



**HAL**  
open science

# Expression, purification et cristallisation de l'aminopeptidase-N humaine (APN ou CD13) : évaluation *in vitro* et *in vivo* d'inhibiteurs sélectifs

Céline Schmitt

► **To cite this version:**

Céline Schmitt. Expression, purification et cristallisation de l'aminopeptidase-N humaine (APN ou CD13) : évaluation *in vitro* et *in vivo* d'inhibiteurs sélectifs. Autre. Université de Haute Alsace - Mulhouse, 2012. Français. NNT : 2012MULH4107 . tel-00812493

**HAL Id: tel-00812493**

**<https://theses.hal.science/tel-00812493>**

Submitted on 12 Apr 2013

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

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

Université de Haute-Alsace  
Ecole Doctorale Jean-Henri Lambert (ED 494)

## THESE

Présentée pour obtenir le grade de

**Docteur de l'Université de Haute-Alsace**

Discipline : Chimie Biologie

**Céline SCHMITT**

Expression, purification et cristallisation de  
l'Aminopeptidase-N humaine (APN ou CD13) ;  
évaluation *in vitro* et *in vivo* d'inhibiteurs sélectifs

EA 4566

Laboratoire de Chimie Organique et Bioorganique  
Ecole Nationale Supérieure de Chimie de Mulhouse

Soutenue publiquement le **18 septembre 2012** devant la commission d'examen :

**Dr. Dominique GUENOT**

**Dr. Isabelle FLORENT**

**Dr. Marc RUFF**

**Dr. Julie CHONG**

**Dr. James DUNFORD**

**Pr. Céline TARNUS**

Président

Rapporteur

Rapporteur

Examineur

Membre invité

Directeur de thèse



*A ma famille,*

*A Christophe,*



# REMERCIEMENTS

Ce manuscrit représente le travail effectué au sein du laboratoire Chimie Organique et Bioorganique de Mulhouse sous la direction du Pr. Céline Tarnus.

Je tiens à remercier le Pr. Céline Tarnus pour sa confiance et son soutien, ainsi que pour l'ensemble des précieux conseils qu'elle a bien voulu me prodiguer durant ces années.

J'exprime ma profonde gratitude au Dr. Isabelle Florent, ainsi qu'au Dr. Marc Ruff, qui ont accepté d'être rapporteur de ce travail.

Toute ma reconnaissance va au Dr. Dominique Guenot, qui me fait l'honneur de juger ce travail et de présider ce jury.

Je remercie également le Dr. Julie Chong pour le temps précieux qu'elle a accordé à la lecture de ce manuscrit et pour sa participation à ce jury.

I sincerely thank Dr. James Dunford for having accepted to be a member of this examining board; I am profoundly grateful to him for taking part into my PhD defence.

Que le Dr. Olivier Zinc trouve ici l'expression de ma profonde gratitude pour l'aide qu'il m'a apportée lors de l'élaboration des stratégies d'expression de la protéine recombinante chez la bactérie et la levure.

Un grand merci à toute l'équipe de l'Institut de Recherche en Hématologie et Transplantation de Mulhouse et tout particulièrement Monsieur le Professeur Philippe Hénon pour l'accueil au sein de son laboratoire : Stéphanie, Céline et Antoinette, merci pour tous vos conseils et votre bonne humeur.

I am deeply grateful to Pr. Udo Oppermann: I was made to feel very welcome in your lab! Both human and professional encounters were so rewarding and fulfilling during these 3 months! Thank you so much James, Carina, Peter, Grazyna and all the members of the team!

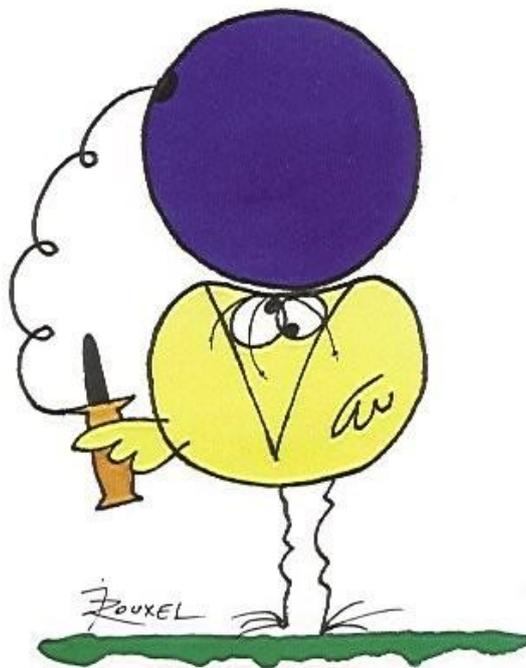
Un merci tout particulier à Emel et Mira pour leur amitié, et pour tous les bon moments passés ensemble depuis le début de nos thèses. Merci à Sandrine, Sarah, Aurélie, Hamid et Lionel pour les moments de détente bien mérités.

Un grand merci à toute l'équipe du laboratoire COB pour leur soutien et leur sympathie qui ont rendu ces dernières années si agréables.

En dernier lieu, je remercie les fondations *Pierre et Jeanne Spiegel* et *Alliance CAIRPSA-CARPRECA*, l'Ecole Nationale Supérieure de Chimie de Mulhouse ainsi que la Ligue nationale contre le Cancer qui ont subventionné ce travail.



## Les devises Shadok



EN ESSAYANT CONTINUUELLEMENT  
ON FINIT PAR RÉUSSIR. DONC:  
PLUS ÇA RATE, PLUS ON A  
DE CHANCES QUE ÇA MARCHE.



# SOMMAIRE

## LISTE DES ABREVIATIONS

## INTRODUCTION GENERALE 1

## INTRODUCTION BIBLIOGRAPHIQUE 5

- 1. L'angiogenèse : une cible thérapeutique contre le cancer 7**
  - 1.1. Informations générales sur le cancer 7
  - 1.2. L'angiogenèse 9
    - 1.2.1. Généralités 9
    - 1.2.2. L'angiogenèse tumorale 11
    - 1.2.3. Le mécanisme de l'angiogenèse tumorale 13
  - 1.3. Les thérapies ciblées anti-angiogéniques 15
    - 1.3.1. Le mécanisme d'action des agents anti-angiogéniques : application en multithérapies 15
    - 1.3.2. Le VEGF : une cible de choix 17
    - 1.3.3. Données cliniques 21
- 2. L'Amino peptidase-N (APN ou CD13) : une cible potentielle 23**
  - 2.1. L'APN : une protéine multifonctionnelle 23
  - 2.2. L'APN : son implication dans la tumorigenèse. 25
- 3. Conception d'inhibiteurs spécifiques de l'Amino peptidase-N 29**
  - 3.1. Les enzymes protéolytiques 29
  - 3.2. Les amino peptidases 31
  - 3.3. La famille M1 : les amino peptidases à un ion zinc essentiel à la catalyse 31
    - 3.3.1. Homologie séquentielle et structurale 31
    - 3.3.2. Mécanisme catalytique et spécificité de substrat 37
    - 3.3.3. L'Amino peptidase-N de mammifères 41
  - 3.4. La famille M17 : les amino peptidases à deux ions zinc 43
    - 3.4.1. La Leucine Amino peptidase cytosolique 43
    - 3.4.2. L'Amino peptidase d'*Aeromonas Proteolytica* 45
  - 3.5. Les inhibiteurs de l'Amino peptidase-N 47
    - 3.5.1. Produits naturels 47
    - 3.5.2. Les inhibiteurs synthétiques 51
    - 3.5.3. Travaux du laboratoire 55

## CHAPITRE 1: ÉVALUATION *IN VITRO* ET *IN VIVO* D'INHIBITEURS PUISSANTS ET SELECTIFS DE L'AMINOPEPTIDASE-N/CD13: INTERET EN ONCOLOGIE 59

- 1. Publication n°1 : A novel amino-benzosuberone derivative is a picomolar inhibitor of mammalian amino peptidase N/CD13 61**
- 2. Publication n°2 : Selective Amino peptidase-N (CD13) – inhibitors of relevance to cancer chemotherapy 81**

## **CHAPITRE 2: Etudes structurales de l'Amino-peptidase-N d'*Escherichia coli* (pepN) en complexe avec les dérivés de la 7-amino-6-benzosubérone** **115**

<b>1. Introduction</b>	<b>117</b>
<b>2. Résultats et discussion</b>	<b>119</b>
2.1. Expression, purification et cristallisation de l'APN d' <i>E. coli</i>	119
2.2. Evaluation <i>in vitro</i> des inhibiteurs	119
2.3. Analyse structurale de l'APN d' <i>E. coli</i> avec les dérivés de la 7-amino-6-benzosubérone	121
i. Conformation de la 7-amino-6-benzosubérone et stéréochimie	121
ii. Etude des relations structure-activité observées	121
iii. Superposition avec la bestatine (analogue de peptide)	123
<b>3. Conclusion</b>	<b>125</b>
<b>4. Matériels et Méthodes</b>	<b>127</b>
4.1. Expression et purification de l'APN d' <i>E. coli</i> (pepN)	127
4.2. Cristallisation	128
4.3. Acquisition des données, résolution et affinement de la structure	129
4.4. Cinétique enzymatique	129

## **CHAPITRE 3: Expression, purification et cristallisation de l'Amino-peptidase-N humaine** **131**

<b>INTRODUCTION</b>	<b>133</b>
<b>RESULTATS ET DISCUSSION</b>	<b>139</b>
<b>1. Expression de la protéine recombinante chez la levure</b>	<b>139</b>
1.1. Clonage	141
1.1.1. Isolement du gène d'intérêt	141
1.1.2. Clonage dans le vecteur d'expression pKLAC2	143
1.1.3. Séquençage de l'ADN	147
1.2. Intégration de la cassette d'expression dans le génome de <i>K. lactis</i>	149
1.3. Tests d'expression de la protéine recombinante	149
1.4. Conclusion et discussion	151
1.4.1. Dégradation par des enzymes protéolytiques	151
1.4.2. Biais d'usage du code	153
<b>2. Expression de la protéine recombinante chez <i>Escherichia coli</i></b>	<b>155</b>
2.1. Clonage	155
2.1.1. Isolement du gène d'intérêt	155
2.1.2. Clonage dans le vecteur d'expression pET-22b(+)	155
2.2. Culture des bactéries et expression de la protéine recombinante	157
2.2.1. Choix des souches bactériennes	157
2.2.2. Expression de la protéine recombinante	159
2.3. Renaturation de la protéine recombinante	163
2.3.1. Purification et solubilisation des corps d'inclusion	163
2.3.2. Renaturation <i>in vitro</i> de la protéine recombinante	163
i. Renaturation par dilution	165

ii.	Renaturation par dialyse	165
2.4.	Purification de la protéine recombinante	167
2.4.1.	Chromatographie d'affinité sur résine de nickel	167
2.4.2.	Chromatographie échangeuse d'anions	169
2.4.3.	Chromatographie d'exclusion	171
2.5.	Conclusion et discussion	171
<b>3.</b>	<b>Expression de la protéine recombinante dans baculovirus</b>	<b>175</b>
3.1.	Clonage	175
3.1.1.	Isolement du gène d'intérêt	177
3.1.2.	Clonage dans les vecteurs baculovirus	177
3.2.	Génération des bacmides recombinants	179
3.3.	Culture cellulaire et expression de la protéine recombinante	179
3.3.1.	Transfection des cellules d'insectes Sf9	179
3.3.2.	Amplification du stock viral	181
3.3.3.	Expression de la protéine recombinante	181
3.4.	Purification de la protéine recombinante	183
3.4.1.	Chromatographie d'affinité	183
3.4.2.	Purification de la protéine hAPN <sub>59-967</sub>	183
3.5.	Caractérisation de la protéine recombinante	187
3.5.1.	Détermination des paramètres cinétiques $K_M$ et $k_{cat}$	187
3.5.2.	Détermination de $K_i$	187
3.6.	Conclusion et discussion	189
<b>4.</b>	<b>Essais de cristallisation</b>	<b>189</b>
	<b>MATERIELS ET METHODES</b>	<b>191</b>
<b>1.</b>	<b>Expression de la protéine recombinante chez la levure</b>	<b>191</b>
1.1.	Principe	191
1.2.	Clonage	194
1.2.1.	Extraction des ARN totaux d'une culture cellulaire	194
1.2.2.	Transcription inverse de l'ARN en ADN complémentaire (ADNc)	196
1.2.3.	Amplification du gène d'intérêt (entier ou tronqué) par PCR	196
1.2.4.	Electrophorèse sur gel d'agarose	197
1.2.5.	Purification du produit de PCR (extraction sur gel d'agarose)	198
1.2.6.	Quantification de l'acide désoxyribonucléique	198
1.2.7.	Digestion enzymatique du produit de PCR	199
1.2.8.	Purification du fragment d'ADN digéré	199
1.2.9.	Digestion enzymatique du vecteur pKLAC2	199
1.2.10.	Déphosphorylation du vecteur pKLAC2	200
1.2.11.	Ligation cohésive par l'ADN ligase du phage T4	200
1.2.12.	Transformation des souches bactériennes XL10 Gold par choc thermique	201
1.2.13.	Criblage des bactéries transformées par PCR	201
1.2.14.	Extraction et purification de l'ADN plasmidique produit par les bactéries : midiprep	203
1.2.15.	Contrôle de l'intégrité de l'ADN plasmidique par PCR	203
1.2.16.	Séquençage de l'ADN	204
1.3.	Intégration de la cassette d'expression dans le génome de <i>K. lactis</i>	205
1.3.1.	Linéarisation du plasmide recombinant	205
1.3.2.	Transformation des cellules compétentes GG799 de <i>K. lactis</i>	205
1.3.3.	Analyse des levures transformées par PCR	206

1.4.	Tests d'expression de la protéine recombinante	208
1.5.	Analyse de l'expression de la protéine recombinante	208
1.5.1.	Electrophorèse sur gel de polyacrylamide (SDS-PAGE)	208
1.5.2.	Transfert de protéines et immunodétection (Western Blot)	209
<b>2.</b>	<b>Expression de la protéine recombinante chez <i>Escherichia coli</i></b>	<b>210</b>
2.1.	Principe	210
2.2.	Clonage	212
2.3.	Culture des bactéries et expression de la protéine recombinante	213
2.3.1.	Souches bactériennes	213
2.3.2.	Transformation par choc thermique dans les souches d'expression	214
2.3.3.	Optimisation des conditions d'expression de la protéine recombinante	214
2.3.4.	Production à grande échelle	215
<b>3.</b>	<b>Expression de la protéine recombinante dans baculovirus</b>	<b>216</b>
3.1.	Principe	216
3.1.1.	Système d'expression Bac-to-Bac®	216
3.1.2.	Clonage sans ligase	217
3.1.3.	Plasmides	218
3.2.	Clonage	221
3.2.1.	Amplification de l'insert par PCR	221
3.2.2.	Digestion enzymatique du vecteur	222
3.2.3.	Traitement par l'ADN polymérase du phage T4.	223
3.2.4.	Ligation et transformation des bactéries XL1-Blue E. coli par choc thermique	223
3.2.5.	Criblage des bactéries transformées par PCR	224
3.2.6.	Extraction et purification de l'ADN plasmidique produit par les bactéries	225
3.3.	Génération des bacmides recombinants	225
3.3.1.	Transformation des bactéries E.coli DHBac10 par choc thermique	225
3.3.2.	Isolement et purification des bacmides recombinants	226
3.3.3.	Analyse des bacmides recombinants par PCR	227
3.4.	Culture cellulaire et expression de la protéine recombinante	228
3.4.1.	Culture des cellules	228
3.4.2.	Transfection des cellules d'insectes Sf9	229
3.4.3.	Amplification du stock viral	229
3.4.4.	Expression de la protéine recombinante	230
<b>4.</b>	<b>Purification de la protéine recombinante</b>	<b>231</b>
4.1.	Extraction de la protéine	231
4.1.1.	Protéine exprimée sous forme de corps d'inclusion dans le cytoplasme des bactéries	231
i.	Isolement et purification des corps d'inclusion	231
ii.	Solubilisation des corps d'inclusion	232
iii.	Renaturation de la protéine par dilution	232
iv.	Renaturation de la protéine par dialyse (Protein Refolding Kit, Novagen)	233
4.1.2.	Protéine exprimée sous forme soluble dans le cytoplasme des cellules d'insectes	234
4.2.	Purification de la protéine recombinante	234
4.2.1.	Suivi de purification	234
i.	Dosage des protéines	234
ii.	Activité enzymatique	235
4.2.2.	Dessalage et concentration par ultrafiltration	236
4.2.3.	Précipitation différentielle au sulfate d'ammonium	237

4.2.4. Chromatographie d'affinité	238
i. Résine de nickel	238
ii. Résine anti-FLAG	239
4.2.5. Chromatographie échangeuse d'ions	240
4.2.6. Chromatographie à interactions hydrophobes (HIC)	241
4.2.7. Chromatographie d'exclusion	241
4.3. Caractérisation de la protéine	242
4.3.1. Détermination des paramètres cinétiques $K_M$ et $k_{cat}$	242
4.3.2. Détermination de $K_i$	243
<b>5. Essais de cristallisation</b>	<b>245</b>
5.1. Principe	245
5.2. Les essais de cristallisation de l'aminopeptidase-N humaine	246
<b>CONCLUSION GENERALE</b>	<b>247</b>
<b>ANNEXES</b>	<b>253</b>



# LISTE DES ABREVIATIONS

(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	Sulfate d'ammonium
Å	Angström
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
<i>amdS</i>	Gène de l'acétamidase
ANG	Angiopoïétine
AP	Aminopectidase
<i>Apaero</i>	Aminopectidase d' <i>Aeromonas proteolytica</i>
APN	Aminopectidase-N
ApR	Résistance à l'ampicilline
Arg	Arginine
ARN	Acide ribonucléique
ARNm	ARN messenger
ARNt	ARN de transfert
AS	Activité spécifique
Asp	Acide aspartique
BET	Bromure d'éthidium
bFGF	Facteur de croissance des fibroblastes basiques
b-HLP-PAS	« basic helic loop helic-Per/Arnt/Sim »
BSA	Albumine de sérum bovin
CD13	Cluster de différenciation 13
CIP	Phosphatase alcaline de l'intestin de veau
ColAP	« Cold-Active » Aminopectidase
CoV	Coronavirus respiratoire de porc
Cys	Cystéine
dATP	Désoxyadénosine triphosphate
DNase I	Désoxyribonucléase I
dNTP	Mélange des 4 désoxyribonucléotides triphosphates
DO	Densité optique
DTT	Dithiothréitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Acide éthylène diamine tétraacétique
ERAP1 ou 2	Aminopectidase du réticulum endoplasmique 1 ou 2
FDA	Agence fédérale américaine des produits alimentaires et médicamenteux
Gly	Glycine
h-APN	Aminopectidase-N humaine
HCl	Acide chlorhydrique
HCMV	Cytomégalovirus humain
HER	Récepteur du facteur de croissance épidermique humain
HIF	Facteur induit par l'hypoxie
HRP	« Horseradish Peroxidase » ou peroxydase de raiford
HUVEC	Cellules endothéliales humaines de veine ombilicale
IC <sub>50</sub>	Concentration inhibitrice médiane
IgG	Immunoglobuline
<i>K. lactis</i>	<i>Kluyveromyces lactis</i>
KCl	Chlorure de potassium
Ki	Constante d'inhibition

K <sub>M</sub>	Constante de Michaelis-Menten
LAPc	Leucine Aminopeptidase cytosolique
LB	Luria Broth
LDS	Dodécylsulfate de lithium
Leu	Leucine
LTA4H	Leucotriène A4 hydrolase
Lys	Lysine
MBP	« Maltose binding protein »
MCS	Site de clonage multiple
MetAP2	Méthionine Aminopeptidase de type 2
MgCl <sub>2</sub>	Chlorure de magnésium
MgSO <sub>4</sub>	Sulfate de magnésium
M-MLV	Moloney Murine Leukemia Virus
MMP	Métallo-protéases matricielles
MOPS	Acide 3-(N- morpholino)- propane sulfonique
mTOR	« mammalian Target Of Rapamycin »
NaCl	Chlorure de sodium
NICE	Institut national pour la santé et l'excellence clinique (Royaume-Uni)
NOD/SCID	« no obese diabetic severe combined immunodeficient »
PAGE	Electrophorèse sur gel de polyacrylamide
pb	Paire de base
PBS	Tampon phosphate salin
PCR	Réaction de polymérisation en chaîne
PDGF	Facteur de croissance dérivé des plaquettes
pepN	Aminopeptidase-N d' <i>Escherichia coli</i>
PfA-M1	M1 Aminopeptidase neutre de <i>Plasmodium falciparum</i>
pl	Point isoélectrique
PNGase F	Peptide N Glycosidase F
Pro	Proline
PVDF	Polyfluorure de vinylidène
RE	Réticulum endoplasmique
RNases	Ribonucléases
RT	Transcriptase inverse
SDS	Dodécylsulfate de sodium
Ser	Sérine
siRNA	petit ARN interférent
β-ME	β-Mercaptoethanol
TBE	Tris, Borate, EDTA
TGF	Facteur de croissance transformant
TIFF3	« Tricorn interacting factor F3 »
TNFα:	Facteur de nécrose tumorale α
TSP-1	Trombospondine-1
UV	Ultra-violet
Val	Valine
VEGF	Facteur de croissance de l'endothélium vasculaire
α-MF	« α-mating factor »

# INTRODUCTION GENERALE



Les progrès considérables réalisés dans la compréhension des mécanismes moléculaires à l'origine de la néo-vascularisation tumorale ont permis le développement, dans les années 2000, des premiers médicaments anti-angiogéniques. Bien que ces thérapies ciblées soient largement utilisées aujourd'hui en oncologie, leur réponse et leur efficacité restent imprévisibles devant la grande diversité des tumeurs.

Les recherches dans ce domaine doivent se poursuivre, car il faut identifier les marqueurs prédictifs de réponse à ces thérapies anti-angiogéniques et développer de nouvelles molécules ciblées plus performantes afin de pouvoir proposer des traitements spécifiques adaptés à chaque patient.

Dans ce contexte, les travaux du laboratoire se sont centrés sur la conception et l'évaluation d'inhibiteurs particulièrement puissants et sélectifs de l'Aminopeptidase-N (APN ou CD13). Cette enzyme protéolytique, surexprimée par les cellules endothéliales lors de la néo-vascularisation tumorale mais également par certains types de cellules tumorales, est décrite comme intervenant dans la morphogenèse et la migration des cellules endothéliales, ainsi que dans la migration des cellules tumorales.

Dans le cadre de ce travail, l'efficacité des inhibiteurs de l'APN synthétisés au laboratoire a été évaluée *in vitro* et *in vivo* dans des modèles d'étude de l'angiogenèse et de la cancérogenèse.

L'objectif à terme étant de co-cristalliser ces inhibiteurs avec l'APN humaine pour en élucider le mode de liaison, il est indispensable de résoudre la structure tridimensionnelle de cette enzyme multifonctionnelle. Dans cette optique, différentes stratégies ont été développées pour exprimer, purifier et cristalliser cette métallo-aminopeptidase.

En parallèle à ces travaux difficiles, et puisque le domaine catalytique des membres de cette famille est conservé, les inhibiteurs ont été co-cristallisés avec l'APN d'*Escherichia coli* (pepN), dont la structure est connue.



## INTRODUCTION BIBLIOGRAPHIQUE



# 1. L'angiogenèse : une cible thérapeutique contre le cancer

## 1.1. Informations générales sur le cancer

Un cancer correspond à une multiplication anarchique de cellules qui échappent aux mécanismes de régulation cellulaire. Cette prolifération anormale donne naissance à une tumeur, qui en grossissant envahit et détruit les tissus sains avoisinants. Les cellules cancéreuses peuvent également migrer de la tumeur initiale vers d'autres tissus et organes, et s'y implanter pour former une nouvelle tumeur : on parle alors de métastases. En détruisant son environnement, le cancer peut devenir un réel danger pour la survie de l'être vivant.

La lutte contre le cancer constitue un enjeu majeur de santé publique. Selon l'Institut de veille sanitaire, le nombre de nouveaux cas diagnostiqués en France au cours de l'année 2011 a été estimé à 365 500, et le nombre de décès associé à 147 500<sup>1</sup>. Le taux de mortalité par cancer est en recul depuis une décennie, et ce, malgré une progression continue de la maladie. Cette diminution s'explique vraisemblablement par l'amélioration de la prévention, des outils de diagnostic et de l'efficacité des traitements disponibles.

Les traitements contre le cancer associent différentes méthodes : la chirurgie, la radiothérapie, et la chimiothérapie cytotoxique en constituent les principales. Les agents cytotoxiques ne sont malheureusement pas spécifiques aux cellules cancéreuses ; ils affectent également les cellules saines en division, et sont, par conséquent, responsables de nombreux effets secondaires.

La compréhension des mécanismes moléculaires impliqués dans la progression tumorale, a permis, ces dernières années, de concevoir de nouvelles stratégies thérapeutiques dirigées contre des cibles moléculaires spécifiques (récepteurs, gènes ou protéines) : ce sont les thérapies ciblées<sup>2,3</sup>. Ces dernières restent, néanmoins, délicates à mettre en place en raison de la grande diversité des cancers.

---

<sup>1</sup> Projection de l'incidence et de la mortalité par cancer en France en 2011 – Synthèse des résultats. Saint-Maurice: Institut de veille sanitaire; 2011. 2 p. <http://www.invs.sante.fr/>

<sup>2</sup> Christoffersen, T., Guren, T. K., Spindler, K. L., Dahl, O., Lønning, P. E., Gjertsen, B. T. (2009). *Eur. J. Pharmacol.* **625**; 6-22.

<sup>3</sup> Rosa, D. D., Ismael, G., Dal Lago, L., and Awada, A. (2008). *Cancer Treat. Rev.* **34**, 61-80.



Les thérapies ciblées ont différents modes d'action qui visent principalement à inhiber :

- la multiplication des cellules cancéreuses en interférant avec les signaux qui stimulent la survie et la prolifération cellulaire (inhibiteurs de HER (récepteurs 1 et 2 du facteur de croissance épidermique humain) : trastuzumab (Herceptin®)<sup>4</sup> ou cétuximab (Erbix®)<sup>5</sup>).
- l'angiogenèse tumorale (anti-angiogéniques : bévacizumab (Avastin®)<sup>6</sup>).

De par leur action ciblée, elles devraient être moins toxiques et entraîner moins d'effets secondaires.

## 1.2. L'angiogenèse

### 1.2.1. Généralités

L'angiogenèse est un mécanisme finement régulé aboutissant à la formation de nouveaux capillaires sanguins par excroissance ou bourgeonnement de vaisseaux préexistants. Elle est particulièrement importante et indispensable au cours de nombreux processus physiologiques tels que l'embryogenèse, la régénération de l'endomètre (cycle menstruel féminin), ou la cicatrisation.

L'angiogenèse est également une composante fondamentale de nombreuses pathologies (la polyarthrite rhumatoïde, le psoriasis)<sup>7</sup> ; son implication dans la croissance tumorale et le développement des métastases est bien décrite et en fait un élément déterminant du pronostic des cancers<sup>8,9</sup>.

Le contrôle de l'angiogenèse s'exerce par le biais de nombreux facteurs endogènes pro ou anti-angiogéniques produits en quantités plus ou moins importantes par les cellules. En général, dans un organisme sain, la production d'inhibiteurs est plus importante que celle d'activateurs ; le phénomène d'angiogenèse est alors bloqué. Cependant, dans certaines conditions (cicatrisation, cycle menstruel féminin, prolifération tumorale), l'angiogenèse est

---

<sup>4</sup> Goel, S., Chirgwin, J., Francis, P., Stuart-Harris, R., Dewar, J., Mileskin, L., Snyder, R., Michael, M., and Koczwara, B. (2011). *The Breast*. **20**, 101-110.

<sup>5</sup> Wong, S-F. (2005). *Clin. Therap.* **27** (6), 684-694.

<sup>6</sup> Shih, T, and Lindley, C. (2006). *Clin. Therap.* **28** (11), 1779-1802.

<sup>7</sup> Pandya, N. M., Dhalla, N. S., and Santani, D. D. (2006). *Vasc. Pharmacol.* **44**, 265-274.

<sup>8</sup> Folkman, J. (2007). *Nat. Rev. Drug Discov.* **6**, 273-285.

<sup>9</sup> Bridges, E. M., and Harris, A. L. (2011). *Biochem. Pharmacol.* **81**, 1183-1191.



déclenchée localement suite à une modification de la balance entre les facteurs pro- et anti-angiogéniques (augmentation de la production d'activateurs ou diminution de la quantité d'inhibiteurs disponibles)<sup>10</sup>.

### 1.2.2. L'angiogenèse tumorale

C'est J. Folkman (1933-2008) qui, en 1971, a émis l'hypothèse que la croissance tumorale était dépendante de la néo-vascularisation<sup>11</sup>. Les cellules cancéreuses peuvent, de manière autonome, former des tumeurs de taille relativement importante. Néanmoins, au-delà de quelques mm<sup>3</sup>, et en l'absence de vascularisation, la vitesse de diffusion de l'oxygène et des nutriments limite leur croissance : le centre de la tumeur est alors soumis à une hypoxie qui s'accompagne d'un phénomène de nécrose. Les cellules cancéreuses répondent à cette hypoxie en sécrétant des facteurs pro-angiogéniques comme le VEGF (facteur de croissance de l'endothélium vasculaire). Ces derniers vont activer les cellules endothéliales des vaisseaux sanguins avoisinants et déclencher la néo-vascularisation tumorale.

L'établissement d'un phénotype angiogénique, également appelé « switch » angiogénique, est une étape clé dans la progression vers les maladies malignes. Ce processus résulte d'une rupture de l'équilibre entre activateurs et inhibiteurs de l'angiogenèse, avec une bascule en faveur des activateurs angiogéniques<sup>10</sup>.

De nombreux facteurs environnementaux et génétiques sont responsables de l'activation de ce processus :

- l'hypoxie intra-tumorale : la réponse à l'hypoxie est principalement régulée par le facteur de transcription HIF (facteur induit par l'hypoxie) ; il s'agit d'un hétérodimère de la famille des protéines bHLH-PAS (« basic helic loop helic-Per/Arnt/Sim »), qui est constitué d'une sous-unité régulatrice  $\alpha$  et d'une sous-unité constitutive  $\beta$ <sup>12</sup>. L'isoforme HIF-1 $\alpha$ , qui est stabilisé en hypoxie, joue un rôle clé dans l'angiogenèse tumorale via l'induction de la transcription de facteurs pro-angiogéniques tels que le VEGF, le PDGF (facteur de croissance dérivé des plaquettes), le TGF $\alpha$  (le facteur de croissance transformant  $\alpha$ ) et l'angiopoïétine-2 (ANG-2)<sup>13</sup>.

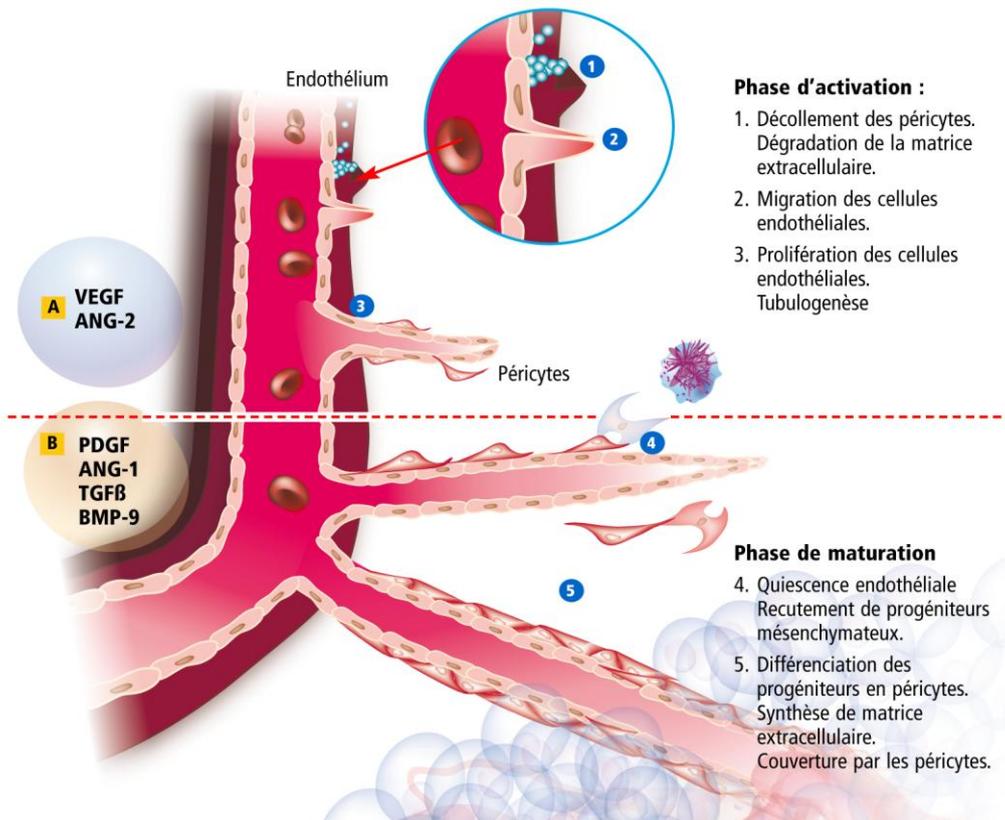
---

<sup>10</sup> Carmeliet, P., and Jain, R. K. (2000). *Nature*. **407**, 249-257.

<sup>11</sup> Folkman, J. (2007), *J. Pediatr. Surg.* **42**, 1-11.

<sup>12</sup> Liao, D., and Johnson, R. S. (2007). *Cancer Metastasis Rev.* **26**, 281-290.

<sup>13</sup> Hickey, M. M., and Simon, M. C. (2006). *Curr. Top. Dev. Biol.* **76**, 217-257.



*Figure 1 : Événements cellulaires majeurs et acteurs moléculaires des différentes phases de l'angiogénèse<sup>14</sup>.*

<sup>14</sup> Feige, J-J. (2009). Collection Thérapie ciblée des cancers, Volume 3- L'Angiogenèse. John Libbey Eurotext et Amgen SAS. pg 2-17.

- l'activation d'oncogènes (par exemple Ras) et/ou l'inactivation de gènes suppresseurs de tumeur (comme p53) ont pour conséquence une augmentation de l'expression de facteurs de croissance comme le VEGF, et une diminution de la production d'inhibiteurs endogènes tels que la thrombospondine-1 (TSP-1)<sup>15,16,17</sup>.

### 1.2.3. Le mécanisme de l'angiogenèse tumorale

Les cellules tumorales en état de stress hypoxique sécrètent des facteurs de croissance qui diffusent dans les tissus adjacents en direction des vaisseaux sanguins existants. Ils se lient à des récepteurs spécifiques (à activité tyrosine kinase) situés à la surface des cellules endothéliales. La première réponse à ces stimuli angiogéniques est la vasodilatation du vaisseau préexistant et le détachement des péricytes, qui induit une augmentation de la perméabilité vasculaire. Les cellules endothéliales activées sécrètent et activent alors des protéases (la plasmine et des métallo-protéases matricielles (MMP)) nécessaires à la dégradation de la membrane basale et de la matrice extracellulaire environnante. Puis, elles migrent et prolifèrent à travers la matrice extracellulaire en direction de la tumeur, et s'organisent en structures tubulaires (sous l'effet du VEGF, de bFGF (facteur de croissance des fibroblastes), et des intégrines). La maturation et la stabilisation des vaisseaux néoformés s'effectuent via la reconstruction de la membrane basale et de la matrice extracellulaire, ainsi que par le recrutement des péricytes par les cellules endothéliales, (sous l'effet de PDGF, TGF- $\beta$  et ANG-1). Un flux sanguin peut dès lors s'établir dans le nouveau vaisseau sanguin ; il fournit l'oxygène et les nutriments nécessaires à la croissance de la tumeur et à sa dissémination métastatique (Figure 1)<sup>18,19</sup>.

---

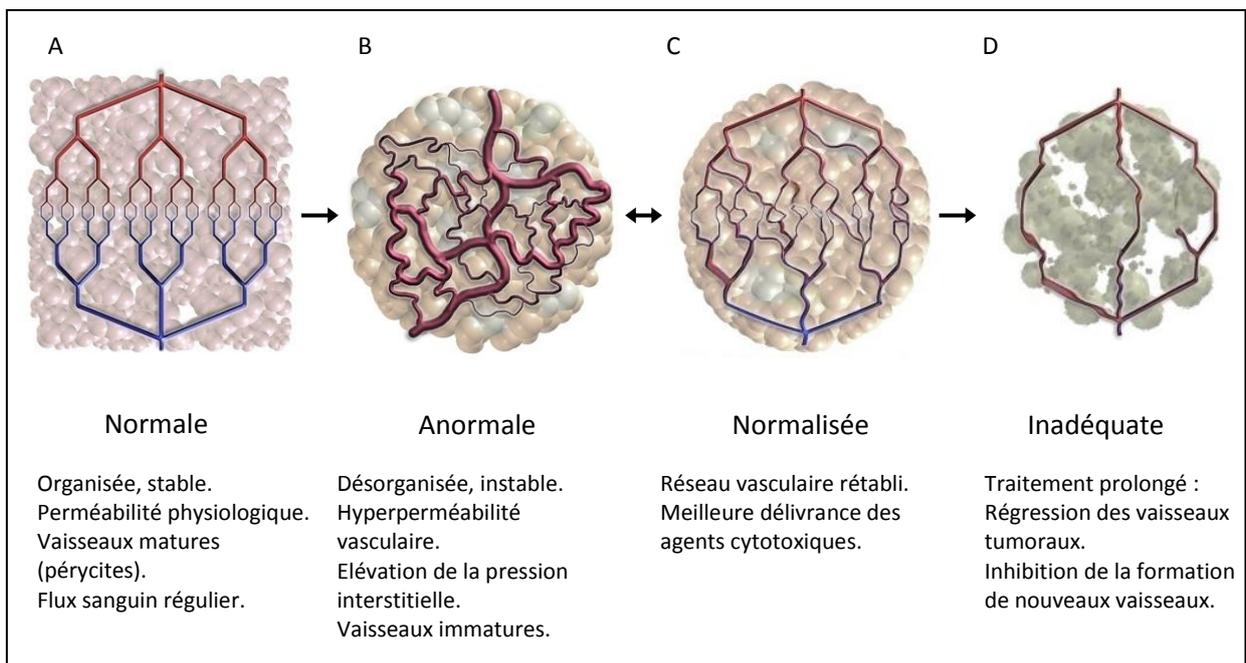
<sup>15</sup> Kranenburg, O., Gebbink, M. F. B. G, and Voest, E. E. (2004). *BBA-Reviews on Cancer*. **1654** (1), 23-37.

<sup>16</sup> Xie, K., Wei, D., Shi, Q., and Huang, S. (2004). *Cytokine & Growth Factor Reviews*. **15**, 297-324.

<sup>17</sup> Ren, B., Yee, K. O., Lawler, J., and Khosravi-Far, R. (2006). *BBA-Reviews on Cancer*. **1765** (2), 178-188.

<sup>18</sup> Wong, M. L. H., Prawira, A., Kaye, A. H., Hovens, C. M. (2009). *J. Clin. Neuroscience*. **16**, 1119-1130.

<sup>19</sup> Otrrock, Z. K., Mahfouz, R. A. R., Makarem, J. A., Shamseddine, A. I. (2007). *Blood Cells Mol Dis*. **39**, 212-220.



*Figure 2 : Concept de normalisation de la vascularisation tumorale<sup>20</sup>.*

<sup>20</sup> Jain, R. K., di Tomaso, E., Duda, D. G., Loeffler, J. S., Sorensen, A. G., and Batchelor, T. T. (2007). *Nat. Rev. Neurosci.* **8**, 610-622.

### 1.3. Les thérapies ciblées anti-angiogéniques

Les thérapies ciblées anti-angiogéniques agissent sur l'environnement de la tumeur ; en inhibant la formation de nouveaux vaisseaux sanguins elles ralentissent la croissance tumorale et le développement de métastases. Comme nous le verrons plus loin, différentes molécules aux propriétés anti-angiogéniques ont été développées ces dernières années. Il s'agit soit d'anticorps monoclonaux dirigés contre des facteurs de croissance ou le domaine extracellulaire de leurs récepteurs membranaires, soit d'inhibiteurs de tyrosine kinase, qui agissent sur le domaine intracellulaire des récepteurs ou sur des protéines kinases impliquées dans les voies de signalisation. Ces médicaments sont principalement utilisés en association avec d'autres traitements comme la chimiothérapie ou la radiothérapie<sup>21,22</sup>

#### *1.3.1. Le mécanisme d'action des agents anti-angiogéniques : application en multithérapies*

Pour expliquer l'effet synergique des agents anti-angiogéniques avec des cytotoxiques, le concept de normalisation des vaisseaux tumoraux a été suggéré. (Figure 2)<sup>23,24</sup>

Contrairement au réseau vasculaire normal, qui est organisé et stable (Figure 2-A), le réseau vasculaire tumoral présente une organisation anarchique et instable (Figure 2-B). L'administration d'un anti-angiogénique induit une normalisation de la vascularisation tumorale via une destruction des cellules endothéliales immatures et un recrutement de pérycites (diminution de la perméabilité vasculaire). Cela conduit à une hausse de l'oxygénation tumorale et à une meilleure délivrance des agents cytotoxiques (Figure 2-C). Un traitement prolongé peut engendrer une inhibition de la formation de nouveaux vaisseaux tumoraux et un ralentissement de la croissance tumorale (Figure 2-D). Malgré l'efficacité initiale du traitement, il est apparu que la normalisation des vaisseaux tumoraux était un phénomène transitoire. Certaines cellules cancéreuses acquièrent rapidement des mécanismes de résistance aux agents anti-angiogéniques, et reprennent leur croissance anarchique (activation de la néo-vascularisation tumorale) (Figure 2-B)<sup>25,26</sup>.

---

<sup>21</sup> Gasparini, G., Longo, R., Fanelli, M., and Teicher, B. A. (2005). *J. Clin. Oncol.* **23** (6), 1295-1311.

<sup>22</sup> Teng, L-S., Jin, K-T., He, K-F., Wang, H-H., Cao, J., and Yu, D-C. (2010). *J. Chin. Med. Asso.* **73** (6), 281-288.

<sup>23</sup> Jain, R. K. (2005). *Science.* 307, 58-62.

<sup>24</sup> Jain, R. K., Duda, D. G., Clark, J. W., and Loeffler, J. S. (2006). *Nat. Clin. Pract. Oncol.* 3 (1),24-40.

<sup>25</sup> Broxterman, H. J., Gotink, K. J., and Verheul, H. M. W. (2009), *Drug Res. Updates.* **12**, 114-126.

<sup>26</sup> Loges, S., Schmidt, T., and Carmeliet, P. (2010). *Genes & Cancer.* **1**, 12-25.

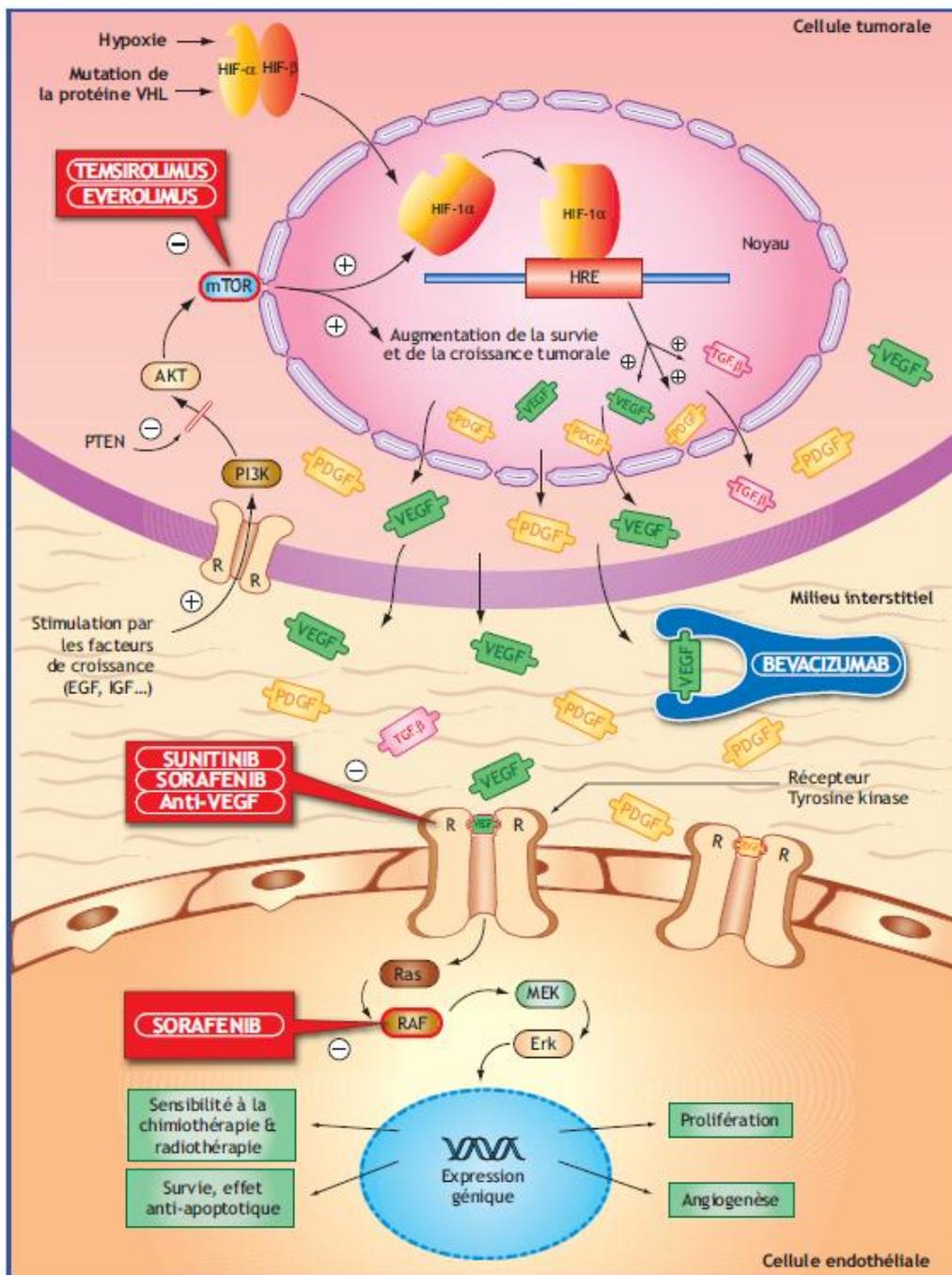


Figure 3 : Cibles des principaux agents anti-angiogéniques approuvés par la FDA<sup>27</sup>

<sup>27</sup> Méjean, A., and Lebret, T. (2008). Progrès en Urologie. Suppl. 7, S228-S233.

### 1.3.2. Le VEGF : une cible de choix

En raison de leur rôle clé dans la néo-vascularisation tumorale (augmentation de la perméabilité vasculaire, migration, prolifération, et survie des cellules endothéliales), le VEGF et ses récepteurs à activité tyrosine kinase sont rapidement apparus comme des cibles de choix pour le développement de molécules aux propriétés anti-angiogéniques<sup>28,29</sup>. Ces médicaments neutralisent le VEGF (le bevacizumab, un anticorps monoclonal humanisé anti-VEGF), ou ses récepteurs membranaires (le sorafenib ou le sunitinib, inhibiteurs de tyrosine kinase). Le sorafenib et le sunitinib sont des inhibiteurs multi-cibles ; ils agissent également sur d'autres récepteurs à tyrosine kinase comme le PDGFR, et sur B-RAF (sorafenib), une protéine kinase de la voie de signalisation Ras/RAF/MEK/ERK (induction de la prolifération et la survie cellulaire, et de la néo-vascularisation tumorale)<sup>30</sup>. D'autres voies de signalisation peuvent être inhibées comme la voie PI3K/AKT/mTOR<sup>31</sup> ; la neutralisation de mTOR par des esters de la rapamycine (temsirolimus ou everolimus) permet d'inhiber à la fois l'angiogénèse (diminution de la transcription du VEGF) et la croissance tumorale (Figure 3).

Le Tableau 1 répertorie les principaux agents anti-angiogéniques, ciblant le VEGF et ses récepteurs membranaires, à avoir été approuvés par la FDA (Agence fédérale américaine des produits alimentaires et médicamenteux). Le bevacizumab/Avastin® (le premier agent anti-angiogénique) a été autorisé en 2004 pour le traitement du cancer colorectal métastatique en association à la chimiothérapie. Ses indications se sont rapidement étendues au traitement d'autres types de cancer. Il s'agit aujourd'hui de l'anti-cancéreux le plus vendu au monde mais également le plus controversé.

---

<sup>28</sup> Backer, M. V., Hamby, C. V., and Backer, J. M. (2009). *Adv. Genet.* **67**, 1-27.

<sup>29</sup> Tugues, S., Koch, S., Gualandi, L., Li, X., Claesson-Welsh, L. (2011). *Mol. Asp. Med.* **32**, 88-111.

<sup>30</sup> Montagut, C., and Settleman, J. (2009). *Cancer Letters.* **283** (2), 125-134.

<sup>31</sup> Jiang, B-H., and Liu, L-Z. (2009). *Adv. Cancer Res.*, **102**, 19-65.



Agent thérapeutique	Type d'inhibiteur	Cible thérapeutique	Approbation de la FDA* et indications**
<b>Bevacizumab</b> <b>Avastin®</b> (Genentech/Roche) <sup>32</sup>	Anticorps monoclonal humanisé de type IgG1 (demi-vie 21 jours)	VEGF	En <b>2004</b> : cancer colorectal métastatique en association à la chimiothérapie (5-FU/acide folinique avec ou sans irinotécan ( <i>traitement de première ligne</i> ), ou fluoropyrimidine ( <i>traitement de deuxième ligne</i> )). En <b>2006</b> : cancer bronchique non à petite cellules avancé et non opérable, et/ou métastatique en association à la chimiothérapie (sels de platine) ( <i>traitement en première ligne</i> ). En <b>2008</b> : cancer du sein métastatique en association au paclitaxel ( <i>traitement en première ligne</i> ). En <b>2009</b> : cancer du rein avancé et/ou métastatique en association à l'interféron alpha-2a ( <i>traitement en première ligne</i> ).
<b>Sorafenib</b> <b>Nexavar®</b> (Bayer/Schering) <sup>33</sup>	Inhibiteur à tyrosine kinase	VEGFR2-3 PDGFR B-RAF	En <b>2005</b> : cancer du rein avancé (carcinome rénal) après échec d'un traitement à base d'interféron alpha ou d'interleukine-2. En <b>2008</b> : cancer du foie (carcinome hépatocellulaire).
<b>Sunitinib</b> <b>Sutent®</b> (Pfizer) <sup>34</sup>	Inhibiteur à tyrosine kinase	VEGFR1-3 PDGFR	En <b>2006</b> : tumeurs stromales gastro-intestinales malignes non résécables, et/ou métastatique après échec d'un traitement par le mésylate d'imatinib dû à une résistance ou à une intolérance. En <b>2006</b> : cancer du rein avancé et/ou métastatique après échec d'un traitement à base d'interféron alpha ou d'interleukine-2.
<b>Pazopanib</b> <b>Votrient®</b> (GlaxoSmithKline) <sup>35</sup>	Inhibiteur à tyrosine kinase	VEGFR1-3 PDGFR	En <b>2009</b> : cancer du rein avancé ( <i>traitement de première ligne</i> ) et chez les patients préalablement traités par des cytokines à un stade avancé de leur maladie.
<b>Aflibercept</b> <b>Zaltrap®</b> (Sanofi) <sup>36</sup>	Protéine de fusion	VEGF PIGF (facteur de croissance placentaire)	Approbation en cours ( <b>prévue en été 2012</b> ): cancer colorectal métastatique en association à la chimiothérapie FOLFIRI (irinotécan-5-fluorouracile-leucovorine).

\*FDA (« Food and Drug Administration ») : Agence fédérale américaine des produits alimentaires et médicamenteux.

\*\* d'après la Haute Autorité de santé, disponible sur le site [www.has-sante.fr](http://www.has-sante.fr).

**Tableau 1** : Les traitements anti-angiogéniques actuellement utilisés en clinique et qui ciblent le VEGF et ses récepteurs membranaires (liste non exhaustive).

<sup>32</sup> Eskens, F. M., and Sleijfer, S. (2008). *EJC*. **44**, 2350-2356.

<sup>33</sup> Sulkes, A. (2010). *Isr. Med. Assoc. J.* **12** (10), 628-632.

<sup>34</sup> Oudard, S., Beuselincq, B., Decoene, J., and Albers, P. (2011). *Cancer Treat Rev.* **37** (3), 178-184.

<sup>35</sup> Pick, A. M., and Nystrom, K. K. (2012). *Clin. Therap.* **34** (3), 511-520.

<sup>36</sup> Gaya, A., and Tse, V. (2012). *Cancer Treat Rev.* **38** (5), 484-493.



### 1.3.3. Données cliniques

L'évaluation clinique des agents anti-angiogéniques en monothérapie ou en association avec d'autres thérapies anti-cancéreuses montre une amélioration de la survie des patients de quelques mois sans progression de la maladie.

Des études récentes ont, toutefois, démontré que ces traitements réduisent, dans un premier temps, la croissance tumorale, puis accélèrent le pouvoir invasif et métastatique des cellules malignes<sup>37,38</sup>.

Par ailleurs, ces thérapies s'accompagnent d'effets secondaires sévères : hypertension artérielle, saignements, mauvaise cicatrisation, réaction cutanée, et dans certain cas perforation gastro-intestinale<sup>39,40</sup>. Le bevacizumab/Avastin® a d'ailleurs été retiré du marché américain par le FDA en novembre 2011 pour le traitement du cancer du sein métastatique, en raison de la gravité des effets secondaires et de son inefficacité en terme de survie<sup>41</sup>. L'institut national pour la santé et l'excellence clinique (NICE) de Grande-Bretagne a pris, depuis le 5 juillet 2012, une décision similaire pour la même pathologie<sup>42</sup>. L'autorisation de commercialisation de l'Avastin® reste toutefois valable dans ces pays pour le traitement des cancers du côlon, du rein, et du poumon.

Les thérapies anti-angiogéniques sont également très onéreuses ; le coût annuel du traitement du cancer du côlon avec l'Avastin® est d'environ 50 000 \$. Pour traiter le cancer du poumon, les doses nécessaires sont plus élevées, et le coût peut atteindre 100 000 \$<sup>43</sup>.

Enfin, la réponse aux thérapies anti-angiogéniques est imprévisible et varie considérablement d'un patient à l'autre, certains étant même totalement résistants. Il est par conséquent nécessaire de concevoir et valider de nouvelles molécules pour proposer des traitements adaptés à chaque patient. C'est dans cette optique que l'Aminopeptidase-N (APN ou CD13) s'est imposée comme une cible de choix. Il s'est avéré que cette dernière interfère avec la morphogénèse et la migration des cellules endothéliales, ainsi que la migration des cellules tumorales.

---

<sup>37</sup> Pàez-Ribes, M., Allen, E., Hudock, J., Takeda, T., Okuyama, H., Vinals, F., Inoue, M., Bergers, G., Hanahan, D., and Casanovas, O. (2009). *Cancer Cell*. **15**, 220-231.

<sup>38</sup> Ebos, J. M. L., Lee, C. R., Cruz-Munoz, W., Bjarnason, G. A., Christensen, J. G., and Kerbel, R. S. (2009). *Cancer Cell*. **15**, 232-239.

<sup>39</sup> Elice, F., and Rodeghiero, F. (2012) *Thrombosis Research*. **129** Suppl. 1, S50-S53.

<sup>40</sup> Verheul, H. M. W., and Pinedo, H. M. (2007). *Nat. Rev. Cancer*. **7**, 475-485.

<sup>41</sup> "Cancer drug Avastin loses US approval". *BBC News*. November 18, 2011.

<sup>42</sup> "Britain rejects Roche's Avastin in advanced breast cancer treatment". *Reuters*. Edition UK. July 6, 2012.

<sup>43</sup> "Prices soar for cancer drugs". *USA Today*. July 11, 2006.



## 2. L'Aminopeptidase-N (APN ou CD13) : une cible potentielle

### 2.1. L'APN : une protéine multifonctionnelle

L'Aminopeptidase-N (APN ou CD13) [EC.3.4.11.2] est une ecto-enzyme ubiquitaire fortement glycosylée appartenant à la classe des aminopeptidases zinc-dépendantes de la famille M1. Elle est exprimée sur les progéniteurs myéloïdes (monocytes, macrophages et cellules dendritiques), à la surface des cellules épithéliales de l'intestin et du rein, ainsi que dans le système nerveux central (synapses et péricytes)<sup>44,45,46</sup>.

L'APN assure une multitude de fonctions au sein de l'organisme (« moonlighting enzyme »), qui dépendent généralement du type cellulaire dans lequel elle est exprimée<sup>47,48</sup>.

En tant qu'enzyme, elle régule l'activité de peptides qui participent à d'importants processus biologiques. Il a été démontré *in vitro* qu'elle est capable de dégrader des protéines de la matrice extracellulaire dont le nidogène<sup>49</sup> et le collagène IV<sup>50</sup>, ou bien d'inactiver des neuropeptides comme les enképhalines et les endorphines (au niveau des vésicules synaptiques)<sup>51</sup>.

Dans le cas de certaines maladies infectieuses, le hasard des reconnaissances moléculaires fait que l'APN est également un récepteur d'endocytose pour de nombreux virus comme le coronavirus humain 229E (HCoV-229E), le coronavirus respiratoire de porc (CoV), ou encore le cytomegalovirus humain (HCMV)<sup>52,53</sup>.

Enfin, il semblerait que l'APN soit impliquée dans la transduction du signal, fonction indépendante de son activité enzymatique. Le domaine cytoplasmique de l'APN étant trop court (7 acides aminés) pour contenir un motif signalétique, il est peu probable que cette fonction soit intrinsèque à l'APN, et elle devrait s'associer à d'autres protéines pour initier une cascade de signalisation (bien que rien n'est clairement été établi à ce jour).

---

<sup>44</sup> Look, A. T., Ashmun, R. A., Shapiro, L. H., and Peiper, S. C. (1989). *J. Clin. Invest.* **83**, 1299-1306.

<sup>45</sup> Riemann, D., Kehlen, A., and Langner, J. (1999). *Immunol. Today.* **20**, 83-88.

<sup>46</sup> Matsas, R., Stephenson, S. L., Hryszko, J., Turner, A. J., and Kenny, A. J. (1985). *Biochem. J.* **231**, 445-449.

<sup>47</sup> Luan, Y., and Xu, W. (2007). *Curr. Med. Chem.* **14**, 639-647.

<sup>48</sup> Mina-Osorio, P. (2008). *Trends Mol. Med.* **14** (8), 361-371.

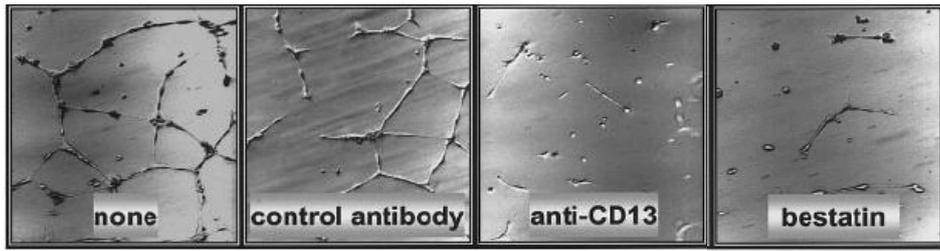
<sup>49</sup> Menrad, A., Speicher, D., Wacker, J., and Herlyn, M. (1993). *Cancer Res.* **53**, 1450-1455.

<sup>50</sup> Saiki, I., Yoneda, J., Azuma, I., Fujii, H., Abe, F., Nakajima, M., and Tsuruo, T. (1993). *Int. J. Cancer.* **54**, 137-143.

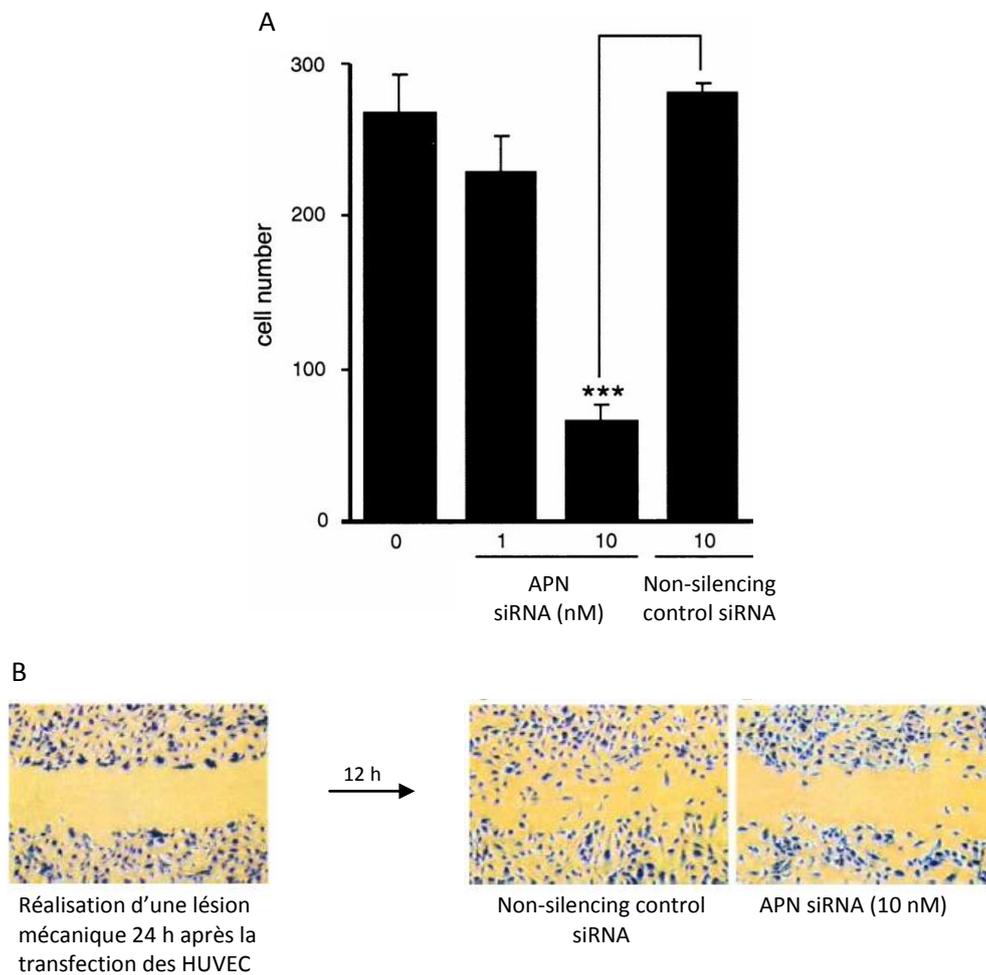
<sup>51</sup> Matsas, R., Stephenson, S. L., Hryszko, J., Kenny, A. J., and Turner, A. J. (1985). *Biochem. J.* **231**, 445-449.

<sup>52</sup> Kolb, A. F., Hegyi, A., Maile, J., Heister, A., Hagemann, M., and Siddell, S. G. (1998). *Adv. Exp. Med. Biol.* **440**, 61-67.

<sup>53</sup> Söderberg, C., Giugni, T. D., Zaia, J. A., Larsson, S., Wahlberg, J. M., and Möller, E. (1993). *J. Virol.* **67**, 6576-6585.



*Figure 4 : Inhibition de l'organisation des cellules endothéliales de veine ombilicale humaine (HUVEC) en tubes capillaires lorsqu'elles sont déposées sur une matrice artificielle (Matrigel) en présence d'inhibiteur (bestatine, 100 µg/l) ou d'anticorps (anti-CD13, 100 µg/l)<sup>54</sup>.*



*Figure 5 : Inhibition de la migration des cellules HUVEC lorsqu'elles sont transfectées avec 10 nM de siRNA de l'APN<sup>59</sup>. (A). Test de migration en chambre de Boyden ; la migration des cellules à travers la membrane microporeuse est induite par la fibronectine (B). Comblement de lésion sur monocouche endothéliale (« Wound-healing assay »).*

## 2.2. L'APN : son implication dans la tumorigenèse.

Il est décrit dans la littérature que l'APN est surexprimée à la surface des cellules endothéliales de vaisseaux angiogéniques, mais est peu présente sur le réseau vasculaire sain. Sa transcription est induite par des stimuli angiogéniques émanant du microenvironnement tumoral (hypoxie et facteurs de croissance comme le VEGF, le bFGF et le TNF $\alpha$  (facteur de nécrose tumorale  $\alpha$ )) ou suite à l'activation de voies de signalisation comme la voie Ras/RAF/MER/ERK<sup>54,55</sup>.

L'APN est également surexprimée à la surface de certaines cellules tumorales : cancer de la peau, des ovaires, de la thyroïde, du poumon, de l'estomac, du colon, du rein, des os et de la prostate<sup>48</sup>. Son expression est liée à l'agressivité des tumeurs, et est un facteur de mauvais pronostic pour le patient<sup>56</sup>. Par ailleurs, il existe une corrélation étroite entre l'élévation de l'expression de l'APN, une activité enzymatique accrue et le pouvoir invasif de nombreux types de cellules tumorales<sup>57,58</sup>.

Selon les données de la littérature, l'APN régule des fonctions généralement altérées en premier lieu dans la tumorigenèse comme l'invasion tumorale, la différenciation, la prolifération et l'apoptose, la motilité cellulaire, l'adhésion et l'angiogenèse<sup>48</sup>. Bien que son activité catalytique extracellulaire soit requise pour ces processus, le mécanisme d'action de cette protéine ainsi que les substrats impliqués ne sont pas identifiés.

L'implication de l'APN dans la morphogenèse<sup>54</sup> (Figure 4), ainsi que dans la migration des cellules endothéliales<sup>59</sup> (Figure 5) a été démontré *in vitro* via l'utilisation d'inhibiteur (bestatine/Ubenimex®), d'anticorps (anti-CD13) ou de siRNA (« small interfering RNA »).

Les travaux de Rangel et al. ont permis de confirmer l'importance de l'APN dans les processus d'angiogenèse<sup>60</sup>. Des souris « knock-out » n'exprimant plus la protéine se développent normalement sans altérations physiologiques, mais présentent une sévère déficience de la néo-vascularisation dans des conditions pathologiques.

---

<sup>54</sup> Bhagwat, S. V., Lahdenranta, J., Giordano, R., Arap, W., Pasqualini, R., and Shapiro, L. H. (2001). *Blood*. **97**, 652-659.

<sup>55</sup> Petrovic, N., Bhagwat, S. V., Ratzan, W. J., Ostrowski, M. C., and Shapiro, L. H. (2003). *J. Biol. Chem.* **278** (49), 49358-49368.

<sup>56</sup> Hashida, H., Takabayashi, A., Kanai, M., Adachi, M., Kondo, K., Kohno, N., Yamaoka, Y., and Miyake, M. (2002). *Gastroenterology*. **122**, 376-386.

<sup>57</sup> Dixon J., Kaklamanis L., Turley, H., Hickson, I. D., Leek, R. D., Harris, A. L., and Gatter, K. C. (1994). *J. Clin. Pathol.* **47**, 43-47.

<sup>58</sup> Fujii, H., Nakajima, M., Saiki, I., Yoneda, J., Azuma, I., and Tsuruo, T. (1995). *Clin. Exp. Metastasis*. **13**, 337-344.

<sup>59</sup> Fukasawa, K., Fujii, H., Saitoh, Y., Koizumi, K., Aozuka, Y., Sekine, K., Yamada, M., Saiki, I., and Nishikawa, K. (2006). *Cancer Lett.* **243**, 135-143.

<sup>60</sup> Rangel, R., Sun, Y., Guzman-Rojas, L., Ozawa, M. G., Sun, J., Giordano, R. J., Van Pelt, C. S., Tinkey, P. T., Behringer, R. R., Sidman, R. L., Arap, W., and Pasqualini, R. (2007). *PNAS*. **104** (11), 4588-4593.

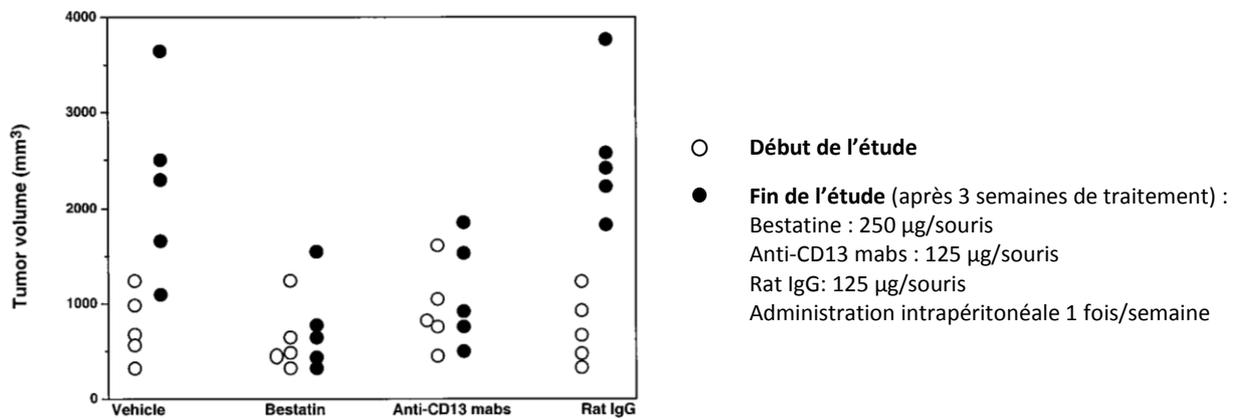


Figure 6 : Effet de la bestatine et d'un anticorps anti-CD13 sur la croissance tumorale dans un modèle de xéno greffe d'une lignée cellulaire de cancer du sein (MDA-MB-435) implanté chez la souris « nude »<sup>61</sup>.

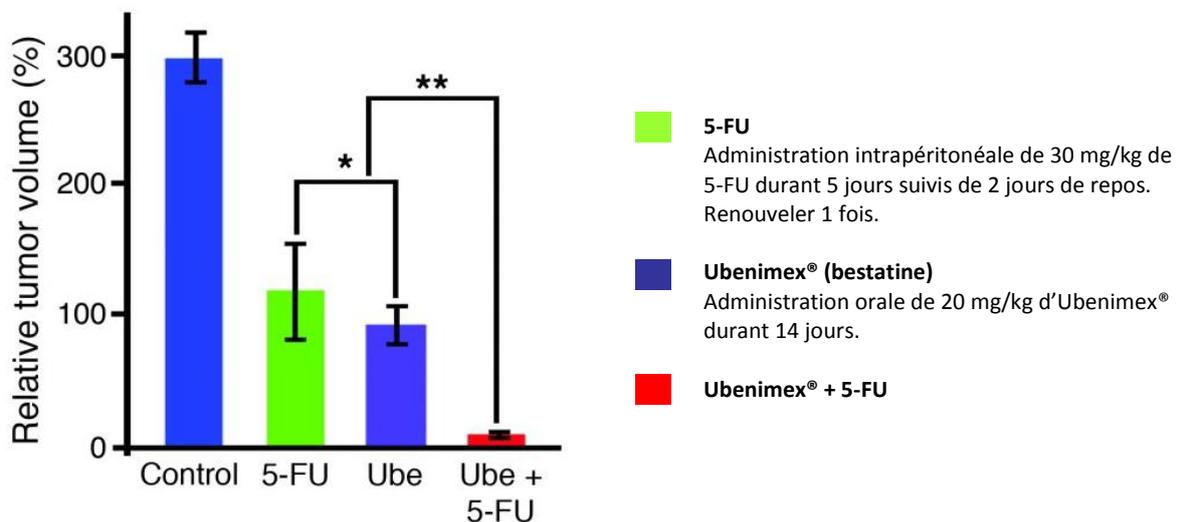


Figure 7 : Effet de la bestatine (Ubemimex®) en monothérapie ou en association à la 5-FU sur la croissance tumorale dans un modèle d'hépatocarcinome (lignée cellulaire PLC/PRF/5) chez la souris NOD/SCID (« no obese diabetic severe combined immunodeficient »)<sup>63</sup>.

Les premières études *in vivo*, dans les années 2000, ont montré que la croissance tumorale d'une xénogreffe de cancer du sein (MDA-MB-435) est inhibée lorsque l'activité enzymatique de l'APN est bloquée avec un inhibiteur (bestatine) ou un anticorps (anti-CD13) (Figure 6)<sup>61</sup>. Par ailleurs, le pouvoir invasif et métastasant de cellules cancéreuses d'ostéosarcome (MNNG/HOS) est réduit de 40 à 60%, dans un modèle murin, lorsque l'expression de l'APN est inhibée (transfection des cellules MNNG/HOS avec un vecteur exprimant l'ARN antisens)<sup>62</sup>. Enfin, dans un modèle d'hépatocarcinome chez la souris, un effet synergique a été observé lorsque la bestatine est utilisée en association avec un agent cytotoxique comme le 5-fluorouracile (5-FU) (Figure 7)<sup>63</sup>.

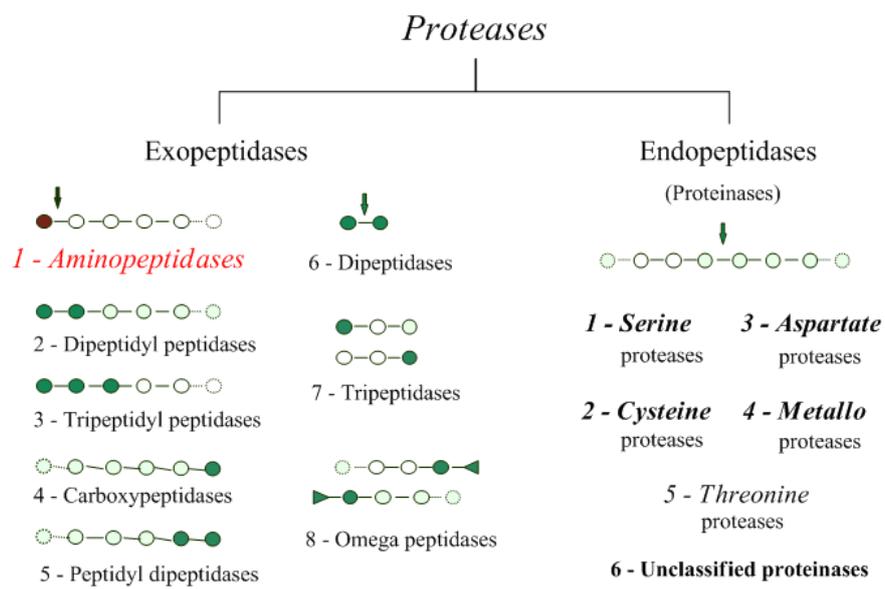
Au regard de l'ensemble des données sur l'APN dans le processus de tumorigenèse, il nous a semblé pertinent de vouloir inhiber son activité catalytique afin de limiter la progression tumorale.

---

<sup>61</sup> Pasqualini, R., Koivunen, E., Kain, R., Lahdenranta, J., Sakamoto, M., Stryhn, A., Ashmun, R. A., Shapiro, L. H., Arap, W., and Ruoslahti, E. (2000). *Cancer Res.* **60**, 722-727.

<sup>62</sup> Kido, A., Krueger, S., Haeckel, C., and Roessner, A. (2003). *Clin. Exp. Metastasis.* **20**, 585-592.

<sup>63</sup> Haraguchi, N., Ishii, H., Mimor, K., Tanaka, F., Ohkuma, M., Kim, H. M., Akita, H., Takiuchi, D., Hatano, H., Nagano, H., Barnard, G. F., Doki, Y., and Mori, M. (2010). *J. Clin. Invest.* **120**, 3326-3339.



*Figure 8 : Classification des protéases*

### 3. Conception d'inhibiteurs spécifiques de l'Amino-peptidase-N

Pour concevoir de manière rationnelle des inhibiteurs hautement spécifiques de l'APN, il faut s'intéresser au mécanisme catalytique ainsi qu'à la spécificité de substrat de cette enzyme ainsi que de celles de la même famille et des familles voisines.

#### 3.1. Les enzymes protéolytiques

Les enzymes protéolytiques (également nommées protéases ou peptidases) catalysent les réactions d'hydrolyse des liaisons amides des peptides ou des protéines<sup>64</sup>. Il s'agit de la réaction irréversible la plus importante dans le monde du vivant.

Les protéases sont naturellement très abondantes et interviennent dans la régulation de nombreux processus biologiques : digestion, recyclage (équilibre entre l'anabolisme et le catabolisme des acides aminés), développement embryonnaire, coagulation, cicatrisation, défense immunitaire, contrôle du cycle cellulaire, différenciation, migration, et apoptose. Elles sont également impliquées dans de nombreuses pathologies telles que les maladies cardiovasculaires et inflammatoires, le cancer, l'ostéoporose et les troubles neurologiques, et sont, par conséquent, des cibles pharmacologiques de choix pour le développement d'inhibiteurs<sup>65,66</sup>.

Deux grandes familles de protéases se distinguent, selon leur mode d'action : les exopeptidases et les endopeptidases (Figure 8). Les premières (exopeptidases), classées en fonction de leur spécificité de substrat, clivent les acides aminés en position amino- ou carboxy-terminale d'un peptide ou d'une protéine. Les secondes (endopeptidases) hydrolysent, quant-à-elles, les liaisons amides à l'intérieur de la chaîne peptidique ou protéique, et sont classées en fonction de la nature du ou des acides aminés impliqués dans la catalyse<sup>67</sup>.

Parmi les différentes classes d'exopeptidases, nous nous intéresserons plus particulièrement aux amino-peptidases.

---

<sup>64</sup> Barrett, A.J., Rawlings, N.D., and Woessner, J.F. (2004). Handbook of proteolytic enzymes. Second Edition. *Elsevier/Academic Press*.

<sup>65</sup> Mentlein, R. (2004). *Int. Rev. Cytol.* **235**, 165-213

<sup>66</sup> Turk, B. (2006). *Nat. Rev. Drug Discov.* **5**, 785-799

<sup>67</sup> Barrett, A. J. (1994). *Methods in Enzymology.* **244**, 1-15.

AP-A : acides aminés acides

AP-B : acides aminés basiques

**AP-N : acides aminés neutre**

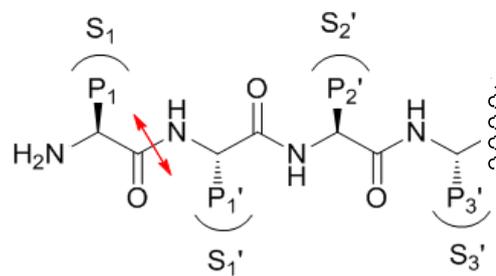
AP-O : Arginine et Acide Aspartique

AP-P : Proline

AP-W : Triptophane

LAPc (Leucine Aminopeptidase cytosolique) : Leucine

MetAP2 : (Méthionine Aminopeptidase de type 2) : Méthionine



*Figure 9* : Exemples de différentes aminopeptidases selon le type d'acide aminé hydrolysé. Représentation schématique d'un substrat peptidique ou protéique au site actif d'une aminopeptidase selon la nomenclature de Schechter et Berger<sup>68</sup> : la liaison peptidique hydrolysée se situe entre les résidus P1 et P1' ; les résidus P et P' interagissent avec les sous-sites S et S' du site catalytique de la protéase.

<sup>68</sup> Schechter, I., Berger, A. (1967). *Biochem. Biophys. Res. Commun.* **27**, 157-162.

## 3.2. Les aminopeptidases

Les aminopeptidases (AP) [EC.3.4.11.-] sont présentes chez les mammifères sous forme soluble ou associées aux membranes ; elles catalysent la réaction d'hydrolyse de l'acide aminé en position N-terminale d'un peptide ou d'une protéine. Dans cette sous classe d'enzymes protéolytiques, il existe encore une grande variété d'aminopeptidases, spécifiques du type d'acide aminé à hydrolyser, c'est-à-dire du résidu P1 (Figure 9). La fonction amine libre est essentielle à la reconnaissance d'un substrat par cette famille d'enzymes protéolytiques.

L'activité catalytique de la plupart de ces métallo-peptidases dépend de la présence dans leur site actif d'un ou deux ions métalliques tels que le zinc (le plus répandu), le cobalt ou le manganèse<sup>69,70</sup>. Les métallo-aminopeptidases de la famille M1 (un atome de zinc) et M17 (deux atomes de zinc organisés en « unité co-catalytique ») seront considérées dans les travaux présentés.

## 3.3. La famille M1 : les aminopeptidases à un ion zinc essentiel à la catalyse

### 3.3.1. Homologie séquentielle et structurale

L'Aminopeptidase-N (APN ou CD13) [EC.3.4.11.2] appartient à la famille M1 des métallo-aminopeptidases zinc-dépendantes<sup>71</sup>. La structure tridimensionnelle de l'enzyme de mammifères n'est pas connue, néanmoins, les structures de certaines enzymes de cette famille ont été résolues récemment. Parmi ces protéines, nous pouvons citer :

- les aminopeptidases du réticulum endoplasmique, ERAP1<sup>72</sup> et ERAP2<sup>73</sup> : ce sont des enzymes multifonctionnelles, impliquées dans la régulation des réponses immunitaires et inflammatoires ainsi que dans une maladie chronique, la spondylarthrite ankylosante (maladie inflammatoire chronique des articulations sacro-iliaques et de la colonne vertébrale)<sup>74</sup>.

<sup>69</sup> Lipscomb, W.N., and Sträter N. (1996). *Chem. Rev.* 96 (7), 2375-2434.

<sup>70</sup> Lowther, W. T., and Matthews, B. W. (2002). *Chem. Rev.* 102, 4581-4607.

<sup>71</sup> Rawlings, N.D., Barrett, A.J., and Bateman, A. (2012). *Nucleic Acids Res.* 40, D343-D350. <http://merops.sanger.ac.uk/>

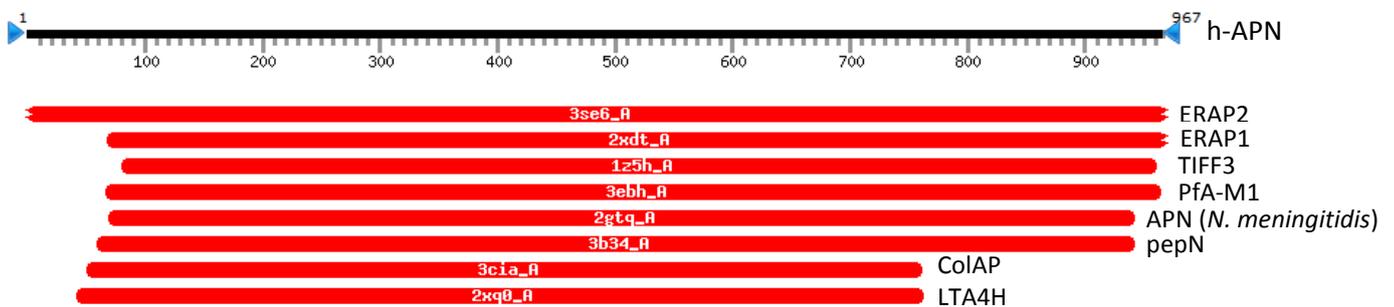
<sup>72</sup> Kochan, G., Krojer, T., Harvey, D., Fischer, R., Chen, L., Vollmar, M., von Delft, F., Kavanagh, K. L., Brown, A. B., Bowness, P., Wordsworth, P., Kessler, B. M., and Oppermann, U. (2011). *PNAS.* 108 (19), 7745-7750.

<sup>73</sup> Birtley, J. R., Saridakis, E., Stratikos, E., and Mavridis, I. M. (2012) *Biochemistry.* 51, 286-295.

<sup>74</sup> Tsui, F. W., Haroon, N., Reveille, J. D., Rahman, P., Chiu, B., Tsui, H. W., and Inman, R. D. (2010). *Ann. Rheum. Dis.* 69, 733-736.

Membre de la famille M1 dont la structure tridimensionnelle a été résolue		Identité de séquence avec h-APN (UniProtKB/Swiss-Prot P15144)
<b>ERAP1</b>	Aminopeptidase 1 humaine du réticulum endoplasmique (UniProtKB/Swiss-Prot ID Q9NZ08)	31 %
<b>ERAP2</b>	Aminopeptidase 2 humaine du réticulum endoplasmique (UniProtKB/Swiss-Prot ID Q6P179)	30 %
<b>TIFF3</b>	« Tricorn Interacting Factor F3 » de <i>Thermoplasma acidophilum</i> (UniProtKB/Swiss-Prot ID O93655)	22 %
<b>PfA-M1</b>	M1 Aminopeptidase neutre de <i>Plasmodium falciparum</i> (UniProtKB/Swiss-Prot ID O96935)	17 %
<b>pepN</b>	Aminopeptidase-N d' <i>Escherichia coli</i> (UniProtKB/Swiss-Prot ID P04825)	14 %
<b>LTA4H</b>	Leucotriène A4 hydrolase humaine (UniProtKB/Swiss-Prot ID P09960)	14 %
<b>ColAP</b>	« Cold-Active » Aminopeptidase de <i>Colwellia psychrerythraea</i> (UniProtKB/Swiss-Prot ID Q7WVY1)	14 %

*Tableau 2* : Identité de séquence entre h-APN et les membres de la famille M1 dont la structure tridimensionnelle a été résolue. L'outil alignement de séquences du serveur UniProt a été utilisé<sup>75</sup>.



*Figure 10* : Identification d'homologues structuraux à l'aide du logiciel HHpred<sup>78</sup>. La séquence protéique de h-APN (en noir) est comparée à l'ensemble des séquences des structures répertoriées dans la PDB. Seules les protéines pour lesquelles la probabilité de l'homologie structurale est de 100% sont représentées (en rouge). Code PDB 3se6 : Aminopeptidase 2 du réticulum endoplasmique humain (ERAP2) ; 2xdt : Aminopeptidase 1 du réticulum endoplasmique humain (ERAP1) ; 1z5h : « Tricorn Interacting Factor F3 » de *Thermoplasma acidophilum* (TIFF3) ; 3ebh : M1 alanyl-Aminopeptidase neutre de *Plasmodium falciparum* (PfA-M1) ; 2gtq : Aminopeptidase N de *Neisseria meningitidis* (APN) ; 3b34 : Aminopeptidase N d'*Escherichia coli* (PepN) ; 3cia : « Cold-active » Aminopeptidase de *Colwellia psychrerythraea* (ColAP) et 2xq0 : Leucotriène A4 Hydrolase de *Saccharomyces cerevisiae* (LTA4H).

<sup>75</sup> Jain, E., Bairoch, A., Duvaud, S., Phan, I., Redaschi, N., Suzek, B.E., Martin, M.J., McGarvey, P., and Gasteiger, E. (2009). BMC Bioinformatics. 10, 136. <http://www.uniprot.org/align/>

- la Leucotriène A4 hydrolase, LTA4H<sup>76</sup> ([EC.3.3.2.6]) : il s'agit d'une enzyme bifonctionnelle qui présente une activité époxyde hydrolase et aminopeptidase<sup>77</sup>. Elle participe (via son activité époxyde hydrolase) à la biosynthèse de la leucotriène B4, un lipide chimio-attractant impliqué dans les réponses immunitaires et inflammatoires (asthme, rhinite allergique, athérosclérose).

L'APN humaine (h-APN) présente peu d'identité de séquence avec les autres membres de la famille (Tableau 2).

Bien que l'homologie séquentielle soit partielle, le repliement est très similaire pour cette famille de métallo-aminopeptidases. L'homologie structurale entre ces enzymes est confirmée en comparant la séquence protéique de h-APN avec l'ensemble des séquences des structures répertoriées dans la PDB (« Protein Data Bank »). L'analyse est effectuée à l'aide du logiciel HHpred<sup>78</sup> ; les résultats sont présentés Figure 10.

Les chaînes polypeptidiques de ces aminopeptidases s'organisent en domaines structuraux distincts (Figure 11). Le domaine catalytique (en vert), contenant l'atome de zinc, essentiel à la catalyse, est très conservé. Son repliement dérive de celui de la Thermolysine ([EC.3.4.24.27]), qui semble être un ancêtre commun à ces métallo-aminopeptidases. Le domaine C-terminal (en magenta) est, à l'inverse, le domaine protéique qui a été le moins bien conservé au cours de l'évolution ; il est, par exemple, constitué de 10 hélices chez LTA4H, 17 chez pepN ou encore 16 chez ERAP1.

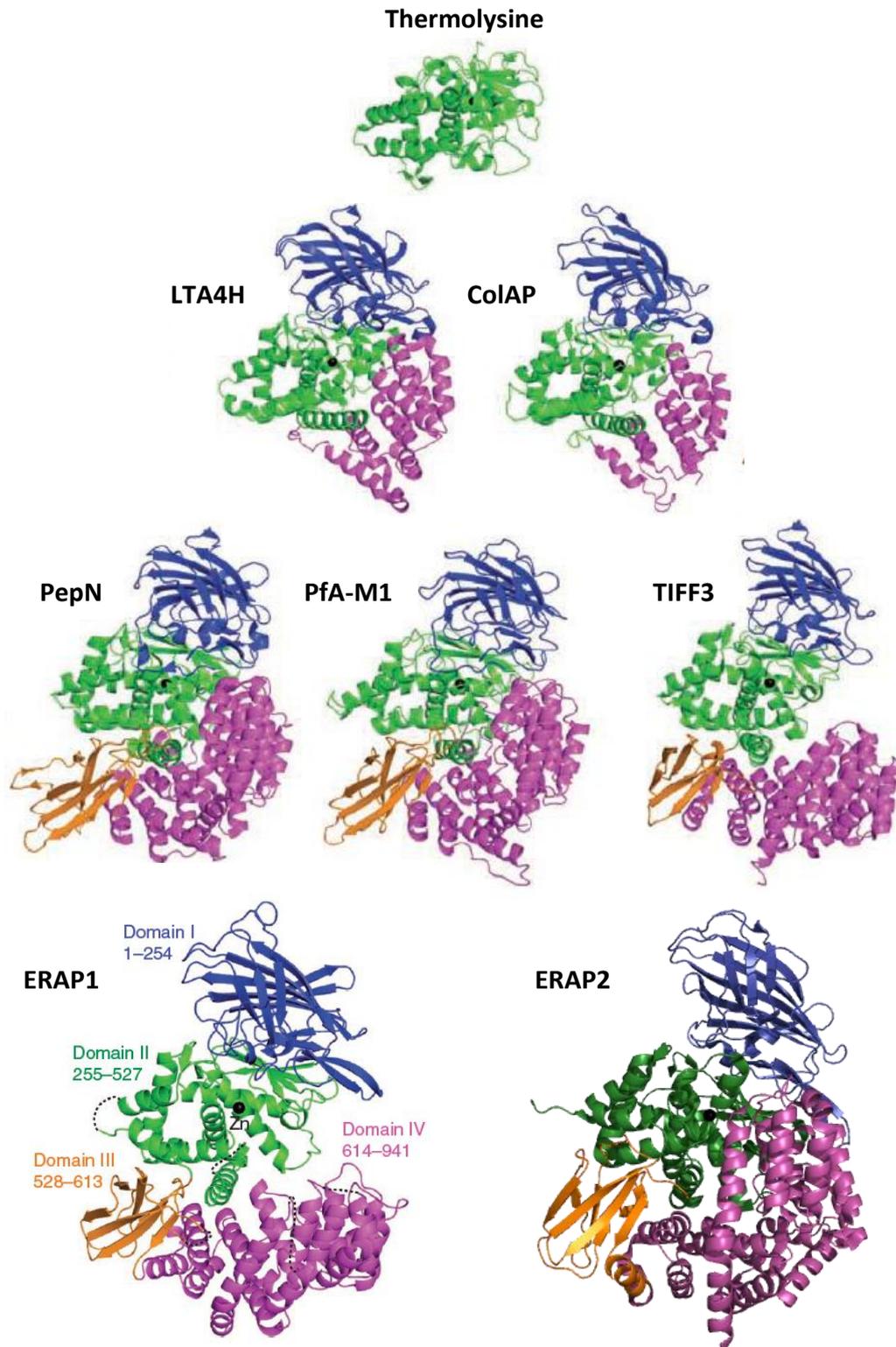
---

<sup>76</sup> Thunnissen, M. M. G. M., Nordlund, P., and Haeggström, J. Z. (2001). *Nat. Struct. Biol.* **8** (2), 131-135.

<sup>77</sup> Rudberg, P. C., Tholander, F., Thunnissen, M. M. G. M., Haeggström, J. Z. (2002). *J. Biol. Chem.* **277** (2), 1398-1404.

<sup>78</sup> Söding, J. (2005). *Bioinformatics*. **21**, 951-960. <http://toolkit.tuebingen.mpg.de/hhpred>





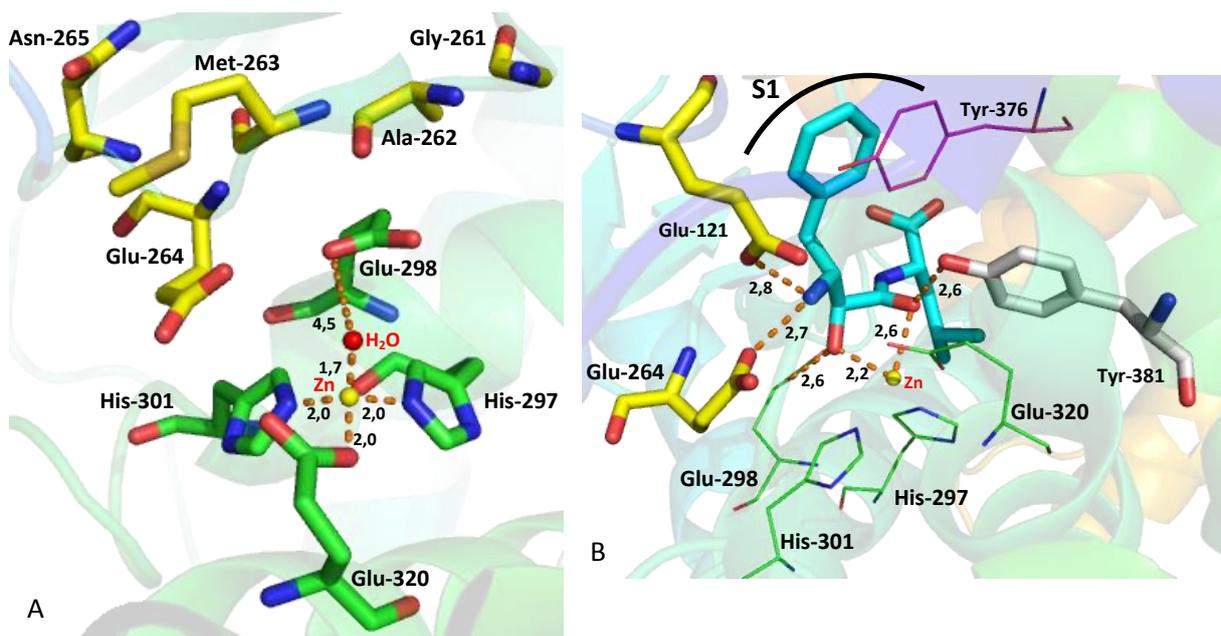
**Figure 11:** Structure tridimensionnelle, représentation en ruban, d'un panel de métallo-aminopeptidases connues de la famille M1 et de la Thermolysine (ancêtre commun)<sup>79,80</sup>.

<sup>79</sup> Nguyen, T. T., Chang, S. C., Evnouchidou, I., York, I. A., Zikos, C., Rock, K. L., Goldberg, A. L., Stratikos, E., and Stern, L. J. (2011). *Nat Struct Mol Biol.* **18** (5), 604-613.

<sup>80</sup> Birtley, J. R., Saridakis, E., Stratikos, E., and Mavridis, I. M. (2012). *Biochemistry.* **51**, 286-295.

hAPN	346	LPDFNAGAMENWGLVITYRENSLLFDPLSSSSSNKERVVTVIAHELAHQWFGNLVTIEWWN
pepN	255	VDFNMGAMENKGLNIFNSKYVLARTDTATDKDYLDIERVIGHEYFHNWVTGNRVTCRDWF
ERAP1	311	IPDFQSAMENWGLTITYRESALLFDAEKSSASSKLGITMTVAHELAHQWFGNLVTMEWWN
PfA-M1	454	VSDFNVMAMENKGLNIFNANSLLASKKNSIDFSYARILTVVGHYFHNVTGNRVTLRDWF
LTA4H	263	PPSFPYGMENPCLTFVTPTLLAGD-----KSLSNVIAHEISHSWTGNLVTNKTWD
		* .*** * . : : .:.* *.: ** ** . *
hAPN	406	DLWLNMGFASYVEYLGADYAEP--TWNLKDLMLVNDVYRVMAVDALASSHPLSTPASEIN
pepN	315	QLSLKEGLTVFRDQEFSSDLGSRVNRINNVRTMR-----GLQFAEDASPMHPIRP-D
ERAP1	371	DLWLNMGFAKFMFVSVSVTHP--ELKVGDFY-FGKCFDAMEVDALNSSHPVSTPV---E
PfA-M1	514	QLTLKEGLTVHRENLFSEEMTKVTTRLSHVDLLR-----SVQFLEDSSPLSHPIRP-E
LTA4H	314	HFWLNMGHTVYLERHICGRLFGKFRHFNALGGWGEL--QNSVKTFGETHPFTKLVD-L
		.: *:* : . : . . .: .: *.:

**Figure 12 :** Alignement partiel des séquences primaires de la région catalytique de quelques métallo-aminopeptidases de la famille M1 : hAPN (Aminopeptidase N humaine, UniProtKB/Swiss-Prot P15144), pepN (Aminopeptidase N d'Escherichia coli, P04825), hERAP1 (Aminopeptidase 1 du réticulum endoplasmique humaine, Q9NZ08), PfA-M1 (M1 Aminopeptidase neutre de Plasmodium falciparum, O96935) et LTA4H (Leucotriène A4 hydrolase humaine, P09960). L'outil alignement de séquences du serveur UniProt a été utilisé<sup>75</sup>.



**Figure 13 :** (A). Aminopeptidase-N d'Escherichia coli code PDB : 2dq6. Résidus impliqués dans la chélation du zinc (le motif HEXXH-X18-E, en vert) et ceux indispensables à la reconnaissance de l'amine libre en position N-terminale d'un substrat peptidique ou protéique (le site GXMEN, en jaune). (B). Aminopeptidase N d'Escherichia coli code PDB : 2dqm. Mode de liaison de la bestatine au site actif de l'enzyme. Les distances entre les hétéroatomes et l'ion métallique sont données en Angström (Å). Le logiciel PyMOL est utilisé pour visualiser la structure tridimensionnelle de la protéine.

Tous les membres de cette famille présentent dans leur séquence primaire le motif HEXXH-X18-E de fixation du zinc, ainsi que le site GXMEN, indispensable à la reconnaissance de l'amine libre en position N-terminale du substrat peptidique ou protéique (Figure 12).

La Figure 13-A illustre les acides aminés du site actif, essentiels à la catalyse chez l'Aminopeptidase-N d'*Escherichia coli* (code PDB : 2dq6). Dans l'enzyme native, l'ion zinc est tétra-coordonné par deux résidus histidine (His-297 et His-301), un résidu acide glutamique (Glu-320) et une molécule d'eau (agent nucléophile dans la réaction d'hydrolyse). Cette dernière est également liée par liaison hydrogène au carboxylate du résidu Glu-298. Le résidu Glu-264 du motif GXMEN participe à la reconnaissance du substrat peptidique : une liaison hydrogène se forme entre le groupement carboxylate de Glu-264 et l'amine primaire du substrat peptidique.

### 3.3.2. Mécanisme catalytique et spécificité de substrat

L'APN présente une spécificité de substrat très large. Elle reconnaît préférentiellement les acides aminés neutres et hydrophobes comme la leucine, la valine et la phénylalanine en position N-terminale, et ce en raison de la présence de nombreux noyaux aromatiques dans le sous-site S1. Mais, elle hydrolyse également les résidus arginine. Dans ce cas une interaction peut s'établir entre la charge positive du résidu Arg et un carboxylate au niveau du sous-site S1. Il est également établi que l'arginine peut interagir par des liaisons cation- $\pi$  avec des noyaux aromatiques présents dans ce sous-site<sup>81,82</sup>.

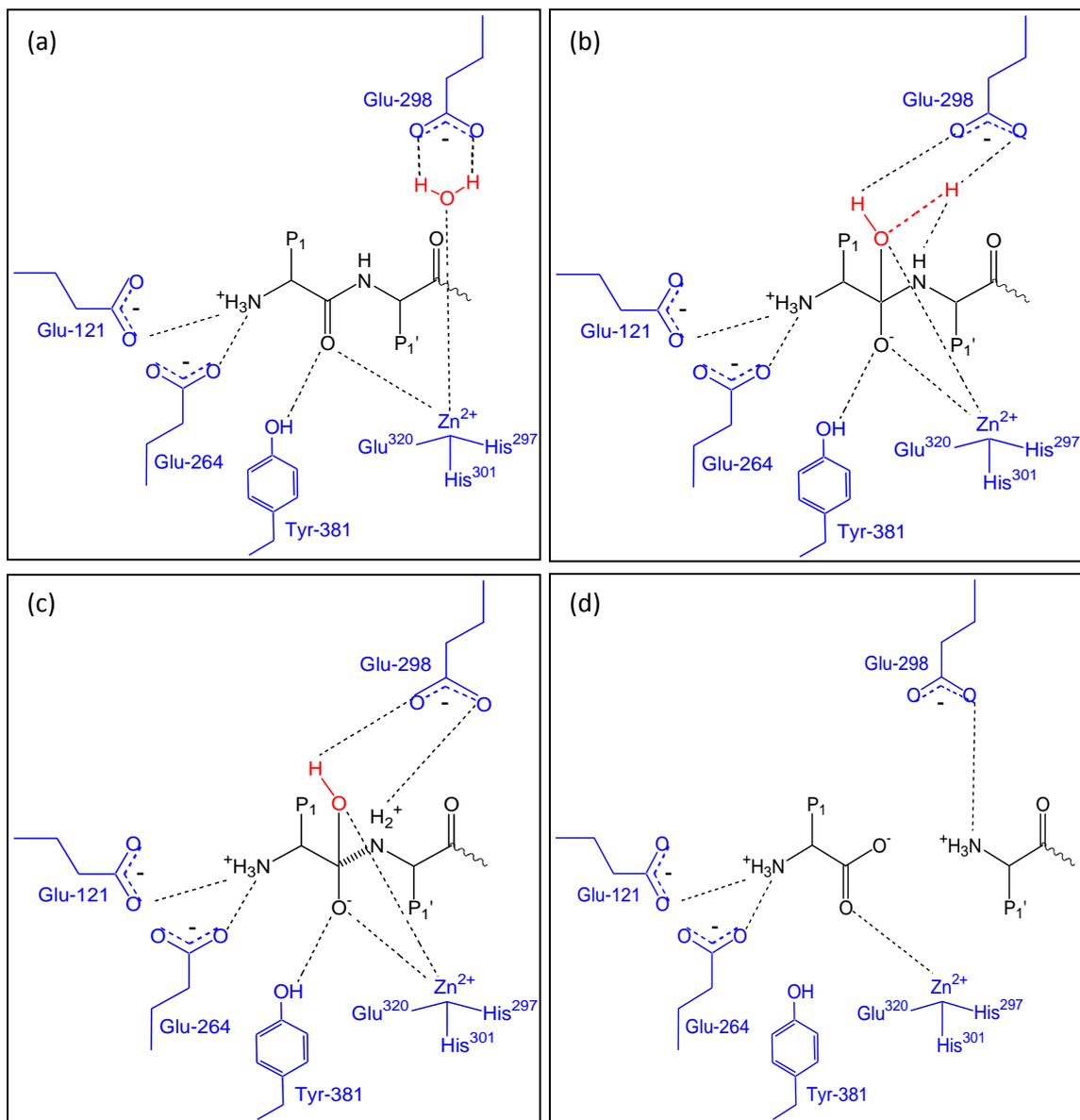
Un mécanisme catalytique a été proposé pour tous les membres de la famille M1 d'après l'analyse du mode de liaison de substrats peptidiques et d'analogues de l'état de transition. La liaison de la bestatine (produit naturel)<sup>83</sup> au site actif de l'Aminopeptidase-N d'*Escherichia coli* (pepN) (code PDB 2dqm) est représentée ci-contre (Figure 13-B)<sup>84</sup> ; sa fixation est similaire à celle d'un substrat peptidique. L'amine primaire de la bestatine interagit avec les résidus Glu-264 (motif GAMEN) et Glu-121 ; la fonction carbonyle se lie à l'ion zinc et forme une liaison hydrogène avec l'hydroxyle du résidu Tyr-381. Le groupement hydroxyle de la bestatine occupe, quant-à-lui, la place de la molécule d'eau (agent nucléophile) ; il se lie au zinc et forme une liaison hydrogène avec le résidu Glu-298 impliqué dans la catalyse.

<sup>81</sup> Flocco, M. M., and Mowbray, S. L. (1994). *J Mol Biol.* **235** (2), 709-717.

<sup>82</sup> Du, Q-S., Meng, J-Z, Liao, S-M, Huang, R-B. (2012). *J. of Mol Graph and Model.* **34**, 38-45.

<sup>83</sup> Umezawa, H. (1981). Small Molecular Immunomodifiers of microbial origin. Fundamental and Clinical Studies of Bestatin. *Japan Scientific Press, Tokyo & Pergamon Press, Oxford, New York, Toronto, Sydney, Paris, Frankfurt.*

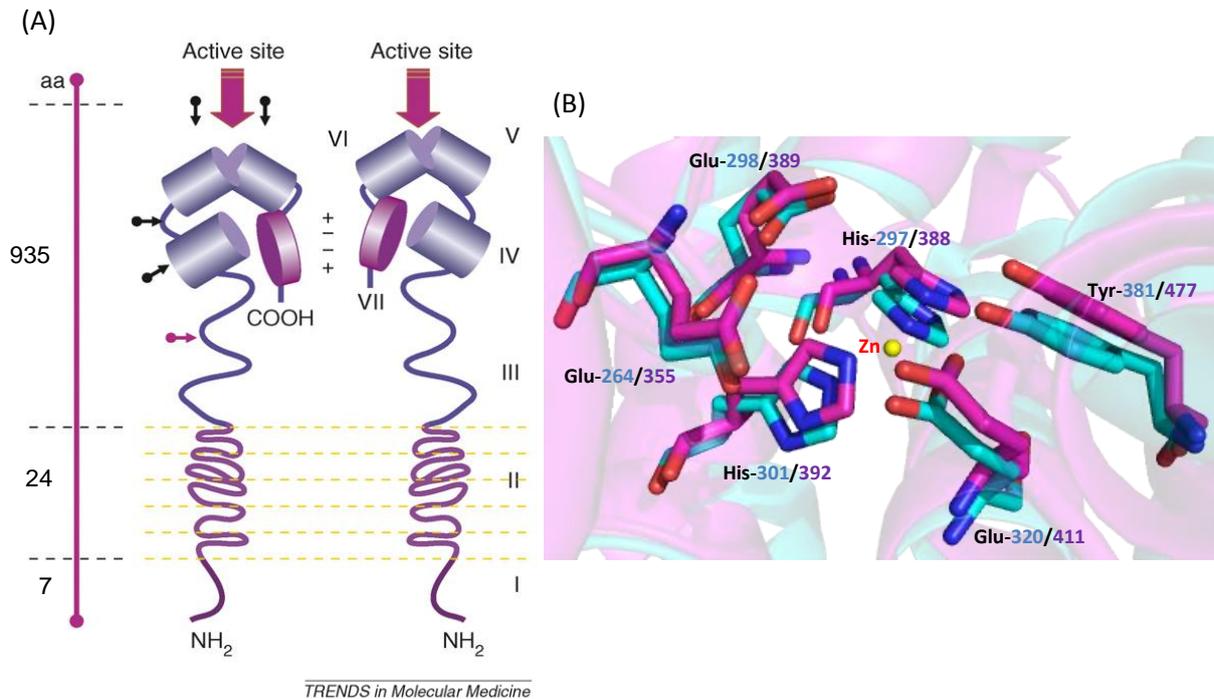
<sup>84</sup> Ito, K., Nakajima, Y., Onohara, Y., Takeo, M., Nakashima, K., Matsubara, F., Ito, T., Shimoto, T. (2006). *J. Biol. Chem.* **281** (44), 33664-33676.



*Figure 14 : Mécanisme catalytique proposé pour les aminopeptidases de la famille M1 d'après l'étude du mode de liaison de substrats peptidiques (en noir). Les acides aminés essentiels du site actif de l'APN d'Escherichia coli sont en bleu.*

Par ailleurs, le noyau aromatique en position N-terminale de la bestatine se dirige dans le sous-site S1 et forme des interactions- $\pi$  avec des phénylalanines (Tyr-376).

La réaction enzymatique décrite ci-contre est une catalyse générale acide/base où le résidu Glu-298 joue le rôle de « navette à proton ». Un substrat peptidique, qui se lie au site actif déplace la molécule d'eau liée au zinc vers l'acide glutamique (Glu-298) ; l'atome de zinc forme ainsi un complexe penta-coordonné (Figure 14-a). La molécule d'eau est activée à la fois par l'influence de l'atome de zinc, qui joue le rôle d'acide de Lewis, et par le résidu Glu-298 (base catalytique). Le pKa de la molécule d'eau est abaissé aux environs de 6, permettant ainsi la formation d'un groupement hydroxyle  $\text{HO}^-$  *in situ*<sup>64</sup>. L'attaque nucléophile de ce dernier sur la fonction carbonyle du substrat conduit à un intermédiaire tétraédrique commun à toutes les enzymes protéolytiques (Figure 14-b). Cet intermédiaire est stabilisé par un réseau de liaisons hydrogène ; le résidu Glu-298 (acide catalytique) protone alors l'atome d'azote de la liaison amide (bon groupe partant) (Figure 14-c), ce qui conduit à l'hydrolyse spontanée de la liaison amide (Figure 14-d).



**Figure 15 :** (A). Modèle schématique de la structure tridimensionnelle de l'APN de mammifère d'après l'analyse des structures secondaires<sup>48</sup>. (B). Modèle tridimensionnelle de l'APN humaine obtenu par homologie de séquence (structure connue d'ERAP1) à l'aide du serveur Swiss-Model (ExPASy). Les acides aminés du site actif essentiels à la catalyse (en magenta) sont superposés avec ceux de l'APN d'*E. coli* (en cyan) (code PDB : 2dq6). Le logiciel PyMOL est utilisé pour visualiser la structure tridimensionnelle de la protéine.

```

1  MAKGFYISKSLGILGILLGVAAVCTIIIALSVVYSQEKKNKNANSSPVASTTTPSASATTPNPASATTLTDQSKAWNRYR
76  LPNTLKPDSYRVTLRPYLTPNDRGLYVFKGSSTVRFCTKEATDVI I IHSKKLNYTLTSQGHRVVLRGVGGSQPPDI
151 DKTELVEPTEYLVVHLKGSVKDSQYEMDSEFEGELADDLAGFYRSEYMEGNVRKVVATTQMQAADARKSFPCTFD
226 EPAMKAEFNITLIHPKDLTALSNNLPGKGPSTPLPEDPNWNVTEFHTTPKMSTYLLAFIVSEFDYVEKQASNGVLI
301 RIWARPSAIAAGHGDYALNVTGPILNFFAGHYDTPYPLPKSDQIGLPDFNAGAMENWGLVITYRENSLLFDPLSSS
376 SSNKERVVTVIAHELAHQWFGNLTVEIWWNDLWLNELGFASYVEYLGADYAEPTWNLKDLMLVNDVYRVMAVDALA
451 SSHPLSTPASEINTPAQISELFDAISYKSGASVLRMLSSFLSEDFVKQGLASYLHTFAYQNTIYLNLDHDLQEA
526 NNRSIQLPPTVRDIMNRWTLQMGFPVITVDTSTGTLSQEHFLDPDSNVTRPSEFNVWVWVPIITSRDGRQQQDYW
601 LIDVRAQNDFLSTSGNEWVLLNLTNTGYRVRVNYDEENWRKIQTQLQRDHSAPVINRAQIINDAFNLASAHKVPV
676 TLALNNTLFLIEERQYMPWEAALSSLSYFKLMFDRSEVYGPKNYLKQVTPLFIFRNNNTNNWREIPENLMDQY
751 SEVNAISTACSNQVPECEEMVSGLFKQWMENPNNNPIHPNLRSTVYCNAAIQGGEEWDFAWEQFRNATLVNEAD
826 KLRALA CSKELWILNRYLSYTLNPDILRKQDATSTIIISITNNVIGQLVWDFVQSNWKKLFNDYGGGSFSFSNL
901 IQAVTRRFSTEYELQQLQLEQFKKDNEETGFGSGTRALEQALEKTKANIKWVKENKEVVLQWFTENSK

```

M : méthionine initiatrice (acide aminé 1).

MAKGFYIS : région cytoplasmique (acides aminés 2 à 8).

KS LGILGILLGV AAVCTIIALS VV : région transmembranaire- repliement hélicoïdal (acides aminés 9 à 32).

YSQEKKNKN ANSSPVASTT... : région extracellulaire (acides aminés 33 à 967).

GAMEN : Résidus impliqués dans la reconnaissance du substrat peptidique.

HEXXH (X<sub>1,3</sub>) E : motif de fixation du zinc.

N : sites potentiels de N-Glycosylation (N-acétylglucosamine ou GlcNAc) (10 sites).

C : Résidus Cystéine

**Figure 16 :** Séquence protéique de l'Aminopeptidase N de mammifère ; UniProtKB/Swiss-Prot ID P15144<sup>85</sup>.

<sup>85</sup> The UniProt Consortium. (2012). *Nucleic Acids Res.* **40**, D71-D75. <http://www.uniprot.org/>

### 3.3.3. L'Aminopeptidase-N de mammifères

L'Aminopeptidase-N de mammifère est une glycoprotéine transmembranaire de type II contenant 967 acides aminés. Elle est exprimée sous la forme d'un homodimère de hauts poids moléculaire (environ 280 kDa) ; les deux sous-unités sont associées de manière non-covalentes du côté C-terminal de la protéine.

Un modèle très schématique a pu être proposé d'après l'analyse des structures secondaires (Figure 15-A)<sup>48,86</sup>. Les chaînes polypeptidiques de cette ectoenzyme sont organisées en différents domaines ; un domaine N-terminal cytoplasmique de 7 acides aminés (I), un domaine transmembranaire constitué d'une hélice  $\alpha$  de 24 résidus hydrophobes (II), ainsi qu'un domaine extracellulaire (III à VII) contenant le site catalytique (résidus 253 à 580) (IV et V) et l'extrémité C-terminale de la protéine (VI et VII).

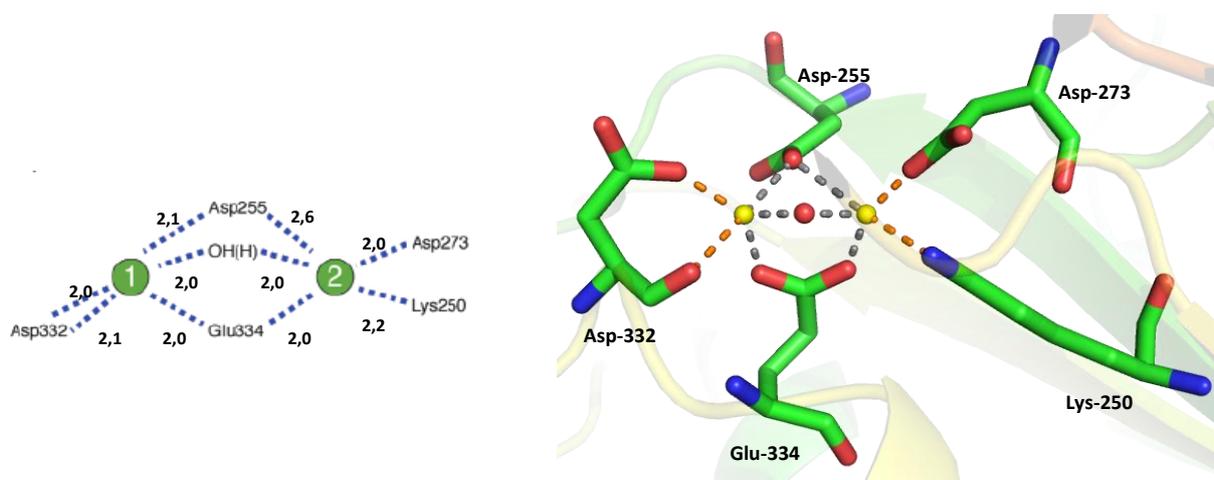
Le serveur Swiss-Model (Expasy) a été utilisé pour construire par homologie de séquence un modèle tridimensionnelle de h-APN, basé sur la structure connue d'ERAP1<sup>87</sup>. La superposition de ce dernier avec la structure de l'APN d'*E. coli* permet d'identifier les acides aminés conservés du site actif essentiels à la catalyse (Figure 15-B).

h-APN possède 10 sites potentiels de N-glycosylation, constituant approximativement 20% de sa masse moléculaire (Figure 16-en cyan). Sa séquence primaire révèle également la présence de 7 résidus Cystéine (Figure 16-en rouge) ; six de ces résidus sont situés dans la partie extracellulaire et peuvent, par conséquent, être impliqués dans la formation de ponts disulfures intramoléculaires, essentiels à la stabilisation de la structure tridimensionnelle de la protéine. Le modèle réalisé, précédemment, à l'aide du serveur Swiss-Model (Expasy) permet de prédire l'établissement d'une liaison covalente entre les résidus cystéines 761-768 et 798-834.

---

<sup>86</sup> Luan, Y., and Xu, W. (2007). *Curr. Med. Chem.* **14**, 639-674.

<sup>87</sup> Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). *Bioinformatics.* 22,195-201. <http://swissmodel.expasy.org/>



*Figure 17 : Leucine Aminopeptidase cytosolique de bovin code PDB : 1lam. Résidus impliqués dans la chélation des ions métalliques. Les distances entre les hétéroatomes et les ions métalliques sont données en Angström (Å). Le logiciel PyMOL est utilisé pour visualiser la structure tridimensionnelle de la protéine.*

### 3.4. La famille M17 : les aminopeptidases à deux ions zinc

Les aminopeptidases de la famille M17 doivent être considérées avec la plus grande attention puisqu'elles catalysent la même réaction d'hydrolyse que les membres de la famille M1.

Ces enzymes possèdent dans leur site actif deux ions zinc organisés en « unité co-catalytique »<sup>71</sup>.

L'homologie séquentielle est très faible entre les enzymes de ces deux familles ; l'APN humaine présente une identité de séquence de 9 % avec la Leucine Aminopeptidase cytosolique (LAPc) de mammifères et de 8 % avec l'Aminopeptidase de la bactérie d'algue marine *Aeromonas Proteolytica* (*Apaero*). Le repliement tridimensionnel adopté par les membres de ces familles est également très différent. Cette faible homologie séquentielle et structurale exclut l'existence d'un ancêtre commun pour les deux familles M1 et M17 des aminopeptidases zinc-dépendantes.

#### 3.4.1. La Leucine Aminopeptidase cytosolique

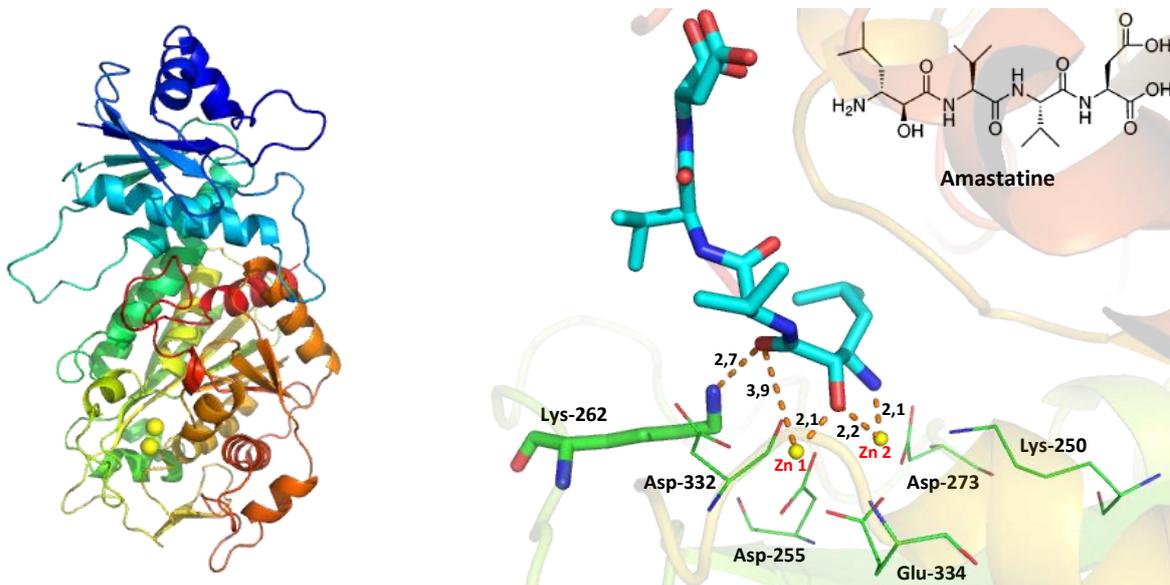
La Leucine Aminopeptidase cytosolique (LAPc) [EC.3.4.11.1] de mammifères est la première métallo-peptidase de cette famille dont la structure tridimensionnelle a été résolue (1990)<sup>88</sup>.

Comme son nom l'indique, elle hydrolyse des résidus leucine en position N-terminale d'un peptide ou d'une protéine ; mais elle reconnaît également des acides aminés hydrophobes comme la phénylalanine et l'alanine.

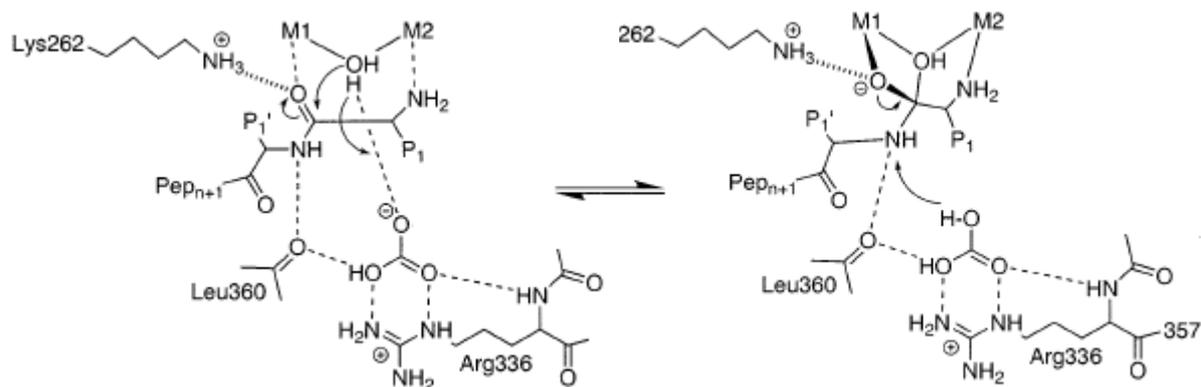
Les résidus impliqués dans la chélation des ions métalliques sont représentés dans la [Figure 17](#)<sup>69</sup>. Un des ions zinc (Zn 1) est lié par le carbonyle et le carboxylate du résidu Asp-332, et le second (Zn 2) interagit avec les groupements  $\epsilon$ -aminé du résidu Lys-250 et carboxylate du résidu Asp-273. Les deux ions métalliques, distants d'environ 3 Å, sont coordonnés aux carboxylates du résidu Glu-334 (ligand bidentate) et Asp-255 (ligand monodentate), ainsi qu'à une molécule d'eau (agent nucléophile).

---

<sup>88</sup> Burley, S.K.; David, P.R.; Taylor, A.; and Lipscomb, W.N. 1990. *Proc. Natl. Acad. Sci. USA* **87**, 6878-6882.



**Figure 18** : Structure tridimensionnelle de la Leucine Aminopeptidase cytosolique de bovin et mode de liaison de l'analogue de transition, l'amastatine (produit naturel)<sup>83</sup> au site actif de l'enzyme: code PDB : 1bl<sup>89</sup>. Les distances entre les hétéroatomes et les ions métalliques sont données en Angström (Å). Le logiciel PyMOL est utilisé pour visualiser la structure tridimensionnelle de la protéine.



**Figure 19** : Mécanisme catalytique proposé de la Leucine Aminopeptidase cytosolique<sup>90</sup>.

L'ion Zn 1 est lié moins fortement au site actif de l'enzyme et peut, par conséquent, être substitué stœchiométriquement par d'autres cations divalents tels que  $Mn^{2+}$ ,  $Mg^{2+}$  et  $Co^{2+}$ <sup>89</sup>. Les échanges au niveau du second site métallique (ZN 2) sont difficiles ; ils sont possibles lorsque les deux sites métalliques sont simultanément inoccupés. Dans ce cas précis, le site métallique peut lier un cation  $Co^{2+}$ . Néanmoins, ces substitutions métalliques affectent de manière significative la constante catalytique  $k_{cat}$  et la constante de Michaelis-Menten  $K_M$ <sup>90</sup>.

Un mécanisme catalytique a été proposé d'après l'étude du mode de liaison d'analogues de l'état de transition, comme par exemple l'amastatine (produit naturel) ([Figure 18](#)) ; il est similaire à celui décrit précédemment pour les aminopeptidases de la famille M1 (catalyse acide-base).

Les membres de la famille M17 ne possèdent pas le motif « GXMEN », l'amine libre du substrat peptidique est reconnue par l'enzyme via son interaction avec l'ion Zn 2. La fonction carbonyle de la liaison peptidique à hydrolyser interagit, quant-à-elle, avec l'ion Zn 1 et forme une liaison hydrogène avec le groupement  $\epsilon$ -aminé du résidu Lys-262 ; ce dernier est impliqué, avec l'ion Zn 1, dans la stabilisation de l'état de transition.

Par ailleurs, des analyses structurales ont révélé la présence, dans le site actif, d'un anion bicarbonate  $HCO_3^-$ , lié au résidu Arg-336. Cette molécule est décrite comme favorisant le transfert d'un proton de la molécule d'eau liée aux ions métalliques vers l'atome d'azote de la liaison amide à cliver (formation d'un bon groupe partant) ([Figure 19](#)).

### 3.4.2. L'Aminopeptidase d'*Aeromonas Proteolytica*

La structure tridimensionnelle de l'Aminopeptidase d'*Aeromonas Proteolytica* (*Apaero*) [EC.3.4.11.10] a été résolue en 1994<sup>91</sup>. Cette métallo-peptidase, difficile à classer, peut être un membre des familles M17 et M28 ; la famille M28 regroupe des aminopeptidases et des carboxypeptidases dans lesquelles les ions zinc sont liés par un acide aspartique<sup>71</sup>.

L'*Apaero* hydrolyse les résidus hydrophobes en position N-terminale d'un peptide ou d'une protéine avec une préférence pour la Leucine.

<sup>89</sup> Kim, H., and Lipscomb, W. N. (1993). *Biochemistry*. 90, 5006-5010.

<sup>90</sup> Lowther, W. T., and Matthews, B. W. (2002). *Chem. Rev.* **102**, 4581-4607.

<sup>91</sup> Chevrier, B., Schalk, C., D'Orchymont, H., Rondeau, J. M., Moras, D., and Tarnus, C. (1994). *Structure*. **2**, 283-291.

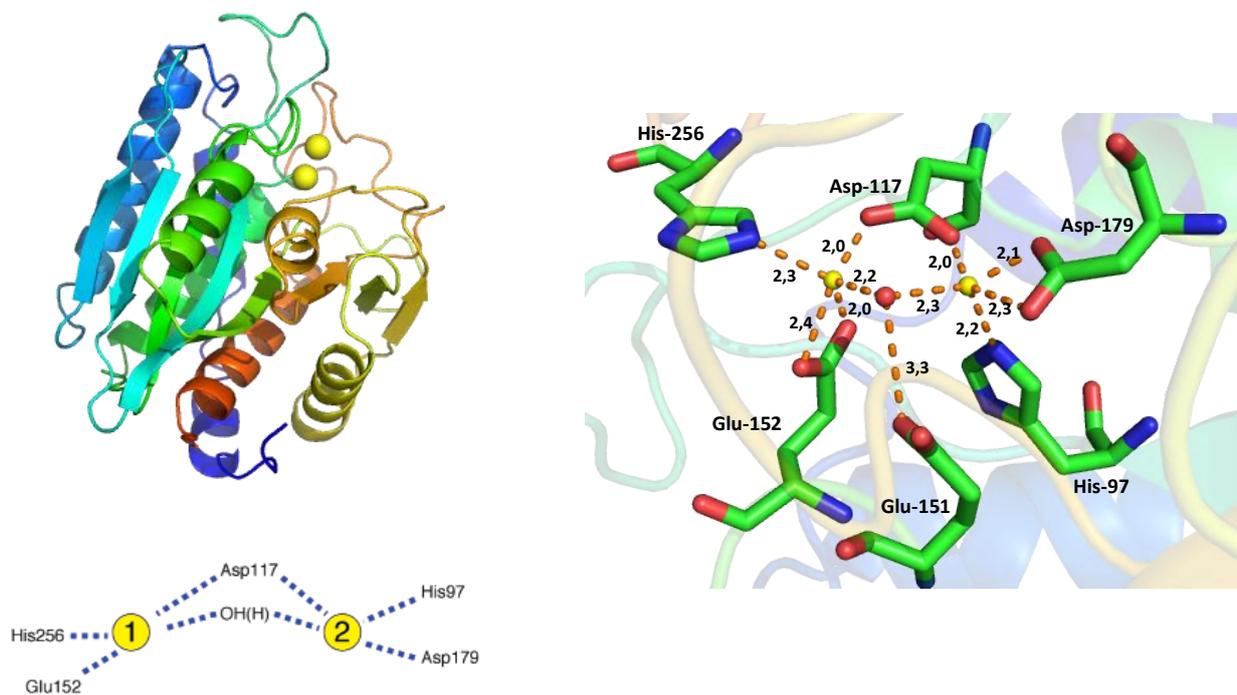


Figure 20 : Structure tridimensionnelle de l'Aminopeptidase d'Aeromonas Proteolytica code PDB : 1amp. Résidus impliqués dans la chélation des ions métalliques. Les distances entre les hétéroatomes et les ions métalliques sont données en Angström (Å). Le logiciel PyMOL est utilisé pour visualiser la structure tridimensionnelle de la protéine.

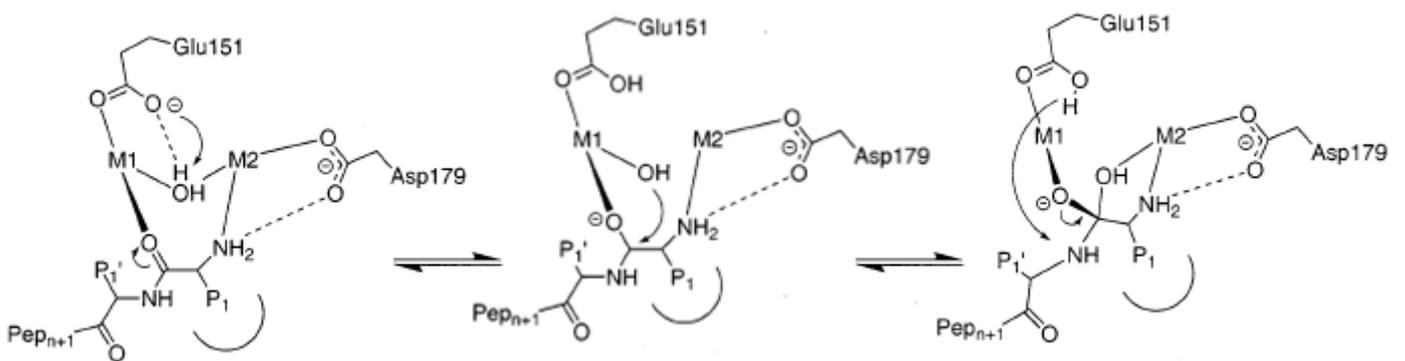


Figure 21 : Mécanisme catalytique proposé de l'Aminopeptidase d'Aeromonas proteolytica<sup>90</sup>.

Les deux ions zinc de l'*Apaero* forment une « unité co-catalytique » symétrique. L'ion Zn 1 est coordonné par les résidus His-256 et Glu-152 et l'ion Zn 2 par les résidus His-97 et Asp-179. Les deux ions métalliques, distants d'environ 3,5 Å, interagissent avec le groupement carboxylate du résidu Asp-117 (ligand bidentate) et avec une molécule d'eau (agent nucléophile). Cette dernière est également liée par liaison hydrogène au carboxylate du résidu Glu-151 (« navette à proton » lors de la catalyse acide-base) ([Figure 20](#)).

Le mécanisme catalytique est similaire à celui décrit précédemment pour la LAPc ([Figure 21](#)). L'ion Zn 2 participe avec le résidu Asp-179 à la reconnaissance du substrat peptidique, alors que l'ion Zn 1 lie la fonction carbonyle du peptide à hydrolyser, et est impliqué dans la stabilisation de l'état de transition. Le résidu Glu-151 favorise la déprotonation de la molécule d'eau liée aux ions métalliques, suivie de la protonation de l'azote de la liaison peptidique à hydrolyser<sup>92</sup>.

### 3.5. Les inhibiteurs de l'Aminopeptidase-N

L'existence de ces différentes familles de métallo-aminopeptidases, ayant des spécificités de substrats très proches ainsi que des mécanismes catalytiques similaires, représente une réelle difficulté pour la conception d'inhibiteurs sélectifs d'une aminopeptidase donnée. Parmi les composés décrits à ce jour, peu présentent une sélectivité intéressante.

#### 3.5.1. Produits naturels

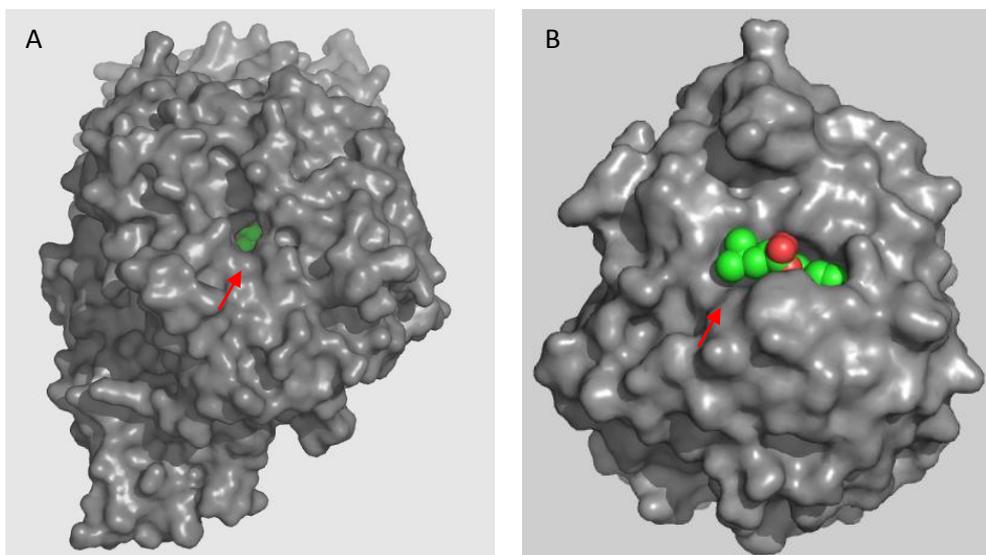
De nombreux produits naturels, isolés de bactéries du genre *Streptomyces*, se sont avérés être des inhibiteurs des aminopeptidases<sup>83</sup>. Ces molécules ne sont cependant pas sélectives de l'APN, et inhibent également les autres aminopeptidases zinc-dépendantes. Une sélection de ces inhibiteurs compétitifs est représentée dans le [Tableau 3](#)<sup>93,94</sup>. Ce sont des analogues partiels de l'état de transition (bestatine, amastatine et probestine) ou des dérivés d'acides hydroxamiques, groupement chélateur du zinc (actinonine).

---

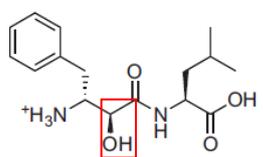
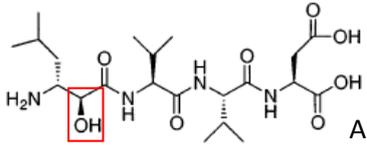
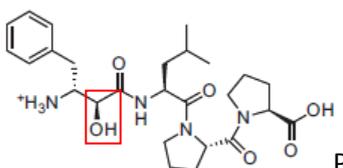
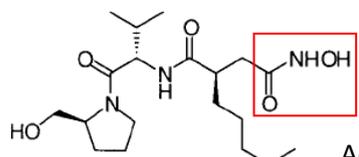
<sup>92</sup> Holz, R. C. (2002). *Coordination Chemistry Reviews*. **232**, 5-26.

<sup>93</sup> Mucha, A., Drag, M., Dalton, J. P., and Kafarski, P. (2010). *Biochimie*. **92**, 1509-1529.

<sup>94</sup> Bauvois, B., and Dauzonne, D. (2006). *Medicinal Research Reviews*. **26** (1), 88-130.



*Figure 22 : Accessibilité des substrats peptidiques au site actif des métallo-protéases de la famille M1 et M17. (A). Structure tridimensionnelle de l'Aminopeptidase N d'Escherichia coli : code PDB 2dqm. (B). Structure tridimensionnelle de l'aminopeptidase d'Aeromonas Proteolytica : code PDB : 1xry. Représentation de la surface des protéines ; le logiciel PyMOL est utilisé pour visualiser la structure tridimensionnelle de la protéine.*

Composés	Constante d'inhibition $K_i$ ( $\mu\text{M}$ )			
	AP-N EC 3.4.11.2	LTA4H EC 3.3.2.6	LAPc EC 3.4.11.1	<i>Apaero</i> EC 3.4.11.10
 Bestatine	3,5	0,5	0,0005	0,0016
 Amastatine	0,05	/	0,03	0,0003
 Probestine	0,02	/	0,06	/
 Actinonine	1	/	1,2	/

\*Constante de dissociation du complexe enzyme-inhibiteur.

**Tableau 3 :** Constantes d'inhibition  $K_i$  d'une sélection de produits naturels inhibant les métallo-protéases à un ion zinc, l'Aminopeptidase N (AP-N, [EC.3.4.11.2]), la Leucotriène A4 Hydrolase (LTA4H, [EC.3.3.2.6]), les aminopeptidases à deux ions zinc organisés en « unité co-catalytique », la Leucine Aminopeptidase (LAPc, [EC.3.4.11.1]) et l'Aminopeptidase d'*Aeromonas Proteolytica* (*Apaero*, [EC.3.4.11.10]).

L'amastatine, la probestine et l'actinonine sont des inhibiteurs compétitifs des deux sous-classes d'aminopeptidases. La bestatine est, quant-à-elle, plus active sur les aminopeptidases à deux ions zinc. Pour comprendre cette sélectivité, le mode de liaison de la bestatine au site actif de pepN (famille M1) et de l'*Apaero* (famille M17) est analysé.

L'accessibilité au site actif est différente pour les deux familles de métallo-peptidases (M1 et M17) (Figure 22). Le site actif des membres de la famille M1 est enfoui dans la protéine, dans un domaine constitué d'hélices  $\alpha$  et de feuillets  $\beta$  ; une fine cavité (8 Å de diamètre) permet au substrat peptidique (de préférence linéaire) de rejoindre le domaine catalytique par l'intermédiaire du domaine C-terminale. Le site actif des aminopeptidases de la famille M17 se situe, quant-à-lui, à la surface de l'enzyme sur des feuillets  $\beta$ , permettant ainsi l'hydrolyse de substrats peptidiques de plus fort encombrement à l'extrémité N-terminale.

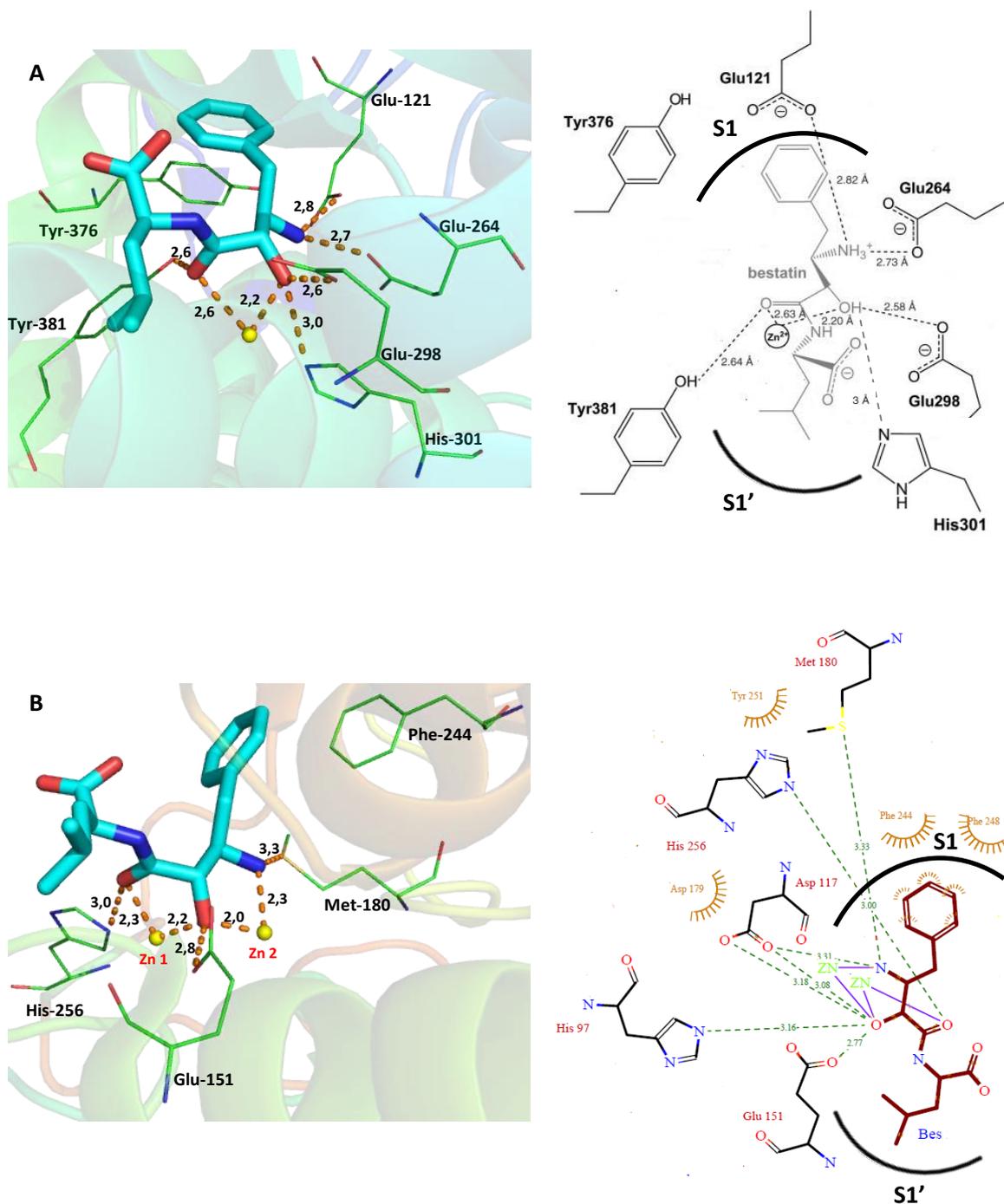


Figure 23 : Mode de liaison de la Bestatine au site actif (A) de l'Amino-peptidase-N d'Escherichia coli (code PDB : 2dqm)<sup>84</sup> et (B) de l'Amino-peptidase d'Aeromonas Proteolytica (code PDB : 1xry)<sup>95</sup>. Les distances entre les hétéroatomes et les ions métalliques sont données en Angström (Å). Le logiciel PyMOL est utilisé pour visualiser la structure tridimensionnelle de la protéine.

<sup>95</sup> Stamper, C.C., Bienvenue, D.L., Bennett, B., Ringe, D., Petsko, G.A., and Holz, R.C. (2004) *Biochemistry*. **43**, 9620-9628.

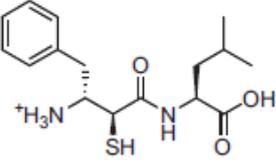
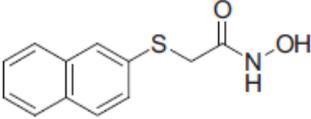
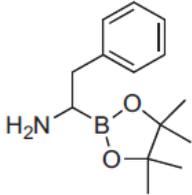
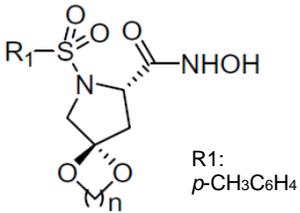
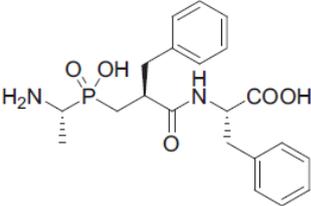
La bestatine est un inhibiteur compétitif, qui mime partiellement l'état de transition formé au cours de la réaction catalytique ; sa fixation au site actif de l'enzyme est similaire à celle d'un substrat peptidique : l'amine primaire de la bestatine interagit avec les résidus du motif GAMEN chez pepN et avec l'ion Zn 2 chez *Apaeo*. La fonction cétone se lie à l'ion métallique (l'ion Zn 1 dans le cas des aminopeptidases à deux ions zinc). Le groupement hydroxyle de la bestatine occupe, quant-à-lui, la place de la molécule d'eau (agent nucléophile) ; il se lie au(x) zinc et forme une liaison hydrogène avec le résidu Acide Glutamique impliqué dans la catalyse (« navette à proton »). Par ailleurs, le noyau aromatique en position N-terminale de la bestatine se dirige dans le sous-site S1 et forme des interactions- $\pi$  avec des phénylalanines (Figure 23).

La bestatine est un inhibiteur à interaction lente (« slow binding ») des aminopeptidases à « unité co-catalytique ». Le complexe enzyme-inhibiteur qui se forme est plus stable car la fonction hydroxyle se positionne au centre des deux ions métalliques, qui interagissent également avec les groupements carboxyle (Zn 1) et amine primaire (Zn 2) de la bestatine (coordination tridentate).

### 3.5.2. Les inhibiteurs synthétiques

Depuis de nombreuses années, plusieurs équipes ont travaillé sur la conception et la synthèse d'inhibiteurs de l'APN (Tableau 4) ; la plupart de ces molécules sont des analogues de l'intermédiaire tétraédrique de l'état de transition ou des dérivés d'acides hydroxamiques (groupement chélateur du zinc)<sup>93</sup>. Beaucoup de ces composés présentent une bonne activité *in vitro*, néanmoins, leur sélectivité vis-à-vis d'autres métallo-aminopeptidases n'est pas toujours bien documentée.



Composés		Constante d'inhibition $K_i$ ( $\mu\text{M}$ ) <sup>a</sup>	
		AP-N EC 3.4.11.2	LAPc EC 3.4.11.1
Dérivés de la Bestatine (Analogue partiel de l'état de transition)		4,4	0,55
Dérivés d'acides hydroxamiques (Chélateur du zinc)		3,4 <sup>b</sup>	/
Dérivés d'acides boroniques (Analogue de l'état de transition)		0,02 <sup>b</sup>	0,05
Dérivés de sulfonamides (Chélateur du zinc)		50 <sup>b</sup>	/
Dérivés d'acides phosphoniques (Analogue de l'état de transition)		0,002	/

<sup>a</sup> Constante d'inhibition  $K_i$  ( $\mu\text{M}$ ).

<sup>b</sup>  $IC_{50}$  ( $\mu\text{M}$ ) : Concentration d'un composé inhibant 50 % de l'effet observé.  
Pour un inhibiteur compétitif  $IC_{50} = (1 + [S]/K_M) \cdot K_i$  donc à  $[S] = K_M \Rightarrow IC_{50} = 2K_i$ .

**Tableau 4 :** Inhibiteurs synthétiques de l'APN. Liste non exhaustive, représentative des principaux concepts d'inhibition.

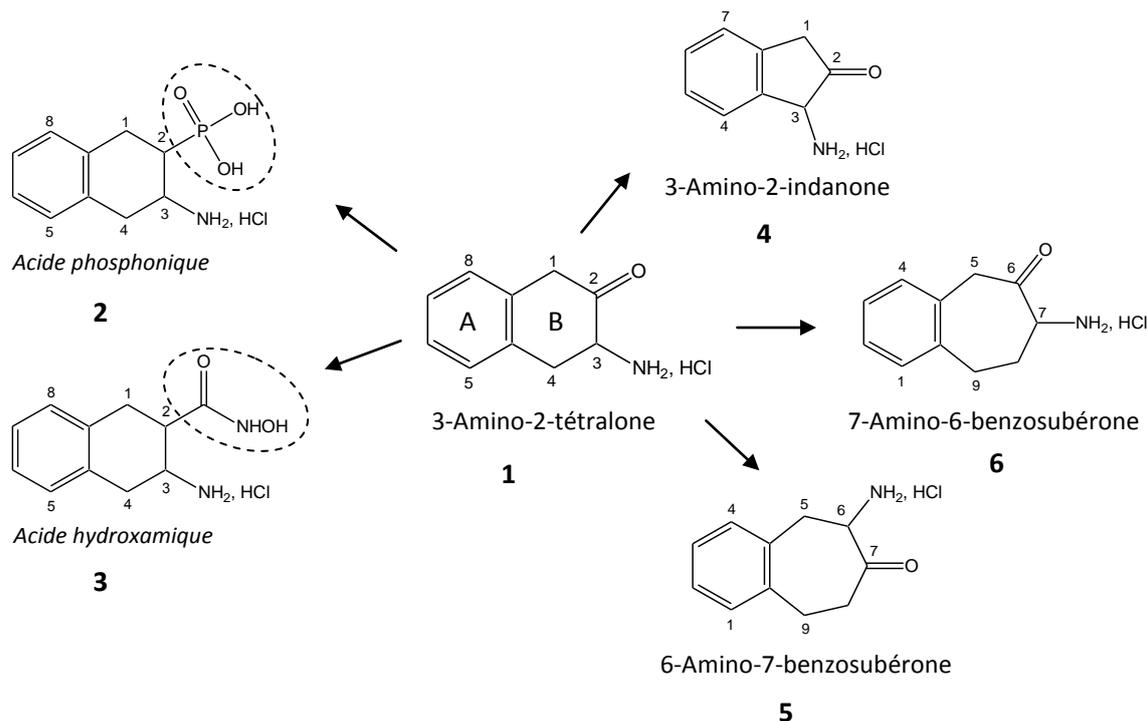


Figure 24 : Sélection d'inhibiteurs compétitifs de l'APN synthésés au laboratoire.

Composés	IC <sub>50</sub> (μM) <sup>a</sup>			
	APN EC 3.4.11.2 (porcin)	LTA4H EC 3.3.2.6 (humain)	LAPc EC 3.4.11.1 (bovin)	Apaero EC 3.4.11.10
1	0,5 <sup>b</sup>	>1000	120	130
2	500	>1000	>1000	150
3	4 <sup>b</sup>	> 1000	10 <sup>b</sup>	4 <sup>b</sup>
4	20 <sup>b</sup>	>1000	>1000	150
5	>1000	>1000	>1000	>1000
6	1 <sup>b</sup>	>1000	>1000	900

<sup>a</sup> Concentration d'un composé inhibant 50 % de l'effet observé. IC<sub>50</sub> déterminée à partir de la représentation de Dixon pour [Substrat]=K<sub>M</sub>. Pour un inhibiteur compétitif IC<sub>50</sub>= (1+[S]/K<sub>M</sub>).K<sub>i</sub> donc à [S]=K<sub>M</sub> => IC<sub>50</sub>= 2K<sub>i</sub>.

<sup>b</sup> K<sub>i</sub> (μM) : Constante de dissociation du complexe enzyme-inhibiteur. K<sub>i</sub> déterminée à partir de la représentation de Dixon pour [S]= K<sub>M</sub>, [S]= 2K<sub>M</sub> et [S]= K<sub>M</sub>/2.

Tableau 5 : Sélection d'inhibiteurs de l'APN synthésés au laboratoire et valeurs des constantes d'inhibition K<sub>i</sub>, obtenues à partir de la représentation de Dixon<sup>96</sup>.

<sup>96</sup> Segel, I. H. (1993). Enzyme Kinetics. John Wiley & Sons, Inc., New-York.

### 3.5.3. Travaux du laboratoire

Les travaux du laboratoire se sont concentrés, dans un premier temps, sur la conception, la synthèse et l'évaluation d'inhibiteurs compétitifs de l'APN, dérivés de la 3-amino-2-tétralone. La structure tridimensionnelle de l'enzyme de mammifères n'étant pas résolue, ces composés ont été conçus d'après l'étude des relations structure-activité (chimie médicinale classique). Ils ont été évalués *in vitro*, au laboratoire, sur les métallo-aminopeptidases des familles M1 (APN et LTA4H) et M17 (LAPc et *Apaero*).

La 3-amino-2-tétralone est un inhibiteur puissant et sélectif de l'APN ( $K_i = 0,5 \mu\text{M}$ )<sup>97</sup>. Ce composé est néanmoins instable en solution aqueuse à pH physiologique. Une aromatisation du cycle B, qui conduit à un composé inactif, est observée. Elle est liée à l'acidité des protons en  $\alpha$  de la cétone, conduisant à une énolisation rapide en position benzylique. Ainsi, de nombreux dérivés ont été synthétisés afin d'améliorer la stabilité (Figure 24)<sup>98,99</sup>.

La fonction cétone, en position 2 de la 3-amino-2-tétralone **1** étant à l'origine de ces modifications irréversibles, nous avons cherché, dans un premier temps, à substituer cette position par différents groupements, comme par exemple l'acide phosphonique **2** ou l'acide hydroxamique **3**, susceptibles de chélater l'ion zinc essentiel à la catalyse. Le dérivé d'acide phosphonique **2** est inactif sur la LTA4H et la LAPc, et de très faible activité sur l'APN et l'*Apaero* (Tableau 5). Par ailleurs, l'introduction d'un acide hydroxamique **3**, puissant chélateur du zinc, conduit à un composé actif mais non sélectif, qui inhibe les enzymes des deux familles M1 et M17.

---

<sup>97</sup> Schalk, C., Dorchymont, H., Jauch, M. F., and Tarnus, C. (1994). *Arch. Biochem. Biophys.* **311**, 42-46.

<sup>98</sup> Albrecht, S., Defoin, A., Salomon, E., Tarnus, C., Wetterholm, A., and Haeggström, J. Z. (2006). *Bioorg. Med. Chem.* **14**, 7241-7257.

<sup>99</sup> Albrecht, S., Al-Lakkis-Wehbe, M., Orsini, A., Defoin, A., Pale, P., Salomon, E., Tarnus, C., and Weibel, J. M. (2011). *Bioorg. Med. Chem.* **19**, 1434-1449.

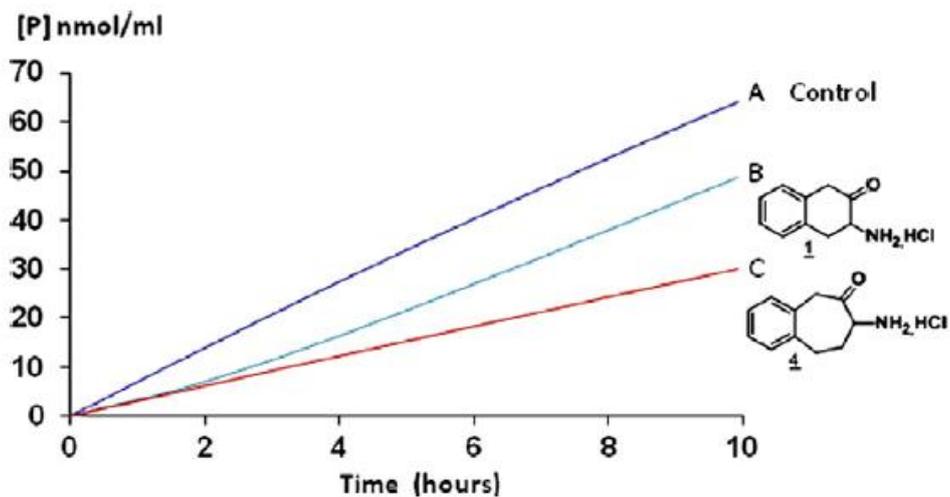


Figure 25 : Etude de stabilité<sup>99</sup> ; mesure de l'activité enzymatique de l'APN en absence (courbe A) ou en présence de 2  $\mu\text{M}$  de la 3-amino-2-tétralone (courbe B) et de 2  $\mu\text{M}$  de la 7-amino-6-benzosubérone (courbe C). La leucine p-nitroanilide est utilisée comme substrat de l'APN ( $K_M = 0,2 \text{ mM}$ ) ; son hydrolyse conduit à la p-nitroaniline qui absorbe dans le visible à 405 nm ( $\epsilon = 10800 \text{ mol}^{-1} \cdot \text{L} \cdot \text{cm}^{-1}$ ). La réaction enzymatique se déroule à 25°C dans du tampon 20 mM Tris-HCl pH 7,5 (volume final de 1ml), et débute après l'ajout de 3mU d'enzyme. La vitesse initiale de la réaction est alors mesurée durant 10 heures en suivant, par spectrophotométrie, la cinétique d'apparition du p-nitroaniline à 405 nm. Courbe A : contrôle ( $V_0 = 6,8 \text{ nmol/h}$ ). Courbe B : durant la première heure, une inhibition de 60% est observée ( $V_i = 3,1 \text{ nmol/h}$ ). Après 8 heures, plus aucune inhibition n'est observée ( $V_i = V_0 = 6,5 \text{ nmol/h}$ ). Courbe C : Une inhibition de 60% est observée durant toute l'expérience ( $V_i = 3,1 \text{ nmol/h}$ ).

L'influence de la taille du cycle B a alors été étudiée. Des dérivés comportant des cycles à 5 et 7 chaînons ont été synthétisés et évalués. La 7-amino-6-benzosubérone **6** s'est révélée le meilleur composé pour poursuivre nos études. Cet inhibiteur tête de série présente une constante d'inhibition  $K_i$  de 1  $\mu\text{M}$  et une efficacité de ligand tout à fait remarquable ( $LE = 0,63$ )<sup>100</sup>. Ce composé est stable en solution à pH physiologique comme le montre l'étude cinétique Figure 25<sup>99</sup>. *In vitro* l'inhibition est maintenue pendant 10 heures (courbe C), alors que la 3-amino-2-tétralone perd son activité inhibitrice après 2 heures (courbe B). De plus, la sélectivité vis-à-vis des enzymes bimétalliques est conservée (Figure 5).

Ainsi, les études qui suivent ont consisté à substituer ce composé tête de série sélectif de l'APN de mammifères afin d'augmenter son affinité pour cette dernière.

---

<sup>100</sup> Hopkins, A. L., Groom, C. R., Alex, A. (2004). *Drug Discov. Today*. **9**, 430-431.



## CHAPITRE 1 :

Evaluation *in vitro* et *in vivo* d'inhibiteurs  
puissants et sélectifs de l'Aminopeptidase-N/CD13 :  
intérêt en oncologie



## 1. Publication n°1 : A novel amino-benzosuberone derivative is a picomolar inhibitor of mammalian aminopeptidase N/CD13

Les travaux du laboratoire sur la 3-amino-2-tétralone ont conduit à la découverte d'un inhibiteur puissant et sélectif de l'APN : la 7-amino-6-benzosubérone. Différents analogues de ce composé ont été synthétisés, et ont fait l'objet d'un brevet (Wo8059141, 2008) (Figure 26).

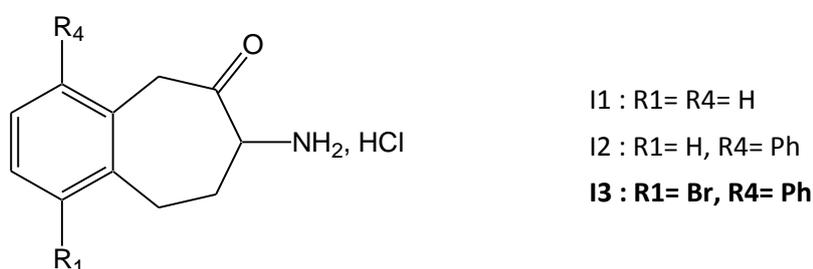


Figure 26 : Exemples de dérivés de la 7-amino-6-benzosubérone synthétisés au laboratoire.

Dans le cadre de cette étude, il m'a été demandé de purifier les enzymes cibles, non commerciales, à partir de tissus, puis, d'évaluer les composés *in vitro* sur l'Aminopeptidase-N de mammifères [EC.3.4.11.2], ainsi que sur les métallo-aminopeptidases de la famille M17, la Leucine Aminopeptidase cytosolique (LAPc) [EC.3.4.11.1], et l'Aminopeptidase d'*Aeromonas proteolytica* (*Apaero*) [EC.3.4.11.10].

Il s'est avéré que ces composés ont des constantes d'inhibition allant du nano au picomolaire, et sont les inhibiteurs les plus puissants et les plus sélectifs de l'APN connus à ce jour (I1 :  $K_i = 1 \mu\text{M}$ , I2 :  $K_i = 7 \text{ nM}$ , et **I3 :  $K_i = 60 \text{ pM}$** ). I3 a été identifié comme le composé tête de série.

Les résultats décrits ci-dessus ont donné lieu à la publication suivante.





## A novel amino-benzosuberone derivative is a picomolar inhibitor of mammalian aminopeptidase N/CD13

Carmen Maiereanu<sup>a</sup>, Céline Schmitt<sup>a</sup>, Nadège Schifano-Faux<sup>b</sup>, Didier Le Nouën<sup>a</sup>, Albert Defoin<sup>a,\*</sup>, Céline Tarnus<sup>a,\*</sup>

<sup>a</sup>Université de Haute Alsace, Laboratoire de Chimie Organique et Bioorganique, EA4466 ENSCMu, 3, rue Alfred Werner, F-68093, Mulhouse Cedex, France

<sup>b</sup>Université de Lille 2, Laboratoire de Chimie Analytique, EA GRIOT 4481 Faculté des Sciences Pharmaceutiques et Biologiques de Lille, 3, rue du Professeur Laguesse, B.P. 83, 59006 Lille cedex, France

### ARTICLE INFO

#### Article history:

Received 21 March 2011

Revised 27 May 2011

Accepted 8 June 2011

Available online 19 July 2011

#### Keywords:

Aminobenzoheptenone

Aminopeptidase-N inhibitor

APN/CD13

### ABSTRACT

A new class of low molecular weight, highly potent and selective non peptidic inhibitors of aminopeptidase N (APN/CD13) is described. We report the synthesis and in vitro evaluation of racemic substituted analogues of 7-amino-benzocycloheptan-6-one **1a**. We investigated various substitutions on the aromatic ring with phenyl and halogen groups. In vitro kinetic studies revealed that these compounds are among the most effective APN/CD13 inhibitors found so far. Hydrophobic substituents placed at position 1 or 4 on the cycloheptanone **1a** led to the potent compounds **1c-h,b'-c',f,h'** with  $K_i$  in the nanomolar range. The key finding of the present work was the observed additive effect of 1,4-disubstitutions which led to the discovery of the picomolar inhibitor **1d'** ( $K_i = 60$  pM). The designed inhibitors retain the selectivity of our lead structure **1a** towards selected members of the aminopeptidase family, combined with an impressive increase in inhibitory potency and a conserved stability.

© 2011 Elsevier Ltd. All rights reserved.

### 1. Introduction

APN/CD13 is a membrane-bound, zinc-dependent homodimeric enzyme and a member of the M1 family of aminopeptidases.<sup>1</sup> Like other members of this family, APN/CD13 possesses the consensus zinc binding motif (HEXXH-(X18)-E) in its extracellular domain, as well as the exopeptidase motif (GXMEN) for binding the free primary amino group of the N-terminal residue of its peptidic substrates. APN/CD13 removes the N-terminal amino acid from unsubstituted oligopeptides, amides or arylamides, with a broad substrate specificity,<sup>1</sup> although a significant preference for hydrophobic residues is observed.<sup>1</sup> This ectoenzyme appears to be a multifunctional protein involved in the regulation of signalling peptides as well as in various cell activation and migration processes.<sup>2</sup> APN/CD13 is emerging as a target of significant biological and medical importance. Indeed, several studies with bestatin,<sup>3</sup> active site-directed anti-APN mAb,<sup>4</sup> siRNA<sup>5</sup> and KO mice,<sup>6</sup> indicate that APN/CD13 is an active player in angiogenesis and tumor metastasis. In addition, overexpression of APN/CD13 has been found to correlate with immunological abnormalities such as chronic inflammatory diseases<sup>7</sup> and autoimmune pathologies,<sup>8</sup> suggesting a role of APN/CD13 in T cell function and activation. However, while a wealth of in vitro data have already been gathered, in vivo data on APN/CD13 blockade remain very scarce. Fur-

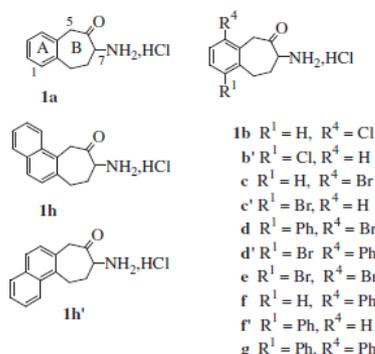
thermore, the exact role played by APN/CD13 in the pathologies mentioned above, as well as the cellular pathways involved, remain to be elucidated. To this end, a small molecular weight, drug-like, selective inhibitor of APN/CD13 would be of immense value for dissecting the biological and pathophysiological roles of this multifunctional enzyme that depend solely on its catalytic activity.

In the recent years, several classes of APN/CD13 inhibitors have been reported, including in particular hydroxamic, phosphoric, sulfonic and boronic acids. Recent reviews provide an excellent compendium of the currently available set of APN/CD13 inhibitors.<sup>9,10</sup> While many of these compounds display high in vitro potency, their selectivity is, regrettably, not always well documented.

We have recently reported the discovery of 7-amino-benzocycloheptan-6-one **1a** as a novel lead structure for selective APN/CD13 inhibition (Scheme 1).<sup>11</sup> This new chemotype displays an excellent ligand efficiency (LE = 0.63, according to the definition by Hopkins et al.<sup>12</sup>) and, therefore, represents a promising starting point for further chemical elaboration. Taking into account the preference of APN/CD13 for hydrophobic N-terminal amino acid residues,<sup>1</sup> we designed hydrophobic analogues of our lead compound **1a**. We first investigated the extension of the aromatic system. This strategy led to the identification of the submicromolar inhibitors **1h** and **1h'**. We then explored substitutions of cycle A in position 1 and/or 4, and discovered compounds **1f** and **1e** with  $K_i$  values in the single-digit nanomolar range. When both positions were optimally substituted, an additive effect was obtained, leading to the picomolar inhibitor **1d'**.

\* Corresponding authors.

E-mail addresses: [Albert.Defoin@uha.fr](mailto:Albert.Defoin@uha.fr) (A. Defoin), [Celine.Tarnus@uha.fr](mailto:Celine.Tarnus@uha.fr) (C. Tarnus).



**Scheme 1.** Chemical structures of the novel APN/CD13 inhibitors.

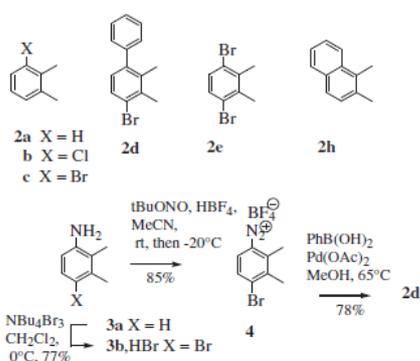
Hereafter, we describe the synthesis of these novel amino-benzosuberone analogues and report their inhibitory activities against a panel of representative metallo-aminopeptidases from the M1 and M17 families,<sup>13,14</sup> including porcine kidney aminopeptidase N/CD13 (EC 3.4.11.2), bovine kidney leucine aminopeptidase (EC 3.4.11.1), *Aeromonas proteolytica* aminopeptidase (EC 3.4.11.10) and the aminopeptidase activity of human leukotriene A4 hydrolase (EC 3.3.2.6). Our data show that the amino-benzocycloheptanone (amino-benzosuberone) derivative **1d'** is the most potent inhibitor of mammalian APN/CD13 known to date ( $K_i = 60$  pM), and that it has a very high selectivity towards the M1 subfamily of one-zinc aminopeptidases, as opposed to the M17 subfamily which requires co-catalytic metal ions for activity.

## 2. Chemistry

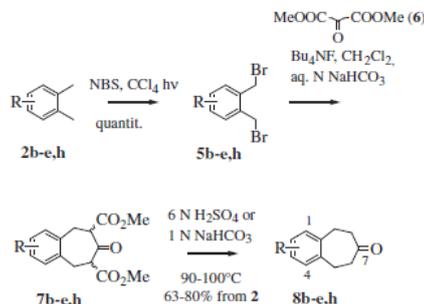
### 2.1. Preparation of the key intermediate ketones **8b–e,h**

We followed the reaction scheme already described for the benzocycloheptanone **1a**<sup>11</sup> from the  $\alpha,\alpha'$ -dibromo derivative of the *o*-xylene derivatives **2a**. In the present work, we used the corresponding commercial *o*-xylene derivatives **2b,c,h** or known dibromo one **2e**<sup>15</sup> as starting material. The bromodimethylbiphenyl **2d** was prepared according to the literature<sup>16</sup> (scheme 2) by coupling the phenylboronic acid with the diazonium salt **4**. This salt was easily obtained from the 2,3-dimethylaniline **3a** in ca. 50% overall yield, by bromination with an ammonium tribromide according to<sup>17</sup> and diazotation.

The general reaction scheme for the synthesis of the ketone intermediates is depicted in Scheme 3. A photochemical bis-bromination of the xylenes **2b–e,h** with two equivalents of *N*-bromosuccinimide



**Scheme 2.** Starting *o*-xylene derivatives and synthesis of the 4-bromo-2,3-dimethylbiphenyl **2d**.



**Scheme 3.** Common synthesis of the intermediate ketones **8b–e,h**.

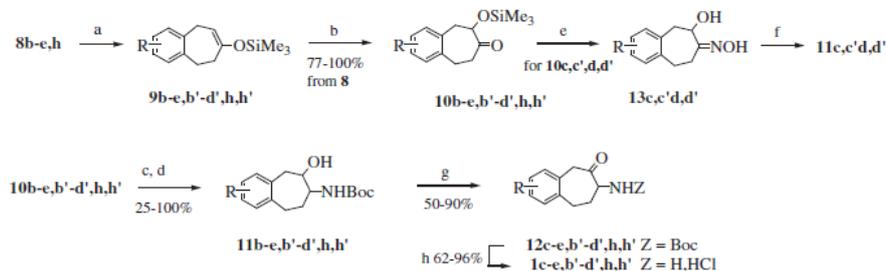
(NBS) gave quantitatively the known  $\alpha,\alpha'$ -dibromoxylenes **5b–c,e,h** and the new **5d**. They were cyclised with the dimethyl acetonedicarboxylate **6** into the benzocycloheptanone diesters **7b–e,h** as 50/50 diastereoisomeric *cis–trans* mixture. Without purification, an acidic or basic hydrolysis/decarboxylation provided easily the substituted benzocyclohepten-7-ones **8b–e,h** in good overall yield (60–83%) from the starting *o*-xylenes **2b–e,h**.

### 2.2. 1,4 Symmetrically-disubstituted series

The commonly used reaction pathway is described in Scheme 4. These reactions led to a desymmetrisation of the molecules, except for the symmetric dibromo derivative **8e**. In this particular case, we used the pathway described previously for the non substituted<sup>11</sup> series. The *O*-silylation reaction of the ketone **8e** into the enol ether **9e** was better performed with  $\text{TfOSiMe}_3/\text{NEt}_3$ <sup>18</sup> than with  $\text{DBU}/\text{ClSiMe}_3$ .<sup>19</sup> The enol ether **9e** was then oxidised with a peracid<sup>20</sup> into the silyloxyketone **10e**. This intermediate was not purified and directly reductively aminated<sup>21</sup> into the amido-alcohol **11e** as a *cis/trans* mixture after *N*-protection, in poor yield (15%) in this case. Finally the isomeric mixture of alcohol derivatives **11e** was oxidised with Dess–Martin periodinane<sup>22</sup> into the amido-ketone **12e**. Acidic deprotection with dry HCl in  $\text{Et}_2\text{O}$ /dioxane gave the aminoketone **2e** as stable hydrochloride in 50% yield from **11e** and in 7% overall yield from the ketone **8e**.

### 2.3. 1,4 Asymmetrically-disubstituted series

In the case of the asymmetric ketones **8b–d,h**, two methods were employed. The reaction described in Scheme 3 led to the formation of a regioisomeric mixture of the intermediate enol ethers and silyloxy-ketones of type **9** and **10** respectively, in good yield from the ketones **8b–d,h** (85–96%). The reductive amination which follows, converted the silyloxy-ketones into the regioisomeric pair of *cis–trans* isomeric mixtures of hydroxy-amides **11b–d,h/11b'–d',h'**. These intermediates were obtained with variable yields, depending on the corresponding ketones. The monohalogeno-derivatives **11b,b'** and **11c,c'** were obtained in good yields (66% and 58%) starting from **8b** and **8c**, respectively. The reaction was less efficient for the bromo-phenyl and benzo-derivatives **11d,d'** and **11h,h'**, obtained in 29% and 37% yield from **8d** and **8h**, respectively. The separation of the four isomers by classic flash-chromatography or by semi-preparative HPLC was fastidious. The isomeric chloro- and bromo- hydroxy-amides **11b/11b'** and **11c/11c'** could hardly be purified in this way. In this series, the 1-chloro-derivative **11b'** was the only compound which was obtained pure as a *cis–trans* mixture. The four isomers of the bromo-derivatives **11c/11c'** *cis–trans* could be separated by semi-preparative HPLC and the **11h/11h'** *cis–trans* by flash-chromatography. The mixture of phenyl-bromo isomers **11d/11d'** was not resolved at this stage of the synthesis. The next step led to intermediates which were easier to isolate.



**Scheme 4.** Synthetic pathway for 1,4 mono- or disubstituted analogues. Reagents and conditions: (a)  $\text{Me}_3\text{SiOTf}$ ,  $\text{NEt}_3$ , toluene, 80–85 °C, 2 h; (b) *m*-CPBA,  $\text{CH}_2\text{Cl}_2$ , 0 °C, 2 h; (c)  $\text{Ti}(\text{O}i\text{Pr})_4$ , 2 M  $\text{NH}_3$  in MeOH, 6 h, then  $\text{NaBH}_4$ , 2 h; (d)  $\text{Boc}_2\text{O}$ ,  $\text{Na}_2\text{CO}_3$ , MeOH, 2 h; (e)  $\text{NH}_2\text{OH}$ , pyridine, 68% for **11c,c'**, 73% for **11d,d'**; (f)  $\text{H}_2$ , Raney-nickel, concd  $\text{NH}_4\text{OH}$  (85–98%, see text); (g) chromatographic separation (see text), then DMP,  $\text{CH}_2\text{Cl}_2$ ; (h) dry HCl 2 M in  $\text{Et}_2\text{O}$ , dioxane.

A second synthetic route was considered for the monobromo and bromo-phenyl series. It started from the oxime mixtures **13c,c'** and **13d,d'**, obtained from the corresponding silyloxy-ketones **10c,c'** and **10d,d'** by classical means. The reduction of the oxime function in amine by hydrogenolysis over Raney-nickel in the presence of ammonia and the following N-protection led to the same hydroxy-amide isomeric mixtures as described in our previous method. In this case, the diastereoisomeric monobromo-oximes **13c** and **13c'** could be easily separated by both fractional crystallisation and flash chromatography, in reasonable yield of 33–35% for each isomer, and their reduction occurred with good yield (85% for **13c**, 98% for **13c'**). By contrast, the phenyl-bromo-oxime mixture **13d,d'** was not resolved, but the global yield of the final hydroxy-amide mixture **11d/11d'** from the ketone **8d** was greatly improved (57% vs 29%).

Then, the oxidation of these various hydroxy-amides by the Dess–Martin periodinane as before, led to the corresponding keto-amide **12b,c,c',h,h'**. The isomeric mixture **11d/11d'** was oxidised in 55% yield into the ketone mixture which could then be resolved by semi-preparative HPLC into the pure isomers **12d** and **12d'**, in 17% and 38% yield, respectively.

All ketoamides **12c,d,h,b'-d',h'** were deprotected with dry HCl in  $\text{Et}_2\text{O}$ /dioxane into the corresponding crystallised amines of type **1**, in good yield (77–95%).

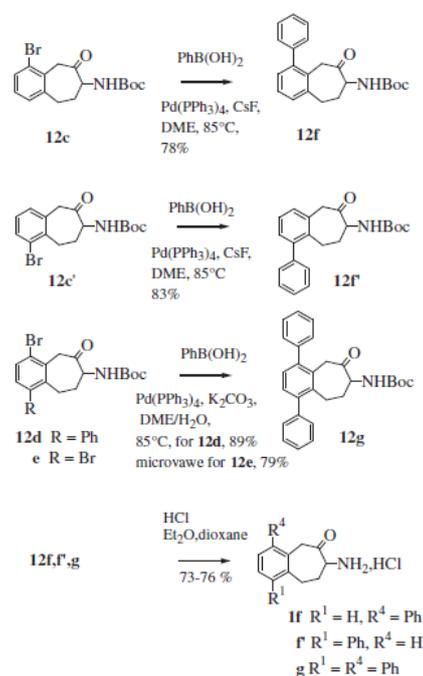
## 2.4. Synthesis of the mono- and di-phenyl keto-amines **1f,f,g**

These compounds were obtained directly from the bromo-keto-amides **12c,c',e** in two chemical steps according to Scheme 5. The Suzuki coupling reaction, performed with the monobromo derivatives **12c,c'** using the phenylboronic acid in dimethoxyethane (DME) as the solvent and cesium fluoride as the base gave the mono-phenyl compounds **12f,f'** in good yield (ca. 80%). The 1,4-di-phenyl derivative **12g** was obtained by two different routes with similar results. The coupling of the phenyl-bromoamide **12d** under the conditions described above, using aqueous potassium carbonate as the base and conventional reflux heating, gave **12g** in excellent yield. The second route started from the dibromo-amide **12e**; a double coupling under microwave irradiation gave **12g** in good yield.

After our standard acidic deprotection step, the final keto-amines **1f,f,g** were obtained in ca. 75% yield from **12c,c',e**, respectively.

## 2.5. Stereostructure and conformation of the hydroxy-amides **11b-d,e,h,b'-d',h'**

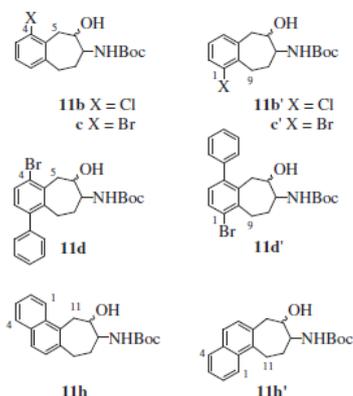
The structure of the different isomers of the hydroxy-amides mentioned in Scheme 6 was easily determined by NMR spectroscopy. The  $^1\text{H}$  NMR spectra of the *trans*-isomers were well resolved



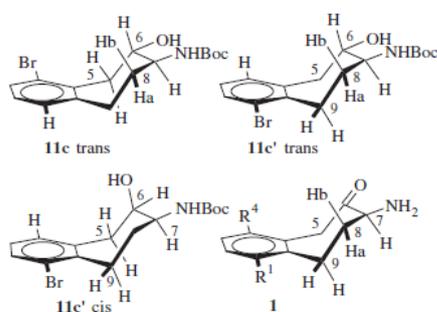
**Scheme 5.** Synthesis of the mono- and di-phenyl compounds **1ff,g**.

at room temperature, so that all H–H coupling could be precisely measured. No significant differences were observed with the various substituents: For all *trans*-isomers, the large value of the coupling between the five adjacent protons: one H-5 and H-6 (ca. 10 Hz), H-6 and H-7 (8.6–10.8 Hz), H-7 and Hb-8 (11.0–12.0 Hz), Hb-8 and one H-9 (11–12 Hz) corresponded to five axial protons in a chair conformation (see Fig. 1 for **11c** and **11c'**). These axial protons in the 5 and 9 positions were generally Hb-5 and Hb-9, but Ha-5 for **11b-trans** and **11c-trans**, and Ha-9 for **11b-trans** and **11c-trans**. In the naphtho series, for **11h-trans** these five protons were respectively Ha-11, H-10, H-9, Hb-8 and Ha-7, and for **11h'-trans** Ha-7, H-8, H-9, Hb-10 and Hb-11.

The *cis*-isomer **11c'** could be fully analysed at 330 K and the weaker values of  $J(6,7)$  and  $J(5b,6)$  (2.2 and 7.2 Hz respectively) were consistent for an equatorial H-6 in a chair conformation. A *trans*-diaxial relation was likely observed between H-7 and Hb-8 and Hb-8 and the axial Hb-9. This chair conformation is identical to the one determined previously for the benzo-unsubstituted analogue of type **11** ( $\text{R}^1 = \text{R}^4 = \text{H}$ ).



**Scheme 6.** Structures of the hydroxy-amides selected for conformational analysis by NMR.



**Figure 1.** Conformation of the hydroxy-amides **11c,c'** and ketone derivative **1**.

*Trans*- and *cis*-isomers were easily characterised by the  $\delta$ -values of the protons NH and H-6, H-7 (respectively H-10 and H-9 for **11h**, H-8 and H-9 for **11h'**): these protons appeared nearly at 5.05, 4.1 and 3.8 ppm for the *cis*-isomers, at 4.5, 3.35 and 3.7 ppm for the *trans*-isomers, respectively.

The structural determination of the regioisomeric hydroxy-amides **11b/11b'**, **11c/11c'**, **11d/11d'**, **11h/11h'** was deduced from the spatial deshielding of the *peri* H-atom caused by the aromatic halogen atom (Cl or Br in position 4 or 1) in the three halogeno-series, or by the naphthalene ring in the last series (Fig. 1 and Scheme 5). The equatorial Ha-5 proton in the 4-halogenated isomers **11b** and **11c**, or the equatorial one Ha-9 in the 1-halogenated regioisomers **11b'** and **11c'**, were spatially close to the halogen atom (see Fig. 1 for **11c** and **11c'**) and thus strongly deshielded with ca. 0.7–0.9 ppm. The same effect was observed for the Ha-11 proton in **11h** and **11h'** in the naphtho series. These equatorial protons appeared at 3.5–3.8 ppm while the corresponding axial ones Hb-5, Hb-9 or Hb-11 appeared at 2.8–3.1 ppm. The phenyl group at position 1 or 4 in **11d,d'** had only a weak effect on these protons.

## 2.6. Conformation of the amino-ketones of type **1**

In contrast to the hydroxy-amides of type **11**, whose regular chair conformation was supported by the numerous H–H *trans*-diaxial relations, the amino-ketones of type **1** seemed to adopt a slightly altered chair conformation. Indeed, the coupling constants between the axial H-7 and axial Hb-8 (respectively H-9 and Hb-8 for **1h**, H-9 and Hb-10 for **1h'**) remained large (11–12 Hz), but the coupling constants between Hb-8 and the axial H-9 (or the cor-

responding Hb-8 and Ha-7 protons for **1h**, Hb-10 and Hb-11 for **1h'**) were weaker (7.6–10.2 Hz), as for the corresponding *trans*-alcohol-amides of type **11**. The chair conformation was favoured by the hydrogen bond between the carbonyl and amine functions and its alteration was due to the steric hindrance between the 1-substituent and the equatorial H-9 and/or between the 4-substituent and the equatorial H-5 (or between the H-1 and equatorial H-11 protons for the naphthalenic derivatives **1h** and **1h'**) (see Fig. 1).

The conformation of the keto-amides of type **12** seemed to be more altered and was not studied here.

## 3. Aminopeptidase inhibition and discussion

All compounds, evaluated as racemic mixtures, behaved as competitive inhibitors of the panel aminopeptidases. The inhibition constants ( $K_i$ ) are reported in Table 1. We have previously reported that the aminobenzosuberone scaffold **1a** was stable in aqueous solutions at the physiological pH used for the kinetic studies.<sup>11</sup> We have also proposed that the ketone function most probably binds in its hydrate form to the zinc ion and to the catalytic glutamic acid residue in the APN/CD13 active site, thus mimicking the transition state that forms during peptide bond hydrolysis. This binding mode would be in line with the excellent ligand efficiency (0.63) of this very small lead structure.<sup>11</sup>

In the present work, we investigated various substituents in position 1 and/or 4 of the benzo aromatic ring A of our lead structure **1a**, in order to assess the effect of electrostatic and steric variations on binding and to determine the best match to the APN/CD13 active site. Since APN has a preference for hydrophobic substrates,<sup>1</sup> we focused on phenyl and halogen groups with large van der Waals radii. Halogen atoms are known to influence structure–activity relationships far beyond the mere steric aspects<sup>23</sup> and many impressive examples of the use of halogen substituents in hit-to-lead conversion have been reported in the recent years.<sup>24</sup>

Our data show that all derivatives of the lead structure **1a** display sub-micromolar inhibition constants towards the 'one zinc' APN/CD13, with an improved selectivity against the 'two zinc' family of enzymes represented here by the mammalian and bacterial LAPc and APAero, respectively. No inhibitory activity was observed up to an inhibitor concentration of 100  $\mu$ M towards the aminopeptidase activity of human LTA<sub>4</sub>H (data not shown).

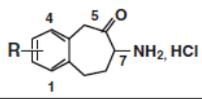
### 3.1. Extension of the aromatic system

We first investigated the extension of the fused ring system with two compounds, **1h** and **1h'**, bearing a fused benzo[3,4] or benzo[1,2] ring, respectively. The observed improvement in inhibitory potency is consistent with our previously published SAR studies in the 3-amino-2-tetralone series.<sup>25</sup> Both extensions improved binding affinity, with compounds **1h** ( $K_i = 0.1 \mu$ M) being 10 times more active than our lead structure **1a**, and **1h'** showing an even larger increase in potency, with a  $K_i$  value of 40 nM.

### 3.2. Substitution in position 1

To assess substitutions at position 1 of the benzo-ring A of our lead compound **1a**, we synthesised the phenyl, chloro and bromo analogues **1f,b',c'**. All modifications led to an increased inhibitory potency, with the following rank order Br > Cl > phenyl ( $K_i$  values of 20, 90 and 250 nM, respectively). This SAR revealed steric tolerance in this position as well as a clear preference for bromide. As a matter of fact, the bromo derivative **1c'** was the most potent inhibitor in this series with a 50-fold increase in potency when compared to **1a**, and was also 5–10 times more active than the phenyl- or chloro-substituted analogues.

**Table 1**  
Inhibition data of aminopeptidase activity<sup>a</sup>

Compounds				$K_i$ ( $\mu\text{M}$ )	
		$R^1$	$R^2$	APN <sup>b</sup> [EC 3.4.11.2] 'One zinc'	LAPc <sup>c</sup> [EC 3.4.11.1] APaero <sup>d</sup> [EC 3.4.11.10] 'Two zinc enzymes'
<b>1a</b>	H	H	1	>100	900
<b>1b'</b>	Cl	H	0.090	>100	>100
<b>1c'</b>	Br	H	0.020	>100	>100
<b>1f'</b>	Ph	H	0.250	>100	>100
<b>1c</b>	H	Br	0.040	>100	213
<b>1f</b>	H	Ph	0.007	>100	28
<b>1d</b>	Ph	Br	0.070	>100	>100
<b>1d'</b>	Br	Ph	0.00006 <sup>e</sup>	70	39
<b>1e</b>	Br	Br	0.006	>100	>100
<b>1g</b>	Ph	Ph	0.040	50	>100
<b>1h</b>	Benzo[3,4]		0.100	>100	>100
<b>1h'</b>	Benzo[1,2]		0.040	>100	>100

<sup>a</sup> All substances were evaluated as racemic mixtures.  $K_i$  ( $\mu\text{M}$ ) values were determined from Dixon plots at a substrate concentration set to the  $K_m$  value for the corresponding enzyme (see Section 5). Inactive compounds were tested up to their solubility limit under the assay conditions that is, 100  $\mu\text{M}$ .

<sup>b</sup> APN: porcine aminopeptidase-N (EC 3.4.11.2).

<sup>c</sup> LAPC: cytosolic leucine aminopeptidase from bovine kidney (EC 3.4.11.1).

<sup>d</sup> APaero: *Aeromonas proteolytica* aminopeptidase (EC 3.4.11.10).

<sup>e</sup> Dixon plot for APN inhibition with **1d'** is reported in Figure 2.

### 3.3. Substitution in position 4

We also selected bromo and phenyl groups for investigating potential substitutions in position 4. Both variations had a pronounced positive effect on binding affinity. In sharp contrast to substitutions in position 1, however, the phenyl derivative **1f** was, with a  $K_i$  value of 7 nM, a more potent inhibitor than the bromo analogue **1c**, which was, comparatively, five times less active ( $K_i = 40$  nM) on APN. Among all monosubstituted analogues in position 1 or 4, **1f** was clearly the best inhibitor.

### 3.4. Disubstitution in positions 1 and 4

Four compounds were designed and synthesised in this series, combining phenyl and/or bromo substituents in positions 1 and 4. The addition of a phenyl moiety in position 1 to the monosubstituted bromo-derivative in position 4 (**1c**,  $K_i = 40$  nM) led to the asymmetrical disubstituted analogue **1d** ( $K_i = 70$  nM) which did not show any improvement in the  $K_i$  value. The symmetric dibromo-derivative **1e** ( $K_i = 6$  nM), however, was about 5 times more active than the monosubstituted analogues **1c'** ( $K_i = 20$  nM) and **1c** ( $K_i = 40$  nM). The other symmetrically disubstituted diphenyl derivative **1g** ( $K_i = 40$  nM) was slightly weaker than the monosubstituted 4-phenyl derivative (**1f**,  $K_i = 7$  nM). Nevertheless, **1g** was more potent than the monosubstituted 1-phenyl derivative **1f'** ( $K_i = 250$  nM).

Based on these results, we expected the asymmetrical disubstitution combining a bromo group in position 1 and a phenyl ring in position 4 to be the most interesting combination, for the corresponding monosubstitutions seemed optimal in both cases. Very gratifyingly, our expectation was fully confirmed by the outstanding inhibitory potency of the disubstituted derivative **1d'** which, with a  $K_i$  value of 60 pM, turned out to be 100 to 1000 times more potent than the corresponding monosubstituted analogues. This new structure is also 20,000 times more active than our starting lead compound **1a**. This spectacular enhancement in potency was achieved through the additive effect obtained by combining the substituents in positions 1 and 4 that fit the APN active site best. Kinetic data for APN inhibition by this particular compound are

reported in Fig. 2 as a Dixon plot<sup>34</sup> which clearly showed that compound **1d'** remains a competitive, reversible inhibitor although its potency is close to the range of tight binding inhibition. This is perfectly in line with the core structure of our compound series, which are cyclic substrate analogues retaining the metal chelating groups.

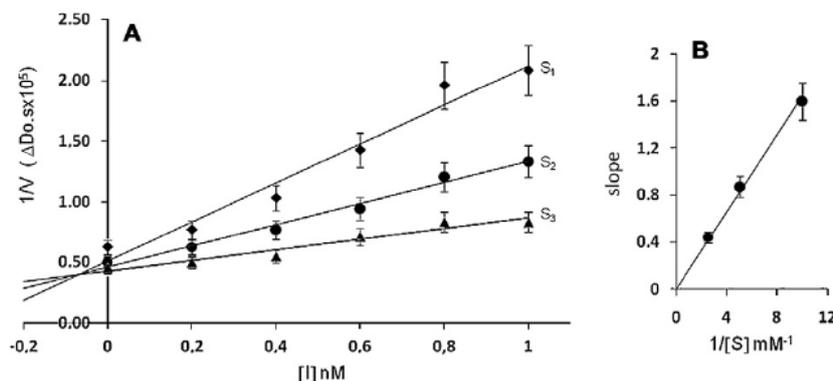
Not only is this novel APN/CD13 inhibitor by far more potent than any other compound investigated in this work, it is also undoubtedly among the most potent and selective non peptidic inhibitors of mammalian APN/CD13 known to date.

## 4. Conclusion

Metallopeptidases constitute a large family of proteolytic enzymes using a transition metal ion at their catalytic center. Small-molecule metallopeptidase inhibitors are generally designed to bind directly to the active site metal,<sup>10</sup> thus achieving a high ligand efficiency, often at the expense, however, of selectivity. The development of highly specific metallopeptidase inhibitors is a technically challenging, yet medically important scientific endeavour, in view of the prominent role played by metallopeptidases in many pathologies. We recently reported the discovery of aminobenzosuberone **1a** as a novel war head showing promise for the selective inhibition of the 'one zinc' mammalian aminopeptidase APN/CD13.<sup>11</sup>

In the present study, a series of highly potent analogues of aminobenzosuberone **1a** is reported. Our data demonstrate that very large improvements in potency can be achieved without compromising selectivity. Moreover, the novel APN inhibitors reported here remain well within the boundaries of Lipinski's rule-of-five which delimits 'drug-like' chemical space.<sup>27</sup> With a molecular weight of only 329 Da, there is still ample room for fine tuning the pharmacological properties of our highly selective, picomolar inhibitor **1d'**. Alternatively, it may be possible to tune the selectivity of this compound series towards related aminopeptidases of pharmaceutical interest, such as *Plasmodium falciparum* aminopeptidase N.<sup>28</sup>

We strongly believe that **1d'**, as well as other potent compounds reported here, will be highly valuable chemical probes for investigating APN/CD13 and for delineating its physiological and pathological roles that require catalytic activity.



**Figure 2.** Kinetic data for APN inhibition by compound **1d**. (A) Dixon plot: the effect of the inhibitor on the enzyme rate is determined at 3 substrate concentrations ( $S_1 = K_m/2$ ,  $S_2 = K_m$  and  $S_3 = 2 K_m$ ) over a range of inhibitor concentrations  $[I]$ , from 0.2 to 1 nM. The concentration of APN (Specific activity, 28 Units per mg) was 0.1 mUnit per assay (12 pM). Data for each substrate concentration fall on a straight line that intersect on  $[I] = K_i = 0.06$  nM. With an average experimental error of 10% ( $n = 3$ ). (B) The replot of the slopes of the Dixon plot is a straight line through the origin, indicating a pure competitive inhibition.<sup>34</sup>

## 5. Experimental part

### 5.1. General

Flash chromatography: silica gel (Merck 60, 230–400 mesh). TLC: Al-roll silica gel (Merck 60,  $F_{254}$ ). Mp: Kofler hot bench, corrected. IR spectra ( $\nu$  in  $cm^{-1}$ ): Nicolet 405 FT-IR.  $^1H$  and  $^{13}C$  NMR (400 MHz and 100.6 MHz resp.) spectra: Bruker Avance 400, tetramethylsilane (TMS), or natrium ( $D_4$ )-trimethylsilylpropionate ( $D_4$ -TSP) in  $D_2O$  ( $^1H$  NMR) and  $CDCl_3$ , or MeOD- $D_4$  [ $\delta(CDCl_3) = 77.0$ ,  $\delta(CD_3OD) = 35.0$  with respect to TMS] ( $^{13}C$  NMR) as internal references;  $\delta$  in ppm and  $J$  in Hertz. High resolution MS were measured on a Bruker MicroTof spectrometer in Institut de Chimie, UMR 7177 CNRS, ULP, Strasbourg, France, or Agilent Technologies 6510 (QToF) spectrometer in ENSCMu, Université de Haute Alsace, Mulhouse, France or Waters Micromass Q-ToF Ultima API, Basilea Pharmaceuticals, Basel. Microanalyses were carried out by the Service Central de Microanalyses du CNRS, F-69390 Vernaison or by the Service de Microanalyse, UMR 7565 CNRS Université Henri Poincaré F-54506 Vandoeuvre-les-Nancy.

### 5.2. Reagents and solvents

5% Pd/C and Raney-nickel were obtained from Fluka, other reactants were purchased from usual provider. Dess–Martin periodinane (DMP) was prepared according to,<sup>22</sup> or purchased in  $CH_2Cl_2$  solution and NMR-titrated by oxidation of benzyl alcohol. Usual solvents were freshly distilled, dry EtOH and MeOH distilled over Mg/MgI<sub>2</sub>, dry THF over Na and benzophenone, dry Et<sub>2</sub>O was distilled and stored over Na,  $CH_2Cl_2$  was distilled over  $P_2O_5$  and kept over  $Na_2CO_3$ .  $NEt_3$  was distilled before use.

## 6. Syntheses of the xylenes **2d,e**, **3b** and **4**

### 6.1. 4-Bromo-2,3-dimethylaniline hydrobromide (**3b**)

To a solution of **3a** (2.5 mL, 20.1 mmol) in dry  $CH_2Cl_2$  (50 mL) was added  $NBu_4Br_3$  (10.2 g, 20.1 mmol, 1 equiv) at 0 °C under Argon, then the solution was stirred at 0 °C for 15 min. The precipitate of **3b**-HBr (4.33 g, 75%) was isolated by filtration, washed with Et<sub>2</sub>O and dried under vacuum.

Compound **3b**-HBr: colorless crystals, mp 258–60 °C. IR (KBr): 2921, 2577, 1531, 1510, 1456, 1180, 1001, 905, 824, 801, 543  $cm^{-1}$ .  $^1H$  NMR ( $CD_3OD$ , 400 MHz): 2.36 (s, 3H, Me-2); 2.48 (s, 3H, Me-3); 7.13 (d, 1H, H-6); 7.58 (d, 1H, H-5);  $J(5,6) = 8.6$  Hz.

$^{13}C$  NMR ( $CD_3OD$ , 100 MHz): 15.6 (Me-2); 20.7 (Me-3); 123.6 (C(6)); 127.3 (C(4)); 130.5 (C(2)); 132.7 (C(5)) 134.4 (C(3)); 140.5 (C(1)). Free base (by stirring with aqueous  $Na_2CO_3$ ),  $^1H$  NMR ( $CDCl_3$ ): same values as in lit.<sup>15</sup>

### 6.2. 4-Bromo-2,3-dimethylbenzediazonium tetrafluoroborate (**4**)

To a solution of **3b**-HBr (4.12 g, 14.7 mmol) in  $CH_3CN$  (10 mL) were added 50% aqueous  $HBF_4$  (5.6 mL, 44 mmol, 3 equiv) and then with stirring at –20 °C dropwise  $tBuONO$  (2.2 mL, 16.1 mmol, 1.1 equiv). The solution was stirred at –20 °C for further 45 min, Et<sub>2</sub>O (20 mL) was then added and the precipitate of **4** (3.46 g, 78%) was isolated by filtration and washing with Et<sub>2</sub>O (10 mL).

Compound **4**: colorless crystals, mp 162–171 °C. (KBr): 3568, 3048, 2256, 1548, 1430, 1386, 1300, 1196, 1083, 1029, 896, 823  $cm^{-1}$ .  $^1H$  NMR ( $CD_3OD$ , 400 MHz): 8.30 (d, 1H,  $J(5,6) = 9.0$  Hz, H-6); 8.07 (d, 1H,  $J(5,6) = 9.0$  Hz, H-5); 2.76 (s, 3H, Me-2); 2.50 (s, 3H, Me-3).  $^{13}C$  NMR ( $CD_3OD$ , 100 MHz): 144.9 (C(1)); 143.5 (C(3)); 141.8 (C(2)); 138.0 (C(4)); 134.7 (C(5)); 131.3 (C(6)); 20.3 ( $CH_3$ -3); 18.3 ( $CH_3$ -2). Anal. Calcd for  $C_9H_8BBF_4N_2 \cdot 0.5H_2O$  (307.88): C, 31.21; H, 2.95; N, 9.10. Found: C, 31.0; H, 2.7; N, 8.8.

### 6.3. 1,4-Dibromo-2,3-dimethylbenzene (**2e**)

A solution of the free base **3b** [8.69 g, 43.5 mmol, obtained by stirring a suspension of **3b** in Et<sub>2</sub>O (20 mL) with  $Na_2CO_3$  (5 g, 50 mmol) and  $H_2O$  (5 mL) and then filtration and evaporation of the solvent] in MeCN (100 mL) was added under Ar to a solution of  $CuBr_2$  (9.71 g, 43.5 mmol, 1 equiv) and  $tBuONO$  (4.93 g, 5.68 mL, 47.8 mmol, 1.1 equiv) in MeCN (200 mL). The mixture was stirred at rt for 16 h, then at 82 °C for 4 h. The mixture was left at rt, diluted with AcOEt (200 mL), washed with brine (2 × 100 mL), dried ( $MgSO_4$ ) and the solvent evaporated to give **2e** (9.79 g, 85%). Same NMR data as in lit.<sup>15</sup>

### 6.4. 4-Bromo-2,3-dimethyl-biphenyle (**2d**)

A solution of **4** (10.9 g, 36.3 mmol) and phenylboronic acid (5.0 g, 41.5 mmol, 1.15 equiv) in MeOH (300 mL) was refluxed under Ar with  $Pd(OAc)_2$  (0.82–0.6 g, 0.1–0.07 equiv) at 65 °C for 2 h. The solution was left at rt, diluted with AcOEt (500 mL), washed with  $H_2O$ , dried ( $SO_4Mg$ ) and evaporated. The residue was purified by flash chromatography (cyclohexane), to give **2d** in two crops after crystallisation in  $iPrOH$  (3.5–5.1 g, 34–54%).

Compound **2d**: colorless crystals, mp 56–57 °C (MeOH). IR (KBr): 2922, 1443, 1006, 823, 765, 703  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 400 MHz): 7.44 (d, 1H, H-5); 7.39 (tm, 2H, Har-m); 7.34 (tm, 1H, Har-p); 7.24 (dm, 2H, Har-o); 6.93 (d, 1H, H-6); 2.46 (s, 3H, Me-3); 2.21 (s, 3H, Me-2);  $J(5,6) = 8.3$ ,  $J(o,m) = 7.5$ ,  $J(m,p) = 7.3$  Hz.  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 100 MHz): 141.8, 141.7 (C(2),C(3)); 136.5, 136.0 (C(1),C(1')); 129.5 (C(5)); 129.3 (C(2'),C(6')); 128.6 (C(6)); 128.1 (C(3'),C(5')); 126.9 (C(4')); 124.6 (C(4)); 20.2 (Me-3); 18.6 (Me-2). Anal. Calcd for  $\text{C}_{14}\text{H}_{13}\text{Br}$  (261.16): C, 64.39; H, 5.02; Br, 30.60. Found: C, 64.5; H, 4.9; Br, 30.3.

## 7. $\alpha,\alpha'$ -Dibromoxylene derivatives **5b–e,h**

*General procedure (a)*: A solution of 1,2-dimethylaryle **2b–e,h** (10 mmol) and finely pulverised N-bromosuccinimide (NBS 3.68 g, 21 mmol, 2.1 equiv) in  $\text{CCl}_4$  (40–80 mL) was irradiated with HPK125 mercury lamp for 1–2 h with good stirring (tlc or  $^1\text{H}$  NMR monitoring). The reaction mixture was diluted with  $\text{CH}_2\text{Cl}_2$ , washed with  $\text{H}_2\text{O}$  or 2 N aqueous  $\text{NH}_4\text{Cl}$  solution and dried over  $\text{MgSO}_4$ . The solvent was evaporated to give quantitatively **5b–e,h** which was used without further purification.

### 7.1. 1-Chloro-2,3-bis(bromomethyl)benzene (**5b**)

*General procedure (a)* with **2b** (2.0 g, 10.8 mmol) and NBS (4.04 g, 22.7 mmol, 2.1 equiv) in  $\text{CCl}_4$  (70 mL) to give **5b**<sup>29</sup> (3.7 g, quant.).

Compound **5b**: yellowish oil.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 400 MHz): 7.38 (dd, 1H, H-6); 7.28 (dd, 1H, H-4); 7.24 (t, 1H, H-5); 4.81 (s, 2H,  $\text{CH}_2\text{Br}$ -2); 4.62 (s, 2H,  $\text{CH}_2\text{Br}$ -3);  $J(4,5) = 7.6$ ,  $J(4,6) = 1.6$ ,  $J(5,6) = 7.8$  Hz.

### 7.2. 1-Bromo-2,3-bis(bromomethyl)benzene (**5c**)

*General procedure (a)* with **2c** (5 g, 27 mmol) and NBS (10.1 g, 56.7 mmol, 2.1 equiv) in  $\text{CCl}_4$  (200 mL) to give **5c** (9.36 g, quant.).

Compound **5c**: yellowish oil.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 400 MHz): 7.57 (dd, 1H, H-6); 7.32 (dd, 1H, H-4); 7.15 (t, 1H, H-5); 4.84 (s, 2H,  $\text{CH}_2\text{Br}$ -2); 4.63 (s, 2H,  $\text{CH}_2\text{Br}$ -3);  $J(4,5) = 7.5$ ,  $J(4,6) = 1.5$ ,  $J(5,6) = 8.1$  Hz. Same data as in lit.<sup>30</sup>

### 7.3. 4-Bromo-2,3-bis(bromomethyl)-biphenyle (**5d**)

*General procedure (a)* with **2d** (5.0 g, 19.1 mmol) and NBS (7.2 g, 40 mmol, 2.1 equiv) in  $\text{CCl}_4$  (200 mL) to give **5d** (9.5 g, quant.).

Compound **5d**: colorless crystals, mp 84–86 °C (cyclohexane). IR (KBr): 556, 613, 659, 704, 759, 827, 1188, 1203, 1222, 1434, 1446, 3032  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 400 MHz): 7.61 (d, 1H,  $J(5,6) = 8.3$  Hz, H-5); 7.45 (m, 5Har); 7.09 (d, 1H,  $J(5,6) = 8.3$  Hz, H-6); 4.97 (s, 2H,  $\text{CH}_2\text{Br}$ -3); 4.55 (s, 2H,  $\text{CH}_2\text{Br}$ -2).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 100 MHz): 28.6 (2- $\text{CH}_2\text{Br}$ ); 29.9 (3- $\text{CH}_2\text{Br}$ ); 125.6 (C(4)); 128.0, 128.4, 128.7 (3 CHar); 132.0 (C(6)); 133.3 (C(5)); 136.2, 136.6 (C(1),C(1')); 139.2 (C(2)); 143.4 (C(3)). Anal. Calcd for  $\text{C}_{14}\text{H}_{11}\text{Br}_3$  (418.96): C, 40.14; H, 2.65. Found: C, 39.9; H, 2.4.

### 7.4. 1,4-Dibromo-2,3-bis(bromomethyl)benzene (**5e**)

*General procedure (a)* with **2e** (1.22 g, 4.62 mmol) and NBS (1.75 g, 9.71 mmol, 2.1 equiv) in  $\text{CCl}_4$  (70 mL) to give **5e** as orange crystals (1.83 g, 94%). Same NMR data as in lit.<sup>31</sup>

### 7.5. 1,2-Bis-bromomethyl-naphthalene (**5h**)

*General procedure (a)* with **2h** (1.0 g, 6.4 mmol) and NBS (2.39 g, 13.4 mmol, 2.1 equiv) in  $\text{CCl}_4$  (50 mL) to give **5h** (2.0 g, quant.).

Compound **5h**: orange crystals, mp 149–150 °C (lit.<sup>32</sup> 148.5–149.5 °C).  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 400 MHz): 8.16 (dm, 1H, H-8); 7.87 (ddd, 1H, H-5); 7.84 (d, 1H, H-4); 7.65 (td, 1H, H-7); 7.55 (td, 1H,

H-6); 7.45 (d, 1H, H-3); 5.12 (s, 2H,  $\text{CH}_2\text{Br}$ -1); 4.78 (s, 2H,  $\text{CH}_2\text{Br}$ -2);  $J(3,4) = 8.4$ ,  $J(5,6) = 8.0$ ,  $J(5,7) = 1.4$ ,  $J(5,8) = 0.6$ ,  $J(6,7) = 6.8$ ,  $J(6,8) = 1.0$ ,  $J(7,8) = 8.6$  Hz. Data in agreement with those of lit.<sup>33</sup>

## 8. Preparation of the ketone diesters **7b–e,h** by reaction with acetone-dicarboxylate **6** and decarboxylation into ketones **8b–e,h**

*General procedure (b), reaction with methyl acetonedicarboxylate (6)*: A solution of 1,2-bis-bromomethylaryle **5a–e,h** (10 mmol) and **6** (1.7 ml, 12 mmol, 1.2 equiv) in  $\text{CH}_2\text{Cl}_2$  (25–40 mL) was added dropwise to a solution of  $\text{NBU}_4\text{Br}$  (2.0 g, 6 mmol, 0.6 equiv) in 1 N aqueous  $\text{NaHCO}_3$  (50–80 mL) and  $\text{CH}_2\text{Cl}_2$  (25–40 mL). The biphasic mixture was vigorously stirred at 40 °C under Argon for 6 h to overnight. The layers were separated, the aqueous layer was extracted with  $\text{CH}_2\text{Cl}_2$  (2  $\times$  50 mL) and the combined organic solutions were evaporated. The residue was dissolved in  $\text{AcOEt}$  and washed with brine (3  $\times$  50 mL), dried over  $\text{MgSO}_4$  and evaporated to give a yellowish resin which was used without further purification (quantitative).

*General procedure (c), decarboxylation in acidic medium*: A vigorously stirred biphasic solution of **7b–e,h** (10 mmol) in 3 M aqueous  $\text{H}_2\text{SO}_4$  (50 mL) and MeCN (15 mL) was refluxed at 90 °C for 16 h. The mixture was diluted with  $\text{AcOEt}$  (100 mL), neutralised with 2 M aqueous NaOH. The organic layer was separated, washed with  $\text{H}_2\text{O}$  or brine, the aqueous layer extracted with  $\text{AcOEt}$ , the combined organic phases were dried over  $\text{MgSO}_4$  and evaporated to give the crude ketone **8**.

*General procedure (d), decarboxylation in basic medium*: A vigorously stirred biphasic solution of 1 N aqueous NaOH (70 ml) and **7b–e,h** (10 mmol) in MeCN solution (20 ml) was refluxed at 90 °C for 2 h. The mixture was left at rt, neutralised with concd HCl and extracted with  $\text{AcOEt}$  (2  $\times$  50 mL). The combined organic phase was washed with brine (4  $\times$  10 mL), dried over  $\text{MgSO}_4$  and evaporated to give the ketone **8** which was purified by flash chromatography.

### 8.1. Dimethyl 1-chloro-7-oxo-5,6,8,9-tetrahydro-benzocyclohepten-6,8-dicarboxylate (**7b**) and 1-chloro-5,6,8,9-tetrahydro-benzocyclohepten-7-one (**8b**)

*General procedure (b)* with  $\text{NBU}_4\text{Br}$  (5.1 g, 16.0 mmol, 0.6 equiv) **5b** (7.6 g, 25.5 mmol), **6** (5.3 g, 30.5 mmol, 1.2 equiv) in 1 M aqueous  $\text{NaHCO}_3$  (100 mL) and  $\text{CH}_2\text{Cl}_2$  (80 mL) for 16 h to give **7b** (7.5 g, 96%) as 50/50 isomeric mixture.

*General procedure (c)* with crude **7b** (7.5 g, 24.2 mmol) in MeCN (20 mL) and aqueous 3 M  $\text{H}_2\text{SO}_4$  (100 mL) to give the crude **8b** (4.5 g, 83% from **3b**).

Compound **7b**: yellowish resin.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 300 MHz): 3.0–4.0 (m, 12H); 7.0–7.4 (m, 3H). HR-MS (ESI-Q-Tof) calcd for  $\text{C}_{15}\text{H}_{15}\text{ClLiO}_5$  [ $\text{M}+\text{Li}$ ] $^+$ : 317.0763; found: 317.0752.

Compound **8b**: yellowish resin. IR (KBr): 2956, 2945, 1699, 1450, 1349, 1187, 880, 794, 786  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 300 MHz): 7.32 (m, 1H, H-2); 7.12 (m, 2H, H-4, H-3); 3.16 (m, 2H,  $\text{CH}_2(9)$ ); 2.96 (m, 2H,  $\text{CH}_2(5)$ ); 2.62 (m, 4H,  $\text{CH}_2(6)$ ,  $\text{CH}_2(8)$ ).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 75 MHz): 210.4 (CO(7)); 142.9, 137.9 (C(9a),C(4a)), 133.7 (C(1)); 128.2, 127.7, 127.6 (C(4),C(3),C(2)); 44.4, 43.2 (C(6),C(8)); 30.9 (C(5)); 25.2 (C(9)). HR-MS (ESI-Q-Tof) calcd for  $\text{C}_{11}\text{H}_{11}\text{ClLiO}$  [ $\text{M}+\text{Li}$ ] $^+$ : 201.0653; found: 201.0636.

### 8.2. Dimethyl 1-bromo-7-oxo-5,6,8,9-tetrahydro-benzocyclohepten-6,8-dicarboxylate (**7c**) and 1-bromo-5,6,8,9-tetrahydro-benzocyclohepten-7-one (**8c**)

*General procedure (b)* with  $\text{NBU}_4\text{Br}$  (5.2 g, 16.2 mmol, 0.6 equiv), **5c** (9.36 g, 27.3 mmol), **6** (5.6 g, 32.4 mmol, 1.2 equiv) in 5%

aqueous NaHCO<sub>3</sub> (120 mL) and CH<sub>2</sub>Cl<sub>2</sub> (100 mL) for 6 h to give **7c** (10.9 g, quant.).

**General procedure (c)** with **7c** (10.9 g, 27 mmol) in MeCN (50 mL) and aqueous 3 M H<sub>2</sub>SO<sub>4</sub> (100 mL). The crude product was purified by flash chromatography (cyclohexane/AcOEt 9:1) to give **8c** (4.0 g, 62% from **3c**).

Compound **7c**: yellowish resin. IR (KBr): 2953, 1744, 1719, 1653, 1437, 1333, 1305, 1226, 1161, 783 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 3.2–4.2 (m, 12H); 7.0–7.2 (m, 2H); 7.4–7.6 (m, 1H). HR-MS (ESI<sup>+</sup>): calcd for C<sub>15</sub>H<sub>16</sub>BrO<sub>5</sub> [M+H]<sup>+</sup>: 355.0176; found: 355.0174; calcd for C<sub>15</sub>H<sub>15</sub>BrNaO<sub>5</sub> [M+Na]<sup>+</sup>: 376.9995; found: 376.9993.

Compound **8c**: colorless crystals, mp 32–34 °C. IR (KBr): 2952, 2907, 1702, 1561, 1448, 1344, 1184, 873, 781 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 7.51 (d, 1H, H-2); 7.16 (d, 1H, H-4); 7.05 (t, 1H, H-3); 3.19 (m, 2H, CH<sub>2</sub>(9)); 2.98 (m, 2H, CH<sub>2</sub>(5)); 2.62 (m, 4H, CH<sub>2</sub>(6), CH<sub>2</sub>(8)); J(1,2) = 8.1, J(2,3) = 7.4 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 210.2 (C(7)); 142.8 (C(9a)), 139.6 (C(4a)); 131.6 (C(2)); 128.4, 128.2 (C(4),C(3)); 124.5 (C(1)); 44.4, 43.1 (C(6),C(8)); 31.1 (C(5)); 28.5 (C(9)). HR-MS (ESI-Q-ToF) calcd for C<sub>11</sub>H<sub>13</sub>BrNaO [M+Na]<sup>+</sup>: 260.9885 and 262.9866; found: 260.9879 and 262.9864.

### 8.3. Dimethyl 1-bromo-4-phenyl-7-oxo-5,6,8,9-tetrahydro-benzocyclohepten-6,8-dicarboxylate (**7d**) and 1-bromo-4-phenyl-5,6,8,9-tetrahydrobenzocycloheptene-7-one (**8d**)

**General procedure (b)** with NBu<sub>4</sub>Br (1.74 g, 5.30 mmol, 0.6 equiv) **5d** (3.7 g, 8.83 mmol), **6** (2.0 mL 13.2 mmol, 1.5 equiv) in 5% aqueous NaHCO<sub>3</sub> (75 mL) and CH<sub>2</sub>Cl<sub>2</sub> (50 mL) for 16 h to give **7d** (3.6 g, quant.) as 50:50 *cis/trans* isomeric mixture.

**General procedure (c)** with **7d** (16.9 g, 39 mmol) in MeCN (115 mL) and H<sub>2</sub>SO<sub>4</sub> 3.8 M (186 mL) at 100 °C for 3 days. Crude ketone was purified by flash chromatography (cyclohexane/AcOEt 9:1) then crystallised from MeOH to give pure **8d** (7.6 g, 64% from **3d**).

**General procedure (d)** with **7d** (3.6 g, 8.8 mmol) in MeCN (20 mL) and aqueous 1 M NaOH (60 mL) for 16 h. The crude product was purified by flash chromatography (cyclohexane/AcOEt 9:1) to give **8d** (1.20 g, 43% from **3d**).

Compound **7d**: colorless crystals, mp 146–152 °C (AcOEt). IR (KBr): 706, 765, 816, 1164, 1231, 1280, 1291, 1305, 1440, 1452, 1641, 1735, 2947 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 7.6–7.4 (m, 4Har); 7.25–7.15 (m, 2Har); 7.2–7.1 (m, 1Har); 3.2–4.0 (m, 12H). HR-MS (ESI<sup>+</sup>): calcd for C<sub>21</sub>H<sub>19</sub>BrNaO<sub>5</sub> [M+Na]<sup>+</sup>: 453.0308 and 455.0287; found: 453.0303 and 455.0287.

Compound **8d**: orange crystals; mp 117–119 °C (iPr<sub>2</sub>O). IR (KBr): 710, 776, 819, 1175, 1451, 1693, 2943 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 7.54 (d, 1H, J(2,3) = 8.4 Hz, H-2); 7.45–7.36 (m, 3H, 2Har-m, Har-p); 7.23 (m, 2Har-o); 7.03 (d, 1H, J(2,3) = 8.4 Hz, H-3); 3.30 (m, 2H, 2 CH<sub>2</sub>(9)); 2.93 (m, 2H, CH<sub>2</sub>(6)); 2.65 (m, 2H, CH<sub>2</sub>(8)); 2.54 (m, 2H, CH<sub>2</sub>(6)). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 210.2 (C(7)); 141.4, 140.9, 140.3, 140.1 ((C(9a), C(4a), C(4), Car-s); 130.7 (C(2)); 129.8 (C(3)); 129.0 (Car-o); 128.4 (Car-m); 127.4 (Car-p); 123.5 (C(1)); 44.3 (C(6)); 43.1 (C(8)); 28.8 (C(9)); 26.4 (C(5)). HR-MS (ESI-Q-ToF) calcd for C<sub>17</sub>H<sub>15</sub>BrLiO [M+Li]<sup>+</sup>: 321.0461 and 323.0446; found: 321.0474 and 323.0461.

### 8.4. Dimethyl 1,4-dibromo-7-oxo-5,6,8,9-tetrahydro-benzocyclohepten-6,8-dicarboxylate (**7e**) and 1,4-dibromo-5,6,8,9-tetrahydrobenzocycloheptene-7-one (**8e**)

**General procedure (b)** with NBu<sub>4</sub>Br (0.84 g, 2.56 mmol, 0.6 equiv), **5e** (1.8 g, 4.27 mmol), **6** (0.92 mL, 6.40 mmol, 1.5 equiv) in 5% aqueous NaHCO<sub>3</sub> (36 mL) and CH<sub>2</sub>Cl<sub>2</sub> (25 mL) for 16 h to give **7e** (1.84 g, quant.) as 50:50 *cis/trans* isomeric mixture.

**General procedure (d)** with **7e** (1.85 g, 4.26 mmol) in MeCN (10 mL) and aqueous 1 M NaOH (30 mL) for 2 h. The crude product was purified by flash chromatography (cyclohexane/AcOEt 9:1) to give **8e** (1.15 g, 85% from **3e**).

Compound **7e**: yellowish oil. IR (KBr): 804, 1159, 1232, 1443, 1648, 1717, 3412 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 3.40–4.20 (m, 12H); 7.26–7.40 (m, 2H). HR-MS (ESI<sup>+</sup>IsoTof) calcd for C<sub>15</sub>H<sub>14</sub>Br<sub>2</sub>NaO<sub>5</sub> [M+Na]<sup>+</sup>: 454.9100, 456.9081 and 458.9063; found: 454.9107, 456.9082 and 458.9067.

Compound **8e**: orange crystals, mp 94 °C. IR (KBr): 3054, 2951, 2878, 1697, 1445, 1213, 1183, 1103, 878, 817, 527 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 7.36 (s, 2H, H-2, H-3); 3.28 (m, 4H, CH<sub>2</sub>(5),CH<sub>2</sub>(9)); 2.62 (m, 4H, CH<sub>2</sub>(6), CH<sub>2</sub>(8)). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 209.3 (C(7)); 141.7 (C(4a),C(9a)); 132.3 (C(2),C(3)); 123.3 (C(1),C(4)); 43.0 (C(6),C(8)); 29.4 (C(5),C(9)). HR-MS (ESI-Q-ToF) calcd for C<sub>11</sub>H<sub>14</sub>Br<sub>2</sub>NO [M+NH<sub>4</sub>]<sup>+</sup>: 333.94376, found: 333.9432.

### 8.5. Dimethyl 9-oxo-7,8,10,11-tetrahydro-naphthocycloheptene-8,10-dicarboxylate (**7h**) and 7,8,10,11-tetrahydro-cyclohepta[a]naphthalen-9-one (**8h**)

**General procedure (b)** with NBu<sub>4</sub>Br (1.32 g, 4 mmol, 0.6 equiv), **5h** (2.0 g, 6.36 mmol), **6** (1.22 g, 1.01 mL, 7 mmol, 1.1 equiv), in 1 N NaHCO<sub>3</sub> solution (30 mL) and CH<sub>2</sub>Cl<sub>2</sub> (50 mL) for 16 h to give **7h** (2.1 g, quant.) as 50:50 *cis/trans* isomeric mixture.

**General procedure (c)** with **7h** (2.0 g, 6.09 mmol), in MeCN (10 mL) and aqueous 3 M H<sub>2</sub>SO<sub>4</sub> (30 mL) to give **8h** (0.9 g, 75%).

Compound **7h**: yellowish resin. IR (KBr): 2953, 1743, 1711, 1648, 1436, 1354, 1293, 1231, 1211, 1166, 817, 748 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 3.2–4.1 (m, 12H); 7.3–7.4 (m, 1H); 7.44–7.55 (m, 2H); 7.65–7.85 (m, 2H); 8.0–8.2 (m, 2H). HR-MS (ESI-Q-ToF) calcd for C<sub>19</sub>H<sub>18</sub>NaO<sub>5</sub> [M+Na]<sup>+</sup>: 349.1046; found: 349.1057; C<sub>19</sub>H<sub>18</sub>LiO<sub>5</sub> [M+Li]<sup>+</sup>: 333.1314; found: 333.1304.

Compound **8h**: yellowish resin. IR (KBr): 2950, 1694, 1597, 1512, 1436, 1384, 1323, 1189, 827, 755 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 8.11 (d, 1H, H-1); 7.86 (dd, 1H, H-4); 7.73 (d, 1H, H-5); 7.54 (ddd, 1H, H-2); 7.47 (ddd, 1H, H-3); 7.36 (d, 1H, H-6); 3.41 (m, 2H, CH<sub>2</sub>(11)); 3.14 (m, 2H, CH<sub>2</sub>(7)); 2.69 (m, 4H, CH<sub>2</sub>(10), CH<sub>2</sub>(8)); J(1,2) = 8.6, J(1,3) = 1.0, J(2,3) = 6.8, J(2,4) = 1.4, J(3,4) = 8.0, J(5,6) = 8.4 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 211.1 (C(7)); 138.1, 135.6, 132.9, 131.4 (4Car); 128.8, 127.8, 127.2, 126.4, 125.1, 122.8 (6Char); 44.1, 43.5 (C(8), C(10)); 30.8 (C(7)); 23.2 (C(11)). HR-MS (ESI-Q-ToF) calcd for C<sub>15</sub>H<sub>14</sub>NaO [M+Na]<sup>+</sup>: 233.0937; found: 233.0941; for C<sub>15</sub>H<sub>14</sub>LiO (M+Li<sup>+</sup>): 217.1200; found: 217.1219.

### 9. Silyl enol ethers **9** or **9/9'** mixture and silyl ether ketones **10** or **10/10'** mixture

**General procedure (e)**: To a solution of **8** (10 mmol, dried by evaporation with toluene) an dry toluene (20–50 mL) and NEt<sub>3</sub> (1.9 mL, 14 mmol, 1.4 equiv) was added at rt dropwise Me<sub>3</sub>SiOTf (2.15 mL, 12 mmol, 1.2 equiv) at rt under Ar. The solution was stirred at 80 °C for 2 h, then left at rt diluted with H<sub>2</sub>O (10 mL) and extracted with cyclohexane (100 mL). The organic phases were washed with brine (20 mL), dried over MgSO<sub>4</sub> and evaporated to give quantitatively **9** or **9/9'** mixture as orange resin which was used without further purification.

**General procedure (f)**: To a solution of **9** or **9/9'** mixture (10 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (10 mL), at 0 °C under Ar, was added portionwise *m*-CPBA (2.5 g, 15 mmol, 1.5 equiv) and stirred for 2 h (tlc monitoring). The solids were discarded by filtration and the organic phase evaporated. The solution of the residue in cyclohexane (100 mL) was washed with aqueous 1 M NaHCO<sub>3</sub> (20 mL) solution

(containing  $\text{Na}_2\text{SO}_3$  or  $\text{Na}_2\text{S}_2\text{O}_3 \cdot 5\text{H}_2\text{O}$  (10 mmol) to reduce the peracid excess) and with brine (20 mL), dried over  $\text{MgSO}_4$  and evaporated to give **10** or **10/10'** mixture as yellowish resin which was used without further purification.

**9.1. 1-Chloro-7-trimethylsilyloxy-6,9-dihydro-5H-benzocycloheptene (9b) and 4-chloro-7-trimethylsilyloxy-6,9-dihydro-5H-benzocycloheptene (9b'); 4-chloro-6-(trimethylsilyloxy)-5,6,8,9-tetrahydrobenzocyclohepten-7-one (10b) and 1-chloro-6-(trimethylsilyloxy)-5,6,8,9-tetrahydrobenzocyclohepten-7-one (10b')**

*General procedure (e)* with **8b** (4.2 g, 21.6 mmol),  $\text{Et}_3\text{N}$  (4.22 mL, 30.3 mmol, 1.4 equiv) in toluene (40 ml) and with  $\text{Me}_3\text{SiOTf}$  (4.7 mL, 26 mmol, 1.2 equiv) to give a 55:45 **9b/9b'** mixture (5.5 g, 96%).

Compound **9b/9b'**: brownish resin, characterised by NMR only.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 400 MHz), major isomer **9b**: 7.20 (m, 1H, H-3); 7.00 (m, 2Har); 5.07 (m, 1H, H-8); 3.51 (d, 2H,  $J = 6.6$  Hz,  $\text{CH}_2(9)$ ); 2.95 (m, 2H,  $\text{CH}_2(5)$ ); 2.31 (m, 2H,  $\text{CH}_2(6)$ ); 0.13 (s, 9H,  $\text{SiMe}_3$ ). Minor isomer **9b'**: 7.20 (m, 1H, H-3); 7.00 (m, 2Har); 5.07 (m, 1H, H-8); 3.30 (d, 2H,  $J = 6.3$  Hz,  $\text{CH}_2(9)$ ); 3.15 (m, 2H,  $\text{CH}_2(5)$ ); 2.31 (m, 2H,  $\text{CH}_2(6)$ ); 0.14 (s, 9H,  $\text{SiMe}_3$ ).

*General procedure (f)* with **9b/9b'** (5.5 g, 20.7 mmol) in  $\text{CH}_2\text{Cl}_2$  (20 mL), *m*-CPBA (5.6 g, 32.6 mmol, 1.5 equiv) for 2 h to give a 55:45 **10b/10b'** mixture (5.2 g, 85% from ketone **8b**).

Compound **10b/10b'**: brownish oil characterised by NMR only.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 400 MHz) major isomer **10b**: 7.20 (m, 1H, H-3); 7.00 (m, 2Har); 4.19 (ddd, 1H, H-6); 3.31 (dd, 1H, Ha-5); 3.15 (dd, 1H, Hb-5); 3.0–2.8 (m, 3H, Ha-8, Ha-9, Hb-9); 2.40 (t, 1H,  $J = 11.0$  Hz, Hb-C8); 0.15 (s, 9H,  $\text{CMe}_3$ );  $J(1,2) = 7.5$ ,  $J(1,3) = 1.2$ ,  $J(2,3) = 8.0$  Hz,  $J(5a,6) = 3.0$ ,  $J(5b,6) = 10.3$ ,  $J(6,8b) = 1.0$  Hz. Minor isomer **10b'**: 7.20 (m, 1H, H-2); 7.00 (m, 2Har); 4.20 (ddd, 1H, H-6); 3.17 (ddd, 1H, Ha-9); 3.14 (dd, 1H, Ha-5); 3.02 (m, 1H, Hb-9); 3.00 (m, 1H, Hb-5); 2.84 (ddd, 1H, Ha-8); 2.41 (ddt, 1H, Hb-8); 0.12 (s, 9H,  $\text{CMe}_3$ );  $J(2,3) = 8.0$ ,  $J(2,4) = 1.2$ ;  $J(3,4) = 7.4$ ,  $J(5a,5b) = 14.2$ ,  $J(5a,6) = 9.5$ ,  $J(5b,6) = 3.2$ ,  $J(6,8b) = 1.0$ ,  $J(8a,8b) = 13.6$ ,  $J(8a,9a) = 8.4$ ,  $J(8a,9b) = 2.3$ ,  $J(8b,9a) = 2.6$ ,  $J(8b,9b) = 11.0$ ,  $J(9a,9b) = 15.0$  Hz.

**9.2. 1-Bromo-7-trimethylsilyloxy-6,9-dihydro-5H-benzocycloheptene (9c) and 4-bromo-7-trimethylsilyloxy-6,9-dihydro-5H-benzocycloheptene (9c'); 4-bromo-6-(trimethylsilyloxy)-5,6,8,9-dihydrobenzocyclohepten-7-one (10c) and 1-bromo-6-(trimethylsilyloxy)-5,6,8,9-dihydrobenzocyclohepten-7-one (10c')**

*General procedure (e)* with **8c** (1.86 g, 7.78 mmol),  $\text{Et}_3\text{N}$  (1.63 mL, 11.7 mmol, 1.5 equiv) in toluene (20 ml) and with  $\text{Me}_3\text{SiOTf}$  (1.76 mL, 9.72 mmol, 1.25 equiv) to give a 55:45 **9c/9c'** mixture (2.35 g, quant.).

Compound **9c/9c'**: brownish resin, characterised by NMR only.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 400 MHz), major isomer **9c**: 7.40 (d, 1H, H-2); 7.10 (d, 1H, H-4); 6.98 (t, 1H, H-3); 5.08 (tt, 1H, H-8); 3.54 (dt, 2H,  $\text{CH}_2(9)$ ); 2.96 (m, 2H,  $\text{CH}_2(5)$ ); 2.31 (m, 2H,  $\text{CH}_2(6)$ ); 0.13 (s, 9H,  $\text{SiMe}_3$ );  $J(2,3) = 8.0$ ,  $J(2,4) = 1.2$ ,  $J(3,4) = 7.3$ ,  $J(6,8) = 1.4$ ,  $J(6,9) = 1.9$ ,  $J(8,9) = 6.5$  Hz. Minor isomer **9c'**: 7.41 (d, 1H, H-3); 7.01 (d, 1H, H-1); 6.94 (t, 1H, H-2); 5.05 (tt, 1H, H-8); 3.31 (dt, 2H,  $\text{CH}_2(9)$ ); 3.18 (m, 2H,  $\text{CH}_2(5)$ ); 2.31 (m, 2H,  $\text{CH}_2(6)$ ); 0.14 (s, 9H,  $\text{SiMe}_3$ );  $J(1,2) = 7.4$ ,  $J(1,3) = 1.2$ ,  $J(2,3) = 8.0$ ,  $J(6,8) = 1.3$ ,  $J(6,9) = 2.1$ ,  $J(8,9) = 6.2$  Hz.  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ , 100 MHz, **9c** M and **9c'** m isomers): 151.6 (C(7)M); 151.4 (C(7)m); 144.5, 143.1, 141.7, 139.8 (C(4a)M+m, C(9a)M+m); 130.6 (C(3)m), 130.5 (C(2)M); 127.42, 127.37, 127.3 (C(4)M, C(3)M, C(2)m); 126.8 (C(1)m); 123.7, 122.8 (C(4)m, C(1)M); 104.2 (C(8)M); 103.5 (C(8)m); 34.2 (C(6)M); 32.6 (C(6)m); 31.4 (C(5)M); 30.7 (C(5)m); 29.1 (C(9)m); 27.8 (C(9)M); 0.04 ( $\text{SiMe}_3$  M+m).

*General procedure (f)* with **9c/9c'** (2.35 g, 7.55 mmol) in  $\text{CH}_2\text{Cl}_2$  (20 mL), *m*-CPBA (2.04 g, 8.31 mmol, 1.1 equiv) for 3 h to give a 55:45 **10c/10c'** mixture (2.32 g, 91% from ketone **8c**).

Compound **10c/10c'**: brownish oil characterised by NMR only.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 400 MHz) isomer **10c**: 7.52 (dd, 1H, H-3); 7.13 (dd, 1H, H-1); 7.05 (t, 1H, H-2); 4.19 (ddd, 1H, H-6); 3.42 (dd, 1H, Ha-5); 3.21 (dd, 1H, Hb-5); 3.0–2.8 (m, 3H, Ha-8, Ha-9, Hb-9); 2.41 (td, 1H,  $J = 11.0$ , 1.0 Hz, Hb-C8); 0.15 (s, 9H,  $\text{CMe}_3$ );  $J(1,2) = 7.5$ ,  $J(1,3) = 1.2$ ,  $J(2,3) = 8.0$  Hz,  $J(5a,5b) = 14.6$ ,  $J(5a,6) = 3.0$ ,  $J(5b,6) = 10.3$ ,  $J(6,8b) = 1.0$  Hz. Isomer **10c'**: 7.51 (dd, 1H, H-2); 7.18 (dd, 1H, H-4); 7.06 (t, 1H, H-3); 4.20 (ddd, 1H, H-6); 3.27 (ddd, 1H, Ha-9); 3.14 (dd, 1H, Ha-5); 3.02 (m, 1H, Hb-9); 3.00 (m, 1H, Hb-5); 2.84 (ddd, 1H, Ha-8); 2.41 (ddt, 1H, Hb-8); 0.12 (s, 9H,  $\text{CMe}_3$ );  $J(2,3) = 8.0$ ,  $J(2,4) = 1.2$ ;  $J(3,4) = 7.4$ ,  $J(5a,5b) = 14.2$ ,  $J(5a,6) = 9.5$ ,  $J(5b,6) = 3.2$ ,  $J(6,8b) = 1.0$ ,  $J(8a,8b) = 13.6$ ,  $J(8a,9a) = 8.4$ ,  $J(8a,9b) = 2.3$ ,  $J(8b,9a) = 2.6$ ,  $J(8b,9b) = 11.0$ ,  $J(9a,9b) = 15.0$  Hz.

**9.3. 1-Bromo-4-phenyl-7-(trimethylsilyloxy)-6,9-dihydro-5H-benzocycloheptene (9d) and 4-bromo-1-phenyl-7-(trimethylsilyloxy)-6,9-dihydro-5H-benzocycloheptene (9d'); 4-bromo-1-phenyl-6-(trimethylsilyloxy)-5,6,8,9-tetrahydrobenzocyclohepten-7-one (10d) and 1-bromo-4-phenyl-6-(trimethylsilyloxy)-5,6,8,9-tetrahydrobenzocyclohepten-7-one (10d')**

*General procedure (e)* with **8d** (1.19 g, 3.78 mmol),  $\text{Et}_3\text{N}$  (0.8 mL, 5.29 mmol, 1.5 equiv) in toluene (15 ml) and with  $\text{Me}_3\text{SiOTf}$  (0.9 mL, 4.53 mmol, 1.2 equiv) for 4 h, to give a 60:40 **9d/9d'** mixture (1.39 g, 95%).

Compound **9d/9d'**: orange resin, characterised by NMR only.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 300 MHz) major isomer **9d**: 7.45–7.33 (m, 5 Har); 7.25 (m, 1Har); 6.95 (d, 1Har,  $J(2,3) = 8.3$  Hz, H-3); 5.11 (tt, 1H, H-8); 3.62 (dt, 2H,  $\text{CH}_2(9)$ ); 2.90 (m, 2H,  $\text{CH}_2(5)$ ); 2.24 (m, 2H,  $\text{CH}_2(6)$ ); 0.15 (s, 9H,  $\text{CMe}_3$ );  $J(6,8) = 1.4$ ,  $J(6,9) = 2.0$ ,  $J(8,9) = 6.6$  Hz. Minor isomer **9d'**: 7.45–7.33 (m, 5 Har); 7.25 (m, 1Har); 6.95 (2 d, 1Har,  $J(2,3) = 8.3$  Hz, H-2); 4.95 (tt, 1H, H-8); 3.29 (m, 2H,  $\text{CH}_2(5)$ ); 3.24 (dt, 2H,  $\text{CH}_2(9)$ ); 2.35 (m, 2H,  $\text{CH}_2(6)$ ); 0.15 (s, 9H,  $\text{CMe}_3$ );  $J(6,8) = 1.4$ ,  $J(6,9) = 2.0$ ,  $J(8,9) = 6.6$  Hz.

*General procedure (f)* with **9d/9d'** (1.46 g, 3.78 mmol) in  $\text{CH}_2\text{Cl}_2$  (25 mL), *m*-CPBA (1.03 g, 5.97 mmol, 1.5 equiv) for 3 h to give a 60/40 **10d/10d'** mixture (1.37 g, 90% from ketone **8d**).

Compound **10d/10d'**: yellowish oil, characterised by NMR only.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 300 MHz) major isomer **10d**: 7.55 (d, 1H, H-3); 7.46–7.36 (m, 4 Har); 7.3–7.2 (m, 1 Har); 7.04 (d, 1H, H-2); 4.22 (ddd, 1H, H-6); 3.57 (dd, 1H, Ha-5); 3.27 (dd, 1H, Hb-5); 3.16–2.75 (m, 3H, Ha-8,  $\text{CH}_2(9)$ ); 2.33 (ddt, 1H, Hb-8); 0.18 (s, 9H,  $\text{CMe}_3$ );  $J(2,3) = 8.2$ ,  $J(5a,6) = 3.3$ ,  $J(5b,6) = 10.2$ ,  $J(5a,5b) = 14.4$ ,  $J(6,8b) = 1.1$ ,  $J(8a,8b) = 13.5$ ,  $J(8b,9a) = 3.7$ ,  $J(8b,9b) = 10.8$  Hz. Minor isomer **10d'**: 7.54 (d, 1H, H-2); 7.46–7.36 (m, 4 Har); 7.3–7.2 (m, 1 Har); 7.05 (d, 1H, H-3); 4.12 (ddd, 1H, H-6); 3.40 (ddd, 1H, Ha-9); 3.16–2.75 (m, 4H,  $\text{CH}_2(5)$ , Ha-8, Hb-9); 2.46 (ddt, 1H, Hb-8); 0.18 (s, 9H,  $\text{CMe}_3$ );  $J(2,3) = 8.2$ ,  $J(5a,6) = 4.0$ ,  $J(5b,6) = 9.0$ ,  $J(6,8b) = 1.1$ ,  $J(8a,8b) = 13.9$ ,  $J(8a,9a) = 7.8$ ,  $J(8b,9a) = 3.1$ ,  $J(8b,9b) = 11.2$ ,  $J(9a,9b) = 14.8$  Hz.

**9.4. 1,4-Dibromo-7-(trimethylsilyloxy)-8,9-dihydro-5H-benzocycloheptene (9e); 1,4-dibromo-6-(trimethylsilyloxy)-5,6,8,9-tetrahydrobenzocyclohepten-7-one (10e)**

*General procedure (e)* with **8e** (6.3 g, 19.8 mmol),  $\text{Et}_3\text{N}$  (3.9 mL, 27.2 mmol, 1.4 equiv) in toluene (100 ml) and with  $\text{Me}_3\text{SiOTf}$  (4.7 mL, 23.8 mmol, 1.2 equiv) to give **9e** (7.79 g, quant.).

Compound **9e**: orange resin, characterised by NMR only.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ , 400 MHz): 7.25 (s, 2H, H-2, H-3); 5.04 (tt, 1H, H-8); 3.58 (dt, 2H,  $\text{CH}_2(9)$ ); 3.24 (m, 2H,  $\text{CH}_2(5)$ ); 2.30 (m, 2H,  $\text{CH}_2(6)$ ); 0.14 (s, 9H,  $\text{SiMe}_3$ );  $J(6,8) = 1.4$ ,  $J(6,9) = 1.9$ ,  $J(8,9) = 6.6$  Hz.

**General procedure (f)** with **9e** (7.79 g, 19.8 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (180 mL), *m*-CPBA (5.17 g, 29.9 mmol, 1.5 equiv) for 3 h to give **10e** (7.75 g, 96%, 90% from ketone **8e**).

Compound **10e**: brownish oil, characterised by NMR only. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 0.15 (s, 9H, SiMe<sub>3</sub>); 2.41 (dddd, 1H, Hb-8); 2.87 (ddd, 1H, Ha-8); 3.13 (ddd, 1H, Hb-9); 3.32 (ddd, 1H, Ha-9); 3.33 (dd, 1H, Hb-5); 3.45 (dd, 1H, Ha-5); 4.16 (ddd, 1H, H-6); 7.37 (m, 2H, H-2, H-3); *J*(5a,5b) = 14.6, *J*(5a,6) = 3.2, *J*(5b,6) = 9.8, *J*(6,8b) = 1.0, *J*(8a,8b) = 14.2, *J*(8a,9a) = 8.2, *J*(8a,9b) = 2.8, *J*(8b,9a) = 2.0, *J*(8b,9b) = 11.0, *J*(9a,9b) = 15.0 Hz.

**9.5. 9-Trimethylsilyloxy-8,11-dihydro-7H-cyclohepta[a]naphthalene (9h) and 9-trimethylsilyloxy-10,11-dihydro-7H-cyclohepta[a]naphthalene (9h'); 10-trimethylsilyloxy-7,8,10,11-tetrahydro-cyclohepta[a]naphthalen-9-one (10h) and 8-trimethylsilyloxy-7,8,10,11-tetrahydro cyclohepta[a]naphthalen-9-one (10h')**

**General procedure (e)** with **8h** (0.74 g, 3.48 mmol), Et<sub>3</sub>N (0.67 mL, 4.87 mmol, 1.4 equiv) in toluene (5 mL) and with Me<sub>3</sub>-SiOTf (0.75 mL, 4.17 mmol, 1.2 equiv) to give a 50:50 **9h/9h'** mixture (0.90 g, 91%).

Compound **9h/9h'**: as orange resin, characterised by NMR only. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz, isomer **9h** M and **9h'** m mixture): 8.14 (d, *J* = 8.6 Hz, 1H M+m, Har-1); 7.85 (2d, 1H M+m, *J* = 8.2 Hz, Har-4); 7.69, 7.66 (2d, 1H M+m, *J* = 8.3 Hz, Har-5), 7.52 (m, 1H M+m); 7.42 (m, 1H M, 2H m); 7.37 (d, 1H M, *J* = 8.3 Hz, H-6); 5.23 (t, 1H M, *J* = 6.4 Hz, H-10), 5.17 (t, 1H m, *J* = 6.2 Hz, H-8 m), 3.78 (m, 2H M+m); 3.50, 3.46 (2 m, 2H M, 4H m); 3.13 (m, 2H M); 0.13 (s, 9H M+m, SiMe<sub>3</sub>).

**General procedure (f)** with **9h/9h'** (0.90 g, 3.16 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (15 mL), *m*-CPBA (0.86 g, 5.0 mmol, 1.6 equiv) for 3 h to give a 50:50 **10h/10h'** mixture as (0.94 g, 96%, 90% from ketone **8h**).

Compound **10h/10h'**: brownish oil, characterised by NMR only. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz) isomer **10h** M and **10h'** m mixture: 8.16–7.33 (m, 6Har M+m); 4.30, 4.27 (2dd, *J* = 3.9, 10.0 Hz, 1H M+m, H-10 M, H-8 m); 3.34 (m, 2H M); 3.14 (m, 2H M+m); 2.93 (m, 2H m); 2.48 (m, 1H M+m); 0.15 (s, 9H M+m, SiMe<sub>3</sub>).

## 10. Hydroxy-amides 11b–e, b'–d', h, h'

**General procedure (g)**: To a solution of **10** (10 mmol) in 2 M NH<sub>3</sub> solution in dry EtOH (50–100 mL), was added Ti(OiPr)<sub>4</sub> (6.2 mL, 20 mmol, 2 equiv) and the mixture stirred for 6 h under Ar. NaBH<sub>4</sub> (0.56 g, 15 mmol, 1.5 equiv) was then added and the mixture stirred further for 2 h and evaporated. The residue was dissolved in AcOEt (50–100 mL) and vigorously stirred with aqueous 1 M NH<sub>4</sub>OH solution (20–50 mL) for 2 h. The solids were filtered out and washed thrice with a mixture of AcOEt (20 mL) and aqueous 1 M NH<sub>4</sub>OH solution (20 mL). The organic phase was separated, the aqueous phase extracted with AcOEt (3 × 20 mL), the combined organic phases dried over MgSO<sub>4</sub> and evaporated to give the crude amine.

A solution of the crude amine (10 mmol) in MeOH (50–100 mL) was stirred with solid NaHCO<sub>3</sub> (1.4 g, 13 mmol, 1.3 equiv) and Boc<sub>2</sub>O (3.3 g, 15 mmol, 1.5 equiv) for 16 h at rt under Ar. The solvent was evaporated, the solution of the residue in AcOEt was washed with H<sub>2</sub>O, dried over MgSO<sub>4</sub> and evaporated to give the crude amide, which was purified by FC (cyclohexane/AcOEt 8:2) to give the amide-alcohol **11** or regioisomer **11/11'** mixture as *cis/trans* mixture.

**10.1. 7-tert-Butoxycarbonylamino-4-chloro-6,7,8,9 tetrahydro-5H-benzocyclohepten-6-ol (cis/trans mixture, 11b) and 7-tert-Butoxycarbonylamino-1-chloro-6,7,8,9 tetrahydro-5H-cyclohepten-6-ol (cis/trans mixture, 11b')**

**General procedure (g)** with **10b/10b'** (5.2 g, 18.5 mmol), in 2 M NH<sub>3</sub> solution in EtOH (120 mL) and with Ti(OiPr)<sub>4</sub> (12.2 mL,

41.3 mmol, 2.2 equiv) for 6 h, then reduction with NaBH<sub>4</sub> (1.2 g, 31 mmol, 1.6 equiv) for 2 h to give the crude amine. Acylation with Boc<sub>2</sub>O (5.4 g, 24.8 mmol, 1.4 equiv), NaHCO<sub>3</sub> in MeOH (100 mL) for 16 h gave the **11b/11b'** mixture (4.45 g, 78%). The isomeric mixture resolution by flash chromatography occurred badly and the *cis/trans* isomeric mixture of **11b'** only was obtained as pure regioisomer.

Isomeric mixture **11b/11b'**: cream crystals, mp 183–184 °C. IR (KBr): 3353, 2980, 2932, 1683, 1525, 1450, 1245, 1170, 779 cm<sup>-1</sup>. HR-MS (ESI-Q-ToF) calcd for C<sub>16</sub>H<sub>22</sub>ClLiNO<sub>3</sub> [M+Li]<sup>+</sup>: 318.1443; found: 318.1419; calcd for C<sub>16</sub>H<sub>22</sub>ClNaNO<sub>3</sub> [M+Na]<sup>+</sup>: 334.1180; found: 334.1159.

*Cis-trans* isomeric mixture **11b'**: colorless crystals, mp 165–168 °C. IR (KBr): 3362, 2982, 2933, 1682, 1665, 1525, 1448, 1391, 1367, 1324, 1248, 1169, 1079, 1042, 779 cm<sup>-1</sup>. HR-MS (ESI<sup>+</sup> IsoToF) calcd for C<sub>16</sub>H<sub>22</sub>ClNaNO<sub>3</sub> [M+Na]<sup>+</sup>: 334.1180; found: 334.1176.

*cis-11b*, <sup>1</sup>H NMR (CDCl<sub>3</sub>, 300 MHz): 7.26 (dd, *J* = 1.5, 7.9 Hz, H-3); 7.08 (dd, *J* = 7.5, 7.9 Hz, H-2); 7.02 (dd, *J* = 1.5, 7.5 Hz, H-1); 5.04 (d, *J* = ca. 8 Hz, NH); 4.18 (broad s, H-6); 3.77 (broad m, H-7, Ha-5); 2.88 (d, *J* = 14.6 Hz, Hb-5); 2.77 (m, CH<sub>2</sub>(9)); 1.99 (m, Ha-8); 1.57 (m, (br m, Hb-8)); 1.45 (s, CMe<sub>3</sub>). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 75 MHz): 155.5 (NCO); 144.8 (C(4a)); 135.7 (C(9a)); 132.4 (C(4)); 128.0, 127.8, 127.4 (C(1), C(2), C(3)); 79.5 (CMe<sub>3</sub>); 68.8 (C(6)); 56.6 (C(7)); 33.7 (C(5)); 32.7 (C(9)); 28.6 (C(8)); 28.4 (CMe<sub>3</sub>).

*trans-11b*, <sup>1</sup>H NMR (CDCl<sub>3</sub>, 300 MHz): 7.25 (dd, H-3); 7.05 (t, H-2); 6.98 (br d, H-1); 4.54 (br d, NH); 3.70 (br m, H-7); 3.64 (dd, Ha-5); 3.35 (br t, H-6); 3.08 (br s, OH); 2.86 (dd, Hb-5); 2.85 (ddd, Ha-9); 2.75 (ddd, Hb-9); 2.21 (m, Ha-8); 1.46 (s, CMe<sub>3</sub>); 1.33 (dq, Hb-8). *J*(1,2) = 7.5, *J*(1,3) = 1.5, *J*(2,3) = 8.1, *J*(5a,5b) = 14.4, *J*(5a,6) = 1.8, *J*(5b,6) = 10.2, *J*(6,7) = 9.5, *J*(NH,7) = ca. 8, *J*(7,8a) = 4.4, *J*(7,8b) = 10.5, *J*(8a,8b) = 13.6, *J*(8a,9a) = 1.6, *J*(8a,9b) = 7.7, *J*(8b,9a) = 12.0, *J*(8b,9b) = 1.8, *J*(9a,9b) = 14.7 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 75 MHz): 155.5 (NCO); 144.6 (C(4a)); 134.6, 134.2 (C(9a), C(4)); 127.8, 127.6, 127.1 (C(1), C(2), C(3)); 80.2 (CMe<sub>3</sub>); 73.6 (C(6)); 60.0 (C(7)); 36.2 (C(5)); 32.7, 32.2 (C(8), C(9)); 28.4 (CMe<sub>3</sub>).

*cis-11b'*, <sup>1</sup>H NMR (CDCl<sub>3</sub>, 300 MHz): 7.28 (m, H-2); 7.06 (m, H-3, H-4); 5.04 (br d, *J* = 7.5 Hz, NH); 4.10 (br s, H-6); 3.81 (broad s, H-7); 3.36 (broad m, Ha-9); 3.06 (m, CH<sub>2</sub>(5)); 2.55 (broad t, *J* = 12.8 Hz, Hb-9); 2.00 (ddd, *J* = 4.4, 8.2, 13.0 Hz, Ha-8); 1.45 (s, CMe<sub>3</sub>); 1.31 (m, Hb-8). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 75 MHz): 155.3 (NCO); 139.8 (C(9a)); 136.8 (C(4a)); 133.5 (C(1)); 130.2, 128.7, 127.2 (C(2), C(3), C(4)); 79.6 (CMe<sub>3</sub>); 69.1 (C(6)); 56.7 (C(7)); 39.6 (C(5)); 28.4 (CMe<sub>3</sub>); 27.7, 26.8 (C(8), C(9)).

*trans-11b'*, <sup>1</sup>H NMR (CDCl<sub>3</sub>, 300 MHz): 7.23 (dd, H-2); 7.09 (dd, H-4); 7.05 (t, H-3); 4.58 (br d, NH); 3.70 (br m, H-7); 3.40 (ddd, Ha-9); 3.34 (dt, H-6); 3.12–2.96 (m, Ha-5, Hb-5); 2.62 (ddd, Hb-9); 2.20 (m, Ha-8); 1.46 (s, CMe<sub>3</sub>); 1.29 (dq, Hb-8); *J*(2,3) = 7.6, *J*(2,4) = 1.5, *J*(3,4) = 7.6, *J*(5a,5b) = 14.0, *J*(5a,6) = ca. 9, *J*(5b,6) = 4.0, *J*(6,7) = 9.0, *J*(NH,7) = ca. 7, *J*(7,8a) = 4.3, *J*(7,8b) = 11.3, *J*(8a,8b) = 13.8, *J*(8a,9a) = 8.0, *J*(8a,9b) = 1.3, *J*(8b,9a) = 1.5, *J*(8b,9b) = 11.5, *J*(9a,9b) = 15.0 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 75 MHz): 156.8 (NCO); 139.4, 139.1 (C(4a), C(9a)); 133.1 (C(1)); 128.8, 128.0, 127.3 (C(2), C(3), C(4)); 80.4 (CMe<sub>3</sub>); 74.5 (C(6)); 59.9 (C(7)); 41.8 (C(5)); 31.6 (C(8)); 28.4 (CMe<sub>3</sub>); 26.3 (C(9)).

**10.2. 4-Bromo-7-tert-butoxycarbonylamino-5,6,8,9-tetrahydro-5H-benzocyclohepten-6-ol (cis/trans mixture, 11c) and 1-bromo-7-tert-butoxycarbonylamino-5,6,8,9-tetrahydro-5H-benzocyclohepten-6-ol (cis/trans mixture, 11c')**

**General procedure (g)** with **10c/10c'** (2.32 g, 7.09 mmol), in 2 M NH<sub>3</sub> solution in EtOH (40 mL) and with Ti(OiPr)<sub>4</sub> (3.87 mL, 14.2 mmol, 2 equiv) for 16 h, then reduction with NaBH<sub>4</sub> (0.4 g, 10.6 mmol, 1.5 equiv) for 2 h to give the crude amine. Acylation with Boc<sub>2</sub>O (2.56 g, 11.7 mmol, 1.5 equiv), Na<sub>2</sub>CO<sub>3</sub> (1.24 g, 10.1 mmol, 1.3 equiv) in MeOH (25 mL) for 3 h gave the **11c/11c'**

mixture (1.63 g, 64%) after crystallisation and washing with  $iPr_2O$  at 0 °C. The isomer separation occurred by semi-preparative HPLC (on  $C_{18}$  Kromasil 100, 5  $\mu$ m, 4.6  $\times$  250 mm) with MeOH/ $H_2O$  75:25 as eluent, to give **11c** as *cis/trans* mixture (250 mg, 10%) and **11c'** as *cis/trans* mixture (1.10 g, 43%) as. The *cis*-**11c** isomer could be obtained pure.

Isomeric mixture **11c/11c'**: cream solid, mp 179–184 °C. IR (KBr): 3358, 2978, 2934, 1681, 1667, 1523, 1446, 1367, 1248, 1166, 1043, 777  $cm^{-1}$ . HR-MS (ESI-Q-ToF) calcd for  $C_{16}H_{22}BrLiNO_3$  [M+Li]<sup>+</sup>: 362.0938 and 364.0919; found: 362.0849 and 364.0829.

*trans*-**11c**: <sup>1</sup>H NMR ( $CDCl_3$ , 400 MHz): 7.44 (dd, 1H, H-3); 7.01 (br d, 1H, H-1); 6.97 (t, 1H, H-2); 4.53 (br d, 1H, NH); 3.69 (br m, 1H, H-7); 3.63 (dd, 1H, Ha-5); 3.35 (br t, 1H, H-6); 3.08 (br s, 1H, OH); 2.95 (dd, 1H, Hb-5); 2.87 (ddd, 1H, Ha-9); 2.76 (ddd, 1H, Hb-9); 2.21 (m, 1H, Ha-8); 1.46 (s, 9H,  $CM_e_3$ ); 1.32 (dq, 1H, Hb-8);  $J(1,2) = 7.4$ ,  $J(1,3) = 1.6$ ,  $J(2,3) = 8.0$ ,  $J(5a,5b) = 14.2$ ,  $J(5a,6) = 1.6$ ,  $J(5b,6) = 10.0$ ,  $J(6,7) = 9.0$ ,  $J(NH,7) = 8.2$ ,  $J(7,8a) = 4.2$ ,  $J(7,8b) = 11.3$ ,  $J(8a,8b) = 13.8$ ,  $J(8a,9a) = 1.8$ ,  $J(8a,9b) = 7.5$ ,  $J(8b,9a) = 11.0$ ,  $J(8b,9b) = 2.0$ ,  $J(9a,9b) = 14.7$  Hz. <sup>13</sup>C NMR ( $CDCl_3$ , 100 MHz): 156.9 (NCO); 144.6 (C(4a)); 136.2 (C(9a)); 131.4 (C(3)); 128.2, 128.0 (C(1), C(2)); 125.7 (C(4)); 80.5 ( $CM_e_3$ ); 73.7 (C(6)); 60.1 (C(7)); 39.7 (C(5)); 32.7 (C(8)); 28.5 (C(9)); 28.5 ( $CM_e_3$ ).

*cis*-**11c**: colorless crystals, mp 183–184 °C. IR (KBr): 775, 1014, 1043, 1085, 1164, 1251, 1367, 1390, 1453, 1524, 1665, 2981, 2933, 3375, 3471  $cm^{-1}$ . <sup>1</sup>H NMR ( $CDCl_3$ , 400 MHz): 7.46 (d, 1H,  $J = 8.0$  Hz, H-3); 7.06 (d, 1H,  $J = 7.3$  Hz, H-1); 7.01 (dd, 1H,  $J = 7.3$ , 8.0 Hz, H-2); 5.04 (br d, 1H,  $J = 8.5$  Hz, NH); 4.19 (br s, 1H, H-6); 3.77 (br s, 2H, H-7, Ha-5); 2.98 (d, 1H,  $J = 14.3$  Hz, Hb-5); 2.79 (m, 2H,  $CH_2(9)$ ); 1.99 (br m, 1H, Ha-8); 1.50 (br m, 1H, Hb-8); 1.45 (s, 9H,  $CM_e_3$ ). <sup>13</sup>C NMR ( $CDCl_3$ , 100 MHz): 155.7 (NCO); 145.0 (C(4a)); 134.7 (C(9a)); 131.5 (C(3)); 128.6, 128.3 (C(1), C(2)); 127.1 (C(4)); 79.7 ( $CM_e_3$ ); 69.1 (C(6)); 57.0 (C(7)); 37.4 (C(5)); 33.2 (C(9)); 28.9 (C(8)); 28.7 ( $CM_e_3$ ). HR-MS (ESI<sup>+</sup>) calcd for  $C_{16}H_{22}BrNNaO_3$  [M+Na]<sup>+</sup>: 378.0675; found: 378.0665.

*trans*-**11c'**: <sup>1</sup>H NMR ( $CDCl_3$ , 400 MHz): 7.43 (dd, 1H, H-2); 7.14 (br d, 1H, H-4); 6.98 (t, 1H, H-3); 4.52 (br d, 1H, NH); 3.70 (br m, 1H, H-7); 3.40 (ddd, 1H, Ha-9); 3.35 (dt, 1H, H-6); 3.06, 3.02 (m, 2H, Ha-5, Hb-5); 2.71 (ddd, 1H, Hb-9); 2.20 (m, 1H, Ha-8); 1.46 (s, 9H,  $CM_e_3$ ); 1.31 (dq, 1H, Hb-8);  $J(2,3) = 8.0$ ,  $J(2,4) = 1.0$ ,  $J(3,4) = 7.4$ ,  $J(5a,5b) = 14.0$ ,  $J(5a,6) = 10.4$ ,  $J(5b,6) = 3.2$ ,  $J(6,7) = 9.0$ ,  $J(NH,7) = ca 7$ ,  $J(7,8a) = 4.4$ ,  $J(7,8b) = 11.2$ ,  $J(8a,8b) = 13.8$ ,  $J(8a,9a) = 8.2$ ,  $J(8a,9b) = 1.6$ ,  $J(8b,9a) = 1.6$ ,  $J(8b,9b) = 11.2$ ,  $J(9a,9b) = 15.0$  Hz.

*cis*-**11c'**: <sup>1</sup>H NMR ( $CDCl_3$ , 400 MHz, 330 K): 7.47 (dd, 1H, H-2); 7.09 (br d, 1H, H-4); 6.98 (t, 1H, H-3); 4.91 (br d, 1H, NH); 4.11 (t, 1H, H-6); 3.81 (m, 1H, H-7); 3.39 (ddd, 1H, Ha-9); 3.12, 3.07 (m, 2H, Ha-5, Hb-5); 2.67 (dd, 1H, Hb-9); 2.00 (ddd, 1H, Ha-8); 1.47 (dq, 1H, Hb-8); 1.46 (s, 9H,  $CM_e_3$ ); 1.28 (br s, 1H, OH);  $J(2,3) = 8.0$ ,  $J(2,4) = 1.0$ ,  $J(3,4) = 7.4$ ,  $J(5a,5b) = 14.4$ ,  $J(5a,6) = 1.8$ ,  $J(5b,6) = 7.2$ ,  $J(6,7) = 2.2$ ,  $J(OH,6) = 7.6$ ;  $J(NH,7) = 8.0$ ,  $J(7,8a) = 4.2$ ,  $J(7,8b) = 11.6$ ,  $J(8a,8b) = 14.0$ ,  $J(8a,9a) = 8.4$ ,  $J(8a,9b) = 1.2$ ,  $J(8b,9a) = 1.5$ ,  $J(8b,9b) = 11.6$ ,  $J(9a,9b) = 14.8$  Hz.

### 10.3. 4-Bromo-7-(*tert*-butoxycarbonylamino)-1-phenyl-6,7,8,9-tetrahydro-5H-benzocyclohepten-6-ol (*cis/trans* mixture, **11d**) and 1-bromo-7-(*tert*-butoxycarbonylamino)-4-phenyl-6,7,8,9-tetrahydro-5H-benzocyclohepten-6-ol (*cis/trans* mixture, **11d'**)

*General procedure (g)* with **10d/10d'** (1.52 g, 3.77 mmol) in 2 M  $NH_3$  solution in EtOH (10 mL) and with  $Ti(OiPr)_4$  (2.3 mL, 7.54 mmol, 2 equiv) for 6 h, then reduction with  $NaBH_4$  (213 mg, 5.65 mmol, 1.5 equiv) for 3 h to give the crude amine (945 mg, 76%). Acylation with  $Boc_2O$  (940 mg, 4.27 mmol, 1.5 equiv),  $NaHCO_3$  (392 mg, 3.70 mmol, 1.3 equiv) in MeOH (20 mL) for 16 h gave the 60:40 mixture of **11d/11d'** (518 mg, 32%) inseparable isomers.

Isomeric mixture **11d/11d'**: colorless crystals, mp 174–176 °C. IR (KBr): 3487, 3362, 2979, 2932, 1664, 1530, 1252, 1168  $cm^{-1}$ .

HR-MS (ESI-Q-ToF) calcd for  $C_{22}H_{26}BrLiNO_3$  [M+Li]<sup>+</sup>: 438.1251 and 440.1233; found: 438.1237 and 440.1221.

*cis*-**11d** or **11d'**: <sup>1</sup>H NMR ( $CDCl_3$ , 400 MHz): 7.49 (d,  $J = 8.3$  Hz, 1Har); 7.45–7.2 (m, 5Har); 7.00 (d,  $J = 8.3$  Hz, 1Har); 5.03 (br s, 1H, NH); 4.24 (br s, 1H, H-6); 3.79 (br s, 2H); 3.06 (br s, 1H); 2.94 (br s, 1H); 2.50 (br s, 1H); 1.90 (br s, 1H); 1.44 (s, 10H).

*cis*-**11d'** or **11d**: <sup>1</sup>H NMR ( $CDCl_3$ , 400 MHz): 7.51 (d,  $J = 8.3$  Hz, 1Har); 7.45–7.2 (m, 5Har); 6.98 (d,  $J = 8.3$  Hz, 1Har); 4.96 (br d,  $J = 8.8$  Hz, 1H, NH); 4.02 (br s, 1H, H-6); 3.79 (br s, 2H); 3.50 (br s, 1H); 3.28 (br s, 1H); 2.84 (br d,  $J = 15$  Hz, 1H); 2.04 (br s, 1H); 1.44 (br s, 10H).

*trans*-**11d**: <sup>1</sup>H NMR ( $CDCl_3$ , 400 MHz): 7.48 (d, 1H, H-3); 7.42–7.33 (m, 3Har); 7.19 (m, 2Har); 6.96 (d, 1H, H-2); 4.54 (br s, 1H, NH); 3.69 (br d, 2H, H-7, Ha-5); 3.43 (br t, 1H, H-6); 3.20 (br s, 1H, OH); 3.05 (dd, 1H, Hb-5); 2.93 (dd, 1H, Ha-9); 2.61 (dd, 1H, Hb-9); 2.10 (m, 1H, Ha-8); 1.44 (s, 9H,  $CM_e_3$ ); 1.29 (br q, 1H, Hb-8);  $J(2,3) = 8.2$ ,  $J(5a,5b) = 14.2$ ,  $J(5a,6) = 1.6$ ,  $J(5b,6) = 10.2$ ,  $J(6,7) = 8.6$ ,  $J(7,8a) = 4.4$ ,  $J(7,8b) = 11.4$ ,  $J(8a,8b) = 13.6$ ,  $J(8a,9a) = 8.0$ ,  $J(8a,9b) = 1.2$ ,  $J(8b,9a) = 1.6$ ,  $J(8b,9b) = 11.0$ ,  $J(9a,9b) = 14.8$  Hz.

*trans*-**11d'**: <sup>1</sup>H NMR ( $CDCl_3$ , 400 MHz): 7.47 (d, 1H, H-2); 7.44–7.33 (m, 3Har); 7.25 (m, 2Har); 6.98 (d, 1H, H-3); 4.56 (br s, 1H, NH); 3.70 (br q, 1H, H-7); 3.50 (dd, 1H, Ha-9); 3.40 (br t, 1H, H-6); 3.19 (br s, 1H, OH); 3.14 (dd, 1H, Ha-5); 2.82 (dd, 1H, Hb-5); 2.78 (dd, 1H, Hb-9); 2.25 (m, 1H, Ha-8); 1.45 (s, 9H,  $CM_e_3$ ); 1.37 (br q, 1H, Hb-8);  $J(2,3) = 8.2$ ,  $J(5a,5b) = 14.4$ ,  $J(5a,6) = 1.8$ ,  $J(5b,6) = 10.2$ ,  $J(6,7) = 10.8$ ,  $J(7,8a) = 4.6$ ,  $J(7,8b) = 11.8$ ,  $J(8a,8b) = 13.6$ ,  $J(8a,9a) = 8.0$ ,  $J(8a,9b) = 1.0$ ,  $J(8b,9a) = 1.6$ ,  $J(8b,9b) = 11.4$ ,  $J(9a,9b) = 15.0$  Hz.

### 10.4. 1,4-Dibromo-7-(*tert*-butoxycarbonylamino)-6,7,8,9-tetrahydro-5H-benzocyclohepten-6-ol (*cis/trans* mixture, **11e**)

*General procedure (g)* with **10e** (7.75 g, 19.1 mmol) in 2 M  $NH_3$  solution in EtOH (60 mL) and with  $Ti(OiPr)_4$  (22.6 mL, g, 76.3 mmol, 4 equiv) for 6 h, then reduction with  $NaBH_4$  (1.08 g, 28.6 mmol, 1.5 equiv) for 3 h to give the crude amine (10.3 g, quant.). Acylation with  $Boc_2O$  (6.48 g, 29.7 mmol, 1.5 equiv),  $NaHCO_3$  (2.73 g, 25.7 mmol, 1.3 equiv) in MeOH (100 mL) for 16 h gave **11e** (1.20 g, 15%) after purification by flash chromatography (cyclohexane/AcOEt 8/2). The *trans*-**11e** isomer could be obtained pure.

*cis*-**11e**: cream resin. IR (KBr): 3450, 3355, 1667, 1529, 1367, 1168  $cm^{-1}$ . <sup>1</sup>H NMR ( $CDCl_3$ , 400 MHz): 7.32, 7.31 (2 d,  $J = 8.6$  Hz, 2H, H-2, H-3); 4.92 (br s, 1H, NH); 4.22 (d, 1H, H-6); 3.75 (m, 1H, H-7); 3.70 (m, 1H, Ha-9); 3.43 (m, 1H, Ha-5); 3.09 (m, 1H, Hb-5); 2.77 (m, 1H, Hb-9); 2.01 (m, 1H, Ha-8); 1.53 (m, 1H, Hb-8); 1.53 (s, 9H,  $CM_e_3$ ). <sup>13</sup>C NMR ( $CDCl_3$ , 100 MHz): 27.1 (C(9)); 28.4 ( $CM_e_3$ ); 31.2 (C(8)); 37.9 (C(5)); 56.1 (C(7)); 68.3 (C(6)); 79.6 ( $CM_e_3$ ); 123.8, 124.6 (C(1), C(4)); 131.8, 132.4 (C(2), C(3)); 141.7, 143.3 (C(4a), C(9a)); 155.2 (NCO).

*trans*-**11e**: cream crystals, mp 175–180 °C. IR (KBr): 3360, 2980, 1677, 1521, 1368, 1324, 1173, 1045  $cm^{-1}$ . <sup>1</sup>H NMR ( $CDCl_3$ , 400 MHz): 7.29, 7.28 (2 d, 2H,  $J = 8.4$  Hz, H-2, H-3); 4.54 (br s, 1H, NH); 3.70 (br s, 1H, H-7); 3.64 (dd, 1H, Ha-5); 3.42 (ddd, 1H, Ha-9); 3.34 (ddd, 1H, H-6); 3.02 (dd, 1H, Hb-5); 2.80 (ddd, 1H, Hb-9); 2.22 (ddd, 1H, Ha-8); 1.27 (dq, 1H, Hb-8); 1.45 (s, 9H,  $CM_e_3$ );  $J(5a,5b) = 14.5$ ,  $J(5a,6) = 1.8$ ,  $J(5b,6) = 10.2$ ,  $J(6,7) = 8.6$ ,  $J(6,8a) = 1.0$ ,  $J(7,8a) = 4.5$ ,  $J(7,8b) = 11.0$ ,  $J(8a,8b) = 13.7$ ,  $J(8a,9a) = 8.0$ ,  $J(8a,9b) = 1.8$ ,  $J(8b,9a) = 1.0$ ,  $J(8b,9b) = 11.5$ ,  $J(9a,9b) = 14.8$  Hz. <sup>13</sup>C NMR ( $CDCl_3$ , 100 MHz): 28.3 ( $CM_e_3$ ); 30.7 (C(9)); 31.3 (C(8)); 40.4 (C(5)); 59.7 (C(7)); 73.5 (C(6)); 80.4 ( $CM_e_3$ ); 122.9, 124.5 (C(4), C(1)); 131.9, 132.1 (C(2), C(3)); 138.2, 143.2 (C(4a), C(9a)); 156.8 (NCO). HR-MS (ESI-Q-ToF) calcd for  $C_{16}H_{21}Br_2NNaO_3$  [M+Na]<sup>+</sup>: 457.9761; found: 457.9753.

**10.5. 9-tert-Butoxycarbonylamino-8,9,10,11-tetrahydro-7-H-cyclohepta[a]naphthalen-10-ol (cis/trans mixture, 11h) and 9-tert-butoxycarbonylamino-8,9,10,11-tetrahydro-7-H-cyclohepta[a]naphthalen-8-ol (cis/trans mixture, 11h')**

*General procedure (g)* with **10h/10h'** (0.9 g, 3.02 mmol) in 2 M NH<sub>3</sub> solution in EtOH (15 mL) and with Ti(OiPr)<sub>4</sub> (1.8 mL, 6.03 mmol, 2 equiv) for 6 h, then reduction with NaBH<sub>4</sub> (171 mg, 4.5 mmol, 1.2 equiv) for 2 h to give the crude amine (560 mg, 81%). Acylation of the crude amine (1.2 g, 5.15 mmol) with Boc<sub>2</sub>O ((1.7 g, 7.73 mmol, 1.5 equiv), NaHCO<sub>3</sub> (562 mg, 6.69 mmol, 1.3 equiv) in MeOH (20 mL) for 16 h gave the 50:50 mixture of **11h/11h'** (0.70 g, 41%) which was resolved by chromatography (AcOEt/cyclohexane/Et<sub>2</sub>O) in the order *cis*-**11h**, *cis*-**11h'**, *trans*-**11h**, *trans*-**11h'**.

*cis*-**11h**: obtained impur and characterised by NMR only. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 8.17 (d, 1H, Har-1); 7.83 (d, 1H, Har-4); 7.72 (d, 1H, Har-5); 7.53 (td, 1H, Har-2); 7.44 (dt, 1H, Har-3); 7.30 (dd, 1H, Har-6); 5.07 (d, 1H, J = 8.3 Hz, NH), 4.27 (dt, 1H, J = 1.6, 8.4 Hz, 1H); 3.95 (m, 2 H); 3.14 (d, 1H, J = 14.9 Hz, Hb-11); 2.95 (m, 2H, CH<sub>2</sub>(7)); 2.08 (m, 1H, Ha-8); 1.46 (m, 1H, Hb-8); 1.42 (s, 9H, CMe<sub>3</sub>); J(1,2) = 8.6, J(1,3) = 1.4, J(2,3) = 6.8, J(2,4) = 1.2, J(3,4) = 8.0, J(5,6) = 8.2 Hz.

*cis*-**11h'**: colorless crystals, mp 195–196 °C. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 8.12 (d, 1H, Har-1); 7.85 (d, 1H, Har-4); 7.70 (d, 1H, Har-5); 7.51 (dt, 1H, Har-2); 7.42 (dt, 1H, Har-3); 7.70 (d, 1H, Har-6); 5.06 (d, 1H, J = 8.5 Hz, NH), 4.17 (br t, 1H, J = 8.0 Hz, H-8); 3.93 (br s, 1H, H-9); 3.64 (dd, 1H, J = 8.0, 14.8 Hz, Ha-11); 3.34 (d, 1H, J = 14.5 Hz, Ha-7); 3.24 (dd, 1H, J = 7.4, 14.5 Hz, Hb-7); 2.78 (dd, 1H, J = 11.4, 14.8 Hz, Hb-11); 2.15 (ddd, 1H, J = 4.2, 8.0, 13.4 Hz, Ha-10); 1.45 (m, 1H, Hb-10); 1.46 (s, 9H, CMe<sub>3</sub>); J(1,2) = 8.6, J(1,3) = 1.2, J(2,3) = 6.8, J(2,4) = 1.4, J(3,4) = 8.0, J(5,6) = 8.1 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 155.4 (NCO); 138.4 (C(6a)); 133.3, 131.5, 131.2 (C(4a),C(11a),C(11b)); 130.0, 128.7, 126.6 (C(4),C(5),C(6)); 126.3 (C(2)); 125.2 (C(3)); 123.1 (C(1)); 79.5 (CMe<sub>3</sub>); 69.0 (C(8)); 56.9 (C(9)); 39.6 (C(7)); 28.5 (CMe<sub>3</sub>); 27.8 (C(10)); 24.4 (C(11)). HRMS (ESI<sup>+</sup>) calcd for C<sub>20</sub>H<sub>25</sub>NNaO<sub>3</sub><sup>+</sup> [M+Na]<sup>+</sup>: 350.1727; found: 350.1720.

*trans*-**11h**: colorless crystals, mp 178–179 °C. IR (KBr): 3367, 2979, 2932, 1681, 1522, 1370, 1316, 1245, 1172, 1001, 743 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 8.23 (d, 1H, Har-1); 7.82 (d, 1H, Har-4); 7.67 (d, 1H, Har-5); 7.52 (td, 1H, Har-2); 7.46 (t, 1H, Har-3); 7.25 (d, 1H, Har-6); 4.53 (d, 1H, NH), 3.81 (d, 2H, Ha-11, H-9); 3.37 (m, 2H, H-10, OH); 3.10 (dd, 1H, Hb-11); 3.07 (dt, 1H, Ha-7); 2.87 (ddd, 1H, Hb-7); 2.25 (m, 1H, Ha-C(8)); 1.47 (s, 9H, CMe<sub>3</sub>); 1.33 (dq, 1H, Hb-8); J(1,2) = 8.6, J(1,3) = 1.0, J(2,3) = 6.8, J(2,4) = 1.4, J(3,4) = 8.0, J(5,6) = 8.3, J(7a,7b) = 14.9, J(7a,8a) = 1.6, J(7a,8b) = 11.6, J(7b,8a) = 7.5, J(7b,8b) = 1.6, J(8a,8b) = 13.5, J(8a,9) = 4.2, J(8b,9) = 11.6, J(9,NH) = 8.5, J(9,10) not determined, J(10,11a) = 1.6, J(10,11b) = 10.2, J(11a,11b) = 14.8 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 157.0 (NCO); 140.0 (C(6a)); 132.8, 132.1, 131.9 (C(4a),C(11a),C(11b)); 128.6, 127.5, 127.0 (C(4),C(5),C(6)); 126.3 (C(2)); 124.9 (C(3)); 123.3 (C(1)); 80.3 (CMe<sub>3</sub>); 74.1 (C(10)); 60.3 (C(9)); 34.1 (C(11)); 32.9, 32.2 (C(7), C(8)); 28.4 (CMe<sub>3</sub>). HRMS (ESI<sup>+</sup>) calcd for C<sub>20</sub>H<sub>25</sub>NNaO<sub>3</sub><sup>+</sup> [M+Na]<sup>+</sup>: 350.1727; found: 350.1720.

*trans*-**11h'**: colorless crystals, mp 174–175 °C. IR (KBr): 740, 815, 1007, 1172, 1244, 1317, 1366, 1523, 1682, 2930, 2982, 3357 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 8.08 (d, 1H, Har-1); 7.83 (dd, 1H, Har-4); 7.68 (d, 1H, Har-5); 7.50 (td, 1H, Har-2); 7.44 (dt, 1H, Har-3); 7.36 (d, 1H, Har-6); 4.55 (br d, 1H, NH), 3.77 (br q, 1H, H-9); 3.62 (dd, 1H, Ha-11); 3.39 (dt, 1H, H-8); 3.28 (dd, 1H, Ha-7, OH-8); 3.13 (dd, 1H, Hb-7); 2.87 (dd, 1H, Hb-11); 2.33 (m, 1H, Ha-10); 1.47 (s, 1H, CMe<sub>3</sub>); 1.34 (q, 1H, Hb-10); J(1,2) = 8.4, J(1,3) = 1.0, J(2,3) = 6.8, J(2,4) = 1.4, J(3,4) = 8.0, J(5,6) = 8.2, J(9,NH) = ca. 8.5, J(7a,7b) = 14.0, J(7a,8) = 10.2, J(7b,8) = 2.0, J(8,9) = 9.0, J(9,10a) = 4.4,

J(9,10b) = 11.4, J(10a,10b) = 13.4, J(10a,11a) = 8.0, J(10a,11b) = 1.2, J(10b,11a) = 1.0, J(10b,11b) = 11.2, J(11a,11b) = 15.0 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 156.9 (NCO); 137.6 (C(6a)); 134.2, 132.9, 131.0 (C(4a),C(11a),C(11b)); 128.8, 128.7 (C(4), C(6)); 126.7, 126.3 (C(5),C(2)); 125.0 (C(3)); 122.8 (C(1)); 80.3 (CMe<sub>3</sub>); 74.4 (C(8)); 60.1 (C(9)); 42.0 (C(7)); 32.1 (C(10)); 28.3 (CMe<sub>3</sub>); 23.9 (C(11)). HRMS (ESI<sup>+</sup>) calcd for C<sub>20</sub>H<sub>25</sub>NNaO<sub>3</sub><sup>+</sup> [M+Na]<sup>+</sup>: 350.1727; found: 350.1718.

**11. Preparation and reduction of the oximes 13c,c',d,d'**

**11.1. 4-Bromo-6-hydroxy-5,6,8,9-tetrahydrobenzocyclohepten-7-hydroxyimine (13c) and 1-bromo-6-hydroxy-5,6,8,9-tetrahydrobenzocyclohepten-8-hydroxyimine (13c')**

A solution of **10c/10c'** (780 mg, 2.38 mmol) in pyridine (4 mL) with NH<sub>2</sub>OH·HCl (199 mg, 2.86 mmol, 1.2 equiv) was stirred under Ar for 5 h at rt. The solvent was evaporated and the residue dissolved in MeCN (1–2 mL) and AcOEt (10 mL) with NBu<sub>4</sub>NF·3H<sub>2</sub>O (0.15 g, 0.5 mmol, 0.2 equiv). The solution was stirred for 1 h at rt, then washed with brine (10 mL), dried (MgSO<sub>4</sub>) and evaporated. The isomer **13c'** crystallised partially by triturating in AcOEt/cyclohexane and was isolated by filtration and washing with MeCN then with Et<sub>2</sub>O. A chromatography of the mother liquor cyclohexane/AcOEt 6:4 (**13c** R<sub>f</sub> = 0.43, **13c'** R<sub>f</sub> = 0.31) gave **13c** (210 mg, 33%) and the remaining **13c'**. Global yield of **13c'**: 230 mg, 35%.

Compound **13c**: colorless crystals, mp 142–143 °C. IR (KBr): 3530, 3441, 3209, 2882, 1448, 1062, 943, 922, 841, 782, 705 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 7.43 (dd, 1H, H-1); 7.14 (dd, 1H, H-3); 7.02 (t, 1H, H-2); 4.39 (dd, 1H, H-6); 3.39 (dd, 1H, Ha-5); 3.34 (dd, 1H, Hb-5); 3.20 (dddd, 1H, Ha-8); 3.04 (ddd, 1H, Ha-9); 2.83 (ddd, 1H, Hb-9); 2.45 (ddd, 1H, Hb-8); J(1,2) = 7.4, <sup>4</sup>J(1,3) = 1.3, J(2,3) = 8.0, J(5a,5b) = 14.3, J(5a,6) = 8.2, J(5b,6) = 4.5, J(8a,8b) = 14.3, J(8a,9a) = 8.2, J(8a,9b) = 4.8, J(8b,9a) = 5.0, J(8b,9b) = 7.9, J(9a,9b) = 14.4 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 161.6 (C=N); 144.9, 137.1 (C(4a),C(9a)); 132.2 (C(3)); 129.4 (C(1)); 128.9 (C(2)); 126.7 (C(4)); 71.4 (C(6)); 40.0 (C(5)); 32.7 (C(9)); 22.9 (C(8)). HR-MS (ESI<sup>+</sup>) calcd for C<sub>11</sub>H<sub>13</sub>BrNO<sub>2</sub> [M+H]<sup>+</sup>: 270.0124 and 272.0103; found: 270.0122 and 272.0105.

Compound **13c'**: colorless crystals, mp 176–177 °C. IR (KBr): 3496, 3213, 2919, 1466, 1447, 1057, 1010, 950, 904, 777, 737 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 7.45 (dd, 1H, H-2); 7.16 (dd, 1H, H-4); 7.01 (t, 1H, H-3); 4.41 (m, 1H, H-6); 3.33 (ddd, 1H, Ha-9); 3.24 (ddd, 1H, Ha-8); 3.15 (d, 2H, CH<sub>2</sub>(5)); 2.93 (ddd, 1H, Hb-9); 2.31 (ddd, 1H, Hb-8); J(2,3) = 8.0, <sup>4</sup>J(2,4) = 1.2, J(3,4) = 7.8, J(6,8) = 1.0, J(8a,8b) = 14.0, J(8a,9a) = 8.8, J(8a,9b) = 3.3, J(8b,9a) = 3.2, J(8b,9b) = 8.8, J(9a,9b) = 14.0 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 161.8 (C=N), 141.5, 140.2 (C(9a),C(4a)), 132.4 (C(2)), 131.4 (C(4)), 128.7 (C(3)), 125.0 (C(1)), 71.8 (C(6)), 42.5 (C(5)), 30.6 (C(9)), 22.1 (C(8)). HR-MS (ESI<sup>+</sup>) calcd for C<sub>11</sub>H<sub>13</sub>BrNO<sub>2</sub> [M+H]<sup>+</sup>: 270.0124 and 272.0103; found: 270.0122 and 272.0105. Anal. Calcd for C<sub>11</sub>H<sub>12</sub>BrNO<sub>2</sub> (270.12): C, 48.91; H, 4.48; N, 5.19. Found: C, 48.9; H, 4.5; N, 5.2.

**11.2. 4-Bromo-6-hydroxy-1-phenyl-5,6,8,9-tetrahydrobenzocyclohepten-7-hydroxyimine (13d) and 1-bromo-6-hydroxy-4-phenyl-5,6,8,9-tetrahydrobenzocyclohepten-8-hydroxyimine (13d')**

A solution of **10d/10d'** (8.6 g, 21 mmol) in pyridine (86 mL) was stirred with NH<sub>2</sub>OH·HCl (1.8 g, 25 mmol, 1.2 equiv) at rt for 16 h. Same work-up as for **13c/13c'**. The crude oxime mixture (7.4 g, quant.) was purified by flash chromatography (cyclohexane/AcOEt 6:4) to give a 60:40 mixture of **13d/13d'** (90:10 E/Z mixtures, 5.4 g, 73%).

Compound **13d/13d'**: colorless crystals, mp 50–52 °C. IR (KBr): 704, 770, 821, 924, 946, 1028, 1064, 1454, 1705, 2918, 3058, 3327 cm<sup>-1</sup>. HR-MS (ESI<sup>+</sup>) calcd for C<sub>17</sub>H<sub>17</sub>BrNO<sub>2</sub> [M+H]<sup>+</sup>: 346.0437 and 348.0418; found: 346.0424 and 348.0400.

Compound **13d**: <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz) major oxime *E*: 7.30–7.45 (m, 4Har); 7.24 (m, 2Har); 6.95 (d, *J* = 8.4 Hz, 1Har); 4.44 (dd, 1H, H-6); 3.52 (dd, 1H, Ha-5); 3.34 (dd, 1H, Hb-5); 2.67–2.97 (m, 4H); *J*(5a,5b) = 14.4, *J*(5a,6) = 4.0, *J*(5b,6) = 8.8 Hz. Oxime *Z*, partial data: 5.46 (dd, 1H, *J* = 3.9, 8.9 Hz, H-6). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz) major oxime *E*: 161.2 (C(7)); 141.1 (C(9a)); 140.9 (Car-s); 140.5 (C(1)); 135.7 (C(4a)); 130.4 (C(3)); 129.8 (C(2)); 129.0 (CHar-m); 128.2 (CHar-o); 127.2 (CHar-p); 124.8 (C(4)); 70.0 (C(6)); 38.9 (C(5)); 26.9 (C(9)); 22.8 (C(8)).

Compound **13d'**: <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz) major oxime *E*: 7.30–7.45 (m, 4Har); 7.24 (m, 2Har); 6.95 (d, *J* = 8.4 Hz, 1Har); 4.32 (dd, 1H, H-6); 3.16 (m, 2H, CH<sub>2</sub>(9)); 3.10 (dd, 1H, Ha-5); 3.04 (dd, 1H, Hb-5); 2.67–2.97 (m, 2H, CH<sub>2</sub>(8)); *J*(5a,5b) = 14.4, *J*(5a,6) = 4.4, *J*(5b,6) = 8.4 Hz. Oxime *Z*, partial data: 5.35 (dd, 1H, *J* = 4.2, 8.5 Hz, H-6). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz) major oxime *E*: 161.6 (C(7)); 142.6 (C(4a)); 140.8 (Car-s); 140.5 (C(4)); 135.5 (C(9a)); 130.7 (C(2)); 129.6 (C(3)); 129.2 (Car-m); 128.1 (Car-o); 127.2 (Car-p); 123.2 (C(1)); 70.8 (C(6)); 36.6 (C(5)); 29.4 (C(9)); 22.2 (C(8)).

### 11.3. Reduction of **13c**

A solution of **13c** (1.01 g, 3.75 mmol) in EtOH (30 mL) and concentrated aqueous NH<sub>4</sub>OH solution (12 mL, 180 mmol, 50 equiv) was hydrogenolysed at rt over wet Raney-nickel (1.8–2.0 g) at rt for 30–50 min with NMR monitoring. When the reduction was complete, the catalyst was discarded by centrifugation or filtration over Celite. The solution was evaporated to give crude amine (0.95 g, ca. 95%) which was directly N-protected. A solution of the crude amine in MeOH (10 mL) was stirred with Boc<sub>2</sub>O (1.22 g, 5.65 mmol, 1.5 equiv) and Na<sub>2</sub>CO<sub>3</sub> (0.5 g, 4.81 mmol, 1.3 equiv) for 16 h at rt. AcOEt was added (50 mL) and the solution washed with brine, dried (MgSO<sub>4</sub>) and evaporated to give **11c** (1.12 g, 85%) as 50:50 *cis/trans* isomeric mixture.

### 11.4. Reduction of **13c'**

Same procedure as for **13c** with **13c'** (0.47 g, 1.75 mmol) in EtOH (10 mL), wet Raney-nickel (0.85 g) and aqueous concentrated NH<sub>4</sub>OH solution (6 mL, 90 mmol, 50 equiv). Same work-up and N-protection with Boc<sub>2</sub>O (0.57 g, 2.6 mmol, 1.5 equiv) and Na<sub>2</sub>CO<sub>3</sub> (0.24 g, 2.26 mmol, 1.3 equiv) gave **11c'** (0.61 g, 98%) as 50:50 *cis/trans* isomeric mixture.

### 11.5. Reduction of **13d,d'**

Same procedure as for **13c** with the mixture **13d,d'** (468 mg, 1.35 mmol) in EtOH (19 mL), aqueous concentrated NH<sub>4</sub>OH solution (5.6 mL, 79 equiv) and wet Raney-nickel (0.9 g) for 1 h at rt. Same work-up and N-protection with Boc<sub>2</sub>O (0.46 g, 2.1 mmol, 1.5 equiv) and NaHCO<sub>3</sub> (0.15 g, 1.8 mmol, 1.3 equiv) in MeOH (7 mL) gave the 60:40 isomeric mixture **11d/11d'** (508 mg, 87%), mp 174–176 °C.

### 12. Keto-amides **12c–e,b'–d',h,h'**

*General procedure (h)*: To a solution of **11** or **11'** (10 mmol) in wet CH<sub>2</sub>Cl<sub>2</sub> (50 mL) was added Dess–Martin periodinane (DMP) (6–9 g, 15–20 mmol, 1.5–2 equiv) and the mixture stirred at rt for 2 h (tlc monitoring). After dilution with Et<sub>2</sub>O (100 mL), the solution was vigorously stirred with aqueous 1 M NaHCO<sub>3</sub> solution

(50 mL, containing Na<sub>2</sub>SO<sub>3</sub> (2.5 g, 20 mmol), or Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub>·5H<sub>2</sub>O (15 g, 60 mmol)) for 20 min, the organic phase was washed with aqueous 1 M NaHCO<sub>3</sub> solution (50 mL) and then with brine (50 mL), dried over MgSO<sub>4</sub> and evaporated to give the crude amide **12** which was crystallised from iPr<sub>2</sub>O.

#### 12.1. 7-*tert*-Butoxycarbonylamino-1-chloro-5,7,8,9-tetrahydro-benzocyclohepten-6-one (**12b**)

*General procedure (h)* with *cis/trans*-**11b'** (0.25 g, 0.80 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (6 mL) and with DMP (510 mg, 1.2 mmol, 1.5 equiv) for 3 h to give **12b'** (0.23 g, 92%).

Compound **12b'**: colorless crystals, mp 132–134 °C. IR (KBr): 2963, 2924, 1724, 1489, 1447, 1447, 1147, 1046, 989, 782 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 300 MHz): 7.31 (dd, 1H, *J* = 1.8, 7.5 Hz, H-2); 7.10 (m, 2H, H-3,H-4); 5.36 (d, 1H, *J* = 7.5 Hz, NH), 4.49 (dt, 1H, *J* = 11.1, 7.5 Hz, H-7), 3.87 (d, 1H, *J* = 16.0 Hz, Ha-5); 3.67 (d, 1H, *J* = 16.0 Hz, Hb-5); 3.14 (m, 2H, CH<sub>2</sub>(9)); 2.61 (m, 1H, Ha-8); 1.48 (m, 1H, Hb-8); 1.41 (s, 9H, CMe<sub>3</sub>). <sup>13</sup>C NMR (CDCl<sub>3</sub>, 75 MHz): 205.0 (CO(6)); 155.0 (NCO); 137.3 (C(9a)); 134.7, 133.8 (C(4a), C(1)); 129.0, 128.3, 128.0 (C(2), C(3), C(4)); 79.9 (CMe<sub>3</sub>); 59.7 (C(7)); 48.2 (C(5)); 32.9 (C(8)); 28.3 (CMe<sub>3</sub>); 26.0 (C(9)). HR-MS (ESI-Q-Tof) calcd for C<sub>16</sub>H<sub>22</sub>ClLiNO<sub>3</sub> [M+Li]<sup>+</sup>: 318.1443; found: 318.1381; calcd for C<sub>16</sub>H<sub>22</sub>ClNaNO<sub>3</sub> [M+Na]<sup>+</sup>: 334.1180; found: 334.1124.

#### 12.2. 4-Bromo-7-*tert*-Butoxycarbonylamino-5,7,8,9-tetrahydro-benzocyclohepten-6-one (**12c**)

*General procedure (h)* with *cis/trans*-**11c** (1.10 g, 3.08 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (10 mL) and with DMP (2.61 g, 6.16 mmol, 2 equiv) for 3 h to give **12c** (816 mg, 75%).

Compound **12c**: colorless crystals, mp 164–166 °C. IR (KBr): 3297, 2978, 1724, 1682, 1542, 1442, 1365, 1298, 1276, 1255, 1187, 1171, 1091, 1056, 1010, 782 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 7.49 (d, 1H, H-3); 7.12 (d, 1H, H-1); 7.06 (t, 1H, H-2); 5.41 (d, 1H, NH); 4.49 (m, 1H, H-7); 4.15 (d, 1H, Ha-5); 3.99 (d, 1H, Hb-5); 2.97 (m, 1H, Ha-9); 2.89 (ddd, 1H, Hb-9); 2.62 (m, 1H, Ha-8); 1.53 (m, 1H, Hb-8); 1.42 (s, 9H, CMe<sub>3</sub>); *J*(1,2) = 7.8, *J*(1,3) = 1.2, *J*(2,3) = 7.4, *J*(5a,5b) = 16.5, *J*(NH,7) = 7.0, *J*(7,8a) = 7.6, *J*(7,8b) = 10.8, *J*(8a,8b) = 12.6, *J*(8a,9a) = 4.8, *J*(8a,9b) = 10.2, *J*(8b,9a) = 6.4, *J*(8b,9b) = 4.6, *J*(9a,9b) = 14.6 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 205.2 (C(6)); 153.3 (NCO); 142.4 (C(9a)); 137.2 (C(4a)); 132.2 (C(3)); 129.5, 128.9 (C(1),C(2)); 125.4 (C(4)); 80.3 (CMe<sub>3</sub>); 59.7 (C(7)); 47.0 (C(5)); 34.5 (C(8)); 31.9 (C(9)); 28.7 (CMe<sub>3</sub>). HR-MS (ESI-Q-Tof) calcd for C<sub>16</sub>H<sub>20</sub>BrLiNO<sub>3</sub> (M+Li)<sup>+</sup>: 360.0781 and 362.0763; found: 360.0740 and 362.0736.

#### 12.3. 1-Bromo-7-*tert*-Butoxycarbonylamino-5,7,8,9-tetrahydro-benzocyclohepten-6-one (**12c'**)

*General procedure (h)* with *cis/trans*-**11c'** (250 mg, 0.7 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (4 mL) and with DMP (0.60 g, 1.4 mmol, 2 equiv) for 3 h to give **12c'** (164 mg, 66%).

Compound **12c'**: colorless crystals, mp 114–116 °C. IR (KBr): 3422, 2969, 2928, 1684, 1654, 1446, 1264, 1166, 1113 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 7.50 (d, 1H, H-2); 7.12 (d, 1H, H-4); 7.04 (dd, 1H, H-3); 5.35 (br d, 1H, NH); 4.49 (dt, 1H, H-7); 3.88 (d, 1H, Ha-5); 3.70 (d, 1H, Hb-5); 3.21 (ddd, 1H, Ha-9); 3.15 (ddd, 1H, Hb-9); 2.62 (m, 1H, Ha-8); 1.50 (m, 1H, Hb-8); 1.42 (s, 9H, CMe<sub>3</sub>); *J*(2,3) = 8.0, *J*(2,4) = 1.2, *J*(3,4) = 7.4, *J*(5a,5b) = 16.1, *J*(NH,7) = 7.4, *J*(7,8a) = 7.6, *J*(7,8b) = 11.0, *J*(8a,8b) = 13.2, *J*(8a,9a) = 4.4, *J*(8a,9b) = 10.4, *J*(8b,9a) = 6.4, *J*(8b,9b) = 4.2, *J*(9a,9b) = 14.8 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 75 MHz): 205.2 (CO(6)); 155.1 (NCO); 139.1 (C(9a)); 134.8 (C(4a)); 132.5 (C(2)); 129.2 (C(4)); 128.5 (C(3)); 124.7 (C(1)); 80.0 (CMe<sub>3</sub>); 59.6 (C(7)); 48.6 (C(5)); 32.9 (C(8)); 29.4 (C(9)); 28.4

(CMe<sub>3</sub>). HR-MS (ESI-Q-ToF) calcd for C<sub>16</sub>H<sub>20</sub>BrLiNO<sub>3</sub> [M+Li]<sup>+</sup>: 360.0781 and 362.0763; found: 360.0781 and 362.0767.

**12.4. 4-Bromo-7-(tert-butoxycarbonylamino)-1-phenyl-5,7,8,9-tetrahydrobenzocyclohepten-6-one (12d) and 1-bromo-7-(tert-butoxycarbonylamino)-4-phenyl-5,7,8,9-tetrahydrobenzocyclohepten-6-one (12d')**

*General procedure (h)* with the isomeric mixture **11d/11d'** (364 mg, 0.84 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (5 mL) and with DMP (1.07 g, 2.53 mmol, 3 equiv) for 2 h to give **12d/12d'** (198 mg, 54%) which were separated by par HPLC (MeOH/H<sub>2</sub>O 7:3) to give **12d** (137 mg, 38%) and **12d'** (61 mg, 17%).

Compound **12d**: colorless crystals, mp 148–154 °C. IR (KBr): 3332, 2984, 2930, 1725, 1678, 1529, 1451, 1371, 1333, 1300, 1274, 1252, 1170, 1051, 1008, 764, 701 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 7.52 (d, 1H, H-3); 7.44–7.35 (m, 3Har); 7.27–7.24 (m, 2Har); 7.05 (d, 1H, H-2); 5.43 (d, 1H, NH); 4.54 (m, 1H, H-7); 4.30 (d, 1H, Ha-5); 3.97 (d, 1H, Hb-5); 2.88 (dt, 1H, Ha-9); 2.77 (ddd, 1H, Hb-9); 2.51 (m, 1H, Ha-8); 1.48 (m, 1H, Hb-8); 1.42 (s, 9H, CMe<sub>3</sub>); J(5a,5b) = 17.6, J(2,3) = 8.3; J(7,NH) = 7.0; J(7,8a) = 8.0, J(7,8b) = 11.0, J(8a,8b) = 13.2, J(8a,9a) = 5.0; J(8a,9b) = 11.2; J(8b,9a) = 5.2; J(8b,9b) = 5.2; J(9a,9b) = 14.7 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 205.2 (C(6)); 154.9 (NCO); 141.7 (C(9a)); 140.4, 139.1 (C(1),Car-s); 132.6 (C(4a)); 131.0 (C(3)); 130.7 (C(2)); 129.0 (Car-o); 128.3 (Car-m); 127.4 (Car-p); 124.1 (C(4)); 79.8 (CMe<sub>3</sub>); 58.5 (C(7)); 47.1 (C(5)); 33.8 (C(8)); 28.3 (CMe<sub>3</sub>); 27.2 (C(9)). Anal. Calcd for C<sub>22</sub>H<sub>24</sub>BrNO<sub>3</sub> (430.33): C, 61.40; H, 5.62; N, 3.25; Br, 18.57. Found: C, 61.2; H, 5.4; N, 3.1; Br, 18.5.

Compound **12d'**: colorless crystals, mp 144–150 °C. IR (KBr): 3301, 2977, 2934, 1725, 1704, 1702, 1675, 1542, 1453, 1366, 1183, 1170, 703. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 7.54 (d, 1H, H-2); 7.45–7.34 (m, 3Har); 7.29 (d, 2Har, J = 7.2 Hz); 7.05 (d, 1H, H-3); 5.37 (d, 1H, NH); 4.48 (m, 1H, H-7); 3.85 (d, 1H, Ha-5); 3.72 (d, 1H, Hb-5); 3.30 (ddd, 1H, Ha-9); 3.17 (ddd, 1H, Hb-9); 2.68 (m, 1H, Ha-8); 1.54 (m, 1H, Hb-8); 1.43 (s, 9H, CMe<sub>3</sub>); J(2,3) = 8.2 Hz, J(5a,5b) = 17.0, J(7,NH) = 7.2, J(7,8a) = 7.9, J(7,8b) = 10.8, J(8a,8b) = 13.0, J(8a,9a) = 4.5, J(8a,9b) = 11.0, J(8b,9a) = 6.0, J(8b,9b) = 4.4, J(9a,9b) = 14.7 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 205.6 (C(6)); 154.8 (NCO); 142.0 (C(9a)); 139.8, 139.1 (C(4),Car-s); 132.3 (C(4a)); 131.6, 130.1, 129.5, 128.4, 127.5 (C(2),C(3), 3Car); 123.6 (C(1)); 79.8 (CMe<sub>3</sub>); 59.1 (C(7)); 44.6 (C(5)); 32.5 (C(8)); 29.8 (C(9)); 28.3 (CMe<sub>3</sub>). HR-MS (ESI<sup>+</sup>) calcd for C<sub>22</sub>H<sub>24</sub>BrNO<sub>3</sub>Li [M+Li]<sup>+</sup>: 436.1095 and 438.1077; found: 436.1086 and 438.1065.

**12.5. 1,4-Dibromo-7-(tert-butoxycarbonylamino)-5,7,8,9-tetrahydrobenzocyclohepten-6-one (12e)**

*General procedure (h)* with *cis/trans*-**11e** (901 mg, 2.07 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (20 mL) and with DMP (1.14 g, 2.69 mmol, 1.3 equiv) for 2 h to give **12e** (710 mg, 79%).

Compound **12e**: yellowish crystals, mp 185–186 °C. IR (KBr): 3277, 2974, 1731, 1675, 1540, 1439, 1365, 1272, 1164, 1052, 978, 808 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 7.38, 7.36 (2 d, 2H, H-2, H-3); 5.36 (d, 1H, NH); 4.41 (q, 1H, H-7); 4.27 (d, 1H, Ha-5); 3.92 (d, 1H, Hb-5); 3.32 (ddd, 1H, Ha-9); 3.02 (ddd, 1H, Hb-9); 2.62 (m, 1H, Ha-8); 1.54 (m, 1H, Hb-8); 1.42 (s, 9H, CMe<sub>3</sub>); J(2,3) = 8.4, J(5a,5b) = 18.0, J(7,NH) = 7.8, J(7,8a) = 7.8, J(7,8b) = 10.4, J(8a,8b) = 12.2, J(8a,9a) = 5.0, J(8a,9b) = 12.2, J(8b,9a) = 4.8, J(8b,9b) = 5.0, J(9a,9b) = 14.7 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 204.9 (C(6)); 154.9 (NCO); 140.6 (C(9a)); 134.3 (C(4a)); 133.2, 132.5 (C(2),C(3)); 124.1, 123.5 (C(1),C(4)); 80.0 (CMe<sub>3</sub>); 58.1 (C(7)); 47.6 (C(5)); 32.2 (C(8)); 30.4 (C(9)); 28.3

(CMe<sub>3</sub>). HR-MS (ESI-Q-ToF) calcd for C<sub>16</sub>H<sub>19</sub>Br<sub>2</sub>NNaO<sub>3</sub> [M+Na]<sup>+</sup>: 455.9604; found: 455.9610.

**12.6. 9-tert-Butoxycarbonylamino-7,8,9,11-tetrahydro-cyclohepta[a]naphthalen-10-one (12h)**

*General procedure (h)* with *cis-* or *trans*-**11h** (50 mg, 0.15 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (5 mL) and with DMP (97 mg, 0.23 mmol, 1.5 equiv) for 3 h to give **12h** (50 mg, quant.).

Compound **12h**: colorless crystals, mp 152–153 °C. IR (KBr): 3352, 2971, 2931, 1720, 1682, 1514, 1367, 1247, 1163, 1058, 982, 819, 743 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 8.12 (d, 1H, H-1); 7.84 (d, 1H, H-4); 7.74 (d, 1H, H-5); 7.55 (dt, 1H, H-2); 7.46 (dt, 1H, H-3); 7.31 (d, 1H, H-6); 5.43 (d, 1H, NH), 4.63 (dt, 1H, H-9); 4.27 (d, 1H, Ha-11); 4.23 (d, 1H, Hb-11); 3.20 (ddd, 1H, Ha-7); 3.05 (ddd, 1H, Hb-7); 2.73 (m, 1H, Ha-8); 1.58 (m, 1H, Hb-8); 1.42 (s, 9H, CMe<sub>3</sub>); J(1,2) = 8.4, J(1,3) = 1.2, J(2,3) = 6.8, J(2,4) = 1.4, J(3,4) = 8.0, J(5,6) = 8.2, J(9,NH) = 7.0, J(7a,7b) = 14.6, J(7a,8a) = 3.4, J(7a,8b) = 8.5, J(7b,8a) = 9.2, J(7b,8b) = 3.5, J(8a,9) = 7.2, J(8b,9) = 11.2, J(8a,8b) = 13.0, J(11a,11b) = 15.0 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 204.8 (CO(10)); 155.0 (NCO); 138.4 (C(6a)); 133.1, 131.5 (C(4a),C(11b)); 128.7 (C(4)); 128.1 (C(5)); 127.7 (C(6)); 127.6 (C(11a)); 126.8 (C(2)); 125.3 (C(3)); 123.2 (C(1)); 79.8 (CMe<sub>3</sub>); 60.8 (C(9)); 41.6 (C(11)); 35.4 (C(8)); 31.6 (C(7)); 28.3 (CMe<sub>3</sub>). HR-MS (ESI-Q-ToF) calcd for C<sub>20</sub>H<sub>23</sub>LiNO<sub>3</sub> [M+Li]<sup>+</sup>: 332.1833; found: 332.1813.

**12.7. 9-tert-Butoxycarbonylamino-7,8,9,11-tetrahydro-cyclohepta[a]naphthalen-8-one (12h')**

*General procedure (h)* with *cis-* or *trans*-**11h'** (90 mg, 0.27 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (6 mL) and with DMP (174 mg, 0.40 mmol, 1.5 equiv) for 3 h to give **12h'** (80 mg, 90%).

Compound **12h'**: colorless crystals, mp 128–129 °C. IR (KBr): 3352, 2971, 2931, 1720, 1682, 1514, 1367, 1246, 1163, 1057, 981, 818, 743 cm<sup>-1</sup>. <sup>1</sup>H NMR (CDCl<sub>3</sub>, 400 MHz): 8.09 (d, 1H, H-1); 7.85 (d, 1H, H-4); 7.72 (d, 1H, H-5); 7.53 (dt, 1H, H-2); 7.48 (dt, 1H, H-3); 7.31 (d, 1H, H-6); 5.39 (d, 1H, NH), 4.55 (dt, 1H, H-9); 4.03 (d, 1H, Ha-7); 3.88 (d, 1H, Hb-7); 3.43 (m, 1H, Ha-11); 3.35 (m, 1H, Hb-11); 2.75 (m, 1H, Ha-10); 1.64 (m, 1H, Hb-10); 1.40 (s, 9H, CMe<sub>3</sub>); J(1,2) = 8.7, J(1,3) = 1.2, J(2,3) = 6.8, J(2,4) = 1.2, J(3,4) = 8.1, J(5,6) = 8.3, J(7a,7b) = 16.3, J(9,NH) = 7.2, J(9,10a) = 7.6, J(9,10b) = 10.8, J(10a,10b) = 12.8, J(10a,11a) = 4.4, J(10a,11b) = 10.8, J(10b,11a) = 6.4, J(10b,11b) = 4.2, J(11a,11b) = 14.8 Hz. <sup>13</sup>C NMR (CDCl<sub>3</sub>, 100 MHz): 205.7 (CO(8)); 154.9 (NCO); 135.5 (C(6a)); 133.3, 131.3 (C(4a),C(11b)); 130.0 (C(11a)); 128.8 (C(4)); 127.7(C(5)); 127.4 (C(6)); 126.5 (C(2)); 125.5 (C(3)); 122.9 (C(1)); 79.7 (CMe<sub>3</sub>); 59.6 (C(9)); 48.5 (C(7)); 33.9 (C(10)); 28.3 (CMe<sub>3</sub>); 23.9 (C(11)). HR-MS (ESI-Q-ToF) calcd for C<sub>20</sub>H<sub>23</sub>NaNO<sub>3</sub> [M+Na]<sup>+</sup>: 348.1570; found: 348.1530.

**13. Keto-amines 1c-e, b'-d', h, h'**

*General procedure (i)* a solution of **12** or **12'** (1 mmol) in dry dioxane (1–2 mL) with 2 N HCl in dry Et<sub>2</sub>O (5 mL) was stirred at rt for 2–4 days. The amine hydrochloride **1** was isolated by filtration or centrifugation and washing with dry Et<sub>2</sub>O (1 mL).

**13.1. 7-Amino-1-chloro-5,7,8,9-tetrahydro-benzocyclohepten-6-one, hydrochloride (1b')**

*General procedure (i)* with **12b'** (80 mg, 0.26 mmol) in dioxane (3 mL) and 2 N HCl in Et<sub>2</sub>O (3 mL) for 2 d to give **1b'** (50 mg, 79%) after recrystallisation in MeOH/Et<sub>2</sub>O.

Compound **1b'**: colorless crystals, mp 240 °C (dec). IR (KBr): 3432, 2962, 2903, 1725, 1582, 1582, 1573, 1509, 1446, 1048,

989, 778  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CD}_3\text{OD}$ , 400 MHz): 7.39 (m, 1H, H-2); 7.22 (m, 1H, H-3, H-4); 4.33 (dd, 1H, H-7); 4.21 (d, 1H, Ha-5); 3.76 (d, 1H, Hb-5); 3.46 (ddd, 1H, Ha-9); 3.30 (ddd, 1H, Hb-9); 2.57 (m, 1H, Ha-8); 1.71 (m, 1H, Hb-8);  $J(5a,5b) = 15.0$ ,  $J(7,8a) = 7.4$ ,  $J(7,8b) = 11.8$ ,  $J(8a,8b) = 13.0$ ,  $J(8a,9a) = 9.3$ ,  $J(8a,9b) = 3.4$ ,  $J(8b,9a) = 3.4$ ,  $J(8b,9b) = 8.6$ ,  $J(9a,9b) = 15.0$  Hz.  $^{13}\text{C}$  NMR ( $\text{CD}_3\text{OD}$ , 100 MHz): 202.1 (CO(6)); 138.4 (C(9a)); 136.0 (C(4a)); 134.8 (C(1)); 130.2 (C(2)); 129.8, 129.7 (C(3), C(4)); 60.1 (C(7)); 48.2 (C(5)); 31.5 (C(8)); 26.1 (C(9)). HR-MS (ESI-Q-ToF) calcd for  $\text{C}_{11}\text{H}_{13}\text{ClNO}$   $[\text{M}+\text{H}]^+$ : 210.0680; found: 210.0690.

### 13.2. 7-Amino-4-bromo-5,7,8,9-tetrahydro-benzocyclohepten-6-one, hydrochloride (1c)

*General procedure (i)* with **12c** (164 mg, 0.46 mmol) in dioxane (1 mL) and 2 N HCl in  $\text{Et}_2\text{O}$  (5 mL) for 3 d to give **1c** (107 mg, 80%) after recrystallisation in *i*PrOH/ $\text{Et}_2\text{O}$ .

Compound **1c**: colorless crystals, mp 212 °C (sublimation). IR (KBr): 2955, 2935, 2204, 2123, 1722, 1444, 1081, 986, 779  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CD}_3\text{OD}$ , 400 MHz, 8:2 mixture with an hemi-acetal): 7.55 (d, 1H, H-3); 7.26 (d, 1H, H-1); 7.15 (t, 1H, H-2); 4.33 (dd, 1H, H-7); 4.28 (d, 1H, Ha-5); 4.20 (d, 1H, Hb-5); 3.09 3.27 (ddd, 1H, Ha-9); 3.09 (ddd, 1H, Hb-9); 2.55 (m, 1H, Ha-8); 1.74 (m, 1H, Hb-8);  $J(1,2) = 7.6$ ,  $J(2,3) = 8.0$ ,  $J(5a,5b) = 15.2$ ,  $J(7,8a) = 7.2$ ,  $J(7,8b) = 11.8$ ,  $J(8a,8b) = 12.8$ ,  $J(8a,9a) = 3.3$ ,  $J(8a,9b) = 8.9$ ,  $J(8b,9a) = 8.8$ ,  $J(8b,9b) = 3.6$ ,  $J(9a,9b) = 14.7$  Hz. Hemiactal, partial data: ca. 1.70 (m, Hb-8); 2.07 (m, Ha-8); 2.94 (m,  $\text{CH}_2(9)$ ); 7.07 (t, H-2); ca. 7.15 (d, H-1); 7.48 ((d, H-3);  $J(1,2) = 7.6$ ,  $J(2,3) = 8.0$  Hz.  $^{13}\text{C}$  NMR ( $\text{CD}_3\text{OD}$ , 100 MHz): 202.0 (CO(6)); 143.8 (C(9a)); 133.5 (C(3)); 133.1 (C(4a)); 130.8, 130.2 (C(1), C(2)); 125.8 (C(1)); 60.6 (C(7)); 46.5 (C(5)); 32.8 (C(8)), 31.8 (C(9)). HR-MS (ESI-Q-ToF) calcd for  $\text{C}_{11}\text{H}_{13}\text{BrNO}$   $[\text{M}+\text{H}]^+$ : 254.0175 and 256.0155; found: 254.0178 and 256.0156.

### 13.3. 7-Amino-1-bromo-5,7,8,9-tetrahydro-benzocyclohepten-6-one, hydrochloride (1c')

*General procedure (i)* with **12c'** (816 mg, 2.3 mmol) in dioxane (3 mL) and 2 N HCl in  $\text{Et}_2\text{O}$  (15 mL) for 3 d to give **1c'** (535 mg, 80%) after recrystallisation in *i*PrOH/ $\text{Et}_2\text{O}$ .

Compound **1c'**: colorless crystals, mp 260–270 °C (sublimation). IR (KBr): 2963, 2924, 1724, 1491, 1443, 1147, 1047, 986, 779  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CD}_3\text{OD}$ , 400 MHz, 8:2 mixture with an hemi-acetal): 7.57 (d, 1H, H-2); 7.25 (d, 1H, H-4); 7.14 (t, 1H, H-3); 4.30 (dd, 1H, H-7); 4.20 (d, 1H, Ha-5); 3.79 (d, 1H, Hb-5); 3.46 (ddd, 1H, Ha-9); 3.35 (ddd, 1H, Hb-9); 2.54 (m, 1H, Ha-8); 1.71 (m, 1H, Hb-8);  $J(2,3) = 8.1$ ,  $J(2,4) = 1.0$ ,  $J(3,4) = 7.6$ ,  $J(5a,5b) = 15.2$ ,  $J(7,8a) = 7.2$ ,  $J(7,8b) = 11.7$ ,  $J(8a,8b) = 13.0$ ,  $J(8a,9a) = 9.6$ ,  $J(8a,9b) = 3.5$ ,  $J(8b,9a) = 3.6$ ,  $J(8b,9b) = 8.2$ ,  $J(9a,9b) = 15.1$  Hz. Hemiactal, partial data: ca. 1.70 (m, 1H, Hb-8); 2.07 (m, 1H, Ha-8); 2.89 (m, 1H, Hb-9); 3.16 (d, 1H, Hb-5); 3.24 (d, 1H, Ha-5); 7.05 (t, 1H, H-3); 7.20 (d, 1H, H-4); 7.48 ((d, 1H, H-2);  $J(2,3) = 8.1$ ,  $J(3,4) = 7.4$ ,  $J(5a,5b) = 15.0$  Hz.  $^{13}\text{C}$  NMR ( $\text{CD}_3\text{OD}$ , 100 MHz): 202.2 (CO(6)); 140.1 (C(9a)); 136.0 (C(4a)); 133.6 (C(2)); 130.5, 130.1 (C(3), C(4)); 125.3 (C(1)); 60.2 (C(7)); 48.4 (C(5)); 31.3 (C(8)); 29.4 (C(9)). HR-MS (ESI-Q-ToF) calcd for  $\text{C}_{11}\text{H}_{13}\text{BrNO}$   $[\text{M}+\text{H}]^+$ : 254.0175 and 256.0155; found: 254.0151 and 256.0130.

### 13.4. 7-Amino-4-bromo-1-phenyl-5,7,8,9-tetrahydrobenzocyclohepten-6-one, hydrochloride (1d)

*General procedure (i)* with **12d** (70 mg, 0.16 mmol) in dioxane (2 mL) and 2 N HCl in  $\text{Et}_2\text{O}$  (4 mL) for 2 d to give **1d** (57.5 mg, 96%) after recrystallisation in *i*PrOH/ $\text{Et}_2\text{O}$ .

Compound **1d**: colorless crystals, mp >250 °C. IR (KBr): 3028, 2866, 1730, 1578, 1507, 1452, 772, 705  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CD}_3\text{OD}$ ,

400 MHz): 7.61 (d, 1H, H-3); 7.49–7.38 (m, 3 Har); 7.29–7.21 (m, 2 Har); 7.12 (d, 1H, H-2); 4.39 (d, 1H, Ha-5); 4.32 (dd, 1H, H-7); 4.25 (d, 1H, Hb-5); 3.06 (t, 2H,  $\text{CH}_2(9)$ ); 2.38 (m, 1H, Ha-8); 1.73 (m, 1H, Hb-8);  $J(2,3) = 8.2$ ,  $J(5a,5b) = 15.9$ ,  $J(7,8a) = 7.6$ ,  $J(7,8b) = 11.8$ ,  $J(8a,8b) = 13.6$ ,  $J(8a,9) = J(8b,9) = 6.1$  Hz.  $^{13}\text{C}$  NMR ( $\text{CD}_3\text{OD}$ , 100 MHz): 202.0 (CO(6)); 143.5, 141.8, 140.6 (C(9a), C(1), Car-s); 133.6 (C(4a)); 132.5 (C(3)); 132.1 (C(2)); 130.1 (Car-o); 129.6 (Car-m); 128.8 (Car-p); 124.8 (C(4)); 59.9 (C(7)); 46.9 (C(5)); 32.1 (C(8)); 27.5 (C(9)). HR-MS (ESI-Q-ToF) calcd for  $\text{C}_{17}\text{H}_{17}\text{BrNO}$   $[\text{M}+\text{H}]^+$ : 330.0488 and 332.0469; found: 330.0453 and 332.0436.

### 13.5. 7-Amino-1-bromo-4-phenyl-5,7,8,9-tetrahydrobenzocyclohepten-6-one, hydrochloride (1d')

*General procedure (i)* with **12d'** (61 mg, 0.14 mmol) in dioxane (1.5 mL) and 2 N HCl in  $\text{Et}_2\text{O}$  (3 mL) for 2 d to give **1d'** (49.4 mg, 95%) after recrystallisation in *i*PrOH/ $\text{Et}_2\text{O}$ .

Compound **1d'**: colorless crystals, mp 185–190 °C (dec). IR (KBr): 3424, 2923, 2891, 1722, 1582, 1579, 1513, 1454, 821, 769, 705  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CD}_3\text{OD}$ , 400 MHz): 7.63 (d, 1H, H-2); 7.48–7.34 (m, 5 Har); 7.12 (d, 1H, H-3); 4.33 (dd, 1H, H-7); 4.03 (d, 1H, Ha-5); 3.86 (d, 1H, H-5); 3.54 (ddd, 1H, Ha-9); 3.43 (ddd, 1H, Hb-9); 2.59 (m, 1H, Ha-8); 1.77 (m, 1H, Hb-8);  $J(2,3) = 8.3$ ,  $J(5a,5b) = 15.3$ ,  $J(7,8a) = 7.6$ ,  $J(7,8b) = 11.8$ ,  $J(8a,8b) = 12.8$ ,  $J(8a,9a) = 9.6$ ,  $J(8a,9b) = 3.6$ ,  $J(8b,9a) = 3.6$ ,  $J(8b,9b) = 8.2$ ,  $J(9a,9b) = 15.0$  Hz.  $^{13}\text{C}$  NMR ( $\text{CD}_3\text{OD}$ , 100 MHz): 202.6 (CO(6)); 143.8, 141.2, 140.5 (C(9a), C(1), Car-s); 133.4 (C(4a)); 133.0 (C(3)); 131.8 (C(2)); 130.8 (Car-o); 129.5 (Car-m); 128.8 (Car-p); 124.5 (C(4)); 60.4 (C(7)); 44.4 (C(5)); 31.1 (C(8)); 30.0 (C(9)). HR-MS (ESI-Q-ToF) calcd for  $\text{C}_{17}\text{H}_{17}\text{BrNO}$   $[\text{M}+\text{H}]^+$ : 330.0488 and 332.0469; found: 330.0458 and 332.0433.

### 13.6. 7-Amino-1,4-dibromo-5,7,8,9-tetrahydrobenzocyclohepten-6-one, hydrochloride (1e)

*General procedure (i)* with **12e** (60 mg, 0.14 mmol) in dioxane (1.5 mL) and 2 N HCl in  $\text{Et}_2\text{O}$  (1.5 mL) for 2 d to give **1e** (32 mg, 63%) after recrystallisation in MeOH/ $\text{Et}_2\text{O}$ .

Compound **1e**: colorless crystals, mp >250 °C. IR (KBr): 3420, 2920, 1728, 1438, 1141, 1445, 809  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CD}_3\text{OD}$ , 400 MHz): 7.49, 7.47 (2 d, 2H, H-2, H-3) 4.39 (d, 1H, H-5); 4.26 (dd, 1H, H-7); 4.21 (d, 1H, H-5); 3.43 (m, 2H,  $\text{CH}_2(9)$ ); 2.54 (m, 1H, Ha-8); 1.75 (m, 1H, Hb-8);  $J(2,3) = 8.7$ ,  $J(7,8a) = 7.9$ ,  $J(7,8b) = 11.2$ ,  $J(5a,5b) = 16.3$  Hz.  $^{13}\text{C}$  NMR ( $\text{CD}_3\text{OD}$ , 100 MHz): 201.6 (CO(6)); 142.1 (C(9a)); 135.4 (C(4a)); 134.6, 134.2 (C(2), C(3)); 125.0 124.6 (C(1), C(4)); 59.6 (C(7)); 47.4 (C(5)); 30.5 (C(9)); 30.6 (C(8)). HR-MS (ESI-Q-ToF) calcd for  $\text{C}_{11}\text{H}_{12}\text{Br}_2\text{NO}$   $[\text{M}+\text{H}]^+$ : 331.9280, 333.9259 and 335.9239; found: 331.9279, 333.9259 and 335.9238.

### 13.7. 9-Amino-7,8,9,11-tetrahydro-cyclohepta[*a*]naphthalen-10-one, hydrochloride (1h)

*General procedure (i)* with **12h** (40 mg, 0.12 mmol) in  $\text{Et}_2\text{O}$  (2 mL) and 2 N HCl in  $\text{Et}_2\text{O}$  (1 mL) for 2 d to give **1h** (27 mg, 84%).

Compound **1h**: colorless crystals, mp 240 °C (dec). IR (KBr): 3440, 2936, 1730, 1719, 1509, 1498, 1114, 1082, 1056, 820, 744  $\text{cm}^{-1}$ .  $^1\text{H}$  NMR ( $\text{CD}_3\text{OD}$ , 400 MHz): 8.22 (d, 1H, H-1); 7.87 (d, 1H, H-4); 7.80 (d, 1H, H-5); 7.58 (dt, 1H, H-2); 7.49 (dt, 1H, H-3); 7.40 (d, 1H, H-6); 4.49 (dd, 1H, H-9); 4.46 (d, 1H, Ha-11); 4.40 (d, 1H, Hb-11); 3.49 (m, 1H, Ha-7); 3.19 (ddd, 1H, Hb-7); 2.67 (m, 1H, Ha-8); 1.80 (m, 1H, Hb-8);  $J(1,2) = 8.4$ ,  $J(1,3) = 1.0$ ,  $J(2,3) = 6.8$ ,  $J(2,4) = 1.3$ ,  $J(3,4) = 8.1$ ,  $J(5,6) = 8.4$ ,  $J(7a,7b) = 14.8$ ,  $J(7a,8a) = 2.6$ ,  $J(7a,8b) = 10.2$ ,  $J(7b,8a) = 8.0$ ,  $J(7b,8b) = 2.8$ ,  $J(8a,9) = 7.2$ ,  $J(8b,9) = 12.0$ ,  $J(8a,8b) = 12.9$ ,  $J(11a,11b) = 14.2$  Hz.  $^{13}\text{C}$  NMR ( $\text{CD}_3\text{OD}$ , 100 MHz): 202.6 (C(10)); 140.2 (C(6a)); 135.3, 133.2 (C(4a), C(11b)); 130.2 (C(4)); 129.9 (C(5)); 129.0 (C(6)); 128.9 (C(11a));

128.3 (C(2)); 127.0 (C(3)); 124.8 (C(1)); 61.6 (C(9)); 41.6 (C(11)); 34.5 (C(8)); 34.0 (C(7)). HR-MS (ESI-Q-ToF) calcd for  $C_{15}H_{16}NO$   $[M+H]^+$ : 226.1226; found: 226.1221.

### 13.8. 9-Amino-7,9,10,11-tetrahydro-cyclohepta[*a*]naphthalen-8-one, hydrochloride (1h')

*General procedure (i)* with **12h'** (50 mg, 0.15 mmol) in  $Et_2O$  (2 mL) and 2 N HCl in  $Et_2O$  (1 mL) for 2 d to give **1h'** (30 mg, 77%).

Compound **1h'**: colorless crystals, mp 232 °C (dec.). IR (KBr): 3422, 2971, 2224, 1721, 1512, 1488, 1458, 818, 772, 737  $cm^{-1}$ .  $^1H$  NMR ( $CD_3OD$ , 400 MHz): 8.22 (d, 1H, H-1); 7.89 (d, 1H, H-4); 7.79 (d, 1H, H-5); 7.58 (dt, 1H, H-2); 7.50 (dt, 1H, H-3); 7.39 (d, 1H, H-6); 4.34 (dd, 1H, H-9); 4.33 (d, 1H, Ha-7); 3.92 (d, 1H, Hb-7); 3.67 (ddd, 1H, Ha-11); 3.57 (ddd, 1H, Hb-11); 2.70 (m, 1H, Ha-10); 1.84 (m, 1H, Hb-10);  $J(1,2) = 8.5$ ,  $J(1,3) = 1.2$ ,  $J(2,3) = 6.8$ ,  $J(2,4) = 1.4$ ,  $J(3,4) = 8.0$ ,  $J(5,6) = 8.4$ ,  $J(7a,7b) = 15.4$ ,  $J(9,10a) = 7.5$ ,  $J(9,10b) = 11.6$ ,  $J(10a,10b) = 12.6$ ,  $J(10a,11a) = 9.5$ ,  $J(10a,11b) = 3.5$ ,  $J(10b,11a) = 3.6$ ,  $J(10b,11b) = 7.8$ ,  $J(11a,11b) = 15.2$  Hz.  $^{13}C$  NMR ( $CD_3OD$ , 100 MHz): 202.8 (C(8)); 136.7 (C(6a)); 135.0, 132.6 (C(4a), C(11b)); 131.1 (C(11a)); 129.9 (C(4)); 129.0 (C(5)); 128.1 (C(6)); 127.9 (C(2)); 126.8 (C(3)); 124.0 (C(1)); 60.3 (C(9)); 48.4 (C(7)); 32.6 (C(10)); 23.9 (C(11)). HR-MS (ESI-Q-ToF) calcd for  $C_{15}H_{16}NO$   $[M+H]^+$ : 226.1226; found: 226.1206.

### 14. Preparation of 12f,f'g by Suzuki coupling

*General procedure (j)* a mixture of bromocetoamide (1 mmol), phenylboronic acid (140 mg, 1.13 mmol, 1.1 equiv), CsF (0.34 g, 2.26 mmol, 2.2 equiv) and  $Pd(PPh_3)_4$  (120 mg, 0.1 mmol) in dry 1,2-dimethoxyethane (DME, 12 mL) was stirred under Ar at 85 °C for 5 h. The reaction mixture was diluted with AcOEt, washed with brine and dried ( $MgSO_4$ ). The solvent was evaporated and the residue purified by flash chromatography (cyclohexane/AcOEt 9/1).

*General procedure (k)* same procedure with  $K_2CO_3$  (0.2 g, 1.5 mmol, 1.5 equiv) as base and in DME (12 mL) and  $H_2O$  (3 mL) as reaction solvent.

### 14.1. 7-tert-Butoxycarbonylamino-4-phenyl-5,7,8,9-tetrahydro-benzocyclohepten-6-one (12f)

*General procedure (j)* with **12c** (90 mg, 0.256 mmol), phenylboronic acid (35 mg, 0.28 mmol), CsF (86 mg, 0.56 mmol) and  $Pd(PPh_3)_4$  (30 mg, 0.026 mmol) in DME (3 mL) was heated at 85 °C for 5 h. The work-up gave **12f** (70 mg, 78%).

Compound **12f**: colorless crystals, mp 173–174 °C. IR (KBr): 3277, 2978, 1725, 1706, 1677, 1554, 1366, 1279, 1189, 1005, 762, 705  $cm^{-1}$ .  $^1H$  NMR ( $CDCl_3$ , 400 MHz): 7.40 (m, 5 Har); 7.21 (m, 3 Har); 5.44 (d, 1H, NH); 4.55 (m, 1H, H-7); 3.80 (d, 1H, Ha-5); 3.71 (d, 1H, Hb-5); 3.06 (m, 1H, Ha-9); 2.97 (ddd, 1H, Hb-9); 2.67 (m, 1H, Ha-8); 1.54 (m, 1H, Hb-8); 1.42 (s, 9H,  $CM_e_3$ );  $J(5a,5b) = 15.0$ ,  $J(NH,7) = 7.0$ ,  $J(7,8a) = 7.2$ ,  $J(7,8b) = 10.0$ ,  $J(8a,8b) = 13.0$ ,  $J(8a,9a) = 3.6$ ,  $J(8a,9b) = 9.0$ ,  $J(8b,9a) = 8.0$ ,  $J(8b,9b) = 3.7$ ,  $J(9a,9b) = 14.6$  Hz.  $^{13}C$  NMR ( $CDCl_3$ , 100 MHz): 205.6 (C(6)); 155.1 (NCO<sub>2</sub>); 142.7 (C(9a)); 140.9, 140.8 (3Car); 130.0 (C(4a)); 130.0, 129.5, 128.6, 128.4, 127.4, 127.3 (6Char); 79.8 ( $CM_e_3$ ); 60.8 (C(7)); 43.6 (C(5)); 34.7 (C(8)); 31.6 (C(9)); 28.5 ( $CM_e_3$ ). HR-MS (ESI-Q-ToF) calcd for  $C_{22}H_{25}LiNO_3$   $[M+Li]^+$ : 358.1989; found: 358.1899;  $C_{22}H_{25}NaNO_3$   $[M+Na]^+$ : 374.1727; found: 374.1635.

### 14.2. 7-tert-Butoxycarbonylamino-1-phenyl-5,7,8,9-tetrahydro-benzocyclohepten-6-one (12f')

*General procedure (j)* with **12c'** (90 mg, 0.26 mmol), phenylboronic acid (35 mg, 0.28 mmol), CsF (86 mg, 0.56 mmol) and

$Pd(PPh_3)_4$  (30 mg, 0.026 mmol) in DME (3 mL) at 85 °C for 5 h. The work-up gave **12f'** (75 mg, 83%).

Compound **12f'**: colorless crystals, mp 187–188 °C. IR (KBr): 3348, 2981, 2931, 1720, 1682, 1524, 1365, 1295, 1253, 1171, 1052, 761, 705  $cm^{-1}$ .  $^1H$  NMR ( $CDCl_3$ , 400 MHz): 7.38 (m, 3 Har); 7.27 (m, 2 Har); 7.20 (m, 2 Har); 5.41 (d, 1H, NH); 4.58 (dt, 1H, H-7); 3.95 (d, 1H, Ha-5); 3.72 (d, 1H, Hb-5); 2.85 (m, 2H,  $CH_2(9)$ ); 2.53 (m, 1H, Ha-8); 1.46 (m, 1H, Hb-8); 1.42 (s, 9H,  $CM_e_3$ );  $J(5a,5b) = 15.6$ ,  $J(NH,7) = 7.2$ ,  $J(7,8a) = 7.2$ ,  $J(7,8b) = 11.2$  Hz.  $^{13}C$  NMR ( $CDCl_3$ , 100 MHz): 205.7 (C(6)); 155.1 (NCO<sub>2</sub>); 142.4, 141.4 (2 Car); 137.5 (C(9a)); 133.3 (C(4a)); 129.9, 129.4, 129.2, 128.3, 127.2, 126.8 (6Char); 79.9 ( $CM_e_3$ ); 60.0 (C(7)); 48.5 (C(5)); 34.5 (C(8)); 28.5 ( $CM_e_3$ ); 26.3 (C(9)). HR-MS (ESI-Q-ToF) calcd for  $C_{22}H_{25}NaNO_3$   $[M+Na]^+$ : 374.1727; found: 374.1720.

### 14.3. 7-tert-Butoxycarbonylamino-1,4-diphenyl-5,7,8,9-tetrahydro-benzocyclohepten-6-one (12g)

1. *General procedure (k)*: A solution of **12e** (30 mg, 0.07 mmol), phenylboronic acid (34 mg, 0.28 mmol, 4 equiv),  $K_2CO_3$  (38 mg, 0.28 mmol, 4 equiv) and  $Pd(PPh_3)_4$  (24 mg, 0.021 mmol) in DME (2 mL) and  $H_2O$  (0.1 mL) were heated under Argon in a microwave heater (for 25 min at 300 W/125 °C/3 bar). The work-up gave **12g** (23 mg, 79%).

2. *General procedure (k)* with **12d** (100 mg, 0.23 mmol), phenylboronic acid (42 mg, 0.34 mmol, 1.5 equiv),  $K_2CO_3$  (48 mg, 0.34 mmol) and  $Pd(PPh_3)_4$  (26.5 mg, 0.023 mmol) in DME (3.3 mL) and  $H_2O$  (0.7 mL) for 3 h at 85 °C. The work-up gave **12g** (88 mg, 89%) after washing with  $iPr_2O$ .

Compound **12g**: colorless crystals, mp 192–196 °C ( $iPr_2O$ ). IR (KBr): 3410, 2972, 2930, 1705, 1492, 1365, 1159, 705  $cm^{-1}$ .  $^1H$  NMR ( $CDCl_3$ , 400 MHz): 1.42 (s, 9H,  $CM_e_3$ ); 1.51 (m, 1H, Hb-8); 2.57 (m, 1H, Ha-8); 2.91 (m, 2H, Hb-9, Ha-9); 3.75 (d, 1H, Hb-5); 3.89 (d, 1H, Ha-5); 4.6 (td, 1H, H-7); 5.43 (d, 1H, NH-7); 7.23 (s, 2H, H-2, H-3); 7.33–7.47 (m, 10 Har);  $J(5a,5b) = 16.4$ ,  $J(7,NH) = 7.6$ ,  $J(7,8a) = 7.6$ ,  $J(7,8b) = 11.2$ ,  $J(8a,8b) = 13.2$  Hz.  $^{13}C$  NMR ( $CDCl_3$ , 100 MHz): 26.7 (C(9)); 28.3 ( $CM_e_3$ ); 34.1 (C(8)); 44.0 (C(5)); 59.6 (C(7)); 79.6 ( $CM_e_3$ ); 127.1, 127.2 (2 Char-p); 128.2 (2 Char-m); 128.5 (C(3)); 129.0 (C(2)); 129.2, 129.8 (2 Char-o); 130.7 (C(4a)); 137.8 (C(9a)); 140.8; 141.4; 141.8 (3 Car); 154.9 (NCO-7); 206.0 (CO(6)). HR-MS (ESI-Q-ToF) calcd for  $C_{28}H_{29}NO_3$   $[M+Na]^+$ : 450.2045; found: 450.2035.

### 15. Keto-amines 1f,f'g

#### 15.1. 7-Amino-4-phenyl-5,7,8,9-tetrahydro-benzocyclohepten-6-one (1f)

*General procedure (i)* with **12f** (50 mg, 0.14 mmol) in dioxane (1 mL) and 2 N HCl in  $Et_2O$  (1 mL) for 48 h to give **1f** (30 mg, 73%).

Compound **1f**: colorless crystals, mp 255–260 °C (dec) (MeOH/ $Et_2O$ ). IR (KBr): 3450, 2898, 2890, 2157, 1715, 1463, 1170, 763, 707  $cm^{-1}$ .  $^1H$  NMR ( $CD_3OD$ , 400 MHz, 9:1 mixture with an hemiacetal). Ketone: 7.41 (m, 5Har); 7.28 (m, 2Har); 7.20 (dd, 1H,  $J = 3.1$ , 5.9 Hz, 1 Har); 4.39 (dd, 1H, H-7); 3.99 (d, 1H, Ha-5); 3.77 (d, 1H, Hb-5); 3.32 (m, 1H, Ha-9); 3.13 (ddd, 1H, Hb-9); 2.60 (m, 1H, Ha-8); 1.76 (m, 1H, Hb-8);  $J(5a,5b) = 13.8$ ,  $J(7,8a) = 7.0$ ,  $J(7,8b) = 12.0$ ,  $J(8a,8b) = 12.6$ ,  $J(8a,9a) = 2.8$ ,  $J(8a,9b) = 8.1$ ,  $J(8b,9a) = 10.0$ ,  $J(8b,9b) = 3.1$ ,  $J(9a,9b) = 15.0$  Hz. Hemi-acetal, partial data: ca. 1.75 (m, 1H, Hb-8); 2.06 (m, 1H, Ha-8); 2.96 (m, 2H,  $CH_2(9)$ ).  $^{13}C$  NMR ( $CD_3OD$ , 100 MHz): 202.9 (CO(6)); 144.1 (C(9a)); 142.1 (2Car); 131.0, 130.7 (2 Char); 130.6 (C(4a)); 129.8, 129.3, 128.7, 128.4 (4Char); 61.3 (C(7)); 43.3 (C(5)); 33.1 (C(8)); 31.6 (C(9)). HR-MS (ESI-Q-ToF) calcd for  $C_{17}H_{18}NO$   $[M+H]^+$ : 252.1383; found: 252.1366.

### 15.2. 7-Amino-1-phenyl-5,7,8,9-tetrahydro-benzocyclohepten-6-one (1f)

General procedure (i) with **12f** (35 mg, 0.1 mmol) in dioxane (1 mL) and 2 N HCl in Et<sub>2</sub>O (1 mL) for 48 h to give **1f** (22 mg, 76%).

Compound **1f**: colorless crystals, mp >250 °C (MeOH/Et<sub>2</sub>O). IR (KBr): 3422, 2966, 1722, 1484, 1459, 762, 704 cm<sup>-1</sup>. <sup>1</sup>H NMR (CD<sub>3</sub>OD, 400 MHz): 7.43 (m, 3Har); 7.28 (m, 4Har); 7.21 (m, 1Har); 4.35 (dd, 1H, H-7); 4.22 (d, 1H, Ha-5); 3.77 (d, 1H, Hb-5); 3.06 (m, 1H, Ha-9); 3.02 (m, 1H, Hb-9); 2.40 (m, 1H, Ha-8); 1.70 (m, 1H, Hb-8); J(5a,5b) = 14.7, J(7,8a) = 7.2, J(7,8b) = 11.8, J(8a,8b) = 13.2, J(8a,9a) = ca. 8.2, J(8a,9b) = ca. 4.2, J(8b,9a) = ca. 4.2, J(8b,9b) = ca. 7.6, J(9a,9b) = 15.0 Hz. <sup>13</sup>C NMR (CD<sub>3</sub>OD, 100 MHz): 202.7 (CO(6)); 143.8, 142.6 (2Car); 138.3 (C(9a)); 134.2 (C(4a)); 130.8, 130.2, 130.2, 129.4, 128.4, 128.2 (6CHar); 60.5 (C(7)); 48.3 (C(5)); 32.7 (C(8)); 26.4 (C(9)). HR-MS (ESI-Q-Tof) calcd for C<sub>17</sub>H<sub>18</sub>NO [M+H]<sup>+</sup>: 252.1383; found: 252.1363.

### 15.3. 7-Amino-1,4-diphenyl-5,7,8,9-tetrahydro-benzocyclohepten-6-one (1g)

General procedure (i) with **12g** (30 mg, 0.07 mmol) in 4 N HCl in Et<sub>2</sub>O (1 mL) and dioxane (1 mL) for 16 h to give **1g** (23 mg, 88%).

Compound **1g**: colorless crystals, mp >200 °C. IR (KBr): 3408, 2923, 2867, 1725, 1509, 1467, 762, 702 cm<sup>-1</sup>. <sup>1</sup>H NMR (CD<sub>3</sub>OD, 400 MHz): 1.77 (m, 1H, Hb-8); 2.44 (m, 1H, Ha-8); 3.07 (ddd, 1H, Hb-9); 3.14 (ddd, 1H, Ha-9); 3.85 (d, 1H, Hb-5); 4.04 (d, 1H, Ha-5); 4.38 (dd, 1H, H-7); 7.25 (s, 2H, H-2,H-3); 7.32 (d, 2Har-o); 7.39–7.49 (m, 8Har); J(5a,5b) = 14.8, J(7,8a) = 7.2, J(7,8b) = 12.0, J(8a,8b) = 12.4, J(8a,9a) = 8.8, J(8a,9b) = 4.0, J(8b,9a) = 4.0, J(8b,9b) = 8.0, J(9a,9b) = 15.0 Hz. <sup>13</sup>C NMR (CD<sub>3</sub>OD, 100 MHz): 26.9 (C(9)); 32.4 (C(8)); 43.8 (C(5)); 60.6 (C(7)); 128.4, 128.5 (2CHar-p); 129.4, 129.5 (2CHar-m); 129.9 (C(3)); 130.2 (C(2), CHar-o); 130.9 (CHar-o); 131.6 (C(4a)); 139 (C(9a)); 142.2, 142.7, 143.1, 143.4 (4CHar); 203.1 (C(6)).

HR-MS (ESI-Q-Tof) calcd for C<sub>23</sub>H<sub>22</sub>NO [M+H]<sup>+</sup>: 328.1701; found: 328.1656.

## 16. Enzyme assays

### 16.1. Enzyme source

Porcine kidney APN and *Aeromonas proteolytica* aminopeptidase were purchased from Sigma Chemical Co. Porcine kidney LAPC was purified according to a published procedure.<sup>26</sup> Human recombinant LTA<sub>4</sub>H was provided by our collaborator J. Z. Haeggström.<sup>25c</sup>

### 16.2. Assay conditions<sup>25c</sup>

(a) All enzymes: Kinetic data were collected with an HP/Agilent UV-Visible, diode array, spectrophotometer 8453 using the software 'HP chemstation' provided with the machine. Typically, spectrophotometric assays were performed with L-leucine-p-nitroanilide as the substrate for APN ( $K_m = 0.2$  mM), LAPC ( $K_m = 2$  mM) and APAero ( $K_m = 0.02$  mM). All kinetic studies were performed at 30 °C and the reactions were started by addition of the enzyme in 1 ml assay medium. (b) APN: 1 mUnits per assay, in 0.02 M Tris-HCl pH 7.5. (c) LAPC: 20 Units per assay in 0.1 M Tris-HCl, 0.1 mM ZnCl<sub>2</sub>, 5 mM MnCl<sub>2</sub>, 1 M KCl, pH 8.0 and (d) APAero 2 mUnits per assays in 0.05 M Tris-HCl pH 7.5.

The release of p-nitroanilide ( $\epsilon = 10,800$  M<sup>-1</sup> cm<sup>-1</sup>) at 405 nm was measured continuously during 30 min to determine initial velocities. Assays were performed in semi microcuvettes (1 cm path).  $K_i$  were determined using Dixon plots.<sup>34</sup>

For the specific evaluation of compound **1d'**, the concentration of APN used in the assay was decreased to 0.1 mUnits (12 pM) per assay and the linear reaction was monitored during at least 5–6 h in order to measure significant velocities. The  $K_i$  value was also determined from a Dixon plot.

## Acknowledgments

The support of the *École Nationale Supérieure de Chimie de Mulhouse* and the *Université de Haute-Alsace* is gratefully acknowledged. We also wish to thank the *Ligue contre le Cancer* for financial support. We thank Professor Patrick Pale for his scientific and material assistance, Dr. Cécile Joyeux for the HR-MS measurements and the students Azely Mirre, Julien Debray, Arnaud Mignatelli and Meral Ilhan for their participation to this work.

## References and notes

- Barret, A. J.; Rawling, N. D.; Woessner, J. F., 2nd ed. In *Handbook of Proteolytic Enzymes*; Elsevier Academic Press: Oxford, 2004; Vol. 2, p 233sq.
- Mina-Osorio, P. *Trends Mol. Med.* **2008**, *14*, 361.
- (a) Bhagwat, S. V.; Lahdenranta, J.; Giordano, R.; Arap, W.; Pasqualini, R. *Blood* **2003**, *101*, 1818; (c) Petrovic, N.; Okamoto, Y.; Shapiro, L. H. *Blood* **2003**, *101*, 1818; (c) Petrovic, N.; Schacke, W.; Gahagan, J. R.; O'Conor, C. A.; Winnicka, B.; Conway, R. E.; Mina-Osorio, P.; Shapiro, L. H. *Blood* **2007**, *110*, 142.
- (a) Terauchi, M.; Kajiyama, H.; Shibata, K.; Ino, K.; Nawa, A.; Mizutani, S.; Kikkawa, F. *BMC Cancer* **2007**, *7*, 140; (b) Langner, J.; Mueller, C.; Riemann, D.; Löhn, M. *Immunol. Lett.* **1997**, *56*, 62.
- (a) Yamashita, M.; Kajiyama, H.; Terauchi, M.; Shibata, K.; Ino, K.; Nawa, A.; Mizutani, S.; Kikkawa, F. *Int. J. Cancer* **2007**, *120*, 2243; (b) Fukasawa, K.; Fujii, H.; Saitoh, Y.; Koizumi, K.; Aozuka, Y.; Sekine, K.; Yamada, M.; Saiki, I.; Nishikawa, K. *Cancer Lett.* **2006**, *243*, 135.
- Rangel, R.; Sun, Y.; Guzman-Rojas, L.; Ozawa, M. G.; Sun, J.; Giordano, R. J.; Van Pelt, C. S.; Tinkey, P. T.; Behringer, R. R.; Sidman, R. L.; Arap, W.; Pasqualini, R. *Proc. Natl. Acad. Sci.* **2007**, *104*, 4588.
- (a) Riemann, D.; Kehlen, A.; Langner, J. *Immunol. Today* **1999**, *20*, 83; (b) Bank, U.; Heimbürg, A.; Helmuth, M.; Stefin, S.; Lendeckel, U.; Reinhold, D.; Faust, J.; Fuchs, P.; Sens, B.; Neubert, K.; Täger, M.; Ansoerge, S. *Int. Immunopharmacol.* **2006**, *6*, 1925; (c) Matteo, P. D.; Arrigoni, G. L.; Alberici, L.; Corti, A.; Gallo-Stampino, C.; Traversari, C.; Dogliani, C.; Rizzardi, G.-P. *Histochem. Cytochem.* **2011**, *59*, 47.
- Reinhold, D.; Biton, A.; Pieper, S.; Lendeckel, U.; Faust, J.; Neubert, K.; Bank, U.; Täger, M.; Ansoerge, S.; Brocke, S. *Int. Immunopharmacol.* **2006**, *6*, 1935.
- (a) Mucha, A.; Drag, M.; Dalton, J. P.; Kafarski, P. *Biochimie* **2010**, *92*, 1509; (b) Bauvois, B.; Dauzonne, D. *Med. Chem. Rev.* **2006**, *26*, 88.
- Cobson, F. E.; Lewis, J. A.; Cohen, S. M. *Chem. Med. Chem.* **2007**, *2*, 152.
- Albrecht, S.; Al-Lakkis-Wehbe, M.; Orsini, A.; Defoin, A.; Pale, P.; Salomon, E.; Tarnus, C.; Weibel, J.-M. *Bioorg. Med. Chem.* **2011**, *19*, 1434.
- Hopkins, A. L.; Groom, C. R.; Alex, A. *Drug Discovery Today* **2004**, *9*, 430.
- Matsui, M.; Fowler, J. H.; Walling, L. L. *Biol. Chem.* **2006**, *387*, 1535.
- Barret, A. J.; Rawling, N. D.; Woessner, J. F., 2nd ed. In *Handbook of Proteolytic Enzymes*; Elsevier Academic Press: Oxford, 2004; Vol. 2, p 253 sq.
- Lai, Y.-H.; Yap, A. H.-T. *J. Chem. Soc., Perkin Trans. 2* **1993**, 1373.
- Sengupta, S.; Bhattacharyya, S. *J. Org. Chem.* **1997**, *62*, 3405.
- Kajigaeshi, S.; Kakinami, T.; Okamoto, T.; Nakamura, H.; Fujikawa, M. *Bull. Chem. Soc. Jpn.* **1987**, *60*, 4187.
- Simchen, G.; Kober, W. *Synthesis* **1976**, 259.
- Taniguchi, Y.; Inanaga, J.; Yamaguchi, M. *Bull. Chem. Soc. Jpn.* **1981**, *154*, 3229.
- (a) Hassner, A.; Reuss, R. H.; Pinnick, H. W. *J. Org. Chem.* **1975**, *40*, 3427; (b) Rubottom, G. M.; Vasquez, M. A.; Pelegrina, D. R. *Tetrahedron Lett.* **1974**, *15*, 4319.
- Miriyala, B.; Bhattacharyya, S.; Williamson, J. *Tetrahedron* **2004**, *60*, 1463.
- (a) Boeckman, R. K., Jr.; Shao, P.; Mullins, J. J. *Org. Synth.* **2000**, *77*, 141; (b) Dess, D. B.; Martin, J. C. *J. Org. Chem.* **1983**, *48*, 4155. [www.SyntheticPages.org/pages/51](http://www.SyntheticPages.org/pages/51).
- Auffinger, P.; Hays, F. A.; Westhof, E.; Shing, H. P. *Proc. Natl. Acad. Sci.* **2004**, *101*, 16789.
- Harnandes, M. Z.; Cavalcanti, S. M. T.; Moreira, D. R. M.; de Azevedo, W. F., Jr.; Leite, A. C. L. *Curr. Drug Targets* **2010**, *11*, 303.
- (a) Schalk, C.; d'Orchymont, H.; Jauch, M.-F.; Tarnus, C. *Arch. Biochem. Biophys.* **1994**, *311*, 42; (b) d'Orchymont, H.; Tarnus, C. Eur. Patent 0378456 A1, 1990; (c) Albrecht, S.; Defoin, A.; Salomon, E.; Tarnus, C.; Wetterholm, A.; Haeggström, J. Z. *Bioorg. Med. Chem.* **2006**, *14*, 7241.
- Himmelhoch, R.; Peterson, S. A. *Biochemistry* **1968**, *7*, 2085.
- Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3.
- Skinner-Adams, T. S.; Stack, C. M.; Trenholme, K. R.; Brown, C. L.; Grembecka, J.; Lowther, J.; Mucha, A.; Drag, M.; Kafarski, P.; McGowan, S.; Whisstock, J. C.; Gardiner, D. L.; Dalton, J. P. *Trends Biochem. Sci.* **2010**, *35*, 53.
- Dey, A. S.; Rosowsky, A.; Modest, E. J. *J. Org. Chem.* **1970**, *35*, 536.

30. (a) McCauley, J. A.; McIntyre, C. J.; Rudd, M. T.; Nguyen, K. T.; Romano, J. J.; Butcher, J. W.; Gilbert, K. F.; Bush, K. J.; Holloway, M. K.; Swestock, J.; Wan, B.-L.; Carroll, S. S.; DiMuzio, J. M.; Graham, D. J.; Ludmerer, S. W.; Mao, S.-S.; Stahlhut, M. W.; Fandozzi, C. M.; Trainor, N.; Olsen, D. B.; Vacca, J. P.; Liverton, N. J. *J. Med. Chem.* **2010**, *53*, 2443; (b) Tsue, H.; Nakashima, S.; Goto, Y.; Tatemitsu, H.; Misumi, S.; Abraham, R. J.; Asahi, T.; Tanaka, Y.; Okada, T.; Mataga, N.; Sakata, Y. *Bull. Chem. Soc. Jpn.* **1994**, *67*, 3067.
31. Wu, A.; Chakraborty, A.; Witt, D.; Lagona, J.; Damkaci, F.; Ofori, M. A.; Chiles, J. K.; Fettinger, J. C.; Isaacs, L. *J. Org. Chem.* **2002**, *67*, 5817.
32. Ried, W.; Bodem, H. *Liebigs Ann. Chem.* **1954**, 589, 55.
33. Ohkawa, S.; DiGiacomo, B.; Larson, D. L.; Takemori, A. E.; Portoghese, P. S. *J. Med. Chem.* **1997**, *40*, 1720.
34. Segel, I. H. *Enzyme Kinetics*; John Wiley & Son: New-York, 1993.

## 2. Publication n°2 : Selective Aminopeptidase-N (CD13) – inhibitors of relevance to cancer chemotherapy

La sélectivité des dérivés de la 7-amino-6-benzosubérone pour l'APN de mammifères ont conduit à évaluer l'efficacité de ces inhibiteurs au niveau cellulaire dans des modèles standards liés à l'angiogenèse et à la cancérogenèse, ainsi qu'à réaliser les premières études pilotes chez la souris.

Dans un premier temps, les composés I1, I2 et I3, ont été testés *in vitro* dans des modèles standards d'étude de l'angiogenèse. Nous avons démontré que ces derniers inhibent de façon dose-dépendante la migration des cellules endothéliales, ainsi que leur organisation en tube capillaire, et ce à des doses non cytotoxiques. L'effet de ces composés sur la migration des cellules tumorales exprimant l'APN a également été mis en évidence.

L'efficacité *in vivo* des inhibiteurs I2 et I3, seuls ou en association avec la rapamycine (inhibiteur de mTOR), a ensuite été étudiée dans des modèles de xénogreffes de tumeurs humaines implantées chez la souris « nude ». Un effet sur la néovascularisation tumorale est observé lorsque les inhibiteurs sont administrés le jour suivant l'injection des cellules cancéreuses, et ce avec une activité proche de celle de l'Avastin®. Ces molécules sont également actives dans des modèles de tumeurs établies (plus proche de la situation clinique) chez la souris. L'effet sur la croissance tumorale dépend, néanmoins, du niveau d'expression de la protéine par les cellules tumorales et non au niveau de la néovascularisation, et est amélioré lorsque ces composés sont utilisés en association avec la rapamycine.

Ces résultats sont prometteurs ; toutefois, des études pharmacocinétiques doivent être rapidement entreprises afin d'optimiser au mieux le traitement (les doses, ainsi que le mode et les fréquences d'injection).

Ces travaux ont été effectués en collaboration avec Manon Voegelin de l'équipe du Dr. Dominique Guenot du laboratoire Physiopathologie et Médecine Translationnelle (EA 4438, Strasbourg). L'ensemble de ces résultats a fait l'objet d'une publication soumise dans le journal « Bioorganic and Medicinal Chemistry ».



## Selective Aminopeptidase-N (CD13) – inhibitors of relevance to cancer chemotherapy

Céline Schmitt<sup>[a],#</sup>, Manon Voegelin<sup>[b],#</sup>, Aurélie Marin<sup>[a]</sup>, Marjorie Schmitt<sup>[a]</sup>, Fabienne Schegg<sup>[a]</sup>, Philippe Hénon<sup>[c]</sup>, Dominique Guenot<sup>[b]\*</sup>, Céline Tarnus<sup>[a]\*</sup>

---

*# authors contributed equally to the work*

*[a] Université de Haute-Alsace, Ecole Nationale Supérieure de Chimie de Mulhouse, Laboratoire de Chimie Organique et Bioorganique EA4566, 3 rue Alfred Werner, 68093 Mulhouse Cedex, France*

*[b] Physiopathologie et Médecine Translationnelle, Université de Strasbourg, Inserm U682, EA 4438, 3 avenue Molière, 67098 Strasbourg Cedex, France*

*[c] Institut de Recherche en Hématologie et Transplantation, Hôpital du Hasenrain, 87 avenue d'Altkirch, 68100 Mulhouse, France*

*\* Corresponding authors :*

*In vivo studies : Dr. Dominique Guenot, EA 4438, 3 avenue Molière, 67098 Strasbourg Cedex, France. Phone : 33-(0)3-88-27-53-62 ; Fax : 33-(0)3-88-26-35-38 ; dominique.guenot@inserm.fr*

*Chemistry and biochemistry : Pr. Céline Tarnus, EA 4566, 3 rue Alfred Werner, 68093 Mulhouse Cedex, France. Phone : 33-(0)3-89-33-68-66 ; Fax : 33-(0)3-89-33-68-60 ; celine.tarnus@uha.fr*

---



## Abstract

---

Aminopeptidase N (APN/CD13) is highly expressed on the surface of numerous types of cancer cells and particularly on the endothelial cells of neoangiogenic vessels during tumourigenesis. This metallo-aminopeptidase has been identified as a potential target for cancer chemotherapy. In this work, we evaluated the efficacy of a novel series of benzosuberone analogues, which were previously reported to be highly potent, selective APN inhibitors with  $K_i$  values in the micromolar to sub-nanomolar range. Endothelial cell morphogenesis as well as cell motility were inhibited *in vitro* in a dose-dependent manner at concentrations that correlated with the potency of the compounds, thus confirming the key role of APN in these established models of angiogenesis. We report toxicity studies in mice showing that these compounds are well tolerated. We report the effects of the compounds, used alone or in combination with rapamycin, on the growth of a select panel of tumours that were subcutaneously xenografted onto Swiss nude mice. Our data indicate that the *in vivo* efficacy of these new APN inhibitors during the initial phase of tumour growth can be ascribed to their anti-angiogenic activities. However, we also provide evidence that these compounds are effective against established solid tumours. For colonic tumours, the anti-tumour effect depends on the level of APN expression in epithelial cells, and APN expression is associated with down-regulation of the transcription factor HIF-1 $\alpha$ . These effects seem to be distinct from those of rapamycin. Our finding that the anti-tumour effect of the inhibitors in the colon requires APN expression strongly suggests that APN plays a crucial function in tumour cells that is distinct from its known role in neovascularisation.

---

**Keywords:** metalloenzyme, aminopeptidase, biological activity, antitumor, rapamycin



## 1. Introduction

Aminopeptidase-N (APN/CD13; EC 3.4.11.2) is an ubiquitous high molecular weight ectopeptidase that belongs to the M1 family of zinc-dependent enzymes<sup>[1]</sup>. It is known to preferentially remove the N-terminal neutral amino acid from unsubstituted oligopeptides and has broad substrate specificity, although a significant preference for hydrophobic residues has been observed<sup>[1,2]</sup>. APN/CD13 is expressed in several human organs, tissues and cell types<sup>[3,4]</sup>, and this broad tissue distribution makes the identification of both its endogenous substrates and its biological functions highly complicated. APN is a multifunctional protein belonging to the moonlighting enzyme family and possesses a catalytic activity that might be linked to complex and systemic effects. Understanding the mechanisms underlying such effects remains a challenge because many of APN's *in vivo* functions have not yet been verified. Nevertheless, interest in APN as a therapeutic target is increasing<sup>[5-7]</sup>.

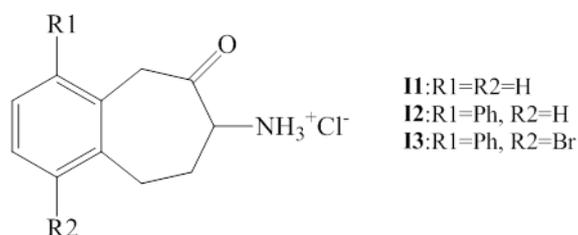
The expression of APN is related to the regulation of signal peptides and is associated with increased migration in various cells<sup>[3,4]</sup>. For example, APN overexpression correlates with immunological abnormalities such as chronic inflammatory diseases<sup>[8,9]</sup> and autoimmune pathologies<sup>[10]</sup>, suggesting a role for APN in T cell function and activation. Several studies using small molecular weight inhibitors of APN<sup>[11]</sup>, active site-directed anti-APN monoclonal antibodies<sup>[11-13]</sup>, APN siRNA<sup>[11]</sup> and APN null mice<sup>[12-15]</sup> have indicated that APN is an active player in angiogenesis, tumour progression and metastasis. Consequently, many of the functions attributed to APN are associated with various aspects of normal development (e.g., that of myeloid progenitor cells) and malignant processes such as tumour cell invasion, differentiation, proliferation, apoptosis, motility and angiogenesis<sup>[16-19]</sup>. With respect to angiogenesis, APN is exclusively found on the endothelial cells of angiogenic vasculature but not those of the normal vasculature<sup>[6]</sup>. Treatment of animals with APN inhibitors significantly decreased retinal neoangiogenesis and xenograft tumour growth. These inhibitors impaired tube formation but not the proliferation of vascular endothelial cells, and APN appears to control endothelial cell morphogenesis<sup>[3,4,6]</sup>. A strong correlation between the expression and catalytic activity of APN and the invasive capacity of numerous tumour cell types was observed during tumourigenesis; moreover, APN has been reported to be either overexpressed or aberrantly activated in numerous carcinomas<sup>[20-22]</sup>. More specifically, APN expression is associated with tumour status in colon cancer, and the overall survival rate of patients with tumours expressing APN was significantly lower than that of patients with APN-negative tumours<sup>[23]</sup>.



Similar results have been described for breast cancer patients<sup>[24]</sup>. APN is a functional marker that could potentially be used to identify dormant liver cancer cells that are resistant to treatment; these cells play a role in liver cancer relapse<sup>[5]</sup>.

The rational design of potent and selective APN inhibitors is of considerable interest because certain functions of APN can be exploited for therapeutic benefit. Over the course of more than three decades, studies have evaluated numerous inhibitors of APN, including bestatin<sup>[25]</sup>, which is best known for its application in the clinical treatment of adult acute non-lymphocytic leukaemia.

Recently, several classes of APN inhibitors have been reported<sup>[26,27]</sup>, including hydroxamic, sulphonic and boronic acid peptide analogues. Recent reviews have provided an excellent compendium of the current available set of APN inhibitors<sup>[28-29]</sup>. While many of these compounds display high *in vitro* potency, their selectivity has unfortunately not always been well documented, and *in vivo* data remains scarce. Therefore, we recently designed 7-amino-benzocycloheptan-6-one derivatives (Scheme 1) as novel lead structures for selective inhibition of APN<sup>[30,31]</sup>.

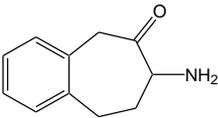
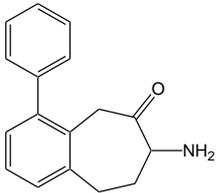
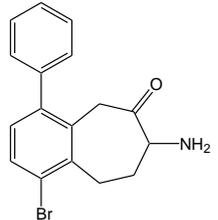


**Scheme 1:** Structure of the new synthesized aminobenzosuberone derivative inhibitors.

These new chemotypes are useful biochemical tools that can be used to characterise new biological functions for APN's catalytic activity. Thus, this work is a continuation of recently published studies<sup>[30,31]</sup>; the purpose of this study was to evaluate the importance of APN in processes that promote cancer progression using standardised *in vitro* and *in vivo* assays. Three compound derivatives with various  $K_i$  values were selected to establish a Structure-Activity Relationship (SAR), thus validating APN as the target. To further evaluate the *in vivo* activity of these compounds, nude xenograft mice were used. The  $K_i$  values of these compounds were evaluated on purified APN of both human and mouse origin because tumorigenic processes such as angiogenesis require interplay between malignant cells (i.e., xenografted human cells) and non-malignant (i.e., mouse) cells within the micro-environment<sup>[14,15]</sup>.

P15415 334-416	PKSDQIALPDFNAGAMENWGLVITYRENALLFDPOSSSISNKERVVTVIAHELAHQWFGNLVTLAWWNDLWLN <del>EG</del> FASSYVEYLG
P15144 339-421	PKSDQIGLPDFNAGAMENWGLVITYRENSLLFDPLSSSSSNKERVVTVIAHELAHQWFGNLVTIEWWNDLWLN <del>EG</del> FASYVEYLG
P97449 338-420	PKSDQIALPDFNAGAMENWGLVITYRESSLVFDSQSSSISNKERVVTVIAHELAHQWFGNLVTVAWWNDLWLN <del>EG</del> FASYVEYLG

*Figure 1: Active site sequence alignment for human, porcine and mouse APN (P15145: Porcine; P15144: Human; P97449: Mouse). The alignment was produced using the Jalview program (<http://www.jalview.org>)*

Ki values and Ligand Efficiency (LE)			
Inhibitor	Porcine APN	Human APN	Mouse APN
<p><b>I1</b></p> 	<p>Ki=1±0.5 μM*</p> <p>LE=0.63**</p>	<p>Ki= 2±0.71μM</p> <p>LE=0.61</p>	<p>Ki=2±1.13μM</p> <p>LE=0.61</p>
<p><b>I2</b></p> 	<p>Ki=7±1.5 nM</p> <p>LE=0.60</p>	<p>Ki= 11±2.4nM</p> <p>LE=0.58</p>	<p>Ki=10±6.2nM</p> <p>LE=0.58</p>
<p><b>I3</b></p> 	<p>Ki=0.06±0.02 nM</p> <p>LE=0.72</p>	<p>Ki= 0.35±0.24nM</p> <p>LE=0.66</p>	<p>Ki=0.2±0.29nM</p> <p>LE=0.67</p>

\*Ki values (n=3).

\*\* Ligand Efficiency (LE=ΔG/N, ΔG=-RTln Ki or IC<sub>50</sub>, Kcal/mol per non-hydrogen atom)

*Table 1: Ki values and Ligand Efficiency (LE) for selected benzosuberone derivatives on human, porcine and murine APNs.*

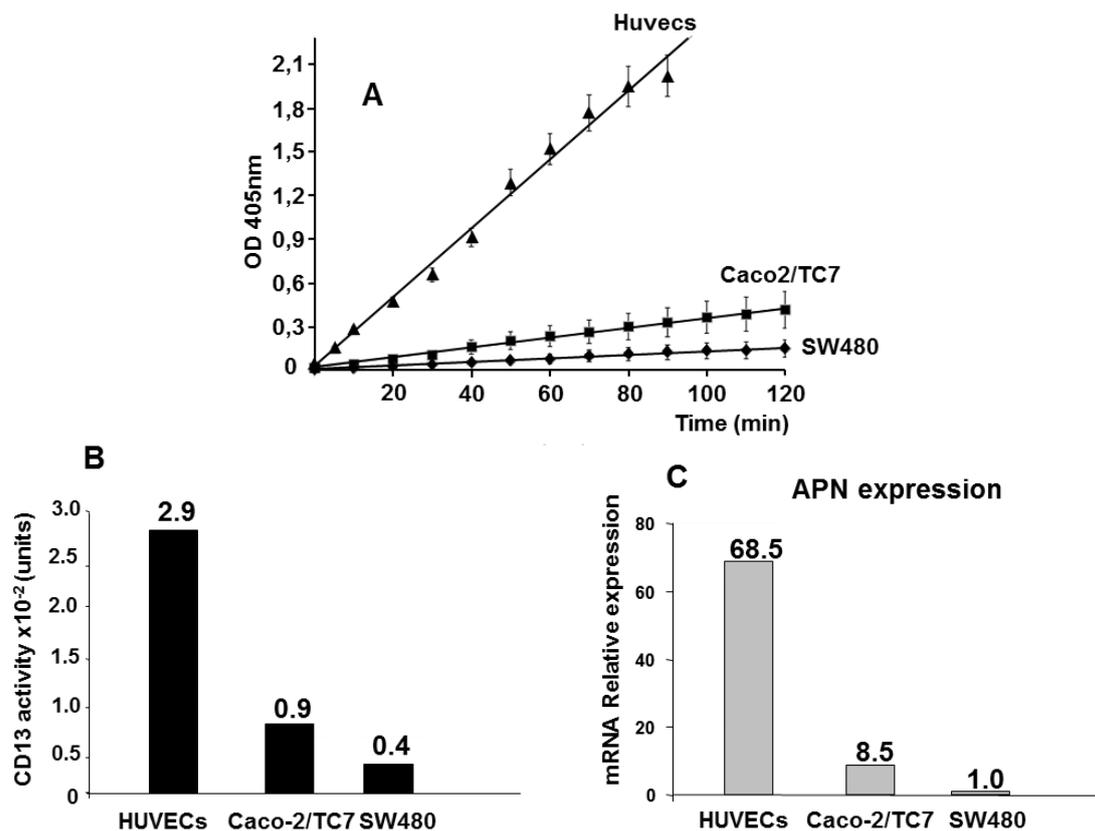
To explore the biological functions of APN, particularly those impacting cell migration and morphogenesis, a series of standard cell-based assays, including in vitro endothelial capillary tube formation and endothelial and tumour cell migration assays, were conducted. Pilot in vivo studies were then conducted using the most promising inhibitors and various tumour types expressing APN. We hypothesised that rapamycin, a well-tolerated macrocyclic lactone and potent mTOR inhibitor/immunosuppressive agent, might cooperate with an APN inhibitor.

## 2. Results and Discussion

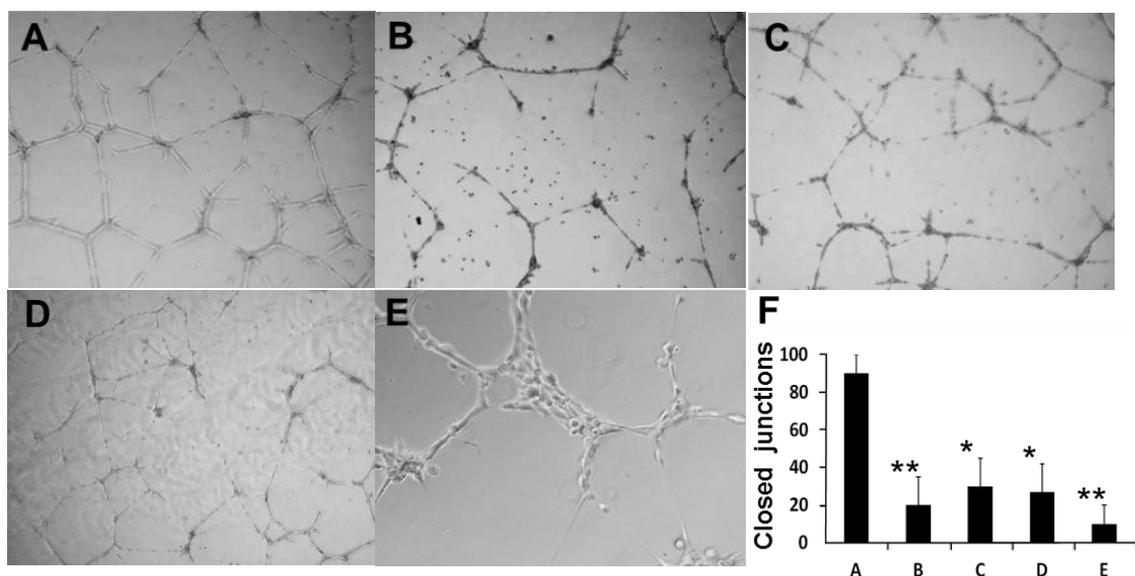
### 2.1. In vitro affinity of APN inhibitors for human and mouse aminopeptidases

Three compounds (I1, I2, I3) with various affinity constants were chosen to investigate a possible structure-activity relationship and dose-dependent response in standard in vitro and in vivo models of tumourigenesis. Synthesis of the inhibitors and their effects on the enzymatic activity of porcine kidney APN have already been published<sup>[30,31]</sup>. In this work, the potency of our new scaffolds was determined for human and mouse APN prior to studies in biological models. The sequence homology of APN from the three species is high; 78% identity was observed between the human and porcine sequences and 76% between the murine and human sequences (<http://www.uniprot.org/>; Sequence Similarity Search (BLAST)). Partial sequence alignments corresponding to the catalytic domain of human (P15144), porcine (P15145) and mouse (P97449) APN are shown in Figure 1. The active sites for all species are highly conserved (98% homology); the consensus zinc binding motif “HEXXH-(X18)-E”, the exopeptidase motif “GXMEN” that binds the free primary amino group of the N-terminal residue of the peptide substrate and the catalytic proton sites are identical.

The sequence similarities explain the close  $K_i$  and Ligand Efficiency (LE) values of these enzymes, which were determined based on purified enzymes (Table 1). I1 and I2 demonstrated very similar activities on enzymes from all three species, with an identical structure-activity relationship. I3 was the most potent compound, with  $K_i$  values of  $0.35 \pm 0.4$  and  $0.2 \pm 0.29$  nM from human and mouse APN, respectively.



**Figure 2:** APN activity and expression. (A) Catalytic activity of APN in SW480, Caco2/TC7 and HUVECs. APN catalytic activity is defined as the fraction of leucine-p-nitroanilide hydrolysis inhibited by the presence of 1 $\mu$ M I3. (B) Quantification of APN activity as enzymatic units per 10<sup>6</sup> cells. (C) APN mRNA expression in SW480 and Caco-2/TC7 cell lines.



**Figure 3:** Capillary tube formation by HUVEC cells on Matrigel under normoxic conditions. (A) Control without inhibitor, (B) Bestatin at a concentration of 300 $\mu$ M (Ki=3.5 $\mu$ M), (C) I1 at concentration of 10 $\mu$ M (Ki=6 $\mu$ M), (D) I2 at a concentration of 1 $\mu$ M (Ki=56nM) and (E) I3 at a concentration of 1 $\mu$ M (Ki=0.35nM). (F) Percent (%) of closed capillary junctions calculated from pictures (A) to (E). \* p<0.05; \*\* p<0.01

The slight changes in the  $K_i$  values for I3 could be a result of subtle differences in the active site; these variations have not yet been identified because the three-dimensional structure of mammalian APN is unresolved. When evaluated *in vitro*, using Human Umbilical Vein Endothelial Cells (HUVECs) and two colon cell lines (Caco2/TC7 and SW480), the  $K_i$  values were very close to the value of the purified human enzyme. The enzyme activity varied according to the cell type (HUVECs > Caco2/TC7 > SW480) (Figure 2A-B) in a manner that correlated with the APN mRNA expression level (Figure. 2C). Finally, commercially available matrix metalloproteinases (MMPs 2, 9, 11, 12 and 13) were not inhibited by I1, I2 and I3 (at concentrations of up to 50 $\mu$ M, data not shown). All of these results suggest that versions of APN derived from different species and cell types can be significantly inhibited *in vitro* and *in vivo* without inhibiting MMPs.

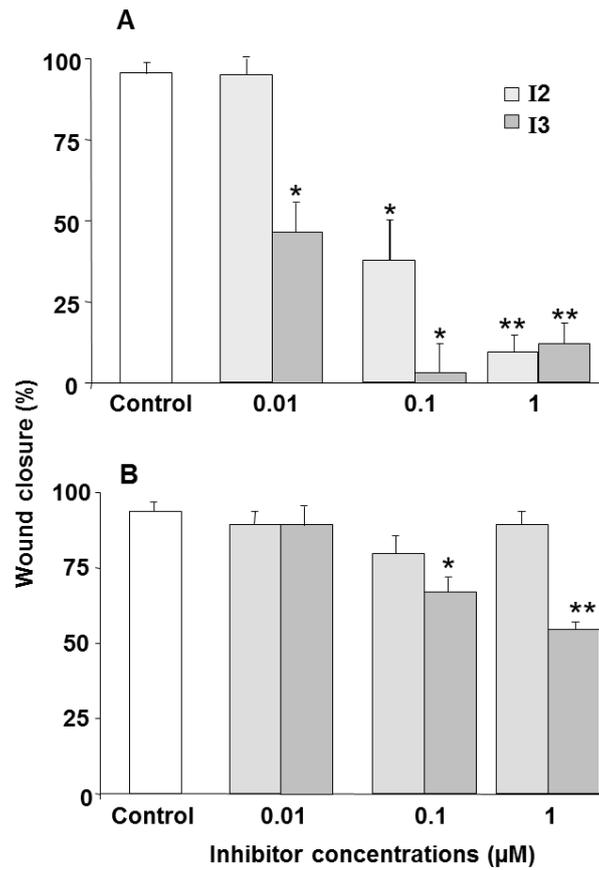
## 2.2. *In vitro* cell-based assays

Several reports suggest that APNs modulate angiogenesis and tumour dissemination; thus, capillary tube formation and cell motility were investigated.

### 2.2.1. Capillary tube formation

We used HUVECs to investigate whether the APN inhibitors inhibited the ability of endothelial cells to spontaneously form three-dimensional capillary tubes.

I1, I2 and I3 inhibited the formation of capillary-like structures, and the few tubes that formed were particularly thin (Figure 3). As expected, the concentrations of I1 and I2 necessary to inhibit this process correlated with their respective  $K_i$  values for APN, indicating that APN participates in the process of new blood vessel formation. Interestingly, I1 ( $K_i=1\mu$ M) is as potent as bestatin, a non-selective standard inhibitor ( $K_i=3.5\mu$ M), but I1 was 70 times more active in this assay. In our experimental conditions, a bestatin concentration of 300 $\mu$ M was necessary to inhibit endothelial tubulogenesis, but only 10 $\mu$ M of I1 and 1 $\mu$ M of I2 and I3 were able to produce a similar effect (60% inhibition of closed junctions). These doses were not cytotoxic; the  $CD_{50}$  values were 100 (I1), 20 (I2) and 10 $\mu$ M (I3). These data corroborate the results of previous studies<sup>[32]</sup> in which APN active site-directed inhibitors, such as bestatin and antibodies, interfered with tube formation but not with the proliferation of primary vascular endothelial cells, suggesting that APN contributes to morphogenesis during vascular tube formation.



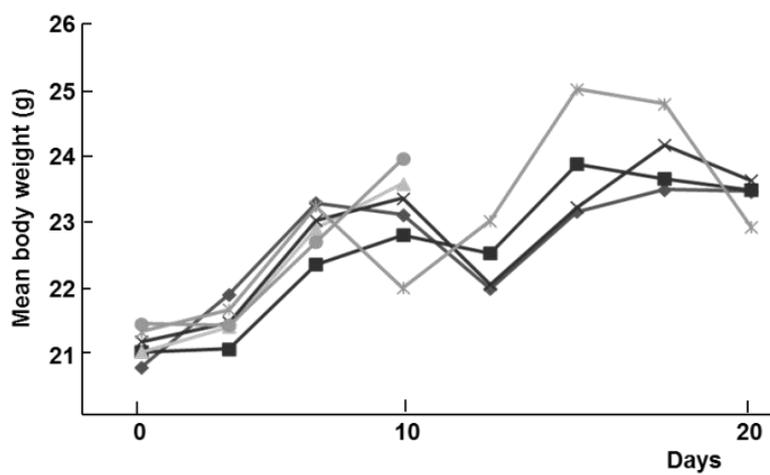
*Figure 4: Wound-healing assay. (A) HUVECs and (B) Caco2/TC7 cells in the presence of I2 and I3 following a 16h incubation. \*  $p < 0.05$ ; \*\*  $p < 0.01$ .*

### 2.2.2. Motility of HUVECs and human colon cells

To investigate whether APN inhibitors affect cell migration, inhibitors were tested in a wound-healing assay. I3 significantly inhibited HUVEC migration at 0.01 $\mu$ M, whereas higher concentrations of both I2 and I3 clearly reduced cell motility (Figure 4A). For Caco2/TC7 cells, I2 had no significant effect at concentrations up to 1 $\mu$ M, but cells treated with 0.1 $\mu$ M or 1 $\mu$ M I3 migrated 30% and 40% more slowly, respectively, than control cells (Figure 4B). We concluded that I3 and, to a lesser extent, I2 inhibit cell migration in a dose-dependent manner, although the inhibitors had a larger effect on endothelial cells than on colon cancer cells. The ability of the inhibitors to reduce cell migration may also be related to the level of APN expression, as HUVECs express more of this enzyme than Caco/TC7 tumour cells (Figure 2). These data are in agreement with data from a previous study demonstrating that levels of APN were correlated with the reduction in cell migration observed following addition of APN inhibitors (APN-targeted siRNA or specific chemical inhibitors) <sup>[33]</sup>. Previous reports have shown that APN increases the motility of several human carcinomas <sup>[16,34,35]</sup> and highly invasive cells express high levels of APN <sup>[35]</sup>. Similarly, we have also observed that I2 and I3, at 0.1 and 1  $\mu$ M, respectively, significantly decrease HUVEC motility by 70%, as observed in a Boyden chamber migration assay with VEGF as the chemo-attractant (data not shown).

### 2.3. Formulation and stability studies

A preclinical formulation, compatible with intraperitoneal administration (i.p.), was developed for I2 and I3. DMSO was chosen as the vehicle to allow for solubilisation of the compounds at a concentration of 5mg/ml in a 5% glucose solution at a ratio of 10/90(v/v). HPLC-MS analysis of storage solutions at -20°C, 4°C and 25°C revealed that the compounds were chemically stable for at least 20 days (at -20, 4 and 25°C). The chromatograms indicated the presence of a single peak for each substance, regardless of the storage conditions (data not shown). Nevertheless, slight precipitation of the compounds appeared over time (24 hours), likely due to the slow formation of a dimeric structure that has already been observed at high concentration <sup>[30]</sup>. Therefore, a fresh solution was prepared from the solid before each injection.



-◆- I2 40mg/kg-Q3D8; -■- I2 70mg/ kg -Q3D8, -▲- I2 100mg/ kg -Q3D4  
 -×- I3 40mg/ kg -Q3D8; -\* - I3 70mg/ kg -Q3D8, -●- I3 100mg/ kg -Q3D4.

*Figure 5: Mean body weights during MTTD determination.*

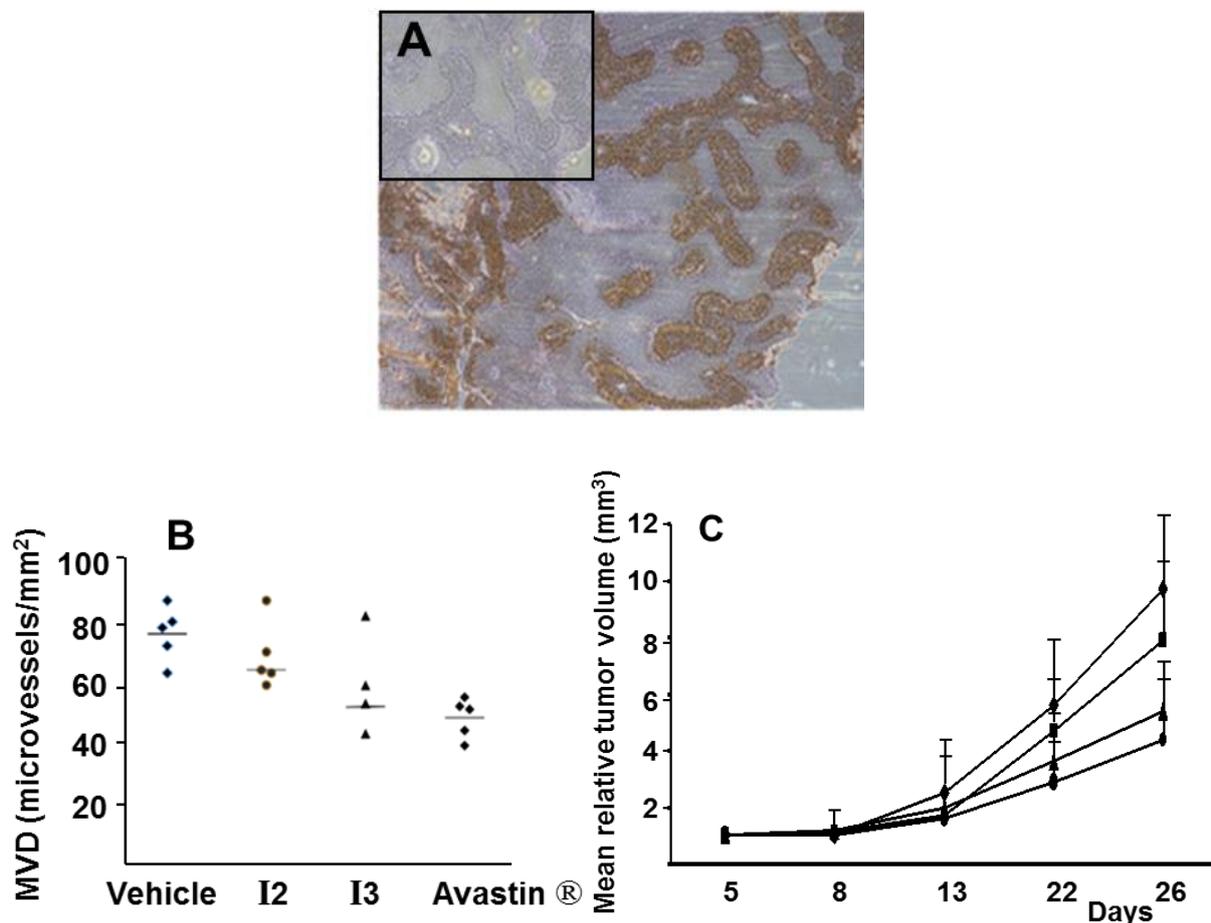
## 2.4. Tolerance study in nude mouse

Toxicity of these new chemical entities was carefully determined in nude mice. The maximum tolerated dose (MTD) and the maximum total tolerated dose (MTTD) were evaluated.

MTD determination: 18 mice were randomised into 6 groups of 3 animals, according to their individual body weight ( $21.2 \pm 1.4\text{g}$ ) on Day 0 (D0). In the I2-treated group, 1 mouse displayed a slight, transient body weight loss between D0 and D2, while the other 2 mice gained weight. All of the I2-treated mice gained weight after D2. The weight loss effect was more pronounced when a dose of 100mg/kg was injected, and all of the mice displayed the transient body weight loss in the first 2 days. At a dose of 150mg/kg, 1 mouse was found dead on D1, and the other 2 mice did not gain weight until day 5 and continued to gain weight until day 11. Similar effects were observed with I3-treated mice. Under these conditions, the MTD was determined to be 100mg/kg.

MTTD determination: Randomisation of 18 mice into 6 groups of 3 animals was conducted according to individual body weights ( $21.2 \pm 0.7\text{g}$ ). Mice were treated on D0 with 40, 70 and 100mg/kg per injection (I2 or I3). Injections were given every 3 days for a total of 8 injections (Q3D8) for the first two dosages and 4 injections for the highest dose. The 40mg/kg/injection dose of I2 and I3 was well tolerated, and no macroscopic sign of toxicity was observed. After the second 70mg/kg I2 injection, a transient catalepsy persisting for two hours was observed. Nonetheless, all of the mice gained weight until the end of treatment. At the highest treatment dose (100mg/kg/injection), convulsions were observed after the second injection, and mice were sacrificed after the third injection on D10 for ethical reasons.

The most active APN inhibitor, I3, was less well tolerated at 70mg/kg/injection because 1 mouse was found dead on day 3, and the other 2 mice displayed marked body weight loss during the course of treatment (Figure 5). Thus, the MTTD was estimated as 40mg/kg for both APN inhibitors when one injection was given every 3 days.

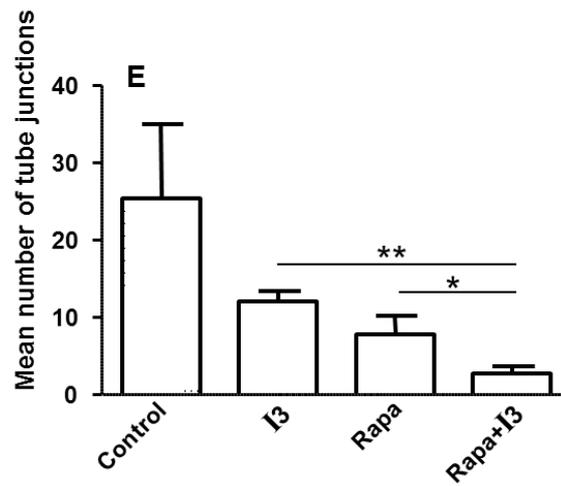
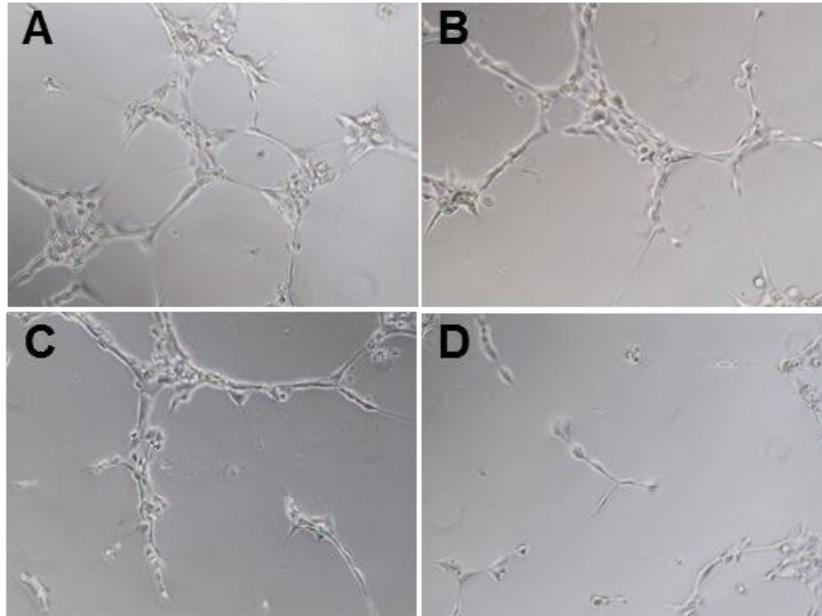


*Figure 6: A-673 rhabdomyosarcoma cell line-derived xenografts. (A) Immunohistochemical expression of CD13 (insert: control without the primary anti-APN antibody). 400X magnification. (B) Vascularisation analysis of A-673 tumours at the end of treatment. (C) Effects of drug treatment on the growth of A-673 tumours subcutaneously xenografted in Swiss nude mice (n=5). Mice were xenografted on D0 via subcutaneous injection of 107 A-673 tumour cells. On D1, mice received one i.p. injection of vehicle (◆), I2 at 40 mg/kg/injection (■) or I3 at 30 mg/kg/injection (▲) every three days for a total of 8 injections, or one i.p. injection of Avastin® (●) at 5 mg/kg/injection twice a week for a total of 7 injections.*

## 2.5. Effects of APN inhibitors on the growth of xenografted tumors

APN is clearly deregulated in human malignancy and confers carcinogenic properties. Therefore, the development of novel anticancer drugs modulating APN expression/function or related signalling pathway(s) could improve treatment options. APN is considered to be a target for anti-angiogenic therapies,<sup>[5]</sup> and although convincing in vitro data supporting this idea are available, in vivo data with specific inhibitors remained scarce<sup>[14,15,20]</sup>. A single aminopeptidase inhibitor (Tosedoxat®, a hydroxamic acid peptide derivative) is currently being evaluated in clinical trials for acute myeloid leukaemia, on the basis of an amino acid deprivation response<sup>[36,37]</sup>.

APN inhibitors also alter the angiogenic process; here, we compared the efficacy of I2, I3 and the anti-angiogenic anti-VEGF antibody Avastin®<sup>[38]</sup> in a xenograft model. We selected a model where the activity of this antibody could be easily observed. Avastin (5mg/kg twice a week for a total of 7 injections) and I2 and I3 (40 mg/kg and 30mg/kg, respectively) were evaluated in a model in which A672 rhabdomyosarcoma cells were implanted into nude mice; these cells express APN (Figure 6A). The mice were treated one day after tumour cell injection and for an additional 26 days. By 26 days, the tumour size was reduced by 55% upon treatment with Avastin® and 44% with I3 (30 mg/kg); the tumours responded poorly to the 40mg/kg dose of I2 (18% growth inhibition) (Figure 6B). Post-treatment tumour sections were analysed and revealed that the density of microvessels was closely related to the impact of each drug on tumour growth. Avastin®, which exhibited the highest growth inhibitory effect, induced a reduction in microvessel density. The effect of I3 on tumour growth and vascularisation of the tumour was similar to that observed with Avastin. I2, which weakly decreased the tumour size, was inefficient at reducing the vessel density within the tumour (Figure 6C). These results indicated that APN inhibitors could behave as antiangiogenics during the initial stages of tumour growth.



*Figure 7: Effect of I3 and rapamycin on HUVEC capillary tube formation. (A) Control without inhibitor. (B) I3 at 1 $\mu$ M. (C) Rapamycin at 20 nM. (D) Rapamycin at 20nM and I3 at 1 $\mu$ M. (E) Quantification of the number of tube junctions according to the treatment conditions. \*  $p < 0.05$ ; \*\*  $p < 0.01$*

## 2.6. APN inhibitors in drug combination

In addition to cancer prevention, an urgent goal of cancer therapy is to treat established tumours and/or prevent the spread of metastases. Therefore, the APN inhibitors were tested on established xenograft tumours, alone and in combination with rapamycin.

To identify a cell type that would respond to APN inhibitors, we focused on human colon tumours because their growth relies on angiogenesis<sup>[39]</sup> and because they express APN at various levels. I3 was tested in combination with rapamycin, an inhibitor of the mTOR pathway; this pathway is associated with tumour vascularisation, cell growth, proliferation and survival<sup>[40]</sup>.

### 2.6.1. Capillary tube formation

Rapamycin inhibits vascular tube morphogenesis<sup>[41]</sup>, and we observed that HUVECs treated with both I3 (1 $\mu$ M) and rapamycin (20nM) did not develop significant self-organised capillary-like structures, as expected for the combination of these two efficient compounds (Figure 7). This result confirmed the impact of combined therapy and encouraged additional *in vivo* investigation.

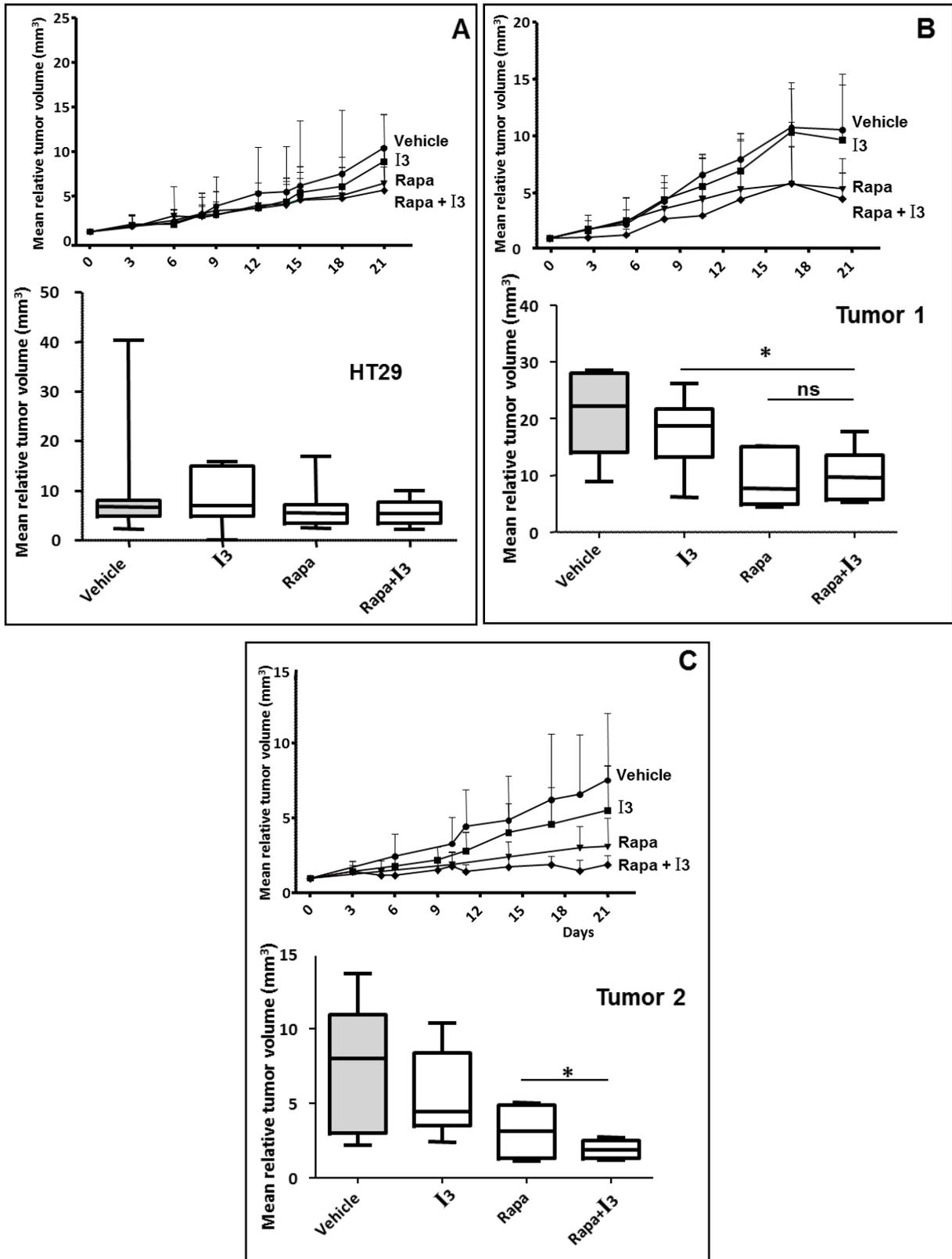
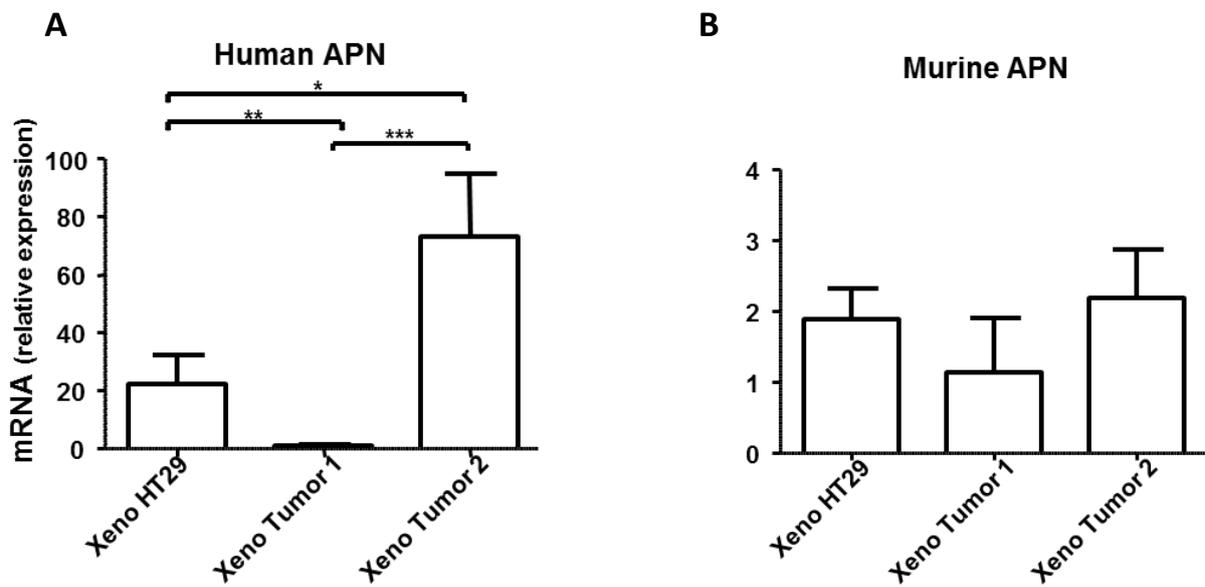


Figure 8: Efficacy of I3 and rapamycin treatment in vivo. Effects of I3 (30 mg/kg) with or without rapamycin (3.3 mg/kg) on the growth of (A) HT29 cell line, (B) colon tumour 1 and (C) colon tumour 2 xenografted in nude mice. \*  $p < 0.05$ .

### 2.6.2. *In vivo* assays

In these pilot experiments, the anti-tumour properties of I3 injected at 30mg/kg (Q3D8) were assessed in xenografts of primary human colon tumours and tumours derived from the human HT29 cell line.

No significant effect on HT29 tumour growth was observed with I3 (15% TGI; Figure 8A). For the primary human colon tumour xenografts, tumour 1 was less sensitive to the APN inhibitor than tumour 2 (8% TGI for tumour 1 and 26% TGI for tumour 2; Figure 8B-C). This effect was independent of the median tumour volume at the start of treatment.



*Figure 9: Relative mRNA expression of APN in human tumour cells and stromal murine cells in xenografts. mRNA was quantified via relative quantitative PCR using primers specific for human and murine APN mRNA. \*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ .*

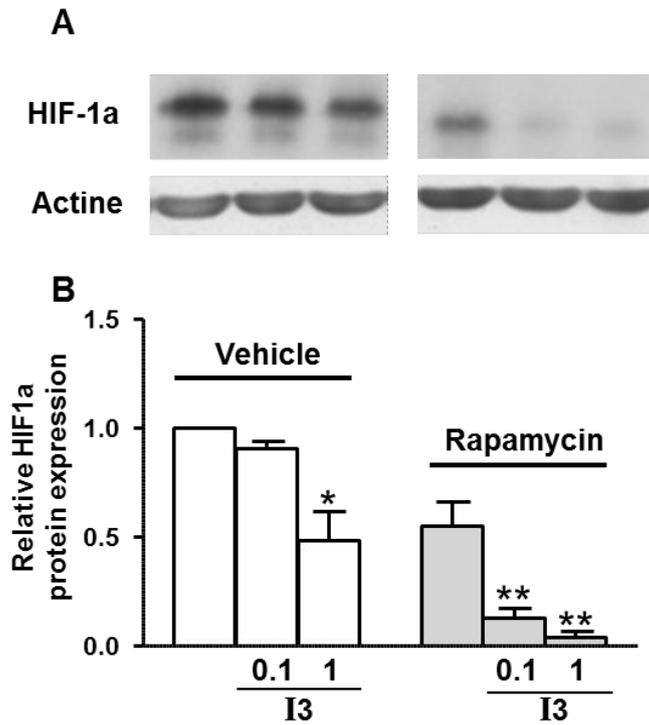
As expected, the inhibitory impact of the APN inhibitor on tumour growth was related to the endogenous expression of APN. Xenografts derived from HT29 cells and tumour 1 had low levels of human APNmRNA expression, whereas tumour 2 expressed human APN mRNA at a high level (70 times more than in tumour 1 and 3.5 times more than in HT29 cells; Figure 9). It is important to note that the xenografts were colonised by the murine vasculature; APN expressed in the host endothelial cells may be involved in the effect. However, the level of murine APN mRNA did not differ among the three samples (Figure 9) most likely because the rate of vascularisation was comparable in the three tumour types. The observed effect may be due to the high expression of APN in the tumour cells.

Although these observations suggest that the efficacy of the inhibitors depends on the level of APN expression, as shown previously<sup>[33]</sup>, the ability of these molecules to inhibit tumour growth was rather weak. Therefore, a combination of these inhibitors with other therapeutic drugs may be necessary to further improve the anti-tumour effect.

A major mechanism mediating adaptive responses to tumour hypoxia is the regulation of genes playing key roles in cancer biology, including Hypoxia-Inducible Factor 1 (HIF-1 $\alpha$ )<sup>[43]</sup>, which regulates angiogenesis<sup>[42]</sup>. Consequently, molecules that modulate the transcriptional activity or the quantity of HIF-1 $\alpha$  have been shown to inhibit the proliferation and/or vascularisation of tumours<sup>[44]</sup>. Published data have demonstrated<sup>[44]</sup> that mTOR inhibitors suppress the steady-state levels of HIF-1 $\alpha$ <sup>[45]</sup>, thus preventing angiogenesis<sup>[41]</sup>. Therefore, we addressed whether rapamycin could enhance I3-mediated tumour growth inhibition.

As anticipated, we observed that rapamycin significantly amplified the anti-tumour effect of the I3 inhibitor only for tumour 2, which expressed a high level of APN (TGI = 57% for tumour 1 vs. 8% with I3 alone; 75% for tumour 2 vs. 26% with I3 alone; Figure 8B-C). Our study strongly suggests that combined therapy with rapamycin and I3 could be a novel approach for the treatment of colon tumours expressing high levels of APN.

Although the combined treatment strongly inhibited *in vitro* capillary tube formation, *in vivo* studies seemed to indicate that the weak but significant effect on tumour growth was likely related to the APN expression level in tumour cells rather than the expression level in endothelial cells. This gap between *in vitro* data obtained in cell lines and the effects observed *in vivo* has been previously described<sup>[46]</sup>. This effect must be thoroughly studied in other tumour types.



*Figure 10: Effects of I3 on protein expression in Caco2/TC7 cells. (A) Western blots for HIF-1a expression. Actin served as the control. (B) Relative changes in the expression of HIF-1a after treatment with I3 (0.1 and 1μM), rapamycin (20nM) or a combination of I3 (0.1 and 1μM) and rapamycin (20nM) for 48h, compared to vehicle-treated cells. \* $p < 0.05$ ; \*\*  $p < 0.01$*

## 2.7. APN inhibition and HIF-1 $\alpha$ expression

APN's mechanisms of action are unknown because it is difficult to identify endogenous substrates of this ubiquitous ectoenzyme. The previously demonstrated biological effects of our specific APN inhibitors on cell migration and tumour growth provided the opportunity to collect data on possible mechanistic events. Due to the known functions of APN in tumourigenesis<sup>[3,4]</sup> and the benefit of rapamycin in combination with the aminobenzosuberone derivatives on tumour growth, the idea that APN regulates mTOR/HIF-1 $\alpha$  pathway was considered. It is well established that HIF-1 $\alpha$  acts as a master transcriptional regulator of the adaptive response to hypoxia and that mTOR inhibition induces the inactivation of HIF-1 $\alpha$  mRNA translation, reducing HIF-1 $\alpha$  protein expression. Subsequently, reduced HIF-1 $\alpha$  transcriptional activity decreases tumour vascularisation, leading to growth arrest. In this work, the I3-induced inhibition of Caco2/TC7 cell migration clearly involved HIF-1 $\alpha$  because its expression decreased with increasing concentrations of I3. Interestingly, the combination of rapamycin and I3 nearly abolished HIF-1 $\alpha$  expression (Figure 10). Therefore, this study demonstrated that I3 targets HIF-1 $\alpha$  and that combining I3 with rapamycin greatly enhances I3's anti-tumour effect in human colon cells.

This type of mechanism has already been proposed in melanoma cells, where HIF-1 $\alpha$  and APN levels are increased following hypoxia; both mRNA and protein levels were elevated<sup>[47]</sup>. However, this is the first insight into a complete mechanism.

Interestingly, inhibition of HIF-1 $\alpha$  protein accumulation in Caco2/TC7 cells following I3 treatment occurred via mTOR-independent mechanisms because the levels of phosphorylated S6RP, an mTOR target, remained unchanged in cells treated with the APN inhibitor (data not shown). Modulation of the mTOR pathway by aminopeptidases has already been described and is cell type dependent. Aminopeptidase inhibition by non-specific inhibitors such as CHR-2797 reduced both mTOR substrate phosphorylation and the rates of protein synthesis in HL-60 cells but not in HuT 78 cells, which are another haematopoietic cell line<sup>[26]</sup>.



Furthermore, rapamycin treatment reduced HIF-1 $\alpha$  expression, an effect that involved the mTOR pathway, as shown by reduced phosphorylation of S6RP following rapamycin treatment. The ability of both the APN inhibitor and rapamycin to decrease HIF-1 $\alpha$  protein levels via independent pathways may explain their combined effect and may bring new opportunities for developing therapeutic strategies. Similar additive effects were shown in a previous report that combined rapamycin and irinotecan <sup>[48]</sup>.

### 3. Conclusion

We have shown that aminobenzosuberone derivatives are potent nanomolar inhibitors of human and mouse APN and that they interfere *in vitro* with cellular processes involved in cancer progression. The level of inhibition correlated with the Ki value of the tested compounds, and dose-dependent effects were demonstrated in standard *in vitro* cell migration assays. Pilot *in vivo* studies were performed, and toxicity was determined. Inhibitors were well tolerated in mice after repeated i.p. injections. Moreover, I3 and Avastin impaired tumour growth to a similar extent, which correlated with their effects on tumour vascularisation in a rhabdomyosarcoma xenograft model. This effect was observed only when treatments occurred one day after tumour cell injections. The effects on established colon tumour growth can be correlated with APN expression, and combined therapies with rapamycin were beneficial. We showed that APN clearly participates in cancer progression, presumably through regulation of HIF-1 $\alpha$ . We also observed that the antitumour effect of the inhibitors relied on APN expression in tumour cells but not on angiogenesis. Further investigations based on these promising results should generate more pharmacological data, including optimal dosing regimens and modes of administration. These data will improve the overall treatment schedules and will allow for the optimisation of aminobenzosuberone derivatives in cancer therapies. Moreover, mechanistic studies will be pursued.

### 4. Experimental Section

#### 4.1. Cell cultures

The *in vitro* tests were carried out in Human Umbilical Vein Endothelial Cells (HUVECs) and human colon carcinoma cells (CaCo2/TC7, SW480 and HT-29). HUVECs (Promocell C-12200) were grown in endothelial cell growth medium (Promocell, C-22010), and

CaCo2/TC7, HT29 and SW480 cells were maintained in DMEM (1g/l glucose) supplemented with 10% fetal calf serum (FCS) and 200U/ml penicillin-streptomycin (Invitrogen). All cell lines were incubated at 37°C and 5% CO<sub>2</sub>. The cells were treated during exponential growth conditions at a confluence of 30%.

## 4.2. Enzyme assays

Human and mouse recombinant APN (R&D Systems) and porcine kidney APN (Sigma Chemical Co.) were used. The data were obtained using an HP/Agilent UV-Visible diode array spectrophotometer 8453 (HP ChemStation software). L-leucine p-nitroanilide ( $K_m=0.2$  mM) was used as the substrate for APN. The tests were conducted on  $1.8 \times 10^5$  HUVEC cells or  $5 \times 10^3$  Caco2/TC7 or SW480 cells grown in 96-well plates. Tests were carried out at 30°C in 10mM Tris-HCl (pH 7.5). The reaction was initiated following the addition of 1mU of porcine APN, 1mU or 0.5mU of human APN and 1 mU of mouse APN. The release of p-nitroanilide ( $\epsilon=10.8$  M<sup>-1</sup>cm<sup>-1</sup>) at 405nm was continuously measured every 10 minutes for 2 hours with a spectrometer (PerkinElmer) to obtain the initial velocities.  $K_i$  was calculated via Dixon plots<sup>[49]</sup>.

## 4.3. Capillary tube formation assay

The capillary tube formation assay was conducted in a 24-well plate coated with Matrigel (BD Bioscience). HUVECs ( $5 \times 10^4$  cells) were seeded and grown in Endothelial Basal Medium supplemented with 0.1% BSA, 1% FCS, 1% DMSO and bestatin (300µM); in the experimental conditions, I2 (10 µM) or I3 (0.2 µM) APN inhibitors were added. After 24 hours, tube formation was periodically observed under a phase contrast microscope (Leica, Germany).

## 4.4. *In vitro* migration assay

### 4.4.1. Boyden Chamber

HUVECs were cultured in Endothelial Basal Medium supplemented with 0.1% BSA and 1% FCS for 5 hours prior to use.  $10^5$  HUVECs were seeded in the upper part of BD Falcon™ Cell Culture Inserts filters (3µm size, BD Biosciences, Mississauga, Canada) in media supplemented with 0.1% BSA and 1% FCS. The lower part of the chamber was filled with media complemented with 100ng/ml VEGF. Inhibitors at various concentrations (0.1 and 0.01µM) were added to both the upper and lower parts of the chamber. After 22 hours of

incubation at 37°C in 5% CO<sub>2</sub>, media were removed from the inserts, and filters were washed with PBS. Plates were then incubated with Calcein (4µg/ml) for one hour at 37°C and then for 30 minutes at room temperature. Finally, the fluorescence was measured (495nm excitation, 515nm emission on a PerkinElmer Victor).

#### 4.4.2. Wound healing assay

Migration potential was assessed using Ibidi cell-culture inserts (Ibidi #81176) according to the manufacturer's protocol. Briefly,  $2.8 \times 10^3$  Caco2/TC7 cells and  $10^5$  HUVECs were seeded in two compartments separated by a 500µm thick wall. When confluence was reached 3 days later (T0), the insert was removed, and cells from each compartment were allowed to migrate in a directed manner. The area between the cell patches was measured in 4 fields per dish at T0, T16 h and T24 h. Treatment with APN inhibitors at 0.01, 0.1 and 1µM or vehicle (DMSO, 0.25% final) was initiated by adding inhibitors (or vehicle) in DMEM containing 1g/l glucose and 1% BSA at T0. Migration was calculated as the surface colonised by cells from T0 to T16 h or T24 h, and results were expressed as the percentage of migration compared to control cells.

#### 4.5. *In vivo* toxicity tests

The toxicity of the APN inhibitors was evaluated by determining the maximum tolerated dose (MTD) and the maximum total tolerated dose (MTTD) on Swiss nude mice. For MTD determination, 6 groups of 3 mice were randomly assigned on day 0, and mice were injected with either I2 or I3 at 50, 100 or 150mg/kg in a volume of 30ml/kg. For MTTD, identical groups were used (6 groups of 3 mice). These mice received an i.p. injection of either I2 or I3 every 3 days (for a total of 8 injections) using 40, 70 or 100mg/kg/inj in a volume of 30ml/kg.

For both MTD and MTTD tests, behavioural and clinical signs were monitored each day, and the mice were weighed twice a week.

Mice were sacrificed on day 11 for MTD tests and day 29 (8 days after the last injection) for MTTD tests. The heart, lungs, spleen, liver, kidneys and gastro-intestinal tract were microscopically examined. The same examinations were also carried out on mice that died during the test period.

#### 4.6. Immunohistochemical staining

Human tissue collection and xenograft experiments were approved by the CCPPRB d'Alsace n°1 Strasbourg ethics committee (agreement number 67-110).

Rhabdomyosarcoma A-673 human tumours were fixed in 10% neutral buffer formalin and embedded in paraffin. Immunohistochemical staining was conducted using the avidin-biotin immunoperoxidase technique (ABC method). The biotinylated secondary antibody is revealed by streptavidin (avidin), which is linked to the peroxidase. In the presence of 3,3'-diaminobenzidine (DAB) and hydrogen peroxide, the positive substrates of peroxidase are visualised according to brown colour. Rabbit polyclonal anti-human APN (H-300, sc-15360, Santa Cruz) and anti-mouse CD31 antibodies were used to determine the microvessel density in the tumours as described previously <sup>[50]</sup>. The most vascularised areas (5 zones of 0.78 mm<sup>2</sup> per histological section) were selected and then studied at 100X magnification (AxioVision software) by two independent experienced biologists.

#### 4.7. *In vivo* tumor xenograft experiments

Female athymic nude mice (nu/nu, 6-8 weeks old) were purchased and maintained under pathogen-free conditions in facilities approved by the French Ethical Committee and under the supervision of authorised investigators. Male athymic nude mice (nu/nu, 6–8 weeks old) were purchased (Charles River, L'Arbresle, France) and maintained under specific pathogen-free conditions. Tissue from human colon tumours was obtained from surgical specimens at the Department of Visceral Surgery (Hôpitaux Universitaires de Strasbourg) under the institutional review of the French ethics committee and with patient consent. Two human colon tumours were minced on ice and injected subcutaneously into the right and left flanks of nude mice as previously described <sup>[51]</sup>. HT29 human tumour cells were injected subcutaneously into 2 points/flank (4 tumours/mouse). The treatment was initiated when the tumours reached a mean volume of 150–300 mm<sup>3</sup>. The mice were randomised into 4 groups (3 mice/group) and treated with i.p. injections of 30mg/kg of I3 alone every 3 days, 3.3 mg/kg of rapamycin alone every 5 days, or both drugs. The control group mice were treated with both vehicles. The individual relative tumour volume (RTV) was defined as  $V_x/V_0$ , where  $V_x$  is the volume in mm<sup>3</sup> at a given time, and  $V_0$  is the volume at the start of treatment. Mean RTV and standard errors of the mean (SEM) were calculated for each group. The tumour growth inhibition was also calculated according to  $TGI = 100 - (RTV_{treated} / RTV_{control}) * 100$ .

## Acknowledgements

This work was supported by The Ligue Nationale contre le cancer, the Université de Strasbourg, the Hôpitaux Universitaires (Strasbourg), and the University of Haute-Alsace (Mulhouse). The authors wish to acknowledge Michèle Lavogez for assistance with animal care. The authors also acknowledge Elisabeth Martin, Christelle Lasthaus and Sophie Billiger for their excellent technical assistance. We thank the staff at Institut de Recherche en Hématologie et Transplantation (Mulhouse) for cell culture assistance.

## References

- [1] A.J. Barret, N.D. Rawling, J.F. Woessner, *Handbook of Proteolytic Enzymes 2<sup>nd</sup> ed., Vol.5*, Elsevier Academic Press, Oxford, **2004**, p. 233.
- [2] M. Drag, M. Bogyo, J.A. Ellman, G.S. Salvesen, *J. Biol. Chem.* **2010**, 285, 3310-3318.
- [3] P. Mina-Osorio, *Trends Mol. Med.* **2008**, 14, 361-371.
- [4] B. Bauvois, *Oncogene.* **2004**, 23, 317-329.
- [5] N. Haraguchi, H. Ishii, K. Mimori, F. Tanaka, M. Ohkuma, H.M. Kim, H. Akita, D. Takiuchi, H. Hatano, H. Nagano, G.F. Barnard, Y. Doki, M. Mori, *J. Clin. Invest.* **2010**, 120 3326-3339.
- [6] M. Wicksröm, R. Larson, P. Nygren, J. Gullbo, *Cancer Sci.* **2011**, 102, 501-508.
- [7] S. Carl-McGrath, U. Lendendeckel, C. Röcken C, *Histol. Histopathol.* **2006**, 21, 339-1353.
- [8] D. Rienmann, A. Kehlen, J. Langner, *Immunology Today.* **1999**, 2, 83-88.
- [9] D. Reinhold, U. Bank U, M. Täger, S. Ansorge, S. Wrenger, A. Thielitz, U. Lendeckel, J. Faust, K. Neubert, S. Brocke, *Front. Biosci.* **2008**, 13, 2356-63.
- [10] D. Reinhold, A. Biton, S. Pieper, U. Lendeckel, J. Faust, K. Neubert, U. Bank, M. Täger, S. Ansorge, S. Brocke, *Int. Immunopharmacol.* **2006**, 6, 1935-42.
- [11] R. Pasqualini, E. Koivunen, R. Kain, J. Lahdenranta, M. Sakamoto, A. Stryhn, R. Ashmun, L.H. Shapiro, W. Arap, E. Ruoslahti, *Cancer Res.* **2000**, 60, 722-727.
- [12] M. Yamashita, H. Kajiyama, M. Terauchi, K. Shibata, K. Ino, A. Nawa, S. Mizutani, F. Kikkawa, *Int. J. Cancer.* **2007**, 120, 2243-2250.
- [13] K. Fukasawa, H. Fujii, Y. Saito, K. Koizumi, Y. Aozuka, K. Sekine, M. Yamada, I. Saiki, K. Nishikawa, *Cancer Lett.* **2006**, 243, 135-143.
- [14] a) R. Rangel, Y. Sun, L. Guzman-Rojas, M.G. Ozawa, J. Sun, R.J. Giordano, C.S. VanPelt, P. T. Tinkey, R.R. Behringer, R.L. Sidman, W. Arap, R. Pasqualini, *Proc. Natl. Acad. Sci. U.S.A.* **2007**, 104, 4588-4593.
- [15] L. Guzman-Rojas, R. Rangel, A. Salameh, J.K. Edwards, E. Dondossola, Y.G. Kim, A. Saghatelian, R.J. Giordano, M.G. Kolonin, F.I. Staquicini, E. Koivunen, R.L. Sidman, W. Arap, R. Pasqualini, *Proc. Natl. Acad. Sci. U.S.A.* **2012**, 109, 1637-1642.
- [16] H. Hashida, A. Takabayashi, M. Kanai, M. Adachi, K. Kondo, N. Kohno, Y. Yamaoka, M. Miyake, *Gastroenterology.* **2002**, 122, 376-386.
- [17] A. Menrad, D. Speicher, J. Wacker, M. Herlyn, *Cancer Res.* **1993**, 53, 1450-1455.
- [18] Y. Mishima, Y. Matsumoto-Mishima, Y. Terui, M. Katsuyama, M. Yamada, M. Mori, Y. Ishizaka, K. Ikeda, J. Watanabe, N. Mizunuma, H. Hayasawa, K. Hatake, *J. Natl. Cancer Inst.* **2002**, 94, 1020-1028.
- [19] N. Petrovic, W. Schacke1, J. Reed Gahagan, C. A. O'Conor1, B. Winnicka1, R.E. Conway, P. Mina-Osorio, L.H. Shapiro, *Blood.* **2007**, 110, 142-150.
- [20] H. Fujii, M. Nakajima, I. Saiki, J. Yoneda, I. Azuma, T. Tsuruo, *Clin. Exp. Metastasis.* **1995**, 13, 337-344.
- [21] A. Kehlen, U. Lendeckel, H. Dralle, J. Langner, C. Hoang-Vu, *Cancer Res.* **2003**, 63, 8500-8506.
- [22] M. Wickström, R. Larsson, P. Nygren, J. Gullbo, *J. Cancer Sci.* **2011**, 102, 501-508.
- [23] B. Bauvois, D. Dauzonne, *Med. Res. Rev.* **2006**, 26, 88-130.
- [24] I. Ranogajec, J.J. Razumovic, V. Puzovic, J. Gabrilovac, *Med. Oncol.* **2012**, 29, 561-9.
- [25] M.B. Harbut, G. Velmourougane, G. Reiss, R. Chandramohanadas, D.C. Greenbaum, *Bioorg. Med. Chem. Lett.* **2008**, 18, 5932-5936.

- [26] D. Krige, L.A. Needham, L.J. Bawden, N. Flores, H. Farmer, L.E.C. Miles, E. Stone, J. Callaghan, S. Chandler, V.L. Clark, P. Kirwin-Jones, V. Legris, J. Owen, T. Patel, S. Wood, G. Box, D. Laber, R. Odedra, A. Wright, L. Michael Wood, S.A. Eccles, E.A. Bone, A. Ayscough, A.H. Drummond, *Cancer Res.* **2008**, *68*, 6669–6679.
- [27] Y. Liu, L. Shang, H. Fang, H. Zhu, J. Mu, Q. Wang, X. Wang, Y. Yuan, W. Xu, *Bioorg. Med. Chem.* **2009**, *17*, 7398–7404.
- [28] A. Mucha, M. Drag, J.P. Dalton, P. Kafarski, *Biochimie.* **2010**, *92*, 1509-1529.
- [29] F.E. Jacobsen, J.A. Lewis, S.M. Cohen, *Chem. Med. Chem.* **2007**, *2*, 152-171.
- [30] S. Albrecht, M. Al-Lakkis-Wehbe, A. Orsini, A. Defoin, P. Pale, E. Salomon, C. Tarnus, J.M. Weibel, *Bioorg. Med. Chem.* **2011**, *19*, 1434-1449.
- [31] C. Maieranu, C. Schmitt, N. Schifano-Faux, D. Le Nouën, A. Defoin, C. Tarnus, *Bioorg. Med. Chem.* **2011**, *19*, 5716-5733.
- [32] S.V. Bhagwat, J. Lahdenranta, R. Giordano, W. Arap, R. Pasqualini, L.H. Shapiro, *Blood.* **2001**, *97*, 652–659.
- [33] A. Lai, A. Ghaffari, A. Ghahary, *Mol. Cell Biochem.* **2010**, *343*, 191-199.
- [34] Y. Saitoh, K. Koizumi, T. Minami, K. Sekine, H. Sakurai, I. Saiki, *Biol. Pharm. Bull.* **2006**, *29*, 709-712.
- [35] Y.W. Chang, S.C. Chen, E.C. Cheng, Y.P. Ko, Y.C. Lin, Y.R. Kao, Y.G. Tsay, P.C. Yang, C.W. Wu, S.R. Roffler, *Int. J. Cancer.* **2005**, *116*, 243-52.
- [36] A.H. Reid, A. Protheroe, G. Attard, N. Hayward, L. Vidal, J. Spicer, H.M. Shaw, E.A. Bone, J. Carter, L. Hooftman, A. Harris, J.S. De Bono, *Clin. Cancer Res.* **2009**, *15*, 4978–4985.
- [37] J.E. Cortes, E.J. Feldman, K.W. Yee, M.J. Toal, A. Richardson, A. Charman, N. Flores and The OPAL Investigators, *J. Clin. Oncol.* **2010**, *28*, abstr TPS277.
- [38] J.R. Tonra, D.J. Hicklin, *Immunological Investigations.* **2007**, *36*, 3-23.
- [39] R.E. Frank, T.J. Saclarides, S. Leurgans, N.J. Speziale, E.A. Drab, D.B. Rubin, *Ann. Surg.* **1995**, *222*, 695-699.
- [40] K. Sheppard, K.M. Kinross, B. Solomon, R.B. Pearson, W.A. Phillips, *Crit. Rev. Oncog.* **2012**, *17*, 69-95.
- [41] Y. Huo, V. Iadevaia, C.G. Proud, *Biochem. Soc. Trans.* **2011**, *39*, 446–450.
- [42] D. Liao D, R.S. Johnson, *Cancer Metastasis Rev.* **2007**, *26*, 281-290.
- [43] G.L. Semenza, *Ann. N-Y. Acad. Sci.* **2009**, *1177*, 2-8.
- [44] G.L. Semenza, *Oncogene.* **2010**, *29*, 625-634.
- [45] C.C. Hudson, M. Liu, G.G. Chiang, D.M. Otterness, D.C. Loomis, F. Kaper, *Mol. Cell Biol.* **2002**, *22*, 7004-7014.
- [46] E. Guerin, W. Raffelsberger, E. Pencreach, A. Maier, A. Neuville, A. Schneider, P. Bachellier, S. Rohr, A. Petitprez, O. Poch, D. Moras, P. Oudet, A.K. Larsen, M-P. Gaub, D. Guenot, *Mol. Med.* **2012**, doi: 10.2119/molmed.2011.00120.
- [47] L.H. Stockwin, J. Blonder, M.A. Bumke, D.A. Lucas, K.C. Chan, T.P. Conrads, H.J. Issaq, T.D. Veenstra, D.L. Newton, S.M. Rybak, *J. Proteome Res.* **2006**, *5*, 2996-3007.
- [48] E. Pencreach, E. Guérin, C. Nicolet, I. Lelong-Rebel, A.C. Voegeli, P. Oudet, A.K. Larsen, M.P. Gaub, D. Guenot, *Clin. Cancer Res.* **2009**, *15*, 1297-1307.
- [49] I.H. Segel, *Enzyme Kinetics*, John Wiley & Son, New-York, **1993**.
- [50] N. Weidner, *Methods Enzymol.* **2008**, *444*, 305-323.
- [51] G. Guenot, E. Guérin, S. Aguilon-Romain, E. Pencreach, A. Schneider, A. Neuville, M.P. Chenard, I. Duluc, S. Du Manoir, C. Brigand, P. Oudet, M. Kedingler, M-P. Gaub, *J. Pathol.* **2006**, *208*, 643-652.

## CHAPITRE 2 :

Etudes structurales de l'Aminopeptidase-N  
d'*Escherichia coli* (pepN) en complexe avec les dérivés  
de la 7-amino-6-benzosubérone

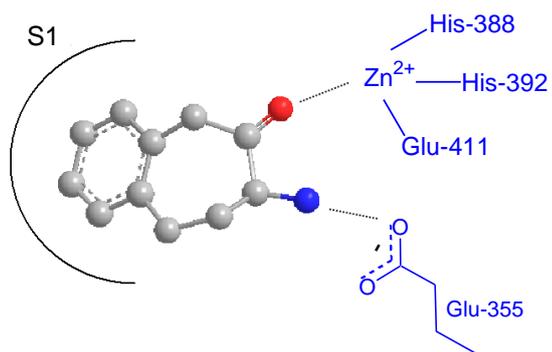


Figure 27 : Mode de liaison proposé des inhibiteurs de type 7-amino-6-benzosubérone au site actif de l'APN humaine.

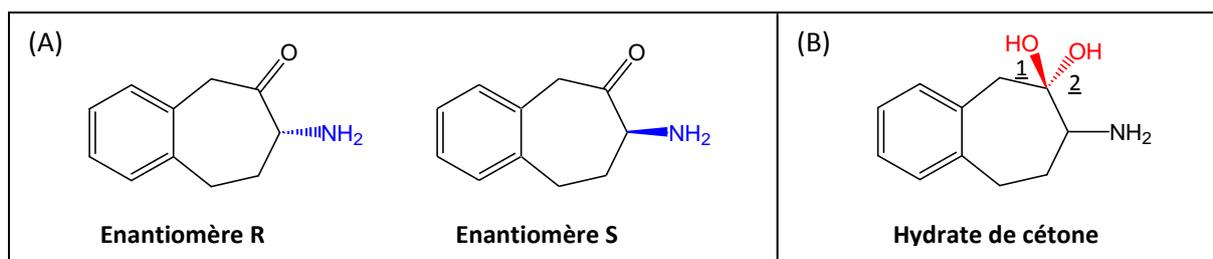


Figure 28 : (A). Enantiomères R et S de la 7-amino-benzosubérone. (B). Composé sous la forme d'un hydrate de cétone.

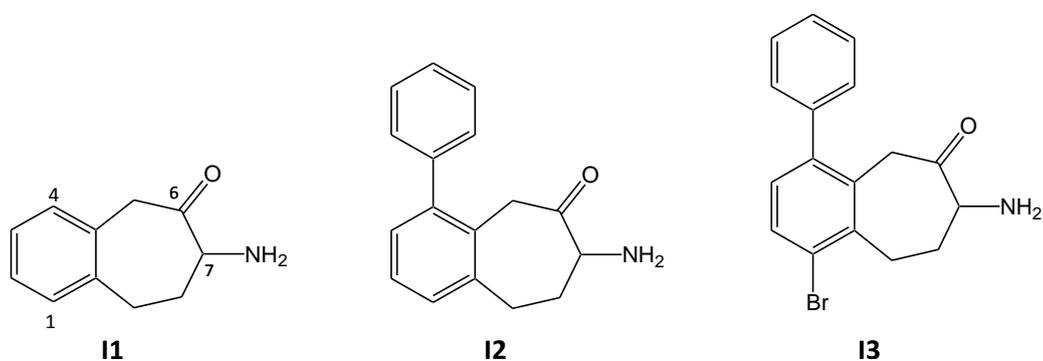


Figure 29 : Dérivés de la 7-amino-6-benzosubérone co-cristallisés avec l'APN d'E. coli.

## 1. Introduction

La structure tridimensionnelle de l'enzyme de mammifère n'étant pas résolue, les inhibiteurs de l'APN ont été conçus d'après l'étude des relations structures-activité, une approche de chimie médicinale classique. Un mode de liaison au site actif de l'enzyme humaine a été proposé<sup>99</sup> ([Figure 27](#)) à partir des structures connues d'autres métallo-aminopeptidases de la famille M1, qui présentent un domaine catalytique très conservé. Dans ce modèle, l'amine primaire interagit avec les résidus du motif GAMEN (indispensables à la reconnaissance de l'amine libre en position N-terminale d'un substrat peptidique) et la fonction carbonyle se lie au zinc.

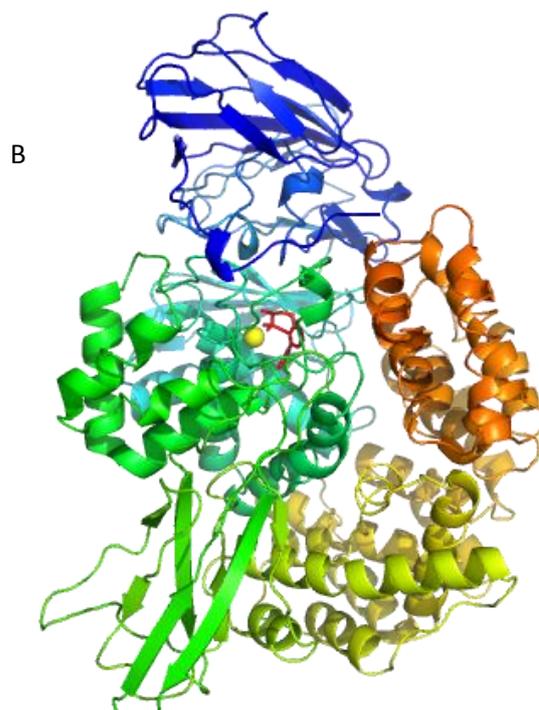
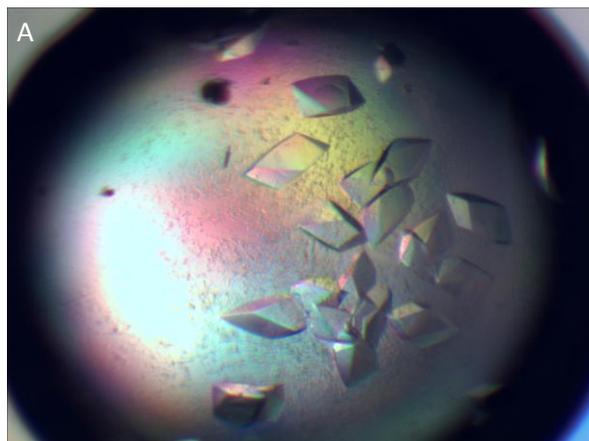
Les inhibiteurs étant synthétisés sous une forme racémique (mélange équimolaire des énantiomères R et S), nous ne savons pas quel est l'énantiomère actif ([Figure 28-A](#)). Par ailleurs, la probabilité que la molécule soit sous la forme d'un hydrate de cétone est élevée en milieu aqueux ([Figure 28-B](#))<sup>99</sup>.

Afin d'obtenir rapidement des informations complémentaires sur le mode de liaison, et confirmer ou infirmer les hypothèses énoncés ci-dessus, les dérivés de la 7-amino-6-benzosubérone (I1, I2 et I3) ([Figure 29](#)) ont été co-cristallisés avec l'Aminopeptidase-N d'*Escherichia coli* (pepN), dont l'expression, la purification et la cristallisation sont décrites dans la littérature<sup>101,102</sup>. L'affinité des composés (I1, I2 et I3) pour l'enzyme recombinante bactérienne a été évaluée avant de commencer les essais de cristallisation.

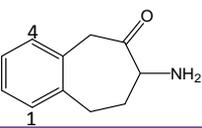
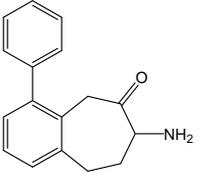
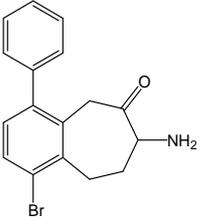
---

<sup>101</sup> Golich, F. C., Han, M., and Crowder, M. W. (2006). *Prot. Exp. Purif.* **47**, 634-639.

<sup>102</sup> Addlagatta, A., Gay, L., and Matthews, B. W. (2006). *PNAS.* **103** (36), 13339-13344.



**Figure 30 :** (A). Cristaux obtenus par la méthode de diffusion en phase vapeur, technique de la goutte pendante. (B). Structure tridimensionnelle de l'Amino-peptidase-N d'Escherichia coli en complexe avec les dérivés de la 7-amino-6-benzosubérone (résolution 1,5 Å). Le logiciel PyMOL est utilisé pour visualiser la structure tridimensionnelle de la protéine.

Inhibiteurs	Constante d'inhibition (Ki) et Efficacité du ligand (LE)		
	APN recombinante d'Escherichia coli	APN de rein de porc (Sigma-Aldrich)	APN recombinante humaine (R&D Systems)
 <b>I1</b>	Ki = 50 μM* LE = 0,46**	Ki = 1 μM LE = 0,63	Ki = 2 μM LE = 0,61
 <b>I2</b>	Ki = 300 nM LE = 0,48	Ki = 7 nM LE = 0,60	Ki = 11 nM LE = 0,58
 <b>I3</b>	Ki = 150 nM LE = 0,48	Ki = 0,060 nM LE = 0,72	Ki = 0,35 nM LE = 0,66

\*Constante d'inhibition  $K_i$  (μM)

\*\*Efficacité du ligand (LE)  $LE = \Delta G/N$  or  $\Delta G = -RT \ln(K_i)$  (kcal/mol/nombre d'atomes non hydrogénés)

**Tableau 6 :** Constantes d'inhibition  $K_i$  et efficacité du ligand de I1, I2 et I3 sur l'APN recombinante d'E.coli et sur l'enzyme de mammifères (APN porcine et humaine).

## 2. Résultats et discussion

### 2.1. Expression, purification et cristallisation de l'APN d'*E. coli*

En se basant sur les travaux de Golich et al. (2006), et d'Addlagatta et al (2006), la protéine a été dans un premier temps exprimée chez *E. coli* et purifiée, puis cristallisée en présence des inhibiteurs. L'acquisition des données cristallographiques, la résolution puis, l'affinement de la structure ont été réalisés en collaboration avec Vincent Olieric du Swiss-Lightsources de l'Institut Paul Scherrer (Villigen, Suisse). Les cristaux ont diffracté à une résolution de 1,5 Å (Figure 30).

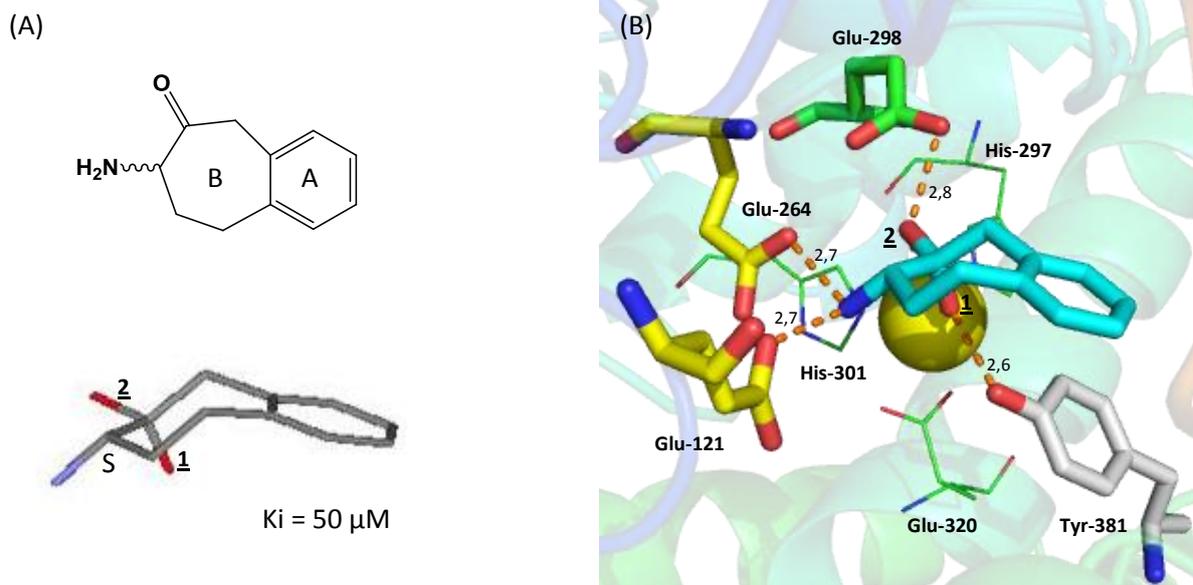
### 2.2. Evaluation *in vitro* des inhibiteurs

Les composés (I1, I2 et I3) ont été évalués sur l'APN d'*E. coli*. Leurs constantes d'inhibition  $K_i$ , déterminées à l'aide de la représentation de Dixon<sup>96</sup>, ont été comparées à celles obtenues sur l'enzyme de mammifères (APN porcine et humaine)<sup>103,104</sup> (Tableau 6). Il s'avère que ces molécules sont moins actives sur l'enzyme de bactérie, et qu'I3 s'est révélé beaucoup plus puissant sur l'APN de mammifères. En effet, l'affinité de l'inhibiteur pour l'enzyme bactérienne n'est pas améliorée suite à la substitution de la position 1 par un groupement brome.

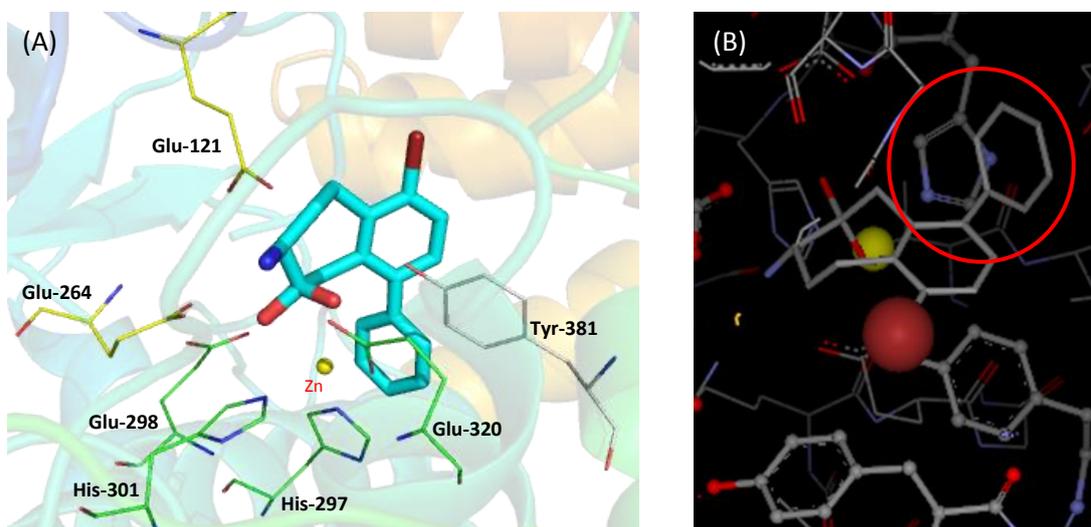
---

<sup>103</sup> Maierneau, C., Schmitt, C., Schifano-Faux, N., Le Nouën, D., Defoin, A., Tarnus, C. (2011). *Bioorg. Med. Chem.* **19**, 5716-5733.

<sup>104</sup> Schmitt, C., Voegelin, M., Marin, A., Schmitt, M., Schegg, F., Hénon, P., Guenot, D., and Tarnus, C. Soumis à *Eur. J. Med. Chem.*



*Figure 31 : (A). Conformation de la 7-amino-6-benzosubérone (B). Structure tridimensionnelle de l'APN d'E. coli avec la 7-amino-6-benzosubérone. Mode de liaison au site actif de l'enzyme. Les distances entre les hétéroatomes sont données en Angström (Å). Le logiciel PyMOL est utilisé pour visualiser la structure tridimensionnelle de la protéine.*



*Figure 32 : (A). Structure tridimensionnelle de l'APN d'E. coli avec un dérivé de la 7-amino-6-benzosubérone (13). Mode de liaison au site actif de l'enzyme. (B). Mise en évidence de l'interaction  $\pi$  (cercle rouge) entre le cycle aromatique et le résidu His-297 (ligand du zinc). Le logiciel PyMOL est utilisé pour visualiser la structure tridimensionnelle de la protéine.*

## 2.3. Analyse structurale de l'APN d'*E. coli* avec les dérivés de la 7-amino-6-benzosubérone

### *i. Conformation de la 7-amino-6-benzosubérone et stéréochimie*

Les structures de l'APN d'*E. coli* en complexe avec les dérivés de la 7-amino-6-benzosubérone apportent les informations suivantes sur le mode de liaison des inhibiteurs au site actif de l'enzyme (Figure 31) :

- le cycle à 7 chaînons (cycle B) adopte une conformation chaise
- l'énantiomère S est le seul composé lié à l'enzyme
- la fonction carbonyle est sous la forme d'un hydrate de cétone, mimant l'intermédiaire tétraédrique de l'état de transition.

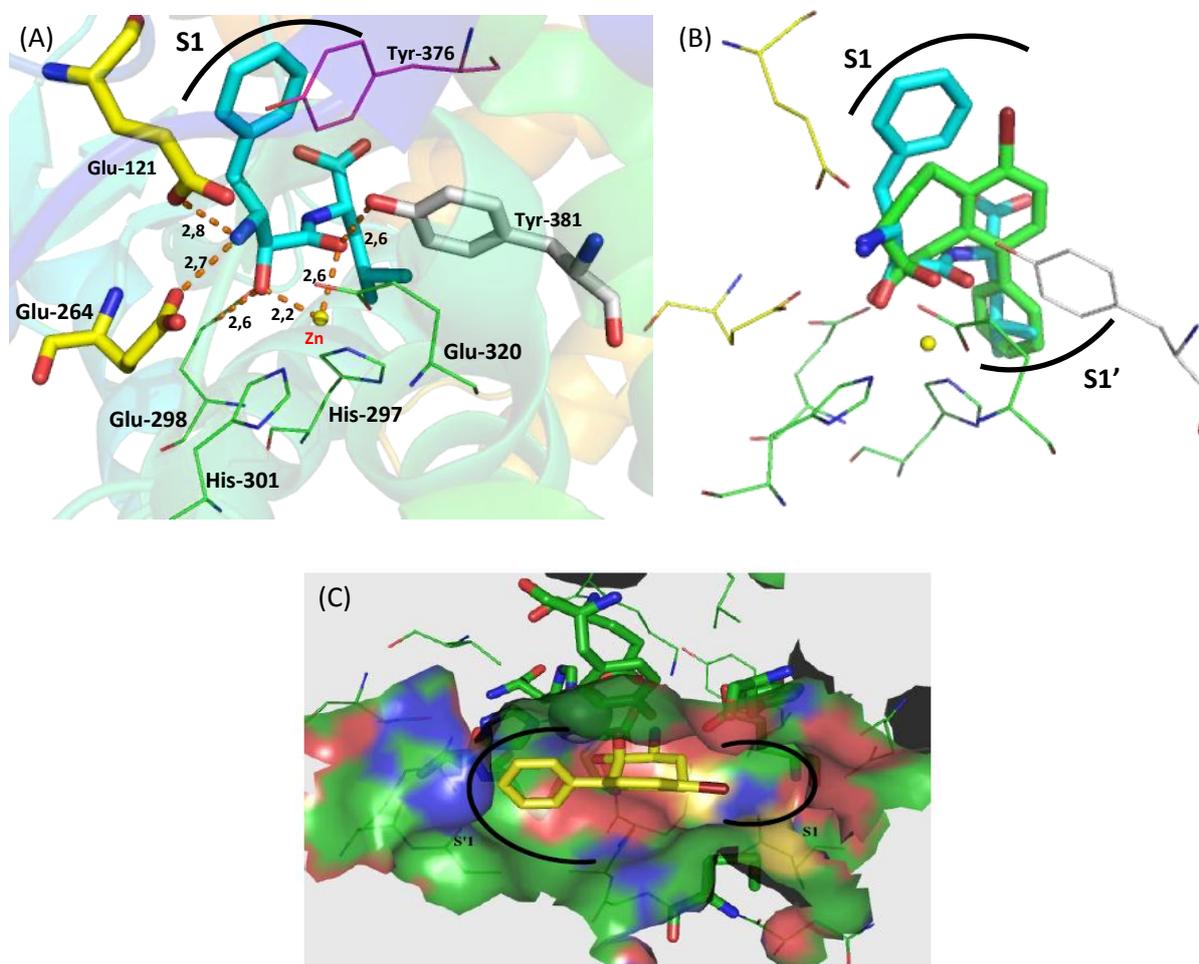
### *ii. Etude des relations structure-activité observées*

L'amine primaire de la 7-amino-6-benzosubérone interagit avec les résidus Glu-264 (motif GAMEN) et Glu-121 (Figure 31). Le groupement hydroxyle **1** se lie à l'ion zinc et forme une liaison hydrogène avec l'hydroxyle du résidu Tyr-381 (acide aminé très conservé dans la famille M1). Le groupement hydroxyle **2** occupe, quant-à-lui, la place de la molécule d'eau (agent nucléophile lors de l'hydrolyse d'un substrat peptidique) ; il se lie au zinc et forme une liaison hydrogène avec le résidu Glu-298 impliqué dans la catalyse (« navette à proton »).

Par ailleurs, l'interaction  $\pi$  entre le résidu Tyr-381 et le cycle à 6 chaînons (cycle A), ainsi que la rigidité de la molécule (gain entropique), confèrent une activité submicromolaire à ce composé de faible poids moléculaire (211,7 g/mol), qui possède une excellente efficacité de ligand (Tableau 6).

L'affinité de l'inhibiteur pour l'enzyme est améliorée lorsqu'un cycle aromatique (groupement phényle) est ajouté en position 4 de la 7-amino-6-benzosubérone (I2). Une interaction  $\pi$  supplémentaire entre le cycle aromatique et le résidu His-297 (ligand du zinc) permet d'expliquer ce gain d'affinité (activité subnanomolaire,  $K_i = 300$  nM) (Figure 32).

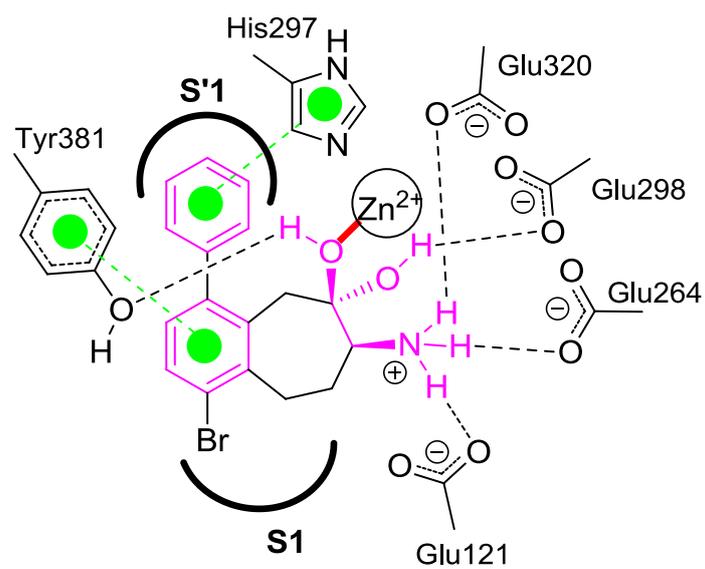
L'ajout d'un groupement brome en position 1 (I3) a peu d'effet sur l'enzyme de bactérie ; aucune interaction n'est observée (Figure 32), ce qui est en accord avec les résultats d'inhibition ( $K_i = 150$  nM) (Tableau 6).



**Figure 33 :** (A). Structure tridimensionnelle de l'Aminopeptidase-N d'*Escherichia coli* code PDB : 2dqm. Mode de liaison de la bestatine (en cyan) au site actif de l'enzyme. (B). Superposition du dérivé de la 7-amino-6-benzosubérone (13) (en vert) et de la bestatine (en cyan) au site actif de l'enzyme de bactérie. (C). Structure tridimensionnelle de l'APN d'*E. coli* avec un dérivé de la 7-amino-6-benzosubérone. Représentation de la surface de la protéine et positionnement des pharmacophores dans les sous-sites du site actif. Les distances entre les hétéroatomes et l'ion métallique sont données en Angström (Å). Le logiciel PyMOL est utilisé pour visualiser la structure tridimensionnelle de la protéine.

*iii. Superposition avec la bestatine (analogue de peptide)*

Le groupe brome en position 1 de la 7-amino-6-benzosubérone se positionne, comme le groupement phényle de la bestatine, dans le sous-site S1 du site actif (Figure 33). Le noyau aromatique en position 4 de la 7-amino-6-benzosubérone se dirige, quant-à-lui, dans le sous-site S1' (superposition avec la leucine de la bestatine).



*Figure 34* : Schéma récapitulatif du mode de liaison d'13 au site actif de l'APN d'*E. coli*.

### 3. Conclusion

La co-cristallisation de l'APN d'*E. coli* avec les composés I1, I2 et I3 a permis d'élucider le mode de liaison des inhibiteurs au site actif des métallo-aminopeptidases de la famille M1, et d'identifier les pharmacophores. Les différentes interactions sont illustrées dans la [Figure 34](#). L'ensemble de ces résultats va donner lieu à une publication.

Les relations structure-activité de ces composés sur l'APN bactérienne et humaine sont similaires. Ces inhibiteurs sont, cependant, bien plus puissants sur les enzymes de mammifères et c'est, en effet, pour ces dernières que l'effet du brome en position 1 est réellement significatif.

La résolution de la structure de l'APN humaine reste donc un enjeu majeur pour déterminer les différences discrètes que comporte son domaine catalytique par rapport à celui de l'enzyme de bactérie, et ainsi expliquer les relations structure-activité observées.



## 4. Matériels et Méthodes

Les protocoles décrits ci-dessous sont basés sur les travaux de Golich et al. (2006), et d'Addlagatta et al (2006)<sup>101,102</sup>.

### 4.1. Expression et purification de l'APN d'*E. coli* (pepN)

Le gène, codant pour la protéine pepN (NCBI, NC\_007779.1)<sup>105</sup> (Figure 35) est amplifié par PCR (réaction de polymérisation en chaîne) à partir de l'ADN génomique d'*E. coli* (souche W3110). Les séquences des amorces utilisées pour l'amplification du gène (2,61 kb) sont les suivantes : amorce sens (5'...TAT CCT CAT ATG CAT CAT CAT CAT CAT GAC GAT GAC GAT AAG ATG ACT CAA CAG CCA CAA GC...3') et antisens (5'...TA GCA TGC GGC CGC GGA TCA AGC CAG TGC TTT AGT TAT C...3'). Le fragment amplifié (~2,7 kb) est cloné dans le vecteur d'expression pET-22b(+) (Novagen®) (Figure 36)<sup>106</sup> entre les sites de restriction uniques NdeI (5'... CATATG...3') et NotI (5'... GCGGCCGC...3'). Afin de faciliter la purification ultérieure de la protéine par chromatographie d'affinité sur résine de nickel, une queue polyhistidine (5'...CATCATCATCATCAT...3') est insérée à l'extrémité N-terminale. Elle pourra ensuite être éliminée par digestion avec l'entérokinase (site de clivage spécifique 5'...GACGATGACGATAAG...3').

Le plasmide recombinant pET-22b(+)-pepN est ensuite introduit, par choc thermique, dans la souche d'expression *E. coli* BL21 (DE3).

Les bactéries ayant intégré le plasmide recombinant sont inoculées dans 10 ml de milieu LB (Luria Broth, 1% de tryptone, 0,5% d'extrait de levure et 0,5% de chlorure de sodium) contenant de l'ampicilline à une concentration de 100 µg/ml. Les cultures sont incubées sous agitation (200 rpm) toute la nuit à 37°C. Ces précultures servent à ensemercer 1 l de milieu LB supplémenté de l'antibiotique. Les cultures sont incubées à 37°C sous agitation. Lorsque la densité optique à 600 nm atteint 0,8-1, la température est abaissée à 30°C et l'expression de la protéine est induite par l'ajout d'IPTG à une concentration de 1 mM. Après 3 h d'incubation à 30°C, les bactéries sont sédimentées par centrifugation (8200 x g pendant 15 min à 4°C), resuspendues dans du tampon 50 mM HEPES, 0,5 M NaCl, pH 7,4, et lysées par sonication.

La fraction soluble après la lyse cellulaire (centrifugation à 23400 x g pendant 15 min à 4°C), est déposée sur une colonne d'affinité (résine de nickel, HisTrap FF 1 ml, GE

<sup>105</sup> Geer, L. Y., Marchler-Bauer, A., Geer, R. C., Han, L., He, J., He, S., Liu, C., Shi, W., and Bryant, S. H. (2010). *Nucleic Acids Res.* 38, D492-D496.

<sup>106</sup> Carte du vecteur pET-22b(+). <http://wolfson.huji.ac.il/expression/commercial-vectors/pet-22b-map.pdf>.

Healthcare), préalablement équilibrée avec du tampon 50 mM HEPES, 0,5 M NaCl, pH 7,4. Les protéines fixées sont ensuite éluées sur 30 fois le volume de la colonne grâce au passage d'un gradient linéaire d'Imidazole (de 0 à 110 mM) à travers la résine. Les fractions contenant la protéine d'intérêt sont réunies et dialysées dans du tampon 20 mM Tris-HCl pH 7,5.

Dans les travaux d'Addlagatta et al la protéine a été cristallisée en présence de l'étiquette polyhistidine<sup>102</sup>. Il a également été décidé d'effectuer les essais de cristallisation en la conservant.

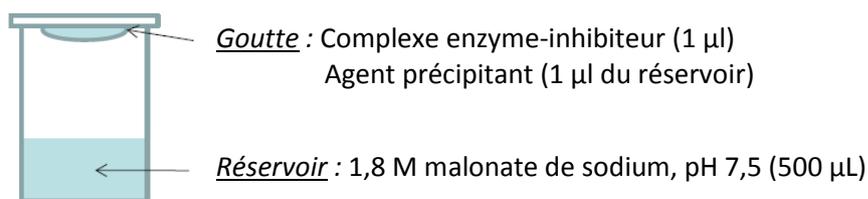
## 4.2. Cristallisation

Les essais de cristallisation sont effectués au laboratoire en utilisant la méthode de diffusion en phase vapeur par la technique de la goutte suspendue. Le malonate de sodium, pH 7,5 est employé comme agent précipitant selon les données de la littérature.

L'échantillon protéique est, au préalable, concentré ( $C_f = 29,4$  mg/ml) par ultrafiltration sur membrane (cellule d'ultrafiltration Amicon Ultra-4 ml, seuil de coupure de 30 kDa). La protéine est ensuite incubée pendant 30 minutes à température ambiante en présence des différents inhibiteurs I1, I2 et I3 (Tableau 7). Puis, 1  $\mu$ l de réservoir (500  $\mu$ l de 1,8 M malonate de sodium, pH 7,5) est ajouté à la goutte contenant le complexe enzyme-inhibiteur (Figure 37). L'enceinte est hermétiquement scellée et les boîtes sont incubées à 20°C durant plusieurs jours.

	Goutte n°1	Goutte n°2	Goutte n°3
<b>Enzyme</b>	0,5 $\mu$ l de pepN à 29,4 mg/ml ( $C_{\text{finale dans la goutte}} = 7,4$ mg/ml)		
<b>Inhibiteur</b>	0,5 $\mu$ l d'I1 à 1 mM, 10 % DMSO ( $C_{\text{finale}} = 250$ $\mu$ M, 2,5 % DMSO)	0,5 $\mu$ l d'I2 à 0,62 mM, 10 % DMSO ( $C_{\text{finale}} = 155$ $\mu$ M, 2,5 % DMSO)	0,5 $\mu$ l d'I3 à 0,5 $\mu$ M, 10 % DMSO ( $C_{\text{finale}} = 125$ $\mu$ M, 2,5 % DMSO)

*Tableau 7 : Complexe enzyme-inhibiteur. Concentration finale dans la goutte après l'ajout de 1  $\mu$ l d'agent précipitant.*



*Figure 37: Méthode de diffusion en phase vapeur, technique de la goutte suspendue. Conditions expérimentales de cristallisation de l'APN d'E. coli en présence des inhibiteurs I1, I2 et I3.*

Les cristaux sont placés dans une solution cryoprotectrice (22 % glycérol, 1,8 M malonate de sodium, pH 7,5, en présence de 250 µM d'I1, ou 155 µM d'I2, ou encore 125 µM d'I3), puis congelés dans de l'azote liquide.

#### 4.3. Acquisition des données, résolution et affinement de la structure

L'acquisition des données de cristallographie, la résolution par remplacement moléculaire, ainsi que l'affinement de la structure sont réalisés par le Dr. Vincent Olieric du Swiss Lightsource de l'Institut Paul Scherrer à Villigen (Suisse).

#### 4.4. Cinétique enzymatique

Les cinétiques enzymatiques sont effectuées à l'aide du spectrophotomètre UV-Visible HP 8453 (Agilent Technologies). La leucine p-nitroanilide est utilisée comme substrat de pepN ( $K_M = 0,2 \text{ mM}$ ) ; son hydrolyse conduit à la p-nitroaniline qui absorbe dans le visible à 405 nm ( $\epsilon = 10800 \text{ mol}^{-1} \cdot \text{L} \cdot \text{cm}^{-1}$  à pH 7,5). La réaction enzymatique se déroule à 30°C dans du tampon 20 mM Tris-HCl pH 7,5 (volume final de 1ml), et débute après l'ajout de 3mU d'enzyme (~30 nM). La vitesse initiale de la réaction est alors mesurée durant 30 minutes en suivant, par spectrophotométrie, la cinétique d'apparition de la p-nitroaniline à 405 nm.

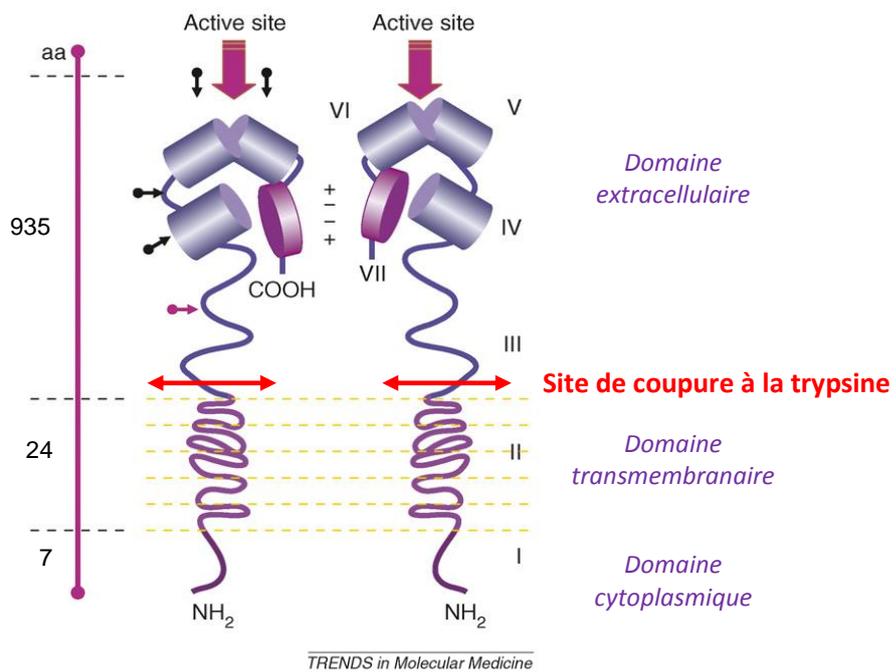
La constante d'inhibition  $K_i$  est déterminée à l'aide de la représentation de Dixon : pour une concentration en substrat donnée ( $[S] = K_M/2$ ,  $[S] = K_M$  et  $[S] = 2K_M$ ), la vitesse initiale de la réaction est mesurée pour différentes concentrations en inhibiteur. Pour chaque concentration en substrat, le droite  $1/V_i$  en fonction de  $[I]$  est tracée. L'intersection des 3 droites correspond à la valeur de  $-K_i$ .

Le composé I3 étant un inhibiteur puissant (« tight binding »), la concentration en enzyme doit être abaissée (0,05mU soit 500pM) pour pouvoir appliquer l'équation de Michaelis-Menten ( $[I] \gg [E]$ ) et déterminer le  $K_i$  via la représentation de Dixon.



## CHAPITRE 3 :

Expression, purification et cristallisation de  
l'Aminopeptidase-N humaine



*Figure 38*: Modèle schématique de la structure tridimensionnelle de l'APN de mammifère d'après l'analyse des structures secondaires<sup>48</sup>.

# INTRODUCTION

Les analyses structurales issues de la co-cristallisation de l'APN d'*Escherichia-coli* avec les dérivés de la 7-amino-6-benzosubérone (I1, I2 et I3) apportent des informations sur le mode de liaisons des inhibiteurs au site actif de l'enzyme. Ces données restent cependant insuffisantes pour expliquer l'affinité plus importante de ces composés vis-à-vis de l'APN de mammifères. La résolution de la structure tridimensionnelle de l'APN humaine constitue donc un enjeu majeur, et est indispensable pour mieux comprendre cette protéine multifonctionnelle et guider la conception d'inhibiteurs plus biodisponibles.

La purification et la cristallisation de l'APN humaine sont d'autant plus difficile que cette dernière est une protéine transmembranaire fortement glycosylée. En effet, les protéines membranaires sont délicates à manipuler car elles sont insolubles en dehors de leur environnement lipidique. L'emploi de détergents est indispensable pour les extraire des membranes et les maintenir en solution sous une forme native lors des différentes étapes de purification<sup>107,108</sup>. De plus, leur cristallisation reste un défi<sup>109</sup> car la présence de détergents empêche souvent la formation d'une structure cristalline de bonne qualité<sup>110</sup>.

Des essais d'extraction et de purification de l'APN à partir de tissus (rein de porc) ont déjà été entrepris au laboratoire. Pour faciliter les étapes de purification, le domaine extracellulaire (résidus 39 à 967) a été isolé par un traitement à la trypsine (coupure après le domaine transmembranaire) ([Figure 38](#)). A l'issue de la purification, les N-glycanes ont été éliminés par digestion enzymatique à la PNGase F (Peptide N Glycosidase F). Des études cinétiques ont permis de démontrer que le domaine soluble, qu'il soit glycosylé ou non, est stable et actif en solution.

Des essais de cristallisation ont été menés, mais en raison de la faible quantité de protéines isolées, seules quelques conditions expérimentales ont pu être testées et aucun cristal ne s'est développé.

---

<sup>107</sup> Seddon, A. M., Curnow, P., and Booth, P. J. (2004). *BBA*. **1666**, 105-117.

<sup>108</sup> Linke, D. (2009). *Methods Enzymol.* **463**, 603-617.

<sup>109</sup> Lacapère, J.-J., Pebay-Peyroula, E., Neumann, J.-M., and Etchebest, C. (2007). *Trends Biochem. Sci.* **32** (6), 259-270.

<sup>110</sup> Gabrielsen, M., Gardiner, A. T., Fromme, P., and Cogdell, R. J. (2009). *Curr. Topics in Membranes*, **63**, 127-149.



Les quantités de protéines obtenues à partir de tissus étant faibles pour pouvoir mener des essais concluants, une plateforme de biologie moléculaire a été créée au sein du laboratoire afin de produire la protéine recombinante. Pour contourner les problèmes liés à l'expression hétérologue de protéines membranaires<sup>111</sup>, et augmenter les chances de succès, l'expression, la purification et la cristallisation du domaine soluble de la protéine h-APN ont également été envisagées. Différentes stratégies de clonage et d'expression ont été développées avec l'aide du Dr. Olivier Zinc, en fonction des infrastructures et des accréditations dont dispose le laboratoire. Ces dernières ont été élaborées en parallèle car la production de protéines recombinantes reste, somme toute, relativement empirique.

*i. Expression en système procaryote*

La protéine sera exprimée chez *Escherichia coli* sous une forme non glycosylée. Toutefois, en l'absence de modifications post-traductionnelles, telles que la glycosylation ou la formation de ponts-disulfures du fait de l'environnement réducteur du cytoplasme de la bactérie, il se pourrait que la protéine s'agrège sous une forme biologiquement inactive. La formation de ces corps d'inclusion demanderait alors des étapes supplémentaires de solubilisation et de renaturation *in vitro* pour obtenir la protéine recombinante sous une forme soluble.

Ce système n'étant pas adapté à l'expression de protéines membranaires eucaryotes (le plus souvent la protéine est toxique pour la bactérie), seul le domaine soluble de h-APN sera produit.

*ii. Expression en système eucaryote*

S'il s'avère que les glycosylations sont importantes pour que la protéine adopte sa conformation active, cette dernière sera exprimée chez la levure *Kluyveromyces lactis* (*K. lactis*). Le profil de N-glycosylation sera cependant différent de celui des mammifères puisque la levure produit des formes peu matures contenant beaucoup de résidus mannoses et conduisant souvent à une hyperglycosylation. A l'issue de la purification, les N-glycanes pourront être éliminés par digestion enzymatique à la PNGase F (Peptide N Glycosidase F) sans pour autant altérer l'activité enzymatique.

---

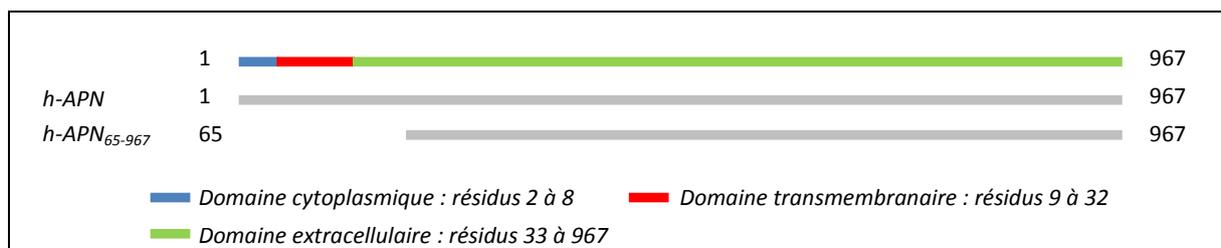
<sup>111</sup> Loll, P. J. (2003). *J. Struc. Biol.* **142**, 144-153.



En cas d'échec, d'autres systèmes d'expression sont disponibles, et la protéine pourra être produite chez les eucaryotes supérieurs comme par exemple, le système baculovirus-cellules d'insectes ou les cellules de mammifères (CHO, Chinese Hamster Ovary). Le laboratoire n'ayant pas les infrastructures et les accréditations nécessaires pour la manipulation d'eucaryotes supérieurs, ces travaux seront effectués en collaboration avec l'équipe du Prof Udo Oppermann du Centre de Recherche Botnar de l'Université d'Oxford.

Les protéines recombinantes seront, ensuite, extraites et purifiées par chromatographie d'affinité grâce à l'introduction d'une queue polyhistidine ou FLAG à l'extrémité N- ou C-terminale de la protéine. Toutefois, d'autres méthodes chromatographiques pourront être associées si le degré de pureté est insuffisant.

Pour finir, des essais de cristallisation seront effectués en partenariat avec le Consortium de Génomique Structurale (SGC) d'Oxford. Différentes conditions seront testées en parallèle afin de trouver celles conduisant à une cristallisation optimale de la protéine. L'acquisition des données cristallographiques, la résolution puis, l'affinement de la structure seront réalisées en collaboration avec des cristallographes du SGC.



*Figure 39* : Expression de la version complète de la protéine (*h-APN*) et d'une forme tronquée à l'extrémité N-terminale (*h-APN<sub>65-967</sub>*) (suppression des domaines cytoplasmique et transmembranaire et d'une partie de la région extracellulaire).

# RESULTATS ET DISCUSSION

## 1. Expression de la protéine recombinante chez la levure

Parmi les systèmes eucaryotes disponibles, la levure (*Saccharomyces cerevisiae*, *Pichia pastoris*, *Kluyveromyces lactis*,...) représente une solution intéressante car son matériel génétique est simple (organisme unicellulaire), sa croissance est rapide, et le rendement de production est élevé (plusieurs centaines de mg/l) pour un coût raisonnable<sup>112</sup>. La levure permet de produire des protéines avec des modifications post-traductionnelles (glycosylations simples, carboxylations, acylations...); cependant, les N-glycosylations ne sont pas aussi bien reproduites qu'avec des systèmes eucaryotes supérieurs (Baculovirus, cellules de mammifères).

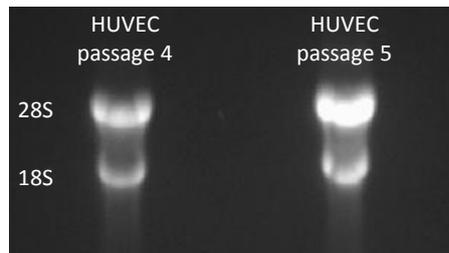
La protéine h-APN est exprimée chez la levure *Kluyveromyces lactis* (*K. lactis*) à l'aide du système « *K. lactis* Protein Expression Kit » commercialisé par New England Biolabs (NEB) (cf. Matériels et Méthodes p191). Contrairement à la levure *Saccharomyces cerevisiae*, *K. lactis* a pour avantage de posséder un système de sécrétion efficace.

Pour pallier aux difficultés rencontrées lors de l'expression hétérologue de protéines membranaires, une forme tronquée (résidus 65 à 967) à l'extrémité N-terminale de la protéine h-APN est également produite (Figure 39).

Le fragment d'ADN (voir Annexe 1) qui code pour la version complète de la protéine (h-APN) ou pour sa forme tronquée (h-APN<sub>65-967</sub>) est cloné dans le vecteur pKLAC2 entre les sites de restriction uniques XhoI et BamHI. La séquence  $\alpha$ -MF, en amont du gène cloné, code pour un peptide signal (précurseur pré-pro du facteur  $\alpha$ ) qui est responsable de l'adressage des protéines recombinantes dans la voie de sécrétion de la levure. Les protéines surexprimées sont ainsi glycosylées. Elles vont soit s'insérer dans la membrane plasmique des cellules hôtes (h-APN), soit être sécrétées dans le milieu de culture sous une forme soluble (h-APN<sub>65-967</sub>).

---

<sup>112</sup> Çelik, E., and Çalik, P. (2011). *Biotechnol. Adv.* doi: 10.1016/j.biotechadv.2011.09.011.



*Figure 40 : Electrophorèse sur gel d'agarose 1% en présence d'un agent dénaturant (formaldéhyde).*

## 1.1. Clonage

### 1.1.1. Isolement du gène d'intérêt

Dans un premier temps, les ARN totaux sont extraits de cellules en culture à l'aide du kit RNAqueous®-4PCR (Ambion). Les cellules endothéliales humaines de veine ombilicale (HUVEC) sont choisies car la protéine est surexprimée à leur surface. L'échantillon d'ARN totaux est ensuite traité à la désoxyribonucléase I (DNase I) afin de dégrader l'ADN génomique résiduel. Les résultats obtenus sont représentés dans le Tableau 8.

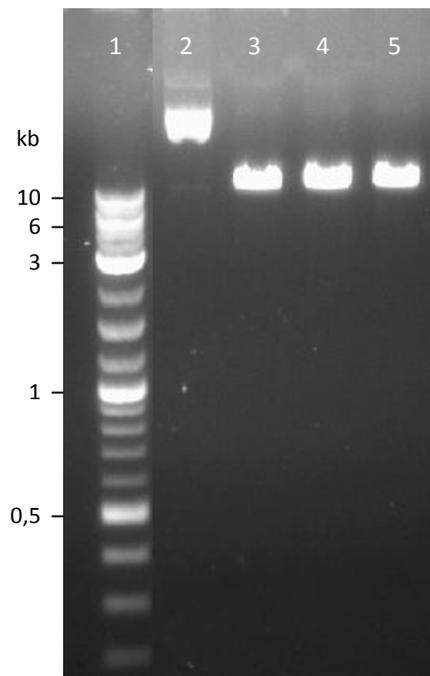
Echantillon	Nombre de cellules	Concentration ARN totaux	260/280	260/230
HUVEC passage 4	~1x 10 <sup>7</sup> cellules	950 ng/μl 724 ng/μl après traitement à la DNase I	2,02	2,13
HUVEC passage 5	~5,4 x 10 <sup>6</sup> cellules	874 ng/μl 611 ng/μl après traitement à la DNase I	2,07	2,02

Tableau 8 : Bilan des extractions d'ARN totaux à partir de cellules en culture (HUVEC).

Les rapports de densité optique 260/280 et 260/230 témoignent de la pureté des échantillons. Aucune contamination par des protéines (ratio 260/280 >1,8) ou par le thiocyanate de guanidine (ratio 260/230 > 1,8) n'est observée.

L'intégrité des ARN totaux est ensuite vérifiée par électrophorèse sur gel d'agarose 1% dénaturant (formaldéhyde) (Figure 40). La présence de deux bandes distinctes, correspondant aux ARN ribosomiaux 28S et 18S, dont le rapport des intensités 28S/18S est d'environ 2, atteste de l'intégrité des échantillons.

Le 1<sup>er</sup> brin d'ADNc est obtenu à partir de l'échantillon d'ARN totaux (HUVEC passage 4) suite à l'action de la transcriptase inverse M-MLV (H-) (Promega); la synthèse est initiée à l'aide d'une amorce oligo-dT qui s'hybride sur la queue polyadénylate des ARN messager (ARNm). L'ADNc sert ensuite de matrice pour la PCR.



*Figure 41*: Electrophorèse sur gel d'agarose 1%. Puits 1 : marqueur de taille moléculaire 0,1-10 kb. Puits 2 : contrôle négatif (vecteur pKLAC2). Puits 3 : vecteur pKLAC2 linéarisé par XhoI. Puits 4 : vecteur pKLAC2 linéarisé par BamHI. Puits 5 : vecteur pKLAC2 clivé simultanément par XhoI et BamHI. La double digestion du vecteur libère un fragment de 66 pb qui n'est pas détectable sur le gel et qui est éliminé lors de la purification sur colonne de silice (QIAquick PCR Purification kit, QIAGEN).

L'ADN polymérase haute-fidélité DyNAzyme™ EXT (Finnzymes) est utilisée pour amplifier par PCR (réaction de polymérisation en chaîne) les ADNc double brin, ANPEP et ANPEP<sub>194-2904</sub>, qui codent pour la version complète de la protéine h-APN et sa forme tronquée h-APN<sub>65-967</sub>. Les sites de restrictions XhoI et BglII sont introduits à chaque extrémité du fragment amplifié pour le clonage dans le vecteur pKLAC2. Afin de faciliter les étapes ultérieures de purification, une queue polyhistidine est introduite en position N-terminale de la protéine ; elle sera ensuite éliminée par digestion avec l'entérokinase. Des fragments de respectivement 2976 pb et 2784 pb sont amplifiés. Ils sont extraits du gel d'agarose et purifiés à l'aide du kit QIAquick Gel Extraction (QIAGEN).

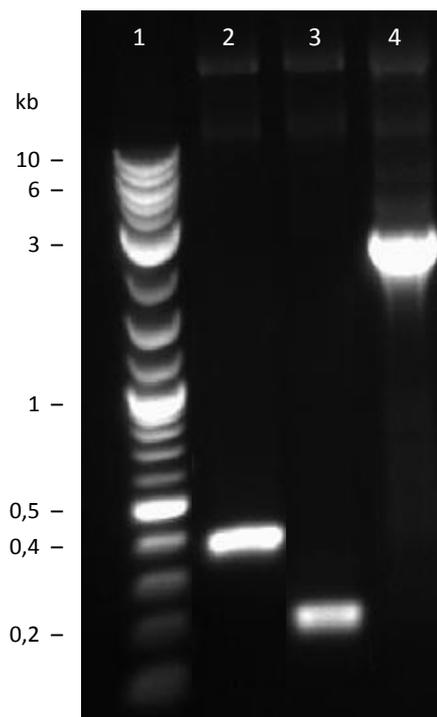
### *1.1.2. Clonage dans le vecteur d'expression pKLAC2*

Les fragments d'ADN amplifiés (ANPEP et ANPEP<sub>194-2904</sub>) sont clivés par les enzymes de restriction XhoI et BglII, qui génèrent des bouts cohésifs à l'extrémité des fragments.

Le vecteur pKLAC2 est, quant à lui, digéré par les enzymes de restriction XhoI et BamHI ; l'enzyme BamHI libère une extrémité cohésive compatible avec celle produite par BglII. Les produits de digestion sont analysés sur gel d'agarose ([Figure 41](#)). Afin d'empêcher la recircularisation du vecteur lors de la ligation par l'ADN ligase T4, le vecteur est ensuite déphosphorylé à l'aide de la phosphatase alcaline de l'intestin de veau (CIP).

Les fragments d'ADN digérés sont purifiés sur des colonnes de silice à l'aide du kit QIAquick PCR purification (QIAGEN).

Différentes proportions de vecteur et d'insert sont mélangées (ratio 1 :1 et 1 :3) en présence d'ADN ligase du phage T4 et d'ATP. Les plasmides recombinants sont ensuite introduits par choc thermique dans les bactéries XL10 Gold (Stratagene) ; les colonies bactériennes qui se sont multipliées sur milieu sélectif solide sont criblées par PCR avec des amorces spécifiques au vecteur. Des fragments de respectivement 3316 pb (pKLAC2-ANPEP) et 3124 pb (pKLAC2-ANPEP<sub>194-2904</sub>) sont amplifiés lorsque l'insert est introduit dans le vecteur.



*Figure 42* : Electrophorèse sur gel d'agarose 1%. Contrôle de l'intégrité du plasmide recombinant pKLAC2-ANPEP. Puits 1 : marqueur de taille moléculaire 0,1-10 kb. Puits 2 et 3 : l'amplification de fragments d'environ 400 et 205 pb démontre que le gène d'intérêt est inséré dans le bon sens. Puits 4 : l'amplification d'un fragment d'environ 2904 pb (utilisation d'amorces spécifiques au gène ANPEP) permet d'affirmer que le gène d'intérêt est bien inséré dans le vecteur. Des résultats similaires sont obtenus pour le plasmide pKLAC2-ANPEP<sub>194-2904</sub>.

Les clones positifs sont inoculés dans un milieu liquide sélectif et remis en culture. Les bactéries sont ensuite lysées par un traitement alcalin, et les ADN plasmidiques sont extraits et purifiés sur une résine échangeuse d'anions (midiprep, QIAGEN). Les résultats obtenus sont indiqués dans le Tableau 9.

ADN plasmidique	Concentration	Volume	Masse	260/280	260/230
pKLAC2-ANPEP	800 ng/μl	100 μl	80 μg	1,9	2,2
pKLAC2-ANPEP <sub>194-2904</sub>	650 ng/μl	100 μl	65 μg	1,9	2,1

*Tableau 9 : Bilan de l'extraction et de la purification de l'ADN plasmidique à l'aide du kit QIAGEN® Plasmid Purification Midiprep (QIAGEN).*

L'intégrité des plasmides recombinants purifiés par midiprep est contrôlée par PCR. Les produits de PCR sont analysés sur gel d'agarose (Figure 42).

HindIII

```
AAGCTTGAAAAAATGAAATTCTCTACTATATTAGCCGCATCTACTGCTTTAATTTCCGTTGTTATGGCTGCTCCAGTTTCTA
CCGAAACTGACATCGACGATCTTCCAAATATCGGTTCCAGAAGAAGCCTTGATTGGATTCATTGACTTAACCGGGGATGAAGTT
TCCTTGTGCTGTAAATAACGGAAACCCACACTGGTATTCTATTCTTAAACACCACCATCGCTGAAGCTGCTTTCCGCTGACAA
GGATGATCTCGAGAAAAGACATATGCATCATCATCATCATGACGATGTCGATAAGATGGCCAAGGGCTTCTATATTTCCA
AGTCCCTGGGCATCTTGGGATCCCTGGGCGTGGCAGCCGTGTGCACAATCATCGCACTGTGAGTGGTGTACTCCAGGAG
AAGAACAAGAACGCCAACAGCTCCCCCGTGGCCTCCACTACCCCGTCCGCCTCAGCCACCACCAACCCCGCTCGGCCACCAC
CTTGGACCAAAGTAAAGCGTGAATCGTTACCGCTCCCAACACGCTGAAACCCGATTCCTACCGGTGACGCTGAGACCGT
ACCTACCCCAATGACAGGGGCTGTACGTTTTTAAGGGCTCCAGCACCGTCCGTTTACCTGCAAGGAGGCCACTGAGCTC
ATCATCATCCACAGCAAGAAGCTCACTACACCCTCAGCCAGGGGCACAGGGTGGTCCCTGCGTGGTGTGGGAGGCTCCCAGCC
CCCCGACATTGACAAGACTGAGCTGGTGGAGCCACCAGGTACCTGGTGGTGCACCTCAAGGGCTCCCTGGTGAAGGACAGCC
AGTATGAGATGGACAGCGAGTTTCGAGGGGAGTTGGCAGATGACCTGGCGGGCTTACCAGCAGGATACATGGAGGGCAAT
GTCAGAAAGGTGGTGGCCACTACACAGATGCAGGCTGCAGATGCCCGGAAGTCCTTCCCATGCTTCGATGAGCCGGCCATGAA
GGCCGAGTTCAACATCAGCTTATCCACCCCAAGGACCTGACAGCCCTGTCCAACATGCTTCCCAAAGTCCCAGCACCCAC
TTCCAGAAGACCCCAACTGGAATGTACTGAGTTCCACACCACGCCAAGATGTCCACGTAATGCTGGCCTTCATTGTGAGT
GAGTTCGACTACGTGGAGAAGCAGGCATCCAATGGTGTCTTGATCCGGATCTGGGCCCGGCCAGTGCCATTGCGGGGGCCA
CGGCATATGCCCAGAAGTACGGGCCCCATCCTTAACTTCTTGTGGTCAATATGACACACCCTACCCACTCCCAAAAT
CAGACCAGATTGGCCTGCCAGACTTCAACGCCGGCCCATGGAGAAGTGGGGACTGGTACCTACCGGGAGAAGTCCCTGCTG
TTCGACCCCTGTCTCCTCCAGCAGCAACAAGGAGCGGGTGGTCACTGTGATTGCTCATGAGCTGGCCACCAGTGGTTCGG
GAACCTGGTGAACATAGAGTGGTGAATGACCTGTGGCTGAACGAGGGCTTCGCCTCCTACGTGGAGTACCTGGGTGCTGACT
ATGCGGAGCCACCTGGAACTTGAAAGACCTCATGGTGTGAATGATGTGTACCGCTGATGGCAGTGGATGCACTGGCCTCC
TCCCACCGCTGTCCACACCCGCTCGGAGATCAACAGCCGGCCAGATCAGTGAGCTGTTTACGCCATCTCCTACAGCAA
GGGCGCTCAGTCTCAGGATGCTCCTCAGCTTCTGTCCGAGGACGATTAAGCAGGGCCTGGCGTCTACCTCCACACCT
TTGCCACAGAACACCATCTACCTGAACCTGTGGGACCCTGCAGGAGGCTGTGAACAACCGGTCCATCCAACCTCCCACC
ACCGTGGGGACATCATGAACCGCTGGACCTGCAGATGGCTTCCCGGTATCAGGTTGGGATACAGCAGCGGGACCCCTTC
CCAGGACACTTCTCCTTGACCCCGATTCCAATGTTACCCGCCCTCAGAATTCAACTACGTGTGGATGTGCCCATCACAT
CCATCAGAGATGGCAGACAGCAGCAGGACTACTGGCTGATGATGTAAGAGCCAGAACGATCTCTTACGACATCAGGCAAT
GAGTGGGTCTGCTGAACCTCAATGTGACGGCTATTACCGGGTGAACACGACGAAGAGAAGTGGAGGAAGATTCAGACTCA
GCTGCAGAGAGACCACTCGGCCATCCCTGTCAATCGGGCACAGATCATTAAATGACGCCTTCAACCTGGCCAGTGGCCATA
AGGTCCCTGTCACTCTGGCGCTGAACAACACCCCTCTTCTGATTGAAGAGAGACAGTACATGCCCTGGGAGGCCGCCCTGAGC
AGCTGAGCTACTTCAAGCTCATGTTTGACCGCTCCGAGGTCTATGGCCCATGAAGAAGTACCTGAAGAAGCAGGTACACC
CCTCTTCACTTCAAGAAATAATACCAACAACCTGGAGGGAGATCCAGAAAACCTGATGGACCAGTACAGCGAGGTTAATG
CCATCAGACCCGCTGCTCCAACGGAGTTCAGAGTGTGAGGAGATGGTCTCTGGCCTTTTCAAGCAGTGGATGGAGAACCC
AATAATAACCCGATCCACCCCAACCTGCGGTCCACCGTCTACTGCAACGCTATCGCCAGGGCGGGAGGAGTGGGACTT
CGCTGGGAGCAGTTCGAAATGCCACACTGGTCAATGAGGCTGACAAGCTCCGGGCAGCCCTGGCCTGCAGCAAAGAGTTGT
GGATCTGAACAGGTACCTGAGCTACACCCTGAACCCGACTTAATCCGGAAGCAGGACGCCACCTTACCATCATCAGCATT
ACCAACAACGTCATTGGGCAAGGTCTGGTCTGGGACTTTGTCCAGAGCAACTGGAAGAAGCTTTTTAACGATTATGGTGGTGG
CTCGTCTCCTTCTCAACCTCATCCAGGCAGTACACGACGATTCTCCACCGAGTATGAGCTGCAGCAGCTGGAGCAGTCA
AGAAGGACAACGAGGAAACAGGCTTCGGCTCAGGCACCCGGGCCCTGGAGCAAGCCCTGGAGAAGACGAAAGCCAACATCAAG
TGGTGAAGGAGAACAAGGAGGTGGTGTCCAGTGGTTCACAGAAAACAGCAAAATAGTCCAGATCCGAATTCCTGCAAGTAA
TTAAATAAAGGCCCT
```

StuI

**ATC...TAG** : séquence codant la protéine h-APN

**CTCGAG** : site de restriction XhoI

**AGATCC** : site BglII-BamHI

**AAAAGA** : site de clivage par la protéase Kex (clivage du domaine pro du facteur  $\alpha$  lors du passage de la protéine de fusion dans l'appareil de Golgi).

**CATCATCATCATCATCAT** : queue polyhistidine (purification ultérieure par chromatographie d'affinité).

**GACGATGTCGATAAG** : séquence codant pour le site de clivage de l'entérokinase (permet d'éliminer la queue polyhistidine en fin de purification).

*Figure 43 : Séquence du plasmide recombinant pKLAC2-ANPEP au niveau du site de clonage multiple. Les mutations ponctuelles sont représentées en rouge.*

### 1.1.3. Séquençage de l'ADN

Le séquençage de l'ADN est réalisé par la société Microsynth AG (Suisse) en utilisant la méthode de Sanger. Les résultats de séquençage sont analysés à l'aide du logiciel Chromaspro<sup>113</sup> et comparés à des bases de données (NCBI Blast)<sup>114</sup>.

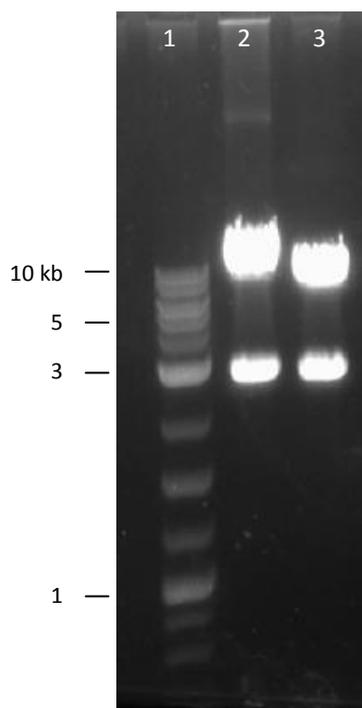
Plusieurs mutations ponctuelles sont détectées dans le plasmide recombinant pKLAC2-ANPEP (Figure 43, ci-contre). Une première mutation est observée au niveau de la séquence codant le site de clivage par l'entérokinase : GAG GAT GTC GAT AAG. La substitution de l'adénine par une thymine dans le codon 3 entraîne une modification de l'acide aminé codé : Asp-Asp-Val-Asp-Lys. Or, la présence de 4 résidus acide aspartique (Asp) est indispensable pour la reconnaissance par la protéase. Par conséquent, la queue polyhistidine en position N-terminale de la protéine ne pourra pas être éliminée en fin de purification si cette dernière est produite. Deux autres mutations sont détectées au niveau du fragment d'ADN qui code pour la protéine h-APN. Le remplacement d'une cytosine (ACC) par une thymine (ACT) au niveau du codon 49 n'engendre aucune modification dans la séquence de la protéine. Cependant, la substitution de l'adénine (ATA) par une guanine (ATG) au niveau du codon 603 provoque un changement dans la séquence protéique : remplacement de l'isoleucine par une méthionine. Ces acides aminés étant tous les deux hydrophobes, cette mutation ne devrait pas être délétère.

Aucune mutation n'est observée dans la séquence d'ADN codant la forme soluble de la protéine h-APN<sub>65-967</sub> (plasmide recombinant pKLAC2-ANPEP<sub>194-2904</sub>).

---

<sup>113</sup> <http://www.technelysium.com.au/ChromasPro.html>.

<sup>114</sup> Geer, L. Y., Marchler-Bauer, A., Geer, R. C., Han, L., He, J., He, S., Liu, C., Shi, W., and Bryant, S. H. (2010). *Nucleic Acids Res.* **38**, D492-496.



*Figure 44* : Electrophorèse sur gel d'agarose 1%. Puits 1 : marqueur de taille moléculaire 0,1-10 kb. Puits 2 : produits de digestion du plasmide recombinant pKLAC2-ANPEP. La digestion par *SacII* libère deux fragments ; la cassette d'expression (9,2 kb) et un fragment d'ADN linéaire (2,8 kb). Il n'est pas nécessaire de séparer ces deux fragments car seule la cassette d'expression peut s'intégrer dans le génome de *K. lactis*. Puits 3 : produits de digestion du plasmide contrôle pKLAC1-malE. La digestion par *SacII* génère deux fragments ; la cassette d'expression (7,2 kb) et un fragment d'ADN linéaire (3,3 kb). Des résultats identiques sont obtenus pour la digestion du plasmide recombinant pKLAC2-ANPEP<sub>194-2904</sub>.

## 1.2. Intégration de la cassette d'expression dans le génome de *K. lactis*

La digestion des plasmides recombinants par l'enzyme de restriction SacII génère un fragment d'ADN linéaire qui est capable de s'intégrer de façon stable dans le génome de *K. lactis*. Les produits de digestion sont analysés par électrophorèse sur gel d'agarose (Figure 44) et purifiés sur colonne de silice.

La cassette d'expression est intégrée par recombinaison homologue dans le génome de *K. lactis* au niveau du locus LAC4. La transposition s'effectue après transformation par choc thermique dans les cellules compétentes GG799 de *K. lactis*. Lors de la transformation, jusqu'à dix copies de la cassette d'expression peuvent se transposer dans le génome de la levure. Les cellules transformées qui se sont multipliées sur une gélose pauvre en azote contenant de l'acétamide sont criblées par PCR à l'aide d'amorces spécifiques au locus LAC4. Un fragment de 2400 pb est amplifié si la cassette d'expression est correctement intégrée dans le génome de la levure. Une seconde PCR permet ensuite d'identifier les clones qui ont incorporés au minimum deux copies de la cassette d'expression (amplification d'un fragment de 2300 pb). Ces derniers ont l'avantage de produire plus de protéines.

## 1.3. Tests d'expression de la protéine recombinante

Des clones positifs, ayant intégré une ou plusieurs copies de la cassette d'expression, sont mis en culture dans un milieu riche de base. L'expression du gène d'intérêt est induite par l'ajout de galactose dans le milieu de culture. La croissance des levures transformées est suivie et comparée à celle de la souche sauvage GG799 ; la surexpression des protéines ne semble pas affecter la croissance des levures (les courbes de croissance sont similaires).

Les culots cellulaires et les surnageants de culture sont analysés pour localiser la protéine d'intérêt et évaluer le niveau d'expression : analyse sur gel SDS-PAGE et par Western Blot.



Les différents essais amènent aux observations suivantes :

- Culture des levures GG799 qui ont intégré la cassette d'expression portant le gène malE (contrôle positif) : l'analyse sur gel SDS-PAGE a permis de démontrer que la protéine de fusion MBP (« maltose binding protein ») est sécrétée dans le milieu de culture (observation d'une bande intense vers 40 kDa).
- Culture des levures GG799 qui ont intégré une ou plusieurs copies de la cassette d'expression contenant le gène ANPEP (codant pour la forme membranaire) : la protéine h-APN devrait être insérée dans la membrane plasmique des levures ; or les analyses par SDS-PAGE et Western-Blot ne permettent pas de déceler la protéine.
- Culture des levures GG799 qui ont intégré une ou plusieurs copies de la cassette d'expression contenant le gène ANPEP tronqué (codant pour la forme soluble de la protéine) : la protéine APN<sub>65-967</sub> devrait être sécrétée sous forme soluble dans le milieu de culture ; cependant aucune expression n'est détectée que ce soit dans le surnageant de culture ou dans la fraction soluble après lyse des levures.

## 1.4. Conclusion et discussion

Dans les conditions d'expression testées, ni la protéine membranaire h-APN ni sa forme soluble (résidus 65 à 967) ne sont produites chez la levure *Kluyveromyces lactis*. Plusieurs hypothèses peuvent être émises pour expliquer ces résultats, la seconde, semblant être la plus plausibles.

### 1.4.1. Dégradation par des enzymes protéolytiques

La dégradation de la protéine hétérologue par des enzymes endogènes de la voie de sécrétion, appartenant à la famille des protéases à acide aspartique, est fréquente chez la levure. NEB commercialise depuis peu des souches mutantes de GG799 déficientes pour une des protéases à acide aspartique (e.g. Yps1, Yps7, Bar1)<sup>115</sup>. Dans certain cas, le rendement de production est amélioré. Toutefois, n'ayant observé aucune expression, il a été décidé de ne pas poursuivre dans cette voie.

---

<sup>115</sup> Ganatra, M.B., Vainauskas, S., Hong, J. M., Taylor, T. E., Denson, JP. M., Esposito, D., Read, J. D., Schmeisser, H., Zoon, K. C., Hartley, J. L., and Taron, C. H. (2011). *FEMS Yeast Res.* **11**(2), 168–178.

**ATG** GCC AAG **GGC** TTC TAT ATT TCC AAG TCC **CTG** **GGC** ATC **CTG** GGG ATC **CTC** **CTG** **GGC** GTG GCA  
 GCC GTG **TGC** ACA ATC ATC GCA **CTG** TCA GTG GTG TAC TCC CAG GAG AAG AAC AAG AAC GCC AAC  
**AGC** TCC **CCC** GTG GCC TCC ACC ACC CCG TCC GCC TCA GCC ACC ACC AAC **CCC** GCC TCG GCC ACC  
 ACC TTG GAC CAA AGT AAA GCG TGG AAT CGT TAC **CGC** **CTC** **CCC** AAC ACG **CTG** AAA **CCC** GAT TCC  
 TAC **CGG** GTG ACG **CTG** AGA CCG TAC **CTC** ACC **CCC** AAT GAC AGG **GGC** **CTG** TAC GTT TTT AAG **GGC**  
 TCC **AGC** ACC GTC CGT TTC ACC **TGC** AAG GAG GCC ACT GAC GTC ATC ATC ATC CAC **AGC** AAG AAG  
**CTC** AAC TAC ACC **CTC** **AGC** CAG GGG CAC AGG GTG GTC **CTG** CGT GGT GTG GGA **GGC** TCC CAG **CCC**  
**CCC** GAC ATT GAC AAG ACT GAG **CTG** GTG GAG **CCC** ACC GAG TAC **CTG** GTG GTG CAC **CTC** AAG **GGC**  
 TCC **CTG** GTG AAG GAC **AGC** CAG TAT GAG ATG GAC **AGC** GAG TTC GAG GGG GAG TTG GCA GAT GAC  
**CTG** GCG **GGC** TTC TAC **CGC** **AGC** GAG TAC ATG GAG **GGC** AAT GTC AGA AAG GTG GTG GCC ACT ACA  
 CAG ATG CAG GCT GCA GAT GCC **CGG** AAG TCC TTC CCA **TGC** TTC GAT GAG CCG GCC ATG AAG GCC  
 GAG TTC AAC ATC CCA CTT ATC CAC **CCC** AAG GAC **CTG** ACA GCC **CTG** TCC AAC ATG CTT **CCC** AAA  
 GGT **CCC** **AGC** ACC CCA CTT CCA GAA GAC **CCC** AAC TGG AAT GTC ACT GAG TTC CAC ACC ACG **CCC**  
 AAG ATG TCC ACG TAC TTG **CTG** GCC TTC ATT GTC AGT GAG TTC GAC TAC GTG GAG AAG CAG GCA  
 TCC AAT GGT GTC TTG ATC **CGG** ATC TGG GCC **CGG** **CCC** AGT GCC ATT GCG GCG **GGC** CAC **GGC** GAT  
 TAT GCC **CTG** AAC GTG ACG **GGC** **CCC** ATC CTT AAC TTC TTT GCT GGT CAT TAT GAC ACA **CCC** TAC  
 CCA **CTC** CCA AAA TCA GAC CAG ATT **GGC** **CTG** CCA GAC TTC AAC GCC **GGC** GCC ATG GAG AAC TGG  
 GGA **CTG** GTG ACC TAC **CGG** GAG AAC TCC **CTG** **CTG** TTC GAC **CCC** **CTG** TCC TCC TCC **AGC** **AGC** AAC  
 AAG GAG **CGG** GTG GTC ACT GTG ATT GCT CAT GAG **CTG** GCC CAC CAG TGG TTC GGG AAC **CTG** GTG  
 ACC ATA GAG TGG TGG AAT GAC **CTG** TGG **CTG** AAC GAG **GGC** TTC GCC TCC TAC GTG GAG TAC **CTG**  
 GGT GCT GAC TAT GCG GAG **CCC** ACC TGG AAC TTG AAA GAC **CTC** ATG GTG **CTG** AAT GAT GTG TAC  
**CGC** GTG ATG GCA GTG GAT GCA **CTG** GCC TCC TCC CAC CCG **CTG** TCC ACA **CCC** GCC TCG GAG ATC  
 AAC ACG CCG GCC CAG ATC AGT GAG **CTG** TTT GAC GCC ATC TCC TAC **AGC** AAG **GGC** GCC TCA GTC  
**CTC** AGG ATG **CTC** TCC **AGC** TTC **CTG** TCC GAG GAC GTA TTC AAG CAG **GGC** **CTG** GCG TCC TAC **CTC**  
 CAC ACC TTT GCC TAC CAG AAC ACC ATC TAC **CTG** AAC **CTG** TGG GAC CAC **CTG** CAG GAG GCT GTG  
 AAC AAC **CGG** TCC ATC CAA **CTC** **CCC** ACC ACC GTG **CGG** GAC ATC ATG AAC **CGC** TGG ACC **CTG** CAG  
 ATG **GGC** TTC CCG GTC ATC ACG GTG GAT ACC **AGC** ACG GGG ACC CTT TCC CAG GAG CAC TTC **CTC**  
 CTT GAC **CCC** GAT TCC AAT GTT ACC **CGC** **CCC** TCA GAA TTC AAC TAC GTG TGG ATT GTG **CCC** ATC  
 ACA TCC ATC AGA GAT **GGC** AGA CAG CAG CAG GAC TAC TGG **CTG** ATA GAT GTA AGA GCC CAG AAC  
 GAT **CTC** TTC **AGC** ACA TCA **GGC** AAT GAG TGG GTC **CTG** **CTG** AAC **CTC** AAT GTG ACG **GGC** TAT TAC  
**CGG** GTG AAC TAC GAC GAA GAG AAC TGG AGG AAG ATT CAG ACT CAG **CTG** CAG AGA GAC CAC TCG  
 GCC ATC CCT GTC ATC AAT **CGG** GCA CAG ATC ATT AAT GAC GCC TTC AAC **CTG** GCC ACC GTG CAC  
 AAG GTC CCT GTC ACT **CTG** GCG **CTG** AAC AAC ACC **CTC** TTC ATT GAA GAG AGA CAG TAC ATG  
**CCC** TGG GAG GCC GCC **CTG** **AGC** **AGC** **CTG** **AGC** TAC TTC AAG **CTC** ATG TTT GAC **CGC** TCC GAG GTC  
 TAT **GGC** **CCC** ATG AAG AAC TAC **CTG** AAG AAG CAG GTC ACA **CCC** **CTC** TTC ATT CAC TTC AGA AAT  
 AAT ACC AAC AAC TGG AGG GAG ATC CCA GAA AAC **CTG** ATG GAC CAG TAC **AGC** GAG GTT AAT GCC  
 ATC **AGC** ACC GCC **TGC** TCC AAC GGA GTT CCA GAG TGT GAG GAG ATG GTC TCT **GGC** CTT TTC AAG  
 CAG TGG ATG GAG AAC **CCC** AAT AAT AAC CCG ATC CAC **CCC** AAC **CTG** **CGG** TCC ACC GTC TAC **TGC**  
 AAC GCT ATC GCC CAG **GGC** GGG GAG GAG TGG GAC TTC GCC TGG GAG CAG TTC CCA AAT GCC  
 ACA **CTG** GTC AAT GAG GCT GAC AAG **CTC** **CGG** GCA GCC **CTG** GCC **TGC** **AGC** AAA GAG TTG TGG ATC  
**CTG** AAC AGG TAC **CTG** **AGC** TAC ACC **CTG** AAC CCG GAC TTA ATC **CGG** AAG CAG GAC GCC ACC TCT  
 ACC ATC ATC **AGC** ATT ACC AAC AAC GTC ATT GGG CAA GGT **CTG** GTC TGG GAC TTT GTC CAG **AGC**  
 AAC TGG AAG AAG CTT TTT AAC GAT TAT GGT GGT **GGC** TCG TTC TCC TTC TCC AAC **CTC** ATC CAG  
 GCA GTG ACA CGA CGA TTC TCC ACC GAG TAT GAG **CTG** CAG CAG **CTG** GAG CAG TTC AAG AAG GAC  
 AAC GAG GAA ACA **GGC** TTC **GGC** TCA **GGC** ACC **CGG** GCC **CTG** GAG CAA GCC **CTG** GAG AAG ACG AAA  
 GCC AAC ATC AAG TGG GTG AAG GAG AAC AAG GAG GTG GTG **CTC** CAG TGG TTC ACA GAA AAC **AGC**  
 AAA **TAG**

Pourcentage de GC : 57,3%

Codons rares Arg : **CGC** (0,6%) et **CGG** (1,45%)

Codon rare Cys : **TGC** (0,6%)

Codon rare Pro : **CCC** (2,9%)

Codons rares Leu : **CTC** (2,2%) et **CTG** (6,1%)

Codon rare Ser : **AGC** (2,5%)

Codon rare Gly : **GGC** (2,9%)

Figure 45 : Séquence codante du gène ANPEP; les codons rarement utilisés chez *K. lactis* sont mis en évidence ; leur fréquence d'apparition dans le gène ANPEP est indiquée entre parenthèse.

#### 1.4.2. Biais d'usage du code

La production de protéines dans les systèmes d'expression hétérologue est souvent limitée par la présence dans la séquence nucléotidique de codons rarement utilisés par la cellule hôte.

Un acide aminé peut être codé par plusieurs codons, cependant l'utilisation préférentielle d'un codon par rapport à leurs synonymes est propre à chaque espèce et dépend de la disponibilité des ARN de transfert (ARNt) dans la cellule (biais de codons)<sup>116</sup>. La présence de codons rares peut devenir un facteur limitant : lorsque le ribosome rencontre un codon rare, la vitesse de traduction est ralentie car l'ARNt associé est peu abondant dans la cellule. Une interruption de la traduction sur plusieurs codons successifs peut conduire à une terminaison prématurée, à un décalage du cadre de lecture et/ou à l'incorporation d'un mauvais acide aminé.

L'étude de l'usage des codons chez *K. lactis* (<http://www.kazusa.or.jp/codon/>) (voir Annexe 2) révèle l'existence dans le gène ANPEP de 8 codons rarement utilisés par la levure : CTC (Leu), CTG (Leu), CCC (Pro), TGC (Cys), CGC (Arg), CGG (Arg), AGC (Ser) et GGC (Gly) (Figure 45, ci-contre).

La proportion (19%) et la succession de plusieurs codons rares sont probablement à l'origine de l'échec de production de la protéine recombinante. L'optimisation des codons via la synthèse du gène semble être la stratégie à développer pour produire la protéine h-APN dans ce système d'expression.

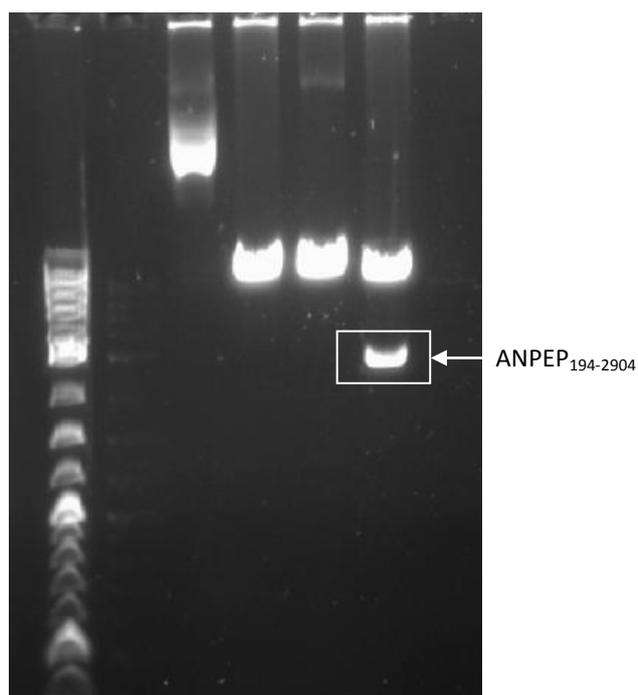
Néanmoins, il serait plus judicieux d'exprimer la protéine chez un eucaryote supérieur comme, par exemple, le système baculovirus-cellules d'insectes. Le laboratoire n'ayant pas les infrastructures et les accréditations nécessaires, une collaboration a été établie avec l'équipe du Prof. Udo Oppermann du Centre de Recherche Botnar de l'Université d'Oxford pour mener à bien ces travaux. Dans un premier temps, les tentatives d'expression de la protéine h-APN chez *Escherichia coli*, réalisées au laboratoire en parallèle à l'expression chez la levure, seront présentées.

---

<sup>116</sup> Gustafsson, C., Govindarajan, S., and Minshull, J. (2004). *TRENDS in Biotechnology*. **22** (7), 346-353.

	<u>HindIII</u>		pKLAC2
1	AAGCTT	AAAAAATGAAATTCTCTACTATATTAGCCGCATCTACTGCTTTAATTTCCGTTGTTATGGCTG	
237		CTTTCGCTGACAAGGATGAT	<span style="color: blue;">CTCGAG</span> <span style="color: green;">AAAAGACATATG</span> <span style="color: purple;">CATCATCATCATCATCAT</span> <span style="color: red;">GACGATGACGATAAG</span>
			<span style="color: blue;">XhoI</span> <span style="color: green;">NdeI</span> <span style="color: purple;">6x HIS</span> <span style="color: red;">entérokinases</span>
308		ATG	GACCAAAGTAAAGCGTGGAATCGTTACCGCCTCCCCAACACGCTGAAACCCGATTCTACCGGGTGAC
			<b>ANPEP tronqué</b>
2990		CTCCAGTGGTTCACAGAAAACAGCAAA	<span style="color: red;">TAC</span> <span style="color: orange;">TCCAGATCC</span> <span style="color: green;">GAATTC</span> <span style="color: green;">CCTGCAGGTAATTAAATAA</span> <span style="color: green;">AGGCCTT</span>
			<span style="color: red;">STOP</span> <span style="color: orange;">BglIII/BamHI</span> <span style="color: green;">StuI</span>

*Figure 46: Séquence simplifiée du plasmide recombinant pKLAC2-ANPEP<sub>194-2904</sub> au niveau du site de clonage multiple.*



*Figure 47: Electrophorèse sur gel d'agarose 1%. Puits 1 : marqueur de taille moléculaire 0,1-10 kb. Puits 2 : contrôle négatif (plasmide pKLAC2-ANPEP<sub>194-2904</sub>). Puits 3 : plasmide pKLAC2-ANPEP<sub>194-2904</sub> linéarisé par NdeI. Puits 4 : plasmide pKLAC2-ANPEP<sub>194-2904</sub> linéarisé par StuI. Puits 5 : plasmide pKLAC2-ANPEP<sub>194-2904</sub> clivé simultanément par NdeI et StuI.*

## 2. Expression de la protéine recombinante chez *Escherichia coli*

La bactérie *Escherichia coli* est la plus connue et la plus utilisée pour la production à haut rendement (plusieurs dizaines de g/L) de protéines recombinantes. Cependant, les protéines produites ne possèdent pas de modifications post-traductionnelles, indispensables à leur maturation, et s'accumulent donc souvent dans le cytoplasme sous forme d'agrégats insolubles (corps d'inclusion). Néanmoins, de nombreuses solutions, que nous discuterons ultérieurement, ont été développées pour favoriser l'expression hétérologue de protéines eucaryotes sous une forme native et soluble.

Le système pET commercialisé par Novagen® est choisi pour exprimer la forme soluble de la protéine h-APN (résidus 65 à 967) chez la bactérie *E. coli* (cf. Matériels et Méthodes p210). Le gène ANPEP<sub>194-2904</sub> qui code pour la forme tronquée de la protéine est extrait du plasmide recombinant pKLAC2-ANPEP<sub>194-2904</sub>, utilisé précédemment chez la levure, et cloné dans le vecteur d'expression pET-22b(+) entre les sites de restriction uniques NdeI et EcoRI. La protéine est exprimée dans le cytoplasme de la bactérie sans aucune modification post-traductionnelle. En raison du faible taux de réussite (protéine toxique pour la bactérie ou agrégation fréquente sous forme de corps d'inclusion), la forme membranaire de h-APN n'est pas produite dans ce système d'expression.

### 2.1. Clonage

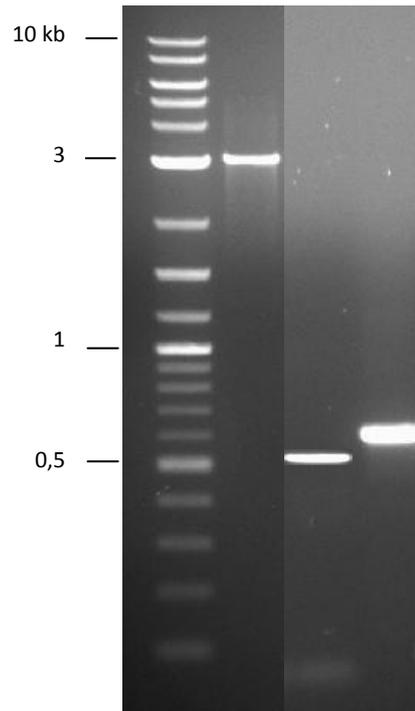
#### 2.1.1. Isolement du gène d'intérêt

Le gène ANPEP<sub>194-2904</sub> qui code pour la forme soluble de la protéine (h-APN<sub>65-967</sub>) est extrait du plasmide pKLAC2-ANPEP<sub>194-2904</sub> par digestion enzymatique avec NdeI et StuI (Figure 46).

Les produits de digestion sont analysés sur gel d'agarose (Figure 47) ; le fragment d'ADN (2786 pb) codant pour la forme tronquée de la protéine h-APN est extrait du gel et purifié sur une colonne de silice.

#### 2.1.2. Clonage dans le vecteur d'expression pET-22b(+)

Le gène dont une des extrémités est cohésive et l'autre franche est cloné dans le vecteur d'expression pET-22b(+) entre les sites de restriction NdeI et EcoRI.



*Figure 48 : Electrophorèse sur gel d'agarose 1%. Contrôle de l'intégrité du plasmide recombinant pET-22b(+)-ANPEP<sub>194-2904</sub>. Puits 1 : marqueur de taille moléculaire 0,1-10 kb. Puits 2 : l'amplification d'un fragment de 2997 pb (utilisation d'amorces spécifiques au vecteur pET-22b(+)) permet d'affirmer que le gène d'intérêt est bien inséré dans le vecteur. Puits 3 et 4 : l'amplification de fragments d'environ 500 et 632 pb démontre que le gène d'intérêt est inséré dans le bon sens.*

Dans un premier temps, le vecteur est digéré par EcoRI. Cette enzyme génère des extrémités cohésives qui sont rendues franches par traitement à la nucléase Mung Bean. Le vecteur est ensuite digéré par l'enzyme NdeI, déphosphorylé à l'aide de la phosphatase alcaline de l'intestin de veau (CIP), et pour finir purifié sur une colonne de silice.

Le gène d'intérêt est mélangé au vecteur linéarisé (ratio molaire 2 :1 et 3 :1) et traité par l'ADN ligase du phage T4 en présence d'ATP. Les plasmides recombinants sont ensuite introduits par choc thermique dans les bactéries XL10 Gold. Les colonies bactériennes qui se sont multipliées sur milieu sélectif solide sont criblées par PCR avec des amorces spécifiques au vecteur. Un fragment de 2997 pb est amplifié lorsque le gène d'intérêt est introduit dans le plasmide pET-22b(+).

Les clones positifs sont mis en culture dans un milieu liquide sélectif ; les ADN plasmidiques produits par les bactéries sont extraits et purifiés sur une résine échangeuse d'anions (midiprep, QIAGEN).

L'intégrité du plasmide recombinant pET-22b(+)-ANPEP<sub>194-2904</sub> (obtenu à une concentration de 1 µg/µl) est vérifiée par PCR ; les produits de PCR sont analysés sur gel d'agarose ([Figure 48](#)).

## 2.2. Culture des bactéries et expression de la protéine recombinante

### 2.2.1. Choix des souches bactériennes

Les souches d'expression décrites ci-dessous sont sélectionnées pour surexprimer la protéine h-APN<sub>65-967</sub> chez la bactérie :

- **BL21 (DE3)** : Les protéines eucaryotes surexprimées chez la bactérie sont rapidement dégradées par des protéases. La souche d'expression BL21 (DE3), qui est déficiente en protéases ompT et lon, permet d'améliorer la stabilité des protéines hétérologues.
- **Rosetta 2 (DE3)** : Dans certains cas, le rendement de production (expression d'une protéine sous une forme native et soluble) peut être amélioré en exprimant la protéine dans une souche bactérienne qui coexprime les ARNt rares<sup>117</sup>.

---

<sup>117</sup> Tegel, H., Tourle, S., Ottosson, J., and Persson, A., 2010. *Protein Expression and Purification*. **69**, 159-167

**ATC** GCC AAG GGC TTC TAT ATT TCC AAG TCC CTG GGC ATC CTG GGG ATC CTC CTG GGC GTG GCA GCC GTG  
 TGC ACA ATC ATC GCA CTG TCA GTG GTG TAC TCC CAG GAG AAG AAC AAG AAC GCC AAC AGC TCC **CCC** GTG  
 GCC TCC ACC ACC CCG TCC GCC TCA GCC ACC AAC **CCC** GCC TGG GCC ACC ACC TTG GAC CAA AGT AAA  
 GCG TGG AAT CGT TAC CGC CTC **CCC** AAC ACG CTG AAA **CCC** GAT TCC TAC **CGG** GTG ACG CTG **AGA** CCG TAC  
 CTC ACC **CCC** AAT GAC **AGG** GGC CTG TAC GTT TTT AAG GGC TCC AGC ACC GTC CGT TTC ACC TGC AAG GAG  
 GCC ACT GAC GTC ATC ATC CAC AGC AAG AAG CTC AAC TAC ACC CTC AGC CAG GGG CAC **AGG** GTG GTC  
 CTG CGT GGT GTG **GGA** GGC TCC CAG **CCC** GAC ATT GAC AAG ACT GAG CTG GTG GAG **CCC** ACC GAG TAC  
 CTG GTG GTG CAC CTC AAG GGC TCC CTG GTG AAG GAC AGC CAG TAT GAG ATG GAC AGC GAG TTC GAG GGG  
 GAG TTG GCA GAT GAC CTG GCG GGC TTC TAC CGC AGC GAG TAC ATG GAG GGC AAT GTC **AGA** AAG GTG GTG  
 GCC ACT ACA CAG ATG CAG GCT GCA GAT **CCC** **CGG** AAG TCC TTC CCA TGC TTC GAT GAG CCG GCC ATG AAG  
 GCC GAG TTC AAC ATC ACG CTT ATC CAC **CCC** AAG GAC CTG ACA GCC CTG TCC AAC ATT GTC **CCC** AAA GGT  
**CCC** AGC ACC CCA CTT CCA GAA GAC **CCC** AAC TGG AAT GTC ACT GAG TTC CAC ACC ACG **CCC** AAG ATG TCC  
 ACG TAC TTG CTG GCC TTC ATT GTC AGT GAG TTC GAC TAC GTG GAG AAG CAG GCA TCC AAT GGT GTC TTG  
 ATC **CGG** ATC TGG GCC **CGG** **CCC** AGT GCC ATT GCG GCG GGC CAC GCC GAT TAT GCC CTG AAC GTG ACG GGC  
**CCC** ATC CTT AAC TTC TTT GCT GGT CAT TAT GAC ACA **CCC** TAC CCA CTC CCA AAA TCA GAC CAG ATT GGC  
 CTG CCA GAC TTC AAC GCC GGC CCG ATG GAG AAC TGG **GGA** CTG ACC TAC **CGG** GAG AAC TCC CTG CTG  
 TTC GAC **CCC** CTG TCC TCC TCC AGC AGC AAC AAG GAG **CGG** GTG GTC ACT GTG ATT GCT CAT GAG CTG GCC  
 CAC CAG TGG TTC GGG AAC CTG GTG ACC **ATA** GAG TGG TGG AAT GAC CTG TGG CTG AAC GAG GGC TTC GCC  
 TCC TAC GTG GAG TAC CTG GGT GCT GAC TAT GCG GAG **CCC** ACC TGG AAC TTG AAA GAC CTC ATG GTG CTG  
 AAT GAT GTG TAC CGC GTG ATG GCA GTG GAT GCA CTG GCC TCC TCC CAC CCG CTG TCC ACA **CCC** GCC TCG  
 GAG ATC AAC ACG CCG GCC CAG ATC AGT GAG CTG TTT GAC GCC ATC TCC TAC AGC AAG GGC GCC TCA GTC  
 CTC **AGG** ATG CTC TCC AGC TTC CTG TCC GAG GAC GTA TTC AAG CAG GGC CTG GCG TCC TAC CTC CAC ACC  
 TTT GCC TAC CAG AAC ACC ATC TAC CTG AAC CTG TGG GAC CAC CTG CAG GAT GTG AAC AAC **CGG** TCC  
 ATC CAA CTC **CCC** ACC ACC GTG **CGG** GAC ATC ATG AAC CCG TGG ACC CTG CAG ATG GGC TTC CCG GTC ATC  
 ACG GTG GAT ACC AGC ACG GGG ACC CTT TCC CAG GAG CAC TTC CTC CTT GAC **CCC** GAT TCC AAT GTT ACC  
 CGC **CCC** TCA GAA TTC AAC TAC GTG TGG ATT GTG **CCC** ATC ACA TCC ATC **AGA** GAT GGC **AGA** CAG CAG CAG  
 GAC TAC TGG CTG **ATA** GAT GTA **AGA** GCC CAG AAC GAT CTC TTC AGC ACA TCA GGC AAT GAG TGG GTC CTG  
 CTG AAC CTC AAT GTG ACG GTG **CGG** GTG AAC TAC GAC GAA GAG AAC TGG **AGG** AAG ATT CAG ACT  
 CAG CTG CAG **AGA** GAC CAC TCG GCC ATC **CCT** GTC ATC AAT **CGG** GCA CAG ATC ATT **AAT** GAC GCC TTC AAC  
 CTG GCC AGT GCC CAT AAG GTC CCT GTC ACT CTG GCG CTG **AAC** AAC ACC CTC TTC CTG ATT GAA GAG **AGA**  
 CAG TAC ATG **CCC** TGG GAG GCC GCC CTG AGC AGC CTG AGC TAC TTC AAG CTC ATG TTT GAC CCG TCC GAG  
 GTC TAT GGC **CCC** ATG AAG AAC TAC CTG AAG AAG CAG GTC ACA **CCC** CTC TTC ATT CAC TTC **AGA** AAT AAT  
 ACC AAC AAC TGG **AGG** GAG ATC CCA GAA AAC CTG ATG GAC CAG TAC AGC GAG GTT AAT GCC **ATC** AGC ACC  
 GCC TGC TCC AAC **GGA** GTT CCA GAG TGT GAG GAG ATG GTC TCT GGC CTT TTC AAG CAG TGG ATG GAG AAC  
**CCC** AAT AAT AAC **CCC** ATC CAC **CCC** AAC CTG **CGG** TCC ACC GTC TAC TGC AAC GCT ATC GCC CAG GGC GGG  
 GAG GAG GAG TGG GAC TTC GCG TGG GAG CAG TTC **CGA** AAT GCC ACA CTG GTC AAT GAG GCT GAC AAC CTC  
**CGG** GCA GCC CTG GCC TGC AGC AAA GAG TTG TGG ATC CTG AAC **AGG** TAC CTG AGC TAC ACC CTG AAC CCG  
 GAC TTA ATC **CGG** AAG CAG GAC GCC ACC TCT ACC ATC ATC AGC ATT ACC AAC AAC GTC ATT GGG CAA GGT  
 CTG GTC TGG GAC TTT GTC CAG AGC AAC TGG AAG CTT TTT AAC GAT TAT GGT GGT GGC TCG TTC TCC  
 TTC TCC AAC CTC ATC CAG GCA GTG ACA **CGA** **CGA** TTC TCC ACC GAG TAT GAG CTG CAG CAG CTG GAG CAG  
 TTC AAG AAG GAC AAC GAG GAA ACA GGC **TTC** **GCG** TCA GGC ACC **CGG** GCC CTG GAG CAA GCC CTG GAG AAG  
 ACG AAA GCC AAC ATC AAG TGG GTG AAG GAG AAC AAG GAG GTG GTG CTC CAG TGG TTC ACA GAA AAC AGC  
 AAA **TAG**

Pourcentage de GC : 57,3%

Codons rares Arg : **AGG** (0,6%), **AGA** (0,8%), **CGG** (1,4%) et **CGA** (0,3%)

Codon rare Gly : **GGA** (0,3%) Codon rare Ile : **ATA** (0,2%) Codon rare Pro : **CCC** (2,9%)

**Figure 49** : séquence nucléotidique du gène ANPEP ; les codons rarement utilisés chez E. coli sont mis en évidence ; leur fréquence d'apparition dans le gène ANPEP est indiqué entre parenthèse.

MAKGFYISKSLGILGILLGVAAV**C**TIIIALSVVYSQEKKNANSSPVASTTPSASATTNPASATTLTLDQSKAWNRYRLPNTLKP  
 SYRVTLRPYLTPNDRGLYVFKGSSTVRFT**C**KEATDVIIISHKKNLNYTSLQGHRVVLRGVGGSQPPDIDKTELVEPT EYLVVHL  
 KGSLVKDSQYEMDSEFEGELADDLAGFYRSEYMEGNVRKVVATTQMQAADARKSFP**C**FDEPAMKAEFNITLIHPKDLTALS  
 LPKGPSTPLPEDPNWNVTEFHTTPKMSTYLLAFIVSEFDYVEKQASNGVLIRIWARPSAIAAGHGDIYALNVTGPILNFFAGHY  
 DTPYPLPKSDQIGLPDFNAGAMENWGLVTYRENSLLFDPLSSSSSNKERVVTVIAHELAHQWFGNLVTIEWWDLWLNEGFAS  
 YVEYLGADYAEPTWNLKDLMLVNDVYRMAVDALASSHPLSTPASEINTPAQISELFDALSYSKGASVLRMLSSFLSEDFVKQ  
 GLASYLHTFAYQNTIYLNLDHDLQEAVNRSIQLPPTVRDIMNRWTLQMGFPVITVDTSTGTLSQEHFLDPDSNVTRPSEFN  
 YVWIVPITSIRDGRQQDYWLIDVRAQNDFSTSGNEWVLLNLNVTGYRVNYDEENWRKIQTQLQRDHSAPVINRAQIIND  
 AFNLASAHKVPVTLALNNTLFLIEERQYMPWEAALSSLSYFKLMFDRSEVYGPKNYLKQVTPFLFIHFRNNTNNWREIPENL  
 MDQYSEVNAISTA**C**SNGVPE**C**EEMVSGLFKQWMENPNNNPIHPNLRSTVY**C**NAIAQGGEEDWFAWEQFRNATLVNEADKLRA  
 ALA**C**SKELWILNRYLSYTLNPD LIRKQDATSTIIISITNNVIGQLVWDFVQSNWKKLFNDYGGGSFSSFNLIQAVTRRFST EY  
 ELQQLEQFKKDNEETGFGSGTRALEQALEKTKANIKWKENKEVVLQWFTENSK

Résidus Cystéine (**C**) en position 24 (domaine transmembranaire), 113, 223, 761, 768, 798 et 834

**Figure 50** : Séquence protéique de la protéine h-APN ; les résidus cystéines sont mis en évidence.

En effet, l'usage du code génétique diffère d'un organisme à l'autre et peut influencer le niveau d'expression de la protéine recombinante. Le gène ANPEP qui code pour la protéine h-APN contient un certain nombre de codons rarement utilisés chez *E. coli* ([Figure 49](#)). Pour limiter le ralentissement ou la terminaison prématurée de la traduction, la protéine est exprimée dans la souche Rosetta 2 (DE3). Cette bactérie, qui dérive des BL21 (DE3), porte un plasmide pRARE2 qui permet la synthèse des ARNt rares : AGG, AGA, AUA, CUA, CCC, GGA et CGG.

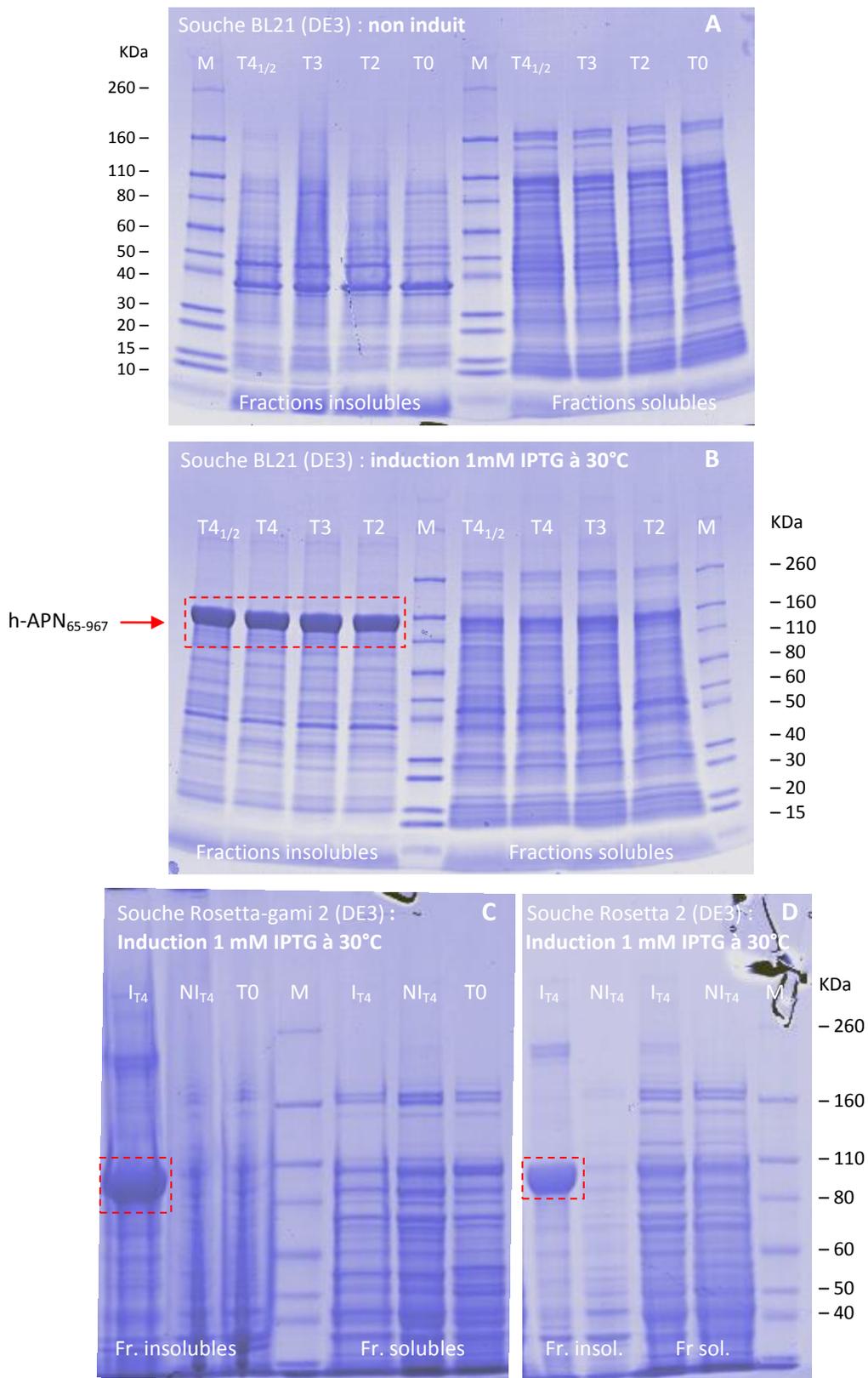
- Rosetta-gami 2 (DE3) : L'environnement réducteur du cytoplasme ne favorise pas la formation de ponts disulfures dont peut dépendre en partie la structure tridimensionnelle d'une protéine. Par conséquent, la protéine ne parvient pas à atteindre sa conformation native et s'accumule sous forme d'agrégats insolubles (biologiquement inactifs). La protéine h-APN contient 7 résidus cystéines ([Figure 50](#)) dont 6 qui peuvent être impliqués dans la formation de ponts disulfures intramoléculaires. La construction par homologie de séquence d'un modèle 3D (serveur Swiss-Model, expasy)<sup>87</sup> permet de prédire l'établissement d'une liaison covalente entre les résidus cystéines 761-768 et 798-834. Pour favoriser la formation de ces 2 ponts disulfures, la protéine est exprimée dans la bactérie *E. coli* Rosetta-gami 2 (DE3)<sup>118</sup> car cette souche possède les mêmes caractéristiques que les Rosetta 2 (DE3) ainsi que des mutations au niveau des gènes *trxB* et *gor*, codant respectivement pour des thiorédoxines réductases et des glutathionnes réductases. L'inhibition de la synthèse de ces enzymes empêche la réduction des thiorédoxines et des glutarédoxines présentes dans le cytoplasme de la bactérie ; sous une forme oxydée, ces dernières sont alors capables de catalyser la formation de ponts disulfures dans les protéines.

### 2.2.2. Expression de la protéine recombinante

Le plasmide recombinant pET-22b(+)-ANPEP<sub>194-2904</sub> est introduit par choc thermique dans les cellules compétentes BL21 (DE3), Rosetta 2 (DE3) et Rosetta-gami 2 (DE3). Les bactéries ayant intégré le plasmide recombinant sont identifiées par PCR. Les clones positifs sont mis en culture dans un milieu nutritif riche en présence des antibiotiques appropriés.

---

<sup>118</sup> Bessette, P. H., Aslund, F., Beckwith, J., and Georgiou, G., 1999. *Biochemistry*. **96**(24), 13703-13708



**Figure 51 :** Analyse de l'expression de la protéine hAPN<sub>65-967</sub> sur gel SDS-PAGE (NuPAGE® 3-8% Tris Acétate, Invitrogen). 30 µg de protéines totales sont déposés dans chaque puits. **A & B :** Expression de la protéine dans la souche BL21 (DE3). **A :** Culture de bactérie non induite (contrôle négatif). **B :** Induction IPTG (1 mM) à 30°C pendant 4h30. **C :** Expression de la protéine dans la souche Rosetta-gami 2 (DE3) ; **D :** Expression de la protéine dans la souche Rosetta 2 (DE3). **C & D :** culture de bactérie non induite (NI) et induction IPTG (1 mM) à 30°C pendant 4h (I). M : marqueur de poids moléculaire.

Différents essais sont réalisés pour identifier les conditions optimales d'expression de la protéine recombinante sous forme native et soluble : variation de la température, de la concentration en IPTG, de la densité optique à 600 nm au moment de l'induction et du temps d'induction (Tableau 10).

Souche	DO <sub>600nm</sub> début induction	Concentration en IPTG (mM)	Température	Temps
BL21 (DE3)	0,6-0,8	1	30°C	4 heures
BL21 (DE3)	0,6-0,8	1	25°C	8 heures
BL21 (DE3)	0,2-0,4	1	25°C	8 heures
BL21 (DE3)	0,6-0,8	0,4	25°C	8 heures
BL21 (DE3)	0,6-0,8	0,4	20°C	1 nuit
Rosetta 2 (DE3)	0,6-0,8	1	30°C	4 heures
Rosetta 2 (DE3)	0,6-0,8	0,4	20°C	1 nuit
Rosetta-gami 2 (DE3)	0,6-0,8	1	30°C	4 heures
Rosetta-gami 2 (DE3)	0,6-0,8	0,4	20°C	1 nuit

Tableau 10 : Tests d'expression de la protéine hAPN<sub>65-967</sub> chez *E. coli* : variation des conditions de culture

Les bactéries sont lysées par sonication, puis les protéines solubles sont séparées des protéines insolubles par centrifugation. Le niveau d'expression et la proportion de protéine soluble sont analysés sur gel SDS-PAGE ; une bande intense est attendue vers 104 kDa (Figure 51).

Les résultats montrent que quelle que soit la souche bactérienne utilisée, la protéine s'accumule dans le cytoplasme sous forme d'agrégats insolubles (corps d'inclusion) : présence d'une bande intense entre 100 et 105 kDa. La proportion de protéine exprimée sous une forme soluble et native est minime voire nulle (absence d'une bande de 104 kDa). Par ailleurs, la variation des conditions de culture, telles que la concentration en IPTG, la densité optique (DO<sub>600nm</sub>) au moment de l'induction et/ou la température et le temps d'induction, ne semble avoir aucune influence sur l'expression de la protéine h-APN<sub>65-967</sub>.

Le fait que la protéine s'agrège en corps d'inclusion a toutefois des avantages, elle est protégée des dégradations par les protéases et elle est facile à purifier car son degré de pureté est élevé. Mais cela nécessite des étapes supplémentaires de solubilisation et de renaturation *in vitro* de la protéine recombinante<sup>119</sup>.

<sup>119</sup> Burgess, R. R., 2009. *Methods in Enzymology*. **463**, 259-282.



## 2.3. Renaturation de la protéine recombinante

### 2.3.1. Purification et solubilisation des corps d'inclusion

La purification des corps d'inclusion n'est pas une étape difficile ; en effet, ces agrégats protéiques sont facilement isolés par centrifugation après la lyse des bactéries.

L'étape suivante de solubilisation est en revanche plus critique et empirique. Elle nécessite l'utilisation d'agent dénaturant, comme l'urée. Différentes concentrations en urée (4, 6 et 8 M) sont testées en parallèle pour déterminer la concentration minimale nécessaire à la solubilisation de l'agrégat protéique. En effet, la renaturation *in vitro* est plus facile lorsque la protéine n'est que partiellement dénaturée. Dans notre cas, une concentration en urée de 8 M est nécessaire pour solubiliser les protéines agrégées. Les essais de solubilisation sont effectués dans deux tampons différents, Tris-HCl pH 8,0 et CAPS pH 11. Par ailleurs, un agent de réduction, le dithiothreitol (DTT), est ajouté dans le tampon pour réduire et/ou éviter la formation de ponts disulfures.

### 2.3.2. Renaturation *in vitro* de la protéine recombinante

L'agent dénaturant doit être éliminé pour favoriser la renaturation *in vitro* de la protéine. Plusieurs techniques sont couramment utilisées : par dialyse, par dilution et/ou par chromatographie d'affinité. Le taux de succès est varié et dépend de la protéine. La concentration en protéine (entre 1 et 10 mg/ml) est un paramètre important et peut influencer le repliement correct au dépend de l'agrégation. Pour favoriser la formation des ponts disulfures, le dithiothreitol ou le couple d'oxydo-réduction Glutathion réduit (GSH)/oxydé (GSSG) sont ajoutés au tampon de renaturation. De nombreux autres additifs peuvent également être utilisés pour aider le repliement de la protéine (mécanisme d'action pas toujours connu) : sels, détergents (Triton X-100), saccharose, polyols, acides aminés (L-Arginine), etc.

Essai	Solubilisation des corps d'inclusion			Renaturation de la protéine par dialyse					
	C (mg/ml)	V (ml)	m (mg)	Dialyse n°	C (mg/ml)	V (ml)	m (mg)	U <sub>T</sub> (μmol/min)	AS (U <sub>T</sub> /mg)
<b>1</b>	10	4	40	1	7,8	4,5	35	0,015	0,0004
<b>2</b>	5,3	8	42	1	4,2	8,8	37	0,004	0,0001
				2	4,1	8	33	0,032	0,0010
				3	4,1	7,5	31	0,083	0,0027
				4	4,1	7,5	31	0,107	0,0035
				5	3,9	7,2	28	0,137	0,0050
				6	3,9	7,2	28	0,148	0,0054
				7	3,9	6,5	25	0,109	0,0047
<b>3</b>	8,6	10,5	90,5	3	8	11,5	92	0,027	0,0003
				7	8,1	11	89	0,026	0,0003
				9	8,1	11	89	0,032	0,0004
				10	8,3	10,8	89	0,075	0,0008

*Tableau 11 : Bilan des essais de renaturation de la protéine hAPN<sub>65-967</sub> par dialyse. Le repliement correct de la protéine est évalué par mesure de l'activité enzymatique.*

### *i. Renaturation par dilution*

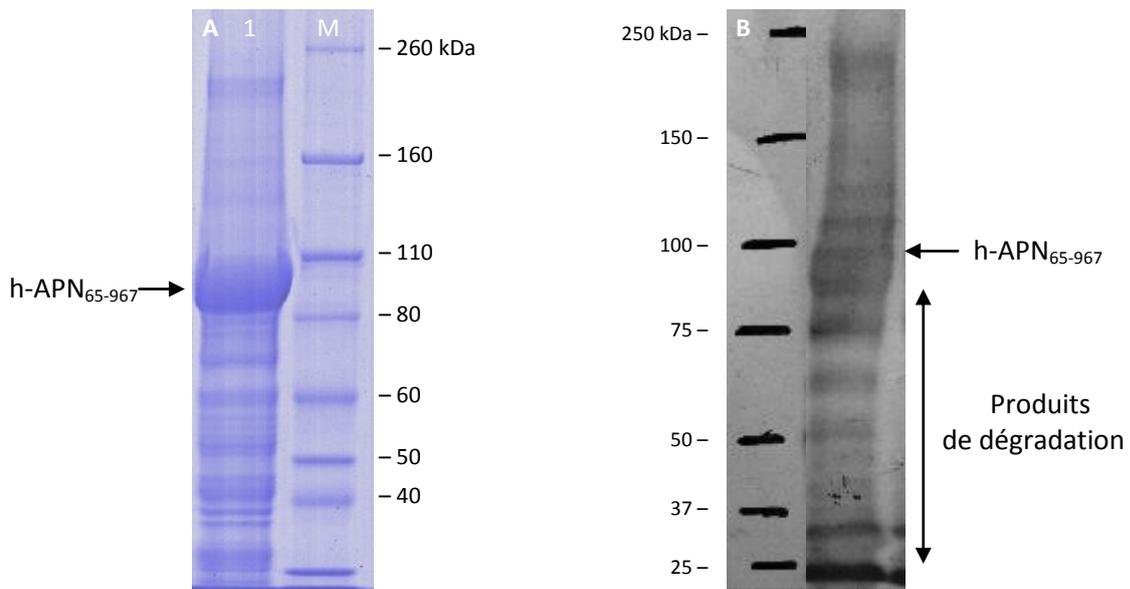
Des essais de renaturation de la protéine h-APN<sub>65-967</sub> par dilution sont effectués. Afin de trouver les conditions optimales, différents tampons sont testés en parallèle (kit QuickFold Protein Refolding, Athena ES). Le repliement correct de la protéine est vérifié par dosage de l'activité enzymatique. L'analyse des résultats permet de conclure quant-à l'utilité ou non de certains composés dans le tampon :

- La solution MES (pH 6,0) ne favorise pas la renaturation de la protéine.
- Il est difficile de tirer des conclusions quant-aux autres conditions de renaturation. (tampon Tris-HCl pH 8,5). L'analyse par SDS-PAGE après concentration de la fraction soluble révèle la présence d'une bande vers 100-105 kDa caractéristique de la protéine hAPN<sub>65-967</sub>. Cependant, aucune activité enzymatique n'est décelée dans ces fractions, ce qui suggère que la protéine n'a pas adoptée sa conformation native.

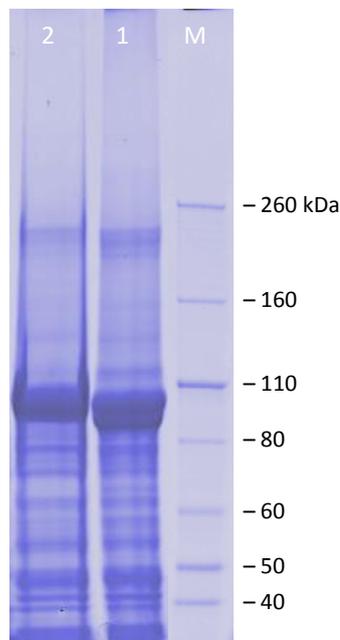
Comme décrit précédemment, l'activité catalytique de la protéine dépend de la présence d'un atome de zinc dans son site actif. C'est pourquoi, des essais sont réalisés en ajoutant du chlorure de zinc (ZnCl<sub>2</sub>) dans les tampons de renaturation. Malgré tout, des résultats similaires aux précédents sont observés ; dans ces conditions le chlorure de zinc n'a aucune influence sur le processus de renaturation de la protéine.

### *ii. Renaturation par dialyse*

En parallèle, des tests de renaturation par dialyse sont réalisés ; la protéine solubilisée dans du tampon CAPS, pH 11,0 est dialysée dans du tampon 20 mM Tris-HCl, pH 8,0, 100 µM ZnCl<sub>2</sub>, 1mM GSH, 0,2 mM GSSG. La mesure de l'activité enzymatique révèle la présence de protéine correctement repliée, mais en faible proportion. Le rendement de renaturation est amélioré en effectuant des dialyses successives (Tableau 11). Cependant, l'activité catalytique de la protéine renaturée reste faible et varie énormément selon les essais (AS  $\sim 10^{-3}$  U/mg alors que l'activité spécifique de l'APN native est de l'ordre de 10 U/mg).



**Figure 52 :** **A :** Analyse sur gel SDS-PAGE (NuPAGE® 3-8% Tris-Acétate). 30 µg de protéines totales sont déposés dans le puits. **M :** Marqueur de poids moléculaire. Puits n°1 : Renaturation de la protéine par dialyse. **B :** Transfert de protéines et immunodétection (Western Blot) à l'aide d'un anticorps polyclonal dirigé contre la protéine h-APN.



**Figure 53 :** Analyse sur gel SDS-PAGE (NuPAGE® 3-8% Tris-Acétate). 22,5 µg de protéines totales sont déposés dans chaque puits. **M :** Marqueur de poids moléculaire. **1 :** Renaturation de la protéine par dialyse. **2 :** Chromatographie d'affinité sur résine de nickel : fractions récoltées lors de la charge.

Après séparation de la fraction soluble et active par centrifugation, le degré de pureté est analysé par électrophorèse sur gel de polyacrylamide ([Figure 52-A](#)). La protéine h-APN<sub>65-967</sub> (bande intense entre 100 et 105 kDa) est facilement identifiée. Cependant, l'échantillon protéique n'est pas pur ; le gel révèle, en effet, la présence de nombreux contaminants protéiques (bandes de poids moléculaire inférieur ou supérieur à la protéine d'intérêt). Une éventuelle dégradation (partielle) de la protéine h-APN<sub>65-967</sub> par des protéases endogènes n'est pas à exclure car, en règle générale, les corps d'inclusions contiennent la protéine recombinante à un degré de pureté très élevé. L'immunodétection des protéines (Western-Blot), réalisée à l'aide d'un anticorps dirigé contre les résidus 668 à 967 de h-APN, permet de confirmer cette hypothèse ([Figure 52-B](#)).

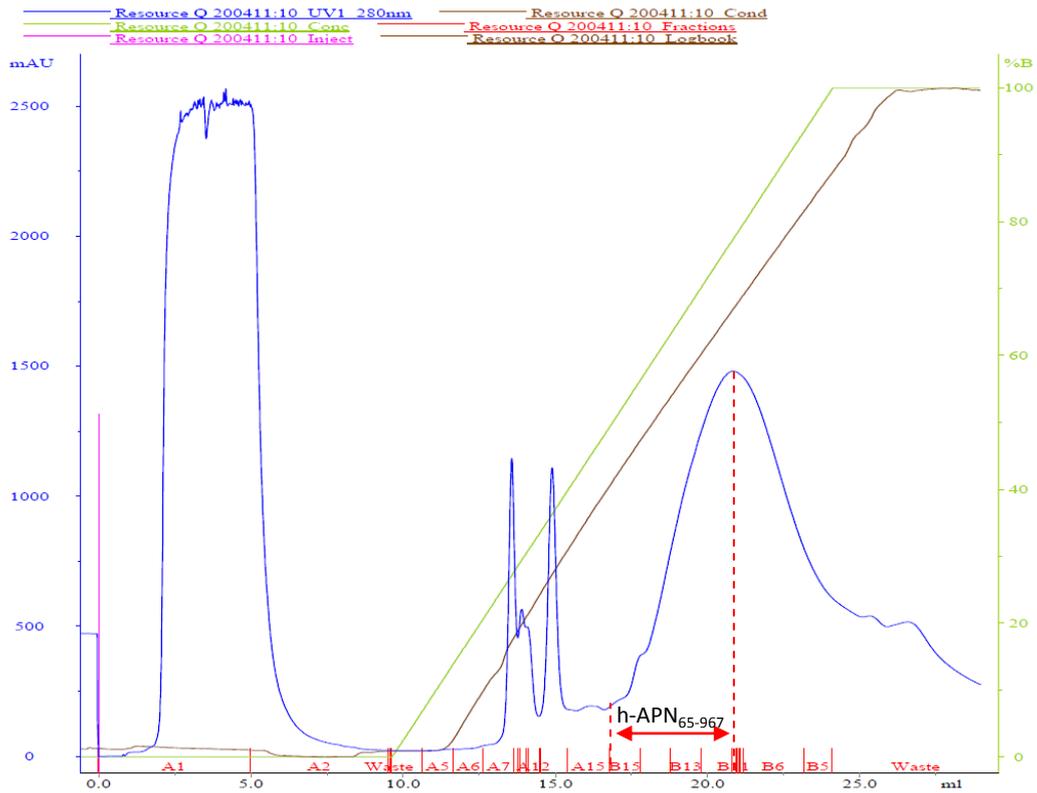
Malgré la faible proportion de protéine active dans les échantillons, différentes techniques de purification sont utilisées pour essayer de purifier la protéine h-APN<sub>65-967</sub>.

## 2.4. Purification de la protéine recombinante

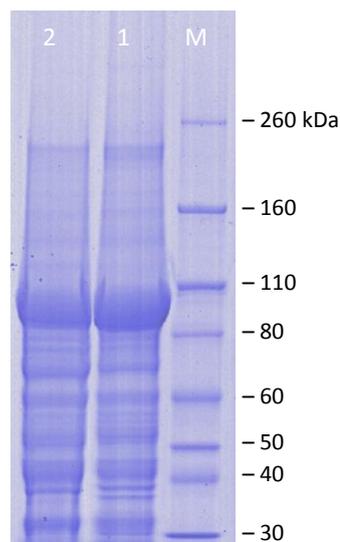
### *2.4.1. Chromatographie d'affinité sur résine de nickel*

La queue polyhistidine fusionnée à l'extrémité N-terminale de la protéine h-APN<sub>65-967</sub> facilite sa purification par chromatographie d'affinité sur résine de nickel : la protéine d'intérêt est séparée de celles qui sont dégradées du côté N-terminale (queue polyhistidine absente).

Dans un premier temps, les additifs (couple d'oxydo-réduction GSH/GSSG et ZnCl<sub>2</sub>) présents dans le tampon de renaturation sont éliminés par dialyse puis, l'échantillon protéique est chargé sur la résine de nickel HisTrap FF 5 ml (GE Healthcare). Les protéines retenues, via les interactions ioniques entre les résidus histidine des protéines et les ions nickel de la résine, sont éluées à l'aide d'un gradient linéaire d'Imidazole (de 0 à 500 mM). La chromatographie est suivie par mesure de l'absorbance à 280 nm. L'analyse par dosage de l'activité enzymatique des fractions récoltées, ainsi que leur dépôt sur gel SDS-PAGE, montrent que la protéine h-APN<sub>65-967</sub> n'est pas retenue sur la résine ([Figure 53](#)). La queue polyhistidine ne semble pas accessible



**Figure 54 :** Chromatographie échangeuse d’anions sur résine Resource Q 1ml (GE Healthcare). La courbe bleue représente la densité optique à 280nm (exprimée en mAU). Les courbes verte et marron représentent respectivement le gradient de sel (en % de NaCl) et la conductimétrie (mS/cm). Les protéines chargées négativement sont retenues sur la résine. Elles sont éluées par un gradient linéaire de sel (de 0 à 500 mM en NaCl).



**Figure 55 :** Analyse sur gel SDS-PAGE (NuPAGE® 3-8% Tris-Acétate). 30 µg de protéines totales sont déposés dans chaque puits. **M** : Marqueur de poids moléculaire. **1** : Renaturation de la protéine par dialyse. **2** : Chromatographie échangeuse d’anions : fractions B15 à B12 récoltées lors de l’éluion.

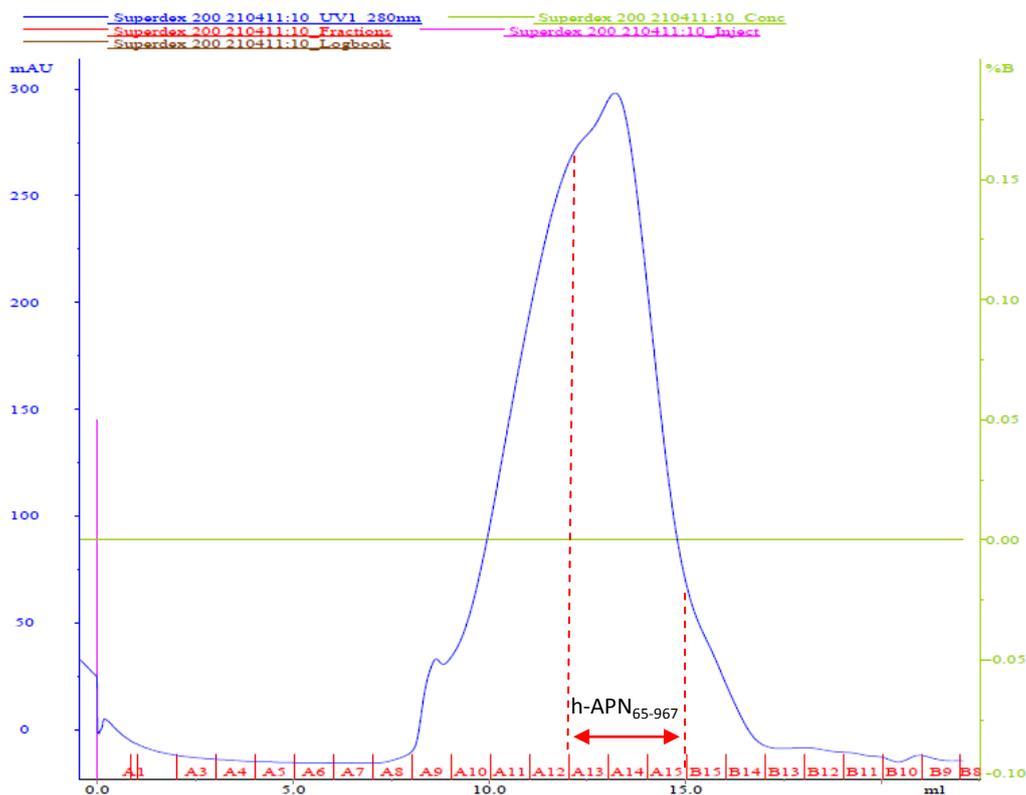
#### 2.4.2. Chromatographie échangeuse d'anions

A un pH supérieur à son point isoélectrique ( $pI=5,3$ ), la protéine h-APN<sub>65-967</sub> est chargée négativement, et peut donc être séparée par chromatographie échangeuse d'anions.

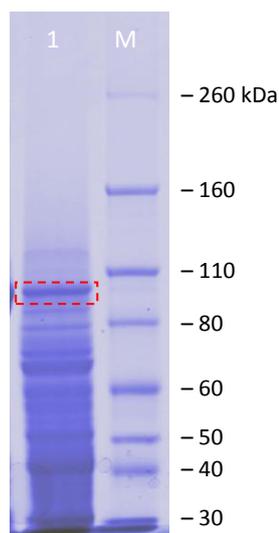
Après dialyse dans un tampon dépourvu de sels, le milieu protéique est chargé sur une résine échangeuse d'anions de type Resource Q 1 ml (GE Healthcare). L'analyse des fractions récoltées par mesure de l'activité enzymatique permet d'identifier la protéine h-APN<sub>65-967</sub> : elle est éluée pour une concentration en sel comprise entre 0,25 et 0,4 M (fractions B15 à B12) (Figure 54).

Les fractions contenant la protéine d'intérêt (B15 à B12) sont mélangées et analysées sur gel SDS-PAGE (Figure 55).

Plusieurs essais avec des gradients différents en sel ont été réalisés, mais aucune amélioration de la séparation n'a été observée. Cette technique ne permettant pas d'éliminer les différents contaminants protéiques, la chromatographie d'exclusion a été choisie pour poursuivre la purification.



*Figure 56* : Chromatographie d'exclusion Superdex 200 10/300 (GE Heathcare). La courbe bleue représente la densité optique à 280nm (exprimée en mAU).



*Figure 57* : Analyse sur gel SDS-PAGE (NuPAGE® 3-8% Tris-Acétate). 22,5 µg de protéines totales sont déposés dans chaque puits. **M** : Marqueur de poids moléculaire. **1** : Chromatographie d'exclusion : fractions A13 à A15 récoltées lors de l'éluion.

### 2.4.3. Chromatographie d'exclusion

Les fractions B15 à B12 (4 ml) sont dessalées et concentrées (volume finale 500 µl). L'échantillon protéique est ensuite déposé sur le gel Superdex 200 10/300 (GE Healthcare) dont la gamme de séparation s'étend de 10 à 600 kDa ; les protéines sont éluées par ordre décroissant de poids moléculaire. L'allure du chromatogramme, obtenu par mesure de l'absorbance à 280 nm, indique que la séparation n'est pas optimale : un seul pic d'éluion est visible. D'après l'analyse des fractions récoltées par dosage de l'activité enzymatique, la protéine h-APN<sub>65-967</sub> est majoritairement présente dans les fractions A13 à A15 (fin du pic d'éluion) (Figure 56).

Les fractions contenant la protéine d'intérêt sont mélangées et déposées sur gel SDS-PAGE (Figure 57). Cette technique permet d'éliminer les contaminants protéiques de haut poids moléculaire mais les fragments protéiques issus de la dégradation de la protéine h-APN<sub>65-967</sub> par des protéases ne sont pas séparés. Au fur à mesure de l'avancé de la purification, il semblerait que la protéine soit dégradée par des protéases.

Après deux étapes (chromatographie échangeuse d'anions et d'exclusion), le rendement de purification est faible ; par ailleurs le degré de pureté est insuffisant pour entamer des essais de cristallisation. D'autres méthodes de purifications ont été utilisées, telles que la précipitation au sulfate d'ammonium ou la chromatographie à interactions hydrophobes, mais des résultats similaires à ceux décrits ci-dessus sont observés.

## 2.5. Conclusion et discussion

Les différentes stratégies, employées pour exprimer le domaine soluble de l'APN sous une forme active chez *E. coli*, n'ont pas abouti. Les protéines insolubles et partiellement dégradées s'accumulent dans les corps d'inclusion. Ce phénomène de dégradation par des protéases au sein même des corps d'inclusion a déjà été mis en évidence<sup>120,121</sup>. Il est difficile, après les étapes de solubilisation et de renaturation *in vitro* de séparer la protéine d'intérêt des fragments protéiques, issus de la dégradation par des protéases.

---

<sup>120</sup> Cubarsi, R., Carrio, M., and Villaverde, A. (2001). *Biochem. Biophys. Res. Commun.* **282**, 436-411.

<sup>121</sup> Carrio, M., Corchero, J., and Villaverde, A. (1999). *Biochim. Biophys. Acta.* **1434**, 170-176.



En parallèle à ces travaux, d'autres stratégies auraient pu être développées pour favoriser l'expression de la protéine sous une forme soluble et active. En effet, plusieurs techniques existent pour minimiser l'agrégation de protéines eucaryotes dans un état biologiquement inactif<sup>122</sup> :

- Expression périplasmique : contrairement au cytoplasme, le périplasma est un milieu oxydant favorable à la formation de ponts disulfures<sup>123</sup>. Le clonage du gène d'intérêt en aval d'une séquence signal permet d'exporter la protéine dans le périplasma de la bactérie. Les résidus cystéine sont maintenus durant leur transport sous une forme réduite par les thiorédoxines et les glutarédoxines cytoplasmiques. Une fois dans le périplasma, les protéines de la famille Dsb vont aider la protéine recombinante à adopter sa conformation native en catalysant la formation de ponts disulfures. Néanmoins, le rendement de la sécrétion est faible, et dans certains cas la protéine s'accumule dans l'espace périplasmique sous forme de corps d'inclusion.
- Protéines de fusion : elles sont souvent employées pour augmenter le rendement de production de protéines recombinantes sous une forme native et soluble. En effet, la fusion de la protéine d'intérêt par son extrémité N-terminale à une protéine soluble, telle que la glutathion S-transférase (GST), favorise parfois le repliement correct de la protéine.
- Co-expression de protéines chaperonnes : elles peuvent influencer et améliorer la solubilité des protéines recombinantes (Hsp60, complexes GroEL/GroES ; Hsp70, DnaK)<sup>124</sup>.

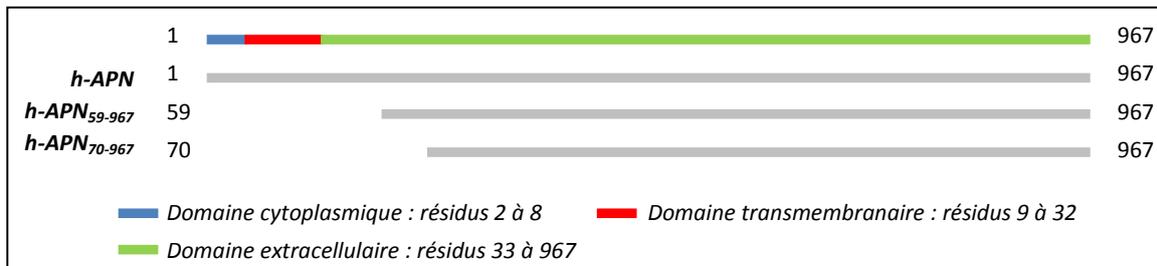
En dépit des avantages que présentent les techniques évoquées ci-dessus, c'est l'expression de la protéine chez les eucaryotes supérieurs qui a été choisie ; ce travail s'est effectué en partenariat avec l'équipe du Prof. Udo Oppermann du Centre de Recherche Botnar de l'Université d'Oxford. Les chercheurs de ce laboratoire travaillent en étroite collaboration avec le Consortium de Génomique Structurale d'Oxford. Ces derniers ont déjà résolu la structure tridimensionnelle de nombreuses protéines à visées thérapeutiques dont ERAP1 (une autre aminopeptidase de la famille M1). Les travaux décrits ci-dessous ont été réalisés lors d'un séjour de 3 mois dans l'équipe du Prof Udo Oppermann.

---

<sup>122</sup> Sorensen, H. P., and Mortensen, K. K., 2005. *Microb Cell Fact.* **4**(1), 1.

<sup>123</sup> Berkmen, M., 2012. *Protein Expression and Purification.* **82**, 240-251.

<sup>124</sup> Kyratsous, C. A., Silverstein, S. J., DeLong, C. R., and Panagiotidis, C. A., 2009. *Gene.* **440**, 9-15.



*Figure 58 : Expression de la protéine membranaire h-APN et de diverses formes solubles.*

### 3. Expression de la protéine recombinante dans baculovirus

Le système baculovirus/cellules d'insectes est un des plus performants pour la production de protéines eucaryotes complexes : sa mise en œuvre est simple et rapide et le taux de production est élevé (centaines de mg/l de culture). Les protéines produites subissent des modifications post-traductionnelles quasi complètes parmi lesquelles l'adressage par des peptides signaux, le repliement, la formation de ponts disulfures ou la glycosylation. Les N-glycosylation produites par les cellules de *Spodoptera frugiperda* Sf9 sont similaires à celles qu'on retrouve sur les cellules de mammifères à la différence près que l'acide sialique est modifié par un hydroxyde.

Le système d'expression « Bac to Bac® Baculovirus Expression System » de Invitrogen est utilisé pour générer les baculovirus recombinants qui vont exprimer la protéine d'intérêt (cf. Matériels et Méthodes p216).

Afin d'augmenter les taux de réussite, diverses formes de h-APN sont surexprimées : une forme membranaire ainsi que deux formes solubles h-APN<sub>59-967</sub> et h-APN<sub>70-967</sub> (suppression des domaines cytoplasmique et transmembranaire et d'une partie de la région extracellulaire). Les formes solubles ont été décrites dans la littérature<sup>125</sup>, elles sont respectivement libérées dans le sérum humain (résidus 59 à 967) et maternel (résidus 70 à 967) suite à l'action d'une protéase (Figure 58).

#### 3.1. Clonage

Les fragments d'ADN, qui codent pour la version complète de la protéine et ses différentes formes tronquées, sont clonés dans les vecteurs d'expression baculovirus pFB-Lic-Bse et pFB-CT10HF-Lic en utilisant la technologie de clonage sans ligase (« Ligation Independent Cloning »). Pour faciliter la purification ultérieure des protéines, les gènes d'intérêt sont clonés en aval d'une séquence codant pour 6 Histidine (pFB-Lic-Bse), ou bien en amont d'une séquence codant pour 6 Histidine et le peptide FLAG (pFB-CT10HF-Lic).

---

<sup>125</sup> Watanabe, Y., Ito, K., Iwaki-Egawa, S., Yamaguchi, R., and Fujimata, Y. (1998). *Molecular genetics and metabolism*. **63**, 289-294.

Plasmide recombinant	Concentration en ng/ $\mu$ L	Volume ( $\mu$ l)	Masse ( $\mu$ g)
pFB-Lic-Bse-ANPEP	451	40	22,5
pFB-Lic-Bse-ANPEP <sub>175-2904</sub>	464		19,5
pFB-Lic-Bse-ANPEP <sub>208-2904</sub>	389		23,2
pFB-CT10HF-Lic-ANPEP	56		2,8
pFB-CT10HF-Lic-ANPEP <sub>175-2904</sub>	264		13,2
pFB-CT10HF-Lic-ANPEP <sub>208-2904</sub>	109		8,1

*Tableau 12 : Bilan de l'extraction et de la purification de l'ADN plasmidique*

### 3.1.1. Isolement du gène d'intérêt

L'ADN polymérase haute-fidélité Platinum® *Pfx* (Invitrogen) est utilisé pour amplifier par PCR le gène d'intérêt qui code pour la protéine membranaire et ses deux formes solubles. La banque d'ADNc, collection MGC/IMAGE ID 6020902) est employée comme matrice ; les extensions LIC sont introduits à chaque extrémité du fragment amplifié pour le clonage sans ligase dans les vecteurs pFB-Lic-Bse et pFB-CT10HF-Lic. Des fragments d'environ 2940 pb, 2765 pb (h-APN<sub>59-967</sub>) et 2735 pb (h-APN<sub>70-967</sub>) sont amplifiés ; ils sont extraits du gel d'agarose et purifiés à l'aide du kit QIAquick Gel Extraction (QIAGEN).

### 3.1.2. Clonage dans les vecteurs baculovirus

Les vecteurs pFB-LIC-Bse et pFB-CT10HF-LIC sont respectivement digérés par les enzymes *BseRI* et *BfuAI*, dont les sites de restriction sont présents au niveau des extensions LIC (séquences d'oligonucléotides dans lesquelles une des bases azotées est absente). Les vecteurs linéaires sont purifiés sur une colonne de silice, puis les extrémités sont rendues cohésives par action de l'ADN polymérase T4 en présence de la base manquante. Les extensions LIC, introduits à chaque extrémité du fragment amplifié, sont également traitées par l'ADN polymérase T4 : des extrémités 5' cohésives complémentaires à celles des vecteurs sont ainsi générées.

La ligation s'effectue sans ajout d'ADN ligase ; l'ADN polymérase du phage T4 présent dans le milieu réactionnel est suffisant à la formation de liaisons phosphodiester entre les extrémités 5' phosphate et 3' hydroxyle des deux brins d'ADN. Les plasmides recombinants sont ensuite introduits par choc thermique dans les bactéries XL1 Blue ; les colonies bactériennes qui se sont multipliées sur milieu sélectif solide sont criblées par PCR avec des amorces spécifiques au gène d'intérêt. Des fragments d'environ 2940 pb, 2765 pb (h-APN<sub>59-967</sub>) et 2735 pb (h-APN<sub>70-967</sub>) sont amplifiés lorsque le gène d'intérêt est introduit dans le vecteur.

Les clones positifs sont inoculés dans du milieu liquide sélectif et remis en culture. Les bactéries sont ensuite lysées par un traitement alcalin, et les ADN plasmidiques sont extraits et purifiés sur une résine échangeuse d'anions (miniprep, QIAGEN). Les résultats obtenus sont indiqués dans le tableau ci-contre.

Bacmide recombinant	Concentration en ng/ $\mu$ L	260/280
pFB-Lic-Bse-ANPEP colonie 1	938	2,00
pFB-Lic-Bse-ANPEP colonie 2	1002	2,05
pFB-Lic-Bse-ANPEP <sub>175-2904</sub> colonie 1	1055	2,04
pFB-Lic-Bse-ANPEP <sub>175-2904</sub> colonie 2	1055	2,02
pFB-Lic-Bse-ANPEP <sub>208-2904</sub> colonie 1	1099	2,02
pFB-Lic-Bse-ANPEP <sub>208-2904</sub> colonie 2	808	2,04
pFB-CT10HF-Lic-ANPEP colonie 1	949	2,0
pFB-CT10HF-Lic-ANPEP colonie 2	790	2,0
pFB-CT10HF-Lic-ANPEP <sub>175-2904</sub> colonie 1	1040	1,92
pFB-CT10HF-Lic-ANPEP <sub>175-2904</sub> colonie 2	1055	2,04
pFB-CT10HF-Lic-ANPEP <sub>208-2904</sub> colonie 1	1078	2,02
pFB-CT10HF-Lic-ANPEP <sub>208-2904</sub> colonie 2	1184	2,03

*Tableau 13 : Bilan de l'extraction et de la purification des bacmides recombinants.*

## 3.2. Génération des bacmides recombinants

La cassette d'expression, contenant le gène d'intérêt, est intégrée dans le génome du baculovirus *AcMNPV* par transposition site-spécifique. L'intégration s'effectue après transformation, par choc thermique, des plasmides recombinants dans les cellules compétentes *E. coli* DHBac10. Ces bactéries hébergent le génome du baculovirus sous la forme d'un bacmide et auquel a été ajouté le gène *lacZ* entre deux sites de transposition Tn7. Cette souche héberge également un plasmide « helper » qui code pour une transposase ; cette dernière permet la transposition spécifique de la cassette d'expression dans le bacmide au niveau du site mini-*att*Tn7, situé dans le gène *lacZ* (qui code pour la  $\beta$ -galactosidase). Les colonies bactériennes qui se sont multipliées sur un milieu sélectif solide en présence d'IPTG et d'un substrat de la  $\beta$ -galactosidase sont sélectionnées par un criblage bleu/blanc. Les clones positifs (colonies blanches) sont inoculés dans du milieu liquide sélectif et remis en culture. Les bactéries sont ensuite lysées par un traitement alcalin, et les bacmides recombinants sont extraits et purifiés. Les résultats obtenus sont indiqués dans le tableau ci-dessous.

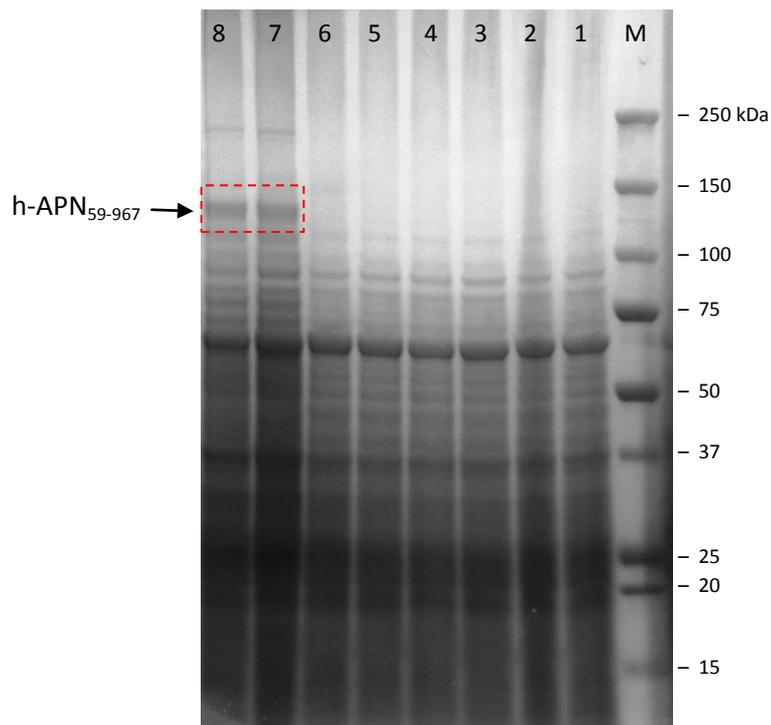
La transposition de la cassette d'expression dans le bacmide est également vérifiée par PCR en utilisant des amorces situées sur le bacmide de part et d'autres du site de transposition mini-*att*Tn7.

Un fragment d'environ 5000 pb est amplifié lorsque la transposition de la cassette d'expression dans le bacmide s'est produite ; dans le cas contraire, seul le site de transposition mini-*att*Tn7 est amplifié (fragment de 300 pb).

## 3.3. Culture cellulaire et expression de la protéine recombinante

### 3.3.1. Transfection des cellules d'insectes Sf9

La lignée cellulaire Sf9, issue de cellules ovariennes du lépidoptère *Spodoptera frugiperda*, est utilisée pour générer les baculovirus recombinants qui vont produire en grande quantité la protéine recombinante. La transfection des différents bacmides recombinants dans les cellules d'insectes, cultivées en monocouche dans un milieu sans sérum, est réalisée avec le kit Nanofectin (PAA). Après 72 heures d'incubation à 27°C, le surnageant est séparé des débris cellulaires par centrifugation : il constitue le stock viral P0 et ne contient qu'une faible quantité de baculovirus recombinant qui doit être amplifiée.



*Figure 59* : Analyse sur gel SDS-PAGE (Criterion™ XT precast gel- 4-12% Bis-Tris). 15 µl de la fraction soluble après lyse cellulaire sont déposés dans chaque puits. **M** : Marqueur de poids moléculaire. **Puits 1 & 2** : clonage dans le vecteur pFB-Lic-Bse, expression de la forme soluble (résidus 59 à 967). **Puits 3 & 4** : clonage dans le vecteur pFB-Lic-Bse, expression de la forme soluble (résidus 70 à 967). **Puits 5 & 6** : clonage dans le vecteur pFB-CT10HF-Lic, expression de la forme soluble (résidus 70 à 967). **Puits 7 & 8** : clonage dans le vecteur pFB-CT10HF-Lic, expression de la forme soluble (résidus 59 à 967).

### *3.3.2. Amplification du stock viral*

Les baculovirus recombinants sont amplifiés en infectant les cellules d'insectes Sf9, cultivées en suspension dans un milieu sans sérum, avec le stock viral P0. Après 72h d'incubation à 27°C sous agitation, le surnageant (stock viral P1) qui contient une concentration plus élevée de baculovirus recombinants est récupéré et conservé à 4°C dans l'obscurité, car ces baculovirus sont sensibles à la lumière. Pour vérifier si l'infection s'est produite, la densité cellulaire est estimée avant et après infection ; en effet, la croissance des cellules d'insectes est interrompue suite à leur infection par un baculovirus recombinant. Le stock viral P1 est à son tour amplifié ; le stock de baculovirus recombinants obtenu (stock viral P2) est suffisamment concentré pour infecter un grand nombre de cellules Sf9 et produire la protéine recombinante.

### *3.3.3. Expression de la protéine recombinante*

Suite à l'infection des cellules Sf9 par le stock de baculovirus recombinant P2, les protéines devraient être exprimées avec des modifications post-traductionnelles soit dans le cytoplasme des cellules (formes soluble) soit ancrées dans la membrane (forme membranaire).

Les cellules d'insectes sont lysées dans un tampon de lyse, puis les protéines solubles sont séparées des protéines insolubles par centrifugation. Le niveau d'expression est analysé sur gel SDS-PAGE ; une bande intense est attendue vers 130-140 kDa pour la protéine membranaire (fraction insoluble) et 120-130 kDa pour les formes solubles h-APN<sub>59-967</sub> et h-APN<sub>70-967</sub> (fraction soluble) (Figure 59).

Les baculovirus recombinants, obtenus après transfection des cellules d'insectes avec le bacmide recombinant pFB-CT10HF-Lic-ANPEP<sub>175-2904</sub>, permettent d'exprimer la protéine recombinante : il s'agit de la forme soluble présente dans le sérum humain (résidus 59 à 967). La protéine h-APN<sub>59-967</sub> (bande intense vers 125 kDa) est facilement identifiée sur le gel. Les autres constructions n'ont pas permis de produire la protéine recombinante. De plus, l'analyse sur gel SDS-PAGE des fractions insolubles après lyse cellulaire ne permet pas de déceler la présence de la protéine membranaire.



Le stock de baculovirus recombinant P2 (pFB-CT10HF-Lic-ANPEP<sub>175-2904</sub>) est utilisé pour infecter des gros volumes de cellules Sf9. Après 96 heures d'incubation, les culots cellulaires sont récupérés par centrifugation et envoyés en France afin d'extraire et de purifier la protéine h-APN<sub>59-967</sub>.

### 3.4. Purification de la protéine recombinante

#### 3.4.1. Chromatographie d'affinité

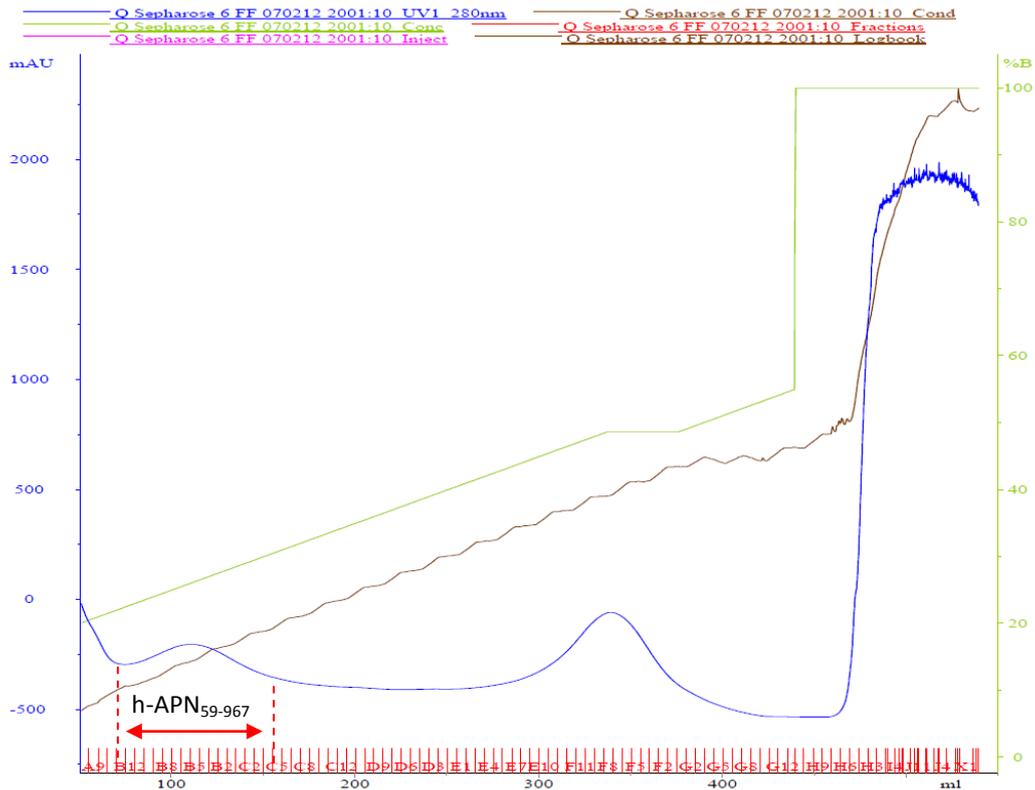
Les étiquettes polyhistidine et FLAG fusionnées à l'extrémité C-terminale de la protéine permettent sa purification par chromatographie d'affinité sur résine de nickel ou sur une résine anti-FLAG. L'efficacité de séparation de ces deux types de résine est étudiée. Les essais sont effectués à partir d'un culot cellulaire, provenant d'une culture de 500 mL. Après lyse cellulaire, la fraction soluble contenant la protéine hAPN<sub>59-967</sub> est séparée des débris insolubles par centrifugation puis, injectée sur une résine de nickel (HisTrap FF 5ml) ou incubée en présence de billes d'agarose recouvertes d'anticorps anti-FLAG. Les protéines fixées sur la résine de nickel sont éluées par un gradient linéaire de sel (de 0 à 500 mM en Imidazole); celles retenues via les interactions spécifiques antigène/anticorps sont détachées par compétition avec un peptide FLAG. L'analyse des fractions récoltées, par dosage de l'activité enzymatique et sur gel SDS-PAGE, indique que la protéine hAPN<sub>65-967</sub> n'est pas retenue sur les résines.

Après de nombreux essais et une longue réflexion, il s'est avéré qu'une erreur a été commise lors de la construction du plasmide recombinant, le codon STOP toujours présent à la fin de la séquence empêche l'expression des étiquettes polyhistidine et FLAG.

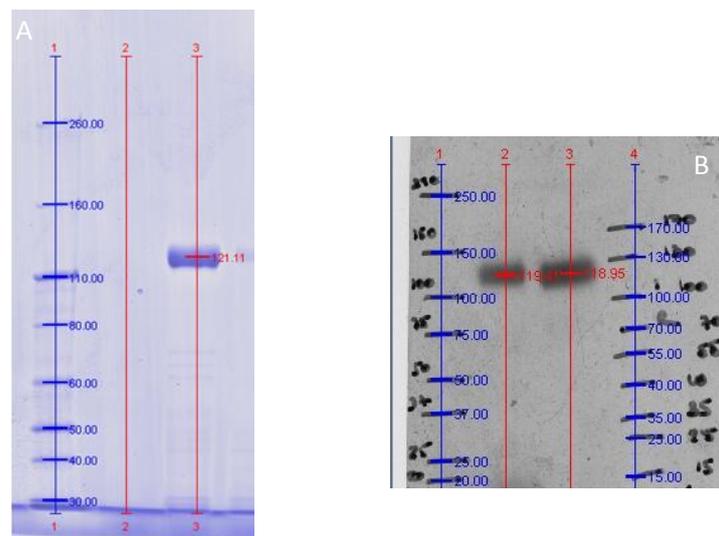
Par conséquent, d'autres techniques de purification sont utilisées. Au terme de plusieurs étapes d'optimisation, la chromatographie échangeuse d'anions s'est révélée être une des techniques les plus efficaces.

#### 3.4.2. Purification de la protéine hAPN<sub>59-967</sub>

La protéine h-APN<sub>59-967</sub> est extraite des cellules d'insectes et purifiée selon le protocole décrit ci-dessous. Durant les différentes étapes de purification, les échantillons sont maintenus sur glace ou à 4°C pour éviter une dégradation par les protéases endogènes.



**Figure 60** : Chromatographie échangeuse d’anions sur résine Q Sepharose 6 FF 50 ml (GE Healthcare). La courbe bleue représente la densité optique à 280nm (exprimée en mAU). Les courbes verte et marron représentent respectivement le gradient de sel (en % de NaCl) et la conductimétrie (mS/cm). Après injection de l’échantillon, la colonne est lavée avec du tampon (50 mM Tris-HCl, pH 7,5) jusqu’au retour du signal de détection UV 280 nm à la ligne de base. Les protéines chargées négativement à pH 7,5 et qui sont fixées sur la résine via des interactions ioniques, sont éluées par un gradient linéaire de sel (de 0 à 0,275 M en NaCl) sur 8 fois le volume de la colonne.



**Figure 61** : **A**. Analyse sur gel SDS-PAGE (NuPAGE® 3-8% Tris-Acétate). 1 µg de l’échantillon protéique est déposé dans le puits. **1** : Marqueur de poids moléculaire : 260, 160, 110, 80, 60, 50, 40 et 30 kDa. **3** : Echantillon protéique obtenu à l’issue des différentes étapes de purification. **B** : Transfert de protéines et immunodétection (Western Blot) à l’aide d’un anticorps polyclonal dirigé contre la protéine h-APN.

Le culot cellulaire, provenant de 5 l de culture, est resuspendu dans du tampon de lyse. Après plusieurs heures d'homogénéisation, la fraction soluble est isolée et concentrée par précipitation au sulfate d'ammonium (50 % de saturation en sel). Le culot, obtenu par sédimentation des protéines précipitées, est repris dans un minimum de tampon 50 mM Tris-HCl, pH 7,5 et dialysé pour éliminer le sulfate d'ammonium résiduel. L'échantillon protéique est injecté sur une résine échangeuse d'anions Q Sepharose (50 ml). Les fractions récoltées sont analysées par dosage de l'activité enzymatique ; celles contenant majoritairement la protéine hAPN<sub>59-967</sub> (fractions A12 à C4) sont réunies (Figure 60). L'excédent de sel est éliminé par ultrafiltration sur une cellule Amicon Ultra qui est composée d'une membrane ayant un seuil de coupure de 100 kDa. Cette dernière étape permet également de concentrer la protéine et de la séparer d'éventuels contaminants protéiques de poids moléculaire inférieur à 100 kDa.

Les résultats sont indiqués dans le tableau de purification ci-dessous : le degré de pureté est amélioré au cours de la purification (activité spécifique finale de 13,4) au détriment du rendement qui diminue (la perte de matériel est inévitable).

Etapes	Dosage des protéines			Activité enzymatique				
	C (mg/ml)	V (ml)	m (mg)	U/ml	UT (μmol/min)	AS (U/mg)	Rendement de purification (%)	Taux de purification
Fraction soluble	23,9	500	11942	0,7	327,1	0,027	/	/
Concentration par précipitation au sulfate d'ammonium	73,3	120	8798	2,5	217,5	0,025	66	0,9
Chromatographie échangeuse d'anions	0,3	85	26	1	83,1	3,2	25,4	118,5
Concentration par ultrafiltration	14,9	0,27	4	199,6	53,9	13,4	16,5	496,3

*Tableau 14 : tableau de purification n°1*

D'après l'analyse sur gel SDS-PAGE, la pureté de la protéine hAPN<sub>59-967</sub> est suffisante (supérieure à 98%) pour entamer des essais de cristallisation (Figure 61-A). L'immunodétection des protéines (Western Blot) permet de confirmer que la protéine purifiée est bien l'APN humaine (Figure 61-B).

Pour disposer d'une grande quantité de protéine pure, deux autres purifications sont réalisées à partir de culots cellulaires, provenant respectivement de cultures de 2,5 et 2L, selon le même protocole que décrit ci-dessus.

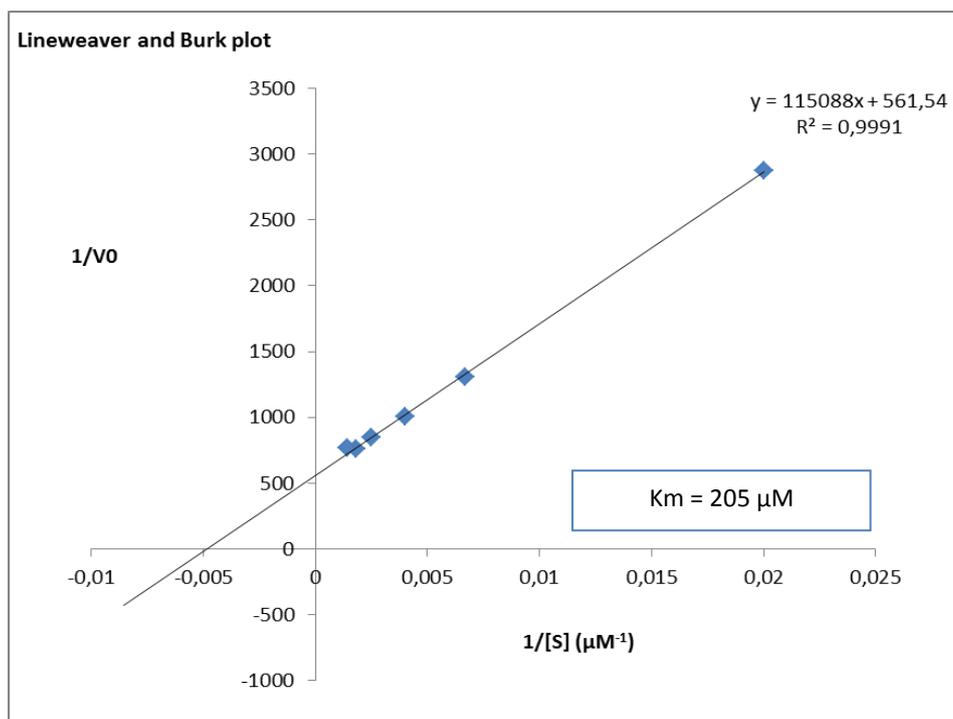


Figure 62 : Représentation en double inverse de Lineweaver et Burk : détermination des paramètres cinétiques  $K_M$  et  $V_{max}$  pour la protéine h-APN<sub>59-967</sub>.

Enzyme	$K_M$ (mM)	$k_{cat} (s^{-1}) = V_{max}/[E_T]$	$k_{cat}/K_M (s^{-1} \cdot M^{-1})$
APN de rein de porc (Sigma-Aldrich)	~0,2	$1,6 \times 10^5$	$8 \times 10^8$
h-APN <sub>59-967</sub> purifiée (AS= 13,4)	~0,2	$2,2 \times 10^5$	$1 \times 10^9$

Tableau 15 : Paramètres cinétiques  $K_M$  et  $k_{cat}$ ; les valeurs déterminés expérimentalement sont proches des valeurs obtenues avec l'enzyme commerciale.

Inhibiteurs	h-APN <sub>59-967</sub> purifiée (AS= 13,4)	AP-N de rein de porc (Sigma-Aldrich)	AP-N recombinante humaine (R&D Systems)	AP-N recombinante de souris (R&D Systems)
<p><b>I1</b></p>	$K_i = 1 \mu\text{M}$ $IC_{50} = 3 \mu\text{M}$	$K_i = 1 \mu\text{M}$ $IC_{50} = 2 \mu\text{M}$	$K_i = 3,6 \mu\text{M}$ $IC_{50} = 9 \mu\text{M}$	$K_i = 2 \mu\text{M}$ $IC_{50} = 4 \mu\text{M}$
<p><b>I2</b></p>	$K_i = 6 \text{ nM}$ $IC_{50} = 10 \text{ nM}$	$K_i = 7 \text{ nM}$ $IC_{50} = 18 \text{ nM}$	$K_i = 13 \text{ nM}$ $IC_{50} = 35 \text{ nM}$	$K_i = 6 \text{ nM}$ $IC_{50} = 15 \text{ nM}$

Tableau 16 : Evaluation de la constante d'inhibition pour 2 inhibiteurs de l'APN ; les valeurs déterminées expérimentalement sont proches des valeurs obtenues avec les enzymes commerciales.

Une étape supplémentaire de chromatographie sur une résine échangeuse d'anions plus résolutive (Mono Q 1 ml) a été nécessaire pour obtenir des échantillons avec une pureté supérieure à 90%.

### 3.5. Caractérisation de la protéine recombinante

#### 3.5.1. Détermination des paramètres cinétiques $K_M$ et $k_{cat}$

La mesure de la constante de Michaelis-Menten  $K_M$ , qui est un paramètre cinétique propre à l'enzyme pour un substrat donné, permet de s'assurer que la protéine purifiée est bien une forme soluble de h-APN. La L-leucine p-nitroanilide est utilisée comme substrat ( $K_M = 0,2 \text{ mM}$ ) ; son hydrolyse par l'enzyme h-APN conduit à la p-nitroaniline qui absorbe à 405 nm. Après avoir évalué la quantité optimum d'enzyme nécessaire à l'hydrolyse de la L-leucine p-nitroanilide, la vitesse initiale de la réaction est mesurée pour différentes concentrations en substrat (entre 0,05 et 0,7 mM) en suivant par spectrophotométrie la cinétique d'apparition du p-nitroaniline à 405 nm. Les paramètres cinétiques sont déterminés selon l'équation en double inverse de Lineweaver et Burk ; cette représentation est une droite qui coupe l'axe des ordonnées au point  $1/V_{max}$  et l'axe des abscisses au point  $-1/K_M$  (Figure 62).

La constante de Michaelis-Menten est calculée à partir de l'équation de la droite; la valeur obtenue ( $K_M = 205 \mu\text{M}$ ) est similaire à celle déterminée sur l'enzyme de porc ( $\sim 200 \mu\text{M}$ ). La constante catalytique  $k_{cat}$  et l'efficacité catalytique de l'enzyme  $k_{cat}/K_M$  sont également évaluées. Les résultats sont en accord avec ceux obtenus pour l'enzyme commerciale de porc (Tableau 15).

#### 3.5.2. Détermination de $K_i$

Des inhibiteurs compétitifs sélectifs de h-APN synthétisés au laboratoire sont évalués sur la protéine recombinante purifiée. La vitesse initiale de la réaction est mesurée pour différentes concentrations d'inhibiteurs et pour une concentration en substrat constante ( $[S] = K_M, K_M/2$  et  $2K_M$ ). La constante d'inhibition  $K_i$  et l' $IC_{50}$  sont déterminés à l'aide de la représentation de Dixon (non représentée). Les valeurs déterminées expérimentalement sont proches des valeurs obtenues avec les enzymes commerciales (Tableau 16). Cet échantillon pourra donc être utilisé pour évaluer les nouveaux inhibiteurs synthétisés au laboratoire.



### 3.6. Conclusion et discussion

Le protocole utilisé pour exprimer la protéine h-APN dans le système baculovirus-cellules d'insectes est similaire à celui mis au point pour ERAP1 (un homologue structural). La stratégie employée a permis de produire une forme soluble de h-APN. Les étapes de purification ont été fastidieuses en raison de l'absence des queues polyhistidines et FLAG à l'extrémité C-terminale de la protéine. Il est difficile en 3 mois d'affiner toutes les étapes du clonage et d'expression de la protéine. A ce stade, de nombreuses questions persistent sur l'effet bénéfique ou non de l'absence des étiquettes polyhistidine et FLAG à l'extrémité C-terminale. De nombreux paramètres doivent encore être optimisés : titration du stock viral, expression de la protéine (temps d'incubation, expression dans les cellules d'insectes Sf21 ou High Five). Néanmoins, une quantité suffisante de protéine pure (près de 6 mg) a été obtenue pour entamer les essais de cristallisation.

## 4. Essais de cristallisation

Les essais de cristallisation sont effectués au Consortium de génomique structurale d'Oxford avec l'aide du Dr. Grazyna Kochan. Dans un premier temps, il a été décidé de conserver les N-glycanes. Pour favoriser le processus de cristallisation, la protéine est co-cristallisée avec un inhibiteur (I3). De nombreuses solutions de cristallisation sont testées en parallèle (kits de criblage HIN et JCSG5, voir annexe 3) et pour chaque condition plusieurs concentrations du mélange protéine-inhibiteur sont expérimentées. Une séparation de phase est observée dans les conditions suivantes: 0,2 M MgCl<sub>2</sub>, 0,1 M Tris, pH 8,5 et 25 % w/v polyéthylène glycol 3350 (H1-kit de criblage HIN) et 0,2 M Li<sub>2</sub>SO<sub>4</sub>, 0,1 M Bis-Tris, pH 5,5, 25 % w/v polyéthylène glycol 3350 (G2- kit de criblage HIN). Néanmoins, au terme des 3 premiers mois aucun cristal ne s'est développé. D'autres essais vont maintenant être menés autour des conditions ayant données lieu à une séparation de phase. En parallèle, des tests seront réalisés sur un échantillon protéique déglycosylé : les N-glycanes seront éliminés par digestion enzymatique à la PNGase F.



# MATERIELS ET METHODES

## 1. Expression de la protéine recombinante chez la levure

### 1.1. Principe

La protéine d'intérêt est exprimée chez la levure *Kluyveromyces lactis* (*K. lactis*) à l'aide du système « *K. lactis* Protein Expression Kit » commercialisé par New England Biolabs (NEB)<sup>126</sup>. Un des avantages de ce système est la sécrétion efficace des protéines dans le milieu de culture (Figure 63). Les étapes ultérieures de purification sont ainsi facilitées.

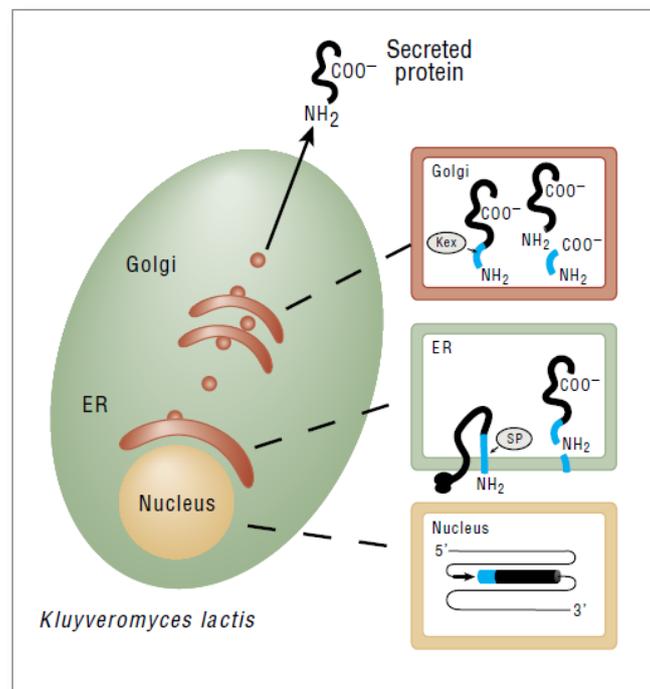


Figure 63 : Voie de sécrétion de la protéine recombinante chez la levure *K. lactis*<sup>127</sup>. Le gène d'intérêt (en noir) est cloné dans le vecteur pKLAC2 en aval de la séquence  $\alpha$ -MF (en bleu) codant pour le précurseur pré-pro du facteur  $\alpha$  (phéromone sexuelle) de *K. lactis*. Le précurseur contient un peptide signal (domaine pré) responsable de l'adressage de la protéine de fusion dans la lumière du réticulum endoplasmique (RE) (première phase de maturation). Le peptide signal est clivé au niveau du RE par une peptidase spécifique. La protéine de fusion est ensuite transportée par des vésicules vers l'appareil de Golgi (seconde phase de maturation) où l'endoprotéase Kex ( $\text{Lys-Arg} \downarrow$ ) clive le domaine pro du facteur  $\alpha$ . Une forme mature de la protéine est transportée par des vésicules de sécrétion vers la membrane plasmique et sécrétée dans le milieu de culture.

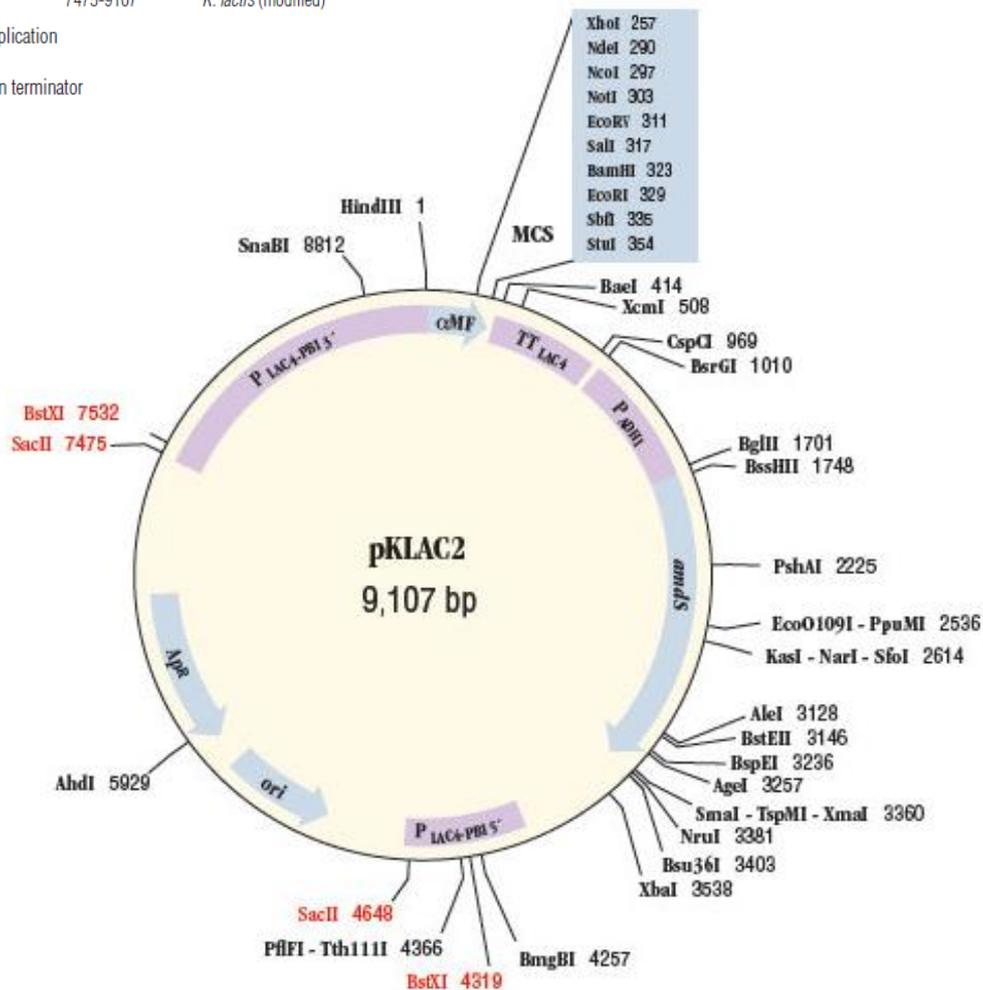
<sup>126</sup> Van Ooyen, A.J.J., Dekker, P., Huang, M., Olsthoorn, M.M.A, Jacobs, D.I., Colussi, P.A., and Taron, C.H. (2006). *FEMS Yeast Res.* **6**, 381-392.

<sup>127</sup> Manuel d'instructions « *K. lactis* Protein Expression Kit ». Version 1.1 (06/2011). New England Biolabs. <http://www.neb.com/nebecomm/manualfiles/manuale1000.pdf>.

Le gène d'intérêt est cloné dans le vecteur pKLAC2 (Figure 64) ; son expression est sous le contrôle d'un mutant du promoteur fort LAC4 (PLAC4-PBI) de *K. lactis*, inductible par le galactose<sup>128</sup>. pKLAC2 est un vecteur intégratif, capable de réplication autonome chez *Escherichia coli* (*E. coli*) ; il possède une origine de réplication bactérienne (pMB1) dérivée de pBR322 et un gène de résistance à l'ampicilline (Ap<sup>R</sup>).

Feature	Coordinates	Source
expression region:		
α-mating factor		
leader sequence	14-349	<i>K. lactis</i>
MCS	257-354	-
LAC4 TT region	371-953	<i>K. lactis</i>
AdH1 promoter region	1010-1712	<i>S. cerevisiae</i>
amdS	1713-3359	<i>A. nidulans</i>
LAC4 promoter		
region (5' end)	4068-4648	<i>K. lactis</i>
origin	5102-5690	pMB1
bla (Ap <sup>R</sup> )	6721-5861	Tn3
LAC4 promoter		
region (3' end)	7475-9107	<i>K. lactis</i> (modified)

ori = origin of replication  
 Ap = ampicillin  
 TT = transcription terminator



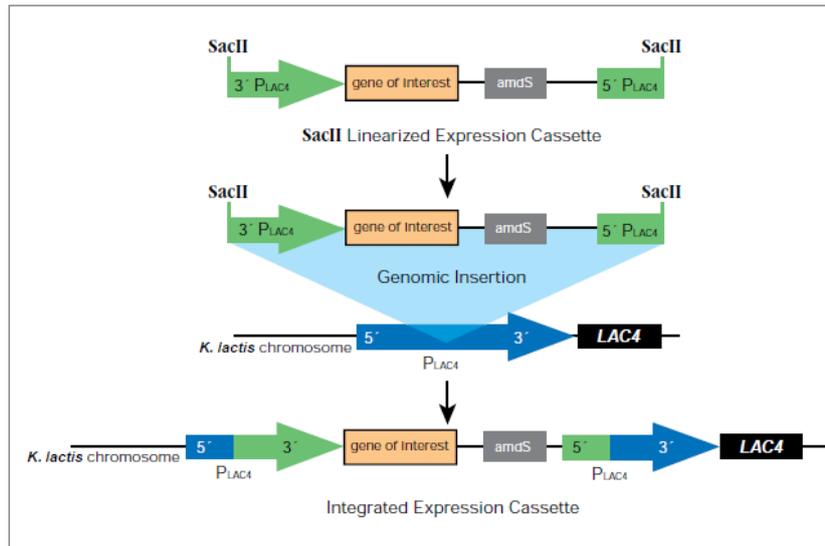
<sup>128</sup> Colussi, P. A., and Taron, C. H. (2005). *Appl. Environ. Microbiol.* **71** (11), 7092-7098.

9009	GAATTGTGAGCGGATAACAAGCTCAACACTTGAAATTTAGGAAAGAGCAGAATTTGGCAA	9068
	<b>HindIII</b>	
9069	AAAAAATAAAAAAAAAATAACACACATACTCATCGAGAAGCTTGAAAAAATGAAATTC	22
	M K F	
23	TCTACTATATTAGCCGCATCTACTGCTTAAATTTCCGTTGTTATGGCTGCTCCAGTTTCT	82
	S T I L A A S T A L I S V V M A A P V S	
83	ACCGAAACTGACATCGACGATCTTCCAATATCGGTTCCAGAAGAAAGCCTTGATTGGATTC	142
	T E T D I D D L P I S V P E E A L I G F	
143	ATTGACTTAACCGGGGATGAAGTTTCCTTGTTCCTGTTAATAACGGAAACCACTGGT	202
	I D L T G D E V S L L P V N N G T H T G	
	<b>XhoI</b>	
203	ATTCTATTCTTAAACACCACCATCGCTGAAGCTGCTTTCGCTGACAAGGATGATCTCGAG	262
	I L F L N T T I A E A A F A D K D D L E	
	<b>NdeI NcoI NotI EcoRV Sall</b>	
263	AAAAGAGAGGCTGAAGCTAGAAGAGCTCATATGTCCATGGCGGCCGCGATATCGTCGAC	322
	K R E A E A R R A H M S M G G R D I V D	
	<b>BamHI EcoRI SbfI StuI</b>	
323	GGATCCGAATTCCTGCAGGTAATTAATAAAGGCCTTGAATCGAGAATTTATACTTAGA	382
	G S E F P A G N *	
383	TAAGTATGTACTTACAGGTATATTTCTATGAGACTGATGTATACATGCATGATAATAT	442
443	TTAAACGGTTATTAGTGCCGATTGTCTTGTGCGATAATGACGTTCTATCAAAGCAATAC	502

**Figure 64:** Carte du vecteur d'expression *pKLAC2* et site de clonage multiple (MCS)<sup>126</sup>.

Cependant, pour pouvoir se répliquer chez *K. lactis*, le vecteur doit s'intégrer de façon stable dans le génome de la levure : l'intégration s'effectue par recombinaison homologue après linéarisation du plasmide recombinant (digestion enzymatique par SacII ou BstXI). Une ou plusieurs copies de la cassette d'expression s'intègrent dans le génome de *K. lactis* au niveau du locus LAC4 (Figure 65). Les levures transformées sont sélectionnées sur un milieu pauvre en azote contenant de l'acétamide via l'expression du gène de l'acétamidase (*amdS*) porté par le vecteur : seules les levures qui expriment le gène *amdS* peuvent décomposer l'acétamide en ammoniac et l'utiliser comme unique source d'azote pour se multiplier<sup>129</sup>.

<sup>129</sup> Read, J. D., Colussi, P. A., Ganatra, M. B., and Taron, C. H. (2007). *Appl. Environ. Microbiol.* **73** (16), 5088-5096.



*Figure 65 : Intégration de la cassette d'expression dans le génome de K. lactis au niveau du locus LAC4 par recombinaison homologue<sup>126</sup>.*

## 1.2. Clonage

### 1.2.1. Extraction des ARN totaux d'une culture cellulaire

Les cellules endothéliales humaines de veine ombilicale (HUVEC), qui expriment fortement la protéine h-APN à leur surface, sont mises en culture pour pouvoir en extraire les ARN totaux.

Les cellules HUVEC sont cultivées à l'Institut de Recherche en Hématologie et Transplantation (Hôpital du Hasenrain à Mulhouse). Elles sont mises en culture dans un milieu spécifique supplémenté permettant une croissance optimale (Endothelial Cells Growth Medium, Promocell). Les flasques de culture sont placés dans un incubateur à 37°C, 5% CO<sub>2</sub>. Lorsque les cellules sont à confluence, elles sont détachées (DetachKit, Promocell) à l'aide d'une solution de Trypsine/EDTA (0,04%/0,03%). Des culots cellulaires d'une densité de 10<sup>5</sup> à 10<sup>7</sup> cellules sont obtenus par centrifugation, rincés dans du tampon PBS 1x et conservés à -80°C.

Les ARN totaux sont extraits de cellules en culture à l'aide du kit RNAqueous®-4PCR (Ambion) en suivant les recommandations du fournisseur. L'extraction débute par une lyse cellulaire (tampon Lysis/Binding) en présence d'un agent chaotropique, le thiocyanate de guanidine (inactive et précipite les protéines dont les ribonucléases (RNases)). Les acides nucléiques présents dans le lysat cellulaire sont précipités avec une solution d'éthanol à 64%

et fixés spécifiquement sur une colonne de silice (en présence de hautes concentrations en sels chaotropiques). Des lavages successives (tampons Wash Solution 1 et Wash Solution 2/3) permettent d'éliminer les contaminants cellulaires. Pour finir, les ARN totaux sont élués avec 60 µl d'eau sans nucléase.

Afin de minimiser la contamination par l'ADN génomique, un traitement à la DNase I est préconisé par le fournisseur : la désoxyribonucléase I (DNase I) hydrolyse les ADN simple et/ou double brin. 2U de DNase I (2U/µl) sont suffisantes pour dégrader l'ADN génomique résiduel présent dans l'échantillon d'ARN totaux. La réaction enzymatique se déroule à 37°C pendant 30 minutes dans la solution tampon 10 mM Tris pH 7,5, 2,5 mM MgCl<sub>2</sub> et 0,5 mM CaCl<sub>2</sub>. L'enzyme est inactivée par ajout d'un inhibiteur de DNase I (DNase Inactivation Reagent).

*La quantité et la qualité* des acides nucléiques sont déterminées par spectrophotométrie à l'aide du NanoDrop ND-1000 (Thermo Scientific). La concentration en ARN est déterminée par mesure de l'absorbance à 260 nm (les bases azotées absorbent fortement à cette longueur d'onde); une unité d'absorbance à 260 nm correspond à 40 µg d'ARN/ml. Le rapport de densité optique entre 260 et 280 nm (maximum d'absorbance des acides aminés) permet de témoigner de la pureté de l'échantillon : l'ARN est considéré comme pur si le rapport A260/A280 est compris entre 1,8 et 2,0 ; un rapport inférieur à 1,8 atteste d'une contamination protéique.

*L'intégrité* des ARN totaux est vérifiée par électrophorèse sur gel d'agarose dénaturant. Le gel d'agarose est préparé comme suit : 1 g d'agarose (Type D-5 DNA Grade, Euromedex) est dissout dans 72 ml d'eau sans nucléase ; après dissolution complète de l'agarose, 10 ml de tampon MOPS 10x (concentration finale 40 mM MOPS, pH 7,0, 10 mM acétate de sodium et 1 mM EDTA) et 18 ml de 37% formaldéhyde (concentration finale 6,6 %) sont ajoutés. L'échantillon d'ARN est mélangé à du tampon de charge (Formaldehyde Load Dye, Ambion) à raison de 0,5 volume de tampon de charge pour 1 µg d'ARN. Du bromure d'éthidium (BET) à une concentration finale de 10 µg/ml est additionné (solution mère de BET à 10 mg/ml, Euromedex). L'échantillon est dénaturé par chauffage à 65-70°C pendant 10 minutes puis, déposé dans le puits du gel. La migration est réalisée dans du tampon MOPS 1x sous voltage constant 5 V/cm. Le profil électrophorétique est révélé sous UV grâce au bromure d'éthidium, qui est un agent intercalant des acides nucléiques. La présence de deux bandes distinctes, correspondant aux ARN ribosomiaux 28S et 18S, dont le rapport des intensités 28S/18S est d'environ 2, atteste de l'intégrité de l'ARN.

Des aliquotes d'ARN totaux sont conservés à -80°C. Pour éviter leur dégradation par les ribonucléases, un inhibiteur de ribonucléases (RNasin Plus RNase Inhibiteur, Promega) est ajouté (40 U d'inhibiteur par µg d'ARN total).

### *1.2.2. Transcription inverse de l'ARN en ADN complémentaire (ADNc)*

La synthèse du premier brin d'ADN complémentaire (ADNc) est catalysée par la transcriptase inverse M-MLV (H-) commercialisée par Promega. Il s'agit d'une version génétiquement modifiée de la transcriptase inverse du virus de la leucémie murine de Moloney (Moloney Murine Leukemia Virus Reverse Transcriptase, M-MLV RT) : son activité ribonucléase H (RNase H) a été supprimée et sa stabilité thermique renforcée. Ces mutations favorisent la synthèse d'ADNc de pleine longueur (>5 kb). Une amorce oligo-dT (synthétisée par Eurofins mwg), qui s'hybride sur la queue polyadénylate des ARN messager (ARNm), est utilisée pour initier la synthèse du premier brin d'ADNc.

1 µg d'ARN total est incubé en présence de 0,5 µg d'oligo-dT pendant 5 minutes à 70°C. La synthèse est réalisée dans du tampon M-MLV RT 1x (50 mM Tris-HCl, pH 8,3, 75 mM KCl, 3 mM MgCl<sub>2</sub> et 10 mM DTT) en présence de 40 U d'inhibiteur de ribonucléases (RNasin Plus RNase Inhibiteur, Promega), 0,5 mM de dNTP (Promega) et 100 U de transcriptase inverse M-MLV (H-) (200 U/µl). Le mélange est incubé 10 minutes à 42°C, puis 50 minutes à 55°C. Pour finir, la transcriptase inverse est inactivée par chauffage à 70°C pendant 15 minutes. Le premier brin d'ADNc synthétisé est conservé à -20°C.

### *1.2.3. Amplification du gène d'intérêt (entier ou tronqué) par PCR*

Le fragment d'ADN (ANPEP ou ANPEP<sub>194-2904</sub>) qui code pour la protéine h-APN ou sa forme tronquée h-APN<sub>65-967</sub> est amplifié par PCR (réaction de polymérisation en chaîne) à l'aide d'amorces spécifiques (Tableau 17). Le premier brin d'ADNc, synthétisé précédemment, sert de matrice à la réaction. Les sites de restriction XhoI (5'...**CTCGAG**...3') et BglII (5'...**AGATCT**...3') sont introduits à chaque extrémité du fragment amplifié pour le clonage dans le vecteur pKLAC2. Le site de clivage spécifique de la protéase Kex (5'...**AAAAG<sup>A</sup>**...3') est inséré dans l'amorce sens : clivage du peptide signal (domaine pro du facteur α) au niveau de l'appareil de Golgi. Afin de faciliter les étapes ultérieures de purification (par chromatographie d'affinité sur résine de nickel), une queue polyhistidine (5'...**CATCATCATCATCATCAT**...3') est introduite en position N-terminale de la protéine. Elle

sera ensuite éliminée par digestion avec l'entérokinase (site de clivage spécifique 5'...GACGATGACGATAAG<sup>v</sup>...3').

Construction	Description	Nom des amorces	Séquence (5'->3')
ANPEP	Amplification du gène codant pour la protéine membranaire h-APN (2904 pb).	Sens : APNS1H Antisens : h-APN_END	TATCCTCTCGAGAAAAGACATATGCATCATCATCATCATCAT GACGATGACGATAAGATGGCCAAGGGCTTCTATATTC TAGCATAGATCTGGACTATTTGCTGTTTTCTGTGAACCAC
ANPEP <sub>194-2904</sub>	Amplification du gène codant pour une forme soluble de la protéine (résidus 65 à 967) (2711 pb).	Sens : APNS2H Antisens : h-APN_END	TATCCTCTCGAGAAAAGACATATGCATCATCATCATCATCAT GACGATGACGATAAGATGGACCAAAGTAAAGCGTGGAA TAGCATAGATCTGGACTATTTGCTGTTTTCTGTGAACCAC

*Tableau 17*: Liste des amorces utilisées pour l'amplification du gène entier (ANPEP) et sa forme tronquée (ANPEP<sub>194-2904</sub>). Les amorces sont synthétisées par la société Eurofins mwg. La zone soulignée correspond à la partie de l'oligonucléotide qui s'hybride spécifiquement au gène. Le site de restriction NdeI (5'...CA<sup>v</sup>TATG...3') est introduit pour cloner, si nécessaire, le gène d'intérêt dans le vecteur d'expression bactérien pET-22b(+). Le site de restriction BglII (5'...A<sup>v</sup>GATCT...3') est, quant à lui, compatible avec le site BamHI (5'...G<sup>v</sup>GATCC...3') des vecteurs d'expression pKLAC2 et pET-22b(+).

Les réactions de PCR sont effectuées dans un volume final de 50 µl contenant du tampon réactionnel (50 mM Tris-HCl, pH 9, 1,5 mM MgCl<sub>2</sub>, 15 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> et 0,1% Triton X-100), 200 µM de dNTP, 0,3 µM de chaque amorce, 2 µl d'ADNc 1<sup>er</sup> brin et 1U de Taq polymérase (DyNAzyme<sup>TM</sup> EXT DNA Polymerase, Finnzymes).

Le thermocycleur MJ Mini<sup>TM</sup> Personal Thermal Cycler (Bio-Rad) est utilisé pour amplifier le gène. Les conditions de PCR consistent en une étape initiale de dénaturation à 94 °C pendant 2 minutes suivie de 30 cycles d'amplification comprenant chacun une étape de dénaturation à 94°C pendant 30 secondes, une étape d'hybridation à 59°C pendant 20 secondes, puis une étape d'élongation à 72 °C pendant 2 minutes et 30 secondes, et enfin une étape d'extension finale à 72 C pendant 10 minutes.

#### 1.2.4. Electrophorèse sur gel d'agarose

Les produits de l'amplification du gène sont analysés par électrophorèse sur gel d'agarose 1% dans du tampon TBE 0,5% (45 mM Tris, 45 mM acide borique, 1 mM EDTA pH 8,0). Le gel d'agarose est préparé dans le même tampon à raison de 1 g d'agarose pour 100 ml de tampon TBE 0,5%, auquel est ajouté du bromure d'éthidium (BET) à une concentration finale de 0,5 µg/ml (solution mère de BET à 10 mg/ml, Euromedex).

L'échantillon d'ADN est mélangé à du tampon de charge 6x (20 mM Tris-HCl pH 8,0, 15% Ficoll 400, 66 mM EDTA, 0,09% bleu de bromophénol, 0,10% de SDS) à raison de 2 µl de tampon de charge pour 10 µl d'ADN et déposé dans les puits du gel. La migration est réalisée sous un voltage constant de 110 V. Le profil électrophorétique est révélé sous UV grâce au bromure d'éthidium. La migration du fragment d'ADN est comparée à un marqueur de taille moléculaire 0,1-10 kb (2-log DNA ladder, NEB).

#### *1.2.5. Purification du produit de PCR (extraction sur gel d'agarose)*

Le fragment d'ADN d'intérêt (entier ou tronqué) est extrait du gel d'agarose et purifié à l'aide d'un kit d'extraction et de purification (QIAquick Gel Extraction Kit, QIAGEN) selon les instructions du fournisseur. Le principe de cette méthode repose sur la propriété que possède la silice de se lier sélectivement à l'ADN en présence de hautes concentrations en sels chaotropiques à pH 7,5.

La bande d'agarose contenant l'ADN est découpée à l'aide d'une lame de scalpel, puis dissout par chauffage dans du tampon QG contenant du thiocyanate de guanidine (sel chaotrope) (3 volumes de tampon QG pour 100 mg de gel d'agarose). L'ADN est ensuite fixé sur une colonne en silice (colonne QIAquick spin) par ultrafiltration (centrifugation à 17 900 g pendant 1 minute à température ambiante). La membrane est rincée avec 500 µl de tampon QG (pour éliminer les excédents d'agarose) et 750 µl de tampon PE (tampon à base d'éthanol). Pour finir, le fragment d'ADN est élué avec 30 ou 50 µl de tampon EB (10 mM Tris-HCl pH 8,5).

#### *1.2.6. Quantification de l'acide désoxyribonucléique*

La concentration en ADN est évaluée par mesure de l'absorbance à 260 nm à l'aide du NanoDrop ND-1000. Une unité d'absorbance à 260 nm correspond à une concentration de 50 µg/ml d'ADN double brin pur. Le fragment d'ADN est considéré comme pur si le rapport A260/A280 est compris entre 1,8 et 2,0. Un rapport inférieur à 1,8 atteste d'une contamination protéique et un rapport supérieur à 2,0 a pour origine une contamination de l'échantillon par de l'ARN.

### *1.2.7. Digestion enzymatique du produit de PCR*

Le produit de PCR (gène ANPEP entier ou tronqué) est digéré simultanément par les enzymes de restriction XhoI et BglII (NEB). Ces endonucléases génèrent des bouts cohésifs à chaque extrémité du fragment amplifié.

La double digestion enzymatique est réalisée dans un volume réactionnel final de 100 µl contenant du tampon NEBuffer 3 (100 mM NaCl, 50 mM Tris-HCl, 10 mM MgCl<sub>2</sub>, 1 mM Dithiothreitol, pH 7,9), 100 µg/ml de BSA (Albumine de sérum bovin), 1 µg du fragment d'ADN d'intérêt et 50 unités de chaque enzyme. Le milieu réactionnel est incubé à 37°C pendant 2 heures. Les enzymes de restriction sont inactivées par un traitement de 20 minutes à 65 °C.

### *1.2.8. Purification du fragment d'ADN digéré*

Le fragment d'ADN digéré est purifié à l'aide d'un kit de purification (QIAquick PCR Purification Kit, QIAGEN) selon les instructions du fournisseur.

L'ADN à purifier est dilué dans du tampon PB contenant du chlorhydrate de guanidine (agent chaotrope) (5 volumes de tampon PB pour 1 volume de produit de PCR), puis fixé sur une colonne en silice (colonne QIAquick spin) par ultrafiltration (centrifugation à 17 900 g pendant 1 min à température ambiante). Les contaminants cellulaires sont éliminés en rinçant la membrane avec 750 µl de tampon PE (tampon à base d'éthanol). Pour finir, l'ADN est élué avec 30 µl de tampon EB (10 mM Tris-HCl pH 8,5).

L'ADN est quantifié par mesure de l'absorbance à 260 nm à l'aide du Nanodrop ND-1000.

### *1.2.9. Digestion enzymatique du vecteur pKLAC2*

Le vecteur d'expression pKLAC2 est digéré simultanément par les enzymes de restriction XhoI et BamHI-HF (NEB). L'enzyme BamHI libère une extrémité cohésive compatible avec celle produite par BglII.

La double digestion enzymatique est réalisée dans un volume réactionnel final de 200 µl contenant du tampon NEBuffer 4 (50 mM acétate de potassium, 20 mM Tris-acétate, 10 mM acétate de magnésium et 1 mM Dithiothreitol, pH 7,9), 100 µg/ml de BSA, 2 µg de vecteur pKLAC2 et 100 unités de chaque enzyme. Le milieu réactionnel est incubé à 37°C pendant 2 heures. Les enzymes de restriction sont inactivées par un traitement de 20 minutes à 65 °C.

En parallèle, l'activité enzymatique de chaque enzyme de restriction est contrôlée séparément et comparée à un contrôle négatif. Les contrôles sont effectués dans un volume réactionnel final de 20 µl contenant du tampon NEBuffer 4, 100 µg/ml de BSA, de l'eau stérile, 0,2 µg de vecteur pKLAC2 et 10 U des enzymes XhoI (contrôle 1) et BamHI-HF (contrôle 2). Les enzymes de restriction sont inactivées en chauffant à 65°C pendant 20 minutes.

Les produits de digestion enzymatique du vecteur sont analysés par électrophorèse sur gel d'agarose 1% dans du tampon TBE 0,5%.

#### *1.2.10. Déphosphorylation du vecteur pKLAC2*

Afin d'empêcher la recircularisation du vecteur lors de la ligation par l'ADN ligase du phage T4, le vecteur est déphosphorylé à l'aide de la phosphatase alcaline de l'intestin de veau (CIP) (NEB). L'enzyme enlève par hydrolyse les groupements phosphates en position 5'. 20 unités de CIP sont directement ajoutées au mélange réactionnel de 200 µl contenant 2 µg de vecteur pKLAC2 digéré par les enzymes XhoI et BamHI-HF dans du tampon NEBuffer 4 (50 mM acétate de potassium, 20 mM Tris-acétate, 10 mM acétate de magnésium et 1 mM Dithiothreitol, pH 7,9). Le milieu réactionnel est incubé 1 heure à 37 °C.

L'ADN est purifié à l'aide du kit de purification (QIAquick PCR Purification Kit, QIAGEN) comme décrit précédemment et dosé par mesure de l'absorbance à 260 nm au Nanodrop ND-1000.

#### *1.2.11. Ligation cohésive par l'ADN ligase du phage T4*

La ligation entre les extrémités cohésives de l'insert (gène ANPEP entier ou tronqué) et du vecteur est effectuée à l'aide de l'ADN ligase du phage T4 (NEB). L'enzyme catalyse la formation d'une liaison phosphodiester entre les extrémités 5' phosphate et 3' hydroxyle des deux brins d'ADN. L'ATP est utilisée comme cofacteur.

La réaction est réalisée à 16°C pendant 1 nuit dans un volume réactionnel final de 20 µl dans du tampon 50 mM Tris-HCl, 10 mM MgCl<sub>2</sub>, 1 mM ATP, 10 mM Dithiothreitol, pH 7,5. 100 ng de vecteur sont additionnés à différentes quantités d'insert (ratio molaire vecteur : insert 1 :1 et 1 :3) en présence de 400 U d'ADN ligase du phage T4.

En parallèle, la réaction est réalisée en absence de l'insert. Ce contrôle permet d'évaluer l'efficacité de la déphosphorylation du vecteur par la phosphatase alcaline de l'intestin de veau.

#### *1.2.12. Transformation des souches bactériennes XL10 Gold par choc thermique*

Les cellules ultra-compétentes d'*Escherichia coli* XL10 Gold (Stratagène) ont été utilisées pour amplifier le plasmide recombinant.

Les bactéries compétentes (100 µl) sont décongelées sur glace, et 4 µl de β-Mercaptoethanol (β-ME) y sont additionnés car il a été démontré que l'efficacité de la transformation augmente en présence de β-ME. Le mélange est incubé sur glace pendant 10 minutes puis réparti délicatement en aliquotes de 10 µl.

1 µl des différentes réactions de ligation est ajouté à des aliquotes de bactéries compétentes et les mélanges sont incubés 30 minutes sur glace. La transformation est réalisée par choc thermique à 42°C pendant 40 secondes. Les tubes sont ensuite placés sur glace pendant 2 minutes. Les bactéries sont resuspendues dans 100 µl de milieu de culture SOC (2% tryptone, 0,5% d'extract de levure, 10 mM NaCl, 2,5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub> et 20 mM glucose), préchauffé à 42°C et incubés 1 heure à 37°C. Pour finir, les suspensions bactériennes sont étalées sur une gélose sélective de LB (Luria Broth ; 1% de tryptone, 0,5% d'extract de levure et 0,5% de NaCl) contenant de l'ampicilline à une concentration de 100 µg/µl. Les boîtes de Pétri sont incubées une nuit à 37°C.

Seules les bactéries transformées pourront se multiplier sur le milieu sélectif car le plasmide porte un gène de résistance à l'ampicilline.

#### *1.2.13. Criblage des bactéries transformées par PCR*

Les colonies bactériennes qui se sont multipliées sur le milieu sélectif peuvent contenir le plasmide recombinant (contenant le gène d'intérêt) ou le vecteur refermé sur lui-même : il est donc important de pouvoir distinguer ces deux populations.

La sélection des bactéries ayant incorporé le plasmide recombinant est effectuée par PCR à l'aide d'amorces spécifiques au vecteur : amorce sens (S1275) 5'...GAAGAAGCCTTGATTGGA...3' et amorce antisens (S1276) 5'... TTATCGCACAAGACAATC...3' (Figure 66). Un fragment de 340 pb + la taille du gène d'intérêt (entier ou tronqué) est amplifié lorsque l'insert est introduit dans le vecteur.

9069	AAAAATAAAAAAAAAAATAAACACACATACTCATCGAGAAGCTTGAAAAAATGAAATTC	22
	HindIII	
	M K F	
23	TCTACTATATTAGCCGCATCTACTGCTTTAATTCGGTTGTTATGGCTGCCAGTTTCT	82
	S T I L A A S T A L I S V V M A A P V S	
83	ACCGAAACTGACATCGACGATCTTCCAATATCGGTTCCA	142
	S1275 →	
	GAAGAAGCCTTGATTGGA TTC	
	T E T D I D D L P I S V P E E A L I G F	
143	ATTGACTTAACCGGGGATGAAGTTTCTTGTGCCTGTTAATAACGGAACCCACACTGGT	202
	I D L T G D E V S L L P V N N G T H T G	
203	ATTCTATTCTTAAACACCACCATCGCTGAAGCTGCTTTCGCTGACAAGGATGATCTCGAG	262
	XhoI	
	I L F L N T T I A E A A F A D K D D L E	
263	AAAAGAGAGGCTGAAGCTAGAAGAGTCATATGTCATGGGCGGCCGATATCGTCGAC	322
	NdeI NcoI NotI EcoRV Sall	
	K R E A E A R R A H M S M G G R D I V D	
323	GGATCCGAATTCCTCGAGGTAATTAATAAAGGCCTTGAATCGAGAATTTACTTAGA	382
	BamHI EcoRI SbfI StuI	
	G S E F P A G N *	
383	TAAGTATGTACTTACAGGTATATTTCTATGAGATACTGATGTATACATGCATGATAATAT	442
443	TAAACGGTTATTAGTGC	502
	← S1276	
	GATTGTCTTGCGATAA TGACGTTCTATCAAAGCAATAC	

Figure 66 : Criblage des bactéries transformées par PCR à l'aide d'amorces spécifiques au vecteur (en rouge)<sup>126</sup>.

Les réactions de PCR sont effectuées dans un volume final de 25 µl contenant du tampon réactionnel (10 mM Tris-HCl, pH 8,3, 50 mM KCl), 2 mM MgCl<sub>2</sub>, 200 µM de dNTPs, 0,5 µM de chaque amorce et 0,5 U de Taq polymérase (NEB). Les colonies bactériennes sont directement inoculées dans le milieu réactionnel de PCR. Durant les premières étapes de dénaturation, les cellules sont lysées et leur ADN est relâché dans le milieu réactionnel, où il peut alors servir de matrice.

Le thermocycleur MJ Mini™ Personal Thermal Cycler (Bio-Rad) est utilisé pour amplifier le gène. Les conditions de PCR consistent en une étape initiale de dénaturation à 94°C pendant 5 minutes suivie de 30 cycles d'amplification comprenant chacun une étape de dénaturation à 94°C pendant 30 secondes, une étape d'hybridation à 55°C pendant 30 secondes, puis une étape d'élongation à 72°C pendant 1 minute et 30 secondes, et enfin une étape d'extension finale à 72°C pendant 7 minutes.

Les produits de PCR sont séparés par migration sur gel d'agarose 1% dans du tampon TBE 0,5X.

#### *1.2.14. Extraction et purification de l'ADN plasmidique produit par les bactéries : midiprep*

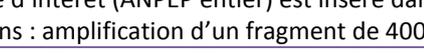
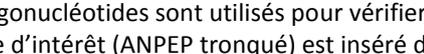
Des clones positifs sont inoculés dans 2 ml de milieu LB contenant 100 µg/ml d'ampicilline et incubés pendant environ 8 heures à 37 °C sous agitation. Ces pré-cultures sont utilisées pour ensemercer (dilution au 1/500e) 25 ml de milieu LB additionné de l'antibiotique. Après une nuit d'incubation à 37°C sous agitation, les bactéries sont sédimentées par centrifugation (6000 g pendant 15 minutes).

Les ADN plasmidiques sont extraits et purifiés à l'aide d'un kit (QIAGEN® Plasmid Purification Midi Kit, QIAGEN) selon les instructions du fournisseur. Les culots bactériens sont resuspendus dans 4 ml de tampon P1 (50 mM Tris-HCl, pH 8,0, 10 mM EDTA) contenant 100 µg/ml de RNase A. Les bactéries sont ensuite lysées par un traitement alcalin (4 ml de tampon P2 (200 mM NaOH, 1% SDS (w/v)) pendant 5 minutes à température ambiante. L'agent de lyse est neutralisé par l'ajout de 4 ml de tampon P3 (3 M acétate de potassium, pH 5,5). Les suspensions bactériennes sont incubées pendant 15 minutes dans un bain de glace. Les débris cellulaires sont séparés par centrifugation (20 000 g pendant 30 minutes à 4 °C). Les surnageants qui contiennent les ADN plasmidiques sont déposés sur une résine échangeuse d'anions en présence de sels en faible concentration. Les contaminants (ARN, protéines et autres petites impuretés) sont éliminés par une série de lavage avec du tampon QC (1 M NaCl, 50 mM MOPS, pH 7,0, 15% isopropanol (v/v)). Pour finir, les ADN plasmidiques sont élués avec 5 ml de tampon salin QF (1,25 M NaCl, 50 mM Tris-HCl, pH 8,5, 15% isopropanol (v/v)) avant d'être précipités à l'isopropanol. Les culots d'ADN plasmidiques sont sédimentés par centrifugation (20 000 g pendant 30 minutes à 4 °C), rincés à l'éthanol 70% et resuspendus dans 100 µl de tampon EB (10 mM Tris-HCl, pH 8,5).

L'ADN plasmidique est quantifié par mesure de l'absorbance à 260 nm au NanoDrop ND-1000.

#### *1.2.15. Contrôle de l'intégrité de l'ADN plasmidique par PCR*

Une réaction de PCR permet de vérifier l'intégrité du plasmide recombinant extrait et purifié par midiprep. Les couples d'amorces ci-dessous sont utilisés (Tableau 18) :

Couple d'amorces	Séquence (5'→3')	Description
Sens : APN1 Antisens : APN_END	ATGGCCAAGGGCTTCTATATTC CTATTTGCTGTTTTCTGTGAACCAC	 <p>Un fragment de 2904 pb est amplifié si le gène d'intérêt (entier) a bien intégré le vecteur.</p>
Sens : APN2 Antisens : APN_END	TGGACCAAAGTAAAGCGTGAA CTATTTGCTGTTTTCTGTGAACCAC	 <p>Un fragment de 2711 pb est amplifié si le gène d'intérêt (tronqué) a bien intégré le vecteur.</p>
Sens : S1275 Antisens : APN_STARTrev	GAAGAAGCCTTGATTGGA TGATTGTGCACACGGCTG	 <p>Ces oligonucléotides sont utilisés pour vérifier que le gène d'intérêt (ANPEP entier) est inséré dans le bon sens : amplification d'un fragment de 400 pb.</p>
Sens : S1275 Antisens : CS32	GAAGAAGCCTTGATTGGA GGTCATCTGCCAACTCCCCT	 <p>Ces oligonucléotides sont utilisés pour vérifier que le gène d'intérêt (ANPEP tronqué) est inséré dans le bon sens : amplification d'un fragment de 562 pb.</p>
Sens : APN_ENDs Antisens : S1276	AACAAGGAGGTGGTGCTC TTATCGCACAAGACAATC	 <p>Ces oligonucléotides sont utilisés pour vérifier que le gène d'intérêt (entier ou tronqué) est inséré dans le bon sens : amplification d'un fragment de 205 pb.</p>

**Tableau 18 :** Liste des couples d'amorces utilisés pour vérifier l'intégrité du plasmide recombinant par PCR. Les amorces sont synthétisées par la société Eurofins mwg

Les réactions de PCR sont effectuées comme décrit précédemment pour le criblage des bactéries transformées. L'ADN plasmidique sert de matrice pour l'amplification des fragments d'ADN.

Les produits de PCR sont analysés sur gel d'agarose 1% dans du tampon TBE 0,5x.

### 1.2.16. Séquençage de l'ADN

Le séquençage de l'ADN est réalisé par la société Microsynth AG en utilisant la méthode de Sanger. Les résultats du séquençage sont analysés à l'aide du logiciel ChromasPro<sup>113</sup> et comparés à des bases de données (NCBI Blast)<sup>114</sup>.

### 1.3. Intégration de la cassette d'expression dans le génome de *K. lactis*

#### 1.3.1. Linéarisation du plasmide recombinant

La digestion du plasmide recombinant par l'enzyme de restriction *SacII* (5'...CCGCGG...3') génère un fragment d'ADN linéaire (cassette d'expression) qui est capable de s'intégrer dans le génome de *K. lactis*.

La digestion enzymatique est réalisée dans un volume réactionnel final de 50 µl. 3 µg de plasmide recombinant sont mélangés avec 20 unités de l'enzyme dans du tampon NEBuffer 4 (50 mM acétate de potassium, 20 mM Tris-acétate, 10 mM acétate de magnésium et 1 mM Dithiothreitol, pH 7,9). Le milieu réactionnel est incubé à 37°C pendant 3 heures. L'enzyme de restriction est inactivée par un traitement de 20 minutes à 65 °C.

En parallèle, le plasmide contrôle pK<sub>LAC1</sub>-malE, fourni avec le kit, est linéarisé dans les mêmes conditions expérimentales.

Les produits de digestion enzymatique sont analysés par électrophorèse sur gel d'agarose 1% dans du tampon TBE 0,5x.

La digestion des plasmides recombinants libère deux fragments : la cassette d'expression et un fragment d'ADN linéaire de 2827 pb. Il n'est pas nécessaire de séparer ces deux fragments car seule la cassette d'expression peut s'intégrer dans le génome de la levure. Les produits de digestion sont purifiés à l'aide du kit QIAquick PCR Purification (QIAGEN).

La concentration en ADN est évaluée par mesure de l'absorbance à 260 nm à l'aide du NanoDrop ND-1000.

#### 1.3.2. Transformation des cellules compétentes GG799 de *K. lactis*

La cassette d'expression est intégrée par recombinaison homologue dans le génome de *K. lactis* au niveau du locus LAC4. La transposition s'effectue après transformation par choc thermique dans les cellules compétentes GG799 de *K. lactis* (NEB).

Les levures *K. lactis* GG799 sont des cellules haploïdes de type sexuel α. Isolée des produits laitiers dans les années 80 (groupe Gist Brocades- DSM), cette souche a initialement été utilisée par les industries alimentaires comme hôte pour la production de protéines

hétérologues. Elle est appréciée pour ses capacités à se développer à haute densité cellulaire et à sécréter efficacement les protéines recombinantes.

Les cellules compétentes sont décongelées sur glace. L'efficacité de la transformation est augmentée en additionnant du réactif pour transformation (620 µl de NEB Yeast Transformation Reagent par aliquote de cellules). 1 µg du plasmide recombinant linéaire est ajouté à un aliquote de cellules compétentes et le mélange est incubé à 30°C durant 30 minutes. La transformation est réalisée par choc thermique à 37°C pendant 1 minute. Les cellules sont ensuite sédimentées par centrifugation (7000 rpm pendant 2 minutes) et resuspendues dans 1 ml de milieu YPGlu (1% d'extrait de levure, 2% de peptone et 2% de glucose). La suspension cellulaire est incubée à 30°C durant 4 heures sous agitation. Les cellules sont récupérées par centrifugation puis, remises en suspension dans du tampon PBS 1x. Différentes dilutions de la suspension cellulaire sont étalées sur une gélose pauvre en azote YCB (NEB) contenant de l'acétamide (5mM). Les boîtes de pétri sont incubées pendant 3-4 jours à 30°C.

En parallèle, le plasmide contrôle pKLAC1-malE est transformé dans les cellules compétentes GG799 dans les mêmes conditions expérimentales.

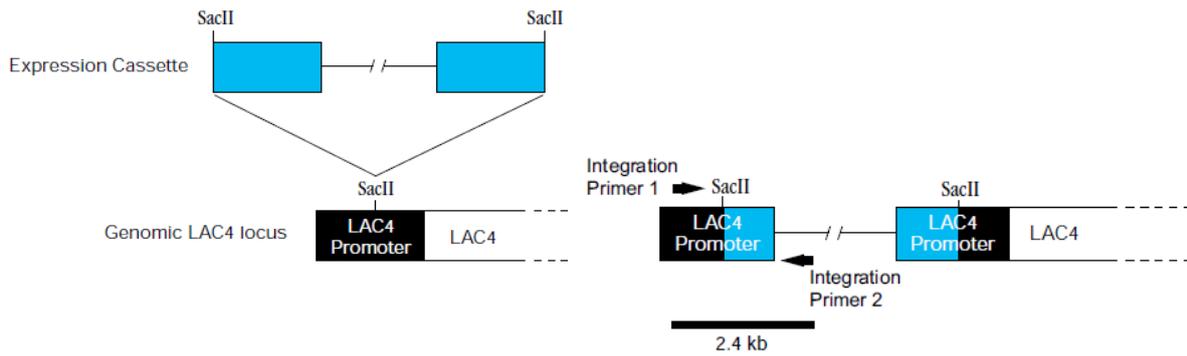
Seules les levures, qui expriment le gène *amdS* (porté par le vecteur), peuvent décomposer l'acétamide en ammoniac et l'utiliser comme unique source d'azote pour se multiplier sur la gélose.

### 1.3.3. Analyse des levures transformées par PCR

Les levures transformées, qui ont intégré la cassette d'expression dans leur génome, sont identifiées par PCR en utilisant les amorces fournies avec le kit :

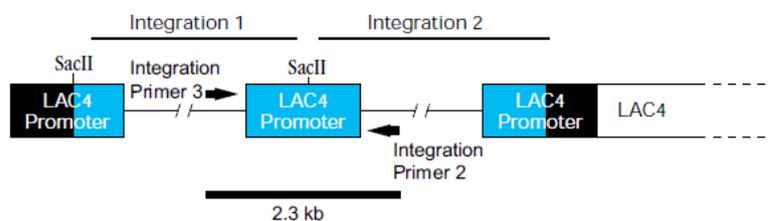
- *Integration primer 1* : 5'...ACACACGTAAACGCGCTCGGT...3'
- *Integration primer 2* : 5'...ATCATCCTTGTCAGCGAAAGC...3'

Un fragment de 2400 pb est amplifié si la cassette d'expression est correctement intégrée dans le génome de *K. lactis*, au niveau du locus LAC4 ([Figure 67](#)).



*Figure 67* : Analyse des levures transformées par PCR ; identification des levures qui ont intégré la cassette d'expression dans leur génome au niveau du locus LAC4<sup>126</sup>.

Lors de la transformation, jusqu'à dix copies de la cassette d'expression peuvent se transposer dans le génome de la levure. Les colonies, qui ont intégré plusieurs copies dans leur génome, produisent plus de protéines. Ces dernières sont identifiées lors d'une seconde réaction de PCR en utilisant le couple d'amorces *Integration primer 2* et *3* (5'...ACCTGAAGATAGAGCTTCTAA...3'). Un fragment de 2300 pb est amplifié lorsqu'au minimum deux copies de la cassette d'expression sont introduites dans le génome de la levure (Figure 68).



*Figure 68* : Analyse des levures transformées par PCR ; identification des levures qui ont intégré plusieurs copies de la cassette d'expression dans leur génome au niveau du locus LAC4<sup>126</sup>.

Les colonies sont inoculées dans 25 µl d'une solution de sorbitol (1M) contenant 2 mg/ml de lyticase (Sigma Aldrich). Les suspensions cellulaires sont incubées à 30°C pendant 1 heure. Une première étape de dénaturation (98°C durant 10 minutes) permet de lyser totalement les cellules ; l'ADN ainsi libéré dans le milieu réactionnel sert de de matrice pour la réaction de PCR.

La réaction de PCR est initiée par l'ajout de 10 µl de chaque amorce (10x), 200 µM de dNTP et 5 U de Taq polymérase (NEB) dans du tampon réactionnel (10 mM Tris-HCl, pH 8,3, 50 mM KCl, 1,5 mM MgCl<sub>2</sub>).

Les 30 cycles d'amplification comprennent chacun une étape de dénaturation à 94°C pendant 30 secondes, une étape d'hybridation à 50°C pendant 30 secondes et une étape d'élongation à 72°C pendant 2 minutes. Pour finir, une étape d'extension finale à 72°C est réalisée pendant 7 minutes.

Les produits d'amplification sont déposés sur gel d'agarose 1% dans du tampon TBE 0,5 x.

#### 1.4. Tests d'expression de la protéine recombinante

Les levures GG799 sont mises en culture à 30°C dans un milieu liquide riche, YPGal, composé de 1% d'extrait de levure, 2% de peptone et 2% de galactose. La transcription du gène est activée suite à l'induction du promoteur LAC4-PBI par le galactose. La croissance des levures est suivie par mesure de la densité optique à 600 nm ( $DO_{600nm}$ ).

Des colonies isolées sont inoculées dans 1 à 10 ml de milieu YPGal et incubées sous agitation (200 rpm) toute la nuit à 30°C. Ces précultures servent à ensemercer 100 à 1000 ml de milieu YPGal (dilution au 1/100<sup>e</sup>). Les cultures sont alors incubées (2-3 jours) à 30°C sous agitation (200 rpm) jusqu'à atteindre une  $DO_{600nm}$  supérieure à 30 (phase stationnaire). Les levures sont ensuite sédimentées par centrifugation à 6000 rpm pendant 10 minutes ; les culots cellulaires et les surnageants de culture sont analysés pour évaluer le niveau d'expression et localiser la protéine recombinante.

#### 1.5. Analyse de l'expression de la protéine recombinante

##### *1.5.1. Electrophorèse sur gel de polyacrylamide (SDS-PAGE)*

Le niveau d'expression de la protéine recombinante est analysé sur gel de polyacrylamide (PAGE) en conditions dénaturantes. Cette technique permet de séparer les protéines en fonction de leur poids moléculaire.

Les échantillons protéiques (culots cellulaires et surnageants de culture) sont mélangés à du tampon de charge 4x (NuPAGE® LDS Sample Buffer 4x, Invitrogen), qui contient du glycérol, des colorants et du lithium dodecyl sulfate (LDS) à pH 8,4. Les protéines sont dénaturées par chauffage à 100°C pendant 5 minutes puis, séparées sur un gel de polyacrylamide en gradient 3-8% (NuPAGE® 3-8% Tris-Acetate gel, Invitrogen). L'électrophorèse est conduite à un voltage constant (150 Volts) durant 1 heure dans du

tampon de migration Tris-Acétate SDS (0,1%), pH 8,4 (NuPAGE® Tris-Acetate SDS Running Buffer, Invitrogen). Une fois la migration terminée, les protéines sont révélées par une coloration au bleu de Coomassie. Le gel est incubé pendant 15 minutes dans du milieu de coloration (SimplyBlue™ SafeStain, Invitrogen). L'excès de colorant est ensuite éliminé par des lavages successifs à l'eau distillée. Le marqueur de poids moléculaire (Novex Sharp 3,5-260 KDa, Invitrogen) est employé comme référence.

### *1.5.2. Transfert de protéines et immunodétection (Western Blot)*

Cette technique est utilisée pour détecter spécifiquement la protéine recombinante. Les protéines séparées par SDS-PAGE sont transférées électriquement sur une membrane en PVDF (polyfluorure de vinylidène) de 0,2 µm (BioRad) pendant 1h15 sous un voltage constant de 100 V dans du tampon de transfert (25 mM Tris, 192 mM glycine, 20% méthanol, pH 8,3). Les sites non spécifiques sont saturés en incubant la membrane dans du tampon PBS- Tween 0,2% contenant 5% (m/v) de lait écrémé, pendant 1 heure à température ambiante. La membrane est alors incubée en présence de l'anticorps primaire anti-CD13 H-300 (anticorps polyclonal de lapin, Santa Cruz Biotechnology) dilué au 1/1000<sup>e</sup> dans du tampon PBS- Tween 0,2% contenant 5% (m/v) de lait écrémé durant 1 nuit à 4°C. Les interactions non spécifiques sont éliminées par une série de lavages dans du tampon PBS-Tween 0,05%. La membrane en PVDF est finalement incubée 1h à température ambiante avec un anticorps secondaire (Cell Signaling Technology) dilué au 1/10000<sup>e</sup> dans du tampon PBS- Tween 0,05% contenant 5% (m/v) de lait écrémé. Ce dernier cible spécifiquement les immunoglobulines (IgG) de lapin et est couplé à une enzyme responsable de la détection de la protéine cible, la HRP (peroxydase de raifort). La membrane est à nouveau rincée dans du tampon PBS-Tween 0,05%. La protéine est révélée par chimiluminescence (Luminata Forte Western HRP substrate, Millipore).

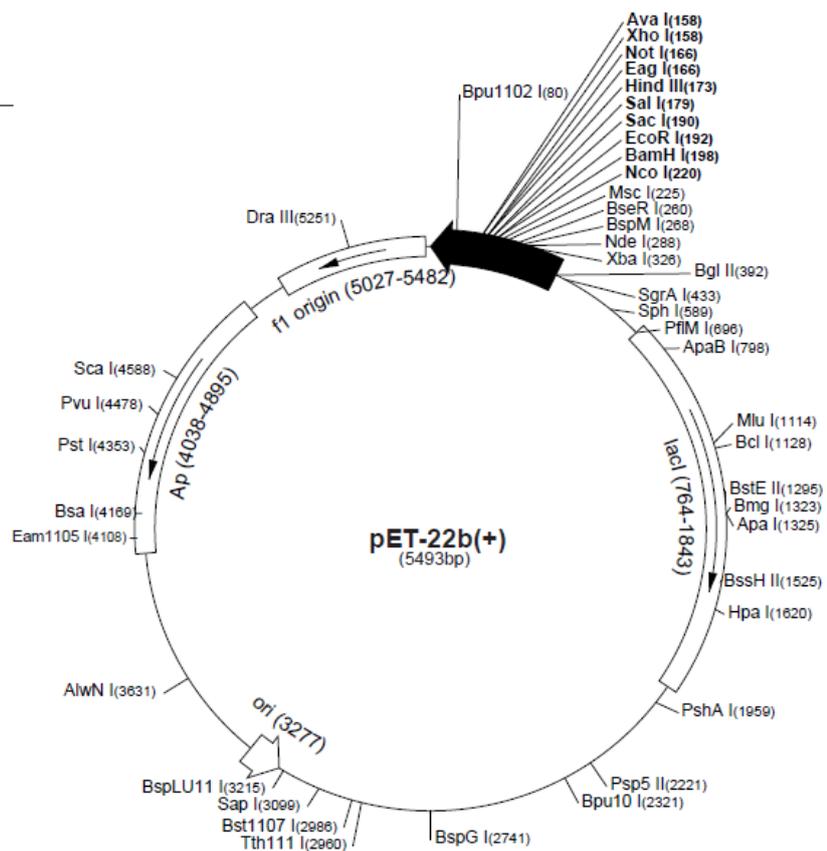
## 2. Expression de la protéine recombinante chez *Escherichia coli*

### 2.1. Principe

La protéine d'intérêt est exprimée chez *Escherichia coli* à l'aide du système pET commercialisé par Novagen®. Le gène d'intérêt est cloné dans le vecteur d'expression pET-22b(+) (Figure 69) ; ce dernier possède une origine de réplication bactérienne pBR322 et porte un gène de résistance à l'ampicilline (ApR). La séquence signal *peIB*, en amont du gène d'intérêt, offre la possibilité d'exporter la protéine recombinante dans le périplasma de la bactérie, dont l'environnement oxydant est favorable à la formation de ponts disulfures.

#### pET-22b(+) sequence landmarks

T7 promoter	361-377
T7 transcription start	360
<i>peIB</i> coding sequence	224-289
Multiple cloning sites ( <i>Nco</i> I - <i>Xho</i> I)	158-225
His•Tag coding sequence	140-157
T7 terminator	26-72
<i>lacI</i> coding sequence	764-1843
pBR322 origin	3277
<i>bla</i> coding sequence	4038-4895
f1 origin	5027-5482



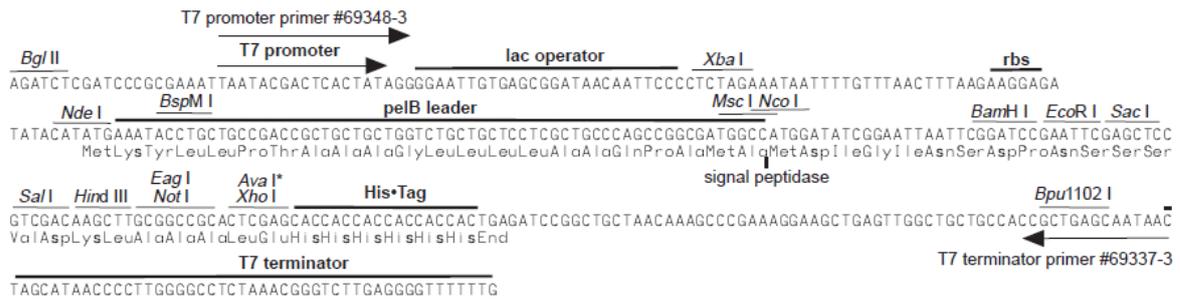


Figure 70 : Carte du vecteur d'expression pET-22b(+)<sup>130</sup> et site de clonage multiple (MCS)<sup>130</sup>.

L'expression du gène d'intérêt est sous le contrôle du promoteur fort T7. Le plasmide recombinant est introduit dans une souche bactérienne infectée par le bactériophage lambda DE3 : ces bactéries contiennent dans leur génome un gène codant pour l'ARN polymérase du bactériophage T7 dépendant du promoteur lac UV5 inducible par l'IPTG (isopropyl-β-D-thiogalactoside) (Figure 71). En absence de lactose (IPTG), le gène régulateur *lacI* code le répresseur *lac* qui va se lier spécifiquement au niveau du promoteur *lac* de l'opéron lactose empêchant ainsi la fixation de l'ARN polymérase au site d'initiation de la transcription. En revanche, en présence de lactose, le répresseur subit un changement conformationnel et perd alors son affinité pour le promoteur *lac*. Le gène de l'ARN polymérase T7 est transcrit puis, l'ARNm est traduit en enzyme ; les molécules d'ARN polymérase T7 produites se fixent alors sur le promoteur T7 et initient la transcription du gène d'intérêt.

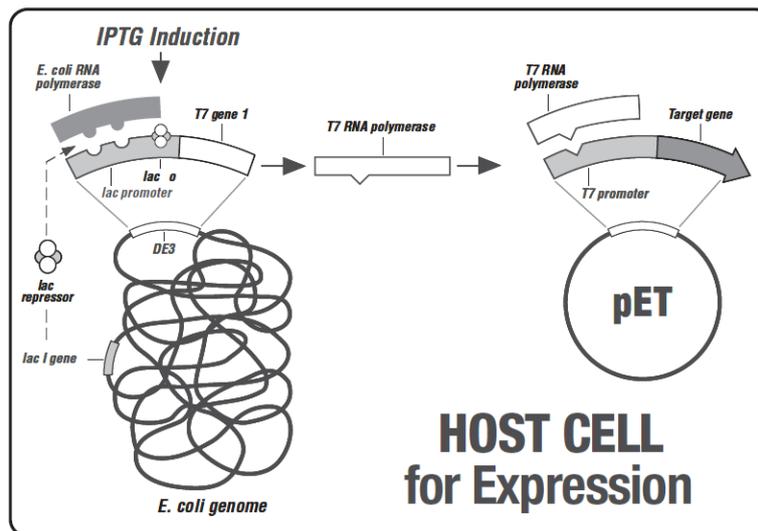


Figure 71 : Système d'expression pET (Novagen); induction de l'expression du gène d'intérêt par l'IPTG<sup>131</sup>.

<sup>130</sup> Carte du vecteur pET-22b(+). <http://wolfson.huji.ac.il/expression/commercial-vectors/pet-22b-map.pdf>.

## 2.2. Clonage

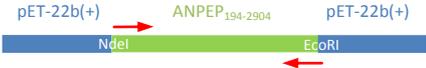
Le plasmide recombinant pKLAC2- ANPEP<sub>194-2904</sub> est digéré simultanément par NdeI (5'...CA<sup>T</sup>TATG...3') et StuI (5'...AGG<sup>C</sup>CT) : 5µg de plasmide recombinant sont incubés durant 2 heure à 37 °C en présence de 100 U de chaque enzyme dans du tampon NEBuffer 4. Les produits de digestion sont déposés sur gel d'agarose 1% dans du tampon TBE 0,5x. Le fragment d'ADN codant pour la forme tronqué de hAPN (résidus 65 à 967) est extrait du gel d'agarose et purifié à l'aide d'un kit (QIAquick Gel Extraction kit, QIAGEN).

Le vecteur pET-22b(+) est dans un premier temps digéré par EcoRI-HF (5'...GA<sup>A</sup>ATTC...3') (10 µg de vecteur sont mélangés à 40 U d'enzyme dans du tampon NEBuffer 4 ; le mélange réactionnel est incubé durant 2 heure à 37°C). Un traitement à la nucléase Mung Bean (NEB) permet ensuite de générer des bouts francs (20 U d'enzyme sont suffisantes pour hydrolyser les extrémités simple-brins ; la digestion enzymatique se déroule à 30°C pendant 1 heure). Pour finir, le vecteur est digéré par NdeI (le vecteur linéarisé est incubé durant 2 heure à 37 °C en présence de 20 U de l'enzyme dans du tampon NEBuffer 4). Afin d'éviter la recircularisation du vecteur lors de la ligation par l'ADN ligase T4, le vecteur est déphosphorylé à l'aide de la phosphatase alcaline de l'intestin de veau (10 U de CIP sont suffisantes pour hydrolyser (en 1 heure à 37°C) les groupements phosphate en position 5').

La ligation entre le gène d'intérêt (ANPEP tronqué) et le vecteur linéarisé est réalisée avec l'ADN ligase du phage T4 à 16°C pendant 1 nuit. Le produit de ligation est ensuite transformé par choc thermique dans les cellules compétentes XL10 GOLD. Les bactéries ayant intégré le plasmide recombinant sont identifiées par PCR à l'aide d'amorces spécifiques au vecteur (Tableau 19). L'ADN plasmidique produit par les bactéries est ensuite extrait et purifié à l'aide d'un kit (QIAGEN® Plasmid Purification Midi Kit, QIAGEN). Pour finir, l'intégrité des plasmides recombinants est contrôlée par PCR avec les couples d'amorces décrits ci-dessous (Tableau 19).

---

<sup>131</sup> Manuel d'instructions « pET system ». 11<sup>ème</sup> Edition (01/2007). Novagen.  
[http://kirschner.med.harvard.edu/files/protocols/Novagen\\_petsystem.pdf](http://kirschner.med.harvard.edu/files/protocols/Novagen_petsystem.pdf).

Couple d'amorces	Séquence (5'→3')	Description
Sens : T7 promoter Antisens : T7 terminator	TAATACGACTCACTATAGGG GCTAGTTATTGCTCAGCGG	 <p>Un fragment de 2997 pb est amplifié si le gène d'intérêt (tronqué) a bien intégré le vecteur.</p>
Sens : APN2 Antisens : APN_END	TGGACCAAAGTAAAGCGTGGAA CTATTTGCTGTTTTCTGTGAACCAC	 <p>Un fragment de 2711 pb est amplifié si le gène d'intérêt (tronqué) a bien intégré le vecteur.</p>
Sens : T7 promoter Antisens : CS32	TAATACGACTCACTATAGGG GGTCATCTGCCAACTCCCCT	 <p>Ces oligonucléotides sont utilisés pour vérifier que le gène d'intérêt (ANPEP entier) est inséré dans le bon sens (amplification d'un fragment de 500 pb).</p>
Sens : CS25 Antisens : T7 terminator	GCCTGGGAGCAGTTCGAAAT GCTAGTTATTGCTCAGCGG	 <p>Ces oligonucléotides sont utilisés pour vérifier que le gène d'intérêt (ANPEP tronqué) est inséré dans le bon sens (amplification d'un fragment de 632 pb).</p>

**Tableau 19 :** Liste des couples d'amorces utilisés pour sélectionner les bactéries qui ont intégré le gène d'intérêt (ANPEP tronqué). Les amorces sont synthétisées par la société Eurofins mwg.

## 2.3. Culture des bactéries et expression de la protéine recombinante

### 2.3.1. Souches bactériennes

Les souches d'expression *E. coli* BL21 (DE3), Rosetta 2 (DE3) et Rosetta-gami 2 (DE3) sont utilisées pour surexprimer la protéine ; ces bactéries sont infectées par le bactériophage lambda DE3.

BL21 (DE3) [ $F^- ompT hsdSB(r_B^- m_B^-) gal dcm$  (DE3)] est la souche de référence pour l'expression de protéines recombinantes. Elle est déficiente en protéases *ompT* (outer membrane protease) et *lon*, ce qui permet une meilleure stabilité des protéines exogènes.

La souche Rosetta 2 (DE3) [ $F^- ompT hsdSB(r_B^- m_B^-) gal dcm$  (DE3) pRARE2<sup>3</sup> (Cam<sup>R</sup>)], dérivée des BL21 (DE3), porte un plasmide pRARE2 (résistant au chloramphénicol), qui permet la synthèse des ARN de transfert (ARNt) correspondant aux codons suivants : AGG, AGA, AUA, CUA, CCC, GGA et CGG. L'expression de protéines eucaryotes contenant ces codons rarement utilisés par *E. coli* est ainsi améliorée.

La souche Rosetta-gami 2 (DE3) [ $\Delta$ (*ara-leu*)7697  $\Delta$ *lacX74*  $\Delta$ *phoA* *PvuII* *phoR* *araD139* *ahpC* *galE* *galk* *rpsL*(DE3)  $F'$ [*lac+ lacI<sup>q</sup> pro*] *gor522::Tn10* *trxB* pRARE2<sup>3</sup> (Cam<sup>R</sup>, Str<sup>R</sup>, Tet<sup>R</sup>)<sup>4</sup>] possède les mêmes caractéristiques que les Rosetta 2 (DE3). Elle contient également des mutations au niveau des gènes *trxB* et *gor*, qui

codent respectivement pour des thiorédoxines réductases et des glutathiones réductases, favorisant ainsi la formation de ponts disulfures dans le cytoplasme des bactéries. Cette souche est résistante au chloramphénicol, à la streptomycine et à la tétracycline.

### *2.3.2. Transformation par choc thermique dans les souches d'expression*

Le protocole utilisé pour la transformation des cellules compétentes BL21 (DE3), Rosetta 2 (DE3) et Rosetta-gami 2 (DE3) est celui préconisé par Novagen®. 200 µl de cellules compétentes sont décongelées sur glace puis répartis délicatement en aliquotes de 20 µl. Les bactéries sont mélangées avec 1 µl de plasmide recombinant (1-10 ng/µl) puis incubées 5 minutes sur glace. La transformation est réalisée par choc thermique à 42 °C pendant 30 secondes. Le tube est ensuite placé sur glace pendant 2 minutes. Les bactéries sont resuspendues dans 80 µl de milieu SOC (préchauffé à 42°C) et incubées 1 heure à 37 °C. Pour finir, les suspensions bactériennes sont étalées sur des géloses LB (Luria Broth) additionnées d'ampicilline (50 µg/ml) (BL21 (DE3)), d'ampicilline et de chloramphénicol (34 µg/ml) (Rosetta 2 (DE3)) et d'ampicilline, de chloramphénicol, de streptomycine (50 µg/ml) et de tétracycline (12,5 µg/ml) (Rosetta-gami 2 (DE3)). Les boîtes de Pétri sont incubées une nuit à 37°C.

Les bactéries ayant intégré le plasmide recombinant sont identifiées par PCR comme décrit précédemment.

### *2.3.3. Optimisation des conditions d'expression de la protéine recombinante*

Les bactéries sont cultivées en milieu liquide Luria Broth (LB) en présence des antibiotiques appropriés. Ce milieu nutritif riche est composé de 1% de tryptone, 0,5% d'extrait de levure et 0,5% de chlorure de sodium. La croissance bactérienne est suivie par mesure de la densité optique à 600 nm ( $DO_{600nm}$ ).

Les conditions optimales d'expression de la protéine recombinante sont identifiées en faisant varier un grand nombre de paramètres : choix des souches bactériennes, de la concentration en IPTG, de la température et du temps d'induction, de la  $DO_{600nm}$  au moment de l'induction.

Le protocole d'expression reste similaire quelle que soit la souche utilisée. Les bactéries ayant intégré le plasmide recombinant sont inoculées dans 2 ml de milieu liquide LB contenant les antibiotiques appropriés. Les cultures sont incubées sous agitation (200 rpm) toute la nuit à 37°C. Ces précultures servent à ensemer 50 à 200 ml de milieu LB supplémenté d'antibiotiques. En fonction des essais, 100 µM de chlorure de zinc ( $ZnCl_2$ ) y

sont additionnés. Les cultures sont incubées à 37°C sous agitation. Lorsque la  $DO_{600nm}$  atteint 0,5-0,8, la température est abaissée (entre 20 et 30°C) et l'expression de la protéine est induite par l'ajout d'IPTG à une concentration comprise entre 0,4 et 1 mM (suivant les essais). Le niveau d'expression de la protéine recombinante est comparé pour différents temps d'induction (de quelques heures à toute une nuit). Au terme de l'expression, les cultures sont centrifugées à 8000 g pendant 10 minutes ; les bactéries sont lysées par sonication, puis les protéines solubles sont séparées des protéines insolubles par centrifugation (17 000 g pendant 30 minutes à 4°C. Le niveau d'expression et la proportion de protéine soluble sont analysés sur gel SDS-PAGE comme décrit précédemment.

#### *2.3.4. Production à grande échelle*

Les bactéries Rosetta 2 (DE3) transformées (conservées à -80°C dans une solution de glycérol 20%) sont inoculées dans du milieu LB supplémentée d'ampicilline et de chloramphénicol. Les précultures sont incubées sous agitation (200 rpm) toute la nuit à 37°C. Elles sont utilisées pour ensemercer 4 l de milieu LB contenant les antibiotiques appropriés et 100  $\mu$ M de chlorure de zinc ( $ZnCl_2$ ). Les cultures sont réalisées dans un fermenteur 5L (Biostat® Aplus, Sartorius). Le volume de préculture à ajouter est calculé de façon à avoir une densité optique à 600 nm de l'ordre de 0,1. Les cultures sont incubées à 30°C sous agitation (200 rpm) jusqu'à une  $DO_{600nm}$  comprise entre 0,5 et 0,8. L'expression de la protéine recombinante est induite par l'addition de 1 mM d'IPTG. Après 4 heures d'induction à 30°C, les bactéries sont sédimentées par centrifugation (8000 xg pendant 10 minutes) et conservées à -20°C.

### 3. Expression de la protéine recombinante dans baculovirus

Ces travaux ont été réalisés lors d'un séjour de 3 mois au Centre de Recherche Botnar de l'université d'Oxford au sein de l'équipe du Prof. Udo Oppermann.

#### 3.1. Principe

##### 3.1.1. Système d'expression Bac-to-Bac®

Le système d'expression « Bac to Bac® Baculovirus Expression System » de Invitrogen est utilisé pour générer des baculovirus recombinants exprimant la protéine d'intérêt. Cette technique est basée sur la transposition site-spécifique d'une cassette d'expression contenant le gène d'intérêt dans un bacmide qui peut se répliquer dans *E. coli* et qui contient l'intégralité du génome de baculovirus AcMNPV (virus de la polyédrose nucléaire multiple d'*Autographa californica*)<sup>132</sup>.

Le gène d'intérêt est cloné dans un vecteur donneur (pFastBac ou un dérivé) entre deux sites de transposition Tn7 (Tn7R et Tn7L). Son expression est sous le contrôle du promoteur fort de la polyhédrine. Le vecteur recombinant, qui possède un marqueur de sélection à la gentamicine, est transformé dans des bactéries compétentes *E. coli* DH10Bac. Ces bactéries hébergent le génome du baculovirus AcMNPV sous la forme d'un bacmide qui porte le gène de résistance à la kanamycine et auquel a été ajouté le gène lacZ entre deux sites de transposition Tn7. Cette souche héberge également un plasmide « helper » qui code pour une transposable et qui contient le gène de résistance à la tétracycline. La transposable permet la transposition spécifique du gène d'intérêt dans le bacmide au niveau du site mini-attTn7, situé dans le gène lacZ. La transfection des cellules d'insectes Sf9 avec l'ADN du bacmide recombinant permet de produire les baculovirus recombinants qui serviront à l'expression de la protéine d'intérêt.

La [Figure 72](#) récapitule les différentes étapes à réaliser pour générer les baculovirus recombinants et exprimer la protéine d'intérêt à l'aide du système d'expression « Bac to Bac® ».

---

<sup>132</sup> Luckow, V.A., Lee, C. S., Barry, G. F., and Olins, P. O. (1993). *J. Virol.* **67**, 4566-4579.

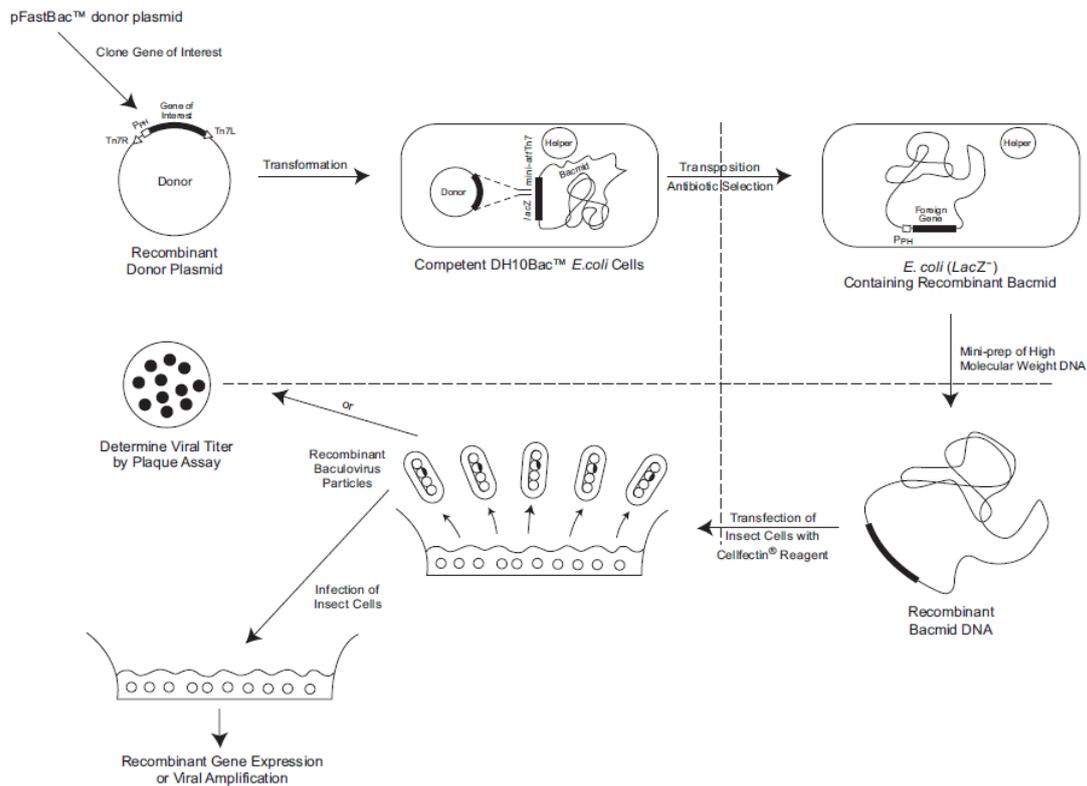


Figure 72 : Schéma récapitulatif des différentes étapes à réaliser pour produire les baculovirus recombinants et exprimer la protéine d'intérêt à l'aide du système « Bac to Bac® Baculovirus Expression System » de Invitrogen<sup>133</sup>.

### 3.1.2. Clonage sans ligase

Pour résoudre rapidement la structure tridimensionnelle d'un grand nombre de protéines à visées thérapeutiques, le Consortium de Génomique Structurale d'Oxford utilise une technologie de clonage directionnel qui permet de s'affranchir des étapes de digestions enzymatiques et des réactions de ligation. Il s'agit d'une procédure rapide et efficace connue sous le nom anglais de « Ligation Independent Cloning »<sup>134</sup> basée sur les propriétés de l'ADN polymérase du phage T4 (Figure 73).

Le vecteur d'expression possède comme sites de clonage des séquences d'oligonucléotides (12 à 16 pb) dans lesquelles une des bases azotées est absente. Le vecteur est digéré par une enzyme de restriction présente aux niveaux de ces sites, puis est traité par l'ADN polymérase du phage T4 en présence de la base manquante. L'activité 3'->5' exonucléase de l'ADN polymérase T4 permet de générer des extrémités 5' cohésives. L'insert

<sup>133</sup> Manuel d'instructions « Bac-to-Bac® Baculovirus Expression System ». Version D (04/2004). Invitrogen. [http://kirschner.med.harvard.edu/files/protocols/Invitrogen\\_bactobacexpression.pdf](http://kirschner.med.harvard.edu/files/protocols/Invitrogen_bactobacexpression.pdf).

<sup>134</sup> Aslanidis, C., and de Jong, P. J. (1990). *Nucleic Acids Res.* **18**, 6069-6074.

est quant à lui amplifié par PCR à l'aide d'amorces spécifiques au gène contenant des séquences homologues aux sites de clonage du vecteur. Le traitement du produit PCR par l'ADN polymérase T4 permet de produire des extrémités 5' cohésives complémentaires à celles du vecteur. L'hybridation des extrémités cohésives de l'insert et du vecteur permet, après transformation dans des cellules compétentes de *E coli*, de générer le vecteur recombinant.

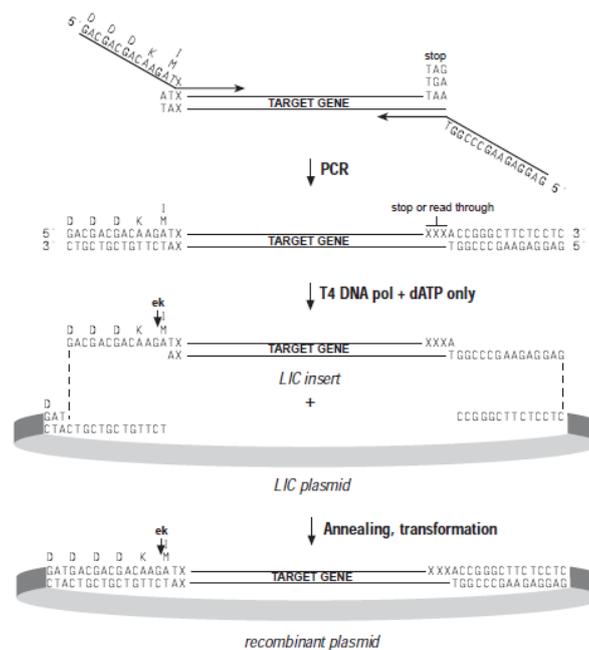


Figure 73 : Stratégie de clonage directionnel sans enzymes de restriction ou ligases (« Ligation Independent Cloning »)<sup>135</sup>

### 3.1.3. Plasmides

Les vecteurs d'expression baculovirus pFB-Lic-Bse (Figure 74) et pFB-CT10HF-LIC (Figure 75) dérivent du vecteur pFastBac commercialisé par Invitrogen. Ils ont été construits respectivement par Opher Gileadi et Grazyna Kochan du Consortium de Génomique Structurale d'Oxford. Ces vecteurs possèdent les extensions LIC 5' et LIC 3' nécessaires pour réaliser un clonage directionnel sans les étapes préalables de digestions enzymatiques et de ligation. Le gène d'intérêt est cloné entre les éléments de transposition Tn7 et son expression est sous le contrôle du promoteur fort de la polyhédrine. Afin d'identifier les bactéries transformées dans *E. coli*, les vecteurs possèdent un marqueur de sélection positif à l'ampicilline et à la gentamicine ainsi qu'un marqueur de sélection négatif à SacB. Pour faciliter les étapes de purification, les vecteurs ont été enrichis avec une queue polyhistidine

<sup>135</sup> Novy, R. E., Yaeger, K. W., and Kolb, K. M. (1996). *InNovations*. 5, 8-10.



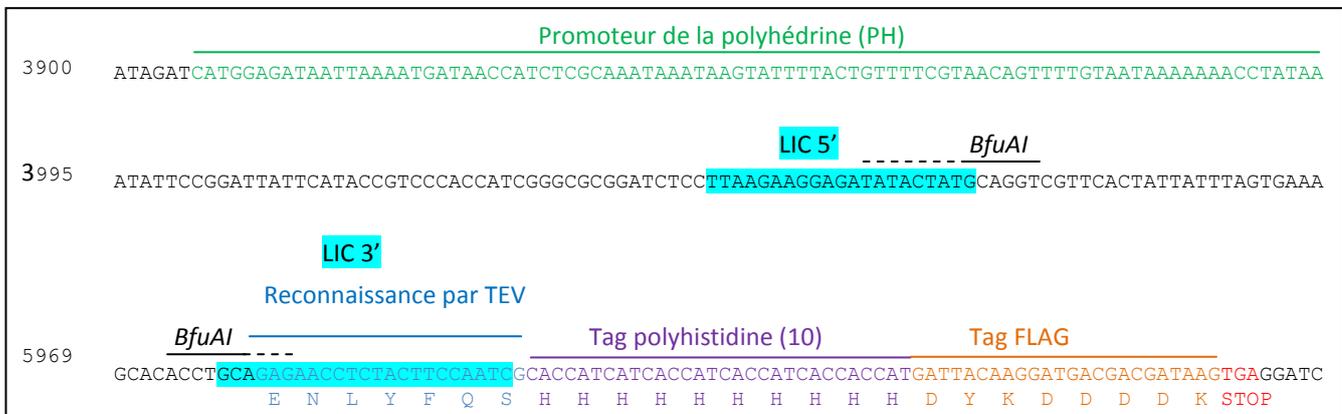
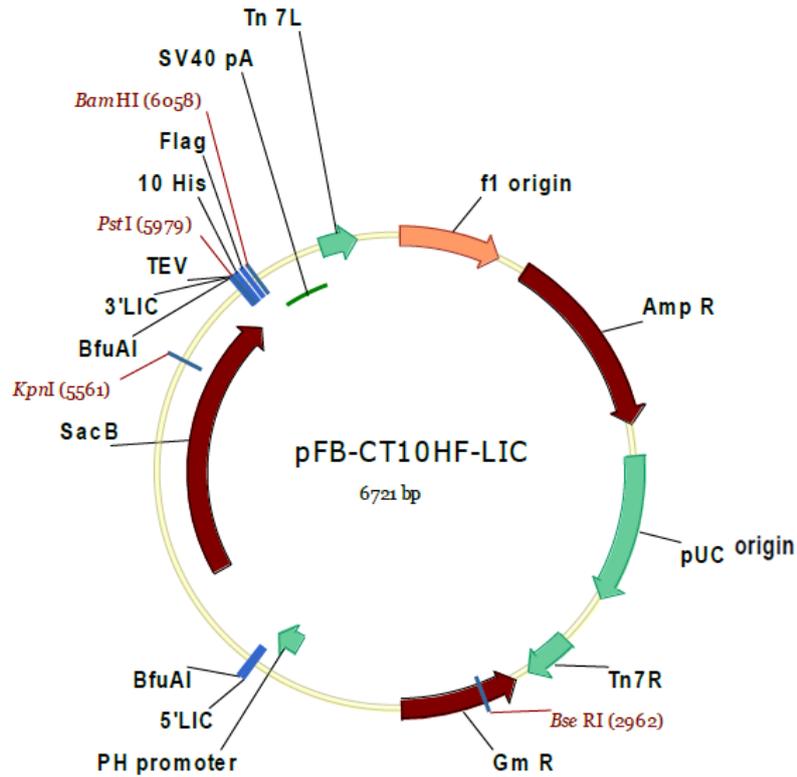


Figure 75 : Carte du vecteur d'expression pFB-CT10HF-LIC<sup>137</sup>.

<sup>137</sup> Carte du vecteur pFB-CT10HF-LIC.

<http://apps.thesgc.org/structures/MM/Vectors/pFB-CT10HF-LIC/pFB-CT10HF-LIC%20Vector%20sheet.pdf>.

## 3.2. Clonage

### 3.2.1. Amplification de l'insert par PCR

Le gène ANPEP (*homo sapiens* alanyl (membrane) aminopeptidase) et ses formes tronquées sont amplifiés par PCR à partir d'une banque d'ADNc (collection MGC/IMAGE ID 6020902).

Les séquences des amorces utilisées sont représentées dans le Tableau 20.

Construction	Description	Nom des amorces	Sequence 5'→3'
ANPEP-a001	Amplification du gène codant pour h-APN. Clonage dans le vecteur pFB-LIC-Bse.	Sens : ANPEP-f001 Antisens : ANPEP-r001	TACTCCCAATCCATG <u>GCCAAGGGCTTCTATATTCC</u> TATCCACCTTTACTGTCACTACTATTTGCTGTTTTCTGTGAACC
ANPEP-a002	Amplification du gène codant pour h-APN <sub>59-967</sub> . Clonage dans le vecteur pFB-LIC-Bse.	Sens : ANPEP-f003 Antisens : ANPEP-r001	TACTCCCAATCCATG <u>CCCGCCTCGGCCACC</u> TATCCACCTTTACTGTCACTACTATTTGCTGTTTTCTGTGAACC
ANPEP-a003	Amplification du gène codant pour h-APN <sub>70-967</sub> . Clonage dans le vecteur pFB-LIC-Bse.	Sens : ANPEP-f005 Antisens : ANPEP-r001	TACTCCCAATCCATG <u>GCGTGGAATCGTTACC</u> TATCCACCTTTACTGTCACTACTATTTGCTGTTTTCTGTGAACC
ANPEP-a011	Amplification du gène codant pour h-APN. Clonage dans le vecteur pFB-CT10HF-LIC.	Sens : ANPEP-f002 Antisens : ANPEP-r002	TTAAGAAGGAGATATACTATG <u>GCCAAGGGCTTCTATATTCC</u> GATTGGAAGTAGAGGTTCTCTGCCTATTTGCTGTTTTCTGTGAACC
ANPEP-a012	Amplification du gène codant pour h-APN <sub>59-967</sub> . Clonage dans le vecteur pFB-CT10HF-LIC.	Sens : ANPEP-f004 Antisens : ANPEP-r002	TTAAGAAGGAGATATACTATG <u>CCCGCCTCGGCCACC</u> GATTGGAAGTAGAGGTTCTCTGCCTATTTGCTGTTTTCTGTGAACC
ANPEP-a013	Amplification du gène codant pour h-APN <sub>70-967</sub> . Clonage dans le vecteur pFB-CT10HF-LIC.	Sens : ANPEP-f006 Antisens : ANPEP-r002	TTAAGAAGGAGATATACTATG <u>GCGTGGAATCGTTACC</u> GATTGGAAGTAGAGGTTCTCTGCCTATTTGCTGTTTTCTGTGAACC

Tableau 20 : Liste des amorces utilisées pour l'amplification du gène ANPEP et ses formes tronquées. Les amorces sont synthétisées par la société Eurofins mwg. La zone soulignée correspond à la partie de l'oligonucléotide qui s'hybride spécifiquement au gène. Les amorces contiennent des séquences homologues aux sites de clonage des vecteurs pFB-LIC-Bse (partie de l'oligonucléotide colorée en bleu) et pFB-CT10HF-LIC (partie de l'oligonucléotide colorée en vert). Ces extensions (LIC 5' et LIC 3') sont nécessaires pour réaliser un clonage directionnel sans les enzymes de restriction et les ligases. Le codon stop est représenté en rouge. Pour les constructions dans le vecteur pFB-CT10HF-LIC, le codon stop est retiré afin de permettre l'expression des queues polyhistidine et FLAG.

Les réactions de PCR sont effectuées dans un volume final de 25 µl contenant du tampon réactionnel (50 mM Tris-HCl, pH 8,0, 50 mM KCl, 1 mM DTT, 0,1 mM EDTA et 50% (v/v) glycérol), 1 mM de MgSO<sub>4</sub>, 300 µM de dNTPs, 0,3 µM de chaque amorce, 6,25 ng d'ADNc et 0,5 U de Taq polymérase (Platinum® Pfx DNA Polymerase, Invitrogen).

Le thermocycleur Thermo Scientific Hybaid PCR (Thermal Cycler) est utilisé pour amplifier le gène. Les conditions de PCR consistent en une étape initiale de dénaturation à 94°C pendant 5 minutes suivie de 30 cycles d'amplification comprenant chacun une étape de dénaturation à 94°C pendant 30 secondes, une étape d'hybridation à 52°C pendant 30 secondes, puis une étape d'élongation à 68°C pendant 3 min, et enfin une étape d'extension finale à 72°C pendant 10 minutes.

Les produits de l'amplification du gène sont analysés par électrophorèse sur gel d'agarose 1% dans du tampon TBE 0,5X. La migration du fragment d'ADN est comparée à un marqueur de taille moléculaire 0,25-10 kb (GeneRuler™ 1kb DNA ladder, Fermentas).

Les fragments d'ADN sont extraits du gel d'agarose et purifiés à l'aide d'un kit d'extraction et de purification (QIAquick Gel Extraction Kit, QIAGEN) selon les instructions du fournisseur et comme décrit précédemment.

### 3.2.2. Digestion enzymatique du vecteur

Les vecteurs pFB-LIC-Bse et pFB-CT10HF-LIC sont respectivement digérés par les enzymes de restriction *BseRI* et *BfuAI* (New England Biolabs). Ces endonucléases, qui coupent les molécules d'ADN double brin au niveau des séquences spécifiques 5'...GAGGAG(N)<sub>10</sub>...3' (*BseRI*) et 5'...ACCTGC(N)<sub>4</sub>...3' (*BfuAI*), génèrent des fragments d'ADN linéaires de 4804 pb et 4805 pb aux extrémités cohésives.

L'enzyme de restriction *BseRI* présente une activité optimale (1 unité permet de digérer 1 µg d'ADN double brin en 1 heure à 37°C dans un volume réactionnel de 50 µl) dans le tampon NEBuffer 4 (50 mM acétate de potassium, 20 mM Tris-acétate, 10 mM acétate de magnésium et 1 mM Dithiothréitol, pH 7,9). L'enzyme *BfuAI* présente quant-à-elle une activité optimale lorsque la réaction a lieu à 50°C dans le tampon NEBuffer 3 (50 mM Tris-HCl, 100 mM NaCl et 10 mM MgCl<sub>2</sub>, pH 7,9).

Les digestions enzymatiques sont réalisées dans des volumes réactionnels finaux de 100 µl contenant du tampon NEBuffer 4 (digestion de pFB-LIC-Bse) ou NEBuffer 3 (digestion de pFB-CT10HF-LIC), 5 µg des vecteurs d'expression et 20 et 25 unités d'enzyme de restriction. Le volume réactionnel est ajusté à 100 µl avec de l'eau distillée stérile. Les milieux réactionnels sont respectivement incubés à 37°C et 50°C au bain-sec pendant deux heures. Les enzymes de restriction sont inactivées par un traitement de 20 minutes à 65 °C.

La digestion des vecteurs libère des fragments d'ADN d'environ 2000 pb. Il n'est pas nécessaire de séparer ces fragments des vecteurs linéarisés.

Les fragments d'ADN digérés sont purifiés à l'aide d'un kit de purification (QIAquick PCR Purification Kit, QIAGEN) selon les instructions du fournisseur et comme décrit précédemment.

### *3.2.3. Traitement par l'ADN polymérase du phage T4.*

Le traitement des produits de PCR et des vecteurs linéaires par l'ADN polymérase T4 permet de générer des extrémités 5' cohésives complémentaires.

5 µl des inserts sont mélangés avec 2,5 U d'ADN polymérase T4 et 2,5 mM de chacun des nucléotides (dCTP pour le clonage dans le vecteur pFB-LIC-Bse et dGTP pour le clonage dans le vecteur pFB-CT10HF-LIC). Le volume réactionnel est complété à 10 µl avec du tampon 67 mM Tris-HCl, pH 8,8, 6,7 mM MgCl<sub>2</sub>, 1mM DTT et 16,7 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>.

50 µl des vecteurs linéaires sont incubés avec 25 U d'ADN polymérase T4 en présence de 2,5 mM de chacun des nucléotides (dGTP pour le vecteur pFB-Lic-Bse et dCTP pour le vecteur pFB-CT10HF-Lic). Le volume réactionnel est complété à 100 µl avec du tampon.

Les milieux réactionnels sont incubés à 22°C pendant 30 minutes. L'activité de l'ADN polymérase T4 est inhibée par un traitement de 20 minutes à 75°C.

### *3.2.4. Ligation et transformation des bactéries XL1-Blue E. coli par choc thermique*

La ligation entre les extrémités cohésives des vecteurs linéaires et des inserts s'effectue sans ajout d'ADN ligase du phage T4. L'ADN polymérase du phage T4 présent dans le milieu réactionnel est suffisant à la formation de liaisons phosphodiester entre les extrémités 5' phosphate et 3' hydroxyle des deux brins d'ADN. La réaction est réalisée à température ambiante pendant 30 minutes. Pour chaque construction 1 µl de vecteur linéaire est additionné à 1 µl d'insert. En parallèle la réaction est réalisée en absence de l'insert ou de vecteur. Ces contrôles permettent de vérifier l'absence de faux-positifs.

Les cellules ultra-compétentes d'Escherichia coli XL1-Blue (Stratagène) ont été utilisées pour amplifier les différents plasmides recombinants.

Les bactéries compétentes (100 µl) sont décongelées sur glace. 1,7 µl de β-Mercaptoethanol (β-ME) y sont additionnés. (Il a été démontré que l'efficacité de la

transformation augmente en présence de  $\beta$ -ME). Le mélange est incubé sur glace pendant 10 minutes.

2  $\mu$ l des différentes constructions sont ajoutés à des aliquotes de bactéries compétentes, et les mélanges sont incubés 30 minutes sur glace. La transformation est réalisée par choc thermique à 42°C pendant 40 secondes. Les tubes sont ensuite placés sur glace pendant 2 minutes. Les bactéries sont resuspendues dans 100  $\mu$ l de milieu de culture SOC préchauffé à 42°C et incubées 1 heure à 37°C. Pour finir, les suspensions bactériennes sont étalées sur une gélose sélective de LB (Luria Broth ; BIO BASIC INC.) contenant de l'Ampicilline à une concentration de 100  $\mu$ g/ $\mu$ l et 5% de sucrose. Les boîtes de pétri sont incubées une nuit à 37°C.

Seules les bactéries transformées pourront se multiplier sur le milieu sélectif car les plasmides portent un gène de résistance contre l'antibiotique. La croissance des bactéries, ayant intégré un plasmide sans le gène d'intérêt, est inhibée via l'expression du gène SacB. Ce dernier code pour une protéine, la levansucrose, qui convertie le sucrose en un produit toxique pour les bactéries à gram négative.

### *3.2.5. Criblage des bactéries transformées par PCR*

La sélection des bactéries ayant incorporées le plasmide recombinant est effectuée par PCR. Les amorces sont identiques à celles ayant servi à l'amplification du gène ANPEP et de ses formes tronquées.

Les réactions de PCR sont effectuées dans un volume final de 20  $\mu$ l contenant du tampon réactionnel (50 mM Tris-HCl, pH 8,0, 50 mM KCl, 1 mM DTT, 0,1 mM EDTA et 50% (v/v) glycérol), 1 mM de  $MgSO_4$ , 300  $\mu$ M de dNTPs, 0,3  $\mu$ M de chaque amorce et 0,5 U de Taq polymérase (Platinum® *Pfx* DNA Polymerase, Invitrogen). Les colonies bactériennes sont directement inoculées dans le milieu réactionnel de PCR. Durant les premières étapes de dénaturation, les cellules sont lysées et leur ADN est libéré dans le milieu réactionnel où il sert alors de matrice.

Le thermocycleur Thermo Scientific Hybaid PCR (Thermal Cycler) a été utilisé pour amplifier le gène selon les conditions décrites précédemment.

Les produits de PCR sont séparés par migration sur gel d'agarose 1% dans du tampon TBE 0,5X.

### 3.2.6. Extraction et purification de l'ADN plasmidique produit par les bactéries

Pour chaque construction un clone positif est inoculé dans 1 ml de milieu LB contenant 100 µg/ml d'ampicilline, et incubé pendant environ 8 heures à 37 °C sous agitation. Ces pré-cultures sont utilisées pour ensemercer (dilution au 1/500<sup>e</sup>) 3 ml de milieu LB additionné de l'antibiotique. Après une nuit d'incubation à 37 °C sous agitation, les bactéries sont sédimentées par centrifugation (6000 g pendant 15 minutes).

Les ADN plasmidiques sont extraits et purifiés à l'aide d'un kit (QIAprep Spin Miniprep Kit) selon les instructions du fournisseur. Les culots bactériens sont resuspendus dans 250 µl de tampon P1 (50 mM Tris-HCl, pH 8,0, 10 mM EDTA) contenant 100 µg/ml de RNase A. Les bactéries sont ensuite lysées par un traitement alcalin (250 µl de tampon P2 (200 mM NaOH, 1% SDS (w/v)) pendant 5 minutes à température ambiante. L'agent de lyse est neutralisé par l'ajout de 350 µl ml de tampon N3. Les débris cellulaires sont séparés par centrifugation (17 900 x g pendant 10 minutes). Les surnageants qui contiennent les ADN plasmidiques sont déposés sur une résine échangeuse d'anions en présence de sels en faible concentration. Les contaminants (ARN, protéines et autres petites impuretés) sont éliminés par une série de lavage avec du tampon PE. Pour finir, les ADN plasmidiques sont élués avec 50 µl de tampon EB (10 mM Tris-HCl, pH 8,5).

La concentration en ADN est évaluée par mesure de l'absorbance à 260 nm à l'aide du NanoDrop ND-1000.

## 3.3. Génération des bacmides recombinants

### 3.3.1. Transformation des bactéries *E.coli* DHBac10 par choc thermique

La transformation des plasmides recombinants dans les cellules compétentes *E. coli* DHBac10 permet de générer des bacmides recombinants par transposition site-spécifique de la cassette d'expression contenant le gène d'intérêt dans le génome du baculovirus.

Les bactéries compétentes DHBac10 (Invitrogen) sont décongelées sur glace (aliquotes de 100 µl). 5 ng de plasmide recombinant y sont ajoutés et les mélanges sont incubés 30 minutes sur glace. La transformation est réalisée par choc thermique à 42°C pendant 45 secondes. Les tubes sont ensuite placés sur glace pendant 2 minutes. Les bactéries sont resuspendues dans 900 µl de milieu de culture SOC préchauffé à 42°C et incubées 4 heures à 37°C, sous agitation. 100 µl de suspension bactérienne sont étalées sur une gélose LB contenant 50 µg/ml de kanamycine, 7 µg/ml de gentamycine, 10 µg/ml de tétracycline, 100

$\mu\text{g/ml}$  de Bluo-gal et  $40 \mu\text{g/ml}$  d'IPTG. La kanamycine permet de sélectionner les bactéries transformées possédant le bacmide, la gentamycine, les bactéries ayant intégré le gène d'intérêt situé entre les éléments transposables Tn7 du vecteur donneur et la tétracycline, les bactéries possédant le plasmide « helper » nécessaire à la transposition. L'IPTG permet d'induire l'expression de la  $\beta$ -galactoside (gène lacZ). Le Bluo-gal est le substrat de la  $\beta$ -galactoside ; il est dégradé en galactose et en un composé bleu. Après 48h d'incubation à  $37^\circ\text{C}$ , si l'insert a été transposé dans le bacmide, les bactéries transformées seront lac- et apparaîtront blanches, alors que si la transposition n'a pas eu lieu, les colonies seront bleues.

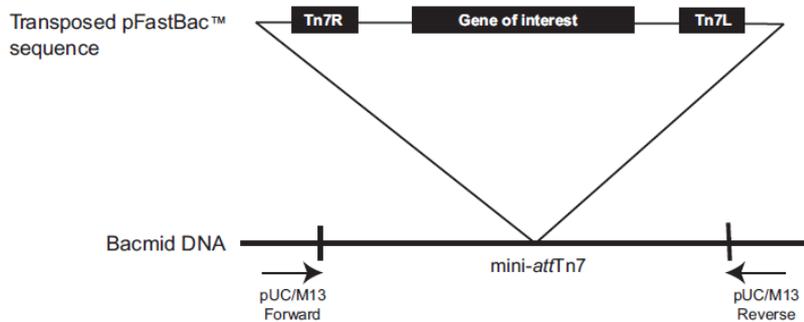
### *3.3.2. Isolement et purification des bacmides recombinants*

Pour chaque construction, trois colonies bien blanches et bien rondes sont prélevées et repiquées sur une gélose LB contenant les mêmes antibiotiques que précédemment. Cette étape permet d'éviter les faux positifs. Après 48 heures d'incubation à  $37^\circ\text{C}$ , trois clones positifs sont inoculés dans 3 ml de milieu LB contenant  $50 \mu\text{g/ml}$  de kanamycine,  $7 \mu\text{g/ml}$  de gentamycine et  $10 \mu\text{g/ml}$  de tétracycline. Après une nuit d'incubation à  $37^\circ\text{C}$  sous agitation, les bactéries sont sédimentées par centrifugation ( $6000 \text{ g}$  pendant 15 minutes). Les culots bactériens sont resuspendus dans  $300 \mu\text{l}$  de tampon  $50 \text{ mM}$  Tris-HCl,  $\text{pH } 8,0$ ,  $10 \text{ mM}$  EDTA contenant  $100 \mu\text{g/ml}$  de RNase A. Les bactéries sont ensuite lysées par un traitement alcalin ( $300 \mu\text{l}$  de tampon  $200 \text{ mM}$  NaOH,  $1\%$  SDS (w/v) pendant 5 minutes à température ambiante. L'agent de lyse est neutralisé par l'ajout de  $300 \mu\text{l}$  ml de tampon  $3 \text{ M}$  acétate de potassium,  $\text{pH } 5,5$ . Les débris cellulaires sont séparés par centrifugation ( $17\,900 \times \text{g}$  pendant 10 minutes). Les bacmides recombinants sont précipités à l'aide d'isopropanol. Le précipité est rincé avec de l'éthanol  $70\%$  puis dissout dans  $40 \mu\text{l}$  de tampon TE ( $10 \text{ mM}$  Tris·Cl,  $\text{pH } 8,0$ ,  $1 \text{ mM}$  EDTA).

La concentration en ADN des bacmides recombinants est évaluée par mesure de l'absorbance à  $260 \text{ nm}$  à l'aide du Nanodrop ND-1000.

### 3.3.3. Analyse des bacmides recombinants par PCR

La transposition de la cassette d'expression dans le bacmide est vérifiée par PCR en utilisant des amorces situées sur le bacmide de part et d'autre de l'insert (Figure 76).



Primer	Sequence
pUC/M13 Forward	5'-CCCAGTCACGACGTTGTA AAAACG-3'
pUC/M13 Reverse	5'-AGCGGATAACAATTTACACAGG-3'

**Figure 76 :** Séquences des amorces pUC/M13 et zones d'hybridation. Les amorces sont synthétisées par la société Eurofins mwg<sup>132</sup>. Elles s'hybrident spécifiquement sur le bacmide au niveau du site de transposition mini-attTn7, situé dans le gène LacZ. Un fragment de 2300 pb + la taille de l'insert est amplifié lorsque la transposition de la cassette d'expression dans le bacmide s'est produite. Dans le cas contraire, seul le site de transposition mini-attTn7 est amplifié (fragment de 300 pb).

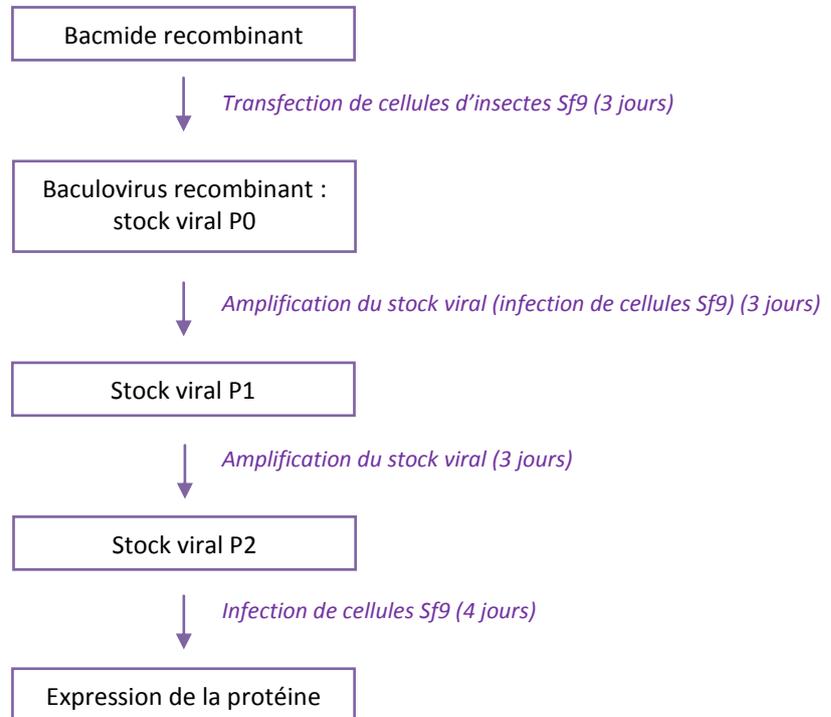
Les réactions de PCR sont effectuées dans un volume final de 25 µl contenant du tampon réactionnel (50 mM Tris-HCl, pH 8,0, 50 mM KCl, 1 mM DTT, 0,1 mM EDTA et 50% (v/v) glycérol), 1 mM de MgSO<sub>4</sub>, 300 µM de dNTPs, 0,3 µM de chaque amorce, 0,5 U de Taq polymérase (Platinum® Pfx DNA Polymerase, Invitrogen) et 0,5 µl de bacmide recombinant.

Le thermocycleur Thermo Scientific Hybaid PCR (Thermal Cycler) est utilisé pour amplifier le gène. Les conditions de PCR consistent en une étape initiale de dénaturation à 94°C pendant 3 minutes suivie de 30 cycles d'amplification comprenant chacun une étape de dénaturation à 94°C pendant 30 secondes, une étape d'hybridation à 50°C pendant 45 secondes, puis une étape d'élongation à 72°C pendant 5 min, suivis d'une étape d'extension finale à 72°C pendant 7 minutes.

Les produits de PCR sont analysés par migration sur gel d'agarose 1% dans du tampon TBE 0,5X.

### 3.4. Culture cellulaire et expression de la protéine recombinante

La Figure 77 résume les étapes de culture cellulaire qui permettent d'aboutir à la production de la protéine.



*Figure 77 : Schéma récapitulatif des différentes étapes de culture cellulaire à réaliser pour exprimer la protéine d'intérêt.*

#### 3.4.1. Culture des cellules

La lignée cellulaire Sf9 (Invitrogen) utilisée dans le système d'expression baculovirus est issue de cellules ovariennes du lépidoptère *Spodoptera frugiperda*. Les cellules sont cultivées en monocouche (transfection) ou en suspension (amplification du virus recombinant et expression de la protéine recombinante) à 27°C dans un milieu de culture sans sérum Sf-900 II SFM (Invitrogen). Leur culture ne nécessite pas l'emploi d'une étuve à CO<sub>2</sub>. Le temps de doublement des populations de cellules est de 24 à 48 heures, et la densité cellulaire varie de 2 à 5x 10<sup>6</sup> cellules/ml.

### *3.4.2. Transfection des cellules d'insectes Sf9*

Le baculovirus recombinant est généré par transfection du bacmide recombinant dans les cellules d'insectes Sf9. La transfection est réalisée avec le kit Nanofectin de PAA selon les instructions du fournisseur. Une nouvelle génération de nanoparticules poreuses fonctionnalisées avec des polymères cationiques est utilisée comme agent de transfection. Le polycation fixe les groupements phosphates négatifs de l'ADN : l'ADN est encapsulé dans les pores des nanoparticules le protégeant ainsi des scissions endonucléosidiques. Ces complexes chargés positivement interagissent directement avec la membrane chargée négativement des cellules, ce qui en facilite l'internalisation par endocytose.

Les cellules Sf9 sontensemencées à une densité  $2 \times 10^5$  cellules/ml dans du milieu Sf-900 II SFM dans une plaque 24 puits (0,5 ml/puits). Après 2 heures d'incubation à 27°C, les cellules forment une monocouche.

1 µg de bacmide recombinant est dilué dans 50 µl de 150 mM NaCl (Diluent). L'agent de transfert Nanofectin (3,2 µl) est dilué dans du tampon 150 mM NaCl (50 µl) puis, additionné à la solution contenant le bacmide recombinant. Le mélange est incubé pendant 30 minutes à température ambiante et ajouté goutte à goutte sur les cellules Sf9 qui ont adhéré dans le puits. Après 2 heures d'incubation à 27°C, 0,5 ml de milieu de culture Sf-900 II SFM sont ajoutés dans le puits. Les cellules sont incubées à 27°C pendant 72 heures supplémentaires. Le surnageant contenant les baculovirus recombinants est séparé des débris cellulaires par centrifugation puis, conservé à 4°C à l'abri de la lumière ; il constitue le stock viral, et ne contient qu'une faible quantité de baculovirus recombinant (stock P0), qui doit être amplifié.

### *3.4.3. Amplification du stock viral*

Les baculovirus recombinants sont amplifiés en infectant des cellules d'insectes Sf9 cultivées en suspension. Les cellules Sf9 sontensemencées à une densité de  $2 \times 10^6$  cellules/ml dans du milieu Sf-900 II SFM dans une plaque 24 puits (3 ml/puits). 100 µl de stock viral P0 sont ajoutés dans le puits. Après 72h d'incubation à 27°C sous agitation, le surnageant qui contient une concentration plus élevée de baculovirus recombinant est récupéré et conservé à 4°C dans l'obscurité. Ce stock viral P1 est à son tour amplifié. 500 µl de stock viral P1 sont additionnés à 50 ml de cellules Sf9ensemencées à une densité de  $2 \times 10^6$  cellules/ml ( $1 \times 10^8$  cellules) dans du milieu Sf-900 II SFM. Les cellules sont cultivées en suspension à 27°C pendant 72 heures. Le stock de baculovirus recombinants obtenu (stock viral P2) est suffisamment concentré pour infecter un grand nombre de cellules Sf9 et produire la protéine recombinante.

#### *3.4.4. Expression de la protéine recombinante*

Des aliquotes (5 ml) de stock viral P2 sont prélevés. Les cellules d'insectes sont sédimentées par centrifugation à 500 g pendant 15 minutes. Afin d'évaluer le niveau d'expression et localiser la protéine recombinante Les culots cellulaires et les surnageants sont analysés sur gel SDS-PAGE comme décrit ci-dessous.

Les échantillons protéiques (culots cellulaires et surnageants de culture) sont mélangés à du tampon de charge 4x (NuPAGE® LDS Sample Buffer 4x, Invitrogen), qui contient du glycérol, des colorants et du lithium dodecyl sulfate (LDS) à pH 8,4. Les protéines sont dénaturées par chauffage à 100°C pendant 5 minutes puis, séparées sur un gel de polyacrylamide en gradient 4-12% Bis-Tris (Criterion™ XT Precast gel, Bio-Rad). L'électrophorèse est conduite à un voltage constant de 170 Volts durant 45 minutes dans du tampon de migration 50 mM MES, 50 mM Tris Base, 0.1% SDS, 1 mM EDTA, pH 7,3. Une fois la migration terminée, les protéines sont révélées par une coloration au bleu de Coomassie. Le gel est incubé pendant 15 minutes dans du milieu de coloration (Instant Blue, Novexin). L'excès de colorant est ensuite éliminé par des lavages successifs à l'eau distillée. Le marqueur de poids moléculaire (Novex Sharp 3,5-260 KDa, Invitrogen) est employé comme référence.

Le stock de baculovirus recombinant P2 (5 ml) est utilisé pour infecter 1000 ml de cellules Sf9ensemencées à une densité de  $2 \times 10^6$  cellules/ml ( $2 \times 10^9$  cellules). Les cellules sont cultivées en suspension à 27°C pendant 96 heures. Les protéines sont exprimées dans le cytoplasme des cellules (formes solubles) ou ancrée dans la membrane (forme membranaire). Les culots cellulaires sont récupérés par centrifugation et lavés avec du tampon PBS 1x avant d'être conservé à -80°C. A la fin du séjour, les culots sont envoyés en France afin d'extraire et de purifier la protéine h-APN.

## 4. Purification de la protéine recombinante

La purification des protéines débute toujours par une phase d'extraction.

### 4.1. Extraction de la protéine

#### *4.1.1. Protéine exprimée sous forme de corps d'inclusion dans le cytoplasme des bactéries*

La forme soluble de la protéine h-APN est accumulée dans le cytoplasme des bactéries sous forme d'agrégats insolubles (corps d'inclusion).

##### *i. Isolement et purification des corps d'inclusion*

Le culot bactérien est remis en suspension dans du tampon 50 mM Tris-HCl pH 8,0 (5 ml de tampon par gramme de bactéries). Les bactéries sont lysées par sonication (Branson Ultrasonics 450 Sonifier) pendant 6 à 9 minutes (pulsations de 15 secondes à une amplitude de 50 %, repos de 1 minute entre les pulsations). Pour éviter la dégradation des protéines pendant la lyse, le tube est placé dans un bain de glace. La lyse est complète lorsque la concentration en protéine dans la fraction soluble (surnageant après centrifugation) est constante.

Lorsque la lyse est complète, 100 µg/ml de Lysozyme et 10 µg/ml de DNase I sont ajoutés et le lysat est incubé 30 minutes à température ambiante. Cette étape permet d'augmenter la pureté des corps d'inclusion en dégradant les résidus polysaccharidiques (Lysozyme) et les acides désoxyribonucléiques (DNase I). Les corps d'inclusion sont sédimentés par centrifugation (17000 g pendant 20 minutes à 4°C).

Les corps d'inclusion sont ensuite purifiés par une série de lavage : 3 fois avec du tampon 50 mM Tris-HCl pH 8,0, 0,5 M NaCl, 1% Triton X-100, 5 mM Dithiothréitol et 2 fois avec du tampon 50 mM Tris-HCl pH 8,0. Après avoir écrasé les corps d'inclusion avec une spatule, ces derniers sont remis en suspension dans du tampon de lavage (10 ml de tampon de lavage par gramme de corps d'inclusion). Le mélange est ensuite homogénéisé à l'aide d'un potter et centrifugé à 17000 g pendant 20 minutes à 4 °C.

Lors du dernier lavage, la concentration en protéine dans la suspension est évaluée par mesure de l'absorbance à 280 nm.

### *ii. Solubilisation des corps d'inclusion*

Les corps d'inclusion sont solubilisés dans du tampon 50 mM Tris-HCl pH 8,0 ou 50 mM CAPS pH 11, contenant 8 M d'Urée (agent dénaturant) et 10 mM de DTT (agent de réduction) : 1 ml de tampon de solubilisation pour 100 mg de corps d'inclusion. La suspension est incubée à température ambiante pendant 1 heure sous agitation. La fraction soluble est séparée des débris insolubles par centrifugation (17000 g pendant 20 minutes à 4 °C).

La concentration en protéine dans la fraction soluble est évaluée par mesure de l'absorbance à 280 nm. La concentration finale en protéine dans la fraction soluble doit être comprise entre 1 et 10 mg/ml.

### *iii. Renaturation de la protéine par dilution*

L'agent dénaturant doit être éliminé pour permettre le repliement de la protéine ; le tampon de renaturation est ajouté goutte à goutte à la solution protéiques solubilisée dans le tampon Tris, pH 8,0 (dilution d'un facteur 20). Afin de trouver les conditions optimales de renaturation in vitro de la protéine, différents tampons sont testés en parallèle ([Figure 78](#)). Après une nuit d'incubation à 4°C, le repliement correct de la protéine est vérifié par mesure de l'activité enzymatique.

QuickFold™ Protein Refolding Buffer Formulation			
Component	Formulation after Preparation (Addition of DTT, GSH, and GSSH)	DTT	GSH/GSSH
Buffer 1	50 mM MES pH 6.0, 9.6 mM NaCl, 0.4 mM KCl, 2 mM MgCl <sub>2</sub> , 2 mM CaCl <sub>2</sub> , 0.75 M Guanidine HCl, 0.5% Triton X-100, 1 mM DTT	●	
Buffer 2	50 mM MES pH 6.0, 9.6 mM NaCl, 0.4 mM KCl, 2 mM MgCl <sub>2</sub> , 2 mM CaCl <sub>2</sub> , 0.5 M arginine, 0.05% polyethylene glycol 3,550, 1 mM GSH, 0.1 mM GSSH		●
Buffer 3	50 mM MES pH 6.0, 9.6 mM NaCl, 0.4 mM KCl, 1 mM EDTA, 0.4 M sucrose, 0.75 M Guanidine HCl, 0.5% Triton X-100, 0.05% polyethylene glycol 3,550, 1 mM DTT	●	
Buffer 4	50 mM MES pH 6.0, 240 mM NaCl, 10 mM KCl, 2 mM MgCl <sub>2</sub> , 2 mM CaCl <sub>2</sub> , 0.5 M arginine, 0.5% Triton X-100, 1 mM GSH, 0.1 mM GSSH		●
Buffer 5	50 mM MES pH 6.0, 240 mM NaCl, 10 mM KCl, 1 mM EDTA, 0.4 M sucrose, 0.75 M Guanidine HCl, 1 mM DTT	●	
Buffer 6	50 mM MES pH 6.0, 240 mM NaCl, 10 mM KCl, 1 mM EDTA, 0.5 M arginine, 0.4 M sucrose, 0.5% Triton X-100, 0.05% polyethylene glycol 3,550, 1 mM GSH, 0.1 mM GSSH		●
Buffer 7	50 mM MES pH 6.0, 240 mM NaCl, 10 mM KCl, 2 mM MgCl <sub>2</sub> , 2 mM CaCl <sub>2</sub> , 0.75 M Guanidine HCl, 0.05% polyethylene glycol 3,550, 1 mM DTT	●	
Buffer 8	50 mM Tris-Cl pH 8.5, 9.6 mM NaCl, 0.4 mM KCl, 2 mM MgCl <sub>2</sub> , 2 mM CaCl <sub>2</sub> , 0.4 M sucrose, 0.5% Triton X-100, 0.05% polyethylene glycol 3,550, 1 mM GSH, 0.1 mM GSSH		●
Buffer 9	50 mM Tris-Cl pH 8.5, 9.6 mM NaCl, 0.4 mM KCl, 1 mM EDTA, 0.5 M arginine, 0.75 M Guanidine HCl, 0.05% polyethylene glycol 3,550, 1 mM DTT	●	
Buffer 10	50 mM Tris-Cl pH 8.5, 9.6 mM NaCl, 0.4 mM KCl, 2 mM MgCl <sub>2</sub> , 2 mM CaCl <sub>2</sub> , 0.5 M arginine, 0.4 M sucrose, 0.75 M Guanidine HCl, 1 mM GSH, 0.1 mM GSSH		●
Buffer 11	50 mM Tris-Cl pH 8.5, 9.6 mM NaCl, 0.4 mM KCl, 1 mM EDTA, 0.5% Triton X-100, 1 mM DTT	●	
Buffer 12	50 mM Tris-Cl pH 8.5, 240 mM NaCl, 10 mM KCl, 1 mM EDTA, 0.05% polyethylene glycol 3,550, 1 mM GSH, 0.1 mM GSSH		●
Buffer 13	50 mM Tris-Cl pH 8.5, 240 mM NaCl, 10 mM KCl, 1 mM EDTA, 0.5 M arginine, 0.75 M Guanidine HCl, 0.5% Triton X-100, 1 mM DTT	●	
Buffer 14	50 mM Tris-Cl pH 8.5, 240 mM NaCl, 10 mM KCl, 2 mM MgCl <sub>2</sub> , 2 mM CaCl <sub>2</sub> , 0.5 M arginine, 0.4 M sucrose, 0.75 M Guanidine HCl, 0.5% Triton X-100, 0.05% polyethylene glycol 3,550, 1 mM GSH, 0.1 mM GSSH		●
Buffer 15	50 mM Tris-Cl pH 8.5, 240 mM NaCl, 10 mM KCl, 2 mM MgCl <sub>2</sub> , 2 mM CaCl <sub>2</sub> , 0.4 M sucrose, 1 mM DTT	●	

**Figure 78 :** Criblage des conditions de renaturation- kit QuickFold Protein Refolding (Athena Enzyme Systems)<sup>138</sup>.

Des essais sont également réalisés en ajoutant du chlorure de zinc (ZnCl<sub>2</sub>) dans les tampons de renaturation.

#### iv. Renaturation de la protéine par dialyse (Protein Refolding Kit, Novagen)

La solution protéique solubilisée dans du tampon CAPS, pH 11,0 est dialysée dans du tampon 20 mM Tris-HCl, pH 8,0, 100 µM ZnCl<sub>2</sub> pendant une nuit à 4°C. Pour favoriser la formation des ponts disulfures, le couple d'oxydo-réduction Glutathion réduit (GSH)/ Glutathion oxydé (GSSG) est additionné au tampon de dialyse (proportion 5 :1 (GSH 1 mM : GSSG 0,2 mM)).

Le repliement de la protéine sous sa configuration native est vérifié par mesure de l'activité enzymatique.

<sup>138</sup> Manuel d'instructions « QuickFold Protein Refolding kit ». Version 2.1. Athena Enzyme Systems. [http://www.athenaes.com/tech\\_documents/ApplicationManual\\_QuickFold.pdf](http://www.athenaes.com/tech_documents/ApplicationManual_QuickFold.pdf).

#### *4.1.2. Protéine exprimée sous forme soluble dans le cytoplasme des cellules d'insectes*

Le culot cellulaire est remis en suspension dans du tampon de lyse : 50 mM Tris-HCl, pH 7,5, 0,5 M NaCl, 5% glycérol et 1% Triton X-100 (100 ml de tampon de lyse par litre de culture). La suspension est homogénéisée et incubée pendant 3 heures à 4°C. La fraction soluble (contenant la protéine d'intérêt) est séparée des débris insolubles par centrifugation (20000 g pendant 30 minutes à 4 °C).

### 4.2. Purification de la protéine recombinante

Différentes techniques de purification (décrites ci-dessous) sont utilisées pour séparer la protéine d'intérêt des contaminants protéiques. Les queues polyhistidine ou FLAG facilitent les étapes de purification par chromatographie d'affinité. Toutefois, si le degré de pureté n'est pas satisfaisant, d'autres méthodes chromatographiques peuvent être associées.

Durant les différentes étapes de purification, les échantillons sont maintenus sur glace ou à 4°C pour éviter une dégradation par les protéases endogènes. Les séparations par chromatographie sont effectuées à 4°C à l'aide du système ÄKTA Purifier de GE Healthcare. La chromatographie est suivie par mesure de l'absorbance à 280 nm.

#### *4.2.1. Suivi de purification*

##### *i. Dosage des protéines*

La concentration des protéines en solution est évaluée par mesure de l'absorbance à 280 nm à l'aide du NanoDrop ND-1000. En effet, le tryptophane absorbe fortement la lumière UV à 280 nm de même que, dans une moindre mesure, la tyrosine. La phénylalanine et la cystéine absorbent aussi faiblement dans cette région de l'UV. Cependant, la présence de contaminants tels que les acides nucléiques, qui absorbent fortement dans l'UV aux alentours de 260nm, fausse la mesure. Il est donc nécessaire, pour déterminer la concentration protéique d'un mélange contenant une proportion donnée d'acides nucléiques, de mesurer l'absorbance à 280 nm et à 260 nm. Ces valeurs peuvent alors s'intégrer dans l'équation suivante qui donne une approximation de la concentration massique en protéine dans l'échantillon :

$$\text{Protéine (mg/ml)} = 1,55 \times A_{280\text{nm}} - 0,76 \times A_{260\text{nm}}$$

Lorsqu'il n'y a plus d'acide nucléique, la formule précédente devient inutile. Une bonne approximation consiste alors à déterminer la concentration en protéine en utilisant le coefficient d'extinction spécifique de la protéine pure  $\epsilon_{1\text{mg/ml}}^{280}$  (ml. mg<sup>-1</sup>. cm<sup>-1</sup>). Une solution à 1 mg/ml de h-APN pur, dont l'absorbance est mesurée à 280 nm, a un coefficient d'extinction égale à 1,803 ml. mg<sup>-1</sup>. cm<sup>-1</sup>.

## ii. *Activité enzymatique*

Le spectrophotomètre UV-Visible HP 8453 (Agilent Technologies) est utilisé pour mesurer l'activité enzymatique après chaque étape de purification. Une unité d'activité enzymatique (U) est définie comme la quantité d'enzyme qui catalyse la transformation de 1 µmole de substrat par minute, en présence d'une concentration saturante en substrat et dans des conditions expérimentales précises (tampon, pH, température). Les cinétiques enzymatiques répondent au modèle de Michaelis et Menten :

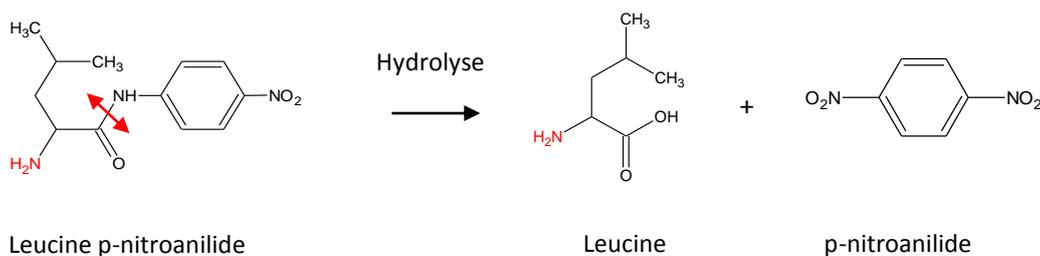


L'équation décrivant la vitesse initiale de formation du produit en fonction de la concentration en substrat est la suivante (équation de Michaelis et Menten) :

$$V_i = \frac{V_{\max} [S]}{K_M + [S]}$$

La constante de Michaelis-Menten  $K_M$  représente l'affinité de l'enzyme pour le substrat et correspond à la concentration en substrat pour laquelle la vitesse de réaction ( $V_i$ ) est égale à la moitié de la vitesse maximum ( $V_{\max}$ ).

La leucine p-nitroanilide est un substrat de h-APN. Son hydrolyse conduit à la p-nitroaniline qui adsorbe dans le visible à 405 nm ( $\epsilon = 10800 \text{ mol}^{-1} \cdot \text{L} \cdot \text{cm}^{-1}$ ) (Figure 79).



**Figure 79 :** Hydrolyse de la Leucine p-nitroanilide par la protéine h-APN

Dans une cuve en PMMA (polyméthacrylate de méthyle) de 1 ml sont introduits 50  $\mu$ l d'une solution de leucine p-nitroanilide à 40 mM ( $[S]=10K_M=2mM$ ), du tampon 20 mM Tris-HCl pH 7,5 et une fraction d'échantillon protéique en quantité adéquate. La vitesse initiale d'hydrolyse du substrat par l'enzyme est mesurée à 30°C en suivant par spectrophotométrie la cinétique d'apparition du p-nitroaniline à 405 nm.

Connaissant la vitesse de formation du p-nitroaniline ( $\Delta Do \cdot \text{min}^{-1}$ ) ainsi que la loi de Beer Lambert  $Do = \epsilon \cdot l \cdot c$ , il est facile de déterminer la concentration en p-nitroaniline (en  $\mu\text{mol} \cdot \text{min}^{-1} \cdot \text{ml}^{-1}$  ou U/ml) ainsi que la quantité d'enzyme active ( $U_{\text{totale}}$ ) dans l'échantillon. A partir du nombre d'unités totales et de la masse de protéines présentes en solution, il est également possible de calculer l'activité spécifique AS, c'est-à-dire la quantité d'enzyme qui catalyse la transformation de 1  $\mu$ mole de substrat en une minute par mg de protéine. C'est une mesure du degré de pureté de l'enzyme ; sa valeur augmente au cours des étapes de purification et est maximale quand l'enzyme est pure.

Chaque étape de purification conduit inévitablement à des pertes de la protéine d'intérêt. Le calcul du rendement de purification, *proportion en pourcentage de la quantité de protéine purifiée (en unité) qui reste par rapport à la quantité initiale*, permet d'effectuer un bilan après chaque étape. Par ailleurs, le gain de pureté est évalué en détermination le taux de purification, *rapport de l'AS mesurée après une étape de purification sur l'AS mesurée initialement*.

#### 4.2.2. Dessalage et concentration par ultrafiltration

La teneur en sel de la solution protéique est modifiée entre les différentes étapes de purification par dialyse ou ultrafiltration.

La *dialyse* permet de séparer les molécules selon leur taille, en utilisant des membranes de cellulose semi-perméables dont les pores ont une taille inférieure aux macromolécules.

Par effet d'osmose et d'agitation moléculaire les sels traversent la membrane selon le gradient de concentration jusqu'à atteindre l'équilibre.

La solution protéique est introduite dans une membrane de dialyse (taille des pores 5000 Da). Cette dernière est plongée dans un volume de tampon 100 à 200 fois supérieure et incubée à 4°C pendant 1 nuit.

L'*ultrafiltration* est un procédé de séparation sélective utilisé pour concentrer ou dessaler des solutions protéiques très diluées par centrifugation. Les cellules d'ultrafiltration Amicon Ultra de capacité 0,5, 4 ou 15 ml (Millipore) sont utilisées ; elles sont composées d'une membrane en cellulose régénérée à faible absorption ayant un seuil de coupure de 30, 50 ou 100 kDa. La solution protéique est déposée dans la partie supérieure du tube, et repose sur la membrane. Le tube est alors centrifugé à basse vitesse (5000 g) pendant 15-30 minutes à 4°C, entraînant ainsi le passage du tampon et des petites molécules à travers la membrane. Les protéines de haut poids moléculaire ne pouvant pas traverser la membrane, restent dans la partie supérieure du tube. La fraction récupérée peut être soit diluée dans le tampon de son choix (dialyse), soit conservée (concentration).

Après les étapes de dialyse ou d'ultrafiltration, la concentration en protéine et l'activité enzymatique sont évaluées.

#### *4.2.3. Précipitation différentielle au sulfate d'ammonium*

La précipitation différentielle au sulfate d'ammonium est une méthode basée sur la séparation des protéines en fonction de leur solubilité.

La solubilité d'une protéine en solution aqueuse est fonction de la concentration en sels dissous ; elle diminue lorsque la force ionique de la solution est élevée. Les protéines précipitent lorsque la concentration en sel est suffisante pour neutraliser les charges ioniques requises en surface pour le maintien de la solubilité, et pour priver les protéines des molécules d'eau disponibles en solution pour les hydrater.

Le sel le plus utilisé pour précipiter les protéines est le sulfate d'ammonium  $(\text{NH}_4)_2\text{SO}_4$  car il est très soluble en solution aqueuse et permet d'atteindre des forces ioniques très élevées. Il a peu d'effet dénaturant et maximise l'obtention de protéines biologiquement actives.

La concentration en sel de la solution protéique est amenée à une valeur tout juste inférieure à celle pour laquelle la protéine d'intérêt précipite. Du sulfate d'ammonium est ajouté progressivement à la solution, et le mélange est agité pendant 1 heure à 4°C. Tous les contaminants protéiques qui précipitent à cette concentration sont éliminés par

centrifugation (10 000 xg pendant 30 minutes à 4°C). La protéine recherchée est ensuite précipitée en augmentant la concentration en sel dans le surnageant. Après centrifugation, le culot est resuspendu dans du tampon 50 mM Tris-HCl, pH 7,5. Ce dernier contient encore une grande quantité de sulfate d'ammonium qu'il faut éliminer par dialyse. Le Tableau 21 indique les quantités de sulfate d'ammonium (en g) à ajouter à 100 ml de solution protéique pour atteindre le pourcentage de saturation souhaité.

The amount of solid ammonium sulphate to be added to a solution to give the desired final saturation at 0°C.																	
<i>Final concentration of ammonium sulphate, % saturation at 0°C</i>																	
	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100
Initial concentration of ammonium sulphate																	
<i>g solid ammonium sulphate to add to 100 ml of solution</i>																	
0	10.7	13.6	16.6	19.7	22.9	26.2	29.5	33.1	36.6	40.4	44.2	48.3	52.3	56.7	61.1	65.9	70.7
5	8.0	10.9	13.9	16.8	20.0	23.2	26.6	30.0	33.6	37.3	41.1	45.0	49.1	53.3	57.8	62.4	67.1
10	5.4	8.2	11.1	14.1	17.1	20.3	23.6	27.0	30.5	34.2	37.9	41.8	45.8	50.0	54.5	58.9	63.6
15	2.6	5.5	8.3	11.3	14.3	17.4	20.7	24.0	27.5	31.0	34.8	38.6	42.6	46.6	51.0	55.5	60.0
20	0	2.7	5.6	8.4	11.5	14.5	17.7	21.0	24.4	28.0	31.6	35.4	39.2	43.3	47.6	51.9	56.5
25		0	2.7	5.7	8.5	11.7	14.8	18.2	21.4	24.8	28.4	32.1	36.0	40.1	44.2	48.5	52.9
30			0	2.8	5.7	8.7	11.9	15.0	18.4	21.7	25.3	28.9	32.8	36.7	40.8	45.1	49.5
35				0	2.8	5.8	8.8	12.0	15.3	18.7	22.1	25.8	29.5	33.4	37.4	41.6	45.9
40					0	2.9	5.9	9.0	12.2	15.5	19.0	22.5	26.2	30.0	34.0	38.1	42.4
45						0	2.9	6.0	9.1	12.5	15.8	19.3	22.9	26.7	30.6	34.7	38.8
50							0	3.0	6.1	9.3	12.7	16.1	19.7	23.3	27.2	31.2	35.3
55								0	3.0	6.2	9.4	12.9	16.3	20.0	23.8	27.7	31.7
60									0	3.1	6.3	9.6	13.1	16.6	20.4	24.2	28.3
65										0	3.1	6.4	9.8	13.4	17.0	20.8	24.7
70											0	3.2	6.6	10.0	13.6	17.3	21.2
75												0	3.2	6.7	10.2	13.9	17.6
80													0	3.3	6.8	10.4	14.1
85														0	3.4	6.9	10.6
90															0	3.4	7.1
95																0	3.5
100																	0

Tableau 21 : Quantités de sulfate d'ammonium (en g) à ajouter à 100 ml de solution protéique pour atteindre le pourcentage de saturation souhaité, à 0°C<sup>139</sup>.

Cette technique peut également être employée pour concentrer les échantillons protéiques très dilués.

#### 4.2.4. Chromatographie d'affinité

##### i. Résine de nickel

La queue polyhistidine fusionnée du côté N ou C-terminal de la protéine recombinante permet sa purification par chromatographie d'affinité sur résine de nickel (Tableau 22). Le principe de cette technique repose sur les interactions ioniques qui se forment entre les noyaux imidazoles des résidus histidines de la protéine et les charges positives des ions

<sup>139</sup> Harris, E.L.V., Angal, S., 1989. Protein purification methods, a practical approach, ED. Oxford University Press, ISBN 0 19 963002 X

nickel de la résine. La protéine est ensuite éluée par un mécanisme compétitif en ajoutant des quantités croissantes d'imidazole.

Nom	Matrice	Taille des billes	Capacité dynamique de liaison
HisTrap FF 5 ml	Billes d'agarose réticulée 6% recouvertes de nickel	90 µm	200 mg de protéine-queue (His) <sub>6</sub>
HisTrap HP 1 ml	Billes d'agarose réticulée 6% recouvertes de nickel	34 µm	40 mg de protéine-queue (His) <sub>6</sub>

*Tableau 22 : Caractéristiques des colonnes d'affinité utilisées pour séparer les protéines fusionnées à une queue polyhistidine*

La colonne est équilibrée avec du tampon A : 50 mM HEPES pH 7,5, 0,5 M NaCl. L'extrait protéique est chargé sur la colonne à un débit de 1 ou 5 ml/min. Après un lavage avec le tampon A (retour du signal de détection UV 280 nm à la ligne de base), la protéine est éluée avec du tampon B (50 mM HEPES pH 7,5, 0,5 M NaCl et 500 mM Imidazole) par gradient sur 10 à 20 fois le volume de la colonne. Les fractions (1 ou 5 ml) récoltées sont analysées par dosage de l'activité enzymatique. Dans des plaques 96 puits sont introduits 270 µl de tampon 20 mM Tris-HCl, pH 7,5 et 15 µl de chaque fraction. La réaction enzymatique débute par ajout de 15 µl de substrat (L-leucine p-nitroanilide) à 40 mM. Après 20 minutes d'incubation à température ambiante, l'absorbance à 405 nm est mesurée à l'aide du lecteur de plaques Victor<sup>3</sup> de PerkinElmer. Les fractions ayant une forte absorption à 405 nm sont réunies et dialysées afin d'éliminer l'excès d'imidazole.

## *ii. Résine anti-FLAG*

L'étiquette FLAG fusionnée à l'extrémité C-terminale de la protéine recombinante (expression dans baculovirus) permet leur purification par chromatographie d'affinité sur une résine recouverte d'anticorps anti-FLAG (Anti-FLAG<sup>®</sup> M2 Affinity gel, Sigma Aldrich).

La résine (1 ml) est équilibrée avec du tampon TBS (50 mM Tris-HCl, 150 mM NaCl, pH 7,4). Après injection de l'échantillon par gravité, la résine est lavée avec du tampon TBS. Les protéines sont ensuite éluées par compétition avec du tampon TBS contenant 100 µg/ml de peptide FLAG (DYKDDDD). Les fractions récoltées sont analysées par SDS-PAGE.

#### 4.2.5. Chromatographie échangeuse d'ions

Cette technique chromatographique permet de séparer les protéines en fonction de leur différence de charge à un pH donné. Elle est basée sur l'affinité que les ions en solution ont pour les ions de charge opposée de la phase stationnaire. La phase stationnaire est généralement une résine sur laquelle sont liés chimiquement des groupements ioniques. Si la phase stationnaire est chargée positivement, il s'agira d'une colonne échangeuse d'anions, sinon (phase stationnaire chargée négativement) d'une colonne échangeuse de cations : les protéines de charge identique au support ne sont donc pas retenues. Les protéines fixées sont, quant à elles, séparées selon leur degré de charge positive ou négative à un certain pH par application d'un gradient salin (NaCl).

A un pH supérieur à son point isoélectrique (pI= 5,31), la protéine hAPN est chargée négativement, et peut donc être séparée sur une résine échangeuse d'anions. Différentes résines sont utilisées (Tableau 23) en fonction du stade de purification : capture initiale, purification intermédiaire ou finale.

Nom	Matrice	Taille des billes	Capacité dynamique de liaison
Q Sepharose 6 FF <i>Capture initiale ou intermédiaire</i>	Billes d'agarose réticulée 6% sur lesquelles sont greffées des amines quaternaires $-(O-CH_2-CHOH-CH_2)_2 N^+(CH_3)_3$	45-165 $\mu m$	120 mg de protéines /ml
Resource Q 1 ml <i>Purification intermédiaire</i>	Billes monodisperses de polystyrène divinyl-benzène sur lesquelles sont greffées des amines quaternaires $-(O-CH_2-CHOH-CH_2)_2 N^+(CH_3)_3$	15 $\mu m$	25 mg de protéine
Mono Q 5/50 GL 1 ml <i>Purification finale</i>	Billes rigides et poreuses de polystyrène divinyl-benzène sur lesquelles sont greffées des amines quaternaires $-CH_2 N^+(CH_3)_3$	10 $\mu m$	50 mg de protéine

*Tableau 23 : Caractéristiques des résines échangeuses d'anions utilisées pour séparer les protéines en fonction de leur différence de charge à un pH donné.*

La résine est au préalable équilibrée avec du tampon A (20-50 mM Tris-HCl, pH 7,5-8,0). Après injection de l'échantillon à un débit de 1 ou 2 ml/min, la colonne est lavée avec du tampon A jusqu'au retour du signal de détection UV 280 nm à la ligne de base. Les protéines fixées sont ensuite éluées sur 10 à 30 fois le volume de la colonne grâce au passage d'un gradient salin (tampon B : 20-50 mM Tris-HCl, pH 7,5-8,0, 0,5 M NaCl) à travers la résine. Des fractions de 0,5 à 5 mL sont récoltées et analysées par dosage de l'activité enzymatique; celles contenant la protéine d'intérêt sont réunies. Le sel (NaCl) est éliminé par dialyse ou ultrafiltration.

#### 4.2.6. Chromatographie à interactions hydrophobes (HIC)

La chromatographie d'interactions hydrophobes est fondée sur l'interaction entre les groupements hydrophobes d'une protéine et ceux du support chromatographique. La phase stationnaire porte un groupement hydrophobe tel qu'un noyau phénol ; la protéine à séparer est fixée sur le support chromatographique en présence d'une forte concentration en sel (sulfate d'ammonium). Le principe est similaire à la précipitation différentielle au sulfate d'ammonium. Les protéines fixées sont ensuite éluées avec un tampon de force ionique décroissante.

Différentes résines sont utilisées ([Tableau 24](#)) en fonction du stade de purification.

Nom	Matrice	Taille des billes	Capacité dynamique de liaison
Phenyl Sepharose 6 FF (high sub) (40 µmol phenyl/ml)	Billes d'agarose réticulée 6% sur lesquelles sont greffées des ligands hydrophobes R-O-CH <sub>2</sub> -CH(OH)-CH <sub>2</sub> -O-C <sub>6</sub> H <sub>5</sub>	45-165 µm	36 mg de protéine/ml
Resource Phe 1 ml	Billes monodisperses de polystyrène divinyl-benzène sur lesquelles sont greffées des ligands hydrophobes R-O-CH <sub>2</sub> -CH(OH)-CH <sub>2</sub> -O-C <sub>6</sub> H <sub>5</sub>	15 µm	25 mg de protéine

*Tableau 24: Caractéristiques des résines à interactions hydrophobes utilisées pour séparer les protéines en fonction de leur différence d'hydrophobicité.*

La résine est au préalable équilibrée avec du tampon A (50 mM Tris-HCl, pH 7,5, 1,7 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>). Après injection de l'échantillon à un débit de 1 ou 2 ml/min, la colonne est lavée avec du tampon A jusqu'au retour du signal de détection UV 280 nm à la ligne de base. Les protéines fixées sont ensuite éluées sur 10 à 30 fois le volume de la colonne en diminuant progressivement la teneur en sel du tampon (tampon B : 50 mM Tris-HCl, pH 7,5). Des fractions de 0,5 à 5 mL sont récoltées et analysées par dosage de l'activité enzymatique; celles contenant la protéine d'intérêt sont réunies, et le sulfate d'ammonium, (NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>, est éliminé par dialyse.

#### 4.2.7. Chromatographie d'exclusion

La chromatographie d'exclusion ou gel filtration est une technique de purification très résolutive qui permet de séparer les protéines en fonction de leur taille. La phase stationnaire est constituée de billes de polymère poreuses. La taille des pores est choisie en fonction de la masse moléculaire des protéines à séparer. Les grosses molécules, dont le

diamètre est supérieur à celui des pores, sont exclues et éluées les premières. Les petites et les moyennes sont éluées plus tardivement, car incluses dans le gel, leur migration est freinée. Les protéines sont donc éluées dans l'ordre inverse des masses moléculaires croissantes.

L'échantillon protéique est concentré par ultrafiltration (volume inférieur à 500 µl) puis, injecté à un débit de 0,4 ml/min sur une colonne d'exclusion Superdex 200 10/300 GL 24 ml (GE Healthcare, [Tableau 25](#)). La phase stationnaire est au préalable équilibrée dans du tampon 20-50 mM Tris-HCl, 0,15 M NaCl, pH 7,5-8,0. Des fractions de 1 ml sont récoltées et analysées par dosage de l'activité enzymatique; celles contenant la protéine d'intérêt sont réunies et concentrées par ultrafiltration.

Nom	Matrice	Taille des billes	Echelle de séparation
Superdex 200 10/300 GL 24 ml	Résine composite d'agarose réticulée et de dextrane	13 µm	10 à 600 kDa

*Tableau 25 : Caractéristiques de la colonne d'exclusion utilisée pour séparer les protéines en fonction de leur taille.*

### 4.3. Caractérisation de la protéine

#### 4.3.1. Détermination des paramètres cinétiques $K_M$ et $k_{cat}$

La détermination de la constante de Michaelis-Menten  $K_M$  (caractéristique d'une enzyme donnée pour un substrat précis) permet de s'assurer que l'enzyme purifiée est bien h-APN ( $K_M=0,2$  mM pour la L-leucine p-nitroanilide).

La mesure de  $K_M$  s'effectue en deux étapes. Dans un premier temps, il faut déterminer la quantité optimum d'enzyme nécessaire à l'hydrolyse de la L-leucine p-nitroanilide (absorbance comprise entre 0,5 et 1). Le protocole à suivre est le même que celui utilisé pour le dosage de l'activité enzymatique. Puis, par cinétique, la vitesse initiale de la réaction pour différentes concentrations de substrat (entre 0,05 et 0,7 mM) est mesurée. Les paramètres cinétiques de l'enzyme sont déterminés selon l'équation en double inverse de Lineweaver et Burk (il s'agit d'une relation linéaire entre  $1/V_i$  et  $1/[S]$ ) :

$$\frac{1}{V_i} = \left( \frac{K_M}{V_{\max}} \right) \frac{1}{[S]} + \frac{1}{V_{\max}}$$

Les paramètres cinétiques d'une enzyme permettent également de mesurer son efficacité catalytique. La constante catalytique est définie par:

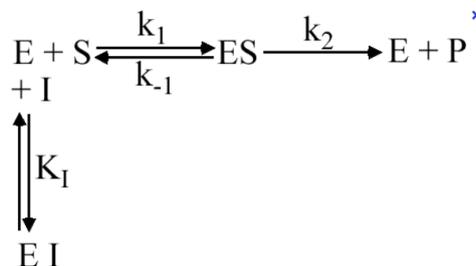
$$k_{cat} = \frac{V_{max}}{[E]_T}$$

Cette valeur est également appelée le « turnover number » d'une enzyme car elle est égale au nombre de molécules de substrat converti en produit par unité de temps par un seul site actif de l'enzyme à saturation. Dans le modèle de Michaelis-Menten:  $k_{cat} = k_2$ .

$k_{cat}/K_M$  correspond à une constante de vitesse pour de basses concentrations en substrat (quand  $[S] \ll K_M$ , très peu de ES est formé, d'où  $[E] = [E]_T$ ) et permet de mesurer l'efficacité catalytique de l'enzyme.

#### 4.3.2. Détermination de $K_i$

Des inhibiteurs compétitifs sélectifs de h-APN synthétisés au laboratoire sont évalués sur la protéine recombinante purifiée.



L'équation décrivant la vitesse initiale de formation du produit en présence d'un inhibiteur compétitif est la suivante :

$$V_i = \frac{V_{max} [S]}{K_M \left( 1 + \frac{[I]}{K_I} \right) + [S]}$$

L'affinité d'un inhibiteur pour une enzyme est donnée par la constante d'inhibition  $K_i$ , qui représente la constante de dissociation du complexe enzyme-inhibiteur. L'affinité d'un inhibiteur pour une enzyme est évaluée de façon précise et plus rapide par le biais de l' $IC_{50}$ ,

qui correspond à la concentration en inhibiteur pour laquelle la vitesse maximale est divisée par deux (50% d'inhibition).

La détermination de la constante d'inhibition  $K_i$  s'effectue en mesurant, par cinétique, la vitesse initiale de la réaction pour différentes concentrations d'inhibiteur (entre 10  $\mu$ M et 1 mM) et pour une concentration en substrat constante ( $[S]=K_M/2$ ,  $K_M$  et  $2K_M$ ). La représentation de Dixon (il s'agit d'une relation linéaire entre  $1/V_i$  et  $[I]$ ) permet ensuite de déterminer l' $IC_{50}$  et la constante d'inhibition  $K_i$  :

$$\frac{1}{V_i} = \frac{K_M}{v_{\max} [S] K_i} \cdot [I] + \frac{1}{v_{\max}} \left( 1 + \frac{K_M}{[S]} \right)$$

## 5. Essais de cristallisation

### 5.1. Principe

La cristallographie aux rayons X est une technique qui permet de déterminer la structure tridimensionnelle d'une protéine. L'étape limitant est l'obtention de cristaux de bonne qualité qui diffractent les rayons X. Un cristal est un arrangement ordonné d'atomes, d'ions ou de molécules, répété de façon périodique dans les trois dimensions de l'espace.

La cristallisation est un changement de phase; la protéine passe d'un état liquide désordonné à un état solide ordonné, le cristal. Lors de la cristallisation, la solubilité de la solution protéique diminue, par un procédé d'évaporation contrôlé, jusqu'à atteindre l'état de sursaturation favorable à la formation de cristaux. Cependant, la solubilité d'une protéine dépend de nombreux paramètres (concentration, pH, température, force ionique, présence d'ions ou d'agents précipitant) et le passage progressif d'un état soluble à un état ordonné peut rapidement basculer vers la formation d'un précipité amorphe.

La technique de cristallisation la plus utilisée est la diffusion en phase vapeur par goutte suspendue ou assise ([Figure 80](#)). La goutte contient un mélange de la protéine et d'un agent précipitant (des sels, des polymères, des solvants organiques ou des composés organiques non volatils) identique à celui présent dans le réservoir, mais la concentration en précipitant y est supérieure. L'enceinte est hermétiquement scellée. Un équilibre s'établit alors entre la goutte et le réservoir par diffusions des espèces volatiles, jusqu'à ce que la tension de vapeurs soit identique à la surface des deux compartiments. L'évaporation d'eau de la goutte vers le réservoir va entraîner une augmentation de la concentration de la protéine et de l'agent précipitant dans la goutte. Si des conditions appropriées ont été choisies, cela va amener la protéine à cristalliser. La méthode reste empirique et les chances de réussite aléatoires tant les paramètres physico-chimiques intervenant sont nombreux.



*Figure 80: Technique de cristallisation par diffusion en phase vapeur : (a) goutte suspendue ou (b) assise*

## 5.2. Les essais de cristallisation de l'aminopeptidase-N humaine

Les essais de cristallisation sont effectués au Consortium de génomique structurale d'Oxford avec l'aide de Grazyna Kochan (groupe Biotechnologie).

Une protéine pure et homogène en grande quantité (10 à 20 mg/ml) est nécessaire pour mener à bien ces essais. Dans certain cas, la présence d'un inhibiteur sélectif dans le site actif de l'enzyme peut favoriser le processus de cristallisation.

La protéine hAPN est au préalable incubée en présence d'un inhibiteur puissant (CBOS3-30,  $K_i = 60$  pm) : à 140  $\mu$ l de la solution protéique (14 mg/ml) est ajouté 1  $\mu$ l de CBOS3-30 à 1 mM (contenant 50% de DMSO). La concentration finale en inhibiteur dans la solution est égale à 7  $\mu$ M (0,35% de DMSO, dose non dommageable à la cristallisation).

Afin de trouver les conditions optimales de cristallisation de la protéine, de nombreuses solutions de cristallisation sont testées. Les essais de cristallisation sont réalisés en boîte 96 puits (3 gouttes par réservoir) par la technique de la goutte assise. Les réservoirs sont remplis avec les kits de criblage HIN et JCSG5 (voir annexe 3) à raison de 20  $\mu$ l de solution de cristallisation par réservoir. Le robot Mosquito (TTP Labtech) est ensuite utilisé pour faire des gouttes d'un volume final de 150 nl. Pour chaque condition de cristallisation plusieurs concentrations du mélange protéine-inhibiteur sont testées ([Tableau 26](#)).

Goutte	Volume de la solution protéine + inhibiteur	Volume du réservoir	Concentration finale dans la goutte
1 Ratio 2/1 P+I/réservoir	100 nl	50 nl	[hAPN]= 9,3 mg/ml [CBOS3-30]= 4,7 $\mu$ M (0,23% DMSO)
2 Ratio 1/1 P+I/réservoir	75 nl	75 nl	[hAPN]= 7 mg/ml [CBOS3-30]= 3,5 $\mu$ M (0,17% DMSO)
3 Ratio 1/2 P+I/réservoir	50 nl	100 nl	[hAPN]= 4,7 mg/ml [CBOS3-30]= 2,3 $\mu$ M (0,12% DMSO)

*Tableau 26*: Mélange réalisé par le robot Mosquito pour faire des gouttes de cristallisation d'un volume final de 150 nl.

Les boîtes sont incubées à 4 et 20°C ; des clichés des différentes gouttes sont pris automatiquement au bout de 24 h, 48 h, 4 jours, 1 semaine, 2 semaines, 1 mois, 2 mois et 3 mois.

CONCLUSION GENERALE



Les thérapies anti-angiogéniques ont connu un essor considérable ces dix dernières années grâce aux progrès réalisés dans la compréhension des mécanismes moléculaires conduisant à la néo-vascularisation tumorale. Différentes molécules aux propriétés anti-angiogéniques, ciblant principalement le VEGF et ses récepteurs, ont ainsi été développées. Bien que les essais cliniques aient permis de prouver l'efficacité de ces molécules, leur réponse reste imprévisible et varie considérablement d'un patient à l'autre pour un même type de cancer. Il est par conséquent nécessaire d'identifier de nouvelles cibles thérapeutiques pour développer les agents anti-angiogéniques de seconde génération, et ainsi pouvoir proposer des traitements adaptés à chaque patient.

Dans ce contexte, l'Aminopeptidase-N s'est avéré une cible thérapeutique potentielle intéressante. Des inhibiteurs de cette enzyme, ayant des constantes d'inhibition allant du nano au picomolaire, ont alors été conçus et évalués au laboratoire ; ces dérivés de type 7-amino-6-benzosubérone sont les inhibiteurs de l'APN les plus puissants et les plus sélectifs décrits à ce jour. *In vivo*, ces molécules sont actives dans des modèles de tumeurs coliques établies chez la souris. Leur effet sur la croissance tumorale dépend du niveau d'expression de la protéine par les cellules tumorales (et non au niveau de la néo-vascularisation), et est significatif lorsqu'ils sont utilisés en association avec la rapamycine. Même si ces résultats sont prometteurs, il reste, toutefois, de nombreuses étapes à franchir avant de pouvoir passer aux tests cliniques. Par ailleurs, des études pharmacocinétiques doivent être rapidement entreprises pour optimiser au mieux le traitement (dosage adéquat, mode d'administration).

En parallèle à ces travaux, et afin d'élucider le mode de liaison des inhibiteurs au site actif de l'enzyme, des essais d'expression, de purification et de cristallisation de l'APN humaine ont été entrepris. Les stratégies employées pour la production de la forme soluble de la protéine dans le système baculovirus-cellules d'insectes se sont avérées les plus fructueuses. Néanmoins, de nombreux paramètres doivent encore être optimisés : titration du stock viral, expression de la protéine (temps d'incubation, expression dans les cellules d'insectes Sf21 ou High Five). Le protocole de purification mis au point a permis d'isoler une quantité suffisante de protéine pour entamer les essais de cristallisation. A ce jour, aucun cristal ne s'est développé dans les conditions expérimentales testées. D'autres essais sont envisagés; pour favoriser le processus de cristallisation, les N-glycanes seront au préalable éliminés par digestion enzymatique à la PNGase F (cela devrait nous permettre d'avoir un échantillon plus homogène). Ces travaux réalisés par le SGC conduiront, nous l'espérons, à la résolution de la structure tridimensionnelle de l'APN humaine.



La co-cristallisation de l'APN d'*Escherichia coli* avec les inhibiteurs a été réalisée avec succès, permettant l'identification des pharmacophores. De nombreuses questions subsistent, comme par exemple, l'origine du gain d'affinité de l'inhibiteur pour l'APN de mammifères observé suite à la substitution de la position 1 de la 7-amino-6-benzosubérone par un groupement brome. La résolution de la structure tridimensionnelle de l'APN humaine demeure donc un enjeu majeur.

Le site actif des aminopeptidases de la famille M1 étant très conservé, il est possible de modifier notre « scaffold » de type amino-benzosubérone afin d'inhiber spécifiquement un autre membre de cette famille. Dans ce contexte, de nouveaux projets ont été développés au sein du laboratoire avec pour cible thérapeutique ERAP1 et ERAP2, des aminopeptidases impliquées dans la spondylarthrite ankylosante, ou encore l'Aminopeptidase de *Plasmodium falciparum*, PfA-M1, ayant une grande importance dans le cycle de développement du parasite responsable de la malaria.



## ANNEXES



## Annexe 1 : Séquence codante du gène ANPEP et séquence protéique de h-APN

Homo sapiens alanyl (membrane) aminopeptidase (ANPEP), mRNA (NCBI Reference Sequence: NM\_001150.2). CDS-293..3196 pb.

```
ATGGCCAAGGGCTTCTATATTTCCAAGTCCCTGGGCATCCTGGGGATCCTCCTGGGCGTGGCAGCCGTGTGCACA
ATCATCGCACTGTCACTGGTGTACTCCCAGGAGAAGAACAAGAACGCCAACAGCTCCCCCGTGGCCTCCACCACC
CCGTCCGCCTCAGCCACCACCAACCCCGCTCGGCCACCACCTTGGACCAAAGTAAAGCGTGGAAATCGTTACCGC
CTCCCCAACACGCTGAAACCCGATTCTACCCGGGTGACGCTGAGACCGTACCTCACCCCAATGACAGGGGCTG
TACGTTTTTAAGGGCTCCAGCACCGTCCGTTTTACCTGCAAGGAGGCCACTGACGTCATCATCATCCACAGCAAG
AAGCTCAACTACACCCTCAGCCAGGGGCACAGGGTGGTCTGCGTGGTGTGGGAGGCTCCAGCCCCCGACATT
GACAAGACTGAGCTGGTGGAGCCCCACCGAGTACCTGGTGGTGCACCTCAAGGGCTCCCTGGTGAAGGACAGCCAG
TATGAGATGGACAGCGAGTTTCGAGGGGGAGTTGGCAGATGACCTGGCGGGCTTCTACCCGAGCGATACATGGAG
GGCAATGTGAGAAAGGTGGTGGCCACTACACAGATGCAGGCTGCAGATGCCCGGAAGTCCCTCCCATGCTTCGAT
GAGCCGGCCATGAAGGCCGAGTTCAACATCACGCTTATCCACCCCAAGGACCTGACAGCCCTGTCCAACATGCTT
CCCAAAGGTCCCAGCACCCCACTTCCAGAAGACCCCAACTGGAATGTCACTGAGTTCCACACCACGCCCAAGATG
TCCACGTACTTGCTGGCCTTCATTGTGAGTGGTTCGACTACGTGGAGAAGCAGGCATCCAATGGTGTCTTGATC
CGGATCTGGGCCCCGGCCAGTGCCATTGCGGGGGCCACGGCGATTATGCCCTGAACGTGACGGGCCCATCCTT
AATTCTTTGCTGGTCAATTATGACACACCCTACCCACTCCAAAATCAGACCAGATTGGCCTGCCAGACTTCAAC
GCCGGCGCCATGGAGAAGTGGGGACTGGTGGTACCTACCGGGAGAAGTCCCTGCTGTTGACCCCTGTCTCTCTCC
AGCAGCAACAAGGAGCGGGTGGTCACTGTGATTGCTCATGAGCTGGCCACCAGTGGTTCGGGAACCTGGTGACC
ATAGAGTGGTGGAAATGACCTGTGGCTGAACGAGGGCTTCGCTCCTACGTGGAGTACCTGGGTGCTGACTATGCG
GAGCCACCTGGAACCTGAAAGACCTCATGGTGTGATGATGTGTACCGCGTGATGGCAGTGGATGCACTGGCC
TCCTCCCACCCGCTGTCCACACCCGCCTCGGAGATCAACACGCCCGGCCAGATCAGTGAGCTGTTTTGACGCCATC
TCCTACAGCAAGGGCGCCTCAGTCCCTCAGGATGCTCTCCAGCTTCTGTCCGAGGACGTATTCAAGCAGGGCCTG
GCGTCTACCTCCACACCTTTGCCTACCAGAACACCATCTACCTGAACCTGTGGGACCACCTGCAGGAGGCTGTG
AACAAACCGGTCCATCCAACCTCCCCACCACCGTGCGGGACATCATGAACCGCTGGACCCTGCAGATGGGCTTCCCG
GTCATCACGGTGGATAACAGCACGGGGACCCCTTTCCAGGAGCACTTCTCTCTGACCCCGATTCCAATGTTACC
CGCCCTCAGAATTAACACTACGTGTGGATTGTGCCCATCACATCCATCAGAGATGGCAGACAGCAGCAGGACTAC
TGGCTGATAGATGTAAGAGCCAGAACGATCTCTTCAGACATCAGGCAATGAGTGGTCCCTGTAACCTCAAT
GTGACGGGCTATTACCGGGTGAACCTACGACGAAGAGAAGTGGAGGAAGATTACAGACTCAGCTGCAGAGAGACCAC
TCGGCCATCCCTGTCAATCGGGCACAGATCATTAATGACGCTTCAACCTGGCCAGTGGCCATAAGTCCCT
GTCACTCTGGCGCTGAACAACACCCTCTTCTGATTGAAGAGAGACAGTACATGCCCTGGGAGGCCGCCCTGAGC
AGCCTGAGCTACTTCAAGCTCATGTTTGACCGCTCCGAGGTCTATGGCCCCATGAAGAAGTACCTGAAGAAGCAG
GTCACACCCCTCTTCACTTCACTTCCAGAAATAATAACCAACAACCTGGAGGGAGATCCCAGAAAACCTGATGGACCAG
TACAGCGAGGTTAATGCCATCAGCACCGCCTGCTCCAACGGAGTTCAGAGTGTGAGGAGATGGTCTCTGGCCTT
TTCAAGCAGTGGATGGAGAACCCCAATAATAACCCGATCCACCCCAACCTGCGGTCCACCGTCTACTGCAACGCT
ATCGCCAGGGCGGGGAGGAGGAGTGGGACTTCGCTGGGAGCAGTTCGAAATGCCACACTGGTCAATGAGGCT
GACAAGCTCCGGGCAGCCCTGGCCTGCAGCAAAGAGTTGTGGATCCTGAACAGGTACCTGAGCTACACCTGAAC
CCGGACTTAATCCGGAAGCAGGACGCCACCTCTACCATCATCAGCATTACCAACAACGTCATTGGGCAAGGTCTG
GTCTGGGACTTTGTCCAGAGCAACTGGAAGAAGCTTTTTAACGATTATGGTGGTGGCTCGTTCTCTCTTCTCCAAC
CTCATCCAGGCAGTGACACGACGATTCTCCACCGAGTATGAGCTGCAGCAGCTGGAGCAGTTCAAGAAGGACAAC
GAGGAAACAGGCTTCGGCTCAGGCACCCGGGCCCTGGAGCAAGCCCTGGAGAAGACGAAAGCCAAACATCAAGTGG
GTGAAGGAGAACAAGGAGGTGGTGTCTCCAGTGGTTACAGAAAACAGCAAATAG
```

>sp|P15144|AMPN\_HUMAN Aminopeptidase N OS=Homo sapiens GN=ANPEP PE=1 SV=4

```
MAKGFYISKSLGILGILLGVAAVCTIIALSVVYSQEKNKNANSSPVASTTPSASATTPNPASATTLTLDQSKAWNRYR
LPNTLKPDSYRVTLRPYLTPNDRGLYVFKGSSTVRFTEKEATDVI I IHSKKNLNYTLSQGHRVVLRLGVGGSQPPDI
DKTELVEPTEYLVVHLKGLVKDSQYEMDSEFEGELADLAGFYRSEYMEGNVRKVAVATTQMQAADARKSFPCFD
EPAMKAEFNITLIHPKDLTALSNNLPKGPSTPLPEDPNWNVTEFHTTPKMSTYLLAFIVSEFDYVEKQASNGVLI
RIWARPSAIAAGHDYALNVTGPILNFFAGHYDTPYPLKSDQIGLPDFNAGAMENWGLVTYRENSLLFDPLSSS
SSNPLRVVTVIAHLEAHQWFGNLVTIEWWNLDLWLNREGFSAYVEYLGADYAEPWNLKDMLMVLNDVYRVMVAVDALA
SSHPLSTPASEINTPAQISELFDIAISYKSGASVLRNGLSSFLSEDFVKQGLASYLHTFAYQNTIYLNLDWHLQEA
VNNRSIQLPTTVRDMNRWTLQMGFPVITVDSTGTLSQEHFLLDPSNVTRPSEFNIVVPIITSIRDGRQQDY
WLI DVRAQNDLDFSTSGNEWVLLNLTGYYRVNYDEENWRKIQTQLQRDHSAPVINRAQIINDAFNLASAHKVP
VTLALNNTLFLIEERQYMPWEAALSSLSYFKLMFDRSEVYGPMKNYLKQVTPLFIFHRNNTNWNREIPENLMDQ
YSEVNAISTACSNVPECEEMVSGLFKQWMENPNNNPIHPNLRSTVYCNAIAQGGEEDWFAWEQFRNATLVNEA
DKLRAALACSKELWILNRYLSYTLNPD LIRKQDATSTIIISITNNVIGQGLVWDFVQSNWKKLFNDYGGGSF SFSN
LIQAVTRRFSTEYELQOLEQFKKDNEETGFGSGTRALEQALEKTKANIKWVKENKEVVLQWFTENSK
```



## Annexe 2 : Usage des codons chez *K. lactis*

***Kluyveromyces lactis* [gbpln]: 301 CDS's (150121 codons)**

fields: [triplet] [frequency: per thousand] ([number])

UUU 19.5 ( 2929)	UCU 26.7 ( 4012)	UAU 20.2 ( 3033)	UGU 9.5 ( 1433)
UUC 23.5 ( 3534)	UCC 12.7 ( 1901)	UAC 17.0 ( 2557)	UGC 3.2 ( 483)
UUA 23.5 ( 3534)	UCA 16.3 ( 2445)	UAA 1.1 ( 163)	UGA 0.6 ( 83)
UUG 33.9 ( 5083)	UCG 7.7 ( 1150)	UAG 0.4 ( 55)	UGG 11.3 ( 1697)
CUU 11.9 ( 1779)	CCU 13.5 ( 2020)	CAU 14.2 ( 2130)	CGU 6.7 ( 1001)
CUC 4.3 ( 649)	CCC 3.8 ( 573)	CAC 6.9 ( 1043)	CGC 1.5 ( 228)
CUA 11.8 ( 1766)	CCA 21.3 ( 3201)	CAA 29.4 ( 4411)	CGA 2.1 ( 318)
CUG 5.1 ( 770)	CCG 4.2 ( 633)	CAG 11.8 ( 1769)	CGG 1.5 ( 224)
AUU 27.5 ( 4123)	ACU 22.9 ( 3444)	AAU 31.5 ( 4735)	AGU 12.8 ( 1917)
AUC 20.9 ( 3138)	ACC 12.8 ( 1923)	AAC 25.5 ( 3829)	AGC 6.3 ( 953)
AUA 15.8 ( 2368)	ACA 15.2 ( 2282)	AAA 37.5 ( 5629)	AGA 24.7 ( 3707)
AUG 21.0 ( 3158)	ACG 5.8 ( 874)	AAG 33.8 ( 5070)	AGG 6.0 ( 902)
GUU 25.9 ( 3893)	GCU 28.1 ( 4217)	GAU 40.7 ( 6116)	GGU 35.5 ( 5335)
GUC 14.2 ( 2138)	GCC 11.8 ( 1778)	GAC 18.4 ( 2762)	GGC 5.6 ( 845)
GUA 10.9 ( 1642)	GCA 15.5 ( 2334)	GAA 47.5 ( 7124)	GGA 11.5 ( 1727)
GUG 10.9 ( 1636)	GCG 4.9 ( 734)	GAG 15.9 ( 2393)	GGG 5.2 ( 788)

Coding GC 39.61% 1st letter GC 45.28% 2nd letter GC 36.75% 3rd letter GC 36.82%

***Homo sapiens* [gbpri]: 93487 CDS's (40662582 codons)**

fields: [triplet] [frequency: per thousand] ([number])

UUU 17.6 (714298)	UCU 15.2 (618711)	UAU 12.2 (495699)	UGU 10.6 (430311)
UUC 20.3 (824692)	UCC 17.7 (718892)	UAC 15.3 (622407)	UGC 12.6 (513028)
UUA 7.7 (311881)	UCA 12.2 (496448)	UAA 1.0 ( 40285)	UGA 1.6 ( 63237)
UUG 12.9 (525688)	UCG 4.4 (179419)	UAG 0.8 ( 32109)	UGG 13.2 (535595)
CUU 13.2 (536515)	CCU 17.5 (713233)	CAU 10.9 (441711)	CGU 4.5 (184609)
CUC 19.6 (796638)	CCC 19.8 (804620)	CAC 15.1 (613713)	CGC 10.4 (423516)
CUA 7.2 (290751)	CCA 16.9 (688038)	CAA 12.3 (501911)	CGA 6.2 (250760)
CUG 39.6 (1611801)	CCG 6.9 (281570)	CAG 34.2 (1391973)	CGG 11.4 (464485)
AUU 16.0 (650473)	ACU 13.1 (533609)	AAU 17.0 (689701)	AGU 12.1 (493429)
AUC 20.8 (846466)	ACC 18.9 (768147)	AAC 19.1 (776603)	AGC 19.5 (791383)
AUA 7.5 (304565)	ACA 15.1 (614523)	AAA 24.4 (993621)	AGA 12.2 (494682)
AUG 22.0 (896005)	ACG 6.1 (246105)	AAG 31.9 (1295568)	AGG 12.0 (486463)
GUU 11.0 (448607)	GCU 18.4 (750096)	GAU 21.8 (885429)	GGU 10.8 (437126)
GUC 14.5 (588138)	GCC 27.7 (1127679)	GAC 25.1 (1020595)	GGC 22.2 (903565)
GUA 7.1 (287712)	GCA 15.8 (643471)	GAA 29.0 (1177632)	GGA 16.5 (669873)
GUG 28.1 (1143534)	GCG 7.4 (299495)	GAG 39.6 (1609975)	GGG 16.5 (669768)

Coding GC 52.27% 1st letter GC 55.72% 2nd letter GC 42.54% 3rd letter GC 58.55%



## Annexe 3 : Kit de cristallisation HIN

Reagent number	[Salt]	[Salt] units	Salt	[Buffer]	[Buffer] units	Buffer	pH	[ppt1]	ppt1 units	ppt1	[ppt2]	[ppt2] units	ppt2			
A1				0,1	M	citric acid	3,5	2,0	M	ammonium sulfate						
A2				0,1	M	sodium acetate trihydrate	4,5	2,0	M	ammonium sulfate						
A3				0,1	M	bis-tris	5,5	2,0	M	ammonium sulfate						
A4				0,1	M	bis-tris	6,5	2,0	M	ammonium sulfate						
A5				0,1	M	HEPES	7,5	2,0	M	ammonium sulfate						
A6				0,1	M	tris	8,5	2,0	M	ammonium sulfate						
A7				0,1	M	citric acid	3,5	3,0	M	sodium chloride						
A8				0,1	M	sodium acetate trihydrate	4,5	3,0	M	sodium chloride						
A9				0,1	M	bis-tris	5,5	3,0	M	sodium chloride						
A10				0,1	M	bis-tris	6,5	3,0	M	sodium chloride						
A11				0,1	M	HEPES	7,5	3,0	M	sodium chloride						
A12				0,1	M	tris	8,5	3,0	M	sodium chloride						
B1				0,1	M	bis-tris	5,5	0,3	M	magnesium formate						
B2				0,1	M	bis-tris	6,5	0,5	M	magnesium formate						
B3				0,1	M	HEPES	7,5	0,5	M	magnesium formate						
B4				0,1	M	tris	8,5	0,3	M	magnesium formate						
B5								1,3	M	sodium dihydrogen phosphate	0,14	M	di-potassium hydrogen phosphate			
B6								0,5	M	sodium dihydrogen phosphate	0,91	M	di-potassium hydrogen phosphate			
B7								0,1	M	sodium dihydrogen phosphate	1,344	M	di-potassium hydrogen phosphate			
B8				0,1	M	HEPES	7,5	1,4	M	tri-sodium citrate dihydrate						
B9								1,8	M	tri-ammonium citrate pH 7.0						
B10								0,8	M	succinic acid pH 7.0						
B11								2,1	M	DL-malic acid pH 7.0						
B12								2,8	M	sodium acetate trihydrate pH 7.0						
C1								3,5	M	sodium formate pH 7.0						
C2								1,1	M	di-ammonium tartrate pH 7.0						
C3								2,4	M	sodium malonate pH 7.0						
C4								35,0	%v/v	tacsimate pH 7.0						
C5								60,0	%v/v	tacsimate pH 7.0						
C6	0,1	M	sodium chloride	0,1	M	bis-tris	6,5	1,5	M	ammonium sulfate						
C7	0,8	M	potassium sodium tartrate	0,1	M	tris	8,5	0,5	%w/v	polyethylene glycol MME 5000						
C8	1,0	M	ammonium sulfate	0,1	M	bis-tris	5,5	1,0	%w/v	polyethylene glycol 3350						
C9	1,1	M	sodium malonate pH 7.0	0,1	M	HEPES	7,0	0,5	%v/v	jeffamine ED-2001 reagent pH 7.0						
C10	1,0	M	succinic acid pH 7.0	0,1	M	HEPES	7,0	1,0	%w/v	polyethylene glycol MME 2000						
C11	1,0	M	ammonium sulfate	0,1	M	HEPES	7,0	0,5	%w/v	polyethylene glycol 8000						
C12	15,00	%w/v	tacsimate pH 7.0	0,1	M	HEPES	7,0	2,0	%w/v	polyethylene glycol 3350						
D1								25,0	%w/v	polyethylene glycol 1500						
D2				0,1	M	HEPES	7,0	30,0	%v/v	jeffamine M-600 reagent pH 7.0						
D3				0,1	M	HEPES	7,0	30,0	%v/v	jeffamine ED-2001 reagent pH 7.0						
D4				0,1	M	citric acid	3,5	25,0	%w/v	polyethylene glycol 3350						
D5				0,1	M	sodium acetate trihydrate	4,5	25,0	%w/v	polyethylene glycol 3350						
D6				0,1	M	bis-tris	5,5	25,0	%w/v	polyethylene glycol 3350						
D7				0,1	M	bis-tris	6,5	25,0	%w/v	polyethylene glycol 3350						
D8				0,1	M	HEPES	7,5	25,0	%w/v	polyethylene glycol 3350						
D9				0,1	M	tris	8,5	25,0	%w/v	polyethylene glycol 3350						
D10				0,1	M	bis-tris	6,5	20,0	%w/v	polyethylene glycol MME 5000						
D11				0,1	M	bis-tris	6,5	28,0	%w/v	polyethylene glycol MME 2000						
D12	0,2	M	calcium chloride	0,1	M	bis-tris	5,5	45,0	%v/v	2-methyl-2,4-pentanediol						
E1	0,2	M	calcium chloride	0,1	M	bis-tris	6,5	45,0	%v/v	2-methyl-2,4-pentanediol						
E2	0,2	M	ammonium acetate	0,1	M	bis-tris	5,5	45,0	%v/v	2-methyl-2,4-pentanediol						
E3	0,2	M	ammonium acetate	0,1	M	bis-tris	6,5	45,0	%v/v	2-methyl-2,4-pentanediol						
E4	0,2	M	ammonium acetate	0,1	M	HEPES	7,5	45,0	%v/v	2-methyl-2,4-pentanediol						
E5	0,2	M	ammonium acetate	0,1	M	tris	8,5	45,0	%v/v	2-methyl-2,4-pentanediol						
E6	0,05	M	calcium chloride	0,1	M	bis-tris	6,5	30,0	%v/v	polyethylene glycol MME 550						
E7	0,05	M	magnesium chloride	0,1	M	HEPES	7,5	30,0	%v/v	polyethylene glycol MME 550						
E8	0,2	M	potassium chloride	0,05	M	HEPES	7,5	35,0	%v/v	pentaerythritol propoxylate (5/4 PO/OH)						
E9	0,05	M	ammonium sulfate	0,05	M	ammonium sulfate	6,5	30,0	%v/v	pentaerythritol ethoxylate (15/4 EO/OH)						
E10				0,1	M	bis-tris	6,5	45,0	%v/v	polypropylene glycol P 400						
E11	0,02	M	magnesium chloride	0,1	M	HEPES	7,5	22,0	%w/v	polyacrylic acid 5100 sodium salt						
E12	0,10	M	cobalt chloride	0,1	M	tris	8,5	20,0	%w/v	polyvinylpyrrolidone K15						
F1	0,20	M	proline	0,1	M	HEPES	7,5	10,0	%w/v	polyethylene glycol 3350						
F2	0,20	M	trimethylamine n-oxide	0,1	M	tris	8,5	20,0	%w/v	polyethylene glycol MME 2000						
F3	5,00	%w/v	tacsimate pH 7.0	0,1	M	HEPES	7,0	10,0	%w/v	polyethylene glycol MME 5000						
F4				0,1	M	HEPES	7,5	12,0	%w/v	polyethylene glycol 3350	0,005	M	nickel (II) chloride	0,005	M	cadmium chloride
F5	0,10	M	ammonium acetate	0,1	M	bis-tris	5,5	17,0	%w/v	polyethylene glycol 10,000	0,005	M	cobalt chloride	0,005	M	magnesium chloride
F6	0,20	M	ammonium sulfate	0,1	M	bis-tris	5,5	25,0	%w/v	polyethylene glycol 3350						
F7	0,20	M	ammonium sulfate	0,1	M	bis-tris	6,5	25,0	%w/v	polyethylene glycol 3350						
F8	0,20	M	ammonium sulfate	0,1	M	HEPES	7,5	25,0	%w/v	polyethylene glycol 3350						
F9	0,20	M	ammonium sulfate	0,1	M	tris	8,5	25,0	%w/v	polyethylene glycol 3350						
F10	0,20	M	sodium chloride	0,1	M	bis-tris	5,5	25,0	%w/v	polyethylene glycol 3350						
F11	0,20	M	sodium chloride	0,1	M	bis-tris	6,5	25,0	%w/v	polyethylene glycol 3350						
F12	0,20	M	sodium chloride	0,1	M	HEPES	7,5	25,0	%w/v	polyethylene glycol 3350						
G1	0,20	M	sodium chloride	0,1	M	tris	8,5	25,0	%w/v	polyethylene glycol 3350						
G2	0,20	M	lithium sulfate	0,1	M	bis-tris	5,5	25,0	%w/v	polyethylene glycol 3350						
G3	0,20	M	lithium sulfate	0,1	M	bis-tris	6,5	25,0	%w/v	polyethylene glycol 3350						
G4	0,20	M	lithium sulfate	0,1	M	HEPES	7,5	25,0	%w/v	polyethylene glycol 3350						
G5	0,20	M	lithium sulfate	0,1	M	tris	8,5	25,0	%w/v	polyethylene glycol 3350						
G6	0,20	M	ammonium acetate	0,1	M	bis-tris	5,5	25,0	%w/v	polyethylene glycol 3350						
G7	0,20	M	ammonium acetate	0,1	M	bis-tris	6,5	25,0	%w/v	polyethylene glycol 3350						
G8	0,20	M	ammonium acetate	0,1	M	HEPES	7,5	25,0	%w/v	polyethylene glycol 3350						
G9	0,20	M	ammonium acetate	0,1	M	tris	8,5	25,0	%w/v	polyethylene glycol 3350						
G10	0,20	M	magnesium chloride	0,1	M	bis-tris	5,5	25,0	%w/v	polyethylene glycol 3350						
G11	0,20	M	magnesium chloride	0,1	M	bis-tris	6,5	25,0	%w/v	polyethylene glycol 3350						
G12	0,20	M	magnesium chloride	0,1	M	HEPES	7,5	25,0	%w/v	polyethylene glycol 3350						
H1	0,20	M	magnesium chloride	0,1	M	tris	8,5	25,0	%w/v	polyethylene glycol 3350						
H2	0,20	M	potassium sodium tartrate					20,0	%w/v	polyethylene glycol 3350						
H3	0,20	M	sodium malonate pH 7.0					20,0	%w/v	polyethylene glycol 3350						
H4	0,20	M	tri-ammonium citrate pH 7.0					20,0	%w/v	polyethylene glycol 3350						
H5	0,10	M	succinic acid pH 7.0					15,0	%w/v	polyethylene glycol 3350						
H6	0,20	M	sodium formate					20,0	%w/v	polyethylene glycol 3350						
H7	0,15	M	DL-malic acid pH 7.0					20,0	%w/v	polyethylene glycol 3350						
H8	0,10	M	magnesium formate					15,0	%w/v	polyethylene glycol 3350						
H9	0,05	M	zinc acetate					20,0	%w/v	polyethylene glycol 3350						
H10	0,20	M	tri-sodium citrate					20,0	%w/v	polyethylene glycol 3350						
H11	0,10	M	potassium thiocyanate					30,0	%w/v	polyethylene glycol MME 2000						
H12	0,15	M	potassium bromide					30,0	%w/v	polyethylene glycol MME 2000						

## Annexe 3 : kit de cristallisation JCSG5

Well	[major] unit	major	[minor] unit	minor	minor(pH)	[other] unit	other	[buff] unit	buff	buff(pH)	Well
A1	50 %	PEG 300	0,20 M	Li <sub>2</sub> SO <sub>4</sub>				0,1 M	acetate	4,5	A1
A2	20 %	PEG 3350						0,1 M	citrate	5,5	A2
A3	20 %	PEG 3350	0,20 M	(NH <sub>4</sub> ) <sub>2</sub> H(cit)	5,0						A3
A4	30 %	MPD	0,02 M	CaCl <sub>2</sub>				0,1 M	acetate	4,6	A4
A5	20 %	PEG 3350	0,20 M	Mg(form) <sub>2</sub>	5,9						A5
A6	20 %	PEG 1K	0,20 M	Li <sub>2</sub> SO <sub>4</sub>				0,1 M	cit/phos	4,2	A6
A7	20 %	PEG 10K						0,1 M	BICINE	9,3	A7
A8	20 %	PEG 3350	0,20 M	Na(form)							A8
A9	20 %	PEG 3350	0,20 M	(NH <sub>4</sub> )Cl	6,3						A9
A10	1,00 M	Li <sub>2</sub> SO <sub>4</sub>	0,50 M	TMAO							A10
A11	50 %	MPD	0,10 M	(NH <sub>4</sub> ) <sub>2</sub> HPO <sub>4</sub>				0,1 M	TRIS	8,5	A11
A12	20 %	PEG 3350	0,20 M	NaNO <sub>3</sub>							A12
B1			0,80 M	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>				0,1 M	citrate	4,0	B1
B2	20 %	PEG 3350	0,20 M	KSCN							B2
B3	20 %	PEG 6K						0,1 M	BICINE	9,0	B3
B4	10 %	PEG 10K	8 %	EtGly				0,1 M	HEPES	7,5	B4
B5	40 %	MPD	5 %	PEG 10K				0,1 M	cacodylate	6,5	B5
B6	40 %	isopropanol	5 %	PEG 1K				0,1 M	citrate/PO <sub>4</sub>	4,2	B6
B7	8 %	PEG 3350						0,1 M	acetate	4,6	B7
B8	10 %	PEG 10K	0,20 M	MgCl <sub>2</sub>				0,1 M	TRIS	7,0	B8
B9	20 %	PEG 6K						0,1 M	citrate	5,0	B9
B10	50 %	PEG 300	0,20 M	MgCl <sub>2</sub>				0,1 M	cacodylate	6,5	B10
B11	2 M	K <sub>3</sub> (cit)									B11
B12	20 %	PEG 3350	0,20 M	K <sub>3</sub> (cit)	8,3						B12
C1	20 %	PEG 10K	0,20 M	NaCl				0,1 M	citrate/PO <sub>4</sub>	4,2	C1
C2	20 %	PEG 6K	1,00 M	LiCl				0,1 M	citrate	4,0	C2
C3	15 %	mPEG 2K	10 %	PVPdne				0,1 M	imidazole	6,5	C3
C4	10 %	PEG 6K						0,1 M	HEPES	7,0	C4
C5	1,60 M	Na/KPO <sub>4</sub>						0,1 M	HEPES	7,5	C5
C6	40 %	PEG 300						0,1 M	citrate/PO <sub>4</sub>	4,2	C6
C7	10 %	PEG 3350	0,20 M	Zn(ac) <sub>2</sub>				0,1 M	acetate	4,5	C7
C8	20 %	isopropanol						0,1 M	TRIS	8,5	C8
C9	25 %	1,2 propandiol	10 %	glycerol				0,1 M	Na/K-PO <sub>4</sub>	6,2	C9
C10	10 %	PEG 10K	2 %	dioxane				0,1 M	BICINE	9,0	C10
C11	2,00 M	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>						0,1 M	acetate	4,6	C11
C12	10 %	PEG 10K	10 %	PEG 1K							C12
D1	24 %	PEG 1K	20 %	glycerol							D1
D2	30 %	PEG 300	0,20 M	MgCl <sub>2</sub>				0,1 M	HEPES	7,5	D2
D3	50 %	PEG 300	0,20 M	NaCl				0,1 M	Na/K-PO <sub>4</sub>	6,2	D3
D4	30 %	PEG 10K	0,20 M	Li <sub>2</sub> SO <sub>4</sub>				0,1 M	acetate	4,5	D4
D5	70 %	MPD						0,1 M	HEPES	7,5	D5
D6	20 %	PEG 10K	0,20 M	MgCl <sub>2</sub>				0,1 M	TRIS	8,5	D6
D7	40 %	PEG 300	0,20 M	Li <sub>2</sub> SO <sub>4</sub>				0,1 M	TRIS	8,5	D7
D8	40 %	MPD						0,1 M	TRIS	8,0	D8
D9	25,5 %	PEG 3350	0,17 M	(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>		15 %	glyc				D9
D10	40 %	PEG 300	0,20 M	Ca(ac) <sub>2</sub>				0,1 M	cacodylate	6,5	D10
D11	14 %	isopropanol	0,14 M	CaCl <sub>2</sub>		30 %	glyc	0,7 M	acetate	4,6	D11
D12	16 %	PEG 10K	0,04 M	Na/KPO <sub>4</sub>		20 %	glyc				D12





Aminopeptidase-N (APN or CD13) [EC.3.4.11.2] is a highly glycosylated type II membrane-bound ectoenzyme that belongs to the M1 family of zinc-dependent aminopeptidases. The members of this family have a thermolysin-like catalytic domain with the consensus HEXXH-X18-E zinc-binding sequence and an exopeptidase motif, GXMEN, in the active site. APN removes the N-terminal amino-acid from unsubstituted oligopeptides, amides or arylamides, with a broad substrate specificity. Although a significant preference for neutral and hydrophobic residues is observed. APN/CD13 is a widespread enzyme, located in many tissues, organs and cells, whose multiple functions dependent on its location. It is overexpressed on the endothelial cells of angiogenic, but not normal, vasculature, as well as on numerous tumor cells. As it was demonstrated that APN plays a critical role in tumor cell angiogenesis and metastasis, this protein was identified as a potential target for cancer therapy.

In this context, highly potent and selective non-peptidic APN inhibitors, with  $K_i$  values ranging from micro to nanomolar, were previously designed and synthesized in the laboratory. *In vitro* and *in vivo* efficacy of these novel amino-benzosuberone derivatives was tested. Endothelial cell morphogenesis as well as cell motility were inhibited *in vitro* in a dose-dependent manner, at concentrations that correlated with the compounds potency. *In vivo* efficacy of these inhibitors was evaluated on human tumor xenograft models. During the early phase of tumor progression, the anti-tumor effect can be ascribed to their anti-angiogenic activity. On established colon tumors, the effects on tumor growth can be correlated with APN expression. Moreover combined therapies with rapamycin (mTOR inhibitor) were proved to bring additional benefit.

In parallel to these works, a new project was started a few years ago which consists to solve the 3D structure of mammalian APN. Co-crystallizations with amino-benzosuberone derivatives would determine the binding mode of these novel inhibitors. Nevertheless solving the structure of a membrane protein like mammalian APN still remains a challenge. Therefore several cloning and expression strategies for human APN production were developed. Recombinant protein expression in both bacteria (*Escherichia coli*) and yeast (*Kluyveromyces lactis*) systems was unsuccessful. The strategy developed by the team of Pr. Udo Oppermann of SGC Oxford in the baculovirus/insect cells system was the most effective. It led to the production, in sufficient quantity, of a soluble and active form of APN.

The protocol below was followed:

- Cloning of the gene of interest into a baculovirus expression vector which derived from pFastBac (Invitrogen)
- Generation of recombinant bacmid DNA
- Transfection of Sf9 insect cells
- Amplification of recombinant baculovirus particules
- Insect cells infection and recombinant protein expression

Finally, anion exchange chromatography was used to purify the recombinant protein. Co-crystallization of pure glycosylated or deglycosylated soluble form of APN with our best inhibitor is currently underway. It should lead to the solving of the 3D structure of the protein as well as the binding mode of our potent inhibitors.

L'Amino-peptidase-N (APN ou CD13) [EC.3.4.11.2] est une ecto-enzyme ubiquitaire fortement glycosylée appartenant à la classe des aminopeptidases zinc-dépendantes de la famille M1. Elle est surexprimée à la surface des cellules endothéliales de vaisseaux angiogéniques, ainsi que sur un certain nombre de cellules tumorales. Son expression est liée à l'agressivité des tumeurs, et est un facteur de mauvais pronostic pour le patient. Par ailleurs, il existe une corrélation étroite entre l'élévation de l'expression de l'APN, une activité enzymatique accrue et le pouvoir invasif de nombreux types de cellules tumorales.

Des inhibiteurs de l'APN, appartenant à la famille des composés de type amino-benzosubérone, ont précédemment été synthétisés au laboratoire. Il s'est avéré que ces composés ont des constantes d'inhibition allant du nano au picomolaire, et sont les inhibiteurs les plus puissants et les plus sélectifs de l'APN connus à ce jour. Ces composés sont actifs *in vitro* dans des modèles standards d'étude de l'angiogenèse et de la migration des cellules tumorales. *In vivo*, dans des modèles de xénogreffes de tumeurs humaines implantées chez la souris « nude », un effet sur la néovascularisation tumorale est observé lorsque les inhibiteurs sont administrés le jour suivant l'injection des cellules cancéreuses. Ces molécules sont également actives dans des modèles de tumeurs établies chez la souris. L'effet sur la croissance tumorale dépend de la quantité de protéines exprimée par les cellules tumorales, mais pas de son expression au niveau de la néovascularisation ; cet effet est d'autant plus important si ces composés sont utilisés en association avec la rapamycine.

En parallèle à ces essais, un nouveau projet a été débuté il y a quelques années, visant à déterminer la structure tridimensionnelle de l'APN humaine. La connaissance de cette structure constitue un enjeu majeur car des co-cristallisations avec ses inhibiteurs permettraient de résoudre le mode de liaison à l'APN de cette nouvelle famille de composés. La difficulté de cette étude réside dans le fait que la technique utilisée pour déterminer la structure tridimensionnelle d'une protéine (la cristallographie des rayons X) nécessite la purification et la cristallisation de cette dernière.

Plusieurs stratégies de clonage et de surexpression de l'APN humaine ont été envisagées, avec pour objectif final, l'obtention d'une protéine cristallisable, glycosylée ou non. Après différents essais d'expression chez la levure et la bactérie *Escherichia coli*, il s'est avéré que le système baculovirus – cellules d'insectes a été le plus performant, conduisant à l'obtention de la protéine en quantité suffisante.

Le protocole ci-dessous a ainsi été suivi :

- clonage dans un vecteur d'expression baculovirus,
- génération du bacmide recombinant,
- transfection des cellules d'insectes Sf9,
- amplification du baculovirus recombinant,
- expression de la protéine recombinante.

Une fois la protéine surexprimée, elle a été purifiée par diverses techniques de chromatographie (d'affinité et échangeuse d'anions). La co-cristallisation de la protéine pure avec un inhibiteur puissant est actuellement en cours, et devrait aboutir à la résolution du mode de liaison de ses inhibiteurs au site actif de l'APN.