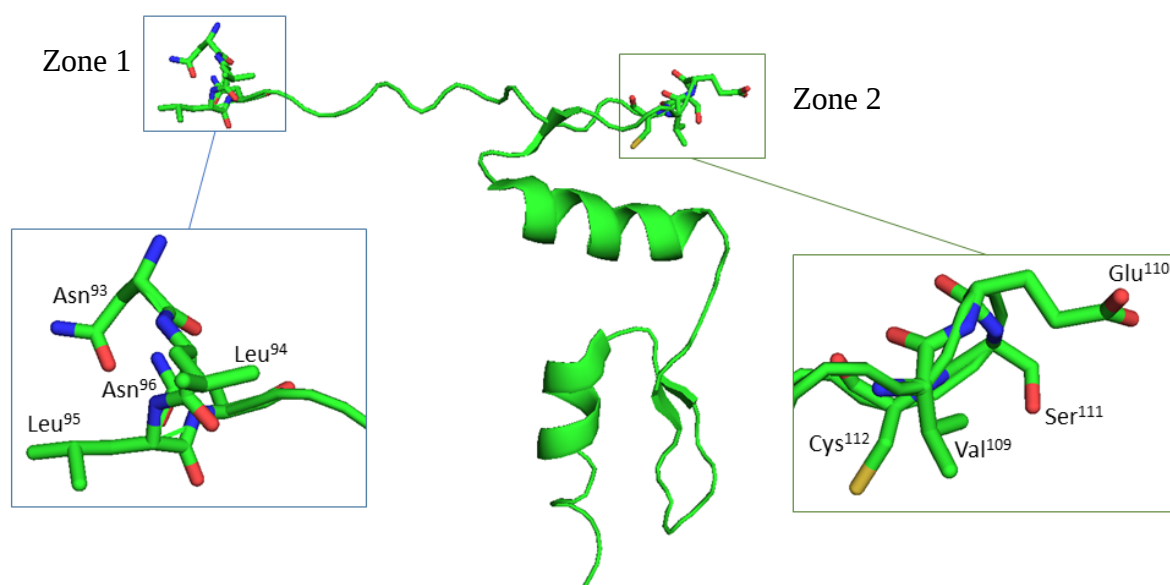


Figure 49: Zones d'introduction des chromophores (Zone 1 pour le DOTA[Tb], et Zone 2 pour le Cs124)

### III Synthèse de la protéine

La méthode la plus commune pour la synthèse de peptides par voie chimique est la synthèse peptidique supportée sur résine (SPPS) en stratégie Fmoc/tBu, détaillée en Figure 50. Cependant, cette stratégie est limitée à la synthèse de peptides n'excédant pas 30-40 acides aminés, de fortes chutes des rendements de couplages étant observées au-delà<sup>[94]</sup>.

La synthèse de la protéine ZIF, n'est donc pas envisageable en une seule fois par SPPS, car longue de 72 a.a.. Il faut par conséquent procéder par assemblages de fragments peptidiques en utilisant une méthode facilement applicable et conduisant à de bons rendements d'assemblage. Pour cela, la réaction de la ligation chimique native (NCL) est une méthode de choix.

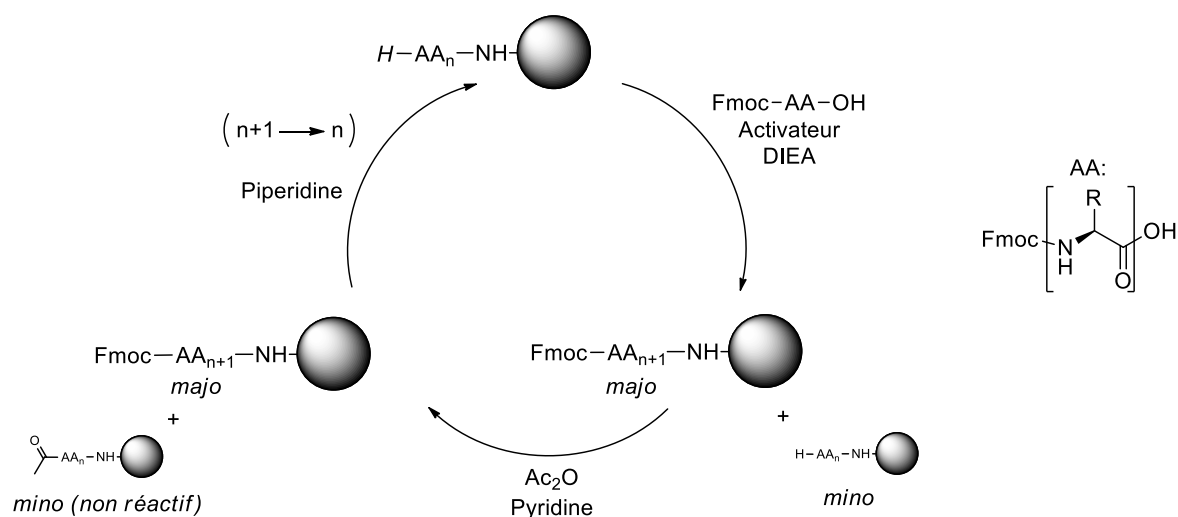


Figure 50: Synthèse classique sur SPPS en stratégie Fmoc/tBu. En première étape, un acide aminé noté Fmoc-AA-OH est couplé sur une fonction chimique supportée sur une résine (dans l'exemple ci-dessus, une résine RinkAmide comportant une fonction amine libre). En étape 2, les fonctions non couplées sont acétylées, empêchant la synthèse conduisant à des séquences tronquées. Ensuite, en troisième étape le groupement protecteur Fmoc est éliminé de l'acide aminé couplé afin de libérer la fonction amine de ce dernier. Le couplage d'un autre acide aminé peut ensuite être réalisé en répétant les mêmes étapes.

## 1) Assemblage par Ligation Chimique Native (NCL)

### A Présentation

La ligation chimique native (NCL) est une réaction développée par Stephen Kent<sup>[95],[96]</sup> permettant d'assembler deux peptides conduisant à la formation d'un lien chimique natif sous forme d'une cystéine. Cette technique couramment utilisée pour la synthèse de longues séquences peptidiques<sup>[97]</sup> peut être réalisée en phase aqueuse<sup>[98]</sup>, ou alors sur résine<sup>[99]</sup>. La méthode la plus couramment utilisée est en phase aqueuse.

### B Formation du lien amide natif

La formation du lien amide natif résulte de la réaction entre un peptide comportant une fonction thioester en C-terminal, et un autre comportant une cystéine en N-terminal, appelé « peptide cystéinyl » (Figure 51). En milieu neutre (pH 6,8-7), le thiol du peptide cystéinyl réagit sur le thioester de l'autre peptide par une réaction de transthioesterification. Cette réaction réversible est l'étape cinétiquement déterminante de la ligation. Le transthioester ainsi formé subit par la suite un réarrangement rapide via la formation d'un cycle à 5 chaînons appelé transfert d'acyl S-N conduisant à la formation du lien natif<sup>[100]</sup>.

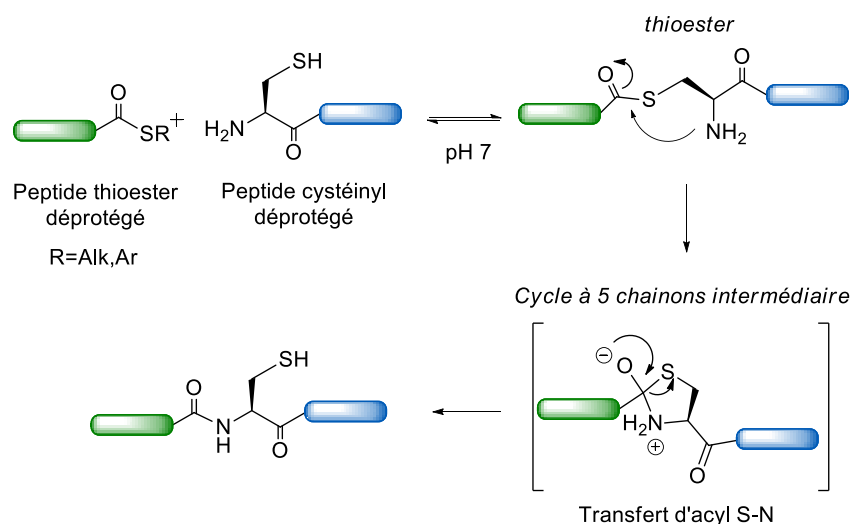


Figure 51: Principe de la réaction de ligation chimique native

### C Cinétique

L'électronégativité du thioester régit la cinétique de la réaction de transthioesterification, et par conséquent de la réaction de ligation. Les alkyles thioesters, ayant un pKa élevé possèdent une cinétique lente pour l'étape de transthioesterification avec la fonction thiol de la cystéine N-terminale du peptide cystéinyl, tandis que les aryles thioesters, aux pKa plus faibles possèdent une cinétique plus rapide<sup>[100]</sup>.

Cependant, l'introduction de thioester sur un peptide n'est pas une tâche aisée et la plupart des méthodes développées à ce jour ne permettent l'introduction que d'alkyl thioester. Pour obtenir des thioesters suffisamment réactifs, une première étape de transthioesterification du peptide avec un thiol aromatique est alors nécessaire, permettant de passer d'un alkyl thioester à un aryl thioester. Ainsi, en conditions de ligation, le peptide comportant le thioester est introduit en présence de larges excès de thiols aromatiques, permettant la formation de l'espèce active<sup>[101]</sup> (Figure 52). L'introduction par la suite du peptide cystéinyl permet une seconde étape de transthioesterification conduisant après un transfert d'acyl N vers S à la formation du lien natif.

Le thiol aromatique le plus couramment utilisé est l'acide 4-mercaptophénylacétique (MPAA), possédant un pKa de 6,6, et soluble dans l'eau à pH neutre<sup>[100]</sup>.

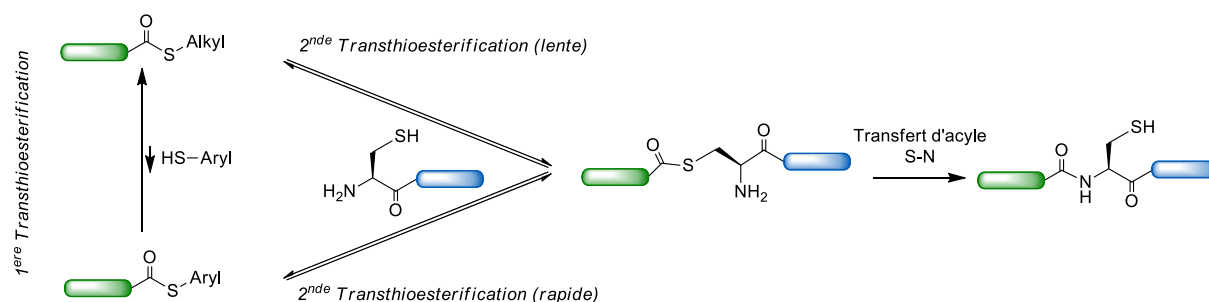


Figure 52: Cinétique de l'étape de transthioesterification entre le peptide N-terminal et le peptide C-terminal

### D Introduction d'un thioester en C-terminal

L'introduction d'un thioester en C-terminal d'un peptide est facilement réalisable par SPPS en stratégie Boc/*t*Bu. Pour cela, la fonction thioester est introduite sur la résine en début de synthèse, avec la possibilité de réaliser l'élongation du peptide à la suite de cette même fonction<sup>[102],[103]</sup>. Toutefois, la dangerosité de HF, réactif utilisé pour le décrochage des peptides synthétisés en stratégie Boc/*t*Bu a conduit ces dernières décennies à l'utilisation de la stratégie Fmoc/*t*Bu en SPPS de façon majoritaire. Cependant, il a été observé en utilisant cette stratégie que les fonctions thioesters introduites sur la résine sont aminolysées lors des traitements répétés à la pipéridine pour la déprotection des groupements Fmoc. Différentes techniques ont été développées afin d'éviter toute source de dégradation, comme l'utilisation de conditions plus douces pour la déprotection du Fmoc<sup>[104]</sup>, l'introduction de la fonction thioester à la fin de la synthèse sur résine<sup>[105]</sup>; ou encore l'utilisation de thioester masqués appelés crypto-thioester<sup>[106],[107]</sup>. Cette dernière est une méthode de choix car permet une certaine facilité d'introduction de la fonction thioester sur un peptide lors de la synthèse.

### 2) Crypto-thioesters

La technique d'introduction de thioesters en C-terminal via l'utilisation de crypto-thioesters s'est développée ces deux dernières décennies et est de plus en plus utilisée pour la synthèse de peptides. Le principe des crypto-thioesters reste toujours le même, une fonction thiol permettant le transfert d'acyl N-vers-S est bloquée par un groupement protecteur (GP) lors de toutes les étapes de la synthèse peptidique. Une fois le peptide décroché de sa résine et purifié, le groupement protecteur de la fonction thiol du crypto-thioester est éliminé, permettant le transfert d'acyl conduisant à la formation d'un thioester en position C-terminal (Figure 53). Quelques exemples de crypto-thioesters sont présentés en Figure 54.

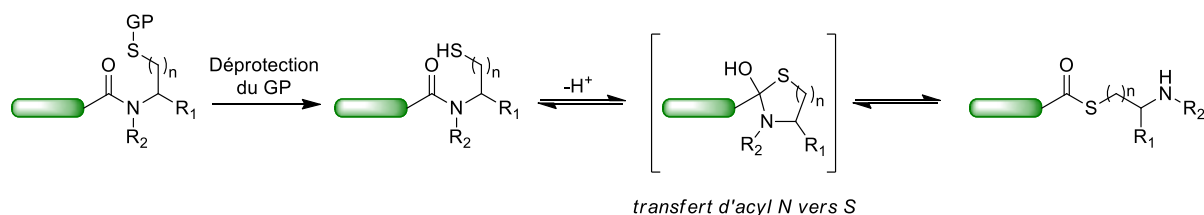


Figure 53: Le crypto-thioester (à gauche) et activé une fois le groupement protecteur du thiol éliminé. Un transfert d'acyl N vers S a ensuite lieu, conduisant à la formation d'un thioester en C-terminal.

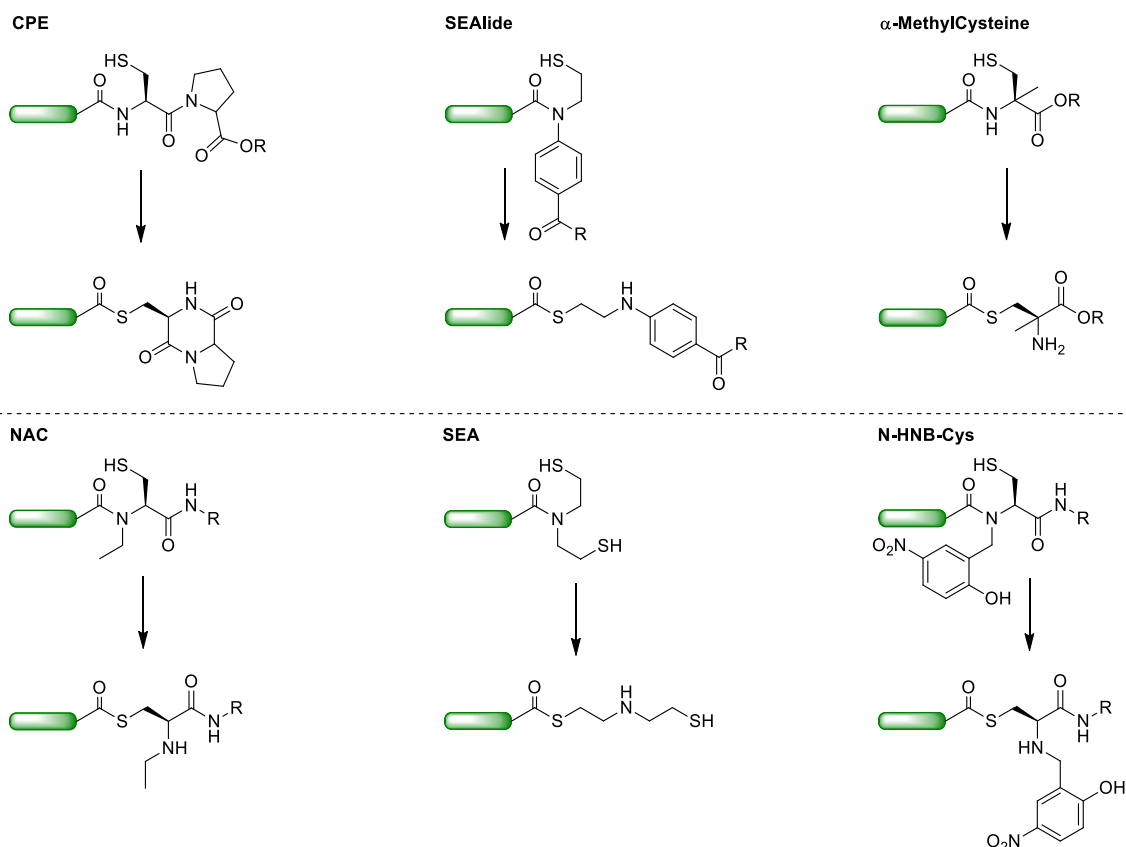


Figure 54: Exemples de crypto-thioesters et leurs thioesters respectifs (CPE<sup>[108]</sup>, SEALide<sup>[109]</sup>, α-MethylCysteine<sup>[110]</sup>, NAC<sup>[111]</sup>, SEA<sup>[112]</sup>, N-HNB-Cys<sup>[113]</sup>)

Le crypto-thioester utilisé au laboratoire est le bis(2-sulfanylethyl)amido (SEA) développé par O. Melnyk<sup>[114]</sup>, commercial sous la forme SEA-2Cl-trt-PS (noté résine SEA). Il est composé d'une amine secondaire sur laquelle peut être couplé le premier acide aminé de la séquence peptidique (Figure 55; noté SEA resin). Une fois l'élongation du peptide achevée, le peptide est décroché de sa résine en milieu acide (TFA), déprotégeant ainsi les fonctions thiols des groupements trityls acido-labiles. Ce dithiol obtenu est en équilibre avec une forme thioester linéaire, et ces deux formes réactives sont appelées SEA<sup>on</sup>. Pour la purification et le stockage du peptide, la forme SEA<sup>on</sup> doit être oxydée en SEA<sup>off</sup>, forme cyclisée du SEA comportant un pont disulfure. Il est possible de tirer avantage des deux formes SEA<sup>off</sup>/SEA<sup>on</sup>, permettant l'assemblage de trois peptides<sup>[115]</sup>.

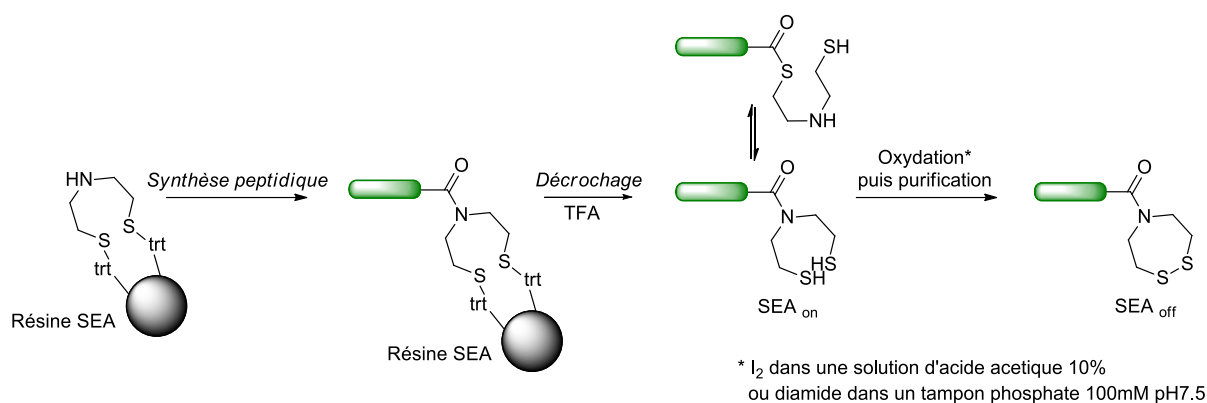


Figure 55: Synthèse, déprotection, et purification d'un peptide sur résine SEA.

Pour prévenir l'oxydation des fonctions thiols présentes sur le peptide lors du passage de la forme SEA<sup>on</sup> à SEA<sup>off</sup>, les cystéines du peptide sont introduites protégées par un groupement non-acidolabile, le plus couramment par des groupements protecteurs StBu. Ce groupement est stable aux conditions de déprotection acide, et peut être éliminé facilement en conditions réductrices.

En conditions de ligation, le pont disulfure du SEA<sup>off</sup> doit être réduit par du TCEP ou du DTT pour former l'espèce ouverte en équilibre avec la forme alkyl thioester. Cette forme n'étant pas suffisamment acide pour apporter une cinétique rapide lors de la ligation, une étape de transthioesterification par un thiol acide est nécessaire<sup>[101]</sup> (Figure 56).

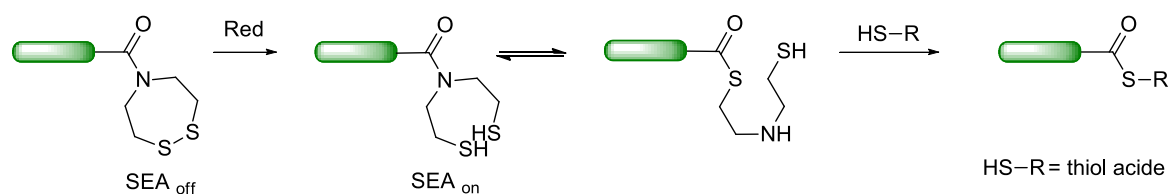


Figure 56: Passage de la forme SEA<sup>off</sup> à un aryl thioester

## IV Stratégies de synthèse des sondes

La protéine à synthétiser (Figure 57) étant composée de doigts de zinc est riche en cystéines (5,5% de cystéine dans la protéine, contre 1,8% en moyenne<sup>[116]</sup>). Plusieurs découpages au niveau de différentes cystéines sont donc possibles pour l'assemblage par ligation native.



Figure 57: Séquence totale de la protéine à synthétiser

## 1) Assemblage en deux fragments

La première stratégie est une stratégie d'assemblage simple en deux fragments. Pour un souci de rendement évident, la séquence du peptide à synthétiser ne doit pas excéder les 40 acides aminés. Or, les découpages au niveau des différentes cystéines de la séquence requièrent la synthèse d'un fragment d'environ 15-20 a.a. et un autre d'une cinquantaine d'a.a. Les découpages représentés sur la Figure 58 (stratégie A et B), nous semblaient les plus intéressants pour obtenir le fragment de 15-20 acides aminés le plus long possible et l'autre fragment d'une cinquantaine d'acides aminés le plus court possible. Par ailleurs, les ligations à des jonctions Glu-Cys possèdent une cinétique plus rapide qu'à des jonctions Ile-Cys<sup>[117]</sup>.



Figure 58: séquence de la protéine. En rouge les 4 cystéines, et en bleu et vert les deux découpages possibles pour l'assemblage en deux fragments.

La stratégie A a été testée, avec la synthèse du fragment de 49 a.a. sur résine Rink Amide PEG par robot de synthèse. Ce fragment comprend en début de séquence le doigt de zinc n°2, avec à sa suite une majeure partie du doigt de zinc n°1. Le suivi de la réaction par spectroscopie UV, permettant de contrôler le couplage de chaque acide aminé en dosant la quantité de Fmoc déprotégé, montre une chute progressive de la charge (Figure 59). Une perte de plus 50% de la charge initiale est observée à partir du 26<sup>ème</sup> acide aminé, premier de la séquence du doigt de zinc n°1. Ceci rend impossible l'obtention du fragment long de 49 a.a avec un rendement acceptable. D'autres protéines synthétisées au laboratoire ont déjà montré cette difficulté de synthèse sur l'enchaînement de plusieurs doigts de zinc.

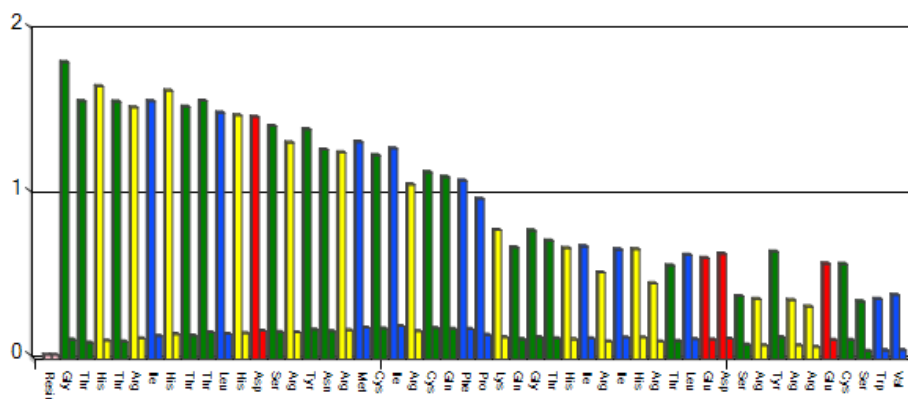


Figure 59: Suivit UV de la synthèse du fragment de 49a.a.

La stratégie B n'a pas été envisagée car le peptide de 48 a.a. devait dans ce cas être synthétisé sur résine SEA-PS, dont la matrice polystyrène est encore moins propice à la synthèse de longs peptides qu'une matrice PEG. L'utilisation d'une résine SEA-PEG<sup>[118]</sup> non-commerciale n'a pas été envisagée.

Un assemblage classique par ligation de deux peptides ne permet donc pas l'obtention des sondes.

## 2) Assemblage en trois fragments

L'assemblage de plusieurs peptides par ligations natives est en plein essor car permettant la synthèse de protéines de plus en plus longues<sup>[119]</sup>. Différentes techniques ont été développées à ce jour<sup>[120]</sup>.

O. Melnyk a décrit en 2012 une méthode d'assemblage de trois fragments peptidiques jouant sur les formes SEA<sup>off</sup>/SEA<sup>on</sup>, permettant d'enchaîner deux réactions de NCL en one-pot<sup>[121]</sup> (Figure 60). La première ligation met en jeu un peptide comportant une fonction thioester de MPA (peptide F1), et un second comportant une fonction SEA<sup>off</sup> en C-terminal ainsi qu'une cystéine(StBu) en N-terminal (peptide F2). Le pont disulfure de la cystéine est réduit sélectivement par le MPAA (servant aussi d'activateur pour la formation d'un aryl thioester), permettant la première ligation entre le peptide F1 et le peptide F2. L'assemblage F1-F2 réalisé, un réducteur plus fort est ensuite introduit afin de passer de la forme SEA<sup>off</sup> à la forme SEA<sup>on</sup>. L'ajout d'un troisième peptide comportant une cystéine en N-terminal (peptide F3) permet finalement une seconde réaction de ligation, conduisant à l'assemblage F1-F2-F3.

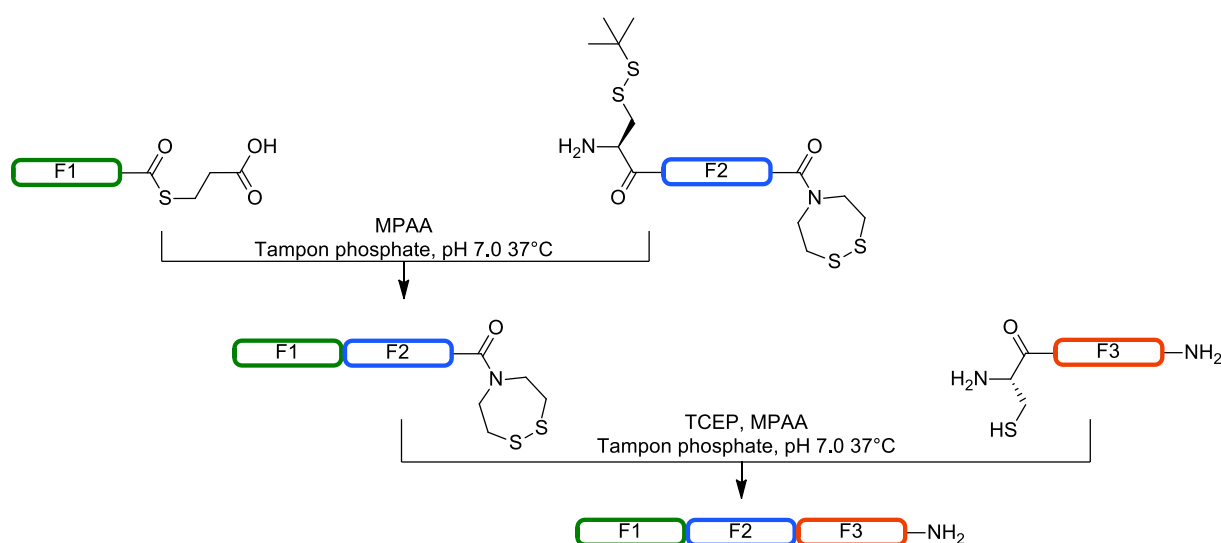


Figure 60: Principe de l'assemblage en trois fragments

Dans notre cas, quatre découpages en trois fragments sont possibles avec la séquence de la protéine ciblée (Figure 61). Les découpages 1 et 2 comportent une jonction Ser-Cys, qu'il est préférable d'éviter à cause d'une épimérisation parfois importante de la sérine lors des réactions de ligations<sup>[122]</sup>. Entre les découpages 3 et 4, la différence réside dans la longueur des peptides F2 et F3. Le second peptide étant synthétisé sur résine SEA-PS et le troisième sur résine RinkAmide PEG, il est préférable que le fragment le plus long soit obtenu par la synthèse sur résine PEG. La stratégie 4 semble la plus appropriée pour la synthèse des fragments peptidiques ainsi que leurs assemblages.

Découpage 1

HPMNNLLNYVVPKMRPYA CPVES CDRRFSRSDDELTRHIRIHTGQKPFQ CRICMRNFSRSDHLTTHIRHTHTGE —NH<sub>2</sub>

Découpage 2

HPMNNLLNYVVPKMRPYA CPVES CDRRFSRSDDELTRHIRIHTGQKPFQ CRICMRNFSRSDHLTTHIRHTHTGE —NH<sub>2</sub>

Découpage 3

HPMNNLLNYVVPKMRPYA CPVES CDRRFSRSDDELTRHIRIHTGQKPFQ CRICMRNFSRSDHLTTHIRHTHTGE —NH<sub>2</sub>

Découpage 4

HPMNNLLNYVVPKMRPYA CPVES CDRRFSRSDDELTRHIRIHTGQKPFQ CRICMRNFSRSDHLTTHIRHTHTGE —NH<sub>2</sub>

Figure 61: Différents découpages possibles pour un assemblage en trois fragments

Avant mon arrivée au laboratoire, Laurent Raibault a utilisé cette stratégie pour obtenir la protéine appelée ZIF(nue) (HPMNNLLNYVVPKMRPYACPVESCDRRFSRSDDELTRHIRIHTGQKPFQCRICMRNFSRSDHLTTHIRHTHTGE-NH<sub>2</sub>) correspondant à la séquence de la protéine ZIF12 décrite par B. Wang et C. Pabo (Figure 62), validant ainsi la stratégie d'assemblage. La protéine non décorée obtenue a été caractérisée par HPLC et ESI/MS (Figure 63), montrant ainsi la faisabilité de cette synthèse.

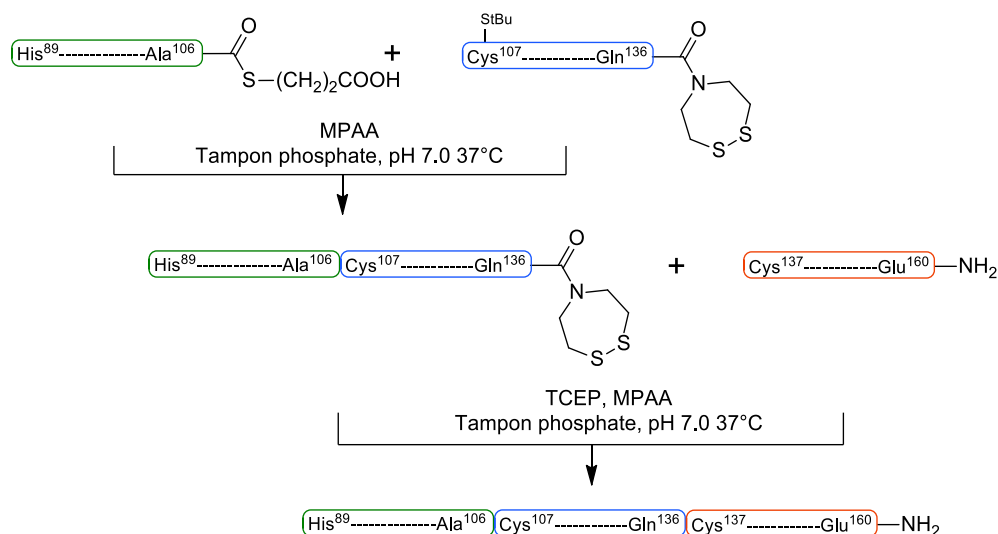


Figure 62: Assemblage de la protéine ZIF(nue)

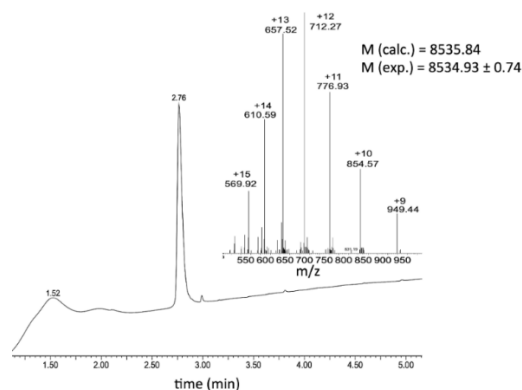


Figure 63: Chromatogramme UPLC et Spectre ESI-MS de la protéine ZIF(nue)

## V Synthèse de la sonde ZIF(Cs124|DOTA[Tb])

Pour la synthèse de la sonde ZIF(Cs124|DOTA[Tb]) (Figure 64 A), nous avons choisi d'assembler les trois fragments F1, F2, F3 suivants (Figure 64 B):

- Un peptide comportant une lysine en position n°93 liée à un DOTA[Tb], et un ester de MPA en C-terminal, noté Ac-F1[Lys(DOTA[Tb])]-MPA
- Un peptide comportant une Cys(StBu) en N-terminal, l'acide aminé Glu(Cs124) en position n°110, et une fonction SEA<sup>off</sup> en C-terminal, noté F2[Cys(StBu)|Glu(Cs124)]-SEA<sup>off</sup>
- Un peptide comportant une cystéine déprotégée en N-terminal, noté F3-NH<sub>2</sub>

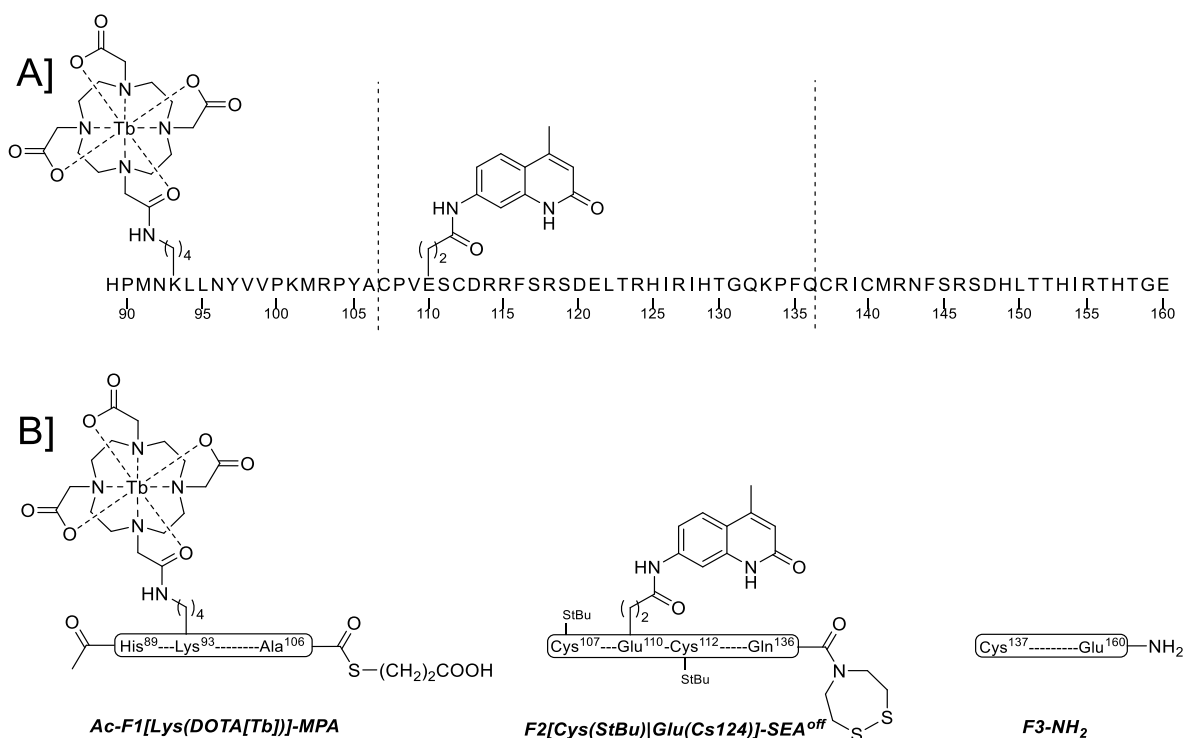


Figure 64: A) Découpage de la protéine ; B) Peptides à synthétiser

### 1) Fragment Ac-F1[Lys(DOTA[Tb])]-MPA

Le peptide Ac-F1[Lys(alloc)]-SEA-2Cl-trt-PS a été synthétisé manuellement en stratégie Fmoc/tBu sur résine SEA-2Cl-trt-PS (Figure 66). Une lysine alloc a été introduite en position n°90 du peptide, puis une fois l'élongation du peptide terminée, la position N-terminale a été acétylée. La fonction alloc de la chaîne latérale de la Lys<sup>90</sup> a ensuite été déprotégée sélectivement en présence de Pd(PPh<sub>3</sub>)<sub>4</sub> et de phénylsilane dans du DCM anhydre et dégazé. Le ligand DO3A-tris(tBu)ester a été greffé par un couplage peptidique classique (PyBOP, DIEA, DMF) durant une nuit. Le peptide a ensuite été décroché de sa résine et ses chaînes latérales déprotégées dans le TFA en présence de pièges (eau, phénylsilane, triisopropylsilane). La fonction SEA<sup>on</sup> a été convertie en SEA<sup>off</sup> par oxydation au diamide dans du tampon phosphate pH 7,5. Le peptide obtenu a été purifié par HPLC préparative, puis le ligand DO3A a été métallé par un excès de TbCl<sub>3</sub> en phase aqueuse à pH 6 suivi d'une purification HPLC afin d'éliminer l'excès de terbium. Finalement la fonction SEA<sup>off</sup> a été réduite par du TCEP en présence d'excès de MPA à pH 4, conduisant à la formation du peptide Ac-F1[Lys(DOTA[Tb])]-MPA, qui a été purifié par HPLC préparative (11,3 % de rendement), lyophilisé et caractérisé par HPLC et ESI-MS (Figure 65).

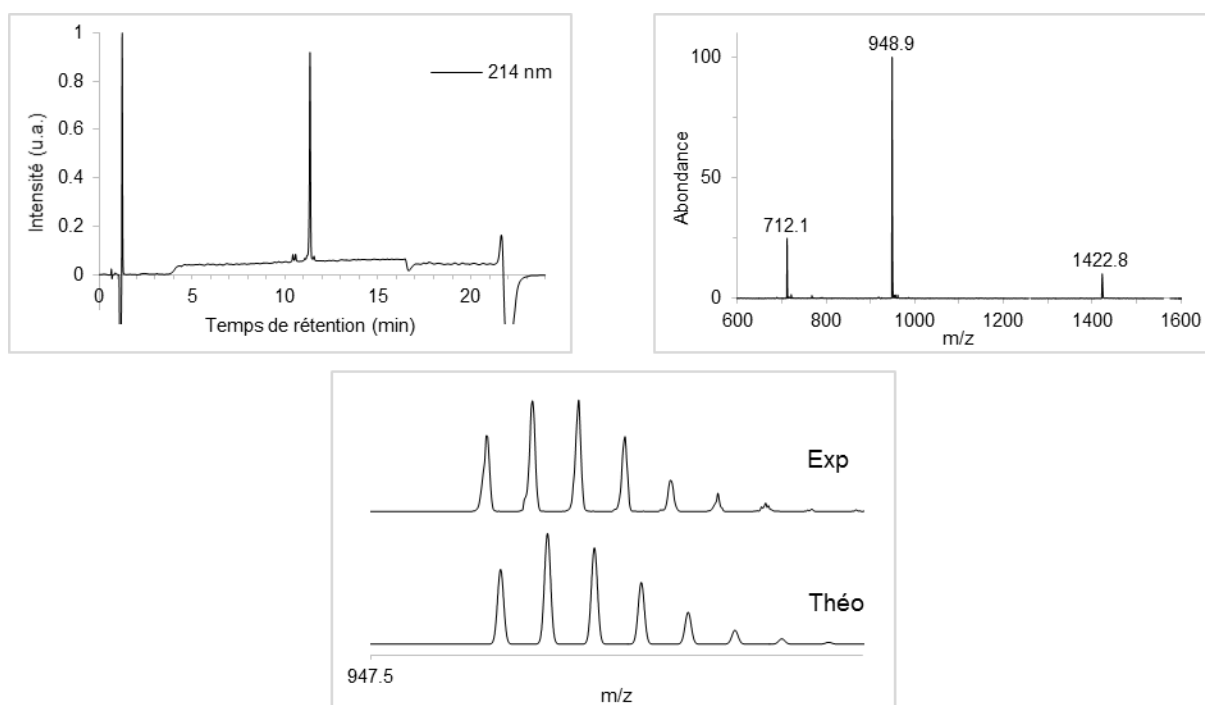


Figure 65: A gauche, chromatogramme HPLC du peptide Ac-F1[Lys(DOTA[Tb])]-MPA; A droite spectre ESI-MS ; Au centre, massif isotopique expérimental et théorique (Masse calculée: 1423,06 [M+2H]<sup>2+</sup>, 949,04 [M+3H]<sup>3+</sup>, 712,03 [M+4H]<sup>4+</sup> pour  $M = C_{121}H_{188}N_{31}O_{32}S_3Tb$ )

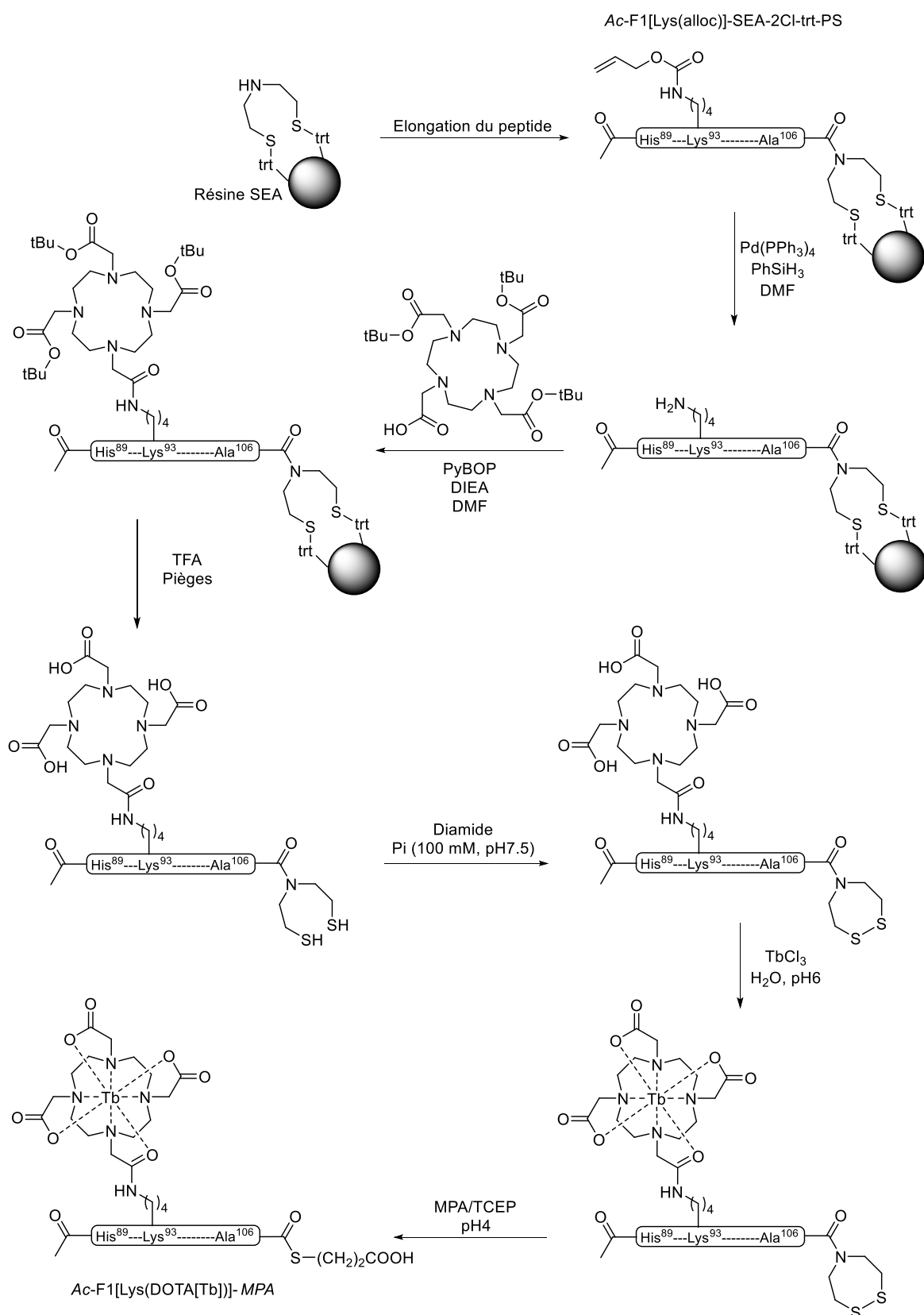


Figure 66: Schéma de synthèse du peptide Ac-F1[Lys(DOTA[Tb])] - MPA

## 2) Fragment F2[Cys(StBu)|Glu(Cs124)]-SEA<sup>off</sup>

Le peptide F2[Cys(StBu)|Glu(Cs124)]-SEA<sup>off</sup> comporte dans sa séquence l'acide aminé modifié Glu(Cs124), qui a été obtenu en suivant le protocole décrit par l'équipe de B. Imperiali<sup>[93]</sup> (Figure 67). Du 1,3 diaminobenzène et de l'éthylacétoacétate ont été chauffés à 150°C lors d'une réaction sans solvant. Une fois la coumarine purifiée par colonne de silice, elle a été couplée sous atmosphère inerte en boîte à gants sur l'acide glutamique Fmoc-Glu(OH)-tBu avec du HATU en présence de DIEA. L'acide aminé Fmoc-Glu(Cs124)-tBu obtenu a ensuite été purifié par chromatographie sur colonne de silice, puis le groupement tBu a été déprotégé dans le TFA en présence de TIS suivi d'une précipitation à l'Et<sub>2</sub>O. L'acide aminé modifié Fmoc-Glu(Cs124)-OH obtenu a été utilisé comme un acide aminé classique en stratégie Fmoc/tBu sans purification supplémentaire (rendement global de synthèse 8%).

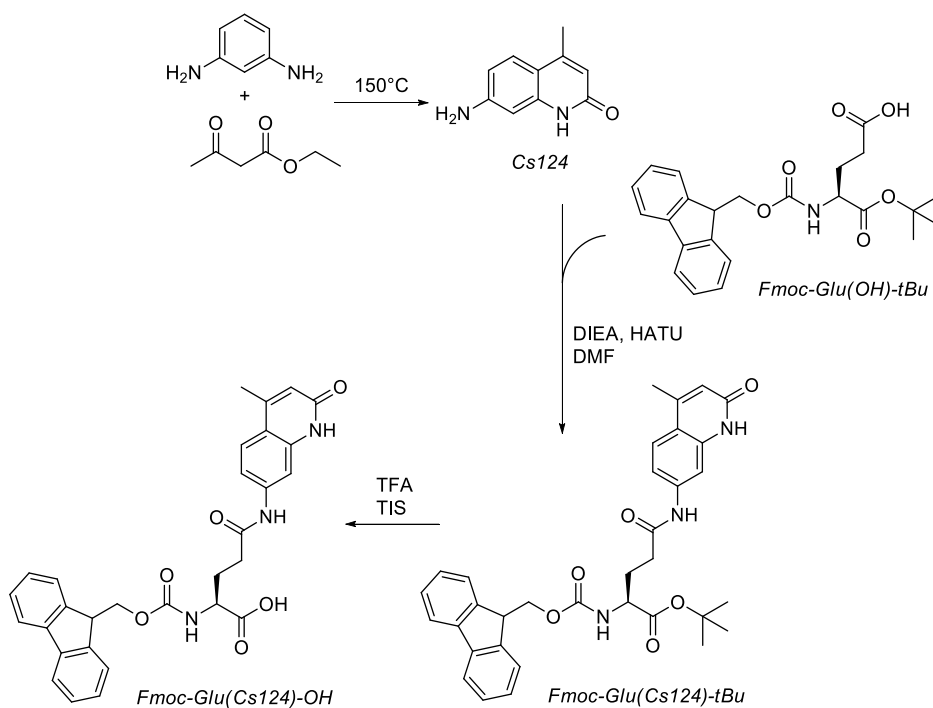


Figure 67: Schéma de synthèse de l'acide aminé modifié Fmoc-Glu(Cs124)-OH

Le peptide F2[Cys(StBu)|Glu(Cs124)]-SEA<sup>off</sup> a été synthétisé en stratégie Fmoc/tBu sur résine SEA-2Cl-trt-PS (Figure 68). Le couplage du premier acide aminé a été réalisé à la main afin de contrôler le taux de greffage, puis la suite de la synthèse a été effectuée en utilisant un robot de synthèse, avec couplage manuel de l'acide aminé modifié Fmoc-Glu(Cs124)-OH en position n°110. Le peptide a ensuite été décroché de sa résine et les chaînes latérales déprotégées dans le TFA en présence de pièges (eau, phénylsilane, triisopropylsilane), puis la fonction SEA<sup>on</sup> a été convertie en SEA<sup>off</sup> par oxydation au diode dans de l'acide acétique à 10%. L'analyse ESI-

MS du brut d'oxydation a montré la présence du peptide attendu, plus deux autres peptides dont l'analyse par spectrométrie de masse révèle une perte de masse de 18 et 121 Da. Un  $\Delta M$  de -18 par rapport à la masse attendue est bien connu, et correspond à la formation d'aspartimides sur le peptide. Le  $\Delta M$  de -121 quant à lui n'est pas renseigné dans la littérature et a dû être identifié.

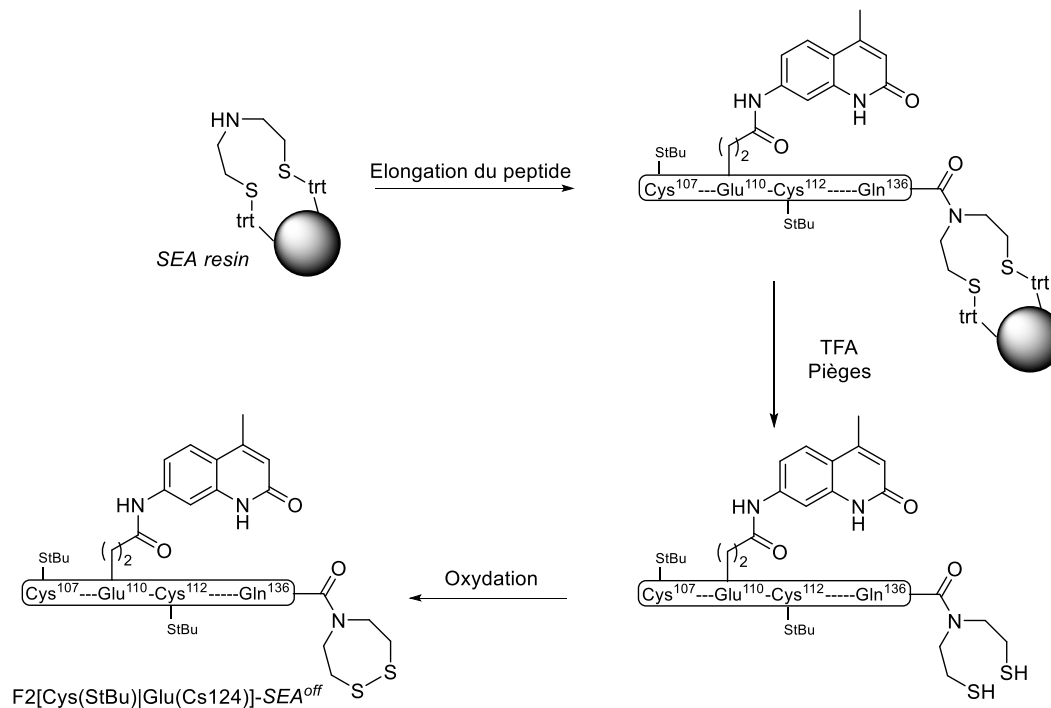


Figure 68: Schéma de synthèse du peptide  $F2[Cys(StBu)|Glu(Cs124)]-SEA^{off}$

### A Aspartimides

La formation d'aspartimides observée par la présence de M-18 sur le peptide  $F2[Cys(StBu)|Glu(Cs124)]-SEA^{off}$  est une des réactions parasites les plus courantes en synthèse peptidique par stratégie Fmoc/tBu<sup>[123],[124],[125]</sup>. Elle est due à l'exposition d'un peptide comportant un acide aspartique à des conditions basiques fortes et répétées, et donc dans notre cas aux traitements à la pipéridine<sup>[126]</sup>.

Le mécanisme de formation d'aspartimides met en jeu la chaîne latérale d'un acide aspartique et la fonction amide liant l'acide aspartique à l'acide aminé le précédent dans la séquence. En milieu basique, la fonction amide est déprotonnée et attaque le carbonyle de la chaîne latérale de l'acide aspartique, conduisant au départ du groupement OtBu (Figure 69). Cette réaction correspond à une déshydratation et donc à une perte de 18Da sur le peptide. Ensuite, lors des étapes de traitements basiques, la position de l'acide aspartique peut être racémisée, le proton en  $\alpha$  de l'acide aspartique étant rendu plus acide par la formation d'un imide.

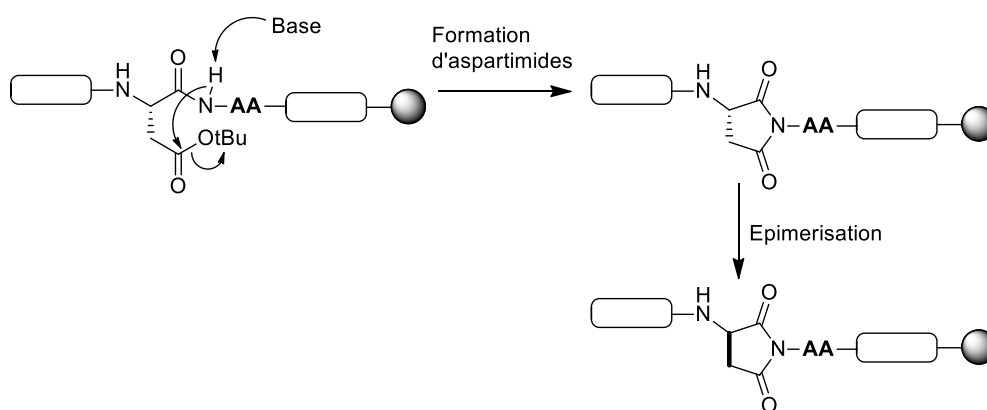


Figure 69: Mécanisme de formation des aspartimides

En conditions basiques, la forme imide racémisée peut être ré-ouverte par réaction d'hydrolyse sur les carbonyles (Voie A et Voie B Figure 70), conduisant à 4 produits possibles possédant la même masse que le produit attendu. La quantité de produit dégradé peut ainsi être sous-évaluée, car non quantifiable par spectrométrie de masse. Toute trace de formation d'aspartimides induit une incertitude sur la quantité de peptide attendu réellement présent.

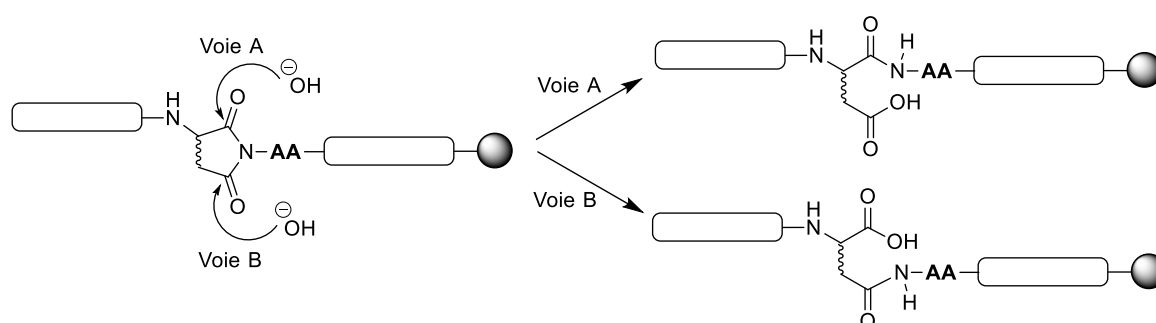


Figure 70: Voies de ré-ouvertures de l'aspartimide

La réaction de cyclisation dépend de l'acidité de l'amide, et par conséquent de la nature de l'acide aminé précédent l'acide aspartique dans la séquence. Une étude a permis de mettre en évidence les acides aminés susceptibles de conduire à la formation d'aspartimides<sup>[127]</sup>. Dans la séquence du fragment 2, deux acides aspartiques sont présents : l'Asp<sup>113</sup> susceptible de former des aspartimides car précédé par une arginine, et l'Asp<sup>121</sup> ne formant normalement pas d'aspartimides car précédé par un acide glutamique.

Il existe trois grands moyens d'empêcher la formation d'aspartimides : soit en protégeant l'amide de l'acide aspartique par un groupement protecteur Hmb<sup>[128]</sup>, soit en protégeant l'acide carboxylique de la chaîne latérale par un groupement encombré en utilisant l'acide aminé Fmoc-Asp(OBno)-OH<sup>[129]</sup>, soit en modifiant les conditions de déprotection de groupements Fmoc<sup>[130]</sup>. Le fait de tamponner la pipéridine permet d'éviter la réaction parasite, avec par

exemple de l'OxymaPur<sup>[131]</sup>, de l'acide formique<sup>[132]</sup>, ou alors en supprimant l'utilisation de pipéridine remplacée par de la pipérazine activée par microondes . Le peptide a été re-synthétisé en testant ces trois différentes conditions compilées dans le Tableau 1.

Tableau 1: Conditions de déprotection du groupement Fmoc lors de la synthèse du peptide F2[Cys(StBu)|Glu(Cs124)]-SEA<sup>off</sup>

| Base                         | additif                                   | -18 Da en ESI-MS |
|------------------------------|-------------------------------------------|------------------|
| Piperidine 20%               | Acide formique (50 mL/L) <sup>[132]</sup> | oui              |
| Piperidine 20%               | OxymaPur (14 g/L) <sup>[131]</sup>        | oui              |
| Piperazine (avec microondes) | /                                         | oui              |

La formation d'aspartimides étant toujours observée, l'acide aspartique en position n°113 a été remplacé par un acide glutamique. Les propriétés de la sonde ne devraient pas être affectées par cette modification, l'Asp<sup>113</sup> étant situé vers l'épingle  $\beta$  du doigt de zinc n°1, n'intervient pas dans le processus de dimérisation, ni dans l'interaction avec l'ADN, et ne joue aucun rôle dans la structuration du doigt de zinc.

La modification a été appliquée lors d'une nouvelle synthèse, où le peptide a été synthétisé sur deux lots. Un premier synthétisé avec des déprotections classiques à la pipéridine, et un second avec de la pipéridine tamponnée à l'acide formique. Sur les deux peptides obtenus, les analyses ESI-MS sont identiques et sans traces d'aspartimides. La réaction parasite avait donc bien lieu uniquement sur l'Asp<sup>113</sup>.

### B Perte d'une masse de -121

Outre la formation d'aspartimides révélée par la présence d'un  $\Delta M$  de -18 par rapport à la masse attendue, un autre produit correspondant à un  $\Delta M$  de -121 a été observé. Le seul découpage permettant d'expliquer une perte de masse de -121 par rapport au peptide attendu correspondrait à la perte de la fonction SStBu sur une cystéine. Deux cystéines protégées par un groupement StBu étant présentes dans la séquence et un seul delta -121 étant identifié en masse, il a fallu identifier la régiosélectivité de la réaction de dégradation.

Cette dégradation a par ailleurs été observée sur d'autres séquences, comme sur le peptide F2[Cys(StBu)|Lys(lissamine)]-SEA<sup>off</sup> (vu au chapitre II p.83) et le peptide SP3A dont la séquence est C(StBu)GKKFTKKEELQKHAKTHTG-NH<sub>2</sub>, utilisé pour un autre projet. Ce dernier comporte une unique cystéine en position N-terminale, laissant présager de manière analogue une perte localisée du groupement SStBu sur la cystéine N-terminale du peptide F2[Cys(StBu)|Glu(Cs124)]-SEA<sup>off</sup>. Ce dernier a été synthétisé une nouvelle fois, avec

caractérisation avant et après introduction de la cystéine finale. L'analyse ESI-MS du brut de déprotection du peptide sans cystéine montre l'absence de masse -121, tandis que celle du peptide avec cystéine en montre la présence, confirmant la localisation de la dégradation sur la cystéine N-terminale (Figure 71).

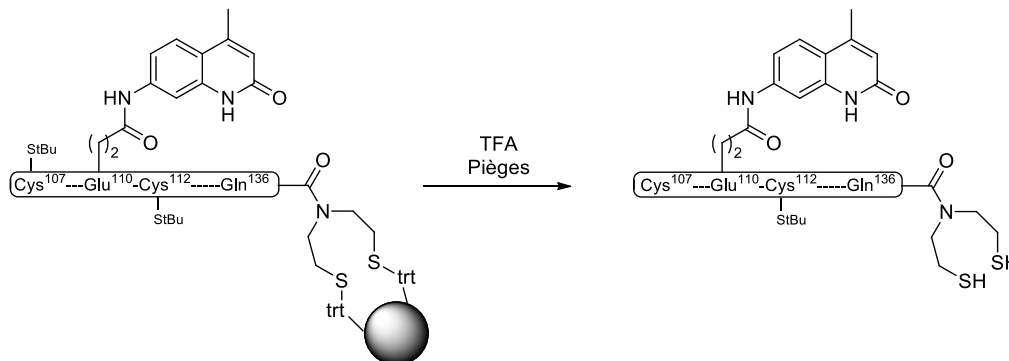


Figure 71: Dégradation régiosélective sur la cystéine N-terminale

Le mécanisme probable de cette réaction de dégradation, et une  $\beta$ -élimination de la cystéine, suivit d'une isomérisation conduisant à l'obtention d'une imine, qui après hydrolyse conduit à l'obtention d'une cétone (Figure 72).

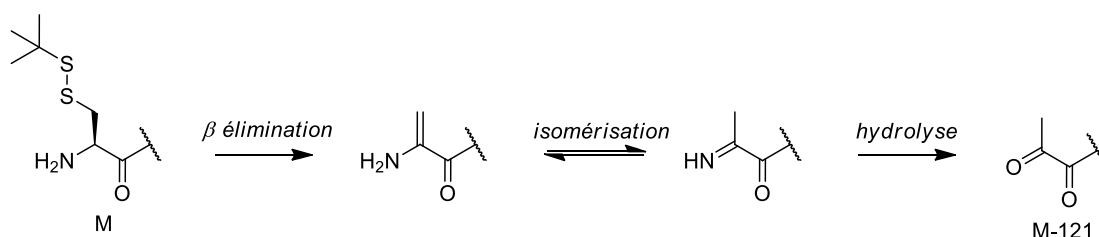


Figure 72: Dégradation de la Cys(StBu)

Pour éviter cette dégradation, il faut comprendre son origine. La cystéine sur laquelle a lieu la dégradation étant le dernier acide aminé à être introduit dans la séquence, aucun post traitement n'est réalisé sur la résine, et le peptide est directement décroché de son support dans le TFA. La dégradation ne peut donc se produire que lors de cette étape de déprotection.

Différents tests ont été réalisés sur le peptide SP3A, facile à synthétiser et peu coûteux car synthétisé sur résine RinkAmide. Un suivi de la réaction par HPLC complété par une analyse par ESI-MS a permis d'identifier, séparer et quantifier le peptide attendu et le peptide dégradé. Dans les conditions classiques de déprotection (2 h, TFA/TIS/Thioanisole/H<sub>2</sub>O ; 9,375 : 0,25 : 0,125 : 0,25), un ratio de 38% en peptide dégradé est mesuré. En conservant le même cocktail de pièges dans le TFA mais avec un temps de déprotection de 30 min, le ratio est descendu à 24%, mais avec une déprotection incomplète du peptide correspondant à 20 % de pertes. En changeant les différents lots de réactifs un à un, le ratio n'a pas non plus été affecté. Finalement,

en remplaçant le TIS par du DTT pendant une déprotection de 30 min, le ratio a été modifié à 15 %, avec 28 % de peptide partiellement déprotégé, ce qui reste loin d'être des conditions idéales.

Une autre stratégie de synthèse a été envisagée, en remplaçant la cystéine finale par une thiazolidine (Thz). Cette dernière peut être déprotégée, une fois le peptide décroché de sa résine et la fonction  $SEA^{on}$  convertie en  $SEA^{off}$ , par un traitement à la méthoxylamine à pH 4 (Figure 73).

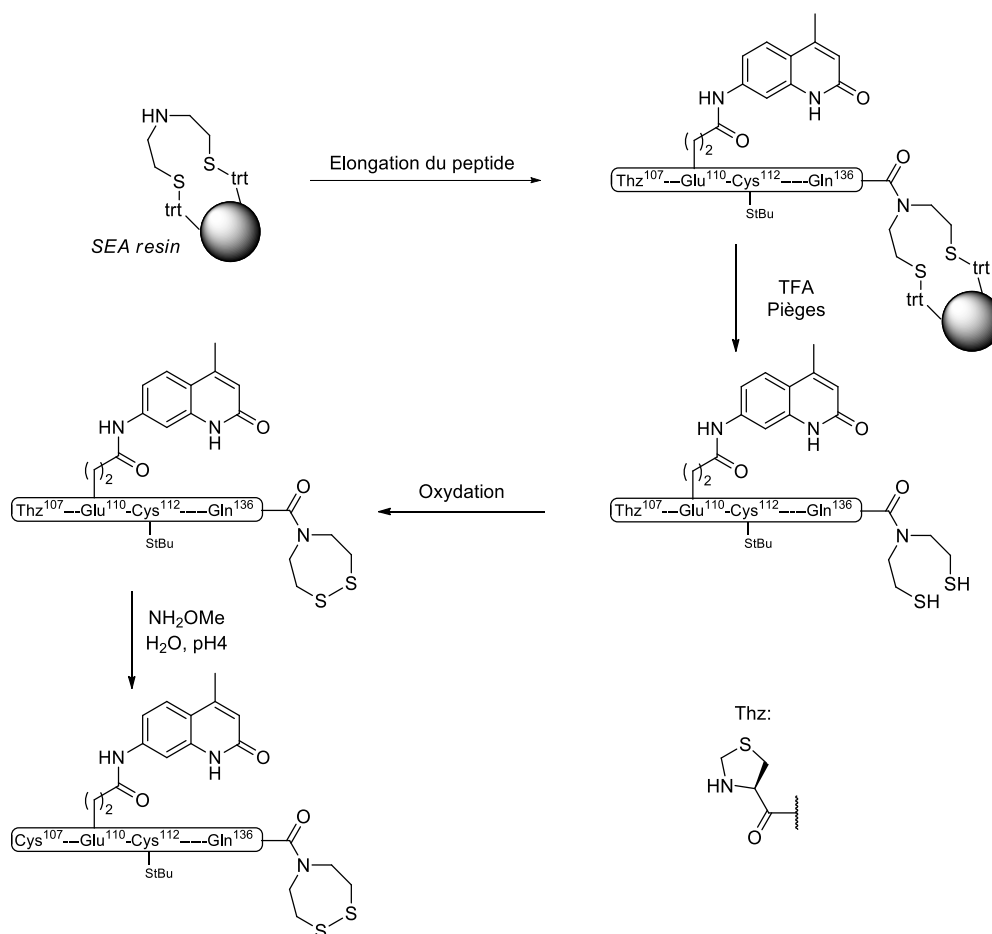


Figure 73: Stratégie de synthèse du peptide F2[Cys|Glu(Cs124)]- $SEA^{off}$

Ce peptide a été synthétisé selon la même approche que le peptide F2[Cys(StBu)|Glu(Cs124)]- $SEA^{off}$ . Cependant, il s'est avéré insoluble en milieu acide lors de l'étape d'oxydation. Il n'a donc pu être oxydé et purifié convenablement (3% de rendement sur l'étape d'oxydation). Cette alternative a été abandonnée.

Le temps de rétention par HPLC du peptide F2[Cys(StBu)|Glu(Cs124)]- $SEA^{off}$  étant légèrement inférieur à celui de son produit de dégradation, il a été choisi de finalement utiliser le groupement protecteur StBu, en faisant le choix de consacrer du temps à l'étape de purification

par HPLC. Le peptide F2[Cys(StBu)|Glu(Cs124)]-SEA<sup>off</sup>, obtenu avec un rendement de 18% a été caractérisé par HPLC et ESI-MS (Figure 74).

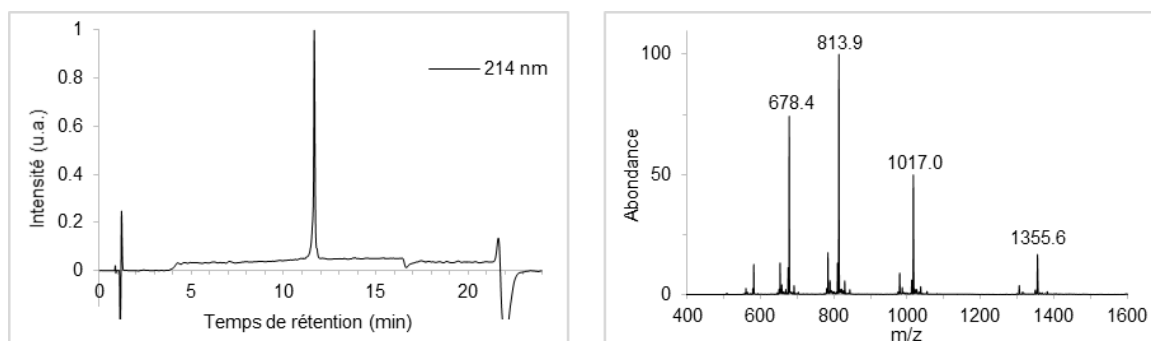


Figure 74: A gauche, chromatogramme HPLC du peptide F2[Cys(StBu)|Glu(Cs124)]-SEA<sup>off</sup>; A droite, spectre ESI-MS (Masse calculée : 1355,61 [M+3H]<sup>3+</sup>, 1016,96 [M+4H]<sup>4+</sup>, 813,77 [M+5H]<sup>5+</sup>, 678,31 [M+6H]<sup>6+</sup> pour M = C<sub>175</sub>H<sub>277</sub>N<sub>55</sub>O<sub>45</sub>S<sub>6</sub>)

### 3) Fragment F3-NH<sub>2</sub>

Le peptide F3-NH<sub>2</sub> a été synthétisé en stratégie Fmoc/tBu sur résine RinkAmide (Figure 75) par robot de synthèse, avec une déprotection finale suivie d'une étape de purification HPLC. L'analyse ESI-MS du peptide brut a mis en évidence une faible proportion de formation d'aspartimides.



Figure 75: Schéma de synthèse du peptide F3-NH<sub>2</sub>

L'unique acide aspartique de la séquence, précédé par une histidine ne correspond pas à un enchainement susceptible de former des aspartimides<sup>[134]</sup>. Le peptide a été re-synthétisé en utilisant de la pipéridine (20% dans le DMF) tamponnée à l'OxymaPur (14g/L)<sup>[131]</sup> pour les traitements de déprotection. Le peptide F3-NH<sub>2</sub> a été obtenu pur avec un rendement de 22,7 %, puis a été caractérisé par HPLC et ESI-MS (Figure 76).

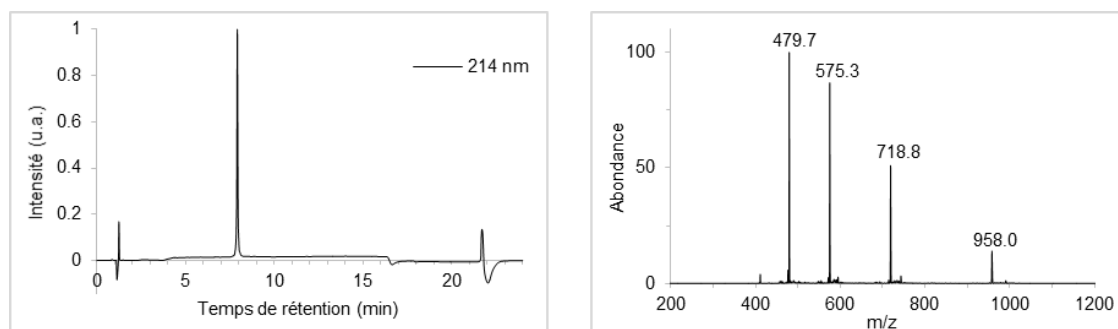


Figure 76: A gauche, chromatogramme HPLC du peptide F3-NH<sub>2</sub>; A droite, spectre ESI-MS (Masse calculée : 958,09 [M+3H]<sup>3+</sup>, 718,82 [M+4H]<sup>4+</sup>, 575,26 [M+5H]<sup>5+</sup>, 479,55 [M+6H]<sup>6+</sup> pour M = C<sub>117</sub>H<sub>192</sub>N<sub>44</sub>O<sub>35</sub>S<sub>3</sub>)

#### 4) Assemblage des différents fragments

Les trois peptides obtenus ont ensuite été assemblés en utilisant les conditions de l'assemblage one-pot décrit par O. Melnyk (Figure 77).

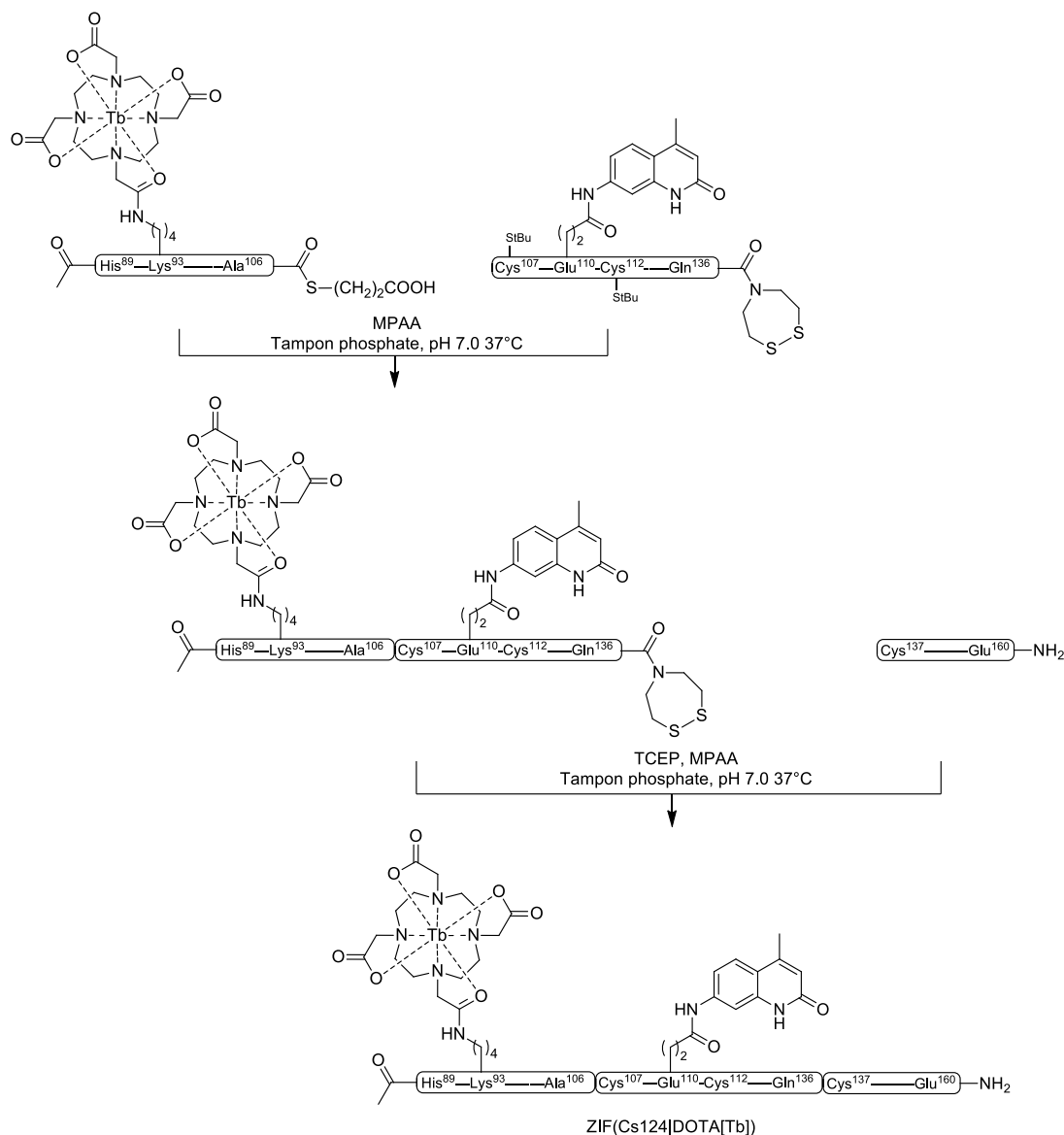


Figure 77: Schéma de la réaction d'assemblage one-pot de la protéine ZIF(Cs124|DOTA[Tb])

L'ensemble de la réaction a été réalisée en boîte à gants, afin d'éviter l'oxydation des cystéines. La première ligation entre le peptide  $\text{Ac-F1[Lys(DOTA[Tb])]-MPA}$  et le peptide  $\text{F2[Cys(StBu)|Glu(Cs124)]-SEA}^{\text{off}}$  a été effectuée à une concentration de 7 mM en peptide dans une solution de MPAA 100 mM dans un tampon phosphate (100 mM pH 7,5) à 37°C. L'avancement de la réaction a été suivi par HPLC, celle-ci étant terminée en 23h. Ensuite, le peptide  $\text{F3-NH}_2$  a été introduit, en présence d'une solution de MPAA/TCEP (100 mM et 200 mM, dans un tampon phosphate 100 mM pH 7,5). Le milieu réactionnel a été agité à 37°C

et l'avancement de la réaction suivi par HPLC, qui s'est complétée en 21h. Le milieu réactionnel a ensuite été acidifié par une solution de TFA 10%, puis le MPAA extrait à l'éther. La phase aqueuse a finalement été purifiée par HPLC préparative et la protéine ZIF(Cs124|DOTA[Tb]) obtenue (rendement 23%, 52 nmol) a été caractérisée par HPLC et ESI-MS (Figure 78).

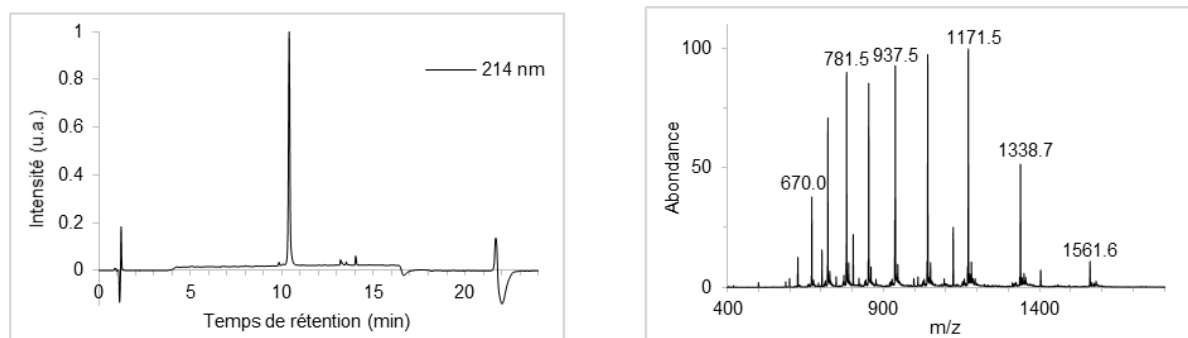


Figure 78: A gauche chromatogramme HPLC de la protéine ZIF(Cs124|DOTA[Tb]); A droite, spectre ESI-MS (masse calculée : 1561,25  $[M+6H]^{6+}$ , 1338,36  $[M+7H]^{7+}$ , 1171,19  $[M+8H]^{8+}$ , 937,15  $[M+10H]^{10+}$ , 781,13  $[M+12H]^{12+}$ , 668,68  $[M+14H]^{14+}$ , for  $M = C_{398}H_{626}N_{129}O_{110}S_7Tb$ )

Les conditions d'O.Melnyk ne sont pas faciles à contrôler, car l'assemblage est décrit par l'introduction des peptides F1, F2, et F3 dans un rapport 1:1:1<sup>[115]</sup>. Le peptide F2[Cys(StBu)|Glu(Cs124)]-SEA<sup>off</sup> étant le plus difficile à obtenir de par son faible rendement, il paraissait logique de l'introduire en tant que réactif limitant lors de la première ligation. Cependant, cette stratégie a laissé la possibilité d'un assemblage F1-F3 après introduction du peptide F3-NH<sub>2</sub>. Cependant, cet assemblage possédant le même temps de rétention que la protéine ZIF(Cs124|DOTA[Tb]), l'étape de purification s'est avérée être compliquée, les deux produits ne pouvant être correctement séparés. Il a donc fallu introduire le peptide F2 en excès lors de la première ligation afin que l'assemblage résiduel soit le peptide F2-F3, possédant un temps de rétention plus long que celui de la protéine.

### 5) Formation des doigts de zinc

La protéine ZIF(Cs124|DOTA[Tb]) synthétisée a été métallée au zinc afin de l'avoir sous une forme active. Les protéines métallées au zinc seront pour l'ensemble de ce manuscrit notées sous la forme ZIF<sup>Zn</sup>(XXX).

La quantité de protéine obtenue a été déterminée par spectrophotométrie UV via la mesure de l'absorbance du Cs124, possédant des maxima à 331 nm ( $\epsilon = 18\,900\text{ cm}^{-1}\text{M}^{-1}$ ) et 341 nm ( $\epsilon = 17\,000\text{ cm}^{-1}\text{M}^{-1}$ )<sup>[88]</sup>, longueurs d'ondes où les protéines n'absorbent pas. La protéine synthétisée a été dosée (Figure 79) dans un tampon HEPES (10 mM, pH 7,5) par mesure de son absorbance à 331 nm, puis 2,2 eq. en ZnCl<sub>2</sub> y ont été additionnés.

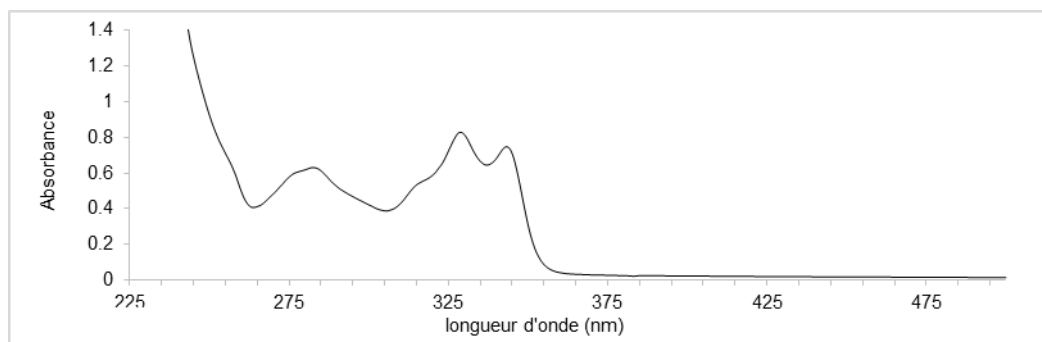


Figure 79: Spectre d'absorbance de la protéine ZIF(DOTA[Tb])/Cs124 dans un tampon HEPES 100 mM, pH 7,5

## VI Interaction de la sonde avec l'ADN

L'interaction de la sonde avec l'ADN a d'abord été étudiée par une technique classique de biologie moléculaire, le retard sur gel (electrophoretic mobility shift assay, EMSA) ou décalage de mobilité électrophorétique, puis par luminescence.

Pour cela, les deux oligonucléotides complémentaires 5'-GGTTGCAGTGGGCGCGCCCACAGTACTTGAACGTAACG-3' et 5'-CGTTACGTTCAAGTACTGTGGGCGCGCCCACTGCAACC-3' ont été hybridés pour former le duplex d'ADN utilisé pour l'ensemble des tests. Ces oligonucléotides ont été synthétisés par voie automatisée et pré-purifiés sur HPLC par Rémy Lartia. La dernière étape de purification a été réalisée par électrophorèse sur gel de polyacrylamide préparatif.

### 1) Mobilité électrophorétique

#### A Principe

La technique de gel retard a pour but de mettre en évidence l'interaction ADN-protéine, en faisant migrer par électrophorèse sur gel non dénaturant l'ADN libre et le complexe ADN-protéine. Une différence de migration est observée car le complexe ADN-protéine plus lourd migre moins vite que l'ADN libre (Figure 80).

Afin de pouvoir visualiser les différentes espèces introduites sur le gel, les brins d'ADN sont préalablement radiomarqués enzymatiquement au P<sup>32</sup>. Les fragments d'ADN marqués sont mis en évidence et quantifiés à l'aide d'un phosphorimager. Le ratio (ADN libre)/(ADN-protéine) peut donc être déterminé.

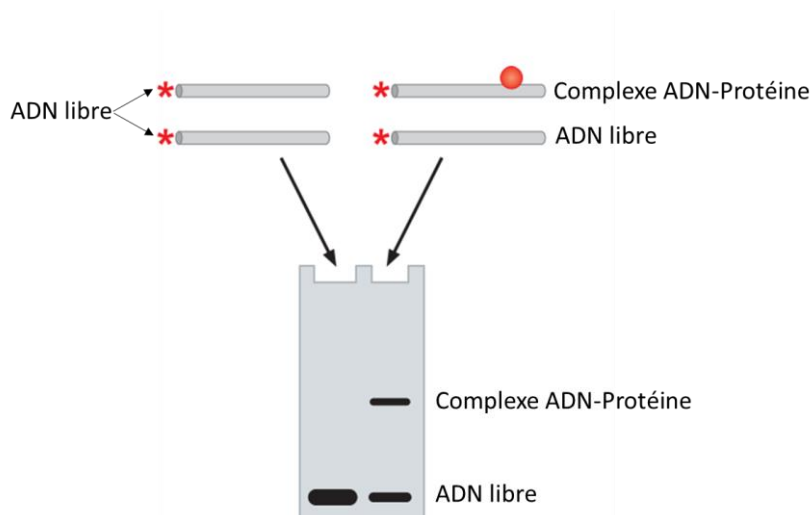


Figure 80: A gauche, échantillon d'ADN seul servant de référence pour la migration. A droite, échantillon d'ADN en présence de protéine. En présence de la protéine, le complexe ADN-protéine est retenu comparé à l'ADN seul

### **B Marquage**

Afin d'effectuer le radiomarquage, un seul des deux brins du duplex d'ADN a été marqué au phosphore 32. Pour cela, l'oligonucléotide a été introduit dans une solution de tampon HEPES pH 7,5 auquel en présence de Kinase et d'ATP  $\gamma$   $^{32}\text{P}$  le groupement phosphate radioactif en position  $\gamma$  est transféré par la kinase sur l'extrémité 5'OH de l'oligonucléotide. Le brin d'ADN marqué a ensuite été hybridé avec son brin complémentaire introduit en léger excès (1,1 éq.) de manière à former le duplex d'ADN cible.

Ce duplex marqué a été mis en présence de protéine puis incubé 18h à 4°C. Les différentes espèces ont été séparées par électrophorèse sur un gel de polyacrylamide à 10 % en conditions non dénaturantes. Le gel a été refroidi par circulation d'eau à 4°C et l'intensité du courant électrique ajustée pour ne pas dépasser une puissance de 3 watts et maintenir le gel à une température inférieure à 10°C. Après 3 heures de migration les bandes radioactives correspondant à l'ADN libre et l'ADN lié ont été quantifiées à l'aide d'un phosphorimager.

### **C Mesures effectuées**

La protéine ZIF<sup>Zn</sup>(nue) identique à la séquence décrite par B. Wang et C. Pabo a été utilisée en premier lieu afin de mettre au point les conditions de migration du gel retard. Dans le cas des travaux de B. Wang et C. Pabo, l'interaction de la protéine avec l'ADN cible dont la concentration est fixée à 10 nM se définit par une constante de dissociation de 15nM.

### i Protéine ZIF<sup>Zn</sup>(nue)

Nous avons réalisé une étude préliminaire en utilisant une gamme de concentration croissante de 0,9 nM, 9 nM, 90 nM, 900 nM, 9  $\mu$ M, de la protéine ZIF<sup>Zn</sup>(nue). Celle-ci a été incubée une nuit avec le duplex d'ADN marqué à une concentration constante de 10 nM dans une solution HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM). La composition du tampon a été simplifiée par rapport à celle utilisée par B. Wang et C. Pabo. Les différents mélanges ADN/Protéine ont été séparés par électrophorèse sur gel de polyacrylamide. L'analyse du gel a révélé l'interaction de la protéine entre une concentration de 90 nM et de 900 nM, avec formation de 100% du complexe ADN/protéine à 9  $\mu$ M (Figure 81). Les conditions d'incubation et de migration n'ont pas été optimisées car cette étude préliminaire avait pour but de prendre en main la technique de gel retard. Nous avons gardé ces conditions pour réaliser une étude préliminaire similaire avec notre protéine ZIF<sup>Zn</sup>(Cs124|DOTA[Tb]).

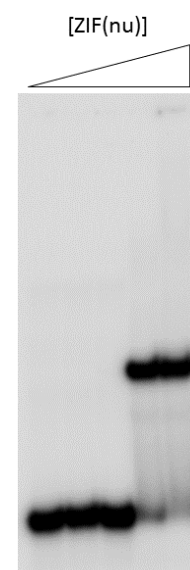


Figure 81: Gel retard de la protéine ZIF<sup>Zn</sup>(nue); de gauche à droite concentration croissante de protéine sur l'ADN

### ii Etude préliminaire sur la protéine ZIF<sup>Zn</sup>(Cs124|DOTA[Tb])

Pour cette étude, une gamme de concentration croissante de 6 nM, 60 nM, 600 nM, 6  $\mu$ M, de la protéine ZIF<sup>Zn</sup>(Cs124|DOTA[Tb]) a été incubée une nuit avec le duplex d'ADN marqué à une concentration constante de 10 nM dans les mêmes conditions qu'avec la protéine ZIF(nue). Le gel de la Figure 82 montre que la formation de complexe ADN/protéine est presque quantitative à une concentration de 600 nM en protéine. Par contre, à 6  $\mu$ M, la bande en haut du gel correspond à la formation d'un précipité qui n'a pas migré. Ceci est sans doute dû à des interactions non spécifiques avec l'ADN est la formation d'espèces insolubles (Figure 82).

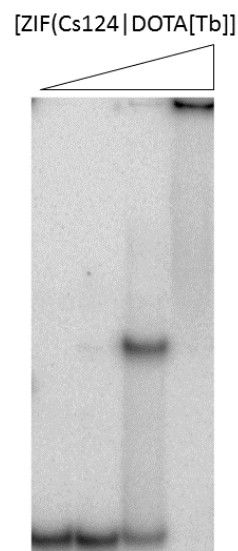


Figure 82: Gel retard de la protéine ZIF<sup>Zn</sup>(Cs124|DOTA[Tb]); de gauche à droite concentration croissante de protéine sur l'ADN

Cet effet n'a pas été observé avec la protéine nue, sans doute plus hydrosoluble.

Il faut remarquer que la protéine ZIF<sup>Zn</sup>(Cs124|DOTA[Tb]) diffère en 4 points de la protéine initialement décrite par Wang et Pabo correspondant à la protéine ZIF<sup>Zn</sup>(nue) :

- L'introduction du Cs124 sur l'acide glutamique n°110
- La modification de l'asparagine n°93 par une lysine conjuguée à un complexe DOTA[Tb]
- Le remplacement de l'acide aspartique n°113 en acide glutamique pour éviter la formation d'aspartimides
- L'acétylation de la protéine en N-ter

En résumé, cette étude montre plusieurs choses importantes :

- La protéine « nue » obtenue par voie synthétique est capable de se lier à l'ADN.
- L'adjonction de chromophores (Cs124 et DOTA[Tb]), l'acétylation N-terminale et la mutation de l'aspartate 113 en glutamate n'empêchent pas la liaison à l'ADN.
- En première approximation, la protéine « nue » et la protéine fonctionnalisée ont des affinités similaires.

## 2) Luminescence

### A Caractérisation de la protéine ZIF<sup>Zn</sup>(Cs124|DOTA[Tb])

Les mesures fluorimétriques de la sonde ZIF<sup>Zn</sup>(Cs124|DOTA[Tb]) ont été réalisées avec une concentration de protéine de 1  $\mu$ M en solution dans un mélange HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM).

Le spectre d'émission de fluorescence, obtenu en excitant à 330 nm, montre une émission de 345 à 445 nm avec un maximum d'émission à 373 nm (Figure 83, droite), caractéristique de la fluorescence du Cs124. Le spectre d'excitation de fluorescence ( $\lambda_{em} = 380$  nm) correspond bien au spectre d'absorption du Cs124 avec les deux maxima à 331 et 344 nm (Figure 83, gauche).

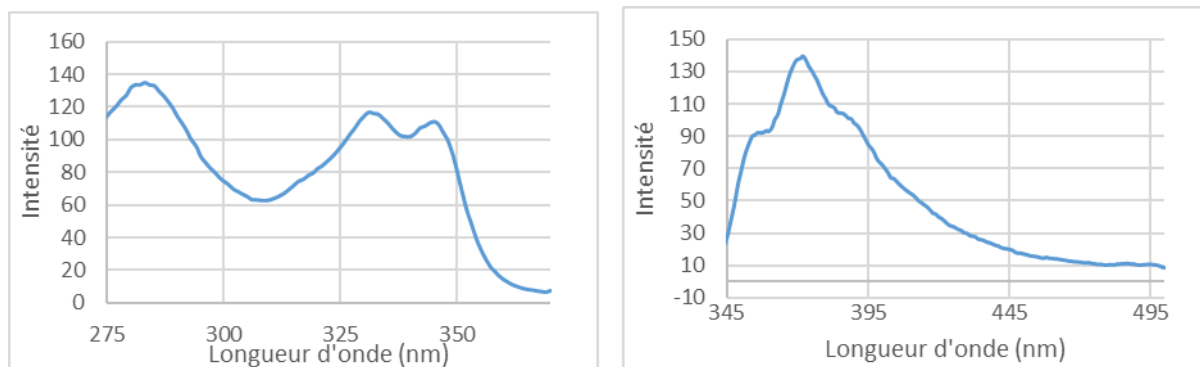


Figure 83: A gauche : spectre d'excitation ( $\lambda_{em}=380$  nm). A droite : spectre d'émission ( $\lambda_{ex}=330$  nm).

Les spectres de phosphorescence ont été mesurés avec une détection résolue en temps correspondant à un délai de 100  $\mu$ s. Le spectre d'émission obtenu en excitant à 348 nm montre les quatre raies caractéristiques des transitions  $^5D_4 \rightarrow ^7F_n$  ( $n=3$  à 6) du terbium à 490, 545, 585 et 623 nm (Figure 84, droite). Le spectre d'excitation ( $\lambda_{em} = 545$  nm) montre deux maxima à 331 et 344 nm, indiquant une sensibilisation du terbium par le Cs124 (Figure 84, gauche).

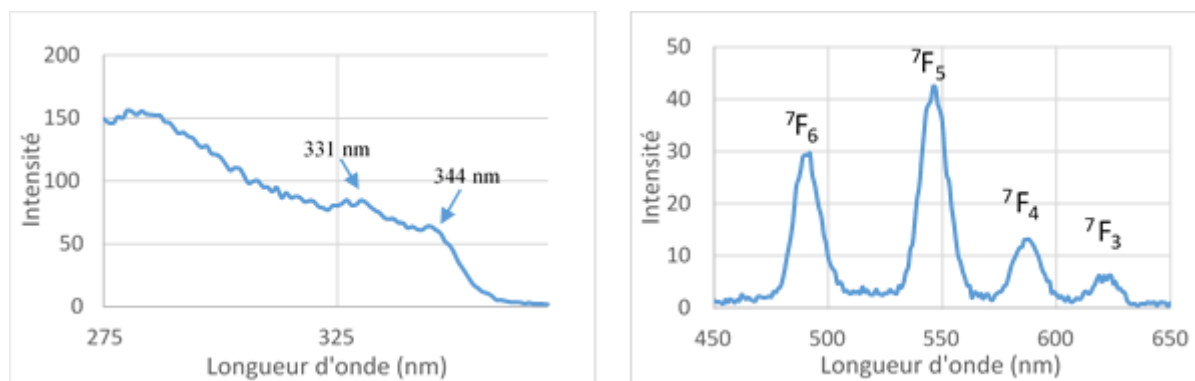


Figure 84: A gauche spectre d'émission ( $\lambda_{ex}=348$  nm); A droite, spectre d'excitation ( $\lambda_{em}=545$  nm)

### B Titration à l'ADN

L'interaction avec l'ADN a été étudiée par spectroscopie d'émission dans un tampon HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM) avec une concentration de sonde  $ZIF^{Zn}(Cs124|DOTA[Tb])$  de 1  $\mu$ M. Les oligonucléotides utilisés 5'-CGCATATGGGCGCGC-CCAATAGCG-3' et 5'-CGCTATTGGGCGCGCCCATATGCG-3' sont plus courts que ceux des expériences de gel retard. D'après les expériences de gel retard, à cette concentration, la fixation de l'ADN à la protéine doit être quasi quantitative. Le titrage a été réalisé sur 3 points : 0 eq., 1 eq., 2 eq., en duplex d'ADN par dimère protéique soit des concentrations de 0, 0,5 et 1  $\mu$ M en ADN. Les spectres d'émission résolus en temps ont été enregistrés avec excitation à 348 nm.

Une légère augmentation de l'émission du terbium est observée entre 0 eq et 1 eq en ADN (Figure 85), puis aucune différence n'est observée entre 1 eq et 2 eq, suggérant la formation du complexe 2 :1 ( $ZIF^{Zn}(Cs124|DOTA[Tb])/ADN$ ) attendu. Néanmoins, la très faible augmentation d'intensité de l'émission du terbium indique aussi que le transfert d'énergie du Cs124 vers le terbium ne varie quasiment pas entre la protéine  $ZIF^{Zn}(Cs124|DOTA[Tb])$  seule et le complexe ADN/protéine, indiquant que la distance entre le Cs124 et le terbium au sein du dimère formé n'est pas significativement différente pour obtenir un transfert d'énergie beaucoup plus efficace.

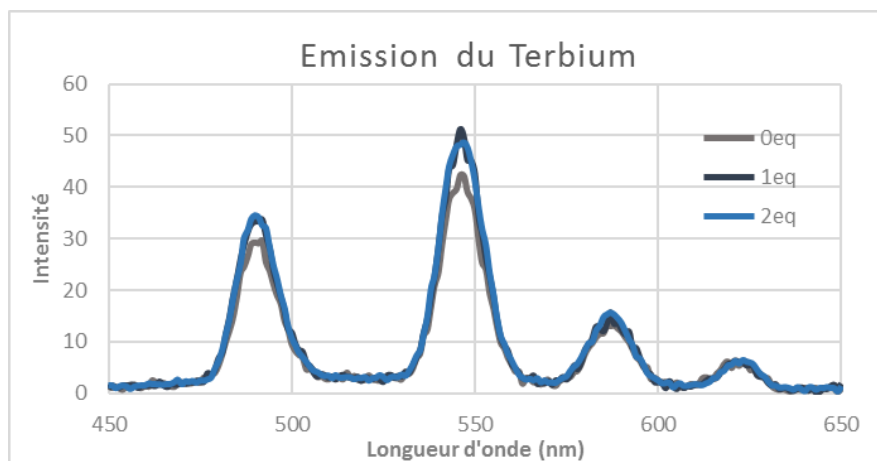


Figure 85: Spectres d'émission du terbium à 0, 1, et 2 eq en ADN

Cela peut être expliqué par le fait que la distance Cs124/Tb, estimée à 10-15Å à partir de la structure RX de B. Wang et C. Pabo, est probablement plus importante du fait de la grande flexibilité des chaînes latérales sur lesquelles sont conjugués les chromophores.

## VII Conclusion

Un système de FRET d'une antenne vers un lanthanide a été implémenté sur la protéine de B. Wang et C. Pabo pour concevoir une première sonde. Malgré les diverses modifications réalisées, l'interaction avec sa séquence d'ADN cible (5'-TGGGCG-3') a été mise en évidence par retard sur gel d'électrophorèse. Cependant, le titrage par luminescence montre une très faible augmentation de l'émission du terbium en présence d'ADN, insuffisante pour envisager de continuer avec cette sonde. Cela étant probablement dû à une distance entre les chromophores supérieure au rayon de Forster du couple Cs124-Tb, il faut soit optimiser leurs positions, soit remplacer le système FRET par un autre plus efficace. Le choix d'autres positions étant très limité, il est plus intéressant de changer de système de FRET en passant à la seconde génération de sondes dont le système repose sur un transfert d'énergie efficace à de plus grandes distances. La synthèse de ces nouvelles sondes sera l'objet du chapitre suivant. Ces résultats mitigés nous ont cependant permis de mettre au point la synthèse par voie chimique de la protéine de B. Wang et C. Pabo fonctionnalisée par des chromophores.



Chapitre II : Sondes de seconde  
génération, FRET entre chromophores  
organiques



# Chapitre II : Sondes de seconde génération, FRET entre chromophores organiques

## I Introduction

### 1) Principe

Le principe des sondes de seconde génération repose sur la formation d'un hétéro-dimère protéique ciblant un duplex d'ADN de séquence choisie et incorporant un système FRET pour la signalisation. Notre but est de concevoir un système FRET pour lequel un complexe de lanthanide est le donneur d'énergie et un fluorophore organique l'accepteur, afin de bénéficier des propriétés des lanthanides exposées dans le chapitre d'introduction. Cependant, cela nécessite de disposer de complexes de lanthanides possédant à la fois de bons rendements quantiques et une fonction de conjugaison à une biomolécule. De tels complexes ne sont pas commerciaux et peu d'exemples sont décrits dans la littérature<sup>[135],[136]</sup>. Nous avons donc décidé d'établir une preuve de concept reposant d'abord sur un FRET entre deux fluorophores organiques. Le système de détection de seconde génération proposé repose sur une unique protéine à laquelle est conjuguée soit le fluorophore donneur, soit le fluorophore accepteur, les deux fluorophores étant introduits sur une même position pour donner les deux composantes de l'hétéro-dimère (Figure 86).

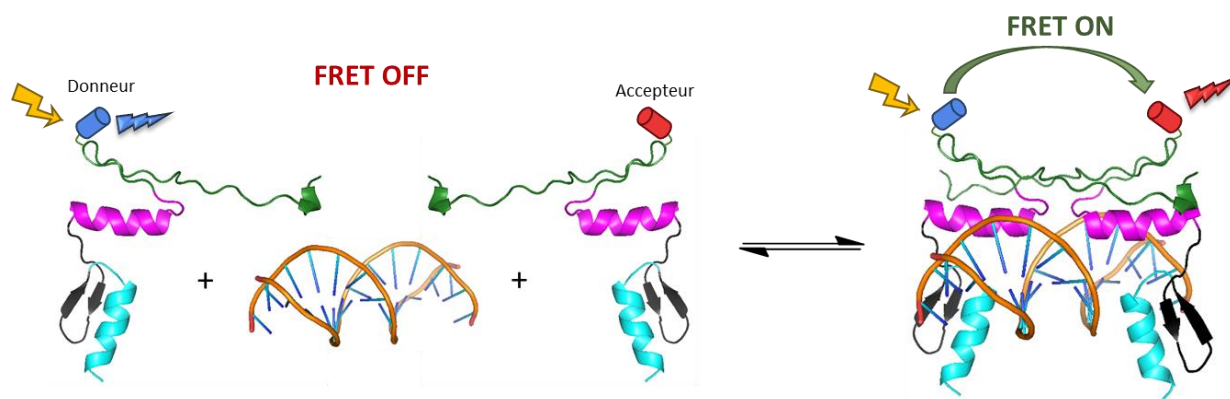


Figure 86: Principe des sondes de seconde génération ; ici le donneur et l'accepteur sont des fluorophores

### 2) Stratégie de synthèse des sondes

La stratégie de synthèse de la protéine est identique à celle de la sonde de première génération, avec le même découpage de fragments peptidiques et un assemblage par ligation native. Parmi les deux positions possibles d'introduction décrites au premier chapitre, nous avons retenu la position n°110 (fragment F2), où le Cs124 était greffé pour la protéine ZIF(DOTA[Tb][Cs124])

vue au premier chapitre. L'introduction du chromophore sur cette position et non sur la position n°93 permet d'avoir une séquence qui se rapproche davantage de la protéine décrite par B. Wang et C. Pabo :

- aucun acide aminé de l'interface de dimérisation n'est modifié,
- l'acétylation en N-terminal de la protéine n'est plus nécessaire.

Ainsi, seuls deux acides aminés sont modifiés par rapport à la séquence initialement décrite par B. Wang et C. Pabo :

- l'acide glutamique n°110 remplacé par une lysine, permettant l'introduction du chromophore sur sa chaîne latérale,
- l'acide aspartique n°113 remplacé par un acide glutamique, empêchant la formation d'aspartimides.

En conservant le découpage adopté pour la sonde de première génération (Figure 87 (A)), le fragment 3 reste inchangé mais les autres fragments doivent être modifiés et synthétisés à nouveau (Figure 87). Pour chaque protéine composant le système FRET, il faudra un fragment central F2 différent incorporant le fluorophore, mais le fragment N-terminal F1 est commun. Cela demande de synthétiser les peptides F1-MPA et F2[Cys(StBu)|Lys(X)]-SEA<sup>off</sup>, où X est le fluorophore, pour l'assemblage final.

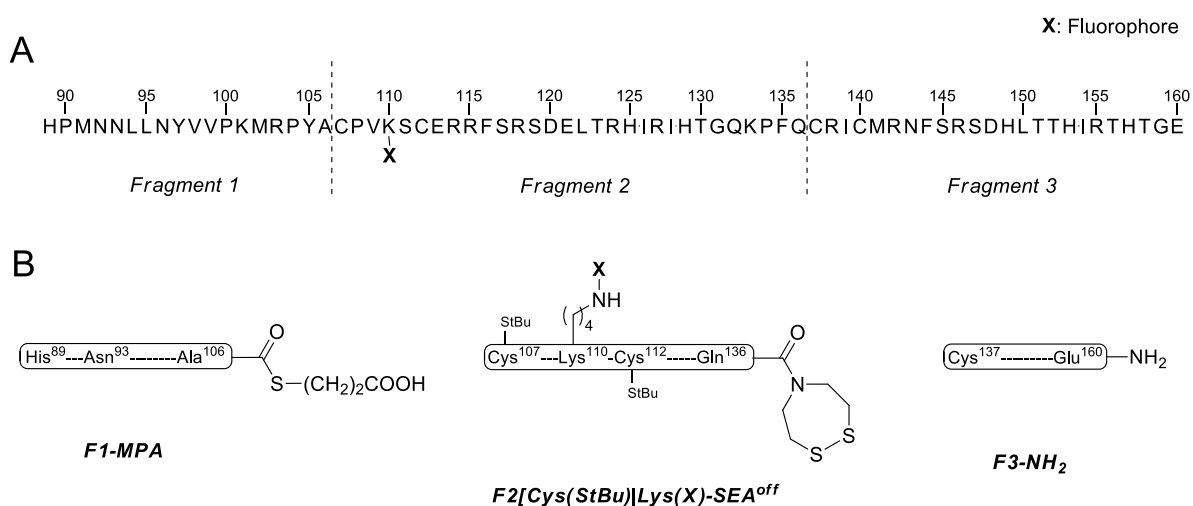


Figure 87: (A) Découpage de la protéine pour les sondes de seconde génération. (B) Fragments peptidiques nécessaires pour l'obtention des sondes de seconde génération.

### 3) Choix des fluorophores du système de FRET

La distance mesurée entre les carbones  $\alpha$  des deux positions n°110 du dimère de la structure RX décrite par B. Wang et C. Pabo est de 44 Å. Des systèmes FRET dont le rayon de Förster

est voisin de 50 Å peuvent facilement être obtenus à l'aide de fluorophores organiques. Le choix des deux fluorophores s'est orienté vers des composés pouvant être utilisés comme accepteurs de complexes de terbium ou d'euprium dans la stratégie de FRET d'un lanthanide vers un fluorophore organique. Nous avons donc sélectionné des fluorophores dont les spectres d'absorption recouvrent les spectres d'émission du terbium ou de l'euprium. Le choix des fluorophores a donc été dicté par plusieurs critères. Ils devaient :

- être compatibles pour un système FRET dont le rayon de Förster est d'environ 50 Å,
- être utilisables comme accepteurs d'énergie d'un lanthanide (terbium ou euprium),
- être photostables,
- être chimiquement compatibles avec la voie de synthèse choisie pour les protéines,
- disposer d'une fonction de conjugaison,
- être commerciaux et peu coûteux ou facilement synthétisables.

De nombreux types de fluorophores ont été développés au long de ces dernières décennies, laissant un large choix dans la gamme de molécules disponibles<sup>[137]</sup>, parmi lesquelles se situent les rhodamines.

Les rhodamines sont des fluorophores appartenant à la famille des xanthènes (Figure 88), dont la photostabilité a conduit à un large engouement dans la synthèse de nombreux dérivés, avec plus de 200 rhodamines recensées en 2019<sup>[138]</sup>. Suivant les groupements et la nature de l'atome en position n°10 (X sur la Figure 88), le maximum d'absorption peut varier en moyenne de 480 à 730 nm avec quelques exceptions allant jusqu'à 880 nm<sup>[139]</sup>. Ce type de fluorophore très performant est généralement commercialisé à des prix très élevés, à quelques exceptions près comme la sulforhodamine B, aussi appelée lissamine (Figure 88). Cette rhodamine est commerciale à un prix modeste (environ 200 € / g, soit 100 fois moins cher que beaucoup de rhodamines commerciales) sous forme d'un chlorure de sulfonyl facilement conjugable à une biomolécule. La lissamine possède des propriétés photophysiques intéressantes<sup>[140]</sup> ( $\lambda_{\text{ex}} = 560 \text{ nm}$ ,  $\lambda_{\text{em}} = 580 \text{ nm}$ ,  $\epsilon_{\text{max}} = 71\,500 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\Phi(\text{H}_2\text{O}) = 0,33$ ). Son spectre d'absorption, recouvrant la zone spectrale de 460 à 600 nm, possède un bon recouvrement avec le spectre d'émission du terbium (Figure 89).

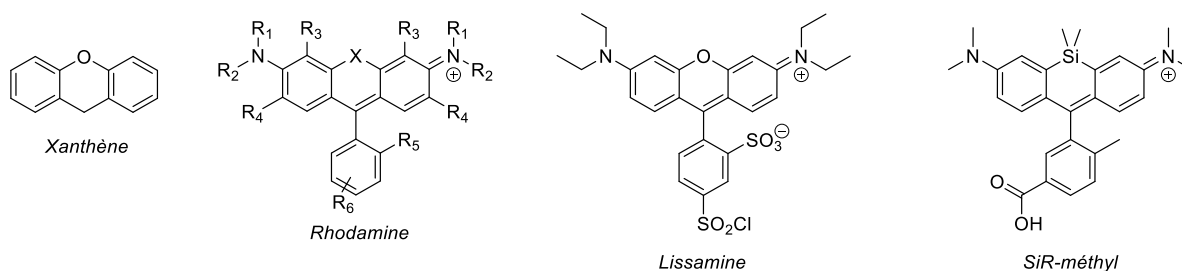


Figure 88: Structure générale du Xanthène et d'une Rhodamine, et structures de la lissamine commerciale sous forme de chlorure de sulfonyle, ainsi que de la SiR-méthyl

Les Si-rhodamines sont des fluorophores qui émettent à des longueurs d'ondes plus élevées que les rhodamines<sup>[141],[142]</sup>. La Si-rhodamine méthyl noté SiR-méthyl (Figure 88), de par ses propriétés photophysiques<sup>[142]</sup> ( $\lambda_{\text{ex}} = 646 \text{ nm}$ ,  $\lambda_{\text{em}} = 660 \text{ nm}$ ,  $\epsilon_{\text{max}} = 100\,000 \text{ M}^{-1} \text{ cm}^{-1}$ ,  $\Phi(\text{PBS pH } 7,4) = 0,3$ ) est un fluorophore pouvant jouer le rôle d'accepteur de la lissamine dans notre système. De plus, son spectre d'absorption en fait un accepteur potentiel pour un FRET avec l'euporium comme donneur, mais aussi avec du terbium dans une efficacité moindre, du fait d'un moins bon recouvrement (Figure 89). Ce composé est disponible au laboratoire par le biais d'une collaboration avec l'équipe de N. McClenaghan (Univ. Bordeaux).

Le rayon de Förster  $R_0$  (en Å) pour le couple lissamine-SiR-méthyl peut être calculé par

$$R_0 = 0.211(\kappa^2 n^{-4} Q_D J(\lambda))^{1/6}$$

où  $\kappa^2$  est un facteur d'orientation entre les deux chromophores (généralement égal à 2/3),  $n$  l'indice de réfraction du milieu,  $Q_D$  le rendement quantique du donneur et  $J(\lambda)$  l'intégrale de recouvrement entre l'émission du donneur et l'absorption de l'accepteur.  $J(\lambda)$  est calculé par

$$J(\lambda) = \int_0^{\infty} F_D(\lambda) \epsilon_A(\lambda) \lambda^4 d\lambda$$

où  $F_D(\lambda)$  est le spectre d'émission du donneur normalisé à une aire sous la courbe de 1 et  $\epsilon_A(\lambda)$  le spectre d'absorption de l'accepteur. Pour le couple lissamine-SiR-méthyl,  $R_0$  peut être estimé à 55 Å, ce qui doit être suffisant pour la sonde que nous voulons développer.

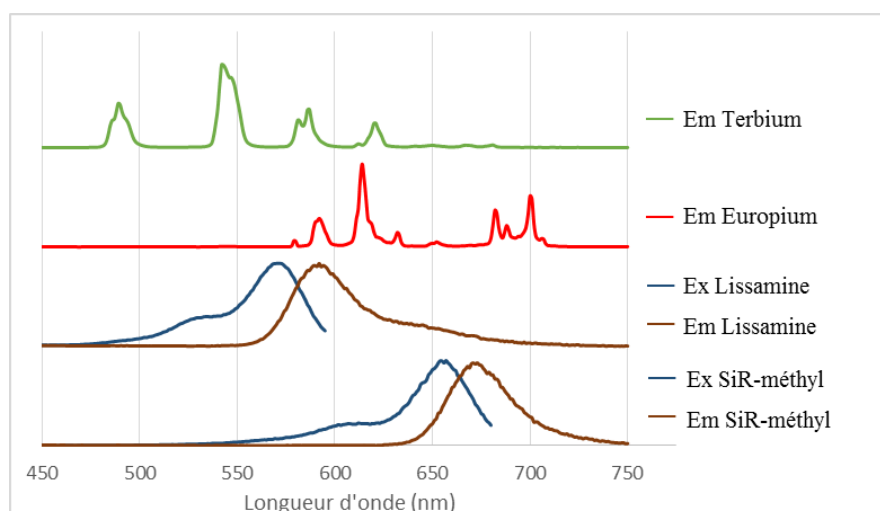


Figure 89: Différents spectres d'excitation et d'émission des chromophores utilisés

## II Synthèse des protéines ZIF<sup>Zn</sup>(lissamine) et ZIF<sup>Zn</sup>(SiR-méthyl)

La sonde permettant un FRET organique requiert l'utilisation de deux protéines métallées au zinc contenant chacune soit la lissamine, soit la SiR-méthyl, dont la notation utilisée pour la suite du manuscrit sera ZIF<sup>Zn</sup>(lissamine) et ZIF<sup>Zn</sup>(SiR-méthyl).

### 1) Synthèse des protéines ZIF(lissamine) et ZIF(SiR-méthyl)

Les fragments présentés en Figure 87 B] comportant en X soit la lissamine, soit la SiR-méthyl permettent après assemblage l'obtention des protéines non métallées appelées ZIF(lissamine) et ZIF(SiR-méthyl)

#### A Fragment F1-MPA

Le peptide *F1-MPA* a été synthétisé de manière analogue au peptide Ac-F1[Lys(DOTA[Tb])]-MPA (Figure 90), avec un rendement de 40% et caractérisé par HPLC et ESI-MS (Figure 91).

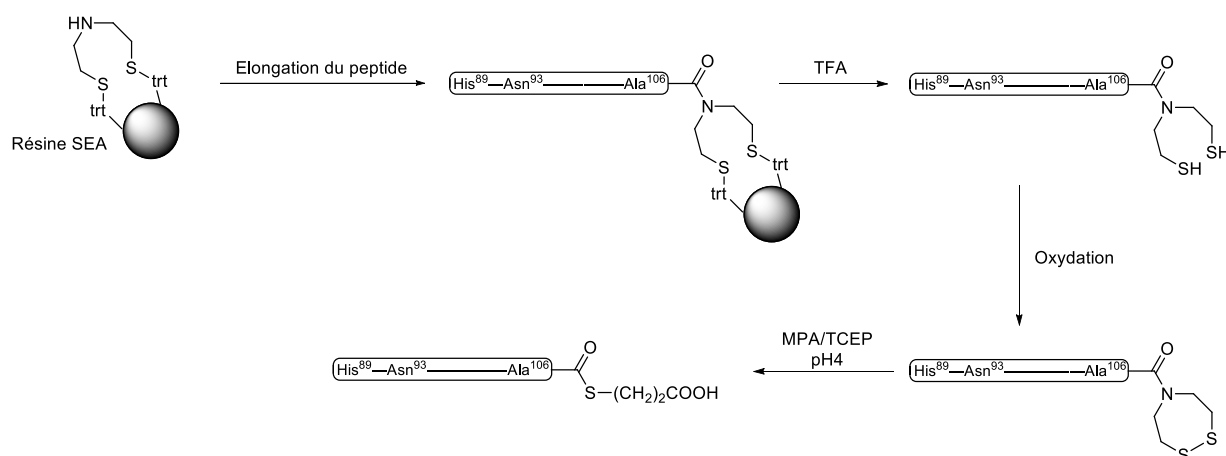


Figure 90: Schéma de synthèse du peptide F1-MPA

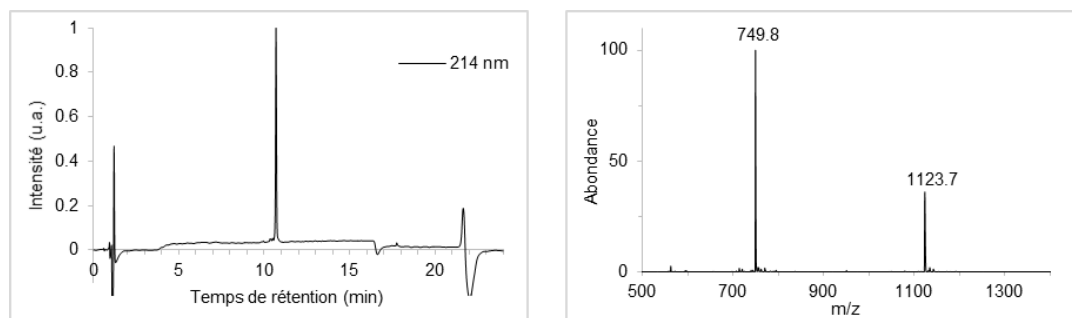


Figure 91: A gauche, chromatogramme HPLC de F1-MPA ; A droite, spectre ESI-MS (masse calculée : 1123,86  $[M+2H]^{2+}$ , 749,57  $[M+3H]^{3+}$  pour  $M = C_{101}H_{157}N_{27}O_{25}S_3$ )

### B Fragments F2

Les deux fragments centraux F2 présentant les fluorophores lissamine et SiR-méthyl ont été obtenus à partir d'un précurseur commun, noté F2[Cys(StBu)Lys(NH<sub>2</sub>)]-SEA-2Cl-trt-PS, contenant un acide aminé Lys(alloc) en position n°110. Le groupement protecteur alloc peut être éliminé sélectivement sur résine pour démasquer la chaîne latérale de la lysine et coupler les fluorophores suivant différentes conditions (Figure 92).

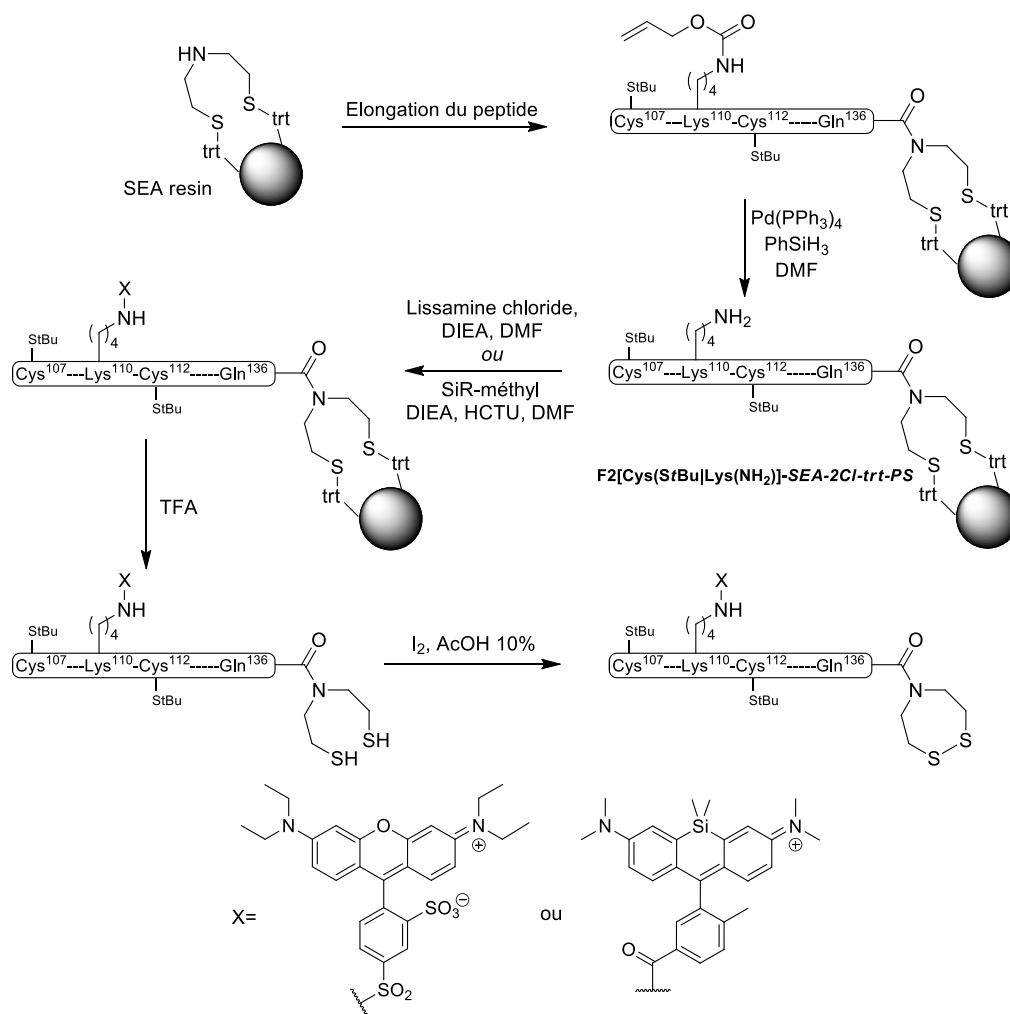


Figure 92: Schéma de synthèse des peptides F2[Cys(StBu)Lys(lissamine)]-SEA<sup>off</sup> et F2[Cys(StBu)Lys(SiR-méthyl)]-SEA<sup>off</sup>

### i Peptide F2[Cys(StBu)]Lys(lissamine)]-SEA<sup>off</sup>

Le peptide F2[Cys(StBu)]Lys(lissamine)]-SEA<sup>off</sup> a été synthétisé avec greffage de la sulforhodamine B sous forme de chlorure de sulfonyl. Le couplage s'est effectué sur la lysine déprotégée du peptide F2[Cys(StBu)Lys(NH<sub>2</sub>)]-SEA-2Cl-trt-PS, en milieu basique, suivi de la déprotection et l'oxydation du groupement SEA (Figure 92). Le profil HPLC du brut de réaction après oxydation au diamide (Figure 93) a montré la présence de nombreux sous-produits, dont le produit de conversion de la Cys(StBu) finale en alanine, observé en ESI-MS. Le rendement de synthèse du peptide a été affecté par une étape de purification difficile à réaliser. Le peptide obtenu (rendement de 4%) a été analysé par HPLC et ESI-MS (Figure 94).

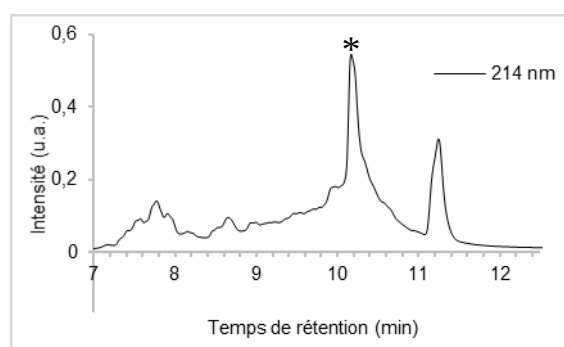


Figure 93: Chromatogramme du brut de déprotection du peptide F2[Cys(StBu)]Lys(lissamine)]-SEA<sup>off</sup> (\*pic du peptide)

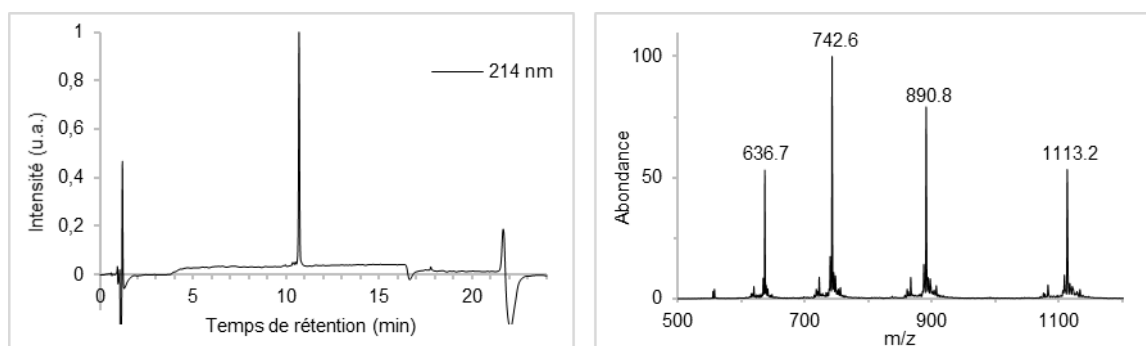


Figure 94: A gauche, chromatogramme HPLC de F2[Cys(StBu)]Lys(lissamine)]-SEA<sup>off</sup> ; A droite, spectre ESI-MS (masse calculée : 1112,84 [M+4H]<sup>4+</sup>, 890,48 [M+5H]<sup>5+</sup>, 742,23 [M+6H]<sup>6+</sup>, 636,34 [M+7H]<sup>7+</sup> pour M = C<sub>193</sub>H<sub>302</sub>N<sub>56</sub>O<sub>49</sub>S<sub>8</sub>)

### ii Peptide F2[Cys(StBu)]Lys(SiR-méthyl)]-SEA<sup>off</sup>

Le peptide F2[Cys(StBu)]Lys(SiR-méthyl)]-SEA<sup>off</sup> a été synthétisé avec greffage de la rhodamine SiR-méthyl dans des conditions de couplage peptidique classique sur le peptide F2[Cys(StBu)Lys(NH<sub>2</sub>)]-SEA-2Cl-trt-PS, suivi de la déprotection et l'oxydation du groupement SEA (Figure 92). Le peptide obtenu a été purifié avec un rendement de 13,3% et caractérisé par HPLC et ESI-MS (Figure 95).

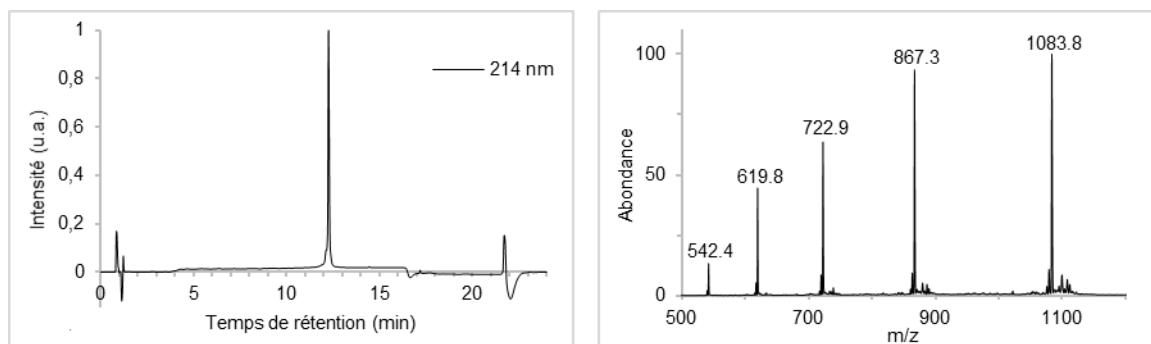


Figure 95: A gauche, chromatogramme HPLC de F2[Cys(StBu)]Lys(SiR-méthyl)]-SEA<sup>off</sup> ; A droite, spectre ESI-MS (masse calculée : 1083,83 [M+4H]<sup>4+</sup>, 867,27 [M+5H]<sup>5+</sup>, 722,89 [M+6H]<sup>6+</sup>, 619,77 [M+7H]<sup>7+</sup>, 542,42 [M+8H]<sup>8+</sup> pour  $M = C_{193}H_{302}N_{56}O_{44}S_6Si$ )

### C Assemblages

Les deux protéines (ZIF(lissamine) et ZIF(SiR-méthyl)) ont été assemblées dans les mêmes conditions que celles décrites en partie « Assemblage des différents fragments » (p.67) par deux réactions de ligation successives, impliquant en premier lieu le fragment 1 et le fragment 2, puis en second lieu le peptide obtenu issu de la première ligation, avec le fragment 3 (Figure 96).

Les protéines ZIF(lissamine) ZIF(SiR-méthyl) ont été obtenues avec respectivement un rendement de 30% et de 8%. Les chromatogrammes HPLC et spectres ESI-MS sont représentés Figure 97 pour ZIF(Lissamine) et Figure 98 pour ZIF(SiR-méthyl).

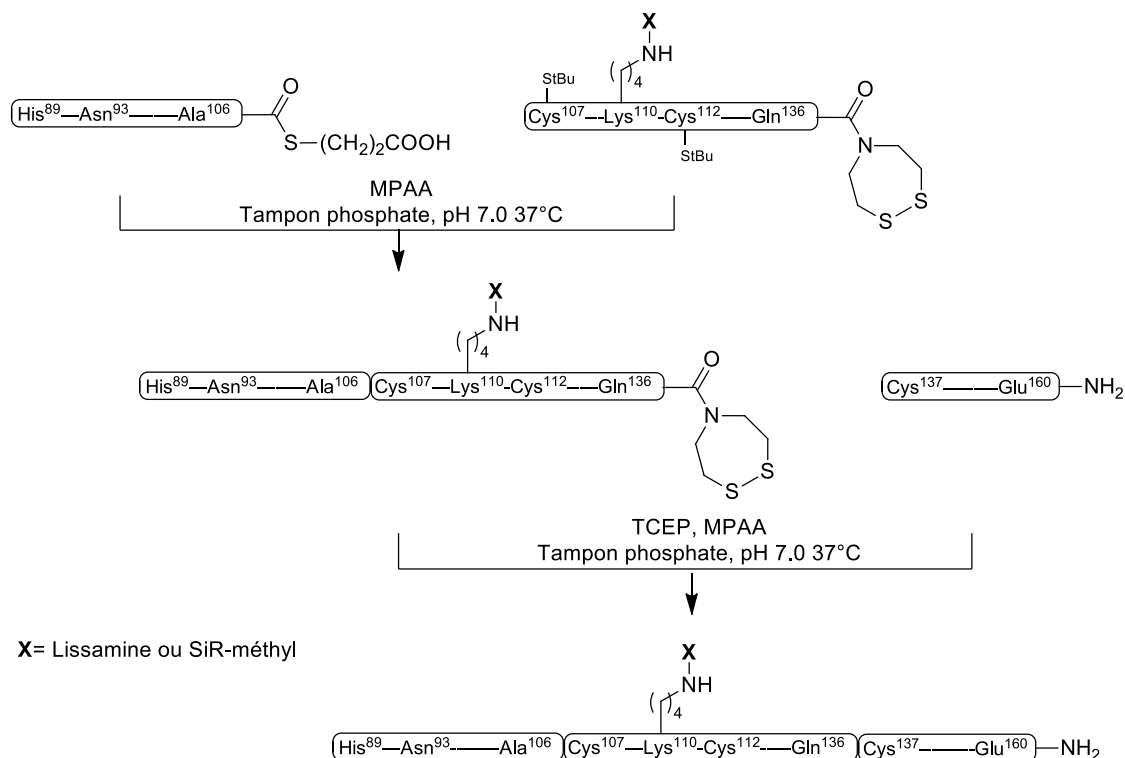


Figure 96: Schéma d'assemblage one-pot pour les protéines ZIF(Lissamine) et ZIF(SiR-méthyl)

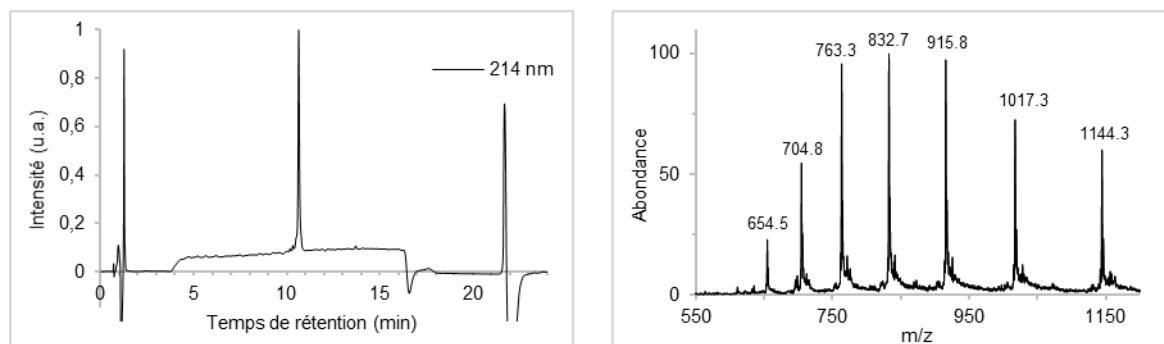


Figure 97: A gauche, chromatogramme HPLC de ZIF(lissamine) ; A droite, spectre ESI-MS (masse calculée : 1144,32  $[M+8H]^{8+}$ , 1017,29  $[M+9H]^{9+}$ , 915,66  $[M+10H]^{10+}$ , 832,51  $[M+11H]^{11+}$ , 763,22  $[M+12H]^{12+}$ , 704,59  $[M+13H]^{13+}$ , 654,33  $[M+14H]^{14+}$ , pour  $M = C_{396}H_{620}N_{126}O_{107}S_9$ )

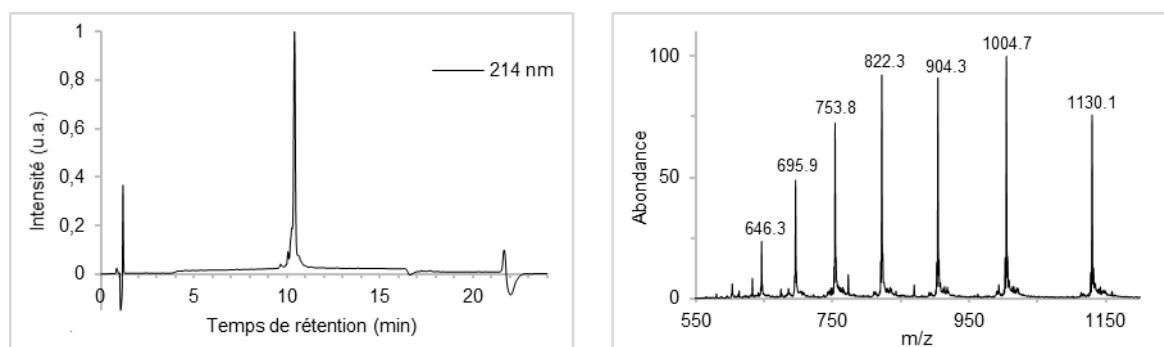


Figure 98: A gauche, chromatogramme HPLC de ZIF(SiR-méthyl) ; A droite, spectre ESI-MS (masse calculée : 1129,82  $[M+8H]^{8+}$ , 1004,40  $[M+9H]^{9+}$ , 904,06  $[M+10H]^{10+}$ , 821,96  $[M+11H]^{11+}$ , 753,55  $[M+12H]^{12+}$ , 695,66  $[M+13H]^{13+}$ , 646,04  $[M+14H]^{14+}$ , pour  $M = C_{396}H_{620}N_{126}O_{102}S_7Si$ )

## 2) Stabilité des protéines, conditions de stockage

Lors de la première synthèse de la protéine ZIF(lissamine), la caractérisation par HPLC (Figure 99 gauche) a montré la présence de sous-produit, laissant supposer une mauvaise purification. L'analyse complémentaire par ESI-MS a montré la présence de la masse du produit attendu notée M et d'autres produits comportant la masse M à laquelle s'ajoute des multiples +16, laissant présager l'oxygénation de la protéine (Figure 99 droite).

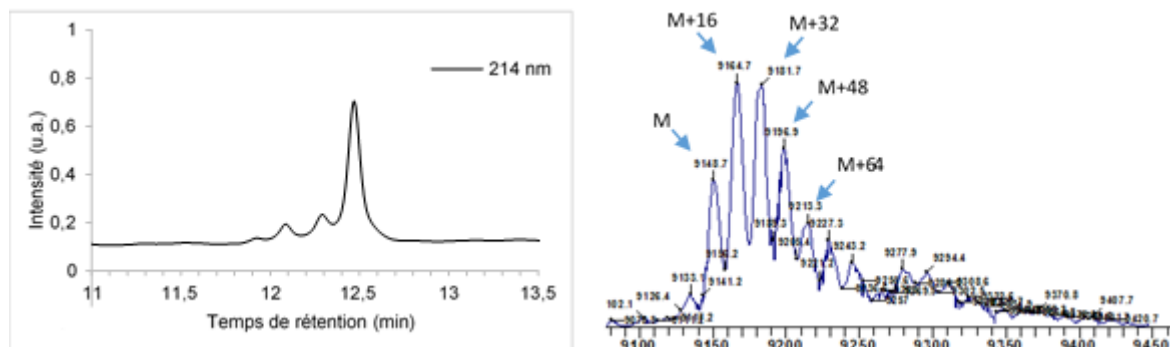


Figure 99: A gauche, chromatogramme HPLC ; A droite, spectre ESI-MS déconvolué de la protéine ZIF(liss)

L'oxydation des cystéines libres ou liées à un métal (protéines à doigts de zinc par exemple) par réaction avec des espèces provenant de la réduction de l'oxygène est un fait connu et décrit dans la littérature depuis de nombreuses années<sup>[143]</sup>. La protéine ZIF(lissamine) n'ayant pas été en contact avec de telles espèces, nous avons plutôt pensé à une oxydation par l'oxygène singulet ( $^1\text{O}_2$ ), état excité de l'oxygène triplet. Il peut être généré par transfert d'énergie depuis un photosensibilisateur, lui-même dans un état triplet excité. La capacité de l'oxygène singulet à oxyder des thiols comme les cystéines, pour donner des dissulfures mais aussi des dérivés oxygénés (sulfinate, sulfonate, thiosulfinate...) a été décrite<sup>[144],[145]</sup>. L'oxydation de cystéines engagées dans des doigts de zinc a été proposée plus récemment<sup>[146]</sup> et a été étudiée en détail au laboratoire ces dernières années<sup>[147],[148],[149]</sup>. Il a notamment été montré que les cystéines liées au zinc pouvaient réagir avec  $^1\text{O}_2$  pour donner des sulfinates ou des sulfonates mais que les doigts de zinc classiques  $\beta\beta\alpha$  sont très peu réactifs en comparaison d'autres doigts de zinc. La réaction des cystéines libres avec  $^1\text{O}_2$  est beaucoup plus rapide que lorsque qu'elles sont engagées dans la coordination au zinc pour former un doigt de zinc classique  $\beta\beta\alpha$ .

Afin de vérifier l'hypothèse de l'oxydation par l'oxygène singulet, la protéine a été synthétisée à nouveau puis directement stockée en boîte à gants après lyophilisation. Après plusieurs semaines, l'analyse par HPLC ne montre que peu de traces d'oxydation (Figure 100 bas droite), indiquant que le dioxygène joue un rôle clé dans l'oxydation. La stabilité a été étudiée dans des conditions aérobies, en comparant la protéine en présence ou absence de zinc, conservée à la lumière ou à l'obscurité (dans le réfrigérateur). Après une semaine, la protéine non métallée au zinc montre une très forte oxydation lorsqu'elle est exposée à la lumière (Figure 100 haut gauche). Dans les mêmes conditions, la protéine métallée au zinc exposée à la lumière présente de faibles traces d'oxydations (Figure 100 haut droite). Stockée au réfrigérateur à l'obscurité, la protéine métallée au zinc n'est que très peu oxydée (Figure 100 bas gauche). Ces observations confirment notre hypothèse de l'oxydation par l'oxygène singulet dont la protéine semble protégée une fois métallée au zinc et conservée à l'obscurité et au froid. L'oxygène singulet est probablement généré par l'état triplet de la lissamine, excitée par exposition à la lumière. Une fois synthétisée, la protéine doit être métallée au zinc dans du tampon HEPES pH 7,5 et conservée en boîte à gants, pour la maintenir sous forme réduite. Métallées au zinc, les protéines ZIF(lissamine) et ZIF(SiR-méthyl) peuvent être sorties de la boîte gants et être manipulées à l'air sans dommages pendant plusieurs heures.

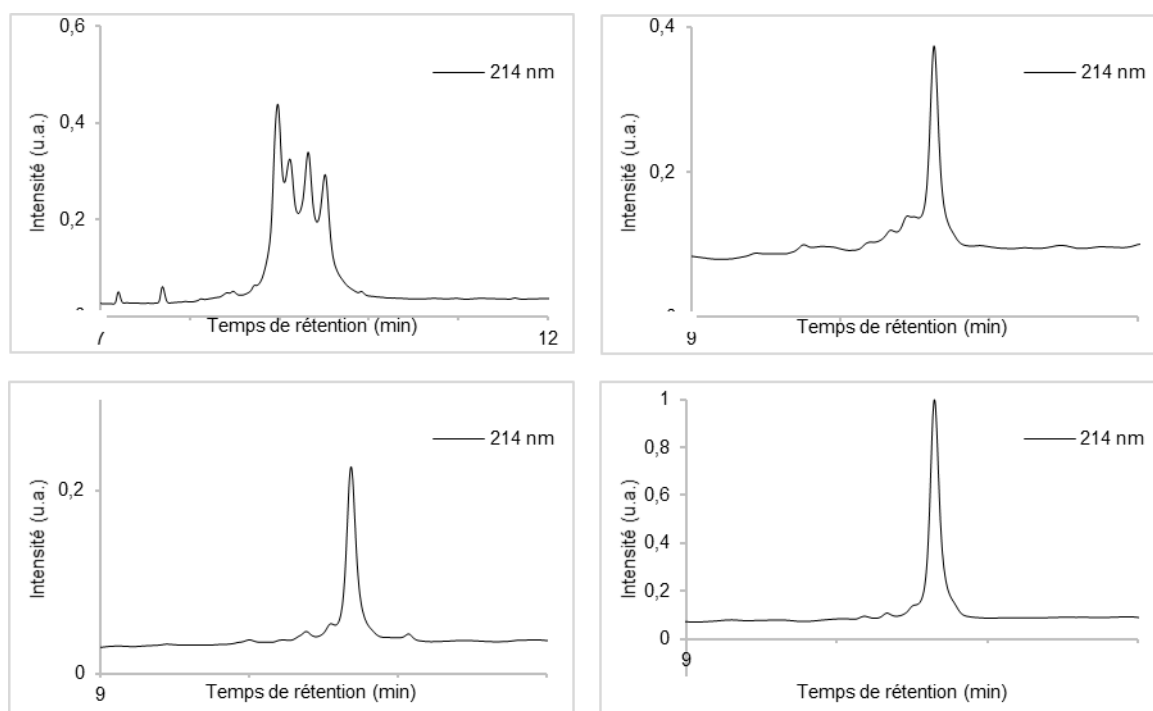


Figure 100: En haut à gauche, chromatogramme de la protéine ZIF(lissamine) stockée à la lumière ; En haut à droite, chromatogramme de la protéine ZIF(lissamine) métallée au zinc et stockée à la lumière ; En bas à gauche, chromatogramme de la protéine ZIF(lissamine) métallée au zinc et stockée au frigo ; En bas à droite, chromatogramme de la protéine ZIF(lissamine) stockée en conditions anaérobie

Le composé ZIF(SiR-méthyl) a présenté la même sensibilité à l'oxygène singulet, mais a aussi montré une autre particularité. Alors que la première réaction d'assemblage s'était déroulée correctement, lorsque nous avons voulu la répéter, nous avons vu que la solution perdait sa couleur bleue, due au chromophore SiR-méthyl, lors de la deuxième étape de ligation. Au final, la protéine assemblée ne présentait plus l'absorption caractéristique du chromophore SiR-méthyl malgré un spectre de masse identique à la protéine ZIF(SiR-méthyl) bleue. Le même phénomène de décoloration a pu être observé en diluant une solution de protéine ZIF(SiR-méthyl) dans l'eau (pH  $\approx$  3), initialement bleue, dans de l'HEPES pH 7,5. En quelques secondes, la couleur bleue avait disparu. Les composés de type rhodamine ou Si-rhodamine, colorés et fluorescents, peuvent réagir avec des nucléophiles (thiols ou autre) sur la position 9 du squelette xanthène pour donner un composé incolore non fluorescent selon la Figure 101. Cette réaction est réversible. Le groupe de Y. Urano a mis à profit cette réactivité pour concevoir une sonde fluorescente pour détecter le glutathion (GSH) avec une Si-Rhodamine<sup>[150]</sup>. Selon leurs travaux, le SiR-méthyl est un des composés les moins réactifs avec une constante de dissociation  $K_d$  de l'adduit SiR-méthyl/GSH de l'ordre de 100 mM. Nous avons pensé à une réaction de ce type pour notre composé. L'analyse ESI-MS de la protéine décolorée ne montrant pas d'addition d'un fragment (MPAA utilisé lors de l'assemblage par exemple), nous avons pensé à une

réaction intramoléculaire entre le SiR-méthyl et une des cystéines de la protéine. Cette réaction serait favorisée lorsque les cystéines se déprotonnent donc lorsque le pH est supérieur à 7. L'addition de zinc sur la protéine décolorée permet de récupérer la coloration bleue en quelques dizaines de secondes. Ceci s'explique par un déplacement de l'équilibre vers la formation du composé coloré par la coordination des cystéines au zinc.

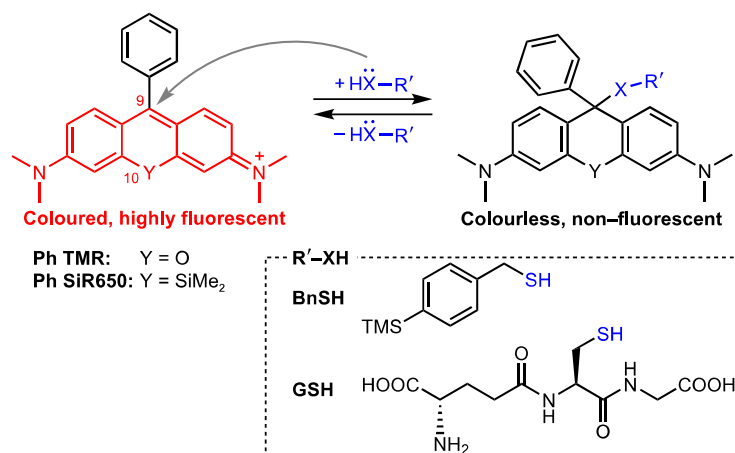


Figure 101: Réaction de la rhodamine avec des nucléophiles<sup>[150]</sup>

### 3) Métallation au zinc

En procédant de manière analogue au chapitre I, les quantités de protéines ont été déterminées par mesure de leurs absorbances dans un tampon HEPES 10 mM pH 7,5, avec pour la protéine ZIF(lissamine) un coefficient d'extinction molaire ( $\epsilon$ ) égal à 71 500 M<sup>-1</sup>cm<sup>-1</sup> à 568 nm<sup>[140]</sup>, et pour la protéine ZIF(SiR-méthyl) un coefficient d'extinction molaire ( $\epsilon$ ) égal à 100 000 M<sup>-1</sup>cm<sup>-1</sup> à 648 nm<sup>[151]</sup>. Une quantité de ZnCl<sub>2</sub> correspondant à un excès de 2,2 à 3eq par rapport aux protéines ont été ajoutés, puis les protéines notées ZIF<sup>Zn</sup>(lissamine) et ZIF<sup>Zn</sup>(SiR-méthyl) ont été stockées au congélateur.

## III Etude spectroscopique du système ZIF<sup>Zn</sup>(lissamine)/ZIF<sup>Zn</sup>(SiR-méthyl)

### A Etude du mélange ZIF<sup>Zn</sup>(lissamine)/ZIF<sup>Zn</sup>(SiR-méthyl)

Les protéines ZIF<sup>Zn</sup>(lissamine) et ZIF<sup>Zn</sup>(SiR-méthyl) ont été introduites dans un mélange HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM) en rapport 1 :1 à une concentration de 1  $\mu$ M. Pour rappel, la lissamine doit servir de donneur dans le système FRET et la SiR-méthyl d'accepteur. Nous avons choisi d'exciter l'échantillon à 540 nm, longueur d'onde où l'absorption de la lissamine est importante ( $\epsilon_{540}^{\text{lissamine}} = 25\,000\text{ M}^{-1}\text{ cm}^{-1}$ ) alors que celle de la SiR-méthyl est très faible ( $\epsilon_{540}^{\text{SiR-méthyl}} = 2\,700\text{ M}^{-1}\text{ cm}^{-1}$ ), afin d'exciter sélectivement

la lissamine. Le spectre d'émission obtenu est représenté sur la Figure 102 (Bas). On observe deux maxima d'émission à 590 nm et 675 nm correspondant à l'émission de la lissamine et de la SiR-méthyl. Les spectres d'excitation obtenus par détection à 600 nm et 675 nm sont en accord avec cette attribution puisqu'ils sont respectivement similaires aux spectres d'absorption de la lissamine et de la SiR-méthyl. On observe cependant une contribution de la lissamine au spectre d'excitation avec détection à 675 nm, qui peut être dû l'émission directe de la lissamine qui traîne jusqu'à 750 nm, ou à un transfert d'énergie de la lissamine à la SiR-méthyl. Il est intéressant de noter qu'en excitant quasi-sélectivement la lissamine à 540 nm, on récupère une émission importante de la SiR-méthyl : la déconvolution des spectres nous donne un ratio de 0,30 entre l'émission de la SiR-méthyl et celle de la lissamine. Ce résultat est trois fois supérieur au ratio de 0,09 attendu à partir des brillances ( $B = \epsilon \times \Phi$ ) à 540 nm pour la lissamine et la SiR-méthyl, qui sont respectivement de  $8250 \text{ M}^{-1} \text{ cm}^{-1}$  et  $810 \text{ M}^{-1} \text{ cm}^{-1}$ . Dans les conditions de concentrations utilisées ici, soit  $1 \mu\text{M}$  pour chaque protéine, il semble donc qu'un transfert d'énergie intermoléculaire lissamine→SiR-méthyl se produise. En diluant le mélange 10 fois ( $0,1 \mu\text{M}$  pour chaque protéine), on peut voir une diminution de l'émission à 675 (environ 13%), ce qui confirme cette hypothèse (Figure 103).

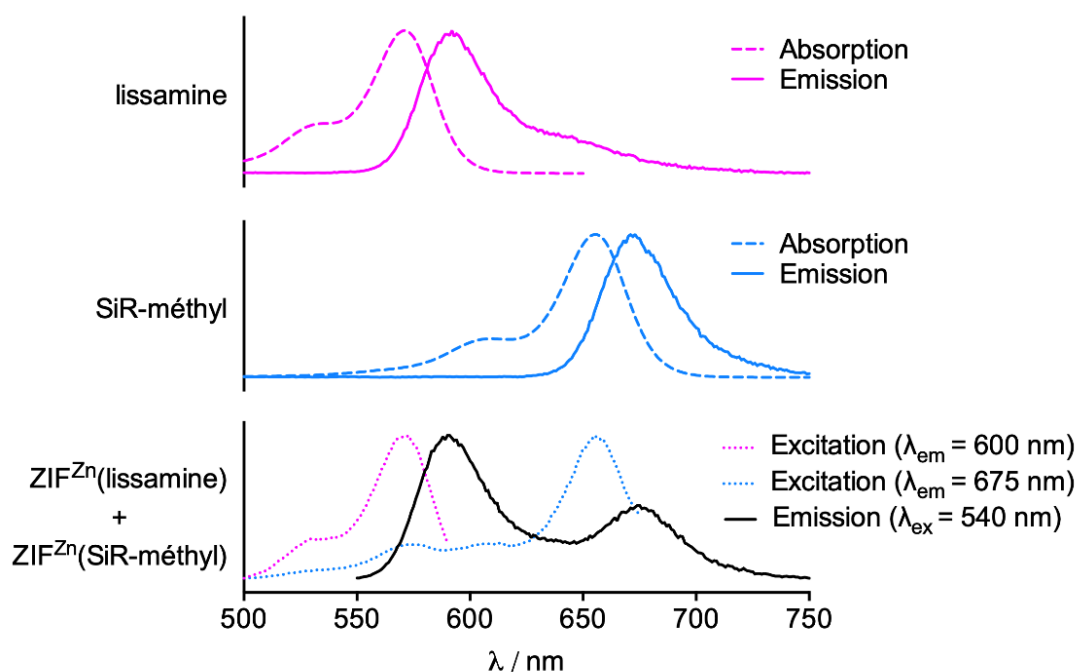


Figure 102: En haut, spectres d'émission et d'absorption de la lissamine ; Au centre, spectres d'émission et d'absorption de la SiR-méthyl ; En bas, spectres d'excitation (à  $\lambda_{em} = 600 \text{ nm}$  et  $\lambda_{em} = 675 \text{ nm}$ ) et d'émission (à  $\lambda_{ex} = 540 \text{ nm}$ ) du mélange  $\text{ZIF}^{2n}(\text{lissamine})$  et  $\text{ZIF}^{2n}(\text{SiR-méthyl})$  (1:1) dans un mélange HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM)

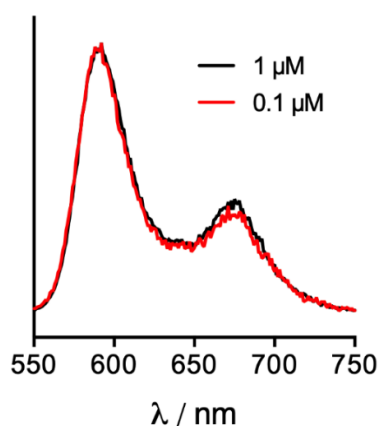


Figure 103: Spectres d'émission normalisés d'un mélange ZIFZn(lissamine)/ZIFZn(SiR-méthyl) (1:1) à 1  $\mu\text{M}$  (noir) et 0,1  $\mu\text{M}$  (rouge) dans un mélange HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM).

### B Etude du mélange ZIF<sup>Zn</sup>(lissamine)/ZIF<sup>Zn</sup>(SiR-méthyl) avec la séquence d'ADN cible

Le titrage par l'ADN a été réalisé en procédant comme pour la sonde ZIF<sup>Zn</sup>(Cs124|DOTA[Tb]), par incubation d'un mélange de ZIF<sup>Zn</sup>(liss) et ZIF<sup>Zn</sup>(SiR-méthyl) en rapport 1:1 à une concentration de 1  $\mu\text{M}$  avec différentes quantités de duplex d'ADN contenant la séquence cible 5'-TGGGCGCGCCCA-3' (0  $\mu\text{M}$  (0 éq.), 1  $\mu\text{M}$  (1 éq.) et 2  $\mu\text{M}$  (2 éq.)) dans un mélange HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM). Le duplex d'ADN utilisé est formé à partir du 38-mère 5'-GGTTGCAGTGGGCGCGCCACAGTACTTG-AACGT AACG-3' hybridé avec son brin complémentaire.

Les spectres d'émission de fluorescence mesurés à  $\lambda_{\text{ex}} = 540 \text{ nm}$  avec les différents équivalents en duplex d'ADN sont superposés dans la Figure 104 (B). L'émission de la lissamine à 590 nm diminue de 40 % lorsque 1 éq. de duplex d'ADN est ajouté. La bande à 675 nm, qui comprend à la fois l'émission de la SiR-méthyl et la queue de bande d'émission de la lissamine, reste quasi constante. Ceci indique que la proportion d'émission de la SiR-méthyl par rapport à la lissamine a augmenté en ajoutant l'ADN suggérant un FRET plus important dû à la formation du complexe ternaire ADN/ZIF<sup>Zn</sup>(lissamine)/ZIF<sup>Zn</sup>(SiR-méthyl). Ceci est clairement mis en évidence sur la Figure 104 (C) où les spectres d'émission ont été normalisés par rapport à l'émission de la lissamine à 590 nm. On peut estimer par déconvolution des spectres que le ratio de l'émission de la SiR-méthyl sur celle de la lissamine passe de 0,3 à 0,69, soit une émission de la SiR-méthyl qui fait plus que doubler lorsque le complexe ternaire se forme. De plus, alors que la proportion d'émission de la lissamine à 590 nm diminue, les spectres d'excitation de la bande à 675 nm correspondants à l'excitation de la SiR-méthyl montrent clairement une

augmentation de la contribution de la lissamine vers 570 nm (Figure 104 (A)). Ainsi, la lissamine sensibilise davantage la SiR-méthyl quand l'ADN est ajouté, confirmant que la formation du complexe ternaire s'accompagne d'un transfert d'énergie lissamine→SiR-méthyl plus important. Finalement, le spectre d'émission reste inchangé après addition d'un deuxième équivalent de duplex d'ADN, ce qui indique que le complexe ternaire est entièrement formé après ajout du premier équivalent de duplex. Ceci valide sans ambiguïté le principe de la détection du duplex palindromique ciblé par notre système de FRET basé sur des protéines à doigts de zinc.

Notons cependant qu'outre le complexe ternaire ADN/ZIF<sup>Zn</sup>(lissamine)/ZIF<sup>Zn</sup>(SiR-méthyl), deux autres complexes ternaires peuvent se former entre l'ADN et les homo-dimères (ZIF<sup>Zn</sup>(lissamine))<sub>2</sub> et (ZIF<sup>Zn</sup>(SiR-méthyl))<sub>2</sub>. Ces trois complexes doivent être dans des proportions 50:25:25. Ceci empêche d'aller plus loin dans la caractérisation du système, notamment la détermination de l'efficacité du transfert et des distances entre les fluorophores, qui nécessite de mesurer précisément les temps de vie du donneur (lissamine) en absence et en présence d'ADN.

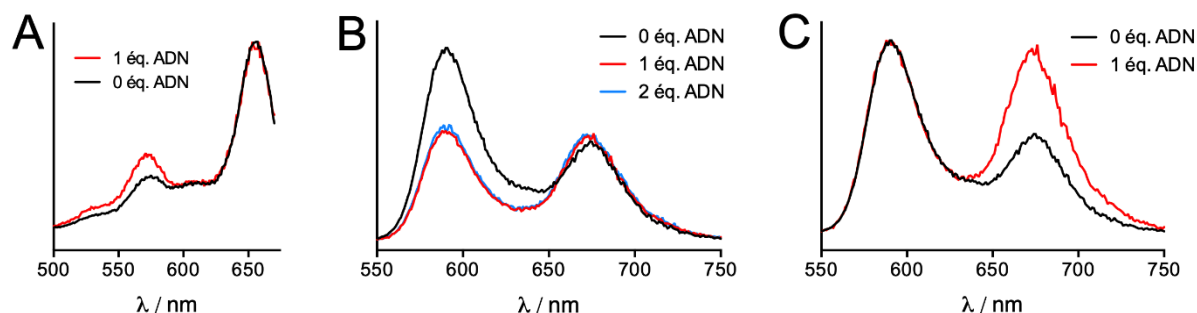


Figure 104: Spectres (A) d'excitation ( $\lambda_{em}=675$  nm) et (B) d'émission ( $\lambda_{ex}=540$  nm) de fluorescence sans et avec 1 ou 2 éq. de duplex d'ADN. (C) Mêmes spectres d'émission normalisés par rapport à la lissamine.

## IV Conclusion

La formation d'un hétéro-dimère  $\text{ZIF}^{\text{Zn}}(\text{lissamine})/\text{ZIF}^{\text{Zn}}(\text{SiR-méthyl})$  complexé au duplex d'ADN contenant la séquence cible 5'-TGGGCGCGCCCA-3' a été observée par luminescence, confirmant la possibilité de concevoir un système de FRET efficace à des distances de plusieurs dizaines d'angströms avec des fluorophores organiques pour détecter l'ADN double brin. Le système développé a permis une augmentation du ratio d'émission du fluorophore accepteur de 0,3 à 0,69, contre 0,83 à 1,2 pour le système de FRET développé par l'équipe d'Hiroshi Ueda<sup>[47]</sup>. Cependant, aux concentrations utilisées (1  $\mu\text{M}$ ), le FRET s'établit en absence d'ADN. Notons que l'ordre choisi dans ce manuscrit pour la présentation des résultats n'est pas chronologique. Les protéines pures n'ont pu être obtenues qu'à la fin de la thèse et nous n'avons pas eu le temps de pousser aussi loin qu'il le faudrait certaines analyses. Notamment, la caractérisation de l'interaction ADN/protéine par des techniques de biologie moléculaire (retard sur gel et foot-printing) n'a pas été réalisée et les études de luminescence pourraient être complétées. Il nous reste encore à mieux caractériser ce système. Il faudrait confirmer la stœchiométrie 1:1:1 du complexe par des titrages plus fins avec des quantités de duplex sub-stœchiométriques, évaluer l'affinité de la sonde pour le duplex cible et vérifier qu'un duplex d'ADN qui ne contient pas la séquence cible n'est pas reconnu.

Néanmoins, ces premiers résultats semblent établir la preuve de concept de la détection d'ADN double brin par fluorescence à l'aide de la protéine Zif12 dont quelques acides aminés ont été modifiés.

Les larges spectres d'absorption et d'émission des deux fluorophores utilisés ainsi que leurs faibles déplacements de Stokes sont des inconvénients intrinsèques au FRET entre fluorophores organiques. Cela diminue la sélectivité de l'excitation et fait que l'interaction avec la cible n'est pas mise en évidence de manière optimale. Les bandes d'émission fines des lanthanides permettraient de gagner en sélectivité spectrale et de réaliser plus facilement de la détection résolue en temps pour améliorer le rapport signal/bruit, ce qui peut être critique quand on travaille à faible concentration en milieu biologique. C'est pourquoi, à ce stade, il semble justifié de continuer le projet en incorporant un complexe de lanthanide comme donneur d'énergie dans le système. Pour cela, il faut synthétiser des complexes de lanthanides avec des propriétés photophysiques intéressantes pour notre système et conjugables à des biomolécules. C'est ce que nous décrivons dans le chapitre suivant.





# Chapitre III : Synthèse de complexes de lanthanides



## Chapitre III : Synthèse de complexes de lanthanides

Pour établir un FRET efficace d'un complexe de lanthanide vers un fluorophore, le complexe de lanthanide doit être fortement luminescent, c'est-à-dire avoir un rendement quantique le plus élevé possible et si possible une absorption intense. Une large gamme de complexes de lanthanides possédant une antenne introduite au sein du complexe ont été développés depuis la découverte de l'effet d'antenne en 1942 par S. Weissman<sup>[152]</sup>.

Les complexes d'euprium et de terbium présentent une émission dans le visible compatible avec de nombreux fluorophores organiques tels que la lissamine ou la SiR-méthyl que nous avons utilisés pour notre première sonde de seconde génération. Ce sont aussi les deux lanthanides pour lesquels il est possible de concevoir aisément des complexes avec de bons rendements quantiques et de longs temps de vie de luminescence.

Les complexes à intégrer sur le système de sonde de seconde génération doivent remplir certains critères de sélection. Ils doivent :

- posséder un rendement quantique le plus fort possible,
- être stables thermodynamiquement, et si possible cinétiquement,
- posséder une fonction de bioconjugaison,
- être solubles dans l'eau, et
- être facilement synthétisables.

Pour ce dernier point, il nous semblait intéressant de limiter l'effort synthétique en réfléchissant à une stratégie de synthèse divergente pour les ligands et leurs complexes.

En plus de tous ces critères, nous en avons ajouté un dernier qui est d'avoir un complexe permettant une absorption bi-photonique<sup>[153]</sup>. Cette propriété n'est pas requise dans l'immédiat pour notre projet et pour la preuve de concept de détection d'ADN double brin séquence spécifique que nous souhaitons apporter. Cependant, elle est utile pour d'autres projets développés dans l'équipe, visant à élaborer des sondes pour la détection par microscopie confocale dans des cellules. Il pouvait donc être intéressant de disposer de tels complexes.

### I Introduction

#### 1) Complexes bioconjuguables

La conjugaison d'un complexe de lanthanide sur une biomolécule requiert l'introduction d'une fonction de couplage lors de la synthèse du ligand. Soit le ligand est couplé puis métallé par la suite, soit le complexe est directement couplé. La fonction de couplage doit donc être chimiquement compatible avec la biomolécule, mais aussi avec toutes les étapes de synthèse réalisées pour le couplage et en post couplage du complexe (ou du ligand).

Deux grandes classes de fonctions de couplages sont utilisées pour la bioconjugaison de complexes de lanthanides. La première classe est composée de fonctions réagissant avec les amines (Figure 105). Par exemple, le complexe noté ici  $[Tb(L-SCN)]$  décrit par K. Mitrinen<sup>[154]</sup> en 1995 comporte une fonction isothiocyanate permettant après réaction avec une amine la formation d'une thio-urée. L'introduction d'une fonction chlorosulfonyl par réaction d'acide chlorosulfonique sur un cycle aromatique, comme utilisé sur le ligand BCPDA décrit par R. Evangelista<sup>[155]</sup>, permet l'obtention d'une fonction réactive avec les amines, le principal défaut de cette technique étant le manque de sélectivité sur la position de fonctionnalisation du cycle aromatique lors de l'introduction de la fonction chlorosulfonyl. Des fonctions telles que la dichlorotriazines ont aussi été décrites comme fonction de bioconjugaison, comme sur le complexe noté ici  $[Tb(L-Cl_2\text{triazine})]$  de V. Mukkala<sup>[156]</sup> conjugable sur une amine. Les esters de N-hydroxysuccinimidyl (NHS) permettent quant à eux le couplage du complexe par réaction avec une amine conduisant à la formation d'un amide, comme avec le complexe  $[TbL-(H_2O)]$  de L. Charbonnière<sup>[157]</sup>.

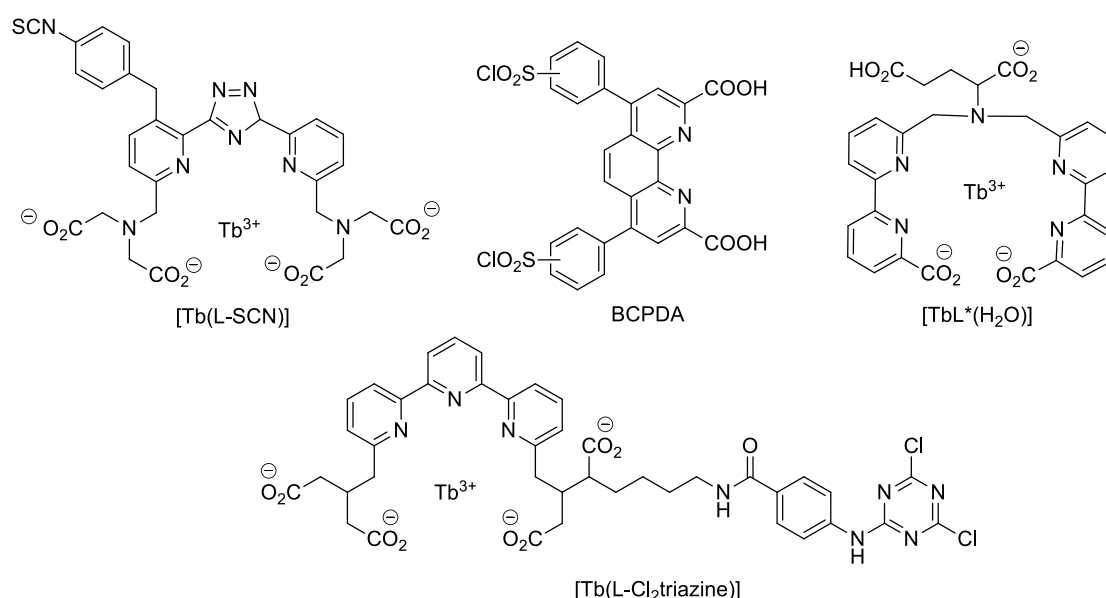


Figure 105: Exemples de ligands et de complexes de lanthanides bioconjugables

La seconde classe de fonction de bioconjugaisons est composée de groupements réagissant avec les thiols. Ces fonctions de couplage n'étant pas utilisables pour le projet à cause des quatre cystéines de la protéine qui doivent coordonner le zinc, nous ne nous y intéresserons pas.

Les limites des complexes bioconjuguables présentés et de ceux décrits dans l'ensemble de la littérature sont en général l'absence d'un des critères de sélection tels que la solubilité en milieu aqueux ou une stabilité cinétique et thermodynamique insuffisante. Par exemple, pour ce dernier point, les polyaminocarboxylates linéaires sont connus pour être moins stables que leurs analogues cycliques<sup>[158]</sup>.

Le développement de complexes de lanthanides correspondant aux critères de sélection est donc nécessaire. Dans l'optique d'une synthèse divergente de différents complexes, des modifications minimales entre les antennes sensibilisant différents lanthanides sont recherchées.

## 2) Complexes d'euprimum à absorption bi-photonique

En 1996, H. Takalo a décrit une antenne à transfert de charge de type pyridine-alcyne-aryl possédant des propriétés d'absorption bi-photonique et permettant la sensibilisation de l'euprimum<sup>[159]</sup> (Figure 106). Depuis, de nombreux dérivés de cette antenne ont été utilisés afin de réaliser l'imagerie confocale cellulaire<sup>[160],[161],[162],[163]</sup>.

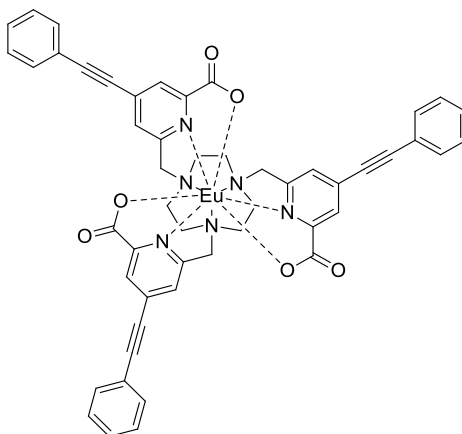


Figure 106: Complexe d'euprimum décrit par H. Takalo

Parmi les complexes d'euprimum décrits dans la littérature, certains des complexes d'O. Maury semblent intéressants, car remplissant tous les critères de sélection excepté la présence d'une fonction de conjugaison.

Le complexe  $[Eu \cdot L^{1b}]$  (Figure 107 gauche) composé d'un macrocycle TACN (1, 4, 7, Triazacyclononane) et de trois antennes de type pyridine-alcyne-aryl comportant en position para un groupement *O-alkyl* possède un rendement quantique de 25%<sup>[164]</sup> dans un mélange

H<sub>2</sub>O/MeOH 3:1. Son analogue [Eu•L<sup>3</sup>(COO)] (Figure 107 droite) comportant une fonction sulfobetaine en position para de l'antenne possède un rendement quantique de 18%<sup>[165]</sup> dans l'eau. Ils présentent tous les deux des propriétés d'absorption bi-photonique, vers 700 nm. Le groupement pyridine-alcyne-aryl est un chromophore de type « push-pull » avec le groupement O-alkyl électro-donneur et la pyridine électro-acceptrice. Il sert d'antenne pour l'euprium. Ce type d'antenne à transfert de charge possède des propriétés d'absorption mono et bi-photonique intéressantes. Les deux complexes ne se différenciant que par la fonction *O-alkyl*, on observe que l'on peut changer le substituant de la fonction éther sans affecter notablement le rendement quantique.

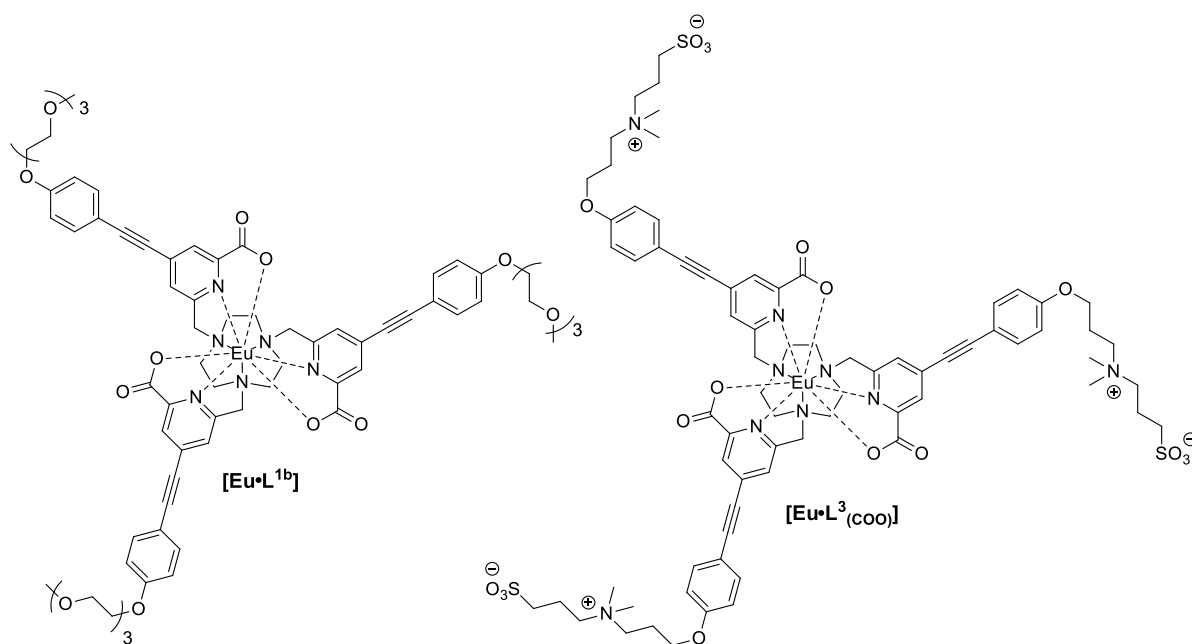


Figure 107: Complexes d'euprium sur base d'un ligand TACN portant des groupements O-donneur décrits par

Un autre complexe d'euprium décrit par O. Maury noté [Na][Eu(L<sup>2</sup>)], composé d'un ligand DO3A et d'une antenne de type pyridine-alcyne-aryl comportant en position para un groupement *S-alkyl* (Figure 108) possède un rendement quantique de 28%<sup>[166]</sup> dans l'eau.

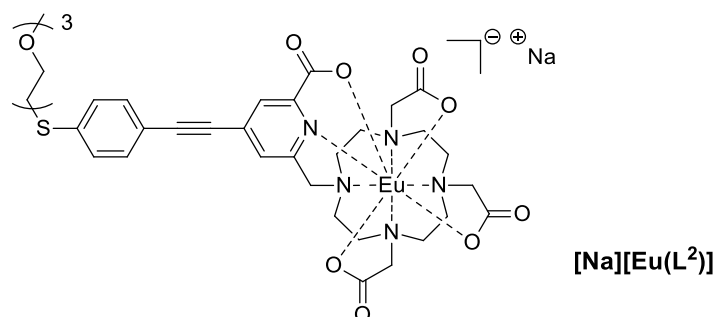


Figure 108: Structure du complexe [Na][Eu(L<sup>2</sup>)]

### 3) Complexes de terbium à absorption bi-photonique

En 2005, V. Werts a décrit une antenne à transfert de charge de type pyridine-aryl permettant de sensibiliser le terbium par absorption bi-photonique<sup>[167]</sup>. D. Parker et O. Maury ont indépendamment introduit cette antenne sur des ligands DOTA et TACN avec des rendements quantiques en phase aqueuse de 5,4%<sup>[168]</sup> pour le complexe TbL<sup>1</sup> (Figure 109 droite), et 30%<sup>[169],[170]</sup> pour le complexe no3A<sup>3</sup>paTb (Figure 109 gauche).

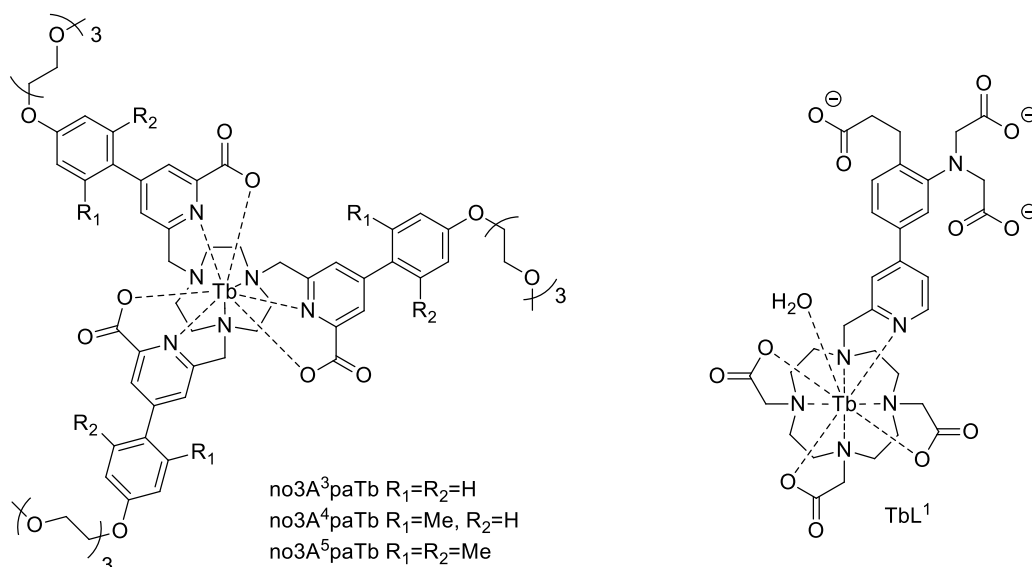


Figure 109: Complexes de terbium décrits par O. Maury (à gauche) et D. Parker (à droite)

Le complexe no3A<sup>3</sup>paTb a été décliné en deux autres complexes no3A<sup>4</sup>paTb et no3A<sup>5</sup>paTb, pour lesquels l'introduction de groupements méthyles en position ortho du phényl permet d'éviter la désexcitation de la molécule en empêchant sa rotation<sup>[170]</sup>. Le complexe (no3A<sup>4</sup>paTb) possède dans l'eau un rendement quantique de 74% tandis que celui du complexe (no3A<sup>5</sup>paTb) est à 65%.

### 4) Conclusion

Les complexes décrits par O. Maury semblent intéressants car possédant de haut rendements quantiques ainsi qu'une stabilité thermodynamique apportée par l'utilisation de ligands TACN ou DO3A. Les trois antennes de ces complexes, notés picAr<sub>1</sub>, picAr<sub>2</sub> et picAr<sub>3</sub> pour la suite de ce manuscrit (Figure 110) doivent donc être introduites au sein de complexes possédant une fonction de bioconjugaison.

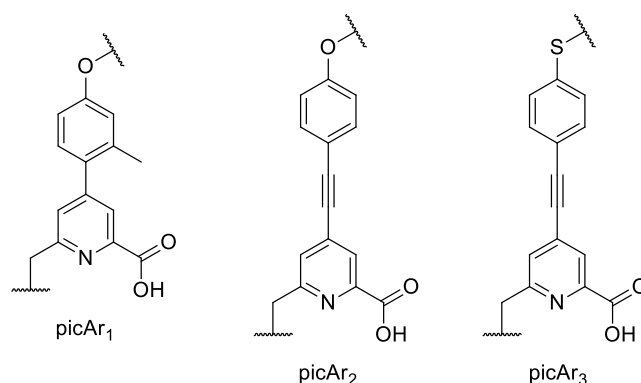


Figure 110: Structures des différentes antennes à introduire dans des complexes de lanthanides

La possibilité de synthétiser deux ligands contenant des antennes de type pyridine-alcyne-aryl ou pyridine-aryl en utilisant une stratégie d'assemblage divergente a été décrite par O. Maury, permettant la synthèse d'une partie commune facilitant ainsi la synthèse (Figure 111).

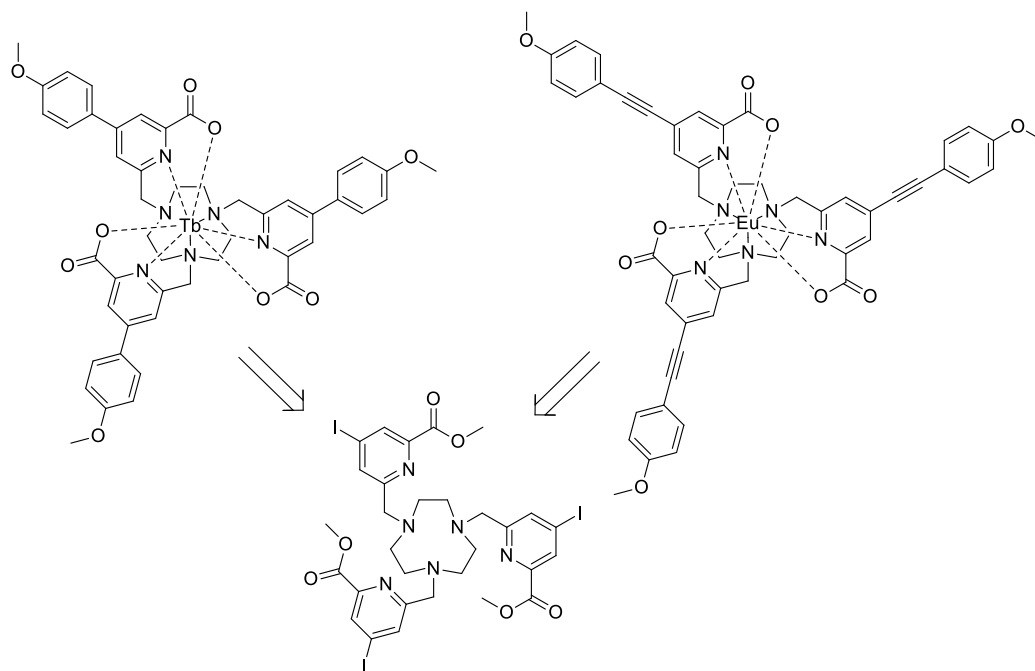


Figure 111: Stratégie de synthèse divergente utilisée par O. Maury permettant la synthèse soit de complexes d'euprimum, soit de complexes de terbium par des réactions de couplage de Sonogashira (pour le complexe d'euprimum), ou de Suzuki-Miyaura (pour le complexe de terbium)

## II Design des complexes de lanthanides, de leurs ligands et stratégie de synthèse

Le complexe de lanthanide  $[\text{Na}][\text{Eu}(\text{L}^2)]$  (cf Figure 108) possède un squelette DO3Apicolinate avec une seule antenne (DO3A = acide 1,4,7,10-Tétraazacyclododécane-1,4,7-triacétique). Par ailleurs, il a été remarqué plus tôt la possibilité de modifier le substituant du groupe électro-

donneur de l'antenne pyridine-alcyne-aryl sans affecter le rendement quantique. Une fonction de conjugaison peut alors y être introduite en remplacement de la chaîne PEG.

### 1) Design des complexes

Nos cibles sont les trois complexes de lanthanides basés sur le squelette DO3Apicolinate, chacun portant une des trois antennes picolinate décrites dans la Figure 110. Nous conjuguerons ces complexes aux protéines par le groupement électro-donneur (Figure 112).

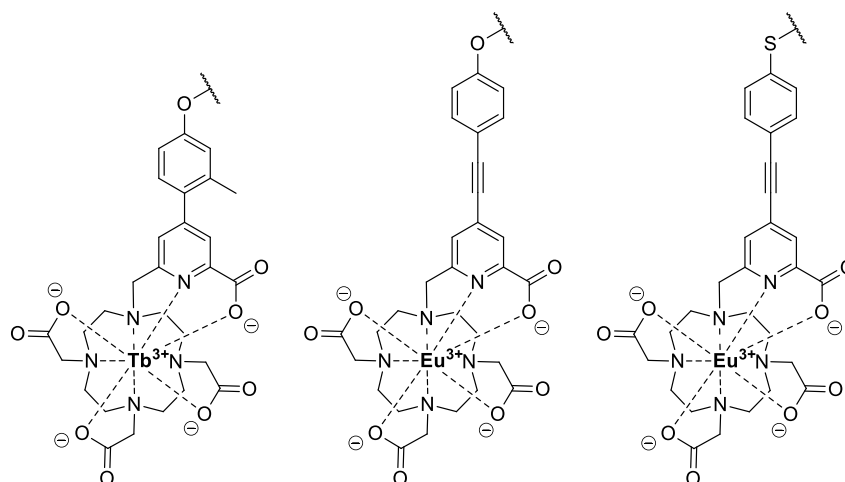


Figure 112: Structures des trois complexes de lanthanides à obtenir, basés sur le complexe  $[Na][Eu(L^2)]$

Deux stratégies peuvent être envisagées pour l'obtention d'un conjugué peptide/complexe de lanthanide : soit l'introduction du complexe (ou du ligand sous forme protégée) est réalisée lors de la synthèse peptidique, sur support ; soit une post-fonctionnalisation du peptide purifié est réalisée par couplage du complexe (ou du ligand) grâce à une réaction chimiosélective dans l'eau.

#### A Couplage par chimie click

La stratégie de couplage par réaction chimiosélective permet la synthèse d'une unique protéine comportant une fonction de couplage. L'introduction du chromophore peut ainsi être réalisée en dernière étape de la synthèse des sondes sur la protéine assemblée, facilitant ainsi la synthèse des peptides (Figure 113).

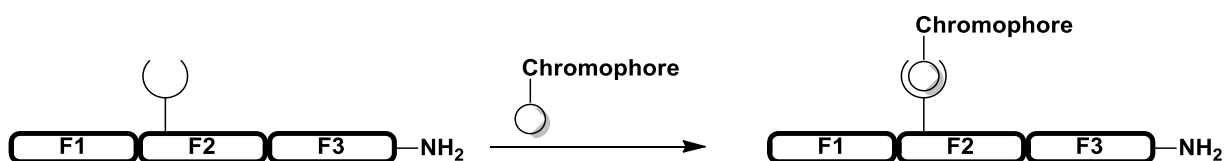


Figure 113: Etape d'introduction du chromophore une fois la protéine formée

Quatre réactions chimiosélectives sont largement utilisées en synthèse peptidique (Figure 114):

- la cycloaddition de Huisgen, catalysée au cuivre permet la formation d'un 1,2,3 triazole par réaction d'un azoture et d'un alcyne vrai<sup>[171]</sup>,
- le couplage thiol-maléimide, qui après une réaction d'addition de Michael entre un thiol et un groupement maléimide, permet la formation d'une imide saturée<sup>[172]</sup>,
- la ligation oxime, impliquant une oxyamine et un carbonyle, qui après réaction de condensation conduit à la formation d'une fonction éther d'oxime<sup>[173]</sup>,
- La ligation tétrazine, basée sur une réaction de Diels-Alder entre une tétrazine et un alcène pauvre en électrons<sup>[174]</sup>.

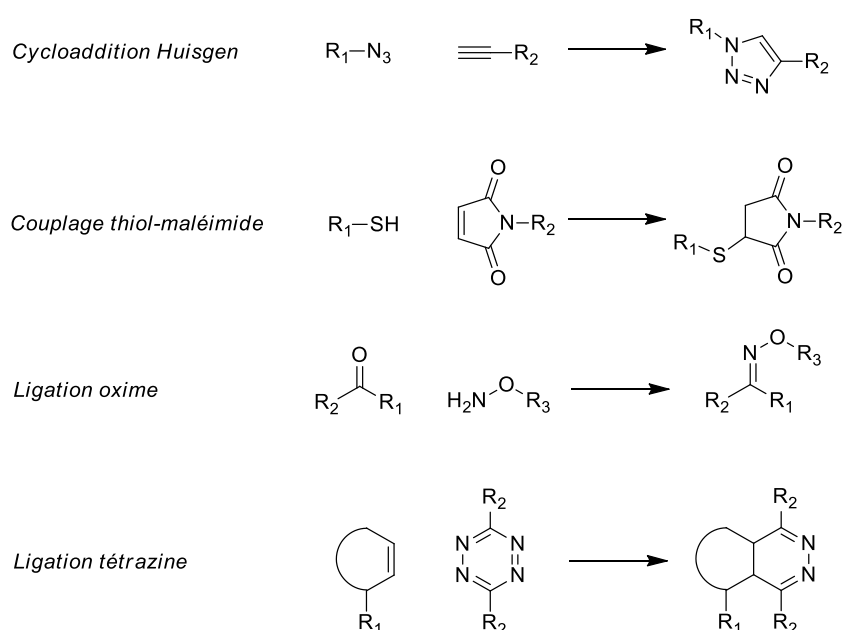


Figure 114: Exemples de réaction de chimie click

La cycloaddition de Huisgen nécessite l'utilisation de cuivre(I), ce qui la rend très difficilement maîtrisable avec les peptides comportant des cystéines, histidines et méthionine comme c'est le cas pour la protéine ZIF que nous utilisons. Pour le couplage thiol-maléimide, du fait de la présence de quatre cystéines pour complexer le zinc dans la protéine ZIF, il apparaît impossible d'introduire une cinquième cystéine et de la faire réagir sélectivement. La ligation oxime a déjà été testée au laboratoire par Céline Cepeda, et s'est avérée incompatible avec le groupement SEA. La dernière méthode de ligation tétrazine, elle, n'a jamais été décrite comme étant compatible avec la ligation native (SEA ou autre), et requiert donc de nombreux tests et optimisations. En effet, la tétrazine n'est pas compatible avec la SPPS en chimie Fmoc/*t*Bu mais l'introduction d'une oléfine au cours de la synthèse peptidique l'est. Sa compatibilité avec les conditions de ligation native reste à démontrer.

### B Couplage lors de la synthèse peptidique

A cause de l'impossibilité ou la difficulté de réaliser un couplage par réaction chimiosélective, le complexe doit être introduit lorsque le peptide est toujours supporté sur résine. La stratégie la plus simple à mettre en œuvre est celle d'un couplage dans des conditions de couplage peptidique classiques. Ainsi, la molécule (ligand ou complexe) à conjuguer peut par exemple comporter une fonction acide carboxylique ou sulfonique (activable) pour réagir avec une amine libre déprotégée sélectivement sur le peptide (obtenue par exemple à partir d'une Lys(Alloc) déprotégée sélectivement au  $\text{Pd}^0$ ). Nous avons retenu l'acide carboxylique comme fonction à introduire au niveau du groupement *O* ou *S-alkyl* électro-donneur et nous avons choisi un espaceur  $\text{CH}_2$ . Après réaction avec le peptide, un lien amide stable sera obtenu (Figure 115).

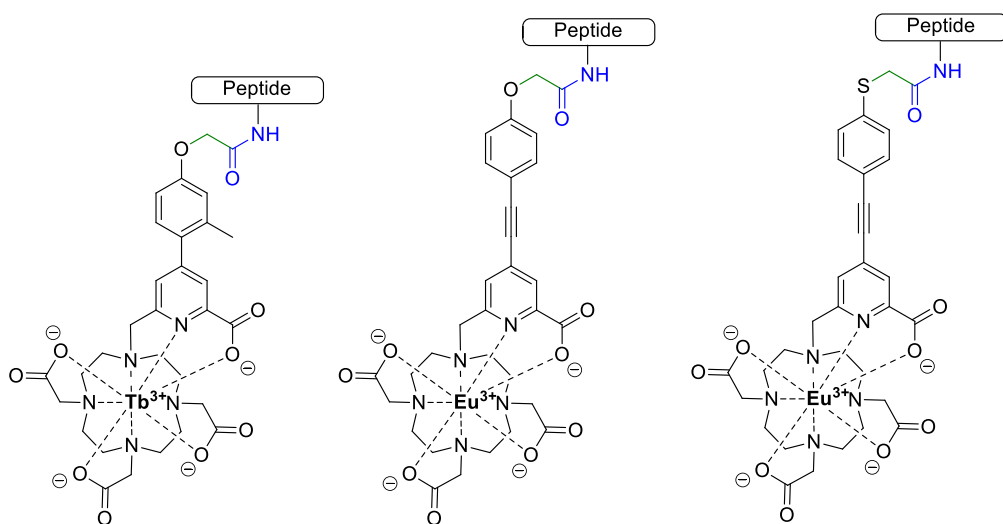


Figure 115: Structure des différents complexes attendus, couplés à un peptide (en bleu, lien amide liant le complexe au peptide ; en vert, espaceur  $\text{CH}_2$ )

Il est à noter que d'autres stratégies sont possibles, comme l'utilisation d'une amine gréffable sur un glutamate ou un aspartate, ou alors l'utilisation d'un acide aminé non-naturel dont la chaîne latérale serait le complexe. Cette dernière option est incertaine du fait de l'encombrement d'un tel acide aminé (de précédents tests avec un acide aminé Lys(DOTA(*t*Bu)<sub>3</sub>)) se sont révélés infructueux<sup>[175]</sup>).

### C Stratégies de couplage

Deux stratégies de couplage peuvent être envisagées pour la formation d'un lien amide entre le complexe et le peptide. La première stratégie consiste à utiliser un pro-ligand ayant des groupements protecteurs (GP) sur l'ensemble des acides carboxyliques du squelette

DO3Apicolinate mais pas sur celui du groupement électro-donneur de l'antenne (Figure 116 A). Une fois le lien amide formé, le peptide est décroché de sa résine puis les groupements GP sont éliminés et le ligand métallé au lanthanide. L'idéal dans ces conditions est la déprotection des groupements GP lors de l'étape de décrochage du peptide. Le groupement protecteur de choix des acides carboxyliques dans ces conditions (TFA + pièges) serait un groupement *t*Bu. Cependant, il a été observé au laboratoire la difficulté d'éliminer ces groupements protecteurs avec les chélates de type DO3A picolinate. Il a été choisi de protéger les acides carboxyliques sous forme d'esters méthyliques (GP=Me), facilement hydrolysables en conditions basiques.

La seconde stratégie est le couplage direct du complexe sur une amine libre du peptide. Les quatre acides carboxyliques du complexe coordonnés au lanthanide devraient être moins réactifs que l'acide du groupement électro-donneur de l'antenne, permettant une activation sélective. Le lanthanide jouerait ainsi le rôle de groupement protecteur (Figure 116 B). Cette stratégie a déjà été utilisée par le groupe de L. Charbonnière<sup>[176]</sup>.

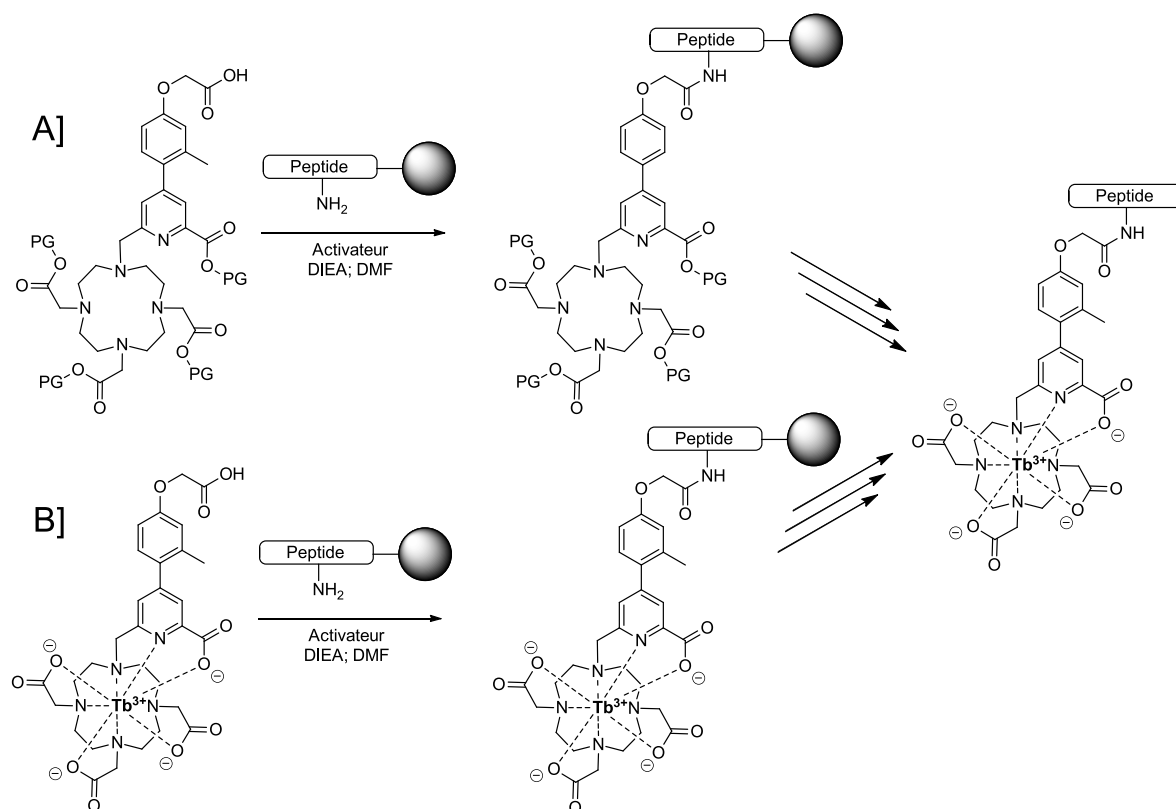


Figure 116: Stratégies de couplage envisagées avec pour exemple le complexe de terbium. A] Couplage du ligand. B] Couplage du complexe de lanthanide.

## 2) Stratégie de synthèse

La synthèse de deux ligands différents est requise suivant la stratégie de couplage choisie. Le couplage par pro-ligand requiert l'introduction d'un groupement protecteur sur l'acide carboxylique de la partie électro-donneur orthogonal aux autres présents sur le reste du ligand. Ces derniers étant des groupements méthyles basico-labiles, l'introduction d'un groupement *t*Bu acido-labile a été choisi (Figure 117).

Pour le couplage direct du complexe, un groupement protecteur méthyle a été choisi afin d'obtenir le ligand libre en une unique étape d'hydrolyse (Figure 117).

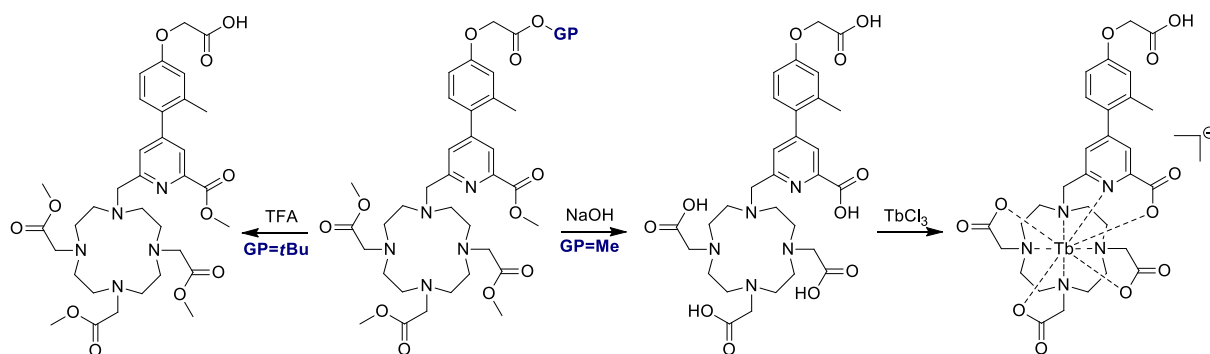


Figure 117: Groupements protecteurs à introduire sur le ligand pour les deux stratégies de couplage

Ces molécules ont en commun le squelette DO3Apicolinate protégé, noté **DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>** et seul le groupement protecteur présent sur l'acide carboxylique du groupement électro-donneur de l'antenne diffère. Six composés différents (trois antennes et deux voies de couplages) peuvent ainsi facilement être synthétisés en utilisant une synthèse divergente sur l'entité **DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>** (Figure 118).

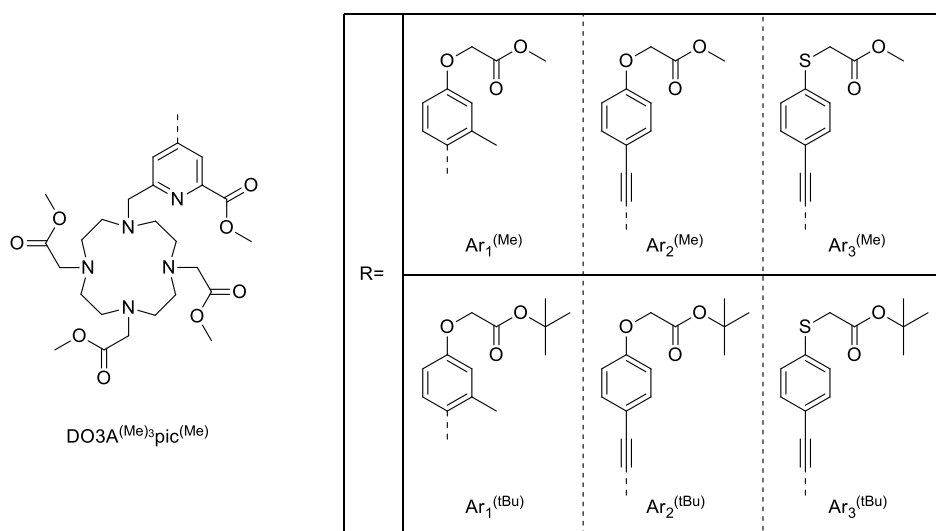


Figure 118: Structures des différents ligands à synthétiser

En suivant l'exemple d'O. Maury (Figure 111), des réactions de couplages de Suzuki-Miyaura ou Sonogashira entre le chélate **DO3A<sup>(Me)3pic<sup>(Me)</sup>I</sup>** (Figure 119) et différents composés aryl ou alcyne-aryl facilement synthétisables (notés **1b**, **1'b**, **2b**, **2'b**, **3b**, **3'b** ; Figure 120) doivent nous donner accès aux composés visés dans la Figure 118.

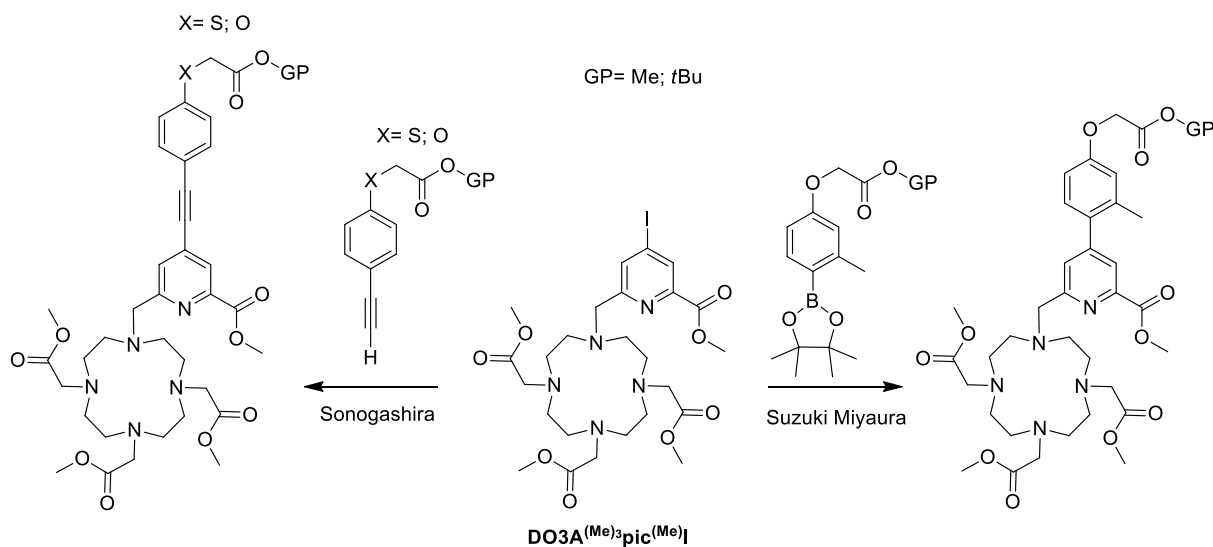


Figure 119: Stratégie de synthèse divergente choisie en partant d'un composé commun noté **DO3A<sup>(Me)3pic<sup>(Me)</sup>I</sup>**

| Suzuki-Miyaura    | Sonogashira       |                   |
|-------------------|-------------------|-------------------|
| <p><b>1b</b></p>  | <p><b>2b</b></p>  | <p><b>3b</b></p>  |
| <p><b>1'b</b></p> | <p><b>2'b</b></p> | <p><b>3'b</b></p> |

Figure 120: Composés aryl et alcyne-aryl à synthétiser

### III Synthèse des ligands

#### 1) Synthèse du macrocycle $\text{DO3A}^{(\text{Me})3}\text{pic}^{(\text{Me})}\text{I}$

Comme illustré sur la Figure 121, la synthèse de la partie commune notée  $\text{DO3A}^{(\text{Me})3}\text{pic}^{(\text{Me})}\text{I}\cdot\text{NaBr}$  est réalisée en 6 étapes, comprenant une dernière étape de synthèse convergente impliquant les composés  $\text{DO3A}^{(\text{Me})3}$  et **7**.

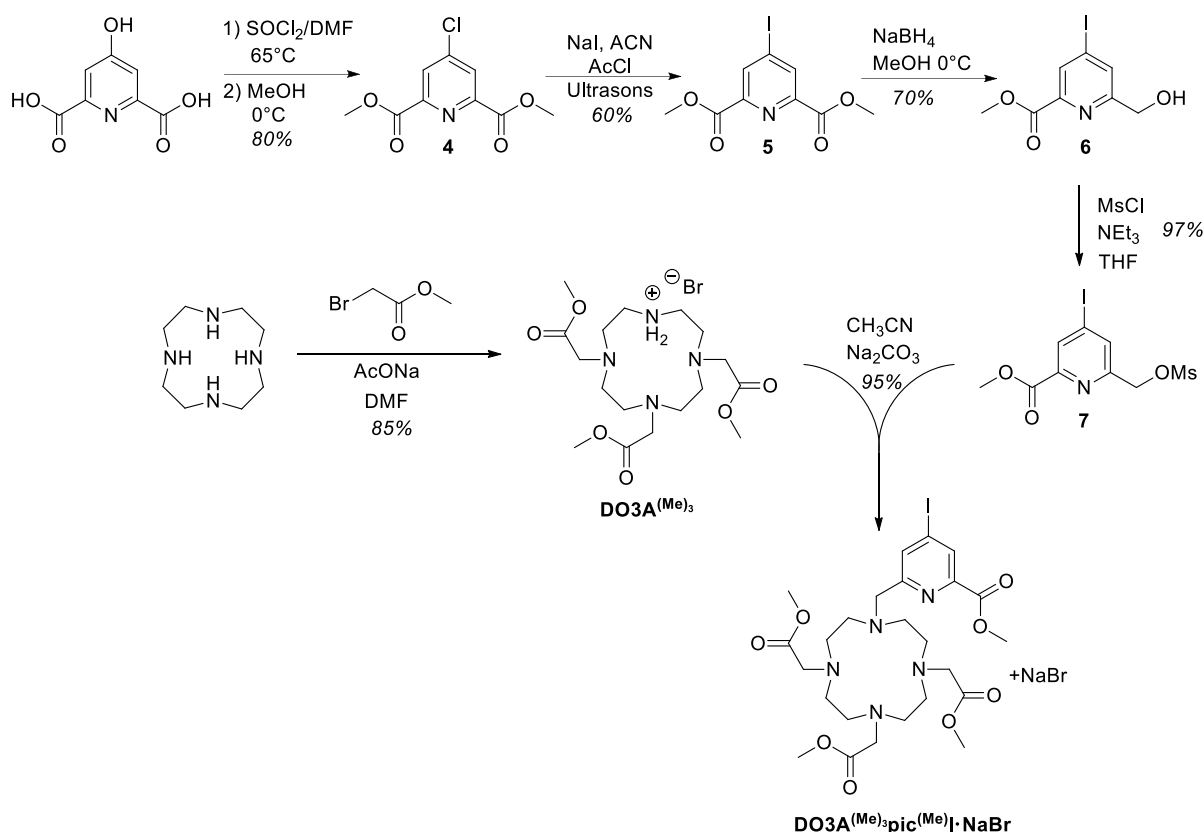


Figure 121: Schéma de synthèse de la partie commune  $\text{DO3A}^{(\text{Me})3}\text{pic}^{(\text{Me})}\text{I}\cdot\text{NaBr}$

#### A Synthèse du composé **7**

Le composé **7** a été obtenu en 4 étapes déjà décrites dans la littérature<sup>[177]</sup>. Une réaction de chloration de l'acide chélidamique par réaction de Vilsmeier-Haack suivie d'une méthanolyse (acide) conduit au diester **4** comportant un atome de chlore en position-4 de la pyridine. Dans une seconde étape, l'atome de chlore est substitué par de l'iode par une réaction de substitution nucléophile aromatique en présence d'excès de chlorure d'acétyle et d'iodure de sodium dans l'acétonitrile sous ultrasons, pour donner le composé **5**. Le composé **6** dissymétrique a été obtenu par une étape de mono-réduction au borohydrure de sodium dans un mélange DCM/MeOH 2:3. Finalement, l'alcool obtenu a été activé par réaction avec du chlorure de

mésyle, conduisant au composé 7. Les étapes de synthèses ont été optimisées, notamment concernant la purification des produits, où désormais seules deux étapes de purifications sont nécessaires. Une première par cristallisation par diffusion du composé 5, puis une seconde par chromatographie sur gel de silice pour le composé 6.

### *B Synthèse du DO3A<sup>(Me)3</sup>*

Le composé **DO3A<sup>(Me)3</sup>** a été synthétisé en suivant le protocole décrit par B. Jagadish<sup>[178]</sup>. Trois équivalents de bromure d'acétyle ont été ajoutés sur du cyclène dissout dans du DMF en présence d'acétate de sodium. La formation de produit tétra-substitué (en faible quantité) et la présence d'acétate de sodium ont requis l'optimisation de l'étape de purification. En dissolvant le brut de réaction obtenu dans de l'acétate d'éthyle, le produit tétra-substitué a pu être éliminé sous forme de précipité par centrifugation. L'acétate de sodium a pu quant à lui être éliminé par l'étape de purification par chromatographie sur gel de silice décrite dans le mode opératoire original.

### *C Couplage final*

La réaction du **DO3A<sup>(Me)3</sup>** sur le composé 7 conduisant au **DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>I•NaBr** a été réalisée dans l'acétonitrile en présence de carbonate de sodium. La purification par chromatographie flash de ce produit a été testée dans de multiples conditions (éluants, phases stationnaires), mais semble impossible dû à l'affinité du composé pour la phase stationnaire silice ou alumine. Finalement, une étape de précipitation dans l'acétate d'éthyle a permis d'éliminer l'excès de **DO3A<sup>(Me)3</sup>** présent dans le brut réactionnel. Le **DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>I•NaBr** a ainsi pu être obtenu avec un rendement global de 26% sur les 6 étapes de synthèse. Des monocristaux de ce composé ont été obtenus par évaporation lente d'une solution dans CDCl<sub>3</sub>. La structure RX (Figure 122) montre le composé **DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>I** complexé à un ion Na<sup>+</sup> par cinq azotes (les quatre azotes du cyclène et celui du picolinate) et trois oxygènes (deux fonctions carbonyles des bras esters et le groupe méthoxy du picolinate). Un anion bromure est présent pour neutraliser la charge.

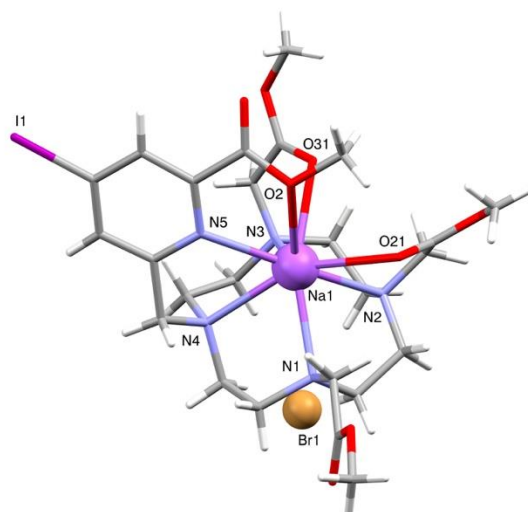
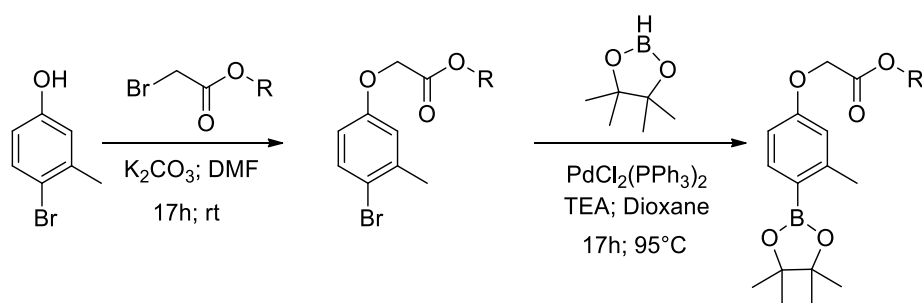


Figure 122 : Structure RX du composé  $\text{DO3A}^{(\text{Me})3}\text{pic}^{(\text{Me})}\text{I}\cdot\text{NaBr}$ . Distances (Å) :  $\text{Na1}\cdots\text{N1}$  2,630 (2),  $\text{Na1}\cdots\text{N2}$  2,532 (2),  $\text{Na1}\cdots\text{N3}$  2,573 (2),  $\text{Na1}\cdots\text{N4}$  2,610 (2),  $\text{Na1}\cdots\text{N5}$  2,612 (2),  $\text{Na1}\cdots\text{O2}$  2,962 (2),  $\text{Na1}\cdots\text{O21}$  2,730 (2),  $\text{Na1}\cdots\text{O31}$  2,585 (2).

## 2) Synthèse des substituants $\text{Ar}_i$ ( $i = 1, 2, 3$ )

### A $\text{Ar}_1$

Les esters boroniques **1b** et **1b'** ont été obtenus en deux étapes (Figure 123). Une première réaction d'alkylation de Williamson a été réalisée entre du 3-méthyl-4-bromophénol et du bromoacétate de méthyle ou de *tert*-butyle en présence de carbonate de potassium, conduisant à la formation quasi-quantitative de **1a** et **1a'**. Une seconde étape de réaction de borylation a été réalisée en présence de palladium bis(triphénylphosphine) dichlorure, de triéthylamine et de pinacolborane à 95°C conduisant aux composés **1b** et **1b'** avec d'excellents rendements.



|                |                  |                  |
|----------------|------------------|------------------|
| R=Me           | <b>1a</b> (92%)  | <b>1b</b> (82%)  |
| R= <i>t</i> Bu | <b>1a'</b> (97%) | <b>1b'</b> (83%) |

Figure 123: Schéma de synthèse des composés **1b** et **1b'**

B Ar<sub>2</sub>

Les composés **2b** et **2b'** ont été synthétisés en deux étapes quasi quantitatives, avec d'abord, comme précédemment une alkylation de Williamson sur l'iodophénol pour donner les composés **2a** et **2a'** (Figure 124), puis un couplage de Sonogashira en présence de palladium bis(triphénylphosphine) dichlorure, de triéthylamine, d'iodure de cuivre(I) et d'éthynyltriméthylsilane. Contrairement à la réaction de borylation, le couplage de Sonogashira a été réalisé à température ambiante, les composés iodés étant plus facilement activables pour ce type de réactions de couplage au palladium(0).

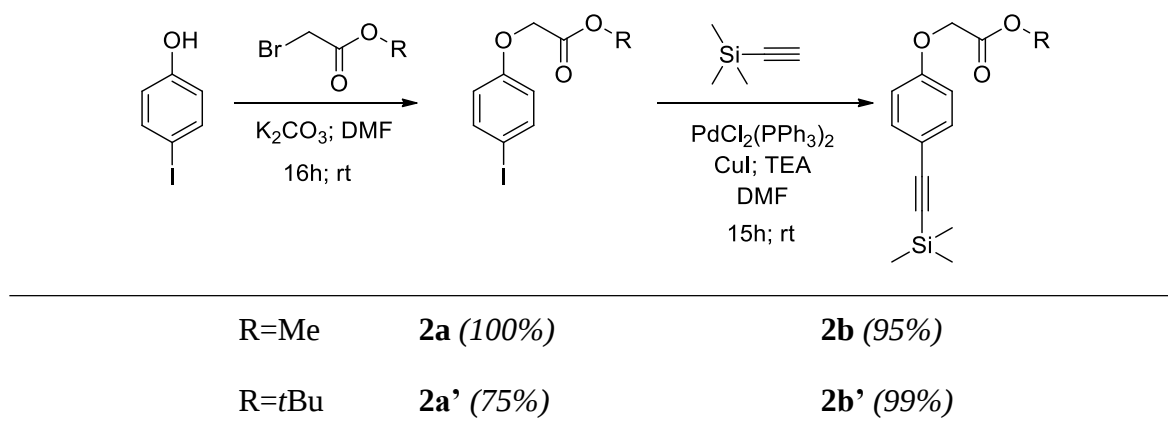
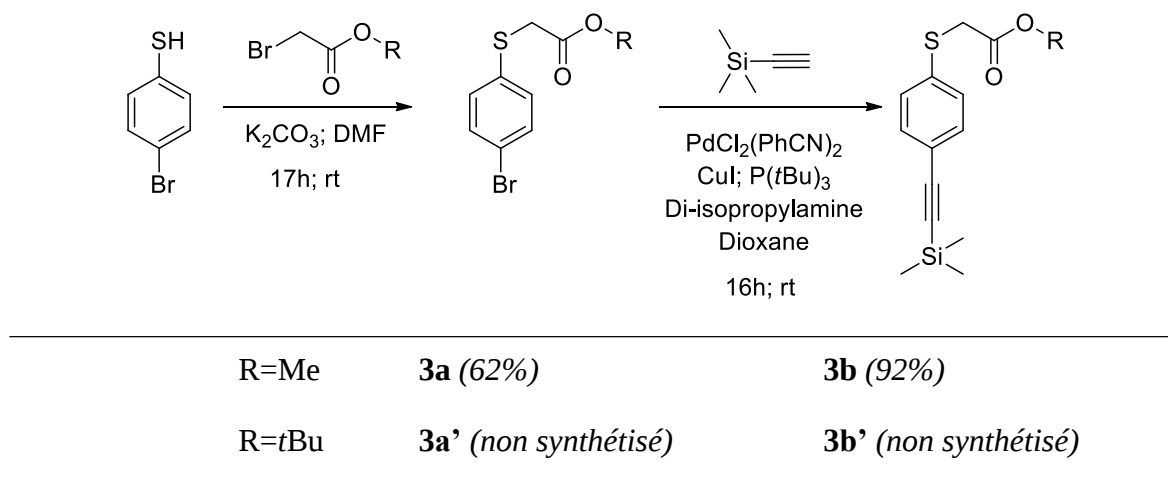


Figure 124: Schéma de synthèse des composés **2b** et **2b'**

C Ar<sub>3</sub>

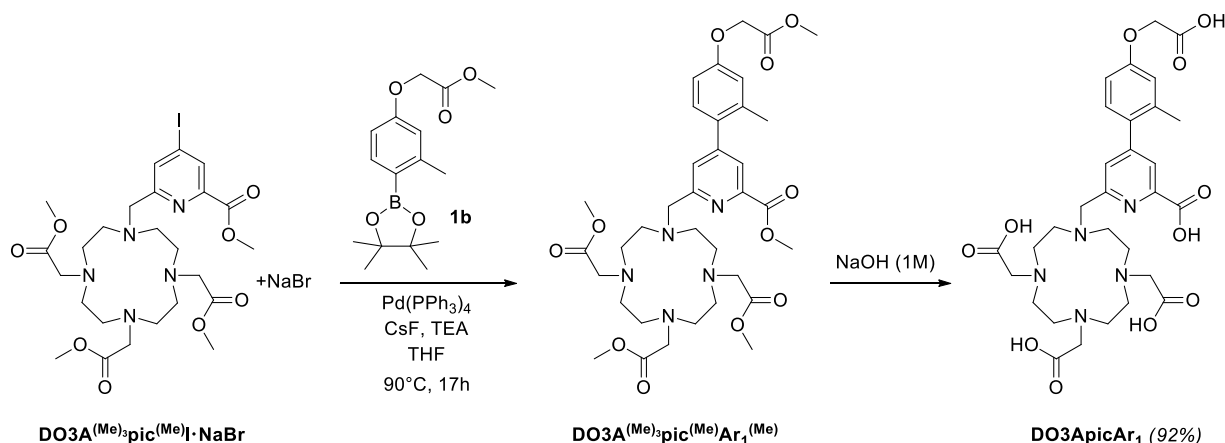
Le composé **3b** a été synthétisé de manière analogue à **2b** et **2b'** avec toutefois quelques différences (Figure 125). La sensibilité à l'oxygène du 4-bromothiophénol en solution, ainsi que le thioéther formé, requièrent des conditions anaérobies en travaillant sous argon. Les conditions du couplage de Sonogashira ont été modifiées par rapport aux conditions utilisées pour les produits **2b** et **2b'**, le composé de départ comportant un atome de brome étant plus difficilement activable pour une étape d'addition oxydante lors de la réaction de couplage. Les conditions de couplage décrites par S. Buchwald<sup>[179]</sup> avec une alkylphosphine riche en électrons ont permis d'obtenir le composé **3b** avec un excellent rendement en réalisant la réaction à température ambiante pour éviter l'oxydation du thioéther.


 Figure 125: Schéma de synthèse des composés **3b** et **3b'**

### 3) Synthèse des ligands DO3ApicAr<sub>1</sub>

#### A Couplages de Suzuki-Miyaura

##### i DO3ApicAr<sub>1</sub>


 Figure 126: Synthèse du ligand **DO3ApicAr<sub>1</sub>**

Le pro-ligand **DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>1</sub><sup>(Me)</sup>** comportant l'antenne pyridine-aryl a été obtenu par un couplage de Suzuki-Miyaura entre les composés **1b** et **DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>I** (Figure 126). Les conditions du couplage ont dû être adaptées, les conditions classiques ne permettant pas l'obtention du produit attendu. Les premiers essais de synthèse ont conduit à l'observation des produits de départ lors de l'analyse par RMN du brut de réaction. L'étape d'addition oxydante sur l'ester boronique n'avait alors certainement pas lieu. Ce manque de réactivité peut être contourné par addition de fluorures dans le milieu réactionnel pour former un trifluoroborate plus réactif que l'ester boronique<sup>[180]</sup>. O. Maury avait déjà observé ce manque réactivité, qu'il

a réussi à contourner en utilisant du fluorure de césium<sup>[181]</sup>. En reprenant ces mêmes conditions, le composé **DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub><sup>(Me)</sup>** a pu être obtenu. Les purifications par chromatographie flash se sont révélées impossibles, le produit restant accroché à la phase stationnaire. Le composé **DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub><sup>(Me)</sup>** étant soluble dans l'eau, il a été extrait du brut de réaction par de l'eau sous ultrasons. Après lyophilisation, une purification par HPLC préparative a permis d'isoler le composé **DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub><sup>(Me)</sup>** pour le caractériser par RMN et ESI/MS. Cependant, cette purification HPLC n'est pas nécessaire pour la suite et le produit peut directement être engagé dans la réaction suivante après lyophilisation. L'hydrolyse basique (NaOH 1 M) suivie d'une purification par HPLC a permis d'obtenir le ligand **DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub>** avec un rendement de 92% (Figure 126).

### ii DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub>

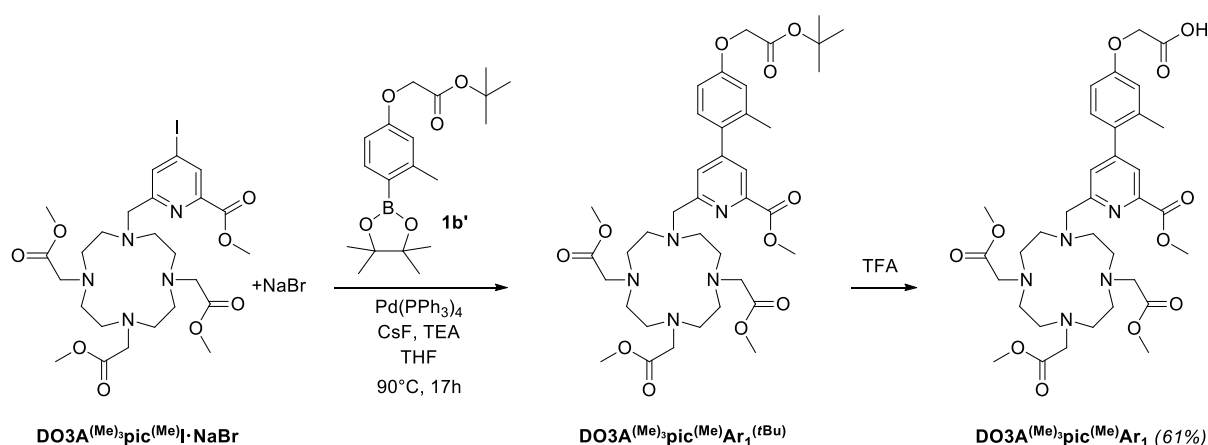
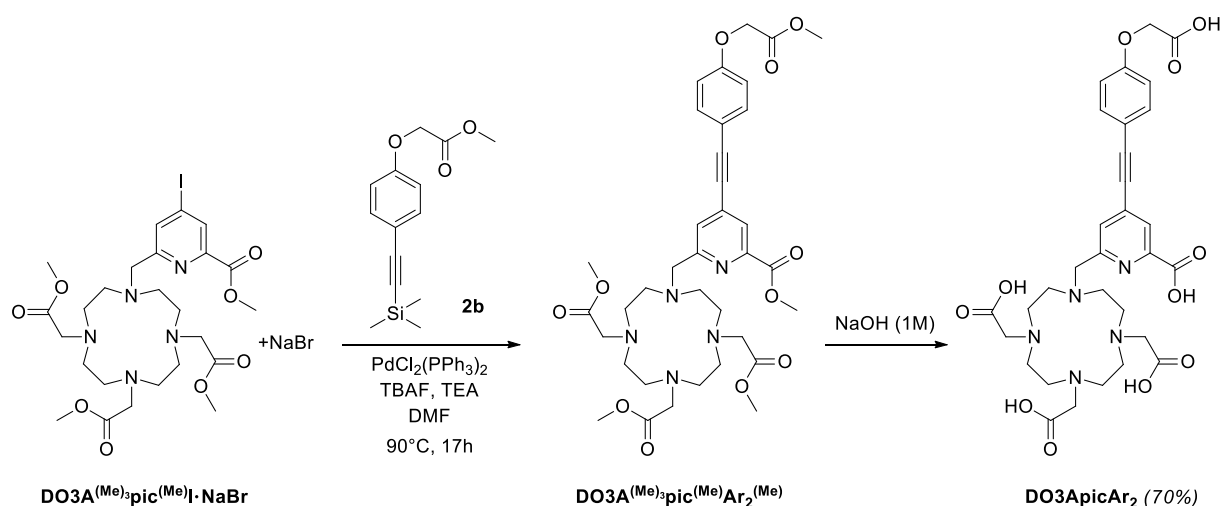


Figure 127: Synthèse du ligand **DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub>**

Le composé **DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub><sup>(tBu)</sup>** a été obtenu par couplage de Suzuki-Miyaura entre **1b'** et **DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>I** dans les conditions données ci-dessus (Figure 127). Une déprotection sélective du groupement protecteur *t*Bu au TFA a été réalisée pour donner après purification par HPLC préparative le ligand **DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub>** avec un rendement de 61% sur les deux étapes. Nous nous sommes aperçu que l'intermédiaire **DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub><sup>(tBu)</sup>** était très sensible aux conditions acides, l'ester *t*Bu du groupe Ar-O-CH<sub>2</sub>COOtBu s'hydrolysant très facilement en milieu même légèrement acide. Ceci a été observé pour chacun des composés de la famille **DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>i</sub><sup>(tBu)</sup>** (*i* = 1, 2).

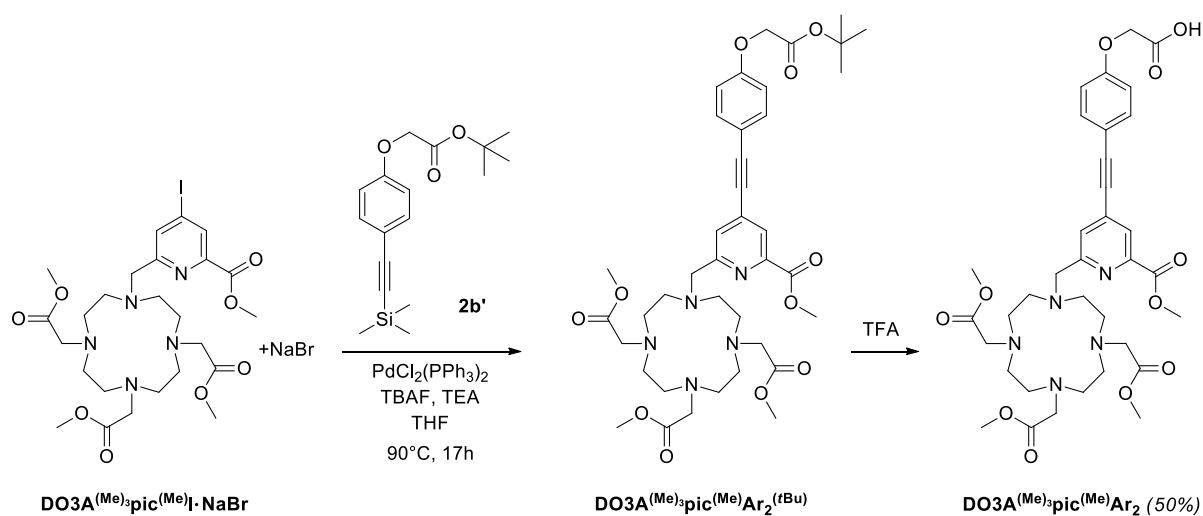
## B Couplages de Sonogashira

### i DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>2</sub>


 Figure 128: Synthèse du ligand **DO3ApicAr<sub>2</sub>**

Le couplage de Sonogashira entre **2b** et **DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>I** (Figure 128) avec désilylation in situ de l'alcyne conduit, dans des conditions de couplage classiques (catalyseur au palladium(0), CuI en quantité catalytique, base), au produit **DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>2</sub><sup>(Me)</sup>** attendu mais avec la coordination du cuivre à l'état Cu<sup>2+</sup> dans le macrocycle. Il n'a pas été possible de séparer la molécule métallée au cuivre. Cu<sup>+</sup> est un co-catalyseur de la réaction de Sonogashira qui favorise la déprotonation de l'alcyne vrai. Le pro-ligand **DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>2</sub><sup>(Me)</sup>** a pu être obtenu en supprimant le co-catalyseur CuI, comme l'a décrit O. Maury pour d'autres composés macrocycliques<sup>[165]</sup>. Après hydrolyse et purification par HPLC, le ligand **DO3ApicAr<sub>2</sub>** a été obtenu avec un rendement de 70%.

## ii **DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>2</sub>**


 Figure 129: Synthèse du ligand **DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>2</sub>**

Le pro-ligand **DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>2</sub>** a été obtenu avec un rendement de 50% par couplage de Sonogashira entre **2b'** et **DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>I** suivi d'une déprotection sélective au TFA (Figure 129).

### iii DO3ApicAr<sub>3</sub>

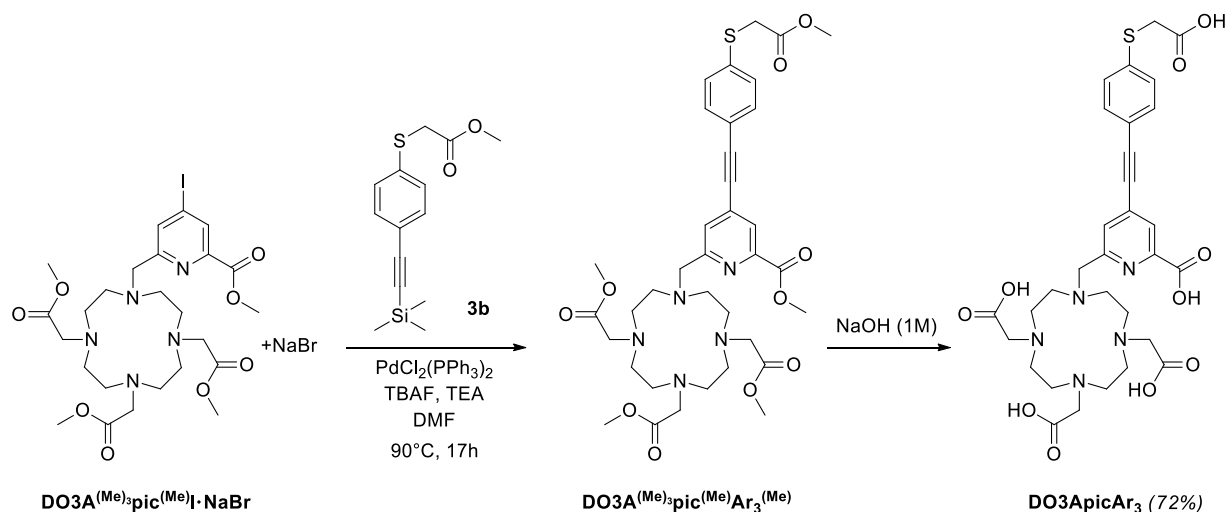


Figure 130: Synthèse du ligand **DO3ApicAr<sub>3</sub>**

Le ligand **DO3ApicAr<sub>3</sub>** a été synthétisé selon le même protocole que le ligand **DO3ApicAr<sub>2</sub>** (Figure 130), avec un rendement de 72% sur les deux étapes.

### 4) Remarque sur la caractérisation des ligands

Tous les ligands ont été caractérisés par ESI-MS, HRMS, et par RMN  $^1\text{H}$  et RMN  $^{13}\text{C}$ . Les spectres RMN  $^1\text{H}$  des ligands obtenus ont premièrement été difficilement interprétables à cause de résonances très larges. Les spectres RMN dans le solvant  $\text{D}_2\text{O}$  ont montré une meilleure définition des pics lorsque la température de l'échantillon augmente (Figure 131). Ainsi, à 353K, tous les protons du composé sont visibles, facilitant notamment l'identification des composés sur des spectres RMN de produits bruts. Les spectres RMN à 353 K des ligands sont donnés en annexes.

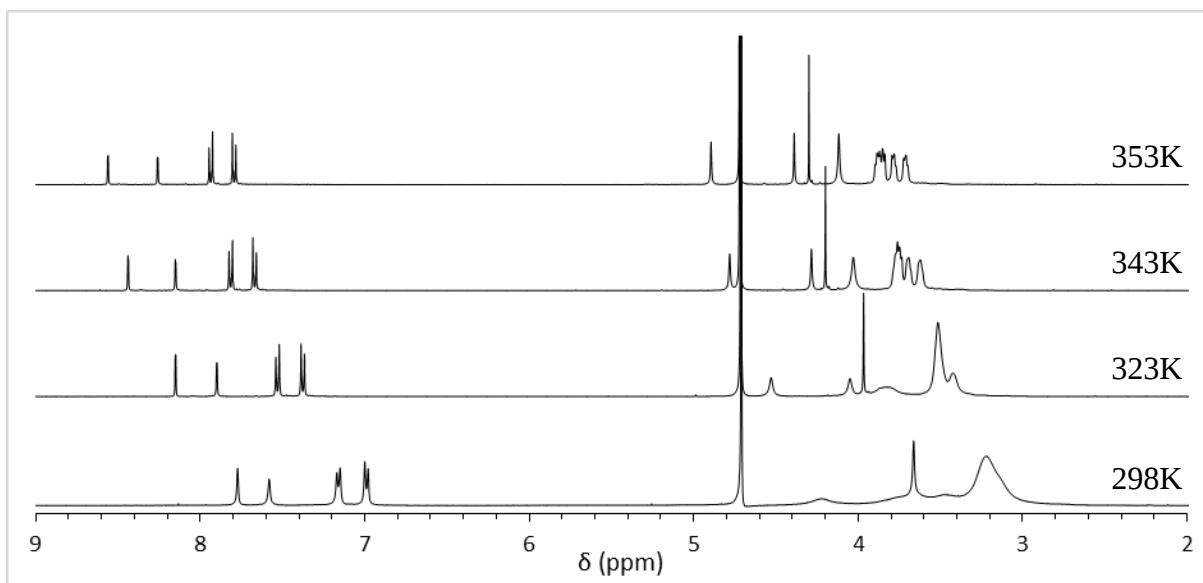


Figure 131: Spectres RMN  $^1\text{H}$  du ligand **DO3ApicAr<sub>3</sub>** à 298, 323, 343 et 353K

## 5) Conclusion

La stratégie de synthèse divergente est validée et permet d'obtenir les ligands avec des rendements satisfaisants. Les différentes étapes ont été optimisées afin de limiter les étapes de purifications notamment après les couplages au palladium, pour lesquels une extraction à l'eau permet désormais d'éliminer la majeure partie des sous-produits en un minimum de temps. Le **DO3A<sup>(Me)<sub>3</sub></sup>pic<sup>(Me)</sup>I** a aussi été obtenu avec un bon rendement, et sa structure RX a été résolue, montrant la présence de sels de bromure de sodium.

## IV Complexes $[\text{Ln}(\text{DO3ApicAr}_i)]^-$

### 1) Synthèse

La métallation du ligand **DO3ApicAr<sub>1</sub>** a été réalisée de la façon suivante. Le ligand est dissout dans de l'eau MilliQ, puis la solution est amenée à pH 9 par ajouts de triéthylamine après ajout d'un excès de  $\text{TbCl}_3$ . Le milieu réactionnel est agité une nuit à température ambiante puis centrifugé. Le surnageant est récupéré et évaporé à sec. Le solide obtenu est précipité dans de l'acétone, trituré aux ultrasons, puis centrifugé. Cette étape est répétée plusieurs fois, puis le solide obtenu est séché. L'analyse ESI-MS du produit obtenu est en accord avec la formation du complexe  $[\text{Tb}(\text{DO3ApicAr}_1)]^-$  (Figure 132).

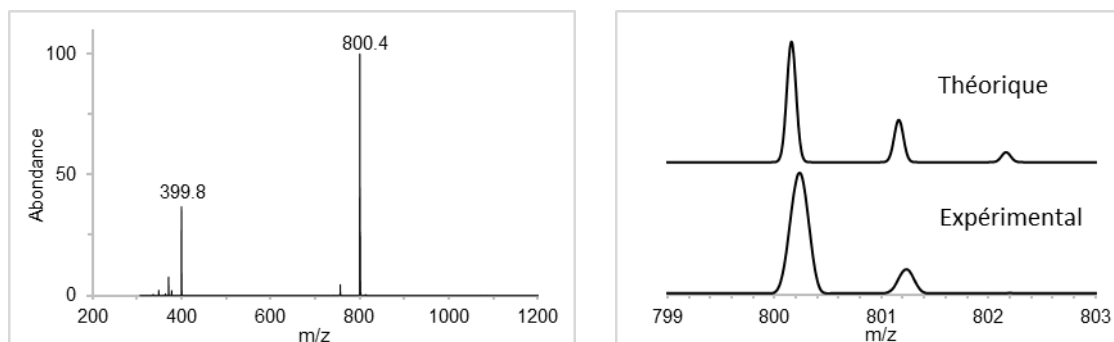


Figure 132 : A gauche, spectre ESI-MS du complexe DO3ApicAr<sub>1</sub>; A droite, massif isotopique expérimental et théorique (Masse calculée: 800,16 [M-H]<sup>-</sup>, 399,58 [M-2H]<sup>2-</sup>, pour M = C<sub>30</sub>H<sub>36</sub>N<sub>5</sub>O<sub>11</sub>Tb).

Des monocristaux de ce complexe ont été obtenus par évaporation lente d'une solution dans un mélange méthanol/eau. La structure cristallographique montre deux complexes [Tb(DO3ApicAr<sub>1</sub>)]<sup>-</sup> pontés par deux ions Na<sup>+</sup> au niveau du groupement carboxylate des picolines (Figure 133). Chacune des unités [Tb(DO3ApicAr<sub>1</sub>)]<sup>-</sup> confirme l'internalisation de l'ion Tb<sup>3+</sup> dans le ligand macrocyclique. Le terbium est nona-coordonné avec dans sa sphère de coordination 5 atomes d'azotes (les quatre atomes d'azotes du cyclène et celui du picolinate) et quatre atomes d'oxygène (trois des bras acétate et 1 du picolinate). Le groupement électro-donneur de l'antenne est sous forme acide carboxylique (protonné). La torsion entre la pyridine et le groupement aryl est de 46°. Cette structure nous confirme donc la formation d'un complexe mononucléaire et la présence d'un contre-ion sodium nous confirme sa nature anionique. Le composé obtenu est donc de formule [Tb(DO3ApicAr<sub>1</sub>)]Na.

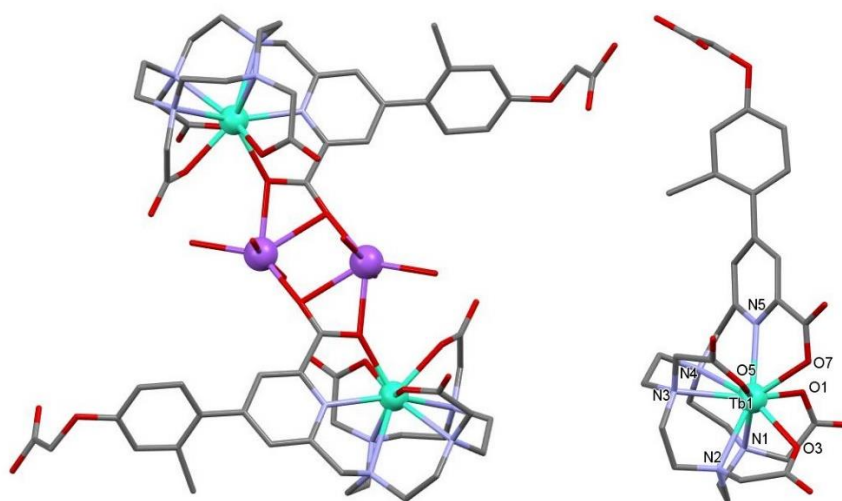


Figure 133 : Structure RX du composé [Tb(DO3ApicAr<sub>1</sub>)]Na. Distances (Å): : Tb1...N1 2,617 (6), Tb1...N2 2,653 (7), Tb1...N3 2,689 (6), Tb1...N4 2,723 (6), Tb1...N5 2,504 (6), Tb1...O1 2,348 (5), Tb1...O3 2,340 (5), Tb1...O5 2,361 (5), Tb1...O7 2,397 (5)

De manière analogue au complexe  $[\text{Tb}(\text{DO3ApicAr}_1)]\text{Na}$ , les complexes  $[\text{Eu}(\text{DO3ApicAr}_1)]\text{Na}$  et  $[\text{Dy}(\text{DO3ApicAr}_1)]\text{Na}$  ont été synthétisés à partir du ligand **DO3ApicAr<sub>1</sub>** (Figure 134 A), O. Maury ayant décrit l'antenne picAr<sub>1</sub> comme pouvant sensibiliser l'euprium ainsi que le dysprosium<sup>[170]</sup>.

Les complexes  $[\text{Eu}(\text{DO3ApicAr}_2)]\text{Na}$  et  $[\text{Sm}(\text{DO3ApicAr}_2)]\text{Na}$  ont été obtenus à partir du ligand **DO3ApicAr<sub>2</sub>** (Figure 134 B) comme les complexes précédents, l'antenne picAr<sub>2</sub> ayant été décrite comme pouvant sensibiliser l'euprium et le samarium<sup>[163]</sup>.

Le dernier complexe  $[\text{Eu}(\text{DO3ApicAr}_3)]\text{Na}$  a été synthétisé à partir du ligand **DO3ApicAr<sub>3</sub>** (Figure 134 C) toujours dans les mêmes conditions mais sous atmosphère inerte.

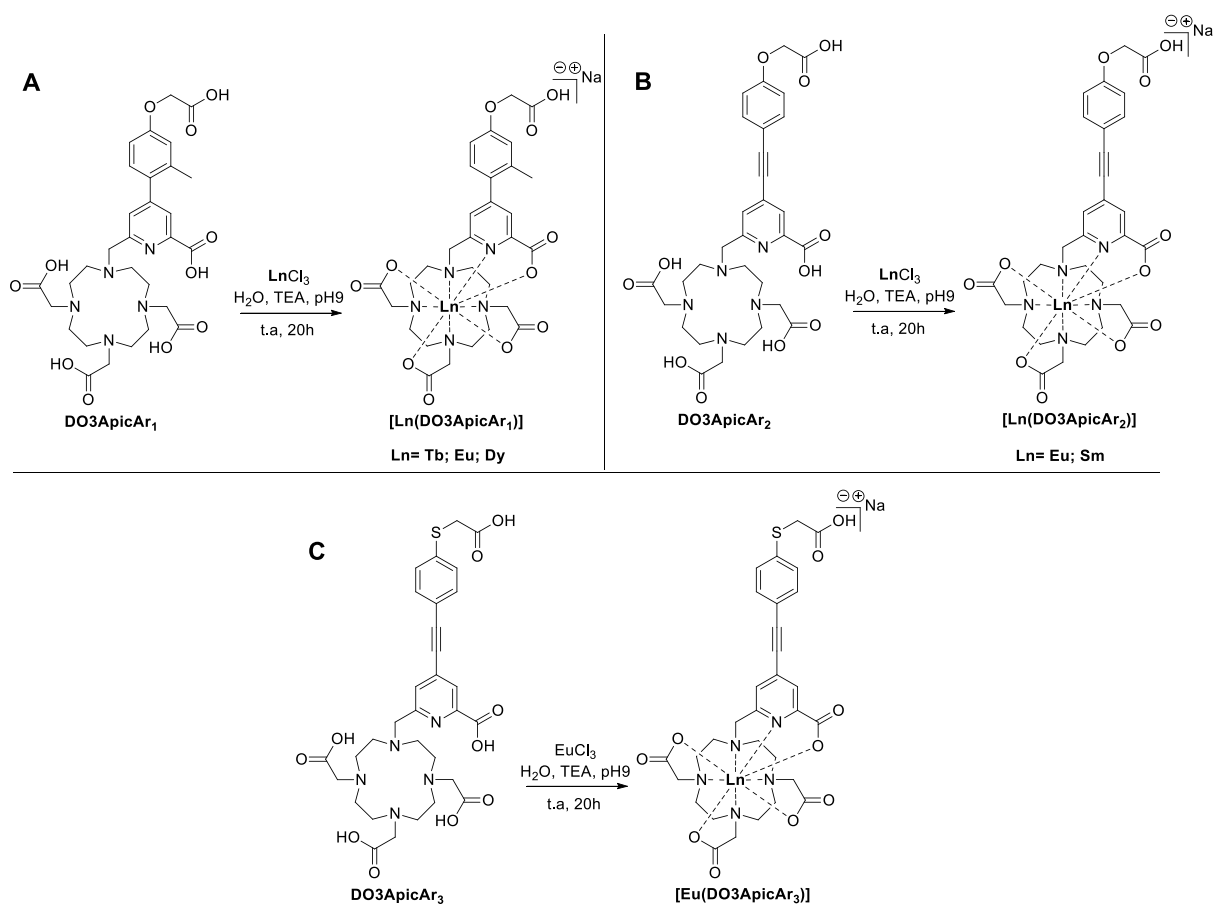


Figure 134: Etape de métallation des différents ligands synthétisés

## 2) Propriétés spectroscopiques des complexes

### A Absorption et coefficients d'extinction molaire

Les coefficients d'extinction molaire des différents complexes ont été déterminés en titrant une solution de ligand par une solution de lanthanide(III) de concentration connue. La Figure 135 montre l'évolution de l'absorption d'une solution de DO3ApicAr<sub>1</sub> lors de l'ajout de Tb<sup>3+</sup>. L'absorbance augmente linéairement jusqu'à ce que l'antenne picolinate soit totalement liée au lanthanide et stagne ensuite. La rupture doit être observée pour une stoechiométrie ligand/lanthanide 1:1, ce qui permet de déterminer la concentration en ligand à partir de la concentration en lanthanide. Ce type de titrage permet de déterminer le coefficient d'extinction molaire au maximum d'absorption. Les longueurs d'ondes du maximum d'absorption,  $\lambda_{\text{max}}$ , et les  $\epsilon$  correspondants sont donnés dans le Tableau 2 pour chacun des complexes [Ln(DO3ApicAr<sub>i</sub>)]Na (i = 1, 2, 3).

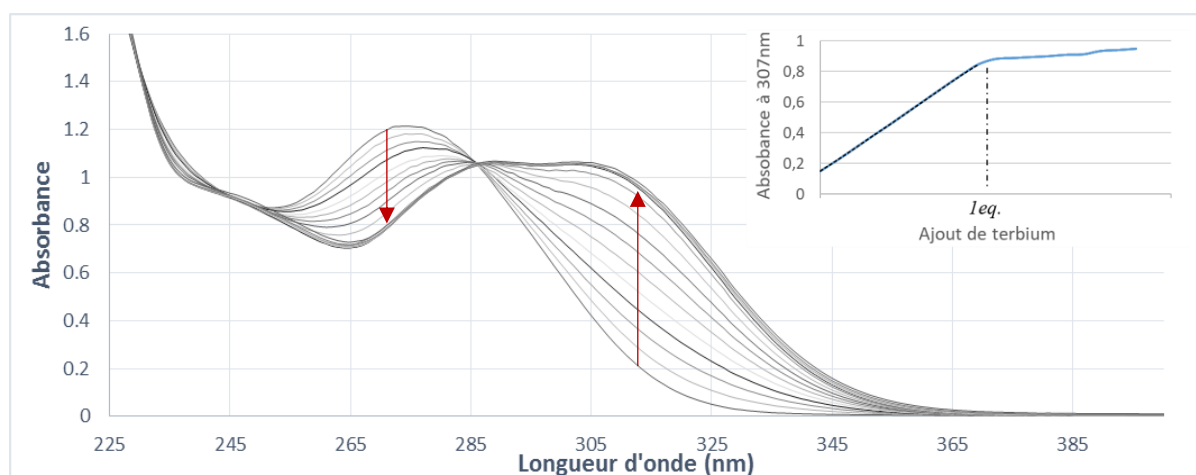


Figure 135 : Titrage du ligand DO3ApicAr<sub>1</sub> par du TbCl<sub>3</sub> suivi par spectrophotométrie UV..

### B Luminescence

Les propriétés de luminescence des complexes en solution dans un tampon HEPES pH 7,5 ou dans l'eau ont été étudiées et sont compilées dans le Tableau 2. La Figure 136 montre les spectres d'émission enregistrés en excitant l'antenne à son maximum d'absorption et les spectres d'excitation enregistrés en détectant au maximum d'émission du lanthanide. Les spectres d'émission présentent les raies caractéristiques des lanthanides (Tb, Eu, Dy et Sm) et l'émission de l'antenne n'est pas détectée, sauf pour le complexe [Sm(DO3ApicAr<sub>2</sub>)]Na. Cette tendance a déjà été observée par le groupe de O. Maury pour des complexes ayant les mêmes

antennes<sup>[163]</sup>. Les spectres d'excitation du lanthanide confirment pour chacun des complexes que l'excitation se fait via l'antenne picolinate. L'absence d'émission de l'antenne indique un transfert d'énergie efficace de l'antenne au lanthanide. Notons que les spectres d'émission des 3 complexes d'euprium sont identiques, indiquant que l'environnement de l'ion  $\text{Eu}^{3+}$  reste le même, ce qui était attendu pour cette famille de ligands. Le temps de vie de luminescence du complexe  $[\text{Tb}(\text{DO3ApicAr}_1)]\text{Na}$ ,  $\tau_{\text{Tb}} = 2,35$  ms, est particulièrement long en comparaison du complexe  $\text{no3A4paTb}$  décrit par O. Maury<sup>[170]</sup> possédant la même antenne et un temps de vie  $\tau_{\text{Tb}} = 1,36$  ms. Le rendement quantique du complexe  $[\text{Tb}(\text{DO3ApicAr}_1)]\text{Na}$  est 62 %, ce qui est particulièrement élevé. Ce rendement quantique s'approche de celui du complexe  $\text{no3A4paTb}$  (74 %) mais la présence de trois antennes  $\text{picAr}_1$  dans ce dernier lui confère une brillance bien supérieure. Les complexes d'euprium et de dysprosium avec ce ligand  $\text{DO3ApicAr}_1$  ont des rendements quantiques plus faibles (6,4 et 1,2 %, respectivement) comme attendu. Le complexe  $[\text{Eu}(\text{DO3ApicAr}_2)]\text{Na}$  a un rendement quantique de 17 % beaucoup plus faible que  $[\text{Tb}(\text{DO3ApicAr}_1)]\text{Na}$  mais sa forte absorption, décalée vers le rouge à 345 nm ( $\epsilon_{\text{max}} = 28\,200 \text{ M}^{-1} \text{ cm}^{-1}$ ) en fait un complexe intéressant. Le temps de vie de luminescence de l'euprium est de 1,10 ms. En 2017, le groupe de G.-L. Law<sup>[182]</sup> a décrit une famille de complexes analogues avec un macrocycle  $\text{DO3A}$  fonctionnalisé par une antenne pyridine-phosphinate substituée par différents groupements alcyne-aryl (Figure 137). Les rendements quantiques associés vont de 6 à 24% et sont de 16 et 20% pour les deux composés comportant une antenne pyridine-alcyne-aryl avec un groupe *O-alkyl* électro-donneur. Le complexe  $[\text{Eu}(\text{DO3ApicAr}_3)]\text{Na}$  avec l'antenne pyridine-alcyne-aryl soufrée présente des propriétés spectroscopiques décevantes ( $\Phi = 6,4$  % et  $\tau_{\text{Eu}} = 0,66$  ms) au regard du composé analogue décrit par O. Maury,  $\text{Na}[\text{Eu}(\text{L}^2)]$  ( $\Phi = 28,4$  % et  $\tau_{\text{Eu}} = 1,39$  ms) possédant un groupement électro-donneur  $\text{S}-(\text{CH}_2\text{CH}_2\text{O})_3-\text{CH}_3$ . Une telle différence n'était pas attendue.

Pour finir, le nombre de molécules d'eau présentes dans l'environnement du lanthanide a été déterminé pour chaque complexe, l'équation de Horrocks permettant de le calculer pour le terbium suivant l'équation

$$q_{\text{Tb}} = 5,0 \times \left[ \left( \frac{1}{\tau_{\text{H}_2\text{O}}} - \frac{1}{\tau_{\text{D}_2\text{O}}} \right) - 0,06 \right]$$

et pour l'euprium suivant l'équation

$$q_{Eu} = 1,2 \times \left[ \left( \frac{1}{\tau_{H_2O}} - \frac{1}{\tau_{D_2O}} \right) - 0,25 - 0,075 \times n_{CONH} \right]$$

,  $\tau_{H_2O}$  correspondant aux temps de vie des lanthanides dans l'eau et  $\tau_{D_2O}$  à celui dans l'eau lourde ( $D_2O$ ).

Sans surprise, le nombre de molécules d'eau calculé pour chaque gamme de complexe est proche de 0, indiquant l'absence de molécules d'eau dans l'environnement du lanthanide et par conséquent la bonne internalisation du lanthanide.

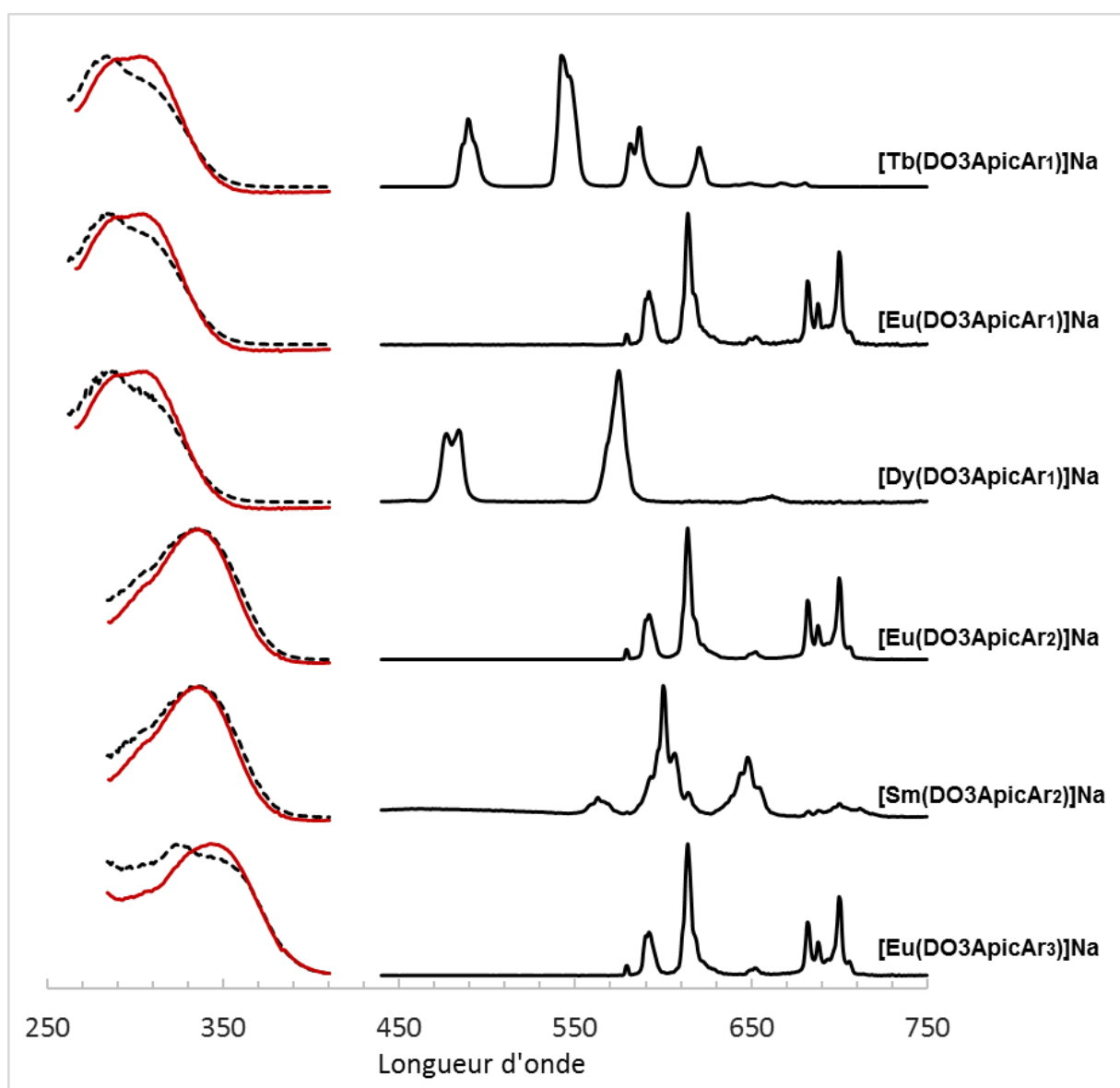


Figure 136 : Spectres d'absorption (en rouge), d'émission (trait plein noir) et d'excitation (tirets noirs) des complexes  $[Ln(DO3ApicAr_i)]Na$  ( $i = 1, 2, 3$ ).

Tableau 2 : Propriétés spectroscopiques des complexes [Ln(DO3Apic<sub>i</sub>)]Na.

| Complexe :                      | $\lambda_{\text{abs}}^{\text{max}}$<br>(nm) | $\epsilon$<br>(M <sup>-1</sup> cm <sup>-1</sup> ) | $\tau_{\text{H}_2\text{O}}$<br>(ms) | $\tau_{\text{D}_2\text{O}}$<br>(ms) | q    | $\phi_{\text{H}_2\text{O}}$<br>(%) | B<br>(M <sup>-1</sup> cm <sup>-1</sup> ) |
|---------------------------------|---------------------------------------------|---------------------------------------------------|-------------------------------------|-------------------------------------|------|------------------------------------|------------------------------------------|
| [Tb(DO3ApicAr <sub>1</sub> )]Na | 306                                         | 12050                                             | 2,35                                | 2,47                                | -0,2 | 62                                 | 7470                                     |
| [Eu(DO3ApicAr <sub>1</sub> )]Na | 306                                         | 12050                                             | 1,10                                | /                                   | /    | 6,4                                | 771                                      |
| [Dy(DO3ApicAr <sub>1</sub> )]Na | 306                                         | 12050                                             | 0,1                                 | /                                   | /    | 1,2                                | 140                                      |
| [Eu(DO3ApicAr <sub>2</sub> )]Na | 345                                         | 28200                                             | 1,09                                | 1,61                                | 0,06 | 17                                 | 4790                                     |
| [Sm(DO3ApicAr <sub>2</sub> )]Na | 345                                         | 28200                                             | 0,02                                | /                                   | /    | 0,4                                | 110                                      |
| [Eu(DO3ApicAr <sub>3</sub> )]Na | 345                                         | 26100                                             | 0,66                                | 0,89                                | 0,18 | 6,4                                | 1670                                     |

| Complex | $\lambda_{\text{abs}}$<br>(max)<br>[nm] | $\tau_{\text{HEPES}}$<br>[ms] | $\tau_{\text{H}_2\text{O}}$<br>[ms] | $\tau_{\text{D}_2\text{O}}$<br>[ms] | $q_{\text{H}_2\text{O}}$ <sup>[a,b]</sup> | $\phi_{\text{HEPES}}$ | $\tau_{\text{rad}}$<br>[ms] | $\Phi_{\text{Eu}}^{\text{Eu}}$ | $\eta_{\text{sens}}$ |
|---------|-----------------------------------------|-------------------------------|-------------------------------------|-------------------------------------|-------------------------------------------|-----------------------|-----------------------------|--------------------------------|----------------------|
| EuL1    | 353                                     | 1.00                          | 1.01                                | 1.44                                | 0.06, 0                                   | 0.069                 | 4.22                        | 0.24                           | 0.29                 |
| EuL2    | 303                                     | 1.27                          |                                     |                                     |                                           |                       |                             |                                |                      |
| EuL3    | 330                                     | 1.27                          | 1.04                                | 1.92                                | 0.23, 0.15                                | 0.19                  | 4.30                        | 0.29                           | 0.64                 |
| EuL4    | 332                                     | 1.23                          | 1.02                                | 1.95                                | 0.26, 0.17                                | 0.24                  | 4.45                        | 0.28                           | 0.87                 |
| EuL5    | 332                                     | 1.25                          | 1.15                                | 1.37                                | 0, 0                                      | 0.20                  | 4.28                        | 0.29                           | 0.69                 |
| EuL6    | 338                                     | 1.25                          | 1.16                                | 1.42                                | 0, 0                                      | 0.16                  | 4.39                        | 0.29                           | 0.54                 |
| EuL7    | 328                                     | 1.24                          | 1.12                                | 1.96                                | 0.16, 0.08                                | 0.23                  | 4.29                        | 0.29                           | 0.79                 |
| EuL8    | 308                                     | 1.20                          | 1.17                                | 1.95                                | 0.11, 0.04                                | 0.20                  | 4.53                        | 0.26                           | 0.75                 |
| EuL9    | 321                                     | 1.25                          | 1.23                                | 1.98                                | 0.07, 0                                   | 0.20 <sup>[d]</sup>   | 4.18                        | 0.30                           | 0.660                |
| EuL10   | 300                                     | 1.27                          | 1.27                                | 2.00                                | 0.04, 0                                   | 0.24 <sup>[d]</sup>   | 4.44                        | 0.29                           | 0.85                 |
| EuL11   | 337                                     | 1.23                          | 1.14                                | 1.90                                | 0.12, 0.05                                | 0.17                  | 4.29                        | 0.29                           | 0.60                 |
| EuL12   | 330                                     | 1.20                          | 1.03                                | 1.91                                | 0.24, 0.15                                | 0.20                  | 4.27                        | 0.28                           | 0.71                 |
| EuL13   | 318                                     | 1.22                          | 1.11                                | 1.99                                | 0.18, 0.10                                | 0.23                  | 4.42                        | 0.28                           | 0.82                 |
| EuL14   | 335                                     | 0.57                          | 0.57                                | 0.81                                | 0.32, 0.23                                | 0.073                 | 4.30                        | 0.13                           | 0.55                 |
| EuL15   | 306                                     | 1.27                          | 1.25                                | 1.98                                | 0.05, 0                                   | 0.18                  | 4.58                        | 0.28                           | 0.65                 |
| EuL16   | 327                                     | 1.09                          | 1.11                                | 1.64                                | 0.05, 0                                   | 0.13                  | 4.32                        | 0.25                           | 0.50                 |

[a] Determined using Parker's equation.<sup>[17]</sup> [b] Determined using Horrocks' equation.<sup>[18]</sup> [c] Not measured due to weak luminescence. [d] Measured against quinine sulfate (0.1 N H<sub>2</sub>SO<sub>4</sub>,  $\lambda_{\text{ex}} = 310 \text{ nm}$ ,  $\phi = 0.53$ ).<sup>[11]</sup> Estimated errors of quantum yield or lifetime are  $\pm 15\%$  and  $\pm 10\%$ , respectively.

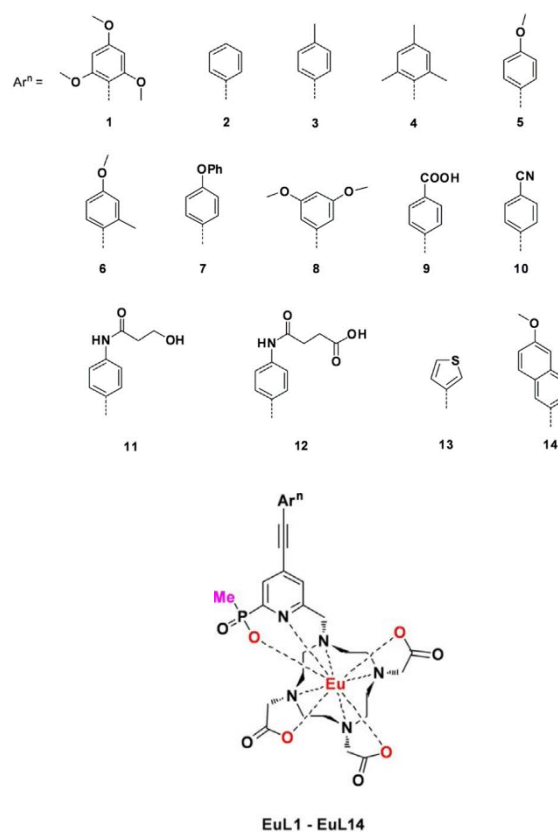


Figure 137: Complexes d'euprium basés sur un chélate DO3A décrits dans la littérature et leurs rendements quantiques

### C Remarque sur la synthèse des complexes

Trois protocoles différents de métallation de ligands de type DO3Apicolinate ont été décrits dans la littérature<sup>[166],[136],[183]</sup>. Avant d'utiliser les conditions données plus tôt (p.117), ces conditions ont été testés pour la métallation du ligand **DO3ApicAr<sub>1</sub>** par du terbium(III) :

- a) hydrolyse des esters du pro-ligand dans une solution aqueuse de NaOH puis ajout d'un excès de sel de lanthanide et métallation à 80°C dans l'eau ajustée à pH 6,0 par ajouts de HCl<sup>[166]</sup> ;
- b) métallation avec un excès de sel de lanthanide, à 90°C dans l'eau à pH 9,0, ajusté avec NaOH puis purification par colonne de phase inverse<sup>[136]</sup> ;
- c) métallation dans l'isopropanol (*i*-PrOH) avec une stœchiométrie ligand :lanthanide 1:1 très précise et de la triéthylamine comme base, suivi d'une précipitation au THF<sup>[183]</sup>.

Pour chaque cas, afin d'évaluer si la métallation a été correctement réalisée, le complexe obtenu a été étudié dans une solution d'HEPES 100 mM à pH 7,5 puis le pH a été abaissé. Le spectre d'émission du terbium en excitant l'antenne à 302 nm a été enregistré ainsi que son temps de vie. Les différentes conditions de synthèse testées n'ont pas conduit à la formation d'un complexe stable aux effets de pH. La Figure 138 montre l'évolution du temps de vie de luminescence du terbium pour ces trois conditions et pour les conditions que nous avons retenues (H<sub>2</sub>O, TEA, pH 9, t.a.). Seules nos conditions ont permis de récupérer et isoler un complexe dont le temps de vie ne change pas avec le pH. Pour les autres conditions, la chute du temps de vie observée en abaissant le pH peut être due à deux facteurs : le terbium n'est pas correctement internalisé dans le ligand et/ou l'excès de terbium(III) utilisé lors de la métallation n'a pas été correctement éliminé. La difficulté à internaliser correctement le lanthanide dans les structures de type DOTA est connue. Un intermédiaire où le lanthanide se lie aux groupements carboxylates et à 2 des 4 atomes d'azote du cyclène se forme rapidement, puis la coordination aux deux atomes d'azote restant est plus lente<sup>[184]</sup>. Il est possible que cela se produise aussi pour le DO3Apicolinate. La deuxième hypothèse a pu être vérifiée en ajoutant un excès de terbium au complexe [Tb(DO3ApicAr<sub>1</sub>)]Na isolé dans nos conditions, dont le temps de vie chute fortement à pH acide. Il est donc possible que nous n'ayons pas réussi à éliminer convenablement l'excès de terbium avec les protocoles de la littérature. Notons que les complexes que nous avons obtenus se sont avérés insolubles dans des conditions acides utilisées pour la séparation par HPLC en phase inverse, ce qui ne nous a pas permis d'utiliser cette technique pour éliminer l'excès de lanthanide. Les conditions que nous proposons permettent

d'utiliser un excès de lanthanide (ce qui est plus facile à gérer avec des ligands hygroscopiques) et de l'éliminer facilement par centrifugation et des lavages à l'acétone, nous affranchissant d'une étape de purification par chromatographie en phase inverse.

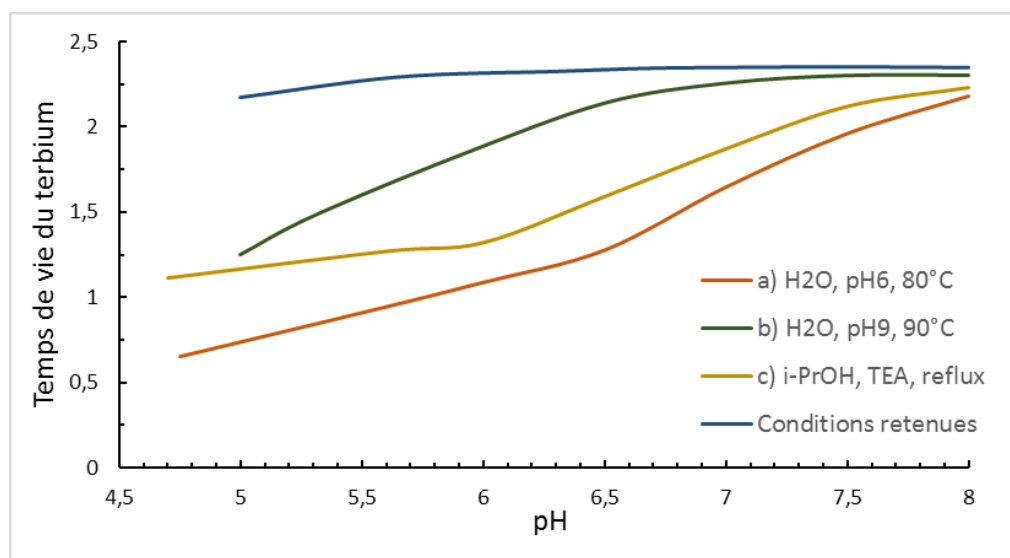


Figure 138: Evolution du temps de vie du terbium en fonction du pH pour les différentes conditions de métallation testées

## V Conclusion

Une nouvelle famille de complexes de lanthanides a été développée. Les efforts synthétiques ont été minimisés par l'utilisation d'une stratégie de synthèse divergente dont les étapes ont été optimisées et requiert désormais le minimum de purifications. Les antennes choisies permettent de sensibiliser différents lanthanides, comme décrit dans la littérature, avec au total 4 lanthanides utilisés (europium, terbium, dysprosium, samarium) permettant de couvrir une gamme de longueurs d'ondes d'émission allant de 470 à 710 nm (l'émission dans le domaine proche infra-rouge du dysprosium et du samarium ne sont pas présentés). Différentes propriétés de luminescence des complexes ont été déterminées (coefficient d'extinction molaire, nombre de molécules d'eau, rendement quantique, temps de vie), ce qui a permis de mettre en valeur deux complexes possédant une forte brillance ( $7470 \text{ M}^{-1}\text{cm}^{-1}$  pour  $[\text{Tb}(\text{DO3ApicAr}_1)]\text{Na}$ , et  $4790 \text{ M}^{-1}\text{cm}^{-1}$  pour  $[\text{Eu}(\text{DO3ApicAr}_2)]\text{Na}$ ), extrêmement intéressant pour leurs utilisations dans des systèmes de FRET lanthanide vers fluorophore. Le design de ces complexes et de leurs ligands permet par ailleurs deux stratégies de conjugaison sur des biomolécules, qui vont être testés dans le prochain chapitre.



Chapitre IV : Sondes de seconde  
génération, FRET d'un lanthanide  
vers un fluorophore



## Chapitre IV : Sondes de seconde génération, FRET d'un lanthanide vers un fluorophore organique

### I Introduction

Comme nous l'avons vu dans le chapitre II, le système reposant sur le FRET entre chromophores organiques comporte certains inconvénients. Nous souhaitons développer un système de FRET pour lequel un complexe de lanthanide serait le donneur d'énergie et pour cela, nous avons conçu une nouvelle famille de complexes conjugables à un peptide ou une protéine, que nous avons décrite dans le chapitre précédent. La Figure 139 montre les spectres d'absorption et d'émission des différents chromophores à notre disposition. On peut noter le bon recouvrement entre l'émission du terbium et l'absorption de la lissamine ainsi qu'entre l'émission de l'euporium et l'absorption de la SiR-méthyl. Ainsi, en nous appuyant sur les protéines ZIF<sup>Zn</sup>(lissamine) et ZIF<sup>Zn</sup>(SiR-méthyl) décrites dans le Chapitre II, nous pouvons proposer deux nouveaux couples pour le FRET. Le premier est ZIF<sup>Zn</sup>(Tb(DO3ApicAr<sub>1</sub>))/ZIF<sup>Zn</sup>(lissamine) où le donneur sera le complexe [Tb(DO3ApicAr<sub>1</sub>)]<sup>-</sup> et l'accepteur la lissamine. Le second est le couple ZIF<sup>Zn</sup>(Eu(DO3ApicAr<sub>2</sub>))/ZIF<sup>Zn</sup>(SiR-méthyl), où le complexe [Eu(DO3ApicAr<sub>2</sub>)]<sup>-</sup> jouera le rôle de donneur et la SiR-méthyl celui d'accepteur. Pour ces deux couples, les rayons de Förster  $R_0$  peuvent être estimés à 54 et 47 Å, respectivement, en utilisant les équations données dans le Chapitre II (p.80). Le couple ZIF<sup>Zn</sup>(Tb(DO3ApicAr<sub>1</sub>))/ZIF<sup>Zn</sup>(SiR-méthyl) peut également être envisagé. Le recouvrement spectral entre l'émission du terbium et l'absorption de la SiR-méthyl est moins bon mais le rendement quantique du donneur [Tb(DO3ApicAr<sub>1</sub>)]<sup>-</sup> est meilleur que celui de [Eu(DO3ApicAr<sub>2</sub>)]<sup>-</sup>, ce qui conduit à un rayon de Förster de 47 Å, similaire à celui du couple ZIF<sup>Zn</sup>(Eu(DO3ApicAr<sub>2</sub>))/ZIF<sup>Zn</sup>(SiR-méthyl).

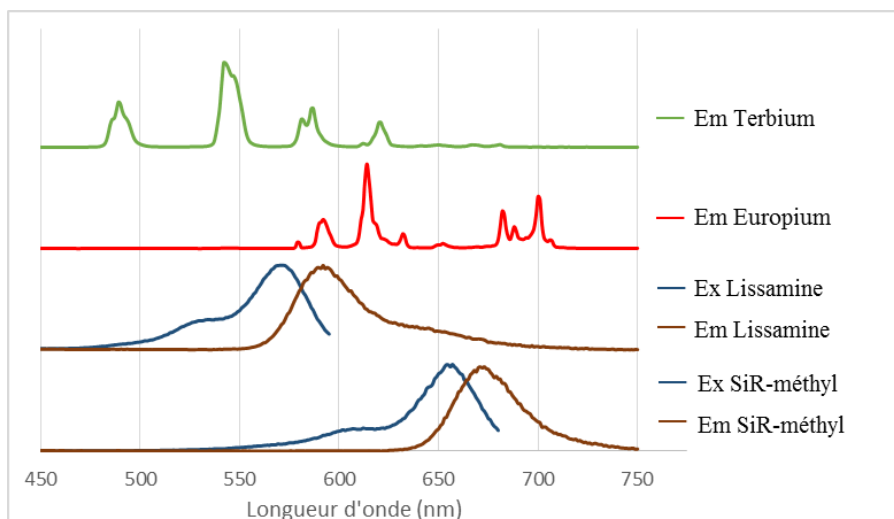


Figure 139: Spectres d'excitations et d'émission des chromophores utilisés

## II FRET du terbium vers un fluorophore

La formation d'un système de FRET d'un lanthanide vers un fluorophore a premièrement été étudié en utilisant le terbium, par combinaison d'une protéine comportant le complexe de terbium  $[\text{Tb}(\text{DO3ApicAr}_1)]^-$ , et les protéines  $\text{ZIF}^{\text{Zn}}(\text{lissamine})$  et  $\text{ZIF}^{\text{Zn}}(\text{SiR-méthyl})$ .

### 1) Synthèse de la protéine $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))$

La protéine  $\text{ZIF}(\text{Tb}(\text{DO3ApicAr}_1))$  a été synthétisée par la même stratégie que pour les deux premières protéines de la sonde de seconde génération avec les fluorophores organiques. Il faut premièrement obtenir le peptide  $\text{F2}[\text{Cys}(\text{StBu})|\text{Lys}(\text{Tb}(\text{DO3ApicAr}_1))]-\text{SEA}^{\text{off}}$  (Figure 140) en utilisant le pro-ligand  $\text{DO3A}^{(\text{Me})_3}\text{pic}^{(\text{Me})}\text{Ar}_1$  ou le complexe  $[\text{Tb}(\text{DO3ApicAr}_1)]\text{Na}$  développés au chapitre précédent.

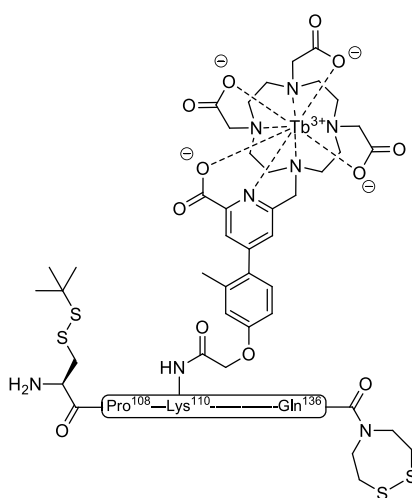


Figure 140: Structure du peptide  $\text{F2}[\text{Cys}(\text{StBu})|\text{Lys}(\text{Tb}(\text{DO3ApicAr}_1))]-\text{SEA}^{\text{off}}$  à obtenir

## A Peptide F2[Cys(StBu)|Lys(Tb(DO3ApicAr<sub>1</sub>))]-SEA<sup>off</sup>

### i Couplage du complexe [Tb(DO3ApicAr<sub>1</sub>)]Na

La stratégie de synthèse la plus simple pour l'obtention du peptide F2[Cys(StBu)|Lys(Tb(DO3ApicAr<sub>1</sub>))]-SEA<sup>off</sup> est par le couplage direct du complexe [Tb(DO3ApicAr<sub>1</sub>)]Na sur le peptide F2[Cys(StBu)|Lys(NH<sub>2</sub>)]-SEA-2Cl-trt-PS synthétisé en partie « Fragments F2 » (p.82), c'est à dire par la stratégie B de la Figure 116 (p.106). Des premiers tests de couplage du complexe de terbium ont été réalisés sur le peptide test C(StBu)GKKFTKK(NH<sub>2</sub>)EELQKHAKTHTG-RinkAmide, par activation au PyBOP et à l'HCTU dans le DMF en présence de DIEA. Après déprotection au TFA, le peptide obtenu analysé par ESI-MS a été identifié comme étant le produit de départ. L'inefficacité de ce couplage peut s'expliquer par une faible solubilité du complexe de terbium dans le DMF, et dans le DCM. Cette voie n'étant pas possible, la stratégie par couplage du pro-ligand correspondant à la stratégie A de la Figure 116 (p.106) a été envisagée.

### ii Couplage du pro-ligand DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>1</sub>

Le peptide F2[Cys(StBu)|Lys(DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>1</sub>)]-SEA<sup>off</sup> a été synthétisé avec greffage du pro-ligand DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>1</sub> dans des conditions de couplage peptidique classique (Figure 142). Le peptide obtenu après déprotection des chaînes latérales des acides aminés et oxydation du groupement SEA a été obtenu avec un rendement de 7,7% après purification (Figure 141).

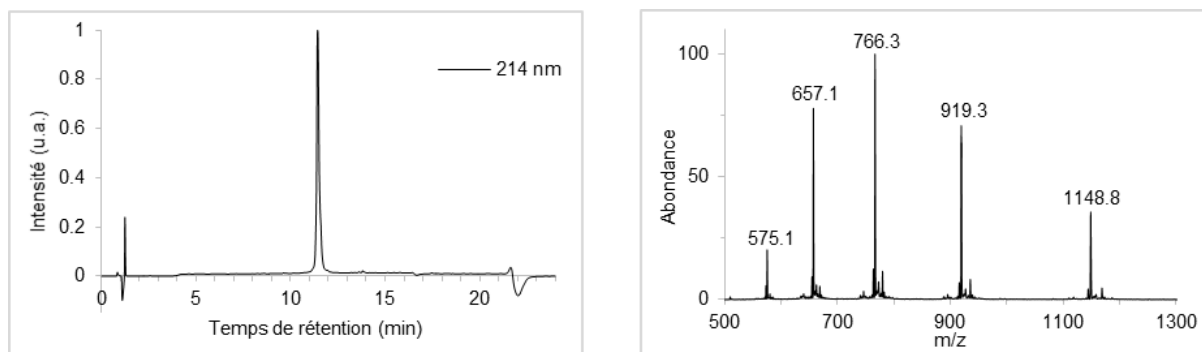
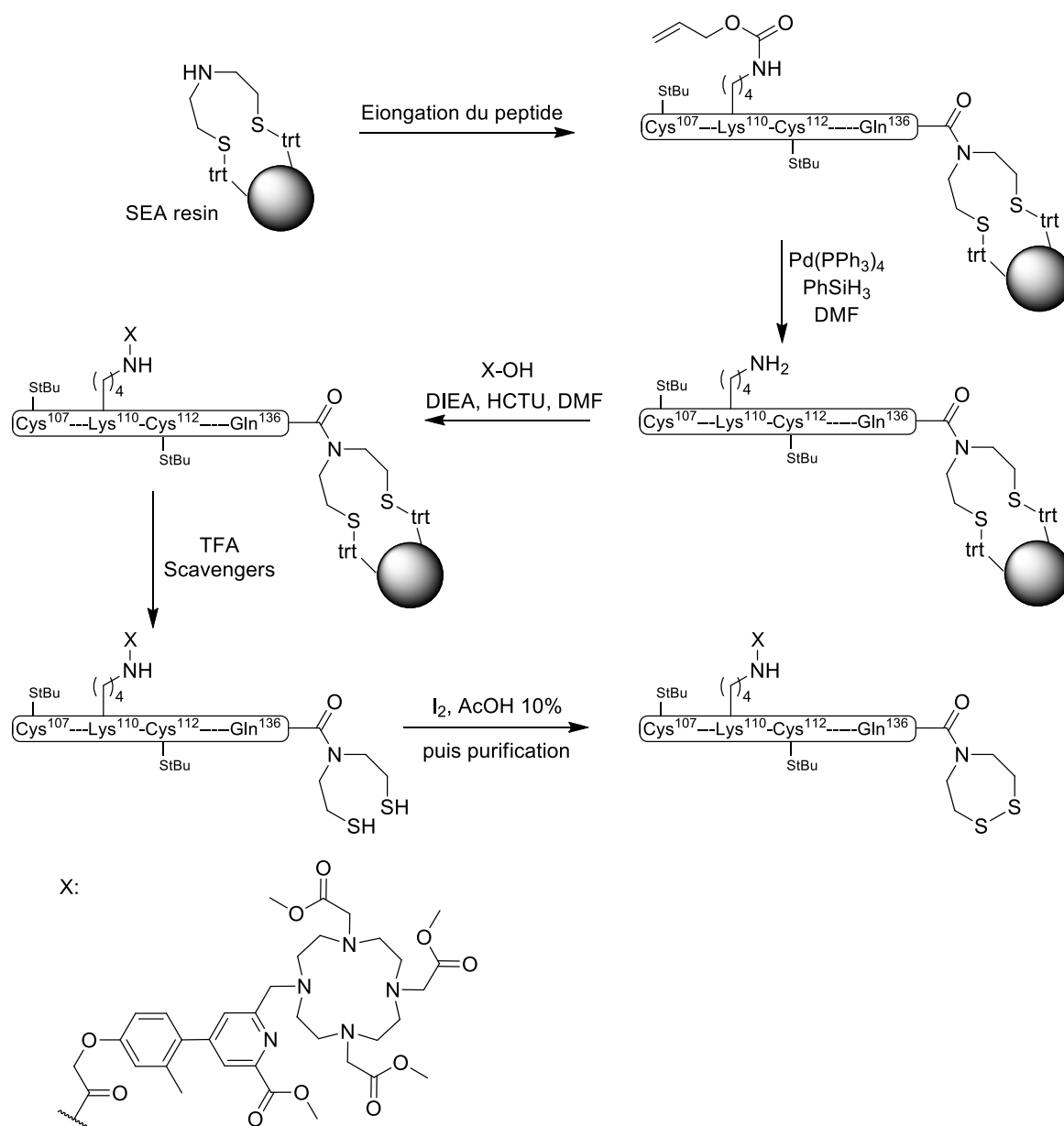


Figure 141: A gauche, chromatogramme HPLC du peptide F2[Cys(StBu)|Lys(DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>1</sub>)]-SEA<sup>off</sup> ; A droite, spectre ESI-MS (masse calculée : 1148,07 [M+4H]<sup>4+</sup>, 918,66 [M+5H]<sup>5+</sup>, 765,71 [M+6H]<sup>6+</sup>, 656,47 [M+7H]<sup>7+</sup>, 574,81 [M+8H]<sup>8+</sup> pour M = C<sub>200</sub>H<sub>319</sub>N<sub>59</sub>O<sub>53</sub>S<sub>6</sub>)



### iii Hydrolyse

L'étape suivante a été l'hydrolyse des esters méthyliques du ligand couplé au peptide (Figure 143).

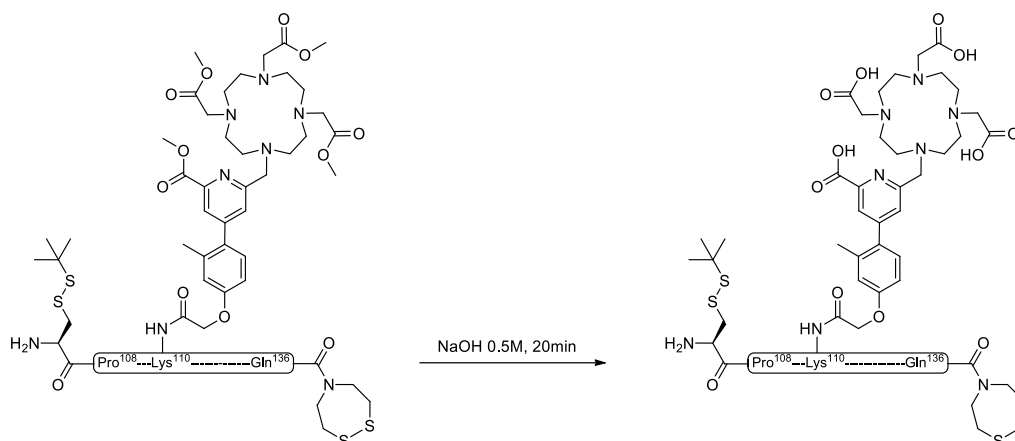


Figure 143: Schéma de l'étape d'hydrolyse du ligand permettant l'obtention du peptide F2[Cys(StBu)]Lys(DO3A<sup>pic</sup>Ar<sub>1</sub>)]-SEA<sup>off</sup>

L'hydrolyse a été testée sur un peptide de séquence C(StBu)GKKFTKK-(DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>1</sub>)EELQKHAKTHTG-NH<sub>2</sub> dans une solution de NaOH 1M pendant 10 min avec suivi de la réaction par HPLC. L'analyse ESI-MS a révélé la formation du produit attendu de masse M, avec un autre produit de masse M+1. Ce sous-produit provient probablement de l'hydrolyse de la fonction amide présente sur la chaîne latérale de la glutamine. Des conditions d'hydrolyse plus « douces » ont donc été nécessaires afin de ne pas dégrader le peptide.

Différentes conditions ont été testées par Ji-Hyung CHOI au laboratoire sur un peptide de séquence CGKKFTKK(DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>1</sub>)EELQKHAKTHTG-NH<sub>2</sub>, en changeant la concentration en NaOH (1M, 0,5M, 0,1M), ainsi que le temps de réaction (5, 10, et 15min) sur le peptide à une concentration de 10mg/mL. Pour une concentration de 0,1M de soude, l'hydrolyse reste très incomplète, même après 15 min de réaction. A une concentration de 1M, l'hydrolyse des amides est observée, même après seulement quelques minutes de réaction. Il faut donc utiliser une concentration intermédiaire de 0,5M pour laquelle différents temps ont été testés. A cette concentration l'analyse ESI-MS montre que l'hydrolyse des esters est totale après 5 min, sans traces d'hydrolyse des amides.

Ces conditions ont été testées sur le peptide F2[Cys(StBu)]Lys(DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>1</sub>)]-SEA<sup>off</sup>, cependant ce dernier s'avère insoluble dans la soude à 0,5M. Des conditions moins basiques ont alors été envisagées en testant l'hydrolyse à 0,1M durant des temps plus longs. Après 30 min de réaction l'hydrolyse restait très incomplète et le peptide toujours insoluble. Nous sommes donc revenus à une concentration de 0,5M en NaOH pour hydrolyser les esters méthyliques. Finalement, quelques tests sur ce peptide avec suivi de la réaction par HPLC ont montré que

malgré sa faible solubilité dans la soude à 0,5M, la réaction avait lieu et que les esters méthyliques étaient entièrement hydrolysés après 20 min de réaction, sans aucunes traces de sous-produits, même si le milieu reste inhomogène. Il semble donc que le peptide soit suffisamment soluble, pour consommer le produit initial et former le produit attendu.

Le protocole retenu a donc été de suspendre le peptide dans du NaOH 0,5M pendant 20 min, puis de dissoudre le produit formé par acidification du milieu réactionnel jusqu'à pH 4. Le chromatogramme du peptide F2[Cys(StBu)|Lys(DO3ApicAr<sub>1</sub>)]-SEA<sup>off</sup> purifié (18% de rendement) et son spectre ESI-MS sont donnés sur la Figure 144.

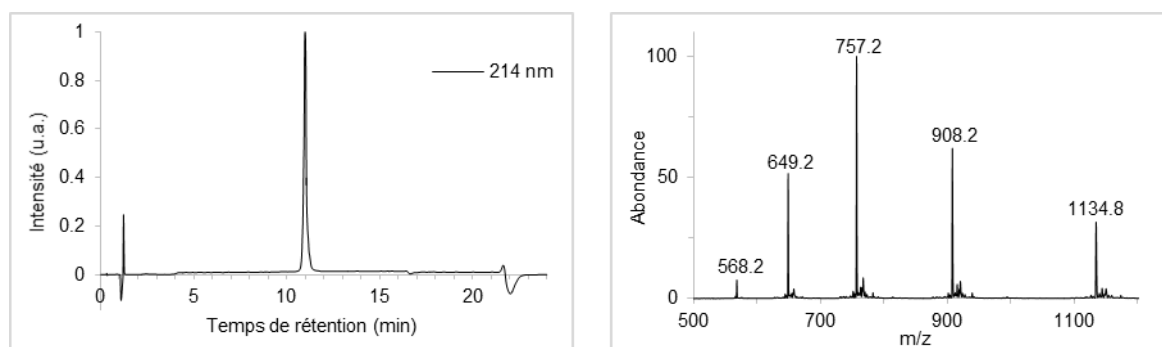


Figure 144: A gauche, chromatogramme HPLC de F2[Cys(StBu)|Lys(DO3ApicAr<sub>1</sub>)]-SEA<sup>off</sup> ; A droite, spectre ESI-MS (masse calculée : 1134,59 [M+4H]<sup>4+</sup>, 907,87 [M+5H]<sup>5+</sup>, 756,73 [M+6H]<sup>6+</sup>, 648,77 [M+7H]<sup>7+</sup>, 567,80 [M+8H]<sup>8+</sup> pour M = C<sub>196</sub>H<sub>311</sub>N<sub>59</sub>O<sub>53</sub>S<sub>6</sub>)

#### iv Métallation

Finalement, le peptide F2[Cys(StBu)|Lys(DO3ApicAr<sub>1</sub>)]-SEA<sup>off</sup> a été métallé à pH 9 en présence d'un excès de chlorure de terbium(III) (Figure 145), puis dessalé sur cartouche Oasis C-18 Waters et caractérisé par HPLC et ESI-MS (Figure 146). Le peptide F2[Cys(StBu)|Lys(Tb(DO3ApicAr<sub>1</sub>))]-SEA<sup>off</sup> a été obtenu avec un rendement de 97%. Le profil isotopique correspond à celui attendu.

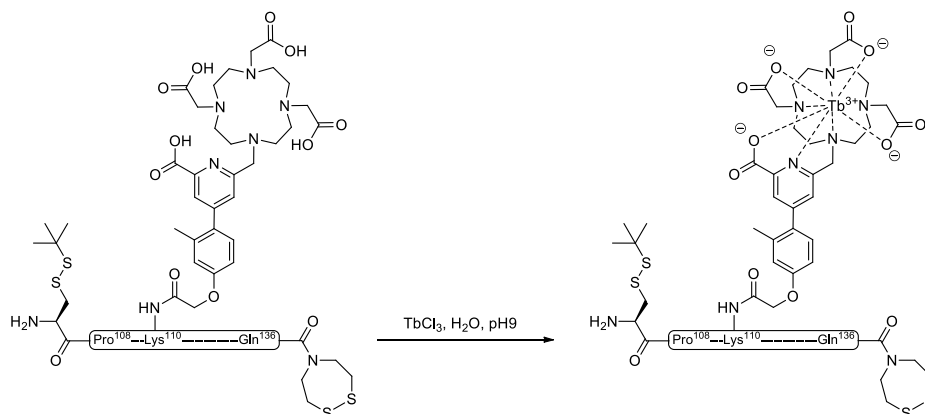


Figure 145: Etape de métallation pour l'obtention du peptide final F2[Cys(StBu)|Lys(Tb(DO3ApicAr<sub>1</sub>))]-SEA<sup>off</sup>

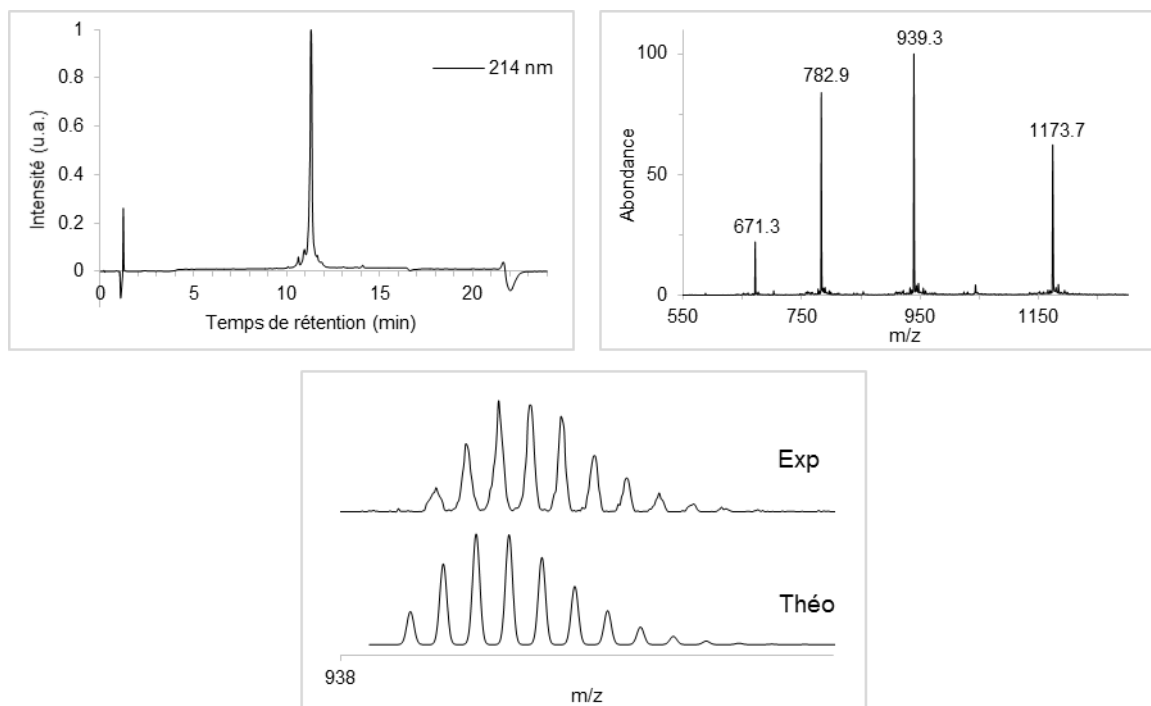


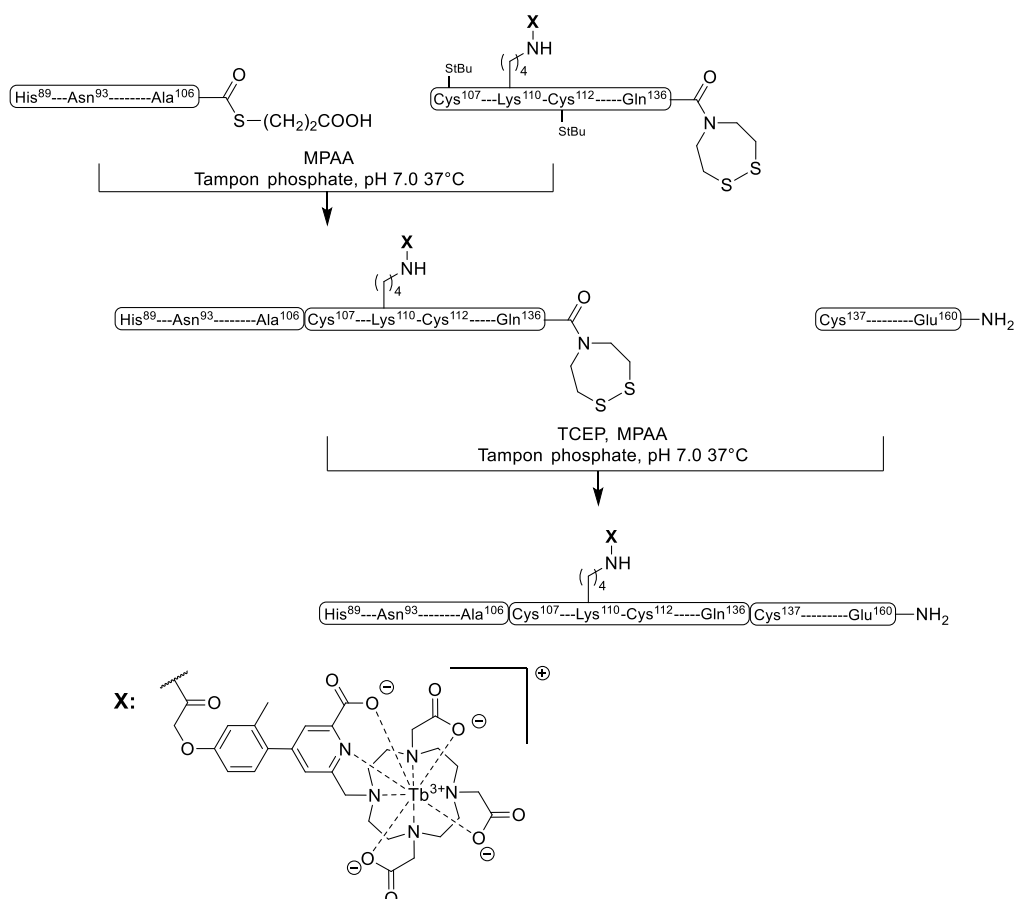
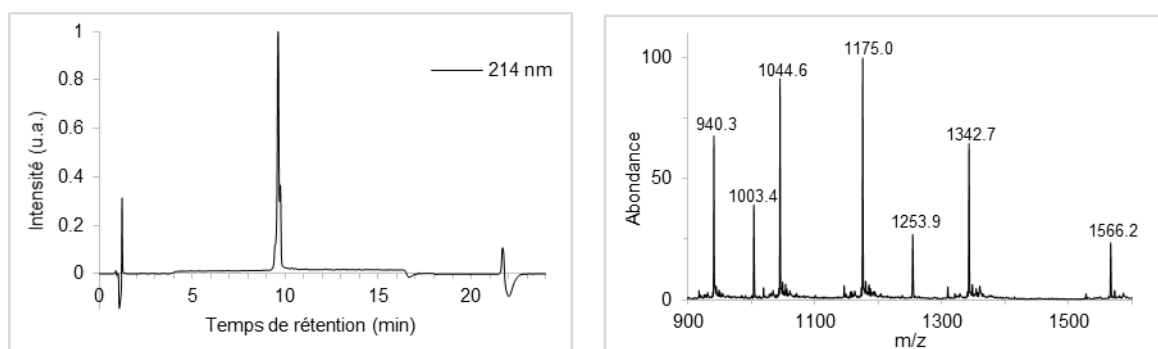
Figure 146: A gauche, chromatogramme HPLC du peptide F2[Cys(StBu)|Lys(Tb(DO3ApicAr<sub>1</sub>))]-SEA<sup>off</sup> ; A droite spectre ESI-MS (Masse calculée : 1173,57 [M+4H]<sup>4+</sup>, 939,06 [M+5H]<sup>5+</sup>, 782,71 [M+6H]<sup>6+</sup>, 671,04 [M+7H]<sup>7+</sup> pour  $M = C_{196}H_{308}N_{59}O_{53}S_6Tb$ ) ; En bas, massif isotopique expérimental et théorique (Formule brute expérimentale =  $C_{196}H_{308}N_{59}O_{53}S_6Tb$  ; Formule brute théorique =  $C_{196}H_{308}N_{59}O_{53}S_6Tb$ )

L'analyse LC-MS montre la présence du ligand métallé à 100%, sans trace de démétallation, confirmant la stabilité du complexe aux conditions d'analyse par HPLC à pH 2, au moins durant le temps de l'élution.

## B Assemblage

La protéine ZIF(Tb(DO3ApicAr<sub>1</sub>)) a été assemblée par deux réactions de ligations successives, impliquant en premier lieu le peptide F1-MPA décrit dans le chapitre II et le peptide F2[Cys(StBu)|Lys(Tb(DO3ApicAr<sub>1</sub>))]-SEA<sup>off</sup>, puis en second lieu le peptide obtenu issu de la première ligation, avec le peptide F3-NH<sub>2</sub> décrit dans le chapitre I (Figure 147). La protéine ZIF(Tb(DO3ApicAr<sub>1</sub>)) a été obtenue avec un rendement de 12,5%, et caractérisée par HPLC et ESI-MS (Figure 148).

Le fragment central F2 étant obtenu avec un faible rendement, il a été introduit en léger défaut lors de l'assemblage permettant la formation de la protéine F1-F3 (cf p.67) dont il n'a pu être séparé lors de l'étape de purification. Cet assemblage est visible sur l'analyse par ESI-MS et HPLC sous la forme d'un léger épaulement sur le pic correspondant au produit attendu.

Figure 147: Schéma d'assemblage one-pot pour la protéine ZIF(Tb(DO3ApicAr<sub>1</sub>))Figure 148: A gauche, chromatogramme HPLC de la protéine ZIF(Tb(DO3ApicAr<sub>1</sub>)) ; A droite, spectre ESI-MS (masse calculée : 1565,92 [M+6H]<sup>6+</sup>, 1342,36 [M+7H]<sup>7+</sup>, 1174,69 [M+8H]<sup>8+</sup>, 1044,28 [M+9H]<sup>9+</sup>, 939,95 [M+10H]<sup>10+</sup>, pour M = C<sub>399</sub>H<sub>626</sub>N<sub>129</sub>O<sub>111</sub>S<sub>7</sub>Tb correspondant à la protéine ZIF(Tb(DO3ApicAr<sub>1</sub>)); et 1253,71 [M+4H]<sup>4+</sup>, 1003,17 [M+5H]<sup>5+</sup>, pour C<sub>215</sub>H<sub>343</sub>N<sub>71</sub>O<sub>58</sub>S<sub>5</sub> correspondant à l'assemblage F1-F3)

### C Métallation au zinc

La quantité de protéine obtenue a été déterminée par mesure de son absorbance dans un tampon HEPES 10 mM pH 7,5, avec un coefficient d'extinction molaire ( $\epsilon$ ) égal à 12 000 M<sup>-1</sup>cm<sup>-1</sup> à 302 nm. Une quantité de ZnCl<sub>2</sub> correspondant à un léger excès (2,2 éq. par rapport à la protéine) a été ajoutée, puis la protéine ZIF<sup>Zn</sup>(Tb(DO3ApicAr<sub>1</sub>)) a été conservée au congélateur.

## 2) Etude de la luminescence

### A Système FRET du terbium vers la lissamine

#### i Etude du mélange $ZIF^{Zn}(Tb(DO3ApicAr_1)) / ZIF^{Zn}(lissamine)$

Les solutions de protéines  $ZIF^{Zn}(Tb(DO3ApicAr_1))$  et  $ZIF^{Zn}(lissamine)$  ont été diluées dans un mélange HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM) en rapport 1:1 à une concentration de 1  $\mu$ M.

Le spectre d'émission mesuré en excitant à  $\lambda_{ex}=306$  nm est présenté sur la Figure 149 (bas). Il montre les raies caractéristiques des transitions du terbium  $^5D_4 \rightarrow ^7F_n$  ( $n=3$  à 6) à 490, 545, 585 et 623 nm, avec superposition d'une bande d'émission large, de 560 à 700 nm, qui correspond à l'émission de la lissamine. Ceci est confirmé par le spectre d'excitation enregistré en détectant à  $\lambda_{em} = 602$  nm, longueur d'onde où le terbium n'émet pas. Ce spectre correspond bien au spectre d'absorption de la lissamine. Il n'est pas étonnant de détecter l'émission de la lissamine en excitant à 306 nm puisque ce fluorophore absorbe faiblement à cette longueur d'onde. Quant au spectre d'excitation du terbium ( $\lambda_{em} = 545$  nm), il correspond bien à l'absorption de l'antenne picAr<sub>1</sub> (cf Chapitre III).

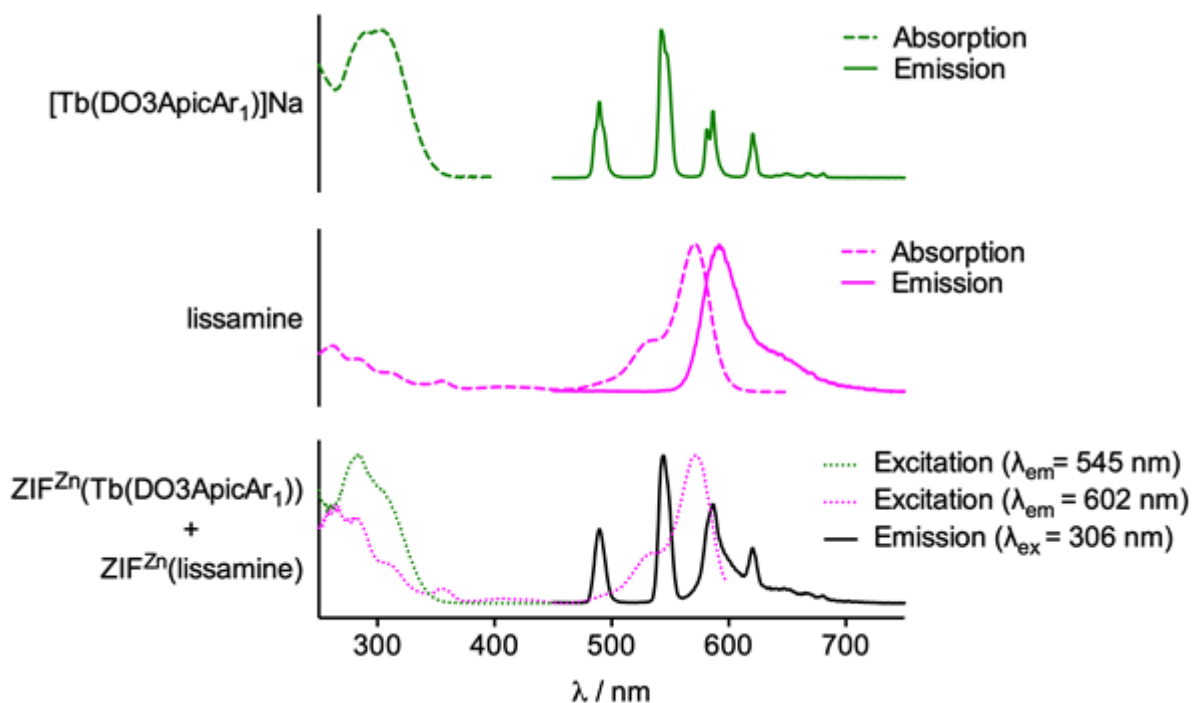


Figure 149 : (Haut) : Spectres d'absorption et d'émission ( $\lambda_{ex}=306$  nm) du complexe  $[Tb(DO3ApicAr_1)]Na$ . (Milieu) Spectres d'absorption et d'émission ( $\lambda_{ex}=580$  nm) de la lissamine. (Bas) Spectre d'émission ( $\lambda_{ex}=306$  nm) et spectres d'excitation ( $\lambda_{em}=545$  nm et 602 nm) du mélange  $ZIF^{Zn}(Tb(DO3ApicAr_1))/ZIF^{Zn}(lissamine)$  (1:1) dans HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM). Les deux protéines sont à une concentration de 1  $\mu$ M.

Le spectre d'émission ( $\lambda_{\text{ex}}=306$  nm) du mélange  $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{lissamine})$  a également été enregistré en temps résolu (délai de 100  $\mu\text{s}$ ). Il est représenté sur la Figure 150. On retrouve les mêmes contributions pour le terbium et la lissamine, mais la proportion d'émission de la lissamine est beaucoup plus faible. Le fait de détecter l'émission de la lissamine en temps résolu avec un délai de 100  $\mu\text{s}$  indique que cette émission a un temps de vie bien supérieur à quelques nanosecondes comme c'est le cas pour un fluorophore organique. Ceci suggère qu'il y a tout de même un peu de transfert d'énergie du terbium à la lissamine en l'absence d'ADN aux concentrations utilisées (1  $\mu\text{M}$  en protéine), à l'instar de ce que nous avons déjà vu avec le FRET entre fluorophores organiques. Ceci est confirmé par le temps de vie du terbium (2,14 ms) pour le mélange  $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{lissamine})$ , légèrement plus faible que celui de la protéine  $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))$  seule (2,35 ms). L'efficacité du transfert est estimée à 9 % par la relation

$$E = 1 - \tau_{\text{DA}}/\tau_{\text{D}}$$

où  $\tau_{\text{D}}$  est le temps de vie du donneur en absence d'accepteur et  $\tau_{\text{DA}}$  le temps de vie du donneur en présence d'accepteur.

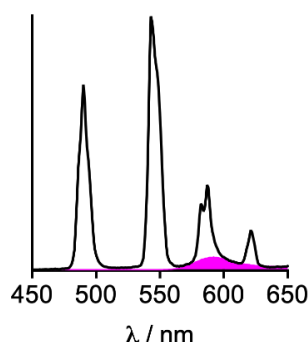


Figure 150 : Spectre d'émission résolu en temps (délai 100  $\mu\text{s}$ ) du mélange  $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{lissamine})$  (1:1) dans HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM). Les deux protéines ont une concentration de 1  $\mu\text{M}$ . L'émission due à la lissamine est mise en évidence en rose.

#### ii Interaction du couple $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{lissamine})$ avec son ADN cible

Avant d'étudier l'interaction du couple  $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{lissamine})$  avec l'ADN, nous avons étudié le comportement de la protéine  $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))$  seule vis-à-vis de l'ADN. Cette protéine a été dissoute (1  $\mu\text{M}$ ) dans un mélange HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM) et les spectres d'émission ainsi que les temps de vie du terbium ont été enregistrés ( $\lambda_{\text{ex}} = 306$  nm) en absence puis en présence du duplex d'ADN 38-mère cible de séquence 5'-GGTTGCAGTGGGCGCGCCACAGTACTTGAACGTAACG-3' (1, puis 2 éq.). Nous n'avons pas observé de modification de l'intensité du spectre d'émission

ni du temps de vie du terbium (2,35 ms), indiquant que la présence du duplex n'affecte pas l'émission du terbium.

L'interaction du couple  $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{lissamine})$  a été étudiée dans le même tampon avec le même duplex (concentration de 1  $\mu\text{M}$  pour chaque protéine). La Figure 151 (A) montre l'évolution des spectres d'émission ( $\lambda_{\text{ex}} = 306 \text{ nm}$ ) pour des quantités croissantes de duplex (0, 0,5, 1 et 2 éq.). On peut voir une diminution de l'émission du terbium à 0,5 puis 1 éq., suivit d'une stagnation à 2 éq. L'émission de la lissamine semble rester quasi-constant. En normalisant les spectres par rapport à l'émission du terbium (Figure 151 (B)), on peut voir qu'en proportion, l'émission de la lissamine augmente légèrement, suggérant clairement un transfert d'énergie entre le terbium et la lissamine. L'évolution du rapport du signal intégré de 560 à 750 nm (correspondant majoritairement à la lissamine) sur le signal intégré de 450 à 560 nm (correspondant au terbium) montre bien une saturation au-dessus de 1 éq. de duplex d'ADN (Figure 152). Tout ceci suggère la formation d'un complexe ternaire  $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{lissamine})/\text{ADN}$ .

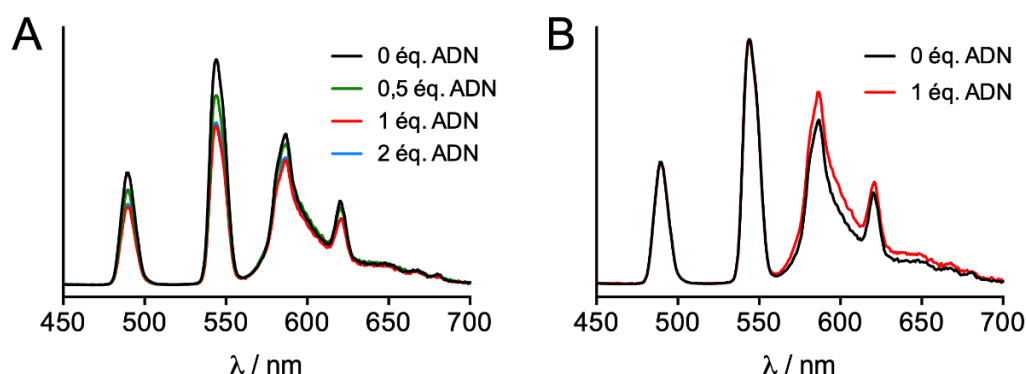


Figure 151 : (A) Spectres d'émission ( $\lambda_{\text{ex}}=306 \text{ nm}$ ) du mélange  $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{lissamine})$  (1:1), 1  $\mu\text{M}$  pour chaque protéine, pour des quantités croissantes de duplex d'ADN dans HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM). (B) Spectres sans et avec 1 éq. de duplex d'ADN normalisés par rapport à l'émission du terbium à 545 nm.

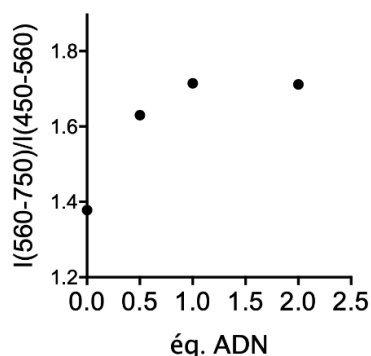


Figure 152 : Evolution du rapport des intensités d'émission (cumulés de 560 à 750 et de 450 à 560 nm) en fonction du nombre d'équivalents de duplex d'ADN.

Pour confirmer cela, les temps de vie du terbium ont été enregistrés ( $\lambda_{em} = 602$  nm). La décroissance de l'émission du terbium peut être modélisée par une monoexponentielle dans chaque cas (Figure 153 Gauche). Le temps de vie de luminescence du terbium, initialement à 2,14 ms passe à 2,01 ms en présence de 0,5 éq. d'ADN puis à 1,96 ms avec 1 éq. et finalement à 1,94 ms avec 2 éq. d'ADN (Figure 153 Droite). La diminution du temps de vie jusqu'à un équivalent suggère la formation du complexe ternaire et l'établissement d'un transfert d'énergie du terbium à la lissamine.

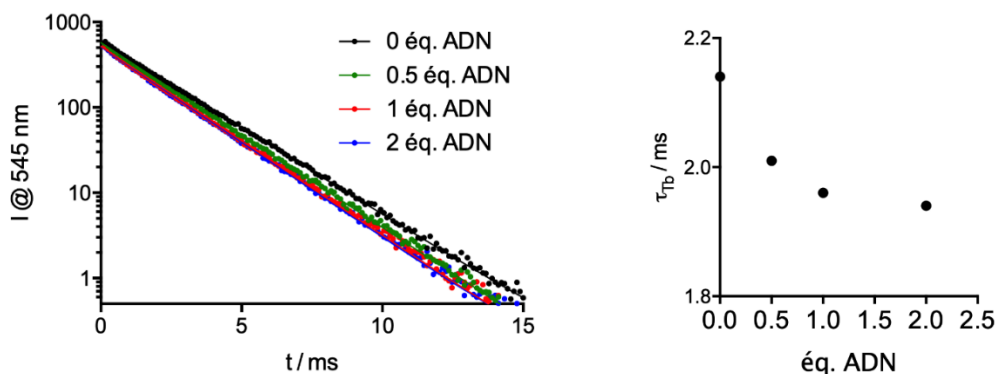


Figure 153 : (Gauche) Déclin du temps de vie de luminescence du terbium ( $\lambda_{ex}=306$  nm) dans le mélange  $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{lissamine})$  (1:1) ( $1 \mu\text{M}$  pour chaque protéine), pour des quantités croissantes de duplex d'ADN dans un mélange HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM). Les ajustements selon une décroissance monoexponentielle sont tracés en trait continu. (Droite) Evolution du temps de vie apparent de la luminescence du terbium à 545 nm en fonction du nombre d'équivalents en duplex d'ADN.

Les spectres d'excitation de la lissamine ( $\lambda_{em} = 602$  nm) ont été enregistrés pour 0 et 1 éq. de duplex d'ADN (Figure 154 A). Nous voyons que l'intensité autour de 300 nm est plus importante en présence de 1 éq. d'ADN alors que le reste du spectre reste inchangé. Le spectre différentiel obtenu en soustrayant le spectre d'excitation en absence d'ADN à celui en présence d'ADN montre clairement le signal supplémentaire vers 300 nm (Figure 154 B). Il correspond parfaitement au spectre d'excitation du terbium pour le complexe  $[\text{Tb}(\text{DO3ApicAr}_1)]\text{Na}$  et donc à l'absorption de l'antenne  $\text{picAr}_1$ . Ceci démontre qu'en présence d'ADN, la lissamine est davantage sensibilisée *via* l'antenne  $\text{picAr}_1$  et donc *via* le complexe de terbium. Le transfert d'énergie du terbium vers la lissamine est plus efficace en présence d'ADN qu'en l'absence d'ADN. Ceci valide le principe de notre sonde. Notons cependant qu'hormis l'hétéro-dimère  $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{lissamine})$ , les homo-dimères  $[\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))]_2$  et  $[\text{ZIF}^{\text{Zn}}(\text{lissamine})]_2$  peuvent se fixer sur l'ADN palindromique, les trois étant dans des proportions 50:25:25. Le système décrit ici n'est donc pas optimal, les spectres et temps de vie mesurés étant la résultante de ces différentes espèces. Il n'est donc pas possible d'évaluer l'efficacité réelle du transfert d'énergie ou la distance entre les chromophores à partir des temps de vie.

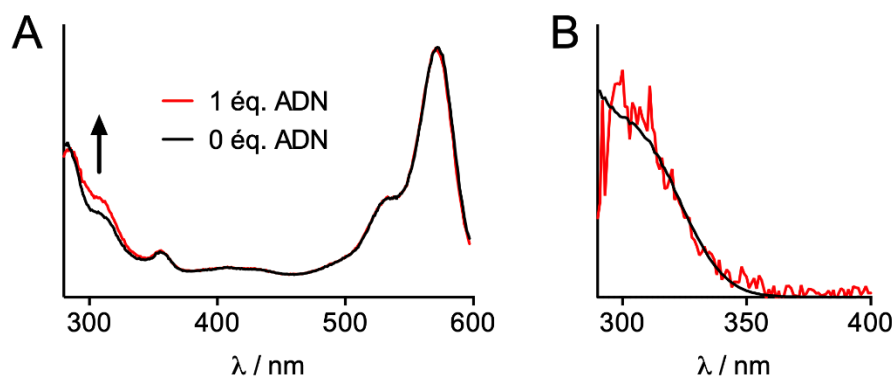


Figure 154 : (A) Spectres d'excitation ( $\lambda_{em}=602$  nm) du mélange  $ZIF^{Zn}(Tb(DO3ApicAr_1))/ZIF^{Zn}(lissamine)$  (1:1), 1  $\mu M$  pour chaque protéine, en absence et présence (1 éq.) du duplex d'ADN dans HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM). (B) Spectre différentiel (rouge) obtenu en soustrayant le spectre d'émission en absence d'ADN à celui en présence d'ADN et spectre d'excitation du complexe  $[Tb(DO3ApicAr_1)]Na$  (noir).

### iii Conclusion

Le principe d'une sonde reconnaissant un duplex d'ADN reposant sur un système FRET avec un complexe de lanthanide comme donneur et un fluorophore organique comme accepteur est maintenant établi. Le couple  $[Tb(DO3ApicAr_1)]/lissamine$  que nous venons de décrire est loin d'être optimisé. En effet :

- la lissamine absorbe fortement aux longueurs d'onde auxquelles le complexe de terbium absorbe, ce qui conduit à une forte excitation directe de la lissamine.
- le changement d'intensité relative entre le terbium et la lissamine est faible à cause d'un transfert d'énergie trop peu efficace, rendant la détection difficile et très peu sensible.

Nous avons alors étudié le couple terbium/SiR-méthyl.

### B Système FRET du terbium vers la SiR-méthyl

Les protéines  $ZIF^{Zn}(Tb(DO3ApicAr_1))$  et  $ZIF^{Zn}(SiR\text{-méthyl})$  ont été dissoutes dans un mélange HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM) en rapport 1:1 à une concentration de 1  $\mu M$  et étudiées comme précédemment. Les résultats sont assez similaires à ceux obtenus avec la lissamine et nous ne détaillerons donc pas autant l'étude.

Le spectre d'émission obtenu en excitant à 306 nm est représenté dans la Figure 155 (Bas). On reconnaît parfaitement l'émission du terbium avec ses 4 bandes caractéristiques à 490, 545, 585 et 623 nm et celle de la SiR-méthyl vers 675 nm. Cette attribution est confirmée sans ambiguïté par les spectres d'excitation obtenus en excitant à 545 et 675 nm. On peut noter que contrairement au couple Tb/lissamine, le couple Tb/SiR-méthyl présente des spectres d'émissions qui se recouvrent peu, ce qui sera un avantage pour la quantification du ratio

d'émission SiR-méthyl/terbium. Comme la lissamine, la SiR-méthyl absorbe à 306 nm et son émission à 675 nm est au moins en partie due à l'excitation directe de ce chromophore. Le spectre d'émission résolu en temps (délai = 100  $\mu$ s) laisse apparaître une faible contribution de la SiR-méthyl vers 675, indiquant une part de transfert provenant du terbium vers la SiR-méthyl (Figure 156 A ; 0 eq.). Ceci est confirmé par le temps de vie de luminescence du terbium (2,18 ms), inférieur à celui de la protéine ZIF<sup>Zn</sup>(Tb(DO3ApicAr<sub>1</sub>)) seule.

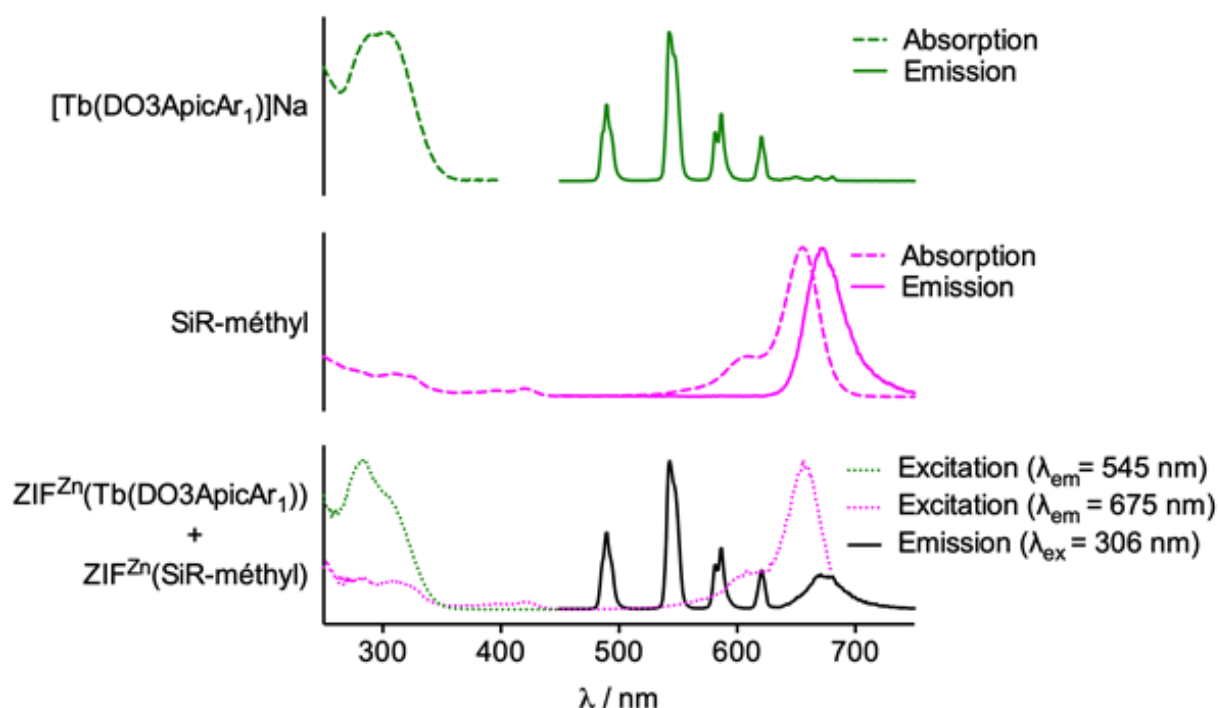


Figure 155 : (Haut) : Spectres d'absorption et d'émission ( $\lambda_{ex}=306$  nm) du complexe  $[Tb(DO3ApicAr_1)]Na$ . (Milieu) Spectres d'absorption et d'émission ( $\lambda_{ex}=660$  nm) de la SiR-méthyl. (Bas) Spectre d'émission ( $\lambda_{ex}=306$  nm) et spectres d'excitation ( $\lambda_{em}=545$  nm et 675 nm) du mélange  $ZIF^{Zn}(Tb(DO3ApicAr_1))/ZIF^{Zn}(SiR-méthyl)$  (1:1) dans HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM). Les deux protéines ont une concentration de 1  $\mu$ M.

L'ajout d'un équivalent de duplex d'ADN 38-mère cible provoque une augmentation du ratio d'émission SiR-méthyl/terbium, visible aussi bien sur les spectres d'émission résolus en temps (Figure 156 A) que sur les spectres d'émission totale sans délai (Figure 156 B), suggérant un transfert d'énergie plus important entre le terbium et la SiR-méthyl en présence d'ADN. Ceci est confirmé par la diminution du temps de vie de luminescence du terbium à 1éq. en ADN (2,04 ms) et les spectres d'excitation de la SiR-méthyl (Figure 157 A), ainsi que leurs différentiels (Figure 157 B).

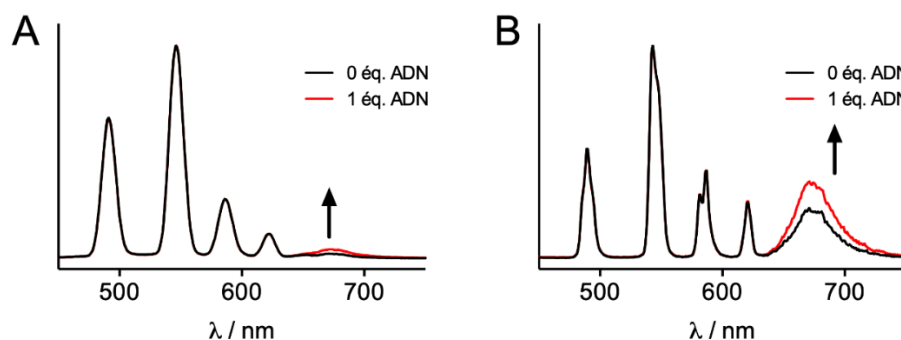


Figure 156 : Spectres d'émission normalisés pour le mélange  $ZIF^{Zn}(Tb(DO3ApicAr_1))/ZIF^{Zn}(SiR\text{-méthyl})$  (1:1) dans HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM) en absence (noir) et en présence (1 éq., rouge) du duplex d'ADN cible. (A) Emission résolue en temps (délai 100  $\mu$ s). (B) Emission totale, sans délai.

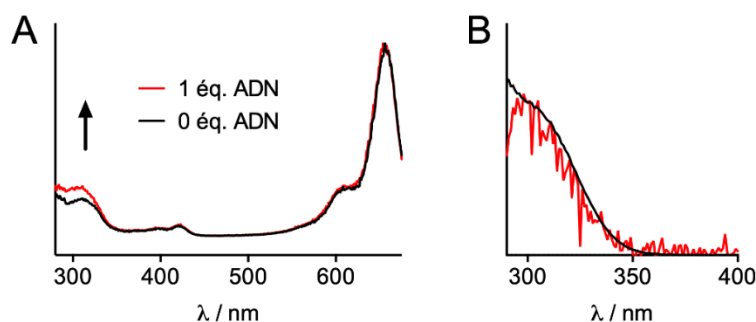


Figure 157 : (A) Spectres d'excitation ( $\lambda_{em}=602$  nm) du mélange  $ZIF^{Zn}(Tb(DO3ApicAr_1))/ZIF^{Zn}(SiR\text{-méthyl})$  (1:1), 1  $\mu$ M pour chaque protéine, en absence et présence (1 éq.) du duplex d'ADN dans HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0,1 g/L, 0,05 %, 100 mM). (B) Spectre différentiel (rouge) obtenu en soustrayant le spectre d'émission en absence d'ADN à celui en présence d'ADN et spectre d'excitation du complexe  $[Tb(DO3ApicAr_1)]Na$  (noir).

### 3) Conclusion

La détection de l'ADN basé sur l'étude d'un FRET du terbium vers la lissamine n'est pas très efficace à cause de l'excitation directe de la lissamine lors de l'excitation du complexe de terbium, conduisant à une forte contribution indépendante de la présence d'ADN au signal de la lissamine. La faible efficacité du transfert d'énergie entre le terbium et la lissamine dans ce système ne permet pas de compenser cela et d'avoir une détection sensible. En remplaçant la lissamine par la SiR-méthyl, qui présente un spectre d'émission à plus basse énergie, de 635 à 735 nm, on améliore la détection de l'ADN. Cependant l'excitation directe de la SiR-méthyl lors de l'excitation du complexe de terbium réduit ici encore la sensibilité du système. Il faudrait donc pouvoir utiliser une longueur d'onde d'excitation plus élevée, vers 350 nm, à laquelle l'absorption de la SiR-méthyl est plus faible. Le complexe d'euporium  $[Eu(DO3ApicAr_2)]$  ( $\lambda_{ex}=345$  nm) pourrait dans ce cas être un donneur parfaitement adapté.

### III FRET de l'euporium vers un fluorophore

Les paramètres photophysiques de l'euporium étant compatibles avec ceux de la SiR-méthyl pour un FRET d'un lanthanide vers un fluorophore, nous avons entrepris la synthèse d'une nouvelle protéine notée ZIF(Eu(DO3ApicAr<sub>2</sub>)).

#### 1) Travaux préliminaires

Avant de travailler avec le complexe [Eu(DO3ApicAr<sub>2</sub>)]<sup>-</sup> et de chercher à l'intégrer sur une protéine servant de sonde à ADN, nous avons travaillé sur d'autres systèmes incorporant le même type d'antenne « picoline-alcyne-aryl » permettant de sensibiliser l'euporium (Figure 158). Les détails du système n'étant pas en rapport avec le sujet de thèse, seules les observations et conclusions qui en ont été tirées seront présentées.

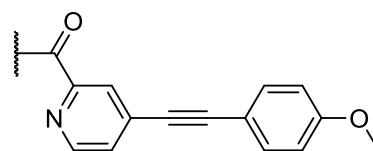


Figure 158: Structure de l'antenne "pyridine-alcyne-aryl" utilisée

##### i Ligation en conditions classiques

Une ligation entre un peptide cystéinyl et un peptide SEA<sup>off</sup> comportant un chromophore « picoline-alcyne-aryl » a été réalisée dans des conditions classiques (MPAA 100mM, TCEP 100mM, tampon phosphate 100mM, pH 7,5), puis analysée par spectrométrie de masse. Le produit de ligation a été obtenu avec une masse M+168, correspondant à l'addition d'une molécule de MPAA sur la fonction alcyne du chromophore (Figure 159). L'utilisation de MPAA pour les ligations avec des molécules comportant les dérivés de l'antenne « picoline-alcyne-aryl » est donc problématique en raison de la réaction d'hydrothiolation des triples liaisons, réaction connue et décrite dans la littérature<sup>[185]</sup>.

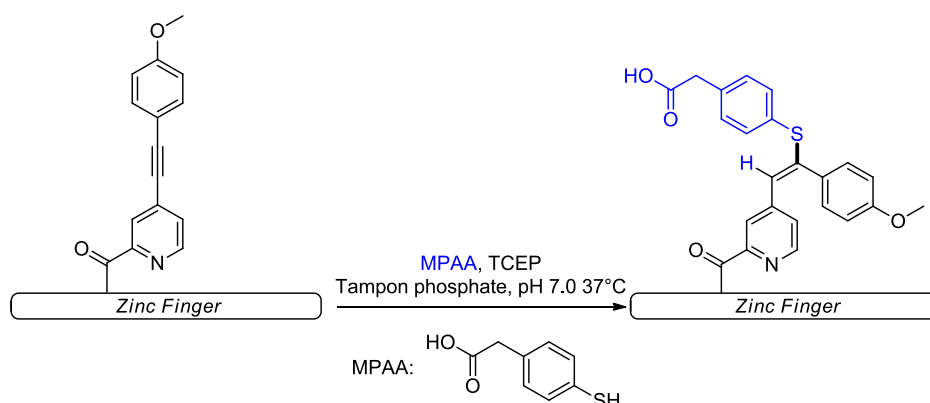


Figure 159: Addition du MPAA sur l'alcyne dans des conditions classiques de ligation

La diminution de la concentration de MPAA conduit à une cinétique de réaction plus lente sans empêcher la formation du produit d'addition de MPAA sur l'alcyne. L'utilisation d'autres thiols moins acides a été envisagée, en testant le MPA ou le MesNa, qui ont montrés la même réactivité vis-à-vis de l'alcyne. On en conclut donc que les chromophores comportant des fonctions alcyne ne doivent en aucun cas être mis en présence de thiols acide et que l'utilisation du complexe  $[\text{Eu}(\text{DO3ApicAr}_2)]^-$  est compromise.

### *ii Changement des conditions de ligation*

Pour contourner ce problème, nous avons repris les travaux de T. Yoshiya, qui a décrit la possibilité de substituer le MPAA par de l'imidazole<sup>[186]</sup> ou du 1,2,4-triazole<sup>[187]</sup>, ce dernier substitut permettant une réaction rapide, l'intermédiaire acyl-triazole étant moins sensible à l'hydrolyse.

L'introduction du peptide conjugué au chromophore « picoline-alcyne-aryl » dans les conditions de ligation décrite par T. Yoshiya (imidazole ou triazole 2,5M dans un tampon phosphate 100 mM pH 7,5), n'a montré aucune trace d'addition d'une quelconque molécule sur le peptide.

La ligation entre le peptide cystéinyl et le peptide SEA<sup>off</sup> comportant un chromophore « picoline-alcyne-aryl » dans les conditions de T. Yoshiya a permis d'assembler les deux fragments peptidiques sans réaction parasite (Figure 160).

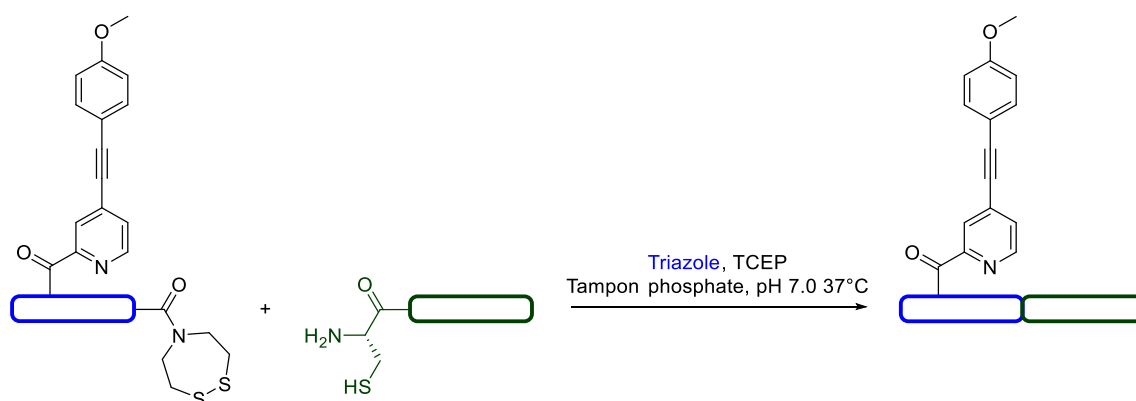


Figure 160: Ligation au triazole avec un peptide comportant un chromophore alcyné

### *iii Conclusion*

Pour mener à bien la synthèse de la protéine désirée, il faut donc pouvoir adapter la stratégie de double ligation one-pot décrite par O. Melnyk<sup>[115]</sup> avec ces nouvelles conditions. Pendant la première ligation, le MPAA sert lors de l'étape de transthioesterification pour former le

thioester de MPAA, mais il joue également le rôle de réducteur sélectif des groupements protecteur *StBu* des cystéines. Sans la déprotection de ces cystéines, aucune réaction de ligation ne peut avoir lieu. La stratégie de O. Melnyk n'est donc pas envisageable en substituant le MPAA par du triazole. La seconde ligation, elle, ne poserait aucun souci, car s'effectuant dans des conditions réductrices. Le talon d'Achille de cet assemblage réside donc dans la protection *StBu* de la cystéine en N-terminal du fragment central.

## 2) Mise en place d'une nouvelle voie de synthèse de protéines

Une nouvelle stratégie est nécessaire, évitant l'utilisation d'une cystéine protégée par un groupement *StBu* pour la synthèse du fragment 2. Pour rappel, la protection de la cystéine finale sur le fragment 2 permet d'éviter son oxydation lors du passage du SEA<sup>on</sup> au SEA<sup>off</sup>. Une alternative pourrait être de changer de groupement protecteur en utilisant par exemple une thioproline. Cependant, comme observé au chapitre I (p.64), le fragment 2 comportant une cystéine cyclisée en N-terminal n'est que très peu soluble lors de l'étape d'oxydation en milieu aqueux acide.

Une alternative à la stratégie SEA est donc nécessaire en changeant la fonction située en C-terminal du fragment 2. Pour cela, nous nous sommes intéressés au crypto-thioester N-HNB-Cys(*StBu*) développé par l'équipe de V. Aucagne<sup>[188]</sup>, dont le principe reste le même que pour le crypto-thioester SEA tout en étant bien moins coûteux.

### A Utilisation de la résine HNB

#### i Principe

Le crypto-thioester N-HNB-Cys(*StBu*) (Figure 161) est composé d'une cystéine dont le thiol est protégé par une fonction *StBu*, et comportant un groupement HydroxyNitroBenzyl (HNB) sur sa fonction amine. Cette cystéine est liée à la résine par une glycine en C-terminal, servant de bras espaceur. Le peptide est synthétisé à partir de l'amine secondaire de la cystéine protégée.

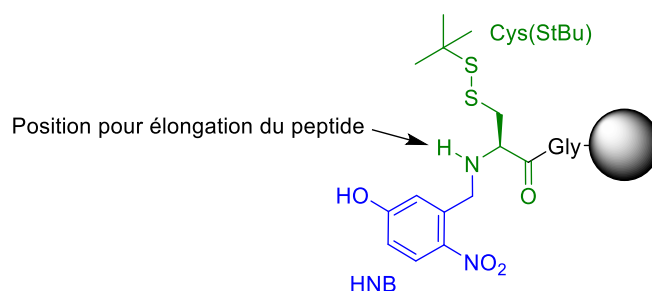


Figure 161: Structure du crypto-thioester HNB

La fonction N-HNB-Cys(StBu) obtenue après décrochage de sa résine est non réactive et est appelée HNB<sub>off</sub> (Figure 162). En conditions réductrices, le groupement protecteur StBu est éliminé conduisant à la fonction N-HNB-Cys(SH) appelée HNB<sub>on</sub>, réactive pour la formation d'un thioester. En conditions basiques un premier transfert d'acyl N-vers-S est alors permis, via la formation d'un cycle à 5 chaînons conduisant à la formation d'un thioester en C-terminal du peptide. Suite à l'introduction d'un peptide contenant une cystéine en N-terminal, une transthioesterification a alors lieu, suivie d'un second transfert d'acyl N-vers-S conduisant à la formation d'un lien chimique natif.

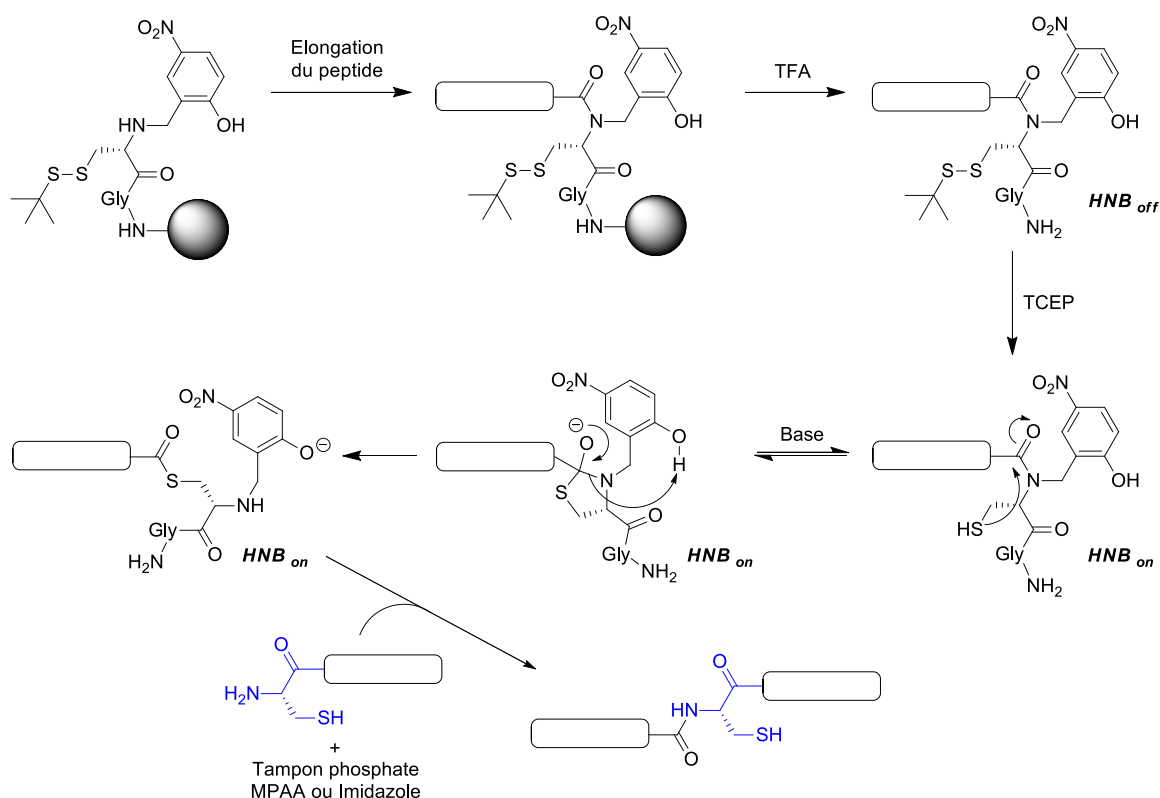


Figure 162: Synthèse d'un peptide sur résine HNB, formation du thioester, et ligation native avec un autre peptide

Contrairement à la stratégie SEA, une fois le peptide décroché de sa résine et purifié, il peut être utilisé en l'état sans réaction post-décrochage (comme c'est le cas avec la stratégie SEA, qui requiert une étape d'oxydation). Les groupements protecteurs des acides aminés introduits lors de la synthèse peptidique peuvent donc être ceux classiquement utilisés en stratégie Fmoc/tBu, permettant l'introduction d'une cystéine N-terminale classique. La synthèse du crypto-thioester est par ailleurs plus facile à mettre en œuvre car réalisable sur résine PEG.

### ii Synthèse du groupement N-HNB-Cys(StBu)

Le crypto thioester est synthétisé sur résine en 2 étapes (Figure 163) en suivant le protocole décrit par le groupe de V. Aucagne<sup>[113]</sup>. Premièrement, la glycine et la cystéine(StBu) sont introduites par couplage peptidique classique, puis une amination réductrice est réalisée par réaction de para-hydroxynitrobenzaldéhyde sur l'amine secondaire de la cystéine, suivi d'un traitement au NaBH<sub>3</sub>CN.

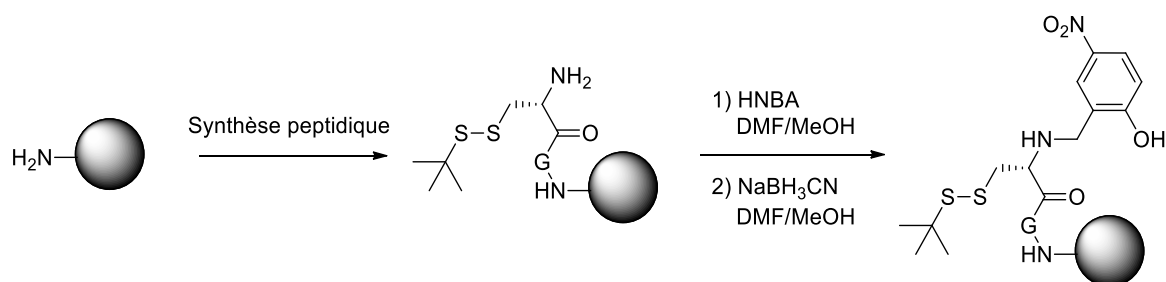


Figure 163: Schéma de synthèse du groupement N-HNB-Cys(StBu)

### iii Réactivité du phénol

Pendant l'élargissement du peptide chaque acide aminé est couplé en N-terminal du peptide par N-acylation, mais la fonction phénol est également acylée, parfois partiellement (Figure 164). Cependant, l'étape de traitement à la pipéridine permet de régénérer la fonction phénol. Une autre possibilité est de réaliser un pré-traitement à l'hydroxylamine, qui permet de libérer sélectivement le phénol en gardant intact le groupement protecteur Fmoc en N-terminal du peptide<sup>[113]</sup>. Ceci présente un intérêt pour le suivi de l'étape de déprotection par spectrophotométrie UV.

Pour éviter de former une espèce O-acylé de manière irréversible, certaines conditions de couplage sont à proscrire lors de l'élargissement du peptide et du greffage des chromophores. Par exemple l'utilisation de PyBOP comme activateur conduit à l'inactivation du phénol par la formation d'une espèce «phénoxy(tris-pyrrolidino-phosphonium)». Selon de protocole décrit par V. Aucagne nous avons utilisés l'HCTU comme activateur tout au long des synthèses des peptides comportant la fonction N-HNB-Cys(StBu).

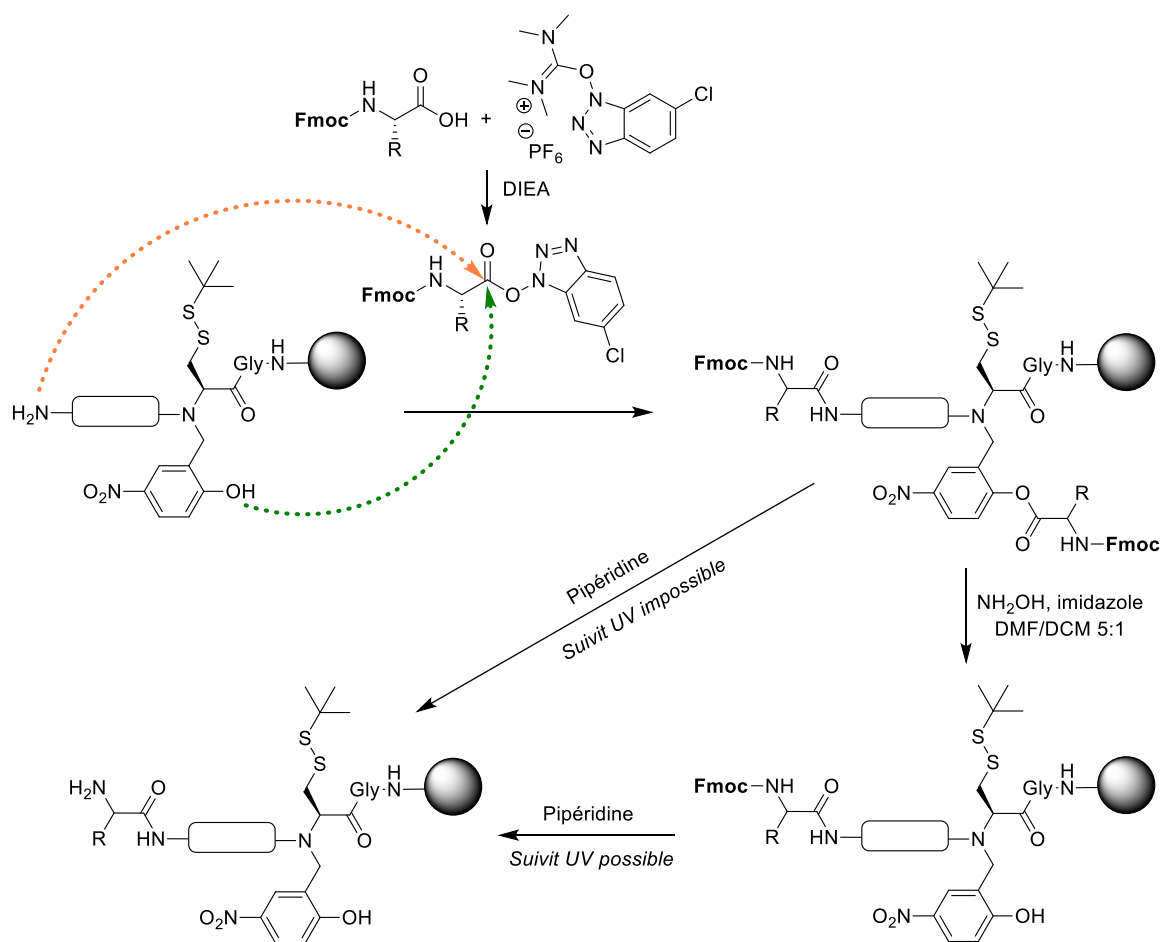


Figure 164: N-acylation (en orange) et O-acylation (en vert), et les deux voies de déprotection du groupement protecteur Fmoc

Le groupement hydroxyle du phénol joue par ailleurs un rôle d'assistance lors du couplage du premier acide aminé de la séquence peptidique sur l'amine secondaire de la cystéine de la fonction N-HNB-Cys(StBu)<sup>[113]</sup>. Le groupement hydroxyle est O-acylé par l'acide aminé à coupler, puis la liaison amide est formée par un transfert d'acyl intramoléculaire O-vers-N (Figure 165).

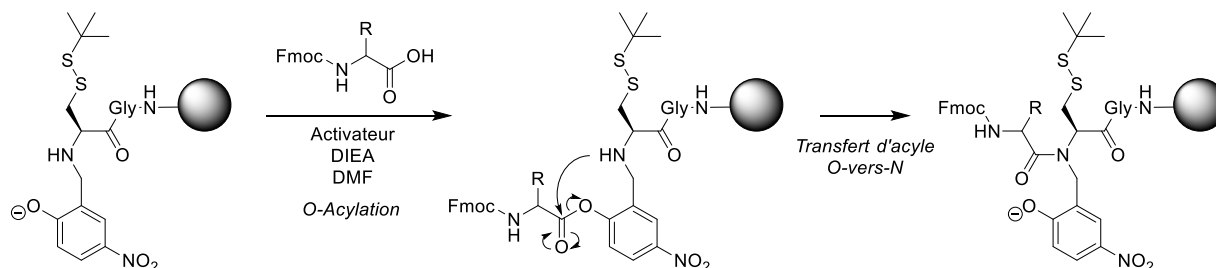


Figure 165: Assistance du phénol pour le couplage du premier acide aminé

## B Test des stratégies d'assemblages

Deux approches sont possibles pour la synthèse de protéines en trois fragments peptidiques avec une fonction N-HNB-Cys(StBu) sur le fragment 2. La première est un assemblage de l'extrémité N-vers-C, et la seconde de l'extrémité C-vers-N. Ces deux stratégies ont été testées pour la synthèse d'une protéine de référence : la protéine ZIF(Cs124|DOTA[Tb]), dont le découpage est resté le même que celui utilisé au premier chapitre.

### i Stratégie d'assemblage de C-vers-N

#### (a) Principe

Le schéma de la stratégie d'assemblage de l'extrémité C-vers-N est donné en Figure 166. Le peptide central F2 et le peptide F3 sont assemblés par une première réaction de ligation en milieu réducteur permettant de passer de la forme  $\text{HNB}^{\text{off}}$  à la forme  $\text{HNB}^{\text{on}}$ , puis la cystéine N-terminale de l'assemblage F2-F3 obtenu est déprotégée. L'introduction du peptide F1 déclenche une seconde réaction de ligation native conduisant à l'assemblage F1-F2-F3. 197

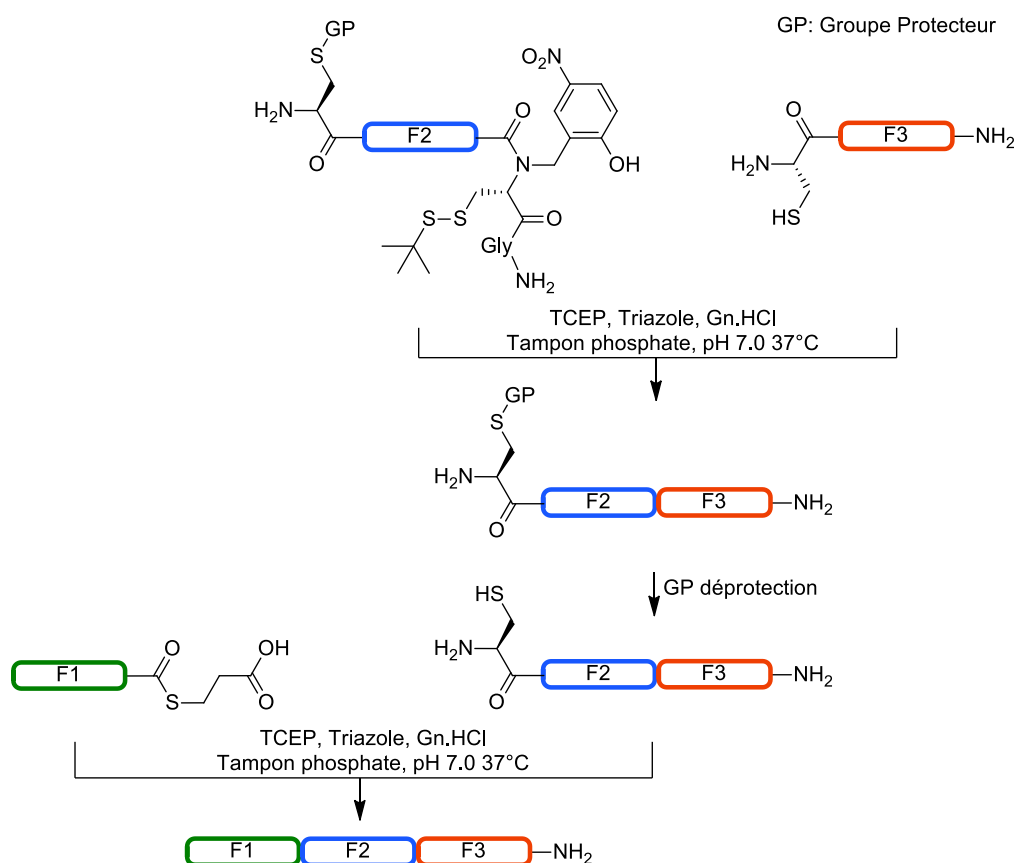


Figure 166: Schéma d'assemblage de trois fragments en stratégie C vers N

La protection de la cystéine N-terminale du peptide F2 peut être réalisée par l'introduction d'une thiazolidine dont la déprotection s'effectue à pH 4 en présence de méthoxylamine.

La synthèse de la protéine ZIF(Cs124|DOTA[Tb]) par la stratégie d'assemblage C-vers-N requiert trois fragments peptidiques :

- un peptide F1 comportant un ester de MPA en C-terminal (F1-MPA),
- un peptide F2 comportant une thiazolidine en N-terminal et une fonction N-HNB-Cys(StBu) en C-terminal (F2[Thz|Glu(Cs124)]-HNB<sup>off</sup>),
- un peptide F3 comportant une cystéine en N-terminal (F3-NH<sub>2</sub>).

Seul le fragment 2 doit être synthétisé, les autres peptides restant les mêmes que ceux utilisés au premier chapitre.

#### (b) Synthèse du peptide F2[Thz|Glu(Cs124)]-HNB<sup>off</sup>

Le peptide F2[Thz|Glu(Cs124)]-HNB<sup>off</sup> a été synthétisé en stratégie Fmoc/tBu sur résine RinkAmide NovaPEG. Les deux premiers acides aminés glycine, et cystéine(StBu) ont été introduits par synthèse manuelle suivie d'une amination réductrice réalisée par réaction de 2-hydroxy-5-nitrobenzaldéhyde, puis de NaBH<sub>3</sub>CN. La suite de la synthèse a été réalisée comme pour le peptide F2[Cys(StBu)|Glu(Cs124)]-SEA<sup>off</sup>. Le peptide F2[Thz|Glu(Cs124)]-HNB<sup>off</sup> a été purifié par HPLC préparative (rendement 9,1%), lyophilisé, et caractérisé par HPLC et ESI-MS (Figure 167).

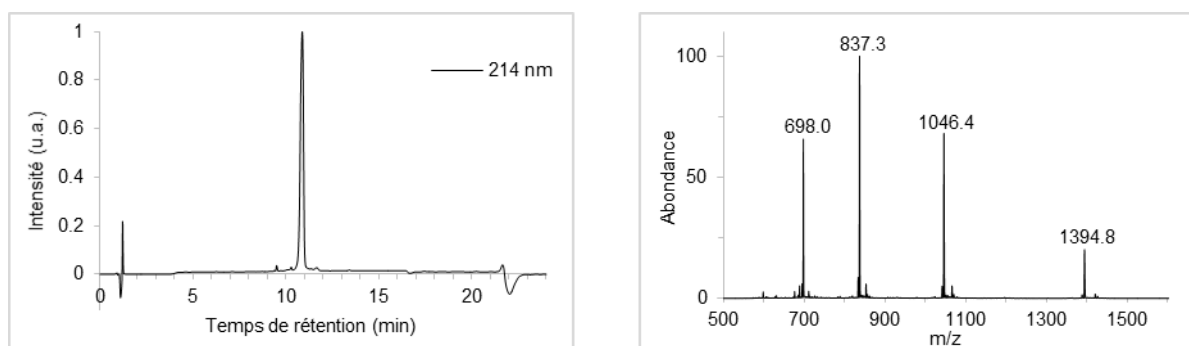


Figure 167: A gauche, chromatogramme HPLC du peptide F2[Thz|Glu(Cs124)]-HNB<sup>off</sup> ; A droite, spectre ESI-MS (masse calculée : 1394,59 [M+3H]<sup>3+</sup>, 1046,19 [M+4H]<sup>4+</sup>, 837,16 [M+5H]<sup>5+</sup>, 697,80 [M+6H]<sup>6+</sup> pour M = C<sub>180</sub>H<sub>276</sub>N<sub>58</sub>O<sub>50</sub>S<sub>4</sub>)

#### (c) Assemblage

L'ensemble de la réaction d'assemblage (Figure 169) a été réalisée sous atmosphère inerte en boîte à gants, afin d'éviter l'oxydation des cystéines. La première ligation entre le peptide F3-NH<sub>2</sub> et le peptide F2[Thz|Glu(Cs124)]-HNB<sup>off</sup> a été effectuée à une concentration de 7 mM en peptide dans une solution de TCEP/triazole/Gn-HCl (200mM, 2,5 M, et 6 M, dans un tampon phosphate 100mM pH7,5) à 37°C. L'avancement de la réaction, qui s'est terminée après 64 h, a été suivi par HPLC (Figure 168). Le milieu réactionnel a été acidifié à pH 4 par ajouts de HCl 2 M, puis de la méthoxylamine a été ajoutée pour obtenir une concentration finale de 0,1 M

dans le milieu. La réaction de déprotection a été suivie par ESI-MS, les temps de rétention pour le peptide avec une thiazolidine ou une cystéine étant identiques. La déprotection a été complétée au bout de 2 h. Le pH de la solution a été ajusté à pH 7 par ajouts de NaOH 2 M, puis le peptide F1-MPA a été ajouté. Le milieu réactionnel a été agité à 37°C et l'avancement de la réaction, qui dure près de 19h, a été suivi par HPLC (Figure 168). Après 120 h, le peptide F1-MPA a été totalement consommé mais davantage de F1-F3 s'est formé rendant la purification plus délicate. Le spectre ESI-MS de la Figure 170 montre que la protéine F1-F3 n'a pas pu être éliminée complètement.

Le suivi de la réaction par HPLC a montré que les temps de rétention pour un même produit pouvaient varier jusqu'à 2 min. Ce décalage n'ayant jamais été observé lors de ligations utilisant le MPAA, nous en avons déduit un effet du triazole sur la rétention de la colonne HPLC, phénomène par ailleurs observé lors de ligations en présence d'imidazole. Nous avons remarqué qu'un rinçage de la colonne d'au moins 15 min avec un mélange 50 :50 TPA/TPB était nécessaire avant l'injection suivante. Ce traitement imparfait nous a permis de suivre les ligations par HPLC, malgré un décalage des temps de rétention dû aux quantités variables de triazole introduites sur la colonne lors des différentes injections. Les chromatogrammes du suivi de la ligation présentés en Figure 168 ont été corrigés par rapport aux temps de rétentions du fragment F3 et de l'assemblage F2-F3.

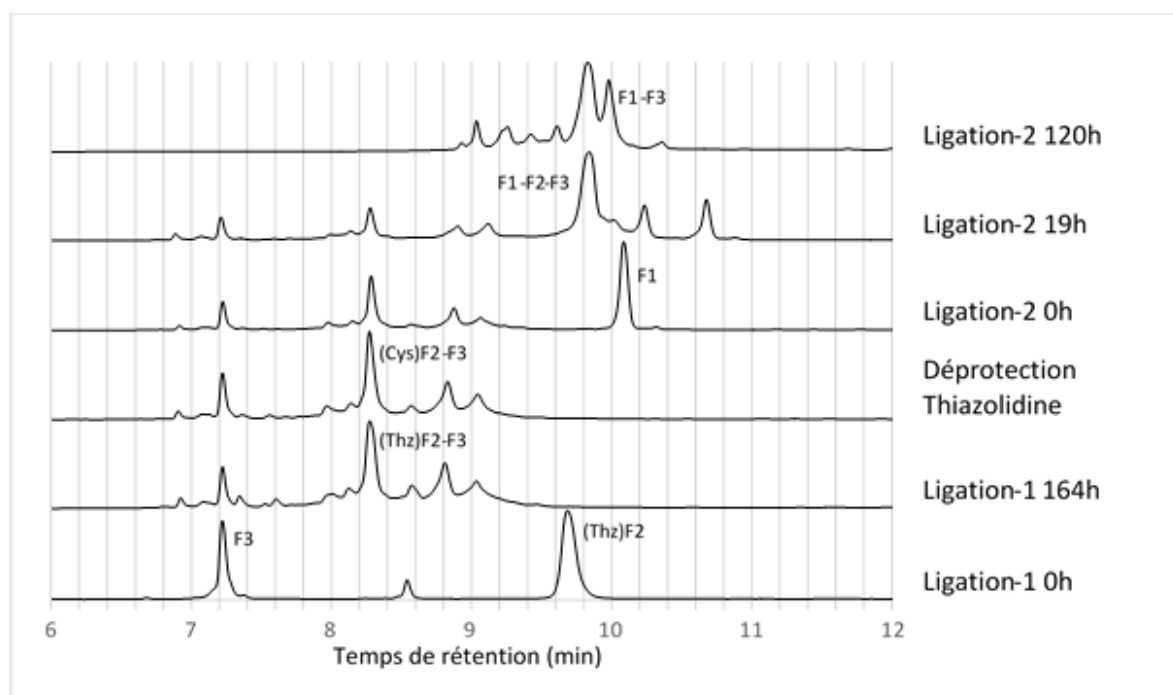


Figure 168: Suivi par HPLC de l'assemblage de la protéine ZIF(Cs124|DOTA[Tb])

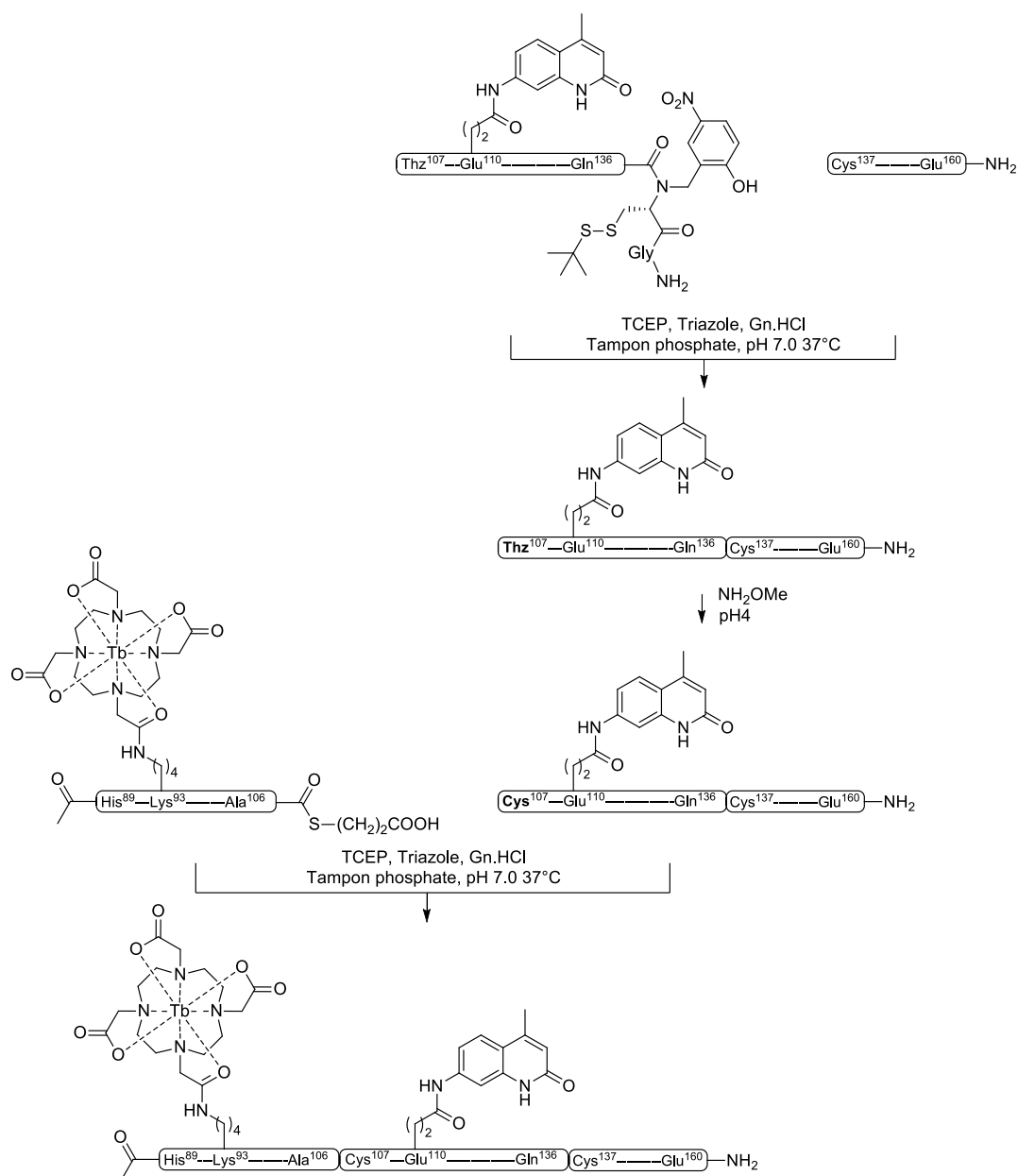


Figure 169: Schéma d'assemblage en stratégie HNB de C-vers-N de la protéine ZIF(Cs124|DOTA[Tb])

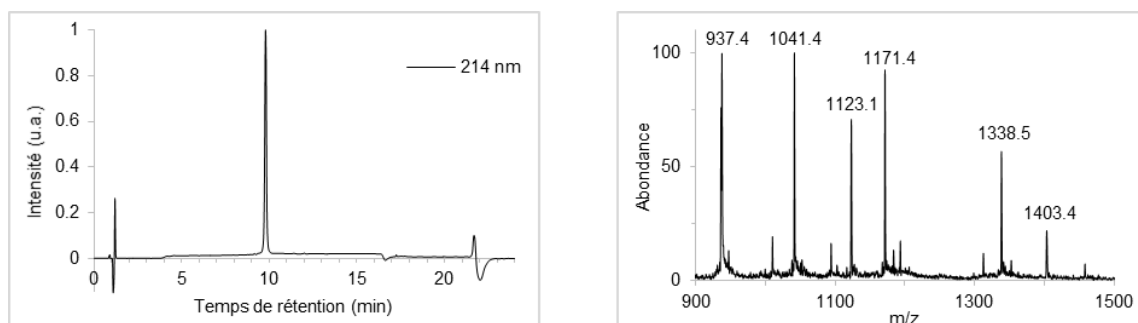


Figure 170: A gauche profil HPLC de la protéine ZIF(Cs124|DOTA[Tb]) obtenue par assemblage HNB de C-vers-N ; A droite, spectre ESI-MS (masse calculée : 1338,36 [M+7H]<sup>7+</sup>, 1171,19 [M+8H]<sup>8+</sup>, 1041,17 [M+9H]<sup>9+</sup>, 937,15 [M+10H]<sup>10+</sup> pour M = C<sub>398</sub>H<sub>626</sub>N<sub>129</sub>O<sub>110</sub>S<sub>7</sub>Tb correspondant à la protéine ZIF(Cs124|DOTA[Tb]); et 1403,31 [M+4H]<sup>4+</sup>, 1122,85 [M+5H]<sup>5+</sup> pour M = C<sub>235</sub>H<sub>374</sub>N<sub>75</sub>O<sub>65</sub>S<sub>5</sub>Tb correspondant à l'assemblage F1-F3)

## (d) Conclusion

La stratégie d'assemblage de C-vers-N a été validée, malgré la formation parasite de l'assemblage F1-F3 difficile à séparer par purification HPLC. Cependant l'étape de déprotection de la thiazolidine avec changement de pH devant s'effectuer en boîte à gants sur de petits volumes de solutions est difficile à maîtriser.

## ii Stratégie d'assemblage de N-vers-C

Le schéma de la stratégie d'assemblage de l'extrémité N vers l'extrémité C est donné en Figure 171. Le peptide central F2 et le peptide F1 sont assemblés par une première réaction de ligation. Ensuite le peptide F3 est introduit en présence d'un réducteur permettant de passer de la forme  $\text{HNB}^{\text{off}}$  à la forme  $\text{HNB}^{\text{on}}$ , déclenchant une seconde réaction de ligation conduisant à l'assemblage F1-F2-F3.

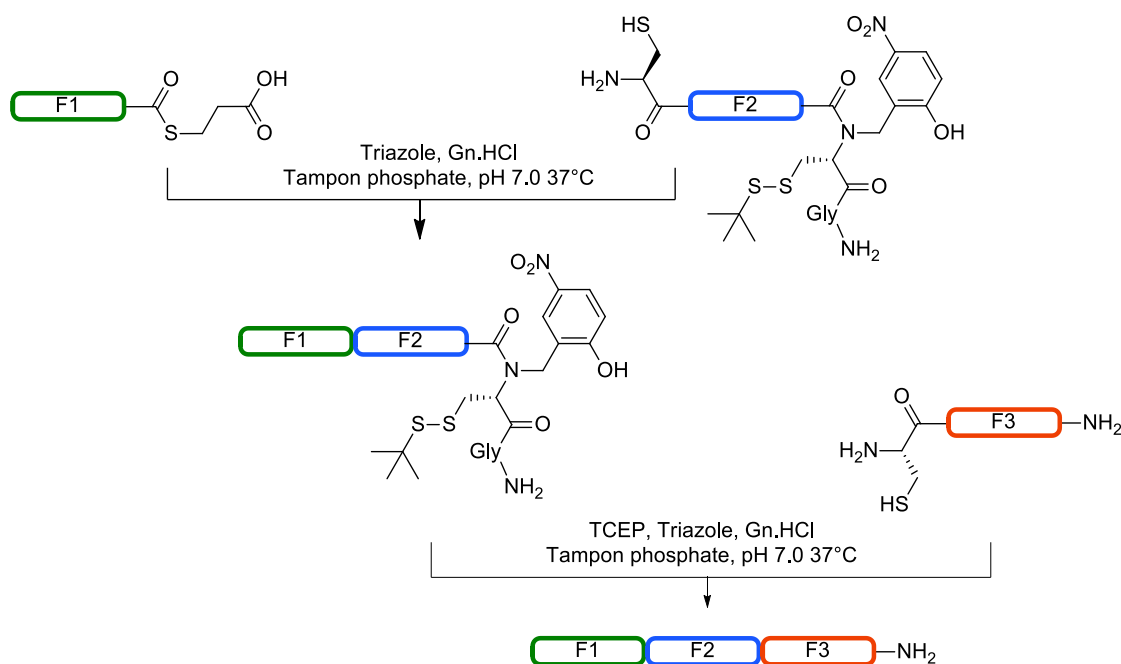


Figure 171: Schéma d'assemblage de trois fragments en stratégie de N vers C

La stratégie d'assemblage de N-vers-C implique l'assemblage de trois peptides :

- un peptide F1 comportant un ester de MPA en C-terminal (F1-MPA),
- un peptide F2 comportant une cystéine en N-terminal et une fonction N-HNB-Cys(StBu) en C-terminal (F2[Cys]-HNB<sup>off</sup>),
- un peptide F3 comportant une cystéine en N-terminal (F3-NH<sub>2</sub>).

De même que pour la stratégie C-vers-N, seul le fragment 2 diffère de la stratégie utilisée au premier chapitre. Le fragment central à synthétiser est noté F2[Cys|Glu(Cs124)]-HNB<sup>off</sup>.

(a) Synthèse du peptide F2[Cys|Glu(Cs124)]-HNB<sup>off</sup>

La synthèse du peptide F2[Cys|Glu(Cs124)]-HNB<sup>off</sup> a été réalisée comme pour le peptide F2[Thz|Glu(Cs124)]-HNB<sup>off</sup>, en remplaçant la thiazolidine par un acide aminé Fmoc-Cys(trt)-OH. Le peptide obtenu a été purifié par HPLC préparative avec un rendement de 11%, puis lyophilisé et caractérisé par HPLC et ESI-MS (Figure 172).

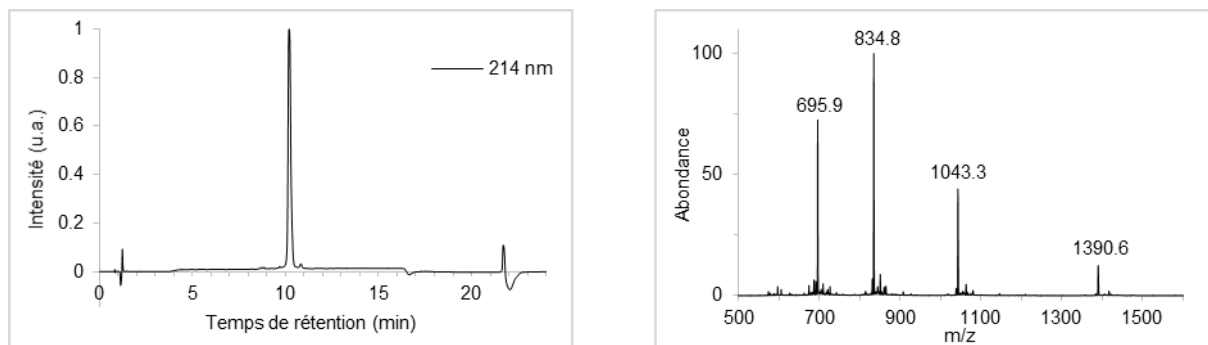


Figure 172: A gauche, chromatogramme HPLC du peptide F2[Cys|Glu(Cs124)]-HNB<sup>off</sup> ; A droite, spectre ESI-MS (masse calculée : 1390,59 [M+3H]<sup>3+</sup>, 1043,19 [M+4H]<sup>4+</sup>, 834,76 [M+5H]<sup>5+</sup>, 695,80 [M+6H]<sup>6+</sup> pour M = C<sub>179</sub>H<sub>276</sub>N<sub>58</sub>O<sub>50</sub>S<sub>4</sub>)

(b) Assemblage

La réaction d'assemblage (Figure 173) se déroule comme vu aux chapitres I&II, avec introduction des peptides F1-MPA et F2[Cys|Glu(Cs124)]-HNB<sup>off</sup> en conditions de ligation non réductrices, suivie de l'introduction du peptide F3-NH<sub>2</sub> et de TCEP comme réducteur. La protéine obtenue a été caractérisée par HPLC, ESI-MS (Figure 176).

De manière analogue à la ligation de C-vers-N, les chromatogrammes du suivi par HPLC ont été recalés ; cette fois sur les temps de rétentions des fragments F2 et F3, ainsi que l'assemblage F1-F2.

La première ligation suivie par HPLC pendant 66h, semble stagner à partir de 40h d'agitation (Figure 174). Les peptides F1 et F2 sont toujours présents, mais ne semblent plus être consommés. De plus, l'analyse par ESI-MS après 40 h montre l'apparition d'un produit d'hydrolyse du thioester de MPA présent en C-terminal du fragment 1. A partir de 66h de ligation, malgré une conversion très incomplète, la seconde ligation a été initiée, par l'ajout de TCEP et du peptide F3. Le suivi après 1h de cette seconde ligation montre l'apparition d'un pic majoritaire avec un temps de rétention entre 10 et 12 min (Figure 174). Le même comportement a été observé sur une autre ligation de N-vers-C lors de l'ajout de TCEP pour la première

ligation. On peut supposer que la première ligation entre F1 et F2 est très rapide en milieu réducteur et ne conduit étonnamment pas à l'assemblage F2-F2. Une hypothèse serait l'assistance du HNB pour cette première ligation, la cystéine de la fonction N-HNB-Cys(SH) permettant une première transthioesterification (Figure 175). Malgré l'incompréhension de ce phénomène, la seconde ligation s'est terminée après 120h d'agitation (Figure 174). La formation d'un pic majoritaire est observée et correspond à l'assemblage F1-F2-F3. L'assemblage résiduel F1-F3 n'est présent qu'à l'état de traces sur le profil HPLC du suivi de la réaction ou de la caractérisation de la protéine finale.

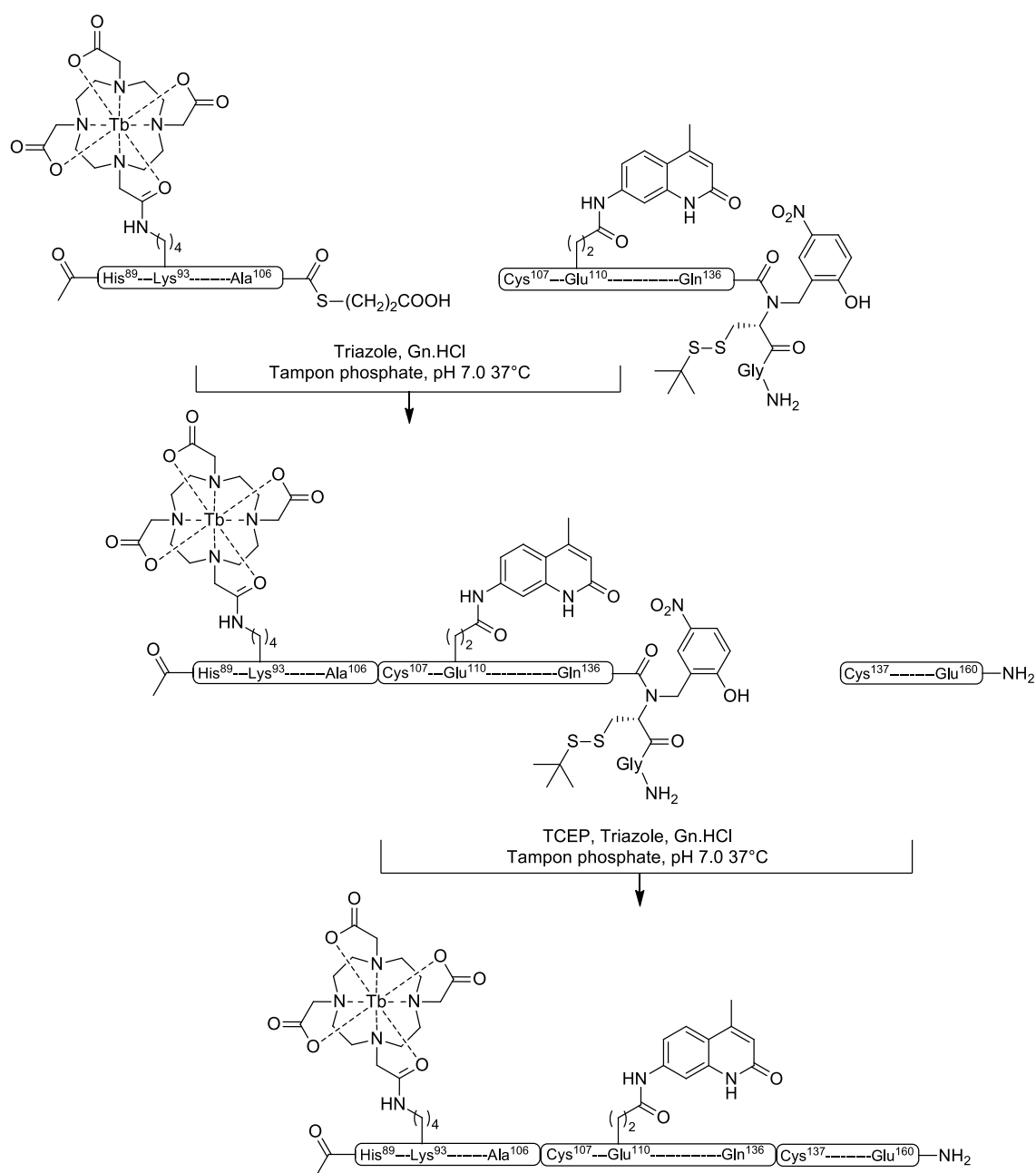


Figure 173: Schéma d'assemblage en stratégie de N vers C de la protéine ZIF(Cs124|DOTA[Tb])

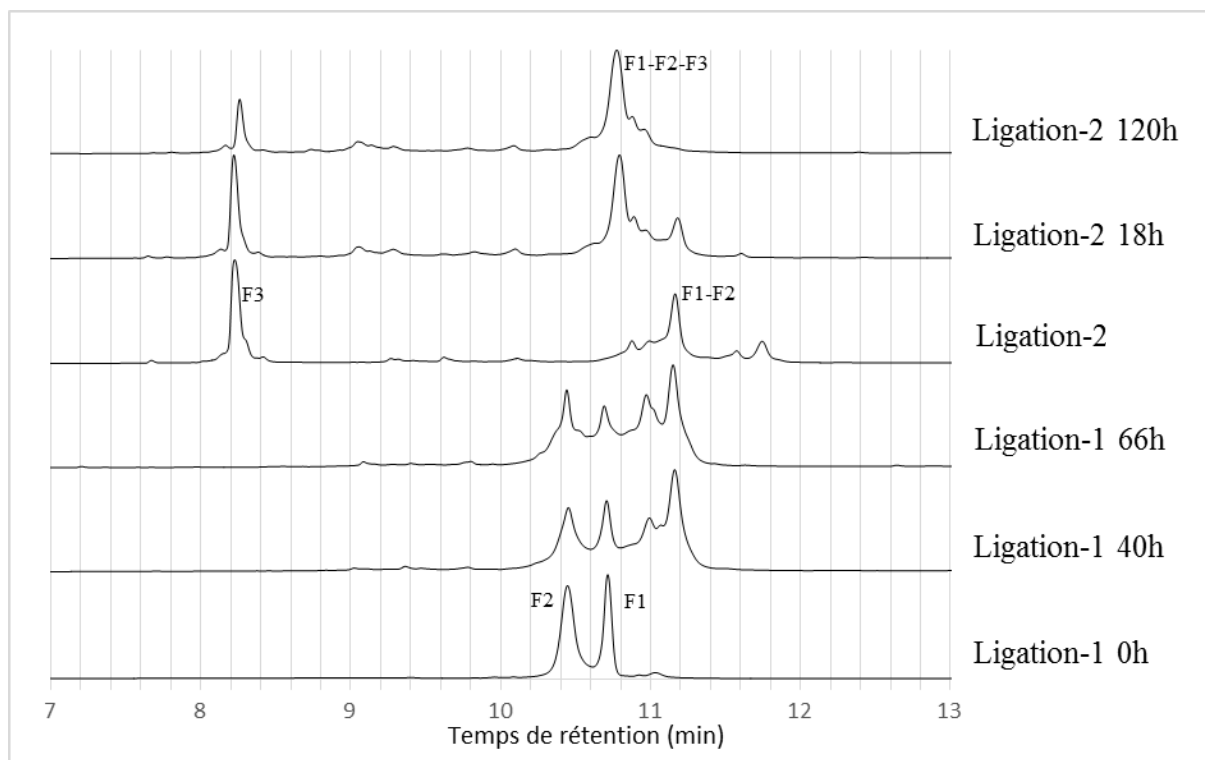


Figure 174: Suivi par HPLC de l'assemblage de la protéine ZIF(Cs124|DOTA[Tb])

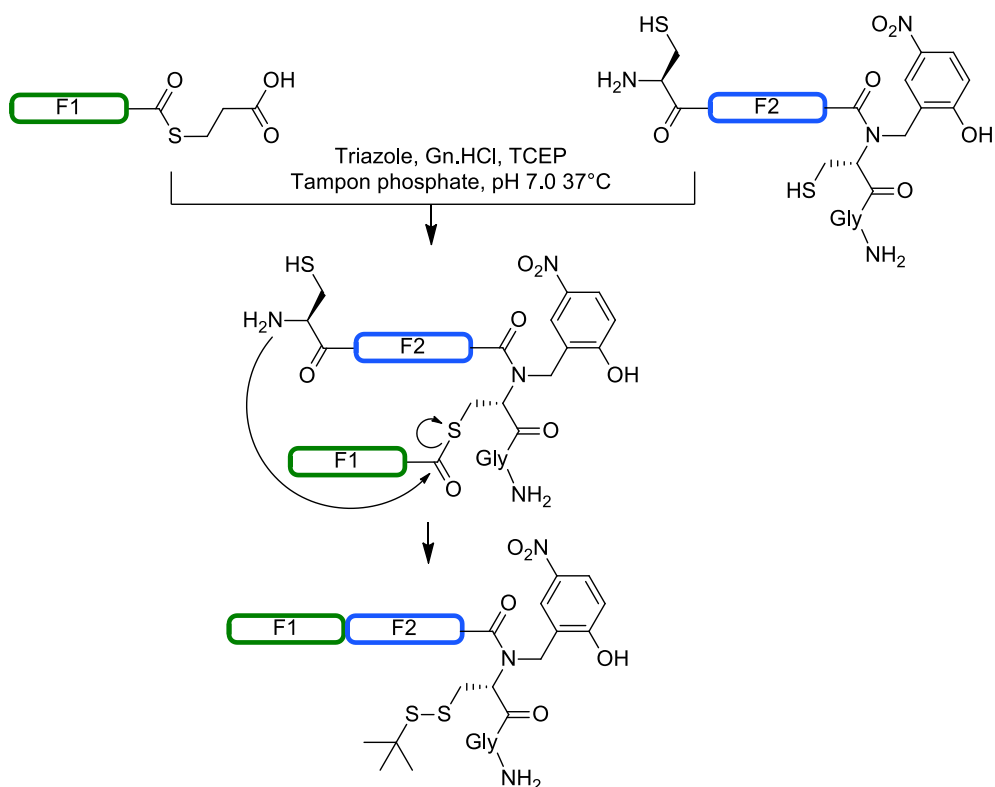


Figure 175: Hypothèse de l'assistance par la cystéine du groupement N-HNB-Cys(SH) pour la première ligation

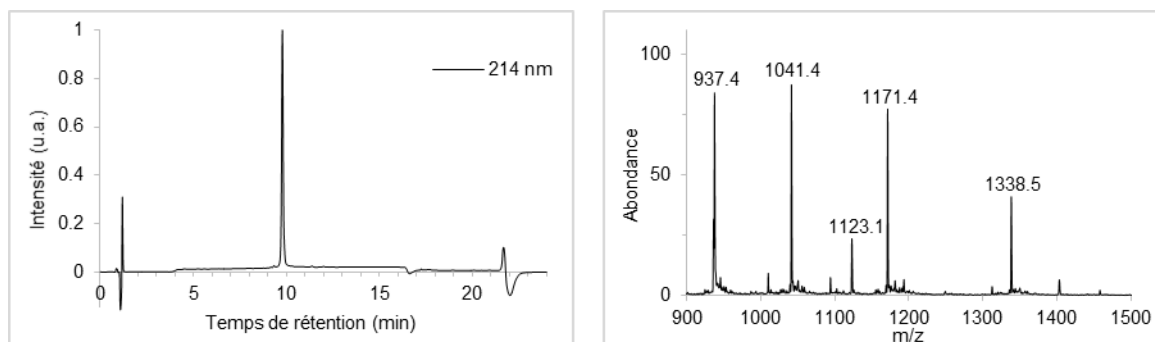


Figure 176: A gauche profil HPLC de la protéine ZIF(Cs124|DOTA[Tb]) obtenue par assemblage HNB N-vers-C ; A droite, spectre ESI-MS (masse calculée : 1338,36  $[M+7H]^{7+}$ , 1171,19  $[M+8H]^{8+}$ , 1041,17  $[M+9H]^{9+}$ , 937,15  $[M+10H]^{10+}$  pour  $M = C_{398}H_{626}N_{129}O_{110}S_7Tb$  correspondant à la protéine ZIF(Cs124|DOTA[Tb]) et 1122,85  $[M+5H]^{5+}$  pour  $M = C_{235}H_{374}N_{75}O_{65}S_5Tb$  correspondant à l'assemblage F1-F3.

### (c) Conclusion

La stratégie d'assemblage de N-vers-C conduit au bon produit d'assemblage, avec peu de sous-produits observés par HPLC. Cependant, la première ligation n'est pas totalement maîtrisée, l'influence du TCEP n'étant pas totalement comprise et l'assemblage n'ayant été réalisé qu'une seule fois. Cette stratégie reste cependant plus simple à mettre en œuvre que l'assemblage de C-vers-N, l'ensemble des réactions se déroulant au même pH.

### iii Conclusion sur les stratégies d'assemblages

Deux nouvelles stratégies d'assemblage de trois fragments peptidiques dans le sens « N-vers-C » et « C-vers-N » combinant une fonction N-HNB-Cys(StBu) en C-terminal du second fragment ont été testées et ont conduit aux protéines attendues. Par manque de temps, les assemblages n'ont pu être réalisés qu'une seule fois, laissant quelques points d'ombre, notamment sur la cinétique de la première ligation de l'assemblage de N-vers-C. On peut tout de même en conclure que la combinaison de la fonction N-HNB-Cys(StBu) et de l'utilisation de triazole en remplacement du MPAA permet l'assemblage des trois fragments peptidiques en one-pot, et devrait être compatible avec la présence d'alcynes et donc avec le complexe  $[Eu(DO3ApicAr_2)]^-$ .

Comme observé sur les profils HPLC des suivis de ligations, la stratégie N-vers-C conduit à moins de sous-produits, et correspond donc à la meilleure stratégie d'assemblage.

### C Synthèse de la protéine ZIF(Eu(DO3ApicAr<sub>2</sub>))

La synthèse de la protéine ZIF(Eu(DO3ApicAr<sub>2</sub>)) par la stratégie d'assemblage HNB de N-vers-C développée dans la partie précédente requiert les peptides F1-MPA et F3-NH<sub>2</sub> précédemment synthétisés (cf Chapitres I&II) et un nouveau peptide central F2[Cys|Lys(Eu(DO3ApicAr<sub>2</sub>))]-HNB<sup>off</sup> (Figure 177).

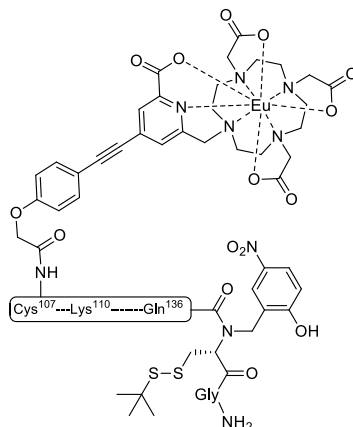


Figure 177: Structure du peptide F2[Cys|Lys(Eu(DO3ApicAr<sub>2</sub>))]-HNB<sup>off</sup>

#### i Synthèse du peptide F2[Cys|Lys(Eu(DO3ApicAr<sub>2</sub>))]-HNB<sup>off</sup>

La synthèse du peptide F2[Cys|Lys(Eu(DO3ApicAr<sub>2</sub>))]-HNB<sup>off</sup> a été réalisée à partir d'un peptide F2[Cys|Lys(Alloc)]-HNB-RinkAmide qui a été synthétisé comme le peptide F2[Cys|Glu(Cs124)]-HNB<sup>off</sup>, en remplaçant l'acide aminé Glu(Cs124) en position 110 par une Lys(Alloc). Le complexe d'euprium peut être introduit sur la lysine une fois déprotégée, soit par le couplage du pro-ligand DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>2</sub> (c'est à dire de manière similaire à la stratégie utilisée pour le peptide F2[Cys(StBu)|Lys(Tb(DO3ApicAr<sub>1</sub>))]-SEA<sup>off</sup>), soit par le couplage direct du complexe [Eu(DO3ApicAr<sub>2</sub>)]Na. Le couplage du pro-ligand ayant démontré son efficacité auparavant, cette stratégie a été choisie.

##### (a) Synthèse par couplage du pro-ligand

###### 1. Couplage du pro-ligand DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>2</sub>

Le pro-ligand DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>2</sub> étant une molécule à forte valeur ajoutée, un couplage dans des conditions stœchiométriques est préférable. Cependant, la réactivité du groupement N-HNB-Cys(StBu) apporte une contrainte puisqu'une O-acylation sur la fonction hydroxyle du pro-ligand est possible (cf partie « Réactivité du phénol » p.148). Il est néanmoins possible de bloquer cette position en l'acétylant avant la déprotection de la Lys(alloc). Une fois la lysine déprotégée et le pro-ligand couplé, la fonction hydroxyle peut être libérée par un traitement à la pipéridine, ou à l'hydroxylamine (Figure 178).

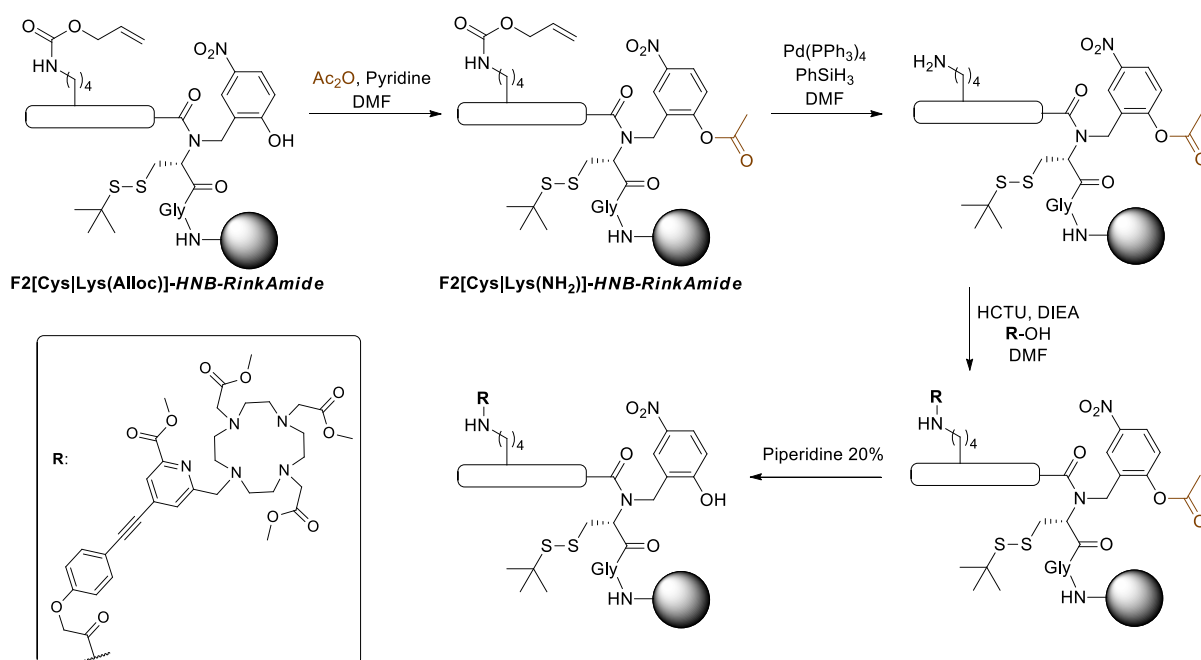


Figure 178: Stratégie pour bloquer/débloquer la fonction hydroxyle du groupement N-HNB-Cys(StBu)

La fonction hydroxyle du peptide **F2[Cys[Lys(Alloc)]-HNB-RinkAmide** a été acétylée, puis le groupement protecteur Alloc de la lysine a été déprotégé, suivant les deux premières étapes de la stratégie présentée en Figure 178. Une analyse par ESI-MS de ce peptide a permis de vérifier la bonne déprotection du groupement protecteur alloc du peptide **F2[Cys[Lys(NH<sub>2</sub>)]-HNB-RinkAmide** obtenu. Une analyse par ESI-MS de ce peptide après traitement à la pipéridine a cependant révélé la persistance d'un groupement acyle, signifiant que soit la déprotection de la fonction hydroxyle à la pipéridine a été inefficace, soit que la lysine a été acétylée.

Le peptide a donc été traité par l'anhydride acétique de manière à obtenir le peptide O-acétylé et N-acétylé, dont l'analyse par ESI-MS a révélé la présence de deux fonctions acétyle sur le peptide. Après traitement de ce peptide à la pipéridine, l'analyse ESI-MS a révélé la présence d'un seul acétyle sur le peptide. Nous en déduisons que le traitement permettant de libérer la fonction hydroxyle est efficace, signifiant que la lysine est acétylée (Figure 179).

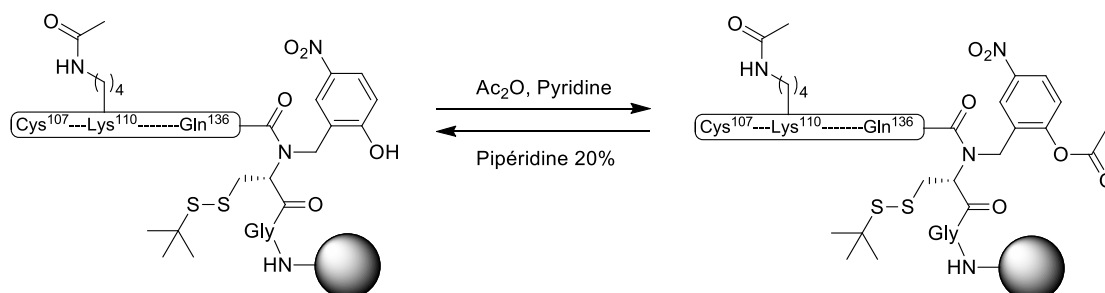


Figure 179: Protection et déprotection de la fonction hydroxyle

Si la lysine est acétylée et non l'hydroxyle, cela signifie qu'un transfert d'acyle O vers N intra ou intermoléculaire a eu lieu (Figure 180). Comme expliqué en partie « Réactivité du phénol » (p.148) le phénol joue un rôle d'assistance pour le couplage du premier acide aminé du peptide, cependant une assistance sur une fonction amine aussi éloignée dans la séquence n'était pas attendue. Le ligand peut donc être introduit en quantité stœchiométrique, le couplage s'effectuant soit directement avec l'espèce activée par l'HCTU, soit par l'assistance du phénol.

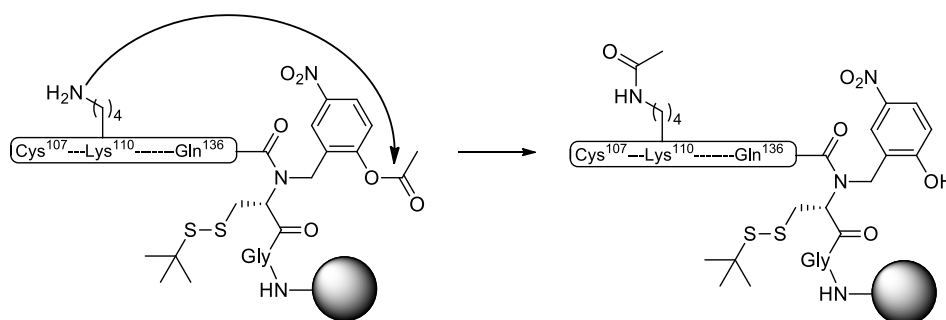


Figure 180: Transfert d'acyl O vers N sur le peptide F2[Cys|Lys(NH<sub>2</sub>)]-HNB<sup>off</sup>

La synthèse du peptide F2[Cys|Lys(NH<sub>2</sub>)]-HNB-RinkAmide a été répétée sans l'étape d'acétylation finale, puis le pro-ligand a été couplé par réaction de 1,1 éq de DO3A<sup>(Me)<sub>3</sub></sup>pic<sup>(Me)</sup>Ar<sub>2</sub> sur le peptide (1éq). Une étape de traitement à la pipéridine a ensuite permis de libérer la fonction hydroxyle du phénol, permettant après déprotection au TFA l'obtention du peptide F2[Cys|Lys(DO3A<sup>(Me)<sub>3</sub></sup>pic<sup>(Me)</sup>Ar<sub>2</sub>)]-HNB<sup>off</sup>.

## 2. Hydrolyse du pro-ligand

Les conditions de l'hydrolyse optimisées pour le peptide F2[Cys(StBu)|Lys(DO3A<sup>(Me)<sub>3</sub></sup>pic<sup>(Me)</sup>Ar<sub>1</sub>)]-SEA<sup>off</sup> en partie « Hydrolyse » (p.132) ont été reprises. L'analyse du produit d'hydrolyse du peptide F2[Cys|Lys(DO3A<sup>(Me)<sub>3</sub></sup>pic<sup>(Me)</sup>Ar<sub>2</sub>)]-HNB<sup>off</sup> a révélé une différence de masse de -123 Da entre le produit obtenu et le produit attendu. Ceci peut correspondre à une  $\beta$ -élimination du groupement S-StBu de la cystéine présente sur le groupement HNB (Figure 181). Pour vérifier cette hypothèse, le peptide issu de l'hydrolyse a été mis en présence de TCEP à pH 7. Aucune différence de masse n'a été observée sur le produit résultant de cette étape de réduction. Le groupement S-StBu n'était donc plus présent sur le peptide, validant notre hypothèse.

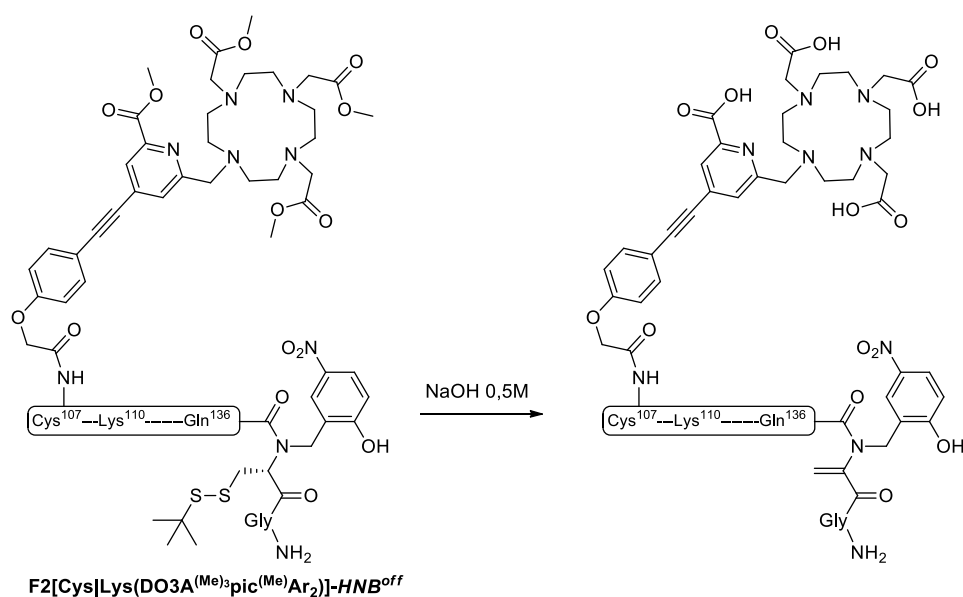


Figure 181: Hydrolyse des esters méthyliques et  $\beta$ -élimination du groupement S-StBu de la cystéine

### 3. Conclusion

La  $\beta$ -élimination du groupement S-StBu de la cystéine du groupement N-HNB-Cys(StBu) lors de l'étape d'hydrolyse du pro-ligand rend impossible l'obtention du peptide final comportant le ligand libre. Le peptide F2[Cys|Lys(Eu(DO3ApicAr<sub>2</sub>))]-HNB<sup>off</sup> doit donc être obtenu par une autre stratégie de synthèse.

#### (b) Couplage direct du complexe [Eu(DO3ApicAr<sub>2</sub>)]Na

Comme décrit en partie « Couplage du complexe [Tb(DO3ApicAr<sub>1</sub>)]Na » (p.131), un essai de couplage direct du complexe [Tb(DO3ApicAr<sub>1</sub>)]Na par de l'HCTU sur le peptide test de séquence C(StBu)GKKFTKK(NH<sub>2</sub>)EELQKHAKTHTG-*RinkAmide* n'a pas conduit à l'espèce attendue. Nous avons tout de même observé une meilleure solubilité dans le DMF du complexe [Eu(DO3ApicoAr<sub>2</sub>)]Na comparé à celle du complexe [Tb(DO3ApicoAr<sub>1</sub>)]Na. Suite aux observations réalisées sur l'assistance de la fonction hydroxyle du phénol par un transfert d'acyle de O-vers-N, un essai de couplage à l'HCTU du complexe [Eu(DO3ApicoAr<sub>2</sub>)]Na sur le peptide F2[Cys|Lys(NH<sub>2</sub>)]-HNB-*RinkAmide* a été réalisé. Le brut de déprotection du peptide analysé par ESI-MS a révélé la présence de deux espèces majoritaires : le peptide comportant le complexe, et le peptide comportant le ligand libre (l'euprium est éliminé lors de l'étape de déprotection au TFA); les espèces minoritaires étant le peptide nu, et des peptides de masse M+42 et M+84. L'utilisation de DIC/HOAt en tant qu'activateur a permis d'obtenir 100% de peptide avec soit le complexe gréffé, soit le ligand (Figure 182). Cela indique que l'assistance par le phénol est certainement possible pour des fonctions difficilement conjuguables. Cette

observation, réalisée en toute fin de thèse, doit tout de même inciter à la prudence, car ces conditions de couplage n'ont pas été testées sur d'autres séquences. Une partie du peptide obtenu possédant un ligand libre à l'issue de l'étape de déprotection, le brut de réaction a été mis en présence d'euprimum à pH 6 afin d'obtenir une seule espèce métallée pour l'étape de purification. Le peptide F2[Cys]Lys(Eu(DO3ApicAr<sub>2</sub>))-HNB<sup>off</sup> a ensuite été purifié, puis analysé par ESI-MS et HPLC (Figure 183).

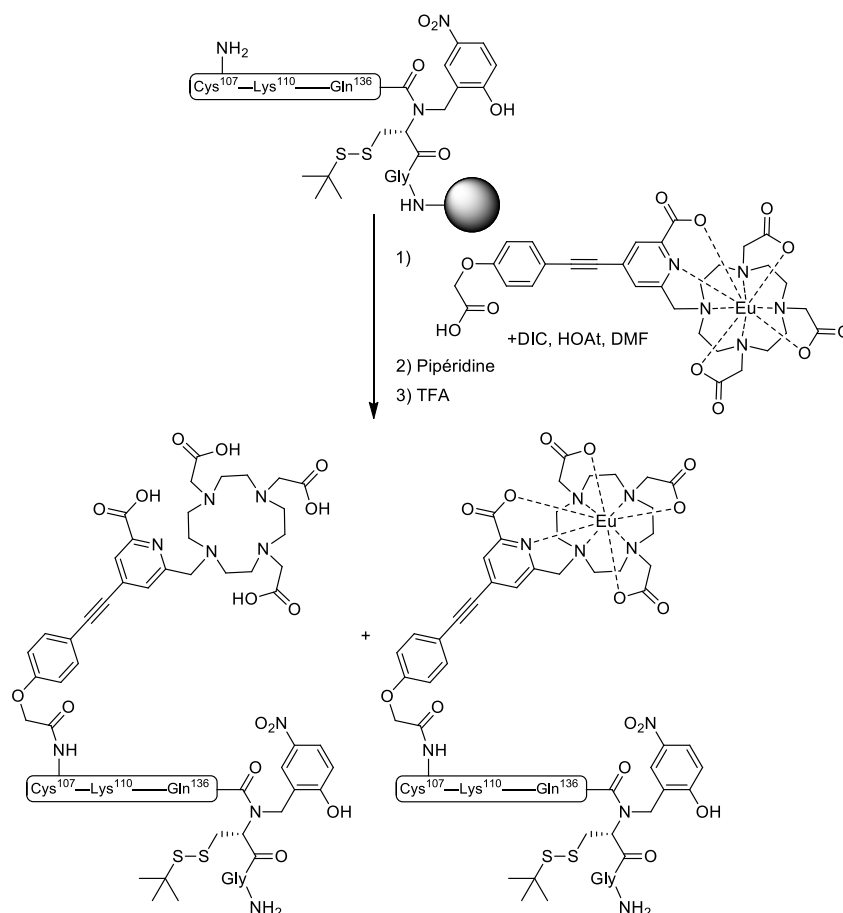


Figure 182: Schéma de synthèse du couplage direct du ligand DO3Apico[Eu]Ar<sub>2</sub> sur le peptide F2[Cys(SH)|Lys(NH<sub>2</sub>)]-HNB<sup>off</sup>

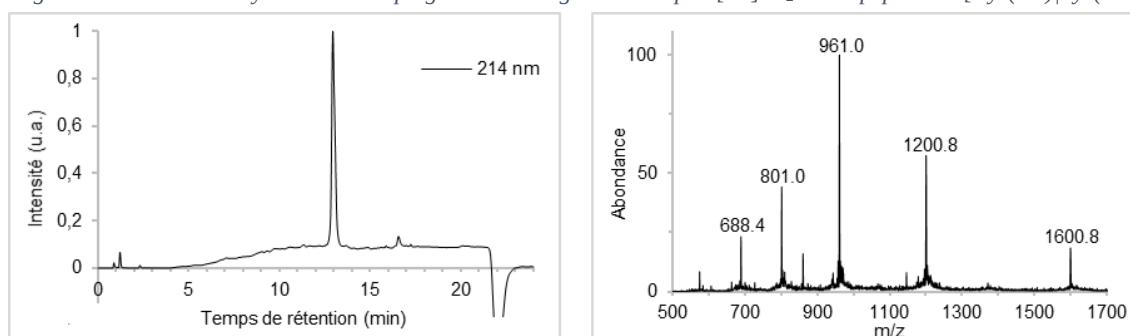


Figure 183: A gauche, chromatogramme HPLC du peptide F2[Cys]Lys(Eu(DO3ApicAr<sub>2</sub>))-HNB<sup>off</sup>; A droite, spectre ESI-MS (masse calculée : 1600,40 [M+3H]<sup>3+</sup>, 1200,55 [M+4H]<sup>4+</sup>, 960,64 [M+5H]<sup>5+</sup>, 800,71 [M+6H]<sup>6+</sup> pour M = C<sub>201</sub>H<sub>304</sub>EuN<sub>62</sub>O<sub>58</sub>S<sub>4</sub>)

## ii Assemblage

La protéine ZIF(Eu(DO3ApicAr<sub>2</sub>)) a été assemblée par la stratégie N-vers-C (Figure 184) par deux réactions de ligations successives, impliquant en premier lieu le peptide F1-MPA décrit dans le chapitre II et le peptide F2[Cys[Lys(Eu(DO3ApicAr<sub>2</sub>))]-HNB<sup>off</sup>, puis en second lieu le peptide obtenu issu de la première ligation, avec le peptide F3-MPA décrit dans le chapitre I. La protéine ZIF(Eu(DO3ApicAr<sub>2</sub>)), obtenue avec un rendement de 12.5%, a été caractérisée par HPLC et ESI-MS (Figure 185).

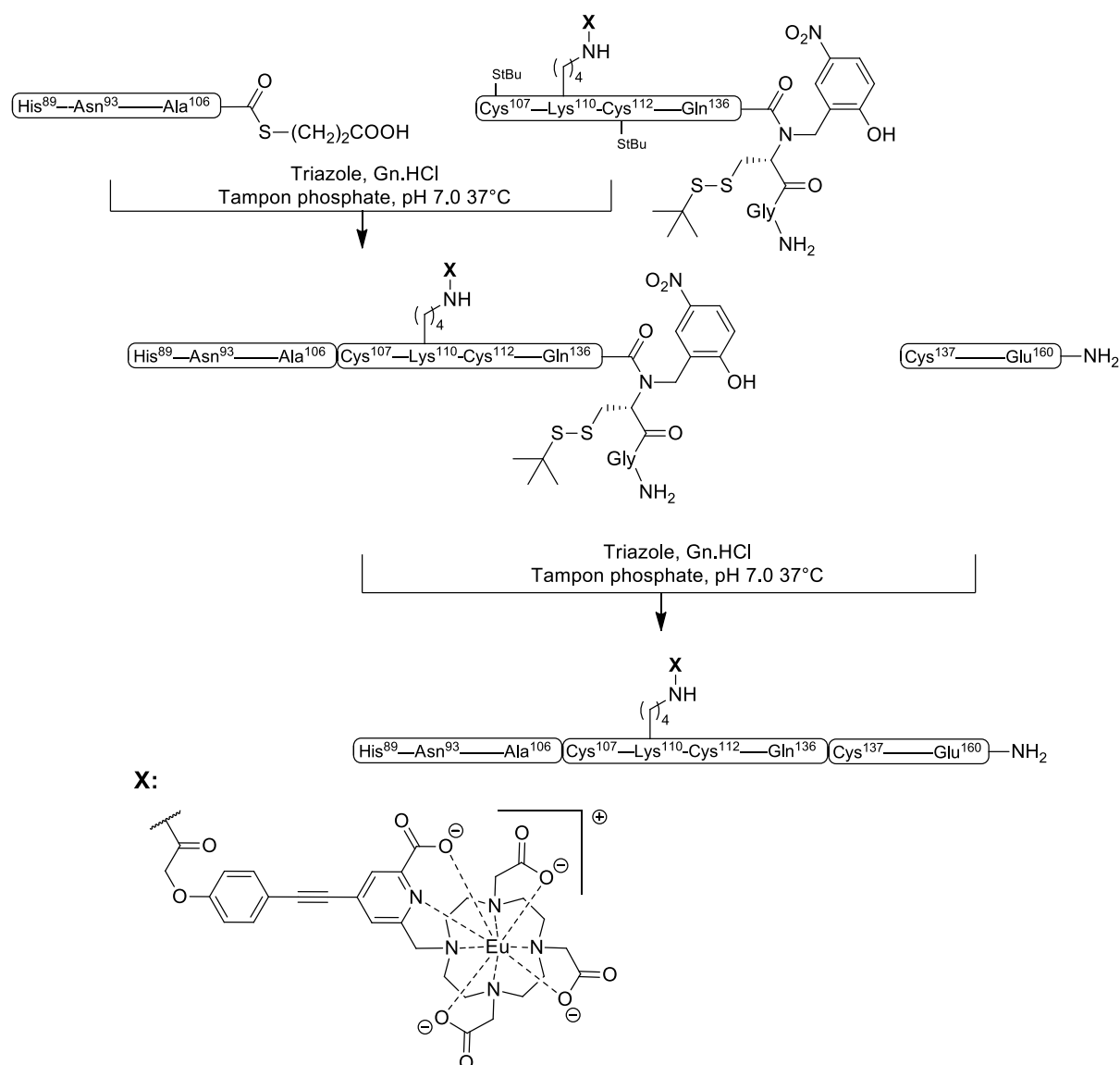


Figure 184: Schéma d'assemblage pour la synthèse de la protéine ZIF(Eu(DO3ApicAr<sub>2</sub>))

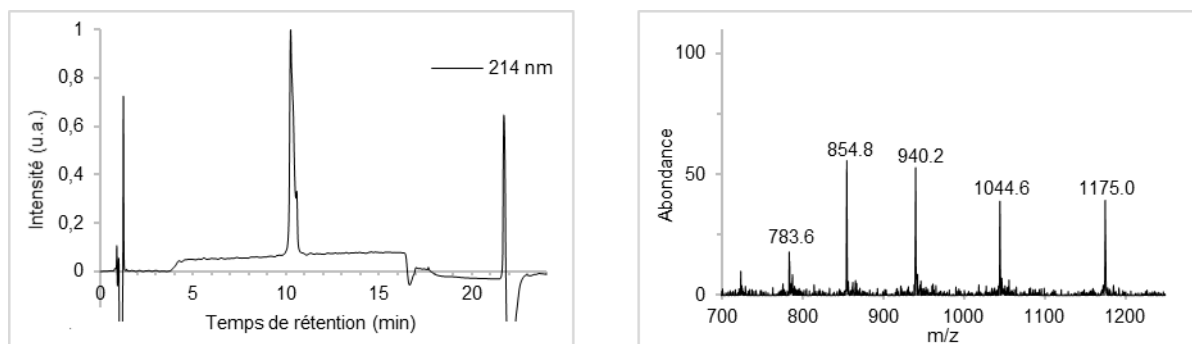


Figure 185: A gauche, chromatogramme HPLC de la protéine ZIF(Eu(DO3ApicAr<sub>2</sub>)); A droite, spectre ESI-MS (masse calculée : 1175,06 [M+8H]<sup>8+</sup>, 1044,62 [M+9H]<sup>9+</sup>, 940,26 [M+10H]<sup>10+</sup>, 854,87 [M+11H]<sup>11+</sup>, 783,71 [M+12H]<sup>12+</sup> pour  $M = C_{400}H_{625}N_{129}O_{111}S_7$ )

Comme observé lors de l'assemblage test, la première ligation a stagné après 40h d'agitation, et ce n'est que lors de l'ajout de TCEP avec agitation pendant 1h que la formation d'un pic majoritaire a été observée. La seconde ligation a conduit à la formation de plusieurs produits après 60h, dont deux majoritaires présentant une absorbance à 365nm, longueur d'onde où seuls les composés HNB et le complexe Eu[DO3ApicoAr<sub>2</sub>] absorbent.

La formation du produit attendu a été confirmée par l'analyse ESI-MS, et les autres produits du milieu réactionnel possèdent des masses plus faibles que la protéine finale, indiquant qu'aucune addition de thiol n'a eu lieu sur la fonction alcyne du complexe.

La stratégie mise en œuvre est donc validée, mais requiert quelques optimisations, la seconde ligation ne conduisant pas à un produit majoritaire.

Par manque de temps, nous n'avons pas pu préparer la quantité de protéine nécessaire pour réaliser l'étude des propriétés spectroscopiques de ce composé, ainsi que l'évaluation d'un effet FRET europium vers SiR-méthyl par titrage avec l'ADN cible.

## IV Conclusion

Le complexe  $[\text{Tb}(\text{DO3ApicAr}_1)]^-$  développé au Chapitre III a été conjugué sur la protéine ZIF en quelques étapes par la voie de couplage du pro-ligand. Ce résultat a validé la stratégie de synthèse, ainsi que l'utilité du complexe développé, gardant ses propriétés spectroscopiques lorsqu'il est introduit sur une biomolécule. La détection de la séquence cible d'ADN 5'-TGGGCGCGCCCA-3' a été mise en évidence par la formation des hétéro-dimères  $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{lissamine})$  et  $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{SiR-méthyl})$ , pour lesquels un transfert d'énergie du terbium vers la lissamine ou la SiR-méthyl a été observé. Une augmentation de la luminescence des fluorophores, ainsi qu'une baisse du temps de vie de luminescence et de l'émission du terbium ont permis de conclure à un transfert d'énergie en présence d'ADN, même si ce dernier reste toutefois modeste. La principale limite de ce système est l'excitation directe des fluorophores à la longueur d'onde d'excitation du complexe de terbium. Pour contourner cette limite, l'introduction du complexe d'euprium sur la protéine ZIF a été envisagée. L'incompatibilité de ce complexe avec la stratégie de synthèse des protéines utilisée jusqu'alors a conduit au développement d'une nouvelle voie d'assemblage. Pour cela, le crypto-thioester N-HNB-Cys(StBu) récemment décrit dans la littérature a permis non seulement le couplage direct du complexe d'euprium lors de la synthèse peptidique, mais aussi l'assemblage des fragments peptidiques sans dégradation du complexe. L'obtention de cette protéine en fin de thèse a permis de valider la nouvelle stratégie de synthèse, cependant aucune étude spectroscopique n'a pu être réalisée.

Comme mentionné dans le chapitre II, il serait intéressant de réaliser un titrage à l'ADN cible afin de mettre en évidence la formation d'un complexe 1 :1 :1 ( $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{fluorophore})/\text{ADN}$ ). Il faudrait par ailleurs évaluer les paramètres des interactions mises en jeu dans ce système ternaire par mobilité électrophorétique, et réaliser un titrage contrôle avec une séquence d'ADN ne comportant pas la séquence cible.





## Conclusion générale et perspectives



## Conclusion générale et perspectives

Ce travail de thèse consistait en la preuve de concept de la détection d'une séquence palindromique d'ADN double brins par luminescence. Pour cela, les sondes ont été basées sur les propriétés d'interaction des protéines à doigt de zinc, qui décorées par des lanthanides permet de réaliser une détection résolue en temps. Pour cela, différents systèmes de FRET ont été introduits sur une protéine décrite comme se dimérisant autour d'une séquence palindromique d'ADN. Les différents systèmes de sondes développés ont été caractérisés par spectroscopie.

La synthèse de la première sonde luminescente basée sur un transfert FRET Cs124 vers terbium développée au premier chapitre a permis de faire face aux différentes contraintes synthétiques concernant la synthèse peptidique, qui une fois résolues ont permis l'obtention de la protéine  $ZIF^{Zn}(Cs124|DOTA[Tb])$ . La synthèse de cette protéine a nécessité son acétylation en position N-terminale, ainsi que la modification de trois acides aminés, dont un pour éviter la formation d'aspartimides, et deux autres pour l'introduction du système de luminescence ; un de ces deux acides aminés étant présent dans l'interface de dimérisation. L'interaction de la sonde avec sa séquence d'ADN cible a été mise en évidence par un test préliminaire de retard sur gel, confirmant la possibilité de modifier certains acides aminés de la protéine initialement décrite par B. Wang et C. Pabo sans en affecter ses propriétés de dimérisation. Cependant, la dimérisation induite par la liaison à l'ADN n'a pas permis d'améliorer le transfert d'énergie du Cs124 vers le terbium, la distance entre le Cs124 et le terbium étant certainement trop importante au sein du dimère formé.

Une seconde génération de sonde basée sur un FRET entre deux fluorophores (lissamine et SiR-méthyl) efficace à de plus fortes distances a ensuite été développée. Cette fois, l'introduction sélective d'un seul fluorophore par protéine a permis de se rapprocher un peu plus de la séquence de la protéine initiale avec la modification de deux acides aminés. Cette génération de sonde a permis de soulever la question de conservation des sondes, ces dernières s'étant montrés instables en conditions aérobies. L'interaction par retard sur gel n'a pas été réalisée, cependant l'étude spectroscopique du système  $ZIF^{Zn}(\text{lissamine})/ZIF^{Zn}(\text{SiR-méthyl})$  a permis de mettre en évidence un transfert FRET de la lissamine vers la SiR-méthyl en présence de la séquence d'ADN cible. Ce résultat a apporté la preuve de concept de la détection d'une séquence d'ADN par ces sondes. Notre objectif étant initialement de concevoir un système

FRET à partir de lanthanides pour bénéficier de leurs longs temps de vie de luminescence, l'introduction de complexes de lanthanides fortement brillants a été recherchée afin de réaliser un transfert FRET d'un lanthanide vers un fluorophore efficace à de fortes distances.

Parmi l'ensemble des complexes de lanthanides décrits dans la littérature, aucun ne correspondait aux critères de sélection que nous nous étions imposés. Les critères limitants étant principalement la présence d'une fonction de bioconjugaison, la possibilité d'une stratégie de synthèse divergente conduisant à différents complexes de lanthanides, et finalement la possibilité d'une absorption bi-photonique ; ce dernier critère n'étant pas recherché pour ce projet, mais pour d'autres développés au laboratoire. Le design de nouveaux complexes a été réalisé en s'appuyant sur les complexes décrits dans la littérature, avec le choix d'une stratégie de synthèse divergente et de deux voies de couplage possibles pour chaque complexe sur une biomolécule. La synthèse de trois ligands a ainsi été optimisée, sur lesquels différents lanthanides ont pu être introduits (terbium, europium, dysprosium, samarium) conduisant à une famille de six complexes. Parmi eux, les complexes  $[\text{Tb}(\text{DO3ApicAr}_1)]\text{Na}$  et  $[\text{Eu}(\text{DO3ApicAr}_2)]\text{Na}$  se sont démarqués de par leurs brillances élevées (respectivement 7470 et 4790  $\text{M}^{-1}\text{cm}^{-1}$ ) ainsi que leurs longs temps de vie de luminescence (2,35 ms pour le terbium et 1,10 ms pour l'euporium). Leur introduction sur la protéine ZIF a alors été recherchée.

La seconde génération de sonde basée sur un FRET d'un lanthanide vers un fluorophore a été développée, avec premièrement l'introduction du complexe de terbium  $[\text{Tb}(\text{DO3ApicAr}_1)]\text{Na}$  sur la protéine ZIF. Cette dernière a été synthétisée par la même stratégie d'assemblage en trois fragments peptidiques utilisée jusqu'alors. Le peptide comportant le complexe de terbium a été obtenu par le couplage d'un pro-ligand, qui après plusieurs étapes a conduit au complexe attendu, validant la voie d'introduction utilisée. L'étude spectroscopique a montré que les systèmes FRET terbium vers lissamine ou terbium vers SiR-méthyl fonctionnaient. En effet, lors de l'ajout d'ADN sur les solutions ( $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{lissamine})$  ou  $\text{ZIF}^{\text{Zn}}(\text{Tb}(\text{DO3ApicAr}_1))/\text{ZIF}^{\text{Zn}}(\text{SiR-méthyl})$ ), une baisse du temps de vie du terbium ainsi que de son émission a été observée, tandis que la luminescence des fluorophores a augmentée (de manière plus prononcée avec la SiR-méthyl qu'avec la lissamine). Cependant, même en l'absence d'ADN l'émission des fluorophores a été observée. Cela provenant de l'excitation directe des fluorophores à la longueur d'onde d'excitation du complexe de terbium, il a fallu rechercher un autre système permettant une excitation à une longueur d'onde à laquelle le fluorophore n'absorberait pas. Le complexe d'euporium  $[\text{Eu}(\text{DO3ApicAr}_2)]\text{Na}$  s'excitant à des

longueurs d'ondes où la lissamine et la SiR-méthyl n'absorbent pas, a été choisi pour être introduit sur la protéine ZIF. La synthèse de cette nouvelle protéine s'est révélé difficile à cause d'une réaction secondaire dégradant le complexe d'euprium dans les conditions d'assemblage de peptides utilisées. Une nouvelle stratégie d'assemblage compatible avec le complexe a alors été développée et validée. Toutefois, par manque de temps la caractérisation spectroscopique de la protéine  $ZIF^{Zn}(Eu(DO3ApicAr_2))$  n'a pu être réalisée.

En perspective, les caractérisations spectroscopiques mériteraient d'être plus approfondies en testant une séquence d'ADN ne comportant pas la séquence cible, afin de confirmer la sélectivité du système. Dans la même idée, les tests par gels retard devraient être réalisés sur les protéines en présence de leurs séquence cible et d'autres séquences.

Toutefois, la preuve de concept a été apportée, et le point à retenir est la possible modification de certains acides aminés de la séquence initialement décrite ne mettant pas en jeu l'intégrité du système.

La preuve de concept ayant été apportée, il est désormais important de simplifier l'aspect synthétique des protéines. Pour cela, la conjugaison finale des chromophores sur la protéine par réaction chimio-sélective est recherchée. Ainsi, la protéine comportant une fonction de conjugaison pourrait être synthétisée à grande échelle (plusieurs milligrammes), puis les chromophores couplés par une simple réaction en milieu aqueux. Pour cela, Thibault Charnay travaille actuellement sur le développement de la ligation tétrazine mentionnée au chapitre III (p.103), qui permettrait de synthétiser la protéine ZIF comportant une oléfine réactive avec une tétrazine liée au chromophore (Figure 186).

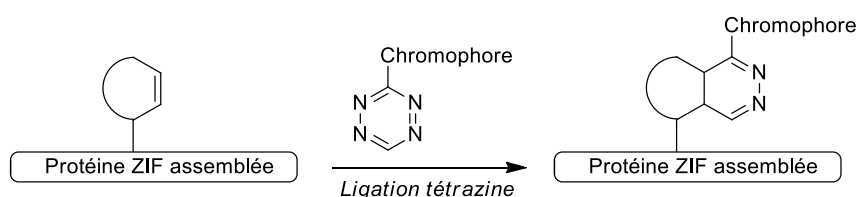


Figure 186: Nouvelle stratégie de synthèse pour le couplage des chromophores

Lorsque l'optimisation sera réalisée, une protéine ciblant une nouvelle séquence d'ADN pourra être synthétisée. Ainsi, une séquence non palindromique de 12 paires de bases d'ADN pourra être reconnue avec les sondes synthétisées.

Finalement, l'introduction d'un troisième doigt de zinc sur chaque monomère permettra la lecture d'une séquence d'ADN de 18 paires de bases lors de la formation d'hétéro-dimères, ce

qui constitue à la cible finale, cette dernière pour rappel étant la détection d'une séquence d'ADN d'au minimum 17 paires de bases.





# Bibliographie



## Bibliographie

- [1] Hélène, C.; Toulmé, J.-J. Control of Gene Expression by Oligodeoxynucleotides Covalently Linked to Intercalating Agents and Nucleic Acid-cleaving Reagents. In *Oligodeoxynucleotides: Antisense Inhibitors of Gene Expression*; Cohen, J.S., Ed. Macmillan Education UK: London, 1989; 137–172.
- [2] Lakowicz, J.R. *Principles of Fluorescence Spectroscopy*. Springer US: , 2006.
- [3] Gillies, R.J. In vivo molecular imaging. *J. Cell. Biochem.* **2002**, 87(S39), 231–238.
- [4] Suganya, S.; Naha, S.; Velmathi, S. A Critical Review on Colorimetric and Fluorescent Probes for the Sensing of Analytes via Relay Recognition from the year 2012–17. *ChemistrySelect* **2018**, 3, 7231–7268.
- [5] Kolanowski, J.L.; Liu, F.; New, E.J. Fluorescent probes for the simultaneous detection of multiple analytes in biology. *Chem. Soc. Rev.* **2018**, 47(1), 195–208.
- [6] Fu, Y.; Finney, N.S. Small-molecule fluorescent probes and their design. *RSC Adv.* **2018**, 8(51), 29051–29061.
- [7] Förster, T. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annalen der Physik* **1948**, 437(1–2), 55–75.
- [8] Stryer, L. Fluorescence Energy Transfer as a Spectroscopic Ruler. *Annu. Rev. Biochem.* **1978**, 47(1), 819–846.
- [9] Yuan, L.; Lin, W.; Zheng, K.; Zhu, S. FRET-Based Small-Molecule Fluorescent Probes: Rational Design and Bioimaging Applications. *Acc. Chem. Res.* **2013**, 46(7), 1462–1473.
- [10] Rowland, C.E.; Brown, C.W.; Medintz, I.L.; Delehanty, J.B. Intracellular FRET-based probes: a review. *Methods Appl. Fluoresc.* **2015**, 3(4), 042006.
- [11] Watson, J.D.; Crick, F.H.C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **1953**, 171(4356), 737–738.
- [12] Yakovchuk, P.; Protozanova, E.; Frank-Kamenetskii, M.D. Base-stacking and base-pairing contributions into thermal stability of the DNA double helix. *Nucleic Acids Res* **2006**, 34(2), 564–574.
- [13] Yang, Y.; Zhao, L. Sensitive fluorescent sensing for DNA assay. *TrAC Trends in Analytical Chemistry* **2010**, 29(9), 980–1003.
- [14] Didenko, V.V. DNA Probes Using Fluorescence Resonance Energy Transfer (FRET): Designs and Applications. *BioTechniques* **2001**, 31(5), 1106–1121.
- [15] Boutorine, A.; Novopashina, D.; Krasheninina, O.; Nozeret, K.; Venyaminova, A. Fluorescent Probes for Nucleic Acid Visualization in Fixed and Live Cells. *Molecules* **2013**, 18(12), 15357–15397.

- [16] Clausen-Schaumann, H.; Rief, M.; Tolksdorf, C.; Gaub, H.E. Mechanical Stability of Single DNA Molecules. *Biophysical Journal* **2000**, 78(4), 1997–2007.
- [17] Goulko, A.A.; Li, F.; Chris Le, X. Bioanalytical applications of aptamer and molecular-beacon probes in fluorescence-affinity assays. *TrAC Trends in Analytical Chemistry* **2009**, 28(7), 878–892.
- [18] Wang, K.; Tang, Z.; Yang, C.J.; Kim, Y.; Fang, X.; Li, W.; Wu, Y.; Medley, C.D.; Cao, Z.; Li, J.; Colon, P.; Lin, H.; Tan, W. Molecular engineering of DNA: molecular beacons. *Angew. Chem. Int. Ed. Engl.* **2009**, 48(5), 856–870.
- [19] Cardullo, R.A.; Agrawal, S.; Flores, C.; Zamecnik, P.C.; Wolf, D.E. Detection of nucleic acid hybridization by nonradiative fluorescence resonance energy transfer. *PNAS* **1988**, 85(23), 8790–8794.
- [20] Vasquez, K.M.; Glazer, P.M. Triplex-forming oligonucleotides: principles and applications. *Quarterly Reviews of Biophysics* **2002**, 35(1), 89–107.
- [21] Doluca, O.; Boutorine, A.S.; Filichev, V.V. Triplex-Forming Twisted Intercalating Nucleic Acids (TINAs): Design Rules, Stabilization of Antiparallel DNA Triplexes and Inhibition of G-Quartet-Dependent Self-Association. *ChemBioChem* **2011**, 12(15), 2365–2374.
- [22] Duca, M.; Vekhoff, P.; Oussedik, K.; Halby, L.; Arimondo, P.B. The triple helix: 50 years later, the outcome. *Nucleic Acids Res* **2008**, 36(16), 5123–5138.
- [23] Dervan, P.B.; Edelson, B.S. Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Current Opinion in Structural Biology* **2003**, 13(3), 284–299.
- [24] Blackledge, M.S.; Melander, C. Programmable DNA-binding small molecules. *Bioorganic & Medicinal Chemistry* **2013**, 21(20), 6101–6114.
- [25] Vaijayanthi, T.; Bando, T.; Pandian, G.N.; Sugiyama, H. Progress and Prospects of Pyrrole-Imidazole Polyamide-Fluorophore Conjugates as Sequence-Selective DNA Probes. *ChemBioChem* **2012**, 13(15), 2170–2185.
- [26] Fechter, E.J.; Olenyuk, B.; Dervan, P.B. Sequence-Specific Fluorescence Detection of DNA by Polyamide–Thiazole Orange Conjugates. *J. Am. Chem. Soc.* **2005**, 127(47), 16685–16691.
- [27] Zhang, W.; Minoshima, M.; Sugiyama, H. Base Pair Recognition of the Stereochemically  $\alpha$ -Substituted  $\gamma$ -Turn of Pyrrole/Imidazole Hairpin Polyamides. *J. Am. Chem. Soc.* **2006**, 128(46), 14905–14912.
- [28] Farkas, M.E.; Li, B.C.; Dose, C.; Dervan, P.B. DNA sequence selectivity of hairpin polyamide turn units. *Bioorganic & Medicinal Chemistry Letters* **2009**, 19(14), 3919–3923.
- [29] Smolina, I.V.; Demidov, V.V.; Soldatenkov, V.A.; Chasovskikh, S.G.; Frank-Kamenetskii, M.D. End invasion of peptide nucleic acids (PNAs) with mixed-base composition into linear DNA duplexes. *Nucleic Acids Res* **2005**, 33(17), e146–e146.

- [30] Nielsen, P.E.; Egholm, M.; Berg, R.H.; Buchardt, O. Sequence-selective recognition of DNA by strand displacement with a thymine-substituted polyamide. *Science* **1991**, *254*(5037), 1497–1500.
- [31] Vilaivan, T. Fluorogenic PNA probes. *Beilstein Journal of Organic Chemistry* **2018**, *14*, 253–281.
- [32] Smolina, I.V.; Kuhn, H.; Lee, C.; Frank-Kamenetskii, M.D. Fluorescence-based detection of short DNA sequences under non-denaturing conditions. *Bioorganic & Medicinal Chemistry* **2008**, *16*(1), 84–93.
- [33] Ma, H.; Reyes-Gutierrez, P.; Pederson, T. Visualization of repetitive DNA sequences in human chromosomes with transcription activator-like effectors. *PNAS* **2013**, *110*(52), 21048–21053.
- [34] Miller, J.C.; Tan, S.; Qiao, G.; Barlow, K.A.; Wang, J.; Xia, D.F.; Meng, X.; Paschon, D.E.; Leung, E.; Hinkley, S.J.; Dulay, G.P.; Hua, K.L.; Ankoudinova, I.; Cost, G.J.; Urnov, F.D.; Zhang, H.S.; Holmes, M.C.; Zhang, L.; Gregory, P.D.; Rebar, E.J. A TALE nuclease architecture for efficient genome editing. *Nature Biotechnology* **2011**, *29*(2), 143–148.
- [35] Urnov, F.D.; Rebar, E.J.; Holmes, M.C.; Zhang, H.S.; Gregory, P.D. Genome editing with engineered zinc finger nucleases. *Nat. Rev. Genet.* **2010**, *11*(9), 636–646.
- [36] Stains, C.I.; Porter, J.R.; Ooi, A.T.; Segal, D.J.; Ghosh, I. DNA Sequence-Enabled Reassembly of the Green Fluorescent Protein. *J. Am. Chem. Soc.* **2005**, *127*(31), 10782–10783.
- [37] Pavletich, N.P.; Pabo, C.O. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* **1991**, *252*(5007), 809–817.
- [38] Klug, A. The Discovery of Zinc Fingers and Their Applications in Gene Regulation and Genome Manipulation. *Annu. Rev. Biochem.* **2010**, *79*(1), 213–231.
- [39] Parraga, G.; Horvath, S.J.; Eisen, A.; Taylor, W.E.; Hood, L.; Young, E.T.; Klevit, R.E. Zinc-dependent structure of a single-finger domain of yeast ADR1. *Science* **1988**, *241*(4872), 1489–1492.
- [40] Michael, S.F.; Kilfoil, V.J.; Schmidt, M.H.; Amann, B.T.; Berg, J.M. Metal binding and folding properties of a minimalist Cys2His2 zinc finger peptide. *Proceedings of the National Academy of Sciences* **1992**, *89*(11), 4796–4800.
- [41] Eom, K.S.; Cheong, J.S.; Lee, S.J. Structural Analyses of Zinc Finger Domains for Specific Interactions with DNA. *Journal of Microbiology and Biotechnology* **2016**, *26*(12), 2019–2029.
- [42] Gersbach, C.A.; Gaj, T.; Barbas, C.F. Synthetic Zinc Finger Proteins: The Advent of Targeted Gene Regulation and Genome Modification Technologies. *Accounts of Chemical Research* **2014**, *47*(8), 2309–2318.
- [43] Persikov, A.V.; Wetzel, J.L.; Rowland, E.F.; Oakes, B.L.; Xu, D.J.; Singh, M.; Noyes, M.B. A systematic survey of the Cys2His2 zinc finger DNA-binding landscape. *Nucleic Acids Research* **2015**, *43*(3), 1965–1984.

- [44] Gersbach, C.A.; Gaj, T.; Barbas, C.F. Synthetic Zinc Finger Proteins: The Advent of Targeted Gene Regulation and Genome Modification Technologies. *Acc. Chem. Res.* **2014**, *47*(8), 2309–2318.
- [45] Wang, B.S.; Pabo, C.O. Dimerization of zinc fingers mediated by peptides evolved in vitro from random sequences. *Proceedings of the National Academy of Sciences* **1999**, *96*(17), 9568–9573.
- [46] Wang, B.S.; Grant, R.A.; Pabo, C.O. Selected peptide extension contacts hydrophobic patch on neighboring zinc finger and mediates dimerization on DNA. *Nature structural biology* **2001**, *8*(7), 589–593.
- [47] Yoshitake, K.; Waki, S.; Ueda, H. Dimerization-based homogeneous fluorosensor proteins for the detection of specific dsDNA. *Biosensors and Bioelectronics* **2008**, *23*(8), 1266–1271.
- [48] Pearson, R.G. Hard and Soft Acids and Bases. *J. Am. Chem. Soc.* **1963**, *85*, 3533–3539.
- [49] Bünzli, J.-C.G. Review: Lanthanide coordination chemistry: from old concepts to coordination polymers. *Journal of Coordination Chemistry* **2014**, *67*(23–24), 3706–3733.
- [50] Yu, D.; Du, R.; Xiao, J.-C.; Xu, S.; Rong, C.; Liu, S. Theoretical Study of pK<sub>a</sub> Values for Trivalent Rare-Earth Metal Cations in Aqueous Solution. *J. Phys. Chem. A* **2018**, *122*(2), 700–707.
- [51] Viswanathan, S.; Kovacs, Z.; Green, K.N.; Ratnakar, S.J.; Sherry, A.D. Alternatives to Gadolinium-Based Metal Chelates for Magnetic Resonance Imaging <sup>†</sup>. *Chem. Rev.* **2010**, *110*(5), 2960–3018.
- [52] Clough, T.J.; Jiang, L.; Wong, K.-L.; Long, N.J. Ligand design strategies to increase stability of gadolinium-based magnetic resonance imaging contrast agents. *Nat Commun* **2019**, *10*(1), 1420.
- [53] Bünzli, J.-C.G.; Eliseeva, S.V. Basics of Lanthanide Photophysics. In *Lanthanide Luminescence*; Hänninen, P., Härmä, H., Eds. Springer Berlin Heidelberg, 2010; Vol. 7, , 1–45.
- [54] Vleck, J.H. Van. The Puzzle of Rare-earth Spectra in Solids. *J. Phys. Chem.* **1937**, *41*(1), 67–80.
- [55] Xu, J.; Gulzar, A.; Yang, P.; Bi, H.; Yang, D.; Gai, S.; He, F.; Lin, J.; Xing, B.; Jin, D. Recent advances in near-infrared emitting lanthanide-doped nanoconstructs: Mechanism, design and application for bioimaging. *Coordination Chemistry Reviews* **2019**, *381*, 104–134.
- [56] Thibon, A.; Pierre, V.C. Principles of responsive lanthanide-based luminescent probes for cellular imaging. *Analytical and Bioanalytical Chemistry* **2009**, *394*(1), 107–120.
- [57] Hanaoka, K.; Kikuchi, K.; Kobayashi, S.; Nagano, T. Time-resolved long-lived luminescence imaging method employing luminescent lanthanide probes with a new microscopy system. *J. Am. Chem. Soc.* **2007**, *129*(44), 13502–13509.
- [58] Beeby, A.; Clarkson, I.M.; Dickins, R.S.; Faulkner, S.; Parker, D.; Royle, L.; Sousa, A.S. de; Williams, J.A.G.; Woods, M. Non-radiative deactivation of the excited states of europium,

terbium and ytterbium complexes by proximate energy-matched OH, NH and CH oscillators: an improved luminescence method for establishing solution hydration states. *J. Chem. Soc., Perkin Trans. 2* **1999**, (3), 493–504.

[59] Supkowski, R.M.; Horrocks, W.DeW. On the determination of the number of water molecules,  $q$ , coordinated to europium(III) ions in solution from luminescence decay lifetimes. *Inorganica Chimica Acta* **2002**, 340, 44–48.

[60] Bettencourt-Dias, A. Small Molecule Luminescent Lanthanide Ion Complexes - Photophysical Characterization and Recent Developments. *Current Organic Chemistry* **2007**, 11(16), 1460–1480.

[61] Sharma, G.; Sehgal, P.; Narula, A.K. Luminescent Lanthanide Sensors and Lanthanide Doped Upconversion Nanoparticles: Current Status and Future Expectations. In *Reviews in Fluorescence 2015*; Geddes, C.D., Ed. Springer International Publishing: Cham, 2016; 269–299.

[62] Weissman, S.I. Intramolecular Energy Transfer The Fluorescence of Complexes of Europium. *The Journal of Chemical Physics* **1942**, 10(4), 214–217.

[63] Bünzli, J.-C.G.; Piguet, C. Taking advantage of luminescent lanthanide ions. *Chem. Soc. Rev.* **2005**, 34(12), 1048–1077.

[64] Bünzli, J.-C.G. On the design of highly luminescent lanthanide complexes. *Coordination Chemistry Reviews* **2015**, 293–294, 19–47.

[65] Crosby, G.A.; Whan, R.E.; Alire, R.M. Intramolecular Energy Transfer in Rare Earth Chelates. Role of the Triplet State. *The Journal of Chemical Physics* **1961**, 34(3), 743–748.

[66] Latvaa, M.; Takalob, H.; Mukkala, V.-M.; Rodriguez-Ubis, J.C.; Kankarea, J. Correlation between the lowest triplet state energy level of the ligand and lanthanide(III) luminescence quantum yield. *Journal of Luminescence* **1997**, 75(2), 149–169.

[67] Leonard, J.P.; Santos, C.M.G. dos; Plush, S.E.; McCabe, T.; Gunnlaugsson, T. PH driven self-assembly of a ternary lanthanide luminescence complex: the sensing of anions using a beta-diketonate-Eu(III) displacement assay. *Chem. Commun. (Camb.)* **2007**, (2), 129–131.

[68] Lee, K.; Dzubeck, V.; Latshaw, L.; Schneider, J.P. De novo designed peptidic redox potential probe: linking sensitized emission to disulfide bond formation. *J. Am. Chem. Soc.* **2004**, 126(42), 13616–13617.

[69] Thibon, A.; Pierre, V.C. A Highly Selective Luminescent Sensor for the Time-Gated Detection of Potassium. *J. Am. Chem. Soc.* **2009**, 131(2), 434–435.

[70] Viguiier, R.F.H.; Hulme, A.N. A sensitized europium complex generated by micromolar concentrations of copper(I): toward the detection of copper(I) in biology. *J. Am. Chem. Soc.* **2006**, 128(35), 11370–11371.

[71] Isaac, M.; Denisov, S.A.; Roux, A.; Imbert, D.; Jonusauskas, G.; McClenaghan, N.D.; Sénèque, O. Lanthanide Luminescence Modulation by Cation- $\pi$  Interaction in a Bioinspired Scaffold: Selective Detection of Copper(I). *Angewandte Chemie* **2015**, 127(39), 11615–11618.

- [72] Halim, M.; Tremblay, M.S.; Jockusch, S.; Turro, N.J.; Sames, D. Transposing Molecular Fluorescent Switches into the Near-IR: Development of Luminogenic Reporter Substrates for Redox Metabolism. *J. Am. Chem. Soc.* **2007**, *129*(25), 7704–7705.
- [73] Pershagen, E.; Nordholm, J.; Borbas, K.E. Luminescent Lanthanide Complexes with Analyte-Triggered Antenna Formation. *J. Am. Chem. Soc.* **2012**, *134*(24), 9832–9835.
- [74] Chen, Y.; Guo, W.; Ye, Z.; Wang, G.; Yuan, J. A europium(III) chelate as an efficient time-gated luminescent probe for nitric oxide. *Chem. Commun.* **2011**, 47(22), 6266–6268.
- [75] Song, B.; Wang, G.; Tan, M.; Yuan, J. A Europium(III) Complex as an Efficient Singlet Oxygen Luminescence Probe. *J. Am. Chem. Soc.* **2006**, *128*(41), 13442–13450.
- [76] Dai, Z.; Tian, L.; Song, B.; Ye, Z.; Liu, X.; Yuan, J. Ratiometric Time-Gated Luminescence Probe for Hydrogen Sulfide Based on Lanthanide Complexes. *Anal. Chem.* **2014**, *86*(23), 11883–11889.
- [77] Smith, D.G.; Pal, R.; Parker, D. Measuring Equilibrium Bicarbonate Concentrations Directly in Cellular Mitochondria and in Human Serum Using Europium/Terbium Emission Intensity Ratios. *Chemistry – A European Journal* **2012**, *18*(37), 11604–11613.
- [78] McMahon, B.K.; Pal, R.; Parker, D. A bright and responsive europium probe for determination of pH change within the endoplasmic reticulum of living cells. *Chem. Commun.* **2013**, 49(47), 5363–5365.
- [79] Law, G.-L.; Pal, R.; Palsson, L.O.; Parker, D.; Wong, K.-L. Responsive and reactive terbium complexes with an azaxanthone sensitiser and one naphthyl group: applications in ratiometric oxygen sensing in vitro and in regioselective cell killing. *Chem. Commun.* **2009**, (47), 7321–7323.
- [80] Sørensen, T.J.; Kenwright, A.M.; Faulkner, S. Bimetallic lanthanide complexes that display a ratiometric response to oxygen concentrations. *Chem. Sci.* **2015**, *6*(3), 2054–2059.
- [81] Bünzli, J.-C.G. Lanthanide Luminescence for Biomedical Analyses and Imaging. *Chem. Rev.* **2010**, *110*(5), 2729–2755.
- [82] Bünzli, J.-C.G. Lanthanide light for biology and medical diagnosis. *Journal of Luminescence* **2016**, *170*, 866–878.
- [83] Sy, M.; Esteban-Gómez, D.; Platas-Iglesias, C.; Rodríguez-Rodríguez, A.; Tripier, R.; Charbonnière, L.J. Spectroscopic Properties of a Family of Mono- to Trinuclear Lanthanide Complexes. *European Journal of Inorganic Chemistry* **2017**, 2017(14), 2122–2129.
- [84] Delbianco, M.; Sadovnikova, V.; Bourrier, E.; Mathis, G.; Lamarque, L.; Zwier, J.M.; Parker, D. Bright, highly water-soluble triazacyclononane europium complexes to detect ligand binding with time-resolved FRET microscopy. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*(40), 10718–10722.
- [85] Li, M.; Selvin, P.R. Luminescent Polyaminocarboxylate Chelates of Terbium and Europium: The Effect of Chelate Structure. *J. Am. Chem. Soc.* **1995**, *117*(31), 8132–8138.

- [86] Kovacs, D.; Lu, X.; Mészáros, L.S.; Ott, M.; Andres, J.; Borbas, K.E. Photophysics of Coumarin and Carbostyryl-Sensitized Luminescent Lanthanide Complexes: Implications for Complex Design in Multiplex Detection. *J. Am. Chem. Soc.* **2017**, *139*(16), 5756–5767.
- [87] Xiao, M.; Selvin, P.R. Quantum Yields of Luminescent Lanthanide Chelates and Far-Red Dyes Measured by Resonance Energy Transfer. *J. Am. Chem. Soc.* **2001**, *123*(29), 7067–7073.
- [88] Krasnoperov, L.N.; Marras, S.A.E.; Kozlov, M.; Wirpsza, L.; Mustaev, A. Luminescent Probes for Ultrasensitive Detection of Nucleic Acids. *Bioconjugate Chem.* **2010**, *21*(2), 319–327.
- [89] Li, L.-L.; Ge, P.; Selvin, P.R.; Lu, Y. Direct Detection of Adenosine in Undiluted Serum Using a Luminescent Aptamer Sensor Attached to a Terbium Complex. *Anal. Chem.* **2012**, *84*(18), 7852–7856.
- [90] Reynolds, A.M.; Sculimbrene, B.R.; Imperiali, B. Lanthanide-Binding Tags with Unnatural Amino Acids: Sensitizing Tb<sup>3+</sup> and Eu<sup>3+</sup> Luminescence at Longer Wavelengths. *Bioconjugate Chem.* **2008**, *19*(3), 588–591.
- [91] Isaac, M.; Raibaut, L.; Cepeda, C.; Roux, A.; Boturyn, D.; Eliseeva, S.V.; Petoud, S.; Sénèque, O. Luminescent Zinc Fingers: Zn-Responsive Neodymium Near-Infrared Emission in Water. *Chem. Eur. J.* **2017**, *23*(46), 10992–10996.
- [92] Raibaut, L.; Vasseur, W.; Shimberg, G.D.; Saint-Pierre, C.; Ravanat, J.-L.; Michel, S.L.J.; Sénèque, O. Design of a synthetic luminescent probe from a biomolecule binding domain: selective detection of AU-rich mRNA sequences. *Chem. Sci.* **2017**, *8*(2), 1658–1664.
- [93] Reynolds, A.M.; Sculimbrene, B.R.; Imperiali, B. Lanthanide-Binding Tags with Unnatural Amino Acids: Sensitizing Tb<sup>3+</sup> and Eu<sup>3+</sup> Luminescence at Longer Wavelengths. *Bioconjugate Chem.* **2008**, *19*(3), 588–591.
- [94] Fahey, R.C.; Hunt, J.S.; Windham, G.C. On the cysteine and cystine content of proteins: Differences between Intracellular and Extracellular Proteins. *J Mol Evol* **1977**, *10*(2), 155–160.
- [95] Dawson, P.E.; Muir, T.W.; Clark-Lewis, I.; Kent, S.B. Synthesis of proteins by native chemical ligation. *Science* **1994**, *266*(5186), 776–779.
- [96] Kent, S.B.H. Total chemical synthesis of proteins. *Chem Soc Rev* **2009**, *38*(2), 338–351.
- [97] Agouridas, V.; El Mahdi, O.; Cargoët, M.; Melnyk, O. A statistical view of protein chemical synthesis using NCL and extended methodologies. *Bioorganic & Medicinal Chemistry* **2017**, *25*(18), 4938–4945.
- [98] Walker, M.A. Protein Synthesis by Chemical Ligation of Unprotected Peptides in Aqueous Solution. *Angewandte Chemie International Edition in English* **1997**, *36*(10), 1069–1071.
- [99] Canne, L.E.; Botti, P.; Simon, R.J.; Chen, Y.; Dennis, E.A.; Kent, S.B.H. Chemical Protein Synthesis by Solid Phase Ligation of Unprotected Peptide Segments. *J. Am. Chem. Soc.* **1999**, *121*(38), 8720–8727.

- [100] Johnson, E.C.B.; Kent, S.B.H. Insights into the Mechanism and Catalysis of the Native Chemical Ligation Reaction. *J. Am. Chem. Soc.* **2006**, *128*(20), 6640–6646.
- [101] Dawson, P.E.; Churchill, M.J.; Ghadiri, M.R.; Kent, S.B.H. Modulation of Reactivity in Native Chemical Ligation through the Use of Thiol Additives. *J. Am. Chem. Soc.* **1997**, *119*(19), 4325–4329.
- [102] Camarero, J.A.; Adeva, A.; Muir, T.W. 3-thiopropionic acid as a highly versatile multidetachable thioester resin linker. *Lett Pept Sci* **2000**, *7*(1), 17–21.
- [103] Bang, D.; Pentelute, B.L.; Gates, Z.P.; Kent, S.B. Direct On-Resin Synthesis of Peptide- $\alpha$ Thiophenylesters for Use in Native Chemical Ligation. *Org. Lett.* **2006**, *8*(6), 1049–1052.
- [104] Bu, X.; Xie, G.; Law, C.W.; Guo, Z. An improved deblocking agent for direct Fmoc solid-phase synthesis of peptide thioesters. *Tetrahedron Letters* **2002**, *43*(13), 2419–2422.
- [105] Futaki, S.; Sogawa, K.; Maruyama, J.; Asahara, T.; Niwa, M.; Hojo, H. Preparation of Peptide Thioesters using Fmoc-Solid-Phase Peptide Synthesis and its Application to the Construction of a Template-Assembled Synthetic Protein (TASP). *Tetrahedron Letters* **1997**, *38*(35), 6237–6240.
- [106] Ohta, Y.; Itoh, S.; Shigenaga, A.; Shintaku, S.; Fujii, N.; Otaka, A. Cysteine-Derived S-Protected Oxazolidinones: Potential Chemical Devices for the Preparation of Peptide Thioesters. *Org. Lett.* **2006**, *8*(3), 467–470.
- [107] Nagaike, F.; Onuma, Y.; Kanazawa, C.; Hojo, H.; Ueki, A.; Nakahara, Y.; Nakahara, Y. Efficient Microwave-Assisted Tandem N- to S-Acyl Transfer and Thioester Exchange for the Preparation of a Glycosylated Peptide Thioester. *Org. Lett.* **2006**, *8*(20), 4465–4468.
- [108] Kawakami, T.; Aimoto, S. Peptide Ligation Using a Building Block Having a Cysteinyll Prolyl Ester (CPE) Autoactivating Unit at the Carboxy Terminus. *Chem. Lett.* **2007**, *36*(1), 76–77.
- [109] Sato, K.; Shigenaga, A.; Tsuji, K.; Tsuda, S.; Sumikawa, Y.; Sakamoto, K.; Otaka, A. N-Sulfanylethylanilide Peptide as a Crypto-Thioester Peptide. *ChemBioChem* **2011**, *12*(12), 1840–1844.
- [110] Burlina, F.; Papageorgiou, G.; Morris, C.; D. White, P.; Offer, J. In situ thioester formation for protein ligation using  $\alpha$ -methylcysteine. *Chemical Science* **2014**, *5*(2), 766–770.
- [111] Asahina, Y.; Nabeshima, K.; Hojo, H. Peptidyl N-alkylcysteine as a peptide thioester surrogate in the native chemical ligation. *Tetrahedron Letters* **2015**, *56*(11), 1370–1373.
- [112] Ollivier, N.; Dheur, J.; Mhidia, R.; Blanpain, A.; Melnyk, O. Bis(2-sulfanylethyl)amino Native Peptide Ligation. *Org. Lett.* **2010**, *12*(22), 5238–5241.
- [113] Terrier, V.P.; Adihou, H.; Arnould, M.; Delmas, A.F.; Aucagne, V. A straightforward method for automated Fmoc-based synthesis of bio-inspired peptide crypto-thioesters. *Chem. Sci.* **2016**, *7*(1), 339–345.

- [114] Ollivier, N.; Dheur, J.; Mhidia, R.; Blanpain, A.; Melnyk, O. Bis(2-sulfanylethyl)amino Native Peptide Ligation. *Org. Lett.* **2010**, *12*(22), 5238–5241.
- [115] Ollivier, N.; Vicogne, J.; Vallin, A.; Drobecq, H.; Desmet, R.; El Mahdi, O.; Leclercq, B.; Goormachtigh, G.; Fafeur, V.; Melnyk, O. A One-Pot Three-Segment Ligation Strategy for Protein Chemical Synthesis. *Angew. Chem. Int. Ed.* **2012**, *51*(1), 209–213.
- [116] Carugo, O. Amino acid composition and protein dimension. *Protein Science* **2008**, *17*(12), 2187–2191.
- [117] Hackeng, T.M.; Griffin, J.H.; Dawson, P.E. Protein synthesis by native chemical ligation: Expanded scope by using straightforward methodology. *Proceedings of the National Academy of Sciences* **1999**, *96*(18), 10068–10073.
- [118] Boll, E.; Drobecq, H.; Ollivier, N.; Blanpain, A.; Raibaut, L.; Desmet, R.; Vicogne, J.; Melnyk, O. One-pot chemical synthesis of small ubiquitin-like modifier protein–peptide conjugates using bis(2-sulfanylethyl)amido peptide latent thioester surrogates. *Nat Protoc* **2015**, *10*(2), 269–292.
- [119] Bang, D.; Pentelute, B.L.; Kent, S.B.H. Kinetically Controlled Ligation for the Convergent Chemical Synthesis of Proteins. *Angewandte Chemie International Edition* **2006**, *45*(24), 3985–3988.
- [120] Kulkarni, S.S.; Sayers, J.; Premdjee, B.; Payne, R.J. Rapid and efficient protein synthesis through expansion of the native chemical ligation concept. *Nat Rev Chem* **2018**, *2*(4), 0122.
- [121] Ollivier, N.; Vicogne, J.; Vallin, A.; Drobecq, H.; Desmet, R.; El Mahdi, O.; Leclercq, B.; Goormachtigh, G.; Fafeur, V.; Melnyk, O. A One-Pot Three-Segment Ligation Strategy for Protein Chemical Synthesis. *Angew. Chem. Int. Ed.* **2012**, *51*(1), 209–213.
- [122] Snella, B.; Diemer, V.; Drobecq, H.; Agouridas, V.; Melnyk, O. Native Chemical Ligation at Serine Revisited. *Org. Lett.* **2018**, *20*(23), 7616–7619.
- [123] Mergler, M.; Dick, F.; Sax, B.; Weiler, P.; Vorherr, T. The aspartimide problem in Fmoc-based SPPS. Part I. *J. Peptide Sci.* **2003**, *9*(1), 36–46.
- [124] Mergler, M.; Dick, F.; Sax, B.; Stähelin, C.; Vorherr, T. The aspartimide problem in Fmoc-based SPPS. Part II. *Journal of Peptide Science* **2003**, *9*(8), 518–526.
- [125] Stathopoulos, P.; Papas, S.; Kostidis, S.; Tsikaris, V.  $\alpha$ - and  $\beta$ - aspartyl peptide ester formation via aspartimide ring opening. *J. Peptide Sci.* **2005**, *11*(10), 658–664.
- [126] Dölling, R.; Beyermann, M.; Haenel, J.; Kernchen, F.; Krause, E.; Franke, P.; Brudel, M.; Bienert, M. Piperidine-mediated side product formation for Asp(OBut)-containing peptides. *J. Chem. Soc., Chem. Commun.* **1994**, (7), 853–854.
- [127] Lauer, J.L.; Fields, C.G.; Fields, G.B. Sequence dependence of aspartimide formation during 9-fluorenylmethoxycarbonyl solid-phase peptide synthesis. *Lett Pept Sci* **1995**, *1*(4), 197–205.

- [128] Quibell, M.; Owen, D.; Packman, L.C.; Johnson, T. Suppression of piperidine-mediated side product formation for Asp(OBut)-containing peptides by the use of N-(2-hydroxy-4-methoxybenzyl)(Hmb) backbone amide protection. *J. Chem. Soc., Chem. Commun.* **1994**, (20), 2343–2344.
- [129] Behrendt, R.; Huber, S.; Martí, R.; White, P. New t-butyl based aspartate protecting groups preventing aspartimide formation in Fmoc SPPS. *J. Pept. Sci.* **2015**, 21(8), 680–687.
- [130] Subirós-Funosas, R.; El-Faham, A.; Albericio, F. Aspartimide formation in peptide chemistry: occurrence, prevention strategies and the role of N-hydroxylamines. *Tetrahedron* **2011**, 67(45), 8595–8606.
- [131] Subirós-Funosas, R.; El-Faham, A.; Albericio, F. Use of Oxyma as pH modulatory agent to be used in the prevention of base-driven side reactions and its effect on 2-chlorotrityl chloride resin. *Peptide Science* **2012**, 98(2), 89–97.
- [132] Michels, T.; Dölling, R.; Haberkorn, U.; Mier, W. Acid-Mediated Prevention of Aspartimide Formation in Solid Phase Peptide Synthesis. *Org. Lett.* **2012**, 14(20), 5218–5221.
- [133] Wade, J.D.; Mathieu, M.N.; Macris, M.; Tregear, G.W. Base-induced side reactions in Fmoc-solid phase peptide synthesis: Minimization by use of piperazine as N<sup>α</sup>-deprotection reagent. *Lett Pept Sci* **2000**, 7(2), 107–112.
- [134] Lauer, J.L.; Fields, C.G.; Fields, G.B. Sequence dependence of aspartimide formation during 9-fluorenylmethoxycarbonyl solid-phase peptide synthesis. *Lett Pept Sci* **1995**, 1(4), 197–205.
- [135] Sy, M.; Nonat, A.; Hildebrandt, N.; J. Charbonnière, L. Lanthanide-based luminescence biolabelling. *Chemical Communications* **2016**, 52(29), 5080–5095.
- [136] O'Malley, W.I.; Abdelkader, E.H.; Aulsebrook, M.L.; Rubbiani, R.; Loh, C.-T.; Grace, M.R.; Spiccia, L.; Gasser, G.; Otting, G.; Tuck, K.L.; Graham, B. Luminescent Alkyne-Bearing Terbium(III) Complexes and Their Application to Bioorthogonal Protein Labeling. *Inorg. Chem.* **2016**, 55(4), 1674–1682.
- [137] Lavis, L.D.; Raines, R.T. Bright Ideas for Chemical Biology. *ACS Chem. Biol.* **2008**, 3(3), 142–155.
- [138] Deng, F.; Xu, Z. Heteroatom-substituted rhodamine dyes: Structure and spectroscopic properties. *Chinese Chemical Letters* **2019**, 30(10), 1667–1681.
- [139] Wang, L.; Du, W.; Hu, Z.; Uvdal, K.; Li, L.; Huang, W. Hybrid Rhodamine Fluorophores in the Visible/NIR Region for Biological Imaging. *Angewandte Chemie International Edition* **2019**, 58(40), 14026–14043.
- [140] Smith, S.N.; Steer, R.P. The photophysics of Lissamine rhodamine-B sulphonyl chloride in aqueous solution: implications for fluorescent protein–dye conjugates. *Journal of Photochemistry and Photobiology A: Chemistry* **2001**, 139(2), 151–156.
- [141] Koide, Y.; Urano, Y.; Hanaoka, K.; Terai, T.; Nagano, T. Development of an Si-Rhodamine-Based Far-Red to Near-Infrared Fluorescence Probe Selective for Hypochlorous

Acid and Its Applications for Biological Imaging. *J. Am. Chem. Soc.* **2011**, *133*(15), 5680–5682.

[142] Koide, Y.; Urano, Y.; Hanaoka, K.; Terai, T.; Nagano, T. Evolution of Group 14 Rhodamines as Platforms for Near-Infrared Fluorescence Probes Utilizing Photoinduced Electron Transfer. *ACS Chem. Biol.* **2011**, *6*(6), 600–608.

[143] Webster, K.A.; Prentice, H.; Bishopric, N.H. Oxidation of Zinc Finger Transcription Factors: Physiological Consequences. *Antioxidants & Redox Signaling* **2001**, *3*(4), 535–548.

[144] Rougee, M.; Bensasson, R.V.; Land, E.J.; Pariente, R. Deactivation of Singlet Molecular Oxygen by Thiols and Related Compounds, Possible Protectors Against Skin Photosensitivity. *Photochemistry and Photobiology* **1988**, *47*(4), 485–489.

[145] Devasagayam, T.P.A.; Sundquist, A.R.; Di Mascio, P.; Kaiser, S.; Sies, H. Activity of thiols as singlet molecular oxygen quenchers. *Journal of Photochemistry and Photobiology B: Biology* **1991**, *9*(1), 105–116.

[146] Klotz, L.-O.; Kröncke, K.-D.; Sies, H. Singlet oxygen-induced signaling effects in mammalian cells. *Photochemical & Photobiological Sciences* **2003**, *2*(2), 88–94.

[147] Lebrun, V.; Tron, A.; Scarpantonio, L.; Lebrun, C.; Ravanat, J.-L.; Latour, J.-M.; McClenaghan, N.D.; Sénèque, O. Efficient Oxidation and Destabilization of Zn(Cys)<sub>4</sub> Zinc Fingers by Singlet Oxygen. *Angew. Chem. Int. Ed.* **2014**, *53*(35), 9365–9368.

[148] Lebrun, V.; Tron, A.; Lebrun, C.; Latour, J.; McClenaghan, N.D.; Sénèque, O. Reactivity of a Zn(Cys)<sub>2</sub>(His)<sub>2</sub> Zinc Finger with Singlet Oxygen: Oxidation Directed toward Cysteines but not Histidines. *Chem. Eur. J.* **2015**, *21*(40), 14002–14010.

[149] Chabert, V.; Lebrun, V.; Lebrun, C.; Latour, J.-M.; Sénèque, O. Model peptide for anti-sigma factor domain HHCC zinc fingers: high reactivity toward <sup>1</sup>O<sub>2</sub> leads to domain unfolding. *Chem. Sci.* **2019**, *10*(12), 3608–3615.

[150] Umezawa, K.; Yoshida, M.; Kamiya, M.; Yamasoba, T.; Urano, Y. Rational design of reversible fluorescent probes for live-cell imaging and quantification of fast glutathione dynamics. *Nature Chemistry* **2017**, *9*(3), 279–286.

[151] Lukinavičius, G.; Umezawa, K.; Olivier, N.; Honigsmann, A.; Yang, G.; Plass, T.; Mueller, V.; Reymond, L.; Corrêa Jr, I.R.; Luo, Z.-G.; Schultz, C.; Lemke, E.A.; Heppenstall, P.; Eggeling, C.; Manley, S.; Johnsson, K. A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins. *Nature Chem* **2013**, *5*(2), 132–139.

[152] Weissman, S.I. Intramolecular Energy Transfer The Fluorescence of Complexes of Europium. *J. Chem. Phys.* **1942**, *10*(4), 214–217.

[153] Auzel, F. Compteur quantique par transfert d'énergie entre deux ions de terres rares dans un tungstate mixte et dans un verre. *C. R. Acad. Sci. Paris* **1966**, *262*, 1016–1019.

[154] Mitrunen, K.; Pettersson, K.; Piironen, T.; Björk, T.; Lilja, H.; Lövgren, T. Dual-label one-step immunoassay for simultaneous measurement of free and total prostate-specific antigen concentrations and ratios in serum. *Clinical Chemistry* **1995**, *41*(8), 1115–1120.

- [155] Evangelista, R.A.; Pollak, A.; Allore, B.; Templeton, A.F.; Morton, R.C.; Diamandis, E.P. A new europium chelate for protein labelling and time-resolved fluorometric applications. *Clinical Biochemistry* **1988**, *21*(3), 173–178.
- [156] Mikkala, V.-M.; Takalo, H.; Liitti, P.; Hemmilä, I. Synthesis and luminescence properties of some Eu(III) and Tb(III) chelate labels having 2,2':6',2''-terpyridine as an energy absorbing part. *Journal of Alloys and Compounds* **1995**, *225*(1), 507–510.
- [157] Weibel, N.; Charbonnière, L.J.; Guardigli, M.; Roda, A.; Ziessel, R. Engineering of Highly Luminescent Lanthanide Tags Suitable for Protein Labeling and Time-Resolved Luminescence Imaging. *Journal of the American Chemical Society* **2004**, *126*(15), 4888–4896.
- [158] Morcos, S.K. Extracellular gadolinium contrast agents: Differences in stability. *European Journal of Radiology* **2008**, *66*(2), 175–179.
- [159] Takalo, H.; Hemmilä, I.; Sutela, T.; Latva, M. Synthesis and Luminescence of Novel Eu(III) Complexing Agents and Labels with 4-(Phenylethynyl)pyridine Subunits. *Helvetica Chimica Acta* **1996**, *79*(3), 789–802.
- [160] Picot, A.; D'Aléo, A.; Baldeck, P.L.; Grichine, A.; Duperray, A.; Andraud, C.; Maury, O. Long-Lived Two-Photon Excited Luminescence of Water-Soluble Europium Complex: Applications in Biological Imaging Using Two-Photon Scanning Microscopy. *J. Am. Chem. Soc.* **2008**, *130*(5), 1532–1533.
- [161] Soulié, M.; Latzko, F.; Bourrier, E.; Placide, V.; Butler, S.J.; Pal, R.; Walton, J.W.; Baldeck, P.L.; Le Guennic, B.; Andraud, C.; Zwieter, J.M.; Lamarque, L.; Parker, D.; Maury, O. Comparative Analysis of Conjugated Alkynyl Chromophore-Triazacyclononane Ligands for Sensitized Emission of Europium and Terbium. *Chem. Eur. J.* **2014**, *20*(28), 8636–8646.
- [162] Kong, H.-K.; Chadbourne, F.L.; Law, G.-L.; Li, H.; Tam, H.-L.; Cobb, S.L.; Lau, C.-K.; Lee, C.-S.; Wong, K.-L. Two-photon induced responsive f–f emissive detection of Cyclin A with a europium-chelating peptide. *Chem. Commun.* **2011**, *47*(28), 8052–8054.
- [163] Bui, A.T.; Grichine, A.; Brasselet, S.; Duperray, A.; Andraud, C.; Maury, O. Unexpected Efficiency of a Luminescent Samarium(III) Complex for Combined Visible and Near-Infrared Biphononic Microscopy. *Chemistry - A European Journal* **2015**, *21*(49), 17757–17761.
- [164] Walton, J.W.; Bourdolle, A.; Butler, S.J.; Soulie, M.; Delbianco, M.; McMahon, B.K.; Pal, R.; Puschmann, H.; Zwieter, J.M.; Lamarque, L.; Maury, O.; Andraud, C.; Parker, D. Very bright europium complexes that stain cellular mitochondria. *Chem. Commun.* **2013**, *49*(16), 1600–1602.
- [165] Placide, V.; Pitrat, D.; Grichine, A.; Duperray, A.; Andraud, C.; Maury, O. Design and synthesis of europium luminescent bio-probes featuring sulfobetaine moieties. *Tetrahedron Letters* **2014**, *55*(7), 1357–1361.
- [166] D'Aléo, A.; Allali, M.; Picot, A.; Baldeck, P.L.; Toupet, L.; Andraud, C.; Maury, O. Sensitization of Eu(III) luminescence by donor-phenylethynyl-functionalized DTPA and DO3A macrocycles. *Comptes Rendus Chimie* **2010**, *13*(6–7), 681–690.

- [167] V. Werts, M.H.; Nerambourg, N.; Pélégry, D.; Le Grand, Y.; Blanchard-Desce, M. Action cross sections of two-photon excited luminescence of some Eu(III) and Tb(III) complexes. *Photochemical & Photobiological Sciences* **2005**, 4(7), 531–538.
- [168] Walter, E.R.H.; Williams, J.A.G.; Parker, D. Solvent polarity and oxygen sensitivity, rather than viscosity, determine lifetimes of biaryl-sensitised terbium luminescence. *Chem. Commun.* **2017**, 53(100), 13344–13347.
- [169] Bui, A.T.; Grichine, A.; Duperray, A.; Lidon, P.; Riobé, F.; Andraud, C.; Maury, O. Terbium(III) Luminescent Complexes as Millisecond-Scale Viscosity Probes for Lifetime Imaging. *J. Am. Chem. Soc.* **2017**, 139(23), 7693–7696.
- [170] Bui, A.T.; Roux, A.; Grichine, A.; Duperray, A.; Andraud, C.; Maury, O. Twisted Charge-Transfer Antennae for Ultra-Bright Terbium(III) and Dysprosium(III) Bioprobes. *Chem. Eur. J.* **2018**, 24(14), 3408–3412.
- [171] Tornøe, C.W.; Christensen, C.; Meldal, M. Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides. *J. Org. Chem.* **2002**, 67(9), 3057–3064.
- [172] Pounder, R.J.; Stanford, M.J.; Brooks, P.; Richards, S.P.; Dove, A.P. Metal free thiol–maleimide ‘Click’ reaction as a mild functionalisation strategy for degradable polymers. *Chem. Commun.* **2008**, (41), 5158–5160.
- [173] Rose, K. Facile synthesis of homogeneous artificial proteins. *J. Am. Chem. Soc.* **1994**, 116(1), 30–33.
- [174] Blackman, M.L.; Royzen, M.; Fox, J.M. Tetrazine Ligation: Fast Bioconjugation Based on Inverse-Electron-Demand Diels–Alder Reactivity. *J. Am. Chem. Soc.* **2008**, 130(41), 13518–13519.
- [175] De León-Rodriguez, L.M.; Kovacs, Z.; Dieckmann, G.R.; Sherry, A.D. Solid-Phase Synthesis of DOTA–Peptides. *Chem. Eur. J.* **2004**, 10(5), 1149–1155.
- [176] Barnhill, H.N.; Claudel-Gillet, S.; Ziessel, R.; Charbonnière, L.J.; Wang, Q. Prototype Protein Assembly as Scaffold for Time-Resolved Fluoroimmuno Assays. *J. Am. Chem. Soc.* **2007**, 129(25), 7799–7806.
- [177] Bourdolle, A.; Allali, M.; Mulatier, J.-C.; Le Guennic, B.; Zwier, J.M.; Baldeck, P.L.; Bünzli, J.-C.G.; Andraud, C.; Lamarque, L.; Maury, O. Modulating the Photophysical Properties of Azamacrocyclic Europium Complexes with Charge-Transfer Antenna Chromophores. *Inorganic Chemistry* **2011**, 50(11), 4987–4999.
- [178] Jagadish, B.; Ozumerzifon, T.J.; Roberts, S.A.; Hall, G.B.; Mash, E.A.; Raghunand, N. Improved Synthesis of 10-(2-Alkylamino-2-oxoethyl)-1,4,7,10-tetraazacyclododecane-1,4,7-triacetic Acid Derivatives Bearing Acid-Sensitive Linkers. *Synthetic Communications* **2014**, 44(3), 441–449.
- [179] Hundertmark, T.; Littke, A.F.; Buchwald, S.L.; Fu, G.C. Pd(PhCN)<sub>2</sub>Cl<sub>2</sub>/P(*t*-Bu)<sub>3</sub>: A Versatile Catalyst for Sonogashira Reactions of Aryl Bromides at Room Temperature. *Org. Lett.* **2000**, 2(12), 1729–1731.

- [180] Wright, S.W.; Hageman, D.L.; McClure, L.D. Fluoride-Mediated Boronic Acid Coupling Reactions. *J. Org. Chem.* **1994**, 59(20), 6095–6097.
- [181] Bui, A.T.; Roux, A.; Grichine, A.; Duperray, A.; Andraud, C.; Maury, O. Twisted Charge-Transfer Antennae for Ultra-Bright Terbium(III) and Dysprosium(III) Bioprobes. *Chem. Eur. J.* **2018**, 24(14), 3408–3412.
- [182] Dai, L.; Lo, W.-S.; Zhang, J.; Law, G.-L. One-Step Reaction for Screening of Chromophores to Improve the Luminescence of Lanthanide Complexes. *Asian Journal of Organic Chemistry* **2017**, 6(12), 1845–1850.
- [183] Regueiro-Figueroa, M.; Bensenane, B.; Ruscsák, E.; Esteban-Gómez, D.; Charbonnière, L.J.; Tircsó, G.; Tóth, I.; Blas, A. de; Rodríguez-Blas, T.; Platas-Iglesias, C. Lanthanide dota-like Complexes Containing a Picolinate Pendant: Structural Entry for the Design of Ln<sup>III</sup> - Based Luminescent Probes. *Inorg. Chem.* **2011**, 50(9), 4125–4141.
- [184] Moreau, J.; Guillon, E.; Pierrard, J.-C.; Rimbault, J.; Port, M.; Aplin-court, M. Complexing Mechanism of the Lanthanide Cations Eu<sup>3+</sup>, Gd<sup>3+</sup>, and Tb<sup>3+</sup> with 1,4,7,10-Tetrakis(carboxymethyl)-1,4,7,10-tetraazacyclododecane (dota)—Characterization of Three Successive Complexing Phases: Study of the Thermodynamic and Structural Properties of the Complexes by Potentiometry, Luminescence Spectroscopy, and EXAFS. *Chem. Eur. J.* **2004**, 10(20), 5218–5232.
- [185] Dondoni, A.; Marra, A. Metal-Catalyzed and Metal-Free Alkyne Hydrothiolation: Synthetic Aspects and Application Trends: Metal-Catalyzed and Metal-Free Alkyne Hydrothiolation. *European Journal of Organic Chemistry* **2014**, 2014(19), 3955–3969.
- [186] Sakamoto, K.; Tsuda, S.; Mochizuki, M.; Nohara, Y.; Nishio, H.; Yoshiya, T. Imidazole-Aided Native Chemical Ligation: Imidazole as a One-Pot Desulfurization-Amenable Non-Thiol-Type Alternative to 4-Mercaptophenylacetic Acid. *Chemistry – A European Journal* **2016**, 22(50), 17940–17944.
- [187] Sakamoto, K.; Tsuda, S.; Nishio, H.; Yoshiya, T. 1,2,4-Triazole-aided native chemical ligation between peptide-N-acyl-N'-methyl-benzimidazolinone and cysteinyl peptide. *Chem. Commun.* **2017**, 53(90), 12236–12239.
- [188] Terrier, V.P.; Adihou, H.; Arnould, M.; Delmas, A.F.; Aucagne, V. A straightforward method for automated Fmoc-based synthesis of bio-inspired peptide crypto-thioesters. *Chem. Sci.* **2016**, 7(1), 339–345.
- [189] Brouwer, A.M. Standards for photoluminescence quantum yield measurements in solution (IUPAC Technical Report). *Pure Appl. Chem.* **2011**, 83(12), 2213–2228.
- [190] Chessa, G.; Canovese, L.; Visentin, F.; Santo, C.; Seraglia, R. Synthesis of poly(pyridylthioether) dendrimers incorporating a Fe<sub>2</sub>(CO)<sub>6</sub> cluster core. *Tetrahedron* **2005**, 61(7), 1755–1763.
- [191] Bourdolle, A.; Allali, M.; Mulatier, J.-C.; Le Guennic, B.; Zwier, J.M.; Baldeck, P.L.; Bünzli, J.-C.G.; Andraud, C.; Lamarque, L.; Maury, O. Modulating the Photophysical

Properties of Azamacrocyclic Europium Complexes with Charge-Transfer Antenna Chromophores. *Inorg. Chem.* **2011**, 50(11), 4987–4999.

[192] Placide, V.; Bui, A.T.; Grichine, A.; Duperray, A.; Pitrat, D.; Andraud, C.; Maury, O. Two-photon multiplexing bio-imaging using a combination of Eu- and Tb-bioprobes. *Dalton Trans.* **2015**, 44(11), 4918–4924.

[193] Jagadish, B.; Ozumerzifon, T.J.; Roberts, S.A.; Hall, G.B.; Mash, E.A.; Raghunand, N. Improved Synthesis of 10-(2-Alkylamino-2-oxoethyl)-1,4,7,10-tetraazacyclododecane-1,4,7-triacetic Acid Derivatives Bearing Acid-Sensitive Linkers. *Synthetic Communications* **2014**, 44(3), 441–449.

[194] He, Y.; Liu, S.; Menon, A.; Stanford, S.; Oppong, E.; Gunawan, A.M.; Wu, L.; Wu, D.J.; Barrios, A.M.; Bottini, N.; Cato, A.C.B.; Zhang, Z.-Y. A Potent and Selective Small-Molecule Inhibitor for the Lymphoid-Specific Tyrosine Phosphatase (LYP), a Target Associated with Autoimmune Diseases. *J. Med. Chem.* **2013**, 56(12), 4990–5008.

[195] Gage, J.L.; Onrust, R.; Johnston, D.; Osnowski, A.; MacDonald, W.; Mitchell, L.; Ürögdi, L.; Rohde, A.; Harbol, K.; Gragerov, S.; Dormán, G.; Wheeler, T.; Florio, V.; Cutshall, N.S. N-Acylhydrazones as inhibitors of PDE10A. *Bioorganic & Medicinal Chemistry Letters* **2011**, 21(14), 4155–4159.

[196] Kathuria, A.; Priya, N.; Chand, K.; Singh, P.; Gupta, A.; Jalal, S.; Gupta, S.; Raj, H.G.; Sharma, S.K. Substrate specificity of acetoxo derivatives of coumarins and quinolones towards Calreticulin mediated transacetylation: Investigations on antiplatelet function. *Bioorganic & Medicinal Chemistry* **2012**, 20(4), 1624–1638.



## Partie expérimentale



## Partie expérimentale

### Materials and methods:

**Solvent and reagents:** N- $\alpha$ -Fmoc-protected amino acids for peptide synthesis, PyBOP and HCTU coupling reagents were obtained from Novabiochem or Iris Biotech. SEA-PS, NovaPEG Rink Amide resin were purchased from X'prochem and Novabiochem. DOTA-*tris*(*t*Bu) ester was purchased from CheMatech. Other reagents for peptide synthesis, solvents, buffers and metal salts were purchased from Sigma-Aldrich. All buffers or zinc solutions were prepared with MilliQ water (Millipore).

**Analytical HPLC:** analytical separations were performed on an Agilent Infinity 1200 system using Waters XBridge BEH130 C18 (2.5  $\mu$ m, 75 mm x 4.6 mm) column at 1 mL/min.

Mobile phase consisted in a gradient of solvent A (0.1% TFA in H<sub>2</sub>O) and B (0.1% TFA in MeCN/H<sub>2</sub>O 9:1). Method Xb70 consisted in 5% B during 1 min followed by a 5 to 70 % B gradient in 14 min at 1 mL/min on the Waters XBridge BEH130 C18 column. Eluate was monitored by electronic absorption at 214, 254, 280, 331, 365, 560, 650 nm.

**Analytical HPLC-MS:** analytical separations follow by mass spectroscopy were performed on an Agilent 1260 Infinity II HPLC system equipped with a 6125 ESI-MS detector.

**Preparative HPLC:** separations for purification were performed on a VWR LaPrep $\Sigma$  system using a Waters XBridge Peptide BEH130 C18 (5  $\mu$ m, 150 mm x 19 mm) column at 14 mL/min, or on a Waters XBridge Peptide BEH130 C18 (5  $\mu$ m, 150 mm x 10 mm) column at 6 mL/min. Mobile phase consisted in a gradient of solvent A (0.1% TFA in H<sub>2</sub>O) and B (0.1% TFA in MeCN/H<sub>2</sub>O 9:1).

**Optical spectroscopy:** UV-Vis absorption spectra were recorded on a Perkin-Elmer Lambda 35 spectrophotometer. Luminescence spectra were recorded on a Varian Cary Eclipse spectrometer or on a modular Fluorolog FL3-22 spectrometer from Horiba-Jobin Yvon-Spex equipped with a double-grating excitation monochromator and an iHR320 imaging spectrometer coupled to a R928P Hamamatsu photomultiplier. Quantum yields were determined by a relative method with quinine sulphate in 0.5M H<sub>2</sub>SO<sub>4</sub> as standard<sup>[189]</sup>. Estimated experimental error for the quantum yield determination is ~10 %.

**NMR:** <sup>1</sup>H, <sup>13</sup>C and DEPT spectra were recorded at 400 MHz on a Varian Avance III 400 spectrometer, at 298 K if no temperature is specify. All shifts for <sup>1</sup>H and <sup>13</sup>C spectra were

referenced to the residual solvent peak ( $\text{CDCl}_3$   $\delta_{\text{H}}$  = 7.26 ppm and  $\delta_{\text{C}}$  = 77.2 ppm;  $\text{D}_2\text{O}$   $\delta_{\text{H}}$  = 4.79 ppm at 298 K and 4.20 ppm at 353 K;  $\text{DMSO-d}_6$   $\delta_{\text{H}}$  = 2.50 ppm). The following abbreviations were used to explain the multiplicities: s = singlet, d = doublet, t = triplet, q = quartet, dd = doublet of doublet, m = multiplet. Coupling constants ( $J$ ) are measured in hertz and are given with 0.5 Hz accuracy and the chemical shift ( $\delta$ ) are measured in ppm. DEPT experiments were used to further characterize organic molecules.

**Mass spectrometry:** ESI-MS analyses were performed on a Thermo Scientific LXQ spectrometer. HRMS analyses were performed on a Thermo Scientific LTQ Orbitrap XL spectrometer.

**X-ray structure determinations and refinements:** Single-crystal X-ray data were collected at 150(2) K on a Rigaku XCalibur S CCD diffractometer equipped with a low temperature system and intensity data were collected using graphite-monochromated Mo  $K\alpha$  radiation ( $\lambda$  = 0.71073 Å). The data integration and reduction were processed with CrysAlis software (CrysAlisPro Software system, version 1.171.37.35, Agilent Technologies UK Ltd, Oxford). An analytical absorption correction was applied (Abspack, a part of the CrysAlisPro Software system, version 1.171.37.35, Agilent Technologies UK Ltd, Oxford). The structures were solved by direct methods using Superflip software<sup>74</sup> and were refined on F<sup>2</sup> by the full-matrix least-squares technique using the SHELXL-2014 program package. Non-hydrogen atoms were refined anisotropically. During the refinement, hydrogens were treated as riding atoms using SHELXL default parameters.

**Automated peptide synthesis:** Peptide synthesis was performed on a CEM Liberty 1 peptide synthesizer.

**Peptide elongation:** Peptide elongation was performed using standard SPPS protocols using Fmoc/*t*Bu chemistry either manually or on an automated peptide synthesizer (CEM Liberty1 Microwave Peptide Synthesizer). Double couplings (30 min) were performed using 4-fold molar excess of Fmoc-L-amino acid, 4-fold molar excess of HCTU and 8-fold molar excess of DIEA at room temperature. A capping step was performed after each coupling with  $\text{Ac}_2\text{O}$ /DIEA in DMF (5 min). Fmoc removal was performed using 20% piperidine in DMF (2×10 min) with Fmoc UV monitoring.

**Chromatography:** Reactions were monitored by thin layer chromatography (TLC) using commercial Pre-coated TLC sheet ALUGRAM SIL G/UV<sub>254</sub> (from Macherey-Nagel). TLC

spots were viewed under ultraviolet light (254 nm). Products were purified by silica gel (Merck silicagel 60; 0.063-0.200 nm) for flash column chromatography.

**Oligonucleotide synthesis & purification:** Oligonucleotide synthesis was performed by an automated 3400 Applied Biosystem synthesizer (reagents provided by Link Technology) using the phosphoramidite synthesis method. Removing from the solid support was performed by treatment with a 15 N solution of ammonia at room temperature. Total deprotection of the sequence was accomplished by heating at 55 ° C for 48 hours. Purification of oligonucleotide was performed using 500 µg of compound, dissolved in 100 µL of a 7 M urea solution and purified by polyacrylamide gel electrophoresis (PAGE) under denaturing conditions in TBE buffer (10.8 g / L Tris, 5.5 g / L of boric acid and 2 mM of EDTA pH 8) at a power of 55W for 10 hours. The PAGE gel was composed by a mixture of : 125 g of urea ; 125 mL of acrylamide / bis-acrylamide 40% ; 25 mL of TBE 10x ; 8.75 mL H<sub>2</sub>O ; 1,675 mL of ammonium persulfate (APS) at 5% ; 100 µL of N, N, N', N'-tetramethylethylenediamine (TEMED). After separation, slice containing the purified product was placed in 4 mL of water overnight at 4 °C to extract the oligonucleotide from acrylamide. After centrifugation, oligonucleotide was desalted on C18 column using water, followed by elution with methanol, then lyophilized.

**Radioactive labeling:** Oligonucleotide was labeled in 5' using the [ $\gamma$ -<sup>32</sup>P] ATP triphosphate (specific activity of 3000 Ci /mmol ; Perkin Elmer supplier) and the T4 polynucleotide kinase (approximately 10 units per 20 pmol of oligonucleotide, MBI supplier. Fermentas) at 37 ° C for 30 min. Activity of the T4 polynucleotide kinase was stopped by incubating the reaction medium at 70° C for 5 min.

**Hybridization:** The <sup>32</sup> P-labeled oligonucleotide was hybridized with a 10% excess of unlabeled complementary strand by incubating at 95 ° C for 3 min followed by a return to room temperature for 1h.

**Electrophoretic Mobility Shift Assay:** The 5'-labeled target duplex (10 nM) was incubated in the presence of various protein concentrations overnight at 4 °C in a final solution volume of 10 µL of HEPES pH 7,5/BSA/Tween/NaCl (15 mM, 0.1 g/L, 0.05 %, 100 mM). After addition of 5 microliters of loading buffer (50% glycerol, 50% water) 10 microliters of the mixture were loaded onto an 8% acrylamide: bisacrylamide EMSA gel (29: 1) in Tris/glycine migration buffer (25 mM Tris, 190 mM glycine) maintained at 4 °C for the duration of the migration (6h at a power not exceeding 3 watts). The gels were finally quantified using a GE Healthcare Typhoon 9410 phosphorimager.

The EMSA gel was composed by a solution of : 6 mL of 40% acrylamide / bis-acrylamide 29:1; 6 mL of 5x Tris /glycine buffer; 18 mL of water. This solution was polymerized by addition of a mixture of 300 µL of APS 10% and 30 µL of TEMED.

F1-SEA<sup>off</sup> : His-Pro-Met-Asn-Asn-Leu-Leu-Asn-Tyr-Val-Val-Pro-Lys-Met-Arg-Pro-Tyr-Ala-SEA<sup>off</sup>

F1-MPA : His-Pro-Met-Asn-Asn-Leu-Leu-Asn-Tyr-Val-Val-Pro-Lys-Met-Arg-Pro-Tyr-Ala-MPA

Ac-F1[Lys(DOTA)]-SEA<sup>off</sup>: Ac-His-Pro-Met-Asn-Lys(DOTA)-Leu-Leu-Asn-Tyr-Val-Val-Pro-Lys-Met-Arg-Pro-Tyr-Ala-SEA<sup>off</sup>

Ac-F1[Lys(DOTA[Tb])] -SEA<sup>off</sup>: Ac-His-Pro-Met-Asn-Lys(DOTA[Tb])-Leu-Leu-Asn-Tyr-Val-Val-Pro-Lys-Met-Arg-Pro-Tyr-Ala-SEA<sup>off</sup>

Ac-F1[Lys(DOTA[Tb])] -MPA: Ac-His-Pro-Met-Asn-Lys(DOTA[Tb])-Leu-Leu-Asn-Tyr-Val-Val-Pro-Lys-Met-Arg-Pro-Tyr-Ala-MPA

F3-NH<sub>2</sub>: Cys-Arg-Ile-Cys-Met-Arg-Asn-Phe-Ser-Arg-Ser-Asp-His-Leu-Thr-Thr-His-Ile-Arg-Thr-His-Thr-Gly-Glu-NH<sub>2</sub>

F2[Cys(StBu)|Glu(Cs124)]-SEA<sup>off</sup> : Cys(StBu)-Pro-Val-Glu(Cs124)-Ser-Cys(StBu)-Glu-Arg-Arg-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln-SEA<sup>off</sup>

F2[Cys(StBu)|Lys(NH<sub>2</sub>)]-SEA-2Cl-trt-PS : Cys(StBu)-Pro-Val-Lys(NH<sub>2</sub>)-Ser-Cys(StBu)-Glu-Arg-Arg-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln-SEA-2Cl-trt-PS

F2[Cys(StBu)|Lys(lissamine)]-SEA<sup>off</sup> : Cys(StBu)-Pro-Val-Lys(lissamine)-Ser-Cys(StBu)-Glu-Arg-Arg-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln-SEA<sup>off</sup>

F2[Cys(StBu)|Lys(SiR-methyl)]-SEA<sup>off</sup> : Cys(StBu)-Pro-Val-Lys(SiR-methyl)-Ser-Cys(StBu)-Glu-Arg-Arg-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln-SEA<sup>off</sup>

F2[Cys(StBu)|Lys(DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>1</sub>)]-SEA<sup>off</sup> : Cys(StBu)-Pro-Val-Lys(DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>1</sub>)-Ser-Cys(StBu)-Glu-Arg-Arg-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln-SEA<sup>off</sup>

F2[Cys(StBu)|Lys(DO3A<sub>pic</sub>Ar<sub>1</sub>)]-SEA<sup>off</sup> : Cys(StBu)-Pro-Val-Lys(DO3A<sub>pic</sub>Ar<sub>1</sub>)-Ser-Cys(StBu)-Glu-Arg-Arg-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln-SEA<sup>off</sup>

F2[Cys(StBu)|Lys(Tb(DO3ApicAr<sub>1</sub>))]-SEA<sup>off</sup> : Cys(StBu)-Pro-Val-Lys(Tb(DO3ApicAr<sub>1</sub>))-Ser-Cys(StBu)-Glu-Arg-Arg-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln-SEA<sup>off</sup>

F2[Cys|Glu(Cs124)]-HNB<sup>off</sup> : Cys-Pro-Val-Glu(Cs124)-Ser-Cys-Glu-Arg-Arg-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln-Cys(StBu)(HNB)-Gly-NH<sub>2</sub>

F2[Thz|Glu(Cs124)]-HNB<sup>off</sup> : Thz-Pro-Val-Glu(Cs124)-Ser-Cys-Glu-Arg-Arg-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln-Cys(StBu)(HNB)-Gly-NH<sub>2</sub>

F2[Cys|Lys(Eu(DO3ApicAr<sub>2</sub>))]-HNB<sup>off</sup> : Cys-Pro-Val-Lys(Eu(DO3ApicAr<sub>2</sub>))-Ser-Cys-Glu-Arg-Arg-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln-Cys(StBu)(HNB)-Gly-NH<sub>2</sub>

ZIF/Cs124|DOTA[Tb] : Ac-His-Pro-Met-Asn-Lys(DOTA[Tb])-Leu-Leu-Asn-Tyr-Val-Val-Pro-Lys-Met-Arg-Pro-Tyr-Ala-Cys-Pro-Val-Glu(Cs124)-Ser-Cys-Glu-Arg-Arg-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln- Cys-Arg-Ile-Cys-Met-Arg-Asn-Phe-Ser-Arg-Ser-Asp-His-Leu-Thr-Thr-His-Ile-Arg-Thr-His-Thr-Gly-Glu-NH<sub>2</sub>

ZIF(lissamine) : His-Pro-Met-Asn-Asn-Leu-Leu-Asn-Tyr-Val-Val-Pro-Lys-Met-Arg-Pro-Tyr-Ala- Cys-Pro-Val-Lys(lissamine)-Ser-Cys-Glu-Arg-Arg-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln-Cys-Arg-Ile-Cys-Met-Arg-Asn-Phe-Ser-Arg-Ser-Asp-His-Leu-Thr-Thr-His-Ile-Arg-Thr-His-Thr-Gly-Glu-NH<sub>2</sub>

ZIF(SiR-methyl) : His-Pro-Met-Asn-Asn-Leu-Leu-Asn-Tyr-Val-Val-Pro-Lys-Met-Arg-Pro-Tyr-Ala- Cys-Pro-Val-Lys(SiR-methyl)-Ser-Cys-Glu-Arg-Arg-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln-Cys-Arg-Ile-Cys-Met-Arg-Asn-Phe-Ser-Arg-Ser-Asp-His-Leu-Thr-Thr-His-Ile-Arg-Thr-His-Thr-Gly-Glu-NH<sub>2</sub>

ZIF(Tb(DO3ApicAr<sub>1</sub>)) : His-Pro-Met-Asn-Asn-Leu-Leu-Asn-Tyr-Val-Val-Pro-Lys-Met-Arg-Pro-Tyr-Ala- Cys-Pro-Val-Lys(Tb(DO3ApicAr<sub>1</sub>))-Ser-Cys-Glu-Arg-Arg-Phe-Ser-Arg-Ser-Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln-Cys-Arg-Ile-Cys-Met-Arg-Asn-Phe-Ser-Arg-Ser-Asp-His-Leu-Thr-Thr-His-Ile-Arg-Thr-His-Thr-Gly-Glu-NH<sub>2</sub>

ZIF(Eu(DO3ApicAr<sub>2</sub>)) : His-Pro-Met-Asn-Asn-Leu-Leu-Asn-Tyr-Val-Val-Pro-Lys-Met-Arg-Pro-Tyr-Ala- Cys-Pro-Val-Lys(Eu(DO3ApicAr<sub>2</sub>))-Ser-Cys-Glu-Arg-Arg-Phe-Ser-Arg-Ser-

Asp-Glu-Leu-Thr-Arg-His-Ile-Arg-Ile-His-Thr-Gly-Gln-Lys-Pro-Phe-Gln-Cys-Arg-Ile-Cys-  
Met-Arg-Asn-Phe-Ser-Arg-Ser-Asp-His-Leu-Thr-Thr-His-Ile-Arg-Thr-His-Thr-Gly-Glu- $NH_2$

F1-SEA<sup>off</sup>:

Peptide elongation was performed manually as described above on a SEA-PS resin (0.16 mmol/g) after introduction of the first amino acid by double manual coupling (2x30 min) using 10 eq. of Fmoc-Ala-OH, 9.5 eq. of HCTU and 20 eq. of DIEA in DMF with pre-activation followed by a capping step using Ac<sub>2</sub>O/DIEA/DCM (2:1:17 by vol., 10 mL, 10 min) and Fmoc removal using 20% piperidine in DMF (3x3min each) with UV monitoring. The resin was washed with DMF, DCM and Et<sub>2</sub>O then dried. Removal of acid-labile side chain protecting groups and cleavage of the peptidyl resin was performed with a mixture of TFA/H<sub>2</sub>O/TIS/thioanisole/octylsulfide (92.5/2.5/2.5/1.25/1.25 by vol., 7.5 mL) during 2h. TFA was evaporated, and the peptide was then precipitated in ice-cold Et<sub>2</sub>O, dissolved in deionized water, and lyophilized. The peptide was dissolved in acetic acid solution (10 %) and treated with 200 mM I<sub>2</sub> in DMSO to oxidize the C-terminal SEA<sup>on</sup> group into SEA<sup>off</sup> group. DTT was finally added to quench the excess of iodine, and the oxidized peptide was immediately purified by HPLC to give F1-SEA<sup>off</sup> (84 mg, 41 %yield, 0.075 mmol scale). HPLC (anal.) *t<sub>R</sub>* = 11.58 min (method Xb70)

ESI-MS: average *m/z* = 1138.4 (2+), 759.3 (3+), 569.7 (4+) / calculated average *m/z* = 1138.41 [M+2H]<sup>2+</sup>, 759.28 [M+3H]<sup>3+</sup>, 569.71 [M+4H]<sup>4+</sup> for M = C<sub>102</sub>H<sub>160</sub>N<sub>28</sub>O<sub>23</sub>S<sub>4</sub>; deconvoluted mass found = 2274.8 / expected mass = 2274.81 (average isotopic composition)

### F1-MPA :

A degassed solution of TCEP/MPA was prepared by dissolving TCEP (283 mg, 0.99 mmol) and MPA (625  $\mu$ L, 7.15 mmol) in a phosphate buffer (100 mM, pH 7.5; 10 mL), adjusted to pH 4 by adding 1 M aq. NaOH. Peptide F1-*SEA<sup>off</sup>* (20 mg, 7.3  $\mu$ mol) was dissolved in this solution (10 mL, peptide concentration 7.12 mM). The reaction mixture was stirred at 37°C for 24 h and then diluted with 10% aq. TFA (3 mL). The solution was extracted with Et<sub>2</sub>O and immediately purified by HPLC to give F1-MPA (8.5 mg, 43 % yield). HPLC (anal.)  $t_R$  = 10.41 min (method Xb70)

ESI-MS: average  $m/z$  = 1123.7 (2+), 749.8 (3+), 562.5 (4+) / calculated average  $m/z$  = 1123.86 [M+2H]<sup>2+</sup>, 749.57 [M+3H]<sup>3+</sup>, 562.43 [M+4H]<sup>4+</sup> for M = C<sub>101</sub>H<sub>157</sub>N<sub>27</sub>O<sub>25</sub>S<sub>3</sub>; deconvoluted mass found = 2246.3 / expected mass = 2245.70 (average isotopic composition)

Ac-F1[Lys(DOTA)]-SEA<sup>off</sup> :

Peptide elongation was performed manually as described above on a SEA-PS resin (0.16 mmol/g), after introduction of the first amino acid by double manual coupling (2x30 min) using 10 eq. of Fmoc-Ala-OH, 9.5 eq. of HCTU and 20 eq. of DIEA in DMF with pre-activation followed by a capping step using Ac<sub>2</sub>O/DIEA/DCM (2:1:17 by vol., 10 mL, 10 min) and Fmoc removal using 20% piperidine in DMF (3x3min each) with UV monitoring . Once the elongation finished, the N-terminal was capped using Ac<sub>2</sub>O/DIEA/DCM (2:1:17 by vol., 10 mL, 10 min) and then the resin was washed with DMF, DCM and Et<sub>2</sub>O then dried. Removal of the Alloc protecting group of the Lys(Alloc) was performed two times, using Pd(PPh<sub>3</sub>)<sub>4</sub> (35 μmol, 0.5 eq., 40 mg) and phenylsilane (1.46 mmol, 25 eq., 0.18 mL) in degassed anhydrous DCM (10 mL) for 1 h in the dark. The resin was then washed successively with DCM (×4), DMF (×4), 1% H<sub>2</sub>O in DMF, DMF (×4), 1% DIEA in DMF, DMF (×4), 0.12 M sodium diethyldithiocarbamate in DMF and DMF. DOTA-tris(tBu) ester (86 μmol, 50 mg, 2 eq.), PyBOP (86 μmol, 45 mg, 2 eq.) and DIEA (1.7 mmol, 30 μL, 4 eq) in solution in DMF (8 mL) was added and stirred overnight at r.t. The resin was washed with DMF, DCM and Et<sub>2</sub>O then dried. Removal of acid-labile side chain protecting groups and cleavage of the peptidyl resin was performed with a mixture of TFA/H<sub>2</sub>O/TIS/thioanisole/octylsulfide (92.5/2.5/2.5/1.25/1.25 by vol., 7.5 mL) during 4h. TFA was evaporated and the peptide was then precipitated in ice-cold Et<sub>2</sub>O, dissolved in deionized water and lyophilized. The peptide was dissolved in phosphate buffer (100 mM, pH 7.5) and diamide was added to oxidize the C-terminal SEA<sup>on</sup> group into SEA<sup>off</sup> group. The oxidized peptide was immediately purified by HPLC to give Ac-F1[Lys(DOTA)]-SEA<sup>off</sup> (42 mg, 27.5 %yield, 0.075 mmol scale). HPLC (anal.) t<sub>R</sub> = 12.04 min (method Xb70)

ESI-MS: average  $m/z$  = 1359.4 (2+), 906.7 (3+), 680.3 (4+) / calculated average  $m/z$  = 1359.66 [M+2H]<sup>2+</sup>, 906.77 [M+3H]<sup>3+</sup>, 680.34 [M+4H]<sup>4+</sup> for M = C<sub>122</sub>H<sub>194</sub>N<sub>32</sub>O<sub>30</sub>S<sub>4</sub>; deconvoluted mass found = 2717.3 / expected mass = 2717.31 (average isotopic composition)

Ac-F1[Lys(DOTA[Tb])]-SEA<sup>off</sup> :

The peptide (10 mg, 3.7 μmol) was dissolved in deionized water (1 mL) and terbium(III) chloride hexahydrate (11 mg, 29 μmol) was added. The pH was adjusted to 6.5 using 1 M aq. NaOH and the solution was degassed 30 min. The pH was controlled and adjusted to 6.5 if required. The mixture was degassed 15 min more and then stirred at r.t. overnight. The solution was diluted with water (2 mL) and purified by HPLC to give Ac-F1[Lys(DOTA[Tb])]-SEA<sup>off</sup> (6 mg, 57 %yield). HPLC (anal.)  $t_R$  = 11.24 min (method Xb70)

ESI-MS: average  $m/z$  = 1437.3 (2+), 958.6 (3+), 719.3 (4+) / calculated average  $m/z$  = 1437.61 [M+2H]<sup>2+</sup>, 958.74 [M+3H]<sup>3+</sup>, 719.31 [M+4H]<sup>4+</sup> for M = C<sub>122</sub>H<sub>191</sub>N<sub>32</sub>O<sub>30</sub>S<sub>4</sub>Tb; deconvoluted mass found = 2873.1 / expected mass = 2873.21 (average isotopic composition)

Ac-F1[Lys(DOTA[Tb])]-MPA :

A degassed solution of TCEP/MPA was prepared by dissolving TCEP (80 mg, 0.28 mmol) and MPA (175 μL, 2.03 mmol) in a phosphate buffer (100 mM, pH 7.5; 3.5 mL), adjusted to pH 4 by adding 1 M aq. NaOH. Peptide Ac-F1-LysDOTA(Tb)-SEA<sup>off</sup> (7 mg, 2.17 μmol) was dissolved in this solution (3.5 mL, peptide concentration 6.9 mM). The reaction mixture was stirred at 37°C for 24 h and then diluted with 10% aq. TFA (3 mL). The solution was extracted with Et<sub>2</sub>O and immediately purified by HPLC to give Ac-F1[Lys(DOTA[Tb])]-MPA (5 mg, 72 %yield). HPLC (anal.)  $t_R$  = 11.18 min (method Xb70)

ESI-MS: average  $m/z$  = 1422.8 (2+), 948.9 (3+), 712.1 (4+) / calculated average  $m/z$  = 1423.06 [M+2H]<sup>2+</sup>, 949.04 [M+3H]<sup>3+</sup>, 712.03 [M+4H]<sup>4+</sup> for M = C<sub>121</sub>H<sub>188</sub>N<sub>31</sub>O<sub>32</sub>S<sub>3</sub>Tb; deconvoluted mass found = 2843.4 / expected mass = 2844.11 (average isotopic composition)

F3-NH<sub>2</sub>:

Peptide elongation was performed as described above by a peptide synthesizer on a Rink Amide resin (0.45 mmol/g), after introduction of the first amino acid by double manual coupling (2x30 min) using 10 eq. of Fmoc-Glu(OtBu)-OH, 9.5 eq. of HCTU and 20 eq. of DIEA in DMF with pre-activation followed by a capping step using Ac<sub>2</sub>O/DIEA/DCM (2:1:17 by vol., 10 mL, 10 min) and Fmoc removal using 20% piperidine and OxymaPure (14g/L) in DMF (3x3min each) with UV monitoring. The resin was washed with DMF, DCM and Et<sub>2</sub>O then dried. Removal of acid-labile side chain protecting groups and cleavage of the peptidyl resin was performed with a mixture of TFA/H<sub>2</sub>O/TIS/thioanisole/octylsulfide (92.5/2.5/2.5/1.25/1.25 by vol., 7.5 mL) during 2h. TFA was evaporated and the peptide was then precipitated in ice-cold Et<sub>2</sub>O, dissolved in deionized water and lyophilized. The peptide was then purified by HPLC to give F3-NH<sub>2</sub> (104 mg, 22.7 % yield, 0.125 mmol scale). HPLC (anal.) *t<sub>R</sub>* = 7.56 min (method Xb70)

ESI-MS: average *m/z* = 958.0 (3+), 718.8 (4+), 575.3 (5+), 479.7 (6+) / calculated average *m/z* = 958.09 [M+3H]<sup>3+</sup>, 718.82 [M+4H]<sup>4+</sup>, 575.26 [M+5H]<sup>5+</sup>, 479.55 [M+6H]<sup>6+</sup> for M = C<sub>117</sub>H<sub>192</sub>N<sub>44</sub>O<sub>35</sub>S<sub>3</sub>; deconvoluted mass found = 2871.6 / expected mass = 2871.25 (average isotopic composition)

F2[Cys(StBu)]Glu(Cs124)]-SEA<sup>off</sup> :

Peptide elongation was performed as described above by a peptide synthesizer on a SEA-PS resin (0.16 mmol/g), after introduction of the first amino acid by double manual coupling (2x30 min) using 10 eq. of Fmoc-Gln-OH, 9.5 eq. of HCTU and 20 eq. of DIEA in DMF with pre-activation followed by a capping step using Ac<sub>2</sub>O/DIEA/DCM (2:1:17 by vol., 10 mL, 10 min) and Fmoc removal using 20% piperidine in DMF (3x3min each) with UV monitoring. The non standard amino acid Fmoc-Glu(Cs124)-OH was introduced manually, by using 4 eq. of amino acid, 3.8 eq. of HCTU and 8 eq. of DIEA in DMF, with pre-activation. The resin was washed with DMF, DCM and Et<sub>2</sub>O then dried.. Removal of acid-labile side chain protecting groups and cleavage of the peptidyl resin was performed with a mixture of TFA/H<sub>2</sub>O/TIS/thioanisole (93.75/2.5/2.5/1.25 by vol., 7.5 mL) during 4h. TFA was evaporated, and the peptide was then precipitated in ice-cold Et<sub>2</sub>O, dissolved in deionized water, and lyophilized. The peptide was dissolved in acetic acid solution (10 %) and treated with 200 mM I<sub>2</sub> in DMSO to oxidize the C-terminal SEA<sup>on</sup> group into SEA<sup>off</sup> group. DTT was finally added to quench the excess of iodine, and the oxidized peptide was immediately purified by HPLC to give F2[Cys(StBu)]Glu(Cs124)]-SEA<sup>off</sup> (11.3 mg, 18 %yield, 12.5 μmol scale). HPLC (anal.) *t<sub>R</sub>* = 11.40 min (method Xb70)

ESI-MS: average *m/z* = 1355.6 (3+), 1017.0 (4+), 813.9 (5+), 678.4 (6+) / calculated average *m/z* = 1355.61 [M+3H]<sup>3+</sup>, 1016.96 [M+4H]<sup>4+</sup>, 813.77 [M+5H]<sup>5+</sup>, 678.31 [M+6H]<sup>6+</sup> for M = C<sub>175</sub>H<sub>277</sub>N<sub>55</sub>O<sub>45</sub>S<sub>6</sub>; deconvoluted mass found = 4064.0 / expected mass = 4063.82 (average isotopic composition)

F2[Cys(StBu)]Lys(NH<sub>2</sub>)]-SEA-2Cl-trt-PS :

Peptide elongation was performed as described above by a peptide synthesizer on a SEA-PS resin (0.075 mmol, 0.16 mmol/g), after introduction of the first amino acid by double manual coupling (2x30 min) using 10 eq. of Fmoc-Gln-OH, 9.5 eq. of HCTU and 20 eq. of DIEA in DMF with pre-activation followed by a capping step using Ac<sub>2</sub>O/DIEA/DCM (2:1:17 by vol., 10 mL, 10 min) and Fmoc removal using 20% piperidine in DMF (3x3min each) with UV monitoring . The Fmoc-Lys(alloc)-OH was introduced manually, by using 4 eq. of amino acid, 3.8 eq. of HCTU and 8 eq. of DIEA in DMF, with pre-activation. The resin was washed with DMF, DCM and Et<sub>2</sub>O then dried. Removal of the Alloc protecting group of the Lys(Alloc) was performed two times, using Pd(PPh<sub>3</sub>)<sub>4</sub> (35 μmol, 0.5 eq., 40 mg) and phenylsilane (1.46 mmol, 25 eq., 0.18 mL) in degassed anhydrous DCM (10 mL) for 1 h in the dark. The resin was then washed successively with DCM (x4), DMF (x4), 1% H<sub>2</sub>O in DMF, DMF (x4), 1% DIEA in DMF, DMF (x4), 0.12 M sodium diethyldithiocarbamate in DMF and DMF. The resin was finally washed with DMF, DCM, Et<sub>2</sub>O and dried to be stocked for further couplings reaction on the free amine.

F2[Cys(StBu)|Lys(lissamine)]-SEA<sup>off</sup>:

A fresh solution of sulforhodamine B (75  $\mu$ mol, 43 mg, 5 eq.), and DIEA (150  $\mu$ mol, 260  $\mu$ L, 10 eq) in DMF (5 mL) was added on F2[Cys(StBu)|Lys(NH<sub>2</sub>)]-SEA-2Cl-trt-PS (15  $\mu$ mol), and stirred overnight in dark at r.t.. The resin was washed with DMF, DCM and Et<sub>2</sub>O then dried. Removal of acid-labile side chain protecting groups and cleavage of the peptidyl resin was performed with a mixture of TFA/H<sub>2</sub>O/TIS/thioanisole (93.75/2.5/2.5/1.25 by vol., 4 mL) during 2h. TFA was evaporated, and the peptide was then precipitated in ice-cold Et<sub>2</sub>O, dissolved in deionized water, and lyophilized. The peptide was dissolved in acetic acid solution (10 %) and treated with 200 mM I<sub>2</sub> in DMSO to oxidize the C-terminal SEA<sup>on</sup> group into SEA<sup>off</sup> group. DTT was finally added to quench the excess of iodine, and the oxidized peptide was immediately purified by HPLC to give F2[Cys(StBu)|Lys(lissamine)]-SEA<sup>off</sup> ( 4.5 mg, 4 %yield). HPLC (anal.)  $t_R$  = 10.41 min (method Xb70)

ESI-MS: average  $m/z$  = 1112.5 (4+), 890.4 (5+), 742.2 (6+), 636.3 (7+) / calculated average  $m/z$  = 1112.84 [M+4H]<sup>4+</sup>, 890.48 [M+5H]<sup>5+</sup>, 742.23 [M+6H]<sup>6+</sup>, 636.34 [M+7H]<sup>7+</sup> for M = C<sub>193</sub>H<sub>302</sub>N<sub>56</sub>O<sub>49</sub>S<sub>8</sub>; deconvoluted mass found = 4447.1 / expected mass = 4447.35 (average isotopic composition)

F2[Cys(StBu)]Lys(SiR-methyl)]-SEA<sup>off</sup> :

A fresh solution of SiR-methyl (18  $\mu$ mol, 9 mg, 1.2 eq.), PyBOP (18  $\mu$ mol, 10 mg, 1.2 eq.) and DIEA (90  $\mu$ mol, 16  $\mu$ L, 6 eq) in DMF (5 mL) was added on F2[Cys(StBu)]Lys(NH<sub>2</sub>)]-SEA-2Cl-trt-PS (15  $\mu$ mol), and stirred overnight in dark at r.t.. The resin was washed with DMF, DCM and Et<sub>2</sub>O then dried. Removal of acid-labile side chain protecting groups and cleavage of the peptidyl resin was performed with a mixture of TFA/H<sub>2</sub>O/TIS/thioanisole (93.75/2.5/2.5/1.25 by vol., 4 mL) during 2h. TFA was evaporated, and the peptide was then precipitated in ice-cold Et<sub>2</sub>O, dissolved in deionized water, and lyophilized. The peptide was dissolved in acetic acid solution (10 %) and treated with 200 mM I<sub>2</sub> in DMSO to oxidize the C-terminal SEA<sup>on</sup> group into SEA<sup>off</sup> group. DTT was finally added to quench the excess of iodine, and the oxidized peptide was immediately purified by HPLC to give F2[Cys(StBu)]Lys(SiR-methyl)]-SEA<sup>off</sup> ( 11 mg, 13 %yield). HPLC (anal.)  $t_R$  = 12.14 min (method Xb70)

ESI-MS: average  $m/z$  = 1083.8 (4+), 867.3 (5+), 722.9 (6+), 619.8 (7+) / calculated average  $m/z$  = 1083.83 [M+4H]<sup>4+</sup>, 867.27 [M+5H]<sup>5+</sup>, 722.89 [M+6H]<sup>6+</sup>, 619.77 [M+7H]<sup>7+</sup> for M = C<sub>193</sub>H<sub>302</sub>N<sub>56</sub>O<sub>44</sub>S<sub>6</sub>Si; deconvoluted mass found = 4331.2 / expected mass = 4331.30 (average isotopic composition)

F2[Cys(StBu)]Lys(DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub>)]-SEA<sup>off</sup>:

A fresh solution of DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub><sup>(OH)</sup> (30  $\mu$ mol, 21 mg, 2 eq.), PyBOP (30  $\mu$ mol, 16 mg, 2 eq.) and DIEA (60  $\mu$ mol, 11  $\mu$ L, 4 eq) in DMF (5 mL) was added on F2[Cys(StBu)]Lys(NH<sub>2</sub>)]-SEA-2Cl-trt-PS (15  $\mu$ mol), and stirred overnight in dark at r.t.. The resin was washed with DMF, DCM and Et<sub>2</sub>O then dried. Removal of acid-labile side chain protecting groups and cleavage of the peptidyl resin was performed with a mixture of TFA/H<sub>2</sub>O/TIS/thioanisole (93.75/2.5/2.5/1.25 by vol., 7.5 mL) during 2h. TFA was evaporated, and the peptide was then precipitated in ice-cold Et<sub>2</sub>O, dissolved in deionized water, and lyophilized. The peptide was dissolved in acetic acid solution (10 %) and treated with 200 mM I<sub>2</sub> in DMSO to oxidize the C-terminal SEA<sup>on</sup> group into SEA<sup>off</sup> group. DTT was finally added to quench the excess of iodine, and the oxidized peptide was immediately purified by HPLC to give F2[Cys(StBu)]Lys(DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub>)]-SEA<sup>off</sup> ( 13 mg, 15 %yield). HPLC (anal.)  $t_R$  = 11.26 min (method Xb70)

ESI-MS: average  $m/z$  = 1148.8 (4+), 919.3 (5+), 766.3 (6+), 657.1 (7+) / calculated average  $m/z$  = 1148.07 [M+4H]<sup>4+</sup>, 918.66 [M+5H]<sup>5+</sup>, 765.71 [M+6H]<sup>6+</sup>, 656.47 [M+7H]<sup>7+</sup> for M = C<sub>200</sub>H<sub>319</sub>N<sub>59</sub>O<sub>53</sub>S<sub>6</sub>; deconvoluted mass found = 4591.7 / expected mass = 4590.44 (average isotopic composition)

F2[Cys(StBu)]Lys(DO3ApicAr<sub>1</sub>)]-SEA<sup>off</sup> :

Peptide F2[Cys(StBu)]Lys(DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub>)]-SEA<sup>off</sup> (9.1 mg,  $\mu\text{mol}$ ) was dissolved in 0.5 M aq. NaOH (900  $\mu\text{L}$ ), sonicated during 2 min then stirred for 20 min. Then, 0.5 M aq. HCl (900  $\mu\text{L}$ ) was added and the peptide was purified by HPLC to give F2[Cys(StBu)]Lys(DO3ApicAr<sub>1</sub>)]-SEA<sup>off</sup> ( 2.3 mg, 18 %yield). HPLC (anal.)  $t_R$  = 11.01 min (method )

ESI-MS: average  $m/z$  = 1134.8 (4+), 908.2 (5+), 757.2 (6+), 649.2 (7+) / calculated average  $m/z$  = 1134.59 [M+4H]<sup>4+</sup>, 907.87 [M+5H]<sup>5+</sup>, 756.73 [M+6H]<sup>6+</sup>, 648.77 [M+7H]<sup>7+</sup> for M = C<sub>196</sub>H<sub>311</sub>N<sub>59</sub>O<sub>53</sub>S<sub>6</sub>; deconvoluted mass found = 4536.5 / expected mass = 4534.34 (average isotopic composition)

F2[Cys(StBu)]Lys(Tb(DO3ApicAr<sub>1</sub>))]-SEA<sup>off</sup> :

The peptide F2[Cys(StBu)]Lys(DO3ApicAr<sub>1</sub>)]-SEA<sup>off</sup> (2.3 mg, 0.41  $\mu\text{mol}$ ) was dissolved in deionized water (0.5 mL), and terbium(III) chloride hexahydrate (0.31 mg, 0.82  $\mu\text{mol}$ ) was added. The pH was adjusted to 9 using 1 M aq. NaOH and the solution was degassed 30 min. The pH was controlled and adjusted to 9 if required. The mixture was degassed 15 min more and then stirred at r.t. overnight. The excess of terbium was removed by using Oasis Cartridge C18, and the peptide lyophilized to give F2[Cys(StBu)]Lys(Tb(DO3ApicAr<sub>1</sub>))]-SEA<sup>off</sup> ( 2.3 mg, 97 %yield). HPLC (anal.)  $t_R$  = 11.32 min (method Xb70)

ESI-MS: average  $m/z$  = 1173.7 (4+), 939.3 (5+), 782.9 (6+), 671.3 (7+) / calculated average  $m/z$  = 1173.57 [M+4H]<sup>4+</sup>, 939.06 [M+5H]<sup>5+</sup>, 782.71 [M+6H]<sup>6+</sup>, 671.04 [M+7H]<sup>7+</sup> for M = C<sub>196</sub>H<sub>308</sub>N<sub>59</sub>O<sub>53</sub>S<sub>6</sub>Tb; deconvoluted mass found = 4691.3 / expected mass = 4690.24 (average isotopic composition)

F2[Cys|Glu(Cs124)]-HNB<sup>off</sup> :

Glycine, then Cysteine were introduced manually as described above on a Rink Amide resin (0.45 mmol/g). The resin was washed with DMF/MeOH/AcOH (9:9:2) during 5 min, then with DMF/MeOH (1:1). A solution of 2-hydroxy-5-nitrobenzyl (125 mM, 10 eq) in DMF/MeOH (1:1, 10 mL) was added to the resin and stirred for 45 min. The resin was washed with DMF/MeOH (1:1) and, without delay, a solution of sodium cyanoborohydride (250 mM, 20 eq) in DMF/MeOH/AcOH (9:9:2) (10 mL) was added. The resin was stirred for 30 min and washed successively with DMF/MeOH/AcOH (9:9:2, 10 mL, ×3), DMF (×4), 5% DIEA in DMF, DMF (×4), DCM (×4). Peptide elongation was then performed as described above by a peptide synthesizer, after introduction of the first amino acid by double manual coupling (2×30 min) using 10 eq. of Fmoc-Gln-OH, 9.5 eq. of HCTU and 20 eq. of DIEA in DMF with pre-activation followed by a capping step using Ac<sub>2</sub>O/DIEA/DCM (2:1:17 by vol., 10 mL, 10 min) and Fmoc removal using 20% piperidine in DMF (3×3min each) with UV monitoring. The non standard amino acid Fmoc-Glu(Cs124)-OH was introduced manually, by using 4 eq. of amino acid, 3.8 eq. of HCTU and 8 eq. of DIEA in DMF, with pre-activation. The resin was washed with DMF, DCM and Et<sub>2</sub>O then dried. Removal of acid-labile side chain protecting groups and cleavage of the peptidyl resin was performed with a mixture of TFA/H<sub>2</sub>O/TIS/thioanisole (93.75/2.5/2.5/1.25 by vol., 7.5 mL) during 2h. TFA was evaporated, and the peptide was then precipitated in ice-cold Et<sub>2</sub>O, dissolved in deionized water, and lyophilized. The peptide was finally purified by HPLC to give F2[Cys|Glu(Cs124)]-HNB<sup>off</sup> ( 11 mg, 11 %yield, 20 μmol scale). HPLC (anal.) *t<sub>R</sub>* = 11.12 min (method Xb70)

ESI-MS: average *m/z* = 1390.6 (3+), 1043.3 (4+), 834.8 (5+), 695.9 (6+) / calculated average *m/z* = 1390.59 [M+3H]<sup>3+</sup>, 1043.19 [M+4H]<sup>4+</sup>, 834.76 [M+5H]<sup>5+</sup>, 695.80 [M+6H]<sup>6+</sup> for M = C<sub>179</sub>H<sub>276</sub>N<sub>58</sub>O<sub>50</sub>S<sub>4</sub>; deconvoluted mass found = 4169.0 / expected mass = 4168.74 (average isotopic composition)

F2[Thz|Glu(Cs124)]-HNB<sup>off</sup>:

Peptide was synthesized using the same protocol as described for F2[Cys|Glu(Cs124)]-HNB<sup>off</sup>, where the Fmoc-Cys(Trt)-OH was replaced by Boc-Thz-OH to provide F2[Thz|Glu(Cs124)]-HNB<sup>off</sup> (9.5 mg, 9 %yield, 20  $\mu$ mol scale). HPLC (anal.)  $t_R$  = 11.52 min (method Xb70)

ESI-MS: average  $m/z$  = 1394.8 (3+), 1046.4 (4+), 837.3 (5+), 698.0 (6+) / calculated average  $m/z$  = 1394.59 [M+3H]<sup>3+</sup>, 1046.19 [M+4H]<sup>4+</sup>, 837.16 [M+5H]<sup>5+</sup>, 697.80 [M+6H]<sup>6+</sup> for M = C<sub>180</sub>H<sub>276</sub>N<sub>58</sub>O<sub>50</sub>S<sub>4</sub>; deconvoluted mass found = 4181.4 / expected mass = 4180.75 (average isotopic composition)

F2[Cys|Lys(Eu(DO3ApicAr<sub>2</sub>))]-HNB<sup>off</sup>:

Glycine, then Cysteine were introduced manually as described above on a Rink Amide resin (0.45 mmol/g). The resin was washed with DMF/MeOH/AcOH (9:9:2) during 5 min, then with DMF/MeOH (1:1). A solution of 2-hydroxy-5-nitrobenzyl (125 mM, 10 eq) in DMF/MeOH (1:1, 10 mL) was added to the resin and stirred for 45 min. The resin was washed with DMF/MeOH (1:1) and, without delay, a solution of sodium cyanoborohydride (250 mM, 20 eq) in DMF/MeOH/AcOH (9:9:2) (10 mL) was added. The resin was stirred for 30 min and washed successively with DMF/MeOH/AcOH (9:9:2, 10 mL,  $\times 3$ ), DMF ( $\times 4$ ), 5% DIEA in DMF, DMF ( $\times 4$ ), DCM ( $\times 4$ ). Peptide elongation was then performed as described above by a peptide synthesizer, after introduction of the first amino acid by double manual coupling (2 $\times$ 30 min) using 10 eq. of Fmoc-Gln-OH, 9.5 eq. of HCTU and 20 eq. of DIEA in DMF with pre-activation followed by a capping step using Ac<sub>2</sub>O/DIEA/DCM (2:1:17 by vol., 10 mL, 10 min) and Fmoc removal using 20% piperidine in DMF (3 $\times$ 3min each) with UV monitoring. The Fmoc-Lys(alloc)-OH was introduced manually, by using 4 eq. of amino acid, 3.8 eq. of HCTU and 8 eq. of DIEA in DMF, with pre-activation. Once the elongation finished, the resin was washed with DMF, DCM and Et<sub>2</sub>O then dried. Removal of the Alloc protecting group of the Lys(Alloc) was performed two times, using Pd(PPh<sub>3</sub>)<sub>4</sub> (35  $\mu$ mol, 0.5 eq., 40 mg) and phenylsilane (1.46 mmol, 25 eq., 0.18 mL) in degassed anhydrous DCM (10 mL) for 1 h in the dark. The resin was then washed successively with DCM ( $\times 4$ ), DMF ( $\times 4$ ), 1% H<sub>2</sub>O in DMF, DMF ( $\times 4$ ), 1% DIEA in DMF, DMF ( $\times 4$ ), 0.12 M sodium diethyldithiocarbamate in DMF and DMF. The resin was finally washed with DMF, DCM, Et<sub>2</sub>O and dried to be stocked for further couplings reaction on the free amine.

The resin (5  $\mu$ mol) was swollen in DCM and washed with DMF. A fresh solution of Eu[DO3ApicAr<sub>2</sub>]Na (37  $\mu$ mol, 30 mg, 7.4 eq.), DIC (37  $\mu$ mol, 5.7  $\mu$ L, 7.4 eq.) and HOAt (37  $\mu$ mol, 5 mg, 7.4 eq) in DMF (10 mL) was added on the resin, and stirred overnight in dark at r.t. The resin was washed with DMF, 20% piperidine in DMF, DMF, DCM and Et<sub>2</sub>O then dried. Removal of acid-labile side chain protecting groups and cleavage of the peptidyl resin was performed with a mixture of TFA/H<sub>2</sub>O/TIS/thioanisole (93.75/2.5/2.5/1.25 by vol., 7.5 mL) during 2h. TFA was evaporated, and the peptide (10 mg) was then precipitated in ice-cold Et<sub>2</sub>O, dissolved in deionized water (1 mL) and europium(III) chloride hexahydrate (6.3 mg, 17  $\mu$ mol) was added. The pH was adjusted to 6.5 using 1 M aq. NaOH and the solution was degassed 10 min and stirred at r.t for 2h. The solution was diluted with water (1 mL) and purified by HPLC to give F2[Cys|Lys(Eu(DO3ApicAr<sub>2</sub>))]-HNB<sup>off</sup> (1.5 mg, 5 %yield, 5  $\mu$ mol scale). HPLC (anal.)  $t_R$  = 12.55 min (method Xb70)

ESI-MS: average  $m/z$  = 1600.8 (3+), 1200.8 (4+), 961.0 (5+), 801.0 (6+) / calculated average  $m/z$  = 1600.40 [M+3H]<sup>3+</sup>, 1200.55 [M+4H]<sup>4+</sup>, 960.64 [M+5H]<sup>5+</sup>, 800.71 [M+6H]<sup>6+</sup> for M = C<sub>201</sub>H<sub>304</sub>EuN<sub>62</sub>O<sub>58</sub>S<sub>4</sub>; deconvoluted mass found = 4799.3 / expected mass = 4798.19 (average isotopic composition)

ZIF(Cs124|DOTA[Tb]):

A solution of MPAA/Gn-HCl was prepared by dissolving MPAA (28.7mg, 0.1mmol) and Gn-HCl (287mg, 3.0 mmol) in 0.5 mL of sodium phosphate buffer (0.1 M, pH 7.5). The pH of the solution was adjusted to 7.0 using 1 M aq. NaOH, then degassed. In a glove box, peptide Ac-F1[Lys(DOTA[Tb])]-MPA (0.7 mg, 0.22  $\mu$ mol, 1.0eq.) and peptide F2[Cys(StBu)|Glu(Cs124)]-SEA<sup>off</sup> (1.3 mg, 0.26  $\mu$ mol, 1.1eq.) were dissolved in the MPAA/Gn-HCl solution (38  $\mu$ L, final peptide concentration 7  $\mu$ M). The solution was stirred at 37°C in a glove box and the ligation was monitored by HPLC. After 23h the ligation was completed. A solution of TCEP was prepared by dissolving TCEP (287 mg, 1.0 mM) in 0.5 mL of sodium phosphate buffer (0.1 M, pH 7.5), adjusted to pH 7.0 using aq. NaOH then degassed. TCEP solution (3.8  $\mu$ L) and peptide F3-NH<sub>2</sub> (1.29 mg, 0.31 $\mu$ mol, 1.5eq.) were added to the reaction mixture in the glove box. The reaction mixture was stirred at 37°C and ligation was monitored by HPLC. After 21h the second ligation was completed. The reaction mixture was diluted with water (1 mL), acidified with 5% aq. TFA (1 mL) and extracted with Et<sub>2</sub>O to remove the excess of MPAA. The crude product was directly purified by HPLC to give ZIF(Cs124|DOTA[Tb]) (5.8x10<sup>-8</sup> mol, 24 %yield). HPLC (anal.)  $t_R$  = 11.24 min (method Xb70) ESI-MS: average  $m/z$  = 1338.7 (7+), 1171.5 (8+), 1041.4 (9+), 937.5 (10+) / calculated average  $m/z$  = 1338.36 [M+7H]<sup>7+</sup>, 1171.19 [M+8H]<sup>8+</sup>, 1041.17 [M+9H]<sup>9+</sup>, 937.15 [M+10H]<sup>10+</sup> for M = C<sub>398</sub>H<sub>626</sub>N<sub>129</sub>O<sub>110</sub>S<sub>7</sub>Tb; deconvoluted mass found = 9364.2 / expected mass = 9361.44 (average isotopic composition)

#### ZIF(lissamine):

A first solution of MPAA/Gn-HCl was prepared by dissolving MPAA (28.7mg, 0.1mmol) and Gn-HCl (287mg, 3.0 mmol) in 0.5 mL of sodium phosphate buffer (0.1 M, pH 7.5). The pH of the solution was adjusted to 7.0 using 1M aq. NaOH, then degassed. In a glove box, peptide F1-MPA (0.59 mg, 0.22  $\mu$ mol, 1.2eq.) and peptide F2[Cys(StBu)|Lys(lissamine)]-SEA<sup>off</sup> (1.0 mg, 0.18  $\mu$ mol, 1.0eq.) were dissolved in the MPAA/Gn-HCl solution (26  $\mu$ L, final peptide concentration 7  $\mu$ M). The solution was stirred at 37°C in the glove box and the ligation was monitored by HPLC. After 17h the ligation was completed. A second solution of TCEP/MPAA/Gn-HCl was prepared by dissolving TCEP (28.7 mg, 0.1 mM), MPAA (28.7mg, 0.1mmol), and Gn-HCl (287mg, 3.0 mmol) in 0.5 mL of sodium phosphate buffer (0.1 M, pH 7.5). The pH of the solution was adjusted to 7.0 using 1 M aq. NaOH then degassed, Peptide F3-NH<sub>2</sub> (1.04 mg, 0.27  $\mu$ mol, 1.5eq.) was dissolved in the second solution (10  $\mu$ L) and was directly added to the reaction mixture in the glove box (final peptide concentration 5  $\mu$ M). The reaction mixture was stirred at 37°C and the ligation was monitored by HPLC. After 20h the second ligation was completed. The reaction mixture was diluted with water (1 mL), acidified with 5% aq. TFA (1 mL) and extracted with Et<sub>2</sub>O to remove the excess of MPAA. The crude product was directly purified by HPLC to give ZIF(lissamine) (5.4x10<sup>-8</sup> mol, 30 %yield). HPLC (anal.)  $t_R$  = 10.40 min (method Xb70)

ESI-MS: average  $m/z$  = 1307.8 (7+), 1144.6 (8+), 1017.6 (9+), 915.9 (10+) / calculated average  $m/z$  = 1307.80 [M+7H]<sup>7+</sup>, 1144.45 [M+8H]<sup>8+</sup>, 1017.40 [M+9H]<sup>9+</sup>, 915.76 [M+10H]<sup>10+</sup> for M = C<sub>396</sub>H<sub>621</sub>N<sub>126</sub>O<sub>107</sub>S<sub>9</sub>; deconvoluted mass found = 9149.3 / expected mass = 9147.56 (average isotopic composition)

ZIF(SiR-methyl):

A first solution of MPAA/Gn-HCl was prepared by dissolving MPAA (28.7mg, 0.1mmol) and Gn-HCl (287mg, 3.0 mmol) in 0.5 mL of sodium phosphate buffer (0.1 M, pH 7.5). The pH of the solution was adjusted to 7.0 using 1 M aq. NaOH, then degassed. In a glove box, peptide F1-MPA (0.60 mg, 0.22  $\mu$ mol, 1.2eq.) and peptide F2[Cys(StBu)]Lys(SiR-methyl)]-SEA<sup>off</sup> (1.0 mg, 0.18  $\mu$ mol, 1.0eq.) were dissolved in the MPAA/Gn-HCl solution (27  $\mu$ L, final peptide concentration 7  $\mu$ M). The solution was stirred at 37°C in the glove box and the ligation was monitored by HPLC. After 17h the ligation was completed. A second solution of TCEP/MPAA/Gn-HCl was prepared by dissolving TCEP (28.7 mg, 0.1 mM), MPAA (28.7mg, 0.1mmol), and Gn-HCl (287mg, 3.0 mmol) in 0.5 mL of sodium phosphate buffer (0.1 M, pH 7.5). The pH of the solution was adjusted to 7.0 using 1 M aq. NaOH then degassed. Peptide F3-NH<sub>2</sub> (1.05 mg, 0.27  $\mu$ mol, 1.5eq.) was dissolved in the second solution (10  $\mu$ L) and was directly added to the reaction mixture in the glove box (final peptide concentration 5  $\mu$ M). The reaction mixture was stirred 37°C and the ligation was monitored by HPLC. After 17h the second ligation was completed. The reaction mixture was diluted with water (1 mL), acidified with 5% aq. TFA (1 mL) and extracted with Et<sub>2</sub>O to remove the excess of MPAA. The crude product was directly purified by HPLC to give ZIF(SiR-methyl) ( $1.49 \times 10^{-8}$  mol, 8 %yield). HPLC (anal.)  $t_R$  = 10.24 min (method Xb70)

ESI-MS: average  $m/z$  = 1291.3 (7+), 1130.1 (8+), 1004.7 (9+), 904.3 (10+) / calculated average  $m/z$  = 1291.22 [M+7H]<sup>7+</sup>, 1129.95 [M+8H]<sup>8+</sup>, 1004.51 [M+9H]<sup>9+</sup>, 904.16 [M+10H]<sup>10+</sup> for M = C<sub>396</sub>H<sub>621</sub>N<sub>126</sub>O<sub>102</sub>S<sub>7</sub>Si; deconvoluted mass found = 9033.5 / expected mass = 9031.52 (average isotopic composition)

ZIF(Tb(DO3ApicAr<sub>1</sub>)):

A solution of MPAA/Gn-HCl was prepared by dissolving MPAA (28.7mg, 0.1mmol) and Gn-HCl (287mg, 3.0 mmol) in 0.5 mL of sodium phosphate buffer (0.1 M, pH 7.5). The pH of the solution was adjusted to 7.0 using 1 M aq. NaOH, then degassed. In a glove box, peptide F1-MPA (0.57 mg, 0.21  $\mu$ mol, 1.0eq.) and peptide F2[Cys(StBu)]Lys(Tb(DO3ApicAr<sub>1</sub>))-SEA<sup>off</sup> (1.0 mg, 0.17  $\mu$ mol, 1.0eq.) were dissolved in the MPAA/Gn-HCl solution (25  $\mu$ L, final peptide concentration 7  $\mu$ M). The solution was stirred at 37°C in the glove box and the ligation was monitored by HPLC. After 20h the ligation was completed. A solution of TCEP was prepared by dissolving TCEP (287 mg, 1.0 mM) in 0.5 mL of sodium phosphate buffer (0.1 M, pH 7.5), adjusted to pH 7.0 using 1 M aq. NaOH then degassed. TCEP solution (2.5  $\mu$ L) and peptide F3-NH<sub>2</sub> (0.99 mg, 0.26  $\mu$ mol, 1.5eq.) were added to the reaction mixture in the glove box. The reaction mixture was stirred at 37°C and ligation was monitored by HPLC. After 21h the second ligation was completed. The reaction mixture was diluted with water (1 mL), acidified with 5% aq. TFA (1 mL) and extracted with Et<sub>2</sub>O to remove the excess of MPAA. The crude product was directly purified by HPLC to give ZIF(Tb(DO3ApicAr<sub>1</sub>)) (2.2x10<sup>-8</sup> mol, 12.5 %yield). HPLC (anal.) *t<sub>R</sub>* = 9.36 min (method Xb70)

ESI-MS: average *m/z* = 1342.7 (7+), 1175.0 (8+), 1044.6 (9+), 940.3 (10+) / calculated average *m/z* = 1342.36 [M+7H]<sup>7+</sup>, 1174.69 [M+8H]<sup>8+</sup>, 1044.28 [M+9H]<sup>9+</sup>, 939.95 [M+10H]<sup>10+</sup> for M = C<sub>399</sub>H<sub>626</sub>N<sub>129</sub>O<sub>111</sub>S<sub>7</sub>Tb; deconvoluted mass found = 9392.8 / expected mass = 9390.45 (average isotopic composition)

ZIF(Cs124|DOTA[Tb]) (with N to C strategy) :

A solution of Triazole/Gn-HCl was prepared by dissolving Triazole (170 mg, 2.5 mmol) and Gn-HCl (573 mg, 6.0 mmol) in 1 mL of sodium phosphate buffer (0.1 M, pH 7.5). The pH of the solution was adjusted to 7.0 using 1 M aq. NaOH then the solution was degassed. In a glove box, peptide Ac-F1[Lys(DOTA[Tb])]-MPA (0.51 mg, 0.16  $\mu$ mol, 1.0 eq.) and peptide F2[Cys|Glu(Cs124)]-HNB<sup>off</sup> (1.0 mg, 0.19  $\mu$ mol, 1.2eq.) were dissolved in the Triazole/Gn-HCl solution (23  $\mu$ L, final peptide concentration 7  $\mu$ M). The solution was stirred at 37°C in the glove box and the ligation was monitored by HPLC. After 66h the ligation was uncompleted, but no advancement were observed. A solution of TCEP was prepared by dissolving TCEP (287 mg, 1.0 mM) in 0.5 mL of sodium phosphate buffer (0.1 M, pH 7.5), adjusted to pH 7.0 using 1M aq. NaOH then degassed. TCEP solution (3.8  $\mu$ L), and peptide F3-NH<sub>2</sub> (1.29 mg, 0.31  $\mu$ mol, 1.5 eq.) in the Triazole/Gn-HCl solution (10 $\mu$ L) were added to the reaction mixture in the glove box. The reaction mixture was stirred at 37°C and ligation was monitored by HPLC. After 120h the second ligation was completed. The reaction mixture was diluted with water (1 mL), and directly purified by HPLC to give ZIF(Cs124|DOTA[Tb]). HPLC (anal.)  $t_R$  = 9.47 min (method Xb70)

ESI-MS: average  $m/z$  = 1338.5 (7+), 1171.4 (8+), 1041.4 (9+), 937.4 (10+) / calculated average  $m/z$  = 1338.36 [M+7H]<sup>7+</sup>, 1171.19 [M+8H]<sup>8+</sup>, 1041.17 [M+9H]<sup>9+</sup>, 937.15 [M+10H]<sup>10+</sup> for M = C<sub>398</sub>H<sub>626</sub>N<sub>129</sub>O<sub>110</sub>S<sub>7</sub>Tb; deconvoluted mass found = 9363.9 / expected mass = 9361.44 (average isotopic composition)

ZIF(Cs124|DOTA[Tb]) (with C to N strategy) :

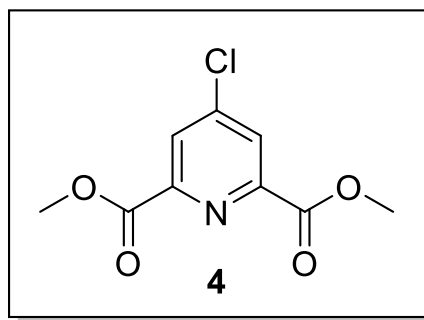
A solution of TCEP/Triazole/Gn-HCl was prepared by dissolving TCEP (29 mg, 0.1 mmol), Triazole (170 mg, 2.5 mmol) and Gn-HCl (573 mg, 6.0 mmol) in 1 mL of sodium phosphate buffer (0.1 M, pH 7.5). The pH of the solution was adjusted to 7.0 using 1 M aq. NaOH then the solution was degassed. In a glove box, peptide F3-NH<sub>2</sub> (0.9 mg, 0.23 μmol, 1.2 eq.) and peptide F2[Thz|Glu(Cs124)]-HNB<sup>off</sup> (1.0 mg, 0.19 μmol, 1.0 eq.) were dissolved in the TCEP/Triazole/Gn-HCl solution (27 μL, final peptide concentration 7 μM). The solution was stirred at 37°C in the glove box and the ligation was monitored by HPLC. After 64h the ligation was completed. A solution of methoxylamine (0.97 M, pH 6.8, 2.8 μL) was added to the mixture, and pH down to 4 by adding HCl 0.1 M. Completeness of thiazolidine hydrolysis was monitored by LC-MS (2 h), then pH was raised to 7 by adding 0.1 M aq. NaOH. Peptide Ac-F1[Lys(DOTA[Tb])]-MPA (0.92 mg, 0.29 μmol, 1.5 eq.) in the TCEP/Triazole/Gn-HCl solution (10 μL) was added to the reaction mixture in the glove box. Reaction mixture was stirred at 37°C and ligation was monitored by HPLC. After 120h the second ligation was completed. The reaction mixture was diluted with water (1 mL), and directly purified by HPLC to give ZIF(Cs124|DOTA[Tb]). HPLC (anal.) *t<sub>R</sub>* = 9.46 min (method Xb70)

ESI-MS: average *m/z* = 1338.5 (7+), 1171.4 (8+), 1041.4 (9+), 937.4 (10+) / calculated average *m/z* = 1338.36 [M+7H]<sup>7+</sup>, 1171.19 [M+8H]<sup>8+</sup>, 1041.17 [M+9H]<sup>9+</sup>, 937.15 [M+10H]<sup>10+</sup> for M = C<sub>398</sub>H<sub>626</sub>N<sub>129</sub>O<sub>110</sub>S<sub>7</sub>Tb; deconvoluted mass found = 9363.8 / expected mass = 9361.44 (average isotopic composition)

ZIF(Eu(DO3ApicAr<sub>2</sub>)) :

A solution of Triazole/Gn-HCl was prepared by dissolving Triazole (170 mg, 2.5 mmol) and Gn-HCl (573 mg, 6.0 mmol) in 1 mL of sodium phosphate buffer (0.1 M, pH 7.5). The pH of the solution was adjusted to 7.0 using 1 M aq. NaOH then the solution was degassed. In a glove box, peptide F1-MPA (0.37 mg, 0.14  $\mu$ mol, 1.0 eq.) and peptide F2[Cys|Lys(Eu(DO3ApicAr<sub>2</sub>))]-HNB<sup>off</sup> (0.8 mg, 0.14  $\mu$ mol, 1.0eq.) were dissolved in the Triazole/Gn-HCl solution (23  $\mu$ L, final peptide concentration 7  $\mu$ M). The solution was stirred at 37°C in the glove box and the ligation was monitored by HPLC. After 120h, the ligation was uncompleted, but no advancement were observed. A solution of TCEP was prepared by dissolving TCEP (287 mg, 1.0 mM) in 0.5 mL of sodium phosphate buffer (0.1 M, pH 7.5), adjusted to pH 7.0 using 1M aq. NaOH then degassed. TCEP solution (3  $\mu$ L) was added to the reaction mixture, and the ligation was monitored by HPLC. After 1h, the first ligation was completed. Peptide F3-NH<sub>2</sub> (0.67 mg, 0.18  $\mu$ mol, 1.3 eq.) was dissolved in the Triazole/Gn-HCl solution (10 $\mu$ L) and was added to the reaction mixture in the glove box. The reaction mixture was stirred at 37°C and ligation was monitored by HPLC. After 44h the second ligation seems to be completed, by the formation of two major compounds at 365nm. The reaction mixture was diluted with water (1 mL), and directly purified by HPLC to give ZIF(Eu(DO3ApicAr<sub>2</sub>)) ( 5.0x10<sup>-9</sup> mol, 3.6 %yield). HPLC (anal.)  $t_R$  = 10.16 min (method Xb70)

ESI-MS: average  $m/z$  = 1175.0 (8+), 1044.6 (9+), 940.2 (10+), 854.8 (11+), 783.6 (12+) / calculated average  $m/z$  = 1175.06 [M+8H]<sup>8+</sup>, 1044.62 [M+9H]<sup>9+</sup>, 940.26 [M+10H]<sup>10+</sup>, 854.87 [M+11H]<sup>11+</sup>, 783.71 [M+12H]<sup>12+</sup> for M = C<sub>400</sub>H<sub>625</sub>N<sub>129</sub>O<sub>111</sub>S<sub>7</sub>; deconvoluted mass found = 9392.2 / expected mass = 9392.48 (average isotopic composition)



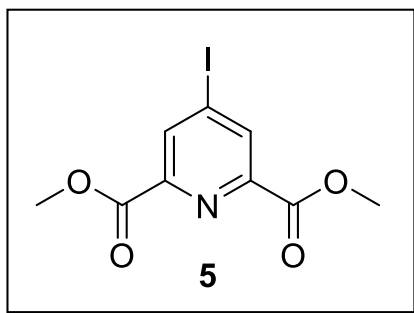
**Dimethyl 4-chloropyridine-2,6-dicarboxylate (4)<sup>[190]</sup>:**

To a suspension of chelidamic acid (11.2 g, 56  $\mu$ mol) in thionyl chloride (81 mL) was added DMF (11 mL). The mixture was heated to reflux for 16 h and solvents were removed under vacuum. Residues were dissolved in DCM (50 mL) and methanol (50 mL) was added dropwise at 0°C. After 1h, solvents were removed under vacuum. The residue was dissolved in DCM and washed with saturated aq. NaHCO<sub>3</sub> and water. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent removed under reduced pressure to afford slightly brown powder (12.4 g, 97 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 8.29 (s, 2H), 4.03 (s, 6H).

<sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 164.2 (C<sub>quat</sub>), 149.6 (C<sub>quat</sub>), 147.0 (C<sub>quat</sub>), 128.4 (CH), 53.6 (CH<sub>3</sub>).

ESI-MS: monoisotopic  $m/z$  = 252.0 [M+Na]<sup>+</sup>, 268.0 [M+K]<sup>+</sup>, 496.6 [2M+K]<sup>+</sup> (calculated monoisotopic  $m/z$  = 252.0 [M+Na]<sup>+</sup>, 267.98 [M+K]<sup>+</sup>, 496.99 [2M+K]<sup>+</sup> for M = C<sub>9</sub>H<sub>8</sub>ClNO<sub>4</sub>).

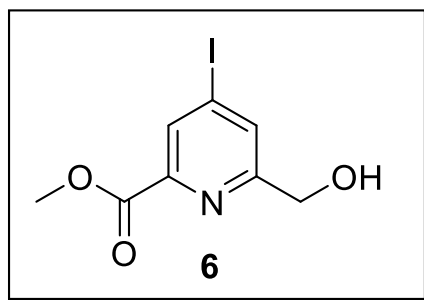
**Dimethyl 4-iodopyridine-2,6-dicarboxylate (5)<sup>[190]</sup>:**

To a solution of **4** (12 g, 52 mmol) in acetonitrile was added sodium iodide (157 g, 105 mmol) and acetyl chloride (11.2 mL, 157 mmol). The mixture was sonicated for 2 h then diluted with DCM and washed successively by saturated aq. NaHCO<sub>3</sub> and aq. sodium thiosulfate (0.1 M) until no color change happens and, finally, with brine. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent removed under reduced pressure. The crude product was dissolved in a minimal amount of DCM, and crystallized by diffusion in methanol to afford white crystals (12.9 g, 77 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 8.62 (s, 2H), 3.98 (s, 6H).

<sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 163.9 (C<sub>quat</sub>), 148.3 (C<sub>quat</sub>), 137.2 (CH), 107.0 (C<sub>quat</sub>), 53.5 (CH<sub>3</sub>).

ESI-MS: monoisotopic  $m/z$  = 344.0 [M+Na]<sup>+</sup>, 360.0 [M+K]<sup>+</sup>, 680.4 [2M+K]<sup>+</sup> (calculated monoisotopic  $m/z$  = 343.94 [M+Na]<sup>+</sup>, 359.92 [M+K]<sup>+</sup>, 680.86 [2M+H]<sup>+</sup> for M = C<sub>9</sub>H<sub>8</sub>INO<sub>4</sub>).



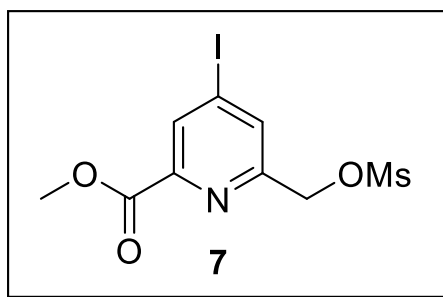
**Methyl 6-(hydroxymethyl)-4-iodopicolinate (6)<sup>[191]</sup>:**

Compound 5 (700 mg, 2.2 mmol) was dissolved in a solution of DCM/MeOH 2:3 (10 mL) cooled at 0°C. NaBH<sub>4</sub> (41 mg, 1.1 mmol) was added by portions and the mixture was stirred at 0°C for 45min. A second amount of NaBH<sub>4</sub> (41 mg, 1.1 mmol) was added and the advancement of the reaction was monitored by TLC. When the di-reduced compound was detected by TLC, the reaction was quenched by addition of 0.1 M aq. HCl (4 mL). The organic solvents were evaporated under reduced pressure and the resulted aq. phase was extracted with AcOEt. The organic phase was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent removed under reduced pressure. The crude product was purified by flash chromatography (silica gel, DCM/MeOH 95:5) to give a white powder (445 mg, 70 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 8.37 (d,  $J_4=1.5\text{Hz}$ , 1H), 7.97 (d,  $J_4=1.5\text{Hz}$ , 1H), 4.81 (s, 2H), 3.98 (s, 3H).

<sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 164.4 (C<sub>quat</sub>), 161.3 (C<sub>quat</sub>), 147.4 (C<sub>quat</sub>), 133.3 (CH), 133.1 (CH), 106.9 (C<sub>quat</sub>), 64.3 (CH<sub>2</sub>), 53.3 (CH<sub>3</sub>).

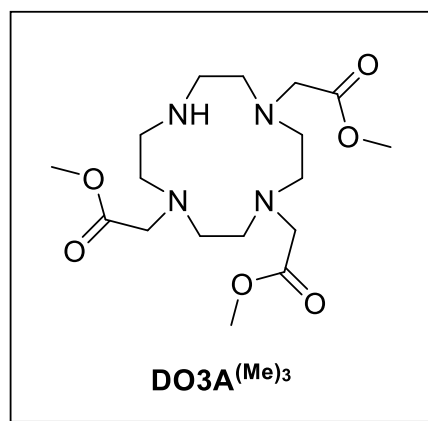
ESI-MS: monoisotopic  $m/z$  = 294.2 [M+H]<sup>+</sup>, 316.1 [M+Na]<sup>+</sup>, 332.1 [M+K]<sup>+</sup> (calculated monoisotopic  $m/z$  = 293.96 [M+H]<sup>+</sup>, 315.94 [M+H]<sup>+</sup>, 331.92 [M+K]<sup>+</sup> for M = C<sub>8</sub>H<sub>8</sub>INO<sub>3</sub>).

**Methyl 4-iodo-6-(((methylsulfonyl)oxy)methyl)picolinate (7)**<sup>[192]</sup>:

To a solution of **6** (1.35 g, 4.37 mmol) in THF (44 mL) was added triethylamine (2.1 mL, 13.1 mmol) and mesylchloride (570  $\mu$ L, 7.0 mmol). The mixture was stirred for 25 min, and solvents were evaporated. The residue was dissolved in DCM and washed by saturated aq.  $\text{NaHCO}_3$  and water. The organic phase was dried over  $\text{Na}_2\text{SO}_4$  and the solvent was removed under reduced pressure to give a white solid (1.59 g, 98 %).

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 8.38 (d,  $J_4=1.5\text{Hz}$ , 1H), 7.98 (d,  $J_4=1.5\text{Hz}$ , 1H), 5.30 (s, 2H), 3.93 (s, 3H), 3.12 (s, 3H).

$^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 163.8 ( $\text{C}_{\text{quat}}$ ), 155.0 ( $\text{C}_{\text{quat}}$ ), 147.7 ( $\text{C}_{\text{quat}}$ ), 134.0 (CH), 107.1 ( $\text{C}_{\text{quat}}$ ), 69.8 ( $\text{CH}_2$ ), 53.2 ( $\text{CH}_3$ ), 38.0 ( $\text{CH}_3$ ).



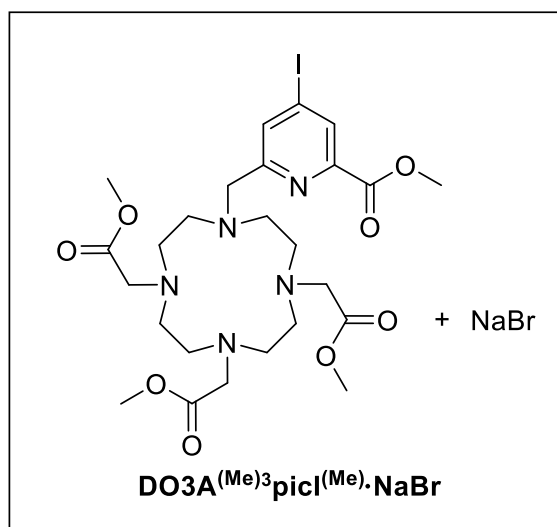
**4,7,10-*tris*(2-methoxy-2-oxoethyl)-1,4,7,10-tetraazacyclododecane hydrobromide (DO3A<sup>(Me)</sup><sub>3</sub>)<sup>[193]</sup>:**

To a suspension of sodium acetate (1.58 g, 19.3 mmol) and cyclen (1.0 g, 5.8 mmol), in DMF (8 mL) at -20°C, was added a solution of methyl bromoacetate (1.7 mL, 17.4 mmol) in DMF (4 mL). The reaction was stirred for 6 h at room temperature. The solution was diluted with DCM (100 mL) and the precipitate was removed by filtration over Celite. Solvents were evaporated under vacuum. The residue was precipitated in AcOEt and centrifuged. The supernatant was evaporated and purified by flash chromatography (silica gel, DCM/MeOH 95:5) to give an orange oil once sonicated on Et<sub>2</sub>O give a white powder (2.31 g, 85 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 9.98 (s, 1H), 3.70 (s, 9H), 3.49 (s, 4H), 3.43 (s, 2H), 3.15-2.81 (m, 16H).

<sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 171.6 (C<sub>quat</sub>), 170.8 (C<sub>quat</sub>), 57.1 (CH<sub>2</sub>), 51.7 (CH<sub>3</sub>), 51.5 (CH<sub>2</sub>), 49.4 (CH<sub>2</sub>), 48.5 (CH<sub>2</sub>) 47.3 (CH<sub>2</sub>).

ESI-MS: monoisotopic  $m/z$  = 389.4 [M+H]<sup>+</sup>, 411.4 [M+Na]<sup>+</sup> (calculated monoisotopic  $m/z$  = 389.24 [M+H]<sup>+</sup>, 411.22 [M+Na]<sup>+</sup> for M = C<sub>17</sub>H<sub>32</sub>N<sub>4</sub>O<sub>6</sub>).

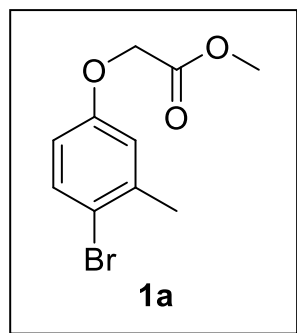
**DO3A<sup>(Me)</sup>3picl<sup>(Me)</sup>.NaBr:**

To a solution of **DO3A<sup>(Me)</sup>3** (2.62 g, 5.6 mmol) in acetonitrile (55 mL), was added Na<sub>2</sub>CO<sub>3</sub> (2.28 g, 21.5 mmol) and then a solution of **7** (1.59 g, 4.3 mmol) in acetonitrile (30 mL). The mixture was heated at reflux for 22 h. The mixture was cooled to room temperature and filtrated. The filtrate was evaporated and the residue was triturated in AcOEt to give a white powder (3.29 g, 99 %). Crystals suitable for X-ray diffraction were obtained by slow evaporation of a MeOH solution. The X-ray structure shows the macrocycle complexed with Na<sup>+</sup> with Br<sup>-</sup> as a counter ion.

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 8.33 (d,  $J_4=1.4$ Hz, 1H), 7.91 (d,  $J_4=1.4$ Hz, 1H), 3.96 (s, 3H), 3.76 (s, 3H), 3.62, (s, 6H), 3.37-2.26 (m, 24H).

<sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 173.8 (C<sub>quat</sub>), 172.7 (C<sub>quat</sub>), 164.4 (C<sub>quat</sub>), 160.0 (C<sub>quat</sub>), 147.8 (C<sub>quat</sub>), 136.7 (CH), 133.0 (CH), 107.9 (C<sub>quat</sub>), 58.9 (CH<sub>2</sub>), 55.4 (CH<sub>2</sub>), 55.2 (CH<sub>2</sub>), 53.2 (CH<sub>3</sub>), 52.2 (CH<sub>3</sub>), 52.1 (CH<sub>3</sub>), 50.6 (CH<sub>2</sub>).

**ESI-MS:** monoisotopic  $m/z$  = 686.3 [M+Na]<sup>+</sup>, 702.3 [M+K]<sup>+</sup> (calculated monoisotopic  $m/z$  = 686.17 [M+Na]<sup>+</sup>, 702.14 [M+K]<sup>+</sup> for M = C<sub>25</sub>H<sub>38</sub>IN<sub>5</sub>O<sub>8</sub>).

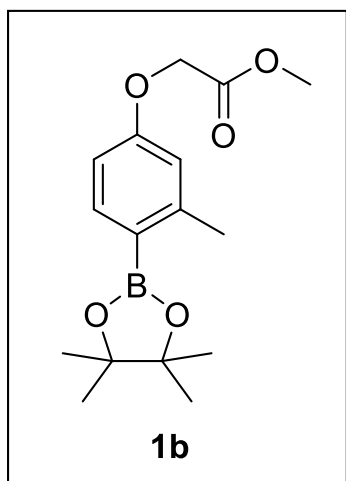


**Methyl 2-(4-bromo-3-methylphenoxy)acetate (1a):**

To a solution of 4-bromo-3-methylphenol (1.0 g, 5.4 mmol) in DMF (10 mL) was added potassium carbonate (2.2 g, 16 mmol) and methyl bromoacetate (610  $\mu$ L, 6.4 mmol). The mixture was stirred at room temperature for 17 h and then filtrated. The solvent was removed under reduced pressure. The oily residue was dissolved in DCM and washed with water. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and evaporated under reduced pressure to give a colorless oil (1.28 g, 92 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 7.39 (d,  $J_3$ =8.7Hz, 1H), 6.79 (d,  $J_4$ =3.0Hz, 2H), 6.58 (dd,  $J_3$ =8.8Hz,  $J_4$ =3.1Hz, 1H), 4.57 (s, 2H), 3.77 (s, 3H), 2.33 (s, 3H).

<sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 169.1 (C<sub>quat</sub>), 157.0 (C<sub>quat</sub>), 139.1 (C<sub>quat</sub>), 132.9 (CH), 117.4 (CH), 116.6 (C<sub>quat</sub>), 113.5 (CH), 65.3 (CH<sub>2</sub>), 52.3 (CH<sub>3</sub>), 23.1 (CH<sub>3</sub>).

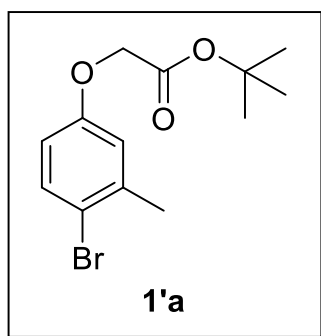
**Methyl 2-(3-methyl-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)acetate (1b):**

To a degassed solution of **1a** (800 mg, 3.1 mmol) and *bis*(triphenylphosphine)palladium(II) dichloride (217 mg, 0.31 mmol) in dioxane (10 mL) was added under an inert atmosphere 4,4,5,5-tetramethyl-1,3,2-dioxaborane (1.4 mL, 9.3 mmol) and triethylamine (2.6 mL, 18.5 mmol). The mixture was heated at 95°C for 17 h. Solvents were removed under reduced pressure. The residue was dissolved in AcOEt and washed with saturated aq. NaHCO<sub>3</sub> and water. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was removed under reduced pressure to give a dark oil. The crude product was purified by flash chromatography (silica gel, DCM) to give a yellow solid (773 mg, 82 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 7.71 (d,  $J_3=8.3$ Hz, 1H), 6.72 (d,  $J_4=2.5$ Hz, 1H), 6.68 (dd,  $J_3=8.2$ Hz,  $J_4=2.6$ Hz, 1H), 4.64 (s, 2H), 3.79 (3H), 2.51 (s, 3H), 1.32 (s, 12H).

<sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>):  $\delta$  (ppm): 169.4(C<sub>quat</sub>), 159.9 (C<sub>quat</sub>), 147.5 (C<sub>quat</sub>), 138.0 (CH), 116.3 (CH), 110.6 (CH), 83.4 (C<sub>quat</sub>), 65.1 (CH<sub>2</sub>), 52.4 (CH<sub>3</sub>), 25.0 (CH<sub>3</sub>), 22.5 (CH<sub>3</sub>).

ESI-MS: monoisotopic  $m/z$  = 329.3 [M+Na]<sup>+</sup>, 345.2 [M+K]<sup>+</sup>, 650.8 [2M+K]<sup>+</sup> (calculated monoisotopic  $m/z$  = 329.15 [M+Na]<sup>+</sup>, 345.13 [M+K]<sup>+</sup>, 651.29 [2M+H]<sup>+</sup> for M = C<sub>16</sub>H<sub>23</sub>BO<sub>5</sub>).



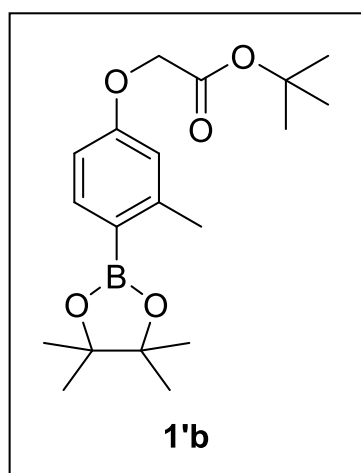
***tert*-butyl 2-(4-bromo-3-methylphenoxy)acetate (1'a):**

To a solution of 4-bromo-3-methylphenol (2.0 g, 10.7 mmol) in DMF (20 mL) was added potassium carbonate (4.44 g, 32.1 mmol) and *tert*-butyl bromoacetate (1.9 mL, 12.8 mmol). The mixture was stirred at room temperature for 17 h and then filtrated. Solvents were removed under reduced pressure. The oily residue was dissolved in DCM and washed with water. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and evaporated under reduced pressure to give a colorless oil (3.12 g, 97 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 7.38 (d,  $J_3=8.7$  Hz, 1H), 6.79 (d,  $J_4=2.9$  Hz, 1H), 6.59 (dd,  $J_4=8.7$  Hz,  $J_3=2.9$ Hz, 1H), 4.47 (s, 2H), 2.34 (s, 3H), 1.48 (s, 9H).

<sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>):  $\delta$  (ppm): 167.8(C<sub>quat</sub>), 157.2 (C<sub>quat</sub>), 139.0 (C<sub>quat</sub>), 132.9 (CH), 117.4 (CH), 116.4 (C<sub>quat</sub>), 113.5 (CH), 82.5 (C<sub>quat</sub>), 65.8 (CH<sub>2</sub>), 28.1 (CH<sub>3</sub>), 23.2 (CH<sub>3</sub>).

ESI-MS: monoisotopic  $m/z$  = 323.0 [M+Na]<sup>+</sup>, 339.0 [M+K]<sup>+</sup> (calculated monoisotopic  $m/z$  = 323.03 [M+Na]<sup>+</sup>, 339.00 [M+K]<sup>+</sup> for M = C<sub>13</sub>H<sub>17</sub>BrO<sub>3</sub>).



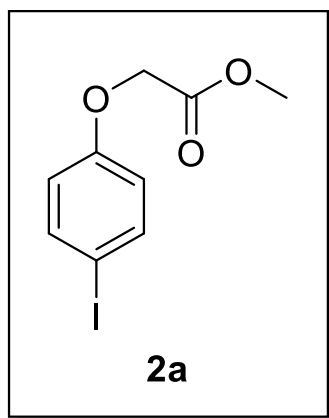
**tert-butyl 2-(3-methyl-4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenoxy)acetate (1'b):**

To a degassed solution of **1'a** (750 mg, 2.5 mmol) and *bis*(triphenylphosphine)palladium(II) dichloride (175 mg, 0.2 mmol) in dioxane (10 mL) was added under an inert atmosphere 4,4,5,5-tetramethyl-1,3,2-dioxaborane (1.1 mL, 7.5 mmol) and triethylamine (2.1 mL, 15 mmol). The mixture was heated at 95°C for 17 h. Solvents were removed under reduced pressure and the residue was dissolved in AcOEt and washed with saturated aq. NaHCO<sub>3</sub> and brine. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was removed under reduced pressure to give a dark oil. The crude product was purified by flash chromatography (silica gel, DCM) to give a yellow oil (720 mg, 83 %) that crystalizes after a few days.

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 7.71 (d,  $J_3=8.2$  Hz, 1H), 6.71 (d,  $J_4=2.5$  Hz, 1H), 6.67 (dd,  $J_3=8.0$  Hz,  $J_4=2.3$  Hz, 1H), 4.51 (s, 2H), 2.51 (s, 3H), 1.48 (s, 9H), 1.32 (s, 12H).

<sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>):  $\delta$  (ppm): 168.1 (C<sub>quat</sub>), 160.1 (C<sub>quat</sub>), 147.4 (C<sub>quat</sub>), 137.9 (CH), 116.3 (CH), 110.7 (CH), 83.3 (C<sub>quat</sub>), 82.4 (C<sub>quat</sub>), 65.5 (CH<sub>2</sub>), 28.2 (CH<sub>3</sub>), 25.0 (CH<sub>3</sub>), 22.5 (CH<sub>3</sub>).

ESI-MS: monoisotopic  $m/z$  = 371.2 [M+Na]<sup>+</sup>, 387.2 [M+K]<sup>+</sup> (calculated monoisotopic  $m/z$  = 371.20 [M+Na]<sup>+</sup>, 387.17 [M+K]<sup>+</sup> for M = C<sub>19</sub>H<sub>29</sub>BO<sub>5</sub>).



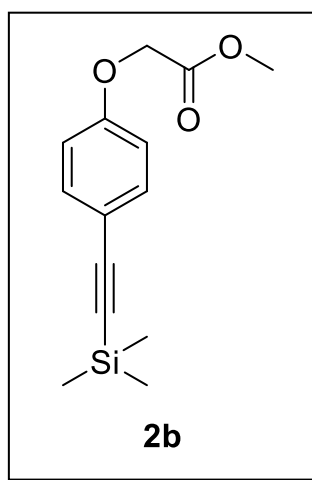
**Methyl 2-(4-iodophenoxy)acetate (2a) (adapted from ref. <sup>[194]</sup>):**

To a solution of 4-iodophenol (1.0 g, 4.55 mmol) in DMF (10 mL) was added potassium carbonate (1.88 g, 13.6 mmol) and methyl bromoacetate (515  $\mu$ L, 5.4 mmol). The mixture was stirred at room temperature for 16 h and then filtrated. The solvent was removed under reduced pressure. The oily residue was dissolved in DCM and washed with water. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was removed under reduced pressure to give an oil (1.34 g, 100 %) that crystalizes after a few days.

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 7.56 (d,  $J_3$ =8.5 Hz, 2H), 6.68 (d,  $J_3$ =8.6 Hz, 2H), 4.60 (s, 2H), 3.80 (s, 3H).

<sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 169.1 (C<sub>quat</sub>), 157.8 (C<sub>quat</sub>), 138.5 (CH), 117.1 (CH), 84.3 (C<sub>quat</sub>), 65.4 (CH<sub>2</sub>), 52.5 (CH<sub>3</sub>).

ESI-MS: monoisotopic  $m/z$  = 315.0 [M+Na]<sup>+</sup>, 330.9 [M+K]<sup>+</sup> (calculated monoisotopic  $m/z$  = 314.95 [M+Na]<sup>+</sup>, 330.92 [M+K]<sup>+</sup> for M = C<sub>9</sub>H<sub>9</sub>IO<sub>3</sub>).



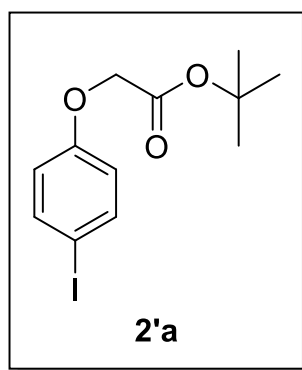
**Methyl 2-(4-((trimethylsilyl)ethynyl)phenoxy)acetate (2b) (adapted from ref. <sup>[194]</sup>):**

To a degassed solution of **2a** (800 mg, 2.74 mmol) and triethylamine (1.2 mL, 8.6 mmol) in DMF (1.2 mL) was added under an inert atmosphere *bis*(triphenylphosphine)palladium(II) dichloride (77 mg, 0.11 mmol), copper iodide (105 mg, 0.55 mmol) and ethynyltrimethylsilane (1.35 mL, 9.6 mmol). The mixture was stirred at room temperature for 15 h. Solvents were removed under reduced pressure. The residue was dissolved in DCM and washed with saturated aq. NH<sub>4</sub>Cl and water. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was removed under reduced pressure to give a dark oil. The crude product was purified by flash chromatography (silica gel, DCM) to give brown crystals (683 mg, 95 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 7.40 (d,  $J_3$  = 8.5 Hz, 2H), 6.82 (d,  $J_3$  = 8.6 Hz, 2H), 4.63 (s, 2H), 3.80 (s, 3H), 0.23 (s, 9H).

<sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>):  $\delta$  (ppm): 169.1(C<sub>quat</sub>), 158.0 (C<sub>quat</sub>), 133.7 (CH), 116.7 (C<sub>quat</sub>), 114.6 (CH), 104.9 (C<sub>quat</sub>), 93.1 (C<sub>quat</sub>), 65.4 (CH<sub>2</sub>), 52.5 (CH<sub>3</sub>), 0.16 (CH<sub>3</sub>).

ESI-MS: monoisotopic  $m/z$  = 285.2 [M+Na]<sup>+</sup>, 301.1 [M+K]<sup>+</sup> (calculated monoisotopic  $m/z$  = 285.09 [M+Na]<sup>+</sup>, 301.07 [M+K]<sup>+</sup> for M = C<sub>14</sub>H<sub>18</sub>O<sub>3</sub>Si).



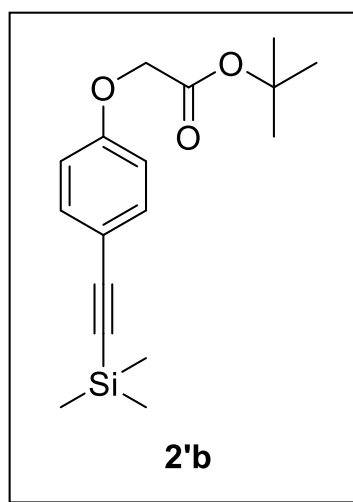
***tert*-butyl 2-(4-iodophenoxy)acetate (2'a):**

To a solution of 4-iodophenol (3.0 g, 13.6 mmol) in DMF (30 mL) was added potassium carbonate (10 g, 106 mmol), and *tert*-butyl bromoacetate (6 mL, 40.5 mmol). The mixture was stirred at room temperature for 16 h and then filtrated. The solvent was removed under reduced pressure. The oily residue was dissolved in DCM and washed with water then brine. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent removed under reduced pressure. The residue was purified by flash chromatography (silica gel, DCM) to give a pale red liquid (3.43 g, 75 %).

<sup>1</sup>H NMR (400MHz, CDCl<sub>3</sub>): δ (ppm) = 7.51 (d, J<sub>3</sub>=9 Hz, 2H), 6.63 (d, J<sub>3</sub>=9 Hz, 2H), 4.44 (s, 2H), 1.45 (s, 9H).

<sup>13</sup>C NMR (100MHz, CDCl<sub>3</sub>): δ (ppm) = 167.6 (C<sub>quat</sub>), 157.8 (C<sub>quat</sub>), 138.3 (CH), 117.0 (CH), 83.8 (C<sub>quat</sub>), 82.6 (C<sub>quat</sub>), 65.7 (CH<sub>2</sub>), 28.0 (CH<sub>3</sub>).

ESI-MS: monoisotopic *m/z* = 357.0 [M+Na]<sup>+</sup>, 373.0 [M+K]<sup>+</sup> (calculated monoisotopic *m/z* = 357.00 [M+Na]<sup>+</sup>, 372.97 [M+K]<sup>+</sup> for M = C<sub>12</sub>H<sub>15</sub>IO<sub>3</sub>).

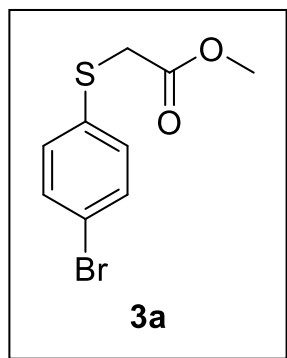


***tert*-butyl 2-(4-((trimethylsilyl)ethynyl)phenoxy)acetate (2'b)**

To a degassed solution of **2'a** (1.0 g, 3.0 mmol) and triethylamine (1.3 mL, 9.3 mmol) in DMF (1.3 mL) was added under an inert atmosphere *bis*(triphenylphosphine)palladium(II) dichloride (84 mg, 0.12 mmol), copper iodide (114 mg, 0.6 mmol) and ethynyltrimethylsilane (1.5 mL, 10.5 mmol). The mixture was stirred at room temperature for 15 h. Solvents were removed under reduced pressure. The residue was dissolved in DCM and washed with saturated aq.  $\text{NH}_4\text{Cl}$  and water. The organic layer was dried over  $\text{Na}_2\text{SO}_4$  and the solvent was removed under reduced pressure to give a dark oil. The crude product was purified by flash chromatography (silica gel, Cyclohexane/AcOEt 95:5) to give brown crystals (900 mg, 99 %).

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 7.39 (d,  $J_3 = 9.0$  Hz, 2H), 6.80 (d,  $J_3 = 9.0$  Hz, 2H), 4.50 (s, 2H), 1.47 (s, 9H), 0.23 (s, 9H).

$^{13}\text{C}$  NMR (100MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm): 167.8( $\text{C}_{\text{quat}}$ ), 158.2 ( $\text{C}_{\text{quat}}$ ), 133.6 (CH), 116.3 ( $\text{C}_{\text{quat}}$ ), 114.6 (CH), 105.0 ( $\text{C}_{\text{quat}}$ ), 92.9 ( $\text{C}_{\text{quat}}$ ), 82.7 ( $\text{C}_{\text{quat}}$ ), 65.8 ( $\text{CH}_2$ ), 28.1 ( $\text{CH}_3$ ), 0.16 ( $\text{CH}_3$ ).

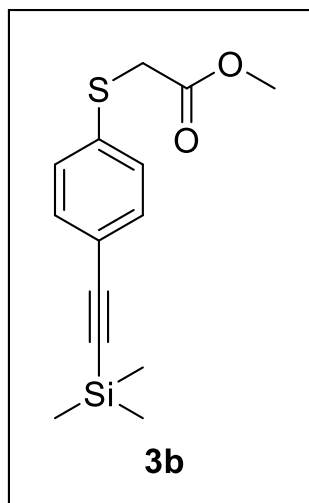


**Methyl 2-((4-bromophenyl)thio)acetate (3a) (adapted from ref. <sup>[195]</sup>):**

To a degassed solution of 4-bromothiophenol (1.0 g, 5.3 mmol) in DMF (10 mL) was added potassium carbonate (2.2 g, 16.0 mmol) and methyl bromoacetate (0.6 mL, 6.3 mmol). The mixture was stirred at room temperature for 17 h and filtrated. The solvent was removed under reduced pressure. The oily residue was dissolved in DCM and washed with water. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent removed under reduced pressure. The residue was purified by flash chromatography (silica gel, cyclohexane/AcOEt 90:10) to give a yellow oil (852 mg, 62 %).

<sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 7.35 (d,  $J_3=8.7$  Hz, 2H), 7.27 (d,  $J_3=8.7$  Hz, 2H), 3.64 (s, 3 H), 3.57 (s, 2 H).

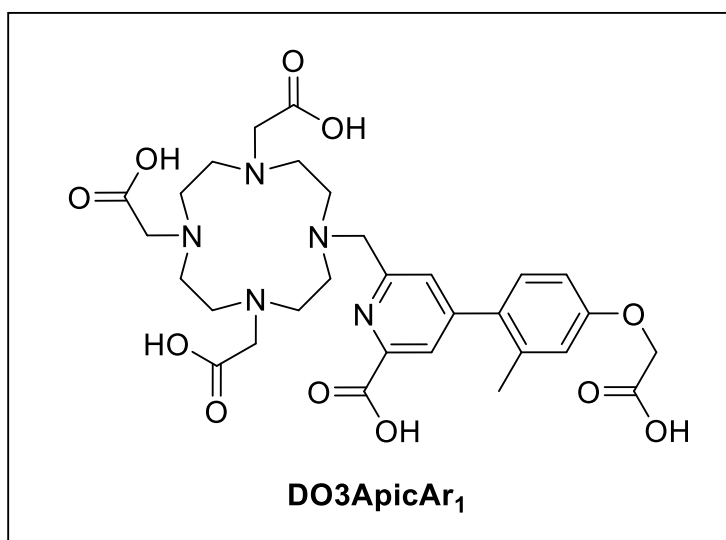
<sup>13</sup>C NMR (100 MHz, CDCl<sub>3</sub>):  $\delta$  (ppm) = 169.7 (C<sub>quat</sub>), 134.2 (C<sub>quat</sub>), 132.1 (CH), 131.3 (CH), 121.0 (C<sub>quat</sub>), 52.6 (CH<sub>3</sub>), 36.3 (CH<sub>2</sub>).

**Methyl 2-((4-((trimethylsilyl)ethynyl)phenyl)thio)acetate (3b):**

To a degassed solution of **3a** (700 mg, 2.68 mmol), copper iodide (21 mg, 0.11 mmol) and bis(benzonitrile)palladium(II) chloride (61 mg, 0.16 mmol) in dioxane (25 mL) was added a solution of tri-tert-butylphosphine (71 mg, 0.35 mmol) in dioxane (1.4 mL), ethynyltrimethylsilane (455  $\mu$ L, 3.22 mmol) and di-isopropylamine (450  $\mu$ L, 3.22 mmol). The mixture was stirred at room temperature for 16 h. Solvents were removed under reduced pressure. The residue was dissolved in DCM and washed with saturated aq.  $\text{NH}_4\text{Cl}$  and water. The organic layer was dried over  $\text{Na}_2\text{SO}_4$  and the solvent was removed under reduced pressure to give a dark oil. The crude product was purified by flash chromatography (silica gel, DCM/Petroleum ether 50:50) to give brown crystals (689 mg, 92 %).

$^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 7.38 (d,  $J_3=8.6$  Hz, 2H), 7.29 (d,  $J_3=8.6$  Hz, 2H), 3.71 (s, 3 H), 3.66 (s, 2 H), 0.24 (s, 9H).

$^{13}\text{C}$  NMR (100 MHz,  $\text{CDCl}_3$ ):  $\delta$  (ppm) = 170.0 ( $\text{C}_{\text{quat}}$ ), 136.1 ( $\text{C}_{\text{quat}}$ ), 132.6 (CH), 128.9 (CH), 121, 6 ( $\text{C}_{\text{quat}}$ ), 104.6 ( $\text{C}_{\text{quat}}$ ), 95.2 ( $\text{C}_{\text{quat}}$ ), 52.8 ( $\text{CH}_3$ ), 36.0 ( $\text{CH}_2$ ), 0.1 ( $\text{CH}_3$ ).



### DO3ApicAr<sub>1</sub> :

To a degassed solution of **DO3A**<sup>(Me)<sub>3</sub>pic<sup>(Me)</sup></sup> (100 mg, 0.131 mmol) and **1b** (37.5 mg, 0.122 mmol) in DMF (1.25 mL) was added *tetrakis*(triphenylphosphine)palladium(0) (14 mg, 12.1 mmol) and cesium fluoride (56 mg, 0.37 mmol). The mixture was stirred at 90°C for 17 h. Solvents were removed under reduced pressure. The oily residue was dissolved in DCM and washed with water. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and evaporated under reduced pressure. Water (16mL) was added to the crude and the suspension was sonicated for 20 min and centrifuged. The supernatant was lyophilized to give DO3A<sup>(Me)<sub>3</sub>pic<sup>(Me)</sup></sup>Ar<sub>1</sub><sup>(Me)</sup> as a yellow oil (102 mg). This product was then used without further purification, but a small amount of ligand was purified by HPLC for characterization. The methyl-protected ligand was dissolved in 1M aq. NaOH (4 mL), and stirred at rt. The reaction was completed in 15 min as monitored by HPLC. The pH was adjusted to 4 by addition of 1 M aq. HCl (4 mL). The solution was purified on HPLC to give DO3ApicAr<sub>1</sub> as a yellow solid (72 mg, 92 %) after lyophilization.

### DO3A<sup>(Me)<sub>3</sub>pic<sup>(Me)</sup></sup>Ar<sub>1</sub><sup>(Me)</sup> :

<sup>1</sup>H NMR (400MHz, D<sub>2</sub>O, 353K): δ (ppm) = 8.12 (d, J<sub>4</sub>=1.4Hz, 1H), 7.85 (d, J<sub>4</sub>=1.4Hz, 1H), 7.25 (d, J<sub>3</sub>=8.5Hz, 1H), 6.97 (d, J<sub>4</sub>=2.3Hz, 1H), 6.93 (dd, J<sub>3</sub>=8.5Hz, J<sub>4</sub>=2.6Hz, 1H), 4.81 (s, 2H), 4.60 (s, 2H), 4.09 (s, 2H), 4.03 (s, 3H), 3.81 (s, 3H), 3.72 (s, 3H), 3.60 (s, 4H), 3.50 (s, 6H), 3.79-3.04 (m, 16H), 2.26 (s, 3H).|

ESI-MS: monoisotopic *m/z* = 716.4 [M+H]<sup>+</sup>, 738.4 [M+Na]<sup>+</sup>, 754.3 [M+K]<sup>+</sup> (calculated monoisotopic *m/z* = 716.35 [M+H]<sup>+</sup>, 738.33 [M+Na]<sup>+</sup>, 754.31 [M+K]<sup>+</sup> for M = C<sub>35</sub>H<sub>49</sub>N<sub>5</sub>O<sub>11</sub>).

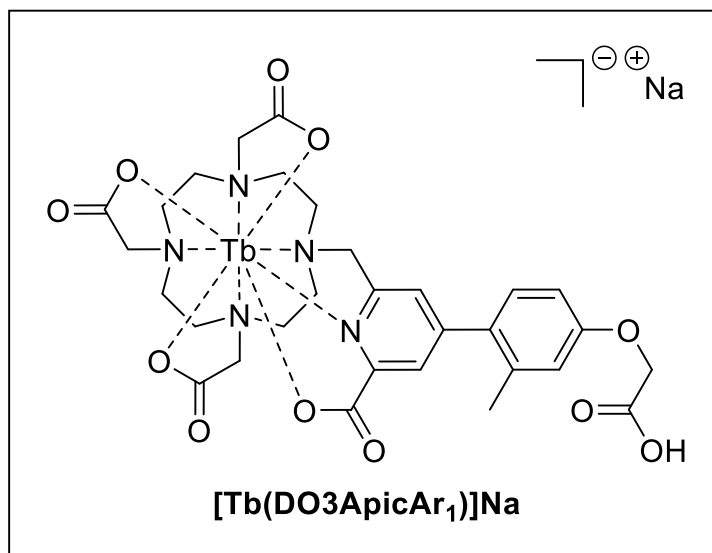
### DO3ApicAr<sub>1</sub> :

$^1\text{H}$  NMR (400MHz,  $\text{D}_2\text{O}$ , 353K):  $\delta$  (ppm) = 8.11 (d,  $J_4=1.5\text{Hz}$ , 1H), 7.82 (d,  $J_4=1.5\text{Hz}$ , 1H), 7.28 (d,  $J_3=8.4\text{Hz}$ , 1H), 6.95 (d,  $J_4=2.7\text{Hz}$ , 1H), 6.92 (dd,  $J_3=8.5\text{Hz}$ ,  $J_4=2.8\text{Hz}$ , 1H), 4.76 (s, 2H), 4.49 (s, 2H), 3.84 (s, 2H), 3.62 (s, 4H), 3.43-3.18 (m, 16H), 2.24 (s, 3H).

$^{13}\text{C}$  NMR (100 MHz,  $\text{D}_2\text{O}$ , 353K):  $\delta$  (ppm) = 173.3 ( $\text{C}_{\text{quat}}$ ), 172.1 ( $\text{C}_{\text{quat}}$ ), 171.0 ( $\text{C}_{\text{quat}}$ ), 166.8 ( $\text{C}_{\text{quat}}$ ), 158.8 ( $\text{C}_{\text{quat}}$ ), 155.2 ( $\text{C}_{\text{quat}}$ ), 151.3 ( $\text{C}_{\text{quat}}$ ), 148.0 ( $\text{C}_{\text{quat}}$ ), 138.5 ( $\text{C}_{\text{quat}}$ ), 131.8 (CH), 131.3 ( $\text{C}_{\text{quat}}$ ), 130.0 (CH), 126.8 (CH), 117.7 (CH), 113.3 (CH), 65.7 ( $\text{CH}_2$ ), 57.8 ( $\text{CH}_2$ ), 55.1 ( $\text{CH}_2$ ), 54.5 ( $\text{CH}_2$ ), 51.1 ( $\text{CH}_2$ ), 50.8 ( $\text{CH}_2$ ), 50.2 ( $\text{CH}_2$ ), 50.1 ( $\text{CH}_2$ ), 20.1 ( $\text{CH}_3$ ).

ESI-MS: monoisotopic  $m/z$  = 646.4  $[\text{M}+\text{H}]^+$ , 668.4  $[\text{M}+\text{Na}]^+$  (calculated monoisotopic  $m/z$  = 646.27  $[\text{M}+\text{H}]^+$ , 668.25  $[\text{M}+\text{Na}]^+$  for  $\text{M} = \text{C}_{30}\text{H}_{39}\text{N}_5\text{O}_{11}$ ).

HRMS (ESI+):  $m/z$  = 646.2701 (+) (calculated  $m/z$  = 646.2719  $[\text{M}+\text{H}]^+$  for  $\text{M} = \text{C}_{30}\text{H}_{39}\text{N}_5\text{O}_{11}$ ).

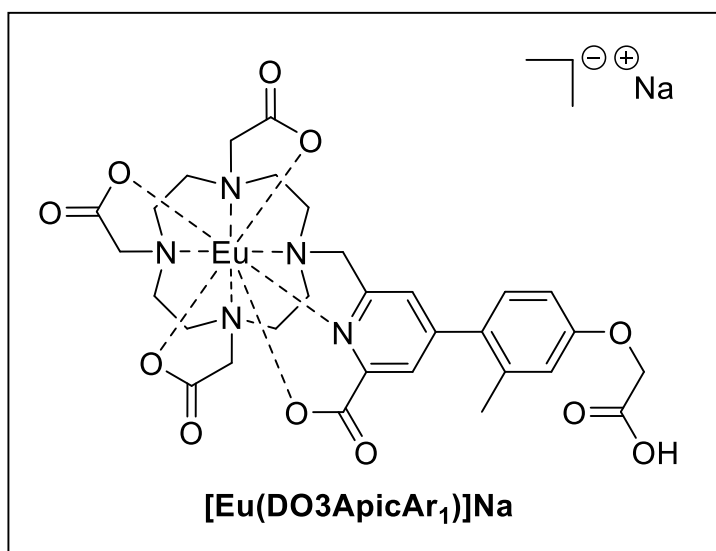


**[Tb(DO3ApicAr<sub>1</sub>)]Na :**

To a solution of **DO3ApicAr<sub>1</sub>** (27.8 mg, 43  $\mu$ mol) in water was added triethylamine to raise pH up to 9. Terbium(III) chloride (17.7 mg, 47  $\mu$ mol) was added and the pH was adjusted to 9 by subsequent addition of triethylamine. The mixture was stirred at r.t for 17 h. The reaction mixture was centrifuged and the supernatant evaporated under reduced pressure. Acetone was added and the suspension was sonicated and then centrifuged. The precipitate was collected and washed twice with acetone to give a white powder (21 mg, 61 %). Crystals suitable for X-ray diffraction were obtained by slow evaporation of a MeOD solution of the compound. The X-ray structure shows the macrocycle complexed with Tb<sup>3+</sup>, with Na<sup>+</sup> as a counter ion.

ESI-MS: monoisotopic  $m/z$  = 800.2 [M-H]<sup>-</sup> (calculated monoisotopic  $m/z$  = 800.16 [M-H]<sup>-</sup> for M = C<sub>30</sub>H<sub>36</sub>N<sub>5</sub>O<sub>11</sub>Tb).

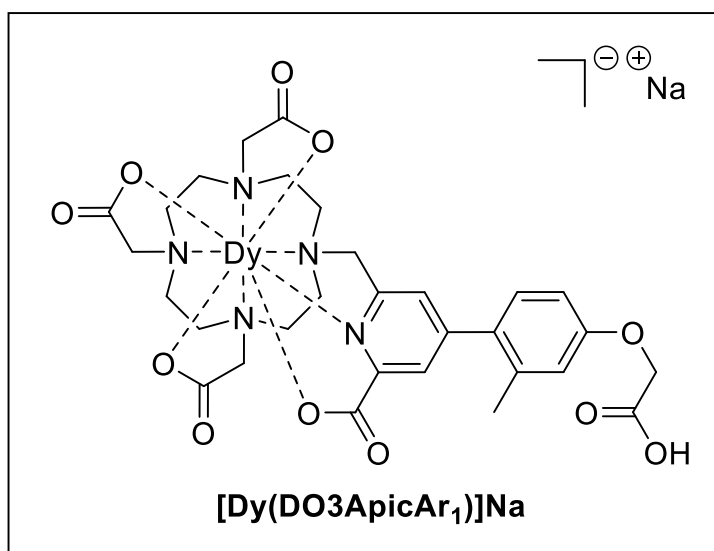
HRMS (ESI+):  $m/z$  = 800.1586 (+) (calculated  $m/z$  = 800.1592 [M+H]<sup>+</sup> for M = C<sub>30</sub>H<sub>36</sub>N<sub>5</sub>O<sub>11</sub>Tb).



The Eu complex **[Eu(DO3ApicAr<sub>1</sub>)]Na** was obtained as described above for the Tb analogue using europium(III) chloride as lanthanide salt.

ESI-MS: monoisotopic  $m/z = 794.2$   $[M-H]^-$  (calculated monoisotopic  $m/z = 794.15$   $[M-H]^-$  for  $M = C_{30}H_{36}EuN_5O_{11}$ ).

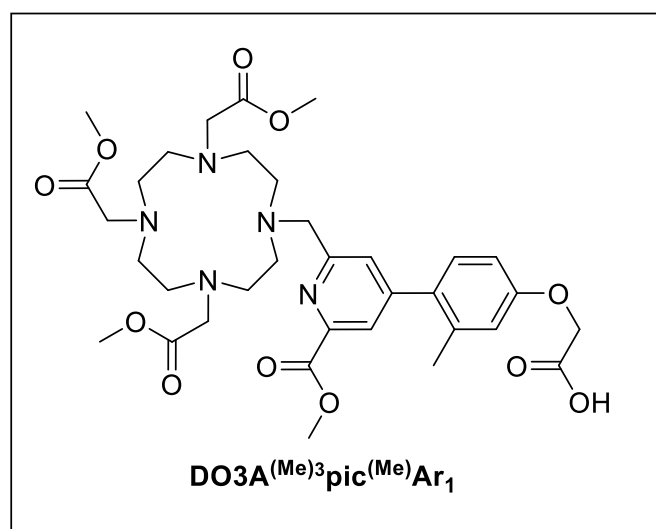
HRMS (ESI+):  $m/z = 794.1548$  (+) (calculated  $m/z = 794.1551$   $[M+H]^+$  for  $M = C_{30}H_{36}EuN_5O_{11}$ ).



The Dy complex **[Dy(DO3ApicAr<sub>1</sub>)]Na** was obtained as described above for the Tb analogue using Dysprosium(III) chloride as lanthanide salt.

ESI-MS: monoisotopic  $m/z = 805.2$   $[M-H]^-$  (calculated monoisotopic  $m/z = 805.16$   $[M-H]^-$  for  $M = C_{30}H_{36}DyN_5O_{11}$ ).

HRMS (ESI+):  $m/z = 805.1640$  (+) (calculated  $m/z = 805.1630$   $[M+H]^+$  for  $M = C_{30}H_{36}DyN_5O_{11}$ ).



### DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub> :

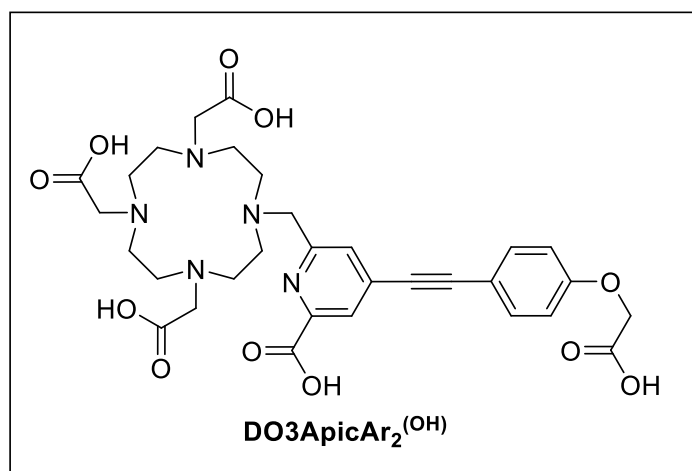
To a degassed solution of DO3A<sup>(Me)3</sup>picI<sup>(Me)</sup> (100 mg, 0.131 mmol) and **1'b** (41 mg, 0.123 mmol) in DMF (1.25 mL) was added *tetrakis*(triphenylphosphin)palladium(0) (14 mg, 12.1 mmol) and cesium fluoride (56 mg, 0.37 mmol). The mixture was stirred at 90°C for 17 h. Solvents were removed under reduced pressure. The oily residue was dissolved in DCM and washed with water. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and evaporated under reduced pressure. Water (16mL) was added to the crude. The suspension was sonicated for 20 min and centrifuged. The supernatant was collected and lyophilized to give DO3A<sup>(Me)3</sup>pic<sup>(Me)</sup>Ar<sub>1</sub>(*t*Bu) as a brown oil (117 mg). This product was identified by LC/MS but not characterized in depth due to the instability of the *tert*-butyl ester. It was used without further purification. Removal of *t*Bu protecting group was performed in TFA (5mL) at rt for 2 h. TFA was evaporated under reduced pressure and the crude was dissolved in water and purified on HPLC to give a slightly yellow oil (52.6 mg, 61 %) after lyophilization.

<sup>1</sup>H NMR (400MHz, D<sub>2</sub>O, 353K): δ (ppm) = 8.06 (d, J<sub>4</sub>=1.5Hz, 1H), 7.83 (d, J<sub>4</sub>=1.5Hz, 1H), 7.21 (d, J<sub>3</sub>=8.5Hz, 1H), 6.95 (d, J<sub>4</sub>=2.7Hz, 1H), 6.90 (dd, J<sub>3</sub>=8.5Hz, J<sub>4</sub>=2.7Hz, 1H), 4.75 (s, 2H), 4.58 (s, 2H), 4.06 (s, 2H), 4.01 (s, 3H), 3.70 (s, 3H), 3.59 (s, 4H), 3.48 (s, 6H) 3.80-2.98 (m, 16H), 2.23 (s, 3H).

<sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O, 353K): δ (ppm) = 173.1 (C<sub>quat</sub>), 172.5 (C<sub>quat</sub>), 168.7 (C<sub>quat</sub>), 166.8 (C<sub>quat</sub>), 158.7 (C<sub>quat</sub>), 153.1 (C<sub>quat</sub>), 152.5 (C<sub>quat</sub>), 148.0 (C<sub>quat</sub>), 138.2 (C<sub>quat</sub>), 131.5 (CH), 131.4 (C<sub>quat</sub>), 129.7 (CH), 126.5 (CH), 117.7 (CH), 113.3 (CH), 65.7 (CH<sub>2</sub>), 58.9 (CH<sub>2</sub>), 54.9 (CH<sub>2</sub>), 54.0 (CH<sub>3</sub>), 53.2 (CH<sub>2</sub>), 53.0 (CH<sub>3</sub>), 52.9 (CH<sub>3</sub>), 52.3 (CH<sub>2</sub>), 52.2 (CH<sub>2</sub>), 49.6 (CH<sub>2</sub>), 49.1 (CH<sub>2</sub>), 20.1 (CH<sub>3</sub>).

ESI-MS: monoisotopic  $m/z = 702.4$   $[M+H]^+$ ,  $724.4$   $[M+Na]^+$ ,  $740.3$   $[M+K]^+$  (calculated monoisotopic  $m/z = 702.33$   $[M+H]^+$ ,  $724.32$   $[M+Na]^+$ ,  $740.29$   $[M+K]^+$  for  $M = C_{34}H_{47}N_5O_{11}$ ).

HRMS (ESI+):  $m/z = 702.3328$  (+) (calculated  $m/z = 702.3345$   $[M+H]^+$  for  $M = C_{34}H_{47}N_5O_{11}$ ).

**DO3ApicAr<sub>2</sub>:**

To a degassed solution of **DO3A<sup>(Me)</sup><sub>3</sub>picI<sup>(Me)</sup>** (100 mg, 130  $\mu$ mol) and **2b** (29 mg, 111  $\mu$ mol) in DMF (1.1 mL) was added triethylamine (425  $\mu$ L, 3.05 mmol), *bis*(triphenylphosphine)palladium(II) dichloride (8 mg, 11.4  $\mu$ mol), and tetrabutylammonium fluoride 1M (130  $\mu$ L, 130  $\mu$ mol). The mixture was stirred at 90°C for 17 h. Solvents were removed under reduced pressure. The oily residue was dissolved in DCM and washed with water. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and evaporated under reduced pressure. Water (16mL) was added to the crude and the suspension was sonicated for 20 min, then centrifuged. The supernatant was lyophilized to give **DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>2</sub><sup>(Me)</sup>** as a brown oil (99 mg). This product was then used without further purification step, but a small amount of ligand was purified by HPLC for characterization. The protected ligand was dissolved in 1 M aq. NaOH (4 mL) and stirred at rt. The reaction was completed in 15 min as monitored by HPLC. The pH was adjusted to 4 by addition of 1 M aq. HCl (4 mL) The solution was purified on HPLC to give **DO3ApicAr<sub>2</sub>** as a yellow/orange solid (51 mg, 70 %), once lyophilized.

**DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>2</sub><sup>(Me)</sup>:**

<sup>1</sup>H NMR (400MHz, D<sub>2</sub>O, 353K):  $\delta$  (ppm) = 8.07 (d,  $J_4$ =1.3Hz, 1H), 7.81 (d,  $J_4$ =1.3Hz, 1H), 7.47 (d,  $J_3$ =9.0Hz, 2H), 6.95 (d,  $J_3$ =9.0Hz, 2H), 4.76 (s, 2H), 4.41 (s, 2H), 4.00 (s, 3H), 3.98 (s, 2H), 3.78 (s, 3H), 3.71 (s, 3H), 3.63 (s, 4H) 3.57 (s, 6H), 3.62-3.58 & 3.43-3.07 (m, 16H).

ESI-MS: monoisotopic  $m/z$  = 726.4 [M+H]<sup>+</sup>, 748.4 [M+Na]<sup>+</sup>, 764.3 [M+K]<sup>+</sup> (calculated monoisotopic  $m/z$  = 726.33 [M+H]<sup>+</sup>, 748.31 [M+H]<sup>+</sup>, 764.29 [M+H]<sup>+</sup> for M = C<sub>36</sub>H<sub>47</sub>N<sub>5</sub>O<sub>11</sub>).

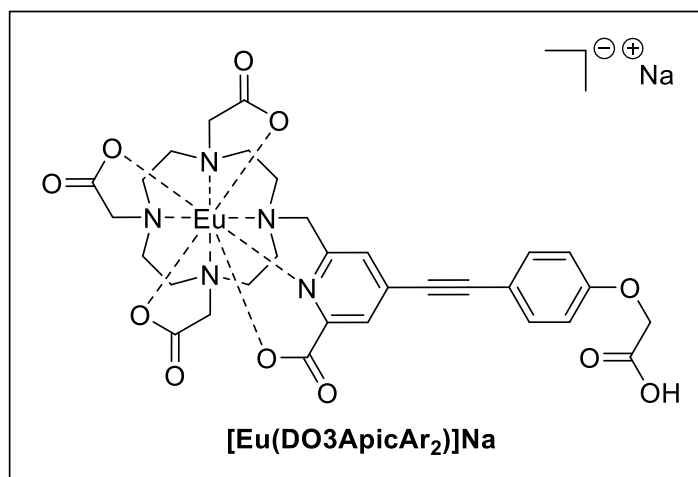
**DO3ApicAr<sub>2</sub>:**

<sup>1</sup>H NMR (400MHz, D<sub>2</sub>O, 353K):  $\delta$  (ppm) = 8.02 (d,  $J_4=1.4\text{Hz}$ , 1H), 7.73 (d,  $J_4=1.4\text{Hz}$ , 1H), 7.46 (d,  $J_3=9.0\text{Hz}$ , 2H), 6.89 (d,  $J_3=9.0\text{Hz}$ , 2H), 4.66 (s, 2H), 4.25 (s, 2H), 3.67 (s, 2H), 3.55 (s, 4H), 3.31-3.08 (m, 16H).

<sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O, 353K):  $\delta$  (ppm) = 173.1 (C<sub>quat</sub>), 172.0 (C<sub>quat</sub>), 171.4 (C<sub>quat</sub>), 166.8 (C<sub>quat</sub>), 159.4 (C<sub>quat</sub>), 152.5 (C<sub>quat</sub>), 148.4 (C<sub>quat</sub>), 136.6 (C<sub>quat</sub>), 134.7 (CH), 130.6 (CH), 127.7 (CH), 115.8 (CH), 114.8 (C<sub>quat</sub>), 98.5 (C<sub>quat</sub>), 85.8 (C<sub>quat</sub>), 65.7 (CH<sub>2</sub>), 57.8 (CH<sub>2</sub>), 55.2 (CH<sub>2</sub>), 54.6 (CH<sub>2</sub>), 50.9 (CH<sub>2</sub>), 50.7 (CH<sub>2</sub>), 50.4 (CH<sub>2</sub>), 50.2 (CH<sub>2</sub>).

ESI-MS: monoisotopic  $m/z$  = 656.4 [M+H]<sup>+</sup> (calculated monoisotopic  $m/z$  = 656.26 [M+H]<sup>+</sup> for M = C<sub>31</sub>H<sub>37</sub>N<sub>5</sub>O<sub>11</sub>).

HRMS (ESI+):  $m/z$  = 656.2599 (+) (calculated  $m/z$  = 656.2562 [M+H]<sup>+</sup> for M = C<sub>31</sub>H<sub>37</sub>N<sub>5</sub>O<sub>11</sub>).

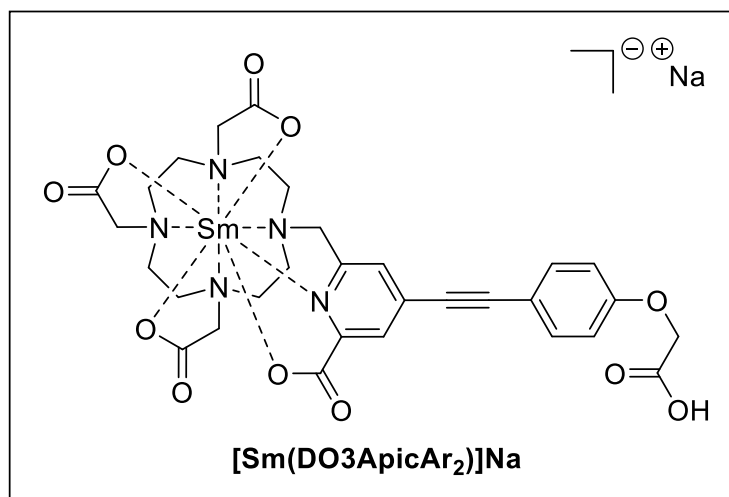


**[Eu(DO3ApicAr<sub>2</sub>)]Na :**

To a solution of DO3ApicAr<sub>2</sub> (30 mg, 46  $\mu$ mol) in water was added triethylamine to adjust the pH at 9. Europium(III) chloride (18.4 mg, 50  $\mu$ mol) was added and the pH was adjusted to 9 by subsequent addition of triethylamine. The mixture was stirred at r.t for 17 h and centrifuged. The supernatant was evaporated under reduced pressure and acetone was added to precipitate the complex. The suspension was sonicated then centrifuged. The precipitate was washed twice with acetone to give a white powder (22 mg, 60 %)

ESI-MS: monoisotopic  $m/z$  = 803.8 [M-H]<sup>-</sup> (calculated monoisotopic  $m/z$  = 804.14 [M-H]<sup>-</sup> for M = C<sub>31</sub>H<sub>34</sub>EuN<sub>5</sub>O<sub>11</sub>).

HRMS (ESI<sup>+</sup>):  $m/z$  = 804.1383 (+) (calculated  $m/z$  = 804.1394 [M+H]<sup>+</sup> for M = C<sub>31</sub>H<sub>34</sub>EuN<sub>5</sub>O<sub>11</sub>).

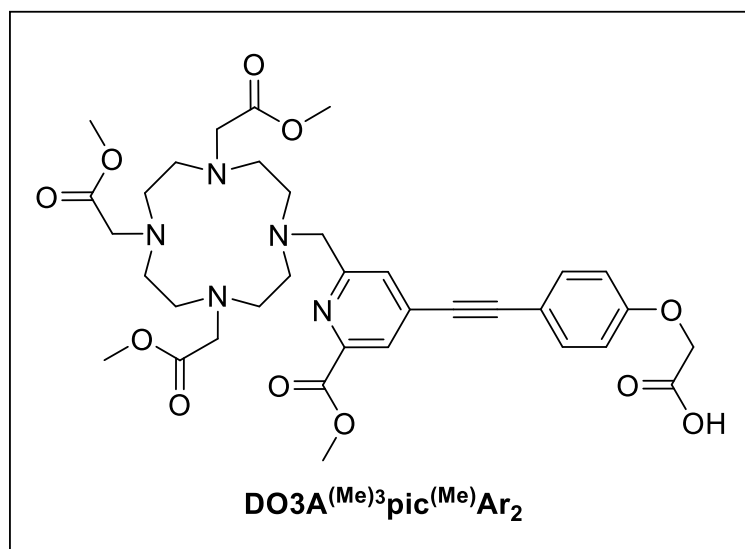


**[Sm(DO3ApicAr<sub>2</sub>)]Na :**

The Sm complex **[Sm(DO3ApicAr<sub>2</sub>)]Na** was obtained as described above for the Eu analogue using Samarium(III) chloride as lanthanide salt.

ESI-MS: monoisotopic  $m/z = 803.2$   $[M-H]^-$  (calculated monoisotopic  $m/z = 803.14$   $[M-H]^-$  for  $M = C_{31}H_{34}N_5O_{11}Sm$ ).

HRMS (ESI+):  $m/z = 803.1364$  (+) (calculated  $m/z = 803.1379$   $[M+H]^+$  for  $M = C_{31}H_{34}N_5O_{11}Sm$ ).

**DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>2</sub> :**

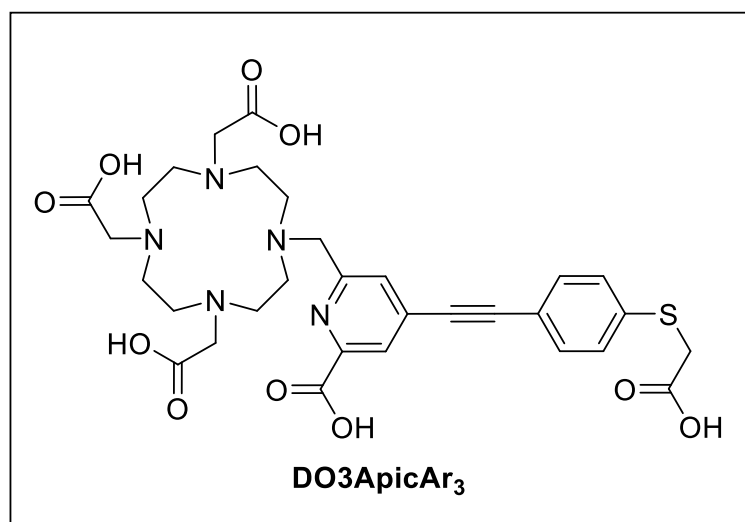
To a degassed solution of **DO3A<sup>(Me)</sup><sub>3</sub>picI<sup>(Me)</sup>** (100 mg, 130  $\mu$ mol) and **2'b** (33 mg, 108  $\mu$ mol) in DMF (1.1 mL) was added triethylamine (425  $\mu$ L, 3.05 mmol), *bis*(triphenylphosphin)palladium(II) dichloride (8 mg, 11.4  $\mu$ mol), and tetrabutylammonium fluoride 1M (130  $\mu$ L, 130  $\mu$ mol). The mixture was stirred at 90°C for 17 h. Solvents were removed under reduced pressure. The oily residue was dissolved in DCM and washed water. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and evaporated under reduced pressure. Water (16mL) was added to the crude and the suspension was sonicated for 20 min, then centrifuged. The supernatant was lyophilized to give DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>2</sub><sup>(tBu)</sup> as a brown oil (121 mg). This product was identified by LC/MS but not characterized in depth due to the instability of the *tert*-butyl ester. It was used without further purification. Removal of tBu was performed in TFA (5mL) at rt for 2h. TFA was evaporated under reduced pressure and the crude was dissolved in water and purified on HPLC to give a brown oil (38.6 mg, 50 %) after lyophilization.

<sup>1</sup>H NMR (400MHz, D<sub>2</sub>O, 353K):  $\delta$  (ppm) = 8.16 (d,  $J_4=0.1$ Hz, 1H), 7.87 (d,  $J_4=1.0$ Hz, 1H), 7.54 (d,  $J_3=8.8$ Hz, 2H), 7.02 (d,  $J_3=8.8$ Hz, 2H), 4.76 (s, 2H), 4.45 (s, 2H), 4.03 (s, 3H), 4.01 (s, 2H), 3.74 (s, 3H), 3.66 (s, 4H), 3.61 (s, 6H), 3.54-3.02 (m, 16H).

<sup>13</sup>C NMR (100 MHz, D<sub>2</sub>O, 353K):  $\delta$  (ppm) = 173.3 (C<sub>quat</sub>), 172.0 (C<sub>quat</sub>), 169.5 (C<sub>quat</sub>), 166.4 (C<sub>quat</sub>), 159.5 (C<sub>quat</sub>), 148.2 (C<sub>quat</sub>), 135.3 (C<sub>quat</sub>), 134.5 (C<sub>quat</sub>), 134.4 (CH), 130.9 (CH), 127.4 (CH), 115.9 (CH), 114.7 (C<sub>quat</sub>), 97.4 (C<sub>quat</sub>), 85.5 (C<sub>quat</sub>), 65.9 (CH<sub>2</sub>), 58.5 (CH<sub>2</sub>), 54.8 (CH<sub>2</sub>), 54.1 (CH<sub>3</sub>), 53.9 (CH<sub>3</sub>), 53.3 (CH<sub>2</sub>), 53.2 (CH<sub>3</sub>), 52.0 (CH<sub>2</sub>), 51.8 (CH<sub>2</sub>), 50.0 (CH<sub>2</sub>), 49.6 (CH<sub>2</sub>).

ESI-MS: monoisotopic  $m/z = 712.6$   $[M+H]^+$ ,  $734.6$   $[M+Na]^+$ ,  $750.6$   $[M+K]^+$  (calculated monoisotopic  $m/z = 712.32$   $[M+H]^+$ ,  $734.30$   $[M+Na]^+$ ,  $750.27$   $[M+K]^+$  for  $M = C_{35}H_{45}N_5O_{11}$ ).

HRMS (ESI+):  $m/z = 712.3174$  (+) (calculated  $m/z = 712.3188$   $[M+H]^+$  for  $M = C_{35}H_{45}N_5O_{11}$ ).

**DO3ApicAr<sub>3</sub> :**

To a degassed solution of **DO3A**<sup>(Me)<sub>3</sub></sup>**picI**<sup>(Me)</sup> (100 mg, 0.13 mmol) and **3b** (30 mg, 0.11 mmol) in DMF (1.1 mL) was added triethylamine (425  $\mu$ L, 3.05 mmol), *bis*(triphenylphosphin)palladium(II) dichloride (7.7 mg, 11  $\mu$ mol), and tetrabutylammonium fluoride 1M (130  $\mu$ L, 0.13 mmol). The mixture was stirred at 90°C for 17 h. Solvents were removed under reduced pressure. The oily residue was dissolved in DCM and washed with water. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and evaporated under reduced pressure. Water (16mL) was added to the crude. The suspension was sonicated for 20 min and then centrifuged. The supernatant was lyophilized to give **DO3A**<sup>(Me)<sub>3</sub></sup>**pic**<sup>(Me)</sup>**Ar<sub>3</sub>**<sup>(Me)</sup> as a brown oil (91 mg). This product was then used without further purification, but a small amount of ligand was purified by HPLC for characterization. The methyl-protected ligand was dissolved in 1 N aq. NaOH (4 mL) and stirred at rt. The reaction was completed in 15 min as monitored by HPLC. The pH was adjusted to 4 by addition of 1 M aq. HCl (4 mL). The solution was purified on HPLC to give **DO3A**<sup>(OH)<sub>3</sub></sup>**pic**<sup>(OH)</sup>**Ar<sub>3</sub>**<sup>(OH)</sup> as an orange solid (53 mg, 72 %) after lyophilization.

**DO3A**<sup>(Me)<sub>3</sub></sup>**pic**<sup>(Me)</sup>**Ar<sub>3</sub>**<sup>(Me)</sup>:

<sup>1</sup>H NMR (400MHz, D<sub>2</sub>O, 353K):  $\delta$  (ppm) = 8.00 (d,  $J_4$ =1.3Hz, 1H), 7.79 (d,  $J_4$ =1.3Hz, 1H), 7.35 (d,  $J_3$ =8.7Hz, 2H), 7.26 (d,  $J_3$ =8.6Hz, 2H), 4.41 (s, 2H), 3.98 (s, 3H), 3.97 (s, 2H), 3.77 (s, 2H), 3.68 (s, 3H), 3.64 (s, 3H), 3.63 (s, 4H) 3.53 (s, 6H), 3.42-3.05 (m, 16H).

ESI-MS: monoisotopic  $m/z$  = 742.4 [M+H]<sup>+</sup>, 764.3 [M+Na]<sup>+</sup>, 781.3 [M+K]<sup>+</sup> (calculated monoisotopic  $m/z$  = 742.31 [M+H]<sup>+</sup>, 764.29 [M+H]<sup>+</sup>, 780.27 [M+H]<sup>+</sup> for M = C<sub>36</sub>H<sub>47</sub>N<sub>5</sub>O<sub>10</sub>S).

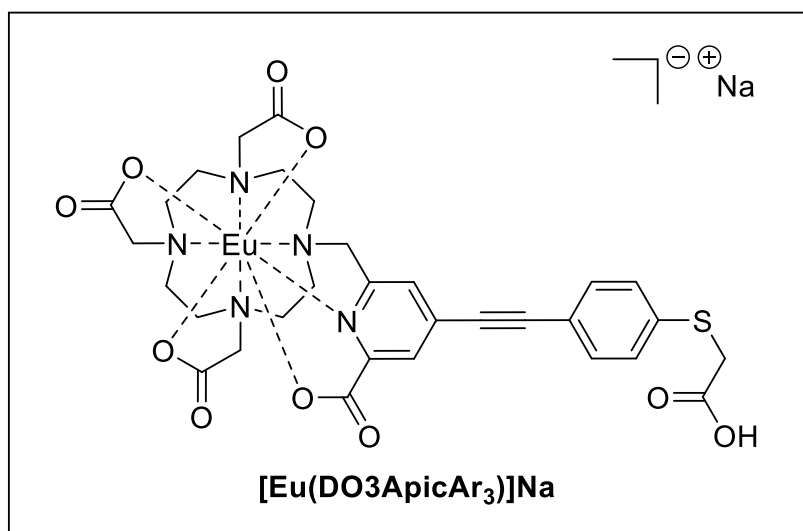
**DO3ApicAr<sub>3</sub>:**

$^1\text{H}$  NMR (400MHz,  $\text{D}_2\text{O}$ , 353K):  $\delta$  (ppm) = 8.04 (d,  $J_4=1.2\text{Hz}$ , 1H), 7.74 (d,  $J_4=1.2\text{Hz}$ , 1H), 7.41 (d,  $J_3=8.6\text{Hz}$ , 2H), 7.27 (d,  $J_3=8.6\text{Hz}$ , 2H), 4.37 (s, 2H), 3.87 (s, 2H), 3.78 (s, 2H), 3.60 (s, 4H), 3.40-3.15 (m, 16H).

$^{13}\text{C}$  NMR (100 MHz,  $\text{D}_2\text{O}$ , 353K):  $\delta$  (ppm) = 173.6 ( $\text{C}_{\text{quat}}$ ), 171.9 ( $\text{C}_{\text{quat}}$ ), 171.5 ( $\text{C}_{\text{quat}}$ ), 166.9 ( $\text{C}_{\text{quat}}$ ), 152.7 ( $\text{C}_{\text{quat}}$ ), 148.6 ( $\text{C}_{\text{quat}}$ ), 138.4 ( $\text{C}_{\text{quat}}$ ), 136.1 ( $\text{C}_{\text{quat}}$ ), 133.3 (CH), 130.6 (CH), 128.9 (CH), 127.7 (CH), 119.5 ( $\text{C}_{\text{quat}}$ ), 97.7 ( $\text{C}_{\text{quat}}$ ), 86.9 ( $\text{C}_{\text{quat}}$ ), 57.8 ( $\text{CH}_2$ ), 55.2 ( $\text{CH}_2$ ), 54.7 ( $\text{CH}_2$ ), 50.9 ( $\text{CH}_2$ ), 50.7 ( $\text{CH}_2$ ), 50.4 ( $\text{CH}_2$ ), 50.2 ( $\text{CH}_2$ ), 35.7 ( $\text{CH}_2$ ).

ESI-MS: monoisotopic  $m/z$  = 672.5  $[\text{M}+\text{H}]^+$ , 694.4  $[\text{M}+\text{Na}]^+$ , 710.4  $[\text{M}+\text{K}]^+$  (calculated monoisotopic  $m/z$  = 672.23  $[\text{M}+\text{H}]^+$ , 694.21  $[\text{M}+\text{Na}]^+$ , 710.19  $[\text{M}+\text{K}]^+$  for  $\text{M} = \text{C}_{31}\text{H}_{37}\text{N}_5\text{O}_{10}\text{S}$ ).

HRMS (ESI+):  $m/z$  = 672.2320 (+) (calculated  $m/z$  = 672.2334  $[\text{M}+\text{H}]^+$  for  $\text{M} = \text{C}_{31}\text{H}_{37}\text{N}_5\text{O}_{10}\text{S}$ ).

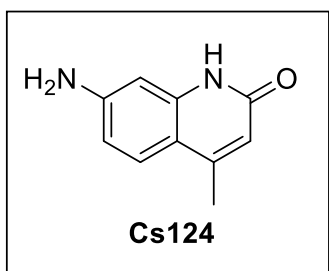


**[Eu(DO3ApicAr<sub>3</sub>)]Na :**

To a degassed solution of DO3ApicAr<sub>3</sub> (35.5 mg, 53  $\mu$ mol) in water was added triethylamine to adjust the pH at 9. Europium(III) chloride (21.3 mg, 58  $\mu$ mol) was added and the pH was adjusted to 9 by subsequent addition of triethylamine.. The mixture was stirred at r.t for 17 h. The reaction mixture was centrifuged, the supernatant evaporate and acetone was added to precipitate the complex. The suspension was sonicated, centrifuged and the precipitate was washed twice with acetone to give a white powder (14 mg, 32 %)

ESI-MS: monoisotopic  $m/z$  = 820.2 [M-H]<sup>-</sup> (calculated monoisotopic  $m/z$  = 820.12 [M-H]<sup>-</sup> for M = C<sub>31</sub>H<sub>34</sub>EuN<sub>5</sub>O<sub>10</sub>S).

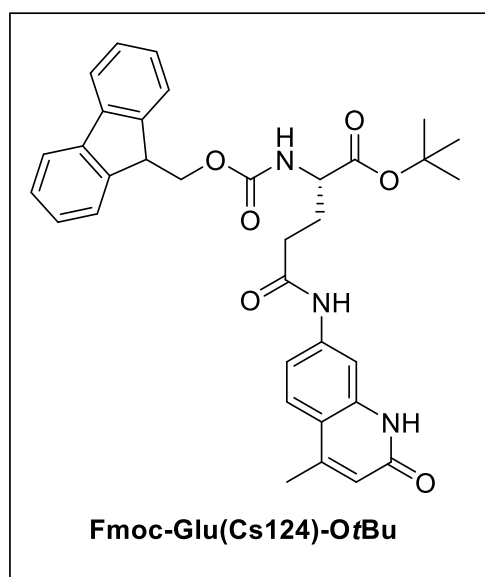
HRMS (ESI+):  $m/z$  = 820.1148 (+) (calculated  $m/z$  = 820.1166 [M+H]<sup>+</sup> for M = C<sub>31</sub>H<sub>34</sub>EuN<sub>5</sub>O<sub>10</sub>S)



**7-amino-4-methylquinolin-2(1*H*)-one (Cs124)** <sup>[196]</sup>:

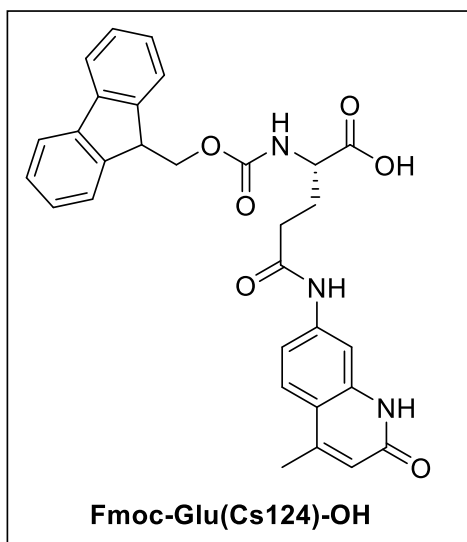
A solution of 1,3-diaminobenzene (3.0 g, 27.7 mmol) in ethylacetate (4.2 mL, 33.3 mmol) was stirred at reflux (150°C) for 20h. It was then poured on 300 mL of cold water, sonicated then stirred for 15 min. The precipitate was filtered and purified by flash chromatography (silica gel, DCM/MeOH 99:1) to give a white powder (1.01 g, 21%).

<sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>):  $\delta$  (ppm) = 7.34 (d,  $J_3=9.0$  Hz, 1H), 6.45 (d,  $J_3=9.0$  Hz, 1H), 6.36 (s, 1H), 5.94 (s, 1H), 5.72 (s, 2H), 2.27 (s, 3H)

**Fmoc-Glu(Cs124)-OtBu** <sup>[93]</sup>:

This product was synthesized adapting the procedure described in the literature. To a degassed solution of Fmoc-Glu-OtBu (1.11 g, 2.60 mmol), HATU (890 mg, 2.34 mmol), and DIEA (550  $\mu$ L, 3.12 mmol) in DMF (8 mL), was added a solution of **Cs124** (500 mg, 2.87 mmol) in DMF (8 mL). The mixture was stirred at r.t. for 15 h, and then solvents were removed under vacuum. The oily residue was dissolved in AcOEt, washed with 0.1 M aq. HCl (3x) and water. The organic layer was dried over Na<sub>2</sub>SO<sub>4</sub> and the solvent was removed under reduced pressure. The crude product was purified by flash chromatography (silica gel, CHCl<sub>3</sub>/MeOH 95:5) and precipitated in AcOEt to give a white powder (587 mg, 43 %).

<sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>):  $\delta$  (ppm) = 10.1 (s, 1H), 7.88 (d,  $J_3$  = 7.5 Hz, 2H), 7.79 (d,  $J_4$  = 1.5 Hz, 1H), 7.72 (d,  $J_3$  = 7.5 Hz, 2H), 7.60 (d,  $J$  = 8.8 Hz, 1H), 7.41 (t,  $J$  = 7.5 Hz, 2H), 7.37 – 7.26 (m, 3H), 6.25 (s, 1H), 4.37–4.15 (m, 3H), 4.06–3.96 (m, 1H), 2.48 – 2.43 (m, 2H), 2.36 (s, 3H), 2.15–2.01 (m, 1H), 1.96–1.78 (m, 1H), 1.40 (s, 9H)



**Fmoc-Glu(Cs124)-OH** <sup>[93]</sup>:

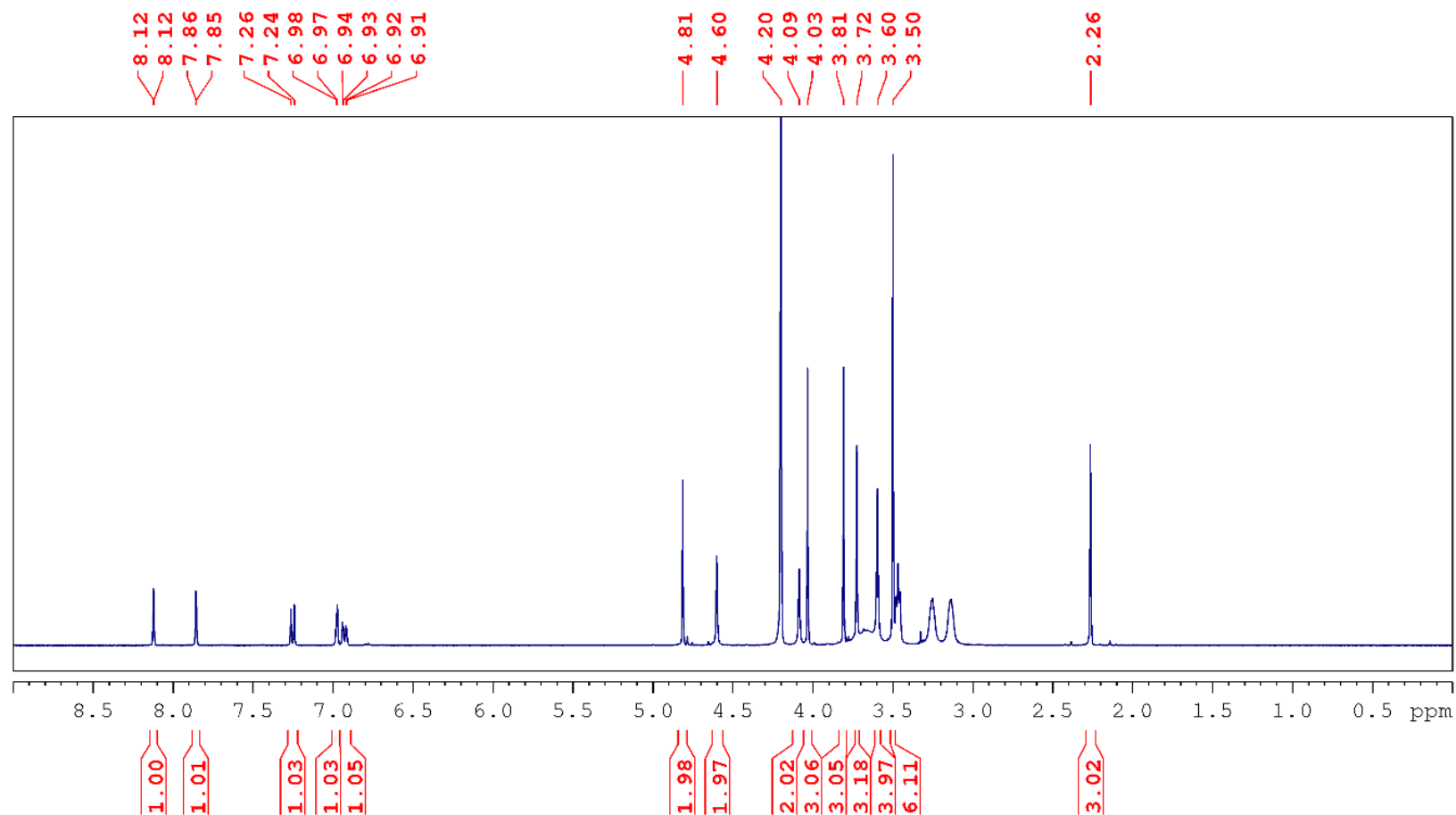
**Fmoc-Glu(Cs124)-OtBu** (560 mg, 0.97 mmol) was dissolved on a mixture of TFA/triisopropylsilane 10:1 (8 mL), and stirred at room temperature for 2h. TFA was evaporated under reduce pressure, and the resulting oil was precipitated on cold Et<sub>2</sub>O, sonicated, and centrifuged. The resulting precipitate was washed twice with Et<sub>2</sub>O, to afford a white powder (452 mg, 88%).

<sup>1</sup>H NMR (400 MHz, DMSO-d<sub>6</sub>):  $\delta$  (ppm) = 10.20 (s, 1H), 7.88 (d,  $J_3$  = 7.5Hz, 2H), 7.80 (d,  $J_4$  = 1.6 Hz, 1H), 7.72 (d,  $J_3$  = 7.5Hz, 2H), 7.60 (d,  $J$  = 8.8 Hz, 1H), 7.42 (t,  $J$  = 7.4 Hz, 2H), 7.37 – 7.24 (m, 3H), 6.25 (s, 1H), 4.34-4.18 (m, 3H), 4.09-3.99 (m, 1H), 2.49 – 2.42 (m, 2H), 2.36 (s, 3H), 2.21-2.05 (m, 1H), 1.97-1.80 (m, 1H)

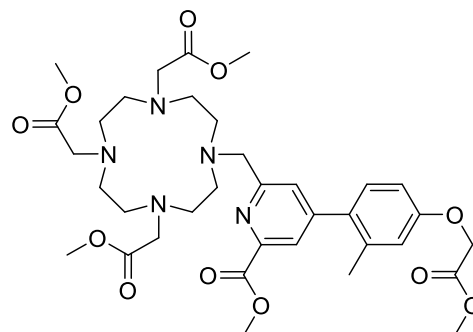


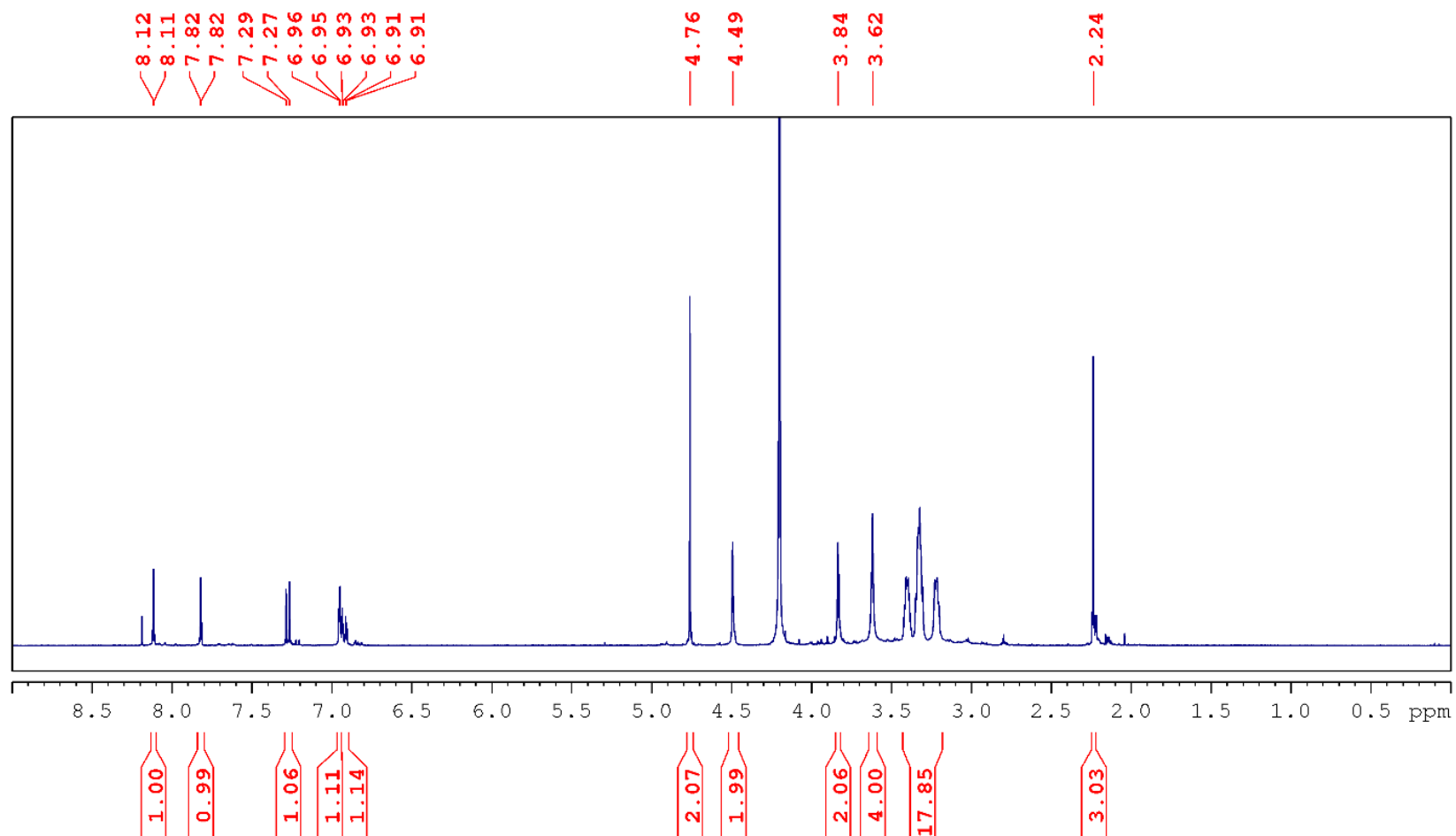
# Annexes



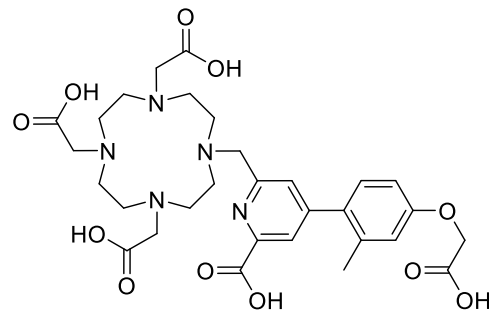


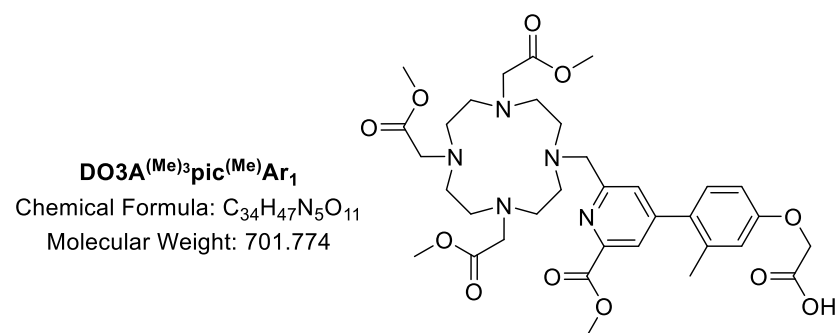
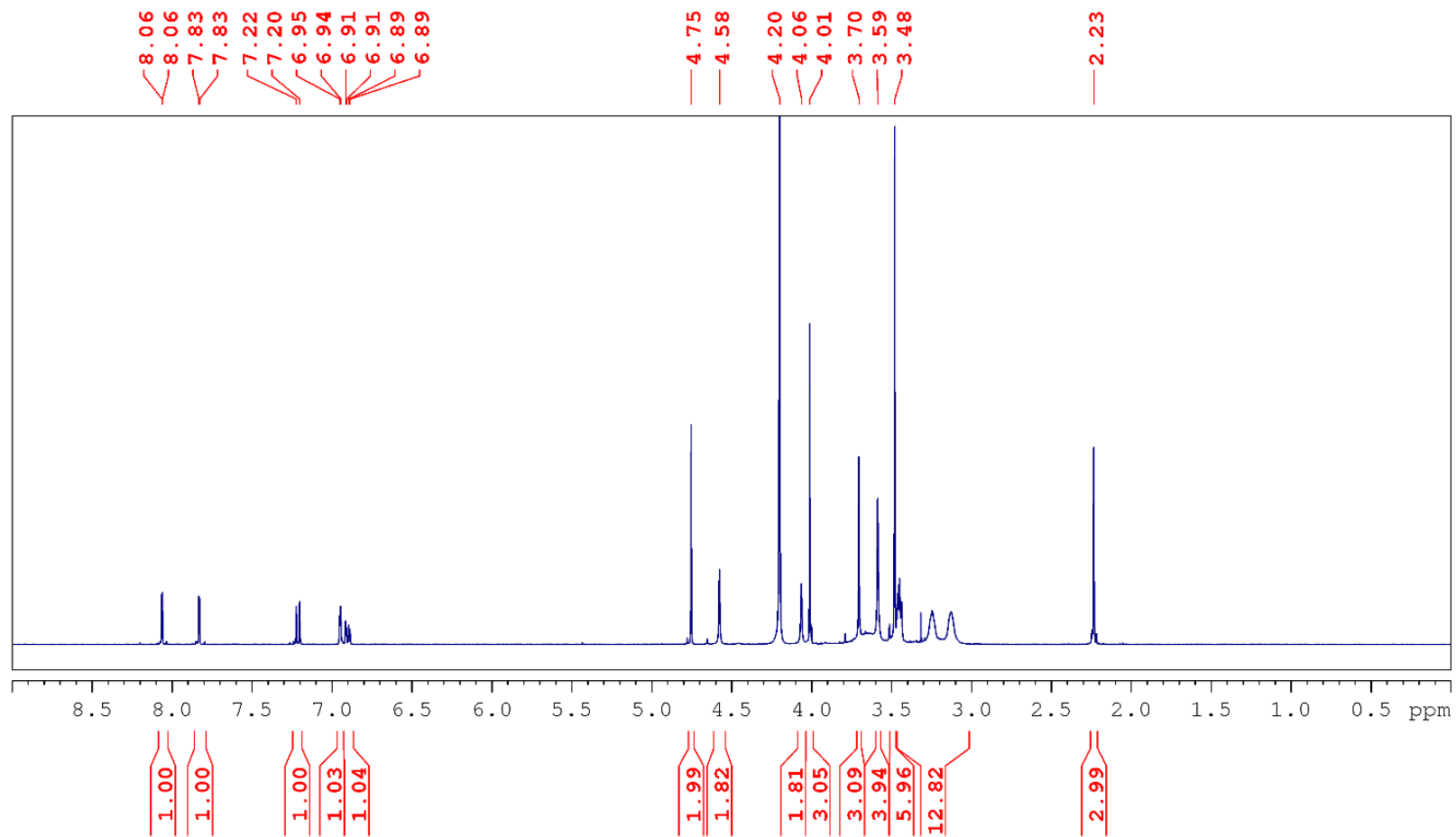
**DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>1</sub><sup>(Me)</sup>**  
 Chemical Formula: C<sub>35</sub>H<sub>49</sub>N<sub>5</sub>O<sub>11</sub>  
 Molecular Weight: 715,79

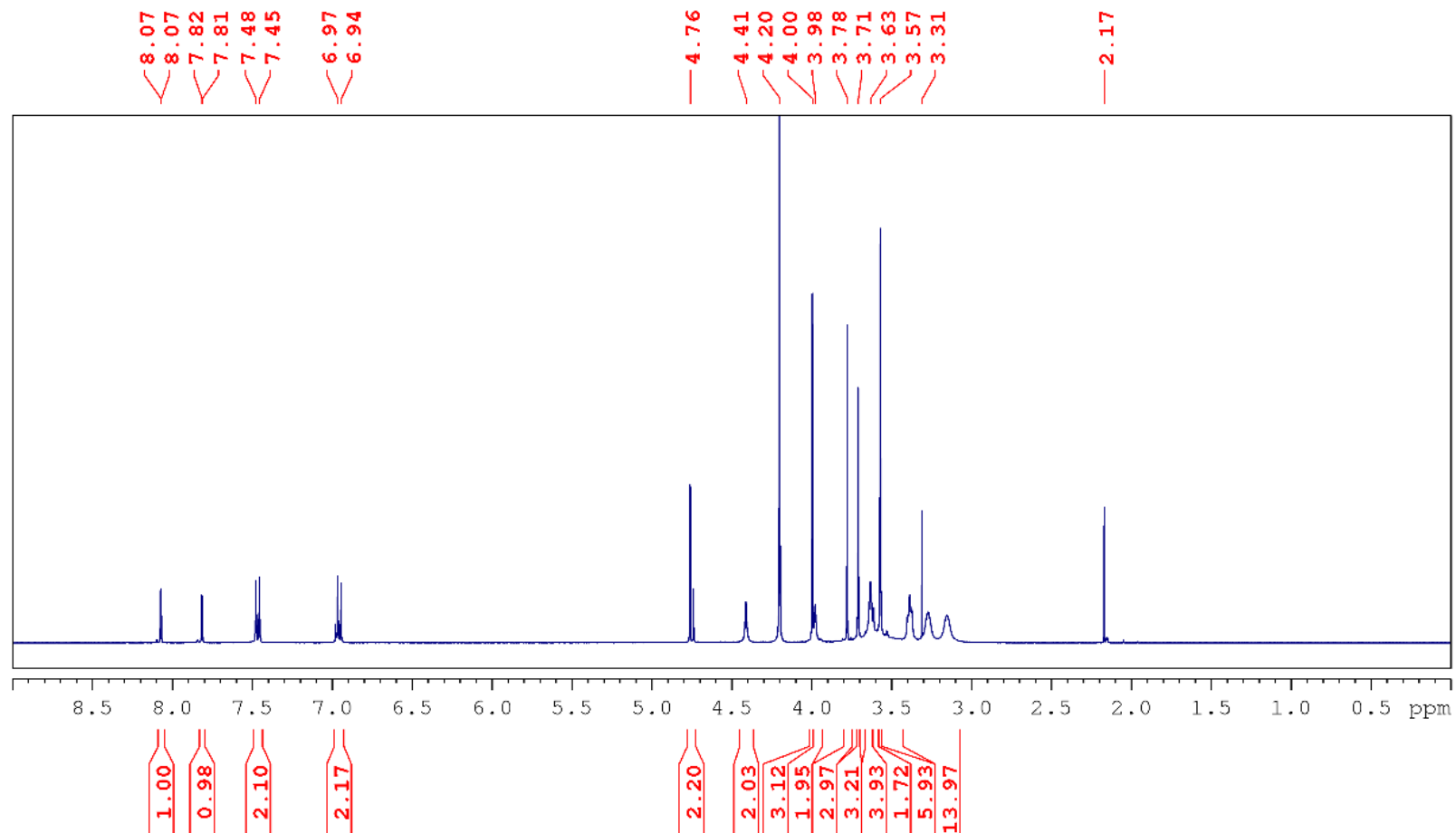




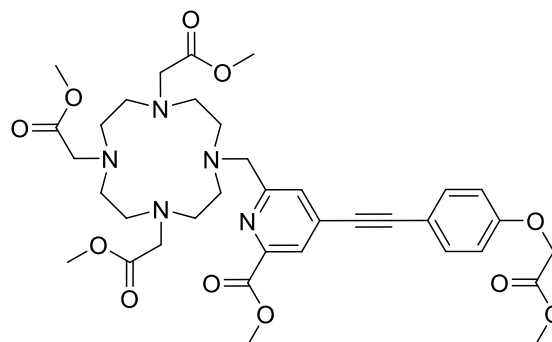
**DO3ApicAr<sub>1</sub>**  
Chemical Formula: C<sub>30</sub>H<sub>39</sub>N<sub>5</sub>O<sub>11</sub>  
Molecular Weight: 645.666

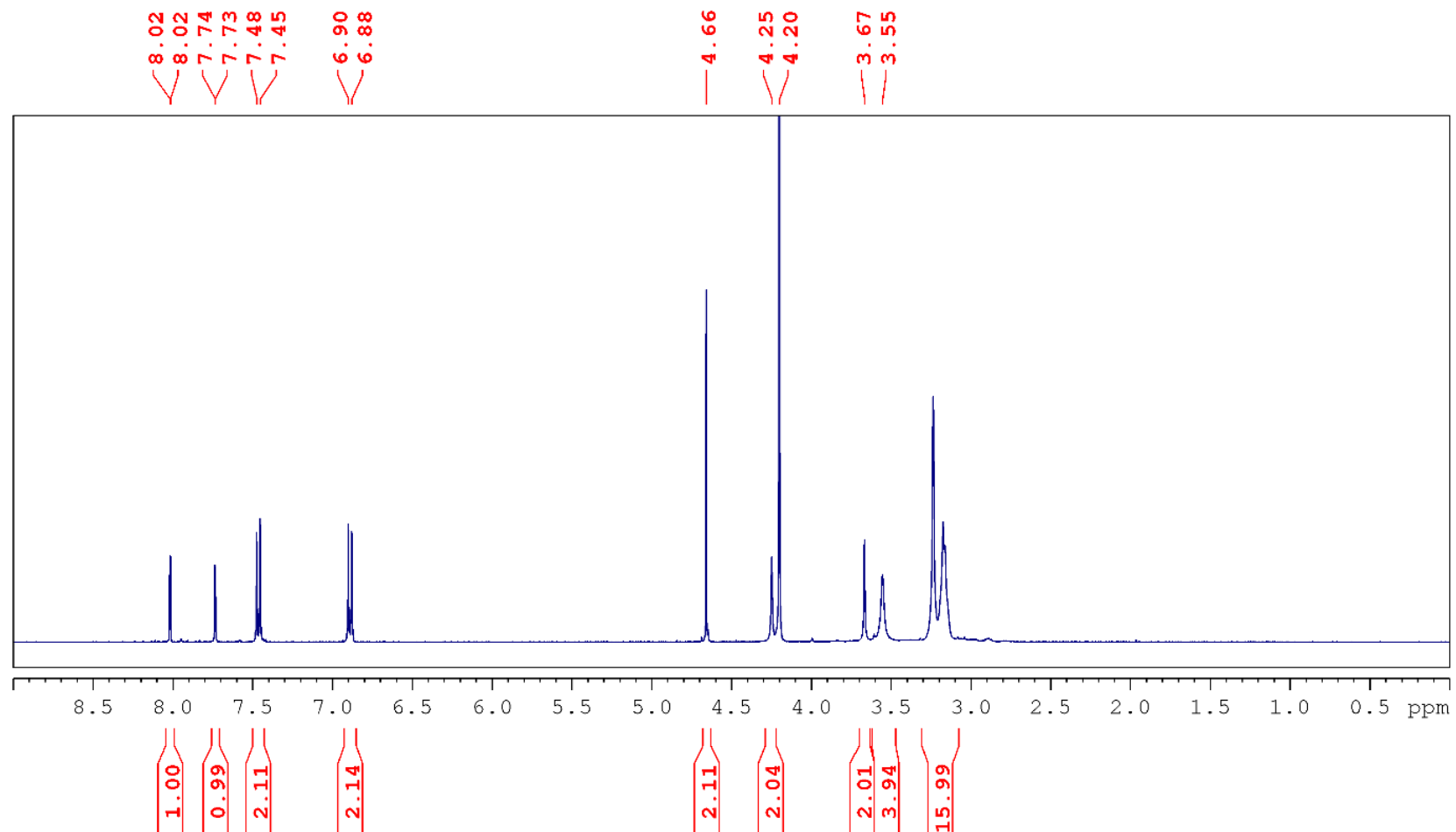




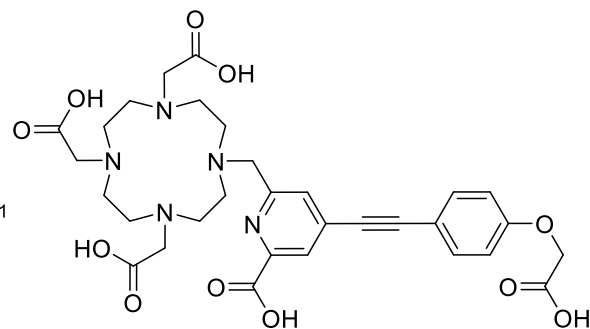


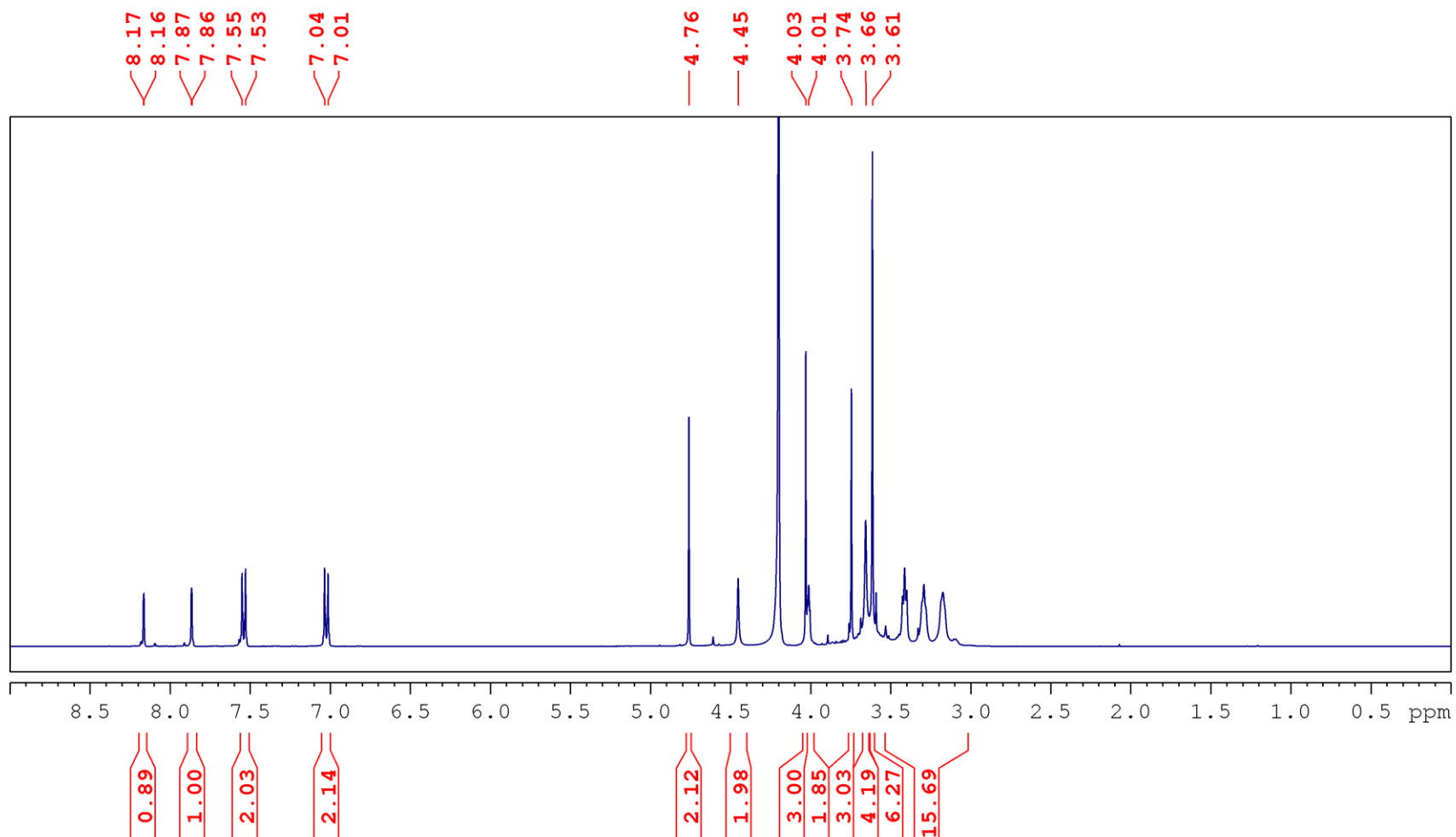
**DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>2</sub><sup>(Me)</sup>**  
 Chemical Formula: C<sub>36</sub>H<sub>47</sub>N<sub>5</sub>O<sub>11</sub>  
 Molecular Weight: 725.796



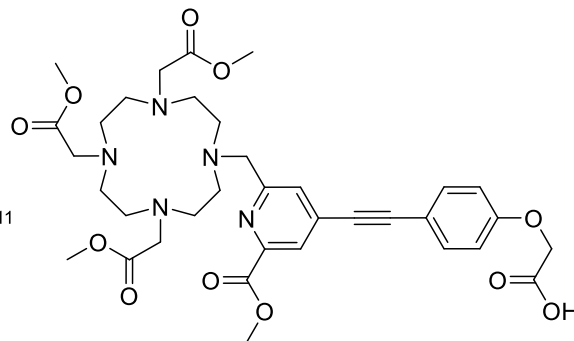


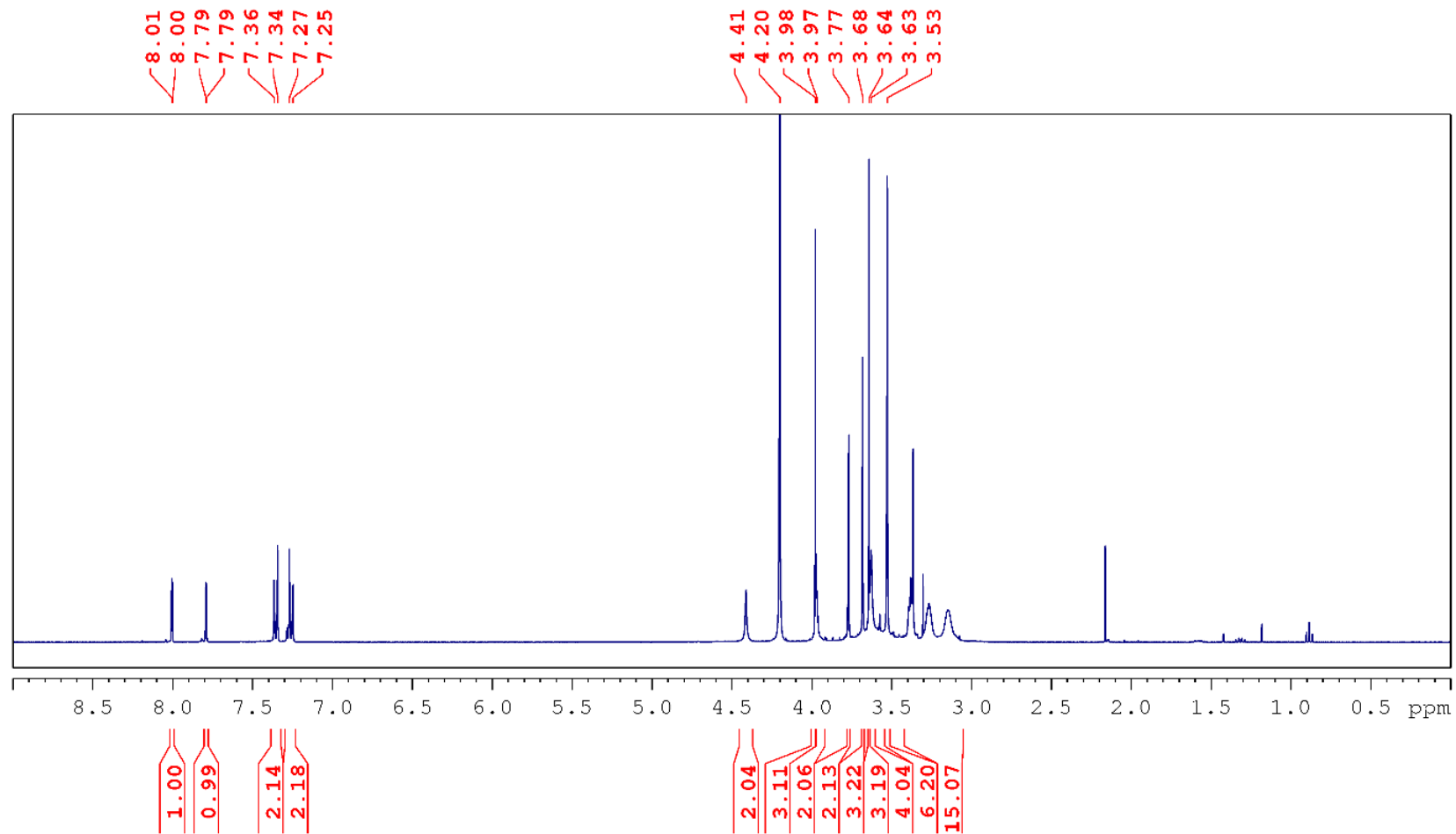
**DO3ApicAr2**  
 Chemical Formula: C<sub>31</sub>H<sub>37</sub>N<sub>5</sub>O<sub>11</sub>  
 Molecular Weight: 655.661



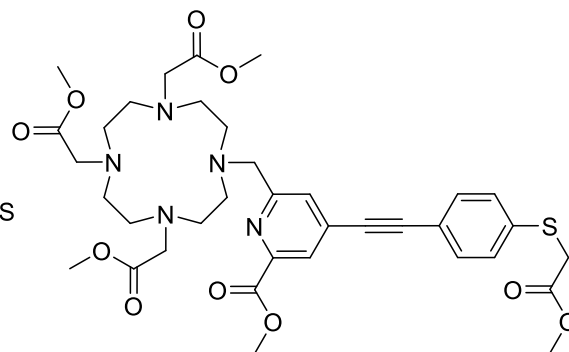


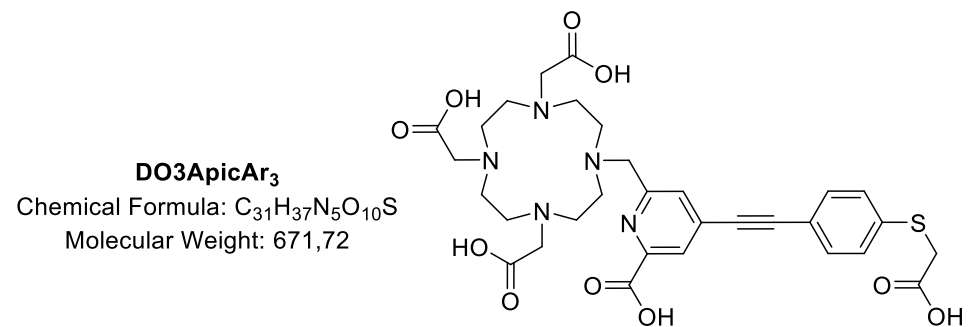
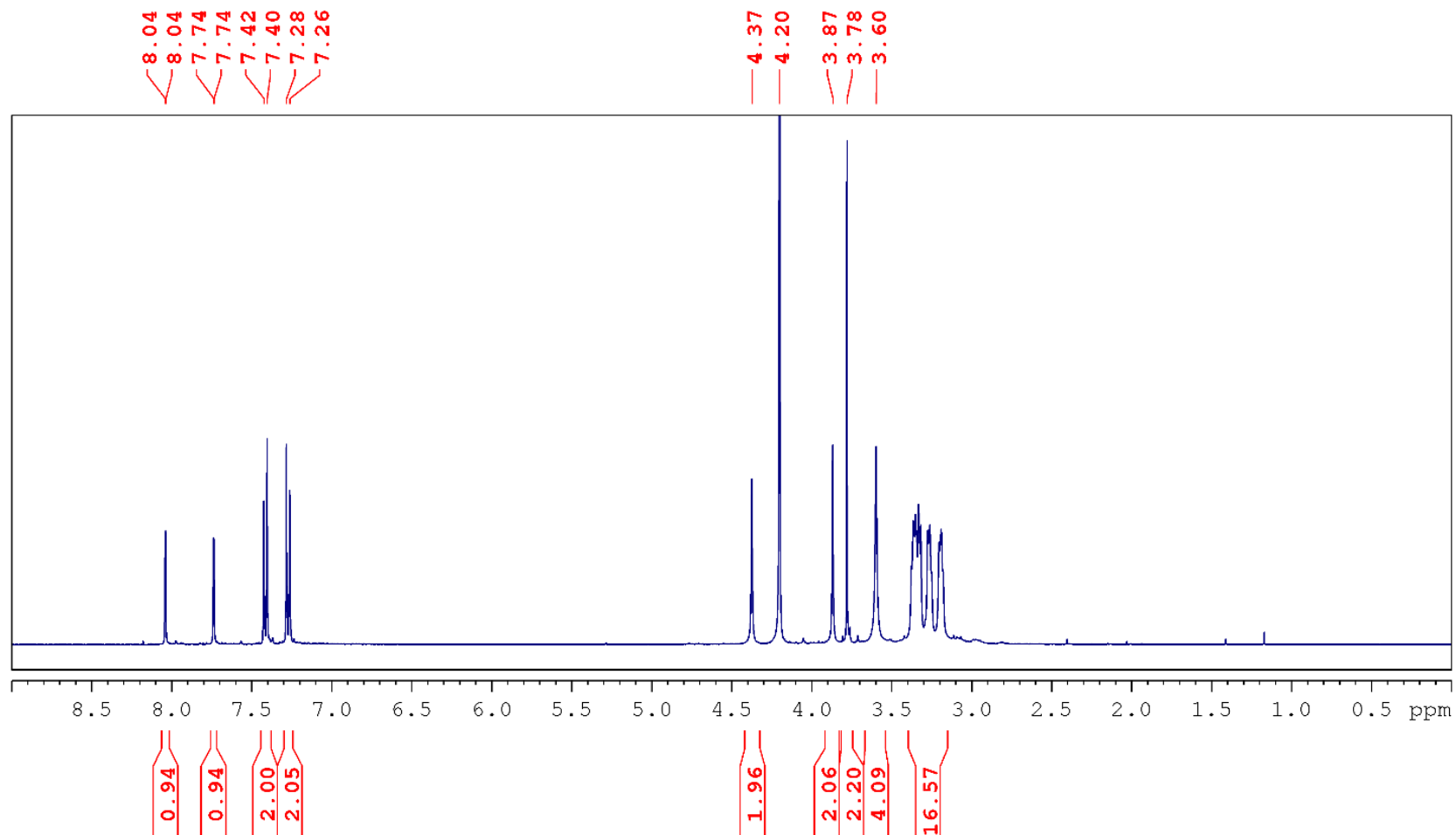
**DO3A(Me)<sub>3</sub>pic(Me)Ar<sub>2</sub>**  
Chemical Formula: C<sub>35</sub>H<sub>45</sub>N<sub>5</sub>O<sub>11</sub>  
Molecular Weight: 711,76





**DO3A<sup>(Me)</sup><sub>3</sub>pic<sup>(Me)</sup>Ar<sub>3</sub><sup>(Me)</sup>**  
 Chemical Formula: C<sub>36</sub>H<sub>47</sub>N<sub>5</sub>O<sub>10</sub>S  
 Molecular Weight: 741,85







## Résumé

L'ADN est le support de l'information de tout être vivant. La possibilité de pouvoir cibler et visualiser *in vivo* une séquence spécifique d'ADN et plus particulièrement un gène est un enjeu de taille pour le suivi médical aussi bien que pour la compréhension du vivant. Pour y parvenir, la détection par luminescence est particulièrement attrayante de par sa facilité de visualisation avec des outils simples.

L'objet de cette thèse était d'apporter la preuve de concept de sondes luminescentes pour la détection séquence spécifique d'ADN double brin, basées d'une part sur les propriétés de luminescence des lanthanides, particulièrement intéressantes pour la détection en milieu biologique, et d'autre part sur les propriétés de reconnaissance de l'ADN par les protéines à doigts de zinc. Nous nous sommes intéressé au ciblage d'un duplex d'ADN palindromique de 12 paires de bases par un couple de protéines à doigts de zinc intégrant un système FRET, où un complexe de lanthanide(III) sur une protéine joue le rôle de donneur et un fluorophore organique sur l'autre protéine joue celui d'accepteur.

Pour cela, une nouvelle famille de complexes de lanthanide(III) bioconjuguables a été élaborée et des protéines à doigts de zinc fonctionnalisées par différents chromophores ont été synthétisées chimiquement par synthèse peptidique supportée sur résine et assemblage par ligation chimique native de trois fragments. Les caractérisations spectroscopiques des systèmes développés ont permis de mettre en évidence l'interaction des sondes avec la séquence d'ADN palindromique et de valider la preuve de concept d'une détection de cette séquence par un FRET basé sur des lanthanides.

### Mots clés :

ADN – Doigts de zinc – Lanthanides – Ligation native – Détection – FRET

## Abstract

DNA is the carrier of genetic information in living organisms. Targeting and visualizing *in vivo* a specific DNA sequence is of particular interest for medical diagnosis and biological research. In this purpose, luminescent detection is very attractive because it can be easily observed with simple tools.

The aim of this work was to establish the proof of concept of luminescent probes for the sequence-specific detection of double stranded DNA based on lanthanide luminescence, which is attractive for biological applications, and on zinc finger proteins for their DNA binding properties. We focused on the detection of a 12-base pair palindromic DNA using a pair of zinc finger proteins, one bearing a lanthanide(III) complex as a the FRET donor and the other an organic fluorophore as an acceptor.

In this purpose, a new family of bioconjugatable lanthanide(III) complexes was developed and zinc finger proteins with various chromophores were synthesized chemically, by combination of solid phase peptide synthesis and assembly of peptidic segments by native chemical ligation. The spectroscopic characterizations of these systems have evidenced the interaction of the probes with the palindromic DNA, thereby validating the proof of concept of luminescence detection of this DNA sequence by a lanthanide-based FRET system.

### Key words :

DNA – Zinc Fingers – Lanthanides – Native Chemical – Detection – FRET