

UNIVERSITÉ DE MONTRÉAL

**Synthèse de composés outils permettant l'approfondissement
des connaissances biologiques dans le domaine oncologique**

Par

Cédric Dicaire-Leduc

Département de chimie, Université de Montréal

Faculté des Arts et des Sciences

Mémoire présenté à la Faculté des études supérieures et postdoctorales

En vue de l'obtention du grade de maître ès sciences (M.Sc.) en Chimie

Février 2021

© Cédric Dicaire-Leduc, 2021

Ce mémoire intitulé

**Synthèse de composés outils permettant l'approfondissement
des connaissances biologiques dans le domaine oncologique**

Présenté par

Cédric Dicaire-Leduc

A été évalué par un jury composé des personnes suivantes

Joelle N. Pelletier

Membre du jury

Stephen Hanessian

Membre du jury

Pierre Thibault

Directeur de recherche

Anne Marinier

Codirectrice de recherche

SOMMAIRE

Le corps humain est composé de plus de 100 000 protéines qui interagissent entre elles afin d'assurer son bon fonctionnement. Néanmoins, il peut arriver que certaines de ces protéines subissent une mutation changeant l'équilibre de fonctions du corps telles que la réplication cellulaire. Ces mutations affectant la réplication cellulaire peuvent mener à des cancers, une maladie qui touche de plus en plus de gens à travers le monde. Afin de développer des médicaments efficaces contre cette maladie, il est nécessaire de bien comprendre les mécanismes biologiques impliqués pour établir une approche thérapeutique. Cette compréhension repose notamment sur l'utilisation d'outils moléculaires, de petites molécules possédant des fonctions chimiques versatiles pouvant donner des informations critiques dans le développement de médicaments. Ce présent mémoire se consacre sur le développement d'outils moléculaires à travers deux projets ayant des objectifs distincts. Le premier projet avait comme principal objectif d'approfondir les connaissances sur l'inhibition de la protéine RAS, une GTPase responsable de plus de 30% des cancers. À cet effet, l'approche privilégiée a été celle de la synthèse de macrocycles peptidiques se liant à la cystéine 118 de RAS d'après la structure du monobody NS1. Les structures cristallines et les données protéomiques obtenues ont permis d'identifier les interactions clés entre la protéine RAS et une série d'inhibiteurs cherchant à émuler les effets biologiques de NS1. Le second projet s'intéresse quant à lui aux effets du composé UM171 sur les cellules souches. En effet, cette petite molécule possède la capacité d'empêcher la différenciation des cellules souches hématopoïétiques et permet donc leur multiplication. Cette propriété est une source d'espoir dans le traitement des leucémies et des transplantations. Cependant, la cible biologique de ce composé reste un mystère à ce jour. Ainsi la seconde partie de ce mémoire mettra de l'avant la synthèse d'outils moléculaire à base de diazirines pour tenter de venir identifier la cible d'intérêt en utilisant la protéomique.

Mots-clés : Synthèse en phase solide, Outils moléculaires, RAS, macrocycles peptidiques, Chimio-protéomique, Sondes diazirines, Cellules souches hématopoïétiques, Cristallographie, Oncologie, Synthèse hétérocyclique

ABSTRACT

The human body is made up of more than 100 000 proteins which interact with each other to keep it functioning properly. However, it can happen that some of these proteins undergo a mutation which changes the balance of body functions such as cell proliferation. These mutations affecting cell proliferation can lead to cancer, a disease that is affecting more and more people around the world. To develop effective drugs against this disease, it is necessary to understand the biological mechanisms involved in order to establish a therapeutic approach. This understanding is particularly based on the use of molecular tools, small molecules with versatile chemical functions that can provide critical information for drug development. This thesis is devoted to the development of molecular tools through two projects with distinct objectives. The main objective of the first project was to deepen knowledge about the inhibition of the RAS protein, a GTPase responsible for more than 30% of cancers. To this end, the preferred approach has been the synthesis of peptide macrocycles binding to RAS cysteine 118 based on the structure of NS1 monobody. The crystal structures and proteomic data obtained have made it possible to identify key interactions between the RAS protein and a series of inhibitors seeking to emulate the biological effects of NS1. The second project focuses on the effects of compound UM171 on stem cells. Indeed, this small molecule can prevent the differentiation of hematopoietic stem cells and therefore allows their multiplication. This property is a source of hope in the treatment of leukemia and transplants. However, the biological target of this compound remains a mystery to this day. Thus, the second part of this thesis will focus on the synthesis of molecular tools based on diazirines in an attempt to identify the target of interest using proteomics.

Keywords : Solid phase synthesis, Molecular tools, RAS, Peptide macrocycles, Chemoproteomic, Diazirine probes, Hematopoietic stem cells, Crystallography, Oncology, Heterocyclic synthesis

TABLE DES MATIÈRES

SOMMAIRE	i
ABSTRACT	ii
TABLE DES MATIÈRES	iii
LISTE DES FIGURES	v
LISTE DES TABLEAUX	ix
LISTE DES ABRÉVIATIONS	x
REMERCIEMENTS	xv
1. Introduction	1
1.1. Préface - Les outils moléculaires	1
1.2. Présentation du projet des macrocycles covalents.....	2
1.2.1 Importance de RAS dans l'oncologie	2
1.2.2 Caractéristiques structurales et isoformes de RAS	7
1.2.3 Approches de la littérature pour l'inhibition de RAS	12
1.2.4 Hypothèses des macrocycles covalents et démonstration du projet	19
1.3. Présentation du projet des sondes photoactivables de UM171.....	24
1.3.1 Cellules souches hématopoïétiques et applications médicales	24
1.3.2 Expansion des cellules souches hématopoïétiques par des petites molécules et UM171	26
1.3.3 Problématique et hypothèses du projet	29
2. Développement de macrocycles covalents cherchant à lier RAS	36
2.1. Synthèses sur support solide des macrocycles covalents	36
2.1.1 Discussions du choix des paramètres pour les macrocycles covalents.....	36
2.1.2 Voie de synthèses des macrocycles covalents	38
2.2. Résultats protéomiques et biochimiques.....	46
2.2.1 Analyse de la réactivité des ogives moléculaires	46
2.2.2 Structure cristalline de UM0140401	54
2.3. Optimisation des macrocycles covalents	56
2.3.1 Détermination de la longueur optimale du bras moléculaire	56
2.3.2 Effets du nombre de degrés de liberté du tour β	62
2.3.3 Structures cristallines des composés UM401, UM692 et UM6934	69
2.3.4 Détermination de l'affinité des macrocycles covalents	73

2.4. Conclusions et perspectives du projet	85
3. Développement de sondes photoactivables de UM171 pour l'identification de cibles protéiques ...	90
3.1 Synthèse des sondes	90
3.1.1 Conception des sondes.....	90
3.1.2 Synthèse des sondes diazirines substituées sur la chaîne aliphatique	94
3.1.3 Synthèse de la sonde diazirine substituée sur un aromatique	103
3.1.4 Détermination de l'activité des sondes.....	111
3.2 Optimisation de la méthode de chimioprotéomique	115
3.2.1 Approche pour l'optimisation de la méthode de chimioprotéomique	115
3.2.2 Incubation des sondes.....	116
3.2.3 Irradiation des sondes	117
3.2.3 Lyse cellulaire et optimisation de la réaction Click	123
3.2.4 Enrichissement et clivage des billes	127
3.3 Procédure de chimioprotéomique avec les sondes de UM171	133
3.3.1. Présentation de la procédure avec les sondes UM171.....	133
3.3.2. Résultats biochimiques et protéomiques de la procédure	136
3.4 Conclusion et perspective du projet	143
4. Conclusion générale.....	148
5. Références	149
6. Annexe A : Partie expérimentale et caractérisation	160
A.1 Organic synthesis section for the first project: Development of covalent macrocycles to mimic the NS1 monobody	161
A.2 Testing of the macrocycles (Biochemistry/Proteomics): Development of covalent macrocycles to mimic the NS1 monobody.....	198
A.3 Organic synthesis section for the second project: Target identification of the UM171 compound	201
A.4 Biochemistry/Proteomics section for the second project: Target identification of the UM171 compound	234

LISTE DES FIGURES

Figure 1 : Voie de signalisation de la voie cellulaire MAPK/ERK..	3
Figure 2 : Illustration du cycle de régulation de l'activité des RAS	5
Figure 3 : Représentation de la proportion des mutations dans les différents acteurs de la voie de signalisation MAPK/ERK parmi les cancers répertoriés.	6
Figure 4 : Caractéristiques structurales des différentes isoformes de RAS	8
Figure 5 : Effet conformationnel de RAS lors de l'activation de SOS.	10
Figure 6 : Mécanismes proposés pour l'activation dimérique de RAF à la membrane par RAS.	11
Figure 7 : Résumé des diverses stratégies thérapeutiques étudiées pour inhiber RAS oncogène	12
Figure 8 : Structure cristalline du site de liaison allostérique de NS1 pour HRAS	17
Figure 9 : Buvardage de Western démontrant l'effet de NS1 sur la phosphorylation en aval de RAS.	18
Figure 10 : Approfondissement des contacts importants de NS1 pour HRAS via la structure cristalline.	19
Figure 11 : Hypothèse synthétique pour mimer la boucle-FG de NS1 en macrocycle.	20
Figure 12 : Illustration de la cystéine 118 se trouvant à proximité de la boucle-FG de NS1	22
Figure 13 : Représentation des différents types cellulaires provenant de la différenciation des cellules souches hématopoïétiques	24
Figure 14 : Essai biologique utilisé par Sophie Corneau pour déterminer l'activité des molécules sur les marqueurs CD34+ et CD45RA-	27
Figure 15 : Procédure d'enrichissement par l'affinité du ligand pour la protéine en deux temps.	30
Figure 16 : Présentation des divers groupements nécessaires pour l'approche ABPP.	31
Figure 17 : Hypothèses synthétiques de la première série de macrocycles covalents se basant sur la structure d'UM0139995.	37
Figure 18 : Comparaison cinétique de diverses ogives de méthylamide réagissant avec la glutathionne.	37
Figure 19 : Séquence synthétique pour l'optimisation des macrocycles covalents en support solide.	39
Figure 20 : Réaction d'ancrage du Fmoc-Lys-Alloc sur la résine 2-chlorotriptyl de chlorure via une E1.	40
Figure 21 : Conditions réactionnelles utilisées pour coupler les différentes ogives aux amines des bras moléculaires.	42
Figure 22 : Procédure utilisée pour détecter les adduits sur les protéines intactes de RAS.	46
Figure 23 : Équation utilisée pour déterminer le pourcentage d'adduits selon l'aire des espèces de protéines en masse intacte	47
Figure 24 : Détermination des cystéines modifiées par des adduits du macrocycle 401 selon une analyse peptidique. La Glu-C a été utilisée pour la digestion de la protéine et une recherche d'adduit de 1455m/z a été effectuée. Les cystéines 80 et 118 ont démontré la présence d' adduits.	50
Figure 25 : Confirmation des résultats de protéomique via un gel SDS-Page par coloration au bleu de Coomassie pour les macrocycles '439 et '440.	53
Figure 26 : Structure cristalline du macrocycle '401 sur la GTPase HRAS selon une résolution de 4°A.	54
Figure 27 : Hypothèses synthétiques de la seconde série de macrocycles covalents cherchant à élucider la longueur optimale du bras moléculaire.	56
Figure 28 : Route synthétique pour l'acide aminé Fmoc-Dab(ivDde)-OH.	57
Figure 29 : Schéma des manipulations expérimentales pour l'analyse de la longueur optimale du bras moléculaire.	59

Figure 30 : Exemples de buvardage de Western pour l'étude d'activité des macrocycles covalents afin de déterminer la longueur optimale du bras moléculaire.	59
Figure 31 : Relation de la longueur optimale des bras moléculaires en fonction de l'emplacement du site d'attaque nucléophile et le nombre d'atomes.....	62
Figure 32 : Plan synthétique des différents tours β qui seront mis de l'avant dans l'étude de la flexibilité	62
Figure 33 : Réaction d'animation réductrice utilisée dans la synthèse sur support solide pour incorporer le tour β Aib réduit.....	64
Figure 34 : Synthèse de l'acide aminé Fmoc-(h-Aib)-OH via l'ajout du groupement protecteur Fmoc.....	64
Figure 35 : Synthèse de l'acide aminé 26 via la formation du carbamate Fmoc suivie d'une oxydation douce	64
Figure 36 : Exemples de buvardage de Western pour l'étude d'activité des macrocycles covalents afin de déterminer la flexibilité optimale du tour- β	66
Figure 37 : Structures complètes des composés pour lesquels un cristal a été obtenu.....	69
Figure 38 : Superposition des bras moléculaires se liant à la cystéine 118 des 3 macrocycles covalents..	70
Figure 39 : Superposition des différents macrocycles covalents, lorsque liés à la protéine HRAS.....	70
Figure 40 : Minimisation des angles et de la longueur des liens atomique pour diminuer la tension interne des macrocycles via la bio-informatique.....	71
Figure 41 : Interaction hydrogène intramoléculaire présente dans la structure cristalline de NS1.....	72
Figure 42 : Équation décrivant le modèle de l'inhibition irréversible.....	73
Figure 43 : Équation décrivant la constante d'affinité apparente dans l'inhibition irréversible	73
Figure 44 : Équation décrivant la constante cinétique de la réaction irréversible	74
Figure 45 : Équation décrivant la constante cinétique observée expérimentalement.....	74
Figure 46 : Régression mathématique obtenue selon le cas d'inhibition irréversible observée expérimentalement.....	75
Figure 47 : Équation décrivant la relation entre le pourcentage de liaison et la constante cinétique observable.....	76
Figure 48 : Graphiques expérimentaux de l'analyse cinétique du macrocycle UM0140401	77
Figure 49 : Graphiques expérimentaux de l'analyse cinétique du macrocycle UM0140440	77
Figure 50 : Approximation qu'il est possible de faire si la constante cinétique de dissociation k_{off} est très grande.....	79
Figure 51 : Plan synthétique afin de transformer les macrocycles covalents actifs en macrocycles covalents inactifs	81
Figure 52 : Analyse de l'importance des interactions non covalentes du macrocycle sur l'affinité via l'utilisation de macrocycles covalents inactifs par gel SDS-Page.	82
Figure 53 : Structure cristalline d'UM0140692 sur HRAS en observant spécifiquement les interactions avec l'hélice alpha 4.....	87
Figure 54 : Suggestions synthétiques cherchant à venir interagir avec l'arginine 135 de HRAS via la tyrosine du macrocycle.	88
Figure 55 : Idée synthétique qui permettrait d'interagir entre l'acide glutamique 162.....	88
Figure 56 : Structure chimique des fonctions photoactivables présentes dans l'approche ABPP ainsi que leur longueur d'onde d'activation et de leur insertion dans les acides aminés.....	91
Figure 57 : Résumé des observations générales concernant l'étude de relation structure-activité des analogues du composé UM171	93

Figure 58 : Approche conceptuelle des sondes photoactivables de UM171 selon les observations décrites par le RSA	93
Figure 59 : Séquence expérimentale pour l'obtention de la diazirine 34 fonctionnalisée par un alcool tosylé	95
Figure 60 : Formation du sel de pyridium en présence d'un tosyl d'alcool rapportée par Sekera.....	96
Figure 61 : Séquence expérimentale pour l'obtention de la diazirine fonctionnalisée par un acide carboxylique	96
Figure 62 : Séquence synthétique pour la formation des sondes diazirines aliphatiques analogues au composé UM171	99
Figure 63 : Spectre RMN focalisé sur la région aromatique afin d'illustrer la perte de l'hydrogène N-H aromatique lors de la réaction avec le tosyloxy d'alcool.....	101
Figure 64 : Valeur de pKa des espèces impliquées dans la réaction de SN ₂ activée par la base carbonate de potassium	101
Figure 65 : Structures corrigées pour les sondes aliphatiques qui devaient être substituées sur l'amine aliphatique par la diazirine mais qui en réalité ont été fonctionnalisées sur le N-H aromatique	103
Figure 66 : Rétro synthèse des grandes étapes pour la formation de la sonde diazirine aromatique	104
Figure 67 : Synthèse pour la formation de l'ester benzylique 53 nécessaire pour la formation du tricycle de la sonde	105
Figure 68 : Synthèse pour la formation du 2-amino-3-amide indole 57 nécessaire pour la formation du tricycle de la sonde.....	106
Figure 69 : Étape synthétique pour la construction de la trifluorométhyle cétone 63, précurseur de la diazirine aromatique	107
Figure 70 : Dernières étapes pour la formation de la sonde diazirine aromatique de UM171.....	109
Figure 71 : Formation d'un produit d'alkylation 70 par le cation t-Butyl.	111
Figure 72 : Suggestion d'une voie alternative pour les sondes diazirines substituées une amine basique sur la chaîne aliphatique	114
Figure 73 : Procédure de chimio-protéomique qui sera utilisée pour l'identification de la cible de UM171.	115
Figure 74 : Manipulations pour la détermination du temps d'incubation pour le composé UM171.....	116
Figure 75 : Spectre d'émission de la lampe B-100AP High Intensity de UVP.....	117
Figure 76 : Cartes de chaleur des diverses sondes et le composé UM171 qui mettent en corrélation les ions fragments en fonction de l'énergie de collision utilisée	119
Figure 77 : Schéma expérimental permettant d'observer des peptides modifiés par les sondes de UM171	120
Figure 78 : Exemple d'observations des ions fragments caractéristiques des sondes sur les spectres MS/MS des peptides de l'hémoglobine modifiées	121
Figure 79 : Distribution statistique de l'insertion du carbène de chaque sonde dans les acides aminés	122
Figure 80 : Détermination du tampon de lyse optimal selon une étude comparative de l'intensité de fluorescence par gel de la BSA	124
Figure 81 : Illustration de la technologie CalFluor pour le composé CalFluor 647	124
Figure 82 : Réponse de fluorescence à 674 nm de la BSA suite à une excitation à 647 nm pour le gel comparant 5 solvants de lyse à diverses concentrations et l'efficacité de la réaction Click.....	125
Figure 83 : Étude cinétique en fonction de la température pour la réaction Click selon une analyse de la fluorescence par gel de la BSA occasionnée par la liaison avec le CalFluor 647.....	126

Figure 84 : Preuve in vitro des conditions de clivages des biotines clivables	129
Figure 85 : Procédure expérimentale démontrant l'efficacité des conditions de clivages in vivo des biotines	130
Figure 86 : Gels démontrant l'efficacité du clivage de la fonction clivable sur la biotine suite à la réaction Click dans divers solvants de lyse	131
Figure 87 : Procédure modifiée de chimio-protéomique qui a été utilisée pour l'identification de la cible de UM171. Un gel SDS-PAGE a été réalisé pour séparer la streptavidine de l'analyse	134
Figure 88 : Coloration à l'argent du gel SDS-PAGE des protéines enrichies par les sondes diazirines de UM171.....	136
Figure 89 : Méthodologie qui a été appliquée pour filtrer les protéines identifiées par Peaks	138
Figure 90 : Carte de chaleur des protéines identifiées selon l'analyse stricte.....	140
Figure 91 : Carte de chaleur des protéines identifiées selon l'analyse souple	141
Figure 92 : Mécanisme réactionnel de la Diels-Alder avec une inversion de demande électronique entre la tétrazine et le transcyclooctène.....	145
Figure 93 : Principe de la technologie du ligand dirigé développé par Pr Hamachi et Dr Tamura	146
Figure 94 : Composition et fonctionnalisation des billes FG permettant d'enrichir les interacteurs d'une molécule d'intérêt.....	147

LISTE DES TABLEAUX

Tableau 1 : Identification et rendement globaux des macrocycles de la première optimisation	44
Tableau 2 : Pourcentages de modifications des protéines HRAS native et HRASC118S mutante par les divers macrocycles covalents quantifiés par spectrométrie de masse	48
Tableau 3 : Pourcentages de modifications des isoformes HRAS, KRAS et NRAS par les divers macrocycles covalents obtenus via la spectrométrie de masse	51
Tableau 4 : Rendements globaux et identification des macrocycles avec un bras moléculaire fait à partir du DAB.....	58
Tableau 5 : Valeurs d'EC ₅₀ pour la conversion des protéines HRAS selon les diverses longueurs de bras moléculaires pour deux familles d'ogives.	60
Tableau 6 : Rendements globaux et identification des macrocycles synthétisés pour l'étude de la flexibilité du tour β	65
Tableau 7 : Valeurs d'EC ₅₀ pour la conversion des protéines HRAS selon les divers tours- β et de longueurs de bras moléculaires.	67
Tableau 8 : Constantes de l'étude cinétique des macrocycles covalents '440 et '401.....	78
Tableau 9 : Rendements globaux et identification des macrocycles inactifs via la substitution de l'Allo-Ile par l'alanine.....	81
Tableau 10 : Efficacité de liaisons des macrocycles covalents actifs et inactifs par spectrométrie de masse sur HRAS	83
Tableau 11 : Diamines propargylés asymétriques synthétisés et leur rendement correspondant.....	98
Tableau 12 : Identification des sondes analogues à la structure de UM171 ainsi que leur activité biologique multiplicative sur les cellules souches hématopoïétiques	112
Tableau 13 : Pourcentage d'efficacité de liaison des sondes selon le ratio des PSM totaux et des PSM modifiés par une sonde.....	122
Tableau 14 : Détermination du nombre de protéines identifiées dans chaque condition selon les deux types d'analyses effectuées	139

LISTE DES ABRÉVIATIONS

ABPP	<i>Activity-based protein profiling</i>
CAN	Acétonitrile
ADPGK	<i>ADP Dependent Glucokinase</i>
ALDH2	Aldéhyde déshydrogénase 2
AIB	Acide 2-aminoisobutyrique
AIBN	Azobisisobutyronitrile
AhR	<i>Aryl hydrocarbon receptor</i>
Al	<i>Acyl imidazole</i>
Alloc	<i>Allyloxycarbonyl</i>
aq.	Aqueux ou aqueuse
BOC	tert-butoxycarbone
BSA	<i>Bovine serum albumin</i>
°C	Degré Celsius
CCM	Chromatographie sur couche mince
CD34/CD45	<i>Cluster of differentiation 34/45</i>
COMT	<i>Catechol-O-methyltransferase</i>
CPT1	<i>Carnitine palmitoyltransferase 1</i>
CMH	Complexe majeur d'histocompatibilité
CSH	Cellules souches hématopoïétiques
CuAAC	Copper-catalyzed alkyne-azide cycloaddition
CYP1A1/CYP1B1	Cytochrome P famille 1, sous-famille A/B et polypeptide 1
D	Doublet
Δ	Déplacement chimique
Da	Dalton
DAB	Acide 2,4-Diaminobutyrique
DADPS	Dialkoxydiphenylsilane
DBU	1,8-diazabicyclo[5.4.0]undec-7-ène
DCM	Dichlorométhane

Dd	Doublet de doublet
Ddd	Doublet de doublet de doublet
Dde	<i>[1-(4,4- dimethyl-2,6-dioxocyclohexylidene)ethyl]</i>
DEPBT	<i>diethoxyphosphoryloxy)-1,2,3-benzotriazin-4(3H)-one</i>
DIPEA	<i>N,N-Diisopropyléthylamine</i>
DMAP	4-diméthylaminopyridine
DMP	<i>Dess-Martin periodinane</i>
DMF	Diméthylformamide
DMSO	Diméthylsulfoxyde
Dt	Doublet de triplet
DTT	<i>Dithiothreitol</i>
EC50	Concentration efficace médiane
EC1,5	Concentration efficace pour un doublement
Eq.	Équivalent
Et	Éthyle
ERK	<i>Extracellular signal-regulated kinases</i>
ESI	<i>Elecytospray ionization</i>
FACS	<i>Fluorescence activated cell sorting</i>
FG loop	Chaîne de peptides reliant les feuillets F et G
FLAG	Séquence peptidique DYKDDDDK marquant les protéines
Fmoc	Fluorénylméthoxycarbonyle
FN3	<i>Fibronectin type III domain</i>
FXR	<i>Farnesoid X receptor</i>
GAP	<i>GTPase-activating protein</i>
GDP	Guanosine diphosphate
GEF	<i>Guanine nucleotide exchange factors</i>
GTP	Guanosine triphosphate
GRB2	<i>Growth factor receptor-bound protein 2</i>
h/hrs	Heure/heures
HATU	<i>Hexafluorophosphate azabenzotriazole tetramethyl uronium</i>

HBTU	<i>Hexafluorophosphate benzotriazole tetramethyl uronium</i>
HLA	<i>Human leucocyte antigen</i>
HPLC	Chromatographie en phase liquide à haute performance
HSQC	<i>Heteronuclear single quantum coherence spectroscopy</i>
Hz	Hertz
iEDDA	Inverse electron demand Diels-Alder
Ile	Isoleucine
ITC	<i>Isothermal titration calorimetry</i>
ITGB1	<i>Integrin beta-1</i>
ivDde	<i>l-(4,4-dimethyl-2,6- dioxocyclohex-1-ylidene)-3-methylbutyl</i>
J	Constante de couplage
kDa	Kilodalton
LCMS	<i>Liquid chromatography-mass spectrometry</i>
LSD1	<i>Lysine-specific histone demethylase 1A</i>
Lys	Lysine
M	Multiplet
MAPK	<i>Mitogen-actived protein kinases</i>
Me	Méthyle
MEK	MAPK/ERK Kinase
Mg	Miligramme
min	Minute
mol	Mole
mmol	Milimole
MS/MS	Spectrométrie de masse en tandem
MST	<i>Microscale thermophoresis</i>
MW	<i>Molecular Weight</i>
NASA	<i>N-Acyl-N-alkylsulfonamide</i>
NCE	Normalized collision energy
NK (cellules)	Cellules lymphocytes cytotoxiques naturels
nM	Nanomolaire

NMO	N-Methylmorpholine N-oxide
NP-40	<i>Nonylphenoxypolyethoxyethanol</i>
Orn	Ornithine
O-NBD	O-Nitrobenzoxadiazole
O-Su	O-Succinimide
PBS	Tampon phosphate salin
PDB	Protein Data Bank
Ph	Phényle
PI3K	Phosphoinositide 3-kinase
pKa	Logarithme négatif de la constante d'acidité
Pont H	Pont hydrogène
ppm	Partie par million
Pro	Proline
PSM	<i>Peptide-spectrum match</i>
Pyr.	Pyridine
PTB	<i>Phosphotyrosine-binding domain</i>
Q	Quadruplet
RALGDS	<i>Ral guanine nucleotide dissociation stimulator</i>
RAS	<i>Rat Sarcoma</i>
RAF	<i>Rapid Accelerated Fibrosarcoma</i>
RBD	<i>Ras-binding domain</i>
RCOR	<i>Rest corepressor</i>
Rdt	Rendement
RIPA	<i>Radioimmunoprecipitation assay buffer</i>
RMN ¹³C	Résonance magnétique nucléaire du carbone
RMN ¹H	Résonance magnétique nucléaire du proton
RSA	Relation structure-activité
RT	Room temperature
RTK	Receptor tyrosine kinase
S	Singulet

<i>SDS-PAGE</i>	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
<i>SH2</i>	Src homology 2
<i>SN2</i>	Substitution nucléophile bimoléculaire
<i>SnAr</i>	Substitution nucléophile aromatique
<i>SOS</i>	<i>Son of Sevenless</i>
<i>SPR</i>	Surface plasmon resonance
<i>SR1</i>	StemRegenin 1
<i>T</i>	Triplet
<i>TAMRA</i>	tétraméthylrhodamine
<i>TBAF</i>	Fluorure de tétra-n-butylammonium
<i>tBu</i>	tert-Butyl
<i>TCEP</i>	Tris(2-carboxyethyl)phosphine
<i>Td</i>	Triplet de doublet
<i>Tdd</i>	Triplet de doublet de doublet
<i>TFA</i>	Acide trifluoroacétique (Trifluoroacetic acid)
<i>TIS</i>	Triisopropyl silane
<i>TMS</i>	Triméthylsilyle
<i>Trp</i>	Tryptophane
<i>Ts</i>	Tosyl
<i>Tyr</i>	Tyrosine
<i>Uv</i>	Ultraviolet

REMERCIEMENTS

Je me considère chanceux et privilégié d'avoir pu réaliser mes études graduées à l'IRIC pour la qualité des scientifiques qui la constitue et la passion qui y est imprégnée. La possibilité d'avoir pu faire une recherche aussi diversifiée a été possible grâce à tous ces scientifiques qui ont généreusement pris le temps de m'expliquer une partie de leur connaissance. Il m'est très difficile de rester succinct dans les remerciements compte tenu du nombre important de personnes qui ont rendu possible ce présent mémoire.

J'attribue une grande partie de la réussite de ce mémoire à la merveilleuse relation que j'ai pu avoir avec mes directeurs de recherche. Tout d'abord, je souhaite remercier chaleureusement Pierre et Anne de m'avoir ouvert grands les bras pour effectuer mes études graduées. Je vais toujours me souvenir de nos discussions sur les projets où j'ai pu voir les scientifiques passionnés que vous êtes. Outre le plan scientifique, vous êtes des mentors au grand cœur, qui avez toujours été motivés par mes avancées et ma réussite. Sur une ligne plus personnelle, je vous remercie Pierre de m'avoir appris à critiquer les résultats et de toujours voir le positif d'une expérience. Je me considère choyé d'avoir bénéficié de votre immense savoir et de votre bienveillance. Anne, je ne saurais jamais vous remercier assez pour votre énergie contagieuse, vous avez toujours été intéressée par les projets et votre empathie est sans égale.

Par la suite, je tiens à envoyer mes plus sincères remerciements au Pre Joelle Pelletier et au Pr Stephen Hanessian qui ont généreusement pris le temps de corriger ce mémoire. C'est un privilège de pouvoir recevoir des commentaires de sommités tels que vous. J'espère de tout cœur que la lecture de ce long manuscrit aura su vous divertir.

Sans un soutien financier, je n'aurais pas pu me permettre de me consacrer pleinement sur ma recherche. Je tiens ainsi à dire un énorme merci aux gouvernements du Canada et du Québec ainsi qu'à l'IRIC, qui ont subventionné ce mémoire et m'ont permis d'avoir la conscience tranquille.

J'ai eu la chance d'être accompagné par des scientifiques d'expériences pour chacun de mes projets, pour qui je suis redevable. Je commence mes remerciements avec l'équipe du projet RAS. C'est avec vous que j'ai appris à me faire confiance et à déployer mes ailes dans le monde de la recherche. Je ne pourrais jamais assez vous remercier pour votre patience et votre soutien dans ce projet. Je pense à vous Ryan, Doris, Alexandre, Amal, Tarun, Marc et Matthew. Je tiens à remercier grandement Hugo et Marc-André pour

m'avoir transmis votre savoir-faire et surtout votre passion. Ça a été un honneur de participer avec vous sur ce projet qui a été stimulant du début à la fin. Marc-André, tu m'as inspiré avec ta patience méthodique et ton organisation hors pair.

Quand je pense au deuxième projet, je pense tout de suite à vous Simon, Jalila, Guy, Éric, Rodriguo, Stéphane, Sophie et Yves. J'ai tellement appris en vous côtoyant ! Jalila et Simon, c'est en essayant de faire de la biologie que j'ai compris à quel point vous êtes des scientifiques extraordinaires. Votre générosité et votre présence sympathique a été un catalyseur pour moi dans ce projet complexe. Rodriguo, Stéphane et Yves, mille merci pour vos commentaires et vos enseignements en synthèse organique.

Une maîtrise peut être 3 ans de douleur ou de plaisir. Sans la présence de collègues qui offrent une atmosphère stimulante, on peut vite être découragé par les échecs constants qui surviennent en science. Je suis éternellement reconnaissant d'avoir pu partager le quotidien avec des collègues en or avec qui la science est restée une partie de plaisir. À mes collègues de synthèse, Éric, Kham, Patrick C., Patrick B., Louis, Guillaume, Fabrice, Simon, Nicole, Gaoqiang, Serge et Laurence, mille merci ! Je ne vous ai pas oublié Yves, Sasmita, Jean-François et Bruno ! Je tenais aussi à vous remercier plus personnellement un à un, car vous avez eu un impact important sur moi. Ma chère Sasmita, je ne pourrais jamais être assez redevable de ta bienveillance à mon égard, tu as toujours eu des bons mots pour moi et pris le temps de m'écouter et prendre de mes nouvelles. Yves, de tes anecdotes farfelues aux apprentissages de la vie, tu as su égayer mes diners quotidiens. Tu es un modèle familial pour moi, père de famille et scientifique hors pair, merci d'être une inspiration. Jean-François, j'ai eu une chance immense de te côtoyer pour tes dernières années de chimistes, tu es un battant qui fonce à contre sens. Tu m'as fait grandir énormément par tes valeurs humaines et ta philosophie de la vie. Je termine par toi mon cher Bruno, pour qui je ne saurais jamais aussi reconnaissant envers la vie de m'avoir fait te rencontrer. Tu es un modèle, une inspiration, un ami et un confident. Je n'ai jamais côtoyé un chimiste aussi talentueux et travaillant que toi. Si je me débrouille en chimie aujourd'hui c'est en grande partie grâce à toi, qui es une encyclopédie sur pattes ! Malgré notre différence d'âge tu ne m'as jamais pris de haut, tu as toujours pris le temps de m'écouter et de m'expliquer les rudiments du métier. Dans les moments plus difficiles, tu as été là et j'en serais toujours reconnaissant. Mille mercis Bruno !

Je dois prendre le temps de remercier maintenant ma deuxième famille de maîtrise, le groupe de protéomique. J'ai eu la chance de rencontrer des amis qui resteront pour bien longtemps. Je vous remercie Chantal, Joël, Mirela, Mathieu, Trent, Simon, Jenna, Francis, Chongyang, Zhaoguan, Charles, Peter, Robin,

Clémence, Éric et Sibylle. Simon, merci de m'avoir fait garder la forme avec ton énergie débordante et ton humour. Mille mercis Mirela pour ta générosité et ton esprit maternel. Je pense à toi Zhaoguan, mon voisin dangereux, avec qui j'ai passé de très bons moments. Éric, tes blagues et ta bonne humeur ont egayé mon quotidien et je ne pourrais jamais te remercier assez pour tes conseils précieux. Je souhaite dire un merci particulier à Sibylle pour m'avoir introduit dans le groupe avec ton côté rassembleur. Ta générosité et ton écoute font de toi une amie précieuse. Merci mille fois à toi aussi, Chongyang, tu es un être d'exception. Ta détermination sans failles est un modèle à suivre! Tu as été un confident précieux et un ami cher.

Je souhaite continuer mes remerciements avec deux étudiants passionnés que j'ai eu la chance de côtoyer. Tout d'abord, merci, Francis, pour ta curiosité et ton implication pour le court stage que tu as fait. Ça a été un plaisir de t'expliquer les rudiments de la chimie organique. Maintenant, Moïse, quel merveilleux étudiant tu as été. J'ai eu des fous rires en ta compagnie et je garde plein de bons souvenirs. Tu m'as démontré ta positivité et ton acharnement au travail pour laquelle je ne te remercierais jamais assez, ah oui ces diazirines !!! Tu vas être un excellent chimiste je n'ai aucun doute!

Je souhaite conclure ces remerciements avec les êtres dans l'ombre qui font partie de mon quotidien et qui sont d'une importance capitale pour moi. Je parle bien sûr de mes parents Line et Jean-Luc, de mon frère Jean-Olivier, de ma grand-maman Laurette et de ma copine Nadège. Vous avez su endurer mes hauts et mes bas tout en me remontant le moral. Je suis éternellement reconnaissant pour votre aide et votre amour. Merci mon frère, de m'avoir supporté tout au long et d'avoir été un correcteur exemplaire. Merci mon amour, de partager ta vie avec moi et de me soutenir au quotidien.

1. Introduction

1.1. Préface - Les outils moléculaires

Les progrès scientifiques ont évolué à vitesse phénoménale depuis les années 2000, grâce à une transmission accélérée des connaissances scientifiques et la venue des nouvelles technologies. Notamment, internet a permis d'offrir une vitrine merveilleuse pour la diffusion des résultats et surtout pour les collaborations entre les chercheurs. La science est dorénavant poussée dans des questionnements excessivement pointus nécessitant une collaboration multidisciplinaire pour répondre et comprendre les résultats expérimentaux. Aujourd'hui, il est possible d'enregistrer son poème préféré dans un ordinateur à ADN¹, où même de prédire si une planète se trouvant à des années lumières possède une atmosphère pouvant y héberger la vie selon une analyse des raies d'absorption atomiques des éléments qui la compose².

Le corps humain a toujours fasciné les scientifiques et reste à ce jour un casse-tête inachevé. L'avenue des sphères dites **omiques** regroupant la génomique, la protéomique, la métabolomique et la transcriptomique a fourni une quantité monumentale d'informations sur les systèmes biologiques du corps humain.³ Ceci a catalysé le développement de nouvelles technologies pouvant interroger ses systèmes biologiques. On notera par exemple le traitement d'une maladie, la compréhension d'un mécanisme cellulaire, le rôle de telle ou telle protéine, la localisation des constituants cellulaires ou même les mécanismes compensatoires. Ainsi, une nouvelle science nommée la chimie biologique, a été grandement sollicitée ces dernières années pour répondre aux nouvelles questions provenant des **omiques**.⁴ La chimie biologique met de l'avant la synthèse organique d'outils moléculaires permettant d'aller explorer les milieux cellulaires. Bien qu'il existe une grande quantité de méthodes génétiques pouvant répondre à ces questions, ces dernières ne permettent pas de sonder le milieu biologique dans son état natif. Cette conséquence peut occasionner des discordances à cause de la réalité faussée des méthodes génétiques.⁵ Ainsi les composés outils développés par la chimie biologique permettent d'offrir une nouvelle approche pour approfondir les connaissances du corps humain. Ce présent mémoire va porter sur l'élaboration de deux types d'outils chimiques allant de leur conception à leur application. Dans un premier cas, il s'agira d'une synthèse de macrocycles covalents ayant pour objectif d'obtenir des informations structurales sur la GTPase RAS dans l'optique de développement d'un

médicament. Tandis que dans le second cas, il s'agira de développement de sondes ayant comme objectif l'identification d'une cible protéique inconnue ayant des propriétés sur l'expansion de cellules souches hématopoïétiques.

1.2. Présentation du projet des macrocycles covalents

1.2.1 Importance de RAS dans l'oncologie

Afin de débiter l'explication du premier projet il est important de se demander qu'est-ce que le cancer ? Le cancer peut être défini de manière simpliste comme un dérèglement cellulaire causant une prolifération non contrôlée menant à la formation de masses bloquant ou limitant des fonctions essentielles du corps humain. Ces dérèglements, d'ordres génétiques et épigénétiques, causent des mutations chez les protéines régissant divers mécanismes qui permettent normalement l'homéostasie cellulaire.⁶ Dans les cellules cancéreuses, il est possible d'observer un dérèglement de divers mécanismes cellulaires tels que la prolifération, la survie (résistance à l'apoptose), la différenciation et la motilité.⁶ Ces divers mécanismes cellulaires sont régis par plusieurs voies de signalisation, dont la voie MAPK/ERK.⁷ Par conséquent, une hyperactivation de cette voie conduit à la transformation oncologique et la progression de divers type de cancers. Afin de comprendre l'approche thérapeutique décrite dans ce présent mémoire, il est nécessaire de présenter cette voie de signalisation (Figure 1)⁷.

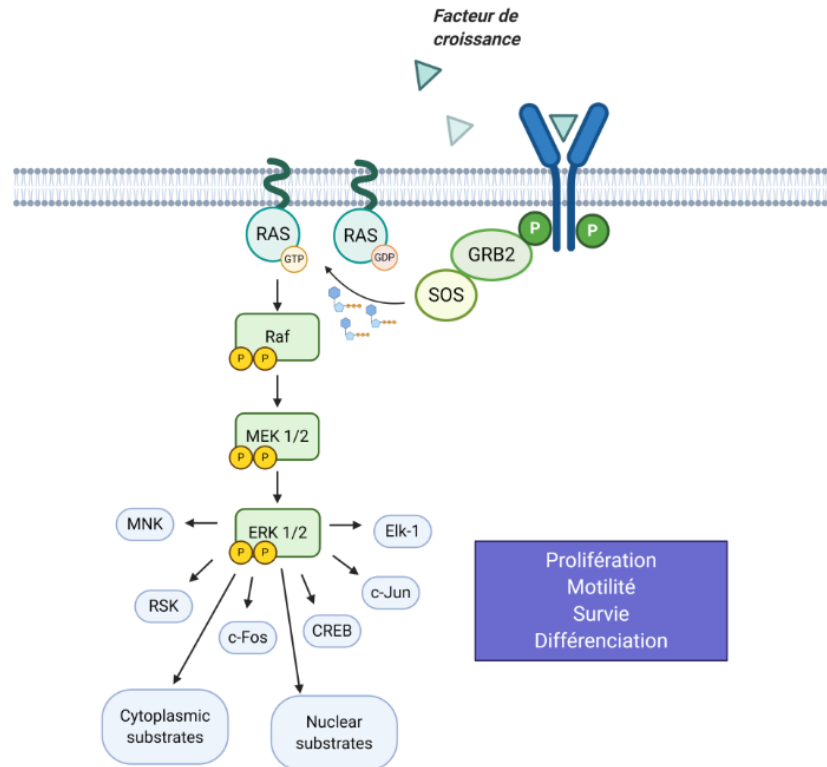


Figure 1 : Voie de signalisation de la voie cellulaire MAPK/ERK. Suite à l'activation des récepteurs tyrosines kinases, il y a autophosphorylation de ces derniers, ce qui permet la liaison de GRB2, qui active SOS et ensuite permet le transfert de nucléotides dans RAS. Une fois RAS-GTP obtenu, une cascade de phosphorylation survient allant de RAS à RAF à MEK et pour finir par ERK. ERK phosphoryle des facteurs de transcriptions (ex. RSK, c-Jun, C-Fos, MNK, CRB, Elk-1). Figure adaptée de la référence (7).

La première étape de ce mécanisme cellulaire est l'activation du récepteur à activité tyrosine kinase (RTK), présent à la membrane cellulaire, par diverses entités telles que les hormones, les chimiokines ou les facteurs de croissances.⁷ L'activation du récepteur se conclut par une phosphorylation des tyrosines qui sont présentes à l'extrémité des deux chaînes C-terminales existantes dans le domaine intracellulaire du récepteur⁸. La seconde étape de la voie MAPK/ERK est la reconnaissance de ces sites phosphorylés par des effecteurs. Le domaine SH2 ou PTB généré par la réponse du récepteur RTK permet un point d'ancrage pour des protéines adaptatrices telles que SHC et GRB2. Ces dernières ont pour objectif de venir recruter la protéine SOS de type GEF dans le cytoplasme afin de l'approcher de la membrane plasmique. Le rôle de SOS est très important : elle va se

lier à RAS (il existe 4 isoformes : NRAS, KRAS4A, KRAS4B et HRAS), une GTPase se trouvant à la membrane plasmique attachée par des chaînes farnésyles ou palmitoyles, pour introduire le GTP. Cette interaction occasionne un changement conformationnel de RAS permettant de libérer le GDP du site actif pour laisser entrer un GTP libre dans la cellule. On dit de RAS que la forme GDP est inactive, tandis que la forme GTP est active. Par conséquent, cette forme active de RAS a la possibilité de se lier à une multitude d'effecteurs. Parmi les 20 effecteurs identifiés confirmés dans la littérature, 3 ont retenu l'attention des chercheurs dans le domaine de l'oncologie : les protéines RAF (il existe 3 isoformes : ARAF, BRAF et CRAF), les protéines RALGDS et les protéines PI3K⁹. Les protéines RALGDS sont impliquées dans l'organisation cytosquelettique et la transformation cellulaire, tandis que RAF et PI3K sont impliquées dans les mécanismes suivants : prolifération cellulaire, survie, différenciation et motilité. En continuant avec la voie MAPK/ERK, RAS-GTP aura comme rôle de recruter les kinases RAF toujours à la membrane et occasionnera l'activation du site catalytique de cette dernière. Le reste de cette voie de signalisation consiste en une cascade de phosphorylations successives entre des kinases, débutant par la phosphorylation des MAPKKK RAF pour les MAPKK (MEK1 et MEK2), qui iront activer des MAPK (ERK1 et ERK2). Lorsque les kinases ERK1 et ERK2 sont activées, elles permettent la phosphorylation d'une multitude de substrats nucléaires et cytoplasmiques (plus de 150 identifiés) impliqués dans diverses réponses cellulaires dont RSK, MNK, Elk-1, CREB, C-Fos et c-Jun.¹⁰

Chaque voie de signalisation possède des signaux et des moyens pour le retour à l'état initial une fois la réponse cellulaire obtenue. Il existe deux classes de réponses : la réponse post-traductionnelle et l'induction de gènes *de novo* qui inhibent spécifiquement la signalisation. La différence majeure entre ces deux types de réponses est le temps nécessaire pour obtenir l'effet. La réponse post-traductionnelle est quasiment instantanée alors que la réponse génétique nécessite la traduction de protéines par le ribosome. Dans le cas d'une réponse d'induction de gènes, une protéine, GAP, est exprimée et permet d'hydrolyser le GTP présent dans RAS pour ainsi former l'espèce inactive RAS-GDP (Figure 2).¹¹ Il existe un équilibre cyclique entre la SOS, qui effectue le

transfert du GDP au GTP, et la GAP, qui accélère l'hydrolyse du GTP par un facteur supérieur à 1000 fois comparativement à l'hydrolyse non catalysée du complexe HRAS-GTP. Une fois la perte de ce phosphate, RAS se trouve à être inactif et ne se lie plus aux effecteurs.

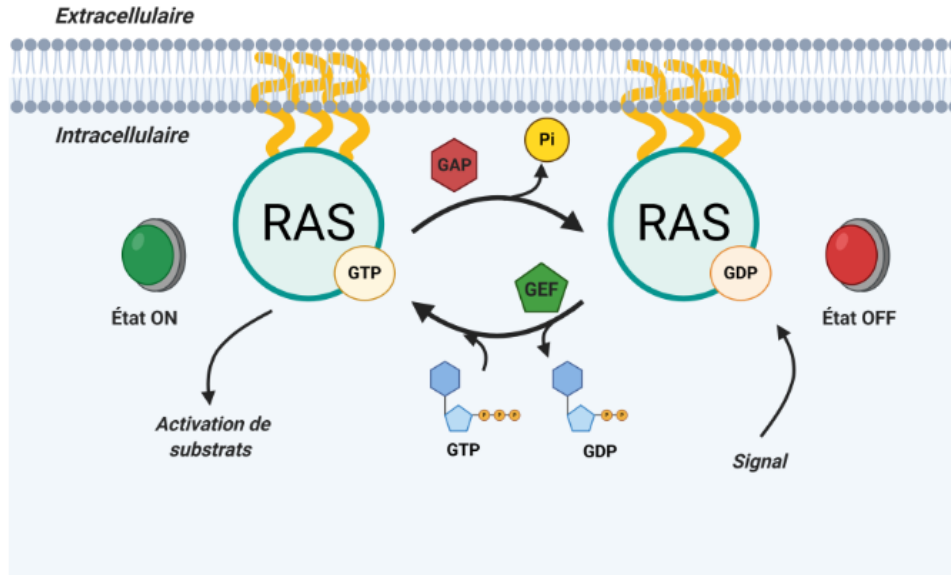


Figure 2 : Illustration du cycle de régulation de l'activité des RAS. La voie de signalisation est régie par le nucléotide lié à RAS. La forme RAS-GDP se trouve inactive et est activée via l'interaction de SOS qui facilite l'échange du GDP au GTP, tandis que la GAP favorise l'hydrolyse du GTP via l'accessibilité du GTP pour les GTPases. Figure adaptée de la référence 11.

Puisque la voie de signalisation MAPK/ERK est impliquée dans une multitude de réponses cellulaires telles que la prolifération et l'inhibition de l'apoptose, une altération de ces mécanismes de régulation négative occasionne un cancer. À la suite de grandes analyses statistiques génétiques et protéomiques de patients atteints de cancer, plusieurs acteurs de la voie de signalisation MAPK/ERK présentaient des mutations responsables d'une hyperactivation de la voie de signalisation. La figure 3 illustre ces principales mutations et les types de cancers où l'on retrouve ces mutations. ⁶

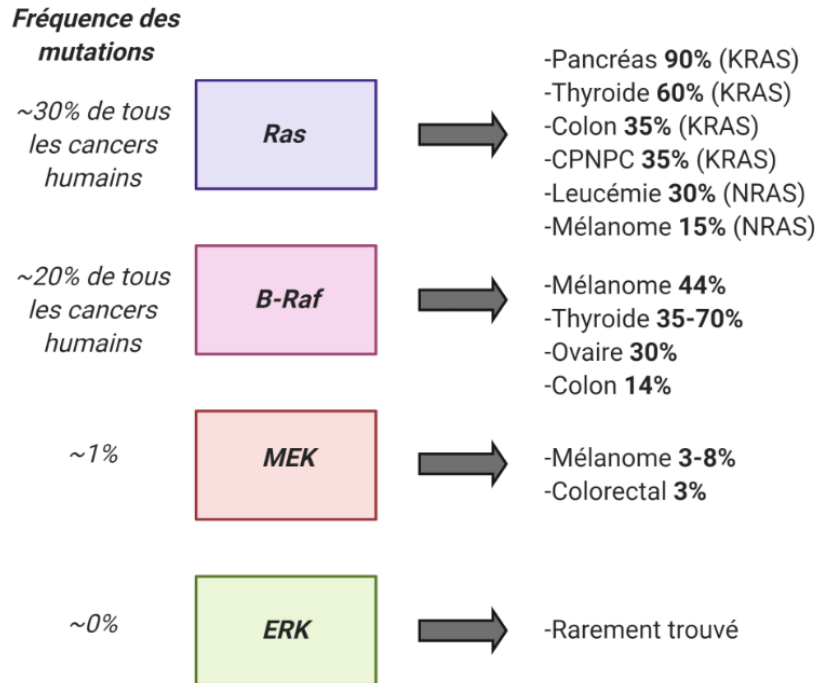


Figure 3 : Représentation de la proportion des mutations dans les différents acteurs de la voie de signalisation MAPK/ERK parmi les cancers répertoriés. KRAS, NRAS et B-RAF sont les principaux mutés, et ce dans divers cancers tels que le pancréas, la thyroïde, le colon et les mélanomes. Figure adaptée de la référence 6.

La plus grande proportion des cancers liés au dérèglement de la voie MAPK/ERK sont causés par des mutations sur RAS et RAF. Il s'avère que des mutations sont majoritairement observées avec les isoformes NRAS et KRAS. Si l'on répartit l'abondance des mutations oncologiques des isoformes de RAS dans les cancers, on retrouve 85 % de mutants pour KRAS, 12 % pour NRAS et moins de 3% pour HRAS.¹² On retrouve une mutation de NRAS majoritairement dans les mélanomes, une mutation de HRAS principalement dans les glandes surrénales et KRAS est mutée le plus souvent dans le pancréas. Une mutation de RAS est si fréquente que près du tiers des tumeurs solides humaines présentent ce type d'anomalie.¹² Dans les GTPases RAS, il a été constaté que la plupart des mutations se trouvaient aux acides aminés 12, 13 et 61. Toutes ces mutations causent un encombrement stérique du "arginine finger" interagissant avec la GAP,

rendant ainsi l'hydrolyse du GTP beaucoup plus difficile.¹² Ceci a pour conséquence de maintenir RAS sous sa forme activée.

Ces chiffres illustrent la grande place qu'occupe la voie de signalisation MAPK/ERK dans le domaine de l'oncologie. Il s'agit d'une importante préoccupation qui mobilise une grande proportion du milieu scientifique pour la recherche de traitements. D'un point de vue de la chimie médicinale, plusieurs stratégies ont été investiguées pour tenter d'inhiber cette voie oncologique. D'un point de vue logique, l'idée la plus simple est de bloquer l'interaction d'une protéine avec son effecteur ou empêcher une protéine de remplir son rôle en bloquant son site catalytique, arrêtant ainsi la suite de la cascade de signalisation. On pourrait ainsi penser à bloquer chaque protéine impliquée dans la voie MAPK/ERK, donc RTK, SOS, RAS, RAF, MEK ou ERK. Cette stratégie a été menée sur plusieurs substrats et a donné naissance à plusieurs médicaments.¹² Cependant ces derniers sont pour la plupart sujets à de la résistance mutationnelle ou à l'activation de voies de signalement parallèles limitant leur effet avec le temps. Pour ce faire il est nécessaire d'imaginer de nouvelles voies de recherche et l'une de ces voies est de bloquer la GTPase RAS.

1.2.2 Caractéristiques structurales et isoformes de RAS

Cette petite protéine est un casse-tête pour la communauté scientifique. Après 3 décennies d'efforts pour développer un candidat clinique, il n'y a pas encore un inhibiteur qui a su produire des résultats convaincants.¹³ La possibilité d'inhiber RAS serait phénoménale pour le traitement du cancer. La position centrale de RAS au sein des voies de signalisation, due à son interaction avec de multiples effecteurs importants tels que RAL et PI3K, fait en sorte que son inhibition permettrait de contrôler plusieurs voies parallèles et de limiter les mécanismes de résistance^{14, 15}. De plus, cette protéine étant mutée dans près de 30 % des cancers, il serait intéressant de pouvoir moduler directement son activité afin de comprendre son fonctionnement et d'offrir de nouvelles

approches thérapeutiques. C'est précisément la raison et la pertinence de ce mémoire : chercher à développer une nouvelle approche thérapeutique sur la GTPase RAS.

Afin de présenter la voie d'inhibition de ce mémoire, il faut se pencher sur le mécanisme des GTPases qui sont la famille RAS. Il existe trois gènes canoniques distincts (Kras, Nras et Hras) qui encodent les protéines de la famille RAS qui sont KRAS4A, KRAS4B, HRAS et NRAS.¹⁶ La protéine KRAS4B est prédominante dans l'épissage du gène Kras. Les séquences ainsi que la structure tridimensionnelle des protéines sont illustrées à la figure 4.¹⁶

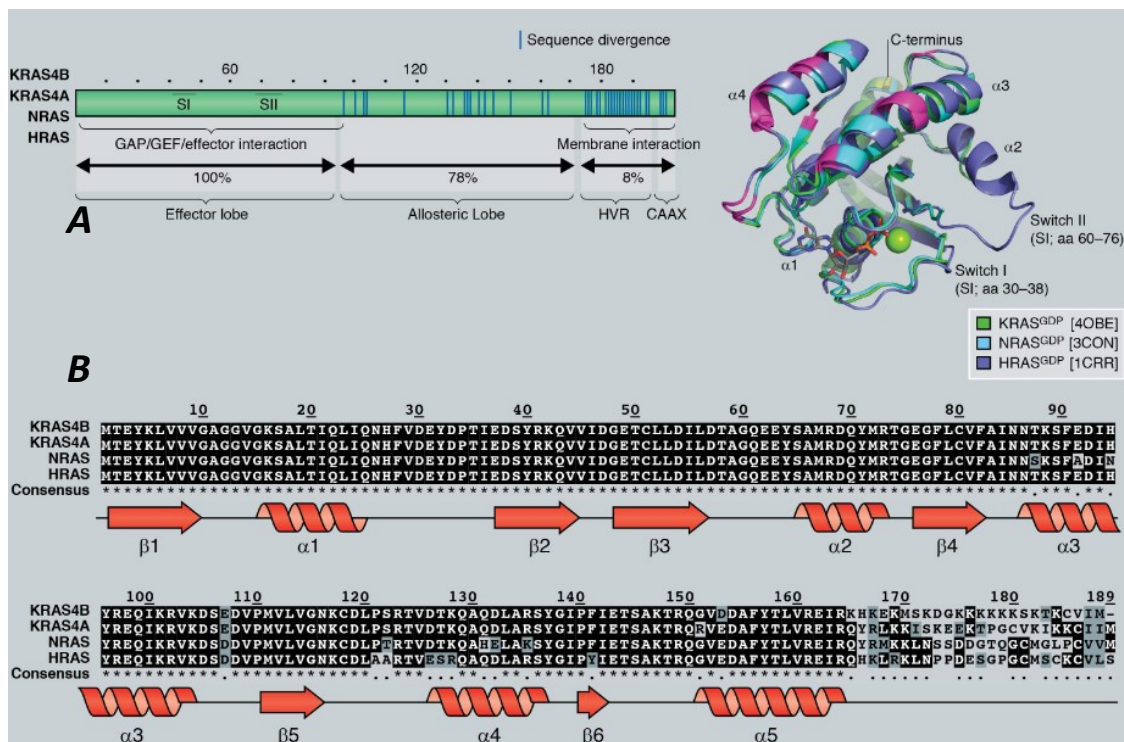


Figure 4 : Caractéristiques structurales des différentes isoformes de RAS. **A)** illustre les différentes régions de la GTPase en fonction de propriétés fonctionnelles. On retrouve le lobe effecteur, le lobe allostérique et la zone d'hypervariabilité. **B)** illustre les séquences respectives des différentes isoformes ainsi que les différents types de structures secondaires présentes dans ces derniers. Figure provenant de la référence 17.

Les protéines RAS se décomposent en trois sections distinctes : le lobe effecteur, le lobe allostérique et la zone d'hypervariabilité. On retrouve 6 feuilletts bêta et 5 hélices alpha dans la protéine, distribuées entre les deux lobes tel qu'illustré par la figure 4A.¹⁶ Le lobe effecteur est composé de deux interrupteurs que sont SI (aa 30-38) et SII (aa 60-76). Tous les effecteurs interagissent sur le lobe effecteur, d'où son nom. Il est nécessaire que le GTP se trouve dans le site actif, afin de produire l'orientation nécessaire pour interagir avec les effecteurs : pour cette raison, on dit de RAS qu'il s'agit d'un interrupteur moléculaire. Le lobe effecteur possède une homologie de 100 % pour toutes les isoformes de RAS. La section de variance entre les isoformes s'observe sur le lobe allostérique et la zone d'hypervariabilité. On observe une homologie des résidus d'uniquement 78 % pour le lobe allostérique et 8 % pour la zone d'hypervariabilité.¹⁶ Finalement, la région hypervariable est une région de forte modification post-traductionnelle qui est nécessaire à la lipitation de RAS, étape primordiale pour que RAS puisse se greffer à la membrane plasmique.¹⁶ Si l'on regarde plus précisément la zone hypervariable des isoformes de RAS, on peut constater des différences importantes. Notamment l'emplacement des cystéines et leur nombre qui donne la sélectivité d'emplacement des isoformes dans les cellules. Ce sont ces cystéines qui subissent la lipitation et permettent de donner le caractère hydrophobe nécessaire pour venir s'accrocher avec la membrane plasmique.

Une fois ancrée à la membrane via la chaîne lipidique, RAS se voit activée par la GEF, SOS. RAS étant instable sans la présence d'un nucléotide dans son site actif, SOS permet de stabiliser RAS dans un état sans nucléotide. Pour ce faire, une hélice alpha provenant de SOS vient s'insérer entre les régions SI et SII. Un mouvement semblable à celui d'un abaisse-langue se produit en ouvrant le site actif du nucléotide permettant son expulsion, lors du déplacement de l'interrupteur SI. (Figure 5)¹⁷ Dans cet état, l'affinité du GTP et GDP est très similaire pour le site actif (pM)⁹. Cependant, comme la concentration libre de GTP dans le milieu cellulaire est largement supérieure (10 fois plus) à celle de GDP, le GTP a statistiquement plus de chance de se retrouver dans le site actif.^{9, 17}

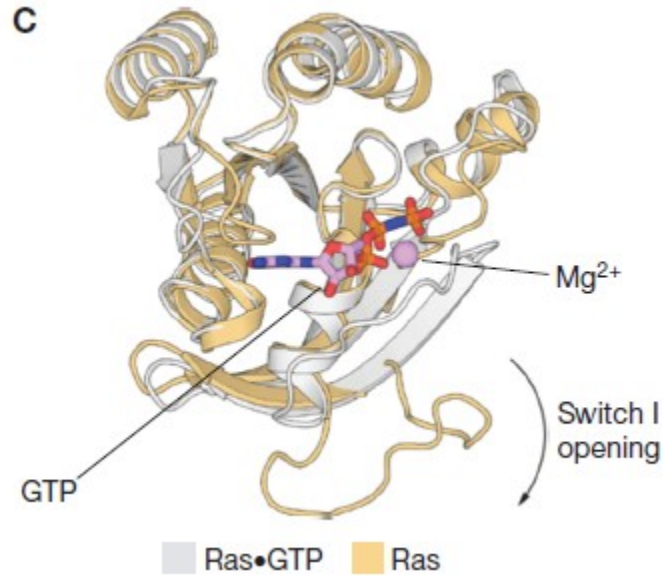


Figure 5 : Effet conformationnel de RAS lors de l'activation de SOS. Lors de l'insertion de l'hélice alpha de SOS dans RAS, il y a ouverture de l'interrupteur 1 de RAS permettant l'échange du nucléotide GDP pour le GTP. En beige est illustré la GTPase RAS en liaison avec SOS et en gris la protéine liée au GTP. Figure provenant de la référence 19.

Comme l'interaction entre SOS et RAS est dynamique, lorsque SOS va se dissocier de RAS, l'hélice alpha va se retirer, le site actif va se refermer et interagir avec le nouveau nucléotide inséré, le GTP. Une fois RAS muni d'un GTP, une interaction avec RAF est possible. À ce jour, il est clair que RAS-GTP interagit avec RAF, bien que le mécanisme de cette interaction reste un grand débat. Il est toutefois reconnu que la dimérisation de RAF est une étape essentielle pour interagir avec MEK.^{18, 19} Les expériences structurales ont démontré que RAF dimérisait selon une interface côté-côté.¹⁹ Cependant, ce qui fait débat, c'est la manière par laquelle l'activation de RAF se produit à la membrane par RAS.²⁰ Deux scénarios sont envisagés et sont illustrés sur le schéma 6.²¹

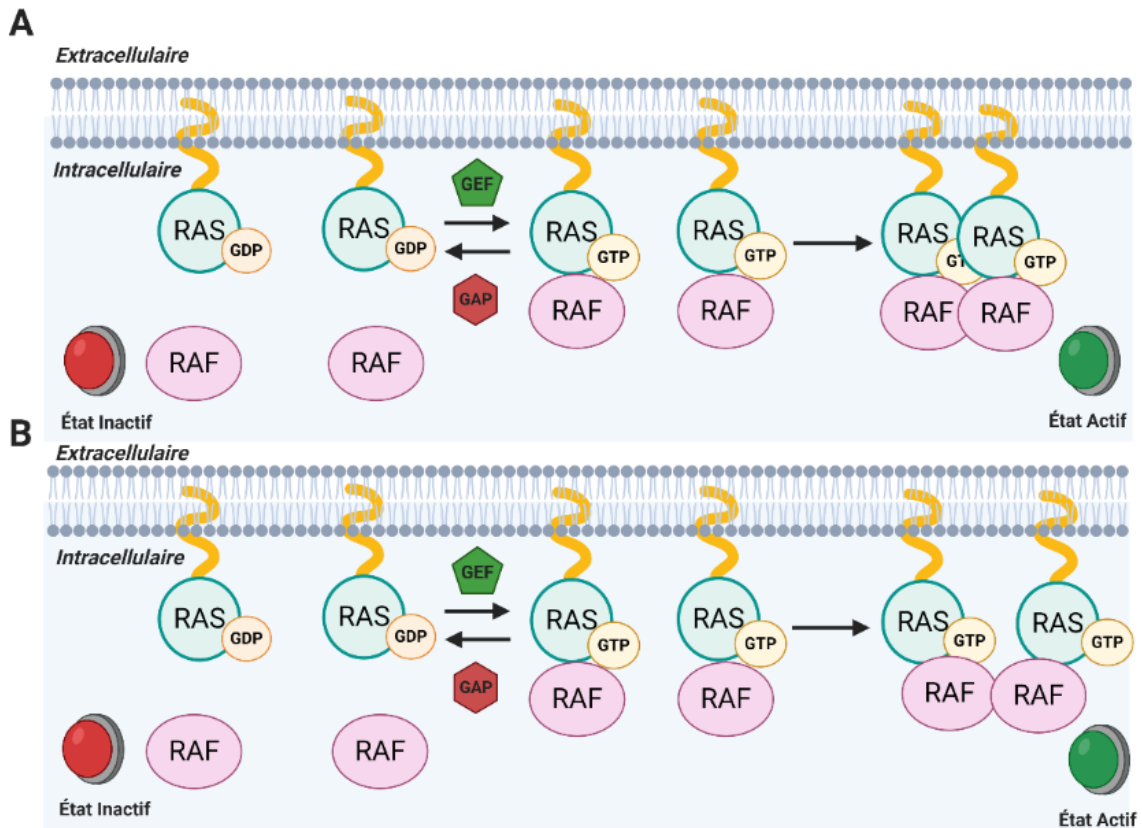


Figure 6 : Mécanismes proposés pour l'activation dimérique de RAF à la membrane par RAS. Une fois RAS se trouvant dans sa conformation active GTP, elle doit activer RAF qui doit être dimérisé. Il est hypothétisé une interaction entre RAS et RAF formant deux complexes monomériques qui ensuite se dimérise en A soit via une interaction entre les RAS ou en B via RAF se trouvant à proximité. Figure adaptée de la référence 23.

D'après le mécanisme A, il y aurait formation d'un dimère entre deux GTPase RAS activées, ce qui permettrait la dimérisation de RAF. Le second cas de figure serait une activation de deux monomères RAS-GTP sur deux protéines RAF, qui pourront ensuite dimériser lorsque situées l'une près de l'autre. Avant d'élaborer la stratégie thérapeutique qui a éclos de cette découverte et dont il est question dans ce mémoire, survolons les recherches d'inhibition de RAS dans la littérature.

1.2.3 Approches de la littérature pour l'inhibition de RAS

Il n'est pas anodin qu'on nomme RAS comme le saint Graal : développer une inhibition de cette protéine serait grandement bénéfique pour le traitement du cancer, compte tenu de la grande place occupée par les mutations de cette GTPase en l'oncologie. Cependant, cette cible thérapeutique est difficile à atteindre et d'innombrables échecs sont survenus au cours de 3 décennies de recherches. Plusieurs raisons expliquent la complexité du développement d'un inhibiteur de RAS. Tout d'abord, RAS possède une structure tertiaire lisse et souple, rendant difficile d'interagir dans une poche.¹³ De plus, son seul site où l'on retrouve une zone possible d'interaction pour une petite molécule se trouve à être son site actif, où se trouve le GTP. Ces difficultés ont grandement ralenti la recherche d'avenues thérapeutiques de cette GTPase, sans empêcher le développement de nouvelles approches alimentées par l'avancement des connaissances et des technologies. La figure 7 illustre les 3 stratégies thérapeutiques récentes dans lesquelles on a tenté l'inhibition de RAS mutée..²²

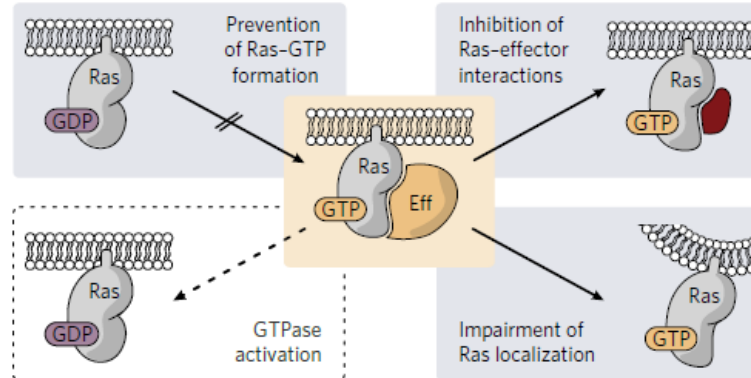
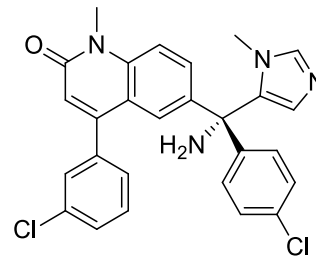
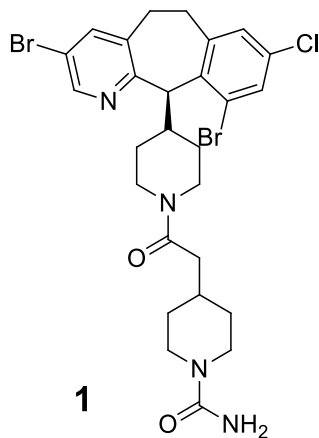


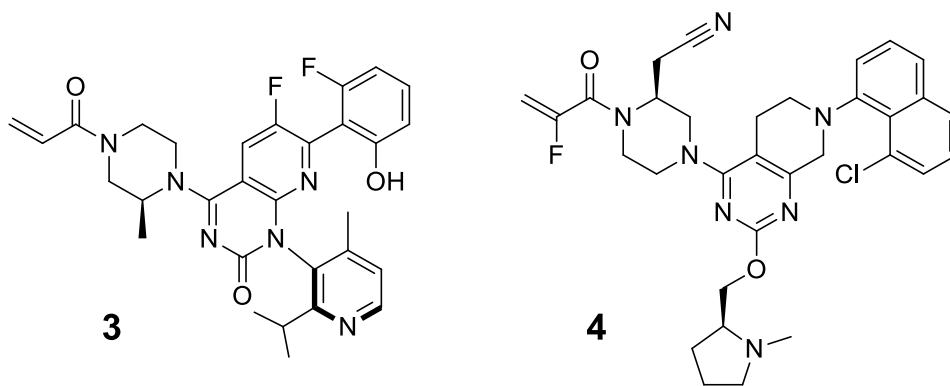
Figure 7 : Résumé des diverses stratégies thérapeutiques étudiées pour inhiber RAS oncogène. Les trois stratégies évaluées sont la prévention de l'activation de RAS-GTP, l'inhibition des interactions de RAS-GTP pour ses effecteurs et l'entrave de RAS durant son transport vers la membrane. Figure provenant de la référence 26.

De façon classique, l'un des moyens d'inhiber une kinase est de venir compétitionner l'ATP dans le site actif avec un inhibiteur ayant une affinité supérieure. L'affinité de l'ATP pour les kinases se situe en moyenne dans l'ordre du micromolaire. Pour obtenir une équivalence avec RAS et son substrat GTP, qui possède une affinité de l'ordre du picomolaire (10 pM), il serait nécessaire d'avoir un inhibiteur dans le femtomolaire²³. Or, ceci relève de l'impossible quand on sait que l'interaction non covalente la plus forte identifiée est de 100 fM (streptavidin pour la biotine) : cette piste de solution n'est donc pas envisageable. L'une des méthodes alternatives consiste à venir bloquer la lipitation des cystéines présentes sur la chaîne d'hypervariabilité. Comme il faut que RAS se trouve à la membrane plasmique pour interagir avec les effecteurs, il faut qu'elle puisse s'ancrer à la membrane via ces lipides. Ainsi l'idée des chercheurs a été de développer des inhibiteurs des diverses enzymes responsables d'effectuer les modifications lipidiques sur les cystéines. Parmi ces dernières, deux enzymes sont importantes : la géranyltransferase I et la farnésyltransférase cytosolique.^{12, 22, 23} Ces inhibiteurs ont fini être abandonnés lors des essais cliniques, incapables de démontrer une efficacité clinique suffisante à cause de mécanismes compensatoires et des effets secondaires (Lonafamib **1**, Tipifarnib **2**)²⁴.

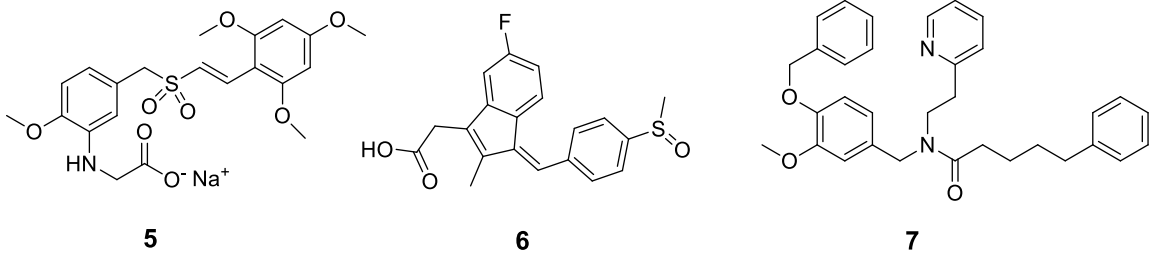


Une seconde approche consiste à développer des inhibiteurs qui permettent de favoriser la forme inactive de RAS(GDP), donc de prévenir l'activation de RAS. L'une des façons de faire est de produire un blocage allostérique de SOS, la GEF venant catalyser le transfert du GDP en GTP, en utilisant des inhibiteurs allostériques comme HBS3 ou DCAI, dont le rôle est d'empêcher l'interaction SOS-RAS et ainsi conserver le RAS-GDP.²²

Des expériences ont démontré une diminution des niveaux de phosphorylation d'ERK1/2 chez les cellules possédant des mutations oncogènes sur KRAS. Cependant, cette stratégie est encore au stade embryonnaire : elle n'a pas encore démontré sa capacité à inhiber le cancer occasionné par les mutations de RAS.²³ Une autre façon de favoriser l'espèce RAS-GDP est de bloquer le changement de conformation qui se produit en venant interagir dans un site allostérique de RAS. Le mutant KRAS G12C possède une poche allostérique accessible uniquement lorsque le GDP est lié à ce mutant KRAS. Une piste de solution consiste à utiliser la cystéine C12 pour développer une approche covalente qui modifie irréversiblement la protéine. Cette modification gêne la protéine, qui n'est plus en mesure de changer de conformation pour échanger le GDP par un GTP. Cette stratégie a présentement 4 composés en essais cliniques en phase I/II (AMG-510 **3**, MRTX849 **4**, JNJ-74699157, LY3499446)^{25, 26}.



Cependant, cette stratégie s'applique uniquement pour une branche très spécifique de cancer KRAS avec la mutation G12C, qui prend avantage de la nouvelle cystéine provenant de cette mutation. Finalement, la dernière approche thérapeutique développée fut l'inhibition des effecteurs de RAS-GTP. L'idée est simple : s'il est possible d'empêcher la réaction de RAS pour les effecteurs en aval, il est possible d'agir sur la voie de signalisation. Par exemple, l'utilisation d'un anticorps monoclonal (Y13-259) bloque le domaine de liaison des effecteurs de RAS. Cependant, cette entité ne possède pas de propriétés médicamenteuses à cause de son inhabilité à pénétrer la cellule.²⁴ Par conséquent, des petites molécules ont été développées pour inhiber l'interaction avec les effecteurs (Rigosertib **5**, Sulindac **6**, MCP110 **7**).



Les mécanismes d'action de Sulindac et MCP110 étant inconnus, une optimisation de ces composés n'a pas été possible. De plus, Sulindac **6** et MCP110 **7**, par leur manque de sélectivité et d'affinité, se sont révélés ne pas être de bons candidats pharmaceutiques. Cependant, d'autres candidats, comme Rigosertib **5**, sont plus prometteurs : en phase clinique III, cet inhibiteur agit sur RAF comme un émule de RAS. Malgré toutes ces approches, peu de composés ont démontré des propriétés intéressantes pour inhiber les isoformes de RAS oncogène et la recherche pour inhiber RAS demeure un défi actuel.

Avec les récentes découvertes sur l'importance de la dimérisation de RAF, une idée thérapeutique a été explorée par un groupe de recherche.²⁷ La problématique était la suivante : est-il possible de concevoir une technologie capable d'influencer les interactions entre RAS (monomérique/dimérique) à la membrane plasmique qui permettrait d'empêcher la dimérisation de RAF? Le groupe de Khoide a utilisé la technologie combinatoire du Phage Display afin de générer ce qu'on nomme des 'monobodies'. Les 'monobodies' sont des protéines synthétiques ayant comme construction initiale le domaine de la fibronectine III (FN3). La FN3 offre la possibilité d'offrir 3 boucles de séquences d'acides aminés hypervariables où il est possible d'effectuer des bibliothèques sans avoir de rejet cellulaire.²⁸ On peut comparer FN3 à un anticorps où les boucles sont l'analogue des zones hypervariables des anticorps qui servent à la reconnaissance. À la différence des anticorps, les 'monobodies' ont la caractéristique de ne pas être affectés par le potentiel redox de leur environnement. Cette caractéristique facilite l'utilisation de ceux-ci en tant qu'inhibiteurs génétiquement encodés.²⁷ Dans le cadre de leur recherche, le groupe de chercheurs a identifié NS1, un 'monobody' ayant la capacité de lier HRAS et KRAS avec une affinité dans l'ordre du nanomolaire. Ce dernier se lie sur le lobe allostérique de RAS, où se trouvent $\alpha 4$, $\beta 5$ et $\alpha 6$. (Figure 8)²⁷ Cependant, la grande majorité des interactions se produisent entre la boucle FG de NS1 et le feuillet $\beta 6$ de RAS. Une interaction clé est la cage formée entre R135 de RAS et les acides aminés E49 et Y71. L'arginine 135 est présente uniquement dans HRAS et KRAS, ce qui explique la sélectivité de NS1 pour ces deux isoformes.

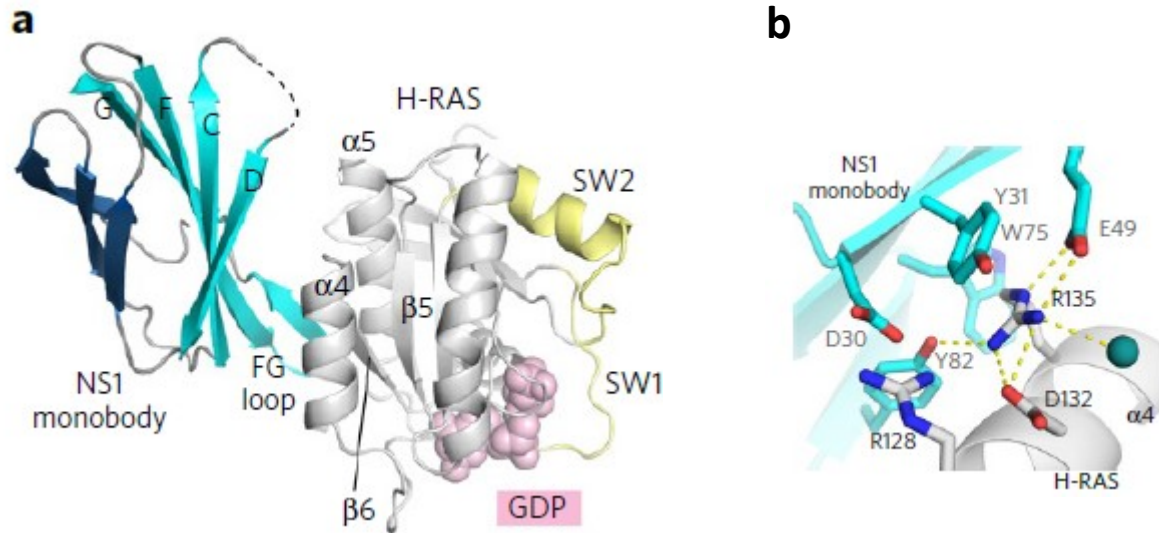


Figure 8 : Structure cristalline du site de liaison allostérique de NS1 pour HRAS. (PDB 4G0N) En a) se retrouve l'interaction du monobody NS1 (bleu) via la boucle FG sur la face des hélices alpha 4 et 5 de RAS en gris. La triade d'interactions ioniques entre la boucle de NS1 et l'arginine 135 de RAS est présentée en b). Figure provenant de la référence 30.

Ce qui fut encore plus intéressant que la capacité à lier un site allostérique de RAS fut sans contredit la capacité de NS1 à influencer la réponse biologique en aval de RAS. Il s'avère que NS1 a la capacité d'interrompre la dimérisation de RAF en gênant l'interaction des monomères de RAS adjacents qu'on pourrait trouver dans les dimères ou dans les nano regroupements monomériques.

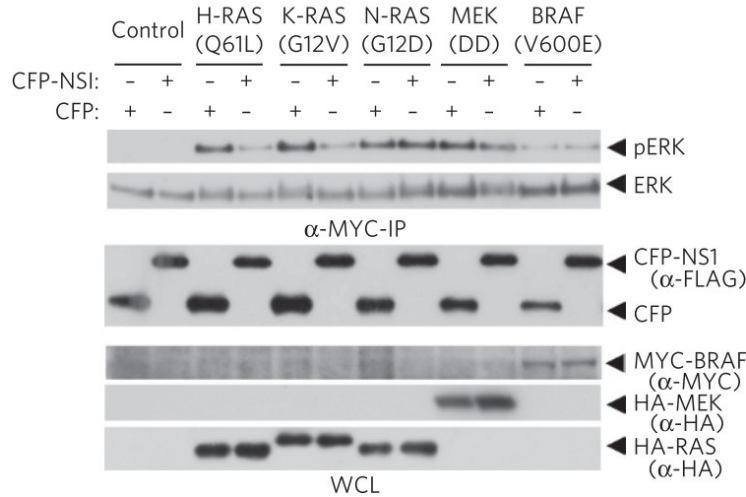


Figure 9 : Buvarbage de Western démontrant l'effet de NS1 sur la phosphorylation en aval de RAS. Ce gel a été effectué via une stimulation EGF sur cellules HEK293, où une construction NS1-GFP avait été exprimée. HRASQ61L, NRASG12D et KRASG12V sont des oncogènes connus dans la littérature et lors d'une utilisation de NS1 il est possible d'observer une diminution de phosphorylation d'ERK pour HRAS et KRAS. Une immunoprécipitation via une étiquette MYC ou HA a été réalisée pour observer les protéines cellulaires. Figure provenant de la référence 30.

Le buvardage de Western illustré à la figure 9 démontre l'effet de NS1 sur la voie de signalisation MAPK/ERK.²⁷ On peut observer que pour les protéines oncologiques HRAS(Q61L) et KRAS(G12V), il y a une diminution flagrante de la phosphorylation d'ERK en présence de NS1, ce qui n'est pas le cas de NRAS(G12D). Ce résultat est très intéressant, puisqu'il illustre une nouvelle avenue thérapeutique sur les oncogènes de RAS, cette GTPase si difficile à atteindre. L'étude de 2016 a pu démontrer par de la microscopie électronique ainsi que du BRET des preuves que le mécanisme d'action de NS1 était bel et bien d'influencer la dimérisation de RAS à la membrane. Cependant, le poids moléculaire d'un 'monobody' créé à partir de la fibronectine III est d'environ 10 kDa, ce qui rend la tâche plus compliquée si l'on souhaite développer un médicament. L'une des premières difficultés des 'monobodies' est d'accéder à RAS : ce dernier est intracellulaire alors que les 'monobodies', ne peuvent traverser la membrane cellulaire.²⁹ Un second inconvénient est la précarité de leurs propriétés pharmacocinétiques et immunogéniques. Comme les 'monobodies' possèdent uniquement des acides aminés

naturels, ils sont sensibles aux mécanismes d'hydrolyse que le corps utilise pour métaboliser les acides aminés²⁹.

1.2.4 Hypothèses des macrocycles covalents et démonstration du projet

Pour ces diverses raisons, des chercheurs de l'IRIC ont eu la brillante idée du développement d'un peptide mimétique offrant la réponse biologique de NS1 et les propriétés permettant le développement d'un médicament. Malgré la grande surface de contact que possède la protéine synthétique NS1, cette dernière interagit avec RAS avec une petite surface de contact. La figure 10 illustre les contacts qu'effectuent NS1 (turquoise) avec RAS (gris). On observe que le point d'ancrage se fait via l'interaction du feuillet β_6 et la Boucle FG de NS1. De plus, il existe des contacts entre NS1 et l'hélice alpha 4 et l'hélice alpha 5. Dans l'optique de développer un mimétique de NS1, l'approche choisie fut le développement de macrocycles peptidiques. Un macrocycle possède les mêmes interactions peptidiques que procure un 'monobody' et possède des propriétés médicamenteuses beaucoup plus intéressantes.

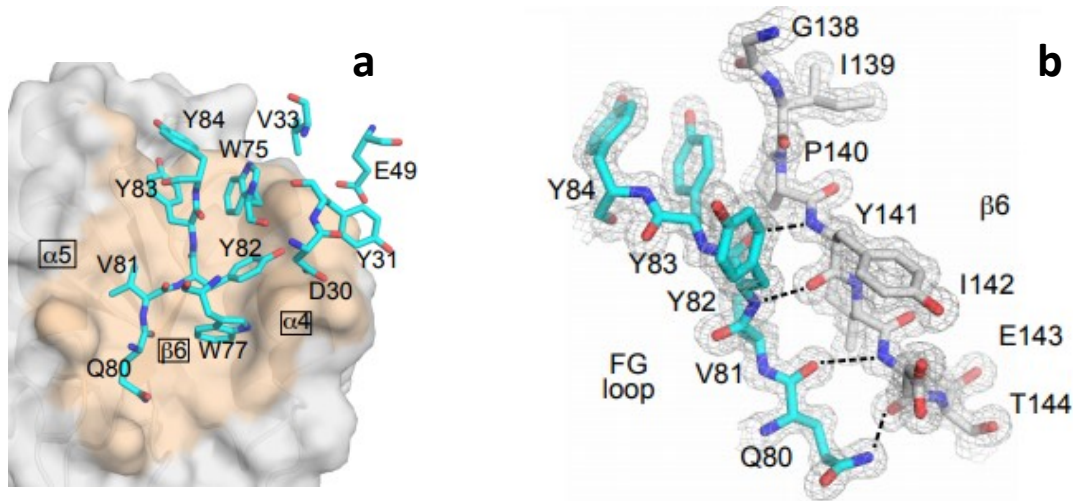


Figure 10 : Approfondissement des contacts importants de NS1 pour HRAS via la structure cristalline. En a) est illustré uniquement les résidus de NS1 se situant dans un rayon de 5Å ou moins qui interagissent avec HRAS. En b) sont illustrées les interactions importantes entre la boucle-FG et le feuillet β_6 de HRAS. Figures provenant de la référence 30.

La collaboration d'une équipe pluridisciplinaire formée de Marc-André Poupart, Hugo Lavoie, Ryan Killoran, Alexandre Beautrait, Anne Marinier, Marc Therrien et Matt Smith a permis l'approfondissement des connaissances sur les macrocycles mimétiques de NS1. Dr Poupart et Dr Beautrait ont développé la chimie de ces derniers via une modélisation informatique de la boucle FG de NS1. De plus, la synthèse de centaines de macrocycles optimisant la dimension du cycle et la composition en acides aminés a permis d'obtenir le composé le plus actif, UM0139995. L'essai qui a été utilisé pour guider le SAR a été mené par Dr Killoran et fut un essai HSQC par résonance magnétique nucléaire. Le principe général de cette technique consiste à observer les déplacements de signaux d'azotes de la protéine HRAS marquée au N¹⁵ en présence d'un inhibiteur. Comme le ligand modifie la densité électronique localement, des signaux spécifiques de la protéine seront influencés.³⁰ La variation d'intensité ou de déplacement chimique des signaux est proportionnelle à l'affinité du ligand pour la protéine, donc du macrocycle pour RAS. Les essais de mutagenèse sur NS1 menés par Dr Lavoie ont été critiques pour illustrer les acides aminés essentiels du macrocycle qui permettent une interaction avec la protéine RAS. La figure 11 illustre la composition de ce macrocycle.

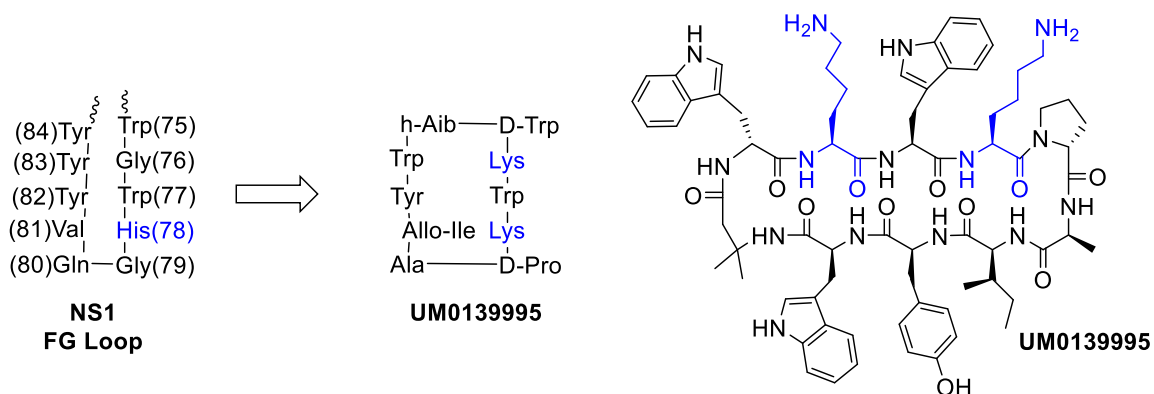


Figure 11 : Hypothèse synthétique pour mimer la boucle-FG de NS1 en macrocycle. La structure d'UM0139995 est illustrée, il s'agit du macrocycle le plus actif obtenu des bibliothèques d'optimisation (IC₅₀ = 100 µM).

Comme on peut le constater, il existe une similitude entre la séquence de la boucle FG de NS1 et le macrocycle UM0139995. L'histidine 78 et la glycine 76 furent remplacées par des lysines afin d'augmenter la solubilité du macrocycle, puisque leur emplacement s'oriente vers le solvant selon la structure cristalline de NS1. L'allo-isoleucine a augmenté l'affinité en comparaison à la valine. Le D-tryptophane se trouve à être essentiel pour obtenir l'orientation nécessaire pour effectuer le tour bêta occasionné par l'homo-Aib et pour conserver des interactions importantes avec RAS. Finalement, la D-Proline et l'homo-Aib furent utilisés afin de permettre une contrainte structurelle pour générer le tour bêta nécessaire qui imite l'épingle de la boucle FG. L'Aib est connu dans la littérature pour générer un tour bêta en présence d'un acide aminé D adjacent, ici le D-Trp, tandis que la proline (D ou L) est connue pour occasionner un tour bêta en présence d'un petit acide aminé tel qu'une alanine et une glycine.^{31, 32} Il est aussi important de mentionner que l'utilisation d'acide aminé D dans un peptide synthétique permet d'offrir une résistance à la protéolyse enzymatique, donc d'augmenter les propriétés pharmacocinétiques du composé.³² Également, l'utilisation d'un tryptophane en remplacement de la tyrosine 83 s'est révélée légèrement bénéfique. Cependant, une tyrosine à cet emplacement est aussi une bonne option, puisque le caractère donneur de ponts H est aussi présent. Toutefois, il s'est avéré que l'activité de ce macrocycle optimisé était uniquement de 100 μM , ce qui est très loin du 20 nM que possède NS1 pour HRAS. Malgré près de 200 macrocycles d'optimisation, l'affinité n'a pas pu être améliorée et le SAR se trouvait être très réticent aux changements exotiques. C'est alors que le projet de développement de composé covalent a été proposé. En observant la structure cristalline de NS1, on peut constater qu'il se trouve une cystéine (C118) à proximité de la boucle FG avec laquelle il serait possible d'interagir (Figure 12). C'est à partir de cette observation qu'est né le premier projet de ce mémoire.

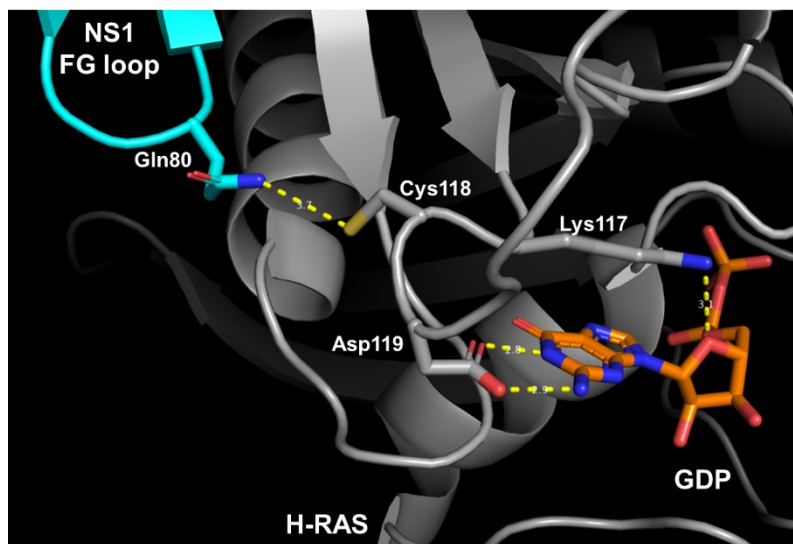


Figure 12 : Illustration de la cystéine 118 se trouvant à proximité de la boucle-FG de NS1 qui pourrait servir d’ancrage pour un lien covalent. La distance entre la glutamine 80 et la cystéine est de 3.7 Å. Selon le cristal PDB 4G0N

Ainsi l’objectif principal de ce projet était de déterminer les raisons pour lesquelles il y a une différence d’affinité de 3 ordres de grandeur entre le ‘monobody’ NS1 et le macrocycle non covalent ayant la meilleure affinité, UM01399995. Afin de répondre à cette interrogation, il est nécessaire d’établir une comparaison visuelle entre les deux entités. Ainsi le projet s’est d’abord orienté vers l’obtention d’une structure cristalline en se servant de la Cys118 de RAS pour effectuer un lien covalent avec le macrocycle, ce qui permet d’obtenir une population de HRAS lié irréversiblement au composé et donne plus de chance pour obtenir ladite structure cristalline. Celle-ci permet d’effectuer la comparaison visuelle désirée en permettant notamment de confirmer si le macrocycle reproduit bien la pose de la boucle FG du monobody NS1 et émule ses interactions. De plus, l’obtention d’une structure cristalline permet de donner de l’information sur des interactions supplémentaires possibles pour augmenter l’affinité de la série non covalente.

Pour réaliser cet objectif il sera question de développer une série de macrocycles covalents se basant sur la structure UM0139995 pour lier la cystéine 118 de RAS. Il sera nécessaire d'optimiser les groupements moléculaires pouvant réagir avec la cystéine ainsi que le bras moléculaire permettant de s'y rendre.

Une fois les résultats de la cristallographie obtenus, un second objectif sera d'explorer des avenues pour améliorer l'affinité des macrocycles en fonction de ces résultats. Ainsi, une optimisation des macrocycles sera envisagée en variant diverses sections des molécules de façon à influencer l'affinité selon des gains enthalpiques ou entropiques.

Un objectif parallèle de cette optimisation sera de déterminer une activité intrinsèque des macrocycles pour RAS, compte tenu que l'essai biochimique utilisé pour quantifier l'affinité, soit un titrage par RMN HSQC, n'a pas été, à notre connaissance, rapporté dans la littérature. Il sera ainsi question de vérifier les valeurs obtenues et de quantifier l'affinité des macrocycles covalents. Une cinétique via un essai biochimique sera envisagée pour répondre à cette incertitude.

Finalement, l'ajout d'une ogive réactive au macrocycle permet par le fait même d'augmenter l'affinité des macrocycles en amenant une dynamique d'irréversibilité pour la cible. Ainsi, il serait possible d'utiliser ces composés pour obtenir une information cellulaire dans le but de reproduire l'effet de NS1, concernant le blocage de la voie MAPK/ERK. Afin de développer les composés covalents, l'utilisation de la spectrométrie de masse sera primordiale pour confirmer la présence des adduits sur les protéines intactes. En résumé, le projet de cette première partie du mémoire est le développement d'une série de macrocycles covalents selon la structure du macrocycle UM0139995 afin de répondre à diverses questions structurales. Des techniques biochimiques et protéomiques seront utilisées pour optimiser et étudier les macrocycles covalents.

1.3. Présentation du projet des sondes photoactivables de

UM171

1.3.1 Cellules souches hématopoïétiques et applications médicales

Dans tout organisme mammifère on sépare les cellules en trois catégories, soit les cellules de la lignée germinale, les cellules somatiques et les cellules souches. La première catégorie regroupe l'ensemble des cellules reproductrices, donc les gamètes.³³ La seconde catégorie regroupe l'ensemble des cellules différenciées du corps, donc les cellules ayant acquis une fonction autre que la reproduction. Il s'agit des cellules musculaires, osseuses, des organes, des tissus, des neurones et du sang.³³ Finalement, les cellules souches sont dites indifférenciées. Elles sont les précurseurs des cellules spécialisées que sont les gamètes et les cellules somatiques. De plus, elles maintiennent l'organisme par division cellulaire. Il est important de mentionner que la spécialisation

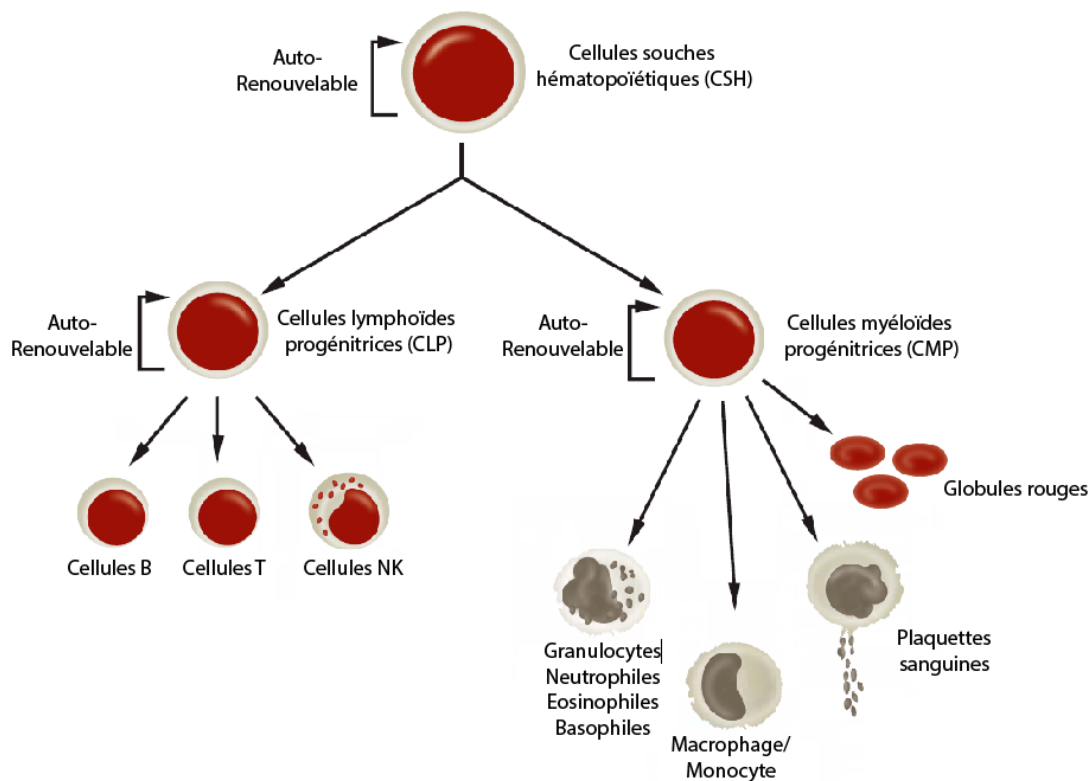


Figure 13 : Représentation des différents types cellulaires provenant de la différenciation des cellules souches hématopoïétiques. Il y a division entre les cellules lymphoïdes et les cellules myéloïdes. Figure provenant de la référence 38

des cellules souches est irréversible. Parmi les cellules souches, il existe une sous-catégorie nommée les cellules souches hématopoïétiques (CSH) qui se spécialisent dans la différenciation des lignées cellulaires présentes dans le sang.³⁴ On retrouve donc les cellules lymphoïdes progénitrices, qui sont les cellules immunitaires (cellules T, cellules B et les cellules NK) et les cellules sanguines ou myéloïdes progénitrices qui sont les plaquettes, les érythrocytes et les leucocytes. (Figure 13).³⁵

La propriété de différenciation des CSH est très intéressante pour les thérapies géniques et les transplantations. Notamment elles sont utilisées pour le traitement de la leucémie myéloïde aigüe, de la leucémie myéloïde chronique, de l'anémie falciforme, la thalassémie, des maladies immunitaires (syndrome Wiskott-Aldrich, adrenaleukodystopie, SCID-X1) et certains cancers solides.^{36, 37} Il existe trois sources possibles pour l'obtention des CSH soit via la moelle osseuse, via le sang périphérique et via le sang de cordon ombilical. Les deux principales sources utilisées sont la moelle osseuse et le sang périphérique. Cependant, dans 30-40% des patients, il existe une incompatibilité immunitaire provenant des antigènes du CMH du donneur avec le receveur et ces patient, malheureusement, n'ont pas la capacité d'obtenir un traitement.³⁸ Il s'avère que le sang de cordon permet une réduction de rejet via une meilleure tolérance des HLA avec les patients à cause du milieu plus primitif du système immunitaire des cordons.³⁸ Cette stratégie est très attrayante, d'autant plus que ce sang est facilement accessible dans les hôpitaux et ne cause pas de problème au niveau éthique puisque les cordons sont pour la plupart jetés. Néanmoins, il existe un problème avec le sang de cordon ombilical, puisque la quantité de cellules souches hématopoïétiques dans ce dernier est faible.³⁹ Il a été relevé dans la littérature que les cellules souches hématopoïétiques possèdent des marqueurs spécifiques ne se retrouvant pas dans d'autres types cellulaires.⁴⁰ Parmi ces derniers on retrouve la phosphosphoglycoprotéine CD34 qui est une protéine transmembranaire et la protéine phosphatase de tyrosine CD45.^{34, 40} Ces protéines uniques aux cellules souches permettent ainsi leur quantification dans le sang.³⁴ Il s'avère que pour qu'une transplantation puisse fonctionner chez un

patient, il faut avoir une quantité d'au moins 2×10^6 cellules CD34⁺ /kg du patient et malheureusement une unité de cordon possède en moyenne seulement 2×10^5 cellules CD34⁺.³⁷ Ceci est insuffisant pour le traitement d'un adulte. Plusieurs approches dans la littérature ont cherché à multiplier les cellules souches recueillies des cordons ombilicaux. On dénote l'utilisation de cytokines, la coculture de cellules stromales mésenchymateuses et l'amélioration des méthodologies de culture basées sur des bioréacteurs.³⁹ Bien que ces méthodes aient permis une multiplication des cellules souches, des problèmes font que ces dernières sont non optimales pour un traitement clinique. On dénote notamment, un délai dans la reconstitution de la réponse immunitaire, une augmentation des infections virales, des rejets de greffes, une trop grande variance entre les essais, l'introduction d'une malignité, un coût non-abordable et une génotoxicité.^{37, 39}

1.3.2 Expansion des cellules souches hématopoïétiques par des petites molécules et UM171

En 2010, Boitano et al. ont publié un papier discutant d'une nouvelle approche utilisant une molécule ayant la capacité de multiplier par 50 fois la quantité de cellules CD34⁺ dans une culture *ex vivo* de sang ombilical.⁴¹ La molécule StemRegenin 1 (SR1) **8** agit de façon antagoniste sur le récepteur d'hydrocarbure aryle impliqué dans les fonctions des cellules souches que sont la maintenance et la différenciation. La capacité d'utiliser cette source de cellules souches permet de limiter les désavantages présentés dans la section précédente tel que les rejets de greffes. Cette découverte a donné un enthousiasme nouveau pour la découverte de molécules permettant la multiplication des CSH aux groupes du Pr Sauvageau et du Pre Marinier. Après une optimisation de conditions permettant un haut débit d'analyse, le criblage de 5280 molécules a été effectué pour l'obtention de molécules ayant la capacité d'augmenter le niveau de cellules souches CD34⁺ CD45RA⁻³⁸ (Voir Figure 14).

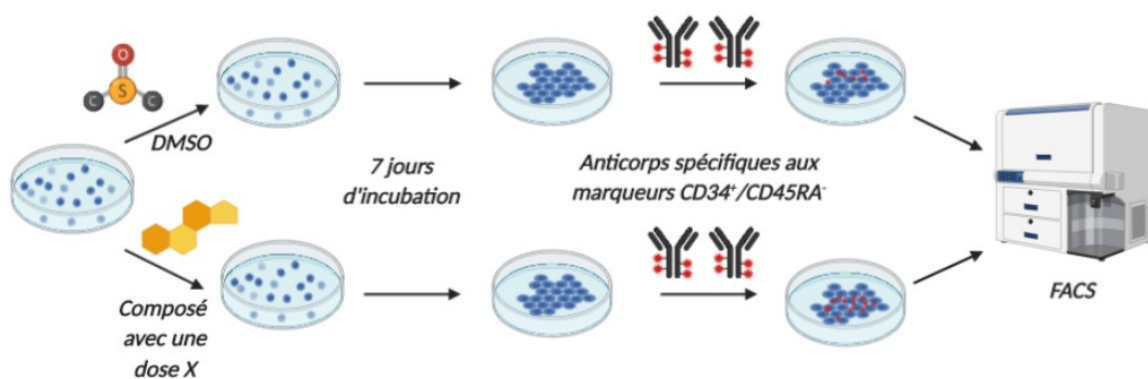
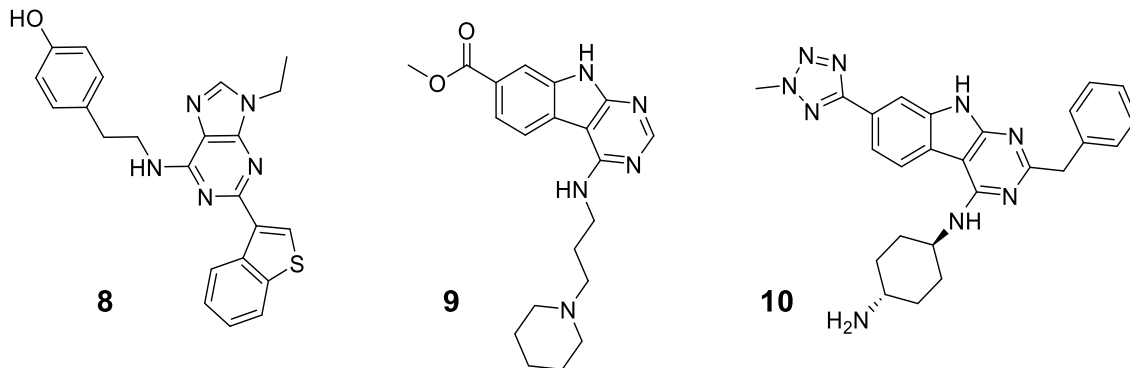


Figure 14 : Essai biologique utilisé par Sophie Corneau pour déterminer l'activité des molécules sur les marqueurs CD34⁺ et CD45RA⁻. Une culture cellulaire est séparée dans divers puits et il y a incubation de diverses concentrations de composés et d'un contrôle négatif. Après 7 jours, un anticorps révèle les cellules souches et il y a mesures quantitatives par FACS.

La logique de l'essai biologique consiste à venir immobiliser des cellules de cordon ombilical dans des canaux. Dans un premier canal, il y a eu ajout de DMSO et dans les autres canaux il y a eu ajout d'une solution du composé à tester dans le DMSO avec différentes doses. Suite à une période d'incubation de 7 jours, il y a eu ajout d'anticorps fluorescents reconnaissant sélectivement les marqueurs CD34⁺ et CD45RA⁻.³⁸ Ceci a permis ensuite de prendre une mesure quantitative des populations cellulaires selon la méthode de FACS (Fluorescence Activated Cell Sorting).⁴² Puisque cette méthode offre un aspect quantitatif en fonction de la réponse de fluorescence, il est possible d'effectuer un ratio entre la culture comportant le composé et la culture contrôle (DMSO). On nomme la dose EC_{1,5}, la dose pour laquelle un composé a occasionné une augmentation de 50% de la population de cellules souches hématopoïétiques par rapport à la population contrôle. Parmi le crible, la molécule UM729 (**9**) possédait l'effet biologique désiré. En se basant sur la structure de cette molécule comme point de départ, une étude relation structure-activité a été réalisée d'après le même essai phénotypique et la molécule UM171 (**10**) est ressortie du lot avec une affinité 10 à 20 fois supérieure à la molécule UM729 en ayant une concentration active de 17-19 nM selon la mesure EC_{1,5}.



Lors de la comparaison des composés UM171 et SR1 d'intéressantes conclusions furent observées. Tout d'abord, les deux composés possèdent un effet additif lorsqu'ils sont utilisés ensemble dans la culture cellulaire. Ceci a pour conséquence de laisser présager une cible différente, sinon l'effet aurait été minime à saturation. Cette observation a été confirmée via un séquençage de l'ARN des cellules traitées avec les molécules qui a démontré un patron génique complètement différent entre les deux composés. En fait, UM171 n'effectue pas une régulation négative des gènes du récepteur d'hydrocarbure aryle, tels que CYP1A1, CYP1B1 et AhRR, contrairement à SR1.³⁸ UM171 agit plus précisément sur des gènes impliqués dans la différenciation des érythroïdes et mégacaryocytes. Il faut savoir qu'il existe deux types de population de cellules souches hématopoïétiques, celles dites de régénération sur le long terme et celles dites de régénération sur le court terme. Celles sur le long terme permettent la régénération de toutes les cellules matures du sang, contrairement à celles sur le court terme qui produisent des progéniteurs de courte durée pour la production des cellules matures du sang et la récupération des neutrophiles.³⁸ Il s'avère qu'UM171 favorise l'expansion des cellules souches hématopoïétiques de longue régénération, ce qui n'est pas aussi accentué pour SR1, confirmant la différence mécanistique entre les deux composés. Cette différence notable offre une nouvelle possibilité pour les transplantations. Diverses études ont illustré les propriétés intéressantes de la molécule UM171 tel que la possibilité de multiplier les cellules souches *ex vivo* et augmenter la transduction génique.^{36, 38, 43, 44} De plus, des essais cliniques de phase 1/2 ont démontré que l'utilisation de sang de cordon amplifié par UM171 comme greffon chez l'humain est sécuritaire.⁴⁵ À ce jour, plus

de 22 patients⁴⁵ ont été traités avec des cordons amplifiés par UM171 et les résultats sont extrêmement prometteurs. Les plus grandes perspectives de cette molécule sont vis-à-vis la thérapie génique et les transplantations pour le traitement de la leucémie et autres maladies graves du sang.

1.3.3 Problématique et hypothèses du projet

Cependant, la cible de la molécule de UM171 reste un mystère à ce jour. Par le fait même, le mécanisme de la molécule n'est pas compris complètement compte tenu que cette molécule provient d'un criblage phénotypique. Les criblages phénotypiques ont plusieurs avantages tel que l'identification de molécules ayant des propriétés physicochimiques désirables *in vivo* simplifiant la preuve de concept du mécanisme d'action et surtout ce type de criblages offre une meilleure propension à développer des médicaments dits premier de classe (cibles inconnues) contrairement à une approche basée sur la cible. Néanmoins, l'approche phénotypique nécessite un travail plus important dans la découverte du mécanisme d'action des molécules développées.⁴⁶ Cet aspect est essentiel pour le développement d'un médicament puisqu'il faut connaître les interactions secondaires possibles et la toxicité qu'engendre ce médicament dans le corps humain. Finalement, il faut démontrer l'effet pharmacologique de ce dernier, ce qui est facilité en connaissant la cible effective. De plus, connaître la cible permet de peaufiner l'affinité du composé actif en optimisant les interactions que peut illustrer une structure cristalline.

Par conséquent, l'identification de cible est un casse-tête important dans le développement de médicaments qui proviennent d'une optimisation sur un phénotype cellulaire. Afin de donner un ordre de grandeur, un être humain possède plus de 10 000 gènes codant ainsi pour plus de 100 000 protéines. Si on se concentre plus spécifiquement à la cellule, on retrouve en moyenne de 12 000 à 40 000 protéines.⁴⁶ Cette seconde partie du mémoire cherchera à venir identifier la cible active de la molécule UM171 parmi la jungle qu'est le protéome humain. Il existe une panoplie d'approches développées dans la littérature dont le but est d'identifier une cible moléculaire, que ce soit de façon

bioinformatique, biochimique, génétique ou protéomique.⁴⁶ La méthode qui a été choisie pour réaliser ce défi est la méthode d'enrichissement basée sur l'affinité du ligand pour la protéine (nommé en anglais ABPP). Cette stratégie tire avantage de l'affinité de la molécule a pour sa cible en milieu cellulaire comparativement aux autres protéines.⁴⁷ (Figure 15).

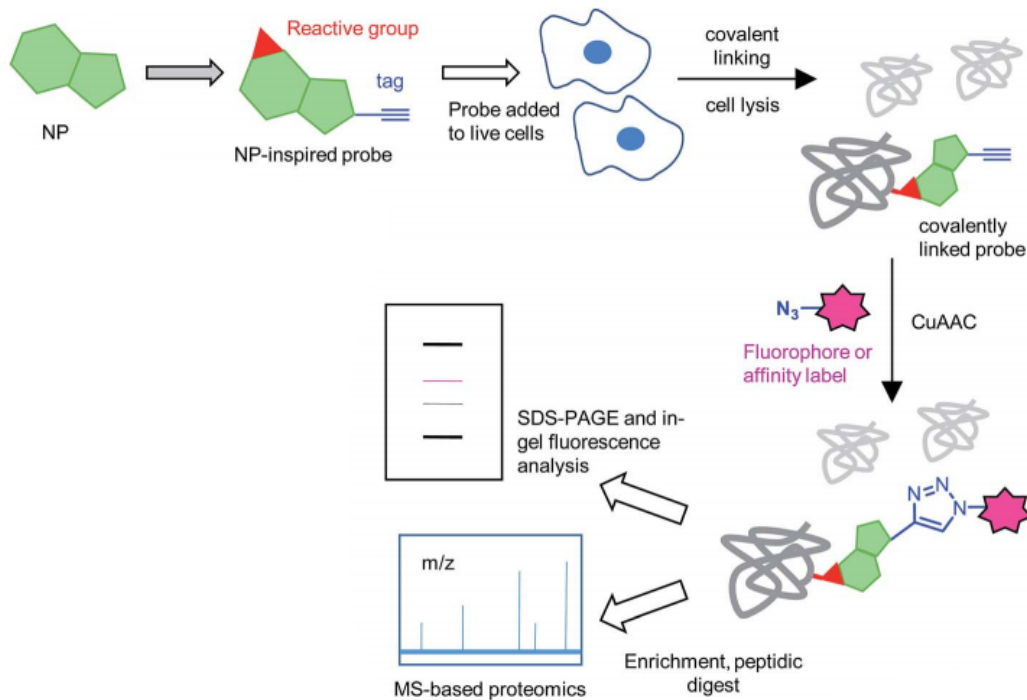


Figure 15 : Procédure d'enrichissement par l'affinité du ligand (par exemple un produit naturel NP) pour la protéine en deux temps. La procédure débute par la modification de la molécule d'intérêt en lui ajoutant un groupement réactif et un alcyne. Ensuite il y a incubation de la sonde pour effectuer le lien covalent, ensuite une réaction cycloaddition de Huisgen pour ajouter l'étiquette d'analyse et finalement il y a analyse selon un gel ou la spectrométrie de masse. Figure provenant de la référence 48

L'idée générale de cette technique consiste à utiliser la molécule pour laquelle la cible est inconnue et de lui ajouter deux fonctionnalités. Une première fonctionnalité sera nécessaire pour permettre l'analyse de la protéine d'intérêt. De multiples possibilités ont

été développées dans la littérature, mais on dénote majoritairement la biotine, un fluorophore, la séquence peptidique FLAG et un atome radioactif.⁴⁷ (Figure 16)

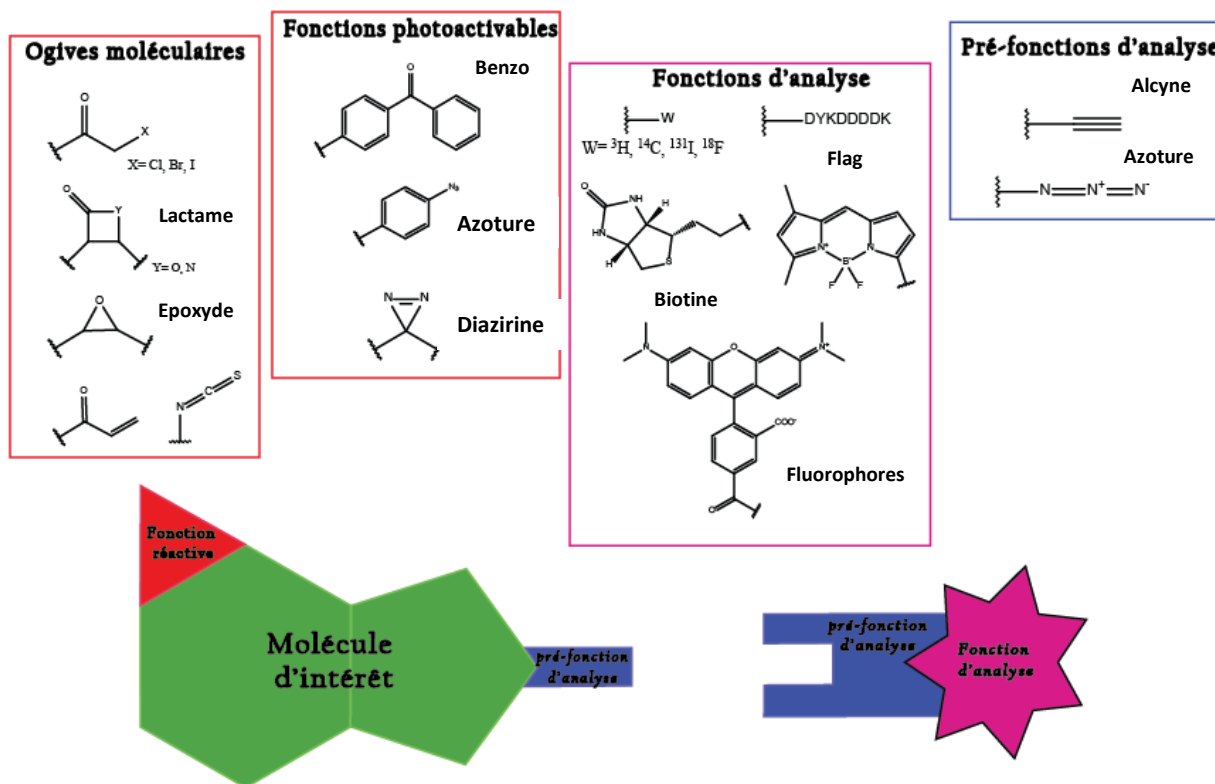


Figure 16 : Présentation des divers groupements nécessaires pour l'approche ABPP. La molécule d'intérêt doit être fonctionnalisée avec une fonction réactive pouvant être une ogive moléculaire ou une fonction photoactivable et une fonction d'analyse. Afin de conserver l'affinité de la molécule d'intérêt il est possible d'effectuer la procédure en deux temps en utilisant une pré-fonction d'analyse.

Cependant l'ajout de ces étiquettes crée souvent des problèmes concernant les propriétés du ligand. Ce sont des groupements volumineux changeant considérablement les propriétés physico-chimiques des substances et pouvant aussi influencer grandement l'affinité du composé pour la cible cellulaire. Pour contrer cette problématique, la procédure se fait en deux temps, il y a d'abord ajout d'un alcyne ou un azoture sur le

ligand et, dans une étape subséquente une réaction de cycloaddition de Huisgen comportant l'étiquette d'analyse avec le partenaire complémentaire pour la cycloaddition.⁴⁷ La réaction de cycloaddition de Huisgen fait partie de la famille des réaction nommée Click.⁴⁸ Pour qu'une réaction soit de type Click, il faut qu'elle soit modulaire, possède des rendements très élevés, ne produisant pas ou peu de sous-produits, qui est stéréospécifique, qui possède une force motrice thermodynamique et finalement qui est physiologiquement stable.⁴⁸ Comme l'alcyne et l'azoture sont de petites fonctionnalités, ils influencent très peu l'affinité et les propriétés du ligand. Cette réaction est très robuste et rapide en milieu aqueux et cellulaire, ce qui en fait une réaction de choix. La seconde fonctionnalité joue le rôle de stabiliser l'interaction protéine-ligand pour permettre de maintenir cette interaction lors des manipulations d'analyses. Pour ce faire, la meilleure façon pour stabiliser une interaction est de former un lien covalent. Ainsi, il existe deux familles pour former le lien covalent, il y a les ogives moléculaires qui sont des électrophiles possédant des groupes partants permettant une attaque nucléophile d'acides aminés spécifiques et la seconde famille consiste en des fonctions photoactivables.⁴⁷ Ces dernières sont des fonctions qui forment des espèces très réactives lors d'une activation lumineuse. On dénote la benzophénone et la diazirine qui forme un carbène et l'azoture qui forme un nitrène. Ces espèces extrêmement instables ont la capacité de s'insérer dans les liens C-H, O-H, N-H et S-H.⁴⁹ Ainsi, les radicaux auront la capacité de réagir avec les acides aminés du site actif de la cible protéique inconnue. L'idée sera ensuite d'utiliser l'étiquette d'analyse pour isoler la cible. Comme l'indique le nom même de l'approche d'affinité du ligand pour la protéine, il est primordial que la sonde conserve une certaine activité biologique avec la cible thérapeutique. Par conséquent, une étude structure-activité est essentielle pour connaître les sites de fonctionnalisation possibles afin de conserver l'effet biologique.

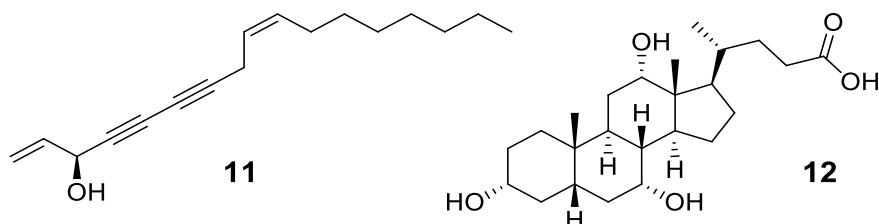
Il existe de multiples étapes pour l'approche d'affinité du ligand pour la protéine et ces dernières sont illustrées à la figure 15.⁴⁹ Lorsque la sonde a été synthétisée, la première étape consiste en une incubation cellulaire pour laisser le temps à cette dernière d'interagir avec la protéine d'intérêt. S'il s'agit d'une sonde munie d'une trappe

électrophile, cette dernière va réagir de manière covalente avec les acides aminés nucléophiles de la protéine d'intérêt et le lien chimique sera formé sans étape supplémentaire. Cependant dans le cas d'une fonction photochimique, la sonde va simplement interagir de façon non-covalente avec la protéine d'intérêt et nécessitera une étape de rayonnement ultraviolet qui formera l'espèce radicalaire créant le lien covalent avec la protéine d'intérêt. L'étape suivante consiste à lyser les cellules pour exposer les protéines liées aux composantes permettant de lier les composantes de la réaction de cycloaddition pour attacher l'étiquette d'analyse à la sonde. À ce stade, le choix de l'analyse va dépendre de l'étiquette d'analyse sélectionnée. Si l'optique est de venir faire un enrichissement, une biotine ou une séquence FLAG sont d'excellents choix. Cependant, si l'optique est d'effectuer un gel SDS-PAGE ou une localisation cellulaire, les fluorophores et les ligands radioactifs sont plus appropriés. Il est possible d'identifier la protéine d'intérêt par une méthode d'immunodétection ou via la spectrométrie de masse. Lorsqu'une cible potentielle est identifiée, il est nécessaire de valider cette dernière selon des essais biochimiques, la SPR (résonance des plasmons de surface), la MST (thermophorèse à micro-échelle) ou de l'ITC (titrage calorimétrique isotherme).⁵⁰

La justification du choix de cette méthode pour l'identification de la cible du composé UM171 s'appuie par la quantité croissante d'articles illustrant un portrait positif pour l'identification de cibles et d'interacteurs utilisant cette technologie. De nombreux articles résument les avancées de la technique, et parmi ces derniers on retrouve plusieurs exemples de réussites, notamment via des sondes photoactivables.^{47, 51-54} Voici quelques exemples qui ont illustré l'efficacité de la méthode d'approche d'affinité du ligand pour la protéine. Le premier exemple provient du groupe du professeur Sieber qui a déterminé la cible inconnue du produit naturel Falcarinol (**11**) qui se retrouve dans les carottes, le persil et le gingembre.⁵⁴ Ce composé possède des propriétés anti-cancéreuses, mais provoque aussi une réaction inflammatoire ressemblant aux allergies. L'importance de connaître le mécanisme d'action et la cible de ce composé est crucial pour expliquer les effets secondaires non-désirables du composé. Une dérivation du composé falcarinol avec le couple diazirine/alcyne a permis d'identifier le substrat ALDH2,

une aldéhyde déshydrogénase. Cette enzyme permet d'oxyder les aldéhydes en acide carboxylique, donc une enzyme essentielle pour la détoxification et le métabolisme des aldéhydes dans le corps. Par conséquent, une inhibition de l'aldéhyde déshydrogénase occasionne une augmentation du stress oxydatif causant des dommages cellulaires menant à l'apoptose.⁵⁴ Ainsi par l'identification de la cible, le groupe du professeur Sieber a été en mesure d'expliquer le mécanisme d'action du composé.

Une seconde étude cherchait à identifier des interacteurs de l'acide biliaire (**12**).⁵³



Bien que cette molécule soit connue pour réguler le métabolisme des lipides et des glucides, elle agit sur d'autres mécanismes cellulaires comme une hormone. Étant donné que de nouvelles pathologies de l'acide biliaire sur des nouveaux substrats sont survenues dans les dernières années, une équipe a souhaité identifier les cibles de l'acide biliaire pour connaître l'étendue des mécanismes physiologiques impliqués. Pour ce faire, ils ont synthétisé une sonde composée d'une diazirine et d'un alcyne et se sont servis de la méthode de chimio-protéomique ABPP. Leur étude de protéomique a réussi à obtenir 331 interacteurs à l'acide biliaire. L'analyse des fonctions où ces protéines sont impliquées a révélé que ces protéines avaient un rôle majoritairement dans la transformation des protéines dans le réticulum endoplasmique, le métabolisme et le transport des protéines. Ils ont obtenu des substrats très bien connus de la littérature tel que les protéines FXR et ITGB1. Cependant, par leur étude ils ont fait la preuve de concept de 3 nouveaux substrats (CPT1A, ADPGK, COMT) par buvardage de Western. Ces protéines offrent de nouvelles réponses et pistes pour la compréhension de mécanisme cellulaire. Cet exemple

démontre aussi que l'approche ABPP permet de déterminer plusieurs interacteurs à la fois et peut même illustrer des interacteurs adjacents.

Par ces divers exemples et l'étendue des résultats positifs de la littérature pour l'identification de cible de petites molécules, la méthodologie de l'ABPP a été choisie pour l'identification de la cible de UM171. Par conséquent, l'hypothèse de ce projet est d'être en mesure d'identifier la cible thérapeutique via l'approche d'affinité du ligand pour la protéine. La sélection de la fonction photoactivable a été privilégiée étant donné une meilleure sélectivité vis-à-vis les ogives moléculaires. L'approche en 2 temps a été aussi choisie pour maximiser la conservation des propriétés physico-chimiques des sondes par rapport au composé UM171.

Afin de réaliser l'objectif d'identifier la cible thérapeutique de UM171, il sera nécessaire dans un premier temps de synthétiser diverses sondes analogues de UM171. Tel que discuté plus haut, il est nécessaire d'avoir une sonde qui possède une bonne affinité, donc il sera fondamental de synthétiser diverses sondes variant l'emplacement des fonctions nécessaires pour l'approche ABPP et en évaluer leur activité biologique par après.

Lorsque ces synthèses seront complétées, il sera nécessaire de valider chaque étape que compose l'approche de l'ABPP. Tout d'abord, pour se faire la main avec la méthode, ensuite pour en comprendre ses faiblesses et les améliorer. Par conséquent, il sera nécessaire de vérifier les paramètres d'incubation des sondes, le solvant de lyse nécessaire, les paramètres de la réaction de cycloaddition de Huisgen et les paramètres d'irradiation ultraviolette. Lorsque la méthodologie sera maîtrisée, il y aura expérience avec les sondes synthétisées pour en déterminer la cible thérapeutique. Les techniques d'analyse qui seront utilisés pour l'analyse de l'approche seront la spectrométrie de masse et les gels SDS-PAGE.

2. Développement de macrocycles covalents cherchant à lier RAS

2.1. Synthèses sur support solide des macrocycles covalents

2.1.1 Discussions du choix des paramètres pour les macrocycles covalents

Lors de la confection d'un outil moléculaire il est important d'optimiser divers paramètres pour maximiser les chances de lier la cible d'intérêt. Dans le cas des macrocycles se liant à RAS, il sera nécessaire d'optimiser les diverses ogives moléculaires pouvant réagir avec la cystéine 118. De plus, il faut connaître la distance atomique optimale entre cette ogive et le résidu thiol. En premier lieu, l'alanine se situant normalement à la position mauve sur la figure 17 dans le composé UM0139995 fut remplacée par un acide aminé possédant une fonction amine. Il a été démontré dans la RSA qu'il y avait une tolérance dans le choix des acides aminés à cette position et la cystéine s'y trouve à proximité d'après la structure cristalline de NS1 (PDB 4G0n).²⁷ Les amines sont idéales puisqu'elles permettent une protection orthogonale lors de la synthèse sur support solide et constituent d'excellents nucléophiles pour l'ajout d'ogives. Ainsi, deux longueurs de chaînes aminées seront choisies pour explorer la longueur optimale du dit bras moléculaire pour aller rejoindre la cystéine 118. Ces acides aminés avec une amine sur la chaîne latérale seront la lysine et l'ornithine, qui ont une différence de 1 méthylène. Ce sont des acides aminés pourvus de chaînes latérales flexibles, donnant alors un maximum de degré de liberté pour venir se lier à la cystéine désirée.

Pour le choix des ogives, deux paramètres furent testés. Dans un premier cas, l'orientation de l'ogive : comme la réaction irréversible varie selon le type d'ogive, elle requiert un angle d'attaque différent de la part du nucléophile.⁵⁵⁻⁵⁷ Par exemple, lorsqu'il y a utilisation de l'acrylamide comme ogive, il est nécessaire que la cystéine soit orientée avec un angle de Bürgi-Dunitz⁵⁸, soit 105-109°, afin de réagir sur le carbone sp² terminal. Cependant, lorsque l'on considère les halogénures en position alpha du carbonyle, l'angle

d'attaque doit être antipériplanaire⁵⁸, 180° par rapport à l'halogène. Finalement, les alcynes requièrent un angle d'attaque obtus avoisinant les 120° à cause de la géométrie des orbitales π et σ .⁵⁹ Ainsi, une large gamme d'angles sera testée par le choix de ces trois familles d'ogives.

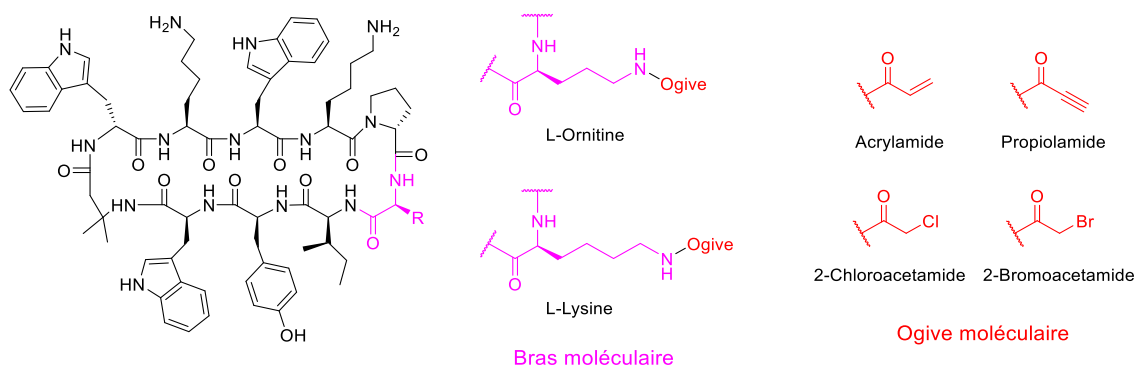


Figure 17: Hypothèses synthétiques de la première série de macrocycles covalents se basant sur la structure d'UM0139995. L'ajout de la section covalente se fera à l'endroit de l'alanine de la structure d'UM0139995. Les bras moléculaires utilisés sont illustrés en mauve et les ogives en rouge.

Le second aspect du choix des ogives est leur réactivité. Une étude a été faite en 2014⁵⁷ qui comparait le temps de demi-réaction de diverses ogives avec le glutathion via RMN et spectrométrie de masse. La figure 18 illustre les temps de demi-réaction de diverses ogives afin de permettre une comparaison.

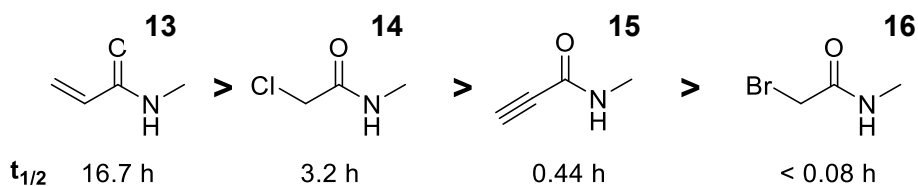


Figure 18 : Comparaison cinétique de diverses ogives de méthylamide réagissant avec le glutathion. Ces temps de demi-vie ont été déterminés par spectrométrie de masse et RMN. Selon la référence 57.

On peut constater qu'il existe de grandes différences en termes de réactivité entre l'acrylamide **13**, qui possède un temps de demi-réaction de 16.7 heures, et le bromoacetamide **16**, qui possède un temps de demi-réaction de moins de 5 minutes. Les différences de réactivité s'expliquent par l'activation du carbone électrophile. Dans le cas du brome, il s'agit d'un excellent groupe partant à cause de son caractère nucléofuge occasionné par son grand nuage électronique. Comparativement, le chlore **14**, qui possède un plus petit nuage électronique, est un moins bon groupe partant. Pour les accepteurs de Michael, la triple liaison **15** est beaucoup plus déformable que l'énone permettant ainsi un système de résonance plus fort avec le carbonyle, activant donc le carbone beta de façon plus importante.⁵⁸ Ainsi, avec ce choix d'ogives comme première tentative, il y a couverture d'un large spectre de réactivité pour réagir avec la cystéine **118**.

2.1.2 Voie de synthèses des macrocycles covalents

Discutons maintenant de l'aspect synthèse de ces macrocycles covalents. La stratégie initiale sera expliquée et les détails techniques seront développés par la suite. La figure 19 met en image l'approche choisie. La chimie sur support solide est extrêmement utile pour la synthèse de peptides et de macrocycles, puisqu'elle permet de contrôler les réactions peptidiques et d'obtenir d'excellents rendements de réactions. Pour élaborer cette synthèse, il a été nécessaire d'utiliser 4 groupes protecteurs orthogonaux différents pour activer les fonctions à des moments précis. Le terme orthogonal signifie la capacité de retirer les groupes protecteurs via différents mécanismes/conditions sans impacter les autres. D'ores et déjà, le choix du groupe protecteur des chaînes latérales des acides aminés s'est arrêté sur les groupements tert-butyle (Tyr) et Boc (Trp, Lys), des fonctionnalités qui se déprotègent aisément en conditions acides. Ensuite, le groupement alloc (Allyloxycarbonyl) fut utilisé pour protéger l'acide carboxylique de la lysine **17** ancrée sur la résine. Ce dernier se déprotège via une hydrogénéolyse. Ensuite, les amines en position alpha de chaque acide aminé, protégées par un groupement Fmoc, sont

libérées en condition basique, plus précisément via des amines secondaires.⁶⁰ Finalement, le groupement ivDde (1-(4,4-diméthyl-2,6-dioxocyclohex-1-ylidène)-3-méthylbutyle) a été utilisé pour protéger l'amine de la chaîne latérale où a été installée l'ogive, cette dernière se déprotègeant via l'utilisation d'hydrazine.⁶¹ Le choix des groupes protecteurs est essentiel afin de choisir sélectivement les acides aminés que l'on veut

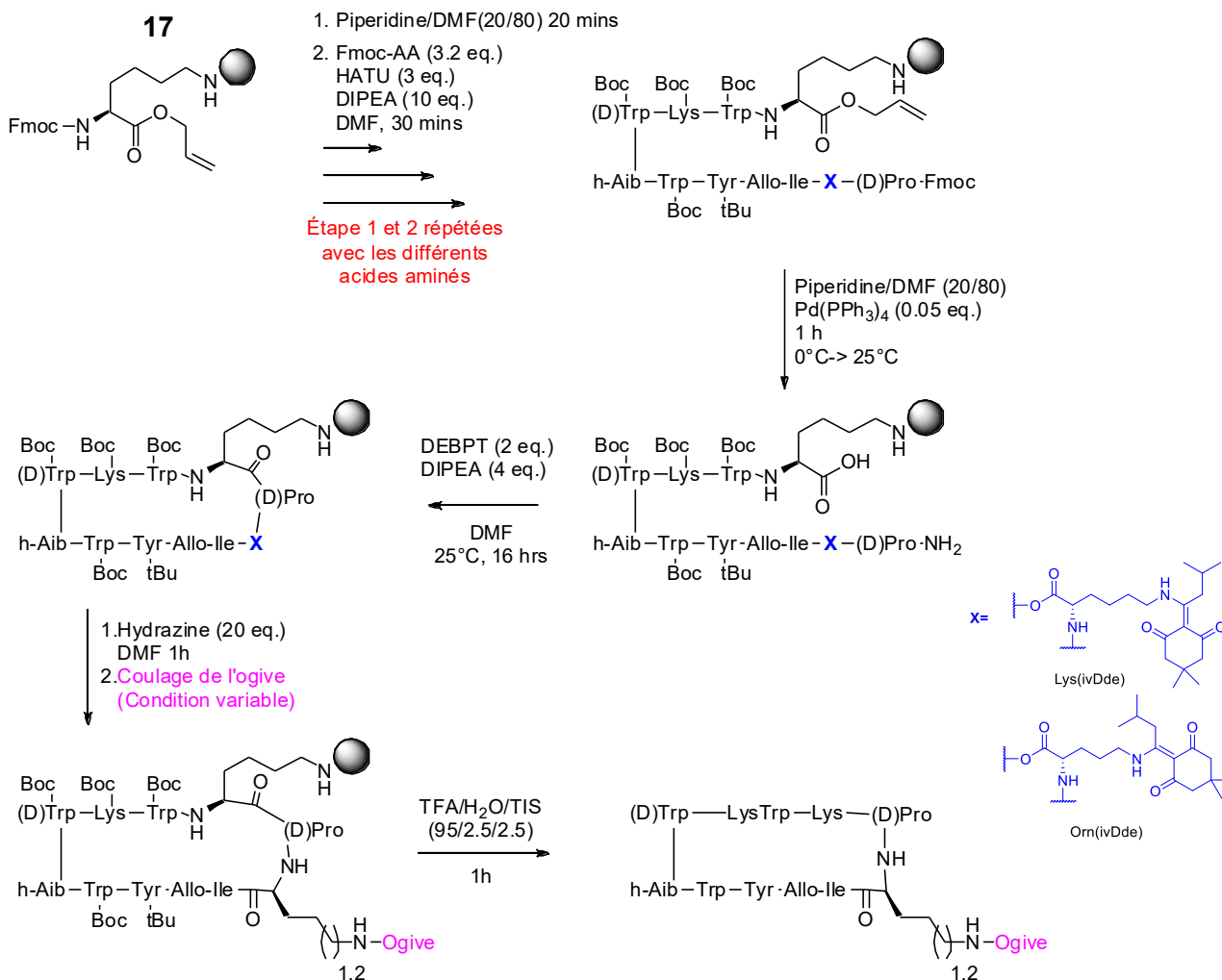


Figure 19 : Séquence synthétique pour l'optimisation des macrocycles covalents en support solide. Une fois la lysine ancrée par l'amine en position epsilon sur la résine, les différents acides aminés de la séquence furent ajoutés via une séquence de déprotection avec la pipéridine et couplage via le HATU. La séquence a divergé via l'ajout du précurseur du bras atomique où l'Orn(ivDde) ou la Lys(ivdde) furent ajoutée. La macrocyclisation a été réalisée via l'agent de couplage DEBPT dans le DMF. Finalement, le groupe protecteur IvDde a été déprotégé par l'hydrazine avant l'instauration de l'ogive moléculaire. La dernière étape de la séquence est le clivage et la déprotection des chaînes latérales du macrocycle par un cocktail de TFA, eau et triisopropylsilane.

faire réagir.

La chimie Fmoc/tBu est idéale pour l'allongement de séquences peptidiques, puisqu'il est possible de déprotéger sélectivement les amines alpha au moment du couplage en laissant les fonctions nucléophiles se trouvant sur les chaînes latérales des acides aminés protégées par des groupements non sensibles aux amines basiques. Ensuite, le groupement ivDde va permettre de fonctionnaliser sélectivement l'amine de la chaîne latérale pour instaurer l'ogive moléculaire sans influencer les autres groupes protecteurs. Finalement, le groupement alloc sera enlevé lors de la macrocyclisation, contrôlant ainsi de manière spécifique l'acide aminé qui va réagir. Les groupements tBu et Boc seront clivés à la toute fin.

La première étape de la synthèse est l'ancrage du premier acide aminé sur le support solide (Figure 20). La stratégie synthétique s'est arrêtée sur l'ancrage de l'amine de la chaîne latérale d'une lysine **19** sur la résine chlorure de 2-chlorotrityle **18**. Cette dernière est idéale pour ancrer une amine et obtenir directement l'amine lors du clivage final. Elle est clivable en milieu acide, donc adéquate pour l'utilisation de la technologie Fmoc/tBu pour la synthèse sur support solide.

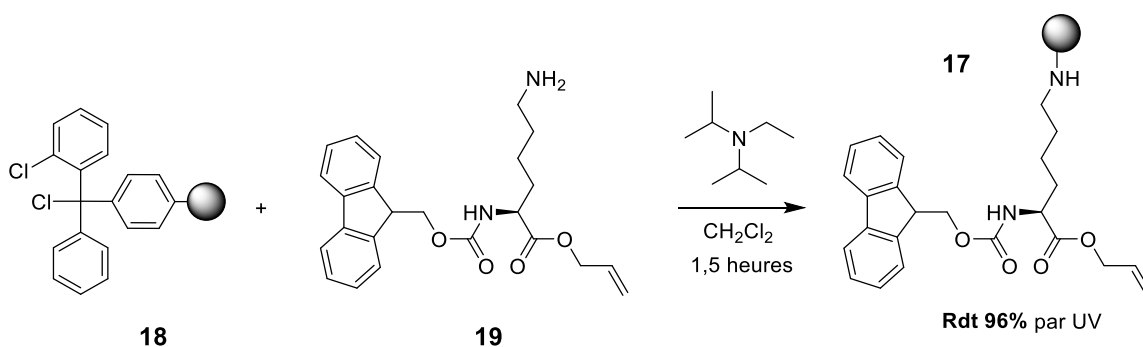


Figure 20 : Réaction d'ancrage du Fmoc-Lys-Alloc sur la résine chlorure de 2-chlorotrityle via une E1

La lysine qui a été utilisée fut préalablement protégée avec un Fmoc sur l'amine alpha et un groupement alloc sur l'acide carboxylique. Cette réaction a donné de bons

rendements, qui a été déterminée via une mesure d'absorption des groupements Fmoc de la littérature .⁶⁰

La seconde étape de la synthèse fut l'ajout des différents acides aminés afin d'obtenir les macrocycles désirés. Les étapes sont toujours les mêmes : déprotection du groupement Fmoc sur l'amine via la pipéridine suivie du couplage de l'acide aminé suivant via l'agent de couplage HATU et la base DIPEA. Ainsi la séquence générale d'acides aminés fut la suivante : (AAX = acide aminé et X est la position)

AA2= Fmoc-Trp(Boc)-OH, **AA3**= Fmoc-Lys(Boc)-OH, **AA4**= Fmoc-(D)-Trp(Boc), **AA5**=Fmoc-Homo-Aib-OH, **AA6**=Fmoc-Tyr(tBu)-OH, **AA8**=Fmoc-Allo-Ile-OH, **AA9**= Fmoc-Lys(ivDde)-OH ou Fmoc-Orn(ivDde)-OH et **AA10**=Fmoc-(D)-Pro-OH.

Après l'ajout de la proline, il y a eu deux réactions conjointement effectuées dans le même milieu réactionnel. Il s'agit de la déprotection du Fmoc sur l'amine alpha de la (D)-proline par l'élimination- β du dibenzofulvene ainsi que le clivage de l'alloc par l'hydrogénolyse. La réaction s'est faite sur une durée de 16 heures à cause de la réactivité ralentie qu'occasionne la catalyse hétérogène nécessaire lors de l'hydrogénolyse. Lorsque ces réactions sont complétées, la macrocyclisation entre la proline et l'amine alpha de la lysine d'ancrage est effectuée en utilisant un couplage peptidique. Cependant, cette fois-ci le DEPBT a été sélectionné contrairement au HATU. Le DEPBT est un réactif organophosphoré réputé pour effectuer des couplages peptidiques avec un très faible taux de racémisation même s'il y a un grand délai avec l'ajout de l'amine couplante. Il a été démontré dans la littérature que pour favoriser une réaction intramoléculaire par rapport à une réaction intermoléculaire, il faut être en condition diluée. Par conséquent, une réaction diluée possède une cinétique plus lente compte tenu que les réactifs ont une probabilité plus faible de réagir ensemble.⁶² Ainsi pour une réaction de macrocyclisation nécessitant un long temps de réaction (16h), l'agent couplant DEPBT est plus adéquat. Finalement, le DEPBT est un excellent réactif pour la cyclisation des peptides,

contrairement aux agents de couplages dérivés de l'urée tels que le HATU et le HBTU. Une étude de 2004 avait démontré la variation de rendements occasionnés uniquement par le choix de l'agent de couplage, le DEPBT s'étant avéré le meilleur.⁶³ L'étape suivante est la déprotection du groupement protecteur ivDde via l'ajout d'hydrazine.

Lorsque l'amine est libérée sur l'acide aminé AA9, il y a eu ajout de diverses ogives. Les différentes conditions sont illustrées à la figure 21.

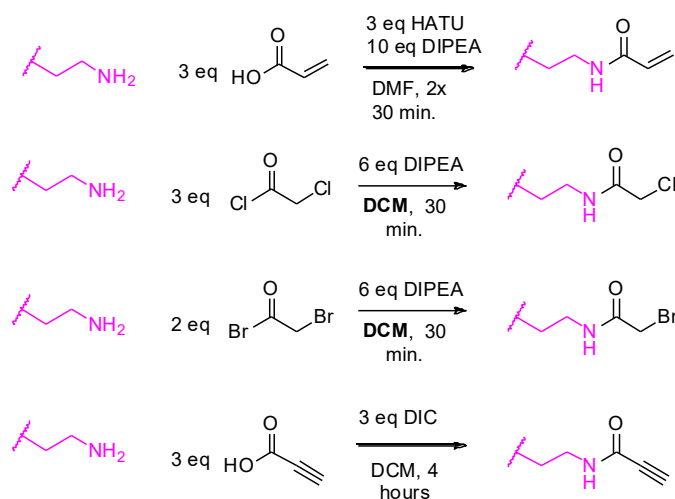


Figure 21 : Conditions réactionnelles utilisées pour coupler les différentes ogives aux amines des bras moléculaires.

Pour ajouter les acrylamides, un couplage activé par le HATU a été effectué pour l'obtention du produit désiré. Afin d'observer si la réaction se produisait, une faible quantité de résine était prélevée et clivée avec un mélange de TFA, d'eau et de triisopropylsilane. Ensuite, l'échantillon était analysé par LCMS et la masse permettait de vérifier l'avancement de la réaction. Pour les chloroacétamides, la réaction qui a fonctionné le mieux était le couplage de l'amine avec le chlorure d'acyle dans le dichlorométhane. L'utilisation du DMF comme solvant pour cette réaction a limité le taux de conversion de la réaction. Cette tendance a aussi été observée pour les deux autres ogives que sont le bromoacétamide et le propiolamide. L'explication de cette observation

est liée à la résine trityle. Il s'agit d'une résine très hydrophobe et comme les macrocycles protégés possèdent un caractère aussi très hydrophobe, ils ont tendance à s'adsorber à la surface des billes et ne sont pas exposés aux réactifs du solvant. Dans la littérature, il a été remarqué que l'utilisation de solvants plus apolaires tels que le toluène, le dichlorométhane et le dioxane permet de mieux solvater ce type de résines, et de conséquemment offrir un meilleur contact des réactifs.⁶⁴ C'est la raison qui explique que l'utilisation du dichlorométhane a permis une conversion complète par rapport au DMF. Pour former le bromoacétamide, la réaction avec le bromure d'acyle a été sélectionnée et finalement pour le propiolamide un couplage avec le N,N'-Diisopropylcarbodiimide et l'acide propiolique a fonctionné.

La dernière étape de la synthèse consiste à la déprotection des chaînes latérales des acides aminés et le clivage de l'amine de la résine trityle. Cette réaction fut effectuée par l'ajout d'un mélange de TFA/H₂O/ TIS (Triisopropyl silane). Cependant, comme le silane est un réducteur, dans les cas des acrylamides et des propiolamides, le TIS a été remplacé par de l'eau par précaution.

Tous les macrocycles ont pu être obtenus avec des rendements globaux variables tels que l'illustre le tableau 1.

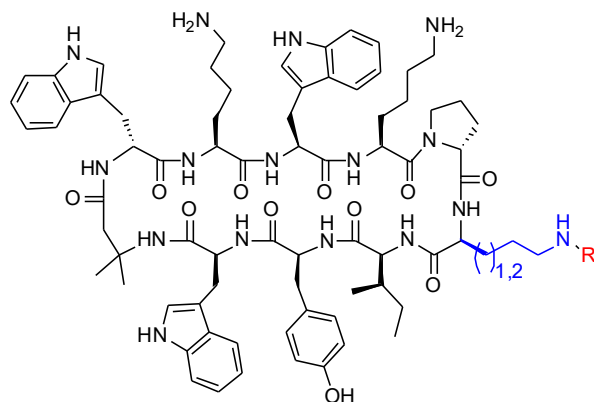
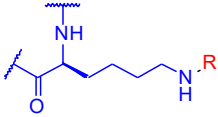
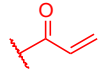
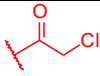
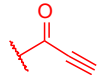
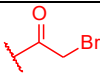
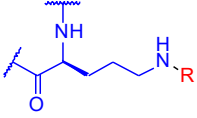
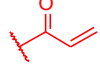
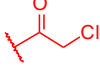
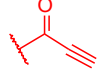
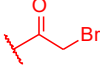


Tableau 1 : Identification et rendement globaux des macrocycles de la première optimisation

Chaîne amine	Ogive R	Numéro UM	Rdt global
L-Lysine 		UM0140184	7.00 mg (12.87%)
		UM0140186	8.75 mg (16.01%)
		UM0140438	9.16 mg (16.87%)
		UM0140439	2.91 mg (5.12 %)
L-Ornithine 		UM0140191	10.65 mg (19.70%)
		UM0140327	9.53 mg (17.42 %)
		UM0140440	12.64 mg (23.50 %)
		UM0140401	11.36 mg (20.16 %)

Il est possible de constater que les rendements globaux pour vont de 5 % à 24 %, et ce, pour un total de 24 réactions pour chacun des 8 composés synthétisés. Un rendement global de 24 % correspond à un rendement moyen de 95 % pour chaque réaction, ce qui démontre l'efficacité de la synthèse sur support solide. Ce rendement peut sembler médiocre au premier regard : or, compte tenu du grand nombre de réactions, la valeur est très acceptable. Bien que les réactions soient suivies par LCMS pour confirmer l'efficacité de la réaction, il peut survenir des pertes tout au long de la synthèse. Cependant, la perte majeure se produit sans aucun doute durant la purification finale des produits. Il y a eu deux purifications par phase inverse sur un système HPLC préparatif, et bien que ce système soit très performant pour la purification, des pertes non négligeables peuvent être encourues. Les pertes sont occasionnées par plusieurs facteurs tels que les fractions impures que l'on doit sacrifier, le délai entre le détecteur et la réponse de récolte des fractions ou même le seuil minimal de réponse nécessaire pour avoir une récolte ainsi que les pertes instrumentales qui peuvent survenir sur la colonne chromatographique.⁶⁵

2.2. Résultats protéomiques et biochimiques

2.2.1 Analyse de la réactivité des ogives moléculaires

Avec les macrocycles en main, l'étape suivante a consisté à vérifier leur liaison sur les protéines RAS. Dans un premier cas, la spectrométrie de masse a été utilisée compte tenu de sa très grande sensibilité. Le schéma de la méthode est illustré à la figure 22.

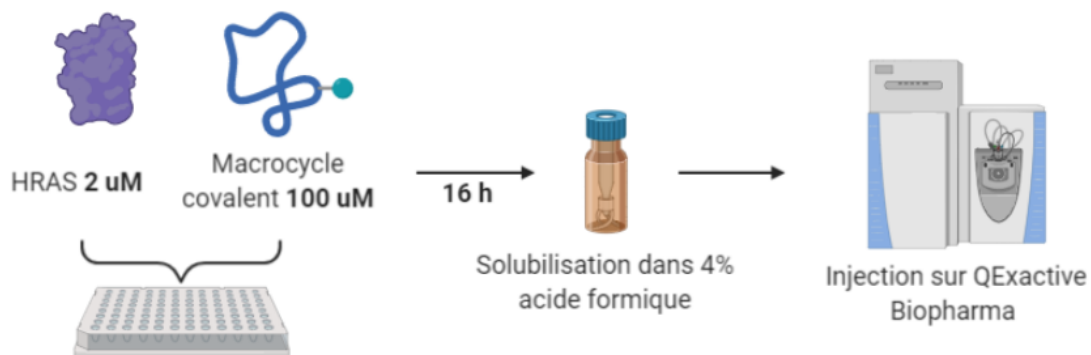


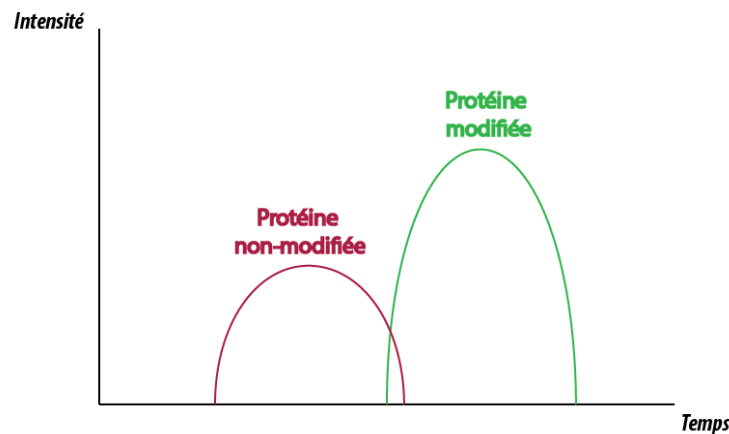
Figure 22 : Procédure utilisée pour détecter les adduits sur les protéines intactes de RAS. Il y a eu réaction en plaque de 96 puits durant 16h avec 2 µM de HRAS et 100 µM de macrocycle dans un milieu tamponné. Pour l'injection, il y a solubilisation avec 4% d'acide formique et injection dans sur une précolonne C8. Une séparation chromatographique sur une colonne capillaire C5 a été effectuée en amont de l'analyse par Orbitrap™.

Les protéines RAS qui ont été utilisées pour cette section furent exprimées et purifiées par Hugo Lavoie du laboratoire de Marc Therrien. Elles ont été exprimées dans des bactéries via une construction à laquelle a été ajoutée une étiquette poly-histidines.

Cette dernière est ajoutée à la fin de la séquence de la protéine afin de permettre une purification par affinité de charge avec une résine d'acide nitrilotriacétique et de Nickel. La première étape de la méthode consiste à faire réagir les macrocycles covalents avec la protéine RAS pour une durée de 16 heures dans un tampon physiologique. Cependant, des expériences préalables ont démontré qu'il n'était pas possible d'utiliser un agent réducteur dans le tampon, telles que le TCEP ou le DTT, puisque ces derniers réagissent

avec les ogives des macrocycles et inhibent grandement les adduits sur les protéines. Un temps de 16 heures a été utilisé pour uniformiser les expériences et surtout limiter la variance cinétique qui se produit entre les différentes ogives moléculaires. Ensuite, les échantillons ont été injectés sur le Q Exactive™ Biopharma de Thermo Fischer.

Il y a deux étapes à l'analyse des données. Dans un premier temps, il y a eu confirmation de la présence d'adduits sur la protéine intacte en effectuant la différence entre la masse exacte observée et la masse de la protéine sans modifications. La seconde étape fut la quantification de la protéine modifiée versus la protéine non modifiée pour avoir un aspect comparatif entre les différents macrocycles. La quantification des adduits a été faite par calcul du ratio des aires sous la courbe des différentes espèces de protéines. (Figure 23) Comme il s'agit de la même protéine avec ou sans modifications, il est possible d'approximer une ionisation suffisamment similaire pour quantifier uniquement via l'intensité des ions de l'enveloppe protéique.

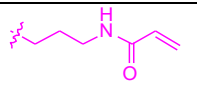
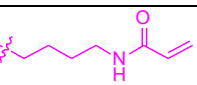
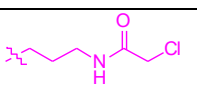
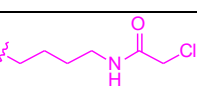
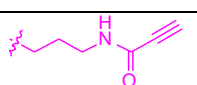
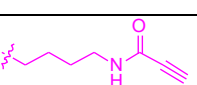
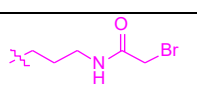
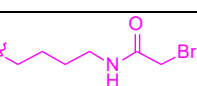


$$\% \text{ d'adduit} = \frac{\text{Aire de la protéine modifiée}}{\text{Aire de la protéine modifiée} + \text{Aire de la protéine non-modifiée}} \times 100$$

Figure 23 : Équation utilisée pour déterminer le pourcentage d'adduits selon l'aire des espèces de protéines en masse intacte

Pour cette première expérience, il y a eu utilisation de la protéine HRAS et la protéine mutée HRAS C118S pour chaque macrocycle. L'avantage d'utiliser la protéine mutée au résidu cystéine 118 consiste à la confirmation de la modification de la bonne cystéine sur RAS, puisque cette dernière possède trois cystéines. Une disparition d'espèces modifiées avec le mutant indiquerait la modification de la bonne cystéine. Les résultats sont présentés dans le tableau 2.

Tableau 2 : Pourcentages de modifications des protéines HRAS native et HRASC118S mutante par les divers macrocycles covalents quantifiés par spectrométrie de masse

Composé		HRAS			HRAS C118S		
Structure ogive	Numéro	%P	%S	%D	%P	%S	%D
	UM0140191	100	0	0	100	0	0
	UM0140184	100	0	0	100	0	0
	UM0140327	100	0	0	100	0	0
	UM0140186	>99	<1	0	100	0	0
	UM0140440	13	87	0	100	0	0
	UM0140438	69	31	0	100	0	0
	UM0140401	0	88	12	76	24	0
	UM0140439	0	85	15	75	25	0

Légende : %P est le pourcentage de protéines non modifiées

%S est le pourcentage de protéines possédant 1 adduit

%D est le pourcentage de protéines possédant 2 adduits

Ce tableau illustre beaucoup de données, pour y détacher les conclusions, il y aura décortiquées des informations. Débutons par l'analyse de la réactivité de différentes ogives sur HRAS. Il y a eu observations d'adduits sur la protéine HRAS avec les ogives propiolamide, chloroacetamide et bromoacetamide. Cependant, il est important de mentionner que dans le cas du chloroacetamide, la quantité de protéine modifiée est très faible. Le pourcentage d'adduits est en accord avec la réactivité intrinsèque des ogives : le bromoacetamide est le plus réactif et possède le plus haut taux d'adduits. En fait, il ne restait plus de protéine HRAS non modifiée après 16h, et ce, peu importe la longueur du bras moléculaire (lysine ou ornithine). De plus, les bromoacetamides (**'401** et **'439**) sont très réactifs au point de former 2 adduits sur la protéine HRAS, donc de se lier à 2 cystéines. On observe que la population majoritaire est celle avec un seul adduit. Finalement, les propiolamides (**'440** et **'438**) sont moins réactifs, au point qu'il reste de la protéine HRAS non modifiée. De surcroît, aucune population de protéines avec deux adduits n'a été observée pour ce type d'ogives. La seconde partie du tableau illustre les adduits sur la protéine mutée HRAS C118S. La logique veut que si l'adduit du macrocycle se fait sur la bonne cystéine, on ne devrait voir aucun adduit sur le mutant. C'est ce qui a été observé avec les propiolamides. Toutefois, pour les bromoacetamides, des adduits ont été observés. Ceci n'est pas une surprise et concorde avec les résultats observés sur la protéine HRAS native, qui possédait une population avec deux adduits. Néanmoins, il est possible de dire que la grande majorité des macrocycles avec un bromoacetamide se lient à la cystéine 118 étant donné la proportion réduite d'adduits sur la protéine mutante. Ces résultats furent très encourageants, car dans un premier cas, les macrocycles se lient sur la protéine et se lient préférentiellement sur la bonne cystéine.

Afin de confirmer la modification des macrocycles sur la cystéine 118, une expérience de détection d'adduits sur les peptides a été réalisée avec le composé **'439**. Cette expérience a été réalisée par Éric Bonneil de la plateforme de protéomique (Figure 24). L'expérience consistait à prendre la protéine modifiée et à la digérer par la protéase Gluc-C. Les peptides protéolytiques ont été ensuite injectés sur le Q Exactive Biopharma. L'objectif a été d'identifier les peptides possédant un adduit du macrocycle et ensuite d'identifier le

site de modification sur le peptide via l'analyse des spectres MS/MS. Les deux cystéines qui avaient une modification étaient la cystéine 118 et la cystéine 80. Cependant, beaucoup plus de peptides modifiés sur la cystéine 118 ont été identifiés, démontrant que cette dernière était le site principal de modification.

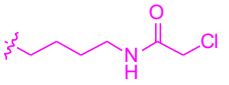
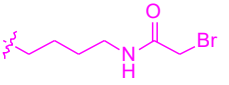
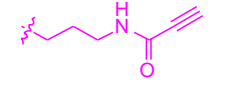
Weight	Sequence	Prob	Pepti...	NTT	Charge	Modifications
1.0	(E)YSAMRDQYMRITGEGFLC V FAINNTKSFE(D)	100%	35.4	2	5	cyclic peptide (+1455)
1.0	(E)YSAMRDQYMRITGEGFLC V FAINNTKSFE(D)	100%	29.2	2	5	cyclic peptide (+1455), Deamidated (+1)
1.0	(E)GFLC V FAINNTKSFE(D)	99%	27.9	2	4	cyclic peptide (+1455)
1.0	(E)GFLC V FAINNTKSFEDIHQYRE(Q)	100%	36.2	2	4	cyclic peptide (+1455)
1.0	(E)GFLC V FAINNTKSFEDIHQYRE(Q)	100%	32.7	2	5	cyclic peptide (+1455)
1.0	(E)QIKRVKDSDDVPMVLVGNK C DLAARTVE(S)	100%	32.8	2	5	cyclic peptide (+1455)
1.0	(E)QIKRVKDSDDVPMVLVGNK C DLAARTVE(S)	100%	31.8	2	4	cyclic peptide (+1455)
1.0	(E)QIKRVKDSDDVPMVLVGNK C DLAARTVE(S)	98%	27.2	2	6	cyclic peptide (+1455)
1.0	(E)QIKRVKDSDDVPMVLVGNK C DLAARTVE(S)	97%	27.8	2	4	Oxidation (+16), Deamidated (+1), cyclic peptide (+1455)
1.0	(E)QIKRVKDSDDVPMVLVGNK C DLAARTVESRQAQDLARSYGIPYIE(T)	100%	48.1	2	6	Oxidation (+16), cyclic peptide (+1455)
1.0	(E)QIKRVKDSDDVPMVLVGNK C DLAARTVESRQAQDLARSYGIPYIE(T)	100%	31.8	2	6	Oxidation (+16), Deamidated (+1), cyclic peptide (+1455)
1.0	(K)RVKDSDDVPMVLVGNK C D(L)	100%	31.5	2	5	Oxidation (+16), cyclic peptide (+1455)
1.0	(R)VKDSDDVPMVLVGNK C D(L)	99%	28.8	2	4	Oxidation (+16), cyclic peptide (+1455)
1.0	(R)VKDSDDVPMVLVGNK C D(L)	99%	27.4	2	5	Oxidation (+16), cyclic peptide (+1455)
1.0	(R)VKDSDDVPMVLVGNK C D(L)	98%	26.6	2	3	Oxidation (+16), cyclic peptide (+1455)
1.0	(D)SDDVPMVLVGNK C DLAARTVE(S)	100%	29.0	2	5	Oxidation (+16), Deamidated (+1), cyclic peptide (+1455)
1.0	(D)SDDVPMVLVGNK C DLAARTVE(S)	100%	28.6	2	4	cyclic peptide (+1455)
1.0	(M)VLVGNK C DLAARTVE(S)	100%	36.2	2	5	cyclic peptide (+1455)
1.0	(M)VLVGNK C DLAARTVE(S)	100%	30.3	2	3	cyclic peptide (+1455)
1.0	(M)VLVGNK C DLAARTVE(S)	99%	28.3	2	3	cyclic peptide (+1455)

Figure 24 : Détermination des cystéines modifiées par des adduits du macrocycle 401 selon une analyse peptidique. La Glu-C a été utilisée pour la digestion de la protéine et une recherche d'adduit de 1455m/z a été effectuée. Les cystéines 80 et 118 ont démontré la présence d' adduits.

Une autre expérience réalisée avec cette série de macrocycles covalents fut d'observer le comportement de ces derniers vis-à-vis les isoformes de HRAS, soit KRAS et NRAS. À titre de rappel, NS1 avait une grande affinité pour HRAS et KRAS, mais était inactif pour NRAS. Chaque isoforme a été exprimée de l'acide aminé 1 à 171, donc sans la zone d'hypervariabilité.

La même procédure expérimentale discutée plus haut a été utilisée, mais cette fois-ci avec les trois isoformes. Ces derniers furent comparés avec les composés covalents '186, '439 et '440. Les résultats de cette expérience sont présentés dans le tableau 3.

Tableau 3 : Pourcentages de modifications des isoformes HRAS, KRAS et NRAS par les divers macrocycles covalents obtenus via la spectrométrie de masse

Composé (100 μ M)		Protéine (4 μ M)		
Structure ogive	Numéro	HRAS	KRAS	NRAS
	UM0140186	%P=83 %S=17 %D=0	%P=100 %S=0 %D=0	%P=100 %S=0 %D=0
	UM0140439	%P=0 %S=72 %D=28	%P=16 %S=58 %D=26	%P=33 %S=43 %D=24
	UM0140440	%P=25 %S=75 %D=0	%P=34 %S=66 %D=0	%P=100 %S=0 %D=0

Légende : %P est le pourcentage de protéines non modifiées

%S est le pourcentage de protéines possédant 1 adduit

%D est le pourcentage de protéines possédant 2 adduits

D'ores et déjà, on peut constater que la tendance est conservée par rapport à la première expérience lorsqu'on analyse HRAS. Les résultats sont reproductibles, alors même qu'on observe un plus grand pourcentage d'adduits avec le '186 que lors de l'expérience précédente. L'intérêt provient des résultats des deux autres isoformes, KRAS et NRAS. Il s'avère que le bromoacetamide ('439) forme des adduits avec KRAS, mais on observe cette fois-ci les trois possibilités en même temps. Il y a une population non modifiée, une modifiée par un adduit et une modifiée par 2 adduits. Ce résultat est similaire à ce qui a été observé avec HRAS, donc les macrocycles semblent interagir avec KRAS aussi, comme

observé pour NS1. Encore une fois l'apparition de doubles adduits avec KRAS illustre qu'une seconde cystéine est modifiée, ce qui est logique avec la très grande réactivité de l'ogive.

Maintenant, l'analyse de NRAS n'a pas été intéressante pour le bromoacetamide, puisque l'on constate la présence d'un grand pourcentage de protéines possédant deux adduits. Il existe le même nombre de cystéines dans chaque isoforme pour la séquence exprimée, soit 3, donc il ne s'agit pas de la présence d'une nouvelle cystéine qui expliquerait ces adduits. Si on compare avec le résultat de NS1, NRAS ne se lie pas avec le monobody, alors que le macrocycle **'439**, semble se lier à cette isoforme. La conclusion de ce résultat est que le bromoacétamide réagit fortement avec les cystéines, ce qui explique une perte de sélectivité. Cependant, lorsque l'on regarde les résultats avec le propiolamide **'440**, on n'observe qu'un seul adduit sur KRAS et aucun adduit sur NRAS. Ce résultat est en accord avec la littérature de NS1.²⁷

Il est toujours bien de confirmer un résultat scientifique via une seconde méthode. Pour ce faire, les résultats de protéomique ont été répétés sur gel SDS-PAGE via une coloration au bleu de Coomassie. Étant donné la grande masse moléculaire des macrocycles, il est aisé de voir une variation de hauteur des bandes sur les gels et ainsi permettre d'évaluer l'ajout d'adduits sur les isoformes de RAS. L'expérience biochimique a uniquement mis de l'avant les composés **'439** et **'440**, étant donné leur bon taux de conversion pour la formation d'adduits.

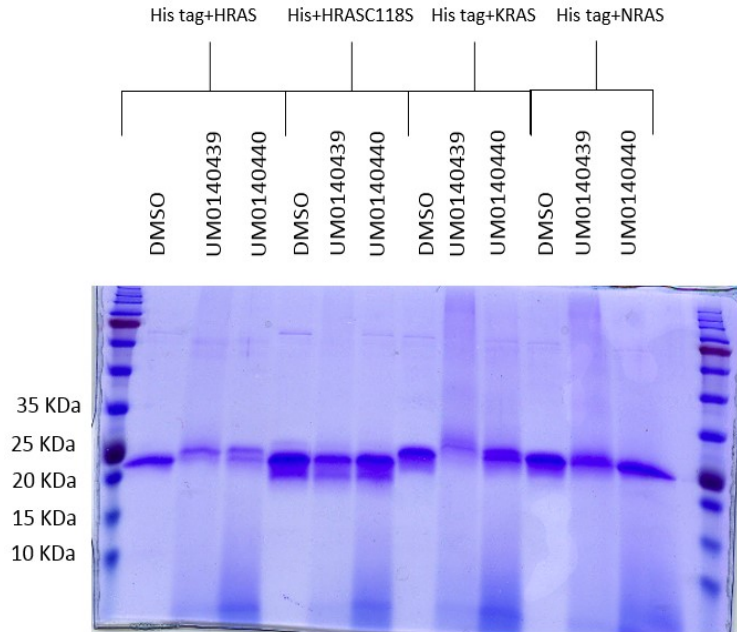


Figure 25 : Confirmation des résultats de protéomique via un gel SDS-Page par coloration au bleu de Coomassie pour les macrocycles **'439** et **'440**. Les protéines HRAS, KRAS et NRAS ont été testées en plus du mutant HRASC118S.

Le gel d'acrylamide a permis de confirmer plusieurs conclusions obtenues par l'analyse des protéines intactes (Figure 25). Tout d'abord, le bromoacétamide **'439** modifie clairement HRAS, KRAS et NRAS, ceci étant détecté par l'observation de bandes supérieures au contrôle. De plus, il y a une faible diminution d'intensité de la bande de HRAS C118S par rapport au contrôle DMSO, ce qui illustre une certaine liaison de **'439** pour le mutant. Maintenant, pour le propiolamide **'440**, on constate un taux de conversion moins important pour HRAS et KRAS contrairement au **'439**. De plus, il y a pratiquement aucun changement pour NRAS et pour le mutant HRAS. Ces résultats confirment les analyses de spectrométrie de masse.

2.2.2 Structure cristalline de UM0140401

Connaissant la propension du bromoacétamide **UM0140401** à modifier aisément HRAS à la cystéine 118, des expériences de cristallographie ont été entamées par Tarun Arya du laboratoire du Pr Therrien. Après de multiples optimisations de conditions, il est parvenu à obtenir des cristaux suffisamment purs pour en déduire une structure cristalline ayant une résolution de 2.75 Å. La figure 26 illustre la superposition du composé covalent '401 et de la boucle FG de NS1. En cristallographie, la résolution correspond à la distance nécessaire pour séparer deux objets distinctement, tels que des atomes.⁶⁶ Ainsi, si la résolution du cristal est de 2.75 Å, cela signifie qu'il est nécessaire d'avoir cette distance pour séparer distinctement deux atomes. Cependant, un lien carbone est de 1,5 Å, donc, avec une résolution de 2.75 Å, il est impossible de distinguer les nuages électroniques d'un lien C-C. Une résolution de 4 Å est très mauvaise et une résolution de 1.2 Å est considérée comme excellente.⁶⁶ Par conséquent, 2.75 Å est une résolution acceptable, puisqu'elle permet d'identifier la disposition du squelette moléculaire du macrocycle ainsi que l'orientation des chaînes latérales des acides aminés.

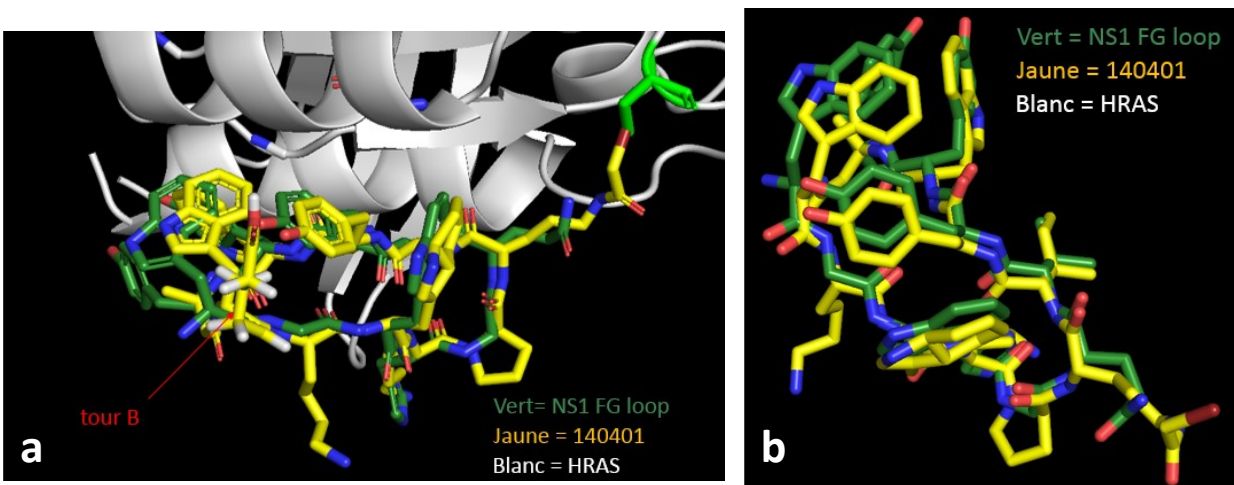


Figure 26 : Structure cristalline du macrocycle '401 sur la GTPase HRAS selon une résolution de 2.75 Å. En a) une vue de côté exposant le site de liaison du macrocycle sur la protéine HRAS en gris, suivie d'une corrélation avec la boucle-FG de NS1. La partie b) de la figure donne une orientation différente pour comparer la superposition de NS1 et 401.

En vert est illustré la boucle FG de NS1, en blanc HRAS et en jaune la macrocycle **'401**. Le premier constat qu'il a été possible d'obtenir de cette structure cristalline, est que le macrocycle se lie à la même place que NS1, ce qui est une très bonne nouvelle. Une seconde observation est que sur la vue de côté (Figure 26a), il est possible de voir la liaison de la cystéine 118 en couleur lime sur le carbone en alpha du carbonyle de l'ogive. Maintenant, si l'on compare la pose de NS1 par rapport au macrocycle **'401**, on peut dégager plusieurs conclusions. Tout d'abord, le macrocycle semble tirer vers la droite en comparaison à NS1, de façon à observer une différence spatiale pour chaque acide aminé. Ce déplacement est de plus en plus accentué lorsque l'on s'éloigne de la cystéine 118. Il existe notamment une différence de 1.7 Å entre NS1 et **'401** par rapport aux atomes de carbones du tryptophane adjacent au tour β . Deux hypothèses furent amenées pour expliquer cette différence. La première est que le bras moléculaire approchant la cystéine 118 est trop court, venant ainsi étirer le macrocycle par une tension. Rappelons que **'401** est le macrocycle ayant une ornithine comme acide aminé en position AA9. Par conséquent, une structure cristalline avec l'analogue lysine permettrait de vérifier l'impact de la tension du bras moléculaire sur la pose du macrocycle. La seconde hypothèse réside dans la flexibilité du tour β . On constate aussi qu'il existe de grandes différences entre le cristal de NS1 et celui du **'401** pour cette région. Il se pourrait que l'homo-Aib utilisé pour effectuer le tour bêta occasionne une contrainte rotationnelle et empêche d'obtenir la même pose que NS1. Outre le décalage du macrocycle, la pose de NS1 est très similaire au composé covalent. Il est logique de se poser les questions suivantes : quelle est la longueur optimale de la chaîne latérale de l'acide aminé AA9 ? Quel est l'impact de la flexibilité du tour bêta sur la pose du macrocycle ? Est-ce que cette distinction spatiale permet d'expliquer la grande différence d'affinité entre les deux entités chimiques ? Ainsi ces questionnements sont la continuité de l'histoire de la prochaine section.

2.3. Optimisation des macrocycles covalents

2.3.1 Détermination de la longueur optimale du bras moléculaire

Cette section sera un méli-mélo des diverses observations sur les macrocycles covalents en tentant de répondre aux diverses questions amenées par la cristallographie. Dans un premier cas, l'un des aspects qu'il était nécessaire de déterminer était la longueur optimale du bras moléculaire pour atteindre la cystéine 118 selon l'ogive moléculaire. La structure cristalline obtenue illustre que le bras semblerait trop court avec l'utilisation du bras moléculaire ornithine dans la structure du UM0140401, où l'ogive était le bromoacétamide. Par conséquent, il est nécessaire de comparer diverses longueurs de bras avec les deux meilleures ogives obtenues, le propiolamide et le bromoacétamide (Figure 27).

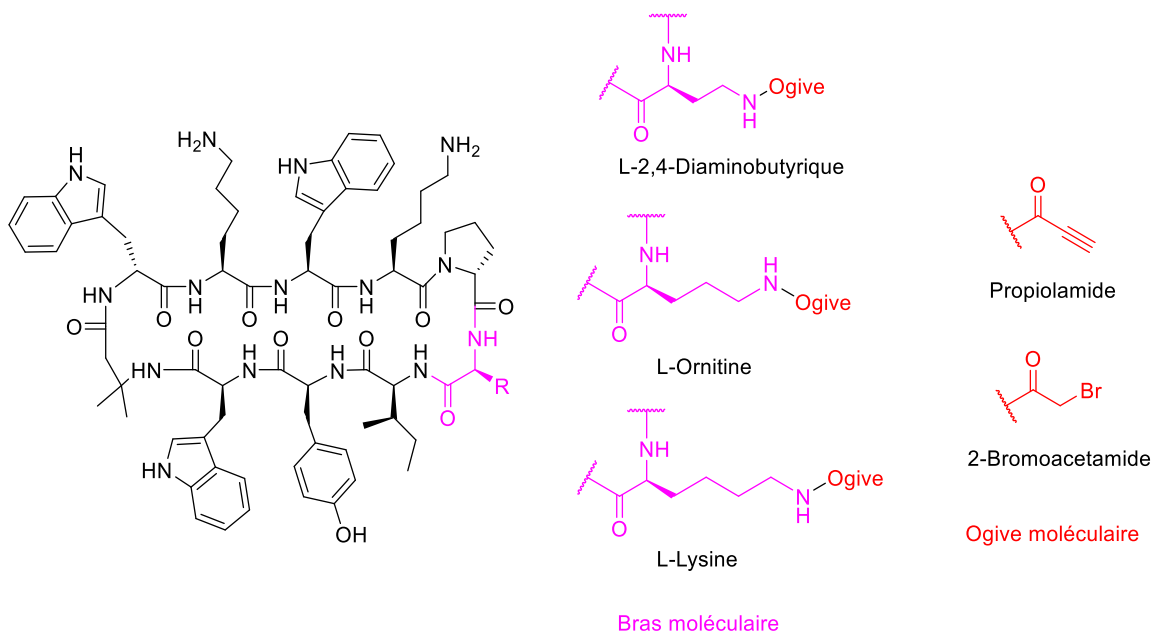


Figure 27 : Hypothèses synthétiques de la seconde série de macrocycles covalents cherchant à élucider la longueur optimale du bras moléculaire. Ces macrocycles se basent toujours sur la structure du macrocycle UM0139995.

Bien que les bras moléculaires à 3 atomes et à 4 carbones aient été synthétisés et déjà discutés dans le chapitre précédent, il était nécessaire de synthétiser l'analogue à deux carbones. Contrairement aux deux derniers acides aminés, l'acide aminé agissant comme bras moléculaire n'était pas disponible commercialement. Par conséquent, il a été requis de le synthétiser, selon les étapes de synthèse décrites à la figure 28.

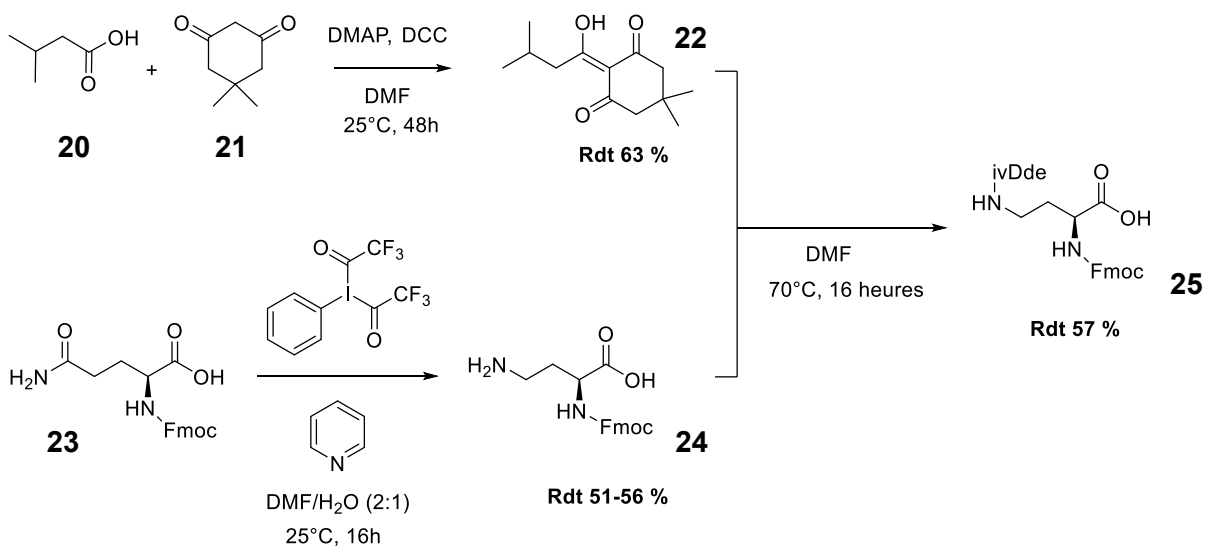
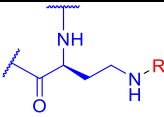




Figure 28 : Route synthétique pour l'acide aminé Fmoc-Dab(ivDde)-OH.

La première partie de la synthèse fut la formation du groupement protecteur, ivDde **22**, qui permet de protéger l'amine sur la chaîne latérale des acides aminés en position AA9. Pour ce faire la procédure décrite par Siri Ram Chhabra a été reproduite.⁶¹ Il s'agit de faire l'acylation de la dimédone **21** avec l'espèce ammonium activée préalablement formée à partir de l'acide isovalérique **20** et du DMAP. Une fois ce groupement obtenu, il a été nécessaire d'obtenir l'acide aminé précurseur. Ainsi, un réarrangement de Hoffman a été réalisé sur la glutamine protégée **23** par le réactif PIFA en milieu basique. La résultante fut l'obtention du Fmoc-DAB-OH **16**. La dernière étape de la séquence fut de venir ajouter le groupement protecteur ivDde sur l'amine beta de l'acide 2,4-diaminobutanoïque (DAB)

24. Pour ce faire, il y a eu condensation de l'amine libre sur l'énol du IvDde. Cette réaction a été plus complexe que prévu. On dénote que la réaction ne se produisait pas et que si on chauffait la réaction, on observait la formation de la lactame correspondante. Suite à une brève optimisation des conditions, il a été possible d'obtenir un rendement de 57 % du produit isolé **25** après utilisation du DMF et chauffage à 70°C sur une durée de 16 heures. Une fois cette entité chimique obtenue, il a été possible d'entamer la synthèse sur support solide des deux macrocycles possédant un bras moléculaire de deux atomes. La procédure utilisée fut la même que celle décrite dans la précédente section (Figure 18).

Tableau 4 : Rendements globaux et identification des macrocycles avec un bras moléculaire fait à partir du DAB

Chaîne amine	Ogive R	Numéro UM	Rdt global
		UM0140706	3.56 mg (6.68 %)
		UM0140690	3.48 mg (6.23 %)

Avec les macrocycles en main, une analyse comparative par gel SDS-PAGE a été effectuée avec l'aide précieuse d'Hugo Lavoie (voir Figure 29). Pour une dose constante de protéines HRAS (2 μM), il y a eu titrage des macrocycles covalents à différentes concentrations (0.8 μM à 50 μM). La concentration où l'on observait 50 % de modifications de la protéine native était considérée comme la valeur de comparaison EC_{50} . Ainsi, pour cet essai, il sera nécessaire d'avoir une concentration supérieure à 1 μM pour observer 50 % d'adduits sur la protéine.

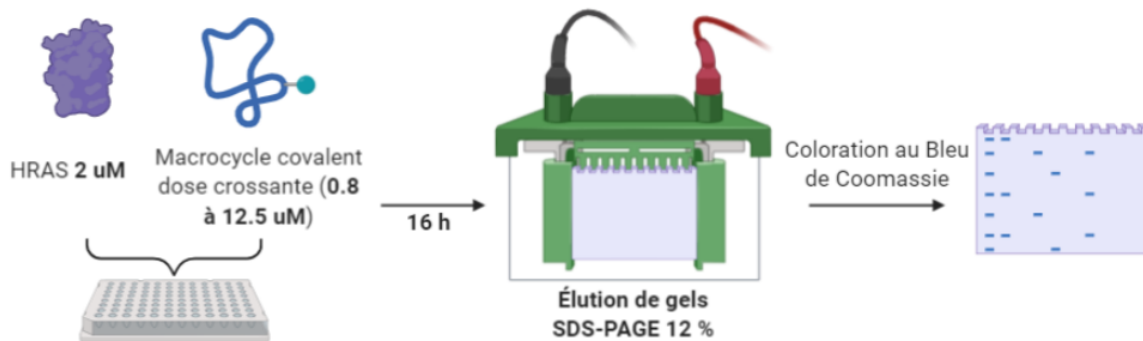


Figure 29 : Schéma des manipulations expérimentales pour l'analyse de la longueur optimale du bras moléculaire.

La figure 30 présente un exemple de gels obtenus parmi les trois réplicats effectués. L'expérience fut réalisée au moins 3 fois pour chaque composé, afin d'avoir une meilleure reproductibilité, considérant qu'il s'agit d'un essai qualitatif. Sur les gels correspondants, la longueur de la chaîne augmente de façon croissante en se déplaçant vers la droite pour chaque famille. Le tableau 5 récapitule les résultats de ces expériences.

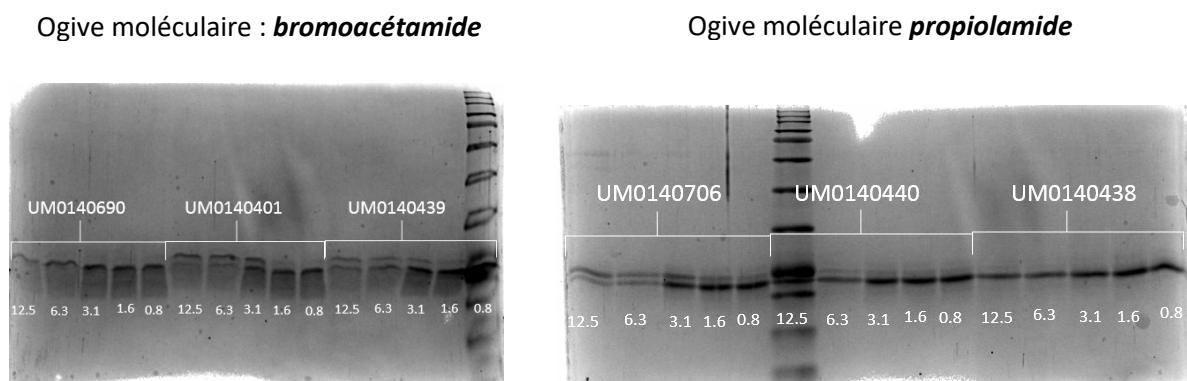
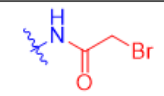
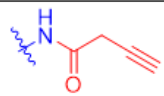
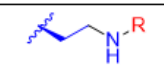
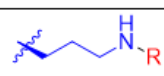
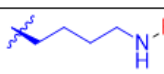


Figure 30 : Exemples gels SDS-PAGE par coloration au Bleu de Coomassie pour l'étude d'activité des macrocycles covalents afin de déterminer la longueur optimale du bras moléculaire.

Tableau 5 : Valeurs d'EC₅₀ pour la conversion des protéines HRAS selon les diverses longueurs de bras moléculaires pour deux familles d'ogives.

Ogive moléculaire Bras moléculaire									
	Numéro UM	EC ₅₀ (μM)	Écart-type(μM)	N=	Valeur p	Numéro UM	EC ₅₀ (μM)	Écart-type(μM)	N=
	UM0140690	>50	N/A	3	N/A	UM0140706	6.0	1.9	4
	UM0140401	3.7	1.9	5		UM0140440	>12.5	N/A	3
	UM0140439	6.8	2.2	6	0.024	UM0140438	>12.5	N/A	3

Pour mieux séparer l'analyse de ces données, discutons des macrocycles par familles, les bromoacétamides et les propiolamides. Tout d'abord, la tendance qui a été observée pour l'ogive bromoacétamide est que la longueur optimale est celle de l'ornithine. Un EC₅₀ se situant près de 3.7 μM a été obtenu pour cette longueur. Cette observation est démontrée à la figure 30, où l'on peut constater qu'à la valeur de 3.1 μM, on observe une double bande pour le composé **'401**, tandis que pour les deux autres analogues, on observe une double bande pour des valeurs plus élevées, soit 6.3 μM pour **'439** et supérieur à 12.5 μM pour **'690**. D'autres gels présentent des incréments de concentration différents pour couvrir une zone plus large, expliquant ainsi les valeurs différentes entre le tableau et les gels. Il pourrait être tentant de croire que les écarts-types soient très grands et dire que la différence n'est pas significative entre l'analogue ornithine et lysine. Cependant, il faut se rappeler que les incréments de concentrations sont très grands (ex. 3.1 μM à 6.3 μM). Toutefois, pour de grands écarts-types, les calculs statistiques démontrent clairement leur significativité lors du calcul de la valeur p qui est inférieure à 5%. La conclusion selon laquelle l'ornithine possède la longueur atomique optimale permet alors d'infirmer l'hypothèse soutenant que le macrocycle semblait tiré de sa pose

à cause d'un bras moléculaire trop court dans la structure cristalline discutée dans le chapitre précédent.

Contrairement à la série bromoacétamide, il s'avère que pour la série propiolamide, la longueur appropriée est obtenue via l'utilisation du DAB avec une valeur d' EC_{50} de 6.0 μ M. Sur le gel présenté, on peut constater que l'utilisation d'une chaîne plus longue du bras moléculaire diminue de beaucoup le nombre d'adduits pour les concentrations supérieures à 12,5 μ M. On peut clairement comprendre que la longueur du bras moléculaire joue un rôle critique dans la liaison du macrocycle à la protéine. Que ce soit dans la série propiolamide ou bromoacétamide, un bras atomique trop long ou trop court réduit de beaucoup l'affinité pour la cystéine 118. Maintenant, si on compare les macrocycles optimaux des deux séries, on observe une différence de près du double entre les deux EC_{50} (3.7 μ M pour **'401** et 6.0 μ M pour **'706**). Ceci est probablement causé par la différence de réactivité intrinsèque des ogives, étant donné que la longueur est optimale cette fois-ci.

Il est intéressant de réfléchir à la justification logique de la longueur des bras moléculaires optimaux des deux séries. Lorsque l'on compare le mécanisme réactionnel de l'attaque du thiol de la cystéine 118 sur les atomes responsables du lien covalent pour les deux ogives, on constate que l'emplacement atomique est le même (Figure 31). L'attaque se produit exactement sur le 6^e atome si l'on débute au carbone adjacent au carbone chiral sur l'acide aminé en position AA9 du macrocycle. Il existe par conséquent une logique parmi les deux séries covalentes, ce qui renforce la crédibilité des résultats. Ce qui a pu être retenu de cette expérience est que la longueur optimale pour les diverses ogives est la suivante : dans le cas propiolamide, le DAB est optimal et pour le bromoacétamide, l'ornithine est optimale.

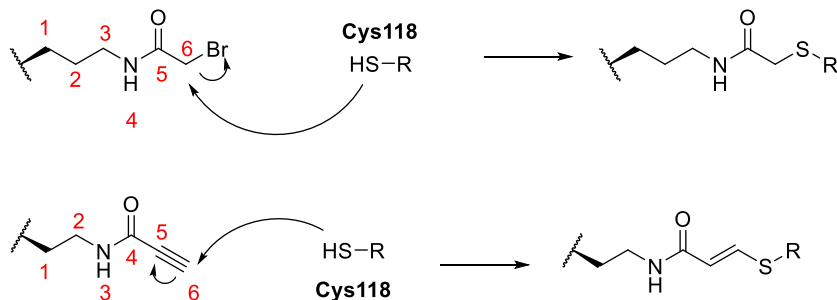


Figure 31 : Relation de la longueur optimale des bras moléculaires en fonction de l'emplacement du site d'attaque nucléophile et le nombre d'atomes.

2.3.2 Effets du nombre de degrés de liberté du tour β

La seconde hypothèse qui pouvait expliquer le décalage spatial des acides aminés entre la boucle FG de NS1 et le macrocycle **401** était la contrainte que pouvait occasionner le tour β . Afin de vérifier cette hypothèse, l'idée a été de synthétiser des macrocycles incorporant 3 tours bêta occasionnant une flexibilité différente. Ces derniers sont l'homo-Aib, l'Aib et la version réduite de l'Aib. Afin d'observer les paramètres constants, uniquement l'ogive bromoacetamide sera synthétisée, et ce, pour deux séries de bras moléculaire, soit la lysine et l'ornithine. Le plan est décrit à la figure 32.

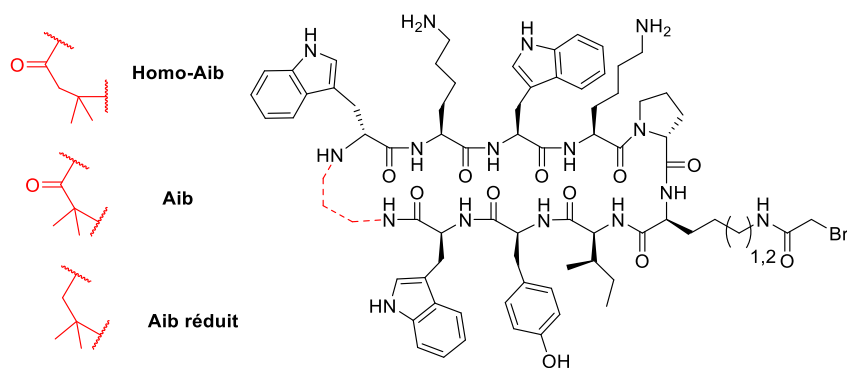


Figure 32 : Plan synthétique des différents tours β qui seront mis de l'avant dans l'étude de la flexibilité

La liaison amide est très intéressante puisqu'elle est planaire et rigide à cause de la résonnance qui survient entre le carbonyle et l'azote. Cette propriété occasionne ainsi une contrainte spatiale en limitant le nombre de degrés de liberté des atomes. Ainsi, on pourrait penser que cette contrainte physique force le macrocycle à adopter une pose précise qui l'empêche d'interagir de façon optimale avec RAS. Dans le macrocycle initial UM0139995, le tour bêta est un homo-Aib : une façon de diminuer le nombre de degrés de liberté consiste donc à éliminer une liaison libre de rotation. Comme la liaison C-C n'a pas de contraintes (sauf en cas d'encombrement stérique), elle permet une rotation de 360°. Ainsi, le remplacement de l'homo-Aib par l'Aib retire le méthylène adjacent à l'amide, ce qui diminue le nombre de degrés de liberté. De plus, ce tour bêta permettra de vérifier l'effet de raccourcir le tour bêta sur la position spatiale des acides aminés adjacents. Finalement, une bonne façon d'augmenter le nombre de degrés de liberté du tour bêta est simplement de retirer la liaison amide. À cette fin, un tour bêta dit Aib réduit sera synthétisé qui offre une rotation NH-C libre.

Une légère différence synthétique a été apportée à la procédure générale de la synthèse des macrocycles en ce qui concerne le tour bêta Aib réduit. Une réaction d'amination réductrice a été insérée lors de la synthèse sur support solide. Pour ce faire, il y a eu déprotection du D-Trp par la pipéridine et ensuite formation de l'imine avec un aldéhyde **26**, qui a ensuite été réduite via le cyanoborohydrure de sodium en condition acide. (Figure 33). Par la suite, le reste de la séquence a été reproduite avec les mêmes conditions. Bien entendu, pour le tour bêta Aib, il y a eu simplement remplacement de l'homo-Aib par l'Aib dans la séquence.

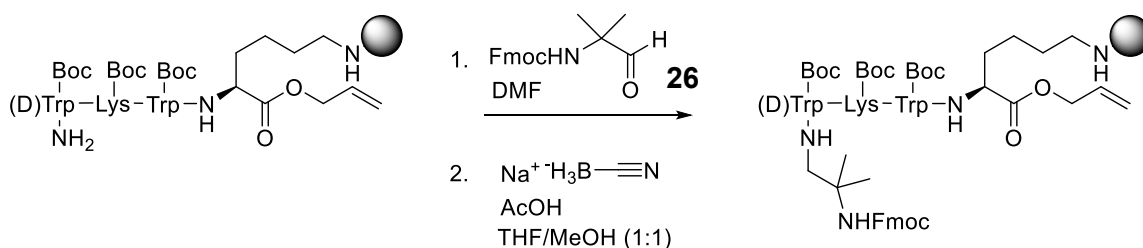


Figure 33 : Réaction d'animation réductrice utilisée dans la synthèse sur support solide pour incorporer le tour β Aib réduit.

À l'exception du Fmoc-Aib-OH, les acides aminés des différents tours bêta n'étaient pas commerciaux : il a donc été nécessaire de les synthétiser. La figure 34 illustre la protection de l'acide 3-amino-3-méthylbutanoïque **27** via le couplage de l'amine et du Fmoc-OSu en présence d'une amine encombrée. Cette réaction a été effectuée dans le dioxane et l'eau durant 16h. Un rendement de 59% a été obtenu où la majorité des pertes est survenue lors de la purification. Cet acide aminé sera désigné comme Fmoc-(h-Aib)-OH **28**.

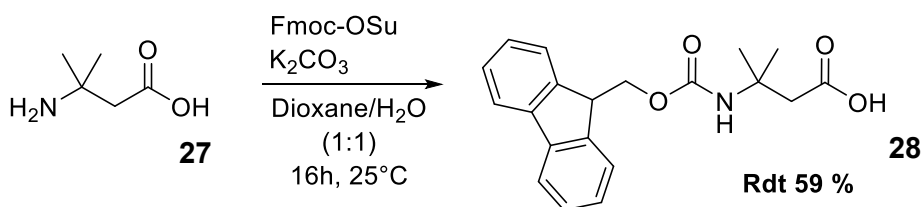


Figure 34 : Synthèse de l'acide aminé Fmoc-(h-Aib)-OH via l'ajout du groupement protecteur Fmoc

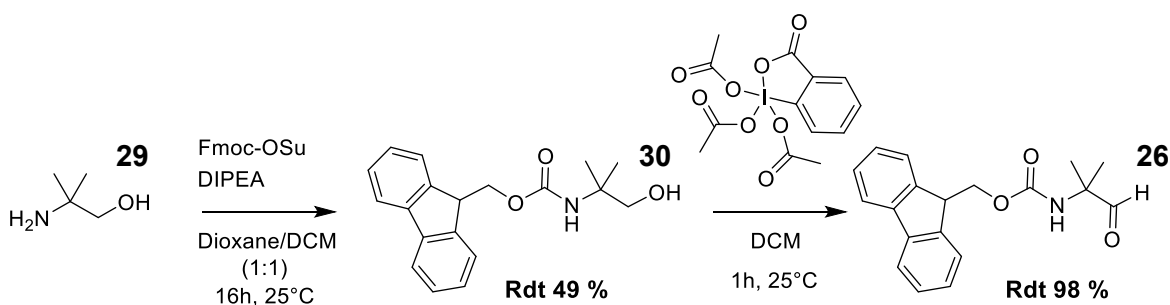
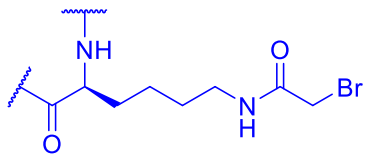
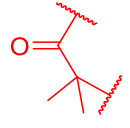
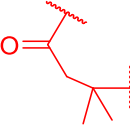
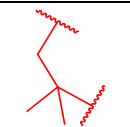
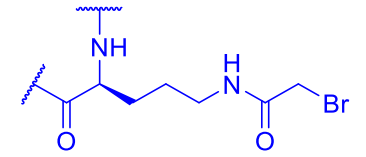
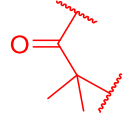
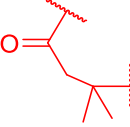
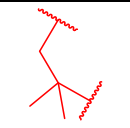


Figure 35 : Synthèse de l'acide aminé **26** via la formation du carbamate Fmoc suivie d'une oxydation douce

Le second réactif qui fut synthétisé est l'aldéhyde précurseur pour l'animation réductive. Pour ce faire le 2-amino-2-méthyl-1-propanol **26** a été protégé par la même réaction discutée plus haut et ensuite une oxydation via le réactif de Dess-Martin a été réalisée pour transformer l'alcool primaire **30** en aldéhyde **26**. Les rendements pour cette séquence étaient nettement meilleurs. Cette molécule a été nommée Fmoc-Aib-H **26**.

Tableau 6 : Rendements globaux et identification des macrocycles synthétisés pour l'étude de la flexibilité du tour β

Chaîne amine et ogive	Tour β	Numéro UM	Rdt global
<p>L-Lysine</p> 		UM0140693	20.4 mg (24.36 %)
		UM0140439	2.91 mg (5.12 %)
		UM0140691	11.8 mg (13.47 %)
<p>L-Ornithine</p> 		UM0140692	19.6 mg (23.62 %)
		UM0140401	11.36 mg (20.16 %)
		UM0140707	7.31 mg (8.89 %)

Le tableau 6 illustre les différents numéros UM des macrocycles synthétisés ainsi que leur rendement. On retrouve ainsi les 3 types de tours β , et ce, sur deux longueurs de bras moléculaire afin d'offrir une meilleure comparaison. Une fois ces macrocycles obtenus, une étude comparative par gel SDS-PAGE a été réalisée de la même façon que dans la section précédente (Figure 36). Encore une fois, l'aide du Dr Lavoie pour la confection de ces gels est à souligner.

Bras moléculaire : *Ornithine*

Bras moléculaire : *Lysine*

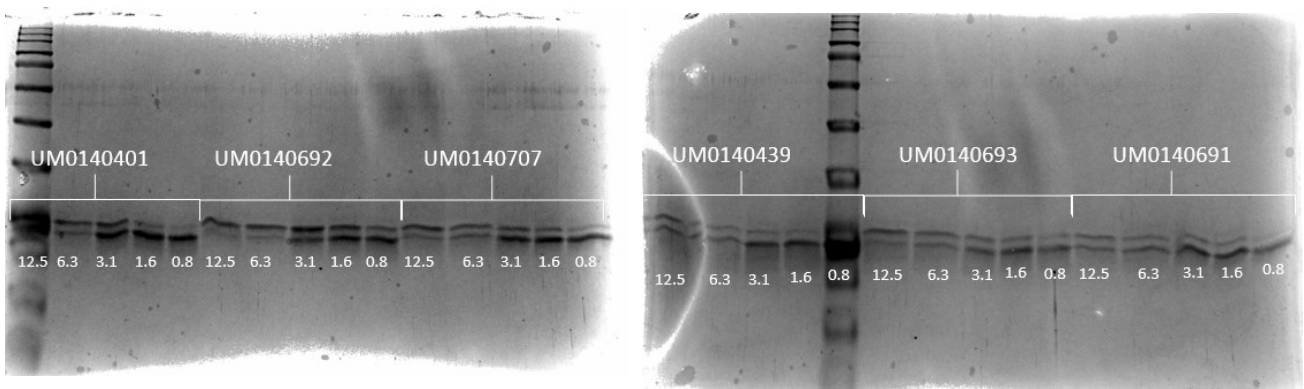
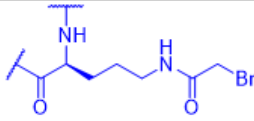
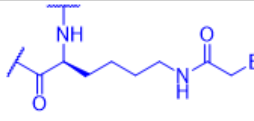

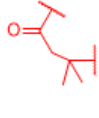
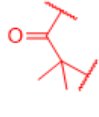


Figure 36 : Exemples SDS-PAGE par coloration au Bleu de Coomassie pour l'étude d'activité des macrocycles covalents afin de déterminer la flexibilité optimale du tour- β

Tableau 7 : Valeurs d'EC₅₀ pour la conversion des protéines HRAS selon les divers tours- β et de longueurs de bras moléculaires.

Bras atomique Tour β										
	Numéro UM	EC ₅₀ (μ M)	Écart-type (μ M)	N=	Valeur p	Numéro UM	EC ₅₀ (μ M)	Écart-type (μ M)	N=	Valeur p
	UM0140707	4.8	1.3	4	0.003	UM0140691	8.2	4.0	4	0.006
	UM0140401	3.7	1.9	5		UM0140439	6.8	2.2	6	
	UM0140692	2.0	0.5	5	0.026	UM0140693	4.1	1.9	4	0.046

La beauté de cette expérience réside dans l'observation d'une même tendance parmi les deux sous-catégories de macrocycles départagés par la longueur du bras moléculaire. Il s'avère que l'ordre d'affinité des macrocycles selon le tour bêta du plus affiné au moins affiné se présente ainsi : Aib > homo-Aib > Aib réduit. Les gels permettent de visualiser cette observation via l'apparition des adduits. Par exemple, pour le composé **UM0140692**, on constate qu'il y a un niveau équivalent d'intensité pour la bande de la protéine et la protéine munie d'un adduit entre 3.1 et 1.6 μ M. Cependant, si l'on regarde pour les deux autres composés, **UM0140401** et **UM0140707**, on observe les doubles bandes à des doses respectives de 3.1 μ M et 6.3 μ M. Ces résultats sont très reproductibles : après plus de 4 gels, la tendance est conservée et les valeurs p démontrent leur significativité. Une autre conclusion rassurante de cette expérience est la différence d'affinité entre la série lysine et ornithine pour le même tour β . Pour chaque tour β , la série ornithine possède une valeur d'EC₅₀ inférieure à la série lysine. Ce résultat va dans le même sens que la conclusion observée dans la section précédente, qui énonçait que la longueur optimale pour l'ogive bromoacétamide est l'ornithine.

En résumé, il est possible de tirer de cette expérience une relation entre degrés de liberté et affinité : plus il y a de degrés de liberté dans le tour bêta, plus l'affinité du macrocycle diminue pour la cible. La perte d'un facteur 2 entre l'Aib et l'espèce exempt du carbonyle illustre clairement cette différence, et ce, dans les deux longueurs de bras moléculaire. On observe cependant une plus mince différence entre l'Aib et son analogue homo-Aib. Il est difficile d'expliquer cette tendance de façon rationnelle, mais on pourrait amener deux explications. Il existe deux grands paramètres très importants pour l'affinité d'une molécule pour une protéine, soit l'entropie et l'enthalpie de liaison.⁶⁷ Quelques tendances simples ont été observées dans la littérature. Tout d'abord, ajouter des interactions hydrophobes ou hydrophiles au composé pour la cible occasionne une enthalpie de liaison favorable.⁶⁷ De plus, augmenter la rigidité d'un composé permet d'augmenter l'entropie de liaison, car la perte de degré de liberté une fois lié dans le site actif sera moins impactante.^{67,68} Par conséquent, il faut trouver un équilibre entre ajouter de la flexibilité dans une molécule pour augmenter les interactions avec la cible et ne pas trop augmenter l'entropie. Ainsi, dans le cas des différents tours β , l'augmentation de la flexibilité n'a pas permis de compenser le gain entropique en allant chercher des contacts enthalpiques. C'est la raison pour laquelle il y a diminution d'affinité. Cet effet est encore plus marqué pour le tour β Aib réduit, car le carbonyle a été éliminé, ce qui empêche la capacité d'interactions avec un donneur de pont H. La conclusion obtenue de cette section est que l'aspect entropique du macrocycle est important pour la liaison, mais il serait intéressant d'avoir une structure cristalline pour observer l'impact de la flexibilité sur la conformation du macrocycle.

2.3.3 Structures cristallines des composés UM401, UM692 et UM6934

Tout en effectuant les analyses biochimiques pour déterminer les impacts des paramètres discutés dans les sections précédentes, les macrocycles ont été remis à Dr Arya pour effectuer de la cristallographie. Avec ses conditions optimisées, il a été en mesure d'obtenir des cristaux suffisamment ordonnés pour être analysés. D'autant plus ces cristaux ont cette fois-ci été envoyés au Synchrotron, qui offre une résolution hors pair pour les structures cristallines. Voici les 3 composés pour lesquelles une structure a été obtenue avec une excellente résolution de 1,5 Å. À cette résolution, il est possible de différencier les atomes de carbones dans un cycle aromatique.

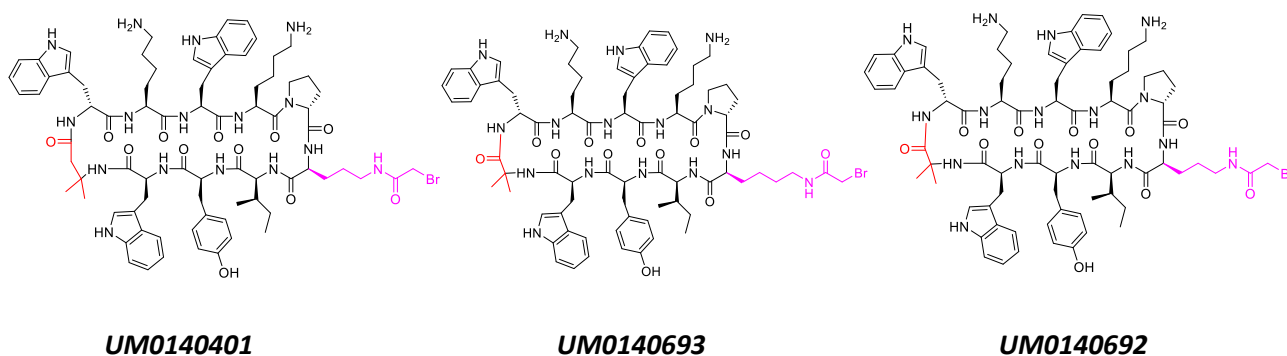


Figure 37 : Structures complètes des composés pour lesquels un cristal a été obtenu. En rose on retrouve les différences de bras moléculaire et en rouge les différences de tour- β entre les macrocycles

Sur la figure 37, les différences notables entre les composés sont soulignées en rouge et en magenta. De façon générale, tous les composés possèdent le bromoacetamide comme ogive, mais présentent une variation dans le bras moléculaire et le tour β . L'acide aminé pour le bras moléculaire est l'ornithine pour le '401 et le '692 et la lysine pour le '693. Finalement, le tour β est l'homo-Aib pour le 401 et l'Aib pour les macrocycles '692 et '693. La première observation qui a été réalisée via la superposition des 3 structures est illustrée à la figure 38.

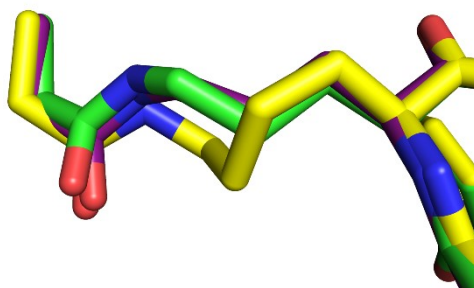


Figure 38 : Superposition des bras moléculaires se liant à la cystéine 118 des 3 macrocycles covalents. La couleur associée à chaque macrocycle est la suivante: En vert le composé '401, en jaune le composé '693 et en magenta le composé '692

La superposition des bras moléculaires illustre très bien pourquoi l'ornithine est idéale pour se lier à la cystéine 118. On constate que le bras est recroquevillé dans le cas de la lysine en jaune, comme si le macrocycle devait s'accommoder de la pose.

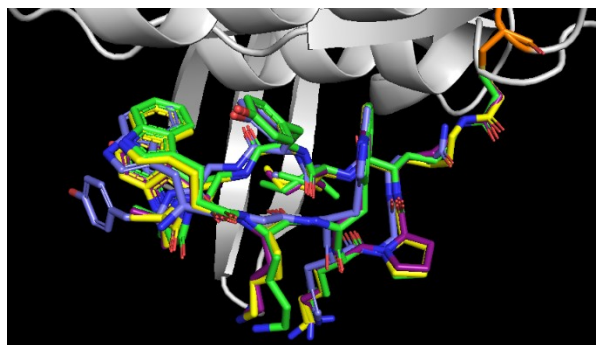


Figure 39 : Superposition des différents macrocycles covalents, lorsque liés à la protéine HRAS. La couleur associée à chaque macrocycle est la suivante: En vert le composé 401, en jaune le composé 693, en magenta le composé 692 et en bleu mauve la boucle FG de NS1.

On peut constater sur la figure 42 qu'il existe toujours un décalage d'environ 1 Å entre le tryptophane de la boucle FG et le tryptophane des macrocycles covalents. Malgré la lysine qui offre un bras atomique plus long, on retrouve quand même le tryptophane décalé. Ceci mène à une conclusion : le décalage ne provient pas du bras atomique trop court,

comme l'a démontré l'ornithine, qui est le bras atomique le plus adapté pour les bromoacetamides.

La seconde hypothèse qui avait été énoncée pour ce décalage était le tour β qui pouvait ne pas être optimal. Malheureusement, le tour β réduit n'a pas réussi à être obtenu pour la cristallographie, mais du moins le Aib et le homo-Aib ont pu être comparés. Grâce au travail de minimisation des angles d'Alexandre Beautrait, il est possible de comprendre pourquoi le Aib est un meilleur tour β .

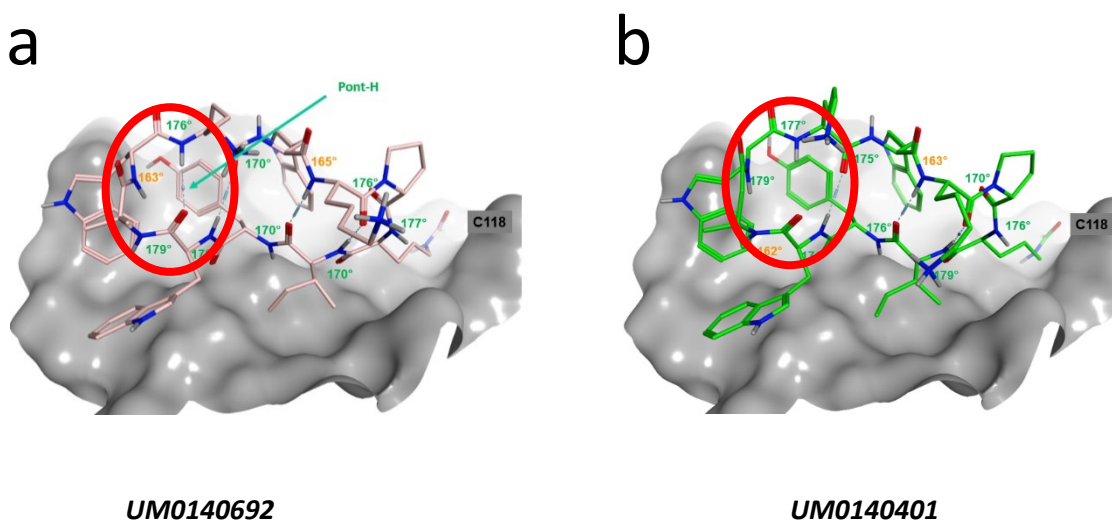


Figure 40 : Minimisation des angles et de la longueur des liens atomique pour diminuer la tension interne des macrocycles via la bio-informatique. Fait en A) avec le composé UM0140692 et en B) avec UM0140401.

Les figures 40a et 40b illustrent les composés covalents **UM0140692** et **UM0140401** avec les angles des liens amides, qui devraient être de 180° pour être optimaux. Avec le composé '692, le Aib permet de faire une liaison hydrogène intramoléculaire sur le carbonyle adjacent, car l'angle est de 179° et la distance est plus courte dans le tour B. Cette interaction hydrogène n'est pas présente dans l'homo-Aib, car le méthylène supplémentaire cause un déplacement dans l'espace et la distance se trouve à être trop grande entre l'azote et le carbonyle. Ce lien hydrogène est très important, car il permet

de réduire considérablement l'entropie du macrocycle et favorise la bonne pose pour interagir avec RAS.

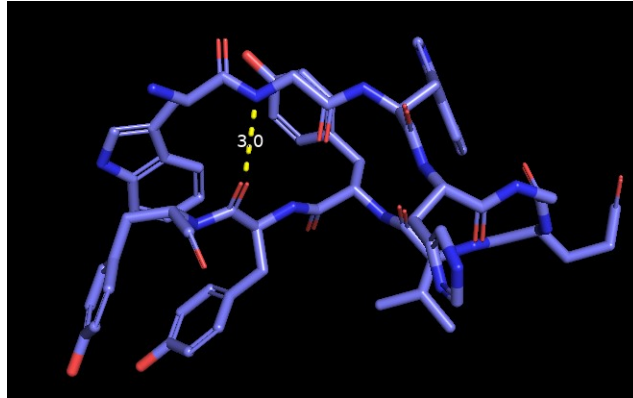


Figure 41 : Interaction hydrogène intramoléculaire présente dans la structure cristalline de NS1

Il s'avère que NS1 possède cette interaction hydrogène, qui semble ainsi primordiale pour obtenir la pose active. (Figure 41) Par conséquent, on peut déduire que la perte du lien amide dans le tour β Aib réduit diminue la rigidité et empêche l'interaction hydrogène, puisque le carbonyle n'est plus présent.

Il est fort à parier que le décalage entre la boucle FG et les macrocycles est lié au tour β . Cependant, il est important de se poser la question suivante : est-ce que cette différence de 1 Å permet d'expliquer la perte de 4 ordres de grandeur d'affinité entre le 'monobody' et le macrocycle? Comme l'illustre la figure 39, tous les autres acides aminés possèdent une excellente corrélation spatiale avec NS1. Il est donc probable que la différence d'affinité provienne d'un facteur différent que le tour β , c'est-à-dire un manque d'interactions et non à cause d'interactions inefficaces.

2.3.4 Détermination de l'affinité des macrocycles covalents

Une question qui a été soulevée durant le processus du développement des macrocycles covalents a été celle de la valeur d'affinité de ces derniers pour la cible HRAS afin d'offrir un aspect comparatif avec les composés non covalents. Cependant, l'analyse de composés covalents est plus complexe et nécessite une attention plus particulière.⁶⁹ L'analogie entre le modèle enzymatique de Michaelis-Menten et le modèle des inhibiteurs irréversible est très similaire. Tout d'abord, il existe deux étapes menant à la liaison irréversible des composés covalents pour leur cible.⁶⁹ (Figure 42)

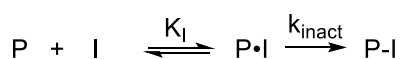


Figure 42 : Équation décrivant le modèle de l'inhibition irréversible. Deux étapes sont à la base du modèle. La première est l'étape d'approche réversible du composé pour la cible (P•I) régie par la constante K_I et la seconde étape est celle de la liaison covalente avec un résidu nucléophile (P-I) qui est décrite par la constante cinétique k_{inact} .

Dans un premier temps, il y a l'approche non covalente du composé pour la protéine P. Ici K_I reflète l'affinité du composé pour la cible en formant le complexe P•I (constante de liaison).

$$K_I = \frac{(k_{off} + k_{inact})}{k_{on}}$$

Figure 43 : Équation décrivant la constante d'affinité apparente dans l'inhibition irréversible où k_{off} est la constante cinétique de dissociation (min^{-1}), k_{on} la constante cinétique d'association ($\text{M}^{-1} \text{min}^{-1}$), K_I est la constante d'affinité d'une sonde covalente non-liée (M) et k_{inact} la constante cinétique de la liaison covalente ($\text{M}^{-1} \text{min}^{-1}$)

Il s'agit alors d'un K_D apparent (constante de dissociation) et non un K_D ! Le K_D est uniquement proportionnel aux constantes cinétiques d'association et de dissociation de la molécule pour la cible.⁶⁹ C'est une distinction primordiale qui implique qu'on ne peut pas comparer une série covalente à une série non covalente. Par conséquent, la valeur de

K_I est influencée par l'affinité non covalente du composé ainsi que par la réactivité de l'ogive et par le temps de la réaction.

La seconde étape du mécanisme est la formation de la liaison covalente décrite par la constante de vitesse k_{inact} . Nécessairement, il s'agit d'une réaction irréversible vers le complexe P-I. (Figure 42) Cette constante cinétique est proportionnelle à la réactivité intrinsèque de l'ogive avec le résidu peptidique faisant la liaison covalente. (Figure 44)

$$v = k_{inact}[P \cdot I]$$

Figure 44 : Équation décrivant la constante cinétique de la réaction irréversible qui est proportionnelle à la constante k_{inact} et la concentration du complexe P•I.

Ainsi les constantes K_I et k_{inact} sont interreliées dans l'analyse de composés covalents. C'est pour cette raison que la meilleure façon de décrire l'efficacité d'un composé covalent est par le ratio k_{inact}/K_I . Plus ce dernier est grand, plus on dit de l'inhibiteur qu'il est efficace.⁶⁹

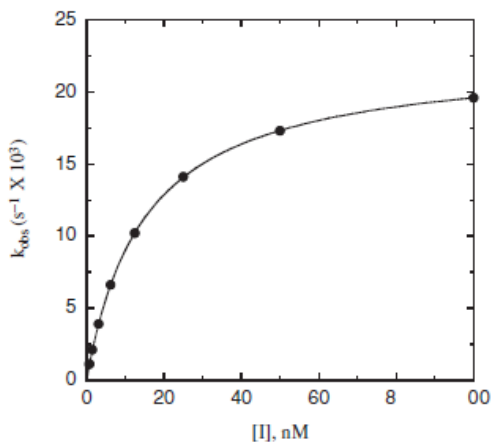
L'équation 45 démontre la relation entre la valeur de la constante cinétique observée expérimentalement et les constantes K_I et k_{inact} .

$$k_{obs} = \frac{k_{inact}[I]}{K_I + [I]}$$

Figure 45 : Équation décrivant la constante cinétique observée expérimentalement qui met en relation les deux variables du modèle de l'inhibition irréversible

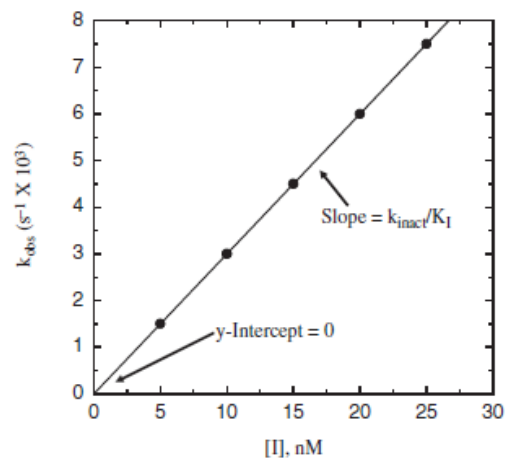
Il est possible d'avoir deux types de scénarios, le premier étant celui où il existe une interaction non covalente précédant l'étape de liaison covalente. Ce scénario est nommé inhibition par affinité. Ce dernier est caractérisé par un graphique exponentiel lors de la

corrélation entre la constante de vitesse expérimentale (k_{obs}) et la concentration d'inhibiteurs. (Figure 46a). Le second scénario, dit non-spécifique, est caractérisé par l'absence de l'étape d'interactions non covalentes et une réaction directe entre l'ogive et les résidus de la cible. Ce cas est observé en présence d'une valeur de K_I très grande, impliquant une très faible affinité. Le graphique de ce scénario donne une régression linéaire lors de la corrélation entre la constante de vitesse expérimentale (k_{obs}) et la concentration d'inhibiteurs. (Figure 46b). C'est de cette façon qu'il est possible d'évaluer la sélectivité de l'inhibiteur. ⁶⁹



Mécanisme A: Inhibition par affinité

$$k_{obs} = \frac{k_{inact}[I]}{K_I + [I]}$$



Mécanisme B: Inhibition non spécifique

$$k_{obs} = \frac{k_{inact}}{K_I} \times [I]$$

Figure 46: Régression mathématique obtenue selon le cas d'inhibition irréversible observée expérimentalement. Le modèle a) illustre une inhibition par affinité, donc implique une dynamique réversible de l'inhibiteur pour sa cible préalable à la liaison covalente. L'équation du modèle prédit une régression logarithmique. Le second modèle b) est celui de l'inhibition non spécifique, qui illustre que l'inhibiteur réagit tellement rapidement avec la cible que l'apport réversible est négligeable. La régression de la constante cinétique est linéaire avec la concentration d'inhibiteurs. Figure proviennent de la référence 68

La façon expérimentale d'obtenir ces graphiques est de mesurer le pourcentage de liaison de la cible avec le composé covalent en fonction du temps, en variant les concentrations. Ainsi, avec l'équation ci-dessous, il est possible d'extraire la valeur de la constante de vitesse expérimentale (k_{obs}).

$$\frac{[Protéine liée]}{[Protéine totale]} \times 100 = \% \text{ de liaison} = 100(1 - e^{-k_{obs} \times t})$$

Figure 47 : Équation décrivant la relation entre le pourcentage de liaison et la constante cinétique observable.

La procédure qui fut utilisée pour obtenir ces constantes expérimentales consistait à effectuer les réactions d'adduits dans des plaques de 96 puits, puis à les transférer dans des gels SDS-PAGE. Ces gels permettaient une quantification selon l'intensité des bandes via le logiciel informatique, Image J. Comme l'intensité des bandes est quantitative, il est possible d'effectuer un ratio et ainsi déterminer un pourcentage de liaison. Les réactions furent débutées en décalage de manière à ce qu'elles se terminent toutes au même moment. Les macrocycles **UM0140401** et **UM0140440** furent testés en parallèle. Pour rappel, ces deux macrocycles possèdent exactement la même séquence d'acides aminés à l'exception de l'ogive qui est différente : bromoacétamide pour **UM0140401** et propiolamide pour **UM0140440**. Les résultats des expériences de cinétique sont illustrés aux figures 48 et 49.

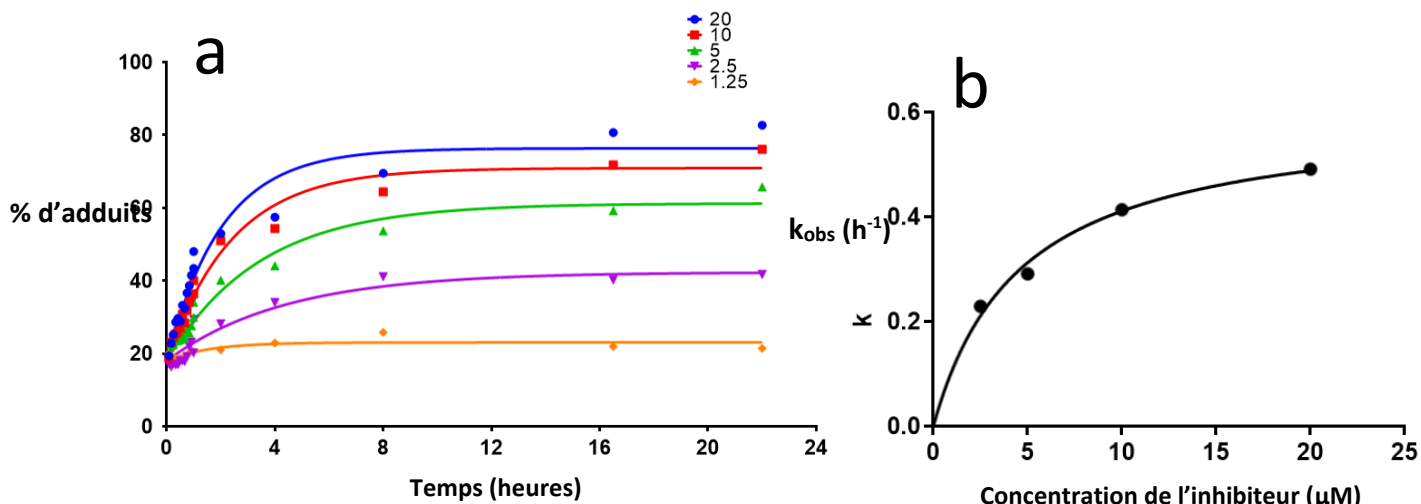


Figure 48 : Graphiques expérimentaux de l'analyse cinétique du macrocycle UM0140401. En a) il s'agit de la détermination des pourcentages d'adduits en fonction du temps et ce pour diverses concentrations de macrocycles (20, 10,5, 2.5 et 1.25 μM). En b) il s'agit des valeurs de k_{obs} de ces diverses courbes de cinétiques en fonction de la concentration de macrocycles.

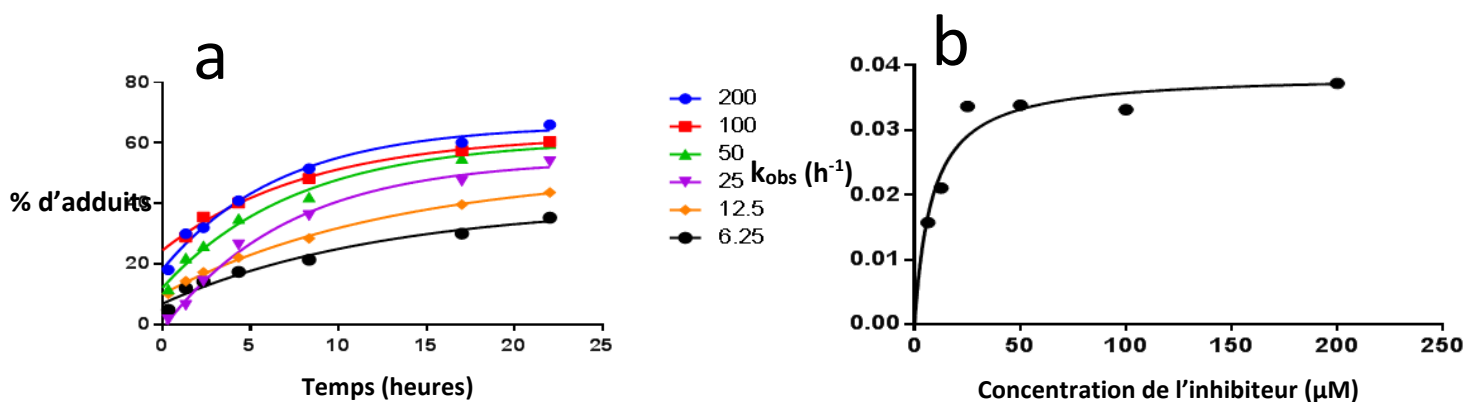


Figure 49 : Graphiques expérimentaux de l'analyse cinétique du macrocycle UM0140440. En a) il s'agit de la détermination des pourcentages d'adduits en fonction du temps et ce pour diverses concentrations de macrocycles (200, 100,50, 25, 12.5 et 6.25 μM). En b) il s'agit des valeurs de k_{obs} de ces diverses courbes de cinétiques en fonction de la concentration de macrocycles.

Tableau 8 : Constantes de l'étude cinétique des macrocycles covalents '440 et '401

Numéro composé	K_i (μM)	k_{inact} (h^{-1})	k_{inact}/K_i ($\text{M}^{-1}\text{s}^{-1}$)
UM0140401	1.2	0.15	36
UM0140440	2.1	0.0096	1.3

Le premier constat est qu'il s'agit de courbes exponentielles. Par conséquent, le scénario cinétique auquel nous faisons face pour ces inhibiteurs covalents est un modèle d'inhibition par affinité. Par la suite, il a été possible d'obtenir les valeurs des constantes du modèle covalent que sont K_i et k_{inact} . Il s'avère que les deux composés possèdent des valeurs très similaires de K_i , mais de très grandes différences de k_{inact} . Le bromoacétamide possède une valeur de 0.15 h^{-1} comparativement à 0.0096 h^{-1} pour le propiolamide. Ceci est en accord avec la littérature et les résultats précédemment énoncés qui indiquent que le bromoacétamide est une ogive plus réactive, donc il est logique que la vitesse de conversion k_{inact} soit 15 fois plus grande.

Passons maintenant à l'analyse des valeurs d'équilibre K_i obtenues pour les deux espèces. On constate que ces valeurs sont extrêmement similaires, soit à $1 \mu\text{M}$ de différence. Ceci semble très surprenant, puisque dans le calcul du K_i on retrouve l'impact du k_{inact} et ces valeurs étaient très différentes entre les deux composés. Ceci mène alors à deux conclusions possibles : la première est que si le k_{on} est très grand, alors l'impact de k_{inact} sur le K_i est négligeable, ce qui explique les valeurs très similaires entre les deux macrocycles. Le second cas de figure possible est une valeur très grande de k_{off} , ce qui résulte à la même conclusion. Si on se réfère aux résultats de titrage obtenue par RMN, le macrocycle non covalent UM0139995 possédait une affinité de $100 \mu\text{M}$, ce qui est très faible. Par conséquent, il est fort probable que le k_{off} soit très grand à cause d'un manque d'interactions entre le macrocycle et la protéine HRAS. Ainsi, ceci explique la raison d'une valeur similaire entre les deux K_i obtenues malgré un facteur de 15 entre les k_{inact} . Maintenant, cherchons à évaluer ce qu'une valeur de $1\text{-}2 \mu\text{M}$ signifie. Il est très dangereux

de comparer directement ces valeurs avec l'affinité obtenue par la série non covalente puisqu'il y a comparaison d'un K_D apparent à un K_D . Si on revient avec l'approximation que le k_{off} est très élevé pour les macrocycles, il est possible d'approximer qu'on est en situation de K_D et non de K_D apparent.

$$K_I = \frac{(k_{off} + k_{inact})}{k_{on}} \quad \text{si } k_{off} \gg k_{inact} \text{ alors } K_I = \frac{k_{off}}{k_{on}}$$

Figure 50 : Approximation qu'il est possible de faire si la constante cinétique de dissociation k_{off} est très grande.

Dans ce scénario, il aurait fallu obtenir un K_I de l'ordre de 100 μM et non de 1-2 μM . Il s'avère que le choix de l'essai joue pour beaucoup. Comme démontré dans la littérature, il existe de grandes variabilités dans le type d'essai utilisé pour déterminer les valeurs de K_I , k_{inact} et K_D .⁷⁰ Il est possible d'obtenir un ordre de grandeur fiable, mais loin d'une valeur exacte entre deux méthodes. Rappelons que le K_I obtenu pour la série non covalente fut réalisé par un titrage RMN et que le K_I obtenu pour la série covalente, par un essai biochimique. Pour la situation covalente, les résultats obtenus se situent dans l'ordre du micromolaire, ce qui correspond avec les résultats précédemment obtenus par l'équipe de recherche pour la situation non-covalente, se situant eux aussi dans le micromolaire.

Finalement si on compare les valeurs du ratio k_{inact}/K_I avec la littérature, il est possible d'arriver aux mêmes constats. Dans une optique d'optimisation de composés covalents se liant à une lysine de HSP72, le groupe du Pr Cheeseman⁷¹ a débuté son évaluation avec un composé très peu affin soit 18 μM et possédant un ratio de 0.86 $\text{M}^{-1}\text{s}^{-1}$. Après de multiples optimisations, ils ont augmenté le ratio de 108 fois, obtenant ainsi une valeur de 4.7 $\text{M}^{-1}\text{s}^{-1}$, qui possède maintenant un K_I de 4.7 μM . Ces valeurs illustrent l'ordre de grandeur auquel il faut s'attendre avec un composé d'affinité dans l'ordre du micromolaire. Maintenant, si l'on prend une étude réalisée par le Pr Murray⁷², qui cette fois-ci étudie des inhibiteurs covalents de récepteurs EGFR nettement plus actifs, les

valeurs de ratio $k_{\text{inact}}/K_{\text{I}}$ se situent dans un ordre de grandeur $10^{5-7} \text{ M}^{-1}\text{s}^{-1}$ pour des composés ayant une affinité dans l'ordre du nanomolaire. Ceci met en relief l'ordre de grandeur qu'il faut s'attendre avec les affinités des composés et leur ratio $k_{\text{inact}}/K_{\text{I}}$. Comme le tableau 8 l'illustre, les valeurs obtenues pour les ratios $k_{\text{inact}}/K_{\text{I}}$ des macrocycles sont de 36 et $1.3 \text{ M}^{-1}\text{s}^{-1}$, donc nécessairement de l'ordre de grandeur discuté par l'étude du Pr Cheeseman. Ceci est une preuve qui appuie la conclusion que l'affinité K_{I} des macrocycles est de l'ordre du micromolaire. Comme le ratio est plus élevé pour le bromoacétamide que le propiolamide, on conclut que le bromoacétamide est plus réactif, ce qui est en accord avec le reste.

Ce que cette expérience a permis de confirmer est que l'affinité du macrocycle est le talon d'Achille pour la liaison de ce dernier à la cible. Sachant que la réactivité intrinsèque du bromoacétamide est extrêmement puissante, obtenir une faible valeur de k_{inact} présage une très faible affinité. Cette donnée est confirmée par un K_{I} similaire entre les deux ogives moléculaires qui illustre un k_{off} très grand et un K_{I} de l'ordre du micromolaire.

Une seconde expérience a été réalisée pour voir l'impact de l'aspect non covalent du macrocycle sur la liaison covalente. Comme les ogives ayant réagi sur la cystéine 118 sont réputées pour être très réactives, il est normal de se demander si les adduits observés résultent d'une liaison non spécifique. Bien que le modèle cinétique démontre que c'est un scénario d'affinité, les résultats de spectrométrie de masse ont démontré plusieurs adduits démontrant un doute sur la spécificité des ogives. D'après des expériences effectuées par Hugo Lavoie et Marc-André Poupart, il a été observé que changer l'Allo-Ile du macrocycle par une simple alanine causait une perte drastique d'affinité du macrocycle. L'idée a donc été de synthétiser les analogues des composés **UM0140401** et **UM0140440**, ayant une alanine à la place de l'Allo-Ile. Par conséquent, si on voit toujours une liaison sur la protéine HRAS, cela signifie que l'ogive est beaucoup trop réactive et que les interactions non covalentes n'ont pas d'impact.

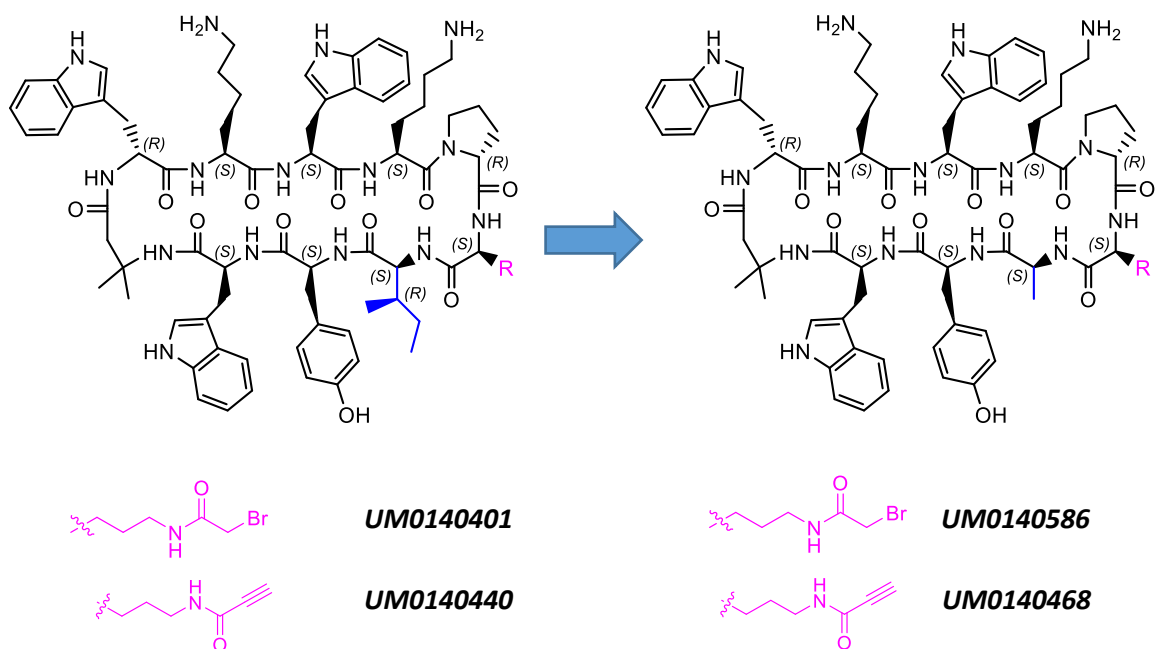


Figure 51 : Plan synthétique afin de transformer les macrocycles covalents actifs en macrocycles covalents inactifs. Le changement de l'Allo-Ile par l'alanine tue l'activité, donc les composés UM0140586 et UM0140468 sont les analogues inactifs des macrocycles UM0140401 et UM0140440.

Les composés **UM0140586** et **UM0140468** furent synthétisés selon la procédure générale décrite dans ce manuscrit en effectuant la substitution de l'Allo-Ile par l'alanine dans la séquence peptidique.

Tableau 9 : Rendements globaux et identification des macrocycles inactifs via la substitution de l'Allo-Ile par l'alanine

Chaîne amine	Ogive	Numéro UM	Rdt global
		UM0140468	1.00 mg (2.96 %)
		UM0140586	1.82 mg (4.74 %)

Le premier test réalisé pour vérifier le résultat fut celui par gel SDS-PAGE. Trois conditions furent analysées pour chaque macrocycle en faisant varier le ratio de concentrations protéine/macrocycle. Les trois ratios évalués sont les suivants (macrocycle/protéine) 25/1, 8.25/1 et 2.5/1 et les réactions ont été réalisées sur 16 heures à température de la pièce. La figure 52 illustre le résultat de cette expérience.

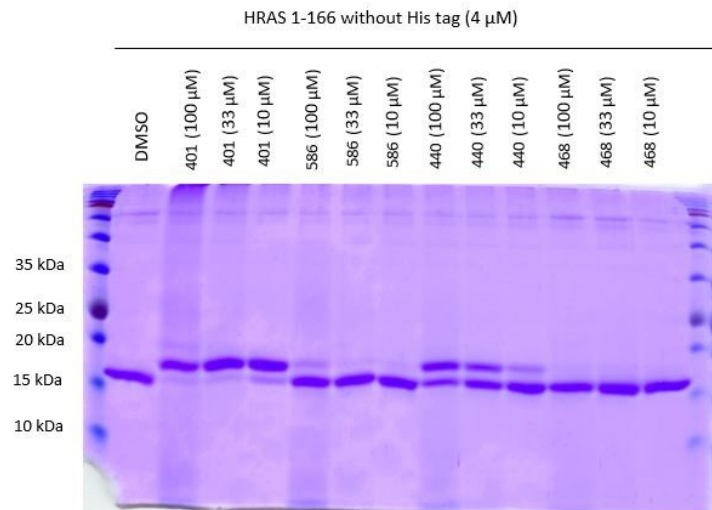


Figure 52 : Analyse de l'importance des interactions non covalentes du macrocycle sur l'affinité via l'utilisation de macrocycles covalents inactifs par gel SDS-Page.

On constate qu'avec le composé 401, il y a conversion complète vers une population HRAS liée pour les ratios 25/1 et 8.25/1, mais il reste un peu de la protéine HRAS non liée avec le plus faible ratio. Si l'on compare avec son analogue inactif, le composé **UM0140586**, il est possible d'observer une faible bande correspondant à HRAS-liée avec le macrocycle avec la condition 25/1. Cependant, pour les deux autres ratios, il ne semble pas y avoir de liaisons. Cette observation est rassurante, puisqu'elle démontre que malgré une ogive très réactive, les interactions non covalentes sont importantes pour l'obtention de la réaction. Cependant, il faut garder en tête qu'il y a aussi un peu de réaction non spécifique. Pour la seconde famille de composés, les propiolamides, on peut constater que l'analogue actif (**UM0140440**) forme des adduits aux trois ratios, mais ce n'est pas une conversion complète. L'analogue inactif du propiolamide ne démontre aucune bande

supérieure aux trois ratios, reproduisant ainsi le résultat observé par les analogues bromoacétamides. Ainsi, le propioloamide est beaucoup plus sélectif que le bromoacétamide.

La même expérience a été reproduite, mais cette fois-ci les échantillons ont été analysés par spectrométrie de masse sur protéines intactes, selon la même procédure décrite et utilisée plus tôt dans ce manuscrit. Ces résultats sont illustrés dans le tableau 10.

Tableau 10 : Efficacité de liaisons des macrocycles covalents actifs et inactifs par spectrométrie de masse sur HRAS

Composé	Ratio [Macrocycle]/[HRAS]	%P	%S	%D
UM0140401	25/1	0	92	8
	8.25/1	0	100	0
	2.5/1	17	83	0
UM0140586	25/1	>75	<25	0
	8.25/1	>90	<10	0
	2.5/1	100	0	0
UM0140440	25/1	18	82	0
	8.25/1	40	60	0
	2.5/1	67	33	0
UM0140468	25/1	100	0	0
	8.25/1	100	0	0
	2.5/1	100	0	0

Légende : %P est le pourcentage de protéines non modifiées

%S est le pourcentage de protéines possédant 1 adduit

%D est le pourcentage de protéines possédant 2 adduits

La même tendance s'observe en gel et en spectrométrie de masse. Il y a une très grande diminution du pourcentage de simple adduits lorsque l'on effectue le changement à l'acide aminé alanine en position AA9. Dans le cas des bromoacétamides, avec le ratio 25/1, il y a une perte de 25 % et de 90 % avec le ratio 8.25/1. Ainsi, la non-spécificité a elle aussi été observée par spectrométrie de masse. Finalement, pour les propiolamides, il y a observation d'aucun adduit avec l'espèce alanine à tous les ratios macrocycle/protéine. La différence de réactivité entre les ogives permet d'expliquer la non-spécificité : il est logique qu'une ogive plus réactive, présente des probabilités plus élevées de réagir avec d'autres sites nucléophiles.

Ce qu'il est possible de conclure de ces expériences, est que les interactions non covalentes du macrocycle sont essentielles pour obtenir la spécificité sur la cible HRAS. Les études cinétiques et par gels ont bien illustré que le modèle d'inhibition est par affinité et que le bromoacétamide est plus réactif que le propiolamide. Cependant, ce dernier est un peu trop réactif au point de procurer de la non-spécificité sur HRAS, et ce, malgré un analogue inactif dans la série de macrocycles non covalents développée par Dr Poupart. Ceci va de pair avec les résultats obtenus avec les mutants HRAS C118S, qui démontraient toujours des adduits malgré la mutation de la cystéine avec les analogues bromoacétamides. Ainsi, dans un projet d'inhibition covalente, il est important de jouer sur plusieurs paramètres pour réduire la non-spécificité. Il faut augmenter l'affinité du composé pour la cible et être en mesure de moduler la réactivité de l'ogive pour être suffisamment réactif pour se lier, sans trop être trop réactif pour causer de la non-spécificité. Ainsi, les composés propiolamides semblent être une meilleure série pour entreprendre des essais cellulaires afin d'évaluer la capacité thérapeutique des macrocycles sur RAS.

2.4. Conclusions et perspectives du projet

La raison d'être de ce projet trouve son origine dans une incohérence d'affinité entre le monobody NS1 et la série des macrocycles non covalents cherchant à émuler l'effet de ce dernier. Afin d'éclaircir cette différence d'affinités, des structures cristallines ont été obtenues grâce à l'utilisation d'outils moléculaires (macrocycles covalents) se basant sur les macrocycles non covalents. Le développement de macrocycles covalents a été fait via une synthèse sur support solide utilisant une approche faisant intervenir 4 groupes protecteurs orthogonaux qui ont permis une fonctionnalisation personnalisée. De cette fonctionnalisation, il a été possible d'obtenir deux ogives moléculaires suffisamment réactives, bromoacétamide et propiolamide, permettant de venir former un lien covalent avec la cystéine 118 de la GTPase HRAS. Ces adduits ont été identifiés par spectrométrie de masse et gels biochimiques. Une optimisation de la longueur du bras moléculaire pour chacune de ces familles d'ogives a illustré que le lieu d'attaque dans l'espace était le même. Autrement dit, comme les deux ogives n'ont pas le même nombre d'atomes il a fallu compenser avec la longueur du bras moléculaire.

Les structures cristallines ont permis de relever une différence spatiale de quelques Angström entre les acides aminés de la boucle FG de NS1 et ceux des macrocycles, laissant présager une perte d'affinité selon la flexibilité de ces derniers. Ainsi, une étude de flexibilité d'un tour β fermant les macrocycles a été réalisée pour voir l'effet du nombre de degrés de liberté sur l'affinité des macrocycles pour la cible. Une augmentation de la rigidité s'est avérée bénéfique en diminuant l'entropie des macrocycles.

Grâce à la synthèse de ces divers macrocycles covalents, il a été possible de réaliser une étude cinétique entre deux macrocycles munis d'ogives différentes. Celle-ci a permis dans de déterminer l'affinité de l'ordre du micromolaire des macrocycles covalents concordant avec les valeurs des macrocycles non covalents préalablement obtenues par le titrage RMN.

L'aspect visuel des structures cristallines a permis d'obtenir plusieurs réponses, notamment que les macrocycles épousent relativement bien la conformation de la boucle

FG dans NS1. Ce résultat démontre que le macrocycle se lie au bon site et surtout que le choix des acides aminés du squelette a été fait judicieusement. Cependant, ce résultat illustre aussi un constat très important : la très grande différence d'affinité entre NS1 et le macrocycle non covalent ne provient pas d'un mauvais mimétisme de la boucle FG, mais plutôt d'un manque d'interactions avec la cible. Les analyses cinétiques et l'expérience du changement de l'allo-Ile par l'alanine ont clairement illustré le caractère sensible des interactions du macrocycle pour HRAS et l'affinité obtenue.

Il faut se rappeler que l'objectif à long terme de ce projet est d'obtenir un composé pouvant émuler les effets de NS1 en cellule, mais possédant des propriétés médicamenteuses plus favorables. Il s'avère qu'à ce stade, la série non covalente ne permet pas de vérifier les effets de NS1 en cellules à cause de sa faible affinité se situant dans l'ordre du micromolaire. L'idée d'utiliser les macrocycles covalents pour voir l'effet dans les cellules fait partie des projets futurs de l'équipe afin de confirmer le potentiel thérapeutique de cette stratégie.

La résultante de ce projet fut très positive compte tenu que le développement des outils moléculaires covalents ont permis de répondre aux questionnements initiaux qui ont mené à ce projet. Les macrocycles covalents ont permis l'obtention de cristaux et de relever le manque d'interactions envers RAS comme maillon faible pour la perte d'affinité vis-à-vis NS1. Cependant, un grand défi qui se pose est la non-sélectivité des meilleurs candidats covalents obtenus. Il s'avère que les ogives utilisées sont très réactives et donc peu sélectives tel qu'ont déjà pu le démontrer les bromoacétamides. Si on souhaite utiliser une série covalente en milieu cellulaire, il faut être en mesure de diminuer grandement la non-sélectivité et, conséquemment, la réactivité des ogives. On se retrouve ainsi à la même problématique : pour diminuer la réactivité des ogives, il faut augmenter l'affinité des macrocycles. Comme discuté dans un chapitre précédent, pour augmenter l'affinité d'un composé, il faut soit augmenter la rigidité ou augmenter le nombre d'interactions hydrophiles/hydrophobes. Par conséquent, le futur de ce projet s'articule autour des pistes permettant l'obtention de nouvelles interactions entre RAS et les macrocycles.

La principale différence entre NS1 et les macrocycles est la perte des interactions ioniques formées par la triade d'acides aminés que sont E49, R135 et Y82. Les interactions ioniques sont les forces les plus fortes et c'est ce qui explique probablement la grande différence d'affinité entre le 'monobody' NS1 et le macrocycle.²⁷

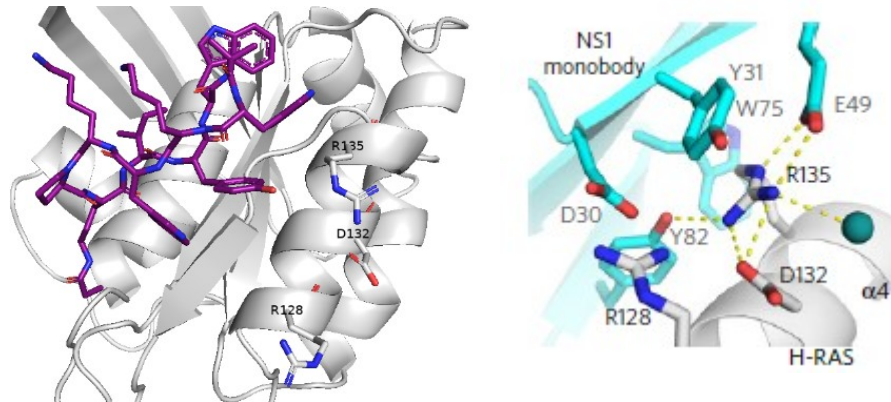


Figure 53 : Structure cristalline d'UM0140692 sur HRAS en observant spécifiquement les interactions avec l'hélice alpha 4. Il manque des interactions non covalentes du macrocycle UM0140692 avec l'arginine 135 de HRAS. Pour rappel, NS1 effectuait une triade d'interactions ioniques avec l'arginine 135.

Tel qu'illustre la figure 53, le macrocycle '692 ne présente aucune interaction entre la tyrosine et l'arginine 135. Comme le macrocycle ne possède pas l'acide glutamique 49 que possède NS1, l'arginine 135 n'est pas orientée en direction de la tyrosine. De plus, il y a aussi une autre interaction ionique entre l'acide glutamique 30 et l'arginine 128 qui est perdue dans les macrocycles. Ces pertes sont considérables dans l'affinité du macrocycle pour HRAS. Il ne serait pas surprenant que ces pertes soient responsables de la perte d'affinité de 4 ordres de grandeur, étant donné qu'une interaction ionique est l'interaction non covalente la plus forte possible.

Une possibilité pour aller chercher l'arginine 135 serait de rapprocher la charge négative de l'arginine 135 à partir de la tyrosine du macrocycle.

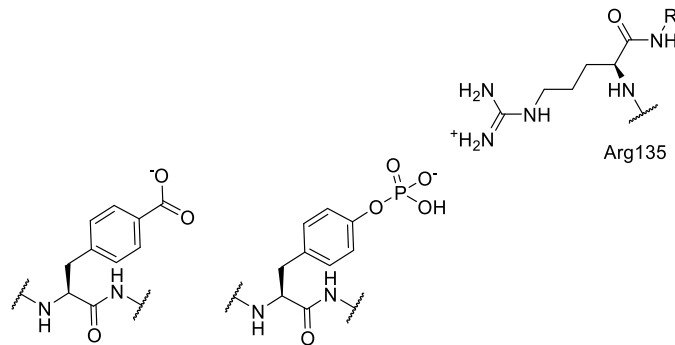


Figure 54 : Suggestions synthétiques cherchant à venir interagir avec l'arginine 135 de HRAS via la tyrosine du macrocycle.

L'idée serait ainsi de venir ajouter un groupement phosphate à la tyrosine qui serait négative et pourrait interagir avec l'arginine. Un second groupement qui pourrait offrir une charge négative serait un carboxylate. Ces deux options seraient de bons choix pour chercher à mimer cette interaction ionique perdue par rapport à NS1. Pour ce qui est de l'interaction avec l'arginine 128, elle est beaucoup plus complexe à optimiser, car elle se situe à une distance très éloignée des chaînes latérales du macrocycle. Pour ce faire, il serait nécessaire d'augmenter la grosseur du macrocycle, donc augmenter l'entropie, ce qui risque de diminuer l'affinité au bout du compte. Une seconde possibilité à exploiter pour augmenter les interactions serait d'interagir avec l'acide glutamique 162 via le tryptophane du macrocycle.

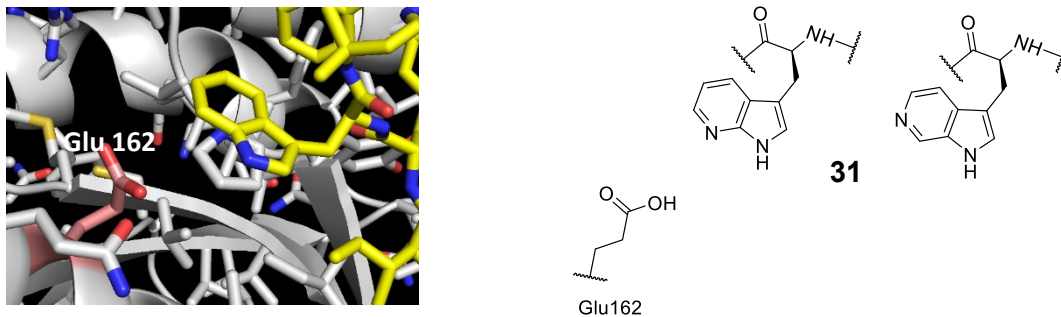


Figure 55 : Idée synthétique qui permettrait d'interagir entre l'acide glutamique 162 (Rose) et un tryptophane du macrocycle selon la structure cristalline de 401. L'introduction d'un azatryptophane serait intéressante à cause de sa polarité.

Avec les avancées de l'évolution dirigée enzymatique, il existe de plus en plus de façon de synthétiser des acides aminés non-protéionogènes. Le tryptophane est synthétisé par un complexe de protéines, tryptophane synthase, en utilisant la sérine et un indole chez les bactéries. Par la méthode développée par Pre Arnold, il a été rapporté dans la littérature qu'il est possible de muter cette enzyme de façon à offrir une flexibilité dans le site actif et à accueillir différent type d'indoles et former l'acide aminé correspondant.⁷³ Des azatryptophanes **31** pourraient être de bons candidats pour offrir un accepteur de pont hydrogène à l'acide glutamique 162 et ainsi offrir une nouvelle interaction.

La clé du succès de ce projet réside dans l'augmentation de l'affinité du macrocycle afin de permettre de vérifier l'effet biologique observé par NS1. À ce moment, la série covalente nécessite des ogives moléculaires trop réactives pour se lier avec HRAS, il est donc important de diminuer la réactivité de ces derniers pour obtenir une sélectivité en cellule. Une cinétique d'échange entre le GTP et le GDP a été réalisée en présence des composés covalents : aucun effet n'a été observé. Ce qui laisse démontrer que la cystéine 118 ne possède pas d'effet stabilisant sur le nucléotide malgré sa proximité. Par conséquent, cette possibilité thérapeutique ne semble pas démontrer un attrait. Aucun essai n'a été tenté concernant la capacité des macrocycles covalents à bloquer une interaction avec des effecteurs *in vitro*. Ces expériences seront dans les priorités futures du groupe de recherche.

3. Développement de sondes photoactivables de UM171 pour l'identification de cibles protéiques

3.1 Synthèse des sondes

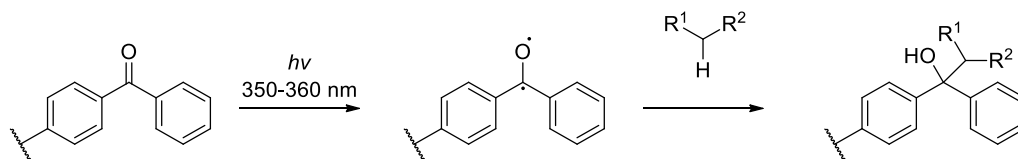
3.1.1 Conception des sondes

Tel que mentionné dans l'introduction, l'approche ABPP nécessite deux fonctions clés pour l'identification d'une cible : une fonction formant la liaison covalente à la cible et une fonction qui permet l'analyse.

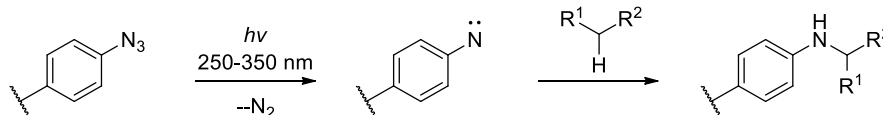
L'utilisation de fonctions photoactivables permet un avantage considérable par rapport aux ogives moléculaires : il est possible d'augmenter la probabilité de lier des protéines d'intérêts, et ce, peu importe les acides aminés présents dans le site actif. L'utilisation d'une ogive moléculaire nécessite un résidu nucléophile à proximité du site actif. Comme la première partie du mémoire a pu le démontrer, il est essentiel d'avoir une position spatiale optimale pour se lier, rendant l'opération complexe. Cette complexité est accrue par le fait que la cible étant inconnue, il n'est pas possible de savoir si le nucléophile sera présent. De plus, l'ajout d'une ogive trop réactive va enrichir beaucoup de protéines non-spécifiques et compliquera l'analyse finale par la spectrométrie de masse.

Pour ces diverses raisons, le choix d'ogives moléculaires n'est pas préférable pour l'identification de cibles. Les fonctions photoactivables offrent un profil plus attrayant par leur spécificité. Parmi les plus communément utilisées, on retrouve la diazirine, la benzophénone et l'azoture. (Figure 56)

Benzophenone



Azoture aromatique



Diazirine

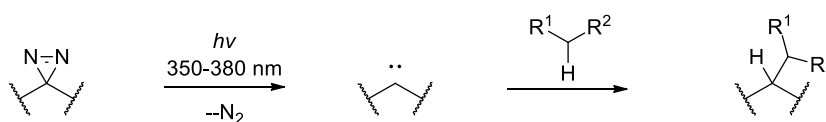


Figure 56 : Structure chimique des fonctions photoactivables présentes dans l'approche ABPP ainsi que leur longueur d'onde d'activation et de leur insertion dans les acides aminés

La fonction diazirine sera choisie par rapport aux deux autres possibilités à cause des avantages qu'elle procure.⁵² La benzophénone est un groupement très encombré pouvant influencer négativement l'affinité du composé pour la cible. De plus, le diradical formé par le rayonnement UV est stabilisé par la résonance des groupements aromatiques, ce qui ralentit la réaction d'insertion. Cette conséquence occasionne plus de liaisons non-spécifiques avec des protéines indésirables. Bien que l'azoture soit un très petit groupement, son activation nécessite une irradiation à 250-350 nm. Il faut savoir qu'une irradiation à cet intervalle de longueur d'onde occasionne des dommages aux matériels cellulaires. De plus, le nitrène a tendance à se réarranger très rapidement, limitant ainsi la disponibilité du radical pour l'insertion dans les acides aminés de la cible. C'est à cause de son excellente stabilité chimique, son encombrement stérique minime, son carbène très réactif qui ne se réarrange pas et sa longueur d'onde élevée que la diazirine a été sélectionnée comme groupement photoactivable de prédilection. Cependant, un inconvénient des diazirines est que le carbène est désactivé rapidement par le solvant.⁷⁴

Finalement, il est possible de choisir une diazirine aliphatique ou une diazirine aromatique. Il a été démontré que les diazirines aliphatiques avaient tendance à réagir intramoléculairement,

mais leur synthèse est beaucoup plus simple.^{51, 75, 76} L'une des façons de contrer cet effet négatif consiste à stabiliser le carbène par un groupement électroattracteur et c'est pour cette raison que les diazirines aromatiques possèdent un groupement CF₃.⁷⁷ Cette propriété leur permet de réagir beaucoup plus sélectivement avec les acides aminés adjacents et le carbène est inactivé beaucoup moins rapidement. Il y a aura utilisation de ces deux types de diazirines pour maximiser l'espace chimique compte tenu qu'aucune information du site actif de la cible est connue.

Toujours dans l'optique de conserver le plus possible l'affinité des sondes pour leur cible, la seconde fonction qui a été sélectionnée est l'alcyne. Cette fonction est très petite et accessible commercialement. De plus, l'alcyne permet la cycloaddition de Huisen qui fait partie de la chimie Click.⁴⁸ L'alcyne a la capacité de réagir sélectivement dans un milieu physiologique avec l'azote au moyen d'une réaction exothermique relarguant de l'azote moléculaire. Les rendements de cette réaction sont excellents et il s'agit d'une réaction robuste. Le choix de l'alcyne s'avère un choix sensé pour le développement des sondes de UM171 pour ces diverses raisons.

Dans un cheminement rationnel pour le choix de l'emplacement des groupements fonctionnels nécessaire pour la chimioprotéomique, il est essentiel d'effectuer une étude de relation structure-activité. Il existe une dualité importante dans la conception d'une sonde lorsque la cible est inconnue. Tout d'abord, il est important de conserver une bonne affinité du composé pour la cible sans quoi il sera nécessaire d'augmenter la concentration cellulaire des sondes pour obtenir une réponse cellulaire similaire. Cette conséquence occasionne une augmentation des protéines non spécifiques conduisant ainsi à une analyse beaucoup plus complexe. La seconde partie de la dualité concerne l'emplacement de la diazirine par rapport à la cible. Si la modification n'influence pas l'affinité de la sonde pour la protéine, il est fort à parier que le groupement photoactivable se trouve dans un endroit loin du site actif, par exemple dans une région exposée au solvant. Cette situation occasionne un second problème : la sonde sera consommée par le solvant et non par la protéine d'intérêt. Ainsi, il est primordial de synthétiser de multiples sondes variant l'emplacement de la fonction photoactivable pour jouer sur cette dualité entre affinité et emplacement spatial. Par chance, une étude structure-activité préalable a été effectuée par divers chimistes de la plateforme qui sont Yves Gareau, Stéphane Gingras, Réjean Ruel et Yves Chantigny. Cette étude a permis d'identifier les sites permissifs qui peuvent accommoder une fonctionnalisation sur la molécule UM171. La figure 57 illustre sommairement les conclusions observées par l'étude RSA.

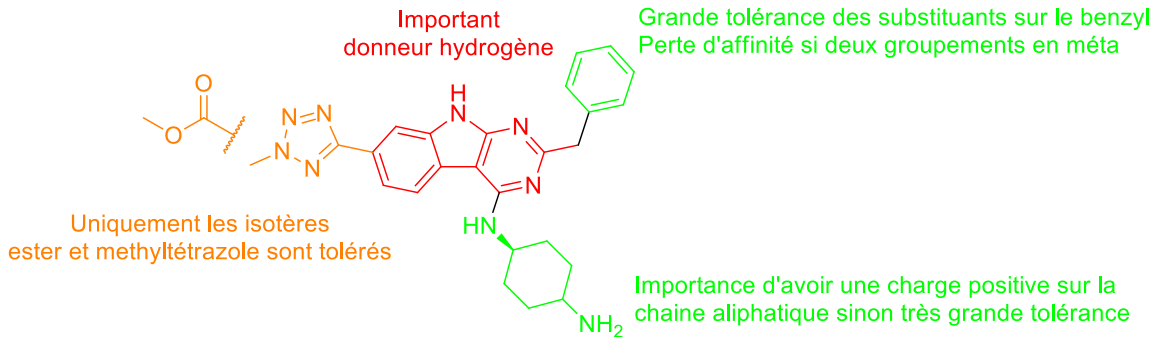


Figure 57 : Résumé des observations générales concernant l'étude de relation structure-activité des analogues du composé UM171. En couleur orange se trouvent les isostères possibles pour ne pas influencer l'activité, en rouge la zone n'offrant aucune liberté de modification et en vert les sites permissifs pour la fonctionnalisation.

Connaissant les possibilités qu'offre la molécule UM171 pour la fonctionnalisation, diverses idées furent mises de l'avant pour la conception de sondes.

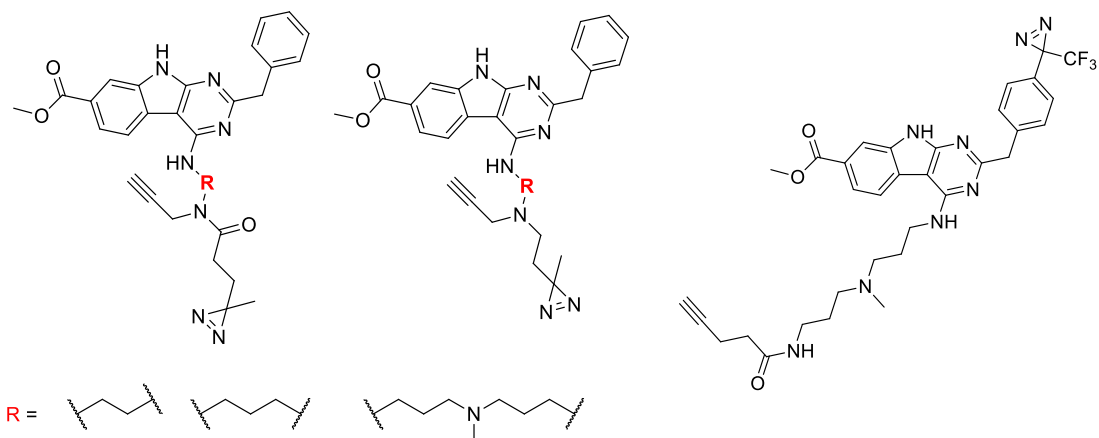


Figure 58 : Approche conceptuelle des sondes photoactivables de UM171 selon les observations décrites par le RSA. La diazirine sera mise soit sur une chaîne aliphatique (structures du centre et de gauche) ou soit sur le groupement benzyle (structure de droite)

Tel que la figure 58 l'illustre, il existe deux approches possibles pour ajouter des fonctions sans trop influencer l'affinité. Il est possible de substituer soit le benzyle ou soit la chaîne aliphatique aminée. L'approche conceptuelle fut initialement de changer le méthyl tétrazole par le méthyl ester pour simplifier la synthèse, d'autant plus que cette modification influence très peu l'affinité du composé. Par la suite, le cyclohexyle sera substitué par une chaîne ouverte aliphatique, notamment pour travailler en milieu achiral tout en modulant plus facilement l'emplacement spatial des fonctions alcyne et diazirine sur les sondes. Dans un cas, l'approche sera de mettre la diazirine sur le benzyl en position para. Cette position permet de limiter la rotation qui pourrait se produire en meta et de maximiser par le fait même les chances que la diazirine soit orientée vers des acides aminés et non dans le solvant. Pour ces sondes, il y aura ajout d'une longue chaîne aminée possédant une amine tertiaire et se terminant avec un alcyne. Dans le cas de figure d'une substitution sur la chaîne aliphatique aminée, il sera question de mettre à la fois la diazirine et l'alcyne sur cette chaîne. Étant donné que cet emplacement est très permissif, il est adéquat d'y ajouter les deux fonctions contiguës l'une de l'autre. Afin de maximiser les chances de réactions de la diazirine pour la protéine d'intérêt, diverses longueurs de chaînes amines seront synthétisées, allant d'une chaîne à 3 carbones à une chaîne à 6 atomes de carbones. Les fonctions diazirine et alcyne seront ajoutées sur l'amine terminale via une liaison amine ou amide.

3.1.2 Synthèse des sondes diazirines substituées sur la chaîne aliphatique

La synthèse de ces sondes a été possible grâce au travail remarquable d'un stagiaire d'été, Moïse Perreault, qui a permis de donner un gros coup de pouce dans leur réalisation. La construction de ces sondes a été séparée en trois étapes distinctes : la synthèse de la diazirine, la synthèse des chaînes aliphatiques et la synthèse de l'hétérocycle 9H-pyrimido[4,5-b] indole.

Synthèse des diazirines

L'une des approches pour ajouter la diazirine à la chaîne aminée consiste à effectuer une réaction de substitution nucléophile entre l'azote et un tosyle. Ainsi, la synthèse de l'intermédiaire de synthèse a été réalisée en suivant la méthodologie développée par le groupe de Radford (Figure 59).⁷⁸ L'idée de cette synthèse de diazirine consiste à former initialement l'imine à partir de la cétone, 4-hydroxy-2-butanone **32**, en utilisant l'ammoniac liquide comme solvant. Puisque l'ammoniac se trouve en excès, l'équilibre cétone/imine est déplacé vers l'imine. Bien que la

plupart des procédures de la littérature utilisent une solution de 7 N d'ammoniac dans le méthanol, la réaction dans l'ammoniac liquide offre une efficacité nettement supérieure compte tenu du grand nombre d'équivalents présent dans la réaction.⁷⁵ Cette réaction nécessite 5 heures d'agitation dans un bain de glace sèche avec acétone pour préserver l'ammoniac sous forme liquide, qui possède un point d'ébullition de 24.7 °C. Une fois cette réaction terminée, il y a eu ajout d'une solution d'acide hydroxylamine-O-sulfonique dans le méthanol. Ce réactif permet de former l'hétérocycle diaziridine, via l'élimination d'acide sulfonique. Ce dernier formera un sel avec l'ammoniac et précipitera du milieu réactionnel favorisant alors la réaction. Il est primordial de solubiliser l'acide hydroxylamine-O-sulfonique dans le méthanol avant de l'ajouter à la réaction, car la réaction est très exothermique et dégage beaucoup de chaleur. Dans l'optimisation de la méthode, lorsque l'acide hydroxylamine-O-sulfonique était ajouté directement sous forme solide à la réaction, une légère explosion était observée, résultant d'une conversion en produit très faible lors de l'analyse du brut réactionnel par RMN. Ainsi, l'acide a été solubilisé dans le méthanol, la réaction a été agitée durant la nuit et l'analyse du brut réactionnel a permis d'observer l'apparition de la diaziridine grâce au déplacement chimique du groupement méthyle. La dernière étape de cette réaction multi-étape était l'oxydation de la diaziridine en diazirine **33** via l'utilisation de l'iode élémentaire et de la triéthylamine. L'apparition d'une couleur orangée persistante indiquait la complétion de la réaction. Il s'agit d'un titrage, l'iode élémentaire est orangé, mais lorsqu'il réagit avec la diaziridine, il y a formation du sel incolore de triéthylammonium et d'iodure. Par conséquent, une persistance de la couleur indique que tous les substrats ont été oxydés. Cette réaction a donné une huile avec un rendement de 51 %, ce qui suit les rendements de la littérature.⁷⁹

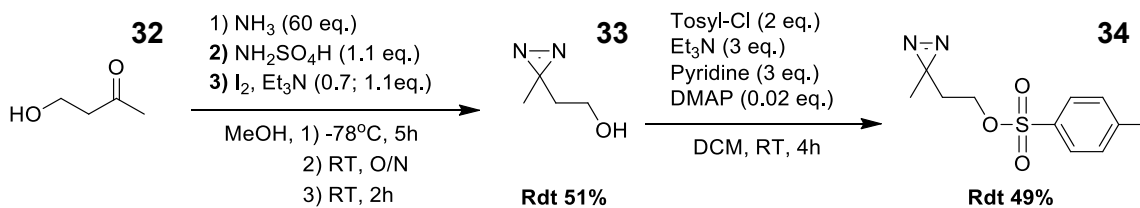


Figure 59 : Séquence expérimentale pour l'obtention de la diazirine **34** fonctionnalisée par un alcool tosylé

La prochaine réaction de cette séquence était la tosylation de l'alcool **33** par le chlorure de tosyloxy. La procédure suivie est celle du groupe du Pr Chen, qui utilise une réaction d'addition-élimination de l'alcool sur le chlorure de tosyloxy.⁸⁰ Il a été observé que la présence d'un solvant chloré tel que le dichlorométhane ou le chloroforme permet de réduire la réaction parasite de la pyridine sur l'ester d'acide sulfonique **35** qui forme le sel de pyridinium **36** (Figure 60).^{81, 82} Le DMAP s'est retrouvé à être un excellent catalyseur pour la tosylation des alcools, réduisant considérablement le temps de réaction en activant l'espèce acylante avec un meilleur groupe partant que le chlore. La triéthylamine permet d'accélérer la réaction en permettant de déprotoner l'alcool ayant réagi sur l'ester sulfonique.

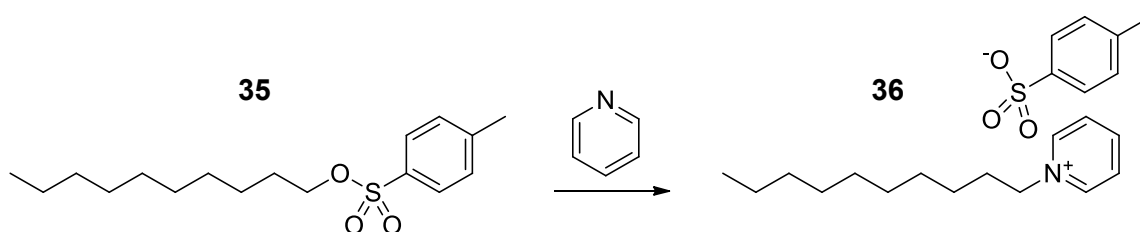


Figure 60 : Formation du sel de pyridium en présence d'un tosyl d'alcool rapportée par Sekera

Cette réaction a donné un rendement de 49 %, ce qui est faible étant donné que cette réaction devrait être quantitative, mais il est probable que la purification sur silice ait décomposé une partie du composé qui est très sensible aux nucléophiles. Le rendement mentionné dans la référence est de 71 %, supérieur à ce qui a été isolé.⁸⁰ Cette diazirine **34** fin prête sera utilisée comme agent alkylant pour fonctionnaliser une amine sur l'échafaudage de UM171.

La seconde diazirine qui a été synthétisée comme intermédiaire de synthèse est un acide carboxylique pour permettre une fonctionnalisation via un couplage amide.

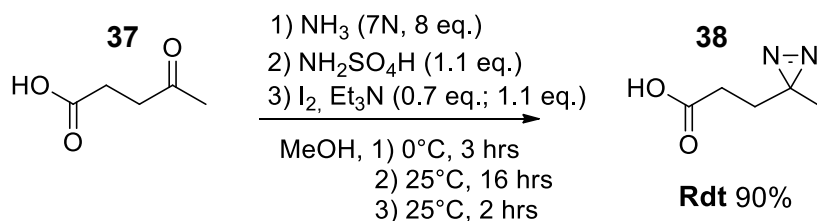


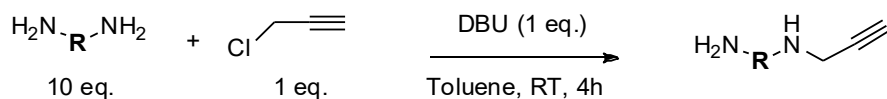
Figure 61 : Séquence expérimentale pour l'obtention de la diazirine fonctionnalisée par un acide carboxylique

Pour synthétiser ce dernier, la méthode du groupe de Pr Jewett a été appliquée.⁸³ Cette méthode utilise l'acide lévulinique **37** et cette fois-ci l'imine a été formée via l'utilisation d'une solution de 7 N d'ammoniac dans le MeOH. Le reste de la méthodologie est la même que celle discutée plus haut, passant par les étapes de la diaziridine via l'acide hydroxylamine-O-sulfonique suivie de l'oxydation par l'iode afin d'obtenir la diazirine **38**. Le rendement obtenu fut de 90% ce qui suit la littérature. Ceci conclut l'obtention des intermédiaires contenant les diazirines.

Synthèse des chaînes aliphatiques

Les seconds blocs de construction qu'il fut nécessaire de synthétiser sont les amines propargylées qui seront ajoutées au corps des sondes de UM171 via une substitution nucléophile aromatique. Le défi de ces synthèses a été de fonctionnaliser asymétriquement une amine sur des substrats diamines. La synthèse sélectionnée s'inspire de la méthodologie du groupe du Pr Nicolas Winssinger, qui utilise un très grand excès de diamines pour une faible quantité de chlorure de propargyle.⁸⁴ De ce fait, le réactif alkylant sera consommé rapidement et les probabilités d'avoir une double alkylation sur la même molécule sont extrêmement faibles. Puisqu'il s'agit d'une question statistique, l'ordre d'ajout des réactifs est primordial dans ce type de réaction. Il faut que l'alkylant soit ajouté dans le réactif en excès et non l'inverse. De plus, l'ajout doit se faire lentement pour avoir une répartition uniforme et limiter les probabilités d'une double addition. Les diamines choisies possédaient diverses longueurs permettant d'offrir un plus grand éventail dans l'espace pour venir lier la cible inconnue. Le tableau 11 renseigne sur les rendements obtenus pour les diverses diamines sélectionnées (**39,40,41**).

Tableau 11 : Diamines propargylés asymétriques synthétisés et leur rendement correspondant



Diamine	Produit	Identification	Rendement
		39	31 %
		40	16 %
		41	40 %

Il est très difficile de suivre cette réaction, puisque l'un des substrats est en large excès et que le second est aliphatique et de petite masse. Pour maximiser les chances de complétion de la réaction, un temps de 4 heures a été suivi aveuglément comme effectué dans la publication.⁸⁴ Pour la purification, il a été nécessaire d'utiliser une distillation de type Kugelrohr pour séparer le toluène, le réactif diamine initial et le produit. Les faibles rendements obtenus sont dus à cette distillation qui n'était pas simple pour la séparation des différents paliers. Des paliers impurs furent obtenus, occasionnant ainsi des pertes de produits. Néanmoins, cette méthode a permis d'obtenir les différentes amines monosubstituées en une seule étape.

Synthèse des sondes aliphatiques

Une fois les différents morceaux obtenus, il était possible d'entamer la synthèse des diverses sondes aliphatiques. La voie de synthèse qui a été suivie est représentée à la figure 62.

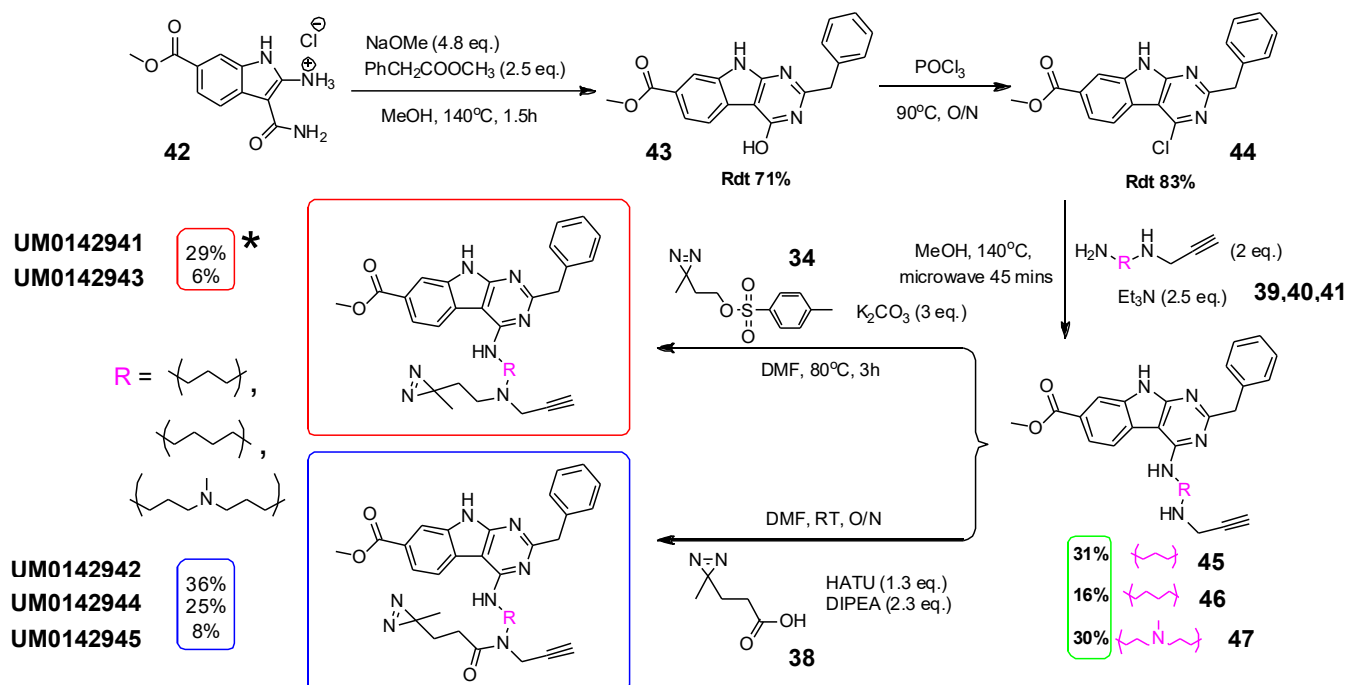


Figure 62 : Séquence synthétique pour la formation des sondes diazirines aliphatiques analogues au composé UM171. (*) Les astérisques signifient que les structures présentées dans le carré rouge sont celles désirées et non les structures réellement obtenues comme il sera expliqué plus loin.

La voie de synthèse de l'échafaudage de UM171 provient du brevet publié par l'équipe des chimistes qui ont développé la chimie pour la synthèse des analogues de UM171.⁸⁵ La séquence est illustrée à la figure 62. La voie qui a été développée débute par la formation de l'hétérocycle pyrimidine qui est fusionné avec l'indole **42** via une condensation entre un méthyl ester et un système amide/amine. Cette réaction se fait dans le micro-onde avec l'ajout du méthyl 2-phénylacetate et le 2-amino-1H-indole-3-carboxamide **42** en présence de méthanoate de sodium. Il était très important d'avoir un milieu anhydre, car l'eau a un impact limitant sur la réaction en déplaçant l'équilibre vers les réactifs. Une fois le corps de la molécule formée **43**, il y a eu chloration de l'alcool de la composante hydroxypyrimidine via le trichlorure de phosphoryle agissant comme solvant et réactif.^{85, 86} Cette réaction se trouve à être très lente à cause de l'acidité du proton présent sur l'alcool nucléophile qui est faible, aspect important pour qu'un anion chlorure puisse effectuer la substitution. La présence du chlore en position adjacente d'un azote permet de faire une réaction clé en synthèse hétérocyclique, soit une substitution nucléophile

aromatique avec les amines synthétisées préalablement. Pour effectuer la réaction de S_NAr , l'amine en question (**39,40** ou **41**) et l'hétérocycle **44** ont été mélangés dans le MeOH et placés dans le micro-onde à 140 °C. L'encombrement stérique va guider la sélectivité de la réaction favorisant l'amine primaire plutôt que l'amine secondaire. Suite à l'obtention de ces produits (**45, 46** ou **47**), la diazine a été ajoutée. Pour ce faire, deux approches ont été utilisées, soit via un acide carboxylique **38** ou soit via un tosylate d'alcool **34**. Dans le premier cas, un couplage avec l'agent activant HATU en présence d'une amine tertiaire, DIPEA, a permis d'obtenir les diverses sondes souhaitées ('**942**, '**944** et '**945**).⁸⁷ Pour l'approche utilisant le tosylate, une réaction de substitution nucléophile S_N2 a été privilégiée dans le DMF avec l'ajout d'une base de carbonate pour effectuer la réaction.⁸⁸ Lors des analyses par LCMS et masses exactes des produits obtenus de cette réaction ('**941** et '**943**), tout semblait illustrer l'obtention des sondes désirées. Cependant, lors de l'analyse RMN une surprise indésirée a été observée : la substitution ne s'est pas produite au bon endroit. La figure 63 illustre une comparaison entre la sonde amide '**942** et la sonde '**941** provenant de la réaction de S_N2 avec le tosylate. Le premier constat fut une absence du proton du NH aromatique lors de l'utilisation de la réaction de substitution nucléophile. Le second constat est que les protons du cycle composé de l'ester méthylique ont subi un déplacement chimique relatif à une sonde où un coupage amide a été effectuée, qui par conséquent possède un N-H pyrrolique non fonctionnalisé. Ces deux indices indiquent une seule conclusion : la diazine a été ajoutée sur l'azote pyrrolique et non sur l'azote aliphatique initialement prévu. Afin de comprendre l'explication de cette réactivité, il faut se pencher sur les pka des diverses espèces présentes dans la réaction, lesquels sont illustrés à la figure 64.

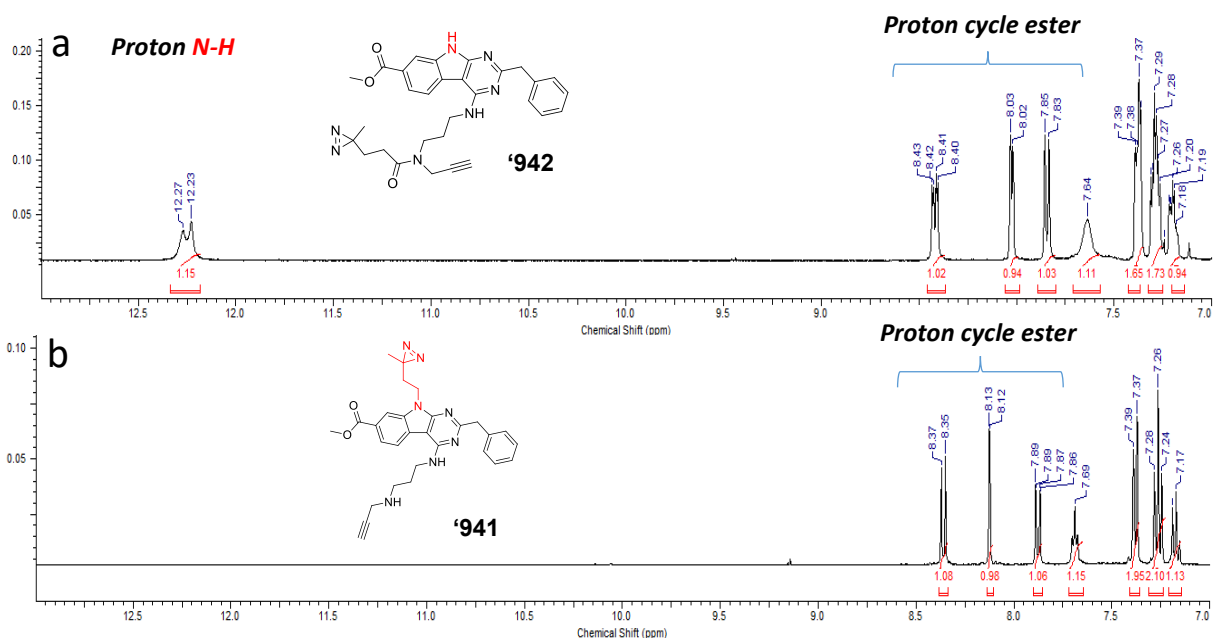
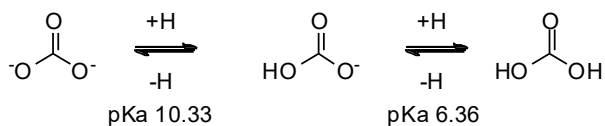
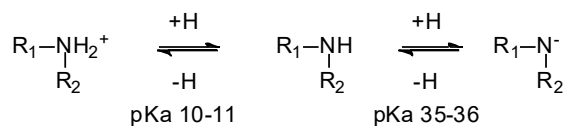


Figure 63 : Spectre RMN focalisé sur la région aromatique afin d'illustrer la perte de l'hydrogène N-H aromatique lors de la réaction avec le tosylo d'alcool. A) Spectre RMN d'une sonde diazirine liée selon un lien amide qui comporte le proton sur l'azote pyrrolique. B) Spectre RMN d'une sonde diazirine provenant de la réaction avec le tosylo d'alcool où il est possible de voir la disparition du proton pyrrolique.

Carbonate



Amine secondaire



Azote aromatique

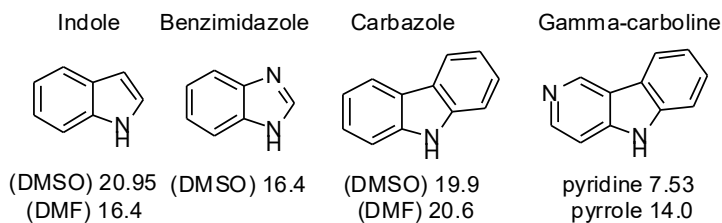


Figure 64 : Valeur de pKa des espèces impliquées dans la réaction de SN₂ activée par la base carbonate de potassium. Le carbonate, l'amine secondaire de la chaîne aliphatique et des analogues hétérocycliques de l'azote pyrrolique se trouvant dans le composé UM171.

L'objectif de l'utilisation de la base carbonate dans la réaction est de venir activer l'azote aliphatique et de s'assurer que l'espèce en présence est l'amine secondaire neutre. L'amine protonée n'est pas réactive pour une réaction de S_N2 , contrairement à l'espèce secondaire neutre. Le carbonate possède un pKa de 10.33 selon la table de pKa⁶² et l'amine secondaire protonée possède un pKa d'environ 10-11 variant en fonction des substituants. Ainsi, la base carbonate permet de s'assurer que le bon état de charge de l'azote pour la réaction soit obtenu. Cependant, si on regarde maintenant le pKa de l'azote aromatique, c'est plus complexe. Il faut faire des analogies avec la littérature pour déduire la valeur du pKa. Si l'on regarde le pKa de l'indole, ce dernier est de 21 dans le DMSO et le fait de rendre le cycle plus pauvre en électrons avec le benzimidazole le descend à 16.4.⁸⁹ L'évolution des pKa dans un système à trois cycles fusionnés va comme suit : dans le cas du carbazole le pKa est légèrement plus acide que celui de l'indole à cause de l'augmentation du degré d'aromaticité. La différence de pKa est encore plus marquée avec l'ajout d'un azote dans un cycle adjacent qui appauvrit la densité électronique et diminue le pKa à une valeur de 14.0 pour l'azote pyrrolique.⁶² Dans un système aromatique, ce qui régit le pKa est la capacité du système à neutraliser la charge négative qu'occasionnera la déprotonation. On peut penser à deux paramètres qui peuvent mieux stabiliser la charge : tout d'abord, l'ajout de fonctions qui peuvent augmenter la résonance, par exemple un ester, un cycle aromatique supplémentaire. Le second aspect est l'ajout de groupements qui retirent de la densité électronique au système, donc qui possède une grande électronégativité, tels que des groupements nitro et trifluorométhyle. Lorsque l'on regarde la structure de base des substrats de UM171, on constate que l'azote pyrrolique se trouve dans un système à trois cycles, avec deux azotes dans le cycle adjacent et possède un ester méthylique dans l'autre cycle adjacent. Bien que le benzyle et l'azote sont des groupements donneurs de densité électronique, leur effet est beaucoup moins important que les autres fonctions qui retirent beaucoup de densité. De ce fait, le pKa de l'azote pyrrolique se trouve à être inférieur à 14.0, comme il a été observé dans le γ -carboline, une espèce possédant une structure analogue au corps de UM171. Ainsi, le carbonate a déprotoné l'azote aromatique : ce dernier, devenu plus nucléophile que l'azote aliphatique, a réagi sur le tosylate. Ces produits ne furent pas désirés à premier abord, mais restent tout de même intéressants puisqu'ils offrent un nouvel emplacement dans l'espace pour la position de la diazirine.

Voici les structures corrigées des composés obtenus :

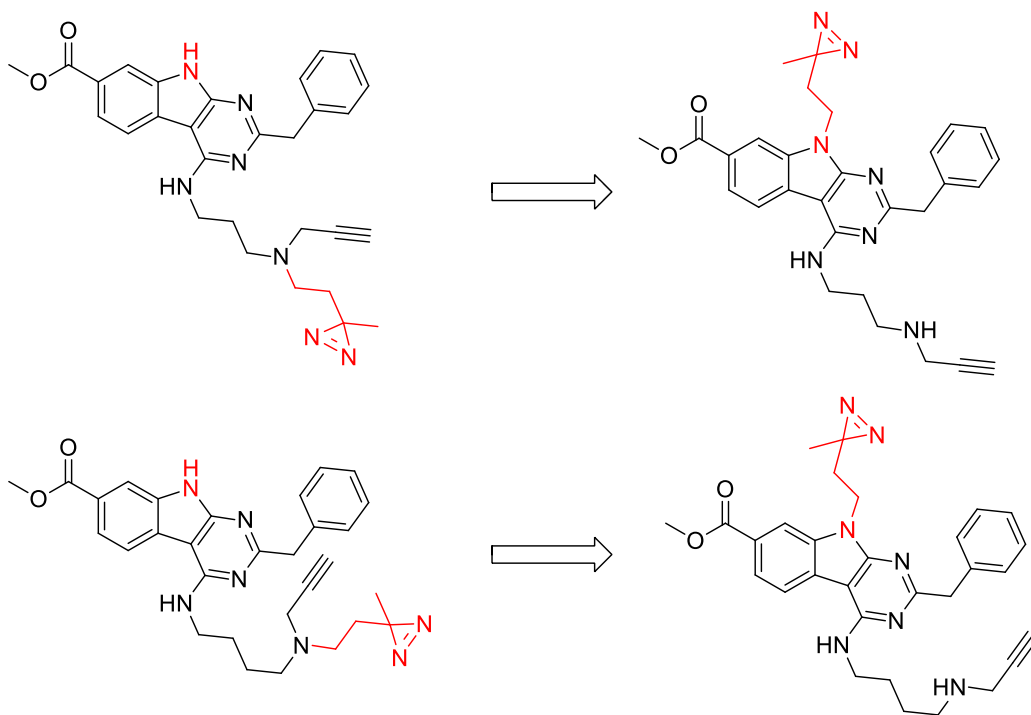


Figure 65 : Structures corrigées pour les sondes aliphatiques qui devaient être substituées sur l'amine aliphatique par la diazirine mais qui en réalité ont été fonctionnalisées sur le N-H aromatique

3.1.3 Synthèse de la sonde diazirine substituée sur un aromatique

La synthèse de cette diazirine a été décomposée en deux étapes. Dans un premier cas, il y a eu pré-fonctionnalisation des espèces formant le corps de UM171 avec un précurseur de diazirine suivie de la condensation formant le cycle. La seconde étape a été de déprotéger la pré-fonctionnalisation et de venir former la diazirine CF_3 . Afin de venir ajouter le groupement trifluorométhyle, il est nécessaire de passer par un carbanion très réactif pouvant réagir avec une multitude d'électrophiles. Par conséquent, il faut ajouter ce groupement en début de synthèse, où les substrats sont simples et non fonctionnalisés. Ensuite, les groupements protecteurs Benzyl et Boc seront nécessaires pour fonctionnaliser sélectivement la molécule.

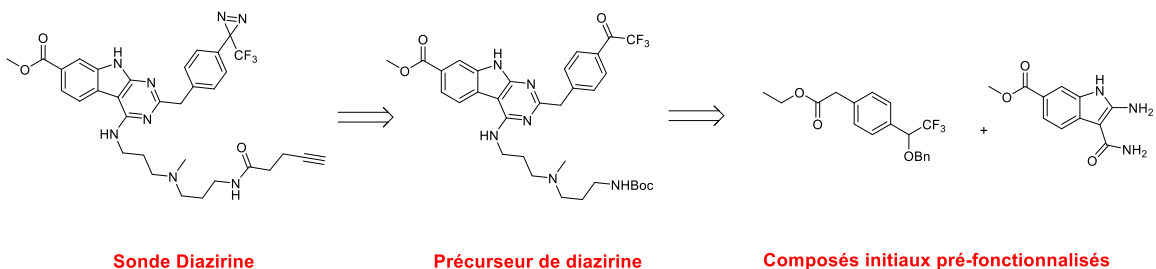


Figure 66 : Rétro synthèse des grandes étapes pour la formation de la sonde diazirine aromatique débutant par la construction de deux morceaux pré-fonctionnalisés suivis du précurseur de diazirine avec le corps du composé formé suivi de la formation de la diazirine

Synthèse des précurseurs du core de UM171

La synthèse du premier précurseur s'est inspirée de la procédure du brevet de UM171 publiée par le groupe de chimie médicinale de l'IRIC.⁸⁵ La synthèse débute par l'acide para-tolylacétique **48** qui a été estérifié **49** selon les conditions de Fischer, soit dans l'éthanol par catalyse avec l'acide sulfurique.⁹⁰ Le rendement fut quantitatif pour cette réaction. La seconde réaction est une bromination radicalaire nécessitant l'utilisation de l'initiateur radicalaire azobisisobutyronitrile.^{62, 91} Le gros problème de cette réaction est le contrôle de la sélectivité du radical brome pouvant réagir sur les deux sites benzyliques. En substituant le tétrachlorométhane par le chloroforme, il a été possible d'augmenter la régiosélectivité et d'obtenir un rendement de 62 % pour la réaction de bromation **50**. La réaction subséquente a été une oxydation de Ganem, qui est en fait la variante de la réaction de Kornblum via l'utilisation d'un N-oxide d'amine tertiaire.^{92, 93} Cette réaction se fait dans un solvant apolaire puisque l'état de transition élimine la charge pour former l'aldéhyde **51**. De ce fait, un solvant apolaire accélère la réaction. Un rendement de 61 % a été obtenu qui peut être expliqué par la présence d'eau dans le dioxane pouvant déprotoner l'amine tertiaire responsable d'agir comme base pour le proton benzylique nécessaire à l'oxydation.⁹⁴ Une fois l'aldéhyde **51** obtenu il y a eu ajout du groupement trifluorométhyle via l'addition du carbanion CF_3^- . Le réactif de Prakash permet la formation du carbanion en présence d'un anion permettant de former un intermédiaire pentavalent sur le silicium.⁹⁵ Il a été démontré par le groupe de Wu que l'utilisation du carbonate de césium permettait d'activer le silicium pour que la réaction sur des aldéhydes et de cétones aromatiques soit possible.⁹⁶ Une fois le carbanion ayant réagi sur l'aldéhyde, l'alcool formé sera *in situ* protégé par le groupement triméthylsilyle. Comme cette synthèse nécessite un groupement protecteur différent du TMS pour la

compatibilité dans les réactions subséquentes, il a été nécessaire d'utiliser une solution de fluorure de TBAF pour déprotéger l'alcool pour introduire un groupe protecteur différent.⁹⁷ Il est important d'effectuer la réaction du carbanion dans un milieu le plus anhydre possible étant donné la grande réactivité de ce dernier vis-à-vis les hydrogènes labiles. De ce fait une présence d'eau neutralise le réactif et un rendement faible sera occasionné. La séquence de l'intermédiaire de synthèse se termine par la protection de l'alcool benzylique **52** nouvellement formé par un groupement benzyle via une réaction SN₂ activée par une base.⁹⁷ La séquence possède un rendement global de 24 % sur 5 étapes avec des pertes plus importantes pour la bromation et l'oxydation de Ganem.

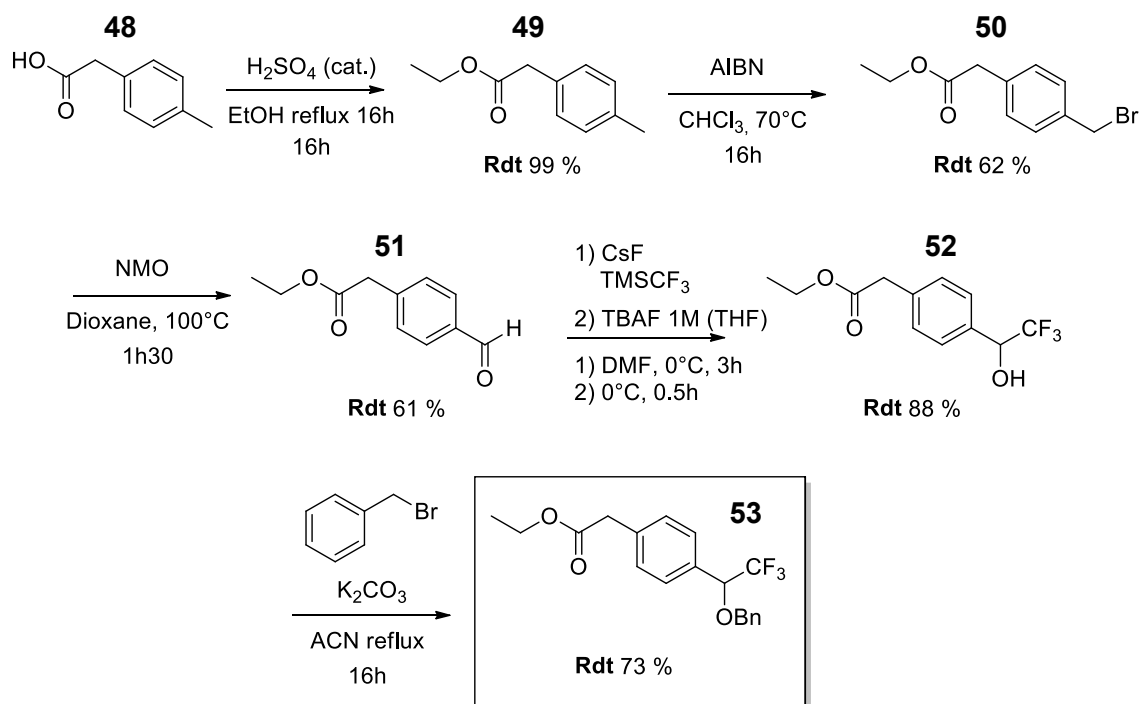


Figure 67 : Synthèse pour la formation de l'ester benzylique **53** nécessaire pour la formation du tricycle de la sonde

Le second précurseur a été synthétisé par un collègue de la plateforme de chimie. Afin d'expliquer la logique de la synthèse, la séquence pour l'obtention de ce dernier est illustrée à la figure 68. Le mérite de cette synthèse appartient à Stéphane Gingras. Cette synthèse est décrite dans le brevet soumis par les chimistes en 2013 pour la synthèse des analogues de UM171.⁸⁵

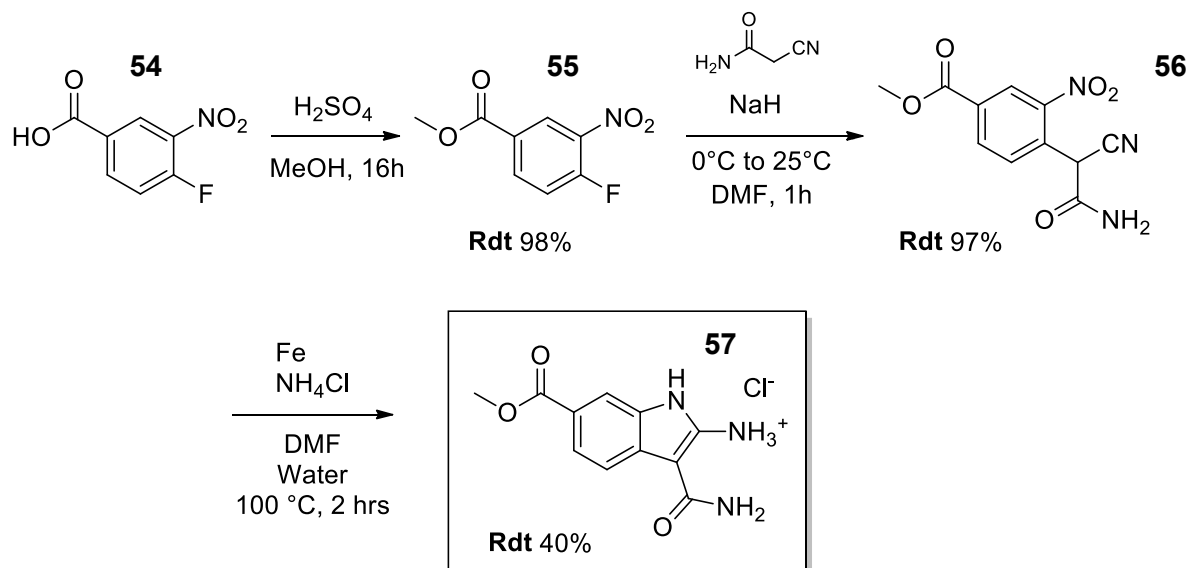


Figure 68 : Synthèse pour la formation du 2-amino-3-amide indole **57** nécessaire pour la formation du tricycle de la sonde

Cette synthèse débute par l'utilisation de l'acide 4-fluoro-3-nitrobenzoïque **54** qui est estérifié par le méthanol en catalyse acide.⁹⁰ L'une des façons traditionnelles de former les 2-amino-indoles consiste à effectuer une réduction cyclique d'un nitro sur un nitrile.⁹⁸ L'intermédiaire 2-amino-3-amide indole est un excellent précurseur pour venir former une pyrimidine qui est nécessaire pour la formation du tricycle que compose les analogues de UM171. Une fois l'ester **55** formé par la réaction de Fisher⁹⁰, il y a eu substitution nucléophile aromatique par l'énolate du 2-cyanoacétamide préalablement formé avec l'hydrure de sodium.^{85, 99} Finalement, une réduction sélective du nitro **56** par la réduction de Bachmann qui utilise le fer et un acide a été réalisée.¹⁰⁰ Ceci a permis la condensation de l'aniline sur le nitrile, permettant la formation du 2-amino-3-amide indole **57**. Cette séquence possède un rendement global de 39 % sur 3 étapes avec une perte importante sur la formation de l'indole.

Synthèse du précurseur de diazine

Une fois les précurseurs **53** et **57** en main, la séquence pour la synthèse est la même que celle décrite dans la section des sondes aliphatiques et dans le brevet de UM171.⁸⁵ Tout d'abord, il y a eu formation de l'hétérocycle par condensation de l'amine du système 2-amino-3-amide indole **57** sur l'ester **53** et l'amide a permis la fermeture du cycle grâce au méthanoate de sodium.

Finalement, il y a eu chloration de l'alcool **58** et substitution nucléophile aromatique sur ce chlorure **59** nouvellement formé par l'amine N,N-Bis(3-aminopropyl)méthylamine.^{85, 86} L'étape suivante de la synthèse a été la protection de l'amine aliphatique **60** par le groupement Boc. Cette réaction se fait via l'utilisation du dicarbonate de di-tert-butyle en présence d'une base, qui fut ici la triéthylamine. La prochaine réaction a été la déprotection du groupement benzyle **61** par une hydrogénolyse par catalyse hétérogène sur palladium.⁹⁷ La dernière réaction pour cette synthèse

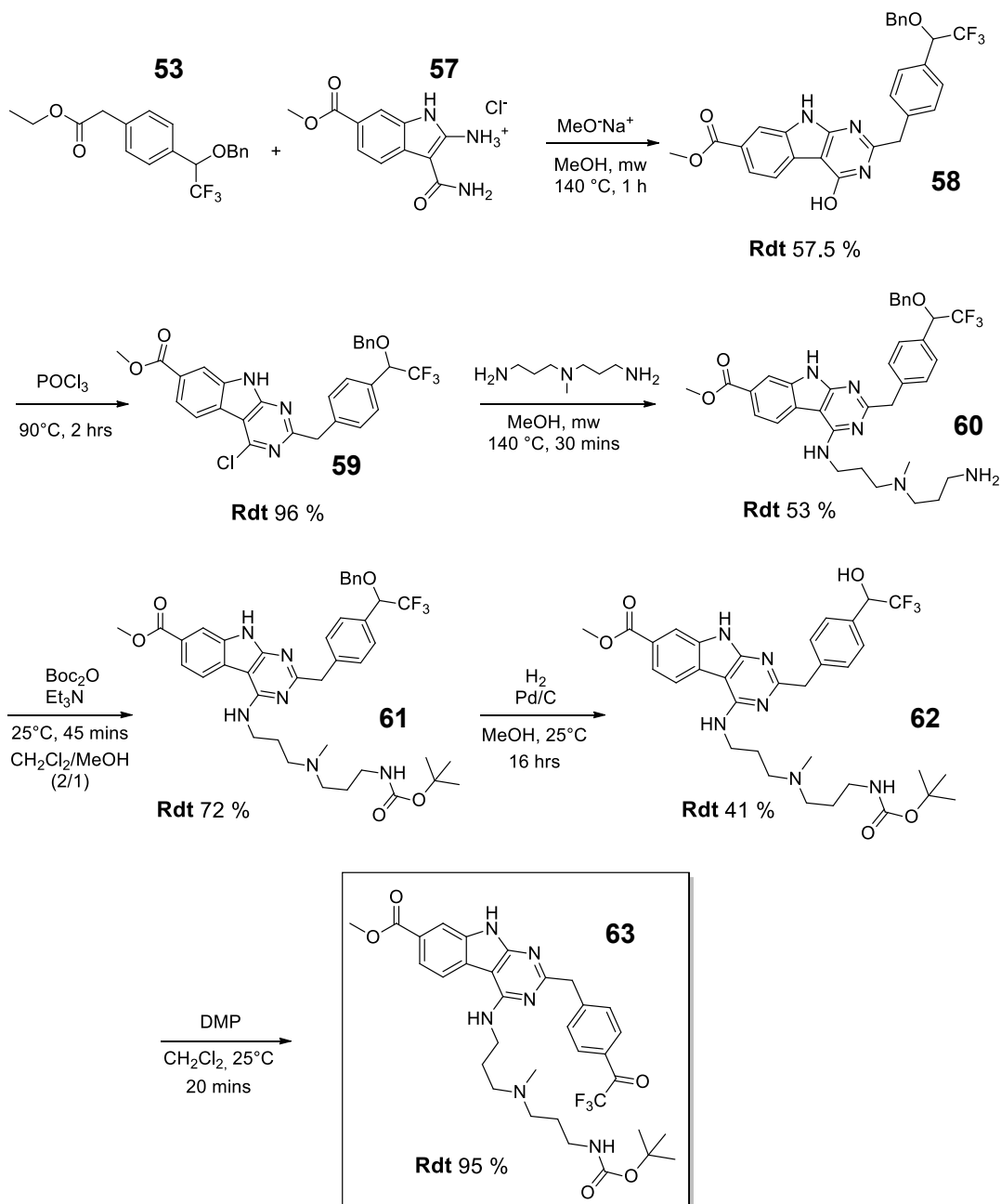


Figure 69 : Étape synthétique pour la construction de la cétone de trifluorométhyle **63**, précurseur de la diazirine aromatique

du précurseur de la diazirine fut d'obtenir la cétone de trifluorométhyle **63**. Pour ce faire, l'oxydation de l'alcool benzylique **62** avec le réactif de Dess-Martin a été effectuée.^{101, 102}

Synthèse de la diazirine aromatique

La dernière section de la synthèse consiste à former la diazirine à partir de la cétone trifluorométhyle **63**. De nombreuses séquences ont été développées pour la synthèse des diazirines aromatiques^{75, 103-105}, cependant la stratégie choisie est celle décrite par le groupe du professeur Zezschwitz.¹⁰⁶ Tout d'abord, la cétone **63** a été transformée en oxime **64** selon une réaction d'addition-élimination sur le carbonyle. Pour catalyser la réaction, la pyridine a été utilisée. Cette dernière est peu nucléophile et va déprotonner la charge de l'hydroxylamine. Lorsque l'oxime **64** est formée, sa fonction hydroxyle sera tosylée par le chlorure de tosyle, transformant ainsi la fonction en excellent groupe partant. Comme cette fonction est sensible, uniquement une extraction liquide-liquide a été réalisée et le brut de réaction a été utilisé pour la formation de la diaziridine. Pour ce faire, il y a eu réaction de l'oxime tosylée **65** avec de l'ammoniac solubilisé dans le dichlorométhane à -78°C. Par une réaction d'addition sur le carbonyle et de substitution nucléophile sur l'azote, il y a eu formation de la diaziridine **66**. Une oxydation de cette dernière par de l'iode et de la triéthylamine a permis la formation de la diazirine aromatique **67** avec un rendement de 75 %.¹⁰⁵ La dernière étape de la synthèse a été la déprotection du groupement Boc par l'acide TFA suivi d'un couplage amine avec l'acide 4-pentynoïque **69**.

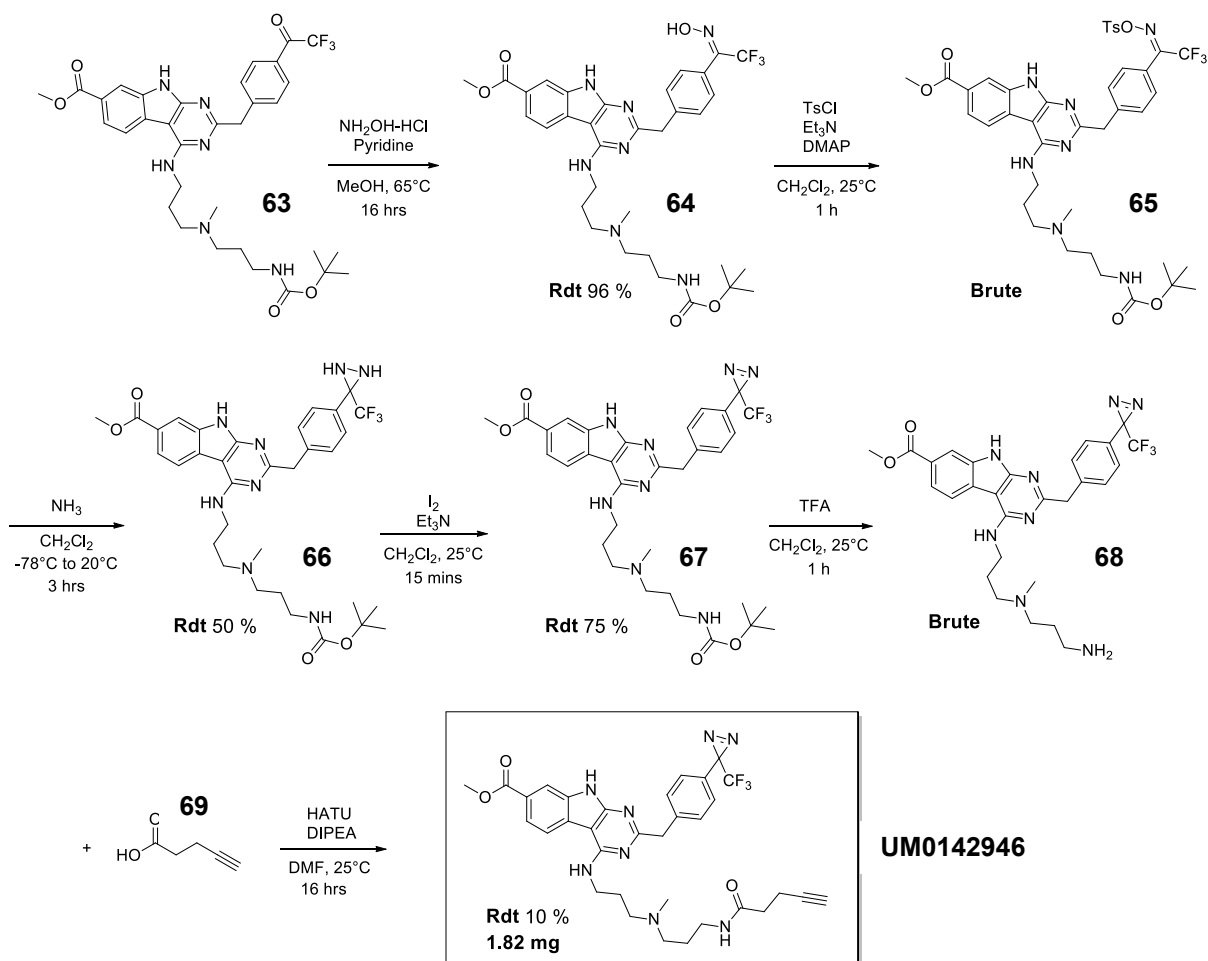
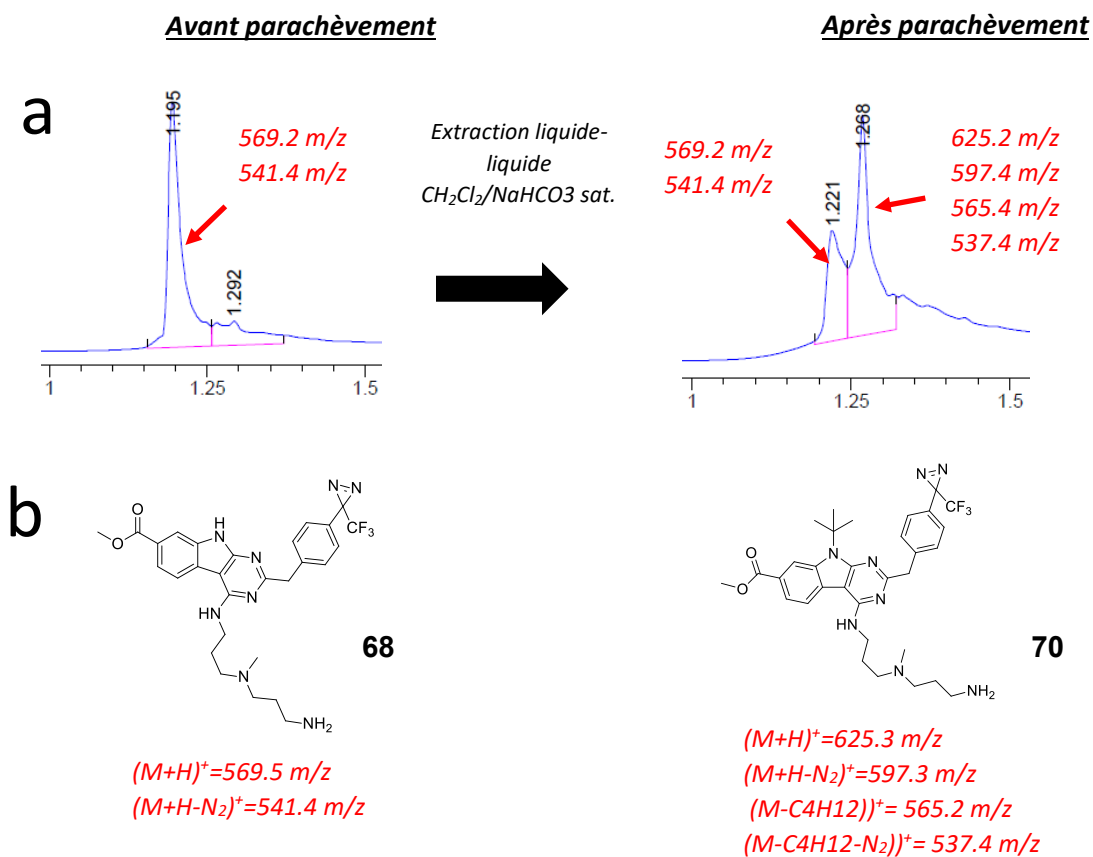


Figure 70 : Dernières étapes pour la formation de la sonde diazirine aromatique de UM171

Si on exclut la synthèse des intermédiaires le rendement global de cette synthèse est de 0.3 % sur 12 étapes. Les étapes qui ont été coûteuses furent la condensation formant l'hétérocycle pyrimidine, la réaction SnAr, l'hydrogénolyse et le couplage de l'amine. Les deux premières réactions sont connues pour l'obtention de bas rendement compte tenu de la difficulté des réactions et du trouble occasionné par l'eau. L'hydrogénolyse est une réaction qui est normalement quantitative, mais la purification sur silice a occasionné une très grande perte. Il s'avère que le composé s'adsorbait sur cette dernière et il a été très difficile de l'extraire malgré des solvants très polaires. Finalement, on pourrait croire que le couplage de l'amine fut désastreux, mais en réalité il s'agit de l'étape de déprotection du Boc qui fut catastrophique pour la synthèse. Une analyse LCMS a été prise avant d'effectuer le parachèvement pour vérifier la

progression de la réaction et uniquement le produit déprotégé a été observé. Par la suite, le solvant a été évaporé et l'huile résultante a subi une neutralisation basique et une extraction liquide-liquide. Cependant une fois la phase organique évaporée, une analyse LCMS du produit final a été effectuée et par mauvaise surprise un nouveau produit s'est formé (Figure 71A). Normalement, le carbocation tert-butyl finit par s'évaporer en formant du 2-méthylpropène, cependant il s'avère qu'en présence de TFA il y a formation d'un ester activé.¹⁰⁷ (Figure 71C). Il a été rapporté dans la littérature que le tryptophane se fait alkyler par un groupement t-Bu lors d'une déprotection au TFA.¹⁰⁷ Il est fort à parier que lors de l'évaporation du solvant acide, il y a eu alkylation de l'azote aromatique étant donné que le milieu était concentré et très acide. Ceci explique la présence du composé secondaire indésirable **70** de la figure 71B. Il aurait été important d'ajouter une molécule piègeuse de carbocation qui agirait comme nucléophile et consommerait cet électrophile. Il existe diverses possibilités telles que le triisopropylsilane et le phénol qui permettent d'empêcher cette réaction non-désirée.¹⁰⁷. Malheureusement la déprotection du Boc a occasionné une perte d'au moins 60 % du produit final.



C

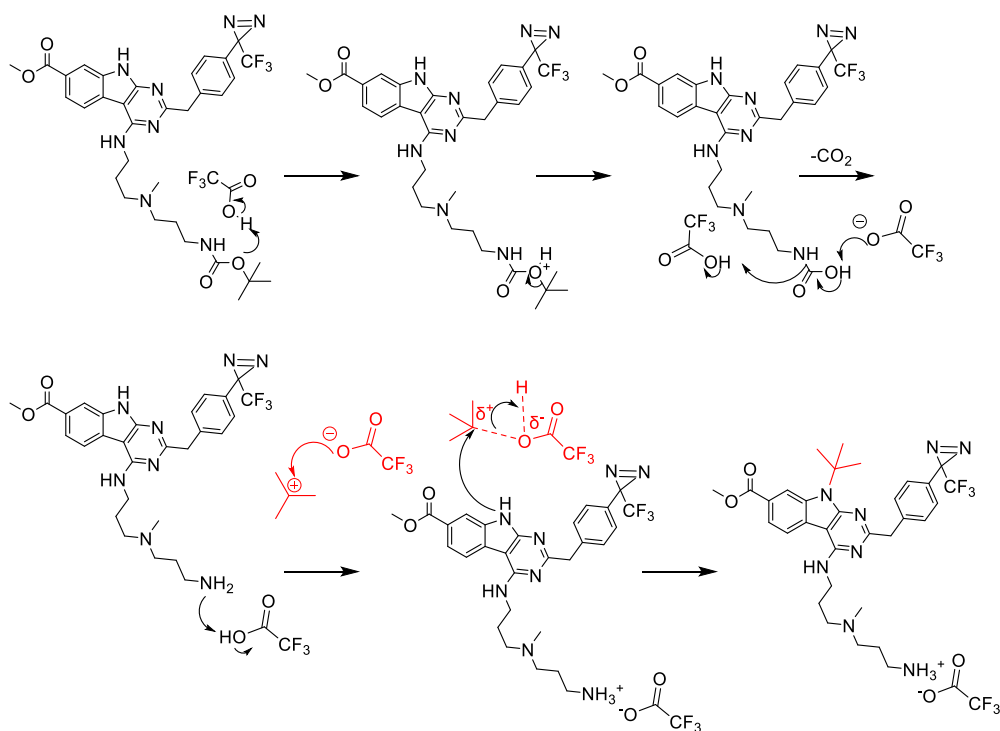
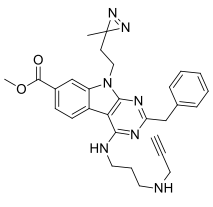
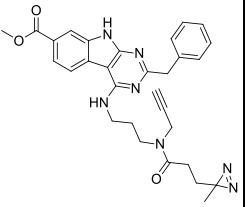
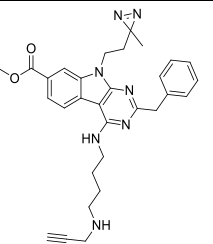
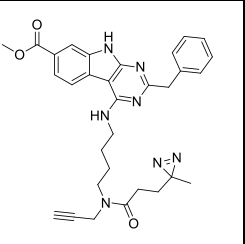
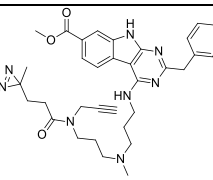
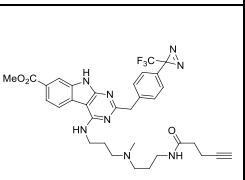
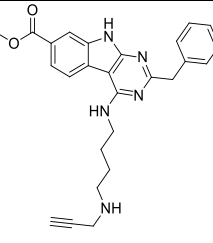
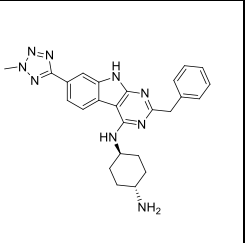


Figure 71 : Formation d'un produit d'alkylation **70** par le cation t-Butyl qui a causé une perte de rendement considérable. a) LCMS du milieu réactionnel avant et après le parachèvement. b) Structure chimique du produit désiré ainsi que de l'impureté. c) Mécanisme réactionnel expliquant la formation du produit secondaire.

3.1.4 Détermination de l'activité des sondes

Lorsque les synthèses ont été complétées, les composés ont été envoyés à Sophie Corneau, une agente de recherche du Pr Sauvageau, qui a déterminé l'activité cellulaire des sondes. Ces activités ont été déterminées par le même essai biologique qui a permis d'effectuer l'étude de relation structure-activité du composé UM171 discutée dans l'introduction, soit via une quantification des protéines CD34⁺ et CD45RA⁻ présent à la surface des cellules souches. Le tableau 12 illustre la structure moléculaire des composés soumis, leur identification et leur activité.

Tableau 12 : Identification des sondes analogues à la structure de UM171 ainsi que leur activité biologique multiplicative sur les cellules souches hématopoïétiques

Composé	Structure chimique	EC _{1,5} (nM) (n=2)	Composé	Structure chimique	EC _{1,5} (nM) (n=2)
UM0142941		4369	UM0142942		280
UM0142943		Inactive	UM0142944		361
UM0142945		36	UM0142946		153
UM0142947		10	UM171		17-19 ³⁸

L'activité qui a été rapportée pour le composé UM171 dans la littérature est de 17-19 nM, connaissant cette valeur et ayant une connaissance du RSA, il est possible de mieux interpréter les valeurs obtenues.³⁸ Le composé **'947** possède une structure similaire au composé UM171 à l'exception de la présence de méthyl ester et d'une chaîne aliphatique aminée distincte. Sans grande surprise ce composé possède une activité du même ordre de grandeur, compte tenu que les changements structuraux conservent les mêmes interactions présentes dans UM171. Maintenant, si l'on regarde les sondes possédant une diazirine liée par un lien amide, soit les composés **'942**, **'944** et **'945**, on peut constater une différence d'affinité entre eux et avec UM171. Pour les sondes **'942** et **'944**, on ne retrouve pas la présence d'une amine basique sur la chaîne aliphatique contrairement au composé **'945**. Cette différence structurale est corrélée par la différence d'affinité entre les sondes **'942/'944** et la sonde **'945** qui est plus affine. Alors la sonde **'945** possède une affinité similaire au composé UM171, puisqu'elle a une amine basique contrairement aux sondes **'942** et **'944**. Cette observation est en accord avec la RSA préalablement effectuée par l'équipe de chimistes de la plateforme de chimie médicinale qui indiquait une augmentation de l'affinité en fonction du pKa de l'amine. On peut faire les hypothèses que cette amine basique permet d'influencer divers aspects tel que la solubilité, les interactions dans le site actif et une localisation favorable pour la molécule et la cible thérapeutique. Si l'on observe les sondes **'941** et **'943**, il est possible de constater que leur activité est très faible. Cette très grande différence est due à l'emplacement de la diazirine sur l'azote central du tricycle de la structure de UM171. Dans la RSA, il avait été démontré qu'une méthylation de cet azote aromatique rendait le composé inactif. On peut conclure que ce dernier permet une liaison hydrogène dans le site actif avec un accepteur. Ce contact est essentiel pour l'activité, il n'est pas surprenant de voir une perte d'affinité lorsque les diazirines sont situées sur ce site. Le composé **'946** possède la diazirine sur le benzyl en position para, selon la RSA, cette position est tolérante à un ajout d'atome tel que le brome et le fluor. Il y a constatation d'une légère perte d'affinité par rapport aux composés UM171 et **'947**. Il est fort à parier que cette perte d'affinité provienne d'encombrement stérique de la fonction diazirine dans le site actif. Ainsi les sondes synthétisées permettent d'offrir trois orientations distinctes pour la diazirine selon leur localisation sur la molécule. Il existe toujours un coût pour l'ajout d'une fonctionnalité sur une molécule comme il a été possible de le constater avec les diverses sondes, cependant l'impact est plus ressenti sur les sondes **'941** et **'943**, contrairement aux autres sondes qui ont perdu au maximum un ordre de grandeur vis-à-vis le composé phare UM171.

Parmi les sondes initialement planifiées, il n'a pas été possible d'obtenir les sondes aliphatiques comportant une diazirine attachée à une amine basique compte tenu d'une réactivité non-planifiée de l'azote indolique lors de la réaction S_N2 sur le tosylate. Une seconde voie synthétique serait possible pour obtenir ce type de substrat. Lors de travaux préliminaires pour la synthèse d'un substrat non-relié à ce projet, une voie de synthèse a été explorée avec Moïse Perreault, qui mettait de l'avant la fonctionnalisation asymétrique d'une diamine avec une diazirine et un alcyne. Débutant avec la diamine **39** décrite dans le mémoire, il est possible de venir protéger sélectivement une amine primaire d'une amine secondaire avec le réactif Boc-ON.¹⁰⁸ Lors du stage de Moïse¹⁰⁹, cette réaction a donné un rendement isolé de 44 % de cette diamine **71**. Comme une amine est nettement plus réactive qu'un amide, l'ajout de la diazirine **34** ayant le groupement partant tosylate a permis d'obtenir sélectivement la réaction S_N2 sur l'amine **71** avec un rendement de 33 %. L'étape suivante est la déprotection du groupement Boc avec de l'acide chlorhydrique dans le dioxane pour se retrouver avec le sel d'amine **73**.⁹⁷ L'hypothèse synthétique serait d'effectuer la substitution nucléophile avec cette amine asymétrique où le seul site réactif sera l'amine primaire. Cette voie de synthèse convergente permet de s'assurer que la diazirine se trouve sur l'amine de la chaîne aliphatique et non l'amine aromatique. La fonctionnalité diazirine est relativement stable à la chaleur, néanmoins l'utilisation du micro-onde pourrait être problématique. Il est possible de faire la réaction S_NAr sur une période de 16h sans utiliser le micro-onde pour obtenir le produit **74**.⁸⁵ De cette façon les sondes auraient pu être obtenues.

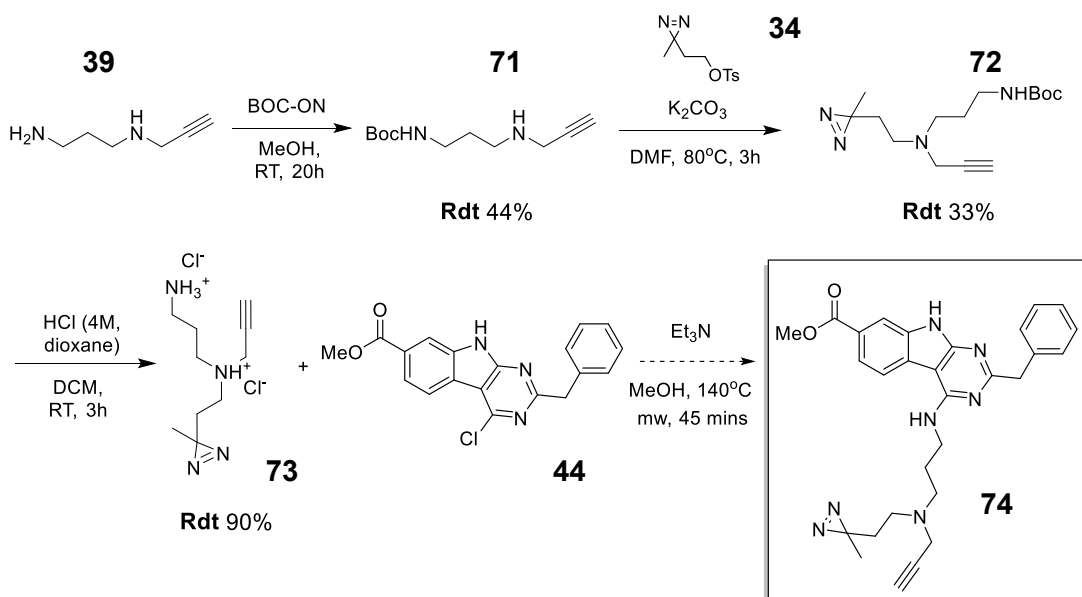


Figure 72 : Suggestion d'une voie alternative convergente pour les sondes diazirines substituants une amine basique sur la chaîne aliphatique

3.2 Optimisation de la méthode de chimioprotéomique

3.2.1 Approche pour l'optimisation de la méthode de chimioprotéomique

Tel que mentionné dans l'introduction l'approche utilisée pour l'identification de la cible de UM171 est la chimioprotéomique via la méthode ABPP. Avant d'effectuer une expérience complète, il est primordial de vérifier et d'optimiser les diverses étapes de la procédure pour en connaître ces limitations et ces étapes critiques. Ainsi cette présente section déclinera les optimisations et vérification de la procédure pour en finir avec une expérience complète avec les sondes synthétisées. La procédure se divise selon les étapes suivantes : il y a incubation de la sonde, suivie d'une irradiation UV pour activer la diazirine qui est ensuite suivie d'une lyse cellulaire pour effectuer une réaction Click venant attacher une biotine aux sondes permettant un enrichissement via des billes de streptavidine. À la suite de lavages, un clivage des billes est finalement effectué avant la digestion trypsique et les peptides sont injectés dans le spectromètre de masse. La figure 73 illustre les diverses étapes schématiquement.

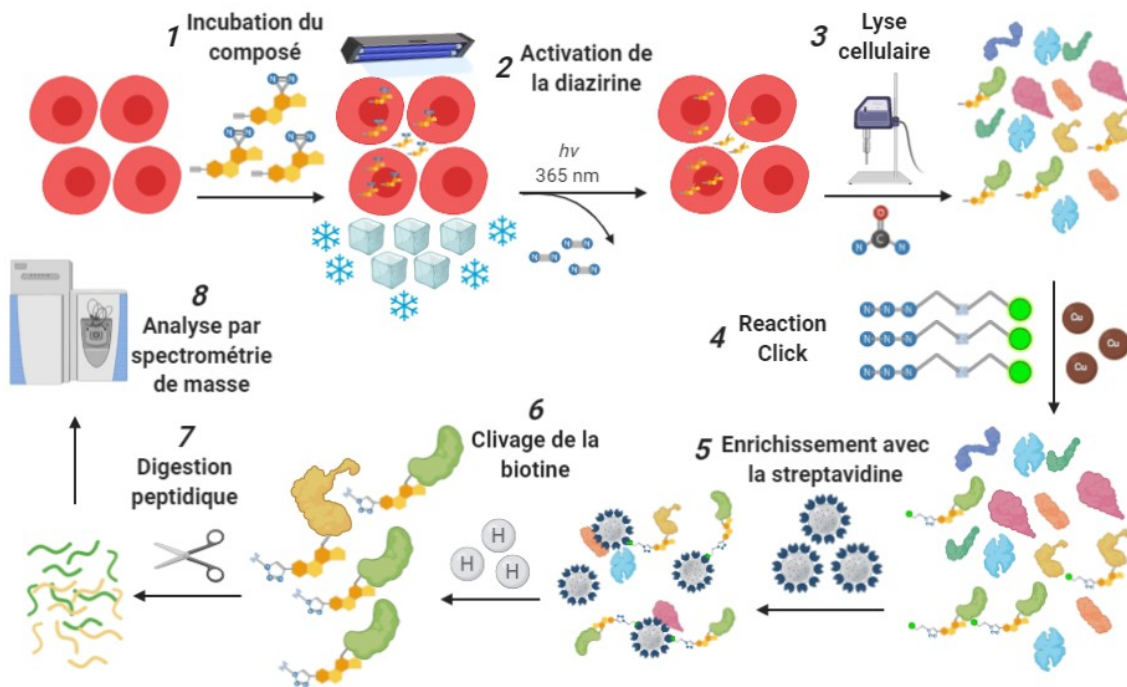


Figure 73: Procédure de chimo-protéomique qui sera utilisée pour l'identification de la cible de UM171. Cette procédure se décline en 8 étapes allant de l'incubation des sondes dans les cellules jusqu'à l'injection des peptides dans le spectromètre de masse.

3.2.2 Incubation des sondes

La première étape de la procédure de chimioprotéomique est l'incubation des sondes dans les cellules. Bien que l'expérience sur l'expansion des cellules souches ait permis de déterminer les valeurs de $EC_{1,5}$, aucune information n'a été obtenue concernant le temps nécessaire pour obtenir une activité cellulaire. Étant donné que cette expérience se déroule sur 7 jours, le temps minimal pour obtenir une réponse cellulaire n'est pas connu. Il est important de connaître cette information pour savoir le moment optimal pour irradier les cellules de façon à maximiser les chances de lier la protéine d'intérêt. Il a été observé par le groupe de Larsson et collaborateurs qu'UM171 interagissait de façon indirecte avec la machinerie du complexe RCOR/LSD1.¹¹⁰ Une des conséquences de cette interaction est la dégradation de RCOR et LSD1 par le protéasome. Il est connu dans la littérature que les inhibiteurs de LSD1 tel que 2-PCPA sont responsables d'empêcher la déméthylation de la lysine (K4) sur l'histone 3. Cette position H3K4 régule la différenciation des gènes de cellules souches hématopoïétiques.^{110, 111} Cependant UM171 n'agit pas sur LSD1 comme a pu le démontrer Larsson. De ce fait, il est possible que la dégradation du complexe LSD1/RCOR puisse déterminer de façon indirecte un temps d'incubation optimal des sondes. Afin de connaître le temps d'incubation pour les sondes les biochimistes du groupe de Pr Guy Sauvageau ont effectué une petite expérience avec le composé UM171 dans les cellules AML1 (Figure 74). L'idée a été d'incuber la molécule à 250 nM et de varier le temps d'incubation. Une fois ces incubations terminées, il y a eu une lyse cellulaire avec un détergent non-ionique (NP40) et préparation d'un gel SDS-PAGE. C'est par un buvardage de Western et l'utilisation d'un anticorps reconnaissant RCOR qu'il a été possible de suivre la dégradation de RCOR dans les cellules AML1.¹¹²

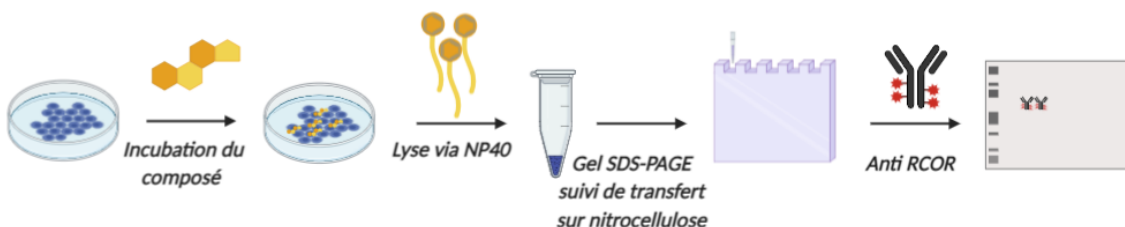


Figure 74 : Manipulations pour la détermination du temps d'incubation pour le composé UM171

L'équipe a déterminé qu'après un temps de 4 heures d'incubation, il y a dégradation complète de RCOR et une légère diminution de LSD1. Ce résultat est très indicatif pour le temps d'incubation et cela s'appuie sur les résultats du groupe de Larsson qui avait observé une dégradation complète de RCOR après 6 heures, mais avec une concentration de 140 nM de composé.¹¹⁰ Une seconde information qui a été déterminée par les biologistes est que la molécule d'UM171 devient toxique lorsque la concentration dépasse 5 μM , causant ainsi de la mortalité cellulaire. En se basant sur les résultats préliminaires des membres du laboratoire du Pr Guy Sauvageau, le choix s'est arrêté sur une incubation de 4 heures avec une concentration maximale de 5 μM pour les sondes photochimiques. Comme les sondes sont moins actives que UM171, une concentration de 5 μM devrait être en mesure de dégrader RCOR tout en limitant l'apparition de toxicité.

3.2.3 Irradiation des sondes

Le second paramètre à vérifier est la réaction de photoactivation de la diazirine avec la lampe. Il est rapporté dans la littérature que la fonction diazirine, qu'elle soit aliphatique ou aromatique, est activée avec une longueur d'onde avoisinant 350 nm.^{113, 114} La première expérience de cette section a été de vérifier l'intensité et la longueur d'onde de la lampe utilisée pour la procédure. Le spectre d'émission est illustré à la figure 75. Il s'avère que la lampe possède un maximum d'émission à 366 nm avec une plage très restreinte à cause d'une émission atomique. L'intensité de ce type de lampe est de 20,000/5,000 $\mu\text{W}/\text{cm}^2$, donc très puissante.¹¹⁵

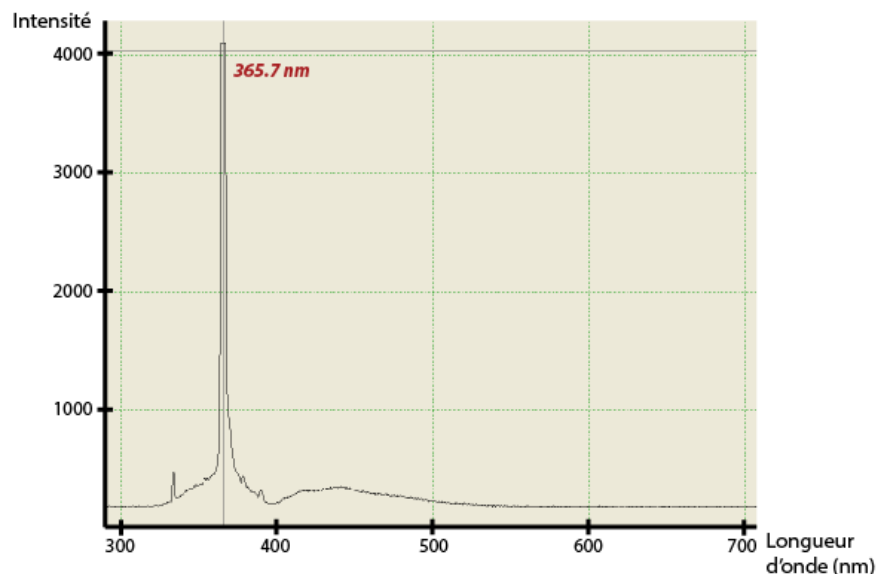


Figure 75: Spectre d'émission de la lampe B-100AP High Intensity de UVP qui possède un maximum d'émission vers 365.7nm.

La prochaine preuve de concept qu'il était nécessaire de démontrer était la capacité même des sondes à lier une entité et d'en déterminer l'efficacité. L'approche qui a été adoptée est celle d'utiliser des peptides et de faire la démonstration d'une modification chimique par les sondes. En 2017, un papier a été publié par le groupe du Pr Schriemer où ils ont mesuré la distribution d'insertion dans les acides aminés selon diverses sondes diazirines.¹¹⁶ Comme preuve de concept la procédure qui a été mise de l'avant pour tester les sondes de UM171 s'est inspirée de ce papier pour avoir une confiance dans les résultats obtenus et avoir un aspect comparatif. Les peptides modifiés par les sondes présentent un patron de fragmentation pouvant rendre difficile la confirmation du peptide lors du MS/MS. Afin de mieux distinguer les faux positifs et les vraies identifications, une étude de fragmentation des sondes a été réalisée préalablement avec l'Orbitrap. L'idée ici fut d'infuser un cocktail composé des sondes et de choisir sélectivement un ion précurseur avec le quadripôle et ensuite faire varier l'énergie de la cellule de collision pour en observer les différents fragments. Avec ce type d'analyse, il est possible de mesurer l'intensité d'ions précis en fonction d'une énergie de collision donnée. Ces analyses offrent de belles mosaïques permettant de retracer les fragments de sondes selon l'énergie donnée au système. Le travail informatique a été possible grâce au coup de main d'un collègue, Mathieu Courcelles. Une couleur verte indique une grande abondance et une couleur rouge une faible abondance. La tendance observée par cette expérience est que pour une faible énergie de collision, on observe des fragments de grandes masses. Au point tel qu'il est possible d'observer la masse intacte dans le cas de 10 NCE (Énergie de collision normalisée), une quantité d'énergie corrélée au rapport masse/charge des différents ions, de telle sorte que les ions avec un rapport plus élevé recevront une plus grande quantité d'énergie que ceux ayant un plus petit rapport.¹¹⁷ La valeur de 10 représente 10% de l'énergie pouvant être fournie par l'appareil et se mesure en eV.^{26, 117} Cependant plus on augmente en énergie plus on voit apparaître des fragments de plus petites masses. Cependant, il a été intéressant d'observer des fragments communs entre les sondes soit 373.17; 345.13; 267.09; 240.08 et 91.05 m/z. L'énergie qui est nécessaire pour avoir une belle fragmentation d'un peptide est entre 20 et 30 NCE, une énergie plus grande offre peu d'ions nécessaires pour l'identification.¹¹⁸ Par conséquent, les fragments des sondes que l'on devrait principalement observer devront se trouver dans cette zone d'énergie de collision.

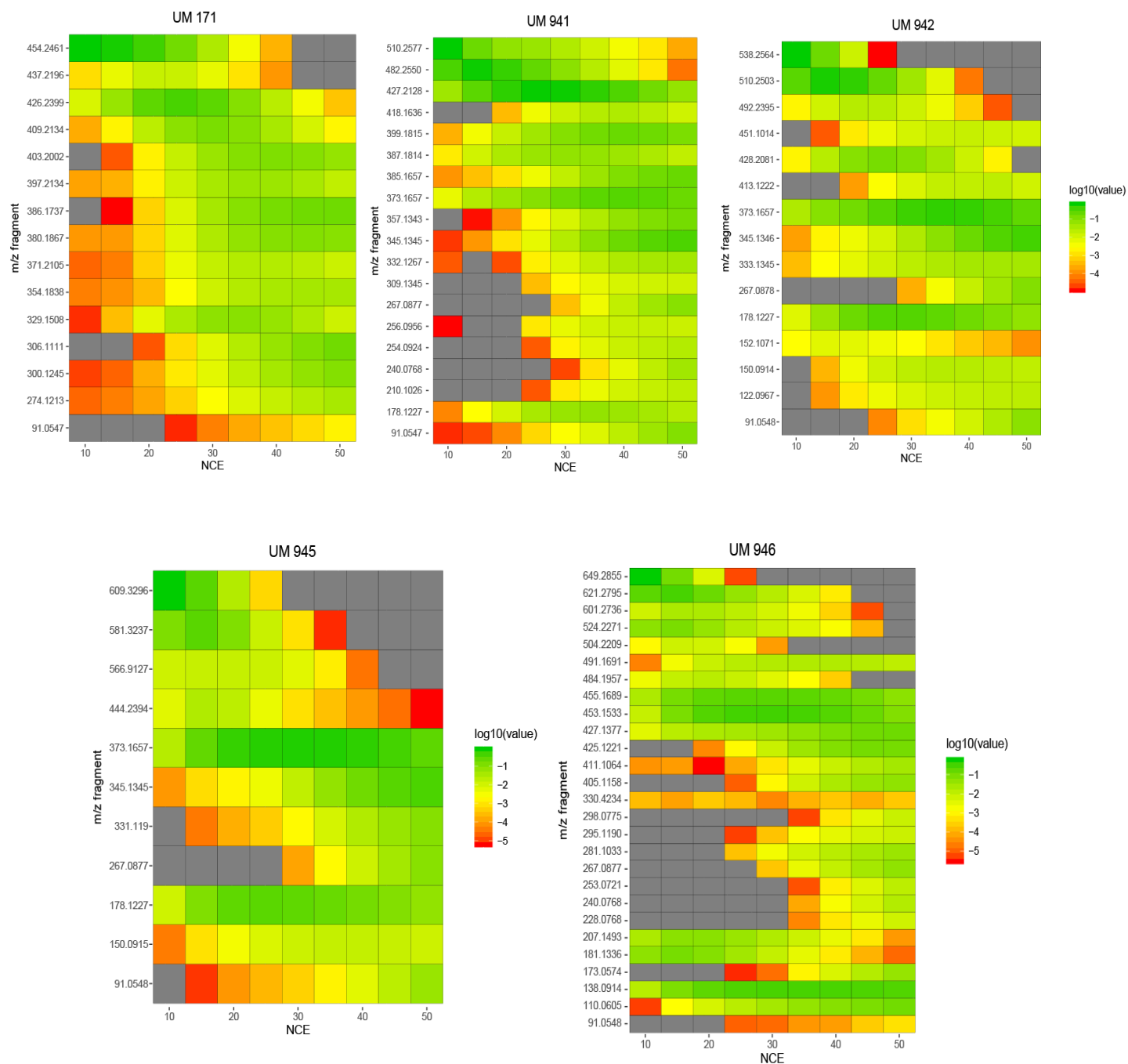


Figure 76 : Cartes de chaleur des diverses sondes et le composé UM171 qui mettent en corrélation les ions fragments en fonction de l'énergie de collision utilisée. Une couleur grise signifie une absence. Les ions sont classés selon une gradation allant du rouge à une couleur verte signifiant une faible intensité à une forte intensité

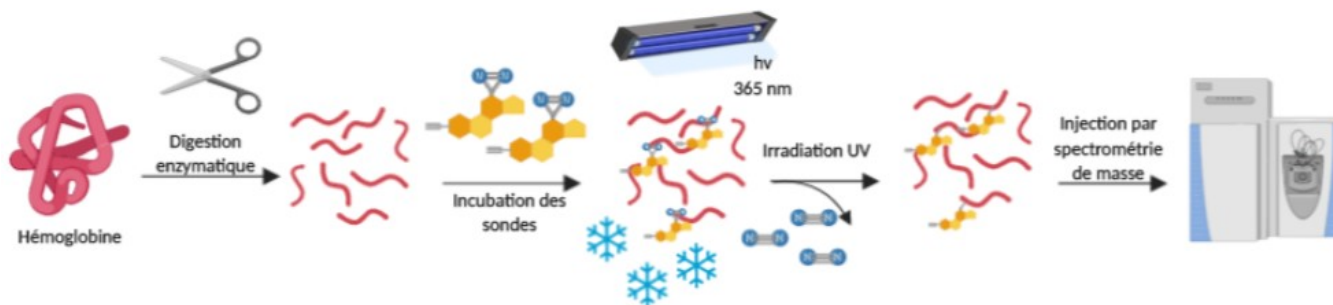


Figure 77 : Schéma expérimental permettant d'observer des peptides modifiés par les sondes de UM171. Les peptides de l'hémoglobine ont été incubés avec les sondes et irradiés sous rayonnement UV avant d'être analysés par spectrométrie de masse.

Une fois cette information en main, il a été possible de faire l'expérience avec les peptides. Pour ce faire l'hémoglobine a été digérée avec la trypsine pour offrir de multiples sites de liaison pour le carbène. La procédure qui a été suivie s'inspire grandement de la procédure décrite par le groupe de Schriemer débutant par la préparation d'une solution de peptides et une seconde préparation de sondes. Ces dernières ont été mélangées dans un petit capillaire et gelées à l'azote liquide puisque l'efficacité d'insertion du carbène est maximisée en phase solide.¹¹⁶ Une irradiation de 30 minutes sous glace sèche a été effectuée avec la lampe UV. Ensuite les peptides ont été solubilisés dans le tampon d'élution et injectés dans un QExactive.

Il a été intéressant de constater la présence de fragments identifiés préalablement dans l'expérience des énergies des sondes intactes sur les peptides modifiés. Si on observe la figure 78, il est possible de retrouver les ions 510,25 et 373,17 m/z dans le spectre ms/ms du peptide FFESFGDLSTPDAVMGNPK de l'hémoglobine. Dans ce présent spectre, la sérine est modifiée par la sonde et la couverture de la séquence est relativement très bonne. Afin de confirmer la présence d'une identification, les deux critères ont été les suivants : la présence d'ions fragments et une couverture supérieure à 75 %. Avec ces deux critères en main, il a été possible d'obtenir les préférences d'insertion de chaque sonde en plus de déterminer le pourcentage d'efficacité de la réaction en comparant le nombre de peptides modifiés sur la totalité identifiée.

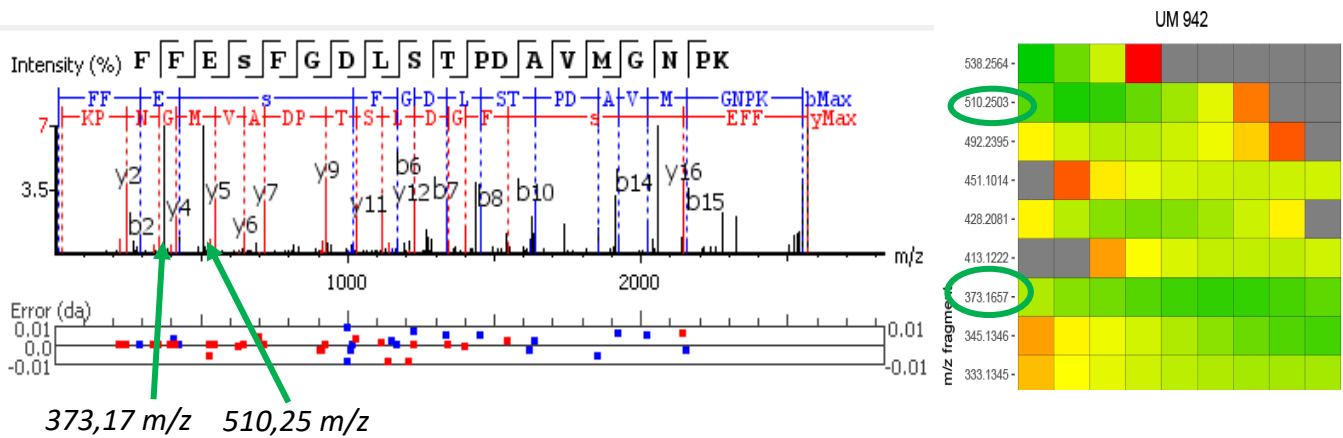


Figure 78 : Exemple d’observations des ions fragments caractéristiques des sondes sur les spectres MS/MS des peptides de l’hémoglobine modifiées. Dans cet exemple, les fragments 373.17 et 510.25 m/z de la sonde ‘942 ont été retrouvés dans la séquence du peptide FFESFGDLSTPDAVMG NPK suite à une modification de la sérine en AA4.

L’une des façons de déterminer l’étendue d’une modification en spectrométrie de masse est de comparer l’appariement des spectres peptidiques (PSM : Peptides-Spectrum Match) entre deux populations. Le logiciel Peaks a comme objectif d’associer une séquence de la base de données avec les séquences observées dans les MS/MS. On décrit un PSM quand le logiciel identifie un spectre MS/MS avec une séquence de la base de données. Ainsi il peut y avoir plus d’un PSM pour un même peptide. Par conséquent, si on effectue le ratio entre le nombre de PSM pour les peptides modifiés par une sonde sur le nombre de PSM totaux on peut avoir un pourcentage de peptides modifiés sur l’ensemble des peptides identifiés. Tel qu’il est possible d’observer dans le tableau 13, le pourcentage de modification des peptides par une sonde est majoritairement inférieur à 20 % à l’exception des sondes ‘944 et ‘945 qui est de 28 et 37 %. Le faible rendement des diazirines est connu dans la littérature compte tenu de leur grande réactivité avec le solvant.⁷⁴ Des expériences non présentées dans ce manuscrit ont prouvé la présence de ces insertions en plus de réactions intramoléculaires dans le cas des sondes aliphatiques. Il n’est donc pas surprenant d’observer un faible taux de modification sur les peptides, il y a beaucoup plus de molécules de solvant que de peptides.

Tableau 13 : Pourcentage d'efficacité de liaison des sondes selon le ratio des PSM totaux et des PSM modifiés par une sonde

Sonde	PSM sonde / PSM totaux	% de modification
'941	193/1347	14
'942	300/2361	13
'943	256/1440	18
'944	416/1473	28
'945	3050/8149	37
'946	1085/6534	17

Le carbène est une espèce très réactive pouvant s'insérer dans la majorité des liens de la chimie organique, soit X-H où X peut être un carbone, un azote, un oxygène ou un soufre. Le groupe de Schriemer a fait une étude pour connaître la distribution de ces insertions parmi les acides aminés et a relevé que les insertions sur les hétéroatomes étaient nettement majoritaires comparativement aux liens C-H.¹¹⁶ La figure 79 illustre le résultat obtenu avec les sondes de UM171 pour une étude similaire.

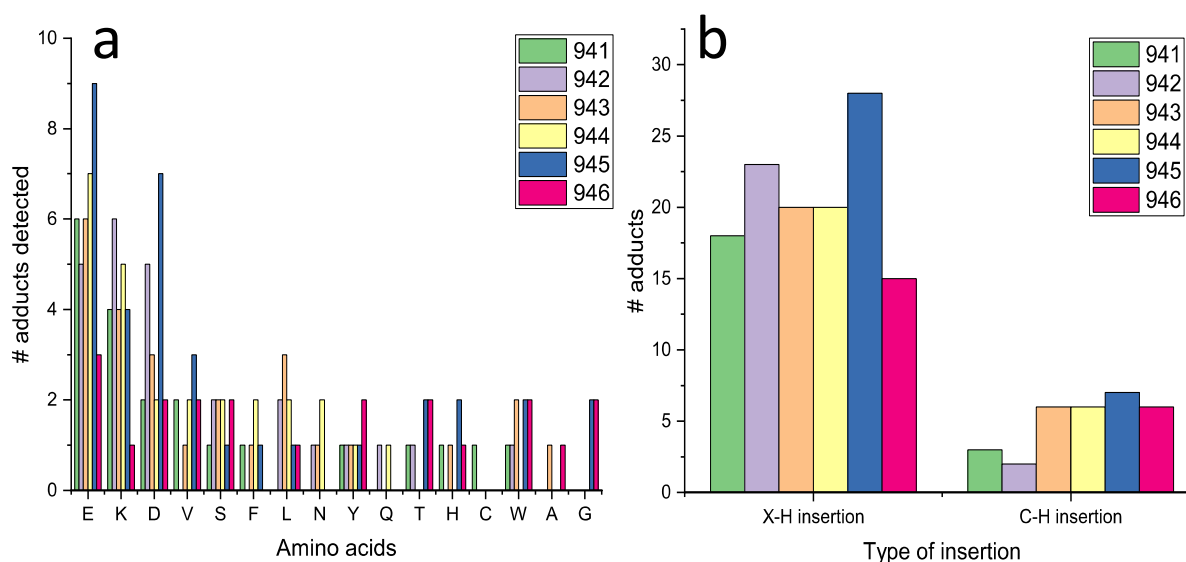


Figure 79 : Distribution statistique de l'insertion du carbène de chaque sonde dans les acides aminés. En a) on retrouve la proportion de modification individuelle sur chaque acide aminé tandis qu'en b) on retrouve la proportion de modification dans un lien hétéroatomique X-H versus un lien carboné C-H.

Les peptides qui furent majoritairement modifiés sont l'acide glutamique, la lysine, l'acide aspartique et la sérine, tous des acides aminés dont les chaînes latérales possèdent des hétéroatomes. Si l'on compare spécifiquement le nombre d'adduits sur des hétéroatomes par rapport au nombre d'adduits sur des atomes de carbone, on observe une nette supériorité des hétéroatomes comme l'illustre la figure 79b. Cette expérience a permis de confirmer le résultat présenté par le groupe de Schriemer, mais surtout de démontrer l'efficacité des diazirines dans la procédure de chimioprotéomique.

3.2.3 Lyse cellulaire et optimisation de la réaction Click

Une fois l'irradiation complétée, la prochaine étape de la procédure est la réaction de cycloaddition [3+2] nécessaire pour l'ajout de la biotine aux sondes. Cette réaction nécessite la présence de cuivre (I), d'un ligand tridenté aminé et le couple alcyne/azoture. Pour cette réaction, il y a eu utilisation du sulfate de cuivre qui est d'oxydation (II). Afin de réduire le cuivre, il y a eu utilisation du réducteur ascorbate de sodium. Le ligand THPTA est composé de trois triazoles augmentant beaucoup la solubilité et il a été rapporté que les ligands aminés accélèrent grandement la réaction en donnant une densité électronique au cuivre pour l'insertion dans l'alcyne.¹¹⁹ Afin d'effectuer cette réaction avec le milieu cellulaire, il faut par conséquent permettre aux réactifs d'entrer en contact avec le contenu cellulaire. Ainsi, pour maximiser l'exposition des réactifs avec les sondes qui se trouvent intracellulaire, il est nécessaire de lyser les cellules. Pour ce faire, il existe de multiples méthodes, que ce soit de façon chimique, mécanique, biologique ou physique.¹²⁰ Cependant le choix s'est arrêté avec la méthode chimique à cause de son efficacité. On retrouve les tampons anioniques (SDS), les tampons non ioniques (Triton, NP40 et Tween) et les tampons chaotropiques (Urée, Guanidine).¹²¹ Chaque famille de tampons possède des différences, que ce soit pour la dénaturation, le bris sélectif de divers compartiments cellulaires et de compatibilité. Toutefois un papier sorti en 2013, illustre la compatibilité de plusieurs tampons de lyse pour la réaction de cycloaddition de Huisgen.¹²² Afin de déterminer le choix du solvant de lyse optimal, une expérience similaire à ce papier a été réalisée. L'idée a été de lyser des cellules avec les concentrations suggérées dans la littérature (1% ou 8M) et de diluer chaque lysat pour obtenir une concentration protéique de 1 mg/mL (Figure 80).¹²² Ensuite, il y a eu ajout de la même quantité de la protéine BSA fonctionnalisée d'une fonction alcyne dans chaque lysat

et la réaction de cycloaddition [3+2] a été réalisée en ajoutant le même nombre d'équivalents de réactifs et le même temps de réaction.

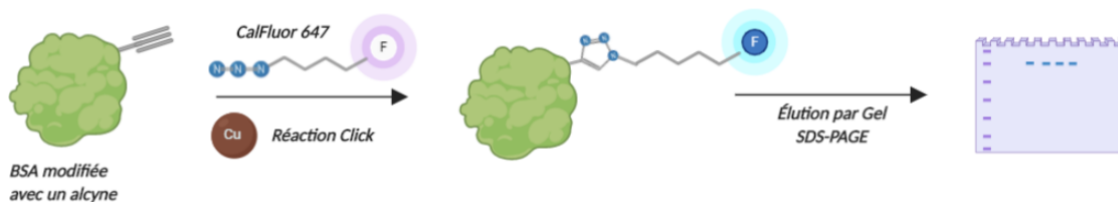


Figure 80 : Détermination du tampon de lyse optimal selon une étude comparative de l'intensité de fluorescence par gel de la BSA. Il y a incubation de la BSA fonctionnalisée par un alcyne avec le Calfluor 647 fonctionnalisé avec un azoture dans divers lysats cellulaires. L'intensité de fluorescence de la BSA est analysée par gel pour comparer les diverses conditions.

Sur le partenaire azoture, il y a eu choix d'un fluorophore permettant d'observer qualitativement par gel l'effet du solvant de lyse sur la réaction Click. La technologie des Calfluor développée par le groupe du Pre Carolyn R. Bertozzi qui permet d'obtenir une réponse fluorescente uniquement si la cycloaddition se produit.¹²³ Le principe de cette technique repose sur l'utilisation d'une molécule fluorescente qui s'auto désactive intramoléculairement à cause du principe de transfert d'électron photoinduit. Dans le cas ici présent, le fluorophore est un système conjugué (fluorescéine) et l'espèce qui vient neutraliser l'énergie d'émission est une charge (azoture). Cependant, lorsque le triazole est formé la charge n'est plus présente et la fluoresceine peut émettre un photon (Figure 81).

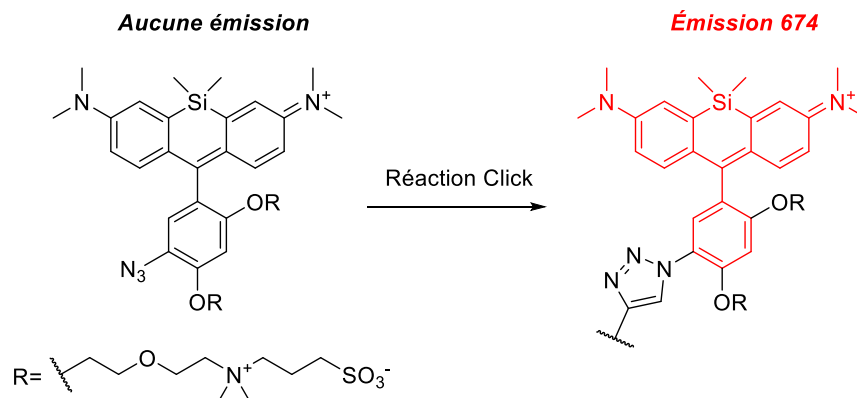


Figure 81 : Illustration de la technologie Calfluor pour le composé CalFluor 647. Le composé azoture n'émet aucune fluorescence suite à une excitation à 655 nm puisque les orbitales du groupement azoture désactivent l'excitation, cependant le triazole formé suite à la réaction Click ne permet pas cette désactivation. Ainsi une émission lumineuse est observée à 674 nm.

Cette technologie permet d'éliminer complètement le bruit de fond non-spécifique qui aurait été observé par la présence de fluorophore non-liée à la protéine. Il a été rapporté que les fluorophores lipophiles ont tendance à s'adsorber sur les protéines et biaiser les résultats.^{123, 124} Par l'utilisation de cette technologie, on s'assure que le signal fluorescent résulte uniquement de la réaction de cycloaddition. Le résultat de cette étude des tampons de lyse est présenté à la figure 82.

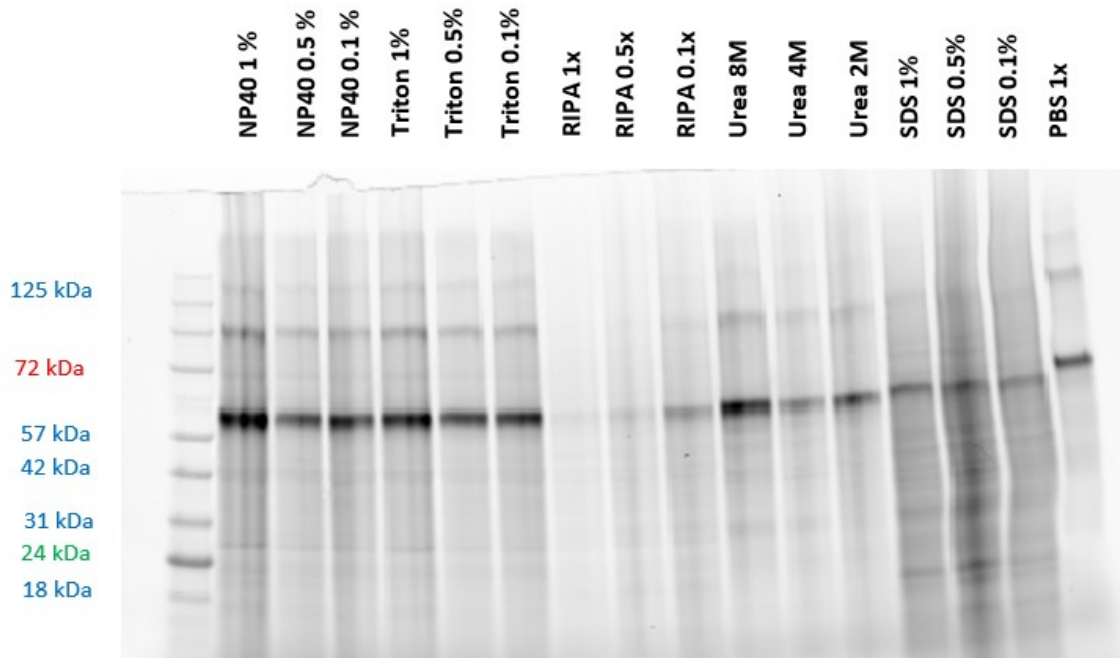


Figure 82 : Réponse de fluorescence à 674 nm de la BSA suite à une excitation à 647 nm pour le gel comparant 5 solvants de lyse à diverses concentrations et l'efficacité de la réaction Click.

La conclusion qui a été rapportée par le groupe de Verhelst fut que les tampons de lyse RIPA et Urée occasionnait une diminution de l'efficacité de la réaction Click, mais qu'en diluant d'un facteur 10 la concentration de ces agents de lyse, il était possible d'améliorer l'efficacité de la réaction.¹²² La figure 82 illustre le résultat d'une étude similaire effectuée avec 5 tampons de lyse à 3 dilutions différentes de l'agent lysant. La BSA possède un poids de 66.5 kDa, donc la bande la plus intense correspond à cette dernière ce qui est démontré par le contrôle PBS.

La première constatation est que les tampons RIPA et SDS diminuent considérablement l'efficacité de la cycloaddition de Huisgen, mais qu'en diluant l'agent lysant, il est possible d'augmenter légèrement cette dernière. Pour ce qui est des autres tampons de lyse, il n'y a pas d'effet négatif sur la réaction de Click. Les résultats concordent tous à ce qu'avait démontré le groupe de Verhelst à l'exception de l'urée qui dans le cas présent n'influence pas la réaction formant le triazole, et ce pour deux réplicats. Le tampon de lyse qui fut sélectionné suite à ce résultat fut l'urée 8 M puisque cet agent chaotropique permet une lyse complète des cellules exposant ainsi tout le contenu des organelles.¹²⁰ Ceci n'est pas le cas lors d'une lyse avec les détergents non-ioniques tels que le Triton et le NP40 qui offrent une lyse dite plus douce. Pour maximiser la lyse et comme la Click ne semble pas être influencée par l'urée, il a été jugé judicieux d'utiliser ce tampon malgré le résultat de la littérature.

Un second aspect qu'il fut nécessaire de vérifier est la température et le temps nécessaire pour effectuer la réaction Click. Comme toute réaction organique, il est important de vérifier la cinétique d'une réaction. Le nombre d'équivalents pour effectuer la réaction était largement en excès pour ne pas en tenir compte dans l'optimisation. Pour déterminer les conditions optimales en ce qui a trait à la température et au temps, le même type d'expérience a été réalisé que celui effectué pour les tampons de lyse, à l'exception que cette fois-ci, le solvant de réaction est le PBS. Les températures qui furent sélectionnées sont 4, 25 et 37°C et les temps de réaction qui furent choisis sont 30 minutes, 2 et 16 heures. Le résultat de cette expérience est présenté à la figure 83.

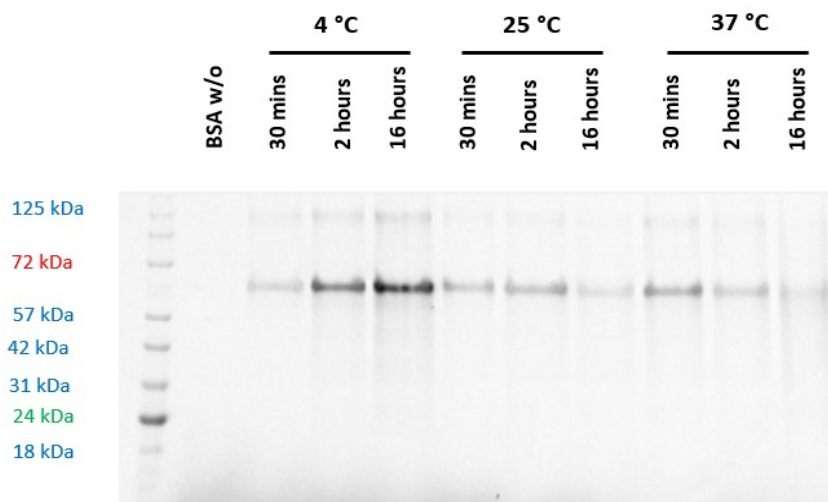


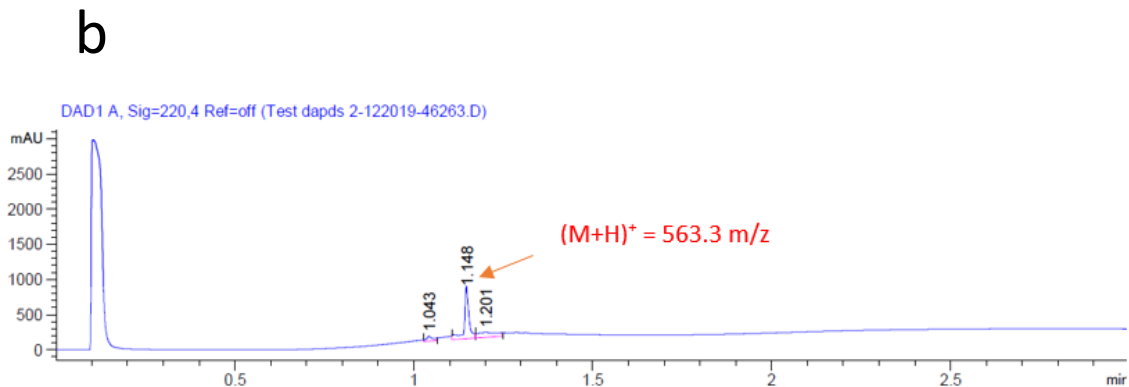
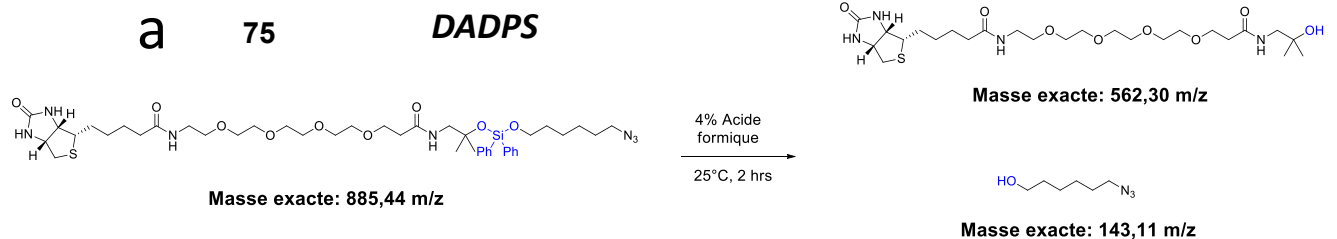
Figure 83 : Étude cinétique en fonction de la température pour la réaction Click selon une analyse de la fluorescence par gel de la BSA occasionné par la liaison avec le CalFluor 647.

Pour le contrôle négatif où la réaction de Click a été faite sans cuivre, on n'observe aucune bande de BSA. Maintenant, deux conclusions peuvent être tirées de cette expérience. Si l'on souhaite effectuer une réaction Huisgen sur une courte période de temps, il est préférable de le faire avec une température élevée. On observe une intensité croissante pour le temps de 30 minutes en fonction de la température. Cependant, il s'avère que pour les temps de 2 heures et de 16 heures cette corrélation n'est plus valable. Il est curieux de penser qu'une réaction chimique puisse être moins efficace en laissant un substrat plus longtemps, à moins que ce dernier se dégrade. Ce que le gel illustre est que pour un temps de réaction long, il est préférable d'utiliser une basse température pour empêcher la dégradation du substrat étiqueté par un fluorophore. Simplement en observant le gel, on constate que la bande de fluorescence est beaucoup plus intense en utilisant une combinaison de température/temps de 4°C/16 heures. Il s'agit de la condition qui fut sélectionnée pour l'expérience de chimioprotéomique.

3.2.4 Enrichissement et clivage des billes

La dernière étape critique de la procédure consiste à la capture des protéines ayant été modifiées par les sondes et qui possèdent maintenant l'étiquette biotine. La façon de faire consiste à utiliser des billes d'agarose fonctionnalisées avec la streptavidine. L'interaction non covalente entre la streptavidine et la biotine sera suffisamment forte pour attraper les protéines modifiées. Ensuite surviennent des lavages et l'isolement des protéines enrichies pour les injecter en spectrométrie de masse. L'une des méthodes pour isoler les protéines consiste à les digérer directement sur les billes par la trypsine et injecter le surnageant.^{113, 125, 126} Cependant, cette technique occasionne deux problèmes majeurs. Le premier est la présence de peptides de streptavidine qui saturent le signal dans l'analyseur compte tenu de l'abondance relative supérieure de cette dernière par rapport aux protéines enrichies qui sont nettement inférieures. La conséquence de ceci est une perte d'identification pour les protéines de faible abondance. Le second problème de cette méthode est la présence de protéines naturellement biotinylées qui se retrouvent dans les données résultantes en spectrométrie de masse. Ces protéines rendent l'analyse finale plus complexe dans le cas de la recherche d'une protéine inconnue. Pour ce faire une nouvelle technologie a vu le jour pour offrir une solution à cette problématique et il s'agit des biotines dites clivables.^{125, 126} L'idée de cette technologie est d'ajouter une fonctionnalité se trouvant sur l'étiquette avec la biotine qu'il est possible de cliver avant la digestion. De multiples groupements

clivables ont été développés dans la littérature en fonction des divers besoins expérimentaux et de la compatibilité des protocoles. Pour ce faire, il existe des groupements pouvant être clivés par des milieux alcalins, acides, réducteurs, oxydants, photoclivants et enzymatiques.¹²⁶ Pour aider avec le choix de groupement clivant pour la procédure, une étude comparative est sortie en 2010 et comparait justement diverses biotines clivables selon l'efficacité du clivage et la sélectivité de l'enrichissement de ces dernières.¹²⁵ Il s'avère que la fonction dialkoxydiphénylsilane **75** possède le portrait le plus attracteur parmi les sondes avec une sélectivité hors pair et un excellent taux de clivage. Cette fonction est aussi très attrayante pour la protéomique à cause de son clivage en milieu acide qui est compatible avec la spectrométrie de masse. Une seconde fonction qui s'est trouvée très intéressante parmi celles proposées dans la littérature est la sonde trifonctionnel Dde-TAMRA **76** qui possède une combinaison avec la fonction clivable 1-(4,4-diméthyl-2,6-dioxacyclohexylidène)éthyl (Dde) par l'hydrazine et un fluorophore, le TAMRA. Ceci permet de suivre visuellement l'enrichissement et le clivage des protéines des billes de streptavidine. L'une des expériences qu'il est important de faire est de vérifier l'efficacité des clivages de ces dernières avec les conditions expérimentales utilisés lors de la procédure de chimioprotéomique. La première étape a été de vérifier par analyse LCMS la résultante d'un clivage des groupements selon les conditions de la littérature. (Figure 84)



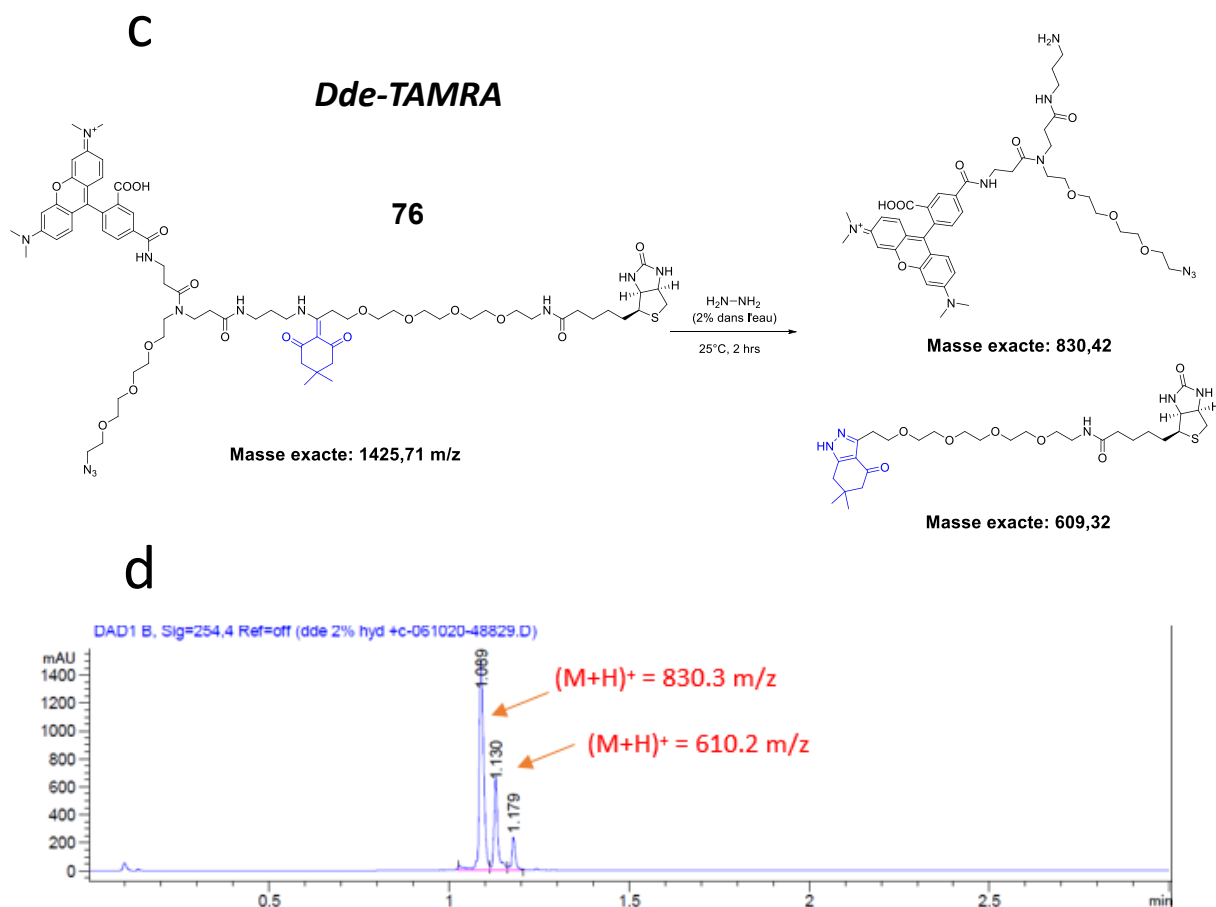


Figure 84 : Preuve *in vitro* des conditions de clivages des biotines clivables. En a) et en c) on retrouve les équations chimiques illustrant les sous-produits des réactions de clivage. En b) et en d) on retrouve les LCMS des réactions de clivages ainsi que les ions observés.

Comme le démontre la figure 84, les biotines clivables ont bel et bien été clivées avec les conditions utilisées. La sonde dialkoxydiphénylsilane **75** est clivée avec une solution de 4% d'acide formique avec une agitation de 2 heures à température pièce. Par l'analyse LCMS, il n'a pas été possible d'observer la présence du deuxième composant compte tenu de sa faible masse et de l'absence de chromophore pour les détecteurs MS et UV. Cependant comme il y avait perte du pic du réactif initial et apparition du composé de masse 562,3 m/z on peut conclure que le clivage a bel et bien eu lieu. Maintenant pour la biotine trifonctionnelle composée du TAMRA **76** et du groupement Dde, cette dernière est clivée avec une solution d'hydrazine 2% durant un temps de 2 heures. Il est possible de constater clairement la présence des deux produits de la réaction de clivage de l'hydrazine sur le composé initial. D'autant plus que le réactif initial a disparu du

chromatogramme. Ces deux indices illustrent un clivage complet de la fonction clivable de l'espèce moléculaire. Ce premier test a permis de démontrer chimiquement que les deux biotines clivables utilisées sont clivées par les conditions utilisées par la littérature. Suite à ce résultat, il était important de vérifier l'efficacité de ces biotines dans un essai plus près des conditions qui seront utilisées expérimentalement, soit dans un lysat cellulaire (Figure 88).

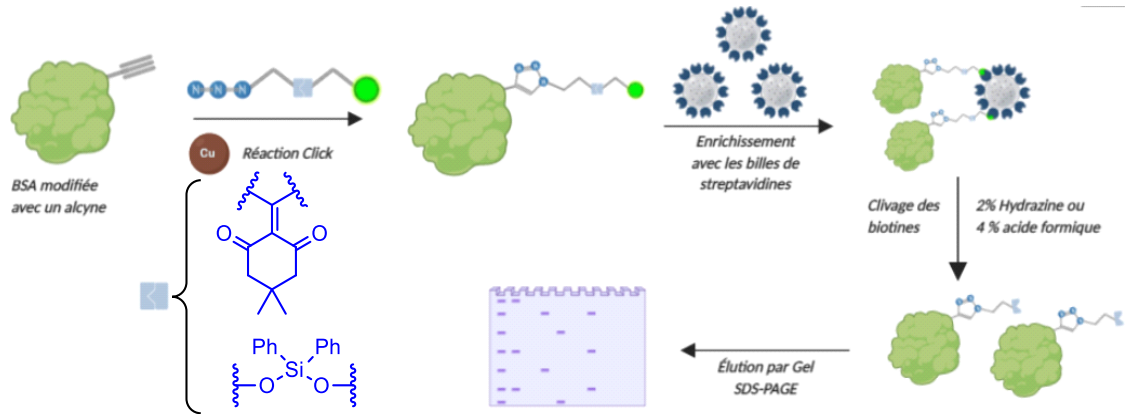


Figure 85 : Procédure expérimentale démontrant l'efficacité des conditions de clivages *in vivo* des biotines. Une réaction Click est réalisée dans un lysat entre la BSA comportant un alcynes et une biotine clivable composé soit du Dde ou du dialcoxydiphénylsilane. Un enrichissement par des billes de streptavidine suivi d'un clivage et d'un gel sont les étapes de la méthode.

Pour effectuer cette optimisation, il y a eu utilisation d'une procédure très similaire à celle discutée dans la section précédente. Il y a eu lyse de cellules avec les trois meilleurs solvants, soit 1% triton, 1% NP40 et 8M urée, suivi d'une dilution de ces lysats par un facteur de 10 de façon à avoir une concentration de 1 mg/mL de protéines. Dans chacun de ces lysats a été ajoutée une quantité fixe de la BSA dotée d'un alcynes. L'idée a été d'utiliser les mêmes conditions de Click dans chaque lysat, mais en variant uniquement la biotine clivable. Le contrôle positif a été de faire la réaction dans le PBS et non dans un lysat complexe. Ensuite, la BSA biotinyllée a été enrichie avec les billes de streptavidine et chaque biotine traitée avec l'agent clivant. Le surnageant a été utilisé pour effectuer un gel qui a été soumis en premier lieu à une mesure de fluorescence et ensuite une coloration à l'argent. Les résultats sont démontrés à la figure 86.

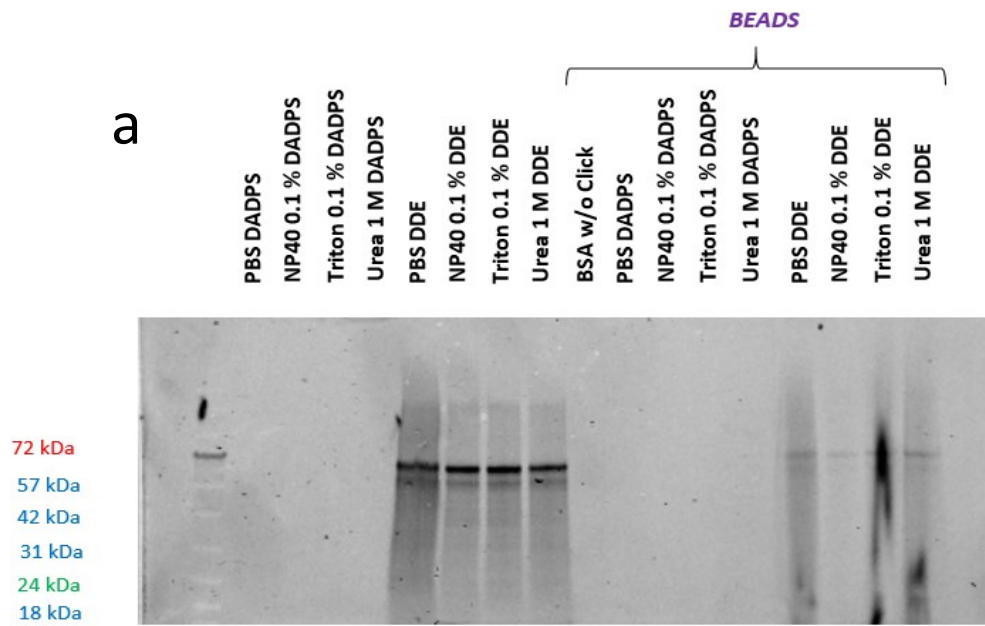


Figure 86 : Gels démontrant l'efficacité du clivage de la fonction clivable sur la biotine suite à la réaction Click dans divers solvants de lyse. En a) on retrouve un gel de fluorescence à 674 nm et en b) on retrouve une coloration à l'argent

Lorsque l'on observe la mesure d'émission à 600 nm on constate la présence de BSA pour chaque condition de lysat uniquement avec la sonde Dde-TAMRA. L'intensité est similaire entre le contrôle positif PBS et les différents lysats démontrant la réponse similaire de la biotine clivable dans lysat cellulaire versus un milieu aqueux simple. La condition de clivage fonctionne mais n'est pas parfaite, comme le démontre la faible réponse observée avec les billes. Cependant, c'est très faible et une réponse fluorescente est très sensible. Sans surprise, aucune réponse fluorescente n'a été observée avec la sonde DADPS en raison de l'absence du fluorophore TAMRA. Afin de vérifier ce dernier, il a été nécessaire d'effectuer une coloration à l'argent. Cette coloration permet une sensibilité de 5 ng, contrairement à une coloration au bleu de Coomassie qui est de l'ordre de 100 ng.¹²⁷ Pour maximiser la capacité de détecter la BSA, il y a eu utilisation d'une coloration plus sensible qu'offre l'argent. Ce qu'il est possible de constater est la présence de BSA dans les échantillons provenant du clivage de la fonction dialkoxydiphénylsilane. Il semble y avoir des différences parmi les solvants de lyse, comme quoi l'urée 1M semble démontrer plus de clivage que les autres solvants de lyse. Ce qu'il est curieux d'observer, est une disparité avec le contrôle positif qui devrait illustrer la meilleure intensité. Il est possible d'observer une quantité de BSA beaucoup moins importante sur les billes avec la biotine clivable DADPS, ce qui démontre une bonne efficacité de clivage. Ce qui est surprenant est la non-correspondance vis-à-vis la réponse de fluorescence et réponse de coloration à l'argent pour la seconde biotine Dde-TAMRA. On observe une faible réponse dans les échantillons du clivage et une plus grande réponse avec les billes, ce qui n'était pas observé avec la fluorescence. Malgré tout, la fluorescence offre une plus grande sélectivité du signal et semble plus fiable pour cette biotine. Toutefois, cette expérience démontre la capacité d'enrichir une protéine biotinylée possédant une fonction clivable, et ce dans un lysat cellulaire. D'après ce gel, la fonction DADPS semble mieux clivée que la fonction Dde avec le signal d'argent.

Après avoir vérifié cette dernière étape, la méthode est optimisée et la procédure de chimioprotéomique est prête à être testée avec les sondes analogues de UM171. Pour résumer : un temps d'incubation de 4 heures est nécessaire pour obtenir une réponse biologique, ensuite une irradiation à 365 nm permet l'activation de la diazirine, ce qui permet ensuite une lyse à l'urée de 8M pour entreprendre la réaction Click à basse température sur une durée de 16 heures. Pour finaliser, il y a un enrichissement des substrats biotinylés avec des billes de streptavidine où uniquement les protéines sélectivement modifiées par des biotines clivables pourront être clivées

avec l'acide ou l'hydrazine selon l'application en aval choisie. Ces protéines pourront soit être mises sur gel ou digérées pour finalement envoyer en spectrométrie de masse.

3.3 Procédure de chimioprotéomique avec les sondes de UM171

3.3.1. Présentation de la procédure avec les sondes UM171

L'optimisation de la procédure de chimioprotéomique a permis, dans un premier temps, de vérifier chaque étape, mais aussi d'en déterminer les meilleurs paramètres possibles pour maximiser les chances de réussir la procédure avec les sondes de UM171. La prochaine étape fut d'effectuer la procédure complète avec les sondes synthétisées préalablement. Il s'avère que plusieurs tentatives furent réalisées avant d'obtenir un résultat satisfaisant. Chacune de ces tentatives fut formatrice et amena de nouvelles complications. Notamment, lors du clivage de la biotine clivable, malgré une attention particulière lors de la prise du surnageant, il y a eu présence en grande quantité de streptavidine lors de la digestion trypsique. Cette conséquence a malheureusement occasionné la saturation du détecteur avec les peptides de la streptavidine et n'a pas permis d'identifier les peptides provenant des protéines enrichies qui sont nettement moins abondants. Pour éviter ce souci, la procédure initiale a été légèrement révisée de façon à séparer la protéine streptavidine des protéines désirées avant l'injection sur l'instrument. Pour ce faire, il y a eu confection d'un gel SDS-PAGE pour effectuer une séparation préalable des protéines et ensuite une digestion trypsique sur gel a été effectuée sur les bandes d'intérêts. (Voir la figure 87). Il s'agit d'une façon simple de se débarrasser de la streptavidine, puisque cette protéine tétramérique se dénature en 4 monomères de 13,2 kDa, qui se trouvent dans le bas du gel et ne masquent pas de protéines.

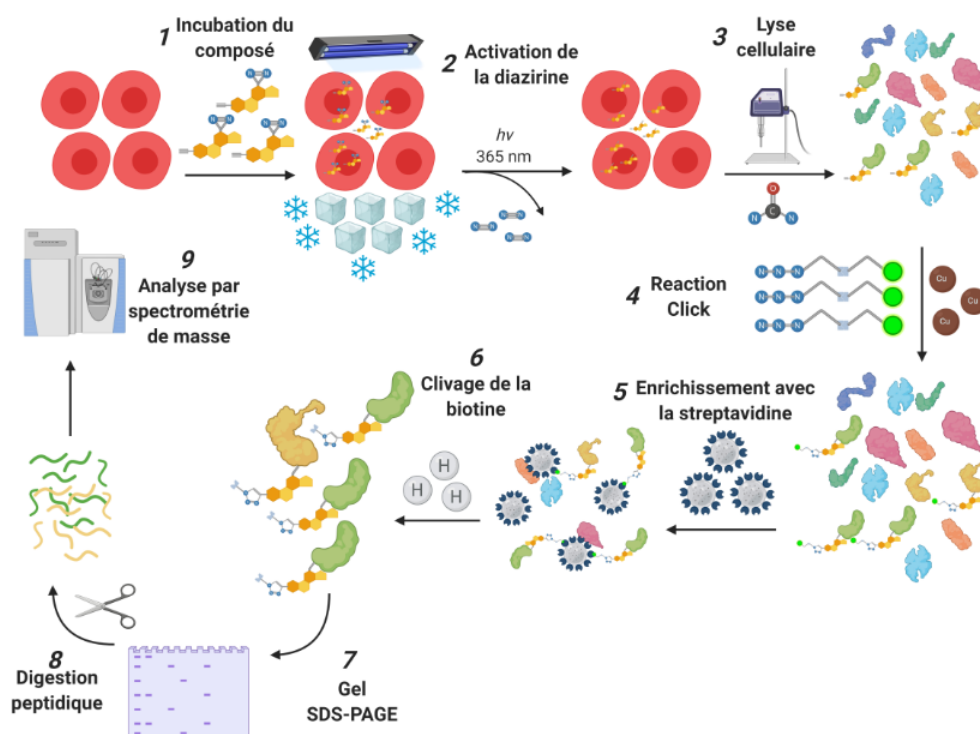


Figure 87 : Procédure modifiée de chimio-protéomique qui a été utilisée pour l'identification de la cible de UM171. Un gel SDS-PAGE a été réalisé pour séparer la streptavidine de l'analyse

La procédure finale a été réalisée selon deux conditions, soit en présence ou non de l'inhibiteur de protéasome MG132. Tel qu'il a été démontré dans la section 3.2.2 et dans la littérature^{110, 112}, la présence du composé UM171 ou des analogues actifs ont pour conséquence la dégradation de RCOR. Étant donné que la cible thérapeutique est inconnue, il est idéal d'utiliser l'inhibiteur de protéasome pour bloquer la machinerie responsable de dégrader les protéines. Il s'agit d'une précaution pour maximiser les protéines enrichies et empêcher leur dégradation si jamais ces dernières sont ubiquitynylées en présence de UM171. Les sondes '941 à '945 furent testés selon ces deux conditions tout comme le contrôle DMSO. Malheureusement la sonde '946 a été consommée en entièreté dans des essais n'ayant pas fonctionné, ainsi cette dernière n'a pas pu être utilisée lors de la procédure finale. Afin de se donner suffisamment de matériel, une quantité de 50 millions de cellules a été utilisée pour chaque condition. Ainsi la procédure a débuté par l'incubation de l'inhibiteur MG132 selon une concentration de 5 μ M durant une durée de 2 heures, qui fut ensuite relayée par l'incubation des sondes selon une concentration de 5 μ M et

ce durant 4 heures. Une fois le temps d'incubation complété, chaque culture cellulaire a été irradiée sous glace durant 30 minutes avec la lampe irradiante à 365,7 nm de la compagnie Jena.¹¹⁵ Étant donné qu'il y a 12 conditions distinctes, il n'était pas possible de les irradier tous au même moment compte tenu du petit diamètre de la lampe. Par conséquent, les moments initiaux d'incubation ont été débutés en décalage de manière à les irradier sans superposition. Lorsque l'irradiation a été complétée, les cellules ont été transférées dans des tubes et lysées avec une solution de 8 M d'urée dans 50 mM d'HEPES. La quantification de protéines obtenue suite à cette lyse donnait une quantité d'environ 4 mg de protéines pour chaque condition. Connaissant cette donnée, il a été possible d'obtenir une concentration de 2 mg/mL, permettant ainsi de diluer la concentration d'urée. Bien qu'aucune différence n'ait été observée dans la section 3.2.3, il a été judicieux de diluer la concentration d'urée pour maximiser l'efficacité de la réaction Click. La réaction de cycloaddition [3+2] a été effectuée sur 16 heures à une température de 4°C pour rester en accord avec l'optimisation effectuée. Le choix de la biotine clivable s'est arrêté sur la dialkoxydiphénylsilane. Par la suite, il y a eu précipitation des protéines avec un mélange de chloroforme/méthanol/eau selon des proportions de (1/4/4). Il existe divers solvants pour précipiter des protéines, telles que l'acétone, l'acide trichloroacétique et le chloroforme/méthanol, cependant il a été observé dans la littérature que la méthode du chloroforme permettait une récupération des protéines nettement supérieure aux autres méthodes.¹²⁸ Comme le méthanol est miscible à l'eau, le caractère hydrophobe du chloroforme augmente considérablement, rendant les protéines peu solubles dans cette phase et comme le chloroforme ne produit pas d'interactions favorables avec les protéines, ces dernières se trouvent à précipiter à l'interphase entre le méthanol/eau et le chloroforme. La phase eau/méthanol permet d'éliminer les composés polaires tels que les sels, l'acide ascorbique, le cuivre et le ligand aminé THPTA. Tandis que la phase chloroforme permet d'éliminer les espèces plus lipophiles telles que l'excès de biotine clivable n'ayant pas réagi. Il est important de retirer ces biotines libres avant l'ajout des billes de streptavidine puisqu'elles viennent occuper les sites de ces dernières et bloquer les éventuelles protéines que l'on souhaite enrichir. La précipitation est essentielle pour maximiser l'enrichissement des protéines biotinyllées par la réaction Click. Une fois la précipitation des protéines complétée pour chaque condition, il y a eu resolubilisation de ces dernières dans un tampon légèrement salé avec 0,5 % de détergents SDS jusqu'à l'obtention d'une solution translucide. L'étape qui a suivi est l'enrichissement des protéines modifiées par les sondes, avec des billes d'agarose fonctionnalisées avec la streptavidine. Cette étape dure 16 heures à 4°C selon

une agitation lente pour maximiser le temps d'interaction entre les billes et le lysat cellulaire. Par la suite, les billes ont été lavées intensément avec diverses conditions dénaturantes et neutres pour éliminer les protéines non spécifiques qui pourraient s'adsorber à la surface des billes. Finalement, les billes ont subi deux traitements subséquents de 4% d'acide formique durant 2 heures pour cliver la fonction dialkoxydiphénylsilane et libérer les protéines des billes de streptavidine. Le surnageant a été combiné et évaporé avec l'appareil Speedvac qui permet une évaporation rapide des solvants en réduisant le point d'ébullition par une diminution de la pression.¹²⁹ Cette évaporation du solvant permet d'éliminer par le fait même l'acide formique utilisé pour libérer les protéines des billes. Une fois les échantillons secs, il y a eu ajout du mélange pour colorer et dénaturer les échantillons nécessaires à la confection d'un gel SDS-PAGE. Ce mélange se compose de bleu de bromophénol, de SDS et de DTT. Le gel a été effectué et une coloration à l'argent a été opérée par la suite.

3.3.2. Résultats biochimiques et protéomiques de la procédure

La coloration à l'argent a permis de révéler les protéines qui ont été enrichies par la procédure de chimioprotéomique. Ce gel est présenté à la figure 88.

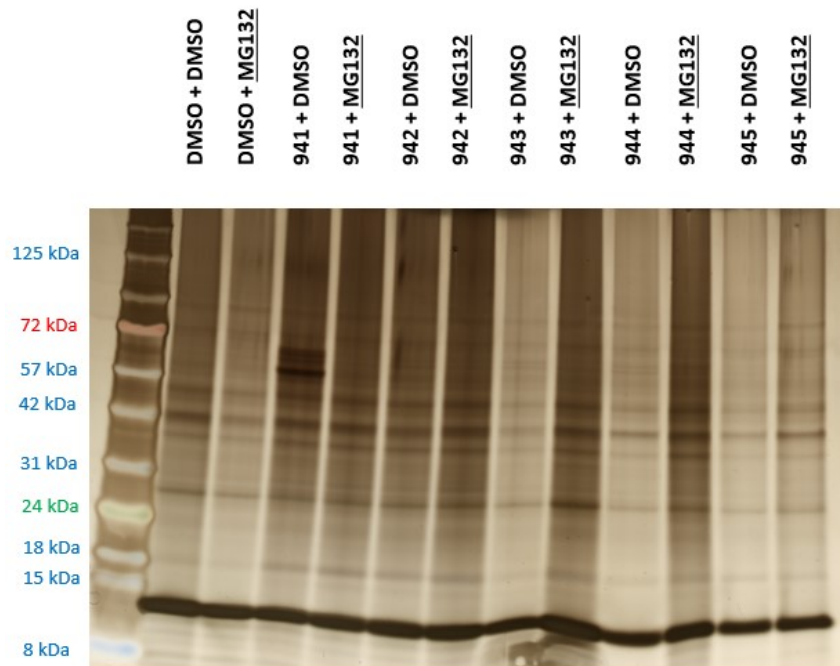


Figure 88 : Coloration à l'argent du gel SDS-PAGE des protéines enrichies par les sondes diazirines de UM171

Le premier constat qu'il est possible de relever sur le gel, est la présence du monomère de streptavidine près du 15 kDa dans tous les échantillons. Ceci permet d'apprécier l'efficacité de séparation qu'offre un gel d'acrylamide. Le second constat concerne l'importance de la présence des contrôles DMSO dans l'expérience qui permettent d'illustrer une grande quantité de bandes non spécifiques provenant des billes. La présence des contrôles permet ainsi d'exclure les bandes intenses que l'on retrouve dans les sondes qui pourraient sembler être intéressantes. L'observation la plus importante consiste en la présence de deux bandes spécifiques dans l'échantillon **941 + DMSO** dans la zone de masses entre 57 et 72 kDa. Ces bandes illustrent que la méthodologie a fonctionné étant donné que des protéines ont pu être enrichies, cependant la logique de cet enrichissement reste à discuter. Il est curieux d'observer uniquement ces bandes dans la condition DMSO et non la condition MG132. Il est possible de faire deux hypothèses pour expliquer ceci. Dans un premier cas, il est possible que ces protéines aient été dégradées par un second mécanisme que le protéosome. Il s'avère que dans le corps humain, il existe deux mécanismes qui régissent la dégradation des protéines, soit le protéasome et la protéolyse lysosomale.¹³⁰ Ces deux mécanismes sont compensatoires, dans le sens que si l'un mécanisme est inhibé, le second est surexploité pour compenser. Étant donné que le protéasome était inhibé, il est possible que la protéolyse lysosomale ait permis de dégrader les protéines comme ce mécanisme était plus favorisé. La seconde hypothèse serait que la sonde diazirine n'a pas eu le même comportement entre les deux conditions et que la résultante soit différente. Le dernier constat de ce gel est que seule la sonde peu active, **'941**, a donné des bandes intéressantes comparativement aux autres sondes qui sont plus actives. Ceci met en question la pertinence des bandes obtenues vis-à-vis l'activité intrinsèque des sondes, il peut s'agir de protéines non spécifiques. Ce gel illustre le comportement spécifique de chaque sonde dans une procédure de chimioprotéomique.

La prochaine étape a été de couper la région d'intérêt, soit de 50 kDa à 72 kDa, dans chacun des échantillons pour y effectuer l'analyse de spectrométrie de masse. La première étape est de décolorer les bandes coupées et ensuite de les couper en petits morceaux. Par la suite, les morceaux de gels sont déshydratés avec de l'acétonitrile qui permet d'extraire le plus de liquide possible. Finalement, les petits morceaux sont digérés avec la trypsine. Le surnageant a été évaporé avec le Speedvac et les échantillons ont été injectés sur l'Orbitrap Exploris™ de Thermo. Les analyses ont été effectuées sur le logiciel Peaks une fois les injections complétées. Comme le document résultant du moteur de recherche permet l'identification d'une multitude de protéines,

il est nécessaire d'effectuer des filtres pour écarter les faux positifs et les impuretés. Les filtres sont répartis selon deux types d'analyses, soit une analyse souple et une analyse stricte, tel que l'illustre la figure 89. Initialement, toutes les protéines qui ont été identifiées uniquement par l'entremise d'un peptide ont été retirées, par la suite la masse moléculaire a été le second filtre. Comme les bandes ont été coupées entre 50 kDa et 72 kDa, toute protéine ne possédant pas une masse dans cet intervalle s'est vu retirer. Finalement, ce qui distingue le caractère rigoureux de l'analyse est par l'application du dernier filtre. L'analyse stricte consistera à éliminer complètement toute protéine qui a été observée dans les conditions sans sonde que sont **DMSO** et **DMSO + MG132**. Afin d'éliminer la malchance qu'une protéine d'intérêt se trouve par un hasard malheureux dans les échantillons contrôles, un ratio d'intensité des protéines a été effectué entre les conditions avec sondes et les conditions sans sonde: il s'agit de l'analyse souple.

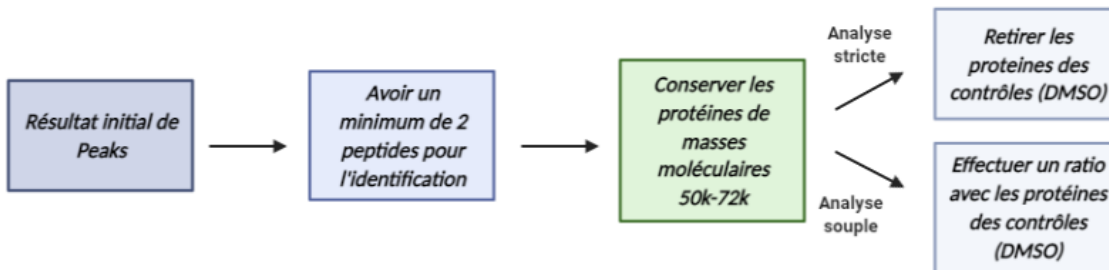


Figure 89 : Méthodologie qui a été appliquée pour filtrer les protéines identifiées par Peaks

Tableau 14 : Détermination du nombre de protéines identifiées dans chaque condition selon les deux types d'analyses effectuées

Condition	Analyse souple se basant sur un ratio d'enrichissement >2	Analyse stricte éliminant les protéines des contrôles
941 DMSO	181	73
941 MG132	72	40
942 DMSO	45	31
942 MG132	132	53
943 DMSO	47	24
943 MG132	18	18
944 DMSO	0	0
944 MG132	36	24
945 DMSO	1	1
945 MG132	9	9

Ce qu'il est possible de constater selon le tableau 13, est que les échantillons **941 DMSO** et **942 MG132** sont les échantillons qui ont permis le plus d'identifications. En spectrométrie de masse, il y a une corrélation entre la quantité de matériel injecté et le nombre d'identifications. Déjà sur le gel d'argent on pouvait entrevoir que ces échantillons étaient plus intenses que les autres. Malheureusement très peu de protéines ont été identifiées pour les sondes 944 et 945, qui sont des sondes très actives. Ce qui importe n'est pas le nombre de protéines identifiées, mais bien de connaître l'identité de ces protéines.

Pour ce faire, il y a eu réalisation de cartes de chaleur qui permettent d'observer visuellement l'abondance relative des protéines dans chaque condition. Le schéma 90 représente la carte de chaleur résultant de l'analyse stricte. Sur cette dernière il est possible de constater que la protéine

beta-hexosaminidase est la protéine la plus abondante identifiée dans l'échantillon **941 DMSO** et qui par le fait même a été aussi identifiée parmi d'autres échantillons. Comme cette protéine n'a pas été obtenue dans les contrôles, ceci augmente le degré de confiance. On constate aussi la présence de plusieurs autres protéines qui sont partagées entre les diverses conditions, comme la sérine hydroxyméthyltransférase et des modificateurs de l'ADN. La seconde carte de chaleur (figure 91) illustre les protéines distinctes qu'il a été possible d'observer selon les paramètres plus souples des filtres.

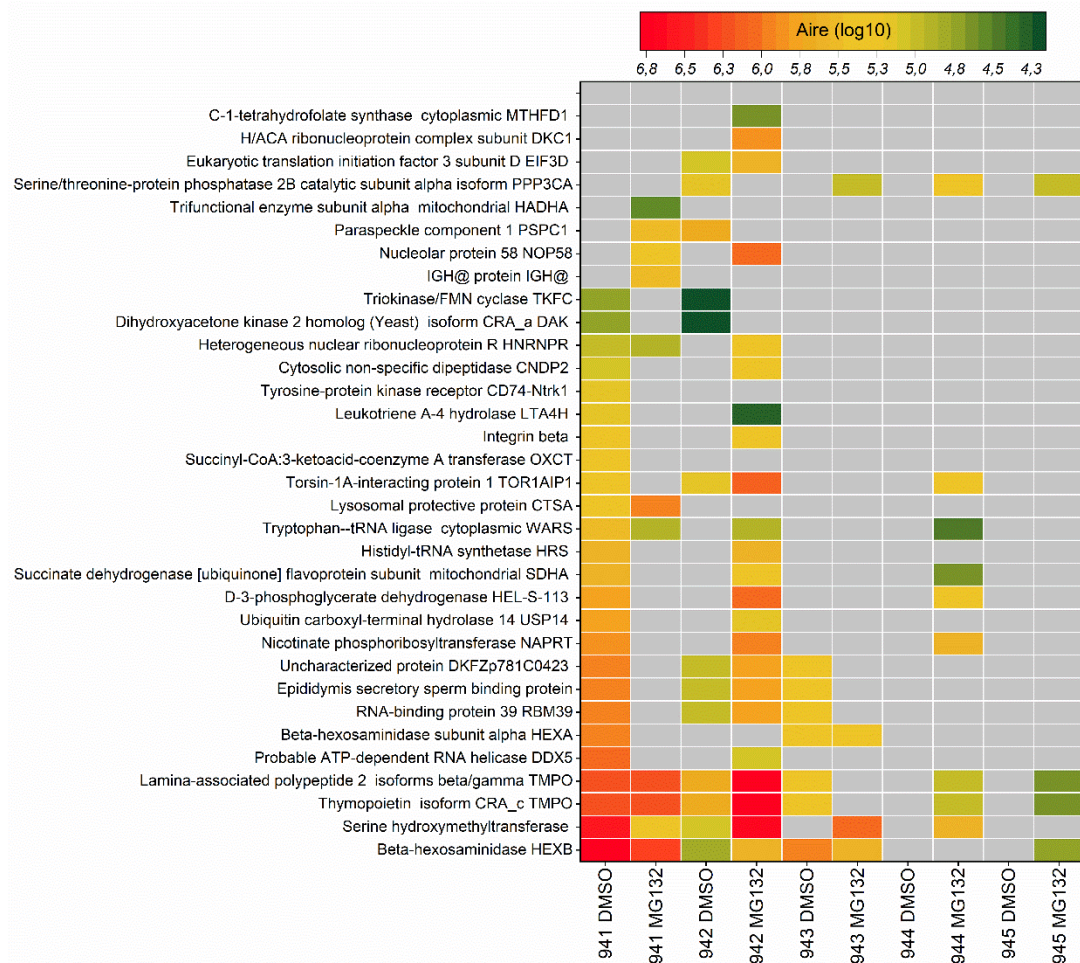


Figure 90 : Carte de chaleur des protéines identifiées selon l'analyse stricte. Les aires sous la courbe des protéines ont été mises au logarithme naturel pour un aspect visuel plus clair

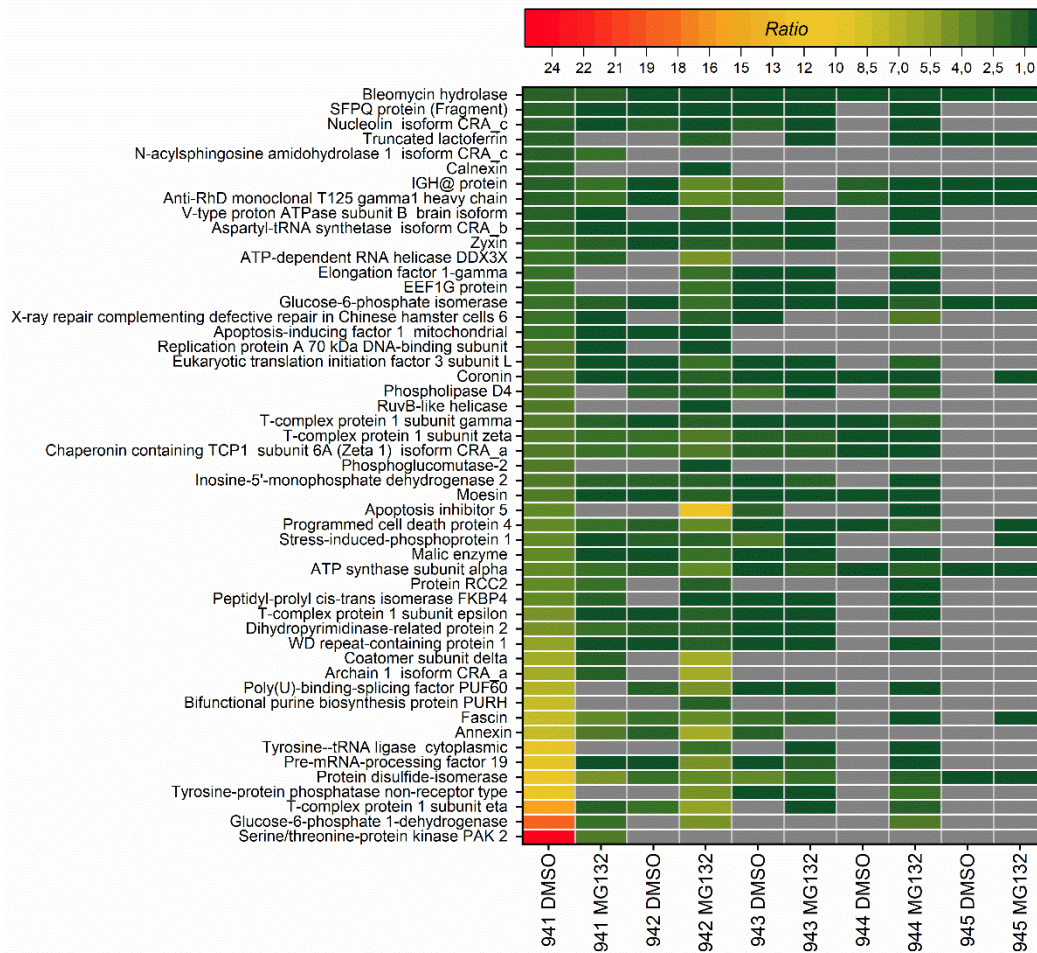


Figure 91 : Carte de chaleur des protéines identifiées selon l’analyse souple. Cette carte présente le ratio entre l’aire de protéines dans les sondes sur l’aire des protéines dans les contrôles

Si l’on fait l’analyse de la localisation générale des protéines observées avec les diverses sondes, on remarque que plusieurs protéines proviennent des vacuoles et lysosomes. Parmi celles-ci il y a les 2 sous-unités de l’hexosaminidase, la protéine protectrice lysosomale et la phosphoribosyltransférase nicotinate. Le fait qu’il y ait eu un enrichissement de ces protéines concorde avec des résultats expérimentaux préalablement observés par l’équipe qui démontrait que les sondes s’accumulaient dans les lysosomes lors d’essais de localisation par fluorophore. Plusieurs hypothèses peuvent expliquer l’accumulation des sondes dans les lysosomes. Tout d’abord, il existe un effet combinatoire qui augmente l’autophagie dans les cellules avec l’incubation des sondes. Il est connu qu’inhiber le protéosome par le MG132 a pour conséquence

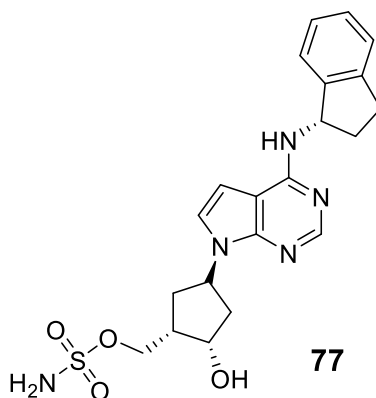
d'induire un stress à la cellule, ce qui résulte par l'augmentation de lysosomes.¹³¹ De plus, un papier connexe a démontré qu'une dégradation de LSD1 a aussi pour effet d'augmenter les effets de l'autophagie.¹³² La molécule UM171 dégrade le complexe RCOR/LSD1, ce qui, conséquemment, augmente la quantité de lysosomes cellulaires. Les lysosomes ont la particularité d'avoir un pH avoisinant 5 comme milieu interne.¹³⁰ Comme les sondes de UM171 ont pour la plupart une amine basique, elles seront piégées dans ces lysosomes en étant chargées et ne pourront plus faire de transport passif à travers la membrane et s'y accumuleront. Par conséquent, il est logique de voir la présence de protéines lysosomales dans l'enrichissement de la chimioprotéomique. Un second composant cellulaire commun aux sondes utilisées est les chaperones, cette machinerie responsable de comprimer l'ADN. On retrouve notamment des membres du complexe T, la protéine RuvBL1 et la protéine STIP1 tous agissant sur la transcription de l'ADN via une interaction avec les histones, comme le fait le complexe CoRest. Dans le même ordre d'idée, il est possible d'observer plusieurs membres du spliceosome (DDX5, PRPF19 et HNRNPR), le complexe agissant sur l'ARN messager pour permettre leur envoi du noyau vers le cytoplasme. Ce qui permet de conclure que les sondes ont réussi à aller dans le noyau cellulaire, compartiment où se situe le complexe RCOR/LSD1. Outre ces mécanismes cellulaires, on retrouve une multitude de protéines faisant partie de divers procédés de la cellule qui n'ont pas de sens avec les connaissances du projet.

Ainsi, il est très difficile de conclure quoi que ce soit avec toutes ces protéines identifiées. À notre grande déception, aucune protéine ne figure dans les interacteurs de la machinerie RCOR/LSD1, soit le complexe offrant la meilleure probabilité d'avoir le candidat qui interagit avec UM171 étant donné la dégradation de RCOR. Si l'on revient à l'identification des bandes spécifiques observées sur le gel d'argent, il est fort à parier qu'il s'agit des protéines Beta-hexosaminidase et de l'hélicase d'ARN DDX5, qui possèdent respectivement des masses de 63 et 69 kDa. Ces masses concordent avec le gel et ces protéines possèdent le plus de peptides identifiés renforçant la thèse de l'identification. Ces protéines sont non spécifiques puisqu'elles n'ont pas de liens avec quelconque interacteur du mécanisme RCOR/LSD1 pour lequel plusieurs études pointent du doigt des candidats possibles liés à ce signal.

3.4 Conclusion et perspective du projet

Ce présent chapitre mettait de l'avant le développement d'outils moléculaires du composé UM171 afin d'en identifier la cible thérapeutique selon l'approche d'enrichissement basée sur l'affinité du ligand. Le premier objectif était de facto la synthèse de sondes diazirines analogues au composé UM171 sans en sacrifier l'affinité. Bien qu'initialement l'objectif était de synthétiser des sondes où la diazirine se situait sur le benzyle et sur la chaîne aliphatique amine, il s'avère qu'une autre position a été exploitée par malchance expérimentale. L'utilisation d'une base lors de l'ajout du tosylate a fonctionnalisé l'azote pyrrolique, donnant ainsi lieu à deux sondes non planifiées. Néanmoins, dans l'ensemble, l'objectif de synthétiser diverses sondes exploitant l'espace chimique de UM171 a été réalisé et 6 sondes diazirines furent obtenues avec des activités variées allant d'inactives à très actives. Le second objectif était de vérifier les diverses étapes qui composent la méthode d'ABPP nécessaire pour l'identification de cible par protéomique. On dénote 4 grandes étapes pour cette méthode, soit l'incubation des sondes, l'activation de la diazirine, la fonctionnalisation de la sonde par une biotine et l'enrichissement des protéines marquées par les sondes. En se référant à la littérature, il a été possible de démontrer le bon fonctionnement de chaque petite étape de façon fragmentée. Par conséquent, l'objectif de vérifier la procédure pour en maîtriser ses rouages a été accompli. Néanmoins, une expérience clé est manquante pour ce mémoire. Il aurait été primordial d'effectuer la procédure de A à Z avec une sonde de la littérature pour y retrouver les mêmes interacteurs et par conséquent démontrer l'efficacité de la méthode. Néanmoins, la procédure de chimioprotéomique a été utilisée avec les diazirines analogues au composé UM171 pour sonder les cellules. Il a été plaisant de constater la présence de bandes spécifiques pour la sonde 941, qui permet de démontrer un enrichissement par la méthode choisie. Des protéines lysosomales et nucléaires ont pu être obtenues, faisant sens avec des observations antérieures, ce qui illustre le fonctionnement de la méthode. Cependant, les protéines qui furent identifiées lors de l'essai n'ont pas de pertinence avec le mécanisme du complexe CoREST, mécanisme responsable de l'effet biologique. Bien que les sondes 941 et 942 aient permis l'identification de protéines, il s'avère que les autres sondes ont enrichi un très faible nombre de protéines, laissant présager une méthode non optimale. Par conséquent, l'identification de la cible de UM171 n'a pas pu être réalisée et d'autres avenues devront être explorées.

Afin de donner des pistes de solutions pour la réalisation de l'identification de cibles, deux options sont envisageables : améliorer la méthode utilisée avec de nouveaux paramètres ou essayer une nouvelle méthode d'identification de cibles. Étant donné que la présence de MG132 favorise l'autophagie, il serait intéressant d'utiliser un inhibiteur différent qui ne produit pas cet effet. Comme le complexe RCOR/LSD1 se fait dégrader par la machinerie du protéosome, il est nécessaire que ces substrats subissent une modification d'ubiquitination. La protéine responsable d'ajouter la modification se nomme une ligase E3 et dans le corps humain, il en existe plus de 700.¹¹² Néanmoins, les dernières observations démontrent que la ligase Cullin3 est impliquée dans la dégradation du complexe CoREST. L'ajout de la modification NEDD8 est nécessaire pour venir activer cette ligase E3, responsable de l'ubiquitination sur le complexe. Ainsi, le groupe de recherche a démontré qu'un inhibiteur des ligases Cullin permet d'empêcher la Neddylation, donc d'empêcher la dégradation du complexe RCOR/LSD1.¹¹² L'inhibiteur MLN4924 (**77**) est par conséquent une excellente alternative au MG132 pour empêcher la dégradation des substrats par UM171, tout en limitant l'amplification des lysosomes dans la cellule.¹¹²



Un second aspect qui pourrait être envisagé pour limiter l'accumulation des sondes dans les lysosomes serait d'incuber les sondes sur une plus courte période. Il serait important de vérifier la cinétique complète de dégradation de RCOR avec les sondes d'UM171. Cette expérience clé aurait dû être réalisée pour chacune des sondes, car leur affinité est différente par rapport à UM171. Il se pourrait que 5 μ M, soit largement suffisant pour dégrader en moins de 4 heures RCOR. En se limitant à un temps d'incubation inférieur, on réduit certes la dégradation de RCOR, mais on permet de voir des interactions qui deviennent peut-être inexistantes avec un temps d'incubation trop long.

Une autre amélioration qu'il serait possible de faire concerne l'utilisation du couple alcyne/azoture. Bien que ces deux fonctions soient très petites, la réaction de cycloaddition CuAAC nécessite du cuivre. Il a été démontré que le cuivre en milieu cellulaire est toxique, pouvant influencer les résultats obtenus.¹³³ De plus, cette réaction nécessite 3 partenaires, soit le cuivre, l'azoture et l'alcyne. Dans un milieu dilué, ces partenaires ont moins de chances de réagir ensemble, ce qui peut limiter la cinétique et les rendements de réaction. Il existe une voie alternative dans la littérature qui exploite la réaction Diels-Alder avec une inversion de demande électronique (iEDDA). Cette réaction se produit entre la tétrazine et le transcyclooctène et permet de former un hétérocycle sans utilisation de cuivre.^{133, 134} Le mécanisme de cette réaction est présenté à la figure 92.

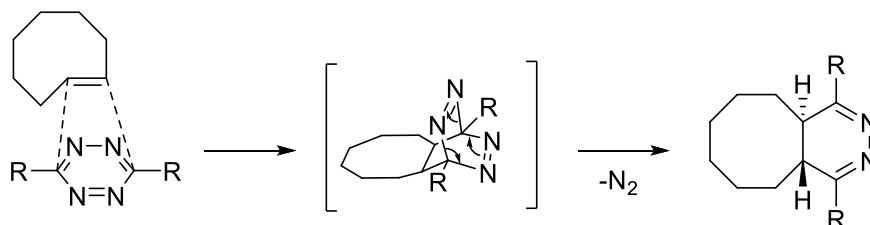


Figure 92 : Mécanisme réactionnel de la Diels-Alder avec une inversion de demande électronique entre la tétrazine et le transcyclooctène

La réaction iEDDA possède 2 avantages notables vis-à-vis la réaction de cycloaddition de l'azoture et l'alcyne. Tout d'abord, elle ne nécessite uniquement deux partenaires, aucun cuivre n'étant nécessaire pour permettre cette réaction. Le second avantage est en lien avec ce constat de la bimolécularité de la réaction : il s'avère que cette dernière est plus rapide de trois ordres de grandeur que la réaction avec le cuivre.¹³³ Ces avantages notables permettent de faire une réaction à basse température et ce pour une durée nettement inférieure en plus de réduire la toxicité dans les cellules. Avec le couple tétrazine/transcyclooctène, il serait ainsi possible de réduire les équivalents de la réaction et d'outrepasser la précipitation des protéines qui est nécessaire de faire pour ne pas saturer les billes de streptavidines avec la biotine libre. Comme dans toute procédure, être en mesure d'éliminer une étape permet de limiter les pertes

expérimentales et maximiser la quantité de matériels obtenus à la fin pour l'identification des protéines enrichies.

Outre l'optimisation de la méthode d'enrichissement ABPP, il existe d'autres avenues dans la littérature qui permettent l'identification de cibles et qui seraient des idées alternatives, notamment la technologie de la chimie du ligand dirigé développée par Itaru Hamachi et Tomori Tamura.⁵ Cette technologie repose sur l'ajout d'un groupement réactif possédant une étiquette d'identification qui est greffé au ligand de cible inconnue. L'idée est qu'une fois la molécule interagissant de manière non covalente avec la cible, des acides aminés nucléophiles à proximité viendront réagir sur la fonction dite réactive et occasionnera un marquage de la protéine d'intérêt avec une étiquette d'identification. Ces étiquettes sont les mêmes présentés en introduction, soit un fluorophore, une biotine ou un élément radioactif. L'objectif ensuite sera de venir enrichir ou visualiser les protéines marquées par le groupement réactif. La figure 93 schématise l'idée de cette technologie.

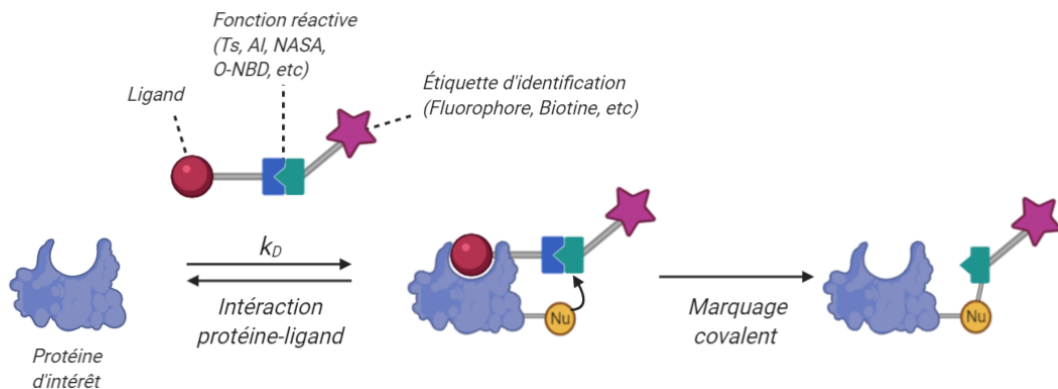


Figure 93 : Principe de la technologie du ligand dirigé développé par Pr Hamachi et Dr Tamura

Il est possible de choisir sélectivement le groupement réactif selon les acides aminés nucléophiles se trouvant à proximité du site actif. Cependant, dans une perspective d'identification, il est nécessaire d'en essayer plusieurs. Cette méthode est attrayante, car il est possible d'accumuler le nombre de protéines modifiées en laissant plus de temps d'incubation ou simplement en ajoutant 2 doses, ce qui n'est pas le cas avec les diazirines.

Finalement, une autre approche qui pourrait être intéressante est celle développée par le professeur Hiroshi Handa, qui met de l'avant la chromatographie d'affinité sur matrice solide.¹³⁵ Avec son groupe, il a développé une résine polymérique composée de styrène et de méthacrylate de glycidyle possédant en son centre des particules de ferrites, une espèce ferromagnétique. Ces billes, nommées FG, sont fonctionnalisées avec divers groupes au choix tels que des amines, des acides carboxyliques, des azotures, des tosyloxy ou des alcools. (Figure 94)

C'est grâce à ces groupements que la molécule d'intérêt sera fixée sur les billes. Ensuite, un lysat cellulaire est mélangé avec les billes et après divers lavages, les protéines qui interagissent avec les molécules seront éluées et analysées par la spectrométrie de masse. Ces billes possèdent des propriétés fort intéressantes compte tenu de leur petit diamètre et de leur caractère hydrophile. Il a été démontré que ces billes FG amènent très peu de protéines non spécifiques. C'est par l'utilisation de ce type de billes que le mécanisme biologique de la thalidomide a pu être identifié.¹³⁵ Il serait ainsi intéressant d'essayer cette nouvelle méthode, qui offre une approche d'affinité sur support solide.

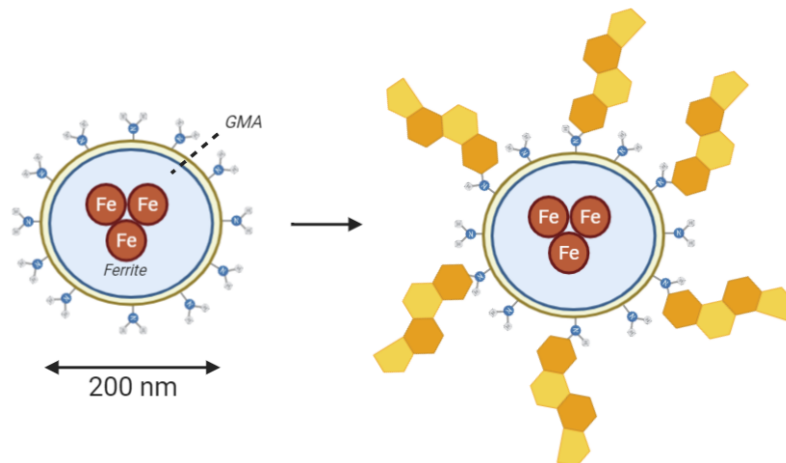


Figure 94 : Composition et fonctionnalisation des billes FG permettant d'enrichir les interacteurs d'une molécule d'intérêt

L'identification de cibles thérapeutiques qui proviennent d'un criblage phénotypique est une tâche très ardue. Il est essentiel de maximiser les approches pour se donner le plus de chances possibles de réussir. Bien que ce projet n'ait pas pu identifier la cible du composé UM171, il a pu développer une expertise dans l'approche de la chimioprotéomique.

4. Conclusion générale

Ce présent ouvrage discute de l'importance des outils moléculaires dans la recherche pharmaceutique. La synthèse des macrocycles covalents a permis de se lier à la cystéine 118 de RAS et d'obtenir des structures cristallines d'une importance capitale pour l'avancement du projet. Grâce à ces dernières, la grande différence d'affinité entre NS1 et les macrocycles non covalents a pu être élucidée par l'équipe et a permis d'indiquer la voie à privilégier pour développer des macrocycles plus affins. Au moment d'écrire ces dernières lignes, les chimistes ont réussi à obtenir près d'un ordre de grandeur en affinité des macrocycles en allant chercher des interactions supplémentaires avec RAS qui ont été identifiées avec les structures cristallines. La prochaine étape de ce projet consistera à établir la preuve biologique que les macrocycles non covalents sont en mesure de répliquer l'effet de NS1 en milieu cellulaire, pouvant ainsi mener à des molécules ayant un potentiel thérapeutique prometteur. Les perspectives d'une inhibition de RAS avec ces macrocycles sont extraordinaires pour le monde du traitement du cancer, cependant le parcours jusqu'au médicament est encore rempli d'embûches.

Le second projet a permis d'offrir des bases importantes pour construire une expertise d'identification de cibles à l'IRIC. Bien qu'il soit malheureux que la procédure n'ait pas réussi à identifier la cible thérapeutique de UM171, les résultats expérimentaux vérifiant chaque étape ont pu donner des informations critiques sur les limites de la méthode d'enrichissement de cibles basée sur l'affinité. Cette expertise est critique à l'institut puisque l'identification de cibles est d'actualité dans plusieurs projets internes de la plateforme de chimie médicinale. Il en reste beaucoup à apprendre sur ce domaine qui est d'une complexité inégalée, mais chaque petit pas est important. L'identification de cibles est essentielle pour optimiser l'approche thérapeutique et l'affinité des composés dans le développement des traitements pharmaceutiques.

5. Références

1. Goldman, N.; Bertone, P.; Chen, S.; Dessimoz, C.; LeProust, E. M.; Sipos, B.; Birney, E., Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature* **2013**, *494* (7435), 77-80.
2. Gonzalez, G., The Chemical Compositions of Stars with Planets: A Review. *Publications of the Astronomical Society of the Pacific* **2006**, *118* (849), 1494-1505.
3. Dey, M.; Weerapana, E., Chemical Tools in Biological Discovery. *Cell Chemical Biology* **2020**, *27* (8), 889-890.
4. Newman, R. H.; Zhang, J., Small molecules and chemical tools at the interface. *Nature Chemical Biology* **2008**, *4* (7), 382-386.
5. Shiraiwa, K.; Cheng, R.; Nonaka, H.; Tamura, T.; Hamachi, I., Chemical Tools for Endogenous Protein Labeling and Profiling. *Cell Chemical Biology* **2020**, *27* (8), 970-985.
6. Frémin, C.; Meloche, S., From basic research to clinical development of MEK1/2 inhibitors for cancer therapy. *Journal of Hematology & Oncology* **2010**, *3* (1), 8.
7. De Luca, A.; Maiello, M. R.; D'Alessio, A.; Pergameno, M.; Normanno, N., The RAS/RAF/MEK/ERK and the PI3K/AKT signalling pathways: role in cancer pathogenesis and implications for therapeutic approaches. *Expert Opinion on Therapeutic Targets* **2012**, *16* (sup2), S17-S27.
8. Butti, R.; Das, S.; Gunasekaran, V. P.; Yadav, A. S.; Kumar, D.; Kundu, G. C., Receptor tyrosine kinases (RTKs) in breast cancer: signaling, therapeutic implications and challenges. *Molecular Cancer* **2018**, *17* (1), 34.
9. Ariel, E.; Julia, M. S., RAS/Effector Interactions from Structural and Biophysical Perspective. *Mini-Reviews in Medicinal Chemistry* **2016**, *16* (5), 370-375.
10. Lake, D.; Corrêa, S. A. L.; Müller, J., Negative feedback regulation of the ERK1/2 MAPK pathway. *Cellular and Molecular Life Sciences* **2016**, *73* (23), 4397-4413.
11. Mattox, E. T.; Chen, X.; Maxuitenko, Y. Y.; Keeton, B. A.; Piazza, A. G., Exploiting RAS Nucleotide Cycling as a Strategy for Drugging RAS-Driven Cancers. *International Journal of Molecular Sciences* **2020**, *21* (1).
12. Degirmenci, U.; Wang, M.; Hu, J., Targeting Aberrant RAS/RAF/MEK/ERK Signaling for Cancer Therapy. *Cells* **2020**, *9* (1).
13. Liu, F.; Yang, X.; Geng, M.; Huang, M., Targeting ERK, an Achilles' Heel of the MAPK pathway, in cancer therapy. *Acta Pharmaceutica Sinica B* **2018**, *8* (4), 552-562.
14. Janes, M. R.; Zhang, J.; Li, L.-S.; Hansen, R.; Peters, U.; Guo, X.; Chen, Y.; Babbar, A.; Firdaus, S. J.; Darjania, L.; Feng, J.; Chen, J. H.; Li, S.; Li, S.; Long, Y. O.; Thach, C.; Liu, Y.; Zariéh, A.; Ely, T.; Kucharski, J. M.; Kessler, L. V.; Wu, T.; Yu, K.; Wang, Y.; Yao, Y.; Deng, X.; Zarrinkar, P. P.; Brehmer, D.; Dhanak, D.; Lorenzi, M. V.; Hu-Lowe, D.; Patricelli, M. P.; Ren, P.; Liu, Y., Targeting KRAS Mutant Cancers with a Covalent G12C-Specific Inhibitor. *Cell* **2018**, *172* (3), 578-589.e17.
15. Canon, J.; Rex, K.; Saiki, A. Y.; Mohr, C.; Cooke, K.; Bagal, D.; Gaida, K.; Holt, T.; Knutson, C. G.; Koppada, N.; Lanman, B. A.; Werner, J.; Rapaport, A. S.; San Miguel, T.; Ortiz, R.; Osgood, T.; Sun, J.-R.; Zhu, X.; McCarter, J. D.; Volak, L. P.; Houk, B. E.; Fakih, M. G.; O'Neil, B. H.; Price, T. J.; Falchook, G. S.; Desai, J.; Kuo, J.; Govindan, R.; Hong,

- D. S.; Ouyang, W.; Henary, H.; Arvedson, T.; Cee, V. J.; Lipford, J. R., The clinical KRAS(G12C) inhibitor AMG 510 drives anti-tumour immunity. *Nature* **2019**, *575* (7781), 217-223.
16. Hobbs, G. A.; Der, C. J.; Rossman, K. L., RAS isoforms and mutations in cancer at a glance. *Journal of Cell Science* **2016**, *129* (7), 1287.
 17. Bandaru, P.; Kondo, Y.; Kuriyan, J., The Interdependent Activation of Son-of-Sevenless and Ras. *Cold Spring Harbor Perspectives in Medicine* **2019**, *9* (2).
 18. Freeman, A. K.; Ritt, D. A.; Morrison, D. K., The importance of Raf dimerization in cell signaling. *Small GTPases* **2013**, *4* (3), 180-185.
 19. Lavoie, H.; Therrien, M., Regulation of RAF protein kinases in ERK signalling. *Nature Reviews Molecular Cell Biology* **2015**, *16* (5), 281-298.
 20. Simanshu, D. K.; Nissley, D. V.; McCormick, F., RAS Proteins and Their Regulators in Human Disease. *Cell* **2017**, *170* (1), 17-33.
 21. O'Bryan, J. P., Pharmacological targeting of RAS: Recent success with direct inhibitors. *Pharmacol Res* **2019**, *139*, 503-511.
 22. Spiegel, J.; Cromm, P. M.; Zimmermann, G.; Grossmann, T. N.; Waldmann, H., Small-molecule modulation of Ras signaling. *Nature Chemical Biology* **2014**, *10* (8), 613-622.
 23. Ostrem, J. M. L.; Shokat, K. M., Direct small-molecule inhibitors of KRAS: from structural insights to mechanism-based design. *Nature Reviews Drug Discovery* **2016**, *15* (11), 771-785.
 24. Roberts, P. J.; Der, C. J., Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene* **2007**, *26* (22), 3291-3310.
 25. Sheridan, C., Grail of RAS cancer drugs within reach. *Nature Biotechnology* **2020**, *38* (1), 6-8.
 26. Lanman, B. A.; Allen, J. R.; Allen, J. G.; Amegadzie, A. K.; Ashton, K. S.; Booker, S. K.; Chen, J. J.; Chen, N.; Frohn, M. J.; Goodman, G.; Kopecky, D. J.; Liu, L.; Lopez, P.; Low, J. D.; Ma, V.; Minatti, A. E.; Nguyen, T. T.; Nishimura, N.; Pickrell, A. J.; Reed, A. B.; Shin, Y.; Siegmund, A. C.; Tamayo, N. A.; Tegley, C. M.; Walton, M. C.; Wang, H.-L.; Wurz, R. P.; Xue, M.; Yang, K. C.; Achanta, P.; Bartberger, M. D.; Canon, J.; Hollis, L. S.; McCarter, J. D.; Mohr, C.; Rex, K.; Saiki, A. Y.; San Miguel, T.; Volak, L. P.; Wang, K. H.; Whittington, D. A.; Zech, S. G.; Lipford, J. R.; Cee, V. J., Discovery of a Covalent Inhibitor of KRASG12C (AMG 510) for the Treatment of Solid Tumors. *Journal of Medicinal Chemistry* **2020**, *63* (1), 52-65.
 27. Spencer-Smith, R.; Koide, A.; Zhou, Y.; Eguchi, R. R.; Sha, F.; Gajwani, P.; Santana, D.; Gupta, A.; Jacobs, M.; Herrero-Garcia, E.; Cobbert, J.; Lavoie, H.; Smith, M.; Rajakulendran, T.; Dowdell, E.; Okur, M. N.; Dementieva, I.; Sicheri, F.; Therrien, M.; Hancock, J. F.; Ikura, M.; Koide, S.; O'Bryan, J. P., Inhibition of RAS function through targeting an allosteric regulatory site. *Nature Chemical Biology* **2017**, *13* (1), 62-68.
 28. Koide, A.; Wojcik, J.; Gilbreth, R. N.; Hoey, R. J.; Koide, S., Teaching an Old Scaffold New Tricks: Monobodies Constructed Using Alternative Surfaces of the FN3 Scaffold. *Journal of Molecular Biology* **2012**, *415* (2), 393-405.
 29. Hantschel, O., Monobodies as possible next-generation protein therapeutics - a perspective. *Swiss Med Wkly* **2017**, *147*, w14545.

30. Becker, W.; Bhattiprolu, K. C.; Gubensäk, N.; Zangger, K., Investigating Protein–Ligand Interactions by Solution Nuclear Magnetic Resonance Spectroscopy. *ChemPhysChem* **2018**, *19* (8), 895-906.
31. Groß, A.; Hashimoto, C.; Sticht, H.; Eichler, J., Synthetic Peptides as Protein Mimics. *Frontiers in Bioengineering and Biotechnology* **2016**, *3*, 211.
32. Khakshoor, O.; Nowick, J. S., Artificial β -sheets: chemical models of β -sheets. *Current Opinion in Chemical Biology* **2008**, *12* (6), 722-729.
33. Reece, J. B.; Meyers, N.; Urry, L. A.; Cain, M. L.; Wasserman, S. A.; Minorsky, P. V.; Jackson, R. B.; Cooke, B. J., *Campbell biology* Pearson: Frenchs Forest, NSW, 2015.
34. Wisniewski, D.; Affer, M.; Willshire, J.; Clarkson, B., Further phenotypic characterization of the primitive lineage– CD34+CD38–CD90+CD45RA– hematopoietic stem cell/progenitor cell sub-population isolated from cord blood, mobilized peripheral blood and patients with chronic myelogenous leukemia. *Blood Cancer Journal* **2011**, *1* (9), e36-e36.
35. Sigma-Aldrich In Vitro Expansion of CD34+/CD38- Human Hematopoietic Progenitor Cells in a Serum-free and Xeno-free Stem Cell Culture Media. <https://www.sigmaaldrich.com/technical-documents/protocols/biology/cell-culture/hematopoietic-stem-cell-media.html#ref> (accessed 2020-11-16).
36. Ngom, M.; Imren, S.; Maetzig, T.; Adair, J. E.; Knapp, D. J. H. F.; Chagraoui, J.; Fares, I.; Bordeleau, M.-E.; Sauvageau, G.; Leboulch, P.; Eaves, C.; Humphries, R. K., UM171 Enhances Lentiviral Gene Transfer and Recovery of Primitive Human Hematopoietic Cells. *Molecular Therapy - Methods & Clinical Development* **2018**, *10*, 156-164.
37. Feng, Y.; Xie, X.-Y.; Yang, Y.-Q.; Sun, Y.-T.; Ma, W.-H.; Zhou, P.-J.; Li, Z.-Y.; Liu, H.-Q.; Wang, Y.-F.; Huang, Y.-S., Synthesis and evaluation of pyrimidoindole analogs in umbilical cord blood ex vivo expansion. *European Journal of Medicinal Chemistry* **2019**, *174*, 181-197.
38. Fares, I.; Chagraoui, J.; Gareau, Y.; Gingras, S.; Ruel, R.; Mayotte, N.; Csaszar, E.; Knapp, D. J. H. F.; Miller, P.; Ngom, M.; Imren, S.; Roy, D.-C.; Watts, K. L.; Kiem, H.-P.; Herrington, R.; Iscove, N. N.; Humphries, R. K.; Eaves, C. J.; Cohen, S.; Marinier, A.; Zandstra, P. W.; Sauvageau, G., Pyrimidoindole derivatives are agonists of human hematopoietic stem cell self-renewal. *Science* **2014**, *345* (6203), 1509.
39. Fares, I.; Rivest-Khan, L.; Cohen, S.; Sauvageau, G., Small molecule regulation of normal and leukemic stem cells. *Current Opinion in Hematology* **2015**, *22* (4).
40. Sidney, L. E.; Branch, M. J.; Dunphy, S. E.; Dua, H. S.; Hopkinson, A., Concise review: evidence for CD34 as a common marker for diverse progenitors. *Stem cells (Dayton, Ohio)* **2014**, *32* (6), 1380-9.
41. Boitano, A. E.; Wang, J.; Romeo, R.; Bouchez, L. C.; Parker, A. E.; Sutton, S. E.; Walker, J. R.; Flaveny, C. A.; Perdew, G. H.; Denison, M. S.; Schultz, P. G.; Cooke, M. P., Aryl Hydrocarbon Receptor Antagonists Promote the Expansion of Human Hematopoietic Stem Cells. *Science* **2010**, *329* (5997), 1345.
42. Sharma, S.; Cabana, R.; Shariatmadar, S.; Krishan, A., Cellular volume and marker expression in human peripheral blood apheresis stem cells. *Cytometry Part A* **2008**, *73A* (2), 160-167.

43. Chagraoui, J.; Lehnertz, B.; Girard, S.; Spinella, J. F.; Fares, I.; Tomellini, E.; Mayotte, N.; Corneau, S.; MacRae, T.; Simon, L.; Sauvageau, G., UM171 induces a homeostatic inflammatory-detoxification response supporting human HSC self-renewal. *PLOS ONE* **2019**, *14* (11), e0224900.
44. Wen, R.; Dong, C.; Xu, C.; Zhao, L.; Yang, Y.; Zhang, Z.; Chen, Y.; Duan, L.; Chen, H.; Yang, Z.; Zhang, B., UM171 promotes expansion of autologous peripheral blood hematopoietic stem cells from poorly mobilizing lymphoma patients. *International Immunopharmacology* **2020**, *81*, 106266.
45. Cohen, S.; Roy, J.; Lachance, S.; Delisle, J.-S.; Marinier, A.; Busque, L.; Roy, D.-C.; Barabé, F.; Ahmad, I.; Bambace, N.; Bernard, L.; Kiss, T.; Bouchard, P.; Caudrelier, P.; Landais, S.; Laroche, F.; Chagraoui, J.; Lehnertz, B.; Corneau, S.; Tomellini, E.; van Kampen, J. J. A.; Cornelissen, J. J.; Dumont-Lagacé, M.; Tanguay, M.; Li, Q.; Lemieux, S.; Zandstra, P. W.; Sauvageau, G., Hematopoietic stem cell transplantation using single UM171-expanded cord blood: a single-arm, phase 1–2 safety and feasibility study. *The Lancet Haematology* **2020**, *7* (2), e134-e145.
46. Comess, K. M.; McLoughlin, S. M.; Oyer, J. A.; Richardson, P. L.; Stöckmann, H.; Vasudevan, A.; Warder, S. E., Emerging Approaches for the Identification of Protein Targets of Small Molecules - A Practitioners' Perspective. *Journal of Medicinal Chemistry* **2018**, *61* (19), 8504-8535.
47. Wright, M. H.; Sieber, S. A., Chemical proteomics approaches for identifying the cellular targets of natural products. *Natural Product Reports* **2016**, *33* (5), 681-708.
48. Kolb, H. C.; Finn, M. G.; Sharpless, K. B., Click Chemistry: Diverse Chemical Function from a Few Good Reactions. *Angewandte Chemie International Edition* **2001**, *40* (11), 2004-2021.
49. Lehmann, J.; Wright, M. H.; Sieber, S. A., Making a Long Journey Short: Alkyne Functionalization of Natural Product Scaffolds. *Chemistry – A European Journal* **2016**, *22* (14), 4666-4678.
50. Chen, X.; Wang, Y.; Ma, N.; Tian, J.; Shao, Y.; Zhu, B.; Wong, Y. K.; Liang, Z.; Zou, C.; Wang, J., Target identification of natural medicine with chemical proteomics approach: probe synthesis, target fishing and protein identification. *Signal Transduction and Targeted Therapy* **2020**, *5* (1), 72.
51. Smith, E.; Collins, I., Photoaffinity labeling in target- and binding-site identification. *Future Medicinal Chemistry* **2015**, *7* (2), 159-183.
52. MacKinnon, A. L.; Taunton, J., Target Identification by Diazirine Photo-Cross-Linking and Click Chemistry. *Current Protocols in Chemical Biology* **2009**, *1* (1), 55-73.
53. Zhuang, S.; Li, Q.; Cai, L.; Wang, C.; Lei, X., Chemoproteomic Profiling of Bile Acid Interacting Proteins. *ACS Central Science* **2017**, *3* (5), 501-509.
54. Heydenreuter, W.; Kunold, E.; Sieber, S. A., Alkynol natural products target ALDH2 in cancer cells by irreversible binding to the active site. *Chemical Communications* **2015**, *51* (87), 15784-15787.
55. Jöst, C.; Nitsche, C.; Scholz, T.; Roux, L.; Klein, C. D., Promiscuity and Selectivity in Covalent Enzyme Inhibition: A Systematic Study of Electrophilic Fragments. *Journal of Medicinal Chemistry* **2014**, *57* (18), 7590-7599.

56. Lonsdale, R.; Burgess, J.; Colclough, N.; Davies, N. L.; Lenz, E. M.; Orton, A. L.; Ward, R. A., Expanding the Armory: Predicting and Tuning Covalent Warhead Reactivity. *Journal of Chemical Information and Modeling* **2017**, *57* (12), 3124-3137.
57. Flanagan, M. E.; Abramite, J. A.; Anderson, D. P.; Aulabaugh, A.; Dahal, U. P.; Gilbert, A. M.; Li, C.; Montgomery, J.; Oppenheimer, S. R.; Ryder, T.; Schuff, B. P.; Uccello, D. P.; Walker, G. S.; Wu, Y.; Brown, M. F.; Chen, J. M.; Hayward, M. M.; Noe, M. C.; Obach, R. S.; Philippe, L.; Shanmugasundaram, V.; Shapiro, M. J.; Starr, J.; Stroh, J.; Che, Y., Chemical and Computational Methods for the Characterization of Covalent Reactive Groups for the Prospective Design of Irreversible Inhibitors. *Journal of Medicinal Chemistry* **2014**, *57* (23), 10072-10079.
58. Smith, M. B., *March's Advanced Organic Chemistry*. Wiley: Newark, UNITED STATES, 2007.
59. Alabugin, I. V.; Gilmore, K.; Manoharan, M., Rules for Anionic and Radical Ring Closure of Alkynes. *Journal of the American Chemical Society* **2011**, *133* (32), 12608-12623.
60. Luna, F. O.; Gomez, J.; Cárdenas, C.; Albericio, F.; Marshall, H. S.; Guzmán, F., Deprotection Reagents in Fmoc Solid Phase Peptide Synthesis: Moving Away from Piperidine? *Molecules* **2016**, *21* (11).
61. Chhabra, S. R.; Hothi, B.; Evans, D. J.; White, P. D.; Bycroft, B. W.; Chan, W. C., An appraisal of new variants of Dde amine protecting group for solid phase peptide synthesis. *Tetrahedron Letters* **1998**, *39* (12), 1603-1606.
62. Smith, M. B., *March's Advanced Organic Chemistry: Reactions, Mechanisms, and Structure, 7th Edition*. 7th Edition ed.; John Wiley & Sons: Hoboken, New Jersey, 2013.
63. Ye, Y. H.; Li, H.; Jiang, X., DEPBT as an efficient coupling reagent for amide bond formation with remarkable resistance to racemization. *Biopolymers* **2005**, *80* (2-3), 172-8.
64. Santini, R.; Griffith, M. C.; Qi, M., A measure of solvent effects on swelling of resins for solid phase organic synthesis. *Tetrahedron Letters* **1998**, *39* (49), 8951-8954.
65. Stevens, J.; Hamstra, A.; Hegeman, T.; Sharrer, G., Factors and Parameters that Directly Affect Recovery of Collected Fractions in Preparative HPLC. *LCGC NORTH AMERICA* **2006**, *24* (2), 190-6.
66. Wlodawer, A.; Minor, W.; Dauter, Z.; Jaskolski, M., Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *The FEBS Journal* **2008**, *275* (1), 1-21.
67. O'Brien Ronan; Markova Natalia; A., H. G., Thermodynamics in Drug Discovery. In *Applied Biophysics for Drug Discovery*, Ltd., J. W. S., Ed. 2017; pp 7-28.
68. Reynolds, C. H.; Holloway, M. K., Thermodynamics of Ligand Binding and Efficiency. *ACS Medicinal Chemistry Letters* **2011**, *2* (6), 433-437.
69. Strelow, J. M., A Perspective on the Kinetics of Covalent and Irreversible Inhibition. *SLAS DISCOVERY: Advancing the Science of Drug Discovery* **2016**, *22* (1), 3-20.
70. Harris, C. M.; Foley, S. E.; Goedken, E. R.; Michalak, M.; Murdock, S.; Wilson, N. S., Merits and Pitfalls in the Characterization of Covalent Inhibitors of Bruton's Tyrosine Kinase. *SLAS DISCOVERY: Advancing the Science of Drug Discovery* **2018**, *23* (10), 1040-1050.

71. Pettinger, J.; Carter, M.; Jones, K.; Cheeseman, M. D., Kinetic Optimization of Lysine-Targeting Covalent Inhibitors of HSP72. *Journal of Medicinal Chemistry* **2019**, *62* (24), 11383-11398.
72. Schwartz, P. A.; Kuzmic, P.; Solowiej, J.; Bergqvist, S.; Bolanos, B.; Almaden, C.; Nagata, A.; Ryan, K.; Feng, J.; Dalvie, D.; Kath, J. C.; Xu, M.; Wani, R.; Murray, B. W., Covalent EGFR inhibitor analysis reveals importance of reversible interactions to potency and mechanisms of drug resistance. *Proceedings of the National Academy of Sciences* **2014**, *111* (1), 173.
73. Romney, D. K.; Murciano-Calles, J.; Wehrmüller, J. E.; Arnold, F. H., Unlocking Reactivity of TrpB: A General Biocatalytic Platform for Synthesis of Tryptophan Analogues. *Journal of the American Chemical Society* **2017**, *139* (31), 10769-10776.
74. Wang, J.; Kubicki, J.; Peng, H.; Platz, M. S., Influence of Solvent on Carbene Intersystem Crossing Rates. *Journal of the American Chemical Society* **2008**, *130* (20), 6604-6609.
75. Hill, J. R.; Robertson, A. A. B., Fishing for Drug Targets: A Focus on Diazirine Photoaffinity Probe Synthesis. *Journal of Medicinal Chemistry* **2018**, *61* (16), 6945-6963.
76. Modarelli, D. A.; Morgan, S.; Platz, M. S., Carbene formation, hydrogen migration, and fluorescence in the excited states of dialkyldiazirines. *Journal of the American Chemical Society* **1992**, *114* (18), 7034-7041.
77. Brunner, J.; Senn, H.; Richards, F. M., 3-Trifluoromethyl-3-phenyldiazirine. A new carbene generating group for photolabeling reagents. *The Journal of biological chemistry* **1980**, *255* (8), 3313-8.
78. Horne, J. E.; Walko, M.; Calabrese, A. N.; Levenstein, M. A.; Brockwell, D. J.; Kapur, N.; Wilson, A. J.; Radford, S. E., Rapid Mapping of Protein Interactions Using Tag-Transfer Photocrosslinkers. *Angewandte Chemie International Edition* **2018**, *57* (51), 16688-16692.
79. Kambe, T.; Correia, B. E.; Niphakis, M. J.; Cravatt, B. F., Mapping the Protein Interaction Landscape for Fully Functionalized Small-Molecule Probes in Human Cells. *Journal of the American Chemical Society* **2014**, *136* (30), 10777-10782.
80. He, D.; Xie, X.; Yang, F.; Zhang, H.; Su, H.; Ge, Y.; Song, H.; Chen, P. R., Quantitative and Comparative Profiling of Protease Substrates through a Genetically Encoded Multifunctional Photocrosslinker. *Angewandte Chemie International Edition* **2017**, *56* (46), 14521-14525.
81. Sekera, V. C.; Marvel, C. S., Higher Alkyl Sulfonates. *Journal of the American Chemical Society* **1933**, *55* (1), 345-349.
82. Kabalka, G. W.; Varma, M.; Varma, R. S.; Srivastava, P. C.; Knapp, F. F., The tosylation of alcohols. *The Journal of Organic Chemistry* **1986**, *51* (12), 2386-2388.
83. Ahad, A. M.; Jensen, S. M.; Jewett, J. C., A Traceless Staudinger Reagent To Deliver Diazirines. *Organic Letters* **2013**, *15* (19), 5060-5063.
84. Chouikhi, D.; Barluenga, S.; Winssinger, N., Clickable peptide nucleic acids (cPNA) with tunable affinity. *Chemical Communications* **2010**, *46* (30), 5476-5478.
85. Sauvageau, G. G., Yves; Ruel, Réjean; Gingras, Stéphane; Fares, Iman PYRIMIDO[4,5-b]INDOLE DERIVATIVES AND USE THEREOF IN THE EXPANSION OF HEMATOPOIETIC STEM CELLS. WO2013110198A1, 2012.

86. Perez, C.; Li, J.; Parlati, F.; Rouffet, M.; Ma, Y.; Mackinnon, A. L.; Chou, T.-F.; Deshaies, R. J.; Cohen, S. M., Discovery of an Inhibitor of the Proteasome Subunit Rpn11. *Journal of Medicinal Chemistry* **2017**, *60* (4), 1343-1361.
87. Anthony J. Pearson, W. R. R., *Activating Agents and Protecting Groups*. Wiley: 1999; p 528.
88. Shi, W.; Nacev, B. A.; Aftab, B. T.; Head, S.; Rudin, C. M.; Liu, J. O., Itraconazole Side Chain Analogues: Structure–Activity Relationship Studies for Inhibition of Endothelial Cell Proliferation, Vascular Endothelial Growth Factor Receptor 2 (VEGFR2) Glycosylation, and Hedgehog Signaling. *Journal of Medicinal Chemistry* **2011**, *54* (20), 7363-7374.
89. Mossine, A. V.; Brooks, A. F.; Henderson, B. D.; Hockley, B. G.; Frey, K. A.; Scott, P. J. H., An updated radiosynthesis of [18F]AV1451 for tau PET imaging. *EJNMMI Radiopharmacy and Chemistry* **2017**, *2* (1), 7.
90. Fischer, E.; Speier, A., Darstellung der Ester. *Berichte der deutschen chemischen Gesellschaft* **1895**, *28* (3), 3252-3258.
91. Kurimoto, A.; Hashimoto, K.; Nakamura, T.; Norimura, K.; Ogita, H.; Takaku, H.; Bonnert, R.; McInally, T.; Wada, H.; Isobe, Y., Synthesis and Biological Evaluation of 8-Oxoadenine Derivatives as Toll-like Receptor 7 Agonists Introducing the Antedrug Concept. *Journal of Medicinal Chemistry* **2010**, *53* (7), 2964-2972.
92. Wang, Z., *Comprehensive Organic Name Reactions and Reagents*. John Wiley & Sons, Inc: 2010.
93. G. Godfrey, A.; Ganem, B., Ready oxidation of halides to aldehydes using trimethylamine N-oxide in dimethylsulfoxide. *Tetrahedron Letters* **1990**, *31* (34), 4825-4826.
94. Griffith, W. P.; Jolliffe, J. M.; Ley, S. V.; Springhorn, K. F.; Tiffin, P. D., Oxidation of Activated Halides to Aldehydes and Ketones by N-Methylmorpholine-N-oxide. *Synthetic Communications* **1992**, *22* (13), 1967-1971.
95. Prakash, G. K. S.; Yudin, A. K., Perfluoroalkylation with Organosilicon Reagents. *Chemical Reviews* **1997**, *97* (3), 757-786.
96. Dong, C.; Bai, X.-F.; Lv, J.-Y.; Cui, Y.-M.; Cao, J.; Zheng, Z.-J.; Xu, L.-W., Cs₂CO₃-Initiated Trifluoro-Methylation of Chalcones and Ketones for Practical Synthesis of Trifluoromethylated Tertiary Silyl Ethers. *Molecules (Basel, Switzerland)* **2017**, *22* (5), 769.
97. Greene, P. G. M. W. T. W., *Greene's Protective Groups in Organic Synthesis, Fourth Edition*. 4th edition ed.; 2006.
98. Xing, R.; Tian, Q.; Liu, Q.; Li, Y., Synthesis of 2-Aminoindole Derivatives with Hantzsch Ester Catalyzed by Pd/C. *Chinese Journal of Chemistry* **2013**, *31* (2), 263-266.
99. Melanie Schultz, K. S., Wolfgang Staehle Autotaxin inhibitors. 2010.
100. Bachmann, W. E.; Hoffman, R. A., The Preparation of Unsymmetrical Biaryls by the Diazo Reaction and the Nitrosoacetylamine Reaction. In *Organic Reactions*, 2011; pp 224-261.
101. Linderman, R. J.; Graves, D. M., Oxidation of fluoroalkyl-substituted carbinols by the Dess-Martin reagent. *The Journal of Organic Chemistry* **1989**, *54* (3), 661-668.
102. Dess, D. B.; Martin, J. C., Readily accessible 12-I-5 oxidant for the conversion of primary and secondary alcohols to aldehydes and ketones. *The Journal of Organic Chemistry* **1983**, *48* (22), 4155-4156.

103. Nassal, M., 4-(1-Azi-2,2,2-trifluoroethyl)benzoic Acid, a Highly Photolabile Carbene Generating Label Readily Fixable to Biochemical Agents. *Liebigs Annalen der Chemie* **1983**, 1983 (9), 1510-1523.
104. Dubinsky, L.; Krom, B. P.; Meijler, M. M., Diazirine based photoaffinity labeling. *Bioorganic & Medicinal Chemistry* **2012**, 20 (2), 554-570.
105. Church, R. F. R.; Weiss, M. J., Diazirines. II. Synthesis and properties of small functionalized diazirine molecules. Observations on the reaction of a diaziridine with the iodine-iodide ion system. *The Journal of Organic Chemistry* **1970**, 35 (8), 2465-2471.
106. Burkard, N.; Bender, T.; Westmeier, J.; Nardmann, C.; Huss, M.; Wiczorek, H.; Grond, S.; von Zezschwitz, P., New Fluorous Photoaffinity Labels (F-PAL) and Their Application in V-ATPase Inhibition Studies. *European Journal of Organic Chemistry* **2010**, 2010 (11), 2176-2181.
107. Lundt, B. F.; Johansen, N. L.; Vølund, A.; Markussen, J., REMOVAL OF t-BUTYL AND t-BUTOXYCARBONYL PROTECTING GROUPS WITH TRIFLUOROACETIC ACID. *International Journal of Peptide and Protein Research* **1978**, 12 (5), 258-268.
108. Corelli, F.; Botta, M.; Lossani, A.; Pasquini, S.; Spadari, S.; Focher, F., Microwave-assisted synthesis and biological evaluation of novel uracil derivatives inhibiting human thymidine phosphorylase. *Il Farmaco* **2004**, 59 (12), 987-992.
109. Perreault, M. *Déconvolution de la cible de UM171 par approche chémoprotéomique*; Université de Montréal: Montréal, 2019; p 62.
110. Subramaniam, A.; Žemaitis, K.; Talkhoncheh, M. S.; Yudovich, D.; Bäckström, A.; Debnath, S.; Chen, J.; Jain, M. V.; Galeev, R.; Gaetani, M.; Zubarev, R. A.; Larsson, J., Lysine-specific demethylase 1A restricts ex vivo propagation of human HSCs and is a target of UM171. *Blood* **2020**, 136 (19), 2151-2161.
111. Magliulo, D.; Bernardi, R.; Messina, S., Lysine-Specific Demethylase 1A as a Promising Target in Acute Myeloid Leukemia. **2018**, 8 (255).
112. Chagraoui, J.; Girard, S.; Spinella, J.-F.; Simon, L.; Bonneil, E.; Mayotte, N.; MacRae, T.; Coulombe-Huntington, J.; Bertomeu, T.; Moison, C.; Tomellini, E.; Thibault, P.; Tyers, M.; Marinier, A.; Sauvageau, G., UM171 Preserves Epigenetic Marks that Are Reduced in Ex Vivo Culture of Human HSCs via Potentiation of the CLR3-KBTBD4 Complex. *Cell Stem Cell* **2021**, 28 (1), 48-62.e6.
113. Mackinnon, A. L.; Taunton, J., Target Identification by Diazirine Photo-Cross-linking and Click Chemistry. *Curr Protoc Chem Biol* **2009**, 1, 55-73.
114. A. Fleming, S., Chemical reagents in photoaffinity labeling. *Tetrahedron* **1995**, 51 (46), 12479-12520.
115. Jena, A. UVP High Intensity UV Lamps <https://www.uvp.com/products/ultraviolet-products/uv-handlamps/uvp-high-intensity-uv-lamps/> (accessed 18-11-2020).
116. Ziemianowicz, D. S.; Bomgarden, R.; Etienne, C.; Schriemer, D. C., Amino Acid Insertion Frequencies Arising from Photoproducts Generated Using Aliphatic Diazirines. *J Am Soc Mass Spectrom* **2017**, 28 (10), 2011-2021.
117. Szabó, D.; Schlosser, G.; Vékey, K.; Drahos, L.; Révész, Á., Collision energies on QToF and Orbitrap instruments: How to make proteomics measurements comparable? *Journal of Mass Spectrometry* **2021**, 56 (1), e4693.

118. Diedrich, J. K.; Pinto, A. F. M.; Yates, J. R., 3rd, Energy dependence of HCD on peptide fragmentation: stepped collisional energy finds the sweet spot. *J Am Soc Mass Spectrom* **2013**, *24* (11), 1690-1699.
119. Liang, L.; Astruc, D., The copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) "click" reaction and its applications. An overview. *Coordination Chemistry Reviews* **2011**, *255* (23), 2933-2945.
120. Shehadul Islam, M.; Aryasomayajula, A.; Selvaganapathy, P. R., A Review on Macroscale and Microscale Cell Lysis Methods. *Micromachines (Basel)* **2017**, *8* (3), 83.
121. Macritchie, F., Effects of temperature on dissolution and precipitation of proteins and polyamino acids. *Journal of Colloid and Interface Science* **1973**, *45* (2), 235-241.
122. Yang, Y.; Yang, X.; Verhelst, S. H. L., Comparative analysis of click chemistry mediated activity-based protein profiling in cell lysates. *Molecules (Basel, Switzerland)* **2013**, *18* (10), 12599-12608.
123. Shieh, P.; Dien, V. T.; Beahm, B. J.; Castellano, J. M.; Wyss-Coray, T.; Bertozzi, C. R., CalFluors: A Universal Motif for Fluorogenic Azide Probes across the Visible Spectrum. *Journal of the American Chemical Society* **2015**, *137* (22), 7145-7151.
124. Yang, N. J.; Hinner, M. J., Getting across the cell membrane: an overview for small molecules, peptides, and proteins. *Methods Mol Biol* **2015**, *1266*, 29-53.
125. Szychowski, J.; Mahdavi, A.; Hodas, J. J.; Bagert, J. D.; Ngo, J. T.; Landgraf, P.; Dieterich, D. C.; Schuman, E. M.; Tirrell, D. A., Cleavable biotin probes for labeling of biomolecules via azide-alkyne cycloaddition. *J Am Chem Soc* **2010**, *132* (51), 18351-60.
126. Yang, Y.; Fonović, M.; Verhelst, S. H., Cleavable Linkers in Chemical Proteomics Applications. *Methods Mol Biol* **2017**, *1491*, 185-203.
127. Goldman, A.; Harper, S.; Speicher, D. W., Detection of Proteins on Blot Membranes. *Curr Protoc Protein Sci* **2016**, *86*, 10.8.1-10.8.11.
128. Fic, E.; Kedracka-Krok, S.; Jankowska, U.; Pirog, A.; Dziedzicka-Wasylewska, M., Comparison of protein precipitation methods for various rat brain structures prior to proteomic analysis. *ELECTROPHORESIS* **2010**, *31* (21), 3573-3579.
129. Atkins, P. W., *Chimie Physique*. 4 ed.; 2013; p 1024.
130. Cooper, G. M., *The Cell*. 2 ed.; 2000.
131. Bao, W.; Gu, Y.; Ta, L.; Wang, K.; Xu, Z., Induction of autophagy by the MG-132 proteasome inhibitor is associated with endoplasmic reticulum stress in MCF-7 cells. *Mol Med Rep* **2016**, *13* (1), 796-804.
132. Ambrosio, S.; Saccà, C. D.; Amente, S.; Paladino, S.; Lania, L.; Majello, B., Lysine-specific demethylase LSD1 regulates autophagy in neuroblastoma through SESN2-dependent pathway. *Oncogene* **2017**, *36* (48), 6701-6711.
133. Kim, E.; Koo, H., Biomedical applications of copper-free click chemistry: in vitro, in vivo, and ex vivo. *Chemical Science* **2019**, *10* (34), 7835-7851.
134. Fang, L.; Chakraborty, S.; Dieter, E. M.; Potter, Z. E.; Lombard, C. K.; Maly, D. J., Chemoproteomic Method for Profiling Inhibitor-Bound Kinase Complexes. *Journal of the American Chemical Society* **2019**, *141* (30), 11912-11922.
135. Sakamoto, S.; Hatakeyama, M.; Ito, T.; Handa, H., Tools and methodologies capable of isolating and identifying a target molecule for a bioactive compound. *Bioorganic & Medicinal Chemistry* **2012**, *20* (6), 1990-2001.

Remarques générales

La structure des molécules, leur nomenclature, leur formule moléculaire (**Formula**) ainsi que leur masse molaire (**Mw**) ont été générées par le logiciel ChemBioDraw Ultra (version 12.0). Les réactions qui ont nécessitées une atmosphère inerte ou des conditions anhydres ont été effectuées sous un flux constant d'azote dans une verrerie sèche. La verrerie a soit été flambée avec un bruleur Bunsen ou séchée préalablement dans une étuve à 200 °C la nuit précédente. Les solvants ont été déshydratés avec l'utilisation de tamis moléculaire 4 Å durant un minimum de 16 heures. Tel qu'il est indiqué par le *Guide de présentation des mémoires et des thèses de la Faculté des études supérieures et postdoctorales de l'Université de Montréal*, une écriture de la partie expérimentale a été effectuée en anglais pour la technicité des termes. Cette partie expérimentale est présentée dans l'Annexe A.

Instrumentation

Le suivi des réactions a été effectué soit par chromatographie sur couche mince ou par LCMS. Les CCM sont des plaques de verre recouvertes d'un gel de silice 60 F₂₄, de la compagnie Merck. Une fois l'élution complétée, les plaques ont été révélées par divers moyens tels qu'un rayonnement UV ou des révélateurs chimiques (solution de KMNO₄, solution de ninhydrine, solution du vert de bromocresol). Les analyses par LCMS ont été effectuées à l'aide d'un système d'Agilent de la série 1260 Infinity II. Le gradient utilisé pour la séparation est séparation était de 0 à 100 % de l'éluant B en 2 minutes, où l'éluant A est 95% H₂O, 5% MeOH et 0.05 % HCOOH et B est 5% H₂O, 95% MeOH et 0.05 % HCOOH. La colonne est une C18 de la compagnie Kinetex avec une dimension de 3.0 mm x 30 mm et des pores de 2.6 µm.

La purification des composés par chromatographie flash sur phase solide a été possible grâce au système automatisé des appareils ISCO CombiFlash RF200 de Teledyne. Les colonnes de 4g, 12g, 20g et 40 g provenaient de la compagnie Canadian Life Science et la séparation des composés a été effectuée avec des solvants spécifiques pour chaque espèce.

La purification effectuée par HPLC-MS a été possible grâce au système d'Agilent 1260 série Infinity, qui couple un système HPLC à un quadrupole. La colonne utilisée pour purifier les composés est un C8 de la compagnie Kinetex possédant les dimensions 21.2 x 50 mm avec des pores de 5 µm. Les gradients et les éluants sont propres pour chaque composé purifié.

La lyophilisation des composés finaux a été effectuée en mélangeant un volume équivalent d'acétonitrile et d'eau jusqu'à solubilisation. Le liquide a été gelé avec de la glace sèche et placée sous vide avec Sentry 2.0 (Virtis) durant au moins 16 heures.

Les analyses spectrales en RMN que ce soit pour le proton (**RMN ¹H**) ou le carbone (**RMN ¹³C**) ont été enregistrés à l'aide d'un appareil de force 400 MHz (Varian) Les solvants deutérés utilisés sont le CDCl₃ et le DMSO-d₆ et proviennent de Sigma-Aldrich. Les spectres RMN ont été analysés à l'aide du logiciel *ACD/NMR processor academic edition*. Les déplacements chimiques des signaux ont été déterminés par rapport au solvant deutéré de référence et sont rapportés en partie par million (δ , ppm). Le logiciel a permis d'obtenir les constantes de couplages entre les mêmes signaux en valeur de hertz (J, Hz) et une valeur d'intensité relative dans le cas des spectres du proton. La multiplicité des signaux a été rapportée selon les termes suivants : s (singulet), d (doublet), t (triplet), q (quadruplet), br. s (singulet large), dt (doublet de triplet), td (triplet de doublet) et m (multiplet).

L'analyse des pouvoirs rotatoires (**$[\alpha]^{25}_D$**) a été effectuée à l'aide d'un polarimètre thermostaté de type Anton Paar MCP 200 de la compagnie Perkins Elmer. Les mesures ont été prises à l'aide d'une cuvette de quartz de 1 mL possédant un parcours de 10 cm. La longueur d'onde de l'analyse était de 589 nm qui est la raie D du sodium. La concentration est donnée en (g/100 mL) et la température T en °C. Les pouvoirs rotatoires sont rapportés selon la séquence suivante : Pouvoir rotatoire (concentration, solvant de mesure).

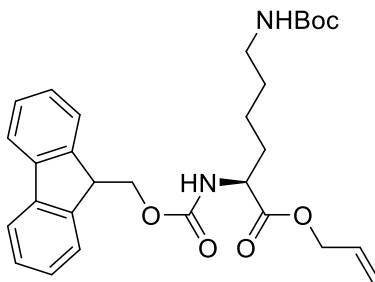
Les analyses de masses exactes (**HRMS**) des composés ont été enregistrées sur un MSD TOF (Agilent) par électronebulisation électrostatique (ESI). Le système de séparation couplé au TOF est un LC 1100 serie d'Agilent. Les données sont décrites selon la différence entre la masse exacte théorique et la masse exacte expérimentale en m/z. Les analyses de chromatographies liquides haute performance (**HPLC**) ont été effectuées sur un appareil Agilent de la série 1200. Le gradient appliqué pour la séparation était de 0 à 100 % de l'éluant B en 2 minutes avec un débit de 3 mL/minutes., où l'éluant A est 95% H₂O, 5% MeOH et 0.5 % TFA et B est 5% H₂O, 95% MeOH et 0.5 % TFA. La colonne utilisée est un C18 de la compagnie Kinetex avec une dimension de 4.6 mm x 30 mm et des pores de 2.6 μ m. Les chromatogrammes sont interprétés par le logiciel Chemstation d'Agilent. Les données sont rapportées selon la nomenclature suivante : type de colonne, gradient utilisé pour la séparation, débit d'élution et le temps de rétention (t_r).

6. Annexe A : Partie expérimentale et caractérisation

A.1 Organic synthesis section for the first project: Development of covalent macrocycles to mimic the NS1 monobody

Synthesis of building blocks

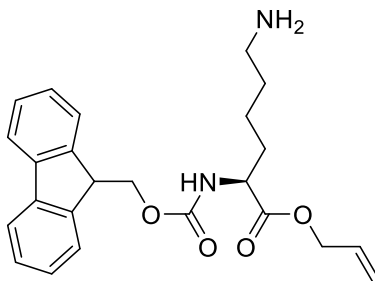
(S)-allyl N2-(((9H-fluoren-9-yl)methoxy)carbonyl)-N6-(tert-butoxycarbonyl)-L-lysinate (19a)



To DMF (20 ml) was firstly added Fmoc-L-Lys(boc)-OH (6.56 g, 14.00 mmol) and secondly K_2CO_3 (1.935 g, 14.00 mmol). It was a pale-yellow solution. After waiting 2 minutes, the neat Allyl bromide (1.212 ml, 14.00 mmol) was added to the solution and the mixture was stirred for 2.5 hours at room temperature. The mixture was diluted in EtOAc (130 ml) and washed with water (3 x 30 ml) and brine (25 ml). The organic layer was dried with Na_2SO_4 , filtered and concentrated under reduced pressure. The white residue was purified by flash chromatography (SiO_2 , 10 % EtOAc/Hex for 6 minutes and 10/70% EtOAc/Hex in 20 minutes.) using a Combiflash Rf 200 (Isco) instrument with a 80 g column (crude material was applied via a SiO_2 20 g). Fractions were combined and concentrated under reduced pressure, and co-evaporated several time with DCM to give allyl N2-(((9H-fluoren-9-yl)methoxy)carbonyl)-N6-(tert-butoxycarbonyl)-L-lysinate (6.797 g, 13.36 mmol, 95 % yield) as a white solid.

Formula: $C_{29}H_{36}N_2O_6$; **MW:** 508,62 g/mol; **1H NMR** ($CDCl_3$, 400 MHz): δ = 7.77 (d, $J=7.8$ Hz, 2 H), 7.61 (d, $J=7.0$ Hz, 2 H), 7.41 (t, $J=7.4$ Hz, 2 H), 7.32 (t, $J=7.4$ Hz, 2 H), 5.85 - 5.98 (m, 1 H), 5.40 - 5.52 (m, 1 H), 5.34 (d, $J=17.2$ Hz, 1 H), 5.27 (dd, $J=10.6, 0.8$ Hz, 1 H), 4.66 (d, $J=5.5$ Hz, 2 H), 4.60 (br. s., 1 H), 4.34 - 4.47 (m, 3 H), 4.23 (t, $J=6.8$ Hz, 1 H), 3.03 - 3.21 (m, 2 H), 1.83 - 1.95 (m, 1 H), 1.67 - 1.77 (m, 1 H), 1.48 - 1.57 (m, 2 H), 1.45 (s, 9 H), 1.28 - 1.41 ppm (m, 2 H); **^{13}C NMR** ($CDCl_3$, 101 MHz): δ = 172.1, 156.0, 155.9, 143.9, 143.7, 141.2, 141.2, 131.4, 127.6, 127.0, 125.1, 119.9, 119.9, 119.0, 79.1, 67.0, 65.9, 53.7, 47.1, 40.0, 32.1, 29.5, 28.4, 22.3 ppm; **HRMS** calcd for $C_{29}H_{36}N_2O_6Na$ $[M+Na]^+$: 531.2471; found 531.2342 (-0.0129 m/z). **$[\alpha]_D^{25}$** -12.2 (c 1.8 mg/mL, EtOH)

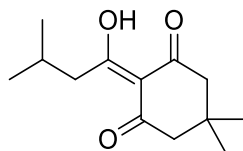
(S)-allyl 2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-6-aminohexanoate (19)



THF(BHT-free) (10 ml) was prepared by distillation of BHT-containing THF on the rotavap, followed by drying on 4A MS. After allyl N2-(((9H-fluoren-9-yl)methoxy)carbonyl)-N6-(tert-butoxycarbonyl)-L-lysinate 13 (2.297 g, 4.502 mmol) and HCl 4 M in dioxane (12.99 ml, 51.9 mmol) were added to THF than stirred at room temperature for 3 hours. The volatiles were evaporated under reduced pressure and co-evaporated 2 times with dioxane. Finally, the (S)-allyl 2-(((9H-fluoren-9-yl)methoxy)carbonylamino)-6-aminohexanoate 14 was pumping for 2 hours and a yellow/white foam was obtained (1.93 g, 96 %).

Formula: C₂₄H₂₈N₂O₄; **MW:** 408,50 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 7.90 (d, J=7.4 Hz, 2 H), 7.92 (br. s, 2 H), 7.81 (d, J=7.8 Hz, 1 H), 7.71 (dd, J=7.4, 3.9 Hz, 2 H), 7.42 (t, J=7.4 Hz, 2 H), 7.33 (t, J=7.4 Hz, 2 H), 5.89 (ddt, J=17.3, 10.5, 5.3 Hz, 1 H), 5.30 (dq, J=17.3, 1.7 Hz, 1 H), 5.20 (dq, J=10.6, 1.6 Hz, 1 H), 4.58 (d, J=5.1 Hz, 2 H), 4.26 - 4.38 (m, 2 H), 4.23 (t, J=7.0 Hz, 1 H), 3.98 - 4.08 (m, 1 H), 2.68 - 2.80 (m, 2 H), 1.63 - 1.79 (m, 2 H), 1.50 - 1.60 (m, 2 H), 1.31 - 1.43 ppm (m, 2 H) ; **¹³C NMR** (DMSO-d₆, 101 MHz): 172.2, 156.3, 143.9, 141.4, 131.5, 127.7, 127.1, 125.4, 120.1, 119.3, 67.3, 66.4, 53.8, 47.3, 39.8, 32.2, 27.8, 22.8 ppm ; **HRMS** calcd for C₂₄H₂₉N₂O₄ [M+H]⁺ : 409.2127 ; found 409.2213 (-0.009 m/z).

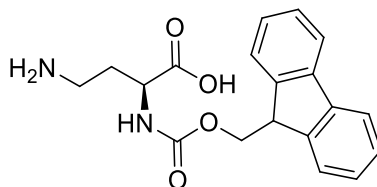
2-(1-hydroxy-3-methylbutylidene)-5,5-dimethylcyclohexane-1,3-dione (22)



In a solution of DMF (50 ml) were dissolved 3-methylbutanoic acid (1.081 ml, 9.79 mmol), 5,5-dimethylcyclohexane-1,3-dione (1.510 g, 10.77 mmol), DCC (1.530 ml, 9.79 mmol) and DMAP (1.196 g, 9.79 mmol). The solution was stirred at room temperature for 48h. Initially the medium was a pink solution, but after the night it became a white suspension. The precipitated DCU was removed by filtration and the filtrate was concentrated under reduced pressure. The orange oil obtained was dissolved in EtOAc (60 ml) and placed in a separating funnel where the organic layer was washed (2x 20 ml) with 1M KHSO₄ (pH = 2) solution and to the organic layer was extracted the compound with NaHCO₃(aq). After the aqueous layers were combined and acidified to pH 3 (the yellow solution became colorless at this pH) and replaced in a separating funnel where the compound was extracted from aqueous layers with DCM. The solvent was then evaporated. A good amount of compounds remained in the initial organic phase (EtOAc), then the organic phases (EtOAc and DCM) were combined to perform flash chromatography. The white residue was purified by flash chromatography (SiO₂, 0 % EtOAc/Hex for 3 minutes and going to 15% EtOAc/Hex in 25 minutes.) using a Combiflash Rf 200 (Isco) instrument with a 80 g column (crude material was applied via a SiO₂ 10 g). Fractions were combined and concentrated under reduced pressure to give 2-(1-hydroxy-3-methylbutylidene)-5,5-dimethylcyclohexane-1,3-dione 22 (1.391 g, 6.20 mmol, 63.3 % yield) as a yellow oil.

Formula: C₁₃H₂₀O₃; **MW:** 224.3 g/mol; **¹H NMR** (CDCl₃, 400MHz): d = 2.92 (d, J=7.0 Hz, 2 H), 2.54 (s, 2 H), 2.36 (s, 2 H), 2.15 (tspt, J=13.7, 6.7 Hz, 1 H), 1.08 (s, 6 H), 0.98 ppm (d, J=6.7 Hz, 6 H) ; **¹³C NMR** (CDCl₃, 101MHz): δ = 204.9, 198.2, 195.0, 112.2, 52.7, 48.5, 47.1, 30.6, 28.2, 25.6, 22.7 ppm; **HRMS** calcd for C₁₃H₂₁O₃ [M+H]⁺ : 225.1491 ; found 225.1317 (-0.0174 m/z).

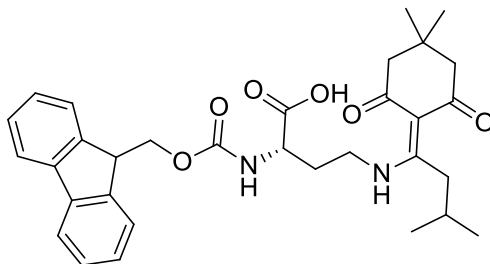
(S)-2-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-4-aminobutanoic acid (24)



Fmoc-Gln-OH (1 g, 2.71 mmol) and [Bis(trifluoroacetoxy)iodo]benzene (1.284 g, 2.99 mmol) were dissolved in a solution of DMF-H₂O 2:1 (6.79 ml). The suspension was stirred at room temperature during 15 minutes before adding the Pyridine (0.439 ml, 5.43 mmol). The resulting white suspension was stirred overnight at room temperature and the suspension became a yellow solution. Firstly, the solvent was removed under reduced pressure. The orange oil was then dissolved in water (20 ml) and the pH was adjusted to 6 with NaOH 2M. After the water was evaporated under reduced pressure and then the resultant orange oil was dissolved in MeOH and 10g of SiO₂ was added before evaporating the MeOH. A flash chromatography was performed in reverse phase (C18, 20% to 60% in 12 column volume and the gradient was maintained at 60 % for the rest of the column) using a Biotage Isolera instrument with a 120 g column (crude material was applied via a solid loading with C18). Fractions were combined and concentrated under reduced pressure, and to give Fmoc-Dab-OH 24 (0.518 g, 1.522 mmol, 56.1 % yield) as a white solid.

Formula: C₁₉H₂₀N₂O₄; **MW:** 340.37 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 7.89 (d, J=7.4 Hz, 2 H), 7.63 - 7.71 (m, 2 H), 7.37 - 7.45 (m, 2 H), 7.28 - 7.37 (m, 2 H), 6.62 (d, J=4.7 Hz, 1 H), 4.16 - 4.31 (m, 3 H), 3.61 (br. s., 1 H), 2.80 - 2.96 (m, 3 H), 1.82 - 1.93 (m, J=10.2 Hz, 1 H), 1.63 - 1.78 ppm (m, 1 H) ; **¹³C NMR** (DMSO-d₆, 101 MHz): δ = 173.6, 156.2, 145.3, 141.1, 127.9, 127.2, 125.4, 120.8, 65.2, 54.5, 47.2, 37.1, 31.2 ppm; **HRMS** calcd for C₁₉H₂₁N₂O₄ [M+H]⁺ : 341.1501 ; found 341.1377 (-0.0124 m/z). **[α]²⁵_D** = + 22.4 (c 0.241; DMSO) **[α]²⁵_D** -12.2 (c 1.8 mg/mL , EtOH)

(S)-2-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-4-((1-(4,4-dimethyl-2,6-dioxocyclohexylidene)-3-methylbutyl)amino)butanoic acid (25)

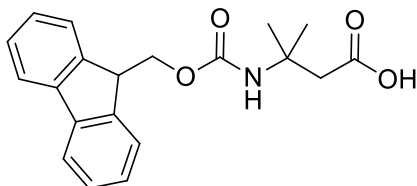


A suspension of Fmoc-Dab-OH (300 mg, 0.881 mmol) and ivDde-OH (962 μ l, 4.41 mmol) in DMF (2938 μ l) was stirred at 70 °C for 24 hours. The solvent was evaporated under reduced pressure. The residue was purified by flash chromatography using Combiflash Rf 200 (Isco) instrument with a 40 g SiO₂ column (crude material was applied via solid handling (10 g of silica). The gradient was 0 % of (20 % MeOH in DCM)/(DCM) for 3 column volume and the gradient go up to 80 % of (20 % MeOH in DCM)/(DCM) in 14 column volume. Fractions were combined and concentrated under reduced pressure to give Fmoc-Dab(ivDde)-OH 25 (275 mg, 0.503 mmol, 57.1 % yield) as a white solid.

Formula: C₃₂H₃₈N₂O₆; **MW:** 546,65 g/mol; **¹H NMR** (DMSO-d₆, 400 MHz): δ = 13.63 (t, J=5.5 Hz, 1 H), 12.88 (br. s., 1 H), 7.89 (d, J=7.8 Hz, 2 H), 7.57 - 7.78 (m, 3 H), 7.42 (t, J=7.4 Hz, 2 H), 7.32 (t, J=7.4 Hz, 2 H), 4.30 (dd, J=7.4, 2.9 Hz, 2 H), 4.22 (t, J=7.0 Hz, 1 H), 3.97 - 4.07 (m, 1 H), 3.47 - 3.63 (m, 2 H), 2.96 (br. s., 2 H), 2.26 (s, 4 H), 1.98 - 2.09 (m, 1 H), 1.85 - 1.97 (m, J=4.5, 4.5 Hz, 1 H), 1.73 - 1.85 (m, J=13.6, 6.7, 6.7 Hz, 1 H), 0.93 (s, 6 H), 0.87 ppm (dd, J=9.4, 7.0 Hz, 6 H) ; **¹³C NMR** (DMSO-d₆, 101MHz): δ = 195.2, 195.2, 175.4, 173.1, 156.1, 143.8, 143.7, 140.7, 140.7, 127.6, 127.0, 125.2, 120.1, 106.3, 65.7, 52.4, 51.4, 46.6, 36.0, 30.7, 29.5, 28.3, 27.8, 22.2, 22.1 ppm; **HRMS** calcd for C₃₂H₃₉N₂O₆ [M+H]⁺ : 547.2808 ; found 547.2635 (-0.0173 m/z). [α]_D²⁵ -6.88 (c 1.6 mg/mL, EtOH)

Formula: C₁₉H₁₉NO₄; **MW:** 325.36 g/mol; **¹H NMR** (DMSO-d₆, 400 MHz): δ = 12.29 (br. s., 1 H), 7.89 (d, J=7.4 Hz, 2 H), 7.72 (d, J=7.4 Hz, 2 H), 7.60 (s, 1 H), 7.38 - 7.46 (m, 2 H), 7.33 (td, J=7.4, 1.2 Hz, 2 H), 4.17 - 4.28 (m, 3 H), 1.35 ppm (s, 6 H); **¹³C NMR** (DMSO-d₆, 101MHz): δ = 175.8, 154.9, 143.8, 140.7, 127.6, 127.1, 125.3, 120.1, 65.3, 55.2, 46.7, 25.1 ppm; **HRMS** calcd for C₁₉H₂₀NO₄ [M+H]⁺ : 326.1392 ; found 326.1279 (-0.0113 m/z).

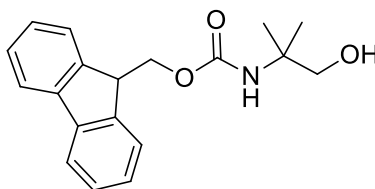
3-(((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-methylbutanoic acid (Fmoc-H-Aib-OH) (28**)**



A white suspension of 3-amino-3-methylbutanoic acid (2.409 g, 20.56 mmol) and K_2CO_3 (2.98 g, 21.59 mmol) in Dioxane/ H_2O 1:1 (40 ml) was firstly done. The Fmoc-OSuccinimide (7.28 g, 21.59 mmol) was then added and the mixture was stirred overnight at room temperature. After the night, the reaction medium became a solution and LCMS show the completion of the reaction. The solution was placed into a 250 ml separating funnel and was diluted with water (50 ml). Solid acid citric was added until a pH of 3. The aqueous layer was extracted with EtOAc (3 x 30 ml), and these organics layers were combined, washed with a brine solution, dried with Na_2SO_4 and concentrated under reduced pressure. The residue was purified by flash chromatography using Combiflash Rf 200 (Isco) instrument with a 80 g SiO_2 column (crude material was applied via solid handling) with a solvent mix of A=DCM B=20% MeOH/DCM. The gradient was 5 % of B for 6 minutes and 5%-80% B in 30 minutes. Fractions were combined and concentrated under reduced pressure and co-evaporated several time with DCM to give a white solid **28**. (4.101 g, 59 % yield)

Formula: $C_{20}H_{21}NO_4$; **MW:** 339.39 g/mol; **1H NMR** (DMSO- d_6 , 400 MHz): δ = 12.04 (br. s., 1 H), 7.88 (d, $J=7.4$ Hz, 2 H), 7.72 (d, $J=7.4$ Hz, 2 H), 7.41 (t, $J=7.4$ Hz, 2 H), 7.32 (t, $J=7.4$ Hz, 2 H), 7.16 (br. s., 1 H), 4.20 (br. s., 3 H), 2.61 (s, 2 H), 1.31 ppm (s, 6 H); **^{13}C NMR** (DMSO- d_6 , 101 MHz): δ = 172.2, 143.9, 140.7, 127.6, 127.0, 125.3, 120.1, 65.0, 50.6, 46.7, 43.6, 26.9 ppm; **HRMS** calcd for $C_{20}H_{22}NO_4$ $[M+H]^+$: 340.1549 ; found 340.1439 (-0.011 m/z).

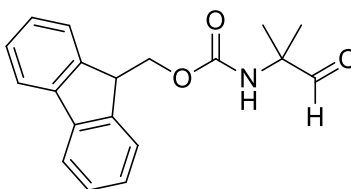
(9H-fluoren-9-yl)methyl (1-hydroxy-2-methylpropan-2-yl)carbamate (30)



Fmoc-OSuccinimide (757 mg, 2,244 mmol) was dissolved in Dioxane/DCM 1:1 (4488 μ l) This suspension was placed under nitrogen atmosphere and treated with 2-amino-2-methylpropan-1-ol (211 μ l, 2,244 mmol) dissolved in Dioxane/DCM (200 μ l) was added rapidly over 2 minutes. With this addition, the suspension became a clear solution. After DIPEA (784 μ l, 4,49 mmol) was added and the reaction was stirred during 16 hours at room temperature. The DCM was evaporated under reduced pressure and the resulting solution was transferred into a separatory funnel. A proportion EtOAc/Water (3/1) was used for the wash. The organic layer was always kept and wash with solvent in this order: 1x water, 1x 1M KHSO₄ aq, 1x water and 1x brine. The organic layer was dried with Na₂SO₄ and concentrated under reduced pressure. The residue was purified by flash chromatography using Combiflash Rf 200 (Isco) instrument with a 24 g SiO₂ column (crude material was applied via solid handling) with a solvent mix of A=Hexanes B=EtOAc. The gradient was 20 % of B for 6 minutes and 5%-90% B in 20 minutes. Fractions were combined and concentrated under reduced pressure to obtain (9H-fluoren-9-yl)methyl (1-hydroxy-2-methylpropan-2-yl)carbamate 30 (342 mg, 1,099 mmol, 49 % yield) as white solid.

Formula: C₁₉H₂₁NO₃; **MW:** 311.37 g/mol; **¹H NMR** (DMSO-d₆, 400 MHz): δ = 7.89 (d, J=7.4 Hz, 2 H), 7.71 (d, J=7.4 Hz, 2 H), 7.41 (t, J=7.2 Hz, 2 H), 7.32 (t, J=7.8 Hz, 2 H), 7.18 (br. s., 1 H), 4.20 (br. s., 3 H), 2.60 (br. s., 2 H), 1.31 ppm (s, 6 H); **¹³C NMR** (DMSO-d₆, 101MHz): δ = 156.3, 144.1, 141.6, 127.9, 127.3, 125.1, 120.2, 70.4, 66.5, 54.7, 47.5, 24.7 ppm; **HRMS** calcd for C₁₉H₂₂NO₃ [M+H]⁺ : 312.1600 ; found 312.1583 (-0.0017 m/z).

(9H-fluoren-9-yl)methyl (2-methyl-1-oxopropan-2-yl)carbamate (26)

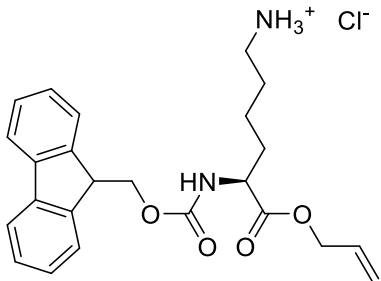
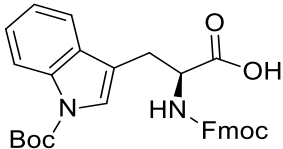
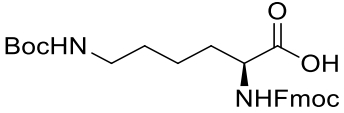
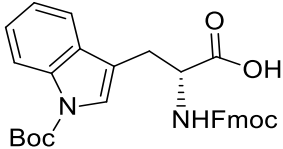
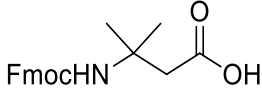


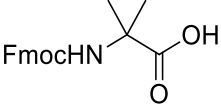
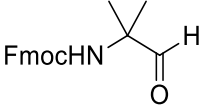
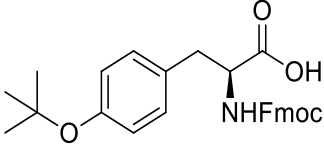
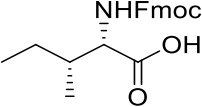
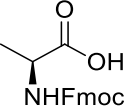
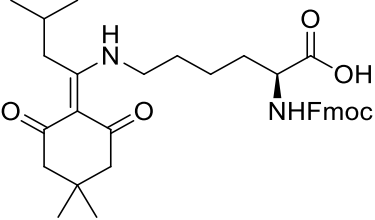
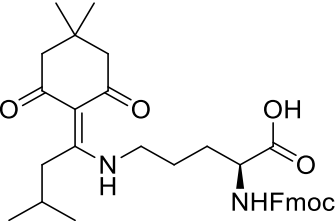
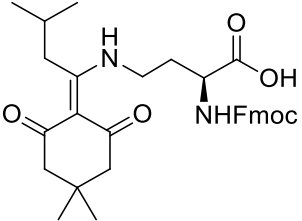
In a 25 ml pointed-shape RBF was dissolved (9H-fluoren-9-yl)methyl (1-hydroxy-2-methylpropan-2-yl)carbamate (300 mg, 0,963 mmol) in DCM (7033 μ l). Then the solution was treated with Dess-Martin Periodinane (613 mg, 1,445 mmol) (the addition was in one portion). The mixture was stirred during 1 hour at room temperature and then the reaction was diluted in EtOAc (10 ml). The white precipitate was filtered over a 0.45 μ m and was washed several times with EtOAc. The organic layer was washed with (2x5 ml) of 5 % sodium thiosulfate solution, (2 x 4 ml) of 5 % sodium carbonate, water (5 ml) and finally 4- brine solution (5 ml). Na_2SO_4 was added to the organic layer, then evaporated under reduced pressure. The residue was co-evaporated 3 times with DCM and finally dried under high vacuum. A white foam Fmoc-Aib-H 26 was obtained (292 mg, 0,944 mmol, 98 % yield).

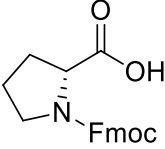
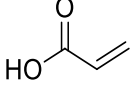
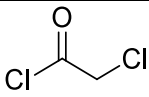
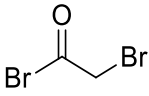
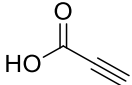
Formula: $\text{C}_{19}\text{H}_{19}\text{NO}_3$; **MW:** 309.36 g/mol; **^1H NMR** (DMSO- d_6 , 400 MHz): δ = 9.30 (s, 1 H), 7.85 - 7.94 (m, 3 H), 7.70 (t, $J=7.4$ Hz, 2 H), 7.42 (t, $J=7.4$ Hz, 2 H), 7.34 (td, $J=7.0, 0.8$ Hz, 2 H), 4.35 (d, $J=6.7$ Hz, 2 H), 4.23 (t, $J=7.0$ Hz, 1 H), 1.15 ppm (s, 6 H); **^{13}C NMR** (DMSO- d_6 , 101MHz): δ = 200.2, 156.3, 143.8, 141.7, 127.8, 127.3, 125.3, 120.3, 69.2, 55.2, 48.2, 25.3 ppm ; **HRMS** calcd for $\text{C}_{19}\text{H}_{20}\text{NO}_3$ $[\text{M}+\text{H}]^+$: 310.1443 ; found 310.1578 (-0.0135 m/z).

Synthesis of macrocycles

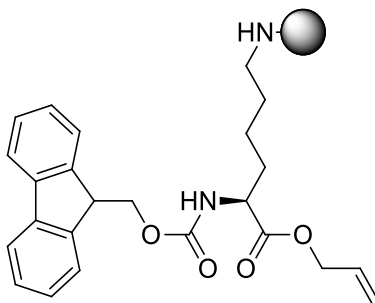
This table summarizes the amino acids used for the synthesis of macrocycles by indicating their origin, name and CAS.

Position	Structure	Name	CAS	Supplier
<u>Amino acids</u>				
AA1		<i>Fmoc-L-Lys-OAll</i>	161580-12-7	In-House
AA2/AA6		Fmoc-L-Trp(Boc)-OH	143824-78-6	Matrix InnovArk Pharma. Inc.
AA3		Fmoc-L-Lys(Boc)-OH	71989-26-9	Matrix InnovArk Pharma. Inc.
AA4		Fmoc-D-Trp(Boc)-OH	163619-04-3	Chem- Impex
AA5		Fmoc-H-Aib-OH	N/A	In-house

AA5		Fmoc-Aib-OH	94744-50-0	Combi-Blocks
AA5		Fmoc-Aib-H	N/A	In-house
AA7		Fmoc-L-Tyr (tBu)-OH	71989-38-3	Chem- Impex
AA8		Fmoc-allo-Ile-OH	251316-98-0	Chem- Impex
AA8		Fmoc-Ala-OH	35661-39-3	Chem- Impex
AA9		Fmoc-Lys(ivDde)-OH	71989-26-9	Ark Pharma. inc.
AA9		Fmoc-Orn(ivDde)-OH	1198321-33-3	Chem impex
AA9		Fmoc-Dab(ivDde)-OH	607366-21-2	In-house

AA10		Fmoc-D-Pro-OH	101555-62-8	Matrix innovation
<u>Warheads</u>				
Warhead		Acrylic acid	79-10-7	Sigma- Aldrich
Warhead		Chloroacetyl chloride	79-04-9	Sigma- Aldrich
Warhead		Bromoacetyl bromide	598-21-0	Sigma- Aldrich
Warhead		Propiolic acid	471-25-0	Sigma- Aldrich

Anchoring of the first amino acid on the resin (17)



Materials:

Peptide synthesis vessels 250 mL: (Chemglass life sciences, CG-1861)

MaxQ 2000 Digital Shaker: (Thermo Scientific, SHKE2000)

Note: Use new bottle of dichloromethane to increase the yield

In a first step, the dry 2-chlorotrityl chloride resin (1.26 mmol/g) (3.54 g, 4.46 mmol) was introduced into the peptide synthesis vessel, then was washed with CH_2Cl_2 (2x 45 ml). Between each wash, the resin containing CH_2Cl_2 was shaken 2 minutes at a speed of 150 rpm on the MaxQ 2000. The next step was to solubilize the Fmoc-L-Lys-OAll (2.380 g, 5.35 mmol) in CH_2Cl_2 (50 ml) and adding DIPEA (3.11 ml, 17.83 mmol) in a 100 mL round-bottom flask. This orange solution was then added to the wet resin and after 5 minutes of shaking, the pH was close to 9. The glass reactor was placed in a horizontal way on the MaxQ200 shaker and was shaken at 150 rpm for 1.5 hours. The liquid was removed from the reactor by the coarse frit and the resin was washed twice with CH_2Cl_2 (2 x 50 ml). In order to cap the resin, there were two treatments with a solution of $\text{CH}_2\text{Cl}_2/\text{MeOH}/\text{DIPEA}$ 80:15:5 (50 ml, 4.46 mmol) for 10 minutes of stirring. The solvent is drained between each treatment. When the second treatment was over, the solution was removed and the resin was washed (1-2 minutes each times) with CH_2Cl_2 (3x50mL), $\text{CH}_2\text{Cl}_2\text{-MeOH}$ 1:1 (3x50mL), $\text{CH}_2\text{Cl}_2\text{-MeOH}$ 0.25:0.75(3x50mL) and MeOH (3x50mL). The yellow resin was dried under a flow of nitrogen and placed under high vacuum overnight to obtain a loading of 96 % (5.799 g).

Measurement of the loading on the resin

The loading procedure comes from the following website: SPS monitoring_S V Ley 1002_ftp page 115; http://www.anaspec.com/html/peptide_tips.html

Materials:

3 volumetric flasks of 10 mL

Uv detector: Ultraspec 2100 Amershaam Biosciences

2 Uv quartz cuvette 3.5 mL: Sigma-Aldrich, Z276669

To obtain a more precise loading value it is necessary to carry out a duplicate, therefore this experiment requires 3 volumetric flasks, 2 for the resin and 1 for the blank. First, an amount of dry resin will be weighed on an analytical balance, aim for about 1-3 mg for each replicate. The resins are transferred into 2 mL vials and for each there is added 1.00 mL of a solution of 20% piperidine/DMF, then shake for 20 minutes. After this period, the supernatant was transferred into 10 mL volumetric flask and the volume was completed to 10 mL with DMF and mix well. For the blank flask, only DMF was added. On the UV detector, there was a calibration with the blank sample by filling the quartz cuvette. Subsequently, the absorbance was taken for the two replicates by filling a second quartz cuvette. The formula used to calculate the loading was as follows:

$$\text{Loading} \left(\frac{\text{mmol}}{\text{g}} \right) = \frac{A_{301\text{nm}} \times V \times 1000 \text{ mg/g}}{7800 \text{ mmol}^{-1}\text{cm}^{-1} \times W \times l}$$

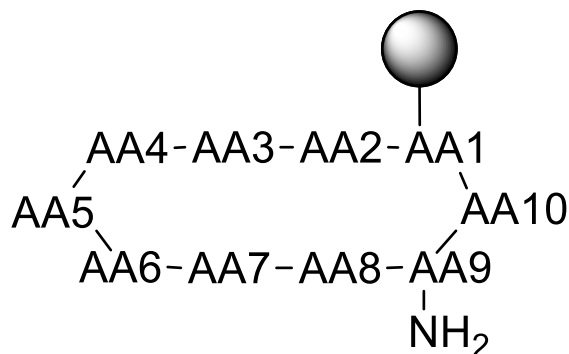
Where **A** is the Absorbance of the solution at 301 nm

V= Total volume of the solution (mL)

W= Mass of the resin weighed (mg)

l= Quartz cuvette length (cm)

General procedure for the synthesis of macrocycles



Materials:

MaxQ 2000 Digital Shaker: Thermo Scientific, SHKE2000

20 mL syringe with a frit: Torviq, SF-1000

3 mL syringe with a frit: Torviq, SF-1000

The general procedure initially consists of anchoring the first amino acid on the resin, followed by an alternation between deprotection of the Fmoc group of the amine and a peptide coupling with the next amino acid fragment. There is coupling up to the tenth amino acid before carrying out the macrocyclization. Finally, there is the addition of the warhead and the release macrocycle from the resin.

1st Fmoc deprotection: The dry yellow resin (1.000 g, 0.739 mmol) was swollen with DMF (8 ml) for 30 minutes. The solvent was then drained out and piperidine 20% / DMF (10.35 ml, 20.69 mmol) was added to the syringe by the bottom and shaken at 150 rpm on the MaxQ 2000. After 5 minutes, the solution was drained out and the resin was treated with another PIPERIDINE 20% / DMF (10.35 ml, 20.69 mmol), but this time the syringe was shaken 20 minutes on the MaxQ2000. After this period, the solvent was released and the resin was washed with several solvent (1-2

minutes each time, 8,0 ml, 150 rpm). These washing steps were DMF (3 x 8mL) then DMF/MeOH 1:1 (2 x 8mL) and then DMF (4 x 8mL).

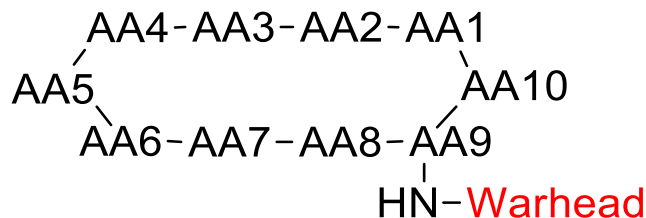
Coupling (general procedure): 2 coupling steps were done for every amino acid fragment. These couplings were performed using 1.5 eq of HATU (0.421 g, 1.109 mmol), 1.6 eq of Fmoc-amino acid and 10 eq. DIPEA (1.291 ml, 7.39 mmol) every 0.5 hour. There was no washing step between the first and the second coupling. In a first time, Fmoc-amino acid and HATU were dissolved in DMF (16 ml total volume solution, 8.0 ml for each coupling). This DMF stock solution was separated in 2 vials (2 x 8 ml). During the moment of the coupling, DIPEA (1.291 ml, 7.39 mmol) was added in the corresponding vial and the resulting solution was quickly stirred and immediately drawn in the syringe. The syringe was shaken for 0.5 hour, then the liquid was removed. This procedure was repeated for the second coupling. After these 30 minutes, the solution was drained and the resin was washed (1-2 minutes each time, 8,0 ml, 150 rpm) DMF (3 x 8mL) then DMF/MeOH 1:1 (2 x 8mL) and then DMF(4 x 8mL).

Fmoc deprotection (general procedure): The wet resin was firstly treated with 50 % piperidine / DMF (10.35 ml, 51.7 mmol) and shaken for 5 minutes. The solution was then drained and another treatment of 50 % piperidine / DMF (10.35 ml, 51.7 mmol) was executed, but this time with 20 minutes of shaking. The liquid was removed from the syringe and the resin was washed (1-2 minutes each time, 8,0 ml, 150 rpm) DMF (3 x 8mL) then DMF/MeOH 1:1 (2 x 8mL) and then DMF (4 x 8mL). After the deprotection, the resin was used immediately for the coupling or left in NMP overnight. **Note:** For the deprotection of AA2 only one deprotection step with 50 % piperidine / DMF (10.35 ml, 51.7 mmol) was executed with a 5 minutes shaking.

Macrocyclization step (general procedure): A quantity of dry resin comprising the 10 coupled amino acids desired was weighed in a 3 mL frit syringe (0.100 g, 0.037 mmol) and was washed (1-2 minutes each times, 1.0 ml, 150 rpm) with DMF (3 x 1mL). In an adjacent vial, a solution of PIPERIDINE 20%/DMF (0.518 ml, 1.036 mmol) was cooled at 10 °C and degassed under vacuum. In a second part, the Pd(Ph₃P)₄ (2.138 mg, 1.850 μmol) was added to this vial and the ice-bath was removed. The mixture was degassed under vacuum for 2 minutes with a magnetic agitation. It's was a yellow solution. The syringe containing the resin was purged several times with nitrogen and the solution with the palladium catalyst was pulled in the syringe. The syringe was placed

horizontally on a MaxQ 2000 shaker and shaken for 1 hour. Another treatment of 20 % piperidine/DMF (0.518 ml, 1.036 mmol) and Pd(Ph₃P)₄ (2.138 mg, 1.850 μmol) was donned and shaken for another 1 hour. After this time, the solution was drained and the resin was washed (1-2 minutes each time, 1,0 ml) with DMF (3 x 1 mL), DMF/MeOH 1:1 (2 x 1 mL), and DMF (4 x 1 mL). The washed resin was neutralized twice with a solution of HOBT hydrate (0.011 g, 0.074 mmol) in DMF (2 x 1 ml) (2 x 2 minutes). The resin was finally washed (1-2 minutes each time, 1,0 ml) with DMF (3 x 1 mL), DMF/MeOH 1:1 (2 x 1 mL), and DMF (4 x 1 mL). In a 4 ml vial was dissolved DEBPT (0.022 g, 0.074 mmol) in DMF (1 ml). To this vial was added DIPEA (0.026 ml, 0.148 mmol) and the solution was aspirated in the 3 mL frit syringe and then shaken overnight on the MaxQ 2000. When the DIPEA was added to the DEBPT solution, the colorless solution became a bright yellow. After this period of time, the liquid was removed from the syringe and the resin was washed with several solvents in this order (1-2 minutes each time, 1 ml) : DMF (3 x 1 mL), DMF/MeOH 1:1 (2 x 1 mL), and DMF (4 x 1 mL) , HOBT hydrate (0.011 g, 0.074 mmol) in DMF (2 x 1 ml) (2 x 2 minutes), DMF (3 x 1 mL), DMF/MeOH 1:1 (2 x 1 mL), and DMF (4 x 1 mL)

ivDde protecting group (general procedure): In a 4 ml vial was dissolved hydrazine hydrate (50 μl, 1.020 mmol) in DMF (1 ml). This solution was added to the wet resin (0.037 mmol) and the 3 ml syringe was shaken during 5 minutes on the MAXQ 2000. The liquid was removed and another treatment with hydrazine hydrate (50 μl, 1.020 mmol) in DMF (1 ml) was donned for 1 hour this time. The resin changed color to grey from yellow after this deprotecting step. After this period, the liquid was removed from the syringe and the resin was washed with several solvents in this order (1-2 minutes each time, 1 ml) : DMF (3 x 1 mL), DMF/MeOH 1:1 (2 x 1 mL), and DMF (4 x 1 mL) , HOBT hydrate (0.011 g, 0.074 mmol) in DMF (2 x 1 ml) (2 x 2 minutes), DMF (3 x 1 mL), DMF/MeOH 1:1 (2 x 1 mL), and DMF (4 x 1 mL).



Warhead

4 warheads were used for the synthesis of macrocycles and they all have a distinct methodology. So these will be described one by one.

Acrylamide warhead (general procedure): With an amount of 0.037 mmol of resin two coupling steps were performed using in total 3 eq of HATU (0.042 g, 0.111 mmol) , 3.2 eq. of acrylic acid (8.13 μ l, 0.118 mmol) and 10 eq. DIPEA (0.065 ml, 0.370 mmol). Initially, half of reagents were injected in 1 ml of DMF and the 3 mL frit syringe was shaken for 15 minutes. The liquid was removed and another treatment with acrylic acid (4.06 μ l, 0.059 mmol), HATU (21 mg, 0.055 mmol) and DIPEA (32 μ l, 0.185 mmol) in DMF (1 ml) was performed for another 15 minutes. After the resin was washed (1-2 minutes each time, 1,0 ml) with DMF (3 x 1 mL), DMF/MeOH 1:1 (2 x 1 mL), and DMF (4 x 1 mL).

Chloroacetamide warhead (general procedure): With an amount of 0.037 mmol of resin a first coupling step was performed using 2-chloroacetyl chloride (8.84 μ l, 0.111 mmol) and 5 eq. DIPEA (0.039 ml, 0.222 mmol) in 1 ml of DCM. The 3 mL frit syringe was shaken for 1 hour, then the liquid was removed. Another coupling was performed with the same condition as described for 1 hour. After the resin was washed (1-2 minutes each time, 1,0 ml) with DCM (3 x 1 mL), DCM/ACN 1:1 (2 x 1 mL), and DCM (4 x 1 mL).

Bromoacetamide warhead (general procedure): First, the resin (0.037 mmol) was contacted with 0.5 ml of cold DCM mixed with DIPEA (0.039 ml, 0.223 mmol) and the syringe was shaken for 3 minutes. In a second vial 0.5 ml of cold DCM was mixed with bromoacetyl bromide (6.46 μ l, 0.074 mmol) and then was added to the 3 mL frit syringe. The reaction was shaken during 30 minutes

at room temperature. After the resin was washed (1-2 minutes each time, 1,0 ml) with DCM (3 x 1 mL), DCM/ACN 1:1 (2 x 1 mL), and DCM (4 x 1 mL).

Propiolamide warhead (general procedure): In a 4 ml vial where dissolved DIC (0.015 ml, 0.096 mmol) and propiolic acid (5.89 μ l, 0.096 mmol) in 1 ml of cold DCM. The solution was agitated during 1 hour at 0°C. It was a yellow solution. After this time, the solution was added to the 3 ml syringe containing the resin (0.032 mmol). The reaction was shaken during 4 hours at room temperature and the resin became more brownish/orange. After the liquid in the syringe was drained out and the resin was washed (1-2 minutes each time, 1,0 ml) with DCM (3x), DCM/ACN 1:1 (2x) and DCM (4x).

Final deprotection and cleavage from the resin (general procedure): In a first step the resin was washed (1-2 minutes each times, 1 ml) with CH₂Cl₂ (3x), CH₂Cl₂: ACN 1:1 (3x) and CH₂Cl₂ (4x). Then the resin (0.037 mmol) was treated with a mix TFA/H₂O/TIS (95/2.5/2.5) (1.0 ml) for 1 hour. Another treatment of TFA/H₂O/TIS (95/2.5/2.5) (0.5 ml) but this time for 5 minutes. The resin turned to a red color when mix with the TFA solution. IMPORTANT the liquid of each treatment was recovery in a round bottom flask of 25 ml. To the RBF was added 5 ml of cold (0 °C) MTBE to help precipitation. Then the solid was filtered and washed with MTBE. The compound was dried under high vacuum for 30 minutes. After the solid was solubilized in 0.6 ml of DMSO-H₂O (9:1) and left for 18 hours to affect the decarboxylation of Boc protecting group. The last step was two purification by prep-HPLC with a collection trigger by mass.

The **first purification** was carried out with these conditions:

Solvent A: MeOH:H₂O 5:95, AcOH 0.1%

Solvent B: MeOH:H₂O 95:5, AcOH 0.1%

Gradient: Start 15% of B for 1.5 min - initial conditions

1.5 to 8.5 min - gradient up to 70% B;

8.5 to 8.6 min ramp to 100% B; maintain up to 10.7 min

10.7 to 10.8 min 15% of B; post run 0.5 min.

Stop time 11.5 min

Column #2: Agilent C8-5 μ m; 21 X 50 mm

Threshold= 1000; collection in negative mode

The **second purification** was carried out with these conditions:

Solvent A: ACN:H₂O 5:95, TFA 0.05%

Solvent B: ACN:H₂O 95:5, TFA 0.05%

Gradient: Start 0% of B for 1.5 min - initial conditions

1.5 to 8.5 min - gradient up to 50% B;

8.5 to 8.6 min ramp to 70% B; maintain up to 10.7 min

10.7 to 10.8 min 0% of B; post run 0.5 min.

Stop time 11.5 min

Column #2: Agilent C8-5 μ m; 21 X 50 mm

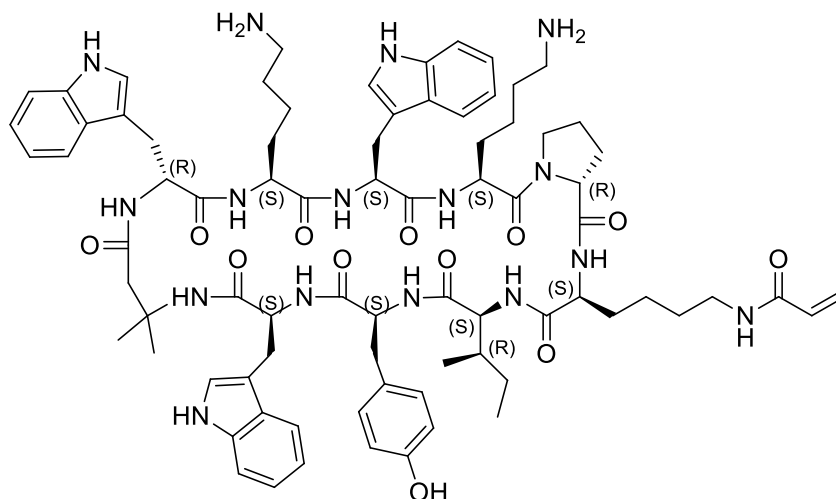
Threshold= 1000; collection in positive mode

After evaporation of the selected tubes, the compound was dissolved in a solution ACN/H₂O (3:7) and transferred to a 4 ml vial for lyophilisation overnight to afford the final compound.

During all coupling steps, an LCMS was performed to check the progress of each reaction. The methodology describing how this was done is presented here:

LCMS sample preparation: Into a glass dropper was placed some glass wool to play a filter role. After on a tip of another dropper was taken some resin and placed on the glass wool with DCM. Then the resin was washed with DCM, MeOH and DCM. Finally, the resin was placed in contact with 100 μ L of TFA for 30 minutes. Liquid was concentrated with a nitrogen flow and then the white solid was solubilized in MeOH and analysed by LCMS.

UM0140184



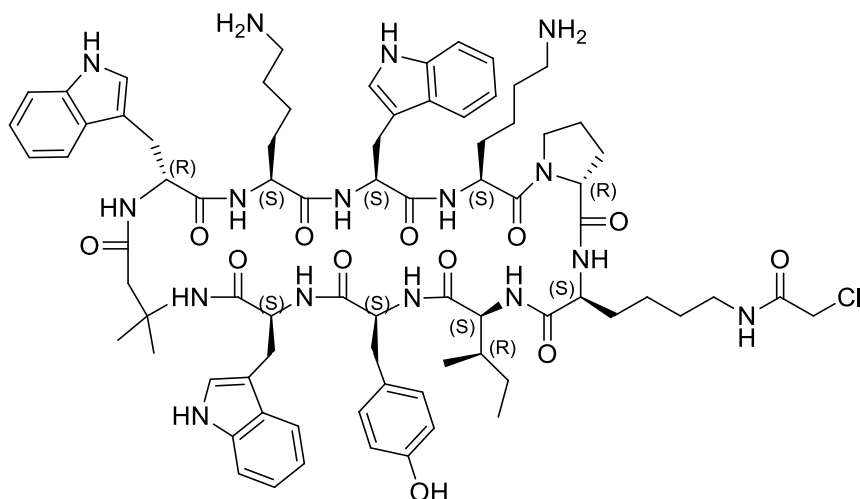
This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA7. Once the 7th coupling was done the resin was separated and the procedure continued with 100 mg of resin (0.037 mmol) until the end of the procedure.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp(Boc)-OH, **AA5**: Fmoc-H-Aib-OH, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr(tBu)-OH, **AA8**: Fmoc-*allo*-Ile-OH, **AA9**: Fmoc-Lys(ivDde)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the ivDde group according to the general procedure, then the coupling of the amine with **acrylic acid** according to the conditions described above. After the acid cleavage and the purification steps an amount of 7 mg (12.87 % yield) of a white solid was obtained, named as UM0140184.

Formula: C₇₉H₁₀₄N₁₆O₁₂; **MW:** 1469.77 g/mol; **HRMS** calcd for : C₇₉H₁₀₅N₁₆O₁₂ [M+H]⁺ : 1469.8020 ; found 1469.8244 (-0.0224 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH:H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.800 min

UM0140186



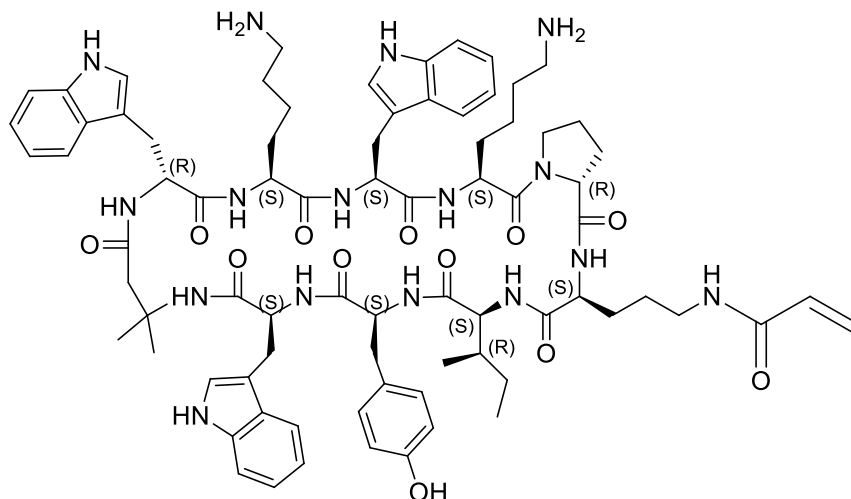
This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA7. Once the 7th coupling was done the resin was separated and the procedure continued with 100 mg of resin (0.037 mmol) until the end of the procedure.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp (Boc)-OH, **AA5**: Fmoc-H-Aib-OH, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr (tBu)-OH, **AA8**: Fmoc-*allo*-Ile-OH, **AA9**: Fmoc-Lys(ivDde)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the ivDde group according to the general procedure, then the coupling of the amine with **2-chloroacetyl chloride** according to the conditions described above. After the acid cleavage and the purification steps an amount of 8.84 mg (16.01 % yield) of a white solid was obtained, named as UM0140186.

Formula: C₇₈H₁₀₃N₁₆O₁₂Cl; **MW**: 1492.21 g/mol; **HRMS** calcd for: C₇₈H₁₀₄N₁₆O₁₂Cl [M+H]⁺ : 1491.7708 ; found 1491.7861 (-0.0153 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH: H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.810 min

UM0140191



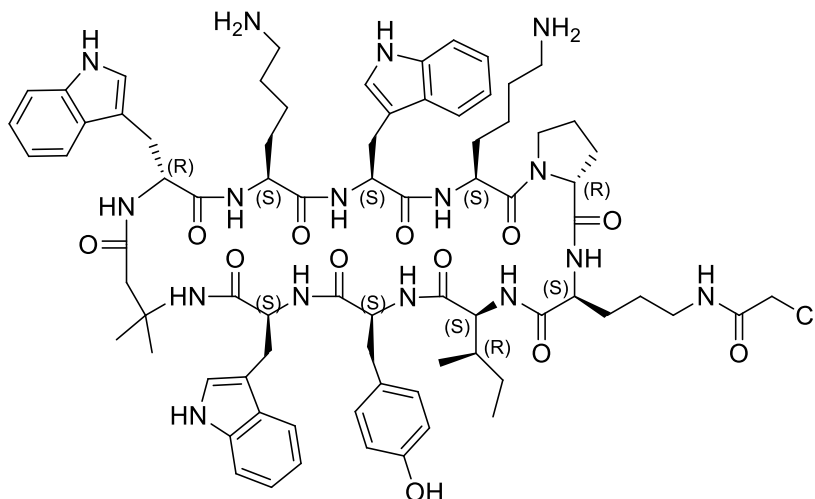
This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA7. Once the 7th coupling was done the resin was separated and the procedure continued with 100 mg of resin (0.037 mmol) until the end of the procedure.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp (Boc)-OH, **AA5**: Fmoc-H-Aib-OH, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr (tBu)-OH, **AA8**: Fmoc-ortho-Ile-OH, **AA9**: Fmoc-Orn(ivDde)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the ivDde group according to the general procedure, then the coupling of the amine with **acrylic acid** according to the conditions described above. After the acid cleavage and the purification steps an amount of 10.65 mg (19.77 % yield) of a white solid was obtained, named as UM0140191.

Formula: C₇₈H₁₀₂N₁₆O₁₂; **MW:** 1455.74 g/mol; **HRMS** calcd for: C₇₈H₁₀₃N₁₆O₁₂ [M+H]⁺ : 1455.7941 ; found 1455.8197 (-0.0256 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH: H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.800 min

UM0140327



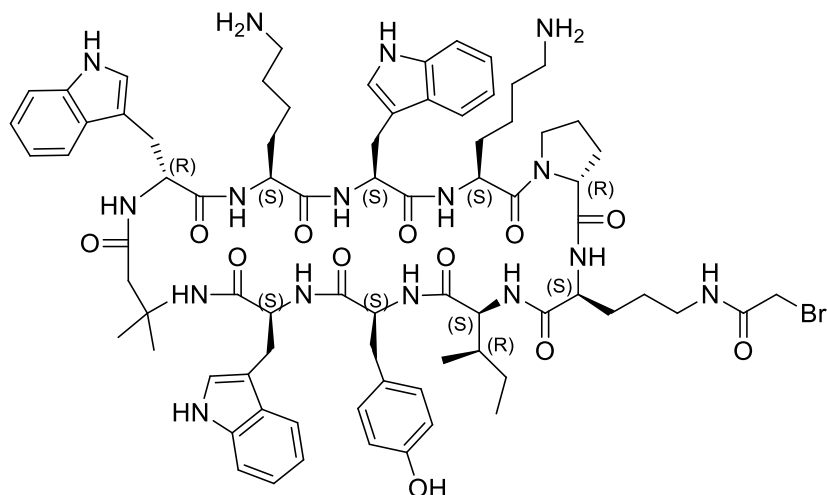
This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA7. Once the 7th coupling was done the resin was separated and the procedure continued with 100 mg of resin (0.037 mmol) until the end of the procedure.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp (Boc)-OH, **AA5**: Fmoc-H-Aib-OH, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr (tBu)-OH, **AA8**: Fmoc-allo-Ile-OH, **AA9**: Fmoc-Orn(ivDde)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the ivDde group according to the general procedure, then the coupling of the amine with **2-chloroacetyl chloride** according to the conditions described above. After the acid cleavage and the purification steps an amount of 9.53 mg (17.42 % yield) of a white solid was obtained, named as UM0140327.

Formula: $C_{77}H_{101}N_{16}O_{12}Cl$; **MW**: 1478.18 g/mol; **HRMS** calcd for : $C_{77}H_{102}N_{16}O_{12}Cl$ $[M+H]^+$: 1477.7552 ; found 1477.7667 (-0.0115 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH:H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 2.064 min

UM0140401



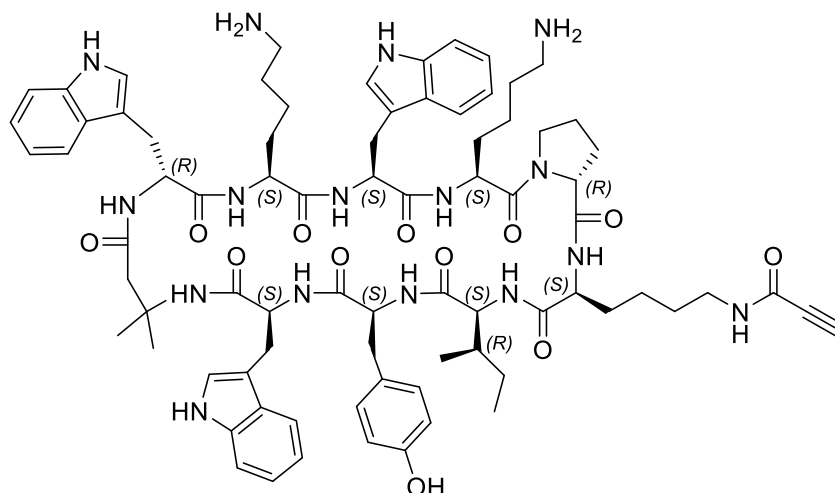
This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA7. Once the 7th coupling was done the resin was separated and the procedure continued with 100 mg of resin (0.037 mmol) until the end of the procedure.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp(Boc)-OH, **AA5**: Fmoc-H-Aib-OH, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr(tBu)-OH, **AA8**: Fmoc-*allo*-Ile-OH, **AA9**: Fmoc-Orn(*ivDde*)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the *ivDde* group according to the general procedure, then the coupling of the amine with **2-bromoacetyl bromide** according to the conditions described above. After the acid cleavage and the purification steps an amount of 11.36 mg (20.16 % yield) of a white solid was obtained, named as UM0140401.

Formula: C₇₇H₁₀₁N₁₆O₁₂Br; **MW**: 1522.63 g/mol; **HRMS** calcd for: C₇₇H₁₀₂N₁₆O₁₂Br [M+H]⁺ : 1521.7047 ; found 1521.7041 (-0.0006 m/z). **HPLC** (Kinetex C18 2.6 microns Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH:H₂O:TFA (5:95:0.5) B MeOH:H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.828 min

UM0140438



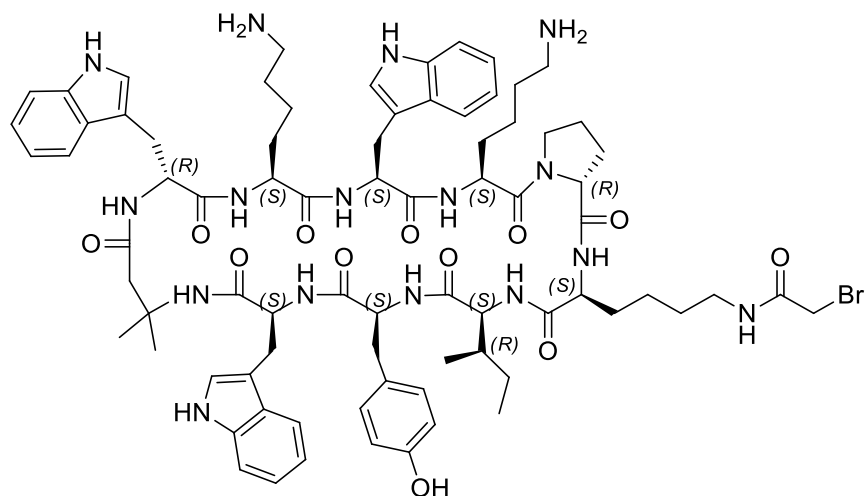
This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA7. Once the 7th coupling was done the resin was separated and the procedure continued with 100 mg of resin (0.037 mmol) until the end of the procedure.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp (Boc)-OH, **AA5**: Fmoc-H-Aib-OH, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr (tBu)-OH, **AA8**: Fmoc-*allo*-Ile-OH, **AA9**: Fmoc-Lys(ivDde)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the ivDde group according to the general procedure, then the coupling of the amine with **propionic acid** according to the conditions described above. After the acid cleavage and the purification steps an amount of 9.16 mg (16.87 % yield) of a white solid was obtained, named as UM0140438.

Formula: C₇₉H₁₀₂N₁₆O₁₂; **MW**: 1467.76 g/mol; **HRMS** calcd for: C₇₉H₁₀₃N₁₆O₁₂ [M+H]⁺ : 1467.7941 ; found 1467.8042 (-0.0101 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH: H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.810 min

UM0140439



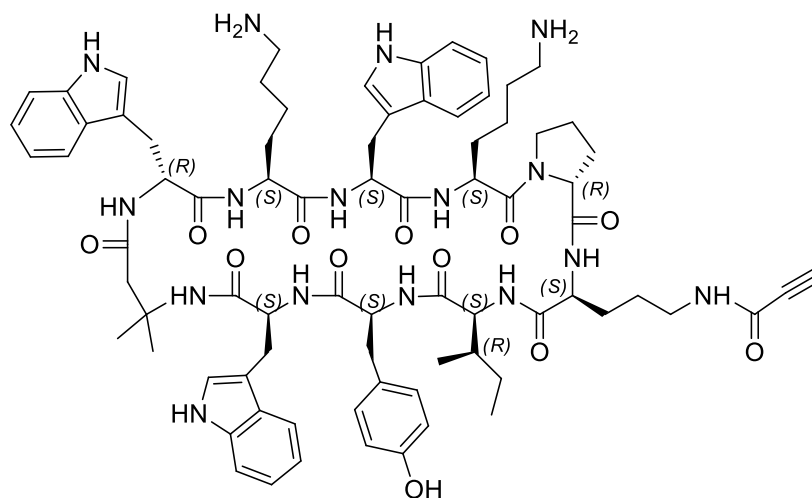
This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA7. Once the 7th coupling was done the resin was separated and the procedure continued with 100 mg of resin (0.037 mmol) until the end of the procedure.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp(Boc)-OH, **AA5**: Fmoc-H-Aib-OH, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr(tBu)-OH, **AA8**: Fmoc-*allo*-Ile-OH, **AA9**: Fmoc-Lys(ivDde)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the ivDde group according to the general procedure, then the coupling of the amine with **2-bromoacetyl bromide** according to the conditions described above. After the acid cleavage and the purification steps an amount of 2.91 mg (5.12 % yield) of a white solid was obtained, named as UM0140439.

Formula: C₇₈H₁₀₃N₁₆O₁₂Br; **MW**: 1536.66 g/mol; **HRMS** calcd for: C₇₈H₁₀₄N₁₆O₁₂Br [M+H]⁺ : 1535.7203 ; found 1535.7380 (-0.0177 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH: H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.836 min

UM0140440



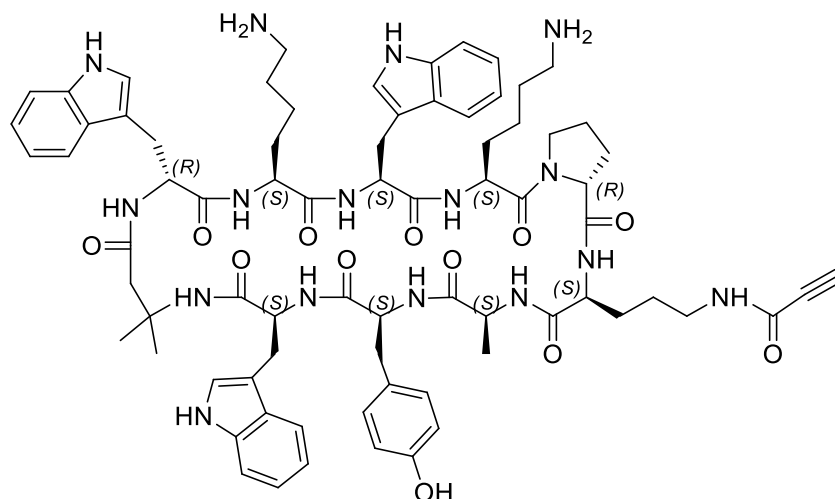
This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA7. Once the 7th coupling was done the resin was separated and the procedure continued with 100 mg of resin (0.037 mmol) until the end of the procedure.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAlI, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp(Boc)-OH, **AA5**: Fmoc-H-Aib-OH, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr(tBu)-OH, **AA8**: Fmoc-*allo*-Ile-OH, **AA9**: Fmoc-Orn(*ivDde*)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the *ivDde* group according to the general procedure, then the coupling of the amine with **propionic acid** according to the conditions described above. After the acid cleavage and the purification steps an amount of 12.64 mg (23.50 % yield) of a white solid was obtained, named as UM0140440.

Formula: C₇₈H₁₀₀N₁₆O₁₂; **MW**: 1453.73 g/mol; **HRMS** calcd for: C₇₈H₁₀₁N₁₆O₁₂ [M+H]⁺ : 1453.7785 ; found 1453.7804 (-0.002 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH: H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.795 min

UM0140468

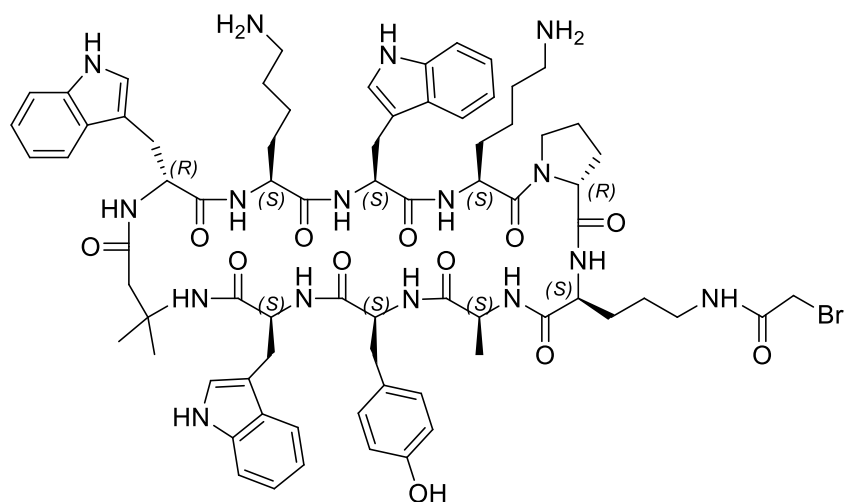


This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA7. Once the 7th coupling was done the resin was separated and the procedure continued with 90 mg of resin (0.026 mmol) until the end of the procedure.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp(Boc)-OH, **AA5**: Fmoc-H-Aib-OH, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr(tBu)-OH, **AA8**: Fmoc-Ala-OH, **AA9**: Fmoc-Orn(ivDde)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the ivDde group according to the general procedure, then the coupling of the amine with **propionic acid** according to the conditions described above. After the acid cleavage and the purification steps an amount of 1.01 mg (2.30 % yield) of a white solid was obtained, named as UM0140468.

Formula: C₇₅H₉₄N₁₆O₁₂; **MW**: 1411.65 g/mol; **HRMS** calcd for: C₇₅H₉₅N₁₆O₁₂ [M+H]⁺ : 1411.7315 ; found 1411.6573 (-0.0742 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH: H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.729 min

UM0140586

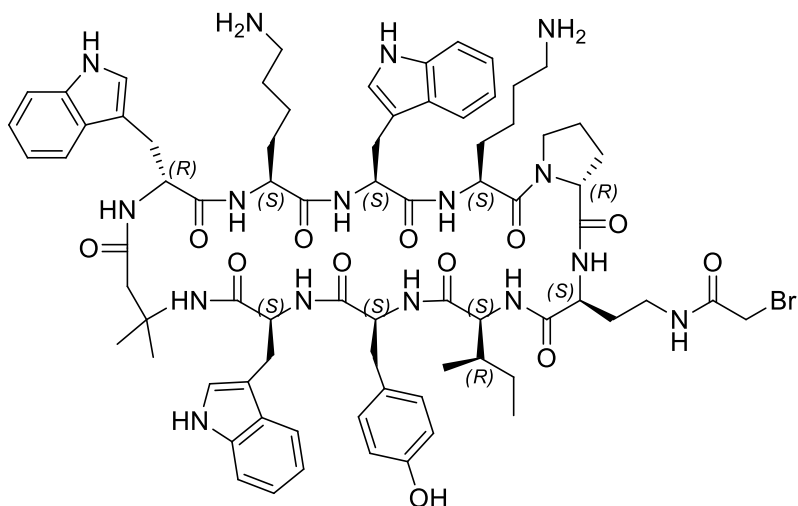
This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA7. Once the 7th coupling was done the resin was separated and the procedure continued with 90 mg of resin (0.026 mmol) until the end of the procedure.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp (Boc)-OH, **AA5**: Fmoc-H-Aib-OH, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr (tBu)-OH, **AA8**: Fmoc-Ala-OH, **AA9**: Fmoc-Orn(ivDde)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the ivDde group according to the general procedure, then the coupling of the amine with **2-bromoacetyl bromide** according to the conditions described above. After the acid cleavage and the purification steps an amount of 1.82 mg (4.74 % yield) of a white solid was obtained, named as UM0140586.

Formula: C₇₄H₉₅N₁₆O₁₂Br; **MW**: 1480.55 g/mol; **HRMS** calcd for: C₇₄H₉₆N₁₆O₁₂Br [M+H]⁺ : 1479.6577 ; found 1479.6394 (-0.0183 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH: H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.742 min

UM0140690



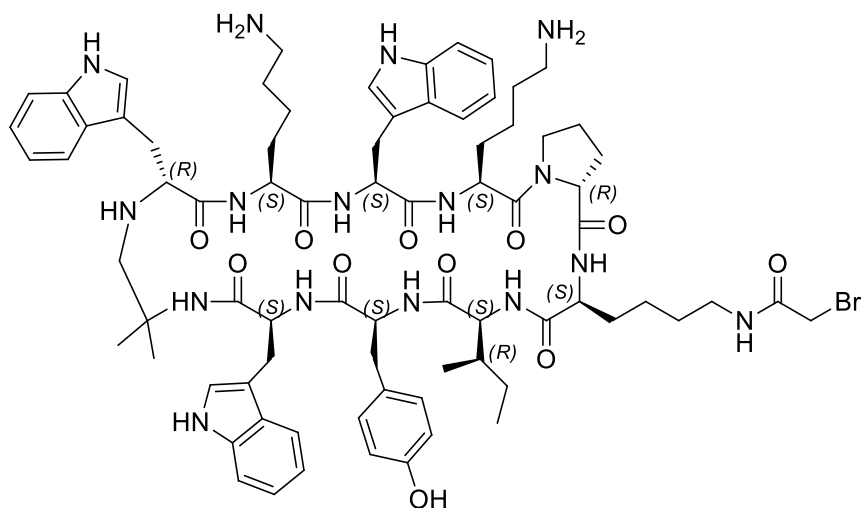
This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA7. Once the 7th coupling was done the resin was separated and the procedure continued with 138 mg of resin (0.037 mmol) until the end of the procedure.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp(Boc)-OH, **AA5**: Fmoc-H-Aib-OH, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr(tBu)-OH, **AA8**: Fmoc-*allo*-Ile-OH, **AA9**: Fmoc-Dab(ivDde)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the ivDde group according to the general procedure, then the coupling of the amine with **2-bromoacetyl bromide** according to the conditions described above. After the acid cleavage and the purification steps an amount of 3.48 mg (6.23 % yield) of a white solid was obtained, named as UM0140690.

Formula: C₇₆H₉₉N₁₆O₁₂Br; **MW**: 1508.65 g/mol; **HRMS** calcd for: C₇₆H₁₀₀N₁₆O₁₂Br [M+H]⁺ : 1507.6890 ; found 1507.6764 (-0.0126 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH: H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.797 min

UM0140691



This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA4. Once the 4th coupling was done the resin was separated and the procedure continued with 504 mg of resin (0.218 mmol) until AA8 where it was split another time. The final amount selected to synthesis this macrocycle was 0.055 mmol of resin having the first 8 amino acids coupled.

There is a variance in the synthesis of this macrocycle. Once the AA4 amino acids has been deprotected by piperidine, there is a reductive amination with the aldehyde (Fmoc-Aib-H) to obtain the primary amine.

Fmoc deprotection of AA4: The wet resin (0.218 mmol) was firstly treated with piperidine 20 % / DMF (3.10 ml, 15.52 mmol) and shaken for 5 minutes. The solution was then drained and another treatment of piperidine 20 % / DMF (3.10 ml, 15.52 mmol) was executed, but this time with 20 minutes of shaking. The liquid was removed from the syringe and the resin was washed (1-2 minutes each time, 5.0 ml) with DMF (3 x 5 mL), DMF/MeOH 1:1 (2 x 5 mL) and DMF (4 x 5 mL).

Reductive amination: The resin was washed several times with THF/MeOH (1:1) (3 x 5 ml, 2 minutes). In a vial of 4 ml was prepared a solution with Fmoc-Aib-H (270 mg, 0.873 mmol) and 4 ml of THF/MeOH (1:1). The solution was split in two (vial 1A and vial 1B). In each was added 1 ml

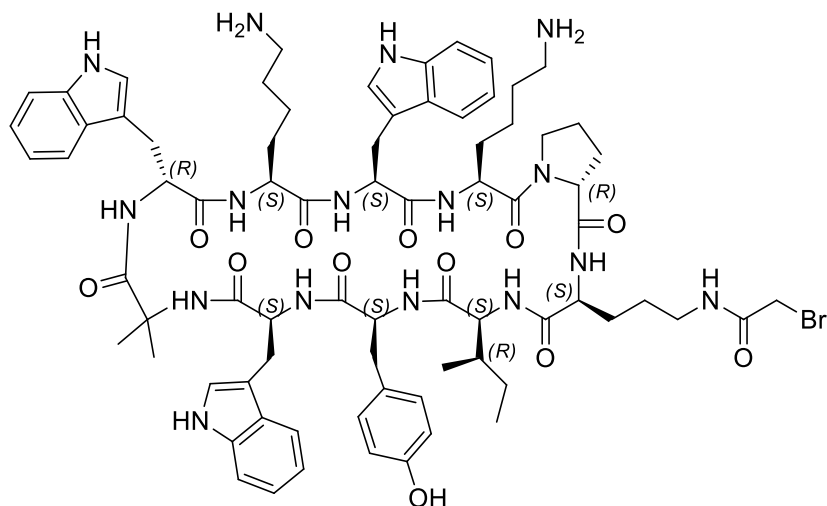
of THF/MeOH (1:1) and ACETIC ACID (0.019 ml, 0.327 mmol) and SODIUM CYANOBOROHYDRIDE (68.6 mg, 1.091 mmol). After the solution was injected into the syringe and shaken for 90 minutes. The treatment was repeated with the other solution (1B). After the liquid was drained out and the resin was washed with several solvent: (5 ml, 1-2 minutes each times) DMF (3 x 5 mL), DMF/MeOH 1:1 (2 x 5 mL) and DMF (4 x 5 mL). After the procedure continued as usual for the other amino acids.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp (Boc)-OH, **AA5**: Fmoc-Aib-H, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr (tBu)-OH, **AA8**: : Fmoc-*allo*-Ile-OH, **AA9**: Fmoc-Lys(ivDde)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the ivDde group according to the general procedure, then the coupling of the amine with **2-bromoacetyl bromide** according to the conditions described above. After the acid cleavage and the purification steps an amount of 11.18 mg (13.47 % yield) of a white solid was obtained, named as UM0140691.

Formula: $C_{77}H_{103}N_{16}O_{11}Br$; **MW**: 1508.60 g/mol; **HRMS** calcd for: $C_{77}H_{104}N_{16}O_{11}Br [M+H (Br^{81})]^+$: 1509.7233 ; found 1509.7442 (-0.0209 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH: H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.679 min

UM0140692

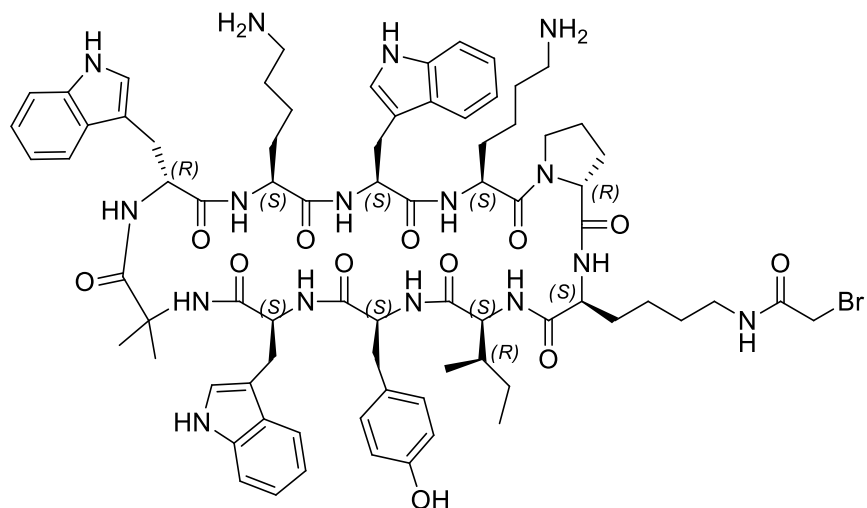


This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA4. Once the 4th coupling was done the resin was separated and the procedure continued with 504 mg of resin (0.218 mmol) until AA8 where it was split another time. The final amount selected to synthesis this macrocycle was 0.055 mmol of resin having the first 8 amino acids coupled.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp (Boc)-OH, **AA5**: Fmoc-Aib-OH, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr (tBu)-OH, **AA8**: : Fmoc-allo-Ile-OH, **AA9**: Fmoc-Orn(ivDde)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the ivDde group according to the general procedure, then the coupling of the amine with **2-bromoacetyl bromide** according to the conditions described above. After the acid cleavage and the purification steps an amount of 19.6 mg (23.62 % yield) of a white solid was obtained, named as UM0140692.

Formula: C₇₆H₉₉N₁₆O₁₂Br; **MW**: 1508.60 g/mol; **HRMS** calcd for: C₇₆H₁₀₀N₁₆O₁₂Br [M+H]⁺ : 1507.6890 ; found 1507.6933 (-0.004 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH: H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.764 min

UM0140693

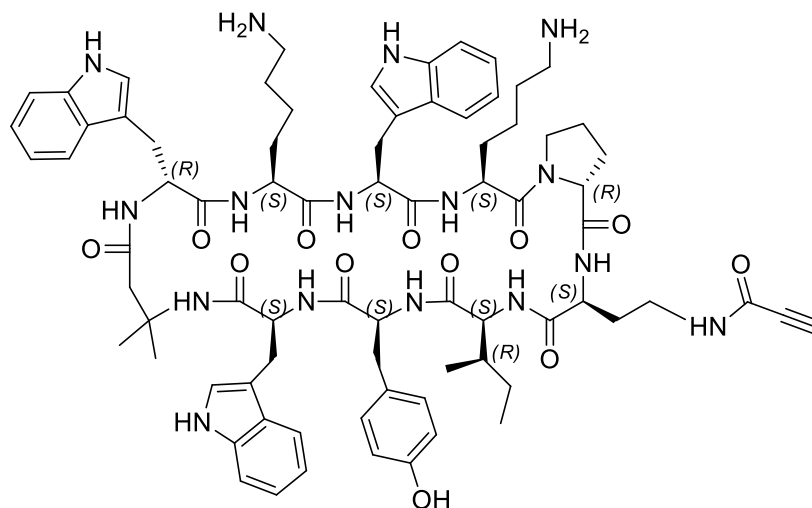
This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA4. Once the 4th coupling was done the resin was separated and the procedure continued with 504 mg of resin (0.218 mmol) until AA8 where it was split another time. The final amount selected to synthesis this macrocycle was 0.055 mmol of resin having the first 8 amino acids coupled.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp (Boc)-OH, **AA5**: Fmoc-Aib-OH, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr (tBu)-OH, **AA8**: : Fmoc-*allo*-Ile-OH, **AA9**: Fmoc-Lys(ivDde)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the ivDde group according to the general procedure, then the coupling of the amine with **2-bromoacetyl bromide** according to the conditions described above. After the acid cleavage and the purification steps an amount of 20.4 mg (24.36 % yield) of a white solid was obtained, named as UM0140693.

Formula: C₇₇H₁₀₁N₁₆O₁₂Br; **MW:** 1522.63 g/mol; **HRMS** calcd for: C₇₇H₁₀₂N₁₆O₁₂Br [M+H]⁺ : 1521.7047 ; found 1521.7015 (-0.003 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH: H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.728 min

UM0140706



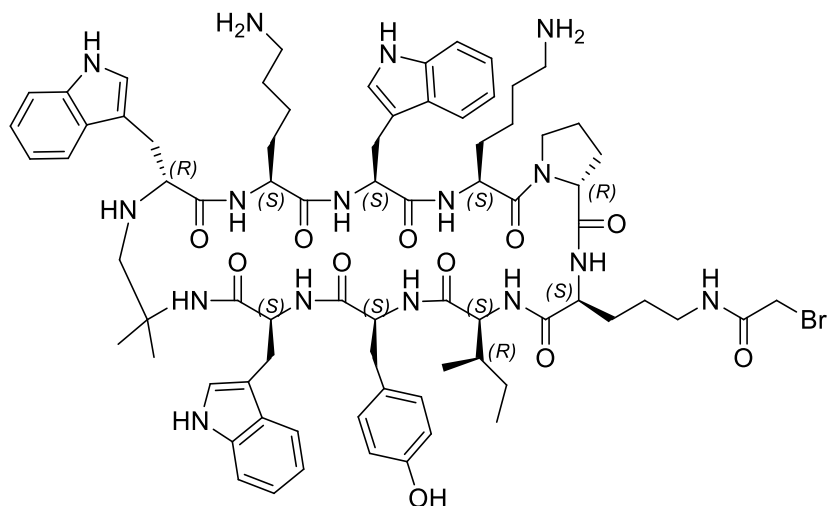
This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA7. Once the 7th coupling was done the resin was separated and the procedure continued with 138 mg of resin (0.037 mmol) until the end of the procedure.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp (Boc)-OH, **AA5**: Fmoc-H-Aib-OH, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr (tBu)-OH, **AA8**: : Fmoc-*allo*-Ile-OH, **AA9**: Fmoc-Dab(ivDde)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the ivDde group according to the general procedure, then the coupling of the amine with **propionic acid** according to the conditions described above. After the acid cleavage and the purification steps an amount of 3.56 mg (6.68 % yield) of a yellow solid was obtained, named as UM0140706.

Formula: C₇₇H₉₈N₁₆O₁₂; **MW**: 1439.70 g/mol; **HRMS** calcd for: C₇₇H₉₉N₁₆O₁₂ [M+H]⁺ : 1439.7628 ; found 1439.7476 (-0.0152 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH: H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.766 min

UM0140707



This macrocycle was synthesized by the following general procedure. The initial amount of anchored resin was (1 g, 0.739 mmol) until the addition of amino acid AA4. Once the 4th coupling was done the resin was separated and the procedure continued with 504 mg of resin (0.218 mmol) until AA8 where it was split another time. The final amount selected to synthesis this macrocycle was 0.055 mmol of resin having the first 8 amino acids coupled.

There is a variance in the synthesis of this macrocycle. Once the AA4 amino acids has been deprotected by piperidine, there is a reductive amination with the aldehyde (Fmoc-Aib-H) to obtain the primary amine.

Fmoc deprotection of AA4: The wet resin (0.218 mmol) was firstly treated with piperidine 20 % / DMF (3.10 ml, 15.52 mmol) and shaken for 5 minutes. The solution was then drained and another treatment of piperidine 20 % / DMF (3.10 ml, 15.52 mmol) was executed, but this time with 20 minutes of shaking. The liquid was removed from the syringe and the resin was washed (1-2 minutes each time, 5.0 ml) with DMF (3 x 5 mL), DMF/MeOH 1:1 (2 x 5 mL) and DMF (4 x 5 mL).

Reductive amination: The resin was washed several times with THF/MeOH (1:1) (3 x 5 ml, 2 minutes). In a vial of 4 ml was prepared a solution with Fmoc-Aib-H (270 mg, 0.873 mmol) and 4

ml of THF/MeOH (1:1). The solution was split in two (vial 1A and vial 1B). In each was added 1 ml of THF/MeOH (1:1) and ACETIC ACID (0.019 ml, 0.327 mmol) and Sodium cyanoborohydride (68.6 mg, 1.091 mmol). After the solution was injected into the syringe and shaken for 90 minutes. The treatment was repeated with the other solution (1B). After the liquid was drained out and the resin was washed with several solvent: (5 ml, 1-2 minutes each times) DMF (3 x 5 mL), DMF/MeOH 1:1 (2 x 5 mL) and DMF (4 x 5 mL). After the procedure continued as usual for the other amino acids.

The amino acid sequence used was the following one: **AA1**: Fmoc-L-Lys-OAll, **AA2**: Fmoc-L-Trp(Boc)-OH, **AA3**: Fmoc-L-Lys(Boc)-OH, **AA4**: Fmoc-D-Trp (Boc)-OH, **AA5**: Fmoc-Aib-H, **AA6**: Fmoc-L-Trp(Boc)-OH, **AA7**: Fmoc-L-Tyr (tBu)-OH, **AA8**: : Fmoc-*allo*-Ile-OH, **AA9**: Fmoc-Orn(ivDde)-OH, **AA10**: Fmoc-D-Pro-OH.

After the macrocyclization step, there was deprotection of the ivDde group according to the general procedure, then the coupling of the amine with **2-bromoacetyl bromide** according to the conditions described above. After the acid cleavage and the purification steps an amount of 7.31 mg (8.89 % yield) of a white solid was obtained, named as UM0140707.

Formula: $C_{76}H_{101}N_{16}O_{11}Br$; **MW**: 1508.60 g/mol; **HRMS** calcd for: $C_{76}H_{102}N_{16}O_{11}Br$ $[M+H]^+$: 1493.7097 ; found 1493.7175 (-0.008 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH: H₂O:TFA (5:95:0.5) B MeOH: H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes, Uv measured at 254 nm): rt = 1.680 min

A.2 Testing of the macrocycles (Biochemistry/Proteomics): Development of covalent macrocycles to mimic the NS1 monobody

Intact mass of macrocycles adducts on RAS isoforms

The isoforms of RAS were expressed and given by Hugo Lavoie from the Lab of Marc Therrien, a special thanks to him. RAS isoforms were expressed with the first 171 amino acids and purified with a HIS-tag. The isoforms used are: HRAS, KRAS, NRAS and HRASC118S.

The procedure was the same for all the isoforms and macrocycles. In a 96-well plate was mixed in the buffer **A** : RAS isoform and macrocycle in a way to obtain a final concentration of 4 μM and 100 μM respectively. The reactions were agitated during 16 hours at room temperature. The next morning 15 μL of the solution was diluted with 15 μL of 4% formic acid solution and then injected on the QExactibe HF Biopharma from Thermo.

Parameter of mass spectrometry: Intact proteins were loaded on C5 precolumn and separated with a home-made C5 column (150 μm ID x 15 cm) made of Jupiter C5, 3 μm , 300 \AA . This separation was possible with a linear gradient of 5-40% of acetonitrile (0.2 % Formic acid) in a 50 min with a flow rate of 600 nL/min. The liquid chromatography was coupled with the Q Exactive Biopharma, with a mass scan range of 500-2000 m/z. The resolution selected was 240 000 and the AGC target values for MS was set to 1e6 (max fill time 200 ms). Computer processing was performed using the software *Protein Deconvolution* from Thermo to identify the exact mass of proteins for the spectras.

Buffer **A**: 1 mM MgCl_2 , 150 mM NaCl and 20 mM HEPES

SDS-PAGE of macrocycles adducts on RAS isoforms

The gels procedure was performed with the same isoforms of RAS expressed by Hugo Lavoie either HRAS, KRAS, NRAS and HRASC118S.

The procedure was the same for all the isoforms and macrocycles. In a 96-well plate was mixed in the buffer **A** : RAS isoform and macrocycle in a way to obtain a final concentration of 4 μM and

100 μ M respectively. The reactions were agitated during 16 hours at room temperature. The next day 12.5 μ L of a **5X loading solution** was added to each sample and was heated at 95°C for 8 minutes. Each condition was loaded on a precast acrylamide gel with gradient 4-12%. The voltage used was 110 V over 1 hour. The gel was revealed with a **Coomassie blue stain** for 3 hours, followed with a **destain solution** until the desired result.

Buffer A: 1 mM MgCl₂, 150 mM NaCl and 20 mM HEPES

5x Loading solution: 3 mL 20 % SDS, 3.75 mL Tris buffer 1 M pH 6.8, 9 mg Bromophenol blue, 1.16 mg DTT, 4.5 mL glycerol, complete to 15 mL with H₂O

Coomassie blue stain: 500 mL methanol, 100 mL glacial AcOH, 1.25 g Brilliant blue G, 400 mL H₂O

Destain solution: 500 mL H₂O, 400 mL MeOH and 100 mL glacial AcOH

Chemicals and materials:

Precast acrylamide gel: 4-12% Criterion XT Bis-Tris Protein Gel from Biorad (3450123)

Bromophenol blue: Sigma-Aldrich, CAS: 115-39-9

DTT: Sigma-Aldrich, CAS: 3483-12-3

Glycerol: Sigma-Aldrich, CAS: 56-81-5

Brilliant blue G: Sigma-Aldrich, CAS: 6104-58-1

Kinetics study of macrocycles

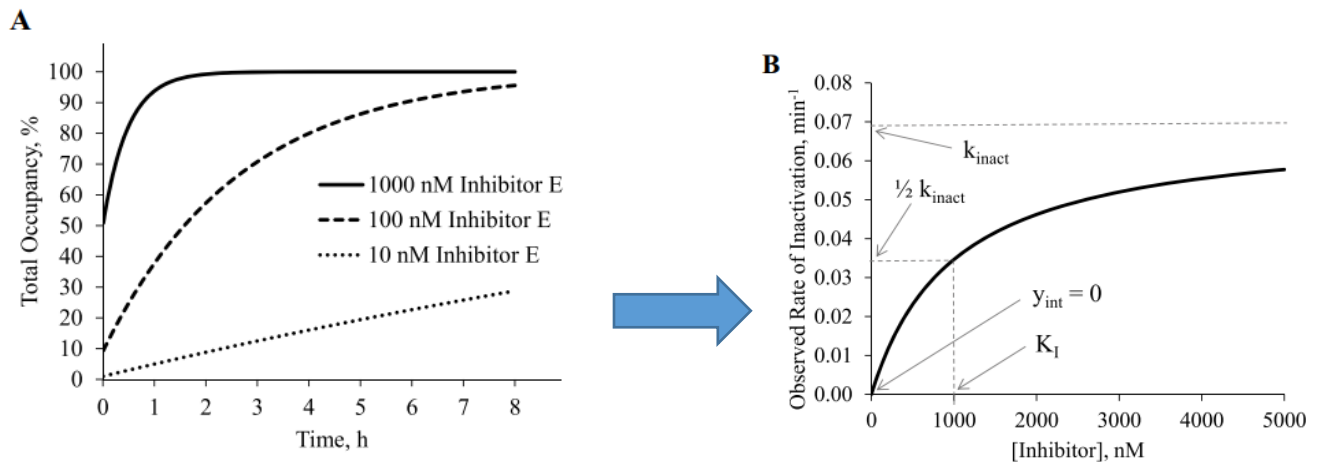
The theory that was followed to perform the kinetic study is described in the paper: Strelow, J. M., A Perspective on the Kinetics of Covalent and Irreversible Inhibition. *SLAS DISCOVERY: Advancing the Science of Drug Discovery* **2016**, 22 (1), 3-20.

The general idea is to be able to have a bioassay that achieves a ratio between the modified protein and the wildtype protein over time. This ratio gives the possibility of obtaining the experimental kinetic constant (k_{obs}) which is proportional to the two constants of the kinetic model (k_{inact} and K_I).

$$\frac{[Modified\ protein]}{[Wildtype\ protein]} \times 100 = \% \text{ total occupancy} = 100(1 - e^{-k_{obs} \times t})$$

$$k_{obs} = \frac{k_{inact}[I]}{K_I + [I]}$$

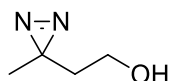
The bioassay that was chosen to obtain the ratios as a function of the time is a gel incubation. As the intensity of a band is proportional to its quantity, it is possible to easily follow the change of the two species over time. The procedure is the same as described in the section: **SDS-PAGE of macrocycles adduct on RAS isoforms**. All incubations were performed in a 96-well plate in a total volume of 50 μ L. HRAS protein was used at a fix concentration of 2 μ M. The kinetics were started in the style of a countdown, so that it all ended at the same time. After that, 12.5 μ L of the 5x loading buffer was added and the samples were heated at 95°C for 8 minutes. Each condition was loaded on a precast acrylamide gel with gradient 4-12%. The voltage used was 110 V over 1 hour. The gel was revealed with a **Coomassie blue stain** for 3 hours, followed with a **destain solution** until the desired result. To obtain the total occupancy at every time, the ratio was calculated with the ImageJ software which measures the light intensity from an image. After which it is possible to obtain a graph of the total occupancy as a function of time for various concentrations of inhibitor (**A**). With the GraphPrism software, it is possible to find the value of the experimental kinetic constant (k_{obs}) for each concentration of inhibitor. Finally, it is possible to trace the relationship between k_{obs} and the concentration of inhibitor to identify the two values that are K_I and k_{inact} (**B**). These values were found with GraphPad.



A.3 Organic synthesis section for the second project: Target identification of the UM171 compound

The experimental procedure for some probes comes from the report submitted to the chemistry department by Moïse Perreault: Perreault, M. *Déconvolution de la cible de UM171 par approche chémoprotéomique*, Université de Montréal, August 26th 2019, CHM 3552

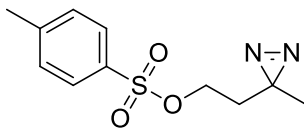
2-(3-methyl-3H-diazirin-3-yl)ethanol (**33**)



4-hydroxybutan-2-one (2,026 ml, 22,70 mmol) was added slowly into stirring liquid NH₃ (30 mL) at -78 °C. After stirring at -78 °C for 5h, a solution of (aminooxy)sulfonic acid (2,82 g, 24,97 mmol) in MeOH (20 ml) was added slowly to the reaction mixture. Then the stirred reaction mixture was allowed to warm up to room temperature overnight. The resulting white precipitate was filtered off and the volume of the filtrate was reduced to 50 mL. The yellow filtrate was cooled to 0 °C; methanol (50 mL) and Et₃N (3,16 ml, 22,70 mmol) were added followed by slow addition of iodine until mixture turns to a persistent brown orange. The reaction mixture was allowed to warm up to room temperature stirred for 2 h. Volume minimized under reduced pressure (80 mbar), then diluted with brine (150 mL) followed by extraction with Et₂O (3x50 mL), dried over anhydrous Na₂SO₄, filtered and concentrated. The diazine was isolated by horizontal distillation with a Kugelrohr apparatus (40 to 70°C at 3 mbar). Diethyl ether was used to recuperate the product, which was concentrated reduced pressure (70 mbar) to give 2-(3-methyl-3H-diazirin-3-yl)ethanol **33** (1,155 g, 11,54 mmol, 50,8 % yield) as a pale yellow oil

Formula: C₄H₈N₂O; **MW:** 100.12 g/mol; **¹H NMR** (CDCl₃, 400MHz): δ = 3.55 (t, J=6.3 Hz, 2 H), 1.66 (t, J=6.3 Hz, 2 H), 1.39 (br. s., 1 H), 1.09 ppm (s, 3 H), **¹³C NMR** (CDCl₃, 101MHz): δ = 57.7, 36.9, 24.2, 20.2 ppm.

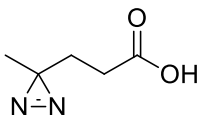
2-(3-methyl-3H-diazirin-3-yl)ethyl 4-methylbenzenesulfonate (34)



2-(3-methyl-3H-diazirin-3-yl)ethanol (520 mg, 5,19 mmol) was dissolved in DCM (3582 μ l) at 0 °C. After addition of pyridine (1260 μ l, 15,58 mmol), Et₃N (2172 μ l, 15,58 mmol) and DMAP (12,69 mg, 0,104 mmol), Tosyl-Cl (1980 mg, 10,39 mmol) dissolved in DCM (1.45M, 6.9mL) was added dropwise for 5 min. Then the reaction was stirred at room temperature overnight. Enough 1 N HCl was added to quench the reaction (30mL), then extracted with Et₂O (2x30 mL), evaporated under reduced pressure and the compound is isolated by ISCO on a 12g SiO₂ column (Hex / EtOAc - 20:1). Fractions are gathered and evaporated under reduced pressure to give 2-(3-methyl-3H-diazirin-3-yl)ethyl 4-methylbenzenesulfonate 34 (649 mg, 2,55 mmol, 49,1 % yield) as a yellow oil.

Formula: C₁₁H₁₄N₂O₃S; **MW:** 254.31 g/mol; **¹H NMR** (CDCl₃, 400MHz): δ = 7.81 - 7.85 (m, J=8.2 Hz, 2 H), 7.35 - 7.40 (m, J=8.6 Hz, 2 H), 3.97 (t, J=6.5 Hz, 2 H), 2.47 (s, 3 H), 1.68 (t, J=6.3 Hz, 2 H), 1.02 ppm (s, 3 H), **¹³C NMR** (CDCl₃, 101MHz): δ = 147.7, 135.4, 132.6, 130.6, 67.7, 36.8, 26.0, 24.3, 22.4 ppm.

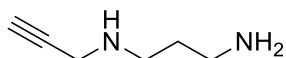
3-(3-methyl-3H-diazirin-3-yl)propanoic acid (**38**)



A solution of 4-oxopentanoic acid (0,437 ml, 4,31 mmol) in MeOH (1.0 mL) was cooled in an ice bath under an argon atmosphere. To this solution was added NH₃ 7N in MeOH (5 ml, 35,0 mmol) via syringe. The mixture was allowed to stir for 3h at 0 °C, then (aminooxy)sulfonic acid (0,536 g, 4,74 mmol) in MeOH (1.5 mL) was added. The reaction mixture was stirred for 16h at RT. The excess ammonia was removed by blowing air through the suspension for 1 h using a glass pipette. The suspension was filtered, and the filtrate concentrated under reduced pressure. The solid mass was dissolved MeOH (0.5 mL) and cooled in an ice bath. Et₃N (0.9 mL, 6.5 mmol) was added, stirred for 5 minutes. Iodine (1.0 g, 7.9 mmol) was added until the orange-brown color remained. The methanol is evaporated under reduced pressure. The oil was diluted with EtOAc (3.0 mL), washed with 1 N HCl (3 mL), a 10% Na₂S₂O₃ solution (3 mL), brine (3 mL), and dried over anhydrous Na₂SO₄. The product is isolated by ISCO on a 24g SiO₂ column (A: Hex / B: EtOAc-Hex-AcOH (90-8-2%)) with a gradient of B from 0 % to 20% over 17 minutes. Product was revealed with bromocresol green stain. Fractions that revealed were gathered, evaporated under reduced pressure to give 3-(3-methyl-3H-diazirin-3-yl)propanoic acid **38** (494mg, 3,86 mmol, 90 % yield) as a pale yellow oil

Formula: C₅H₈N₂O₂; **MW:** 128.13 g/mol; ¹H NMR (CDCl₃, 400MHz): δ = 2.25 (t, J=7.6 Hz, 2 H), 1.74 (t, J=7.6 Hz, 2 H), 1.06 ppm (s, 3 H), ¹³C NMR (CDCl₃, 101MHz): δ = 178.7, 29.3, 28.5, 25.0, 19.6 ppm.

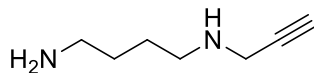
N1-(prop-2-yn-1-yl)propane-1,3-diamine (**39**)



A solution of 3-chloroprop-1-yne (1,428 ml, 13,82 mmol) and DBU (2,083 ml, 13,82 mmol) in toluene (8mL) is added to another solution (propane-1,3-diamine (11,64 ml, 138 mmol) in toluene (8mL)) at 0°C. The reaction is warmed at room temperature and stirred for 4 hours. The compound is isolated by distillation with a Kugelrohr apparatus (35°C to 50°C at 3.5 mbar for 30 minutes to flush impurities; compound isolated between 70 to 140°C at 3.5 mbar). The compound is further isolated by ISCO with a 24g SiO₂ column by solid handling (A: DCM / B: DCM-MeOH-NH₄OH (78-20-2%)) with a gradient of B from 0 % to 100% over 20 minutes. Fractions are gathered to give N1-(prop-2-yn-1-yl)propane-1,3-diamine **39** (477 mg, 4,25 mmol, 30,8 % yield) as a yellow oil.

Formula: C₆H₁₂N₂; **MW:** 112.17 g/mol; **¹H NMR** (CDCl₃, 400MHz): δ = 3.44 (d, J=2.3 Hz, 2 H), 2.79 (q, J=7.0 Hz, 4 H), 2.22 (t, J=2.5 Hz, 1 H), 1.74 - 1.83 (m, 1 H), 1.65 (quin, J=6.8 Hz, 2 H), 1.29 ppm (br. s., 2 H), **¹³C NMR** (CDCl₃, 101MHz): δ = 82.2, 71.2, 46.5, 40.4, 38.2, 33.5 ppm

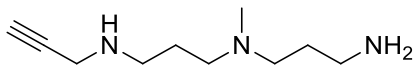
N1-(prop-2-yn-1-yl)butane-1,4-diamine (**40**)



A solution of butane-1,4-diamine (1.389 ml, 13.8 mmol) and DBU (2,083 ml, 13.8 mmol) in toluene (8mL) is added to another solution (butane-1,4-diamine (1.389 ml, 13.8 mmol) in toluene (8mL)) at 0°C. The reaction is warmed at room temperature and stirred for 4 hours. The compound is isolated by distillation with a Kugelrohr apparatus (35°C to 50°C at 3.5 mbar for 30 minutes to flush impurities; compound isolated between 70 to 140°C at 3.5 mbar). The compound is further isolated by ISCO with a 24g SiO₂ column by solid handling (A: DCM / B: DCM-MeOH- NH₄OH (78-20-2%)) with a gradient of B from 0 % to 100% over 20 minutes. Fractions are gathered to give N1-(prop-2-yn-1-yl)butane-1,4-diamine **40** (279 mg, 2,211 mmol, 16,00 % yield) as a yellow oil.

Formula: C₇H₁₄N₂; **MW:** 126.20 g/mol; **¹H NMR** (CDCl₃, 400MHz): δ = 3.44 (d, J=2.7 Hz, 2 H), 2.65 - 2.78 (m, 4 H), 2.22 (t, J=2.5 Hz, 1 H), 1.48 - 1.57 (m, 4 H), 1.40 ppm (br. s, 3 H), **¹³C NMR** (CDCl₃, 101MHz): δ = 82.2, 71.2, 48.5, 42.0, 38.1, 31.4, 27.2 ppm.

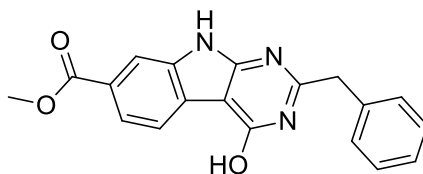
N1-(3-aminopropyl)-N1-methyl-N3-(prop-2-yn-1-yl)propane-1,3-diamine (41)



A solution of N1-(3-aminopropyl)-N1-methylpropane-1,3-diamine (2.228 ml, 13.8 mmol) and DBU (2,083 ml, 13.8 mmol) in toluene (8mL) is added to another solution N1-(3-aminopropyl)-N1-methylpropane-1,3-diamine (2.228 ml, 13.8 mmol) in toluene (8mL)) at 0°C. The reaction is warmed at room temperature and stirred for 4 hours. The compound is isolated by distillation with a Kugelrohr apparatus (35°C to 50°C at 3.5 mbar for 30 minutes to flush impurities; compound stayed in the round-bottom flask). The compound is further isolated by ISCO with a 24g SiO₂ column by solid handing (A: DCM / B: DCM-MeOH- NH₄OH (78-20-2%)) with a gradient of B from 0 % to 100% over 20 minutes. Fractions are gathered to give N1-(3-aminopropyl)-N1-methyl-N3-(prop-2-yn-1-yl)propane-1,3-diamine **41** (1,012 g, 5,52 mmol, 40,0 % yield) as an orange oil.

Formula: C₁₀H₂₁N₃; **MW:** 183.29 g/mol; **¹H NMR** (CDCl₃, 400MHz): δ = 3.41 - 3.49 (m, 1 H), 3.39 (d, J=2.3 Hz, 1 H), 2.71 (td, J=6.7, 2.0 Hz, 3 H), 2.37 (q, J=6.8 Hz, 4 H), 2.15 - 2.22 (m, 4 H), 1.57 - 1.66 (m, 4 H), 1.41 ppm (br. s, 2 H), **¹³C NMR** (CDCl₃, 101MHz): δ = 82.0, 70.8, 55.9, 55.3, 47.1, 41.9, 40.3, 37.9, 30.8, 27.0 ppm.

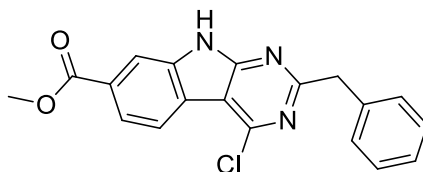
Methyl 2-benzyl-4-hydroxy-9H-pyrimido[4,5-b]indole-7-carboxylate (43)



In a 10 mL microwave vial was added 3-carbamoyl-6-(methoxycarbonyl)-1H-indol-2-aminium chloride (500 mg, 1,854 mmol), methyl 2-phenylacetate (392 μ l, 2,78 mmol) and sodium methanolate 5.4 M (1030 μ l, 5,56 mmol) in MeOH (5618 μ l) to give a brown suspension. After heating in the microwave (140°C, 45min), sodium methanolate 5.4 M (618 μ l, 3,34 mmol) and methyl 2-phenylacetate (261 μ l, 1,854 mmol) were added to complete the reaction in the microwave (140°C, 45min). AcOH (743 μ l, 12,98 mmol) was added and the black solution turns to a light brown suspension. Stirred the resulting suspension at room temperature overnight. Filtered to afford methyl 2-benzyl-4-hydroxy-9H-pyrimido[4,5-b]indole-7-carboxylate 43 (436 mg, 1,308 mmol, 70,5 % yield) as a purple-brown solid

Formula: C₁₉H₁₅N₃O₃; **MW:** 333.54 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 12.49 (br. s., 1 H), 7.99 - 8.03 (m, 2 H), 7.83 (dd, J=8.2, 2.0 Hz, 1 H), 7.39 (dd, J=8.2, 1.6 Hz, 2 H), 7.33 (td, J=7.4, 1.2 Hz, 2 H), 7.25 (tt, J=7.4, 2.3 Hz, 1 H), 4.03 (s, 2 H), 3.87 (s, 3 H), 1.86 ppm (s, 1 H), **¹³C NMR** (DMSO-d₆, 101MHz): δ = 133.4, 114.6, 98.6, 96.9, 95.0, 92.7, 91.5, 91.0, 90.7, 90.3, 90.1, 89.3, 33.7, 23.1, 3.4 ppm.

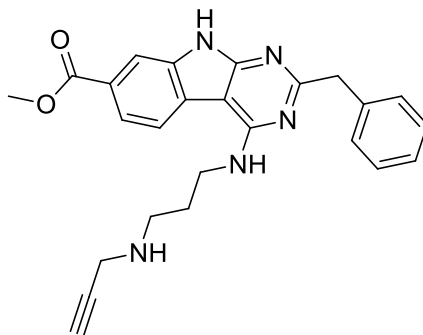
Methyl 2-benzyl-4-chloro-9H-pyrimido[4,5-b]indole-7-carboxylate (44)



In a 25 mL round-bottom flask was added methyl 2-benzyl-4-hydroxy-9H-pyrimido[4,5-b]indole-7-carboxylate (420 mg, 1,260 mmol) in POCl₃ (10,200 ml, 109 mmol) to give a brown suspension. The reaction was heated at 90°C overnight, then concentrated under reduced pressure to give a brown foam. The foam was suspended in sat. NaHCO₃ solution (10 mL) and stirred for 1 hour. The solid was collected on Buchner, washed with water (3x2 mL) and the compound was dried under high vacuum to give methyl 2-benzyl-4-chloro-9H-pyrimido[4,5-b]indole-7-carboxylate 44 (381 mg, 1,083 mmol, 83 % yield) as a brittle brown solid.

Formula: C₁₉H₁₄N₃O₂Cl; **MW:** 351.79 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 12.94 (br. s., 1 H), 8.30 (d, J=8.2 Hz, 1 H), 8.07 - 8.13 (m, 1 H), 7.96 (dd, J=8.2, 1.6 Hz, 1 H), 7.33 (dt, J=13.5, 7.0 Hz, 4 H), 7.22 (tt, J=7.0, 2.0 Hz, 1 H), 4.30 (s, 2 H), 3.91 ppm (s, 3 H), **¹³C NMR** (DMSO-d₆, 101MHz): δ = 166.6, 166.1, 157.6, 152.5, 138.1, 138.0, 129.2, 128.5, 128.4, 126.4, 122.1, 122.0, 121.5, 113.1, 108.4, 52.4, 44.9 ppm.

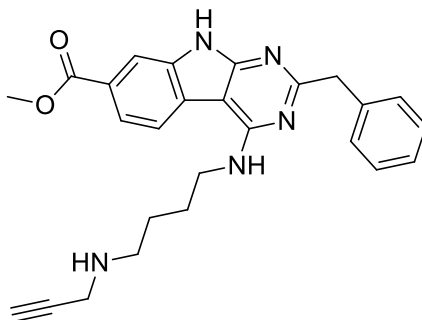
Methyl 2-benzyl-4-((3-(prop-2-yn-1-ylamino)propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (45)



A mixture of methyl 2-benzyl-4-chloro-9H-pyrimido[4,5-b]indole-7-carboxylate (120 mg, 0,341 mmol), N1-(prop-2-yn-1-yl)propane-1,3-diamine (963 μ l, 2,047 mmol) and Et₃N (95 μ l, 0,682 mmol) in MeOH (1503 μ l) was heated in the microwave (140°C, 40min). Solvent was removed under reduced pressure and the product was isolated by ISCO on a 4g SiO₂ column (A: DCM / B: DCM-MeOH-NH₄OH (78-20-2%)) with a gradient of B from 0 % to 100% over 16 minutes. Fractions are collected to give methyl 2-benzyl-4-((3-(prop-2-yn-1-ylamino)propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate **45** (45 mg, 0,105 mmol, 30,9 % yield) as a yellowish solid.

Formula: C₂₅H₂₅N₅O₂; **MW:** 427.5 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 12.05 (s, 1 H), 8.31 (d, J=8.2 Hz, 1 H), 7.98 (d, J=1.6 Hz, 1 H), 7.82 (dd, J=8.6, 1.6 Hz, 1 H), 7.60 (t, J=5.7 Hz, 1 H), 7.37 (d, J=6.7 Hz, 2 H), 7.24 - 7.32 (m, 2 H), 7.18 (tt, J=7.0, 2.0 Hz, 1 H), 4.04 (s, 2 H), 3.88 (s, 3 H), 3.66 (q, J=6.3 Hz, 2 H), 3.35 (d, J=2.3 Hz, 2 H), 3.04 (t, J=2.3 Hz, 1 H), 2.64 - 2.71 (m, 2 H), 2.33 (quin, J=2.0 Hz, 1 H), 1.78 ppm (quin, J=6.7 Hz, 2 H), **¹³C NMR** (DMSO-d₆, 101MHz): δ = 195.3, 166.9, 166.9, 157.6, 157.1, 139.4, 135.8, 129.3, 128.3, 126.2, 124.8, 123.8, 121.0, 120.4, 112.1, 93.5, 83.2, 73.7, 52.2, 46.3, 46.0, 37.6, 28.9 ppm.

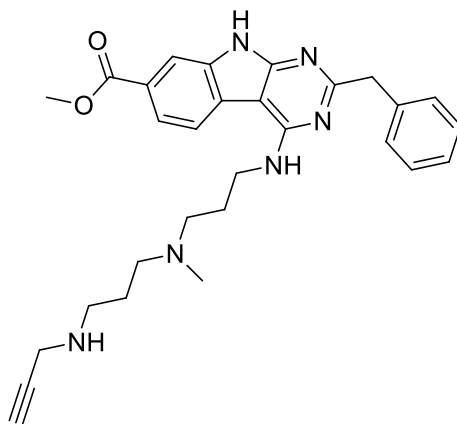
Methyl 2-benzyl-4-((4-(prop-2-yn-1-ylamino)butyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (46) (UM0142947)



A mixture of methyl 2-benzyl-4-chloro-9H-pyrimido[4,5-b]indole-7-carboxylate (100 mg, 0,284 mmol), of N1-(prop-2-yn-1-yl)butane-1,4-diamine (71,7 mg, 0,569 mmol) and TRIETHYLAMINE (79 μ l, 0,569 mmol) in MeOH (1252 μ l) was heated in the microwave for 40min at 140°C. Solvent was removed under reduced pressure and the product was isolated by ISCO (DCM / DCM-MeOH-NH₄OH (78-20-2%)) on a 4g SiO₂ column. Fractions are collected to give methyl 2-benzyl-4-((4-(prop-2-yn-1-ylamino)butyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate **46** (24mg, 0,054 mmol, 15,94 % yield) as a yellowish solid.

Formula: C₂₆H₂₇N₅O₂; **MW:** 441.52 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 11.98 - 12.07 (m, 1 H), 8.32 - 8.39 (m, 1 H), 7.95 - 8.01 (m, 1 H), 7.76 - 7.83 (m, 1 H), 7.32 - 7.41 (m, 2 H), 7.22 - 7.32 (m, 2 H), 7.13 - 7.21 (m, 1 H), 4.03 (s, 2 H), 3.88 (s, 3 H), 3.55 - 3.67 (m, 2 H), 3.29 (d, J=2.3 Hz, 2 H), 3.02 (t, J=2.3 Hz, 1 H), 2.58 (t, J=7.0 Hz, 2 H), 1.56 - 1.74 (m, 2 H), 1.39 - 1.53 ppm (m, 2 H), **¹³C NMR** (DMSO-d₆, 101MHz): δ = 166.7, 166.6, 157.5, 156.9, 139.2, 135.6, 129.1, 128.0, 126.0, 124.6, 123.6, 120.6, 120.4, 111.9, 93.3, 82.8, 79.0, 73.6, 52.0, 47.7, 45.8, 37.4, 27.1, 26.5 ppm.

Methyl 2-benzyl-4-((3-(methyl(3-(prop-2-yn-1-ylamino)propyl)amino)propyl)amino)propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (47)

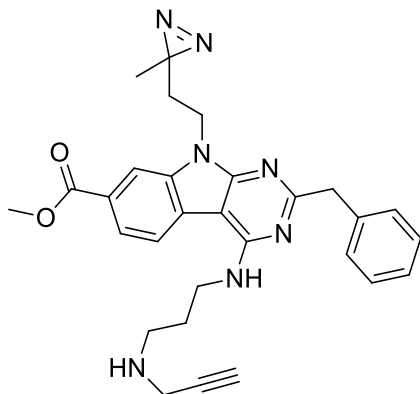


A mixture of methyl 2-benzyl-4-chloro-9H-pyrimido[4,5-b]indole-7-carboxylate (120 mg, 0,341 mmol), of N1-(3-aminopropyl)-N1-methyl-N3-(prop-2-yn-1-yl)propane-1,3-diamine (247 μ l, 0,682 mmol) and TRIETHYLAMINE (95 μ l, 0,682 mmol) in MeOH (1503 μ l) was heated in the microwave for 40min at 140°C. Solvent was removed under reduced pressure and the product was isolated by ISCO (DCM / DCM-MeOH-NH₄OH (78-20-2%)) on a 4g SiO₂ column. Fraction was collected to give methyl 2-benzyl-4-((3-(methyl(3-(prop-2-yn-1-ylamino)propyl)amino)propyl)amino)propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate **47** (36 mg, 0,072 mmol, 21,17 % yield) as a yellowish solid.

Formula: C₂₉H₃₄N₆O₂; **MW:** 498.62 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 12.08 (br. s., 1 H), 8.35 (d, J=8.6 Hz, 1 H), 7.95 - 7.98 (m, 1 H), 7.78 (dd, J=8.4, 1.4 Hz, 1 H), 7.50 (t, J=5.5 Hz, 1 H), 7.29 - 7.36 (m, 2 H), 7.22 - 7.28 (m, 2 H), 7.11 - 7.20 (m, 1 H), 4.02 (s, 2 H), 3.84 (s, 3 H), 3.62 (d, J=6.3 Hz, 2 H), 3.44 - 3.48 (m, 1 H), 3.40 (br. s., 2 H), 3.18 (s, 1 H), 2.69 - 2.82 (m, 3 H), 2.60 - 2.68 (m, 2 H), 1.88 - 1.97 (m, 2 H), 1.87 (s, 2 H), 1.66 - 1.74 ppm (m, 2 H), **¹³C NMR** (DMSO-d₆, 101MHz): δ = 172.0, 166.7, 157.4, 156.8, 139.2, 135.6, 129.1, 128.1, 126.0, 124.7, 123.5, 120.7, 120.4, 112.0, 93.4, 79.3, 79.0, 78.6, 75.2, 54.2, 54.0, 52.0, 45.8, 45.1, 38.2, 36.6, 25.0, 23.5, 21.1 ppm.

Methyl

2-benzyl-9-(2-(3-methyl-3H-diazirin-3-yl)ethyl)-4-((3-(prop-2-yn-1-ylamino)propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (UM0142941)

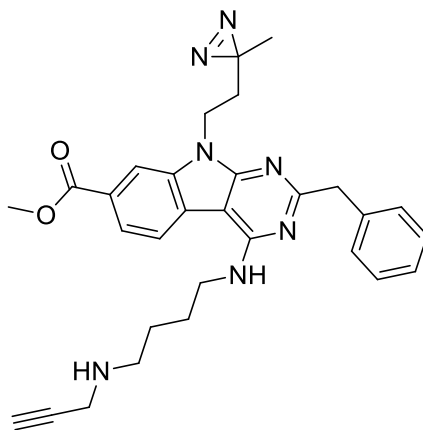


To a solution of 2-(3-methyl-3H-diazirin-3-yl)ethyl 4-methylbenzenesulfonate (103 μ l, 0,051 mmol) in DMF (94 μ l) were added K₂CO₃ (19,40 mg, 0,140 mmol) and methyl 2-benzyl-4-((3-(prop-2-yn-1-ylamino)propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (20 mg, 0,047 mmol) was stirred overnight (80°C). The compound was isolated by HPLC prep (A: MeOH:H₂O 5:95, AcOH 0.1% / B: MeOH:H₂O 95:5, AcOH 0.1 %) with a gradient of B from 25 % to 100 % over 15.5 minutes. Fractions are concentrated under reduced pressure to give Methyl 2-benzyl-4-((3-((2-(3-methyl-3H-diazirin-3-yl)ethyl)(prop-2-yn-1-yl)amino)propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (6,98mg, 0,014 mmol, 29,3 % yield) as a yellowish solid.

Formula: C₂₉H₃₁N₇O₂; **MW:** 509.60 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 8.36 (d, J=8.2 Hz, 1 H), 8.11 - 8.14 (m, 1 H), 7.88 (dd, J=8.4, 1.4 Hz, 1 H), 7.69 (s, 1 H), 7.38 (d, J=7.0 Hz, 2 H), 7.26 (t, J=7.4 Hz, 2 H), 7.18 (d, J=7.0 Hz, 1 H), 4.49 (t, J=6.7 Hz, 2 H), 4.08 (s, 2 H), 3.89 (s, 3 H), 3.67 (q, J=6.3 Hz, 2 H), 3.36 (d, J=2.3 Hz, 2 H), 3.04 (t, J=2.3 Hz, 1 H), 2.68 (t, J=6.5 Hz, 2 H), 1.74 - 1.83 (m, J=6.8, 6.8 Hz, 4 H), 0.93 ppm (s, 3 H), **¹³C NMR** (DMSO-d₆, 101MHz): δ = 166.7, 166.6, 156.8, 139.1, 135.8, 129.1, 128.0, 126.0, 124.8, 123.6, 121.2, 120.4, 110.4, 93.1, 82.9, 79.3, 78.6, 73.6, 52.1, 46.1, 45.8, 37.4, 36.5, 33.4, 28.6, 24.5, 18.9 ppm **HRMS** calcd for : C₂₉H₃₂N₇O₂ [M+H]⁺ : 510.2617 ; found 510.2577 (-0.004 m/z).

Methyl

2-benzyl-9-(2-(3-methyl-3H-diazirin-3-yl)ethyl)-4-((4-(prop-2-yn-1-ylamino)butyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (UM0142943)

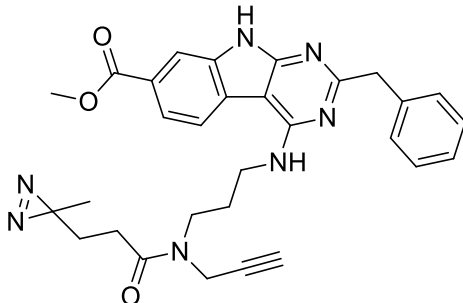


To a solution of 2-(3-methyl-3H-diazirin-3-yl)ethyl 4-methylbenzenesulfonate (21,54 mg, 0,085 mmol) in DMF (154 μ l) were added K_2CO_3 (31,9 mg, 0,231 mmol) and methyl 2-benzyl-4-((4-(prop-2-yn-1-ylamino)butyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (34 mg, 0,077 mmol) was stirred overnight (80°C). The compound was isolated by HPLC prep (A: MeOH:H₂O 5:95, AcOH 0.1% / B: MeOH:H₂O 95:5, AcOH 0.1%) with a gradient of B from 25% to 100% over 15.5 minutes. Fractions are concentrated under reduced pressure to give Methyl 2-benzyl-9-(2-(3-methyl-3H-diazirin-3-yl)ethyl)-4-((4-(prop-2-yn-1-ylamino)butyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (2.5 mg, 0,014 mmol, 6.2% yield) as a yellowish solid.

Formula: C₃₀H₃₃N₇O₂; **MW:** 523.63 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 8.41 (d, J=8.6 Hz, 1 H), 8.10 - 8.14 (m, 1 H), 7.86 (dd, J=8.4, 1.4 Hz, 1 H), 7.46 (t, J=5.5 Hz, 1 H), 7.38 (d, J=7.4 Hz, 2 H), 7.26 (t, J=7.4 Hz, 2 H), 7.15 - 7.20 (m, 1 H), 4.49 (t, J=6.7 Hz, 2 H), 4.07 (s, 2 H), 3.89 (s, 3 H), 3.60 (dd, J=13.1, 6.5 Hz, 3 H), 3.29 (d, J=2.0 Hz, 2 H), 3.02 (t, J=2.3 Hz, 1 H), 2.55 - 2.61 (m, 2 H), 1.78 (t, J=6.7 Hz, 2 H), 1.65 (dd, J=14.3, 7.2 Hz, 2 H), 1.42 - 1.50 (m, 2H), 0.94 ppm (s, 3 H), **¹³C NMR** (DMSO-d₆, 101MHz): δ = 166.8, 166.6, 156.9, 156.9, 139.2, 135.8, 129.1, 128.1, 126.1, 124.8, 123.6, 121.2, 120.6, 110.4, 93.2, 83.2, 73.4, 52.1, 47.8, 45.9, 37.4, 36.6, 33.4, 27.2, 26.7, 24.6, 18.9 ppm **HRMS** calcd for : C₃₀H₃₄N₇O₂ [M+H]⁺ : 524.2774 ; found 524.2926 (-0.0152 m/z).

Methyl

2-benzyl-4-((3-(3-(3-methyl-3H-diazirin-3-yl)-N-(prop-2-yn-1-yl)propanamido)propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (UM0142942)

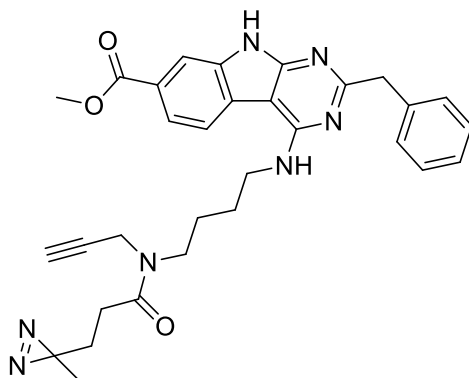


HATU (23,13 mg, 0,061 mmol) was added to a solution of 3-(3-(3-methyl-3H-diazirin-3-yl)propanoic acid (7,19 mg, 0,056 mmol) in DMF (440 μ l). 5 minutes later, methyl 2-benzyl-4-((3-(3-(3-methyl-3H-diazirin-3-yl)-N-(prop-2-yn-1-yl)propanamido) propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (8,69mg, 0,016 mmol, 34,6 % yield) in DMF (1240 μ l) was added, followed by DIPEA (18,79 μ l, 0,108 mmol). After letting the reaction go overnight, the compound was isolated by HPLC prep. (A: MeOH:H₂O 5:95, AcOH 0.1% / B: MeOH:H₂O 95:5, AcOH 0.1 %) with a gradient of B from 15 % to 100 % over 13.5 minutes. Fractions are collected to give methyl 2-benzyl-4-((3-(3-(3-methyl-3H-diazirin-3-yl)-N-(prop-2-yn-1-yl)propanamido)propyl) amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (8,69mg, 0,016 mmol, 34,6 % yield) as a pale yellow solid

Formula: C₃₀H₃₁N₇O₃; **MW:** 537.61 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 12.25 (d, J=16.8 Hz, 1 H), 8.42 (dd, J=8.2, 3.1 Hz, 1 H), 8.02 (d, J=5.1 Hz, 1 H), 7.84 (d, J=8.2 Hz, 1 H), 7.64 (br. s., 1 H), 7.35 - 7.41 (m, 2 H), 7.25 - 7.32 (m, 2 H), 7.16 - 7.23 (m, 1 H), 4.19 (d, J=2.0 Hz, 1 H), 4.13 (d, J=2.0 Hz, 1 H), 4.10 (s, 2 H), 3.88 (s, 3 H), 3.59 - 3.66 (m, 2 H), 3.39 - 3.48 (m, 2 H), 3.34 (t, J=2.3 Hz, 1 H), 2.33 (t, J=7.4 Hz, 1 H), 2.10 (t, J=7.6 Hz, 1 H), 1.93 - 2.02 (m, 1 H), 1.81 - 1.91 (m, 1 H), 1.63 (t, J=7.2 Hz, 1 H), 1.45 (t, J=7.4 Hz, 1 H), 0.98 - 1.04 (m, 1 H), 0.79 ppm (s, 1 H), **¹³C NMR** (DMSO-d₆, 101MHz): δ = 173.0, 171.2, 170.0, 166.6, 166.6, 147.2, 135.7, 129.1, 128.2, 126.2, 123.5, 120.4, 93.5, 93.5, 79.9, 74.8, 52.1, 44.5, 43.6, 36.9, 33.7, 29.3, 29.1, 26.8, 26.4, 25.7, 19.5, 19.1 ppm **HRMS** calcd for : C₃₀H₃₂N₇O₃ [M+H]⁺ : 538.2567 ; found 538.2564 (-0.0003 m/z).

Methyl

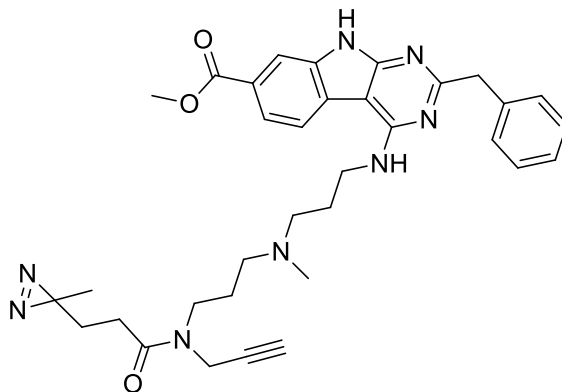
2-benzyl-4-((4-(3-(3-methyl-3H-diazirin-3-yl)-N-(prop-2-yn-1-yl)propanamido)butyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (UM142944)



HATU (33,6 mg, 0,088 mmol) was added to a solution of 3-(3-methyl-3H-diazirin-3-yl)propanoic acid (10,45 mg, 0,082 mmol) in DMF (440 μ l). 5 minutes later, methyl 2-benzyl-4-((4-(prop-2-yn-1-ylamino)butyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (30 mg, 0,068 mmol) in DMF (1240 μ l) was added, followed by Hunig's base (27,3 μ l, 0,156 mmol). After letting the reaction go overnight, the compound was directly purified by HPLC prep (2 injections). Solvent A: MeOH:H₂O 5:95, AcOH 0.1% Solvent B: MeOH:H₂O 95:5, AcOH 0.1%. Gradient 25% to 100% over 15.5 minutes. Fractions are collected to give methyl 2-benzyl-4-((4-(3-(3-methyl-3H-diazirin-3-yl)-N-(prop-2-yn-1-yl)propanamido)butyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (9,1mg, 0,016 mmol, 24,28% yield) as a pale yellow solid.

Formula: C₃₁H₃₃N₇O₃; **MW:** 551.64 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 12.05 (br. s., 2 H), 8.36 (dd, J=8.2, 2.3 Hz, 1 H), 7.99 (s, 1 H), 7.81 (d, J=8.2 Hz, 1 H), 7.37 (d, J=7.8 Hz, 3 H), 7.28 (t, J=7.2 Hz, 2 H), 7.12 - 7.22 (m, 1 H), 4.12 - 4.16 (m, 1 H), 4.09 (d, J=2.3 Hz, 1 H), 4.04 (d, J=4.7 Hz, 2 H), 3.84 - 3.91 (m, 3 H), 3.63 (dd, J=12.9, 5.9 Hz, 3 H), 3.28 (t, J=2.0 Hz, 1 H), 3.08 (t, J=2.2 Hz, 1 H), 2.28 (t, J=7.4 Hz, 1 H), 2.12 (t, J=7.2 Hz, 1 H), 1.47 - 1.71 (m, 7 H), 0.99 (s, 2 H), 0.89 ppm (s, 2 H), **¹³C NMR** (DMSO-d₆, 101MHz): δ = 195.1, 170.6, 170.1, 166.7, 166.7, 157.5, 157.0, 156.9, 139.3, 135.7, 129.2, 129.1, 128.1, 126.0, 123.6, 120.7, 120.4, 111.9, 93.3, 80.3, 80.1, 74.6, 73.5, 52.1, 45.8, 33.9, 29.2, 26.3, 25.6, 25.4, 19.4 ppm **HRMS** calcd for : C₃₁H₃₄N₇O₃ [M+H]⁺ : 552.2723 ; found 552.2894 (-0.0171 m/z).

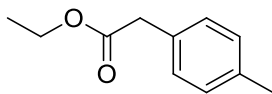
Methyl 2-benzyl-4-((3-(methyl(3-(3-(3-methyl-3H-diazirin-3-yl)-N-(prop-2-yn-1-yl)propanamido)propyl)amino)propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate
(UM0142945)



HATU (19,83 mg, 0,052 mmol) was added to a solution of 3-(3-methyl-3H-diazirin-3-yl)propanoic acid (6,17 mg, 0,048 mmol) in DMF (440 μ l). 5 minutes later, methyl 2-benzyl-4-((3-(methyl(3-(prop-2-yn-1-ylamino)propyl)amino)propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (20mg, 0,040 mmol) in DMF (1240 μ l) was added, followed by Hunig's Base (16,11 μ l, 0,092 mmol). After letting the reaction go overnight, the compound was directly purified by HPLC prep (2 injections). Solvent A: MeOH:H₂O 5:95, AcOH 0.1% Solvent B: MeOH:H₂O 95:5, AcOH 0.1 %. Gradient 25 % to 90 % over 12 minutes. Tubes were combined and concentrated under reduced pressure to give methyl 2-benzyl-4-((3-(methyl(3-(3-(3-methyl-3H-diazirin-3-yl)-N-(prop-2-yn-1-yl)propanamido)propyl)amino)propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (2,05mg, 3,37 μ mol, 8,40 % yield) as white solid.

Formula: C₃₄H₄₀N₈O₃; **MW:** 608.73 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 12.04 (s, 1 H), 8.27 (t, J=7.4 Hz, 1 H), 7.97 (d, J=1.2 Hz, 1 H), 7.80 (dd, J=8.2, 1.6 Hz, 1 H), 7.42 - 7.51 (m, 1 H), 7.35 (d, J=7.4 Hz, 2 H), 7.25 (t, J=7.6 Hz, 2 H), 7.12 - 7.20 (m, 1 H), 4.05 - 4.18 (m, 2 H), 4.02 (s, 2 H), 3.86 (s, 3 H), 3.55 - 3.72 (m, 3 H), 3.21 - 3.27 (m, 2 H), 3.04 (t, J=2.3 Hz, 1 H), 2.28 - 2.40 (m, 2 H), 2.13 - 2.28 (m, 4 H), 1.75 - 1.87 (m, 2 H), 1.63 - 1.74 (m, 2 H), 1.52 (dt, J=13.1, 7.5 Hz, 2 H), 0.86 - 0.98 ppm (m, 3 H), **¹³C NMR** (DMSO-d₆, 101MHz): δ = 195.0, 170.2, 166.7, 157.4, 156.9, 139.3, 135.6, 129.1, 129.1, 128.1, 126.0, 124.7, 123.6, 120.7, 120.2, 112.0, 93.4, 80.3, 73.5, 52.0, 45.8, 29.2, 29.2, 26.8, 26.3, 25.7, 19.5, 19.4 ppm **HRMS** calcd for : C₃₄H₄₁N₈O₃ [M+H]⁺ : 609.3302 ; found 609.3428 (-0.0126 m/z).

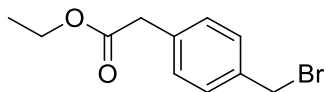
Ethyl 2-(p-tolyl)acetate (49)



In a 250 were dissolved 2-(p-tolyl)acetic acid (10 g, 66.6 mmol) in ethanol (101 ml) and then SULFURIC ACID (0.505 ml, 9.47 mmol) was added. The reaction was stirred over-weekend with reflux. The solvent was concentrated under reduced pressure. The residual yellow oil was transferred into a separating funnel were 200 ml of Et₂O and 70 ml of NaHCO₃ were added. The organic layer was washed 2 times with a solution of NaHCO₃ sat. The organic layer was then washed with brine, dried with Na₂SO₄ and filtered. The yellow solution was concentrated under reduced pressure and give a yellow oil ethyl 2-(p-tolyl)acetate 49 (11.90 g, 66.8 mmol, 100 % yield).

Formula: C₁₁H₁₄O₂; **MW:** 178.23 g/mol; **¹H NMR** (CDCl₃, 400MHz): δ = 7.19 (d, J=8.22 Hz, 2 H), 7.14 (d, J=7.83 Hz, 2 H), 4.15 (q, J=7.04 Hz, 2 H), 2.34 (s, 3 H) 3.58 (s, 2 H), 1.26 (t, J=7.04 Hz, 3 H) ppm, **¹³C NMR** (CDCl₃, 101MHz): δ = 171.79, 136.61, 129.22, 129.07, 60.77, 41.01, 21.06, 14.17 ppm **LCMS(ESI)** [M+H]⁺ : 179.2 m/z

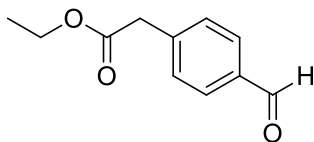
Ethyl 2-(4-(bromomethyl)phenyl)acetate (50)



Initially ethyl 2-(p-tolyl)acetate (10 g, 56.1 mmol), NBS (10.98 g, 61.7 mmol) and AIBN (0.461 g, 2.81 mmol) were mixed together in CHCl₃ (56.1 ml) depending the reaction. The suspension was brought to reflux at 70 °C overnight then cooled at 5°C. The reaction was filtered, and the filtrate was concentrated under reduced pressure. The compound was than purified by flash chromatography using Combiflash Rf 200 (Isco) instrument with a 80 g SiO₂ column (crude material was applied via solid handing) with a solvent mix of A=Hexanes B=EtOAc. The gradient was 0 % of B to 20% B in 35 minutes. Tubes were mixed together and ceoncentrated under reduced pressure to give ethyl 2-(4-(bromomethyl)phenyl)acetate 50 (10.78 g, 34.8 mmol, 62.0 % yield) as a white solid.

Formula: C₁₁H₁₃O₂Br; **MW:** 257.12 g/mol; **¹H NMR** (CDCl₃, 400MHz): δ = 7.36 (d, J=8.22 Hz, 2 H), 7.28 (d, J=8.20 Hz, 2 H), 4.50 (s, 2 H), 4.16 (q, J=7.04 Hz, 2 H), 3.62 (s, 2 H), 1.26 ppm (t, J=6.7 Hz, 3 H) **¹³C NMR** (CDCl₃, 101MHz): δ = 171.4, 137.3, 135.2, 129.6, 61.35, 41.5, 33.2, 14.2 ppm

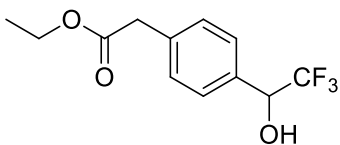
Ethyl 2-(4-formylphenyl)acetate (51)



A mixture of ethyl 2-(4-(bromomethyl)phenyl)acetate (4.35 g, 16.92 mmol) and 4-methylmorpholine 4-oxide hydrate (2.287 g, 16.92 mmol) in Dioxane (56.4 ml) was heated to 100 °C for 90 mins. The reaction was followed by TLC. When the reaction was over, 2/3 of the solvent was evaporated under reduced pressure and the rest transferred into a separating funnel with EtOAc and water. Organic layer (200 ml) was washed with 50 ml of water. The water contained compound so 3 extraction with EtOAc was performed. The combined organic layers were dried over MgSO₄, filtered and solvent removed. The red oil was solubilized in DCM and silica (12 g) was added. The compound was then purified by flash chromatography using Combiflash Rf 200 (Isco) instrument with a 80 g SiO₂ column (crude material was applied via solid handling) with a solvent mix of A=Hexanes B=EtOAc. The gradient was 0 % of B to 25% B in 30 minutes. Tubes were concentrated under reduced pressure to obtain ethyl 2-(4-formylphenyl)acetate 51 (1.99 g, 10.35 mmol, 61.2 % yield) as a white foam.

Formula: C₁₁H₁₂O₃; **MW:** 192.21 g/mol; **¹H NMR** (CDCl₃, 400MHz): δ = 10.01 (s, 1 H), 7.86 (d, J=8.2 Hz, 2 H), 7.47 (d, J=7.8 Hz, 2 H), 4.18 (q, J=7.2 Hz, 2 H), 3.71 (s, 2 H), 1.27 ppm (t, J=7.4 Hz, 3 H), **¹³C NMR** (CDCl₃, 101MHz): δ = 191.5, 170.3, 140.7, 135.0, 129.7, 129.6, 60.9, 41.2, 13.8 ppm
LCMS(ESI) [M+H]⁺ : 193.2 m/z

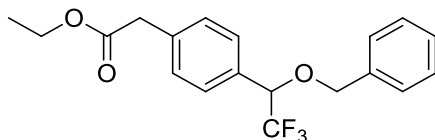
Ethyl 2-(4-(2,2,2-trifluoro-1-hydroxyethyl)phenyl)acetate (52)



In DMF anhydrous (7.43 ml) were dissolved ethyl 2-(4-formylphenyl)acetate (1 g, 5.20 mmol) and cesium fluoride (0.055 g, 0.364 mmol) to give a colorless solution (under N₂). The medium was cooled to 0-5 °C with ice bath and trimethyl(trifluoromethyl)silane (1.000 ml, 6.76 mmol) was added and the reaction was stirred at this temperature for 3 hours. When the trimethyl(trifluoromethyl)silane was added the reaction turned to a bright orange color. TLC with KMnO₄ showed the completion of the reaction. After this time TBAF 1M in THF (5.20 ml, 5.20 mmol) was added and the resulting solution was stirred during 30 minutes at 0-5 °C. TLC with KMnO₄ showed the completion of the reaction. The reaction was transferred in a separating funnel with water (5 ml) and MTBE (5 ml). The aqueous layer was extracted 3 times with MTBE and the combined organic layers were washed with brine, then dried with MgSO₄, filtered and concentrated under reduced pressure. The compound was then purified by flash chromatography using Combiflash Rf 200 (Isco) instrument with a 40 g SiO₂ column (crude material was applied via solid handling) with a solvent mix of A=Hexanes B=EtOAc. The gradient was 0 % of B to 25% B in 30 minutes. Tubes were concentrated under reduced pressure to obtain the ethyl 2-(4-(2,2,2-trifluoro-1-hydroxyethyl)phenyl)acetate 52 (0.974 g, 3.71 mmol, 71.4 % yield) as a white solid.

Formula: C₁₂H₁₃F₃O₃; **MW:** 262.23 g/mol; **¹H NMR** (CDCl₃, 400MHz): δ = 7.32-7.37 (m, J=8.6 Hz, 2 H), 5.02 (q, J=6.7 Hz, 1 H), 4.17 (q, J=7.2 Hz, 2 H), 3.64 (s, 2 H), 2.58 (br. s., 1 H), 1.26 ppm (t, J=7.0 Hz, 3 H), **¹³C NMR** (CDCl₃, 101MHz): δ = 171.2, 135.6, 132.8, 129.5, 127.6, 125.6, 122.8, 72.6, 61.0, 41.1, 29.7, 14.1 ppm

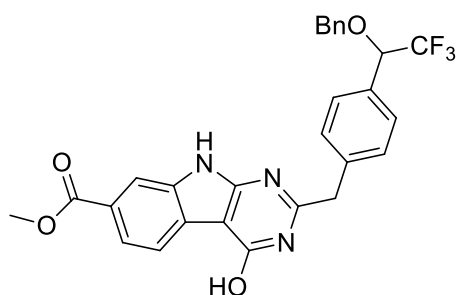
Ethyl 2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)phenyl)acetate (53)



In a 150 ml sealed tube were added ethyl 2-(4-(2,2,2-trifluoro-1-hydroxyethyl)phenyl)acetate (1.5 g, 5.72 mmol) and K_2CO_3 (2.451 g, 17.73 mmol) in Acetonitrile (38.1 ml) to give a white suspension. After the benzyl bromide (0.748 ml, 6.29 mmol) was added and the reaction was stirred at 85°C for 24 hours. Then the reaction was cooled to room temperature and filtered on a buchner. The solid was washed (2 x 5 ml) with acetonitrile and the filtrate was concentrated under reduced pressure. The orange oil was purified by flash chromatography using Combiflash Rf 200 (Isco) instrument with a 24 g SiO_2 column (crude material was applied via solid handling) with a solvent mix of A=Hexanes B=EtOAc. The gradient was 0 % of B to 10% B in 18 minutes. Tubes were concentrated under reduced pressure to obtain ethyl 2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)phenyl)acetate 53 (1.461 g, 4.15 mmol, 72.5 % yield) as a colorless oil.

Formula: $C_{19}H_{19}F_3O_3$; **MW:** 352.35 g/mol; **1H NMR** ($CDCl_3$, 400MHz): δ = 7.29 - 7.44 (m, 9 H), 4.69 (d, J=12.1 Hz, 1 H), 4.66 (q, J=6.7 Hz, 1 H), 4.46 (d, J=12.1 Hz, 1 H), 4.19 (q, J=7.2 Hz, 2 H), 3.66 (s, 2H), 1.29 ppm (t, J=7.0 Hz, 3 H) **^{13}C NMR** ($CDCl_3$, 101MHz): δ = 171.2, 136.8, 136.0, 131.6, 129.9, 128.9, 128.9, 128.5, 128.2, 78.5, 78.2, 77.5, 71.8, 61.3, 41.4, 30.0, 14.5 ppm **LCMS(ESI)** $[M+H]^+$: 353.1 m/z

Methyl 2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)benzyl)-4-hydroxy-9H-pyrimido[4,5-b]indole-7-carboxylate (58)

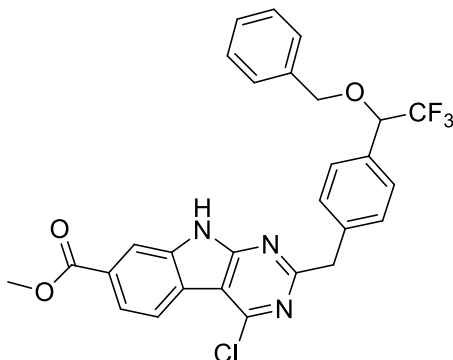


In a 10 mL microwave vial was added 3-carbamoyl-6-(methoxycarbonyl)-1H-indol-2-aminium chloride (600 mg, 2.225 mmol), ethyl 2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)phenyl)acetate (1176 mg, 3.34 mmol) and sodium methanolate 5.4 M (1236 μ l, 6.67 mmol) in MeOH (6742 μ l) to give a brown suspension. The vial was placed in the microwave and heated to 140 °C for 45 mins. LCMS showed only a conversion of 39 % of desired compound and 61 % of starting material. To complete the reaction sodium methanolate 5.4 M (742 μ l, 4.00 mmol) and ethyl 2-(4-bromophenyl)acetate (541 mg, 2.225 mmol) were added and the reaction was placed for another 45 mins in microwave. AcOH (892 μ l, 15.57 mmol) was added and the black solution turns to a light brown suspension. Stirred the resulting suspension at 20°C for 2 hours. Filtered to afford a purple solid methyl 2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)benzyl)-4-hydroxy-9H-pyrimido[4,5-b]indole-7-carboxylate 58 (569 mg, 1.091 mmol, 49.0 % yield).

The filtrate was purified by flash chromatography using Combiflash Rf 200 (Isco) instrument with a 24 g SiO₂ column (crude material was applied via solid handling) with a solvent mix of A=DCM B=EtOAc. The gradient was 0 % of B to 40% B in 18 minutes. Tubes were concentrated under reduced pressure to obtain methyl 2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)benzyl)-4-hydroxy-9H-pyrimido[4,5-b]indole-7-carboxylate 58 (99 mg, 0.190 mmol, 8.53 % yield) as a brown solid.

Formula: C₂₈H₂₂F₃N₃O₄; **MW:** 521.49 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 12.54 (s, 1 H), 12.45 (s, 1 H), 7.99 - 8.05 (m, 2 H), 7.84 (dd, J=8.2, 1.6 Hz, 1 H), 7.48 (s, 4 H), 7.30 - 7.35 (m, 5 H), 5.21 (q, J=7.3 Hz, 1 H), 4.54 (d, J=3.1 Hz, 2 H), 4.08 (s, 2 H), 3.87 ppm (s, 3 H) **LCMS(ESI)** [M+H]⁺ : 522.1 m/z

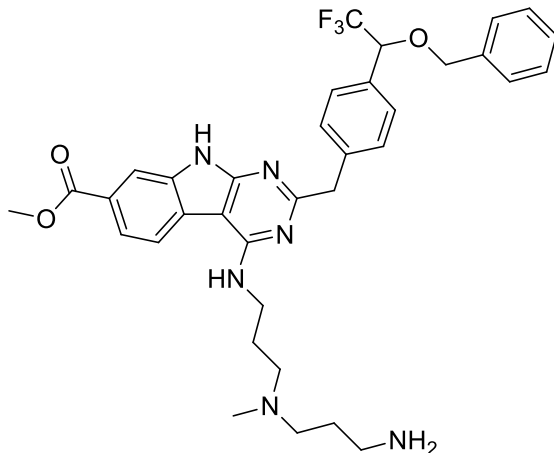
Methyl 2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)benzyl)-4-chloro-9H-pyrimido[4,5-b]indole-7-carboxylate (59)



In a 25 mL round-bottomed flask was added methyl 2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)benzyl)-4-hydroxy-9H-pyrimido[4,5-b]indole-7-carboxylate (550 mg, 1.055 mmol) in POCl₃ (8258 μ l, 89 mmol) to give a brown suspension. The reaction was heated at 90°C for two hours after which the reaction was concentrated under reduced pressure to give a brown foam. The foam was suspended in sat. NaHCO₃ (20 mL for) and stirred for 1 hour. The solid was collected on buchner, washed with water (3 x 2 mL) and the compound was dried 3 hours at 40°C under high vacuum. Isolated methyl 2-(3-(1-(benzyloxy)-2,2,2-trifluoroethyl)benzyl)-4-chloro-9H-pyrimido[4,5-b]indole-7-carboxylate **59** (375 mg, 0.695 mmol, 92 % yield) as a tan solid.

Formula: C₂₈H₂₁F₃N₃O₃Cl; **MW:** 539.93 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 12.99 (s, 1 H), 8.35 (d, J=8.2 Hz, 1 H), 8.14 (d, J=1.6 Hz, 1 H), 8.00 (dd, J=8.2, 1.6 Hz, 1 H), 7.43 - 7.48 (m, 4H), 7.28 - 7.37 (m, 5 H), 5.20 (q, J=7.0 Hz, 1 H), 4.54 (d, J=3.5 Hz, 2 H), 4.36 (s, 2 H), 3.92 ppm (s, 3 H)
LCMS(ESI) [M+H]⁺ : 540.2 m/z

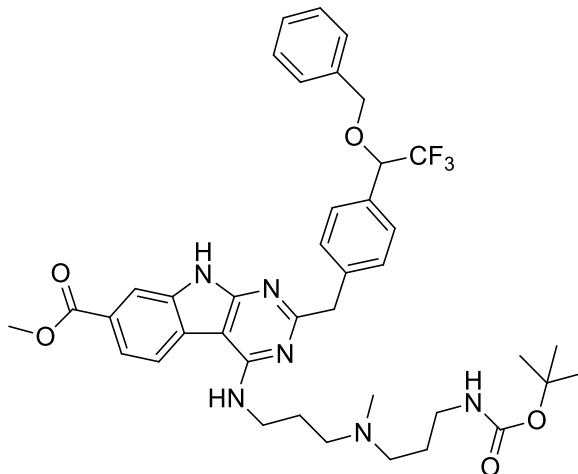
Methyl 4-((3-((3-aminopropyl)(methyl)amino)propyl)amino)-2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (60)



In a 10-20 mL microwave vial was added methyl 2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)benzyl)-4-chloro-9H-pyrimido[4,5-b]indole-7-carboxylate (570 mg, 1.055 mmol) and N1-(3-aminopropyl)-N1-methylpropane-1,3-diamine (1.191 ml, 7.39 mmol) in MeOH (12.000 ml). (brown suspension) The vial was placed in the microwave and heated to 140°C for 30 mins. After this time, LCMS showed completion of the reaction. The solvent was removed by reduced pressure, and the residue was purified by flash chromatography. The filtrate concentrated was purified by flash chromatography using Combiflash Rf 200 (Isco) instrument with a 24 g SiO₂ column (crude material was applied via solid handling) with a solvent mix of A=DCM B=20% MeOH/DCM. The gradient was 0 % of B to 100% B in 18 minutes. Tubes were concentrated under reduced pressure and methyl 4-((3-((3-aminopropyl)(methyl)amino)propyl)amino)-2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate **60** (363 mg, 0.560 mmol, 53.0 % yield) was isolated.

Formula: C₃₅H₃₉F₃N₆O₃; **MW:** 648.72 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 8.28 (d, J=8.6 Hz, 1 H), 7.99 (d, J=1.6 Hz, 1 H), 7.83 (dd, J=8.4, 1.4 Hz, 1 H), 7.58 (br. s., 1 H), 7.39 - 7.51 (m, 4 H), 7.25 - 7.38 (m, 5 H), 5.17 (q, J=7.3 Hz, 1 H), 4.54 (d, J=11.3 Hz, 1 H), 4.50 (d, J=12.1 Hz, 1 H), 4.09 (s, 3 H), 3.84 - 3.91 (m, 3 H), 3.64 (q, J=6.3 Hz, 2 H), 2.57 (br. s., 2 H), 2.52 (br. s., 2 H), 2.40 (t, J=6.8 Hz, 2 H), 2.34 (t, J=7.0 Hz, 2 H), 2.16 (s, 2 H), 1.79 (quin, J=6.7 Hz, 2 H), 1.51 ppm (dt, J=14.4, 7.1 Hz, 2 H) **LCMS(ESI) [M+H]⁺** : 649.3 m/z

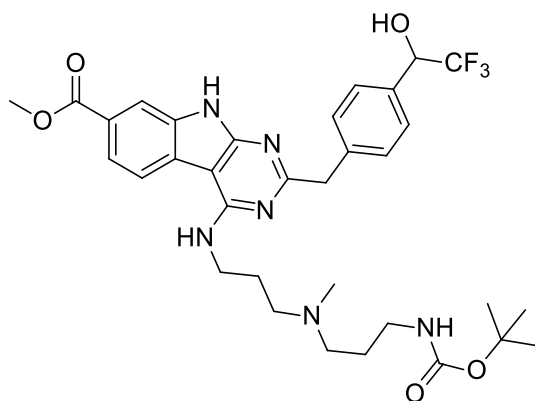
Methyl 2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)benzyl)-4-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (61**)**



In a 25 mL round-bottomed flask were added methyl 4-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (363 mg, 0.560 mmol) and Et₃N (94 μ l, 0.671 mmol) in CH₂Cl₂ (3780 μ l, 58.8 mmol) and MeOH (2521 μ l, 62.3 mmol) to give a yellow solution. A solution of di-tert-butyl dicarbonate (156 μ l, 0.671 mmol) in CH₂Cl₂ (1260 μ l, 19.58 mmol) was added slowly. Some gas was released with the addition of this solution. Stirred at 20 °C. After 45 minutes, reaction was finished by LCMS. Concentrated to dryness to give a yellow foam. The residue was purified on ISCO using a RediSep Gold 24 g column with a solvent mix of A=DCM B=(20/78/2) MeOH/DCM/NH₄OH. The gradient was 20 % of B to 80% B in 16 minutes. The tubes were concentrated under reduced pressure and methyl 2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)benzyl)-4-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate **61** (300 mg, 0.401 mmol, 71.6 % yield) was isolated as a yellow foam.

Formula: C₄₀H₄₇F₃N₆O₅; **MW:** 748.83 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 12.06 (s, 1 H), 8.27 (d, J=8.2 Hz, 1 H), 7.99 (d, J=1.2 Hz, 1 H), 7.83 (dd, J=8.2, 1.6 Hz, 1 H), 7.54 (t, J=5.5 Hz, 1 H), 7.45 - 7.50 (m, 2 H), 7.40 - 7.45 (m, 2 H), 7.27 - 7.36 (m, 5 H), 6.77 (t, J=5.5 Hz, 1 H), 5.16 (q, J=7.0 Hz, 1 H), 4.54 (d, J=11.3 Hz, 1 H), 4.50 (d, J=11.7 Hz, 1 H), 4.09 (s, 2 H), 3.88 (s, 3 H), 3.64 (q, J=6.7 Hz, 2 H), 2.93 (q, J=6.5 Hz, 2 H), 2.40 (br. s., 2 H), 2.32 (br. s., 2 H), 2.16 (br. s., 3H), 1.78 (quin, J=6.3 Hz, 2 H), 1.54 (quin, J=7.4 Hz, 2 H), 1.33 ppm (s, 9 H) **LCMS(ESI)** [M+H]⁺ : 749.3 m/z

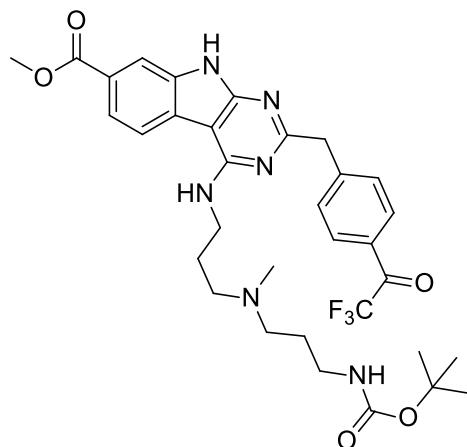
Methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(2,2,2-trifluoro-1-hydroxyethyl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (62)



In a 50 mL round-bottomed flask were added methyl 2-(4-(1-(benzyloxy)-2,2,2-trifluoroethyl)benzyl)-4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-9H-pyrimido[4,5-b]indole-7-carboxylate (0.300 g, 0.401 mmol) and Pd-C 10%wt. (50% wet) (0.298 g, 0.140 mmol) in MeOH (9.72 ml, 240 mmol) to give a black solution. The flask was flushed with Hydrogen (3 vacuums + refill cycles). Stirred at 20 °C under an hydrogen atmosphere overnight. The mixture was filtered over celite and the filtrate was concentrated. The oil was purified on ISCO using a RediSep Gold 12 g column with a solvent mix of A=DCM B=(20/78/2) MeOH/DCM/NH₄OH. The gradient was 20 % of B to 80% B in 16 minutes. The tubes were concentrated under reduced pressure and methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(2,2,2-trifluoro-1-hydroxyethyl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate 62 (107 mg, 0.162 mmol, 40.5 % yield) was isolated as a white foam.

Formula: C₃₃H₄₁F₃N₆O₅; **MW:** 658.71 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 12.10 (s, 1 H), 8.28 - 8.40 (m, 1 H), 7.99 (d, J=1.2 Hz, 1 H), 7.82 (dd, J=8.2, 1.6 Hz, 1 H), 7.53 (t, J=5.5 Hz, 1 H), 7.34 - 7.47 (m, 4 H), 6.86 (br. s., 1 H), 6.76 (d, J=5.9 Hz, 1 H), 5.05 - 5.13 (m, 1 H), 4.07 (s, 2 H), 3.88 (s, 3 H), 3.66 (d, J=5.9 Hz, 2 H), 2.96 (d, J=5.9 Hz, 2 H), 2.74 (br. s., 4 H), 2.21 - 2.48 (m, 3 H), 1.91 (br. s., 2 H), 1.64 (br. s., 2 H), 1.34 ppm (s, 9 H) **LCMS(ESI)** [M+H]⁺ : 659.3 m/z

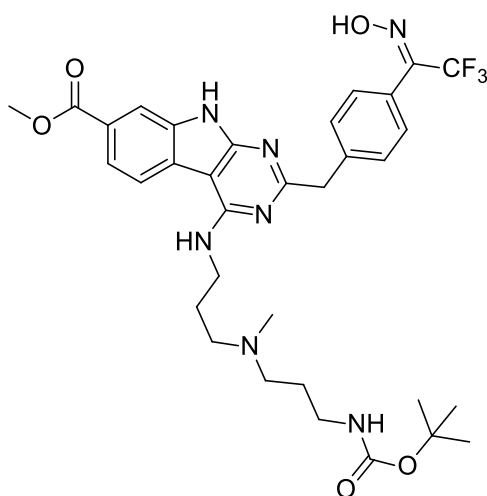
Methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(2,2,2-trifluoroacetyl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (63)



In a 25 mL round-bottomed flask were added methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(2,2,2-trifluoro-1-hydroxyethyl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (93 mg, 0.141 mmol) and Dess-Martin periodinane (263 mg, 0.621 mmol) in CH₂Cl₂ (3488 μl, 54.2 mmol) to give a white suspension. After 2 minutes, an orange suspension was obtained and was stirred at 20 °C for 40 mins. The residue was purified on ISCO using a RediSep 12 g column (CH₂Cl₂/Acetone with Et₃N 0.5% in each). The crude product was dissolved and loaded onto a pre-column (medium, 6g SiO₂, used crude reaction mixture) and a gradient of 0% to 100 % of Acetone was used. Collected fractions were concentrated and methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(2,2,2-trifluoroacetyl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate **63** (92 mg, 0.140 mmol, 99 % yield) as a yellow foam.

Formula: C₃₃H₃₉F₃N₆O₅; **MW:** 656.70 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 12.08 (d, J=13.7 Hz, 1 H), 8.24 - 8.33 (m, 1 H), 7.96 - 8.03 (m, 1 H), 7.80 - 7.85 (m, 1 H), 7.66 (d, J=8.6 Hz, 1H), 7.47 - 7.57 (m, 2 H), 7.34 - 7.46 (m, 2 H), 6.80 (br. s., 1 H), 4.06 (s, 2 H), 3.87 (s, 3 H), 3.57 - 3.71 (m, 2 H), 3.38 (s, 2 H), 2.84 - 3.09 (m, 4 H), 2.28 (br. s., 4 H), 1.85 (br. s., 2 H), 1.57 (br. s., 2 H), 1.33 ppm (d, J=3.9 Hz, 9 H) **LCMS(ESI)** [M+H]⁺ : 657.3 m/z

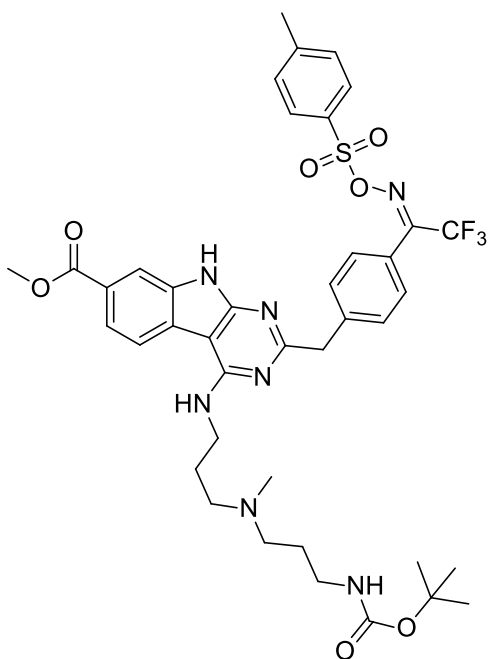
(E)-methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(2,2,2-trifluoro-1-(hydroxyimino)ethyl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (64**)**



In a 2 - 5 mL microwave vial was added methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(2,2,2-trifluoroacetyl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (0.092 g, 0.140 mmol), hydroxylamine hydrochloride (0.012 g, 0.168 mmol) and pyridine (0.181 ml, 2.242 mmol) in MeOH (0.969 ml, 23.94 mmol) to give a yellow solution. The vial was sealed and heated to 65 °C overnight. After 24 hrs, hydroxylamine hydrochloride (2.92 mg, 0.042 mmol) was added again and the reaction was continued heating at 65 °C for 6 hours. The reaction mixture was concentrated to dryness to give a light-yellow foam. The foam was dissolved into CH₂Cl₂ (50 mL) and MeOH (10 mL). Washed with sat. NaHCO₃ (15 mL). The aqueous layer was backextracted with CH₂Cl₂ (35 mL) and MeOH (10 mL). The organic layer was dried over MgSO₄, filtered and concentrated to dryness. The yellow foam **64** was used as such for the next reaction (crude).

Formula: C₃₃H₄₀F₃N₇O₅; **MW:** 671.71 g/mol; **LCMS(ESI) [M+H]⁺ :** 672.3 m/z

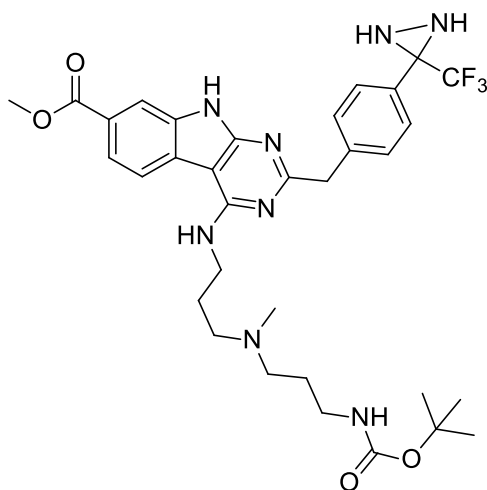
(E)-methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(2,2,2-trifluoro-1-((tosyloxy)imino)ethyl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (65)



In a 25 mL round-bottomed flask were added methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(2,2,2-trifluoro-1-(hydroxyimino)ethyl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (0.094 g, 0.140 mmol), DMAP (1.710 mg, 0.014 mmol) and TRIETHYLAMINE (0.023 ml, 0.167 mmol) in CH₂Cl₂ (5.87 ml, 91 mmol) to give a yellow solution. Ts-Cl (0.0307 g, 0.154 mmol) was added portion-wise. After 1 hr the reaction was finished. Then the mixture was diluted with CH₂Cl₂ (12 mL) and was washed with water (3 x 12 mL). The organic layer was dried over MgSO₄, filtered, and concentrated to dryness. The compound **65** was used as such for the next reaction.

Formula: C₄₀H₄₆F₃N₇O₇S; **MW:** 825.90 g/mol; **LCMS(ESI) [M+H]⁺ :** 826.2 m/z

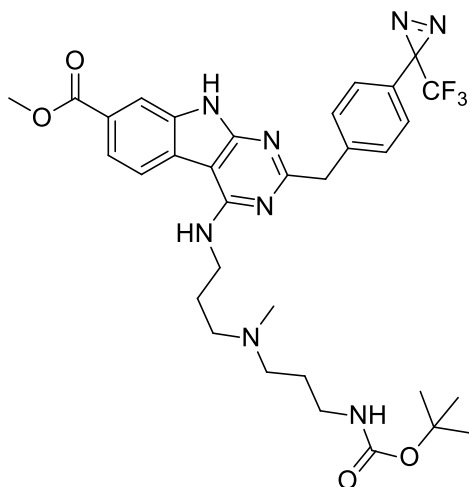
Methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(3-(trifluoromethyl)diaziridin-3-yl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (66)



In a 25 mL sealed tube was added methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(2,2,2-trifluoro-1-((tosyloxy)imino)ethyl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (0.098 g, 0.119 mmol) in CH₂Cl₂ (2.290 ml, 35.6 mmol) to give a orange solution. The mixture was cooled to -78 °C. AMMONIA (0.770 ml, 35.6 mmol) was condensed into the tube at -78 °C. Slowly warmed to 20 °C and stirred for 3 hours. Cooled to -78 °C and placed a septa with gas outlet. Slowly warmed to 20 °C to evaporate ammonia. After 3 hours the reaction was completed, then concentrated to dryness and left in freezer overnight. The next day the residue was purified on ISCO using a RediSep 6 g column with a solvent mix of A=DCM B=(20/78/2) MeOH/DCM/NH₄OH. The gradient was 20 % of B to 80% B in 16 minutes. The tubes were concentrated under reduced pressure and methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(3-(trifluoromethyl)diaziridin-3-yl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate **66** (30 mg, 0.045 mmol, 37.7 % yield) was isolated as a white foam.

Formula: C₃₃H₄₁F₃N₈O₄; **MW:** 670.73 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 12.07 (s, 1 H), 8.29 (d, J=8.2 Hz, 1 H), 7.99 (d, J=1.6 Hz, 1 H), 7.82 (dd, J=8.2, 1.6 Hz, 1 H), 7.54 (t, J=5.3 Hz, 1 H), 7.37 - 7.50 (m, 4 H), 6.81 (br. s., 1 H), 4.02 - 4.14 (m, 3 H), 3.88 (s, 3 H), 3.86 (br. s., 1 H), 3.61 - 3.70 (m, 2 H), 2.90 - 2.99 (m, 2 H), 2.60 (br. s., 2 H), 2.22 - 2.36 (m, 3 H), 1.85 (br. s., 2 H), 1.60 (br. s., 2 H), 1.35 ppm (s, 9 H) **LCMS(ESI)** [M+H]⁺ : 671.3 m/z

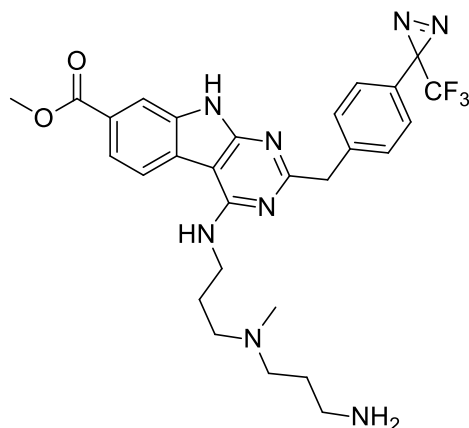
Methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (67)



In a 5 mL round-bottomed flask protected from light were added methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(3-(trifluoromethyl)diazirin-3-yl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (30 mg, 0,045 mmol) and Et₃N (18,70 μ l, 0,134 mmol) in CH₂Cl₂ (857 μ l, 13,32 mmol) to give a light yellow solution. Iodine (11,92 mg, 0,047 mmol) was added. Stirred at 20 °C. After 15 minutes, the reaction was completed and the residue was purified on ISCO using a RediSep 4 g column with a solvent mix of A=DCM B=(20/78/2) MeOH/DCM/NH₄OH. The gradient was 20 % of B to 80% B in 12 minutes. Collected fractions were concentrated under reduced pressure and the methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate 67 (21 mg, 0,031 mmol, 70,2 % yield) was isolated as a light yellow foam.

Formula: C₃₃H₃₉F₃N₈O₄; **MW:** 668.71 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): d = 12.07 (s, 1 H), 8.31 (d, J=8.2 Hz, 1 H), 7.99 (d, J=1.2 Hz, 1 H), 7.82 (d, J=8.6 Hz, 1 H), 7.45 - 7.56 (m, 3 H), 7.21 (d, J=7.8 Hz, 2 H), 6.83 (br. s., 1 H), 4.10 (s, 2 H), 3.88 (s, 3 H), 3.62 (q, J=6.7 Hz, 1 H), 2.95 (q, J=6.3 Hz, 2 H), 2.61 (br. s., 2 H), 2.38 (br. s., 3 H), 1.85 (br. s., 2 H), 1.54 - 1.68 (m, 2 H), 1.33 ppm (s, 9 H) **LCMS(ESI)** [M+H]⁺ : 669.4 m/z

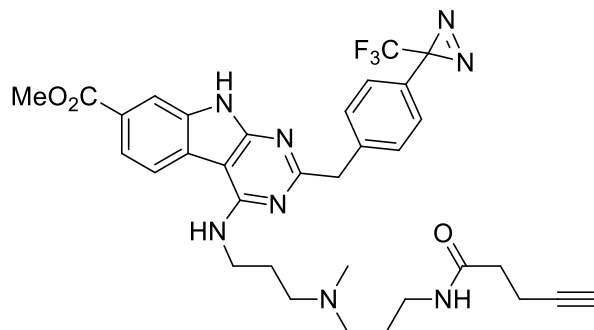
Methyl 4-((3-((3-aminopropyl)(methyl)amino)propyl)amino)-2-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (68**)**



In a 25 mL round-bottomed flask was added methyl 4-((3-((3-((tert-butoxycarbonyl)amino)propyl)(methyl)amino)propyl)amino)-2-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (0.021 g, 0.031 mmol) in CH₂Cl₂ (1.312 ml, 20.40 mmol) to give a light yellow solution. Trifluoroacetic acid (0.131 ml, 1.702 mmol) was added. Stirred at 20 °C. After 1 hour, the reaction was finished. Then the reaction mixture was diluted with CH₂Cl₂ (15 mL) and the organic layer washed with sat. NaHCO₃ (10 mL). The aqueous layer was backextracted with CH₂Cl₂ (10 mL). The organic layer was dried over MgSO₄, filtered and concentrated to dryness. The light-yellow foam **68** was used directly to the next reaction.

Formula: C₂₈H₃₁F₃N₈O₂; **MW:** 568.59 g/mol; **LCMS(ESI)** [M+H]⁺ : 569.2 m/z

Methyl 4-((3-(methyl(3-(pent-4-ynamido)propyl)amino)propyl)amino)-2-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (UM142946)



HATU (8,07 mg, 0,034 mmol) was added to a solution of pent-4-ynoic acid (3.11 mg, 0.032 mmol) in DMF (0.125 ml). 5 min later, methyl 4-((3-((3-aminopropyl)(methyl)amino)propyl)amino)-2-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (0.015 g, 0.026 mmol) in DMF (0.35 ml) was added followed by Hunig'sBase (10.60 μ l, 0.061 mmol). After letting the reaction go overnight, the compound was directly purified by HPLC-MS (2 injections). Solvent A: MeOH:H₂O 5:95, AcOH 0.1% Solvent B: MeOH:H₂O 95:5, AcOH 0.1 %. Gradient 25 % to 90 % over 12 minutes. Tubes were combined and concentrated under reduced pressure to give methyl 4-((3-(methyl(3-(pent-4-ynamido)propyl)amino)propyl)amino)-2-(4-(3-(trifluoromethyl)-3H-diazirin-3-yl)benzyl)-9H-pyrimido[4,5-b]indole-7-carboxylate (1,82 mg, 2.81 μ mol, 10,64 % yield) as orange solid.

Formula: C₃₃H₃₅F₃N₈O₃; **MW:** 648.68 g/mol; **¹H NMR** (DMSO-d₆, 400MHz): δ = 12.07 (br. s., 1 H), 8.27 (d, J=8.2 Hz, 1 H), 7.99 (d, J=1.6 Hz, 1 H), 7.84 - 7.87 (m, 1 H), 7.83 (dd, J=8.4, 1.4 Hz, 1 H), 7.55 (t, J=5.5 Hz, 1 H), 7.50 (d, J=8.6 Hz, 2 H), 7.20 (d, J=7.8 Hz, 1 H), 4.10 (s, 2 H), 3.88 (s, 2 H), 3.55 - 3.64 (m, 2 H), 3.02 - 3.10 (m, 2 H), 2.72 (t, J=2.5 Hz, 1 H), 2.28 - 2.41 (m, 6 H), 2.19 - 2.25 (m, 2 H), 2.14 (s, 3 H), 1.67 - 1.83 (m, 3 H), 1.51 - 1.60 ppm (m, 2 H) **HRMS** calcd for : C₃₃H₃₆F₃N₈O₃ [M+H]⁺ : 649.2862 ; found 649.2855 (-0.0007 m/z). **HPLC** (Kinetex C18 2.6 microns 4.6 x 30 mm, A MeOH:H₂O:TFA (5:95:0.5) B MeOH:H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes (end time 3 minutes), measured at 254 nm): rt = 2.148 min **HPLC** (Kinetex Eclipse XDB C5 2.6 microns, A

MeOH:H₂O:TFA (5:95:0.5) B MeOH:H₂O:TFA (95:5:0.5) 0 to 100% of B in 2 minutes (end time 3 minutes), Uv measured at 254 nm): rt = 2.147 min

A.4 Biochemistry/Proteomics section for the second project: Target identification of the UM171 compound

Cell culture

The cells used for experiments were from a source of an acute meloid leukemia patient, named AML5. To grow them it was used the Duplecco's Modified Eagle Medium containing 10 % heat inactivated FBS, 1% Penicillin/Streptomycin antibiotic and the growth factor GMCSF (10 ng/mL). The cells were incubated with a 5 % CO₂ atmosphere and a temperature of 37 °C in flask. The cells were counted with the trypan blue under microscope with cell counting chamber slides. When the desired amount of cells was obtained, they were centrifuged at 1000 g for 5 minutes at 4 °C. Cell pellets were then washed with cold PBS 1x and used as such for future manipulation or frozen at -20°C.

Chemicals:

DMEM: Wisent 319-051-CL
FBS: Hyclone SH30396
Penicillin/Streptomycin: Gibco 15140122
GMCSF: Shenandoah 100-08-20ug
PBS 1x: Wisent 311-010-CL
Tryptan Blue: Gibco 15250061

Wavelength determination

The emission spectrum was taken by simply irradiating a spectrophotometer with the Jena lamp and there was real-time acquisition of this emission. The resulting spectrum comes from an average recording over 20 seconds. The wavelengths probed were 200 to 800 nm.

Materials:

Uv lamp: UVP High Intensity UV Lamps B-100AP Analytik Jena

Spectrophotometer: Ocean Optics Spectrophotometer USB2000

Peptides fragments identifications

For each compound, 500 μL of a solution of 10 μM in MeOH was performed prior to injection. After that this solution was drawn into a 500 μL syringe. To this syringe was connected directly to the source via a polyethylene tube. The sample was injected on the Q Exactiveplus under the continuous mode by direct infusion with a flow rate of 1 $\mu\text{L}/\text{min}$. The choice of collision energies as well as the selection of the transmitted masses were selected manually. The selected collision energies were from 10 eV to 50 eV and the analysis zone for the ions of the MS2 from 90 to 700 m/z. Each spectrum was recorded continuously, until a full overview of the collision energies. After that a correlation of the most abundant ions according to the collision energies was performed with R software to draw the heat map.

Chemicals and materials:

UM171's probes (941, 942, 945, 946 and the compound UM171)

Mass spectrometer: Q ExactivePlus Hybrid Quadrupole-Orbitrap

Thermo Scientific HPLC Syringe 500 μL

Peptides modifications with UM171's probes

The procedure was inspired from this paper: Ziemianowicz, D. S.; Bomgarden, R.; Etienne, C.; Schriemer, D. C., Amino Acid Insertion Frequencies Arising from Photoproducts Generated Using Aliphatic Diazirines. *J Am Soc Mass Spectrom* **2017**, *28* (10), 2011-2021

Initially, 100 μL of a solution of Hb was prepared with a concentration of 400 μM in the 50 mM ABC buffer. To this mixture were added 1 μL of TCEP 0.5 M and the solution was mixed at 37°C during 30 minutes on the thermal shaker at 600 rpm. After, 1.5 μL of a solution of chloroacetamide 1M was added to the protein solution and then agitated for another 30 minutes at 37°C. Following this step, the protein was digested with 4 μL of a solution of trypsin at 114 μM overnight at 37°C. The reaction was stopped by adding 8 μL of formic acid into the solution. The peptides were desalted using HLB cartridge 30 mg following this procedure: The column was initially washed with a solution of ACN/Water/FA (80/20/0.1), then washed 2 times with Water/FA (99.8/0.2), then the sample was added. The sample was added on the column, washed 2 times with Water/FA (99.8/0.2) and the peptides were eluted with the solution ACN/Water/FA (80/20/0.1). The filtrate was evaporated under reduced pressure using a SpeedVac and reconstituted in DMSO in a way to

have a final concentration of 240 μM of Hb. In a TLC capillary tube were mix 1 μL of the solution probe (10 mM) with 1 μM of the protein solution (240 μM). The liquid was frozen with dry ice and was irradiated with UV lamp under dry ice for 30 minutes. The liquid was solubilised in 18 μL of 4% FA in water and injected on the QExactive HF from Thermo.

Parameter of mass spectrometry: Trypsic peptides were loaded on C18 precolumn and separated with an home-made C18 column (150 μm ID x 15 cm) made of Jupiter C18, 3 μm , 300 \AA . This separation was possible with a linear gradient of 5-40% of acetonitrile (0.2 % Formic acid) in a 50 min with a flow rate of 600 nL/min. The liquid chromatography was coupled with the Q Exactive HF, with a mass scan range of 650-2000 m/z. The resolution selected was 120 000 and the AGC target values for MS was set to 5e6 (max fill time 50 ms) and for MS/MS it was set to 5e4 (max fill time of 750 ms). The neutral energy of collision was fixed to 25 and with an exclusion of charge state below 2 and higher than 7. Tandem mass spectra were searched upon the Uniprot databases with specific modification of the probe. The search was performed with the software Peaks (Bioinformatics Solutions Inc) using a FDR of 3%.

Chemicals and materials:

ABC 50 mM buffer (pH = 7.8)

TCEP 0.5 M solution: Thermo scientific, #77720

Chloroacetamide: Sigma-Aldrich, CAS: 79-07-2

Hemoglobin Human: Sigma-Aldrich, CAS: 9008-02-0

Sequencing Grade Modified Trpsin: Promega, V5111

Dry ice

Capillary tube: Fischer Scientific, CG118202

Uv lamp: UVP High Intensity UV Lamps B-100AP Analytik Jena

Trypsin from bovine pancreas: Sigma-Aldrich, T1426

Thermal Shaker: VWR Thermal Shake touch, 89232-908

Acetonitrile: Fisher Chemical, CAS: 75-05-8

Formic acid: Sigma-Aldrich, CAS: 64-18-6

HLB cartridge: Waters (Oasis), WAT094225

Vacuum Manifolds: Thermo Scientific, 60104-233

SpeedVac Vacuum Concentrator: Thermo Fisher, SPD131DDA

Determination of the best lysis buffer for Click reaction

The procedure was inspired from this paper: Yang, Y.; Yang, X.; Verhelst, S. H. L., Comparative analysis of click chemistry mediated activity-based protein profiling in cell lysates. *Molecules (Basel, Switzerland)* **2013**, *18* (10), 12599-12608

The initial step was to culture 40 M of AML5 cells following the procedure describe in the section **cell culture**. The idea was to split this numbers of cells in 5 equal number to test every lysis buffer. The next step was the preparation of these buffer prio to the experience. The buffers were prepare as such:

NP40 1% in 50 mM Hepes (5 mL) : 50 uL NP40, 250 uL HEPES 1M pH 7.5, 4,7 mL H₂O

Triton X-100 1 % in 50 mM Hepes (5 mL) : 50 uL triton X-100, 250 uL HEPES 1M pH 7.5, 4,7 mL H₂O

RIPA 1x (5 mL): 50 uL triton X-100, 50 mg sodium deoxycholate, 50 uL (10 % SDS), 250 uL HEPES 1M pH 7.5, 43.8mg NaCl, 4.65 ml H₂O.

8M Urea in 50 mM Hepes (5 mL): 2.4 g urea, 250 uL HEPES 1M pH 7.5, adjust until 5 mL of H₂O.

SDS 1% in 100 mM ABC pH 8.5 (5 mL) : 39.5 mg ABC, 50 uL (10 % SDS), 4,95 mL H₂O.

Each cells were lyse with 100 uL of the corresponding buffer, after the addition of the liquid the 5 sample were incubate under ice for a total of 30 minutes. Every 10 minutes, every sample was sonicated for 5 secondes. Thereafter, cells were spin at 15 000 rpm at 4°C for 20 minutes. The supernatant was transfered in new 1.5 mL tubes. The concentration of every sample was determined by the protein assay procedure from Bio-Rad. Depending of the concentration obtained, all the conditions were prepared by dilution in a way to obtain 1 mg/mL of protein. If the concentration of buffer should be reduced, water was added either way initial buffer was added to dilute the protein.

The concentration of every buffer that was used for this experiment are present in this table

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
NP-40 1%	NP-40 0.5%	NP-40 0.1%	Triton 1%	Triton 0.5%	Triton 0.1 %	RIPA 1 x	RIPA 0.5 x	RIPA 0.1 x	Urea 8M	Urea 4M	Urea 0.8M	SDS 1%	SDS 0.5%	SDS 0.1%	PBS

Thereafter reagent of the click chemistry was added to every condition. The final concentration/quantity of every reagent per reaction was 300 ng of BSA alkyne, 0.2 mM CuSO₄, 5 mM NaAsc, 1 mM THPTA and 50 μM of CalFluor 647. Each reaction was performed on the thermomixer at 37°C with 800 rpm for 1 hour. Reactions were stopped by adding 240 uL of methanol, 80 uL of chloroform and 320 uL of water. These mixtures were spin at 15 000 rpm at 4°C for 5 minutes, then the upper layer was removed. Another 150 uL of MeOH was added and tubes were spin during 30 minutes at 15 000 rpm. The supernatant was removed, and the precipitate was dry with air to remove the methanol. Proteins were resolubilized in the loading buffer 1x, heated at 95°C for 8 minutes and loaded on a precast acrylamide gel with gradient 4-12%. The voltage used was 110 V over 1 hour. The gel was revealed with the FLA 9500.

Chemicals and materials:

BSA alkyne: Click chemistry tools, 1536

PBS 1x: Wisent 311-010-CL

MeOH: Sigma-Aldrich, CAS : 67-56-1

CHCl₃: Sigma-Aldrich, CAS: 67-66-3

CuSO₄ : Sigma-Aldrich, CAS: 7758-98-7

CalFluor 647: Click chemistry tools, 1372

THPTA: Sigma-Aldrich, CAS:760952-88-3

Sodium L-Ascorbate: Sigma-Aldrich, CAS: 134-03-2

NP40 alternative: Sigma-Aldrich, CAS: 9016-45-9

Triton X-100: Sigma-Aldrich, CAS: 9002-93-1

Urea: Fisher scientific, CAS: 57-13-6

10 % SDS solution: Wisent, 880-550-CL

Bio-Rad Protein Assay: Bio-Rad, 5000006

96-well plate clear flat bottom: Corning, 353072

Sonic Dismembrator Model 100: Fisher Scientific

Centrifuge: Sorvall Legend RT and Fisher Scientific, Accuspin micro 17

Typhoon FLA 9500: Ge Healthcare

Precast acrylamide gel: 4-12% Criterion XT Bis-Tris Protein Gel from Biorad (3450123)

Determination of the best temperature for Click reaction

In order to determine the ideal temperature to perform the Click reaction, 10 conditions were prepared. These conditions are presented in the following table:

	1	2	3	4	5	6	7	8	9	10
Temperature	Negative control (No copper)	4 °C	4 °C	4 °C	25 °C	25 °C	25 °C	37 °C	37 °C	37 °C
Time		0.5 h	2 h	16 h	0.5 h	2 h	16 h	0.5 h	2 h	16 h

In order to prepare the conditions, the various components were mixed so as to form a final solution in PBS (for every tube) with the following concentrations: 100 ng of BSA alkyne, 0.2 mM CuSO₄, 5 mM NaAsc, 1 mM THPTA and 50 µM of CalFluor 647. There is an exception for the negative control where copper has not been added. The reactions were started off in a staggered fashion so that they all ended at the same time. So the 16 hours reaction was started the day before. Reactions were stopped by adding 240 µL of methanol, 80 µL of chloroform and 320 µL of water. These mixtures were spin at 15 000 rpm at 4°C for 5 minutes, then the upper layer was removed. Another 150 µL of MeOH was added and tubes were spin during 30 minutes at 15 000 rpm. The supernatant was removed, and the precipitate was dry with air to remove the methanol. Proteins were resolubilized in the loading buffer 1x, heated at 95°C for 8 minutes and loaded on a precast acrylamide gel with gradient 4-12%. The voltage used was 110 V over 1 hour. The gel was revealed with the FLA 9500.

Chemicals and materials:

BSA alkyne: Click chemistry tools, 1536

PBS 1x: Wisent 311-010-CL

MeOH: Sigma-Aldrich, CAS : 67-56-1

CHCl₃: Sigma-Aldrich, CAS: 67-66-3

CuSO₄ : Sigma-Aldrich, CAS: 7758-98-7

CalFluor 647: Click chemistry tools, 1372

THPTA: Sigma-Aldrich, CAS:760952-88-3

Sodium L-Ascorbate: Sigma-Aldrich, CAS: 134-03-2

Centrifuge: Sorvall Legend RT and Fisher Scientific, Accuspin micro 17

Typhoon FLA 9500: Ge Healthcare

Precast acrylamide gel: 4-12% Criterion XT Bis-Tris Protein Gel from Biorad (3450123)

Evaluation of the cleavage reaction by LCMS

This experiment includes 2 cleavable biotins, therefore two conditions. In a first Eppendorf there was solubilization of 1 uL of the DADPS biotin alkyne (10 mM in DMSO) in 100 uL of water. In the second Eppendorf, 1 uL of the Dde TAMRA biotin alkyne was dissolved in 100 uL of water. The two samples were injected before their cleavage reaction on an Agilent LCMS with the following conditions: A gradient of 0 to 100 % of B in 2 minutes, where A is 95% H₂O, 5% MeOH and 0.05 % HCOOH and B is 5% H₂O, 95% MeOH and 0.05 % HCOOH. The column used was a C18 of Kinetex (3.0 mm x 30 mm and pores 2.6 μm).

Then for the condition with the DADPS biotin there was addition of 4 uL of formic acid to the solution and stirring at 25°C for 2 hours. For the TAMRA biotin, there was addition of 4 uL of a 50 % wt hydrazine solution to the Eppendorf and the mixture was mixed under the same shaking method. Thereafter, both samples were injected on the LCMS.

Chemicals and materials:

Formic acid: Sigma-Aldrich, CAS: 64-18-6
DADPS biotin alkyne: Click chemistry tools, #1331
Dde TAMRA biotin Alkyne: Click chemistry tools, #1368
Hydrazine: Sigma-Aldrich, CAS: 10217-52-4
LCMS 1260 infinity II : Agilent

Evaluation of cleavable biotins in cellular media

With the same procedure described in section ***Determination of the best lysis buffer for Click reaction***, there was preparation of the lysis solution with concentration of 1 mg/mL for each buffer. As there are 3 different lysis conditions there was a preparation of 600 uL, i.e. 200 uL with 0.8 M urea, 200 uL with 0.1 % of NP40 and 200 uL with 0.1% of Triton X-100. The click conditions are those presented in the table below:

	1	2	3	4	5	6	7	8
Concentration lysis buffer	No lysate (PBS)	NP40 0.1%	Triton X-100 0.1 %	Urea 0.8 M	No lysate (PBS)	NP40 0.1%	Triton X-100 0.1 %	Urea 0.8 M
Cleavable biotin	DADPS	DADPS	DADPS	DADPS	Dde-TAMRA	Dde-TAMRA	Dde-TAMRA	Dde-TAMRA
Volume	100 uL	100 uL	100 uL	100 uL	100 uL	100 uL	100 uL	100 uL

Samples 1 and 5 are the positive controls and there was no lysate in them, it was 100 uL of PBS. For the other samples, 100 uL of the dilute lysis buffers were used following the table. For each sample, reagents were added to perform the Click reaction. So, 300 ng of BSA alkyne, 1 uL of CuSO₄ (20 mM), 5 uL of NaAsc (100 mM), 2 uL of THPTA (50 mM) and 0.5 uL of the corresponding cleavable biotin (10 mM). The reactions were mixed on the thermomixer during 1 hour at 37°C. Reactions were stopped by adding 240 uL of methanol, 80 uL of chloroform and 320 uL of water. These mixtures were spin at 15 000 rpm at 4°C for 5 minutes, then the upper layer was removed. Another 150 uL of MeOH was added and tubes were spin during 30 minutes at 15 000 rpm. The supernatant was removed, and the precipitate was dry with air to remove the methanol. Subsequently, the proteins were solubilized using a solution of 1% SDS in PBS (100 uL). To this solution 15 uL of the streptavidin beads slurry was added and the mixture was let overnight at 4°C on the shaker. The next day, the supernatant was removed from each sample and each bead was washed multiples times: 3 times with a solution of 1% SDS (3x100 uL) and 4 times with PBS (4x100 uL). The supernatant was removed, and the cleavage solution was added to the corresponding sample. DADPS samples receive 100 uL of a 4 % formic acid solution and for the Dde-TAMRA samples 100 uL of 2% hydrazine solution. The beads were stirred for 1 hour at 37°C and the supernatant was gently removed and transferred to a second tube. The cleavage step was repeated a second time. For the PBS controls, only water was added. Supernatants from the same samples were combined and the liquid was evaporated with the SpeedVac. Proteins were resolubilized in the loading buffer 1x, heated at 95°C for 8 minutes and loaded on a precast acrylamide gel with gradient 4-12%. The voltage used was 110 V over 1 hour. For the fluorescence image, the gel was revealed with the FLA 9500 and for the silver staining the procedure provided by the retailer was followed.

Chemicals and materials:

BSA alkyne: Click chemistry tools, 1536

PBS 1x: Wisent 311-010-CL

MeOH: Sigma-Aldrich, CAS : 67-56-1

CHCl₃: Sigma-Aldrich, CAS: 67-66-3

CuSO₄ : Sigma-Aldrich, CAS: 7758-98-7

THPTA: Sigma-Aldrich, CAS:760952-88-3

Sodium L-Ascorbate: Sigma-Aldrich, CAS: 134-03-2

NP40 alternative: Sigma-Aldrich, CAS: 9016-45-9

Triton X-100: Sigma-Aldrich, CAS: 9002-93-1

Urea: Fisher scientific, CAS: 57-13-6

Centrifuge: Sorvall Legend RT and Fisher Scientific, Accuspin micro 17

Typhoon FLA 9500: Ge Healthcare

Silverstain kit: SilverQuest Invitrogen, LC6070

Formic acid: Sigma-Aldrich, CAS: 64-18-6

DADPS biotin alkyne: Click chemistry tools, #1331

Dde TAMRA biotin Alkyne: Click chemistry tools, #1368

Hydrazine: Sigma-Aldrich, CAS: 10217-52-4

Pierce High Capacity Streptavidin Agarose: Thermo Scientific, #20359

Precast acrylamide gel: 4-12% Criterion XT Bis-Tris Protein Gel from Biorad (3450123)

Chemoproteomics procedure with the UM171 probes

The procedure was inspired from this paper: MacKinnon, A. L.; Taunton, J., Target Identification by Diazirine Photo-Cross-Linking and Click Chemistry. *Current Protocols in Chemical Biology* **2009**, *1* (1), 55-73

By the same procedure described in section **Cell culture**, 600 M of AML5 cells were cultured. The cells were combined and centrifuged at 1000 rpm. The idea was to have a volume of 20 mL per condition, so a total volume of 120 mL. Thus, each condition was assigned 50 M cells in 20 mL, in 50 mL petri dishes. The conditions are as follows:

	1	2	3	4	5	6	7	8	9	10	11	12
Compound	DMSO	DMSO	941	941	942	942	943	943	944	944	945	945
Condition	DMSO	MG132	DMSO	MG132	DMSO	MG132	DMSO	MG132	DMSO	MG132	DMSO	MG132
Number of cells	50 M	50 M	50 M	50 M	50 M	50 M	50 M	50 M	50 M	50 M	50 M	50 M

Because the lamp does not allow more than 2 petris to be irradiated at a time, it was necessary to start incubations of the probes off-lines. To do this, the incubations were carried out with a 40 minutes interval. The incubation sequence was the same for each condition. Initially, 15 μ L of MG132 (10 mM) or DMSO dissolved in 5 mL of DMEM is added. This incubation is carried out for 2 hours, before the addition of 15 μ L of probe (10 mM) or DMSO which is also dissolved in 5 mL of DMEM. The incubation of the probe takes place over a period of 4 hours with a 5 % CO₂ atmosphere and a temperature of 37 °C in the petri dishes. Once the probe incubation was completed, the petris were irradiated for 30 minutes on ice. The lamp was as close as possible to the sample. The cells were transferred to 50 mL falcons and centrifuged at 1000 rpm for 5 minutes. The supernatant was removed, then the cells were transferred to 2 mL Eppendorf with PBS 1x. Cells were spin down another time to remove the PBS. Subsequently, cells were lysed with 750 μ L of a solution of 8 M urea in 50 mM HEPES. Protein quantification was performed with the Bio-Rad kit, giving an average concentration of 5 mg/mL for each sample. In order to have a concentration of 1.5 mg/mL of protein for each sample there was addition of 1.5 mL of PBS, consequently the diluted lysates were transferred into 15 mL tubes. For each sample, reagents were added to perform the Click reaction. So, 25 μ L of CuSO₄ (20 mM), 125.2 μ L of NaAsc (100 mM), 45 μ L of THPTA (50 mM) and 12 μ L of the DADPS alkyne cleavable biotin (10 mM). The reactions were mixed on the thermomixer during 16 hours at 4°C. Reactions were stopped by adding 11 mL of methanol, 2.5 mL of chloroform and 5 mL of water. Theses mixtures were spin at 4000 rpm at 4°C for 15 minutes, then the upper layer was removed. Another 11 mL of MeOH was added and tubes were spin during 30 minutes at 15 000 rpm. The supernatant was removed, and the precipitate was dry with air to remove the methanol. Thereafter the proteins were resolubilized in 1.5 mL of 0.5 % SDS in PBS (20 minutes) and transferred to 2 mL tubes. For each tube, 150 μ L of streptavidin beads was added. The latter were stirred slowly for 16 hours at 4°C with the protein's media. The next day, the supernatant was removed from each sample and each bead was washed multiples times: 3 times with a solution of 1% SDS (3x500 μ L) and 4 times with PBS (4x500 μ L). The supernatant was removed and 500 μ L of a 4 % formic acid solution was added to all samples. The beads were stirred for 2 hours at 37°C and the supernatant was gently removed and transferred to a second tube. The cleavage process was repeated a second time. Supernatants from the same samples were combined and the liquid was evaporated with the SpeedVac. Proteins were resolubilized in the loading buffer 1x, heated at 95°C for 8 minutes and loaded on a precast acrylamide gel with gradient 4-12%. The gel was revealed with silver staining

following the procedure provided by the retailer. Subsequently, interesting proteins bands were cut into small cubes with a blade and put into new tubes. The cubes have undergone the manufacturer's destain procedure to become translucent. Then, the small cubes are deshydrated with 2 washes with ACN (200 uL) for 10 minutes each time. Once the pieces had turned white, an alkylation solution (200 uL) composed of 5 mM TCEP and 1.5 mM Chloroacetamide was added to each sample. Thereafter, the small cubes are deshydrated another times with 2 washes with ACN (200 uL). Finally, the gel pieces were digested with 1 ug of trypsin dissolved in 50 mM ABC buffer for 16 hours at 37°C on the thermomixer. The next day, the supernatant was put in another tubes, and the gel pieces were dehydrated two times with 200 uL of ACN. Supernatants from the same sample were combined and dried out on the SpeedVac. The peptides were solubilised in 20 uL of 4% FA in water and injected on the Exploris 480 from Thermo.

Parameter of mass spectrometry: Trypsic peptides were loaded on C18 precolumn and separated with a home-made C18 column (150 µm ID x 15 cm) made of Jupiter C18, 3 µm, 300 Å. This separation was possible with a linear gradient of 5-40% of acetonitrile (0.2 % Formic acid) in a 50 min with a flow rate of 600 nL/min. The liquid chromatography was coupled with the Exploris 480 , with a mass scan range of 300-1200 m/z. The resolution selected was 240 000 and the AGC target values for MS was set to 5e6 (max fill time 50 ms) and for MS/MS it was set to 5e4 (max fill time of 750 ms). The neutral energy of collision was fixed to 25 and with a charge state between 2 and 6. The database searches were doned unsing the Peak software against the Uniprot human database. The parametters selected were 10 ppm and 0.01 Da for the error windows for the precursor and fragment ion mass, respectively. The maximum mass clivage allowed was set to 2 and the PTM selected were phosphorylation, oxidation, deamidation and carbamidomethylation. The heat map was drawn by the Origin software.

Chemicals and materials:

Petri dishes: Fisher Brand, FB012924

DMEM: Wisent 319-051-CL

MG132: Sigma-Aldrich, CAS: 133407-82-6

DMSO: Fisher Scientific, CAS: 67-68-5

PBS 1x: Wisent 311-010-CL

Uv lamp : UVP High Intensity UV Lamps B-100AP Analytik Jena

MeOH: Sigma-Aldrich, CAS : 67-56-1

CHCl₃: Sigma-Aldrich, CAS: 67-66-3

CuSO₄ : Sigma-Aldrich, CAS: 7758-98-7

THPTA: Sigma-Aldrich, CAS:760952-88-3
Sodium L-Ascorbate: Sigma-Aldrich, CAS: 134-03-2
NP40 alternative: Sigma-Aldrich, CAS: 9016-45-9
Triton X-100: Sigma-Aldrich, CAS: 9002-93-1
Urea: Fisher scientific, CAS: 57-13-6,
HEPES: Sigma-Aldrich, CAS: 7365-45-9
Bio-Rad Protein Assay: Bio-Rad, 5000006
96-well plate clear flat bottom: Corning, 353072
Centrifuge: Sorvall Legend RT and Fisher Scientific, Accuspin micro 17
Silverstain kit: SilverQuest Invitrogen, LC6070
Formic acid: Sigma-Aldrich, CAS: 64-18-6
DADPS biotin alkyne: Click chemistry tools, #1331
10 % SDS solution: Wisent, 880-550-CL
Pierce High Capacity Streptavidin Agarose: Thermo Scientific, #20359
Precast acrylamide gel: 4-12% Criterion XT Bis-Tris Protein Gel from Biorad (3450123)
Sequencing Grade Modified Trpsin: Promega, V5111
ACN: Fisher Scientific, CAS: 75-05-8